

UNIVERSIDAD POLITÉCNICA DE VALENCIA

Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos

Departamento de Ingeniería Hidráulica y Medio Ambiente



**CARACTERIZACIÓN ESTOCÁSTICA DE
LA REALIZACIÓN ESPACIO-TEMPORAL
DE EVENTOS HIDROLÓGICOS
EXTREMOS DE SEQUÍAS**

Tesis Doctoral,

Presentada por D. Javier González Pérez

Dirigida por Dr. D. Juan Bautista Marco Segura

Valencia, Enero de 2005

*A Isabel,
por su amor y comprensión.*

Agradecimientos

Mis agradecimientos se dirigen en primer lugar a mi Director de Tesis, el Doctor Juan Bautista Marco Segura, por su apoyo, consejos y confianza mostrada, no sólo durante el desarrollo de esta Tesis, sino desde que nos conocemos. Su colaboración ha estado presente en toda mi formación científica e investigadora.

Mis agradecimientos van también dirigidos a las tres universidades que me han acogido y han participado en mi formación:

- La Universidad Politécnica de Valencia, y en concreto el Departamento de Ingeniería Hidráulica y Medio Ambiente, con mayor énfasis en la unidad de Ingeniería Hidráulica y Hidrología, destacando al Doctor Félix Francés García, y la unidad de Ingeniería de Recursos Hídricos, destacando allí al Doctor Ignacio Escuder Bueno y todo su grupo.
- La Universidad de Arizona, y en concreto el Departamento de Ingeniería Civil e Ingeniería Mecánica, el Departamento de Hidrología y Recursos Hídricos, y el Centro para la Sostenibilidad Hidrológica de zonas semi-Áridas y zonas de Ribera. En especial debo agradecer toda su contribución en mi formación al Doctor Juan Bernardo Valdés, quien ante todo es un gran amigo. A él se unen todo un grupo de compañeros, que hacen mis estancias allí siempre deseables.
- La Universidad de Castilla-La Mancha, y en concreto la E.T.S.I. Caminos, Canales y Puertos, donde se hace posible mi vocación docente, por su confianza y el apoyo prestado. El entusiasmo por construir una escuela reconocida, produce la unión de la pequeña familia que la compone.

Quisiera también expresar mis agradecimiento a los comentarios, sugerencias y consejos de todas aquellas otras personas que de forma desinteresada han contribuido a conseguir un mejor trabajo.

Por último, y siendo lo más importante, agradezco a mi familia y amigos el apoyo prestado para la consecución de esta Tesis.

Resumen

Las sequías se consideran como uno de los principales riesgos naturales del planeta. Sus efectos se cuantifican en graves daños sobre al medio ambiente, la economía o la sociedad. Las realizaciones de sequías afectan a todos los sistemas, naturales y artificiales, usuarios del agua. En este sentido, son múltiples las facetas que las sequías tienen y muy variados los daños que pueden producir, dependiendo de la vulnerabilidad de cada sistema. Sin embargo, todas las afecciones producidas tienen un origen común, encontrado en la realización de un evento natural extraordinario, que se repite con poca frecuencia en el tiempo, y provoca una reducción en la presencia de agua, persistente y extensa. Este fenómeno, dada la incapacidad para su predicción, encuentra en el tratamiento estadístico y estocástico de fenómenos aleatorios extremos su vía natural de análisis.

Aquí se presenta una metodología para la caracterización estocástica de los eventos hidrológicos extremos de sequía de acuerdo a su naturaleza aleatoria. La metodología se basa en la caracterización de las desviaciones persistentes extremas de una variable aleatoria, con respecto a su régimen de variación normal. Esta caracterización se cuantifica en la forma de la frecuencia media de recurrencia, y da lugar a la definición de un nuevo índice de sequías: el Índice de Frecuencia de Sequías (IFS). El índice permite el estudio y evaluación de sequía sobre cualquier variable hidro-climática aleatoria, en la que se manifieste el efecto de las sequías. La metodología se complementa con una sistemática que permite el análisis de las sequías en su desarrollo espacio-temporal, a través de la elaboración de mapas de IFS. Ello hace posible caracterizar, desde un punto de vista estocástico, la realización de eventos de sequía, incorporando en el análisis el carácter extraordinario manifestado en su comportamiento persistente y extenso. El IFS no establece en su definición el inicio y fin de las sequías, sino que caracteriza la desviación que está teniendo lugar en términos de su frecuencia media de recurrencia. Para cada aplicación concreta, sobre la escala del IFS puede ser definida la situación de sequía, estableciendo la frecuencia umbral de acuerdo a la vulnerabilidad particular. Esto posibilita uniformizar sobre una misma escala las definiciones de sequías.

Al desarrollo de la metodología se añade el análisis de las principales limitaciones en la caracterización de sequías con las metodologías disponibles y su comparación con la aquí propuesta. Son estudiadas las problemáticas existentes entorno a la definición de las magnitudes críticas y la definición del estado de sequía en la aplicación de modelos basados en la Teoría de Rachas, el problema de la escala temporal en el análisis de sequías, y la comparabi-

lidad de los índices de sequías evaluados sobre diferentes climas y variables. Por último, se presentarán un conjunto de aplicaciones que demuestran el potencial uso de la metodología en el monitoreo de sequías, usando el caso de las precipitaciones sobre la cuenca del río Colorado (EEUU), y en la gestión de los recursos hídricos, mostrando aplicaciones en el sistema del río Júcar, sobre la que se analiza una política de operación en base al IFS, y el sistema del río Conchos (México), en la que se muestra su aplicación en la definición de conceptos tales como sequía extraordinaria.

Palabras clave

Sequías, caracterización estocástica, análisis multivariado, análisis espacial, teoría de rachas, índices de sequías, mapas de sequías.

Resum

Les sequeres es consideren com un dels principals riscos naturals del planeta. Els seus efectes es quantifiquen en greus danys en el medi ambient, l'economia o la societat. Les realitzacions de sequera afecten a tots els sistemes, naturals i artificials, usuaris de l'aigua. En aquest sentit, són múltiples les facetes que les sequeres tenen i molt variats els danys que poden produir, depenent de la vulnerabilitat de cada sistema. No obstant això, totes les afeccions produïdes tenen un origen comú, trobat en la realització d'un esdeveniment natural extraordinari, que es repeteix amb poca freqüència en el temps, i provoca una reducció de la presència d'aigua, persistent i extensa. Aquest fenomen, donada la incapacitat per a la seva predicció, troba en el tractament estadístic y estocàstic de fenòmens aleatoris extrems la seva via natural d'anàlisi.

Ací es presenta una metodologia per a la caracterització estocàstica dels esdeveniments hidrològics extrems de sequera d'acord amb la seva naturalesa aleatòria. La metodologia es basa en la caracterització de les desviacions persistents extremes d'una variable aleatòria, pel que fa al seu règim de variació normal. Aquesta caracterització es quantifica en la forma de la freqüència mitja de recurrència, i dóna lloc a la definició d'un nou índex de sequeres: l'índex de freqüència de sequeres (IFS). L'índex permet l'estudi i avaluació de sequeres sobre qualsevol variable hidro-climàtica aleatòria, en la qual es manifeste l'efecte de les sequeres. La metodologia es completa amb una sistemàtica que permet l'anàlisi de les sequeres en el seu desenvolupament espai-temporal, a través de l'elaboració de mapes d'IFS. Això fa possible caracteritzar, des d'un punt de vista estocàstic, la realització d'esdeveniments de sequera, incorporant en l'anàlisi el caràcter extraordinari manifestat en el seu comportament persistent i extens. L'IFS no estableix en la seva definició el començament i la fi de les sequeres, sinó que caracteritza la desviació que està tenint lloc en termes de la seva freqüència mitja de recurrència. Per a cada aplicació concreta, sobre l'escala de l'IFS pot ser definida la situació de sequera, d'acord a la vulnerabilitat particular. Això possibilita uniformitzar sobre una mateixa escala les definicions de sequeres.

Al desenvolupament de la metodologia s'afegeix l'anàlisi de les principals limitacions en la caracterització de sequeres amb les metodologies disponibles i la seva comparació amb la proposada ací. Són estudiades les problemàtiques existents entorn a la definició de les magnituds crítiques i la definició de l'estat de sequera en l'aplicació de models basats en la Teoria de Ratxes, el problema de l'escala temporal en l'anàlisi de sequeres, i la

possible comparació dels índexs de sequeres avaluats sobre diferents climes i variables. Finalment, es presentaran un conjunt d'aplicacions que demostren l'ús potencial de la metodologia en el seguiment de sequeres, utilitzant el cas de les precipitacions sobre la conca del riu Colorado (EEUU), i en la gestió dels recursos hídrics, mostrant aplicacions en el sistema del riu Xúquer, sobre la qual s'analitza una política d'operació sobre la base de l'IFS, i el sistema del riu Conchos (EEUU), en la qual es mostra la seva aplicació en la definició de conceptes tals com sequera extraordinària.

Palaures clau

Sequeres, caracterització estocàstica, anàlisi multivariant, anàlisi espacial, teoria de ratxes, índexs de sequeres, mapes de sequeres.

Abstract

Droughts are one of the world's main natural risks. They produce serious damages over the natural environment, economy and society. Drought realizations affect water dependent systems, both natural and artificial. Thus, droughts may be viewed from many points of view, and their consequences might be highly diverse depending on systems vulnerability. However, all drought effects have a common origin: the realization of an extraordinary natural event, which causes a persistent and wide water availability shortage. Because droughts are unpredictable, statistic and stochastic analyses of extreme events are the most suitable tools handling droughts.

Presented here is an approach for the stochastic characterization of extreme hydrologic droughts according to their random nature. The approach is based on the characterization of random variable extreme persistent deviations, referred to as the variable's normal variation regime. The characterization is quantified in terms of the mean frequency of recurrence, providing the basis of a new drought index: the Drought Frequency Index (DFI). This index allows analyzing and evaluating droughts over any random hydro-climatic variable affected by droughts. To complement the approach, a methodology is proposed to analyze a drought spatial-temporal progress over a region by making DFI maps and to characterize droughts from a stochastic point of view, based on their extraordinary persistence and areal extent. In the computation, the DFI does not assume a drought from beginning to end, but it characterizes a dry deviation in terms of the mean frequency of recurrence. Each application may use the DFI scale to define drought state, setting the threshold frequency according to its vulnerability, which will allow homogenizing drought definitions on a unique scale.

In addition to the approach development, the main limitation in drought characterization using other approaches is analyzed and compared with the DFI. The magnitude selection or drought state definition issue using run analyses, the time scale issue in drought evaluations, and the drought indices spatial comparability issues are studied. The potential of the index is evaluated for drought monitoring using the annual Colorado River Basin (USA) distributed precipitation; for water resources management in the Júcar River System (Spain) analyzing a DFI-based operational policy; and, for the Conchos River System (Mexico) defining an extraordinary drought situation.

Key words

Drought, stochastic characterization, multivariate analysis, spatial analysis, runs analysis, drought indices, drought maps

Índice general

1. Introducción	1
2. El concepto de sequía y sus diferentes acepciones funcionales	5
2.1. La definición de sequía	6
2.2. Clasificación de sequías	9
2.3. Indicadores de sequía	11
2.3.1. El Índice de Severidad de Sequías de Palmer	11
2.3.2. Los décimos de población como indicadores de sequía	13
2.3.3. El Índice de Humedad de Cultivos	14
2.3.4. El Índice de Disponibilidad de Agua Superficial . . .	14
2.3.5. El Índice de Precipitación Estandarizado	15
2.3.6. El Índice de Condiciones de la Vegetación	17
2.3.7. El Índice de Disponibilidad de Agua para la Vegetación	17
2.3.8. El Índice Agregado de Sequías	18
3. La modelación estocástica de ocurrencia de sequías	19
3.1. Eventos puntuales en el tiempo frente a eventos no puntuales	19
3.2. Modelación de sequías mediante modelos estocásticos de procesos puntuales	21
3.2.1. Definición y clasificación de procesos puntuales	21
3.2.2. Modelación estocástica de procesos puntuales	23
3.3. La Teoría de Rachas y la caracterización de eventos no puntuales en el tiempo	24
3.3.1. Caracterización univariada de eventos de sequía . . .	27
3.3.2. Caracterización multivariada de eventos de sequía . .	30
3.4. Caracterización espacial de eventos de sequía	34
4. Planteamiento de la problemática y objetivos de la tesis	39
5. Caracterización estocástica de secuencias de realizaciones	43
5.1. Criterios para la caracterización de una secuencia de realizaciones	43

5.1.1.	La función de persistencia extremal	46
5.1.2.	Análisis comparativo de los criterios de caracterización de secuencias	54
5.2.	Propiedades de recurrencia de desviaciones persistentes extremas	59
5.2.1.	Modelación de la recurrencia de desviaciones persistentes extremas a partir de la Teoría de Rachas	59
5.2.2.	Propiedades de las funciones de probabilidad de las duraciones de rachas.	62
5.2.3.	La función de probabilidad acumulada de persistencias extremas	67
5.2.4.	El período de retorno de desviaciones persistentes hacia la región extremal	72
5.3.	Consideraciones en poblaciones con correlación temporal	75
5.3.1.	Transformación de Rosenblatt en variables ARMA	76
5.3.2.	Cálculo de probabilidades conjuntas mediante integración múltiple	80
5.4.	Tratamiento de señales estacionales	82
6.	El Índice de Frecuencias de Sequías	85
6.1.	Definición del Índice de Frecuencias de Sequías	85
6.2.	Análisis comparativo con otros índices y procedimientos	89
6.2.1.	Principales deficiencias metodológicas en la caracterización de sequías	89
6.2.2.	Análisis de sensibilidad en el uso de la Teoría de Rachas	91
6.2.3.	El efecto de la escala temporal	102
6.2.4.	La comparabilidad espacial en los índices de sequías	107
7.	Caracterización espacio-temporal de sequías	113
7.1.	Criterios para la caracterización de una realización espacial aleatoria	114
7.1.1.	Evaluación de la <i>f.a.e.c.</i> mediante análisis de componentes principales	117
7.2.	Caracterización estocástica de realizaciones espacio-temporales de sequías	127
7.2.1.	Algoritmo para la construcción de los mapas de IFS	128
8.	Aplicaciones	133
8.1.	Herramienta para el control espacio-temporal de sequías	133
8.1.1.	El control de las sequías en la cuenca del río Colorado (EEUU)	133
8.2.	Apoyo a la toma de decisiones en la gestión de Recursos Hídricos	152

8.2.1.	El IFS en la política de restricciones a suministros. La gestión de RRHH en la cuenca del río Júcar	153
8.2.2.	El IFS como indicador de sequía extraordinaria. La gestión de RRHH en la cuenca del río Conchos (México)	174
9.	Conclusiones y futuras líneas de desarrollo	193
9.1.	Futuras líneas de desarrollo	197
A.	Métodos numéricos para el cálculo de integrales múltiples	199
A.1.	Métodos numéricos en integrales múltiples de alta dimensión	200
A.1.1.	El método de Monte Carlo	201
A.1.2.	Los métodos de Quasi-Monte Carlo	202
A.2.	Los métodos de Lattice en integrales múltiples de alta dimensión	206
A.2.1.	Reglas de Lattice de Rango-1. El método de los buenos puntos de lattice.	207
A.3.	Variante e implementación del método de lattice para integrales múltiples de <i>f.d.p.</i>	211
A.3.1.	Variante del los métodos de Lattice	211
A.3.2.	Implementación de la variante de los métodos de Lattice	212
B.	Mapas de <i>IFS</i> sobre la cuenca del Río Colorado (EEUU)	215
	Bibliografía	273

Índice de figuras

2.1. Secuencia de los efectos de la sequía, asociados a las sequías meteorológicas, agrícolas e hidrológicas (Adaptada de: National Drought Mitigation Center, EEUU)	10
5.1. Localización de las regiones extremas, en distribuciones de variables aleatorias con dominio de variación infinito y distribuciones gaussianas (a), y variables con dominio de variación semi-infinito y distribuciones exponenciales (b)	48
5.2. Función de probabilidad extremal (ec. 5.27), para diferentes valores de W , y la <i>f.p.a.</i> (en línea discontinua) de una variable $N(\mu = 30, \sigma^2 = 30)$	52
5.3. Dependencia de la <i>f.p.</i> de la duración de periodos \mathcal{N} con d y W , estimada a partir de resolución numérica, para $\hat{F}_C=0.8$	63
5.4. Representación doble logarítmica del cociente entre la <i>f.p.</i> y la <i>f.p.a.</i> de la duración de los períodos \mathcal{S} versus la duración; obtenido mediante resolución numérica, con $\hat{F}_C = 0,8$ y $W = 5, 10, 15, 20$ pasos.	65
5.5. Representación doble logarítmica del cociente entre la <i>f.p.</i> y la <i>f.p.a.</i> de la duración de los períodos \mathcal{S} versus la duración; obtenido mediante resolución numérica, con $\hat{F}_C = 0,2, 0,5, 0,9, 0,99$ y $W = 10$ pasos.	66
5.6. Función <i>f.a.p.e.</i> para diferentes valores de W , en función de \hat{F}_C	71
5.7. Comparación en escala doble logarítmica de los resultados de la estimación del periodo de retorno a partir de analizar la serie sintética, incluyendo el intervalo de confianza del 99% (rojo), frente a la evaluación de la ec. 5.64, para un conjunto de W en el rango $[1,700]$ y de \hat{F}_C en el rango $[0.5,0.9999]$	74
6.1. Diagrama de flujo para la cálculo del IFS en un instante t	86
6.2. Aportaciones trimestrales en el Río Ebro a su paso por Zaragoza y análisis de sequías mediante el IFS.	88
6.3. Aportaciones anuales en el río Blue (Colorado, EEUU).	92

6.4.	Estimación e intervalo de confianza del 95 % de los parámetros de las distribuciones marginales de severidad y factor de uniformidad fijada la duración, conjunto con el parámetro de la distribución de duraciones de sequías, y la parametrización final completa (río Blue, con umbral Q50 %).	95
6.5.	Caracterización trivariada (duración, severidad y coeficiente uniformidad) de las sequías a partir del periodo de retorno (expresado en años) obtenido mediante la Teoría de Rachas, sobre la serie de aportaciones anuales en el río Blue, en función del umbral de definición del estado de sequía (señalados los umbrales de cuantiles del 5 % al 50 %, en intervalos de 5 %)	100
6.6.	Análisis IFS sobre la serie de aportaciones anuales del río Blue.	101
6.7.	Análisis SPI sobre la serie de precipitaciones de la División Climática 5 de California (río San Joaquín), para los períodos de agregación de 3, 6 y 12 meses.	103
6.8.	Análisis <i>IFS</i> de las precipitaciones medias areales sobre la División Climática 5 de California (EEUU), para las escalas temporales trimestral, semestral y anual.	104
6.9.	Detalle de la posición de los ríos Sacramento y San Joaquín sobre EEUU	108
6.10.	Análisis PDSI de las sequías producidas sobre las Divisiones Climáticas 5 y 2 de California, EEUU (fuente: <i>National Climatic Data Center, USA</i>).	109
6.11.	Análisis SPI sobre la serie de precipitaciones de la División Climática 2 de California (río Sacramento), para los períodos de agregación de 3, 6 y 12 meses.	110
6.12.	Análisis IFS de la serie de precipitaciones trimestral sobre la División Climática 2 de California (EEUU)	111
7.1.	Discretización regular de una cuenca donde realizar el cálculo del mapa de IFS.	129
7.2.	Evolución de las agrupaciones en el caso de $C^{-1-}.IFS_t^o > S.IFS_t^o$, ampliándose \mathcal{A}^{-1-} .	130
7.3.	Evolución de las agrupaciones en el caso de $S.IFS_t^o > C^{-1-}.IFS_t^o$, creándose una nueva agrupación \mathcal{A}^{-1-} .	131
7.4.	Evolución de las agrupaciones, con $C^{-1-}.IFS_t^o > C^{-2-}.IFS_t^o$ y $S.IFS_t^o$, y siendo $\{\mathcal{A}^{-1-} \cup \mathcal{A}^{-2-} \cup C^{-i-}\}.IFS_t^o < C^{-1-}.IFS_t^o$. No se produce la unión, ampliándose \mathcal{A}^{-1-} .	132
7.5.	Evolución de las agrupaciones, con $C^{-1-}.IFS_t^o > C^{-2-}.IFS_t^o$ y $S.IFS_t^o$, y siendo $\{\mathcal{A}^{-1-} \cup \mathcal{A}^{-2-} \cup C^{-i-}\}.IFS_t^o > C^{-1-}.IFS_t^o$. Se produce la unión de las agrupaciones \mathcal{A}^{-1-} , \mathcal{A}^{-2-} y la celda C^{-1-} , resultando la agrupación \mathcal{A}^{-3-} .	132

8.1. Cuenca del río Colorado y río Bravo, y su intersección con los estados de EEUU y México.	134
8.2. Mapa de precipitación anual media sobre la cuenca del río Colorado (EEUU), en el período de 1895 a 2003.	135
8.3. Cuencas hidrográficas dentro de la Confederación Hidrográfica del Júcar (Fuente CHJ).	154
8.4. Esquema considerado del Sistema de RRHH del río Júcar.	155
8.5. Sueltas producidas en un sistema de RRHH gestionada mediante una Regla de Operación Normal (a) y mediante una Regla de Operación con Restricciones (b).	162
8.6. Aportaciones naturales anuales históricas en el sistema del río Júcar.	164
8.7. Índice de Frecuencia de Sequías sobre la series de aportaciones históricas totales en los puntos de derivación de suministro del río Júcar.	165
8.8. Muestreo de la función objetivo en el proceso de optimización para la serie histórica de aportaciones, en el sistema del río Júcar, siguiendo el algoritmo SCEM-UA.	166
8.9. Función de distribución posterior de los parámetros a optimizar sobre la serie de aportaciones históricas en el sistema del río Júcar, siguiendo el algoritmo SCEM-UA.	167
8.10. Índice de Frecuencia de Sequías de la serie de aportaciones sintéticas en los puntos de derivación a suministro del sistema del río Júcar.	169
8.11. Muestreo de la función objetivo en la búsqueda del óptimo sobre la serie sintética de 500 años en sistema del río Júcar.	170
8.12. Funciones de distribución posterior de los parámetros a optimizar sobre la serie sintética de 500 años en el sistema del río Júcar.	171
8.13. Muestreo de la función objetivo en la búsqueda del óptimo sobre la serie sintética de 500 años en sistema del río Júcar, con las demandas de riego incrementada en un 20%.	172
8.14. Funciones de distribución posterior de los parámetros a optimizar sobre la serie sintética de 500 años en el sistema del río Júcar, con las demandas de riego incrementadas un 20%.	173
8.15. Precipitación media anual en la cuenca del río Conchos (México).	175
8.16. Cuenca y Red Hidrográfica de río Conchos (México)	176
8.17. Distritos de riego en la cuenca del río Conchos.	179

8.18. Esquema del Sistema de RRHH del río Conchos y esquema de optimización a dos escalas, mensual y quinquenal, con posibilidad de transposición de déficit entre balances quinquenales.	184
8.19. Índice de Frecuencia de Sequías producidos en la situación en los que se ha transpuesto un déficit a EEUU al quinquenio consecutivo en el sistema del río Conchos.	186
8.20. Resultados de la optimización de la regla de operación del sistema del río Conchos, con restricciones función del IFS y un único parámetro.	188
8.21. Resultados de la optimización de la regla de operación del sistema del río Conchos, con restricciones constantes y un único parámetro.	189
8.22. Muestreo de la funciones objetivos en el proceso de optimización de la regla de operación del sistema del río Conchos, con restricciones función del IFS y tres parámetro.	191
8.23. Función de distribución posterior en el proceso de optimización de la regla de operación del sistema del río Conchos, con restricciones función del IFS y tres parámetro	192

Índice de tablas

2.1. Clasificación de los períodos húmedos y secos de acuerdo al PDSI (Palmer, 1965).	13
2.2. Clasificación de los períodos húmedos y secos de acuerdo las Décimas (Gibbs y Maher, 1967).	14
2.3. Clasificación de los períodos húmedos y secos de acuerdo al SPI (McKee, 1993).	16
5.1. Ejemplos de aplicación de los diferentes criterios para la caracterización de secuencias, en una variable lognormal, con umbral para los criterios de rachas en el primer cuartil (30.19). 56	
6.1. Periodo de retorno de eventos de sequía (umbral cuantil medio) sobre las aportaciones anuales del río Blue, en el período 1916-2002.	97
8.1. Estadísticos principales de las aportaciones anuales naturales en el río Júcar de la serie de 1940-1995 (Fuente: PHN, 2000). 156	
8.2. Capacidades de almacenamiento (Hm3) estacional de los embalses principales en el río Júcar (Fuente: PHN, 2000). . . . 156	
8.3. Demandas principales consideradas en el río Júcar (Fuente: PHN, 2000).	158
8.4. Fuentes que componen las disponibilidades de recursos para las distintas demandas de riego en el sistema del río Júcar. . 161	
8.5. Valores óptimos de los coeficiente de restricciones sobre la serie histórica en el sistema del río Júcar.	165
8.6. Garantías al abastecimiento urbano en el sistema del río Júcar obtenidas tras la optimización de los coeficientes de restricciones sobre la serie histórica.	166
8.7. Frecuencia de cumplimiento de las condiciones del criterio de Utah en el suministro a las zonas de regadío del sistema del río Júcar tras la optimización de los coeficientes de restricciones sobre la serie histórica.	166
8.8. Embalses y subcuencas principales en el sistema del río Conchos.	177

8.9. Parámetros del modelo de Temez para cada una de las subcuencas principales en del río Conchos.	178
8.10. Coeficientes de cultivos y área media ocupada por mes en el distrito de riego de río Florido.	180
8.11. Coeficientes de cultivos y área media ocupada por mes en el distrito de riego de Las Delicias.	180
8.12. Coeficientes de cultivos y área media ocupada por mes en el distrito de riego de Bajo Río Conchos.	180
8.13. Demanda media anual computada por cada distrito de riego en el río Conchos.	181

Capítulo 1

Introducción

De acuerdo con la Organización Mundial de Meteorología, sequía se define como una deficiencia en la precipitación, sostenida y extensa. (WMO, 1986). Sin embargo, esta definición únicamente recoge una acepción de las múltiples definiciones de sequía que se encuentran en la literatura científica. Por ejemplo, Wilhite y Glantz (1985) presentaron un análisis sobre más de 150 definiciones. La falta de consenso con respecto a la definición de sequía propicia una dispersión en los trabajos de investigación que tratan este fenómeno natural, reduciendo los avances en la consecución de un tratamiento general y unificado (Yevjevich, 1967). La dispersión en las múltiples acepciones de sequía está motivada fundamentalmente por la particularización a la percepción concreta que cada afectado tiene. Los fenómenos de sequía tienen su impacto sobre múltiples sistemas reales, todos aquellos que dependen de la presencia de agua. Así las primeras afecciones se producen en la vegetación y la agricultura, por la reducción en la humedad del suelo, seguido progresivamente de reducción en los caudales naturales en ríos, disminución de los almacenamientos en lagos y embalses, retroceso de humedales, descenso de los niveles en acuíferos, etc. Esta afección progresiva produce efectos sobre el medio ambiente, los suministros para consumo humano, industrial, o agrícola, sistemas de producción de energía, etc. Sus efectos pueden ser reversibles o incluso irreversibles, como es el caso de los procesos de desertificación natural. Esto coloca a los fenómenos de sequía como uno de los mayores riesgos naturales del planeta.

La aparición de un fenómeno de sequía puede venir originada de forma natural, considerado como un proceso aleatorio, pero también su realización puede hacerse más frecuente como consecuencias de acciones antrópicas, no ya referida a la forma de gestión de los recursos, sino por los cambios de uso sobre las cuencas, que modifican las interacciones entre la circulación atmosférica y la superficie terrestre. El aumento progresivo de las demandas de agua, añadido a la transformación del territorio está produciendo que

cada vez sea más frecuente la aparición de situaciones deficitarias de agua, consideradas como sequías.

Sin embargo, este aumento de la frecuencia en la aparición de sequías no es sino una modificación del régimen de variación de los recursos, añadido a un aumento de las demandas. En este sentido, no deben ser consideradas como tales, entendiendo sequía como un fenómeno hidrológico extremo, sino que deben ser entendidas como variaciones permanentes en las fuentes de suministros que hacen más vulnerables a los sistemas usuarios. Es esta mayor vulnerabilidad de los sistemas usuarios lo que está motivando un esfuerzo a nivel internacional por el tratamiento de las sequías, con el objeto de mitigar sus efectos. Como ejemplo particular en España, la ley del Plan Hidrológico Nacional aprobada en 2001 (10/2001, de 5 de julio de 2001) dedica su Artículo 27 a la Gestión de Sequías. En dicho artículo se atiende a la necesidad de establecer unos indicadores de sequías, que permitan la alerta y declaración de sequía, y de unos Planes Especiales y de Emergencia de contingencia frente a sequías.

En este marco general, la Tesis Doctoral que aquí se introduce, con título "Caracterización Estocástica de la Realización Espacio-Temporal de Eventos Hidrológicos Extremos de Sequías", pretende contribuir al adecuado tratamiento del fenómeno, realizando un mayor énfasis en el análisis del proceso aleatorio extremo que tiene lugar, y que origina graves efectos de diversa naturaleza. Esta Tesis tiene como principal contribución la presentación de una metodología que permita la caracterización estocástica de eventos extremos como son las sequías, caracterizando su desviación persistente con respecto al régimen de variación estadística y expresándolo en términos de su recurrencia media de aparición. La metodología se concreta en un nuevo índice de sequía, denominado Índice de Frecuencia de Sequía (*IFS*), que permite analizar y evaluar la ocurrencia de sequía tanto de la situación que en cada momento está teniendo lugar, como de los eventos pasados registrados en las series históricas. El índice es aplicable sobre cualquier variable hidroclimática que en su señal manifieste la realización de una sequía como una desviación temporal respecto de su régimen de variación normal. Así mismo, manteniendo un tratamiento estocástico, se completa la metodología con una sistemática para la elaboración de mapas de *IFS* que permitan representar y caracterizar el desarrollo espacial y temporal de las sequías como eventos aleatorios extremos.

La exposición de la metodología se complementa con el análisis de las principales limitaciones existentes en la caracterización general de sequía con las metodologías disponibles. Así se analizan los efectos de la selección de diferentes magnitudes y definiciones del estado de sequía en la aplicación de metodologías de caracterización estocástica basadas en la Teoría de Rachas. Los resultados del análisis de sensibilidad en estas aplicaciones

son comparados con el uso del *IFS*. Por otro lado, se analiza el efecto de la escala temporal sobre los análisis de sequía, comparando los resultados que se obtienen con el Índice de Precipitación Estandarizado, que trata cada escala temporal por separado, con la aplicación del *IFS*, que muestra menos sensibilidad a variaciones en la escala. Por último, se trata el problema de la comparabilidad espacial de los resultados de caracterización, demostrando las deficiencias de otros índices ampliamente utilizados como el Índice de Severidad de Sequías de Palmer, frente a índice de naturaleza probabilista, como es el *IFS*.

A continuación de la exposición y análisis se abordará la presentación de aplicaciones prácticas, con las que se pretende poner de manifiesto la aplicabilidad de los resultados obtenidos en casos prácticos reales. Así se muestra el uso potencial de los mapas *IFS* en el seguimiento espacio-temporal de sequías, utilizando el caso de la cuenca del río Colorado (EEUU) donde se produce una gran diversidad climática, comparando resultados con otros sistemas de control de sequías. En su potencial aplicación a sistemas de recursos hídricos (RRHH), se presenta el caso de aplicación sobre la cuenca del río Júcar, de múltiples embalses y demandas, sobre la que se formula una política de operación de los embalses en la que entra en juego el valor del *IFS* en los principales puntos de suministro. El último caso de aplicación es el de la cuenca del río Conchos (México) sobre el que además de definirse una política de operación de sus embalses en función del *IFS*, se propone el índice como escala de definición del concepto de sequía extraordinaria que contempla el tratado bilateral de aguas internacionales entre EEUU y México del año 1941 al que está sometida la gestión, pero que no define.

La estructura de esta memoria de Tesis comienza con una revisión de la literatura científica y del estado del arte entorno al tratamiento de las sequías y las metodologías disponibles para su caracterización en los Capítulos 2 y 3. El Capítulo 3 incluye además la generalización de un modelo estocástico de recurrencia siguiendo la Teoría de Rachas para el caso multivariado. El Capítulo 4 aborda la concreción de los objetivos de la Tesis en base al planteamiento de la problemática. El Capítulo 5 presenta la metodología para la caracterización de desviaciones persistentes extremas, que da lugar a la definición del *IFS* en el Capítulo 6. En este Capítulo son además analizadas las limitaciones en la evaluación y caracterización de sequías. En el Capítulo 7 se expone la metodología para la caracterización espacial de sequías a partir de la elaboración de mapas *IFS*. Las aplicaciones del conjunto de la metodología expuesta se presentan en el Capítulo 8. Por último, el Capítulo 9 recoge las conclusiones del trabajo de investigación desarrollado y propone futuras líneas de desarrollo.

Capítulo 2

El concepto de sequía y sus diferentes acepciones funcionales

En la naturaleza se producen múltiples fenómenos que llevan asociados riesgos para la población humana e impactos en el medio ambiente. Así destacan entre los principales fenómenos naturales de riesgo las inundaciones, las sequías, los sismos o las erupciones volcánicas. Los elevados daños que producen están en gran medida provocados por la difícil anticipación a su realización. Para el estado actual de conocimientos, estos fenómenos aparecen de modo aleatorio, requiriendo en su tratamiento del análisis estadístico y estocástico. Las sequías pueden verse como los fenómenos naturales de riesgo asociados a una excepcional reducción en la presencia de agua durante un tiempo prolongado y afectando una amplia extensión. No existe sin embargo una definición consensuada de sequía, resultando un fenómeno complejo y difícil de definir en forma precisa. La complejidad de las sequías se fundamenta en la complejidad del sistema o sistemas sobre el que tiene su efecto.

Aunque no es el objeto de esta Tesis el análisis en profundidad del concepto de sequía y sus múltiples acepciones, en este capítulo se pretende establecer el contexto sobre el que se desarrolla el trabajo investigación, lo que permite, así mismo, exponer su motivación. En primer lugar, se realizará un análisis de la definición de sequía, presentando las diferentes acepciones que esta puede tener y aquella por la que será entendida aquí. Seguidamente se expondrán diferentes clasificaciones de sequías que permiten profundizar en la complejidad del concepto. Por último, se desarrollará una revisión de los principales indicadores de sequía que son encontrados en la literatura científica.

2.1. La definición de sequía

Aunque no existe una definición universalmente aceptada del concepto de sequía, sí puede encontrarse un conjunto de aspectos generales que la caracterizan. Consultando algunos de los principales diccionarios se encuentran las siguientes definiciones:

- Diccionario de la Real Academia de la Lengua Española: Tiempo seco de larga duración
- Diccionario Oxford: período prolongado de baja precipitación poco frecuente (*prolonged period of abnormally low rainfall*)
- Diccionario Merriam-Webster:
 - 1 : período prolongado de sequedad, que causa extenso daños a los cultivos y evita su completo desarrollo (*a period of dryness especially when prolonged that causes extensive damage to crops or prevents their successful growth*)
 - 2 : una escasez prolongada o crónica de algo esperado o deseado (*a prolonged or chronic shortage or lack of something expected or desired*)

Aunque no coinciden entre si ninguna de las definiciones, pueden encontrarse ideas repetidas: en todos los casos la sequía se relaciona con una reducida presencia de agua; este estado de reducida disponibilidad viene referido a una situación que se puede considerar normal, resultando el bajo nivel alcanzado poco frecuente; y siempre se pone como condición que esa situación sea persistente, es decir, se mantenga de forma notoria en el tiempo.

Por otro lado, la Organización Mundial de Meteorología define la sequía como una deficiencia en la precipitación, sostenida y extensa (*a sustained, extended deficiency in precipitation*, (WMO, 1986)). Asimismo, como una definición ampliamente referida, destaca la propuesta por Huschke (1959): tiempo anormalmente seco, para un área específica, cuya escasez de agua es suficientemente prolongada para producir serios desequilibrios hidrológicos (*abnormal dry weather for a specific area that is sufficiently prolonged for the lack of water to cause serious hydrological imbalance*). En estas definiciones aparecen dos nuevas ideas: por un lado, se hace referencia a la extensión del fenómeno, entendiéndose que ocupa una amplia superficie; y por otro lado se establece el carácter específico para cada área, dependiendo de la zona las condiciones que indican una situación de sequía pueden ser diferentes. Esta última idea se puede considerar similar a la reseña a una situación normal,

específica de cada lugar, que se toma como referencia para manifestar su grado de excepcionalidad.

Por lo tanto, con carácter general puede decirse que el concepto de sequía lleva asociadas las siguientes ideas:

- Supone una reducida presencia de agua
- El nivel reducido se refiere con respecto a las condiciones que son consideradas normales, esperadas o más probables en el lugar donde se produce
- Este estado se mantiene durante un período prolongado de tiempo, y ocupa una extensa área, resultando extraordinaria su realización

Todas o casi todas estas ideas se encuentran incluidas en diferente manera en las múltiples definiciones realizadas de sequía, la mayoría de ellas de tipo particular, para un lugar y fin determinado. Wilhite y Glantz (1985) realizaron un análisis sobre más de 150 definiciones de eventos sequía, clasificándolas en diferentes formas. De ellas las principales se encuentran recogidas en los trabajos de Dracup *et al.* (1980a), y Tate y Gustard (2000). La amplia variedad en definiciones de sequías tiene dos orígenes fundamentalmente. El primero de ellos es la **particularización para cada lugar**. Así, un período de varios días consecutivos sin producirse precipitación en un clima húmedo resulta poco frecuente y susceptible de considerarse sequía. Si una definición de este tipo quiere trasladarse a un clima semiárido, el período a considerar deberá ser de meses para que resulte un fenómeno poco frecuente. Incluso puede completarse la definición incluyendo la dependencia estacional del régimen de precipitación. Por otro lado, la definición se puede particularizar al encontrarse **asociada a un determinado uso** o función del agua. Esto es, definiciones que están orientadas hacia el efecto o la manifestación de la sequía, y no a la realización como fenómeno físico. Ejemplos de este tipo de criterio se encuentra en la referencia de sequía como situación en la que la disponibilidad de agua por los cultivos no es la suficiente como para permitir un completo desarrollo; o en el caso de estar extrayéndose un caudal de un río natural sin regulación artificial, aquellos períodos donde el caudal que circula es inferior al caudal demandado.

En algunas ocasiones, una definición de sequía referida a la aplicación a una determinada función del agua que se extrae de un sistema natural, se confunde o se solapa con los conceptos de aridez, y escasez o insuficiencia. El concepto de aridez, desde un punto de vista climatológico se refiere a la ineficiencia de la precipitación, relacionada con el ratio entre la altura de precipitación y la evapotranspiración potencial, y por tanto referida a un uso, el uso del agua por la vegetación. Un período es árido cuando la

relación entre la altura de precipitación y la evapotranspiración potencial es menor que la unidad. Los períodos de aridez se producen en los climas áridos y semiáridos de forma periódica, durante la estación o estaciones con menor precipitación. Estos períodos de tiempo en algunas ocasiones se refieren a períodos de sequía, cuando en realidad no mantienen el carácter extraordinario. En lo que respecta al concepto de escasez o insuficiencia, se produce cuando una demanda se sitúa por encima de una disponibilidad de recurso frecuente. Por ejemplo, si en el caso de las extracciones de un río natural sin regulación artificial la demanda se iguala al caudal medio, entendiendo que el caudal será una variable aleatoria, solo podrá satisfacerse la totalidad de la demanda durante el 50 % del tiempo, el resto del tiempo se producirá déficit. Esta es una situación general de escasez o insuficiencia del recurso, aunque en algunos casos también son denominados los períodos de déficit como períodos de sequía, sin que tampoco contengan el carácter de situación excepcional.

Buscando rigor en la definición, estas acepciones relativas a aridez y escasez no van a ser consideradas dentro del concepto de sequía. Por otro lado, el abanico de definiciones que supone la particularización para cada lugar y régimen de variación concreto y la asociación a la insatisfacción de uso determinado del agua tienen un origen común general para cualquier definición: su carácter extraordinario. Al encontrarse siempre el origen del agua en un sistema natural, la cantidad en la que se presenta es una variable aleatoria. En cada caso, el régimen de variación de la variable podrá ser diferente, pero el uso de cuantías probabilistas permite la valoración de su carácter excepcional. Por este motivo, la particularización para cada caso encuentra una vía de homogenización concretando la definición en términos de probabilidad, mediante caracterización estadística, o en su reflejo sobre la escala temporal en términos de frecuencia de recurrencia, mediante caracterización estocástica. Las referencias en las múltiples definiciones propuestas a los daños excepcionalmente producidos por la insatisfacción de una determinada demanda, apartando los casos de insuficiencia del recurso, también mantienen como un exponente común la realización de un episodio excepcional de reducida presencia de agua. En ese sentido, la concreción de estas definiciones pasa por el uso riguroso de las probabilidades o frecuencias de aparición. La vulnerabilidad del sistema usuario fijará el nivel para el que se considera sequía, como aquél que provoca daños significativos.

La realización de una reducida presencia de agua podrá ser extraordinaria, no sólo por su bajo valor, sino tanto más por persistir esa situación de forma prolongada en el tiempo. Además, en un análisis espacial podrá tener un carácter más excepcional cuanto mayor sea la extensión espacial sobre la que se manifiesta. De este modo, y a partir del concepto general de sequía antes expuesto, la definición en una base uniforme de un estado de sequía

pasa por una caracterización espacio-temporal o bien estadística, o bien estocástica, de los períodos extremos que suponen las sequías. Esto motiva el desarrollo de esta Tesis, a lo largo de la cual se justificará la conveniencia de una caracterización estocástica frente a una estadística.

2.2. Clasificación de sequías

Las sequías tienen su origen dentro del ciclo hidrológico en una deficiente precipitación, lo cual provoca de forma directa una disminución en la presencia de agua en todos los estados del ciclo hidrológico (Brass, 1990). A esta deficiente precipitación puede sumarse factores como una alta temperatura, la presencia de fuertes vientos, una baja humedad relativa o una reducida nubosidad con una radiación solar terrestre alta. Todos ellos tienen en común el aumento en la evaporación y transpiración que acentúa la disminución en la presencia de agua en la componente terrestre del ciclo hidrológico.

De este origen común se derivan múltiples efectos sobre los sistemas naturales o artificiales asociados, que dan lugar a las distintas definiciones de sequía. Se han propuesto diferentes clasificaciones que catalogan estas definiciones orientadas hacia el efecto producido. Por ejemplo, Wilhite y Glantz (1985) propusieron diferenciar entre definiciones de sequía de tipo **conceptual** y **operacional**. Las primeras son aquellas formuladas en términos generales, no aplicables directamente a la evaluación de la situación de sequía. Las segundas, por contra, si permiten de forma práctica evaluar una situación de sequía, y en muchos casos permiten la delimitación de su inicio, fin e incluso grado de severidad.

Otra clasificación extensamente utilizada es la propuesta por Dracup *et al.* (1980a). En este caso las sequías son clasificadas desde un punto de vista disciplinar:

- **Sequía meteorológica:** descrita como una deficiente precipitación, con respecto a una situación normal, y manifestada de forma sostenida en el tiempo.
- **Sequía agrícola:** analizada desde su efecto sobre la humedad disponible en el suelo para las plantas. En muchos casos valorada de acuerdo con los daños que provoca sobre las producciones agrícolas.
- **Sequía hidrológica:** evaluada sobre los flujos o almacenamientos del ciclo hidrológico terrestre, al margen del almacenamiento en forma de humedad en el suelo. Generalmente referida a sus efectos sobre los puntos donde tiene lugar el uso del agua o es traído el recurso.

Estando el origen del fenómeno en las sequías meteorológicas, la subsecuente aparición de los otros tipos de sequía depende del grado en el que la sequía meteorológica se manifieste y de la vulnerabilidad del sistema usuario. Por vulnerabilidad se entiende la gravedad de los daños que puedan producirse en el sistema a causa de un recurso insuficiente, durante un período prolongado. Los sistemas agrícolas son generalmente los más vulnerables, y es en los que primero se manifiesta una sequía sobre los sistemas usuarios. La limitada capacidad de almacenamiento de agua en las primeras capas del suelo, acompañado de los flujos a los que está sometido en forma de evapotranspiración y percolación, convierten los sistemas agrícolas en un sistema vulnerable a falta de fuentes adicionales de humedad como pueden ser los riegos. Por otro lado, dentro del resto de elementos que pueden encontrarse en el ciclo hidrológico (lagos, cauces, acuíferos, humedales,...) su vulnerabilidad depende de su capacidad de almacenamiento en relación al flujo medio que reciben. Los últimos elementos que sufren la influencia de la sequía son los acuíferos, que suelen mantener una inercia elevada. La propagación de una sequía meteorológica hacia las otras formas de sequía no solamente depende de la persistencia del fenómeno, sino también de la extensión sobre la que tiene lugar. La afección a la disponibilidad de agua en los puntos donde tiene lugar su uso o donde se produce su extracción puede provocar impactos económicos, sociales, y/o medio ambientales (figura 2.1).

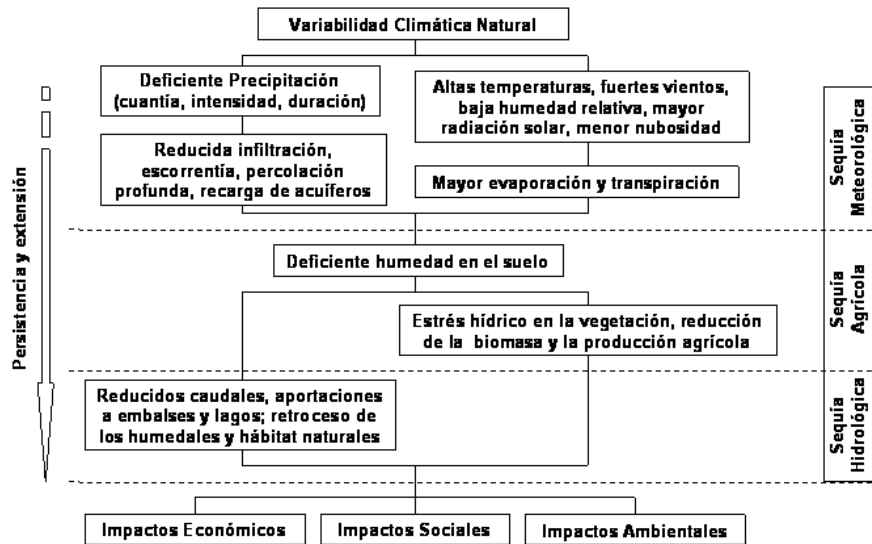


Figura 2.1: Secuencia de los efectos de la sequía, asociados a las sequías meteorológicas, agrícolas e hidrológicas (Adaptada de: National Drought Mitigation Center, EEUU)

De forma complementaria a esta clasificación, Wilhite y Glantz (1985) añadieron las sequías **socio-económicas**. En este grupo se engloban aquellas definiciones de sequías que están directamente orientadas hacia la satisfacción de una determinada demanda que produce un bien económico o social. La realización de una sequía socio-económica puede venir producida a consecuencia de una sequía meteorológica, como la realización de un fenómeno aleatorio extraordinario, o motivada de modo artificial por la gestión inadecuada de los recursos.

Una clasificación adicional, también disciplinar y de gran similitud con la anterior, es la propuesta por Tate y Gustard (2000) que divide las sequías en climatológicas, agrometeorológicas, fluviales y subterráneas.

2.3. Indicadores de sequía

Las múltiples definiciones operativas de situación de sequía, cada una de utilidad en el lugar y para el fin con el que fue creada, producen una amplia variedad en las definiciones operativas de sequía. Esta amplia variedad dificulta la tarea de análisis y comparación de criterios por su falta de homogeneidad. Esto ha motivado la creación de índices o indicadores de sequía que buscan proporcionar una escala sobre la que uniformizar los criterios. Los índices de sequía pueden ser función de una o varias variables, generalmente de tipo hidroclimáticas (e.g. precipitación, temperatura, aportaciones,...), y no están directamente relacionados con el valor de una demanda. A continuación se presenta una relación de los índices de sequías más destacados. Sobre los más generales se realizará en el capítulo 6 un análisis de los principales problemas en la caracterización de sequías a partir de índices.

2.3.1. El Índice de Severidad de Sequías de Palmer

Entre los índices de sequía más conocidos destaca como el más utilizado el Índice de Severidad de Sequías de Palmer (*Palmer Drought Severity Index*, PDSI). El índice se clasifica como un índice de sequías meteorológicas, y está basado en un modelo hidrológico conceptual simplificado del proceso de almacenamiento del agua en el suelo, producción de escorrentía y evapotranspiración. El PDSI cuantifica en una escala adimensional las desviaciones en la humedad del suelo con respecto de unas condiciones consideradas estándares (Palmer, 1965), de modo que pueden realizarse comparaciones entre valores obtenidos en diferentes lugares, u obtenidos en diferentes meses.

En el cómputo del índice entran en juego no solamente la precipitación, sino también la temperatura y el contenido de humedad disponible en el

suelo (*Available Water Content*, AWC). Adicionalmente, para el cálculo del índice se requiere la fijación de un parámetro local que indica la capacidad máxima de almacenamiento de humedad en el suelo. En el índice no entran en juego referencias a usos del agua o demandas, ni se consideran en el balance los impactos que la actividad humana pueda tener.

Palmer buscó incluir en el índice el efecto de la duración de la sequía y su desarrollo temporal. Así entre sus objetivos para la definición del índice se encontraba que un mes anormalmente húmedo dentro del desarrollo de una sequía de gran duración, no suponga un gran impacto en el índice; o tras una severa sequía, una serie de meses con unas precipitaciones próximas a las condiciones normales no supongan la declaración del fin de la sequía. De este modo el valor del índice en cada momento depende no sólo de las condiciones de humedad presentes, sino principalmente de las condiciones previamente producidas, y en menor medida de las inmediatas futuras. El cálculo se concreta a partir del algoritmo propuesto por Palmer (Palmer, 1965). En análisis en tiempo real, aparece como derivación simplificada del PDSI un índice de tipo hidrológico, que caracteriza la situación de sequedad a partir de los flujos y almacenamiento en el modelo hidrológico conceptual, pero sin tener en cuenta la tendencia pasada. Este índice se denomina el Índice de Sequías Hidrológicas de Palmer (*Palmer Hydrological Drought Index*, PHDI) (Karl & Knight, 1985).

El PDSI varía de forma aproximada entre $-6,0$ y $+6,0$, si bien sus límites no están acotados. Palmer selecciono de forma arbitraria la escala de clasificación de las condiciones de humedad basada en su área de estudio en el momento de la definición del índice, localizada en la zona centro del estado de Iowa y el oeste de Kansas (EEUU). La tabla 2.1 recoge la clasificación de los períodos secos y húmedos en función del PDSI. De acuerdo a su diseño, idénticos valores de PDSI producidos en dos lugares diferentes del área de estudio indican el mismo grado de desvío respecto del régimen normal de variación de la humedad. El PDSI se calcula en general a escala mensual, aunque puede calcularse a otras escalas.

El índice de Palmer se utiliza ampliamente para el control de la realización de sequía a lo largo de todo EEUU. Así mismo ha sido aplicado en análisis de sequías en otros países (Cabrinha & Santo, 2000; Loukas *et al.*, 2002). Alley (1984) identificó tres motivos que contribuyen a la popularidad del PDSI: proporciona una medida del grado de excepcionalidad de las condiciones recientes en una región, permite su comparación con las diferentes condiciones que se han producido a lo largo de la historia y facilita la comparación temporal y espacial de las sequías históricas. A pesar de su popularidad, existen serias limitaciones en el uso del PDSI. Entre las principales se encuentran (Alley, 1984; Karl, 1986): la falta de rigor científico en la definición de la escala del índice, ajustada experimentalmente

PDSI	Clasificación
4 ó mayor	Extremadamente húmedo
3.0 a 3.99	Muy húmedo
2.0 a 2.99	Moderadamente húmedo
1.0 a 1.99	Ligeramente húmedo
0.5 a 0.99	Incipiente período húmedo
0.49 a -0.49	Normal
-0.5 a -0.99	Incipiente período seco
-1.0 a -1.99	Ligeramente seco
-2.0 a -2.99	Moderadamente seco
-3.0 a -3.99	Muy seco
-4 ó menor	Extremadamente seco

Tabla 2.1: Clasificación de los períodos húmedos y secos de acuerdo al PDSI (Palmer, 1965).

para la región de estudio utilizada en su definición; el índice es sensible a la AWC de cada tipo de suelo, por tanto su uso a una gran escala espacial puede resultar muy general; el modelo hidrológico conceptual utilizado es muy simplificado, y puede no resultar una representación precisa de una zona; la evapotranspiración potencial se estima haciendo uso del método de Thorthwaite, que resulta sólo una aproximación simple.

Otras limitaciones fueron expuestas por diferentes autores. Entre ellos destacan cuestiones sobre la aplicabilidad del índice fuera de EEUU, sugiriendo que el PDSI no es aplicable en regiones donde se producen fuerte variabilidad en precipitación y escorrentía (Willeke *et al.*, 1994). Adicionalmente se encuentra que el índice clasifica eventos de sequía como extremos o severos con diferente frecuencia en aplicaciones sobre diferentes regiones, incluso dentro de EEUU, lo que supone una limitación para su uso en comparaciones espaciales (Guttman *et al.*, 1992).

2.3.2. Los décimos de población como indicadores de sequía

Agrupando las precipitaciones mensuales ocurridas en diez grupos, por orden de magnitud, permite la catalogación de desviaciones con respecto de un comportamiento normal. Cada grupo, o cada décimo (*decile*), está compuesto por un 10 % de la población, de acuerdo a su distribución estadística (Gibbs & Maher, 1967). El primer décimo se corresponde con la precipitación que no se excede por más del 10 % de la precipitación ocurrida. El último décimo se encuentra sobre el mayor de las precipitaciones ocurridas. Por definición, el quinto décimo coincide con la mediana. Los décimos se

Décimas	Clasificación
Décimas 1-2 (20 % menor)	Muy por debajo de lo normal
Décimas 3-4 (20 % siguiente)	Por debajo de lo normal
Décimas 5-6 (20 % medio)	Normal
Décimas 7-8 (20 % siguiente)	Por encima de lo normal
Décimas 9-10 (20 % mayor)	Muy por encima de lo normal

Tabla 2.2: Clasificación de los períodos húmedos y secos de acuerdo las Décimas (Gibbs y Maher, 1967).

agrupan en cinco clasificaciones (tabla 2.2).

Este método fue seleccionado por el Sistema de Alerta de Sequías de Australia (*Australian Drought Watch System*) por su simplicidad y el menor número de hipótesis asumidas que el PDSI. De acuerdo con este sistema, los agricultores pueden solicitar ayudas gubernamentales si se produce una sequía que ocurre una vez cada 20-25 años (décimos 1 y 2) y persiste durante más de 12 meses (White & O’Meagher, 1995).

2.3.3. El Índice de Humedad de Cultivos

El Índice de Humedad de Cultivos (*Crop Moisture Index*, CMI) fue propuesto por Palmer (1968) para el control de las condiciones de humedad en el suelo con un paso temporal semanal. El desarrollo del índice se encuentra dentro del procedimiento seguido para el cálculo del PDSI. El índice está basado en la temperatura media y la precipitación producida en cada semana dentro de la región de análisis, así como del valor del CMI producido en la semana previa. El CMI responde rápidamente a cambios, y es ponderado para cada localización y época del año, de modo que permite la construcción de mapas semanales de las condiciones de humedad para la comparación espacial.

La rapidez con la que cambia el sentido del índice no lo hacen adecuado para su uso en el seguimiento de sequías persistentes. Otro aspecto del índice que limita su aplicación en análisis de largas rachas de sequías es que el índice tiende a cero en los comienzos y finales de los períodos de crecimiento de los cultivos. Esta limitación lo convierte en inadecuado para su evaluación fuera del período de desarrollo de los cultivos.

2.3.4. El Índice de Disponibilidad de Agua Superficial

El Índice de Disponibilidad de Agua Superficial (*Surface Water Supply Index*, SWSI) fue desarrollado por Shafer y Dezman (1982) para complementar el PDSI en el estado de Colorado (EEUU). El PDSI es aplicable en regiones relativamente homogéneas, pero no fue diseñado para su uso sobre

regiones montañosas y no tiene en cuenta las acumulaciones de nieve y su influencia posterior en las aportaciones. El objetivo del SWSI fue englobar los aspectos climatológicos e hidrológicos en un único índice adaptando el PDSI para las principales cuencas del estado de Colorado (EEUU). Los valores del índice están estandarizados, permitiendo la comparación de las condiciones que se producen en diferentes cuencas. El cómputo del índice depende de la estación del año. Así durante el invierno el índice depende de la nieve acumulada, las aportaciones y el almacenamiento en lagos y embalses; mientras que en el resto del año la nieve acumulada se sustituye en el cómputo por las aportaciones.

El procedimiento para la determinación del SWSI en una determinada cuenca es el siguiente: a partir de datos mensuales, se agregan con la frecuencia deseada los datos de precipitación, niveles en embalses, y altura de nieve ó aportaciones. Cada componente agregada se normaliza haciendo uso de un análisis de frecuencias, a partir de una serie que proporcione una muestra representativa. La probabilidad de no excedencia se determina para cada componente a partir del análisis de frecuencia. Cada componente tiene un peso asignado de acuerdo a su contribución al volumen de agua superficial dentro de la cuenca, de modo que una suma ponderada determina el valor del SWSI representativo del conjunto de la cuenca. De forma similar al PDSI, el SWSI está centrado en el cero, y varía en el rango entre -4.2 a 4.2.

Varias características del índice limitan su aplicación, entre ellas se encuentra que debido a que para cada cuenca difiere la ecuación que proporciona el SWSI, es difícil la comparación de valores del índice obtenidos en diferentes cuencas o regiones. Adicionalmente, en una cuenca de estudio, cambios en el conjunto de puntos de medida considerados (e.g. embalses) implica la necesidad de nuevas distribuciones de frecuencia. Además el índice está influenciado por la gestión producida sobre la cuenca, cuya modificación puede requerir de reajustes en los pesos asignados a cada componente. Esto hace que sea difícil mantener una serie temporal del SWSI homogénea (Heddinghaus & Sabol, 1991).

2.3.5. El Índice de Precipitación Estandarizado

La idea de que un déficit en precipitación produce diferentes impactos sobre acuíferos, lagos, humedad en el suelo, acumulaciones de nieve o las aportaciones condujo a proponer a Mckee *et al.* (1993) el Índice de Precipitación Estandarizado (*Standardized Precipitation Index*, SPI). El SPI fue diseñado para cuantificar deficiencias de precipitación en múltiples escalas temporales, cada una de ellas reflejo del impacto de una sequía en la disponibilidad de agua en las diferentes fuentes de recursos hídricos. De

SPI	Clasificación
2 ó mayor	Extremadamente húmedo
1.5 a 1.99	Muy húmedo
1.0 a 1.49	Moderadamente húmedo
-0.99 a 0.99	Normal
-1.0 a -1.49	Moderadamente seco
-1.5 a -1.99	Muy seco
-2 ó menor	Extremadamente seco

Tabla 2.3: Clasificación de los períodos húmedos y secos de acuerdo al SPI (Mckee, 1993).

este modo, la condición de humedad en el suelo responde a anomalías en la precipitación en una escala temporal relativamente corta. Por otro lado, el nivel de almacenamiento en acuíferos, lagos, o las aportaciones son un reflejo de anomalías en la precipitación en una escala temporal mayor. Ello motivó que en la definición original del SPI fuera propuesto su cálculo para períodos de 3, 6, 12, 24 y 48 meses de duración.

El cálculo del SPI se realiza a partir de un registro suficientemente largo de precipitaciones que permita la caracterización estadística de la precipitación caída en los diferentes períodos de agregación. Para cada uno de estos períodos de agregación, se ajusta una distribución estadística a la muestra, y con ella se transforman los valores que componen la serie en la variable aleatoria equivalente correspondiente a una distribución normal estándar. Valores positivos del SPI indican precipitaciones por encima de la media, mientras que valores negativos reflejan precipitaciones por debajo de la media. El SPI permite el análisis de los períodos secos y húmedos, debido a la simetría de la variable transformada.

A partir de los valores del SPI, es posible clasificar los períodos húmedos y secos resultantes en los análisis a las diferentes escalas (tabla 2.3). Un evento de sequía ocurre cuando durante un período el SPI es continuamente negativo, alcanzando un valor de -1 ó inferior. El evento termina cuando el índice pasa a ser positivo. De este modo, cada evento de sequía tiene un principio y un final, y una intensidad por cada paso temporal dentro del evento. La suma de los valores producidos en cada paso por el SPI puede ser utilizada como una magnitud de la sequía.

A partir del análisis de registros producidos en estaciones de la cuenca de Colorado, Mckee determinó que el SPI aparece como normal el 24 % del tiempo, moderadamente seco el 9.2 %, muy seco el 4.4 % y extremadamente seco el 2.3 % (McKee *et al.* , 1993). Estos porcentajes coinciden con los esperados de una distribución normal del SPI, al haber sido estandarizado. Situaciones extremadamente secas aparecen el 2.3 % del tiempo, mientras que a partir del uso del PDSI esta categoría resulta más del 10 % del tiempo

en las regiones utilizada para la diseño del PDSI. La estandarización del SPI permite cuantificar el grado de excepcionalidad de una sequía, así como la probabilidad de producirse la precipitación necesaria para dar por terminada una sequía en una situación determinada (McKee *et al.* , 1993).

2.3.6. El Índice de Condiciones de la Vegetación

En la aplicación de nuevas fuentes de información en el control de sequías destaca el Índice de Condiciones de la Vegetación (*Vegetation Conditions Index*, VCI) que se basa en el uso de la información obtenida por el satélite AVHRR. El índice utiliza los datos de radiación (visible y próxima a infraroja), ajustados de acuerdo al clima de la superficie, la ecología y las condiciones meteorológicas (Kogan, 1995). La detección y la catalogación de las sequías se basa en la alta dependencia de la vegetación sobre la meteorología, permitiendo medir la duración, intensidad e impacto sobre la vegetación. En la actualidad se usa principalmente para el control de sequías en EEUU.

2.3.7. El Índice de Disponibilidad de Agua para la Vegetación

El Índice de Disponibilidad de Agua para la Vegetación (*Water Supplying Vegetation Index*, WSVI) representa otra aplicación importante de la información disponible a través de los satélites meteorológicos para el control de sequías. En este caso, su aplicación se realiza en China. El WSVI se fundamenta en la conjunción de dos efectos de las sequías sobre la superficie terrestre. En condiciones de sequía, la densidad de vegetación disminuye, lo que es detectado y cuantificado a través del Índice de Variación en la Vegetación Normalizado (*Normalized Difference Vegetation Index*, NDVI), que en tales circunstancias decrece. Por otro lado, la sequía produce un aumento en la temperatura sobre la superficie y la vegetación por encima de las condiciones normales. Utilizando ambas fuentes de información, el WSVI se define como el cociente entre NDVI y temperatura superficial. De este modo el índice detecta las condiciones de sequía por una disminución de su valor. Este tipo de índices basados en el efecto de la sequía sobre la vegetación son de utilidad desde el punto de vista del control de las sequías agrícolas, pero no permiten apreciar directamente la realización de una sequía de mayor persistencia (e.g. con duraciones superiores a un año).

2.3.8. El Índice Agregado de Sequías

Como último ejemplo de índice de sequías propuesto en la literatura científica, dentro de los más recientes se encuentra el Índice Agregado de Sequías (*Aggregate Drought Index*, ADI), propuesto por Keyantash y Dracup (2004) para la evaluación agregada de la ocurrencia de sequías en todas sus formas (meteorológicas, hidrológicas y agrícolas). El índice es aplicable sobre una región, y tiene en cuenta en su formulación seis variables: precipitación, evapotranspiración, caudales, almacenamientos en lagos y embalses, humedad en el suelo y altura de nieve almacenada. De este modo el índice está relacionado con las componentes superficiales del ciclo hidrológico de la región de análisis. El cómputo del índice pasa por la transformación estadística inicial de cada variable a su equivalente normal estándar. Sobre el conjunto de variables se realiza un análisis de componentes principales, coincidiendo el ADI con los pesos tipificados de la primera componente. El índice mantiene similitudes con el SWSI, lo que le hace compartir también sus limitaciones, especialmente las referentes a la falta de homogeneidad que puede aparecer en la serie.

Capítulo 3

La modelación estocástica de ocurrencia de sequías

En este capítulo se revisan las diferentes metodologías que más habitualmente se utilizan en la modelación estocástica de sequías. En primer lugar se expone la distinción entre procesos aleatorios puntuales y aquellos no puntuales, tal y como en muchas ocasiones son tratados los eventos de sequía. Secuidamente se presentan algunos modelos de procesos aleatorios puntuales utilizados en la caracterización de sequías. Por último se introduce la Teoría de Rachas. Esta teoría es la que con mayor frecuencia se utiliza cuando se analiza la ocurrencia de eventos de sequía, motivado principalmente por que la sequía mantiene una duración en el tiempo significativa, y su análisis mediante procesos puntuales presenta importantes limitaciones. Como aplicación de la Teoría de Rachas se desarrolla en forma general un modelo para la caracterización univariada, extendiéndose también para el caso multivariado. Finalmente, siendo la sequía un evento extremo cuya extensión en el espacio tiene mucha importancia, se revisan al final del Capítulo los métodos utilizados para la caracterización de su comportamiento espacial.

3.1. Eventos puntuales en el tiempo frente a eventos no puntuales

El análisis estocástico de las realizaciones de eventos de sequía en la literatura científica se ha tratado principalmente haciendo uso de dos tipos de modelos conceptuales: modelos de procesos puntuales (MPP, *Point Process Models*) y modelos de rachas (MR, *Runs Models*). Los modelos MPP son utilizados en hidrología para la modelación de la ocurrencia de eventos aleatorios, principalmente para la modelación de la presentación de eventos

de precipitación o inundaciones (Le Cam, 1961; Kavvas & Delleur, 1976), los cuales tienen una ocurrencia en el tiempo muy localizada. Estos modelos permiten simular las distribuciones estocásticas de los tiempos entre aparición de los eventos, y con ellos derivar propiedades estocásticas asociadas de interés o realizar análisis de riesgo. Los modelos MPP pueden completarse secuencialmente, incluyendo otras características de los procesos, como pueden ser la propia duración del evento, o la magnitud del mismo (e.g., altura de precipitación caída, caudal máximo, etc.). Además pueden incluirse modelos de agrupaciones (*cluster*) para modelar el desarrollo del evento ocurrido (Neyman & Scott, 1958). Sin embargo, aunque realizaciones muy complejas podrían a priori ser modeladas con este tipo de modelos conceptuales, su principal limitación se encuentra en el número de parámetros que aparece en esas aplicaciones (Ramírez & Brass, 1985; Rodríguez-Iturbe *et al.*, 1987b; Rodríguez-Iturbe *et al.*, 1987a; Islam *et al.*, 1988) y en su inconsistencia cuando se realizan validaciones a escalas temporales distintas de las de calibración (Rodríguez-Iturbe *et al.*, 1984; Koepsell & Valdés, 1991). La aplicación de este tipo de modelos en el análisis de la ocurrencias de sequías se encuentra por ejemplo en análisis de aportaciones diarias de un río, analizando los períodos con aportaciones por debajo de un determinado umbral (Abi-Zeid *et al.*, 2004).

Por otro lado, los modelos de rachas tienen como base fundamental, la asignación en cada paso de tiempo de un determinado estado (en aplicaciones de sequías son usados dos: seco y húmedo) y el análisis de las duraciones, u otras propiedades de interés, de los períodos en los que se mantiene el mismo estado (Yevjevich, 1963). Los cambios instantáneos de estados son modelados, por ejemplo, mediante modelos de Markov (Norris, 1997). Los modelos de rachas permiten así mismo estimar el período de retorno de la ocurrencia de un evento extremo, entendido como el tiempo medio entre aparición de eventos.

En las siguientes secciones se describen cada una de estas concepciones, enfocadas principalmente a la modelación de sucesos de sequía. No obstante, debe notarse que existe una gran afinidad entre ambos tipos de modelos, compartiendo en muchos casos simplificaciones conceptuales de la realidad, y apareciendo modelos donde no resulta clara su clasificación en uno de estos grupos. Sin embargo, podría afirmarse que los modelos MPP en general trabajan a una escala de tiempo menor que los modelos de rachas, normalmente diaria o semanal, pudiendo describir la realización del evento en mayor detalle. A estos modelos, en ocasiones se les denomina modelos de eventos. Mientras que los modelos de rachas, suelen trabajar a escalas temporales superiores (mensual, estacional, anual) y modelan con menor detalle la realización del evento, quedándose en muchos casos como único descriptor su duración. Resultan sin embargo más adecuados para carac-

terizar eventos de sequía extremas, las cuales tienen lugar durante períodos de tiempo prolongados y por definición corresponde a eventos de frecuencia muy baja. Los modelos de rachas en ocasiones se denominan modelos de estado.

3.2. Modelación de sequías mediante modelos estocásticos de procesos puntuales

En los MPP, también denominados modelos de eventos, se considera la realización de cada evento como un proceso aleatorio. En el caso de las sequías, cuando por ejemplo se están analizando caudales diarios en un río, los cuales deben satisfacer una determinada demanda, se pueden distinguir dos tipos de eventos puntuales: el producido durante la transición entre el estado de satisfacción de la demanda al estado de déficit, y viceversa. Evidentemente, estos procesos no sólo dependen de la variable aleatoria analizada (caudales diarios) sino también de la demanda establecida. Los MPP pueden ser usados para modelar la ocurrencia de esas transiciones, derivando a partir del modelo propiedades estadísticas de sus realizaciones. A diferencia de las aplicaciones de MPP al análisis de hietogramas de tormentas o avenidas, donde la duración del evento tiene menor importancia relativa frente a otras características, en el caso de sequía la duración del período entre una transición desde el estado de satisfacción a estado deficitario y su retorno a estado de satisfacción es una de las variables más importantes. Para ello es necesario combinar en el modelo la ocurrencia de estos dos tipos de eventos.

3.2.1. Definición y clasificación de procesos puntuales

Un proceso puntual está basado en la noción de la ocurrencia de un evento. Un evento se caracteriza fundamentalmente por el instante en el que se produce su realización. Así un proceso puntual puede venir definido por (Cox & Isham, 1980):

- La distribución conjunta del número de eventos $N(A)$ en un intervalo de tiempo A
- La distribución conjunta de los intervalos $\{X_j\}$ entre eventos sucesivos
- La función de intensidad $\rho(t, H_t)$, donde $\rho(t, H_t) \cdot \Delta t$ es la probabilidad de realización instantánea del evento en el intervalo $(t, t + \Delta t)$, y H_t es la historia de los eventos previos ocurridos hasta el instante t

En lo referente a la clasificación de los procesos puntuales, existe una amplia familia de procesos (Snyder, 1975). A continuación se describen aquellos más usuales en el análisis de sequías, dejando fuera por ejemplo los procesos de agrupación (*cluster*), que aunque son muy usados en hidrología resultan más comunes en análisis de otro tipo de eventos (Waymire & Gupta, 1981).

Procesos de Poisson

Los procesos de Poisson homogéneos son los procesos puntuales más sencillos. Se caracterizan por una función de intensidad constante ρ , donde $\rho \cdot \Delta t$ es la probabilidad de ocurrencia instantánea de un evento en el intervalo de tiempo Δt . En este tipo de procesos son asumidas las hipótesis de proceso estacionario, independiente, y no múltiple (i.e. en cada instante puede producirse hasta un evento). Este modelo ha sido utilizado por ejemplo para modelar la ocurrencia de sequías sobre aportaciones anuales (Madsen & Rosbjerg, 1995).

Los procesos de Poisson no homogéneos son una extensión de los procesos homogéneos, donde la hipótesis de proceso estacionario se reemplazan por una función de intensidad que depende del tiempo en el que se evalúe $\rho(t)$. Este tipo de procesos resultan conveniente cuando se trabaja a escalas temporales inferiores a la anual, donde la hipótesis de proceso estacionario no es asumible (Tallaksen *et al.*, 1997). Una forma general de la función de intensidad cuando la función es periódica viene dada por ec. 3.1 (Abi-Zeid *et al.*, 2004).

$$\rho(t) = \lambda(t) = \exp \left[\beta + \sum_{j=1}^m K_m \cdot \sin(\varpi \cdot t + \theta_m) \right] \quad (3.1)$$

donde ϖ representa la frecuencia de cada armónico, resultando para procesos anuales y escala de temporal diaria $\varpi = 2\pi/365$.

Procesos de renovación (*renewal process*)

Estos procesos se diferencian de los procesos de Poisson en que su función de intensidad sigue una determinada función de distribución, la cual depende del tiempo transcurrido desde la última realización $\rho(t, H_t) = \rho(t - u)$, donde u es el instante en el que ocurrió la última realización. Este tipo de modelos mantienen las hipótesis de homogeneidad, independencia, y no-multiplicidad. Ejemplo de aplicación puede encontrarse en la modelación de sequías definidas sobre caudales medios anuales (Lee *et al.*, 1986).

Cuando la función de intensidad de un proceso de renovación también depende del tiempo $\rho(t, H_t) = \rho(t, t - u)$, reemplazándose la hipótesis de

homogeneidad, se trata entonces de modelos de procesos de renovación no homogéneos. En este caso la forma general (ec. 3.1) debe ser completada para tener en cuenta la no homogeneidad de la función (e.g, ec. 3.2).

$$\rho(t, H_t) = \varphi(t - u) \cdot \lambda(t) \quad (3.2)$$

Procesos de Poisson condicionales y no homogéneos

Cuando la función de intensidad depende de una variable externa, así como del instante en el que se evalúa la probabilidad, el proceso se denomina proceso de Poisson condicional no homogéneo. Ejemplo de aplicación puede encontrarse en la modelación de sequías en caudales diarios, utilizando como variables externas la precipitación y la temperatura (Abi-Zeid *et al.*, 1998). Así el proceso real por el cual los caudales de un río se ven disminuidos tras períodos prolongados sin precipitaciones, puede incluirse en (ec. 3.1) con:

$$\rho(t, H_t) = \exp \left[\beta + K \cdot nd + \sum_{j=1}^m K_m \cdot \sin(\varpi \cdot t + \theta_m) \right] \quad (3.3)$$

donde nd representa el número de días transcurridos desde la última precipitación.

3.2.2. Modelación estocástica de procesos puntuales

La modelación estocástica de procesos puntuales depende del tipo de proceso analizado, siendo muy conveniente el conocimiento de los procesos físicos que los genera. No existe un procedimiento general para la modelación de cualquier proceso, siendo necesario siempre adecuarse a las particularidades de los procesos usando aquellos modelos que pudieran resultar más adecuados. Sin embargo se pueden dar una serie de pasos principales que en toda modelación se debe seguir (Abi-Zeid *et al.*, 2004):

- Definición del evento que quiere ser analizado. En general, tratando con sequías ello supone la definición de un umbral en la variable de análisis. Con ello se puede definir dos eventos: el evento puntual con el que da comienzo la sequía, en el cual se pasa de valores de la variable superiores al umbral a valores inferiores al umbral; y el evento puntual por el que termina la sequía, volviendo las realizaciones de la variable analizada a estar por encima del umbral.
- El análisis de los tiempos entre eventos y la selección del proceso puntual aleatorio más conveniente para su modelado, adoptando formas tales como las mencionadas en la sección anterior. Debe notarse que la

selección del umbral que define los eventos puede producir variaciones en la adecuación de los procesos teóricos.

- Estimación de los parámetros que definen las funciones de intensidad. Para ellos están disponibles estimadores estadísticos basados en criterios de máxima verosimilitud, mínima suma de cuadrados, L-moments, etc. En general resulta el criterio de máxima verosimilitud el más adecuado.
- Validación del modelo mediante aplicación de test estadísticos. Existen test disponibles para la validación estadística de los diferentes procesos teóricos homogéneos (Kavvas & Delleur, 1976). En los casos en los que el proceso sea no homogéneo, es necesaria su transformación para conseguir un proceso homogéneo. Así mismo deben ser comprobadas las hipótesis asumidas en los procesos teóricos.

Los modelos para sequías basados en procesos puntuales permiten resolver preguntas desde un punto de vista probabilista relacionadas con el riesgo de presentación de una sequía a partir de una determinada situación, o el tiempo que tardará la variable indicadora en abandonar una situación de sequía. En general, para la caracterización de la recurrencia de sequías de una determinada magnitud, es más común el uso de la Teoría de Rachas.

3.3. La Teoría de Rachas y la caracterización de eventos no puntuales en el tiempo

La Teoría de Rachas (TR) ha sido usada para el análisis y caracterización estocástica de las sequías desde que Yevjevich (1967) propusiera la definición de eventos de sequía basada en la TR: Un evento de sequía es definido como el período durante el cual la variable indicadora de la disponibilidad de agua, x , (aportaciones, lluvia, humedad del suelo,...) se encuentra por debajo de un determinado umbral, $X_t \leq U$. En esta definición se presupone que la variable indicadora de la disponibilidad de recurso hídrico indica situaciones de poca disponibilidad con valores bajos de la variable. El inicio de un evento de sequía se produce por tanto con el paso de una situación en la que la variable indicadora se encuentra por encima del umbral, a situarse en el siguiente paso temporal por debajo del umbral, $X_{t_0-1} > U$, $X_{t_0} \leq U$. Y de forma inversa se produce el final del evento de sequía, $X_{t_f} \leq U$, $X_{t_f+1} > U$. Entre cada dos períodos sucesivos de sequía se encuentra un período durante el cual la variable indicadora se mantiene por encima del umbral, a estos períodos se les suele denominar períodos húmedos. Así una serie de realizaciones de la variable indicadora, X_0, \dots, X_t, \dots , se divide

en rachas, cada una de las cuales representa un período durante el cual la variable se mantiene en una determinada situación (seco / húmedo). *TR* realiza un análisis de las propiedades estadísticas de las rachas, que permiten caracterizar cada evento con una determinada probabilidad de ocurrencia, así como estimar un período de retorno que caracteriza su recurrencia. El análisis de estas propiedades requiere de la selección previa de aquella(s) magnitud(es) sobre la(s) que se quiere realizar el análisis.

La magnitud que con mayor frecuencia se analiza por la *TR* es la duración, y es la única que se analiza en el caso de las rachas de períodos húmedos, por ser la única que resulta de interés. Respecto a los períodos secos, a la duración se le puede añadir otras como pueden ser: la severidad, definida como la suma acumulada de las diferencias entre el umbral y los valores de la variable; la intensidad máxima, definida como la máxima de las diferencias entre el umbral y las realizaciones que componen el evento; o la intensidad media, definida como el cociente entre la severidad y la duración. No obstante, las magnitudes más analizadas en la literatura son la duración y la severidad (Dracup *et al.*, 1980b; Dracup *et al.*, 1980a). Así Downer *et al.* (1967) derivó analíticamente los momentos de la duración y la severidad bajo la hipótesis de independencia e idéntica distribución (*iid*) de la variable aleatoria. Llamas y Siddiqui (1969) mostraron que bajo la hipótesis de *iid* la distribución de probabilidad de las duraciones sigue una distribución geométrica. Sen (1976; 1977) abordó el problema de derivación de la función de distribución de probabilidad de las características de las sequías, utilizando un modelo de Markov, de rezago unidad, y el primer y segundo momento de las severidades.

Otros autores han tratado el problema de analizar la máxima duración de un evento y la máxima severidad en un número limitado de rachas. Millan y Yevjevich (1971) desarrollaron ecuaciones para la máxima duración de sequías en un número de años determinado, basadas en simulaciones, suponiendo un modelo autoregresivo de primer orden para la variable indicadora. Asumiendo una distribución de Poisson para el número de sequías en un determinado intervalo de tiempo, Gupta y Duckstein (1975) analizaron la distribución de las máximas duraciones, y Sen (1980b) analizó adicionalmente la máxima severidad. Guven (1983) y Sharma (1995) derivaron expresiones similares que las obtenidas por Sen, asumiendo una función gamma para las severidades, y una distribución normal truncada, respectivamente.

A pesar de las investigaciones desarrolladas, aun permanece sin resolver la derivación analítica de las funciones de distribución de probabilidad de la severidad, o la función bivariada de las duraciones y severidades conjuntamente. En este sentido, han sido utilizadas funciones de distribución teóricas univariadas, que de acuerdo con la distribución empírica de du-

raciones, y/o severidades, mejor se ajustaban a la muestra (Sen, 1980b; Guven, 1983; Zelenhasic & Salvai, 1987; Mathier *et al.*, 1992; Sharma, 1995; Shiau & Shen, 2001). Como ejemplo, Mathier *et al.* (1992) ajustaron la distribución geométrica y la distribución gamma a las duraciones y a las severidades, respectivamente, para eventos de sequía identificados en una serie de aportaciones mensuales, a partir de la definición de un umbral en forma periódica. Debe notarse que aunque las magnitudes de duración y severidad están altamente correlacionadas en general, las caracterizaciones univariadas basadas en cada una de ellas por separado pueden diferir significativamente (Frick *et al.*, 1990; González & Valdés, 2003).

Una de las ventajas que mantiene *TR* frente a *PPM* es la posibilidad de derivación del período de retorno de eventos que cumplen una determinada condición. El período de retorno ha sido definido por diferentes autores de diferentes formas. Así por ejemplo, puede definirse como el período de tiempo medio entre la ocurrencia de unos eventos específicos, definidos por el cumplimiento de unas determinadas condiciones (Lloyd, 1970; Fernández & Salas, 1999). Una definición alternativa utilizada es el período de tiempo hasta la ocurrencia del siguiente evento (Brass, 1990; Fernández & Salas, 1999). Fernández y Salas (1999) derivaron la expresión que estima el período de retorno, suponiendo un proceso de Markov de dos estados estacionario. Chung y Salas (2000) extendieron la formulación para los casos en que la duración de las secuencias de rachas secas y húmedas pueden ser modeladas con un modelo AutoRegresivo de Media Móvil Discreto (*Discrete Autoregressive Moving Average, DARMA*). Manteniendo el análisis del período de retorno en un caso univariado, Shiau y Shen (2001) derivaron el período de retorno de eventos de una duración fija, con severidad superior a una crítica. Tratando así mismo las severidades, Bonaccorso *et al.* (2003) expresaron el período de retorno de las severidades (suponiendo que sigue una distribución lognormal) en función de la severidad media, de su coeficiente de variación, así como del umbral seleccionado.

En la misma línea, se han realizado trabajos sobre una escala temporal inferior a la anual, manifestándose propiedades de periodicidad. Así Fernández y Vergara (1998) estimaron el período de retorno de sequías caracterizadas por su duración, utilizando un modelo de Markov periódico. Este modelo ha sido utilizado también para derivar la función de distribución de probabilidad de duraciones de sequías, que permiten hacer análisis de riesgo con escalas temporales inferiores a la anual (Cancelliere & Salas, 2004).

Como ha sido mencionado anteriormente, la selección de la magnitud de análisis (e.g. la duración o la severidad) puede conferir una alta sensibilidad a los resultados, con variaciones importantes en el período retorno asociado. La posibilidad de realizar un análisis de recurrencia con un modelo multivariado, incluyendo más de una magnitud, para evitar este desacuerdo

en los casos en los que ambas magnitudes resultan de interés, se encuentra principalmente limitada por la disponibilidad de series de suficiente longitud que permitan de modo consistente ajustar un modelo probabilista. A diferencia de otros eventos hidrológicos extremos como las avenidas, la sequía lleva implícita una duración significativa, manifestación de su persistencia, lo que produce que en las series instrumentales disponibles se cuente con un limitado número de eventos para su caracterización. Sin embargo, han sido realizados trabajos en este sentido, buscando obtener una caracterización de las sequías más completa. Kim et al. (2003b) presentaron una metodología para estimación del período de retorno de sequías usando un estimador kernel no paramétrico en procesos univariados y bivariados. Shiau y Shen (2001) analizaron las distribuciones condicionales de la severidad, para una duración fija, proponiendo, con base empírica, una distribución gamma para la función de distribución condicional, cuyos parámetros resultan función lineal de la duración. Avanzando en este resultado, González y Valdés (2003) desarrollaron la metodología para la derivación del período de retorno de un evento de sequía caracterizado en base a su duración y severidad simultáneamente. Para la aplicación práctica del mismo, se presentó además una metodología para la inclusión de reconstrucciones paleoclimáticas en el análisis de recurrencia de sequías extremas, de modo que el aumento en el número eventos permite el ajuste estadístico de un modelo de recurrencia bivariado sin infringir el Principio de Parsimonia.

A continuación se presenta en un formato genérico la teoría de TR que permite la caracterización univariada de sequías. Este formato es completo para el caso de la caracterización multivariada, extendiendo el trabajo de González y Valdés (2003) a cualquier caso multivariado. Ello será utilizado en un caso práctico en el Capítulo 6 para comparar los resultados obtenidos con esta teoría con los resultados del Índice de Frecuencia de Sequías.

3.3.1. Caracterización univariada de eventos de sequía

A partir de la Teoría de Rachas, a cada evento de sequía le pueden ser asociados diferentes magnitudes sobre las que realizar su caracterización. Estas pueden ser la duración, la severidad, la intensidad media, la intensidad máxima, etc. La caracterización basada en el valor directo de cualquiera de estas magnitudes puede ser útil para comparación entre sequías producidas en el mismo lugar y analizadas con la misma variable indicadora. Sin embargo, no es posible su comparación con otro tipo de variables o entre análisis resultantes en lugares diferentes, donde la comparación puede aparecer inconsistente. Una escala de comparación de sequías que resulta más universal, y que permite la comparación espacial, sería una cuantía

probabilista (Guttman *et al.*, 1992). La TR puede ser utilizada en la construcción de un modelo estocástico de recurrencia de sequías que permita la estimación del período de retorno de eventos extremos, y de ese modo caracterizar en una escala independiente del lugar, donde se produce cada evento de sequía. Se presenta una formulación genérica del modelo de recurrencia basada en la TR que permite estimar el período de retorno, bajo la hipótesis de distribución estacionaria e independencia entre los períodos secos y húmedos, donde es usado el concepto de procesos alternos de reanudación (*alternating renewal process*). La definición del período de retorno utilizado es el tiempo medio entre inicios de eventos de sequías superiores o iguales a la crítica. Para la definición de la sequía crítica son usadas magnitudes asociadas a los eventos. En el caso de utilizarse una única magnitud, la caracterización sería univariada, y en el caso de utilizarse más de una, resultaría una caracterización multivariada. Sin embargo, para la estimación del período de retorno es necesario el análisis y la inclusión en el modelo de recurrencia de las propiedades estadísticas de la duración de los períodos de sequías, por lo que la caracterización univariada suele basarse en duraciones, aunque puede derivarse para otra magnitud (Shiau & Shen, 2001).

Sea una variable aleatoria x , la cual será utilizada como indicador de sequía (aportación, precipitación, humedad, etc.). La realización de esta variable durante el período de control, ha resultado la serie X_1, X_2, \dots, X_N . Para la definición de los períodos o rachas de sequía se fija un determinado umbral U , que puede considerarse constante, o función del tiempo $U(t)$ (e.g. periódico, de período anual). La intersección de este umbral con la serie de realizaciones produce una subdivisión de la misma en períodos secos (\mathcal{S}) y húmedos (\mathcal{H}). Del análisis de los períodos secos pueden ser obtenidas las diversas magnitudes asociadas. Aquella imprescindible para la estimación del período de retorno es la duración $d_{\mathcal{S}}$, que será aquí la única considerada. Así mismo, la serie de períodos húmedos proporciona una serie de duraciones de estos períodos $d_{\mathcal{H}}$. El análisis estadístico de la muestra de duraciones de períodos secos y húmedos proporciona la posibilidad de ajustar una función de distribución teórica para la distribución poblacional de las duraciones: $f_{d_{\mathcal{S}}}(d)$, $f_{d_{\mathcal{H}}}(d)$. Las funciones que habitualmente mejor se ajustan a estas poblaciones son las funciones de distribución de probabilidad geométrica, en el caso de considerar como variable discreta la duración, o la función de distribución de probabilidad exponencial, en el caso de considerar la variable duración continua. Así mismo, y bajo la hipótesis de independencia entre las duraciones de los períodos secos y húmedos, estas funciones de distribución pueden ser derivadas del modelado de los cambios de un período seco a un período húmedo, y viceversa, por modelos de Markov, con procesos de renovación estacionarios, o no estacionarios –como puede ser el caso si se trabaja a una escala temporal inferior a la anual (Cancelliere & Salas, 2004)–

. De este modo el tiempo entre comienzos de dos eventos sucesivos de sequía, será suma de dos variables aleatorias independientes:

$$L = d_S + d_{\mathcal{H}} \quad (3.4)$$

Utilizando la duración de los eventos de sequía como magnitud para la definición de la sequía crítica, el tiempo transcurrido entre los inicios de los eventos sucesivos de sequía con duraciones mayores o iguales que la duración crítica D , se puede escribir como la suma de las diferentes rachas ocurridas entre ellas:

$$T_d = L_{d_S \geq D} + \sum_{i=1}^{N_d} L_{d_S < D} \quad (3.5)$$

donde N_d representa el número de eventos de sequía ocurridos entre dos eventos sucesivos superiores o iguales al crítico, y

$$L_{d_S \geq D} = d_{S|d_S \geq D} + d_{\mathcal{H}} \quad (3.6a)$$

$$L_{d_S < D} = d_{S|d_S < D} + d_{\mathcal{H}} \quad (3.6b)$$

A partir de la hipótesis de independencia de las duraciones de los períodos alternos, y entre las duraciones correspondientes a los períodos secos y húmedos (i.e. $\rho_{d_S, d_{\mathcal{H}}}, \rho_{d_S, d_S}, \rho_{d_{\mathcal{H}}, d_{\mathcal{H}}} = 0$) el valor esperado del tiempo T_d resulta:

$$E(T_d) = E(L_{d_S \geq D}) + [E(N_d) - 1] \cdot E(L_{d_S < D}) \quad (3.7)$$

Sustituyendo ec. 3.6

$$E(T_d) = E(N_d) \cdot E(d_{\mathcal{H}}) + E(d_{S|d_S \geq D}) + [E(N_d) - 1] \cdot E(d_{S|d_S < D}) \quad (3.8)$$

De nuevo haciendo uso de las hipótesis de independencia, el número de eventos de sequía hasta la ocurrencia de un evento superior al crítico sigue una distribución geométrica de valor medio

$$E(N_d) = \frac{1}{\mathcal{P}(d_S \geq D)} = \frac{1}{\int_{\ell=D}^{\infty} f_{d_S}(\ell) \cdot d\ell} \quad (3.9)$$

Dado que se cumple que $\mathcal{P}(d_S \geq D) + \mathcal{P}(d_S < D) = 1$, se tiene

$$E(N_d) - 1 = \frac{\mathcal{P}(d_S < D)}{\mathcal{P}(d_S \geq D)} \quad (3.10)$$

Por último, el desarrollo de los términos de valores esperados de duraciones de ec. 3.8 resultan

$$E(d_S | d_S \geq D) = \frac{\int_{\ell=D}^{\infty} \ell \cdot f_{d_S}(\ell) \cdot d\ell}{\mathcal{P}(d_S \geq D)} \quad (3.11)$$

$$E(d_S | d_S < D) = \frac{\int_{\ell=0}^{D-1} \ell \cdot f_{d_S}(\ell) \cdot d\ell}{\mathcal{P}(d_S < D)} \quad (3.12)$$

donde el símbolo $[\cdot]^*$ es indicativo para el caso de considerar la variable duración como variable discreta. Sustituyendo en ec. 3.8, las ec. 3.9, 3.10, 3.11 y 3.12, queda:

$$\begin{aligned} E(T_d) &= E(N_d) \cdot E(d_{\mathcal{H}}) + \frac{\int_{\ell=0}^{\infty} \ell \cdot f_{d_S}(\ell) \cdot d\ell}{\mathcal{P}(d_S \geq D)} = \\ &= E(N_d) \cdot [E(d_{\mathcal{H}}) + E(d_S)] \quad (3.13) \end{aligned}$$

donde el período de retorno $E(T_d)$ vendría expresado en las unidades de tiempo manejadas en el análisis. Las hipótesis de independencia permiten expresar de este modo el período de retorno de eventos superiores o iguales al crítico, como el producto del número medio de eventos de sequía entre dos eventos superiores o iguales al crítico, por la duración media entre comienzos de dos eventos sucesivos de sequía. En el caso desarrollado se ha utilizado como magnitud caracterizadora del evento la duración del mismo. Podría ser utilizada otra magnitud, pero al ser requisito necesario la caracterización de las duraciones de eventos de sequía, deberían disponerse de las funciones de distribución conjunta de la duración y la magnitud adicional. Ello pasa pues por un análisis multivariado, el cual se describe de forma genérica a continuación.

3.3.2. Caracterización multivariada de eventos de sequía

Cuando la discriminación entre eventos de sequía quiere ser referida a más de una magnitud, la caracterización resultante es multivariada. Entre las magnitudes consideradas debe incluirse la duración, o en caso de no ser conocida, sí debe serlo la distribución marginal de duraciones conocidas el resto de variables. Para extender la exposición previa al caso multivariado,

el símbolo $\{m\} = \{m_1, \dots, m_n\}$ representa el conjunto de magnitudes (n), excluida la duración, que son utilizadas como criterios de caracterización de los eventos de sequía. Los valores críticos de las magnitudes vienen representados por \mathcal{M} , siendo la condición $\{m\} \geq \mathcal{M}$ representación del cumplimiento simultáneo de la inecuación para cada una de las magnitudes, $\{m_1 \geq M_1, \dots, m_n \geq M_n\}$. En el caso multivariado, el período de retorno se define como el tiempo entre inicios de dos eventos de sequía sucesivos superiores o iguales al crítico, donde la igualdad o superioridad viene manifestada por la simultaneidad de las condiciones $\{m_1 \geq M_1, \dots, m_n \geq M_n\}$, añadiéndose la condición $d_S \geq D$ cuando la duración sea también una magnitud crítica. Seguidamente se incluye el desarrollo general considerando la duración como una magnitud crítica. A partir del mismo resulta directo particularizarlo para el caso en que no sea incluida.

Al igual que ocurría en el caso univariado, el uso de múltiples magnitudes genera, a partir de la serie de realizaciones y de su análisis de rachas, una muestra de la población de cada magnitud. Magnitudes tales como duración, severidad, intensidad media, etc. presentan una significativa correlación entre sí, por lo tanto para su descripción estadística es necesario el ajuste de la función de distribución conjunta de todas las magnitudes consideradas críticas relacionadas con los eventos de sequía. En particular resulta conveniente expresar la función de distribución de probabilidad conjunta como producto de la función de distribución de probabilidad de la duración, multiplicada por la función de distribución de probabilidad del resto de magnitudes condicionada al valor de la duración:

$$f_{d,m_S}(d, m) = f_{d_S}(d) \cdot f_{m|d_S}(m) \quad (3.14)$$

Ahora la variable aleatoria que representa el tiempo entre comienzos de episodios de sequía iguales o superiores a la crítica resulta:

$$T_d = L_{\{d_S \geq D \cap m \geq M\}} + \sum_{i=1}^{N_d} L_{\{d_S < D \cup m < M\}} \quad (3.15)$$

donde

$$L_{\{d_S \geq D \cap m \geq M\}} = d_{S|\{d_S \geq D \cap m \geq M\}} + d_{\mathcal{H}} \quad (3.16a)$$

$$L_{\{d_S < D \cup m < M\}} = d_{S|\{d_S < D \cup m < M\}} + d_{\mathcal{H}} \quad (3.16b)$$

Con esto, junto con las hipótesis de independencia, la ec. 3.8 quedará en el caso multivariado:

$$E(T_d) = E(N_d) \cdot E(d_{\mathcal{H}}) + E(d_{S|\{d_S \geq D \cap m \geq M\}}) + [E(N_d) - 1] \cdot E(d_{S|\{d_S < D \cup m < M\}}) \quad (3.17)$$

donde ahora

$$E(N_d) = \frac{1}{\mathcal{P}(d_S \geq D \cap m \geq M)} = \frac{1}{\int_{\ell=D}^{\infty} f_{aS}(\ell) \cdot \mathcal{P}(m \geq M / d_S = \ell) \cdot d\ell} \quad (3.18)$$

y dado que se cumple que $\mathcal{P}(d_S \geq D \cap m \geq M) + \mathcal{P}(d_S < D \cup m < M) = 1$, se tiene

$$E(N_d) - 1 = \frac{\mathcal{P}(d_S < D \cup m < M)}{\mathcal{P}(d_S \geq D \cap m \geq M)} \quad (3.19)$$

resultando así mismo:

$$E(d_{S|\{d_S \geq D \cap m \geq M\}}) = \frac{\int_{\ell=D}^{\infty} \ell \cdot f_{aS}(\ell) \cdot \mathcal{P}(m \geq M / d_S = \ell) \cdot d\ell}{\mathcal{P}(d_S \geq D \cap m \geq M)} \quad (3.20)$$

$$E(d_{S|\{d_S < D \cup m < M\}}) = \frac{\int_{\ell=0[1]^*}^{D[D-1]^*} \ell \cdot f_{aS}(\ell) \cdot d\ell + \int_{\ell=D}^{\infty} \ell \cdot f_{aS}(\ell) \cdot \mathcal{P}(m < M / d_S = \ell) \cdot d\ell}{\mathcal{P}(d_S < D \cup m < M)} \quad (3.21)$$

Sustituyendo en ec. 3.17, las ec. 3.18, 3.19, 3.20 y 3.21, quedará:

$$\begin{aligned} E(T_d) &= E(N_d) \cdot E(d_{\mathcal{H}}) + \frac{\int_{\ell=0[1]^*}^{\infty} \ell \cdot f_{aS}(\ell) \cdot d\ell}{\mathcal{P}(d_S \geq D \cap m \geq M)} = \\ &= E(N_d) \cdot [E(d_{\mathcal{H}}) + E(d_S)] \quad (3.22) \end{aligned}$$

De este modo se obtiene que la expresión 3.13 puede ser generalizada para el caso multivariado 3.22, como ya demostró González y Valdés (2003) para un caso bivariado. La estructura final de la expresión que determina el período de retorno está compuesta del valor esperado del número de eventos de sequía hasta la realización de una nueva sequía superior o igual a la crítica, cuya expresión, bajo las hipótesis de independencia, sigue una distribución geométrica multiplicada por el valor esperado del tiempo medio entre inicios de eventos de sequía sucesivos. La diferencia fundamental de esta expresión, respecto de la que por ejemplo son usadas en análisis de avenidas máximas anuales, considerando como variable de análisis la precipitación o el caudal máximo anual, es la multiplicación al número esperado

de realizaciones hasta la realización del evento superior o igual al crítico, por el valor esperado del tiempo medio entre inicios. Este último término es reflejo de la importancia de la duración del evento, mientras que en el caso de avenidas máximas anuales, su duración no es relevante para la caracterización estocástica.

Con la misma simplificación conceptual realizada utilizando la TR y el concepto de procesos de renovación es posible extraer otros estadísticos, como por ejemplo los indicativos del riesgo de aparición de una sequía bajo unas determinadas condiciones iniciales. De igual forma, cuando la escala de trabajo temporal inhabilite la hipótesis de proceso estacionario en los procesos de renovación, el modelo puede ser adaptado para su inclusión (Cancelliere & Salas, 2004).

La limitación principal en el uso de estos modelos estocásticos para caracterización de sequías no se encuentra en la formulación de los modelos, sino en la disponibilidad de series suficientemente extensas como para poder realizar un análisis con el número de magnitudes críticas deseado. Las series instrumentales generalmente permiten la realización de una caracterización univariada. Pero el limitado número de eventos de sequía que puede encontrarse en series de longitud entorno a 100 años inhabilita el uso de una caracterización multivariada, la cual requiere de muestras de la población de tamaños mayores. Para ello se ha propuesto en la literatura de métodos indirectos (Yevjevich *et al.*, 1983), en los que es ajustado un modelo estocástico a la variable indicadora de sequías, con el cual son generadas series sintéticas de realizaciones de la variables de la longitud que sean necesarias. A partir de estas series sintéticas y del análisis mediante la TR de los eventos de sequía, es posible el ajuste estadístico de un modelo estocástico multivariado de recurrencia de sequías. Sin embargo esta técnica no aporta nueva información para la caracterización, manteniéndose la limitación de que la muestra de eventos de sequía es relativamente pequeña. Para aumentar la información disponible sobre las propiedades estocásticas de realización de eventos de sequía una fuente de información valiosa son las reconstrucciones paleoclimáticas. González y Valdés (2003) presentaron una metodología para la inclusión de las reconstrucciones, que generalmente reproducen una limitada variabilidad de la variable respecto de la total, junto con las información de registros instrumentales para el ajuste de modelos estocásticos de recurrencia de sequías. Las reconstrucciones proporcionan información sobre el comportamiento de la variable indicadora especialmente dentro de las frecuencias bajas, que son por otro lado las peor representadas por la serie instrumental. En el Capítulo 6 se presenta los resultados de aplicación de esta metodología en la construcción de modelos con hasta tres magnitudes, utilizando la reconstrucción paleoclimática de las aportaciones anuales del río Blue en la cabecera del río Colorado (EEUU). Como en esta aplicación

se pone de manifiesto, la variación del conjunto de magnitudes a considerar críticas, o la modificación del umbral para la delimitación de las rachas, produce una alta sensibilidad en los resultados, lo cual hace que para el objeto de una caracterización general de sequías, la aplicación de la TR no sea el procedimiento más indicado.

3.4. Caracterización espacial de eventos de sequía

El concepto de sequía tiene como atributo adicional a una persistente reducción de la disponibilidad de recurso hídrico, su extensión espacial. Así, los daños que ocasiona un evento de sequía están altamente ligados a la cobertura espacial del evento. Ello conduce a que la caracterización del desarrollo espacial de los eventos de sequía resulte de especial importancia, conjunto con su desarrollo temporal. Sin embargo, dada la mayor complejidad de los análisis espaciales, sólo una reducida parte del esfuerzo investigador se ha dedicado a la caracterización espacial de las sequías, siendo más abundantes los trabajos sobre la caracterización temporal. Los trabajos realizados se centran en la definición de una magnitud adicional, referida a la extensión espacial de las sequías, incorporándose ésta junto con otras magnitudes (e.g. duración, intensidad, severidad,...) en un análisis multivariado (Tase, 1976; Sen, 1980a). Santos (1983) trata por primera vez el problema de la definición de **sequía regional**, estableciendo finalmente como sequía regional aquella que produce un déficit superior a un determinado umbral, en un área superior a una crítica. A partir de la definición, establece tres importantes características de las sequías: la duración de las sequías regionales, el déficit total acumulado regional y la extensión media areal de la sequía. Para estas magnitudes, y bajo precisas hipótesis, desarrolla analíticamente modelos estocásticos e incluso analiza propiedades de correlación, utilizando finalmente como caso de aplicación las precipitaciones sobre Portugal.

Rossi *et al.* (1992) ponen de manifiesto la necesidad de profundizar en el uso de procedimientos analíticos y técnicas de Monte Carlo para extraer la información disponible de las series históricas y la caracterización estadística de la función de distribución de probabilidad de características de sequías regionales.

Bogardie *et al.* (1994) presentaron un modelo hidroclimatológico de generación espacial de sequías en el este de Nebraska (EEUU) a partir de patrones de circulación atmosférica de gran escala. Haciendo uso del Índice de Sequía de Bhalme-Mooley (Bahlme & Mooley, 1980) a escala mensual, fue construido un modelo para la generación espacial distribuida de este

índice a partir de los patrones de circulación atmosférica mediante análisis de componentes principales y análisis factorial. Ello permitía predecir las diferencias en probabilidad de ocurrencias de sequías provocadas por cambios climáticos, quedando caracterizada las realizaciones espaciales de sequías a partir de los modos principales de circulación atmosférica.

Henriques y Santos (1999) desarrollaron una metodología para la construcción de las curvas severidad-área-frecuencia que permite la caracterización estocástica de sequías regionales. En esta línea, un importante trabajo en el estudio y análisis de sequías regionales en Europa se encuentra en el programa ARIDE, *Assesment of the Regional Impact of Drought in Europe*. El análisis de sequías se llevó a cabo separadamente sobre los patrones espaciales de aportaciones hidrológicas (Zaidman & Rees, 2000), y sobre los patrones espaciales de precipitaciones (Álvarez & Estrela, 2000). En el caso del análisis de las sequías regionales hidrológicas, fueron analizadas series diarias de caudales en una malla regular. La definición de sequía vendría dada por la fijación de un umbral (media menos dos veces la desviación típica) a partir de la cual son analizados los patrones espaciales de realización de las sequías mediante diversos parámetros relativos a duraciones, déficit, y época del año en la que se produce. La forma en la que fueron analizadas las sequías regionales meteorológicas (Álvarez & Estrela, 2000) fue haciendo uso de técnicas de agrupación (basada en el coeficiente de correlación de Pearson) de sub-regiones, definidas cada una de ellas por una serie de precipitaciones a escala mensual, lo que da lugar a la delimitación de regiones dentro de cada una de las cuales se observan similares patrones de variación. A partir de esta clasificación, en cada región es analizada mediante la TR la ocurrencia de sequías, definiendo un umbral estacional y criterios mínimos sobre la severidad y la duración para la definición de sequía. Ello permite el control de forma cuantitativa de la ocurrencia de sequías históricas sobre Europa.

Loukas *et al.* (2002) presentaron otro ejemplo de aplicación de análisis de sequías regionales mediante la aplicación de técnicas de agrupación y análisis de factores en Grecia, utilizando como variable indicadora de sequía el índice-Z de anomalías de humedad del Índice de Severidad de Sequías de Palmer (Palmer, 1965) a escala mensual. Tras la delimitación de las regiones homogéneas, en cada una de ellas fue analizada la ocurrencia de sequías mediante el uso de diversas técnicas de análisis estadístico univariado.

A diferencia de estos análisis de sequías regionales, en los cuales no se mantiene un análisis de forma distribuida, sino que por el contrario, se agregan regiones homogéneas sobre cada una de las cuales se realizan análisis univariados estadísticos y estocásticos, aparecen otros trabajos donde se vuelve a analizar en mayor profundidad la extensión espacial como una magnitud adicional (Kim *et al.* , 2003a; Hisdal & Tallaksen, 2003), como ya

presentase Santos (1983) . En estos trabajos se exponen métodos para la construcción de curvas que relacionan frecuencia de ocurrencia y extensión de la sequía. En el caso de Hisdal y Tallaksen (2003), fue utilizado como caso de estudio la región de Dinamarca. Sobre esta región, con un conjunto de registros pluviométricos y fononómicos que por separado cubrían toda la región, fue aplicada la técnica de transformación mediante Funciones Ortogonales Empíricas (*FOE*). Además, se desarrollaron funciones de interpolación para el cálculo de los valores de los pesos de estas *FOE* en cada una de las celdas de una malla regular, y para los casos de precipitación y aportaciones a nivel mensual. A partir de esto, se construyeron modelos de generación espacial de los pesos de estas funciones, de forma que mediante transformación inversa, podían ser generadas en cada celda series sintéticas de precipitaciones y aportaciones. Estas series permitían el análisis de la recurrencia de sequías, introduciendo el carácter de su extensión espacial. Mediante la selección del conjunto de sequías cuya extensión espacial estaba dentro de un intervalo (e.g. 1-10 %, 11-20 %, 21-30 %, ...), fueron analizadas para cada una de ellas la recurrencia estocástica de sequías de una determinada severidad, mediante análisis univariado, analizándose el diferente comportamiento en función de la extensión espacial, y en función del tipo de sequía tratada: meteorológica o hidrológica. Los resultados muestran la mayor probabilidad de ocurrencia de sequías meteorológicas frente a las sequías hidrológicas, resultando de mayor persistencia las sequías hidrológicas.

Como resumen de la revisión de los trabajos realizados desde el punto de vista de la caracterización espacial de eventos de sequía, las metodologías se encuentran en definitiva con limitaciones similares al caso del análisis del desarrollo temporal de sequías. La limitada longitud de las series instrumentales no permite en muchos casos análisis de mayor dimensión que los univariados. En ese sentido, las técnicas de agrupación estadística permiten simplificar el problema. Por otro lado, en los casos en los que se adopta la extensión espacial como una magnitud adicional, la inclusión de esta magnitud en el análisis lo hace más complejo, obteniéndose únicamente relaciones marginales, fruto de fijar el rango de variación de todas las variables magnitudes excepto una. La interrelación entre las diferentes magnitudes de una sequía (e.g. duración, severidad, intensidad, extensión,...) conjunto con los patrones espaciales de desarrollo de las sequías (e.g. posición en el espacio, formas de evolución espacial,...) son aspectos de alta complejidad cuando el objeto es su análisis conjunto. Estos últimos aspectos están muy ligados a los patrones climáticos globales, en análisis regionales de gran escala, y patrones locales, de menor escala. Es en esta línea, pueden ser desarrollados análisis completos de la ocurrencia y desarrollo de sequías en su dimensión espacial y temporal introduciendo modelos de clima globales. Por tanto requiere la conjugación de las disciplinas de las ciencias atmosféricas conjunto

con las ciencias hidrológicas, para analizar el impacto que sobre los sistemas hídricos producen.

No obstante, y previamente a la realización de cualquier análisis es necesaria la delimitación de los criterios sobre los que se va a realizar el análisis de los eventos de sequía, como ocurría en el caso de los desarrollo temporales, repitiéndose las consideraciones acerca del alto grado de indefinición del problema cuando se trata de un análisis general de sequías donde todas las magnitudes que puedan presentar un carácter extremo son importantes.

Capítulo 4

Planteamiento de la problemática y objetivos de la tesis

La revisión del estado del arte presentada en los dos capítulos anteriores pone de manifiesto la complejidad que puede alcanzar el análisis y la evaluación de situaciones de sequía. De la revisión se deduce como principales conclusiones las siguientes:

- La dispersión existente en las múltiples definiciones propuestas de sequía viene motivada, por un lado, por su evaluación sobre los efectos que produce y no por el fenómeno o fenómenos que la origina. Así mismo, la particularización para cada lugar y fin concreto multiplican el abanico de definiciones.
- Todas las definiciones de sequía persiguen de forma directa o indirecta detectar la realización de un evento que implica una reducida presencia de agua excepcional, referida a un régimen de variación normal, y mantenida de forma persistente y extensa.
- Esto propicia que de acuerdo a la naturaleza aleatoria de las variables que cuantifican la presencia de agua, la vía que permite la evaluación de forma uniforme de la situación de sequía pase por una caracterización bien estadística (en términos de probabilidad) o bien estocástica (en términos de frecuencia).
- La vulnerabilidad de cada sistema usuario a situaciones de sequía delimita para cada aplicación el nivel de excepcionalidad para el que es considerado sequía.

- La importancia del aspecto persistente en el concepto de sequía colocan en segundo orden de adecuación a la caracterización estadística, frente a la caracterización estocástica, para evaluar en una escala uniforme el carácter extraordinario de un evento tenido lugar. La caracterización estadística requiere siempre de la especificación de la escala temporal utilizada en el análisis, pues distintas escalas pueden suponer distintas propiedades de recurrencia para el mismo nivel de probabilidad.
- De entre los índices de sequía propuestos en la literatura, y utilizados en la práctica, pueden encontrarse índices basados en el tratamiento estadístico del fenómeno, destacando el SPI. Sin embargo, no se encuentra ningún índice basado en un tratamiento estocástico riguroso.
- De las metodologías disponibles para realizar caracterizaciones estocásticas de eventos de sequías destaca las aplicaciones basadas en la Teoría de Rachas. Su uso mantiene como más seria limitación la ambigüedad del análisis cuando se pretende realizar una caracterización general de sequías, donde no está impuesta *a priori* una magnitud, o un conjunto de magnitudes, que defina el criterio de caracterización de los eventos. La selección de diferentes criterios puede provocar diferentes resultados. Asimismo, la metodología requiere la definición del estado de sequía, cuya fijación es fuente de ambigüedad adicional para la caracterización (Capítulo 6).
- En lo que se refiere a metodologías de análisis del desarrollo completo espacio-temporal de eventos de sequía, las técnicas revisadas de análisis regionales de sequía tratan de discretizar la región de estudio en sub-regiones sobre las que la realización de una sequía puede considerarse homogénea. Esto lo obtienen a partir de técnicas de análisis multivariado. Una vez obtenida la discretización, a cada sub-región le es aplicado técnicas de análisis de sequía sobre una serie temporal. La discretización en pocas sub-regiones que puede resultar con este tipo de análisis dificulta la completa visualización del desarrollo de las sequías. Así mismo no se ha encontrado ningún caso donde sea propuesta la evaluación desde un punto de vista estocástico.
- Otros métodos de análisis espacial se basan en la aplicación de un índice de sequía que permita uniformizar las diferentes variabilidades climáticas existentes en la región. Este tipo de técnicas mantienen la limitación que presenta los índices propuestos en la capacidad de comparación espacial, siendo más aplicables en regiones quasi-homogéneas desde un punto de vista climatológico.

De acuerdo a la necesidad detectada de una metodología que permita la caracterización estocástica de sequías sobre la que poder evaluar sus realizaciones espacio-temporales, la Tesis que aquí se presenta mantiene como objetivos principales los que a continuación se enumeran:

- Desarrollar una metodología de análisis de desviaciones persistentes extremas producidas en secuencias de realizaciones de variables aleatorias que permita la caracterización estocástica de estas desviaciones, y como caso particular, la caracterización de sequías.
- La definición a partir de la metodología anterior de un índice de sequía que permita la evaluación de la situación de sequía en cada instante con las siguientes premisas:
 - La escala sobre la que se define el índice debe ser independiente del régimen de variación de la variable aleatoria utilizada para el análisis de las sequías, permitiendo la comparación de los valores obtenidos por el índice sobre realizaciones de la variable representativa del fenómeno en cualesquiera dos lugares.
 - La escala deberá estar relacionada con la recurrencia media con la que se produce el fenómeno, de modo que proporcione una cuantificación directa de su carácter extraordinario.
 - El índice debe depender únicamente de la variable aleatoria representativa utilizada, sin influir criterios relacionados con un uso o demanda, proporcionando una caracterización general de la desviación producida.
 - La vulnerabilidad que un sistema usuario tenga a las realizaciones de sequías se pondrá de manifiesto en cada aplicación particular en el valor del índice para el que se defina la situación de sequía en ese uso determinado.
 - Deberá ser posible calcular el índice a partir de las series de registros instrumentales generalmente disponibles, entendiendo que éstos son representativos de la distribución estadística de su población
- Completar la metodología definiendo una sistemática que permita analizar y caracterizar el desarrollo espacio-temporal de la realización de los eventos de sequía a partir del índice anterior

El desarrollo de estos tres objetivos principales se presenta en los Capítulos 5, 6 y 7, respectivamente. El análisis de las principales limitaciones en la caracterización general de sequías, encontradas al hacer uso de los índices

más generales y las metodologías disponibles de caracterización, y la mejora del tratamiento con el nuevo índice es también un objetivo general del trabajo de investigación (Capítulo 6). Por último, otro objetivo general y que también supone la justificación de este trabajo, es la utilidad práctica de sus resultados. Su aplicabilidad es analizada en el Capítulo 8 sobre varios casos prácticos reales.

Capítulo 5

Caracterización estocástica de secuencias de realizaciones con desviaciones persistentes extremas

Este capítulo describe la metodología propuesta en esta Tesis Doctoral para la caracterización de secuencias de realizaciones, de una variable aleatoria, respecto a su aproximación persistente hacia una región extremal de la función de distribución estadística de la variable. En primer lugar se describen los diferentes criterios que pueden ser utilizados para la caracterización de secuencias de realizaciones, y el criterio aquí adoptado. A partir de éste, se analizan las propiedades de recurrencia que se derivan de su aplicación en la caracterización, bajo la hipótesis de estacionariedad e independencia temporal de la variable aleatoria analizada. Dado que muchas de las variables hidrológicas presentan una significativa dependencia temporal en sus realizaciones para determinadas escalas temporales, se describen además consideraciones para su tratamiento en la metodología expuesta. Asimismo, como también es habitual, el no cumplimiento de la hipótesis de estacionariedad puede ser obviado mediante la aplicación de técnicas clásicas de estacionarización. Estas técnicas son revisadas al final de este capítulo.

5.1. Criterios para la caracterización de una secuencia de realizaciones

Una secuencia de realizaciones, $X_i, X_{i+1}, \dots, X_{i+w-1}$, de una variable aleatoria de rango continuo y estacionaria, x , con una función de densidad

de probabilidad (*f.d.p.*) $f_x(x)$, puede ser caracterizada de formas muy distinta en función del aspecto que se pretenda analizar en ella. Para el caso de variables de tipo hidrológico (e.g. precipitaciones, aportaciones, niveles piezométricos en acuíferos, humedad del suelo, etc.) y con el objeto de analizar fenómenos de sequías se pueden destacar como los criterios más popularmente utilizados los siguientes:

- El valor acumulado de la secuencia de realizaciones, que para un determinado período de agregación, w , proporciona una cuantificación de la tendencia general del período. Este criterio es el utilizado por ejemplo en la definición del Índice de Precipitaciones Estandarizado (*Standardized Precipitation Index*, SPI) (McKee *et al.* , 1993) para los diferentes períodos de agregación.

$$Y_{i,w} = \sum_{j=i}^{i+w-1} X_j \quad (5.1)$$

- El valor medio de la secuencia de realizaciones, que cuantifica la tendencia general que en la secuencia se ha producido con independencia de su duración.

$$\bar{X}_{i,w} = \frac{Y_{i,w}}{w} \quad (5.2)$$

- El valor mínimo producido en la secuencia, que cuantifica la desviación extrema observada dentro de la secuencia.

$$X_{\min,i,w} = \min_{j=i}^{i+w-1} (X_j) \quad (5.3)$$

En el caso de análisis de rachas, en donde se define un umbral, U , para el cual el conjunto de valores consecutivos de realizaciones por debajo de este umbral componen una racha,

$$X_r, X_{r+1}, \dots, X_{r+D-1} / X_j \leq U, j \in [r, \dots, r + D - 1]; X_k > U, \\ k = r - 1, r + D \quad (5.4)$$

se pueden definir los siguientes estadísticos de caracterización (Dracup *et al.* , 1980a):

- La duración del período, D , durante el cual las realizaciones se han mantenido por debajo del umbral. Proporciona una cuantificación directa de la persistencia de la racha.
- La intensidad máxima, definida como la mayor de las diferencias entre el umbral y las realizaciones producidas en la racha. También definida como el déficit máximo producido en una realización de las que compone la racha. Caracteriza la racha por su mayor desviación del umbral, sin notar la distribución del resto de realizaciones que forman la racha.

$$I_{\text{máx}} = \text{máx}(U - X_j) \quad (5.5)$$

- La intensidad media, o el déficit promedio producido durante la racha. Se define como el promedio de las desviaciones con respecto al umbral U . De forma similar al valor promedio de una secuencia, proporciona la tendencia general en la racha. Su cálculo puede derivarse del valor medio de una secuencia, sustituyendo w por D :

$$\bar{I} = \frac{\sum_{j=r}^{r+D-1} (U - X_j)}{D} = U - \bar{X}_{r,D} \quad (5.6)$$

- La severidad, o déficit acumulado, definida como la acumulada de las diferencia entre el umbral y las realizaciones. Está relacionada con la duración y la intensidad media, cuantificando el déficit total producido durante la racha.

$$S = \sum_{j=r}^{r+D-1} (U - X_j) = \bar{I} \cdot D \quad (5.7)$$

Adicionalmente a estos estadísticos, clásicamente utilizados, puede incluirse un factor para tener en cuenta la distribución temporal de los déficit:

- El factor de uniformidad, coeficiente adimensional, definido para cuantificar el grado de uniformidad en la desviación producida durante la racha. Se define para valores de la duración superiores a 1, $D > 1$, y varía entre 0, para déficit concentrados en el tiempo, y $+\infty$ para desviaciones totalmente uniformes en el tiempo. Mantiene una mayor independencia con S y D que la que mantiene $I_{\text{máx}}$.

$$F_U = -\log \left(\frac{\frac{I_{\text{máx}}}{S} - \frac{1}{D}}{1 - \frac{1}{D}} \right) \quad (5.8)$$

Cada una de estas magnitudes o estadísticos, proporciona una cuantificación de una propiedad que puede conferir el adjetivo de extremo a una determinada secuencia de realizaciones o a una determinada racha, cuando un umbral es definido. Cuando el objeto del análisis de sequías es particular con un determinado objetivo, puede ser seleccionada la magnitud o magnitudes más apropiadas para el objetivo buscado, y a partir de esto, realizar la caracterización de los diferentes eventos ocurridos en una serie, o incluso realizar un análisis de su recurrencia (Capítulo 3). Así puede ser el caso del problema de la sequía agrícola, con una parcela donde se requiere el análisis de la longitud de los períodos con humedad del suelo por debajo del punto de marchitez de la planta, y cuya duración excesiva puede producir la pérdida de una plantación. En tal situación, la aplicación de la Teoría de Rachas es lo más adecuado para analizar la distribución y recurrencia de sus duraciones, definiendo el umbral como la humedad que iguala el punto de marchitez, por debajo de la cual se produce el fallo. En el caso que sea posible la aplicación de riegos de emergencia para evitar la muerte de las plantas, también se requerirá conocer no sólo la duración de los períodos de sequía agrícola, sino también los volúmenes de aguas a aplicar necesarios para su mitigación. Ante este problema un análisis bivariado, duración–severidad de los eventos de sequías permitiría fijar el volumen a disponer para limitar el riesgo hasta el nivel deseado.

Sin embargo, ninguna de estas magnitudes por separado proporciona una caracterización general de los episodios de sequía, con el cual caracterizar de forma genérica una serie de episodios. Así mismo, para la aplicación de la Teoría de Rachas, es necesaria la definición del umbral. En una caracterización genérica, donde no hay una definición concreta del fallo, la arbitrariedad en la definición del umbral puede introducir fuertes sensibilidades en el análisis. Adicionalmente, los diferentes regímenes de variación que pueden producirse en diferentes localizaciones geográficas de una determinada variable, puede hacer que no sean comparables los valores de las magnitudes utilizadas en cada lugar. Estos efectos son analizados de forma práctica en el Capítulo 6. Con el objeto de cubrir la necesidad de un criterio de caracterización de carácter general, aquí ha sido desarrollado un nuevo criterio de naturaleza probabilística, que busca evaluar la desviación hacia una región extremal del rango de variación de una variable aleatoria durante una secuencia. A continuación se introducirá este criterio, representado por la función de persistencia extremal.

5.1.1. La función de persistencia extremal

Una región extremal de la función de distribución estadística de una variable aleatoria, se define aquí como aquella para la que la probabilidad,

o la densidad de probabilidad, en cada uno de sus puntos es relativamente pequeña, pero no nula. En gran número de variables aleatorias naturales, de dominio de variación infinito (o considerado infinito a efectos prácticos), pero cuya distribución se concentra entorno a un rango de valores, con una forma de campana de Gauss o sesgada hacia una determinada dirección, las regiones extremales se encuentran en las colas de sus distribuciones. En otros muchos casos, el dominio de variación de la variable no es infinito, sino semi-infinito, pero con una reducida probabilidad de realizaciones entorno al límite de su rango de variación y estando el rango de valores con mayor probabilidad alejados de este límite. Para estos casos puede considerarse en la práctica una distribución teórica de rango infinito, apareciendo dos regiones extremales en su distribución. En el caso de distribuciones con un dominio de variación semi-infinito, y donde el valor que limita el dominio mantiene una alta probabilidad, sólo se encuentra una región extremal, situada hacia la apertura del dominio en la dirección del límite infinito (figura 5.1).

La ocurrencia de un episodio de sequía, analizada sobre los valores observados de una variable hidroclimática, se presenta como una secuencia persistente de realizaciones de una variable aleatoria con tendencia hacia una región extremal. En el caso de sequías, la región extremal será aquella asociada a reducida presencia de recurso hídrico (i.e. altura de precipitación, aportaciones, humedad en el suelo, etc.). Sin pérdida de generalidad, será asumido que la región extremal analiza en el caso de las sequías, será aquella que se extiende hacia el límite impropio $-\infty$ (o en el caso de una variable con un dominio de variación limitado, su límite inferior). A este límite se le considera el polo del análisis. Con el objeto de cuantificar de forma probabilista la proximidad de una realización de la variable aleatoria, x , al polo, se hace uso de la función de probabilidad acumulada (*f.p.a.*) de su población (ecuación 5.9).

$$F_x(X) = \mathcal{P}(x \leq X) \quad (5.9)$$

Valores bajos de la *f.p.a.* corresponden a realizaciones próximas a la región extremal objeto de análisis. En una secuencia de realizaciones, $X_i, X_{i+1}, \dots, X_{i+w-1}$, su proximidad a la región extremal podría cuantificarse mediante una generalización de la *f.p.a.* en un espacio multidimensional, de dimensión w , con la aquí llamada función extremal (*f.e.*):

$$\check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) = \mathcal{P}(x_i, x_{i+1}, \dots, x_{i+w-1}/x_i \leq X_i, x_{i+1} \leq X_{i+1}, \dots, x_{i+w-1} \leq X_{i+w-1}) \quad (5.10)$$

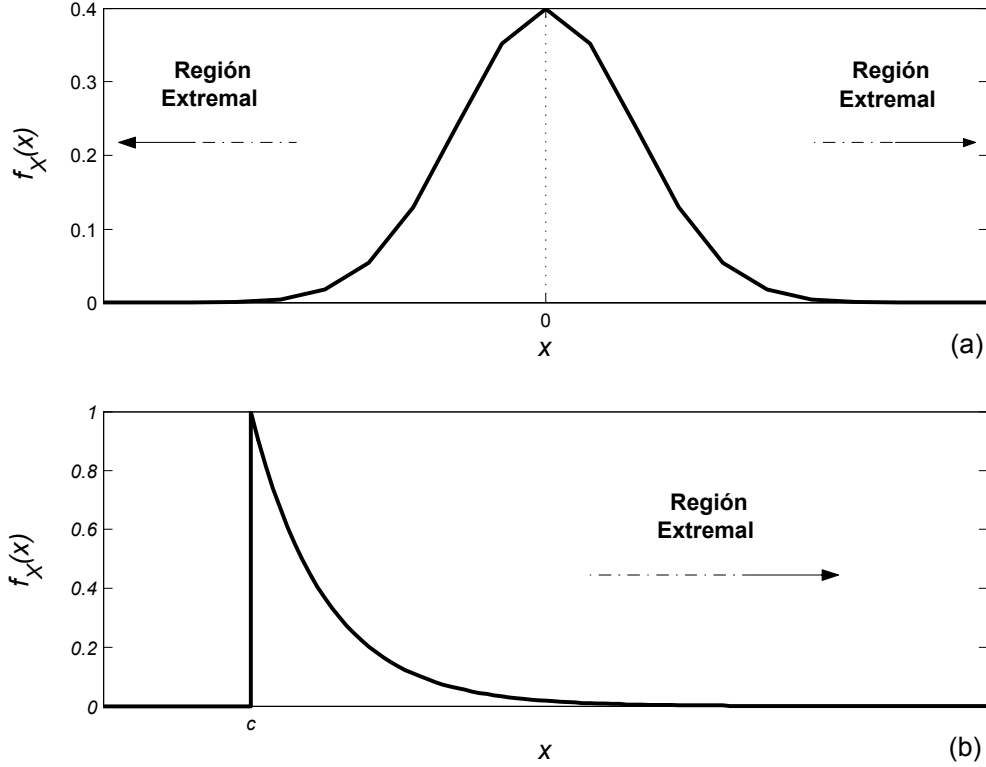


Figura 5.1: Localización de las regiones extremales, en distribuciones de variables aleatorias con dominio de variación infinito y distribuciones gaussianas (a), y variables con dominio de variación semi-infinito y distribuciones exponenciales (b)

Bajo la hipótesis de independencia temporal de las realizaciones, lo que implica que el valor producido en una nueva realización es independiente del conjunto de realizaciones previas, la función extremal se puede escribir como el producto de las funciones *f.p.a.* de las realizaciones que componen la secuencia (ec. 5.11).

$$\check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) = \prod_{j=i}^{i+w-1} F_x(X_j) \quad (5.11)$$

Cuanto menor resulta el valor de la función extremal, mayor es la desviación producida en la secuencia hacia el polo de análisis. Sin embargo la función extremal tiene una alta dependencia de la dimensión de la secuencia, w . En general valores de la función extremal correspondiente a secuencias de diferente longitud (dimensión) no son comparables, pues la función extremal para secuencias de mayor número de realizaciones tiende a producir valores

más bajos. Así la longitud de la secuencia es importante en el análisis, dado que secuencias más largas de realizaciones igualmente extremas (i.e. igual valor de sus *f.p.a.*) resultan menos probables. Por ejemplo, la realización de una secuencia de w valores, todos iguales y correspondientes a un valor de la *f.p.a.* de $F_x(X_j) = 0,5$, $j = i, i+1, \dots, i+w-1$, produce un valor de la función extremal de $\check{F}_{x,w} = 0,5^w$. El aumento en una unidad de la longitud de la secuencia reduce a la mitad su función extremal, cuando realizaciones con *f.p.a.* de $0,5$ no pueden ser consideradas extremas. Para reducir el efecto de la dimensión sobre el criterio de análisis se define la función de persistencia extremal (*f.p.e.*) sobre una secuencia de realizaciones de una variable aleatoria $x, X_i, X_{i+1}, \dots, X_{i+w-1}$, como la función de probabilidad acumulada de secuencias de longitud w , con valores de la función extremal mayores o iguales que $\check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1})$:

$$\hat{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) = \mathcal{P} \left[x_i, x_{i+1}, \dots, x_{i+w-1} / \check{F}_{x,w}(x_i, x_{i+1}, \dots, x_{i+w-1}) \geq \check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) \right] \quad (5.12)$$

El incremento de la *f.p.e.* supone una mayor desviación de la secuencia hacia el polo. Para facilitar el cómputo de esta función, a partir de funciones conocidas, se aplica la siguiente transformación a la variable aleatoria:

$$v = -\ln(F_x(x)) \quad (5.13)$$

donde \ln representa el logaritmo neperiano. La variable aleatoria resultante de la transformación, tiene el dominio de variación $v \in (0, +\infty)$, siendo su función de distribución estadística la función teórica exponencial, con valor medio $\mu_v = 1$. Las ecuaciones 5.14 y 5.15 describen la *f.p.a.* y la *f.d.p.* de la variable transformada v , respectivamente.

$$F_v(v) = 1 - \exp(-v) \quad (5.14)$$

$$f_v(v) = \exp(-v) \quad (5.15)$$

A partir de la transformación de la secuencia de realizaciones producida ($X_i, X_{i+1}, \dots, X_{i+w-1} \implies V_i, V_{i+1}, \dots, V_{i+w-1}$), la ecuación 5.11 puede escribirse como:

$$\ln \left(\check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) \right) = \sum_{j=i}^{i+w-1} V_j \quad (5.16)$$

De aquí se deriva que ec. 5.12 pueda escribirse en términos de la variable transformada como:

$$\hat{F}_{v,w}(V_i, V_{i+1}, \dots, V_{i+w-1}) = \mathcal{P} \left(v_i, v_{i+1}, \dots, v_{i+w-1} / \sum_{j=i}^{i+w-1} v_j < V_C \right) \quad (5.17)$$

$$V_C = \sum_{j=i}^{i+w-1} V_j \quad (5.18)$$

El parámetro V_C condensa la información de la secuencia, a efectos del computo de la *f.p.e.* Desarrollada como una ecuación integral, y haciendo uso de 5.15, la ecuación 5.17 resulta:

$$\begin{aligned} \hat{F}_{v,w}(V_C) = & \int_0^{V_C} \exp(-t_1) \cdot \int_0^{V_C-t_1} \exp(-t_2) \dots \\ & \dots \int_0^{V_C - \sum_{k=1}^{w-1} t_k} \exp(-t_w) \cdot dt_w \cdot \dots \cdot dt_2 \cdot dt_1 \quad (5.19) \end{aligned}$$

La solución de la ecuación integral 5.19 tiene un desarrollo en serie, en la siguiente forma:

$$\hat{F}_{v,w}(V_C) = 1 - \left(\sum_{k=1}^w \frac{V_C^k}{k!} \right) \cdot \exp(-V_C) \quad (5.20)$$

La demostración de la relación 5.20 se puede realizar mediante el método de inducción, del siguiente modo:

$$\hat{F}_{v,w=1}(V_C) = \int_0^{V_C} \exp(-t_1) \cdot dt_1 = 1 - \exp(-V_C) \quad (5.21)$$

$$\hat{F}_{v,w=2}(V_C) = \int_0^{V_C} \exp(-t_1) \cdot \int_0^{V_C-t_1} \exp(-t_2) \cdot dt_2 \cdot dt_1 = 1 - (1 + V_C) \cdot \exp(-V_C) \quad (5.22)$$

Suponiendo que para $w = n - 1$ se cumple 5.20 y sustituyendo V_C por $V_C - t$ en ec. 5.19, para $w = n$ la ec. 5.19 podría escribirse como:

$$\hat{F}_{v,w=n}(V_C) = \int_0^{V_C} \exp(-t) \cdot \hat{F}_{v,w=n-1}(V_C - t) \cdot dt \quad (5.23)$$

$$\hat{F}_{v,w=n}(V_C) = \int_0^{V_C} \exp(-t) \cdot \left[1 - \left(\sum_{k=1}^{n-1} \frac{(V_C - t)^k}{k!} \right) \cdot \exp(-V_C - t) \right] \cdot dt \quad (5.24)$$

$$\hat{F}_{v,w=n}(V_C) = 1 - \exp(-V_C) - \left(\sum_{k=1}^{n-1} \frac{(V_C)^{k+1}}{k! \cdot (k+1)} \right) \cdot \exp(-V_C) \quad (5.25)$$

$$\hat{F}_{v,w=n}(V_C) = 1 - \left(\sum_{k=1}^n \frac{V_C^k}{k!} \right) \cdot \exp(-V_C) \quad (5.26)$$

lo cual coincide con ec. 5.20 para $w = n$.

Adicionalmente, la función $\hat{F}_{v,w}$ puede escribirse en la forma de la función gamma incompleta (Chaudhry & Zubair, 2002):

$$\hat{F}_{v,w}(V_C) = \Gamma(w, V_C) = \frac{1}{\Gamma(w)} \cdot \int_0^{V_C} t^{w-1} \cdot \exp(-t) \cdot dt \quad (5.27)$$

La función *f.p.e.* en función de V_C , y para diferentes valores de w , se muestra en la figura 5.2. Por ser una función de probabilidad acumulada, *f.p.e.* varía entre 0 y 1. Valores de la *f.p.e.* próximos a 1 indican secuencias con una desviación significativa hacia la región extremal, definida por el polo de análisis. La variable aleatoria V_C resulta suma de una serie finita de variables aleatorias V , con función de distribución estadística exponencial (ec. 5.14), cuya media y varianza son uno. La *f.p.a.* de V_C , para un determinado w , sigue la ec. 5.27, que para valores altos de w , por ser suma de w variables aleatorias *iid*, de acuerdo al Teorema Límite Central, tiende a la función de distribución de probabilidad de una variable aleatoria normal $N(\mu = w, \sigma^2 = w)$ (figura 5.2).

Es la función de probabilidad extremal la que se adopta como criterio para una caracterización de secuencias de realizaciones de una variable aleatoria. A continuación se realiza un análisis comparativo de los resultados prácticos obtenidos utilizando los diferentes criterios expuestos para mostrar la mejor adecuación de *f.p.e.* como criterio de caracterización general, de acuerdo con los objetivos formulados.

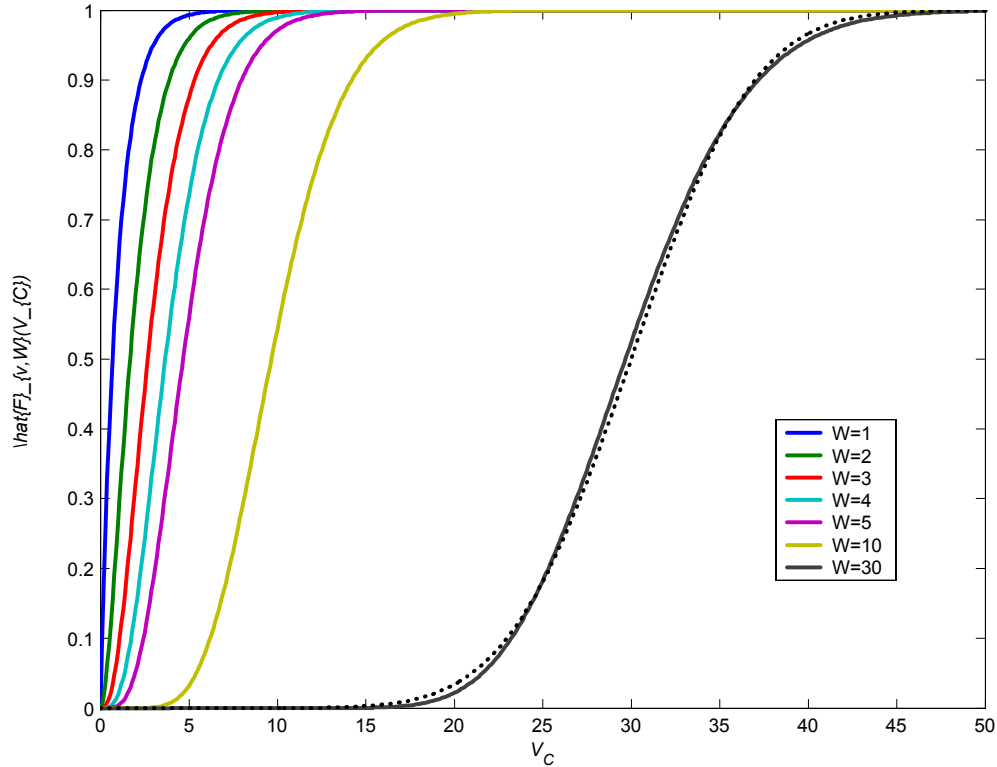


Figura 5.2: Función de probabilidad extremal (ec. 5.27), para diferentes valores de W , y la *f.p.a.* (en línea discontinua) de una variable N ($\mu = 30, \sigma^2 = 30$)

Una propiedad que será notable en la siguiente sección, es la reducción de la capacidad de variación de la *f.p.e.* para un incremento unitario de t con el aumento de w . A medida que valores más altos de w van siendo adoptados, se reduce la capacidad de variación de los valores de la función $\hat{F}_{x,w}$ en cada paso. Esto es debido a que la función $\hat{F}_{x,w}$ mantiene una derivada menor, para mayores valores de w (figura 5.2). Y simultáneamente, para el sumatorio del que depende $\hat{F}_{x,w}$, de acuerdo a su expresión en la forma de la variable transformada (ec. 5.17), los estadísticos de su variación con respecto a variaciones en t no depende de w , por lo que se manifiesta la disminución de la derivada con w . Esta idea puede ser cuantificada de forma analítica analizando los momentos estadísticos de primer y segundo orden de la variación de la función $\hat{F}_{x,w}$ con una variación Δt en t , y para un valor fijado de w :

$$\frac{\Delta \hat{F}_{x,w}(x_{t-w+1}, x_{t-w+2}, \dots, x_t)}{\Delta t} = \frac{\Delta \hat{F}_{v,w} \left(v_C = \sum_{j=t-w+1}^t v_j \right)}{\Delta t} \quad (5.28)$$

$$\frac{\Delta \hat{F}_{v,w}(v_C)}{\Delta t} = \frac{\partial \hat{F}_{v,w}(v_C)}{\partial v_C} \cdot \frac{\Delta v_C}{\Delta t} \quad (5.29)$$

Al ser la función $\partial \hat{F}_{v,w}(v_C)/\partial v_C$ una función determinística, los momentos estadísticos de $\Delta \hat{F}_{x,w}/\Delta t$ vienen dados por los momentos estadísticos de la función aleatoria $\Delta v_C/\Delta t$, afectada por el valor determinístico de $\partial \hat{F}_{v,w}(v_C)/\partial v_C$. Ello conduce a desarrollar los momentos de primer y segundo orden en la siguiente forma:

$$E \left[\frac{\Delta \hat{F}_{v,w}(v_C)}{\Delta t} \right] = \frac{\partial \hat{F}_{v,w}(v_C)}{\partial v_C} \cdot E \left[\frac{\Delta v_C}{\Delta t} \right] \quad (5.30)$$

$$E \left[\left(\frac{\Delta \hat{F}_{v,w}(v_C)}{\Delta t} - E \left[\frac{\Delta \hat{F}_{v,w}(v_C)}{\Delta t} \right] \right)^2 \right] = \frac{\partial \hat{F}_{v,w}(v_C)}{\partial v_C} \cdot E \left[\left(\frac{\Delta v_C}{\Delta t} - E \left[\frac{\Delta v_C}{\Delta t} \right] \right)^2 \right] \quad (5.31)$$

donde, de acuerdo a la hipótesis de independencia de las realizaciones x_t , y consecuentemente v_t , y la distribución estadística de cada variable aleatoria v_t , de media y varianza unidad, se deduce a partir de ec. 5.32 y 5.33, las expresiones 5.34 y 5.35 para los momentos estadísticos de primer y segundo orden, respectivamente.

$$\begin{aligned} E \left[\frac{\Delta v_C}{\Delta t} \right] &= E \left[\frac{\Delta \left(\sum_{j=t-w+1}^t v_j \right)}{\Delta t} \right] = E \left[\frac{\sum_{j=t+1}^{t+\Delta t} v_j - \sum_{j=t-w+1}^{t-w+\Delta t} v_j}{\Delta t} \right] = \\ &= \frac{E \left[\sum_{j=t+1}^{t+\Delta t} v_j \right] - E \left[\sum_{j=t-w+1}^{t-w+\Delta t} v_j \right]}{\Delta t} = \frac{\Delta t - \Delta t}{\Delta t} = 0 \quad (5.32) \end{aligned}$$

$$\begin{aligned}
E \left[\left(\frac{\Delta v_C}{\Delta t} - E \left[\frac{\Delta v_C}{\Delta t} \right] \right)^2 \right] &= E \left[\left(\frac{\Delta \left(\sum_{j=t-w+1}^t v_j \right)}{\Delta t} \right)^2 \right] = \\
&= E \left[\frac{\left(\sum_{j=t+1}^{t+\Delta t} v_j - \sum_{j=t-w+1}^{t-w+\Delta t} v_j \right)^2}{\Delta t^2} \right] = \frac{E \left[\left(\sum_{j=t+1}^{t+\Delta t} v_j + \Delta t - \sum_{j=t-w+1}^{t-w+\Delta t} v_j - \Delta t \right)^2 \right]}{\Delta t^2} = \\
&= \frac{E \left[\left(\sum_{j=t+1}^{t+\Delta t} v_j + \Delta t \right)^2 \right] + E \left[\left(\sum_{j=t-w+1}^{t-w+\Delta t} v_j + \Delta t \right)^2 \right]}{\Delta t^2} + \\
&\quad + \frac{2 \cdot E \left[\left(\sum_{j=t+1}^{t+\Delta t} v_j + \Delta t \right) \cdot \left(\sum_{j=t-w+1}^{t-w+\Delta t} v_j + \Delta t \right) \right]}{\Delta t^2} = \\
&= \frac{\Delta t + \Delta t + 2 \cdot 0}{\Delta t^2} = \frac{2}{\Delta t} \quad (5.33)
\end{aligned}$$

$$E \left[\frac{\Delta \hat{F}_{v,w}(v_C)}{\Delta t} \right] = 0 \quad (5.34)$$

$$E \left[\left(\frac{\Delta \hat{F}_{v,w}(v_C)}{\Delta t} - E \left[\frac{\Delta \hat{F}_{v,w}(v_C)}{\Delta t} \right] \right)^2 \right] = \frac{\partial \hat{F}_{v,w}(v_C)}{\partial v_C} \cdot \frac{2}{\Delta t} \quad (5.35)$$

Las relaciones 5.34 y 5.35, demuestran que la capacidad de variación unitaria de la *f.p.e.* con respecto a un cambio unitario en t , sólo depende de la derivada de la *f.p.e.* con respecto a v_C , la cual es menor para valores crecientes de w . Ello se ve reflejado en un aumento de la persistencia en los valores de la *f.p.e.* con el aumento de la w .

5.1.2. Análisis comparativo de los criterios de caracterización de secuencias

Con el fin de comparar los resultados obtenidos en la caracterización con los diferentes criterios tradicionales, y compararlos con el nuevo criterio presentado, se ha supuesto un simple ejemplo:

Supóngase una variable aleatoria normal, en la que no se observa dependencia temporal significativa, y que sigue una distribución lognormal, de media $\mu = 50 \text{ ud}$ y desviación típica $\sigma = 35 \text{ ud}$. Sobre esta variable, se va a suponer la producción de una serie de secuencias de realizaciones de valores relativamente bajos, que podrían considerarse períodos de sequía. Para incluir en el análisis los criterios derivados del análisis de rachas, se adoptará un umbral igual al valor del primer cuartil de la variable aleatoria, $U = 30,192$. Así se presenta la tabla 5.1 con los resultados obtenidos para cada uno de los criterios antes descritos.

Caso	w	$Y_{i,w}$	$\bar{X}_{i,w}$	$X_{\min,i,w}$	D	I_{\max}	\bar{I}	S	F_U	\hat{F}	$\frac{1}{1-\hat{F}}$	Secuencia
1	6	143	23,8	15	6	15,2	6,4	38,2	0,56	$2 \cdot 10^{-6}$	122	[30 20 15 29 19 30]
2	6	145	24,2	15	4	15,2	9,4	37,8	0,69	$2 \cdot 10^{-6}$	110	[31 20 15 29 19 31]
3	6	145	24,2	15	–	–	–	–	–	$2 \cdot 10^{-6}$	110	[30 20 15 31 19 30]
4	3	65	21,7	15	3	15,2	8,5	25,6	0,41	$4 \cdot 10^{-4}$	59	[30 20 15]
5	2	48	24	19	2	11,2	6,2	12,4	0,09	$2 \cdot 10^{-2}$	13	[19 30]
6	6	143	23,8	20,75	6	9,4	6,4	38,2	1,01	$3 \cdot 10^{-6}$	76	[30 20,75 20,75 20,75 20,75 30]
7	6	143	23,8	8	6	22,2	6,4	38,2	0,3	$1 \cdot 10^{-7}$	782	[30 29 29 17 8 30]
8	6	169	28,2	19	6	11,2	2,0	12,2	0,04	$6 \cdot 10^{-5}$	13	[30 30 30 30 19 30]

Tabla 5.1: Ejemplos de aplicación de los diferentes criterios para la caracterización de secuencias, en una variable lognormal, con umbral para los criterios de rachas en el primer cuartil (30.19).

En la tabla 5.1 se presentan un conjunto de casos, generados modificando una secuencia de realizaciones de longitud 6. Este tipo de longitudes es habitual en análisis de fenómenos de sequía, con una escala temporal anual, donde los valores de la secuencia podrían tratarse de aportaciones anuales, por ejemplo. Se han analizado todos los criterios descritos, incluido el uso de la *f.e.* y la *f.p.e.* Para el caso de la función de persistencia extrema, se ha tabulado la inversa de su complementaria, con el objeto de mostrar mejor su sensibilidad. El caso 1 presenta una secuencia con dos valores mínimos locales. La definición del umbral puede producir alta sensibilidad en alguna de las magnitudes relacionadas con rachas. Comparando el caso 1 con el caso 2, el cambio se produce en los valores inicial y final, que en el segundo caso se encuentran por encima del umbral, y por tanto fuera de la racha. La principal magnitud que sufre cambios es la duración, seguida por la intensidad media y el factor de uniformidad. Las magnitudes o criterios relativos a la caracterización de la secuencia, como son la suma acumulada, el valor promedio, o el valor mínimo, no sufren cambios significativos. Lo mismo ocurre con la severidad y la intensidad máxima en el resto de magnitudes sobre rachas. Respecto a la sensibilidad sobre la *f.e.* y la *f.p.e.*, no son sensibles a estos cambios, pues desviación de las secuencias 1 y 2 hacia la región extrema son prácticamente iguales. En el caso 3, se ha aumentado el cuarto valor por encima del umbral, de modo que de acuerdo a la teoría de rachas, en la secuencia se incluirían dos eventos de sequía, analizados en el caso 4 y 5. Al igual que ocurría en la comparación entre los casos 1 y 2, ninguno de las magnitudes referidas en el análisis de la secuencia sufren variación significativa, sin embargo si sufren fuerte variación los criterios derivados del análisis de rachas, a excepción de la intensidad máxima en el

caso 4. En los casos 4 y 5 se puede observar la influencia de la longitud de la secuencia en el valor de la *f.e.*, de órdenes de magnitud diferentes a las secuencias de longitud 6. Igualmente, al reducirse la el número de valores próximos a la región extremal, el uso de la *f.p.e.* califica los casos como de menor desviación que los anteriores. Finalmente, se han incluido tres casos adicionales, para comprobar el comportamiento de la función *f.p.e.* En el caso 6, se ha supuesto una secuencia con un valor medio idéntico al del caso 1, pero con una mayor uniformidad. Esto tiene su reflejo en el factor de uniformidad, que alcanza el máximo de entre todos los casos, al igual que en el uso de la *f.p.e.*, que lo cuantifica como una secuencia de menor desviación hacia la región extremal que en los casos 1 a 3. Si en la secuencia se introduce un valor notablemente más extremo, ajustándose los vecinos para tener un valor promedio igual que en el caso 1, la *f.p.e.* discrimina fuertemente este caso, al igual que produce sensibilidad en la *f.e.*, el valor mínimo, la intensidad máxima y el factor de uniformidad (caso 7). El resto de criterios no captan la influencia de un valor significativamente más extremos que los antes presentados. Por último, el caso 8 está basado en el caso 5, en el que se ha aumentado la longitud de la secuencia mediante valores próximos al cuantil primero, que pueden ser considerados no significativamente extremos. Los resultados muestran la no variación de la *f.p.e.* con respecto al caso 5, por producirse el mismo valor extremo y no aparecer en ninguno de los casos efectos de persistencias, como puede encontrarse en el caso 6. Si el caso 8 hubiese sido completado, con respecto del caso 5, con valores alejados de la región extremal, el valor de la *f.p.e.* hubiera cuantificado la desviación global en la secuencia como una desviación menos extrema, al diluirse el efecto de realizaciones próximas con realizaciones alejadas de la región extremal.

A partir de los casos presentados, se puede concluir que el único criterio, entre los estudiados, que proporciona una caracterización general adecuada es el que hace uso de la *f.p.e.*, el resto de parámetros o bien son insensibles a determinados cambios significativos, o de forma inversa, presentan una fuerte sensibilidad a cambios no significativos. La *f.p.e.* grava tanto la ocurrencia de valores extremos, así como la persistencia de valores próximos a la región extremal. A estas conclusiones hay que añadir las ventajas que supone el uso de un criterio de naturaleza probabilística, en lo que se refiere a la posibilidad de comparación entre resultados de aplicar la metodología en lugares con regímenes de variación diferentes.

5.2. Propiedades de recurrencia de desviaciones persistentes extremas en poblaciones con independencia temporal

La caracterización de desviaciones persistentes hacia regiones extremas, en secuencias de realizaciones de una variable aleatoria, a partir de la *f.p.e.* proporciona un criterio general adecuado para la caracterización de realizaciones de eventos de sequías extremos. Este criterio facilita una cuantificación de tipo estadístico referida a la significancia de la desviación que se ha producido en una secuencia. Sin embargo, de su cuantificación no se deduce de forma directa propiedades de recurrencia de desviaciones de similar o mayor significancia. En la práctica, la caracterización de eventos extremos en términos de propiedades de su recurrencia, por ejemplo la recurrencia media, proporciona una valoración mucho más directa y fácil de interpretar que los valores de probabilidades acumuladas, además de poseer mejor comportamiento frente a cambios en la escala temporal (Capítulo 6). Así sería deseable establecer una relación entre la caracterización probabilista expuesta en la sección anterior, y las propiedades estocásticas de ocurrencia. Ello permitiría realizar la caracterización en términos estocásticos, en lugar de en términos estadísticos.

En esta sección se analizan las propiedades de recurrencia de desviaciones persistentes de realizaciones de una variable aleatoria, caracterizadas de acuerdo a la *f.p.e.*, bajo la hipótesis de independencia de las realizaciones.

5.2.1. Modelación de la recurrencia de desviaciones persistentes extremas a partir de la Teoría de Rachas

La recurrente aparición de eventos considerados como desviaciones persistentes extremas, puede ser modelada haciendo uso de la Teoría de Rachas (Yevjevich, 1963). Defínase un determinado nivel de significancia de desviaciones persistentes hacia la región extremal objeto de análisis, a partir de un valor crítico de la *f.p.e.* (\hat{F}_C). Para una determinada duración de las secuencias analizadas, w , en una serie de realizaciones, se podrían observar períodos en los que las secuencias de longitud w arrojan un valor de la *f.p.e.* inferior al crítico, y períodos en los que resulta superior al crítico. En lo sucesivo, definida una longitud máxima de las secuencias a analizar W , en un determinado instante, t , si alguna de las secuencias terminadas en dicho instante, de longitud menor o igual que W , $w \leq W$, mantiene un valor de la *f.p.e.* mayor o igual que el crítico, dicho instante se considera que

la señal está en estado de stress (cumple la condición dada en ec. 5.36). A los períodos formados por las realizaciones que cumplen la condición en ec. 5.36 se les denominará como períodos de stress (\mathcal{S}), mientras que aquellos períodos que no la cumplen, se les denominará períodos de no-stress (\mathcal{N}).

$$\hat{F}_{x,w}(X_{t-w+1}, X_{t-w+2}, \dots, X_t) \geq \hat{F}_C \quad \text{para algún } w \leq W \quad (5.36)$$

La longitud máxima de las secuencias a analizar, W , va a influir en los resultados del análisis, indicando el máximo nivel retrospectivo alcanzado. Así mismo, el efecto notado en la sección previa, del aumento de la persistencia en los valores de la *f.p.e.* con mayores valores de w , va a tener influencia en las propiedades de recurrencia de las desviaciones persistentes extremas caracterizadas con un valor fijado de W .

A partir de este modelo conceptual, el análisis de las propiedades de recurrencia de los períodos \mathcal{S} para un determinado valor crítico, \hat{F}_C , y longitud máxima de las secuencias, W , pasa por el análisis de las propiedades estadísticas de las longitudes de los períodos de \mathcal{S} y \mathcal{N} , así como las propiedades de correlación entre las duraciones de las diferentes rachas. De este modo, la función de probabilidad (*f.p.*) de la duración de los períodos \mathcal{N} , $d_{\mathcal{N}}$, puede ser expresada haciendo uso del Teorema de Bayes (Kolmogorov, 1956):

$$f_{\mathcal{N}}(W, \hat{F}_C, d) = \frac{\mathcal{P}(\mathcal{S}_0 \cap \mathcal{N}_1 \cap \dots \cap \mathcal{N}_d \cap \mathcal{S}_{d+1})}{\mathcal{P}(\mathcal{S}_0 \cap \mathcal{N}_1)} \quad (5.37)$$

con

$$\mathcal{P}(\mathcal{S}_k) = \mathcal{P}\left(\bigcap_{w=1}^W \hat{F}_{x,w}(X_{k-w+1}, X_{k-w+2}, \dots, X_k) \geq \hat{F}_C\right) \quad (5.38)$$

$$\mathcal{P}(\mathcal{N}_k) = \mathcal{P}\left(\bigcup_{w=1}^W \hat{F}_{x,w}(X_{k-w+1}, X_{k-w+2}, \dots, X_k) < \hat{F}_C\right) \quad (5.39)$$

y $k = 0, 1, \dots, d+1$. Donde el símbolo \cap indica el necesario cumplimiento de ambas condiciones conectadas (suceso intersección), y el símbolo \cup indica el suceso donde se cumple al menos alguna de las condiciones conectadas (suceso unión). De forma similar, en el caso de las duraciones de los períodos \mathcal{S} su *f.p.* resultaría:

$$f_{\mathcal{S}}(W, \hat{F}_C, d) = \frac{\mathcal{P}(\mathcal{N}_0 \cap \mathcal{S}_1 \cap \dots \cap \mathcal{S}_d \cap \mathcal{N}_{d+1})}{\mathcal{P}(\mathcal{N}_0 \cap \mathcal{S}_1)} \quad (5.40)$$

El desarrollo de las funciones expresadas en 5.37, y 5.40 conduce a expresiones compuestas por suma de términos, que cada uno de ellos se desarrolla en la forma de integrales múltiples. Como un ejemplo, y dado que será referida posterior, se muestra la resolución analítica en forma de integral múltiple resultante de los numeradores y denominadores de las funciones 5.37 y 5.40 para $W = 2$ y $d = 1$ en las expresiones 5.41–5.42 y 5.43–5.44, respectivamente. La resolución analítica de dichas integrales múltiples, aunque mantiene solución conocida en todos los casos, no resulta práctico a efectos computacionales, por su elevada complejidad a medida que los valores de W , y d aumentan. Por esta razón, en lugar de su resolución analítica, se ha optado por su resolución numérica, aplicando metodos para la resolución de integrales múltiples de gran dimensión. El Apéndice A contiene una descripción de los métodos disponibles para el cálculo numérico de integrales múltiples, así como el método aquí adoptado. El análisis de los resultados de las resoluciones numéricas permite extraer una serie de propiedades, que se describen en la siguiente sección.

$$\begin{aligned} \mathcal{P}(\mathcal{S}_0 \cup \mathcal{N}_1 \cup \mathcal{S}_2) = & \\ & \int_{V_{C1}}^{V_{C2}-V_{C1}} \exp(-t_1) \cdot \int_{V_{C2}-t_1}^{V_{C1}} \exp(-t_2) \cdot \int_0^{V_{C2}-t_2} \exp(-V_{C2}) \cdot dt_3 \cdot dt_2 \cdot dt_1 + \dots \\ & \dots + \exp(-V_{C1}) \cdot \hat{F}_C \quad (5.41) \end{aligned}$$

$$\begin{aligned} \mathcal{P}(\mathcal{S}_0 \cap \mathcal{N}_1) = & \int_{V_{C1}}^{V_{C2}-V_{C1}} \exp(-t_1) \cdot \int_0^{V_{C1}} \exp(-t_2) \cdot \int_{V_{C2}-t_2}^{+\infty} \exp(-t_3) \cdot dt_3 \cdot dt_2 \cdot dt_1 + \dots \\ & \dots + \int_0^{V_{C2}-V_{C1}} \exp(-t_1) \cdot \int_{V_{C1}}^{V_{C2}-t_1} \exp(-t_2) \cdot dt_2 \cdot dt_1 + \dots \\ & \dots + \int_{V_{C2}-V_{C1}}^{V_{C2}} \exp(-t_1) \cdot \int_0^{V_{C2}-t_1} \exp(-t_2) \cdot \int_{V_{C2}-t_2}^{+\infty} \exp(-t_3) \cdot dt_3 \cdot dt_2 \cdot dt_1 \\ & \dots \quad (5.42) \end{aligned}$$

$$\begin{aligned}
\mathcal{P}(\mathcal{N}_0 \cap \mathcal{S}_1 \cap \mathcal{N}_2) = & \\
& \int_{V_{C1}}^{V_{C2}-V_{C1}} \exp(-t_1) \cdot \int_{V_{C2}-V_{C1}}^{V_{C1}} \exp(-t_2) \cdot \int_{V_{C2}-t_2}^{V_{C1}} [1 - \exp(-V_{C2} + t_3)] \cdot \exp(-t_3) \cdot dt_3 \cdot dt_2 \cdot dt_1 + \dots \\
& \dots + \int_{V_{C2}-V_{C1}}^{V_{C1}} \exp(-t_1) \cdot \int_{V_{C2}-V_{C1}}^{V_{C2}-t_1} \exp(-t_2) \cdot \int_{V_{C2}-t_2}^{V_{C1}} [1 - \exp(-V_{C2} + t_3)] \cdot \exp(-t_3) \cdot dt_3 \cdot dt_2 \cdot dt_1 + \dots \\
& \dots + \int_0^{V_{C2}-V_{C1}} [\exp(-V_{C1}) - \exp(-V_{C2} + t_3)] \cdot \exp(-t_1) \cdot \hat{F}_C \cdot dt_3 \cdot dt_2 \cdot dt_1
\end{aligned} \tag{5.43}$$

$$\begin{aligned}
\mathcal{P}(\mathcal{N}_0 \cap \mathcal{S}_1) = & \\
& \int_{V_{C2}-V_{C1}}^{V_{C1}} \exp(-t_1) \cdot \int_{V_{C2}-t_1}^{V_{C1}} \exp(-t_2) \cdot \int_0^{V_{C2}-t_2} \exp(-t_3) \cdot dt_3 \cdot dt_2 \cdot dt_1 + \int_{V_{C1}}^{+\infty} \exp(-t_1) \cdot \hat{F}_C \cdot dt_1
\end{aligned} \tag{5.44}$$

con

$$V_{Ck} / \hat{F}_{v,w=k}(V_{Ck}) = \hat{F}_C \tag{5.45}$$

5.2.2. Propiedades de las funciones de probabilidad de las duraciones de rachas.

A partir de la estimación numérica de las expresiones 5.37 y 5.40, mediante métodos numéricos para la resolución de integrales múltiples (vease Apéndice A), se ha explorado las propiedades que las funciones de probabilidad de las duraciones que períodos \mathcal{N} y \mathcal{S} tienen.

Función de probabilidad de la duración de períodos \mathcal{N} Las estimaciones numéricas de la *f.p.* de la duración de períodos \mathcal{N} muestra una relación lineal cuando son representadas en escala semi-logarítmica las estimaciones de la *f.p.* contra la duración (figura 5.3). Esto motiva inferir que la distribución teórica seguida por la duración de los períodos \mathcal{N} es una distribución geométrica (ec. 5.46), con un único parámetro $p_{\mathcal{N}}$.

$$f_{\mathcal{N}}(W, \hat{F}_C, d) = p_{\mathcal{N}}(W, \hat{F}_C) \cdot [1 - p_{\mathcal{N}}(W, \hat{F}_C)]^{d-1} \tag{5.46}$$

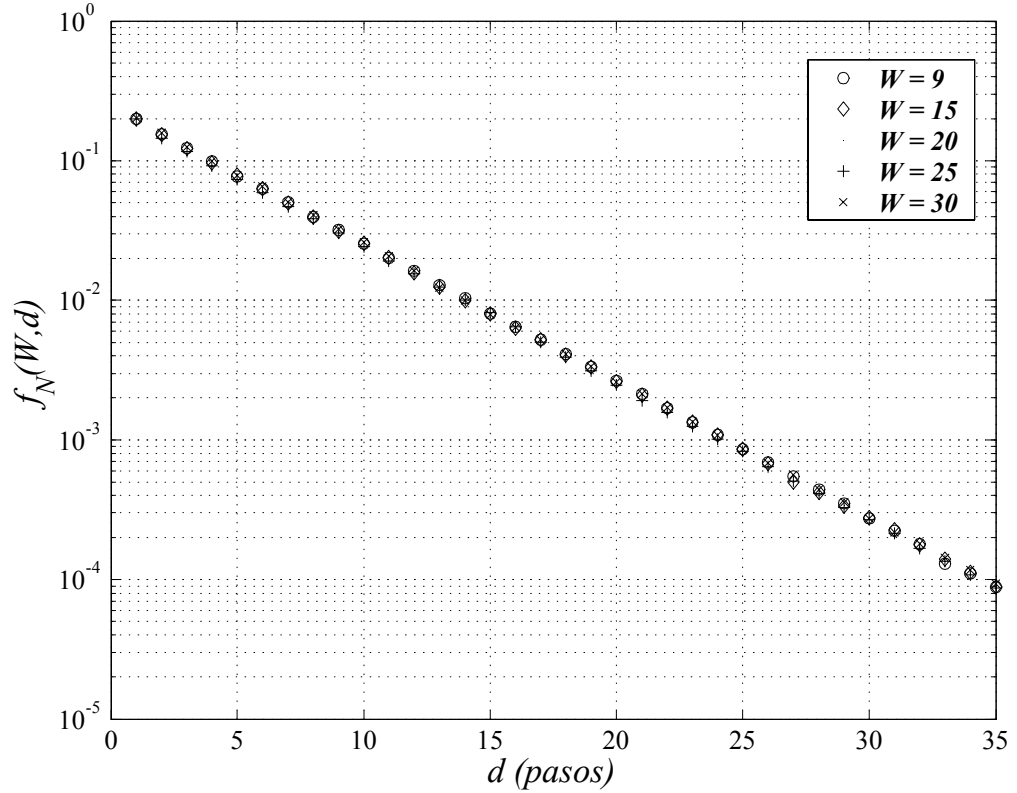


Figura 5.3: Dependencia de la *f.p.* de la duración de periodos \mathcal{N} con d y W , estimada a partir de resolución numérica, para $\hat{F}_C=0.8$.

Así mismo, la representación gráfica simultánea de los puntos estimados para los diferentes valores de W refleja la invarianza del parámetro de la distribución frente a cambios en W (figura 5.3). Por otro lado, sí que se observa dependencia con el valor de \hat{F}_C adoptado, resultando \hat{F}_C la única variable de la que depende $p_{\mathcal{N}}$, y por tanto $f_{\mathcal{N}}$.

$$f_{\mathcal{N}}(\hat{F}_C, d) = p_{\mathcal{N}}(\hat{F}_C) \cdot [1 - p_{\mathcal{N}}(\hat{F}_C)]^{d-1} \quad (5.47)$$

De acuerdo a estas propiedades, la *f.p.* de la duración de los periodos \mathcal{N} puede determinarse para un valor de \hat{F}_C a partir de 5.41 y 5.42, que determinan en 5.37 el valor de la *f.p.* de la duración para $W = 2$ y $d = 1$, cuyo valor permite estimar $p_{\mathcal{N}}(\hat{F}_C)$ mediante ec. 5.47. De este modo, la distribución estadística de las duraciones de periodos \mathcal{N} sigue el modelo de distribución estadística geométrica, cuyo parámetro puede ser estimado a partir de 5.41–5.42 y 5.47.

El valor del parámetro de la *f.p.* de la duración de períodos \mathcal{N} disminuye con el incremento de \hat{F}_C . Dado que la media de la distribución geométrica (ec. 5.47) coincide con la inversa de su parámetro $p_{\mathcal{N}}$, el valor esperado de la duración de los períodos \mathcal{N} aumenta al incrementarse \hat{F}_C .

Función de probabilidad de la duración de períodos \mathcal{S} . De igual forma a la exploración realizada con la *f.p.* de la duración de períodos \mathcal{N} , ha sido llevado a cabo el análisis para la *f.p.* de la duración de períodos \mathcal{S} . Los resultados para un valor fijo de \hat{F}_C , y variaciones de W , son mostrados en la figura 5.4, donde se ha representado en escala doble logarítmica los ratios de la *f.p.* y la función de probabilidad acumulada (*f.p.a.*), contra la duración d . Los resultados muestran una dependencia de las propiedades de las distribuciones en W , con una propiedad invariante: Para un valor fijo de \hat{F}_C , los ratios *f.p./f.p.a.* de la duración de los períodos \mathcal{S} para $d < W$ siguen una relación lineal con d . La ecuación de esta recta es independiente de W , variable que únicamente influye en el rango de aplicación de la relación. Así pues el modelo teórico seguido en el rango $0 < d < W$, con $W > 1$, viene dado por la siguiente distribución:

$$\frac{f_{\mathcal{S}}(W, \hat{F}_C, d)}{F_{\mathcal{S}}(W, \hat{F}_C, d)} = \frac{1}{d^{\alpha(\hat{F}_C)}}, \text{ para } 1 < d < W \quad (5.48)$$

$$f_{\mathcal{S}}(W, \hat{F}_C, d) = \zeta(\hat{F}_C), \text{ para } d = 1 \quad (5.49)$$

Ec. 5.48 y 5.49 derivan para $W > 2$ a la *f.p.* de la duración de períodos \mathcal{N} definida en 5.50.

$$f_{\mathcal{S}}(W, \hat{F}_C, d) = \zeta(\hat{F}_C) \cdot \frac{\prod_{n=1}^{d-1} n^{\alpha(\hat{F}_C)}}{\prod_{n=2}^d [n^{\alpha(\hat{F}_C)} - 1]}, \text{ para } 1 < d < W \quad (5.50)$$

La distribución tiene dos parámetros, α y ζ , ambos función de \hat{F}_C . El primer parámetro, α , está relacionado con la forma de la función, mientras que el segundo parámetro, ζ , es un parámetro de escala. El parámetro de escala puede ser estimado a partir de ec. 5.43 y 5.44, mediante ec. 5.40, siguiendo ec. 5.49. Mientras que el parámetro de forma, α , requiere para su estimación directa el cálculo numérico de al menos dos puntos en la *f.p.* de duraciones de períodos \mathcal{S} (ec. 5.40). La figura 5.5 representa los ratios *f.p./f.p.a.* en función de d , para distintos \hat{F}_C y mantenido constante

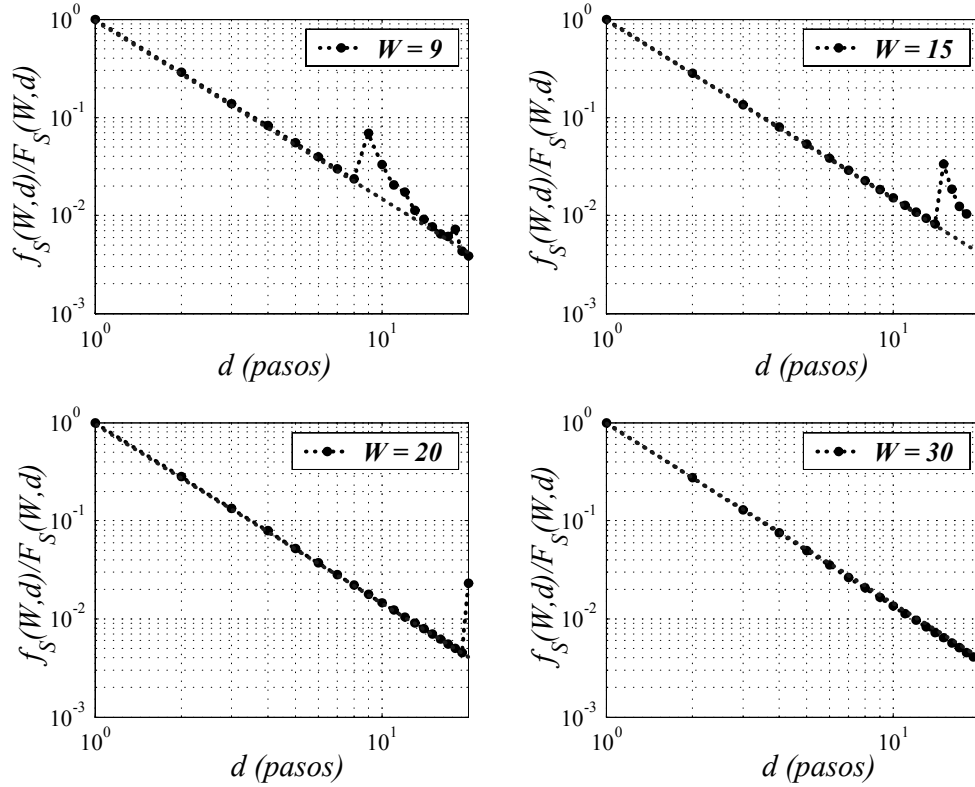


Figura 5.4: Representación doble logarítmica del cociente entre la *f.p.* y la *f.p.a.* de la duración de los períodos \mathcal{S} versus la duración; obtenido mediante resolución numérica, con $\hat{F}_C = 0,8$ y $W = 5, 10, 15, 20$ pasos.

W , obtenidos a partir de la evaluación directa de la función analítica mediante resolución numérica. Alternativamente, podría pensarse en fijar el parámetro α mediante la condición de que suponiendo que el rango de validez de 5.50 fuera ilimitado ($W \rightarrow +\infty$), la hipotética *f.p.* definida por 5.49 y 5.50 debería cumplir la condición $\sum_{d=1}^{+\infty} f_S(\hat{F}_C, d) = 1$. Sin embargo, la suma infinita de la serie en d , formada por 5.49 y 5.50 converge con $\alpha > 1$ (Knopp, 1956), por lo que esta debería ser una condición a cumplir por α (cond. 1). Asumiendo que se cumpliera la cond. 1, la suma infinita de la serie que estima la media de la duración $\left(\sum_{d=1}^{+\infty} d \cdot f_S(\hat{F}_C, d) \right)$ requiere para su convergencia que $\alpha > 2$ (cond. 2). Los resultados numéricos muestran que el rango de valores que adopta el parámetro de forma, α , cumple la cond. 1, pero no la cond. 2 (figuras 5.4 y 5.5). Este análisis permite realizar una serie de conclusiones:

- Para $W \rightarrow +\infty$, la *f.p.* de la duración de los períodos \mathcal{S} tiende a la distribución dada por ec. 5.49 y 5.50, por lo que el valor esperado de la duración del período \mathcal{S} es impropio, al no cumplirse la cond. 2. Ello está relacionado con el aumento de la persistencia de la función $\hat{F}_{x,w}$ para valores creciente de w .
- Para valores finitos de W , el valor esperado de la duración de los períodos \mathcal{S} es finito, y por tanto se tiene la convergencia de la probabilidad acumulada y el valor medio de la *f.p.* en el rango $d \geq W$, cuya estructura no se ha llegado a ser identificada.

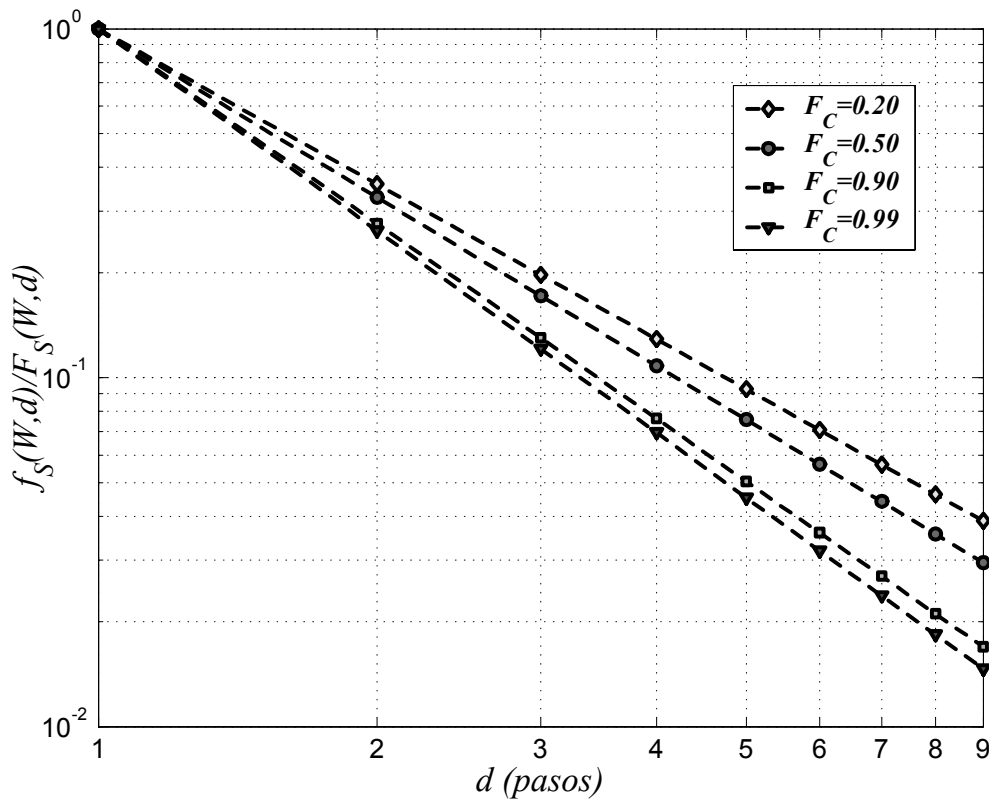


Figura 5.5: Representación doble logarítmica del cociente entre la *f.p.* y la *f.p.a.* de la duración de los períodos \mathcal{S} versus la duración; obtenido mediante resolución numérica, con $\hat{F}_C = 0,2, 0,5, 0,9, 0,99$ y $W = 10$ pasos.

De acuerdo con estas conclusiones, a diferencia de lo que ocurriría en el caso de la determinación de la distribución estadística de la duración de

períodos \mathcal{N} , para períodos \mathcal{S} no ha sido posible identificar completamente la distribución estadística de su duración. No obstante, el análisis muestra la importante influencia que produce W en la estimación del valor esperado de la duración de los períodos \mathcal{S} , para un determinado valor de \hat{F}_C . Valores mayores de W suponen una menor capacidad de variación de la función $\hat{F}_{x,w}$ para un incremento de t fijo, lo que se deriva en un valor esperado de la duración de los períodos \mathcal{S} mayor.

A partir de los análisis realizados de las propiedades estadísticas de las variables aleatorias duraciones de los períodos \mathcal{N} y \mathcal{S} , a continuación se busca una función que permita relacionar la recurrencia esperada de períodos \mathcal{S} , con los valores de W y \hat{F}_C adoptados. Para ello se definirá la función de probabilidad acumulada de persistencias extremas, y a partir de estimaciones de la recurrencia esperada de períodos \mathcal{S} , obtenida a partir de la generación de series sintética, se establecerá una relación aproximada entre esta función y los valores de la recurrencia esperada.

5.2.3. La función de probabilidad acumulada de persistencias extremas

Dada una secuencia de realizaciones de una variable aleatoria x , *i.i.d.*, de longitud W , $X_{t-W+1}, X_{t-W+2}, \dots, X_t$, se define como el valor característico de dicha secuencia, desde el punto de vista de las desviaciones persistentes hacia la región extremal, el valor máximo alcanzado por la *f.p.e.* de las realizaciones $X_{t-w+1}, X_{t-w+2}, \dots, X_t$, con $w \leq W$. A partir de este valor característico, se define la función de probabilidad acumulada de persistencias extremas (*f.a.p.e.*) de esta secuencia de realizaciones, como la probabilidad acumulada de ocurrencia de secuencias $x_{t-w+1}, x_{t-w+2}, \dots, x_t$, con $w \leq W$, con *f.p.e.* menor que el valor característico (ec. 5.51).

$$\mathfrak{S}_{x,W}(X_{t-W+1}, X_{t-W+2}, \dots, X_t) = \mathcal{P}\{x_{t-w+1}, x_{t-w+2}, \dots, x_t / w \leq W, \hat{F}_{x,w}(x_{t-w+1}, x_{t-w+2}, \dots, x_t) \leq \hat{F}_C(X_{t-w+1}, X_{t-w+2}, \dots, X_t)\} \quad (5.51)$$

con

$$\hat{F}_C(X_{t-w+1}, X_{t-w+2}, \dots, X_t) = \max \left[\hat{F}_{x,w}(X_{t-w+1}, X_{t-w+2}, \dots, X_t) / w \leq W \right] \quad (5.52)$$

Valores de la *f.a.p.e.* próximos a 1, son reflejo de una secuencia en la que se observa una importante desviación persistente hacia la región extremal, caracterizada a partir de la *f.p.e.* Esta función, o la probabilidad que la

define, tiene un desarrollo analítico en la forma de una expresión recursiva. La expresión analítica de la ec. 5.51 en la forma de integral múltiple adopta la forma , en términos de la variable aleatoria transformada v .

$$\mathfrak{S}_{v,W}(\hat{F}_C) = \int_0^{V_{C1}} \exp(-t_1) \cdot \int_0^{V_{C2}-t_1} \exp(-t_2) \cdot \dots \cdot \int_0^{V_{CW}-\sum_{k=1}^{W-1} t_k} \exp(-t_W) \cdot dt_W \cdot \dots \cdot dt_2 \cdot dt_1 \quad (5.53)$$

donde V_{Ck} es obtenido de ec. 5.45, $k = 1, \dots, W$.

La ec. 5.51 puede resolverse de forma recurva, siguiendo 5.54.

$$\begin{aligned} \mathfrak{S}_{v,W=1}(\hat{F}_C) &= \Gamma(1, V_{C1}) \\ \mathfrak{S}_{v,W=2}(\hat{F}_C) &= \Gamma(2, V_{C2}) - V_{C1} \cdot \exp(-V_{C1}) \cdot \Gamma(1, V_{C2} - V_{C1}) \\ \mathfrak{S}_{v,W}(\hat{F}_C) &= \Gamma(W, V_{CW}) - \sum_{k=0}^{W-2} (q_k \cdot \exp(-V_{Ck+1}) \cdot \Gamma(W - k, V_{CW} - V_{Ck+1})), \\ &\text{con } W > 2 \end{aligned} \quad (5.54)$$

siendo $\Gamma(i, V_{Ci})$ la función gamma incompleta (ec. 5.27), y viniendo dado q_k por la a su vez expresión recursiva en ec. 5.55.

$$q_0 = 1 \quad (5.55a)$$

$$q_1 = V_{C1} \quad (5.55b)$$

$$q_k = \frac{V_{Ck}^k}{k!} - \sum_{j=0}^{k-2} q_j \cdot \frac{(V_{Ck} - V_{Cj+1})^{k-j}}{(k-j)!}, \text{ para } k > 1 \quad (5.55c)$$

Para demostrar la expresión 5.54, se debe realizar en primer lugar, el siguiente cambio de variable múltiple:

$$y_j = \sum_{k=j}^W t_k \quad (5.56)$$

La matriz Jacobiana del cambio de variable múltiple (ec. 5.56) resulta ec. 5.57, de donde el Jacobiano de la transformación, determinante de la matriz Jacobiana, es idéntico a la unidad.

$$J = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_W \end{bmatrix} \cdot \left[\frac{\partial}{\partial t_1} \quad \frac{\partial}{\partial t_2} \quad \cdots \quad \frac{\partial}{\partial t_W} \right] = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad (5.57)$$

Con ello, tras el cambio de variable (ec. 5.56), la ec. 5.53 adopta la forma dada en la ec. 5.58.

$$\mathfrak{S}_{v,W}(\hat{F}_C) = \int_0^{V_{CW}} \exp(-y_1) \cdot \int_{\max(0, y_1 - V_{C1})}^{y_1} \cdot \int_{\max(0, y_1 - V_{C2})}^{y_2} \cdots \int_{\max(0, y_1 - V_{C_{W-1}})}^{y_{W-1}} dy_W \cdots dy_3 \cdot dy_2 \cdot dy_1 \quad (5.58)$$

La resolución de ec. 5.58 se puede desarrollar de la siguiente forma:

$$\mathfrak{S}_{v,W}(\hat{F}_C) = \sum_{k=1}^W G(k, \hat{F}_C) \quad (5.59)$$

con

$$G(k, \hat{F}_C) = \int_{V_{C_{k-1}}}^{V_{C_k}} \exp(-y_1) \cdot \int_{y_1 - V_{C_1}}^{y_1} \cdots \int_{y_1 - V_{C_{k-1}}}^{y_{k-1}} \cdot \int_0^{y_k} \cdots \int_0^{y_{W-1}} dy_W \cdots dy_{k-1} \cdot dy_k \cdots dy_2 \cdot dy_1 \quad (5.60)$$

Por convención se asume que $V_{C_0} = 0$. La resolución de cada término 5.60, en ec. 5.59, y a través de sucesivas simplificaciones algebraicas, conduce a ec. 5.61.

$$\begin{aligned} \mathfrak{S}_{v,W}(\hat{F}_C) &= \int_0^{V_{C1}} \frac{y_1^W}{W!} \cdot \exp(-y_1) \cdot dy_1 - q_0 \cdot \int_{V_{C1}}^{V_{C2}} \frac{(y_1 - V_{C1})^W}{W!} \cdot \exp(-y_1) \cdot dy_1 - \\ &- q_1 \cdot \int_{V_{C2}}^{V_{C3}} \frac{(y_1 - V_{C2})^{W-1}}{(W-1)!} \cdot \exp(-y_1) \cdot dy_1 - \cdots - q_{W-2} \cdot \int_{V_{C_{W-1}}}^{V_{CW}} \frac{(y_1 - V_{C_{W-1}})^2}{2!} \cdot \exp(-y_1) \cdot dy_1 \end{aligned} \quad (5.61)$$

Aplicando la definición de la función gamma incompleta (ec. 5.27), ec. 5.61 se puede escribir como ec. 5.62, la cual coincide con la ec. 5.54.

$$\begin{aligned} \mathfrak{S}_{v,W}(\hat{F}_C) = & \Gamma(W, V_{CW}) - q_0 \cdot \exp(-V_{C1}) \cdot \Gamma(W-1, V_{CW} - V_{C1}) - \\ & - q_1 \cdot \exp(-V_{C1}) \cdot \Gamma(W-2, V_{CW} - V_{C2}) - \dots - q_{W-2} \cdot \exp(-V_{C1}) \cdot \Gamma(2, V_{CW} - V_{C_{W-1}}) \end{aligned} \quad (5.62)$$

La forma de la *f.a.p.e.*, para diferentes W , se muestra en la figura 5.6. Puede apreciarse que a medida que aumenta W , la distribución del valor característico de una secuencia (ec. 5.52) tiende a distribuirse en valores más próximos a 1, partiendo de una distribución uniforme para $W = 1$. Ello es reflejo de que cuanto mayor es el valor de W , mayor es la probabilidad de encontrar en la secuencia $X_{t-W+1}, X_{t-W+2}, \dots, X_t$, una subsecuencia $X_{t-w+1}, X_{t-w+2}, \dots, X_t$, con $w \leq W$, para el que la *f.p.e.* resulte mayor que un determinado umbral. La *f.a.p.e.* va a permitir establecer una relación que estime la recurrencia esperada de desviaciones persistentes hacia la región extremal, caracterizadas a partir de la *f.p.e.*

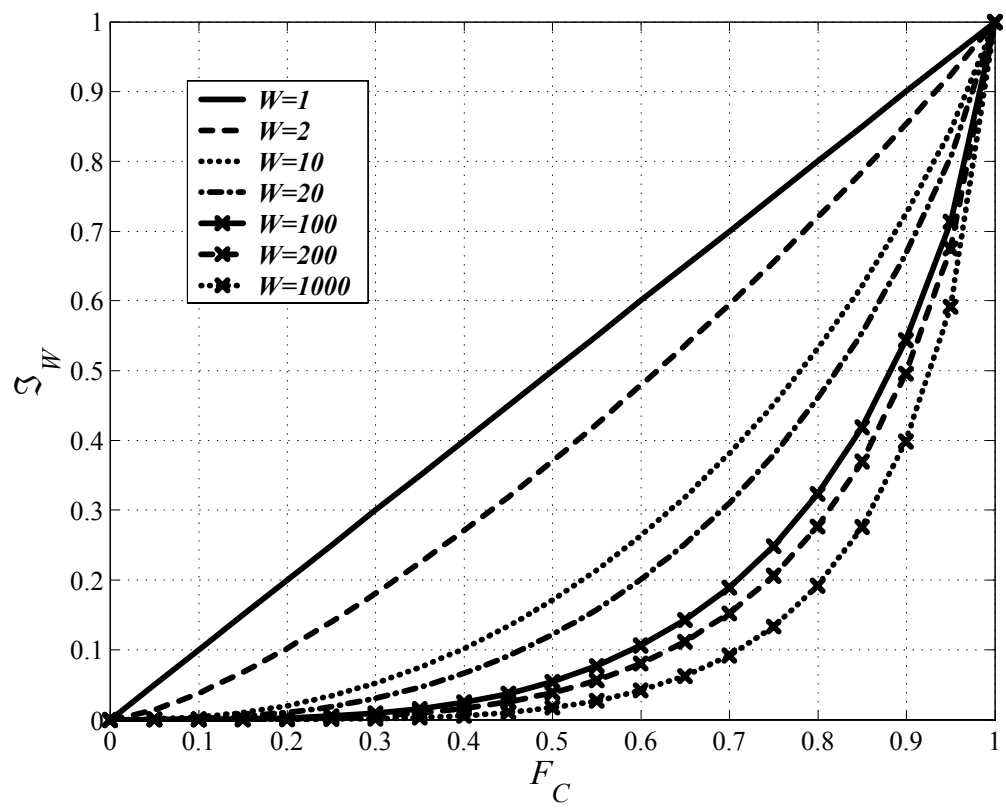


Figura 5.6: Función *f.a.p.e.* para diferentes valores de W , en función de \hat{F}_C .

5.2.4. El período de retorno de desviaciones persistentes hacia la región extremal

El análisis de las distribuciones estadísticas de las duraciones de los períodos \mathcal{N} y \mathcal{S} , a partir de la resolución numérica directa de las ecuaciones que definen las distribuciones, ha permitido deducir algunas propiedades de estas distribuciones. Sin embargo, la recurrencia esperada de desviaciones persistentes hacia la región extremal no se deriva a partir de ellas. En esta subsección se va a establecer una relación analítica que permite su estimación aproximada a partir de la *f.a.p.e.*

Para el análisis de la recurrencia esperada, se utilizará el concepto de período de retorno, definido aquí como el tiempo medio entre los instantes de inicio de dos períodos \mathcal{S} sucesivos (i.e. separados únicamente por un período \mathcal{N}). Para su estimación, se ha utilizado la generación de una serie sintética de la gran longitud (10^6 realizaciones) de una variable uniforme, *i.i.d.* Esta generación se basa en la aplicación de algoritmos de generación de números aleatorios, los cuales producen en realidad una serie determinista, cuyo valor depende del valor de la semilla (o estado) utilizado por el algoritmo, que cambia en cada realización. El alto rango de valores que puede adoptar la semilla produce que a efectos prácticos, para la mayoría de las aplicaciones, y siempre que no se intervenga el valor que va adoptando la semilla, el comportamiento de las generaciones producidas sobre las aplicaciones satisfagan los test estadísticos referidos a las propiedades de *i.i.d.* De acuerdo a los objetivos aquí buscados, el número de valores posible de la semilla marca el límite de realizaciones posibles que puede componer la serie generada, ya que al generarse una única serie, no existiría posibilidad de evaluar en más de una ocasión el algoritmo para la misma semilla. En el caso del algoritmo de generación de números aleatorios implementado en Matlab, aquí utilizado, este límite se encuentra en 10^{10} .

A partir de dicha generación, para diferentes valores de W , se ha calculado el valor característico de las secuencias (ec. 5.52) en cada instante. Con ello, previa adopción de un valor umbral \hat{F}_C , puede ser descompuesta la serie en períodos \mathcal{N} y \mathcal{S} , analizando de forma indirecta la distribución de sus duraciones. Se comprueba las propiedades analizadas de forma directa (sección 5.2.2), y se analizan las correlaciones entre las duraciones de períodos consecutivos. Los valores encontrados de los coeficientes de correlación de las duraciones de períodos consecutivos \mathcal{N} y \mathcal{S} , $\rho_{\mathcal{N}-\mathcal{S}}$, períodos sucesivos \mathcal{S} y \mathcal{S} , $\rho_{\mathcal{S}-\mathcal{S}}$, y períodos sucesivos \mathcal{N} y \mathcal{N} , $\rho_{\mathcal{N}-\mathcal{N}}$ no muestran evidencia estadística de la existencia de correlaciones, con un nivel de significancia del 5%. Ello motiva la adopción de la hipótesis de distribución *i.i.d.* para cada una de las duraciones de los períodos \mathcal{N} y \mathcal{S} . De acuerdo con la cual, el período de retorno de una situación \mathcal{S} podría escribirse como suma de los

valores esperados de las duraciones de los períodos \mathcal{N} y \mathcal{S} , para el valor de W y \hat{F}_C adoptados.

$$T(W, \hat{F}_C) = E[d_{\mathcal{N}} + d_{\mathcal{S}}] = E[d_{\mathcal{N}}] + E[d_{\mathcal{S}}] \quad (5.63)$$

El efecto de W se pone de manifiesto en su influencia sobre el valor esperado $E[d_{\mathcal{S}}]$, dado que $E[d_{\mathcal{N}}]$ no depende de W . Ambos términos depende de \hat{F}_C . La divergencia observada en $E[d_{\mathcal{S}}]$ para valores de $W \rightarrow +\infty$, con ec. 5.63, implica la divergencia también del período de retorno, al estar $E[d_{\mathcal{N}}]$ siempre acotado. A partir de la serie sintética generada, y de los análisis efectuados sobre ella, se puede realizar la estimación del período de retorno para un conjunto de parejas (W, \hat{F}_C) . Se observa que para valores altos de la *f.a.p.e.*, puede escribirse una expresión aproximada del período de retorno en función de *f.a.p.e.*, como se muestra en la figura 5.7, siguiendo ec. 5.64, cuyos parámetros óptimos resultan $\tau = 1,13$ y $\kappa = 0,20$.

$$T \approx \frac{\tau \cdot (W - 1)^\kappa}{1 - \mathfrak{S}_W(\hat{F}_C, W)} \quad (5.64)$$

Para $W = 1$, la relación se simplifica hasta $T \approx 1 / (1 - \mathfrak{S}_W(\hat{F}_C, W))$. En la figura 5.7 se incluye así mismo el intervalo de confianza de la estimación correspondiente al 95 % de probabilidad. La convergencia aproximada hacia la ec. 5.64 resulta asintótica para valores crecientes de la *f.a.p.e.* Esta relación permite estimar el período de retorno asociado al valor característico de una secuencia, para un determinado W , para valores de la *f.a.p.e.* suficientemente altos.

El conjunto de estos resultados permiten caracterizar las desviaciones persistentes extremas de una secuencia a partir de la función *f.p.e.*, y estimar el período de retorno asociado a la caracterización. Este será el procedimiento en el que se base la definición del Índice de Frecuencia de Sequías, que será definido en el capítulo siguiente. Previamente, y con el fin de complementar estos resultados, se incluyen procedimientos que permiten aplicar esta metodología a señales en las que no se cumple la hipótesis de *i.i.d.*

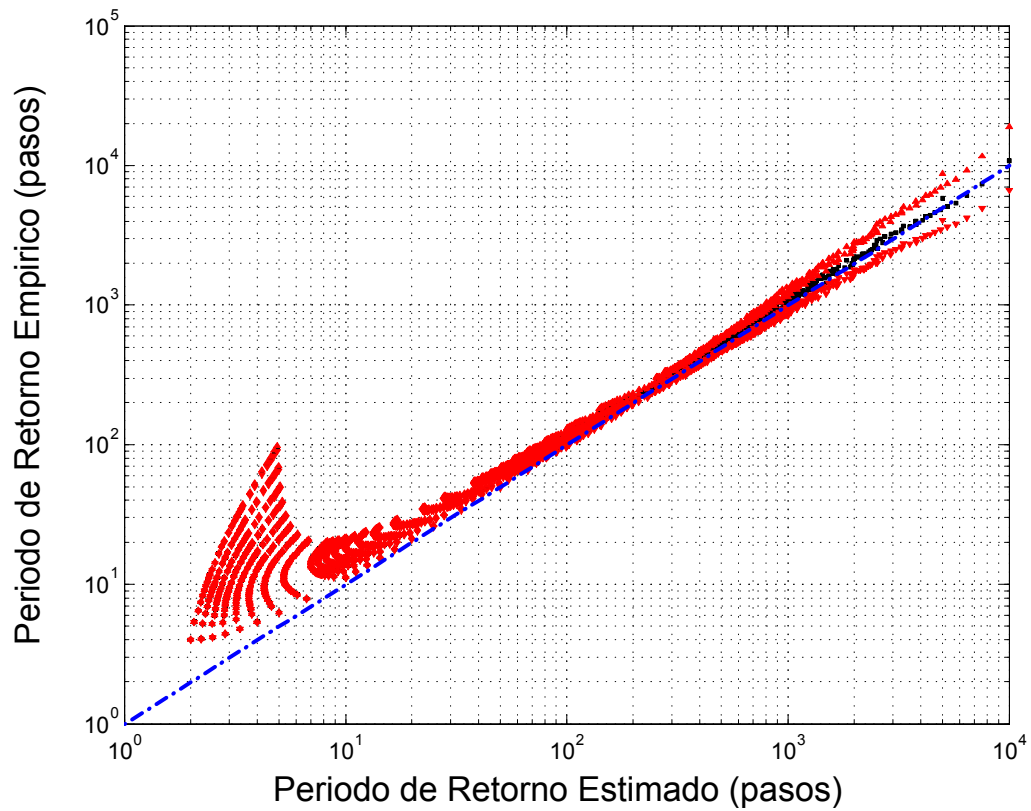


Figura 5.7: Comparación en escala doble logarítmica de los resultados de la estimación del periodo de retorno a partir de analizar la serie sintética, incluyendo el intervalo de confianza del 99 % (rojo), frente a la evaluación de la ec. 5.64, para un conjunto de W en el rango $[1,700]$ y de \hat{F}_C en el rango $[0.5,0.9999]$.

5.3. Consideraciones en poblaciones con correlación temporal

Multitud de ejemplos pueden encontrarse en registros naturales de variables hidro-climáticas en las que existe evidencia estadística del no cumplimiento de la hipótesis de independencia temporal. Hipótesis que en toda la metodología expuesta para la caracterización de secuencias, contenida en este capítulo 5, ha sido asumida. Esta sección se dedica a la descripción de un procedimiento que posibilite el cálculo, y de ese modo la aplicación de la metodología descrita, de funciones tales como las definidas en ec. 5.10, 5.12 y 5.51 ante la falta de cumplimiento de la hipótesis de independencia de las realizaciones.

En aquellos casos en los que existe una significativa dependencia de las realizaciones sobre las producidas previamente, la evaluación de probabilidades conjuntas, como son las que definen las ec. 5.10, 5.12 ó 5.51, requiere la aplicación del Teorema de Multiplicación (Kolmogorov, 1956), que se obtiene por inducción generalizando el Teorema de Bayes. La ec. 5.65 proporciona la expresión que condensa el teorema, donde A_i , $i = 1, \dots, n$, representa eventos o suceso, el símbolo \cap representa el evento conjunto o suceso intersección, y $\mathcal{P}_B(C)$ representa la probabilidad condicional de realización del suceso C , bajo la condición de haber ocurrido el suceso B .

$$\mathcal{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathcal{P}(A_1) \cdot \mathcal{P}_{A_1}(A_2) \cdot \mathcal{P}_{A_1 \cap A_2}(A_3) \cdot \dots \cdot \mathcal{P}_{A_1 \cap A_2 \cap \dots \cap A_{n-1}}(A_n) \quad (5.65)$$

La aplicación del Teorema de Multiplicación (ec. 5.65) exige la evaluación de las probabilidades condicionales. En el caso de analizar la probabilidad de ocurrencia de una secuencia, X_1, X_2, \dots, X_n , bajo determinadas condiciones, donde se manifiesta una dependencia en las realizaciones, la evaluación de las probabilidades condicionales se puede realizar a partir de la transformación de las variables, mediante la Transformación de Rosenblatt (Rosenblatt, 1952):

$$\begin{aligned} U_1 &= F(X_1) \\ U_2 &= F(X_2|X_1) \\ U_3 &= F(X_3|X_1, X_2) \\ &\dots = \dots \\ U_n &= F(X_n|X_1, X_2, \dots, X_{n-1}) \end{aligned} \quad (5.66)$$

donde F representa la función de probabilidad acumulada (ec. 5.9), y U_i representa la variable aleatoria transformada, que sigue una distribución

uniforme $\mathcal{U}(0, 1)$. La realización de la Transformación de Rosenblatt depende de la estructura de dependencia estocástica que mantengan las variables. Para ello se requiere la construcción de un modelo que simule esta estructura. En el caso de variables de tipo hidrológico, el modelo más habitualmente utilizado para la modelación de la estructura de dependencia de una variable aleatoria estacionaria, son los modelos AutoRegresivos y de Media Móvil (*ARMA*) (Box & Jenkins, 1976). A continuación se detalla un procedimiento para la realización de la Transformación de Rosenblatt en una secuencia de una variable que sigue un modelo estocástico *ARMA*. A partir de esta transformación, es posible mediante la aplicación de ec. 5.65 el computo de probabilidades como las definidas en ec. 5.10, 5.12 ó 5.51, pero ello exige la evaluación de integrales múltiples, lo que se detalla en el punto 5.3.2. La complejidad numérica que resulta de este modo la evaluación de ec. 5.12 y máxime ec. 5.51 limitaran la aplicación de esta metodología para variables con una alta dependencia temporal.

5.3.1. Transformación de Rosenblatt en variables ARMA

Los modelos estocásticos autoregresivos de media móvil (*AutoRegressive Moving Average, ARMA*) (Box & Jenkins, 1976) se definen de la siguiente forma general:

Sea una variable aleatoria estacionaria, x , que se distribuye siguiendo una distribución normal estandar $N(0, 1)$, y sea a una variable aleatoria estacionaria e independiente, que sigue una distribución normal $N(0, \sigma_a^2)$, la variable x se dice que sigue un modelo *ARMA*(p, q) cuando se cumple:

$$X_t - \phi_1 \cdot X_{t-1} - \phi_2 \cdot X_{t-2} - \dots - \phi_p \cdot X_{t-p} = a_t - \theta_1 \cdot a_{t-1} - \theta_2 \cdot a_{t-2} - \dots - \theta_q \cdot a_{t-q} \quad (5.67)$$

donde ϕ_i y θ_j , $i = 1, \dots, p$ y $j = 1, \dots, q$, definen los parámetro del modelo autoregresivo de orden p , y de media móvil de orden q , respectivamente.

En el caso que $p = 0$, el modelo se simplifica a un modelo de media móvil (*Moving Average, MA*), y en el caso que $q = 0$, el modelo se simplifica hasta un modelo autoregresivo (*AutoRegressive, AR*). En los casos en los que la variable se distribuya mediante una distribución normal $N(\mu, \sigma_x^2)$, la construcción del modelo descrito en la forma de ec. 5.67, requiere la previa tipificación de la variable, mediante la transformación a una variable

$N(0, 1)$: $z = (x - \mu) / \sigma_x$. En aquellos casos en los que la variable no se distribuya de acuerdo a una distribución normal, su transformación a una distribución normal estándar puede obtenerse mediante varios procedimientos:

- Si la variable sigue una distribución lognormal, la transformación logarítmica, o en general siguiendo ec. 5.68 produce una variable normal, donde ξ es el límite inferior de la variable aleatoria y .

$$x = \ln(y - \xi) \quad (5.68)$$

- De forma alternativa puede aplicarse una aplicación de tipo potencial, siguiendo la transformación Box-Cox (Box & Cox, 1964), en la siguiente forma:

$$x = \begin{cases} \frac{(y^\lambda - 1)}{\lambda} & \lambda \neq 0 \\ \ln(y) & \lambda = 0 \end{cases} \quad (5.69)$$

- En los casos en los que sea conocida una distribución teórica seguida por la variable, la transformación a una variable normal estándar equivalente sería la más apropiada.

$$x = \Phi(F_y(y)) \quad (5.70)$$

Debe notarse la existencia de modelos estocásticos en los que se supone una distribución teórica de la variable diferente a la distribución normal. Por ejemplo el modelo de procesos autoregresivo gamma (GAR), que supone la distribución gamma para la variable analizada (Fernández & Salas, 1990).

La adecuación del modelo *ARMA*, para la modelación de la estructura estocástica seguida por la variable aleatoria x , exige la satisfacción de las hipótesis y condiciones del modelo:

- La variable aleatoria x debe seguir una distribución $N(0, 1)$
- La variable aleatoria a debe mostrar independencia temporal, e independencia de la variable x , y seguir una distribución $N(0, \sigma_a^2)$
- Con $p > 0$, la parte autoregresiva del modelo se dice que es estable, cuando se tiene que todas las raíces del polinomio característico de la parte autoregresiva (ec. 5.71) cumplen que $|\lambda_i| < 1$, $i = 1, \dots, p$.

$$\lambda^p - \phi_1 \cdot \lambda^{p-1} - \phi_2 \cdot \lambda^{p-2} - \dots - \phi_p = 0 \quad (5.71)$$

Bajo la condición de estabilidad se asegura que la varianza de la variable aleatoria x que simula el modelo esté acotada, y el modelo *ARMA* pueda ser escrito de forma equivalente como un modelo de media móvil (*MA*) de orden finito.

- Con $q > 0$, el modelo se dice que es invertible, si se tiene que todas las raíces del polinomio característico de la parte de media móvil (ec. 5.72) cumplen que $|v_j| < 1$, $j = 1, \dots, q$.

$$v^q - \theta_1 \cdot v^{q-1} - \theta_2 \cdot v^{q-2} - \dots - \theta_p = 0 \quad (5.72)$$

La condición de invertibilidad implica la existencia de un modelo equivalente autoregresivo (*AR*) de orden finito, para el que la varianza de la variable a_t está acotada.

Bajo estas hipótesis y condiciones, la variable aleatoria x_t producida en el instante t , que sigue un modelo $ARMA(p, q)$, depende hasta el orden p de las variables x_{t-i} , y hasta el orden q de las variables a_{t-j} . En el supuesto de ser conocidas todas las variable x_i e a_j hasta el instante $t - 1$, cada nueva realización de x_t viene dada por la realización de a_t (ec. 5.67). De este modo, a partir de ec. 5.67, toda condición impuesta sobre x_t puede ser escrita de forma equivalente sobre a_t . La evaluación de la probabilidad del cumplimiento de condiciones sobre a_t resulta sencilla, al ser ésta una variable independiente.

En el caso de las probabilidades condicionales requeridas para la Transformación de Rosenblatt, resulta de especial utilidad la transformación del modelo $ARMA(p, q)$ en el modelo $AR(n)$ equivalente, asegurado bajo el cumplimiento de las condiciones de invertibilidad. De este modo, conocidos los n valores previos de x , a la ocurrencia de x_t , es posible expresar cualquier condición sobre la realización de x_t de forma equivalente sobre a_t , y con ello evaluar la probabilidad de su ocurrencia (ec. 5.73).

$$a_t = x_t - I_1 \cdot x_{t-1} - I_2 \cdot x_{t-2} - \dots - I_n \cdot x_{t-n} \quad (5.73)$$

Cuando son conocidos valores previos a x_t hasta un desfase inferior $m < n$, no es posible la aplicación de la transformación dada por ec. 5.73 para evaluar probabilidades condicionadas sobre la nueva realización x_t . Sin embargo, si la variable sigue un modelo $ARMA$, permanece el carácter lineal de la dependencia de x_t sobre realizaciones previas. Así la relación de una nueva realización x_t sobre realizaciones previas hasta el desfase m debe seguir un modelo general, dado por ec. 5.74, equivalente a un modelo de Regresión Lineal Múltiple (RLM) tomando como variables predictoras las m realizaciones previas. En este modelo, la variable residuo ε , siguiendo la nomenclatura habitualmente utilizada en modelos RLM, es independiente de todas las variables predictoras y se distribuye de acuerdo a una distribución $N(0, \sigma_{\varepsilon_m}^2)$.

$$x_t = \varepsilon_{m,t} + I_{m,1} \cdot x_{t-1} + I_{m,2} \cdot x_{t-2} + \dots + I_{m,m} \cdot x_{t-m} \quad (5.74)$$

Así pues, conocidas las m realizaciones de x previas a x_t , y conocida la relación dada por ec. 5.74, puede ser escrita cualquier condición sobre x_t en términos de la variable transformada $\varepsilon_{m,t}$, con la que evaluar la probabilidad de ocurrencia condicional. Para la estimación de los parámetros del modelo de *RLM* en ec. 5.74, el método de máxima verosimilitud, bajo la hipótesis de distribución normal de las variables, conduce a la ec. 5.75.

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \vdots \\ \gamma_m \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{m-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{m-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \ddots & \gamma_{m-3} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \gamma_{m-1} & \gamma_{m-2} & \gamma_{m-3} & \cdots & \gamma_0 \end{bmatrix} \cdot \begin{bmatrix} I_{m,1} \\ I_{m,2} \\ I_{m,3} \\ \vdots \\ I_{m,m} \end{bmatrix} \quad (5.75)$$

donde γ_i , $i = 0, \dots, m-1$, es la autocorrelación de desfase i :

$$\gamma_i = E[(x_t - \mu_x) \cdot (x_{t-i} - \mu_x)] \quad (5.76)$$

y la varianza de la variable transformada ε_m resulta:

$$\sigma_{\varepsilon_m}^2 = \sigma_x^2 \cdot \left(1 + \sum_{i=1}^m I_i^2\right) + 2 \cdot \sum_{i=0}^{m-1} \left(I_i \cdot \sum_{j=i+1}^m I_j \cdot \gamma_{j-i}\right) \quad (5.77)$$

con $I_0 = -1$. La estimación de la función de autocorrelación a partir de estimadores muestrales produce generalmente un resultado sesgado. Por ello, una vez modelada la estructura estocástica de correlación de la variable aleatoria x mediante un modelo *ARMA*(p, q), para mantener coherencia con este modelo, se puede utilizar en ec. 5.75 la función de autocorrelación teórica que resulta derivada del modelo teórico *ARMA* ajustado. Esta función de autocorrelación, para modelos *ARMA*(p, q), para $p > q$, viene dada por (Pandit & Wu, 2001):

$$\gamma_k = d_1 \cdot \lambda_1^k + d_2 \cdot \lambda_2^k + \dots + d_p \cdot \lambda_p^k \quad (5.78)$$

donde λ_i , $i = 1, \dots, p$, son las raíces del polinomio característico de la parte autoregresiva de *ARMA*(p, q) (ec. 5.71), y d_i , $i = 1, \dots, p$, se obtiene mediante:

$$d_i = \left(\frac{g_i \cdot g_1}{1 - \lambda_i \cdot \lambda_1} + \frac{g_i \cdot g_2}{1 - \lambda_i \cdot \lambda_2} + \dots + \frac{g_i \cdot g_p}{1 - \lambda_i \cdot \lambda_p} \right) \cdot \sigma_a^2 \quad (5.79)$$

con

$$g_i = \frac{(\lambda_i^{p-1} - \theta_1 \cdot \lambda_i^{p-2} - \dots - \theta_{p-1})}{(\lambda_i - \lambda_1) \cdot (\lambda_i - \lambda_2) \cdot \dots \cdot (\lambda_i - \lambda_{i-1}) \cdot (\lambda_i - \lambda_{i+1}) \cdot \dots \cdot (\lambda_i - \lambda_p)} \quad (5.80)$$

Nótese que en los casos en los que $q < p - 1$ la evaluación de ec. 5.80 se realiza suponiendo $\theta_j = 0$, para $j = q + 1, \dots, p - 1$. Con este procedimiento, para $m = n$ la ec. 5.74 coincide con 5.73, asimilando ε_n con a_t .

La transformación obtenida, en lugar de proporcionar una variable aleatoria distribuida uniformemente, como enuncia la Transformación de Rosenblatt original (ec. 5.66), proporciona una variable normal, sobre la que fácilmente se pueden imponer las condiciones equivalentes que resultan transformación de las impuestas sobre x_t , y evaluar de ese modo las probabilidades condicionales.

5.3.2. Cálculo de probabilidades conjuntas mediante integración múltiple

A partir de la transformación de las variables $x_t \in N(0, 1)$, a variables $\varepsilon_{t,m} \in N(0, \sigma_{\varepsilon_m}^2)$, que resultan independientes de las realizaciones previas conocidas, mediante ec. 5.74, es posible la evaluación de funciones tales como las definidas en ec. 5.10, 5.12 ó 5.51 mediante la aplicación del Teorema de Multiplicación (ec. 5.65).

Sea en primer lugar la definición de la *f.e.* (ec. 5.10), la aplicación del Teorema de Multiplicación produciría:

$$\begin{aligned} \check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) &= \mathcal{P}(x_i \leq X_i) \cdot \mathcal{P}(x_{i+1} \leq X_{i+1} | x_i \leq X_i) \cdot \dots \\ &\dots \cdot \mathcal{P}(x_{i+w-1} \leq X_{i+w-1} | x_i \leq X_i, x_{i+1} \leq X_{i+1}, \dots, x_{i+w-2} \leq X_{i+w-2}) \end{aligned} \quad (5.81)$$

ecuación que resulta equivalente a 5.82, tras la transformación definida en 5.83, transformaciones del tipo de ec. 5.74.

$$\begin{aligned} \check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) &= \mathcal{P}(x_i \leq X_i) \cdot \mathcal{P}(\varepsilon_{1,i+1} \leq \Upsilon_{i+1} | x_i \leq X_i) \cdot \dots \\ &\dots \cdot \mathcal{P}(\varepsilon_{w-1,i+w-1} \leq \Upsilon_{i+w-1} | x_i \leq X_i, x_{i+1} \leq X_{i+1}, \dots, x_{i+w-2} \leq X_{i+w-2}) \end{aligned} \quad (5.82)$$

con

$$\begin{aligned} x_{i+1} &= \varepsilon_{1,i+1} + I_{1,1} \cdot x_i \\ &\dots = \dots \\ x_{i+w-1} &= \varepsilon_{w-1,i+w-1} + I_{w-1,1} \cdot x_i + I_{w-1,2} \cdot x_{i+1} + \dots + I_{w-1,w-1} \cdot x_{i+w-2} \end{aligned} \quad (5.83)$$

$$\begin{aligned}
\Upsilon_{i+1} &= X_{i+1} - I_{1,1} \cdot x_i \\
&\dots = \dots \\
\Upsilon_{i+w-1} &= X_{i+w-1} - I_{w-1,1} \cdot x_i - I_{w-1,2} \cdot x_{i+1} - \dots - I_{w-1,w-1} \cdot x_{i+w-2}
\end{aligned} \tag{5.84}$$

Ec. 5.82 tiene su desarrollo en la forma de integral múltiple, en la ec. 5.85, en donde $\psi(x, \sigma)$ es la función de densidad de probabilidad de una variable aleatoria $N(0, \sigma^2)$, ec. 5.86.

$$\begin{aligned}
\check{F}_{x,w}(X_i, X_{i+1}, \dots, X_{i+w-1}) &= \int_{-\infty}^{X_i} \psi(x_i, 1) \cdot \int_{-\infty}^{\Upsilon_{i+1}} \psi(\varepsilon_{1,i+1}, \sigma_{\varepsilon_1}) \cdot \dots \\
&\dots \int_{-\infty}^{\Upsilon_{i+w-1}} \psi(\varepsilon_{w-1,i+w-1}, \sigma_{\varepsilon_{w-1}}) \cdot d\varepsilon_{w-1,i+w-1} \cdot \dots \cdot d\varepsilon_{1,i+1} \cdot dx_i
\end{aligned} \tag{5.85}$$

$$\psi(x, \sigma) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{x}{\sigma} \right)^2 \right] \tag{5.86}$$

La ec. 5.85 puede ser resuelta de este modo de forma numérica, mediante la aplicación del método de lattice de integración múltiple, descrito en el Apéndice A.1.

Similar procedimiento podría aplicarse sobre las definiciones de las funciones *f.p.e.* y *f.a.p.e.* (ec. 5.12 y 5.51). Aunque la resolución numérica de ec. 5.85 es costosa desde el punto de vista computacional (en función de w), la capacidad de cálculo convencional de la que se dispone actualmente lo permite. Sin embargo, la resolución del conjunto de integrales múltiples correspondientes que resultarían de desarrollar ec. 5.12, y en mayor medida, ec. 5.51, inhabilitan el procedimiento para la evaluación de las funciones *f.p.e.* y *f.a.p.e.* En este punto la metodología explicada encuentra una limitación para su implementación de forma rigurosa en análisis de variables aleatorias con dependencia temporal importante. De forma práctica, la dependencia temporal de las variables hidrológicas depende de la escala temporal de trabajo, así para escala mayores (i.e, mayores pasos temporales) la dependencia se reduce. Desde este punto de vista, la limitación de aplicación del procedimiento se encontraría en la escala temporal de trabajo mínima que puede ser utilizada en el análisis. Cuando la dependencia temporal sea débil, puede aplicarse el procedimiento de forma aproximada evaluando la *f.e.* a partir de resolver ec. 5.85, mientras que para evaluar las *f.p.e.* y *f.a.p.e.* se aplicarían las funciones analíticas ec. 5.27 y ec. 5.51. El límite práctico de la menor escala temporal a utilizar con este procedimiento se puede

localizar partiendo de una escala en la que no exista evidencia estadística de dependencia temporal, y comparando los resultados en los eventos de desviaciones persistentes extremas más acusados, a medida que la escala es reducida. Esto se ilustra en el punto 6.2.3, al analizar la sensibilidad del Índice de Frecuencias de Sequías con la escala temporal.

5.4. Tratamiento de señales estacionales

Variabes de tipo hidro-climático, definidas a escalas temporales menores de la escala anual suelen contener características de estacionalidad, es decir, su valor o su distribución de probabilidad depende de la estación en la que se produzca. Esto se manifiesta en una señal con un carácter periódico, cuyo período depende del ciclo al que responda la estacionalidad. Como ejemplo, alturas de precipitación o aportaciones acumuladas mensuales, mantienen variaciones de sus distribuciones estadísticas periódicas, de ciclo anual. En otros casos, como la temperatura de la superficie terrestre, además del ciclo anual, se observa superpuesto un ciclo diario.

Este tipo de señales requiere de una transformación para la aplicación de la metodología recogida en este capítulo, pues la hipótesis de suponer la señal estacionaria no es cumplida. La estacionalidad puede manifestarse en la variación periódica de determinados estadísticos, como puede ser la estacionalidad en la media, o en la varianza. En el caso de la estacionalidad de la media, ésta puede eliminarse mediante la sustracción en cada período del valor medio, transformado la variable original en una variable de media cero (ec. 5.87).

$$x'_{i,\tau} = x_{i,\tau} - \mu_\tau \quad (5.87)$$

donde τ representa la estación dentro del ciclo de período ω , $\tau = 1, \dots, \omega$; $x_{i,\tau}$ indica la variable x observada en el ciclo i , durante la estación τ ; $x'_{i,\tau}$ es la variable transformada; y μ_τ representa el valor de la media de x cuando la estación coincide con τ . Así mismo, cuando adicionalmente se observa estacionalidad en la varianza de las variables, ésta se puede sustraer mediante la transformación dada por ec. 5.88.

$$x''_{i,\tau} = \frac{x_{i,\tau} - \mu_\tau}{\sigma_\tau} \quad (5.88)$$

donde σ_τ representa el valor de la desviación típica de x cuando la estación coincide con τ . La estimación de μ_τ y σ_τ puede realizarse a partir de estimadores muestrales, debiendo mostrar generalmente variaciones suaves de estación a estación. Para reducir el número de parámetros, y aumentar la fiabilidad de la estimación, pueden ser ajustadas funciones periódicas a μ_τ

y σ_τ . Un procedimiento habitual es el uso de ajustes tipo desarrollo en serie de Fourier, en el que la serie de Fourier que pasa por todos los puntos, es truncada aproximando la función con un menor número de parámetros.

La transformación dada por ec. 5.87, o ec. 5.88, es suficiente cuando la variable sigue distribuciones normales. En los casos que la variable siga una distribución diferente, los parámetros que resultan del ajuste de la distribución a cada estación τ pueden producir también una variación suave periódica, siendo conveniente entonces parametrizarlos mediante el ajuste de series de Fourier truncadas.

Cuando además de manifestarse una estacionalidad en los momentos estadísticos, se manifiesta estacionalidad en la autocorrelación, es necesario el análisis de su modelación con modelos estocásticos estacionales. Los modelos estocásticos estacionales son aquellos en los que sus parámetros que representan las características de autocorrelación presentan una variación periódica. Un ejemplo se encuentra en el modelo Periódico AutoRegresivo de orden p , $PAR(p)$, (Salas *et al.*, 1980) que sigue la siguiente forma:

$$x_{t,\tau} = \mu_\tau + \sum_{j=1}^p \phi_{j,\tau} \cdot (x_{t,\tau-j} - \mu_{\tau-j}) + \varepsilon_{t,\tau} \quad (5.89)$$

donde $x_{t,\tau}$ es una variable aleatoria independiente que sigue una distribución $N(\mu_\tau, \sigma_{x_\tau}^2)$, y $\varepsilon_{t,\tau}$ es una variable aleatoria independiente que sigue una distribución $N(0, \sigma_{\varepsilon_\tau}^2)$. Si $\tau - j \leq 0$, entonces $x_{t,\tau-j}$ y $\mu_{\tau-j}$ representan $x_{t-1, \omega+\tau-j}$ y $\mu_{\omega+\tau-j}$, respectivamente. Así mismo se pueden incluir los términos correspondientes al modelo de media móvil, formulando los modelos Periódicos AutoRegresivos de Media Móvil, $PARMA(p, q)$ (Salas *et al.*, 1980):

$$x_{t,\tau} = \mu_\tau + \sum_{j=1}^p \phi_{j,\tau} \cdot (x_{t,\tau-j} - \mu_{\tau-j}) + \varepsilon_{t,\tau} - \sum_{j=1}^q \theta_{j,\tau} \cdot \varepsilon_{t,\tau-j} \quad (5.90)$$

donde además de las consideraciones anteriores, si $\tau - j \leq 0$, entonces $\varepsilon_{t,\tau-j}$ representa $\varepsilon_{t-1, \omega+\tau-j}$.

Este tipo de modelos estocásticos periódicos son utilizados con pequeños órdenes de p y q , para los que existen expresiones derivadas que permiten la estimación de sus parámetros. Para órdenes de p y q superiores a 1, el cumplimiento del Principio de Parsimonia puede no conseguirse por el elevado número de parámetros.

Desde el punto de vista del tratamiento de señales estacionales cuando se desea aplicar procedimientos de análisis como el expuesto anteriormente, en los casos en los que no existe periodicidad en la estructura estocástica de la señal, las transformaciones de la señal mediante ec. 5.87 o ec. 5.88 puede

ser adecuado. En los casos en los que no siga una distribución normal, la parametrización de la curva que define los parámetros de la distribución en función de la estación, permite modelar la estacionalidad. Sin embargo, en los casos en los que se manifieste una estacionalidad en la estructura estocástica de la señal, sería necesario la aplicación de modelos tipo *PAR* o *PARMA*, incluyendo las consideraciones realizadas en 5.3.1. Ello puede limitar en la práctica el procedimiento para los casos en que la señal siga un modelo estacional estocástico de orden superior a 1, en donde el número de parámetros que definirían transformaciones como la dada en ec. 5.74 conjunto con la complejidad en el cálculo de las integrales múltiples derivadas, reducen la conveniencia de su aplicación.

Capítulo 6

El Índice de Frecuencias de Sequías

En el Capítulo 5 se ha expuesto el criterio y la metodología para la caracterización de secuencias con desviaciones persistentes hacia una región extremal. La caracterización expuesta es de naturaleza probabilista. Además, se analizó su relación con las propiedades estocásticas, de tal modo que puede ser expresada dicha caracterización, de modo aproximado, en términos del período de retorno medio entre ocurrencias, constituyendo por tanto una cuantificación de tipo estocástico. En este Capítulo 6 se utiliza la metodología expuesta en el Capítulo 5 para establecer una sistemática de aplicación de la misma, dando como resultado el Índice de Frecuencias de Sequías. Este índice aplica la caracterización de secuencias respecto de su desviación persistente hacia una región extremal, para controlar el carácter excepcional de una desviación producida sobre una determinada variable hidro-climática de interés, usada como indicador, y expresándolo en términos de la recurrencia media.

6.1. Definición del Índice de Frecuencias de Sequías

El Índice de Frecuencias de Sequías (IFS) –o en lengua inglesa: *Drought Frequency Index (DFI)*– se define para una determinada variable, usada como indicador, y en un determinado instante como el valor resultante de ec. 5.64 para el mayor valor de la *f.p.e.* obtenido de las secuencias que terminan en ese instante. De forma más concreta, en cada instante t se analizan cada una de las secuencias de realizaciones de la variable aleatoria x que terminan en dicho instante, $X_{t-w+1}, X_{t-w+2}, \dots, X_t$, con $w \geq 0$. Para

cada una de ellas se evalúa la *f.p.e.*, buscando aquella secuencia pasada que resulte más extrema, cuya longitud se asignará a W . La forma de expresar el máximo grado de desviación extremal alcanzado por la variable hasta dicho instante, se realiza mediante la estimación de la frecuencia media de recurrencia de desviaciones de dicho grado, dada por ec. 5.64. Este es un valor aproximado de estimación del período de retorno, cuya convergencia se puede situar para períodos de $T \geq 100$ realizaciones. La figura 6.1 muestra de forma gráfica el diagrama de flujo seguido para el cálculo del IFS en cada instante.

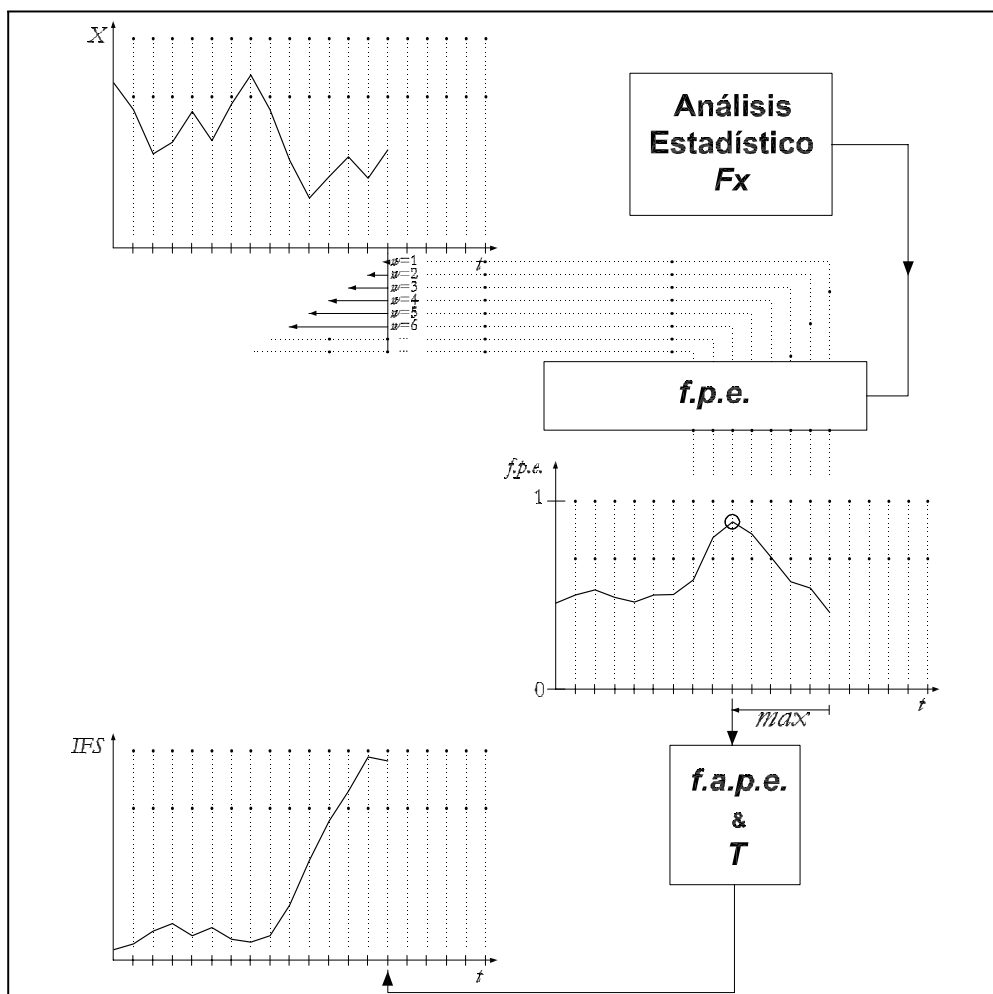


Figura 6.1: Diagrama de flujo para la cálculo del IFS en un instante t

De acuerdo a su definición, el IFS es un índice retrospectivo, que caracteriza cada instante con aquella secuencia de realizaciones producidas hasta el momento que resulte con la mayor desviación persistente extrema, evaluada con la *f.p.e.* Ello responde al hecho de que las sequías no son procesos aleatorios puntuales en el tiempo, sino que tienen un desarrollo temporal, en el que pueden aparecer diferentes aspectos que la adjetivicen como evento extremo (persistencia, severidad, intensidad máxima, intensidad media, factor de uniformidad, etc.). La *f.p.e.* es usada en el IFS para englobar todos estos aspectos particulares, en un criterio de carácter general de cuantificación con naturaleza probabilista. En lugar de tratar cada uno de los aspectos particulares que pueden calificar a un evento de extremo, se realiza un análisis directo de la secuencia de realizaciones producidas, desde el punto de vista de su aproximación a la región extremal, y se evalúa de manera conjunta con la *f.p.e.* Debe ser esperado que todos los eventos que resulten extremos bajo uno o varios de los aspectos particulares, tengan su reflejo en valores significativamente altos en la *f.p.e.* Dado su carácter general, no se realiza un escalado de la importancia de cada aspecto particular en la valoración global, sino que se trata de establecer un criterio que responda a las ideas comunes que aparecen en las definiciones generales de sequía.

La figura 6.2 muestra el resultado del análisis mediante el IFS sobre la serie registrada de 1912 a 2002 en el río Ebro, a su paso por Zaragoza. El IFS no sólo caracteriza la situación tras cada realización, sino que puede ser usado para caracterizar eventos. Así, en el caso del evento de sequía producido entorno a 1990, se observa un ascenso del índice a medida que el evento va teniendo lugar, hasta llegar a un punto de persistencia máxima, a partir del cual la ocurrencia de realizaciones más alejadas de la región extremal produce una recuperación del índice hacia valores más frecuentes, y por tanto situaciones no extremas. El índice no sólo está señalando la realización de valores próximos a la región extremal, como puede ser el caso del evento ocurrido en 1934, sino que también tiene en cuenta la proximidad o la persistencia de estos valores más alejados del régimen normal de variación de la variable aleatoria. El ejemplo permite, así mismo, ilustrar el comportamiento del índice en análisis de series con manifiesta estacionalidad, distinguiendo la desviación que se produce en cada momento del año de acuerdo a su distribución poblacional. Adicionalmente, esta serie presenta una significativa correlación, la cual es tenida en cuenta en el cómputo del índice para cuantificar la desviación máxima producida en cada instante. En la formulación del IFS no se incluye ningún criterio que pueda tener un grado de arbitrariedad para un análisis general. Así por ejemplo, no existe la definición de un umbral; por lo cual, el IFS no determina de forma explícita lo que considera el principio y final de un evento, sino que estos límites,

cuando sean necesarios, deberían ser adoptados en una aplicación práctica particular para un determinado nivel del índice que mejor se acomode a objetivos particulares.

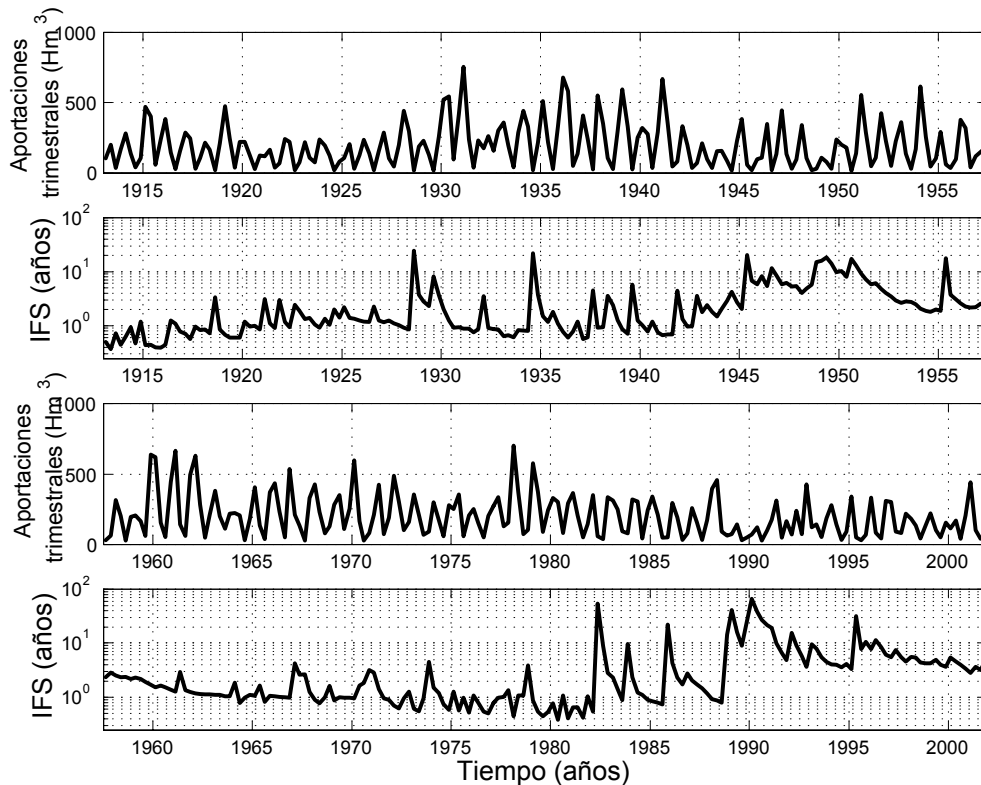


Figura 6.2: Aportaciones trimestrales en el Río Ebro a su paso por Zaragoza y análisis de sequías mediante el IFS.

La evaluación del IFS, de acuerdo a la figura 6.1, permite la caracterización estocástica de realizaciones de eventos extremos de sequías en el tiempo, manifestada sobre una determinada variable aleatoria usada como indicador. Esta variable generalmente será de tipo hidrológico y puede representar un proceso considerado puntual en el espacio, o englobar bajo algún criterio de agregación, la característica deseada de una realización espacial en la región de estudio. Ejemplo puede ser la altura de precipitación media espacial recogida en una determinada cuenca, o las aportaciones en un punto de la red hidrográfica de una cuenca. En el Capítulo 7 se amplía la metodología para la construcción de mapas que permitan la caracterización estocástica

espacio-temporal de realizaciones de eventos hidrológicos extremos, caracterizando bajo determinados criterios la realización distribuida no sólo en el tiempo, sino también en el espacio.

La siguiente sección de este capítulo analiza en mayor detalle las propiedades de comportamiento práctico del IFS, presentando sus resultados ante circunstancias que generalmente suponen una debilidad en la aplicación de otros índices o procedimientos de caracterización.

6.2. Análisis comparativo con otros índices y procedimientos

Como ya ha sido puesto de manifiesto, en la literatura pueden encontrarse un gran número de acepciones de sequía y de criterios e índices para su análisis. Sin embargo, ninguno de ellos se encuentra universalmente aceptado, y es más, posiblemente no pueda diseñarse un indicador universal, dada la multitud de definiciones que existen de sequías, ajustada cada una a un problema práctico. No obstante, sí se encuentran un conjunto de indicadores, que con objetivos generales de caracterización, son utilizados para el análisis de realizaciones de eventos extremos de sequías. Entre los indicadores puede destacarse el Índice de Severidad de Sequías de Palmer (*Palmer Drought Severity Index, PDSI*), o el Índice de Precipitación Estandarizado (*Standardized Precipitation Index, SPI*). Por otro lado, como ya ha sido introducido en el Capítulo 3, entre las metodologías más extendidas de análisis de la recurrencia estocástica de los eventos de sequía destaca el uso de la Teoría de Rachas (TR), la cual, mediante la construcción de un modelo de recurrencia, permite estimar propiedades tales como el período de retorno de eventos de sequía. A pesar de que estas metodologías están extensamente utilizadas a lo largo de la literatura, mantienen una serie de deficiencias que hacen poco recomendable su aplicación en determinadas circunstancias o bajo ciertos objetivos. A continuación son presentadas las principales deficiencias de estos métodos, comparándolas con el resultado de aplicar el IFS en diversos casos prácticos.

6.2.1. Principales deficiencias metodológicas en la caracterización de sequías

Tratando en primer lugar los métodos que permite realizar un análisis estocástico de recurrencia, la primera deficiencia que se encuentra en los métodos implementados en la literatura científica es la limitación en el número de magnitudes a tratar en el análisis de forma simultánea. Generalmente, las series histórica con las que se cuenta contienen un número

de eventos de sequía reducido que no permite la formulación de modelos de recurrencia con un alto número de parámetros a ajustar con criterios estadísticos. Ello provoca que ante el objetivo de realizar un análisis de sequías general, donde más de una magnitud pueda ser de importancia, no sea posible pasar de la selección de una única magnitud para la construcción del modelo. Esta magnitud suele ser la duración, pues es para la que con menor número de parámetros suele ajustarse un modelo. Aunque las diferentes magnitudes que pueden ser utilizadas en el análisis (duración, severidad, intensidad media, intensidad máxima, factor de uniformidad,...) suelen presentar una significativa correlación, el uso de diferentes magnitudes en la construcción de modelos de recurrencia arroja resultados que pueden ser muy diferentes (Frick *et al.*, 1990; González & Valdés, 2003).

Otro aspecto que debe ser fijado cuando se pretende utilizar la TR en el análisis de recurrencia de sequías es la selección del umbral que define los períodos considerados secos. Existen casos en los que su fijación es clara, por ejemplo en el análisis del suministro a una demanda a partir de una derivación en azud, sin capacidad de almacenamiento. Es tal situación, el umbral coincide con la demanda estudiada, ya sea ésta constante o variable. En otros caso donde son analizados los eventos a partir del *PDSI*, el umbral se encuentra en el $PDSI = -1$, donde el índice define el comienzo de sequía. Sin embargo en otros muchos casos no existe un condicionante que la defina, y en ellos es común la adopción de determinados cuantiles como umbrales en el análisis (e.g. el cuantil del 25 %, donde existe una probabilidad de 0.75 de ser excedido). Cuando el umbral no viene condicionado, el análisis puede mostrar importante sensibilidad a variaciones en su posición, lo que hace que con objetivos generales el análisis no sea el más adecuado.

En cuanto a los indicadores o índices de sequía, el *PDSI* puede considerarse como el más popularmente utilizado. El índice expresa de forma adimensional la severidad de la desviación de la situación normal del volumen disponible de agua en una determinada región, partiendo de un modelo conceptual de balance hídrico regional simplificado. El índice caracteriza las sequías meteorológicas, aunque se han presentado variantes para el análisis de sequías hidrológicas. Dado su uso, también se encuentra en la literatura un significativo número de trabajos que tratan de evaluar su aptitudes como índice general (Karl, 1983; Alley, 1984; Guttman, 1991; Heddinghaus & Sabol, 1991). Las principales deficiencias encontradas en el índice provienen de su algoritmo de cálculo, donde son introducidos un número de coeficientes cuyo valor fue seleccionado de forma directa buscando la mejor adecuación del índice a la región de estudio donde quería ser aplicado, el sur de EEUU (Palmer, 1965). A pesar de este grado de arbitrariedad en la definición del índice, creado inicialmente para su aplicación en una región particular de clima fundamentalmente árido y semiárido, el *PDSI* ha sido y está siendo

usado de forma extensiva a lo largo del globo incluso en regiones de clima húmedo. Posiblemente la causa que motiva su extendido uso es su popularidad, sin embargo el grado de arbitrariedad en su definición provocan que la comparación de resultados del índice entre regiones con diferentes regímenes climáticos no sea siempre adecuada (Guttman *et al.*, 1992).

Es precisamente esta deficiencia en la comparabilidad espacial del *PDSI* lo que motiva la recomendación de Guttman (1992) del tratamiento de las sequías en términos probabilísticos. De ese modo, la diferencia de regímenes climáticos no resultaría una limitación ante la necesidad de comparar el carácter extraordinario de sequías que afectan diferentes regiones. En este sentido se encuentra el *SPI* (McKee *et al.*, 1993), que caracteriza la realización de una sequía meteorológica en términos de la distribución estadística de precipitaciones agregadas en una determinada escala temporal (período de agregación, e.g. mes, trimestre, año,...), expresado como el valor de la variable normal estándar equivalente. Es precisamente con esta transformación a una variable normal estándar equivalente como se posibilita la comparación de resultados entre poblaciones distintas, esto es, en diferentes regiones. La formulación del índice indica la utilidad de repetir el análisis para diferentes períodos de agregación, utilizando así períodos de 1, 3, y 6 meses, y 1, 2, y 4 años. Ello permite la caracterización de eventos correspondientes a realizaciones de diferente duración, como pueden ser una sequía agrícola de pocos meses o una sequía severa de varios años. Sin embargo, es precisamente este análisis separado de los eventos producidos en las diferentes duraciones lo que provoca una deficiencia en el método. Operando de este modo, no es posible la comparación de eventos que resultan cada uno significativos bajo diferentes períodos de agregación o escalas temporales. El método no posibilita su comparación, salvo la comparación directa de los valores del *SPI* obtenidos. Pero, cuando estos valores del índice, aunque iguales, provienen de diferentes períodos de agregación, la frecuencia con la que se realizan son distintas.

Seguidamente son ilustrados cada una de estas deficiencias con ejemplos prácticos. Así mismo, son comparados los resultados de aplicar el IFS sobre esos casos, donde otros métodos resultan inadecuados.

6.2.2. Análisis de sensibilidad en el uso de la Teoría de Rachas

Caracterización Uni-, Bi-, y Trivariada de las sequías en las aportaciones anuales del río Blue (EEUU)

El río Blue se encuentra en la cabecera del río Colorado (EEUU), en el estado del mismo nombre. Sobre este río se cuenta con los datos de aporta-

ciones anuales medidos de forma instrumental en el período 1916 a 2002, que se muestra en la figura 6.3. Las aportaciones siguen una distribución normal, con una media de $280Hm^3$ y una desviación típica de $74Hm^3$. El río Blue es utilizado aquí para la aplicación de la TR en la caracterización estocástica de sequías.

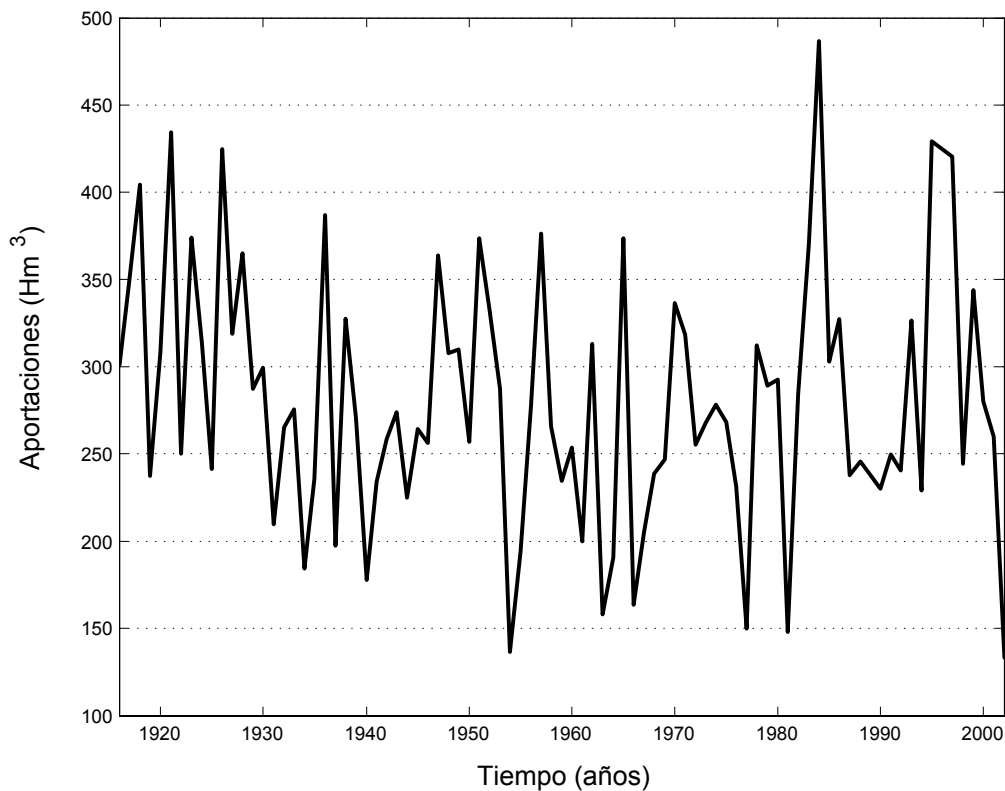


Figura 6.3: Aportaciones anuales en el río Blue (Colorado, EEUU).

Los 87 años de datos de aportaciones anuales contienen un número de eventos de sequías, que siguiendo la TR, depende de la selección del umbral. Así si es utilizado un umbral coincidente con la media, o en este caso el cuantil $Q_{50\%}$, el número de eventos de sequía disponibles para el análisis es de 16 eventos. Este limitado número de eventos permite el ajuste de modelos de recurrencia simples, como pueden resultar aquí el que correspondiente a la selección de la duración como magnitud de análisis. Para posibilitar la aplicación de modelos de recurrencia bivariados y trivariados se puede hacer uso de la información disponible en reconstrucciones paleo-

climáticas, que prolonga hacia el pasado las series instrumentales a partir de otro tipo de registros de variables relacionadas con la variable a reconstruir, de las cuales se encuentran en la naturaleza series más largas. Entre las variables usadas en las reconstrucciones paleoclimáticas destacan las series dendrocronológicas de anchos de anillos de los que se componen los troncos de árboles, cuyos espesores están relacionados con la disponibilidad de humedad por la planta durante el período anual de crecimiento. El análisis de las relaciones entre los anchos de anillos y variables a reconstruir, como puede ser las aportaciones anuales de un río, permite la extensión de la serie instrumental en el pasado.

En el caso del río Blue, se cuenta con una reconstrucción de la serie de aportaciones anuales hasta el año 1383, obtenida a partir de series dendrocronológicas por Woodhouse et al. (2004). Uno de los aspectos que caracteriza estas reconstrucciones cuando se comparan con la serie instrumental es su variación más suave. La reconstrucción generalmente atiende a las oscilaciones de menor frecuencia de la serie, reproduciendo sólo un porcentaje de la varianza. En el caso del río Blue, la reconstrucción reproduce el 75 % de la varianza de la serie instrumental, quedando un 25 % sin reconstruir, no captado sobre la variable predictora. Sin embargo, las variaciones que corresponden al porcentaje no reconstruido, responden a variaciones de corta frecuencia, no sensibles en las series dendrocronológicas, pero que generalmente quedan bien recogidas en la serie instrumental de la variable estudiada. González y Valdés (2003) presentaron una metodología para introducir la información disponible en las series reconstruidas en el ajuste de modelos de recurrencia basados en la TR. El procedimiento se basa en el análisis de los errores de la reconstrucción durante el período de calibración, donde se solapan la serie instrumental y las series dendrocronológicas, el cual sigue generalmente un modelo estocástico. A partir de este modelo, es posible el completado de la serie reconstruida, mediante la generación de series de errores que se añaden a la reconstrucción. La generación de múltiples series de errores, y de aquí, múltiples reconstrucciones posibles con una varianza completa, permite analizar los eventos de sequías que pudieron haber ocurrido en el pasado, y añadirlos a los ya disponibles en la serie instrumental. Siguiendo esta metodología se ajustó un modelo ARMA(3,2) a los residuos a partir del cual fueron generadas un total de 1,000 posibles series de aportaciones que pudieron ser realizadas en el pasado y producir las mismas dendrocronologías.

Fijando el umbral en el cuantil del 50%, $Q_{50\%}$, es posible para cada generación, al igual que para la serie instrumental, la delimitación de los períodos de sequía (\mathcal{S}) y los períodos húmedos (\mathcal{H}). Del conjunto de los períodos húmedos lo que interesa es la distribución de su duración, $f_{d\mathcal{H}}(d)$, que para este caso se comprueba que sigue una distribución binomial (ec.

6.1).

$$f_{d\mathcal{H}}(d) = P_{\mathcal{H}\mathcal{S}} \cdot (1 - P_{\mathcal{H}\mathcal{S}})^{d-1} \quad (6.1)$$

Para los períodos de sequía, son analizados hasta tres magnitudes: duración, severidad y factor de uniformidad. Del análisis de los mismos se encuentra que las correlaciones entre las duraciones de los períodos secos y los siguientes períodos húmedos no son estadísticamente significativas, con el 5 % de nivel de significancia. Así mismo, tampoco que encuentran correlacionadas las duraciones de eventos secos sucesivos o de eventos húmedos. Con ello es posible la aplicación de un modelo de recurrencia multivariado como el descrito de en Capítulo 3. Para ello, al igual que ocurre con la distribución de duraciones de períodos húmedos, la función de distribución de duraciones de períodos secos sigue también una distribución binomial (ec. 6.2). Esto da lugar a que la aparición y duración de los eventos de sequía pueda modelarse como un modelo de Markov simple de dos estados (Norris, 1997).

$$f_{d\mathcal{S}}(d) = P_{\mathcal{S}\mathcal{H}} \cdot (1 - P_{\mathcal{S}\mathcal{H}})^{d-1} \quad (6.2)$$

Así mismo, es necesaria la construcción de la función de distribución conjunta de las magnitudes que pueden caracterizar un evento de sequía. Se encuentra que para cada duración, la distribución marginal conjunta de las magnitudes severidad y factor de uniformidad sigue una distribución teórica bigamma. La expresión teórica de la función de distribución bivariada seleccionada es la propuesta por Moran (1969):

$$f(x, y) = \frac{1}{\sqrt{(1 - \rho^2)}} \cdot f_X(x; \alpha_X, \lambda_X) \cdot f_Y(x; \alpha_Y, \lambda_Y) \cdot \exp \left[-\frac{(\rho \cdot x')^2 - 2 \cdot \rho \cdot x' \cdot y' + (\rho \cdot y')^2}{2 \cdot (1 - \rho^2)} \right] \quad (6.3)$$

$$x' = \Phi^{-1} [F_X(x, \alpha_X, \lambda_X)]$$

$$y' = \Phi^{-1} [F_Y(x, \alpha_Y, \lambda_Y)]$$

donde f_X y f_Y representan las funciones *f.d.p.* de la función de distribución gamma de las variables X e Y , F_X y F_Y representan las funciones *f.p.a.* de las mismas variables, $\alpha_X, \lambda_X, \alpha_Y, \lambda_Y$ representan los parámetros de la distribución:

$$f_Z(x; \alpha_Z, \lambda_Z) = \frac{1}{\Gamma(\lambda_Z)} \cdot \alpha_Z^{\lambda_Z} \cdot z^{\lambda_Z-1} \cdot \exp(-z \cdot \alpha_Z) \quad (6.4)$$

Φ^{-1} simboliza la inversa de la función de distribución normal estándar, y ρ representa el coeficiente de correlación de las variables normalmente distribuidas x' e y' . De acuerdo a esta función, las distribuciones marginales y condicionales de cada una de las variables resultarían funciones gamma. El ajuste de los parámetros para cada duración y su intervalo de confianza del 95 % puede observarse en la figura 6.4. A medida que la duración aumenta, el tamaño de la muestra disponible se reduce y el intervalo de confianza de la estimación crece. Se comprueba que el parámetro de la distribución marginal de severidades α_S se puede considerar constante, y los parámetros $\lambda_Y, \alpha_U, \lambda_U$ linealmente dependientes de la duración. El coeficiente de correlación también aparece como constante e independiente de la duración.

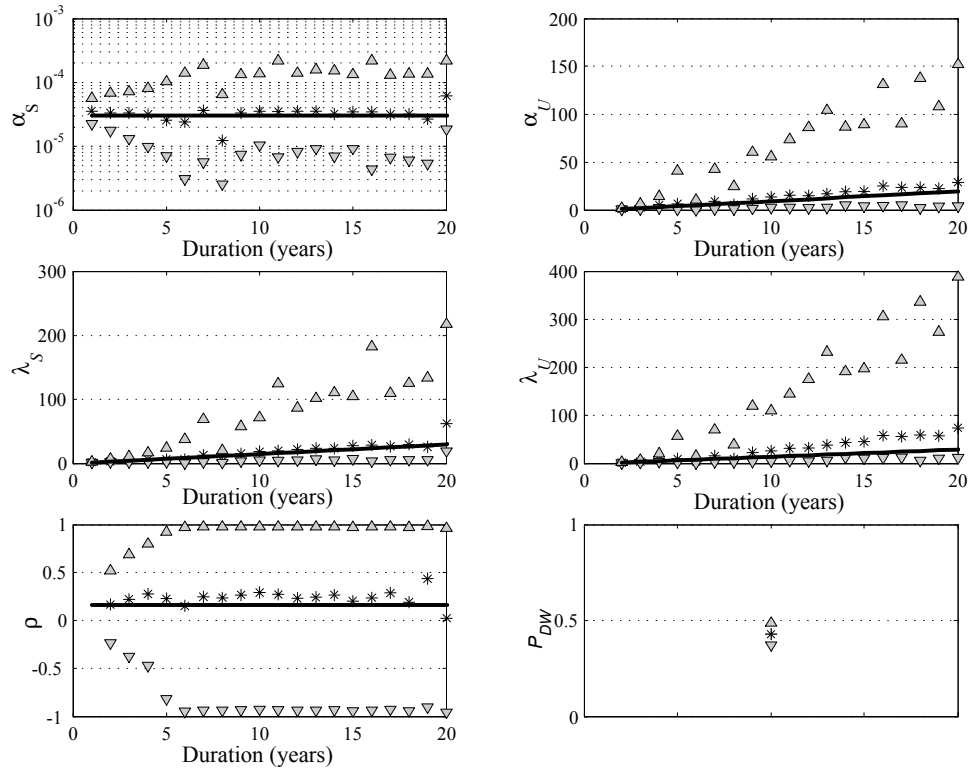


Figura 6.4: Estimación e intervalo de confianza del 95 % de los parámetros de las distribuciones marginales de severidad y factor de uniformidad fijada la duración, conjunto con el parámetro de la distribución de duraciones de sequías, y la parametrización final completa (río Blue, con umbral Q50 %).

Siguiendo esta parametrización, y ajustando sus parámetros ($P_{SH}, \alpha_S, b_S,$

, a_U, b_U, ρ) por el método de máxima verosimilitud se tiene la función de distribución conjunta de las magnitudes de sequía en la siguiente forma:

$$f_{D,S,U}(d, s, u) = f_{dS}(d) \cdot \frac{1}{\sqrt{(1-\rho^2)}} \cdot f_S(s; \alpha_S, \lambda_S) \cdot f_U(u; \alpha_U, \lambda_U) \cdot \exp \left[-\frac{(\rho \cdot s')^2 - 2 \cdot \rho \cdot s' \cdot u' + (\rho \cdot u')^2}{2 \cdot (1-\rho^2)} \right] \quad (6.5)$$

$$\lambda_S = d \cdot b_S \quad (6.6)$$

$$\alpha_U = (d-1) \cdot a_U \quad (6.7)$$

$$\lambda_U = (d-1) \cdot b_U \quad (6.8)$$

A partir de esta distribución se comprueba el ajuste y se completan los elementos que componen el modelo de recurrencia multivariado descrito en el Capítulo 3. Los valores de los parámetros obtenidos resultan: $P_{\mathcal{H}S} = 0,46$, $P_{\mathcal{S}\mathcal{H}} = 0,43$, $\alpha_S = 3,04e - 5$, $b_S = 1,50$, $a_U = 1,03$, $b_U = 1,51$, $\rho = 0,17$. Obteniendo las distribuciones marginales de las magnitudes deseadas es posible la caracterización de la recurrencia de eventos de sequía en los casos uni-, bi- y trivariados.

Sensibilidad en la selección de la/s magnitud/es crítica/s

De acuerdo con el modelo de recurrencia multivariado construido (Capítulo 3), cuyas hipótesis fueron comprobadas en el apartado anterior, la función de distribución de las duraciones de períodos húmedos y la función de distribución conjunta de las magnitudes duración, severidad y coeficiente de uniformidad de los períodos secos, anteriormente ajustada, es posible la caracterización de cada uno de los eventos de sequía ocurridos en la serie instrumental. El análisis puede hacerse de forma multivariada, ya sea uni-, bi-, o trivariada, según el número de magnitudes seleccionadas. En cada caso, la función de distribución de las magnitudes es obtenida a partir de la función de distribución conjunta completa. Considerando la situación de sequía para el umbral posicionado en el cuantil medio ($Q_{50\%}$), los resultados de análisis de recurrencia, para las diferentes combinaciones posible de magnitudes críticas seleccionadas, se expresan en en la tabla 6.1.

La correlación existente entre las magnitudes analizadas es significativa en todos los casos ($\rho_{DS} = 0,86$, $\rho_{DU} = 0,42$, $\rho_{SU} = 0,76$). Sin embargo, los resultados pueden discrepar en función de la magnitud o el conjunto de magnitudes seleccionadas. Así destaca el caso del evento ocurrido en el período

Evento	T_D	T_S	T_U	T_{DS}	T_{DU}	T_{SU}	T_{DSU}
1919	5	6	5	6	5	6	6
1922	5	6	5	6	5	5	5
1925	5	6	5	6	5	6	6
1931/35	42	25	15	51	96	61	121
1937	5	8	5	8	5	8	8
1939/46	226	38	15	234	571	96	598
1950	5	5	5	5	5	5	5
1954/56	14	25	29	28	73	185	225
1958/61	24	15	17	28	65	43	80
1962/63	8	21	11	22	11	30	30
1966/1969	24	33	14	43	48	73	92
1972/1977	74	26	35	80	1480	260	1767
1981	5	11	5	11	5	11	11
1987/1992	74	27	8	81	74	29	82
1994	5	6	5	6	5	6	6
1998	5	6	5	6	5	6	6

Tabla 6.1: Período de retorno de eventos de sequía (umbral cuantil medio) sobre las aportaciones anuales del río Blue, en el período 1916-2002.

1939/46, de una duración de 8 años, lo que le confiere un período de retorno del orden de los 200 años. Sin embargo, este evento analizado de acuerdo a su severidad o factor de uniformidad se encuentra en un período de retorno de un orden de magnitud inferior. Como un caso diferente, puede destacarse el evento ocurrido entre 1954/56, de una duración de 3 años que supone un período de retorno de 14 años, pero en el que el uso de su severidad y factor de uniformidad como magnitudes críticas le asigna una recurrencia media de 185 años. Por supuesto, a medida que son considerados un mayor número de magnitudes en el análisis, el período de retorno resulta igual o superior, pues se añade una condición adicional a cumplirse en la formulación probabilista. En el caso del evento de 1987/1992, la consideración de las magnitudes duración y severidad caracteriza al evento con un período de retorno de 81 años. La consideración adicional del factor de uniformidad no modifica sensiblemente su caracterización. En este sentido, aunque diversas magnitudes pueden expresar el carácter extraordinario de los eventos de sequías, es esperable que dada la correlación que entre las magnitudes existe, a medida que el número de magnitudes consideradas aumente el período de retorno asociado se mantenga de forma asintótica. Sin embargo, el problema fundamental aparece cuando se pretende realizar un análisis multivariado se encuentra en la longitud de la serie de datos disponible, que debe contener un número suficiente de eventos de sequía para realizar el ajuste estadístico con suficiente significación.

Sensibilidad en la selección del umbral

Los resultados obtenidos en el análisis anterior, a partir del modelo de recurrencia construido de acuerdo a la TR, vienen dados a partir de la fijación del umbral que define la sequía en el cuantil medio ($Q_{50\%}$). Sin embargo, en un análisis general de sequías este umbral no está prefijado por ningún condicionante externo. No sería el mismo caso si el objetivo fuera realizar el análisis de los déficit de suministros a una demanda, en un sistema no regulado y con una demanda conocida. En este caso, el umbral sería igual a la demanda, ya fuera ésta constante o variable.

Para estudiar la sensibilidad en la posición del umbral de definición del estado de sequía en un estudio de caracterización aplicando la TR se ha repetido el análisis expuesto anteriormente para los umbrales posicionados en los cuantiles de 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 % y 50 %. Así, la figura 6.5 presenta en una escala de colores los resultados que se obtendrían sobre la serie instrumental de aportaciones anuales en el río Blue (EEUU) utilizando un análisis multivariado, adoptando como magnitudes críticas la duración, severidad y factor de uniformidad.

Como puede apreciarse en la figura 6.5 el rango de variación del período de retorno resultante de la caracterización, en función del umbral seleccionado, puede variar hasta en tres ordenes de magnitud. Incluso se encuentran eventos en los que variaciones del umbral en un porcentaje del $\pm 5\%$ implican cambio en un orden de magnitud en los resultados (e.g. evento de sequía de 1972/77). También puede resultar de significativa importancia el hecho de que ante un descenso en el umbral, un evento definido durante un período de tiempo se vea convertido en dos eventos de diferente caracterización. Así, en un evento compuesto por varios mínimos muy próximo, puede resultar muy difícil la caracterización general de acuerdo con la TR.

Comparación con el uso del IFS

Para comparar los resultados obtenidos mediante la caracterización de sequías a partir del período de retorno, obtenida mediante un modelo de recurrencia estocástico, la figura 6.6 muestra los valores obtenidos aplicando el IFS a la serie de aportaciones anuales en el río Blue (EEUU). En él destacan cuatro eventos, los ocurridos en 1972/77 y 1981, con valor del índice entorno a los 70 años, y los ocurridos en 1954/56 y 1962/63 del orden de los 30 años.

El análisis con el IFS muestra que el evento de 1972/77 puede resultar el que mayor desviación persistente produce. Sin embargo, la realización próxima de una nueva desviación (la producida en 1981) vuelve a producir

en el índice similares valores ya que por medio no ha habido un período de clara recuperación. Por otro lado, el evento de 1954/56 aunque alcanza el valor mínimo en la serie, la persistencia de valores bajos, próximos a la región extremal, resulta menor comparada con el evento de 1972/77, lo que le confiere un valor del índice menor que el de éste. En lo que se refiere a la comparación entre los eventos producidos en 1954/56 y 1962/63 resultan conclusiones similares a las obtenidas entre 1972/77 y 1981. Las condiciones previas a la realización del evento de 1954/56 resultan de menor desviación que las ocurridas cuando se produce el evento de 1962/63, lo que produce un valor del índice similar.

Analizando las diferencias con la aplicación de la TR, en su análisis multivariado, la principal apreciación resulta sobre los menores valores resultantes con el índice, recordado que estos valores están relacionados con su recurrencia media, comparados con los que resultan de la caracterización multivariada a partir de la TR. Esta diferencia se da principalmente cuando son usados umbrales altos. En ellos no puede decirse que se este produciendo una desviación significativa respecto del régimen de variación normal. Sin embargo, sí se comprueban mayores similitudes en ordenes de magnitud entre las dos metodologías para los umbrales más bajos, aunque la caracterización que se obtiene con la definición de rachas no tiene en cuenta todas las condiciones que se han ido produciendo hasta el momento de su evaluación, por lo que la caracterización de la persistencia, así entendida, no puede ser obtenida para umbrales muy bajos.

Profundizando en el motivo de las diferencias, la aplicación de la TR está muy condicionada por la definición del umbral, y por tanto de los déficit que se producirían sin regulación artificial. Sin embargo, el concepto del IFS no está asociado a ninguna definición de umbral, considerando que desviaciones que de acuerdo a la TR pueden resultar significativas, son desviaciones dentro del régimen de variación normal de la variable aleatoria, y por tanto no susceptibles de ser consideradas como extremas (e.g el evento de 1987/92). De aquí, puede concluirse que frente al objetivo de realizar una caracterización general de sequías, donde no exista un condicionante que defina el umbral de sequías, resulta no recomendable el uso de la TR, pues el análisis puede resultar ambiguo. Por otro lado, cuando tampoco es definida la consideración en el análisis de una o varias magnitudes críticas sobre las que desarrollar un modelo de recurrencia, la selección de la magnitud puede introducir mayor grado de ambigüedad en el análisis. Ello podría evitarse si se utilizan un análisis multivariado, pero su realización viene condicionada con la disponibilidad de información suficiente para la caracterización completa de la variabilidad en los eventos.

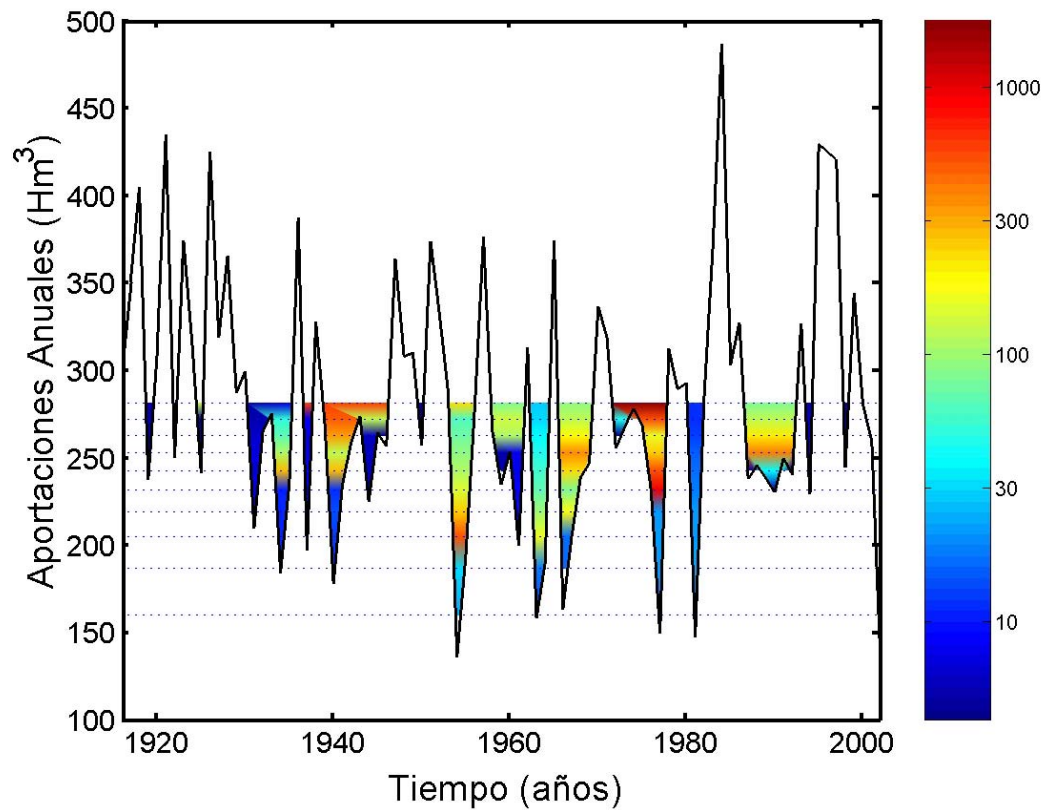


Figura 6.5: Caracterización trivariada (duración, severidad y coeficiente uniformidad) de las sequías a partir del periodo de retorno (expresado en años) obtenido mediante la Teoría de Rachas, sobre la serie de aportaciones anuales en el río Blue, en función del umbral de definición del estado de sequía (señalados los umbrales de cuantiles del 5% al 50%, en intervalos de 5%)

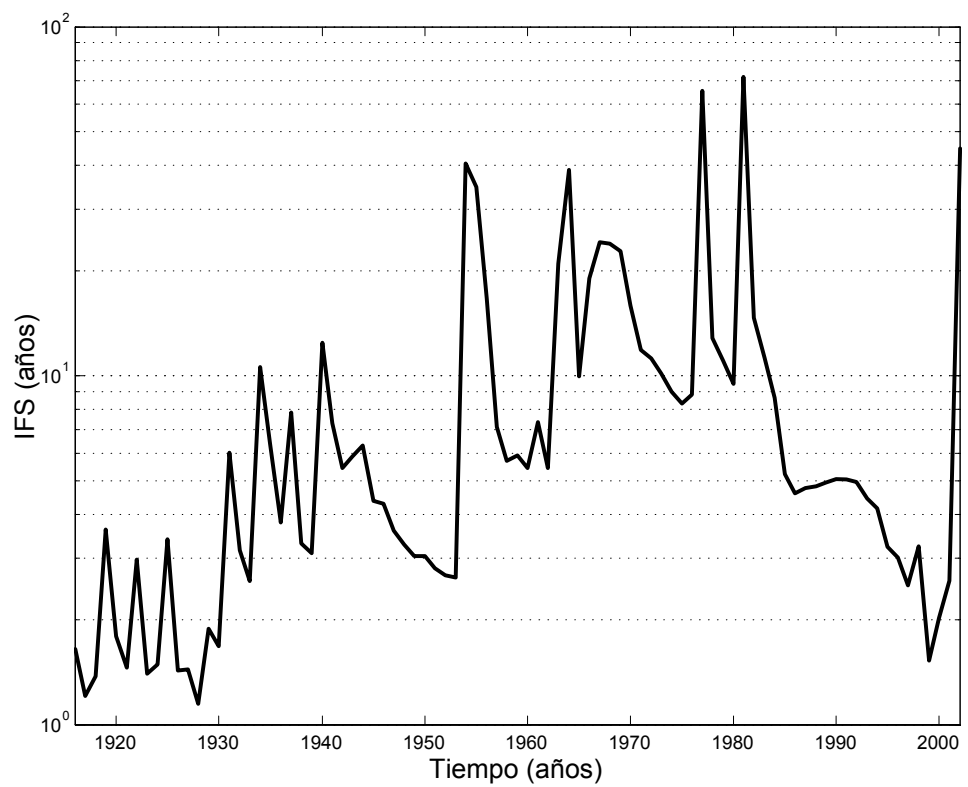


Figura 6.6: Análisis IFS sobre la serie de aportaciones anuales del río Blue.

6.2.3. El efecto de la escala temporal

La escala temporal utilizada en el análisis de caracterización de sequías puede resultar en algunos casos un impedimento para la completa caracterización de un conjunto de eventos de sequías. La escala temporal influye cuando se analizan períodos de agregación, sobre los que son acumulados los valores realizados de la variable hidroclimática estudiada. Así puede ocurrir con los valores de precipitaciones, analizados a escala mensual, trimestral o anual, o de igual manera en el caso de aportaciones. El análisis sobre diferentes escalas de agregación puede producir discrepancias sobre la ocurrencia o no de una situación de sequía en un instante.

El Índice de Precipitación Estandarizado y la escala temporal de sequías

El Índice de Precipitación Estandarizado (*Standardized Precipitation Index*, SPI), introducido en la sección 2.3, caracteriza una serie de precipitaciones indicando en cada instante el valor de la variable normal estandar equivalente que mantiene el mismo valor de su función de probabilidad acumulada. De este modo, la naturaleza del índice es probabilista, expresando las posibles desviaciones con respecto de su distribución en una escala adimensional e independiente de la población analizada. En la definición del SPI se hace notar la importancia de la escala temporal a la que se realiza el análisis, proponiendo repetir el análisis para diferentes escalas de agregación de modo que puedan ser estudiadas diferentes tipos de sequías, correspondientes a desviaciones de diferente duración (McKee *et al.*, 1993). Como un ejemplo de aplicación, la figura 6.7 muestra los resultados de realizar un análisis SPI sobre los registros de precipitaciones de la División Climática 5 de California (EEUU) en el período 1895 a 2003. Esta región climática coincide sensiblemente con la extensión cubierta por la cuenca del río San Joaquín en California.

En el análisis se han utilizado tres escalas temporales, precipitación trimestral, semestral y anual. Los resultados muestran como eventos de corta duración, que resultan con un valor del índice significativamente bajo en un análisis trimestral, como son ocurridos durante la primera década del siglo XX, desaparece su significancia cuando son utilizados análisis semestrales o anuales. De forma recíproca ocurre con eventos como el que tiene lugar entorno al año 1990, donde su significancia es notoria especialmente en el análisis anual, y sin detectarse una desviación significativa en el análisis trimestral.

La necesidad de realizar análisis a diferentes escalas para la caracterización de las sequías producidas en diferentes escalas temporales provoca

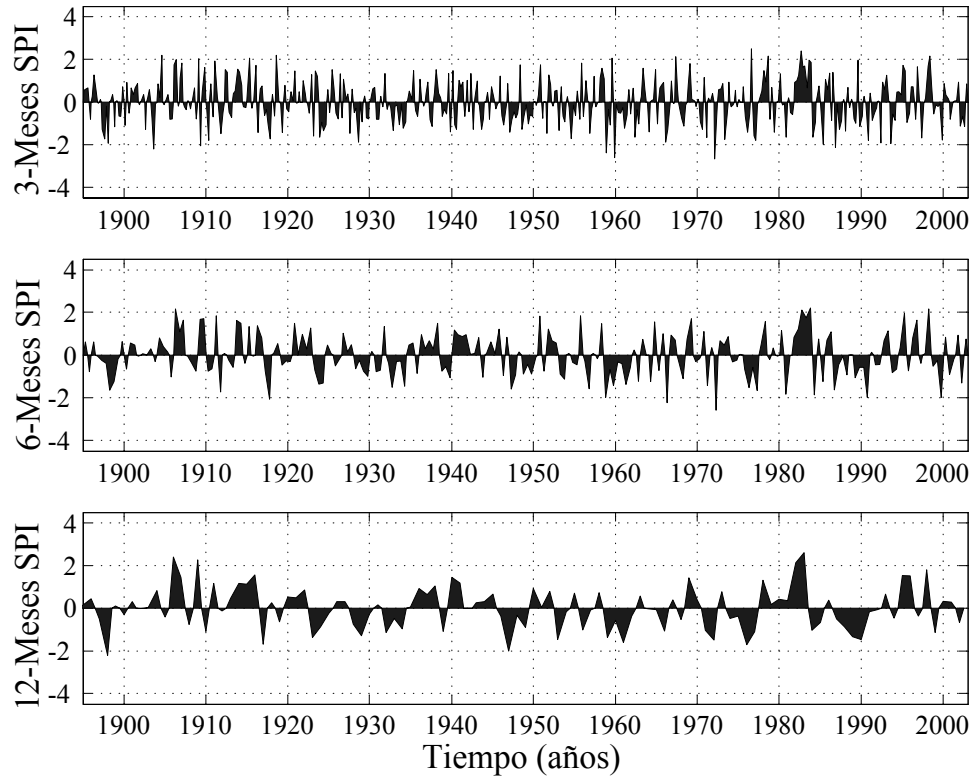


Figura 6.7: Análisis SPI sobre la serie de precipitaciones de la División Climática 5 de California (río San Joaquín), para los períodos de agregación de 3, 6 y 12 meses.

una limitación en el uso del SPI. Mediante la aplicación de este índice no es posible la comparación del carácter extraordinario entre dos eventos de sequías significativos en diferentes escalas temporales. En principio podrían ser comparados los valores mínimos obtenidos por el índice sobre cada evento, pero si estos valores proceden de análisis con escalas temporales distintas, su comparación no puede ser realizada de forma directa. Valores similares del índice correspondientes a diferentes escalas no indican el mismo grado de excepcionalidad del evento, pues si se analizaran las frecuencias con las que se producen, su frecuencia dependerá directamente de la escala temporal. Así serán menos frecuentes cuando mayor sea la escala temporal donde resultan significantivos. De este modo, el problema de la escala temporal en la caracterización de eventos de sequía extrema no se encuentra completamente resuelto en el uso del SPI.

Análisis comparativo con el IFS

Para comparar el comportamiento del IFS frente a cambios en la escala temporal, con respecto de que ocurren en las aplicaciones del SPI, la figura 6.8 presenta los resultados del análisis IFS sobre las series de precipitaciones areales trimestral, semestral y anual en la División Climática 5 de California.

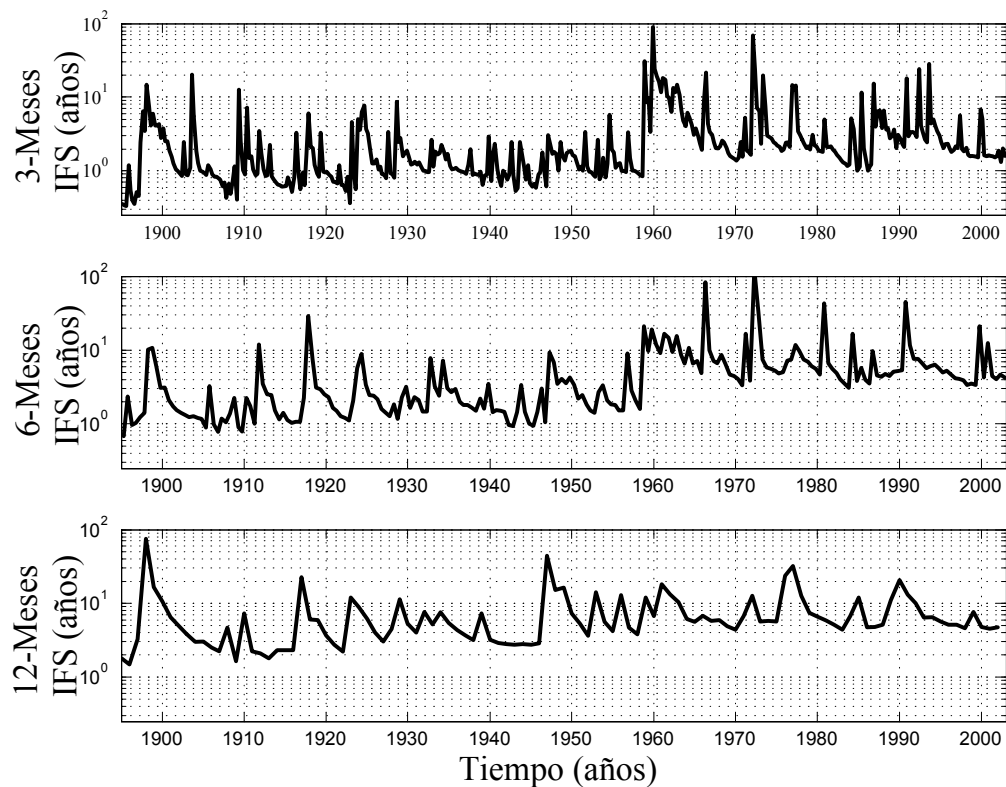


Figura 6.8: Análisis *IFS* de las precipitaciones medias areales sobre la División Climática 5 de California (EEUU), para las escalas temporales trimestral, semestral y anual.

La comparación entre las figuras correspondientes a las diferentes escalas temporales provoca como primera apreciación la suavización de la gráficas a medida que la escala temporal aumenta. La agregación continuada de la serie de realizaciones provoca que desviaciones significativas en unas determinadas escalas no sean detectables al realizar una agregación con otras realizaciones. Sin embargo, lo que si se comprueba en la mayoría de los casos es la persistencia de la tendencia general de la gráfica frente al cam-

bio de escala. Así, el uso de un período de agregación largo no permite la detección de eventos cortos, pero los resultados con períodos de agregación cortos proporcionan una tendencia similar a lo que ocurriría con los períodos largos. Ello provoca que haciendo uso del análisis IFS a escala trimestral sean comparable por ejemplo la realización de eventos cortos, como los ocurridos en la primera década del siglo XX, con el evento de mayor duración producido entorno a 1990, ambos con un valor del índice entorno a los 20 años.

Aunque es notorio el mejor comportamiento del índice frente a la escala temporal, no puede afirmarse la invarianza del índice frente a la longitud de los períodos sobre los que se realiza la agregación aditiva. Eventos como el producido a finales de 1940 no resulta especialmente significativos hasta que no es realizado un análisis de la serie anual. Ello es debido a que en el diseño del índice no se ha utilizado un criterio relacionado con la adición directa, sino que lo que se analiza es la población resultante a la escala prefijada. El cambio de la escala produce cambios en la población aleatoria estudiada, por lo que los resultados pueden variar.

Esto indica un efecto "lupa" en los análisis de sequías, en la siguiente forma:

- Aumentando la escala temporal, o ampliando la duración del período de agregación, pueden quedar ocultos desviaciones significativas producidas durante períodos de duración menor.
- Disminuyendo la escala temporal, o reduciendo la duración del período de agregación, puede ocurrir que en una desviación que resulta significativa a una gran escala, resulte de menor significancia cuando se analiza a una escala mucho más pequeña sin aparecer ninguna desviación puntual altamente significativa.

Frente a este efecto, el IFS presenta la ventaja de que viene expresado en escala relacionada con la escala de frecuencia de ocurrencia, por lo que valores resultantes de diferentes escalas son comparables directamente. Por otro lado, el índice permite la detección de fenómenos persistentes, por lo que mantiene mayor constancia de sus resultados frente al cambio en la escala temporal de lo que se produce con el SPI. En este sentido, resulta más conveniente la aplicación en estudios de caracterización de sequías la aplicación del IFS. La escala a la que debería aplicarse puede depender del tipo de eventos que se desee estudiar. Así, si en particular se desea estudiar eventos de sequías agrícolas, no resulta adecuado el uso de la escala anual. Cuanto menor sea la escala utilizada en el análisis más completa puede resultar la caracterización, siendo recomendable realizar un análisis de sensibilidad a otras escalas. Sin embargo, el uso de escalas muy pequeñas, puede provocar

que se esté analizando una variable con una alta y compleja correlación temporal, donde el uso de esta metodología puede resultar complejo y costoso desde el punto de vista computacional (sección 5.3.2).

6.2.4. La comparabilidad espacial en los índices de sequías

Las sequías se manifiestan mediante desviaciones del régimen normal de variación de una variable hidrológica relacionada con la disponibilidad de agua. En diferentes regiones puede producirse que el régimen de variación de un mismo tipo de variable hidrológica sea diferente, por lo que no resulte adecuada la comparación directa de los valores realizados sobre la variable analizada. Los índices de sequía con carácter general buscan la cuantificación del grado de excepcionalidad de una sequía mediante su expresión en una escala adimensional, independiente de la región analizada. Entre los índices de sequías de carácter general se encuentra el Índice de Severidad de Sequías de Palmer (*Palmer Drought Severity Index*, PDSI), presentado en la sección 2.3, y cuya comparabilidad espacial es analizada a continuación.

El Índice de Severidad de Sequías de Palmer

El PDSI fue diseñado para la caracterización de sequías meteorológicas. En su justificación se busca la obtención de un índice que permita la caracterización cuantitativa de las sequías producidas sobre la mitad sur de EEUU (Palmer, 1965), donde se producen los climas árido y semiáridos. La particularidad de las regiones para las que fue diseñado se pone de manifiesto en los coeficientes que son utilizados en su cómputo, que son ajustados empíricamente con este fin. Sin embargo, el uso del PDSI se ha visto extendido, hasta el punto que es aplicado en la actualidad a la totalidad de los EEUU e incluso en múltiples regiones a lo largo del planeta. El uso del índice en diferentes regiones, y con climatologías muy distintas de las que fueron usadas en su diseño, provoca que se ponga en entredicho su aplicabilidad. Sin embargo, posiblemente la falta de un criterio o índice mejor y tan conocido motivan que se siga aplicando en la actualidad (Alley, 1984).

Para ilustrar este punto, son utilizados los casos de las Divisiones Climáticas 5 y 2 de California (EEUU). La primera de ellas, como ya se mencionó anteriormente, coincide sensiblemente con la cuenca del río San Joaquín, mientras que la segunda coincide en este caso con la cuenca del río Sacramento (figura 6.9). Estas dos cuencas, aunque próximas, presentan climas muy distintos. La cuenca del río San Joaquín es una cuenca de clima semiárido, de acuerdo con la clasificación de Köppen, mientras que la cuenca del río Sacramento corresponde a un clima húmedo. La figura 6.10 muestra los resultados del análisis PDSI de sequías meteorológicas sobre ambas cuencas.

Los resultados muestran en general la coincidencia en el tiempo de los

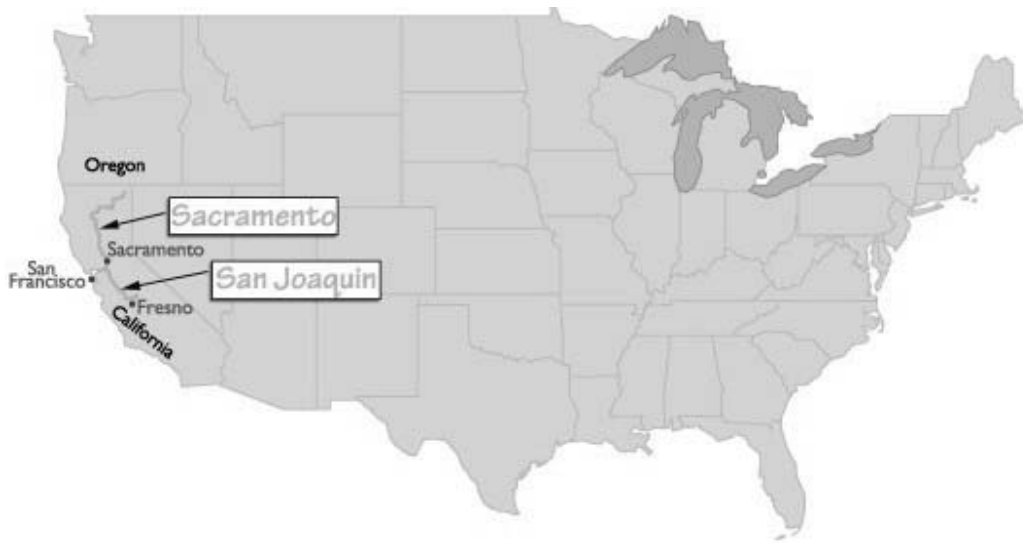


Figura 6.9: Detalle de la posición de los ríos Sacramento y San Joaquín sobre EEUU

fenómenos de sequías más importantes producidos sobre ambas cuencas, indicados por un valor de índice por debajo de -1 y de mayor severidad cuanto menor resulte el índice. Así mismo se aprecia que los valores mínimos alcanzados en cada evento son menores en el caso de la cuenca del río San Joaquín (CaCD5) que los producidos sobre la cuenca del río Sacramento (CaCD2). De aquí podría deducirse que la severidad de las sequías más importantes (eventos entorno a 1898, 1925 y 1977) sobre la cuenca del río San Joaquín son siempre mayores que las producidas sobre la cuenca del río Sacramento. Esta conclusión, a partir de los resultados del PDSI, es analizada con mayor detalle en el siguiente punto, comparando con los resultados de otros índices.

Análisis de la comparabilidad espacial del PDSI

Para comprobar la bondad de los resultados anteriores, obtenidos al utilizar el PDSI en la caracterización de eventos de sequía entre diferentes regiones, han sido comparados los resultados con el uso del SPI. La figura 6.11 muestra los resultados de aplicar el SPI sobre la División Climática 2 de California (EEUU), coincidente con la cuenca del río Sacramento. Comparando los resultados obtenidos con el uso del SPI sobre ambas cuencas se tiene que episodios de sequía como el ocurrido entorno a 1925 y 1977 resultan más significativo sobre la cuenca del río Sacramento que sobre la cuenca del río San Joaquín (comparando los SPI anuales), mientras que otros eventos como los ocurridos entorno a 1898 ó 1990 puede resultar de

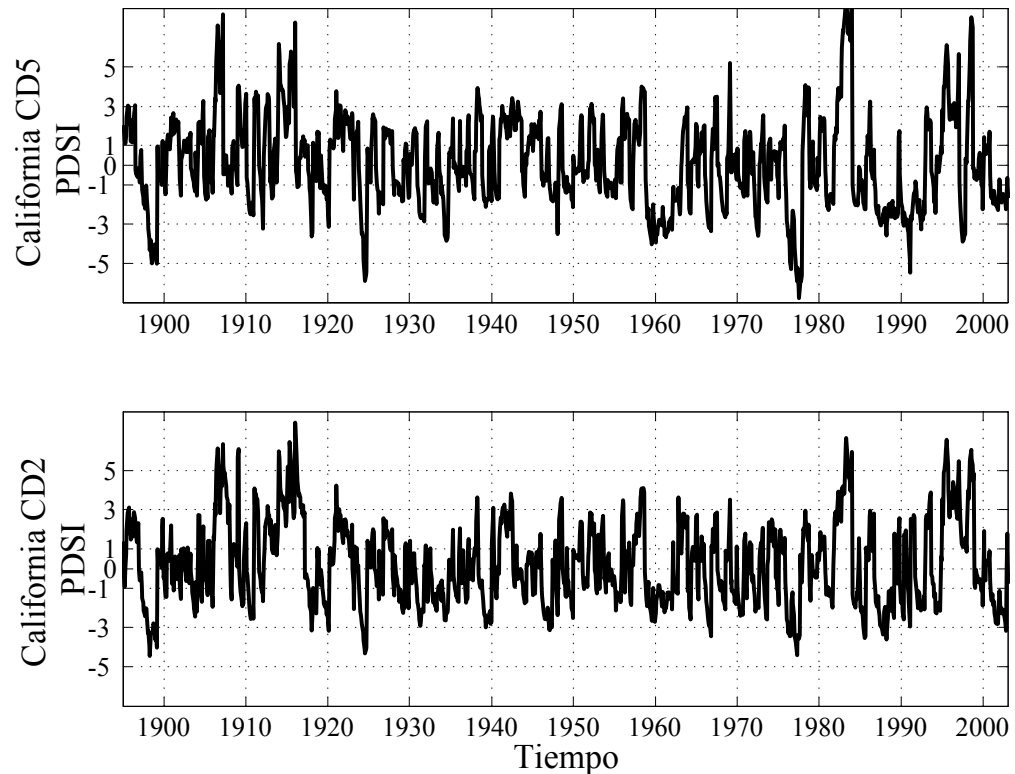


Figura 6.10: Análisis PDSI de las sequías producidas sobre las Divisiones Climáticas 5 y 2 de California, EEUU (fuente: *National Climatic Data Center, USA*).

similar significancia.

El índice SPI es de naturaleza estadística, a diferencia del PDSI que mantiene una naturaleza empírica. Desde este punto de vista, la posibilidad de comparación de valores de cada uno de estos índices en diferentes regiones está más justificada para el caso del SPI, por usar una escala absolutamente independiente. Sin embargo, no lo está para el caso del PDSI, especialmente cuando alguna de las regiones donde se realiza el análisis no coincide con las regiones para la que se ajuste el índice. De este modo, la contradicción que se ha puesto de manifiesto entre las aplicaciones sobre las cuencas del río San Joaquín y Sacramento son fruto del sesgo del PDSI. Ésto ya fue expuesto por Guttman (Guttman, 1998), quien recomienda el uso de indicadores de naturaleza probabilista para la caracterización de fenómenos de sequías, de modo que permitan la comparación entre resultados

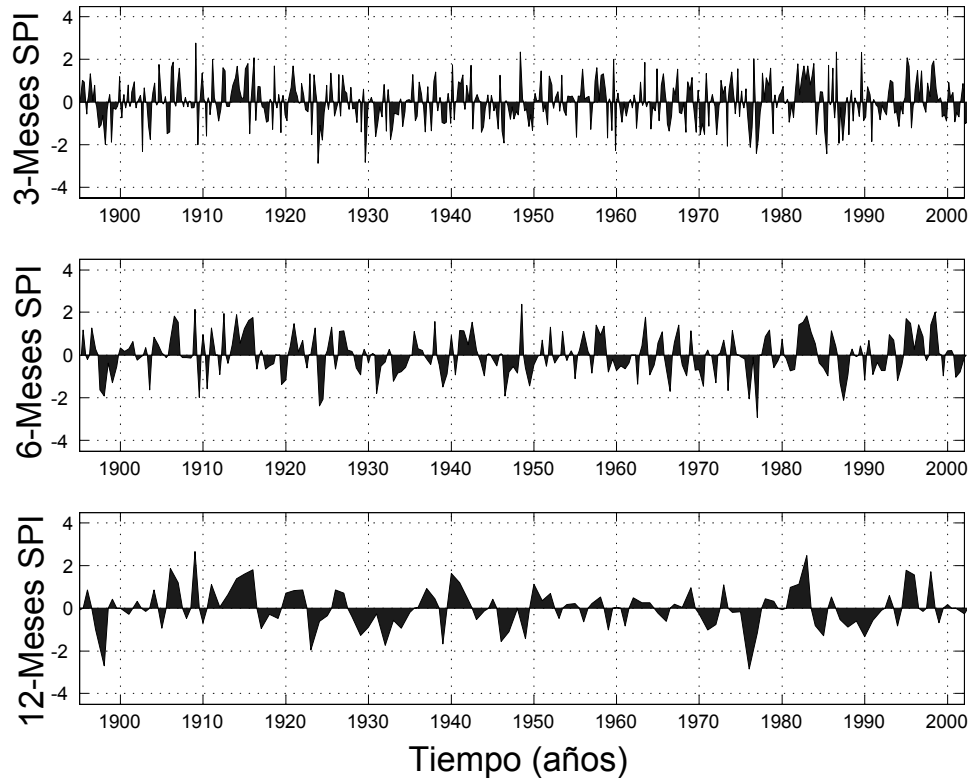


Figura 6.11: Análisis SPI sobre la serie de precipitaciones de la División Climática 2 de California (río Sacramento), para los períodos de agregación de 3, 6 y 12 meses.

procedentes de diferentes regiones climáticas. Para soportar esta afirmación es realizado también el análisis IFS sobre la serie de precipitaciones trimestrales de la División Climática 2 de California (EEUU). A partir de él es posible comparar los resultados del análisis IFS sobre ambas cuencas.

Comparando la figura 6.12 con la primera de las gráficas en la figura 6.8, se comprueba que los eventos producidos entorno a 1898 y 1990 resultan de similar significancia las cuencas de los río San Joaquín y Sacramento, mientras que los eventos producidos entorno a 1925 y 1977 resulten más significativa la realización de las sequías sobre la cuenca del río Sacramento. Esto está en clara discrepancia con lo manifestado por el PDSI, para los que resultan siempre más severos los efectos de estas sequías sobre la cuenca del río San Joaquín. Su justificación, como ya se ha hecho notar, se encuentra en el empirismo residente en la definición del índice PDSI, ajustado para regiones áridas y semiáridas, y cuya aplicación en regiones húmedas produce

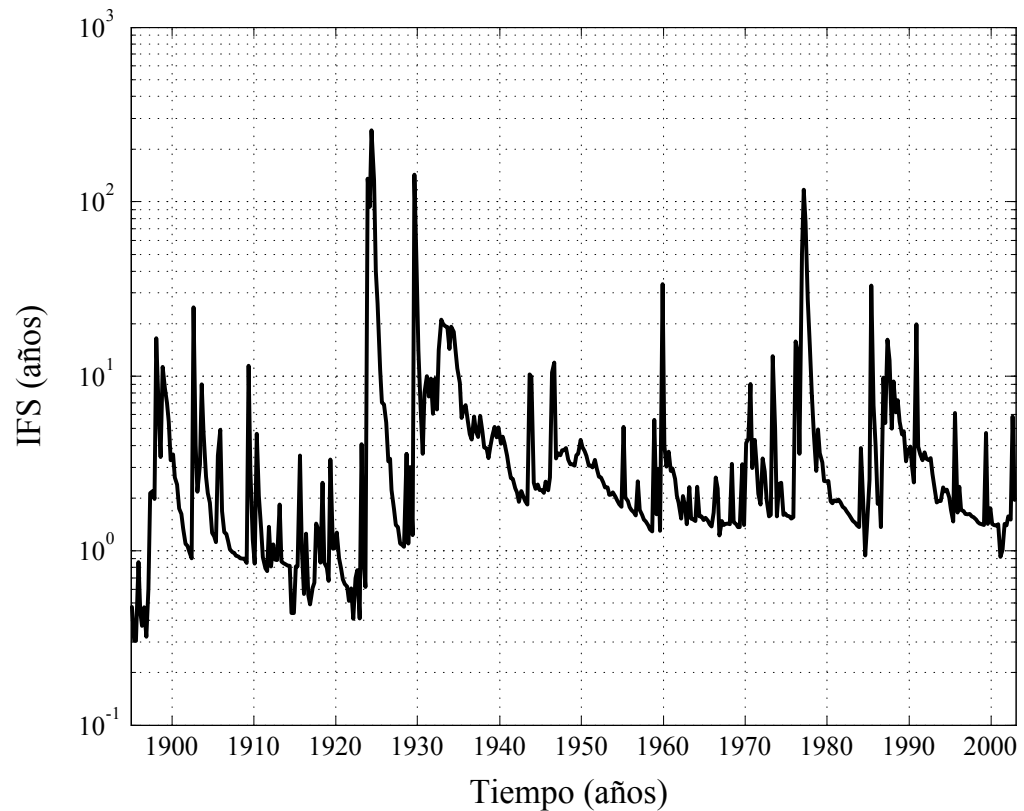


Figura 6.12: Análisis IFS de la serie de precipitaciones trimestral sobre la División Climática 2 de California (EEUU)

siempre resultados menos significantes que los producidos con índices de naturaleza probabilista.

Este ejemplo ilustra la conveniencia de la aplicación de índices como el SPI o el IFS cuando se pretende realizar comparaciones de las manifestaciones de episodios de sequías sobre regiones con diferente climatología o variables de diferente regímenes de variación. A este aspecto se añaden la conveniencia del uso del IFS sobre el SPI en lo que respecta a los problemas relacionados con la escala temporal del análisis, y el mejor comportamiento del IFS frente a la aplicación de modelos estocásticos basados en la TR en una caracterización general.

Capítulo 7

Caracterización espacio-temporal de sequías

Las sequías, además de suponer una reducción en la disponibilidad de recursos hídricos prolongada en el tiempo, se le asocia como característica adicional que le puede conferir un carácter extremo, su extensión en el espacio. El IFS, tal y como ha sido definido en el Capítulo 6, basado en la metodología presentada en el Capítulo 5, permite la caracterización estocástica de los eventos extremos de sequías a partir del análisis de las realizaciones de una variable aleatoria, que puede resultar como agregación de determinados procesos ocurridos de forma distribuida en el espacio. Sin embargo, de forma general, el modo en el cual se ha distribuido en el espacio la realización de un determinado evento extremo de sequía es de tanta importancia como la que puede ser su distribución temporal. Ello motiva la necesidad de complementar la metodología de análisis expuesta en los capítulos anteriores, con un procedimiento que permita incluir en la caracterización el análisis de la evolución que ha sufrido en el espacio la realización de un episodio de sequía extrema.

Al igual que en la caracterización en el tiempo, la caracterización de realizaciones aleatorias en el espacio requiere la definición del criterio adoptado. Muchos pueden ser los criterios que se adopten, siendo su mayor o menor adecuación función únicamente del objetivo buscado con el análisis. A continuación se presentan un conjunto de criterios de caracterización, algunos de los cuales son convenientes con objetivos concretos, y otro son más adecuados para una caracterización general. Dentro de los criterios generales se presenta como criterio probabilístico la función de probabilidad acumulada de probabilidad conjunta. Es este último el que se desarrolla en este capítulo como criterio de análisis de las realizaciones espacio-temporales de sequías.

7.1. Criterios para la caracterización de una realización espacial aleatoria

La realización en el espacio de un evento de sequía se manifiesta en la distribución espacial de diferentes variables hidro-climáticas (por ejemplo precipitación, humedad del suelo, nivel piezométrico, producción de escorrentía superficial, etc.). Tomando alguna de estas variables como indicador, puede ser analizada en cada instante t , la distribución espacial de la realización de x , en una región Ω , mediante diferentes criterios. Los expuestos en esta sección se pueden clasificar en tres grupos: criterios de asignación directa, criterios derivados de la definición de un umbral, criterios de tipo probabilístico.

Los criterios de asignación directa, mantienen similitudes con los criterios para la caracterización de secuencias temporales, analizados en el punto 5.1. En estos criterios, se asigna a la realización en una región Ω una magnitud que proviene directamente de los valores realizados:

- El valor acumulado del proceso de realización en la región Ω .

$$Y_{t,\Omega} = \iint_{\Omega} X_t \cdot d\Omega \quad (7.1)$$

- El valor medio de la realización espacial en la región Ω , que cuantifica la tendencia general que se ha producido en el espacio con independencia de su extensión.

$$\bar{X}_{t,\Omega} = \frac{\iint_{\Omega} X_t \cdot d\Omega}{\iint_{\Omega} d\Omega} \quad (7.2)$$

- El valor mínimo producido en la realización, dentro de la región Ω , que asigna al conjunto, el valor mínimo realizado en la región.

$$X_{\min,t,\Omega} = \min_{\Omega} (X_t) \quad (7.3)$$

Estos criterios simples pueden componerse en criterios más complejos, como puede ser el valor mínimo medio, producido en el instante t , de una sub-región dentro de Ω , con extensión $\alpha \cdot \iint_{\Omega} d\Omega$, con $\alpha < 1$.

Los criterios derivados de la definición de un umbral, U , resultan en muchos casos como generalización de aquellos criterios homónimos utilizados en la caracterización de rachas:

- La intensidad máxima espacial, definida como la mayor de las diferencias entre el umbral y las realizaciones producidas en la región Ω . Caracteriza la realización espacial por su mayor desviación puntual con respecto al umbral, sin notar la distribución del resto de realizaciones en la región.

$$I_{\text{máx},\Omega} = \underset{\Omega}{\text{máx}} (U - X_t) \quad (7.4)$$

- La intensidad media espacial, o el déficit promedio producido en la región Ω . Se define como el promedio espacial de las desviaciones con respecto al umbral U . De forma similar al valor medio de la realización espacial, proporciona la tendencia general en la región Ω . Su cálculo puede derivarse del valor medio de la realización espacial:

$$\bar{I}_{\Omega} = \frac{\iint_{\Omega} (U - X_t) \cdot d\Omega}{\iint_{\Omega} d\Omega} = U - \bar{X}_{t,\Omega} \quad (7.5)$$

- La severidad espacial, o déficit acumulado espacialmente, definida como la acumulada de las diferencia entre el umbral y las realizaciones producidas en la región Ω . Está relacionada con la extensión de Ω y la intensidad media espacial, cuantificando el déficit total producido en la región Ω .

$$S_{\Omega} = \iint_{\Omega} (U - X_t) \cdot d\Omega = \bar{I}_{\Omega} \cdot \iint_{\Omega} d\Omega \quad (7.6)$$

- El factor de uniformidad espacial, que resulta como generalización del coeficiente de uniformidad de una racha, definido para cuantificar el grado de uniformidad en la desviación producida en la región Ω , con respecto al umbral U .

$$F_{U,\Omega} = -\log \left(\frac{\frac{I_{\text{máx},\Omega}}{S_{\Omega}} - \frac{1}{\iint_{\Omega} d\Omega}}{1 - \frac{1}{\iint_{\Omega} d\Omega}} \right) \quad (7.7)$$

Por último, los criterios de tipo probabilísticos son aquellos que asignan una probabilidad a cada realización producida en el espacio. Cualquiera de los criterios anteriores puede ser derivado en un criterio probabilístico, caracterizando la realización espacial en lugar de por el valor de la magnitud

generada, por la probabilidad de realización de dicha magnitud en las realizaciones sobre la región Ω . A continuación se presentan tres criterios, el primero de ellos se deriva del valor medio de la realización espacial, y los dos últimos producen una caracterización probabilística como agregación probabilística directa de la distribución de realizaciones producidas, sin tener en cuenta ninguna magnitud concreta.

- El análisis estadístico de la serie de valores medios espaciales (ec. 7.2) producidos en cada instante en una región Ω , permite caracterizar la realización producida en esta región, asignándole en cada instante t , el valor de su función de probabilidad acumulada (*f.p.a.*) (ec. 7.8). Esta probabilidad representa la de ocurrencia de realizaciones con media espacial en la región Ω inferior o igual a la producida en ese instante.

$$F_{x_\Omega}(\bar{X}_{t,\Omega}) = \mathcal{P}(\bar{x}_\Omega \leq \bar{X}_{t,\Omega}) \quad (7.8)$$

- Si en lugar de utilizar el valor medio de las realizaciones, se calcula el valor medio de las *f.p.a.* producidas en cada punto de la región, se obtiene el valor medio de la *f.p.a.* en la región Ω . Así en el caso que en todos los puntos de la región Ω , la variable aleatoria x mantenga la misma *f.p.a.*, el valor medio de las *f.p.a.* producidas resulta:

$$\bar{F}_{x,\Omega}(X_t) = \frac{\iint_{\Omega} F_x(X_t) \cdot d\Omega}{\iint_{\Omega} d\Omega} \quad (7.9)$$

En los casos que la *f.p.a.* no sea uniforme en la región, F_x dependerá además del punto donde se evalúa dentro de la región.

Este criterio no trata de modo natural la agregación de las probabilidades, de acuerdo a la Teoría Clásica de Probabilidad (Kolmogorov, 1956), ya que sucesos ocurridos simultáneamente se agregan en términos probabilísticos como el producto de sus probabilidades (cuando éstos son independientes; de otro modo, siguiendo el Teorema de Multiplicación).

- En el caso que la realización producida en cada punto del espacio sea independiente del resto de realizaciones, se define la función de probabilidad acumulada espacial conjunta (*f.a.e.c.*) como:

$$F_{x,\Omega}(X_t) = \exp \left\{ \frac{\iint_{\Omega} \ln [F_x(X_t)] \cdot d\Omega}{\iint_{\Omega} d\Omega} \right\} \quad (7.10)$$

Esta función caracteriza la ocurrencia de una realización calculando la probabilidad acumulada conjunta de ocurrencia de una realización dentro de una región del espacio Ω . El divisor dentro del exponente permite eliminar el efecto de escala que supone la introducción de la dimensión espacial de la región.

El criterio de la *f.a.e.c.* se muestra como el de mayor coherencia seguida, de acuerdo con la Teoría Clásica de Probabilidad, para caracterizar una realización espacial. Sin embargo, la hipótesis de independencia que permite aplicar la ec. 7.10 en su cálculo rara vez se puede considerar. Por otro lado, el criterio basado en la *f.p.a.* de los valores medios resulta el más simple.

A continuación se presenta una metodología que permite calcular la *f.a.e.c.* en una región, en aquellos casos en los que no es asumible la hipótesis de independencia espacial en las realizaciones. Con ello se hará posible la aplicación común del concepto de la *f.a.e.c.* para la caracterización espacio-temporal de sequías.

7.1.1. Evaluación de la *f.a.e.c.* mediante análisis de componentes principales

En la mayoría de las aplicaciones prácticas donde se produce el seguimiento espacial distribuido de variables aleatorias de carácter hidro-climático, el control se realiza de manera puntual en un conjunto de punto, uniforme o no-uniformemente distribuidos en el espacio. Para representar la distribución espacial de esas realizaciones puntuales, habitualmente o bien se asigna a cada punto de observación una sub-región dentro de la región de estudio, o bien se construye una función de interpolación que permite estimar el valor producido en cualquier punto de la región. Aquí se supondrá que se tiene una red (regular o irregular) compuestas por celdas que discretizan la región espacial de interés. Cada una de las celdas contiene una serie de realizaciones del valor medio areal de la variable indicadora en la celda. Si cuando se analizan las correlaciones entre las series de celdas vecinas, se observa una significativa correlación, no es posible aplicar la ec. 7.10 para el cálculo de la *f.a.e.c.*, pues hay evidencia del no cumplimiento de la hipótesis de independencia. En tal circunstancia pueden ser aplicadas técnicas como el Análisis de Componentes Principales (ACP) para la evaluación de la *f.a.e.c.*

Análisis de componentes principales de una región espacial discretizada en una red

Sea una región Ω , discretizada en una red ordenada $\Omega_{\#}$ (regular o irregular) de N celdas Ω_i , $i = 1, \dots, N$, que discretizan la región Ω , sobre cada

una de las cuales se observa la realización distribuida de una variable aleatoria $x_{t,i}$, donde t representa el instante en el que se observa, $t = 1, \dots, L$, e i representa el índice ordinal de la celda en la red $\Omega_{\#}$. La matriz $[X]_{t,\Omega}$ organiza el conjunto de las series de realizaciones sobre cada celda, donde cada columna representa una celda, y en filas se indica la realización temporal:

$$[X]_{t,\Omega} = \begin{bmatrix} X_{t=1,\Omega_1} & X_{t=1,\Omega_2} & \cdots & X_{t=1,\Omega_N} \\ X_{t=2,\Omega_1} & X_{t=2,\Omega_2} & \cdots & X_{t=2,\Omega_N} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t=L,\Omega_1} & X_{t=L,\Omega_2} & \cdots & X_{t=L,\Omega_N} \end{bmatrix}_{L \times N} \quad (7.11)$$

Cada una de estas variables aleatoria que es observada en cada celda de $\Omega_{\#}$ seguirá una distribución de probabilidad poblacional $F_{i,\Omega_{\#}}$. Si cada una de las realizaciones $X_{t,i}$ es transformada al valor de la variable normal estándar equivalente, $Z_{t,i} = \Phi^{-1} [F_{i,\Omega_{\#}} (X_{t,i})]$, se obtiene la matriz $[Z]_{t,\Omega}$ de realizaciones transformadas. A partir de esta matriz, puede ser calculada la matriz de varianzas-covarianzas de las variables $Z_{t,i}$ siguiendo ec. 7.12. Al distribuirse todas las variables transformadas según una $N(0, 1)$, todos los elementos de la diagonal de la matriz $[S]_{\Omega}$ resultan igual a la unidad, resultando su traza igual al número de celdas de la red, $\text{traza}([S]_{\Omega}) = N$.

$$[S]_{\Omega} = \frac{1}{L-1} \cdot [Z]_{t,\Omega}^T \cdot [Z]_{t,\Omega} \quad (7.12)$$

La aplicación del ACP a la matriz simétrica $[S]_{\Omega}$ produce la matriz de vectores principales ortonormal $[\mathcal{U}]_{\Omega}$, $\sum_{j=1}^N \mathcal{U}_{j,i}^2 = 1$ $i = 1, \dots, N$, representando cada componente por columnas, y la matriz diagonal de valores principales $[\Lambda]_{\Omega}$. Cada vector principal representa un modo principal de variación. Se supondrá que el orden en el que se encuentran los modos principales es en el orden decrecientes de su valor principal, $\Lambda_1 \geq \Lambda_2 \geq \dots \geq \Lambda_N$. El conjunto de vectores principales forman una base en el espacio algebraico real N -dimensional. Cada valor principal, Λ_i , $i = 1, \dots, N$, coincide con la varianza de la serie formada por las componentes de $[Z]_{t,\Omega}$ proyectada sobre el vector principal $[\mathcal{U}_{\cdot,i}]$ (ec. 7.13 y 7.14), cumpliéndose que $\sum_{i=1}^N \Lambda_i = N$, al ser la traza de una aplicación lineal un invariante en los cambios de base.

$$[Y]_{t,\Omega} = [Z]_{t,\Omega} \cdot [\mathcal{U}]_{\Omega} \quad (7.13)$$

$$\sigma^2([Y_{\cdot,i}]) = \Lambda_i, \quad i = 1, \dots, N \quad (7.14)$$

Del hecho de que cada serie original, $[Z]_{t,\Omega}$, proveniente de la transformación de las series de realizaciones en cada celda, mantiene varianza

unidad, la relación Λ_i/N , $i = 1, \dots, N$, para cada valor propio, indica en tanto por uno la proporción de la varianza total de $[Z]_{t,\Omega}$ explicada por la componente i . Por tanto el ACP supone una desagregación de la varianza total de $[Z]_{t,\Omega}$ en modos de variación principales sobre el conjunto de las celdas que forman $\Omega_{\#}$. Seleccionando aquellos modos principales que representan la mayor varianza, la varianza del conjunto de variables puede ser aproximada con un número limitado de modos de variación, mucho menor que N . Si cada modo explica una proporción Λ_i/N de la varianza total N , que coincide con el número de celdas, se puede asimilar que cada modo de variación principal explica proporciones de la varianza en cada celda hasta un total de Λ_i . Así, la serie producida en cada celda se puede descomponer en la suma de los N modos principales de variación:

$$[Z_{\cdot,j}] = [Y]_{t,\Omega} \cdot [\mathcal{U}_{j,\cdot}]^T \quad (7.15)$$

Dado que cada una de las series de componentes $[Y_{\cdot,i}]$, $i = 1, \dots, N$, resultan mutuamente independientes, por ser componentes de una base ortogonal, la varianza de cada celda también puede ser descompuesta como suma de las varianzas de cada modo de variación principal (ec. 7.16).

$$\sigma^2([Z_{\cdot,j}]) = \sum_{i=1}^N \sigma^2([Y_{\cdot,i}]) \cdot \mathcal{U}_{j,i}^2 = \sum_{i=1}^N \Lambda_i \cdot \mathcal{U}_{j,i}^2 \quad (7.16)$$

El modo de variación principal i explica en la celda j una porción $\Lambda_i \cdot \mathcal{U}_{j,i}^2$ de su varianza. Agregando la proporción explicada en cada celda para el conjunto de la red $\Omega_{\#}$, se tiene nuevamente la varianza total explicada por el modo de variación principal i : $\sum_{j=1}^N \Lambda_i \cdot \mathcal{U}_{j,i}^2 = \Lambda_i \cdot \sum_{j=1}^N \mathcal{U}_{j,i}^2 = \Lambda_i$. A partir de esta descomposición en cada celda de $\Omega_{\#}$, sería posible el cálculo de la *f.a.e.c.* (ec. 7.10) en la región Ω , asignándole a cada modo de variación principal en cada celda, una proporción del área igual a la proporción de la varianza explicada por ese modo en la celda, siguiendo ec. 7.17.

$$F_{x,\Omega}(X_t) = \exp \left\{ \frac{\sum_{i=1}^N \left\{ \ln [F_{y,i}(Y_{i,t})] \cdot \sum_{j=1}^N \left(\Lambda_i \cdot \mathcal{U}_{j,i}^2 \cdot \iint_{\Omega_j} d\Omega_j \right) \right\}}{\sum_{j=1}^N \iint_{\Omega_j} d\Omega_j} \right\} \quad (7.17)$$

Las expresiones ec. 7.16 y 7.17 pueden ser aproximadas incluyendo en el sumatorio, en lugar del total de los modos, N , despreciar aquellos que resultan de poca potencia, incluyendo sólo los más significativos ($\Lambda_i \geq \Lambda_C$). Sin embargo, en una descomposición tal y como la formulada aplicada al

cálculo de la *f.a.e.c.* (ec. 7.17) se pierde la discriminación de la situación de sequía, al no estar asegurada el sentido físico de cada una de los modos principales. Las *f.p.a.* de cada modo principal, $F_{y,i}$, permiten cuantificar en términos probabilísticos la aproximación a valores bajos de la componente del modo de variación principal i en un determinado instante. Pero ello en general no supone una aproximación a una situación de sequía extrema en las realizaciones sobre las celdas, pues cada modo de variación así obtenido es carente de sentido físico. Por ello, previamente a la aplicación de los resultados del ACP para el cálculo de *f.a.e.c.*, es necesaria la introducción de un procedimiento que seleccione los modos de variación sin que se produzca la pérdida total de su sentido físico. Para ello se dispone de procedimientos basados en la rotación de la matriz $[\mathcal{U}]_{\Omega}$, hasta conseguir la maximización de un criterio seleccionado. Existen una variedad de criterios que permiten búsquedas alternativas de modos de variación (orthomax, pasimax, pattern, procluster, promax, varimax,...), cada una de las cuales de acuerdo al objetivo buscado. En el caso de la caracterización de sequías el criterio más adecuado sería aquel en el que se buscan los modos que agrupan celdas con modos de variación similares. Así una variación de la componente de un modo de variación en un sentido, indica la variación en ese sentido de las realizaciones producidas en las celdas donde predomina ese modo. El criterio que mantiene un objetivo similar a este, y que resulta más popularmente utilizado por su eficiencia, es el criterio varimax, que se describe a continuación.

El criterio varimax y la rotación de componentes principales

El ACP permite la desagregación de los modos de variación de un conjunto de datos, proporcionando los modos de variación principales, cada uno de los cuales explica por sí solo la mayor proporción de la varianza total. Como técnica derivada se encuentra el Análisis de Factores (*AF*), técnica en la que en lugar de hacerse el mayor énfasis en la varianza, se hace en la covarianza, buscando modos de variación comunes en las variables que componen el conjunto de datos analizados. El procedimiento matemático utilizado en el *AF* se basa en la aplicación del ACP al conjunto de datos a analizar, con lo que se consigue la matriz de vectores principales $[\mathcal{U}]$. Sobre esta matriz se realiza una rotación buscando maximizar una función objetivo que exprese matemáticamente los fines buscado. La rotación de $[\mathcal{U}]$ se realiza mediante una matriz de rotación $[R]$, la cual puede ser ortonormal o puede permitirse la posibilidad de que sea oblicua, en función del criterio y algoritmo utilizado (pattern, procluster, o promax permiten rotaciones oblicuas). Muchos de los criterios, o los algoritmos que los implementan, producen resultados similares, siendo más eficientes cada uno de ellos para

aplicaciones particulares. Aquí ha sido seleccionado el criterio varimax por ser un criterio profusamente utilizado en la práctica, existiendo algoritmos de optimización muy eficientes, y cuya función objetivo se ajusta a los criterios buscado para el cálculo de la *f.a.e.c.*

En el criterio varimax (Kaiser, 1958) el objetivo es conseguir una rotación ortogonal de modo que cada modo de variación tenga una alta correlación con unas variables, reducida en otras, e intermedia en un número mínimo de variables. Ello, en el caso del análisis de la matriz $[Z]_{t,\Omega}$, implica que cada modo de variación explicaría la práctica totalidad de la varianza en un número dado de celdas de $\Omega_{\#}$, no tendría efecto sobre la varianza de otras celdas, y explicaría de forma compartida con otros modos de variación la varianza de un número mínimo de celdas. Para presentar la función objetivo del criterio varimax, se define previamente la matriz de pesos (*loading factors matrix*), como la matriz de correlaciones entre las variables analizadas, $[Z_{\cdot,j}]$ $j = 1, \dots, N$, y las componentes de los modo de variación obtenidos tras la rotación, $[H_{\cdot,i}]$ $i = 1, \dots, N$, que sigue:

$$[P]_{\Omega} = \frac{1}{N-1} [Z]_{t,\Omega}^T \cdot [H]_{t,\Omega} \cdot [\sigma_H]_{\Omega}^{-1} \quad (7.18)$$

$$[H]_{t,\Omega} = [Z]_{t,\Omega} \cdot [U]_{\Omega} \cdot [R] \quad (7.19)$$

donde $[\sigma_H]_{\Omega}$ representa la matriz diagonal compuesta por las desviaciones típicas de las componentes de los factores, $[H_{\cdot,i}]$ $i = 1, \dots, N$. En el caso en el que no se produce rotación, $[R] = [I]$, se tiene $[\sigma_H]_{\Omega} = \sqrt{[\Lambda]_{\Omega}}$. Con esto, la función objetivo del criterio varimax resulta:

$$V([P]) = \sum_{k=1}^N \left[\frac{1}{N} \sum_{j=1}^N p_{j,k}^4 - \left(\frac{1}{N} \sum_{j=1}^N p_{j,k}^2 \right)^2 \right] \quad (7.20)$$

La maximización de ec. 7.20 implica la maximización de la varianza de cada columna de los cuadrados de los componentes de la matriz de pesos $[P]$. De este modo el criterio varimax optimiza la matriz de rotación ortogonal $[R]$ para maximizar la función objetivo 7.20, lo que resulta generalmente en que cada columna de $[P]$ esté compuesta por un número de elementos con valores absolutos altos, y el resto de elementos próximos a 0. El criterio varimax permite así desagregar la variación total incluida en las variables analizadas en modos de variación, que ya no son principales, y corresponden a modos comunes a un número de celdas en los que explica una alta proporción de su varianza. A diferencia de los modos de variación principales, los modos de variación obtenidos tras la rotación-varimax no resultan ortogonales entre sí, sino que en general mantienen covarianzas diferentes a 0. La ventaja de aplicar la rotación-varimax tras el ACP para la estimación de

la *f.a.e.c.* en análisis de sequías radica en que los modos resultantes tienen mayor sentido físico que los modos principales, ya que las componentes en cada modo indican la tendencia general observada en el conjunto de celdas a las que principalmente representa, en las que explicar una alta proporción de su varianza. Esto facilita la interpretación física de una desviación en las componentes de un factor, ya que indicaría que se produce esa misma tendencia en el conjunto de las celdas a las que representa. Debe notarse que otros métodos de análisis multidimensional como las técnicas de agrupación (*cluster analysis*) podrían aplicarse a este problema. Sin embargo éstas no resultan tan adecuadas por no considerar la estructura completa de interdependencia entre las variables.

Metodología resultante para la estimación de la *f.a.e.c.* mediante ACP y rotación-varimax

La estimación de la *f.a.e.c.*, definida en la ec. 7.10, sobre la realización de una variable aleatoria en una región del espacio, requiere el análisis de la estructura de intercorrelación de las realizaciones puntuales existente en la región de estudio Ω . En este sentido, la técnica de ACP permite determinar los modos principales de variación que existen dentro de la red de observaciones que discretizan la realización espacial. A partir del resultado del ACP pueden ser seleccionadas aquellos modos que resulten más significativos, es decir que explican la mayoría de la varianza. En la literatura pueden encontrarse multitud de criterios para la selección de los modos significativos: la prueba de dimensionalidad de Barlett (1951), la prueba de Velicer (1976), o la prueba de validación cruzada (Monteller & Wallace, 1963), entre otros. Por tratarse de una operación que en la práctica debe ser repetida un gran número de veces, se ha buscado un método simple, pero robusto. El método seleccionado es el de la prueba de Broken Stick (Jolliffe, 1986), que está basada en la distribución de las longitudes de los fragmentos producidos mediante división aleatoria en p fragmentos de un segmento de longitud unidad. La longitud esperada del k -ésimo fragmentos mayor resulta:

$$g_k = \frac{1}{p} \sum_{j=k}^p \frac{1}{j} \quad (7.21)$$

Así, si la proporción de varianza explicada por cada modo principal, Λ_k/N , es superior al estadístico g_k , entonces el modo principal correspondiente es seleccionado.

A partir de la selección de los n modos principales significativos ($n \leq N$), con lo que se reduce de forma eficiente el número de variables con las que continuar el análisis, se realiza la rotación-varimax para conferir un mayor sentido físico a los modos en los que se descompone la varianza total. Como

resultado de la rotación-varimax resulta un número n de modos de variación, cada uno de los cuales representa un modo de variación común a un número de celdas, en las que se explica la mayor parte de su varianza. De esa forma, los modos permiten conceptualmente agrupar en n grupos las celdas de la red $\Omega_{\#}$, la mayor parte de las celdas pertenecerán sólo a uno de los grupos, y pueden existir un número reducido de celdas que pertenezcan a varios grupos (celdas frontera). Las componentes por cada uno de los modos de variación rotados resultan de la ec. 7.22:

$$[H^n]_{t,\Omega} = [Z]_{t,\Omega} \cdot [\mathcal{U}_{:,1:n}] \cdot [R^n] \quad (7.22)$$

donde $[R^n]$ representa la matriz de rotación ortonormal que maximiza el criterio varimax, $[\mathcal{U}_{:,1:n}]$ representa la matriz formada por las n primeras columnas de la matriz de vectores principales $[\mathcal{U}]_{\Omega}$ y $[H^n]_{t,\Omega}$ representa las componentes de los modos de variación en Ω obtenidos tras la rotación de los n modos principales significativos. Cada uno de los modos rotados tendrá una determinada varianza, y por no ser principales, pueden mantener en general cierta correlación con el resto de modos, resultando que la matriz de varianzas-covarianzas de los modos de variación rotados sigue:

$$[S^n] = [R^n]^T \cdot [\Lambda_{1:n}] \cdot [R^n] \quad (7.23)$$

donde $[\Lambda_{1:n}]$ representa la matriz diagonal formada por los valores principales de los n modos principales significativos, y $[S^n]$ la matriz de varianzas-covarianzas de estos n modos rotados. Por tratarse de una rotación ortonormal, la producida en ec. 7.23, las trazas de las matrices varianzas-covarianzas coinciden, $\text{traza}([S^n]) = \sum_{j=1}^n \Lambda_j$. Dado que los modos principales significa-

tivos explican una alta proporción de la varianza total de $[Z]_{t,\Omega}$, $\sum_{j=1}^n \Lambda_j \approx$

$\sum_{j=1}^N \Lambda_j = N$, los modos principales rotados explicarán esta misma proporción.

De acuerdo a la función objetivo utilizada para obtener la rotación-varimax (ec. 7.20), tras la rotación, el patrón de variación en la mayoría de las celdas vendrá explicado principalmente por un único modo de variación, existiendo un menor número de celdas en las que su variación sea combinación de varios modos. Como la varianza en cada celda es unidad, la varianza de cada modo principal rotado, σ_{i,H^n}^2 $i = 1, \dots, n$, está relacionada con el número de celdas en las que predomina ese modo de variación. Esto permite asignar una porción del área total de la región Ω a cada modo. Así mismo las componentes en cada uno de los modos principales rotados tienen un sentido físico, indicando una desviación en las componentes de un modo hacia $-\infty$, una desviación hacia la región extremal de las realizaciones producidas

en las celdas a las que representa. Sin embargo, las componentes de cada uno de los modos principales rotados no resultan en general independientes. Atendiendo a la relación lineal que se establece entre las distintas variables, la ec. 7.24 permite escribir cada componente $H_{t,q}^n$ como combinación lineal de las componentes producidas en los modos anteriores $k = 1, \dots, q - 1$, más un término independiente con distribución normal $\varepsilon_{t,q}^n = N(0, \sigma_{\varepsilon_q}^2)$, donde $\sigma_{\varepsilon_q}^2$ proviene de ec. 7.26, para $q = 2, \dots, n$, y $\sigma_{\varepsilon_1}^2 = \sigma_{1,H^n}^2$.

$$\begin{bmatrix} H_{t,1}^n \\ H_{t,2}^n \\ H_{t,3}^n \\ \vdots \\ H_{t,n}^n \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \beta_{2,1} & 0 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \beta_{n-1,1} & \beta_{n-1,2} & \ddots & 0 & 0 \\ \beta_{n,1} & \beta_{n,2} & \cdots & \beta_{n,n-1} & 0 \end{bmatrix} \cdot \begin{bmatrix} H_{t,1}^n \\ H_{t,2}^n \\ H_{t,3}^n \\ \vdots \\ H_{t,n}^n \end{bmatrix} + \begin{bmatrix} \varepsilon_{t,1}^n \\ \varepsilon_{t,2}^n \\ \varepsilon_{t,3}^n \\ \vdots \\ \varepsilon_{t,n}^n \end{bmatrix} \quad (7.24)$$

con

$$\begin{bmatrix} \sigma_1^2 & \gamma_{1,2} & \cdots & \gamma_{1,q-1} \\ \gamma_{1,2} & \sigma_2^2 & \cdots & \gamma_{2,q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1,q-1} & \gamma_{2,q-1} & \cdots & \sigma_{q-1}^2 \end{bmatrix}_{H^n} \cdot \begin{bmatrix} \beta_{q,1} \\ \beta_{q,2} \\ \vdots \\ \beta_{q,q-1} \end{bmatrix} = \begin{bmatrix} \gamma_{1,q} \\ \gamma_{2,q} \\ \vdots \\ \gamma_{q-1,q} \end{bmatrix}_{H^n} \quad \text{para } q = 2, \dots, n \quad (7.25)$$

$$\sigma_{q,\varepsilon^n}^2 = \begin{bmatrix} -\beta_{q,1} & \cdots & -\beta_{q,q-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} \sigma_1^2 & \gamma_{1,2} & \cdots & \gamma_{1,q} \\ \gamma_{1,2} & \sigma_2^2 & \cdots & \gamma_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1,q} & \gamma_{2,q} & \cdots & \sigma_q^2 \end{bmatrix}_{H^n} \cdot \begin{bmatrix} -\beta_{q,1} \\ \vdots \\ -\beta_{q,q-1} \\ 1 \end{bmatrix} \quad (7.26)$$

De aquí resulta que las componentes de los n modos principales rotados producidos en cada instante t pueden ser expresadas como combinación lineal de n variables aleatorias independientes, siguiendo ec. 7.27. Debe notarse que la expresión de las componentes de los modos principales rotados como combinación lineal de variables independiente también podría haberse conseguido expresándolos en función de las componentes de los modos principales (ec. 7.22 y 7.13). Sin embargo, mientras las componentes de los modos principales no mantienen en general un sentido físico adecuado para la caracterización de sequías en cada celda, la expresión en términos de $\varepsilon_{t,q}^n$ (ec. 7.27) sí lo mantiene.

$$\begin{bmatrix} H_{t,1}^n \\ H_{t,2}^n \\ \vdots \\ H_{t,n-1}^n \\ H_{t,n}^n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ \alpha_{2,1} & 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha_{n-1,1} & \alpha_{n-1,2} & \ddots & 1 & 0 \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_{t,1}^n \\ \varepsilon_{t,2}^n \\ \vdots \\ \varepsilon_{t,n-1}^n \\ \varepsilon_{t,n}^n \end{bmatrix} \quad (7.27)$$

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ \alpha_{2,1} & 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha_{n-1,1} & \alpha_{n-1,2} & \ddots & 1 & 0 \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\beta_{2,1} & 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ -\beta_{n-1,1} & -\beta_{n-1,2} & \ddots & 1 & 0 \\ -\beta_{n,1} & -\beta_{n,2} & \cdots & -\beta_{n,n-1} & 1 \end{bmatrix}^{-1} \quad (7.28)$$

La ec. 7.27 permite descomponer la varianza total de $[H^n]_{t,\Omega}$ como suma de la varianza producida por las variables independientes $[\varepsilon^n]_{t,\Omega}$:

$$\sum_{i=1}^n \sigma_{i,H^n}^2 = \sum_{i=1}^n \Theta_i \quad (7.29)$$

con

$$\Theta_i = \begin{cases} \left(1 + \sum_{j=i+1}^n \alpha_{j,i}^2 \right) \cdot \sigma_{i,\varepsilon^n}^2 & i < n \\ \sigma_{n,\varepsilon^n}^2 & i = n \end{cases} \quad (7.30)$$

La relación $\Theta_i / \sum_{j=1}^n \Theta_j$ indica la proporción de la varianza total en la región Ω producida por cada variable independiente $\varepsilon_{t,i}^n$, $i = 1, \dots, n$. A partir de esto, y aplicando el Teorema de Multiplicación derivado del Teorema de Bayes (Kolmogorov, 1956), se puede escribir la *f.a.e.c.* en el caso de una región Ω discretizada por una red regular de celdas $\Omega_{\#}$, todas ellas de la misma superficie, como:

$$F_{x,\Omega}(X_t) \approx \exp \left\{ \Theta_1 \cdot \ln [\mathcal{P}(h_1^n \leq H_{1,t}^n)] + \sum_{i=2}^n \Theta_i \cdot \ln [\mathcal{P}(h_i^n \leq H_{i,t}^n | h_k^n \leq H_{k,t}^n, k = 1, \dots, i-1)] \right\} \quad (7.31)$$

donde h_i^n , $i = 1, \dots, n$, representa la variable aleatoria componente del modo principal rotado i , que se distribuye según $h_i^n \in N(0, \sigma_{i,H^n}^2)$, y mantiene

una estructura de correlación con las componentes del resto de modos dada por $[S^n]$. La transformación logarítmica convierte la serie multiplicativa de ec. 5.65 en un sumatorio, representando el coeficiente multiplicador de cada sumando al porcentaje del área de la región en la que predomina cada modo.

Para evaluar las probabilidades condicionales en ec. 7.31, y haciendo uso de la expresión de las componentes en cada modo en función de las variables independientes, puede escribirse:

$$F_{x,\Omega}(X_t) \approx \exp \left\{ \Theta_1 \cdot \ln [\mathcal{P}(\varepsilon_1^n \leq \Upsilon_{t,1})] + \sum_{i=2}^n \Theta_i \cdot \ln [\mathcal{P}(\varepsilon_i^n \leq \Upsilon_{t,i} | h_k^n \leq H_{k,t}^n, k = 1, \dots, i-1)] \right\} \quad (7.32)$$

con

$$\begin{aligned} \Upsilon_{t,1} &= H_{t,1}^n \\ \Upsilon_{t,2} &= H_{t,2}^n - \alpha_{2,1} \cdot \varepsilon_{t,1}^n \\ \dots &= \dots \\ \Upsilon_{t,n} &= H_{t,n}^n - \alpha_{n,1} \cdot \varepsilon_{t,1}^n - \dots - \alpha_{n,n-1} \cdot \varepsilon_{t,n-1}^n \end{aligned} \quad (7.33)$$

Ec. 7.32 puede desarrollarse a través de las integrales múltiples que definen las probabilidades condicionales, en términos de variables aleatorias normales independientes:

$$\begin{aligned} F_{x,\Omega}(X_t) \approx \exp \left\{ (\Theta_1 - \Theta_2) \cdot \ln \left[\int_{-\infty}^{\Upsilon_{t,1}} \psi(\varepsilon_1^n, \sigma_{1,\varepsilon^n}^2) \cdot d\varepsilon_1^n \right] + \sum_{i=2}^{n-1} (\Theta_i - \Theta_{i+1}) \cdot \ln \left[\int_{-\infty}^{\Upsilon_{t,1}} \psi(\varepsilon_1^n, \sigma_{1,\varepsilon^n}^2) \cdot \dots \cdot \int_{-\infty}^{\Upsilon_{t,i}} \psi(\varepsilon_i^n, \sigma_{i,\varepsilon^n}^2) \cdot d\varepsilon_i^n \cdot \dots \cdot d\varepsilon_1^n \right] + \Theta_n \cdot \ln \left[\int_{-\infty}^{\Upsilon_{t,1}} \psi(\varepsilon_1^n, \sigma_{1,\varepsilon^n}^2) \cdot \dots \cdot \int_{-\infty}^{\Upsilon_{t,n}} \psi(\varepsilon_n^n, \sigma_{n,\varepsilon^n}^2) \cdot d\varepsilon_n^n \cdot \dots \cdot d\varepsilon_1^n \right] \right\} \quad (7.34) \end{aligned}$$

donde $\psi(x, \sigma)$ es la función de densidad de probabilidad de una variable aleatoria $N(0, \sigma^2)$.

La ec. 7.34 puede ser evaluada mediante la resolución numérica de las integrales múltiples que contiene. El Apéndice A contiene una descripción de los métodos numéricos disponibles para la estimación de estas integrales múltiples, y el método implementado en esta Tesis. Al producirse una secuencia de integrales múltiples tal que se van incluyendo nuevos integrandos,

manteniéndose los anteriores y sus límites, el coste computacional no resulta la suma de los costes de cada integral múltiple, sino aproximadamente el coste de evaluar la integral de mayor dimensión.

La ec. 7.34 permite evaluar la *f.a.e.c.* en cada instante t , en una región Ω , discretizada por una red de celdas $\Omega_{\#}$, en los casos en los que no pueda asumirse independencia entre las realizaciones de cada celda. Con lo cual puede ser caracterizadas las desviaciones persistentes extremas de la secuencia de $F_{x,\Omega}(X_t)$ a partir de la función *f.p.e.*, y estimar el período de retorno asociado a la caracterización. Ello se desarrolla en la siguiente sección, donde se expone los criterios para la elaboración de mapas de IFS que permitan no solo caracterizar la realización de una sequía en el tiempo, sino también su desarrollo en el espacio.

7.2. Caracterización estocástica de realizaciones espacio-temporales de sequías

A partir del uso de la *f.p.e.* para la caracterización temporal de un evento de sequía (Capítulo 5), y de la selección de un criterio para la caracterización del desarrollo espacial de un evento, es posible la realización de un análisis completo de caracterización espacio-temporal dando como principales resultados los mapas de IFS. El criterio de caracterización espacial de mayor consistencia dentro de un análisis probabilístico es el que proporciona la *f.a.e.c.* (sección 7.1). Este es el criterio aquí adoptado, con el objetivo de una caracterización general. Sin embargo cualquier otro criterio de tipo probabilístico puede ser utilizado, variando únicamente la metodología en la función a evaluar.

La metodología que se presenta permite la conjunción práctica de los resultados anteriores para conseguir la caracterización de la situación de sequía en cada punto de una región e instante de tiempo, adoptando como principios para el análisis, las calidades de:

- **Retrospectivo:** en cada instante se busca aquella secuencia terminada entonces para la que la desviación resulte mayor.
- **Conexo:** en cada punto se busca la agrupación conexas de puntos que lo incluye para la que desviación sea máxima

Ello se concreta en la siguiente definición del valor que caracteriza, en cada instante de tiempo y en cada punto del espacio de la región analizada, la desviación persistente de las realizaciones de una variable hidro-climática, utilizada como indicador, hacia la región extremal asociada a la situación de sequía:

Para cada instante de tiempo y en cada punto del espacio se caracteriza su situación de sequía con el valor la *f.p.e.* obtenido de la secuencia terminada en ese instante de la *f.a.e.c.* evaluada sobre la agrupación conexas que contiene al punto de análisis y se sitúa dentro de la región de estudio, cuya longitud de secuencia, posición y tamaño de la agrupación maximizan la *f.p.e.*

Esta caracterización puede ser expresada en términos de la estimación asintótica del período de retorno asociado (ec. 5.64) resultando como producto el mapa de IFS.

La obtención del valor de la *f.p.e.* que caracteriza en cada instante de tiempo y en cada punto del espacio de una región la situación de sequía requiere el cómputo de las múltiples combinaciones posibles entre las que se puede producir este máximo. Cuando el número de puntos de la red que discretizan la región de análisis sea significativo, el número de combinaciones que se pueden producir motiva la utilización de un algoritmo de búsqueda del máximo, que no requiera la evaluación de todas las combinaciones posibles.

A continuación se presenta un algoritmo que permite la búsqueda de estos máximos para cada instante de tiempo y en cada punto del espacio de una región, facilitando el cálculo de los mapas de IFS.

7.2.1. Algoritmo para la construcción de los mapas de IFS

Sea una región espacial Ω , sobre la que se ha producido la realización espacial de la variable aleatoria x , representada por un conjunto de series de realizaciones $[X]_{t,\Omega}$, de longitud L , definidas sobre una red de N celdas que constituyen una discretización regular de la región $\Omega_{\#}$. Cada una de estas series puede ser transformada a una serie de una variable normal estándar equivalente, $Z_{t,k} = \Phi^{-1} [F_{k,\Omega_{\#}}(X_{t,k})]$, $k = 1, \dots, N$, resultando la matriz $[Z]_{t,\Omega}$ la matriz de las transformadas de las realizaciones. Si la variable aleatoria x presenta estacionalidad, deberá realizarse así mismo la transformación pertinente hasta conseguir una variable estacionaria.

A partir de la matriz $[Z]_{t,\Omega}$ puede ser calculada, de forma independiente para cada celda, las series de IFS, $[IFS]_{t,\Omega}$, entendiéndose aquí expresado el IFS en términos de la *f.p.e.* asociada (sin realizar su transformación hacia la estimación asintótica de su período de retorno). Estas series analizan las desviaciones persistentes extremas hacia la región extremal producida en cada celda de la discretización. El símbolo \cdot será utilizado como indicador del carácter puntual del valor, mientras que el símbolo \circ indicará el carácter espacial del valor.

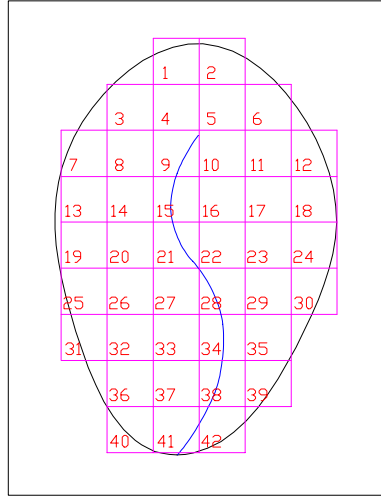


Figura 7.1: Discretización regular de una cuenca donde realizar el cálculo del mapa de IFS.

A partir de los valores puntuales, para realizar la construcción del mapa de $[IFS^o]_{t,\Omega}$, en cada instante t , se seguirán los siguientes pasos:

1. La celda donde se produzca el mayor valor de los IFS_t^o en el instante t , se convertirá en el primer germen de agrupación \mathcal{A}^{-1-} , donde el valor del mapa de IFS en la celda que lo forma coincide con su valor puntual, $\mathcal{A}_{(1)}^{-1-}.IFS_t^o = IFS_t^o(\mathcal{A}_{(1)}^{-1-}.k)$, siendo $\mathcal{A}_{(1)}^{-1-}.k$ el ordinal que define la celda dentro de la red $\Omega_{\#}$.
2. Entre las celdas vecinas a la agrupación \mathcal{A}^{-1-} se selecciona como candidata a unirse a la agrupación aquella con el valor puntual de IFS mayor: C^{-1-} . Para ella, se calcula el valor de $C^{-1-}.IFS_t^o$ que se produce como caracterización espacio-temporal de la sub-región formada por la celda que componen \mathcal{A}^{-1-} junto con la celda candidata C^{-1-} , $C^{-1-}.IFS_t^o = \{\mathcal{A}^{-1-} \cup C^{-1-}\}.IFS_t^o$. Así mismo se busca entre todas las celdas que no forman parte de agrupaciones, aquella con mayor IFS_t^o , la cual constituirá una semilla de nueva agrupación S , con $S.IFS_t^o = IFS_t^o(S.k)$.
3. Si $C^{-1-}.IFS_t^o > S.IFS_t^o$ entonces la agrupación \mathcal{A}^{-1-} se amplía, incluyendo la celda candidata, $\mathcal{A}_{(n^{-1-}+1)}^{-1-} = C^{-1-}$, $n^{-1-} = n^{-1-} + 1$, siendo n^{-1-} el número de celdas que componen la agrupación \mathcal{A}^{-1-} , y con ella se calcula la nueva celda candidata de la agrupación C^{-1-} . Si no, entonces se forma una nueva agrupación formada por la semilla $\mathcal{A}^{-2-} = S$, siendo $m = 2$ el número de agrupaciones, y se calcula la nueva celda semilla S .

4. Siguiendo de forma recurrente, entre las m celdas candidatas a ampliar las agrupaciones existentes y la semilla de agrupación, se busca aquella con mayor IFS_t^o . Si el máximo se produce en una celda candidata, $C^{-i-}.IFS_t^o = \max\{C^{-1-}.IFS_t^o, \dots, C^{-m-}.IFS_t^o, S.IFS_t^o\}$, y esta celda $C^{-i-}.k$ no es candidata de otra agrupación, entonces la agrupación \mathcal{A}^{-i-} se amplía incluyendo la celda candidata, $\mathcal{A}_{(n^{-i-}+1)}^{-i-} = C^{-i-}$, $n^{-i-} = n^{-i-} + 1$, calculándose la nueva celda candidata de la agrupación C^{-i-} . Si el máximo se produce con la celda semilla $S.IFS_t^o = \max\{C^{-1-}.IFS_t^o, \dots, C^{-m-}.IFS_t^o, S.IFS_t^o\}$, entonces se crea una nueva agrupación formada por la celda semilla, $\mathcal{A}^{-m+1-} = S$, y $m = m + 1$, calculándose la nueva celda semilla no incluida en ninguna agrupación existente.

Si siendo $C^{-i-}.IFS_t^o = \max\{C^{-1-}.IFS_t^o, \dots, C^{-m-}.IFS_t^o, S.IFS_t^o\}$, resulta que la celda $C^{-i-}.k$ es también candidata de otra agrupación, $C^{-i-}.k = C^{-j-}.k$, entonces se calcula el IFS_t^o de la sub-región formada por las agrupaciones \mathcal{A}^{-i-} y \mathcal{A}^{-j-} , y la celda candidata C^{-i-} . Si $\{\mathcal{A}^{-i-} \cup \mathcal{A}^{-j-} \cup C^{-i-}\}.IFS_t^o > C^{-i-}.IFS_t^o$ entonces se forma una nueva agrupación unión de las anteriores $\mathcal{A}^{-m+1-} = \{\mathcal{A}^{-i-} \cup \mathcal{A}^{-j-} \cup C^{-i-}\}$, $m = m + 1$, y se calcula la celda candidata de esta nueva agrupación.

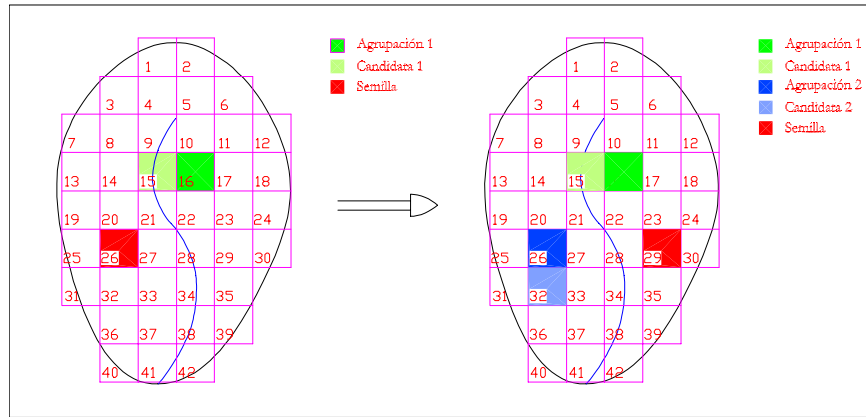


Figura 7.2: Evolución de las agrupaciones en el caso de $C^{-1-}.IFS_t^o > S.IFS_t^o$, ampliándose \mathcal{A}^{-1-} .

De este modo van siendo completados los valores del mapa de IFS, caracterizando cada celda por el valor de la *f.p.e.*, IFS_t^o , correspondiente a la agrupación de entre las que la incluye y la secuencia que producen una mayor desviación persistente hacia la región extremal. Con ello, a cada celda le

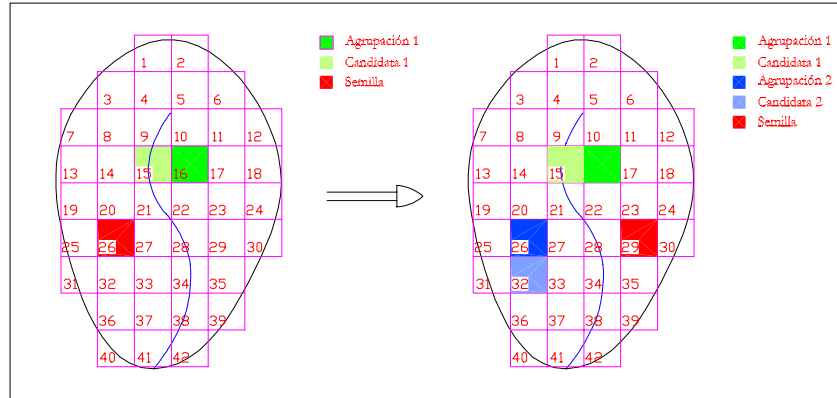


Figura 7.3: Evolución de las agrupaciones en el caso de $S.IFS^o_t > C^{-1-}.IFS^o_t$, creándose una nueva agrupación \mathcal{A}^{-1-} .

es asignado un valor diferente, apareciendo mapas con uno o varios máximos del IFS, disminuyendo el índice a medida que aumenta la distancia al centro donde se concentra la realización extrema. El mapa de índices, puede ser expresado en términos de la *f.p.e.*, o preferiblemente transformado mediante ec. 5.64 hacia la estimación del período de retorno asociado (para lo cual es necesario conocer en cada celda adicionalmente la longitud de la secuencia que produce el máximo en *f.p.e.*).

El algoritmo aquí presentado no asegura la obtención exacta en cada celda del valor que define su caracterización espacio-temporal de realización de sequías. Sin embargo, aunque éste no sea conseguido de forma rigurosa, el resultado de implementar el algoritmo puede usarse como una aproximación suficiente a la definición exacta, con un menor coste computacional. En el capítulo siguiente se muestra ejemplo de aplicación práctica del mapa de IFS de las precipitaciones anuales, en la cuenca del río Colorado (EEUU).

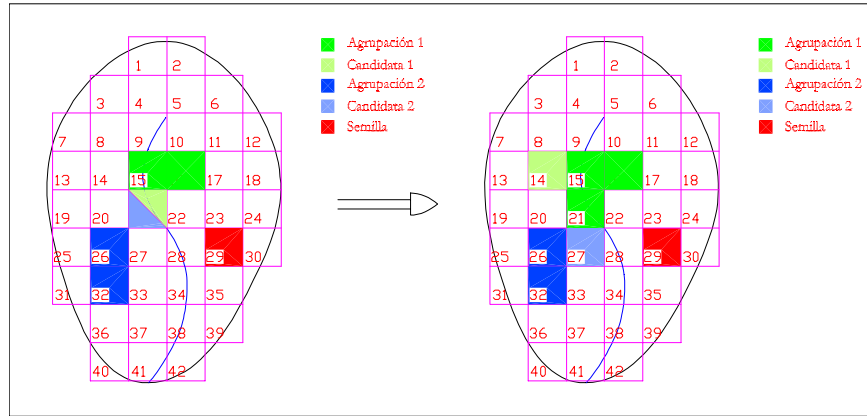


Figura 7.4: Evolución de las agrupaciones, con $C^{-1-}.IFS_t^o > C^{-2-}.IFS_t^o$ y $S.IFS_t^o$, y siendo $\{\mathcal{A}^{-1-} \cup \mathcal{A}^{-2-} \cup C^{-i-}\}.IFS_t^o < C^{-1-}.IFS_t^o$. No se produce la unión, ampliándose \mathcal{A}^{-1-} .

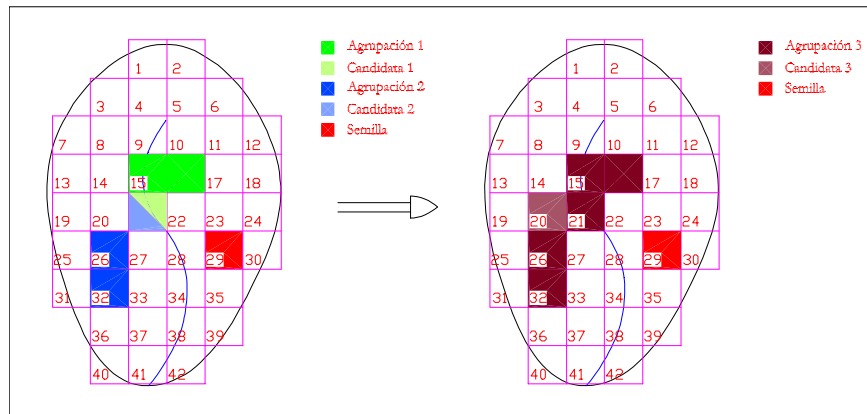


Figura 7.5: Evolución de las agrupaciones, con $C^{-1-}.IFS_t^o > C^{-2-}.IFS_t^o$ y $S.IFS_t^o$, y siendo $\{\mathcal{A}^{-1-} \cup \mathcal{A}^{-2-} \cup C^{-i-}\}.IFS_t^o > C^{-1-}.IFS_t^o$. Se produce la unión de las agrupaciones \mathcal{A}^{-1-} , \mathcal{A}^{-2-} y la celda C^{-1-} , resultando la agrupación \mathcal{A}^{-3-} .

Capítulo 8

Aplicaciones

8.1. Herramienta para el control espacio-temporal de sequías

8.1.1. El control de las sequías en la cuenca del río Colorado (EEUU)

La cuenca del río Colorado se encuentra en el sur-oeste de EEUU, cubriendo una superficie de más de $630,000\text{km}^2$, y extendiéndose sobre 7 estados: Wyoming, Nevada, UTAH, Colorado, California, Arizona y Nuevo México (figura 8.1).

La precipitación media anual sobre la cuenca varía desde la zona de alta montaña de cabecera, con unos valores máximos superiores a 1100mm/año , hasta la zona suroeste de la cuenca, con una precipitación anual en algunos puntos inferior a los 100mm/año (figura 8.2). La alta variabilidad en la precipitación media es un reflejo también de la importante variabilidad climatológica existente, desde el clima húmedo en el noreste hasta el clima desértico árido en el suroeste, próximo a su desembocadura.

La elevada variabilidad climática que se produce en la cuenca del río Colorado la hacen de especial interés para la aplicación sobre la cuenca de un análisis espacial de sequías, que permita la caracterización, para cada punto de la cuenca y cada instante, de la situación de sequedad en la que se encuentre. El uso de la metodología para la construcción de mapas de IFS, con un fundamento meramente probabilístico, permite la comparación entre cualquier par de puntos del espacio, en una escala independiente del régimen de variación del punto analizado. Los mapas de IFS sobre la cuenca proporcionan un análisis del desarrollo temporal y espacial de las sequías producidas. El análisis que a continuación se presenta es realizado sobre las precipitaciones anuales en una red uniforme que discretiza la cuenca. Así



Figura 8.1: Cuenca del río Colorado y río Bravo, y su intersección con los estados de EE.UU. y México.

el análisis se referirá a las sequías meteorológicas producidas sobre la misma. En este sentido cabe destacar que una importante sequía meteorológica tendrá su reflejo en el régimen hidrológico. Sin embargo, la significancia sobre la serie de aportaciones hidrológicas dependerá en gran medida de la zona sobre la que se produce la sequía, en especial de la contribución de la precipitación de la zona a la aportación total. Como fuente de datos de precipitaciones para la realización del análisis espacio-temporal de sequías sobre la cuenca del río Colorado se hace uso del modelo PRISM de interpolación espacial de alturas de precipitación, que proporciona datos de precipitación media areal a escala mensual en una malla regular de 5×5 minutos, cubriendo la totalidad de EE.UU. A continuación se realiza una breve revisión de los métodos de análisis espacial y se describe en mayor detalle el modelo PRISM.

El modelo PRISM de interpolación espacial de alturas de precipitación

Históricamente pueden distinguirse dos categorías principales para clasificar los métodos utilizados en el análisis espacial de variables climáticas a partir de registros puntuales: los métodos geográficos y los métodos estadís-

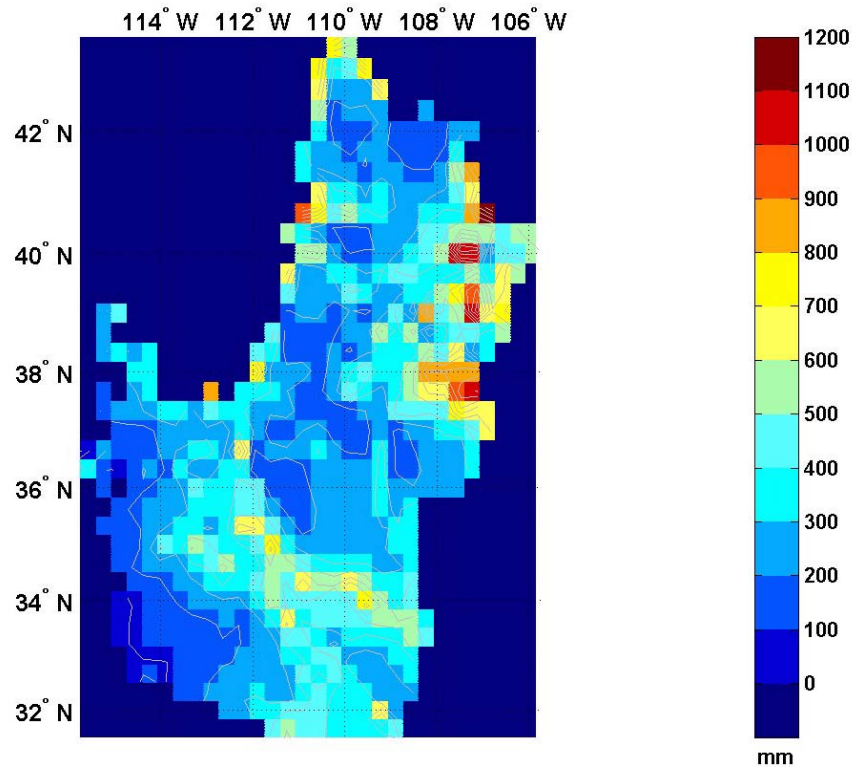


Figura 8.2: Mapa de precipitación anual media sobre la cuenca del río Colorado (EEUU), en el período de 1895 a 2003.

ticos. Los métodos geográficos dominan durante más de la primera mitad del siglo XX. Sobre ellos destaca la elaboración manual de mapas de precipitaciones, a partir de registros pluviométricos, buscando tener en cuenta la orografía de la zona (Thiessen, 1911; Peek & Brown, 1962). Estos mapas tenían en general un alto grado de subjetividad, estando apoyados en reglas y análisis simples, como el estudio de la correlación entre la precipitación y la altitud (Spren, 1947; Schermerhorn, 1967). A partir de los años setenta, con el uso de las computadoras y el aumento progresivo en la disponibilidad de información, comienzan a ser implementados métodos con un mayor carácter estadístico, como pueden ser algoritmos de interpolación ponderada basada en distancias, métodos de krigearje, métodos spline, o análisis multivariado.

Como método híbrido entre los métodos geográficos y estadísticos para la interpolación de precipitaciones aparece el modelo PRISM, (*Parameter-elevation Regressions on Independent Slopes Model*), (Daly & Neilson, 1992;

Daly *et al.*, 1994). El modelo PRISM hace uso de un modelo de elevación digital de la región de estudio, conjunto con los registros puntuales de precipitaciones, para analizar y tener en cuenta en la interpolación de la precipitación la orografía existente. PRISM no resuelve de forma sistemática un conjunto fijo de ecuaciones, sino que representa un modelo conceptual, un grupo de reglas, y herramienta de análisis que apoyan a la toma de decisiones en la creación de un mapa de precipitaciones por un experto climatólogo. El modelo conceptual está fundado en tres aspectos principales:

- El efecto de la elevación en la precipitación. Para ello adopta un modelo lineal como la relación más adecuada, de modo que en el modelo es asumido que la precipitación sobre una ladera crece linealmente con la altura hasta la cumbre, a no ser que la información de un registro puntual lo contradiga. Está hipótesis es válida generalmente en zonas de latitud media.
- La escala espacial de los efectos orográficos. La escala espacial de trabajo, definida como el tamaño de las celdas que discretizan la región espacial sobre las que se pretende interpolar su precipitación media, es de especial importancia en los análisis orográficos. La altitud de las estaciones donde son recogidos los registros puede presentar valores de correlación con la precipitación bastante bajos. Sin embargo, al asignarle a cada estación la altitud media de la celda dentro de la que se encuentra, se puede obtener mejores correlaciones en función de la escala espacial. Esto es consecuencia de que los efectos orográficos se manifiestan en un rango de escalas espaciales. El valor de la escala óptima de trabajo no está resuelto de forma universal, pero para el caso del modelo PRISM se adoptó un tamaño de celda del modelo de elevación digital de 5×5 minutos, como valor que mejor se adecua en análisis regionales de precipitaciones.
- Los patrones espaciales de regímenes orográficos de precipitación sobre topografía compleja. Las relaciones entre precipitación y altitud pueden variar significativamente de una ladera a otra por motivos de variaciones en la pendiente, la orientación, la existencia de barreras orográficas, etc. Ello motiva la necesidad de delimitar en la región de análisis sub-regiones topográficas donde la pendiente y la orientación es razonablemente constante, analizada esta en la escala espacial donde se manifiestan los efectos orográficos sobre la precipitación. Es en esta delimitación donde es especialmente importante el trabajo del experto climatólogo.

Con ello el modelo PRISM realiza la interpolación de precipitaciones en cada celda de acuerdo a los siguientes pasos:

1. Estima la elevación orográfica (referida a la escala espacial adoptada) de cada estación de precipitación haciendo uso de un modelo de elevación digital de 5×5 minutos latitud-longitud.
2. Delimita las regiones topográficas y les asigna las estaciones pluviométricas que se encuentra en su interior.
3. Para cada celda, PRISM estima la precipitación promedio mediante técnicas de ventana que relacionan linealmente la precipitación en la celda, con las precipitaciones registradas en las estaciones más próximas dentro de la sub-región topográfica.

De este modo, el modelo ajusta continuamente la relación precipitación-orografía integrando la información orográfica y el conjunto de registros pluviométricos. Además, el modelo permite la interpolación de otras variables climáticas como la temperatura, al igual que lo hace con la precipitación. El modelo PRISM ha sido implementado para la totalidad de EEUU continental proporcionando una base de datos de precipitaciones y temperaturas interpoladas en una red regular de 5×5 minutos, a escala mensual, desde 1895 hasta la actualidad (Daly *et al.* , 2002).

Red de monitoreo de precipitaciones anuales en el río Colorado

Partiendo de la red original en la que se encuentra disponible los datos del modelo PRISM, resulta que la extensión de la cuenca del río Colorado queda cubierta por una malla uniforme de 292×243 celdas, cada una de las cuales de tamaño 5×5 minutos latitud-longitud. Con esta discretización, el número de celdas situadas en el interior de la cuenca asciende a unas 39,000 celdas. Al resultar este número de celdas impracticable desde un punto de vista computacional, la discretización es reducida con un factor de 8, de modo que se obtiene una malla de 36×30 celdas, cada una de las cuales de tamaño 40×40 minutos latitud-longitud. Para esta red, el número de celdas que cubren la cuenca se sitúa en 593 celdas, que resulta admisible desde el punto de vista de tiempos de cálculo, y mantiene una distribución representativa de todas las climatologías disponibles en la cuenca, con un área cubierta por cada celda del orden de los $1000km^2$. La longitud de la serie que será analizada es de 109 años, desde 1895 a 2003.

El número de celdas que discretizan el área de estudio resulta de gran importancia, pues debe cumplir el compromiso de mantener una representación de la variabilidad climatológica espacial en la cuenca, y por otro lado se encuentra limitada por la capacidad de cálculo disponible, dado que el número de combinaciones de agrupaciones de celdas que deben ser analizadas crece de forma exponencial con el número total de celdas.

Mapas de IFS sobre la cuenca del río Colorado

Para la construcción de los mapas de IFS sobre la cuenca del río Colorado se sigue la metodología presentada en el Capítulo 7. En primer lugar se realiza un análisis de la distribución estadística de la precipitaciones media anual en cada celda. En la totalidad de las celdas la distribución gamma no es rechazada por el test estadístico de Kolmogorov-Smirnov, con un 5 % de significancia. Si se utiliza la distribución normal, el test Chi-cuadrado o el test Kolmogorov-Smirnov producen rechazo en un 25 % y 20 % de las celdas, respectivamente.

En segundo lugar se analiza en cada celda la correlación temporal sobre la serie de precipitaciones medias anuales. Aplicando el Criterio de Información de Akaike (Akaike, 1924), del análisis resulta que en un 96 % de las celdas no se encuentra significancia estadística de la correlación temporal, para los rezagos de 1 a 10 años, y para una nivel de significancia del 5 %. En el 4 % de celdas donde la correlación resulta significativa, se da únicamente significancia para el rezago de 1 año, y con un valor máximo del coeficiente de correlación de 0.35. Por este motivo, dado el alto coste que produciría en un análisis a gran escala de este tipo el tratamiento completo de una señal correlacionada temporalmente, y a la vista de la reducida significancia de la correlación temporal sobre las precipitaciones anuales en la cuenca, la hipótesis de independencia temporal de la variable en cada celda se extiende al total de la red.

A partir de este análisis preliminar, se sigue la metodología propuesta para la construcción de mapas de IFS, que se describe con detalle en el Capítulo 7, en la cuenca del río Colorado. Los mapas resultantes del análisis se incluyen en el Apéndice B. Los resultados de la aplicación se comentan a continuación.

Comportamiento espacial de las sequías Entre los principales episodios de sequías que se muestran en los mapas de IFS sobre la precipitación anual en la cuenca del río Colorado destacan los siguientes:

- La sequía producida entre 1901 a 1906, en la zona centro y norte de la cuenca, con una dirección de desarrollo siguiendo SO-NE, de extensión variable y dos núcleos principales, uno situado en la zona centro y el segundo en la zona norte. Por la posición que ocupa esta sequía, sobre regiones de importante contribución a las aportaciones totales del río Colorado, esta sequía meteorológica tiene su manifestación en una importante sequía hidrológica producida en el Colorado a principios de siglo XX.
- En el período de 1909 a 1916 se produce una sequía concentrada en la

zona norte de la cuenca, con una extensión y posición muy persistente. En la misma forma se produce un nuevo evento de menor duración de 1921 a 1923.

- En la década de los años 1930, se desarrolla la más importante sequía de EEUU producida en el siglo XX. El centro de la misma se sitúa especialmente sobre la zona centro del país, pero también se desarrolla sobre la cuenca del Colorado. Destacan entorno a este período los mapas de los años 1928 a 1930 donde se produce la realización de una sequía en el suroeste de la cuenca próximo a la desembocadura.
- Durante la década de 1950, también se desarrolla un importante episodio en el sur y suroeste de EEUU, centrada sobre la zona centro del país. Su realización se manifiesta en los mapas de los años de 1956 a 1957, con una intensa sequía en el bajo Colorado, que en el año 1958 se desplaza en dirección norte.
- En el período de 1963 a 1965 se repite de nuevo un episodio de sequía en la zona centro y norte de la cuenca, con una dirección de desarrollo siguiendo SO-NE.
- Entre 1974 a 1977 se sitúa una sequía similar a la producida en el período 1909 a 1916, pero posicionada algo más al sur.
- En el período de 1986 a 1990 tiene lugar un nuevo evento de sequía en la zona centro y norte de la cuenca, coincidente con una de las más importantes sequías hidrológicas producidas en el siglo XX en la cuenca.
- Durante el período de 1998 a 2000 se concentra una sequía sobre la zona noreste de la cuenca, que ha producido un reflejo importante sobre las aportaciones al río Colorado. Esta sequía se desplaza en la dirección suroeste, manifestándose con mayor intensidad en la zona sur de la cuenca en 2002.

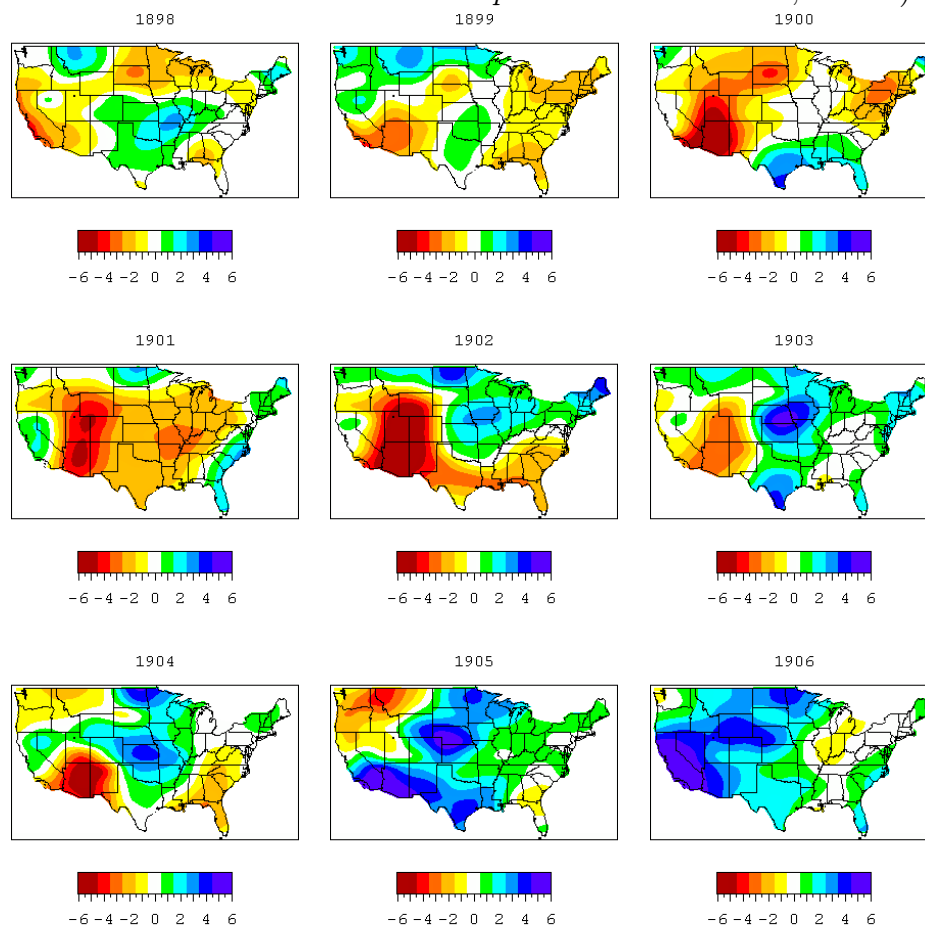
Cabe señalar un patrón de desarrollo de las sequías que se repite con mucha frecuencia sobre los mapas obtenidos. Este es el desarrollo alineado siguiendo la dirección SO-NE, a lo largo de la cual se producen muchos de los episodios de sequías. Esto mantiene relación con los frentes de humedad que alimentan las precipitaciones sobre la cuenca, que proceden principalmente del Océano Pacífico. En este sentido están siendo desarrollados un gran número de trabajos relacionados con la teleconexión existentes entre las precipitaciones o aportaciones en el río Colorado y los índices de circulación atmosférica de gran escala sobre el Océano Pacífico (Redmond & Koch, 1991; Dracup & Kahya, 1994; Kanya & Dracup, 1993).

Comparación con otros sistemas de control

Con el objetivo de comparar los resultados que se obtienen mediante el control de las sequías sobre la cuenca del río Colorado a partir de los mapas de IFS, se presentan a continuación los mapas que se obtienen cuando es utilizado el índice PDSI, que es el indicador de sequía más utilizado para estudiar la realización de eventos sobre EEUU. Por otro lado, el PDSI es un indicador de sequía meteorológica, el cual utiliza como principal información la pluviométrica, aunque también depende de la temperatura y la capacidad de almacenamiento medio de humedad de la cuenca. En este sentido, resulta adecuado comparar mapas de PDSI y mapas de IFS realizados sobre registros de precipitación. En primer lugar se incluyen los mapas de PDSI elaborados a escala anual en los períodos en los que se manifiestan sequías significativas que cubren la cuenca de estudio. Estos mapas son realizados con una red de 2×3 grados, latitud-longitud.

Analizando el caso de la sequía que se manifiesta en la cuenca a principios del siglo XX, en ambos tipos de mapas se refleja la realización de la sequía, siguiendo la misma orientación SO-NE. En el caso de los mapas PDSI la sequía aparece distribuida a lo largo de toda la cuenca. Sin embargo, en los mapas de IFS la sequía está más concentrada en la zona centro y el cuadrante nororiental de la cuenca, precisamente coincidente con las zonas de mayor pluviosidad media. El hecho de que no sea detectada de igual manera la distribución de la sequía por el PDSI viene motivado en gran medida por la discretización más gruesa que utilizan sus mapas, pero también por el sesgo que en general se pone de manifiesto en el índice cuando se comparan regiones húmedas y áridas o semiáridas.

Mapas de PDSI anual sobre EEUU en el período 1898 a 1906
(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



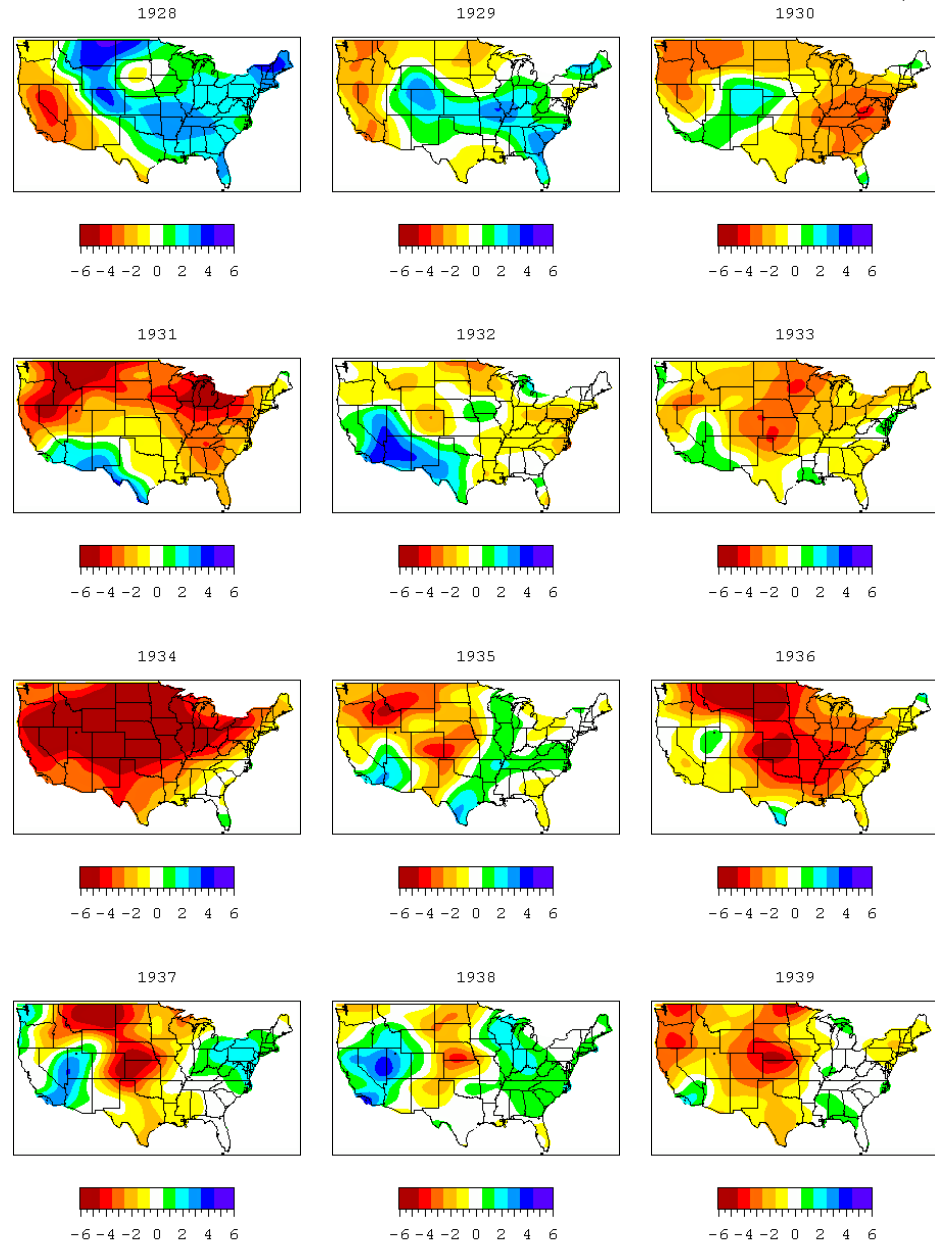
En el caso de la sequía producida sobre el suroeste de la cuenca en el período 1928 a 1930, ambos mapas coinciden en señalarla, si bien en el caso del PDSI da por terminada la sequía en el año 1930, mientras que en los mapas de IFS en ese año aun aparece una región con desviaciones significativa. Esta diferencia únicamente depende de la propia definición del índice, que en el caso del PDSI la definición del estado de sequía se realiza a partir de un balance de humedad, por lo que se alcanza suficiente precipitación en un año puede salirse inmediatamente de la situación de sequía. En el caso del IFS no se parte de una definición de sequía, sino que se cuantifica la desviación producida en los últimos años más desfavorable, manifestándose en una escala de aproximación asintótica a la frecuencia media de recurrencia. En este sentido, no hay un inicio y fin de sequía prefijado.

El diferente tratamiento de las sequías con cada índice puede provocar diferencias en la representación del desarrollo de las sequías como las que

se observa en la realización sobre la cuenca del Colorado de la sequía más importante de EEUU ocurrida en el siglo XX (denominada, *Dust Bowl*) (Schubert *et al.* , 2004). Según los mapas PDSI, la cuenca se encuentra en un zona fronteriza de la zona de desarrollo del episodio, a excepción del año 1934, donde la sequía alcanza su mayor intensidad y extensión, cubriendo el núcleo de la sequía la zona alta de la cuenca. Si bien este año es en el que se desarrolla una mayor intensidad, un aspecto que destaca del análisis de los mapas PDSI es el rápido cambio que se produce en la extensión e intensidad de las sequías, notorio en la secuencia presentada. Cabría esperar que la realización de una extraordinaria sequía como la producida en la década de 1930 produjera en el indicador una zona sometida a este estado de forma persistente, donde se concentran los efectos de la sequía, y no una variación intermitente.

El uso de los mapas de IFS no manifiestan una realización extendida a lo largo de toda la cuenca durante ese período, sino que los efectos se centran principalmente en zonas próximas a la divisoria de la mitad norte de la cuenca, con una presentación persistente.

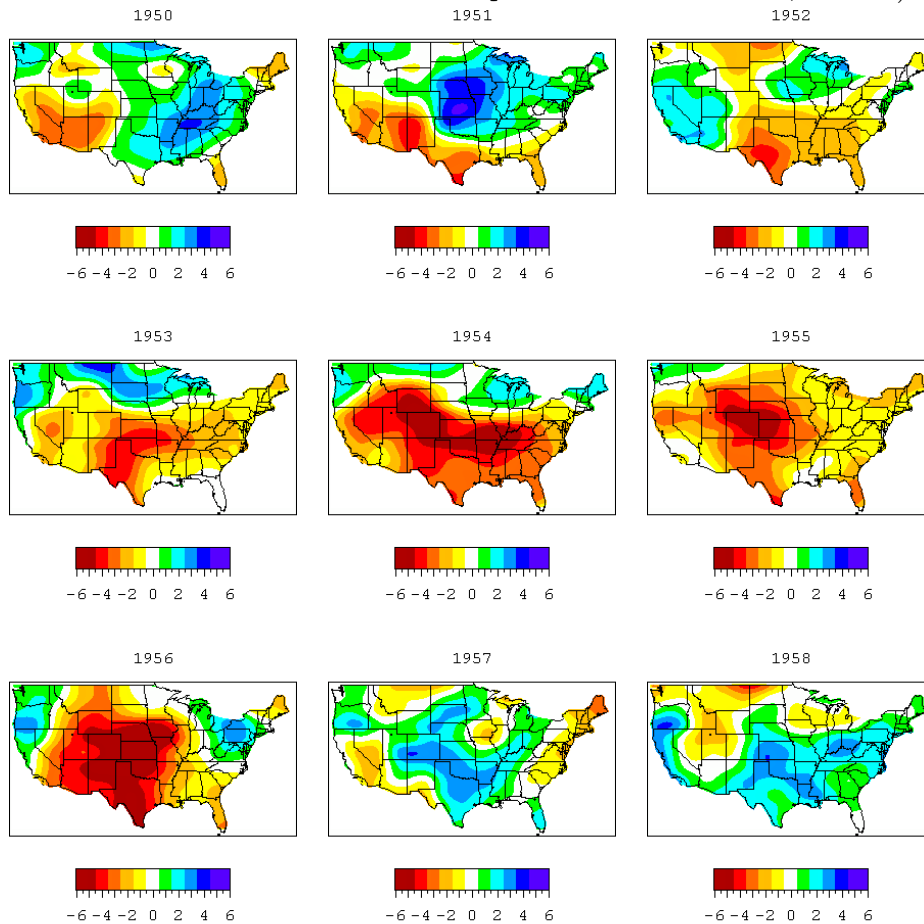
Mapas de PDSI anual sobre EEUU en el período 1928 a 1940
(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



Durante la década de 1950 se produce un episodio de sequía, cuya mayor intensidad se manifiesta en 1956, como coinciden en representar los mapas de PDSI y IFS. Sin embargo, en el caso de los mapas de PDSI se produce un cambio brusco en el sentido del índice al siguiente año, pasando en algunas zonas del rango de PDSI más seco al más húmedo. En el caso del IFS la variación es más gradual, y se puede observar el desplazamiento de la sequía hacia el norte. Esto resulta más difícil su apreciación sobre los mapas de

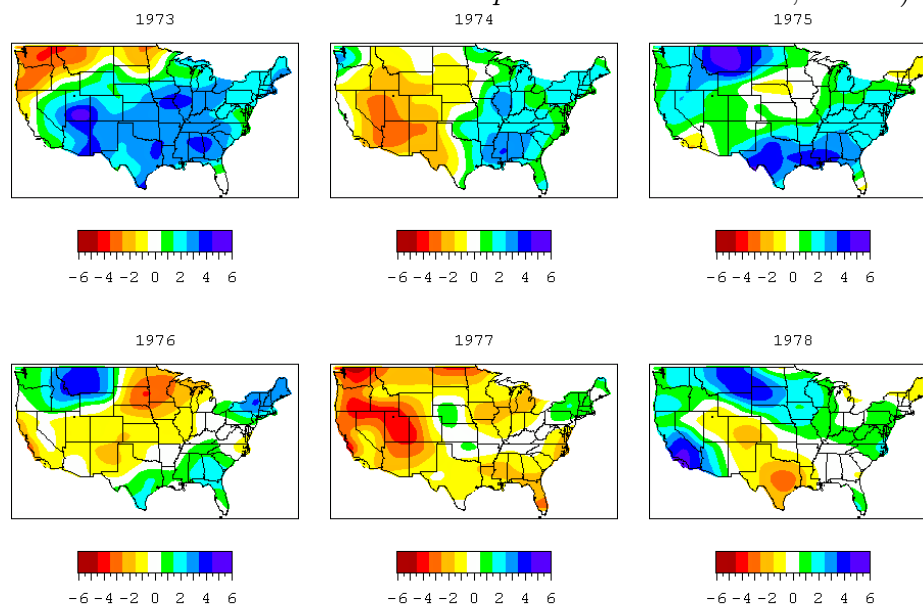
PDSI.

Mapas de PDSI anual sobre EEUU en el período 1950 a 1958
(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



Por último se incluyen los gráficos relativos al episodio de sequía producido durante la década de 1970, que se manifiesta principalmente durante los años de 1974 a 1977, de acuerdo con los mapas de IFS, ocupando el alto Colorado, con una zona centro posicionada siempre en la misma posición. En el caso de los mapas de PDSI se observa nuevamente la alternancia del estado de sequía durante los mapas de este período, estando ausente la percepción de persistencia del fenómeno.

Mapas de PDSI anual sobre EEUU en el período 1973 a 1978
(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



Otros tipos de mapas que son utilizados para el control de sequías a partir del PDSI son los que hacen uso del índice calculado a escala mensual, y en este caso por divisiones climáticas sobre EEUU continental. Para realizar la comparación con los mapas de IFS se incluyen cuatro de los mapas mensuales, por cada año, durante el período 1999 a 2002.

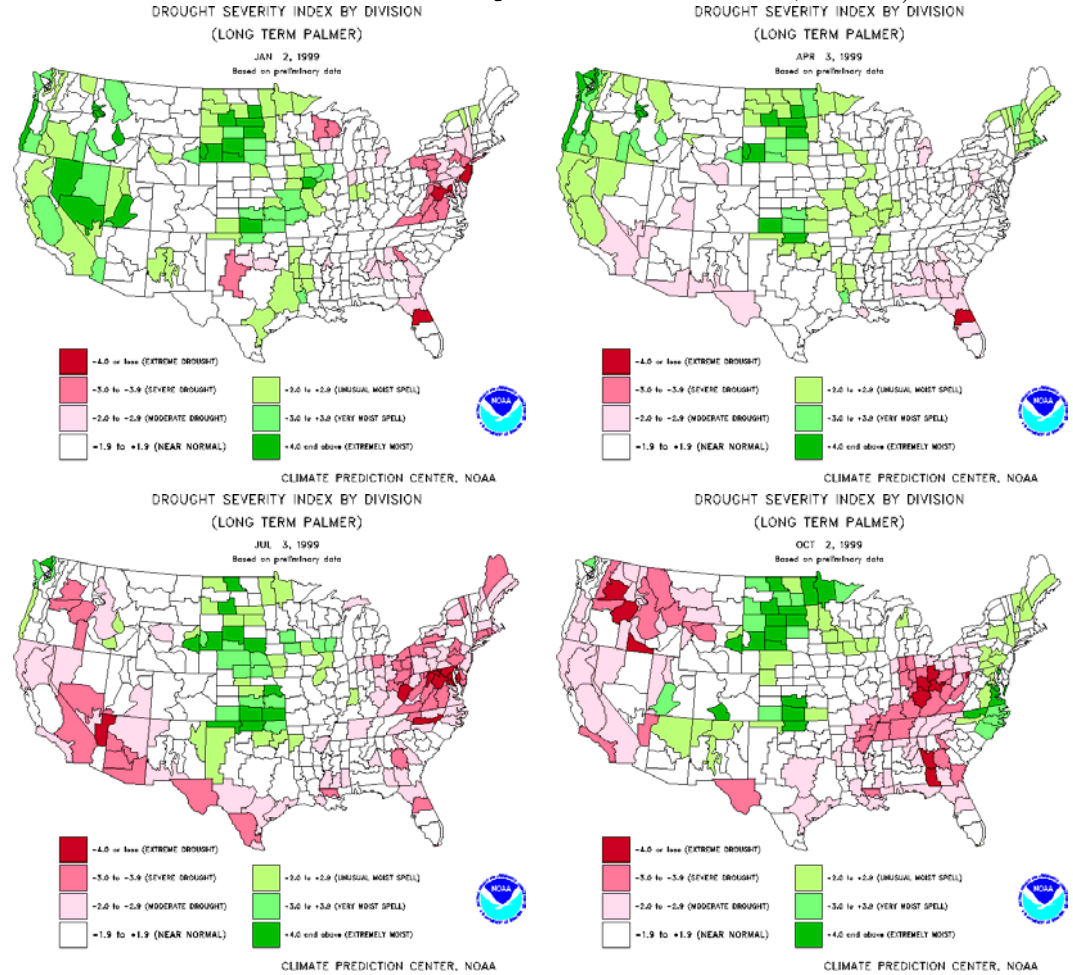
De acuerdo con los mapas de IFS se observa la realización de una sequía en la cuenca del Colorado durante el período 1998 a 2000, centrada en la zona noreste y centro de la cuenca, acompañada de un evento de menor significancia durante ese período en el suroeste, pero que alcanza mayor intensidad durante 2002.

Cuando se analiza la representación de estos eventos sobre los mapas PDSI mensuales destaca en primer lugar que la ya mencionada alternancia del signo del índice en los mapas PDSI anuales también se produce en los mapas PDSI mensuales, pero esta vez con evoluciones más cortas de tiempo, variando a lo largo del ciclo anual el signo de índice. Por otro lado, destaca también que el evento producido en la zona noreste y centro de la cuenca, coincidente con zonas de clima húmedo, no se pone de manifiesto con la misma significancia por el PDSI. Por ejemplo, los mapas PDSI indican en el año 1999 la situación de sequía en esta zona durante algunos meses y con menor severidad de como la cuantifica en el sureste. En el 2000 sí se indica el mayor nivel de severidad de sequía en la zona centro y noreste, pero sólo durante algunos meses. El episodio que ambos tipos de mapa coinciden

en señalar es el episodio producido en el suroeste, cuyo mayor efecto se concentra en el año 2002.

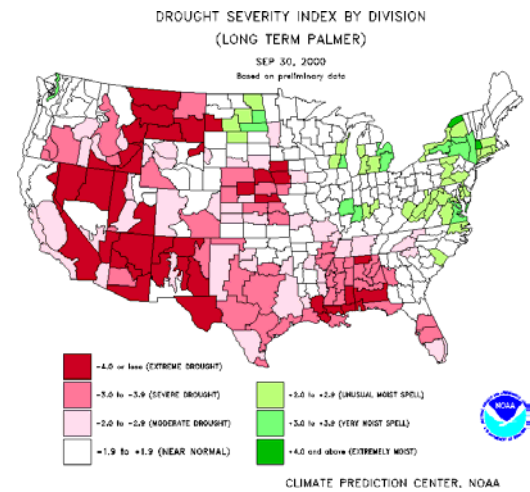
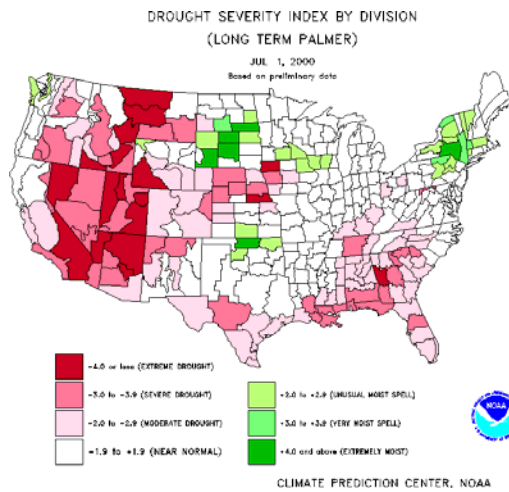
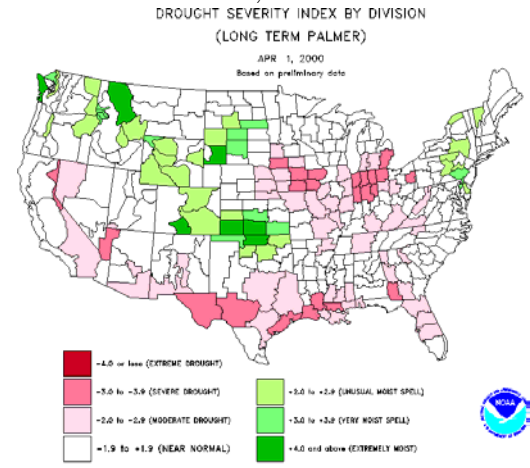
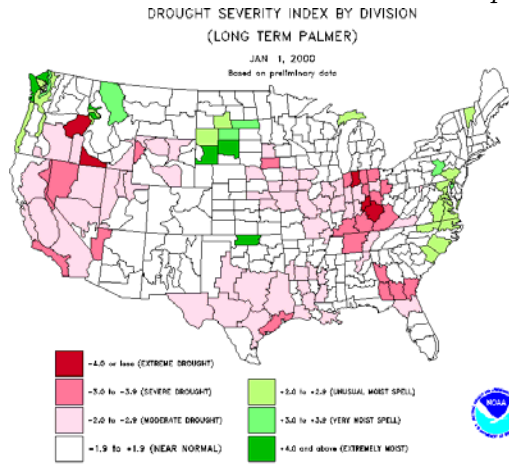
Mapas de PDSI mensual sobre EEUU en el año 1999

(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



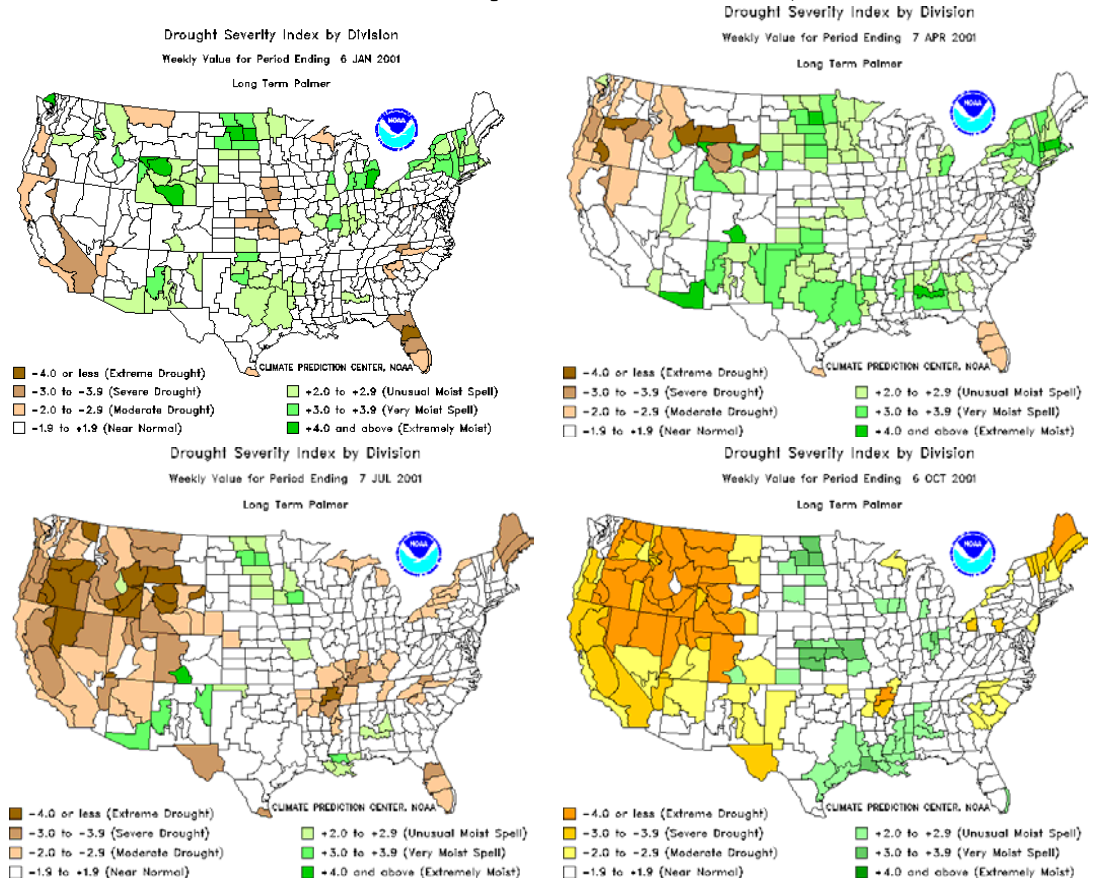
Mapas de PDSI mensual sobre EEUU en el año 2000)

(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



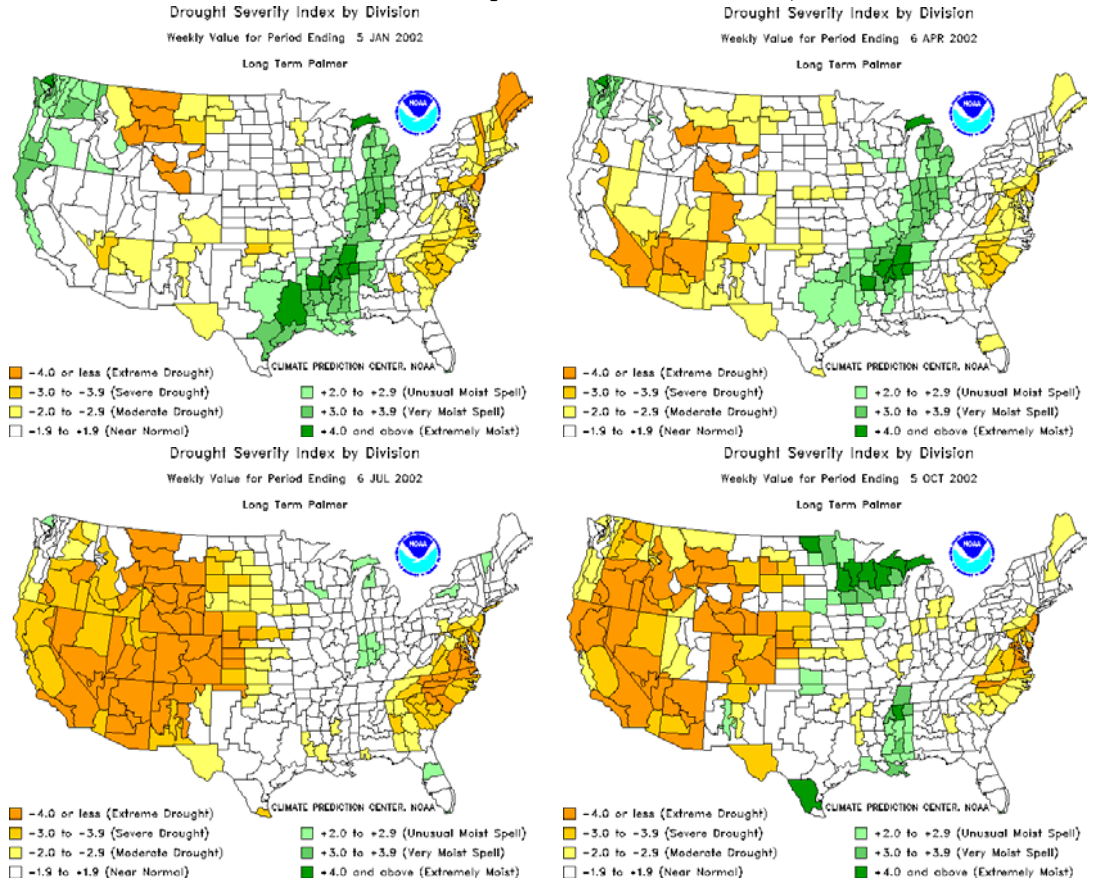
Mapas de PDSI mensual sobre EEUU en el año 2001)

(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



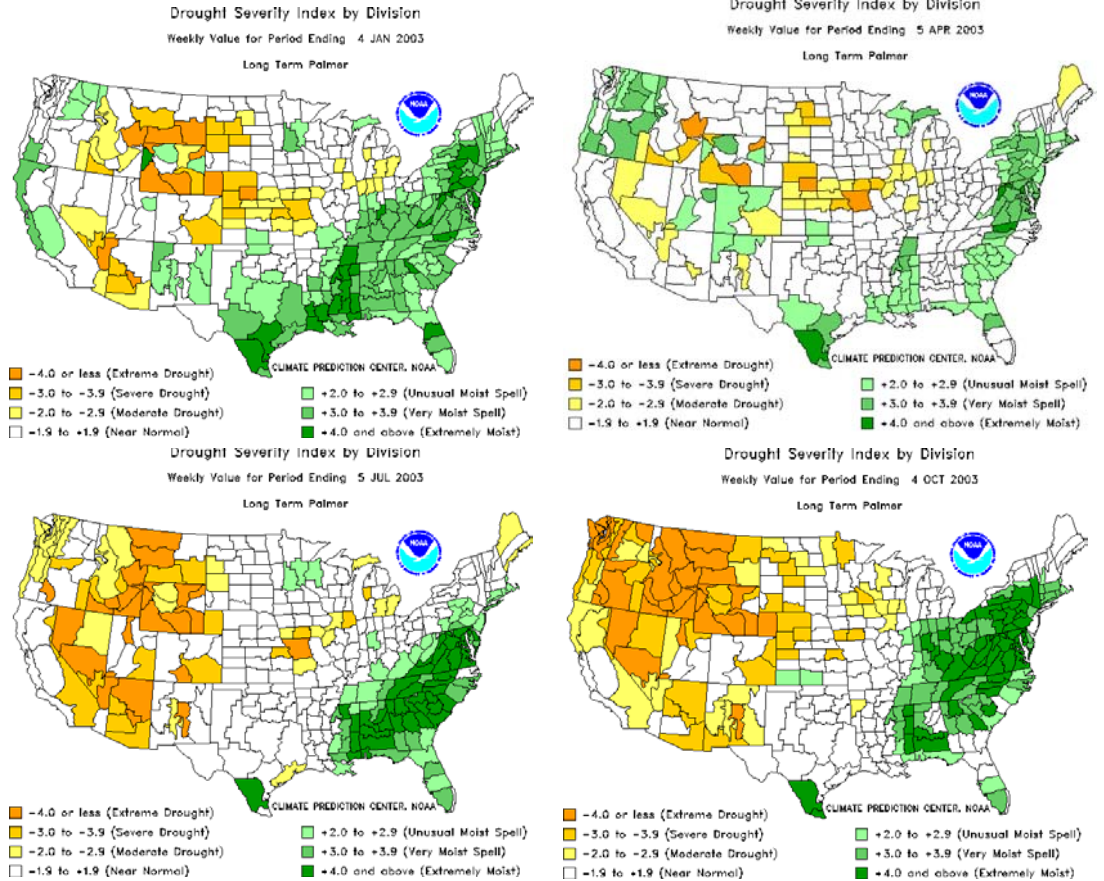
Mapas de PDSI mensual sobre EEUU en el año 2002)

(Fuente: *National Oceanic and Atmospheric Administration*, EEUU



Mapas de PDSI mensual sobre EEUU en el año 2003)

(Fuente: *National Oceanic and Atmospheric Administration, EEUU*)



A partir de esta aplicación y de su análisis comparativo con otros sistemas de control espacio-temporal de sequías, se ponen de manifiesto la adecuación del uso de los mapas de IFS en el control del desarrollo espacial de las sequías sobre una región. Destacan principalmente su independencia del régimen o regímenes climáticos que puedan producirse sobre la cuenca, permitiendo realizar comparaciones entre cualquier pareja de puntos. Ello posibilita el análisis y caracterización del desarrollo espacio-temporal de las sequías en la región de estudio, sin necesidad de discretizar la región en sub-regiones con climatologías homogénea, como ha sido propuesto por otros autores (Capítulo 3), describiendo inclusive la realización concreta dentro de cada sub-región homogénea. Por último, el índice refleja de las sequías su persistencia en el tiempo y en el espacio, resultando mayor significación de la sequía cuanto mayor es esta persistencia. Ésto es un requisito necesario en un problema de caracterización general de sequía, pues se incluye dentro de las connotaciones asociadas con el fenómeno. La metodología y algoritmo propuesto permiten el cálculo de estos mapas de IFS con un grado de detalle suficiente para contener en una gran cuenca o extensión la representación de todas las climatologías que en ella se producen, con un coste computacional adecuado.

8.2. Apoyo a la toma de decisiones en la gestión de Recursos Hídricos

El IFS, como indicador cuantitativo del estado de sequedad de una determinada variable hidro-climática, puede tener interesantes aplicaciones en la gestión de Recursos Hídricos (RRHH), donde los episodios de sequía extrema producen unas condiciones de funcionamiento límites en el sistema. Aquí son utilizados dos casos de aplicación para ilustrar la aplicación potencial del índice en la gestión de RRHH. En el primero de ellos, se analiza la cuenca del río Júcar, sobre la que se encuentra un sistema complejo de RRHH con múltiples embalses y demandas, para cada una de las cuales se debe mantener diversos requisitos de garantías. La distribución de los recursos para la satisfacción de las diferentes demandas resulta de máxima importancia, especialmente cuando se está produciendo un evento de sequía. La aplicación hace notar para que tipos de demandas el uso del IFS puede ayudar en la definición del criterio de suministro más adecuado.

En segundo lugar, se estudia la cuenca del río Conchos (México), principal tributario del bajo río Bravo, tratándose también de un sistema complejo, con múltiples demandas entre las que destaca la demanda impuesta por el tratado internacional de aguas entre EEUU y México. Como se analiza, una de las peculiaridades que dificulta la aplicación del tratado es la

indefinición del concepto de sequía extrema, recogido en el tratado como una situación excepcional. El análisis de la gestión en la cuenca muestra como puede usarse el IFS para su definición, y del mismo modo que el caso de la cuenca del río Júcar, puede utilizarse el índice en la definición de la política de operación óptima de los embalses.

8.2.1. El IFS en la política de restricciones a suministros. La gestión de RRHH en la cuenca del río Júcar

La cuenca del río Júcar es utilizada aquí como caso de estudio en el análisis de aplicaciones del IFS en la gestión de sistemas complejos de recursos hídricos. Su uso principal se encuentra en la posibilidad de su incorporación dentro de la definición de una regla de operación función del nivel de sequedad en el que se encuentren las aportaciones de la cuenca, cuantificado por el índice, y estableciendo las restricciones en función del mismo. Tratándose de un ejemplo donde se pretende principalmente analizar el potencial uso de IFS, el sistema de recursos hídricos del Júcar se ha simplificado para poner de manifiesto su adecuación de forma clara. La implementación práctica de la política de operación pasaría por una modelación del sistema en mayor detalle, y un estudio de la regla de operación más pormenorizado.

Descripción de la cuenca del río Júcar

La cuenca del río Júcar se encuentra en el levante español, ocupando parte de las Comunidades Autónomas de Aragón, Castilla-La Mancha y Valencia, y siendo su sistema hídrico perteneciente al ámbito administrativo de la Confederación Hidrográfica del Júcar, que incluye dentro de su jurisdicción un número de cuencas vertientes adicionales (figura 8.3). La precipitación media en la cuenca se encuentra entorno a los 500 mm/año, con variaciones entre 800 mm/año en la zona de alta de las cuencas, hasta valores próximos a 300 mm/año en la zona de meseta. Así en la cuenca se pueden distinguir regiones de clima semiárido, sub-húmedo y húmedo. El río Júcar cuenta como afluente principal el río Cabriel que nace de la misma zona de montaña, el Sistema Ibérico, y en el que se sitúa el embalse de Cabriel. En el río Júcar se encuentran como principales elementos reguladores el embalse de Alarcón y el embalse de Tous. Sobre la cuenca se sitúa un alto porcentaje de suelo ocupado por la actividad agrícola (>50%), parte de ella de regadío. Además de las demandas de regadío, el sistema del río Júcar abastece a las poblaciones de Valencia, Sagunto y Albacete.



Figura 8.3: Cuencas hidrográficas dentro de la Confederación Hidrográfica del Júcar (Fuente CHJ).

Recursos Hídricos y elementos de regulación El esquema representado en la figura 8.4 muestra el sistema del río Júcar tal y como va a ser considerado aquí de forma simplificada. El sistema cuenta con tres embalses principales para la gestión: Alarcón, en el río Júcar, con una capacidad de $1112Hm^3$; Contreras, en el río Cabriel, con una capacidad de almacenamiento de $463Hm^3$; y el embalse de Tous, en el río Júcar aguas abajo de la confluencia del río Cabriel, con una capacidad de almacenamiento máxima de $340Hm^3$. Se van a considerar dentro del sistema cinco puntos de incorporación de aportaciones naturales: aportaciones al embalse de Alarcón, aportaciones en el sitio de Molinar (aguas abajo de Alarcón), aportaciones en el embalse de Contreras, aportaciones al embalse de Tous (aguas abajo de Contreras y Molinar) y aportaciones en Cullera (aguas abajo de Tous). Los estadísticos básicos anuales de estas aportaciones quedan reflejados en la tabla 8.1. Cabe destacar los diferentes regímenes de variación entre las aportaciones en Alarcón y Contreras, frente a las aportaciones en Molinar y Tous, siendo estas últimas unas aportaciones con una componente subte-

rránea mucho más importante, lo que motiva una menor desviación típica, y un mayor sesgo fruto de la superposición de las componentes subterránea y superficial.

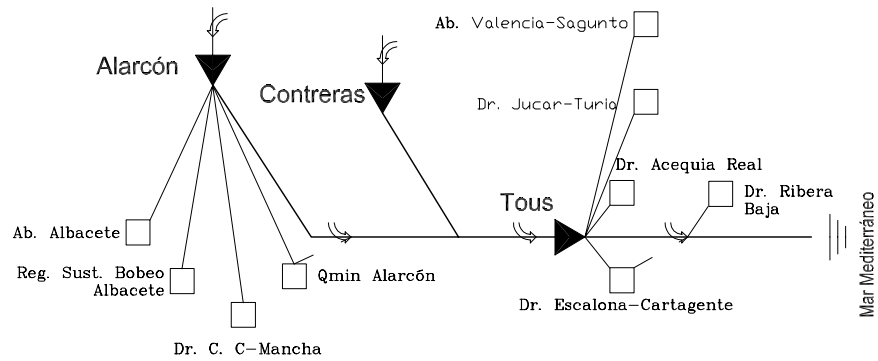


Figura 8.4: Esquema considerado del Sistema de RRHH del río Júcar.

Punto de Aforo	Aportación Media Anual	Desviación Típica-Aportación Anual	Sesgo-Aportación Anual
	Hm^3	Hm^3	-
Alarcón	424	236	0.7
Molinar	354	86	1.4
Contreras	383	199	0.7
Tous	280	98	1.8
Cullera	348	224	1.5

Tabla 8.1: Estadísticos principales de las aportaciones anuales naturales en el río Júcar de la serie de 1940-1995 (Fuente: PHN, 2000).

Embalse	OCT	NOV	DIC	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	MIN
Alarcón	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	1112	30
Contreras	463	463	463	463	463	463	463	463	463	463	463	463	15
Tous	72	72	378	378	378	378	378	378	378	378	378	195	72
													30

Tabla 8.2: Capacidades de almacenamiento (Hm^3) estacional de los embalses principales en el río Júcar (Fuente: PHN, 2000).

Se observa que la mayor capacidad de regulación se encuentra en el embalse de Alarcón, en la cabecera de la cuenca, que proporciona la principal componente de regulación interanual. El resto de embalses mantienen principalmente una función de regulación intranual, a lo que se añade en el caso del embalse de Tous, la función laminadora de avenidas que motiva su variación estacional de su capacidad de almacenamiento máxima 8.2.

Las demandas y los criterios de garantías En lo que se refiere a las demandas consideradas en el sistema, la tabla 8.3 muestra las demandas principales a las que el sistema sirve, describiéndose su distribución a lo largo del año y los retornos al sistema, en el caso de éstos existir. Se consideran tres tipos principales de demandas: demandas de abastecimiento urbano, demandas de regadío y caudales mínimos aguas abajo de embalses o caudales ecológicos. La demanda total de uso consuntivo se sitúa en un valor de $1374Hm^3/año$, siendo la aportación media anual de $1789Hm^3/año$.

En lo que respecta a las garantías de suministros de estas demandas, las exigencias que mantienen las demandas de abastecimiento y de regadíos son marcadamente diferentes. Mientras las primeras mantienen niveles altos de garantías temporales y volumétricas, siendo las de mayor prioridad, las segundas mantienen una mayor flexibilidad frente a la existencia de déficit. Los criterios de garantía que con mayor frecuencia son utilizados en el caso de las demandas urbanas son criterios de garantía mensual, definiendo la ocurrencia de fallo cuando el suministro sea menor del 80 % de la demanda. Bajo este criterio, los niveles de garantías manejados habitualmente oscilan entorno a 0.96-0.98.

En el caso de las demandas de regadío pueden encontrarse en la literatura multitud de criterios de garantías para cuantificar el grado de satisfacción a la demanda dentro de un sistema de RRHH, para una revisión de los criterios disponibles vease Estrada (Estrada, 1994). En general los criterios relativos a suministros a demandas de regadío, incorporan requisitos de suministro de la demanda en diferentes agregaciones temporales, permitiendo distintos niveles de déficit en diferentes períodos de agregación. Así como uno de los criterios más conocidos, es el criterio del *Utah Department of Water Resources* (Lall & Miller, 1988), el cual considera satisfactorio el suministro si se cumplen simultáneamente las siguientes condiciones:

- El déficit en un año no excede el 50 % de la demanda anual
- El déficit de dos años consecutivos no excede el 75 % de la demanda anual
- El déficit en diez años consecutivos no excede el 100 % de la demanda anual

Demanda	Annual (Hm^3)	Distribución Mensual de la Demanda (%)												Ret. (%)		
		OC	NV	DC	EN	FB	MR	AB	MY	JN	JL	AG	SP			
Ab/Reg/Qmin																
Ab. Albacete	31	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.4	8.4	8.4	8.4	-
Ab. Valencia	189	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.4	8.4	8.4	8.4	-
Ab. Sagunto	31	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.4	8.4	8.4	8.4	-
Reg. Sust. Bomb. Alb.	80	7	0	0	0	1	4	4	6	8	24	29	17			-
Reg. Dtr. Acuífero	185	8	7	8	8	8	9	7	9	7	9	11	9			-
Reg. C. C-Mancha	100	7	0	0	0	1	4	4	6	8	24	29	17			-
Reg. Júcar-Turia	125	10	2	1	0	0	0	5	8	16	21	21	16			-
Reg. Acequia Real	300	8	7	2	4	5	7	9	12	11	14	13	8			-
Reg. Escalona-Carcagente	54	7	6	2	3	4	3	8	11	16	17	15	8			48
Reg. Ribera Baja	279	8	8	8	6	3	5	7	16	12	12	11	4			-
Qmin Alarcon	60	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.4	8.4	8.4	8.4	8.4	50
Qmin Contreiras	12	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.4	8.4	8.4	8.4	8.4	100
Qmin Tous	19.2	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.4	8.4	8.4	8.4	8.4	100
Total	1465															

Tabla 8.3: Demandas principales consideradas en el río Júcar (Fuente: PHN, 2000).

Este criterio es de aplicación sobre una determinada serie histórica de demandas y suministros, generalmente obtenida por simulación, y se trata de un criterio de aceptación o rechazo. Estadísticamente, la probabilidad de cumplimiento de las condiciones que impone el criterio disminuye con el aumento de la longitud de la serie, siendo necesaria una modificación del criterio cuando se pretende realizar análisis sobre simulaciones de series de gran longitud. Una forma conveniente de introducir cuantificación estadística bajo criterios de este tipo, es la cuantificación por separado de las garantías anuales de cumplimiento de cada una de las condiciones, llegando en lugar de un criterio de aceptación o rechazo, a un valor de garantía para cada período de agregación.

La adopción de un criterio como el de Utah, con déficit permisibles en diferentes períodos de agregación, frente a criterios más exigentes y menos flexibles, como los criterios de garantía mensual, implica también la conveniencia de diferentes políticas de gestión. En este sentido el suministro a demandas urbanas cuando la fuente de suministro es única suele seguir una política de operación similar a la Regla de Operación Normal (figura 8.5). Sin embargo, en el caso de las demandas de riego, donde son permisibles mayores déficit, y sobre todo es importante su distribución en el tiempo, resulta en algunos casos más convenientes la aplicación de políticas de gestión basadas en provocar déficit cuando las disponibilidades de recursos son bajas, aunque todavía suficientes para el suministro al 100 %. Con ello se consigue reducir la probabilidad de producirse déficit futuros extraordinarios. A continuación se expone una política de gestión para el sistema del río Júcar, que produce restricciones al suministro en las demandas agrícolas en función de la disponibilidad del recursos y del IFS, la cual ha sido optimizada para una función objetivo diseñada que engloba la probabilidad de ocurrencia de los diferentes fallos.

Política de gestión basada en restricciones al suministro función del IFS

Como ha sido introducido en el párrafo anterior, las diferentes exigencias de garantías en distintos tipos de demanda hacen convenientes diferentes políticas de gestión de los recursos. En el caso del sistema del río Júcar, se han considerado tres tipos de demandas. La primera, y de mayor prioridad, es la demanda de abastecimiento urbano; la segunda, es la demanda de regadío; y la tercera, es la demanda de caudal mínimo aguas abajo de los embalses. Para la satisfacción de estas demandas, se construye la siguiente política de operación con paso temporal mensual:

1. En primer lugar se decide el suministro a las demandas urbanas, aplicándoles una regla tipo normal, es decir, suministrando todo el

recurso disponible hasta su satisfacción. Esta es la regla que con mayor garantía satisface este tipo de demandas, pues aunque admite cierto déficit hasta la declaración del fallo (20 %), al ser esta holgura pequeña, una política con restricciones no introduciría notorias ventajas. La fuente de recurso para el suministro de estas demandas es en el caso de Ab. Albacete el embalse de Alarcón, única fuente posible. En el caso de las demandas de Ab. Valencia y Sagunto, se adopta como orden para la extracción del recurso el embalse de Tous, el embalse de Contreras, y en último lugar, cuando no exista recurso suficiente en los dos anteriores, el embalse de Alarcón. Ello responde a la idea de la gestión de embalses en serie, y de mantener el embalse de Alarcón como embalse de reserva, por tener éste una capacidad de regulación significativamente mayor, y por tanto, menor probabilidad de llenado que Contreras.

2. En segundo lugar, se calculan los volúmenes a suministrar a las demandas de regadío. Para ello se establecen tres grupos de demandas de regadío: las demandas suministradas desde el embalse de Alarcón, las demandas con suministro derivado desde el embalse de Tous, y la demanda de la Ribera Baja. Para cada uno de estos grupos de demandas de regadío se propone realizar una restricción al suministro en función de la disponibilidad de recurso y del IFS que se esté produciendo. La forma más sencilla de establecer restricciones al suministro cuando la disponibilidad de recurso comienza a ser pequeña es modificando la pendiente del primer tramo de la curva que define la Regla de Operación Normal (figura 8.5). Asignándole una pendiente menor de 1 al primer tramo, donde se producen los déficit, se adelanta la aparición de estos déficit, dando lugar a la Regla de Operación con Restricciones (figura 8.5). La pendiente del primer tramo de la regla debe asignarse de tal modo que se reduzcan los déficit máximos, pero no se introduzcan en exceso nuevos déficit pequeños consecutivos.

El resultado de esta regla sobre un criterio de garantía tipo Utah es positivo cuando se presenta una situación de sequía, donde de forma persistente van a existir déficit en el suministro. Sin embargo, el inconveniente se encuentra cuando se tratan de variaciones normales dentro de la aleatoriedad del recurso hídrico, las cuales producen un déficit puntual en el tiempo con una rápida recuperación del sistema. En esos casos, este tipo de reglas incrementan el déficit total producido.

De acuerdo a estas consideraciones, parece oportuno adoptar una pendiente del primer tramo de la regla en función del estado de sequedad, de modo que en situaciones de extrema sequía la pendiente disminuya, procurando reducir los daños que unos recursos nulos producirían, y en

los casos en los que no sea notoria la situación de sequía se mantengan unas restricciones si las disponibilidades son bajas de forma cautelar. Ello puede expresarse asignando la siguiente función a la pendiente del primer tramo de la regla:

$$\rho = \begin{cases} 1 & IFS \leq \exp(C) \\ \frac{C}{\ln(IF S)} & IFS > \exp(C) \end{cases} \quad (8.1)$$

La regla de operación así definida, con restricciones función del estado de sequedad, permite la modulación de los déficit en función del estado de sequía que se esté produciendo. Además, al finalizar la sequía, mientras el índice recupera valores bajos, las restricciones se disminuyen o se anulan siempre que las aportaciones hayan sido suficientes para conseguir valores de disponibilidad de recurso dentro de los rangos normales. La forma en la que se ha implementado estas reglas en el sistema Júcar de múltiples embalses y demandas es considerando como disponibilidades para cada demanda la suma de todo el volumen disponible para satisfacer su demanda, es decir:

Demanda de riego	Disponibilidad de recurso (F)
Reg. Sust. Bomb. Alb. Reg. C. C-Mancha	Volumen Almacenado Alarcón +Aportaciones Alarcón
Reg. Júcar-Turia Reg. Acequia Real Reg. Escalona-Carcagente	Volumen Almacenado Alarcón +Aportaciones Alarcón +Volumen Almacenado Contreras +Aportaciones Contreras +Volumen Almacenado Tous +Aportaciones Molinar +Aportaciones Tous
Reg. Ribera Baja	Volumen Almacenado Alarcón +Aportaciones Alarcón +Volumen Almacenado Contreras +Aportaciones Contreras +Volumen Almacenado Tous +Aportaciones Molinar +Aportaciones Tous +Aportaciones Cullera

Tabla 8.4: Fuentes que componen las disponibilidades de recursos para las distintas demandas de riego en el sistema del río Júcar.

En cuanto a la variable sobre la que se computa el IFS para la evaluación de cada restricción, se adopta la suma de las aportaciones naturales que se produciría en cada fuente de suministro de las demandas de riego. Con ello quedaría para cada caso, *dem = Alarcón, Tous, Ribera Baja*, la siguiente restricción:

$$R = \begin{cases} F_{dem} \cdot \rho_{dem} & F_{dem} \leq \frac{D_{dem}}{\rho_{dem}} \\ D_{dem} & F_{dem} > \frac{D_{dem}}{\rho_{dem}} \end{cases} \quad (8.2)$$

donde R es el volumen a suministrar, F_{dem} es el volumen disponible para cada tipo de demanda, D_{dem} es el volumen demandado. Estos volúmenes a suministrar han sido detraídos de las fuentes de suministros siguiendo un orden de toma de cada embalse similar al que se produciría en el abastecimiento a la demanda urbana.

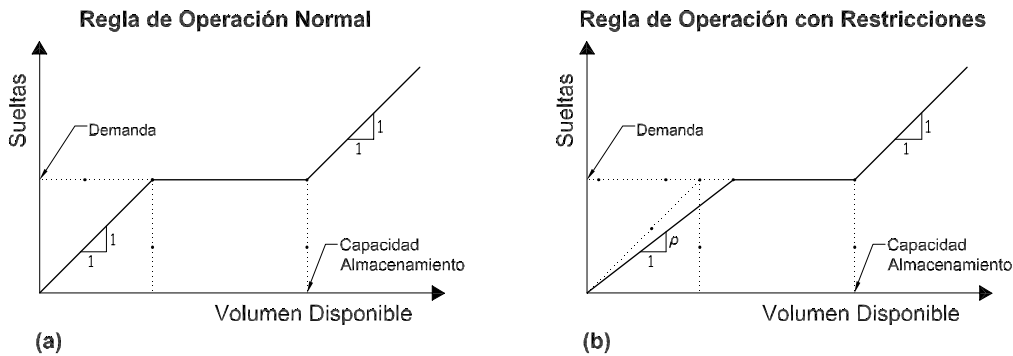


Figura 8.5: Sueitas producidas en un sistema de RRHH gestionada mediante una Regla de Operación Normal (a) y mediante una Regla de Operación con Restricciones (b).

3. Por último, es comprobada la satisfacción de las demandas relativas a caudales mínimos aguas abajo de los embalses a partir de las sueltas que han sido necesarias realizar de cada embalse, y en caso de no cumplirse, son efectuados los desembalses necesarios para alcanzarlo, reajustándose los transvases realizados anteriormente, siempre que exista disponibilidad de recurso.

Con una política de operación como la descrita, donde se ha obviado en la descripción el balance adecuado en cada uno de los embalses, teniendo en cuenta las capacidades estacionales y las pérdidas por evaporación producidas en los vasos, su funcionamiento queda en función de la selección de los tres parámetros que definen las restricciones sobre las demandas de riego: C_{dem} , $dem = Alarcón, Tous, Ribera Baja$. La fijación de estos parámetros debe realizarse por simulación de una serie de aportaciones, pudiendo ser utilizada una función objetivo a optimizar. A continuación se exponen los resultados de realizar una optimización sobre una serie histórica de aportaciones naturales en el sistema, y seguidamente, los resultados sobre una serie sintética de mayor longitud como análisis de sensibilidad.

Política de operación óptima para la serie histórica de aportaciones.

En este caso ha sido analizada la serie histórica de aportaciones naturales en el sistema del río Júcar entre los años hidrológicos de 1940-1941 a 2000-2001, la cual se presenta en la figura 8.6.

A partir de estas aportaciones, las series resultantes del IFS en los tres puntos de derivación de suministro quedarían como representa la figura 8.7, con unos valores máximos generalizado en toda la cuenca a mediados de la década de los noventa, de valor entorno a los 100 años.

La función objetivo a optimizar se construye mediante producto de las garantías de suministro a las demandas de abastecimiento y la frecuencia de cumplimiento de las tres condiciones del criterio de UTAH para cada una de las demandas de riego agrupadas por origen de suministro:

$$O(C_{Alarcón}, C_{Tous}, C_{Rib Baja}) = \prod_i G_i^m \cdot \prod_j \left[\prod_{k=1}^3 P_{j,k} \right] \quad (8.3)$$

donde $i = Ab Albacete, Ab Valencia - Sagunto$, n es un coeficiente para ponderar en mayor medida la satisfacción a las demandas urbanas,

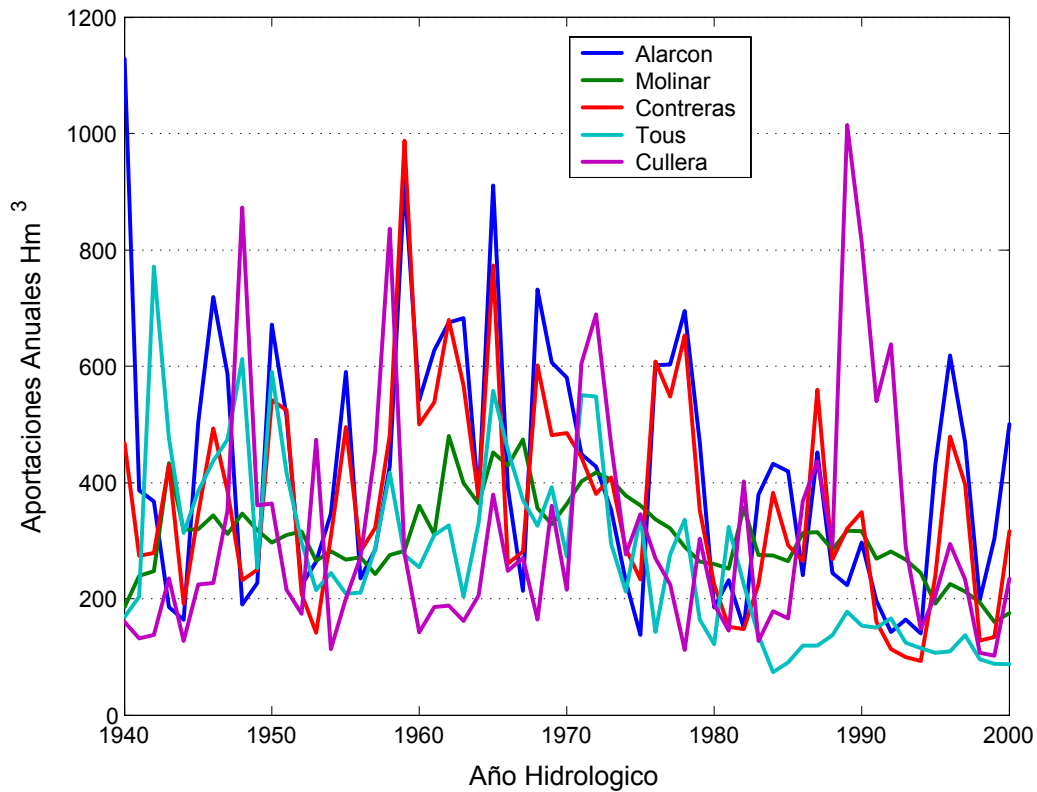


Figura 8.6: Aportaciones naturales anuales históricas en el sistema del río Júcar.

$j = Rg$ desde Alarcón, Rg desde Tous, Rg Ribera Baja, y $P_{j,k}$ indica la frecuencia de cumplimiento de la condición $k = 1, 2, 3$ en el criterio de UTAH. Para la búsqueda de la optimización de esta función objetivo ha sido utilizado el algoritmo de optimización *Shuffled Complex Evolution Metropolis Algorithm* (SCEM-UA) de la Universidad de Arizona (Vrugt *et al.*, 2003). La implementación del algoritmo sobre este problema arroja como gráficos de muestreo de la función objetivo la figura 8.8. Así funciones de distribución posteriores de los parámetros a optimizar tras el resultado de la búsqueda quedan representadas en la figura 8.9. Los valores muestreados con mayor frecuencia se corresponden con los valores próximos al óptimo.

De acuerdo a los resultados mostrados en las figuras 8.8 y 8.9, el parámetro que mejor definida tiene su posición entorno al óptimo es el coeficiente de restricción de Tous, seguido por el de la Ribera Baja, y con mucha menor definición el coeficiente de restricción de Alarcón. Los valores óptimos encontrados para la serie histórica quedan como sigue, acompañándose los valores asociados de garantías y frecuencia de cumplimiento de las condiciones del criterio de Utah:

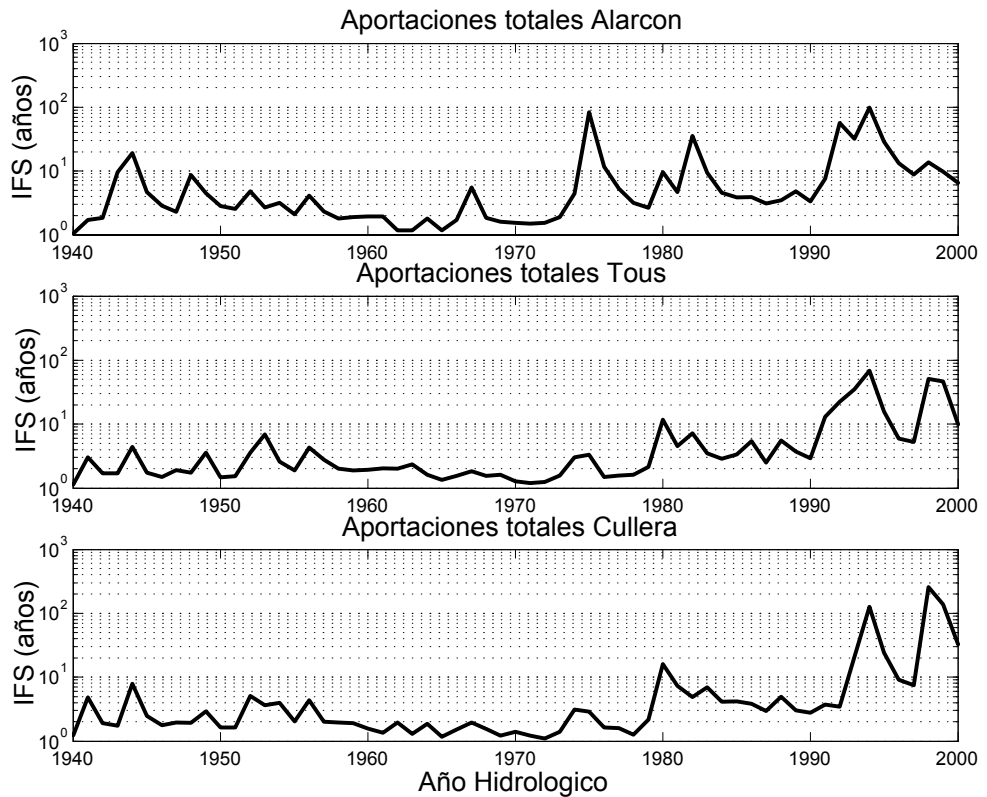


Figura 8.7: Índice de Frecuencia de Sequías sobre la series de aportaciones históricas totales en los puntos de derivación de suministro del río Júcar.

Esta optimización muestra la alta sensibilidad de los resultados a las restricciones impuestas en los suministros desde Tous, donde se concentra el mayor volumen de consumos de regadío, seguido por Alarcón. El caso de la Ribera Baja, la poca sensibilidad sobre las restricciones impuestas es debida fundamentalmente a que esta demanda es satisfecha en gran medida directamente con las aportaciones en el último tramo del río Júcar y los retorno de otros riegos, siendo pequeño el volumen adicional que necesita para su completa satisfacción. A continuación se analiza la sensibilidad de estos resultados a la serie considerada, utilizándose en la optimización una serie sintética de mayor longitud.

$C_{Alarcón}$	C_{Tous}	$C_{Ribera Baja}$
0.5	1	3

Tabla 8.5: Valores óptimos de los coeficiente de restricciones sobre la serie histórica en el sistema del río Júcar.

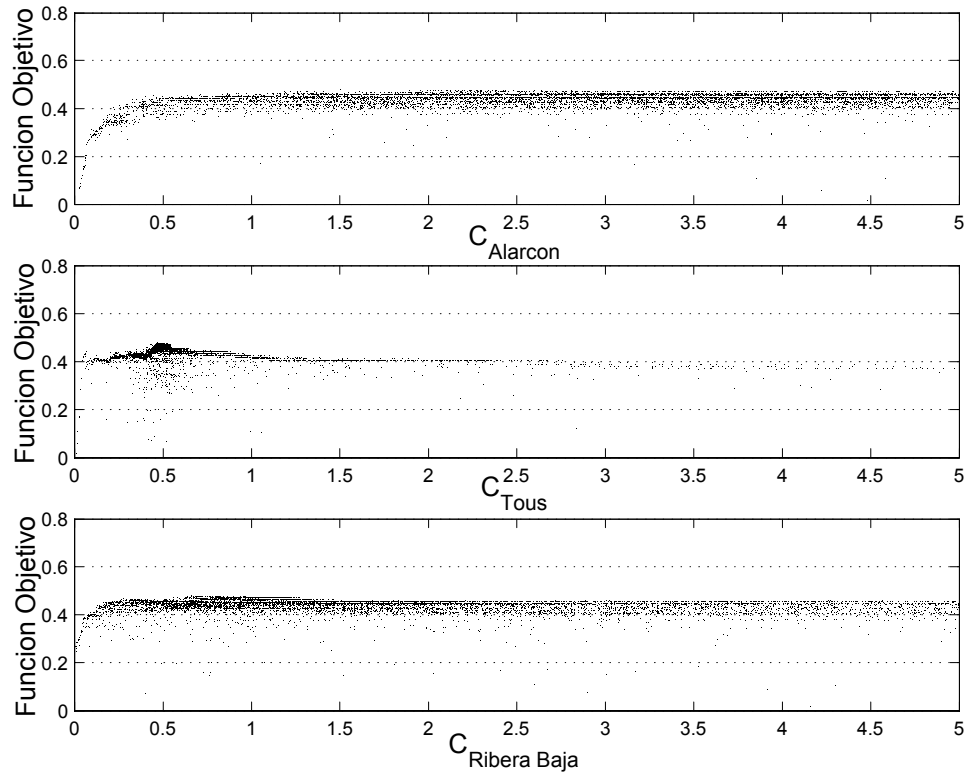


Figura 8.8: Muestreo de la función objetivo en el proceso de optimización para la serie histórica de aportaciones, en el sistema del río Júcar, siguiendo el algoritmo SCEM-UA.

	Ab. Albacete	Ab. Valencia-Sagunto
Garantía	1	0.99

Tabla 8.6: Garantías al abastecimiento urbano en el sistema del río Júcar obtenidas tras la optimización de los coeficientes de restricciones sobre la serie histórica.

Frecuencia cumplimiento	Rg. desde Alarcón	Rg. desde Tous	Rg. Ribera Baja
Condición en 1 año	0.95	0.90	1.00
Condición en 2 años	0.95	0.92	1.00
Condición en 10 años	0.91	0.90	1.00

Tabla 8.7: Frecuencia de cumplimiento de las condiciones del criterio de Utah en el suministro a las zonas de regadío del sistema del río Júcar tras la optimización de los coeficientes de restricciones sobre la serie histórica.

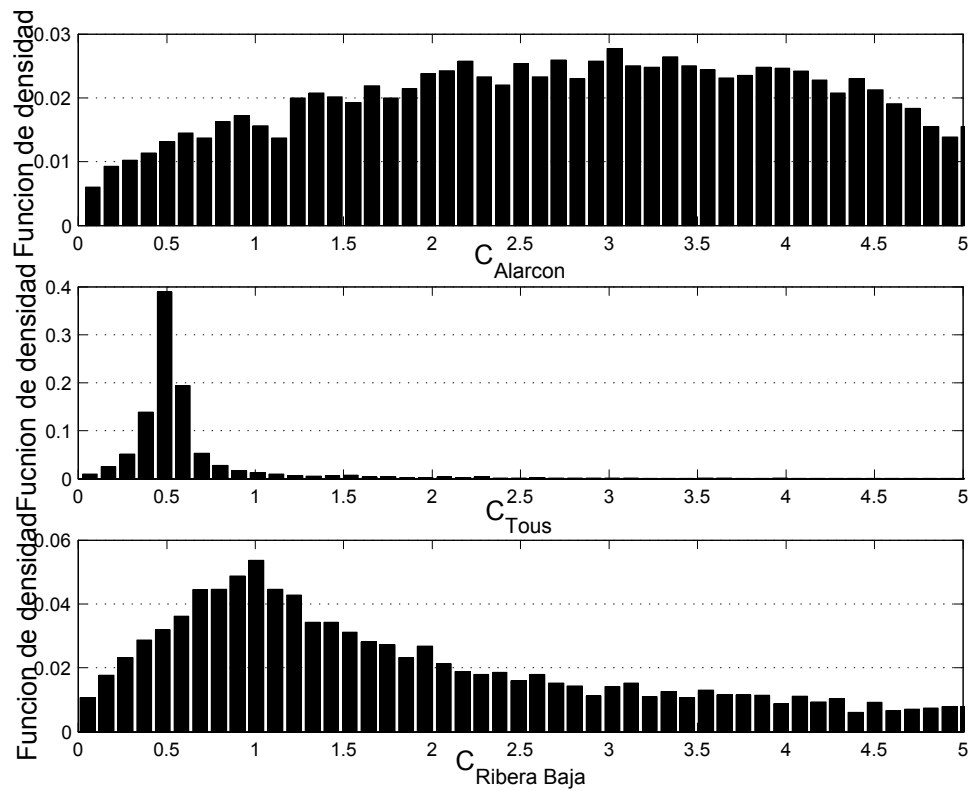


Figura 8.9: Función de distribución posterior de los parámetros a optimizar sobre la serie de aportaciones históricas en el sistema del río Júcar, siguiendo el algoritmo SCEM-UA.

Sensibilidad de la política de operación óptima Para el análisis de sensibilidad del resultado de optimización obtenido sobre la serie histórica se ha construido un modelo de generación de aportaciones el sistema del río Júcar. El modelo estocástico adoptado es el modelo multivariado de tipo ARMA(1,1) (Salas, 1992), al cual se ajustan las aportaciones históricas, tras su transformación, estacionarización y tipificación. La transformación logarítmica (ec. 8.4) aplicada sobre la muestra de aportaciones produce una muestra normal, para las aportaciones al embalse de Alarcón, y Contreras, mientras que para el caso de Molinar y Tous es necesario aplicar una transformación logarítmica con tres parámetros (ec. 8.5), y en el caso de Cullera se ajusta la doble transformación logarítmica (ec. 8.6).

$$Y_i = \ln(X_i + 1) \quad (8.4)$$

$$Y_i = \ln(X_i - m) \quad (8.5)$$

$$Y_i = \ln[\ln(X_i + 1) + 1] \quad (8.6)$$

Con el modelo construido de aportaciones, ha sido generada una serie sintética de 500 años sobre la que analizar la sensibilidad de los resultados anteriores. La figura 8.10 presenta los valores que resultan del IFS sobre las series generadas, observándose la existencia de eventos de sequía más extremos. Cuando se aplica el algoritmo de optimización sobre esta serie resultan los valores de la función objetivo que se presentan en la figura 8.11.

Las funciones de distribución posterior (figura 8.12) muestran la indefinición de los parámetros óptimos para el caso de la serie sintética generada, resultando poco sensible la función objetivo a variaciones en los parámetros que definen la política de restricciones en el suministro. Este resultado es debido principalmente a que se trata de un sistema donde las garantías de suministros y las frecuencias de cumplimientos de las condiciones son relativamente altas, por lo que en estos casos resulta más conveniente la aplicación de una regla de operación de tipo normal. Esto se pone de manifiesto si son aumentados en un 20 % el conjunto de las demandas de riego, con lo que los resultados de la optimización vendrían dados en las figuras 8.13 y 8.14.

Puede observarse que tras el incremento de las demandas de riego, el parámetro que define las restricciones sobre las demandas suministradas desde el embalse de Tous se define de nuevo entorno al valor de $C_{Tous} = 0,5$.

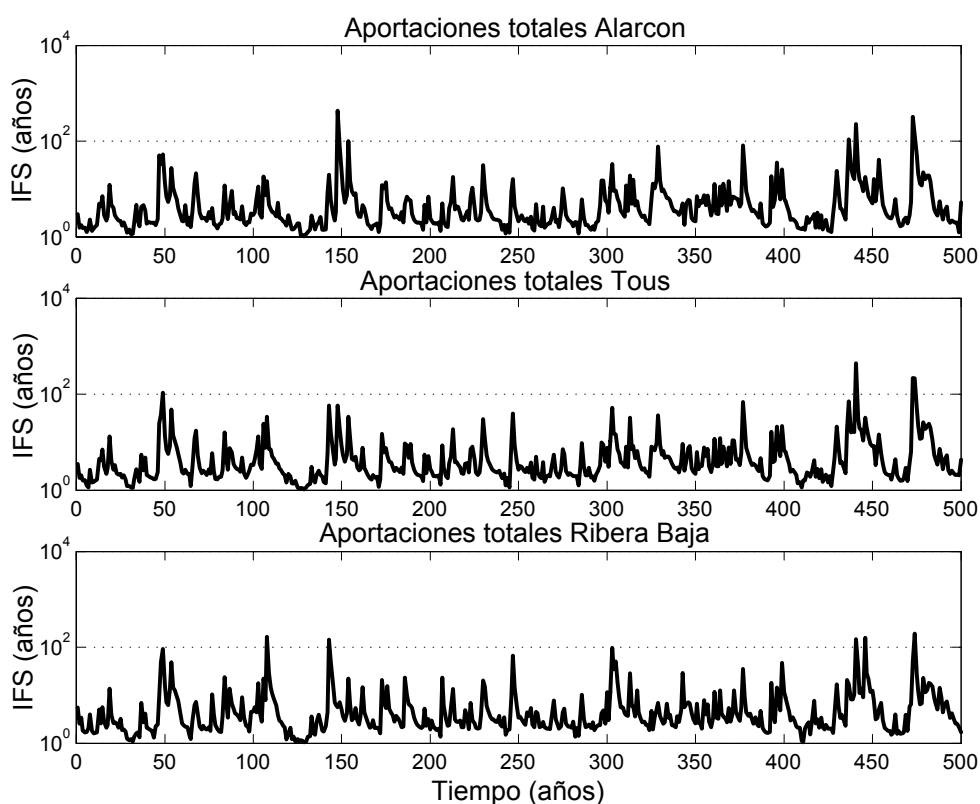


Figura 8.10: Índice de Frecuencia de Sequías de la serie de aportaciones sintéticas en los puntos de derivación a suministro del sistema del río Júcar.

Todo ello apunta que si bien se comprueba que la serie sintética generada mantiene los estadísticos básicos no significativamente diferentes de la muestra representada en la serie histórica, su introducción en el modelo de simulación produce resultados diferentes de los obtenidos con la serie histórica, manteniéndose el sistema con mayores niveles de satisfacción a la demanda que los que se obtendrían con la serie histórica. La confirmación de los resultados de la serie sintética como diagnóstico del sistema Júcar, requería el análisis en mayor detalle del modelo de aportaciones, generando nuevas series, ajustando nuevos modelos de generación, etc. Sin embargo, ello queda fuera del objeto de este trabajo. No obstante, este ejemplo de aplicación permite ilustrar el potencial uso del IFS en la definición de políticas de gestión de RRHH, siendo éste de utilidad cuando se trata de demandas en las que el grado de satisfacción no es elevado. En caso contrario, una política de restricciones no presentaría grandes ventajas, siendo más próxima a la política óptima una política tipo regla de operación normal (Bayazit & Ünal, 1990).

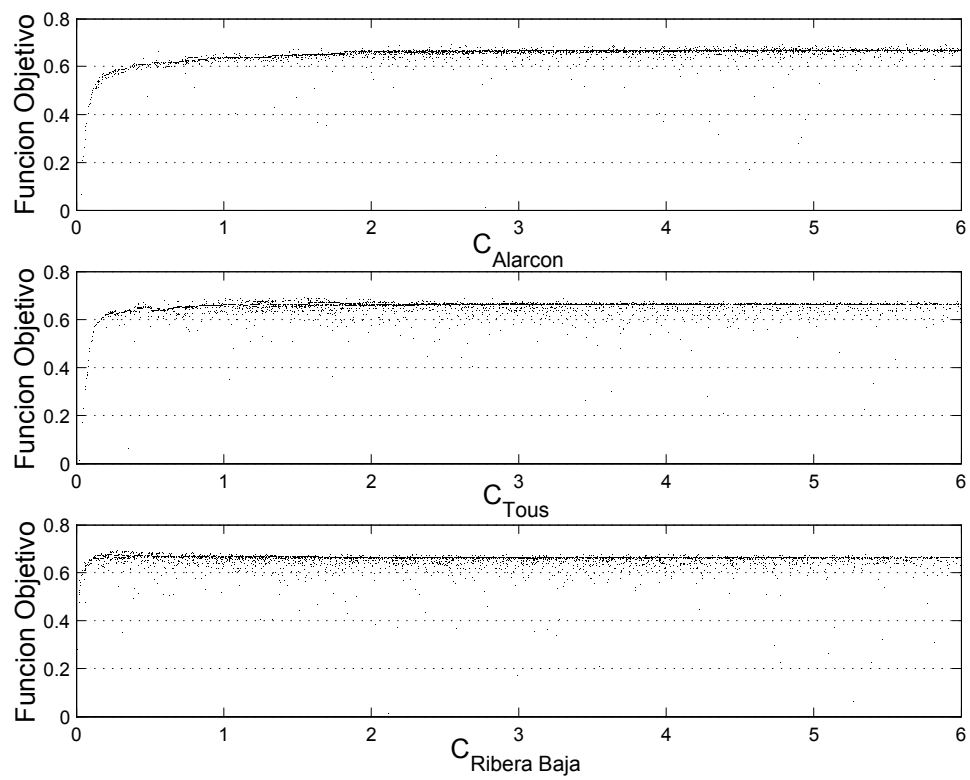


Figura 8.11: Muestreo de la función objetivo en la búsqueda del óptimo sobre la serie sintética de 500 años en sistema del río Júcar.

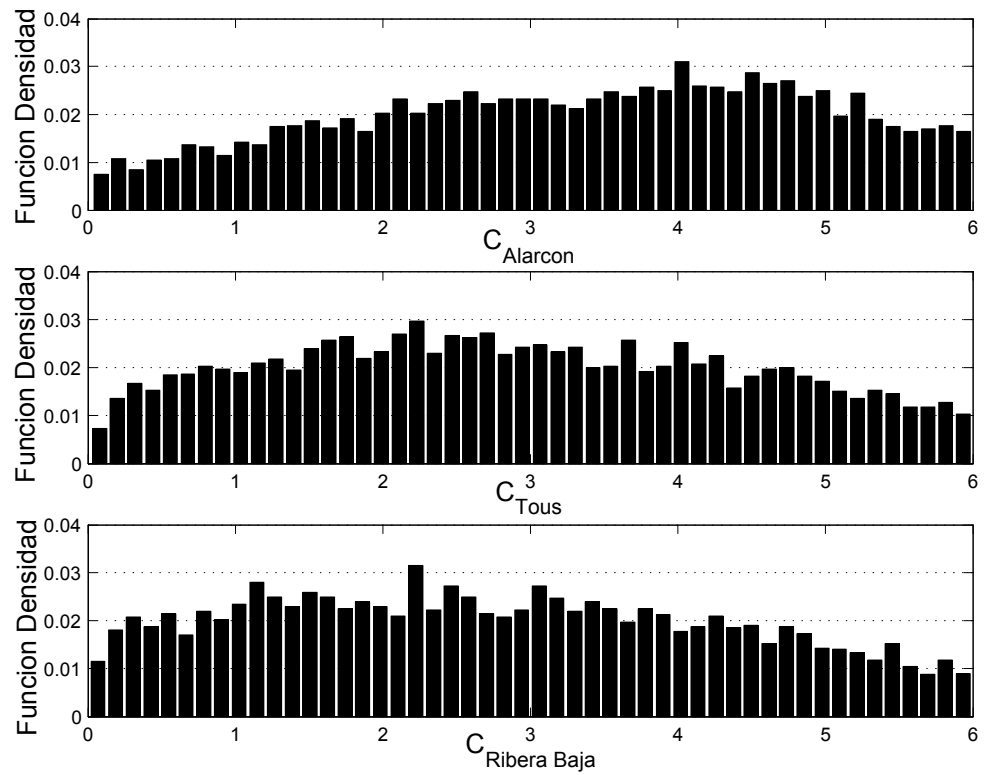


Figura 8.12: Funciones de distribución posterior de los parámetros a optimizar sobre la serie sintética de 500 años en el sistema del río Júcar.

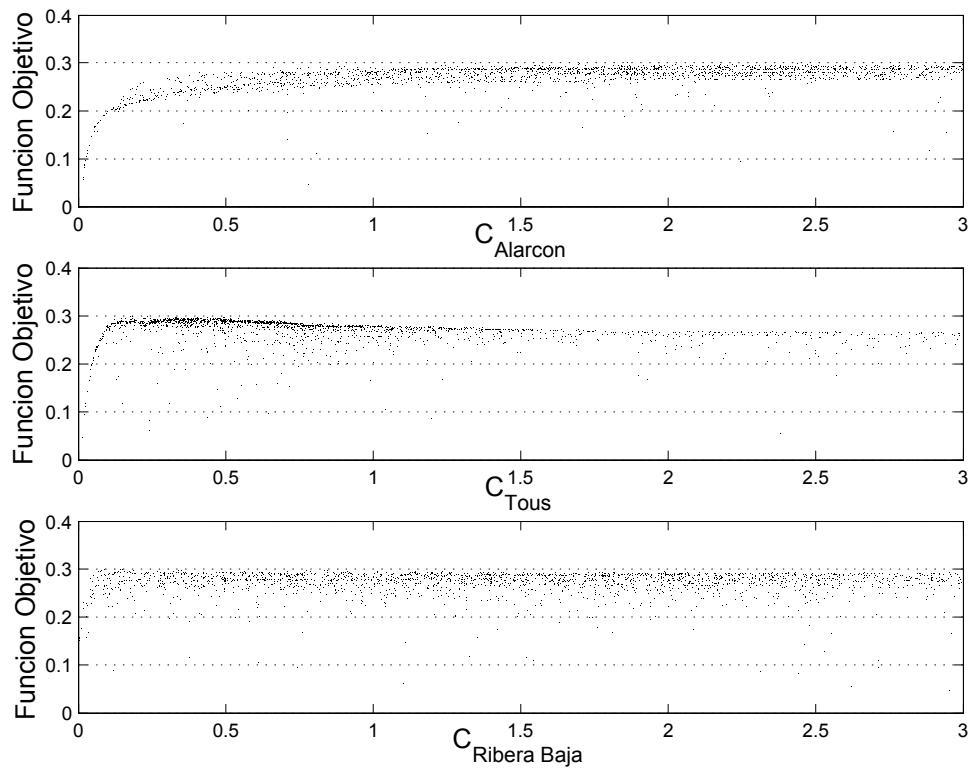


Figura 8.13: Muestreo de la función objetivo en la búsqueda del óptimo sobre la serie sintética de 500 años en sistema del río Júcar, con las demandas de riego incrementada en un 20 %.

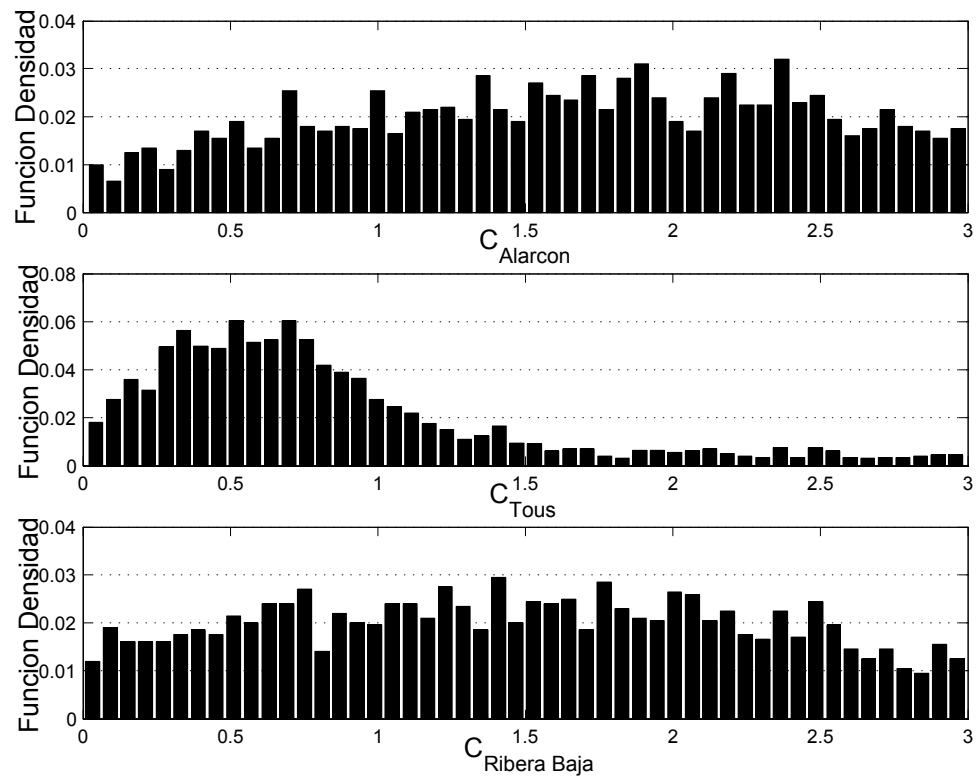


Figura 8.14: Funciones de distribución posterior de los parámetros a optimizar sobre la serie sintética de 500 años en el sistema del río Júcar, con las demandas de riego incrementadas un 20 %.

8.2.2. El IFS como indicador de sequía extraordinaria. La gestión de RRHH en la cuenca del río Conchos (México)

Como último ejemplo de aplicación se ha utilizado el río Conchos, situado al norte de México. Siendo el principal afluente del río Bravo, río internacional, la gestión de los RRHH sobre el río Conchos es de especial importancia, teniendo que cumplir con los condicionantes del tratado internacional de aguas entre México y EEUU y, al mismo tiempo, siendo este río la fuente de recurso hídrico para el riego de importantes extensiones en zonas áridas y semiáridas del valle de Chihuahua (México), con una economía principalmente agrícola. Esta aplicación permite analizar la gestión de los RRHH en la cuenca del río Conchos, y poner de manifiesto la utilidad que el IFS puede tener en la definición de una política de gestión óptima en la cuenca.

Descripción de la cuenca del río Conchos

La cuenca del río Conchos, el más importante río de la zona norte de México, nace de la Sierra Madre Occidental y riega amplias superficies de regadío en zonas áridas del valle central de Chihuahua. La cuenca ocupa una superficie de unos 68000km^2 y vierte sus aguas al río Grande, del cual es el principal tributario en su zona baja, con una aportación entorno al 40 % del total (Collado *et al.*, 2001). La precipitación media anual es de unos 400mm/año , con unos valores máximos en la cabecera de cuenca del orden de 800mm/año y unos valores mínimos en la zona del valle, y próximo a la desembocadura de 150mm/año (figura 8.15). Los gradientes que se producen en precipitaciones, así como en régimen de temperaturas, permite que en la cuenca se observe un rango de climas desde semi-húmedo hasta árido, variando de sur a norte.

El uso de los RRHH dentro de la cuenca del Río Conchos son en un 90 % de carácter agrícola, manteniendo un pequeño porcentaje de uso destinado al abastecimiento urbano y la industria. Las aportaciones totales que llegarían de forma natural a la desembocadura en el Río Grande, en Ojinaga, son del orden de 2750Hm^3 . Al tratarse el río Grande de un río internacional, la gestión de las aguas del río Conchos está sometida al cumplimiento de un tratado internacional, lo que introduce en la gestión de un río con una fuerte irregularidad natural, importantes restricciones.

Recursos Hídricos y elementos de regulación El río Conchos cuenta como principales afluentes el río Florido, y el río San Pedro (figura 8.16). El sistema de recursos hídricos que los compone cuenta con cinco embalses principales, cuyas capacidades aparecen descritas en la tabla 8.8. En ella destaca

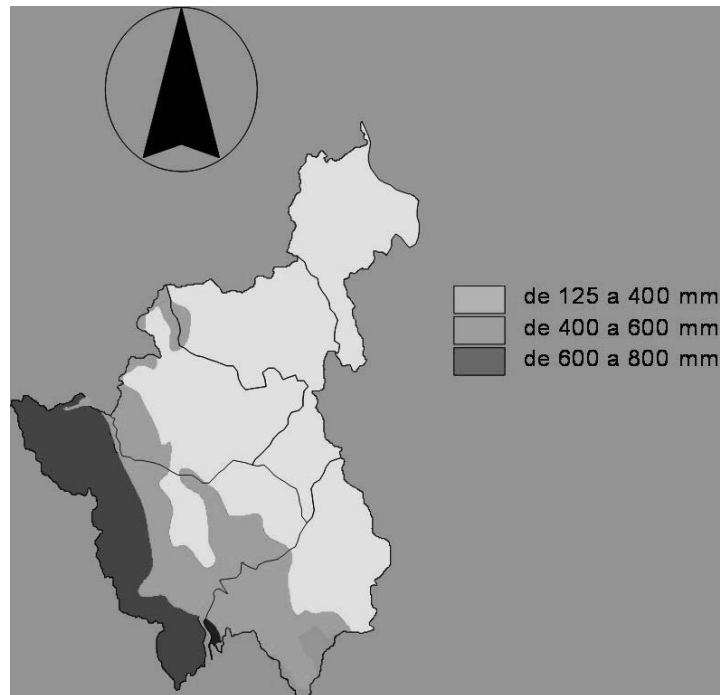


Figura 8.15: Precipitación media anual en la cuenca del río Conchos (México).

el embalse de la Boquilla, con una capacidad superior a los $2700 Hm^3$, el cual posibilita la regulación interanual en la cuenca. El menor de los embalse es el de Pico del Águila, que tiene una reducida capacidad de regulación, y es fundamentalmente utilizado en la laminación de avenidas. El resto de embalses tienen fundamentalmente una función reguladora intranual. Para el análisis de los recursos en la cuenca se cuenta con series históricas de precipitaciones mensuales y temperaturas medias en 11 estaciones sobre la cuenca, en el período de 1954 a 1998. Asimismo, se cuenta con series de aportaciones mensuales en cada subcuenca principal para el mismo período. Especialmente los datos de aportaciones naturalizadas resultan poco fiables, por lo que ha sido considerado conveniente el ajuste de un modelo hidrológico en la cuenca para el análisis de sus recursos hídricos, que también permita la generación de series sintéticas de mayor longitud.

Para ello, en primer lugar han sido analizados los datos de precipitaciones. Se han obtenido para cada subcuenca las precipitaciones medias areales, obteniéndose 6 series de precipitaciones, sobre las cuales se ha ajustado un modelo estocástico. Se ha encontrado adecuada la distribución gamma para cada una de las poblaciones de precipitaciones mensuales. Se han ajustado la variación estacional de cada uno de sus parámetros de cada serie mediante funciones sinusoidales (ajuste de Fourier). Resulta importante en este tipo de series de precipitaciones en regiones áridas el tratamiento de los ceros. Ello es notorio en las series de las subcuencas del Emb. de San Gabriel,



Figura 8.16: Cuenca y Red Hidrográfica de río Conchos (México)

Pico del Aguila y Ojinaga, con probabilidades de precipitación nula en los meses de primavera superior al 10%. Para tenerlo en cuenta, ha sido truncada su función de distribución, asignándole una probabilidad mensual de ocurrencia de precipitación nula, siguiendo los valores distintos de cero la distribución gamma (Salas, 1992). Para la fijación del valor de esta probabilidad de cero, dada su variación estacional, ha sido obtenida mediante ajuste de una función sinusoidal.

El análisis de las propiedades de autocorrelación y correlación cruzada de las series se realiza previa transformación de sus valores a los de una distribución normal estándar equivalente. La aplicación del análisis de componentes principales a la matriz de varianzas-covarianzas de las series estandarizadas permite descomponer la realización de las series inter-correlacionadas, en un conjunto de series no inter-correlacionadas obtenidas mediante transformación lineal. Sobre cada una de las series no inter-correlacionadas, compuestas por los pesos en las componentes principales, se analiza su estructura de au-

Subcuenca	Capacidad Almacenamiento (Hm^3)	Superficie subcuenca (km^2)	Aportaciones medias ($Hm^3/año$)	Precipitación media ($mm/año$)
Emb. San Gabriel	245	15.4	165	495
Emb. Pico del Aguila	45	15.4	165	495
Emb. La Boquilla	2745	208	1250	370
Emb. Francisco I. Madero	298	107	394	414
Emb. Luis L. Leon	260	243	497	422
Ojinaga	-	87	277	298

Tabla 8.8: Embalses y subcuencas principales en el sistema del río Conchos.

tocorrelación, siguiendo tres de las series un modelo autoregresivo de orden 1, y resultando el resto no autocorrelacionadas. A partir de este análisis, adoptando los modelos AR(1) de las series autocorrelacionadas y la transformación lineal inversa, resultante del análisis de componentes principales, es posible la generación de series sintéticas de precipitaciones transformadas a una variable normal estándar. Para la obtención final de las series de precipitaciones, se aplica la transformada inversa desde poblaciones normales estándar, a las poblaciones gamma estacionales, en algunos casos truncadas. El modelo es validado mediante la comparación entre los valores de la serie histórica y 100 series sintéticas de la misma extensión de los siguientes estadísticos:

- Media anual en cada subcuenca
- Desviación típica anual en cada subcuenca
- Sesgo anual en cada subcuenca
- Rango anual en cada subcuenca
- Autocorrelación anual en cada subcuenca
- Intercorrelación anual entre subcuencas
- Autocorrelación mensual en cada subcuenca
- Intercorrelación mensual entre subcuencas

Con el objeto de completar el modelo hidrológico, se ha realizado un modelo de precipitación escorrentía. Para ello se opta por un modelo agregado en cada subcuenca, que permita la generación de aportaciones a cada embalse. Sin embargo, dada la incertidumbre en la fiabilidad de los datos de aportaciones, el ajuste de los modelos de cada subcuenca se realiza agregando el conjunto de las aportaciones aguas arriba, ajustando el modelo desde

aguas arriba hacia aguas abajo, y comparando las aportaciones totales naturales que llegarían a cada embalse o punto de control. De este modo los errores de restitución al régimen natural de las aportaciones son filtrados. Como modelos de precipitación-escorrentía se analizan el modelo abcd (Thomas, 1981; Thomas *et al.*, 1983) y el modelo de Téméz (Téméz, 1977). Ambos tipos de modelos son ajustados para el conjunto de puntos de control. En el ajuste es utilizada la metodología propuesta por Sankarasubramanian and Vogel (Sankarasubramanian & Vogel, 2002) para el ajuste inicial, con una posterior homogeneización de los parámetros en las diferentes subcuencas cuando estos son próximos. Para la estimación de la evapotranspiración potencial en el modelo, son utilizados los datos de temperaturas de las estaciones meteorológicas, suponiendo un ciclo anual que se repite, a partir de los cuales se obtienen las evapotranspiraciones potenciales haciendo uso del procedimiento propuesto por la FAO ((Allen *et al.*, 1998)). Los resultados muestran similares aptitudes al ajuste por ambos modelos, si bien el modelo de Téméz parece mantener un mejor ajuste, especialmente en la modelación de la aportaciones en meses de reducida precipitación, conseguido mediante la modelación de la componente subterránea mediante un embalse lineal. Así se adopta el modelo de Téméz para cada una de las subcuencas, con los siguientes parámetros:

Subcuenca	Hmax (<i>mm</i>)	C (-)	Imax (<i>mm/mes</i>)	α (<i>1/mes</i>)
Emb. San Gabriel	180	0.3	10	0.05
Emb. Pico del Aguila	180	0.3	10	0.05
Emb. La Boquilla	156	0.3	20	0.10
Emb. Francisco I. Madero	242	0.3	10	0.05
Emb. Luis L. Leon	290	0.3	80	0.10
Ojinaga	142	0.3	80	0.10

Tabla 8.9: Parámetros del modelo de Téméz para cada una de las subcuencas principales en del río Conchos.

Las demandas y el tratado internacional EEUU-México de las aguas del río Bravo Las demandas a las que atiende el sistema de RRHH del río Conchos son en un 90 % demandas agrícolas. El resto lo forman la demanda urbana y la demanda industrial. En este análisis del sistema, dada la menor importancia de las demandas urbanas e industriales, y sobre todo los pocos datos con los que se cuenta sobre su valor, van a ser tenidos en cuenta únicamente en el análisis las demandas agrícolas. En el caso del uso agrícola, el sistema abastece a tres zonas principales de riego: el distrito de riego de las río Florido, las Delicias, y Bajo río Conchos (figura 8.17). En la estimación de sus demandas, se cuenta con los registros de áreas cultivadas por tipo de cultivo en cada distrito en el período 1970-2001. En él se observa una variación importante en el área dedicada a cada tipo de cultivo a lo largo del período, debido a diversos condicionantes. Con el fin de evaluar el volumen total que en media cada distrito demanda y su distribución estacional, son obtenidos para cada distrito la extensión media que se dedica a cada cultivo en función de la estación del año. Con ellos, y asignado a cada cultivo su curva correspondiente de coeficientes de cultivos, es posible la obtención del coeficiente de cultivo agregado por cada distrito de riego y mes del año.

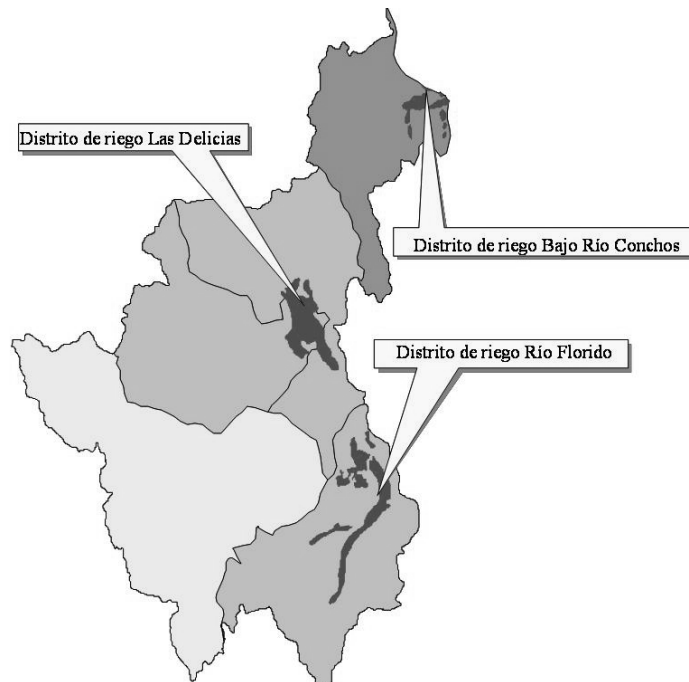


Figura 8.17: Distritos de riego en la cuenca del río Conchos.

	OC	NV	DC	EN	FB	MR	AB	MY	JN	JL	AG	SP
Kc	0.24	0.32	0.44	0.51	0.56	0.42	0.13	0.25	0.46	0.51	0.49	0.33
Área (ha)	1106	1472	2016	2367	2572	1932	606	1150	2126	2371	2267	1500

Tabla 8.10: Coeficientes de cultivos y área media ocupada por mes en el distrito de riego de río Florido.

	OC	NV	DC	EN	FB	MR	AB	MY	JN	JL	AG	SP
Kc	0.21	0.32	0.49	0.59	0.65	0.46	0.13	0.22	0.34	0.40	0.40	0.30
Área (ha)	14226	21426	32862	39647	43274	30607	8776	14902	22947	26684	26664	19708

Tabla 8.11: Coeficientes de cultivos y área media ocupada por mes en el distrito de riego de Las Delicias.

A partir de los coeficientes de cultivos, la evapotranspiración potencial y la precipitación producida en cada mes es posible la estimación del déficit de humedad de la planta (ec. 8.7). Asumiendo una eficiencia total en el transporte, distribución y sistema de aplicación del riego, $Ef = 25\%$, se obtiene el volumen total demandado por cada distrito de riego en cada mes:

$$Def = \text{máx}(0, ETP_0 \cdot K_C - P) \quad (8.7)$$

$$D = \frac{Def \cdot A}{Ef} \quad (8.8)$$

donde A es el área de cada distrito de riego. Este proceso de cálculo de las demandas de riego da lugar a una demanda variable cada mes, que depende de la precipitación producida sobre cada distrito de riego, considerándose que la evapotranspiración potencial (ETP_0) es constante con variación mensual para cada año. Los valores de la demanda media anual que son obtenidos para cada distritos se indican en la tabla 8.13. los valores obtenidos se validan con los valores publicados en balances anuales (Kelly, 2001).

Además de la demanda agrícola hasta aquí analizada, el sistema del río Conchos tiene como demanda la impuesta por el tratado bilateral entre EEUU y México de 1944 sobre el reparto de las aguas de las cuencas transfronterizas entre los dos países. La cuenca del río Conchos se encuentra de forma integra en territorio mexicano, sin embargo, al ser este río tributario

	OC	NV	DC	EN	FB	MR	AB	MY	JN	JL	AG	SP
Kc	0.18	0.26	0.40	0.47	0.50	0.35	0.16	0.26	0.50	0.57	0.52	0.27
Área (ha)	980	1406	2165	2569	2727	1924	851	1406	2707	3141	2858	1500

Tabla 8.12: Coeficientes de cultivos y área media ocupada por mes en el distrito de riego de Bajo Río Conchos.

	Demanda de riego (Hm^3)
DR río Florido	84
DR Las Delicias	1224
DR Bajo río Conchos	91

Tabla 8.13: Demanda media anual computada por cada distrito de riego en el río Conchos.

del río Bravo, las derivaciones al río Bravo están reguladas por este tratado. En concreto el tratado impone las siguientes consideraciones a este respecto:

- México deberá derivar desde el río Conchos al río Grande al menos un tercio de las aportaciones que se producirían de forma natural, con un volumen mínimo de $432Hm^3/año$. Esta condición deberá cumplirse en períodos quinquenales.
- Si bajo la realización de una sequía extraordinaria no fuera posible la derivación del volumen establecido en el punto anterior, el déficit producido deberá reponerse en el siguiente quinquenio.

Las restricciones que impone este tratado en el período de tiempo sobre los que debe comprobarse el balance de aguas, proporciona limitaciones al aprovechamiento del recurso en períodos secos en los distritos de riego del sistema. Incluso el tratado contempla la posibilidad de producirse un episodio de sequía extraordinaria, ante la cual permite la transposición del déficit. Ello sin embargo, queda indefinido, pues no está especificado el concepto de sequía extraordinaria dentro del tratado. Así mismo, tampoco existe una definición universalmente aceptada en la que se declare la situación de sequía extraordinaria. En ese sentido en los volúmenes derivados históricamente al río Grande se observa que, por ejemplo, en el quinquenio terminado en octubre de 1997 existía un déficit de $1240Hm^3$. Este déficit creció en $480Hm^3$ en el balance de febrero del 2000 (Kelly, 2001). Ante tal situación, y la indefinición de la situación de sequía extraordinaria, no es posible la correcta comprobación del cumplimiento del tratado. En la siguiente sección se acompaña un análisis del sistema que permite proponer un criterio operativo para la definición de la situación de sequía extraordinaria a partir del IFS.

El IFS y la condición de sequía extraordinaria

Las condiciones a las que está comprometida la gestión de los RRHH en el sistema del río Conchos hace necesaria la construcción de un modelo de gestión del sistema y su análisis para el mejor conocimiento de las implicaciones que puede tener en la gestión. La demanda de agua correspondiente al

cumplimiento del tratado internacional a los efectos de este análisis son las de mayor prioridad, siendo el objeto de una correcta gestión el cumplimiento de este tratado con el mejor aprovechamiento de los recursos restantes en la cuenca del conchos. Resulta entonces de interés la delimitación de aquellas situaciones en las que sería deseable la transposición de un déficit no satisfecho en un período quinquenal al período siguiente. Para el análisis de cuando estos períodos pueden producirse se plantea la gestión del sistema como un problema de optimización.

La gestión del sistema puede simplificarse para convertirla en un problema de optimización lineal, de modo que se disponga de herramientas de resolución eficientes. Dado una serie de aportaciones a cada subcuenca o embalse, y de demandas de los distritos de riego, el objetivo del problema de optimización debe ser el cumplimiento del tratado bilateral, y el mejor aprovechamiento de los recursos para la satisfacción de las demandas de riego. Así se propone el esquema representado en la figura 8.18, donde el esquema de gestión del sistema se repite cada mes, con un balance de cada quinquenio previo para el vertido a EEUU. Además dada la posibilidad que contiene el tratado de transponer déficit al quinquenio siguiente, existe una conexión entre quinquenios consecutivos que posibilita el ajuste del balance. La función objetivo lineal que permite la obtención de la gestión buscada vendría dada por la siguiente expresión:

$$\begin{aligned} O(S_{DrRioFlorido}, S_{DrLasDelicias}, S_{DrBajoRioConchos}, S_{EEUU}) &= \\ &= \sum_i \sum_t S_{Dr,i,t} \cdot a + \sum_t S_{EEUU,t} \cdot b - \sum_t T_t \cdot c - \sum_t Vert_t \cdot d \quad (8.9) \end{aligned}$$

donde $S_{Dr,i,t}$ es el suministro en el mes t al distrito de riego i , $i = RioFlorido, LasDelicias, BajoRioConchos$, $S_{EEUU,t}$ es el volumen derivado al río Bravo en el mes t para el cumplimiento de las condiciones del tratado, $Vert_t$ es el volumen excedente derivado al río Bravo en el mes t que no es necesario para su cumplimiento y no es posible almacenar en los embalses, y T_t son los volúmenes de déficit quinquenales transpuestos. Los parámetros a, b, c, d definen la función objetivo, y la fijación de sus valores viene dada por las condiciones que se desean cumplir:

- El cumplimiento de las condiciones del tratado bilateral debe ser prioritario frente al suministro a las demandas de riego

$$a < b$$

- En el caso de producirse déficit en el suministro a la demanda de regadío en un quinquenio y excedente en el quinquenio consecutivo,

debe resultar más beneficioso la transposición de un déficit en el primer quinquenio, para su satisfacción en el siguiente:

$$b - d < a + b - c < 2 \cdot b$$

- No debe admitirse la transposición de déficit en más de un quinquenio consecutivo

$$b + a - 2 \cdot c < -d$$

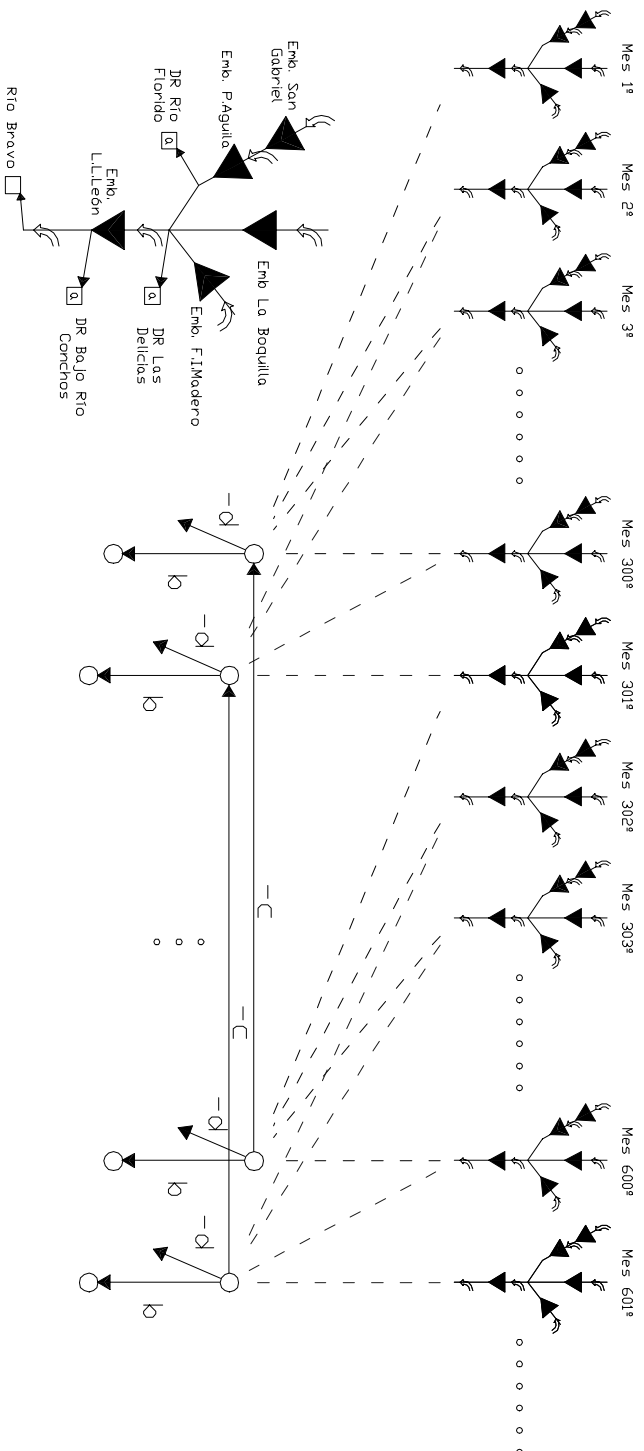


Figura 8.18: Esquema del Sistema de RRHH del río Conchos y esquema de optimización a dos escalas, mensual y quinquenal, con posibilidad de transposición de déficit entre balances quinquenales.

Cualquier combinación de parámetros que cumplan con estas restricciones simultáneamente producirían una función objetivo (ec. 8.9) que permite la búsqueda de aquellas situaciones donde la transferencia de déficit es conveniente. Por supuesto, la simplificación del problema hasta definirlo como un problema de programación lineal supone funciones de pérdidas/beneficios lineal. Esta simplificación no se ajusta a la realidad, ya que las demandas de regadío permiten pequeños déficit con reducidas pérdidas, pero las pérdidas aumentan exponencialmente al aumentar los déficit. Por otro lado, técnicas de linealización de funciones de pérdidas cóncavas permitirían poner de manifiesto este aspecto. Sin embargo, las técnicas de linealización aumentan el número de variables a optimizar y aquí una aproximación lineal se considera suficiente a estos efectos, donde lo que se busca es definir cuándo serían convenientes la transposición de déficit.

Para aplicar la metodología se han generado de forma sintética un total de 10,000 años de aportaciones y correspondientes demandas de riego, en las cuales han sido optimizado su gestión en períodos de 1000 años. Los valores adoptados de los parámetros que definen la función objetivo han sido: $a = 0,9$, $b = 1$, $c = 1,25$, $d = 0,5$; con ellos se comprueba las condiciones anteriores. A partir de los resultados de la optimización pueden delimitarse los períodos en los que se transpone un déficit al siguiente quinquenio. Cuando se comparan estos déficit con los valores del IFS del instante en el que se produce, calculado sobre la serie de aportaciones totales naturales del río Conchos, se tiene la figura 8.19.

A partir de la figura se puede afirmar que los principales déficit se producen tras la ocurrencia de sequías con un IFS superior a 70 años. Para IFS menores, los déficit que se producen son de menor consideración. Por tanto, desde un punto de vista operativo, se puede establecer como definición de sequía extraordinaria aquella que produce un IFS superior a 70-100 años. Este límite es absolutamente operativo, y relativo al problema en el sistema del río Conchos, de acuerdo al escenario de demandas considerados. Una variación de estas demandas podría hacer variar la posición del límite operativo. Este ejemplo ilustra la posible aplicación del índice en el establecimiento de una definición para el término de sequía extraordinaria. Por tratarse de un índice de naturaleza probabilista, relacionado con la recurrencia del evento, el índice permite de forma directa asociar una cuantía del grado de excepcionalidad del evento. Como desventaja puede encontrarse que el índice responde a un criterio de caracterización de sequías que no está de forma directa y explícitamente relacionado con los volúmenes de aportaciones producidos, aunque sí mantiene una relación clara de forma indirecta. Puede ocurrir que la definición de un criterio sobre otra magnitud, como pudiera ser la aportación media en el último quinquenio, produjera un criterio de definición de la sequía operativa más directamente relacionado con el

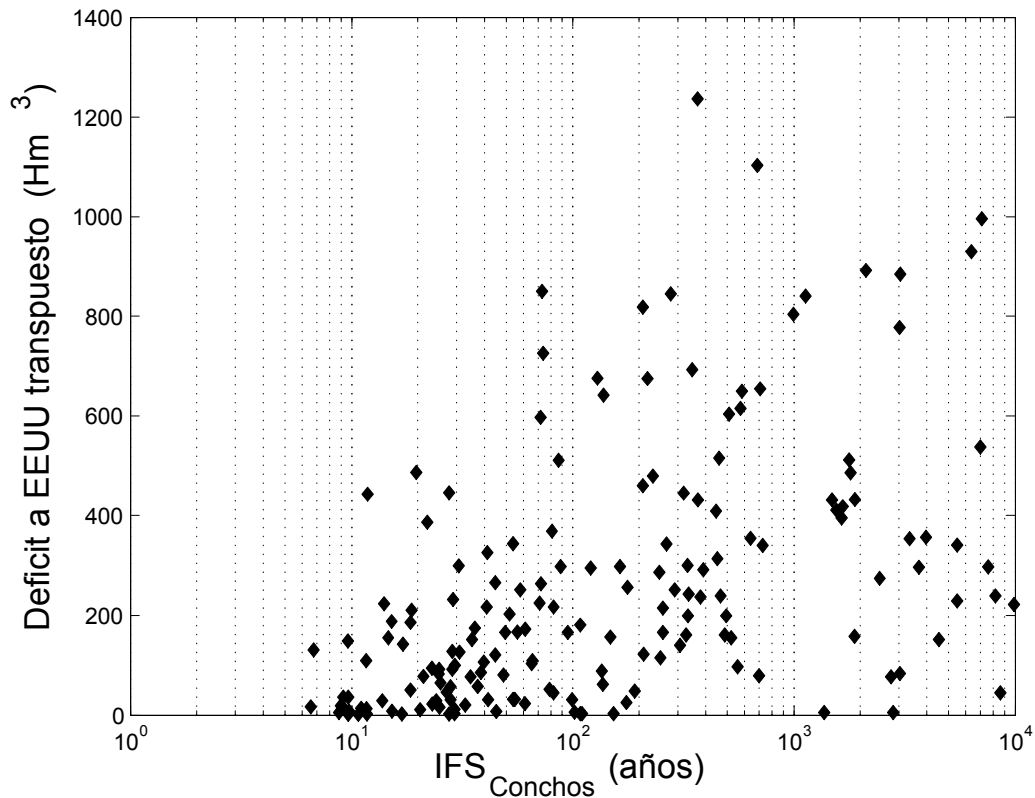


Figura 8.19: Índice de Frecuencia de Sequías producidos en la situación en los que se ha transpuesto un déficit a EEUU al quinquenio consecutivo en el sistema del río Conchos.

problema. Sin embargo, esto presentaría la desventaja de la particularidad de la definición, de difícil comparación con otros casos.

La política de gestión óptima

El análisis anterior permite la definición de situaciones en las que es permisible la transposición de déficit quinquenales. Aquí se analiza, a partir de estos resultados, la política de gestión óptima que permitiría de forma práctica la gestión de los RRHH en el sistema del río Conchos. En este sentido, se plantea una gestión del suministro a las demandas de regadío similares a las analizadas en el caso del sistema del río Júcar, con restricciones al suministro en función del IFS (ec. 8.2). Para ello se propone unas reglas de operación con la zona deficitaria de pendiente función del IFS (ec. 8.1) para cada uno de los suministros a las demandas de riego, *dem = Río Florido, Las Delicias, Bajo Río Conchos*. La forma de ges-

tionar los embalses, una vez definidos los volúmenes objetivos a suministrar, es a partir de las siguientes reglas:

- Para el suministro de la demanda de río Florido se hace uso del sistema de embalses en serie formado por los embalses de San Gabriel y Pico de Águila, este último, situado aguas abajo, se vacía siempre en primer lugar.
- El suministro a Las Delicias se realiza desde los embalses de la Boquilla y Francisco I. Madero, vaciando en primer lugar este último embalse, por tener mucho mayor probabilidad de llenado, y dejando el embalse de la Boquilla en reserva.
- La demanda del Bajo Río Conchos se satisface en primer lugar por las aguas del embalse de Luis L. León, situado aguas abajo del sistema y por lo tanto el primero que debe ser vaciado, y seguidamente por los embalses de la Boquilla y Francisco I. Madero.
- En cuanto al suministro de los volúmenes derivados a EEUU se realiza en primer lugar por el embalse de Luis L. León, seguido por el sistema de los embalses de la Boquilla y Francisco I. Madero, y si es necesario por el último sistema de San Gabriel y Pico de Águila.

Como función objetivo a optimizar para la fijación de los parámetros que definen la ley, y al igual que ocurría en caso del sistema del río Júcar, se utiliza el producto de las frecuencias de cumplimiento de las condiciones del criterio de Utah sobre cada una de las demandas. Definida la regla de operación, y utilizando en primer lugar la misma función de restricciones para todas las demandas, se busca el óptimo haciendo uso del algoritmo de SCEM-UA. La figura 8.20 muestra los resultados del muestreo de la función objetivo y la función de distribución posterior resultante.

Como puede observarse el parámetro que define la regla tendría marcado un valor óptimo entorno a $C_{Conchos} = 0,3$. Si se realiza el mismo proceso de optimización utilizando unas reglas de operación con restricciones de pendiente constante, α , se obtendría los resultados que muestra la figura 8.21. Puede apreciarse que bajo esta política de operación el valor alcanzado de la función objetivo en el punto óptimo es menor que el que se producía con restricciones función del IFS.

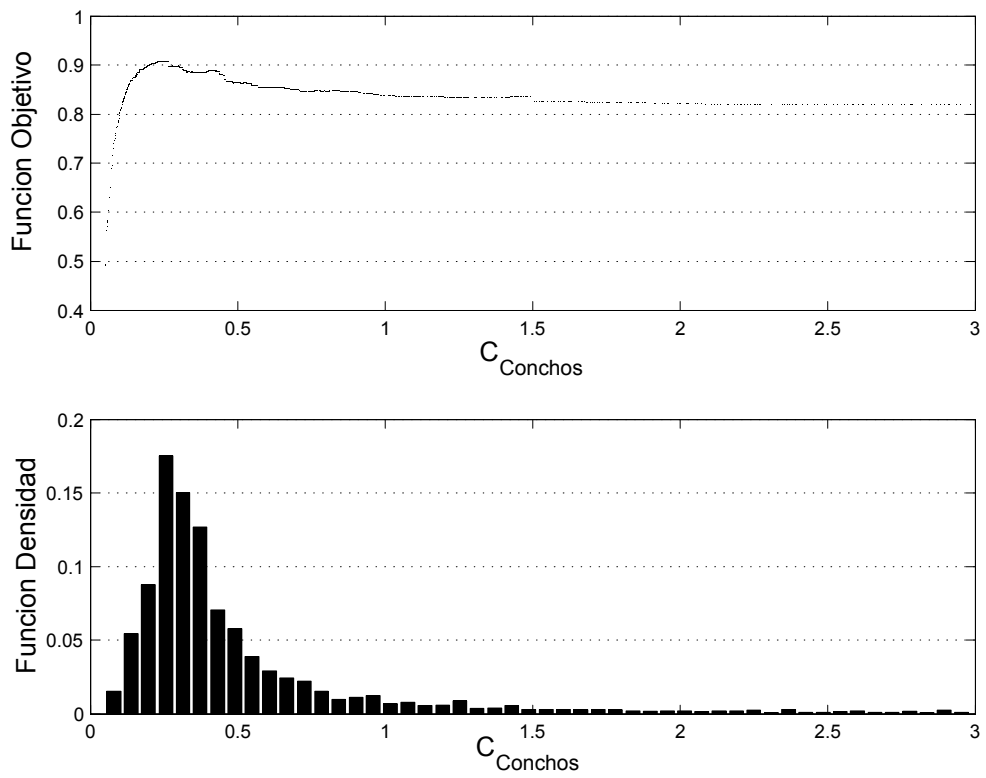


Figura 8.20: Resultados de la optimización de la regla de operación del sistema del río Conchos, con restricciones función del IFS y un único parámetro.

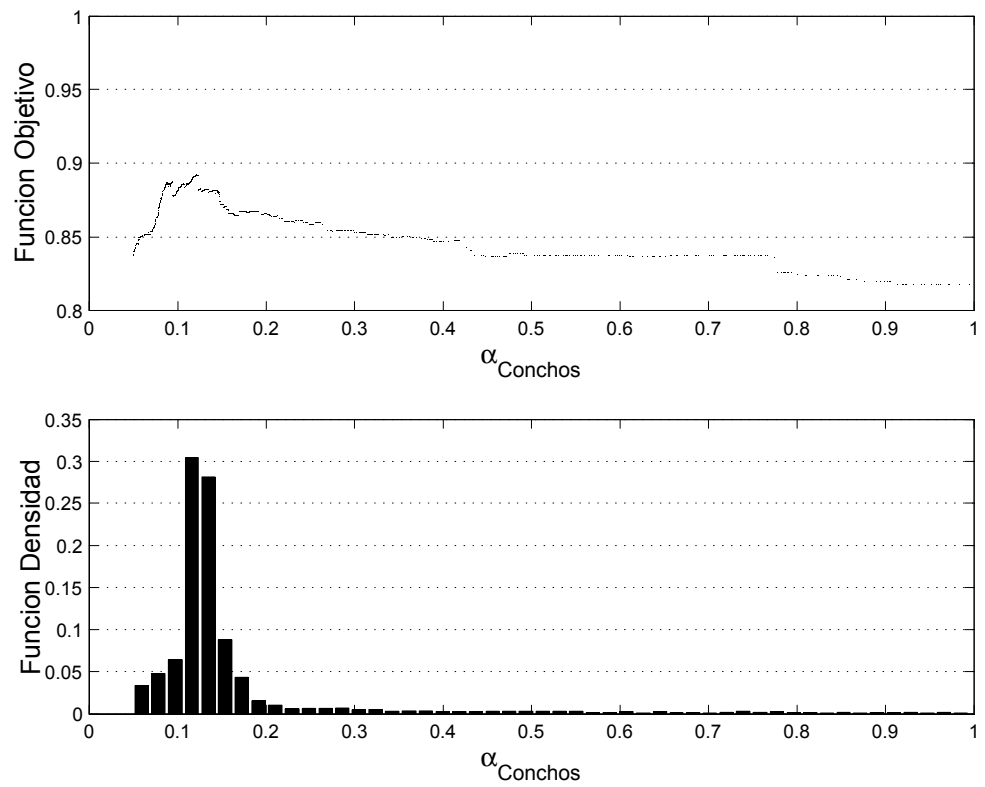


Figura 8.21: Resultados de la optimización de la regla de operación del sistema del río Conchos, con restricciones constantes y un único parámetro.

A partir de esta aproximación al óptimo, se analiza el problema completo, asignando un coeficiente de restricción distinto a cada demanda de riego, obteniendo los resultados que muestran las figuras 8.22 y 8.23. En este caso se aprecia que la sensibilidad en la definición de la regla se encuentra principalmente en el parámetro que define las restricciones al distrito de riego de Las Delicias. Ello se ajusta a la idea de que este distrito es el de mayor importancia en el sistema, con una demanda superior al 80 % de la demanda total. En el resto de distritos, la imposición de restricciones no produce ventajas apreciables, manifestándose únicamente empeoramiento cuando las restricciones son muy altas. Debe notarse en cualquiera de los escenarios simulados que siempre se garantiza en primer lugar el cumplimiento del tratado bilateral, de acuerdo a la definición de sequía extraordinaria establecida. De aquí podría concluirse que la política de operación óptima, de acuerdo a los criterios objetivos establecidos, se alcanza introduciendo una restricción al suministro de todas las demandas con un coeficiente de restricción de $C_{Conchos} = 0,3$.

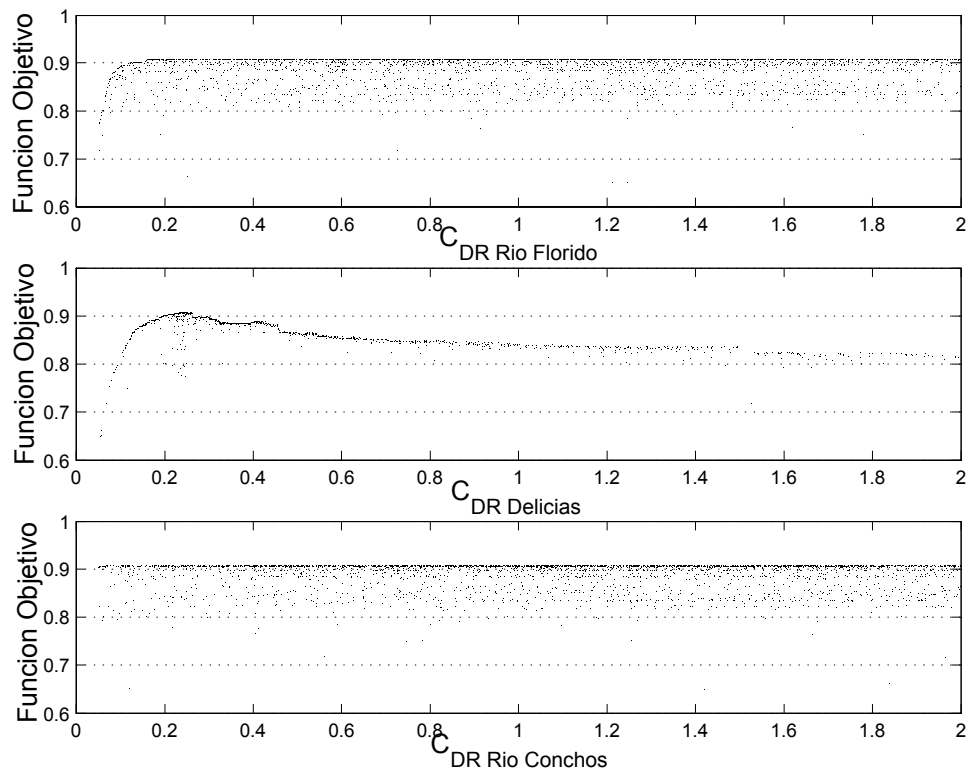


Figura 8.22: Muestreo de la funciones objetivos en el proceso de optimización de la regla de operación del sistema del río Conchos, con restricciones función del IFS y tres parámetro.

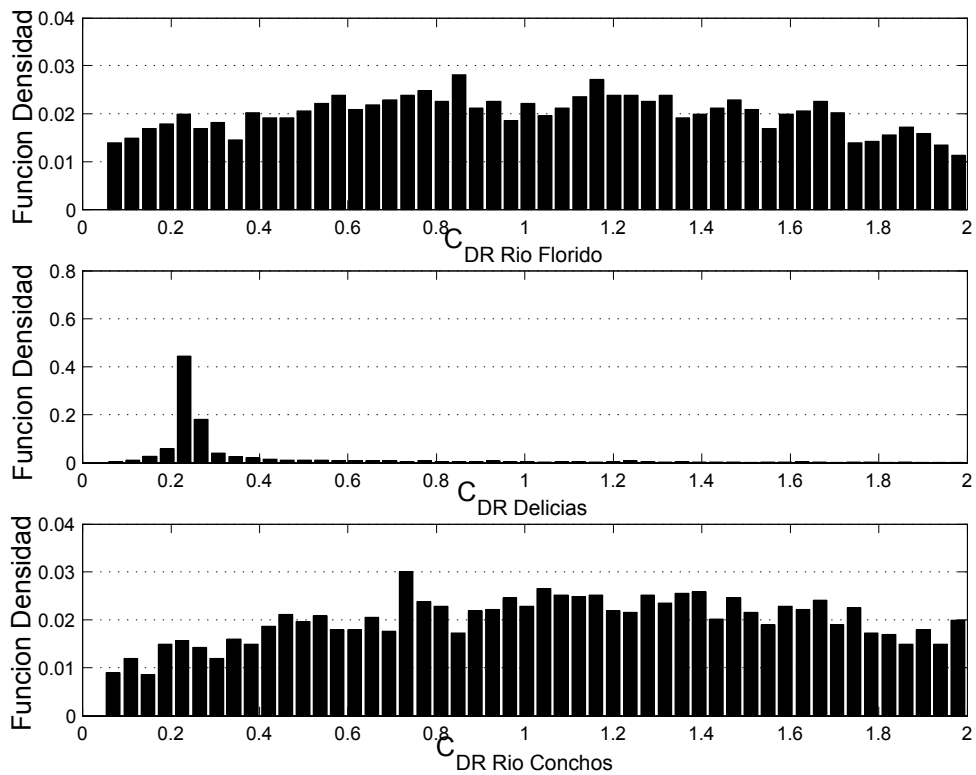


Figura 8.23: Función de distribución posterior en el proceso de optimización de la regla de operación del sistema del río Conchos, con restricciones función del IFS y tres parámetro

Capítulo 9

Conclusiones y futuras líneas de desarrollo

La primera y principal conclusión de esta Tesis Doctoral, y cuya formalización y justificación vertebran el trabajo de investigación y definen su contribución, se encuentra en la adecuación de la caracterización estocástica para el análisis y evaluación de las realizaciones espacio-temporales de los eventos hidrológicos extremos de sequías.

El concepto de sequía recoge las ideas de una deficiente presencia de agua excepcional, referida a una variabilidad considerada normal, que persiste en el tiempo y se muestra de forma extensa en el espacio. De los efectos de este fenómeno aleatorio extremo se derivan una amplia variedad de definiciones de sequía para cada lugar, condiciones y uso concreto. El abanico de definiciones tiene como denominador común la realización de un evento extraordinario. La posibilidad de uniformizar en una escala general todos los criterios de definición de sequía y tratar el fenómeno de forma coherente con su naturaleza aleatoria está por tanto en la caracterización del grado de excepcionalidad con el que se presenta el evento.

Para este fin ha sido desarrollada una metodología de caracterización de desviaciones persistentes extremas en secuencias de realizaciones de variables aleatorias consideradas independientes. El criterio de caracterización se basa en el uso de la función de persistencia extremal como medida de la aproximación persistente de las realizaciones hacia la región extremal considerada. A partir de esta función, se analiza la relación existente entre el valor de la función y la recurrencia media con la que se alcanza un valor similar o más extremo. Esta relación mantiene una expresión analítica aproximada dependiente de la función de probabilidad acumulada de persistencias extremas. La aproximación es asintótica, con aproximación creciente para valores más extremos. Cuando la hipótesis de independencia no se mantiene, la aplicación de la metodología pasa por la resolución de

expresiones integrales múltiples, donde en general no se encuentra una solución analítica. En estos casos, las expresiones resultantes deben ser resueltas a partir de métodos numéricos de integración múltiple de gran dimensión. El coste computacional que esto puede suponer limita la aplicabilidad de la metodología a señales con significativa correlación temporal. En los casos en los que la dependencia temporal es pequeña, la metodología puede aplicarse evaluando de forma numérica la expresión que proporciona la función de persistencia extremal, y adoptando de forma aproximada las expresiones analíticas derivadas para el caso independiente que relacionan el valor de la función con la recurrencia media. Esta limitación supone en la práctica que la metodología sea aplicable hasta una determinada escala temporal, por debajo de la cual la variable resultante de una escala temporal menor (menor paso temporal) mantiene una correlación inadmisiblemente alta.

Esta metodología de caracterización estocástica se adopta para la definición del Índice de Frecuencia de Sequías. El índice caracteriza en cada instante la situación de sequedad a partir de evaluar la recurrencia media de las secuencias terminadas en ese instante respecto a la desviación persistente hacia la región extremal correspondiente con el fenómeno de sequía. El período de retorno de la secuencia que produce un mayor valor de la función de persistencia extremal es el asignado a ese instante. El índice permite evaluar en cada momento la situación de sequía, utilizando como escala la frecuencia media de recurrencia. El IFS, así definido, proporciona una cuantificación natural y directa del grado de excepcionalidad del evento que ha tenido o está teniendo lugar. El análisis se dirige sobre el fenómeno que se produce, y no sobre el efecto en un uso particular. Esto proporciona una herramienta para la homogeneización de las definiciones de sequías en cada aplicación concreta, función también de la vulnerabilidad del sistema usuario.

El IFS, a diferencia de otros índices general de sequía como el PDSI, no contiene en su definición ningún criterio arbitrario ni está diseñado para su aplicación a una región y fin concreto. La escala independiente y directamente relacionada con la frecuencia del fenómeno validan su aplicación no sólo en comparaciones de episodios producidos en un lugar y sobre una variable determinada, sino también sobre realizaciones producidas sobre lugares diferentes, sean cual fueren sus regímenes de variación (e.g. la climatología en el caso de analizar precipitaciones). Incluso el índice permite la comparación de los efectos producidos sobre las diferentes variables hidroclimáticas o elementos del ciclo hidrológico, por expresarse en una escala independiente. Así permite comparar la recurrencia de las sequías meteorológicas, con la de sequías hidrológicas analizando la serie producida de aportaciones, almacenamientos en acuíferos o lagos.

La caracterización de sequías a partir del IFS frente a otras metodologías

estocásticas, como se encuentra en aplicaciones de la Teoría de Rachas, presenta su mayor adecuación en análisis generales de sequías. La indefinición de una magnitud, o un conjunto de magnitudes, que resulten de especial interés en un análisis general para la caracterización de las sequías puede provocar importantes discrepancias en los resultados de los modelos estocásticos de recurrencia. Además, la aplicación de la Teoría de Rachas requiere de la definición del estado de sequía, la cual si no viene impuesta al análisis puede producir fuerte sensibilidad en los resultados frente a variaciones en su definición.

La escala temporal con la que se realiza un análisis de sequías, definida por el paso temporal utilizado en la serie, debe ser ajustada de acuerdo a los objetivos perseguidos con el análisis. El aquí llamado efecto "lupa" produce que análisis con una pequeña escala temporal (reducido paso temporal) puedan producir una incorrecta detección de una sequía con desarrollo temporal muy prolongado. De modo inverso, el uso de una gran escala temporal (largo paso temporal) produce que desviaciones cortas puedan no ser detectadas, o pobremente caracterizadas. Ésto se recoge en la aplicación del índice SPI mediante la repetición del análisis estadístico para diferentes escalas temporales. Sin embargo, este modo de operar presenta el inconveniente que valores resultantes del SPI correspondientes a análisis con escalas temporales diferentes no son comparables, pues aunque fueran iguales, cada uno de ellos puede corresponder a frecuencia de recurrencia diferentes. Es este sentido, un mejor tratamiento se consigue con el uso del IFS, que cuantifica el carácter excepcional de una sequía en términos de frecuencia de recurrencia, y no en la forma de probabilidad de ocurrencia. El índice presenta menor sensibilidad frente a cambios en la escala temporal, aunque no pueda considerarse invariante, y se ve afectado por el efecto "lupa". Es este sentido, el IFS es más adecuado que otros índices en análisis donde la escala temporal no esté fijada por los objetivos.

El carácter extenso asociado al concepto de sequía se sitúa también en la idea de sequía como un fenómeno extraordinario. Realizaciones de deficiente presencia de agua sobre diferentes extensiones resultan menos frecuentes cuanto mayor es la extensión ocupada. De este modo, este aspecto de la sequía tiene también su valoración en el marco de una caracterización estocástica. La aplicación del IFS ha sido generalizada para la elaboración de mapas de IFS que permita el análisis del desarrollo espacio-temporal de las sequías de acuerdo a su caracterización estocástica. Esto se concreta en un algoritmo que permite la caracterización de la realización espacial utilizando la función de probabilidad acumulada espacial conjunta. A cada punto del espacio le es asignado el valor del IFS correspondiente a la sub-región que lo contiene, cuya secuencia de valores de la función de probabilidad acumulada espacial conjunta produce un mayor valor de la función de persistencia

extremal. Los mapas de IFS permiten el análisis de la evolución en el tiempo de las sequías y de su desarrollo sobre espacial. El carácter extenso de la sequía puede ser analizado indirectamente mediante el análisis IFS sobre una variable hidrológica que esté influenciada por el desarrollo espacial de la sequía, como pueden ser las aportaciones cuando la región de análisis coincide con una cuenca hidrográfica. Sin embargo, esta simplificación resulta inconveniente cuando la distribución de precipitaciones sobre la cuenca no es homogénea, como puede venir provocado por la orografía.

Las aplicaciones realizadas del IFS son un reflejo del uso potencial del índice en la práctica. La elaboración de los mapas de IFS sobre las precipitaciones anuales en la cuenca del río Colorado (EEUU) es un ejemplo de la capacidad de la metodología para tratar sobre una escala uniforme una amplia variedad de regímenes climáticos. Los mapas permiten analizar la posición que ocupa la sequía en cada momento y el grado de desviación que está teniendo lugar, tratando de forma uniforme las desviaciones que son producidas en climas húmedos de alta montaña, con las producidas en climas desérticos. Además, la comparación con mapas obtenidos a partir del PDSI, o los que resultarían con el control de las condiciones de la vegetación (e.g VIC o WSVI), muestra en los mapas de IFS una mejor representación del carácter persistente de las sequías, resultando los mapas de IFS una adecuada representación del desarrollo del fenómeno.

Por último, el uso particular del índice IFS bajo objetivos concretos, como puede ser la gestión de los RRHH, puede proporcionar significativas ventajas como medida de la gravedad de la situación que está teniendo lugar. Así se pone de manifiesto en el análisis de la gestión de los RRHH en la cuenca del río Júcar. En este sistema de múltiples embalses y múltiples demandas, puede ser definida una política de operación óptima función del valor del índice calculado sobre las aportaciones totales naturales que se producirían en los principales puntos de toma. El análisis de optimización muestra la sensibilidad de la política a la presencia de una situación de sequía. Otro ejemplo de aplicación es el uso del IFS sobre la cuenca del río Conchos (México), donde la gestión de los RRHH está condicionada por el tratado internacional de aguas entre EEUU y México. En este tratado se definen unas condiciones excepcionales de requerimientos de sueltas cuando se produce una situación de sequía extraordinaria, que no es definida. El análisis del sistema del río Conchos muestra el potencial uso del IFS como indicador para la definición de las condiciones de sequía extraordinaria, además de permitir la definición de una política de operación del sistema.

Con todo lo presentado, el trabajo de investigación expuesto en esta Tesis Doctoral permite la aplicación de la caracterización estocástica en el análisis y evaluación de sequías, tanto en su desarrollo temporal como espacial, y proporciona una herramienta aplicable en la práctica sobre los

diferentes campos donde las sequías muestran su efecto. Así pretende ser una contribución al mejor tratamiento de estos fenómenos hidrológicos extremos, que facilite la mitigación de sus consecuencias.

9.1. Futuras líneas de desarrollo

Entre las posibles líneas de desarrollo de este trabajo se encuentran:

- Estudio del efecto de la escala temporal en el análisis de sequías, buscando conocer para cada escala temporal de análisis el tipo de desviaciones persistentes que es posible caracterizar y aquellas que se producen durante un período excesivamente corto o largo para ser correctamente evaluadas.
- Estudio del tratamiento riguroso de señales con manifiesta correlación temporal, de modo que permita la caracterización de sus realizaciones de forma práctica complementando la metodología expuesta.
- Estudio de la sensibilidad de los mapas de IFS al tamaño de la discretización espacial. El tamaño de la discretización tiene notoria influencia en el tiempo de cálculo de los mapas de IFS. Sin embargo, la discretización debe ser lo suficientemente fina para ser representativa de la variedad climática existente en la región de análisis, así como permitir detectar la posición concreta sobre la que se desarrollan las sequías, sin diluir su manifestación en una discretización muy gruesa.
- Conjugación de la metodología de caracterización de desviaciones persistentes extremas y la teledetección o el control de las sequías a partir de su efecto sobre la vegetación, detectado desde la información de satélites meteorológicos. De este modo se buscaría conseguir una representación completa del fenómeno, aprovechando la información distribuida proporcionada por las técnicas de teledetección y la caracterización del desarrollo espacio-temporal de la sequía mediante la caracterización estocástica. Ello permitiría controlar la realización de una sequía reflejando su persistencia y su extensión en el espacio, lo que serviría como sistema de alerta de sequías en su afección a sistemas de mayor inercia como pueden ser aportaciones en grandes cuencas ó acuíferos.
- Estudio de la relación entre los valores del IFS y la cuantía de los daños producidos por los efectos de cada sequía, para un determinado usuario, expresado en términos económicos, por ejemplo. De este modo, si se comprobase una relación aproximadamente biunívoca en

un determinado lugar y aplicación entre los daños provocados y el valor del índice alcanzado facilitarían la implementación del índice en aplicaciones reales, permitiendo plantear el análisis en términos de riesgo.

- Aplicación de la metodología de caracterización de desviaciones persistentes extremas a otras variables aleatorias donde la caracterización de las realizaciones de estos eventos sea de interés. Ejemplos pueden encontrarse en otras disciplinas, como puede ser la aplicación al análisis de ecosistemas acuáticos naturales, donde un descenso prolongado del contenido de oxígeno disuelto en el agua puede producir daños severos en el sistema.

Apéndice A

Métodos numéricos para el cálculo de integrales múltiples

En la literatura pueden encontrarse numerosos métodos para el cálculo numérico de integrales simples (i.e., de dimensión unidad) propias o impropias. Cada método resulta más apropiado dependiendo fundamentalmente de la función a integrar. Así, se puede pasar desde la simple regla del trapecio, poco eficiente en general, hasta los métodos de Cuadratura de Gauss, que componen una familia de métodos en los que bajo ciertas propiedades del integrando y los límites de integración, el método resulta óptimo (Davis & Rabinowitz, 1984). Cuando la integral a resolver es múltiple, de pequeña dimensión y la función integrando presenta una variación suave, la implementación iterativa de los métodos unidimensionales resulta adecuada, respondiendo al Teorema de Fubini, siendo conveniente en estos casos el uso de la notación tensorial.

A medida que la dimensión de la integral crece, el uso iterativo de métodos unidimensionales resulta ineficiente desde el punto de vista computacional, siendo entonces conveniente el uso de métodos numéricos para integrales múltiples de alta dimensión. La frontera a partir de la cual resultan ineficientes los métodos clásicos, y comienzan a resultar más adecuados métodos de integración múltiple de alta dimensión, no es fija, sino que depende de la función a integrar. Asimismo, y desde un punto de vista práctico, depende de la capacidad computacional de cálculo disponible.

El tipo de integrales múltiples que motiva la necesidad de utilizar aquí métodos numéricos de alta dimensión son las que aparecen para la evaluación de la *f.p.e.*, en los casos en los que la variable no es independiente (ec. 5.85), al igual que aparecen cuando es necesaria la evaluación de la *f.a.e.c.* en la construcción de mapas de *IFS* (ec. 7.34). A estas se añaden las ecuaciones integrales que resultan como desarrollo de las *f.p.* de las duraciones $d_{\mathcal{N}}$, $d_{\mathcal{N}}$.

A continuación se presentan una breve introducción a los diferentes métodos de integración múltiple para alta dimensión. De entre ellos destaca por su conveniencia para el tipo de integrales que ha sido necesario resolver en esta Tesis, los métodos de Lattice. Estos son descritos con mayor detalle. Finalmente se incluye la variante aquí adoptada de los métodos de Lattice y el modo en el que ésta es implementada para la resolución numérica de integrales múltiples de *f.d.p.*

A.1. Métodos numéricos en integrales múltiples de alta dimensión

El problema fundamental a resolver por lo métodos de integración múltiple es la resolución numérica de:

$$I = \int_{[0,1]^d} f(x) \cdot dx \quad (\text{A.1})$$

donde d es la dimensión de la integral a resolver. En este problema pueden ser incluidos mediante convenientes cambios de variables integración sobre regiones no rectangulares, integraciones con respecto distribuciones de probabilidad no uniformes, variantes diseñadas para mejorar la precisión como en el caso de periodización (i.e. transformación a una función periódica), etc.

Entre los métodos numéricos considerados de alta dimensión se encuentran el Método de Monte Carlo (*MC*) y los métodos de equidistribución o Quasi-Monte Carlo (*QMC*). La diferencia fundamental entre los métodos de *MC* y *QMC* se encuentra en la generación de los puntos donde es evaluada la función. Mientras que en el método de *MC*, la generación se realiza con funciones pseudo-determinísticas, los métodos de *QMC* están basados en la generación de una familia determinística de puntos en el supercubo $[0, 1]^d$. Los métodos de *QMC* aseguran las propiedades de equidistribución en sus generaciones, mientras que el método de *MC* no las cumple en la práctica, presentado sus generaciones huecos o agrupaciones de mayor densidad de puntos dentro del supercubo. Así, los métodos de *QMC* resultan más eficientes que el método de *MC* para grandes simulaciones; sin embargo, a medida que la dimensión d aumenta, las diferencias disminuyen. En problemas de muy alta dimensión, la implementación de métodos como *MC* o *QMC* resulta ineficiente, resultando muy costosa la construcción de generaciones equiprobables. Para esos casos aparecen técnicas como el Método de Muestreo Latin en el Supercubo (*Latin Supercube Sampling, LSS*) (Owen,

1998) que descomponen el problema en un conjunto de problemas de menor dimensión en los que es aplicable técnicas de alta dimensión. A continuación se describen cada una de estas familias con mayor detalle.

A.1.1. El método de Monte Carlo

El método de Monte Carlo se basa en la aproximación de la integral I (ec. A.1) mediante:

$$\hat{I} = \hat{I}_n = \frac{1}{n} \cdot \sum_{i=1}^n f(X_i) \quad (\text{A.2})$$

seleccionando los puntos de muestreo de la función $X_i = (X_i^1, \dots, X_i^d) \in [0, 1)^d$ de forma aleatoria desde una distribución uniforme en el supercubo $[0, 1)^d$. En casos especiales puede no darse un peso uniforme a todos los puntos, asumiendo determinadas propiedades de la función como puede ser su forma periódica.

La forma en la que habitualmente son generados el conjunto de puntos aleatorios es mediante el uso de funciones de generación pseudo-aleatorias, las cuales resultan implementadas en algoritmos de generación. Estos algoritmos arrojan valores pseudo-determinísticos en función del valor de la semilla, repitiéndose una secuencia determinística. Sin embargo, si la longitud de la secuencia es muy grande, en comparación con el número de realizaciones solicitadas al algoritmo, el resultado práctico se percibe como números aleatorios (la probabilidad de repetición de la semilla es pequeña). La secuencia debe de cumplir ciertas propiedades para asegurar su percepción como equidistribuidas, para lo cual existen convenientes test estadísticos.

Bajo condiciones medias de comportamiento de la función a integrar, el estimador \hat{I} es una variable aleatoria, cuyo valor esperado coincide con I , y cuya varianza resulta σ^2/n , donde:

$$\sigma^2 = \int_{[0,1)^d} [f(x) - I]^2 \cdot dx \quad (\text{A.3})$$

En la práctica puede ser estimado mediante ec. A.4, y haciendo uso del teorema del límite central, puede ser aproximado un intervalo de confianza de I .

$$\hat{\sigma}^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (f(X_i) - \hat{I})^2 \quad (\text{A.4})$$

De este modo, el error en la estimación de I con ec. A.2 es del orden de $n^{1/2}$ en probabilidad. Métodos clásicos pueden alcanzar ordenes mejores para pequeña d y funciones integrando de variación suave. Sin embargo, el orden alcanzado con MC tiene la ventaja que se mantiene para cualquier dimensión d .

A.1.2. Los métodos de Quasi-Monte Carlo

La precisión alcanzada con los métodos basados en técnicas de MC se ve altamente afectadas por la aparición en las generaciones de puntos, de huecos o agrupaciones de alta densidad de punto dentro del supercubo dominio de integración. Para evitar estos problemas, los métodos de equidistribuciones, o los métodos de QMC (Niederreiter, 1992), se basan en la construcción determinística del conjunto de puntos donde evaluar la función integrando, de modo que se evite la existencia de huecos o agrupamientos en la distribución de los puntos de evaluación sobre el dominio de integración.

Para cuantificar la uniformidad en la distribución de un conjunto de puntos, existen diferentes definiciones de distancias entre la distribución uniforme continua en $[0, 1]^d$ y un conjunto discreto de puntos uniformemente distribuidos X_i cada uno con una probabilidad $1/n$, $i = 1, \dots, n$. La distancia más comúnmente usada en la discrepancia estrella:

$$D_n^* = D_n^*(X_1, \dots, X_n) = \sup_{0 \leq c_j < 1} \left| \prod_{j=1}^d c_j - \frac{1}{n} \cdot \sum_{i=1}^n \prod_{j=1}^d \delta(0 \leq X_i^j < c_j) \right| \quad (\text{A.5})$$

donde $\delta(0 \leq X_i^j < c_j)$ representa la delta de Kronecker, que resulta 1 cuando se cumple la condición, y 0 en otro caso. La discrepancia estrella compara las distribución uniforme continua y discreta en los hiperrectángulo con un vértice en el origen.

La discrepancia estrella está relacionada con la precisión de integración siguiendo la inecuación de Koksma-Hlawka (Niederreiter, 1992):

$$\left| \widehat{I} - I \right| \leq D_n^*(X_1, \dots, X_n) \cdot V_{HK}(f) \quad (\text{A.6})$$

donde $V_{HK}(f)$ es la variación total de f en el sentido de Hardy y Krause (Niederreiter, 1992).

Es posible la construcción de una secuencia infinita de puntos X_1, X_2, \dots a lo largo de la cual se tiene $D_n^* = \mathcal{O}\left(n^{-1} \cdot (\log n)^d\right)$. Ello prueba que puede ser alcanzado de forma asintótica precisiones mejores que las alcanzadas con MC , al menos para funciones integrando con $V_{HK}(f) < \infty$. La acotación del error aumenta con n hasta que se cumple $n \geq \exp(d)$, lo que implica que

para muy alta dimensión d la mejora sobre el método de MC resulta muy costosa, o inalcanzable de forma práctica. No obstante, estudios empíricos sugieren que QMC resulta más preciso que MC en problemas reales. Pero esta ventaja de los métodos de QMC sobre MC disminuye al aumentar d (Morokoff & Caflisch, 1994; Morokoff & Caflisch, 1995).

Entre los métodos que permiten la construcción determinista de los conjuntos de puntos equi-distribuidos que muestrean el dominio de integración se encuentran el Método de Muestreo Latin en el Hiper cubo (*Latin Hypercube Sampling, LHS*), el método de (t,m,s)-redes y (t,s)-secuencias, o los Métodos de Lattice. Seguidamente se presentan los dos primeros, exponiéndose el último en la siguiente sección.

Uno de los inconvenientes con los que cuenta el método de QMC frente al método de MC es que no existe en la práctica forma de encontrar una cota superior de $|\hat{I} - I|$ a partir de la evaluación de las funciones $f(X_1), \dots, f(X_n)$. La estimación de $V_{HK}(f)$ es extremadamente difícil y, en todo caso, la inecuación A.6 resulta bastante conservativa. Diversas variantes se han presentado, asociadas a las diferentes metodologías de generación de puntos, en las que se introduce un cierto grado de aleatoriedad a QMC , manteniendo las propiedades de equi-distribución del conjunto de puntos. Con ello, si la resultante estimación de I permanece no sesgada, repitiendo de forma independiente un número de veces la generación de puntos y la estimación de la integral, puede ser estimada también la varianza de la estimación. A dichas variantes se las denomina métodos de Quasi-Monte Carlo Aleatorio (*Randomized quasi-Monte Carlo, RQMC*).

Generación mediante el método Latin de Muestreo en el Hiper cubo Unidad

La regla del punto medio, en integrales de una dimensión, realiza una estimación de la integral mediante la evaluación de la función en los puntos:

$$A_i = \frac{i - 0,5}{n}, \quad i = 1, \dots, n \quad (\text{A.7})$$

donde se representa con A_i los puntos de un método de integración que son usados para construir otro método de integración. La regla del punto medio puede ser presentada como un método QMC para integrales unidimensionales. Si la regla del punto medio es variada en la forma de:

$$A_i = \frac{i - V_i}{n}, \quad i = 1, \dots, n \quad (\text{A.8})$$

donde V_i son variables aleatorias independientes uniformemente distribuida $U(0, 1]$, este método, llamado método de muestreo estratificado, aparece como un método tipo RQMC.

Una variante del método de Latin Hypercube Sampling puede ser obtenida con d reglas del punto medio, convirtiendo en aleatorio su orden:

$$X_i^j = A_{\pi_j(i)} = \frac{\pi_j(i) - 0,5}{n} \quad (\text{A.9})$$

donde π_j representa permutaciones aleatorias uniformes de los enteros $1, \dots, n$. Cada una de las $n!$ posibles permutaciones tienen igual probabilidad de ser usada.

La definición original del método de *LHS* (McKay *et al.*, 1979) utiliza d métodos de muestreo estratificado independientes con variaciones del orden en cada uno de ellos:

$$X_i^j = A_{\pi_j(i)} = \frac{\pi_j(i) - V_{\pi_j(i)}^j}{n} \quad (\text{A.10})$$

El conjunto de puntos proporcionado por *LHS* posee proyecciones sobre cada eje del espacio d -dimensional prácticamente uniformes. Así, si el integrando es próximo a la suma de funciones univariadas dependientes de las coordenadas de X , el método de *LHS* resulta muy adecuado. Si el integrando está alejado de una suma de funciones univariadas, el método de *LHS* no es esperable que proporcione mejores estimaciones que el método de *MC*, pero tampoco peores. De hecho, se cumple que para cualquier f con $\int f^2 dX < +\infty$, se tiene

$$\sigma_{LHS}^2(\hat{I}) \leq \frac{n}{n-1} \cdot \sigma_{MC}^2(\hat{I}) \quad (\text{A.11})$$

con lo que un conjunto de $n > 1$ puntos procedentes de *LHS* no es nunca peor que $n - 1$ puntos procedentes de *MC* (Owen, 1997).

Generación mediante (t,m,s)-redes y (t,s)-secuencias

Una red (*net*) se describe como una secuencia (*sequence*) de puntos en el hipercubo $[0, 1]^s$. En una secuencia, un entero $b \geq 2$ es usado para representar la base de los puntos en $[0, 1)$. De modo que $X_i^j = \sum_{h=1}^{\infty} x_{ijh} \cdot b^{-h}$, donde x_{ijh} son enteros con $0 \leq x_{ijh} < b$. Aquí se utiliza para designar la dimensión de la regla de integración a la variable s , siendo d la dimensión del dominio completo sobre el que se efectúa la integración. La distinción está motivada en la literatura por ser habitualmente utilizada esta técnica de integración como subordinada de otras técnicas que transforman una integral múltiple en un dominio de dimensión d en una serie de integrales múltiples con dominios de menor dimensión s .

Un intervalo elemental en $[0, 1]^s$ en base b se define como un conjunto de la forma:

$$B = \prod_{j=1}^s \left[\frac{t_j}{b^{k_j}}, \frac{t_j + 1}{b^{k_j}} \right) \quad (\text{A.12})$$

con enteros no negativos k_j y $t_j < b^{k_j}$. El intervalo elemental B resulta un hiperrectángulo de volumen b^{-m} , donde $m = \sum_{j=1}^s k_j$.

Idealmente cada intervalo elemental de volumen V debe contener de un conjunto de puntos equidistribuidos $n \cdot V$ puntos. Sean $t \geq 0$ y $m \geq 0$ enteros. Una secuencia finita de $X_1, \dots, X_n \in [0, 1]^s$ con $n = b^m$ es una (t,m,s)-red en base b si cada intervalo elemental de volumen b^{t-m} contiene exactamente b^t puntos de la secuencia. La secuencia infinita $X_1, X_2, \dots \in [0, 1]^s$ es una (t,s)-secuencia en base b si para todos los $m \geq 0$ y todos los $k \geq 0$, cada secuencia finita $X_{k \cdot b^m + 1}, \dots, X_{(k+1) \cdot b^m}$ es una (t,m,s)-red en base b . Para casos particulares se demuestra la existencia de (t,m,s)-redes y (t,s)-secuencias, así como existen algoritmos para su cálculo (Niederreiter, 1992).

Pequeños valores de t implican mejor propiedades de equi-distribución de tanto (t,m,s)-redes como (t,s)-secuencias. Una ventaja del uso de la red formada por los $n = b^m$ primeros puntos de una (t,s)-secuencias, es que posteriormente puede ser incrementado el número de puntos, con una secuencia de $n = \lambda \cdot b^m$ $1 \leq \lambda < b$ puntos, y producirse que todos los puntos usados en la estimación \hat{I}_{b^m} son también usados en $\hat{I}_{\lambda \cdot b^m}$. De este modo, a medida que aumenta λ , y con ello n , a cada intervalo elemental le es añadido el mismo número de puntos. Las integrales numéricas realizadas a partir de una (t,m,s)-red tienen un error del orden $\mathcal{O}(n^{-1} \cdot (\log n)^{s-1})$ (Niederreiter, 1992).

Existen métodos eficientes en los que se introduce un grado de aleatoriedad a la generación de las (t,m,s)-redes, manteniendo sus propiedades. Con ello es posible una estimación de la varianza de la estimación de la integral, cumpliéndose (Owen, 1995)

$$\sigma_{RNET}^2(\hat{I}) = \mathcal{O}(n^{-3} \cdot (\log n)^{s-1}) \quad (\text{A.13})$$

La principal limitación práctica para la aplicación de integraciones múltiples basadas en (t,m,s)-redes es que no son conocidos algoritmos de generación para cualquier dimensión d ni base, reduciéndose su versatilidad para una aplicación general donde la dimensión de la integral es variable.

A.2. Los métodos de Lattice en integrales múltiples de alta dimensión

Las reglas de integración basadas en lattices, son reglas designadas para aproximar la integral de una función f en el supercubo $[0, 1]^d$, en situaciones en las que la función a integrar tiene unas variaciones suaves y es periódica, con periodo unidad, en cada una de las componentes de cada dirección:

$$f(\mathbf{x}) = f(\mathbf{x} + \mathbf{z}) \text{ para } \forall \mathbf{z} \in \mathbb{Z}^d \text{ y } \forall \mathbf{x} \in \mathbb{R}^d \quad (\text{A.14})$$

En el problema unidimensional, la regla del rectángulo permite la estimación de la integral A.1 en la siguiente forma:

$$\hat{I}_R = \frac{1}{n} \cdot \sum_{j=0}^{n-1} f\left(\frac{j}{n}\right) \quad (\text{A.15})$$

La propiedad de periodicidad de la función permitiría hacer uso de forma equivalente de la regla del trapecio:

$$\hat{I}_T = \frac{1}{2 \cdot n} \cdot f(0) + \frac{1}{n} \cdot \sum_{j=1}^{n-1} f\left(\frac{j}{n}\right) + \frac{1}{2 \cdot n} \cdot f(1) \quad (\text{A.16})$$

La generalización de la regla del rectángulo en un problema de dimensión $d > 1$ podría expresarse de la siguiente forma:

$$\hat{I}_{R^d} = \frac{1}{n^d} \cdot \sum_{j_1=0}^{n-1} \cdots \sum_{j_d=0}^{n-1} f\left(\frac{j_1}{n}, \dots, \frac{j_d}{n}\right) \quad (\text{A.17})$$

Sin embargo, el aumento de tipo potencial en el número de puntos donde evaluar la función, hace de esta regla altamente ineficaz para grandes dimensiones.

Como alternativa aparece la regla de buenos puntos de lattice –*good lattice points*– (Korobov, 1959), aproximándose la integral mediante la siguiente regla:

$$\hat{I}_Q = \frac{1}{n} \cdot \sum_{j=0}^{n-1} f\left(\frac{j}{n} \cdot \mathbf{z}\right) \quad (\text{A.18})$$

donde n es el número de puntos a evaluar por la cuadratura, y \mathbf{z} es vector de números enteros seleccionado apropiadamente. Matemáticamente pueden resultar puntos que aparecen fuera del supercubo dominio de integración, sin embargo dada la propiedad de periodicidad de la función, se entiende

como que a cada uno de estos puntos se le ha sustraído un vector de enteros, de modo que el punto resultante quede dentro del dominio. De este modo aparece la definición de lattice (Sloan & Joe, 1994):

Una lattice en \mathbb{R}^d es un subconjunto discreto de \mathbb{R}^d el cual es formado mediante adición y resta.

Así considérese $\mathbf{z} = (1, 2)$ y $n = 5$, el conjunto de puntos en $[0, 1]^2$ generado por \mathbf{z} en una regla de $n = 5$ puntos serán $\{(0, 0), (\frac{1}{5}, \frac{2}{5}), (\frac{2}{5}, \frac{4}{5}), (\frac{3}{5}, \frac{1}{5}), (\frac{4}{5}, \frac{3}{5})\}$. Estos puntos forman un lattice en $[0, 1]^2$.

Como generalización de la regla de buenos puntos de lattice, expresada en forma canónica (Sloan & Joe, 1994), resulta ec. A.19, donde $N = n_1 \cdot \dots \cdot n_r$, el orden de la regla de integración, $n_1, \dots, n_r > 1$ son enteros tales que n_k es múltiplo de n_{k+1} , $1 \leq r \leq d$ es el rango de la regla, y $\{z_1, \dots, z_r\}$ son vectores enteros linealmente independientes.

$$\hat{I}_Q = \frac{1}{N} \cdot \sum_{j_1=0}^{n_1-1} \dots \sum_{j_r=0}^{n_r-1} f\left(\frac{j_1}{n_1} \cdot \mathbf{z}_1 + \dots + \frac{j_r}{n_r} \cdot \mathbf{z}_r\right) \quad (\text{A.19})$$

Cuando el orden de la regla es unidad, ec. A.19 deriva en la regla de *good lattice points*. El objetivo buscado por una regla como la dada en ec. A.19 es la obtención de un conjunto de puntos que permitan muestrear el dominio de integración de la forma más eficiente. En este sentido, la selección de los generadores de la regla $\{\mathbf{z}_1, \dots, \mathbf{z}_r\}$ debe ser cuidadosa.

Desde un punto de vista práctico, el uso de reglas de lattice con rango superior a la unidad no introduce significativas ventajas sobre el uso de reglas de rango unidad, o reglas de buenos puntos de lattice. Incluso podría pensarse que resulta menos ventajosa, por dificultar la selección de los generadores más adecuados para la regla. A la descripción de un criterio para la búsqueda de este generador óptimo para reglas de buenos puntos de lattice, está dedicada el siguiente punto.

A.2.1. Reglas de Lattice de Rango-1. El método de los buenos puntos de lattice.

El objetivo fundamental en la implementación de una regla de lattice de rango unidad, es la selección del vector generador de la regla. El generador óptimo depende de la función a integrar, pero su búsqueda para cada función particular requiere el conocimiento previo del resultado de la integral múltiple a resolver. Por ello es necesaria la definición de un criterio de medida general para medir la adecuación de la regla, al menos limitada a funciones pertenecientes a una determinada clase. Para definir una clase

de funciones, será introducido primero algunas propiedades de interés que resultan de aplicar la regla de lattice a funciones desarrollables en series de Fourier, en términos de cuyo desarrollo puede expresarse el error de la regla de lattice en la estimación de la integral.

Al igual que ocurre con la regla del rectángulo, es fácil expresar el error mantenido por una regla de lattice en términos de los coeficientes que definen la serie de Fourier que aproxima una función f , suponiendo que esta función tiene un desarrollo en serie de Fourier absolutamente convergente.

$$f(\mathbf{x}) = \sum_{\mathbf{h} \in \mathbb{Z}^d} \hat{f}(\mathbf{h}) \cdot \exp(2 \cdot \pi \cdot i \cdot \mathbf{h} \cdot \mathbf{x}) \quad (\text{A.20})$$

donde $\mathbf{h} \cdot \mathbf{x} = h_1 \cdot x_1 + \dots + h_d \cdot x_d$, y

$$\hat{f}(\mathbf{h}) = \int_{\mathbb{C}^d} \exp(-2 \cdot \pi \cdot i \cdot \mathbf{h} \cdot \mathbf{x}) \cdot f(\mathbf{x}) \cdot d\mathbf{x}, \quad \mathbf{h} \in \mathbb{Z}^d \quad (\text{A.21})$$

Aplicando ahora la regla de lattice sobre el desarrollo en serie de Fourier, queda:

$$\hat{I}_Q(f) = f(\mathbf{x}) = \sum_{\mathbf{h} \in \mathbb{Z}^d} \hat{f}(\mathbf{h}) \cdot \hat{I}_Q[\exp(2 \cdot \pi \cdot i \cdot \mathbf{h} \cdot \mathbf{x})] \quad (\text{A.22})$$

El siguiente paso es estudiar el comportamiento de la regla de lattice sobre el término exponencial. Ello resulta fácil utilizando el concepto de lattice dual:

Dada una lattice L , el lattice dual L^\perp es el conjunto $\{\mathbf{h} \in \mathbb{R}^d : \mathbf{h} \cdot \mathbf{x} \in \mathbb{Z}, \forall \mathbf{x} \in L\}$

Entonces se tiene que (Sloan & Joe, 1994):

$$\hat{I}_Q[\exp(2 \cdot \pi \cdot i \cdot \mathbf{h} \cdot \mathbf{x})] = \begin{cases} 1 & h \in L^\perp \\ 0 & h \notin L^\perp \end{cases} \quad (\text{A.23})$$

Con ello se concluye el siguiente teorema

Sea Q una regla de lattice de dimensión d , y sea L la lattice asociada. Supóngase que f es una función con desarrollo en serie de Fourier absolutamente convergente. Entonces se cumple que:

$$\hat{I}_Q(f) - I(f) = \sum_{\mathbf{h} \in L^\perp} \hat{f}(\mathbf{h}) \quad (\text{A.24})$$

Para definir ahora de forma precisa esa clase de funciones, sea $c > 0$ y $\alpha > 1$, es definido $E_\alpha(c)$ la clase de funciones cuyos coeficiente de Fourier satisfacen:

$$\left| \widehat{f}(h) \right| \leq \frac{c}{(\bar{h}_1 \cdot \dots \cdot \bar{h}_s)^\alpha} \quad (\text{A.25})$$

donde $\bar{h} = \max(1, |h|)$. La velocidad de decaimiento de los coeficientes de Fourier de una función está relacionada con el grado de suavidad en la variación de una función. Así $E_\alpha(c)$ representa una familia de funciones con un determinado grado de suavidad en su variación. Basado en esto, se demuestra que para una regla de rango-1 aplicada sobre una función $f \in E_\alpha(c)$, existe una cota de su error:

$$\left| \widehat{I}_{Q[\mathbf{z}, N]}(f) - I(f) \right| \leq c \sum_{\mathbf{h} \cdot \mathbf{x} \equiv 0 \pmod{N}} \frac{1}{(\bar{h}_1 \cdot \dots \cdot \bar{h}_s)^\alpha} \quad (\text{A.26})$$

Con ello, se define como criterio de medida de la adecuación de una regla a la función P_α con $\alpha > 1$

$$P_\alpha(Q) = \sum_{\mathbf{h} \cdot \mathbf{x} \equiv 0 \pmod{N}} \frac{1}{(\bar{h}_1 \cdot \dots \cdot \bar{h}_s)^\alpha} \quad (\text{A.27})$$

En la práctica, la función P_α es calculada mediante

$$P_\alpha(Q) = I_Q \left[\prod_{\mathbf{h} \neq \mathbf{0}} F_\alpha(x_k) - 1 \right] \quad (\text{A.28})$$

donde

$$F_\alpha(x) = 1 + \sum_{\mathbf{h} \neq \mathbf{0}} \frac{\exp(-2 \cdot \pi \cdot i \cdot \mathbf{h} \cdot \mathbf{x})}{|\mathbf{h}|^\alpha} \quad (\text{A.29})$$

Las expresiones para $F_{\alpha=2}(x)$, $F_{\alpha=4}(x)$, $F_{\alpha=6}(x)$ con $x \in [0, 1]$ resultan:

$$F_{\alpha=2}(x) = 1 + 2 \cdot \pi^2 \cdot \left(x^2 - x + \frac{1}{6} \right) \quad (\text{A.30})$$

$$F_{\alpha=4}(x) = 1 + \frac{\pi^4}{45} \cdot (1 - 30 \cdot x^2 \cdot (1 - x^2)) \quad (\text{A.31})$$

$$F_{\alpha=6}(x) = 1 + \frac{2 \cdot \pi^6}{945} \cdot (1 - 21 \cdot x^2 + 105 \cdot x^4 - 126 \cdot x^5 + 42 \cdot x^6) \quad (\text{A.32})$$

De este modo, dado $n > 1$ y $\alpha > 1$, la regla de buenos puntos de lattice se define como la regla Q dada por ec. A.18 con $\mathbf{z} \in \mathbb{Z}^d$ tal que minimiza P_α .

En la práctica la búsqueda del vector entero generador se limita a vectores de la forma (Korobov, 1959)

$$\mathbf{z}(\ell) = (1, \ell, \ell^2 \bmod n, \dots, \ell^{d-1} \bmod n), \quad 1 \leq \ell \leq \left\lfloor \frac{n}{2} \right\rfloor \quad (\text{A.33})$$

limitándose así la búsqueda a $\left\lfloor \frac{n}{2} \right\rfloor$ vectores generadores.

Estimación del error en el método de los buenos puntos de lattice

La medida P_α (ec. A.27) ofrece una cota superior del error de implementar la regla de buenos puntos de lattice sobre una familia de funciones de una clase. El siguiente teorema muestra la convergencia de las reglas de buenos puntos de lattice (Korobov, 1960):

Si $N > 2$ es un entero y $s \geq 2$, entonces existe un $\mathbf{z} \in \mathbb{Z}^d$ tal que

$$P_\alpha(\mathbf{z}, N) \leq \frac{(2 \cdot \log N)^{\alpha \cdot s}}{N^\alpha} + \mathcal{O}\left(\frac{(\log N)^{\alpha \cdot s - 1}}{N^\alpha}\right) \quad (\text{A.34})$$

Sin embargo la cota P_α resulta muy conservadora en la práctica. Para estimar el error que se está cometiendo al implementar una regla de buenos puntos de lattice es conveniente proporcionar una estructura a la regla, de modo que puedan ser extraídos resultados parciales con los que aproximar el error. Así, puede ser usada la secuencia de reglas de lattice en la forma Q_0, Q_1, \dots, Q_d , dados por

$$\widehat{I}_{Q_r}(f) = \frac{1}{2^r n} \sum_{k_1=0}^1 \cdots \sum_{k_r=0}^1 \sum_{j=0}^{n-1} f\left(\frac{j}{n} \cdot \mathbf{z} + \frac{(k_1, \dots, k_r, 0, \dots, 0)}{2}\right) \quad (\text{A.35})$$

El primero de estas reglas, Q_0 , coincide con la regla de buenos puntos de lattice, mientras que los siguientes tienen la propiedad de que cada regla contiene todos los puntos de las reglas anteriores, mientras que añade el mismo número de puntos a la cuadratura. Se trata de una secuencia embebida de reglas. La regla final Q_d es la que utiliza un mayor número de puntos, y es la que proporciona la mejor aproximación a $I(f)$. El error de la estimación puede aproximarse por:

$$\left| \widehat{I}_{Q_d}(f) - \widehat{I}_{Q_{d-1}}(f) \right| \quad (\text{A.36})$$

Otra forma de obtener una estimación del error que se está cometiendo al implementar una regla de buenos puntos de lattice es introduciendo cierto grado de aleatoriedad en la implementación y analizar periódicamente sus resultados. Existen diversas formas de realizar una implementación de este modo, a continuación se presenta la que ha sido diseñada como más conveniente para la resolución de las integrales múltiples en esta Tesis.

A.3. Variante e implementación del método de lattice para integrales múltiples de *f.d.p.*

Han sido presentados los métodos numéricos más popularmente utilizados para la resolución de integrales múltiples de alta dimensión. Cada uno de ellos presenta una variedad de circunstancias sobre el problema a resolver ante las cuales el método resulta el más adecuado desde el punto de vista de eficiencia computacional. En el caso de las integrales múltiples que resultan de la definición de la *f.p.e.* o la *f.a.e.c.* se trata de integrales múltiples con las siguientes propiedades principales:

- La función a integrar está compuesta por el producto de d *f.d.p.*
- Los límites de integración no son constantes, sino que están subordinados al resto de variables sobre las que operan las integrales previas
- La dimensión resultante de la integral múltiple no es constante, sino que puede variar indefinidamente, aunque un límite práctico puede situarse entorno a $\text{máx } d \approx 30 - 50$, dependiendo de la capacidad de cálculo

Bajo estos condicionantes, se ha encontrado como método más adecuado el método de lattice para integrales múltiples, por tener una alta versatilidad y similar eficacia que otros métodos. Dentro del método de lattice, y dada la necesidad de controlar el error con el que se están produciendo los resultados del método, y por tanto, las estimaciones de las funciones de probabilidad, se ha planteado una nueva variante del método que facilita el proceso de chequeo periódico de la precisión en la estimación, hasta conseguir una adecuada aproximación. Además se ha mejorado la eficiencia de la variante teniendo en cuenta en su implementación el tipo de funciones objeto de integración.

A.3.1. Variante del los métodos de Lattice

El error de una regla de Lattice puede ser aproximado confiando aleatoriedad a la regla (Cranley & Patterson, 1976). La posibilidad de generar múltiples estimaciones de la integral permite extraer la estadística de esas estimaciones. El método por el que se confiere aleatoriedad a una regla debe permitir por un lado realizar un gran número de estimaciones, de modo que a medida que éstas se van produciendo, el control de su varianza y la definición de una tolerancia, proporcionan un criterio de parada. Por otro lado,

el método debe de asegurar la propiedad de equi-distribución de cada conjunto de puntos sobre los que evalúa una función para la estimación de su integral. La variante del método de lattice propuesta para acomodar estos requisitos tiene la siguiente expresión:

$$\widehat{I}_{RQ}(f) = \frac{1}{N} \cdot \sum_{k=0}^m \sum_{j=0}^{n-1} f\left(\frac{j}{n} \cdot \mathbf{z} + \frac{1}{2} \cdot \mathbf{b}_{j,k}\right) \quad (\text{A.37})$$

donde $N = n \cdot m$ es el número de puntos de evaluación con los que se realiza la estimación, con $m \leq 2^d$, y $\mathbf{b}_{j,k}$ es un vector binario de dimensión d . Con esta forma pueden ser realizadas hasta 2^d aproximaciones de la integral, cada una de ellas con n puntos adicionales respecto a la anterior. Tomando $q > 1$ entero, y seleccionando m como un múltiplo de q , $m = (i + 1) \cdot q$, $i = 0, \dots, \left[\frac{2^d}{q} - 1\right]$, incrementando i puede irse aumentando la precisión de la aproximación de $I(f)$. Así mismo, puede ser estimada esta precisión mediante:

$$\sigma_{RQ} = \sqrt{\frac{\sum_{\ell=1}^q \left(\widehat{I}_{RQ_\ell}(f) - \widehat{I}_{RQ}(f)\right)^2}{q - 1}} \quad (\text{A.38})$$

donde $\widehat{I}_{RQ}(f)$ refiere a ec. A.37, y

$$\widehat{I}_{RQ_\ell}(f) = \frac{1}{n} \cdot \sum_{t=0}^i \sum_{j=0}^{n-1} f\left(\frac{j}{n} \cdot \mathbf{z} + \frac{1}{2} \cdot \mathbf{b}_{j,t \cdot q + \ell}\right) \quad (\text{A.39})$$

Con esta construcción, el intervalo de confianza de la aproximación puede ser estimado haciendo uso de la inecuación de Chebyshev (Lindgren, 1976):

$$\mathcal{P} \left[\left| \widehat{I}_{RQ}(f) - I(f) \right| < u \cdot \sigma_{RQ} \right] \geq 1 - \frac{1}{u^2}; u > 1 \quad (\text{A.40})$$

A.3.2. Implementación de la variante de los métodos de Lattice

En la implementación de la variante antes presentada de los métodos de Lattice a las funciones del tipo de *f.p.e.* o *f.a.e.c.*, debe notarse la forma peculiar que mantienen las integrales múltiples que definen dichas funciones. Cada integral que las compone está formada por la integral de una función de densidad, $\psi(x_i)$, que pondera a una función definida así mismo por integrales múltiples, $g(x_i)$. Además los límites de integración dependen de las variables que están siendo integradas en un orden superior.

$$\int_{a(x_1, \dots, x_{i-1})}^{b(x_1, \dots, x_{i-1})} \psi(x_i) \cdot g(x_i) \cdot dt \quad (\text{A.41})$$

La forma más eficiente para resolver esta integral, para una función genérica $g(x_i)$ es notando que se está haciendo una media ponderada de la función $g(x_i)$ de acuerdo a la función de densidad de probabilidad $\psi(x_i)$. Así una estimación de la integral vendría dada por:

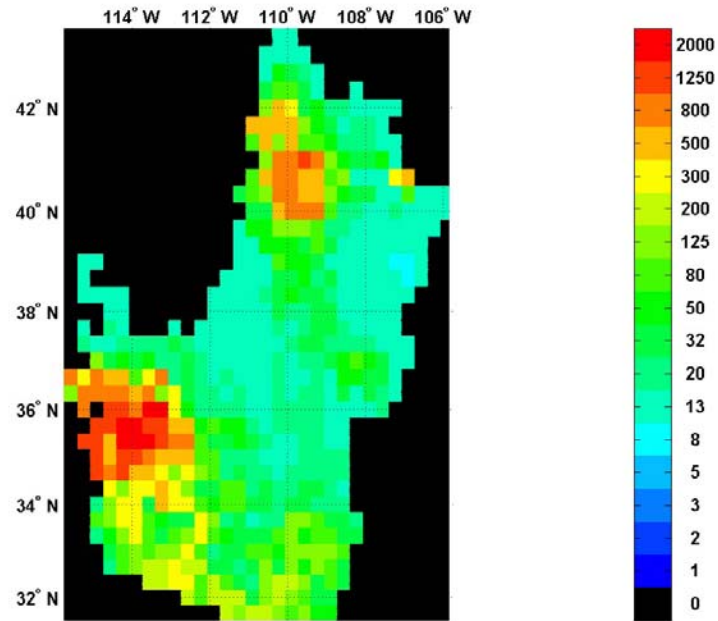
$$\int_{a(x_1, \dots, x_{i-1})}^{b(x_1, \dots, x_{i-1})} \psi(x_i) \cdot g(x_i) \cdot dt \approx [\Psi(b(x_1, \dots, x_{i-1})) - \Psi(a(x_1, \dots, x_{i-1}))] \cdot \frac{1}{n} \cdot \sum_{j=1}^n g(u_j) \quad (\text{A.42})$$

donde $\Psi(x)$ es la *f.p.a.* correspondiente a la *f.d.p.* $\psi(x)$, y u_j son números aleatorios $u_j \in U(0, 1)$, $j = 1, \dots, n$. En esta aplicación ha sido adoptado $n = 1$. Cuanto mayor sea el valor adoptado para n mayor es la precisión con la que realiza la estimación, sin embargo la eficacia computacional se reduce rápidamente al aumentar la dimensión, ya el número de puntos a evaluar crece según n^d .

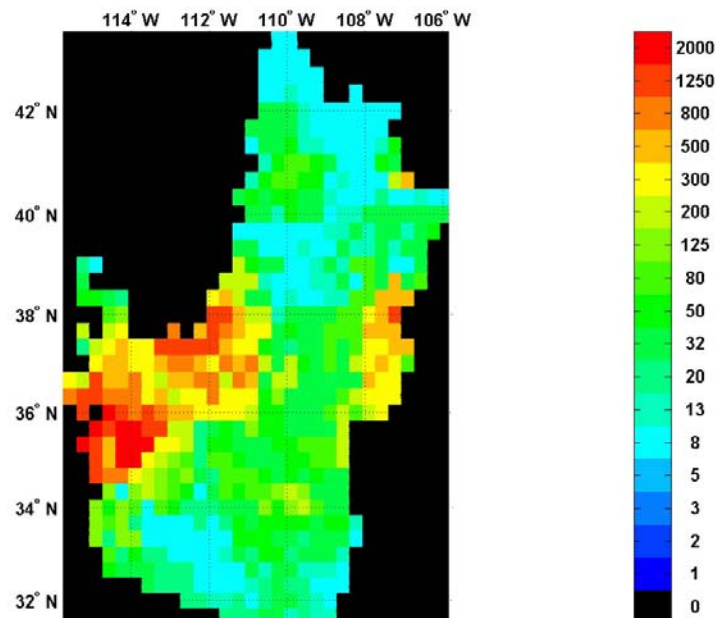
Por otro lado, puede notarse que los métodos de Lattice condicionan la función a integrar a ser periódica en cada una de sus variables, con periodo unidad. Ello podría pensarse como una limitación para su aplicación en el tipo de integral que aparecen con *f.p.e.* o *f.a.e.c.*, cuyos integrandos no siempre pueden considerarse como periódicos. Sin embargo, esta condición viene impuesta por el tratamiento que realiza el método de la evaluación de la función en los vértices. El método utiliza siempre la evaluación de la función en el vértice origen, con un peso $1/N$, mientras que no es incluido ningún otro vértice por resultar la función en esos puntos idéntica a la que resulta en el origen. Si la función no es periódica la regla estaría dando sistemáticamente mayor peso al origen que al resto de vértices. Al aumentar el número de puntos a evaluar, este efecto va siendo irrelevante. Así mismo, de acuerdo a la variante propuesta (ec. A.37), no en todas las aplicaciones es evaluada la función en el origen, teniendo una probabilidad de ser evaluada de 2^{-d} . Ello justifica que el incumplimiento en general de la condición de periodicidad de la función a integrar, no inhabilite el uso práctico del método.

Apéndice B

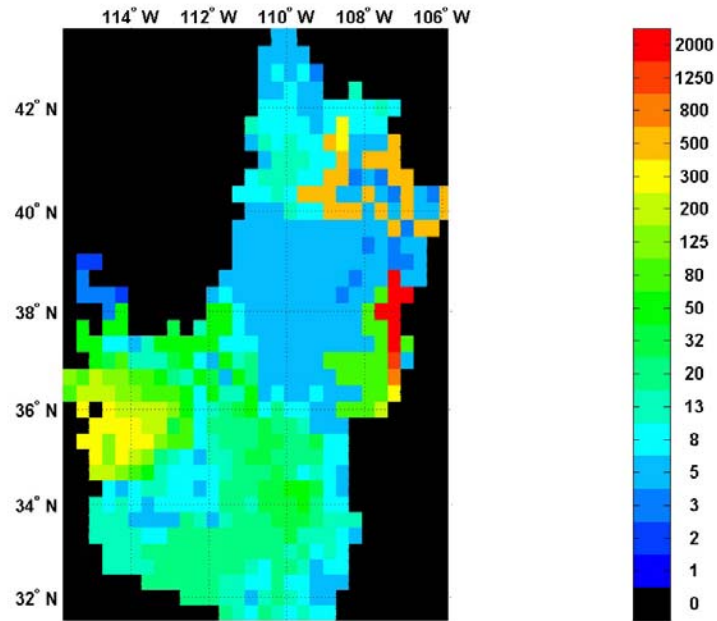
Mapas de *IFS* sobre la cuenca del Río Colorado (EEUU)



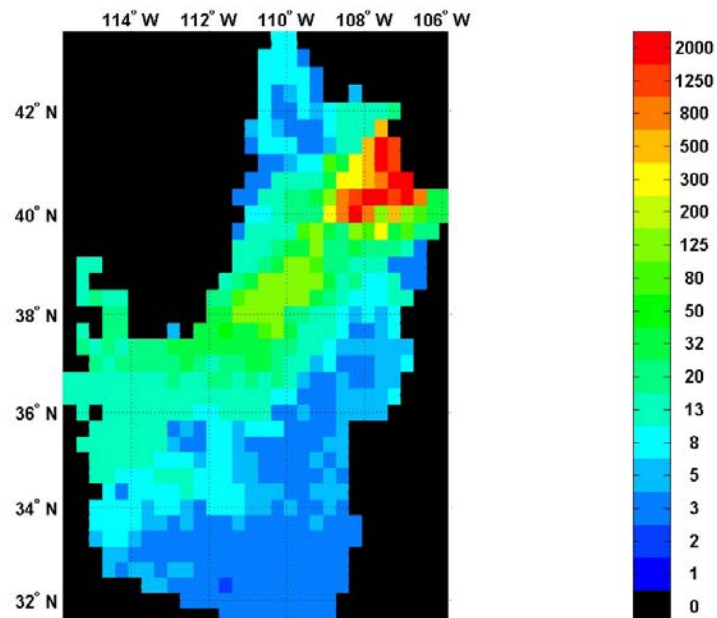
Mapa de *IFS* (años): precipitaciones anuales, año 1895 (Colorado, EEUU).



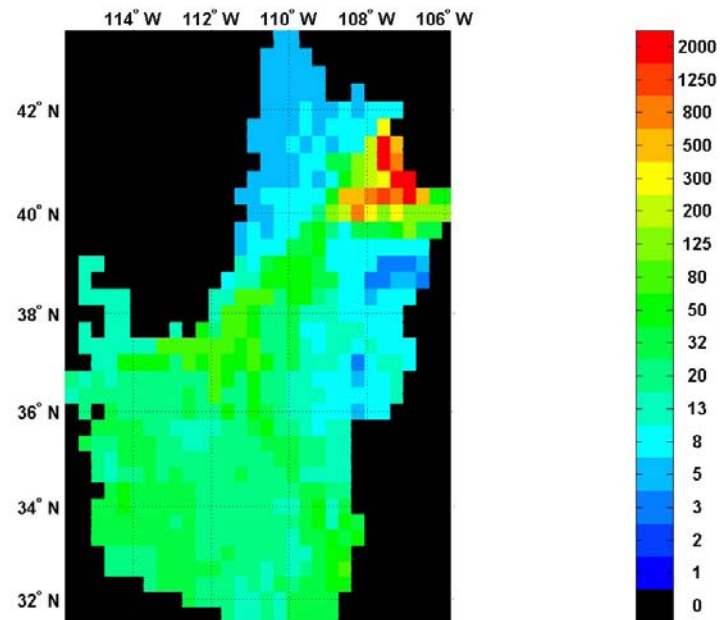
Mapa de *IFS* (años): precipitaciones anuales, año 1986 (Colorado, EEUU).



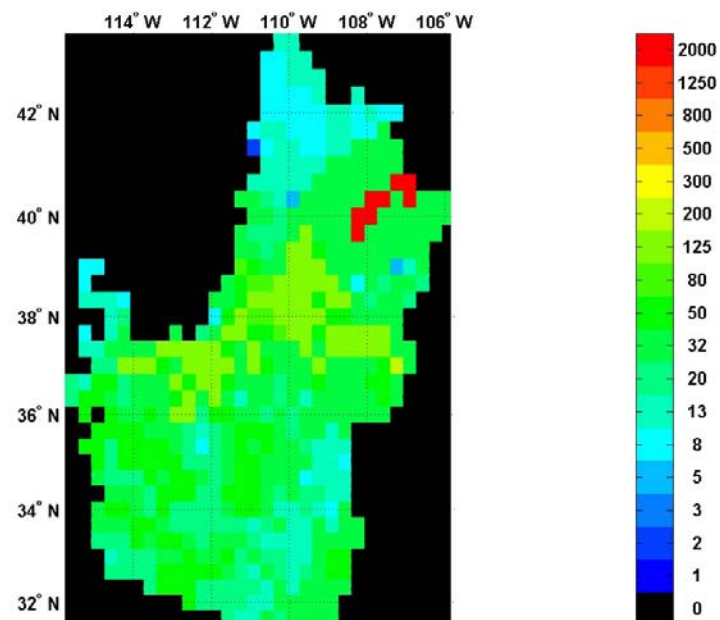
Mapa de *IFS (años)*: precipitaciones anuales, año 1897 (Colorado, EEUU).



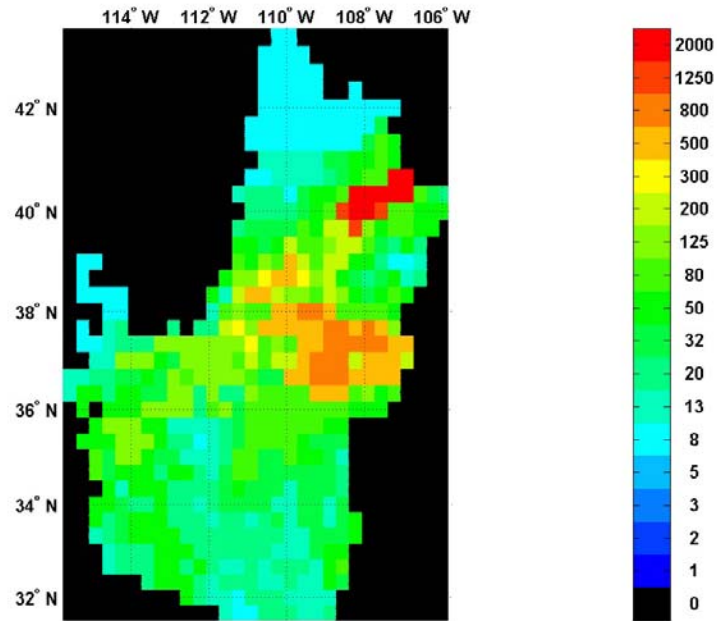
Mapa de *IFS (años)*: precipitaciones anuales, año 1898 (Colorado, EEUU).



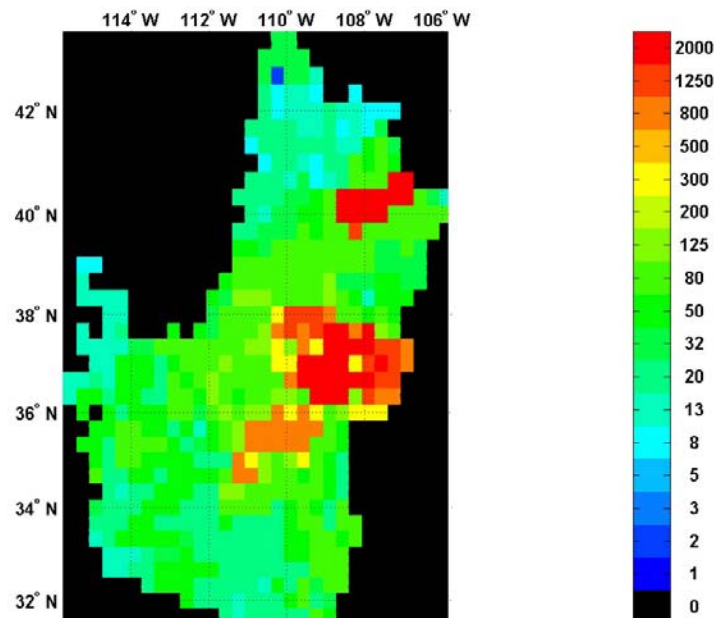
Mapa de *IFS (años)*: precipitaciones anuales, año 1899 (Colorado, EEUU).



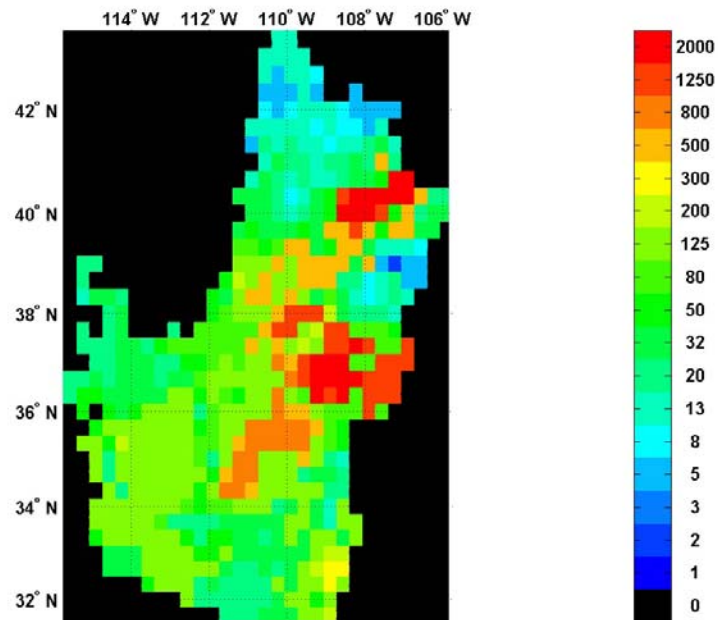
Mapa de *IFS (años)*: precipitaciones anuales, año 1900 (Colorado, EEUU).



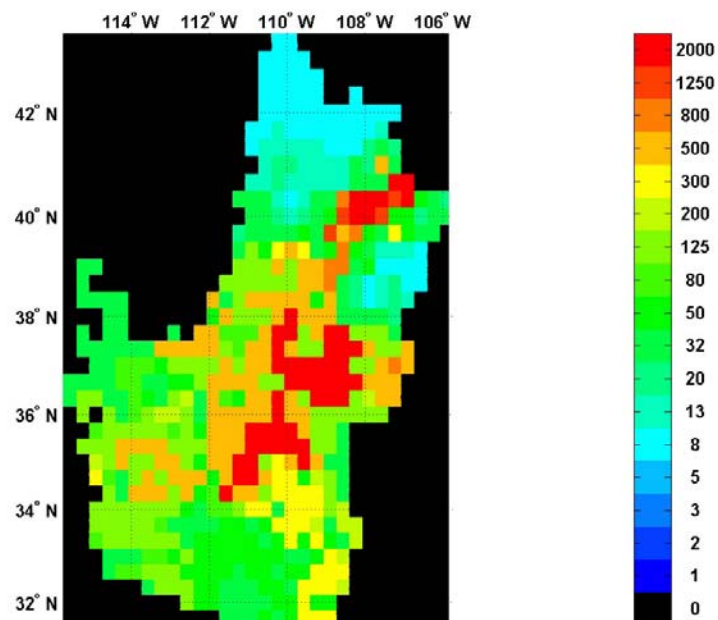
Mapa de *IFS* (años): precipitaciones anuales, año 1901 (Colorado, EEUU).



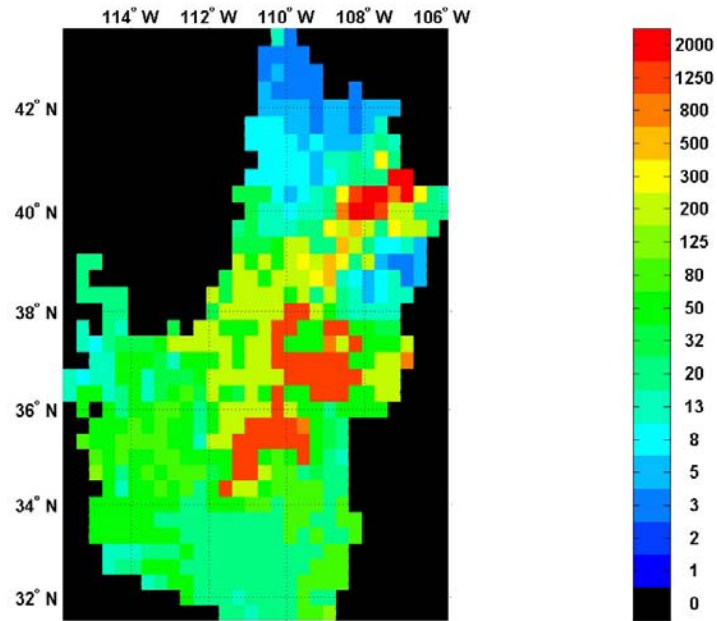
Mapa de *IFS* (años): precipitaciones anuales, año 1902 (Colorado, EEUU).



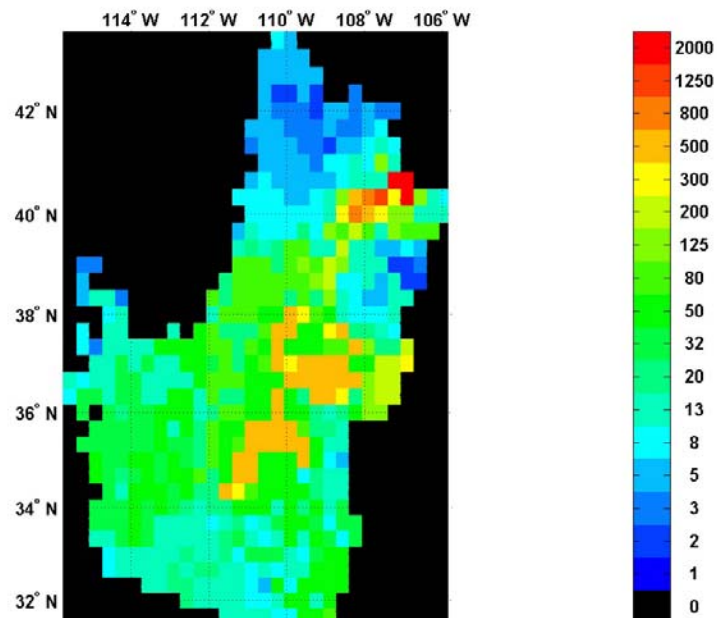
Mapa de *IFS (años)*: precipitaciones anuales, año 1903 (Colorado, EEUU).



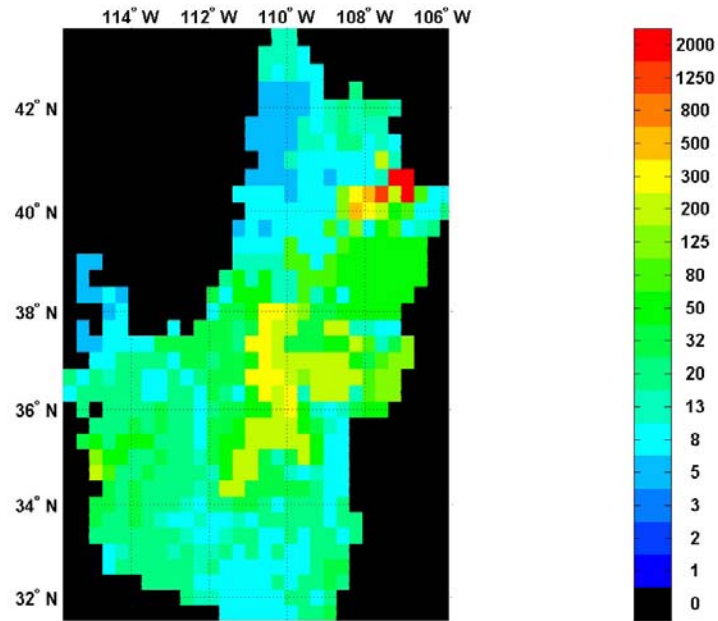
Mapa de *IFS (años)*: precipitaciones anuales, año 1904 (Colorado, EEUU).



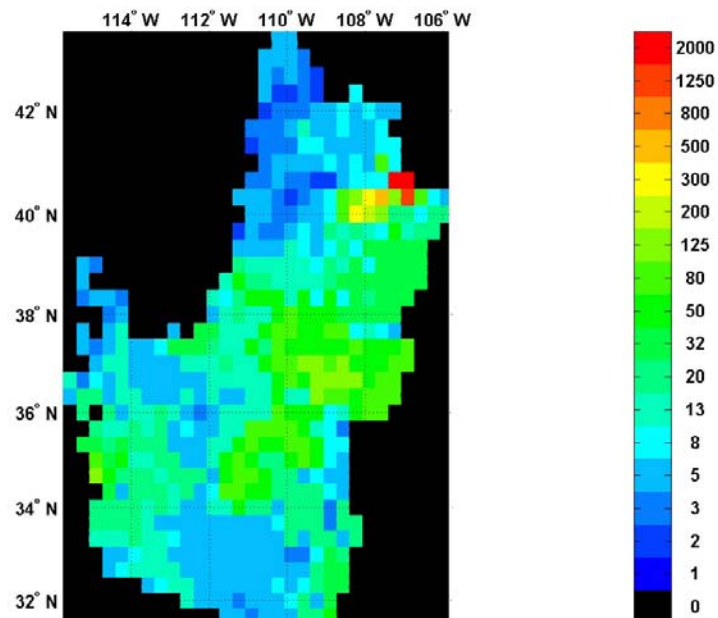
Mapa de *IFS (años)*: precipitaciones anuales, año 1905 (Colorado, EEUU).



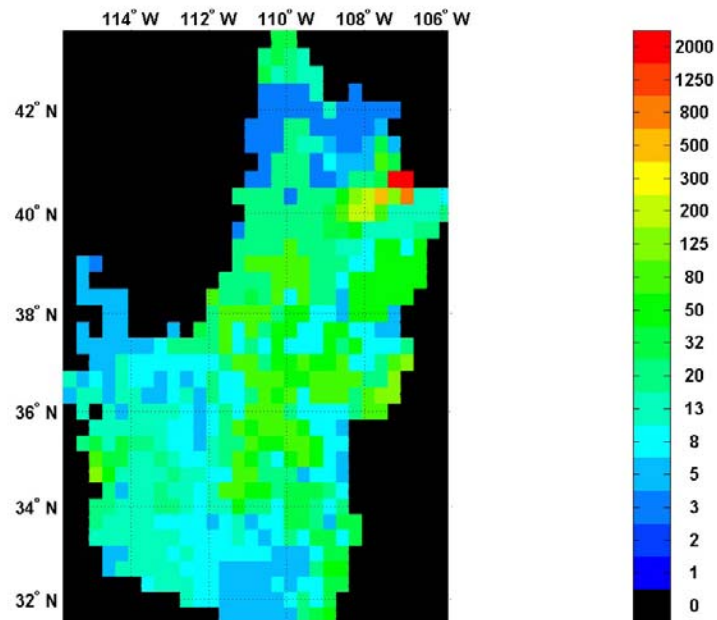
Mapa de *IFS (años)*: precipitaciones anuales, año 1906 (Colorado, EEUU).



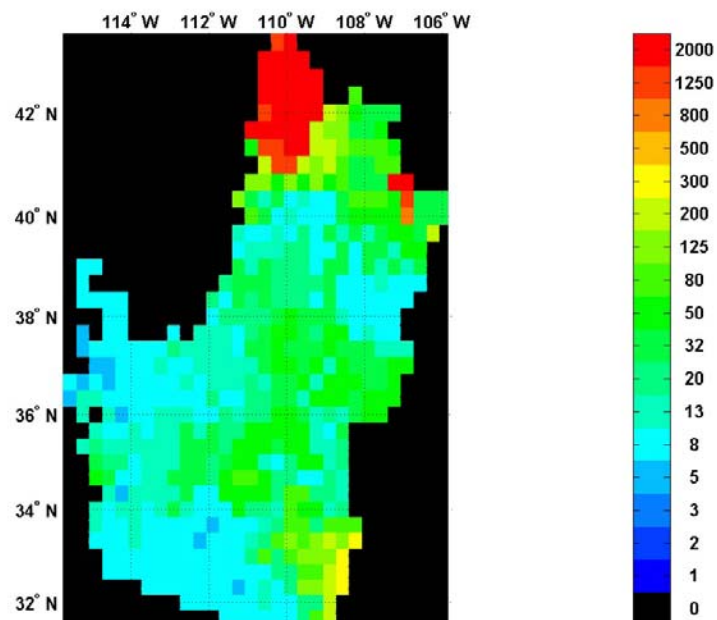
Mapa de *IFS (años)*: precipitaciones anuales, año 1907 (Colorado, EEUU).



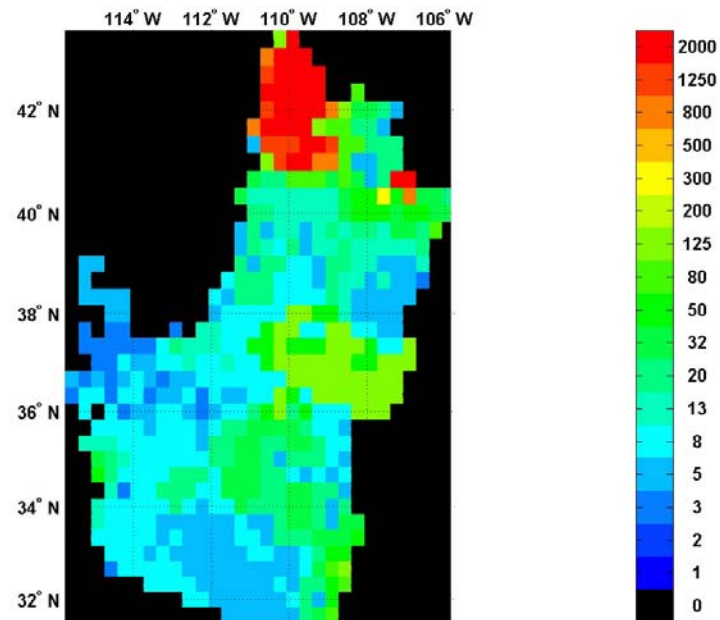
Mapa de *IFS (años)*: precipitaciones anuales, año 1908 (Colorado, EEUU).



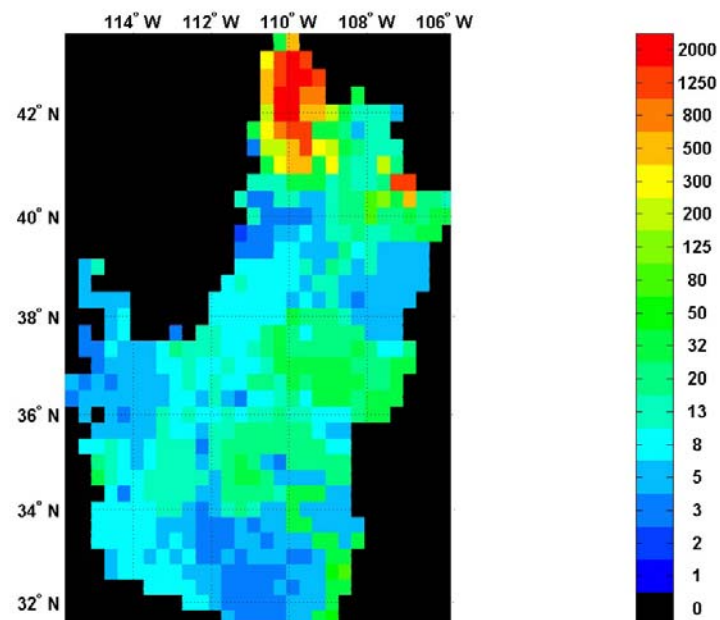
Mapa de *IFS (años)*: precipitaciones anuales, año 1909 (Colorado, EEUU).



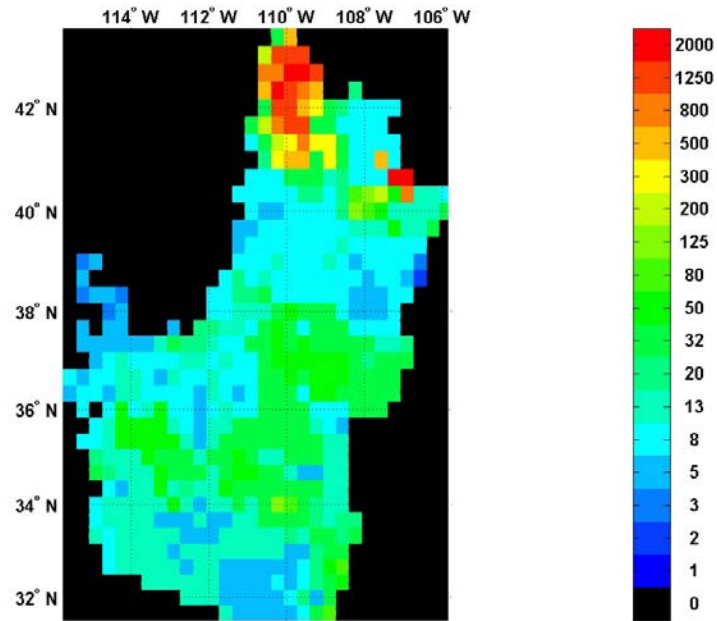
Mapa de *IFS (años)*: precipitaciones anuales, año 1910 (Colorado, EEUU).



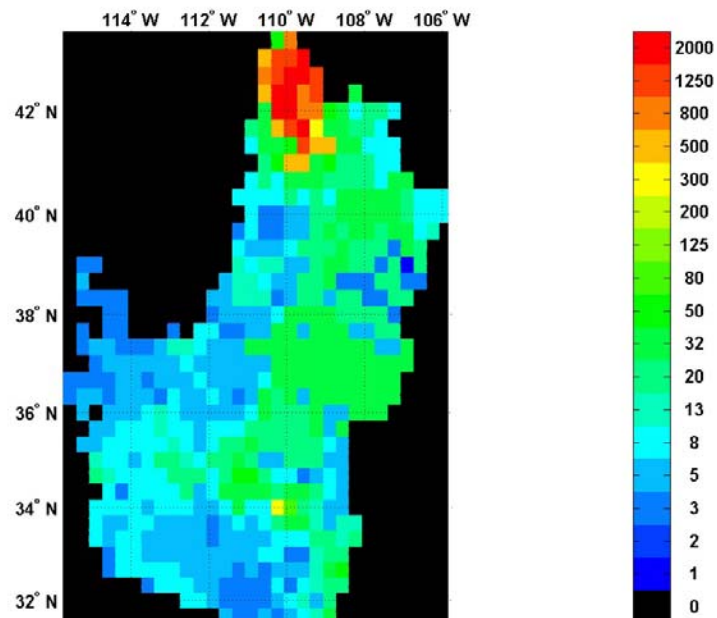
Mapa de *IFS (años)*: precipitaciones anuales, año 1911 (Colorado, EEUU).



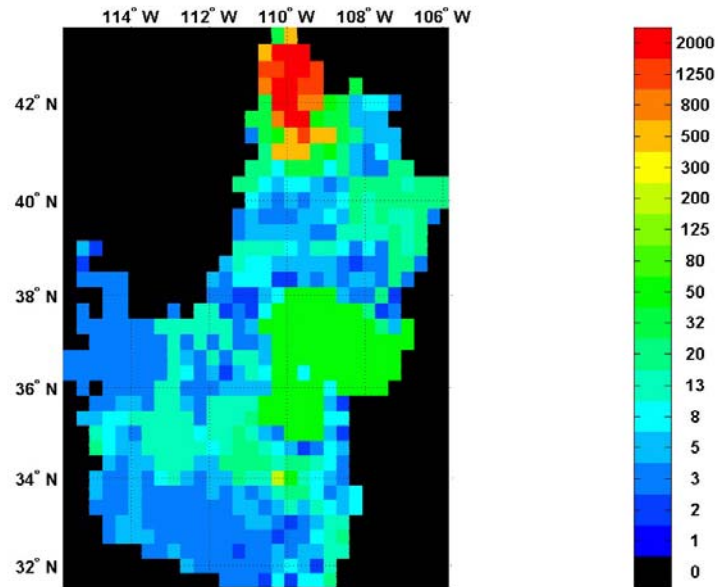
Mapa de *IFS (años)*: precipitaciones anuales, año 1912 (Colorado, EEUU).



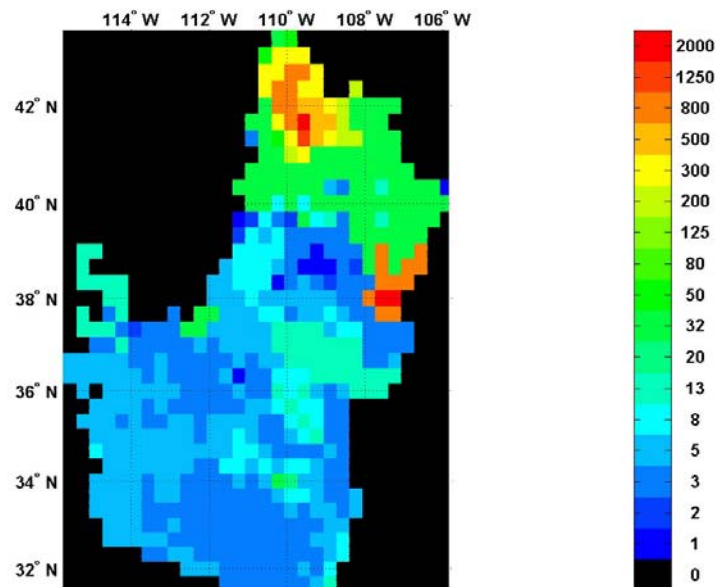
Mapa de *IFS (años)*: precipitaciones anuales, año 1913 (Colorado, EEUU).



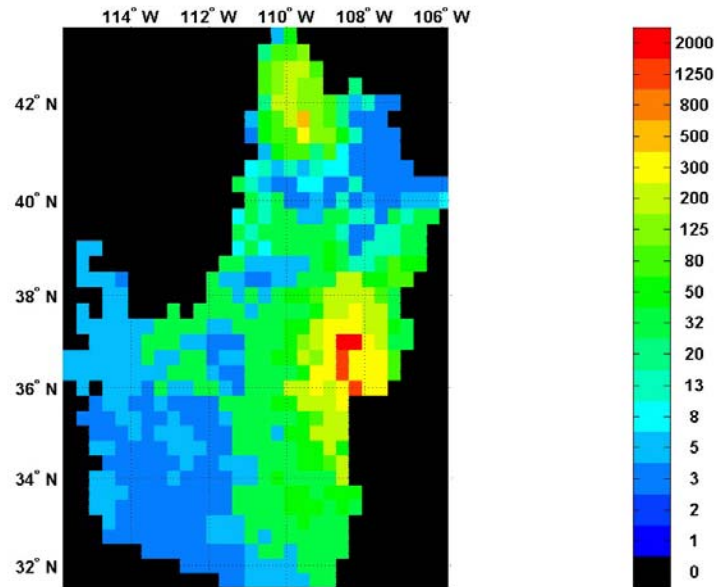
Mapa de *IFS (años)*: precipitaciones anuales, año 1914 (Colorado, EEUU).



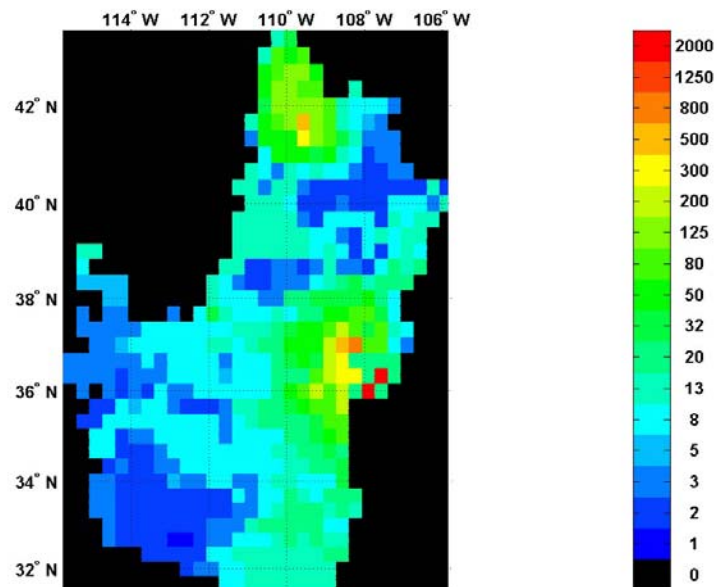
Mapa de *IFS (años)*: precipitaciones anuales, año 1915 (Colorado, EEUU).



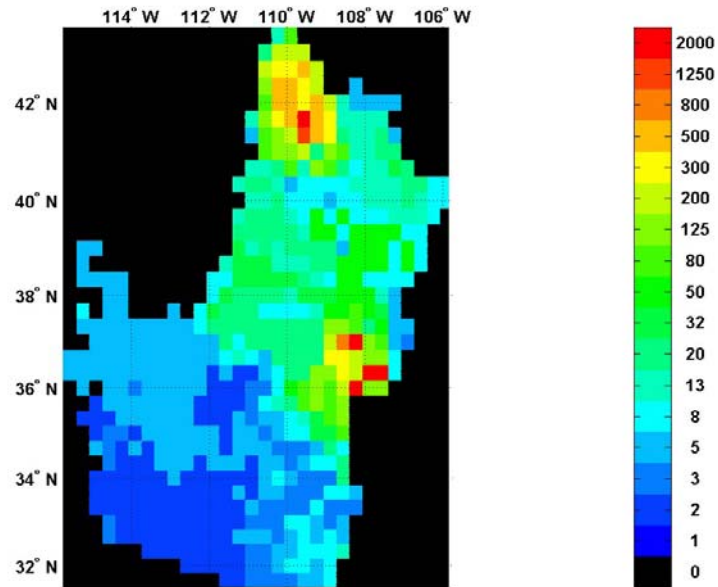
Mapa de *IFS (años)*: precipitaciones anuales, año 1916 (Colorado, EEUU).



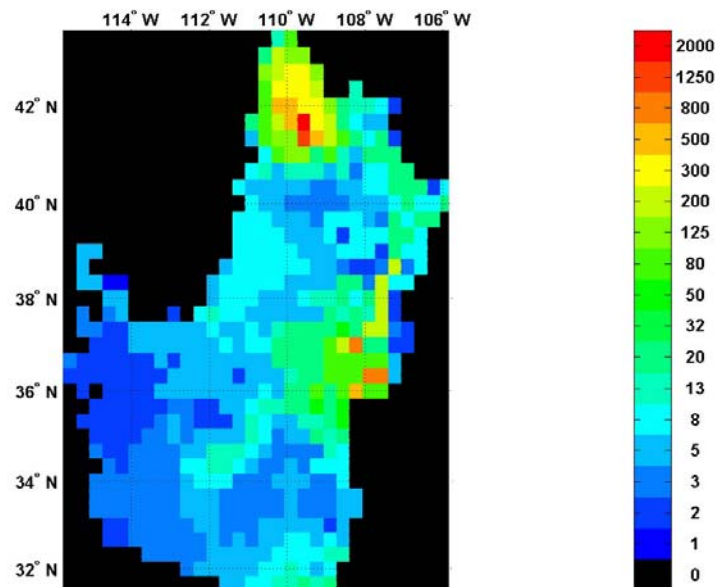
Mapa de *IFS (años)*: precipitaciones anuales, año 1917 (Colorado, EEUU).



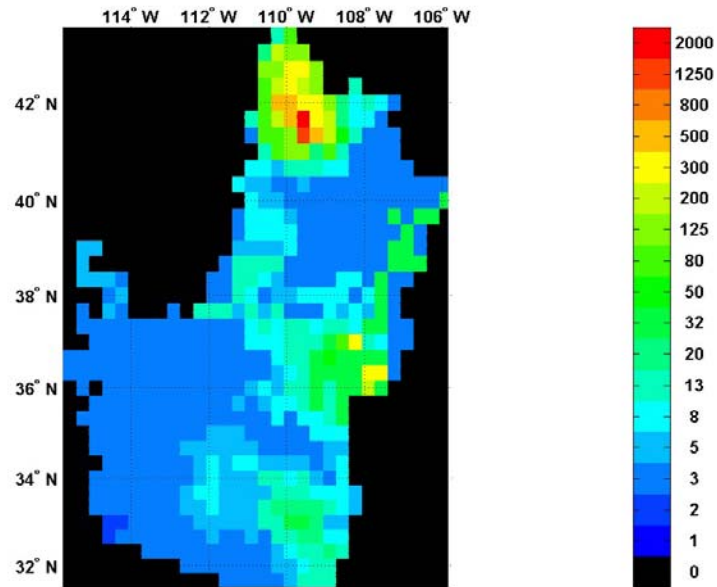
Mapa de *IFS (años)*: precipitaciones anuales, año 1918 (Colorado, EEUU).



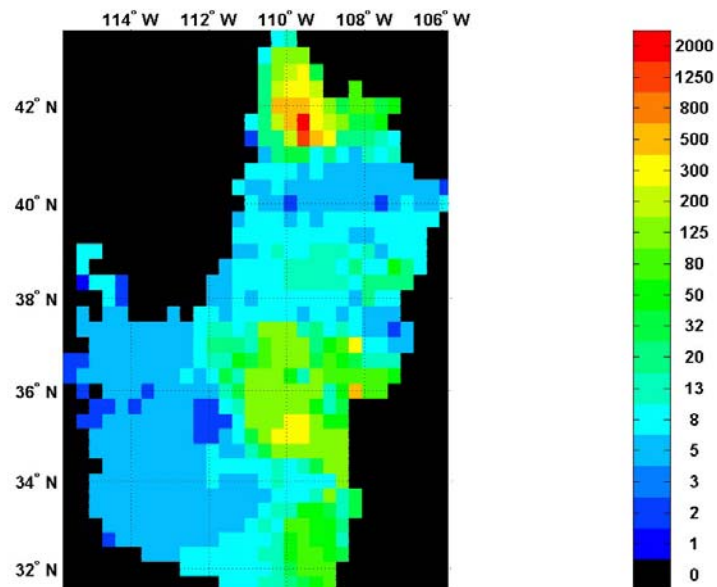
Mapa de *IFS (años)*: precipitaciones anuales, año 1919 (Colorado, EEUU).



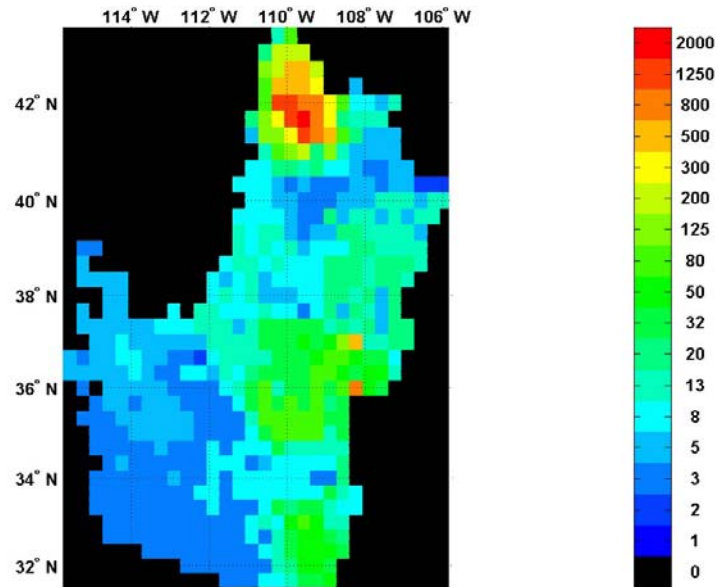
Mapa de *IFS (años)*: precipitaciones anuales, año 1920 (Colorado, EEUU).



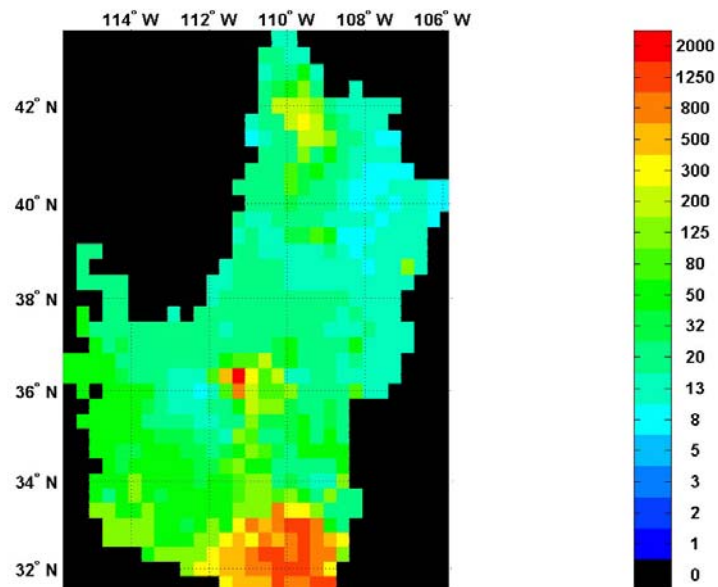
Mapa de *IFS (años)*: precipitaciones anuales, año 1921 (Colorado, EEUU).



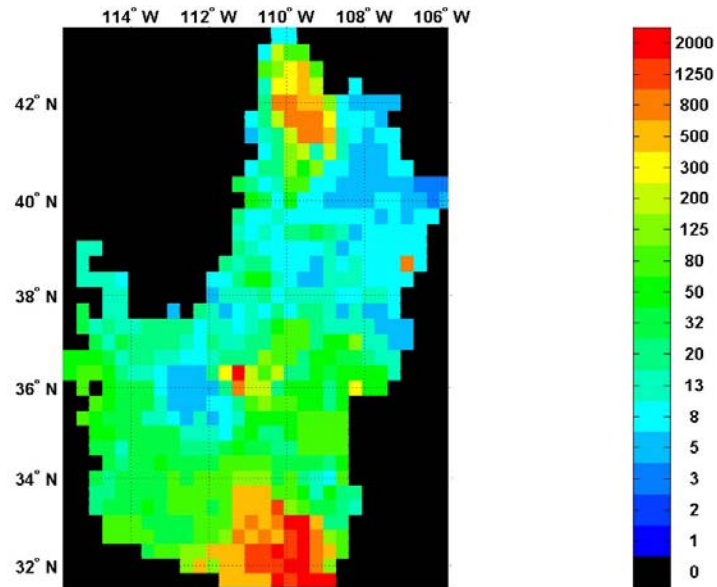
Mapa de *IFS (años)*: precipitaciones anuales, año 1922 (Colorado, EEUU).



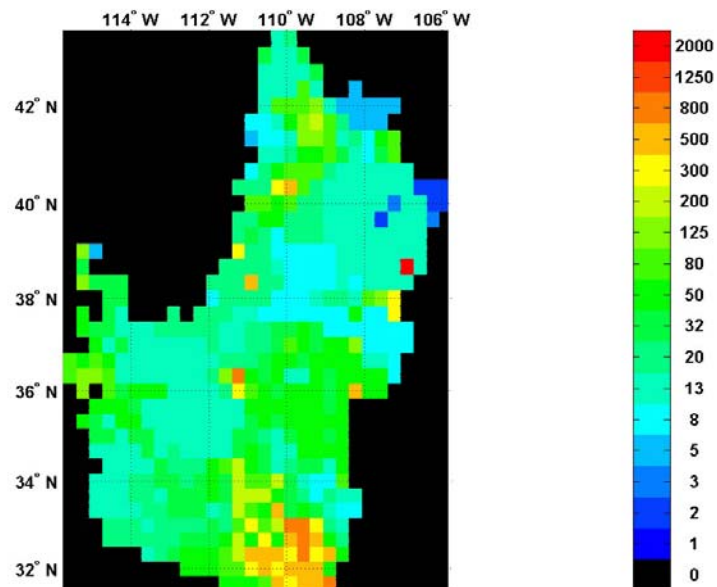
Mapa de *IFS (años)*: precipitaciones anuales, año 1923 (Colorado, EEUU).



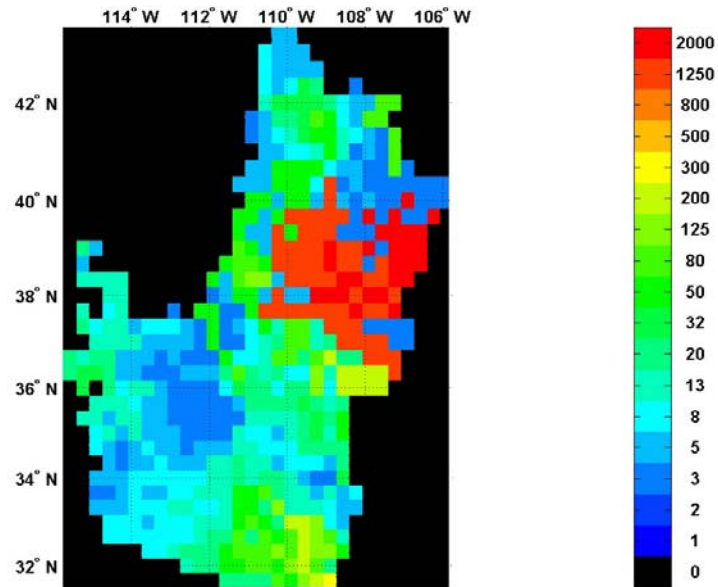
Mapa de *IFS (años)*: precipitaciones anuales, año 1924 (Colorado, EEUU).



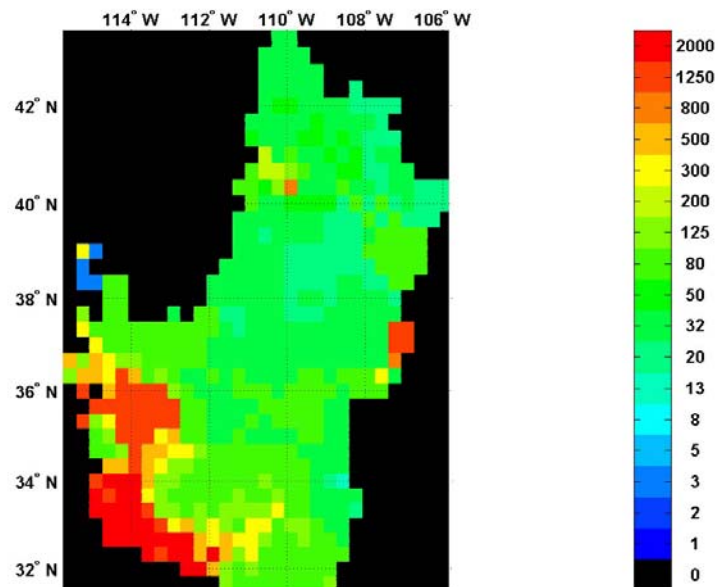
Mapa de *IFS* (años): precipitaciones anuales, año 1925 (Colorado, EEUU).



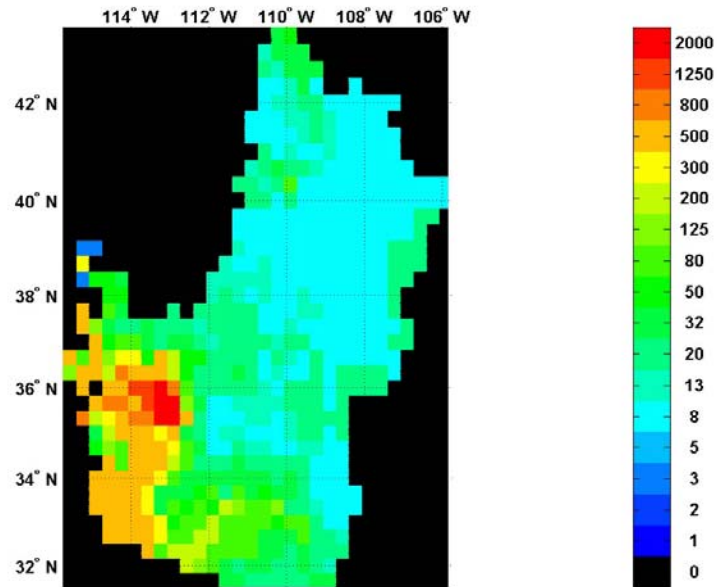
Mapa de *IFS* (años): precipitaciones anuales, año 1926 (Colorado, EEUU).



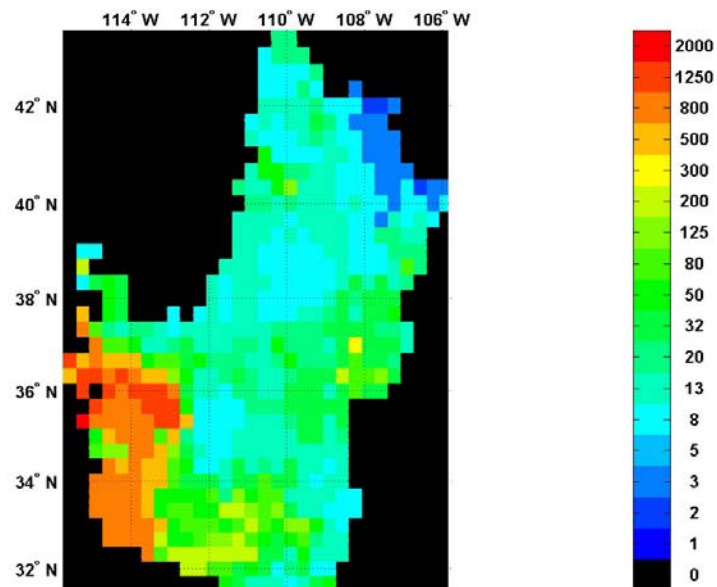
Mapa de *IFS (años)*: precipitaciones anuales, año 1927 (Colorado, EEUU).



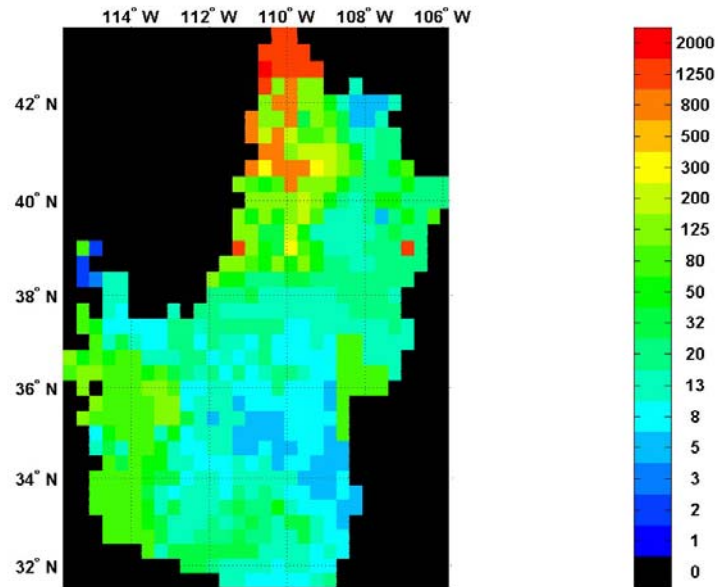
Mapa de *IFS (años)*: precipitaciones anuales, año 1928 (Colorado, EEUU).



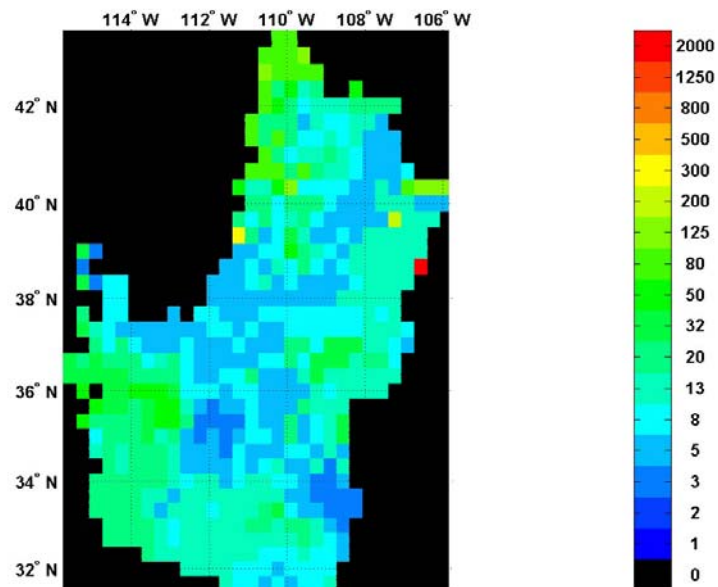
Mapa de *IFS* (*años*): precipitaciones anuales, año 1929 (Colorado, EEUU).



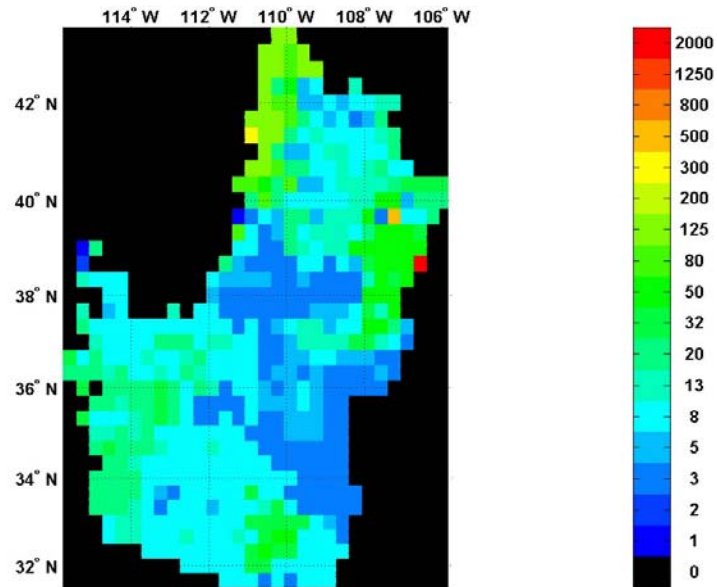
Mapa de *IFS* (*años*): precipitaciones anuales, año 1930 (Colorado, EEUU).



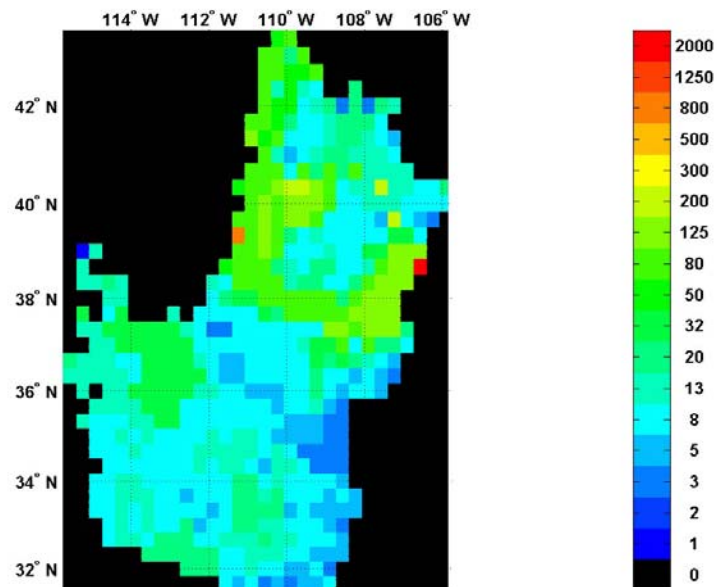
Mapa de *IFS (años)*: precipitaciones anuales, año 1931 (Colorado, EEUU).



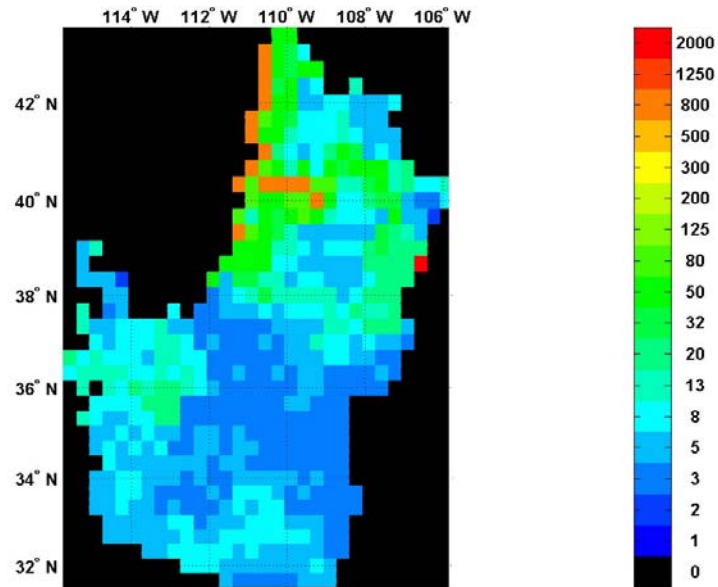
Mapa de *IFS (años)*: precipitaciones anuales, año 1932 (Colorado, EEUU).



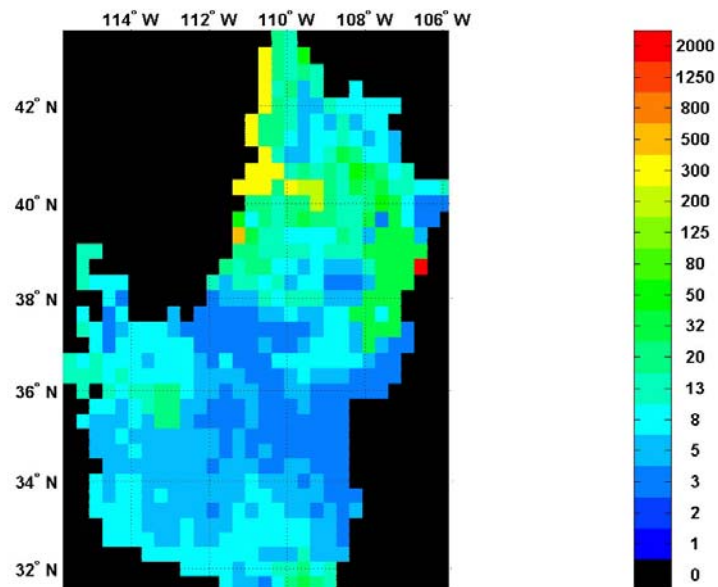
Mapa de *IFS* (años): precipitaciones anuales, año 1933 (Colorado, EEUU).



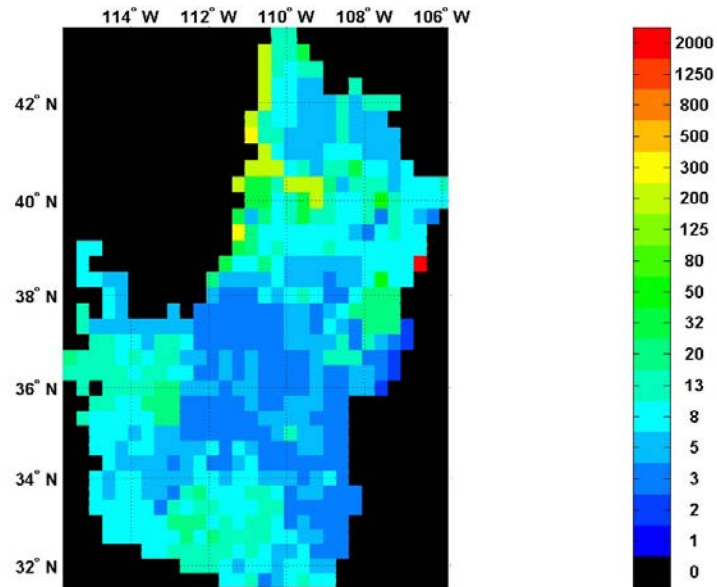
Mapa de *IFS* (años): precipitaciones anuales, año 1934 (Colorado, EEUU).



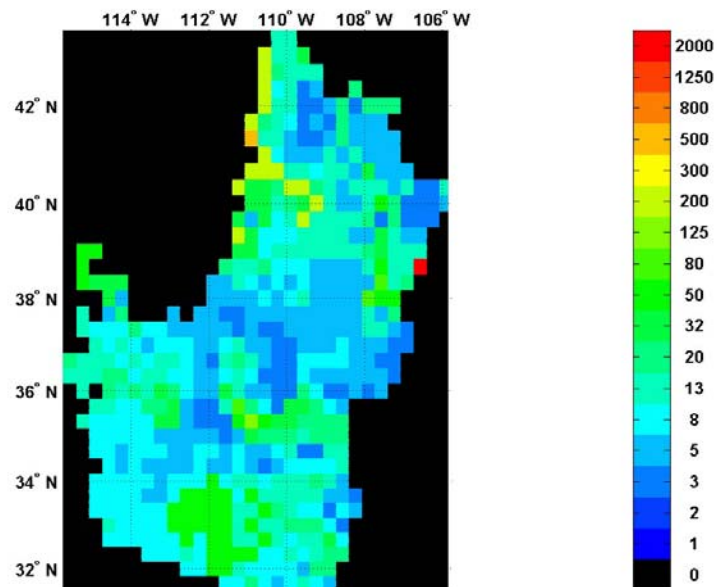
Mapa de *IFS (años)*: precipitaciones anuales, año 1935 (Colorado, EEUU).



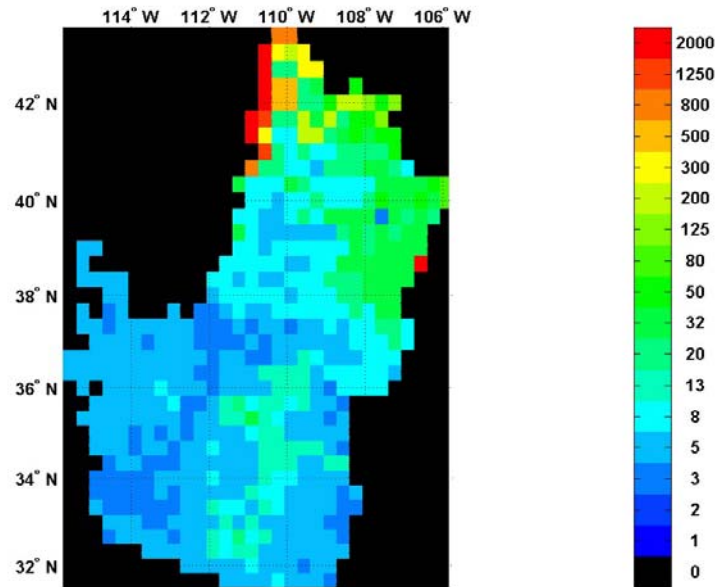
Mapa de *IFS (años)*: precipitaciones anuales, año 1936 (Colorado, EEUU).



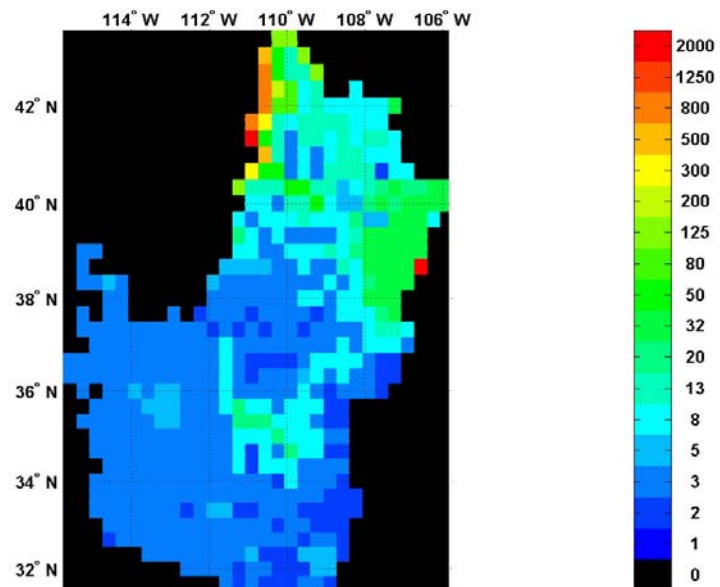
Mapa de *IFS* (años): precipitaciones anuales, año 1937 (Colorado, EEUU).



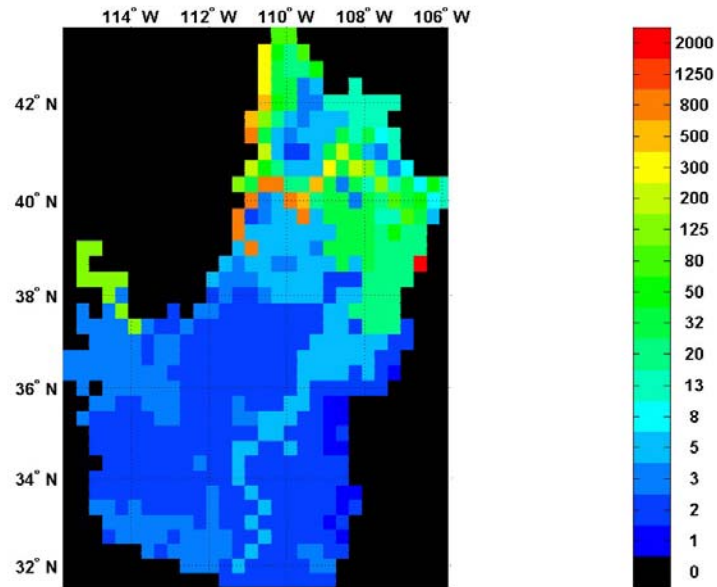
Mapa de *IFS* (años): precipitaciones anuales, año 1938 (Colorado, EEUU).



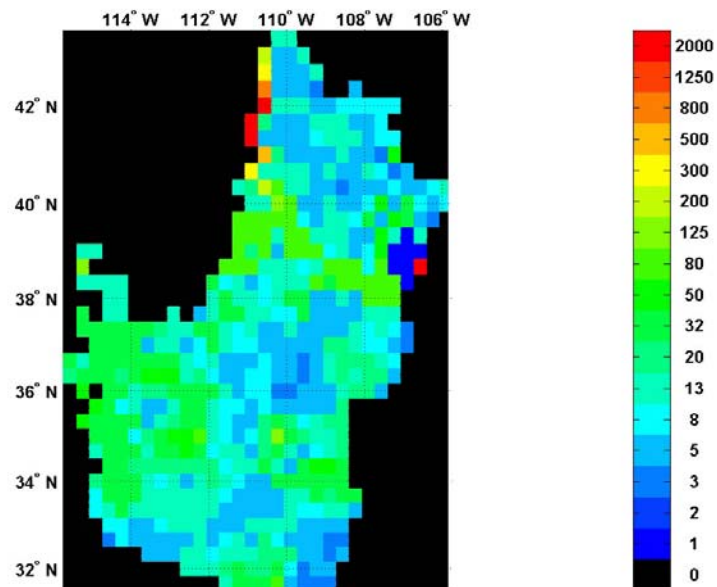
Mapa de *IFS (años)*: precipitaciones anuales, año 1939 (Colorado, EEUU).



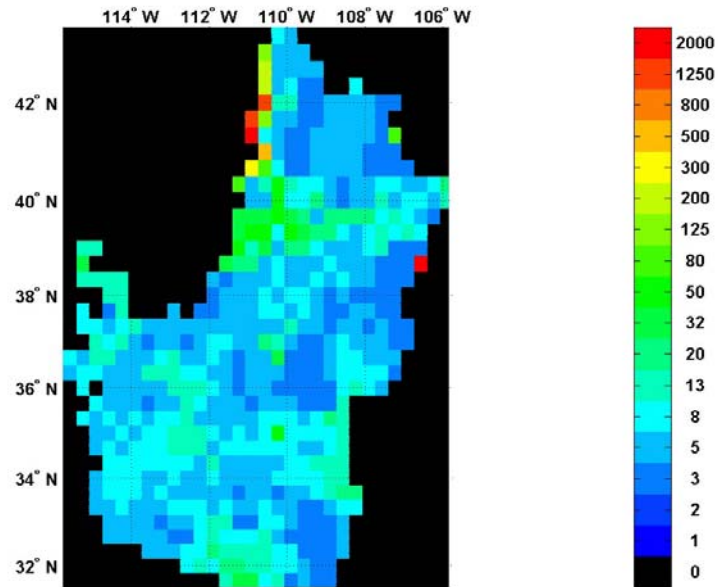
Mapa de *IFS (años)*: precipitaciones anuales, año 1940 (Colorado, EEUU).



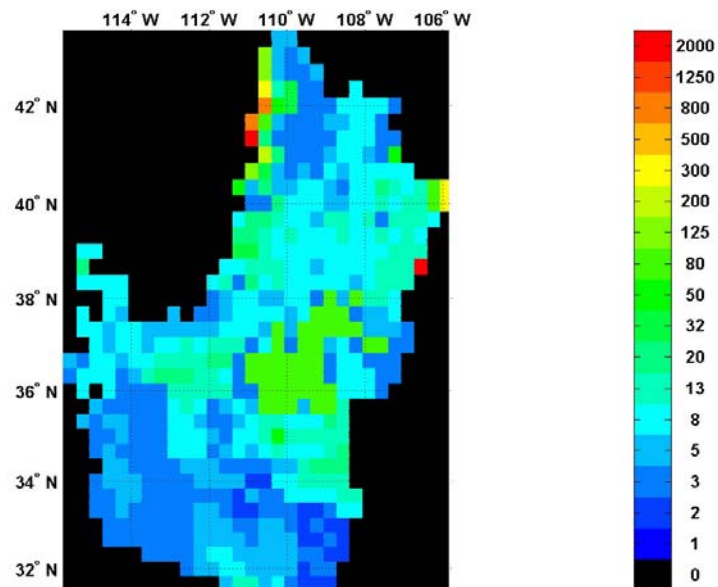
Mapa de *IFS (años)*: precipitaciones anuales, año 1941 (Colorado, EEUU).



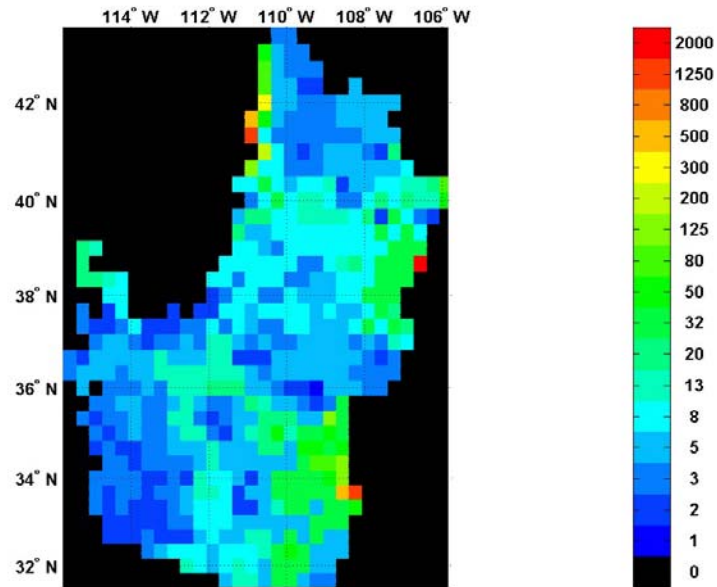
Mapa de *IFS (años)*: precipitaciones anuales, año 1942 (Colorado, EEUU).



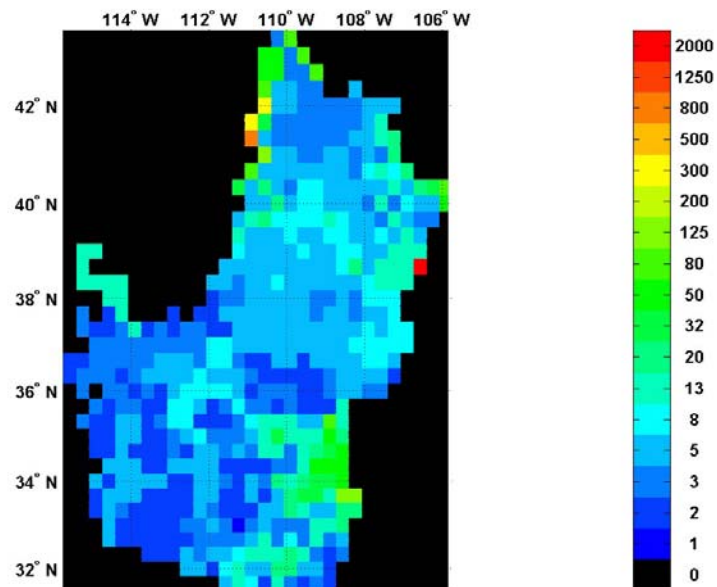
Mapa de *IFS (años)*: precipitaciones anuales, año 1943 (Colorado, EEUU).



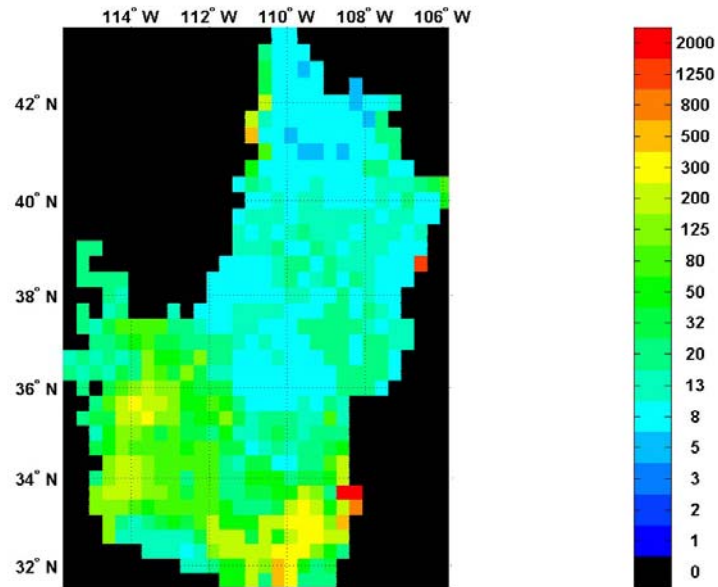
Mapa de *IFS (años)*: precipitaciones anuales, año 1944 (Colorado, EEUU).



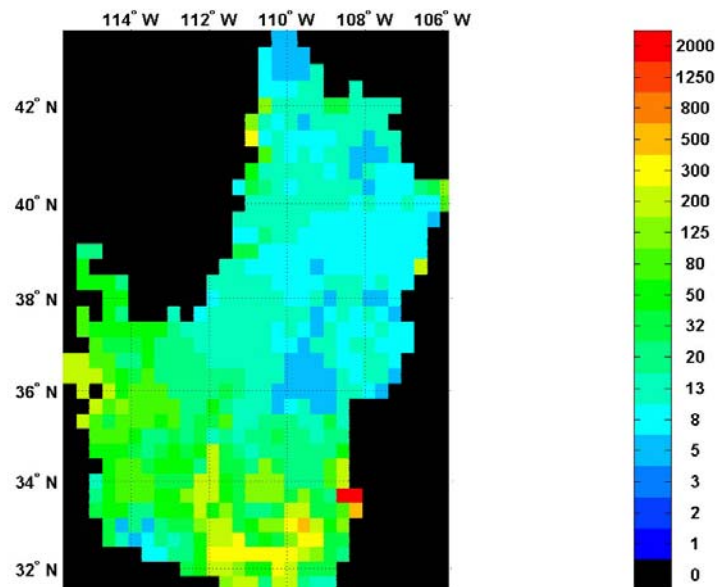
Mapa de *IFS* (años): precipitaciones anuales, año 1945 (Colorado, EEUU).



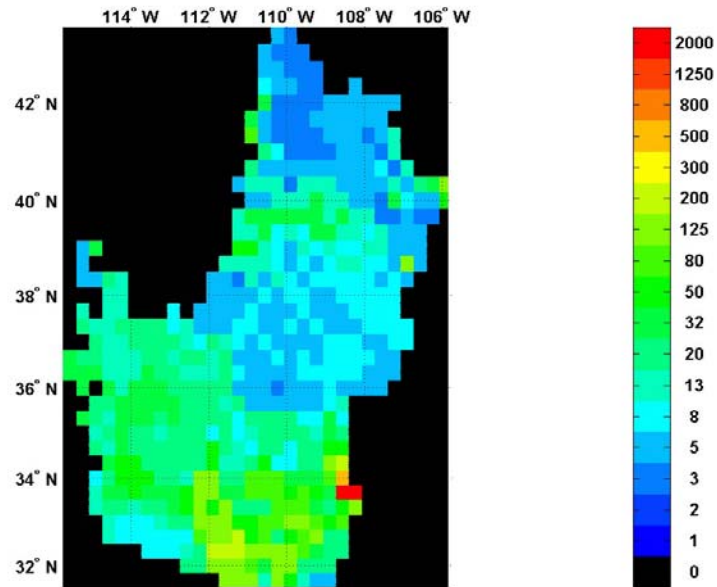
Mapa de *IFS* (años): precipitaciones anuales, año 1946 (Colorado, EEUU).



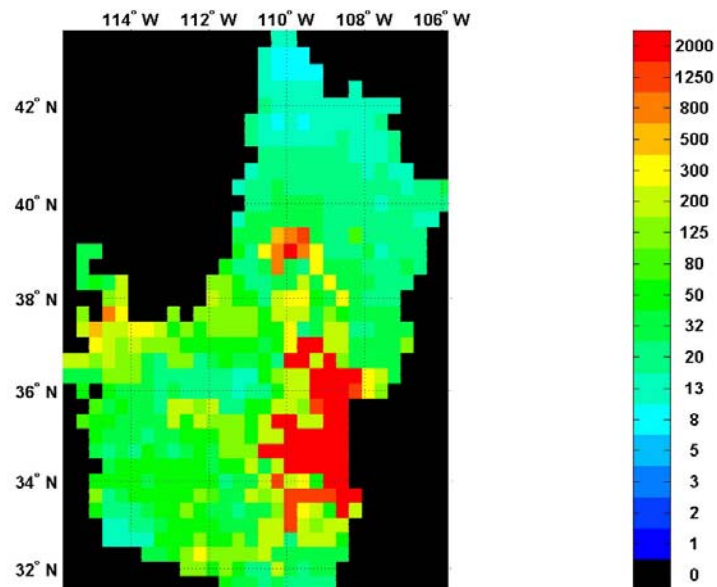
Mapa de *IFS (años)*: precipitaciones anuales, año 1947 (Colorado, EEUU).



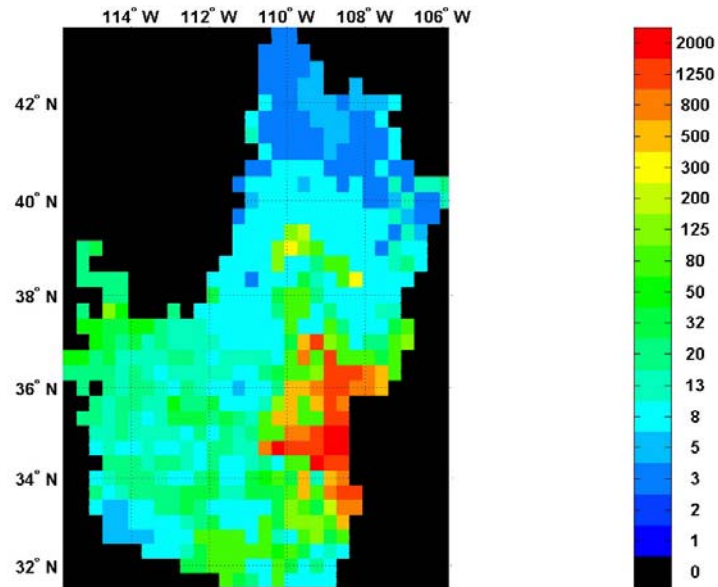
Mapa de *IFS (años)*: precipitaciones anuales, año 1948 (Colorado, EEUU).



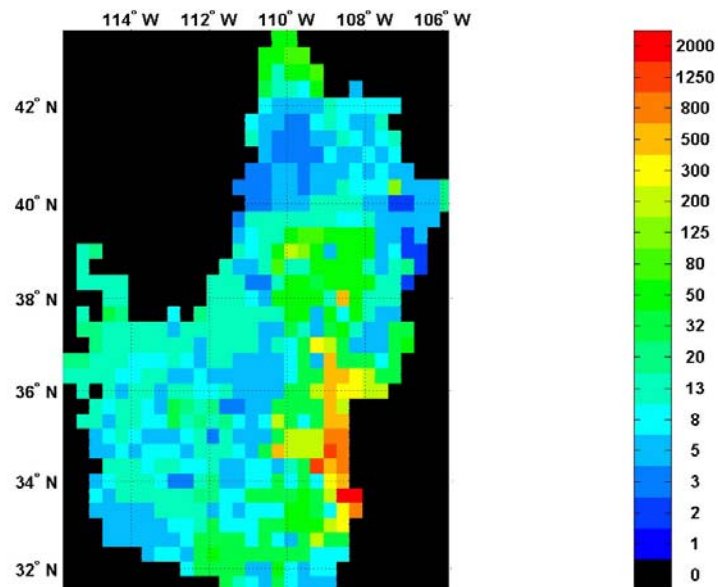
Mapa de *IFS (años)*: precipitaciones anuales, año 1949 (Colorado, EEUU).



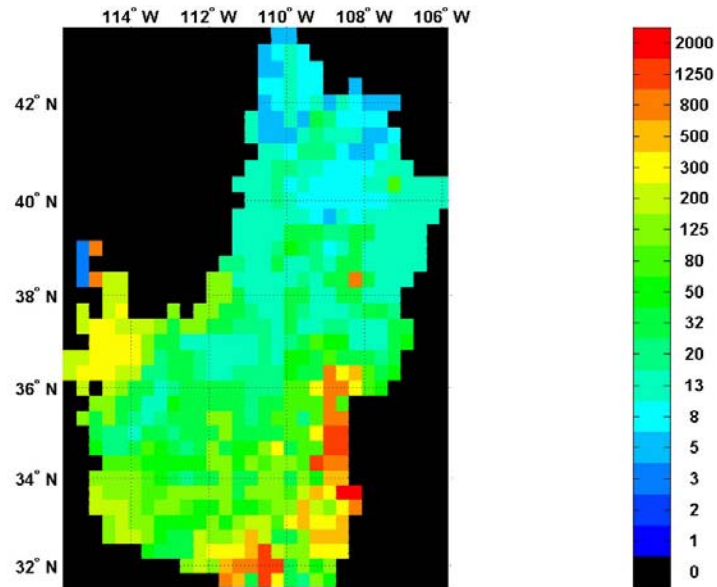
Mapa de *IFS (años)*: precipitaciones anuales, año 1950 (Colorado, EEUU).



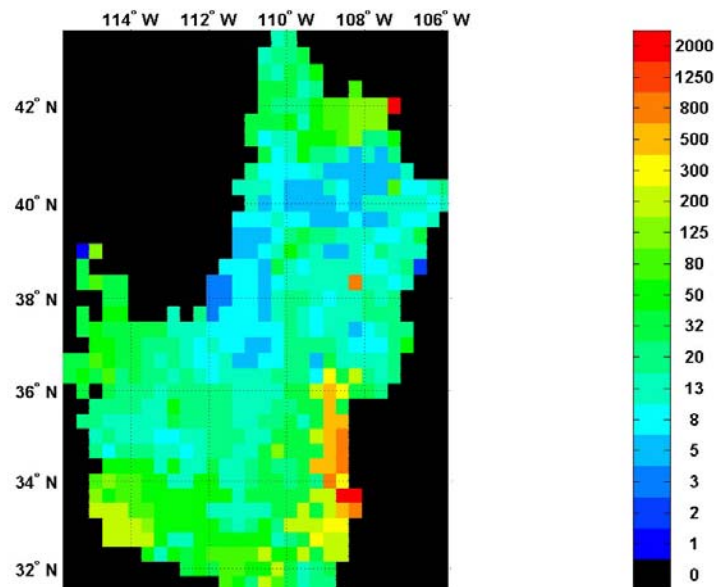
Mapa de *IFS (años)*: precipitaciones anuales, año 1951 (Colorado, EEUU).



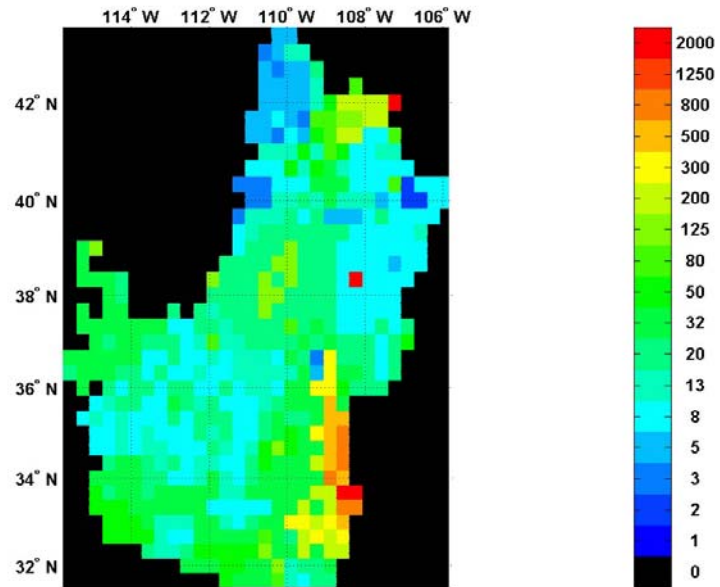
Mapa de *IFS (años)*: precipitaciones anuales, año 1952 (Colorado, EEUU).



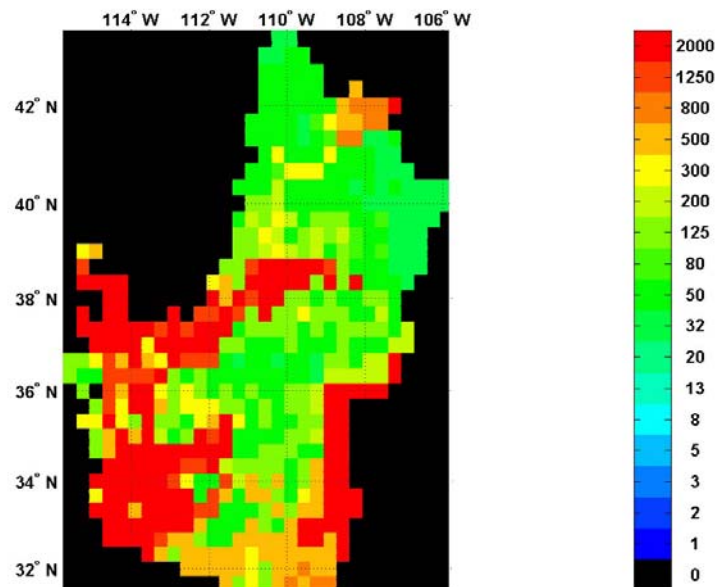
Mapa de *IFS (años)*: precipitaciones anuales, año 1953 (Colorado, EEUU).



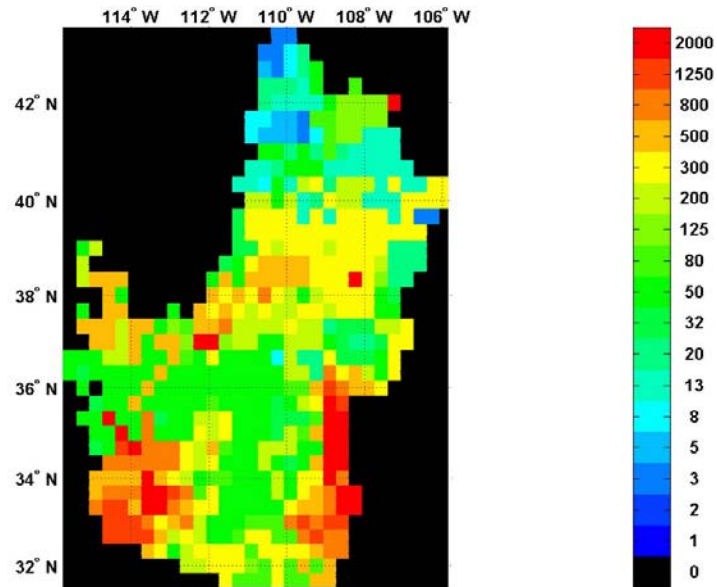
Mapa de *IFS (años)*: precipitaciones anuales, año 1954 (Colorado, EEUU).



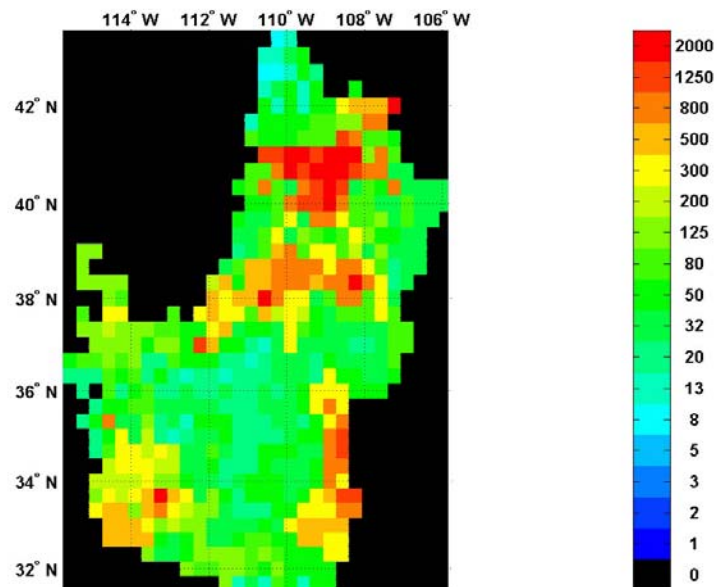
Mapa de *IFS (años)*: precipitaciones anuales, año 1955 (Colorado, EEUU).



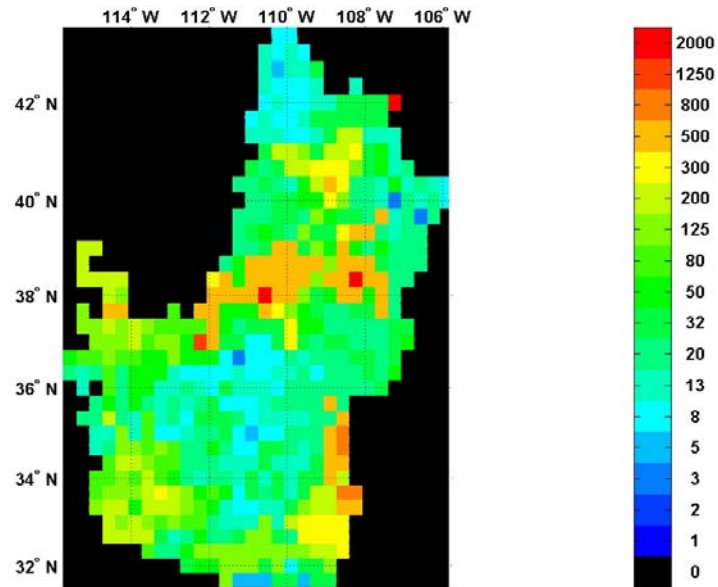
Mapa de *IFS (años)*: precipitaciones anuales, año 1956 (Colorado, EEUU).



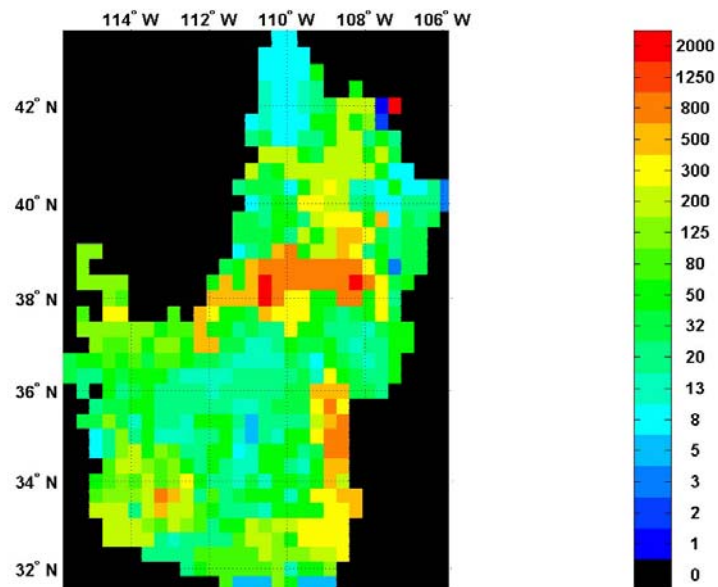
Mapa de *IFS (años)*: precipitaciones anuales, año 1957 (Colorado, EEUU).



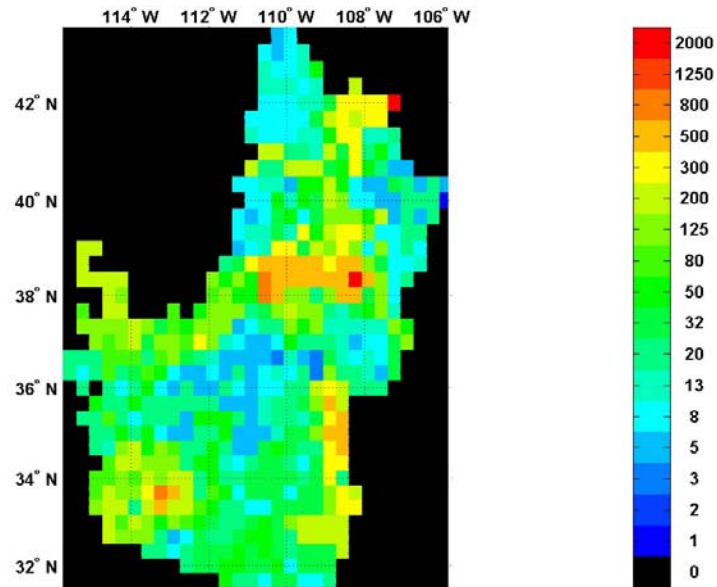
Mapa de *IFS (años)*: precipitaciones anuales, año 1958 (Colorado, EEUU).



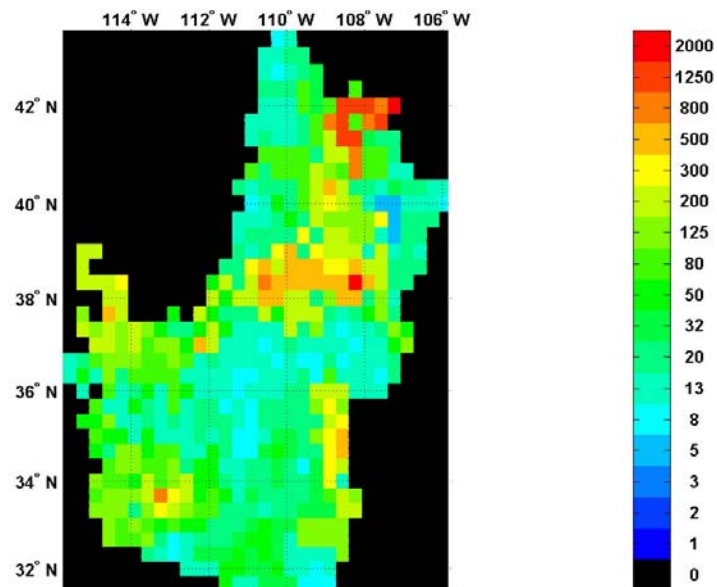
Mapa de *IFS (años)*: precipitaciones anuales, año 1959 (Colorado, EEUU).



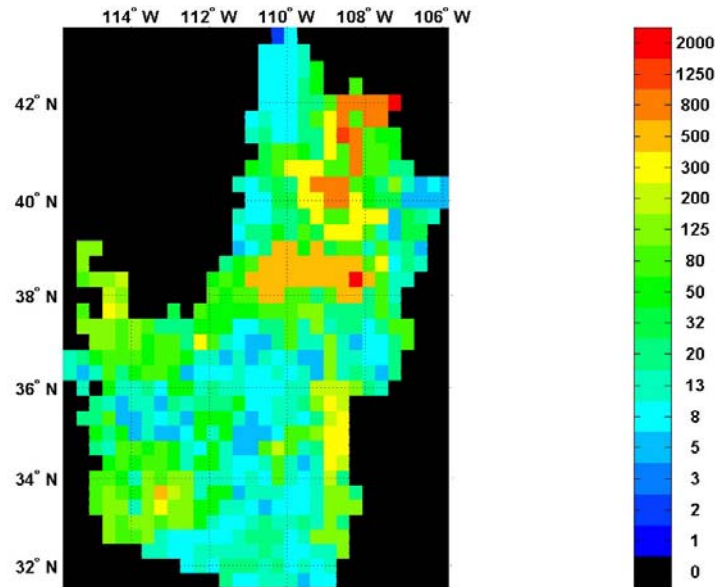
Mapa de *IFS (años)*: precipitaciones anuales, año 1960 (Colorado, EEUU).



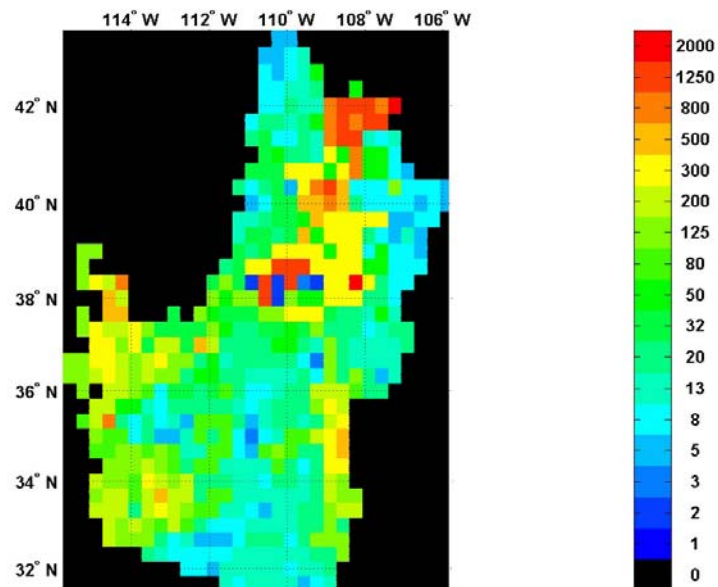
Mapa de *IFS* (años): precipitaciones anuales, año 1961 (Colorado, EEUU).



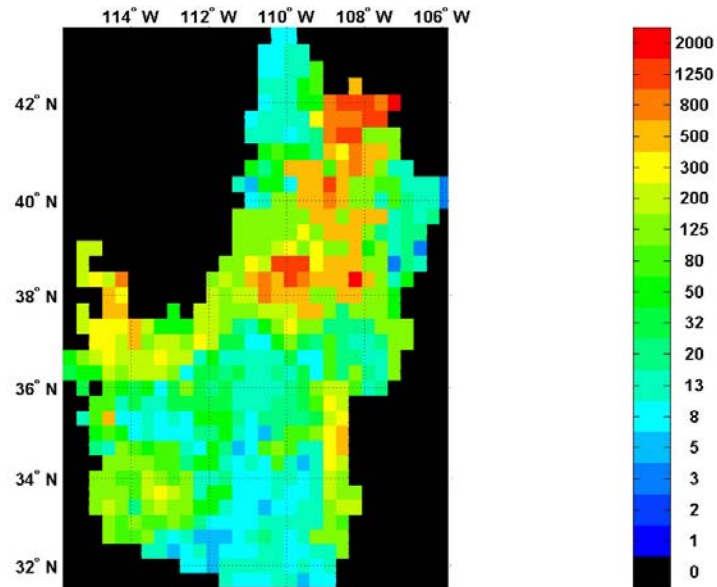
Mapa de *IFS* (años): precipitaciones anuales, año 1962 (Colorado, EEUU).



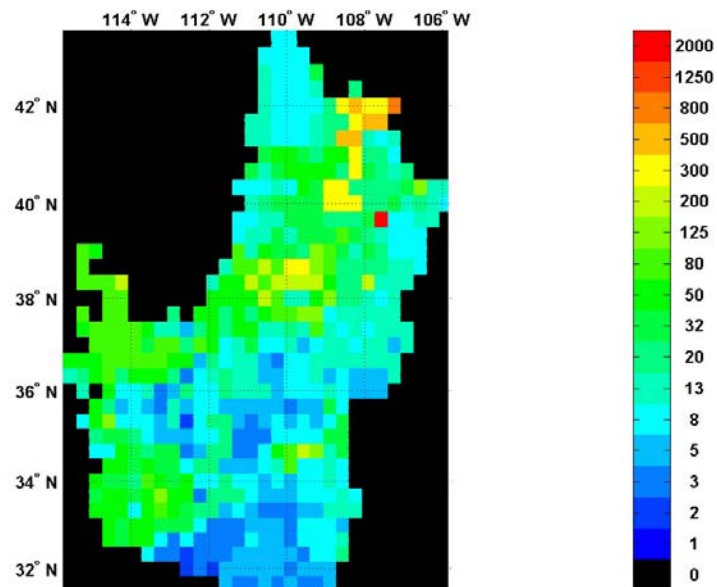
Mapa de *IFS (años)*: precipitaciones anuales, año 1963 (Colorado, EEUU).



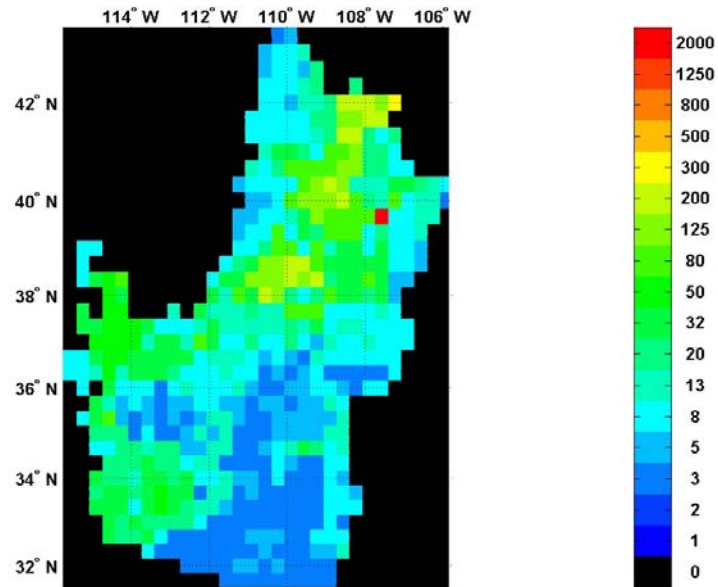
Mapa de *IFS (años)*: precipitaciones anuales, año 1964 (Colorado, EEUU).



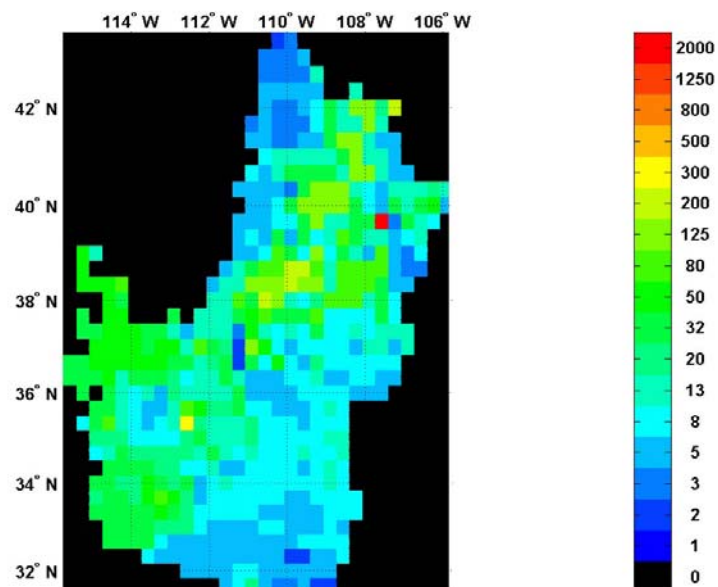
Mapa de *IFS* (años): precipitaciones anuales, año 1965 (Colorado, EEUU).



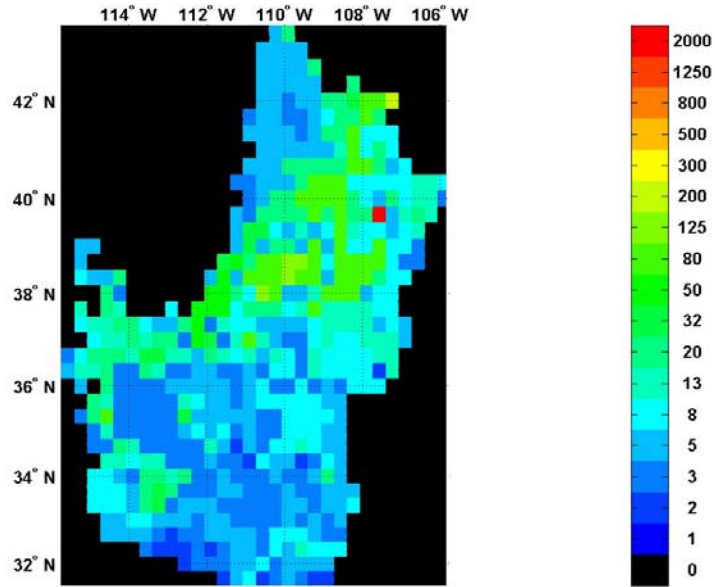
Mapa de *IFS* (años): precipitaciones anuales, año 1966 (Colorado, EEUU).



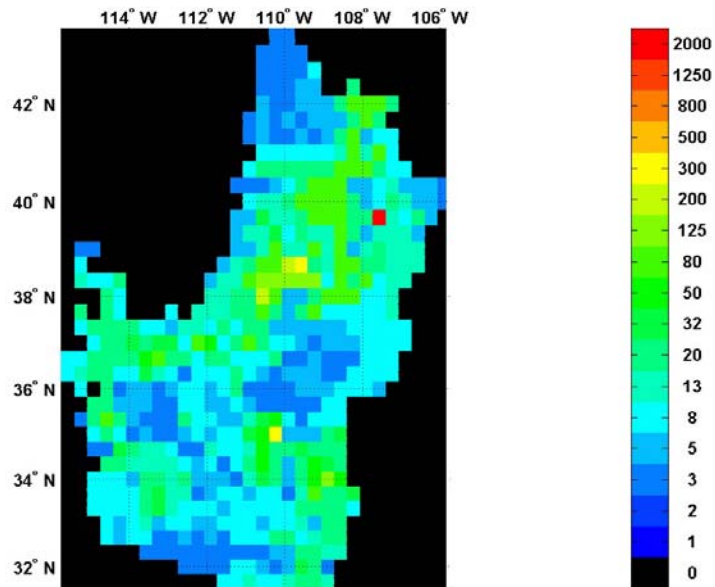
Mapa de *IFS (años)*: precipitaciones anuales, año 1967 (Colorado, EEUU).



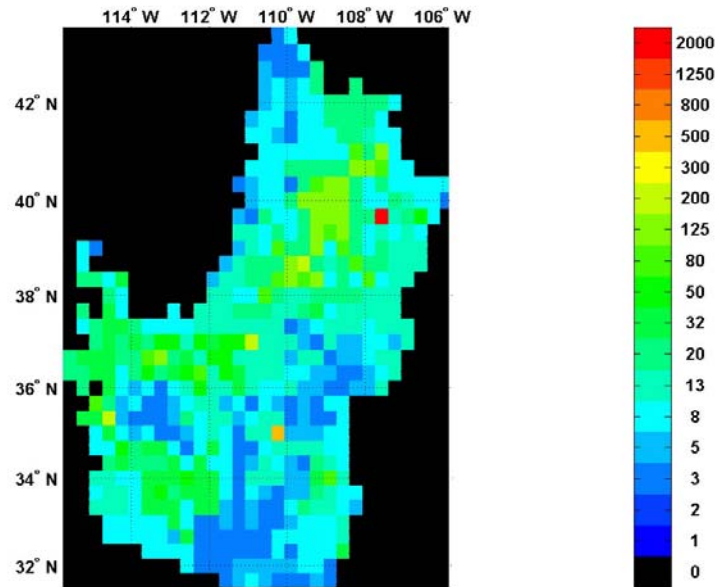
Mapa de *IFS (años)*: precipitaciones anuales, año 1968 (Colorado, EEUU).



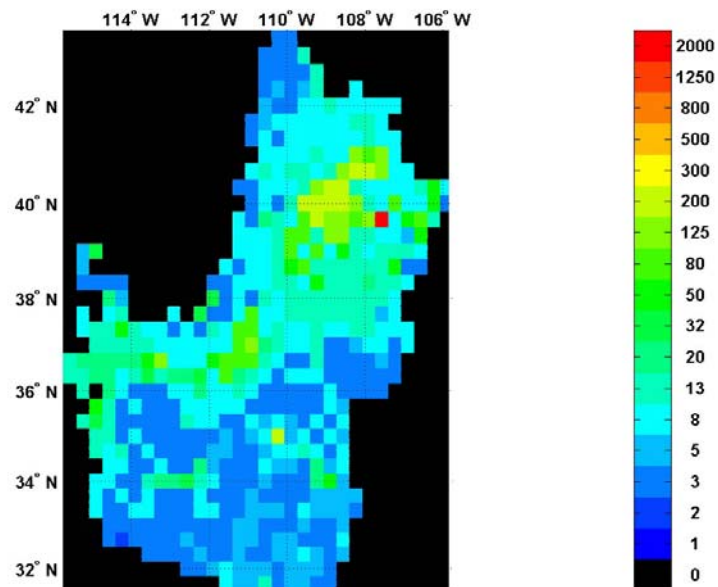
Mapa de *IFS* (años): precipitaciones anuales, año 1969 (Colorado, EEUU).



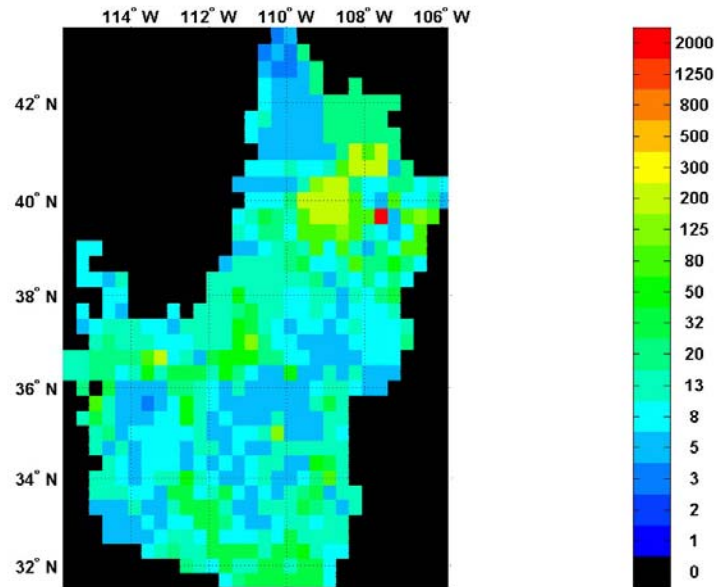
Mapa de *IFS* (años): precipitaciones anuales, año 1970 (Colorado, EEUU).



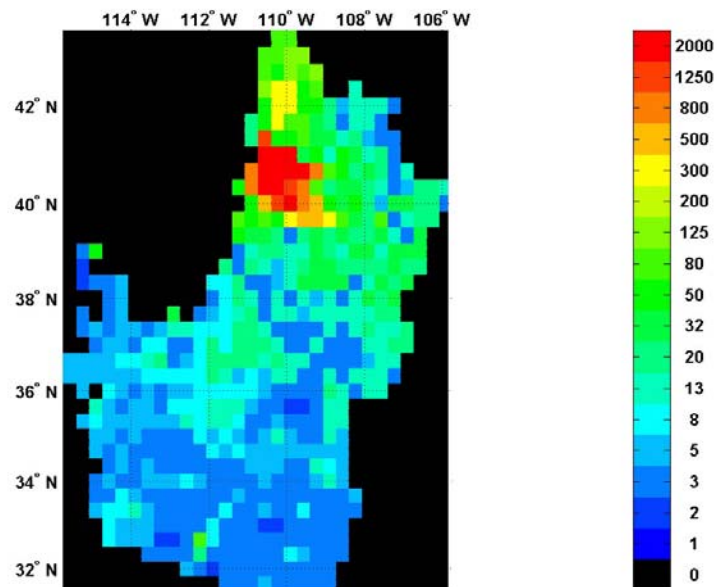
Mapa de *IFS (años)*: precipitaciones anuales, año 1971 (Colorado, EEUU).



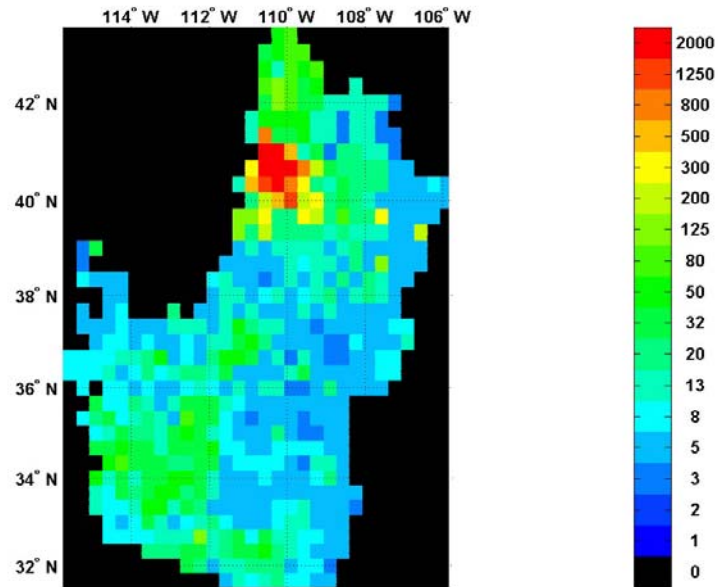
Mapa de *IFS (años)*: precipitaciones anuales, año 1972 (Colorado, EEUU).



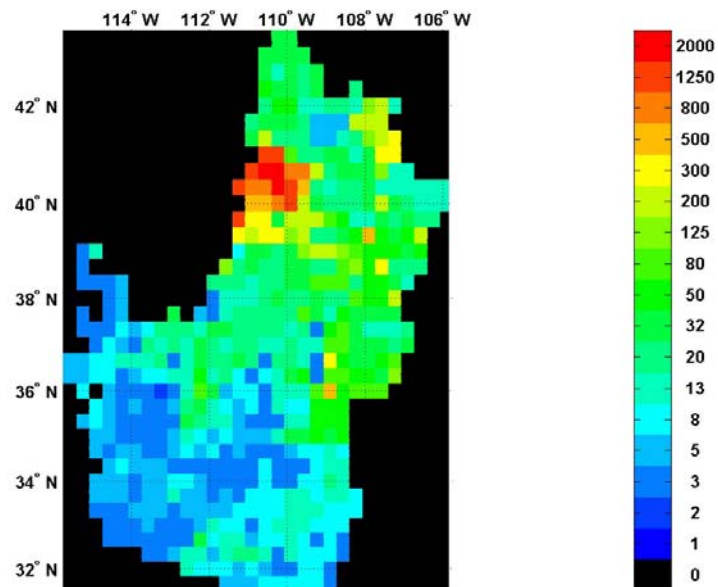
Mapa de *IFS* (años): precipitaciones anuales, año 1973 (Colorado, EEUU).



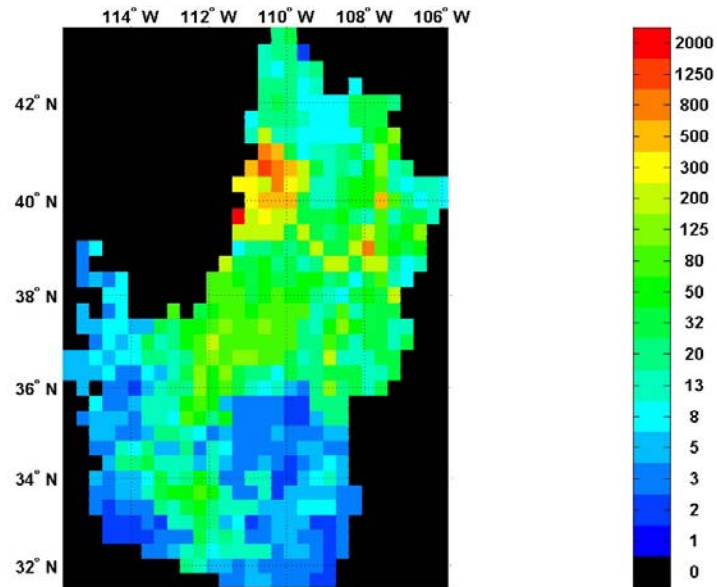
Mapa de *IFS* (años): precipitaciones anuales, año 1974 (Colorado, EEUU).



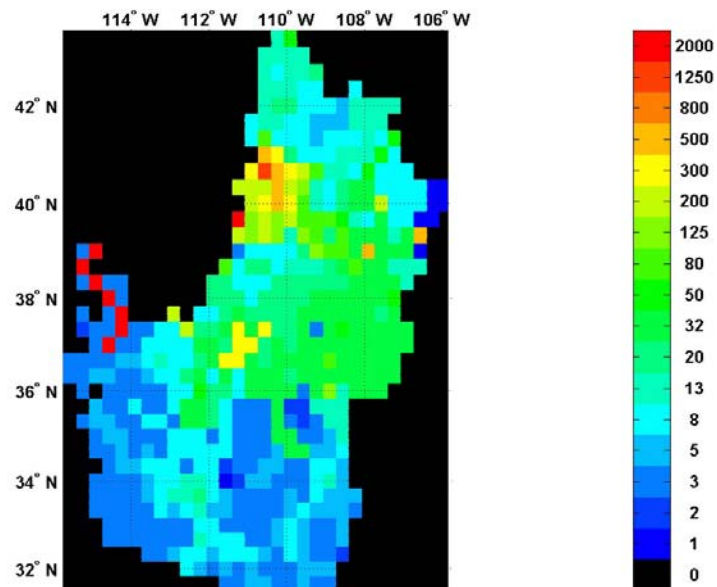
Mapa de *IFS (años)*: precipitaciones anuales, año 1975 (Colorado, EEUU).



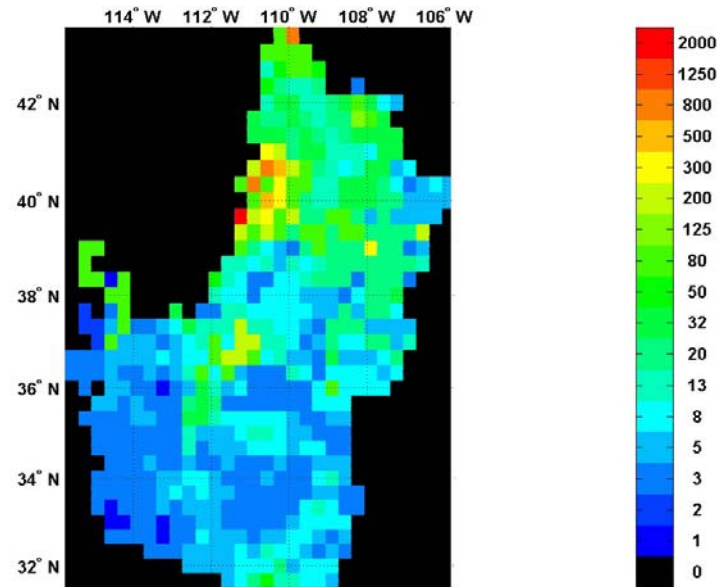
Mapa de *IFS (años)*: precipitaciones anuales, año 1976 (Colorado, EEUU).



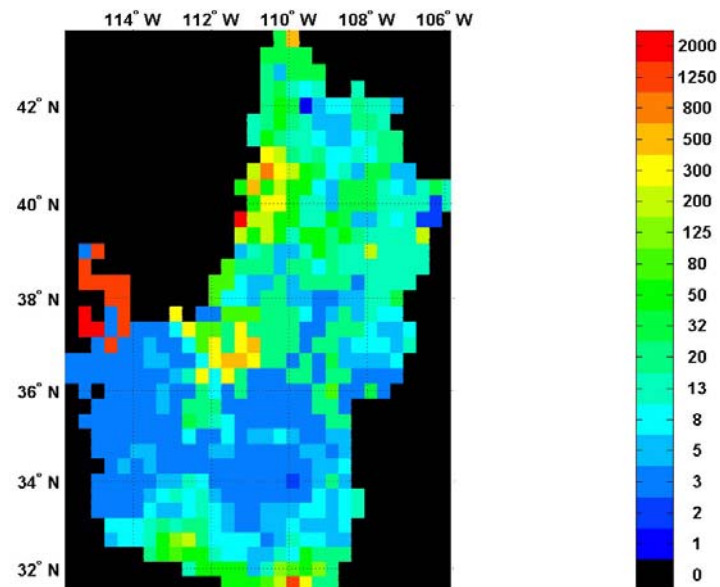
Mapa de *IFS* (años): precipitaciones anuales, año 1977 (Colorado, EEUU).



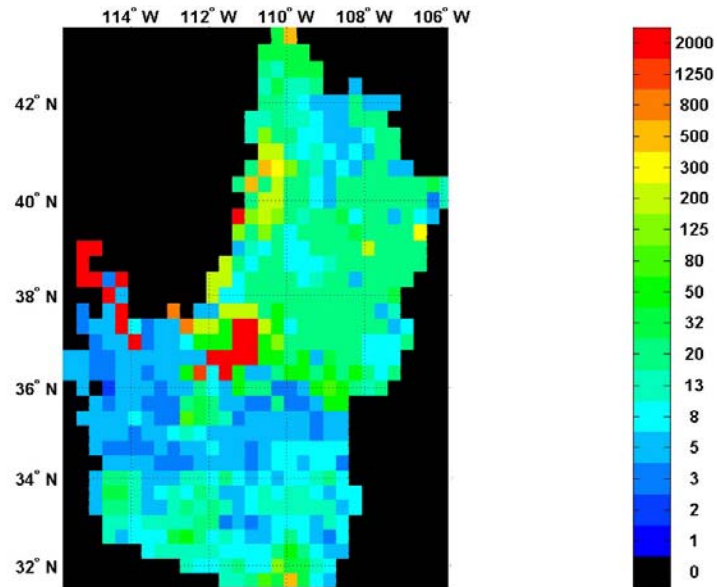
Mapa de *IFS* (años): precipitaciones anuales, año 1978 (Colorado, EEUU).



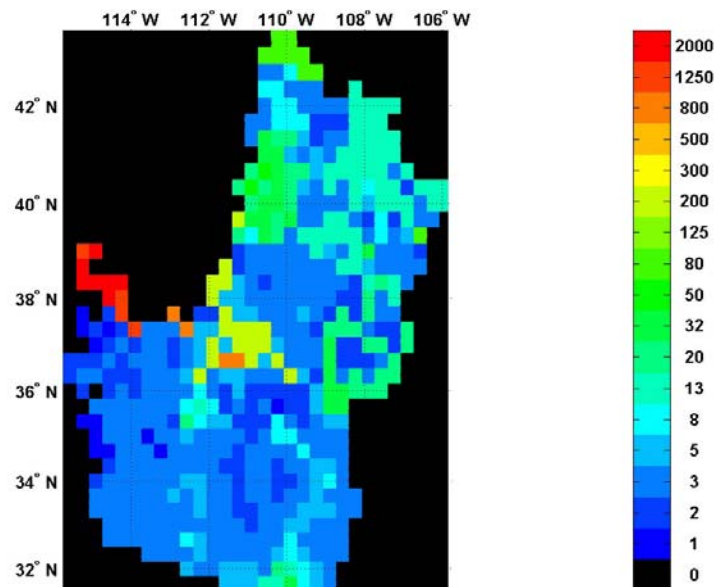
Mapa de *IFS (años)*: precipitaciones anuales, año 1979 (Colorado, EEUU).



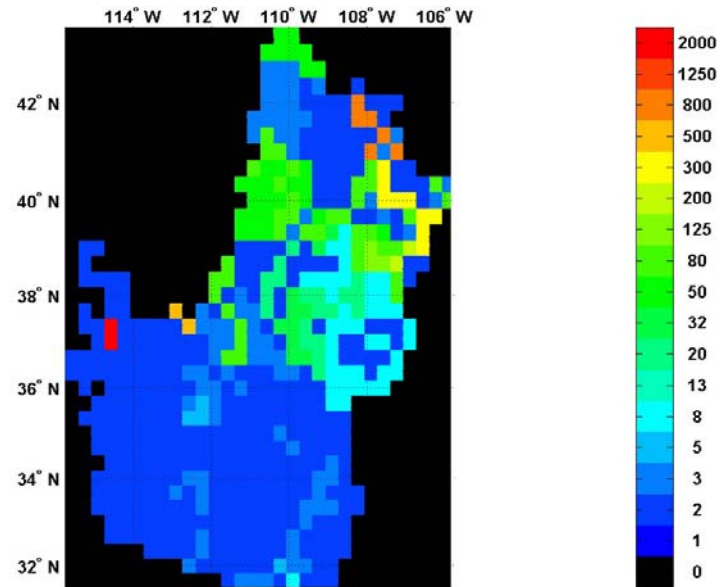
Mapa de *IFS (años)*: precipitaciones anuales, año 1980 (Colorado, EEUU).



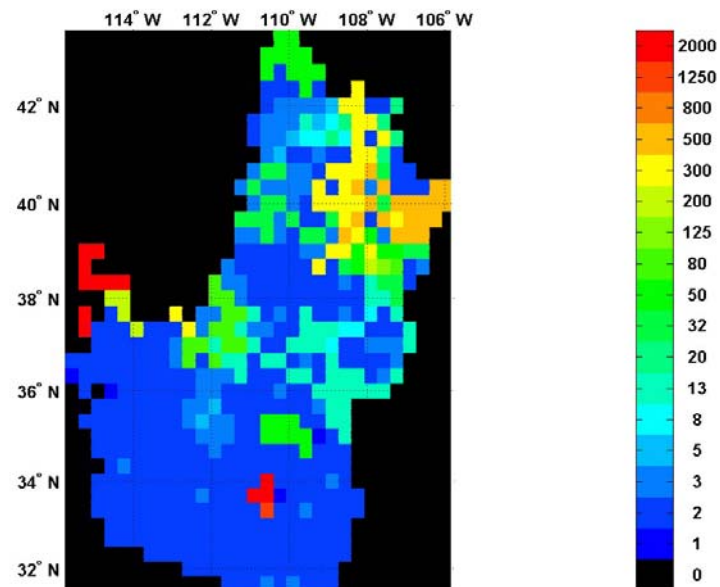
Mapa de *IFS* (años): precipitaciones anuales, año 1981 (Colorado, EEUU).



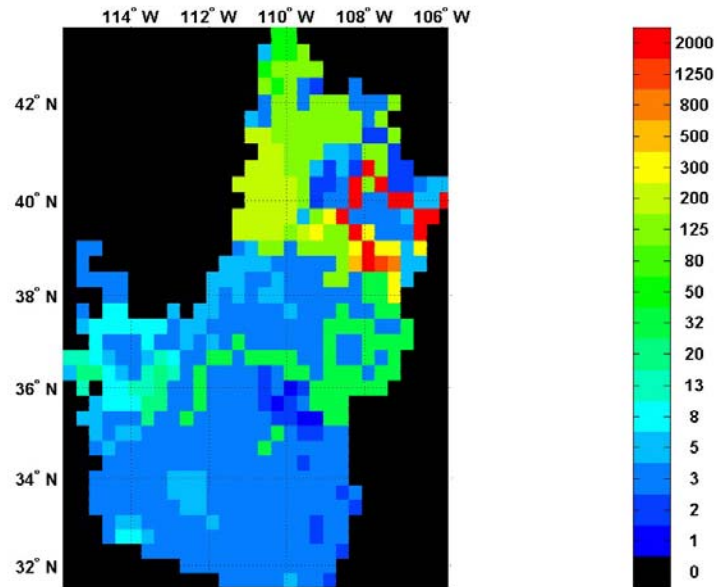
Mapa de *IFS* (años): precipitaciones anuales, año 1982 (Colorado, EEUU).



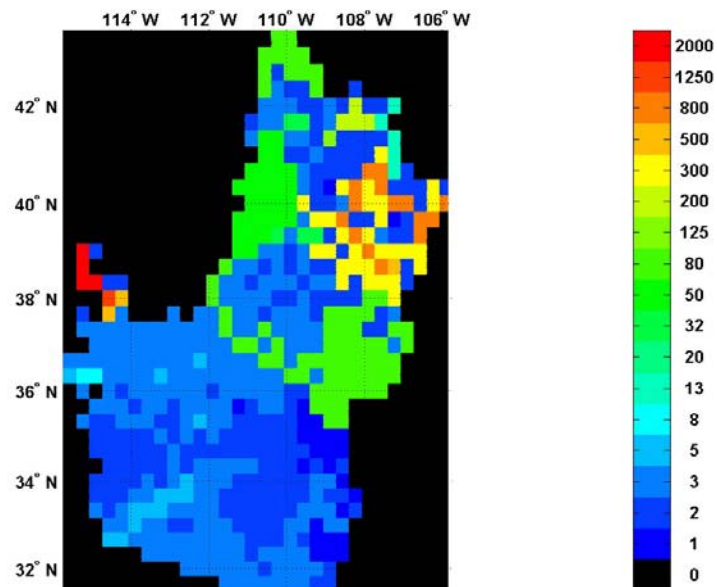
Mapa de *IFS (años)*: precipitaciones anuales, año 1983 (Colorado, EEUU).



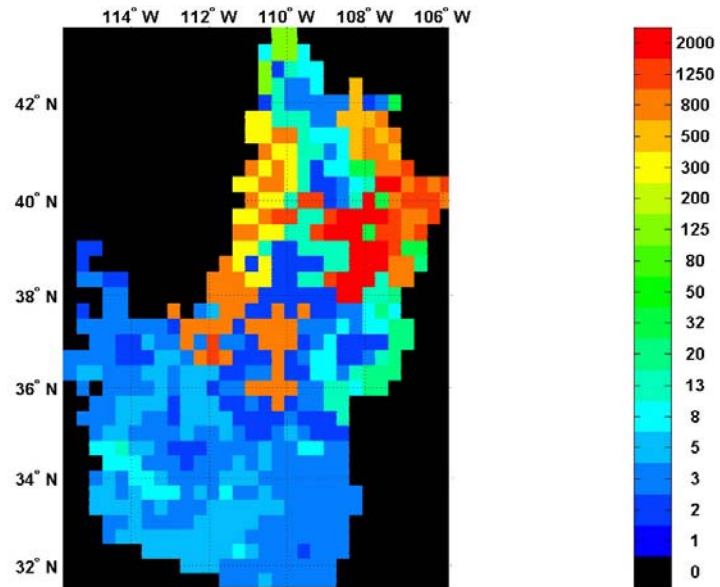
Mapa de *IFS (años)*: precipitaciones anuales, año 1984 (Colorado, EEUU).



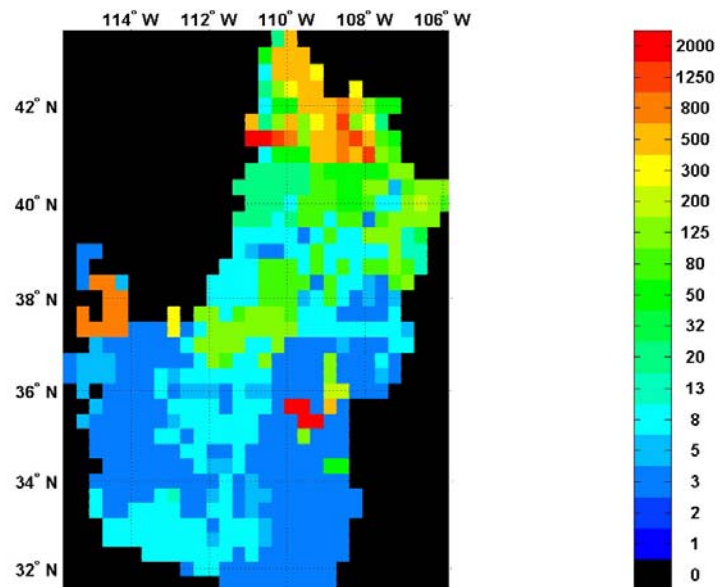
Mapa de *IFS (años)*: precipitaciones anuales, año 1985 (Colorado, EEUU).



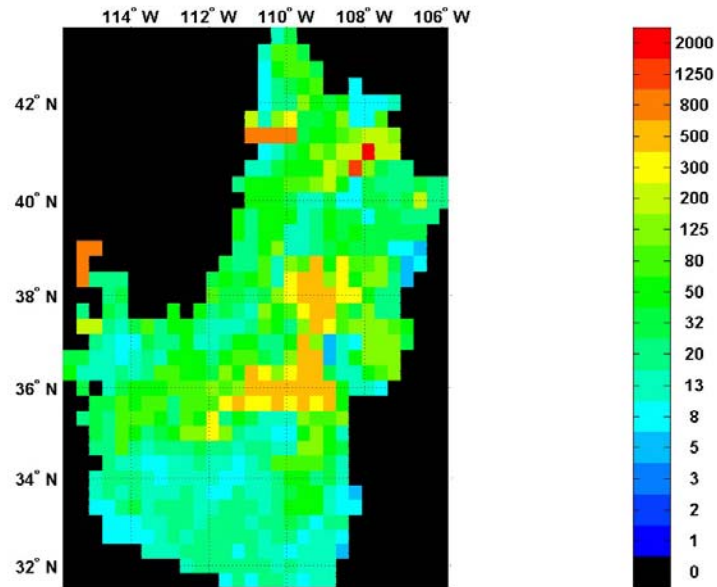
Mapa de *IFS (años)*: precipitaciones anuales, año 1986 (Colorado, EEUU).



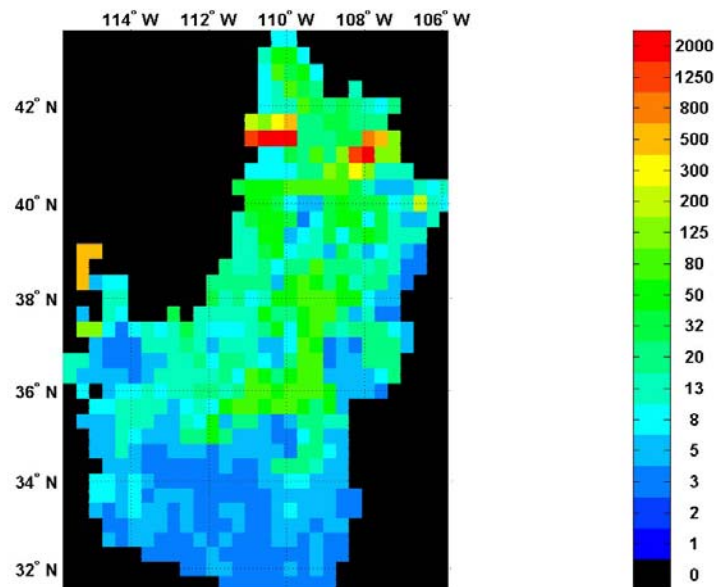
Mapa de *IFS (años)*: precipitaciones anuales, año 1987 (Colorado, EEUU).



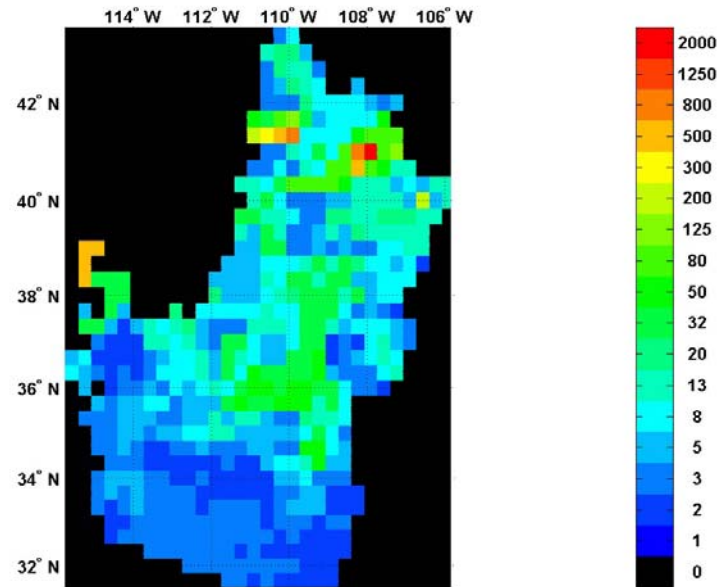
Mapa de *IFS (años)*: precipitaciones anuales, año 1988 (Colorado, EEUU).



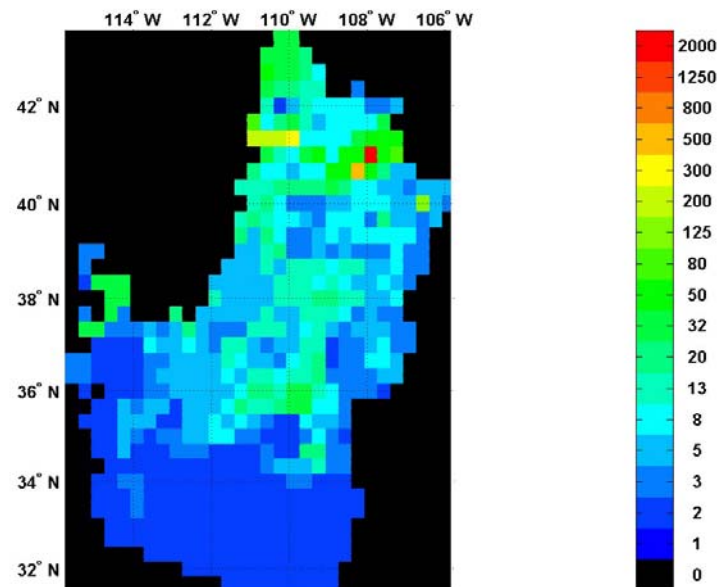
Mapa de *IFS (años)*: precipitaciones anuales, año 1989 (Colorado, EEUU).



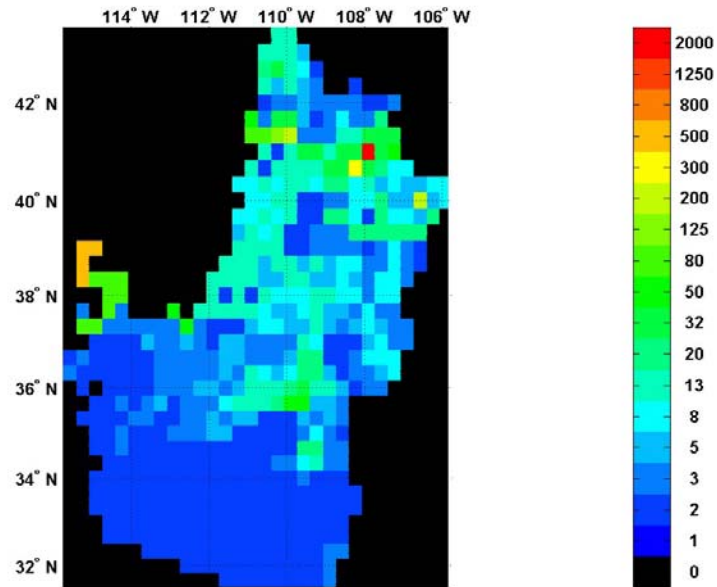
Mapa de *IFS (años)*: precipitaciones anuales, año 1990 (Colorado, EEUU).



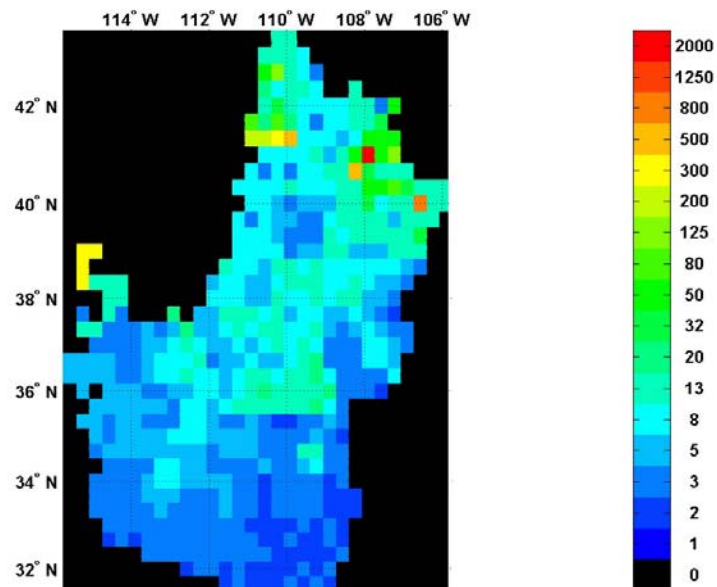
Mapa de *IFS (años)*: precipitaciones anuales, año 1991 (Colorado, EEUU).



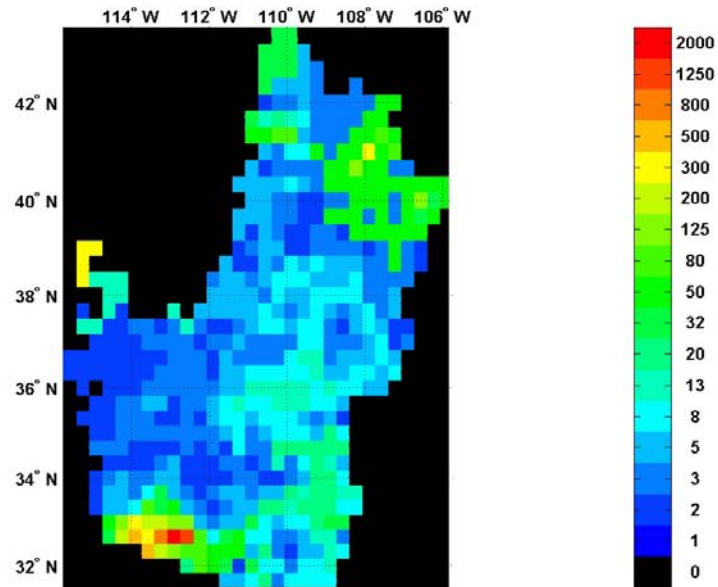
Mapa de *IFS (años)*: precipitaciones anuales, año 1992 (Colorado, EEUU).



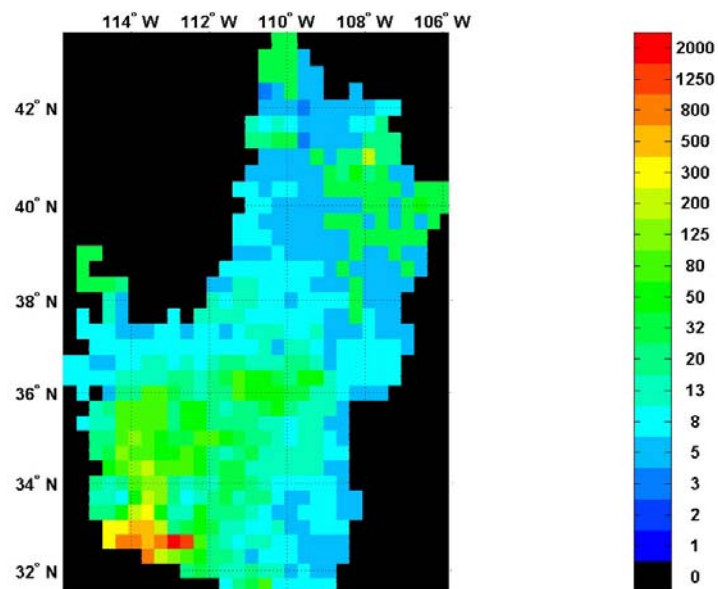
Mapa de *IFS* (*años*): precipitaciones anuales, año 1993 (Colorado, EEUU).



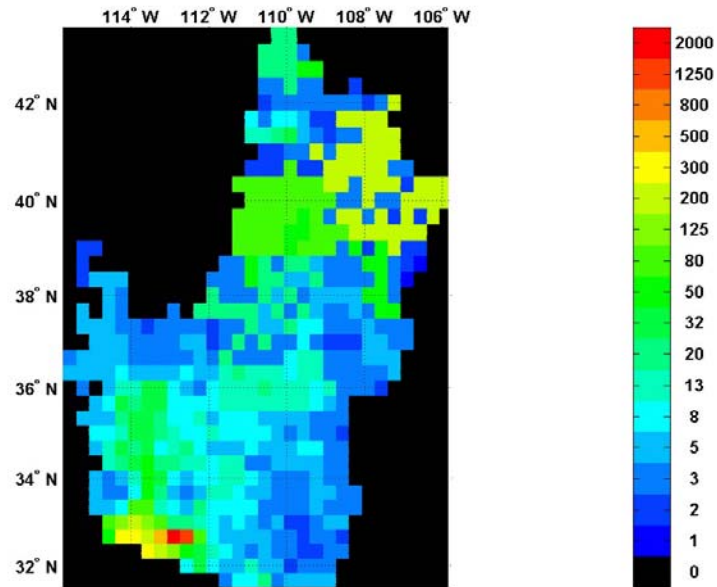
Mapa de *IFS* (*años*): precipitaciones anuales, año 1994 (Colorado, EEUU).



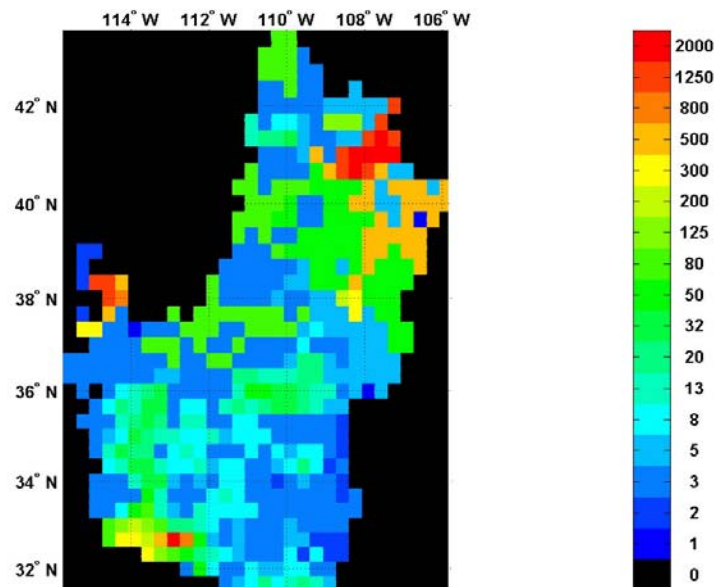
Mapa de *IFS (años)*: precipitaciones anuales, año 1995 (Colorado, EEUU).



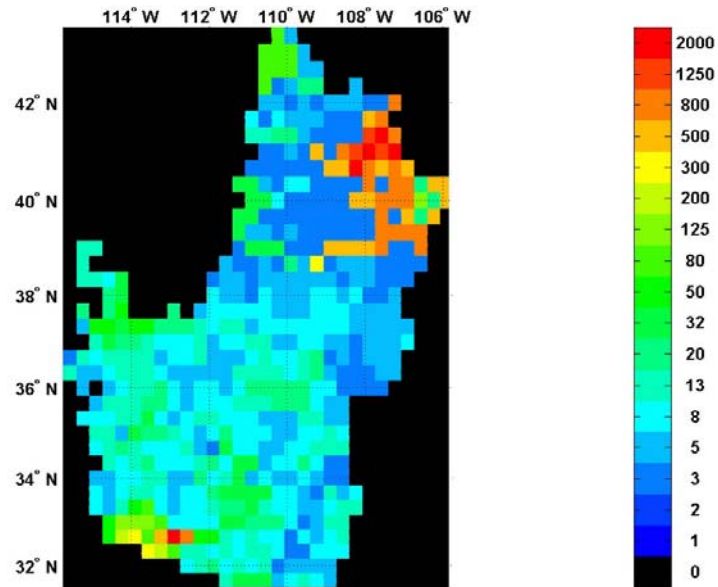
Mapa de *IFS (años)*: precipitaciones anuales, año 1996 (Colorado, EEUU).



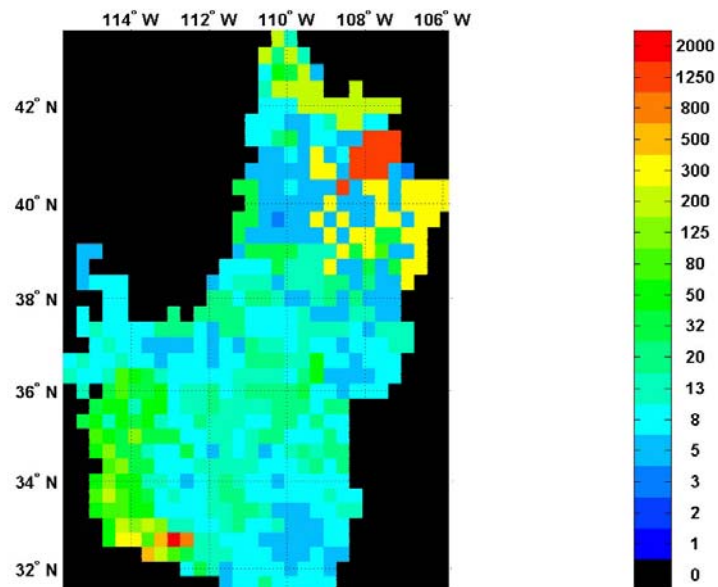
Mapa de *IFS* (años): precipitaciones anuales, año 1997 (Colorado, EEUU).



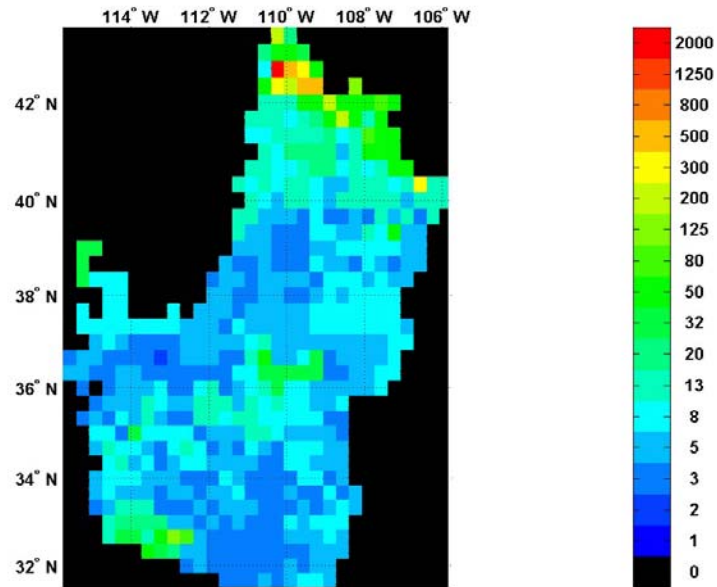
Mapa de *IFS* (años): precipitaciones anuales, año 1998 (Colorado, EEUU).



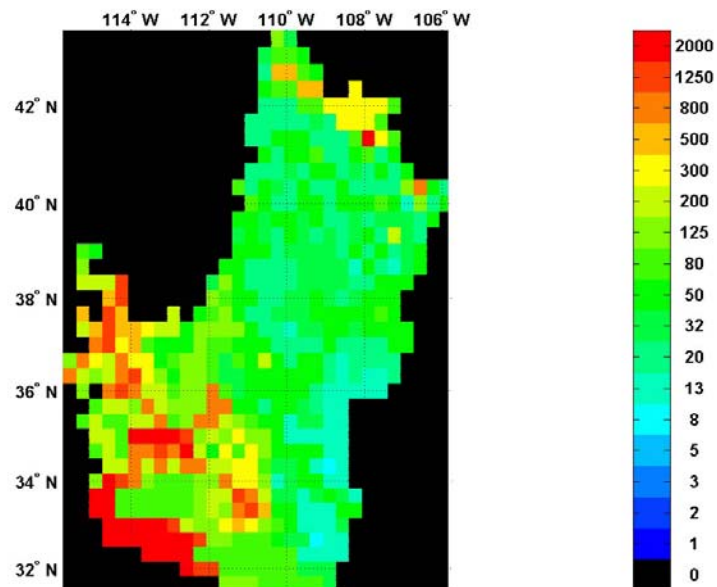
Mapa de *IFS (años)*: precipitaciones anuales, año 1999 (Colorado, EEUU).



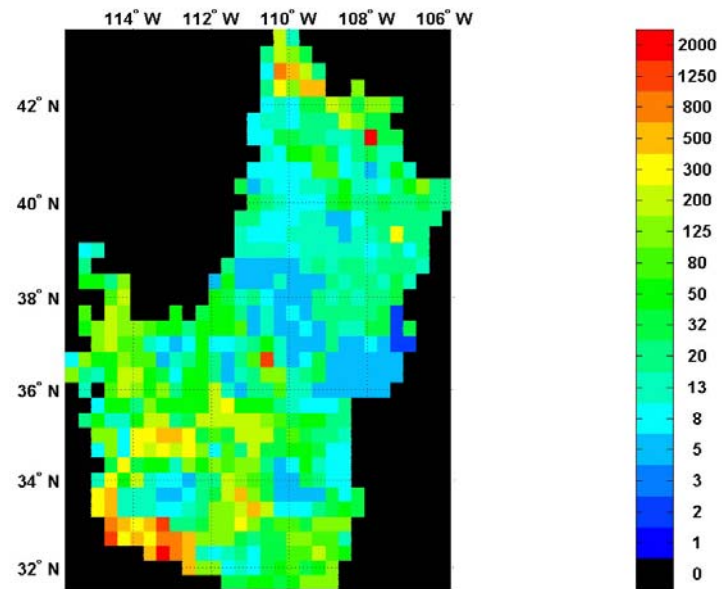
Mapa de *IFS (años)*: precipitaciones anuales, año 2000 (Colorado, EEUU).



Mapa de *IFS* (años): precipitaciones anuales, año 2001 (Colorado, EEUU).



Mapa de *IFS* (años): precipitaciones anuales, año 2002 (Colorado, EEUU).



Mapa de *IFS (años)*: precipitaciones anuales, año 2003 (Colorado, EEUU).

Bibliografía

- ABI-ZEID, I., PARENT, E., & BOBÉE, B. 1998. Computing risk indices based on drought durations. *Páginas 175–184 de: International symposium on water supply and treatment.*
- ABI-ZEID, I., PARENT, E., & BOBÉE, B. 2004. The stochastic modeling of low flows by the alternating point processes approach: Methodology and application. *Journal of hydrology*, **285**, 41–61.
- AKAIKE, H. 1974. A new look at the statistical model identification. *Ieee trans. automat. contr.*, **AC-19**, 716–723.
- ALLEN, R.G., PEREIRA, L.S., RAES, D., & SMITH, M. 1998. *Crop evapotranspiration - guidelines for computing crop water requirements*. FAO Irrigation and drainage paper 56. Food and Agriculture Organization of the United Nations.
- ALLEY, W.M. 1984. The palmer drought severity index: Limitations and assumptions. *J. climate appl. meteor.*, **23**, 1100–1109.
- ÁLVAREZ, J., & ESTRELA, T. 2000. *Regionalization and drought characterisation in europe*. ARIDE. Technical Report 9. Institute of Hydrology, Freiburg, Germany.
- BAHLME, H.N., & MOOLEY. 1980. Large-scale drought/floods and monsoon circulation. *Mon. weather rev.*, **108**, 1197–1211.
- BARLETT, M.S. 1951. A further note on tests of significance in factor analysis. *Brit. j. psych. stat. sec.*, **4**, 1–2.
- BAYAZIT, M., & ÜNAL, N.E. 1990. Effects of hedging on reservoir performance. *Water resources research*, **26**(4), 713–719.
- BOGARDI, I., MATYASOVSKY, I., BARDOSSY, A., & DUCKSTEIN, L. 1994. A hydroclimatological model of areal drought. *Journal of hydrology*, **153**, 245–264.

- BONACCORSO, B., CANCELLIERE, A., & ROSSI, G. 2003. An analytical formulation of return period of drought severity. *Stochastic environ. res. risk assess.*, **17**, 157–174. doi:10.1007/s00477-003-0127-7.
- BOX, G.E.P., & COX, D.R. 1964. An analysis of transformations. *Journal of the royal statistical society: Series b*, **26**, 211–252.
- BOX, G.E.P., & JENKINS, G.M. 1976. *Time series analysis, forecasting, and control*. 2nd edn. San Francisco: Holden-Day.
- BRASS, R.L. 1990. *Hydrology: An introduction to hydrologic science*. Reading, Mass: Addison-Wesley-Longman.
- CABRINHA, V., & SANTO, F.E. 2000. Drought tendencies in mainland portugal. *En: Drought early warning systems for drought preparedness and drought management*. Lisbon, Portugal: World Meteorological Organization.
- CANCELLIERE, A., & SALAS, J.D. 2004. Drought length properties for periodic stochastic hydrologic data. *Water resources research*, **40**, W02503, doi:10.1029/2002WR001750.
- CHAUDHRY, M.A., & ZUBAIR, S.M. 2002. *On a class of incomplete gamma functions with applications*. Florida: Chapman & Hall/CRC.
- CHUNG, C., & SALAS, J.D. 2000. Drought occurrence probability and risks of dependent hydrologic processes. *J. hydrologic eng.*, **5**(3), 259–268.
- COLLADO, J., COLLADO, W., & ANA, I. 2001 (Septiembre). *Proyecto de reglamento de distribución y uso de las aguas superficiales de la cuenca del río bravo*. Tech. rept. Instituto Mexicano de Tecnología del Agua (IMTA).
- COX, D.R., & ISHAM, V. 1980. *Point processes*. Chapman & Hall.
- CRANLEY, R., & PATTERSON, T.N.L. 1976. Randomization of number theoretic methods for multiple integration. *Siam journal on numerical analysis*, **13**, 904–914.
- DALY, C., & NEILSON, R.P. 1992. A digital topographic approach to modeling the distribution of precipitation in mountainous terrain. *Páginas 437–454 de: JONES, M.E., & LAENEN, A. (eds), Interdisciplinary approaches in hydrology and hydrogeology*. American Institute of Hydrology.

- DALY, C., NEILSON, R.P., & PHILLIPS, D.L. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of applied meteorology*, 140–158.
- DALY, C., GIBSON, W., & TAYLOR, G. 2002. *103-year high-resolution precipitation climate data set for the conterminous united states*. Tech. rept. Spatial Climate Analysis Service, Corvallis, Oregon. <ftp://ftp.ncdc.noaa.gov/pub/data/prism100>.
- DAVIS, P.J., & RABINOWITZ, P. 1984. *Methods of numerical integration*. 2nd edn. San Diego: Academic Press.
- DOWNER, R., SIDDIQUI, M., & YEVJEVICH, V. 1967. Applications of runs to hydrologic droughts. *En: International hydrology symposium*. Fort Collins, CO: Colorado State University.
- DRACUP, J.A., & KAHYA, E. 1994. The relationships between u.s. stream-flow and la niña. *Water resources research*, **30**(7), 2133–2141.
- DRACUP, J.A., LEE, K.S., & PAULSON, E.G. 1980a. On the definition of droughts. *Water resources research*, **16**(2), 297–307.
- DRACUP, J.A., LEE, K.S., & PAULSON, E.G. 1980b. On the statistical characteristics of drought events. *Water resources bulletin*, **26**(5), 823–829.
- ESTRADA, F. 1994. *Garantías en los sistemas de explotación de los recursos hidráulicos*. Madrid, España: CEDEX, Ministerio de Obras Públicas, Transportes y Medio Ambiente.
- FERNÁNDEZ, B., & SALAS, J.D. 1990. Gamma-autoregressive models for streamflow simulation. *Journal of hydraulic engineering*, **116**(11), 1403–1414.
- FERNÁNDEZ, B., & SALAS, J.D. 1999. Return period and risk of hydrologic events i: Mathematical formulation. *Journal of hydrologic engineering*, **4**(4), 297–307.
- FERNÁNDEZ, B., & VERGARA, A. 1998. Risk of scarcity of monthly precipitation and streamflows in semiarid regions. *Hydrol. sci. j.*, **43**(5), 759–773.
- FRICK, D.M., BODE, D., & SALAS, J.D. 1990. Effect of drought on urban water supplies i: Drought analysis. *J. hydraul. eng.*, **116**(6), 733–753.

- GIBBS, W.J., & MAHER, J.V. 1967. *Rainfall deciles as drought indicators*. Bureau of Meteorology Bulletin 48. Commonwealth of Australia, Melbourne.
- GONZÁLEZ, J., & VALDÉS, J.B. 2003. Bivariate drought recurrence analysis using tree ring reconstructions. *J. hydrologic eng.*, **8**(5), 247–258.
- GUPTA, V., & DUCKSTEIN, L. 1975. Droughts as extreme distributions from a point rainfall process. *Water resources research*, **11**(2), 221–228.
- GUTTMAN, N.B. 1991. A sensitivity analysis of the palmer hydrologic drought index. *Journal of the american water resources association*, **25**(4), 797–807.
- GUTTMAN, N.B. 1998. Comparing the palmer drought index and the standardized precipitation index. *Journal of the american water resources association*, **34**(1), 1029–1039.
- GUTTMAN, N.B., WALLIS, J.R., & HOSKING, J.R.M. 1992. Spatial comparability of the palmer drought severity index. *Journal of the american water resources association*, **28**(6), 1111–1119.
- GUVEN, O. 1983. A simplified semiempirical approach to probabilities of extreme hydrologic droughts. *Water resources research*, **19**(2), 441–453.
- HEDDINGHAUS, T.R., & SABOL, P. 1991. A review of the palmer drought severity index and where do we go from here? *Páginas 242–246 de: 7th conf. on applied climatology*. Boston, Massachusetts: Amer. Meteor. Soc.
- HENRIQUES, A.G., & SANTOS, M.J.J. 1999. Regional drought distribution model. *Physics and chemistry of the earth. part b. hydrology, oceans and atmosphere. european water resources and climate change processes. european geophysical society*, **24**(1/2), 19–22.
- HISDAL, H., & TALLAKSEN, L.M. 2003. Estimation of regional meteorological and hydrological drought characteristics: A case study for denmark. *J. hydrology*, **281**, 230–247.
- HUSCHKE, R.E. 1959. *Glossary of meteorology*. American Meteorological Society.
- ISLAM, S., BRASS, R.L., & RODRÍGUEZ-ITURBE, I. 1988. Multidimensional modeling of cumulative rainfall: Parameter estimation and model adequacy through a continuum of scales. *Water resources research*, **24**(7), 985–992.

- JOLLIFFE, I.T. 1986. *Principal components analysis*. New York: Springer-Verlag.
- KAISER, H.F. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187–200.
- KANYA, E., & DRACUP, J.A. 1993. U.s. streamflow patterns in relation to the Niño/Southern oscillation. *Water resources research*, **29**(8), 2491–2503.
- KARL, T.R. 1983. Some spatial characteristic of drought duration in the united states. *J. climate appl. meteor.*, **22**, 1356–1366.
- KARL, T.R. 1986. The sensitivity of the palmer drought severity index and palmer's a-index to their calibration coefficients including potential evapotranspiration. *J. climate appl. meteor.*, **25**, 77–86.
- KARL, T.R., & KNIGHT, R.W. 1985. Atlas of monthly palmer hydrological drought indices (1931–1983) for the contiguous united states. *En: Historical climatology series 3-7*. Asheville, North Carolina: National Climatic Data Center.
- KAVVAS, M.L., & DELLEUR, J.W. 1976. Methodology for the selection and application of probability models for the simulations of daily rainfall and runoff. *Hydrological sciences bulletin*, **XXI**, 99–111.
- KELLY, M.E. 2001. *The río conchos: A preliminary overview*. Tech. rept. Texas Center for Policy Studies.
- KEYANTASH, J.A., & DRACUP, J.A. 2004. An aggregate drought index: Assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water resources research*, **40**, **W09304**, doi: **10.1029/2003WR002610**.
- KIM, T.W., VALDÉS, J.B., & APARICIO, J. 2003a. Frequency and spatial characteristic of drought in the conchos river basin, mexico. *Water intern.*, **27**(3), 420–430.
- KIM, T.W., VALDÉS, J.B., & YOO, C. 2003b. A nonparametric approach for estimating return periods of drought in arid regions. *J. hydrologic eng.*, **8**(5), 237–246.
- KNOPP, K. 1956. *Infinite sequences and series*. Dover Publications Inc.
- KOEPSSELL, R.W., & VALDÉS, J.B. 1991. Multidimensional rainfall parameter estimation from a sparse network. *J. hydraulic eng.*, **117**(7), 832–850.

- KOGAN, E.N. 1995. Droughts of the late 1980's in the united state as derived from NOAA polar-orbiting satellite data. *Bulletin of the american meteorological society*, **76**, 655–668.
- KOLMOGOROV, A.N. 1956. *Foundations of the theory of probability*. Second english edn. New York: Chelsea Publishing Company.
- KOROBOV, N.M. 1959. The approximate computation of multiple integrals. *Doklady akademii nauk sssr*, **124**, 1207–1210. Russian.
- KOROBOV, N.M. 1960. Properties and calculation of optimal coefficients. *Doklady akademii nauk sssr*, **132**, 1009–1012.
- LALL, U., & MILLER, C.W. 1988. An optimization model for screening multipurpose reservoir systems. *Water resources research*, **24**(7), 953–968.
- LE CAM, L.A. 1961. A stochastic description of precipitation. *Páginas 165–186 de: NEWMAN, J. (ed), 4th berkeley symp. mathematics, statistics and probability*. Berkeley: University of California Press.
- LEE, K.S., SADEGHIPOUR, J., & DRACUP, J.A. 1986. An approach for frequency analysis of multiyear drought durations. *Water resources research*, **22**(5), 655–662.
- LINDGREN, B. W. 1976. *Statistical theory*. 3rd edn. New York: Collier Macmillan.
- LLAMAS, J., & SIDDIQUI, M. 1969. *Runs of precipitation series*. Hydrol. Report 33. Colorado State University, Fort Collins, CO.
- LLOYD, E.H. 1970. Return period in the presence of persistence. *Journal of hydrology*, **10**(3), 202–215.
- LOUKAS, A., VASILIADES, L., & DALEZIOS, N.R. 2002. Hydroclimatic variability of regional droughts in greece using the palmer moisture anomaly index. *Nordic hydrology journal*, **33**(5), 425–442.
- MADSEN, H., & ROSBJERG, D. 1995. On the modelling of extreme droughts. *Iahs publ.*, **231**, 377–385.
- MATHIER, L., PERREAULT, L., & BOBÉE, B. 1992. The use of geometric and gamma related distributions for frequency analysis of water deficit. *Stochastoc. hydrol. hydraul.*, **6**, 239–254.

- MCKAY, M.D., BECKMAN, R.J., & CONOVER, W.J. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 239–245.
- MCKEE, T.B., DOESKEN, N.J., & KLEIST, J. 1993. Drought monitoring with multiple timescales. *Páginas 179–184 de: Proc. 8th conference on applied climatology*. Anaheim, CA: American Meteorological Society.
- MILLAN, J., & YEVJEVICH, V. 1971. *Probability of the observed droughts*. Hydrol. Pap. 50. Colorado State University, Fort Collins, CO.
- MONTELLER, F., & WALLACE, D.L. 1963. Inference in a autorship problem. *J. amer. stat. assoc.*, **58**, 275–309.
- MORAN, P.A.P. 1969. Statistical inference with bivariate gamma distributions. *Biometrika*, **56**, 627–634.
- MOROKOFF, W., & CAFLISCH, R. 1994. Quasi-monte carlo sequences and their discrepancies. *Siam j. sci. stat. comput.*, **15**, 1251–1279.
- MOROKOFF, W., & CAFLISCH, R. 1995. Quasi-monte carlo integration. *J. comp. phys.*, **122**, 218–230.
- NEYMAN, J., & SCOTT, E.L. 1958. Statistical approach to problems of cosmology. *J. roy. statistical soc., serie b*, **20**(1), 1–43.
- NIEDERREITER, H. 1992. *Random number generation and quasi-monte carlo methods*. Philadelphia: SIAM.
- NORRIS, J.R. 1997. *Markov chains*. New York: Cambridge University Press.
- OWEN, A.B. 1995. Randomly permuted (t,m,s)-nets and (t,s)-sequences. *Páginas 299–317 de: NIEDERREITER, H., & SHIUE, P.J.-S. (eds), Monte carlo and quasi-monte carlo methods in scientific computing*. New York: Springer-Verlag.
- OWEN, A.B. 1997. Monte carlo variance of scrambled equidistribution quadrature. *Siam journal of numerical analysis*, **34**(5), 1984–1910.
- OWEN, A.B. 1998. Latin supercube sampling for very high dimensional simulations. *Acm transaction on modeling and computer simulation*, **8**(1), 71–102.
- PALMER, W.C. 1965. *Meteorological drought*. Research Paper 45. U. S. Department of Commerce, Weather Bureau, Washington, D.C.

- PALMER, W.C. 1968. Keeping track of crop moisture conditions, nationwide: The new crop moisture index. *Weatherwise*, **21**, 156–161.
- PANDIT, S.M., & WU, S.-M. 2001. *Time series and system analysis with applications*. Malabar, Florida: Krieger P. C.
- PEEK, E.L., & BROWN, M.J. 1962. An approach to the development of isohyetal map for mountainous areas. *J. geophys. res.*, **67**, 681–694.
- RAMÍREZ, J.A., & BRASS, R.L. 1985. Conditional distributions of neyman-scott models for storm arrivals and their use in irrigation control. *Water resources research*, **21**(3), 317–330.
- REDMOND, K.T., & KOCH, R.W. 1991. Surface climate and streamflow variability in the wester united states and their relationship to large-scale circulation indices. *Water resources research*, **27**(9), 2381–2399.
- RODRÍGUEZ-ITURBE, I., GUPTA, V.K., & WAYMIRE, E. 1984. Scale considerations in the modeling of temporal rainfall. *Water resources research*, **20**(11), 1611–1619.
- RODRÍGUEZ-ITURBE, I., FEBRES DE POWER, B., & VALDÉS, J.B. 1987a. Rectangular pulses point process models for rainfall: Analysis of empirical data. *J. geophys. res.*, **92**(D8), 9645–9656.
- RODRÍGUEZ-ITURBE, I., COX, D.R., & ISHAM, V. 1987b. Some models for rainfall based on stochastic point processes. *Proc. royal soc. london (a)*, **410**, 269–288.
- ROSENBLATT, M. 1952. Remarks on a multivariate transformation. *Ann. math. stat.*, **23**(3), 470–472.
- ROSSI, G., BENEDINI, M., TSAKIRIS, G., & GIAKOUMAKIS, S. 1992. On regional drought estimation and analysis. *Water resources management*, **6**, 249–277.
- SALAS, J.D. 1992. Analysis and modeling of hydrologic time series. *Chap. 19 de: MAIDMENT, D.R. (ed), Handbook of hydrology*. McGraw-Hill.
- SALAS, J.D., DELLEUR, J.R., YEVJEVICH, V., & LANE, W.L. 1980. *Applied modeling of hydrologic time series*. Littleton, Colorado: Water Resources Publications.
- SANKARASUBRAMANIAN, A., & VOGEL, R.M. 2002. Annual hydroclimatology of the united states. *Water resources research*, **38**(6), 1083, 10.1029/2001WR000619.

- SANTOS, M.A. 1983. Regional droughts: A stochastic characterization. *Journal of hydrology*, **66**, 183–211.
- SCHERMENHORN, V.P. 1967. Relations between topography and annual precipitation in western oregon and washington. *Water resources research*, **3**, 707–711.
- SCHUBERT, S.D., SUZREZ, M.J., PEGION, P.J., KOSTER, R.D., & BACMEISTER, J.T. 2004. On the cause of the 1930s dust bowl. *Science*, **303**, 1855–1859.
- SEN, Z. 1976. Wet and dry periods of annual flows series. *J. hydr. div., asce*, **102**(10), 1503–1514.
- SEN, Z. 1977. Run-sums of annual flow series. *Journal of hydrology*, **35**, 311–324.
- SEN, Z. 1980a. Regional drought and flood frequency analysis theoretical consideration. *Journal of hydrology*, **46**, 265–279.
- SEN, Z. 1980b. Statistical analysis of hydrologic critical droughts. *J. hydr. div., asce*, **106**(1), 99–115.
- SHAFER, B.A., & DEZMAN, L.E. 1982. Development of a surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas. *Páginas 164–175 de: Proceedings of the western snow conference*. Colorado State University, Fort Collins, Colorado.
- SHARMA, T. 1995. Estimation of drought severity on independent and dependent hydrologic series. *Water resources management*, **11**, 35–49.
- SHIAU, J., & SHEN, H.W. 2001. Recurrence analysis of hydrologic drought of differing severity. *J. of water resources planning and management*, **127**(1), 30–40.
- SLOAN, I.H., & JOE, S. 1994. *Lattice methods for multiple integration*. New York: Oxford University Press Inc.
- SNYDER, L. 1975. *Random point processes*. New York: Wiley.
- SPREEN, W.C. 1947. A determination of the effect of topography upon precipitation. *Trans. am. geophys. union*, **28**, 285–290.
- TALLAKSEN, L.K., MADSEN, H., & CLAUSEN, B. 1997. On the definition and modelling of streamflow drought duration and deficit volume. *Hydrological sciences bulletin*, **42**(1), 15–33.

- TASE, N. 1976. *Area-deficit-intensity characteristics of drought*. Hydrology Papers 87. Colorado State University, Fort Collins, USA.
- TATE, E.L., & GUSTARD, A. 2000. Drought definition: A hydrological perspective. *Páginas 23–48 de: VOGT, J.V., & SOMMA, F. (eds), Drought and drought mitigation in europe*. Netherlands: Kluwer Academic Publishers.
- TEMEZ, J.R. 1977. Modelo mensual precipitación-aportación. *Asinel*.
- THIESSEN, A.H. 1911. Precipitation averages for large areas. *Mon. wea. rev.*, **39**, 1082–1084.
- THOMAS, H.A. 1981. *Improved methods for national water assessment*. Report, contract wr 15249270. US Water Resources Council, Washington, DC, USA.
- THOMAS, H.A., MARIN, C.M., BROWN, M.J., & FIERING, M.B. 1983. *Methodology for water resources assessment*. Report ntis 84-124163. US Geological Survey, National Tech. Info. Serv., Springfield, Virginia, USA.
- VELICER, W.F. 1976. Determining the number of components from the matrix of partial correlations. *Psychometrika*, **41**, 321–327.
- VRUGT, J.A., GUPTA, H.V., BOUTEN, W., & SOROOSHIAN, S. 2003. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters. *Water resources research*, **39**(8), 1201, doi:10.1029/2002WR001642.
- WAYMIRE, R., & GUPTA, V.J. 1981. The mathematical structure of rainfall representations. 1. a review of the stochastic rainfall models. *Water resources research*, **17**(5), 1261–1272.
- WHITE, D.H., & O'MEAGHER, B. 1995. Coping with exceptional droughts in australia. *Drought network news*, **7**(2), 13–17.
- WILHITE, D.A., & GLANTZ, M.H. 1985. Understanding the drought phenomenon: The role of definitions. *Water international*, **10**(3), 111–120.
- WILLEKE, G., HOSKING, J.R.M., WALLIS, J.R., & GUTTMAN, N.B. 1994. *The national drought atlas*. Report 94-NDS-4. Institute for Water Resources, U.S. Army Corps of Engineers.
- WMO. 1986. *Report on drought and countries affected by drought during 1974–1985*. Tech. rept. World Meteorological Organization, Geneva.

- WOODHOUSE, C.A. 2004. *Tree ring colorado streamflow reconstructions*. IGBP Pages / World Data Center for Paleoclimatology Data Contribution Series 2004-029.NOAA / NGDC Paleoclimatology Program.
- YEVJEVICH, V. 1963. *Fluctuation of wet and dry years, part i, research data assembly and mathematical models*. Hydrology paper 1. Colorado State University, Fort. Collins, CO.
- YEVJEVICH, V. 1967. *An objective approach to definitions and investigations of continental hydrologic droughts*. Hydrology paper 23. Colorado State University, Fort. Collins, CO.
- YEVJEVICH, V., DA CUNHA, L., & VLACHOS, E. 1983. *Coping with droughts*. Water Resources Publications. Eds.
- Z Aidman, M., & REES, G. 2000. *Spatial patterns of streamflow drought in western europe 1960-1995*. ARIDE. Technical Report 8. Institute of Hydrology, Freiburg, Germany.
- Zelenhasic, E., & Salvai, A. 1987. A method for streamflow drought analysis. *Water resources research*, **23**(1), 156–168.