



Universitat Politècnica de València

Departamento de Ingeniería Hidráulica y Medio Ambiente

Programa de Doctorado de Ingeniería del Agua y Medioambiental

**Inferencia Bayesiana conjunta de
modelos hidrológicos y modelos de
error generalizados, para la
evaluación de las incertidumbres
predictiva y de los parámetros**

AUTOR: MARIO RAMÓN HERNÁNDEZ LÓPEZ

DIRECTOR: DR. FÉLIX FRANCÉS GARCÍA

VALENCIA, SEPTIEMBRE 2017

Por y para mis *papás* Mario y María del Carmen (*I.G.D.*)

A Noelia

A mis pequeñas *posteriors* Sarah y Rebeca

Tu problema puede ser modesto, pero si desafía tu curiosidad,
te lleva a jugar con tus capacidades inventivas
y lo resuelves por tus propios medios,
vas a experimentar la tensión y a disfrutar del triunfo del descubrimiento
George Pólya, matemático (1887-1985)

El científico es libre y debe ser libre para hacer cualquier pregunta,
para dudar de cualquier aseveración, para buscar cualquier evidencia.
Para corregir cualquier error
J. Robert Oppenheimer, físico teórico (1904-1967)

AGRADECIMIENTOS

A la Universitat Politècnica de València la cual, mediante su Programa de Apoyo a la Investigación y el Desarrollo (PAID), para la formación de personal investigador, ha hecho posible que pudiese continuar y concluir mis estudios de doctorado.

Al director de esta Tesis, Félix Francés García, por haberme permitido hacer realidad lo que siempre he considerado un sueño.

A mis compañer@s del Grupo de Investigación de Modelación Hidrológica y Ambiental (GIMHA), especialmente a l@s que han contribuido a que disfrute tanto de este ratito.

INDICE DE CONTENIDOS

Agradecimientos	7
Lista de figuras	13
Lista de tablas	19
Resumen	21
Summary	25
Resum	29
Capítulo 1. Introducción	33
1.1. Introducción General.....	35
1.2. Objetivos principales de la Tesis	38
1.2.1. Motivación	38
1.2.2. Objetivos principales	41
Capítulo 2. Estado de la Cuestión	43
2.1. Modelación ambiental	45
2.1.1. Tipos de modelos	46
2.1.2. Complejidad de los modelos.....	48
2.1.3. Identificación de modelos: equifinalidad y sensibilidad	50
2.1.4. Incertidumbre de los parámetros inferidos	52
2.1.5. Revisión de métodos de inferencia de los parámetros.....	57
2.2. Inferencia Bayesiana.....	62
2.2.1. Estimación Bayesiana de los parámetros y su incertidumbre	65
2.2.2. Métodos de muestreo de la posterior.....	69
2.2.3. Incertidumbre estructural mediante la aproximación Bayesiana	80
2.3. Formulación del problema de inferencia en modelación hidrológica.....	82
2.3.1. Inferencia en el caso ideal. Método de los mínimos cuadrados	83

2.3.2. Inferencia en los casos no ideales	87
2.3.3. Relajación de la hipótesis de homocedasticidad de los errores	90
2.3.4. Relajación de la hipótesis de independencia de los errores.....	95
2.3.5. Función de verosimilitud generalizada (GL)	97
2.3.6. Corrección de Evin et al. a la función GL	101
2.4. Incertidumbre de la predicción	104
2.4.1. Evaluación de la incertidumbre total de la predicción.....	104
2.4.2. Evaluación de la capacidad predictiva	106
Capítulo 3. El modelo de error	111
3.1. Motivación.....	113
3.1.1. Primeros experimentos.....	114
3.1.2. Inspiración	116
3.2. Las Leyes Totales.....	122
3.2.1. Derivación	122
3.2.2. Casuística en la aplicación de las Leyes Totales	124
3.3. Un nuevo modelo de error.....	126
3.3.1. Definición del problema de inferencia	126
3.3.2. Derivación de la función de verosimilitud	133
3.4. Inclusión de las Leyes Totales en el modelo de error.....	136
3.4.1. Inclusión de las TLs al relajar la hipótesis de media nula.....	137
3.4.2. Inclusión de las TLs al relajar la hipótesis de homocedasticidad. Modelo de varianza del error.....	138
3.4.3. Inclusión de las TLs al relajar simultáneamente las hipótesis de media nula y homocedasticidad	140
3.4.4. Sobre el método WLS.....	142
3.5. Estimación de la incertidumbre predictiva a partir de los modelos hidrológico y de error conjuntamente inferidos.....	144

Capítulo 4. Descripción de Experimentos	147
4.1. Motivación.....	149
4.2. Modelos de error.....	152
4.3. Modelación hidrológica agregada.....	155
4.3.1. Descripción de los modelos considerados.....	155
4.3.2. Cuenca de estudio.....	157
4.4. Modelación hidrológica distribuida	159
4.4.1. Descripción del modelo considerado	159
4.4.2. Descripción de la cuenca.....	160
Capítulo 5. Consecuencias de la (no) aplicación de las TLs en la inferencia conjunta	163
5.1. SLS.....	165
5.1.1. Validez del modelo de error: cumplimiento de sus hipótesis.....	166
5.1.2. Identificabilidad e incertidumbre de los parámetros	168
5.1.3. Predicción y estimación de su incertidumbre	171
5.1.4. SLS. Conclusiones	172
5.2. WLS y WLS-NTL.....	174
5.2.1. La paradoja en los resultados de WLS-NTL	175
5.2.2. Validez del modelo de error WLS: cumplimiento de sus hipótesis ..	178
5.2.3. Identificabilidad e incertidumbre de los parámetros	179
5.2.4. Predicción y estimación de su incertidumbre	180
5.2.5. WLS y WLS-NTL. Conclusiones	180
5.3. GL++ y GL++NTL	181
5.3.1. Validez del modelo de error: cumplimiento de sus hipótesis.....	181
5.3.2. Efectos de la (no) aplicación de las Leyes Totales	185
5.3.3. GL++ y GL++NTL. Conclusiones.....	194

Capítulo 6. Aplicación del nuevo modelo de error completo: GL++Bias 197

- 6.1. Aplicación en modelación agregada..... 199
 - 6.1.1. Validez del modelo de error: cumplimiento de sus hipótesis 200
 - 6.1.2. Identificabilidad e incertidumbre de los parámetros con GR4J..... 202
 - 6.1.3. Predicción y estimación de su incertidumbre 203
 - 6.1.4. Sobre la estimación y el sesgo de los parámetros 207
 - 6.1.5. Comparación de los experimentos con modelos agregados 211
- 6.2. Aplicación en modelación distribuida 215
 - 6.2.1. Validez del modelo de error: cumplimiento de sus hipótesis 216
 - 6.2.2. Identificabilidad e incertidumbre de los parámetros 218
 - 6.2.3. Predicción y estimación de su incertidumbre 220
 - 6.2.4. Influencia del modelo de error empleado en la inferencia sobre los procesos simulados por el modelo 225

Capítulo 7. Conclusiones 227

- 7.1. Epílogo..... 229
- 7.2. Conclusiones 232
- 7.3. Líneas futuras 236

8. Referencias..... 239

LISTA DE FIGURAS

*Figura 1.1. Hidrología e incertidumbre. Evolución anual del número de publicaciones que relacionan ambos términos (panel superior). Evolución anual del número de citas sobre publicaciones que relacionan ambos términos (panel inferior). Fuente: Web Of Science.*37

Figura 2.1 Relaciones entre sensibilidad e incertidumbre52

Figura 2.2 Esquema clásico de calibración automática. Fuente: Vrugt et al. (2008)60

Figura 2.3. Izquierda: Superficie con posibles pdf's (muestrales) de la variable de interés, asumiendo una prior plana del parámetro. Derecha: Superficie de probabilidad conjunta de los parámetros (un solo parámetro θ en este caso) y la variable de interés. En el plano lateral se representa la prior del parámetro. Fuente: Bolstad (2010).....65

Figura 2.4 Izquierda: Función de verosimilitud del parámetro, dada la observación de la variable de interés. Derecha: La pdf posterior del parámetro dada la observación de la variable de interés. La función de verosimilitud y la posterior son el resultado de cortar con el mismo hiperplano, determinado por las observaciones, dos superficies diferentes. Fuente: Bolstad (2010).....67

Figura 2.5 Evolución de dos cadenas diferentes, durante el muestreo MCMC de un parámetro. Izquierda: buen mezclado de la cadena representativo de haber alcanzado el estado de equilibrio. Derecha: mal mezclado de la cadena76

Figura 2.6. Test de convergencia MCMC, mediante la inspección de las trazas de tres parámetros, con sus correspondientes densidades posterior muestreadas.....77

Figura 2.7. Densidades de la pdf SEP estándar (media=0 y std.dev=1), para varios valores de sus dos parámetros, curtosis (β) y asimetría (ξ). Fuente: Schoups y Vrugt (2010)99

Figura 2.8. Distribuciones predictivas sobre 3 cuencas. En los paneles de la izquierda, obtenidas mediante la aproximación de Schoups y Vrugt (2010). En los paneles de la derecha obtenidas con la corrección de Evin et al. (2013). Fuente: Evin et al. (2013)103

Figura 2.9. Diagramas interpretativos de los PP-plots. Se representan diferentes combinaciones de las propiedades de exactitud (bias) y precisión (incertidumbre) de la distribución predictiva108

Figura 3.1. Distribución muestral conjunta de $P(\mathbf{Y}_M, \mathbf{Y})$, en la cual se representan la marginal (en rojo) $P(\mathbf{Y})$ y una de las distribuciones condicionales $P(\mathbf{Y}|\mathbf{Y}_M)$, condicionada, en este caso, al valor $y_{jM} = y_M^*$117

Figura 3.2. Incoherencia estadística debida al desacople entre las distribuciones condicionales inferidas (en azul) y la distribución marginal (en verde) de la nube errores inferida $\{y_{jM}, e_j\}$ 119

Figura 3.3. Explicación sobre la tendencia al aumento de las varianzas condicionales de los errores, durante una inferencia conjunta121

Figura 3.4. Esquemas de descomposición de la distribución conjunta $P(\mathbf{Y}_M, \mathbf{Y})$ en sus distribuciones condicionales (azul) y marginal (verde), bajo la asunción de 4 hipótesis diferentes, para los modelos de varianza y sesgo de los errores.....125

Figura 5.1. Comprobación de las hipótesis del modelo de error SLS, con los modelos hidrológicos CRR (columna izquierda) y GR4J (columna derecha). Normalidad de los errores estándar observados (paneles superiores). Evaluación de la homocedasticidad (paneles intermedios). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (paneles inferiores).....167

Figura 5.2. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada con SLS para el modelo hidrológico CRR (panel superior) y GR4J (panel inferior). Como se puede apreciar la contribución de los parámetros es, en este caso, despreciable171

Figura 5.3. PP-plots de todas las distribuciones predictivas inferidas con los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior)173

Figura 5.4. Comprobación de las hipótesis del modelo de error GR4J, con el modelo hidrológico GR4J. Evaluación de la homocedasticidad (panel superior). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (panel inferior).....179

Figura 5.5. Comprobación de las hipótesis del modelo de error GL++, con los modelos hidrológicos CRR (columna izquierda) y GR4J (columna derecha). Ajuste a SEP de los errores estándar observados (paneles superiores). Evaluación de la homocedasticidad (paneles intermedios). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (paneles inferiores).....182

Figura 5.6. Elementos de los modelos de error GL++ y GL++NTL, para los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior). Las gráficas muestran las nubes de error en rojo para GL++ (aplica TVL) y en gris para GL++NTL. Superpuestas aparecen las líneas inferidas que representan a los modelos de varianza, para GL++ en negro, y para GL++NTL en gris.....186

Figura 5.7. Modelo hidrológico CRR. Distribución posterior del parámetro de linealidad del modelo de varianza, distribución posterior del parámetro de autocorrelación y nube de valores muestrales entre ambos parámetros. Se representa el caso con el modelo de error GL++ (panel izquierdo) y el caso con el modelo de error GL++NTL (panel derecho)188

Figura 5.8. Matriz de correlación para los parámetros del modelo GR4J inferidos con el modelo de error GL++189

Figura 5.9. Matriz de correlación para los parámetros del modelo GR4J inferidos con el modelo de error GL++NTL.....190

Figura 5.10. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada para el modelo hidrológico CRR, con los modelos de error GL++ (panel superior) y GL++NTL (panel inferior). Como se puede apreciar la contribución de los parámetros es mínima, aunque superior a la obtenida con SLS y WLS192

Figura 5.11. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada para el modelo hidrológico GR4J, con los modelos de error GL++ (panel superior) y GL++NTL (panel inferior). Como se puede apreciar, la contribución de los parámetros es mínima este caso.....193

Figura 6.1. Elementos del modelo de error GL++Bias, para los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior). Las gráficas muestran las nubes de error en rojo. Superpuestas aparecen las líneas inferidas que representan a los modelos de varianza (línea negra) y sesgo (línea azul discontinua). Las ventanas pequeñas muestran el comportamiento del modelo de sesgo para los caudales bajos.....200

Figura 6.2 Comprobación de las hipótesis del modelo de error GL++Bias, con los modelos hidrológicos CRR (columna izquierda) y GR4J (columna derecha). Ajuste a SEP de los errores estándar observados (paneles superiores). Evaluación de la homocedasticidad (paneles intermedios). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (paneles inferiores).....201

Figura 6.3. Matriz de correlación para los parámetros del modelo GR4J inferidos con el modelo de error GL++Bias203

Figura 6.4. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada con el modelo de error GL++Bias, para los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior). Como se puede apreciar, la contribución de los parámetros a la incertidumbre total es mínima para el modelo CRR, sin embargo con GR4J, es la mayor contribución observada en todos los experimentos205

Figura 6.5. Cambios en la distribución posterior (box-plots) de los 4 parámetros del modelo GR4J, según el modelo de error utilizado en la inferencia. Se ilustran los experimentos con SLS, WLS, GL++ y GL++Bias208

Figura 6.6. Diagrama Fiabilidad-Resolución para la comparación de las distribuciones predictivas inferidas con los diferentes modelos de error. Para el modelo hidrológico CRR se utilizan trayectorias con línea negra continua. Para el modelo GR4J se utilizan líneas azules discontinuas. Las trayectorias a los modelos de error que no contemplan Leyes Totales se indican con líneas rojas, continuas para CRR y discontinuas para GR4J212

Figura 6.7. Elementos de los modelos de error GL++ (panel izquierdo) y GL++Bias (panel derecho). Las gráficas muestran las nubes de error en rojo. Superpuestas aparecen las líneas inferidas que representan a los modelos de varianza (línea negra) y sesgo (línea azul discontinua). La ventana pequeña en el panel derecho muestra el comportamiento del modelo de sesgo para caudales bajos216

Figura 6.8. Ajuste de las innovaciones estandarizadas (puntos rojos) a la distribución SEP teórica (línea azul). Periodo de calibración en paneles superiores, Periodo de validación en paneles inferiores. Inferencia con GL++ en columna izquierda. Inferencia con GL++Bias en columna derecha217

Figura 6.9. Evaluación de la estabilización de la varianza mediante la comparación de la distribución de los errores (puntos rojos) y las innovaciones (puntos negros), en función de la magnitud del caudal simulado durante el periodo de validación. Las líneas representan la varianza móvil (60 datos) de los errores (línea roja) y de las innovaciones (línea gris). Inferencia con GL++ (panel superior) e inferencia con GL++Bias (panel inferior)218

Figura 6.10. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Superpuestas a la banda de incertidumbre, se muestran las

observaciones (puntos rojos), el resultado de la simulación del modelo (línea azul) y la predicción media (línea verde). Los paneles superiores corresponden a la inferencia con GL++. Los paneles inferiores corresponden a la inferencia con GL++Bias. La columna derecha muestra el detalle para un tramo de caudales bajos.222

Figura 6.11. Fiabilidad de las distribuciones predictivas inferidas con GL++ (líneas grises) y con GL++Bias (líneas azules). Para el periodo de calibración las líneas son discontinuas, y paa el periodo de validación son continuas.....223

Figura 6.12. Banda (95%) de incertidumbre total (gris). Superpuestas a la banda de incertidumbre, se muestran las observaciones (puntos rojos), el resultado de la simulación del modelo (línea azul) y la predicción media (línea verde). Los paneles superiores corresponden a la validación temporal del modelo inferido con GL++. Los paneles inferiores corresponden a la validación temporal del modelo inferido con GL++Bias.224

Figura 6.13. Contribución (%) de los tres componentes de la producción de escorrentía a la respuesta total del sistema, en función de la magnitud de dicha respuesta durante el periodo de calibración. La figura muestra los resultados para la inferencia con GL++ (panel izquierdo) y para la inferencia con GL++Bias (panel derecho). Los componentes mostrados son el flujo base (línea negra) el interflujo (línea verde) y escorrentía superficial (línea azul). La línea roja vertical en el panel derecho indica la posición del umbral de caudal inferido para el modelo de sesgo.....226

LISTA DE TABLAS

<i>Tabla 4.1. Parámetros del modelo TETIS y del modelo de error. Descripción y rangos...</i>	161
<i>Tabla 5.1. Índices de desempeño sobre diferentes aspectos de los resultados, con los principales modelos de error de los experimentos, y para ambos modelos hidrológicos. La primera fila muestra la Log-verosimilitud de los parámetros inferidos. Los índices de las tres siguientes filas, se relacionan con la bondad de la predicción media obtenida. Los índices de las cuatro últimas filas evalúan los modelos de sesgo y varianza inferidos, así como la exactitud y precisión de la distribución predictiva respecto de las observaciones. En verde se marcan las mejores inferencias. En negrita se marcan los mejores valores</i>	166
<i>Tabla 5.2. Modelo hidrológico CRR: Máximo de la posterior (MAP) y coeficientes de variación (CV) de los parámetros. CVs mayores o iguales a 0.2 se marcan en rojo. Parámetros restringidos por las TLs, se somborean en azul.....</i>	169
<i>Tabla 5.3. Modelo hidrológico GR4J. Máximo de la posterior (MAP) y coeficientes de variación (CV) de los parámetros. CVs mayores o iguales a 0.2 se marcan en rojo. Parámetros restringidos por las TLs, se somborean en azul.....</i>	170
<i>Tabla 6.1. Máximo de la posterior (MAP) y coeficientes de variación (CV) de los parámetros. CVs mayores o iguales a 0.2 se marcan en rojo. Parámetros restringidos por las TLs, se somborean en azul.....</i>	219
<i>Tabla 6.2. Índices de desempeño sobre diferentes aspectos de los resultados en los periodos de calibración y validación, de las inferencias del modelo distribuido TETIS, con los modelos de error GL++ y GL++Bias. La primera fila muestra la Log-verosimilitud de los parámetros inferidos. Los índices de las tres siguientes filas, se relacionan con la bondad de la predicción media obtenida. Los índices de las dos últimas filas evalúan la exactitud y precisión de la distribución predictiva respecto de las observaciones.....</i>	221

RESUMEN

A lo largo de los años, el método de los mínimos cuadrados (SLS) ha sido el método de inferencia comúnmente aplicado en modelación hidrológica, a pesar de que sus hipótesis no son respetadas por los errores en los resultados de la modelación. La concienciación sobre el hecho de que el proceso de modelación hidrológica es afectado por más, y más importantes, fuentes de incertidumbre que la puramente observacional, única fuente de error considerada por el SLS, ha contribuido a la aparición de publicaciones que sugieren la necesidad de aplicar métodos de inferencia más adecuados para los modelos hidrológicos, y en general, los modelos ambientales. La adecuación de los métodos de inferencia pasa por considerar todas las fuentes de error, o sus efectos, que influyen en el proceso de modelación. Solamente de esa manera es posible la obtención de unos parámetros fiables, una predicción no sesgada y una estimación correcta de la incertidumbre de ambos, siendo estos los objetivos principales de esta Tesis Doctoral. Para ello, esta Tesis plantea la inferencia conjunta, siguiendo el paradigma de los métodos de inferencia Bayesianos, de los parámetros hidrológicos y los parámetros de un modelo de error generalizado, el cual proporciona la flexibilidad necesaria para relajar todas las hipótesis (errores Gaussianos con media nula, independientes e idénticamente distribuidos), que inhabilitan al modelo de error SLS para inferir modelos hidrológicos. La principal aportación de la Tesis es la proposición de la metodología a seguir para la correcta aplicación de la modelación directa (sin previa transformación de las variables) de la varianza de los errores. Dicha metodología se fundamenta en la necesidad de considerar el acoplamiento, durante la inferencia conjunta, entre las variaciones de la distribución marginal de los errores y las variaciones de las distribuciones condicionales, las cuales son modeladas por el modelo de error. Dicho acoplamiento se garantiza mediante el cumplimiento de las Leyes Totales (TLs) de la esperanza y de la varianza. Para la comprobación de la viabilidad de los aspectos teóricos deducidos en la Tesis, se realiza una serie de experimentos de inferencia en los que se combina 2 modelos hidrológicos agregados y uno distribuido, con dos modelos de error clásicos (SLS y WLS) y dos modelos de error generalizados propuestos en esta Tesis (GL++ y GL++Bias). Los resultados muestran, una vez más, que las inferencias con SLS y WLS no son aplicables a modelos hidrológicos, puesto que los errores generados no cumplen con sus hipótesis. Igualmente, en base a los resultados obtenidos, esta Tesis se puede considerar como la afirmación culminada de la hipótesis en ella planteada, esto es: la no aplicación de las TLs en la modelación directa de la varianza y el sesgo de los errores produce una incorrecta estimación de los parámetros hidrológicos, así como de su incertidumbre y una errónea estimación de la distribución predictiva.

SUMMARY

Over the years, the method of least squares (SLS) has been the method of inference commonly applied in hydrological modeling, although its hypotheses are not respected by the modeling errors. Awareness of the fact that the hydrological modeling process is affected by more, and more important, sources of uncertainty than the purely observational, the only source of error considered by SLS, has contributed to the appearance of publications that suggest the need for applying more appropriate inference methods on hydrological models, and in general, on environmental models. The adequacy of inference methods involves considering all sources of error, or their effects, which influence the modeling process. Only in this way it is possible to obtain reliable parameters, a non-biased prediction and a correct estimation of the uncertainty of both, these being the main objectives of this Doctoral Thesis. To this end, this thesis proposes the joint inference, following the Bayesian inference paradigm, of hydrological parameters and the parameters of a generalized error model, which provides the necessary flexibility to relax all hypotheses (Gaussian errors with null mean, independent and identically distributed), which disable the SLS error model to infer hydrological models. The main contribution of the thesis is the proposition of the methodology to follow, for the correct application of the direct modeling (without previous transformation of the variables) of the variance of the errors. This methodology is based on the need to consider the coupling, during the joint inference, between the variations of the marginal distribution of the errors and the variations of their conditional distributions, which are modeled by the error model. Such coupling is guaranteed by the fulfillment of the Total Laws (TLs) of the expectation and the variance. In order to verify the feasibility of the theoretical aspects deduced in the Thesis, a series of inference experiments is performed in which two lumped and one distributed hydrological models are combined with two classic error models (SLS and WLS) and two generalized error models proposed in this Thesis (GL++ and GL++Bias). The results show, once again, that inferences with SLS and WLS are not applicable to hydrological models, since the generated errors do not fulfill their hypotheses. Likewise, based on the results obtained, this thesis can be considered as the culminated affirmation of the hypothesis in it, that is to say: the non application of the TLs in the direct modeling of the variance and the bias of the errors produces an incorrect estimation of the hydrological parameters, as well as their uncertainty and an erroneous estimation of the predictive distribution.

RESUM

Al llarg dels anys, el mètode dels mínims quadrats (SLS) ha sigut el mètode d'inferència comunament aplicat en modelació hidrològica, a pesar que les seues hipòtesis no són respectades pels errors en els resultats de la modelació. La conscienciació sobre el fet de que el procés de modelació hidrològica és afectat per més, i més importants, fonts d'incertesa que la purament observacional, única font d'error considerada pel SLS, ha contribuït a l'aparició de publicacions que suggerixen la necessitat d'aplicar mètodes d'inferència més adequats per als models hidrològics, i en general, els models ambientals. L'adequació dels mètodes d'inferència passa per considerar totes les fonts d'error, o els seus efectes, que influïxen en el procés de modelació. Només d'eixa manera és possible l'obtenció d'uns paràmetres fiables, una predicció no esbiaixada i una estimació correcta de la incertesa d'ambdós, sent estos els objectius principals d'esta Tesi Doctoral. Per a això, esta Tesi planteja la inferència conjunta, seguint el paradigma dels mètodes d'inferència Bayesianos, dels paràmetres hidrològics i els paràmetres d'un model d'error generalitzat, el qual proporciona la flexibilitat necessària per a relaxar totes les hipòtesis (errors Gaussianos amb mitja nul·la, independents i idènticament distribuïts), que inhabiliten el model d'error SLS per a inferir models hidrològics. La principal aportació de la Tesi és la proposició de la metodologia que s'ha de seguir per a la correcta aplicació de la modelació directa (sense prèvia transformació de les variables) de la varianza dels errors. La dita metodologia es fonamenta en la necessitat de considerar l'adaptament, durant la inferència conjunta, entre les variacions de la distribució marginal dels errors i les variacions de les distribucions condicional, les quals són modelades pel model d'error. El dit adaptament es garanteix per mitjà del compliment de les Lleis Totals (TLs) de l'esperança i de la variància. Per a la comprovació de la viabilitat dels aspectes teòrics deduïts en la Tesi, es realitza una sèrie d'experiments d'inferència en què es combina 2 models hidrològics agregats i un distribuït, amb dos models d'error clàssics (SLS i WLS) i dos models d'error generalitzats proposats en esta Tesi (GL++ i GL++Bias). Els resultats mostren, una vegada més, que les inferències amb SLS i WLS no són aplicables a models hidrològics, ja que els errors generats no complixen amb les seues hipòtesis. Igualment, basant-se en els resultats obtinguts, esta Tesi es pot considerar com l'afirmació culminada de la hipòtesi en ella plantejada, açò és: la no aplicació de les TLs en la modelació directa de la variància i el biaix dels errors produïx una incorrecta estimació dels paràmetres hidrològics, així com de la seua incertesa i una errònia estimació de la distribució predictiva.

CAPÍTULO 1. INTRODUCCIÓN

Apparently, what makes the system alive
is the same agent that creates the uncertainty

Only dead systems are certain

-Demetris Koutsoyiannis (1955)-

1.1. Introducción General

Incertidumbre

La Directiva Marco del Agua (DMA, 2000/60/CE) tiene entre sus principios básicos, velar por la protección y preservación de los sistemas acuáticos, en términos de cantidad y calidad y promover el uso racional de los recursos naturales basándose en el principio de precaución y prevención. En la medida en que el principio de precaución tiene como objetivo proteger a los humanos y al medio ambiente contra los riesgos inciertos, por medio de medidas preventivas, dicha tarea no puede llevarse a cabo sin la incorporación de la evaluación de la incertidumbre en el proceso de toma de decisiones. La evaluación de la incertidumbre de los modelos de simulación es, por lo tanto, importante cuando los modelos se utilizan para apoyar las decisiones sobre la gestión del agua.

Incertidumbre es un término usado en formas sutilmente diferentes en una serie de campos tan distintos como la filosofía, la estadística, la economía, la ingeniería y la ciencia. Se aplica a predicciones de eventos futuros, a medidas físicas ya realizadas, o a lo desconocido. No hay una definición universal aplicable al concepto de incertidumbre. La Real Academia Española (RAE) define el término, como *falta de certidumbre o certeza*. Atendiendo a la definición que la RAE otorga al término *certeza*, el concepto de incertidumbre se puede definir de las siguientes maneras: 1. *Falta de conocimiento seguro y claro de algo*. 2. *Adhesión insegura de la mente a algo conocible, con temor de errar*. Ya en el ámbito científico, Refsgaard et al. (2007) aportan una acepción generalista, que define a la incertidumbre como una propiedad de la información empleada para tomar decisiones, o como una manifestación de la confianza que se tiene en dicha información. En el contexto de la Hidrología, la incertidumbre es entendida como una indicación cuantitativa de la fiabilidad de una determinada magnitud hidrológica, ya sea observada o inferida (Montanari, 2011). Si bien son

diversas las taxonomías que sobre la incertidumbre se encuentran en la literatura (Matott et al., 2009), la incertidumbre puede ser clasificada, de forma general, en dos categorías principales: Aleatoria y Epistémica.

La Incertidumbre es inevitable en modelación hidrológica (ambiental)

La evaluación cuantitativa de la incertidumbre en Hidrología, sobre bases teóricas, es considerada un tema relativamente nuevo (Montanari, 2011), si bien los hidrólogos han sido conscientes de su existencia desde los primeros estudios y aplicaciones hidrológicas, en los que era común el empleo de técnicas empíricas (e.g. factores de seguridad) para compensar la falta de fiabilidad de los modelos. Desde hace más de tres décadas (Beck, 1987) se han discutido en la literatura multitud de temas relacionados con la incertidumbre inherente a la aplicación de modelos ambientales. Sin embargo, en la actualidad y por diversas razones (Pappenberger y Beven, 2006), la estimación de la incertidumbre en modelación hidrológica todavía no es una práctica común (Beven, 2009), si bien en los últimos años las publicaciones de investigación sobre el tema están aumentando considerablemente (Figura 1.1).

En el pasado cercano, la modelación hidrológica ha obviado (fuera del ámbito de la investigación) el tema de la incertidumbre, centrándose en la forma de obtener el mejor conjunto de parámetros del modelo, que mejor cumpla con las necesidades y objetivos de la modelación. No obstante, el hecho de trabajar con modelos cada vez más complejos, que requieren mayores y más diversas cantidades de información, ha sacado a la luz la evidencia de que ni modelos ni información están libres de incertidumbre, la cual debe ser tenida en consideración. La paradigmática paradoja de la incertidumbre en modelación hidrológica viene precisamente de la mano de los modelos deterministas físicamente basados. El espíritu de estos modelos es hacer una descripción puramente determinista de los sistemas hidrológicos, mediante un mejor entendimiento de las dinámicas internas que rigen sus procesos. Sin embargo, son tantos los factores que afectan a la descripción determinista, que en la práctica, solo resulta posible abordar su tratamiento de forma probabilística (Koutsoyiannis, 2009). La incertidumbre es una **propiedad inherente** a los sistemas naturales, y por tanto es inevitable en modelación hidrológica. Por ello, los esquemas de modelación deben de adaptarse y explicitar su presencia (Montanari y Koutsoyiannis, 2012).

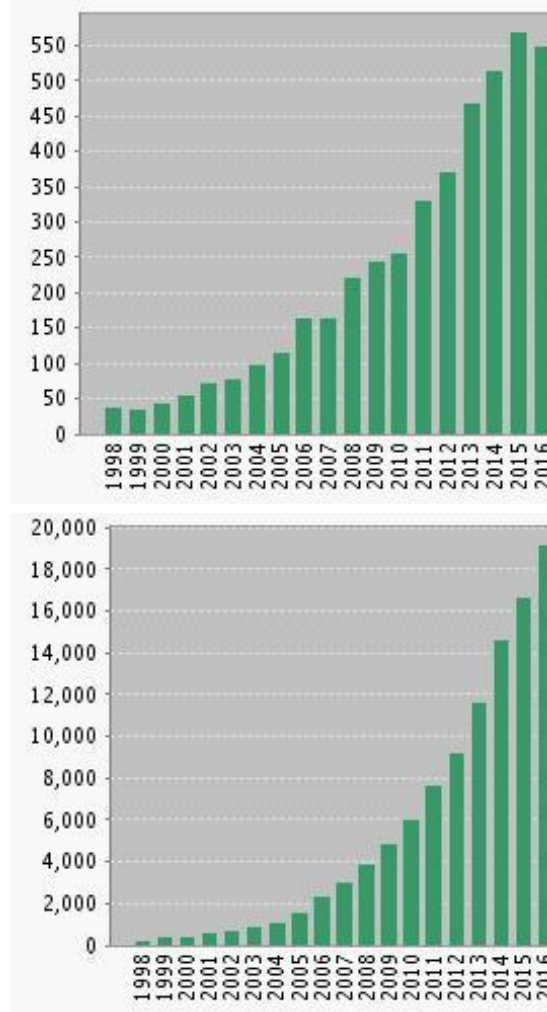


Figura 1.1. Hidrología e incertidumbre. Evolución anual del número de publicaciones que relacionan ambos términos (panel superior). Evolución anual del número de citas sobre publicaciones que relacionan ambos términos (panel inferior). Fuente: Web Of Science

Incetidumbre: causas y efectos en modelación hidrológica

¿Por qué hay incertidumbre en modelación? El proceso de modelación supone un camino que conduce de la realidad percibida, a la formulación teórica de las hipótesis que rigen los procesos observados de esa realidad y finalmente a la cuantificación de las magnitudes que los caracterizan. En todo ese camino desde la realidad a lo cuantitativo, pasando por lo cualitativo, se van perdiendo sutiles

piezas de información, sin las cuales el modelo pierde correspondencia con el sistema que trata de predecir. Las **fuentes de incertidumbre** son los lugares o momentos del camino, donde se pierden las piezas de información más importantes. Las principales fuentes de incertidumbre en modelación están en los datos suministrados al modelo y en la propia estructura del modelo. La incertidumbre de las fuentes acaba propagándose a los resultados del modelo, y en mayor o menor medida, a los parámetros inferidos durante el proceso de estimación de parámetros. Por tanto, los parámetros inferidos pueden llegar a ser un sumidero de incertidumbre, más que una fuente, contrariamente a lo que ocurre con los parámetros cuyo valor es asignado a priori, con su correspondiente incertidumbre.

La inferencia como factor determinante en la correcta evaluación de la incertidumbre

Adicionalmente a las fuentes de incertidumbre mencionadas, existe otra causa que afecta a la estimación de los parámetros, de las predicciones y de sus respectivas incertidumbres. Esta causa es el método de inferencia empleado. El método de inferencia añade más incertidumbre al proceso de modelación, produciendo estimaciones erróneas de valores (parámetros y predicciones) y de sus incertidumbres. Es conocido, que la incertidumbre se reduce conforme se añade más información no redundante o independiente al modelo. Sin embargo esa incertidumbre podrá estar (infra) sobreestimada, cuando el método de inferencia sea incorrectamente especificado (Mantovan y Todini, 2006).

El método de inferencia adoptado en esta Tesis tiene base probabilística. Concretamente, se van a seguir los procedimientos del **paradigma Bayesiano**, por sus propiedades y ventajas, frente a los clásicos métodos frecuentistas y de optimización. Entre esas ventajas, destaca el tratamiento de la incertidumbre de los parámetros de forma nativa, así como la posibilidad de considerar cualquier fuente de incertidumbre durante la inferencia.

1.2. Objetivos principales de la Tesis

1.2.1. Motivación

Desde hace años (Kavetski et al., 2003; Kuczera y Parent, 1998; Sorooshian y Dracup, 1980), se ha tomado conciencia en la literatura científica de que, en

modelación hidrológica, el método de inferencia es un importante factor de incertidumbre en la estimación de los parámetros, de la predicción y de sus respectivas incertidumbres.

Durante muchos años, los métodos de inferencia más comunes han estado basados en procedimientos de optimización de una función objetivo derivada o relacionada con la suma de los errores cuadráticos. Es decir, procedimientos basados en el clásico método de los mínimos cuadrados (SLS). La base teórica probabilística de dicho método se basa en la realización de ciertas hipótesis sobre las propiedades que tienen las desviaciones entre las observaciones y los resultados del modelo (los errores). Hipótesis, que la experiencia ha demostrado no cumplirse en el caso de la modelación hidrológica: Media nula, Normalidad, independencia y distribución constante. Ciertamente es que estas hipótesis, con las que se justifica el método SLS, se cumplen en vastos campos de la investigación. Como también es cierto, que solo y exclusivamente en el caso de cumplirse esas hipótesis, concurren dos objetivos: que el modelo reproduzca lo más fielmente posible las observaciones y al tiempo, se mantenga la coherencia teórica o estadística, atribuible al método de inferencia.

Este es el contexto en el cual se enmarca la presente Tesis Doctoral, en línea con las publicaciones que, en la literatura científica relacionada, sugieren la necesidad de realizar la inferencia de modelos hidrológicos con modelos de error más acordes a los errores producidos en este tipo de modelación. Esta necesidad viene motivada por el hecho de que, solamente cuando los errores son modelados adecuadamente por el modelo de error, es posible conseguir una estimación insesgada de los parámetros, una correcta predicción, así como una correcta estimación de la incertidumbre de ambos (Kavetski et al., 2003; Sorooshian y Dracup, 1980; Sorooshian y Gupta, 1983). Dicho de otra manera, el objetivo de la inferencia de los parámetros en modelación hidrológica no debería ser obtener el conjunto de parámetros óptimo que consiga el mejor ajuste entre observados y simulados, sino el conjunto de parámetros que proporcione los errores, que mejor cumplan con las hipótesis del modelo de error.

Al tener que inferir con modelos de error más adecuados, aumenta su complejidad, así como su número de parámetros, los cuales habrán de ser condicionados, al igual que los parámetros hidrológicos, a los datos que se les presente durante el proceso de inferencia. Como consecuencia, es aconsejable

que la inferencia de ambos conjuntos de parámetros, hidrológicos y de error, sea ejecutada mediante el procedimiento de **inferencia conjunta**, con el fin de no impedir las posibles interacciones entre parámetros y evitar así el sesgo, tanto en sus estimaciones, como en la predicción del modelo (Evin et al., 2014).

La presente Tesis Doctoral, al igual que los recientes trabajos de otros autores (Evin et al., 2013, 2014; Koskela et al., 2012; Scharnagl et al., 2015), se inspira fundamentalmente en el trabajo de Schoups y Vrugt (2010). En dicho trabajo se propone el uso de un modelo de error generalizado, que aporta una gran flexibilidad a la inferencia, al poder relajar las ya mencionadas hipótesis que limitan al método SLS. Así mismo, en dicho trabajo proponen la inferencia Bayesiana conjunta de los parámetros hidrológicos y de error.

El otro gran aporte del trabajo de Schoups y Vrugt (2010), consiste en la modelación de la varianza de los errores (modelación de la heterocedasticidad), de forma directa, esto es, sin la aplicación previa de un método de transformación (Sakia, 1992) sobre las variables del problema, con el fin de “Normalizar” los errores y estabilizar su varianza. Hasta donde se ha revisado la literatura en esta Tesis, no hay trabajo anterior a 2010, que haya intentado la modelación directa de la varianza de los errores. Todos los trabajos hallados, tratan de eliminar la heterocedasticidad mediante los citados métodos de transformación. Hay que remontarse hasta 1980, para encontrar un trabajo (Sorooshian y Dracup, 1980) en el que la modelación de la varianza se aborda mediante un método, que esta Tesis ha acuñado como *método híbrido*. El método consiste en modelar la varianza de los errores de forma directa, mediante un modelo de varianza cuyos parámetros, se estiman a partir de las variables previamente transformadas.

Sin embargo, en el trabajo de Schoups y Vrugt (2010), así como en los recientes trabajos que en él se inspiran, aparecen problemas con la inferencia de los parámetros, especialmente cuando esta inferencia se realiza de manera conjunta. Los problemas son diversos y a priori, no parecen seguir un patrón fijo que permita identificar su origen. Evin et al (2014) llegan a afirmar que posiblemente, no se pueda realizar la inferencia conjunta de los parámetros hidrológicos y del modelo de error, debido a las interacciones que entre ellos se producen.

1.2.2. Objetivos principales

Esta Tesis Doctoral continúa la línea de investigación sobre incertidumbre en el Grupo de Investigación GIMHA-IIAMA-UPV. La Tesis, en su origen, plantea dos objetivos principales de carácter hidrológico cuya consecución, sin embargo, solo es posible con el cumplimiento necesario de un tercer objetivo. Por tanto, los tres son objetivos principales de la Tesis y se describen a continuación.

El primer objetivo que se marca esta Tesis es la estimación de los parámetros hidrológicos, de la forma menos sesgada posible, así como la correcta evaluación de su incertidumbre mediante el paradigma de los métodos de inferencia Bayesianos. La consecución de este objetivo solo es posible mediante la inferencia conjunta con un modelo de error adecuado, cuyas hipótesis sean respetadas por los errores de modelación producidos en la inferencia. Para ello, se han escogido dos modelos hidrológicos agregados (GR4J y CRR) y un modelo hidrológico distribuido (TETIS), que permiten comprobar la viabilidad de los métodos Bayesianos con diferentes estructuras conceptuales (modelos) y con diferentes dimensiones del espacio de parámetros.

El segundo objetivo, es el objetivo fundamental de la modelación hidrológica: la obtención de una correcta predicción, junto con la evaluación de su incertidumbre total. Al igual que para la estimación de los parámetros y su incertidumbre, este objetivo se lleva a cabo con la aplicación de los tres modelos hidrológicos mencionados. Para conseguir este objetivo, es necesario haber realizado previamente la inferencia con el mejor modelo de error posible.

El tercer objetivo, si bien cronológicamente ha sido el primero, es el “diseño” del modelo de error que permita la consecución de los dos objetivos principales. Para ello, esta Tesis se ha basado en el modelo de error generalizado empleado por Schoups y Vrugt en 2010. Sin embargo, al igual que les ocurriese a estos autores, y a otros ya mencionados que siguieron su metodología, esta Tesis se encontró con problemas en la aplicación de la inferencia conjunta con el modelo de error generalizado. Problemas que ha habido que resolver, para poder seguir con la idea de la inferencia conjunta. Es por ello, que la búsqueda e investigación de la solución a los problemas sobrevenidos, se haya convertido finalmente en el objetivo, si no principal, sí más crítico de la Tesis, sin cuya resolución no se podrían haber realizado los restantes objetivos, al menos en la forma en la que se han llevado a cabo. Sirva como adelanto para el lector, la

idea de que el origen de los problemas con la inferencia conjunta del modelo de error generalizado estaba en la incorrecta aplicación del método directo de modelar la varianza de los errores. Como producto de esta investigación, se ha obtenido la metodología que ha permitido diseñar un nuevo modelo de error generalizado (capítulo 3), con el cual ya no aparecen los problemas en la inferencia conjunta de modelos de error generalizados, referenciados en la literatura reciente.

Seguidamente se da una breve descripción del contenido de los capítulos de esta Tesis. En el capítulo 2 se abordan de forma general los conceptos e ideas fundamentales que permiten ubicar al lector en la Cuestión tratada por la Tesis. El capítulo 3 presenta, en primer lugar, las claves que dan lugar a la formulación de la hipótesis principal defendida en la Tesis para, a continuación, desarrollar la teoría que permitirá probar dicha hipótesis. Los desarrollos teóricos realizados dan forma a un nuevo modelo de error generalizado. En el capítulo 4 se describen los materiales y métodos con los que se va a realizar los experimentos que, en primer lugar, demuestran la utilidad del nuevo modelo de error desarrollado y, en segundo lugar, permiten demostrar la hipótesis central de la Tesis. En los capítulos 5 y 6 se muestra y analiza en detalle, los resultados de los experimentos realizados. Las conclusiones, se van obteniendo de forma progresiva a lo largo de estos dos capítulos, si bien son finalmente recopiladas y relacionadas en el capítulo 7.

CAPÍTULO 2. ESTADO DE LA CUESTIÓN

We live on an island surrounded by a sea of ignorance

As our island of knowledge grows, so does the shore of our ignorance

-John Archibald Wheeler(1911-2008)-

2.1. Modelación ambiental

Los sistemas ambientales pueden ser concebidos como conjuntos de pequeños y grandes ecosistemas encajados unos dentro de otros en una jerarquía de escalas espaciales. Así, se constituye una compleja red de todo tipo de interacciones (físicas, químicas, biológicas) a varios niveles de escala espacial y temporal. Ello dificulta en varios aspectos (coste, tecnología), la posibilidad de medir directamente todos los procesos que sería necesario medir, para comprender completamente el funcionamiento del sistema.

En el estudio de cualquier sistema ambiental es posible, desde el punto de vista de su comprensión científica, percibir más detalles y mucha más complejidad de la que es posible representar de forma matemática, con el propósito de hacer predicciones cuantitativas sobre algún aspecto de dicho sistema (Beven, 2009). Por tanto, el paso del modelo ideal del sistema basado en nuestra percepción de toda su complejidad, el **modelo perceptible**, hacia la descripción matemática del mismo, el **modelo formal**, requiere inevitablemente la introducción de hipótesis simplificadoras, siendo esta, la parte más crítica y subjetiva del proceso de análisis.

Con el modelo formal (o matemático), el investigador persigue reproducir, lo más fielmente posible, el comportamiento del sistema analizado. Sin embargo, el modelo es una simplificación del sistema, lo que teóricamente podría no ser un problema, siempre que, en sus diferentes aplicaciones, no se olvide qué complejidades del sistema fueron ignoradas en la conceptualización, si es que esto pueda llegar a conocerse por completo. Adicionalmente, cabe considerar las simplificaciones producidas en el paso del modelo formal, al llamado **modelo procedimental**, compuesto por el conjunto de algoritmos y adaptaciones necesarias para convertir el modelo matemático en modelo computacional.

Los modelos computacionales son una herramienta excepcional, no solo para explicar lo que se puede observar del sistema. Además, y más importante,

permiten explicar lo que no se puede observar de forma directa, descubrir nuevo conocimiento y ampliar la comprensión del funcionamiento del sistema, pudiendo incluso predecir su comportamiento en el espacio y en el tiempo.

2.1.1. Tipos de modelos

Un vasto número de estructuras de modelos se han desarrollado e implementado computacionalmente, desde principios de los años 60 del siglo pasado (Devia et al., 2015; Rodda y Robinson, 2015; Todini, 1988, 2007). En general, todas ellas son una combinación de funciones lineales y no-lineales. Posiblemente, de entre todas las clasificaciones de modelos que se han hecho, la más común emplea tres clases en las que se engloba todos los tipos de modelos (Wheater et al., 1993). Estas son: i) Modelos métricos (conducidos por los datos medidos), también conocidos como **empíricos**, dirigidos por datos (*data-driven*) o de tipo “caja negra”; ii) Modelos paramétricos, también llamados **conceptuales** o de “caja gris” y iii) Modelos **mecanicistas**, físicamente basados o de “caja blanca”. No obstante, se dan casos de modelos que no podrían ser encajados exclusivamente, en una de las tres clases. Las principales características de cada clase se exponen brevemente en lo que sigue.

Los modelos empíricos (Elshorbagy et al., 2010a, 2010b; Remesan y Mathew, 2015) se basan exclusivamente en la información de la que se dispone, para definir su estructura y el valor de sus parámetros. No consideran ningún conocimiento a priori sobre el sistema que simulan, del cual tampoco encierran ninguna interpretación física. Usualmente son espacialmente agregados. Los más conocidos son las redes neuronales artificiales (Dibike y Solomatine, 2001; García-Bartual, 2002; Govindaraju y Rao, 2000; Shrestha et al., 2009)) y las funciones de transferencia (Young, 1998, 2002, 2011, 2013)

Los modelos conceptuales (Beven, 2012; Bowles y O’Connell, 1991; Frances et al., 2007) emplean elementos de almacenamiento (depósitos o tanques) para conformar la estructura del modelo, la cual es predefinida por el modelador, en base a su comprensión del sistema a modelar. Dichos depósitos son llenados y vaciados por los distintos flujos del sistema que se quiere representar (precipitación, infiltración, percolación, evapotranspiración real, escorrentía directa, flujo subsuperficial y flujo subterráneo). Las leyes (lineales y no-lineales) de llenado y vaciado de los depósitos tratan de reproducir la dinámica del sistema, cuyo modo de funcionamiento viene finalmente determinado por el valor

de sus parámetros. Es común, que muchos de estos parámetros agreguen (en espacio y/o en tiempo) varios de los procesos del sistema, por lo que no son parámetros directamente medibles en campo, es decir son magnitudes **inconmensurables**. Como explica Beven (2009), son parámetros que si bien tienen el mismo nombre que la propiedad física a la que representan, adoptan diferente valor. Por tanto, los valores de dichos parámetros han de ser determinados, al igual que en los modelos empíricos, en base a los datos observados de alguna(s) de las variables del sistema, mediante un proceso de estimación, inferencia, calibración o modelación inversa.

Los modelos mecanicistas (Beven, 1989; Freeze y Harlan, 1969) se fundamentan en ecuaciones de conservación de la masa, del momento y de la energía del sistema analizado. Su aplicación comenzó en los años 80 del siglo pasado, como consecuencia de la mejora en el poder de computación. Se esperaba que el grado de realismo físico con el que se construían dichos modelos, fuese suficiente para relacionar sus parámetros con las características físicas reales y medibles del sistema (Todini, 1988), evitando así el proceso de calibración de los mismos. Sin embargo, el resultado fue bien distinto. Estos modelos requieren gran cantidad de información para su funcionamiento, difícil de obtener en su mayor parte. Así mismo, son modelos que presentan una ingente cantidad de parámetros. El empleo de estos modelos hizo evidente la existencia de problemas de escala, entre el proceso que ocurre realmente, lo que se mide en campo y lo que el modelo calcula. En definitiva, muchos de los parámetros del modelo, no pudieron ser estimados directamente por medición en campo, debiendo de ser también inferidos.

Métodos de asignación de valores a los parámetros

Resumiendo, en cualquiera de los tipos de modelo descritos, se va a necesitar asignar valores a los parámetros, que no pueden ser medidos de forma directa. En la literatura, aparecen diferentes aproximaciones para la asignación de valores a los parámetros. En McInerney et al. (2017) se hace una clasificación de métodos, distinguiendo entre parámetros del modelo hidrológico y parámetros del modelo de error. Sin embargo, tal distinción no es necesaria, pudiendo generalizarse la clasificación que ellos proponen mediante la consideración de todos los parámetros, sin distinguir si pertenecen al modelo hidrológico o al de error. Siguiendo ese criterio, se distinguen principalmente dos formas de asignación de valores a los parámetros, con la que se obtienen parámetros de

valor fijo y parámetros inferidos. Los **parámetros de valor fijo**, reciben su valor en base a la experiencia previa, y al no formar parte del proceso de calibración, no interactúan con el resto de parámetros. Respecto a los parámetros inferidos, se pueden dar tres tipos de inferencia: en posproceso, iterativa y conjunta. La **inferencia en posproceso** consiste en separar los parámetros en dos bloques, llámense bloque principal y bloque condicionado (al principal). Se infieren los parámetros del bloque principal. A continuación, dejando ya fijos los valores del bloque principal, se infieren los parámetros del bloque de parámetros condicionado. En la **inferencia iterativa**, también se dividen los parámetros en dos bloques. En cada iteración se fija uno de los bloques de parámetros, mientras se infieren los parámetros del segundo bloque. En las diferentes iteraciones se van alternando el bloque fijo y el que se infiere, hasta que se alcanza el óptimo, es decir, hasta que los parámetros ya no cambian. Por último, la **inferencia conjunta** trata de inferir todos los parámetros al mismo tiempo, permitiendo sus posibles múltiples interacciones. Teóricamente, la inferencia conjunta consigue una evaluación completa (no sesgada) de la incertidumbre de los parámetros, así como de la incertidumbre de la predicción, puesto que incluye todas las posibles interacciones entre los parámetros (Evin et al., 2014). Los tres métodos de inferencia se han ordenado de menor a mayor corrección, si se quiere evitar la obtención de parámetros sesgados. Por otra parte, es posible que el lector pudiese pensar que la inferencia iterativa y la inferencia conjunta van a llegar a la misma solución óptima de parámetros. Quizá esto pueda ser cierto en la inferencia de problemas lineales y/o con espacios paramétricos de baja dimensión. Sin embargo, no va a ser necesariamente así en la inferencia de problemas con fuerte no linealidad, que además presentan un elevado número de parámetros.

En conclusión, el mayor o menor acierto en la asignación de valores a los parámetros, es la clave que nos proporcionará mayor o menor éxito en el empleo de cualquiera de los tipos de modelos como herramienta de predicción.

2.1.2. Complejidad de los modelos

Cualquier parametrización es una simplificación del modelo perceptible, realizada durante la conceptualización y diseño del modelo matemático. Visto desde el punto de vista del modelo matemático, su complejidad es directamente proporcional a su número de parámetros libres, los cuales le dan flexibilidad para

generar resultados que se ajusten a las observaciones. Al aumentar la flexibilidad del modelo, conforme se van añadiendo más parámetros, menores son los errores de la predicción, si bien la varianza de esta aumenta debido a que cada parámetro añade ruido al modelo, tanto más, cuanto menos soporte le dan los datos. La literatura refiere este hecho como el dilema entre el bias (o sesgo) y la varianza (incertidumbre) de la predicción (Geman et al., 1992). Desde antiguo, es conocida la preferencia natural por modelos simples y con pocos grados de libertad. Esta circunstancia queda recogida en el llamado **Principio de Parsimonia** (posiblemente sus sinónimos *parquedad* o *conciación*, serían más adecuados o mejor entendidos en lengua castellana), también conocido como **la navaja de Occam** por Williams Occam (1287-1374), filósofo y teólogo inglés quien estableció que “*las explicaciones nunca deben multiplicar las causas sin necesidad*”, lo que viene a significar que cuando dos o más explicaciones se ofrecen para un fenómeno, la explicación completa más simple es preferible. Las teorías (hipótesis) más simples requieren menos suposiciones y por tanto, para una cantidad limitada de información que las soporte tienen menos incertidumbre.

Con el aumento de la potencia de computación disponible, los modelos han podido ver incrementada su complejidad con la consiguiente adición de más y más parámetros (modelos sobreparametrizados). Pero ¿en qué medida el aumento en la complejidad en los modelos está realmente reflejando la complejidad del sistema real que representan? En la práctica, los modelos sobreparametrizados presentan suficientes grados de libertad para ser capaces, mediante la calibración de sus parámetros, de obtener un “excelente” **sobreajuste** a las observaciones. Generalmente, el aparente buen funcionamiento de un modelo sobreparametrizado, no se está produciendo por las razones correctas (Kirchner, 2006). Sus parámetros estimados no son robustos, perdiendo validez en condiciones (espacio-temporales) diferentes a las de calibración, lo que fue denominado por Sage y Melsa (1971) y Sorooshian y Dracup (1980), como fenómeno de **divergencia del modelo**. Este fenómeno consiste en un aumento significativo en la varianza de los errores de las predicciones, fuera del periodo de calibración y/o del punto de localización para el que los parámetros fueron estimados. En relación con la complejidad de los modelos, merece ser referenciada una curiosidad publicada por Mayer et al. (2010), quienes reproducen una anécdota en la que el Nobel Enrico Fermi (1901-1954) criticaba la excesiva complejidad del modelo de un colega,

mediante unas palabras originales de John von Neumann (1903-1957): “*con cuatro parámetros puedo ajustar un elefante y con cinco, puedo lograr que mueva su trompa*”.

2.1.3. Identificación de modelos: equifinalidad y sensibilidad

Según Brun et al. (2001), un modelo es *identificable* si es teóricamente posible conocer el valor de sus parámetros, después de obtener un número infinito de resultados del mismo, para unos datos dados. Dicho de otra forma, dado un modelo, sus parámetros serán identificables si pueden ser estimados de forma única a partir de los valores observados y simulados de una o más variables del modelo.

Los métodos clásicos de calibración de modelos, basados en la optimización de una función objetivo (también llamada función de pérdida o de coste), tienen en la falta de identificabilidad de los parámetros, a uno de sus principales problemas. Esta falta de identificación se manifiesta de dos formas posibles: i) Mediante la ausencia de un claro valor óptimo del parámetro sobre la superficie de respuesta de la función objetivo y ii) Mediante la existencia de múltiples valores óptimos del parámetro. Ambos efectos, talón de Aquiles de los algoritmos de optimización, son el origen de lo que se ha dado en llamar **equifinalidad** de los parámetros (Beven, 2001, 2002, 2006, 2000; Beven y Binley, 1992), concepto que expresa el hecho de que diferentes conjuntos de parámetros (o como generalización, diferentes estructuras del modelo) proporcionan similares valores de la función objetivo.

El problema surge principalmente en modelos sobreparametrizados, siendo los modelos físicamente basados el mayor exponente de esta problemática, la cual ha sido ampliamente tratada en la literatura. Así, Beck (1987) asegura que la principal razón de la pobre identificabilidad de los parámetros (y por tanto de los modelos) reside en que “*Lo que nos gustaría conocer sobre la descripción interna del sistema [...] es de un orden muy superior a lo que se puede observar sobre la descripción externa del mismo*”. O lo que es equivalente, el número de dimensiones de flexibilidad generadoras del espacio de parámetros, excede en mucho la capacidad de los datos para constreñirlas (Kirchner, 2006). Según Beven (2006), la **sobreparametrización** es un riesgo endémico de los modelos ambientales.

Yapo et al. (1996) enuncian las posibles razones por las que los parámetros del modelo pueden ser difícilmente identificables: i) La presencia de colinealidad (dependencia lineal) y en general, cualquier tipo de interacción entre los parámetros. Las interacciones producen un efecto de compensación entre las diferentes variaciones de los parámetros afectados y es condición suficiente aunque no necesaria, para la falta de identificabilidad; ii) No estacionariedad de los parámetros: las medidas de localización y dispersión de los parámetros se correlacionan con variables de estado o con los datos; iii) Falta de sensibilidad del modelo a los inputs: los datos no contienen el tipo y/o la cantidad de información necesaria para constreñir los parámetros del modelo; iv) El criterio de evaluación escogido (e.g. la función objetivo) no extrae adecuadamente la información contenida en los datos. Los métodos de optimización multiobjetivo alivian este problema y v) Falta de sensibilidad del modelo a variaciones en el valor de los parámetros.

Todas estas causas se resumen en tres: problemas estructurales del modelo, ausencia de datos con suficiente calidad y procedimientos de inferencia no adecuados. Lo que tradicionalmente se ha dado en llamar falta de identificabilidad o equifinalidad, responde al concepto más amplio de **incertidumbre de los parámetros** (inferidos).

Análisis de sensibilidad

El problema de identificación de modelos (de sus parámetros) ha sido tradicionalmente acoplado con las metodologías de análisis de sensibilidad, las cuales se basan en la evaluación del impacto de los cambios en los parámetros, los inputs o las condiciones iniciales, sobre los resultados del modelo (Matott et al., 2009; Pianosi et al., 2016; Pianosi y Raso, 2012; Pianosi y Wagener, 2015; Saltelli, 2004; Saltelli et al., 2008; Sorooshian y Gupta, 1995; Wagener et al., 2003; Wagener y Kollat, 2007). El análisis de sensibilidad permite identificar los parámetros que menos influyen en la respuesta del modelo, condicionado a los datos que se le presentan, lo cual aporta perspectivas e intuición para poder diagnosticar las razones por las que un modelo no funciona de forma correcta (Brun et al., 2001; Pianosi y Wagener, 2016). Así mismo, facilita la labor de discernir cual es el valor más adecuado de un parámetro, de entre todos los plausibles, aliviando en parte el problema de la no identificabilidad. El análisis de sensibilidad puede ser realizado **localmente**, modificando secuencialmente el valor de cada parámetro, bien en la cercanía de un punto dado del espacio de

parámetros o bien sobre una región amplia de valores plausibles. Así mismo, el análisis de sensibilidad puede ser **global**, modificando de forma simultánea todos los parámetros y asignándoles valores en todo su dominio. El análisis global es adecuado para cuantificar la influencia media de los parámetros y sus interacciones, sobre el resultado del modelo (Brun et al., 2001).

Hay una estrecha relación entre los conceptos de sensibilidad, incertidumbre de los parámetros e incertidumbre de la predicción. La sensibilidad del modelo a un parámetro determina la intensidad con la que dicho parámetro propaga su propia incertidumbre a la predicción del modelo. La Figura 2.1 expresa gráficamente esta idea.

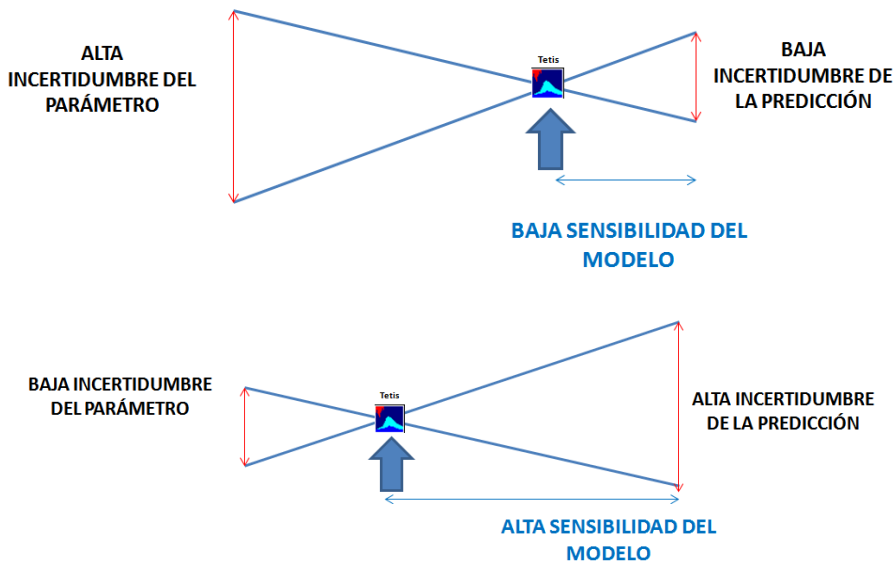


Figura 2.1 Relaciones entre sensibilidad e incertidumbre

2.1.4. Incertidumbre de los parámetros inferidos

Kuczera y Parent (1998) afirman que ningún hidrólogo debería confiar en un único conjunto determinista de los valores de los parámetros. La incertidumbre de los parámetros inferidos es el resultado de alguna de las siguientes causas, las cuales no son excluyentes: i) La ausencia de datos de calidad; ii) La

inadecuada estructura/parametrización del modelo y iii) Un método de inferencia inadecuado. Por tanto, se deduce que para evaluar correctamente cual es la incertidumbre real de los parámetros, habrá que considerar mediante el método de inferencia adecuado, tanto la incertidumbre estructural, como la debida a los datos.

En relación a **la calidad de los datos**, es una de las fuentes de incertidumbre a la que, con frecuencia, se le supone un papel marginal en la contribución a la incertidumbre total. Hay una gran variedad de tipos de observaciones empleadas en modelación hidrológica. Las más comunes son la precipitación y el caudal en el río. Sin embargo, también son de frecuente uso, temperatura, humedad del suelo, datos geomorfológicos, geológicos, de radiación solar, velocidad del viento, etc. Con carácter general, la incertidumbre en la precipitación es considerada dominante, por su variabilidad espacio-temporal. La medición de caudales también puede suponer un significativo nivel de error, debido a errores en la determinación o en el uso de las curvas de aforo.

Respecto a **la estructura del modelo**, es comúnmente aceptado que ningún modelo es perfecto para simular el sistema que representa, debido a las hipótesis simplificadoras (conocidas o no) que lo definen. Sin embargo, es difícil conocer para cada caso, donde están los fallos de especificación que motivan la imperfección. La solución comúnmente adoptada ha sido cuantificar el resto de fuentes de incertidumbre, dejando la incertidumbre del modelo como un componente incluido en el error total que no es explícitamente tratado (Gourley y Vieux, 2006). Esta forma de tratar la incertidumbre podría tener resultados adecuados, siempre que el modelo de error que trata las incertidumbres no consideradas de forma explícita sea el adecuado, y lo más general posible (Schoups y Vrugt, 2010), para tratar de modelar los efectos de la incertidumbre estructural sobre los errores (e.g. sesgo de la predicción, dependencia serial de los errores).

Otra alternativa para la consideración de la incertidumbre estructural, que está tomando fuerza en los últimos años, es considerar la modelación como una comprobación de la bondad de múltiples hipótesis (Clark et al., 2011), con el propósito de seleccionar la más adecuada o bien ponderar probabilísticamente todas ellas. Sobre estos métodos se dará una breve explicación en la sección 2.2.3. En cualquier caso, es costoso comprobar si el conjunto de hipótesis o modelos plausibles ha sido completamente explorado (Refsgaard et al., 2007),

por lo que la estimación de la incertidumbre estructural quedará limitada o condicionada a las hipótesis contempladas. Muchas otras estrategias están siendo propuestas para considerar la incertidumbre estructural del modelo, por ejemplo mediante su representación jerarquizada, empleando variaciones estocásticas de algunos de los parámetros del modelo (Renard et al., 2010, 2011). Cual, de todas las aproximaciones mencionadas, proporciona una adecuada descripción de los errores estructurales, es una cuestión que todavía no está totalmente resuelta. Sin embargo, la mayor o menor adecuación de cada metodología, es algo que podrá ser evaluado en cada caso mediante: i) La inspección de la matriz de covarianza de los parámetros, comprobando que la covarianza entre ellos sea baja; ii) La inspección y el diagnóstico de los errores y iii) La evaluación de las incertidumbres parciales y total de la predicción (Renard et al., 2011).

Finalmente, la última de las causas que afecta a la incertidumbre de los parámetros es **el método de inferencia** empleado. Tanto en los métodos de inferencia clásicos (frecuentistas) como en los Bayesianos, la incertidumbre de los parámetros se reduce conforme se añade más información al modelo. Sin embargo, dicha incertidumbre podrá estar (infra) sobreestimada si el método de inferencia (e.g. función de verosimilitud) es incorrectamente especificado (Mantovan y Todini, 2006). Los métodos de inferencia en modelación hidrológica se tratarán en la siguiente sección 2.1.5.

Cabe tener en consideración, el concepto de **incertidumbre (aleatoria) inherente** en el comportamiento de los sistemas naturales, el cual se mencionó en el capítulo 1. Esta aleatoriedad contribuye de forma inevitable (irreducible), tanto a la incertidumbre de los datos, como a la incertidumbre derivada de la estructura del modelo. Contrariamente, el método de inferencia empleado es la única causa que aporta incertidumbre a los parámetros que, en teoría, puede ser totalmente eliminada por su origen puramente epistémico. Esa es la motivación de los grandes esfuerzos que durante décadas, han sido (y siguen siendo) dedicados a la investigación sobre metodologías de estimación de parámetros y de su incertidumbre asociada (Duan et al., 1992, 2003; Sorooshian et al., 1993).

En modelación hidrológica, los primeros métodos realistas de evaluación de la incertidumbre de los parámetros se basaron principalmente en las hipótesis de la teoría clásica estadística (aproximación de primer orden): linealidad del modelo respecto de los parámetros y asunción de una distribución multinormal para los

mismos (Kuczera, 1983; Kuczera y Mroczkowski, 1998). Así, se pudieron aproximar los intervalos de confianza del conjunto de parámetros óptimo, mediante la estimación de su **matriz de covarianza de los parámetros** $\mathbb{C}(\boldsymbol{\theta})$.

En un modelo lineal, la matriz $\mathbb{C}(\boldsymbol{\theta})$ puede ser obtenida a partir de la matriz de sensibilidad del modelo respecto de los parámetros, con el llamado método del inverso de la Hessiana (Kavetski et al., 2003).

Se llama **matriz de sensibilidad** del modelo respecto de los parámetros, a la matriz Jacobiana de la función objetivo $\mathbb{J}_{\boldsymbol{\theta}} = \left[\frac{\partial f_{obj}}{\partial \boldsymbol{\theta}} \right]$. Por su parte, la matriz

Hessiana de la función objetivo se define como $\mathbb{H}_{\boldsymbol{\theta}} = \left[\frac{\partial^2 f_{obj}}{\partial_{\boldsymbol{\theta}} \partial_{\boldsymbol{\theta}}} \right]$. La matriz de

covarianza $\mathbb{C}(\boldsymbol{\theta})$ es proporcional al inverso de la matriz Hessiana de la función objetivo f_{obj} , es decir $\mathbb{C}(\boldsymbol{\theta}) \propto \mathbb{H}_{\boldsymbol{\theta}}^{-1}$. Para darle sentido a esta ecuación, nótese que $\mathbb{H}_{\boldsymbol{\theta}}$ es una medida de la tasa de variación de la sensibilidad de la f_{obj} frente a cambios en $\boldsymbol{\theta}$. Supóngase que $\boldsymbol{\theta}_{opt}$ es la estimación que optimiza f_{obj} . Si la tasa de variación de la sensibilidad ($\mathbb{H}_{\boldsymbol{\theta}}$) es grande en el punto $\boldsymbol{\theta}_{opt}$, la f_{obj} no admitirá la más pequeña variación respecto a $\boldsymbol{\theta}_{opt}$, para evitar alejarse rápidamente del objetivo. Es decir, $\boldsymbol{\theta}_{opt}$ no admitirá mucha incertidumbre: $\mathbb{C}(\boldsymbol{\theta})$ con valores pequeños. La situación contraria se entiende de igual forma. Bajos valores en la matriz $\mathbb{H}_{\boldsymbol{\theta}}$ indicarán que la f_{obj} no cambia demasiado al movernos del punto $\boldsymbol{\theta}_{opt}$, con lo que los parámetros admitirán rangos más amplios de valores: $\mathbb{C}(\boldsymbol{\theta})$ con valores grandes. En general, los parámetros a los que la función objetivo resulta muy sensible, serán parámetros poco inciertos, si bien su poca incertidumbre afectará significativamente a la incertidumbre de la predicción (Figura 2.1, diagrama inferior).

Para modelos con dependencia lineal respecto de los parámetros, la matriz de covarianza calculada como $\mathbb{C}(\boldsymbol{\theta}) \propto \mathbb{H}_{\boldsymbol{\theta}}^{-1}$ es exacta. Además, cuando hay

linealidad, dicha matriz de covarianza es constante e independiente del punto del espacio de parámetros para el que es evaluada.

En condiciones de cuasi-linealidad, se podrá aplicar la aproximación de primer orden, calculando la matriz de covarianza mediante la aplicación del mencionado método del Hessiano inverso, pero necesariamente aplicado sobre la solución óptima del espacio de parámetros (sobre la moda de la distribución Gaussiana del parámetro). Cuanto más lejos se esté de la hipótesis de linealidad del modelo y/o menos cierta sea la hipótesis de Gaussianidad del parámetro, menos válida será la aproximación de la matriz de covarianza, y por tanto, la estimación de la incertidumbre de los parámetros.

Para los modelos en los que existe alta no-linealidad respecto de los parámetros (los modelos hidrológicos son así en su mayoría), la Hessiana y por tanto la matriz de covarianza deben ser aproximados numéricamente. En estos casos, la matriz de covarianza varía considerablemente sobre el dominio de los parámetros, a diferencia del caso lineal en el que es constante. Duan et al. (1992) observan que, en este tipo de modelos, la superficie de densidad mapeada en el espacio de parámetros, se desvía marcadamente de la suavidad cuadrática de la multinormal. Estas desviaciones incluyen óptimos locales, discontinuidades no derivables y crestas multidimensionales. Jakeman y Hornberger (1993) apuntan que muchos de esos problemas pueden deberse al empleo de modelos altamente sobreparametrizados respecto a la cantidad de datos disponible para la calibración. Por su parte, Kuczera y Parent (1998) y Vrugt y Bouten (2002) apuntan que es común en modelos con alta no-linealidad respecto de los parámetros, que la incertidumbre de estos sea pobremente descrita por una distribución multinormal, debido a fuertes interacciones no lineales entre los parámetros, donde es incluso posible la presencia de múltiples modas de densidad de probabilidad. En definitiva, en modelación hidrológica, donde los modelos presentan usualmente una fuerte no-linealidad, no es posible obtener una expresión analítica explícita para las funciones de densidad marginales (o la conjunta) de los parámetros. Otras técnicas estadísticas para estimar la incertidumbre de los parámetros, basadas en la teoría Bayesiana, serán la solución que se hará más popular a partir de la última década del siglo pasado, si bien su origen se remonta a la mitad de dicho siglo.

2.1.5. Revisión de métodos de inferencia de los parámetros

La Real Academia Española (RAE) define el término ***inferir*** como, *deducir algo o sacarlo como conclusión de otra cosa*. La estimación (o inferencia) de los parámetros de un modelo, consiste en la utilización de las observaciones de una (o varias) variable de interés simulada por el modelo, para deducir cuales son los parámetros que mejor le hacen cumplir con el objetivo establecido, reduciendo así el conjunto de posibles modelos (hipótesis) que pueden explicar la variable analizada, y por tanto, reduciendo también su incertidumbre (Beven, 2009).

Según Kavetski et al. (2003), un **método de inferencia** se puede entender como un conjunto de instrucciones o reglas, que el modelo utiliza para extraer información contenida en los datos que se le presentan. La información extraída queda almacenada en los parámetros. Forman parte de esas instrucciones al modelo, la adecuada consideración de las fuentes de error (incertidumbre) de cada uno de los componentes que intervienen en el proceso de modelación, principalmente los datos y el propio modelo. Así mismo, la posible información previa que se tenga sobre el valor de los parámetros también debería ser considerada para, al menos, establecer el rango de valores posibles que pueden adoptar.

Dados una estructura de modelo y unos datos, las reglas proporcionadas al modelo por el método de inferencia van a ser determinantes en la fiabilidad (exactitud + precisión) de los parámetros estimados y por tanto, en la fiabilidad de las predicciones del modelo. Las hipótesis adoptadas en el conjunto de reglas han de comprobarse a posteriori, verificando que la inferencia es compatible con dichas reglas. Si la verificación resulta negativa, los parámetros y resultados inferidos no serán fiables.

Desde hace tiempo (Kuczera, 1983; Sorooshian y Dracup, 1980), y de manera más intensa en este siglo (Ajami et al., 2007; Evin et al., 2013, 2014; Del Giudice et al., 2013; Honti et al., 2013; Kavetski et al., 2003, 2006a, 2006b; Reichert y Mieleitner, 2009; Renard et al., 2011, 2010; Scharnagl et al., 2015; Schoups y Vrugt, 2010; Smith et al., 2010, 2015; Thyer et al., 2009; Yen et al., 2014), la literatura científica está poniendo un foco crítico en la manera en que tradicionalmente, los datos observados sobre el sistema real son mostrados al modelo de simulación. Dicho de otro modo, han surgido inquietudes sobre si los métodos de inferencia empleados tradicionalmente son los adecuados para

estimar los valores de los parámetros de los modelos hidrológicos, así como para obtener su incertidumbre. El problema radica en que, usualmente, la calibración de modelos ha estado basada en la optimización de funciones objetivo, diseñadas ad-hoc, en lugar de basarse en métodos robustos que atiendan a las propiedades estadísticas de los datos observados, de los resultados del modelo y de la confrontación de ambos, mediante el empleo de funciones de verosimilitud apropiadas.

Por tanto, el objetivo de la inferencia de los parámetros de un modelo hidrológico ha sido tradicionalmente malentendido al considerarlo como el proceso de obtención de los parámetros que proporcionan el mejor ajuste observado-simulado. Siendo eso lo más comúnmente aceptado, el objetivo teóricamente correcto es, sin embargo, obtener los parámetros que proporcionan el mejor ajuste de los errores de modelación a las hipótesis que sobre ellos se han establecido y en base a las cuales se ha efectuado la inferencia. El origen de la confusión está en que, bajo ciertas hipótesis, ambos objetivos son el mismo. Esas hipótesis son las que subyacen a los métodos de ajuste por mínimos cuadrados, que la presente Tesis trata más adelante bajo la denominación de método de inferencia, en el “caso ideal”. Normalmente, los modelos hidrológicos no cumplen con las hipótesis de ese “caso ideal”.

Los principales métodos de inferencia de parámetros que se emplean en modelación hidrológica se pueden clasificar en: métodos de optimización, métodos Bayesianos y métodos de asimilación de datos.

1- MÉTODOS DE OPTIMIZACIÓN

Sus características generales son:

- a. Se basan en optimizar una o varias funciones objetivo (multiobjetivo).
- b. Los métodos multiobjetivo (Efstratiadis y Koutsoyiannis, 2010; Fenicia et al., 2007; Gupta et al., 1998; Kollat et al., 2012; Madsen, 2000; Sorooshian et al., 1993; Vrugt et al., 2003; Vrugt y Robinson, 2007a; Yapó et al., 1998) no aportan una única solución óptima al valor de los parámetros. Las posibles soluciones, llamadas “no-dominadas”, quedan ubicadas en un espacio nf -dimensional sobre una $(nf-1)$ -variedad denominada Frente de Pareto, siendo nf el número de funciones objetivo. El Frente de Pareto está conformado por todas las soluciones

Pareto-óptimas, que se caracterizan por no poderse mejorar ninguno de los objetivos, sin empeorar alguno(s) de los objetivos restantes.

- c. Los parámetros se consideran deterministas durante el proceso de optimización. Es decir, la solución es un único conjunto de parámetros (o varios conjuntos discretos, en la aproximación multiobjetivo).
- d. Los métodos de búsqueda de la solución pueden ser métodos de optimización local o global.
- e. La estimación de la incertidumbre de los parámetros se realiza a posteriori, bajo la hipótesis de linealidad del modelo respecto a los parámetros y de distribución multinormal de los mismos.

Los métodos de optimización clásicos, o **de búsqueda local** (Duan et al., 2003; Gupta y Sorooshian, 1985; Hendrickson y Sorooshian, 1988), están diseñados para encontrar el máximo (mínimo) de la función objetivo, guiados mediante el cálculo de derivadas respecto de los parámetros (Jacobianas, Hessianas), sobre la superficie de respuesta de la función objetivo (e.g. métodos de Newton, Gauss-Newton, Levenberg-Marquardt etc.). Como excepción, un método local que no emplea derivadas es el Simplex (Nelder, 1965). En modelos altamente no lineales, como es el caso de los modelos hidrológicos, la función objetivo puede definir una compleja y rugosa superficie de respuesta, plagada de mínimos locales, discontinuidades no derivables, planicies y múltiples regiones de atracción, que en un espacio de parámetros hiperdimensional, hacen complicado encontrar eficiente y eficazmente el óptimo global. Ninguno de los métodos locales mencionados, está libre de tener problemas de convergencia en la optimización de estos modelos.

Con el propósito de abordar la compleja optimización de los modelos no lineales, surgen los métodos de optimización de **búsqueda global**, que generalmente se basan en una búsqueda aleatoria del óptimo, a diferencia de la búsqueda guiada por gradiente de los métodos locales. Los principales algoritmos globales son los de tipo estocástico (e.g. muestreo por Monte Carlo (Eckhardt, 1987; Metropolis, 1949, 1987), Recocido Simulado (Kirkpatrick et al., 1983)) y los de base (Meta)Heurística (e.g. algoritmos evolutivos (Holland, 1975), algoritmos evolutivos diferenciales (Storn y Price, 1997), o los algoritmos basados en inteligencia de enjambres (Kennedy y Eberhart, 1995; Shi y Eberhart, 1998)).

En el ámbito de la modelación hidrológica, las mencionadas consideraciones acerca de los problemas de convergencia de los métodos de optimización locales, inspiraron a Duan et al. (1992) a desarrollar un potente, robusto y eficiente método de optimización de búsqueda global: el algoritmo Shuffled Complex Evolution (SCE-UA). Numerosos casos de estudio han demostrado su eficiencia y efectividad en la localización de los parámetros “óptimos” con diversos modelos hidrológicos. Sin embargo, tal y como ya se ha mencionado en apartados previos, en ciertos casos resultaba imposible encontrar un único conjunto de parámetros cuya medida de verosimilitud, resultara significativamente diferente de otros conjuntos de parámetros (equifinales), generalmente próximos en el espacio de parámetros. El problema de la pobre identificabilidad de los parámetros seguía existiendo. El algoritmo SCE-UA ha sido considerado durante años como un algoritmo de optimización en la vanguardia del estado del arte. Sin embargo, el trabajo de Vrugt et al. (2009a) demuestra que, en algunos casos de estudio, el SCE-UA termina el proceso de optimización en un mínimo local, mientras que el nuevo algoritmo Bayesiano de inferencia (DREAM) presentado en ese trabajo, sí alcanza el óptimo global.

La Figura 2.2 muestra el esquema clásico de un problema de calibración por optimización automática. Los parámetros del modelo son ajustados de forma iterativa, de forma que los resultados del modelo f (línea gris continua), se aproximen lo más posible a la respuesta observada del sistema (línea discontinua).

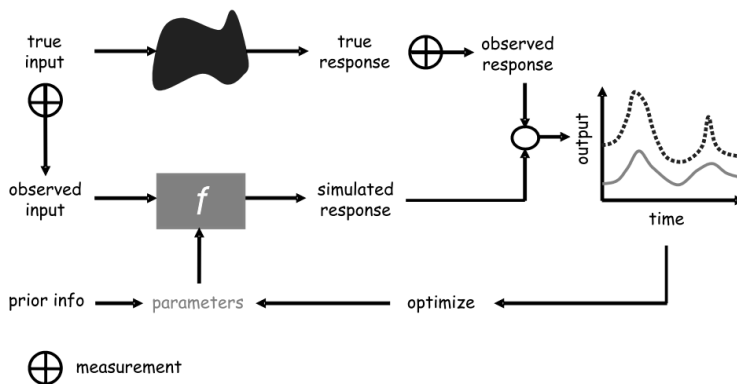


Figura 2.2 Esquema clásico de calibración automática. Fuente: Vrugt et al. (2008)

2- MÉTODOS BAYESIANOS

Sus características generales son:

- a. Realizan una inferencia de base probabilística, por lo que las “funciones objetivo” se definen en términos de probabilidad o verosimilitud. Así, se evalúa la probabilidad de cada conjunto de parámetros, dadas las observaciones.
- b. Los parámetros se consideran inciertos por naturaleza (variables aleatorias). Por tanto, el resultado de la inferencia Bayesiana no es un conjunto de parámetros (un d -vector), sino una distribución d -variada de densidad de probabilidad, a la que se llamará **distribución posterior**, como se explica más adelante.
- c. El equivalente a los métodos de búsqueda empleados en optimización, son los métodos de muestreo que permiten obtener la distribución posterior, puesto que normalmente con modelos complejos, no es posible su derivación analítica.
- d. La inferencia Bayesiana obtiene de forma directa la incertidumbre de los parámetros, permitiendo además el manejo de otras fuentes de incertidumbre.

Según el tipo de función de verosimilitud empleada, se pueden distinguir dos metodologías: los **métodos Formales Bayesianos** y los **métodos Bayesianos aproximados**. En los métodos Formales, las funciones de verosimilitud tienen base probabilística y se construyen a partir del modelo de los errores de modelación (diferencia entre observados y simulados) que se haya asumido. Los métodos Formales son validables mediante el análisis a posteriori de los errores y permiten evaluar la incertidumbre de la predicción de forma inmediata, como se explicará más adelante (sección 2.4). En cuanto a los métodos aproximados, no hacen uso de una función de verosimilitud con base probabilística, sino que emplean evaluadores ad-hoc de la verosimilitud (índices de eficiencia, *firmas* de la cuenca, etc.). Son útiles cuando no se tiene la información apropiada para construir una función de verosimilitud formal. Con los métodos aproximados, generalmente solo se pueden obtener distribuciones posteriores de los parámetros muy difusas (dispersas), y no es posible obtener la incertidumbre de la predicción. Estos métodos se agrupan bajo el nombre de ABC (en inglés *Approximate Bayesian Computation*) (Csilléry et al., 2010; Marin et al., 2012; Nott et al., 2012; Sadegh y Vrugt, 2013; Vrugt y Sadegh, 2013).

3- MÉTODOS DE ASIMILACIÓN DE DATOS

En el marco del cuestionamiento de la utilidad del paradigma clásico de estimación de parámetros, se han incorporado los métodos secuenciales de asimilación de datos (métodos DA). Dichos métodos, datan de los años sesenta del siglo pasado (Filtro de Kalman), si bien no fueron introducidos en Hidrología operacional hasta principios de los 80, siendo empleados para mejorar las predicciones de avenidas en tiempo real.

Los métodos DA proporcionan un marco general que permite tratar explícitamente con las incertidumbres de los inputs, la incertidumbre estructural y los errores de observación. Dicho entorno permite que los inciertos resultados del modelo sean secuencial y continuamente fusionados con las observaciones, de una forma optimizada. Dos son las importantes diferencias con los métodos clásicos de estimación de parámetros. La primera es que los métodos DA corrigen en cada paso temporal, las variables de estado del modelo en base a las observaciones, mejorando así la predicción. La segunda es que los filtros DA asumen como conocidos y generalmente estáticos, los parámetros del modelo hidrológico.

Para extender la utilidad de los métodos secuenciales DA, han surgido varias aproximaciones (Vrugt et al., 2013, 2005; Zhou et al., 2014), que permiten tanto la actualización secuencial de las variables de estado del modelo, como la reestimación de sus parámetros. Los beneficios que, a priori, se pueden atribuir a esta aproximación a la estimación de los parámetros son principalmente tres: i) Una estimación fiable de la incertidumbre de la predicción; ii) Unos parámetros estimados que no se corrompen intentando corregir los problemas del modelo, inputs, etc., puesto que la modificación de la variable de estado ya soluciona estos problemas y iii) La serie temporal de correcciones, que se hacen sobre las variables de estado, puede permitir la investigación y dar luz a los problemas que están afectando al proceso de modelación.

2.2. Inferencia Bayesiana

Una aproximación a la inferencia de los parámetros, que es capaz de tratar con espacios paramétricos de alta dimensión, a la que no le afecta la marcada no linealidad de los modelos, o la no Gaussianidad de los parámetros, son los métodos de la estadística Bayesiana (Clark, 2005). Dichos métodos, pueden

incorporar información a priori sobre los parámetros. Así mismo, no necesitan la identificabilidad de los parámetros para poder ser empleados (Omlin y Reichert, 1999; Reichert y Omlin, 1997). Además, lo más importante, al estar estrictamente basados en la teoría estadística, sus hipótesis pueden (y deben) ser verificadas a posteriori. La contrapartida es que dichos métodos presentan, comparativamente, un alto coste computacional.

En la **estadística clásica o frecuentista**, el parámetro es considerado un valor fijo pero desconocido. La inferencia del parámetro se basa en una función de densidad de probabilidad (en adelante pdf) en el espacio muestral, que contiene el conjunto de todos los valores posibles de la variable observada. La probabilidad se interpreta como una frecuencia relativa en el largo plazo, sobre todos los valores en el espacio muestral, dado el desconocido parámetro.

En la **estadística Bayesiana**, los parámetros son considerados como variables aleatorias, y la inferencia del parámetro se basa en la pdf posterior, situada en el espacio de parámetros. Todo se basa en el conocido teorema (también referido como regla o ley) descubierto por el reverendo Thomas Bayes (1701-1761). Si bien fue publicado póstumamente sin tener demasiada trascendencia (Stigler, 1983), algunos años más tarde, el teorema fue redescubierto por Laplace (1749-1827). Sin embargo, hasta principios del siglo pasado, con el establecimiento del sistema axiomático de la teoría de la probabilidad por Kolmogorov (1903-1987), el teorema no toma la consistencia matemática que hoy tiene. Siendo la piedra angular de la teoría moderna de la probabilidad, resulta consistente con ambas teorías de aproximación al conocimiento: la frecuentista y la subjetivista (o Bayesiana).

El teorema de Bayes tiene su base teórica estadística, en la definición matemática de probabilidad condicional. Así, el teorema especifica la relación entre la probabilidad de dos entidades o eventos, y sus respectivas probabilidades condicionales. En lo práctico, el teorema permite actualizar las probabilidades de eventos no observados, dados otros eventos relacionados que sí que han sido observados. Es decir, el teorema de Bayes es un algoritmo de actualización, que nos proporciona la pdf posterior de los parámetros, mediante la combinación de dos fuentes de información sobre dichos parámetros: La distribución de probabilidad a priori de los parámetros (en adelante prior), y los datos observados. Por tanto, el teorema de Bayes resume en la pdf posterior, todo el conocimiento que se tiene sobre los parámetros, y no solo el originado

por las observaciones. El proceso de inferencia Bayesiano tiene muchos elementos en común con el proceso de aprendizaje humano, en el cual los conocimientos, comportamientos, creencias o preferencias existentes, son continuamente modificados o reforzados, en respuesta a la información recibida.

Según Laine (2008), el verdadero test de un modelo es su habilidad para predecir las observaciones, viejas y nuevas. La inferencia Bayesiana no es más que una herramienta que persigue ese objetivo ajustándose al paradigma falsacionista de Popper (1957): *“Las observaciones refutarán tanto un mal modelo, como unas mal seleccionadas priors”*. El Falsacionismo (o Racionalismo Crítico) surge como el modelo estándar del método científico, en respuesta al método de inducción y establece que las teorías deben de ser falsificadas más que comprobadas. Según Beck (1987), los modelos hidrológicos solo deberían ser usados para testear hipótesis, siendo estas útiles, cuando pueden ser descartadas en base a los resultados del modelo. En palabras del Nobel Enrico Fermi (1901-1954): *“Si el resultado confirma la hipótesis, has hecho una medida. Si el resultado es contrario a la hipótesis, has hecho un descubrimiento”*. Sin embargo, en modelación hidrológica no ha sido inusual considerar las evidencias en contra de las hipótesis, como *outliers* (Beven, 2009), excluyéndolas del conjunto de observaciones.

Los estadísticos han conocido desde hace tiempo, las claras ventajas que ofrece la aproximación Bayesiana a la estadística inferencial, respecto a la aproximación frecuentista. Sin embargo, no ocurre lo mismo con los algoritmos computacionales que han permitido el muestreo de las distribuciones de los parámetros, en problemas que eran intratables de forma analítica. De hecho, no fue hasta mediados de los 90 del siglo pasado, cuando gracias a los trabajos de Müller (1991) y Tierney (1994) se estimuló el interés en el mundo estadístico por algoritmos que se habían estado utilizando en física desde 1953 (Hastings, 1970; Metropolis et al., 1953). Las soluciones Bayesianas, fáciles en la teoría, se vuelven complicadas de llevar a la práctica en problemas de modelación de sistemas reales y complejos. La estadística Bayesiana computacional cambió (y sigue cambiando) esto, gracias al empleo de algoritmos que permiten muestrear la desconocida distribución posterior de forma eficiente y usar sus muestras, como base para la inferencia.

Kuczera y Parent (1998) fueron los primeros en introducir los métodos de estadística Bayesiana computacional, en el mundo de la modelación hidrológica.

Desde mitad de la primera década del presente siglo, comienzan a emplearse profusamente, sumándose así la Hidrología, a otras ramas de la ciencia que están en la vanguardia de la investigación (Astronomía, Cosmología, Medicina, Genética, etc.).

2.2.1. Estimación Bayesiana de los parámetros y su incertidumbre

Considérese un vector de n mediciones de una variable \mathbf{Y} , cuyo análisis es de interés $\tilde{\mathbf{Y}} = \{\tilde{y}_1, \dots, \tilde{y}_n\}$, observadas de forma discreta en los tiempos $t = \{1, \dots, n\}$. La inferencia Bayesiana considera que tanto las variables (e.g. \mathbf{Y}), como los parámetros del modelo que se quiere inferir $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_d\}$, tienen carácter aleatorio, formando ambos una pdf conjunta $P(\boldsymbol{\theta}, \tilde{\mathbf{Y}})$ llamada el **modelo Bayesiano completo** (Bolstad, 2010). Considérese la ley de probabilidades condicional, mediante la que se puede escribir dicha pdf conjunta (Figura 2.3, panel derecho) según la ecuación (2.1)

$$P(\boldsymbol{\theta}, \tilde{\mathbf{Y}}) = P(\tilde{\mathbf{Y}} | \boldsymbol{\theta})P(\boldsymbol{\theta}) \quad (2.1)$$

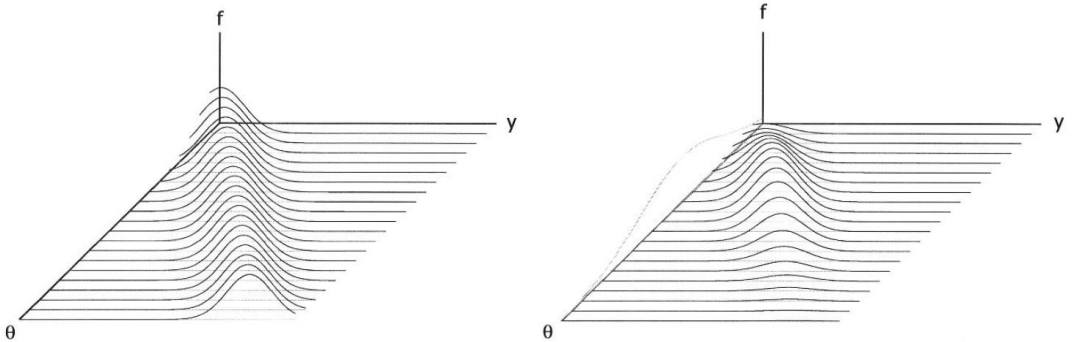


Figura 2.3. Izquierda: Superficie con posibles pdf's (muestrales) de la variable de interés, asumiendo una prior plana del parámetro. Derecha: Superficie de probabilidad conjunta de los parámetros (un solo parámetro θ en este caso) y la variable de interés. En el plano lateral se representa la prior del parámetro. Fuente: Bolstad (2010)

A la probabilidad $P(\tilde{\mathbf{Y}}|\boldsymbol{\theta})$, se le llama **pdf muestral de las observaciones**, cuando los parámetros se consideran fijos (Figura 2.3, panel izquierdo). Sin embargo, cuando son los datos los que se consideran fijos, la misma expresión analítica se denomina **función de verosimilitud** (de los parámetros), la cual suele expresarse como $\ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$. La función de verosimilitud (Figura 2.4, panel izquierdo) no es una pdf, dado que su integral no suma uno. La verosimilitud cuantifica mediante reglas probabilísticas, la adecuación de los valores de los parámetros, dados (condicionados a) el modelo y las n observaciones de la variable de interés. En la práctica, es más intuitivo expresar la verosimilitud de los simulados (o de los parámetros que los generan) en función de los errores (diferencia entre observados y simulados), en lugar de expresarla en función de los propios simulados y de los observados. Es por esto, que el paso previo para definir una expresión formal (con base probabilística) de la función de verosimilitud, es la definición de las propiedades de los errores, es decir, la estructura del modelo de error.

Por otra parte, el término $P(\boldsymbol{\theta})$ de la ecuación (2.1) es la **pdf prior** de los parámetros, la cual contiene el conocimiento que se tiene acerca de ellos, antes de presentar los nuevos datos al modelo para inferirlos. En una interpretación directa de la inferencia Bayesiana, la información a priori puede representar la honesta (y posiblemente subjetiva) opinión que el investigador tiene sobre lo desconocido (Laine, 2008). El trabajo de Berger et al. (2009) describe diferentes técnicas para construir la distribución Prior. En general, las pdf prior pueden ser construidas sobre la base de conocimiento adquirido sobre los parámetros en libros, publicaciones, experiencia previa de campo o laboratorio, e incluso datos basados en información cualitativa y conocimiento experto. Las pdf prior son clasificadas como informativas o no-informativas. Una prior no-informativa asume que todos los valores posibles del parámetro son equiprobables. Así mismo, proporciona el rango de valores que adoptará el parámetro. Por tanto, aunque escasa, la pdf prior aporta algo de información. En la literatura aparecen denominaciones más adecuadas que la de “no-informativa”: *difusa*, *débil*, *uniforme* o *plana*.

Volviendo al tema de la estimación Bayesiana de los parámetros, el teorema de Bayes se puede deducir a partir del hecho de que la ecuación (2.1) es

equivalente a la ecuación $P(\boldsymbol{\theta}, \tilde{\mathbf{Y}}) = P(\boldsymbol{\theta} | \tilde{\mathbf{Y}})P(\tilde{\mathbf{Y}})$ y por tanto, igualando y reordenado ambas expresiones se llega a la ecuación (2.2), expresión del **teorema de Bayes**.

$$P(\boldsymbol{\theta} | \tilde{\mathbf{Y}}) = \frac{\ell(\boldsymbol{\theta} | \tilde{\mathbf{Y}})P(\boldsymbol{\theta})}{P(\tilde{\mathbf{Y}})} \quad (2.2)$$

El teorema expresa, en una simple fórmula, la relación entre el conocimiento (o creencia) previo sobre los parámetros, el efecto de la información nueva y la actualización del conocimiento después de haber considerado esa nueva información, actualización que queda recogida en la pdf posterior de los parámetros $P(\boldsymbol{\theta} | \tilde{\mathbf{Y}})$ (Figura 2.4, panel derecho).

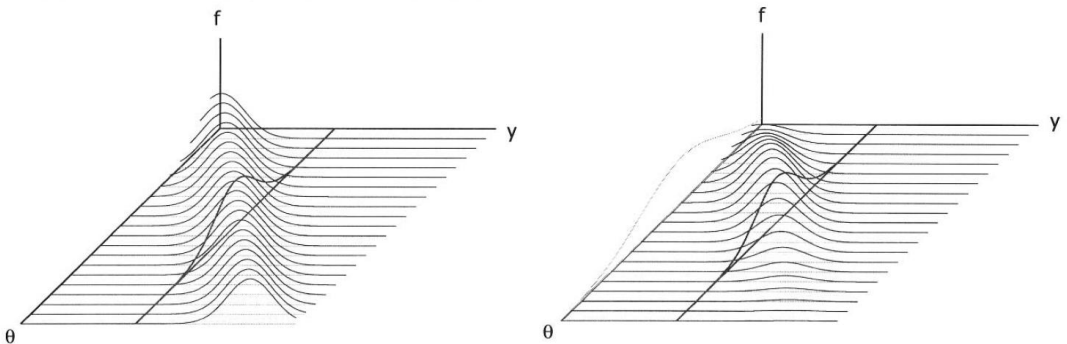


Figura 2.4 Izquierda: Función de verosimilitud del parámetro, dada la observación de la variable de interés. Derecha: La pdf posterior del parámetro dada la observación de la variable de interés. La función de verosimilitud y la posterior son el resultado de cortar con el mismo hiperplano, determinado por las observaciones, dos superficies diferentes. Fuente: Bolstad (2010)

La posterior es el objetivo de inferencia en la aproximación Bayesiana, mientras que en la aproximación frecuentista coexisten varios tipos de inferencia: la estimación puntual, la estimación de intervalo y los test de hipótesis. La pdf $P(\tilde{\mathbf{Y}})$, es la verosimilitud marginal, también llamada **evidencia Bayesiana del modelo**. Dicho término, dentro del teorema de Bayes, actúa como una constante de normalización (factor de escala), de manera que la integral de $P(\boldsymbol{\theta} | \tilde{\mathbf{Y}})$ sea

uno, a diferencia de lo que ocurre con la verosimilitud $\ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$. La expresión matemática de $P(\tilde{\mathbf{Y}})$ viene dada por la ecuación (2.3).

$$P(\tilde{\mathbf{Y}}) = \int_{\Theta} P(\boldsymbol{\theta}, \tilde{\mathbf{Y}}) d\boldsymbol{\theta} = \int_{\Theta} \ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) P(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathbb{E}_{\Theta} [\ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}})] \quad (2.3)$$

Expresada en palabras, la verosimilitud marginal es la esperanza de las verosimilitudes condicionales, sobre el espacio de parámetros plausible $\boldsymbol{\theta} \in \Theta \in \mathbb{R}^d$. Mientras que el cálculo de $P(\tilde{\mathbf{Y}})$, es necesario en la realización de test de hipótesis para la selección del modelo más verosímil de entre un conjunto de posibles modelos (ver sección 2.2.3), no será necesario para realizar la inferencia de los parámetros, tal y como se verá en la sección siguiente. Por tanto, la obtención de la pdf posterior de los parámetros puede ser realizada mediante la forma no escalada del teorema de Bayes, dada por la ecuación (2.4)

$$P(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) \propto \ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) P(\boldsymbol{\theta}) \quad (2.4)$$

Es decir, para la implementación numérica del teorema de Bayes, con el objetivo de inferir los parámetros de un modelo, solamente es necesaria la especificación de la pdf prior de los parámetros y la construcción de la función de verosimilitud. Es importante reseñar, que cuando se asume una pdf prior “plana” o no-informativa, solo los datos determinan, a través de la función de verosimilitud, la pdf posterior, puesto que $P(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) \propto \ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$. Este caso podría ser el equivalente a la estimación de los parámetros por máxima verosimilitud (MLE), si bien, dentro del paradigma Bayesiano es posible obtener la pdf posterior completa (pdf d -variada), en lugar de obtener solamente el “mejor” conjunto de parámetros (d -vector) o estimación máximo-verosímil.

Un tema que cabe mencionar es la influencia que la pdf prior puede ejercer sobre la pdf posterior. El empleo de pdf's prior resulta especialmente ventajoso cuando, ante la falta de suficientes datos con los que calibrar el modelo, la pdf posterior que se infiere resulta muy dispersa (parámetro muy incierto). El empleo de una pdf prior complementa la información de los datos y contribuye a reducir la incertidumbre del parámetro. No obstante, es importante tener en cuenta que la distribución asintótica de la moda de la posterior (también conocida como

máximo a posteriori o **MAP**) depende exclusivamente de la matriz de información de Fisher y no de la prior (Vexler et al., 2016).

Como aclaración al lector, cabe decir que la **matriz de información de Fisher** \mathbb{F} mide la cantidad de información sobre los parámetros de un modelo, que contienen las observaciones de una variable aleatoria, simulada por dicho modelo. Esta medida queda representada por la curvatura (derivada de segundo orden de dicha variable respecto del parámetro) que presenta la superficie de respuesta de dicha variable del modelo, en el entorno del valor estimado óptimo del parámetro. Una medida baja de curvatura (valores bajos en la matriz de Fisher) indicaría que hay poca información en las observaciones de la variable, para apoyar el valor MAP del parámetro. Una curvatura pronunciada indicaría que los datos observados de la variable apoyan fuertemente el valor MAP del parámetro. Adicionalmente, cabe mencionar un resultado interesante, relacionado con la matriz de Fisher. La llamada **cota inferior de Cramer-Rao**, que establece que la inversa de la matriz de información de Fisher es una cota inferior de la matriz de covarianza de los parámetros $\mathbb{C}(\boldsymbol{\theta}) \geq \mathbb{F}(\boldsymbol{\theta}, \tilde{\mathbf{Y}})^{-1}$. Esta desigualdad es una expresión matemática, que confirma la relación inversa que existe entre el contenido de información de las observaciones y la incertidumbre de los parámetros. Así mismo puede resultar interesante el ejercicio de relacionar este resultado, con lo explicado en la sección 2.1.4 sobre la relación entre la matriz de covarianza de los parámetros y la sensibilidad del modelo a los parámetros. De este ejercicio, surge la importante interrelación entre, la información que hay en los datos aprovechable por un parámetro, medida por $\mathbb{F}(\boldsymbol{\theta}, \tilde{\mathbf{Y}})$, la información que está realmente “atrapando” el parámetro, medida por $\mathbb{C}(\boldsymbol{\theta})$ y cómo de influyente en el modelo, es esa información almacenada en el parámetro, medida por la sensibilidad $\mathbb{H}_{\boldsymbol{\theta}}$.

2.2.2. Métodos de muestreo de la posterior

Desafortunadamente, en la mayoría de problemas que requieren la aplicación del teorema de Bayes, según la ecuación (2.2), no es posible obtener una expresión analítica de la pdf posterior. El problema radica en el cálculo del factor de escala de la posterior, es decir, en la integral de dimensión d representada en la ecuación (2.3). Los pocos casos en que su resolución analítica es posible

requieren que el espacio paramétrico sea de baja dimensión, que las observaciones (y por tanto la verosimilitud) se distribuyan según una pdf de la familia de las exponenciales (e.g. Gaussiana) y que la pdf prior pertenezca a una familia *conjugada* de la pdf de las observaciones. Por ejemplo, si la verosimilitud es Gaussiana y se emplea una pdf prior también Gaussiana para la media, se asegura que la pdf posterior sea también Gaussiana. Las pdf prior conjugadas se elegían de forma deliberada a fin de posibilitar el tratamiento estadístico analítico. En el resto de ocasiones, la integral de la ecuación (2.3) debe ser abordada numéricamente siendo, incluso así, difícil de obtener en el caso de espacios paramétricos de elevada dimensión. Por tanto, si bien es fácil la obtención de la posterior no escalada, el problema surge al querer obtener la verdadera posterior.

Los métodos bayesianos computacionales se basan en el desarrollo de algoritmos, que permiten extraer muestras de la verdadera posterior, incluso cuando solo se conoce su versión no escalada (ecuación (2.4)). La aproximación computacional a la estadística bayesiana no sigue la vía de integrar el factor de escala numéricamente. En su lugar emplea las **técnicas de Monte Carlo** (Bolstad, 2007, 2010; Eckhardt, 1987; Lindley, 1972; Metropolis, 1987; Robert, 2007) para la generación de muestras aleatorias, las cuales aproximarán la verdadera posterior, cuando el tamaño de la muestra sea suficientemente grande. La ventaja de esta forma de atacar el problema es no estar limitado al empleo de modelos, cuya formulación deba ser sencilla y tratable analítica o numéricamente, en perjuicio de su buena aproximación a la realidad subyacente. Los inconvenientes que plantean estas metodologías son, a grandes rasgos, los siguientes: i) ¿Cómo generar las muestras de forma eficiente? Un generador aleatorio no es factible, puesto que no genera las muestras con una adecuada distribución de densidad, similar a la pdf posterior y ii) ¿Qué número de muestras se necesita para representar una distribución multivariada cuya forma no se conoce?

De forma general, los métodos Monte Carlo se pueden agrupar en dos grandes bloques: los métodos directos y los métodos guiados por cadenas de Markov (MCMC en inglés).

Métodos de muestreo directos

Los métodos de Monte Carlo directos se basan en la extracción de muestras aleatorias de una pdf fácilmente muestreable, llamada distribución propuesta,

candidata o “de salto” que sirve de catalizador para aproximar la pdf posterior de los parámetros. Dichas muestras aleatorias son alteradas (sesgadas), mediante la aceptación de solamente algunos de sus elementos en la muestra final. Los elementos aceptados serán una muestra aleatoria de la pdf posterior (en la literatura Bayesiana, también se la llama **pdf objetivo**). La forma en que se realiza el sesgado de la muestra, es la que diferencia a un método de otro. En general, son métodos ineficientes (tienen bajo ratio de aceptación), en los que se requieren muchas muestras no sesgadas, para conseguir una muestra sesgada suficientemente grande, que permita hacer inferencia sobre la posterior. El problema de la ineficiencia se va agravando en espacios paramétricos de dimensión mayor que uno. Destacan dos métodos directos (Bolstad, 2010): los llamados de aceptación-rechazo (acceptance-rejection sampling en inglés) y los de muestreo por importancia o ponderación (importance sampling en inglés). Los métodos de muestreo por importancia suponen una considerable mejora sobre los métodos de aceptación-rechazo.

Métodos Monte Carlo, guiados por cadenas de Markov (MCMC)

Se define a una **cadena de Markov**, como un proceso estocástico discreto en el tiempo, que evoluciona “visitando” un conjunto de posibles valores o estados, llamado espacio de estados. La principal propiedad de estas cadenas de valores es la falta de memoria, es decir, cada estado solo depende de su inmediato anterior. Así mismo, las cadenas de Markov, si cumplen determinadas propiedades (Bolstad, 2010), alcanzan tras un periodo de tiempo suficientemente largo un régimen estacionario en el cual todos los estados de la cadena son visitados, cada uno con una frecuencia de visita característica e invariante en el tiempo y que responde a la llamada **distribución de equilibrio** de la cadena.

Para la explicación que sigue, resultará de ayuda al lector recordar que, un proceso estocástico es **ergódico** cuando sus propiedades estadísticas se pueden deducir de una sola, aunque suficientemente larga, realización del proceso. Se define como cadena de Markov ergódica, a aquella que tiene las siguientes propiedades: i) Alcanza el equilibrio o régimen estacionario (posee por tanto una única distribución de equilibrio); ii) Desde un estado cualquiera puede alcanzar todos los demás (propiedad de irreducibilidad) y iii) El tiempo entre visitas a un estado es finito (propiedad de recurrencia positiva). La ergodicidad es una propiedad necesaria que deben cumplir las cadenas de Markov en los métodos MCMC.

La otra propiedad para garantizar la consecución del objetivo de los algoritmos MCMC (obtener la posterior), es la condición de **reversibilidad de la cadena** una vez alcanzado el equilibrio. Esta condición viene a expresar que, dados dos estados de la cadena, la probabilidad de movimiento entre ellos debe ser la misma, independientemente de cuál de los dos sea el estado del que se parta. En una cadena de Markov reversible, se cumple la llamada **ecuación de balance detallado**, en la cual se relacionan la distribución de equilibrio de la cadena (probabilidad de permanecer en cada estado) con la probabilidad de salto entre dos estados. En resumen, si una cadena de Markov ergódica mantiene además el balance detallado, una vez alcanzado el equilibrio, con total seguridad visitará cada estado con una frecuencia proporcional a la distribución de equilibrio de la cadena. Todos estos términos y conceptos de apariencia abstracta cobran sentido en la aplicación de los algoritmos MCMC, para la obtención de la distribución posterior de los parámetros, puesto que la distribución posterior se corresponde con la mencionada distribución de equilibrio de una cadena de Markov ergódica y reversible.

Por tanto, los métodos MCMC, se basan en la simulación de una **cadena de Markov ergódica y reversible**, que genera una trayectoria aleatoria a través del espacio de parámetros, visitando de forma sucesiva diferentes puntos (antes llamados estados) de dicho espacio. Una vez la cadena ha alcanzado su régimen estacionario, la frecuencia con la cual visita cada punto o valor del parámetro, se corresponde con la frecuencia (densidad de probabilidad en parámetros con distribución continua) de dicho valor, según su distribución posterior. La simulación de la cadena se realiza mediante el muestreo de una distribución auxiliar $q(\theta_{i-1} \rightleftharpoons \theta_p)$, llamada distribución candidata, distribución propuesta o **distribución de salto**. Partiendo de un punto del espacio de parámetros θ_{i-1} , en el que se ubica centrada la distribución de salto, se muestrea aleatoriamente otro punto propuesto θ_p . El punto propuesto podrá ser rechazado o aceptado. Si es aceptado pasará a formar parte de la cadena y la distribución de salto se moverá ubicando su centro en él. Si el punto es rechazado, la distribución de salto permanecerá en el punto inicial, el cual quedará como un punto duplicado dentro del conjunto de puntos previos de la cadena. Sobre este procedimiento general, han surgido y siguen surgiendo multitud de variaciones, que afectan a, el criterio de aceptación de cada nuevo punto, a como se actúa ante el rechazo del punto o incluso, afectan a las

características de la distribución de salto, adaptando su escala y orientación (su matriz de covarianza) en función de lo aprendido al recorrer el espacio de parámetros.

El primer algoritmo MCMC fue el algoritmo de paseo aleatorio de Metropolis (**RWM** en inglés), introducido por Metropolis et al. (1953). Dicho algoritmo es la semilla desde la que ha evolucionado la inmensa variedad de algoritmos MCMC existentes hoy en día (Matthew Richey, 2010). RWM emplea una distribución de salto simétrica. Hastings (1970) hizo el algoritmo más general (**RWMH**), permitiendo el uso de distribuciones de salto asimétricas, lo que amplía considerablemente las posibilidades del método. El algoritmo RWMH o **MH** (Metropolis-Hastings) es considerado como uno de los diez algoritmos más importantes del siglo XX.

El problema con el que se encontraron Metropolis et al. (1953), fue que la mayoría de distribuciones no satisfacían la condición de reversibilidad, con lo que al no cumplirse la ecuación de balance no se podía garantizar que la cadena estuviese muestreando exactamente la posterior $P(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$. La expresión matemática del balance detallado viene dada por la ecuación (2.5)

$$\begin{aligned} P(\boldsymbol{\theta}_{i-1}|\tilde{\mathbf{Y}}) q(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) P_{\text{acep}}(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) &= \\ = P(\boldsymbol{\theta}_p|\tilde{\mathbf{Y}}) q(\boldsymbol{\theta}_{i-1} \leftarrow \boldsymbol{\theta}_p) P_{\text{acep}}(\boldsymbol{\theta}_{i-1} \leftarrow \boldsymbol{\theta}_p) & \end{aligned} \quad (2.5)$$

En la ecuación, $P(\bullet)$ representa a las probabilidades de estar en un cierto estado (es decir a la distribución posterior), $q(\bullet)$ representa a la probabilidad de moverse desde el estado actual hasta otro estado (es decir, a la distribución de salto), mientras que $P_{\text{acep}}(\bullet)$ representa a la probabilidad de que finalmente se acepte el movimiento, en base a algún criterio de convergencia. Metropolis et al. (1953) descubrieron la forma de equilibrar la ecuación de balance, actuando sobre la probabilidad de aceptación del movimiento $P_{\text{acep}}(\bullet)$. Así, decidieron calcular dicha probabilidad de aceptación mediante el llamado **ratio de Metropolis**, dado por la ecuación (2.6) que corresponde a la generalización de Hastings (1970) para distribuciones de salto asimétricas.

$$P_{\text{acep}}(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) = \min \left[1, \frac{P(\boldsymbol{\theta}_p | \tilde{\mathbf{Y}}) q(\boldsymbol{\theta}_{i-1} \leftarrow \boldsymbol{\theta}_p)}{P(\boldsymbol{\theta}_{i-1} | \tilde{\mathbf{Y}}) q(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p)} \right] \quad (2.6)$$

Aplicando esta regla de aceptación, a los puntos propuestos por la distribución de salto, se equilibra la ecuación de balance y, por tanto, se asegura la convergencia de la cadena hacia el muestreo de la distribución posterior de los parámetros. En el caso de emplear distribuciones de salto simétricas, la expresión se simplifica mediante la cancelación de los términos $q(\boldsymbol{\theta}_{i-1} \rightleftharpoons \boldsymbol{\theta}_p)$, puesto que ambas probabilidades coinciden. Así, escribiendo la ecuación (2.6) para el caso simétrico y sustituyendo las probabilidades posteriores por su ecuación no escalada, se obtiene la ecuación (2.7)

$$P_{\text{acep}}(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) = \min \left[1, \frac{\ell(\boldsymbol{\theta}_p | \tilde{\mathbf{Y}}) P(\boldsymbol{\theta}_p)}{\ell(\boldsymbol{\theta}_{i-1} | \tilde{\mathbf{Y}}) P(\boldsymbol{\theta}_{i-1})} \right] \quad (2.7)$$

A la vista de la ecuación (2.7), se entiende el motivo de que esta regla de aceptación permita trabajar con las distribuciones no escaladas de la posterior de los parámetros, como se explica a continuación. Si se considerase la constante de escalamiento $P(\tilde{\mathbf{Y}})$, al aparecer en el numerador y denominador de la ecuación (2.7), quedaría automáticamente cancelada.

Otra muy interesante propiedad de esta regla de aceptación, que permite evolucionar la búsqueda por el espacio de parámetros, es la siguiente. Con el propósito de simplificar la explicación se considerará una prior uniforme, de forma que la ecuación (2.7) se reduce a la ecuación (2.8)

$$P_{\text{acep}}(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) = \min \left[1, \frac{\ell(\boldsymbol{\theta}_p | \tilde{\mathbf{Y}})}{\ell(\boldsymbol{\theta}_{i-1} | \tilde{\mathbf{Y}})} \right] \quad (2.8)$$

De la ecuación (2.8), se deduce que cuando el punto propuesto por la distribución de salto $\boldsymbol{\theta}_p$, es más verosímil que el punto del estado actual de la cadena $\boldsymbol{\theta}_{i-1}$, la probabilidad de aceptación será $P_{\text{acep}}(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) = 1$, por lo que el movimiento siempre se producirá cuando mejora la verosimilitud. Sin embargo,

cuando la verosimilitud del nuevo punto es menor que la del punto actual, también se le da una oportunidad al movimiento, el cual se producirá con una probabilidad de aceptación calculada como $P_{\text{acep}}(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) = \ell(\boldsymbol{\theta}_p | \tilde{\mathbf{Y}}) \ell^{-1}(\boldsymbol{\theta}_{i-1} | \tilde{\mathbf{Y}})$.

Es decir, la ratio de Metropolis permite de forma ocasional, movimientos que “empeoran” el estado actual. Esta es la gran innovación de estos métodos Bayesianos de estimación de parámetros. Mientras que los tradicionales algoritmos de optimización están sujetos a una degeneración progresiva del espacio de búsqueda, los algoritmos que siguen la regla de Metropolis promulgan la diversificación necesaria, que permite la exploración de espacios de parámetros difíciles, como los que planteaban problemas de no convergencia o convergencia en mínimos locales con los clásicos métodos de optimización. El desarrollo del algoritmo MH puede encontrarse en multitud de libros y artículos científicos (e.g. Bolstad (2010)).

Es importante tener en cuenta las siguientes consideraciones. En primer lugar, con carácter general, cuanto más se parezca la distribución de salto escogida, a la posterior de los parámetros, más valores propuestos serán aceptados (mayor será la ratio **de aceptación**) y antes se alcanzará la convergencia de la cadena. En segundo lugar, la elección de la distribución de salto, de la cual depende en gran medida la eficiencia del algoritmo MCMC, conlleva asumir una solución de compromiso. Por una parte, cuando la distribución de salto es muy estrecha, la mayor parte de los puntos candidatos serán aceptados, si bien la distancia cubierta en cada salto será pequeña y muchos los saltos necesarios para explorar todo el espacio de parámetros, siendo lenta la convergencia de la cadena al régimen estacionario. Por otro lado, con distribuciones de salto demasiado anchas, muchos puntos candidatos serán rechazados de forma consecutiva, provocando que la cadena no se mezcle bien y por tanto, ralentizando el proceso de convergencia. Se dice que una cadena presenta buenas **propiedades de mezclado** (Figura 2.5), si puede alcanzar en pocos saltos cualquier punto del espacio de parámetros, sea cual sea el punto inicial de partida de la cadena.

En tercer lugar, hasta ahora se ha estado hablando de alcanzar el estado estacionario, o de equilibrio o la convergencia de la cadena de Markov, tras un número suficiente de simulaciones. Sin embargo, a priori, se desconoce cuál es ese número suficiente de simulaciones, el cual debe ser determinado “sobre la marcha” durante el proceso de evolución de la cadena. El tiempo que transcurre

hasta alcanzar la convergencia, se llama **periodo de calentamiento** de la cadena. La determinación del número de simulaciones de calentamiento se realiza mediante el empleo de uno o más diagnósticos de convergencia aplicados sobre los valores ya simulados por la cadena.

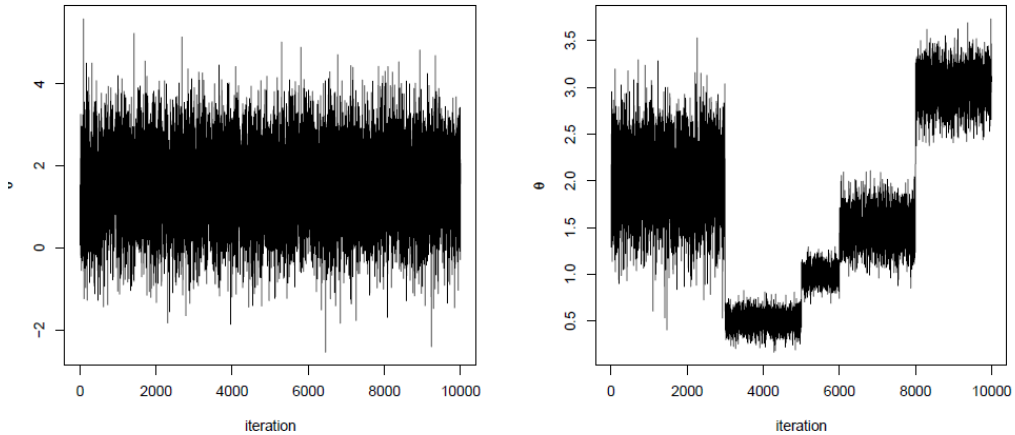


Figura 2.5 Evolución de dos cadenas diferentes, durante el muestreo MCMC de un parámetro. Izquierda: buen mezclado de la cadena representativo de haber alcanzado el estado de equilibrio. Derecha: mal mezclado de la cadena

Se puede encontrar un gran volumen de investigación desarrollada en la última década del siglo pasado (Cowles y Carlin, 1996; Gelman y Rubin, 1992; Gelman y Shirley, 2011; Geweke, 1992; Raftery y Lewis, 1992) sobre los diagnósticos que ayudan a determinar si la cadena de Markov puede haber alcanzado el equilibrio. Una colección de los mejores diagnósticos fue recopilada por Best et al. (1995) y convertida en el software CODA (del que hay versión para el paquete estadístico R). Por otro lado, también se emplean los test de convergencia de carácter visual, como la visualización de la traza (Figura 2.6) de valores de la cadena, su estacionariedad y sus propiedades de mezclado.

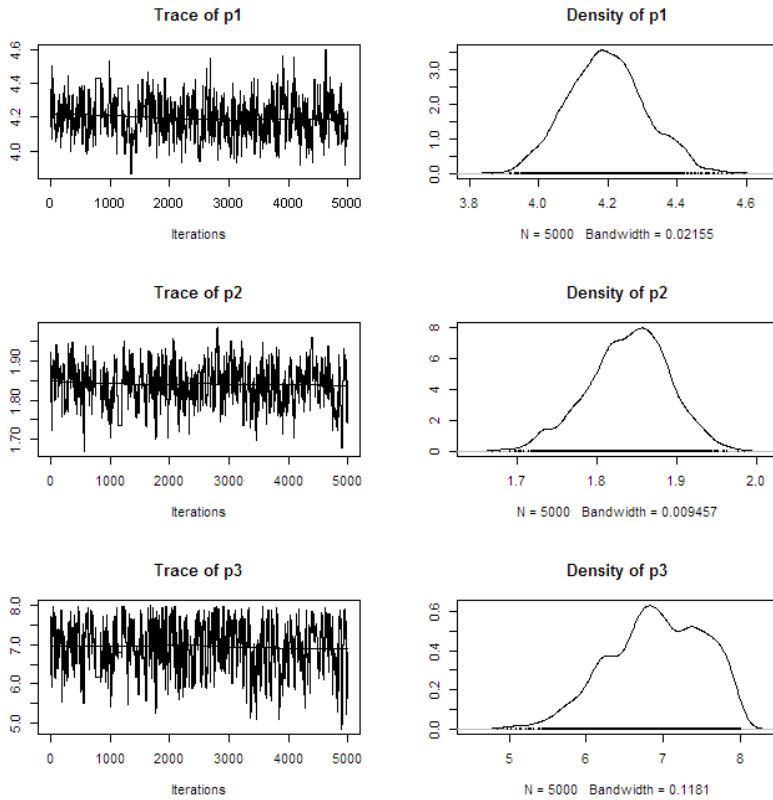


Figura 2.6. Test de convergencia MCMC, mediante la inspección de las trazas de tres parámetros, con sus correspondientes densidades posterior muestreadas

Siguiendo con las consideraciones a tener en cuenta, la cuarta sería que para incluir a los puntos visitados y aceptados, como muestras de la posterior de los parámetros, la cadena de Markov debe haber alcanzado necesariamente la convergencia. Así, la muestra de la posterior será la obtenida a partir del primer punto aceptado tras el periodo de calentamiento. La quinta y última consideración es que la muestra de la posterior obtenida, tal cual se ha explicado, no es una muestra independiente (tiene correlación de orden 1 o Markoviana), por lo cual deberá ser procesada y reducida convenientemente (“**thinning process**”) antes de poder realizar los oportunos cálculos de inferencia sobre ella. Esto no afecta a la estimación de la moda o MAP, pero si a la estimación de los intervalos de credibilidad y los momentos de la posterior.

Métodos MCMC adaptativos y de cadena múltiple

En la primera década del presente siglo, se propusieron diferentes aproximaciones con el propósito de incrementar la eficiencia de los métodos MCMC, mejorando el algoritmo original MH. Estas mejoras pueden ser agrupadas en dos grupos: mejoras sobre los métodos adaptativos de cadena simple y desarrollo de los métodos multicadena.

Los **algoritmos adaptativos de cadena simple** más conocidos son (Roberts y Rosenthal, 2009), el de la distribución propuesta (o de salto) adaptativa o algoritmo AP (Haario et al., 1999), el algoritmo adaptativo de Metropolis o algoritmo AM (Haario, 2001) y el algoritmo AM con rechazo retrasado o algoritmo DRAM (Haario et al., 2006). Todos estos algoritmos MCMC simulan una sola trayectoria (cadena) mediante la extracción de muestras de la distribución de salto, tal y como hace RWMH. La innovación que introducen es la adaptación de la matriz de covarianza de la distribución de salto, cada cierto número de extracciones de la cadena. Dicha adaptación se realiza en base a lo que se aprende de las muestras aceptadas en el pasado de la cadena. Esta adaptación implica la progresiva alineación de la distribución de salto, con la orientación y la escala de la distribución objetivo (posterior de los parámetros). La adaptación mejora, a veces de forma dramática, la velocidad de convergencia de la cadena. Sin embargo, cualquier tipo de adaptación “sobre la marcha” presenta un importante inconveniente: deja de satisfacerse la ecuación de balance detallado (ecuación (2.5)). Esto se explica fácilmente. La probabilidad condicional del salto hacia adelante $q(\boldsymbol{\theta}_{i-1} \rightarrow \boldsymbol{\theta}_p) = q(\boldsymbol{\theta}_p | \boldsymbol{\theta}_{i-1})$ realizado con la matriz de covarianza previa a la adaptación, no es igual a la probabilidad condicional del salto hacia atrás $q(\boldsymbol{\theta}_{i-1} \leftarrow \boldsymbol{\theta}_p) = q(\boldsymbol{\theta}_{i-1} | \boldsymbol{\theta}_p)$, el cual se realiza ya con la nueva matriz de covarianza adaptada. El balance es rápidamente restaurado durante las siguientes simulaciones hasta la siguiente adaptación. Sin embargo, la cadena resultante, a causa de esas “interrupciones”, no es estrictamente Markoviana. El hecho de que la cadena pueda converger a la verdadera distribución posterior es debido, según Roberts y Rosenthal (2007), al hecho de que conforme crece el número de muestras de la cadena, la diferencia entre sucesivas adaptaciones de la matriz de covarianza es cada vez menor. En cualquier caso, limitando la adaptación al periodo de calentamiento de la cadena (cuyas muestras se desechan), las siguientes muestras ya en el equilibrio serán completamente Markovianas.

Los **métodos MCMC de cadena múltiple** simulan diferentes trayectorias, explorando simultáneamente el espacio de parámetros. Dichos métodos surgieron ante la dificultad con la que se encontraban los métodos de cadena simple al explorar complicadas distribuciones objetivo, con múltiples modas, con largas colas o muy asimétricas. La experiencia ha mostrado que con una sola cadena, la eficiencia en la exploración decrece considerablemente en espacios paramétricos de muy alta dimensión, en el que pueden aparecer más fácilmente varias regiones de atracción (varias modas) y muchos más mínimos locales. Es decir, cuando el número de parámetros comienza a ser considerablemente alto, los algoritmos MCMC de cadena simple muestran problemas de convergencia similares a los que afectaban a los métodos de optimización. Por ejemplo, Vrugt et al. (2009a) muestran el mal desempeño del algoritmo DRAM (Haario et al., 2006), de los más avanzados de cadena simple, en la inferencia de un modelo hidrológico con 13 parámetros. Dado que, en estos métodos, cada cadena comienza su trayectoria en puntos diferentes del espacio de parámetros y alejados entre sí, se obtiene una protección robusta ante la convergencia prematura, puesto que la solución de convergencia debe obviamente, ser la misma para todas las cadenas. En relación con la convergencia del algoritmo MCMC, en los métodos de cadena simple los test de convergencia son herramientas que solo sirven para evidenciar la no convergencia y no dan pruebas robustas que aseguren que la cadena esté muestreando la verdadera distribución posterior del parámetro. Sin embargo, con los métodos de cadena múltiple se dispone de más herramientas y más poderosas (Brooks y Gelman, 1998; Brooks y Giudici, 2000; Gelman y Rubin, 1992), para evaluar estadísticamente que el conjunto de cadenas está muestreando simultáneamente la verdadera posterior.

Ter Braak (2006) presentó un innovador, al tiempo que simple, algoritmo MCMC adaptativo llamado DE-MC (del inglés Differential evolution Markov Chain). Dicho algoritmo combina los algoritmos de tipo RWM adaptativos con la heurística de los algoritmos genéticos. Es decir, el algoritmo evoluciona varias cadenas de Markov en paralelo, generando una población de valores de los parámetros mediante evolución diferencial (Price et al., 2005; Storn y Price, 1997), a partir de los valores generados y aceptados previamente. Además, emplea la regla de Metropolis para decidir cuáles de los nuevos puntos propuestos en cada generación deberían reemplazar a sus padres. El algoritmo es capaz de saltar de forma directa entre diferentes modas inconexas de la posterior (Ter Braak, 2006;

Ter Braak y Vrugt, 2008; Vrugt et al., 2008) lo cual resulta ventajoso para el eficiente muestreo de las distribuciones posterior multimodales. Así mismo, permite ir ajustando tanto la orientación como la escala de la distribución de salto, adaptándola progresivamente a la de la distribución posterior. En el trabajo de (Vrugt et al., 2009a) se propone el algoritmo DREAM (siglas de Differential Evolution Adaptive Metropolis) basado en DE-MC con la inclusión de mejoras, como el muestreo en subespacios aleatorios (e.g. Metropolis-within-Gibbs), que aceleran considerablemente la convergencia a la distribución posterior. Recientemente, aparecen en la literatura sucesivos trabajos que van proponiendo una familia de algoritmos cuyo núcleo central es el algoritmo DREAM y que incluyen mejoras o variaciones sobre el algoritmo principal (Vrugt, 2016). Uno de estos algoritmos MCMC es el **DREAM_ZS** (Laloy y Vrugt, 2012; Schoups y Vrugt, 2010; Vrugt et al., 2009b), el cual ha servido de herramienta para la inferencia de los parámetros en la presente Tesis Doctoral.

2.2.3. Incertidumbre estructural mediante la aproximación Bayesiana

Hasta ahora, todo lo visto en el presente capítulo sobre inferencia se ha centrado en un solo modelo. Por tanto, los resultados sobre incertidumbre de la predicción que de ello se derivasen, no tendrían carácter general, sino que estarían condicionados al empleo de ese único modelo, con sus defectos y virtudes.

Como se mencionó en la introducción de este capítulo, los modelos deben cumplir con el Principio de Parquedad y al tiempo, deben obtener el más fiable (exacto y preciso) ajuste a las observaciones. Estos dos requerimientos antagonicos, son también referidos en la literatura como el dilema entre el sesgo (falta de exactitud) y la varianza (falta de precisión) de la predicción (Geman et al., 1992). Cumplir con ambas premisas, simplicidad y fiabilidad, implica la elección, de entre varios candidatos, del modelo que minimiza el error cuadrático total de la predicción (varianza más sesgo²).

En las dos últimas décadas, han sido muchos los autores que, reconociendo la incertidumbre conceptual (estructural, epistémica) del modelo, como la principal contribuyente a la incertidumbre predictiva total, han empleado herramientas para considerarla como parte integral del proceso de inferencia (Ajami et al., 2007; Burnham et al., 2011; Burnham y Anderson, 2002; Clark et al., 2011; Gupta et al., 2012; Refsgaard et al., 2006; Renard et al., 2010; Vrugt y Robinson, 2007b; Wöhling y Vrugt, 2008). Las diferentes herramientas para considerar la

incertidumbre estructural se agrupan en tres grupos. Por una parte, están los **métodos de selección** y los **métodos de promediado** de modelos, métodos que consideran que los modelos ya tienen estimados sus parámetros de forma adecuada. Por otra parte, el tercero de los métodos es el llamado **método desagregado** para el tratamiento de la incertidumbre (Kuczera et al., 2007; Renard et al., 2011), en el cual se considera por separado la incertidumbre de diferentes componentes del modelo, a los cuales se les acopla ciertas variables latentes de carácter estocástico, las cuales aportan aleatoriedad (incertidumbre) a esos componentes que a priori eran deterministas. La presente sección, trata sobre los dos primeros métodos. Mientras los métodos de selección tratan de quedarse con el mejor modelo de entre un grupo de modelos candidatos, los métodos de promediado se basan en emplear todos los modelos candidatos de forma ponderada. La forma de ponderar los modelos es la que genera las diversas variantes de estos métodos, conocidos como métodos Bayesianos de promediado de modelos (**BMA** en inglés).

Entre los criterios más conocidos para la selección de modelos, cabe destacar los basados en los clásicos test de hipótesis (e.g. test de las ratios de verosimilitud) y los basados en la teoría de la información, como el criterio de información de Akaike (AIC) y sus variantes, o el Bayesian information criterion (BIC). Para más detalles se pueden consultar Banks et al. (2014), Sun y Sun (2015), o Wilks (2011) entre muchos otros.

En el ámbito de la investigación científica, se extiende cada vez más la idea de evitar la selección del “mejor modelo”, en favor de la utilización de diversos modelos (lo más estructuralmente diferentes posible) promediándolos mediante las técnicas BMA. Según Burnham y Anderson (2002), “*la modelación es tan solo un ejercicio de aproximación al contenido de información explicable que hay en los datos*”. Diferentes estructuras de modelo serán capaces de extraer más información de los datos, que un solo modelo con sus limitaciones. Por otra parte, la utilización de varios modelos de forma ponderada parece estar de acuerdo con el hecho de que intentar optimizar la simplicidad y la fiabilidad de un modelo, al ser un problema multiobjetivo, teóricamente no puede tener una única solución. Los métodos BMA tienen base probabilística Bayesiana, e implícitamente aplican el criterio de la *navaja de Occam*. El principal problema de estos métodos es de índole computacional, ya que requieren evaluar la evidencia de los datos observados, también llamada evidencia Bayesiana del modelo, según cada modelo considerado. Recordando la expresión de la

posterior de los parámetros (ecuación (2.2)), la evidencia de los datos es el término del denominador $P(\tilde{\mathbf{Y}})$, el cual no se necesita evaluar en la inferencia de los parámetros (en la que es suficiente emplear la posterior no escalada), pero es imprescindible en la evaluación (o ponderación) de las diferentes estructuras de modelo (hipótesis) candidatas. Su definición matemática viene dada por la ecuación (2.3), que evalúa la media de las verosimilitudes, calculadas sobre un gran número de valores muestreados de los parámetros.

Puesto que la presente Tesis Doctoral se centra en la evaluación de la incertidumbre de los parámetros y de la predicción, condicionados a un solo modelo, no se profundizará más en estos temas. Baste decir que: i) La evaluación de $P(\tilde{\mathbf{Y}})$ de una forma eficiente es todavía un tema no resuelto, objeto de numerosas contribuciones en la literatura reciente y ii) La elección del método de evaluación influye en el resultado de la ponderación de los modelos. Schöniger et al. (2014) presentan un minucioso trabajo de comparación de nueve metodologías de evaluación de la evidencia Bayesiana del modelo.

2.3. Formulación del problema de inferencia en modelación hidrológica

Considérese de nuevo, el vector de observaciones de una variable de interés $\tilde{\mathbf{Y}} = \{\tilde{y}_1, \dots, \tilde{y}_n\}$, que fue definido en la sección 2.2.1. Dicho vector resume la respuesta de un sistema natural analizado, llámese \mathfrak{S} , frente a la excitación producida por k señales o inputs variables en el tiempo, $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, donde los vectores columna se definen para cada t como $\mathbf{b}_t = \{b_{t_1}, \dots, b_{t_k}\}$. Los inputs son el equivalente a los regresores, en un análisis clásico de regresión. Para explicar el sistema analizado, en base a los datos observados sobre él, se emplea un modelo matemático computacional, llámese $\mathcal{M}(\cdot)$. De forma simplificada, se puede resumir el proceso de análisis mediante la ecuación (2.9)

$$\tilde{\mathbf{Y}} \leftarrow \mathcal{M}(\boldsymbol{\theta}, \mathbf{x}_0, \tilde{\mathbf{B}}) + \mathbf{E} \quad (2.9)$$

en la que $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_d\}$ es el $d \times 1$ -vector de parámetros del modelo, \mathbf{x}_0 es el vector que almacena los valores de las variables de estado del sistema, al comienzo de la simulación ($t=0$), $\tilde{\mathbf{B}}$ es la matriz de dimensiones $k \times n$ que contiene las mediciones en el tiempo, de los k inputs que excitan el sistema y $\mathbf{E} = \{e_1, \dots, e_n\}$ es un vector de dimensión n con errores aditivos. Si bien la hipótesis de errores aditivos es la más común, también se podría manejar la hipótesis de errores multiplicativos mediante la transformación logarítmica de las variables de la ecuación (2.9) (Beven, 2009).

En según qué publicaciones de la literatura, se puede observar que, a los errores, tal como se han definido, se les llama residuos. Sin embargo, en este documento se ha convenido en llamar residuos (o indistintamente innovaciones) a aquellos errores que son exclusivamente generados por un proceso estocástico de tipo **ruido blanco** (independientes, idénticamente distribuidos y de media nula). Entonces, el vector de errores depende del modelo $\mathcal{M}(\cdot)$ que se ha asumido, depende de los valores de sus parámetros, de los estados iniciales de las variables de estado del sistema y naturalmente, de los inputs que excitan o “fuerzan” al modelo. Esto queda expresado por la ecuación (2.10).

$$\mathbf{E}(\boldsymbol{\theta}, \mathbf{x}_0, \tilde{\mathbf{B}}) = \tilde{\mathbf{Y}} - \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \mathbf{x}_0, \tilde{\mathbf{B}}) \quad (2.10)$$

donde, $\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \mathbf{x}_0, \tilde{\mathbf{B}}) = \{y_{1,\mathcal{M}}, \dots, y_{n,\mathcal{M}}\}$ son los resultados de la simulación que el modelo proporciona para cada t sobre la variable de interés \mathbf{Y} .

2.3.1. Inferencia en el caso ideal. Método de los mínimos cuadrados

Supóngase por conveniencia para la exposición de este apartado, que el modelo asumido $\mathcal{M}(\cdot)$ es perfecto y que los datos relativos a los inputs del modelo presentan una calidad excelente, siendo por tanto muy fiables. Supóngase también, que los parámetros del modelo estimados a priori, si los hubiere, presentan valores adecuados. Y finalmente, considérese que los estados iniciales del sistema adoptan los valores correctos, como producto de un periodo de *calentamiento* del modelo, de longitud suficiente. En este caso, que esta

Tesis denomina **caso ideal**, los errores producidos diferencia entre las observaciones y la simulación que proporciona el modelo, son únicamente debidos a los errores de medición en las observaciones, cuya estructura comúnmente aceptada es básicamente aleatoria de tipo *ruido blanco* (i.i.d. y con media nula), aunque ocasionalmente puedan presentar algún sesgo o bias, de fácil identificación y eliminación. El lector puede darse cuenta de que, en resumen, la hipótesis de este caso ideal de modelación es que la única fuente de incertidumbre procede de la medición de los observados. Esta es una fuerte hipótesis, mediante la cual se está dando una forma muy particular a la estructura de error. Siendo este el caso, el vector de errores se puede reescribir según la ecuación (2.11)

$$\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}) = \{e_1(\boldsymbol{\theta}), \dots, e_n(\boldsymbol{\theta})\} \quad (2.11)$$

Esta ecuación necesita de un proceso de inferencia, con el fin de determinar el vector de parámetros $\boldsymbol{\theta}$. Al ser *ruido blanco*, los residuos son serialmente independientes, con lo que su verosimilitud conjunta $\ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$ se puede escribir según la ecuación (2.12)

$$\ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) = P(e_1(\boldsymbol{\theta})) \times \dots \times P(e_n(\boldsymbol{\theta})) = \prod_{j=1}^n P(e_j(\boldsymbol{\theta})) \quad (2.12)$$

donde $P(\bullet)$ es una función de densidad de probabilidad (pdf) con media cero, evaluada en \bullet . En el marco del caso ideal tratado, las n funciones de densidad son idénticas, siendo por tanto la desviación típica de todas y cada una de ellas la misma σ . A esta propiedad se le llama **homocedasticidad**, del griego *homos* (igual) y *skedasis* (dispersión). En contraposición, se define el término **heterocedasticidad** para denominar al caso en el cual las distintas pdf presentan una desviación típica diferente. Con el fin de evaluar la verosimilitud conjunta de los residuos, es necesario asumir una determinada función $P(\bullet)$. La decisión más habitual, en el marco del caso ideal, es asumir que los residuos se distribuyen según una distribución Gaussiana, lo que se puede escribir como $e_j(\boldsymbol{\theta}) \stackrel{\mathcal{D}}{\sim} \mathcal{N}(0, \sigma^2) \quad \forall j \in \{1, \dots, n\}$.

Con esta última hipótesis, queda completamente definido el modelo de error que corresponde al caso ideal de modelación analizado. El único parámetro de este modelo de error es la varianza de los errores σ^2 . Si se sustituye $\mathcal{N}(0, \sigma^2)$ en la ecuación (2.12), la verosimilitud conjunta de los n residuos, o lo que es equivalente, de los parámetros que los generan, se puede evaluar según la ecuación (2.13)

$$\ell(\boldsymbol{\theta} | \tilde{\mathbf{Y}}, \sigma^2) = \prod_{j=1}^n (2\pi)^{-\frac{1}{2}} \frac{1}{\sigma} \exp\left(-\frac{1}{2} \left(\frac{\tilde{y}_j - y_j(\boldsymbol{\theta})}{\sigma}\right)^2\right) \quad (2.13)$$

Por razones de estabilidad numérica en la evaluación de la ecuación (2.13), la cual puede resultar crítica cuando el ajuste del modelo a las observaciones es muy deficiente, es decir, cuando $|e_t(\boldsymbol{\theta})| \gg \sigma$ para muchos valores e_t , resulta común trabajar con el logaritmo de la verosimilitud. De esta manera se llega a la ecuación (2.14)

$$\mathcal{L}(\boldsymbol{\theta} | \tilde{\mathbf{Y}}, \sigma^2) = -\frac{1}{2} n \log(2\pi) - \frac{1}{2} n \log(\sigma^2) - \frac{1}{2\sigma^2} \text{SRC}(\boldsymbol{\theta}) \quad (2.14)$$

donde **SRC** es la abreviatura de *Suma de Residuos Cuadrados*. La función

$\text{SRC}(\boldsymbol{\theta}) = \sum_{j=1}^n (\tilde{y}_j - y_j(\boldsymbol{\theta}))^2 = \sum_{j=1}^n (e_j(\boldsymbol{\theta}))^2$ es usada comúnmente en la calibración de

modelos como función objetivo a minimizar, mediante los clásicos métodos de optimización, dando lugar al conocido y ampliamente utilizado *método de los mínimos cuadrados* (en adelante **SLS**), el cual fue descubierto de modo independiente y de forma empírica por Legendre (1805) y Gauss (1795?, 1809), si bien adquirió sus bases estadísticas teóricas gracias al teorema de Gauss-Markov (1929). Según el método SLS, en vista de la ecuación (2.14), cuanto mejor se ajusta el modelo a los datos observados, mayor es el valor de la verosimilitud de los parámetros, o equivalentemente, menor es el valor de la función $\text{SRC}(\boldsymbol{\theta})$, es decir, más pequeños son los residuos (cuadráticos). Por tanto, se observa que, en el caso ideal de modelación, se da la equivalencia entre la estimación de parámetros por el método de máxima verosimilitud (en adelante **ML**) y su estimación mediante el método SLS.

A modo de recordatorio para el lector, las condiciones que han de cumplir los errores (o residuos en este caso) para la equivalencia de ambos métodos de estimación de parámetros son: i) Los errores deben ser variables aleatorias independientes; ii) Deben tener media nula; iii) Los errores deben estar idénticamente distribuidos, lo que implica que han de ser homocedásticos y iv) La distribución de errores debe ser Gaussiana. En tres palabras: **ruido blanco Gaussiano**.

Con carácter general, cualquier parámetro del modelo de error (o del modelo hidrológico) se puede estimar a priori dejándolo fuera del proceso de calibración automática. Lo correcto cuando un parámetro puede interactuar durante el proceso de inferencia, con uno o más de los demás parámetros, es inferirlos conjuntamente (Evin et al., 2014). Por otra parte, es habitual considerar que los parámetros del modelo de error (e.g. σ) son **parámetros nuisance** en la inferencia, lo que significa que en sí mismos no tienen gran interés para el objetivo principal de la modelación, aunque son necesarios para el modelo probabilístico y por tanto, para el cálculo de la función de verosimilitud. En este caso ideal, no es necesario inferir σ para poder inferir θ , puesto que ambos no interactúan durante la inferencia. Esto se puede deducir a través de la expresión no normalizada de la ecuación (2.14). Esta expresión se obtiene mediante la sustitución de σ^2 en (2.14), por el estimador muestral insesgado de la varianza poblacional $s^2 = (n-1)^{-1} \text{SRC}(\theta)$. En la expresión resultante, se eliminan los términos que no dependen de θ y el resultado es la ecuación (2.15)

$$\mathcal{L}(\theta | \tilde{\mathbf{Y}}) \propto -\frac{n}{2} \log(\text{SRC}(\theta)) \quad (2.15)$$

De esta ecuación se deriva que la verosimilitud de los parámetros dadas las observaciones $\ell(\theta | \tilde{\mathbf{Y}})$, es inversamente proporcional a $\text{SRC}(\theta)^{n/2}$.

Con esto, quedan expuestas cuales son las hipótesis que dan validez a la inferencia mediante el método SLS.

2.3.2. Inferencia en los casos no ideales

En el anterior apartado, se han mostrado las condiciones que han de cumplir los errores cuando se adopta la hipótesis de que la única fuente de incertidumbre que afecta a la modelación reside en los errores de observación (o de medida) de la variable de interés. Bajo dicha hipótesis, el método de estimación de parámetros SLS queda justificado, pues como se ha mostrado, su criterio de optimización coincide exactamente con el criterio del método de estimación probabilística de máxima verosimilitud: la minimización de $SRC(\theta)$.

El caso de los modelos hidrológicos

El método SLS ha sido, y probablemente sigue siendo, el método de estimación de parámetros más utilizado en modelación hidrológica. Sin embargo, como ya se avanzó anteriormente, es bien conocido desde hace tiempo (Kuczera, 1983; Sorooshian y Dracup, 1980) que los errores generados por los modelos hidrológicos no son en general un *ruido blanco* Gaussiano.

La dependencia serial de los errores se debe principalmente a problemas estructurales del modelo (Honti et al., 2013; Koskela et al., 2012). De hecho, el empleo de modelos de error autorregresivos es una forma de tomar en consideración de forma agregada, la parte estructural o sistemática (bias) de los errores (Reichert y Mieleitner, 2009). La autocorrelación en los errores es un indicador de que todavía contienen información (no son ruido) que el modelo no ha sido capaz de procesar y almacenar en sus parámetros, como resultado de su mala especificación e incapacidad para simular correctamente los procesos. Los errores en los inputs del modelo son el otro factor de importancia, que provoca el apartamiento de las hipótesis del método SLS.

Consecuencias de inferir con el método SLS

Inferir los parámetros de un modelo hidrológico con SLS conlleva, generalmente, tres importantes consecuencias (Kavetski et al., 2003; Schoups y Vrugt, 2010; Thyer et al., 2009). La primera consecuencia ocurre sobre el valor esperado de los parámetros del modelo hidrológico, el cual queda sesgado para compensar los problemas estructurales del modelo y/o los errores de los inputs, con el propósito de poder ajustarse lo más posible a las observaciones. En palabras de Kirchner (2006): “*Los modelos suelen ser marionetas matemáticas que bailan con la melodía de los datos de calibración*”.

El sesgo de los parámetros implica una pérdida de significado o coherencia conceptual de los mismos, que también sería pérdida de significado físico en el caso, poco probable, de que ya no lo hubiesen perdido debido a los efectos de inconmensurabilidad (Beven, 2009) o efectos de escala, los cuales son la causa de los llamados **parámetros efectivos**. Es importante tener en cuenta, que aunque los parámetros de los modelos conceptuales sean *efectivos*, no por ello deben sacrificar la coherencia hidrológica con la que fueron concebidos, en aras de conseguir un ajuste observado-simulado en calibración excepcionalmente bueno. El valor absoluto de los parámetros efectivos, generalmente, no va a proporcionar ninguna o poca información, pero sus magnitudes relativas deben seguir manteniendo una coherencia. En este sentido, esta Tesis considera como **“valor correcto”** de los parámetros, al valor efectivo con coherencia conceptual que tendrían los parámetros, en el caso ideal en el que la única fuente de incertidumbre que afecta a la inferencia fuese la medición de los observables. El efecto que produce tanto el SLS como cualquier modelo de error incorrecto, sobre los parámetros del modelo hidrológico, puede considerarse equivalente al efecto que se produce en los parámetros de un modelo del tipo “caja negra”: el modelo forzaría cualquier valor del parámetro, con la única consideración de que el ajuste de simulados y observados sea el óptimo.

La segunda de las consecuencias que produce SLS sobre los parámetros, es la infraestimación de su incertidumbre. El método SLS está obviando la existencia de errores diferentes al de observación, que son incluso de mayor magnitud. Esto lo hace forzando a que los parámetros estimados tengan la mínima varianza (incertidumbre) posible, sobreestimando por tanto la cantidad de información real que contienen almacenada en ellos. La confirmación de que esto ocurre se tiene al analizar los errores producidos tras la calibración. Generalmente, son errores con fuerte autocorrelación, lo que indica que contienen la información que el modelo hidrológico no ha sido capaz de procesar y almacenar en sus parámetros.

Finalmente, la tercera consecuencia que produce el SLS es la sobreestimación de la incertidumbre total de la predicción del modelo. Todas estas consecuencias serán empíricamente demostradas en el capítulo 5.

Cabe reseñar, por si al lector le quedan dudas, que cuando se habla de método SLS, se están incluyendo todos los métodos basados en minimizar una función

objetivo, relacionada con el estadístico $SRC(\theta)$. Póngase como ejemplo, por ampliamente utilizado como función objetivo en la calibración de modelos hidrológicos, el índice NSE (Nash y Sutcliffe, 1970). Cheng et al. (2014) ponen de manifiesto la equivalencia entre minimizar SRC y maximizar NSE, obteniendo la expresión $SRC(\theta) = (1 - NSE(\theta))(n-1)\sigma_Y^2$, siendo σ_Y^2 la varianza de las observaciones.

Soluciones adoptadas en la literatura para la correcta inferencia en modelación hidrológica

Conociendo las consecuencias que puede tener la adopción de SLS, como método de inferencia en modelación hidrológica, han sido varias las soluciones adoptadas en la literatura con el fin de eliminar o paliar en lo posible esas consecuencias. Dado que el problema está en las diversas fuentes de incertidumbre (o de error) que afectan a la modelación hidrológica, la solución está en afrontar la consideración de esos errores en el proceso de inferencia, mediante su modelación. Las estrategias empleadas para este fin son dos: El método agregado de modelación del error y el método desagregado. Estos últimos ya se mencionaron en la sección 2.2.3 al hablar sobre la estimación de la incertidumbre estructural de los modelos.

El **método desagregado** (Kavetski et al., 2006a, 2006b; Kuczera et al., 2006; Reichert y Mieleitner, 2009; Renard et al., 2010, 2011; Vrugt et al., 2008) trata de modelar la incertidumbre que proviene de las distintas fuentes de incertidumbre (estructural, de los Inputs, de las observaciones) de forma separada, con el propósito de evaluar cómo y cuanta es la contribución de cada fuente a la incertidumbre total. Aunque la idea es muy prometedora, aun resulta complicado identificar la contribución de las distintas fuentes a la incertidumbre total, a menos que se hagan hipótesis fuertes y subjetivas, o se obtengan observaciones adicionales de otras variables del sistema que permitan un condicionamiento más robusto y coherente de los parámetros (Montanari, 2011).

Por otra parte, el **método agregado**, en el cual se basa la investigación de esta Tesis, trata de modelar los errores del modelo hidrológico mediante la consideración, de forma agregada, de los efectos (bias, autocorrelación y heterocedasticidad) que en ellos producen todas las fuentes de incertidumbre. Por tanto, el método no pretende discernir en qué medida contribuye cada fuente

de incertidumbre al error total. La ejecución del método requiere la inclusión de todo el conocimiento que se tenga sobre los errores de modelación, en una función de verosimilitud adecuada, lo cual puede ser una tarea no sencilla (Smith et al., 2015). Si se cuenta con dicha función de verosimilitud y con el número de observaciones necesarias para determinar las desconocidas cantidades del modelo (los parámetros), condición denominada en Gelb (2001) como **observabilidad del sistema**, será entonces posible una correcta estimación de los parámetros hidrológicos (Sorooshian y Gupta, 1983).

Muy importante es la siguiente consideración. Como se ha visto anteriormente, con el método SLS solamente es necesario estimar los parámetros del modelo hidrológico, puesto que el único parámetro del modelo de error, la varianza de los errores, puede ser estimado por separado al no interaccionar con los parámetros del modelo hidrológico. Sin embargo, en los modelos de error que se han de plantear para tomar en consideración los efectos de todas las fuentes de incertidumbre, aparecen nuevos parámetros e hiperparámetros (parámetros de los estadísticos), que no pueden inferirse de forma independiente debido a sus posibles interacciones con los parámetros del modelo hidrológico. Por tanto, la inferencia del modelo hidrológico y del modelo de error debe ser una **inferencia conjunta**. De no hacerse así, los parámetros estimados quedarán sesgados (Evin et al., 2014). A pesar de ello, es difícil encontrar en la literatura ejemplos de inferencia conjunta de los parámetros de los modelos hidrológico y de error, por dos causas principales: i) La mayor parte de trabajos emplean el método SLS y ii) Los trabajos recientes con modelos de error más sofisticados que el SLS, se encuentran con problemas de convergencia y/o ciertos problemas espurios, que les hacen optar por otras alternativas a la inferencia conjunta.

Las aproximaciones al método agregado que se han propuesto en la literatura, han ido relajando las hipótesis sobre los errores en las que se apoya el método SLS, tal y como se explica en lo que sigue.

2.3.3. Relajación de la hipótesis de homocedasticidad de los errores

La aproximación más común en la literatura para el tratamiento de la heterocedasticidad de los errores ha sido el empleo de métodos en los que previamente a la inferencia, se hace una transformación de los datos (Box y Cox, 1964; Sakia, 1992). Así, mediante la apropiada transformación previa de observados y simulados, se obtiene, en la mayoría de los casos, la estabilización

de la varianza de los errores e incluso una reducción significativa de la asimetría de su distribución (si la hubiere), tras lo cual, la inferencia se lleva a cabo con el tradicional SLS. Cabe considerar que el problema de la posible distribución de errores con colas *pesadas*, es difícilmente corregido por las citadas transformaciones (Bates y Campbell, 2001; Yang et al., 2007).

Los **métodos transformacionales** han sido los más utilizados en la literatura hidrológica (Bates y Campbell, 2001; Del Giudice et al., 2013; Kuczera, 1983; Reichert y Mieleitner, 2009; Vrugt et al., 2009b, etc.). Un inconveniente de estos métodos es la obtención de parámetros con unos valores descontextualizados, aunque esto solo es un inconveniente en caso de querer comparar diferentes casos de estudio con y sin transformaciones, o en el caso de querer regionalizar los parámetros. Por otra parte, eliminar el problema mediante una transformación no ayuda a entenderlo, lo cual impide la mejora del conocimiento (Evin et al., 2013; Schoups y Vrugt, 2010) y por ende, la mejora de los modelos.

Por otra parte están los **métodos directos**, los cuales en lugar de evitar el problema de una varianza no constante de los errores, tratan de modelarlo. Su aplicación es, en principio, muy sencilla. Considérese la ecuación (2.13) que evalúa la verosimilitud conjunta, de los n errores del periodo analizado. Si se asume que las n distribuciones condicionales de los errores son Gaussianas pero con varianza no constante, esto es $\mathcal{N}(0, \sigma^2)$, siendo $\sigma^2 = \{\sigma_{e|y_{1,M}}^2, \dots, \sigma_{e|y_{n,M}}^2\}$ el vector que almacena el valor de esas varianzas, entonces la Log-verosimilitud queda formulada según la ecuación (2.16)

$$\mathcal{L}(\boldsymbol{\theta} | \tilde{\mathbf{Y}}, \boldsymbol{\sigma}^2) = -\frac{1}{2} n \log(2\pi) - \sum_{j=1}^n \left\{ \log(\sigma_{e|y_{j,M}}) \right\} - \frac{1}{2} \text{SRC}_w \quad (2.16)$$

donde $\text{SRC}_w = \sum_{j=1}^n (w_j e_j(\boldsymbol{\theta}))^2$ es la suma cuadrática de los errores ponderados,

siendo $w_j = \sigma_{e|y_{j,M}}^{-1}$ los coeficientes de ponderación de los errores. Como puede observarse, y en general es así, cada error se pondera de forma inversamente proporcional a la magnitud de su varianza (incertidumbre) condicional. Es decir, a diferencia del método SLS en el que a todas las observaciones (o a sus errores) se les atribuye la misma influencia sobre la función objetivo, según la

ecuación (2.16), tienen más importancia las observaciones con menos incertidumbre, cuyo error presenta menos varianza. Por lo demás, el método es similar al SLS. Este método es conocido como el método de los mínimos cuadrados ponderados (**WLS** en inglés).

Generalmente, el cálculo de estos pesos nunca ha sido un problema en inferencias donde la repetición de mediciones, en igualdad de condiciones del experimento, es posible. Desafortunadamente, en modelación hidrológica (y en general en modelación ambiental) los problemas de inferencia se enfrentan con datos seriales observados sin la replicación de medidas, ya que la Naturaleza no ofrece esa oportunidad. Este fue el motivo que propició el enorme éxito de los métodos transformacionales de estabilización de la varianza, anteriormente explicados. Mediante la transformación, ya no es necesario estimar estos pesos. Otra alternativa menos extendida, fue la estimación de los pesos mediante reglas empíricas, muchas veces con cierto carácter subjetivo y pérdida de rigor estadístico.

El método híbrido de Sorooshian y Dracup

Sorooshian y Dracup (1980) fueron los primeros que intentaron salvar la dificultad de obtener los pesos de los errores, sin el empleo de un método de transformación y procurando no entrar en conflicto con los métodos objetivos de la teoría probabilística. Su método, siendo en espíritu un método directo para estabilizar la varianza, se apoya sin embargo en los métodos de transformación, con lo que podría catalogarse como un método a medio camino, o **método híbrido** entre los métodos directos y los transformacionales. La justificación del método se expondrá brevemente en lo que sigue, no porque tenga demasiado interés por sí misma, sino por conveniencia para mostrar que el resultado al que llega este método híbrido, tiene propiedades similares a la metodología directa propuesta en esta Tesis.

El método de Sorooshian y Dracup (1980) se basa en las siguientes hipótesis. Primero, considera que las varianzas condicionales de las observaciones $\tilde{\mathbf{Y}}$, son función de su valor esperado, el cual es precisamente el valor simulado por el modelo $\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})$, dado que los errores se consideran sin sesgo. Por tanto, se puede escribir $\mathbb{V}[\tilde{\mathbf{Y}} | \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})] = f(\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}))$ o lo que es lo mismo, escrita en función de los errores, $\boldsymbol{\sigma}^2 = \mathbb{V}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})] = f(\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}))$

Como las varianzas σ^2 no son constantes, hay que aplicar una transformación apropiada $g(\cdot)$ para estabilizarlas. Dicha transformación, a priori desconocida, pertenece a la familia de las transformaciones paramétricas propuestas por Box y Cox (1964). Aplicando la transformación $g(\cdot)$ a la ecuación (2.11) se llega a la ecuación (2.17)

$$\mathbf{E}^* = g(\tilde{\mathbf{Y}}; \lambda) - g(\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}); \lambda) \quad (2.17)$$

donde λ es el parámetro de la transformación Box-Cox y el asterisco indica que es una variable derivada de la transformación. Si la transformación es la adecuada, los errores \mathbf{E}^* deben tener varianza constante σ^{2*} , y por tanto se cumplirá que la varianza marginal de los errores transformados y sus varianzas condicionales tendrán el mismo valor (homocedasticidad). En forma de ecuación $\sigma^{2*} = \mathbb{V}[\mathbf{E}^* | \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})] = \mathbb{V}[\mathbf{E}^*]$. Sin embargo, la idea de Sorooshian y Dracup (1980) no era utilizar el método transformacional, sino obtener una expresión directa del tipo $\sigma^2 = f(\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}))$ para la varianza de los errores no transformados. Para llegar a ese objetivo, se aplica la expansión en serie de Taylor de primer orden a la ecuación (2.17) para obtener la ecuación (2.18)

$$\begin{aligned} \mathbf{E}^* &= g'(\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}))(\tilde{\mathbf{Y}} - \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})) + \mathcal{O}\left(\left(\tilde{\mathbf{Y}} - \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})\right)^2\right) = \\ &= g'(\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}))\mathbf{E}(\boldsymbol{\theta}) + \mathcal{O}\left(\left(\mathbf{E}(\boldsymbol{\theta})\right)^2\right) \end{aligned} \quad (2.18)$$

Atendiendo a la ecuación (2.18), la varianza de los errores transformados se puede escribir según la ecuación (2.19)

$$\sigma^{2*} = \mathbb{V}[\mathbf{E}^*] \approx \left(g'(\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}))\right)^2 \mathbb{V}[\mathbf{E}(\boldsymbol{\theta})] = \left(\frac{d\mathbf{Y}_{\mathcal{M}}^*}{d\mathbf{Y}_{\mathcal{M}}}\right)^2 \sigma^2 \quad (2.19)$$

Teniendo en consideración, que según la transformación Box-Cox, el vector de simulaciones en función de sus respectivas transformadas se expresa como $\mathbf{Y}_{\mathcal{M}} = (\lambda \mathbf{Y}_{\mathcal{M}}^* + 1)^{1/\lambda}$, y resolviendo la derivada en (2.19), Sorooshian y Dracup

(1980) obtuvieron la expresión que buscaban, representada por la ecuación (2.20)

$$\sigma^2 = \sigma^{2*} \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})^{2(1-\lambda)} \quad (2.20)$$

Con la ecuación (2.20) se puede calcular cada una de las varianzas condicionales de los errores a partir de las simulaciones del modelo $\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})$, de un factor λ que permite una transformación estabilizadora de la varianza de los errores y del correspondiente valor de esa varianza estabilizada σ^{2*} . Considerando las desviaciones estándar en lugar de las varianzas, se puede escribir la ecuación (2.21)

$$\sigma = \sigma^* \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})^{1-\lambda} \quad (2.21)$$

Calculadas las desviaciones condicionales σ , es posible calcular los pesos de cada observación como $w_j = \sigma_{elY_{j,\mathcal{M}}}^{-1} = \sigma^{*-1} y_{j,\mathcal{M}}^{\lambda-1} \quad \forall j \in \{1, \dots, n\}$, con el fin de ser introducidos en la función de verosimilitud de la ecuación (2.16). La función de verosimilitud $\mathcal{L}(\boldsymbol{\theta}, \lambda | \tilde{\mathbf{Y}})$ incorpora en este caso el nuevo parámetro Box-Cox.

Las ecuaciones (2.20) o (2.21) resultan de trascendental importancia, puesto que son el primer método directo con el que se trató el problema de la heterocedasticidad en modelación hidrológica. La ecuación (2.21) se puede escribir de una forma más general como

$$\sigma = \alpha + \kappa \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta})^\psi \quad (2.22)$$

donde en este caso, $\alpha = 0$, $\kappa = \sigma^*$ y $\psi = 1 - \lambda$.

Después del trabajo de Sorooshian y Dracup (1980), no se volverá a utilizar el método directo en investigación hidrológica, hasta el trabajo de Schoups y Vrugt (2010). Sin embargo, en ese y sucesivos trabajos se omite una importante cuestión: los parámetros κ y ψ de la ecuación (2.22) estimados por Sorooshian y Dracup (1980), si bien son parámetros de un método directo para el tratamiento de la varianza de los errores, fueron inferidos a partir de las variables transformadas (método híbrido). Esta particular forma de proceder, dota a los

parámetros inferidos de unas características que no tendrían si se hubieran inferido a partir de las variables originales. Quédele el lector en la memoria con la ecuación (2.23), que relaciona el parámetro $\kappa = \sigma^*$ con la varianza marginal de los errores transformados.

$$\kappa^2 = \sigma^{2*} = \mathbb{V}[\mathbf{E}^*] \quad (2.23)$$

En los posteriores trabajos que aplican el método directo desde 2010, se asume un modelo de varianza similar al de la ecuación (2.22). Sin embargo, y esta es la cuestión clave, sus parámetros se infieren a partir de los datos sin transformar, a diferencia del método híbrido de 1980. Esta circunstancia dará lugar a una serie de problemas reportados por diversos autores, en la literatura de los últimos años, entre otros Evin et al. (2013, 2014), Scharnagl et al. (2015) y Schoups y Vrugt (2010). En la propuesta de solución de esos problemas, basa esta Tesis su principal aportación.

2.3.4. Relajación de la hipótesis de independencia de los errores

Supóngase que los errores definidos por la ecuación (2.11) muestran correlación serial, también llamada autocorrelación. Este efecto es debido a errores en la especificación del modelo, que provocan que la información que el modelo no procesa adecuadamente, se vaya propagando por sus elementos de almacenamiento, hasta quedar finalmente repartida en los errores a lo largo de los diferentes pasos de tiempo de la simulación (Kavetski et al., 2003). La autocorrelación es, de hecho, una forma de sesgo en los errores (Reichert y Mieleitner, 2009). Dada una serie estocástica estacionaria y ergódica de errores, la autocorrelación entre dos errores solo depende de la diferencia entre sus posiciones en la serie, llamada desfase (*lag*), y no de su posición absoluta en dicha serie. Una herramienta importante para el análisis de la correlación entre errores es la llamada función de autocorrelación, en la cual se grafica el valor de correlación lineal que presentan los elementos de la serie de errores, en función del desfase entre dichos elementos. A la correlación lineal entre elementos que tienen un desfase K , se le denomina correlación lineal de orden K .

Si bien el cuerpo teórico que hay detrás del análisis de series temporales es amplio y diverso (Box et al., 1994; Brockwell, 2010; De Gooijer y Hyndman, 2006; Koutsoyiannis y Montanari, 2007; Prado y West, 2010; Shumway y Stoffer,

2011; Young, 2011), lo más habitual en la literatura es encontrar que la modelación de la dependencia serial de los errores, se realiza mediante modelos lineales autorregresivos (AR) de orden 1. Algunas excepciones se pueden encontrar, en las cuales se emplean modelos AR de orden 2 (Datta y Bolisetti, 2013) o incluso modelos AR que incluyen dependencia respecto de variables exógenas (ARX) (Pianosi y Raso, 2012).

Una vez definido el modelo de dependencia de los errores, este puede ser incluido en la función de verosimilitud, relajando así la hipótesis de independencia del caso ideal. Supóngase el caso común de asumir un modelo autorregresivo lineal de orden 1 para modelar la dependencia de los errores, como el de la ecuación (2.24)

$$e_j(\boldsymbol{\theta}) = \phi_1 e_{j-1}(\boldsymbol{\theta}) + z_j(\boldsymbol{\theta}, \phi_1) \quad (2.24)$$

donde los errores son homocedásticos con varianza σ^2 , se distribuyen según $e_j(\boldsymbol{\theta}) \stackrel{\mathcal{D}}{\sim} \mathcal{N}(0, \sigma^2) \quad \forall j \in \{1, \dots, n\}$ y su decorrelación produce las innovaciones o residuos $z_j(\boldsymbol{\theta}, \phi_1)$, que son un *ruido blanco* con varianza $\sigma_z^2 = \sigma^2(1 - \phi_1^2)$ y distribución $z_j(\boldsymbol{\theta}, \phi_1) \stackrel{\mathcal{D}}{\sim} \mathcal{N}(0, \sigma_z^2) \quad \forall j \in \{1, \dots, n\}$.

Siguiendo el mismo razonamiento que en casos anteriores, obtener la función de verosimilitud de los parámetros implica calcular la expresión dada por la ecuación (2.25)

$$\begin{aligned} \ell(\boldsymbol{\theta}, \phi_1 | \tilde{\mathbf{Y}}) &= P(z_1(\boldsymbol{\theta}, \phi_1)) \times \dots \times P(z_n(\boldsymbol{\theta}, \phi_1)) = \\ &= \prod_{j=1}^n P(z_j(\boldsymbol{\theta}, \phi_1)) = \prod_{j=1}^n P(e(\boldsymbol{\theta})_j - \phi_1 e(\boldsymbol{\theta})_{j-1}) \end{aligned} \quad (2.25)$$

Sin embargo, puesto que no se conoce el valor $e(\boldsymbol{\theta})_0$ (para $j=1$), la ecuación (2.25) se calculará $\forall j \in \{2, \dots, n\}$, obteniendo el equivalente a la ecuación (2.14), en la ecuación (2.26)

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}, \phi_1 | \{\tilde{y}_2, \dots, \tilde{y}_n\}, \sigma^2) &= -\frac{1}{2}(n-1)\log(2\pi) \\ &- \frac{1}{2}(n-1)\log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=2}^n (e_j(\boldsymbol{\theta}) - \phi_1 e_{j-1}(\boldsymbol{\theta}))^2 \end{aligned} \quad (2.26)$$

Para una serie de datos cuyo número de elementos n es de moderado a grande, lo habitual en modelación hidrológica, el cálculo de la ecuación (2.26) se puede extender para $j \in \{1, \dots, n\}$, siendo una buena aproximación la ecuación (2.27)

$$\mathcal{L}(\boldsymbol{\theta}, \phi_1 | \tilde{\mathbf{Y}}, \sigma^2) \approx -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n (z_j(\boldsymbol{\theta}, \phi_1))^2 \quad (2.27)$$

2.3.5. Función de verosimilitud generalizada (GL)

En las dos secciones anteriores, se ha mostrado como se generaliza parcialmente la función de verosimilitud correspondiente al método SLS, considerando por separado el caso de tener errores heterocedásticos y el caso de tener errores con dependencia serial. En la literatura relacionada anterior al año 2010, se observa la siguiente casuística de relajación de las hipótesis del SLS: i) Se considera por primera vez la heterocedasticidad de los errores, considerados independientes, mediante un método híbrido (Sorooshian y Dracup, 1980); ii) Se considera por primera vez mediante un AR(1), la dependencia serial de los errores, considerados homocedásticos (Sorooshian y Dracup, 1980); iii) Se considera la homocedasticidad de los errores, previa transformación de las variables y la independencia de los errores (Thiemann et al. (2001) entre otros) y iv) Se considera la homocedasticidad de los errores, previa transformación de las variables, y la dependencia de los errores se modela bien mediante un modelo AR (Bates y Campbell, 2001), bien mediante un modelo ARMA (Kuczera, 1983).

Schoups y Vrugt (2010) publican un trabajo en el cual se propone una función de verosimilitud de los parámetros, desarrollada a partir de un modelo de error en el cual se han relajado de forma simultánea las hipótesis del SLS, permitiendo así la modelación de errores heterocedásticos, serialmente dependientes y no Gaussianos. El modelo de error propuesto y por tanto la función de verosimilitud

generalizada (**GL**, del inglés General Likelihood), se basa en las siguientes hipótesis: i) El modelo de dependencia de los errores es un modelo lineal autorregresivo de primer orden AR(1), como el de la ecuación (2.24); ii) La heterocedasticidad se modela mediante un modelo de varianza lineal de dos parámetros $\{\alpha, \kappa\}$, similar al de la ecuación (2.22) pero asumiendo $\psi = 1$. Este modelo de varianza se aplica sobre las innovaciones y no sobre los errores, esto es $\sigma_z = \alpha + \kappa \mathbf{Y}_M(\boldsymbol{\theta})$, y iii) La distribución de probabilidad según la cual se distribuyen las innovaciones, es una doble exponencial que permite asimetría y exceso/defecto de curtosis. Dicha función es una variación de la pdf simétrica *Exponential Power* (EP) (Box y Tiao, 1992), mediante el método de conversión de funciones simétricas en asimétricas de Fernandez y Steel (1998). Su nombre es *Skew Exponential Power Distribution* (SEP).

La SEP puede adoptar una variedad de formas, incluyendo la Gaussiana, la de Laplace y la consideración de distribuciones de densidad cuyas colas son pesadas (Figura 2.7). Esta flexibilidad es constreñida por las observaciones $\tilde{\mathbf{Y}}$, durante la inferencia de los parámetros. Su expresión estandarizada (media cero y desviación estándar unidad) tiene dos parámetros $\{\beta, \xi\}$, que representan a la curtosis y la asimetría de la distribución, respectivamente. La expresión analítica de la $P_{\text{SEP}}(0, 1, \xi, \beta)$ viene dada por la ecuación (2.28)

$$P_{\text{SEP}}(a | \xi, \beta) = \frac{2\sigma_\xi}{\xi + \xi^{-1}} w_\beta \exp\left(-c_\beta \left| \frac{\mu_\xi + \sigma_\xi a}{\xi^{\text{sgn}(\mu_\xi + \sigma_\xi a)}} \right|^{\frac{2}{1+\beta}}\right) \quad (2.28)$$

donde los factores $c_\beta, w_\beta, \mu_\xi, \sigma_\xi$ son función de los parámetros $\{\beta, \xi\}$, según Fernandez y Steel (1998) o Schoups y Vrugt (2010). En la ecuación (2.28), la variable a representa a las innovaciones estandarizadas, es decir $a_j = \sigma_{z_j}^{-1} z_j$ $\forall j \in \{1, \dots, n\}$ siendo $z_j = e_j - \phi_1 e_{j-1}$. Como se observa, la desviación típica de las innovaciones σ_{z_j} no se considera constante. Nótese que en el trabajo previo de Sorooshian y Dracup (1980), el único hasta la fecha que había modelado de forma directa la varianza, se modela la varianza de los errores y no de las

innovaciones. En la siguiente sección, se explica cómo Evin et al. (2013) prueban que esta forma de estabilizar la varianza, actuando sobre las innovaciones y no sobre los errores, produce inestabilidades en la inferencia conjunta de los parámetros.

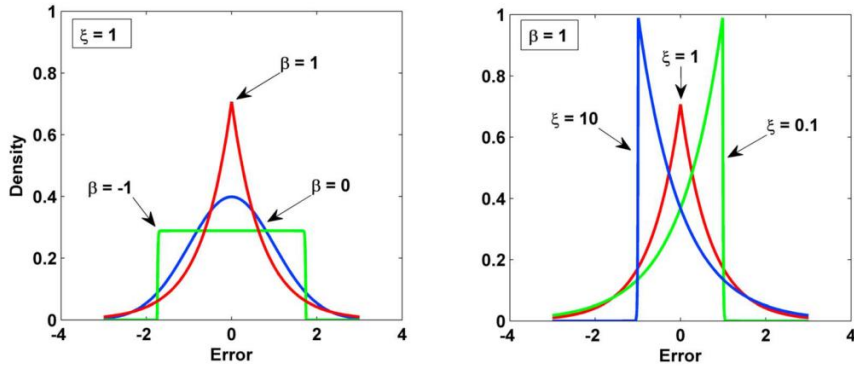


Figura 2.7. Densidades de la pdf SEP estándar (media=0 y std.dev=1), para varios valores de sus dos parámetros, curtosis (β) y asimetría (ξ). Fuente: Schoups y Vrugt (2010)

En resumen, la función de verosimilitud GL que se propone, incorpora cinco parámetros correspondientes al modelo de error $\Omega = \{\phi_1, \alpha, \kappa, \beta, \xi\}$ los cuales, y esto es importante, se infieren de manera conjunta con los parámetros θ del modelo hidrológico. La verosimilitud GL deducida en Schoups y Vrugt (2010) viene expresada por la ecuación (2.29)

$$\ell(\theta, \Omega | \tilde{Y}) = \prod_{j=1}^n \sigma_{z_j}^{-1} P_{SEP}(a_j(\theta, \Omega)) \quad (2.29)$$

Aplicando logaritmos y tomando en consideración la ecuación (2.28) de la SEP, se llega a la expresión (2.30), derivada por Schoups y Vrugt (2010)

$$\mathcal{L}(\theta, \Omega | \tilde{Y}) \approx n \log \frac{2\sigma_\xi w_\beta}{(\xi + \xi^{-1})} - \sum_{j=1}^n \log \{\sigma_{z_j}\} - c_\beta \sum_{j=1}^n \left| \frac{\mu_\xi + \sigma_\xi a_j}{\xi^{sgn(\mu_\xi + \sigma_\xi a_j)}} \right|^{1+\beta} \quad (2.30)$$

Problemática de los casos de estudio

Schoups y Vrugt (2010) presentan dos casos de estudio para demostrar la viabilidad de la metodología propuesta. La principal diferencia entre ambos casos es el tipo de cuenca que emplean. Ambas cuencas pertenecen al experimento MOPEX (Duan et al., 2006) teniendo la propiedad de ser la más seca (Guadalupe River (GR)) y la más húmeda (French Broad (FB)) del MOPEX. En ambos casos de estudio emplean un modelo hidrológico llamado CRR (Schoups et al., 2010). Puesto que, en uno de los casos de estudio de la presente Tesis, se emplea el mismo modelo y las mismas cuencas, más adelante se describirán todos ellos con más detalle.

En el primer caso de estudio, aplicado sobre la cuenca FB, realizan la inferencia conjunta de los parámetros mediante la verosimilitud GL. Sin embargo, solo consideran la inferencia conjunta de los parámetros $\{\theta, \phi_1, \alpha, \kappa, \beta\}$, dejando fijo con valor $\xi = 1$ el parámetro correspondiente a la asimetría de la SEP. El caso de estudio, tal cual lo presentan, no les genera ningún inconveniente o problema.

El segundo caso de estudio, aplicado sobre la cuenca efímera GR, lo plantean en principio con la inferencia conjunta de todos los parámetros hidrológicos y del modelo de error. Al realizar la inferencia encuentran resultados “sin sentido” como se expresa en el artículo. Obtienen unas bandas de incertidumbre de la predicción exageradamente anchas. El trabajo también reporta un coeficiente de correlación inferido $\phi_1 \approx 1$, lo cual es síntoma de que la serie de errores está cerca de no ser o no es estacionaria. Los autores describen varios intentos para solucionar el problema. Entre otras medidas, emplean modelos AR de orden superior (hasta de orden 4) para comprobar si el problema estaba en la simplicidad del modelo AR(1) para modelar los errores. Sin embargo, no se consigue arreglar el problema. Finalmente, deciden fijar el valor de dicho coeficiente $\phi_1 = 0.4$, en base a datos muestrales de errores obtenidos en alguna calibración previa del modelo, procediendo así con éxito a la inferencia conjunta del resto de los parámetros $\{\theta, \alpha, \kappa, \beta, \xi\}$. Estos problemas descritos en el segundo caso de estudio, se reportan de idéntica forma con otro modelo hidrológico y otra cuenca en el trabajo de Scharnagl et al. (2015), quienes tampoco encuentran una solución.

En resumen, de los 5 parámetros de error de la verosimilitud GL, Schoups y Vrugt (2010) solamente infieren conjuntamente 4, en ambos casos de estudio. En el primer caso, no se explica el por qué, se deja de inferir el parámetro de asimetría ξ . En el segundo caso, no infieren el modelo de dependencia ϕ_1 , por los problemas sobrevenidos.

2.3.6. Corrección de Evin et al. a la función GL

Evin et al. (2013) muestran, con una nota técnica, el posible origen de los problemas encontrados por Schoups y Vrugt (2010) en su segundo caso de estudio. La hipótesis que plantean es la que sigue. En inferencias conjuntas del modelo hidrológico, con modelos de error que incluyen un modelo de dependencia y un modelo de heterocedasticidad (método directo sin transformación previa), la aplicación del modelo de heterocedasticidad sobre las innovaciones resultantes de aplicar el modelo de dependencia puede producir inestabilidad en la inferencia, mostrando un “comportamiento pobre” del modelo de error, con un exagerado agrandamiento de las bandas de incertidumbre. El problema, según los autores, parece residir en el orden en el que se aplican el modelo de dependencia y el modelo de varianza. Por ello, proponen evitar la citada inestabilidad mediante la aplicación del modelo de varianza sobre los errores, en primer lugar. A continuación, sobre los errores con la varianza ya estabilizada, se podría aplicar sin problemas el modelo de dependencia, cuyo resultado será la obtención de unas innovaciones ya independientes y con varianza uniforme.

La hipótesis y solución planteadas son comprobadas empíricamente por sus autores, mediante 3 casos de estudio sobre diferentes cuencas (2 húmedas y una efímera), empleando el modelo hidrológico GR4J (Perrin et al., 2003). En dichos casos de estudio no emplean la SEP como distribución de las innovaciones, sino que asumen la Normalidad de estas. Por tanto, su inferencia conjunta presenta tan solo los parámetros $\{\theta, \alpha, \kappa, \phi_1\}$.

Uno de los casos de estudio, se realiza sobre la cuenca húmeda FB, también empleada por Schoups y Vrugt (2010), aunque ambos estudios difieren en el modelo de error y en el modelo hidrológico empleado. Evin et al. (2013) demuestran de forma empírica, sobre la cuenca FB, como al aplicar el modelo de varianza sobre las innovaciones, se producen los ya mencionados problemas de

ensanchamiento de las bandas de incertidumbre. Sin embargo, al aplicar su solución, esto es, el modelo de varianza sobre los errores y a continuación el modelo de dependencia, el problema del ensanchamiento de las bandas desaparece. La justificación teórica que subyace a la diferencia de comportamiento entre ambas aproximaciones, las cuales aparentemente solo difieren en el orden de aplicación de los dos elementos del modelo de error, la encuentran analizando las propiedades de autocorrelación de primer orden que presentan los errores, en cada una de las dos aproximaciones. Cuando se emplea la aproximación propuesta como solución, la autocorrelación de primer orden es constante y coincide exactamente con el coeficiente de autocorrelación, es decir, $\rho(e_j, e_{j-1}) = \phi_1$ resultado conocido para los procesos lineales autorregresivos de primer orden (Box et al., 1994). Sin embargo, en la aproximación que siguen Schoups y Vrugt (2010), el resultado es más complejo, obteniendo el siguiente resultado,

$$\rho(e_j, e_{j-1}) = \phi_1 \frac{\sigma_{e|y_{j-1,M}}}{\sigma_{e|y_{j,M}}} = \phi_1 \frac{\alpha + \kappa y_{j-1,M}}{\alpha + \kappa y_{j,M}} \approx \phi_1 \frac{y_{j-1,M}}{y_{j,M}} = \phi_1 \frac{y_{j-1,M}}{y_{j-1,M} + \Delta y_{j,M}} \quad (2.31)$$

De la ecuación (2.31) se pueden deducir tres posibles comportamientos para el valor de $\rho(e_j, e_{j-1})$, en función del valor que adopte el término $\Delta y_{j,M}$, que expresa el aumento en el valor de la variable del modelo $y_{j,M}$ respecto al resultado inmediato anterior $y_{j-1,M}$. Así, en el caso de que $\Delta y_{j,M} \geq 0$, entonces $\rho(e_j, e_{j-1}) \leq \phi_1$, ocurriendo estas situaciones en las ramas de crecida o en zonas llanas en la gráfica de la variable $y_{j,M}$ (e.g. del hidrograma). En el caso de que $\Delta y_{j,M} < 0$, según la ecuación (2.31), se obtendrá un resultado $\rho(e_j, e_{j-1}) > \phi_1$, en el que la correlación será tanto mayor cuanto más brusco sea el decremento $\Delta y_{j,M}$. Esta situación ocurre en las ramas de recesión de la variable $y_{j,M}$. El aumento de la correlación entre valores consecutivos del error es un aumento de la “memoria” entre dichos valores de la serie. Este efecto aparecerá, exclusivamente, después de los picos en la variable $y_{j,M}$, para los cuales la varianza del error es máxima (máxima incertidumbre) y a partir de los cuales

$\Delta y_{j,M} < 0$. Si bien Evin et al. (2013) argumentan la teoría del problema de modo diferente, si que expresan de forma explícita el hecho de que el excesivo ensanchamiento de las bandas se produce exclusivamente en las recesiones, y de forma más exagerada en las dos cuencas húmedas que en la efímera.

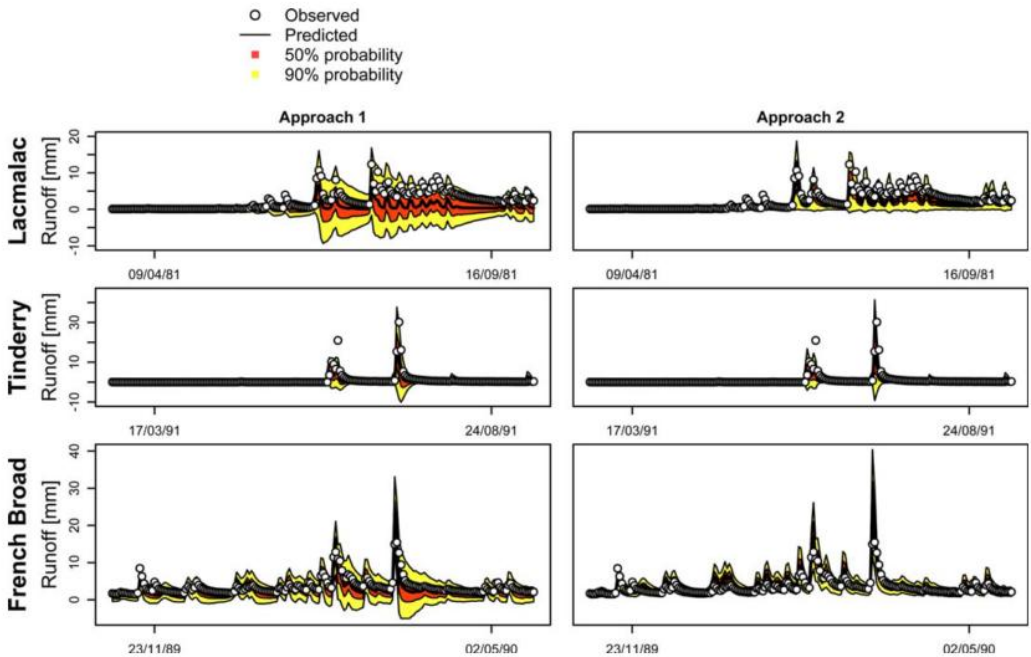


Figura 2.8. Distribuciones predictivas sobre 3 cuencas. En los paneles de la izquierda, obtenidas mediante la aproximación de Schoups y Vrugt (2010). En los paneles de la derecha obtenidas con la corrección de Evin et al. (2013). Fuente: Evin et al. (2013)

Esto último, responde al hecho de que en las cuencas húmedas la recesión es más suave y extendida en el tiempo (con la consiguiente prolongación del efecto memoria en las bandas de incertidumbre), mientras que en la cuenca efímera (Tinderry) la recesión es abrupta y el efecto de ensanchamiento de las bandas no es tan apreciable (Figura 2.8). Sin embargo, a pesar de haber resuelto (aparentemente) el problema del ensanchamiento de las bandas de incertidumbre, los autores refieren nuevos problemas que “enturbian” la fiabilidad de la solución propuesta. Dichos problemas están relacionados con la fuerte correlación que detectan entre ciertos parámetros del modelo hidrológico y los parámetros del modelo de error. Por otra parte, en vista de los resultados, los

autores también reconocen que la hipótesis de Normalidad de las innovaciones es, al menos, cuestionable puesto que comprueban la presencia de asimetrías y exceso de curtosis, en los tres casos de estudio.

2.4. Incertidumbre de la predicción

La incertidumbre predictiva de un modelo dado (o de un conjunto de modelos en el marco de las técnicas BMA), se define (Montanari y Koutsoyiannis, 2012) como la incertidumbre en el proceso de predicción del verdadero valor de una variable hidrológica. Se cuantifica mediante la distribución de probabilidad de la variable predicha y generalmente, se expresa en forma de límites del rango de la variable, referidos al valor de simulación del modelo, dentro de los cuales estaría el verdadero valor de la variable, con una probabilidad igual al nivel de confianza deseado.

2.4.1. Evaluación de la incertidumbre total de la predicción

Según Todini (2007), en modelación hidrológica se pueden diferenciar dos objetivos bien diferentes: la estimación de parámetros y la predicción hidrológica. La estimación de parámetros, sobre la que ya ha tratado esta Tesis, persigue la obtención del conjunto de parámetros más verosímil del modelo hidrológico, con el mejor modelo de error posible y para el conjunto de datos dado.

Con respecto al problema de hacer predicciones sobre una variable de interés, se puede resolver sin necesidad de la correcta estimación de los parámetros hidrológicos. La incertidumbre predictiva $P(\mathbf{Y}|\mathbf{Y}_M)$, se puede obtener mediante la modelación directa de de la distribución conjunta $P(\mathbf{Y}, \mathbf{Y}_M)$ de la variable a predecir y su simulación proporcionada por el modelo hidrológico (Coccia y Todini, 2011; Todini, 2013).

La consecución simultánea de ambos objetivos, el de la estimación correcta de los parámetros y el de la incertidumbre predictiva, es posible si en el proceso de inferencia se emplea el mejor modelo de error posible, que permita la correcta modelación de los errores generados. Obtener tal modelo de error puede, sin embargo, ser una tarea complicada (Smith et al., 2015). La inferencia de los

parámetros, con base probabilística (basada en un modelo de error), se conoce con el nombre de **inferencia Formal** (Bayesiana).

Dados los modelos hidrológico y de error, una definición formal de **incertidumbre predictiva** (Mantovan y Todini, 2006) viene dada por la ecuación (2.32)

$$P\left(\mathbf{Y}|\mathbf{Y}_{\mathcal{M}}\left(\tilde{\mathbf{Y}}, \tilde{\mathbf{B}}, \tilde{\mathbf{x}}_0\right)\right)=\int_{\Pi} P\left(\mathbf{Y}|\mathbf{Y}_{\mathcal{M}}\left(\boldsymbol{\pi}, \tilde{\mathbf{B}}\right)\right) P\left(\boldsymbol{\pi}|\tilde{\mathbf{Y}}, \tilde{\mathbf{B}}, \tilde{\mathbf{x}}_0\right) d\boldsymbol{\pi} \quad (2.32)$$

donde el término de la izquierda es la pdf de la variable a predecir \mathbf{Y} , condicionada a sus n observaciones $\tilde{\mathbf{Y}}$, a las condiciones iniciales de las variables de estado del sistema $\tilde{\mathbf{x}}_0$, a las series observadas de las variables que excitan el sistema $\tilde{\mathbf{B}}$ y necesariamente al propio modelo \mathcal{M} .

Por tanto, en esta expresión de la incertidumbre predictiva, solamente se ha marginalizado el efecto de la incertidumbre de los parámetros $\boldsymbol{\pi}=\{\boldsymbol{\theta}, \boldsymbol{\Omega}\}$ $\boldsymbol{\pi} \in \Pi \in \mathbb{R}^d$, mediante la integral sobre su dominio de existencia Π . Dicho de otra forma, la evaluación de la incertidumbre predictiva, tal como está expresada por la ecuación (2.32), requiere considerar todas las simulaciones $\mathbf{Y}_{\mathcal{M}}$ del modelo hidrológico correspondientes a las múltiples realizaciones muestreadas a partir de la posterior de los parámetros, en lugar de considerar solamente la simulación correspondiente al conjunto de parámetros óptimo Bayesiano o MAP. Lo mismo ocurre al simular el modelo de error para cada simulación $\mathbf{Y}_{\mathcal{M}}$. La simulación del modelo de error requerirá la consideración del muestreo de múltiples realizaciones de la posterior de sus parámetros.

El primer término en la parte derecha de la ecuación $P\left(\mathbf{Y}|\mathbf{Y}_{\mathcal{M}}\left(\boldsymbol{\pi}, \tilde{\mathbf{B}}\right)\right)$, es la pdf de la variable a predecir condicionada a las series observadas de las variables que excitan el sistema y a un único conjunto de parámetros de los modelos hidrológico y de error. El segundo término en la derecha de la ecuación $P\left(\boldsymbol{\pi}|\tilde{\mathbf{Y}}, \tilde{\mathbf{B}}, \tilde{\mathbf{x}}_0\right)$ corresponde a la distribución posterior conjunta de los parámetros hidrológicos y del modelo de error, dadas las series observadas de la variable que se quiere predecir, de las variables excitadoras del sistema y de las

condiciones iniciales de las variables de estado. La distribución posterior de los parámetros se puede obtener mediante el proceso de inferencia Bayesiana, tal y como se ha explicado en las secciones 2.2.1 y 2.2.2.

2.4.2. Evaluación de la capacidad predictiva

En la literatura existen múltiples métodos para la evaluación de la capacidad de predicción que posee la distribución predictiva obtenida (Datta y Bolisetti, 2013; Gneiting et al., 2007; Laio y Tamea, 2006; Li et al., 2011).

La metodología seguida en esta Tesis está basada en el trabajo de Laio y Tamea (2006), por ser la más empleada en las publicaciones relacionadas con la presente investigación. Concretamente se ha seguido el mismo procedimiento que en las publicaciones de Renard et al. (2010) o Thyer et al. (2009), las cuales introducen pequeñas variaciones respecto al trabajo de Laio y Tamea (2006). El procedimiento se basa en las gráficas denominadas PP-plots, en las cuales se enfrentan dos series de probabilidades correspondientes a dos distribuciones de probabilidad, con el fin de compararlas. Según la aplicación, hay casos en los que ambas series de probabilidad enfrentadas son desconocidas y de carácter empírico (muestrales) y en otros casos se enfrenta una serie de probabilidades empírica, frente a una serie de probabilidades teórica de una distribución conocida. A este último caso corresponde el tipo de PP-plots empleados para evaluar la fiabilidad de la distribución predictiva. Seguidamente se expone el fundamento del método.

Una **distribución predictiva correcta** es aquella que representa a las verdaderas distribuciones marginales de todas y cada una de las observaciones (Laio y Tamea, 2006). Ocurre que estas distribuciones marginales no son conocidas, es más, usualmente solo se tiene una muestra observada de cada una de ellas. Aún así, es posible comprobar la mencionada representatividad mediante un test de hipótesis. La explicación es como sigue. Piense el lector, en una de las observaciones de la variable predicha $\tilde{y}_j \in \tilde{\mathbf{Y}} \quad \forall j \in \{1, \dots, n\}$. Por otra parte, sea $F_{j,\mathcal{M}}(Y_{\mathcal{M}})$, la distribución empírica de probabilidad acumulada en j , $\forall j \in \{1, \dots, n\}$, obtenida a partir del *ensemble* generado de simulaciones del modelo $\{y_{j,\mathcal{M}}\}_{\Pi} \in \{\mathbf{Y}_{\mathcal{M}}\}_{\Pi}$ como resultado de marginalizar la incertidumbre de los

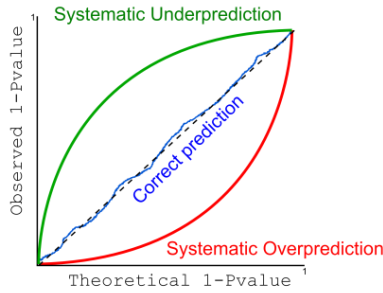
parámetros mediante la integración sobre todo su rango probable de valores Π , es decir, muestreando la posterior. Considérese también, el cálculo de la probabilidad acumulada de la observación \tilde{y}_j , obtenida según la distribución empírica del mencionado *ensemble* de simulaciones, esto es $z_j = F_{j,\mathcal{M}}(\tilde{y}_j)$. El test de coincidencia, consiste en comprobar que las probabilidades $z_j \forall j \in \{1, \dots, n\}$, se distribuyen de forma uniforme, esto es, comprobar que las $z_j \stackrel{\mathcal{D}}{\sim} \mathcal{U}(0,1) \quad \forall j \in \{1, \dots, n\}$.

El significado del test es comprobar si las observaciones son una muestra aleatoria de la distribución predictiva generada a partir de los modelos, hidrológico y de error. Por tanto, si se construye un PP-plot en el que se sitúan en abscisas las probabilidades acumuladas teóricas de la distribución $u_j \stackrel{\mathcal{D}}{\sim} \mathcal{U}(0,1)$, y en ordenadas las probabilidades acumuladas observadas $z_j \forall j \in \{1, \dots, n\}$, el resultado positivo del test de representatividad se producirá cuando la posición de los pares $\{u_j, z_j\}$ se ubique sobre el bisector del diagrama. La asignación de cada variable a los ejes de ordenadas y abscisas es indiferente y no afecta al método. Sí que afecta, sin embargo, a las interpretaciones del diagrama obtenido. A continuación, se explica cómo interpretar los diagramas, considerando la configuración de ejes descrita.

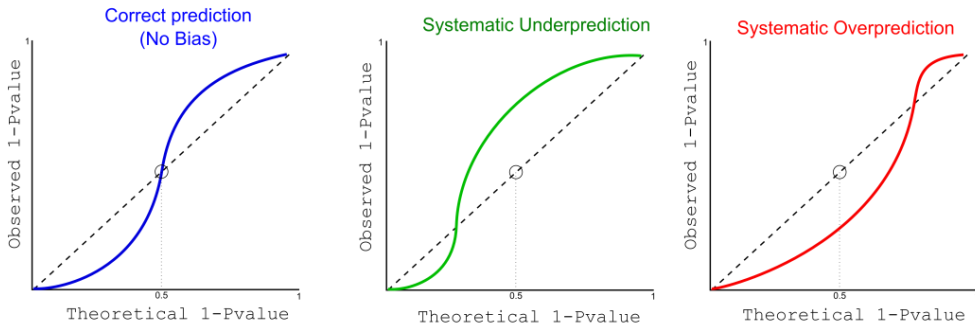
Interpretación de los PP-plots

Los PP-plots son una herramienta capaz de mostrar dos aspectos diferentes, aunque relacionados, sobre la distribución predictiva que se ha inferido. El primero de ellos se relaciona con el sesgo de la distribución, es decir, con el desplazamiento relativo entre la distribución predictiva y las observaciones. Solamente cuando la curva del PP-plot cruza el bisector en el valor de probabilidad 0.5, la distribución predictiva es insesgada. Si el cruce se produce en otro punto del bisector o no hay cruce, se tendrá una (infra) sobrepredicción sistemática. El segundo de los aspectos mostrados está relacionado con la corrección del ancho de las bandas de incertidumbre, es decir, con la bondad de la estimación de la incertidumbre.

Correct Uncertainty estimation



Uncertainty Underestimation



Uncertainty Overestimation

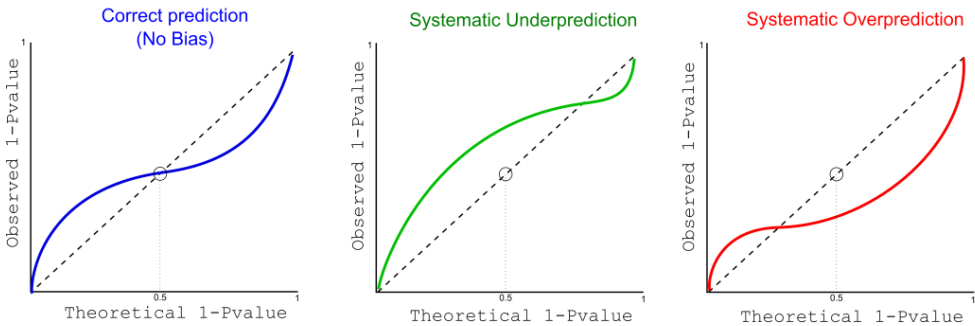


Figura 2.9. Diagramas interpretativos de los PP-plots. Se representan diferentes combinaciones de las propiedades de exactitud (bias) y precisión (incertidumbre) de la distribución predictiva

Así, una curva cuya forma es de “S” invertida (típicamente obtenida en inferencias con el modelo de error SLS), indica una sobreestimación de la incertidumbre predictiva. Por el contrario, una curva con forma de “S” está indicando una infraestimación de la incertidumbre. Con estos criterios generales, se ha dibujado en la Figura 2.9, unos esquemas con los casos posibles que pueden darse en la interpretación de estos diagramas.

Adicionalmente a la identificación visual de las características de la distribución predictiva inferida, es posible realizar la cuantificación de las mismas mediante dos índices llamados **Fiabilidad** y **Resolución**. El índice de fiabilidad cuantifica en qué medida se parece la distribución predictiva inferida, a la distribución (desconocida) de la cual son una muestra aleatoria las observaciones. Su rango de valores es $[0,1]$. Cuando el PP-plot se superpone al bisector del diagrama, ambas distribuciones son la misma y en ese caso el índice de fiabilidad es 1. Cuanto más se diferencia la distribución predictiva inferida de la distribución de las observaciones, más se aleja el PP-plot del bisector. En ese caso el índice de fiabilidad es menor que uno, tanto menor cuanto más aumenta el área encerrada entre el PP-plot y el bisector. El índice de resolución es una medida de la precisión media de la distribución predictiva. De hecho, es la esperanza de los inversos de los coeficientes de variación de las distintas distribuciones condicionales que definen la distribución predictiva. Su rango de valores teórico es $]0,+\infty[$, siendo más precisas (menos anchas) las distribuciones predictivas con mayor índice de resolución.

En resumen, el método aconsejado por esta Tesis para caracterizar la distribución predictiva pasa por identificar visualmente a cuál de los posibles diagramas de la Figura 2.9 se adapta mejor el PP-plot obtenido. A efectos de comparación cuantitativa entre distribuciones predictivas, serán más correctas las distribuciones que tengan su índice de fiabilidad más próximo a 1. El índice de resolución, a diferencia del de fiabilidad, no es un indicador de la corrección de la distribución predictiva inferida, sino más bien, un indicador medio de la magnitud de la incertidumbre de la predicción media, que se deriva de la distribución predictiva inferida. Por tanto, si el índice de fiabilidad indica corrección en la distribución predictiva, pero se obtiene un bajo índice de resolución (e.g. inferior a 2.5-3.0), lo que se está indicando es que las predicciones obtenidas presentan, en promedio, una alta incertidumbre. Para más detalles sobre la evaluación de la capacidad predictiva, se puede consultar los trabajos de Renard et al. (2010) o Thyer et al. (2009).

CAPÍTULO 3. EL MODELO DE ERROR

Imagination is more important than knowledge. For knowledge is limited to all we now know and understand, while imagination embraces the entire world, and all there ever will be to know and understand

-Albert Einstein (1879-1955)-

3.1. Motivación

Las conclusiones más importantes sobre el estado de la Cuestión, las cuales motivan tanto los objetivos de esta Tesis, como la forma, herramientas y métodos con los que se ha alcanzado dichos objetivos, se resumen en los siguientes puntos:

1. Los métodos clásicos de optimización de parámetros se han quedado obsoletos para ser aplicados en modelación hidrológica, por diversas causas entre las cuales destacan: i) No son capaces de tratar con parámetros no identificables del modelo; ii) No aseguran el mínimo global, dentro del rango de búsqueda; iii) No proporcionan una estimación fiable de la incertidumbre de los parámetros y iv) Todos los métodos basan su estrategia en la optimización de funciones objetivo diseñadas Ad-Hoc, siendo la mayoría, variantes del método SLS cuyas hipótesis no son válidas, con carácter general, en modelación hidrológica.
2. La inferencia Bayesiana demuestra ser una herramienta útil que es capaz de solucionar todos los inconvenientes mencionados sobre los métodos de optimización. De entre todos los algoritmos computacionales Bayesianos, diseñados para el muestreo de la posterior de los parámetros, en la actualidad son los métodos MCMC adaptativos de cadena múltiple los que presentan mayor ventaja respecto a sus alternativas.
3. En 1980, Sorooshian y Dracup publican un trabajo en el que se pone de manifiesto la necesidad de modelar los errores procedentes de la modelación hidrológica, con hipótesis adecuadas a su naturaleza, las cuales son contrarias a las hipótesis que subyacen al método SLS, el más empleado hasta entonces en calibración de modelos. Dichos autores proponen por vez primera un método directo para la modelación de la heterocedasticidad de los errores, como alternativa a los métodos de transformación de las variables empleados hasta entonces. Así mismo, obtienen la función de verosimilitud que permite la consideración de la dependencia serial de los

- errores. Ambas propuestas las realizan por separado, es decir, no aúnan en una sola función de verosimilitud ambas soluciones.
4. En 2010, 30 años después del trabajo de Sorooshian y Dracup, los autores Schoups y Vrugt publican un trabajo en el cual retoman la metodología directa de la modelación de la varianza, si bien la aplican sobre las innovaciones, en lugar de sobre los errores. Su aporte es innovador en dos sentidos: i) Proponen una única función de verosimilitud, a la que denominan Generalizada (GL), mediante la cual es posible considerar la heterocedasticidad, la dependencia y la no Normalidad de los errores y ii) Realizan de forma conjunta la inferencia Bayesiana de todos los parámetros del problema, es decir, los hidrológicos junto con los del modelo de error. Sin embargo, en uno de los dos casos de estudio que publican, se encuentran con problemas en la inferencia del parámetro de autocorrelación de los errores, problema que les genera unas bandas de incertidumbre exageradamente anchas. Los autores no encuentran solución al problema y deciden asignar un valor fijo al coeficiente de autocorrelación, sacándolo así de la inferencia conjunta, lo cual les permite obtener resultados aceptables.

3.1.1. Primeros experimentos

Estando así la Cuestión, esta Tesis queda encuadrada como una aproximación Formal de inferencia Bayesiana, con modelo de error agregado, para la obtención de la incertidumbre de los parámetros y la incertidumbre total de la predicción. La metodología a seguir es, en principio, similar a la del trabajo de Schoups y Vrugt (2010). El primer experimento realizado, sigue los pasos del primer caso de estudio de estos autores: misma cuenca (FB), mismo modelo hidrológico (CRR), mismo periodo de estudio, misma función de verosimilitud (GL) y mismo algoritmo MCMC. Tan solo una diferencia, que en principio pasa inadvertida para el que escribe: en el experimento se incluyen en la inferencia conjunta los 5 parámetros del modelo de error, esto es $\Omega = \{\phi, \alpha, \kappa, \beta, \xi\}$, a diferencia del trabajo original, en el que no se incluyó el parámetro de asimetría ξ de la SEP. El resultado es que el problema no alcanza la convergencia después de la enorme cantidad de 1.5E6 iteraciones del algoritmo MCMC. Tras el primer experimento fallido se prueba la inferencia de forma exacta a como la realizan Schoups y Vrugt (2010), con los 4 parámetros del modelo de error

$\Omega = \{\phi_1, \alpha, \kappa, \beta\}$. La inferencia concluye de forma normal, obteniendo similares resultados al trabajo original.

Pocas semanas después del fallido primer experimento, Evin et al. (2013) publican una nota técnica (ver sección 2.3.6) en la cual dicen haber encontrado el origen y la solución del problema de las, inexplicablemente anchas, bandas de incertidumbre reportadas en el segundo caso de estudio de Schoups y Vrugt (2010). Empíricamente demuestran que la solución que proponen es buena para evitar la inflación de las bandas, sin embargo, surgen otros problemas relacionados con la alta correlación entre parámetros del modelo hidrológico y los parámetros del modelo de error. Así mismo, reconocen que la hipótesis de errores Gaussianos, empleada por ellos, no es adecuada en ninguno de sus casos de estudio. Del trabajo de Evin et al. (2013) se obtienen dos ideas importantes. La primera es que la función de distribución de las innovaciones a utilizar no es la Gaussiana, por lo que la SEP es una buena vía con la que continuar la investigación. La segunda idea es que posiblemente su aportación resuelva el problema con el que se ha topado el primer experimento de esta Tesis.

De la fusión de los trabajos de Schoups y Vrugt (2010) y Evin et al. (2013) surge el segundo de los experimentos. En este segundo experimento, se incluyen de nuevo los 5 parámetros del modelo de error, como en el primer experimento, pero se aplica la modificación al método de Schoups y Vrugt propuesta por Evin et al. Esta vez, la convergencia se produce en tan solo 0.1E6 iteraciones MCMC. Sin embargo, los resultados muestran unas bandas de incertidumbre de la predicción con una amplitud excesiva.

Si bien la solución de Evin permite la convergencia con los cinco parámetros de error en el segundo experimento, no resuelve el problema de la inflación de las bandas de incertidumbre, que su publicación decía resolver. Nótese, que las diferencias entre ambos trabajos de Schoups y Vrugt (2010) y Evin et al. (2013), distinto modelo hidrológico, modelos de error diferentes con diferente número de parámetros, no permiten identificar claramente, por comparación, donde puede estar el origen de los problemas.

Otro detalle a tener en cuenta por intrigante es, por qué el primer experimento realizado converge o deja de hacerlo, por la simple eliminación o adición a la inferencia de un solo parámetro de error, ξ , la asimetría de la SEP.

3.1.2. Inspiración

Después de los primeros experimentos, infructuosos, las intenciones de seguir por la vía de la inferencia conjunta pasaron por momentos de duda. Fue momento de parar y volver a buscar en trabajos de otros autores. Especialmente inspiradores fueron los trabajos de Montanari y Brath (2004) y Todini (2007). En el trabajo de Todini aparece la siguiente idea: “[...] let us introduce the concept of the joint probability distribution of the real quantity of interest, the predictand Y (namely the discharge [...]), and the model forecast Y_M . Unless the model is so exceptionally accurate, thus perfectly matching the observations of Y , a scatter $\{y_{j_M}, \tilde{y}_j\} \forall j \in \{1, \dots, n\}$ will always be observed in the $Y_M - Y$ plane. This scatter is a representation of the joint sample frequency of $P(\mathbf{Y}, \mathbf{Y}_M)$, that can be used to estimate this joint probability density”.

La idea considera que la variable a predecir \mathbf{Y} y su simulado \mathbf{Y}_M conforman una pdf conjunta (Figura 3.1). Esta es una idea fundamental, siendo uno de los pilares en los que se basan los métodos de estimación directa de la incertidumbre predictiva total de los modelos hidrológicos. Recuérdese (sección 2.4.1) que estos métodos asumen que los parámetros del modelo ya están calibrados, sin atender a cómo se ha realizado la inferencia. El primero de estos métodos de estimación directa fue el Procesador de Incertidumbre Hidrológica (HUP, siglas en inglés) (Krzysztofowicz y Kelly, 2000), que es uno de los componentes del Sistema de Predicción Bayesiana (BFS, siglas en inglés) de Krzysztofowicz (1999, 2001, 2002), Krzysztofowicz y Herr (2001) y Krzysztofowicz y Kelly (2000). El método HUP evolucionaría años más tarde, dando lugar al método del Procesador Condicional de Modelos (MCP en inglés) (Todini, 2008, 2013).

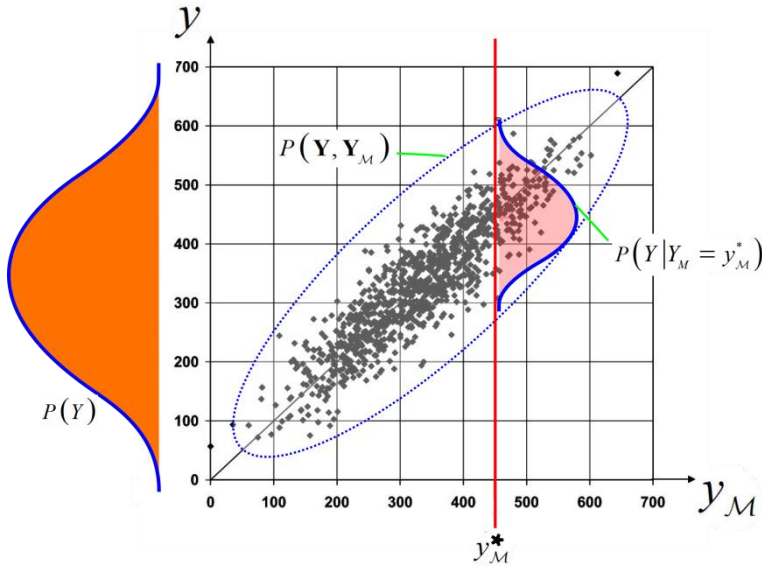


Figura 3.1. Distribución muestral conjunta de $P(\mathbf{Y}, \mathbf{Y}_M)$, en la cual se representan la marginal (en rojo) $P(Y)$ y una de las distribuciones condicionales $P(Y|Y_M)$, condicionada en este caso al valor $y_{jM} = y_M^*$

En la Figura 3.1, la nube de puntos $\{y_{jM}, \tilde{y}_j\} \forall j \in \{1, \dots, n\}$ es una muestra de la pdf bivariada $P(\mathbf{Y}, \mathbf{Y}_M)$. Krzysztofowicz (1999) sugiere la idea de que la marginal $P(\mathbf{Y})$ de la distribución conjunta, contiene toda la variabilidad con la que se observa la variable de interés \mathbf{Y} . Sin embargo, la consideración de la información de la que disponemos, típicamente contenida en un modelo hidrológico, permite reducir la incertidumbre sobre \mathbf{Y} , cuantificándola en términos de una pdf condicional a los resultados que el modelo proporciona, es decir, la pdf condicional $P(\mathbf{Y}|\mathbf{Y}_M = y_M^*)$ es la incertidumbre predictiva total de las simulaciones $\{y_{jM}\} \forall j \in \{1, \dots, n\}$ cuya magnitud es y_M^* .

El otro trabajo especialmente inspirador fue el trabajo de Montanari y Brath (2004). En lo fundamental el trabajo sigue la línea del HUP (Krzysztofowicz y Kelly, 2000), sin embargo, este trabajo propone la estimación de la pdf de los

errores, condicionados a la variable predicha (el caudal en este caso), trabajando la parte estadística en el espacio Meta-Gaussiano (Kelly y Krzysztofowicz, 1997; Krzysztofowicz y Herr, 2001). Por tanto, el trabajo considera la existencia de la distribución bivariada compuesta por los errores en la predicción de la variable y las correspondientes simulaciones del modelo.

Nótese, que bajo la hipótesis de un modelo de error aditivo, se puede realizar una traslación vertical de la nube de puntos de la Figura 3.1 de la siguiente forma $\{y_{j_M}, \tilde{y}_j\} \Rightarrow \{y_{j_M}, \tilde{y}_j - y_{j_M}\} \Rightarrow \{y_{j_M}, e_j\} \forall j \in \{1, \dots, n\}$, tras la cual se cumple que $P(\mathbf{Y}|\mathbf{Y}_M) = P(\mathbf{E}|\mathbf{Y}_M)$, es decir, sin afectar a las densidades de probabilidad condicionales a \mathbf{Y}_M .

Hasta aquí la inspiración de la que surgen preguntas, tal vez, llevadas por la intuición ¿Qué similitudes o diferencias puede tener el modelo de error representado por la GL de Schoups y Vrugt (2010) en relación con las aproximaciones directas de Krzysztofowicz, Montanari o Todini? ¿Se puede generar con el modelo de varianza condicional de los errores de la ecuación (2.22) lo que se está representando en la Figura 3.1?

Del fracaso de los primeros experimentos realizados y del análisis de sus resultados, surgió con fuerza la siguiente idea. El esquema de la Figura 3.2 representa la nube de errores que es muestra de la bivariada $P(\mathbf{E}, \mathbf{Y}_M)$. Sobre ella se han dibujado las distribuciones condicionales $P(\mathbf{E}|\mathbf{Y}_M)$ cuya varianza creciente de forma lineal con y_{j_M} simula haberse modelado con la ecuación (2.22) con un $\psi = 1$, mismo modelo de varianza empleado por Schoups y Vrugt (2010) y por Evin et al. (2013). Se representa a la izquierda, en línea discontinua azul, la distribución marginal que se corresponde con las pdf's condicionales dibujadas. Se representa (en verde) la distribución marginal correcta $P(\mathbf{E})$, que correspondería a la nube de errores representada. La cuestión que surge es la siguiente. La definición de un modelo lineal de varianza condicional de los errores, con dos parámetros $\{\alpha, \kappa\}$, sin ninguna otra restricción o consideración, ¿Garantiza que no se puedan obtener resultados estadísticamente incoherentes como el de la Figura 3.2?

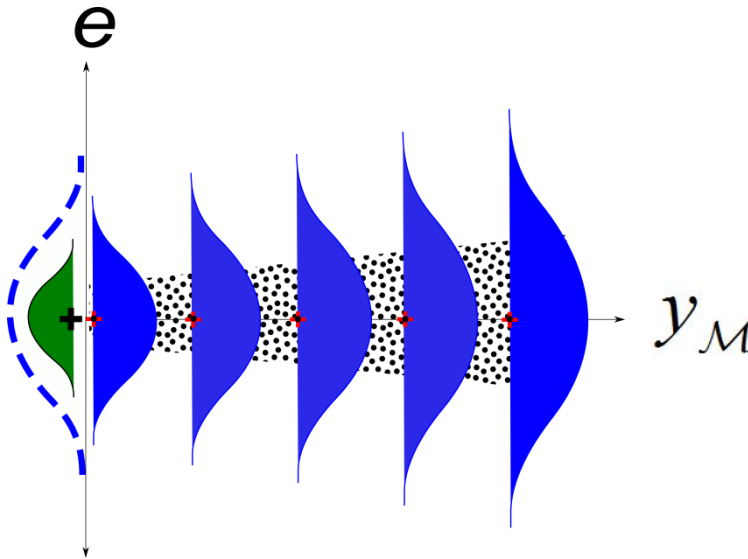


Figura 3.2. Incoherencia estadística debida al desacople entre las distribuciones condicionales inferidas (en azul) y la distribución marginal (en verde) de la nube errores inferida $\{y_{j_M}, e_j\}$

En dicha figura se observa que las distribuciones condicionales, calculadas con el modelo de varianza, y su correspondiente pdf marginal no mantienen ninguna coherencia estadística con la verdadera pdf marginal (en verde) de los errores producidos durante la inferencia, representados por la nube de puntos. En el caso de la figura, las distribuciones condicionales presentan más varianza de la que sería compatible con la distribución marginal correcta. También se podría dar el caso contrario, es decir, unas condicionales con menor varianza que la compatible con la correcta marginal. La ocurrencia de uno u otro caso podría venir exclusivamente impuesta por el valor de los parámetros $\{\alpha, \kappa\}$ que maximizan la función de verosimilitud (e.g. ecuación (2.29)). Según esa ecuación, en la evaluación de la verosimilitud de un error actúan dos “fuerzas” opuestas, el efecto minorador o varianza condicional del error y la propia probabilidad de ocurrencia del error. La fuerza que predomine es la que producirá uno de los dos casos que se explican a continuación.

El primero de los casos es el que, a priori, resulta más común. La función de verosimilitud será maximizada con los menores valores de las varianzas condicionales. Es el mismo mecanismo que se produce en la calibración con los

modelos de error SLS o WLS. Sin embargo, en el caso del SLS la varianza marginal y las condicionales son la misma, con lo cual durante la calibración se reducen simultáneamente manteniendo la coherencia estadística. En el caso de modelar la varianza condicional (e.g. en el caso del WLS), su disminución supone favorecer la tendencia a valores pequeños de los parámetros $\{\alpha, \kappa\}$, aunque sin tener en cuenta el valor de la varianza marginal de los errores.

El segundo caso es el representado en la Figura 3.2 y su funcionamiento resulta menos evidente. La explicación se apoyará en el esquema de la Figura 3.3. En él se representan dos distribuciones condicionales para un mismo valor del modelo y_{j_M} . La distribución roja $P_R(e_j | y_{j_M})$ y con menor varianza, la distribución azul $P_B(e_j | y_{j_M})$. Se representan también dos errores $\{e_1, e_2\}$ correspondientes a dos observaciones de diferente magnitud $\{\tilde{y}_1, \tilde{y}_2\}$ para las que el modelo generó idéntico resultado. Del esquema de la figura se obtiene que $P_R(e_1 | y_{j_M}) > P_B(e_1 | y_{j_M})$, del mismo modo que también se observa que $P_R(e_2 | y_{j_M}) \gg P_B(e_2 | y_{j_M}) \approx 0$. Si se piensa, podría darse el caso en el que un aumento moderado en la varianza condicional, producido durante la inferencia, provocase un aumento importante de la probabilidad total de un conjunto de errores, el cual no fuese contrarrestado de forma suficiente por el moderado aumento de varianza, aumentando de esta manera el valor de la verosimilitud, según la ecuación (2.29). Si esta circunstancia se produce con cierta frecuencia durante la inferencia, podría favorecerse la tendencia a valores grandes de $\{\alpha, \kappa\}$, y como en el primer caso, sin tener en cuenta el valor de la varianza marginal de los errores.

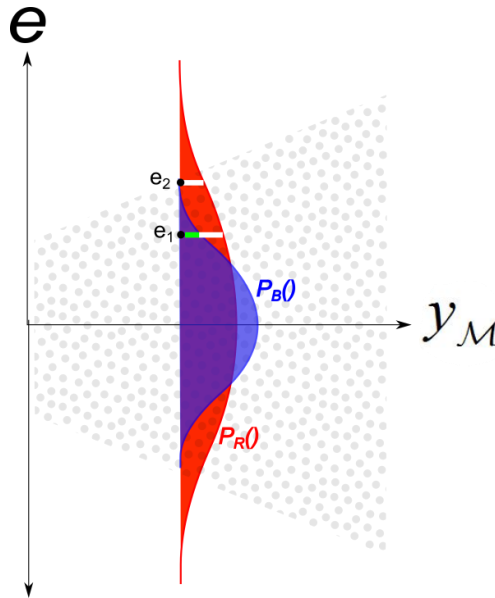


Figura 3.3. Explicación sobre la tendencia al aumento de las varianzas condicionales de los errores, durante una inferencia conjunta

En definitiva, los parámetros $\{\alpha, \kappa\}$ del modelo de varianza podrían seguir unos mecanismos de optimización que no están perfectamente acoplados con el mecanismo de optimización de los parámetros hidrológicos. Tratando de explicarlo de forma visual, durante la inferencia, los movimientos de la nube de errores, que es la que define cómo es la distribución marginal, son guiados principalmente por el modelo hidrológico. Mientras, los movimientos de las distribuciones condicionales no dependen, al menos por completo, del movimiento de la nube de errores, teniendo una influencia importante las dos “fuerzas” contrapuestas que se han mencionado. Piense el lector, que en la calibración con el método SLS, la nube de errores define al mismo tiempo las distribuciones condicionales y la marginal, por lo que en ese caso el acoplamiento es completo.

Este efecto de desacoplamiento puede verse agravado con el uso de modelos de error sofisticados. Un pequeño incremento de la varianza en una distribución condicional asimétrica puede aumentar la probabilidad de ocurrencia en un mayor número de errores. Incluso sin aumentar la varianza condicional, el

aumento de asimetría produciría el mismo efecto. Lo mismo puede ocurrir en el caso de modelar el sesgo y la varianza condicionales de forma conjunta.

La clave del problema está en que **la dinámica de variación de las distribuciones condicionales y marginales de los errores, están desacopladas durante la inferencia, lo cual provoca una pérdida de la coherencia estadística** similar a la de la Figura 3.2.

Por tanto, esta Tesis plantea como hipótesis principal la siguiente:

Un modelo de varianza condicional de los errores que siga el método directo, es decir, sin previa transformación de los datos, debe tomar en consideración la relación que existe entre las distribuciones condicionales y la correspondiente distribución marginal de los errores. La variación de ambas distribuciones, durante el proceso de inferencia conjunta, debe ocurrir de forma acoplada, manteniendo así la coherencia estadística.

El acople del comportamiento entre ambas distribuciones, es sinónimo de coherencia estadística, y ha de venir forzado por la relación matemática que existe entre ambos tipos de distribuciones, cuando pertenecen a una única distribución de probabilidad conjunta. Esa relación matemática viene definida por las Leyes Totales.

3.2. Las Leyes Totales

3.2.1. Derivación

Según la ley de probabilidades condicionales se puede escribir la distribución conjunta de las simulaciones y sus errores como muestra la ecuación (3.1)

$$P(\mathbf{E}, \mathbf{Y}_M) = P(\mathbf{E} | \mathbf{Y}_M) P(\mathbf{Y}_M) \quad (3.1)$$

donde $P(\mathbf{Y}_M)$ es la pdf marginal de las simulaciones del modelo. Por otra parte, la distribución marginal de los errores se puede obtener mediante la integración de $P(\mathbf{E}, \mathbf{Y}_M)$, como se indica en la ecuación (3.2)

$$P(\mathbf{E}) = \int_{\mathbf{Y}_{\mathcal{M}}} P(\mathbf{E}, \mathbf{Y}_{\mathcal{M}}) d\mathbf{Y}_{\mathcal{M}} = \int_{\mathbf{Y}_{\mathcal{M}}} P(\mathbf{E} | \mathbf{Y}_{\mathcal{M}}) P(\mathbf{Y}_{\mathcal{M}}) d\mathbf{Y}_{\mathcal{M}} \quad (3.2)$$

El último término integral de la ecuación (3.2) se puede escribir además como muestra la ecuación (3.3)

$$\int_{\mathbf{Y}_{\mathcal{M}}} P(\mathbf{E} | \mathbf{Y}_{\mathcal{M}}) P(\mathbf{Y}_{\mathcal{M}}) d\mathbf{Y}_{\mathcal{M}} = \mathbb{E}[P(\mathbf{E} | \mathbf{Y}_{\mathcal{M}})] \quad (3.3)$$

Por tanto, a partir de las ecuaciones (3.2) y (3.3) se llega a la relación que existe entre la pdf marginal y las pdf's condicionales de una distribución bivariada.

$$P(\mathbf{E}) = \mathbb{E}[P(\mathbf{E} | \mathbf{Y}_{\mathcal{M}})] \quad (3.4)$$

De esta relación entre las distribuciones de densidad de probabilidad, se pueden derivar, al menos, las relaciones que hay entre sus dos primeros momentos, relaciones que reciben el nombre de Ley de la Esperanza Total (**TEL**) y Ley de la Varianza Total (**TVL**).

La TEL se deduce de forma directa tomando en consideración la ecuación (3.4). Su expresión viene dada por la ecuación (3.5)

$$\mathbb{E}[\mathbf{E}] = \mathbb{E}_{\mathbf{Y}_{\mathcal{M}}}[\mathbb{E}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}]] \quad (3.5)$$

Respecto a la TVL, su expresión se puede obtener mediante la aplicación de la TEL, a la definición de varianza dada por la ecuación $\mathbb{V}[\mathbf{E}] = \mathbb{E}[\mathbf{E}^2] - \mathbb{E}^2[\mathbf{E}]$. Mediante el desarrollo y reordenación de términos se llega a la ecuación (3.6) que define la TVL.

$$\mathbb{V}[\mathbf{E}] = \mathbb{E}_{\mathbf{Y}_{\mathcal{M}}}[\mathbb{V}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}]] + \mathbb{V}_{\mathbf{Y}_{\mathcal{M}}}[\mathbb{E}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}]] \quad (3.6)$$

En el caso específico en el que los errores no presenten sesgo en sus distribuciones condicionales, es decir cuando $\mathbb{E}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}] = 0$, la TVL se reduce a la expresión de la ecuación (3.7)

$$\mathbb{V}[\mathbf{E}] = \mathbb{E}_{\mathbf{Y}_{\mathcal{M}}}[\mathbb{V}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}]] \quad (3.7)$$

En resumen, las llamadas Leyes Totales (TLs) definidas por las ecuaciones (3.5) y (3.6) o (3.7), son relaciones que deben mantenerse en una pdf bivariada, entre las distribuciones condicionales y su respectiva marginal.

3.2.2. Casuística en la aplicación de las Leyes Totales

Esta sección tiene por objeto mostrar las implicaciones de las TLs en cuatro casos generales de distribución de la nube de errores $\{y_{j_{\mathcal{M}}}, e_j\} \forall j \in \{1, \dots, n\}$ representada en el plano $Y_{\mathcal{M}} - Y$. Esto se va a realizar mediante la representación y explicación de una serie de cuatro esquemas. En cada esquema se representan las distribuciones condicionales (en azul) obtenidas a partir del correspondiente modelo de error que se está asumiendo. Para obtener el máximo entendimiento de la figura, hay que considerar que la distribución conjunta $P(\mathbf{E}, \mathbf{Y}_{\mathcal{M}})$, se define a partir de la modelación del sesgo y la varianza de sus distribuciones condicionales $P(\mathbf{E} | \mathbf{Y}_{\mathcal{M}})$, mediante un modelo de error en el que algunos de sus parámetros van a quedar restringidos por el cumplimiento de las TLs, con el propósito de mantener la coherencia entre pdf's condicionales y la pdf marginal (en verde).

En el primero de los casos, Figura 3.4-a, se muestra un esquema de los modelos de varianza y sesgo que representan al método de los mínimos cuadrados (**SLS**), en el que los errores son homocedásticos con media nula. Según este esquema, las distribuciones condicionales son copia de la distribución marginal. Lo cual implica que $\mathbb{E}[\mathbf{E}] = \mathbb{E}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}] = 0$ y que $\mathbb{V}[\mathbf{E}] = \mathbb{V}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}]$, es decir, con las hipótesis del método SLS, las Leyes Totales se cumplen de forma natural.

En el segundo esquema, Figura 3.4-b, se representa el caso de unos modelos de sesgo nulo y varianza no constante, que corresponden al método **WLS** (visto en la sección 2.3.3). ¿Cómo se cumplen las TLs en este esquema? Si bien la TEL se cumple teóricamente por defecto, como en el SLS, para el cumplimiento

de la TVL el modelo de varianza debe ser tal, que permita el cumplimiento de la igualdad dada por la ecuación (3.7), es decir, $\mathbb{V}[\mathbf{E}] = \mathbb{E}[\mathbb{V}[\mathbf{E} | \mathbf{Y}_M]]$.

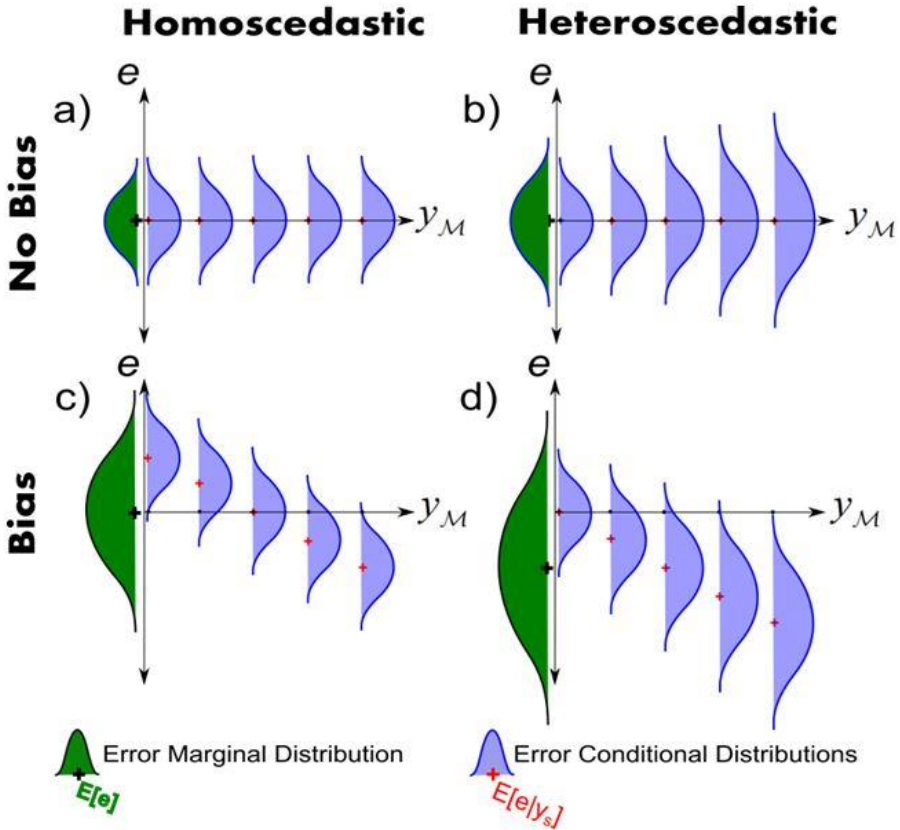


Figura 3.4. Esquemas de descomposición de la distribución conjunta $P(\mathbf{Y}_M, \mathbf{Y})$ en sus distribuciones condicionales (azul) y marginal (verde), bajo la asunción de 4 hipótesis diferentes, para los modelos de varianza y sesgo de los errores

En el esquema de la Figura 3.4-c, se representa el caso de un modelo de error en el que la varianza de los errores es constante, no así el sesgo. Se puede observar como el soporte de la distribución marginal es mucho mayor que el de las condicionales, lo que es debido a la existencia del sesgo, que desplaza las pdf's condicionales respecto al cero. Nótese, que aunque este modelo de error presenta sesgo, ello no impide que la marginal de los errores no lo tenga (su media es nula) debido a un engañoso efecto de autocompensación. Respecto a

las TLs, el cumplimiento de la TEL necesita que el modelo de sesgo de los errores permita el cumplimiento de la ecuación (3.5). El cumplimiento de la TVL necesita que el modelo de sesgo de los errores permita el cumplimiento de la igualdad $\mathbb{V}[\mathbf{E}] = \mathbb{V}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}] + \mathbb{V}_{\mathbf{Y}_{\mathcal{M}}}[\mathbb{E}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}]]$.

Finalmente, en el esquema de la Figura 3.4-d se muestra el caso más general en el cual ni la varianza ni el sesgo de los errores son constantes. En este caso, los respectivos modelos deben permitir que se cumplan las TLs en su forma completa, dadas por las ecuaciones (3.5) y (3.6). En la sección 3.3.3 se describirá como acondicionar los modelos de varianza y/o sesgo de los errores para que permitan el cumplimiento de las Leyes Totales.

3.3. Un nuevo modelo de error

En esta sección se van a definir las ecuaciones que forman parte de los diferentes elementos que componen el modelo de error que propone esta Tesis. El guión que sigue el diseño de este nuevo modelo de error es similar al guión que conduce al desarrollo del resto de modelos de error sofisticados (e.g. GL de Schoups y Vrugt (2010)), esto es, la relajación de algunas o todas las hipótesis del método SLS. Esta relajación es necesaria debido a que además de los errores aleatorios de observación ya contemplados por el SLS, existen otras fuentes de incertidumbre que se deben considerar, cuyos errores son, con diferencia, mucho más importantes que los meramente observacionales. Siendo el mismo guión, las diferencias entre distintos modelos de error surgen por las distintas formas con que se lleva a cabo la relajación de las hipótesis.

Con el propósito de dejar el capítulo lo más autocontenido posible, se volverá a definir el problema de inferencia, permitiendo así un mejor entendimiento de la nomenclatura empleada, sin necesidad de retroceder a capítulos anteriores.

3.3.1. Definición del problema de inferencia

Considérese el vector de observaciones de una variable de interés $\tilde{\mathbf{Y}} = \{\tilde{y}_1, \dots, \tilde{y}_n\}$. Dicho vector resume la respuesta de un sistema natural analizado, llámese \mathfrak{S} , frente a la excitación producida por k señales o inputs

variables en el tiempo $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, donde los vectores columna se definen para cada t como $\mathbf{b}_t = \{b_{t_1}, \dots, b_{t_k}\}$. Los inputs son el equivalente a los regresores, en un análisis clásico de regresión. Para explicar el sistema analizado, en base a los datos observados sobre él, se emplea un modelo matemático computacional, llámese $\mathcal{M}(\cdot)$. De forma simplificada, se puede resumir el proceso de análisis mediante la ecuación (3.8).

$$\tilde{\mathbf{Y}} \leftarrow \mathcal{M}(\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{x}_0, \tilde{\mathbf{B}}) + \mathbf{E} \quad (3.8)$$

en la que $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_d\}$ es el $d \times 1$ -vector de parámetros del modelo hidrológico, $\boldsymbol{\Omega} = \{\Omega_1, \dots, \Omega_r\}$ es el $r \times 1$ -vector de parámetros del modelo de error, \mathbf{x}_0 es el vector que almacena los valores de las variables de estado del sistema al comienzo de la simulación ($t=0$), $\tilde{\mathbf{B}}$ es la matriz de dimensiones $k \times n$ que contiene las mediciones en el tiempo de los k inputs que excitan el sistema y \mathbf{E} es un vector de dimensión n con errores aditivos. El vector de errores depende, por tanto, del modelo $\mathcal{M}(\cdot)$, de los valores de sus parámetros, de los estados iniciales de las variables de estado del sistema y naturalmente, de los inputs que excitan o “fuerzan” al modelo.

Sin embargo, en un caso real, otros factores que influyen en los errores podrían ser considerados. Por ejemplo, el modelo $\mathcal{M}(\cdot)$ puede presentar errores de especificación o estructurales $\mathbf{E}_{\mathcal{M}}$. Así mismo los inputs $\tilde{\mathbf{B}}$ pueden contener errores $\mathbf{E}_{\tilde{\mathbf{B}}}$, que van más allá de los puramente observacionales. Los parámetros del modelo estimados a priori, si los hubiere, pueden presentar incertidumbre en su valor. Las condiciones iniciales podrían tener error en su estimación $\mathbf{E}_{\mathbf{x}_0}$. Con todo ello, el vector de errores se representaría por la ecuación (3.9)

$$\mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{E}_{\tilde{\mathbf{Y}}}, \mathbf{x}_0 + \mathbf{E}_{\mathbf{x}_0}, \tilde{\mathbf{B}} + \mathbf{E}_{\tilde{\mathbf{B}}}, \mathbf{E}_{\mathcal{M}}) = \tilde{\mathbf{Y}}(\mathbf{E}_{\tilde{\mathbf{Y}}}) - \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{x}_0 + \mathbf{E}_{\mathbf{x}_0}, \tilde{\mathbf{B}} + \mathbf{E}_{\tilde{\mathbf{B}}}, \mathbf{E}_{\mathcal{M}}) \quad (3.9)$$

donde, $\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{x}_0 + \mathbf{E}_{\mathbf{x}_0}, \tilde{\mathbf{B}} + \mathbf{E}_{\tilde{\mathbf{B}}}, \mathbf{E}_{\mathcal{M}}) = \{y_{1,\mathcal{M}}, \dots, y_{n,\mathcal{M}}\}$ son los resultados de la simulación, que el modelo proporciona para cada t , sobre la variable de interés \mathbf{Y} .

Considérese por conveniencia, que los estados iniciales del sistema adoptan los valores correctos como producto de un periodo de *calentamiento* del modelo, de longitud suficiente. Así mismo asúmase la hipótesis de que las observaciones $\tilde{\mathbf{Y}}$ presentan un error de medición despreciable de carácter puramente aleatorio. Supóngase así mismo, que todos los parámetros del modelo son inferidos, con lo cual, la incertidumbre que presenten será función del resto de fuentes de incertidumbre y del método de inferencia utilizado. Siendo este el caso, el vector de errores se puede reescribir según la ecuación (3.10)

$$\mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega}, \tilde{\mathbf{B}} + \mathbf{E}_{\tilde{\mathbf{B}}}, \mathbf{E}_{\mathcal{M}}) = \tilde{\mathbf{Y}} - \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{x}_0, \tilde{\mathbf{B}} + \mathbf{E}_{\tilde{\mathbf{B}}}, \mathbf{E}_{\mathcal{M}}) \quad (3.10)$$

Esta ecuación necesita de un proceso de inferencia conjunta, con el fin de determinar los vectores de parámetros $\{\boldsymbol{\theta}, \boldsymbol{\Omega}\}$. Por simplicidad, la ecuación (3.10) y sus términos, se escribirán en adelante como $\mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega}) = \tilde{\mathbf{Y}} - \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \boldsymbol{\Omega}) = \{e_1(\boldsymbol{\theta}, \boldsymbol{\Omega}), \dots, e_n(\boldsymbol{\theta}, \boldsymbol{\Omega})\}$, sin dejar de considerar por ello que los errores se ven afectados por las fuentes de incertidumbre estructural y de los inputs.

Relajación de la hipótesis de errores con media nula

Los errores $\mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega})$, que representan de forma agregada a todas las incertidumbres del proceso de inferencia, pueden ser descompuestos en dos componentes, como expresa la ecuación (3.11)

$$\mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega}) = \boldsymbol{\mu} + \boldsymbol{\varepsilon} \quad (3.11)$$

El componente $\boldsymbol{\mu} = \{\mu_1(\boldsymbol{\theta}, \boldsymbol{\Omega}), \dots, \mu_n(\boldsymbol{\theta}, \boldsymbol{\Omega})\}$, al que se llamará sesgo, representa a la parte sistemática de los errores modelada generalmente de forma determinista. El segundo componente $\boldsymbol{\varepsilon} = \{\varepsilon_1(\boldsymbol{\theta}, \boldsymbol{\Omega}), \dots, \varepsilon_n(\boldsymbol{\theta}, \boldsymbol{\Omega})\}$, representa a la parte aleatoria de los errores debida a la aleatoriedad inherente a los procesos simulados, así como a la información que no quede explicada por el término de

sesgo y que el modelo tampoco ha sido capaz de almacenar en sus parámetros. Esta información remanente es la causa de que persista la dependencia serial en ε .

Por tanto, los errores y sus dos componentes tienen las siguientes propiedades:

- 1- $\mathbb{E}[\mathbf{E}|\mathbf{Y}_{\mathcal{M}}] = \boldsymbol{\mu}$
- 2- $\mathbb{E}[\mathbf{E}] = \mathbf{0} \not\approx \boldsymbol{\mu} = \bar{\mathbf{0}}$
- 3- $\mathbb{E}[\boldsymbol{\varepsilon}|\mathbf{Y}_{\mathcal{M}}] = \bar{\mathbf{0}}$
- 4- $\mathbb{E}[\varepsilon_j \varepsilon_{j+1}] \neq 0$

Modelo de sesgo de 3 parámetros

En el caso de la inferencia que ocupa esta Tesis, se define un sencillo modelo de sesgo de los errores compuesto por dos leyes lineales cuya variable independiente son los resultados del modelo $\mathbf{Y}_{\mathcal{M}}$. Dicho modelo de sesgo se define $\forall j \in \{1, \dots, n\}$ según la ecuación (3.12)

$$\begin{aligned} \mu_j &= \gamma & \forall y_{j_{\mathcal{M}}} \leq y_{\mathcal{M}}^* \\ \mu_j &= \gamma + \tau (y_{j_{\mathcal{M}}} - y_{\mathcal{M}}^*) & \forall y_{j_{\mathcal{M}}} > y_{\mathcal{M}}^* \end{aligned} \quad (3.12)$$

donde $\{\gamma, \tau, y_{\mathcal{M}}^*\}$ son los parámetros del modelo de error que se deben inferir. Esta doble función pretende considerar, de forma sencilla, los dos diferentes comportamientos del error que se espera que ocurran con caudales bajos y con caudales altos.

Modelo de sesgo de 4 parámetros

Así mismo, se define un segundo modelo de sesgo de los errores algo más flexible (4 parámetros en lugar de 3), compuesto igualmente por dos leyes lineales cuya variable independiente son los resultados del modelo $\mathbf{Y}_{\mathcal{M}}$. Dicho modelo de sesgo se define $\forall j \in \{1, \dots, n\}$ según la ecuación (3.13)

$$\begin{aligned} \mu_j &= \gamma + \delta y_{j_M} & \forall y_{j_M} &\leq y_M^* \\ \mu_j &= \gamma + (\delta - \tau) y_M^* + \tau y_{j_M} & \forall y_{j_M} &> y_M^* \end{aligned} \quad (3.13)$$

donde $\{\gamma, \delta, \tau, y_M^*\}$ son los parámetros del modelo de error que se deben inferir.

Relajación de la hipótesis de errores con varianza constante

Al igual que hicieron Sorooshian y Dracup (1980) y años más tarde Schoups y Vrugt (2010) o Evin et al. (2013), el modelo de error que se está definiendo va a considerar la modelación de la varianza no constante de los errores, mediante un método directo sin transformación de las variables originales, por las ventajas que ello conlleva, las cuales se explicaron en la sección 2.3.3 del capítulo anterior.

Para ello, se define la desviación estándar de los errores $\mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega})$ como función lineal del resultado del modelo hidrológico \mathbf{Y}_M , siendo de hecho el mismo modelo de varianza expresado por la ecuación (2.22), si bien aquí se asume un valor $\psi = 1$, con lo cual el modelo de desviación estándar, queda expresado por la ecuación (3.14)

$$\boldsymbol{\sigma} = \boldsymbol{\alpha} + \boldsymbol{\kappa} \mathbf{Y}_M(\boldsymbol{\theta}, \boldsymbol{\Omega}) \quad (3.14)$$

donde $\{\boldsymbol{\alpha}, \boldsymbol{\kappa}\}$ son parámetros del modelo de error que se deben inferir.

Relajación de la hipótesis de errores independientes

Para la consideración de la estructura de dependencia serial de los errores se asume, como es común en la literatura, un modelo lineal autorregresivo de primer orden, es decir, un modelo AR(1). El modelo de dependencia, tradicionalmente, se ha definido sobre unos errores cuya varianza se ha estabilizado previamente mediante un método transformacional. Sin embargo, Schoups y Vrugt (2010) aplicaron el modelo AR(1) sobre los errores heterocedásticos sin estabilizar (sección 2.3.5).

Evin et al. (2013) demostraron que al aplicar un modelo AR(1) sobre una serie de errores con varianza no constante, se producen unos incrementos indeseables de la autocorrelación, especialmente en las ramas de recesión de los

hidrogramas (sección 2.3.6). Este aumento de correlación, según los autores, es el origen de la exagerada inflación de las bandas de incertidumbre de la predicción. Por tanto, Evin et al. (2013) proponen la estabilización previa de la varianza de los errores aplicando un método directo, es decir, ponderando los errores mediante unos factores o pesos $w_j = \sigma_{el y_{j,M}}^{-1}$, de igual manera a como se hace en el método WLS, visto en el capítulo 2. Proponen, por tanto, la estandarización de los errores $\mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega})$ mediante su desviación típica $\boldsymbol{\sigma}$, dada por la ecuación (3.14). Tras la estandarización se aplicará el modelo AR(1) sobre los errores estandarizados. Dado que en esta Tesis también se está relajando, en su modelo de error completo, la hipótesis de media nula de los errores, la estandarización va a considerar además la eliminación del sesgo $\boldsymbol{\mu}$ de los errores. Así, se definen los errores estandarizados $\boldsymbol{\eta} = \{\eta_1(\boldsymbol{\theta}, \boldsymbol{\Omega}), \dots, \eta_n(\boldsymbol{\theta}, \boldsymbol{\Omega})\}$, mediante la ecuación (3.15)

$$\boldsymbol{\eta} = \boldsymbol{\sigma}^{-1}(\mathbf{E} - \boldsymbol{\mu}) = \boldsymbol{\sigma}^{-1}\boldsymbol{\varepsilon} \quad (3.15)$$

Sobre los errores estandarizados se aplica el modelo lineal de autorregresión de primer orden. Según Box et al. (1994), los modelos lineales autorregresivos de orden p se pueden expresar de forma compacta mediante la ecuación (3.16)

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right) \boldsymbol{\eta} = \phi_p(B) \boldsymbol{\eta} = \mathbf{z} \quad (3.16)$$

Donde ϕ_i es el coeficiente autorregresivo de orden i , $\phi_p(B)$ es el operador polinomial de retardos, B^i es el operador que produce i -retardos en el término de la serie sobre el que actúa, por ejemplo $B^i \eta_j = \eta_{j-i}$. El vector $\mathbf{z} = \{z_1(\boldsymbol{\theta}, \boldsymbol{\Omega}), \dots, z_n(\boldsymbol{\theta}, \boldsymbol{\Omega})\}$ almacena las innovaciones no estandarizadas. Las innovaciones \mathbf{z} son errores con media nula, varianza constante de valor σ_z^2 , e independientes, es decir, son *ruido blanco*. El valor de σ_z^2 se calcula según Box et al. (1994), mediante la ecuación (3.17).

$$\sigma_z^2 = \sigma_\eta^2 \left(1 - \sum_{i=1}^p \rho_i \phi_i \right) \quad (3.17)$$

donde ρ_i son los coeficientes de correlación entre dos errores estandarizados $\eta_j(\boldsymbol{\theta}, \boldsymbol{\Omega}), \dots, \eta_{j+i}(\boldsymbol{\theta}, \boldsymbol{\Omega})$, separados por un desfase (*lag*) i . El término σ_η^2 es la varianza marginal de los errores estandarizados.

El valor σ_z^2 ha de considerarse como un parámetro más del modelo de error, necesario para la generación de las bandas de incertidumbre (sección 3.5), tanto en el periodo de calibración, como en periodos en los que no se disponga de observaciones de la variable de interés \mathbf{Y} .

Puesto que el modelo de error que se propone va a contemplar un modelo lineal AR(1), las ecuaciones (3.16) y (3.17) se particularizan para $p=1$ obteniendo las ecuaciones (3.18) y (3.19) respectivamente.

$$\eta_j - \phi_1 \eta_{j-1} = z_j \quad \forall j \in \{1, \dots, n\} \quad (3.18)$$

$$\sigma_z^2 = \sigma_\eta^2 (1 - \phi_1^2) \quad (3.19)$$

Relajación de la hipótesis de Normalidad

Finalmente, se considera la posibilidad de que las innovaciones (o residuos) \mathbf{z} se distribuyan de forma no Gaussiana. Para ello se adopta la función $P_{\text{SEP}}(0,1,\xi,\beta)$ que proponen Schoups y Vrugt (2010), dada por la ecuación (2.28). Puesto que $P_{\text{SEP}}(0,1,\xi,\beta)$ considera las innovaciones ya estandarizadas, se ha de estandarizar \mathbf{z} , según expresa la ecuación (3.20)

$$\mathbf{a} = \sigma_z^{-1} \mathbf{z} \quad (3.20)$$

El vector $\mathbf{a} = \{a_1(\boldsymbol{\theta}, \boldsymbol{\Omega}), \dots, a_n(\boldsymbol{\theta}, \boldsymbol{\Omega})\}$ almacena una muestra de *ruido blanco* estandarizado, el cual se va a considerar que se distribuye según una pdf $P_{\text{SEP}}(0,1,\xi,\beta)$.

Resumen del nuevo modelo de error

El modelo de error configurado permite la consideración de errores \mathbf{E} , con media no nula, con varianza no constante σ^2 , dependencia serial lineal y distribución diferente a la Normal (aunque si los datos así lo sugieren, la SEP puede adoptar la forma Gaussiana). El vector de parámetros de error, según sea el modelo de sesgo adoptado, es de dimensión $r=8$, $\mathbf{\Omega} = \{\gamma, \tau, y_{\mathcal{M}}^*, \alpha, \kappa, \phi_1, \xi, \beta\}$ o de dimensión $r=9$, $\mathbf{\Omega} = \{\gamma, \delta, \tau, y_{\mathcal{M}}^*, \alpha, \kappa, \phi_1, \xi, \beta\}$.

La relación existente entre errores e innovaciones, según el modelo de error propuesto, queda definida por la ecuación (3.21)

$$\mathbf{a} = \sigma_z^{-1} \phi_p(B) (\sigma^{-1} (\mathbf{E} - \boldsymbol{\mu})) \quad (3.21)$$

3.3.2. Derivación de la función de verosimilitud

Como ya se explicó en la sección 2.2.2, para la implementación numérica del teorema de Bayes, con el objetivo de inferir la distribución posterior de los parámetros de un modelo, solamente es necesaria la especificación de la pdf prior de los parámetros y la construcción de la función de verosimilitud. Puesto que en esta Tesis se asume una pdf prior plana, solamente los datos, a través de la función de verosimilitud, condicionan la distribución posterior de los parámetros, es decir, se cumple que $P(\boldsymbol{\theta}, \mathbf{\Omega} | \mathbf{Y}) \propto P(\mathbf{Y} | \boldsymbol{\theta}, \mathbf{\Omega}) = \ell(\boldsymbol{\theta}, \mathbf{\Omega} | \mathbf{Y})$.

Bajo la hipótesis de un modelo de error aditivo, siguiendo el mismo razonamiento de la sección 3.1.2, se pueden escribir la pdf posterior y la función de verosimilitud en función de los errores como muestra la ecuación (3.22)

$$P(\boldsymbol{\theta}, \mathbf{\Omega} | \mathbf{E}) \propto P(\mathbf{E} | \boldsymbol{\theta}, \mathbf{\Omega}) = \ell(\boldsymbol{\theta}, \mathbf{\Omega} | \mathbf{E}) \quad (3.22)$$

Por conveniencia para los desarrollos que siguen, el vector de errores \mathbf{E} se dividirá en dos vectores, siendo $\mathbf{E}_1 = \{e_1\}$ y $\mathbf{E}_2 = \{e_2, \dots, e_n\}$. Aplicando la ley de probabilidades condicionales, se puede escribir la ecuación (3.23)

$$P(\mathbf{E} | \boldsymbol{\theta}, \mathbf{\Omega}) = P(\mathbf{E}_1 | \boldsymbol{\theta}, \mathbf{\Omega}) P(\mathbf{E}_2 | \mathbf{E}_1, \boldsymbol{\theta}, \mathbf{\Omega}) \quad (3.23)$$

La pdf $P(\mathbf{E}|\boldsymbol{\theta}, \boldsymbol{\Omega})$ viene definida por el modelo de error que se ha configurado en la sección anterior, con el cual se llega a la obtención de unos residuos o innovaciones de tal forma que $a_j(\boldsymbol{\theta}, \boldsymbol{\Omega}) \stackrel{D}{\sim} P_{\text{SEP}}(0, 1, \xi, \beta) \forall j \in \{1, \dots, n\}$, donde cada $a_j(\boldsymbol{\theta}, \boldsymbol{\Omega})$ es obtenido mediante la ecuación (3.21). Por tanto, la transformación $\mathbf{E} \rightarrow \mathbf{a}$ debe aplicarse también a sus pdf's. En primer lugar, se realizará la transformación $\mathbf{E} \rightarrow \boldsymbol{\varepsilon} \rightarrow \boldsymbol{\eta}$ sobre las pdf's de \mathbf{E}_1 y \mathbf{E}_2 . Se procede con \mathbf{E}_1 , en la ecuación (3.24)

$$P(e_1 - \mu_1 | \boldsymbol{\theta}, \boldsymbol{\Omega}) = P(\varepsilon_1 | \boldsymbol{\theta}, \boldsymbol{\Omega}) = \left| \frac{\partial \eta_1}{\partial \varepsilon_1} \right| P(\eta_1 | \boldsymbol{\theta}, \boldsymbol{\Omega}) = \sigma_1^{-1} P(\eta_1 | \boldsymbol{\theta}, \boldsymbol{\Omega}) \quad (3.24)$$

Procediendo de igual forma con \mathbf{E}_2 , se escribe la ecuación (3.25)

$$\begin{aligned} P(\mathbf{E}_2 - \boldsymbol{\mu}_2 | e_1 - \mu_1, \boldsymbol{\theta}, \boldsymbol{\Omega}) &= P(\boldsymbol{\varepsilon}_2 | \varepsilon_1, \boldsymbol{\theta}, \boldsymbol{\Omega}) = \det|\mathbb{J}| P(\boldsymbol{\eta}_2 | \eta_1, \boldsymbol{\theta}, \boldsymbol{\Omega}) = \\ &= \prod_{j=2}^n \left| \frac{\partial \eta_j}{\partial \varepsilon_j} \right| P(\boldsymbol{\eta}_2 | \eta_1, \boldsymbol{\theta}, \boldsymbol{\Omega}) = \prod_{j=2}^n \sigma_j^{-1} P(\boldsymbol{\eta}_2 | \eta_1, \boldsymbol{\theta}, \boldsymbol{\Omega}) = \\ &= \prod_{j=2}^n \sigma_j^{-1} P(\eta_j | \eta_{1:(j-1)}, \boldsymbol{\theta}, \boldsymbol{\Omega}) \end{aligned} \quad (3.25)$$

En la ecuación (3.25), el término $\det|\mathbb{J}|$ es el valor absoluto del determinante de la matriz Jacobiana, correspondiente a la función de transformación de estandarización de la ecuación (3.15).

Sustituyendo las ecuaciones (3.24) y (3.25) en la ecuación (3.23) se escribe la ecuación (3.26)

$$P(\mathbf{E} - \boldsymbol{\mu} | \boldsymbol{\theta}, \boldsymbol{\Omega}) = P(\boldsymbol{\varepsilon} | \boldsymbol{\theta}, \boldsymbol{\Omega}) = \sigma_1^{-1} P(\eta_1 | \boldsymbol{\theta}, \boldsymbol{\Omega}) \prod_{j=2}^n \sigma_j^{-1} P(\eta_j | \eta_{1:(j-1)}, \boldsymbol{\theta}, \boldsymbol{\Omega}) \quad (3.26)$$

Tomando en consideración la relación de dependencia entre dos errores estandarizados $\{\eta_{j-1}(\boldsymbol{\theta}, \boldsymbol{\Omega}), \eta_j(\boldsymbol{\theta}, \boldsymbol{\Omega})\}$, definida por la ecuación (3.18), se procede a la transformación $\boldsymbol{\eta} \rightarrow \mathbf{z}$, tal como muestra la ecuación (3.27)

$$P(\boldsymbol{\varepsilon} | \boldsymbol{\theta}, \boldsymbol{\Omega}) = \sigma_1^{-1} P(\eta_1 | \boldsymbol{\theta}, \boldsymbol{\Omega}) \prod_{j=2}^n \sigma_j^{-1} P(z_j | \boldsymbol{\theta}, \boldsymbol{\Omega}) \quad (3.27)$$

La ecuación (3.27) se puede aproximar por la ecuación (3.28) para valores de n de moderados a grandes (Schoups y Vrugt (2010)), como es el caso de las series de modelación hidrológica.

$$P(\boldsymbol{\varepsilon} | \boldsymbol{\theta}, \boldsymbol{\Omega}) \approx \prod_{j=1}^n \sigma_j^{-1} P(z_j | \boldsymbol{\theta}, \boldsymbol{\Omega}) \quad (3.28)$$

Finalmente, se aplica la transformación de estandarización de las innovaciones $\mathbf{z} \rightarrow \mathbf{a}$, mediante la ecuación (3.29), en la cual ya se incorpora la hipótesis de que $a_j(\boldsymbol{\theta}, \boldsymbol{\Omega}) \stackrel{\mathcal{D}}{\sim} P_{\text{SEP}}(0, 1, \xi, \beta) \forall j \in \{1, \dots, n\}$.

$$P(\boldsymbol{\varepsilon} | \boldsymbol{\theta}, \boldsymbol{\Omega}) = \ell(\boldsymbol{\theta}, \boldsymbol{\Omega} | \boldsymbol{\varepsilon}) \approx \prod_{j=1}^n \sigma_j^{-1} \sigma_z^{-1} P_{\text{SEP}}(a_j | \boldsymbol{\theta}, \boldsymbol{\Omega}) \quad (3.29)$$

Sustituyendo en (3.29) la expresión de la $P_{\text{SEP}}(0, 1, \xi, \beta)$, se escribe la función de verosimilitud en la ecuación (3.30)

$$\ell(\boldsymbol{\theta}, \boldsymbol{\Omega} | \boldsymbol{\varepsilon}) \approx \sigma_z^{-n} \prod_{j=1}^n \sigma_j^{-1} \frac{2\sigma_\xi}{\xi + \xi^{-1}} w_\beta \exp \left(-c_\beta \left| \frac{\mu_\xi + \sigma_\xi a_j}{\xi^{\text{sgn}(\mu_\xi + \sigma_\xi a_j)}} \right|^{\frac{2}{1+\beta}} \right) \quad (3.30)$$

Para evitar los problemas numéricos, ya explicados en el capítulo anterior, en la inferencia se suele emplear el logaritmo de la verosimilitud. Tomando logaritmos en (3.30), se llega a la ecuación (3.31)

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Omega} | \boldsymbol{\varepsilon}) \approx n \log \frac{2\sigma_\xi w_\beta}{\sigma_z (\xi + \xi^{-1})} - \sum_{j=1}^n \log \sigma_j - c_\beta \sum_{j=1}^n \left| \frac{\mu_\xi + \sigma_\xi a_j}{\xi^{\text{sgn}(\mu_\xi + \sigma_\xi a_j)}} \right|^{\frac{2}{1+\beta}} \quad (3.31)$$

donde los factores $c_{\beta}, w_{\beta}, \mu_{\xi}, \sigma_{\xi}$ son función de los parámetros $\{\beta, \xi\}$ de la SEP, según Fernandez y Steel (1998) o Schoups y Vrugt (2010), σ_z se obtiene según la ecuación (3.19) y σ_j se obtiene de la ecuación (3.14).

A la función de verosimilitud generalizada representada por la ecuación (3.31) se le denominará **GL++NTL**, para diferenciarla de la función **GL** de Schoups y Vrugt (2010). Las principales diferencias visibles entre las ecuaciones (2.30) y (3.31) se deben a la modificación propuesta por Evin et al. (2013). Fruto de esa modificación aparece la constante σ_z (varianza constante de las innovaciones) en el primer término de la parte derecha de la ecuación. Por la misma causa, aparece en el segundo término de la ecuación, la variable σ_j (varianza de los errores), en lugar de la variable σ_{z_j} (varianza de las innovaciones) que aparece en GL (ecuación (2.30)).

3.4. Inclusión de las Leyes Totales en el modelo de error

La hipótesis que origina todo el desarrollo de este capítulo 3 consiste en considerar que la aplicación del método directo de modelar la varianza de los errores, tal cual fue concebido por los trabajos de 2010 y posteriores, no garantiza que al inferir los parámetros de error $\{\alpha, \kappa\}$, no se puedan producir situaciones paradójicas como las de la Figura 3.2, en las que no se está manteniendo la coherencia estadística. Lo que ocurre con el modelo de varianza, ocurre de igual manera con un modelo de sesgo de las distribuciones condicionales. La estimación de los parámetros de error, con la coherencia estadística necesaria mostrada en la Figura 3.4, pasa por la consideración de las Leyes Totales junto con los modelos que se consideren para la varianza y el sesgo de los errores.

Por tanto, **el modelo de error definido en la sección 3.3.1 y su correspondiente función de verosimilitud GL++NTL definida en la sección 3.3.2, quedan incompletos sin los resultados obtenidos en la presente sección 3.4.** Con la adición de los resultados de esta sección, se obtienen los nuevos modelos de error que propone esta Tesis, **GL++**, en el caso de no incluir el modelo de sesgo y **GL++Bias**, en el caso de incluirlo.

3.4.1. Inclusión de las TLs al relajar la hipótesis de media nula.

Modelo de sesgo de 3 parámetros

El modelo de sesgo $\boldsymbol{\mu}$ propuesto en la ecuación (3.12), cuyos parámetros son $\{\gamma, \tau, y_{\mathcal{M}}^*\}$, permite modelar la posición central de las diferentes distribuciones de densidad condicional del error, pertenecientes a la bivariada $P(\mathbf{E}, \mathbf{Y}_{\mathcal{M}})$, esto es $\boldsymbol{\mu} = \mathbb{E}[\mathbf{E} | \mathbf{Y}_{\mathcal{M}}]$. Si las pdf condicionales $P(\mathbf{E} | \mathbf{Y}_{\mathcal{M}})$ han de formar parte de la bivariada, se debe cumplir la TEL (ecuación (3.5)), la cual considerando el modelo de sesgo asumido se puede reescribir como $\mathbb{E}[\mathbf{E}] = \mathbb{E}[\boldsymbol{\mu}]$. Si se aplica $\mathbb{E}[\boldsymbol{\mu}]$ al modelo de sesgo definido en (3.12), se obtienen las ecuaciones de (3.32)

$$\begin{aligned} \mathbb{E}[\boldsymbol{\mu}_1] &= \gamma & \boldsymbol{\mu}_1 &= \{\mu_j\} & \forall j / y_{j_{\mathcal{M}}} \leq y_{\mathcal{M}}^* \\ \mathbb{E}[\boldsymbol{\mu}_2] &= \gamma + \tau \left(\mathbb{E}[\mathbf{Y}_{2_{\mathcal{M}}}] - y_{\mathcal{M}}^* \right) & \boldsymbol{\mu}_2 &= \{\mu_j\}, \mathbf{Y}_{2_{\mathcal{M}}} = \{y_{j_{\mathcal{M}}}\} & \forall j / y_{j_{\mathcal{M}}} > y_{\mathcal{M}}^* \end{aligned} \quad (3.32)$$

Siendo tres los parámetros y dos las ecuaciones, solamente un parámetro será independiente. Nótese entonces, que no todos los parámetros del modelo de sesgo son libremente inferidos. Si se deja al parámetro umbral de caudal $y_{\mathcal{M}}^*$, como parámetro independiente y se aplica la TEL a (3.32), esto es, $\mathbb{E}[\mathbf{E}] = \mathbb{E}[\boldsymbol{\mu}]$, los otros dos parámetros adoptan los valores de la ecuación (3.33)

$$\begin{aligned} \gamma &= \mathbb{E}[\mathbf{E}_1] & \mathbf{E}_1 &= \{e_j\} & \forall j / y_{j_{\mathcal{M}}} \leq y_{\mathcal{M}}^* \\ \tau &= \frac{\mathbb{E}[\mathbf{E}_2] - \gamma}{\mathbb{E}[\mathbf{Y}_{2_{\mathcal{M}}}] - y_{\mathcal{M}}^*} = \frac{\mathbb{E}[\mathbf{E}_2] - \mathbb{E}[\mathbf{E}_1]}{\mathbb{E}[\mathbf{Y}_{2_{\mathcal{M}}}] - y_{\mathcal{M}}^*} & \mathbf{E}_2 &= \{e_j\}, \mathbf{Y}_{2_{\mathcal{M}}} = \{y_{j_{\mathcal{M}}}\} & \forall j / y_{j_{\mathcal{M}}} > y_{\mathcal{M}}^* \end{aligned} \quad (3.33)$$

En resumen, de los tres parámetros del modelo de sesgo definido por la ecuación (3.12), los parámetros $\{\gamma, \tau\}$ quedan restringidos por la TEL adoptando los valores dados por la ecuación (3.33). Solamente el parámetro $y_{\mathcal{M}}^*$ se infiere libremente.

Modelo de sesgo de 4 parámetros

En relación al modelo de sesgo definido por la ecuación (3.13), cuyos parámetros son $\{\gamma, \delta, \tau, y_{\mathcal{M}}^*\}$, la aplicación de TEL de idéntica forma a como se ha expuesto previamente, conduce a la restricción de los parámetros $\{\gamma, \tau\}$, cuyas expresiones vienen definidas por las ecuaciones (3.34)

$$\begin{aligned} \gamma &= \mathbb{E}[\mathbf{E}_1] - \delta \mathbb{E}[\mathbf{Y}_{1,\mathcal{M}}] & \mathbf{E}_1 &= \{e_j\}, \mathbf{Y}_{1,\mathcal{M}} = \{y_{j,\mathcal{M}}\} \quad \forall j / y_{j,\mathcal{M}} \leq y_{\mathcal{M}}^* \\ \tau &= \frac{\left(\mathbb{E}[\mathbf{E}_2] - \mathbb{E}[\mathbf{E}_1] + \delta \left(\mathbb{E}[\mathbf{Y}_{1,\mathcal{M}}] - y_{\mathcal{M}}^*\right)\right)}{\mathbb{E}[\mathbf{Y}_{2,\mathcal{M}}] - y_{\mathcal{M}}^*} & \mathbf{E}_2 &= \{e_j\}, \mathbf{Y}_{2,\mathcal{M}} = \{y_{j,\mathcal{M}}\} \quad \forall j / y_{j,\mathcal{M}} > y_{\mathcal{M}}^* \end{aligned} \quad (3.34)$$

Por tanto, de los cuatro parámetros del modelo de sesgo definido por la ecuación (3.13), los parámetros $\{\gamma, \tau\}$ quedan restringidos por la TEL adoptando los valores dados por la ecuación (3.34) mientras que los parámetros $\{\delta, y_{\mathcal{M}}^*\}$ se infieren libremente.

3.4.2. Inclusión de las TLs al relajar la hipótesis de homocedasticidad. Modelo de varianza del error.

El modelo de varianza (realmente es de desviación estándar σ) propuesto con la ecuación (3.14), cuyos parámetros son $\{\alpha, \kappa\}$, permite modelar la dispersión de las diferentes distribuciones de densidad condicional del error, pertenecientes a la bivariada $P(\mathbf{E}, \mathbf{Y}_{\mathcal{M}})$, esto es, $\sigma = \left(\mathbb{V}[\mathbf{E}|\mathbf{Y}_{\mathcal{M}}]\right)^{\frac{1}{2}}$. Si las pdf condicionales $P(\mathbf{E}|\mathbf{Y}_{\mathcal{M}})$ han de formar parte de la bivariada se debe cumplir la TVL, la cual considerando el modelo de varianza asumido se puede reescribir como $\mathbb{V}[\mathbf{E}] = \mathbb{E}[\sigma^2]$. Siendo dos los parámetros y habiendo una sola ecuación, uno de los dos parámetros será independiente. Para llegar a la expresión que vincula ambos parámetros, se procede como sigue. Por la definición de varianza, se cumple la relación $\mathbb{V}[\sigma] = \mathbb{E}[\sigma^2] - \mathbb{E}^2[\sigma]$, de la cual se puede despejar el

término $\mathbb{E}[\boldsymbol{\sigma}]$ y aplicando la TVL dada por la ecuación (3.7), escribir la ecuación (3.35)

$$\mathbb{E}[\boldsymbol{\sigma}] = \left(\mathbb{E}[\boldsymbol{\sigma}^2] - \mathbb{V}[\boldsymbol{\sigma}] \right)^{\frac{1}{2}} \stackrel{\text{TVL}}{\rightarrow} = \left(\mathbb{V}[\mathbf{E}] - \mathbb{V}[\boldsymbol{\sigma}] \right)^{\frac{1}{2}} \quad (3.35)$$

Por otra parte, la esperanza de la función de varianza $\mathbb{E}[\boldsymbol{\sigma}]$, se puede calcular mediante la definición general de esperanza de una función $g(x)$, esto es,

$\mathbb{E}[g(x)] = \int_{-\infty}^{+\infty} g(x)P(x)dx$. Calculando así la esperanza de la función $\boldsymbol{\sigma}$, se obtiene la ecuación (3.36)

$$\mathbb{E}[\boldsymbol{\sigma}] = \alpha + \kappa \mathbb{E}[\mathbf{Y}_{\mathcal{M}}] \quad (3.36)$$

De manera similar, el término $\mathbb{V}[\boldsymbol{\sigma}]$ de la ecuación (3.35) se puede calcular mediante la definición general de varianza de una función $g(x)$, esto es,

$\mathbb{V}[g(x)] = \int_{-\infty}^{+\infty} (g(x) - \mathbb{E}[g(x)])^2 P(x)dx$. Calculando así la varianza de la función de desviación estándar $\boldsymbol{\sigma}$, se llega a la ecuación (3.37)

$$\mathbb{V}[\boldsymbol{\sigma}] = \kappa^2 \mathbb{V}[\mathbf{Y}_{\mathcal{M}}] \quad (3.37)$$

De las ecuaciones (3.35), (3.36) y (3.37) se obtiene la ecuación (3.38), que define el parámetro α en función del parámetro independiente κ .

$$\alpha = \left(\mathbb{V}[\mathbf{E}] - \kappa^2 \mathbb{V}[\mathbf{Y}_{\mathcal{M}}] \right)^{\frac{1}{2}} - \kappa \mathbb{E}[\mathbf{Y}_{\mathcal{M}}] \quad (3.38)$$

Resumiendo, de los dos parámetros del modelo de varianza, α está restringido por la TVL adoptando el valor dado por la ecuación (3.38). Solamente el parámetro κ se infiere libremente.

3.4.3. Inclusión de las TLs al relajar simultáneamente las hipótesis de media nula y homocedasticidad

Cuando el modelo de error contempla la relajación de ambas hipótesis de forma simultánea, se tiene según el modelo de sesgo elegido, un total de cinco $\{\gamma, \tau, y_{\mathcal{M}}^*, \alpha, \kappa\}$ o seis $\{\gamma, \delta, \tau, y_{\mathcal{M}}^*, \alpha, \kappa\}$ parámetros de error relacionados con estas relajaciones. Los resultados obtenidos anteriormente para los parámetros de los respectivos modelos de sesgo no cambian, es decir, $\{\gamma, \tau\}$ quedan condicionados por la TEL según las ecuaciones (3.33) o (3.34), mientras que los parámetros $\{\delta, y_{\mathcal{M}}^*\}$ se infieren libremente.

Sí que cambian, sin embargo, los cálculos para el modelo de varianza. Si se retrocede al apartado de inclusión de las TLs, cuando se relaja la hipótesis de homocedasticidad, hay un punto en el que se dice: por la definición de varianza, se cumple la relación $\mathbb{V}[\boldsymbol{\sigma}] = \mathbb{E}[\boldsymbol{\sigma}^2] - \mathbb{E}^2[\boldsymbol{\sigma}]$, de la cual se puede despejar el término $\mathbb{E}[\boldsymbol{\sigma}]$ y posteriormente aplicar la TVL. Así como previamente se aplicaba la TVL dada por la ecuación (3.7), ahora se debe aplicar su expresión completa dada por la ecuación (3.6). Se puede escribir así la ecuación (3.39)

$$\mathbb{E}[\boldsymbol{\sigma}] = \left(\mathbb{E}[\boldsymbol{\sigma}^2] - \mathbb{V}[\boldsymbol{\sigma}] \right)^{\frac{1}{2}} \stackrel{\text{TVL}}{\rightarrow} = \left(\mathbb{V}[\mathbf{E}] - \mathbb{V}[\boldsymbol{\mu}] - \mathbb{V}[\boldsymbol{\sigma}] \right)^{\frac{1}{2}} \quad (3.39)$$

Modelo de sesgo de 3 parámetros

En (3.39), el único término no calculado previamente es $\mathbb{V}[\boldsymbol{\mu}]$. Aplicando la definición general de varianza de una función, sobre las funciones del modelo de sesgo definidas en (3.12) se obtienen las ecuaciones de (3.40)

$$\begin{aligned} \mathbb{V}[\boldsymbol{\mu}_1] &= 0 & \boldsymbol{\mu}_1 &= \{\mu_j\} & \forall j / y_{j_{\mathcal{M}}} \leq y_{\mathcal{M}}^* \\ \mathbb{V}[\boldsymbol{\mu}_2] &= \tau^2 \mathbb{V}[\mathbf{Y}_{2_{\mathcal{M}}}] & \boldsymbol{\mu}_2 &= \{\mu_j\}, \mathbf{Y}_{2_{\mathcal{M}}} = \{y_{j_{\mathcal{M}}}\} & \forall j / y_{j_{\mathcal{M}}} > y_{\mathcal{M}}^* \end{aligned} \quad (3.40)$$

Aprovechando la circunstancia de que el modelo de varianza es idéntico para cada uno de los dos tramos del modelo de sesgo, el parámetro α se puede definir considerando de forma independiente cada uno de esos dos tramos.

Tomando en consideración la función de sesgo en su primer tramo, α se puede obtener a partir de las ecuaciones (3.36), (3.37), (3.39) y (3.40), según la ecuación (3.41)

$$\alpha = \left(\mathbb{V}[\mathbf{E}_1] - \kappa^2 \mathbb{V}[\mathbf{Y}_{1,M}] \right)^{\frac{1}{2}} - \kappa \mathbb{E}[\mathbf{Y}_{1,M}] \quad (3.41)$$

Si se considera la función de sesgo en su segundo tramo, α se obtiene también a partir de las ecuaciones (3.36), (3.37), (3.39) y (3.40), según la ecuación (3.42)

$$\alpha = \left(\mathbb{V}[\mathbf{E}_2] - (\tau^2 + \kappa^2) \mathbb{V}[\mathbf{Y}_{2,M}] \right)^{\frac{1}{2}} - \kappa \mathbb{E}[\mathbf{Y}_{2,M}] \quad (3.42)$$

Si se igualan ambas ecuaciones, se obtiene una expresión de la que es posible resolver en forma iterativa el parámetro κ . Para los casos en que $\alpha \approx 0$, una aproximación explícita del parámetro κ viene dada por la ecuación (3.43)

$$\kappa = \left(\frac{\mathbb{V}[\mathbf{E}_1] - \mathbb{V}[\mathbf{E}_2] + \tau^2 \mathbb{V}[\mathbf{Y}_{2,M}]}{\mathbb{E}[\mathbf{Y}_{1,M}^2] - \mathbb{E}[\mathbf{Y}_{2,M}^2]} \right)^{\frac{1}{2}} \quad (3.43)$$

Es decir, los dos parámetros del modelo de varianza $\{\alpha, \kappa\}$ quedan, en este caso, condicionados por las dos Leyes Totales, TEL y TVL. El parámetro α puede adoptar su valor, aplicando indistintamente las ecuaciones (3.41) o (3.42). El parámetro κ adopta el valor dado por la ecuación (3.43).

Modelo de sesgo de 4 parámetros

Si se parte de la ecuación (3.39), mediante idénticas consideraciones a las realizadas previamente, aplicadas ahora sobre el modelo de sesgo definido en (3.13), se llega a la restricción de los parámetros $\{\alpha, \kappa\}$ del modelo de varianza. Tomando en consideración la función de sesgo en su primer tramo, α se puede calcular según la ecuación (3.44)

$$\alpha = \left(\mathbb{V}[\mathbf{E}_1] - (\delta^2 + \kappa^2) \mathbb{V}[\mathbf{Y}_{1,M}] \right)^{\frac{1}{2}} - \kappa \mathbb{E}[\mathbf{Y}_{1,M}] \quad (3.44)$$

Si se considera la función de sesgo en su segundo tramo, el cálculo de α se realiza de nuevo mediante la ecuación (3.42).

Si se igualan ambas ecuaciones, se obtiene una expresión de la que es posible resolver en forma iterativa el parámetro κ . Para los casos en que $\alpha \approx 0$, una aproximación explícita del parámetro κ viene dada por la ecuación (3.45)

$$\kappa = \left(\frac{\mathbb{V}[\mathbf{E}_1] - \mathbb{V}[\mathbf{E}_2] - \delta^2 \mathbb{V}[\mathbf{Y}_{1,M}] + \tau^2 \mathbb{V}[\mathbf{Y}_{2,M}]}{\mathbb{E}[\mathbf{Y}_{1,M}^2] - \mathbb{E}[\mathbf{Y}_{2,M}^2]} \right)^{\frac{1}{2}} \quad (3.45)$$

Como resumen, en este caso de doble relajación de hipótesis, de los cinco (o seis) parámetros que incluyen los modelos de sesgo y varianza, solamente se infieren libremente los parámetros $\{\delta, y_M^*\}$, mientras que los parámetros $\{\gamma, \tau, \alpha, \kappa\}$ quedan condicionados por las dos Leyes Totales. Una importante disminución de grados de libertad en el problema de inferencia.

Para finalizar, es importante remarcar de nuevo la importancia de la aplicación de las Leyes Totales, la cual permite la coherencia estadística de los parámetros inferidos. Como se ha visto, también supone una reducción importante del número de parámetros libres a inferir. Para el caso en el que se aborde la inferencia conjunta del nuevo modelo de error que propone esta Tesis, en su versión completa, de entre todos los posibles parámetros que lo componen $\Omega = \{\gamma, \delta, \tau, y_M^*, \alpha, \kappa, \phi_1, \xi, \beta\}$, solamente se infieren libremente (mediante un algoritmo MCMC) los parámetros $\{\delta, y_M^*, \phi_1, \xi, \beta\}$.

3.4.4. Sobre el método WLS

De los tres casos que se acaban de tratar sobre la incorporación de las TLs al modelo de error, aquel en el que se relaja únicamente la hipótesis de varianza constante corresponde al llamado método WLS, el cual ya se trató en el capítulo 2. Recuérdese la problemática de dicho método, en el caso de no contar con la posibilidad de la repetición de mediciones con las que poder estimar el vector

$\sigma^2 = \left\{ \sigma_{e|y_{1,M}}^2, \dots, \sigma_{e|y_{n,M}}^2 \right\}$, que a su vez permite el cálculo de los pesos $w_j = \sigma_{e|y_{j,M}}^{-1}$

que ponderan a las observaciones. Este problema con el WLS en modelación hidrológica fue el origen del éxito de los métodos transformacionales que evitaban el problema. También fue la motivación que llevó a Sorooshian y Dracup (1980) a desarrollar su método híbrido para modelar la varianza de los errores, un método directo cuyos parámetros se infieren a partir de los datos transformados. Todo esto se explicó con detalle en la sección 2.3.3 del capítulo 2. En dicha sección se le pedía al lector que retuviese en memoria la ecuación (2.23). No obstante, se reproduce aquí, de nuevo, aquella breve ecuación $\kappa^2 = \sigma^{2*} = \mathbb{V}[\mathbf{E}^*]$. Lo especial de la ecuación, es la relación que se establece entre el parámetro κ del modelo de varianza y la varianza marginal de los errores transformados.

Préstese de nuevo atención a la ecuación (3.38) del modelo de error propuesto en esta Tesis. Si en dicha ecuación se asigna un valor $\alpha = 0$, como en el trabajo de Sorooshian y Dracup (1980) y se despeja κ^2 , se obtiene el resultado de la ecuación (3.46)

$$\kappa^2 = \frac{1}{\mathbb{E}[\mathbf{Y}_{\mathcal{M}}^2]} \mathbb{V}[\mathbf{E}] \propto \mathbb{V}[\mathbf{E}] \quad (3.46)$$

No parece casual, que en dos métodos que siguen procedimientos tan diferentes para la modelación directa de la varianza condicional de los errores, se llegue a resultados en la estimación de κ , que guardan tan estrecha relación. Por un lado, el método híbrido de 1980, el cual infiere el parámetro κ a partir de datos previamente transformados para estabilizar la varianza de los errores. Su estimación de κ se relaciona (coincide de hecho) con la varianza marginal de los errores transformados. Por otro lado, el método directo puro propuesto en esta Tesis, que infiere el parámetro κ a partir de los datos originales sin transformar, pero que en la inferencia toma en consideración la ley de la varianza total (TVL). Su estimación de κ también queda relacionada con la varianza marginal de los errores, como se observa en la ecuación (3.46).

La conclusión es que, en ambos procedimientos, siendo bien diferentes, la modelación de las varianzas condicionales de los errores requiere que el parámetro κ tenga en cuenta, de alguna manera, la distribución marginal de los

mismos. Ninguno de los trabajos existentes en la literatura, en los que se aplica un método directo de modelación de la varianza condicional de los errores, después de 1980 hasta el presente, ha tenido en cuenta estas consideraciones.

Especialmente representativo es el caso del trabajo de Evin et al. (2014), en el cual se emplea el método WLS en uno de sus casos de estudio. El trabajo se limita a modelar la varianza condicional de los errores mediante la ecuación (3.14), sin ninguna otra consideración. Es posible, que de la inferencia así realizada, se obtengan unos pesos $w_j = \sigma_{e|y_j, \mathcal{M}}^{-1}$ con los que ponderar las observaciones. Sin embargo, se puede poner en duda que los parámetros $\{\alpha, \kappa\}$ así inferidos, mantengan la coherencia estadística de la que se ha estado hablando a lo largo de este capítulo 3. Porque, de no ser así, ¿Qué razón de ser tienen el método híbrido de Sorooshian y Dracup (1980), o el método propuesto en esta Tesis? En los experimentos diseñados en el capítulo 4 cuyos resultados se muestran en los capítulos 5 y 6, se volverá a tratar el modelo de error WLS, demostrando que si su aplicación tiene por objeto la determinación estadísticamente coherente de los parámetros $\{\alpha, \kappa\}$, necesariamente ha de considerar las TLs.

3.5. Estimación de la incertidumbre predictiva a partir de los modelos hidrológico y de error conjuntamente inferidos

Una vez terminada la inferencia conjunta de los modelos, hidrológico y de error, es posible evaluar la incertidumbre de la predicción del modelo hidrológico. El método de obtención de la distribución predictiva (PD) se fundamenta en la definición teórica expuesta en la sección 2.4.1. En la práctica, el procedimiento de cálculo es el que se explica a continuación.

Primero, se define la llamada **ecuación de muestreo predictivo**, la cual permite generar un conjunto aleatorio de múltiples simulaciones del modelo hidrológico (llamado *ensemble* en la literatura). Este conjunto de simulaciones posibles está considerando tanto la incertidumbre almacenada en los parámetros (de error e hidrológicos) como la incertidumbre total que afecta a cada simulación del modelo hidrológico y que queda representada por el modelo de error calibrado conjuntamente. La ecuación de muestreo predictivo se define según la ecuación (3.47)

$$\hat{\mathbf{Y}} = \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \boldsymbol{\Omega}) + \mathbf{E}(\boldsymbol{\theta}, \boldsymbol{\Omega}) = \mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}, \boldsymbol{\Omega}) + \boldsymbol{\mu} + \boldsymbol{\sigma} \sigma_z \phi_p^{-1}(B) \mathbf{a} \quad (3.47)$$

Primero, a partir de las distribuciones posteriores inferidas, se muestrea \mathbb{S} conjuntos de parámetros hidrológicos y de error $\{\boldsymbol{\theta}_i, \boldsymbol{\Omega}_i\} \forall i \in \{1, \dots, \mathbb{S}\}$. Seguidamente, se realizan simulaciones del modelo hidrológico con los parámetros $\boldsymbol{\theta}_1$, obteniéndose así una simulación de la variable de interés $\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}_1, \boldsymbol{\Omega})$. A continuación, con el conjunto de parámetros de error $\boldsymbol{\Omega}_1$ y una muestra aleatoria de la SEP $\mathbf{a} = \{a_j\} \stackrel{\mathcal{D}}{\sim} P_{\text{SEP}}(0, 1, \xi_1, \beta_1) \forall j \in \{1, \dots, n\}$, se entra en la ecuación (3.47) para obtener una estimación de la variable de interés $\hat{\mathbf{Y}}(\boldsymbol{\theta}_1, \boldsymbol{\Omega}_1) = \{\hat{y}_1, \dots, \hat{y}_n\}$. Con el fin de considerar, con suficiente resolución, la incertidumbre total de $\mathbf{Y}_{\mathcal{M}}(\boldsymbol{\theta}_1, \boldsymbol{\Omega})$, la cual está “almacenada” en el modelo de error, se repite el muestreo aleatorio de la SEP obteniendo un total de \mathbb{K} muestras $\{\mathbf{a}_1, \dots, \mathbf{a}_{\mathbb{K}}\}$. Con las muestras SEP, entrando \mathbb{K} veces en la ecuación (3.47) se obtienen otras tantas estimaciones de la variable de interés $\{\hat{\mathbf{Y}}_1(\boldsymbol{\theta}_1, \boldsymbol{\Omega}_1), \dots, \hat{\mathbf{Y}}_{\mathbb{K}}(\boldsymbol{\theta}_1, \boldsymbol{\Omega}_1)\}$. Este *ensemble* de \mathbb{K} estimaciones, es una muestra de la distribución predictiva $P(\mathbf{Y} | \mathbf{Y}_{\mathcal{M}}, \boldsymbol{\theta}_1, \boldsymbol{\Omega}_1)$. Para eliminar la condicionalidad de la predicción a los parámetros, hay que repetir todo el proceso con los parámetros muestreados $\{\boldsymbol{\theta}_i, \boldsymbol{\Omega}_i\} \forall i \in \{2, \dots, \mathbb{S}\}$. Al final del procedimiento, se obtiene el *ensemble* $\{\hat{\mathbf{Y}}\}_{\mathbb{S} \times \mathbb{K}}$, el cual es una muestra de la distribución predictiva, incondicional a los parámetros $P(\mathbf{Y} | \mathbf{Y}_{\mathcal{M}})$. A partir de esta muestra se pueden realizar las inferencias necesarias, siendo lo más habitual obtener la esperanza de la predicción y la banda de incertidumbre acotada entre dos cuantiles determinados (e.g. la banda del 95%, acotada por las probabilidades de no excedencia de 2.5% y 97.5%). Esta Tesis adopta como estimador de la esperanza de la predicción $\mathbf{Y}^* = \{y_1^*, \dots, y_n^*\}$, a la media del *ensemble* $\{\hat{\mathbf{Y}}\}_{\mathbb{S} \times \mathbb{K}}$, calculada según la ecuación (3.48)

$$\mathbb{E}[\mathbf{Y} | \mathbf{Y}_{\mathcal{M}}, \tilde{\mathbf{Y}}, \mathbf{x}_0, \tilde{\mathbf{B}}, \mathcal{M}] \approx \mathbf{Y}^* = \frac{1}{\mathbb{S}\mathbb{K}} \sum_{i=1}^{\mathbb{S}} \sum_{j=1}^{\mathbb{K}} \{\hat{\mathbf{Y}}\}_{ij} \quad (3.48)$$

Nótese, que a pesar de las hipótesis hechas al principio, para facilitar el manejo y desarrollo del modelo de error, los resultados obtenidos y su incertidumbre van a estar siempre condicionados a los datos $\{\tilde{\mathbf{Y}}, \mathbf{x}_0, \tilde{\mathbf{B}}\}$, al modelo \mathcal{M} y a las herramientas (e.g. algoritmos computacionales Bayesianos) con las que se ha realizado la inferencia. Todo influye en la capacidad de predecir del modelo, así como en su incertidumbre.

CAPÍTULO 4. DESCRIPCIÓN DE EXPERIMENTOS

Philosophers have said that if the same circumstances don't always produce the same results, predictions are impossible and science will collapse [...] that's the way it is: Nature permits us to calculate only probabilities. Yet science has not collapsed

-Richard Feynman (1918-1988)-

4.1. Motivación

Esta Tesis propone que la inferencia de los parámetros de un modelo hidrológico sea una inferencia conjunta con los parámetros del nuevo modelo de error propuesto en el capítulo 3. El modelo de error propuesto es más apropiado que los modelos de error clásicos, a los errores que se derivan del proceso de inferencia de modelos hidrológicos. Así mismo, recomienda el empleo de técnicas de inferencia Bayesiana las cuales, entre otras bondades, permiten tratar de forma nativa la incertidumbre de los parámetros.

El nuevo modelo de error se basa en el modelo de error propuesto por Schoups y Vrugt (2010), al que se le ha añadido la necesaria modificación que proponen Evin et al. (2013), siendo la primera vez que se define la función de verosimilitud resultado de la fusión de ambos trabajos (sección 3.3.2). Por tanto, respecto al clásico método del SLS, el modelo de error propuesto está relajando todas las hipótesis, esto es, media nula, varianza constante, independencia de los errores y Normalidad de los mismos.

Todo lo mencionado supone una novedad, aunque relativa. Schoups y Vrugt (2010) ya relajaban con su GL, todas las hipótesis mencionadas. Evin et al. (2013), si bien no relajan las hipótesis de Normalidad y media nula, mejoran el método de Schoups y Vrugt (2010) evitando (al menos en sus casos de estudio) los problemas espurios con la inflación de las bandas de incertidumbre. Pero lo cierto es que, en el comienzo de esta Tesis, se comprobó que el método Schoups-Evin seguía sin funcionar, como se explica en la sección 3.1.1 en la que se describen los dos primeros experimentos de esta Tesis. El estudio y la intuición condujeron a formular la hipótesis central de la Tesis. Esta hipótesis, que está expresada para la modelación de las varianzas condicionales, es igualmente aplicable a la modelación de las esperanzas condicionales de los errores.

Un modelo de varianza condicional de los errores que siga el método directo, es decir, sin previa transformación de los datos, debe tomar en consideración la relación que existe entre las distribuciones condicionales y la correspondiente distribución marginal de los errores. La variación de ambas distribuciones, durante el proceso de inferencia conjunta, debe ocurrir de forma acoplada, manteniendo así la coherencia estadística.

La hipótesis requiere la misma asunción que se hace en los métodos Bayesianos directos para el cálculo de la incertidumbre predictiva (e.g. Todini, 2013), esto es, la existencia de una distribución conjunta entre las simulaciones de la variable de interés $\mathbf{Y}_{\mathcal{M}}$ y la propia variable de interés, representada muestralmente por sus observaciones $\tilde{\mathbf{Y}}$. En este caso, asumiendo que los errores son aditivos, la distribución conjunta $P(\mathbf{Y}, \mathbf{Y}_{\mathcal{M}})$ es equivalente a la distribución $P(\mathbf{E}, \mathbf{Y}_{\mathcal{M}})$.

Como se explicó en el capítulo 2, la mayor parte (si no todos) los trabajos hallados en la literatura tratan el tema de la heterocedasticidad mediante los métodos transformacionales. Sin embargo, a partir de 2010 y siguiendo el trabajo de Schoups y Vrugt (2010) son varios los trabajos que han optado por aplicar el método directo (o explícito, según algún autor) de modelar la varianza.

En cierta medida, resulta sorprendente que en esos trabajos, excepto en cuatro (Evin et al., 2013, 2014; Scharnagl et al., 2015; Schoups y Vrugt, 2010), se decide no relajar la hipótesis de independencia de los errores, es decir, no inferir el modelo autorregresivo conjuntamente con el de varianza (e.g. (Westra et al., 2014)). Otra forma de evitar la inferencia conjunta del modelo de varianza con los otros componentes del modelo de error, es la que emplean los autores Han y Zheng (2016) y Zheng y Han (2016). Estos autores aplican un método de inferencia iterativo (ver los métodos de asignación de valores a los parámetros en la sección 2.1.1), en el cual alternan la inferencia mediante MCMC de los parámetros hidrológicos, con la inferencia mediante estimación ML de los parámetros de error, evitando así la interacción entre ambos grupos de parámetros. Dada la alta dimensión y no linealidad de los problemas de inferencia tratados generalmente en modelación hidrológica, cabe la duda razonable de si el resultado de la estimación de parámetros mediante una solución condicional iterativa en la que se evita su interacción coincidirá, en

cualquier caso, con el resultado de la estimación mediante una inferencia conjunta.

Sin embargo, no resulta tan sorprendente que la mayor parte de los autores rehuya el empleo de la inferencia conjunta del modelo de varianza, si se presta atención al resultado que obtuvieron los cuatro trabajos que sí la acometen.

Schoups y Vrugt (2010) solo consiguen inferir conjuntamente y sin aparentes problemas, 4 parámetros de error (de un total de 5) en cada caso de estudio, esto es, $\Omega_{\text{caso-1}} = \{\alpha, \kappa, \phi_1, \beta\}$ y $\Omega_{\text{caso-2}} = \{\alpha, \kappa, \xi, \beta\}$.

Evin et al. (2013), con su cambio metodológico (ver sección 2.3.6) consiguen inferir conjuntamente y sin problemas tres parámetros de error $\Omega = \{\alpha, \kappa, \phi_1\}$, los dos del modelo de varianza y el del modelo autorregresivo AR(1). Comprueban que, sin aplicar su mejora, se obtienen bandas de incertidumbre con inflación. Además, observan muy alta correlación entre los parámetros $\{\kappa, \phi_1\}$ e igualmente entre $\{\theta_2, \phi_1\}$ donde θ_2 es un parámetro del modelo hidrológico que controla fuertemente el balance de agua en la cuenca.

Evin et al. (2014), en un exhaustivo trabajo de investigación sobre 12 cuencas del experimento MOPEX (Duan et al., 2006), llegan a la siguiente conclusión: *“Despite the theoretical attraction of the joint inference approach, the empirical results suggest that joint inference of hydrological parameters and the residual error autocorrelation parameter can be non-robust. In particular, the joint approach can produce unrealistically wide predictive limits”*. Su modelo de error completo contempla los mismos tres parámetros $\Omega = \{\alpha, \kappa, \phi_1\}$ que en el trabajo de Evin et al. (2013) y también aplican la corrección de Evin (sección 2.3.6). En este trabajo se encuentran con diferentes cuencas en las que les vuelve a aparecer el problema de la inflación de las bandas, y de nuevo, altas correlaciones entre $\{\kappa, \phi_1\}$ y $\{\theta_2, \phi_1\}$. De las 12 cuencas analizadas, solamente queda libre de aparentes problemas, la cuenca French Broad (FB), la misma del primer caso de estudio de Schoups y Vrugt (2010). Entre las recomendaciones finales de estos autores está el uso de la técnica de inferencia en posproceso (ver los métodos de asignación de valores a los parámetros en la sección 2.1.1),

a pesar de las desventajas o inconvenientes que supone y que ellos mismos reconocen.

El cuarto de los trabajos que aborda el método directo de modelar la varianza es el de Scharnagl et al. (2015). Este trabajo es el más parecido al que se va a abordar en esta Tesis, por el número de parámetros y similitud de su modelo de error. En primer lugar, su modelo de varianza, formado por un polinomio cúbico interpolador de Hermite, presenta 4 parámetros. Su modelo de dependencia es el común AR(1), el cual se aplica sobre los errores con la varianza ya estabilizada (según la corrección de Evin et al.). No incluyen un modelo de sesgo. Su hipótesis de distribución de las innovaciones asume una distribución Student asimétrica con dos parámetros, que como en la SEP, representan la asimetría y el exceso/defecto de curtosis. En total 7 parámetros de error libres $\Omega = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \phi_1, \beta, \xi\}$, que infieren conjuntamente con los parámetros hidrológicos. En su trabajo reportan obtener resultados sin sentido, entre los que llama la atención la aparición de la inflación en las bandas de incertidumbre y un extremadamente alto valor MAP (0.99) del parámetro de autocorrelación ϕ_1 , como en el segundo caso de estudio de Schoups y Vrugt (2010). Valores tan extremos de este parámetro son sinónimo de que la posterior del parámetro no es identificable, puesto que con esos valores tan próximos a la unidad, el proceso estocástico AR(1) no es estacionario (Box et al., 1994), por lo que no representa a un *ruido blanco*.

4.2. Modelos de error

En esta sección se van a describir las pruebas o experimentos que permiten comprobar si los problemas aparecidos en la literatura, guardan alguna relación con la forma (incorrecta) en la que se está tratando de modelar la varianza de los errores por el método directo. Dichos experimentos serán ejecutados tanto con modelos de error clásicos, que sirven de referencia, como con el nuevo modelo de error que se postula como una posible solución para modelar correctamente y de forma directa la varianza de los errores.

Los experimentos que se va a realizar consideran los siguientes modelos de error. En el capítulo 5 se realizará la inferencia conjunta con los modelos de error SLS, WLS y GL++, aplicados junto con dos modelos hidrológicos agregados.

Además, con los modelos WLS y GL++ que incluyen la modelación directa de la varianza de los errores, con la consideración de las TLs, se comprobará la hipótesis principal de esta Tesis mediante la comparación con sus respectivas versiones WLS-NTL y GL++NTL, que no consideran las TLs. En la primera parte del capítulo 6, se realizará la inferencia conjunta del modelo de error completo GL++Bias definido en el capítulo 3, en su variante de 8 parámetros (modelo de sesgo de 3 parámetros), junto con los mismos dos modelos agregados del capítulo 5. En la segunda parte del capítulo 6, se aplicarán con un modelo hidrológico distribuido el modelo de error GL++ y el modelo GL++Bias en su variante de 9 parámetros (modelo de sesgo de 4 parámetros). Seguidamente se describirán las características de todos los modelos de error.

El modelo de error **SLS**, que ya se definió como el adecuado exclusivamente en los casos de inferencia ideales (sección 2.3.1), es el modelo de error más empleado en los análisis de regresión, y aunque cada vez menos, en la calibración de los modelos hidrológicos. Su consideración como experimento, es la de tener una referencia con la que comparar el resto de modelos de error.

El modelo de error **WLS** se definió en la sección 2.3.3 del capítulo 2, y es el resultado del relajamiento de la hipótesis de varianza constante del modelo SLS. En este modelo de error podría estar realmente la clave de toda la Tesis, a pesar de su sencillez, o gracias a su simplicidad. Recuérdese, que hasta que no se publicó el trabajo de Sorooshian y Dracup de 1980, e incluso después de su publicación, la forma de tratar la heterocedasticidad de los errores era evitarla mediante la transformación de las variables. En el citado trabajo, se aborda por primera vez el método directo de modelar la varianza de los errores, esto es, modelarla como función de las variables originales sin transformar (e.g. el caudal). La estimación de los parámetros de ese modelo directo de la varianza, se realizó a partir de una transformación de las variables. 30 años después en 2010, Schoups y Vrugt retoman el método directo de modelar la varianza. Sin embargo, se olvidan de que el método directo tal cual se concibió en 1980, no está completo sin estimar los parámetros a partir de las variables transformadas. Todos los trabajos posteriores a 2010, que tratan de modelar de forma directa la varianza de los errores, caen en el mismo olvido.

Para comprobar la hipótesis de esta Tesis, se van realizar dos experimentos relacionados con este modelo de error. El primero, llamado **WLS-NTL**, solo

aplicará el modelo de varianza con dos parámetros libres $\Omega = \{\alpha, \kappa\}$ según la ecuación (3.14), sin más consideraciones. El segundo experimento, llamado **WLS**, aplicará el mismo modelo de varianza, pero esta vez considerando la aplicación de la Ley de la Varianza Total (TVL), que implica la restricción de uno de los dos parámetros del modelo de varianza, según la ecuación (3.38). Dicha restricción permite mantener la coherencia estadística de la que se habló en el capítulo 3. El modelo WLS-NTL tiene su equivalente en la literatura en el trabajo de Evin et al. (2014), donde se le llama WLS.

El tercero de los modelos de error que se va aplicar es el nuevo modelo de error **GL++**. En este modelo de error se han relajado todas las hipótesis del modelo SLS, con excepción de la hipótesis de errores con media nula. Como consecuencia, el modelo contempla el tratamiento de la varianza no constante (ecuación (3.14)), la dependencia serial de los errores (ecuación (3.18)) y la distribución no Gaussiana de los residuos (o innovaciones) mediante la pdf SEP descrita por la ecuación (2.28). Así mismo, considera la corrección de Evin et al. (sección 2.3.6) para la estandarización de los errores con el modelo de varianza, previa a la aplicación del filtro AR(1). Hasta aquí, se tiene al que previamente se llamó modelo **GL++NTL** (GL más la corrección de Evin), que define la función de verosimilitud (3.31) con 5 parámetros libres en el modelo de error $\Omega = \{\alpha, \kappa, \phi_1, \xi, \beta\}$. Sin embargo, para comprobar la hipótesis planteada, es adicionalmente necesaria la aplicación de la TVL, restringiendo el valor de uno de los parámetros del modelo de varianza mediante la ecuación (3.38). Por tanto, el nuevo modelo de error GL++ queda con 4 parámetros libres $\Omega = \{\kappa, \phi_1, \xi, \beta\}$ y uno restringido por la TVL.

Cabe la aclaración, de que el modelo GL++NTL es el mismo experimento número 2, de los dos experimentos iniciales de la Tesis (sección 3.1.1). Por tanto, se va a demostrar que GL++ soluciona el problema que aparecía en dicho experimento. Así mismo, GL++NTL es también el experimento equivalente al WLS-AR1 del trabajo de Evin et al. (2014), si bien ellos no relajan la hipótesis Gaussiana, por lo que su modelo de error tiene dos parámetros menos.

Finalmente, el cuarto modelo de error que se aplicará, **GL++Bias**, resulta de añadir al modelo GL++ la relajación de la hipótesis de errores con media nula del método SLS. Este modelo representa al modelo de error completo definido en la

sección 3.3, con 8 parámetros libres $\Omega = \{\gamma, \tau, y_M^*, \alpha, \kappa, \phi_1, \xi, \beta\}$ cuando, de forma errónea, no se consideran las TLs (TVL y TEL). Al considerarlas, los parámetros libres se reducen a 4. Así, los dos parámetros del modelo de varianza definido por la ecuación (3.14) dejan de ser libres, y se infieren mediante la ecuación (3.41) (o indistintamente la (3.42)) y la ecuación (3.43). De los tres parámetros del modelo de sesgo definido por la ecuación (3.12), se infiere libremente el parámetro umbral de caudal y_M^* , mientras que los otros dos parámetros se infieren mediante las ecuaciones de (3.33). En resumen, el modelo de error GL++Bias mantiene 4 parámetros libres $\Omega = \{y_M^*, \phi_1, \xi, \beta\}$ y otros 4 restringidos por las TLs.

En el caso de la variante de 9 parámetros $\Omega = \{\gamma, \delta, \tau, y_M^*, \alpha, \kappa, \phi_1, \xi, \beta\}$ del modelo **GL++Bias**, aplicada exclusivamente junto con el modelo hidrológico distribuido y cuyos resultados se muestran en el capítulo 6, se inferirán libremente 5 parámetros $\Omega = \{\delta, y_M^*, \phi_1, \xi, \beta\}$, quedando restringidos por aplicación de las TLs los parámetros $\{\gamma, \tau, \alpha, \kappa\}$.

4.3. Modelación hidrológica agregada

4.3.1. Descripción de los modelos considerados

Se han escogido dos modelos lluvia-escorrentía agregados, para ilustrar los aspectos teóricos que se quiere comprobar o demostrar. Dichos modelos han sido empleados en muchas de las publicaciones consultadas en la literatura. Los modelos son el **CRR**, empleado en la publicación de Schoups y Vrugt (2010), y el **GR4J** (Perrin et al., 2003) utilizado profusamente en la literatura y también en los trabajos de Evin et al. (2013, 2014).

El modelo CRR es un modelo conceptual espacialmente agregado, para la simulación diaria de los procesos de generación de la escorrentía a partir de los inputs de lluvia. El modelo tiene 7 parámetros a ser calibrados: la intercepción máxima (I_{max} , mm), la capacidad de almacenamiento del suelo (S_{max} , mm), la máxima tasa de percolación (Q_{smax} , mm/d), un parámetro adimensional del

submodelo de evaporación (α_e), un parámetro adimensional del submodelo de generación de escorrentía (α_f), un parámetro del tanque de propagación rápida (K_f , días) y un parámetro del tanque de propagación lenta (K_s , días). Por tanto, el modelo considera los procesos de interceptación, transcolación, evapotranspiración, generación de escorrentía directa, percolación y propagación del flujo superficial y subterráneo, hasta el punto de desagüe de la cuenca.

El modelo GR4J es también un modelo espacialmente agregado que fue desarrollado con el objetivo de proporcionar el mejor comportamiento medio sobre un amplio rango de cuencas con diferentes condiciones hidrológicas (Perrin et al., 2003). El GR4J solo tiene 4 parámetros a ser calibrados: la capacidad de almacenamiento del suelo (θ_1 , mm), el coeficiente de intercambio de flujo subterráneo (θ_2 , mm), la máxima capacidad de almacenamiento diario del tanque de propagación (θ_3 , mm) y el tiempo base del hidrograma unitario que emplea el modelo (θ_4 , días).

La inferencia de los parámetros hidrológicos va a ser una inferencia Bayesiana (ver sección 2.2) conjunta con los parámetros del modelo de error que se asuma en cada experimento. En todas las inferencias se asumirá unas distribuciones prior de los parámetros planas (o uniformes), puesto que el número de datos diarios empleados $n=1825$, correspondiente a 5 años, es suficiente para soportar la inferencia de los parámetros.

Respecto a la suficiencia del número de datos requerido para la calibración, se ha elegido un periodo de 5 años por ser exactamente el mismo periodo (en longitud y rango de fechas) empleado en el trabajo de Schoups y Vrugt (2010). Estos autores sugieren que 5 años de datos diarios, contienen la información necesaria para determinar los parámetros de un modelo hidrológico conceptual y agregado. Por otra parte, Brigode et al. (2013) afirman que no hay un claro consenso en la determinación de la longitud mínima del periodo de calibración, la cual deberá ser establecida en función de la especificidad del modelo y la cuenca analizados. Por concretar, Ancill et al. (2004) obtuvieron buenos resultados con el modelo GR4J con periodos de calibración de entre 3 y 5 años de datos diarios. Perrin et al. (2007) muestran que la calibración del GR4J proporciona resultados aceptables con tan solo un año de datos diarios. Finalmente, la experiencia del grupo de investigación del que escribe, muestra

que incluso con modelos distribuidos más complejos, y con más parámetros implicados, es posible la estimación de parámetros con un año de datos diarios (Frances et al., 2007).

Una cuestión importante a ser tomada en consideración. Ninguno de los dos modelos empleados modelan los procesos de precipitación en forma de nieve, y lógicamente, tampoco el proceso de generación de escorrentía por su fusión. Sin embargo, dichos procesos ocurren en la parte alta de una de las cuencas (FB), aunque su influencia es poco significativa en el balance hidrológico global. En cualquier caso, no deja de ser un error de especificación de ambos modelos, que contribuirá junto con el resto de fuentes de incertidumbre, a la incertidumbre de los parámetros y de la predicción.

4.3.2. Cuenca de estudio

Para la implementación de los experimentos con los modelos agregados se ha escogido la cuenca **French Broad River (FB)**, localizada en Carolina del Norte (USA), con un área de 2448 km². La cuenca FB es una cuenca cuyos datos se pueden obtener libremente del experimento MOPEX (Duan et al., 2006). Es una cuenca húmeda representativa, en la que la evapotranspiración real es limitada, siendo la escorrentía un proceso predominante. La cantidad de precipitación anual media es de 1495 mm y la de evapotranspiración potencial media es 820 mm.

Como se ha comentado previamente, para la estimación de los parámetros de los modelos se han empleado 5 años (desde 01/01/1962 a 31/12/1966) de datos diarios observados de precipitación, evapotranspiración potencial, y caudal en el punto de control de la cuenca. Adicionalmente a los 5 años de calibración, se han empleado dos años previos como periodo de calentamiento del modelo, con el propósito de que la incertidumbre en las condiciones iniciales afecte lo menos posible a los resultados.

Como fue comentado en la sección 3.1.1, la cuenca FB ya fue utilizada en el primer caso de estudio de Schoups y Vrugt (2010), así como en uno de los casos de estudio de los trabajos de Evin et al. (2013, 2014). En el trabajo de Schoups y Vrugt la cuenca no mostró problemas en su inferencia. En el trabajo de Evin et al. (2013), sí que dio problemas sin considerar la corrección que proponen. En el trabajo de Evin et al. (2014), aplicando la corrección de Evin, la cuenca tampoco

generó problemas. En los experimentos número 1 y número 2 de esta Tesis, si que dio problemas como se comenta a continuación.

En los dos primeros experimentos de esta Tesis (sección 3.1.1) no fue posible la correcta inferencia de la cuenca FB. El primer experimento finalizó con la no convergencia del problema, con tan solo añadir el parámetro de asimetría de la SEP al caso de estudio 1 de Schoups y Vrugt (2010). Por lo demás, era el mismo caso de estudio, es decir, misma cuenca, mismos datos, mismo modelo hidrológico (CRR) y mismo modelo de error (excepto el parámetro de asimetría). Un solo parámetro de error marcaba la diferencia entre el éxito aparente y el fracaso evidente de la inferencia.

Evin et al. (2013) emplearon el modelo hidrológico GR4J (3 parámetros menos que el CRR de Schoups y Vrugt) y un modelo de error con 3 parámetros (dos parámetros menos que Schoups y Vrugt, debido a la no relajación de la hipótesis Gaussiana de los residuos). La inferencia les dio problemas de inflación en las bandas de incertidumbre, sin aplicar la corrección de Evin. Con la citada corrección desapareció la inflación, aunque detectaron una fuerte interacción entre los parámetros hidrológicos y de error.

En el segundo experimento de esta Tesis, la corrección de Evin et al. permitió alcanzar la convergencia no conseguida en el experimento 1, pero los resultados mostraban el problema de inflación de la banda de incertidumbre. Este experimento con el modelo hidrológico CRR, se va a reproducir con el nombre GL++NTL.

En resumen, el fracaso de los dos primeros experimentos de la Tesis, motiva a seguir con la cuenca FB, para intentar dilucidar el origen de los problemas que aparecen y desaparecen sobre una misma cuenca, aparentemente en función de solo 2 factores: el modelo hidrológico, y el número de parámetros de error o, equivalentemente, el nivel de relajación de hipótesis del SLS que contempla el modelo de error.

4.4. Modelación hidrológica distribuida

4.4.1. Descripción del modelo considerado

La aplicación de los nuevos modelos de error **GL++** y **GL++Bias** (versión de 9 parámetros) a la modelación hidrológica distribuida, se realiza con el modelo hidrológico distribuido conceptual **TETIS** (Bussi et al., 2013; Frances et al., 2007; Salazar et al., 2012; Vélez et al., 2007), el cual tiene 8 parámetros con sentido físico (que no físicamente basados) para la producción de escorrentía. Dichos parámetros han de ser estimados para cada celda del modelo de la cuenca. Para la propagación en cauce, el modelo utiliza la aproximación de Onda Cinemática Geomorfológica (Frances et al., 2007) basada en la clásica simplificación de onda cinemática para el transporte de flujo, con la incorporación de relaciones geomorfológicas para la estimación de las características hidráulicas, en cualquier celda del modelo. Los inputs del modelo incluyen series temporales observadas de caudal, precipitación y evapotranspiración potencial, las cuales se definen en determinados puntos del modelo (puntos de control y estaciones de medida). La precipitación es interpolada espacialmente, considerando en el cálculo de cada punto (celda), un número limitado de estaciones influyentes, así como una corrección de la precipitación interpolada, basada en la altitud (parámetro FP1). La información espacial se introduce al modelo mediante mapas raster.

La principal característica del modelo TETIS es su estructura separada de los parámetros (Frances et al., 2007; Francés y Casado, 1995). Esta característica facilita considerablemente la inferencia de los parámetros de un modelo distribuido, convirtiendo un problema que en general, presenta espacios paramétricos de muy alta dimensión, en un problema con un número razonable de parámetros, que permite su determinación condicionada a las observaciones. La estructura separada considera la estimación de los parámetros efectivos en dos fases. En una primera fase se estima el patrón espacial (los mapas) de los parámetros de interés, empleando para ello toda la información ambiental disponible. Una vez definida la variabilidad espacial de la información, se estiman mediante inferencia los llamados factores correctores o factores de regularización (Pokhrel y Gupta, 2010), los cuales modifican globalmente los mapas estimados previamente, manteniendo los patrones de variabilidad espacial. En el caso de aplicación de esta Tesis, se infieren un total de 9

parámetros (factores correctores): 7 parámetros corresponden a los mecanismos de producción de escorrentía, 1 parámetro corresponde al mecanismo de propagación y 1 último parámetro corresponde al modelo de modificación de la precipitación con la altitud. La descripción de los parámetros con sus rangos se muestra en la Tabla 4.1.

4.4.2. Descripción de la cuenca

La cuenca seleccionada para la aplicación distribuida es una subcuenca de la parte alta del Río Oria, en el País Vasco (España). La estación de medición de caudal de la subcuenca modelada se llama “Alegia” (337 km²). La precipitación anual media es 1534 mm, por lo que se considera una cuenca húmeda. La temperatura media es de 13° C y la evapotranspiración potencial anual es 837 mm. La resolución espacial del modelo (tamaño de celda) es de 500 m x 500 m. En relación a la geología de la zona, predomina la alternancia de estratos de origen detrítico. La vegetación predominante está formada en un 60% por pastos y bosques.

Para la estimación de los 9 parámetros del modelo, se utilizan 2 años de datos históricos diarios (del 1 octubre de 1998 al 30 de septiembre de 2000). Adicionalmente se modela un año previo para el calentamiento del modelo. Así mismo, para la validación temporal del modelo se utiliza un periodo de 3 años de datos históricos diarios (del 1 de octubre de 1995 al 30 de septiembre de 1998), con la adición de un año previo para el calentamiento del modelo.

Tabla 4.1. Parámetros del modelo TETIS y del modelo de error. Descripción y rangos

Hydro. Model	CORRECTION FACTORS DESCRIPTION	MIN	MAX
F1	Maximum Static storage	0.8	1.3
F2	Evapotranspiration	0.8	1.3
F3	Infiltration capacity	0	0.5
F4	Slope velocity	0	0.5
F5	Percolation capacity	0	0.5
F6	Interflow	800	10000
F8	Connected aquifer flow	50	2000
F9	Flow velocity	0	0.5
FP1	Rainfall correction by altitude	0	0.1
Error. Model	DESCRIPTION	MIN	MAX
α	Intercept of variance model		
κ	Slope Coeficient of variance model	0	1
β	SEP Distribution Kurtosis Parameter	-1	1
ξ	SEP Distribution Skewness Parameter	0.1	10
ϕ_1	Autorregressive Model Parameter	0	0.99
γ^*	Flow Threshold for the Bias Model	1	80
γ	Intercept of Bias Linear Model for Low Flows		
δ	Slope of Bias Linear Model for Low Flows	-1	1
τ	Slope of Bias Linear Model for High Flows		

CAPÍTULO 5. CONSECUENCIAS DE LA (NO) APLICACIÓN DE LAS TLs EN LA INFERENCIA CONJUNTA

You can recognize truth by its beauty and simplicity [...]

the truth always turns out to be simpler than you thought

-Richard Feynman (1918-1988)-

En los experimentos de esta Tesis, se considera el caudal diario en el punto de control de la cuenca como la variable de interés a ser predicha. La nomenclatura empleada en los experimentos presenta los siguientes cambios respecto a los desarrollos teóricos de los capítulos previos. La variable de interés simulada por el modelo \mathbf{Y}_M , es el caudal simulado y se representa por \mathbf{q}_s . Las observaciones de la variable de interés $\tilde{\mathbf{Y}}$, son los caudales observados que se representan por $\tilde{\mathbf{q}}$. Finalmente, la predicción \mathbf{Y}^* es el caudal predicho, representado por \mathbf{q}^* .

5.1. SLS

Como se explicó en detalle en la sección 2.3.1, el objetivo del método SLS es minimizar el error cuadrático medio de la predicción, esto es, obtener residuos con media nula y mínima varianza. Ese objetivo tiende a provocar unos parámetros con poca incertidumbre siempre que el modelo no esté sobreparametrizado. De forma engañosa, la poca incertidumbre de los parámetros da a entender que la información de las observaciones apoya fuertemente su valor, cuando lo que realmente ocurre es que la función objetivo (función de verosimilitud) no está teniendo en cuenta todas las fuentes de error que afectan la inferencia. Se está sobrevalorando la información de los datos y/o la capacidad del modelo para procesarlos y obtener resultados ajustados a las observaciones. Al tiempo que presentan poca incertidumbre, los parámetros han sido forzados para que el modelo, aún con sus posibles errores de especificación, obtenga el mejor ajuste a los observados. Los parámetros así forzados, generalmente estarán muy alejados de su *valor correcto* (definición de *valor correcto* en sección 2.3.2). En resumen, los parámetros estimados con SLS, generalmente serán sesgados y con poca incertidumbre.

5.1.1. Validez del modelo de error: cumplimiento de sus hipótesis

Los resultados de la inferencia con el modelo de error SLS muestran, como era esperable, la menor varianza marginal de los errores de todos los experimentos realizados. Tanto con el modelo hidrológico CRR, $\sigma_E^2 = 0.34 \text{ mm}^2/\text{d}^2$, como con el modelo GR4J, $\sigma_E^2 = 0.43 \text{ mm}^2/\text{d}^2$.

Tabla 5.1. Índices de desempeño sobre diferentes aspectos de los resultados, con los principales modelos de error de los experimentos, y para ambos modelos hidrológicos. La primera fila muestra la Log-verosimilitud de los parámetros inferidos. Los índices de las tres siguientes filas, se relacionan con la bondad de la predicción media obtenida. Los índices de las cuatro últimas filas evalúan (GoES) los modelos de sesgo y varianza inferidos, así como la fiabilidad (exactitud) y precisión de la distribución predictiva respecto de las observaciones. En verde se marcan las mejores inferencias. En negrita se marcan los mejores valores

	CRR				GR4J			
	SLS	WLS	GL++	GL++Bias	SLS	WLS	GL++	GL++Bias
Log-L	-1625.5	-461.3	759.0	773.3	-1819.9	-726.1	742.2	749.9
NSE*	0.90	0.87	0.25	0.76	0.87	0.85	0.82	0.80
RMSE*	0.59	0.66	1.59	0.91	0.66	0.70	0.78	0.83
VE* (%)	5.5	2.7	32.9	0.0	0.8	-3.6	2.8	0.0
Std.dev[η]	---	0.88	0.76	1.04	---	0.79	0.95	1.5
$E[\eta]$	---	-0.09	-0.73	0.04	---	0.07	-0.26	-0.28
Fiabilidad	0.79	0.91	0.50	0.96	0.81	0.96	0.81	0.88
Resolución	3.68	5.66	4.70	3.94	3.13	4.45	4.30	4.40

* Valores para la predicción media

En la Tabla 5.1 se pueden comparar los valores de RMSE para todos los experimentos, los cuales son un buen estimador muestral de la desviación típica marginal de los errores. Otro resultado interesante es el valor de la media marginal de los errores, cuyo valor no es cero, luego no cumple con la hipótesis de media nula de la inferencia. Con el modelo CRR, la media marginal es $\mu_E = -0.11 \text{ mm/d}$, algo mayor que con el modelo GR4J en el que $\mu_E = -0.02 \text{ mm/d}$.

Tampoco se cumplen ninguna de las otras tres hipótesis en las que se basa el modelo de error, lo cual se hace evidente en la Figura 5.1.

Respecto a la hipótesis de **Normalidad** (Figura 5.1, paneles superiores), en los dos modelos hidrológicos se aprecia una distribución de errores con exceso de curtosis. En relación a la hipótesis de varianza constante u **homocedasticidad**, se observa en ambos modelos hidrológicos (paneles intermedios) cómo la nube de errores tiene la típica forma de embudo, en la que la dispersión de la nube crece al aumentar la magnitud del caudal simulado q_s .

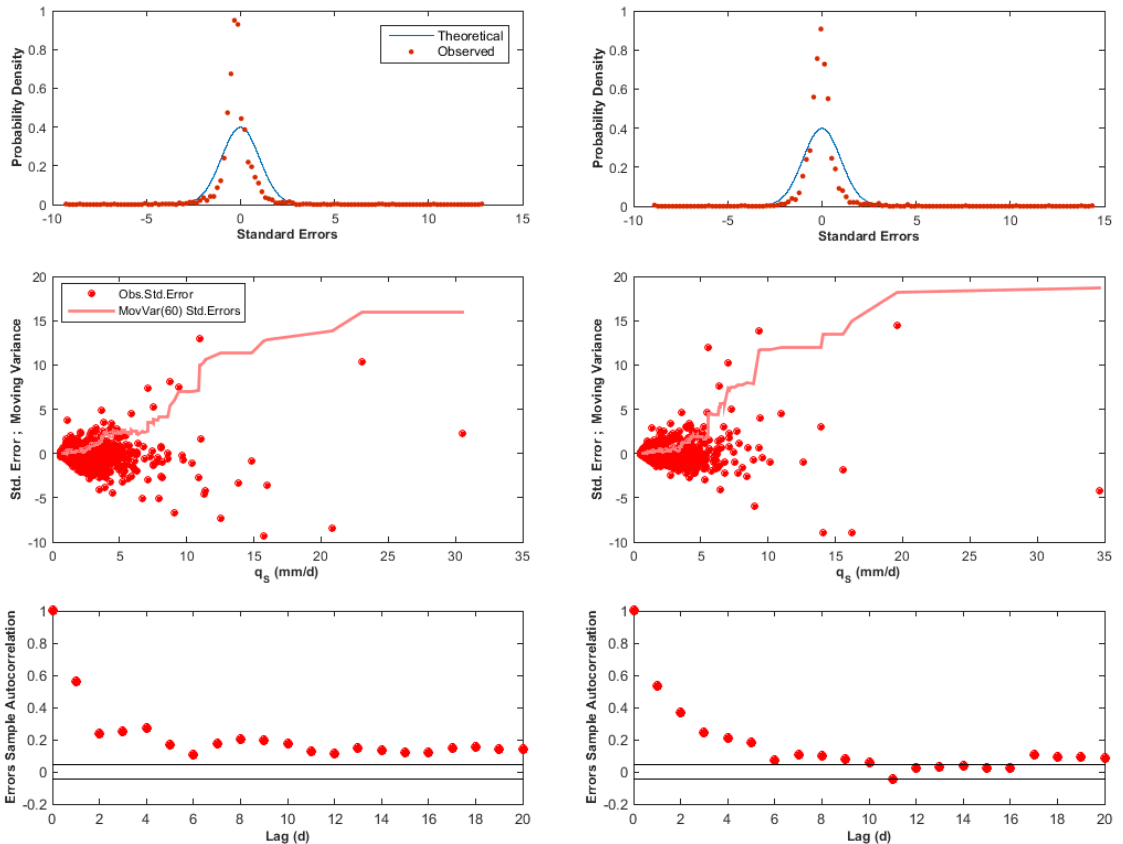


Figura 5.1. Comprobación de las hipótesis del modelo de error SLS, con los modelos hidrológicos CRR (columna izquierda) y GR4J (columna derecha). Normalidad de los errores estándar observados (paneles superiores). Evaluación de la homocedasticidad (paneles intermedios). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (paneles inferiores)

Para evaluar la magnitud de la heterocedasticidad de forma cuantitativa, se ha calculado una varianza móvil de los errores mediante una ventana móvil con 60 datos de longitud, sobre la cual se calcula la varianza. El resultado del cálculo se superpone a la nube de errores de la figura en forma de línea roja. La horizontalidad de esta línea indicaría una varianza condicional (a \mathbf{q}_s) de los errores constante, y como se observa, la línea representada está lejos de ser horizontal. La cuarta de las hipótesis del modelo de error SLS, es la **independencia** serial de los errores. Esta característica puede ser evaluada mediante la función de autocorrelación de los errores, mostrada para ambos modelos hidrológicos en los paneles inferiores de la Figura 5.1. Las gráficas muestran además, los límites de Anderson (Anderson, 1942) en cuyo interior quedan, con un 95% de confianza, los errores con autocorrelación nula. Como se observa, los errores de ambos modelos están autocorrelacionados al menos durante los primeros 10 lags.

5.1.2. Identificabilidad e incertidumbre de los parámetros

No cumpliéndose ninguna de las hipótesis del modelo de error con el que se infieren los parámetros, sus estimaciones no pueden ser correctas (Sorooshian y Gupta, 1983). Si se presta atención a la Tabla 5.1, se pueden analizar los valores de verosimilitud (log-verosimilitud para ser exactos) de los parámetros estimados en cada uno de los experimentos. De forma clara, los peores valores de verosimilitud son para el modelo de error SLS, con el que los parámetros del modelo CRR obtienen una $\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Omega} | \boldsymbol{\varepsilon}) = -1625$ y los parámetros del modelo GR4J una $\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Omega} | \boldsymbol{\varepsilon}) = -1819$. Sin embargo, como se observa en la Tabla 5.2 y en la Tabla 5.3, el modelo de error SLS estima los parámetros con menor incertidumbre, cuantificada mediante el coeficiente de variación (CV) de sus distribuciones posteriores. Como se dijo anteriormente, eso es generalmente válido, siempre que el modelo no esté sobreparametrizado, en cuyo caso la incertidumbre de los parámetros en exceso puede ser importante, llegando incluso a ser parámetros poco identificables. El modelo CRR muestra síntomas en ese sentido, ya que de los 7 parámetros que tiene, SLS (e incluso en mayor medida los otros modelos de error) muestra dos de ellos $\{\alpha_e, \alpha_f\}$ con un alto coeficiente de variación, es decir, poco identificables.

Tabla 5.2. Modelo hidrológico CRR: Máximo de la posterior (MAP) y coeficientes de variación (CV) de los parámetros. CVs mayores o iguales a 0.2 se marcan en rojo. Parámetros restringidos por las TLs, se somborean en azul

	WLS						GL++						GL++ Bias					
	SLS			NTL			TL			NTL			TL			TL		
	MAP	CV		MAP	CV		MAP	CV		MAP	CV		MAP	CV		MAP	CV	
CRR																		
I_{max}	10.00	0.06		2.32	0.18		2.07	0.17		0.49	0.47		0.48	0.42		0.96	0.42	
S_{max}	207	0.05		369	0.07		363	0.06		37	0.80		34	0.13		36	0.09	
Q_{smax}	7.60	0.07		3.67	0.08		3.76	0.07		16	0.76		12	0.12		8.84	0.11	
α_e	97	0.20		12	0.09		12	0.09		89	0.33		90	0.28		85	0.29	
α_f	-0.12	1.77		-0.64	0.43		-0.60	0.38		-0.36	3.88		-0.768	0.72		-1.26	0.20	
K_f	2.46	0.02		3.14	0.03		3.15	0.03		1.98	0.04		2.05	0.04		1.83	0.03	
K_s	91	0.05		46	0.07		47	0.07		62	0.10		66	0.08		90	0.11	
α	0.58	0.06		-0.08	0.08		-0.09	0.17		-0.29	0.09		-0.24	0.05		-0.21	0.42	
κ	----	----		0.24	0.03		0.27	0.02		0.52	0.06		0.42	0.03		0.33	0.09	
β	----	----		----	----		----	----		1.00	0.01		1.00	0.01		0.99	0.01	
ξ	----	----		----	----		----	----		1.03	0.03		1.04	0.03		0.95	0.02	
φ₁	----	----		----	----		----	----		0.98	0.00		0.95	0.00		0.94	0.00	
q₀	----	----		----	----		----	----		----	----		----	----		1.96	0.05	
γ	----	----		----	----		----	----		----	----		----	----		-0.25	0.11	
τ	----	----		----	----		----	----		----	----		----	----		-0.38	0.29	

Tabla 5.3. Modelo hidrológico GR4J. Máximo de la posterior (MAP) y coeficientes de variación (CV) de los parámetros. CVs mayores o iguales a 0.2 se marcan en rojo. Parámetros restringidos por las TLs, se somborean en azul

	WLS						GL++						GL++ Bias						
	SLS			NTL			TL			NTL			TL			TL			
	MAP	CV		MAP	CV		MAP	CV		MAP	CV		MAP	CV		MAP	CV		
GR4J	θ_1	991	0.03	1252	0.03	0.03	1254	0.03	0.03	1847	0.08	2210	0.08	2329	0.08	0.08	2329	0.08	0.08
	θ_2	-0.71	0.09	-1.36	0.07	0.07	-1.37	0.07	0.07	-4.92	0.10	-0.55	0.21	2.89	0.16	0.16	2.89	0.16	0.16
	θ_3	153	0.03	162	0.04	0.04	163	0.04	0.04	79	0.04	81	0.04	65	0.05	0.05	65	0.05	0.05
	θ_4	1.55	0.02	1.66	0.03	0.03	1.67	0.03	0.03	1.54	0.01	1.67	0.02	1.58	0.02	0.02	1.58	0.02	0.02
	α	0.655	0.00	-0.001	12.14	-0.003	5.20	0.00	1.07	-0.17	0.07	-0.21	0.09	0.09	0.09	0.09	-0.21	0.09	0.09
	κ	---	---	0.23	0.06	0.29	0.04	0.66	0.12	0.35	0.04	0.15	0.14	0.14	0.14	0.14	0.15	0.14	0.14
	β	---	---	---	---	---	---	1.00	0.01	1.00	0.01	1.00	0.01	1.00	0.01	0.01	1.00	0.01	0.01
	ξ	---	---	---	---	---	---	0.82	0.02	0.99	0.03	0.94	0.02	0.94	0.02	0.02	0.94	0.02	0.02
	ϕ_1	---	---	---	---	---	---	0.97	0.01	0.93	0.01	0.93	0.00	0.93	0.00	0.00	0.93	0.00	0.00
	ρ_b	---	---	---	---	---	---	---	---	---	---	1.57	0.13	0.13	0.13	0.13	1.57	0.13	0.13
	γ	---	---	---	---	---	---	---	---	---	---	-1.00	0.20	0.20	0.20	0.20	-1.00	0.20	0.20
	τ	---	---	---	---	---	---	---	---	---	---	-0.32	0.13	0.13	0.13	0.13	-0.32	0.13	0.13

5.1.3. Predicción y estimación de su incertidumbre

En cuanto a la exactitud de la predicción media, la Tabla 5.1 muestra para SLS los índices de Nash más altos (NSE=0.90 para CRR, y NSE=0.87 para GR4J). Otro aspecto que se puede observar en la Tabla 5.1 es el error en volumen de la predicción media, definido como $VE = (\mathbf{V}_s - \tilde{\mathbf{V}}) \tilde{\mathbf{V}}^{-1}$. Sus valores se muestran coherentes con los valores, antes comentados de la media marginal de los errores: mayor para CRR que para GR4J.

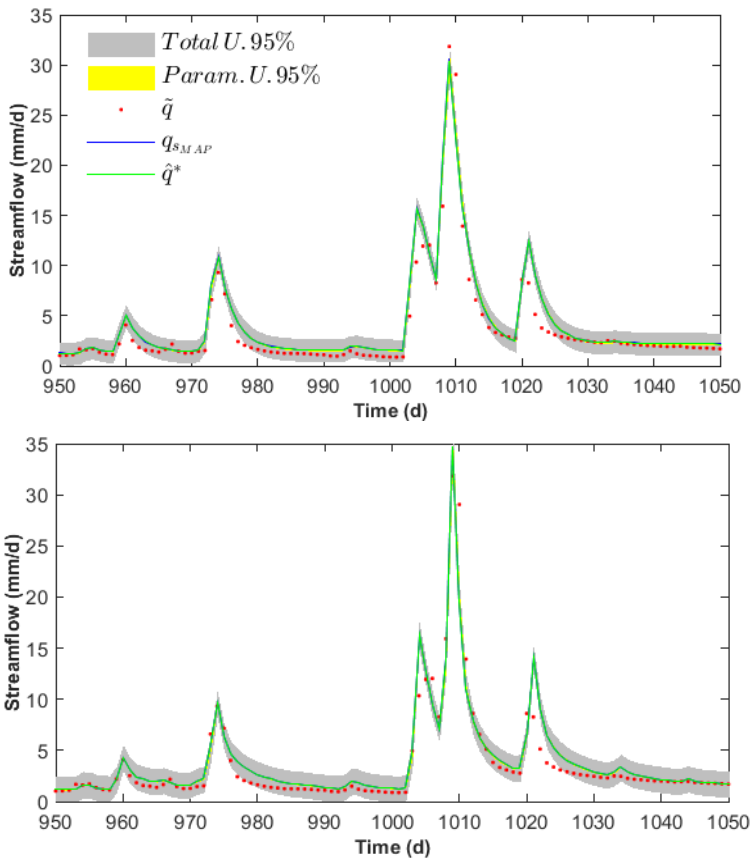


Figura 5.2. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada con SLS para el modelo hidrológico CRR (panel superior) y GR4J (panel inferior). Como se puede apreciar la contribución de los parámetros es, en este caso, despreciable

En la Figura 5.2 se representan las observaciones \tilde{q} , la predicción media q^* y la simulación del modelo hidrológico q_s (estas dos últimas coinciden en SLS) para el conjunto de parámetros óptimo (MAP). En la figura se observa el buen ajuste del modelo CRR (panel superior) al máximo pico del periodo de calibración, aunque el ajuste a los picos secundarios no es tan bueno y en general, los caudales bajos son sobreestimados. El modelo GR4J (panel inferior) obtiene, en líneas generales, unos resultados con similares características.

También se pueden observar en la Figura 5.2, la banda (95% de probabilidad) de incertidumbre de la predicción media, dibujada en color gris. Con el modelo de error SLS, la hipótesis de varianza de los errores constante se traduce en una banda de incertidumbre cuya anchura es constante para cualquier magnitud de caudal q_s . No obstante, debido a un efecto puramente visual, la gráfica no reproduce fielmente esta propiedad. Nótese, que en la figura también se está representando (en color amarillo) la banda de incertidumbre debida exclusivamente a los parámetros. Su extrema delgadez indica, que con SLS, es escasa la contribución de la incertidumbre de los parámetros a la incertidumbre total de la predicción, lo que principalmente se debe a la escasa incertidumbre estimada para los parámetros.

De manera más formal, se puede evaluar en términos de exactitud y precisión la incertidumbre de la predicción media obtenida. Esta comprobación se realiza mediante las gráficas de tipo PP-plot (Laio y Tamea, 2006; Renard et al., 2010; Thyer et al., 2009) y se interpreta según lo explicado en la sección 2.4.2. Se observa en la Figura 5.3, para el modelo CRR (panel superior), una distribución predictiva obtenida a partir del modelo de error SLS (línea azul) con un moderado sesgo sistemático de sobrepredicción, además de una importante sobreestimación de la incertidumbre. Respecto del modelo GR4J (panel inferior), la distribución predictiva obtenida con SLS, a penas muestra sobrepredicción, pero sí una considerable sobreestimación de la incertidumbre.

5.1.4. SLS. Conclusiones

Con los resultados obtenidos, se ha mostrado lo paradójico que resulta calibrar con SLS un modelo hidrológico. De entre todos los experimentos que se realizan

en este capítulo, SLS obtiene los parámetros con menor incertidumbre, al tiempo que dichos parámetros presentan la peor verosimilitud. Se obtiene el mejor ajuste de la predicción media a las observaciones, pero se está sobreestimando su incertidumbre.

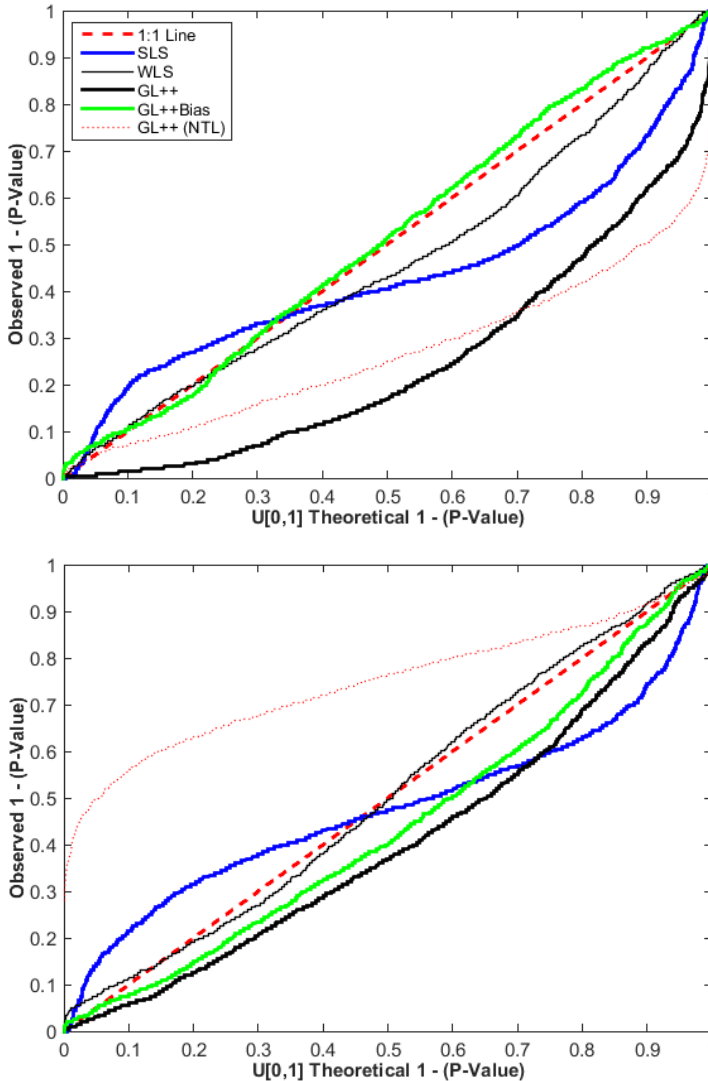


Figura 5.3. PP-plots de todas las distribuciones predictivas inferidas con los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior)

El SLS queda invalidado como modelo de error por estar asumiendo las 4 hipótesis del caso ideal de inferencia (sección 2.3.1), mientras que los resultados evidencian que no se cumple ninguna. En este sentido, los autores Sorooshian y Dracup (1980), Schoups y Vrugt (2010), Sorooshian y Gupta (1983) o Thyer et al. (2009) entre otros, afirman que la calibración de modelos con el método SLS, cuando se produce la violación de sus hipótesis, sesga los valores estimados de los parámetros, de la predicción y afectan negativamente a la estimación de sus respectivas incertidumbres. Estos efectos negativos ocurren con SLS, y en mayor o menor medida, con cualquier modelo de error que no cumpla con sus hipótesis.

5.2. WLS y WLS-NTL

El modelo de error **WLS** (sección 2.3.3) es el resultado del relajamiento de la hipótesis de varianza constante del modelo SLS. Si bien no es un modelo de error válido en modelación hidrológica, por incumplir como el SLS algunas de sus hipótesis, merece especial atención. De la lectura de la sección 2.3.3 se desprende la relativa importancia de este modelo de error. Este método de estimación de parámetros se está aplicando de forma errónea en la literatura reciente, por la incorrecta aplicación del modelo de varianza de los errores. Lo que vuelve el problema sinuoso y retorcido es el hecho de que la aplicación incorrecta del WLS, no produce efectos sobre los únicos resultados de la inferencia que hasta ahora interesaban, esto es, los parámetros y la predicción del modelo hidrológico.

No sería de trascendencia la incorrección en la aplicación del modelo de error WLS, de no afectar a la inferencia conjunta de otros modelos de error más sofisticados, que aplican de la misma forma incorrecta el modelo de varianza, pero en los cuales sí se produce la alteración de cualquiera de sus resultados de inferencia (estimación de parámetros, predicción e incertidumbres). Aclarando el problema, sin consecuencias prácticas, de la mala aplicación del WLS, se resolverán los problemas aparecidos en inferencias con otros modelos de error.

5.2.1. La paradoja en los resultados de WLS-NTL

El primero de los experimentos de esta sección, llamado **WLS-NTL**, trata de modelar la varianza condicional de los errores mediante el método directo. Para ello el modelo de error aplica el modelo de varianza definido por la ecuación (3.14) sin ninguna otra consideración o restricción sobre los parámetros del modelo de error $\Omega = \{\alpha, \kappa\}$. Este modelo de error, así definido, tiene su equivalente en la literatura en el trabajo de Evin et al. (2014), donde se le llama WLS. Cabe mencionar, que en la literatura no se ha encontrado una aplicación del WLS calculando los pesos de forma directa, como la efectuada en ese trabajo.

El segundo de los experimentos aplica el modelo de error **WLS**, mediante la restricción del parámetro α según la ecuación (3.38). Los resultados son los esperados. Los dos parámetros del modelo de error inferidos resultan diferentes en WLS y WLS-NTL. Sin embargo, ese resultado por sí solo no aporta demasiado, sin la confirmación esperada. Para el modelo hidrológico CRR, cuando no se le aplica la TVL al modelo de error (WLS-NTL), la varianza marginal de los errores es $\sigma_E^2 = 0.44 \text{ mm}^2/d^2$, mientras que las varianzas condicionales tienen una media de $\mathbb{E}[\sigma^2] = 0.33 \text{ mm}^2/d^2$. Este resultado confirma el desacople entre las varianzas condicionales y la marginal de los errores. Cuando se aplica la TVL (WLS) la varianza marginal y las condicionales coinciden $\sigma_E^2 = \mathbb{E}[\sigma^2] = 0.44 \text{ mm}^2/d^2$.

Como muestran los valores, la varianza marginal no ha cambiado al variar el modelo de error. Es la varianza condicional la que se ha acoplado a la marginal cuando se ha considerado la TVL. Este hecho ejemplifica la discusión acerca del desacople, expuesta en la sección 3.1.2. La distribución marginal, definida por la nube de errores, ha estado condicionada únicamente por el modelo hidrológico y no por el modelo de error. Es decir, el cambio de modelo de error de WLS a WLS-NTL no ha afectado a la distribución marginal de los errores y por tanto, se deduce que tampoco al modelo hidrológico. Dicho de otra forma, la inferencia con el modelo de error WLS es similar a una inferencia en posproceso (sección 2.1.1): no hay interacción entre los parámetros de error y los parámetros

hidrológicos. Respecto a la estimación de los parámetros de error, con el modelo WLS-NTL se estima un $\kappa_{\text{WLS-NTL}}=0.24$, valor estimado inferior al del modelo WLS con el que $\kappa_{\text{WLS}}=0.27$. La magnitud de estos resultados va en concordancia con las magnitudes de $\mathbb{E}[\sigma^2]$ obtenidas con cada modelo de error y mostradas previamente.

Los resultados para el modelo hidrológico GR4J van en el mismo sentido. Con el modelo de error WLS-NTL la varianza marginal es $\sigma_{\text{E}}^2 = 0.49 \text{ mm}^2/\text{d}^2$, mientras que la media de las varianzas condicionales es $\mathbb{E}[\sigma^2] = 0.33 \text{ mm}^2/\text{d}^2$. Cuando se aplica la TVL (WLS) la varianza marginal y las condicionales coinciden $\sigma_{\text{E}}^2 = \mathbb{E}[\sigma^2] = 0.49 \text{ mm}^2/\text{d}^2$, con lo que se confirma el mismo efecto de acople unilateral de las condicionales a la marginal. En la estimación de los parámetros de error, el modelo WLS-NTL estima un $\kappa_{\text{WLS-NTL}}=0.23$, valor estimado inferior al modelo WLS en el que $\kappa_{\text{WLS}}=0.29$.

Si se observan los parámetros de los modelos hidrológicos (Tabla 5.2 y Tabla 5.3), se puede comprobar que sus estimaciones no difieren significativamente entre los modelos WLS y WLS-NTL, lo que confirma el hecho de que al modelo hidrológico no le afecta el empleo de uno u otro modelo de error, a pesar de que WLS-NTL no mantiene la coherencia estadística. Con los valores óptimos de la función de verosimilitud ocurre lo mismo, no cambian significativamente al cambiar el modelo de error.

Este es un hecho paradójico, que probablemente ocurra muchas veces en ciencia: algo funciona, a pesar de ser teóricamente incorrecto. Su explicación, en este caso, acaba siendo lógica y sencilla. El método WLS, si se recuerda, trata de estimar unos pesos que le dan mayor o menor importancia a cada observación, habitualmente en función de su incertidumbre estimada a partir de su varianza $w_j = \sigma_{e|y_j, \mathcal{M}}^{-1} \quad \forall j \in \{1, \dots, n\}$. No es difícil ver, que dada la solución estadísticamente coherente del valor de esos pesos llámese $\{w_j\}_{\text{ML-TL}}$, cualquier

otra solución $\{w_j\}_{ML}$ que sea proporcional, esto es $\{w_j\}_{ML} \propto \{w_j\}_{ML-TL}$, deberá tener el mismo efecto en la función de verosimilitud, dado que los pesos relativos entre observaciones no habrán cambiado. Trasladando esta idea al modelo de error, se cumple que son válidos todos los conjuntos de parámetros tales que $\{\alpha, \kappa\}_{ML} = c^{-1} \{\alpha, \kappa\}_{ML-TL}$ con $c > 1$, donde el subíndice ML indica “estimador de máxima verosimilitud”, y el subíndice TL indica “considerando las Leyes Totales”. En la Tabla 5.2 y la Tabla 5.3 se puede comprobar el valor de esa constante de proporcionalidad en la inferencia con ambos modelos hidrológicos $c_{CRR} = 1.125$ y $c_{GR4J} = 1.26$.

La otra coincidencia entre los resultados de WLS y WLS-NTL, se produce en la generación de la banda de incertidumbre, lo cual también resulta paradójico, siendo que $\kappa_{ML-TL} > \kappa_{ML}$. La explicación vuelve a estar relacionada con un efecto de, llámese proporcionalidad, compensación de términos, etc., que se explica en lo que sigue. Supóngase que se han inferido los parámetros $\{\alpha, \kappa\}$ del modelo de varianza. A partir de ellos se pueden obtener los errores estandarizados $\boldsymbol{\eta}$, mediante la ecuación (3.15). Se calcula la desviación típica de esos errores como $\sigma_{\boldsymbol{\eta}} = (\mathbb{V}[\boldsymbol{\eta}])^{0.5}$. Para entrar en la función de verosimilitud dada por la ecuación (3.31) es necesario estandarizar previamente los residuos, de forma que $\mathbf{a} = \boldsymbol{\eta} \sigma_{\boldsymbol{\eta}}^{-1} \stackrel{\mathcal{D}}{\sim} P_{SEP}(0, 1, \xi, \beta)$. Supóngase ahora que se consideran otros parámetros de error, proporcionales a los anteriores, tal que $\{\alpha, \kappa\}^* = c \{\alpha, \kappa\}$, con $c > 1$, por lo que $\boldsymbol{\eta}^* = \boldsymbol{\eta} c^{-1}$ y por tanto $\sigma_{\boldsymbol{\eta}}^* = \sigma_{\boldsymbol{\eta}} c^{-1}$. Los nuevos residuos estandarizados, para entrar en la función de verosimilitud serán $\mathbf{a}^* = \boldsymbol{\eta}^* \sigma_{\boldsymbol{\eta}}^{*-1}$. Y de esta expresión se puede obtener mediante las apropiadas sustituciones $\mathbf{a}^* = \boldsymbol{\eta} c^{-1} \sigma_{\boldsymbol{\eta}}^{-1} c = \mathbf{a}$. Por tanto, cualquier conjunto de parámetros proporcional $\{\alpha, \kappa\}^* = c \{\alpha, \kappa\}$ proporciona los mismos residuos $\mathbf{a}^* = \mathbf{a}$. Como consecuencia, las bandas de incertidumbre, calculadas como se explica en la sección 3.5, serán idénticas.

5.2.2. Validez del modelo de error WLS: cumplimiento de sus hipótesis

Como ya se ha comentado, aun estimando de forma estadísticamente coherente los parámetros del WLS, no es un modelo de error apropiado en modelación hidrológica, por el incumplimiento de las hipótesis que lo sustentan. A la vista de los resultados, la inferencia no cumple con la hipótesis de Normalidad, ni con la de independencia. El modelo de error sí consigue estabilizar la varianza de los errores, cumpliéndose por tanto la hipótesis de varianza constante. También se cumple, en este experimento y de forma aproximada, la hipótesis de errores con media nula.

La correcta estabilización de la varianza se produce con ambos modelos hidrológicos. La Figura 5.4 (panel superior) muestra la estabilización de la varianza para el modelo GR4J. En la figura se representan la nube de errores (en rojo) y la nube de errores estabilizados (negro), así como las líneas respectivas que cuantifican la evolución de la varianza, tal como se explicó en el experimento con el SLS. Como se observa, la varianza móvil calculada para los errores presenta la misma apariencia que en el experimento SLS. Sin embargo, la varianza móvil calculada para los errores estabilizados es prácticamente una línea horizontal, representativa de una varianza constante.

El panel inferior de la Figura 5.4, en el que se representan las funciones de autocorrelación de los errores (en rojo) y de los errores estabilizados (azul), muestra como la estabilización de la varianza provoca un importante aumento en la dependencia serial de los errores. Mientras WLS está resolviendo el problema de la heterocedasticidad, al mismo tiempo está agravando el problema de la dependencia. Cabe recordar que: i) Los errores que exhiben un alto grado de autocorrelación están indicando un error de especificación en el modelo que los genera y ii) Los errores de especificación disminuyen el poder predictivo de los modelos, aunque aparentemente muestren un buen desempeño en calibración, debido al sobreajuste a las observaciones mediante el forzamiento del valor de sus parámetros.

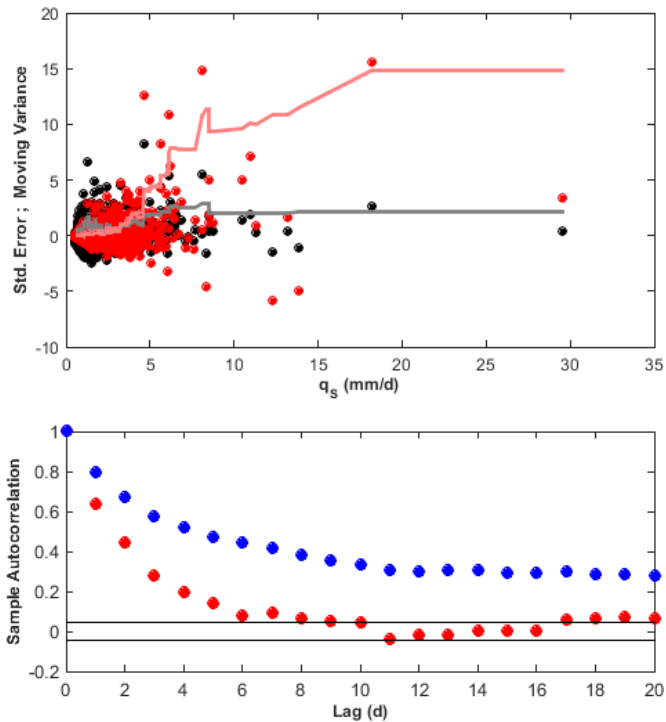


Figura 5.4. Comprobación de las hipótesis del modelo de error WLS, con el modelo hidrológico GR4J. Evaluación de la homocedasticidad (panel superior). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (panel inferior)

5.2.3. Identificabilidad e incertidumbre de los parámetros

Con WLS la incertidumbre de los parámetros no presenta, en general, diferencias significativas respecto de la mostrada para el modelo de error SLS. Esto se debe a que ambos modelos de error tienen, un objetivo similar en lo fundamental (minimizar la suma de errores cuadráticos), aunque no es similar en la forma debido a la ponderación selectiva que aplica el WLS sobre los residuos. Respecto a la contribución de la incertidumbre de los parámetros, a la incertidumbre total de la predicción, ocurre lo mismo que con SLS: su efecto es despreciable.

5.2.4. Predicción y estimación de su incertidumbre

La predicción media obtenida con WLS presenta un ajuste a las observaciones con una eficacia muy similar a la obtenida con el SLS. En la Tabla 5.1 WLS muestra un índice NSE ligeramente inferior al conseguido por el SLS. Esta reducción se debe a la menor ponderación que durante la calibración, WLS otorga a los picos de caudal (mayores errores), en beneficio de una mayor ponderación de los caudales bajos. Esta circunstancia produce un peor ajuste a los picos (muy influyentes en el índice NSE), pero también un porcentaje de error en volumen (VE) menor que en el SLS.

Respecto a la incertidumbre de la predicción, WLS mejora la estimación de la incertidumbre de la predicción respecto a SLS. Esto puede observarse en los PP-plots de la Figura 5.3. No obstante, a juicio del que escribe, tiene poco o ningún sentido valorar cual de las incertidumbres predictiva es la más correcta, de entre los modelos de error SLS y WLS. Ambos modelos de error incumplen una de las hipótesis más importantes: la independencia de los errores. Una inferencia que produce residuos con dependencia serial, está indicando que hay problemas estructurales en el modelo hidrológico y/o errores en los datos. Problemas que no son tenidos en cuenta por el modelo de error. Si el modelo de error es incorrecto y es el generador de la distribución predictiva (sección 3.5), la incorrección de uno producirá la incorrección en la otra. Dicho de otra forma, así como es absurdo decir que SLS es mejor modelo de error que WLS por obtener un mejor índice de Nash, carece igualmente de sentido decir que WLS es mejor modelo de error que SLS por obtener una distribución predictiva con mejores índices de fiabilidad. Ambos son modelos de error incorrectos, para ser usados en la inferencia de modelos hidrológicos.

5.2.5. WLS y WLS-NTL. Conclusiones

El primero de los criterios para evaluar un modelo de error debería ser el cumplimiento de sus hipótesis. No deberían darse por válidos modelos de error cuyos residuos (o innovaciones) no cumplan al menos, con dos de las hipótesis de un *ruido blanco*: la de independencia y la de homocedasticidad. Incluir al SLS y WLS, modelos que no cumplen sus hipótesis, en el conjunto de experimentos de esta Tesis, tiene el único valor de haber probado modelos de error que han

sido y son referencia en la calibración de modelos hidrológicos. Incluir en los experimentos el WLS-NTL, tiene además el valor de mostrar la raíz de todos los problemas que van ser mostrados en los siguientes experimentos, con modelos de error más sofisticados. El modelo de error WLS-NTL es un modelo de error estadísticamente incoherente, cuya inferencia de los parámetros de error es un problema mal planteado, por no poderse identificar un único conjunto de parámetros que sea solución del problema de inferencia. Sin embargo, siendo incoherente WLS-NTL, permite la obtención de unas soluciones válidas, al menos, para resolver el problema que originalmente resolvía el WLS: la obtención de unos pesos. Ahora bien, trasladar esa incoherencia estadística fuera del ámbito del WLS, es algo que no ha resultado inocuo, como muestran diferentes aportaciones a la literatura ya mencionadas en esta Tesis y como se mostrará a continuación.

5.3. GL++ y GL++NTL

5.3.1. Validez del modelo de error: cumplimiento de sus hipótesis

Los resultados de las inferencias realizadas sobre ambos modelos hidrológicos, con el modelo de error GL++ muestran de forma aceptable, el cumplimiento de la mayoría de sus hipótesis. Como se observa en la Figura 5.5 (paneles superiores), los residuos obtenidos de la inferencia cumplen adecuadamente con el ajuste a la pdf SEP, si bien el modelo CRR (columna izquierda) presenta unas colas más pesadas que el modelo GR4J. Respecto a la dependencia serial (paneles inferiores), la figura muestra como la autocorrelación de las innovaciones (puntos azules) está prácticamente eliminada en comparación con la de los errores (puntos rojos). No obstante, la eliminación no es perfecta, lo cual puede ser debido a la necesidad de usar un modelo AR de orden superior, o incluso, a que haya dependencias de tipo no lineal.

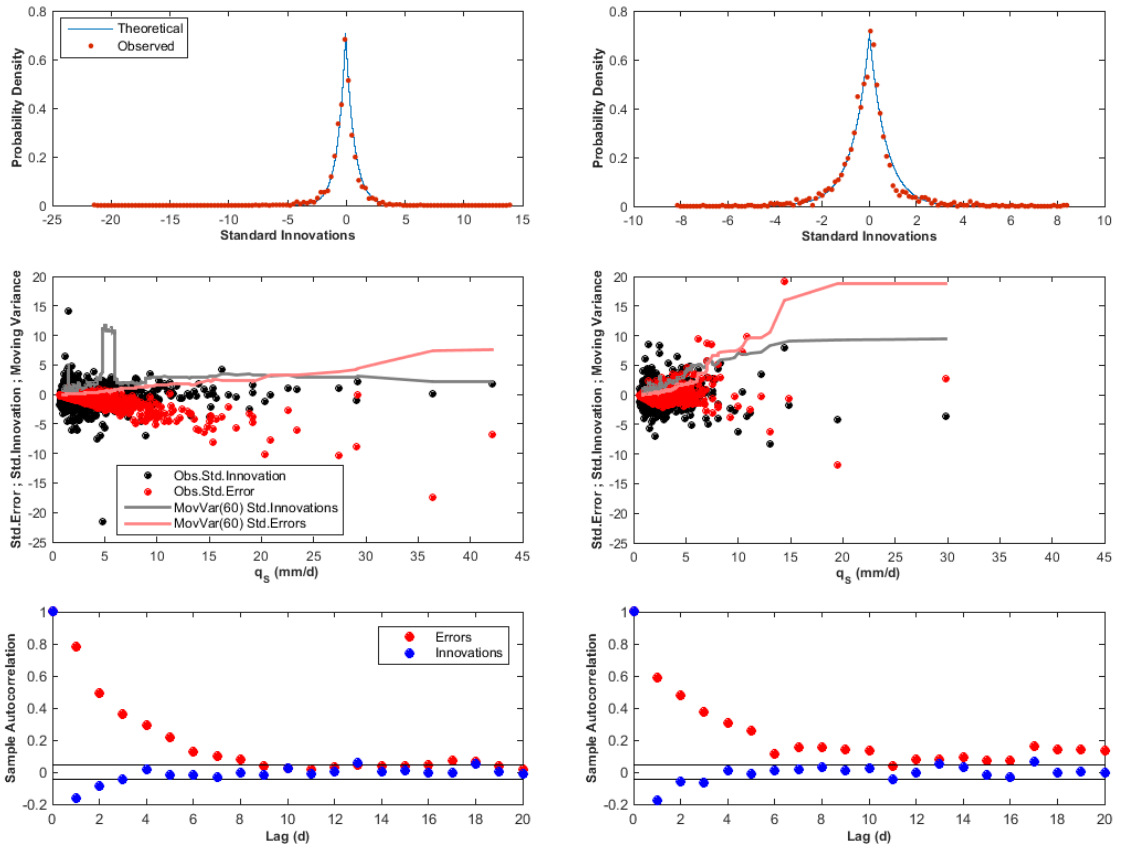


Figura 5.5. Comprobación de las hipótesis del modelo de error GL++, con los modelos hidrológicos CRR (columna izquierda) y GR4J (columna derecha). Ajuste a SEP de los errores estándar observados (paneles superiores). Evaluación de la homocedasticidad (paneles intermedios). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (paneles inferiores)

En lo referente a la estabilización de la varianza, la figura muestra (paneles intermedios) cómo las varianzas condicionales de errores e innovaciones son muy similares para la mayor parte del rango de caudales simulados. El modelo de error GL++ no está siendo capaz de estabilizar la varianza de forma efectiva, siendo esto más evidente con el modelo GR4J (columna derecha) que con el modelo CRR. La razón de este pobre resultado en la estabilización de la varianza, que se explicará en un párrafo posterior, tiene su origen en que el modelo de error GL++ provoca la aparición de un importante sesgo en los

errores, cuya modelación no es contemplada por las hipótesis de este modelo de error. El sesgo generado provoca una sobrepredicción de los caudales. Esto se puede comprobar en la Figura 5.5 (paneles intermedios), siendo especialmente evidente con el modelo CRR (columna izquierda), para el que se observa como la media de los valores de la nube de errores, cuya esperanza marginal es $\mu_E = -0.69$ mm/d, exhibe un crecimiento negativo, al aumentar la magnitud del caudal simulado. El sesgo generado en la inferencia con GL++ también puede ser observado en los PP-plots de la Figura 5.3 donde las curvas correspondientes al modelo de error GL++ tienen una pronunciada curvatura, especialmente en el modelo CRR, bastante alejada de las curvas en forma de “S” o “S” invertida. La Tabla 5.1 muestra los efectos del sesgo. En ella se aprecia cómo GL++ sufre un fuerte deterioro en el índice de Nash y en el error en volumen, respecto a SLS y WLS.

Nótese, cómo GL++ produce una significativa diferenciación en el rendimiento de los dos modelos hidrológicos, cuya consecuencia es el importante deterioro en el desempeño del modelo CRR, en comparación con el modelo GR4J, que todavía mantiene un desempeño aceptable. En las inferencias con los modelos de error SLS y WLS, no había diferencias significativas en el desempeño de ambos modelos hidrológicos. Esta circunstancia apunta a que GL++ es capaz de juzgar de forma diferente el funcionamiento de modelos estructuralmente diferentes, lo que es positivo si se quiere comparar el desempeño real de diferentes modelos hidrológicos. En este sentido, el presente experimento muestra que el modelo GR4J, con solo 4 parámetros, parece ser mejor modelo que el CRR, con 7 parámetros. El hecho de que GL++ provoque menos sesgo en la predicción de GR4J, podría estar indicando una menor presencia de problemas estructurales en este modelo hidrológico.

Al haber sesgo, GL++ no cumple con su hipótesis de errores con media nula. Tampoco consigue una correcta estabilización de la varianza de los errores, debido al pobre comportamiento del modelo de varianza, en su labor de estabilización. La razón está relacionada con el sesgo, como se indicó anteriormente. El modelo de error GL++, por hipótesis, considera a las distribuciones condicionales con media cero (sin sesgo). Sin embargo, los resultados muestran su importante desplazamiento respecto a dicho valor. No es posible estimar correctamente la amplitud (varianza) de las distribuciones

condicionales, cuando el modelo de error no está estimando cual es la posición de dichas distribuciones.

Dicho lo anterior, comparando con SLS y WLS, el modelo de error GL++ presenta un mejor cumplimiento de sus hipótesis, por lo que se considera un mejor modelo de error. Según numerosos autores (Schoups y Vrugt, 2010; Sorooshian y Dracup, 1980; Sorooshian y Gupta, 1983; Thyer et al., 2009; Vrugt et al., 2005) la correcta modelación de la heterocedasticidad, la dependencia y la no Normalidad de los errores, teóricamente proporciona parámetros más robustos (menos sesgados) así como distribuciones predictivas más robustas, que los modelos clásicos de error. La robustez que pueda proporcionar GL++ a sus resultados, se atribuye a los siguientes factores: i) Modelar la heterocedasticidad concede menos peso a los picos de caudal, lo que se traduce en unos parámetros cuya estimación es menos sensible a esos picos en diferentes periodos de validación: ii) La distribución SEP puede modelar colas más pesadas que la Normal, permitiendo mayor número de errores grandes, lo que induce robustez frente a posibles *outliers* en los caudales observados y iii) Modelar la autocorrelación, caso especial de sesgo, permite el filtrado de errores de tipo estructural y errores sistemáticos en los datos.

Predicción

Cuando hay problemas con el modelo y/o con los datos, no es posible conseguir al mismo tiempo un buen ajuste a las observaciones y unos parámetros robustos. Para conseguir una cosa, se ha de sacrificar la otra. A la vista de los resultados de GL++, parece razonable que un sesgo en la predicción del modelo hidrológico, presente como contrapartida la mayor robustez o menor sesgo de los parámetros. Una estimación de parámetros más plausible mostraría en forma de sesgo, las deficiencias existentes en el modelo hidrológico y/o en los datos. Mientras SLS sacrifica la robustez de los parámetros, GL++ sacrifica el buen (aunque engañoso) ajuste a las observaciones. Si se asume como bueno el razonamiento previo, el modelo de error GL++ está actuando como una (muy general) herramienta de diagnóstico del modelo hidrológico, mostrando de forma más realista lo que el modelo es capaz de ofrecer. Así, GL++ está diciendo: "Estos son los mejores resultados que se pueden conseguir con estos datos y

este modelo hidrológico. Si se quieren mejores predicciones hay que mejorar el modelo hidrológico, los datos o ambos”.

Incertidumbre de los parámetros

Con GL++ también se produce un aumento en la incertidumbre de los parámetros (Tablas 5.2 y 5.3), lo cual va en concordancia con la idea de que los problemas estructurales y/o los errores en los datos introducen incertidumbre en los parámetros. Puesto que SLS y WLS no consideran esas fuentes de incertidumbre, sobreestiman la información capturada por los parámetros, mostrando así una aparente certidumbre.

5.3.2. Efectos de la (no) aplicación de las Leyes Totales

Desacople entre la distribución marginal de los errores y las condicionales

En el experimento con el modelo de error WLS-NTL, se comprobó como su incorrección en el plano teórico no generaba consecuencias en lo práctico. Se dejó entrever, sin embargo, que el mantenimiento de esa incorrección en la aplicación de un modelo de varianza, con modelos de error más sofisticados, podría no quedar sin consecuencias. El modelo GL++ es buen ejemplo de ello. Como ya se ha explicado, GL++ tiene 4 parámetros libres $\Omega = \{\kappa, \phi_1, \xi, \beta\}$, más el parámetro α restringido por las TLs. El mismo modelo sin aplicar las TLs, GL++NTL, tiene por tanto los cinco parámetros libres $\Omega = \{\alpha, \kappa, \phi_1, \xi, \beta\}$ y se corresponde con el modelo del segundo de los 2 primeros experimentos (fallidos) de la Tesis. Los diferentes resultados obtenidos por ambos modelos de error confirman lo esperado.

La Figura 5.6 muestra para los dos modelos hidrológicos, las nubes de error generadas con GL++ (en rojo) y con GL++NTL (en gris). Sobre las nubes de error se han superpuesto las rectas que representan el modelo lineal de varianzas condicionales, para GL++ (en negro) y para GL++NTL (en gris). Las gráficas muestran una sobreestimación del modelo de varianzas condicionales en el caso GL++NTL, ligeramente mayor con el modelo hidrológico GR4J (panel inferior).

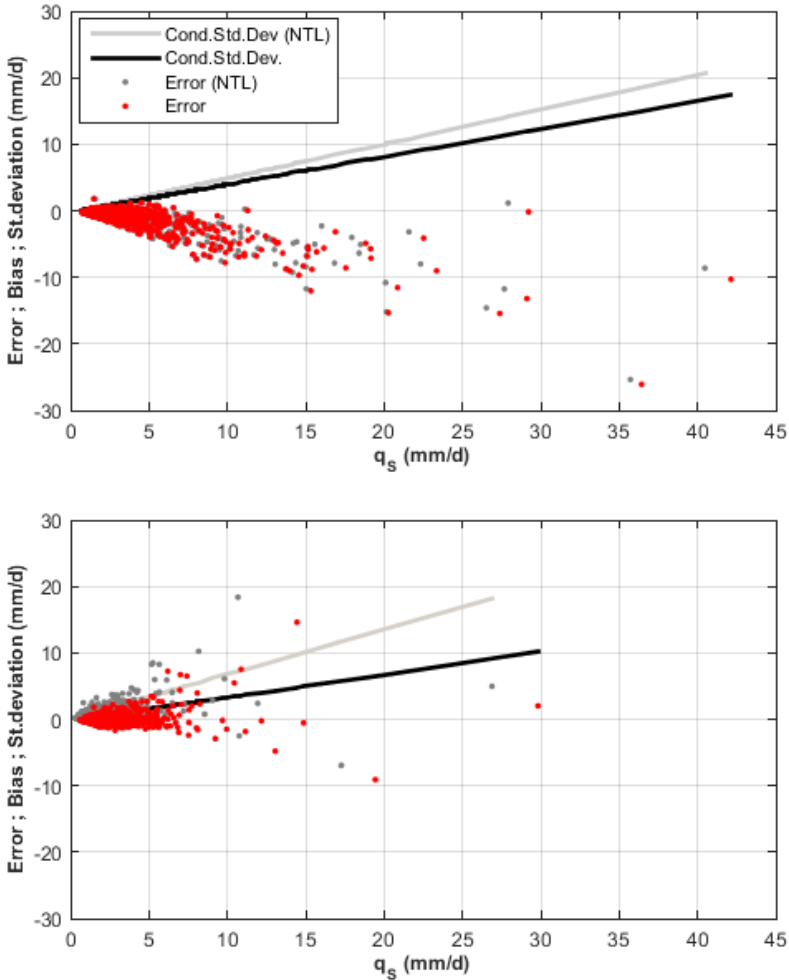


Figura 5.6. Elementos de los modelos de error GL++ y GL++NTL, para los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior). Las gráficas muestran las nubes de error en rojo para GL++ (aplica TVL) y en gris para GL++NTL. Superpuestas aparecen las líneas inferidas que representan a los modelos de varianza, para GL++ en negro, y para GL++NTL en gris

El desacople entre la varianza marginal de los errores y las varianzas condicionales que se produce con el modelo de error GL++NTL, es muy importante. Por ejemplo, para el modelo hidrológico CRR, la varianza marginal de los errores es $\sigma_E^2 = 1.97 \text{ mm}^2/\text{d}^2$, mientras que las varianzas condicionales

tienen una media de $\mathbb{E}[\sigma^2] = 3.31 \text{ mm}^2/\text{d}^2$. Cuando se aplica el modelo de error GL++, hay coherencia entre ambos tipos de distribución, obteniéndose que $\sigma_E^2 = \mathbb{E}[\sigma^2] = 2.25 \text{ mm}^2/\text{d}^2$.

Como muestran esos resultados numéricos, hay dos diferencias respecto a lo que ocurría con WLS y WLS-NTL. En primer lugar, mientras WLS-NTL infraestimaba las varianzas condicionales respecto de la marginal de los errores, GL++NTL las sobreestima. Por tanto, cada modelo de error representa a uno de los dos casos expuestos en la sección 3.1.2. En segundo lugar, a diferencia de lo que ocurre con WLS y WLS-NTL, la varianza marginal de los errores no es la misma con GL++ y GL++NTL. De ello se deduce que el acople de la marginal y las condicionales es bilateral en este caso, es decir, el modelo hidrológico y el modelo de error han interactuado, cambiando ambos el valor que tienen sus respectivos parámetros con GL++NTL (ver Tablas 5.2 y 5.3) para converger hacia la solución estadísticamente coherente dada por GL++.

Identificabilidad, correlación e incertidumbre de los parámetros

Otros aspectos importantes de la inferencia son, el de la identificabilidad, la incertidumbre y la correlación entre los parámetros inferidos. Los resultados van a mostrar distintos efectos espurios (solamente se producen en la solución estadísticamente incoherente) relacionados con esas tres propiedades. La Figura 5.7 muestra dos ejemplos en la inferencia del modelo hidrológico CRR.

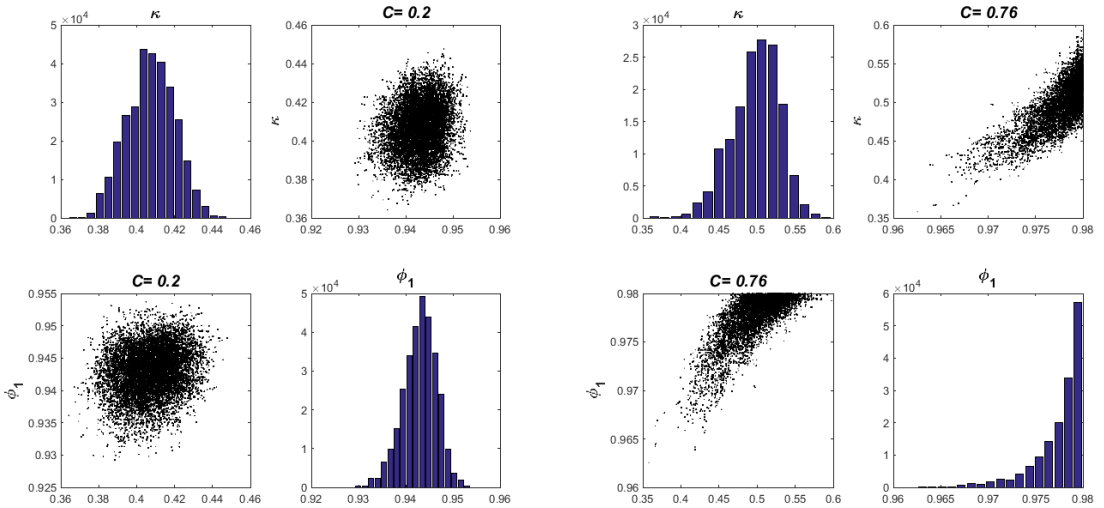


Figura 5.7. Modelo hidrológico CRR. Distribución posterior del parámetro de linealidad del modelo de varianza, distribución posterior del parámetro de autocorrelación y nube de valores muestrales entre ambos parámetros. Se representa el caso con el modelo de error GL++ (panel izquierdo) y el caso con el modelo de error GL++NTL (panel derecho)

El primero de ellos está relacionado con la correlación entre el parámetro de autocorrelación ϕ_1 y la pendiente κ del modelo lineal de varianza de los errores. En la izquierda de la figura, se representan las posteriores de estos parámetros estimadas con GL++: como se observa, la correlación entre ambos parámetros es poco significativa. En la derecha de la figura, se representan las posteriores de ambos parámetros estimadas con GL++NTL: con este modelo de error la correlación ha aumentado considerablemente, teniendo los parámetros un coeficiente de correlación lineal $C=0.76$. Evin et al. (2013) reportan idéntico efecto entre los mismos parámetros, después de aplicar la corrección metodológica que proponen, es decir, aplicando un modelo de error como GL++NTL, pero con la hipótesis de residuos Gaussianos.

El segundo de los ejemplos está relacionado con la identificabilidad del parámetro ϕ_1 . Se observa que con GL++ (izquierda) el parámetro queda perfectamente identificado con un valor MAP estimado de $\phi_1 = 0.945$,

ciertamente alto. Sin embargo, con GL++NTL (derecha) el valor MAP estimado para el parámetro es $\hat{\phi}_1 \approx 0.99$. En la introducción de este capítulo (sección 4.1) ya se referenció el trabajo de Scharnagl et al. (2015), en el cual se reporta el problema de haber obtenido un valor MAP similar para el parámetro $\hat{\phi}_1$.

Igualmente, en el trabajo de Evin et al. (2014) con el modelo de error WLS-AR1 (que es el mismo modelo de Evin et al. (2013)), dos de las doce cuencas analizadas presentan el mismo problema. Más casos con similares y nuevos efectos espurios, se producen en la inferencia con el modelo hidrológico GR4J.

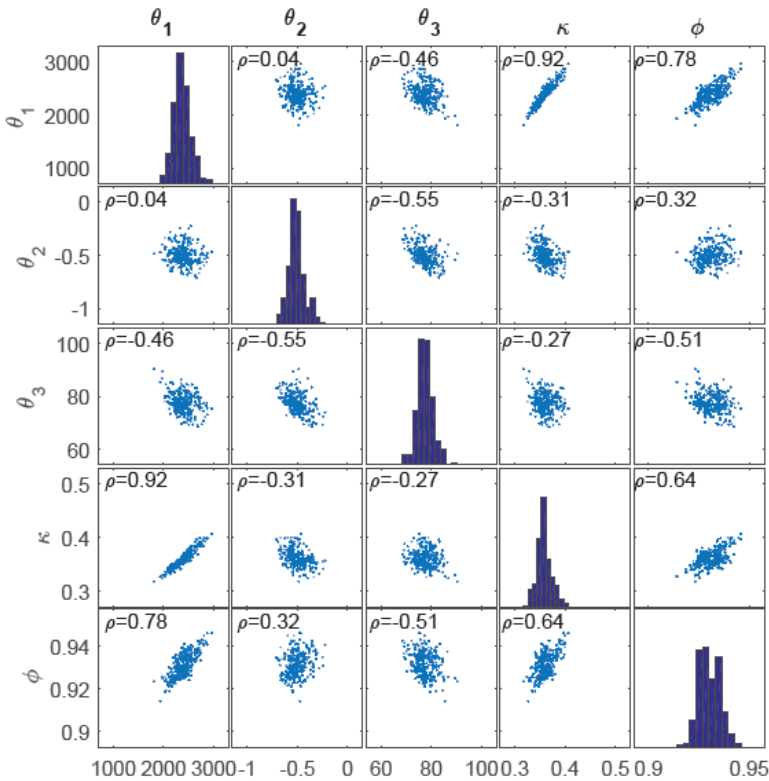


Figura 5.8. Matriz de correlación para los parámetros del modelo GR4J inferidos con el modelo de error GL++

Obsérvese la Figura 5.8 que muestra de forma gráfica la matriz de correlación de los parámetros, así como sus posteriors, para GL++. Los principales aspectos a considerar en la figura son:

- 1- No hay correlación significativa entre los parámetros hidrológicos
- 2- Las correlaciones más altas de la matriz se producen entre los parámetros $\{\theta_1, \kappa\}$ y $\{\theta_1, \phi_1\}$. Sin embargo, todos los parámetros son perfectamente identificables.
- 3- La correlación $\{\kappa, \phi_1\}$ presenta un valor moderado de $C=0.64$

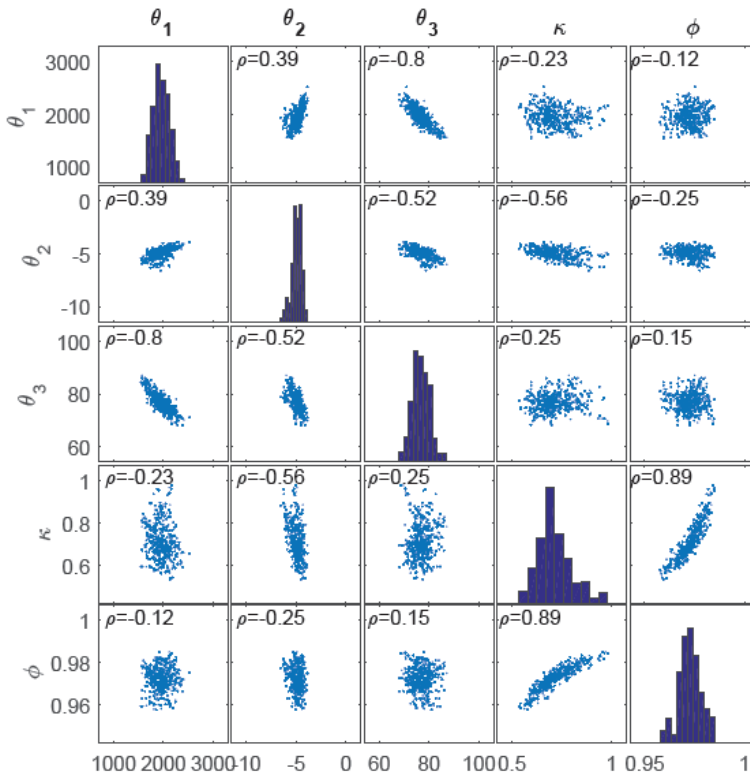


Figura 5.9. Matriz de correlación para los parámetros del modelo GR4J inferidos con el modelo de error GL++NTL

Obsérvese ahora la Figura 5.9. Esta figura muestra la matriz de correlación de los parámetros y sus posteriors, para el modelo de error GL++NTL. Los principales aspectos a considerar en la figura son:

- 1- El modelo de error GL++NTL ha provocado un incremento muy significativo en las correlaciones entre los parámetros hidrológicos: $C\{\theta_1, \theta_2\}$ se incrementa de 0.04 a 0.39 y $C\{\theta_1, \theta_3\}$ se incrementa de 0.46 a 0.80. Las altas correlaciones entre los parámetros de un modelo generalmente indican problemas estructurales en el modelo. Sin embargo, con GL++ no aparecen esas correlaciones, con lo que la explicación hay que buscarla en otro sitio. Una posible explicación se da en el trabajo de Jakeman y Hornberger (1993). Según estos autores, los incrementos en la matriz de covarianza de los parámetros son un indicador de que los parámetros están capturando menos información de los datos. Desde este punto de vista, podría deducirse que la inferencia GL++NTL es menos eficaz que la inferencia GL++, puesto que con ella los parámetros están capturando menos información.
- 2- La correlación $\{\kappa, \phi_1\}$ presenta el valor más alto de la matriz $C=0.89$, como ocurría con el modelo hidrológico CRR. Sin embargo, con GR4J el parámetro de autocorrelación ϕ_1 sí resulta identificable.
- 3- Con GL++NTL, la incertidumbre del parámetro κ ha aumentado considerablemente ($CV=0.04$ a $CV=0.12$). De igual manera el valor MAP prácticamente se duplica (ver Tabla 5.3).

Predicción y estimación de su incertidumbre

Las diferencias entre GL++ y GL++NTL, es decir, entre aplicar o no aplicar las TLs al modelo de varianza, no terminan con lo que se ha mostrado hasta ahora. Falta analizar el efecto de la inflación de las bandas de incertidumbre de la predicción, cuya causa podría estar también en la incoherencia estadística que se está cometiendo al no considerar las TLs.

En la Figura 5.10 se muestran la predicción media y su banda de incertidumbre (95%), para el modelo hidrológico CRR, cuando se infiere con el modelo de error GL++ (panel superior) y con el modelo GL++NTL (panel inferior). Con ambos modelos de error se observa la importante sobrepredicción (sesgo), que ya fue

mencionada al comienzo de esta sección: en ambas gráficas se ve como la predicción media (línea verde) se sitúa frecuentemente sobre las observaciones (puntos rojos). No obstante, lo más llamativo de ambas gráficas es la importante diferencia en el aspecto de la banda de incertidumbre (gris). La comparación de ambas gráficas pone de manifiesto cómo con el modelo de error GL++NTL se produce la ya referenciada por otros trabajos como inflación “sin sentido” de la banda de incertidumbre. La inflación es debida a la errónea y estadísticamente incoherente estimación del modelo de varianza.

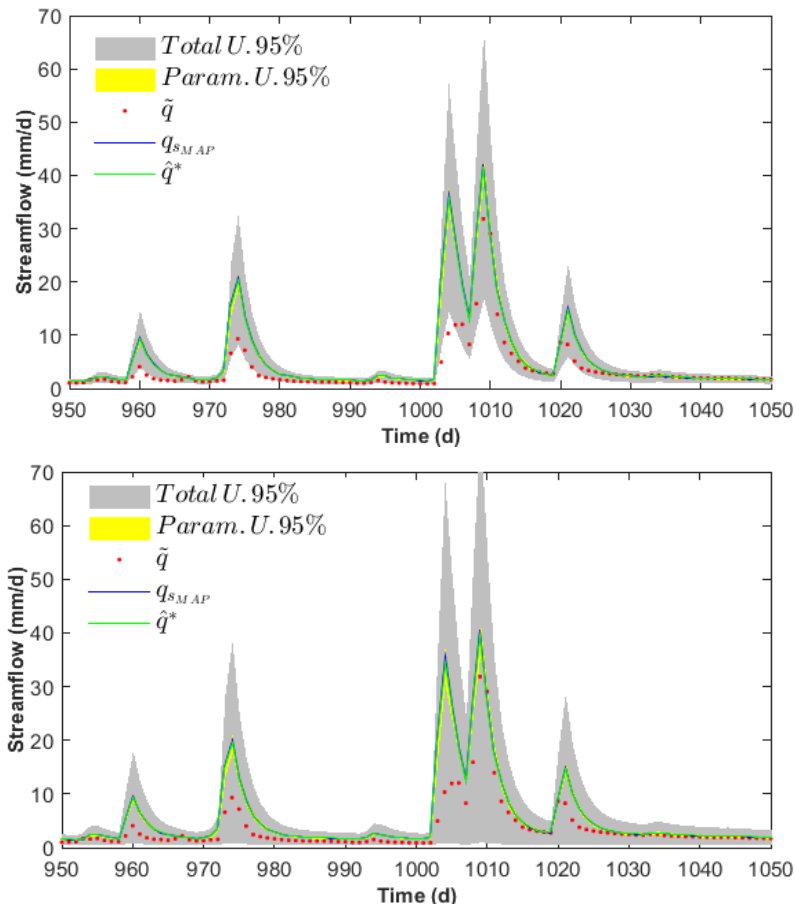


Figura 5.10. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada para el modelo hidrológico CRR, con los modelos de error GL++ (panel superior) y GL++NTL (panel inferior). Como se puede apreciar la contribución de los parámetros es mínima, aunque superior a la obtenida con SLS y WLS

Con respecto a la contribución de la incertidumbre de los parámetros a la incertidumbre total, con GL++ se aprecia una mayor contribución que con SLS y WLS, posiblemente debida al propio aumento en la incertidumbre de los parámetros.

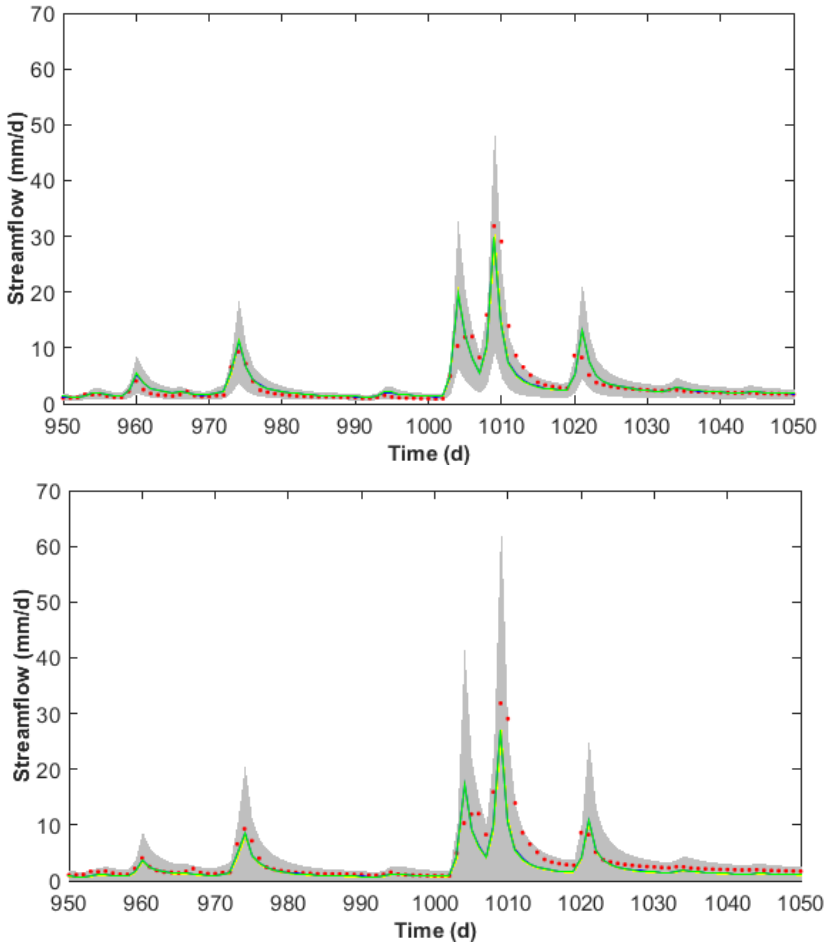


Figura 5.11. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada para el modelo hidrológico GR4J, con los modelos de error GL++ (panel superior) y GL++NTL (panel inferior). Como se puede apreciar, la contribución de los parámetros es mínima este caso

Obsérvese la Figura 5.11 para el modelo hidrológico GR4J. En dicha figura se puede observar que la banda de incertidumbre obtenida con GL++ (panel

superior) es más estrecha que la correspondiente en el modelo CRR (Figura 5.10, panel superior), es decir, las predicciones con CRR son más inciertas que con GR4J. Respecto a la inferencia de GR4J con GL++NTL (Figura 5.11, panel inferior), nada cambia respecto a lo que ocurría con CRR, puesto que se genera una banda de incertidumbre con el mismo problema de inflación.

Mediante los PP-plots de la Figura 5.3 se puede comprobar cómo la evaluación de la incertidumbre obtenida por GL++ es correcta (no sobreestima, ni infraestima la incertidumbre) para ambos modelos hidrológicos. En la figura también se observa como los PP-plots correspondientes a GL++NTL muestran el empeoramiento, para ambos modelos hidrológicos, de la estimación de la incertidumbre de la predicción media. Esta sobreestimación de la incertidumbre está en concordancia con la inflación de las bandas observada. Llama fuertemente la atención el PP-plot correspondiente a GL++NTL para el modelo GR4J, el cual muestra una fuerte infrapredicción del modelo hidrológico, mientras que con GL++ muestra, contrariamente, sobrepredicción. Este comportamiento no se observa con el modelo hidrológico CRR.

5.3.3. GL++ y GL++NTL. Conclusiones

El modelo de error GL++ mejora considerablemente a los modelos SLS y WLS, debido a las relajaciones añadidas (respecto a WLS) de las hipótesis de independencia de los errores y Normalidad de su distribución. El análisis a posteriori de los errores muestra un cumplimiento aceptable de la mayoría de las hipótesis. No obstante, no se cumple la hipótesis de media nula en los errores y es mejorable la estabilización de la varianza cuya eficacia se ve comprometida, precisamente, como consecuencia de la hipótesis de media nula en los errores.

GL++ parece ser una buena herramienta de diagnóstico (general), mostrando que el modelo hidrológico tiene defectos, aunque no se especifica cuáles son. En este sentido, el modelo hidrológico CRR parece proporcionar una predicción media con más sesgo y más incertidumbre que GR4J. Ambos modelos presentan un sesgo de sobrepredicción inducido por el modelo de error GL++. La estimación de la incertidumbre inferida por GL++, es correcta para ambos modelos hidrológicos.

Los resultados han mostrado que, como se sospechaba, la inclusión en el modelo de error, de un modelo de varianza de los errores, requiere tomar en consideración las TLs. Ignorar este requerimiento no generaba problemas en los resultados con el modelo de error WLS. Sin embargo, con un modelo de error más sofisticado y flexible (con más grados de libertad) como GL++, sí que aparecen problemas. Los problemas mostrados con GL++NTL se resumen en: i) Muy altas correlaciones entre parámetros y correlación espuria entre parámetros hidrológicos; ii) Aparición de parámetros no identificables; iii) Inflación de las bandas de incertidumbre (estimación errónea de la incertidumbre) y iv) Estimación errónea de la predicción media.

CAPÍTULO 6. APLICACIÓN DEL NUEVO MODELO DE ERROR COMPLETO: GL++BIAS

Life is uncertain. Eat dessert first

-Ernestine Ulmer (1892-1987)-

6.1. Aplicación en modelación agregada

En el capítulo 5, los resultados con el modelo de error GL++ muestran que la inferencia con él realizada introduce un sesgo en la distribución predictiva, para ambos modelos hidrológicos. Por tanto, GL++ no cumple con su hipótesis de modelar errores con media nula. Con el modelo de error GL++Bias se relaja también esta hipótesis, la última de las condiciones del SLS que quedaba por liberar. Para ello se emplea el modelo de sesgo de tres parámetros definido por la ecuación (3.12), el cual permite modelar un comportamiento de los errores diferente entre caudales bajos y altos.

La Figura 6.1 representa, para ambos modelos hidrológicos CRR (panel superior) y GR4J (panel inferior), las nubes de errores inferidas sobre las que se ha superpuesto la representación de los modelos inferidos de varianza (líneas negras) y de sesgo (líneas discontinuas azules). Nótese como el modelo de sesgo resultante con cada modelo hidrológico, presenta una importante diferencia, resaltada en las pequeñas ventanas de ampliación de la gráfica principal. Por una parte, con el modelo CRR se distingue claramente el diferente comportamiento del sesgo entre caudales bajos y altos, cada uno representado por su diferente pendiente. El caudal umbral inferido, que separa ambos comportamientos, es $y_{\mathcal{M}}^* = 1.96$ mm/d. Por otra parte, para el modelo GR4J la inferencia con GL++Bias no muestra que haya distinción entre el sesgo con caudales bajos y el sesgo con caudales altos. Esta conclusión se deriva del hecho de que el caudal umbral inferido $y_{\mathcal{M}}^* = 1.57$ mm/d está prácticamente en el límite inferior del rango de valores del caudal simulado. En cualquier caso, el modelo de sesgo muestra un buen desempeño, puesto que sus leyes representadas (líneas discontinuas azules) se ubican en la posición central de la nube de errores, indicando dónde se encuentran las medias de las distribuciones condicionales de error.

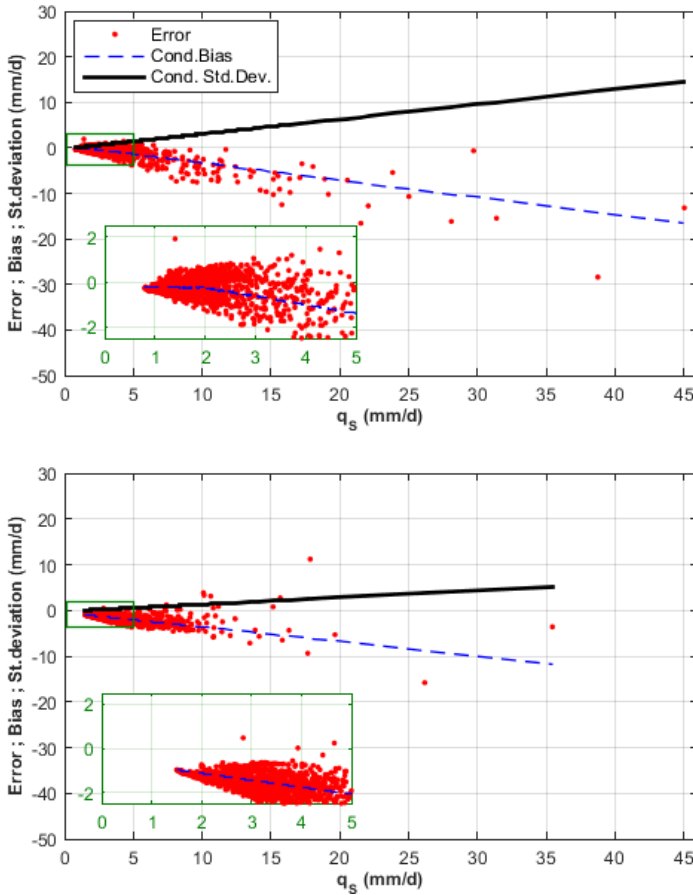


Figura 6.1. Elementos del modelo de error GL++Bias, para los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior). Las gráficas muestran las nubes de error en rojo. Superpuestas aparecen las líneas inferidas que representan a los modelos de varianza (línea negra) y sesgo (línea azul discontinua). Las ventanas pequeñas muestran el comportamiento del modelo de sesgo para los caudales bajos

6.1.1. Validez del modelo de error: cumplimiento de sus hipótesis

La inferencia con el modelo de error GL++Bias presenta un cumplimiento razonable de sus hipótesis acerca de los errores que está modelando. En cualquier caso, un mejor comportamiento que las inferencias realizadas con los modelos de error previamente considerados en la Tesis. No obstante, todavía

quedan cuestiones mejorables que se comentan a continuación. Respecto a la hipótesis de distribución de las innovaciones, la Figura 6.2 muestra (paneles superiores), para ambos modelos hidrológicos, una densidad de probabilidad observada de las innovaciones con valor cero, superior a la que la pdf SEP puede reproducir. Es decir, la curtosis observada excede a la curtosis teórica inferida. En lo referente a la hipótesis de independencia de los errores, GL++Bias al igual que GL++ eliminan de forma significativa la autocorrelación de los errores, si bien las funciones de autocorrelación (paneles inferiores) muestran una pequeña correlación remanente para los dos primeros lags.

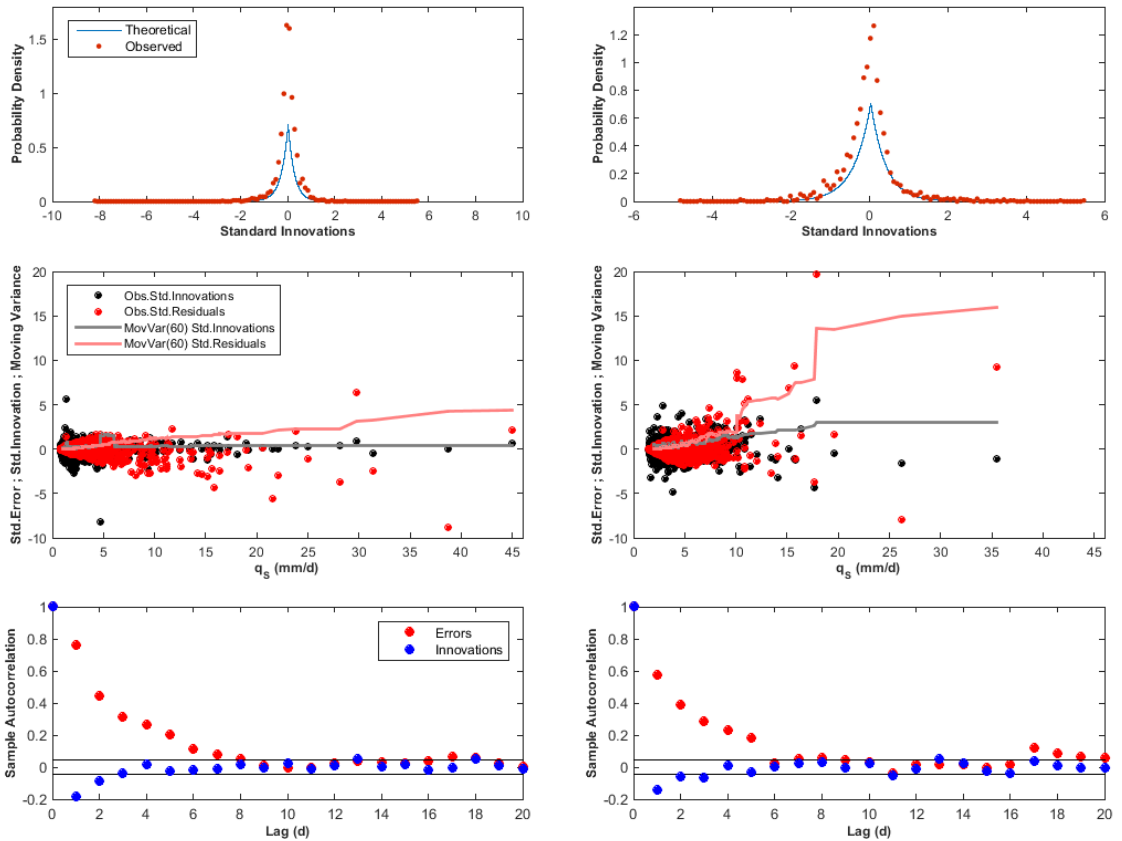


Figura 6.2 Comprobación de las hipótesis del modelo de error GL++Bias, con los modelos hidrológicos CRR (columna izquierda) y GR4J (columna derecha). Ajuste a SEP de los errores estándar observados (paneles superiores). Evaluación de la homocedasticidad (paneles

intermedios). Evaluación de la independencia mediante la función de autocorrelación y los límites (95%) de Anderson (paneles inferiores)

Con relación a la estabilización de la varianza (paneles centrales), se observa que la varianza móvil de las innovaciones (líneas grises) tiende a ser prácticamente constante para todo el rango de caudales simulados (con CRR mayor horizontalidad que con GR4J), es decir, las innovaciones de GL++Bias cumplen aceptablemente la hipótesis de homocedasticidad. Por tanto, se confirma la explicación dada en el capítulo 5, sobre el motivo de la deficiente estabilización de la varianza conseguida con el modelo de error GL++: **cuando se incumple la hipótesis de tener errores con media cero**, como ocurría con el modelo de error GL++, el modelo de varianza no se puede inferir de forma correcta, lo que **impide una correcta estabilización de la varianza** de los errores.

6.1.2. Identificabilidad e incertidumbre de los parámetros con GR4J

En la Figura 6.3 se muestra de forma gráfica, para el modelo hidrológico GR4J, la matriz de correlación de los parámetros, así como sus posteriores, obtenidos con el modelo de error GL++Bias.

Los principales aspectos a considerar en la figura son:

- 1- No hay correlación significativa entre los parámetros hidrológicos, siendo además todos ellos identificables, como ocurría con GL++
- 2- Las correlaciones más altas de la matriz se producen entre el parámetro θ_2 del modelo hidrológico y los parámetros del modelo de error relacionados con la varianza y el sesgo, esto es $C\{\theta_2, \kappa\} = 0.96$, $C\{\theta_2, \gamma\} = 0.95$ y $C\{\theta_2, \tau\} = 0.91$.

La explicación a la alta correlación entre el parámetro θ_2 y los parámetros del modelo de error es, básicamente, que todos ellos están participando en el mismo proceso dentro del problema de modelación que se está infiriendo. No obstante, en la sección 6.1.4 donde se comparan los valores de los parámetros hidrológicos estimados con cada modelo de error, se dará una explicación más detallada de este hecho.

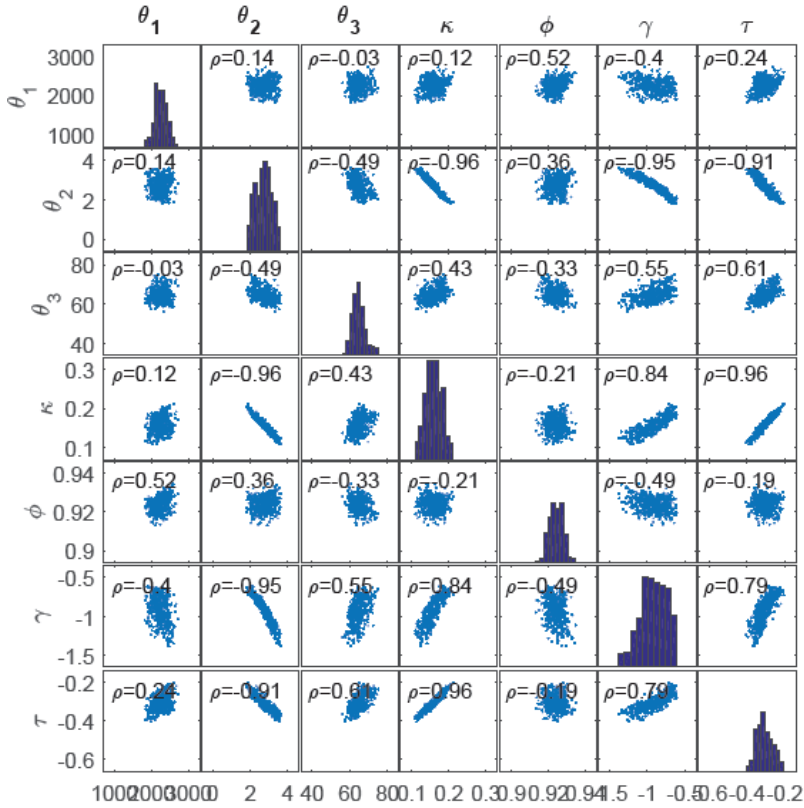


Figura 6.3. Matriz de correlación para los parámetros del modelo GR4J inferidos con el modelo de error GL++Bias

6.1.3. Predicción y estimación de su incertidumbre

El efecto de inferir con GL++Bias es muy diferente para el modelo hidrológico CRR y para el GR4J. Como se explicó en la inferencia con GL++, de los dos modelos hidrológicos analizados, el modelo hidrológico CRR resultaba tener la predicción más incierta y con más sesgo.

La inferencia con GL++Bias es capaz de eliminar prácticamente la totalidad del sesgo que presentaba CRR en la inferencia con GL++. En la Figura 6.4 (panel superior) se puede observar el buen ajuste de la predicción media (línea verde) a

las observaciones. La eliminación del sesgo en CRR puede también ser reconocida en los PP-Plots de la Figura 5.3 en la que se observa, en el panel superior, que la línea correspondiente al modelo de error GL++Bias prácticamente se solapa con el bisector del diagrama, es decir, presenta un resultado excepcional. Por su parte, la Tabla 5.1 muestra como la predicción media del modelo CRR obtiene un aceptable NSE=0.76 (con GL++ el índice era de 0.25) con un error relativo de volumen=0.0%, lo cual va en consonancia con el excelente comportamiento del modelo de sesgo.

El punto negativo del modelo CRR es la gran incertidumbre que presenta su predicción, como muestra la banda (95%) de incertidumbre de la Figura 6.4 (panel superior). Esta banda de incertidumbre tan amplia, no se debe a ningún error espurio (basta comprobar el PP-plot para ver su fiabilidad), como en los casos en los que no se aplican las TLs. El motivo de la amplitud es que la información contenida en los datos no apoya fuertemente a la estructura del modelo CRR y sus parámetros. Esto también se observa en la Tabla 5.1, que muestra un empeoramiento de la precisión (o resolución) entre GL++ y GL++Bias, desde 4.70 a 3.94. En resumen, en base al aceptable cumplimiento por los errores, de las hipótesis de GL++Bias y al correcto PP-plot de su distribución predictiva, se puede afirmar que CRR, con un total de 7 parámetros, presenta una predicción exacta, aunque muy incierta (muy poco precisa). Estas propiedades de los resultados del modelo CRR parecen ser coherentes con las ideas expresadas en la sección 2.1.2 sobre el **dilema sesgo-varianza** de la predicción y el **Principio de Parquedad**, que indican que CRR podría estar sobreparametrizado.

El comportamiento de GL++Bias con el modelo hidrológico GR4J no parece, a priori, tan efectivo. GR4J ve mejorada su predicción media con una reducción en su sesgo, aunque mucho más modesta (Figura 5.3) que en el modelo CRR. La predicción media proporciona un índice NSE=0.80, empeorando ligeramente el obtenido en la inferencia con GL++ (NSE=0.82). El error relativo en volumen (0.0%) mejora respecto al obtenido con GL++. A la vista de la Figura 6.4 (panel inferior), se puede decir que GR4J, aunque con un ligero sesgo en su predicción, muestra una banda de incertidumbre más estilizada, lo que es señal de ofrecer más precisión que CRR. Esto también es confirmado por el valor de precisión (o resolución) de la Tabla 5.1, en la que CRR presenta un valor de 3.94, menor que

el valor 4.40 obtenido por GR4J. En resumen, la inferencia de GR4J con GL++Bias cumple aceptablemente las hipótesis del modelo de error. La predicción media en conjunto permanece similar a la que se obtiene al inferir con GL++.

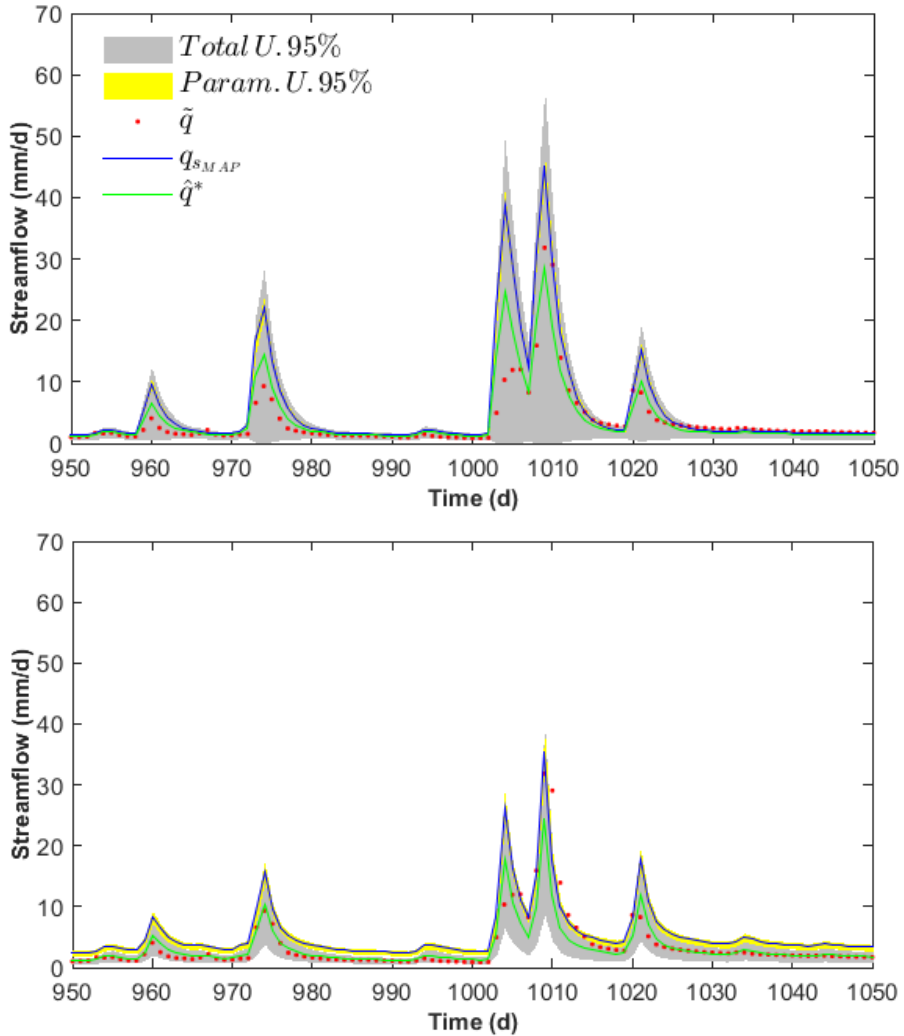


Figura 6.4. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Inferencia realizada con el modelo de error GL++Bias, para los modelos hidrológicos CRR (panel superior) y GR4J (panel inferior). Como se puede apreciar, la contribución de los

parámetros a la incertidumbre total es mínima para el modelo CRR, sin embargo, con GR4J, es la mayor contribución observada en todos los experimentos

En lo relativo a la estimación de la incertidumbre, la Tabla 5.1 muestra que hay una muy leve mejora (de 4.30 a 4.40) en la precisión respecto a la inferencia con GL++. En conjunto, la inferencia de GR4J con GL++Bias muestra que GR4J presenta una predicción ligeramente más exacta y mucho más precisa que CRR, aunque para GR4J la diferencia entre las inferencias con GL++ y GL++Bias es muy leve: esto indica que el modelo de sesgo no es el más adecuado para el modelo GR4J.

Por otra parte, prestando atención al panel inferior de la Figura 6.4, llama la atención la amplitud de la banda (95%) de incertidumbre de la predicción debida a la incertidumbre de los parámetros (en amarillo). Es la mayor amplitud que esta banda muestra en todos los experimentos realizados. Para la explicación de este hecho, es interesante recordar lo explicado en la sección 2.1.3 y la Figura 2.1. Los actores involucrados en esta explicación son, el modelo hidrológico GR4J, su parámetro θ_2 , y el modelo de inferencia (modelo de error) GL++Bias. La incertidumbre del parámetro se evalúa mediante el coeficiente de variación (CV) de su posterior en la Tabla 5.3. En ella se observa que la incertidumbre de θ_2 en su inferencia con GL++Bias es CV=0.16. Significativamente más pequeña que la incertidumbre que tiene en la inferencia con GL++ cuyo valor es CV=0.21. Sin embargo, siendo θ_2 más incierto, propaga escasamente su incertidumbre a la predicción. Compruébese esto con la observación de la Figura 5.11 (panel superior) en la que al inferir con GL++ no se aprecia la banda amarilla debida a los parámetros. Contrariamente, la Figura 6.4 (panel inferior) muestra la amplia banda amarilla de la inferencia con GL++Bias, debida a un θ_2 que, sin embargo, tiene menos incertidumbre.

En correspondencia con el diagrama inferior de la Figura 2.1, la situación se está produciendo debido a la alta sensibilidad que GR4J presenta a θ_2 , cuando la inferencia se realiza con GL++Bias. Recuérdese, que la alta sensibilidad de un modelo respecto a un parámetro es un indicador de que los datos apoyan fuertemente el valor óptimo (MAP) de ese parámetro, por lo que pequeñas variaciones del mismo, afectan considerablemente a la respuesta del modelo. En

conclusión, el parámetro θ_2 está capturando más información de los datos con GL++Bias que con GL++, por lo que con el modelo de error GL++Bias θ_2 se infiere de modo más eficiente.

6.1.4. Sobre la estimación y el sesgo de los parámetros

Se ha comentado en secciones previas, el hecho de que la inferencia con un modelo de error (o función de verosimilitud) incorrecto puede generar “buenas” predicciones medias en calibración, pero una estimación de parámetros sesgada, así como una incertidumbre de la predicción incorrecta. El paradigma de esta idea está en los modelos de error SLS y WLS, cuyo principal objetivo es el mejor ajuste a las observaciones sin atender a otras consideraciones, llámense fuentes de error.

En el lado opuesto, una consideración, si no correcta, al menos apropiada de los errores generados en la inferencia, produce una estimación de parámetros menos sesgada, así como una predicción hidrológica realista, dadas las herramientas y los datos con los que se realiza. Ese “realismo” supone que, si la inferencia se realiza con datos erróneos y/o modelos poco adecuados al problema, los resultados predictivos van a ser deficientes, en el sentido de un pobre ajuste a las observaciones, como no puede ser de otra manera. Sin embargo, esa deficiencia en los resultados de la predicción media se va a ver compensada por una estimación realista de la predicción, una estimación correcta de la incertidumbre de la predicción y una estimación correcta de los parámetros y su incertidumbre.

Los experimentos realizados cuentan, al menos, con el problema estructural en los dos modelos hidrológicos, de no considerar los procesos de nieve y fusión. Además, puede haber otros problemas no identificados en los modelos o en los datos. Estos problemas han de generar, necesariamente, errores que se apartan de las hipótesis de Normalidad, media nula, independencia y homocedasticidad. En las secciones anteriores se ha mostrado como cambia el resultado de la inferencia en lo relativo a la predicción media y su incertidumbre, cuando se infiere con diferentes modelos de error, desde el clásico SLS hasta el nuevo modelo propuesto GL++Bias. En esta sección se va a mostrar la otra cara de la

moneda: ¿cómo cambia la estimación de parámetros al cambiar el modelo de error? Para ilustrar el efecto sobre los parámetros se ha escogido el modelo de error GR4J por su menor número de parámetros, todos ellos identificables en todos los experimentos realizados.

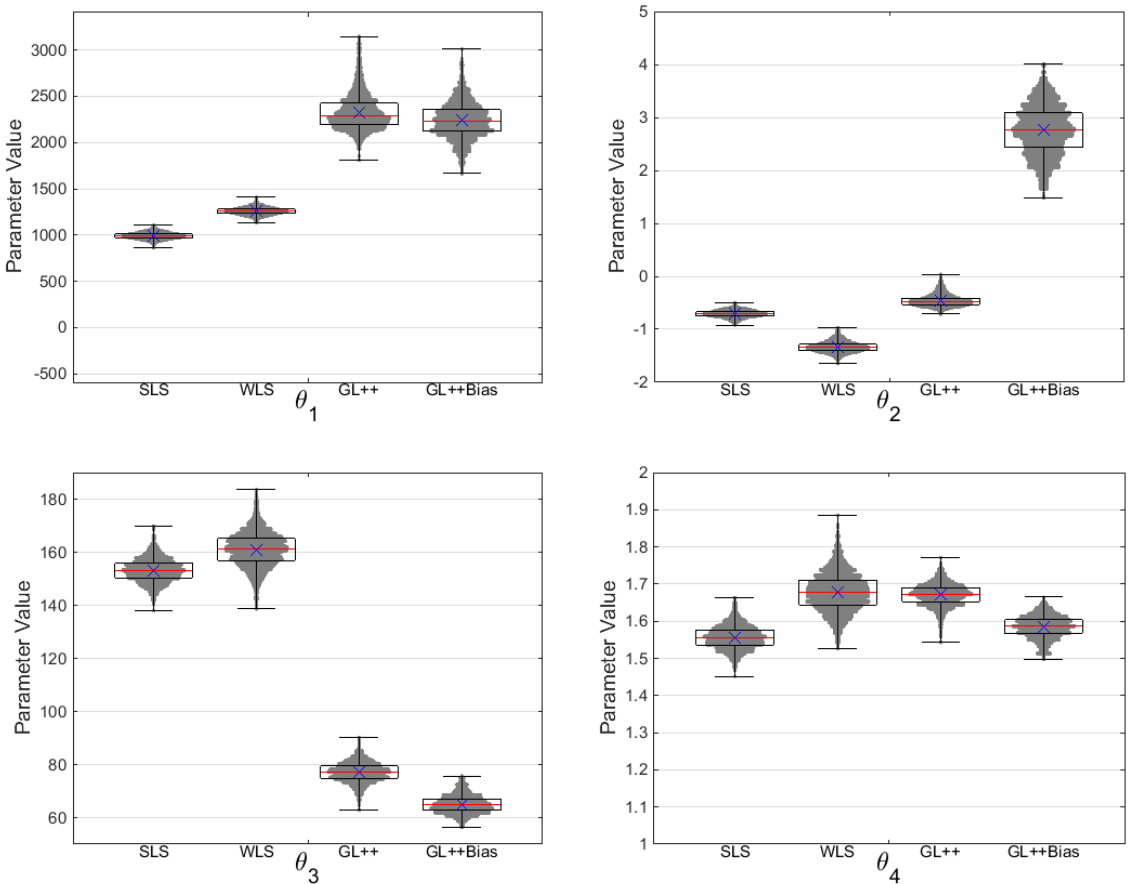


Figura 6.5. Cambios en la distribución posterior (box-plots) de los 4 parámetros del modelo GR4J, según el modelo de error utilizado en la inferencia. Se ilustran los experimentos con SLS, WLS, GL++ y GL++Bias

La Figura 6.5 muestra las marginales posteriores de los 4 parámetros del modelo GR4J, representando cada parámetro en un panel diferente. Dentro de cada uno

de los 4 paneles, se muestran los cambios del parámetro al ser inferido con cada uno de los 4 modelos de error considerados: SLS, WLS, GL++ y GL++Bias.

Una rápida vista a la Figura 6.5 evidencia que las diferencias en los valores MAP de los parámetros no son despreciables, al menos, para tres de los cuatro parámetros de GR4J $\{\theta_1, \theta_2, \theta_3, \}$. También se pueden apreciar diferencias importantes en la incertidumbre de su estimación (altura del box-plot), especialmente en los parámetros $\{\theta_1, \theta_2, \}$, los cuales presentan la tendencia de aumentar su incertidumbre al realizar la inferencia con un modelo de error más sofisticado, según se vio en la explicación de resultados de cada modelo de error.

Sin embargo, el caso más paradigmático que muestra la necesidad de inferir con modelos de error apropiados se produce con el parámetro θ_2 . Con este parámetro se puede deducir lo bien o mal que lo está haciendo la inferencia, debido a que conocemos con relativa certeza el valor que el parámetro debería tener. Por esto, el ejemplo tiene gran valor didáctico. El parámetro θ_2 es el coeficiente de intercambio de agua subterránea. Su misión en GR4J es regular el trasvase de agua entre la cuenca y un acuífero profundo (regional). De manera que, valores negativos del parámetro indican que la cuenca exporta agua hacia el acuífero profundo. Por el contrario, valores positivos del parámetro están indicando una importación de agua desde el acuífero profundo (agua cuyo origen puede trascender los límites de la cuenca superficial). Por tanto, θ_2 es un parámetro excepcionalmente importante e influyente en el modelo GR4J, puesto que él solo es capaz de cerrar el balance de la cuenca.

Si se observa la Figura 6.5 (panel superior derecho), muestra importantes cambios en la inferencia de θ_2 según el modelo de error utilizado. Incluso un cambio en su signo. Estos cambios tienen profundas consecuencias en la conceptualización del modelo de la cuenca analizada. En base al conocimiento adquirido sobre la geología de la cuenca FB, esta cuenca no presenta ningún acuífero profundo con el cual intercambiar agua. Toda su estructura hidrogeológica queda delimitada por el mismo contorno de la cuenca superficial y su capacidad de almacenamiento es reducida, considerándose por tanto que su

respuesta es bastante rápida. Siendo esto así, el valor conceptualmente válido en esta cuenca para el parámetro θ_2 debería ser $\theta_2 = 0$. Por tanto, el balance de agua de la cuenca debería ser cerrado sin pérdidas y sin contribuciones diferentes a la evapotranspiración real y la precipitación, respectivamente. Si la inferencia con un modelo de error proporciona valores muy diferentes a $\theta_2 = 0$, ello se debe a deficiencias en el modelo y/o en los inputs, que el parámetro θ_2 trata de compensar y como se ha explicado antes, puede hacerlo al tener el control sobre el balance de la cuenca.

Los valores inferidos en los experimentos son: $\theta_2 = -0.71$ con SLS; $\theta_2 = -1.37$ con WLS; $\theta_2 = -0.55$ con GL++; $\theta_2 = 2.89$ con GL++Bias. Como curiosidad, $\theta_2 = -4.92$ con GL++NTL, el más negativo de todos los valores del parámetro inferidos. Cuanto más correcto es el modelo de error, atendiendo al cumplimiento de sus hipótesis, menos negativo es el valor inferido para θ_2 . Como ejemplo, GL++NTL es incluso peor que WLS o SLS. Otro aspecto interesante a advertir es que todos los valores de θ_2 inferidos, excepto el inferido con GL++Bias, necesitan drenar la cuenca (exportar agua) si bien el que la drena con menor intensidad es el inferido con GL++.

Si se toma por buena la hipótesis de que el valor correcto de θ_2 debería ser el cero: i) De las inferencias realizadas, la más correcta es la realizada con el modelo de error GL++, puesto que infiere el θ_2 más próximo a cero; ii) La inferencia realizada con el modelo GL++Bias es la única que obtiene un valor de θ_2 positivo. Por tanto un modelo de sesgo permite mover θ_2 hacia valores más positivos, desde el valor adoptado con GL++; iii) El modelo GL++Bias utilizado no es correcto en su funcionamiento con GR4J (lo cual ya se dedujo anteriormente en este capítulo) obteniendo un valor excesivamente alto para θ_2 y iv) Un modelo de sesgo más adecuado permitiría, en teoría, la correcta inferencia del parámetro con un valor $\theta_2 \approx 0$.

¿Cómo debería ser ese modelo de sesgo adecuado? Baste fijarse en las inferencias SLS, WLS y GL++. Todas estiman θ_2 con el propósito de que drene la cuenca. Por tanto, el modelo de sesgo debe extraer agua de la cuenca, liberando a θ_2 de esa función que, en este caso, no le corresponde. El modelo de inferencia GL++ intenta no asignar esa función a θ_2 reduciendo de hecho su valor, a costa de obtener una predicción muy sesgada que sobrepasa (por exceso de agua no drenada del sistema) a las observaciones. Los modelos de error SLS y WLS, sin embargo, subsanan el problema del drenaje forzando el valor de todos los parámetros.

¿Por qué el modelo de sesgo empleado no es correcto? Si el modelo fuera el correcto, debería haber estimado $\theta_2 \approx 0$. Sin embargo, la inferencia está estimando un valor positivo y alejado de 0, que indica que la cuenca necesita traer agua. El motivo, es que el modelo de sesgo de GL++Bias está cumpliendo la función prevista de extraer agua de la cuenca, aunque lo está haciendo en exceso. Tanto, que se está forzando a θ_2 para traer el agua que el modelo de sesgo se está llevando. Durante la inferencia, los parámetros del modelo de sesgo y los del modelo de varianza (forzosamente relacionados por la TVL) están ocupándose de extraer agua, mientras θ_2 se ocupa de traerla. Al ocuparse todos estos parámetros del mismo proceso, cerrar el balance de agua en la cuenca, deben estar altamente correlacionados. Esta es la explicación, que fue pospuesta con anterioridad, de por qué en la Figura 6.3 hay correlaciones muy altas entre los parámetros mencionados.

6.1.5. Comparación de los experimentos con modelos agregados

En la Figura 6.6 se representa un diagrama Fiabilidad-Resolución, muy útil para comparar las distribuciones predictivas inferidas con diferentes modelos de error. Los límites entre zonas de clasificación son subjetivos, basados en la experiencia y su única utilidad es la de tener una guía de referencia. Las líneas del diagrama son trayectorias orientadas desde el modelo clásico más simple que sirve de referencia, el SLS, hasta los modelos más sofisticados empleados en los experimentos. Con fines de comparación, también se representan los

modelos de error en los que no se han aplicado las Leyes Totales (modelos NTL) cuyas líneas de trayectoria se han representado en color rojo.

Los aspectos más importantes del diagrama se exponen en los siguientes puntos.

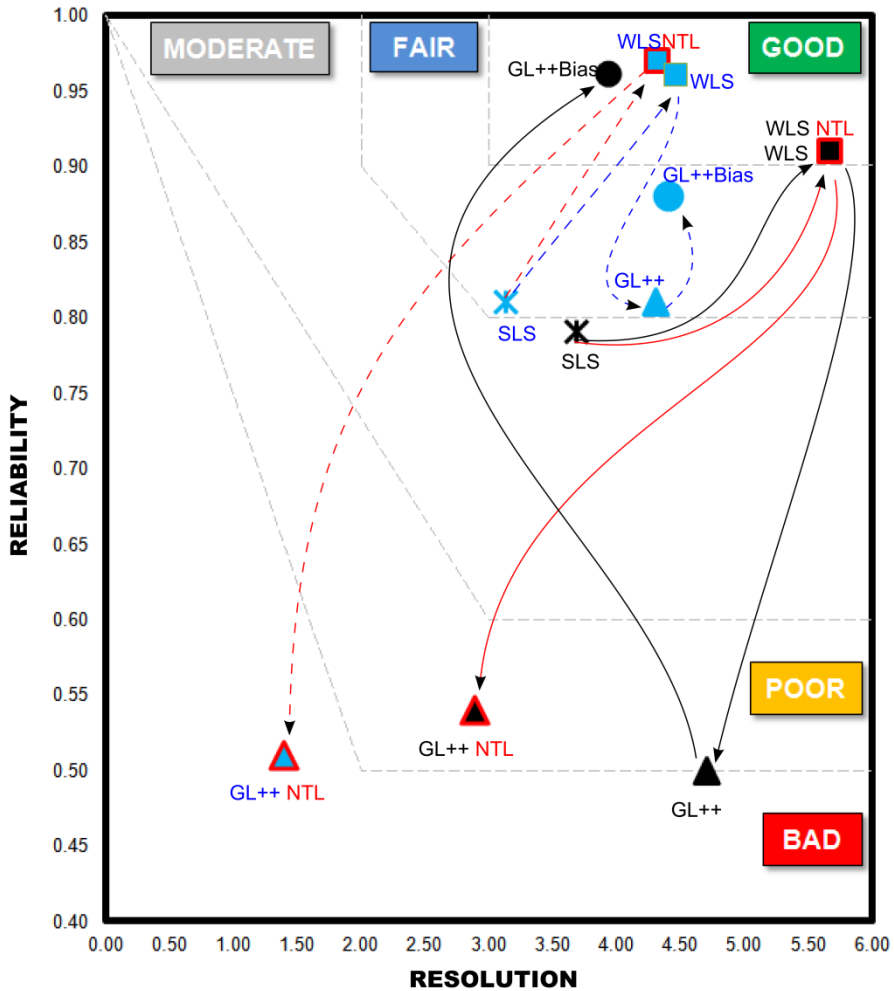


Figura 6.6. Diagrama Fiabilidad-Resolución para la comparación de las distribuciones predictivas inferidas con los diferentes modelos de error. Para el modelo hidrológico CRR se utilizan trayectorias con línea negra continua. Para el modelo GR4J se utilizan líneas azules discontinuas. Las trayectorias a los modelos de error que no contemplan Leyes Totales se indican con líneas rojas, continuas para CRR y discontinuas para GR4J

- 1- Los modelos SLS y WLS se muestran en el diagrama a modo de referencia, puesto que su no cumplimiento de las hipótesis en que se basan, los descarta como modelos de error válidos. El diagrama demuestra que, a pesar de ser modelos inválidos para los experimentos realizados, obtienen una distribución predictiva con propiedades que, al menos en este caso, resultan muy aceptables. Incluso las inferencias con WLS se ubican en lo más alto del diagrama. No obstante, esa circunstancia es tan engañosa como los altos índices de Nash que proporcionan estos mismos modelos de error. Si el modelo de error inferido no cumple con sus hipótesis, lo que de él se obtenga (e.g. la distribución predictiva) no podrá ser correcto. Por otra parte, puede observarse como las propiedades de la distribución predictiva para WLS y WLS-NTL son las mismas, tal como se esperaba y fue explicado en la sección 5.2.
- 2- Se puede observar que las inferencias con GL++ deterioran las propiedades de la distribución predictiva, debido a la aparición del sesgo en la predicción. Se puede razonar que la aparición de este sesgo en la predicción es motivada por la aplicación del filtro autorregresivo sobre los errores, el cual es capaz de convertir el sesgo que hay en los parámetros, en un sesgo en la predicción. Sería razonable pensar que cuanto más distancia haya entre los puntos representativos de WLS (o SLS) y los de GL++, más sesgados están los parámetros inferidos con WLS (o SLS) y más sesgo tendrá la predicción de la inferencia con GL++. Si se sigue esta idea, se puede advertir que la distancia entre WLS y GL++ es más pequeña para el modelo GR4J que para el CRR, de lo cual se deduce que GR4J presenta menos problemas estructurales y es mejor modelo hidrológico que CRR, al menos para el caso de estudio analizado.
- 3- Como se puede observar, las inferencias sin Leyes Totales obtienen una distribución predictiva cuyas propiedades son incluso peores, en fiabilidad y resolución (precisión), que las obtenidas con el modelo de error SLS.
- 4- La adición del modelo de sesgo a las inferencias GL++, es decir las inferencias con GL++Bias, corrigen en mayor o menor medida el empeoramiento de la distribución predictiva producida al inferir con GL++. Sin embargo, como se ha podido observar, un mismo modelo de error no es igualmente efectivo con todos los modelos hidrológicos. El modelo de error, con cada uno de sus componentes, debe ser capaz de reproducir la estructura de error generada, dadas las observaciones y el modelo hidrológico. ¿Por qué el modelo de error GL++Bias obtiene resultados tan diferentes al inferir CRR o GR4J? En lo que sigue se dará una explicación.

En el buen o mal desempeño de un modelo de error al modelar los errores van a influir, lógicamente, todos los componentes que dan forma al modelo de error. Dos de los más importantes son el modelo de varianza y el modelo de sesgo. ¿Cómo saber si estos componentes del modelo son apropiados? El primer test que pasan es la comprobación de las hipótesis en las que se basa el modelo. Sin embargo, esta comprobación puede no ser del todo determinante o clarificadora, para saber cómo de buenos son los modelos de sesgo y varianza. Es posible obtener más pistas al respecto mediante la evaluación de la estandarización de los errores, realizada mediante la ecuación (3.15). Si los modelos de sesgo y varianza son los correctos, los errores estandarizados η deben tener $\mathbb{E}[\eta]=0$ y $\mathbb{V}[\eta]=1$. La obtención de valores significativamente diferentes indica que se está escalando o trasladando la pdf de los errores \mathbf{E} , pero no realizando su correcta estandarización. Esta forma de evaluar la adecuación de los modelos de sesgo y varianza es denominada por esta Tesis como comprobación de la *bondad de estandarización de los errores* (GoES, Goodness of Errors Standardization). Las filas quinta y sexta de la Tabla 5.1 recogen la evaluación de $\mathbb{V}[\eta]^{1/2}$ y $\mathbb{E}[\eta]$ para cada una de las inferencias que ha modelado la varianza y/o el sesgo. Los resultados muestran que para el modelo hidrológico CRR, los modelos de sesgo y varianza consiguen una excelente estandarización de los errores en la inferencia con GL++Bias. Para el modelo hidrológico GR4J, la mejor estandarización se consigue en la inferencia con GL++, con lo que se vuelve a incidir en la idea de que el modelo de sesgo no ha funcionado bien con GR4J.

Nótese que, a pesar de la importancia de considerar las Leyes Totales, éstas son condición necesaria pero no suficiente para obtener una inferencia correcta. La aplicación de las TLs sobre modelos de varianza o sesgo (o AR, etc.) no apropiados, no va a proporcionar los buenos resultados que se esperan.

En ciertos trabajos de la literatura (e.g. Evin et al. (2013, 2014) o Scharnagl et al. (2015)), se ha asumido que tras la estandarización de los errores se obtiene en cualquier caso $\mathbb{V}[\eta]=1$, cuando esto no es necesariamente cierto incluso aplicando las TLs y es con seguridad falso cuando estas no se aplican, siendo este último, el caso de los citados trabajos. Como ejemplo, para el modelo GR4J

la inferencia con GL++ proporciona una $\mathbb{V}[\boldsymbol{\eta}]^{1/2} = 0.95$, muy próxima al valor teórico. Cuando se infiere con GL++NTL el resultado $\mathbb{V}[\boldsymbol{\eta}]^{1/2} = 0.20$ es bien diferente.

6.2. Aplicación en modelación distribuida

En el capítulo 5 y en la primera parte de este capítulo 6, se han mostrado y analizado los resultados de las inferencias de los modelos hidrológicos agregados con los modelos de error GL++ y GL++Bias. En esta última parte del capítulo 6 se van a mostrar los resultados de la inferencia del modelo hidrológico distribuido TETIS, con ambos modelos de error. En este caso, se emplea el modelo de error GL++Bias en su versión de 4 parámetros, definido por la ecuación (3.13).

La Figura 6.7 representa para ambos modelos de error GL++ (panel izquierdo) y GL++Bias (panel derecho), las nubes de errores inferidas, sobre las que se ha superpuesto la representación de los modelos inferidos de varianza (líneas negras) y de sesgo (líneas discontinuas azules). Respecto a la inferencia con GL++, se puede observar una ligera asimetría de la nube de errores hacia los valores negativos, que supone un sesgo marginal de los errores de $\mu_{\mathbf{E}} = -0.43$ m³/s, por tanto, GL++ no cumple con su hipótesis de errores con media nula.

Nótese (recuadro pequeño en panel derecho) cómo en el modelo GL++Bias se distingue entre el sesgo de la predicción para caudales bajos y el sesgo para caudales altos. El valor umbral inferido para la magnitud de caudal que separa ambos comportamientos es $y_{\mathcal{M}}^* = 6.01$ m³/s, con un coeficiente de variación moderado de 0.18 (Tabla 6.1). Por otra parte, si se observa la nube de errores obtenida de la inferencia con GL++Bias, su forma parece sugerir un modelo no lineal de sesgo para el tramo alto de caudales. Esta falta de correspondencia entre el modelo teórico y los errores inferidos debe estar afectando (sesgando) la estimación del resto de parámetros.

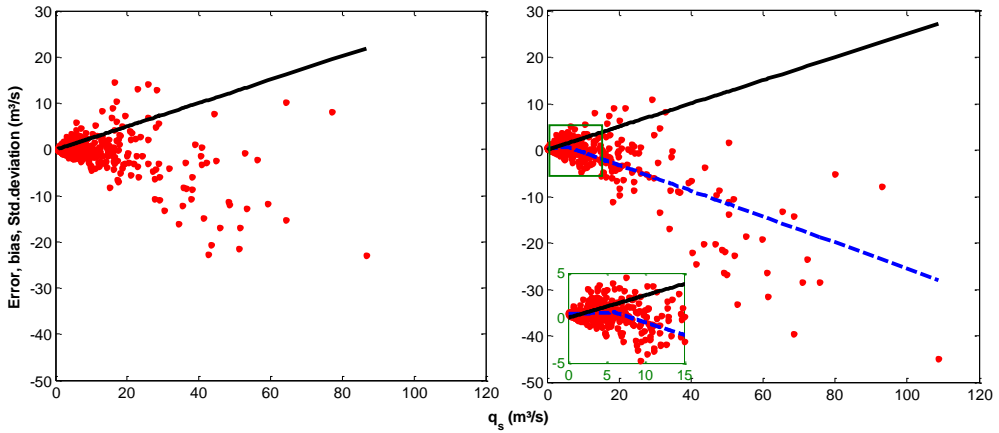


Figura 6.7. Elementos de los modelos de error GL++ (panel izquierdo) y GL++Bias (panel derecho). Las gráficas muestran las nubes de error en rojo. Superpuestas aparecen las líneas inferidas que representan a los modelos de varianza (línea negra) y sesgo (línea azul discontinua). La ventana pequeña en el panel derecho muestra el comportamiento del modelo de sesgo para caudales bajos

6.2.1. Validez del modelo de error: cumplimiento de sus hipótesis

La Figura 6.8 muestra el ajuste de las innovaciones estandarizadas a la pdf teórica SEP. Se puede observar un ajuste excelente con ambos modelos de error, para el periodo de calibración (paneles superiores). En el periodo de validación (paneles inferiores), la bondad del ajuste disminuye, especialmente para la inferencia con GL++, en la que se observan unos errores cuya distribución tiene mucha más curtosis de la que la pdf SEP es capaz de reproducir. Cabe mencionar también, la visible asimetría de las distribuciones, especialmente en la inferencia con GL++. La Tabla 6.1 también muestra esta propiedad en los valores del parámetro ξ de la SEP. La mayor asimetría inferida en GL++, trata de compensar la falta de un modelo de sesgo que desplace convenientemente las distribuciones condicionales de error. En relación con la independencia de los residuos, las funciones de autocorrelación (no mostradas) indican, tanto en calibración como en validación y con ambos modelos de error, la eliminación de la mayor parte de la autocorrelación de los errores, si bien, al igual que ha ocurrido en los experimentos con modelo agregados, permanece una pequeña autocorrelación (inferior a 0.2) para un periodo de desfase (*lag*).

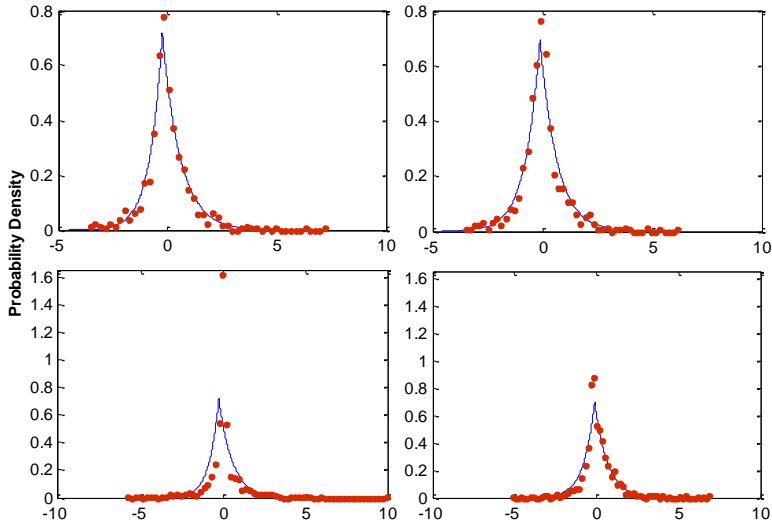


Figura 6.8. Ajuste de las innovaciones estandarizadas (puntos rojos) a la distribución SEP teórica (línea azul). Periodo de calibración en paneles superiores, Periodo de validación en paneles inferiores. Inferencia con GL++ en columna izquierda. Inferencia con GL++Bias en columna derecha

En cuanto a la estabilización de la varianza, la Figura 6.9 muestra la comparación entre los dos modelos de error GL++ y GL++Bias, para el periodo de validación. Las gráficas son similares a las presentadas para los experimentos con los modelos agregados (figuras 5.1, 5.4, 5.5 y 6.2). Ahora, al igual que ocurrió en esos experimentos, la estabilización de la varianza que consigue GL++ (panel superior) es mucho peor que la obtenida por GL++Bias (panel inferior). La razón, como ya se explicó, está en que cuando se incumple la hipótesis de tener errores con media cero (caso del GL++), el modelo de varianza no se puede inferir de forma correcta, lo que impide una correcta estabilización de la varianza de los errores.

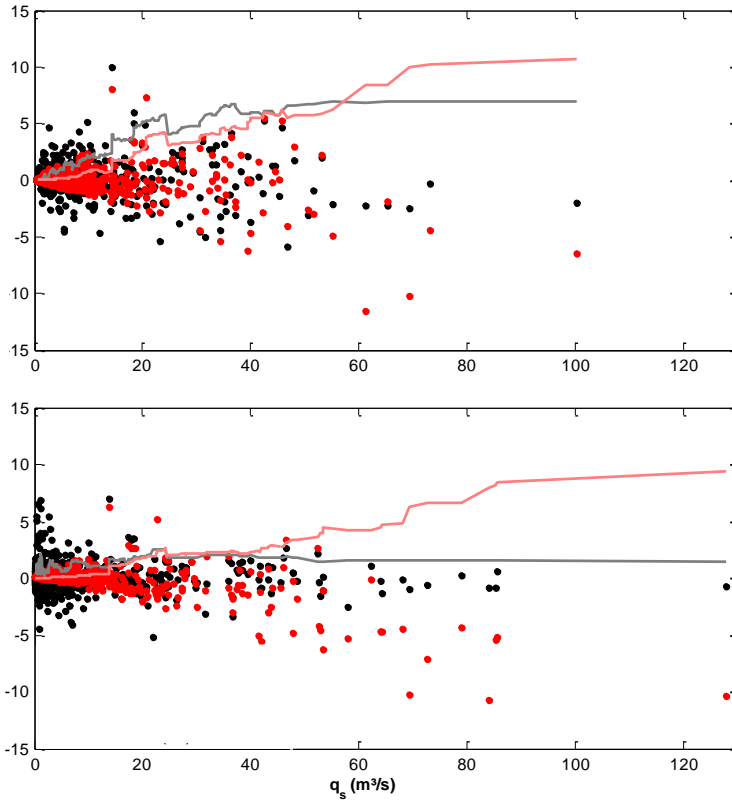


Figura 6.9. Evaluación de la estabilización de la varianza mediante la comparación de la distribución de los errores (puntos rojos) y las innovaciones (puntos negros), en función de la magnitud del caudal simulado durante el periodo de validación. Las líneas representan la varianza móvil (60 datos) de los errores (línea roja) y de las innovaciones (línea gris). Inferencia con GL++ (panel superior) e inferencia con GL++Bias (panel inferior)

6.2.2. Identificabilidad e incertidumbre de los parámetros

Según se muestra en la Tabla 6.1, la inferencia con ambos modelos de error ha podido identificar todos los parámetros del modelo hidrológico, siendo en general y como es esperado, más inciertos los parámetros inferidos con GL++Bias que los inferidos con GL++. La inferencia con GL++ estima como parámetro hidrológico más incierto al F8 (factor corrector del flujo en el acuífero). Por su parte, la inferencia con GL++Bias estima al parámetro F4 (relacionado con la velocidad en ladera) como el parámetro hidrológico más incierto. Por tanto, si

bien el modelo de sesgo permite una mejor estimación del parámetro de la componente subterránea del flujo, no acaba de conseguir una correcta inferencia para el factor relacionado con la respuesta rápida del sistema. Estos resultados están en consonancia con las explicaciones que se dan en el último punto de esta sección, en relación con la Figura 6.13.

Tabla 6.1. Máximo de la posterior (MAP) y coeficientes de variación (CV) de los parámetros. CVs mayores o iguales a 0.2 se marcan en rojo. Parámetros restringidos por las TLs, se somborean en azul

PARAMETERS	GL++		GL++Bias	
	MAP	CV	MAP	CV
F1	1.25	0.07	0.85	0.11
F2	1.11	0.01	1.20	0.02
F3	0.21	0.07	0.26	0.16
F4	0.008	0.17	0.018	0.38
F5	0.07	0.06	0.21	0.11
F6	886	0.08	5145	0.18
F8	52	0.24	1003	0.11
F9	0.10	0.15	0.09	0.15
FP1	0.001490	0.01	0.001480	0.01
α	-0.11	0.07	0.02	0.11
κ	0.25	0.04	0.25	0.02
β	1.00	0.03	0.96	0.03
ξ	1.19	0.05	1.08	0.05
ϕ_1	0.77	0.03	0.67	0.06
y^*	----	----	6.01	0.18
γ	----	----	0.42	0.16
δ	----	----	0.04	1.82
τ	----	----	-0.28	0.38

Por otra parte, en relación con los parámetros de los modelos de error, para GL++ todos los parámetros inferidos presentan muy poca incertidumbre (Tabla 6.1). No ocurre lo mismo en la inferencia con GL++Bias. En esta inferencia, dos

de los parámetros del modelo de sesgo $\{\delta, \tau\}$ presentan una importante incertidumbre, especialmente el parámetro δ . La posible razón de esta pobre inferencia de los parámetros del modelo de sesgo está en la inadecuación de la hipótesis de sesgo lineal, tal y como se comentó al principio de la presente sección 6.2.

Por último, resulta interesante advertir la reducción en el valor estimado del parámetro de autocorrelación ϕ_1 , cuando se realiza la inferencia con el modelo de sesgo. Si bien esta reducción en ϕ_1 era mínima en la inferencia del modelo agregado CRR y nula en la inferencia de GR4J, el hecho de que con un mejor modelo hidrológico (distribuido) sí que se produzca, confirma la idea de que autocorrelación y sesgo son dos expresiones del mismo problema y por tanto, cuanto mejor sea el modelo de sesgo, menor será la autocorrelación remanente en los errores.

6.2.3. Predicción y estimación de su incertidumbre

Para la evaluación de la predicción media obtenida en cada una de las dos inferencias, en los periodos de calibración y validación, se han utilizado los índices mostrados en la Tabla 6.2: el índice NSE, el índice RMSE y el error relativo en volumen $VE(\%) = (\mathbf{V}_s - \tilde{\mathbf{V}}) \tilde{\mathbf{V}}^{-1}$. La predicción obtenida en la inferencia con GL++ resulta peor que la obtenida al inferir con GL++Bias. De hecho, al igual que en los experimentos con los modelos agregados CRR Y GR4J, la inferencia GL++ está actuando como herramienta de diagnóstico del modelo hidrológico, mostrando de forma coherente los resultados que el modelo es capaz de obtener, condicionados a los problemas en la estructura del modelo y/o a los errores en los datos. Es el motivo por el que GL++ muestra un importante sesgo sobrepredictivo, el cual se puede observar en la Tabla 6.2 en forma de un importante error en volumen, en el PP-plot correspondiente de la Figura 6.11 o en el panel superior derecho de la Figura 6.10.

La predicción en calibración obtenida de la inferencia con GL++Bias, muestra una anulación completa del error en volumen. Lo mismo se conseguía en las

inferencias con el modelo GL++Bias, en los experimentos con los modelos agregados.

El empeoramiento de los índices, producido entre los periodos de calibración y validación, es razonablemente pequeño con ambos modelos de error, lo cual es un indicador de robustez en los parámetros inferidos.

Tabla 6.2. Índices de desempeño sobre diferentes aspectos de los resultados en los periodos de calibración y validación, de las inferencias del modelo distribuido TETIS, con los modelos de error GL++ y GL++Bias. La primera fila muestra la Log-verosimilitud de los parámetros inferidos. Los índices de las tres siguientes filas se relacionan con la bondad de la predicción media obtenida. Los índices de las dos últimas filas evalúan la exactitud y precisión de la distribución predictiva respecto de las observaciones.

	CALIBRACIÓN		VALIDACIÓN	
	GL++	GL++Bias	GL++	GL++Bias
Log-L	-456	-417	-1351	-748
NSE*	0.86	0.91	0.83	0.85
RMSE*	3.59	2.92	3.91	3.60
VE* (%)	10.5	0.00	4.30	-3.70
Fiabilidad	0.85	0.93	0.90	0.93
Resolución	4.79	5.02	5.21	5.12

*** Valores para la predicción media**

En la Figura 6.10 se muestran para el periodo de calibración, la predicción media (línea verde), la predicción MAP del modelo (línea azul) y las bandas (95%) de incertidumbre total (gris) y de los parámetros (amarillo). En los paneles superiores se muestra el resultado para la inferencia con GL++, mientras que en los paneles inferiores se muestra el resultado con GL++Bias. En la columna de la derecha se muestra un detalle de las gráficas de la columna izquierda, el cual permite visualizar con claridad los resultados en un tramo con caudales bajos.

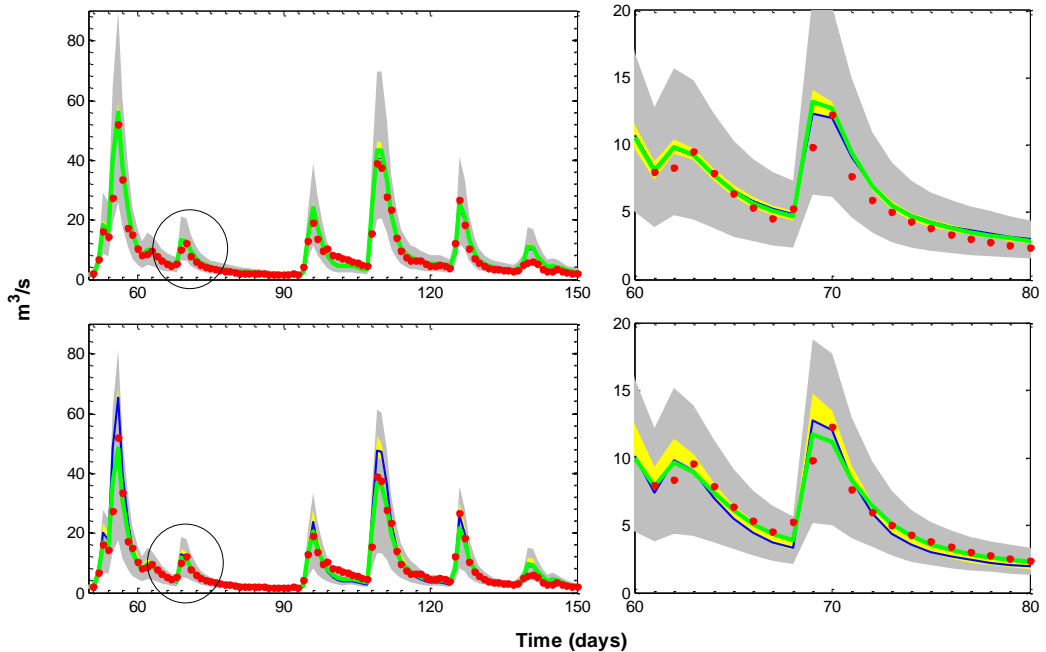


Figura 6.10. Banda (95%) de incertidumbre total (gris) y la correspondiente a los parámetros (amarillo). Superpuestas a la banda de incertidumbre, se muestran las observaciones (puntos rojos), el resultado de la simulación del modelo (línea azul) y la predicción media (línea verde). Los paneles superiores corresponden a la inferencia con GL++. Los paneles inferiores corresponden a la inferencia con GL++Bias. La columna derecha muestra el detalle para un tramo de caudales bajos.

Los principales puntos a destacar en la Figura 6.10 son:

- 1- El aspecto esbelto de la banda de incertidumbre total, con ambos modelos de error, está en concordancia con las altas resoluciones mostradas en la Tabla 6.2, si bien la resolución de la distribución predictiva es mayor en GL++Bias.
- 2- GL++ en la columna derecha, muestra como la predicción media (línea verde) se ubica frecuentemente por encima de las observaciones (puntos rojos), debido al sesgo sobrepredictivo mencionado anteriormente. Este sesgo es la principal causa del bajo valor de fiabilidad mostrado en la Tabla 6.2. Además, se observa la no coincidencia de la predicción media y la predicción MAP del modelo, puesto que esta última corresponde a la

- mediana de la distribución predictiva (cuando no se modela el sesgo). La no coincidencia de media y mediana ocurre en distribuciones asimétricas, como la SEP inferida con GL++, como se comentó previamente.
- 3- En la inferencia con GL++Bias se puede observar el correcto desempeño del modelo de sesgo, prestando atención al panel inferior derecho, en el que se aprecia como la predicción media (línea verde) tiende, en general, a un mejor ajuste a las observaciones (puntos rojos), tratando de corregir la infrapredicción del modelo hidrológico en los caudales bajos ($< 6 \text{ m}^3/\text{s}$) y la sobrepredicción en los altos.
 - 4- Como se ha observado en el resto de experimentos de la Tesis, la contribución de la incertidumbre de los parámetros a la incertidumbre total no es demasiado importante. Se puede observar, no obstante, que la inferencia con el modelo GL++Bias propaga con más efectividad la incertidumbre de los parámetros a la predicción, generando una banda amarilla ligeramente más ancha que con GL++. Este resultado puede ser debido al incremento con GL++Bias, de la incertidumbre de la mayoría de los parámetros hidrológicos y/o a un aumento de la sensibilidad del modelo a algunos de esos parámetros.

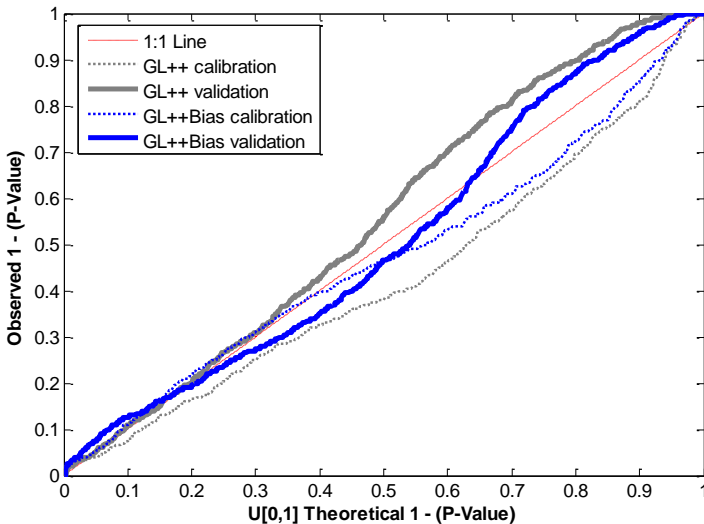


Figura 6.11. Fiabilidad de las distribuciones predictivas inferidas con GL++ (líneas grises) y con GL++Bias (líneas azules). Para el periodo de calibración las líneas son discontinuas, y para el periodo de validación son continuas

La Figura 6.12 es similar a la anterior Figura 6.10, si bien está representando los resultados en el periodo de validación. Con carácter general, se puede decir que la distribución predictiva inferida se está comportando de forma muy aceptable en el periodo de validación. Esta afirmación viene respaldada por los altos índices de fiabilidad y resolución mostrados, con ambos modelos de error, en la Tabla 6.2, así como por los PP-plots de la Figura 6.11. Los elementos observados en la Figura 6.12 no son en general diferentes, a los ya vistos para el periodo de calibración.

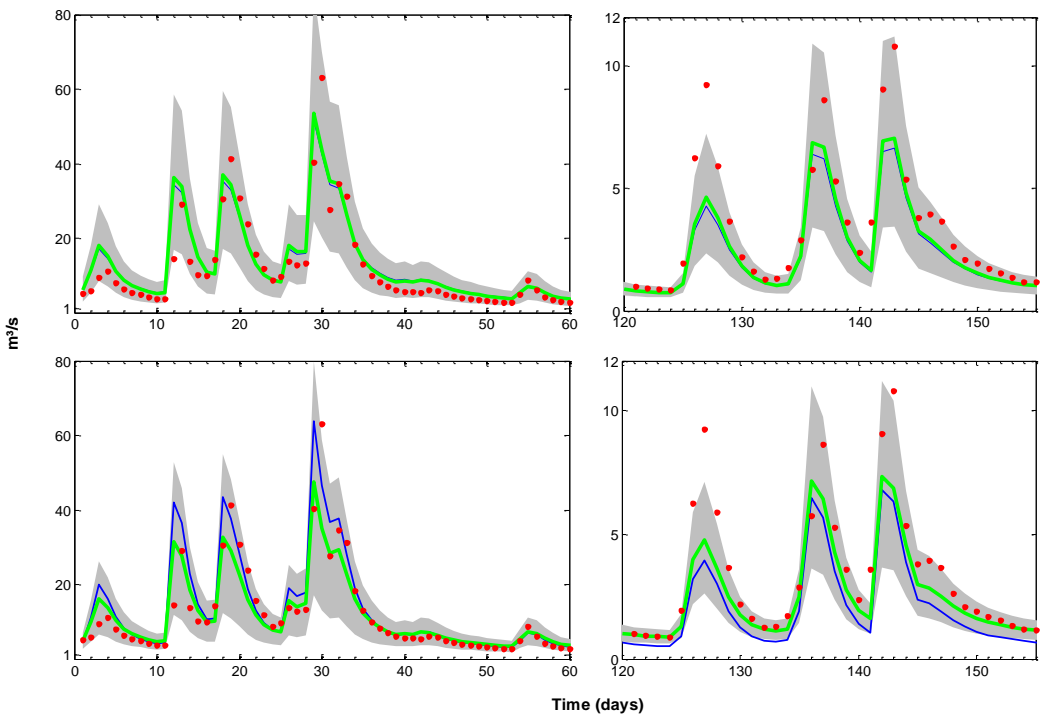


Figura 6.12. Banda (95%) de incertidumbre total (gris). Superpuestas a la banda de incertidumbre, se muestran las observaciones (puntos rojos), el resultado de la simulación del modelo (línea azul) y la predicción media (línea verde). Los paneles superiores corresponden a la validación temporal del modelo inferido con GL++. Los paneles inferiores corresponden a la validación temporal del modelo inferido con GL++Bias.

6.2.4. Influencia del modelo de error empleado en la inferencia sobre los procesos simulados por el modelo

En el transcurso de los capítulos que han mostrado los experimentos realizados en la Tesis, los resultados han mostrado en qué medida la elección del modelo de error con el que se realiza la inferencia influye, tanto en la estimación de los parámetros, como en el cálculo de la predicción y en la estimación de sus respectivas incertidumbres. Sin embargo, en los experimentos realizados con los modelos CRR y GR4J no se analizaron otros elementos de la modelación diferentes a la respuesta total de la cuenca en el punto de control. Con el modelo distribuido TETIS es posible obtener las series temporales de cualquiera de los flujos que suceden en el sistema, en cualquier celda del modelo. Con ese nivel de detalle en los resultados del modelo, es posible analizar la credibilidad de los mecanismos que el modelo está reproduciendo. De este análisis, se puede deducir si los resultados del modelo, además de parecer correctos, lo son por las razones correctas (Kirchner, 2006).

En la Figura 6.13 se representan dos gráficas que muestran la importancia de cada uno de los tres mecanismos que intervienen en la producción de escorrentía: el flujo base, el flujo subsuperficial o interflujo y el flujo superficial. Es decir las respuestas, lenta, intermedia y rápida del sistema. La importancia de cada componente, tal cual se muestra en las gráficas, está relativizada a la magnitud de la respuesta total.

Para la inferencia con el modelo GL++ (Figura 6.13, izquierda) se puede observar una contribución del flujo base a la respuesta total menor del 20%, para caudales mayores de 4 m³/s (cuya probabilidad de excedencia es del 40%). Para caudales alrededor de 2 m³/s (cuya probabilidad de excedencia es del 60%), el flujo base y el interflujo contribuyen cada uno en un 40% a la respuesta total del sistema. Por el conocimiento que se tiene de la cuenca analizada, la escasa contribución del flujo base estimada por el modelo hidrológico inferido con GL++, no parece compatible con la hidrología observada en la cuenca.

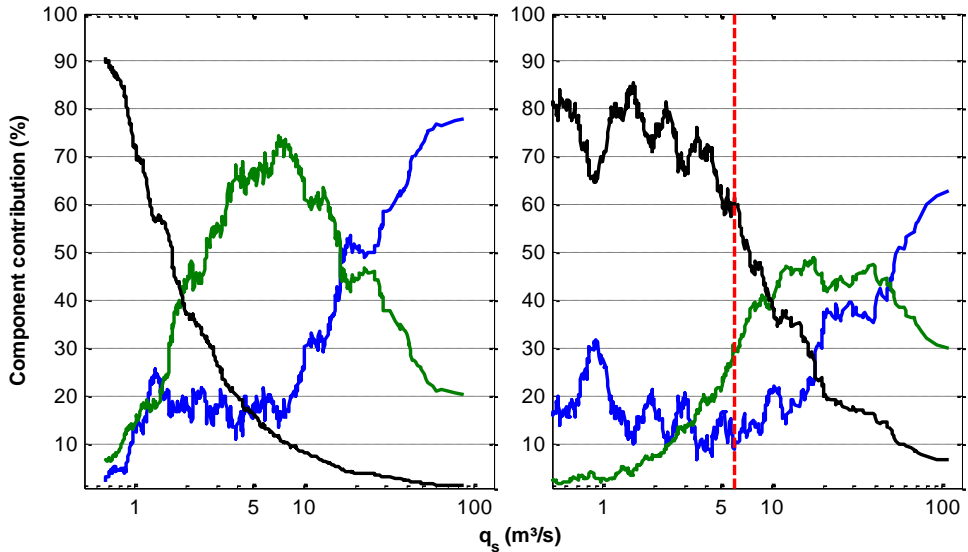


Figura 6.13. Contribución (%) de los tres componentes de la producción de escorrentía a la respuesta total del sistema, en función de la magnitud de dicha respuesta durante el periodo de calibración. La figura muestra los resultados para la inferencia con GL++ (panel izquierdo) y para la inferencia con GL++Bias (panel derecho). Los componentes mostrados son el flujo base (línea negra) el interflujo (línea verde) y escorrentía superficial (línea azul). La línea roja vertical en el panel derecho indica la posición del umbral de caudal inferido para el modelo de sesgo.

La Figura 6.13 en su panel derecho muestra para la inferencia con GL++Bias, la contribución de los tres componentes de la respuesta del sistema. Esta gráfica muestra valores más razonables de la contribución del flujo base. Por ejemplo, para caudales mayores de 4 m³/s la contribución del flujo base es mayor al 60%, y con caudales inferiores a 2 m³/s la contribución media se sitúa por encima del 70%. Resulta de interés observar la posición (línea vertical roja) del umbral de caudal que distingue el sesgo de la predicción del modelo con caudales altos, del sesgo que se produce con caudales bajos. Se observa que dicho umbral marca, de forma grosera, la magnitud de caudal para la cual el flujo base comienza a perder importancia en favor del flujo superficial y la contribución del interflujo comienza a estabilizarse. Por tanto, los mayores responsables del sesgo sobrepredictivo producido en la respuesta total del modelo son sus componentes intermedia y rápida.

CAPÍTULO 7. CONCLUSIONES

De algún modo lo veo, pero no sé cómo declararlo sino diciendo que,
todo lo que comienza a ser y deja de ser,
entonces comienza y entonces acaba,
cuando en la razón eterna en la que nada empieza ni acaba,
se conoce que debió comenzar o debió acabar
-San Agustín (354-430)-

7.1. Epílogo

En todo proceso de modelación se pierden piezas de información, en el camino que hay entre la realidad percibida y la cuantificación de sus magnitudes. Los modelos hidrológicos, y en general los modelos ambientales, no son una excepción sino más bien el paradigma. Las fuentes de incertidumbre son los puntos de ese camino en los que más información se pierde. Los modelos hidrológicos tienen como fuentes principales de incertidumbre a los datos con que se alimentan y a su propia estructura o conceptualización. En el pasado cercano, la modelación hidrológica ha obviado, fuera del ámbito de la investigación, el tema de la incertidumbre. El hecho de trabajar con modelos cada vez más complejos, que requieren mayores y más diversas cantidades de información, ha sacado a la luz la evidencia de que ni los modelos hidrológicos ni la información que los alimenta están libres de error, de incertidumbre. Se ha tomado conciencia de que la incertidumbre es una propiedad inherente a los sistemas naturales, por lo cual la modelación de estos debe incorporar de forma explícita su presencia.

En cualquiera de los tipos de modelos empleados en modelación hidrológica, se va a necesitar asignar valores a los parámetros que no pueden ser medidos de forma directa. El mayor o menor acierto en esa asignación de valores es la clave que nos proporcionará mayor o menor éxito, en el objetivo de emplear los modelos como herramienta de predicción. La forma (o método) en la que se

realice el proceso de inferencia de los parámetros será decisiva en la consecución del citado objetivo.

Por tanto, es necesario que el método de inferencia de los parámetros considere, de forma apropiada, todas las incertidumbres (o su efecto) que están afectando al proceso de modelación, causantes de la diferencia entre lo que observamos en el sistema estudiado y el resultado que nos proporciona el modelo (los errores). De lo contrario, el propio método de inferencia se convertirá en otra fuente de incertidumbre, a añadir a las ya existentes. Durante muchos años, la inferencia de los modelos hidrológicos se ha basado en procedimientos cuyo fundamento es el clásico método de los mínimos cuadrados (SLS), cuyo objetivo fundamental es obtener el mejor ajuste posible entre las observaciones y los resultados del modelo. Estos métodos solamente consideran una única fuente de incertidumbre: los errores de medición en las observaciones de la variable que el modelo está prediciendo. Estos errores tienen como característica distintiva, ser un *ruido blanco* Gaussiano. Los modelos hidrológicos, como se ha comentado, tienen más y más importantes fuentes de incertidumbre y, en consecuencia, los errores generados en su aplicación están lejos de ser un *ruido blanco* Gaussiano.

La presente Tesis Doctoral, en línea con recientes publicaciones de la literatura científica relacionada, se ha marcado dos objetivos. El primer objetivo es la estimación de los parámetros hidrológicos, de la forma menos sesgada posible, así como la correcta evaluación de su incertidumbre mediante el paradigma Bayesiano. El segundo objetivo, es el objetivo fundamental de la modelación hidrológica: la obtención de una correcta predicción, junto con la evaluación de su incertidumbre total. Para la consecución, simultánea y con éxito, de ambos objetivos es necesario realizar la inferencia con el mejor modelo de error posible. Además, esta inferencia debe realizarse, de forma conjunta, sobre los parámetros de los modelos hidrológico y de error. Así, el objetivo de la inferencia ya no es obtener el conjunto de parámetros que consiga el mejor ajuste entre los observados y los simulados de la variable de interés, sino la obtención del conjunto de parámetros que genere los errores (diferencia entre observados y simulados) que mejor respeten las hipótesis del modelo de error. En el caso de cumplirse las hipótesis del SLS, ambos objetivos son el mismo.

Para llevar a cabo los mencionados objetivos, la Tesis se ha inspirado fundamentalmente, al igual que los recientes trabajos de otros autores (Evin et al., 2013, 2014; Koskela et al., 2012; Scharnagl et al., 2015), en el trabajo de Schoups y Vrugt (2010), quienes proponen realizar la inferencia, de forma conjunta, mediante un modelo de error generalizado, el cual aporta una gran flexibilidad a la inferencia al poder relajar todas las hipótesis que limitan el método SLS. Sin embargo, en el trabajo de Schoups y Vrugt (2010), así como en los recientes trabajos que en él se han inspirado, aparecen problemas relacionados con la inferencia de los parámetros, especialmente cuando esta inferencia se realiza de manera conjunta. Los problemas son diversos y a priori, no parecen seguir un patrón fijo que permita identificar su origen. Evin et al (2014) llegaron a afirmar que posiblemente, no pueda realizarse la inferencia conjunta de los parámetros hidrológicos y del modelo de error, debido a las fuertes interacciones que entre ellos se producen.

En esta Tesis, se postula que la clave del problema está en el incorrecto procedimiento seguido en la modelación directa de la varianza de los errores. La forma en que los citados trabajos están aplicando el modelo de varianza origina que, durante la inferencia, la dinámica de variación de las distribuciones condicionales y marginales de los errores esté desacoplada, lo cual provoca una pérdida de la coherencia estadística similar a la de la Figura 3.2. Por tanto, se propone (capítulo 3) como hipótesis principal de la Tesis, la siguiente:

Un modelo de varianza condicional de los errores que siga el método directo, es decir, sin previa transformación de los datos, debe tomar en consideración la relación que existe entre las distribuciones condicionales y la correspondiente distribución marginal de los errores. La variación de ambas distribuciones, durante el proceso de inferencia conjunta, debe ocurrir de forma acoplada, manteniendo así la coherencia estadística.

Esta hipótesis, enunciada para los modelos de varianza condicional, también es de aplicación a los modelos de sesgo condicional de los errores. El acople del comportamiento entre ambas distribuciones, marginal y condicional, es sinónimo de coherencia estadística y ha de venir forzado por la relación matemática que

existe entre ambos tipos de distribuciones, cuando pertenecen a una única distribución de probabilidad conjunta. Esa relación matemática viene definida por las **Leyes Totales de la esperanza (TEL) y de la varianza (TVL)**.

Para comprobar que la hipótesis es correcta, y que con la solución propuesta desaparecen los problemas espurios referidos en la literatura, se ha realizado una serie de experimentos de inferencia conjunta (capítulos 5 y 6) en los que se ha combinado dos modelos hidrológicos agregados (**CRR y GR4J**) y uno distribuido (**TETIS**), con diferentes modelos de error: los clásicos **SLS y WLS** y los nuevos modelos de error propuestos en esta Tesis **GL++ y GL++Bias**. Los resultados obtenidos con los experimentos hacen válido, en un sentido pragmático, lo esencial de la hipótesis propuesta, por lo que la presente Tesis se puede considerar como la afirmación culminada de la hipótesis planteada.

Sin el cumplimiento de la hipótesis propuesta, no hubiera sido posible alcanzar, de forma exitosa, los dos objetivos principales de la Tesis, no al menos, de la forma en la que se han alcanzado. Por ello, la formulación y verificación de la hipótesis acerca de los modelos de varianza y sesgo condicional de los errores, puede considerarse como la principal aportación y conclusión de esta Tesis.

Adicionalmente, el trabajo de investigación ha obtenido multitud de resultados que permiten elaborar las siguientes y no menos importantes conclusiones.

7.2. Conclusiones

Sobre la verificación de la hipótesis principal

Se ha comprobado la incoherencia estadística, en forma de desacople entre la distribución marginal de los errores y sus condicionales, de los modelos que no consideran la aplicación de las TLs: WLS-NTL y GL++NTL. Si bien dicha incoherencia en el plano teórico, no tiene efectos prácticos en WLS-NTL, si los tiene en GL++NTL. Esos efectos son los diferentes problemas que han sido reportados por varios trabajos (citados varias veces en el documento) desde 2010 hasta el presente. Estos efectos espurios son: i) Muy altas correlaciones entre parámetros de error e hidrológicos y correlación entre parámetros hidrológicos; ii) Parámetros de varianza y/o de autocorrelación no identificables;

iii) Inflación de las bandas de incertidumbre (estimación errónea de la incertidumbre) y iv) Estimación errónea de la predicción media.

Sobre la aparentemente errática aparición de los efectos espurios

Como se ha podido comprobar en la Tesis, es diversa la casuística observada en la literatura, respecto de la aparición de los efectos espurios en la inferencia conjunta. El modelo hidrológico, la flexibilidad del modelo de error (n° de parámetros) y las posibles interacciones entre ambos parecen ser los factores determinantes de esta variabilidad en el comportamiento de la inferencia. Por una parte, el modelo hidrológico es el factor que determina la mayor o menor complejidad de la estructura de los errores, es decir, de la distribución bivariada $P(\mathbf{E}, \mathbf{Y}_{\mathcal{M}})$. Por otra parte, el modelo de error, con sus grados de libertad, debe ser capaz de adaptarse a esta estructura de error que se va construyendo durante la inferencia. Un exceso en los grados de libertad del modelo de error convierte al problema de inferencia conjunta, en un problema indeterminado, incapaz de constreñir el valor de sus parámetros principales de error, los cuales definen la forma de la distribución bivariada, esto es: i) El parámetro κ del modelo de varianza, define la tasa de aumento de la dispersión de los errores, al aumentar la magnitud de $\mathbf{Y}_{\mathcal{M}}$; ii) El parámetro ϕ_1 del modelo de dependencia de los errores que, al igual que un modelo de sesgo, actúa sobre la posición de las distribuciones condicionales de los errores y iii) El parámetro de asimetría ξ , que, como los dos anteriores, también influye en la posición de la distribución bivariada. La aplicación de las TLs, además de ser un requerimiento teórico que proporciona consistencia estadística, desde un punto de vista pragmático proporciona una eventual disminución de los grados de libertad del problema y permite así su resolución.

Sobre los modelos SLS y WLS

El primero de los resultados que, aunque ya conocido, ha vuelto a ser refrendado, es el de la invalidez de los modelos de error SLS y WLS para ser empleados en la inferencia de modelos hidrológicos. Sin embargo, no está tan extendida en la literatura, según se deduce de alguna de las publicaciones revisadas, la siguiente idea. De igual manera que un buen índice NSE de la

predicción, obtenido con el modelo de error SLS, no lo hace válido, tampoco hacen válido al modelo de error WLS, unos excepcionales índices de evaluación de su distribución predictiva (Figura 6.6). Los errores en modelación hidrológica no respetan, en general, las hipótesis de SLS o WLS y esto los invalida, independientemente de cualquier otra consideración.

Respecto al modelo de error WLS y de forma específica, respecto a la modelación directa de la varianza de los errores, uno de los resultados reveladores obtenidos de esta investigación, es el vínculo subyacente que existe entre el método híbrido de Sorooshian y Dracup (1980) y la metodología propuesta en esta Tesis para modelar la varianza de los errores de forma directa, sin la transformación previa de las variables. La semejanza en los resultados teóricos, de dos metodologías tan diferentes, es una prueba de fiabilidad de la metodología propuesta, esto es, modelar la varianza de los errores de forma directa requiere vincular el modelo de varianza de los errores (varianzas condicionales) con la distribución marginal de estos. Las ecuaciones (3.46) y (2.23) apoyan esta afirmación.

GL++ o la revelación de las verdaderas capacidades del modelo hidrológico

El mayor o menor acierto en la asignación de valores a los parámetros, es la clave que nos proporcionará mayor o menor éxito en el empleo de un modelo, como herramienta de predicción. En lo práctico, esto se traduce en que los valores de los parámetros deben hacer funcionar el modelo tal cual es, es decir, las razones por las que el modelo muestre un buen o mal funcionamiento han de ser las correctas y esto solo se consigue con un procedimiento de inferencia adecuado que considere, de algún modo, las posibles fuentes de error (o incertidumbre) que afectan al proceso de modelación. En la sección 6.1.4 se expuso un claro y didáctico análisis sobre las distintas inferencias posibles, con los modelos de error probados en la Tesis, de uno de los parámetros hidrológicos del modelo GR4J. Este ejemplo evidencia la necesidad de hacer la inferencia con modelos de error adecuados, si se quiere comprender cuál es la verdadera dinámica de los sistemas analizados.

El modelo de error GL++ se ha revelado como una herramienta que permite comprobar las capacidades reales del modelo hidrológico. En primer lugar, ha mostrado ser capaz de diferenciar el rendimiento de los dos modelos

hidrológicos agregados empleados en esta Tesis, lo cual no ha sido posible ni con SLS ni con WLS. Por otra parte, las predicciones con GL++ han mostrado un sesgo que no aparece al inferir con SLS o WLS. La explicación más razonable es que el sesgo (desconocido) que se produce en los parámetros, al realizar la inferencia con SLS o WLS, se transfiere a la predicción cuando la inferencia se hace con un modelo de error más adecuado a los errores generados. Recuérdese, que los errores son el reflejo de las fuentes de error que afectan al proceso de modelación. La siguiente frase resume estas ideas: cuando hay errores en los datos o en la especificación del modelo, no es posible obtener de forma simultánea unos parámetros insesgados y una predicción insesgada, sacrificándose uno de los dos, según sea el modelo de error empleado en la inferencia.

GL++Bias. Modelar $P(\mathbf{E}|\mathbf{Y}_{\mathcal{M}})$ correctamente, exige modelar su varianza y su sesgo

Con la comparación de las inferencias con GL++ y GL++Bias se ha comprobado que no es posible inferir correctamente el modelo de varianza de los errores sin inferir, de forma igualmente correcta, su modelo de sesgo. Lo contrario también es cierto, lo cual es comprensible, a la luz de la expresión de la ley de la varianza total, cuya ecuación (3.6) incluye ambos términos de varianza y sesgo. El modelo de error GL++Bias ha mostrado ser más adecuado para el modelo hidrológico CRR que para el modelo GR4J. Así, la inferencia de CRR con GL++Bias ha permitido observar unos resultados de la predicción muy precisos pero con una alta incertidumbre, circunstancia que es propia de los modelos sobreparametrizados. Por otra parte, algunos de los parámetros de CRR resultan poco identificables, lo cual también es compatible con el diagnóstico de sobreparametrización del modelo CRR. Este tipo de conclusiones no sería posible con modelos de error como el SLS o el WLS.

Leyes Totales: Necesarias pero no suficientes

A pesar de la importancia de considerar las Leyes Totales al modelar la varianza (y/o el sesgo) de los errores, su aplicación es condición necesaria pero no suficiente para obtener una inferencia correcta. La aplicación de las TLs sobre modelos de varianza o sesgo (o AR, etc.) no apropiados, no va a proporcionar

los buenos resultados que se esperan de la inferencia. Entre los objetivos de esta Tesis no ha estado el de encontrar el mejor modelo de error posible para los experimentos analizados. Sin embargo, sí se ha querido determinar en qué medida los modelos de error han sido más o menos adecuados. El primer test de adecuación que pasan los modelos de error es la comprobación de las hipótesis en las que se basa el modelo. Si su incumplimiento se hace evidente, no deberían emplearse para inferir modelos hidrológicos. Sin embargo, si esta comprobación no es del todo determinante o clarificadora, es posible obtener más pistas sobre su adecuación, mediante la evaluación de la estandarización de los errores, realizada mediante la ecuación (3.15). Si los modelos de sesgo y varianza son los correctos, los errores estandarizados deberían tener los valores teóricos de una media nula y una desviación estándar igual a uno. La obtención de valores significativamente diferentes indica que se está escalando o trasladando la pdf de los errores, pero no realizando su correcta estandarización. Esta forma de evaluar la adecuación de los modelos de sesgo y varianza es denominada por esta Tesis como comprobación de la bondad de estandarización de los errores (GoES, Goodness of Errors Standardization). De los experimentos realizados, la comprobación de GoES muestra (sección 6.1.5 y Tabla 5.1) que la inferencia más apropiada de CRR se realiza con el modelo de error GL++Bias, lo cual también se hace patente al inspeccionar el correspondiente PP-plot de la Figura 5.3, en el que se observa el solapamiento con el bisector del diagrama. La inferencia más apropiada del modelo GR4J ha sido, sin embargo, la realizada con el modelo de error GL++, conclusión compatible con la deducida al analizar los valores de los parámetros inferidos (sección 6.1.4).

7.3. Líneas futuras

Desarrollo de metodologías que muestren como es, y qué variables condicionan, la estructura de error

La inferencia conjunta de los modelos hidrológicos, con modelos de error apropiados, ha mostrado ser útil y necesaria para la obtención de modelos hidrológicos que proporcionen predicciones realistas y robustas. Sin embargo, el gran potencial que presenta la metodología ensayada en la Tesis queda fuertemente condicionado a la idoneidad del modelo de error con el que se lleve

a cabo. Por tanto, el que escribe considera que la continuación de esta investigación tiene que poner el foco en el desarrollo de metodologías que permitan dilucidar cuál es la estructura de error que se presenta en cada caso de inferencia, con el fin de poder realizar la inferencia conjunta con el que ya se sepa que es el mejor modelo de error, adecuado al caso considerado.

A tal fin, se propone el empleo de los métodos directos de modelación de la incertidumbre predictiva (Coccia y Todini, 2011; Todini, 2013). Dichos métodos proporcionan, generalmente, una estimación adecuada de la distribución predictiva de un modelo hidrológico, sin atender al método de inferencia (e.g. SLS) con el que fueron estimados sus parámetros. Partiendo de la distribución predictiva obtenida con estos métodos es posible deducir la estructura de los errores, siguiendo el camino inverso al que se sigue en la metodología de inferencia conjunta desarrollada en esta Tesis. En los experimentos ya realizados en este sentido, se ha observado que, tanto el modelo de varianza como el de sesgo de los errores, muestran una dependencia que no se limita únicamente a la variable predicha, el caudal en esta Tesis, puesto que, para un mismo valor de esa variable, se obtienen diferentes valores de varianza y/o de sesgo. Por tanto, debe de haber otras variables que influyen en estas dos propiedades de los errores. En este sentido, merecería el esfuerzo considerar la relación que pueda existir entre el sesgo y varianza de los errores, no solo con variables de estado del modelo, sino con *firmas* hidrológicas que no dependan de los datos de entrada al modelo. Esto permitiría dilucidar cómo y en qué medida influye exclusivamente la estructura del modelo (sus deficiencias) en las propiedades de los errores.

El desarrollo de la metodología apropiada para la determinación de qué variables (o *firmas*) influyen sobre los modelos de varianza y sesgo, sería el paso final para la consecución del objetivo planteado de obtener el mejor modelo de error posible.

Aplicación de métodos de selección de modelos

Como criterio adicional, para la selección del mejor modelo de error dado el modelo hidrológico, parece interesante la exploración de los métodos (BMA) existentes para la selección de modelos, en este caso, aplicando esas metodologías al conjunto modelo hidrológico–modelo de error. Parecen

prometedores los métodos que basan esta evaluación en la estimación de la evidencia Bayesiana del modelo, los cuales fueron mencionados en la sección 2.2.3 de esta Tesis.

Combinación de las metodologías agregada y desagregada

Separándose ligeramente de lo propuesto en esta Tesis, podría ser interesante explorar la vía de la consideración de forma desagregada (secciones 2.2.3 y 2.3.2) de las fuentes de incertidumbre en combinación con los modelos de error. Con esta combinación el modelo de error se ocuparía de la incertidumbre no contemplada de forma explícita o desagregada. Resultaría de interés analizar los cambios que se producen en los parámetros, en las predicciones y en sus estimaciones de incertidumbre, cuando se combinan ambas metodologías respecto a los resultados obtenidos en esta Tesis.

8. REFERENCIAS

Ajami, Duan and Sorooshian, S.: An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour.*, 43(1), 1–19, doi:10.1029/2005WR004745, 2007.

Anctil, F., Perrin, C. and Andréassian, V.: Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models, *Environ. Model. Softw.*, 19(4), 357–368, doi:10.1016/S1364-8152(03)00135-X, 2004.

Anderson, R. L.: Distribution of the serial correlation coefficient, *Ann. Math. Stat.*, 13(1), 1–13, doi:10.1214/aoms/1177731638, 1942.

Banks, H., Hu, S. and Thompson, W.: *Modeling and Inverse Problems in the Presence of Uncertainty*, CRC Press., 2014.

Bates, B. C. and Campbell, E. P.: A Markov chain Monte-Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling, *Water Resour. Res.*, 37(4), 937–947, doi:10.1029/2000WR900363, 2001.

Beck, M. B.: Water quality modeling: A review of the analysis of uncertainty, *Water Resour. Res.*, 23(8), 1393–1442, doi:10.1029/WR023i008p01393, 1987.

Berger, J. O., Bernardo, J. M. and Sun, D.: The formal definition of reference priors, *Ann. Stat.*, 37(2), 905–938, doi:10.1214/07-AOS587, 2009.

Best, N., Cowles, M. and Vines, S.: *CODA Manual version 0.30*, MRC Biostat. Unit, Cambridge, UK, 1995.

Beven, K.: Changing ideas in hydrology-The case of physically based models, *J. Hydrol.*, 105(1–4), 157–172, doi:10.1016/0022-1694(90)90161-P, 1989.

Beven, K.: How far can we go in distributed hydrological modelling?, *Hydrol. Earth Syst. Sci. [Hydrol. Earth Syst. Sci.]*. Vol. 5, 5(1), 1–12, doi:10.5194/hess-5-1-2001, 2001.

Beven, K.: Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system, *Hydrol. Process.*, 16(2), 189–206, doi:10.1002/hyp.343, 2002.

Beven, K.: A manifesto for the equifinality thesis, *J. Hydrol.*, 320(1–2), 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K.: *Environmental Modelling: An Uncertain Future?*, 2009.

Beven, K.: *Rainfall-Runoff Modelling: The Primer: Second Edition.*, 2012.

Beven, K. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction., *Hydrol. Process.*, 6(May 1991), 279–298, doi:10.1002/hyp.3360060305, 1992.

Beven, K. J.: Uniqueness of place and process representation in hydrological modelling., *Hydrol. Earth Syst. Sci.*, 4, 203–213, doi:10.5194/hess-4-203-2000, 2000.

Bolstad, W. M.: *Introduction to Bayesian Statistics.*, 2007.

Bolstad, W. M.: *Understanding Computational Bayesian Statistics*, Wiley., 2010.

Bowles, D. S. and O’Connell, P. E.: *Recent advances in the modeling of hydrologic systems.*, 1991.

Box, G. E. P. and Cox, D. R.: An analysis of transformations, *J. R. Stat. Soc. Ser. B (Methodological)*, 211–252, doi:10.2307/2287791, 1964.

Box, G. E. P. and Tiao, G. C.: *Bayesian Inference in Statistical Analysis*, Wiley., 1992.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C.: *Time Series Analysis: Forecasting & Control.*, 1994.

Ter Braak, C. J. F.: A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: Easy Bayesian computing for real parameter spaces, *Stat. Comput.*, 16(3), 239–249, doi:10.1007/s11222-006-8769-1, 2006.

Ter Braak, C. J. F. and Vrugt, J. A.: Differential Evolution Markov Chain with snooker updater and fewer chains, *Stat. Comput.*, 18(4), 435–446, doi:10.1007/s11222-008-9104-9, 2008.

Brigode, P., Oudin, L. and Perrin, C.: Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change?, *J. Hydrol.*, 476, 410–425, doi:10.1016/j.jhydrol.2012.11.012, 2013.

Brockwell, P. J.: *Time Series Analysis*, *Int. Encycl. Educ.*, 474–481, doi:10.1016/B978-0-

08-044894-7.01372-5, 2010.

Brooks, S. P. and Gelman, A.: General Methods for Monitoring Convergence of Iterative Simulations, *J. Comput. Graph. Stat.*, 7(4), 434–455, doi:10.1080/10618600.1998.10474787, 1998.

Brooks, S. P. and Giudici, P.: Markov Chain Monte Carlo Convergence Assessment via Two-Way Analysis of Variance, *J. Comput. Graph. Stat.*, 9(2), 266–285, doi:10.2307/1390654, 2000.

Brun, R., Reichert, P. and Ku, H.: Practical identifiability analysis of large environmental simulation models, *Water Resour. Res.*, 37(4), 1015–1030, 2001.

Burnham, K. P. and Anderson, D. R.: *Model Selection and Multimodel Inference.*, 2002.

Burnham, K. P., Anderson, D. R. and Huyvaert, K. P.: AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons, *Behav. Ecol. Sociobiol.*, 65(1), 23–35, doi:10.1007/s00265-010-1029-6, 2011.

Bussi, G., Rodriguez-Lloveras, X., Frances, F., Benito, G., Sanchez-Moya, Y. and Sopeña, A.: Sediment yield model implementation based on check dam infill stratigraphy in a semiarid Mediterranean catchment, *Hydrol. Earth Syst. Sci.*, 17(8), 3339–3354, doi:10.5194/hess-17-3339-2013, 2013.

Cheng, Q. B., Chen, X., Xu, C. Y., Reinhardt-Imjela, C. and Schulte, A.: Improvement and comparison of likelihood functions for model calibration and parameter uncertainty analysis within a Markov chain Monte Carlo scheme, *J. Hydrol.*, 519(PB), 2202–2214, doi:10.1016/j.jhydrol.2014.10.008, 2014.

Clark, J. S.: Why environmental scientist are becoming Bayesians, *Ecol. Lett.*, 8, 2–14, 2005.

Clark, M. P., Kavetski, D. and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resour. Res.*, 47(9), n/a-n/a, doi:10.1029/2010WR009827, 2011.

Coccia, G. and Todini, E.: Recent developments in predictive uncertainty assessment based on the model conditional processor approach, *Hydrol. Earth Syst. Sci.*, 15(10), 3253–3274, doi:10.5194/hess-15-3253-2011, 2011.

Cowles, M. K. and Carlin, B. P.: Markov chain Monte Carlo convergence diagnostics: a

comparative review, *J. Amer. Stat. Assoc.*, 91(434), 883–904, doi:10.2307/2291683, 1996.

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E. and François, O.: Approximate Bayesian Computation (ABC) in practice, *Trends Ecol. Evol.*, 25(7), 410–418, doi:10.1016/j.tree.2010.04.001, 2010.

Datta, A. R. and Bolisetti, T.: Second-order autoregressive model based likelihood function for calibration and uncertainty analysis of SWAT model, *J. Hydrol. Eng.*, 131003223048007, doi:10.1061/(ASCE)HE.1943-5584.0000917, 2013.

Devia, G., Ganasri, B. and Dwarakish, G.: A Review on Hydrological Models, *Aquat. Procedia*, doi:10.1016/j.aqpro.2015.02.126, 2015.

Dibike, Y. B. and Solomatine, D. P.: River flow forecasting using artificial neural networks, *Phys. Chem. Earth, Part B Hydrol. Ocean. Atmos.*, 26(1), 1–7, doi:10.1016/S1464-1909(01)85005-X, 2001.

Duan, Q., Sorooshian, S. and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28(4), 1015–1031, 1992.

Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N. and Turcotte, R.: Calibration of Watershed Models., 2003.

Duan, Q., Schaake, J., Andrassian, V., Franks, S., Goteti, G., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T. and Wood, E. F.: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, in *Journal of Hydrology*, vol. 320, pp. 3–17., 2006.

Eckhardt, R.: Stan Ulam, John von Neumann, and the Monte Carlo Method, *Los Alamos Sci., Special Is*, 131–141, 1987.

Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration approaches in hydrological modelling: a review, *Hydrol. Sci. J.*, 55(1), 58–78, doi:10.1080/02626660903526292, 2010.

Elshorbagy, A., Corzo, G., Srinivasulu, S. and Solomatine, D. P.: Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: Concepts and methodology, *Hydrol. Earth Syst. Sci.*, 14(10), 1931–1941, doi:10.5194/hess-14-1931-2010, 2010a.

Elshorbagy, A., Corzo, G., Srinivasulu, S. and Solomatine, D. P.: Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 2: Application, *Hydrol. Earth Syst. Sci.*, 14(10), 1943–1961, doi:10.5194/hess-14-1943-2010, 2010b.

Evin, G., Kavetski, D., Thyer, M. and Kuczera, G.: Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration, *Water Resour. Res.*, 49(7), 4518–4524, doi:10.1002/wrcr.20284, 2013.

Evin, G., Thyer, M., Kavetski, D., McInerney, D. and Kuczera, G.: Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity, *Water Resour. Res.*, 50(3), 2350–2375, doi:10.1002/2013WR014185, 2014.

Fenicia, F., Savenije, H. H. G., Matgen, P. and Pfister, L.: A comparison of alternative multiobjective calibration strategies for hydrological modeling, *Water Resour. Res.*, 43(3), 1–16, doi:10.1029/2006WR005098, 2007.

Fernandez, C. and Steel, M. F. J.: On Bayesian Modeling of Fat Tails and Skewness, *J. Am. Stat. Assoc.*, 93(441), 359–371, doi:10.1080/01621459.1998.10474117, 1998.

Frances, F., Velez, J. I. and Velez, J. J.: Split-parameter structure for the automatic calibration of distributed hydrological models, *J. Hydrol.*, 332(1–2), 226–240, doi:10.1016/j.jhydrol.2006.06.032, 2007.

Francés, F. and Casado, J.: La modelación distribuida con pocos parámetros de las crecidas, *Ing. del Agua*, 2(diciembre), 7–24, 1995.

Freeze, R. A. and Harlan, R. L.: Blueprint for a physically-based, digitally-simulated hydrologic response model, *J. Hydrol.*, 9(3), 237–258, doi:10.1016/0022-1694(69)90020-1, 1969.

García-Bartual, R.: Short term river flood forecasting with neural networks, in *International Congress on Environmental Modelling and Software.*, Iemss., 2002.

Gelb, A.: *Applied Optimal Estimation*, MIT Press., 2001.

Gelman, A. and Rubin, D. B.: Inference from Iterative Simulation Using Multiple Sequences, *Stat. Sci.*, 7(4), 457–472, doi:10.2307/2246093, 1992.

Gelman, A. and Shirley, K.: *Inference from simulations and monitoring convergence.*,

2011.

Geman, S., Doursat, R. and Bienenstock, E.: Neural Networks and the Bias-Variance Dilemma, *Neural Comput.*, 4, 1–58, doi:10.1162/neco.1992.4.1.1, 1992.

Geweke, J.: Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, *Bayesian Stat.* 4, 169–193, doi:citeulike-article-id:1176289, 1992.

Del Giudice, D., Honti, M., Scheidegger, A., Albert, C., Reichert, P. and Rieckermann, J.: Improving uncertainty estimation in urban hydrological modeling by statistically describing bias, *Hydrol. Earth Syst. Sci.*, 17(10), 4209–4225, doi:10.5194/hess-17-4209-2013, 2013.

Gneiting, T., Balabdaoui, F. and Raftery, A. E.: Probabilistic forecasts, calibration and sharpness, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 69(2), 243–268, doi:10.1111/j.1467-9868.2007.00587.x, 2007.

De Gooijer, J. G. and Hyndman, R. J.: 25 Years of Time Series Forecasting, *Int. J. Forecast.*, 22(3), 443–473, doi:10.1016/j.ijforecast.2006.01.001, 2006.

Gourley, J. J. and Vieux, B. E.: A method for identifying sources of model uncertainty in rainfall-runoff simulations, *J. Hydrol.*, 327(1–2), 68–80, doi:10.1016/j.jhydrol.2005.11.036, 2006.

Govindaraju, R. S. and Rao, A. R.: *Artificial Neural Networks in Hydrology*, edited by R. S. Govindaraju and A. R. Rao, Springer Netherlands, Dordrecht., 2000.

Gupta, H. V., Sorooshian, S. and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, 34(4), 751–763, doi:10.1029/97WR03495, 1998.

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G. and Ye, M.: Towards a comprehensive assessment of model structural adequacy, *Water Resour. Res.*, 48(8), n/a--n/a, doi:10.1029/2011WR011044, 2012.

Gupta, V. K. and Sorooshian, H.: The automatic calibration of conceptual catchment models using derivative based optimization algorithms, *Water Resour. Res.*, 21(4), 473–485, 1985.

Haario, H.: An adaptive Metropolis algorithm, *Bernoulli*, 7(2), 223–242, doi:10.2307/3318737, 2001.

Haario, H., Saksman, E. and Tamminen, J.: Adaptive proposal distribution for random walk Metropolis algorithm, *Comput. Stat.*, 14(3), 375, doi:10.1007/s001800050022, 1999.

Haario, H., Laine, M., Mira, A. and Saksman, E.: DRAM: Efficient adaptive MCMC, *Stat. Comput.*, 16(4), 339–354, doi:10.1007/s11222-006-9438-0, 2006.

Han, F. and Zheng, Y.: Multiple-response Bayesian calibration of watershed water quality models with significant input and model structure errors, *Adv. Water Resour.*, 88, 109–123, doi:10.1016/j.advwatres.2015.12.007, 2016.

Hastings, W. K.: Monte Carlo sampling methods using Markov chains and their applications, *Biometrika Vol*, 57, 97–109, 1970.

Hendrickson, J. D. and Sorooshian, S.: Comparison of Newton-Type and Direct Search Algorithms for Calibration Rainfall-Runoff Models, *Water Resour. Res.*, 24(5), 691–700, 1988.

Holland, J. H.: *Adaptation in Natural and Artificial Systems.*, 1975.

Honti, M., Stamm, C. and Reichert, P.: Integrated uncertainty assessment of discharge predictions with a statistical error model, *Water Resour. Res.*, 49(8), 4866–4884, doi:10.1002/wrcr.20374, 2013.

Jakeman, A. J. and Hornberger, G. M.: How Much Complexity Is Warranted in a Rainfall-Runoff Model ? are good predictors of streamflow and, *Water Resour. Res.*, 29(8), 2637–2649, doi:10.1029/93WR00877, 1993.

Kavetski, D., Franks, S. W. and Kuczera, G.: Confronting input uncertainty in environmental modelling, in *Water Science and Application*, vol. 6, pp. 49–68, AGU, Washington, DC., 2003.

Kavetski, D., Kuczera, G. and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.*, 42(3), n/a-n/a, doi:10.1029/2005WR004368, 2006a.

Kavetski, D., Kuczera, G. and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resour. Res.*, 42(3), n/a-n/a, doi:10.1029/2005WR004376, 2006b.

Kelly, K. and Krzysztofowicz, R.: A bivariate meta-Gaussian density for use in hydrology,

Stoch. Hydrol. Hydraul., 11(1), 17–31, doi:10.1007/BF02428423, 1997.

Kennedy, J. and Eberhart, R.: Particle swarm optimization, Neural Networks, 1995. Proceedings., IEEE Int. Conf., 4, 1942–1948 vol.4, doi:10.1109/ICNN.1995.488968, 1995.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resour. Res., 42(3), n/a--n/a, doi:10.1029/2005WR004362, 2006.

Kirkpatrick, S., Gelatt Jr, C. and Vecchi, M.: Optimization by Simulated Annealing, Science (80-.), 220(4598), 671–680, doi:10.1126/science.220.4598.671, 1983.

Kollat, J. B., Reed, P. M. and Wagener, T.: When are multiobjective calibration trade-offs in hydrologic models meaningful?, Water Resour. Res., 48(3), W03520, doi:10.1029/2011WR011534, 2012.

Koskela, J. J., Croke, B. W. F., Koivusalo, H., Jakeman, A. J. and Kokkonen, T.: Bayesian inference of uncertainties in precipitation-streamflow modeling in a snow affected catchment, Water Resour. Res., 48(11), n/a--n/a, doi:10.1029/2011WR011773, 2012.

Koutsoyiannis, D.: A random walk on water, Hydrol. Earth Syst. Sci. Discuss., 6(5), 6611–6658, doi:10.5194/hessd-6-6611-2009, 2009.

Koutsoyiannis, D. and Montanari, A.: Statistical analysis of hydroclimatic time series: Uncertainty and insights, Water Resour. Res., 43(5), doi:10.1029/2006WR005592, 2007.

Krzysztofowicz, R.: Bayesian theory of probabilistic forecasting via deterministic hydrologic model, Water Resour. Res., 35(9), 2739–2750, doi:10.1029/1999WR900099, 1999.

Krzysztofowicz, R.: The case for probabilistic forecasting in hydrology, J. Hydrol., 249(1–4), 2–9, doi:10.1016/S0022-1694(01)00420-6, 2001.

Krzysztofowicz, R.: Bayesian system for probabilistic river stage forecasting, J. Hydrol., 268(1–4), 16–40, doi:10.1016/S0022-1694(02)00106-3, 2002.

Krzysztofowicz, R. and Herr, H. D.: Hydrologic uncertainty processor for probabilistic river stage forecasting: precipitation-dependent model, J. Hydrol., 249(1–4), 46–68, doi:10.1016/S0022-1694(01)00412-7, 2001.

Krzysztofowicz, R. and Kelly, K. S.: Hydrologic uncertainty processor for probabilistic river

stage forecasting, *Water Resour. Res.*, 36(11), 3265–3277, doi:10.1029/2000WR900108, 2000.

Kuczera, G.: Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty, *Water Resour. Res.*, 19(5), 1151–1162, doi:10.1029/WR019i005p01151, 1983.

Kuczera, G. and Mroczkowski, M.: Assessment of hydrologic parameter uncertainty and the worth of multiresponse data, *Water Resour. Res.*, 34(6), 1481–1489, doi:10.1029/98WR00496, 1998.

Kuczera, G. and Parent, E.: Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, *J. Hydrol.*, 211(1–4), 69–85, doi:10.1016/S0022-1694(98)00198-X, 1998.

Kuczera, G., Kavetski, D., Franks, S. and Thyer, M.: Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters, *J. Hydrol.*, 331(1–2), 161–177, doi:10.1016/j.jhydrol.2006.05.010, 2006.

Kuczera, G., Kavetski, D., Renard, B. and Thyer, M.: Bayesian Total Error Analysis For Hydrologic Models: Markov Chain Monte Carlo Methods To Evaluate The Posterior Distribution, *Mod. Simul. Soc. Aust. NZ*, 2466–2472, 2007.

Laine, M.: *Adaptive MCMC Methods with Applications in Environmental and Models.*, 2008.

Laio, F. and Tamea, S.: Verification tools for probabilistic forecasts of continuous hydrological variables, *Hydrol. Earth Syst. Sci. Discuss.*, 3(4), 2145–2173, doi:10.5194/hessd-3-2145-2006, 2006.

Laloy, E. and Vrugt, J.: High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing, *Water Resour. Res.*, 48(1), 1–18, doi:10.1029/2011WR010608, 2012.

Li, L., Xu, C. Y., Xia, J., Engeland, K. and Reggiani, P.: Uncertainty estimates by Bayesian method with likelihood of AR (1) plus Normal model and AR (1) plus Multi-Normal model in different time-scales hydrological models, *J. Hydrol.*, 406(1–2), 54–65, doi:10.1016/j.jhydrol.2011.05.052, 2011.

Lindley, D. V: *Bayesian statistics: a review.*, Society for Industrial and Applied

Mathematics (SIAM)., 1972.

Madsen, H.: Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *J. Hydrol.*, 235(3–4), 235,276-288, doi:10.1016/S0022-1694(00)00279-1, 2000.

Mantovan, P. and Todini, E.: Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology, *J. Hydrol.*, 330(1–2), 368–381, doi:10.1016/j.jhydrol.2006.04.046, 2006.

Marin, J.-M., Pudlo, P., Robert, C. P. and Ryder, R. J.: Approximate Bayesian computational methods, *Stat. Comput.*, 22, 1167–1180, doi:10.1007/s11222-011-9288-2, 2012.

Matott, L. S., Babendreier, J. E. and Purucker, S. T.: Evaluating uncertainty in integrated environmental models: A review of concepts and tools, *Water Resour. Res.*, 45(6), 1–14, doi:10.1029/2008WR007301, 2009.

Matthew Richey: The Evolution of Markov Chain Monte Carlo Methods, *Am. Math. Mon.*, 117(5), 383, doi:10.4169/000298910x485923, 2010.

Mayer, J., Khairy, K. and Howard, J.: Drawing an elephant with four complex parameters, *Am. J. Phys.*, 78(6), 648–649, doi:10.1119/1.3254017, 2010.

McInerney, D., Thyer, M., Kavetski, D., Lerat, J. and Kuczera, G.: Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors, *Water Resour. Res.*, 1–41, doi:10.1002/2016WR019168, 2017.

Metropolis: The Monte Carlo Method, ... *Am. Stat. Assoc.*, (April 2013), 37–41, 1949.

Metropolis, N.: The beginning of the Monte Carlo method, *Los Alamos Sci.*, 15, 125–130, doi:10.1128/JCM.05092-11, 1987.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E.: Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 21(6), 1087–1092, doi:http://dx.doi.org/10.1063/1.1699114, 1953.

Montanari, A.: Uncertainty of hydrological predictions, *Treatise Water Sci.*, 2, 459–478, doi:10.1016/B978-0-444-53199-5.00045-2, 2011.

Montanari, A. and Brath, A.: A stochastic approach for assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, 40(1), 1–11, doi:10.1029/2003WR002540, 2004.

Montanari, A. and Koutsoyiannis, D.: A blueprint for process-based modeling of uncertain hydrological systems, *Water Resour. Res.*, 48(9), n/a-n/a, doi:10.1029/2011WR011412, 2012.

Müller, P.: A generic approach to posterior integration and Gibbs sampling., 1991.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I—a discussion of principles, *J. Hydrol.*, 10(3), 282–290, 1970.

Nelder, J. A.: A simplex method for function minimization, *Comput. J.*, 1965.

Nott, D. J., Marshall, L. and Ngoc, T. M.: The ensemble Kalman filter is an ABC algorithm, *Stat. Comput.*, 22(6), 1273–1276, doi:10.1007/s11222-011-9300-x, 2012.

Omlin, M. and Reichert, P.: A comparison of techniques for the estimation of model prediction uncertainty, *Ecol. Modell.*, 115(1), 45–59, doi:10.1016/S0304-3800(98)00174-4, 1999.

Pappenberger, F. and Beven, K. J.: Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resour. Res.*, 42(5), n/a-n/a, doi:10.1029/2005WR004820, 2006.

Perrin, C., Michel, C. and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, 279(1–4), 275–289, doi:10.1016/S0022-1694(03)00225-7, 2003.

Perrin, C., Oudin, L., ANDREASSIAN, V., ROJAS-SERNA, C., MICHEL, C. and MATHEVET, T.: Impact of limited streamflow data on the efficiency and the parameters of rainfall—runoff models, *Hydrol. Sci. J.*, 52(1), 131–151, doi:10.1623/hysj.52.1.131, 2007.

Pianosi, F. and Raso, L.: Dynamic modeling of predictive uncertainty by regression on absolute errors, *Water Resour. Res.*, 48(3), n/a--n/a, doi:10.1029/2011WR010603, 2012.

Pianosi, F. and Wagener, T.: A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, *Environ. Model. Softw.*, 67, 1–11, doi:10.1016/j.envsoft.2015.01.004, 2015.

Pianosi, F. and Wagener, T.: Understanding the time-varying importance of different uncertainty sources in hydrological modelling using global sensitivity analysis, *Hydrol. Process.*, 30(22), 3991–4003, doi:10.1002/hyp.10968, 2016.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B. and Wagener, T.: Sensitivity analysis of environmental models: A systematic review with practical workflow, *Environ. Model. Softw.*, 79, 214–232, doi:10.1016/j.envsoft.2016.02.008, 2016.

Pokhrel, P. and Gupta, H. V.: On the use of spatial regularization strategies to improve calibration of distributed watershed models, *Water Resour. Res.*, 46(1), 1–17, doi:10.1029/2009WR008066, 2010.

Popper, K.: *Philosophy of science: A personal report*, Br. Philos. mid-century, 1957.

Prado, R. and West, M.: *Time series: modeling, computation, and inference.*, 2010.

Price, K. V., Storn, R. M. and Lampinen, J. A.: *Differential Evolution - A Practical Approach to Global Optimization*, edited by G. Rozenberg, T. Bäck, A. E. Eiben, J. N. Kok, and H. P. Spauk, Springer Berlin Heidelberg., 2005.

Raftery, A. E. and Lewis, S.: How many iterations in the Gibbs sampler?, *Bayesian Stat.*, 763--773, 1992.

Refsgaard, J. C., van der Sluijs, J. P., Brown, J. and van der Keur, P.: A framework for dealing with uncertainty due to model structure error, *Adv. Water Resour.*, 29(11), 1586–1597, doi:10.1016/j.advwatres.2005.11.013, 2006.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L. and Vanrolleghem, P. A.: Uncertainty in the environmental modelling process - A framework and guidance, *Environ. Model. Softw.*, 22(11), 1543–1556, doi:10.1016/j.envsoft.2007.02.004, 2007.

Reichert, P. and Mieleitner, J.: Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters, *Water Resour. Res.*, 45(10), n/a--n/a, doi:10.1029/2009WR007814, 2009.

Reichert, P. and Omlin, M.: On the usefulness of overparameterized ecological models, *Ecol. Modell.*, 95(2–3), 289–299, doi:10.1016/S0304-3800(96)00043-9, 1997.

Remesan, R. and Mathew, J.: *Hydrological Data Driven Modelling*, Springer International Publishing, Cham., 2015.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M. and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resour. Res.*, 46(5), W05521, doi:10.1029/2009WR008328, 2010.

Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G. and Franks, S. W.: Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation, *Water Resour. Res.*, 47(11), doi:10.1029/2011WR010643, 2011.

Robert, C. P.: *The Bayesian Choice*, Springer New York, New York, NY., 2007.

Roberts, G. O. and Rosenthal, J. S.: Coupling and ergodicity of adaptive Markov chain Monte Carlo algorithms, *J. Appl. Probab.*, 44(2), 458–475, doi:10.1239/jap/1183667414, 2007.

Roberts, G. O. and Rosenthal, J. S.: Examples of Adaptive MCMC, *J. Comput. Graph. Stat.*, 18(2), 349–367, doi:10.1198/jcgs.2009.06134, 2009.

Rodda, J. C. and Robinson, M.: *Progress in Modern Hydrology: Past, Present and Future.*, 2015.

Sadegh, M. and Vrugt, J. A.: Bridging the gap between GLUE and formal statistical approaches: Approximate Bayesian computation, *Hydrol. Earth Syst. Sci.*, 17(12), 4831–4850, doi:10.5194/hess-17-4831-2013, 2013.

Sage, A. and Melsa, J.: *Estimation theory with applications to communications and control*, McGraw-Hill., 1971.

Sakia, R. M.: The Box-Cox Transformation Technique: A Review, *J. R. Stat. Soc. Ser. D (The Stat.)*, 41(2), 169–178, doi:10.2307/2348250, 1992.

Salazar, S., Frances, F., Komma, J., Blume, T., Francke, T., Bronstert, A. and Blöschl, G.: A comparative analysis of the effectiveness of flood management measures based on the concept of “retaining water in the landscape” in different European hydro-climatic regions, *Nat. Hazards Earth Syst. Sci.*, 12(11), 3287–3306, doi:10.5194/nhess-12-3287-2012, 2012.

Saltelli, A.: *Global sensitivity analysis: an introduction*. [online] Available from: http://sensitivity-analysis.jrc.it/Events/SAMO2005/Saltelli_tutorial.pdf, 2004.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S.: *Global sensitivity analysis: the primer.*, 2008.

Scharnagl, B., Iden, S. C., Durner, W., Vereecken, H. and Herbst, M.: Inverse modelling of in situ soil water dynamics: accounting for heteroscedastic, autocorrelated, and non-Gaussian distributed residuals, *Hydrol. Earth Syst. Sci. Discuss.*, 12(2), 2155–2199, doi:10.5194/hessd-12-2155-2015, 2015.

Schöniger, A., Wöhling, T., Samaniego, L. and Nowak, W.: Model selection on solid ground: Rigorous comparison of nine ways to evaluate Bayesian model evidence, *Water Resour. Res.*, 50(12), 9484–9513, doi:10.1002/2014WR016062, 2014.

Schoups, G. and Vrugt, J.: A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, *Water Resour. Res.*, 46(10), 1–17, doi:10.1029/2009WR008933, 2010.

Schoups, G., Vrugt, J. A., Fenicia, F. and Van De Giesen, N. C.: Corruption of accuracy and efficiency of Markov chain Monte Carlo simulation by inaccurate numerical implementation of conceptual hydrologic models, *Water Resour. Res.*, 46(10), W10530, doi:10.1029/2009WR008648, 2010.

Shi, Y. and Eberhart, R.: A modified particle swarm optimizer, 1998 IEEE Int. Conf. Evol. Comput. Proceedings. IEEE World Congr. Comput. Intell. (Cat. No.98TH8360), 69–73, doi:10.1109/ICEC.1998.699146, 1998.

Shrestha, D. L., Kayastha, N. and Solomatine, D. P.: A novel approach to parameter uncertainty analysis of hydrological models using neural networks, *Hydrol. Earth Syst. Sci.*, 13(7), 1235–1248, 2009.

Shumway, R. H. and Stoffer, D. S.: *Time Series Analysis and its Applications.*, 2011.

Smith, T., Sharma, A., Marshall, L., Mehrotra, R. and Sisson, S.: Development of a formal likelihood function for improved Bayesian inference of ephemeral catchments, *Water Resour. Res.*, 46(12), n/a–n/a, doi:10.1029/2010WR009514, 2010.

Smith, T., Marshall, L. and Sharma, A.: Modeling residual hydrologic errors with Bayesian inference, *J. Hydrol.*, 528(Sept-2015), 29–37, doi:10.1016/j.jhydrol.2015.05.051, 2015.

Sorooshian, S. and Dracup, J. A.: Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour. Res.*, 16(2), 430–442, doi:10.1029/WR016i002p00430, 1980.

Sorooshian, S. and Gupta, H. V.: Model calibration. Chapter 2. Computer models of watershed hydrology, in *Computer models of watershed hydrology*, edited by V. P. Singh, pp. 23–68, Water Resources Publications. Littleton, CO., 1995.

Sorooshian, S. and Gupta, V. K.: Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness, *Water Resour. Res.*, 19(1), 260–268, doi:10.1029/WR019i001p00260, 1983.

Sorooshian, S., Duan, Q. Y. and Gupta, V. K.: Calibration of Rainfall-Runoff Models - Application of Global Optimization to the Sacramento Soil-Moisture Accounting Model, *Water Resour. Res.*, 29(4), 1185–1194, doi:10.1029/92WR02617, 1993.

Stigler, S. M.: Who Discovered Bayes Theorem, *Am. Stat.*, 37(4a), 290–296, doi:10.1080/00031305.1983.10483122, 1983.

Storn, R. and Price, K.: Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces, *J. Glob. Optim.*, 11(4), 341–359, doi:10.1023/A:1008202821328, 1997.

Sun, N.-Z. and Sun, A.: *Model Calibration and Parameter Estimation*, Springer New York, New York, NY., 2015.

Thiemann, M., Trosset, M., Gupta, H. and Sorooshian, S.: Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, 37(10), 2521–2535, doi:10.1029/2000WR900405, 2001.

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W. and Srikanthan, S.: Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, *Water Resour. Res.*, 45(12), n/a–n/a, doi:10.1029/2008WR006825, 2009.

Tierney, L.: Markov Chains for Exploring Posterior Distributions, *Ann. Stat.*, 22(4), 1701–1728, doi:10.1214/aos/1176325750, 1994.

Todini, E.: Rainfall-runoff modeling - Past, present and future, *J. Hydrol.*, 100(1–3), 341–352, doi:10.1016/0022-1694(88)90191-6, 1988.

Todini, E.: Hydrological catchment modelling: past, present and future, *Hydrol. Earth Syst. Sci.*, 11(1), 468–482, doi:10.5194/hess-11-468-2007, 2007.

Todini, E.: A model conditional processor to assess predictive uncertainty in flood

forecasting, *Int. J. River Basin Manag.*, 6(2), 123–137, doi:10.1080/15715124.2008.9635342, 2008.

Todini, E.: FROM HUP TO MCP., 2013a.

Todini, E.: From HUP to MCP: Analogies and extended performances, *J. Hydrol.*, 477, 33–42, doi:10.1016/j.jhydrol.2012.10.037, 2013b.

Vélez, J. J., López Unzu, F., Puricelli, M. and Francés, F.: Parameter extrapolation to ungauged basins with a hydrological distributed model in a regional framework, *Hydrol. Earth Syst. Sci. Discuss.*, 4(2), 909–956, doi:10.5194/hessd-4-909-2007, 2007.

Vexler, A., Hutson, A. D. and Chen, X.: *Statistical Testing Strategies in the Health Sciences*, edited by Taylor & Francis, CRC Press., 2016.

Vrugt, J. and Bouten, W.: Validity of First-Order Approximations to Describe Parameter Uncertainty in Soil Hydrologic Models, *Soil Sci. Soc. Am. J.*, 66(6), 1740–1751, doi:10.2136/sssaj2002.1740, 2002.

Vrugt, J. and Robinson, B.: Improved evolutionary optimization from genetically adaptive multimethod search., *Proc. Natl. Acad. Sci. U. S. A.*, 104(3), 708–711, doi:10.1073/pnas.0610471104, 2007a.

Vrugt, J., ter Braak, C. J. F., Clark, M. P., Hyman, J. M. and Robinson, B.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, *Water Resour. Res.*, 44, 1–52, doi:10.1029/2007WR006720, 2008.

Vrugt, J., ter Braak, C. J. F., Diks, C. G. H., Robinson, B., Hyman, J. M. and Higdon, D.: Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling, *Int. J. Nonlinear Sci. Numer. Simul.*, 10(3), 273–290, doi:10.1515/IJNSNS.2009.10.3.273, 2009a.

Vrugt, J., ter Braak, C., Diks, C. G. and Schoups, G.: Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications, *Adv. Water Resour.*, 51, 457–478, doi:10.1016/j.advwatres.2012.04.002, 2013.

Vrugt, J. A.: Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation, *Environ. Model. Softw.*, 75(March), 273–316, doi:doi.org/10.1016/j.envsoft.2015.08.013, 2016.

Vrugt, J. A. and Robinson, B. A.: Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging, *Water Resour. Res.*, 43(1), 1–15, doi:10.1029/2005WR004838, 2007b.

Vrugt, J. A. and Sadegh, M.: Toward diagnostic model calibration and evaluation: Approximate Bayesian computation, *Water Resour. Res.*, 49(7), 4335–4345, doi:10.1002/wrcr.20354, 2013.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W. and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, 39(8), 1–19, doi:10.1029/2002WR001746, 2003.

Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W. and Verstraten, J. M.: Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, 41(1), 1–17, doi:10.1029/2004WR003059, 2005.

Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V. and Robinson, B. A.: Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, *Stoch. Environ. Res. Risk Assess.*, 23(7), 1011–1026, doi:10.1007/s00477-008-0274-y, 2009b.

Wagener, T. and Kollat, J.: Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox, *Environ. Model. Softw.*, 22(7), 1021–1033, doi:10.1016/j.envsoft.2006.06.017, 2007.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S. and Gupta, H. V.: Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrol. Process.*, 17(2), 455–476, doi:10.1002/hyp.1135, 2003.

Westra, S., Thyer, M., Leonard, M., Kavetski, D. and Lambert, M.: A strategy for diagnosing and interpreting hydrological model non-stationarity, *Water Resour. Res.*, 1–24, doi:10.1002/2013WR014719. Received, 2014.

Wheater, H. S., Jakeman, A. J. and Beven, K. J.: Progress and directions in rainfall-runoff modelling, in *Modelling Change in Environmental Systems*, edited by A. J. Jakeman, M. B. Beck, and M. J. McAllen, pp. 101–132, John Wiley, Chichester., 1993.

Wilks, D. S.: *Statistical methods in the atmospheric sciences.*, 2011.

Wöhling, T. and Vrugt, J. A.: Combining multiobjective optimization and Bayesian model averaging to calibrate forecast ensembles of soil hydraulic models, *Water Resour. Res.*,

44(12), 1–18, doi:10.1029/2008WR007154, 2008.

Yang, J., Reichert, P. and Abbaspour, K. C.: Bayesian un certainty analysis in distributed hydrologic modeling: A case study in the Thur River basin (Switzerland), *Water Resour. Res.*, 43(10), 1–13, doi:10.1029/2006WR005497, 2007.

Yapo, P. O., Gupta, H. V. and Sorooshian, S.: Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data, *J. Hydrol.*, 181(1–4), 23–48, doi:10.1016/0022-1694(95)02918-4, 1996.

Yapo, P. O., Gupta, H. V. and Sorooshian, S.: Multi-objective global optimization for hydrologic models, *J. Hydrol.*, 204(1–4), 83–97, doi:10.1016/S0022-1694(97)00107-8, 1998.

Yen, H., Wang, X., Fontane, D. G., Harmel, R. D. and Arabi, M.: A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling, *Environ. Model. Softw.*, 54, 211–221, doi:10.1016/j.envsoft.2014.01.004, 2014.

Young, P.: Data-based mechanistic modelling of environmental, ecological, economic and engineering systems, *Environ. Model. Softw.*, 13(2), 105–122, doi:10.1016/S1364-8152(98)00011-5, 1998.

Young, P. C.: *Advances in real-time flood forecasting.*, *Philos. Trans. A. Math. Phys. Eng. Sci.*, 360(1796), 1433–1450, doi:10.1098/rsta.2002.1008, 2002.

Young, P. C.: *Recursive Estimation and Time-Series Analysis*, Springer Berlin Heidelberg, Berlin, Heidelberg., 2011.

Young, P. C.: Data-based mechanistic modelling and the simplification of environmental systems, in *Environmental Modelling Finding Simplicity in Complexity*, pp. 371–388, John Wiley & Sons, Chichester., 2013.

Zheng, Y. and Han, F.: Markov Chain Monte Carlo (MCMC) uncertainty analysis for watershed water quality modeling and management, *Stoch. Environ. Res. Risk Assess.*, 30(1), 293–308, doi:10.1007/s00477-015-1091-8, 2016.

Zhou, H., Gómez-Hernández, J. J. and Li, L.: Inverse methods in hydrogeology: Evolution and recent trends, *Adv. Water Resour.*, 63(Liangping Li), 22–37, doi:10.1016/j.advwatres.2013.10.014, 2014.