

UNIVERSIDAD NACIONAL DE CÓRDOBA

Facultad de Ciencias Exactas Físicas y Naturales

Doctorado en Ciencias de la Ingeniería

Tesis Doctoral



Optimización multi-objetivo como herramienta de
decisión frente a diversos escenarios de pronóstico
hidrológico

Autor: Facundo José ALONSO

Director: Prof. Dr. Juan Carlos BERTONI

Julio de 2018

Optimización multi-objetivo como herramienta de decisión frente a diversos escenarios de pronóstico hidrológico

por

Facundo José ALONSO

Prof. Dr. Juan Carlos BERTONI

Director

COMISIÓN ASESORA:

Prof. Dr. Juan Carlos BERTONI

FCEFYN - UNC

Prof. Dr. Andrés RODRÍGUEZ

FCEFYN - UNC

Esta Tesis fue enviada a la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de Córdoba para cumplimentar los requerimientos de obtención del grado académico de Doctor en Ciencias de la Ingeniería.

Córdoba, Argentina
Julio de 2018



ACTA DE EXAMENES

Libro: 00001 Acta: 04361 Hoja 01/01
LLAMADO: 1 20/07/2018
CATEDRA - MESA:

DI002 TESIS DOCTORADO EN CIENCIAS DE LA INGENIERIA

NUMERO	APELLIDO Y NOMBRE	DOCUMENTO INGRESO COND.	NOTA	FIRMA
26180168	ALONSO, FACUNDO JOSE	DNI: 26180168 2013 T	Aprobado	

TEIXEIRA GURBINDO, Luis - FERNANDEZ FERREYRA, Damián - Gironas Leon, Jorge Alfredo - GARCÍA RODR

Observaciones:

Sergio Elaskar
DIRECTOR
Doctorado en Ciencias de
la Ingeniería
F.C.E.F. y N. - U.N.C.

Córdoba, ___/___/___.

Certifico que la/s firma/s que ha/n sido puesta/s en la presente Acta pertenece/n a: Damián Fernández

1 0 1 0 1
Inscriptos Ausentes Examinados Reprobados Aprobados
04/07/2018 09:40:35

*Dedico esta tesis a mi mujer, Sofi
y a mis hijos, Justito, Clarita y Félix...
...por resignar generosamente ese tiempo que no pudimos compartir...*

En la persona de mi tercer hijo, Félix María, que todavía se refugia en la panza de su mamá, dedico también esta tesis a los niños por nacer, y abogo por el buen criterio y coherencia (*) de nuestra querida Patria en la protección de esas pequeñas vidas que luchan silenciosamente por subsistir.

(*) La Organización Meteorológica Mundial (WMO, 2014), en su “Atlas de la mortalidad y las pérdidas económicas provocadas por fenómenos meteorológicos, climáticos e hidrológicos extremos”, señala que entre 1970 y 2012 se registraron 1,94 millones de víctimas mortales ocasionadas por eventos meteorológicos extremos, de los cuales el 55% fueron debido a tormentas y crecidas. Es decir, en el orden de 25400 personas por año a nivel mundial. Por otra parte, según datos de la Organización Mundial de la Salud (WHO, 2018) entre 2010 y 2014 hubo en el mundo un promedio anual de 56 millones de abortos provocados. De donde se deduce que por cada persona que muere por un fenómeno hidrometeorológico extremo, son asesinados 2204 niños en el vientre de su madre.

“...la amenaza más grande que sufre la paz hoy en día es el aborto”

Madre Teresa de Calcuta (Washington D.C., 3 de febrero de 1994)

Agradecimientos

No tengo del todo claro si fueron mis virtudes de la perseverancia y la responsabilidad, o mis defectos de la terquedad y la obstinación que me llevaron a terminar definitivamente la tesis de doctorado. El resultado me encontró en un momento de la vida en el cual dejó de valer más la tesis completa que una sola hora contemplando la sonrisa de algún hijo... Era el momento de concluir o abandonarla definitivamente. Felizmente, prevaleció la primera opción y hoy siento que todo esfuerzo (mío y ajeno) tuvo sentido...

Por eso, aunque tal vez no sea académicamente correcto, quiero empezar por agradecer a mi mujer Sofía, y a mis hijos, Justito, Clarita y Félix, a quienes además dedico la tesis, por regalarme cada hora de sus días, esos días que ya no vuelven, en esas edades que ya no tendrán..., para pagar el precio de alcanzar este objetivo, que lo hicieron propio...

En segundo lugar, agradezco a mis padres, Carlos y Patricia, por sentir la alegría de mi logro, como si fuera de ellos, con el amor que sólo siente quien pone la vida ajena delante de la suya propia. De manera especial, agradezco a mi hermano Carlos y mi cuñada Cati, por estar siempre ahí, dándome su apoyo incondicional, sin el cual no habría podido llegar a la meta.

Agradezco a mi director de tesis, Juan Carlos Bertoni, por haber modelado mi formación, y ¡lo hago responsable por haberme contagiado de la pasión por la hidrología y la optimización! Reconozco su apoyo en la conformación de proyectos conjuntos con los centros de investigación de Francia a partir de los cuales, se terminó de configurar el programa de trabajo que concluyó en la tesis. En este sentido, destaco especialmente las contribuciones de Nicolás Le Moine al desarrollo de la metodología adoptada, y sus aportes a la interpretación de los resultados que permitieron la organización de los capítulos decisivos de esta tesis. ¡Gracias Nico por tu compromiso de colega y amigo! También agradezco al resto del equipo de la Universidad "Pierre et Marie Curie", de París, por su afectuosa acogida durante mis estancias allí, y su generoso aporte técnico y científico, representados principalmente en las personas de los profesores Pierre Ribstein y Ludovic Oudin. Asimismo, destaco un especial reconocimiento a los miembros del Instituto IRSTEA Antony, que reflexionaron conmigo en las múltiples reuniones y encuentros que fortalecieron los objetivos de la tesis: Charles Perrin, María Helena Ramos, Vazken Andréassian y Lionel Berthet. Valoro particularmente la buena predisposición de Damien Lilas por el tiempo dedicado en explicarme los pormenores de programación del modelo Génie Rural, así como a Florent Lobligeois y Andrea Ficchi por brindarme las ricas series de datos de las cuencas francesas. Destaco la asistencia generosa de Antoine Patalano en el armado del proyecto HYRADIER, dentro del programa STIC-AMSUD que fortaleció el vínculo institucional con las contrapartes francesas.

Quiero también agradecer a Andrés Rodríguez, quien sigue y empuja mi recorrido académico desde los orígenes, como director de mi trabajo final de la carrera de grado de Ingeniero Civil, y de quién, como miembro de mi comisión asesora, recibí el auxilio oportuno para lograr el cierre

definitivo de la tesis doctoral. En esta tarea de aportar elementos para mejorar la calidad de la tesis, agradezco la actuación de los miembros del tribunal, quienes leyeron minuciosamente el trabajo, e hicieron aportes de gran valor. Especialmente destaco el dictamen del Dr. Gironás cuya contribución ordenó notablemente la presentación de algunos capítulos.

En un lugar especialmente destacado quiero poner en relevancia el apoyo que recibí de mi querido país, que financió con la beca interna doctoral del CONICET una parte de mi formación, y a través de programas de cooperación internacional, me permitió complementar mis estudios en centros extranjeros de alto prestigio científico internacional. Soy consciente de que este apoyo me imprime una deuda pública a mis compatriotas que con sus impuestos pagaron mi formación, especialmente hacia aquellos que viven en la periferia de estas oportunidades, y para los cuales llegar hasta esta instancia es tal vez un objetivo humanamente imposible.

A los fines de aportar resultados en nuestro país realicé ingentes esfuerzos por aplicar las metodologías desarrolladas en esta tesis a cuencas hidrográficas argentinas. Más de la mitad del tiempo invertido en esta carrera del doctorado se puede atribuir sin duda alguna a la búsqueda de series confiables y procesamiento de datos hidrometeorológicos con paso de tiempo horario en cuencas argentinas. En este sentido, destaco particularmente el apoyo de las siguientes personas durante mi búsqueda: Marcelo García, Carlos Catalini, César García y Laura Colladón, por el río San Antonio (Córdoba); Luis Lenzi y Juan Bianchi, por el río Gualeguay (Entre Ríos); Eduardo Zamanillo, Nicolás Vaiman, Eduardo Díaz y Silvina San Miguel, por el río Gualeguaychú (Entre Ríos); y Sergio Menajovsky, por el río Salado santafesino. También agradezco la colaboración de mis estudiantes en el procesamiento de parte de las series, María José Martín, Nereo Juárez y Nicolás Bustos. Lamentablemente, la calidad de las series de datos conformadas, especialmente para los periodos de caudales máximos, fue absolutamente deficiente, y no permitió realizar experimentos confiables de donde puedan extraerse conclusiones útiles. Por esta razón, me vi forzado, a los fines de darle un cierre a los trabajos, de aplicar los estudios en cuencas francesas.

Quiero agradecer específicamente a mi querida Universidad Nacional de Córdoba, que me brindó gratuitamente sus instalaciones y su ciencia durante toda mi formación. En ella agradezco también a mis compañeros profesores de los departamentos de Hidráulica y Construcciones Civiles, que de un modo afectuoso me manifestaron su genuino sentimiento por verme alcanzar este importante logro, como José Li Gambi, Roberto Terzariol, Santiago Reyna, Viviana Rodríguez, Mariano Corral, Teresa Reyna, Alejandro Maza, Juan Alippi, Magalí Carro Pérez, Rodolfo Giordana, Santiago Pérez Corti, Daniela Villanueva, Leonardo Cocco, Alejo Testa, Gustavo Andrade, Julito Capdevila, Marcos Montoro, Horacio Herrero y Franco Francisca.

Al Instituto Superior de Estudios Ambientales (ISEA), que me ofreció un lugar de trabajo, y al Centro de Estudios y Tecnología del Agua (CETA), que acompañó institucionalmente mis proyectos de cooperación internacional.

No quiero dejar de agradecer a todas las demás personas que enriquecieron mi vida, le sumaron color a los momentos y acompañaron mi crecimiento humano de la mano del profesional, como mis hermanas Eugenia, Jimena y Macarena; mi cuñado y amigo Nacho; mis queridos suegros Fermín y Alicia; al Ichi y la Mer por su apoyo y compañía. A mi ahijada Bernardita y mis demás sobrinos. Finalmente, a mis queridos amigos, por tantos momentos divertidos y recuerdos impagables.

RESUMEN

El acelerado crecimiento demográfico de las poblaciones ubicadas a las márgenes de los ríos, sumado a un mundo con escenarios climáticos futuros cada vez menos predecibles en términos fenomenológicos, conforma una combinación compleja que implica nuevos desafíos en el campo de la prevención de daños contra inundaciones. Los modelos de pronóstico hidrológico constituyen potentes herramientas capaces de predecir con suficiente anticipación los valores de caudales máximos esperados, y el tiempo en el cual se espera que ocurran dichos picos de crecida. Sin embargo, se han identificado dos situaciones típicas de operación que los vuelven vulnerables. En primer lugar, estos modelos se caracterizan por disponer de una estructura altamente dependiente de la información a tiempo real de los valores de caudal medidos en la sección de interés. Ahora bien, durante un evento de gran magnitud, existe un riesgo potencial de falla del sistema de transmisión de datos, produciendo en ocasiones una degradación parcial o una pérdida total de dicha información. En estos eventos, el modelo se vuelve simplemente de simulación, ejercicio para el cual otro modelo (o el mismo parametrizado distinto) podría dar resultados más confiables. En segundo lugar, el complejo entramado de los modernos sistemas de alerta temprana frente a inundaciones requiere la predicción de valores de caudal para distintos tiempos de antelación. En estos casos, el modelo es forzado a emitir valores de caudal para horizontes de pronóstico en los cuales no fue calibrado. En síntesis, no existe en la actualidad una estrategia de adaptación de los modelos que les permita operar en condiciones distintas para los cuales fueron calibrados. En esta tesis se aporta una herramienta que permite al operador del modelo de pronóstico adaptar su funcionamiento a diversos escenarios típicos de operación no contemplados a partir de la calibración convencional. Para ello, se desarrolló una versión mejorada del algoritmo MOCOM-UA de optimización multi-objetivo como metodología de confrontación de los diversos escenarios alternativos, lo cual permitió ofrecer soluciones de compromiso capaces de abordar a un mismo tiempo más de una situación posible. Los experimentos fueron llevados a cabo para la pequeña cuenca del río Moros al noroeste de Francia de sólo 20,97 km². Los resultados mostraron que, para un gran número de casos, existe un juego de valores para los parámetros capaz de optimizar simultáneamente ambos escenarios confrontados, volviendo la optimización muti-objetivo no estrictamente necesaria. Sin embargo, para el resto de los casos, la determinación del óptimo para cada escenario es decisiva en relación al éxito del valor estimado, convirtiendo al frente de Pareto en una herramienta fundamental para la toma de decisiones del operador del sistema de pronóstico hidrológico. Finalmente, el método MOCOM-UA adaptado en esta tesis, demostró mayor robustez en alcanzar en los sucesivos experimentos el mismo frente de Pareto, por lo cual se recomienda su uso, en lugar del método original, para los experimentos de esta naturaleza.

ABSTRACT

The accelerated demographic growth of the populations located at the river banks, coupled with a world with future climatic scenarios every time less predictable from a phenomenological point of view, forms a complex combination that involves new challenges in the field of flood damage prevention. The flood forecasting models are powerful tools capable of predicting with sufficient anticipation the values of expected maximum flow rates, and the time in which such flood peaks are expected to occur. However, two typical operating situations have been identified that make them vulnerable. First, these models are characterized by having a structure highly dependent on the real-time information of the discharge values measured in the section of interest. However, during a major event, there is a potential risk of the data transmission system failure, sometimes resulting in partial degradation or a total loss of such information. In these events, the model simply becomes a simulation model, exercise for which another model (or the same model with another set of parameters) could give more reliable results. Secondly, the complex structure of modern flood warning systems requires the prediction of flow values for different times in advance. In these cases, the model is forced to emit discharge values for forecast horizons in which it was not calibrated. In summary, there is currently no strategy for adapting the models in order to allow them to operate under different conditions with respect to which they were calibrated. This thesis provides a tool that allows the operator of the forecast model to adapt its operation to different typical scenarios of operation not contemplated from the conventional calibration. To this end, an improved version of the multi-objective optimization algorithm, MOCOM-UA, was developed as a methodology for confronting the various alternative scenarios, which allowed us to offer compromise solutions capable of addressing at the same time more than one possible situation. The experiments were carried out for the small basin of the Moros river to the northwest of France of only 20.97 km². The results showed that for a large number of cases, there is a set of values for the parameters capable of simultaneously optimizing both confronted scenarios, making multi-objective optimization not strictly necessary. However, for the rest of the cases, the determination of the optimum for each scenario is decisive in relation to the success of the estimated value, turning the Pareto front into a fundamental tool for the decision making of the operator of the hydrological forecasting system. Finally, the MOCOM-AU method, adapted in this thesis, showed a greater robustness in reaching the same Pareto front in successive experiments, so its use, instead of the original method, is recommended for experiments of this nature.

RÉSUMÉ

La croissance démographique accélérée des populations situées au bord des fleuves, ajoutée à un monde aux scénarios climatiques futurs de moins en moins prévus, forment une combinaison complexe qui implique de nouveaux défis dans le domaine de la prévention des dommages causés par les inondations. Les modèles de prévision hydrologique sont des outils puissants capables de prédire à l'avance les valeurs des débits maximaux attendus et le temps pendant lequel ces pics sont attendus. Deux situations d'exploitation typiques les rendant vulnérables ont toutefois été identifiées. En premier lieu, ces modèles se caractérisent par une structure fortement dépendante de l'information en temps réel des valeurs de débit mesurées dans la section d'intérêt. Cependant, lors d'un événement de grande ampleur, il existe un risque potentiel de défaillance du système de transmission de données, qui peut parfois entraîner une dégradation partielle ou une perte totale de ces informations. Dans ces cas, le modèle devient simplement un modèle de simulation, un exercice pour lequel un autre modèle (ou le même paramétré distinct) pourrait donner des résultats plus fiables. Deuxièmement, le réseau complexe de systèmes modernes d'alerte rapide en cas d'inondation nécessite la prévision préalable des valeurs de débit pour différentes périodes. Dans ces cas, le modèle est obligé d'émettre des valeurs de débit pour des horizons de prévision dans lesquels il n'a pas été calé. En résumé, il n'existe actuellement aucune stratégie pour adapter les modèles qui leur permettent de fonctionner dans différentes conditions pour lesquelles ils ont été calés. Cette thèse fournit un outil qui permet à l'opérateur du modèle de prévision d'adapter son fonctionnement à divers scénarios de fonctionnement typiques non prises en compte par le calage conventionnel. À cette fin, une version améliorée de l'algorithme d'optimisation multi-objectifs MOCOM-UA a été développée en tant que méthodologie permettant de comparer les différents scénarios alternatifs, ce qui permettait d'offrir des solutions de compromis capables de gérer simultanément plusieurs situations possibles. Les expériences ont été réalisées pour le petit bassin versant de la rivière Moros, dans le nord-ouest de la France, de 20,97 km² seulement. Les résultats ont montré que, dans un grand nombre de cas, il existe un ensemble de valeurs pour les paramètres capables d'optimiser simultanément les deux scénarios confrontés, rendant l'optimisation multiobjectif non strictement nécessaire. Cependant, pour le reste des cas, la détermination de l'optimum pour chaque scénario est déterminante pour le succès de la valeur estimée, faisant du front de Pareto un outil fondamental pour la prise de décision de l'opérateur du système de prévision hydrologique. Enfin, la méthode MOCOM-UA adaptée à cette thèse a montré une plus grande robustesse pour atteindre dans les expériences successives le même front de Pareto, pour lequel son utilisation est recommandée, au lieu de la méthode originale, pour des expériences de cette nature.

ESTRUCTURA DE LA TESIS

Resumen

PARTE 1: INTRODUCCIÓN

- Cap. 1: Contexto
- Cap. 2: Problemática
- Cap. 3: Objetivos y organización de la tesis

PARTE 2: REVISIÓN DE ANTECEDENTES

- Cap. 4: Modelación lluvia-caudal (P-Q)
- Cap. 5: Pronóstico hidrológico
- Cap. 6: Optimización multi-objetivo
- Cap. 7: Criterios de ajuste y performance

PARTE 3: METODOLOGÍA Y DATOS

- Cap. 8: El modelo de pronóstico hidrológico GRP
- Cap. 9: El algoritmo de optimización multi-objetivo MOCOM-UA
- Cap. 10: Descripción de la cuenca y series de datos disponibles
- Cap. 11: Propuesta metodológica

PARTE 4: RESULTADOS

- Cap. 12: Adaptación del Método MOCOM-UA para mejorar su robustez.
- Cap. 13: Evaluación cuantitativa de la mejora de la performance del Q a tiempo real.
- Cap. 14: Desempeño del modelo frente a distintos horizontes de pronóstico.

- Cap. 15: Conclusión General

- Cap. 16: Referencias bibliográficas

INDICE

1.	Contexto	27
1.1.	Inundaciones: un problema creciente	28
1.2.	¿Qué beneficios ofrece buen pronóstico hidrológico?	32
1.3.	¿Qué riesgos acarrea un deficiente pronóstico hidrológico?	33
1.4.	¿Qué ocurre en la realidad?	34
2.	Problemática	37
2.1.	Marco conceptual de la tesis	37
2.2.	El interrogante científico	38
3.	Objetivos y organización de la tesis	45
4.	Modelación lluvia-caudal (P-Q)	51
4.1.	Sistemas, conceptos y clasificación	51
4.2.	El sistema (ciclo) hidrológico	53
4.3.	El proceso de transformación lluvia-caudal	54
4.4.	¿Qué es un modelo y para qué sirve?	57
4.5.	Una clasificación de los modelos P-Q	60
4.6.	Evolución de los modelos P-Q	62
4.7.	Modelos "a reservorios" P-Q	71
4.8.	Aspectos complementarios de modelación	74
4.8.1.	Variables y parámetros	74
4.8.2.	Memoria del modelo, condiciones iniciales y periodo de calentamiento	76
5.	Pronóstico hidrológico	81
5.1.	Los sistemas de Alerta temprana	81
5.2.	Diversas metodologías de pronóstico	82
5.3.	Respecto del escenario futuro de precipitaciones	84
5.4.	Modelos P-Q de Pronóstico hidrológico	85
5.5.	Respecto de las incertidumbres	89
6.	Optimización multi-objetivo	91
6.1.	Algoritmos genéticos	94
6.1.1.	Introducción a los Algoritmos Genéticos	95

6.1.2.	Conceptos básicos de biología	99
6.1.1.	La técnica de los Algoritmos Genéticos.....	105
6.2.	Técnicas de calibración.....	109
6.2.1.	Clasificación de las técnicas de calibración	110
6.2.1.	Técnicas de optimización global.....	119
6.3.	Evolución de los algoritmos de optimización para modelos P-Q.....	121
7.	Criterios de ajuste y evaluación de la performance del modelo.....	133
7.1.	Restricciones.....	135
7.2.	Criterios de ajuste basados en el error de simulación	135
7.3.	Criterios de performance	140
8.	El modelo de pronóstico hidrológico GRP.....	153
8.1.	Evolución de los modelos GR	153
8.2.	Estructura del modelo GR4P	156
8.2.1.	Aspectos de la estructura que adaptan el modelo al paso de tiempo diario y horario	164
8.2.1.	Algoritmo de actualización de estado vinculado al pronóstico hidrológico	166
9.	El algoritmo de optimización multi-objetivo MOCOM-UA.....	169
9.1.	Descripción del método	170
9.1.1.	Generación de la población inicial	173
9.1.2.	Clasificación y ordenamiento de los individuos de la población.....	176
9.1.3.	Selección basada en el ranking.....	178
9.1.4.	Evolución multi-objetivo compleja.....	179
9.1.5.	Evolución Multi-Objetivo del Simplex (MOSIM).....	181
9.1.6.	Un esquema del proceso evolutivo del MOCOM-UA.....	182
10.	Descripción de la cuenca y series de datos disponibles.....	185
11.	Propuesta metodológica	189
11.1.	Análisis retrospectivo	191
11.2.	Periodo de adecuación o "calentamiento" del estado del modelo	192
11.3.	Criterio de partición de las series de datos	193
11.4.	Procedimiento de calibración y control multi-objetivo.....	194
11.5.	El criterio de ajuste.....	196

11.6.	Respecto de los límites de los parámetros del modelo GR.....	197
12.	Adaptación del Método MOCOM-UA para mejorar su robustez.....	201
12.1.	Problemática	201
12.2.	Funciones objetivo y criterio de evaluación del ajuste	202
12.3.	Optimización del modelo GR4P	203
12.4.	Modificaciones al MOCOM-UA para mejorar la robustez	206
12.4.1.	Población inicial mejorada	206
12.4.2.	Adimensionalización de parámetros.....	207
12.4.3.	Búsqueda local mejorada	208
12.4.4.	Condición de parada	210
12.5.	Resultados y discusiones.....	212
12.6.	Conclusiones	213
13.	Evaluación cuantitativa de la mejora de la performance del Q a tiempo real.	215
13.1.	Problemática	215
13.2.	Funciones objetivo y criterio de evaluación del ajuste	217
13.3.	Optimización del modelo GR4P	218
13.4.	Síntesis de los resultados y discusiones	223
13.5.	Conclusiones	224
14.	Desempeño del modelo frente a distintos horizontes de pronóstico.	227
14.1.	Problemática	227
14.2.	Funciones objetivo y criterio de evaluación del ajuste	228
14.3.	Optimización del modelo GR4P	228
14.4.	Síntesis de los resultados y discusiones	240
14.5.	Conclusiones	241
15.	Conclusión General	243
16.	Referencias bibliográficas	249

PARTE 1: INTRODUCCIÓN

En esta primera parte introductoria se presenta el contexto general del fenómeno de las inundaciones, abordado con valores numéricos que grafican el impacto que generan en las distintas regiones del planeta, y en particular en la Argentina. Asimismo, se vincula la herramienta del pronóstico hidrológico como soporte preventivo de este tipo de manifestaciones hidroclimáticas. Y en ese contexto, se inserta la problemática de esta tesis orientada a enriquecer los elementos disponibles para la toma de decisiones. A partir de estos elementos, se cierra la primera parte de la tesis con una definición del interrogante científico que le dio nacimiento, y se plantean sus objetivos.

1. Contexto

Este capítulo presenta el marco contextual dentro del cual tiene lugar la presente tesis. Describe desde el punto de vista práctico cuál es la problemática que le da origen, y cuantifica numéricamente la gravedad del problema.

A menudo digo que cuando puedes medir lo que estás hablando y expresarlo en números, sabes algo al respecto; pero cuando no puedes medirlo, cuando no puedes expresarlo en números, tu conocimiento es escaso e insatisfactorio.

Lord Kelvin

El planeta Tierra se encuentra en constante evolución, configurando de manera infatigable las formas de las montañas y los cursos de agua, alterando día tras día las composiciones químicas de las sustancias que lo integran, interactuando incansablemente con la atmósfera que lo cubre y lo protege. Existen innumerables procesos de flujo de materia y energía entre los diversos ambientes geológicos, biológicos e hidrológicos que lo componen. El hombre, frente a la necesidad de adecuarse a un planeta cambiante, ha procurado interpretar su comportamiento, y sobre esa base propone un conjunto de herramientas que le permiten estimar la evolución del mismo.

La comprensión gradual de los procesos evolutivos y el comportamiento de los diversos sistemas que se interrelacionan en el globo terráqueo, permite a la humanidad ir generando líneas de acción y estableciendo las bases para la toma de decisión frente a diversos escenarios presentes o probables hacia el futuro. Dichas líneas de acción se van perfeccionando de generación en generación a medida que es posible ampliar la comprensión (¡siempre limitada!) que el hombre posee de su

entorno. En este sentido, y bajo esa premisa, se considera que todo aporte y contribución a ampliar el conocimiento del medio habitable por la humanidad, así como el refinamiento de las herramientas disponibles, es un paso adelante en dirección al mejoramiento de la calidad de vida de las personas. En particular, la coexistencia de la civilización y los sistemas fluviales desde los orígenes de los tiempos, pone al descubierto la fragilidad y vulnerabilidad humana frente a las fuerzas de la naturaleza. Las crecidas repentinas de los ríos, y las inundaciones de las ciudades conforman una amenaza permanente en este tipo de ecosistemas.

En ese contexto, emerge la necesidad siempre latente de anticiparse al fenómeno natural a partir de herramientas competentes. En nuestros días, los sistemas de alerta temprana vinculados a los modelos de pronóstico hidrológico son el medio adecuado para abordar esta casuística, por lo que todo esfuerzo por perfeccionar su desempeño es un nuevo aporte a la calidad de vida de las civilizaciones.

1.1. Inundaciones: un problema creciente

Se denomina comúnmente inundación al fenómeno por el cual una parte de la superficie terrestre queda cubierta temporalmente por el agua, ante una subida extraordinaria del nivel de ésta (Bertoni, 2004a). El fenómeno de la inundación en sí mismo está previsto en el ciclo natural de los niveles de los ríos, y en muchos ecosistemas forma parte de una condición cíclica que permite el equilibrio de las comunidades de especies que los conforman.

Conforme se grafica en Bertoni (2004a), el cauce de un curso de agua está compuesto por el lecho menor y el lecho mayor (Figura 1-1). El lecho menor es aquel ocupado totalmente por el río con un período de retorno que oscila entre 2 y 5 años. Tal condición se denomina de "márgenes llenas" ("bankfull" en la literatura de lengua inglesa). La crecida máxima anual (u ordinaria) es contenida totalmente por el lecho menor. El lecho mayor de los ríos, comúnmente llamado "llanura de

inundación”, es ocupado por las aguas con una recurrencia (o tiempo de retorno) mayor.

Dado el fluctuante comportamiento que manifiestan los niveles de los cursos de agua a lo largo de su historia en función de las también cambiantes condiciones climáticas a las cuales se ven sometidas sus cuencas hidrográficas, las llanuras de inundación son tierras normalmente secas pero ricas orgánicamente. El hombre, en una tendencia natural por ubicarse próximo a la fuente de abastecimiento de agua, ha ido propiciando a lo largo de su historia la conformación de asentamientos y comunidades a las márgenes de los ríos. Este proceso de “urbanización” es favorecido por las excelentes condiciones de siembra que estas tierras en general manifiestan.

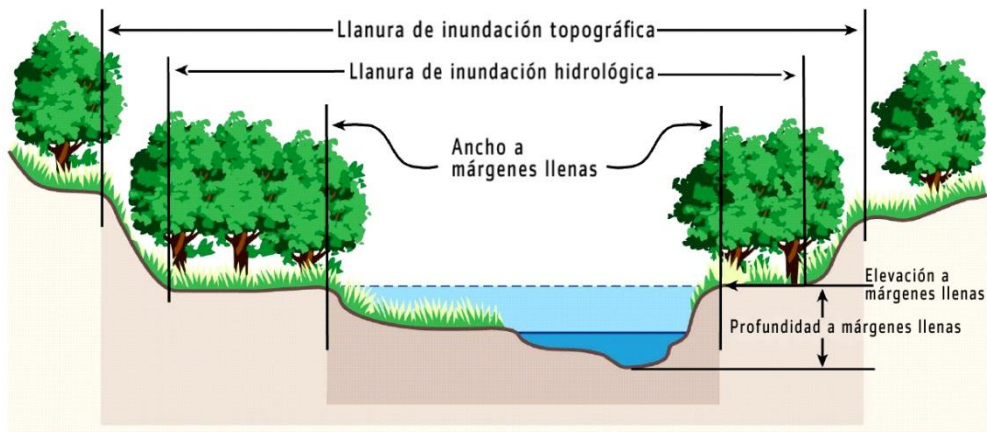


Figura 1-1: Cauce y llanuras de inundación en ríos (Fuente: Bertoni, 2004a)

Ahora bien, cuando el hombre se asienta sobre las márgenes de los ríos, ocupando periódicos o temporales cauces naturales, y el agua vuelve ocasionalmente a reclamar el lugar que le pertenece, produce una afectación directa sobre las personas y sus bienes. Es decir que la inundación es concebida como un desastre cuando involucra daños materiales sobre los bienes o las personas.

Las inundaciones figuran en primer lugar dentro de las catástrofes naturales en el mundo, causando alrededor de 20.000 muertes por año (Berthet, 2010). Al respecto

son alarmantes las estadísticas de la OMM (2009), según las cuales las inundaciones y crecidas en los ríos son responsables del 54% de los daños significativos, del 65% de las personas afectadas y del 29% de las muertes humanas causadas por desastres naturales. Las inundaciones fueron responsables del 50% de los 2200 desastres hidrometeorológicos computados a nivel mundial en el periodo 1990 – 2001, y del 15% de los decesos ocasionados por desastres naturales. Aproximadamente 66 millones de personas sufrieron daños causados por las inundaciones entre 1973 y 1997 (OMM, 2009).

Según la Agencia Europea del Ambiente, entre 1998 y 2007, las inundaciones sobrevenidas en Europa han causado alrededor de 700 muertes, medio millón de personas evacuadas y al menos 25.000 millones de euros de pérdidas económicas cubiertas por las aseguradoras de riesgo. Francia, por ejemplo, cuenta con unas 13.000 comunas inundables y más de 2 millones de personas viviendo en zonas inundables. Se han registrado en Francia 72 muertos debido a las diez más importantes crecidas de estos últimos diez años. El 80% del costo de daños imputables a los riesgos naturales son causados por las inundaciones, que corresponde en promedio a unos 250 millones de euros por año (Berthet, 2010).

Entre 1987 y 1997 las inundaciones causaron en Asia pérdidas económicas del orden de los 136.000 millones de dólares. Las pérdidas económicas derivadas de las catástrofes naturales han aumentado de 30.000 a 70.000 millones de dólares de los Estados Unidos entre 1990 y 1999. Estas cifras subestiman la verdadera magnitud de las pérdidas, que son probablemente dos o más veces mayores (UNESCO, 2003). Durante las últimas cuatro décadas, las inundaciones han causado en Brasil más de 4,3 mil millones de dólares en daños materiales y se llevaron consigo la vida de centenas de personas (VAISALA, 2012).

Con respecto a nuestro país, Bertoni (2004a) indica que, de acuerdo al Banco Mundial, Argentina, dentro del contexto mundial, se encuentra entre los 14 países

más afectados por catástrofes de inundaciones rurales y/o urbanas, alcanzando pérdidas superiores al 1,1 % del PBI nacional. La frecuencia de ocurrencia de afectaciones importantes por inundaciones de todo tipo es calificada como elevada, siendo superior a un evento importante cada diez años. Desde 1957 han ocurrido 12 episodios severos (uno cada cuatro años), ocasionando muertes y elevadas pérdidas en infraestructura, producción agropecuaria, bienes privados y actividades económicas. Diversos factores explican la creciente incidencia de inundaciones catastróficas, entre ellos, el crecimiento de las poblaciones, la densa ocupación de llanuras aluviales y otras áreas propensas a inundarse, y la ocupación imprudente de terrenos en cauces de agua (WMO, 2009).

Como se ha expresado previamente, el crecimiento no planificado de las ciudades muchas veces se orienta hacia las márgenes de los ríos. Por esta razón existe normalmente una correlación entre el crecimiento poblacional y la vulnerabilidad de las ciudades frente a las inundaciones. Pengue (2012) señala que actualmente más del 50 % de la población mundial vive en ciudades, y que en las próximas cuatro décadas el total del crecimiento de la población mundial será absorbido por áreas urbanas. Este proceso migratorio, viene acompañado de un incremento en la vulnerabilidad de las poblaciones próximas a los ríos. En González (2009) se presentan datos del Programa Mundial de Evaluación de los Recursos Hídricos (en inglés, el World Water Assessment Programme – WWAP) en 2006, de la UNESCO, donde se observa el vertiginoso crecimiento de los desastres por inundaciones a nivel mundial, entre 1960 y 2004. En el periodo 1960-1964 ocurrieron menos de cien desastres por inundación en el mundo, y entre 2000 y 2004, casi ochocientos (CRED, 2012).

Bertoni (2004a) apunta que América del Sur encabeza la tendencia de urbanización a nivel mundial, y que, en este continente, la concentración urbana en las principales ciudades tiene el ritmo de crecimiento más alto entre las regiones del mundo y el

mayor de todos los tiempos. Argentina particularmente posee el 90 % de su población localizada en áreas urbanas, por lo cual se ubica entre los países que detentan la concentración urbana más alta de América del Sur (Tucci, 2006).

De acuerdo a cifras de DesInventar (www.desinventar.org/es, 2004) en Argentina durante el periodo 1970-2004, las inundaciones superaron en número la suma de todos los otros tipos de desastres relacionados con el agua.

1.2. ¿Qué beneficios ofrece buen pronóstico hidrológico?

Si bien las medidas de protección estructural (obras) son capaces de atenuar los daños causados por eventos extremos, están lejos de poder preservar integralmente las poblaciones y las infraestructuras humanas de los efectos de las inundaciones (Thielen et al., 2009). Adicionalmente, dentro del cambio climático global que experimenta el planeta, se producen alteraciones en los patrones de precipitaciones que requerirían una adecuación permanente de dichas obras. Por otro lado, un diseño inadecuado o incompleto de un sistema estructural de protección puede volverse una trampa mortal y altamente destructiva, como ocurrió con el caso de la catastrófica inundación de Santa Fe en abril del 2003, debida a la crecida del río Salado (Bertoni, 2004b).

A pesar de que es imposible impedir las inundaciones, o cualquier otro desastre, la capacidad de acción y reacción de los servicios de emergencia ha mejorado notablemente (UNESCO, 2003). Los sistemas de pronóstico hidrológico constituyen actualmente importantes herramientas para reducir, al menos parcialmente, las pérdidas humanas y materiales frente a la ocurrencia de una crecida de gran magnitud. Cada vez más, los Sistemas de Alerta Temprana (SAT) están siendo reconocidos en los más altos niveles políticos como una herramienta crítica para salvar vidas y bienes (UCAR, 2012).

En Estados Unidos, gracias a la información brindada por el Sistema de Pronóstico Hidrológico del NWS (National Weather Service) para el manejo de embalses al USACE (U.S. Army Corps of Engineers) y al USBR (U.S. Bureau of Reclamation), se estima un ahorro anual en promedio de más de 1000 millones de dólares (NHWC, 2002), sólo en bienes materiales.

Krzysztofowicz y Davis (1984), muestran que el incremento en el tiempo de horizonte de pronóstico de un valor preciso, redundará en una reducción del daño.

1.3. ¿Qué riesgos acarrea un deficiente pronóstico hidrológico?

Un modelo de pronóstico hidrológico, que alimenta un SAT puede volverse gradualmente un arma mortal si no funciona con aceptable precisión, o bien, no se utiliza adecuadamente. En efecto, un SAT constituye una estructura metodológica aplicada en una sociedad que debe aprender gradualmente a interpretarla.

Si el modelo de pronóstico anuncia valores considerablemente inferiores a los que ocurre luego en la realidad, y se emite una alarma que subestima los niveles de inundación alcanzados luego por el fenómeno real, podría afectar a zonas desprevenidas que ven llegar la inundación a sus hogares sin tomar recaudo alguno.

El mensaje inverso también es peligroso, puesto que, si se anuncia un nivel superior al real de manera sistemática por el SAT, la población se vuelve gradualmente escéptica a las señales de alerta. En consecuencia, luego de varios procesos fallidos de desalojo de sectores supuestos vulnerables, se irá gestando una creciente resistencia por parte de la sociedad de actuar en consecuencia de la alerta, con el peligro potencial permanente de que en algún episodio se cumpla el pronóstico.

No se trata pues de exigirle al modelo de pronóstico hidrológico que emita una señal exacta del nivel que tomará el cauce en determinada sección, sino de mantener dentro de cierto rango aceptable de error sus valores pronosticados.

1.4. ¿Qué ocurre en la realidad?

Un sistema de pronóstico hidrológico está compuesto por un conjunto de elementos necesarios para garantizar su adecuado funcionamiento. Pero una inundación es, en ocasiones, un episodio hidrometeorológico extremo, por lo cual aumentan las probabilidades de ocurrencia de imprevistos, como la falla o destrucción de parte del sistema de colección y transmisión de información.

En Espinosa (2012) se presenta la crecida máxima histórica correspondiente al día 8 diciembre de 2010 en el río Chagres, que es el principal río tributario del Lago Gatun, el cual integra el sistema del Canal de Panamá. Durante la crecida quedó inhabilitada la estación limnigráfica “Chico”, en la parte superior de la cuenca. Esta estación transmite información a tiempo real al modelo de pronóstico hidrológico, por lo cual se perdió totalmente la capacidad de pronóstico del sistema, provocando un elevado grado de incertidumbre en las decisiones sobre la operación del embalse, que es una tarea crítica en la seguridad de las estructuras que forman el canal. Como parte de esto, durante el paso de esta crecida por el embalse, fue inhabilitada la operación del canal de navegación, con los inmensos daños económicos que eso implica.

El 19 de octubre de 2012, un fenómeno tormentoso excepcional en la cabecera del cauce del Arba de Riguel, en la provincia de Zaragoza (España), ha generado una avenida extraordinaria con un periodo de retorno de 100 años (Heraldo.es, 2012). La misma, arrasó la estación que mide los niveles instalada en Sádaba, dejando de ofrecer información a tiempo real al sistema de Alerta Temprana de la región. Adicionalmente, existe también un factor de falta de mantenimiento y conservación de los elementos del sistema, y dado que los episodios de las crecidas más destructivas son excepcionales, es altamente probable que cuando se produce la inundación, algunas de las partes del sistema no hayan sido verificadas y estén funcionando mal, o fuera de funcionamiento. Por esta razón, es factible suponer que el modelo hidrológico pueda quedarse en determinado momento del episodio

lluvioso, sin información de niveles en la sección de interés, y en ese caso, deberá continuar con su pronóstico con tanta precisión como se lo permita la escasa cantidad o calidad de datos disponibles.

2. Problemática

En este capítulo, se redefine el problema planteado en el capítulo anterior, pero desde un punto de vista científico. Se incursiona en la génesis del problema, y los elementos disponibles que se consideran aptos para abordarlo. De allí se deduce y se plantea formalmente el interrogante científico que se pretende responder. La complejidad del problema admite como axioma “a priori” que no habrá una solución única, sino que el análisis estará dado en los términos de un conjunto de soluciones posibles e incertidumbres asociadas a ellas.

Cuando los fenómenos son tan complejos, los pronósticos no pueden señalar nunca un solo camino, y se puede incurrir en el error de esperar consecuencias demasiado rápidas e inevitables de causas que acaso no son todas las aplicables al problema.

John Maynard Keynes

2.1. Marco conceptual de la tesis

Ha quedado plasmado previamente el tenor de los impactos que las inundaciones producen en nuestro planeta. También se ha puesto en evidencia la importancia de los modelos de pronóstico hidrológico como herramienta operativa para salvaguardar las personas y los bienes de una buena parte de los daños que pueden ocasionar estos fenómenos. Por otra parte, se ha remarcado la importancia de contar con un modelo confiable en términos del impacto en la respuesta social que generan imprecisiones en los valores pronosticados. Finalmente, se destacaron algunas fallas imprevistas que en un sistema operativo real pueden manifestarse y poner en riesgo la capacidad del modelo de emitir valores con aceptable precisión.

En síntesis, el desafío en la construcción y operación de los modelos de pronóstico hidrológico radica en contar con una herramienta aceptablemente precisa aun frente a las adversidades en la disponibilidad de información a tiempo real.

Esta tesis se ubica en el marco de los estudios orientados a la reducción de incertidumbre asociada al pronóstico hidrológico de los modelos lluvia-caudal, como herramienta de un sistema de alerta temprana contra inundaciones.

2.2. El interrogante científico

En el capítulo relativo a la definición y descripción de los modelos de pronóstico hidrológico, es tratada con profundidad la cuestión relativa a los componentes de la estructura interna de este tipo de modelos. Pero a los fines de presentar la motivación de esta tesis en este capítulo, es necesario anticipar que los modelos de pronóstico hidrológico, en general, son altamente dependientes de la información a tiempo real de los caudales producidos en la sección de interés. Es decir, en la propia sección en la cual nos interesa pronosticar el caudal futuro, es menester contar con un instrumento de medición capaz de ofrecer un valor de nivel, traducido a caudal, que en tiempo real comunica al modelo para ser incorporado por este. Como se explica entonces en el capítulo referido, el caudal que está ocurriendo en la sección de interés es normalmente un dato clave de entrada al modelo para el éxito del pronóstico que se pretende realizar.

Esto es debido a que ese valor permite adecuar el estado interno del modelo a las condiciones existentes, y prepararlo para una nueva emisión. Por lo tanto, para maximizar la calidad del pronóstico en el horizonte elegido, es deseable partir del mejor estado posible en el instante de la emisión, y es por lo tanto conveniente “re-ajustar” el modelo, si es posible, sobre la base de los últimos caudales observados a lo largo de algunos pasos de tiempo precedentes. Este procedimiento es descrito esquemáticamente en la Figura 2-1.

En efecto, un modelo de pronóstico se diferencia de un modelo de simulación, en que el primero actualiza su estado periódicamente durante el proceso de pronóstico, lo cual le confiere normalmente mayor precisión.

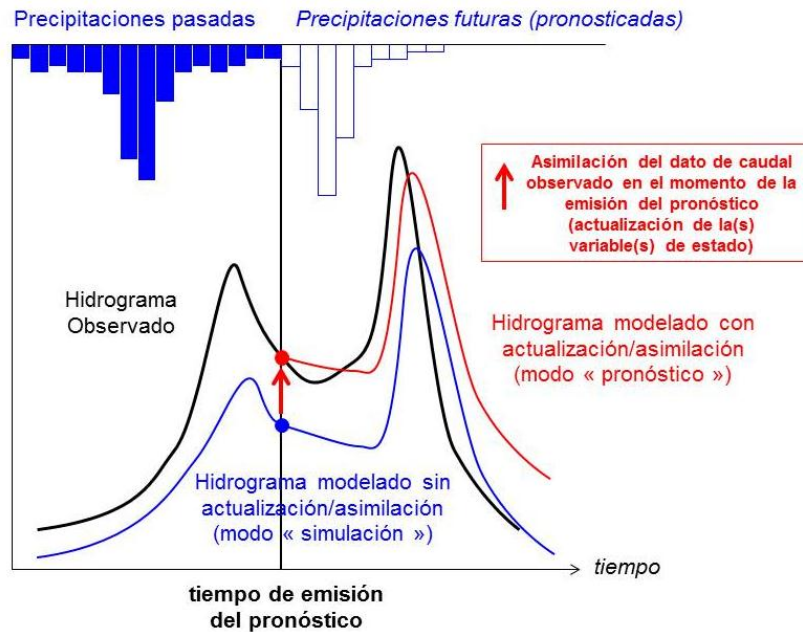


Figura 2-1: Hidrograma simulado y pronosticado

A raíz de esto, si la información a tiempo real de los últimos datos de caudales observados son perdidos (daño en la estación hidrométrica, en la red de comunicación, etc.) y no puede ser más asimilado, el modelo hidrológico se vuelve en una situación típica de simulación, ejercicio para el cual otro modelo (o el mismo modelo, pero parametrizado distinto) podría dar resultados más confiables.

En la calibración típica de modelos hidrológicos, usualmente no se tienen en cuenta escenarios no convencionales de operación, a pesar de tener alguna probabilidad de ocurrencia, como por ejemplo el caso descrito. En consecuencia, si estos escenarios se presentan, normalmente la calidad del pronóstico cae abruptamente.

Otro aspecto que cobra especial importancia en fase operativa durante el proceso de pronóstico, está vinculado a la necesidad por parte de distintos sectores de la población de conocer los valores pronosticados con distinto tiempo de antelación.

Por ejemplo, el órgano de ejecución de las acciones de evacuación de viviendas de la parte de la población potencialmente afectada por la crecida prevista, tendrá necesidad de conocer con mayor antelación la estimación de los niveles que los responsables de evacuar a los turistas de las márgenes de los ríos. Esto es debido a que este último proceso referido es relativamente rápido, mientras que la migración de un sector de la población a otra parte más segura, implica un aparato logístico complicado y con mayor grado de complejidad, lo cual se traduce en mayores tiempos de ejecución. En estos casos, normalmente el modelo es forzado a emitir pronósticos fuera del horizonte para el cual fue calibrado.

Como puede verse, en un proceso operativo de pronóstico hay diversas condiciones reinantes que configuran diversos escenarios de decisión.

En estos casos, nos preguntamos si ¿existe alguna manera de adaptar “a tiempo real” el modelo de pronóstico para que se adecue a estas nuevas condiciones de operación, para las cuales debe seguir emitiendo valores pronosticados confiables?

Dado que son los parámetros quienes juegan un rol fundamental en la “adaptación” del modelo a las distintas situaciones, el primer interrogante científico que proponemos es:

¿Es posible controlar en tiempo real, mediante las combinaciones de valores paramétricos del modelo, su adaptabilidad a los diversos escenarios de pronóstico?

Detrás de este interrogante, subsiste una cuestión encubierta, y es el hecho de que normalmente los modelos de pronóstico hidrológico (como por ejemplo el modelo GRP) tienen una estructura mucho más sencilla que los de simulación hidrológica simplemente (como el modelo GR4J). Esto es porque los modelos de pronóstico pueden prescindir de varios procesos simulados por un modelo de simulación, ya que a la hora del pronóstico el componente de su estructura que tiene mayor

importancia (y al cual mayor peso le atribuye el modelador), es al que asimila el dato de caudal observado a tiempo real. En otras palabras, un modelo de pronóstico hidrológico es altamente dependiente de este valor, por lo tanto, en la mayor parte de los casos, si se quedara sin ese valor durante el pronóstico, no habría manera de lograr un buen desempeño aún con un buen ajuste de sus parámetros.

Por esta razón, si el propósito es mejorar el pronóstico del caudal a un cierto horizonte, aún sin el dato de caudal observado a tiempo real, se debería emplear un modelo de simulación (como el GR4J) calibrado para realizar pronósticos. Si ese fuera el caso, la “adaptación” al nuevo escenario de pronóstico implicaría cambiar no sólo de parámetros, sino también de modelo.

A los fines de simplificar la propuesta operativa, en esta tesis se trabajó con los modelos hidrológicos de la familia Génie Rural (GR). Dentro de la misma, se optó por conservar la estructura completa del modelo de simulación GR4J, y se incorporó sólo el algoritmo de actualización de estado del modelo de pronóstico GRP, con el propósito de convertir el modelo de simulación en un modelo de pronóstico. Así se obtuvo el modelo GR4P, propuesto en esta tesis, todo lo cual fue programado en el lenguaje de programación FORTRAN 90/95. En el capítulo referido a la definición y descripción del modelo GR4P, se explica que el modelo GRP es la versión de pronóstico de la familia de modelos GR, siendo su estructura más sencilla que la versión de simulación GR4J. Es decir, conserva elementos básicos de su estructura y le da mucho peso al valor de caudal a tiempo real a través de un mecanismo de actualización de estado, volviéndose muy dependiente de este último. Entonces, la explicación hecha previamente en este apartado aclara que se adopta para este trabajo el modelo GR4J, o sea la versión más completa de la estructura de los modelos GR, pero se le incorporaron los mecanismos de actualización de estado que tiene GRP. Así se da nacimiento a una versión denominada GR4P.

Por todo lo dicho, para responder al primer interrogante científico arriba expresado, en primer lugar, nos aseguramos de que el modelo de pronóstico elegido para parametrizar conserve también la estructura de un buen modelo de simulación.

Para no incurrir en un problema semántico, en el informe de tesis nos referimos usualmente al modelo GR4P funcionando en “modo pronóstico” o en “modo simulación”, pero en el fondo es como si fueran dos modelos distintos, ya que en un caso emplea el algoritmo de actualización de estado, y en el otro lo ignora.

Ahora bien, dado que el modelo con el cual se decidió trabajar tiene la estructura adecuada para funcionar bien en modo simulación y en modo pronóstico, nos preguntamos si habrá algún juego de valores de parámetros que sea suficientemente bueno para funcionar en ambos a la vez con aceptable precisión. Aun cuando ese juego de valores de parámetros no es el óptimo en ninguno de los escenarios, podría ser que sea “aceptablemente bueno” en ambos. De ese modo, no habría que cambiar entre juegos de parámetros según el escenario, sino que, con un juego de parámetros bien elegido, sería suficiente para enfrentar cualquiera de las situaciones.

De este planteo, surgió el segundo interrogante científico:

¿Existe algún juego de valores para los parámetros capaz de satisfacer a un mismo tiempo diversos escenarios con aceptable precisión?

Todo el razonamiento hecho previamente que, de algún modo, enfrenta “modelo en modo simulación” Vs “modelo en modo pronóstico”, puede realizarse de manera análoga para confrontar la performance del “modelo pronosticando a un horizonte” Vs “modelo pronosticando a otro horizonte”.

Para ambos casos de confrontación, emerge como necesaria alguna herramienta capaz de optimizar ambas situaciones extremas, y ofrecer opciones intermedias que permita a un tomador de decisiones, posicionarse en una solución de compromiso.

De allí surge el tercer interrogante científico:

¿Es posible contar con alguna herramienta capaz de ofrecer anticipadamente al tomador de decisión, un conjunto de soluciones de compromiso?

Sin embargo, la bibliografía consultada presenta un espacio de debilidad en relación a la capacidad de los modelos de pronóstico hidrológico a adecuarse a las condiciones cambiantes del proceso operativo.

En consecuencia, fue identificada un área de vacancia en los siguientes interrogantes científicos:

1. ¿Es posible controlar en tiempo real, mediante las combinaciones de valores paramétricos del modelo, su adaptabilidad a los diversos escenarios de pronóstico?
2. ¿Existe algún juego de valores para los parámetros capaz de satisfacer a un mismo tiempo diversos escenarios con aceptable precisión?
3. ¿Es posible contar con alguna herramienta capaz de ofrecer anticipadamente al tomador de decisión, un conjunto de soluciones de compromiso?

En el apartado relativo a la propuesta metodológica, se da detalle del procedimiento establecido para dar respuesta a dichos interrogantes.

3. Objetivos y organización de la tesis

En este capítulo se presenta el objetivo general de la tesis y se da detalle de los objetivos particulares. En consecuencia, el desarrollo de la tesis respeta rigurosamente el abordaje de los experimentos parciales que van conduciendo a los resultados intermedios, los cuales constituyen pasos esenciales para justificar las conclusiones finales.

La mayor parte de nuestros fracasos provienen por querer adelantar la hora de nuestros éxitos.

Amado Nervo

El objetivo general de esta tesis reside en: explorar la aptitud de un método de optimización multi-objetivo de aportar elementos de decisión en relación a la capacidad de un modelo de pronóstico hidrológico para adaptarse con buena performance a diversos escenarios típicos de operación, a partir de una adecuación de sus valores paramétricos.

Este objetivo general pretende ser abordado a partir de un conjunto de objetivos particulares, que se describen a continuación:

- Describir adecuadamente escenarios típicos que se manifiestan en modo operativo.
- Programar todas las herramientas de trabajo en el lenguaje FORTRAN 95 para interactuar y adaptar las metodologías aplicadas en la tesis.

- Adaptar un método de optimización global multi-objetivo para explorar poblaciones optimizadas de parámetros capaces de ofrecer buenas performances en los objetivos propuestos.
- Proponer distintos frentes de Pareto (poblaciones optimizadas) que permitan al operador contar con una herramienta de decisión frente a los escenarios de pronóstico con y sin actualización de estado.
- Proponer distintos frentes de Pareto que permitan al operador contar con una herramienta de decisión frente a los escenarios de pronóstico con distintos horizontes de pronóstico.
- Proponer al operador un procedimiento que le permita tomar decisiones en el curso de un pronóstico.

A los fines de ir abordando estos objetivos, la tesis fue estructurada en cuatro grandes bloques. Una primera parte introductoria en la cual se presenta el contexto general del fenómeno de las inundaciones, abordado con valores numéricos que grafican el impacto que generan en las distintas regiones del planeta, y en particular en la Argentina. Asimismo, se vincula la herramienta del pronóstico hidrológico como soporte preventivo de este tipo de manifestaciones hidroclimáticas. Y en ese contexto, se inserta la problemática de esta tesis orientada a enriquecer los elementos disponibles para la toma de decisiones. A partir de estos elementos, se cierra la primera parte de la tesis con una definición del interrogante científico que le dio nacimiento, y se plantean claramente sus objetivos.

En la segunda parte de este trabajo, se realiza una revisión bibliográfica completa de todas las temáticas involucradas. Estas incluyen la modelación matemática de los procesos de transformación lluvia-caudal (en particular, de los modelos de pronóstico hidrológico), las técnicas de optimización global (y en particular, los

métodos multi-objetivo), y un desarrollo de los criterios adoptados para medir la bondad de los resultados.

La tercera parte de la tesis se focaliza concretamente en los métodos y técnicas empleados en este trabajo, realizando una descripción detallada y pormenorizada de las herramientas de trabajo. Asimismo, se realiza también una descripción de la cuenca de aplicación y los datos disponibles. Esta parte se cierra con una explicación de la metodología propuesta para orientar los trabajos a la concreción de los objetivos.

La cuarta parte de la tesis consta de tres capítulos, el primero de los cuales se focaliza en la adecuación de la herramienta de trabajo a nuevos escenarios para los cuales previamente no ha sido utilizado. Los otros dos capítulos que complementan este bloque, constituyen aplicaciones prácticas de la metodología propuesta en el bloque anterior, y la consecuente producción de resultados que permitieron la obtención de conclusiones específicas para cada abordaje.

Finalmente, en el capítulo 15 se ofrece una conclusión general de toda la tesis, y una síntesis de las conclusiones particulares que fueron remarcándose en los capítulos previos respectivos.

PARTE 2: REVISIÓN DE ANTECEDENTES

En esta segunda parte de la tesis, se realiza una revisión bibliográfica completa de todas las temáticas involucradas. Estas incluyen la modelación matemática de los procesos de transformación lluvia-caudal (en particular, de los modelos de pronóstico hidrológico), las técnicas de optimización global (en particular, los métodos multi-objetivo), y un desarrollo de los criterios adoptados para medir la bondad de los resultados.

Dado que se pretende que el presente informe sea auto-explicativo, en términos de facilitarle al lector los conocimientos necesarios para interpretar los estudios realizados en el contexto de la tesis, se expone a continuación una reseña teórica de las principales áreas del conocimiento que aborda este trabajo:

- Modelación lluvia-caudal (P-Q)
- Pronóstico hidrológico
- Algoritmos genéticos
- Calibración Automática (Optimización)

4. Modelación lluvia-caudal (P-Q)

Este capítulo está destinado a brindar los elementos conceptuales necesarios para interpretar la modelación lluvia-caudal (P-Q) a la luz de la hidrología y ofrecer una aproximación a los principios básicos de construcción de modelos hidrológicos. Existe una vasta experiencia en modelación matemática en general y en modelos aplicados a la hidrología en particular, por lo que resulta esencial realizar una revisión exhausta de dicha experiencia para poder justificar el uso de la herramienta más adecuada en relación a los fines perseguidos en este trabajo.

Nuestras clases dominantes han procurado siempre que los trabajadores no tengan historia, no tengan doctrina, no tengan héroes y mártires. Si la experiencia colectiva se pierde y las lecciones se olvidan, cada lucha debe empezar de nuevo, separada de las luchas anteriores...

Arturo Jauretche

4.1. Sistemas, conceptos y clasificación

Habidas cuentas de que un modelo es una representación simplificada de un sistema real, es necesario aportar una breve introducción previa al concepto de sistema y en particular, del sistema (ciclo) hidrológico.

Existen numerosas definiciones del término sistema. A los fines de este trabajo se consideró adecuada la que ofrece Dooge (1973), quién entiende por **sistema** a *toda estructura, esquema o procedimiento, real o abstracto, que, en un tiempo de referencia dado, interrelaciona una entrada, causa o estímulo de energía, materia o información, con una salida, efecto, o respuesta, de energía, materia o información.*

Matemáticamente se puede representar un sistema como:

$$y(t) = \mu(t) \psi x(t) \quad (3.1)$$

Donde:

$y(t)$ es la función de salida

$\mu(t)$ es la función que representa el proceso

$x(t)$ es la función de entrada

ψ vincula la función de entrada con la del proceso

Los distintos tipos de problemas derivados del análisis de sistemas se asocian principalmente a la incógnita planteada. Según este criterio, los problemas pueden clasificarse en (Bertoni, 2003):

- a. de predicción, en los cuales la incógnita es la salida $y(t)$
- b. de identificación, cuya incógnita es la función que representa el proceso $\mu(t)$.
- c. de detección, que tienen por incógnita la entrada $x(t)$.
- d. de síntesis, donde la incógnita es el sistema mismo.

El problema de síntesis es poco común y el más difícil de imaginar. En este tipo de problemas, se presenta el desafío de diseñar el sistema mismo, con el objetivo de que para una cierta entrada se obtenga una determinada salida, conociendo la función de los procesos que ocurren en el sistema. Por ejemplo, el proyectista de un sistema de drenaje urbano, que debe modelar una cuenca, con sus materiales, pendientes, etc, se enfrenta a un auténtico problema de síntesis.

Los sistemas pueden clasificarse como Lineales, cuando la salida de un proceso es lineal a la entrada, o No lineales, cuando no lo sea. Los lineales cumplen con las propiedades de proporcionalidad y superposición.

Pueden ser Concentrados o Distribuidos, de acuerdo a que las variables que intervienen en el proceso que ocurre en el sistema, no varíen en el espacio, o sí lo hagan, respectivamente.

También pueden clasificarse entre Continuos y Discretos. Siendo los primeros aquellos sistemas cuyos procesos incluyen variables de variación continua, y los segundos, cuyos procesos incluyen variables de variación discreta.

Finalmente, se habla de los sistemas Causales y No causales. Siendo los causales aquellos sistemas en los cuales el efecto ocurre en un tiempo posterior o en el mismo tiempo que la causa que le dio origen. No causales son aquellos en los cuales lo hacen en un tiempo anterior.

4.2. El sistema (ciclo) hidrológico

De los innumerables sistemas que ofrece la naturaleza, en particular nos interesa el denominado: ciclo hidrológico (Figura 4-1), considerado el concepto fundamental de la hidrología. De acuerdo con la clasificación mencionada anteriormente, se puede decir que el Ciclo Hidrológico, constituye un sistema de tipo no lineal, distribuido, continuo y causal. Como se verá posteriormente, las limitaciones inherentes al hombre obligan a representar el alto grado de complejidad que propone la naturaleza mediante un modelo lineal, concentrado y discreto.

Como todo ciclo, no tiene principio ni fin, y su descripción puede comenzar en cualquier punto, se elige arbitrariamente un punto de comienzo para su explicación.

El agua que se encuentra sobre la superficie terrestre o muy cerca de ella se evapora bajo el efecto de la radiación solar y el viento. El vapor de agua, que así se forma, se eleva y se transporta por la atmósfera en forma de nubes hasta que se condensa y cae hacia la tierra en forma de precipitación. Durante su trayecto hacia la superficie de la tierra, el agua precipitada puede volver a evaporarse o ser interceptada por las

plantas o las construcciones, luego fluye por la superficie hasta las corrientes o se infiltra.

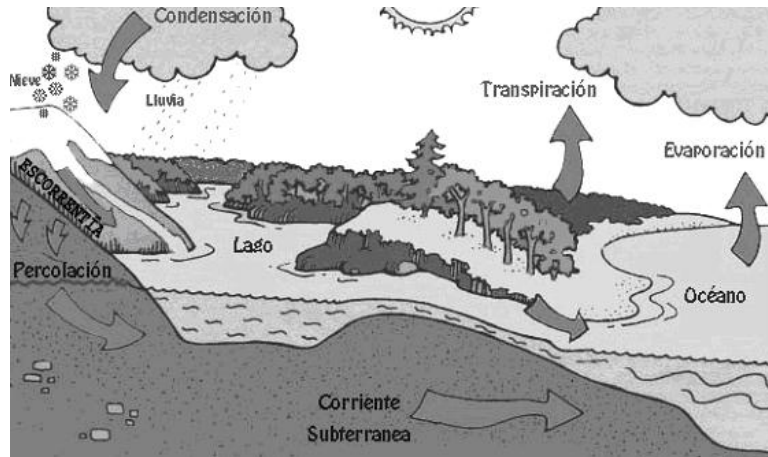


Figura 4-1: El ciclo hidrológico (climaestres.blogspot.com.ar)

El agua interceptada, más una parte de la infiltrada y de la que corre por la superficie se evapora nuevamente. De la precipitación que llega a las corrientes, una parte se infiltra y otra llega hasta los océanos y otros grandes cuerpos de agua, como embalses y lagos. Del agua infiltrada, una fracción es absorbida por las plantas y posteriormente es transpirada, casi en su totalidad, hacia la atmósfera, y otra fluye bajo la superficie de la tierra hacia las corrientes, el mar u otros cuerpos de agua, o bien hacia zonas profundas del suelo (percolación) para ser almacenada como agua subterránea y después aflorar en manantiales, ríos o el mar. Una parte de esta agua es nuevamente evaporada, y así comienza nuevamente el ciclo hidrológico.

4.3. El proceso de transformación lluvia-caudal

El ciclo hidrológico descrito anteriormente, se desarrolla a nivel global, alimentado por una innumerable cantidad de subprocesos que ocurren en forma simultánea en todo el planeta. En la ingeniería hidrológica, rara vez interesa un estudio global del ciclo hidrológico. En cambio, tiene mucha aplicación el estudio hidrológico de cuencas, dentro de las cuales ocurre el proceso de transformación lluvia-caudal. En otras palabras, la ingeniería hidrológica, prefiere estudiar el proceso por el cual el

agua de lluvia que cae sobre la cuenca, se mueve sobre y dentro de ella, generando hidrogramas de flujo subterráneo y superficial, cuyos volúmenes y picos totales interesa conocer.

El escurrimiento se define como el agua proveniente de la precipitación que circula sobre o bajo la superficie terrestre y que llega a una corriente para finalmente ser drenada hasta la salida de la cuenca. El agua proveniente de la precipitación que llega hasta la superficie terrestre, una vez que una parte ha sido interceptada y evaporada, sigue diversos caminos hasta llegar a la salida de la cuenca. Conviene dividir estos caminos en tres clases: escurrimiento superficial, escurrimiento subsuperficial y escurrimiento subterráneo. La Figura 4-2 ilustra el proceso.

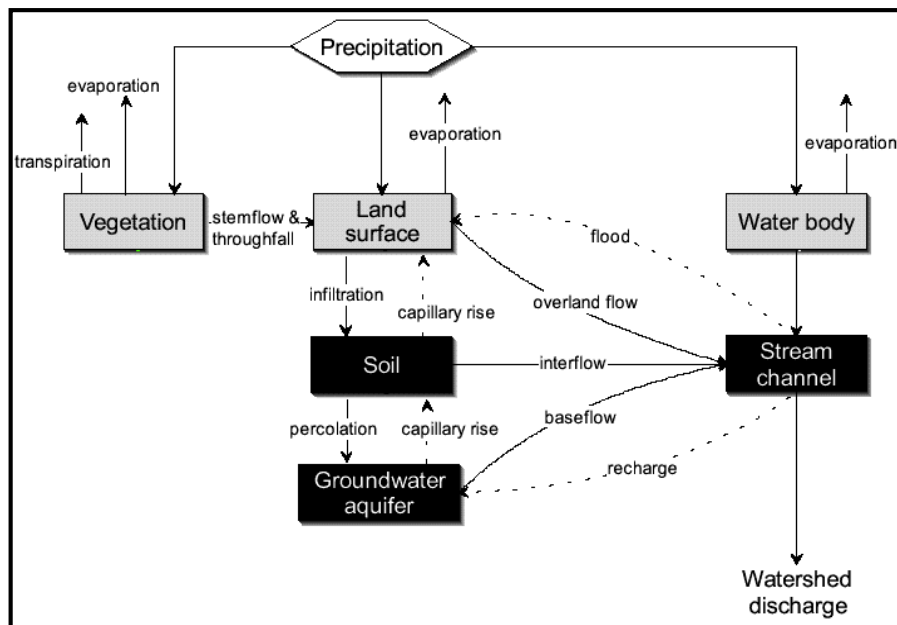


Figura 4-2: Modelo hidrológico de transformación lluvia-caudal (HEC-HMS, 2000)

Una vez que la precipitación alcanza la superficie del suelo, se infiltra hasta que las capas superiores del mismo se saturan. Posteriormente, se comienzan a llenar las depresiones del terreno y, al mismo tiempo, el agua comienza a escurrir sobre su superficie. Este escurrimiento, llamado flujo en la superficie del terreno, se produce mientras el agua no llegue a cauces bien definidos (se entiende por cauces definidos

a aquellos que no desaparecen entre dos tormentas sucesivas). En su trayectoria hacia la corriente más próxima, el agua que fluye sobre el terreno se sigue infiltrando, e incluso se evapora en pequeñas cantidades. Una vez que llega a un cauce bien definido se convierte en escurrimiento en corrientes.

El flujo sobre el terreno, junto con el escurrimiento en corrientes, forma el escurrimiento superficial. Una parte del agua de precipitación que se infiltra escurre cerca de la superficie del suelo y más o menos paralelamente a él. A esta parte del escurrimiento se le llama escurrimiento subsuperficial; la otra parte, que se infiltra hasta niveles inferiores al freático, se denomina escurrimiento subterráneo.

De los tres tipos de escurrimiento, el superficial es el que llega más rápido hasta la salida de la cuenca. Por ello está relacionado directamente con una tormenta particular y entonces se dice que proviene de la precipitación en exceso o efectiva y que constituye el escurrimiento directo. El escurrimiento subterráneo es el que de manera más lenta llega hasta la salida de la cuenca (puede tardar años en llegar), y, en general, difícilmente se le puede relacionar con una tormenta particular, a menos que la cuenca sea demasiado pequeña y su suelo muy permeable. Debido a que se produce bajo el nivel freático, es el único que alimenta a las corrientes cuando no hay lluvias y por eso se dice que forma el escurrimiento base.

El escurrimiento subsuperficial puede ser casi tan rápido como el superficial o casi tan lento como el subterráneo, dependiendo de la permeabilidad de los estratos superiores del suelo; por ello es difícil distinguirlo de los otros dos. Cuando es relativamente rápido se le trata junto con el escurrimiento superficial, y cuando es relativamente lento se le considera parte del subterráneo.

La clasificación anterior, aunque ilustrativa, no deja de ser arbitraria. El agua puede comenzar su viaje hacia la corriente como flujo superficial e infiltrarse en el camino, terminando como escurrimiento subsuperficial o subterráneo. A la inversa, el

escurrimiento subsuperficial puede emerger a la superficie si en su camino se encuentra con un estrato muy permeable que aflora en una ladera (vertiente). Lo importante en realidad es la rapidez con que una cuenca responde a una tormenta, pues esto es lo que determina la magnitud de las correspondientes avenidas.

4.4. ¿Qué es un modelo y para qué sirve?

La hidrología trata de los fenómenos naturales complejos encontrados en el ciclo hidrológico. Los procesos, como la precipitación, evaporación, infiltración y el escurrimiento en ríos, dependen de un gran número de factores que dificulta el análisis cuantitativo y cualitativo de los mismos.

Tucci (2005), define **modelo** como la representación de algún objeto o sistema, en un lenguaje o forma de fácil acceso y uso, con el objetivo de entenderlo y buscar sus respuestas para diferentes entradas.

Cuanto más complejos son los sistemas, más necesarios y más desafiantes son los modelos. Según destaca Tucci (2005), el modelo hidrológico es una de las herramientas que la ciencia desarrolló para entender y representar mejor el comportamiento de la cuenca hidrográfica y prever condiciones diferentes de las observadas.

Dada la gran complejidad que caracteriza los procesos implicados en el ciclo del agua, los hidrólogos enfrentan hoy en día sus estudios parciales. En efecto, la representación de todo o parte de ese ciclo (proceso de modelización), para distintos objetivos, se enfrenta con la dificultad de captar, interpretar y describir los fenómenos y sistemas estudiados. En consecuencia, la representación de estos sistemas, es decir los modelos, son necesariamente simplificaciones, reducciones de la complejidad natural, y por lo tanto groseramente inexactos. El modelador captura del sistema los elementos y aspectos que juzga ser los más pertinentes para tender hacia una solución lo menos inexacta posible.

El desarrollo de modelos reposa generalmente sobre tres elementos (Perrin, 2000):

1. El sistema observado y su discretización espacial y temporal, que definen el objeto y los límites (espaciales o temporales). En nuestro caso, el sistema es típicamente una cuenca hidrográfica.
2. El objetivo de la modelación, por el cual el modelo es desarrollado. El modelo es construido para responder a algún interrogante, pudiendo ser al mismo tiempo una herramienta de representación y un instrumento de interpretación del sistema representado. En nuestro caso, nos interesa la representación de la transformación P-Q.
3. La elección de una formulación de la realidad. Una vez definidos los puntos precedentes, lo esencial del proceso de modelación consiste en encontrar la formulación de la realidad tan satisfactoria como sea posible en términos de los objetivos fijados. En efecto, el modelo es, en general, el resultado de un compromiso entre generalidad, realismo y precisión (Kauark-Leite y Nascimento, 1993). La formulación de un modelo es condicionada por los conocimientos anteriores de los procesos o de los sistemas considerados, por las ideas e imaginación del modelador y por las hipótesis formuladas. En este trabajo sólo se utiliza un modelo matemático, es decir, la formulación de la realidad es a través de ecuaciones matemáticas.

El estudio de transformación P-Q es una de las numerosas disciplinas que incluye la hidrología, y es de la cual nos ocupamos en esta tesis. A excepción de los modelos de cálculo del flujo del agua a una escala continental, normalmente acoplado a modelos climáticos de circulación general, el objeto de un estudio de modelos P-Q es la cuenca hidrográfica. La cual está vinculada a un punto sobre un curso de agua, en términos de la porción de tierra que drena por una red hidrográfica situada aguas arriba de este punto. Así, una gota de agua de lluvia que caiga sobre esta superficie

acabará por pasar por este punto de concentración referido del río, a menos que sea evaporado o bien infiltre hacia capas más profundas. La definición de los límites de una cuenca hidrográfica es esencialmente topográfica. También la composición subyacente geológica puede producir una diferencia entre la configuración superficial topográfica y la superficie real que aporta al punto de concentración.

En el contexto espacial de la cuenca hidrográfica, los problemas típicos de modelación hidrológica cuantitativa pueden clasificarse en (Perrin, 2000):

- Simulación de caudales, para relleno de lagunas en series de datos, la reconstitución de caudales históricos (los datos de lluvia son normalmente disponibles en periodos mucho mayores que los caudales), o para tratamiento estadístico.
- Predeterminación de caudales de crecida o de recesión: se procura saber con qué frecuencia los caudales de crecida (superiores a algún caudal crítico, por ejemplo) o mínimos (por debajo de un caudal de reserva, por ejemplo) amenazan producirse y por cuánto tiempo. Acá nos ubicamos en un desarrollo de tipo análisis frecuencial. Este conocimiento puede permitir el dimensionamiento de obras y embalses o adecuación de cursos de agua para transporte de caudales.
- Pronóstico de crecidas: se trata de evaluar anticipadamente (con una antelación de algunas horas a algunos días), conociendo el estado de la cuenca, los caudales de las crecidas susceptibles de presentar riesgo de inundación o los caudales de recesión que podrían requerir la activación de alguna gestión particular del uso del recurso (por ejemplo, para presas de reserva) para asegurar el aprovisionamiento. Este estudio se inscribe dentro de un enfoque continuo de la modelación.

- Influencia en cambios en la hidrología: se procura predecir los cambios de la respuesta de la cuenca a partir de modificaciones de las características de la cuenca, de origen humano (antropogénicos) o cambios ambientales.

Estas problemáticas conducen a dar respuesta a dos aspectos importantes: la evaluación del riesgo y la gestión del recurso. La competencia de la respuesta que somos capaces de aportar está condicionada por la del modelo en su representación de la cuenca en relación a los objetivos fijados.

4.5. Una clasificación de los modelos P-Q

Un modelo puede ser clasificado según diversos criterios. Un sistema también puede ser clasificado bajo los mismos criterios, por lo que la clasificación que se hace a continuación, vale para ambos, y los menciona indistintamente. Sin embargo, no siempre un sistema es representado por un modelo de las mismas características, debido a las simplificaciones adoptadas para estos últimos (Tucci, 2005). Por ejemplo, el escurrimiento superficial en la cuenca es un proceso no lineal, pero es representado frecuentemente por el hidrograma unitario, que es un modelo lineal. Esto es debido a que el error asociado a la simplificación, no justifica un modelo no lineal.

Como lo indica Perrin (2000), se ha dado una gran cantidad de clasificaciones de modelos, por lo que a continuación se presentan algunas concepciones y criterios tomados de diversos autores:

Concentrado (global), semi-distribuido o distribuido. Un modelo es concentrado cuando no tiene en cuenta la variabilidad espacial, y todas sus variables y parámetros son promediadas en el área de la cuenca. Un modelo es distribuido cuando sí tiene en cuenta esta variabilidad. Como destaca Tucci (2005), en la práctica no existen modelos puramente distribuidos, ya que son utilizadas discretizaciones numéricas, que de alguna forma vuelven el modelo distribuido en concentrado, en una pequeña

subdivisión. En efecto, un modelo es considerado como semi-distribuido cuando divide la cuenca en subcuencas, o bien en partes identificadas conceptual o geométricamente, y a cada una de sus partes las toma como un elemento concentrado en el cual reinan variables y parámetros promedios. Luego integra estas partes para resolver globalmente el problema.

Estocástico o determinístico. De acuerdo a Chow (1964) un modelo es estocástico cuando tiene en cuenta en su formulación el concepto de probabilidad asignado a la ocurrencia de la variable. Por lo tanto, indica que es determinístico cuando el modelo sigue una ley definida, no de probabilidades. De acuerdo a Dooge (1974) un modelo es determinístico cuando para una misma entrada, el sistema produce siempre la misma salida. Pero es estocástico cuando la relación entre la entrada y la salida es estadística.

Tucci (2005) remarca que, aun cuando las variables de entrada y salida de un sistema sean aleatorias, el sistema puede tener comportamiento determinístico o ser representado por un modelo determinístico. Hace notar además que el simple examen de la respuesta de un sistema puede llevar a conclusiones erradas, ya que un sistema con comportamiento aparentemente aleatorio también puede ser determinístico. En efecto, cuando el sistema es no lineal y altamente dependiente de sus condiciones iniciales, la respuesta puede presentar características de una variable aleatoria y pasar todas las pruebas estadísticas y estocásticas. Este concepto llevado al extremo ha sido denominado por la bibliografía como “caos determinístico”.

Físico, conceptual y empírico. Un modelo es denominado con base física cuando ha sido formulado sobre la base de los procesos físicos intervinientes. Los modelos empíricos o de “caja negra”, son aquellos que prescinden completamente en la base física, y sólo persiguen el objetivo de aproximar el caudal simulado al observado. La definición de modelo con base física es algo artificial pues la mayor parte de los

modelos fuertemente sustentados en la física de los procesos simulados, involucran también formulaciones empíricas dentro de sus funciones (como las ecuaciones de Darcy o Horton, por ejemplo). Este tipo intermedio de modelo, que relaciona características del proceso, pero que mantiene un razonable empirismo en los parámetros de las ecuaciones involucradas, se denomina conceptual.

Existe un gran número de modelos hidrológicos, y no necesariamente cada uno de ellos está involucrado en uno de los grupos detallados previamente. La clasificación anterior ha pretendido servir de marco conceptual para caracterizar los modelos hidrológicos desde el punto de vista de su concepción. Aunque hay modelos hidrológicos bien definidos, que se podrían identificar plenamente con uno de esos grupos, existen múltiples matices que incluyen formulaciones intermedias con componentes de varios de estos tipos.

Además, por su propia naturaleza, hay una tendencia a agruparse en función de sus características. Así, por ejemplo, los modelos físicos, normalmente son determinísticos y distribuidos; los estadísticos, son empíricos; y los conceptuales, suelen ser modelos globales.

4.6. Evolución de los modelos P-Q

Hasta los años 30 del siglo XX, la modelización hidrológica había sido una ciencia llena de empirismos y con una vaga comprensión de los procesos implicados. Entonces, Robert E. Horton en 1933 introdujo el rol de la infiltración en el ciclo hidrológico (Beven, 2004), iniciando un marco de estudio más consistente para cuantificar el caudal superficial, o mejor dicho para determinar la parte de lluvia que efectivamente produce escorrentía (lo que se llama función de pérdidas o función de producción), que además de la infiltración incorporaría los procesos de evaporación y retención.

Por la misma época, en 1932, Sherman introdujo el concepto del hidrograma unitario HU, que con el tiempo se vio que se podía integrar dentro de la teoría de sistemas, donde se asume que la cuenca es un sistema lineal, causativo e invariante en el tiempo. A raíz de estos desarrollos, clásicamente han sido utilizados los llamados modelos globales o agregados, que tratan la cuenca como si fuera una sola entidad con una única entrada de lluvia (lluvia media), y el caudal de salida se reproduce a partir de una dinámica global del sistema.

La identificación del HU no es sencilla. A partir de información hidrológica disponible de la cuenca (muchos datos de lluvia y caudal), el HU se puede derivar de la solución del problema inverso, como lo presenta Chow et al. (1988). Pero estos datos no siempre están disponibles con la suficiente cantidad y calidad, y además arrastran la descripción espacial limitada de la lluvia.

En el intento de evitar este problema, ha tenido lugar la conceptualización de la cuenca en sistemas más sencillos de manera que se pueda derivar un hidrograma unitario que dependa de pocos parámetros, los cuales pueden ser estimados por técnicas estadísticas. Estos son los llamados modelos agregados conceptuales, cuyo exponente más conocido es el método de la cascada de depósitos propuesto por Nash (1957).

Un paso adelante en este sentido han sido todos los esfuerzos hechos en relación a atar los parámetros del HU conceptual con las relaciones geomorfológicas de Horton (1945) y Strahler (1952), y que han dado lugar a la aparición del hidrograma Unitario Geomorfológico (Rodríguez-Iturbe y Valdes (1979); Bras y Rodríguez-Iturbe (1989); García Bartual (1989). Todos estos trabajos han sido encaminados a encontrar el HU a partir únicamente de las características físicas de la cuenca, sin necesidad de disponer de datos hidrológicos.

En un intento de representar los procesos con una base física mayor, en los años 60 aparecieron unos modelos que podríamos llamar modelos agregados con base física o mejor aún, modelos de componentes con base física. Estos consisten en una sucesión de elementos conceptuales interconectados que representan la respuesta de diferentes subsistemas del ciclo hidrológico (evaporación, escorrentía en medio saturado, infiltración, flujo en canal, etc). Estos modelos respondían a las necesidades operativas del momento (Corral, 2004): la aplicación de los modelos en simulación continua; disminuir la dependencia de los datos hidrológicos; evitar la complejidad y subjetividad de la separación del caudal base; aplicación en cuencas complejas con gran variedad de suelos, vegetación y pendientes. Se pensaba que, si la descripción de la estructura del modelo era muy parecida a la real, entonces los parámetros óptimos podrían obtenerse de las características fisiográficas de la cuenca (Manley, 1975). Entre los más conocidos de aquella época están el modelo propuesto por Dawdy y O'Donnell (1965), el modelo Stanford Watershed IV (Crawford y Linsley, 1966) y el modelo Sacramento (Burnash et al., 1973), y con posterioridad podemos destacar el modelo Xinanjiang (Zhao, 1992) y el ARNO (Todini, 1996). Generalmente, sin embargo, el número de parámetros involucrados es muy grande, la descripción es demasiado simplificada para conseguir los objetivos propuestos, y los valores de los parámetros que pueden dar resultados aceptables no tienen sentido físico, haciendo que no exista demasiada diferencia entre este tipo de modelos y los modelos agregados conceptuales (Beven, 1989). Por eso, la dificultad principal de los modelos físicos ha residido siempre en la variabilidad espacial de los procesos, y en la dificultad de estimar sus parámetros (Tucci, 2005), como lo corrobora el trabajo de Mendoza et al. (2012).

Viendo las limitaciones de los modelos de componentes físicas, y de la mano con los desarrollos matemáticos de la época, a partir de los años 70 hicieron aparición un tipo de modelos con un planteamiento totalmente diferente, que pretendían

modelar el comportamiento de la cuenca sin introducir ningún tipo de conocimiento sobre ella.

Estos modelos (que con rigor no pueden catalogarse como modelos globales dentro de la clasificación realizada hasta ahora), extraen la información que hay entre las entradas y las salidas del modelo, y la estructura del modelo se adapta a esta estrategia. Por este motivo, son llamados modelos de caja negra, ya que es muy difícil encontrar una interpretación que relacione el modelo con el funcionamiento físico de la cuenca. Los modelos empíricos tuvieron gran aceptación inmediata para el pronóstico a tiempo real, debido a que sus características permiten una actualización más fácil de sus parámetros a lo largo del pronóstico (Tucci, 2005).

Los primeros fueron los modelos autorregresivos (modelos AR, ARMA) basados en el análisis de series temporales (Box y Jenkins (1970). Actualmente existe una línea de trabajo basada en los modelos de redes neuronales artificiales (Abrahart, 2004; Shamseldin, 1997), que simulan el funcionamiento de una red de neuronas con la posibilidad de aprendizaje a medida que hay mayor número de datos, y permiten la introducción de múltiples entradas (lluvia, temperatura, humedad, datos históricos de caudal ...). Aunque han existido casos con buenos resultados, a veces mejores que con otros modelos, parece que todavía queda un largo camino para ver la capacidad de este tipo de modelos. En todo caso, no están claros los criterios en que se puede basar la construcción de la estructura del modelo, lo que dificulta la utilización del modelo en otras cuencas y la regionalización de los parámetros en cuencas sin datos. Y como son casi incapaces de aportar información sobre los procesos físicos, hay quien considera que son científicamente poco útiles de cara al desarrollo racional de la hidrología.

Sea como fuere, la mayoría de modelos descritos hasta aquí han partido de la limitación que supone el conocimiento poco exhaustivo de muchas de las variables que intervienen en el sistema, así como de la imposibilidad de manejar grandes

cantidades de información. Sin embargo, la tecnología actual permite conocer ciertas variables con una buena resolución espacial (por ejemplo, utilizando la teledetección), y las computadoras actuales ofrecen la posibilidad de tratar mucha información en un tiempo reducido y de manera operativa (por ejemplo, haciendo uso de los sistemas de información geográfica SIG). Esto ha hecho pensar que un modelo que mire de incorporar más cuidadosamente la variabilidad espacial de las diferentes variables y reproducir más fielmente los procesos que tienen lugar dentro de la cuenca, podrá ofrecer mejores resultados que uno de agregado. Estos son los llamados modelos distribuidos, que además pueden proporcionar información en puntos interiores de la cuenca (y no sólo en la salida).

Freeze y Harlan (1969) propusieron un esquema para diseñar los modelos hidrológicos del futuro, poniendo la base de los modelos físicos (o modelos distribuidos con base física, o modelos totalmente distribuidos). Estos modelos plantean la descripción de los fenómenos a escala fina (100-500 m), aplicando las ecuaciones características de los diferentes procesos (flujo en superficie, infiltración en suelos no saturados, flujo subterráneo, intercambio suelo-atmósfera, etc.), que son relativamente bien conocidas. Posteriormente se integran las salidas de los diferentes procesos de cada celda con las vecinas, fenómenos que son menos conocidos y mucho más difíciles de describir. Esta estructura deriva en modelos muy complejos que necesitan la integración numérica de las ecuaciones diferenciales y condiciones de contorno implicadas, requieren una gran cantidad de información para poder ser aplicados, y hay que calibrar un número enorme de parámetros en caso de no poder estimar todas las variables. El modelo SHE (Système Hydrologique Européen, Abbott et al., 1986a y Abbott et al., 1986b) ha sido el proyecto más ambicioso (fue fruto de la cooperación entre los institutos más prestigiosos de Europa), y quizá se ha convertido en el modelo más complejo de los actualmente

existentes, pero de momento su aplicación queda limitada a algunas cuencas experimentales de pequeño tamaño.

Debido a que describir con gran detalle los procesos hidrológicos de la cuenca, y conocer todas las variables que intervienen con exactitud, son tareas actualmente inabordables, estos modelos tan complejos muestran con claridad el paradigma de la hidrología. Reducir la escala del problema no es condición suficiente para mejorar la descripción de los procesos, y cualquier intento de incrementar la complejidad del modelo en este sentido puede convertirse en un fracaso, sobre todo si los datos disponibles no están en consonancia con la estructura del modelo (Beven, 1989; Jakeman y Hornberger, 1993).

En este sentido, la información habitual con el que comparar las salidas del sistema (medidas de caudales y del estado de saturación en algunos puntos de la cuenca, en los mejores casos) no parece una información suficientemente completa para poder conocer el estado hidrológico de la cuenca con la calidad y resolución que piden estos modelos.

Es de suponer que en el futuro se dispondrá de unos medios técnicos y de unos conocimientos que permitirán mejorar el grado de fiabilidad de los modelos hidrológicos. En todo caso, sin embargo, parece que una modelización que pudiera describir exactamente todos los procesos de manera determinista, y que se basara en una medida cuidadosa de las variables que intervienen, queda muy lejos del alcance científico-técnico y en general de la realidad en que nos encontramos en estos momentos.

Esto explica (Michel, 1991) que el margen operacional dejado al enfoque físico-determinístico, es decir inspirado estrechamente por la física de los fenómenos, es débil. También agrega que, en los problemas de predeterminación, donde es necesario estimar la distribución estadística de los caudales, este enfoque no

convergerá verosímilmente jamás ya que, frente a cualquier descripción detallada homogénea y coherente del sistema físico representativo de la cuenca hidrográfica, será necesario disponer de un proceso multivariable que permita simular las entradas (lluvias, temperaturas, energía radiactiva, humedad de la atmósfera, viento, etc). En consecuencia, el interés es principalmente, y continúa siendo, hacia los enfoques estadísticos, estudiando directamente los caudales o estableciendo las relaciones causales que unen los caudales con tal o cual dato disponible sobre las largas series.

Una alternativa más acorde con las posibilidades actuales consiste en un punto a medio camino entre los dos extremos, de modo que algunas de las variables pueden ser descritas con buena resolución, incorporando la vertiente espacial distribuido, sin que el modelo sea extremadamente complejo, siendo computacionalmente eficiente. Los modelos más comúnmente utilizados son los llamados modelos de parámetros distribuidos, que se caracterizan por la división de la cuenca en subunidades menores más homogéneas, la respuesta global de la cuenca se compone a partir de las contribuciones de las diferentes subunidades. Una de las ideas ha sido integrar la filosofía de los modelos de componentes con base física dentro de una estructura distribuida (por ejemplo, el modelo Tetis, de Vélez, 2001). Los modelos distribuidos no garantizan necesariamente mejor resultados que los modelos conceptuales, pero permiten analizar con mejor fundamento la modificación de las características del sistema, lo cual es casi imposible a través de los modelos conceptuales, debido al exceso de empirismo, y a las características concentradas de su formulación (Tucci, 2005). Ahora bien, buena parte de los modelos de este tipo lo que hacen es aplicar un modelo agregado sencillo a cada subunidad.

En una primera clasificación, este tipo de modelos se suelen diferenciar por la manera de hacer la división en subunidades, que puede ser en subcuencas (donde

no se pierde el sentido de cuenca hidrológica) o en subunidades artificiales siguiendo algún criterio establecido. Entre los primeros los más conocidos son el paquete de modelos HEC-1 (Hydrologic Engineering Center, 1990) y el modelo TR20 (del Soil Conservation Service, 1973). Entre los segundos ha habido un buen número de propuestas, pero se puede destacar la división en rectángulos equivalentes (Engman y Rogowski, 1974), las laderas triangulares del modelo SHIFT (Palacios-Vélez y Cuevas-Renaud, 1992), y el modelo MIT (Maddaus y Eagleson, 1969) que hace una división en elementos conceptuales basados en depósitos y conexiones de canales lineales.

El hecho de que a cada subunidad se aplique un modelo global hace que este tipo de modelos se denominen también modelos distribuidos, pero conceptualmente son modelos globales (Corral, 2004). Y es que estos modelos parten de la suposición de que la complejidad de los fenómenos se reduce cuando se reduce la escala del problema, ya que el modelo se aplica sobre unidades teóricamente más homogéneas. A menudo, sin embargo, no se analiza si el modelo agregado representa la realidad a escalas más finas, y en general los parámetros involucrados siguen sufriendo una falta de sentido físico, lo que se ve agravada por el hecho de que el número de parámetros a calibrar aumenta con las subdivisiones. Además, la composición de las aportaciones entre subunidades puede no representar adecuadamente la realidad, lo que puede hacer poco rentable el anterior esfuerzo de descomposición.

Una variante de los modelos distribuidos, que intenta superar los problemas que padecen los modelos anteriores (sobre todo el problema del número de parámetros), es aquella que formula una integración entre las subunidades de manera que quede representada la espacialidad de las variables de forma sencilla. En esta línea cabe destacar Topmodel (Beven y Kirkby (1979)), que incorpora la información topográfica de la cuenca a partir de la proposición de un índice

topográfico (relacionado con la capacidad de saturación de cada punto), y describe globalmente los caudales a partir del estado de saturación de la cuenca. En una línea similar, apareció una interesante teoría introducida por Wood et al. (1988) que intenta encontrar una escala adecuada donde las heterogeneidades pueden ser tratadas estadísticamente, definiendo una "área representativa elemental" (REA). Estos conceptos han sido posteriormente desarrollados a partir de la introducción de una serie de ecuaciones y restricciones basadas en las leyes de la termodinámica, definiendo una "cuenca representativa elemental" (REW), por Reggiani et al. (1999) y Reggiani et al. (2005). En esta unidad se podrían aplicar las ecuaciones de manera relativamente sencilla y con una buena base física (o con una mayor comprensión de los procesos globales), quedando una adecuada integración entre las diferentes unidades para obtener la aportación total.

Esta parece ser la tendencia actual, como se pone de manifiesto en los intentos que se están haciendo con el modelo totalmente distribuido Topkapi (Ciarapica y Todini, 2002) para incrementar la escala de trabajo sin perder la base que permite mantener la descripción física de los parámetros (Liu, 2008).

Se puede pensar sin embargo que los desarrollos futuros escapan cada vez más al dilema "enfoque físico"- "enfoque estadístico" para enfocarse en una modelización conceptual muy simple, pero nueva y liberada de prejuicios en materia de deducciones físicas injustificadas debido al desconocimiento del sistema. De la misma manera que la hidráulica a superficie libre reposa más sobre la ley de Manning-Strickler que, sobre una descripción fina de la turbulencia, la hidrología deberá cada vez más recurrir a los esquemas globales y simples para efectuar progresos prácticos (Michel, 1991).

4.7. Modelos "a reservorios" P-Q

Dado que en esta tesis se emplea un modelo "a reservorio", resulta pertinente desarrollar con mayor detalle la concepción y estructura de estos modelos, a los fines de una mejor comprensión en su posterior descripción.

Los modelos de esta clase difieren de aquellos llamados "de caja negra" en el hecho de que descomponen la transformación de la lluvia en caudal en sub-procesos que serán tratados de manera global (contrariamente a los modelos basados en la física, los cuales se sustentan en las leyes de la física, como se ve más adelante). Estos procedimientos toman usualmente la forma de reservorios que se llenan y se vacían a lo largo del tiempo. Estos permiten, en consecuencia, al sistema tener una memoria de las condiciones anteriores, que no se limitan exclusivamente a las lluvias y caudales anteriores (como en el caso de los modelos autorregresivos por ejemplo). Estos modelos son esquemáticamente constituidos de reservorios interconectados que aseguran la transformación de lluvia en caudal. Por lo tanto, son menos abstractos que los modelos de caja negra en su representación de la transformación de lluvia en caudal, aun cuando su interpretación física no es generalmente inmediata. Estos modelos, se estructuran a partir de dos componentes principales:

- Un módulo de producción responsable de la determinación de los balances de agua, es decir, de la repartición de la lluvia bruta en lluvia neta (alimentando el caudal), en cantidades evaporadas y en cantidades almacenadas. A veces presentes en estos modelos, las funciones de intercambio de agua que permiten simular las pérdidas y/o los aportes desde y hacia el exterior, forman también parte de este módulo.
- Un módulo de transferencia o "ruteo" que permite asegurar la repartición temporal de la cantidad de agua que transita en el curso de agua.

A medida que el paso de tiempo de estudio aumenta, la importancia de la transferencia disminuye y el modelo tiende a reducirse a una función de producción destinada a administrar el balance de agua. A medida que el paso de tiempo disminuye, la transferencia tiene tendencia de volverse cada vez más sofisticada. Producción y transferencia no son siempre claramente dissociables en el modelo: un reservorio puede jugar un rol para cada una de estas funciones. Estas funciones dependen de parámetros cuyos valores deben ser determinados para cada cuenca hidrográfica. Rara vez, medibles en forma directa, la mayor parte de los parámetros deben ser optimizados por un procedimiento de calibración numérica.

Como lo remarca Perrin (2000), la diferencia subyacente entre los adjetivos "conceptual" y "empírico" en relación a la clasificación de los modelos, hace exclusiva referencia a la manera con la cual los modelos fueron desarrollados:

- La gran mayoría de los modelos a reservorios existentes son modelos conceptuales puestos a punto a partir de conceptos sobre la manera con la cual ocurre la transformación P-Q. Se representa entonces en el modelo de manera simplificada los procesos juzgados pertinentes para la modelación. El diseñador del modelo introduce las ideas a priori, generalmente fundadas sobre los conocimientos físicos y las observaciones de la cuenca. Beven (1987) califica por lo tanto este enfoque de "perceptual", subrayando que se vuelve personal y está dominado por el modo en que el modelador percibe la realidad.
- Mucho más marginales que los precedentes, los modelos empíricos, son construidos solamente a partir de las observaciones hidro-meteorológicas (correspondiendo a las entradas y salidas del modelo), sin reparar sobre hipótesis previas respecto de la naturaleza de los procesos dominantes o las ideas preconcebidas sobre la estructura del modelo. El modelo es entonces construido a partir de una estructura elemental, que se volverá

progresivamente más compleja si es necesario para una mejora en la eficacia. El modelador busca, en efecto, descubrir cuales elementos deben intervenir en esta estructura y como deben ser interrelacionados, no en función de una cadena de fenómenos de la cual tiene conocimiento en el medio natural, sino en función del mayor o menor éxito del modelo para reproducir los caudales observados.

Varios estudios fueron consagrados a comparar los modelos fundados en la física con relación a los de "caja negra" o a los modelos a reservorios.

Lindström et al. (1997) entienden que la filosofía básica de construcción de un modelo debe estar fundada en los siguientes principios:

- el modelo se basa en un fundamento científico sólido;
- en la mayoría de las cuencas, la necesidad de datos debe estar de acuerdo con la disponibilidad;
- la complejidad del modelo debe ser justificada por su performance;
- el modelo debe estar debidamente validado;
- el modelo debe ser comprensible por los usuarios.

A lo cual, Perrin (2000) agrega que el modelo debe ofrecer las respuestas satisfactorias a los interrogantes planteados y debe presentar las cualidades de generalidad y robustez. Indica que un modo de evaluar el grado de satisfacción que esos modelos ofrecen en sus respuestas a las cuestiones planteadas es a través de estudios comparativos.

Según agrega Perrin (2000) una conclusión del estudio de Loague y Freeze (1985) es que la aplicación de los modelos fundados sobre la física presenta el problema de la medida de la variabilidad espacial de las entradas del modelo (lluvia) y sobre todo de

las características de la cuenca (propiedades del suelo, por ejemplo). No obteniendo mejores resultados que con los modelos bien simples, ellos critican duramente estos enfoques tan complejos. Michaud y Sorooshian (1994) aplicaron, sobre una cuenca experimental norteamericana bajo un clima semi-árido, un modelo basado en la física y dos versiones, una distribuida y la otra global, del modelo del SCS acoplado a un hidrograma unitario. Los resultados muestran que cuando se dispone de todos los datos necesarios, se encuentran leves mejoras del modelo más complejo respecto del simple.

Perrin (2000) muestra varios ejemplos bibliográficos más, y concluye que queda bastante en evidencia que los modelos basados en la física son todavía muy poco convenientes, ya que el grado de realismo que demandan, en relación al esfuerzo de construirlos y disponer de los datos que requieren, debería darles mucho mayor precisión, cosa que no se ha puesto todavía en evidencia en ningún estudio comparativo.

4.8. Aspectos complementarios de modelación

4.8.1. Variables y parámetros

Como fue descrito previamente, los modelos a reservorios tienen una estructura basada en reservorios, que van administrando el balance de agua a lo largo del proceso de transformación P-Q, en medio de dos grandes bloques conceptuales que son el de producción y transferencia.

Esta estructura involucra las variables de entrada que están normalmente representadas por la lluvia (P) en primer lugar, y la evapotranspiración potencial (EP) también en la mayoría de los casos. Análogamente, se define como variable de salida el caudal (Q). El caudal puede ser también una variable de entrada en el caso de un modelo de pronóstico hidrológico, como se verá más adelante.

Adicionalmente, complementan al modelo las variables de estado, que son aquellas que determinan justamente en cada paso de tiempo de simulación el "estado" del sistema, siendo las que establecen normalmente la cantidad de agua que conservan los reservorios en las funciones de producción y transferencia. La variable típica de estado es, por ejemplo, la humedad del suelo, pudiendo ser también alguna otra variable que no tenga por ejemplo un sentido físico demasiado interpretable pero que representa el nivel de alguno de los reservorios del modelo.

El conjunto de cantidades fijas que caracteriza al sistema recibe el nombre de parámetros del sistema (Tucci, 2006). Los parámetros de un sistema, normalmente no pueden medirse, razón por la cual deben estimarse.

A los fines de ejemplificar los conceptos anteriores, bastaría considerar la cuenca hidrográfica como el sistema en el cual ocurre el proceso de transformación lluvia a caudal. Este proceso, a su vez, involucra principalmente los procesos de retención, infiltración, evaporación, escurrimiento superficial y flujo en canales. El modo de ocurrencia de cada uno de estos procesos estará gobernado por ciertos parámetros de la cuenca, como son el coeficiente de rugosidad, la permeabilidad del suelo, la pendiente de la cuenca, etc.

En consecuencia, el hidrograma generado a la salida de la cuenca (salida del sistema), depende de las variables de entrada a la cuenca (entrada al sistema), que es la *precipitación*, y de los parámetros que caracterizan al sistema. Por esta razón, todo modelo debería involucrar en su estructura, al menos los parámetros más significativos numéricamente que caracterizan el sistema que representa. La lluvia puede ser medida y también el caudal escurrido, aunque sea en forma aproximada. Sin embargo, los parámetros no pueden medirse en forma directa. Normalmente se procede a estimar estos parámetros en función de ciertas características observables. El gran desafío en la modelación matemática de un sistema se basa en la estimación de los parámetros que lo caracterizan.

Estos parámetros estimados y las variables de entrada medidas, se agregan al modelo y se simula el proceso que ocurre en el sistema, obteniendo, como resultado, las variables de salida. Las variables simuladas de salida deberían coincidir con las medidas, al menos dentro del rango de aplicación del modelo. En general, no se logra tal coincidencia, por lo cual es necesario modificar sensiblemente los valores de los parámetros hasta lograr un resultado aceptablemente cercano al buscado. Esta técnica de ajustar los parámetros del modelo recibe el nombre de calibración del modelo matemático (Bertoni, 2003). El éxito de un modelo hidrológico conceptual de transformación lluvia-caudal depende en gran medida de la bondad de su calibración (Duan et al., 1992).

4.8.2. Memoria del modelo, condiciones iniciales y periodo de calentamiento

Dado que los modelos hidrológicos cumplen el rol de simulación de un sistema real (en esencia se proponen predecir un estado futuro), existe necesariamente una influencia de la situación o variables actuales respecto del futuro. Dooge (1973), define el concepto de memoria como el intervalo de tiempo, en el pasado, durante el cual, la entrada afecta el estado presente del sistema. Memoria cero, significa que la entrada afecta el sistema solamente en el tiempo en el que ella ocurre. La memoria infinita existe cuando el sistema depende de todo tiempo pasado. Una memoria es finita cuando el sistema depende de la entrada ocurrida en un periodo finito en el pasado.

Según señala Vélez y Francés (2008), es bien conocido y aceptado que los resultados de los modelos hidrológicos dependen fuertemente de las condiciones antecedentes de humedad, puesto que la evapotranspiración, la capacidad de infiltración, la escorrentía y el almacenamiento en depresiones del suelo se ven afectados por el estado de humedad al inicio del evento.

El efecto de las condiciones iniciales de humedad (siguiendo con Vélez y Francés, 2008) puede ser disminuido si se trabaja a escalas temporales mensuales o anuales

ya que en pocos meses o años el balance se restablece. Sin embargo, este efecto es significativo en modelación de evento y tiene que ser incluido si se desea realizar predicción en tiempo real.

Tangara (2005), alerta sobre los problemas que trae normalmente el desconocimiento de las condiciones iniciales de la cueca. Pero agrega que si se trabaja con un modelo de simulación continua (refiriéndose al modelo GRP), se reduce la importancia de conocer a priori esas condiciones iniciales. ¿A qué se debe? Un modelo hidrológico conceptual P-Q presenta en su estructura variables de estado que están asociadas al estado de la cuenca. Cuando uno comienza la simulación es necesario fijar el valor de dichas variables de estado en el tiempo cero. Si fuera posible establecerlas con precisión se dispondría de un estado inicial perfecto para comenzar el proceso de simulación. Ahora bien, normalmente dicho estado no se conoce, o es muy difícil de determinar ya que se estima normalmente sobre la base de la experiencia. En un modelo de simulación de eventos cortos, la estimación de dichas variables es la única alternativa. Pero en un modelo de simulación continua, cuando se disponen de series de considerable extensión de datos de calibración (varios años), el mismo modelo va ordenando su balance interno procurando alcanzar las condiciones reales de simulación. Con este propósito, es destinado un primer periodo (primeros años) de la serie disponible al ajuste interno de sus variables de estado dejándolas liberadas sin obligar a los caudales simulados a adoptar el valor de los caudales observados (durante ese periodo). Este proceso de ajuste del balance interno es forzado por el proceso de calibración que se realiza con el resto de la serie, en el cual sí es forzado el hidrograma a acercarse tanto como sea posible al hidrograma observado. En otras palabras, el ajuste de los parámetros obtiene un hidrograma simulado tan cercano al observado como sea posible en toda la serie excepto en el primer tramo, el cual es dejado libre a evolucionar el estado interno del modelo. Este primer tramo es denominado "periodo de calentamiento"

pues va "precalentando" al modelo durante un tiempo hasta que el mismo se vuelve "en condiciones" de comenzar a simular.

Por esta razón en los modelos de simulación continua, con una serie suficientemente larga que permita destinar un primer tramo al "calentamiento" del modelo, y el resto a la calibración de los parámetros, las condiciones iniciales pierden importancia. Por esta razón, como se verá más adelante, el modelo GRP no requiere como información de entrada las condiciones iniciales de sus variables de estado.

En el caso de un modelo de pronóstico, que cuenta con la información del caudal observado en cada paso de tiempo, la parte de la "historia de la serie" de mayor importancia es la "historia reciente", es decir lo que acaba de ocurrir. Como veremos más adelante, el modelo de pronóstico se alimenta de los últimos valores de caudales observados, para ajustar su estado, por esta razón Vélez y Francés (2008) denomina "periodo de calentamiento" al periodo del tiempo inmediatamente anterior de la emisión del pronóstico. Durante este periodo, el modelo recibe la información en tiempo real sobre estaciones de aforo, lluvia o temperatura, por ejemplo. Por lo tanto, el periodo de calentamiento se orienta a ajustar automáticamente las condiciones iniciales de humedad al mismo tiempo que la calibración automática procura calibrar los parámetros, todo orientado a obtener el mejor ajuste posible entre los caudales observados y simulados.

Es necesario definir la longitud del periodo del tiempo anterior al pronóstico según el tamaño de la cuenca y de la escala temporal de los datos. Dicho periodo de tiempo de anticipación normalmente se adopta al menos en un ciclo hidrológico completo, y si fuera posible dos. Depende normalmente de la longitud de la serie disponible. Si se dispone de una serie considerablemente larga, reservar dos años para "calentamiento" no impide contar con una serie suficientemente larga como para calibrar.

Por otra parte, Berthet et al. (2009) comentan que, por muchos años, se ha recomendado el enfoque de modelación continua como una solución rigurosa a la estimación de las condiciones iniciales (citando a Linsley, 1982). En efecto, Kitanidis y Bras (1980) alertan que los errores en las simulaciones son muchas veces introducidos por el desconocimiento de las condiciones iniciales del estado del sistema, pero que el enfoque de simulación continua en pronóstico hidrológico vuelve de poca importancia este error, ya que la memoria del sistema tiene alcance finito. Sin embargo, Berthet et al. (2009) agregan que se debe reconocer que a veces todavía es preferida la modelación de eventos en aplicaciones operacionales a tiempo real (citando a Lamb y Kay, 2004). Los modelos basados en evento pueden ser más simples porque normalmente no necesitan incluir todos los procesos necesarios en un modelo continuo. Es decir, menos requerimiento de datos que alivian el modelo y lo hacen más fácil de usar e implementar. Otra razón subyace en la dificultad de muchos países en mantener y validar una red automática de mediciones. Esta es una situación frecuente cuando se pretende tener alta resolución temporal en las series (es decir, horaria). Para superar ese obstáculo Nalbantis (1995) propone basarse en series de datos más gruesos (diarios) para deducir condiciones iniciales con series de datos más finos (horarios). El problema también es cultural, ya que los usuarios finales, acostumbrados a usar métodos hidráulicos basados en eventos, se resisten a emplear los modelos hidrológicos continuos. Otra razón a favor de los modelos de evento es cuando los modelos no son empleados sólo para el pronóstico de crecidas, sino también para otros usos (como simulación de inundaciones torrenciales, por ejemplo).

5. Pronóstico hidrológico

Este capítulo discurre en las técnicas y métodos antecedentes en materia de pronóstico hidrológico. Casi indefectiblemente todo modelo de pronóstico está basado en la historia del fenómeno que pretende predecir. Una interpretación criteriosa de dicha historia ofrecerá la estructura adecuada para modelar y los elementos necesarios para ajustar dichas estructuras a las condiciones reinantes.

El buen historiador tiene mucho de profeta: como conoce y comprende el pasado, puede conocer y comprender el presente, y hasta predecir el futuro.

José María Rosa

5.1. Los sistemas de Alerta temprana

Un Sistema de Alerta Temprana (SAT) comprende la suma de las políticas, estrategias, instrumentos y acciones particulares referidos a la identificación y monitoreo de amenazas, vulnerabilidades y riesgo, el diseño e implementación de alertas o alarma relacionada con la ocurrencia inminente de eventos peligrosos; los preparativos para la respuesta a emergencias y la ejecución de los mismos.

Los Sistemas de Alerta Temprana ante inundaciones son las herramientas más avanzadas que existen en el ámbito de la previsión de crecidas (Quintero, 2011). Estos sistemas combinan el potencial de métodos avanzados para la estimación y previsión de la precipitación junto con el de modelos de simulación de caudales. Un gran número de instituciones han adoptado estos sistemas para producir predicciones de caudal en tiempo real (Quintero, 2011) sobre cuencas de tamaños medianos a grandes (del orden de cientos de kilómetros cuadrados, a miles de

kilómetros cuadrados), lo que ha llevado a un avance significativo en la reducción del número de víctimas de este desastre, gracias a su adecuada alerta y previsión (Meon, 2006; Armien, 2001). Se han desarrollado un gran número de manuales e informes orientados a estandarizar los procedimientos de acción (WMO, 2011).

La importancia de un SAT, radica en que permite conocer anticipadamente y con cierto nivel de certeza, en que tiempo y espacio, una amenaza de inundación puede desencadenar situaciones potencialmente peligrosas. Por lo cual las alertas deben difundirse con suficiente anticipación (Armién, 2001).

Los sistemas automatizados se basan en la observación y monitoreo mediante la utilización de redes telemétricas, estaciones de lluvia y niveles de los ríos, modelos hidrológicos computarizados, sensores remotos, y/o satélites; con lo cual se vigila la cantidad de lluvia, los niveles de los ríos, para finalmente pronosticar crecidas con cierto grado de precisión.

El funcionamiento de un Sistema de Alerta Temprana consiste básicamente en los siguientes pasos (Armién et al., 2001): (i) lectura y registro de la medición de los instrumentos, (ii) transmisión de los datos registrados, (iii) procesamiento e interpretación de la información hidrológica y formulación del pronóstico, (iv) establecimiento del nivel y tipo de alerta, v) difusión del nivel de alerta, y (vi) activación de un Plan de Emergencias o Evacuación. En esta tesis se aborda exclusivamente el paso (iii) referido previamente.

5.2. Diversas metodologías de pronóstico

Existen diversas maneras de pronosticar un evento de crecida hidrológica en la sección de un río. La experiencia de los observadores es tal vez la primera y más intuitiva de todas. Con el paso del tiempo, el conocimiento del sistema físico y las herramientas matemáticas han ido permitiendo al hidrólogo formular metodologías conceptuales capaces de representar numéricamente al sistema a través de modelos

matemáticos. Desde este punto de vista, existen dos tipos de problemas de pronóstico de caudales: los problemas “caudal-caudal”, y los problemas “lluvia-caudal” (Michel, 1991).

El primer caso normalmente es más adecuado para cuencas grandes, en las cuales normalmente se dispone de información de caudal en secciones aguas arriba del tramo de interés, y por lo tanto se puede modelar la onda de crecida entre ambas secciones a través de alguna herramienta hidráulica de propagación de la onda (Failache et al., 2010).

El segundo caso, en cambio, es aplicable principalmente a los ríos con cuencas hidrográficas pequeñas y medianas, de gran velocidad de respuesta y crecidas violentas. En estos casos, el pronóstico de caudal es realizado sobre la base de la disponibilidad de los datos de lluvia anteriores (Michel, 1991). Gaume, et al. (2009) presentan un compendio de cuencas hidrográficas europeas que responden a este criterio, y para el cual sería imprescindible este tipo de pronóstico.

Según Corral (2004), las cuencas donde tiene sentido este tipo de objetivo son las de tamaño medio, aproximadamente entre 50 y 1000 km². Esto se debe a que, en cuencas más pequeñas, las crecidas normalmente no representan un peligro importante, o el tiempo de respuesta puede ser demasiado pequeño para poder actuar a tiempo con este tipo de modelos, siendo entonces más efectivo concentrar los esfuerzos en un buen modelo de previsión de lluvia. Mientras que, en cuencas más grandes, puede hacerse de manera más fiable a partir de puntos de control aguas arriba. En cambio, es en las cuencas de tamaño medio donde la transformación lluvia-escorrentía juega un papel importante para poder anticipar las crecidas.

5.3. Respecto del escenario futuro de precipitaciones

Conforme lo aclara Berthet (2010), es evidente que, para evaluar la capacidad de pronóstico de un modelo de estas características, debemos hacerlo correr "como si" no conociéramos los datos posteriores al instante de la emisión del caudal. En efecto, el modelo (y sus actualizaciones) no debe(n) disponer de los valores de caudal posteriores al instante de previsión. Sin embargo, el modelo sí requiere de los datos de precipitación entre la fecha t y $t+H$ (o sea, a lo largo de todo el horizonte "H" de pronóstico).

En condiciones operacionales, el pronosticador ignora las cantidades de agua que lloverán a lo largo del horizonte de pronóstico. A lo sumo, dispondrá de un pronóstico cuantitativo de precipitaciones (PQP), que presentan una gran incertidumbre.

Existen métodos de análisis estadístico en función de los pronósticos probabilísticos de precipitación, que permiten alargar los horizontes de alerta (Norbiato et al., 2008), asignando una probabilidad determinada a que cierta precipitación sea igualada o superada, y de allí pronosticar el caudal que será igualado o superado mediante un modelo P-Q. Por esta razón, desde hace varios años, se están desarrollando modelos hidrológicos P-Q capaces de acoplar la información de lluvia pronosticada (Georgakakos, 1984 y 1985) por modelos de pronóstico meteorológico. Cabe aclarar que estos modelos también tienen asociados al pronóstico un grado de incertidumbre. Sin embargo, estas herramientas son todavía muy deficientes y presentan grandes dificultades, tanto en el nivel de modelización de ellas mismas, como en la coordinación con el pronóstico cuantitativo acoplado de precipitaciones. Adicionalmente, incluso en los países con mayor desarrollo tecnológico disponen de muy poca información al respecto (Berthet, 2010).

Como el objetivo de este trabajo es la evaluación de la performance del modelo de pronóstico (y no de todo el sistema de pronóstico), preferimos trabajar con las observaciones (no disponibles en condiciones reales durante el pronóstico) a lo largo del horizonte de pronóstico. Este escenario, muy optimista, por cierto, permite eliminar las precipitaciones como fuente de incertidumbre.

En síntesis, este trabajo se ubica en el marco del pronóstico con el empleo de un modelo matemático P-Q bajo la hipótesis de precipitación conocida durante todo el horizonte de pronóstico.

5.4. Modelos P-Q de Pronóstico hidrológico

En un contexto general, el pronóstico hidrológico puede ser definido como “aquella parte de la ciencia y de la ingeniería que trata del análisis y asimilación de datos e información hidrometeorológica, así como también del ingreso de esta información en modelos hidrológicos y procesos de predicción, a fin de formular las previsiones relativas a los estados futuros de varios de los componentes del ciclo hidrológico, especialmente de los flujos escurridos a través de los cursos de agua” (Georgakakos y Hudlow, 1984).

Muchos autores han explorado características y variantes de los modelos de pronóstico hidrológico para mejorar su performance. Entre otros, Georgakakos et al. (1990) presentan un modelo de pronóstico hidrológico que actualiza variables de estado aprovechando la información a tiempo real. En Tilford et al. (2003) se da una síntesis enumerativa de todos los aspectos asociados a la problemática de la modelación de pronóstico hidrológico. Moore et al. (2000), por ejemplo, proveen una guía para la elección de un modelo para diferentes tipos y dimensiones de cuenca y para diferentes situaciones meteorológicas. Del conjunto de modelos comparados se hizo una preselección de 8 modelos sobre 9 cuencas de distintos tamaños, uso de suelo, relieve y litología. Para tres de las cuencas se empleó

pronóstico de lluvia por radares como alternativa de los pronósticos por medición de lluvias. Evaluaron los modelos en modo simulación y en modo pronóstico. El rango de los modelos incluyó diferentes métodos de actualización basados en predicción de error, corrección de estado y ajuste de parámetro. Luego juzgaron la performance empleando estadística. Moore et al. (2001 a y b) describen los modelos de pronóstico hidrológico de mayor uso y se realizó una valoración comparativa entre ellos.

Por norma general, la finalidad última de un modelo hidrológico de crecidas es poder orientar la toma de decisiones de cara a las alertas en zonas sensibles a inundaciones y en embalses (tanto por defensa frente a avenidas como para gestión del agua). Por tanto, el modelo debe ser capaz de funcionar en tiempo real y en continuo, debe poder ser utilizado con la información disponible en el entorno, y debe proporcionar unas previsiones medianamente fiables (Corral, 2004).

Liu y Gupta (2007) ofrecen una visión sistémica para describir un modelo hidrológico, asumiéndolo como un operador f , caracterizado por un juego de parámetros Θ , invariante en el tiempo. El estado del sistema es caracterizado por un vector de estado X_t . El modelo f tiene por lo tanto en cuenta las entradas e_t (que para nosotros son la lluvia, la evapotranspiración potencial y las otras entradas eventuales del modelo) para modificar el vector de estado y determinar las salidas del modelo y_t (que para nosotros sería el vector de caudales). La actualización objetiva a modificar uno de esos elementos basándose en la información tomada de las observaciones z_t .

Refsgaard (1997) distingue cuatro tipos de actualizaciones:

1. Actualización de las entradas (lluvia). Esta actualización se justifica por la mediocridad de nuestro conocimiento de las precipitaciones en la cuenca. Tienen por objetivo mejorar los pronósticos futuros mediante la modificación

de las precipitaciones pasadas de modo de obtener a posteriori los mejores pronósticos en el pasado.

2. Actualización de los parámetros. Conforme a la concepción clásica que fue definida más arriba, los parámetros del modelo para una cuenca dada son teóricamente valores invariantes temporales durante el proceso de transformación P-Q y deben describir el funcionamiento de la cuenca hidrográfica. En consecuencia, estos parámetros son normalmente calibrados sobre una serie histórica de datos: incluso haciendo abstracción del error intrínseco del modelo, hay incertidumbre en los parámetros. Pero también es razonable buscar una (aunque sea leve) modificación de los valores paramétricos recalibrándolos empleando la última información hidrológica del modelo. Se dice en este caso que el modelo funciona en modo "adaptativo".
3. Actualización de los estados. Este enfoque viene a suponer que, debido a las incertidumbres de los datos de entrada, de los parámetros y de la estructura del modelo, ciertos estados no tienen el valor que deberían tener en el instante de la emisión del pronóstico. Esta actualización procura corregir esos valores.
4. Actualización de la salida o corrección del error. Aquí, simplemente se procura corregir el error de salida del modelo sin reparar en explicar su origen. Consiste básicamente en complementar un modelo conceptual de pronóstico con otro empírico de tratamiento de salidas (generalmente un modelo de naturaleza estadística -Troutman, 1985- o de señal para tener en cuenta los errores observados en la calibración.

En esta tesis se trabaja con un modelo perteneciente al "caso 3" previamente descrito. En efecto, como se describe más adelante, el modelo de pronóstico GRP

actualiza su estado en el intervalo de tiempo en el cual emite el pronóstico empleando como información de ajuste del estado la información del último valor de caudal observado.

A causa de esta etapa, denominada de asimilación (que corrige el estado del modelo), un buen modelo de pronóstico no es necesariamente un buen modelo de simulación, y viceversa. Debido al elevado peso asignado a las últimas observaciones, la estructura de un modelo hidrológico de pronóstico puede típicamente ser más simple que aquella de un modelo de simulación resultando potencialmente bastante desconcertante tanto por los desarrolladores como por los utilizadores de estos modelos (Berthet, 2010).

Sin embargo, desde el punto de vista del desarrollo de los modelos hidrológicos, la confrontación entre estos dos objetivos (simulación y pronóstico) no es necesariamente tan clara como parece. En efecto, como lo recuerda Duband (2000), "para una utilización operacional en pronóstico, se necesita de una modelización específica notoriamente distinta a las necesidades de la hidrología aplicada orientada a la comprensión y a la reconstrucción del ciclo del agua: simular no es pronosticar. De todos modos, el progreso en el conocimiento puede contribuir tanto a uno como al otro objetivo". En efecto, para el caso de un modelo de pronóstico hidrológico, altamente dependiente de la actualización de estado a cada paso de tiempo, si la cadena de transmisión de datos hidrometeorológicos (red pluviométrica e hidrométrica) que alimenta el sistema de pronóstico falla, la performance real puede volverse muy por debajo de aquella que uno esperaría en un caso ideal.

Liu et al. (2012) sintetizan el estado del arte en métodos de asimilación de datos (DA), indicando en primer lugar que todavía hay varios desafíos teóricos y matemáticos que se requiere desarrollar para que los métodos de DA puedan beneficiar completamente al pronóstico operativo. Por otra parte, aclaran que el

éxito de un esquema de DA depende críticamente de la caracterización de las incertidumbres.

5.5. Respecto de las incertidumbres

Dado que existen varios elementos incorporados en la estructura de un modelo de pronóstico hidrológico que introducen incertidumbre a la emisión del valor pronosticado, actualmente hay una tendencia general a asignar una función de probabilidades al valor pronosticado (Arduino et al., 2005). Por su parte, Krzysztofowicz (2001) enfatiza sobre la importancia de utilizar modelos probabilísticos en vez de determinísticos para realizar pronósticos, dado que el valor asignado al pronóstico siempre importa un grado de incertidumbre que debe intentarse de evaluar. Verkade y Werner (2011) demuestran que el pronóstico con mayor tiempo de antelación no necesariamente conduce a una reducción en el riesgo de daños por inundación, dado que existe una probabilidad de éxito en la estimación que implica costos asociados de diversa magnitud. Vélez y Francés (2008) indican que, para realizar un análisis de sensibilidad completo del modelo, es necesario realizar un análisis de estimación de la incertidumbre por verosimilitud generalizada. Es decir, emplear el GLUE (Generalized Likelihood Uncertainty Estimator – GLUE), propuesto por Beven y Binley (1992).

Arduino et al. (2005) en una publicación que sintetiza los principales aportes técnicos y científicos de la "Conferencia Internacional de Pronóstico de Inundaciones en Europa", realizada en marzo de 2003 en la ciudad de Rotterdam, indican que las fuentes de incertidumbre y su propagación a lo largo del proceso de pronóstico, son múltiples, siendo principalmente la estimación de las precipitaciones, seguida por las incertidumbres atribuibles al estado interno de los modelos hidrológicos e hidráulicos, la condiciones iniciales y los procesos relevantes de parametrización. Esas incertidumbres dependen, desde luego, de las condiciones particulares de cada

modelo. Concluyen Arduino et al. (2005) que es necesario profundizar el vínculo entre los modeladores y los tomadores de decisión, es decir, usuarios finales de los modelos, ya que la falta de conocimiento de los aspectos claves en el pronóstico son las fuentes principales de los errores en el pronóstico.

6. Optimización multi-objetivo

Este capítulo constituye una revisión de los antecedentes vinculados con metodologías de optimización multi-objetivo. Particularmente se profundiza en los métodos de algoritmos genéticos, ya que es la herramienta que se emplea en esta tesis. Dichas metodologías esencialmente admiten que el algoritmo de búsqueda incurriere eventualmente en zonas con valores paramétricos que degradan drásticamente el desempeño del modelo, puesto que dichas caídas de performance son en repetidas oportunidades ocasión de encontrar las combinaciones óptimas globales de los experimentos numéricos.

Bendita la crisis que te hizo crecer, la caída que te hizo mirar al cielo, el problema que te hizo buscar a Dios.

Padre Pio de Pietrelcina

La optimización de modelos matemáticos, tradicionalmente ha sido tomada como una de las etapas de un proceso más amplio, que se ha dado en llamar la Investigación de Operaciones, la cual se refiere a la conducción y coordinación de operaciones (o actividades) dentro de una organización. Actualmente, el campo de la optimización ha crecido tanto que puede tratarse como un campo de investigación en sí mismo, que cumple la función de desarrollar herramientas que sean útiles en un proceso más complejo, que le otorga utilidad real.

El proceso involucrado en la Investigación de Operaciones comienza por la observación cuidadosa y la formulación del problema incluyendo la recolección de los datos pertinentes. El siguiente paso es la construcción de un modelo científico (por lo general matemático) que intenta abstraer la esencia del problema real. En

este punto se propone la hipótesis de que el modelo es una representación lo suficientemente precisa de las características esenciales de la situación como para que las conclusiones (soluciones) obtenidas sean válidas también para el problema real. Después, se llevan a cabo los experimentos adecuados para probar esta hipótesis, modificarla si es necesario y eventualmente verificarla. Con frecuencia este paso se conoce como validación del modelo. A continuación, se analizan alternativas de decisión, y se procura seleccionar la solución óptima al problema planteado a partir de un criterio de jerarquía previamente establecido. Finalmente, se realiza la toma de decisión e implementación de las medidas adoptadas, en función de los resultados de la modelación.

Resumidamente casi todos los autores (Castillo et al., 2002; Zoraida Omaña, 2006) parecen coincidir con estas etapas, cuyo orden no es necesariamente el mencionado arriba, sino que las etapas se van desarrollando en paralelo y retroalimentándose unas con otras. Conviene remarcar que una de las características principales de la investigación de operaciones es la de intentar encontrar una mejor solución (llamada solución óptima) para el problema bajo consideración. Se dice una mejor solución, y no la mejor solución, porque pueden existir muchas soluciones que empaten con la mejor.

Los problemas que son objeto de la optimización pueden clasificarse de acuerdo a su linealidad o no linealidad, de acuerdo a que sean continuas, discretas o mixtas, y que tengan, o no, restricciones. En Alonso (2008) se ofrece un amplio abanico de problemas clásicos de optimización, y se presenta una descripción de la evolución de esta ciencia a lo largo del tiempo. De la descripción referida, merece destacarse particularmente que muchos problemas de optimización no pueden ser abordados por métodos exactos, ya sea, por su alto grado combinatorio o por la dificultad de generar un modelo basado en programación matemática que represente exactamente una situación real. Para situaciones de ésta naturaleza se han venido

generando desde la década de los sesenta métodos conocidos como heurísticos, capaces de encontrar soluciones de buena calidad, pero en muchos casos aproximada a la solución óptima.

En los primeros tiempos se generaron métodos orientados específicamente a la resolución de cada problema, gran parte de estos métodos fueron generados inspirándose en la resolución de problemas de fácil representación, pero de muy difícil solución como lo son: el Problema del Vendedor Viajero; el Problema de la Mochila; el Problema de los Conjuntos de Cobertura; etc. Por la naturaleza diferente de estos problemas los métodos que se generaron eran útiles apenas para el problema en el cual habían sido inspirados. Pero a partir de la década del 80, se han generado una familia de métodos conocidos como meta-heurísticos que ahora tienen la capacidad de ser aplicables a problemas de diversa naturaleza. Es decir, una misma plantilla algorítmica puede ser utilizada para resolver problemas que provienen de diversos sectores.

Los métodos meta-heurísticos más conocidos son: Búsqueda Tabú, Simulated Annealing, Algoritmos Genéticos, Redes Neuronales, Métodos basados en Trayectoria de Hormiga y Métodos basados en Inteligencia Artificial. En esta tesis, sólo se profundiza el método de los Algoritmos Genéticos (AG). Una explicación detallada de cualquiera de estos métodos, pueden encontrarse en la literatura especializada (Palahí y Pukkala, 2004; Kuri y Galaviz, 2002, Viveros et al., 1998, entre otros).

El formato general del problema de optimización abordable mediante métodos heurísticos es el siguiente: La función $f(x)$, es una función matemática que se desea minimizar y las soluciones posibles deben pertenecer a un espacio W , los métodos de búsqueda heurística recorren el espacio W tratando de identificar la solución que genera el mejor valor para $f(x)$, el menor en el caso de minimización o el mayor en el caso de maximización.

6.1. Algoritmos genéticos

Dado que el método de optimización empleado en esta tesis, MOCOM-UA, está basado en la técnica de los Algoritmos Genéticos (AG), resulta necesario presentar una breve descripción de la teoría que subyace a este método heurístico, así como el léxico que predomina en su conformación.

Los algoritmos genéticos (AG) fueron desarrollados por John Holland, sus colegas, y sus estudiantes en la Universidad de Michigan. Los principales logros de sus investigaciones fueron: (1) lograr una explicación abstracta y rigurosa de los procesos de adaptación de los sistemas naturales, y (2) diseñar sistemas artificiales que retengan los mecanismos importantes de los sistemas naturales. El foco principal de investigación de los AG ha sido la robustez, y el balance entre la eficiencia y la eficacia necesaria para sobrevivir en diferentes ambientes.

Según establece Goldberg (1989), los algoritmos genéticos son algoritmos de búsqueda basados en los mecanismos de selección natural y genética natural. Estos combinan, continúa Goldberg (1989), la supervivencia del más apto entre estructuras de cadenas (que representan distintos individuos) con alguna otra estructura de información cambiante de manera aleatoria, para formar un algoritmo de búsqueda con ciertos instintos de búsqueda innovadores respecto de la búsqueda humana.

En cada generación, se crea un nuevo juego de individuos artificiales (cadenas) que emplean algunos elementos y piezas que eligen de entre los más aptos de los antiguos individuos, probando ocasionalmente alguna pieza nueva que pudiera resultar buena. Aunque es una técnica aleatoria, aclara Goldberg (1989), los AG no son un simple camino aleatorio, sino que exploran eficientemente toda la información histórica para especular en el siguiente punto de búsqueda con una performance mejorada.

6.1.1. Introducción a los Algoritmos Genéticos

Recientemente, un grupo de ingenieros que diseñaba un submarino se percató de que el pingüino posee una forma tal que, al fluir el agua a su alrededor, prácticamente no genera turbulencias en las que se pierda la energía invertida por el ave al nadar y al mismo tiempo, posee un volumen cercano al máximo, lo que permite al pingüino acumular gran cantidad de grasa que lo aísla del frío exterior y constituye su reserva de energía (Kuri y Galaviz, 2002). Al moverse, una hormiga deposita una sustancia denominada feromona, marcando el camino que fue recorrido. Las siguientes hormigas detectan la feromona ya depositada y deciden con alta probabilidad seguirla. El camino más corto recibirá mayor cantidad de feromonas por unidad de tiempo, y pronto todas las hormigas estarán recorriendo la trayectoria óptima que une el nido con el alimento (Viveros et al., 1998). Ejemplos como estos del pingüino y las hormigas, abundan en la bibliografía (Hincapié et al., 2004). A medida que el hombre va conociendo el mundo que lo rodea, comprende que los mecanismos de optimización están al alcance de nuestros sentidos, sólo basta con interpretarlos.

Fabricius ab Aquapendente, afirmaría alrededor del año 1600, anticipándose a Darwin: “La naturaleza perpetúa aquello que resulta mejor.” (De Motu Locali Animalium). Casi un siglo después, tal vez una de las mentes más brillantes que ha dado a luz la humanidad, el alemán Gottfried Wilhelm Leibniz, diría en su Discours de Metaphysique, refiriéndose a la perfección de la naturaleza, que “[...] cuanto más informados e iluminados estemos acerca de las obras de Dios, más inclinados estaremos a encontrarlas excelentes y totalmente conformes a cuanto se hubiera podido desear”. Tal como lo expresa Leibniz, la naturaleza es una gran máquina que se ocupa incansablemente de perfeccionar las especies que en ella viven. El 24 de noviembre del año 1859, Darwin publicó su famoso y revolucionario libro: El origen de las especies, selección natural, en el cual expresa que “La selección natural obra

solamente mediante la conservación y acumulación de pequeñas modificaciones heredadas, provechosas todas al ser conservado”.

Las formas vivientes que vemos hoy en día es el resultado acumulado de miles de experimentos exitosos que han ido refinando paulatinamente alguna creación primigenia (Kuri y Galaviz, 2002). Los experimentos fallidos, probablemente algunos órdenes de magnitud más numerosos que los exitosos, no los vemos. Los individuos resultantes perecieron compitiendo al lado de otros más aptos para sobrevivir. Cada individuo debe desarrollar sus capacidades de adaptación, para sobrevivir en su ambiente. Para ello debe adquirir características que le permitan alimentarse y sobrevivir a las amenazas del medio. Así, por ejemplo, el oso polar dispone de las características necesarias para sobrevivir en su entorno, que en otro medio perecería. Su pelaje blanquecino le permite confundirse con el paisaje y así capturar con mayor facilidad a su presa, su espesa capa de grasa, le permite protegerse del frío. Así como estas, existen innumerables características visibles y no visibles, que hacen a cada animal un exponente ideal para sobrevivir en su entorno. El tiempo permite a la naturaleza continuar con su proceso de perfección de las especies para adaptarlas al medio en que viven, y a las modificaciones que el mismo experimenta a lo largo de las miles y miles de generaciones.

Estas “modificaciones heredadas”, señaladas por Darwin como las generadoras de organismos mejores, son llamadas mutaciones hoy en día y constituyen el motor de la evolución. Un organismo mutante ha sufrido una modificación que lo hace diferente al resto de sus congéneres. Esta modificación puede ser un inconveniente para él (la falta de algún miembro útil de su cuerpo, por ejemplo), en cuyo caso perecerá por incompetente frente a la lucha por el alimento, o bien a la lucha por escapar de un depredador hambriento. Pero puede ocurrir también que le confiera alguna cualidad que le permita sobrevivir más fácilmente que al resto de los individuos de su especie. Este organismo tendrá mayor probabilidad de reproducirse

y sus descendientes heredarán la característica que le dio ventaja. Con el tiempo, gracias a la competencia, los organismos que en un principio eran raros se volverán comunes a costa de la desaparición del “modelo anterior” (Kuri y Galaviz, 2002). Se habrá dado entonces un paso en el proceso evolutivo.

Cada individuo de cada especie que habita el planeta tiene características que lo identifican, como su color de ojos, de pelo, estatura y tipo sanguíneo, por ejemplo. Estas características “externas”, aunque algunas de ellas no se puedan ver, como el tipo sanguíneo, constituyen lo que se denomina el fenotipo de un individuo. Cada una de estas características es igual a la correspondiente de alguno de los antecesores del individuo, es decir, nos son dadas por herencia, o por lo menos nos es dada cierta predisposición a ella (como la diabetes, por ejemplo). El fenotipo es el resultado de la interacción del medio ambiente en que se desarrolla un individuo y la herencia que éste recibe de sus ancestros (Kuri y Galaviz, 2002). De una generación a la siguiente, el fenotipo estará dado exclusivamente por la herencia que recibe de sus padres, ya que los genes de estos, serán los que determinen las características externas de sus hijos. Pero al cabo de un gran número (miles o millones) de generaciones el medioambiente habrá intervenido lenta pero decididamente a la modificación sensible del fenotipo del individuo.

Evidentemente las características externas de un individuo no se modificarán de una generación para otra. De ninguna manera se piense que la adaptación de las especies al medio se produce durante la vida del individuo, de tal suerte que una variación genética durante su vida, fue heredada posteriormente por su hijo que luego la continúa desarrollando y, a su vez, la transmite a su hijo, cadena que continúa hasta que las especies logren la característica que requiere para sobrevivir en el medio. No es así. La teoría evolutiva de las especies sostiene que la mutación se produce en el proceso de la gestación de la nueva especie, y con modificaciones infinitesimales. Cada vez que se produce la gestación de un nuevo individuo, existe una pequeña

probabilidad de que se genere algún error en la codificación genética. Ese error, que tiene bajas probabilidades de ocurrir, producirá en el individuo engendrado alguna característica que no es propia ni del padre ni de la madre, sino una malformación genética, que podrá ser negativa, en caso de que merme sus probabilidades de subsistencia en el medio respecto de los demás individuos de la misma especie. Pero también puede ser positiva, en caso de que lo beneficie.

Si fuese negativa, existen altas probabilidades de que ese individuo muera sin tener descendencia, o sea sin transmitir esa anomalía a la próxima generación. Pero si fuese positiva, ese individuo sobrevivirá con mayor probabilidad que los demás individuos de la misma especie, pudiendo desarrollarse y procrear. El hijo de este individuo podrá heredar la característica que le da ventaja respecto de sus pares, y con mayor probabilidad sobrevivirá y procreará, transmitiendo nuevamente a sus hijos esta característica. Al cabo de unas cuantas generaciones, ya serán varios los individuos que podrán haber heredado esa característica que le brinda superioridad. Ese proceso continuará hasta que la totalidad de los individuos hayan adquirido esa característica. Entonces se habrán extinguido los que no la tienen por falta de adaptación al medio frente a sus semejantes (Kuri y Galaviz, 2002).

Un nuevo error en ventaja, dará lugar a un nuevo aporte para mejorar la especie, proceso que continúa hasta que la especie ha optimizado esa característica. Todas las características de los individuos se van perfeccionando en paralelo con las innumerables gestaciones que se producen al cabo de miles de millones de generaciones de esa especie. Así se irá llegando a una generación de individuos todos casi perfectamente aptos para el medio en el cual les ha tocado desarrollarse.

Esta cualidad del proceso natural de la evolución (generar organismos óptimos sobre los que influyen infinidad de variables), llamó la atención de algunos científicos en las décadas de los cincuenta y sesenta. Un alemán, de apellido Rechenberg, introdujo en 1965 lo que denominó *evolutionsstrategie*, o estrategias evolutivas, como un

método para optimizar funciones de varias variables que definían dispositivos tales como perfiles de alas de avión. En 1966 los investigadores Fogel, Owens y Walsh se dieron a la tarea de dejar evolucionar máquinas de estados finitos sometiendo sus diagramas de transición a cambios aleatorios (mutaciones), creando con ello lo que se conoce como programación evolutiva.

También en los sesenta, John Holland, junto con algunos de sus colegas y alumnos de la Universidad de Michigan, desarrolló lo que se conoce como algoritmos genéticos (AGs). Recién a mediados de los años setenta se produce la “explosión” de los algoritmos genéticos a través del conocido trabajo de Holland (1975) denominado: “Adaptation in natural and artificial systems”. En la actualidad, los AGs son preferentemente utilizados como métodos de búsqueda de soluciones óptimas que simulan la evolución natural y han sido usados con éxito en la solución de problemas de optimización combinatoria, optimización de funciones reales y como mecanismos de aprendizaje de máquina (machine learning). Esto último les ha ganado un lugar en el campo de la inteligencia artificial (Kuri y Galaviz, 2002).

6.1.2. Conceptos básicos de biología

A continuación, se abordan algunos conceptos básicos de la genética que son estrictamente necesarios para situar al lector en el contexto preciso en el cual se gestan los algoritmos genéticos. El texto presentado posee limitaciones, ya que de ninguna manera se pretende en esta breve sección describir completamente el complejísimo y riquísimo misterio de la vida y consecución de las especies vivientes.

El cuerpo humano está formado por tejidos, los cuales se componen de millones de células. Cada célula del cuerpo humano contiene toda la información genética de esa persona. Hay dos tipos de células básicamente: las somáticas, que son células diploides, denominadas así porque poseen la dotación completa del número de cromosomas de la especie; y los gametos o células sexuales, que son células haploides, porque poseen la mitad de la dotación. En el caso del ser humano, las

células diploides poseen 23 pares de cromosomas, mientras que las haploides, sólo 23 cromosomas. Un cromosoma no es otra cosa que una larga molécula de ADN en su máximo estado de empaquetamiento (o compresión). El ADN es una molécula muy compleja de la cual -a los fines de esta tesis- sólo interesa destacar que allí está codificada la información genética de la persona. Cada cromosoma posee innumerables características de la persona; entre los 23 cromosomas incluyen la totalidad de las características de la persona (Alberts et al., 2002).

En las células diploides, cada par de cromosomas, está formado por uno heredado de la madre y otro heredado del padre. En cada par de cromosomas, la información de ambos es respecto de la misma característica, sin embargo, esta información puede ser diferente entre sí. Por ejemplo, si en un par de cromosomas está almacenado el color de ojos de la persona (entre otras características), entonces uno de los cromosomas puede indicar color azul (por ej. el de la madre), y el otro negro (el del padre). De la confrontación entre la información que tiene cada cromosoma respecto de cada característica, surgen las características dominantes y las recesivas. Siguiendo con el ejemplo, si se confronta el color de ojos azules que aporta la madre, con el negro del padre, triunfará el negro por orden de jerarquía. Es importante aclarar que el hecho de que el padre haya aportado negro no significa que tenga ojos negros, ni la madre, azules.

La molécula de ADN (Ácido Desoxirribonucleico) está formada por cuatro distintos compuestos más simples llamados bases o nucleótidos: adenina (A), guanina (G), citosina (C) y timina (T). Cada subcadena de tres nucleótidos codifica un aminoácido diferente. A las subcadenas de tres nucleótidos se les llama codones. Al unirse varios codones, forman una proteína. Al conjunto de nucleótidos que codifican una proteína completa se les llama genes. El valor que posee un gen determinado se denomina alelo (Alberts et al., 2002). Por ejemplo, un gen puede tener el valor “azul” para la característica “color de ojos”.

Cuando en un par de cromosomas de una célula diploide, ambos tienen el mismo alelo se dice que son homocigos, si no, son heterocigos, y en este último caso sólo uno de los alelos se manifestará en el fenotipo. El alelo que se manifiesta se denomina dominante, y su homólogo que no se manifiesta, se llama recesivo. El conjunto de todos los cromosomas, es decir, toda la información genética de un individuo se llama genoma y el conjunto de genes contenidos en el genoma, genotipo. Es éste el que determina, en buena medida, el fenotipo del individuo.

En síntesis, el fenotipo de cada individuo está determinado por las proteínas que produce, y esto a su vez está definido en la información genética de cada una de sus células.

La reproducción comienza con el encuentro entre el óvulo y el espermatozoide. Ambos son células de tipo haploides, por lo tanto, poseen solamente 23 cromosomas cada una de ellas. Al unirse en la fecundación, se fusionan el núcleo del espermatozoide con el núcleo del óvulo. De este modo, se obtiene una nueva célula denominada huevo o cigota, de tipo diploide, ya que posee los 23 cromosomas de la madre y los 23 del padre.

Esa célula es la primera expresión de vida y se denomina embrión. En ese momento se ha producido la fecundación. La célula comenzará a efectuar el proceso de mitosis, que es el fenómeno por el cual a partir de una célula madre, se originan dos células hijas con la misma información genética. Resumidamente, la mitosis comienza con la duplicación de la información genética de cada cromosoma, y luego se parte en dos células que se llevan, cada una de ellas, uno de cada cromosoma duplicado, logrando así dos células exactamente iguales a la célula que le dio origen. Luego, se producirá nuevamente la mitosis en cada una de las dos células recientemente formadas, con lo cual se tendrán cuatro células iguales. En cada una de las cuatro, también ocurrirá la mitosis para volverse ocho, y luego dieciséis, y así sucesivamente.

De una célula original, se habrán logrado millones de copias exactamente iguales a ella.

Este proceso de sucesivas mitosis continuará hasta que se alcance la etapa de mórula. Hasta la mórula, el embrión tiene la totalidad de las células exactamente iguales entre sí. A partir de aquí comienza la etapa de diferenciación, en la cual se comienzan a manifestar diferentes genes en cada célula.

Cuando comienza esta etapa, algunas de estas células dan origen a las gametas. Es decir, siendo diploides aún, darán origen a la formación de células haploides, a través de un proceso denominado Meiosis. La Meiosis es el proceso de división celular mediante el cual se originan, a partir de una célula madre, cuatro células hijas con la mitad del número cromosómico que la célula madre que les dio origen.

Ese tipo de células (diploides), se denomina espermatogonia (cuando se producen en el hombre) y ovogonia (mujer). Las espermatogonias (diploides) generarán los espermatozoides (haploides), y las otras, los óvulos (o más precisamente los ovocitos secundarios, ya que el término óvulo debe aplicarse cuando el ovocito ha sido fecundado). Las etapas de la Meiosis (Kuri y Galaviz, 2002) se grafican en la Figura 6-1, y son:

1. Se observa una célula diploide que dará origen a dos haploides a través del proceso de Meiosis. Esta célula tiene los 23 cromosomas del padre, y los 23 de la madre. En este esquema sólo se representa lo que ocurre en cada par de cromosomas. Así por ejemplo se puede suponer que la línea color negro representa uno de los 23 cromosomas que aporta el padre, y la verde uno de los 23 cromosomas que aporta la madre. En cada par de cromosomas se producirá lo que se describe aquí para uno sólo. Esta célula recibe el nombre de espermatogonia (u ovogonia) y dará origen a dos gametas.
2. Se produjo la duplicación del material genético en cada cromosoma.

3. Comienza el proceso de Crossing Over, en el cual hay un entrecruzamiento entre uno de los cromosomas duplicados de la madre en contacto con uno de los cromosomas duplicados del padre de cada par. Se denomina quiasma a los puntos de contacto entre los cromosomas.
4. Ya están entrecruzados. De este modo quedará cada par de cromosomas formado por dos pares de cromátidas hermanas. Una dada por un cromosoma puro del padre y otro del padre con componentes de la madre, y la otra dada por un cromosoma puro de la madre, y otro de la madre con componentes del padre.
5. Comienza a producirse una polarización dentro de la célula que atrae a los cromosomas del padre a un polo, y a los de la madre, al otro.
6. La célula se divide en dos células nuevas, cada una de las cuales posee 23 cromosomas duplicados, dando finalización a la llamada Meiosis 1.
7. Se separan las cromátidas hermanas, dando origen cada célula recientemente formada, a dos nuevas células, con la mitad del número de cromosomas. Por lo que resultan cuatro células.
8. Se han formado las cuatro células hijas.

Es preciso remarcar que entre el paso uno y el dos, se ha producido una réplica de los cromosomas existentes. Para ello existe una enzima encargada de copiarlos, denominada ADN polimerasa. La molécula de ADN tiene forma de una doble hélice, como una escalera de caracol. La enzima abre por el medio los “escalones” de esta hélice y ensambla en cada mitad los nucleótidos que debe ensamblar (Figura 6-2).

Ocasionalmente esta enzima comete un error, que puede ser causado por radiaciones energéticas externas o sustancias extrañas.

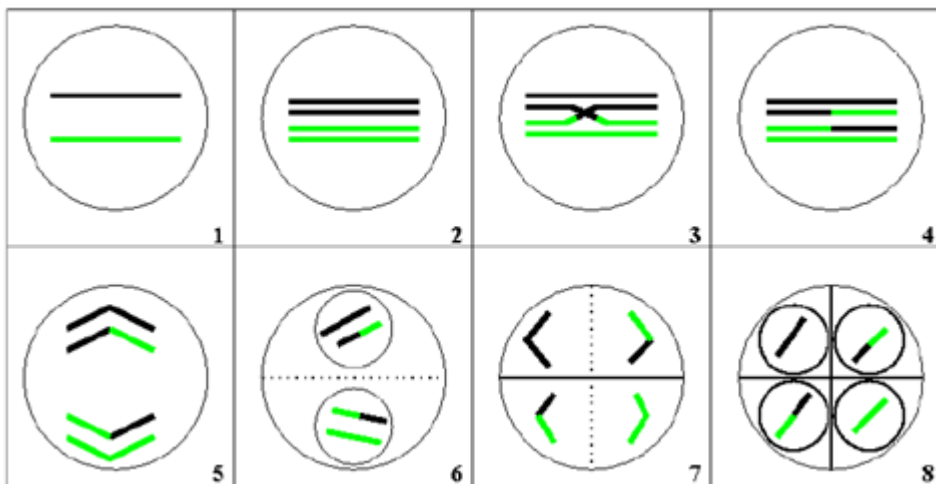


Figura 6-1: Etapas de la Meiosis (Kuri y Galaviz, 2002)

La alteración de la molécula de ADN original constituye una mutación que puede manifestarse en el fenotipo y hacer al individuo diferente del resto de sus congéneres. Es muy poco probable que cambiar al azar un trozo de información que la naturaleza ha refinado cuidadosamente a lo largo de millones de años resulte en algo bueno. Por lo general las mutaciones son desfavorables, incluso letales, para el organismo mutante. Pero ocasionalmente pueden no serlo y conferirle a dicho organismo alguna ventaja que le permita sobrevivir más fácilmente en su medio. Esta característica será transmitida a sus descendientes y un pequeño paso evolutivo se habrá dado (Kuri y Galaviz, 2002).

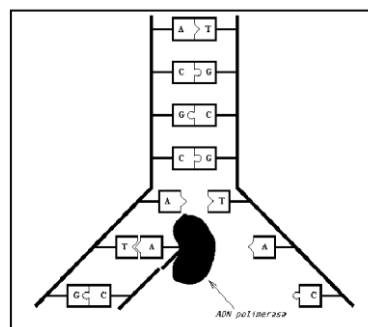


Figura 6-2: Esquema de trabajo de la enzima ADN polimerasa (Kuri y Galaviz, 2002)

6.1.1. La técnica de los Algoritmos Genéticos

Los algoritmos genéticos son algoritmos computacionales inspirados en los procesos de evolución natural. Por eso, cuando se realiza la descripción de los mecanismos de optimización basados en algoritmos genéticos, se introducen muchos términos que han sido tomados de la biología genética. A continuación, se presentan los mecanismos más destacados de los algoritmos genéticos, que tienen base en la biología, conforme fueron descritos por Kuri et al. (2002), y por Gomez Rocha et al. (2006).

Codificación del dominio

Como se ha expresado anteriormente, las características de cada individuo dependen de las proteínas. Las proteínas son codificadas por los genes, que constituyen una combinación de nucleótidos. Existe una innumerable cantidad de combinaciones posibles entre los millones de nucleótidos que forman cada molécula de ADN, cada una de las cuales sería capaz de codificar una proteína distinta. En otras palabras, las características de cada individuo responden a un código genético. A lo largo de miles y millones de generaciones, ese código genético ha ido optimizando las especies para su perfecta adaptación al medio. Se puede decir que todos los posibles códigos genéticos, son posibles soluciones a la formación de un nuevo individuo. Este conjunto de todas las posibles soluciones constituye el dominio dentro del cual el código puede tomar valores.

Del mismo modo, en un algoritmo genético lo primero que se requiere es determinar en qué espacio se encuentran las posibles soluciones al problema que se pretende resolver. Es decir, es necesario codificar de alguna manera el dominio del problema para obtener estructuras manejables que puedan ser manipuladas por el AG. Cada una de estas estructuras constituye el equivalente al genotipo de un individuo en términos biológicos. El individuo del dominio del problema al que se mapea este

genotipo es el análogo al fenotipo. Es frecuente que el código de los elementos del dominio del problema utilice un alfabeto binario (ceros y unos).

Una vez que se ha definido la manera de codificar los individuos del dominio del problema y se conoce la forma de pasar de un individuo a su código y viceversa, es necesario fijar un punto de partida. Los algoritmos genéticos manipulan conjuntos de códigos en generaciones sucesivas. Nuevamente haciendo una analogía manipulan poblaciones de códigos. En éstas un código puede aparecer más de una vez. El algoritmo se encargará de favorecer la aparición en la población de códigos que correspondan a elementos del dominio que estén próximos a resolver el problema. En resumen, el algoritmo recibirá como entrada una población de códigos y a partir de ésta generará nuevas poblaciones, donde algunos códigos desaparecerán mientras que otros, que representan mejores soluciones posibles, aparecen con más frecuencia hasta que se encuentra una satisfactoria o hasta que se cumpla alguna otra condición de terminación.

Adaptación

En una población, hay algunos individuos más hábiles para sobrevivir en el medio que otros. Estos individuos tienen mejor adaptación al medio (fitness) respecto del resto de los individuos. De la misma manera, en los AG existen códigos que resultan en individuos mejor adaptados para resolver el problema que se pretende optimizar. En general, este calificativo de adaptación se mide a través de un real no negativo. El objetivo de este número es que permita distinguir propuestas de solución buenas de aquéllas que no lo son. Si el problema a resolver consiste en minimizar una función, entonces tanto mejor calificado estará el individuo cuanto menor sea dicho valor.

Evidentemente, al hablar de que a cada individuo de la población se le asigna una y sólo una calificación, se está hablando de una función que se denomina función de

adaptación, cuya evaluación puede no ser sencilla y es, de hecho, lo que en la mayoría de los casos consume más tiempo en la ejecución de un algoritmo genético. Hay que tener en cuenta que se evalúa una vez en cada individuo de cada generación. Si un AG es ejecutado con una población de tamaño 100 durante 100 generaciones, la función es evaluada 10.000 veces. Además, puede darse el caso de que la función de evaluación no tenga una regla de correspondencia explícita, esto es, una expresión algebraica, y puede ocurrir incluso que la función cambie de generación en generación.

Selección

Una vez calificados todos los individuos de una generación, el algoritmo debe, al igual que lo hacen la naturaleza y el hombre, seleccionar a los individuos más calificados, mejor adaptados al medio, para que tengan mayor oportunidad de reproducción. De esta forma se incrementa la probabilidad de tener individuos “buenos” (con alta calificación) en el futuro.

En principio podría parecer que es conveniente tener una estrategia de selección estricta para que mejore rápidamente la población y converja el algoritmo, es decir, que la población se acumule alrededor de un genotipo óptimo. Esto no es cierto. Lo que ocurrirá es que la población se acumulará rápidamente alrededor de algún individuo que sea bueno, comparativamente con el resto de los individuos considerados a lo largo de la ejecución del algoritmo, pero este individuo puede no ser el mejor posible. A esto se le suele llamar convergencia prematura. Se habrá obtenido un óptimo local. Por esta razón, los algoritmos genéticos están diseñados para fomentar la supervivencia de los individuos aptos, pero no elimina la posibilidad de supervivencia y procreación por parte de los malos. Esta característica permite que se incursione una nueva región del dominio que podría converger hacia un óptimo local mejor que los anteriores. En la naturaleza ocurre igual. Los individuos menos favorecidos tienen menor probabilidad de sobrevivir y procrear, pero nada

impide que ocurra cada tanto e impongan sus genes, en principio, desfavorables. En los algoritmos genéticos debe buscarse en todo momento la variabilidad de individuos. Mientras que se tiende al óptimo debe permitirse que sobrevivan varios malos o poco aptos.

Cruzamiento

Durante la meiosis ocurre el proceso de producción de gametos. El código genético de los padres de un individuo se mezcla para producir gametos cuyo contenido genético es híbrido, es decir, una mezcla. De esta manera es posible que un individuo herede a sus descendientes las características mezcladas de sus propios padres, por ejemplo: el color de ojos del padre y el de cabello de la madre. La cruce de los códigos genéticos de individuos exitosos favorece la aparición de nuevos individuos que hereden de sus ancestros características deseables.

En el contexto de los algoritmos genéticos reproducirse significa que, dados dos individuos seleccionados en función de su grado de adaptación, éstos pasen a formar parte de la siguiente generación o, al menos, mezclen sus códigos genéticos para generar hijos que posean un código híbrido. Es decir, los códigos genéticos de los individuos se cruzan. Existen muchos mecanismos de cruzamiento.

Mutación

Algunas veces, muy pocas de hecho, la ADN-polimerasa (la enzima encargada de replicar el código genético), se equivoca y produce una mutación, una alteración accidental en el código genético de los seres vivos. Ocasionalmente algunos elementos del código de ciertos individuos de un algoritmo genético se alteran a propósito. Éstos se seleccionan aleatoriamente en lo que constituye el símil de una mutación. El objetivo es generar nuevos individuos, que exploren regiones del dominio del problema que probablemente no se han visitado aún. Esta exploración no presupone conocimiento alguno, no es sesgada.

Aleatoriamente se buscan nuevas soluciones posibles que quizá superen las encontradas hasta el momento. Esta es una de las características que hacen aplicables los algoritmos genéticos a gran variedad de problemas: no presuponer conocimiento previo acerca del problema a resolver ni de su dominio, no sólo en la mutación sino en el proceso total. De hecho, el problema a resolver sólo determina la función de evaluación y la manera de codificar las soluciones posibles (la semántica de los códigos genéticos de los individuos). El resto de los subprocesos que constituyen el algoritmo son independientes y universalmente aplicables.

6.2. Técnicas de calibración

Una de las etapas fundamentales para la aplicación de un modelo matemático de tipo lluvia-caudal a una cuenca dada es la calibración de sus parámetros tomando como información de ajuste los datos de caudales observados. Esta calibración es realizada modificando los valores de los parámetros hasta que el hidrograma calculado reproduzca al hidrograma observado con buena precisión. Sin embargo, un ajuste cuantitativo del modelo, dado por una minimización aceptable de la FO, no basta para considerar buena la simulación, sino que es necesario agregar la contribución del método cualitativo. Este último consiste básicamente en una evaluación visual de los resultados, que se presentan usualmente en gráficos. En esta evaluación visual del ajuste, se presta especial atención a la parte del hidrograma que presenta mayor interés para el propósito del modelo, como puede ser el pico o la curva de recesión del hidrograma, o, por ejemplo, el volumen total escurrido.

Sumado al ajuste cualitativo-cuantitativo, resulta necesaria una interpretación de los valores obtenidos para los parámetros por parte de un hidrólogo experimentado. Esto se debe a que no existe un conjunto único de parámetros capaz de representar todos los procesos hidrológicos involucrados en la cuenca, debido a las incertidumbres inherentes de los datos, a las simplificaciones introducidas en el

desarrollo de los modelos y a la limitada representatividad de los parámetros involucrados (Collischonn et al., 2001). Existen varias soluciones posibles que, de forma adecuada darán la respuesta deseada (Beven y Binley, 1992), dentro del concepto de equifinalidad. Es decir, existen varios juegos de parámetros que satisfacen las ecuaciones planteadas, debido a que el número de incógnitas supera al de las ecuaciones. Sin embargo, solo algunos de esos juegos de parámetros tendrán un sentido físico lógico en el problema. A la inversa, será posible establecer varios juegos de parámetros con sentido físico lógico en el problema, pero que sólo algunos satisfacen las ecuaciones planteadas. A la vista de los resultados, la experiencia del modelador debe dar la última palabra respecto del juego de parámetros que mejor representa las condiciones modeladas.

6.2.1. Clasificación de las técnicas de calibración

Las técnicas de calibración empleadas en modelos hidrológicos se pueden clasificar de acuerdo a los siguientes criterios (Bertoni, 2003):

- a. desde el punto de vista operativo: técnicas de calibración automática y manual.
- b. desde el punto de vista de la cantidad de objetivos que se propone optimizar: técnicas de calibración que utilizan funciones mono y multi-objetivo.
- c. desde el punto de vista del enfoque sobre el dominio analizado: técnicas de calibración local y global.

Calibración automática y calibración manual

Básicamente existen dos formas de calibrar un modelo matemático hidrológico desde el punto de vista operativo: (a) ajuste manual mediante prueba y error y (b) ajuste o calibración automática. La primera es la manera más clásica, que ha sido empleada por los hidrólogos desde los comienzos de la aplicación de las técnicas de calibración. Esta consiste básicamente en la aplicación del conocido método de

prueba y error. Cuando el modelo utilizado tiene un gran número de parámetros y el usuario tiene poca experiencia en el proceso de calibración, este proceso puede volverse demasiado lento y tedioso. Además de ello, esta técnica no brinda al usuario un criterio objetivo de parada (Sorooshian y Gupta, 1995), sino que, por lo general, usuarios diferentes generalmente encuentran un conjunto diferente de parámetros con el mismo modelo, en la misma cuenca y en el mismo periodo de tiempo (Beven y Binley, 1992).

La calibración automática consiste en el empleo de las computadoras digitales para llevar a cabo el proceso de ajuste de los parámetros. Surgió como una alternativa para acotar el grado de error de la calibración manual ejecutada por los usuarios sin experiencia, para acelerar el proceso de calibración y para restarle subjetividad (Sorooshian y Gupta, 1995). De modo que se convierte en una herramienta que intenta determinar el conjunto óptimo de los valores de los parámetros de un modelo hidrológico con la mínima interferencia del usuario. La experiencia del usuario es sustituida, en gran parte, por una función objetivo sometida a una técnica iterativa de ajuste. El objetivo de la calibración automática es, por lo tanto, minimizar (o maximizar) el valor de una función objetivo que mide la calidad del ajuste de los hidrogramas. Sin embargo, enfatizan Sorooshian y Gupta (1995), el estado del arte de los métodos de calibración automática no ha madurado al punto de sustituir enteramente a la calibración manual.

Calibración con funciones mono y multi-objetivo

Como se explicó previamente, generalmente la FO representa la diferencia entre los valores medidos y los calculados por el modelo. Pero también se aclaró que esa diferencia debía venir dada por la característica del hidrograma que se pretenda ajustar con mayor precisión, en virtud del propósito para el cual será empleado el modelo una vez que se haya calibrado. Así, por ejemplo, si el modelo hidrológico ha

sido pensado para predecir caudales máximos de diseño, la FO debería marcar la diferencia entre los picos de los hidrogramas medidos y simulados. Si el propósito del modelo hidrológico estuviese asociado al balance hídrico, convendría asegurar un correcto ajuste de volúmenes escurridos, por lo que la FO debería representar la diferencia entre el volumen escurrido medido y el simulado.

Ahora bien, cuando se desean ajustar varias características del hidrograma de salida, como caudal pico, volumen escurrido, curva de recesión, etc., es necesario proponer la minimización simultánea de varias funciones objetivos, cada una de las cuales representa la diferencia entre la variable medida y la simulada de cada una de las características del hidrograma de salida. Este tipo de ajuste se denomina multi-objetivo. En una calibración automática multi-objetivo el propósito es optimizar, al mismo tiempo, varias funciones que evalúan varios aspectos de calidad de ajuste entre hidrogramas calculados y observados.

En numerosos procesos hidrológicos resulta insuficiente el ajuste de mono-objetivo. Por ejemplo, en la simulación continua se requiere de una característica multi-objetivo para el uso del modelo hidrológico, que permita representar correctamente los periodos lluviosos y de estiaje y no simplemente de uno de ellos en detrimento de los otros. La optimización multi-objetivo es un nuevo estado del uso de estas técnicas en el ajuste de los modelos hidrológicos.

Dadas las imperfecciones de los modelos y de los datos de entrada, es prácticamente imposible encontrar un conjunto de parámetros que produzca, al mismo tiempo, valores óptimos de varias funciones objetivo diferentes. Casi siempre, la optimización adecuada de una función objetivo se logra a expensas de la inadecuada optimización de otra. Esta es la situación que normalmente se enfrenta en la calibración de modelos lluvia-caudal, cuando un conjunto de parámetros resulta en un buen ajuste de los picos de caudal, pero como consecuencia presenta resultados mediocres durante los periodos de estiaje (o viceversa).

La principal característica de un problema de calibración multi-objetivo, es que la solución, en general, no será única (Collischonn et al., 2003). Es decir, un conjunto A de valores de los parámetros, podrá corresponder al óptimo de la función F1, y un conjunto B (distinto de A), corresponderá al óptimo de la función F2. Existirán muchos conjuntos de parámetros entre A y B que no optimizan F1 ni F2, pero pueden ser considerados aceptables. Por ejemplo, si C y D son dos conjuntos de valores de parámetros entre A y B, serían considerados aceptables si, partiendo de la solución C hacia la solución D, un cambio en mejoría para una función objetivo será acompañada de un deterioro en la otra, y no existirá ninguna solución E que sea al mismo tiempo mejor que C y D en ambas funciones objetivo F1 y F2. La Figura 6-3 ilustra este concepto para el supuesto caso de un modelo de un parámetro. Como puede verse, no existe un valor de ese parámetro que minimice simultáneamente ambas FO, sino que cuando el parámetro toma el valor de 6 minimiza FO1, pero cuando toma el valor de 10, minimiza FO2. Entre 6 y 10 existe un rango de valores para el parámetro tal que, a medida que se acerca a 6, mejora FO1, pero empeora FO2, y viceversa. Si se toman dos valores en el rango entre 6 y 10, como, por ejemplo, 7 y 9 puede verse claramente que no existe un valor entre ellos que sea mejor que ambos en ambas FO. Entonces 7 y 9 son considerados aceptables. Del mismo modo, todos los valores entre 6 y 10 también lo son.

Si no hubiere ninguna información adicional que permita dar mayor importancia a una de las dos funciones objetivo (Collischonn y Tucci, 2003), todas las soluciones intermedias deben ser consideradas buenas. Debido a estas características, la solución de un problema de optimización multi-objetivo no es un punto, sino que se trata de una región de soluciones posibles, todas ellas aceptables, denominada región de Pareto, o región de soluciones no inferiores o no dominadas (Yapo et al., 1998). Más adelante, en esta tesis, se explica acabadamente en concepto de dominancia, pero a los fines de interpretar esta revisión de antecedentes conviene

adelantar el criterio de dominancia, según el cual son definidos puntos Dominados aquellos tales que existe otro punto que es mejor que ellos en todas las FO, siendo los restantes puntos No Dominados.

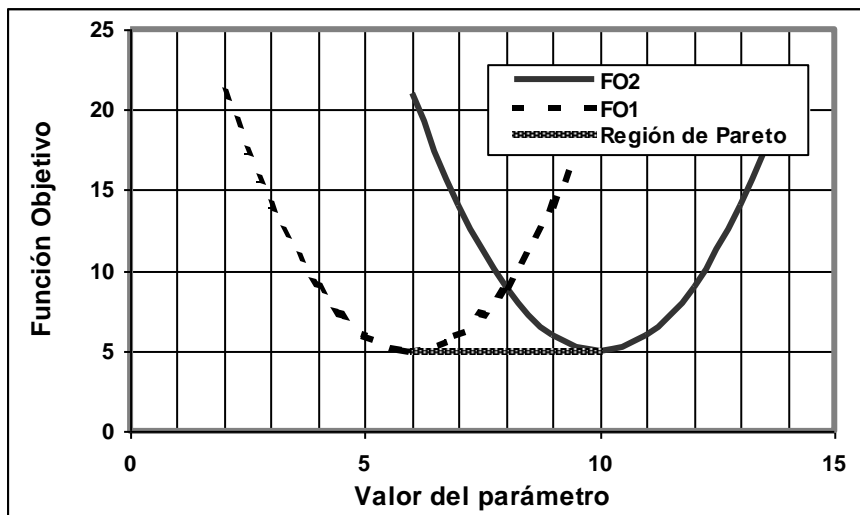


Figura 6-3: Esquema de la Región de Pareto para dos funciones objetivo

Este último enfoque responde al concepto actual en la calibración de modelos matemáticos que se orienta a buscar un conjunto de soluciones posibles y aceptables, en vez de una solución óptima. Esto se debe a que existe una gran incerteza asociada a la estimación de los parámetros del modelo. Las técnicas recientes en calibración de modelos matemáticos prefieren asignar un conjunto de soluciones posibles con una cierta probabilidad de éxito en la calibración. De este moderno concepto proviene la importancia del análisis de incertidumbre en las técnicas de calibración.

El juego de soluciones óptimas es denominado el conjunto o juego de soluciones de Pareto (en inglés se denomina "Pareto optimal set", o simplemente "Pareto set"), y se denota como \bar{P} . Mientras que la imagen de \bar{P} que se desarrolla en el espacio de la superficie de respuesta de la funciones objetivo F , recibe el nombre de Región de Pareto o Frente de Pareto (en inglés, "Pareto front"), como lo aclaran Kraube et al. (2012).

La región de Pareto es desde el punto de vista teórico un conjunto continuo de soluciones en un espacio (una región) óptimo donde se desarrolla una sucesión ininterrumpida de soluciones no dominadas. Dado que el juego de soluciones de Pareto rara vez consiste en un número finito de soluciones, la mayor parte de las técnicas tienen a identificar un número contable de distintas soluciones distribuidas dentro de la región de Pareto. De modo tal que se pueda inferir a partir de dichas soluciones la región que ocupa la solución verdadera. En Gupta et al. (1998) se explica que los métodos clásicos para encontrar este conjunto de soluciones finitas, se pueden clasificar en: métodos "a posteriori", métodos "a priori", y métodos "interactivos". En esta publicación se presentan a su vez varias publicaciones en las cuales se da detalle y ejemplos de cada uno de estos tipos de métodos.

Métodos de Calibración local y global

Dado que las técnicas de calibración automática son casos particulares de los métodos de Optimización, estas técnicas también pueden clasificarse entre local y global, de acuerdo al tipo de búsqueda que haga. El gráfico de la Figura 6-4 muestra el caso hipotético de un modelo con un solo parámetro. Como puede verse, la FO tiene varios mínimos locales. Se denomina mínimo global, al mínimo valor de los mínimos locales. En el ejemplo, los puntos A, B y C son mínimos locales, ya que en la vecindad de dichos puntos los demás son superiores a ellos, pero sólo el punto B es mínimo global, ya que es el menor de los mínimos locales, dentro del dominio de validez del único parámetro del modelo. Este concepto es fácilmente extrapolado a modelos de varios parámetros, como son la mayoría de los modelos matemáticos utilizados en hidrología.

Las primeras técnicas desarrolladas que se encuentran en la bibliografía estaban basadas en algoritmos de búsqueda local. Debido a la complejidad de los modelos hidrológicos, las funciones objetivo, son expresiones implícitas. En otras palabras, no es posible despejar los valores de los parámetros asociados al mínimo valor de la FO,

sino que resulta necesario tender a éste mínimo mediante aproximaciones sucesivas basadas en iteraciones direccionadas de prueba y error.

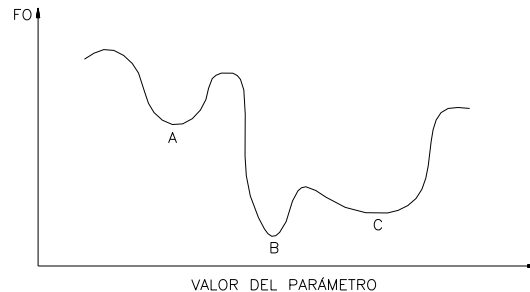


Figura 6-4: Proyección plana de la superficie de respuesta de una FO

Estos métodos iterativos de búsqueda local poseen cuatro características básicas (Tucci, 2005), que son: a) definición del punto de partida, b) dirección de la búsqueda, c) espaciamiento de cada prueba, y d) criterio de parada.

La bibliografía es pródiga en metodologías de optimización local (Beveridge y Schechter, 1979), sin embargo, sólo algunos de estos métodos, fueron preferidos para calibración de modelos hidrológicos.

Una de las primeras propuestas metodológicas de calibración local, que posteriormente fueron empleadas en calibración de modelos hidrológicos (Beard, 1966), surge de la mano de dos economistas (Friedman y Savage, 1947) con el método Univariacional. Para ejemplificar este método resulta ilustrativo suponer un modelo con dos parámetros y luego extrapolar la idea a n parámetros.

Supóngase que se tienen dos parámetros X_1 y X_2 . La superficie de respuesta de la FO se desarrolla en tres dimensiones, dos de las cuales son los valores de ambos parámetros del modelo, siendo la tercera, el valor de la FO para cada par de valores $FO=f(X_1, X_2)$. El método se inicia en un punto de partida elegido por el usuario, y luego se elige una dirección de búsqueda, por ejemplo, la dirección de X_1 . Se realiza un paso DX_1 , en algún sentido del eje X_1 . Si en la nueva ubicación se obtiene un valor de FO menor que la anterior, realiza un nuevo paso en este sentido, de lo contrario,

prueba con un $DX1$ en sentido contrario. Si en esta nueva ubicación la FO es inferior continúa en este sentido. Una vez que se ha verificado que para la dirección $X1$, en cualquier sentido se obtiene un valor superior al actual, entonces se comienza la búsqueda a lo largo de $X2$, para lo cual se elige un sentido de búsqueda y se da un paso $DX2$. Nuevamente se procura reducir el valor de FO, de lo contrario se invierte el sentido.

En cualquier caso, se continúa el sentido de búsqueda siempre que la FO siga reduciéndose con cada nuevo paso, y se invierte en caso de que aumente. Cuando en ambos sentidos FO verifica un aumento, se toma la siguiente dirección. Así se continúa hasta que se agoten todas las direcciones, posterior a lo cual se dice que se ha cumplido un ciclo, y luego se comienza con la primera dirección nuevamente. En este ejemplo hemos supuesto dos direcciones, pero si se tratara de un modelo de n parámetros, tendríamos un espacio hiperdimensional con n ejes ortogonales entre sí. Si bien no puede graficarse este espacio, la idea del método es exactamente igual. Así se continúa, hasta que la FO no perciba reducciones en ningún sentido de las n direcciones, o bien, las reducciones sean aceptablemente despreciables para el caso en estudio.

El método Univariacional tiene una estructura muy convincente pero poco eficiente, ya que no tiene ninguna dirección preferencial de búsqueda. Simplemente prueba en cada una de las direcciones correspondientes al sistema de coordenadas de referencia de la superficie de respuesta. Rosenbrock (1960) propuso el método de Rotación de Coordenadas, que comienza del mismo modo que el método Univariacional, hasta el final del primer ciclo. Posteriormente, traza una recta imaginaria entre el primer punto y el último del primer ciclo, y la establece como dirección preferencial de búsqueda para el parámetro $X1$, ajustando las restantes direcciones a esta rotación. Las pruebas son realizadas del mismo modo en cada ciclo, provocando una rotación de los ejes de coordenadas al cabo de cada uno. Otra

diferencia con respecto al método Univariacional es la incorporación de un par de coeficientes de aceleración (A valor mayor que 1) y retardo (B valores entre 0 y 1). Estos coeficientes de aceleración y retraso se multiplican por el valor del espaciamiento del paso de avance, de acuerdo a la secuencia de éxitos (reducción de la FO) o fracasos (aumento de la FO) que experimente el avance en cierta dirección.

Ibbitt (1970) modificó el procedimiento de búsqueda en cada dirección del método de rotación de coordenadas. El criterio adoptado incluye una aproximación cuadrática para estimar la tendencia de minimización de la superficie de respuesta. Una explicación más detallada puede encontrarse en Tucci (2005).

El método simplex (Nelder y Mead, 1965), dio un paso considerablemente superior en la eficiencia de búsqueda, ya que este incursiona libremente en cualquier dirección del espacio hiperdimensional de validez de los parámetros. En otras palabras, no se limita a las direcciones de los ejes de coordenadas de los parámetros, sino que elige la dirección en la cual parece más probable que se minimice la FO. Además de eso, no parte de un punto, sino de un conjunto de puntos, lo que permite un direccionamiento más preciso y eficiente en la búsqueda. Este algoritmo, que no debe confundirse con el célebre método simplex para programación lineal, se explica en detalle en el capítulo referido a la Metodología.

Todos estos métodos de optimización local, dependen fuertemente del punto de partida o conjunto de puntos elegidos inicialmente (Johnston y Pilgrim, 1976), y por lo tanto caen comúnmente en punto óptimos locales. Diferentes puntos de partida, podrían conducir a diferentes óptimos locales, lo que genera diferentes juegos de valores de los parámetros del modelo, aun cuando se utilice la misma serie de datos en la misma cuenca, para todos los diferentes puntos de partida. Esto se debe a que las superficies de respuesta de las funciones objetivos no siempre son suaves y continuas, sino que presentan discontinuidades, y varios óptimos locales (Duan et

al., 1992). Un complemento a estos métodos de búsqueda local podría estar dado por una iniciación múltiple del método, en diferentes puntos de partida, y la selección del mínimo de los mínimos locales encontrados.

Sorooshian y Gupta (1995) probaron la eficiencia del algoritmo Simplex, mediante una implementación de múltiples partidas, denominado MSX. En esta prueba quedó demostrado que ni siquiera este algoritmo de búsqueda local con múltiples partidas, fue capaz de encontrar eficazmente el óptimo global de una superficie de respuesta. Inicializando 50 veces para cada optimización, el algoritmo MSX no fue capaz de encontrar el óptimo global de los parámetros del modelo Sacramento (Burnash, 1995) en ninguna de las diez pruebas realizadas.

Dadas las limitaciones de las metodologías de búsqueda local, surgen las técnicas de optimización global, que serán tratadas a continuación.

6.2.1. Técnicas de optimización global

Actualmente existe una innumerable producción técnico-científica en el campo de la optimización global. Se realizan todos los años decenas de congresos y eventos científicos relativos a la optimización, que pueden encontrarse en múltiples páginas de Internet dedicadas al tema exclusivamente. Una larga lista de revistas, focalizadas en la optimización, clasifican los trabajos más relevantes de entre los centenares que reciben diariamente. Todos estos trabajos se difunden entre las más variadas áreas del conocimiento. Una gran variedad de aplicaciones en el campo del diseño, ciencias de decisión, e investigación de operaciones han sido formuladas como un problema de la optimización global (Wang, 2001). Tales aplicaciones incluyen redes neuronales, señal digital y procesamiento de imágenes, optimización estructural, ingeniería de diseño, diseño asistido por computadora, operación y diseño mecánico, control de proceso del producto químico, etc. En el campo de la ingeniería hidráulica, las principales aplicaciones están asociadas a la optimización en el aprovechamiento de un embalse o sistema de embalses (Bravo, 2006), calibración

de modelos hidrológicos conceptuales para el pronóstico de crecidas a tiempo real o predicción de crecidas extraordinarias (Collischonn y Tucci, 2003); planificación de abastecimiento de agua potable a localidades mediante redes de agua (Sherali y Smith, 1997; Shoemaker et al., 2006; Abebe y Solomatine, 1998); y optimización de plantas de tratamientos de aguas y efluentes (Moles et al., 2003).

En sus versiones primitivas, estos métodos consistían esencialmente en la aplicación sucesiva de la técnica de búsqueda local. Surgen los Métodos de Búsqueda Aleatoria (Random Search, RS). Entre otros, se comienza a desarrollar la técnica de Búsqueda Aleatoria Uniforme (URS) que consiste en asignar una distribución probabilística uniforme para cada variable a calibrar. Como ejemplo, se puede suponer dos parámetros dados X_1 y X_2 a calibrar en el modelo. Los mismos se van arrojando aleatoriamente según una distribución probabilística uniforme para cada uno y luego con cada par X_1 , X_2 se evalúa la función objetivo. A continuación, se comienzan a aplicar las técnicas de optimización local en cada punto X_1 , X_2 . Una vez obtenido un óptimo para cada par X_1 , X_2 se elige el óptimo entre éstos, que corresponde al óptimo global. Brazil (1988) propone la Técnica de búsqueda aleatoria adaptativa, que asigna una distribución de probabilidades a cada variable. Por lo demás, es igual al método anteriormente descrito. Kuczera (1997), aporta el Método de búsqueda local con múltiples partidas (o método combinado – Múltiples start local search, MSLS), que utiliza métodos aleatorios para fijar los puntos de comienzo de búsqueda.

Dando principio a una nueva generación de métodos de optimización, Holland (1975), formaliza los principios de la técnica de los Algoritmos Genéticos (Genetic Algorithm, GA) que han sido tratados en sección aparte.

6.3. Evolución de los algoritmos de optimización para modelos P-Q

La superficie descrita por la FO en el espacio de validez de los parámetros, recibe el nombre de *superficie de respuesta*. Un algoritmo de optimización es un procedimiento lógico empleado para buscar dentro de la superficie de respuesta, restringida a los rangos permitidos para los parámetros, el juego de valores de los parámetros del modelo que optimizan (minimizan o maximizan, según corresponda) el valor numérico de la FO (Sorooshian y Gupta, 1995). El procedimiento es típicamente implementado en una computadora digital para permitir una búsqueda rápida.

Uno de los primeros antecedentes que se encuentran en la bibliografía, respecto del empleo de las técnicas de optimización antes citadas para llevar a cabo la calibración automática de modelos hidrológicos, se remonta a la década del 60. Beard (1966) empleó la calibración automática para ajustar el modelo hidrológico HEC1, predecesor del conocido modelo hidrológico HEC-HMS. Ibbitt (1970) empleó la calibración automática para ajustar un modelo hidrológico para diferentes combinaciones de series con y libre de errores, y estudió los efectos de los errores sobre el ajuste. Posteriormente, (Ibbitt, 1972) utilizó un modelo conceptual para generar datos de caudales sintéticos libre de errores de información de precipitación y evaporación potencial. Luego introdujo errores en las tres series de datos. Mediante ajuste automático y objetivo, ajustó el modelo para diferente combinación de series libre de errores y con errores, y estudió los efectos de los errores sobre el ajuste. Encontró que los valores del criterio de ajuste dependen fuertemente de los errores en los datos de caudal, ya que las variaciones en los valores finales de los parámetros para el ajuste a los datos con errores en las otras dos series mostraron no ser tan grandes como a los datos libres de errores. Johnston and Pilgrim (1976) organizaron la información dispersa hasta el momento en materia de calibración automática de modelos hidrológicos y presentaron una serie de recomendaciones.

Clarke (1973) se introduce en la representatividad de la FO respecto de la bondad del ajuste. Sugiere que la estimación de los parámetros del modelo mediante una función objetivo de mínimos cuadrados, representa una aplicación de la conocida teoría de regresión no lineal, a situaciones en las cuales las simplificaciones normalmente realizadas en esta teoría, son rara vez válidas. Propone como una corrección a esta limitación, el uso de funciones no necesariamente basadas en la suma de cuadrados, sino de otras más realistas basadas en las características estocásticas de los residuos del modelo. Agrega que la interdependencia entre los parámetros del modelo, necesita un estudio más detallado, de la superficie de respuesta de la función objetivo de la suma de cuadrados, en la vecindad del mínimo, aun cuando las hipótesis de regresión sean válidas.

Al respecto, Sorooshian y Farid (1982) estudiaron el rol del análisis de sensibilidad de los parámetros para estudio de post-calibración de los modelos conceptuales de transformación lluvia-caudal. Para esto, desarrollaron dos medidas de sensibilidad, la concentricidad y la interacción, las cuales caracterizan la forma de la superficie de respuesta de la función objetivo en la vecindad del punto de convergencia. Cuando ambas medidas son empleadas conjuntamente, ofrecen alguna información útil relativa a la geometría de la superficie de respuesta en subespacios de dos dimensiones. Demostraron el rol que desempeñan estas medidas para el caso de un modelo sencillo de cuatro parámetros, empleando una serie sintética de datos. Sorooshian y Gupta (1983) examinaron las razones para la incapacidad de obtener un juego único y conceptualmente realista de parámetros en modelos conceptuales. A partir de estudios sobre la superficie de respuesta, demostraron que la elección de una FO que explique algunas de las propiedades estocásticas de los errores de la salida del modelo redundaba en una mejor y más suave forma de la superficie de respuesta. Luego, son mejoradas las chances de lograr un juego de parámetros único y más realista. Sin embargo, su estudio indica que parte del problema es también

por estructura inadecuada del modelo. Ilustran sus resultados con el modelo de conteo de humedad acumulada en el suelo del sistema de pronóstico de ríos del National Weather Service.

Los errores en los modelos conceptuales PQ complejos pueden ser analizados mediante ubicación de ellos dentro de un marco estadístico. Esto tiene importancia en el sentido de tratar los errores como variables aleatorias y definir la estructura probabilística de los errores. Mediante el empleo de tal marco, un amplio paquete de técnicas, muchas de las cuales han sido presentadas en la literatura estadística, se vuelven disponibles para el modelador para analizar y cuantificar las diferentes fuentes de error. Troutman (1985) revisa algunas de esas técnicas, con especial atención a las peculiaridades de los modelos hidrológicos. Aplica particularmente un número de metodologías conocidas para estimación (calibración) de parámetros para obtener estimaciones con sentido físico y para explicar de qué modo el sesgo en la predicción de los caudales causados por los errores del modelo y los errores de los datos de entrada pueden contribuir al sesgo en la estimación del parámetro. Sorooshian y Gupta (1985) estudiaron los conceptos y herramientas necesarias para evaluar si la estructura de un modelo hidrológico es tal que sus parámetros pueden ser claramente identificados. Presentan una definición de la identificabilidad estructural que está basada puramente en las propiedades del modelo y es independiente de la naturaleza estocástica de los errores observados de salida del modelo. Utilizan este enfoque para desarrollar una medida de identificabilidad estructural local, la cual es útil en el análisis de sensibilidad de los parámetros del modelo. Demuestran la utilidad de los conceptos discutidos en ese paper, mediante un ejemplo de simulación utilizando un modelo conceptual lluvia-caudal.

Hendrichson et al. (1988) llevan a cabo un examen de los aspectos de la calibración de los modelos conceptuales de transformación lluvia-caudal, empleando el modelo Sacramento (de almacenamiento de humedad en el suelo) y comparan el algoritmo

del tipo de Newton con el de optimización de búsqueda directa. Los resultados indican que el de búsqueda directa es más robusto debido a la sensibilidad que presenta el de Newton a condiciones pobres de la superficie de respuesta. Estudios gráficos de la superficie de respuesta del espacio de los parámetros del modelo confirmaron la presencia de discontinuidades y superficies ásperas, particularmente en las derivadas. Wang (1991) emplea algoritmos genéticos para calibrar un modelo conceptual lluvia caudal de siete parámetros para datos de una cuenca en particular. Sus resultados demuestran que el algoritmo genético puede ser fuerte y robusto.

Beven y Binley (1992) describen una metodología para la calibración y la estimación de la incertidumbre de los modelos distribuidos. El procedimiento GLUE (siglas en inglés: Generalized Likelihood Uncertainty Estimation) trabaja con múltiples juegos de valores de parámetros y permite que, dentro de las limitaciones de la estructura de un cierto modelo y los errores en las condiciones de borde y observaciones de campo, diferentes sets de valores puedan ser igualmente probables como simuladores de una cuenca. Ilustran la metodología mediante una aplicación del Institute of Hydrology Distributed Model a datos de la cuenca experimental GWY en Plynlimon, mid-Wales.

Duan et al. (1992) presentan claramente la naturaleza de los problemas que asocia a la calibración la existencia de múltiples óptimos locales, en el intento de calibración del modelo hidrológico PQ denominado SIXPAR. Estos resultados sugieren que la optimización de los modelos PQ es más difícil de lo que se pensaba previamente y que los procedimientos corrientes de búsqueda local tienen una muy baja probabilidad de éxito a la hora de encontrar el juego óptimo de parámetros. Por consiguiente, evalúan la performance de tres algoritmos de búsqueda local aplicados sobre el modelo SIXPAR. Finalmente, presentan un nuevo procedimiento de optimización global, denominado SCE-UA (Shuffled Complex Evolution method), el cual fue capaz de encontrar consistentemente el óptimo global del modelo SIXPAR,

y aparece como capaz de resolver efectiva y eficientemente el problema de optimización de un modelo PQ.

Sorooshian, Duan y Gupta (1993) enfrentan la tarea de probar el algoritmo SCE-UA. Investigan respecto de la consistencia con que dos métodos, el SCE-UA y el método Simplex Multistar (MSX), son capaces de encontrar el juego de parámetros óptimo durante la calibración del modelo Sacramento (SAC-SMA), del National Weather Service River Forecast System (NWSRFS). En la primera fase de este estudio emplean datos sintéticos libres de errores para conducir el estudio a un plano comparativo entre los algoritmos bajo condiciones ideales. En 10 pruebas independientes de cada algoritmo en los que fueron optimizados simultáneamente 13 parámetros del modelo SAC-SMA, el algoritmo SCE-UA logró 100% de éxito en cuanto a la velocidad de ubicación del óptimo global correcto (es decir, de los verdaderos valores de los parámetros), mientras que el MSX falló para todas las pruebas, aun duplicando el número de evaluaciones de la FO. En la segunda fase emplearon datos históricos de escurrimiento superficial del Leaf River, para conducir la evaluación comparativa de ambos algoritmos bajo condiciones reales, empleando dos criterios de estimación diferentes, DRMS y HMLE. Encontraron que el algoritmo SCE-UA obtuvo consistentemente valores menores de la función y estimaciones de la función mejor agrupados, aun empleando un tercio menos de evaluaciones de la función que el algoritmo MSX. Duan et al. (1994) indican que experiencias con el modelo han demostrado que la efectividad y eficiencia de los algoritmos están influenciadas por la selección de los parámetros algorítmicos. Entonces, presentan criterios de selección de valores adecuados de los parámetros del algoritmo SCE-UA para calibración automática de modelos hidrológicos, sobre la base de resultados obtenidos de varios estudios experimentales en donde el modelo NWSRFS-SMA, empleado por la National Weather Service para el pronóstico de crecidas, fue calibrado usando diferentes parámetros algorítmicos.

Gan y Biftu (1996) probaron, con cuatro modelos conceptuales aplicados a ocho cuencas (32 combinaciones), la calidad de calibración de tres métodos de optimización: SCE-UA, Múltiple Start Simplex (MSX), y Simplex local. Aún con semejantes estructuras entre los modelos, algunos juegos de parámetros para las mismas cuencas dieron valores sorprendentemente diferentes deducidos por los tres métodos de optimización. Adicionalmente, el juego de parámetros del SCE-UA o el MSX, los cuales normalmente producen valores marginalmente mejores que el Simplex local en la etapa de calibración, podían terminar con peores valores en la etapa de validación. Aparentemente, dada la inherente limitación de los datos de calibración, las inadecuadas estructuras del modelo, y la identificabilidad de los problemas, es imposible lograr la convergencia global en la búsqueda de los parámetros. Sin embargo, fuera de las cuencas secas, tales como Ihimbu o Bird Creek, el juego de parámetros obtenidos es generalmente adecuado. Tanto el SCE-UA, como el Simplex local son herramientas de optimización viables, mientras que el MSX es ineficiente computacionalmente. El SCE-UA puede completar la búsqueda del parámetro en una sola corrida, mientras que el método Simplex local a menudo requiere varias operaciones para obtener buenos resultados.

Cooper et al. (1997) estudiaron la performance de tres técnicas probabilísticas de optimización para la calibración automática del modelo conceptual PQ Tank Model. Estos métodos fueron el SCE-UA, Algoritmos Genéticos (GA) y Simulated Annealing (SA). Encontraron que la performance dependió de la elección de la FO considerada, y también de la posición del comienzo de la búsqueda en relación al óptimo global. De los tres métodos de optimización global, el SCE produjo mejores estimaciones de la solución óptima que los métodos GA y SA. Con respecto a la eficiencia, expresada como el número de iteraciones necesarias para converger, el ranking en orden decreciente de performance fue SCE, GA y SA.

Todos estos métodos son muy costosos en términos computacionales, sobre todo cuando la búsqueda es conducida sobre un gran hipercubo. Más aún, las especificaciones del hipercubo puede ser problemática, particularmente si hay una fuerte interacción entre los parámetros. Kuczera (1997) busca reducir el esfuerzo computacional mediante confinamiento de la búsqueda al subespacio dentro del cual el óptimo global es probable de ser encontrado. La técnica incluye ubicación de un óptimo local empleando una búsqueda orientada por un gradiente de óptimo local. Presenta un caso de estudio que involucra un modelo conceptual lluvia-caudal de cinco parámetros. Demuestra que esta búsqueda estratégica tiene mayor eficiencia (normalmente dos, pero llega hasta 5 veces de mayor eficiencia) que la búsqueda sobre el hipercubo. Compara cuatro algoritmos de búsqueda probabilística: SCE-UA, algoritmo genético empleando el crossover tradicional, y múltiples inicializaciones empleando los algoritmos de búsqueda local Simplex y el de cuasi-Newton. En el caso de estudio, el SCE demostró ser robusto y más eficiente. El algoritmo genético, aunque presentó tasas de convergencia inicial superiores al SCE, tendió con torpeza al óptimo y no demostró confianza para la ubicación del óptimo global.

Franchini et al. (1998) presentan los resultados de la comparación entre tres algoritmos: el Algoritmo genético acoplado con la Programación Cuadrática Secuencial (GA-SQP), la Búsqueda Patrón también acoplada con la Programación Cuadrática Secuencial (PS-SQP), y el SCE-UA. Realizaron los estudios empleando un modelo conceptual PQ aplicado a una cuenca simple y a otra más compleja. Para ambos tipos de cuencas, fue considerado un caso teórico sin errores en modelo y datos, en el cual los verdaderos valores de los parámetros son conocidos a priori, y varios casos reales donde existen errores de modelo y datos. Con relación a la cuenca simple, el SCE-UA fue el más confiable mientras que los otros dos algoritmos dieron resultados equivalentes al SCE-UA en los casos teóricos, pero en los casos reales

mostraron una tendencia creciente (particularmente el PS-SQP) de quedar atrapado en un mínimo local. Con referencia a la cuenca compleja, ninguno de los tres algoritmos identificó la solución exacta en el caso teórico. Sin embargo, el SCE-UA fue el único que sistemáticamente se aproximó al valor correcto mejor que los otros. En el caso real, sus soluciones fueron estables pero caracterizadas por muchos juegos de valores de parámetros en los límites de sus propios rangos. Los otros dos algoritmos produjeron un set de parámetros muy inestables. Un gran número de trabajos posteriores probaron la eficacia y capacidad de búsqueda del método SCE-UA (Thyer, et al., 1999; Gupta et al., 1999; Abdulla et al., 1999; Madsen, 2000; Eckhardt y Arnold, 2001); incluso Siqueira et al. (2016) lo emplearon recientemente para optimizar el modelo HEC-RAS en un tramo de la cuenca baja del río Iguazú (Brasil), a los fines de complementar la mejora del pronóstico hidrológico.

En esta línea de investigación, Yapo et al. (1998) introducen la necesidad de calibrar con múltiples objetivos. Indican que las experiencias prácticas realizadas con calibración de modelos sugieren que ninguna función mono-objetivo es adecuada para identificar las formas en que el modelo falla para reunir las características importantes de los datos observados. Dado que algunos de los últimos modelos hidrológicos simulan varias características del flujo de salida (como por ej., agua, energía, constituyentes químicos), sugieren que existe una necesidad de procedimientos de calibración multi-objetivo eficientes y efectivos a los fines de explotar toda la información útil respecto del sistema físico contenida en series de datos medidos. Entonces, presentan el algoritmo MOCOM-UA, que es una metodología eficiente y efectiva para resolver el problema de la optimización global multi-objetivo. El método es una extensión del exitoso algoritmo mono-objetivo de calibración global SCE-UA. A través de un simple modelo hidrológico estudian las características y capacidades de calibración del MOCOM-UA.

Gupta et al. (1998) concuerdan con la presunta desaparición de los problemas de calibración basados en el incremento y mejora de la disponibilidad de medidas de campo, y enfatizan en la creciente complejidad del problema de la calibración debido a la sofisticación de los modelos de escorrentía más recientes basados en los procesos físicos. Sugieren que la emergencia de un nuevo y más poderoso paradigma de calibración debe incluir el reconocimiento de la naturaleza multi-objetivo inherente al problema y debe reconocer explícitamente el rol del error del modelo. Presentan los resultados de algunos estudios preliminares y demuestran mediante la ilustración de un caso de estudio que el enfoque multi-objetivo no sólo es relativamente simple de implementar, sino que puede también proveer información útil respecto de las limitaciones del modelo.

Boyle et al. (2000) continúan el trabajo de Gupta et al. (1998), poniendo en relevancia la gran agilidad de la calibración automática, pero aceptan que estos métodos no proveen de estimaciones de parámetros (y su consecuente simulación de hidrogramas) aceptables para las operaciones de pronóstico, y atribuyen a esta deficiencia la razón por la cual no se ha generalizado su uso. En contraste, destacan que las técnicas manuales, si bien son capaces de lograr una calibración excelente de los modelos, son complicadas y requieren de mucha labor, y la experiencia adquirida por un individuo no es fácilmente transferible a otra persona (o modelo). Entonces proponen una aproximación híbrida que combina las ventajas de ambos. Emplean una formulación multicriterio para “modelar” las técnicas de evaluación empleando estrategias de calibración manual y algoritmos computarizados. La nueva aproximación provee un más fuerte test de performance del modelo que los métodos que usan un simple estadístico global para agregar los errores del modelo sobre un largo rango de comportamientos hidrológicos.

Gupta et al. (2003), señalan que el método MOCOM-UA presenta algunos problemas de convergencia e indican que sus principales debilidades son su tendencia a

concentrar las soluciones no dominadas y el arribo prematuro a la región de Pareto (antes de encontrar los valores óptimos de cada función objetivo). Vrugt et al. (2003) concluyen que estas fallas se deben a la probabilidad de ser elegido un individuo de la población es la misma para los individuos de una misma posición dentro del ranking de Pareto. Como aclara Yang (2015), los miembros de un mismo nivel dentro del ranking no son distinguidos entre ellos cuando se asigna la probabilidad de ser elegidos. Y la segunda razón que expresan Vrugt et al. (2003) es la alta dimensionalidad del espacio de búsqueda, que impide al método explorar adecuadamente. Por eso, propusieron una versión mono-objetivo de optimización, que constituye la versión modificada del algoritmo original de optimización global SCE-UA desarrollado por Duan et al. (1992). Se trata de un manipulador de muestras (sampler) del tipo Markov Monte Carlo (MCMC), denominado Shuffled Complex Evolution Metropolis (SCEM-UA), ya que incorpora conceptos del método de Metrópolis. Posteriormente, en otra publicación, Vrugt et al. (2003) lanzan el mismo año la versión multi-objetivo del SCEM-UA, denominado MOSCEM-UA (Multiobjective Shuffled Complex Evolution Metropolis). Los autores insisten en que la experiencia práctica con la calibración de modelos hidrológicos sugiere que ninguna función mono-objetivo, independientemente de cuan cuidadosamente haya sido elegida, es capaz de representar por sí misma, adecuadamente todas las características de los datos considerados importantes. Entonces proponen el MOSCEM, como un algoritmo capaz de resolver el problema de optimización multi-objetivo para modelos hidrológicos, empleando el concepto de dominancia de Pareto (más que el de evaluación directa de función mono-objetivo) para evolucionar la población inicial de puntos hacia un conjunto de soluciones estancadas con distribución estable (set de Pareto). Por último, los autores comparan la eficacia del MOSCEM-UA con la del algoritmo MOCOM-UA para tres casos de estudio de modelaciones hidrológicas que se incrementan en complejidad. Yilmaz (2007) confirma que algunos estudios reportan que el MOCOM-UA tiene una

tendencia a concentrar las soluciones en el centro de la región de Pareto y que puede requerir una gran cantidad de corridas para converger.

Sintetizando la última década en métodos de optimización, Efstratiadis et al. (2010) comentan que el progreso ha sido significativo en relación a diferentes componentes de los problemas de la calibración multi-objetivo, incluyendo: (a) la manipulación del algoritmo; (b) la formulación de los objetivos; (c) la interpretación de las soluciones no-dominadas y la orientación hacia una mejor solución de compromiso, y (d) el vínculo con el enfoque de la evaluación de incertidumbres. Todavía, subsisten varios asuntos abiertos reconocidos luego de la experiencia obtenida por el empleo del enfoque multi-objetivo en un amplio espectro de aplicaciones. En este sentido, aclaran que los problemas de calibración de modelos hidrológicos ciertamente presentan dificultades no existentes en otras áreas tecnológicas. En primer lugar, el elevado tiempo que demanda la corrida de un modelo hidrológico, no permite que la optimización multi-objetivo se realice con razonable esfuerzo. En segundo lugar, hay muy poca experiencia en superficies de respuesta multi-dimensionales de funciones objetivo. Siguiendo con Efstratiadis et al. (2010) se sostiene que un enfoque multi-objetivo no garantiza necesariamente la detección de calibraciones aceptables para fines hidrológicos.

Krauße et al. (2012) indican que muchos estudios prácticos sugieren que una sola función objetivo, sin importar cuán cuidadosamente fue elegida, es a menudo insuficiente para representar todas las características del comportamiento del sistema. Agrega que el error medio cuadrático puede ser un buen estimador para un estudio de balance, sin embargo, pero que es probablemente inadecuado para medir la performance del modelo para pronóstico de crecidas donde es crucial una correcta estimación del caudal pico y su tiempo de ocurrencia. En consecuencia, indican que un enfoque mono-objetivo de calibración no es considerado aceptable en muchos casos por hidrólogos con experiencia. Krauße et al. (2012), además hacen énfasis en

que el objetivo actual de una buena calibración de modelo no debería ser encontrar los vectores de parámetros que sólo ajustan bien el periodo calibrado, sino que dichos parámetros sean robustos, lo cual implica tener buena performance en fase de validación (también denominadas verificación o control). De este modo, ese juego de parámetros podrá ser transferido a otros periodos. Es por esto que resulta una condición deseable que los parámetros sean tan insensibles como sea posible, de modo que pequeñas modificaciones en los valores paramétricos no impliquen grandes variaciones de performance del modelo. Ahora bien, el modelo puede ser sensible a modificaciones de los parámetros, pero lo importante es que en las proximidades del óptimo no lo sea. También agrega que una buena estrategia de desarrollo de un modelo robusto, debería implicar una buena implementación del principio de parsimonia en el mismo, es decir que represente con la estructura deseada, pero empleando el menor número de parámetros posible.

Reed et al. (2013) realizan una evaluación rigurosa del estado el arte en algoritmos evolutivos multi-objetivo (MOEAs, por sus siglas en inglés: MultiObjective Evolutionary Algorithms) y remarcan claves que los nuevos desarrollos en el campo de los recursos hídricos deben considerar. El diagnóstico mide la efectividad, eficiencia, confiabilidad y "controlabilidad" de diez MOEAs de referencia ("benchmark") en aplicaciones de recursos hídricos, que incluye calibración de modelos P-Q, entre otros. Los autores remarcan los aspectos claves que deberán afrontar en el futuro los MOEAs orientados a las herramientas necesarias para los recursos hídricos. Indican que es importante profundizar en los algoritmos auto-adaptativos que adaptan su búsqueda de acuerdo a los rasgos de un problema dado con una mínima entrada del usuario, esto fue inicialmente implementado por Vrugt et al. (2009) en su algoritmo AMALGAM.

7. Criterios de ajuste y evaluación de la performance del modelo

Este capítulo introduce sobre aspectos vinculados a la calidad del ajuste del modelo, presentando las expresiones clásicas para cuantificar la bondad de la calibración, y dirime respecto de la representatividad del valor numérico obtenido por cada una.

Sobre las cosas que no se conocen siempre se tiene mejor opinión

Gottfried W. Leibniz

Una sintética definición de la función objetivo (FO) es apuntada por Sorooshian y Gupta (1995) estableciéndola como “una ecuación usada para computar una medida numérica de la diferencia entre la salida simulada por el modelo (típicamente el hidrograma de escurrimiento) y el hidrograma observado (medido)”. Es decir que la FO representa el criterio de ajuste. Agregan estos autores que el propósito de la calibración del modelo es, por consiguiente: “encontrar esos valores de los parámetros del modelo que optimizan (minimizan o maximizan, como sea apropiado) el valor numérico de la función objetivo”. Asimismo, expresan que la FO es uno de los pilares sobre los cuales se apoya la calibración automática de modelos matemáticos, siendo los otros tres el algoritmo de optimización, el criterio de parada y los datos de calibración.

Una vez realizada la calibración, se establece una pauta de referencia que nos indica la calidad del ajuste realizado a través de un valor numérico. Es decir que se debe establecer un criterio de performance.

Tradicionalmente, la FO es adoptada como el error dado por la diferencia numérica entre los valores simulados por el modelo matemático (en nuestro caso, caudales

simulados), y los valores observados en la realidad (caudales observados), mediante algún tipo de estimador, típicamente el error medio cuadrático según apuntan Sorooshian y Gupta (1995). Por esta razón, la metodología clásica de calibración automática tenía por objetivo minimizar el valor numérico de la FO. Posteriormente se adoptaba un criterio de performance que nos daba la pauta objetiva del grado de calidad del ajuste realizado, como es por ejemplo el estimador de Nash y Sutcliffe (1970).

Actualmente, hay una tendencia a emplear el mismo criterio de performance como FO, es decir como criterio de ajuste. En este caso, evidentemente el método de calibración automática debe maximizar el valor numérico de la FO. Es decir, el mismo coeficiente que antes se usaba para evaluar el ajuste, ahora directamente se emplea para ajustar, y la obtención de un buen valor del mismo, ya implica buena performance. En la Figura 7-1 se muestra la forma típica que tiene el frente de Pareto en caso de que la FO esté planteada como un error, la cual es cóncava ya que el método minimiza su valor. A su lado se observa la forma convexa que presenta el frente de Pareto cuando la FO es una medida de la performance del modelo, ya que se procura maximizar su valor.

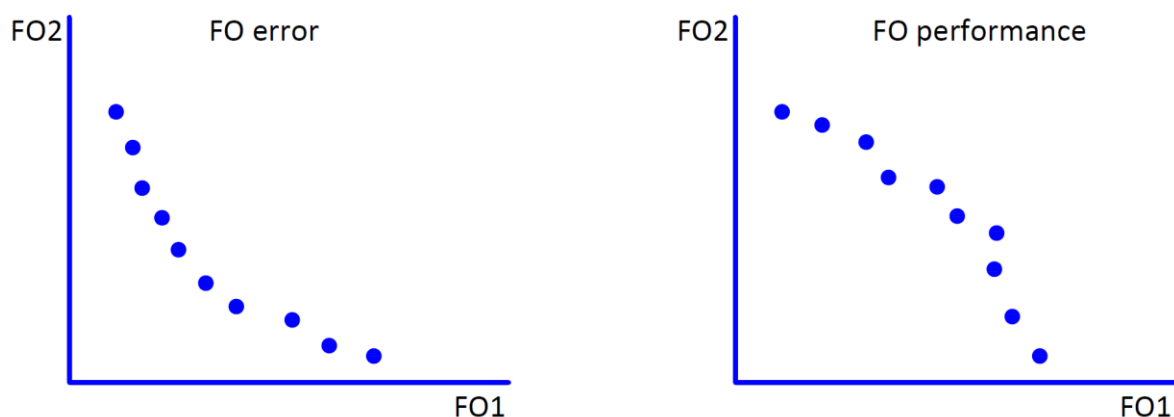


Figura 7-1: Ejemplo de frente de Pareto según el tipo de FO empleado.

Debido a la delicada función que cumple la FO en el proceso de calibración, la misma conviene que sea elegida cuidadosamente para obtener el resultado buscado.

Existen algunas pautas generales para propender a la elección de la FO más conveniente para cada problema, como por ejemplo si el propósito es mejorar la estimación de los picos de los hidrogramas, la FO por excelencia es el error cuadrático entre los hidrogramas observados y simulados, pues maximiza justamente los errores en los picos. A la inversa, si se desea mejorar la estimación de los valles, podría ser la función inversa la que mejore dicha estimación, como se verá seguidamente.

A pesar de haber pautas generales, varios autores (Sorooshian et al, 1983; Gan y Burges, 1990; Brath y Rosso, 1993; Vélez y Francés, 2004) señalan que no está claro cuál es la función objetivo de mejor desempeño. La FO elegida debe tener una estructura matemática sencilla, para poder identificar con claridad el sentido físico del número que nos arroja.

7.1. Restricciones

El proceso de calibración del modelo matemático puede estar sometido a restricciones o no. En el primer caso, el valor de los parámetros debe moverse en un intervalo previamente establecido por las restricciones, por lo que la FO puede tomar valores en una región definida por los rangos de validez de los parámetros del modelo. En el segundo caso, los parámetros del modelo pueden tomar cualquier valor en el proceso de calibración. Tanto la FO, como sus restricciones son funciones matemáticas que podrán tener forma lineal o no lineal (cuadrática, logarítmica, etc.). Las restricciones vienen expresadas, normalmente, como desigualdades.

7.2. Criterios de ajuste basados en el error de simulación

Como se expresó previamente existen diversos tipos de FO, de acuerdo a su finalidad. Sin embargo la FO mayormente empleada en la teoría de optimización (Sorooshian y Gupta, 1995), y por lo tanto en calibración automática, ha sido

diferentes formas de la Función de Mínimos Cuadrados Ponderada, presentada en la Ec. 7-1.

$$FO = \sum_{i=1}^n w_i (QO_i - QC_i)^2 \quad \text{Ec. 7-1}$$

Que puede simplificarse a la denominada Función Cuadrática, como la denomina Tucci (2005), simplemente igualando los pesos de todos los intervalos, como se muestra en la Ec. 7-2.

$$FO = \sum_{i=1}^n (QO_i - QC_i)^2 \quad \text{Ec. 7-2}$$

Donde:

QO_i: valor de caudal observado en el intervalo i.

QC_i: valor de caudal calculado en el intervalo i.

w_i: peso relativo de la diferencia cuadrática en el intervalo i.

n: número de intervalos de los hidrogramas comparados.

Según Tucci (2005), esta función asigna mayor peso a los caudales mayores. En el ajuste de un periodo continuo de caudales la tendencia del ajuste será el de ajustar mejor las crecidas. Esto se debe a que, para un porcentaje igual de error, la magnitud de dicho error es mayor cuando los caudales son mayores. Por ejemplo, supongamos un caudal en recesión de 20 m³/s y en crecida de 200 m³/s para un determinado río. Un error del 5% en recesión significa sólo 1 m³/s, pero en crecida corresponde a 10 m³/s. Es decir, que elevando al cuadrado el error en recesión corresponde a 1 m³/s, y en crecida a 100 m³/s. Con este ejemplo se observa que, si el método de calibración automática está diseñado para minimizar los errores, y la FO que mide dicho error es el error medio cuadrático, cuando alcance una precisión de 1 m³/s habrá ajustado al 5% los caudales en recesión y al 0,5% los caudales en crecida.

La FO de mínimos cuadrados tiene gran aplicación por su sencilla formulación y fácil interpretación. Sin embargo, recuerdan Sorooshian y Gupta (1995) que la elección de alguna forma de la Función de Mínimos Cuadrados está típicamente basada en un criterio subjetivo, de acuerdo a dos supuestos referidos a la distribución probabilística de los errores, que se desprenden de la teoría de Máxima Verosimilitud:

- La probabilidad conjunta de los errores sobre la serie de datos disponibles, es Gaussiana con media cero.
- Los errores son independientes unos de otros.

Adicionalmente, Clarke (1973) enfatizó que, dado que los datos hidrológicos pueden frecuentemente violar esas hipótesis, los efectos de asumirlas pueden resultar bastante serios. En consecuencia, Sorooshian y Dracup (1980) desarrollaron funciones objetivo basadas en la Máxima Verosimilitud, para contabilizar apropiadamente tanto la autocorrelación (carencia de independencia), como la heterocedasticidad (varianza cambiante) de las series de datos de caudales. En muchos casos, el criterio de Máxima Verosimilitud (MV) dio pruebas de superioridad con respecto a la Función de Mínimos Cuadrados (FMC). Estudios posteriores de Sorooshian y Gupta (1983) revelaron que entre las razones que justifican el éxito del criterio de MV, fueron las propiedades de convergencia superiores de las superficies de las funciones respuestas generadas por esta.

La forma más exitosa del criterio de Máxima Verosimilitud ha sido una función llamada HMLE (Heteroscedastic Maximum Likelihood Estimator) propuesta por Sorooshian y Gupta (1995) que contabiliza la no estacionariedad de la varianza en los errores de caudal medidos. El estimador HMLE es el estimador de máxima verosimilitud, mínima varianza, asintóticamente no sesgado, cuando los errores en los datos de salida son Gaussianos, con media cero, no correlacionada, y tiene una

varianza relativa a la magnitud de los caudales generados. La no estacionariedad de la varianza de los errores se cree que son comunes en los datos de caudal. Además, el estimador HMLE está relacionado con la FMC, ya que aquella se reduce a esta en el caso especial en que la varianza se hace constante en los errores de los datos de caudal.

El estimador HMLE, tiene la forma de la Ec. 7-3.

$$\min HMLE = \frac{\frac{1}{n} \sum_{i=1}^n w_i \varepsilon_i^2}{\left[\prod_{i=1}^n w_i \right]^{1/n}} \quad \text{Ec. 7-3}$$

Donde:

$\varepsilon_i = \text{QO}_i - \text{QC}_i$, es el residual del modelo en el tiempo i , siendo QO y QC los caudales observados y calculados por el modelo, respectivamente.

w_i es el peso asignado al tiempo i , computado como sigue:

$$w_i = f_i^{2(\lambda-1)}$$

Donde:

f_i es el valor verdadero de caudal en el intervalo de tiempo i . Se aproxima mediante QO_i .

n es el número de datos de la serie.

λ es el parámetro desconocido de transformación que estabiliza la varianza.

Tucci (2005), resume las principales funciones empleadas en la hidrología para calibración automática de modelos hidrológicos, además de la Función de Mínimos Cuadrados: la Función Módulo, la Función Inversa y la Función Relativa, presentadas respectivamente en las ecuaciones Ec. 7-4, Ec. 7-5 y Ec. 7-6. Indica Tucci (2005) que

la Función Módulo también asigna mayor peso a los caudales mayores, mientras que la Función Inversa busca ajustar mejor los caudales menores. Mientras que la Función Relativa, empareja las prioridades dividiendo por el valor del caudal observado en cada intervalo.

$$FO = \sum_{i=1}^n |QO_i - QC_i| \quad \text{Ec. 7-4}$$

$$FO = \sum_{i=1}^n \left(\frac{1}{QO_i} - \frac{1}{QC_i} \right)^2 \quad \text{Ec. 7-5}$$

$$FO = \sum_{i=1}^n \left(\frac{QO_i - QC_i}{QO_i} \right)^2 \quad \text{Ec. 7-6}$$

Tassi et al. (2006), presentan las tres FO anteriores, más la Función Cuadrática (Ec. 7-7), pero dividiéndolas por el número de datos n , y obteniéndole la raíz cuadrada. Así, le asigna un sentido conceptual al valor numérico hallado, convirtiendo la Función Cuadrática en la expresión de la Raíz Media Cuadrática, o sea, una forma de error promedio. Además, emplean también una nueva FO, que se muestra en la Ec. 7-7. Esta forma de FO también es elegida por Yapó et al. (1998), quienes adoptan además el criterio de HMLE para presentar y probar el desempeño del algoritmo multi-objetivo MOCOM-UA.

$$FO = \sum_{i=1}^n \left(\frac{QO_i - QC_i}{QO_i} \right) \quad \text{Ec. 7-7}$$

Como puede verse, la FO adoptada para calibrar el modelo hidrológico es altamente dependiente del uso al cual será destinado el mismo. Si el objetivo es ajustar el modelo para simulación de crecidas, debe procurarse minimizar los errores en la simulación de los caudales mayores, para lo cual sería adecuado aplicar la función cuadrática o módulo. Si, en cambio, el modelo será usado para simulación de una secuencia de caudales, para la operación o para dimensionamiento de reservorios,

el ajuste busca que las frecuencias de los caudales observados y calculados sean semejantes, en cuyo caso sería recomendable la función relativa (Tucci, 2005).

Madsen (2000) resume los objetivos mayormente considerados en una apropiada calibración automática, en los siguientes cuatro:

1. Un buen acuerdo entre los volúmenes totales escurridos y simulados por el modelo.
2. Una concordancia global aceptable entre las formas de ambos hidrogramas.
3. Una buena concordancia entre los picos de hidrogramas, en términos de tiempo de pico, pendiente de la curva y volumen.
4. Una buena aproximación de los caudales bajos.

Agrega Madsen (2000) que resulta necesario traducir el objetivo global de la calibración en términos más operacionales, de donde se desprenden las formas más comunes de las funciones objetivos. Propone en su trabajo algunas otras formas de FO que pueden resultar interesante, por lo cual el lector interesado podrá remitirse a la publicación referida.

7.3. Criterios de performance

Con el propósito de verificar la bondad de la calibración realizada, han sido desarrolladas algunas expresiones matemáticas que cuantifican la performance del modelo.

Eficiencia de Nash y Sutcliffe (NSE)

En la bibliografía de modelos PQ ha ganado notable protagonismo el clásico estimador de Nash y Sutcliffe (1970), usualmente simbolizado como NSE presentado

en la Ec. 7-8, que representa una reinterpretación del coeficiente de determinación, empleado en regresión lineal.

$$NSE = 1 - \frac{F^2}{F_o^2} = \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \quad \text{Ec. 7-8}$$

Donde $\overline{Q_{obs}}$ es el promedio de los caudales del hidrograma observado.

Notar que el criterio de Nash-Sutcliffe (NSE) es una expresión matemática que cuantifica la performance del modelo sobre la base de restar a la unidad una magnitud de error, que se normaliza relacionando el error de simulación respecto de la media del hidrograma observado. Lo que importa destacar acá es que el criterio de NSE resta a la unidad un cociente que representa un error al cuadrado que está normalizado respecto del promedio de las observaciones de la serie de caudales observados. Es decir, que la magnitud de referencia es la media de caudales observados de la serie histórica de datos. En efecto, al normalizar se puede comparar simulaciones entre sí, aunque pertenezcan a distintas cuencas, períodos, etc. Si observamos el denominador, notaremos que mide la distancia al cuadrado entre los caudales y la media, lo cual nos da un orden de magnitud de las dimensiones de los caudales medidos. Entonces promediando esas diferencias elevadas al cuadrado, da como resultado la varianza de esa distribución, o sea ofrece un orden de magnitud de las variaciones. Ahora bien, cuando se calcula el cuadrado de las diferencias reales de los Q simulados con respecto a los observados, el resultado es la varianza del error real, que se normaliza al dividirlo por la varianza de las magnitudes verdaderas. Eso es muy valioso porque, por ejemplo, cometer un error de simulación de 10 m³/s en un río que tiene caudales de 5 a 20 m³/s es muy grosero, pero es insignificante para un río que tiene fluctuaciones entre 2000 y 5000 m³/s. Entonces si se calcula la magnitud de esas fluctuaciones en el denominador con un valor como el propuesto por Nash-Sutcliffe, dado por la varianza de los caudales observados respecto de la media observada, se tiene un orden de magnitud de los valores que alcanzan las

fluctuaciones reales en ese río. Si luego se divide la varianza de los caudales simulados respecto de los observados, se tiene un número que representa la magnitud del error. La relación entre ambos, da un coeficiente adimensional que indica la relación entre el error de simulación y la magnitud promedio de las fluctuaciones de la variable, que en este caso es el caudal. Eso da una idea del error al cuadrado normalizado, es como la varianza normalizada. De modo que es posible comparar ríos con módulos del orden de los 10 m³/s con otros de 5000 m³/s. Notar que ese valor no puede ser inferior a cero porque el numerador y el denominador no pueden ser negativos, por lo tanto, siempre el resultado de ese cociente es positivo. En otras palabras, el valor de NSE no puede ser mayor que 1. Pero sí puede ser mucho mayor el denominador que el numerador, en cuyo caso es un error pequeño, o bien el numerador que el denominador, en cuyo caso es un error grande y por lo tanto se obtendría un cociente mucho mayor que uno, por lo que NSE sería negativo. En teoría si el error fuera inmenso NSE puede tender incluso a menos infinito.

A pesar de su futura popularidad, un año más tarde de su clásica publicación el mismo equipo de trabajo (Garrick et al., 1978) publica que el criterio NSE es notoriamente optimista ya que produce valores elevados que no refleja apropiadamente la calidad del ajuste.

A este respecto Perrin (2000) reflexiona con ejemplos muy sencillos. Llamando ε a la relación F/F_0 , que representa un error relativo, se puede entonces sintetizar que el NSE es $(1-\varepsilon^2)$. Visto de ese modo, cuando ε toma el valor de 0,2 representa un error medio cuadrático del 20% en relación al error de referencia de los caudales observados. Pero en ese caso, el criterio de NSE daría 0,96, lo cual parece indicar una elevada performance del modelo. Visualmente podría suponerse que se cometió un error del 4%, cuando en realidad fue del 20%. Si el error relativo fuera del 40%, el estimador NSE daría 0,84, nuevamente sobreestimando la performance del modelo.

Y agrega Perrin (2000) que la asociación abusiva que se hace del criterio de NSE a un coeficiente de determinación tiende a hacer creer que el modelo es bueno, aunque no sea el caso.

Gupta et al. (2009) renuevan las críticas al estimador NSE indicando que hay una cierta incapacidad para comparar cuencas entre sí, que fueron superadas por el estimador KGE, como se explica más adelante.

En este sentido puede observarse que dos cuencas con el mismo caudal promedio histórico de sus series de datos, pero con distintas magnitudes de sus caudales picos, o con distintas cantidades de eventos extremos a lo largo de sus series de datos, pueden alterar la posibilidad de compararlos entre sí.

C2M

Mathevet et al. (2006), atentos a las críticas de Garrick et al. (1978) respecto del popular criterio de NSE, proponen un criterio denominado C2M, que presenta de un modo más realista la performance del modelo, al mismo tiempo que presenta un rango de valores posibles del modelo entre -1 y 1, en vez de $-\infty$ a 1, como es el rango de NSE. Esto último, aclaran que debería permitir una mejor interpretación del significado estadístico de la performance del modelo. La relación entre C2M y NSE se expresan en la Ec. 7-9, y en la Figura 7-2 se presenta gráficamente dicha relación.

$$C_{2M} = \frac{NS}{2 - NS}$$

Ec. 7-9

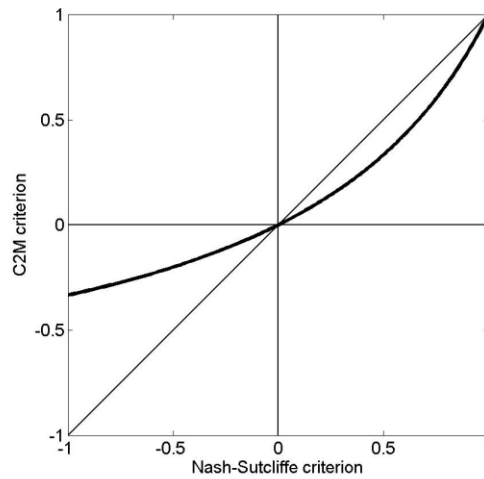


Figura 7-2: Relación entre los criterios de NSE y C2M (Fuente: Mathevet et al., 2006)

Eficiencia de Kling y Gupta (KGE)

Un estimador que ha cobrado gran relevancia y popularidad en los últimos años es el denominado Eficiencia de Kling y Gupta (KGE), desarrollado y publicado por Gupta et al. (2009), y está dado por la relación indicada en la Ec. 7-10.

$$KGE = 1 - ED \tag{Ec. 7-10}$$

Donde:

$$ED = \sqrt{(r - 1)^2 + (\alpha - 1)^2(\beta - 1)^2} \tag{Ec. 7-11}$$

Siendo cada término:

$$\alpha = \frac{D(Q_{sim})}{D(Q_{obs})}$$

$$\beta = \frac{\mu_s}{\mu_o}$$

$$r = \frac{E(Q_{obs} \cdot Q_{sim}) - E(Q_{sim}) \cdot E(Q_{obs})}{D(Q_{sim}) \cdot D(Q_{obs})}$$

A su vez, estas expresiones representan:

$$\mu_s = E(Q_{sim})$$

$$\begin{aligned}\mu_o &= E(Q_{obs}) \\ D(Q_{sim}) &= \sqrt{V(Q_{sim})} \\ D(Q_{obs}) &= \sqrt{V(Q_{obs})} \\ V(Q_{sim}) &= E(Q_{sim}^2) - [E(Q_{sim})]^2 \\ V(Q_{obs}) &= E(Q_{obs}^2) - [E(Q_{obs})]^2 \\ E(Q_{sim}) &= \frac{\sum Q_{sim}}{n} \\ E(Q_{obs}) &= \frac{\sum Q_{obs}}{n} \\ E(Q_{sim}^2) &= \frac{\sum(Q_{sim}^2)}{n} \\ E(Q_{obs}^2) &= \frac{\sum(Q_{obs}^2)}{n} \\ E(Q_{sim} \cdot Q_{obs}) &= \frac{\sum(Q_{sim} \cdot Q_{obs})}{n}\end{aligned}$$

ED: Distancia Euclidiana del punto ideal.

β : relación entre el cual medio simulado y el caudal medio observado.

E(Qsim): esperanza de Qsim. (Idem Qobs).

V(Qsim): varianza de Qsim. (Idem Qobs).

D(Qsim): desvío de Qsim. (Idem Qobs).

Gupta et al. (2009) descompusieron el coeficiente NSE, y remarcaron problemas asociados con el uso de este criterio para ser usado como indicador de performance e incluso como función objetivo para calibrar modelos hidrológicos. Luego demostraron que el criterio de KGE puede sobreponerse a varios de esos inconvenientes, y ajusta mejor los modelos, sobre todo cuando es empleado el mismo coeficiente como criterio de calibración.

Coeficiente de Persistencia (PC)

El estimador basado en la persistencia (PME), también denominado coeficiente de Persistencia (PC) o bien Índice de Persistencia (PI), es un coeficiente que ganó protagonismo en el campo del pronóstico hidrológico, dado que relaciona la magnitud del error de pronóstico con una magnitud de referencia dada por la variación de los caudales en el horizonte de pronóstico.

El mismo es propuesto por Kitadinis y Bras (1980) para comparar la performance del modelo empleando el valor observado del paso de tiempo previo como la predicción del paso de tiempo actual. Estos autores indican que el coeficiente de persistencia compara las predicciones del modelo con las predicciones obtenidas por asumiendo que se trata de un proceso de tipo Wiener, es decir que la varianza se incrementa linealmente con el tiempo, en cuyo caso, el mejor estimador para el futuro está dado por la última medición.

Moriasi et al. (2007) expresan que la estimación PC es una evaluación estadística normalizada del modelo que cuantifica la magnitud relativa de la variancia residual (ruido) respecto de la variancia del error obtenida por el uso de un simple modelo de persistencia.

El rango de variación del estimador es entre 0 y 1, siendo $PC=1$ el valor óptimo. Conforme lo sentencian Gupta et al. (1999), el PC debería ser al menos mayor que 0 para indicar una performance mínimamente aceptable para el modelo. De acuerdo con Gupta et al. (1999), PC es capaz de indicar claramente la baja performance de un modelo, pero dado su escasa referencia en la bibliografía, no se cuenta con un rango disponible de su representatividad en relación a la performance.

La expresión del Coeficiente de Persistencia tiene la forma que se presenta en la Ec. 7-12, y se representa gráficamente en la Figura 7-3.

$$PC = 1 - \frac{\sum \varepsilon_m^2}{\sum \varepsilon_r^2}$$

Ec. 7-12

Donde:

ε_m : representa el error en la estimación.

ε_r : es una diferencia de magnitud entre $Q_{obs}(i)$ y el valor de $Q_{obs}(i+L)$, donde L es el horizonte de pronóstico.

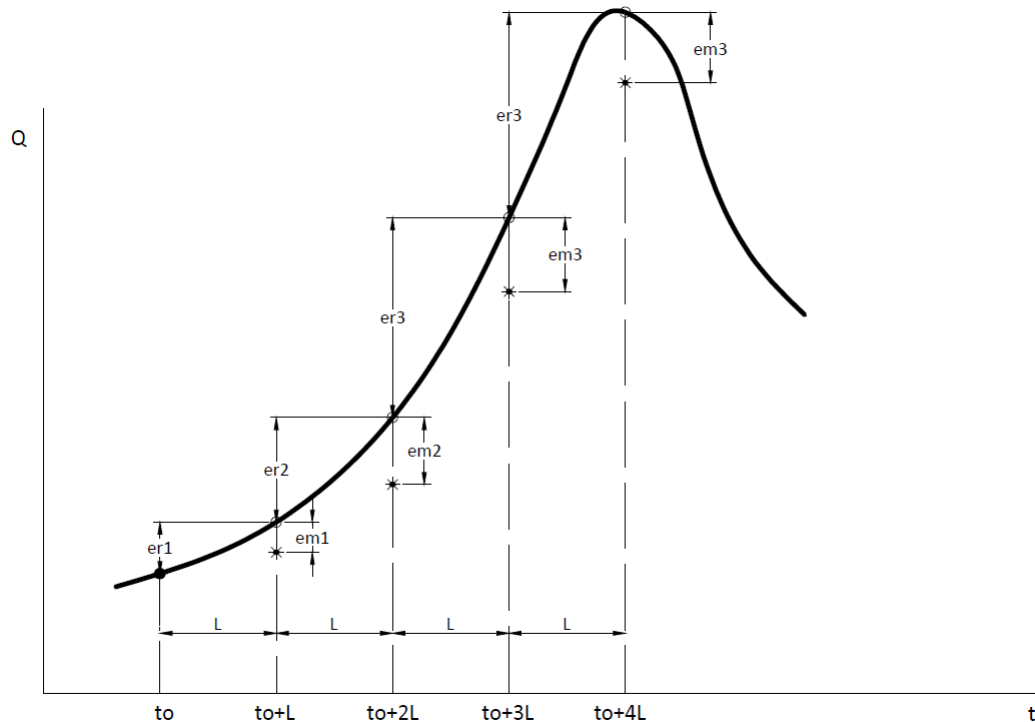


Figura 7-3: Ilustrativo de las variables que intervienen en la expresión del Coeficiente de Persistencia.

A pesar de las observaciones y críticas apuntadas previamente sobre la expresión de Nash, ha demostrado ser un indicador aceptable en la bibliografía. Sin embargo, su bondad está más vinculada a representar de manera normalizada la calidad del ajuste de un modelo de simulación, y no necesariamente para un modelo aplicado al pronóstico. Esto es porque en pronóstico la medida comparativa que establece la escala de referencia no es el caudal promedio (o la diferencia promedio de caudales) de todo el hidrograma, sino que interesa más específicamente el tramo de hidrograma en el paso en que se está abordando el pronóstico. En este aspecto

emerge como más apropiado el concepto de “persistencia” como criterio de evaluación de la calidad de pronóstico.

Si observamos la Figura 7-3, puede verse que existe una dimensión εm , que mide el error, y otra dimensión εr , que indica la medida de referencia con la cual se adimensionaliza el error de pronóstico. Esta segunda dimensión, si bien es denominada también con la letra ε , en rigor no es un error, sino una diferencia simplemente, que indica la medida con respecto a la cual se normalizan los errores de pronóstico. Como εr representa la diferencia de los valores de caudal entre el tiempo actual y el tiempo futuro pronosticado, ofrece un orden de magnitud de los saltos que da el hidrograma en ese intervalo. Por ejemplo, si se comete un error de pronóstico de $1\text{ m}^3/\text{s}$ en un tiempo de pronóstico dado en que la variación del caudal del hidrograma fue de $100\text{ m}^3/\text{s}$, el error es insignificante, mientras que no lo es si la variación de caudal fuera de $10\text{ m}^3/\text{s}$.

Notar que también podría haberse usado el criterio de Nash, pues tiene en cuenta el promedio de los caudales, es decir algo así como el orden de magnitud del módulo del río. Pero el criterio de Nash no tiene en cuenta cómo varían los caudales durante el momento en que se está pronosticando. Aclaremos con un ejemplo. Si tenemos un hidrograma que tiene un módulo de unos $2000\text{ m}^3/\text{s}$, durante los tramos de crecida puede tener fluctuaciones de $1000\text{ m}^3/\text{s}$ en un intervalo de pronóstico, mientras que en los de recesión, puede rondar los $10\text{ m}^3/\text{s}$ por ejemplo. Pero el error cometido en pronóstico durante el periodo de crecida, no puede ser medido con la misma vara con que se mide un error en recesión, que es el problema que tiene el criterio de Nash para ser aplicado en pronóstico. En efecto, si se cometiera un error de $1\text{ m}^3/\text{s}$, de acuerdo al criterio de Nash, es insignificante como error al lado de lo que significa el módulo del río, pero si consideramos que las variaciones de los caudales durante el periodo de recesión rondan los $10\text{ m}^3/\text{s}$, de acuerdo al criterio

de persistencia, se vuelve importante. En cambio, un error de $1\text{m}^3/\text{s}$ en crecida, el mismo criterio de persistencia también lo considera insignificante, pues en ese caso lo contrastaría con el salto de $1000\text{ m}^3/\text{s}$ que se produce durante el periodo de crecida.

En síntesis, el criterio de Nash es menos exigente pues no atiende la medida de los caudales en el momento en que se está pronosticando, sino que mira la medida del error sin importar en qué momento se produce. El criterio de persistencia, en cambio, compara ese error con el orden de magnitud de las variaciones de los caudales en el momento en el que se está pronosticando. Está diseñado este criterio para fijar el grado de precisión que se requiere en cada parte del hidrograma, y mide el error cometido contrastándolo con ese grado de precisión.

Error del punto ideal (IPE)

El IPE (por sus siglas en inglés: Ideal Point Error) es presentada por los autores (Dawson et al., 2012) como una medida simple e integrada de performance especialmente orientada a modelos de pronóstico hidrológico, que permite una inter-comparación de modelos más directa.

$$IPE = \left[0.25 \left(\left(\frac{RMSE_i}{\max(RMSE)} \right)^2 + \left(\frac{MARE_i}{\max(MARE)} \right)^2 + \left(\frac{ME_i}{\max|ME|} \right)^2 + \left(\frac{R_i - 1}{1/\max(R)} \right)^2 \right) \right]^{1/2} \quad \text{Ec. 7-13}$$

Donde:

RMSE: Raíz media cuadrática del error.

MARE: Error relativo medio absoluto.

ME: Error medio.

R: Coeficiente de Correlación.

PARTE 3: METODOLOGÍA Y DATOS

Esta tercera parte del informe se focaliza concretamente en los métodos y técnicas empleados en esta tesis, incluyendo una descripción detallada y pormenorizada de las herramientas de trabajo. La totalidad de las herramientas presentadas en esta tesis fueron programadas en FORTRAN 90/95 con el propósito de permitir la manipulación de las estructuras de los modelos y algoritmos empleados. Asimismo, se realiza también una descripción de la cuenca de aplicación y los datos disponibles. Esta parte se cierra con una explicación de la metodología propuesta para orientar los trabajos a la concreción de los objetivos.

8. El modelo de pronóstico hidrológico GRP

En este capítulo se realiza una amplia descripción del modelo de pronóstico hidrológico Génie Rural Prévision (GRP). Tal como se viene expresando en los capítulos anteriores, todo modelo tiene asociada una incertidumbre en los valores estimados, relacionada con las fuentes de error que intervienen en la determinación de los valores pronosticados. Por esta razón, la descripción de la estructura del modelo de manera detallada permite mejorar la comprensión del alcance y tomar sensibilidad de la confiabilidad que presentan los resultados obtenidos.

Es preciso recordar que siempre hay un futuro, futuro que es inexorablemente incierto. Con esto no quiero dar un mensaje de pesimismo, ni mucho menos, ya que la incertidumbre puede terminar tanto negativa como positivamente, con la ventaja de que el ser humano tiene facultades para inclinar la balanza a su favor.

Roberto Gómez Bolaños (el chavo del 8)

8.1. Evolución de los modelos GR

La simulación del caudal en la salida de una cuenca hidrográfica es necesaria, y en ocasiones, indispensable para las numerosas aplicaciones de ingeniería y de gestión de los recursos hídricos, tales como el dimensionamiento y la gestión de obras, el pronóstico de crecidas o de estiaje, la detección del impacto en el uso del suelo o en los cambios climáticos. Para responder a esta cuestión, el instituto IRSTEA (ex CEMAGREF) comenzó a desarrollar a comienzos de los años 1980 (Michel, 1983) una generación de modelos hidrológicos denominados genéricamente “Génie Rural” (GR). Como todo modelo hidrológico, los modelos GR operan una triple globalización (Perrin et al., 2007):

A nivel de proceso: se propone una relación comportamental simple a la escala de cuenca hidrográfica, puesta a punto empírica, y sin vínculo conceptual con los procesos físicos subyacentes. Es decir, la puesta a punto se focaliza en el juego de parámetros que mejor ajusta al modelo para las condiciones planteadas.

A nivel espacial: Los modelos GR, son de tipo global, es decir que trabajan con información promediada de toda la cuenca hidrográfica. A pesar de ello, algunas de sus versiones han sido calibradas con éxito en cuencas hidrográficas desde unos pocos Km², hasta varios miles de Km² (Perrin, 2000), como, por ejemplo, la cuenca hidrográfica del río Sena a la entrada a París, de unos 43.800 km².

A nivel temporal: los modelos GR han sido desarrollados para distintos pasos de tiempo específicos: (i) Versión anual denominado GR1A (Mouelhi et al., 2006^a), (ii) Versión mensual denominado GR2M (Mouelhi et al., 2006^a), (iii) Versión con paso de tiempo diario, denominada GR4J (Perrin et al., 2003), (iv) Versión horaria, GR3H (Berthet, 2010). Adicionalmente, existe una versión reformulada para la aplicación en pronóstico hidrológico, denominada GRP, orientada al paso de tiempo diario (Tangara, 2005), y horario (Berthet, 2010). Para conocer en detalle la estructura de cada modelo, remitirse a la referencia citada.

La especificidad del paso de tiempo, a pesar de que muchas veces es despreciado, es importante, ya que una misma estructura de modelo no presenta la misma performance para todos los pasos de tiempo. Globalmente, las estructuras son más simples cuanto más largo es el paso de tiempo considerado.

Aunque estos modelos aparentan ser conceptuales, debido a que fueron concebidos como estructuras vinculadas de reservorios sobre la base de los conceptos físicos de los procesos subyacentes, en realidad son modelos empíricos, ya que su construcción fue desarrollada sobre la base de grandes cantidades de datos. Esta técnica ha permitido ir descubriendo progresivamente la estructura del modelo que

mejor reproduce el comportamiento hidrológico de la cuenca hidrográfica (es decir, su respuesta a las precipitaciones).

Los siguientes conceptos han sido progresivamente impuestos en el transcurso del desarrollo de estos modelos para permitir obtener modelos confiables y robustos (Mathevet, 2005):

- Representación global de la cuenca hidrográfica.
- Enfoque empírico del desarrollo, sin recurrir a priori a la física del escurrimiento.
- Aumento progresivo de la complejidad de la estructura del modelo partiendo de estructuras simples.
- Justificación de la complejidad de la estructura de un modelo por sus performances.
- Estudio de estructuras generales del modelo (aplicables a diversas cuencas hidrográficas).
- Utilización de grandes cantidades de cuencas hidrográficas para probar las performances del modelo (Andréassian et al., 2006).
- Evaluación de un modelo mediante la comparación de otras estructuras de modelos.

Estos modelos disponen de ciertas características que le atribuyen versatilidad para ser aplicados a los múltiples escenarios que pueden presentarse en un estudio regional asociado a la Gestión Integrada de los Recursos Hídricos.

A lo largo de su desarrollo, los modelos GR han sido todos probados sobre una larga muestra de varias centenas de cuencas hidrográficas, particularmente de Francia,

pero también de Estados Unidos, Australia, Brasil, México, Argelia, y Costa de Marfil (Perrin et al., 2007). Diferentes equipos de investigación, consultoras, y servicios operacionales han utilizado los modelos en condiciones muy variadas. Sólo Berthet (2010) aplicó el modelo sobre 1024 cuencas hidrográficas francesas con paso de tiempo horario.

Estos modelos no requieren, para su funcionamiento, más datos que las series continuas de precipitación y evapotranspiración potencial, pudiendo ser esta última estimada con base en una curva promedio interanual (Oudin et al., 2005a y b).

Los modelos GR tienen un pequeño número de parámetros, lo cual permite limitar los problemas de sobreparametrización y les confiere una buena robustez (Perrin et al., 2000). Los parámetros son determinados por calibración, minimizando el error del modelo mediante una serie de caudales observados. Si estos últimos no están disponibles para la calibración, la estimación de los parámetros puede realizarse mediante las fórmulas regionales que vinculan los valores de los parámetros con los descriptores físicos de la cuenca, empleando los modelos calibrados en las cuencas vecinas, o mejor aún mediante la explotación de algunas medidas puntuales en periodos de caudales importantes (Rojas-Serna, 2005).

8.2. Estructura del modelo GR4P

La filosofía general de construcción de modelos hidrológicos que sigue el equipo de trabajo que desarrolló la familia de modelos “GR”, consiste en reducir tanto como sea posible la complejidad de la estructura del modelo, procurando conservar la performance y eficiencia del mismo.

Por esta razón, las versiones indicadas previamente no difieren mucho entre sí, a pesar de estar orientadas a distintos pasos de tiempo. La estructura general de este tipo de modelos “GR” se conserva, y sólo existen pequeñas diferencias según sea el paso de tiempo (anual, diario u horario) y la finalidad de simulación o pronóstico.

En esta tesis se emplea un modelo, que hemos denominado GR4P, que permite realizar pronóstico hidrológico en paso de tiempo diario y horario, según la serie de datos introducida. Esta versión combina:

- La estructura del modelo GR4J (Perrin, 2000), desarrollado para correr a paso de tiempo diario y que cuenta con 4 parámetros.
- El modelo GRP (Tangara, 2005), que corresponde a la versión de pronóstico del modelo GR4J a paso de tiempo diario, pero diseñado con sólo 3 parámetros. Es decir, es una estructura simplificada para realizar pronóstico.
- Una adaptación de dicho modelo para correr también a paso de tiempo horario (Berthet, 2010) que dispone de 3 parámetros.

A partir de estas bases, se diseñó una estructura que tuvo por objetivo conservar la versión de simulación GR4J a paso de tiempo diario, tan completa como fuera posible, e incorporarle las variantes que le permiten correr a paso de tiempo horario como se verá más adelante, obteniendo de este modo un modelo GR4J/H.

Sobre la base de este nuevo modelo, se incorporó el algoritmo de actualización de estado para convertirlo en un modelo de pronóstico capaz de correr en paso de tiempo diario y horario.

Por esta razón, la descripción que se hace a continuación es básicamente la que corresponde a la versión GR4J, es decir, con paso de tiempo diario, aclarando luego las pequeñas modificaciones relativas al paso de tiempo horario. Finalmente, se describe la función que cumple el algoritmo de actualización de estado.

El modelo de pronóstico GR4P es entonces un modelo empírico, con una estructura de modelo conceptual de reservorios, que dispone de un procedimiento de control de las condiciones de humedad del reservorio que tiene en cuenta las condiciones anteriores y para garantizar el funcionamiento en simulación continua. Su estructura

combina un reservorio de producción y uno de transferencia, un par de hidrogramas unitarios, así como una función de apertura hacia el exterior no atmosférico (es decir, subterráneo), permitiendo simular el comportamiento global de la gran cuenca hidrológica, que incluye el flujo subterráneo.

En la Figura 8-1, se muestra un diagrama de la estructura del modelo empleado en esta tesis. En la misma, “Pk” es la lluvia del intervalo de tiempo “k”, y “E” es la evapotranspiración potencial media del mismo intervalo. Si bien la variable evapotranspiración potencial podría estar mejor representada por sus siglas “EP”, en vez de “E”, en este trabajo se adopta esta última dado que es la notación empleada por las publicaciones originales de este modelo, permitiendo de este modo una coherencia entre los trabajos de esta tesis y las publicaciones bibliográficas antecedentes.

- Neutralización

El primer paso es la neutralización entre “Pk” y “E” para determinar una lluvia neta “Pn”, y la evapotranspiración neta “En”. Estas se calculan por las siguientes ecuaciones, según corresponda:

$$\text{Si } P_k \geq E, \text{ entonces} \quad P_n = P_k - E \quad \text{y} \quad E_n = 0 \quad \text{Ec. 8-1}$$

$$\text{Si } P_k < E, \text{ entonces} \quad P_n = 0 \quad \text{y} \quad E_n = E - P_k \quad \text{Ec. 8-2}$$

- Función de salida

En el caso en que Pn es diferente de cero, una parte de Pn, denominada Ps, es asignada al tanque de producción. Esta porción es calculada por:

$$P_s = \frac{X_1 \left(1 - \left(\frac{S_K}{X_1} \right)^2 \right) * \tanh \left(\frac{P_n}{X_1} \right)}{1 + \frac{S_K}{X_1} * \tanh \left(\frac{P_n}{X_1} \right)} \quad \text{Ec. 8-3}$$

Donde X_1 (mm) es la capacidad máxima del reservorio de producción, y S_k el nivel S del reservorio de producción al comienzo del intervalo "k".

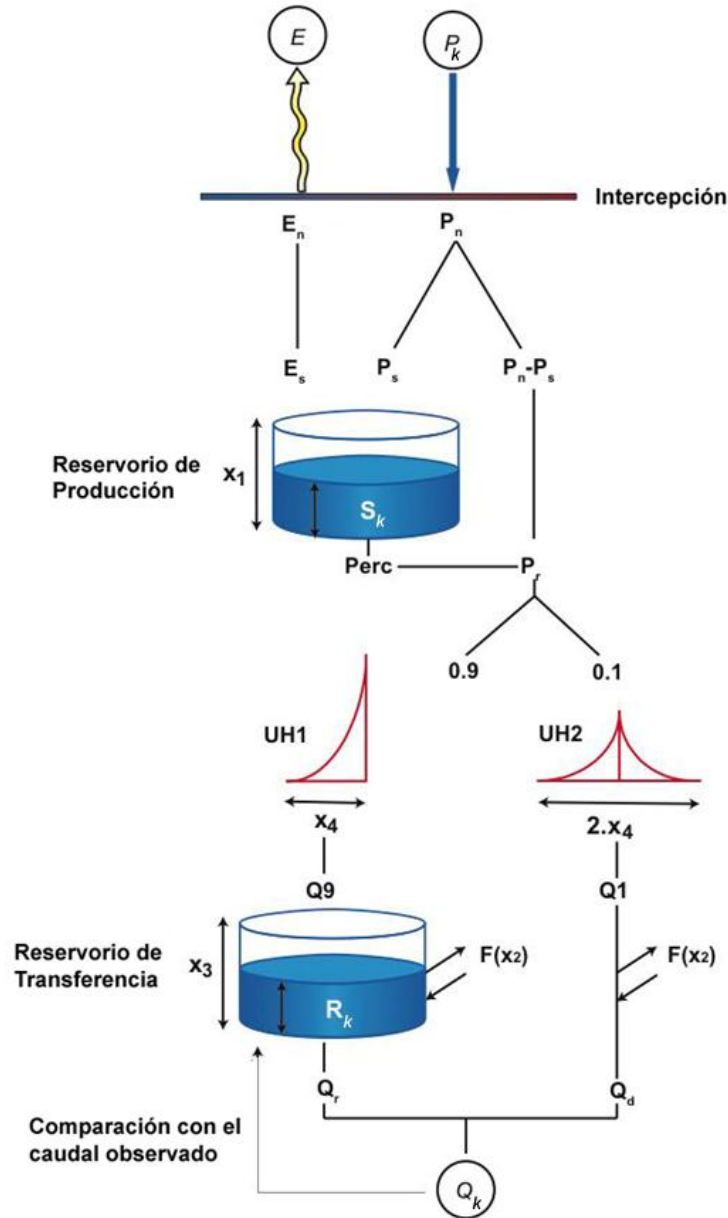


Figura 8-1: Estructura del modelo GR4P.

Cuando E_n es distinto de cero, una parte de E_n , denominada E_s , abandona el tanque de producción. Esta cantidad, se calcula según la siguiente ecuación:

$$E_S = \frac{S_K \left(2 - \frac{S_K}{X_1}\right) * \tanh\left(\frac{E_n}{X_1}\right)}{1 + \left(1 - \frac{S_K}{X_1}\right) * \tanh\left(\frac{E_n}{X_1}\right)} \quad \text{Ec. 8-4}$$

El contenido del reservorio resultante, está dado por:

$$S' = S_k + P_s - E_s \quad \text{Ec. 8-5}$$

Percolación

La Percolación ("Perc") que se produce desde el tanque de producción se calcula con la siguiente ecuación:

$$Perc = S' * \left\{ 1 - \left[1 + \left(\frac{4}{9} * \frac{S'}{X_1} \right)^4 \right]^{\frac{1}{4}} \right\} \quad \text{Ec. 8-6}$$

Como resultado, el contenido del tanque se convierte en:

$$S_k + 1 = S' - Perc \quad \text{Ec. 8-7}$$

La cantidad de agua Pr que finalmente llega a la etapa de transferencia del modelo está dada por:

$$Pr = Perc + (P_n - P_s) \quad \text{Ec. 8-8}$$

- Hidrogramas unitarios

Pr se divide en dos componentes de flujo, el 90% describe un hidrograma unitario, HU1, y se almacena en un depósito y el 10 % describe otro hidrograma unitario, HU2, el cual es simétrico. HU1 y HU2 dependen del mismo parámetro X4, que es el tiempo en la base de HU1 expresado en intervalos de tiempo.

Las ordenadas de los hidrogramas se calculan a partir de las curvas "S" denominadas SH1 y SH2, respectivamente, que corresponden a las funciones acumuladas de los hidrogramas.

SH1 se define por una función de tiempo:

$$\text{Para } t \leq 0, \quad SH1(t) = 0 \quad \text{Ec. 8-9}$$

$$\text{Para } 0 < t < X_4, \quad SH1(t) = \left(\frac{t}{X_4}\right)^{\frac{5}{2}} \quad \text{Ec. 8-10}$$

$$\text{Para } t \geq X_4 \quad SH1(t) = 1 \quad \text{Ec. 8-11}$$

SH2 se define de manera similar:

$$\text{Para } t \leq 0, \quad SH2(t) = 0 \quad \text{Ec. 8-12}$$

$$\text{Para } 0 < t < X_4, \quad SH2(t) = \frac{1}{2} * \left(\frac{t}{X_4}\right)^{\frac{5}{2}} \quad \text{Ec. 8-13}$$

$$\text{Para } 0 < t < 2 X_4, \quad SH2(t) = 1 - \frac{1}{2} * \left(2 - \frac{t}{X_4}\right)^{\frac{5}{2}} \quad \text{Ec. 8-14}$$

$$\text{Para } t \geq X_4 \quad SH2(t) = 1 \quad \text{Ec. 8-15}$$

Para el modelo con paso de tiempo horario, en lugar de utilizar un índice de potencia igual a 5/2, este se transforma a 5/4. Este pequeño cambio, pero altamente significativo para mejorar el desempeño del modelo frente a un cambio en el paso de tiempo empleado para la simulación, proviene de los trabajos realizados por Berthet, 2010, por lo que una explicación detallada que justifica dicho cambio puede buscarse en la fuente referida.

Los valores en las ordenadas de HU1 Y HU2 se calculan por medio de las ecuaciones:

$$UH1(j) = SH1(j) - SH1(j - 1) \quad \text{Ec. 8-16}$$

$$UH2(j) = SH2(j) - SH(j - 1) \quad \text{Ec. 8-17}$$

Donde j es un número entero.

En cada paso de tiempo k , los caudales Q_9 y Q_1 salidos de los dos hidrogramas corresponden a la convolución de las lluvias anteriores por el esquema de repartición dado por el hidrograma discretizado y son calculados por:

$$Q_9(k) = 0.9 * \sum_{j=1}^l UH1(j) * Pr(k-j+1) \quad \text{Ec. 8-18}$$

$$Q_1(k) = 0.1 * \sum_{j=1}^m UH2(j) * Pr(k-j+1) \quad \text{Ec. 8-19}$$

Donde

$$l = \text{int}(X_4) + 1$$

$$m = \text{int}(2.X_4) + 1$$

Siendo que $\text{int}(\cdot)$ denota la parte entera.

Función de intercambio con el exterior no atmosférico. El intercambio de agua subterránea se calcula a través de:

$$F = X_2 * \left(\frac{R_k}{X_3}\right)^{\frac{7}{2}} \quad \text{Ec. 8-20}$$

Donde R_k es el nivel en el reservorio al inicio del paso de tiempo, X_3 es la capacidad del tanque a un intervalo, y X_2 es el coeficiente de intercambio de agua que puede ser positivo en el caso que aporte o negativo en el caso de pérdidas a hacia las aguas subterráneas o bien puede ser nulo. La interpretación física de esta función de cambio, no es directa.

- Reservorio de transferencia

El nivel en el depósito de transferencia se modifica mediante la adición de Q_9 que proviene del hidrograma HU1 y F:

$$R' = \max(0; R_k + Q_9(k) + F)$$

El caudal Q_r que se obtiene a la salida del depósito de transferencia es:

$$Q_r = R' * \left\{ 1 - \left[1 + \left(\frac{R'}{X_3} \right)^4 \right]^{\frac{1}{4}} \right\} \quad \text{Ec. 8-21}$$

Entonces el nivel del reservorio se convierte en:

$$RK + 1 = R' - Qr \quad \text{Ec. 8-22}$$

La salida Q1 del hidrograma HU2 es sometida al intercambio para dar la componente de flujo Qd:

$$Qd = \max (0 ; Q1(k) + F)$$

El caudal total Q resulta entonces:

$$QK = Qr + Qd \quad \text{Ec. 8-23}$$

- Parámetros:

El modelo propuesto se compone de cuatro parámetros de calibración:

X1: Es la capacidad del reservorio de producción en mm.

X2: Es el coeficiente de intercambio subterráneo, en mm.

X3: Es la capacidad del reservorio de transferencia en mm

X4: Es el tiempo de base del hidrograma unitario HU1.

De una amplia muestra de cuencas hidrográficas el IRSTEA dispone de valores límites y medios de los parámetros (Perrin et al. 2007) para el modelo GR4J, es decir que pueden ser tomados como referencia sólo con el modelo funcionando con paso de tiempo diario, como se muestra en la Tabla 8-1.

Tabla 8-1: Parámetros modelo GR4J, (Perrin et al. 2007)

Parámetro	Media	Intervalo de confianza al 80%
X1 (mm)	350	100 a 1200
X2 (mm)	0	-5 a 3
X3 (mm)	90	20 a 300
X4 (días)	1.7	1.1 a 2.9

8.2.1. Aspectos de la estructura que adaptan el modelo al paso de tiempo diario y horario

Como se ha expresado repetidamente, el modelo GR4J (Perrin, 2000) es un modelo diseñado para trabajar en simulación con paso de tiempo diario. La estructura de ese modelo fue la que se conservó para los trabajos realizados en esta tesis, agregándole ciertas variantes que lo adaptan a poder trabajar en paso de tiempo horario, más un algoritmo que le permite trabajar en pronóstico.

A continuación, se detallan estas variantes relacionadas con su adaptación al paso de tiempo horario. Nuevamente se insiste en que para ver detalles de los justificativos que dieron origen a estos cambios remitirse a la tesis de Berthet (2010). Otro análisis completo del problema temporal del paso de tiempo y su impacto en la estructura de los modelos Génie Rural se presenta detalladamente en Ficchi (2017).

- NH: es una variable de programación que representa el máximo tiempo base del Hidrograma Unitario UH1. Se usa esta variable para dimensionar el vector que aloja las ordenadas del UH1, es decir es el máximo valor que puede adoptar el parámetro X4, que justamente representa el tiempo base del Hidrograma Unitario UH1.
 - En paso de tiempo diario, NH = 20 días.
 - En paso de tiempo horario, NH = 240 horas.

En términos de programación, las sentencias que establecen esta condición son simplemente:

NH=20

IF(Pdh==2)NH=240 !Pdh=1 (diario) Pdh=2 (horario)

- D: es el exponente adimensional involucrado en la expresión matemática del hidrograma unitario acumulado SH1 y SH2.
 - En paso de tiempo diario, $D = 5/2 = 2,50$.
 - En paso de tiempo horario, $D = 5/4 = 1,25$.

En términos de programación, las sentencias que establecen esta condición son simplemente:

D=2.5

IF(Pdh==2)D=1.25 !Pdh=1 (diario) Pdh=2 (horario)

- El factor de la expresión de la percolación, denominado a nivel de código del programa computacional como "Cop". La ecuación de la percolación (Ec. 8-24), que abajo se transcribe, también puede tener la forma que aparece a la derecha del signo igual, en la cual vemos que aparece una variable "Cop", que es justamente la que se modifica de acuerdo a que se esté trabajando con una serie con paso de tiempo diario u horario.

$$Perc = S' * \left\{ 1 - \left[1 + \left(\frac{4}{9} * \frac{S'}{X_1} \right)^4 \right]^{\frac{1}{4}} \right\} = S' * \left\{ 1 - \left[1 + \left(\frac{S'}{Cop * X_1} \right)^4 \right]^{\frac{1}{4}} \right\} \quad \text{Ec. 8-24}$$

En la expresión presentada, para paso de tiempo diario, el factor vale 4/9, pero multiplicando a X1 en el denominador, su forma se invierte valiendo 9/4. Ese factor vale:

- En paso de tiempo diario, $D = 9/4 = 2,25$.
- En paso de tiempo horario, $D = 21/4 = 5,25$.

En términos de programación, las sentencias que establecen esta condición son simplemente:

Cop=2.25

IF(Pdh==2)Cop=5.25 !Pdh=1 (diario) Pdh=2 (horario)

$$\text{PERC} = S * (1. - (1. + (S / (\text{Cop} * X1))^{**4.})^{**(-0.25)})$$

8.2.1. Algoritmo de actualización de estado vinculado al pronóstico hidrológico

El modelo en modo pronóstico, incorpora al proceso de simulación un algoritmo de actualización de datos de estado, que mejora el pronóstico aprovechando los propios datos de caudal medidos en tiempo real. En la misma Figura 8-1 se observa que existe un procedimiento de ajuste del reservorio de transferencia en función de los datos de caudales observados a tiempo real.

Como fue presentado previamente, un modelo de simulación utiliza como datos de entrada sólo la precipitación y la evapotranspiración potencial. Con una serie de estas variables, se estima una serie de caudales simulados, que representa la salida del modelo.

El modelo en modo pronóstico agrega a estas variables de entrada, también el caudal observado a tiempo real. De este modo, es posible estimar en cada paso de tiempo, cuál debería ser el estado del reservorio de transferencia para producir caudales como los que son observados a tiempo real.

De este modo, el modelo va ajustando (se podría decir también, calibrando) la variable de estado en cada paso de tiempo en función de los caudales que el sistema de colección de datos a tiempo real va arrojando a la computadora.

Sintetizando, para aclarar conceptos se puede decir que en el tiempo "t0" el modelo recibe como datos de entrada P(t0) y EP(t0), y genera el hidrograma para t1 (siendo t1 = t0 + Dt), es decir el valor Q1, y para todo valor de t > t0 (es decir Q2, Q3...). Cuando transcurre un Dt, y ya estamos en t1, es posible verificar si el caudal generado Q1 cuando estábamos en t0, coincide con el valor actual Q1' de caudal (conocido). Si no coincide, el modelo modifica su variable de estado R0 (para el paso de tiempo anterior), ajustando de este modo el estado del sistema para que Q1 sea

igual al observado $Q1'$. Es decir, el modelo arroja el valor que tendría que haber tenido R en el paso de tiempo anterior, para haber arrojado un valor de caudal $Q1$, igual al medido $Q1'$. Una vez que ajustó R , vuelve a correr y calcula $Q2$, con los datos de entrada $P1$ y $EP1$. Y así sucesivamente con todo el periodo de tiempo simulado.

Las versiones del modelo de pronóstico GRP desarrollado por Tangara (2005), y modificado por Berthet et al. (2009) presentan además otra función de asimilación (actualización) en su estructura de pronóstico. Dicha función captura la información del último error del modelo para actualizar la salida del modelo a través de un modelo multiplicativo ARIMA (Box y Jenkins, 1976). Pero esta segunda función referida no es incorporada al modelo empleado en esta tesis.

9. El algoritmo de optimización multi-objetivo MOCOM-UA

En este capítulo se presenta en forma detallada el funcionamiento del método de optimización MOCOM-UA, empleado en esta tesis. Por ser una metodología multiobjetivo, el proceso de optimización deberá acabar con un conjunto de individuos, todos ellos buenos, capaces de representar mejor que cualquier otro conjunto de individuos el comportamiento de la cuenca para pronosticar caudales. Dichos individuos estarán dotados de la historia hidrológica de la cuenca que representan.

A cada época la salva un pequeño puñado de hombres que tienen el coraje de ser inactuales.

G.K. Chesterton

Las investigaciones en métodos de optimización han conducido al uso de estrategias de búsqueda basadas en hacer evolucionar una población en dirección al óptimo (e.g. Brazil and Krajewski, 1987; Brazil, 1988; Wang, 1991; Duan et al., 1992, 1993; Sorooshian et al., 1993, entre otros). En este sentido, el algoritmo de optimización global mono-objetivo Shuffled Complex Evolution (SCE-UA) ha demostrado ser consistente, efectivo y eficiente en la localización del juego de parámetros del modelo hidrológico óptimo globalmente con respecto a alguna función objetivo dada, como lo prueban un gran número de trabajos publicados (Duan et al., 1992, 1993; Sorooshian et al., 1993; Luce and Cundy, 1994; Duan et al., 1994; Gan and Biftu, 1996; Tanakamaru, 1995; Tanakamaru and Burges, 1996; Yapo et al., 1996; Thyer et al., 1999; Lidén y Harlin, 2000; Senarath et al., 2000; Madsen et al., 2002; Vélez, 2003; entre otros).

Debido al inmediato éxito de este algoritmo evolutivo mono-objetivo SCE-UA, fue desarrollada su versión multi-objetivo MOCOM-UA por Yapó et al. (1998) en la cual se conserva la estrategia de búsqueda basada en algoritmos genéticos, que es el método de optimización que se utiliza en este trabajo.

9.1. Descripción del método

El método MOCOM-UA es una estrategia de optimización global diseñada para ser efectiva y eficiente frente a un amplio espectro de problemas. La estrategia MOCOM-UA combina las fortalezas de la búsqueda aleatoria controlada (Price, 1987), con la evolución competitiva (Holland, 1975), el ranking de Pareto (Goldberg, 1989) y la recientemente desarrollada estrategia de la búsqueda local multi-objetivo Simplex (extensión multi-objetivo del método de búsqueda local de Nelder y Mead (1965). A continuación, se realiza una descripción detallada conforme es descrito en su trabajo original (Yapó et al., 1998).

Sintéticamente ilustrado en la Figura 9-1 y la Figura 9-2 (Yapó et al., 1998), se parte de una muestra inicial de "s" puntos aleatoriamente generada en el espacio n-dimensional de los parámetros " $U^0(\Theta)$ ", que represente la incertidumbre inicial de estos parámetros (espacio de validez asumido "a priori"). Para cada punto, se computa un vector de funciones objetivo $F(\Theta)$, en nuestro caso sólo dos FO, luego se genera un ranking de la población y a continuación se realiza una selección de individuos empleando el ranking de Pareto propuesto por Goldberg (1989). Se seleccionan simplex de (n+1) puntos de la población de acuerdo a un método de selección de puntos robusto basado en el ranking (Whitley, 1989). Cada Simplex evoluciona luego en una dirección de mejora usando una estrategia de búsqueda que resulta ser una extensión multi-objetivo del método Simplex de Nelder y Mead. El procedimiento evolutivo genera nuevos puntos, la descendencia, que son en promedio mejor que los puntos originales, padres. Una aplicación iterativa del

procedimiento del ranking y evolución orienta la población en dirección del juego óptimo de Pareto. El procedimiento concluye automáticamente cuando todos los puntos de la población se vuelven mutuamente no-dominados.

Al igual que el método SCE-UA, el MOCOM-UA trata la búsqueda global como un proceso de evolución natural. La muestra de "s" puntos constituye la población. Cada miembro de la población es un padre potencial con la habilidad de participar en el proceso de reproducción. El "Simplex" elegido de la población es como un par de padres, con la diferencia de que en el simplex hay más de dos miembros (excepto para el caso trivial de $n=1$). Para asegurar que el proceso evolutivo es competitivo, se requiere que la probabilidad de que los "mejores" padres contribuyan a la generación de descendencia con mayor probabilidad que los "peores". Se usa para ellos una distribución triangular de probabilidades para asegurar esta competitividad. Luego se aplica la estrategia de búsqueda multi-objetivo local de cada simplex para generar la descendencia. La misma información contenida en el simplex es usada para dirigir la evolución en una dirección de mejora. Cada nuevo descendiente reemplaza al peor punto del subcomplejo. Esto asegura que cada padre tiene al menos una oportunidad de contribuir al proceso de reproducción antes de ser reemplazado o descartado; es decir, ninguna información de la muestra es ignorada.

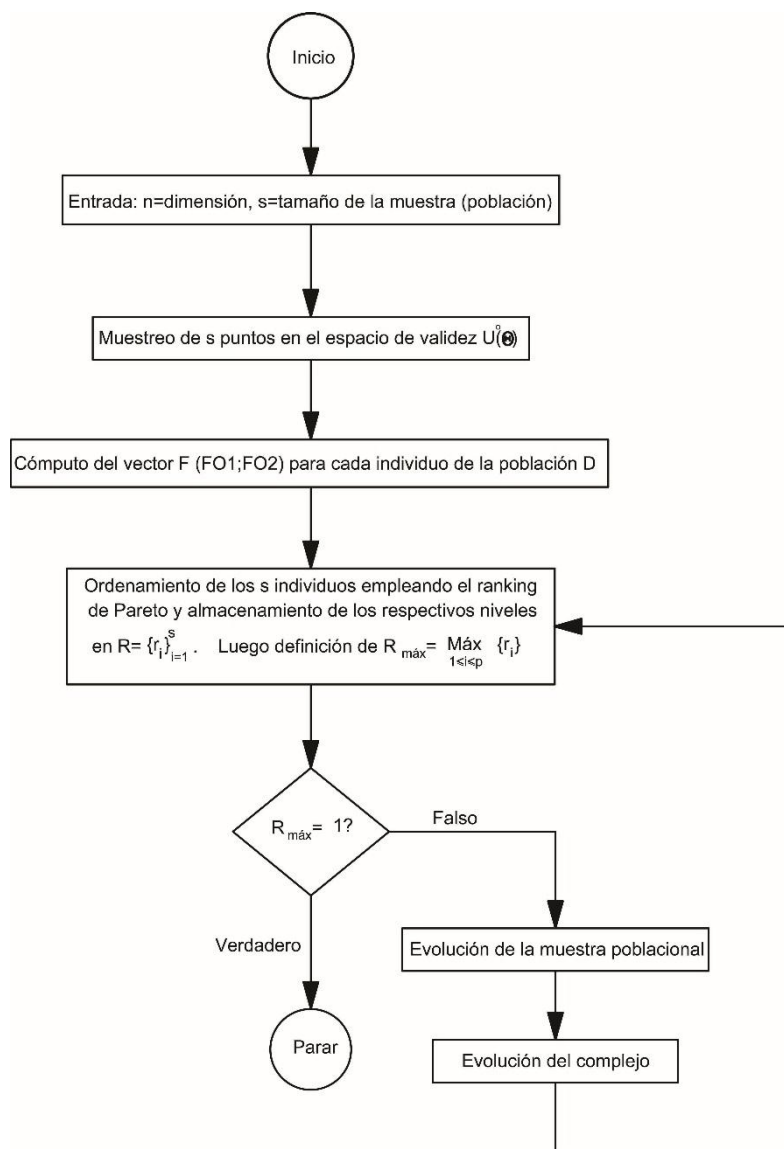


Figura 9-1: Diagrama de flujo del algoritmo MOCOM-UA (Traducido de Yapo et al., 1998).

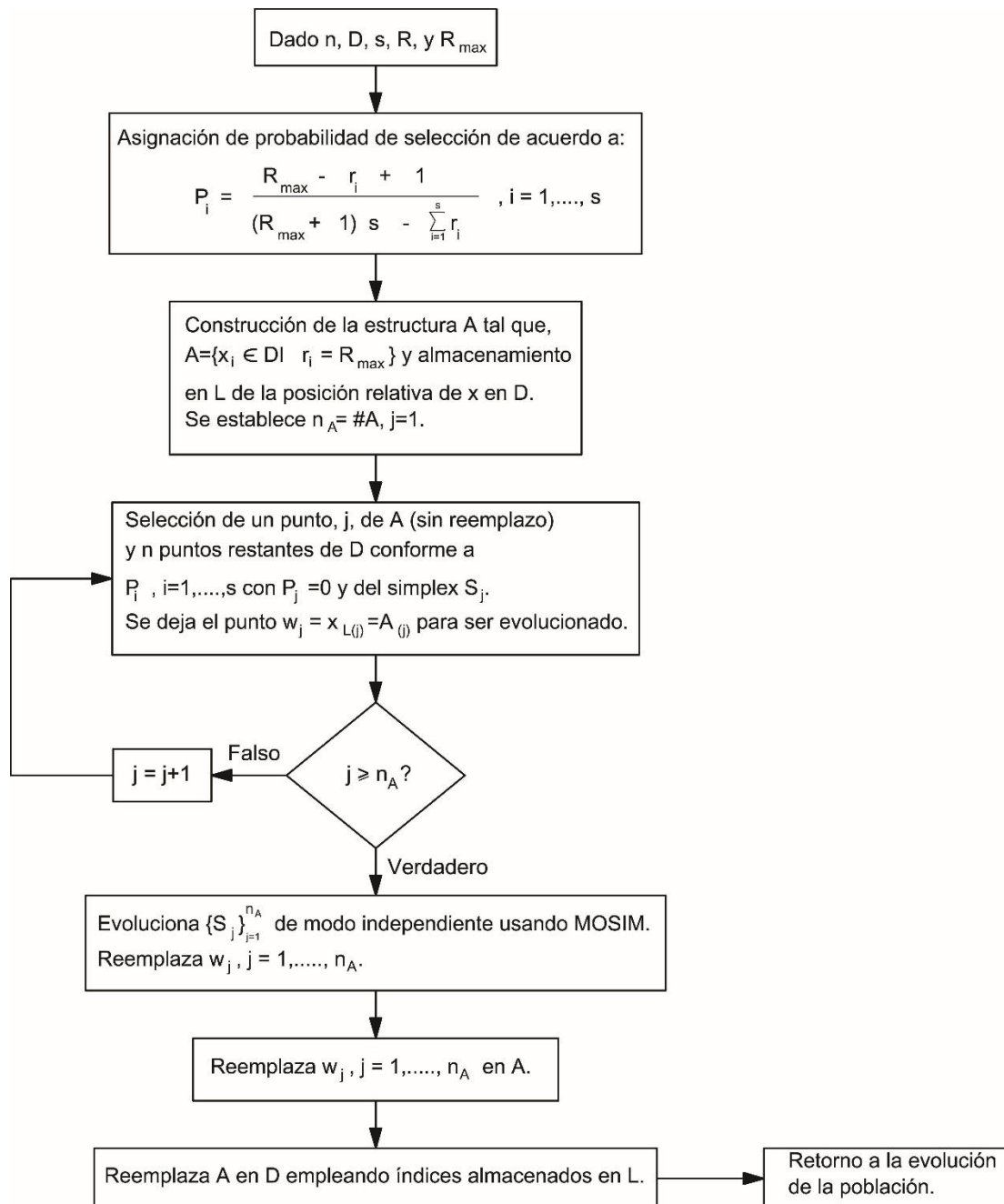


Figura 9-2: Evolución de los complejos en el MOCOM-UA (Traducido de Yapó et al., 1998).

9.1.1. Generación de la población inicial

Se establece el rango de validez de los parámetros del modelo matemático, que define un hipercubo en el espacio de los parámetros. Luego, se genera aleatoriamente una población de puntos en su interior. En ausencia de información

"a priori" respecto de la ubicación del conjunto de Pareto óptimo, se emplea distribución probabilística uniforme en la generación de puntos.

Por ejemplo, si se estuviese calibrando un modelo de dos parámetros ($P1$ y $P2$), cada uno de ellos tendría su rango de validez, de tal modo que se cumpla para cualquier punto ($P1$, $P2$) de la población inicial que:

$$P1_{\min} \leq P1 \leq P1_{\max} \quad \text{Ec. 9-1}$$

$$P2_{\min} \leq P2 \leq P2_{\max} \quad \text{Ec. 9-2}$$

De tal modo que la zona de validez de los valores de los parámetros está dada por un espacio de dos dimensiones (plano), como lo muestra la Figura 9-3.

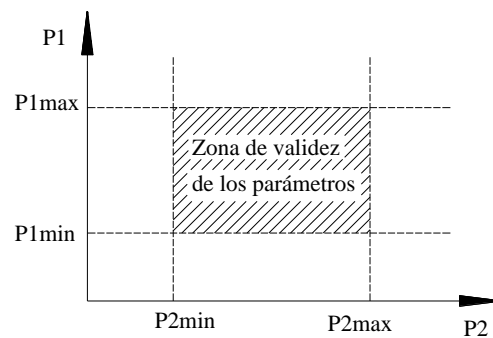


Figura 9-3: Zona de validez de los parámetros del modelo.

En la Figura 9-4, pueden observarse cuatro puntos A, B, C y D generados aleatoriamente con distribución uniforme en este espacio bidimensional asociado al supuesto modelo hidrológico de dos parámetros $P1$ y $P2$. Cada punto generado tiene asociado un valor de cada parámetro, así el punto B por ejemplo tiene asociado un valor de $P1$, que se ha denominado $P1B$, y uno de $P2$, denominado $P2B$. Como puede verse, el punto B satisface las desigualdades 9-1 y 9-2, que constituyen las restricciones del problema en cuestión, del mismo modo que lo hace el punto C. Mientras que los puntos A y D, están fuera de la zona de validez determinada por las

restricciones (4-79) y (4-80), por lo tanto, no son puntos válidos y no pueden ser considerados en la población.

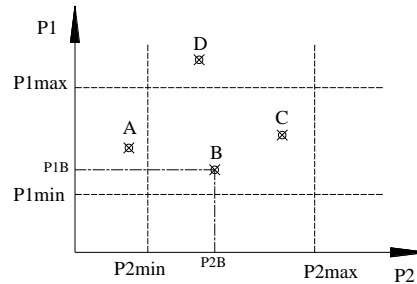


Figura 9-4: Puntos válidos y no válidos.

Si en vez de tratarse de un modelo de dos parámetros, fuese de tres, la zona de validez de los parámetros estaría definida entonces por un espacio de tres dimensiones, tal que los puntos en su interior cumplirían con las condiciones 9-1 y 9-2, más la siguiente:

$$P3_{\min} \leq P3 \leq P3_{\max}$$

Ec. 9-3

Extrapolando esta idea a cuatro, cinco y más parámetros, se puede imaginar la existencia de un hipercubo de n dimensiones (que naturalmente no puede visualizarse) dentro del cual deben caer los puntos aleatoriamente generados. Por ejemplo, el modelo hidrológico GR4P tiene cuatro parámetros, cada uno de los cuales tiene un rango dentro del cual puede tomar valores posibles o con sentido físico. Una generación aleatoria de puntos para los parámetros del modelo GR4P, debería caer dentro del hipercubo limitado por los rangos de variación admisibles de estos cuatro parámetros.

De este modo, cada punto generado en el interior del hipercubo de validez de los parámetros, tiene asociado un juego de valores de los parámetros del modelo a calibrar, que constituyen las coordenadas de este punto. Por ejemplo, si el modelo tiene cuatro parámetros, cada punto tiene cuatro valores asociados (que son sus

coordenadas en este espacio cuatri-dimensional), cada uno de los cuales corresponde a cada uno de los parámetros. Si hemos asignado a cada parámetro un valor dentro de su rango de validez, entonces los cuatro valores definirán un punto que caerá indudablemente en el interior del hipercubo imaginario.

La generación de puntos en el hiperespacio sigue, en este método, una ley uniforme de distribución probabilística, pero también podría pensarse en una generación aleatoria con distribución normal, o triangular, por ejemplo. Para ello se emplea, por ejemplo, el Método de Monte Carlo.

Cada punto, de este modo generado, es un individuo de una población primitiva y original que se dispone a evolucionar. Es decir, esta población irá evolucionando mediante un reemplazo de sus individuos por otros generados de algún modo. Luego de la evolución, la población resultará ser otra población nueva con otros, pero igual número, de individuos finales.

9.1.2. Clasificación y ordenamiento de los individuos de la población.

Para realizar la clasificación se evalúa cada una de las FO para cada uno de los puntos de la población. A los fines explicativos, se asume en esta descripción que la FO consiste en una medida de error (no en un valor de performance), de modo tal que el propósito de la optimización es minimizarla. Luego de la evaluación resultarán puntos que tienen menor valor que otros en ambas funciones objetivo, en cuyo caso no cabe duda que estos puntos son mejores que los otros. Pero el problema se plantea cuando se observa que hay puntos que tienen menor valor que otro en algunas FO, pero mayor en otras FO. No puede establecerse un orden de prioridad si hay puntos mejores que otros con el criterio de una de las FO, pero peores con el criterio de la otra. Por ejemplo, supóngase que se tiene dos FO, denominadas FO1 y FO2. Si un punto A tiene menor valor de FO1 que otro B, pero mayor valor de FO2, conforme se representa en la Figura 9-5, entonces no puede decirse que A es mejor que B, ni tampoco que B es mejor que A.

Es decir, por tratarse de un problema multi-objetivo, varios objetivos deben ser optimizados simultáneamente, y no es posible establecer una superioridad escalar entre los individuos, ya que cada uno tiene más de un componente, o sea más de una FO. Debe entonces establecerse una selección basada en un criterio multi-objetivo. El método MOCOM-UA utiliza el método de selección de Pareto (Goldberg, 1989). Surgen aquí el criterio de dominancia, según el cual son definidos puntos Dominados aquellos tales que existe otro punto que es mejor que ellos en todas las FO. Y, por consiguiente, los restantes puntos son No Dominados.

A partir de este concepto, el procedimiento continúa mediante una selección de todos los puntos no dominados de la población y se dice que forman el Nivel 1. Una vez que se retiran de la población esos puntos no dominados, habrá ahora nuevos puntos que son dominados, y otros no dominados. Entonces se retiran los nuevos puntos no dominados y se dice que pertenecen al Nivel 2. Una vez retirados estos puntos no dominados, se toman los restantes y se identifican los nuevos puntos no dominados, y así sucesivamente, hasta que resulte un conjunto de puntos tal que todos ellos son no dominados.

———— FO1
 - - - - - FO2

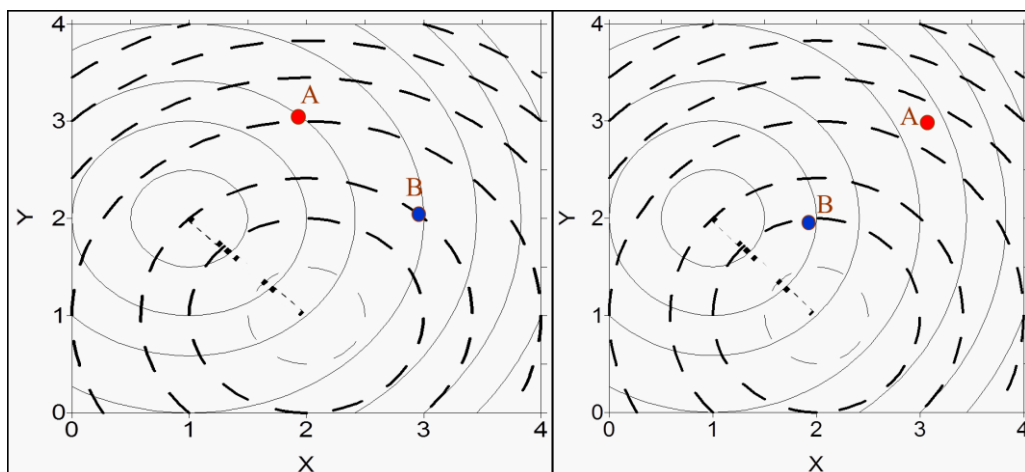


Figura 9-5: Criterio de dominancia en el espacio de los parámetros.

De acuerdo al esquema de jerarquización antes explicado, en cada nivel habrá un conjunto de puntos, como se ilustra en la Figura 9-6. Como ya no pueden organizarse los puntos de menor a mayor valor de FO (como se jerarquiza en un método de optimización mono-objetivo) puesto que hay más de una FO, se dice que los puntos del nivel 1, son todos mejores que los puntos del nivel 2, y estos mejores que los del nivel 3, y así sucesivamente. Pero no hay manera de identificar cuál de los puntos del nivel 1 es el mejor. Puede suponerse que todos los puntos de un mismo nivel tienen la misma jerarquía. El ranking de este modo logrado se denomina *Ranking de Pareto*.

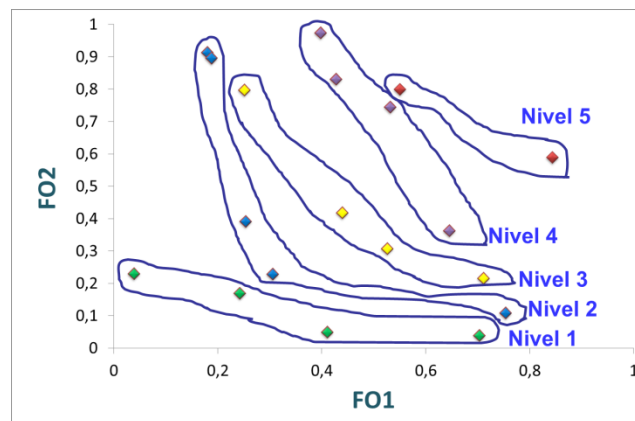


Figura 9-6: Niveles de dominancia en el ranking de Pareto.

En una población de "s" individuos, los niveles asignados son 1, 2, 3,..., R_{max} , donde $R_{max} \leq s$. Los individuos incluidos en los niveles inferiores del ranking están próximos al juego óptimo de Pareto, mientras que aquellos de mayor nivel están más alejados.

9.1.3. Selección basada en el ranking

En esta etapa se diseña un criterio para asignarle mayor o menor probabilidad de ser elegidos los puntos de la población para pasar a la etapa reproductiva. Dicho criterio está asociado a su posición en el ranking de Pareto. Sea "ri" el nivel del punto i. La asignación de la probabilidad de ser elegido está dada por el ajuste de la distribución de probabilidades de la población presentada en la Ec. 9-4.

$$P_i = \frac{(R_{max} - r_i + 1)}{\sum_j^s (R_{max} - r_j + 1)} = \frac{(R_{max} - r_i + 1)}{s^*(R_{max} + 1) - \sum_j^s r_j} \quad \text{Ec. 9-4}$$

El denominador de la Ec. 9-4 asegura que P_i para $i=1$ a s es una función de masa de probabilidad (pmf) tal que $\sum_i^s P_i = 1$. Ya fue demostrado por Whitley (1989) que la selección de puntos basado en niveles del ranking es un esquema más robusto que la selección basado en función para aplicaciones multi-objetivo.

9.1.4. Evolución multi-objetivo compleja

Una vez que fue asignado un nivel a todos los individuos, la evolución continúa mediante la selección de los simplex dentro de la población. Se eligen tantos simplex como el número de puntos del último nivel del ranking, o sea NRmax. Cada Simplex contiene "n+1" puntos, donde un punto es elegido del conjunto de NRmax peores puntos, y los remanentes "n" puntos son elegidos aleatoriamente de los "s - NRmax" puntos remanentes de acuerdo a la función de probabilidad de la Ec. 9-4.

La Figura 9-7 ilustra el procedimiento de la selección del Simplex para un problema simple de dos dimensiones con dos FO a ser minimizadas ($n=2$) y una población de siete puntos. El número entre paréntesis próximo a cada punto indica su nivel dentro del ranking de Pareto; los puntos A, B, y C tienen nivel 1, los puntos D y E tienen nivel 2, y los puntos F y G tienen nivel 3 ($R_{max}=3$). Dado que hay dos puntos con el peor ranking, se forman dos Simplex para evolucionar. Cada Simplex contiene uno de los puntos con peor nivel, y dos (n) de los puntos mejor "rankeados".

Esta estrategia de generación de Simplex permite identificar los peores puntos de acuerdo a un ranking global de la población (en vez de un ranking local del Simplex) y determina por lo tanto las direcciones globales de mejora. Dado que los simplex son hechos evolucionar de modo independiente, el procedimiento de evolución es bien apto para procesamiento en paralelo aprovechando las ventajas de tiempo computacional en caso de programar dicho proceso en esta forma.

Cuando todos los simplex han dado exactamente un paso evolutivo, los nuevos puntos reemplazan todos los peores puntos de la población original, creando en consecuencia una nueva población también denominada como la "próxima generación".

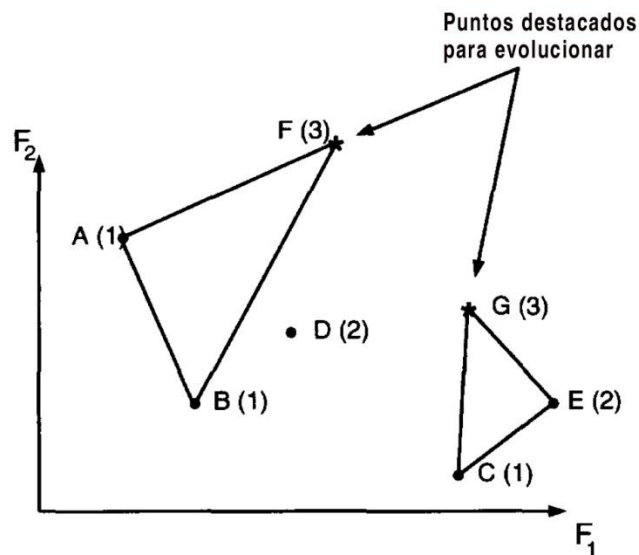


Figura 9-7: Esquema de selección del Simplex en un problema de dos objetivos (Fuente: Yapo et al, 1998).

Todos los puntos de la población original que no pertenecían al último (peor) nivel del ranking realizado, pasan a la nueva generación del mismo modo que en la población original, es decir, no cambian. De modo que la nueva población resulta ser igual a la original en la que sólo cambian los puntos del último nivel.

Los pasos del ranking de Pareto, generación de los simplex, y la evolución de estos simplex constituyen un loop principal del algoritmo MOCOM; iteraciones a través de estos pasos evoluciona la población en dirección del juego óptimo global de Pareto.

El algoritmo termina de un modo natural cuando todos los puntos de la población tienen ranking igual a uno (vale decir que no existen puntos inferiores que otros en la población), lo cual indica que no es posible encontrar mediante esta estrategia nuevas direcciones globales de mejora.

9.1.5. Evolución Multi-Objetivo del Simplex (MOSIM)

El procedimiento de evolución empleado para mejorar el peor punto es una extensión multi-objetivo del método Simplex (Nelder y Mead, 1965) y por lo tanto fue denominado MOSIM (Yapo et al., 1998).

El método MOSIM genera un nuevo punto que reemplaza el punto del Simplex ubicado en el último nivel del ranking global.

El nuevo punto es generado a partir de la aplicación de dos operaciones al simplex, denominadas "Reflexión" y "Contracción", según la Ec. 9-5.

$$S_{new} = \gamma S_g + (1 - \gamma) S_w \quad \text{Ec. 9-5}$$

Donde S_{new} es el nuevo punto, S_w es el peor punto "rankeado" dentro del Simplex, y S_g es la ubicación del centroide de los "n" mejores "rankeados" puntos del Simplex. Cuando $\gamma=2$, es obtenido un punto de Reflexión y, cuando $\gamma=0.5$, es obtenido un punto de Contracción.

La regla para elegir sea un punto de Reflexión (S_{ref}) o de Contracción (S_{con}) está basada en el concepto de dominancia multi-objetivo. El punto de Reflexión S_{ref} es aceptado si (y sólo si) es no-dominado con respecto a los "n" puntos que fueron usados para computar el centroide. Sin embargo, si el punto de Reflexión resultó ser dominado, es descartado, y se acepta en cambio automáticamente el punto de Contracción S_{con} . Este procedimiento produce que cada Simplex evolucione exactamente una vez y produzca un solo descendiente.

En la Figura 9-8 se ilustra el procedimiento MOSIM para un problema de dos FO (F_1 , F_2) con dos parámetros. Los objetivos a ser minimizados son la función cuadrática centrada en (0,0) y (1,0) y está representada por las líneas de contorno; las líneas continuas corresponden a F_1 , y las líneas de trazo corresponden a F_2 . El juego de Pareto está indicado por una delgada línea gris. En el ejemplo es dado un Simplex {a

b c}, a partir del cual se prueba en primer lugar un punto de Reflexión; el mismo es descartado porque es dominado por el punto {a}. Por lo tanto, el punto de Contracción es aceptado automáticamente y pasa a reemplazar al punto {c} en el Simplex original.

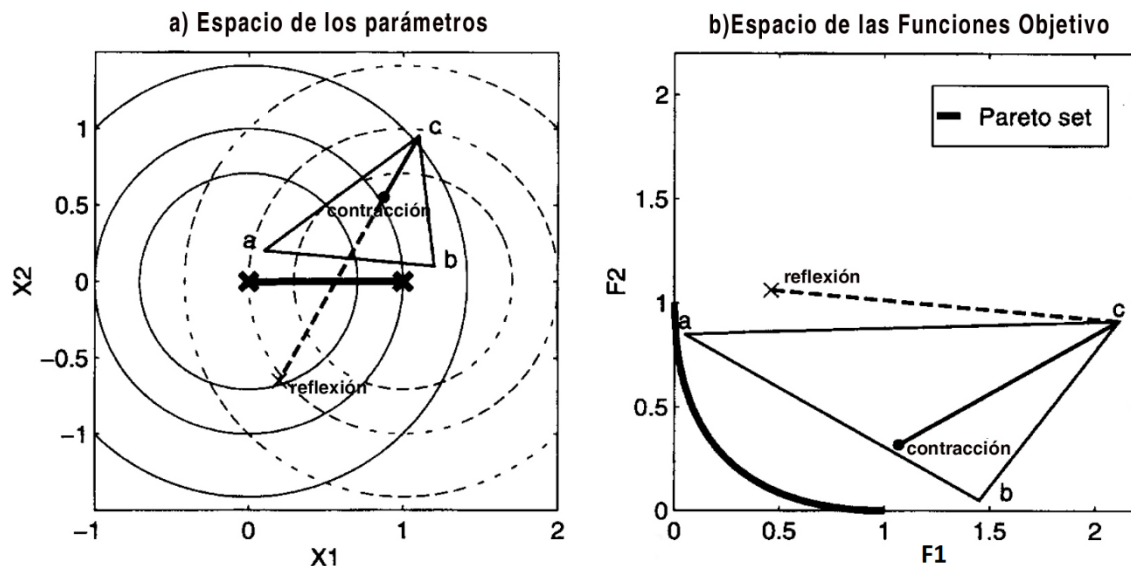


Figura 9-8: Esquema del procedimiento MOSIM (Fuente: Yapo et al., 1998).

9.1.6. Un esquema del proceso evolutivo del MOCOM-UA

La progresión del algoritmo MOCOM-UA es ilustrado por Yapo et al. (1998) en la Figura 9-9 para un problema sencillo de dos dimensiones y dos FO. Se muestra una población inicial de 30 puntos aleatoriamente generados en el espacio de validez de los parámetros. Los cuatro puntos negros son no-dominados, mientras que los 26 círculos denotan puntos dominados. Luego de un loop (gráficos c y d), pasan a haber siete individuos no-dominados (puntos negros). Luego de una secuencia de loops, el frente de Pareto se aproxima con los puntos negros hacia el verdadero frente de Pareto tanto en espacio de los parámetros como de las FO. Los puntos finales son todos mutuamente no-dominados y proveen una idea muy próxima de la región de Pareto (gráfico e), así como de la curva de compromiso.

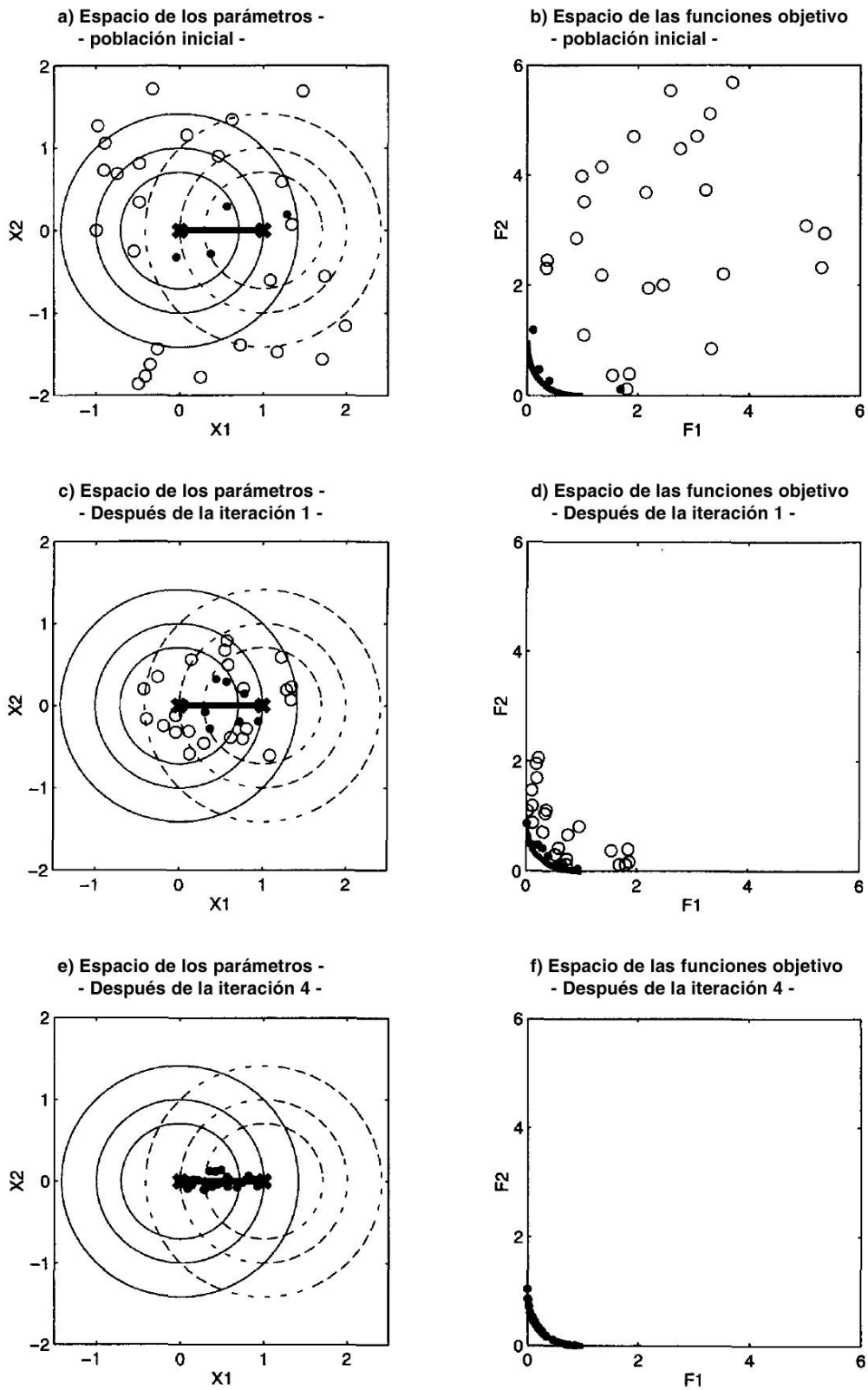


Figura 9-9: Esquema del MOCOM-UA para un problema de dos objetivos. (Fuente: Yapo et al., 1998)

10. Descripción de la cuenca y series de datos disponibles

Este capítulo presenta la cuenca hidrográfica en la cual tienen lugar las aplicaciones realizadas en esta tesis, y se describen los datos empleados que constituyen la información de base para convalidar los trabajos. Si bien se contó para ello con una rica serie de datos hidrometeorológicos, es importante remarcar que cualquiera sea la cantidad de datos empleados, jamás dejará de ser un conjunto de fotografías tomadas sobre la compleja naturaleza, lo cual devuelve ineludiblemente un conjunto de errores que incorporan incertidumbre a los resultados.

El auténtico conocimiento es conocer la extensión de la propia ignorancia.

Confucio

Los trabajos realizados en esta tesis están orientados principalmente a establecer herramientas de utilidad para el operador de un sistema de pronóstico hidrológico. El proceso de desarrollo de esas herramientas, prevé un procedimiento de convalidación del modelo hidrológico de pronóstico que depende fuertemente de la calidad de los datos empleados.

Por la razón expuesta, se consideró conveniente realizar los estudios empleando la cuenca del río Moros con su cierre a la altura de la localidad de Concarneau, en la Bretaña francesa. Dicha cuenca dispone de suficiente información hidrométrica, y además cuenta con experiencia de empleo del modelo GRP en la misma.

El río Moros nace en las proximidades de Rosporden, en la Bretaña francesa según la publicación digital de la asociación "Eau & Rivières de Bretagne" (<http://www.eau-et-rivieres.fr>, 2016) a una altitud de 140 m sobre el nivel del mar, y desemboca en

las proximidades del puerto de Concarneau, recorriendo su cauce principal unos 18,8 km de longitud. Es decir que se trata de un curso de agua con una pendiente moderada. En su desembocadura en el puerto de Concarneau, la cuenca hidrográfica tiene un área de 45 km², en la cual viven alrededor de 3000 habitantes.

En este estudio se cuenta con información de caudal medida por un limnógrafo ubicado en el "Moulin du Henant", que corresponde a una sección del río ubicada a unos 3,8 km antes de la desembocadura del río en el puerto, y a unos 21 m de altura sobre el nivel del mar, concentrando allí un área de aporte correspondiente a sólo 20,97 km².

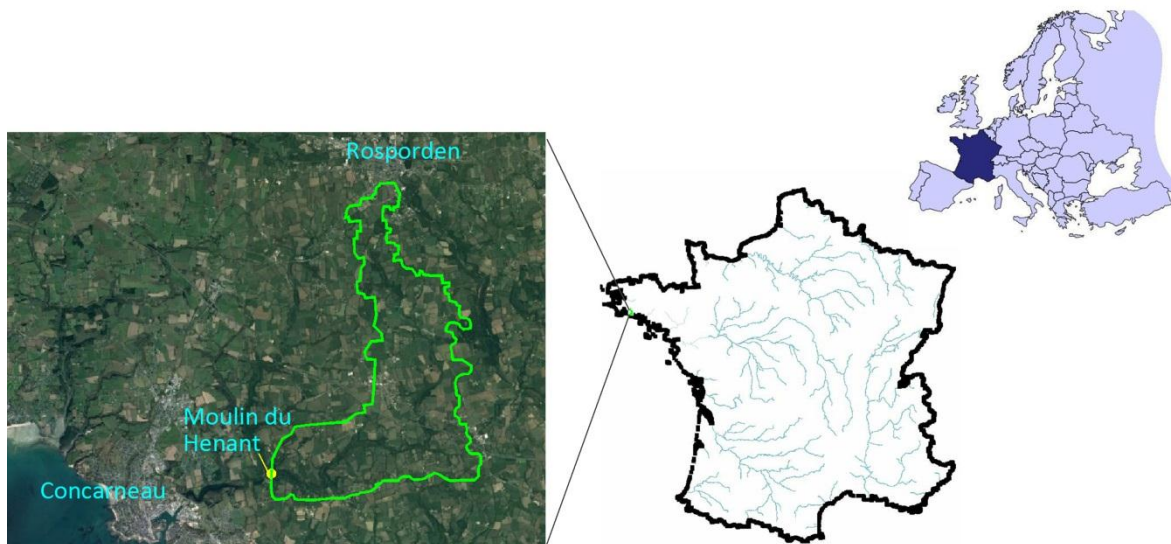


Figura 10-1: Ubicación de la cuenca "Moros" en Concarneau.

En la sección referida se dispone de un registro de información de caudales desde 1967 a 2010, y meteorológicos desde 1958 a 2010, según datos publicados en Cemagref (2012). De acuerdo a este banco de datos, el río tiene un módulo de 1,35 mm/día, que equivale (escalado al área de la cuenca de aporte) a 0,33 m³/s. Este pequeño módulo es identificado adicionalmente en la geomorfología del cauce que presenta secciones reducidas, como se observa en las fotos de la Figura 10-2, tomadas en las proximidades de la sección de aforo (Moulin du Hénant). La

precipitación media anual de la cuenca es de 1077 mm, con un aporte nival despreciable (0,5%). La temperatura media de la cuenca es de 11,4 °C, y según el método de Penman posee una Evapotranspiración Potencial que ronda los 637 mm/año.

En el periodo indicado de medición, se registró un caudal máximo el día 28/01/1995 de 4,56 m³/s, seguido por el evento del 05/01/2001 con 3,48 m³/s. Así mismo, el 05/07/1991 se registró la máxima precipitación del periodo de datos en 55 mm/día, seguido por el 07/07/2004 con una precipitación de 52 mm durante ese día.



Figura 10-2: Sección de aforo en las proximidades del "Moulin du Hénant".

Si bien en Cemagref (2012) se indica que se cuenta con registros de caudales para esta cuenca desde 1967 hasta 2010, un procesamiento y depuración de la información existente por parte del equipo de investigadores y técnicos del IRSTEA (ex Cemagref), rescataron la serie a partir de 2003. Sin embargo, desde 2009 se poseen valores de precipitación con paso de tiempo horario, por lo que en este trabajo se tomó la serie desde 2009. También se dispuso de datos hasta 2013, provisto por el personal del Irstea, a pesar de no haber sido publicados formalmente todavía.

En consecuencia, se rescató una serie de datos de precipitación media de la cuenca con paso de tiempo horario y caudales medios horarios de la sección de aforo, desde el 01/08/2009 a las 6 hs., hasta el 01/08/2013 a las 5 hs, que representan un total de 35.066 datos durante 4 años.

Mayores detalles de la cuenca hidrográfica pueden encontrarse en la página de internet del Ministerio de Ecología de Francia (<http://hydro.eaufrance.fr/stations/J4514010>), o a partir de las fichas técnicas actualizadas del instituto IRSTEA (Antony). De todos modos, a los fines de capturar una concepción general de su comportamiento, se presentan a continuación la curva de frecuencia acumulada de caudal y precipitación media de la cuenca (Figura 10-3).

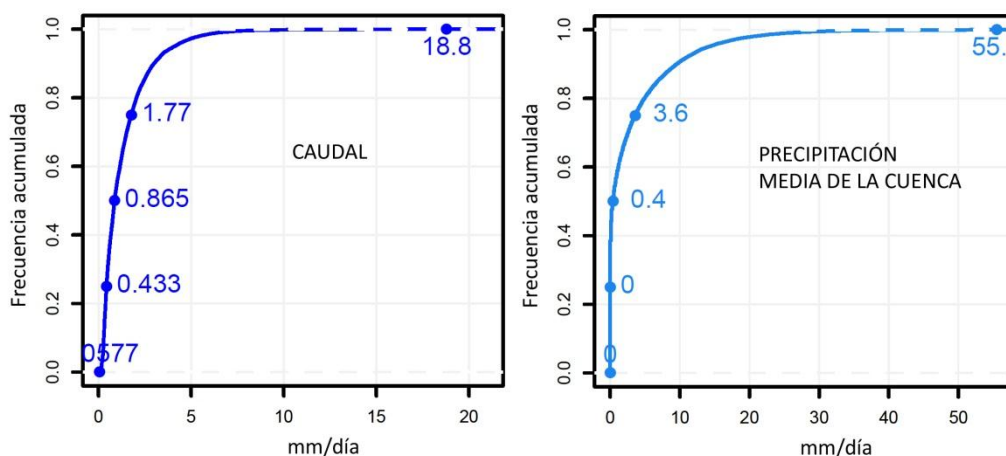


Figura 10-3: Curvas de frecuencia acumulada de caudal y precipitación media de la cuenca.

11. Propuesta metodológica

En este capítulo se especifica la metodología empleada, incluyendo el tratamiento de origen que se dará a los datos, las restricciones de modelación y simulación, así como los criterios de desempeño adoptados para los experimentos numéricos. Se plantea entonces un protocolo de trabajo para obtener resultados comparativos.

Somos lo que hacemos día tras día. La excelencia no es un acto, sino un hábito.

Aristóteles

En el capítulo relativo a la "Problemática", se describió un conjunto de escenarios operativos en pronóstico hidrológico que interfieren en el adecuado desempeño de un modelo hidrológico P-Q que fue calibrado para una condición de trabajo específico. Allí se planteó que un modelo P-Q de pronóstico hidrológico calibrado en condiciones ideales, con datos de caudal a tiempo real disponibles (permitiendo la actualización de estado en cada paso de tiempo), y para un determinado horizonte de pronóstico, está limitado a esas condiciones ideales.

En consecuencia, la calidad del pronóstico del mismo modelo sin disponibilidad de caudales a tiempo real, o bien para horizontes de pronóstico distintos al empleado para calibrarlo, es altamente cuestionable. En efecto, estos nuevos escenarios dan origen a un conjunto de nuevas incertidumbres.

Del contraste entre los escenarios extremos planteados (con datos de Q a tiempo real, y sin datos; un horizonte de pronóstico, frente a otro), se identifica la importancia de establecer una metodología de cuantificación de situaciones intermedias a través de la confrontación de dichos escenarios.

Del planteo de ese estado de situaciones, surgieron los siguientes interrogantes científicos:

- ¿Es posible controlar en tiempo real, mediante las combinaciones de valores paramétricos del modelo, su adaptabilidad a los diversos escenarios?
- ¿Existe algún juego de valores para los parámetros capaz de satisfacer a un mismo tiempo diversos escenarios con aceptable precisión?
- ¿Es posible contar con alguna metodología capaz de ofrecer anticipadamente al tomador de decisión, un conjunto de soluciones de compromiso?

En consecuencia, en este trabajo se realizaron tres acciones orientadas a aproximar una respuesta en forma conjunta a estos interrogantes, que permita ofrecer elementos concretos para disponer de una herramienta útil y práctica en el servicio operativo de pronóstico.

En primer lugar, se llevó a cabo la aplicación de una metodología binaria simulación/pronóstico (Alonso et al., 2015), como enfoque de una visión más matizada y continua entre dos situaciones de pronóstico: i) en contexto bien instrumentado (escenario 1), donde la información a tiempo real puede ser plenamente puesta al servicio de la asimilación, y ii) poco instrumentado (escenario 2), donde el modelo está levemente restringido ya que carece de información de caudal a tiempo real, dejado por lo tanto libre de evolucionar en condiciones próximas a las de un modelo de simulación.

En segundo lugar, se llevó adelante un análisis semejante al anterior, pero confrontando la performance del modelo de pronóstico para un horizonte de antelación (escenario 1) respecto de su performance para otro horizonte distinto, adoptado como escenario 2 (Alonso et al., 2016).

Para llevar a cabo ambas metodologías se hizo uso del método de optimización multi-objetivo MOCOM-UA (Yapo et al., 1998) aplicado sobre el modelo de pronóstico hidrológico GR4P (Tangara, 2005 y Berthet, 2010), sometiendo la capacidad del método de encontrar juegos de valores de parámetros capaces de alcanzar a la vez buenas performances para ambos escenarios descriptos.

Ahora bien, en el proceso de optimización se observó una aparente incapacidad por parte del método de encontrar de manera robusta en todos los experimentos de optimización el mismo frente de Pareto.

En consecuencia, una tercera acción abordada en este trabajo puso foco principal sobre este aspecto, por lo que se llevó a cabo ciertas modificaciones en el algoritmo de búsqueda, orientadas a mejorar su eficacia para llegar sucesivamente al mismo frente de Pareto, o semejante (Alonso, et. al, 2017). Se procuró de este modo tender a un método de optimización multi-objetivo (que podríamos llamar “MOCOM adaptado”) más robusto que su versión original.

11.1. Análisis retrospectivo

Los estudios realizados en este trabajo tienen una metodología de tipo retrospectiva. En el contexto del pronóstico de las series de tiempo, el concepto de análisis retrospectivo hace referencia al proceso de evaluación de la precisión del método de pronóstico mediante el uso de datos históricos existentes. El análisis retrospectivo se utiliza para calcular la precisión futura esperada de un método de pronóstico, lo que resulta útil para evaluar qué modelo de pronóstico debería considerarse más preciso, o bien qué juego de valores de los parámetros son los más adecuados para un modelo determinado.

En términos prácticos, el procedimiento de calibración y verificación del modelo se realiza con valores históricos conocidos, de la siguiente manera. Se dispone de una serie de datos de lluvia y de caudal medidos en la cuenca de estudio. Se toma un

tramo de la serie histórica disponible, y se realiza el pronóstico hacia el futuro del tramo adoptado (pero todavía en el pasado respecto del presente actual). Es decir, el valor pronosticado por el modelo también se encuentra en el dominio temporal del "pasado". Sin embargo, dado que conocemos el valor verdadero, es posible calibrar el modelo cotejando con estos valores de referencia.

11.2. Periodo de adecuación o "calentamiento" del estado del modelo

Conforme se ha expresado en el apartado referido a la descripción de los "aspectos complementarios de modelación", es altamente recomendado por varios autores (Dooge, 1973; Vélez y Francés, 2008; Tangara, 2005; Linsley, 1982; Kitanidis y Bras, 1980) un ajuste de los valores iniciales de las variables de estado. Por otra parte, Berthet et al. (2009) demuestran que las condiciones iniciales no son altamente influyentes en los resultados, sin embargo recomiendan considerarlas cuando es posible.

El modelo hidrológico empleado dispone de un conjunto de variables de estado, a las cuales les son asignados valores iniciales por el modelador, con algún criterio asociado normalmente al sentido conceptual que representa la variable y/o a la experiencia en el uso del modelo en cuanto al valor inicial conveniente para cada variable. Cuando se comienza a simular una serie de datos, el modelo procesa internamente esa información, y va modificando los valores de las variables de estado, que le fueron asignados arbitrariamente por el modelador. Si bien ese valor inicial fue asignado con algún criterio, se irá adecuando en virtud de la estructura del modelo y en función de los datos de entrada. Por esta razón, conviene dejar al modelo libre de evolucionar antes de tomar sus datos de salida como ciertos. En otras palabras, conviene dejar una primera parte de la serie para que las variables de estado se acomoden, y empiece entonces el modelo a producir valores de salida confiables.

Para lograr ese objetivo, se toma para la calibración toda la serie menos un tramo inicial de ajuste de variables de estado. Supongamos que disponemos de una serie de unos 10 años. Entonces para esta determinación de las variables de estado, se pueden invertir los primeros 2 años, y dejar los siguientes 8 años para la calibración. O sea, se consideran los caudales de salida de los últimos 8 años para comparar con los caudales observados y calibrar el modelo con esta comparación. En la práctica, el modelo computacional arroja un vector de caudales simulados (salida del modelo), que representa el hidrograma durante el periodo simulado, del cual se toman los últimos ocho años, por ejemplo, para calibración.

11.3. Criterio de partición de las series de datos

La calibración del modelo matemático es seguida de un proceso de validación (también denominado de verificación o control). Para este fin, el tratamiento de la serie para llevar a cabo ambos procesos fue la recomendada por Klemes (1986): para cada cuenca se divide la serie de datos en dos o más tramos independientes y no superpuestos (Oudin et al., 2005; Berthet et al., 2010).

En esta tesis, la serie de tiempo disponible de datos es dividida en dos partes. La primera es empleada para calibración, y la segunda para verificación. Luego se invierte, y se emplea la segunda para calibración y la primera para verificación. Ahora bien, como además se necesita disponer de un rango de estabilización de las variables de estado, como se explicó previamente, se reserva la primera parte de cada tramo para ese fin. Para la cuenca en estudio, se dispone de una serie de 4 años de registro ininterrumpido de datos de calibración, que corresponde a 35066 pasos de tiempo horarios. Dicha serie es dividida en dos tramos superpuestos, pero teniendo la precaución de que el periodo de superposición corresponde al periodo reservado para inicialización de las variables de estado (periodo de "calentamiento" del modelo). Como puede verse en la Figura 11-1, el primer periodo adoptado es el

tramo 1, que posee 22.026 pasos de tiempo (pdt) horarios, del cual los primeros 9000 pdt son destinados a la inicialización del modelo. Luego el tramo 2, posee 22.040 pdt, siendo los 9000 pdt primeros superpuestos a los 9000 últimos del tramo 1. Sin embargo, ya que son destinados a la inicialización del modelo, no son tomados como superpuestos ya que no son empleados en el proceso de calibración.

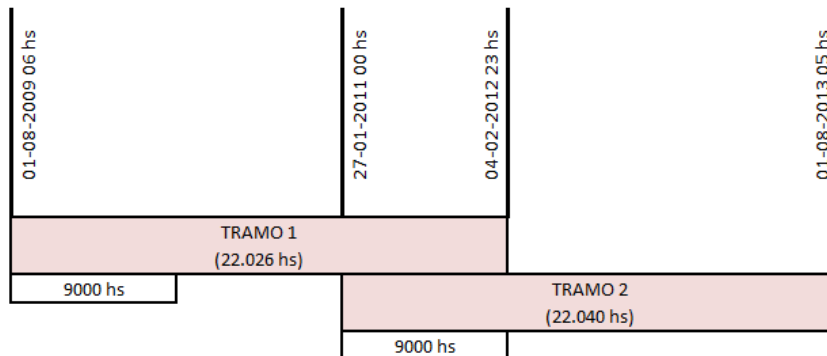


Figura 11-1: Periodos de ajuste y verificación de los datos empleados.

11.4. Procedimiento de calibración y control multi-objetivo

La metodología abordada en esta tesis se basa en la aplicación de un método de optimización multi-objetivo. Debido a esto el proceso de calibración y control tiene algunas particularidades que no lo hacen totalmente adaptable al método de calibración mono-objetivo. En efecto, en optimización mono-objetivo, se realiza primeramente un proceso de calibración, encontrando el juego de valores de parámetros que mejor ajusta el modelo para el tramo de serie reservada para calibración. Luego, se toma ese juego de valores de parámetros, y se corre el modelo en el tramo de la serie reservado para verificación. Entonces, se verifica que ese juego de valores de parámetros también optimiza el ajuste en verificación. Este procedimiento es trivial ya que emplea una sola FO, y por lo tanto el propósito es que el método minimice (en caso de que la FO cuantifica un error) o bien maximice (en caso de que la FO cuantifica una performance) el valor de la única FO.

Ahora bien, en optimización multi-objetivo no se tiene un solo juego de valores de parámetros, sino un conjunto de juegos de valores de parámetros. Entonces, la verificación consiste en estimar para el tramo de la serie reservado a este fin, si los valores de los parámetros de todos los puntos de la región de Pareto, producen FO con posiciones equivalentes en la región de Pareto. Es decir, si de acuerdo al proceso de calibración, encontramos que un juego de valores de parámetros minimiza la FO1, y otro la FO2, entonces deberíamos encontrar que, con esos mismos juegos de valores de parámetros empleados en el tramo de serie de verificación, también se minimizan los mismos criterios respectivamente. Y no sólo eso, sino que los juegos de valores de parámetros de la región de Pareto, entre los que minimizan FO1 y FO2, deberían minimizar de modo similar el tramo de verificación. Si se diera una situación así, implicaría que el procedimiento de calibración se desarrolló robustamente, o sea que variando la serie sobre la cual se aplica, siempre los juegos de valores de la región de Pareto responden equivalentemente a satisfacer del mismo modo las FO respectivas.

En forma sintética, se puede decir que el procedimiento consta de una primer fase, en la cual se toma el "Tramo 1" y se realiza la optimización multi-objetivo, obteniendo como resultado de esta etapa la región de Pareto, la cual se supone que contiene los juegos de valores de parámetros que mejor ajustan FO1 y FO2, más la familia de juegos de valores de parámetros que representan soluciones intermedias entre ambos extremos. Luego, en una segunda fase, se evalúan las FO1 y FO2 para todos los individuos (recientemente obtenidos por optimización) pero empleando el "Tramo 2", en un proceso que emula la validación mono-objetivo. Aquí es donde se puede evaluar si se ha realizado una optimización "robusta" del modelo. Se supone que cada juego de valores de parámetros (o sea, cada individuo) debería ocupar un lugar semejante a la región de Pareto optimizada por el tramo 1. En otras palabras,

ambas regiones de Pareto (la obtenida por el tramo 1 y la verificada en el tramo 2) deberían parecerse bastante.

Idealmente, el individuo que maximiza la FO1 en el tramo 1, debería también maximizar la FO1 en el tramo 2. Lo mismo para la FO2, y para las soluciones intermedias de compromiso (Figura 11-2).

Por último, la siguiente prueba consiste en repetir el experimento, pero invirtiendo los "roles" de los tramos. En este caso, se optimiza con el segundo tramo y se valida con el primero. También en este caso deberíamos obtener como resultado Regiones de Pareto semejantes.

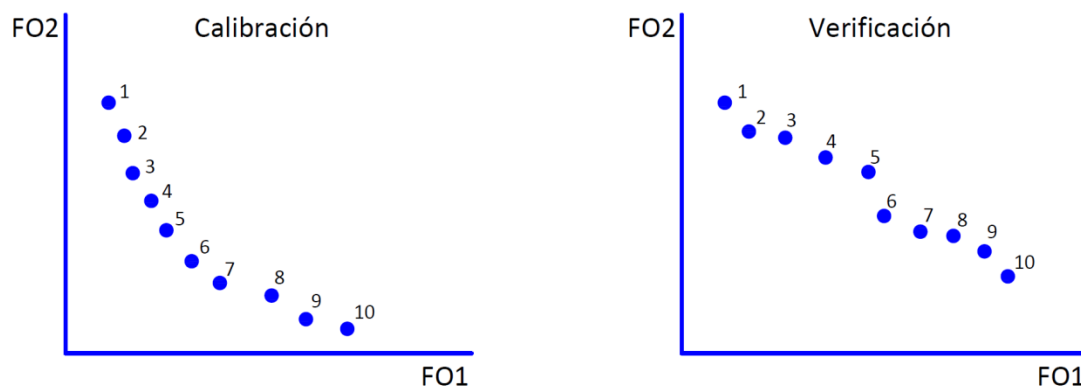


Figura 11-2: Ejemplo de frentes de Pareto en Calibración y Verificación respectivamente.

11.5. El criterio de ajuste

Como se ha expresado previamente, en este trabajo se emplea el modelo corriendo en dos modos distintos: simulación y pronóstico. Normalmente, para hacer la optimización de un modelo hidrológico en simulación, se emplea alguna FO clásica, como Mínimos Cuadrados por ejemplo, y luego se verifica con algún estimador de la calidad del ajuste, como por ejemplo el coeficiente de Nash-Sutcliffe. En pronóstico, también a veces se realiza el ajuste minimizando el coeficiente de mínimos cuadrados, y luego se verifica mediante el Coeficiente de Persistencia, por ejemplo.

Sin embargo, desde hace algunos años, se ha empezado a popularizar en la bibliografía la técnica de emplear el mismo coeficiente para calibrar que para verificar el modelo. Así, por ejemplo, actualmente es habitual emplear directamente el criterio de Nash-Sutcliffe para calibración y verificación del modelo en modo simulación, o bien, el empleo del criterio de Persistencia en calibración y verificación, para el ajuste del modelo en modo pronóstico.

En este trabajo se emplea el modelo en modo simulación y en modo pronóstico, pero en ambos casos el propósito es evaluar su desempeño para realizar pronóstico. Por esta razón, un buen indicador es el criterio de persistencia. De todos modos, a los fines de estudiar integralmente la performance del modelo, se emplea sucesivamente el estimador de Nash & Sutcliffe (NSE), de Kling & Gupta (KGE), y el de Persistencia.

11.6. Respecto de los límites de los parámetros del modelo GR

Se ha descrito previamente en este trabajo el modelo GRP, y se han sugerido valores límites de los parámetros del modelo GR4J. A partir de esta información, y a los fines de comprender rangos mayores para abarcar completamente los límites del modelo en la historia de aplicación del mismo, se emplearon los siguientes límites para las aplicaciones realizadas en este trabajo:

X1 (Capac. máx del reserv de producc): [0 ; 1000]

X2 (Coef. de intercambio): [-10 ; 5]

X3 (Capac a un día del reservorio de propagac): [0 ; 500]

X4 (Base del hidrograma unitario):

En el caso que sea con paso de tiempo diario: [0 ; 10] (o sea $NH/2 = 20/2$)

En el caso que sea con paso de tiempo horario: [0 ; 120] (o sea $NH/2 = 240/2$)

PARTE 4: RESULTADOS

Esta cuarta parte de la tesis consta de tres capítulos, el primero de los cuales se focaliza en la adecuación del método MOCOM-UA a los nuevos escenarios propuestos en este trabajo para los cuales previamente no ha sido utilizado. Los siguientes dos capítulos que complementan este bloque, constituyen aplicaciones prácticas de la metodología propuesta en el bloque anterior, que ponen a prueba la capacidad del método de optimización multi-objetivo para enriquecer la información disponible por el operador de un modelo de pronóstico hidrológico a los fines de reducir las incertidumbres asociadas a la toma de decisiones en tiempo real.

12. Adaptación del Método MOCOM-UA para mejorar su robustez

12.1. Problemática

En el capítulo relativo al planteo de la problemática general de esta tesis se argumentó respecto de la importancia de contar con una metodología binaria simulación/pronóstico, como enfoque de una visión más matizada y continua entre dos situaciones de pronóstico: i) en contexto bien instrumentado (escenario 1), donde la información a tiempo real puede ser plenamente puesta al servicio de la asimilación, y ii) poco instrumentado (escenario 2), donde el modelo está levemente restringido ya que carece de información de caudal a tiempo real, dejado por lo tanto libre de evolucionar en condiciones próximas a las de un modelo de simulación.

Dicha metodología se instrumentó a través de un método de optimización multi-objetivo aplicado sobre un modelo de pronóstico hidrológico, sometiendo la capacidad del método de encontrar juegos de valores de parámetros capaces de alcanzar a la vez buenos desempeños del modelo para ambos escenarios descritos.

Ahora bien, en el proceso de optimización se observó una aparente incapacidad por parte del método de encontrar de manera robusta en todos los experimentos de optimización el mismo frente de Pareto. En consecuencia, en este capítulo se puso foco principal sobre este aspecto, por lo que se llevó a cabo ciertas modificaciones en el algoritmo de búsqueda del método, orientadas a mejorar su eficacia para llegar sucesivamente al mismo frente de Pareto, o semejante.

Cabe aclarar que en este capítulo se entiende por "desempeño del modelo" a la capacidad del mismo de representar el hidrograma observado. Así, por ejemplo, se entiende que el desempeño del modelo es adecuado cuando los hidrogramas pronosticados por el mismo se aproximan aceptablemente al hidrograma observado.

12.2. Funciones objetivo y criterio de evaluación del ajuste

En este experimento, también la expresión matemática adoptada como FO para calibrar el modelo y también para evaluar su desempeño fue el coeficiente de Nash-Sutcliffe (NSE), cuya descripción puede verse en Nash y Sutcliffe (1970). La expresión de NSE fue presentada y desarrollada en el capítulo 7, y se transcribe a continuación, donde Q_{obs} y Q_{calc} , representan las ordenadas del hidrograma observado y calculado respectivamente. En consecuencia, NSE cuantifica el error de simulación respecto de la media del hidrograma observado. Nótese que se le resta a la unidad un cociente que representa la magnitud de un error, dado este cociente por la relación entre la varianza de la simulación respecto de la observación, y la varianza de la observación respecto de la media observada. En otras palabras, NSE es una medida del desempeño del modelo que puede tomar valores desde $-\infty$ a 1. Siendo el valor 1 un ajuste perfecto, degradándose la calidad del ajuste a medida que se reduce su valor.

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \quad \text{Ec. 12-1}$$

El trabajo se plantea en el contexto de los dos escenarios factibles durante el proceso operativo de pronóstico:

Escenario 1: el sistema de transmisión de datos de Q_{obs} a tiempo real funciona adecuadamente, por lo que el operador del modelo de pronóstico cuenta con la posibilidad de actualizar el estado del modelo en cada paso de tiempo a partir del valor conocido del caudal observado.

Escenario 2: el sistema es incapaz de enviar la información de caudal a tiempo real, por lo que el modelo funciona como un simple modelo de simulación, ya que no puede actualizar su estado. El operador debe pronosticar con el modelo en esa situación.

A partir de estos dos escenarios, se definen las FO de la siguiente manera: FO1: NSE para la calibración del modelo funcionando en condiciones del escenario 1 (sin Qobs a tiempo real). FO2: NSE para la calibración del modelo funcionando en condiciones del escenario 2 (con Qobs a tiempo real).

12.3. Optimización del modelo GR4P

A los fines de poner a prueba la optimización multi-objetivo como una herramienta de utilidad práctica en la operación de los modelos hidrológicos de crecidas, se llevaron a cabo algunas optimizaciones del modelo hidrológico GR4P en la cuenca del río Moros en Concarneau.

En la Figura 12-1 se puede ver una secuencia de 7 pruebas de optimización. Se destaca que prácticamente no hay coincidencia entre dos regiones de Pareto asociadas a distintas optimizaciones. En todos los casos el criterio de parada consistió en alcanzar la región de Pareto, o bien el número máximo de 20.000 iteraciones.

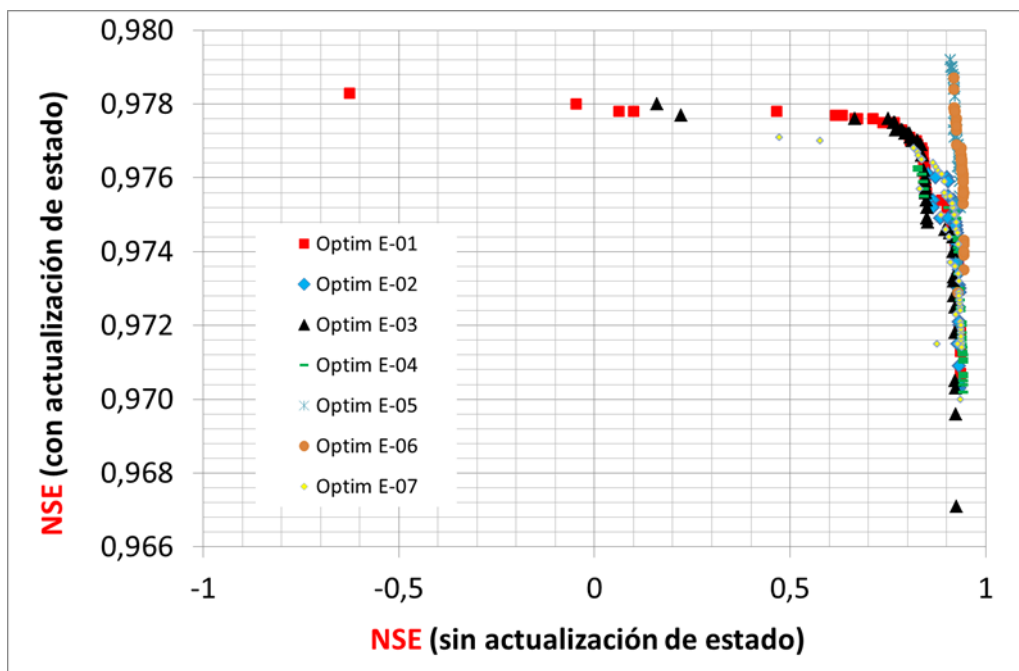


Figura 12-1: Pruebas de optimización con el MOCOM en su versión original.

En la Figura 12-2 se observa con mayor detalle que cada optimización condujo a un frente de Pareto distinto. Incluso, algunas optimizaciones no alcanzaron dicho Pareto ya que se destacan individuos dominados en la nube de puntos.

En efecto en las optimizaciones 2, 6 y 7 se cortó prematuramente el proceso de optimización ya que se alcanzó el número máximo de iteraciones que se dejó realizar al algoritmo de búsqueda (Tabla 12-1), quedando, por lo tanto, la población obtenida con algunos individuos todavía dominados. A pesar de eso, se observa que estas poblaciones optimizadas tienen todos individuos buenos a juzgar por el valor de NSE.

En Alonso et al. (2015) se muestra que en algunos procesos de optimización puede alcanzarse una población optimizada en la cual exista al menos un individuo capaz de satisfacer con aceptable desempeño del modelo ambos objetivos optimizados simultáneamente. Por esta razón, el frente de Pareto no sólo tiene importancia en sus extremos, donde se alcanza el mejor valor para cada criterio de ajuste, sino también en sus soluciones de compromiso, donde existen estos individuos “buenos” para ambos criterios al mismo tiempo.

En consecuencia, dada la incapacidad del método para llegar en todos los experimentos a un mismo (o muy semejante) frente de Pareto, subsisten los siguientes interrogantes: i) ¿cuál es la verdadera FO1 optimizada?, ii) ¿cuál es la verdadera FO2 optimizada?, y iii) ¿existe alguna solución de compromiso capaz de satisfacer simultáneamente ambos criterios con aceptable desempeño del modelo para ambas funciones objetivo?, y ¿cuál es dicha solución?

Tabla 12-1: Experimentos de Optimización, valores numéricos

Optimización	Cantidad de Iterac.	FO1	FO2
Opt. 01	13858	0,9357	0,9783
Opt. 02	20000	0,9338	0,9761
Opt. 03	6997	0,9250	0,9780
Opt. 04	2311	0,9425	0,9763
Opt. 05	9844	0,9350	0,9792
Opt. 06	20000	0,9447	0,9787
Opt. 07	20000	0,9386	0,9771

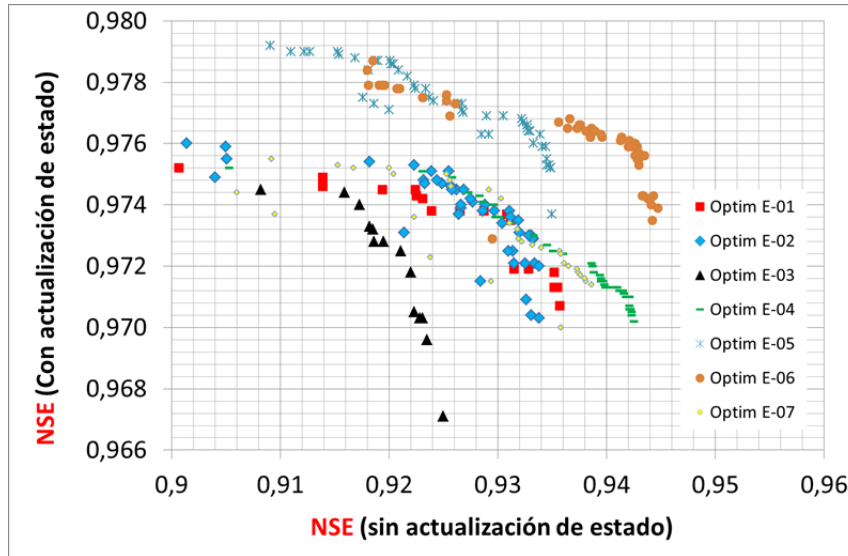


Figura 12-2: Pruebas de optimización con el MOCOM en su versión original. Detalle.

Si bien el método MOCOM-UA dispone de una estructura con cierto grado de aleatoriedad como fue descrito previamente, la cual enriquece el método, una sistemática incongruencia en las poblaciones optimizadas obtenidas aumenta la incertidumbre de los resultados. Además, fue detectado un tiempo excesivo en el desarrollo de cada optimización.

Por las razones expuestas, emerge como necesario mejorar la técnica de búsqueda y optimización procurando una metodología más robusta, capaz de alcanzar en las sucesivas prácticas de optimización regiones de Pareto semejantes entre sí.

12.4. Modificaciones al MOCOM-UA para mejorar la robustez

Atendiendo a la problemática previamente planteada, se decidió implementar conjuntamente una serie de medidas correctivas al método orientadas a mejorar su robustez, en términos de su capacidad para encontrar sucesivamente semejantes poblaciones optimizadas, y reducir su tiempo de optimización.

Dichas medidas correctivas implementadas fueron:

1. Población inicial mejorada
2. Adimensionalización de parámetros.
3. Diseño de una búsqueda local mejorada.
4. Cambiar la condición de parada.

12.4.1. Población inicial mejorada

En su versión original, el MOCOM-UA parte de una población inicial obtenida a partir de una generación aleatoria con distribución probabilística uniforme de puntos en el dominio de validez de los parámetros. Esa forma de establecer la población inicial es razonable cuando no se dispone información a-priori sobre el comportamiento de los parámetros.

Aprovechando la ventaja de disponer de una larga librería de valores de los parámetros del modelo GR4(J/H) calibrado para series de datos diarios y horarios a partir de una rica historia de aplicación del mismo por parte del Irstea, se propuso establecer como punto de partida del método una población inicial que tome individuos de la librería de parámetros del Irstea, del siguiente modo. En primer lugar, se dispuso de 480 juegos de valores de parámetros. A estos individuos se los sometió a la clasificación de Pareto y se eligieron los Nip mejores puntos como población inicial.

De este modo se consigue un doble propósito: i) sesgar la búsqueda hacia el dominio de mayor probabilidad de la existencia del óptimo, y ii) ahorrar el recurso computacional (menor tiempo) focalizando la búsqueda en la región apropiada.

12.4.2. Adimensionalización de parámetros

Dado que el rango numérico de validez de los parámetros es bien distinto entre sí (Perrin et al., 2007), resulta conveniente normalizarlos, a los fines de conseguir una exploración homogénea en todas las direcciones del dominio de búsqueda.

Se adoptaron los criterios de adimensionalización propuestos por el Irstea de los parámetros de calibración para el modelo aplicado sobre una serie de datos con paso de tiempo horario (Tabla 12-2) y diario (Tabla 12-3).

Tabla 12-2: Relaciones de adimensionalización de los parámetros en paso de tiempo horario

$XV(1) = e^{X(1)}$	$X(1) = \text{Ln}[XV(1)]$
$XV(2) = \frac{1}{24} \cdot 2.21^{3.5} \cdot \sinh[X(2)]$	$X(2) = \text{asinh} \left[\frac{24 \cdot XV(2)}{2.21^{3.5}} \right]$
$XV(3) = 2.21 \cdot e^{X(3)}$	$X(3) = \text{Ln} \left[\frac{XV(3)}{2.21} \right]$
$XV(4) = 24 \cdot [5 + 0.499 \cdot X(4)]$	$X(4) = \frac{1}{0.499} \cdot \left[\frac{XV(4)}{24} - 5 \right]$

Tabla 12-3: Relaciones de adimensionalización de los parámetros en paso de tiempo diario

$XV(1) = e^{X(1)}$	$X(1) = \text{Ln}[XV(1)]$
$XV(2) = \sinh[X(2)]$	$X(2) = \text{asinh}[XV(2)]$
$XV(3) = e^{X(3)}$	$X(3) = \text{Ln}[XV(3)]$
$XV(4) = 5 + 0.499 \cdot X(4)$	$X(4) = \frac{1}{0.499} \cdot [XV(4) - 5]$

Esta normalización tiene en cuenta que los parámetros tienen unidades y rangos de validez notoriamente diferentes. Así, por ejemplo, en calibración del modelo sobre series de tiempo con paso diario, el valor del parámetro X1 ronda entre los 100 y 1200 mm; X2 entre -5 à 3; X3 entre 20 y 300 mm; y X4 entre 1.1 y 2.9 días. Además,

representan distintos aspectos dentro de la estructura del modelo, y en consecuencia los hidrogramas generados tienen distintos grados de sensibilidad a variaciones en los diferentes parámetros del mismo. Un modo de reducir la dispersión en la sensibilidad de los distintos parámetros, es normalizar sus valores entre 0 y 1, a través de las expresiones previamente indicadas.

12.4.3. Búsqueda local mejorada

Como fue expresado previamente, en la versión original del método MOCOM-UA, se propone un punto de reflexión, y en caso de ser dominado, se adopta directamente el punto de contracción.

La propuesta de modificación llevada adelante en este trabajo consistió en las variantes que se indican a continuación:

- 1) Cuando se genera el punto de reflexión Rf_1 , puede ocurrir que caiga dentro del dominio de validez de los parámetros, en cuyo caso se pasa al paso siguiente. Pero si cae fuera del dominio de validez de los parámetros, se genera un nuevo punto de reflexión Rf_2 que se obtiene como el promedio espacial entre el punto de reflexión recientemente probado (Rf_1) y el baricentro "Sg" de los $n-1$ mejores puntos del simplex. Así sucesivamente, este paso se realiza Rf_{int} veces, al cabo de las cuales si nuevamente se obtiene un punto fuera del espacio de validez se pasa al paso "iii", pero si es válido se sigue en el paso "ii".
- 2) Una vez que el punto de reflexión cae dentro del dominio de validez de los parámetros, se evalúa en las FO1 y FO2. Si es un punto no dominado, reemplaza al peor punto del simplex. Pero si es dominado, se genera un nuevo punto de reflexión como el promedio entre el punto de reflexión y el baricentro referido, del mismo modo que se explicó en el paso i. Esta prueba se realiza hasta que se encuentre un punto no dominado que reemplace al

peor punto del simplex, o bien hasta que se haya probado N_{ajr} veces, al cabo de las cuales se pasa al paso "iii".

- 3) Se genera un punto de contracción "Ct1". Si el mismo es no dominado, reemplaza al peor punto del simplex (punto "c"), pero si es dominado, se vuelve a generar un punto de contracción "Ct2" como el promedio entre este punto "Ct1" y el baricentro "g" referido. Si nuevamente es dominado se vuelve a generar un punto promedio hasta que sea no dominado, en cuyo caso este nuevo punto reemplaza al peor punto, o bien hasta que se alcancen las N_{ajc} pruebas, en cuyo caso automáticamente el último punto de contracción probado reemplaza al peor punto, sea o no dominado.

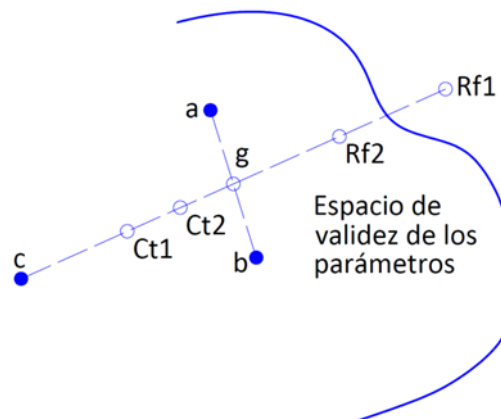


Figura 12-3: Esquema que ilustra las pruebas en Reflexión (Rf1, Rf2,...) y en Contracción (Ct1, Ct2,...) en el método de búsqueda local del MOCOM-UA.

En síntesis, la modificación que se produjo al método de búsqueda local consiste básicamente en: i) probar R_{fint} puntos de reflexión para caer dentro del espacio de validez de los parámetros, ii) probar N_{ajr} puntos de reflexión (en vez de una vez) en búsqueda de un punto no dominado, y iii) probar N_{ajc} puntos de contracción (en vez de una vez) en búsqueda de un punto no dominado.

12.4.4. Condición de parada

La versión original del MOCOM-UA establece como condición de parada que la totalidad de los puntos de la población sean no-dominados. Las pruebas efectuadas nos demostraron que en múltiples casos el método se aproxima rápidamente a la “zona óptima” pero tarda un gran número de iteraciones en ir logrando el criterio de no-dominancia de la totalidad de sus puntos. En la Figura 12-4, por ejemplo, se muestra que el método llega rápidamente a posicionar los 100 puntos de la población en la zona óptima de los parámetros para optimizar una serie de 15.000 datos con paso de tiempo diario de la cuenca L'Ilh en Altkirch, Francia. Pero sin embargo, en la Figura 12-5 se pone en evidencia que la población optimizada solamente dispone de 8 puntos no dominados. Llevar esta población optimizada a la totalidad de puntos no-dominados implicaría un costo computacional altísimo y tal vez innecesario.

Es decir, el objetivo de lograr una región de Pareto pura de puntos no-dominados es muy costoso en términos computacionales. Además, este criterio no es en sí mismo garantía de haber arribado a buenos valores de función objetivo, lo cual quedó demostrado en varias pruebas en las cuales se obtuvieron prematuramente regiones de Pareto con muy malas funciones objetivo.

Por ello, se consideró necesario exigir mayor calidad de FO, por una parte, y por otra mayor celeridad en la conclusión del proceso de optimización, para lo cual se proponen los siguientes criterios de parada:

- I. Entre una iteración evolutiva y la siguiente ambas FO no deben tener una variación superior a cierta tolerancia pre-establecida.
- II. El número total de puntos no-dominados de la población optimizada debe ser igual o superior a un porcentaje pre-establecido respecto del total de los puntos de la población.

Los criterios antes mencionados deben cumplirse simultáneamente para que el método se detenga y asuma la población obtenida como población optimizada. Por último, también se establece un número máximo de pasos iterativos que objetivan detener el algoritmo si no es posible cumplir hasta entonces con los criterios anteriores.

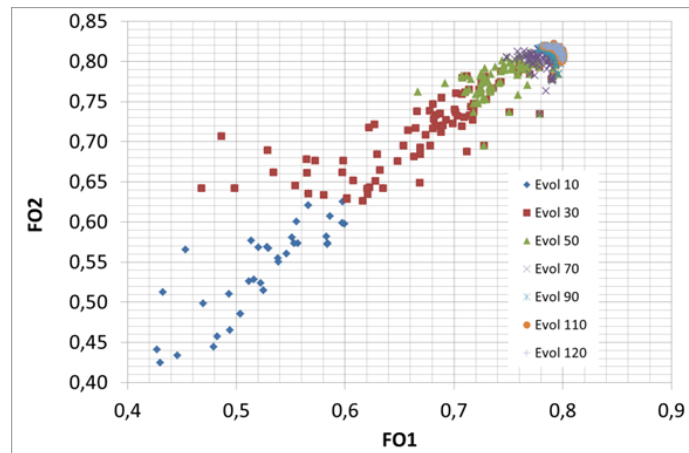


Figura 12-4: Ejemplo de optimización realizada con la cuenca L'III à Altkirch.

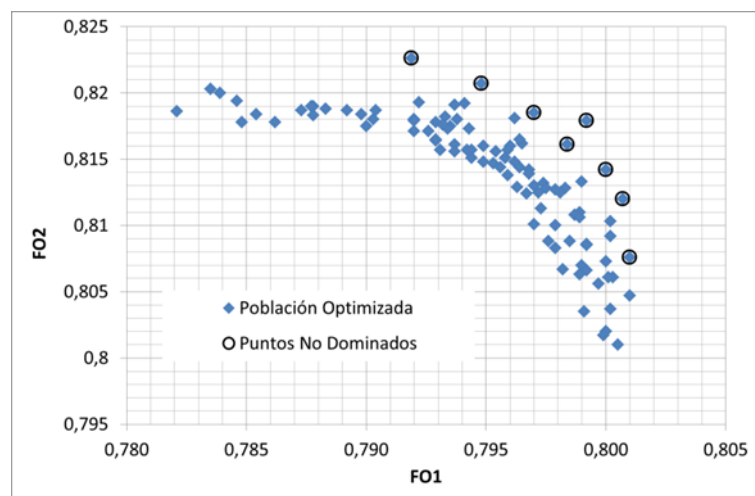


Figura 12-5: Población "optimizada" y puntos no dominados.

12.5. Resultados y discusiones

Una vez adaptado el método de optimización sobre la base de las modificaciones descritas previamente, se realizaron nuevas optimizaciones, estableciendo esta vez sólo 5000 iteraciones admisibles como límite máximo de ciclos evolutivos.

En la Figura 12-6 se observa que el método llegó robustamente en las cuatro pruebas realizadas a un mismo frente de Pareto. A los fines de permitir la comparación con las pruebas realizadas previamente con el método MOCOM en su versión original, se adoptó la misma escala que la Figura 12-1.

En la Figura 12-7 se pone de relevancia con mayor detalle este aspecto, realizando un ajuste de la escala de los ejes que permita su comparación con la Figura 12-2.

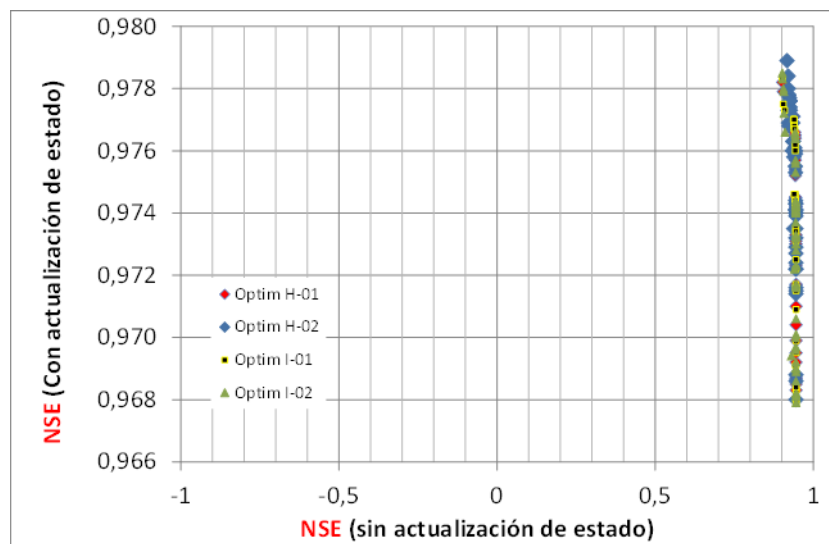


Figura 12-6: Pruebas de optimización con el MOCOM en su versión modificada.

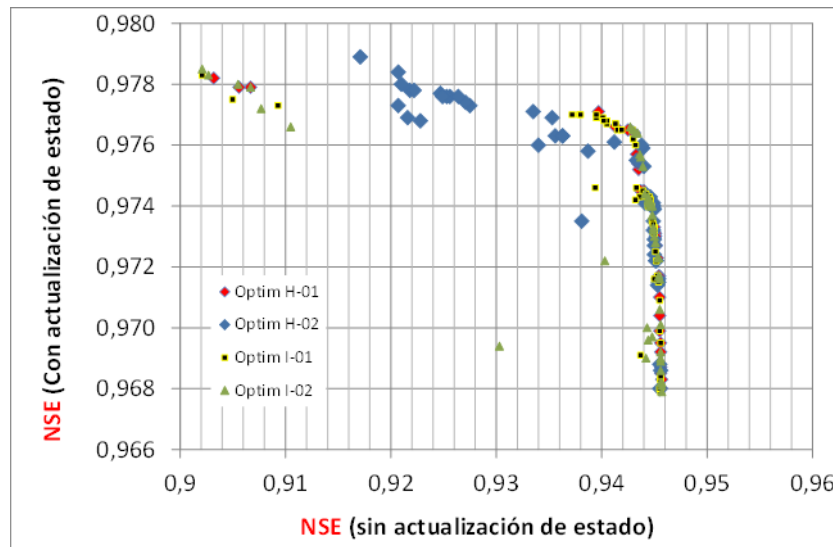


Figura 12-7: Pruebas de optimización con el MOCOM en su versión original (detalle).

Por su parte, en la Tabla 12-4 se pone de manifiesto numéricamente que el método MOCOM modificado ha alcanzado prácticamente el mismo valor de la FO1 y FO2 en todas las pruebas de optimización. También se observa que sólo en una prueba se terminó el proceso de optimización cuando se alcanzó el máximo número de 5000 iteraciones. Sin embargo, en las otras tres pruebas el óptimo fue alcanzado en menor número de iteraciones que todas las pruebas realizadas con el MOCOM original según puede verse en la Tabla 1.

Tabla 12-4: Experimentos de Optimización, valores numéricos

Optimización	Cantidad de Iterac.	FO1	FO2
Opt. H-01	3926	0,9457	0,9782
Opt. H-02	5000	0,9456	0,9789
Opt. I-01	4435	0,9456	0,9783
Opt. I-02	2175	0,9457	0,9785

12.6. Conclusiones

Si bien el método MOCOM-UA tiene una probada trayectoria en la bibliografía, no ha sido capaz de encontrar de manera robusta y eficiente el frente de Pareto para

un problema de optimización múltiple que involucre como objetivos el desempeño de un modelo de pronóstico con y sin actualización de estado.

En este trabajo fueron planteadas algunas modificaciones a la estructura original del método, que objetivaron llegar con mayor precisión y constancia a la misma zona óptima, configurando frentes de Pareto semejantes y en el menor número de iteraciones posibles.

Los resultados demostraron que, si bien no se alcanzaron desempeños del modelo altamente superiores a las obtenidas con el método MOCOM-UA original, pudo notarse una leve mejora y fundamentalmente una mejora de la robustez en la estimación del frente de Pareto, así como en el consumo computacional manifestado en el menor número de iteraciones requeridas para esto.

Se considera que este aporte al campo de la optimización de modelos matemáticos, contribuye levemente al desempeño del modelo en su proceso de calibración, pero aporta elementos sólidos a la reducción de la incertidumbre en la estimación del juego de valores óptimos para la simulación y pronóstico de modelos hidrológicos.

13. Evaluación cuantitativa de la mejora de la performance del Q a tiempo real.

13.1. Problemática

Como se ha expresado previamente, en un modelo hidrológico de pronóstico, para maximizar la calidad del pronóstico en el horizonte elegido, es deseable partir del mejor estado posible en el instante de la previsión, y es por lo tanto conveniente “re-ajustar” el modelo, si es posible, sobre la base de los últimos caudales observados a lo largo de algunos pasos de tiempo precedentes al instante en el cual debe ser emitido el pronóstico.

A causa de esta etapa, denominada de asimilación (que corrige el estado del modelo), un buen modelo de pronóstico no es necesariamente un buen modelo de simulación, y viceversa. Debido al elevado peso asignado a las últimas observaciones, la estructura de un modelo hidrológico de pronóstico puede típicamente ser más simple que aquella de un modelo de simulación resultando potencialmente bastante desconcertante tanto por los desarrolladores como por los usuarios de estos modelos (Berthet, 2010).

Sin embargo, desde el punto de vista del desarrollo de los modelos hidrológicos, la confrontación entre estos dos objetivos (simulación y pronóstico) no es necesariamente tan clara como parece. En efecto, como lo recuerda (Duband, 1999), "para una utilización operacional en pronóstico, se necesita de una modelización específica notoriamente distinta a las necesidades de la hidrología aplicada orientada a la comprensión y a la reconstrucción del ciclo del agua: simular no es pronosticar. De todos modos, el progreso en el conocimiento puede contribuir tanto a uno como al otro objetivo". En efecto, para el caso de un modelo de pronóstico hidrológico, altamente dependiente de la actualización de estado a cada paso de tiempo, si la cadena de transmisión de datos hidrometeorológicos (red pluviométrica

e hidrométrica) que alimenta el sistema de pronóstico falla, las performances reales pueden volverse muy por debajo de aquella que uno esperaría en un caso ideal.

En particular, si la información a tiempo real de los últimos datos de caudales observados son perdidos (daño en la estación hidrométrica, en la red de comunicación, etc.) y no puede ser más asimilado, el modelo hidrológico se vuelve en una situación típica de simulación, ejercicio para el cual otro modelo (o el mismo modelo, pero parametrizado distinto) podría dar resultados más confiables.

El especial interés de los datos de radar para el pronóstico hidrológico es debido a la posibilidad de contar con información más fina que la proveniente de la red telemétrica de sensores pluviográficos sobre la distribución espacial de las precipitaciones, así como sobre las velocidades y direcciones de desplazamiento de las células de lluvia.

Esto es particularmente notorio cuando el pronóstico y la alerta temprana deben ser hechos para cuencas hidrográficas de área entre pequeña y mediana, con tiempos de respuesta inferiores a 6-8 hs, máxime para zonas fuertemente urbanizadas. Sin embargo, el radar debe ser calibrado empleando datos pluviométricos locales, siendo por lo tanto dependiente de la calidad de la información pluviométrica existente en los sitios de aplicación del pronóstico.

El estudio de incertidumbre asociada al acople radar-modelización hidrológica y la evaluación de los riesgos de falsa alarma pueden aportar una contribución importante para la concepción y la gestión de los sistemas de alerta temprana de crecidas.

En este capítulo, que forma parte del proyecto HYRADIER, una cooperación científica entre Argentina, Brasil y Francia, financiado por el programa STIC-AmSud, se propone una metodología binaria simulación/pronóstico, como enfoque de una visión más matizada y continua entre las situaciones de pronóstico en contexto bien

instrumentado, donde la información a tiempo real puede ser plenamente puesta al servicio de la asimilación, y poco instrumentado, donde el modelo está levemente restringido ya que carece de información de caudal a tiempo real, dejado por lo tanto libre de evolucionar en condiciones próximas a las de un modelo de simulación.

En particular, se pretende saber si es siempre preferible calibrar un modelo de pronóstico utilizando específicamente los criterios (parámetros) de pronóstico y en modo pronóstico, o si puede ser interesante dar un peso a los criterios más cercanos a un enfoque de simulación. El estudio de la determinación de la combinación óptima, que dependerá eventualmente de los contextos fisiográficos y meteorológicos, fue efectuado en este trabajo con series de datos sobre la cuenca del río Moros en su desembocadura en Concarneau, región de la Bretaña al noroeste de Francia.

13.2. Funciones objetivo y criterio de evaluación del ajuste

La expresión matemática adoptada como FO para calibrar el modelo y también para evaluar su performance fue el coeficiente KGE (Gupta, 2009) descrito en el capítulo de esta tesis respectivo a los criterios de ajuste y evaluación del modelo.

El trabajo se plantea en el contexto de los dos escenarios factibles durante el proceso operativo de pronóstico:

Escenario 1: el sistema de transmisión de datos de Qobs a tiempo real funciona adecuadamente, por lo que el operador del modelo de pronóstico cuenta con la posibilidad de actualizar el estado del modelo en cada paso de tiempo a partir del valor conocido del caudal observado.

Escenario 2: el sistema es incapaz de enviar la información de caudal a tiempo real, por lo que el modelo funciona como un simple modelo de simulación, ya que no

puede actualizar su estado. El operador debe pronosticar con el modelo en esa situación.

A partir de estos dos escenarios, se definen las FO de la siguiente manera: FO1: KGE para la calibración del modelo funcionando en condiciones del escenario 1 (sin Qobs a tiempo real). FO2: KGE para la calibración del modelo funcionando en condiciones del escenario 2 (con Qobs a tiempo real).

13.3. Optimización del modelo GR4P

Se realizó la optimización del modelo hidrológico GR4P para horizontes de previsión $L=1$, $L=6$, y $L=24$ horas, como se muestran en la Figura 13-1, tomando el primer periodo de datos que fue definido previamente. A los fines de provocar un ahorro de consumo computacional, se permitió al método de optimización concluir el proceso antes de alcanzar el frente de Pareto final, a condición de obtener al menos el 85% de los 60 individuos de la población no-dominados, y una tolerancia al cambio del valor de FO de 0,001 entre dos generaciones poblacionales. Por esta razón, se observa en la gráfica que no todos los individuos de la población optimizada son no-dominados.

Para los tres horizontes de pronóstico (L) se llega al mismo "mejor valor" de FO1. En cambio, el mejor valor de FO2, se degrada cuanto mayor es el valor de L . Por otra parte, para $L = 1h$, el "mejor valor de FO2" es mayor que el "mejor valor de FO1", por lo que la actualización de estado mejora la performance; para $L = 6h$, la performance es indiferente a la actualización de estado; y para $L = 24h$, parece que la actualización de estado desmejora la performance. Es decir, hay una notoria tendencia de desmejorar la performance debido a la actualización de estado, a medida que aumenta el horizonte L . Esto puede ser explicado porque el procedimiento de actualización del modelo, obliga a que a cada paso de tiempo el valor pronosticado

tome exactamente el valor observado en cada Dt, quitándole de este modo grados de libertad en la selección del juego óptimo de parámetros.

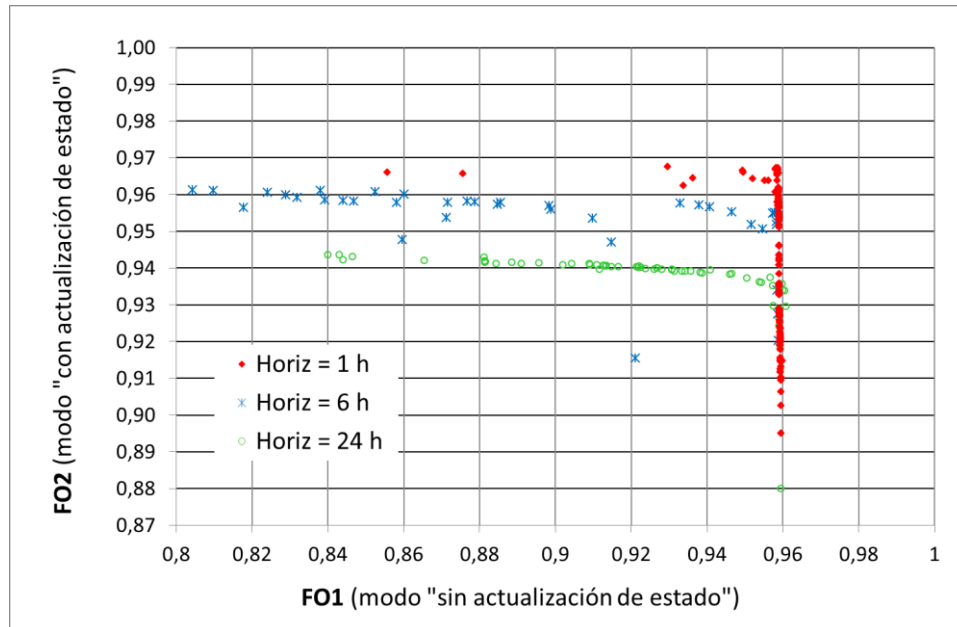


Figura 13-1: Optimización del Modelo GR4P para distintos horizontes de pronóstico.

Con la población optimizada, se realizó el control de la calidad de la calibración, tomando el segundo periodo de datos de la cuenca definido previamente. En la Figura 13-2 se muestran 30 individuos representativos de la población optimizada, donde se observa el valor del coeficiente KGE para cada uno en la fase de calibración (primer periodo de datos), y control (segundo periodo), tanto para el modelo "sin actualización" (escenario 1), y "con actualización" (escenario 2), para un horizonte de pronóstico (L) de 1 hora.

Los individuos son ordenados de mejor a peor valor de la FO2 en calibración. Esto quiere decir que el individuo 1 tiene el mejor valor de la FO2 en calibración, y el individuo "30", el peor valor. Se observa que, si bien la población no arribó al frente de Pareto final, presenta individuos con buena performance tanto en calibración como en control, lo cual demuestra que hay ciertas combinaciones de parámetros robustas ya que son capaces de adaptarse con buena performance a cualquiera de

los escenarios. Se puede ver también que los valores de las FO1 mejoran al mismo tiempo que los valores de la FO2 se degradan. Finalmente, se puede ver que los puntos que tienen valores de la FO1 buenos en calibración, tienen también buenos valores en control.

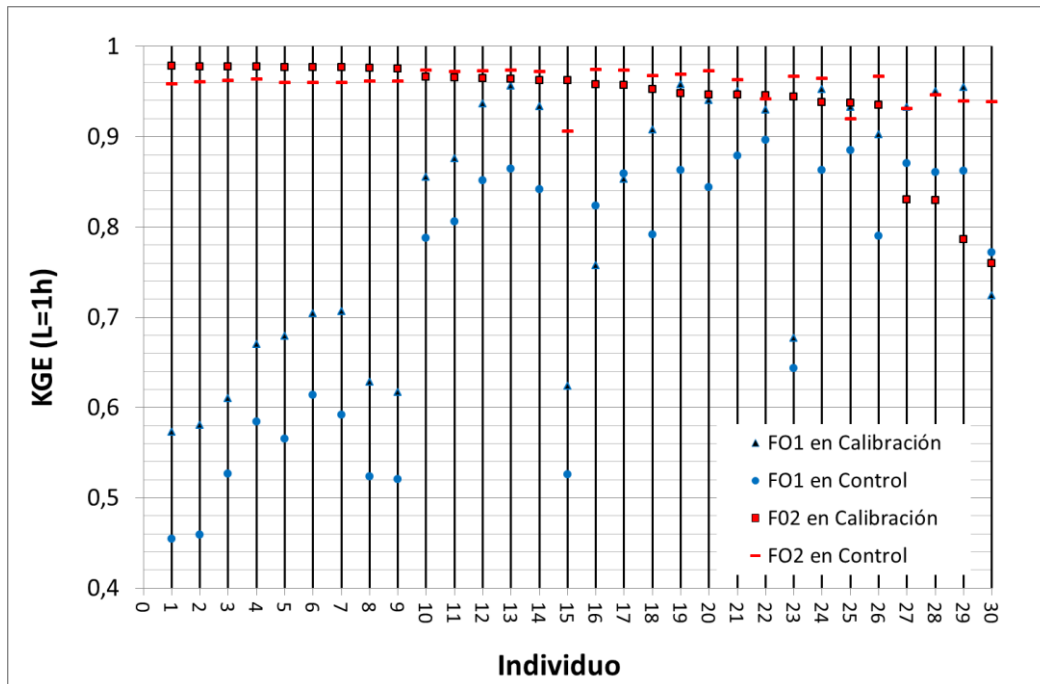


Figura 13-2: Performance de la población optimizada para los distintos escenarios.

En la Figura 13-3 se muestran en puntos la performance de la población óptima calibrada en el primer periodo de la serie de datos, y en cruces las performances de las simulaciones de los mismos individuos en el segundo periodo de la serie (control).

Como se reflexionó previamente, se puede identificar puntos buenos para ambos periodos (punto 22, por ejemplo). Realizando un cambio de escala de los ejes del gráfico, se puede mirar con mejor detalle en la Figura 13-2, Figura 13-3, y en la Figura 13-4 la posición del individuo 22 y sus valores del coeficiente de KGE para cada escenario explicado.

Si bien el punto "22" tiene buen valor para las FO1 y FO2, hay puntos que pueden ser mejores en la FO1 en calibración y en control, y otros puntos que son mejores en

FO2 también en calibración y en control. Por ejemplo, el punto "4" tiene mejor valor que el punto "22" tanto en calibración como en control para la FO2. Pero en FO1 es inferior.

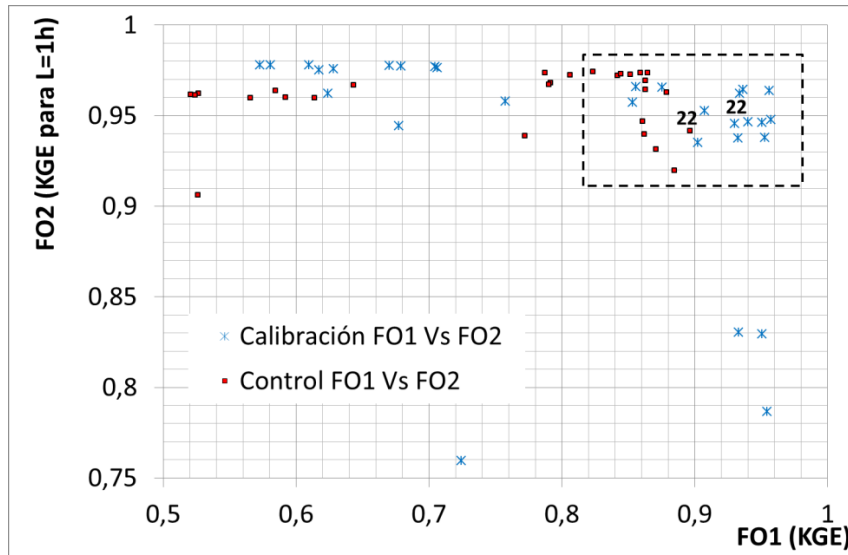


Figura 13-3: Población optimizada para L=1h.

Ahora bien, la mejora en FO2 que tiene el punto 4 respecto del punto 22, justifica abandonar la ventaja de contar con un único juego de valores de parámetros que interprete al modelo. En efecto, al adoptar dos soluciones buenas (una para FO1 y otra para FO2), se requiere en modo operativo dos corridas en paralelo del modelo, pero calibrado con dos juegos de parámetros distintos.

En la Tabla 13-1 se muestran los valores de los parámetros correspondientes al punto 4, y al punto 22, en comparación con los valores de los parámetros optimizados por el IRSTEA para la cuenca en estudio.

Esto quiere decir que los valores del punto 4 no son buenos más que para L=1h, pero el punto 22 es más robusto, como se muestra en la Figura 13-5.

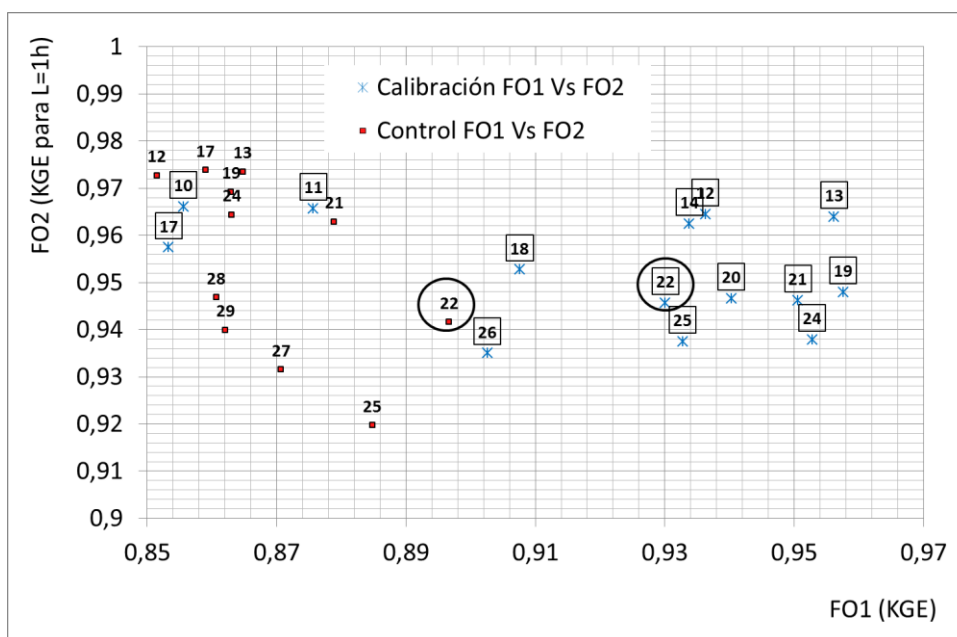


Figura 13-4: Detalle de la población optimizada para L=1h.

Se puede observar que el punto 22 está más próximo a los valores propuestos por el IRSTEA (que tiene una vasta experiencia en la operación del modelo de pronóstico en la cuenca en estudio), y se puede ver también que la performance del punto 4 se degrada rápidamente al mismo tiempo que el horizonte de pronóstico se prolonga.

Tabla 13-1: Performance del modelo para diversos escenarios

Posición	Parámetros				Calibración		Control				
	X1	X2	X3	X4	KGE sin actualiz	KGE con act. (1 h)	KGE sin act.	KGE con act.	KGE con act.	KGE con act.	KGE con act.
								L (h)	L (h)	L (h)	L (h)
								1	6	12	24
4	882,3	-0,903	41,2	6,017	0,6701	0,9777	0,585	0,964	0,652	0,424	0,179
22	482,1	-0,698	444,5	5,909	0,9301	0,9457	0,897	0,942	0,932	0,890	0,802
IRSTEA 01	417,6	-1,375	561,1	7,857							
IRSTEA 02	419,7	-0,762	498,0	7,751							

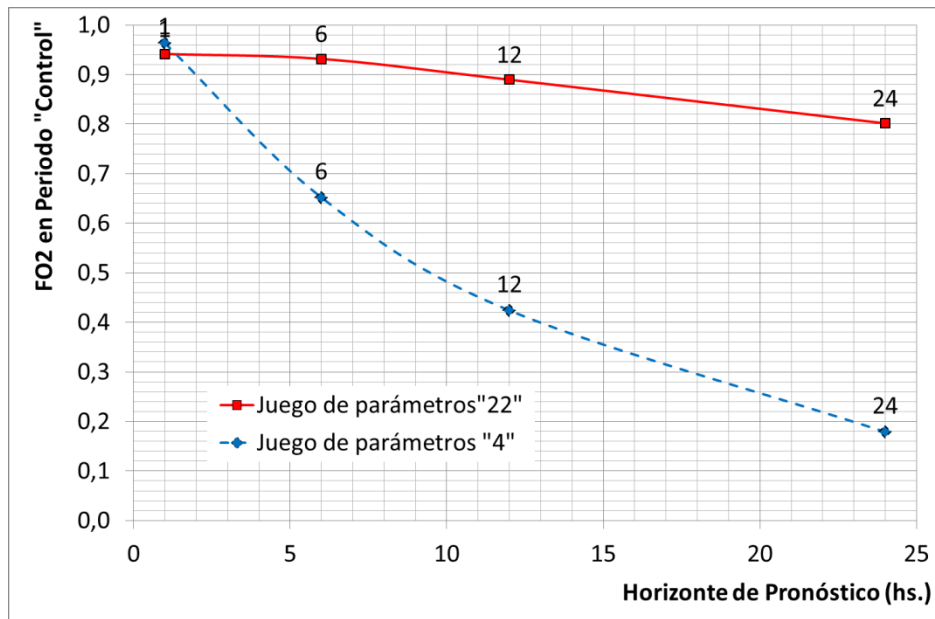


Figura 13-5: Sensibilidad de la performance (KGE) frente a variaciones en el horizonte pronosticado (L).

13.4. Síntesis de los resultados y discusiones

Teniendo en cuenta los resultados presentados, desde el punto de vista operacional se puede formular la siguiente pregunta: frente a la posibilidad de quedarse (en cualquier momento) sin los caudales observados durante el pronóstico (es decir, aquellos valores requeridos para realizar la actualización del estado del modelo), ¿cuál de los dos criterios siguientes será el más adecuado para adoptar?

Opción 1: Elegir de la población optimizada, el juego de parámetros que parece mejor, porque es bastante bueno para simular la salida del modelo en modo “con” y “sin” la actualización (por ejemplo, el punto 22) y simular en cualquier escenario.

Opción 2: Elegir de la población optimizada el juego de parámetros con mejor performance en pronóstico (por ejemplo, el punto 1), y frente a la eventual carencia de la información de caudal observado, cambiar al modo “sin la actualización”, donde el modelo corre con el mejor juego de parámetros “sin la actualización” (por ejemplo el punto 25).

Si existe algún punto capaz de representar todos los escenarios posibles con una performance aceptable, parece razonable adoptarlo como solución única (Opción 1), pero si la mejora de la performance en modo “sin la actualización”, significa la degradación de la performance en modo “con la actualización” (o a la inversa), resulta de gran valor la disposición de las soluciones óptimas extremas (mejor FO1 y mejor FO2).

En consecuencia, para analizar cuál opción elegir en operación, eso dependerá en gran medida de la cuenca hidrográfica, ya que cada cuenca tiene su propia población óptima particular. En efecto, de la Tabla 13-1 se puede extraer conclusiones únicamente para esa cuenca.

13.5. Conclusiones

El modelo de pronóstico hidrológico GR4P es una poderosa herramienta para la estimación del caudal con un determinado horizonte de previsión. Sin embargo, la exactitud de la estimación es altamente dependiente de la disponibilidad a tiempo real del caudal observado, ya que le permite actualizar el estado del modelo. La carencia, durante la fase operativa, del caudal observado a tiempo real, produce que el modelo pase a funcionar como un simple modelo de simulación. El empleo de la optimización multi-objetivo para la calibración de la cuenca “Moros en Concarneau” entre ambos escenarios, permitió alcanzar un conjunto de soluciones con distintas performances en ambas FO.

Se observó que hay una notoria tendencia de desmejorar la performance debido a la actualización de estado, a medida que aumenta el horizonte L. Esto puede ser explicado porque el procedimiento de actualización del modelo, obliga a que a cada paso de tiempo el valor pronosticado tome exactamente el valor observado en cada Dt, quitándole de este modo grados de libertad en la selección del juego óptimo de parámetros.

Se puso en evidencia que la bondad de un juego de parámetros no sólo depende de la performance del modelo frente al horizonte de pronóstico de interés, sino también a la sensibilidad de la performance a variaciones en el tiempo de horizonte establecido, a medida que se aleja del horizonte con el cual fue calibrado.

Finalmente, se propone como herramienta operativa la optimización multi-objetivo de un modelo, a los fines de contar con mayor información de base frente a la decisión de adoptar uno o dos juegos de valores de parámetros capaces de adaptarse a los dos escenarios posibles durante operación: disponibilidad o no disponibilidad de los valores de cual observado a tiempo real.

14. Desempeño del modelo frente a distintos horizontes de pronóstico.

14.1. Problemática

La robustez de un modelo hidrológico de pronóstico está dada por la versatilidad que manifiesta el mismo en adaptarse a diversas condiciones de pronóstico manteniendo considerablemente la calidad en los resultados.

A los fines de evaluar la robustez del modelo GRP, se pone a prueba la capacidad del mismo de adaptarse a múltiples escenarios frecuentes durante la operación de pronóstico. En el capítulo precedente se estudió la capacidad del modelo de responder con aceptable precisión a escenarios imprevistos asociadas a la pérdida o carencia de la información de actualización del estado del modelo a tiempo real.

En este capítulo se aborda un escenario de pronóstico en un contexto bien instrumentado en el cual se dispone de la información de actualización a tiempo real, pero se plantea la problemática de someter al modelo al ejercicio de adaptarse a diversos horizontes de pronóstico.

En efecto, habitualmente los destinatarios del pronóstico son organismos estatales que se nutren de esta información para diseñar un plan de emergencias capaz de ser puesto en marcha de manera inmediata frente a una posible próxima inundación. Dicho plan de emergencia comúnmente requiere conocer en distintos horizontes de tiempo futuro el valor estimado de los caudales (transformados en niveles) en ciertas secciones de interés del curso de agua.

Los trabajos abordados en este capítulo fueron ordenados en una secuencia de experimentos cuyos resultados ofrecen al operador del modelo algunos elementos de contraste que permiten tomar ciertas decisiones respecto de la incertidumbre asociada a la selección de diversos horizontes de pronóstico.

14.2. Funciones objetivo y criterio de evaluación del ajuste

Se adoptó el estimador KGE (Gupta et al., 2009) como FO para calibrar el modelo, así como para evaluar la performance del ajuste para ambos objetivos a optimizar. Cabe remarcar que los dos objetivos son precisamente el valor de la FO para cada horizonte de pronóstico respectivamente.

14.3. Optimización del modelo GR4P

A los fines de comenzar con los experimentos referidos, se adoptó la cuenca ya descrita "Le Moros", en la sección de Concarneau, con una superficie de 22,97 km². La serie empleada, con paso de tiempo horario, corresponde al tramo desde el 01/08/2009 al 19/05/2010 (7000 hs).

Por otra parte, el horizonte de pronóstico definido para la FO1 (es decir, Hp1) fue de 6hs; mientras que el Hp2 corresponde a 12hs. Se fijó un tramo de inicialización del modelo de 1000 hs.

Finalmente, con relación a los criterios de optimización, se estableció una población de 30 individuos a optimizar, obligando al 85% de la misma a alcanzar la condición de "no dominado", y aplicando para ambas FO una tolerancia de fuga de 0,0005. Esto significa que si entre una iteración de subdivisión de la población en complejos, y la siguiente, se verifica que la FO no modifica (es decir, no mejora) su valor en más del 0,05%, se considera que esa FO ha alcanzado el óptimo. Por lo tanto, cuando ambas FO cumplen con esa condición, y además el 85% de la población representan individuos no dominados, el método sale del proceso de optimización, y se toma la población por optimizada.

El proceso de optimización se desarrolló de acuerdo a la secuencia que muestra el gráfico de la Figura 14-1. Allí se observa que la población original tiene una tendencia

a satisfacer simultáneamente ambas FO, poniendo en evidencia que una mejora en la performance del pronóstico para un horizonte de 6hs equivale a una mejora en la performance para un horizonte de 12hs.

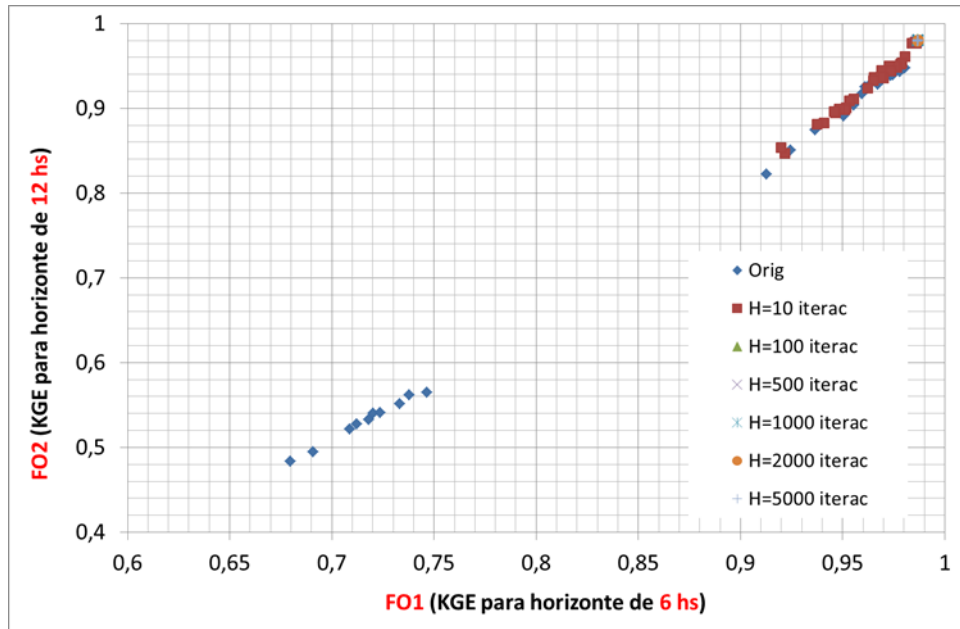


Figura 14-1: "Región de Pareto" entre distintos Hp (horizontes de previsión), 6hs Vs 12hs..

También se observa que en las primeras 10 iteraciones la población migra hacia una zona con mejor performance en conjunto, pero se conserva la relación entre los individuos, los que se ubican siguiendo una línea. De allí se deduce que un individuo con mejor FO1, también tiene mejor FO2.

Esta evidencia parece indicar que **para los casos ensayados sólo sería suficiente la optimización mono-objetivo, tomando como criterio reducir el error en un horizonte de previsión**, ya que está garantizada la reducción también en el otro horizonte.

La Figura 14-2 es idéntica a la Figura 14-1, a la cual se ha cambiado la escala de los ejes para mejor interpretar los resultados. Se observa que continúa la tendencia previamente referida, y también se hace notar que a medida que aumenta el número de iteraciones, la población se va concentrando en el extremo superior derecho del

gráfico, es decir, en la zona de mejor performance para ambas FO. **Se destaca que en las primeras iteraciones la población se aproxima rápidamente al óptimo**, y en el resto de las iteraciones, va progresando la performance en pequeñas mejoras diferenciales.

En efecto, en la Figura 14-3 se observa que el proceso de optimización concentra, en la iteración número 100, los puntos hacia un sector muy reducido. Sin embargo, haciendo un zoom en este sector, se observa en la Figura 14-4 que ya no se conserva dicha relación de performance según la cual cuando mejora la performance en FO1 también lo hace en FO2, es decir que pueden haber puntos que son mejores que otro en la FO1, y peores en la FO2, y viceversa.

Se observa en una primera vista que, **para diferencias grandes de los valores de los parámetros, una mejora en la performance del pronóstico en un Hp trae aparejado una mejora en los otros, pero una vez obtenido un juego de valores de parámetros suficientemente optimizados, la mejora en la performance de un Hp, puede arrastrar la desmejora en el otro**, como es habitual en un proceso multi-objetivo de optimización.

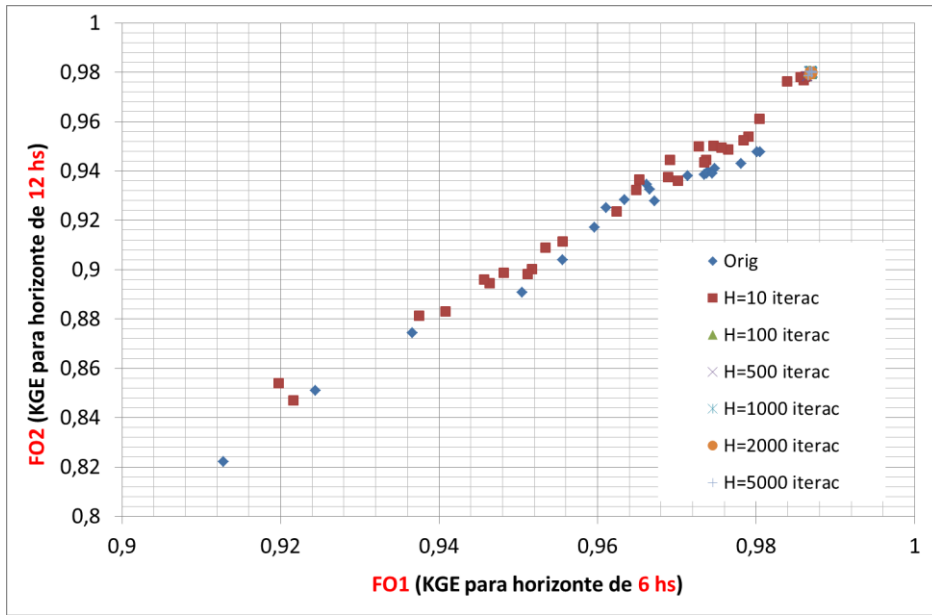


Figura 14-2: Ajuste de escala de los ejes de la Figura 14-1 para mejor interpretación.

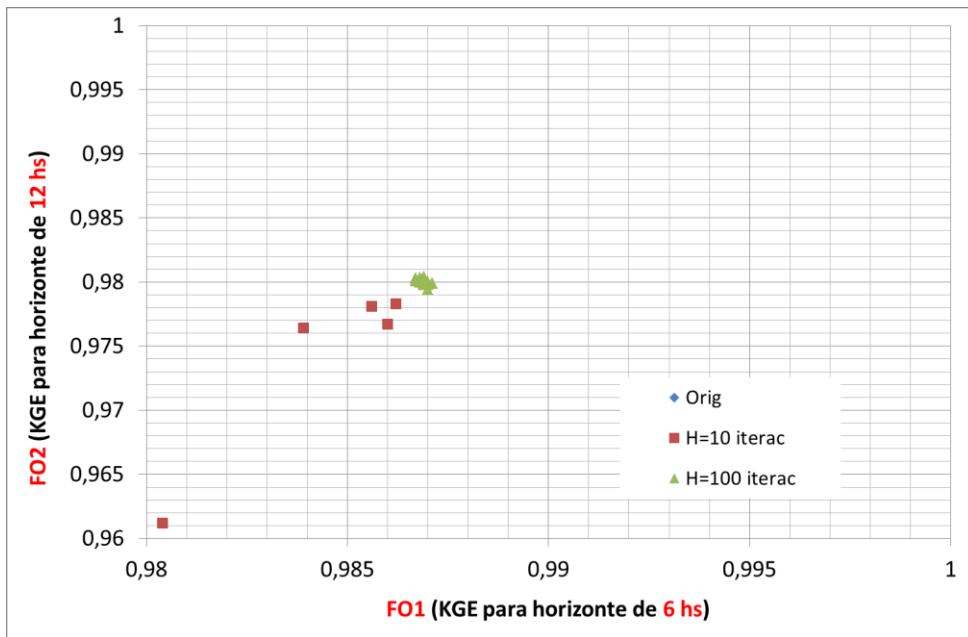


Figura 14-3: Concentración de puntos en la región optimizada.

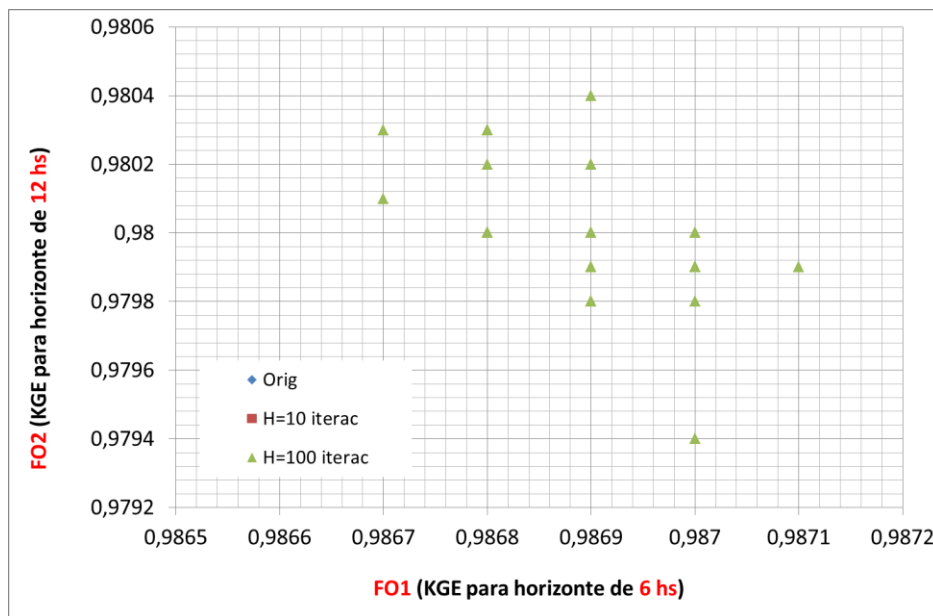


Figura 14-4: Zoom en el sector de concentración de puntos de la iteración 100 de la Figura 14-3.

Resulta interesante analizar si la tendencia referida (de satisfacer simultáneamente las FO asociadas a dos distintos H_p) se manifiesta de modo espontáneo en una población cualquiera. Para esto, se tomó la población original, que fue concebida a partir de la base de datos de parámetros del Irstea, y se calculó el criterio KGE para cada individuo para los siguientes cuatro horizontes de pronóstico: 1, 6, 12 y 24 hs. Luego se graficó en la Figura 14-5 la performance del modelo para 1h Vs la performance del modelo para 6, 12 y 24 hs.

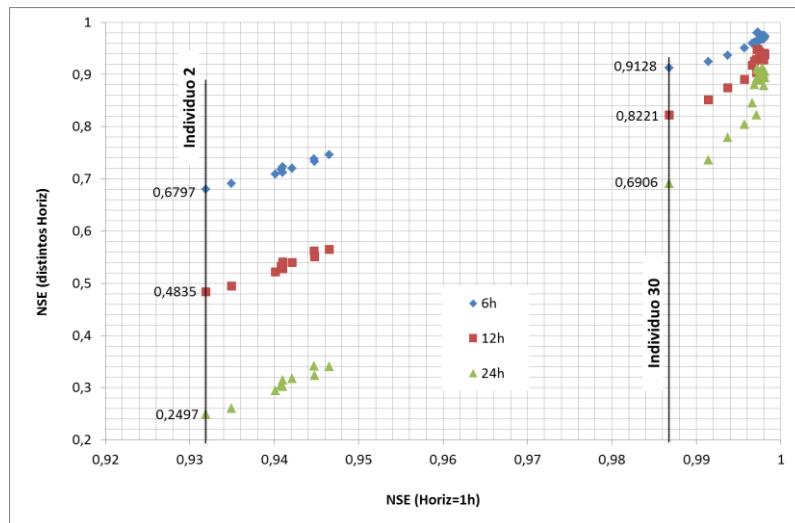


Figura 14-5: Relación entre las performances de todos los individuos de la población original para distintos Hp.

Resulta evidente que los individuos de mejor performance para un Hp de 1h, se corresponden con mejores performances en horizontes de pronóstico de 6, 12 y 24hs. Sirva de ejemplo cómo el individuo 30 supera al individuo 2 en la performance de todos sus horizontes.

En ese mismo gráfico se destaca también que a medida que los individuos tienen mejores performances los valores de estas se vuelven más cercanas entre sí. Así por ejemplo el individuo 30 tiene las performances de todos sus Hp más próximos en valor que el individuo 2.

Si hacemos un zoom en la zona superior derecha del gráfico observaremos en la Figura 14-6 que la relación creciente de todas las performances en conjunto se comienza a hacer difusa aproximadamente a partir del individuo 29 en adelante. En efecto, el individuo 29 supera al 21 en todas las performances, pero de allí en adelante no se puede establecer con facilidad un patrón de respuesta.

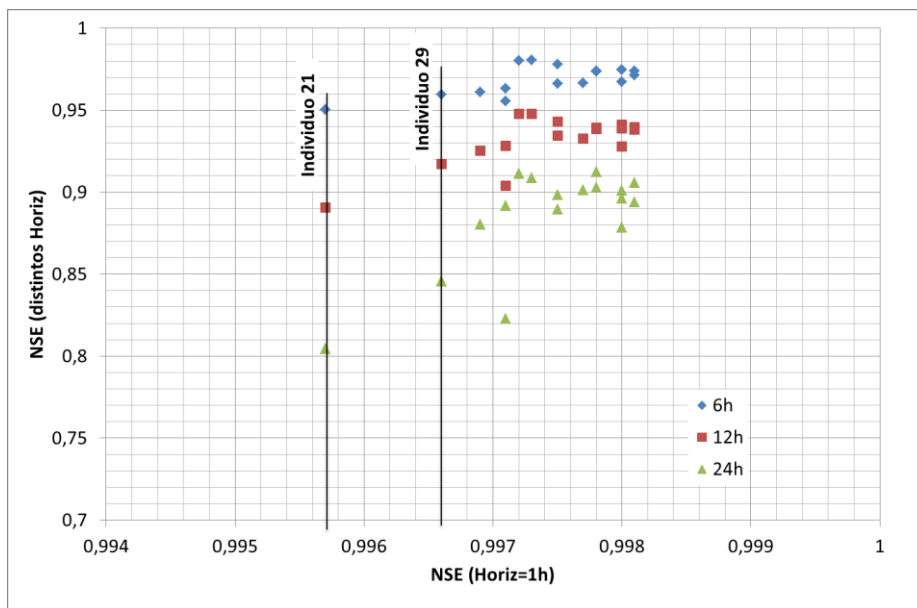


Figura 14-6: Zoom de la región más óptima del gráfico de la Figura 14-5. Se pierde la relación entre las mejoras de las performances.

En la Figura 14-7 se observa que luego de la iteración 100, la mejora de la performance fue muy escasa, hasta la última iteración (la número 5000). Además, la iteración 1000, 2000 y 5000 son prácticamente coincidentes. Esto quiere decir que **en pocas iteraciones relativamente se llega a una zona de parámetros optimizados con alta performance.**

Es importante destacar que el nivel de precisión de la performance se ha establecido en 4 decimales en la rutina de salida del programa, por esa razón todos los valores son redondeados en cuatro decimales y se encuentran en el gráfico, todos en las mismas líneas verticales encolumnados. No tiene sentido darle mayor precisión que 4 decimales porque en general dos performances que se diferencian entre ellas en 0.0001 no debería tener un impacto perceptible en los hidrogramas generados, ni tampoco responden normalmente a grandes variaciones en los valores de los parámetros, sino que son casi el mismo valor de parámetros. Es decir, prácticamente podríamos decir que todos los individuos que observamos en la zona optimizada equivalen a un punto optimizado que se podría haber buscado (como se ha dicho antes) mediante la optimización mono-objetivo.

En cualquier caso, puede verse que todos los puntos de la población se encuentran muy próximos, con performances casi idénticas. A continuación veremos la diferencia de valor de los parámetros que los conforman.

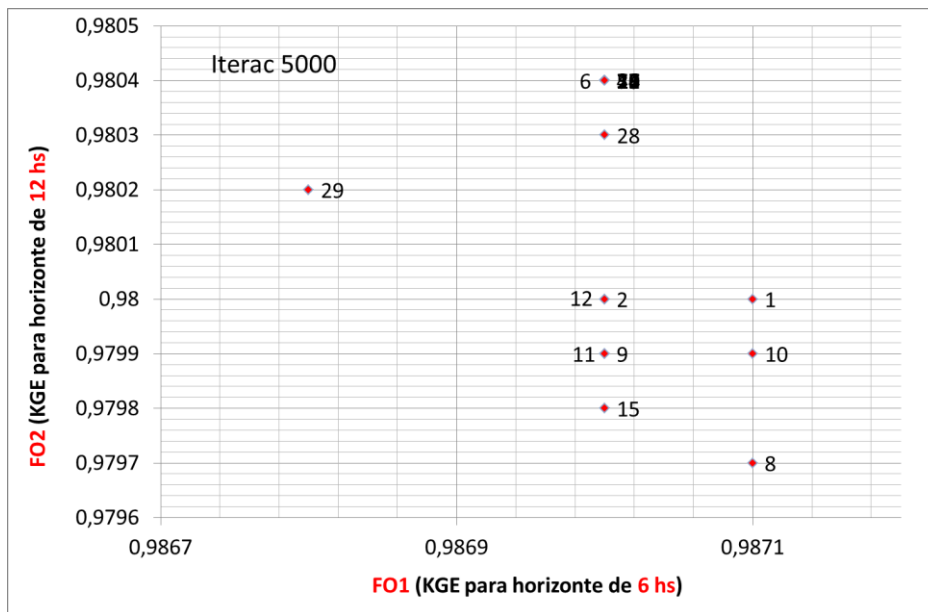


Figura 14-8: Población optimizada.

En la Tabla 14-2 se presenta el juego de valores de los parámetros para la población optimizada en la iteración 5000, con los valores de performances respectivos FO1 y FO2, para horizontes de pronóstico de 6 y 12 hs respectivamente. Vemos que los valores de los parámetros son muy semejantes entre sí. En la Tabla 14-1 se muestran los valores de los parámetros calculados por el IRSTEA (Ficchi, 2017) para la cuenca “Le Moros à Concarneau” empleando un método de calibración automática mono-objetivo, con el modelo en modo simulación. Es decir, si bien los parámetros fueron estimados para ajustar el modelo en modo simulación, son valores referenciales que ayudan a interpretar si se obtuvieron resultados semejantes.

Tabla 14-1: Parám y FO de dos juegos de parámetros “Le Moros à Concarneau” (Ficchi, 2017)

X1	X2	X3	X4	FOsim
417,57	-1,375	561,10	7,857	0,9548
419,75	-0,762	497,98	7,751	0,9688

Se observa que los valores de X2 y X4 calculados en este trabajo, son semejantes a los estimados por Irstea, pero no así los de X1 y X3 que presentan ciertas diferencias.

Tabla 14-2: Parámetros y FO de la Población optimizada en la iteración 5000

Indiv.	X1	X2	X3	X4	FO1	FO2
1	211,18	-0,671	1794,84	8,36	0,9871	0,98
2	209,94	-0,657	1800,59	8,32	0,987	0,98
3	215,08	-0,677	1714,96	8,34	0,987	0,9804
4	215,07	-0,677	1714,84	8,34	0,987	0,9804
5	215,10	-0,677	1714,86	8,34	0,987	0,9804
6	215,10	-0,677	1715,01	8,34	0,987	0,9804
7	215,10	-0,677	1714,86	8,34	0,987	0,9804
8	228,91	-0,695	1573,16	8,21	0,9871	0,9797
9	225,22	-0,701	1566,87	8,31	0,987	0,9799
10	209,30	-0,668	1811,44	8,37	0,9871	0,9799
11	224,41	-0,699	1563,88	8,36	0,987	0,9799
12	218,73	-0,686	1650,14	8,34	0,987	0,98
13	215,08	-0,677	1714,95	8,34	0,987	0,9804
14	214,87	-0,677	1715,27	8,35	0,987	0,9804
15	223,63	-0,700	1572,73	8,33	0,987	0,9798
16	215,09	-0,677	1714,93	8,34	0,987	0,9804
17	215,10	-0,677	1715,06	8,35	0,987	0,9804
18	215,08	-0,677	1714,99	8,34	0,987	0,9804
19	215,09	-0,677	1714,90	8,34	0,987	0,9804
20	214,89	-0,677	1714,71	8,35	0,987	0,9804
21	215,07	-0,677	1714,97	8,35	0,987	0,9804
22	215,63	-0,677	1714,68	8,33	0,987	0,9804
23	215,63	-0,677	1714,72	8,33	0,987	0,9804
24	215,59	-0,677	1714,77	8,33	0,987	0,9804
25	215,45	-0,677	1715,08	8,33	0,987	0,9804
26	215,50	-0,677	1714,60	8,33	0,987	0,9804
27	215,26	-0,677	1714,76	8,34	0,987	0,9804
28	215,56	-0,677	1715,18	8,33	0,987	0,9803
29	214,79	-0,676	1725,08	8,34	0,9868	0,9802
30	215,90	-0,677	1709,76	8,33	0,9866	0,9786

Sin embargo, en la Tabla 14-3 se observa que los valores máximos y mínimos de cada uno de los parámetros según la historia de datos del Irstea (Perrin, et al., 2007), contiene completamente a los valores obtenidos en este trabajo.

Tabla 14-3: Valores extremos de los parámetros según Irstea

RESUMEN	X1	X2	X3	X4
MIN	21,37	-29,78	5,13	0,13
MED	352,56	-1,42	153,18	20,41
MAX	3494,10	5,38	3006,77	239,88

Con respecto a los hidrogramas, se procede a estudiar el comportamiento de los mismos para un punto representativo elegido. Tomemos por ejemplo el punto 6, y grafiquemos los hidrogramas para distintos horizontes de pronóstico.

En la Figura 14-9 se observa que el modelo representa con aceptable semejanza al hidrograma observado, mientras que en la Figura 14-10, se observa con mayor detalle que a mayor tiempo de antelación, mayor es el error cometido en el pronóstico. En el mismo puede verse que, a pesar de que el punto 6 tiene muy buenos coeficientes de KGE (0,987 para 6 hs, y 0,9804 para 12 hs), se observan errores de consideración a la hora de pronosticar el pico del hidrograma, tanto para 6 como para 12 hs de tiempo de antelación.

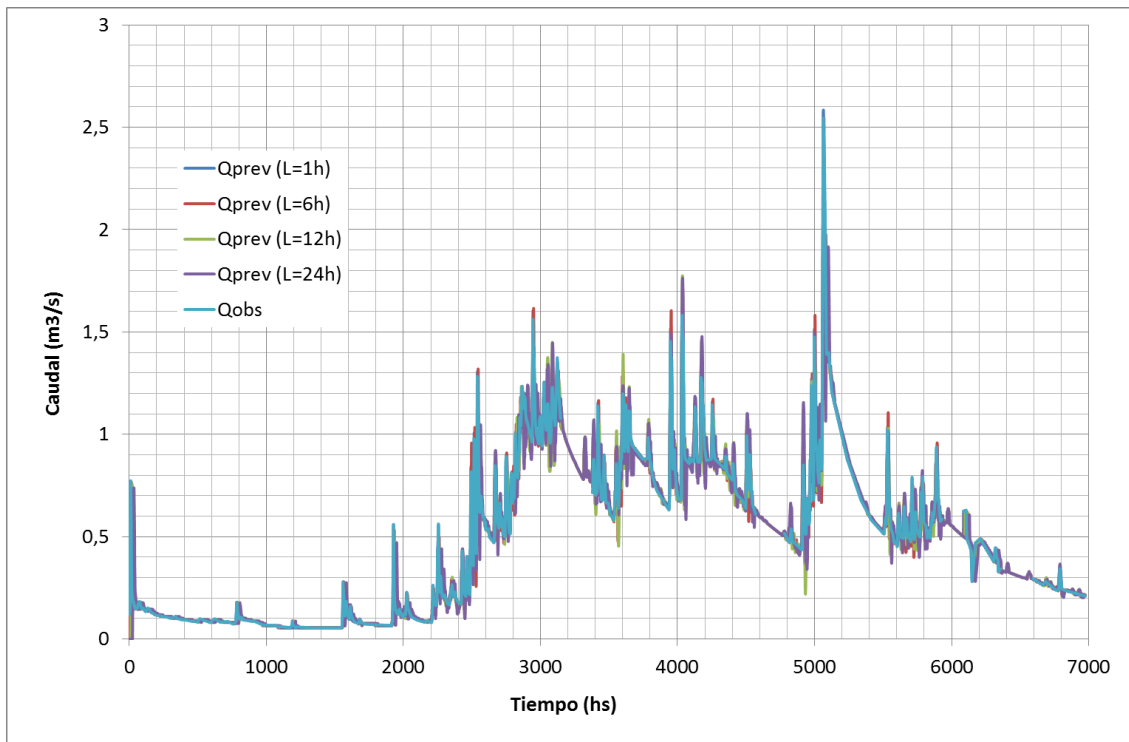


Figura 14-9: Hidrograma para la serie simulada con los parámetros del individuo 6.

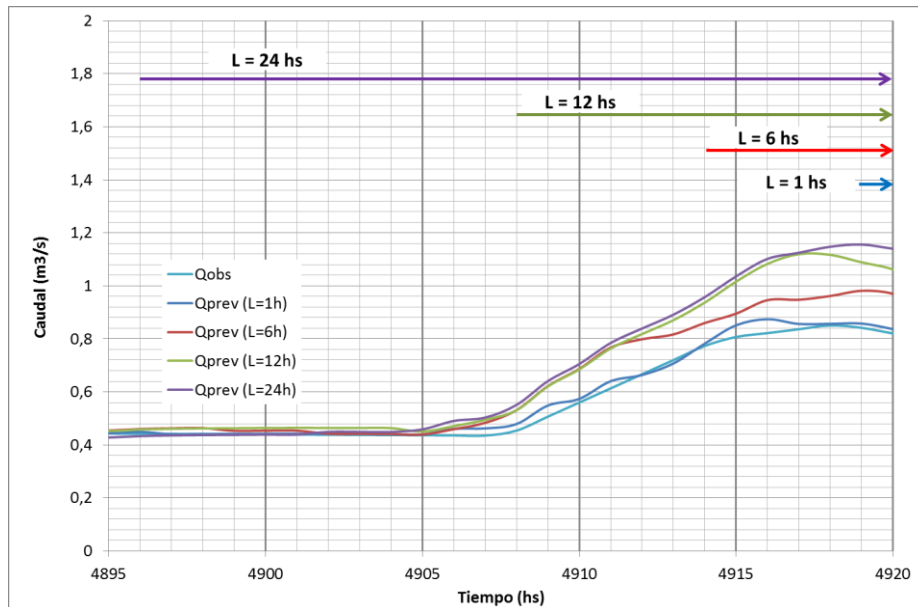


Figura 14-10: Hidrograma pronosticado para distintos horizontes (individuo 6), haciendo zoom en un evento.

En el gráfico de la Figura 14-11 vemos que dos puntos como el 6 y el 8, ubicados en la zona óptima, son ambos casi idénticos, en términos de los hidrogramas que generan.

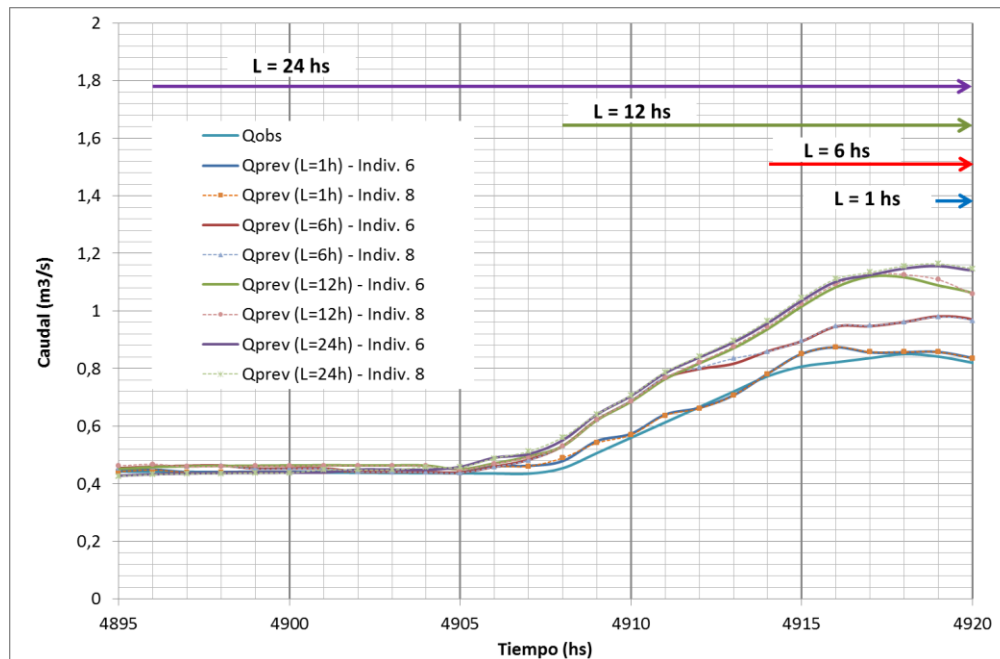


Figura 14-11: Comparación de hidrogramas para los individuos 6 y 8.

14.4. Síntesis de los resultados y discusiones

Los experimentos realizados muestran que el proceso de optimización se aproxima rápidamente a la zona óptima, satisfaciendo solidariamente ambas FO en cada mejora. Luego de la primer centena de iteraciones, el método sigue optimizando la población, pero en pequeñas mejoras diferenciales y comienza a perderse el efecto de mejora simultánea de ambas FO.

En esta región, se comienza a desarrollar un frente de Pareto, que no ha podido completarse satisfactoriamente en todos los procesos de optimización. Es decir, fueron obtenidos frentes de Pareto "incompletos" que al menos mostraron una tendencia de desarrollo en una pequeña región de elevados valores de FO. Sin embargo, estos frentes "incompletos" en la mayor parte de los casos satisficieron un mínimo de condiciones de desarrollo establecidas como condiciones de salida (pequeña sensibilidad de las FO a nuevas variaciones en los parámetros, y un

porcentaje mínimo de individuos no dominados), que representaron la silueta aproximada de la región óptima.

Los hidrogramas obtenidos para los puntos integrantes del frente de Pareto dieron testimonio del alto grado de semejanza entre los individuos de la población optimizada, ya que mostraron desarrollos prácticamente equivalentes.

Sin embargo, a pesar de obtener frentes de Pareto con individuos homogéneos en términos de los hidrogramas generados, y con FO elevadas, se destaca una profunda caída de la performance en términos gráficos a medida que se incrementa el horizonte de pronóstico.

14.5. Conclusiones

En este capítulo se evaluó la capacidad del modelo de adaptarse a distintos escenarios durante el proceso de pronóstico. Concretamente, se estudió la aptitud de mantener niveles de performance aceptables a medida que se modifican los horizontes de pronóstico estipulados. En este sentido, se empleó la técnica de optimización multi-objetivo para confrontar FO que representen las performances del modelo para dos distintos horizontes de pronóstico.

A partir de los resultados obtenidos en las pruebas y experimentos realizados, se puede concluir que existe una fuerte correlación entre las FO asociadas a los diversos horizontes de pronóstico. Esto indicaría que todo juego de valores de parámetros que mejore el pronóstico del modelo frente a un horizonte de pronóstico, también lo hace en los otros horizontes. Expresado de modo inverso, los valores paramétricos que optimizan la performance del modelo son insensibles al horizonte de pronóstico. Desde esta perspectiva, se concluye que no parece tener asidero fáctico la

confrontación de FO que representen performances del modelo frente a diversos horizontes de pronóstico.

Ahora bien, en las proximidades del óptimo deja de tener validez la regla anterior, ya que en esa región se observa que las FO comienzan a competir desarrollando frentes de Pareto en espacios hiperdimensionales muy reducidos.

La prueba de calidad del ajuste de los hidrogramas probó que dos individuos integrantes del frente de Pareto, distantes entre sí, pero dentro del espacio reducido referido, presentan performances semejantes, para iguales horizontes de pronóstico. Por lo tanto, se concluye que la optimización mono-objetivo podría reemplazar a la optimización multi-objetivo sin sacrificar fácticamente la performance de los pronósticos realizados.

15. Conclusión General

La ciencia, muchacho, está hecha de errores, pero de errores útiles de cometer, pues poco a poco, conducen a la verdad.

Julio Verne

Esta tesis se ha desarrollado sobre la base de una problemática actual de elevada importancia a nivel local y global, como son las crecidas e inundaciones fluviales. En el universo de los múltiples factores y actores que participan de este tipo de fenómenos, se destaca como una herramienta de particular utilidad para la prevención oportuna de la catástrofe y mitigación parcial de daños humanos y materiales, los modelos de pronóstico hidrológico.

Tradicionalmente, el uso de modelos de pronóstico hidrológico está arraigado en el clásico método de calibración y empleo de dichos modelos a partir del mejor juego (único) de valores de parámetros.

Sin embargo, en la práctica habitual el modelo es sometido a un sinnúmero de escenarios cambiantes durante el proceso operativo para los cuales no ha sido probado, y por lo tanto no es posible asegurar el éxito de su desempeño, o al menos la incertidumbre en sus valores pronosticados.

Por las razones expuestas, esta tesis ha centrado su foco en el estudio del comportamiento de un modelo de pronóstico hidrológico frente a diversos escenarios de aplicación práctica, evaluando su performance a partir de Funciones Objetivo que cuantifican el error en las estimaciones. Para abordar este propósito se hizo uso de una técnica de optimización multi-objetivo capaz de confrontar numéricamente los escenarios propuestos.

A partir de una importante revisión de antecedentes, se seleccionó el método MOCOM-UA como un procedimiento robusto para confrontar objetivos. Ahora bien, el empleo del MOCOM demostró una cierta incapacidad para llegar en forma rápida y precisa al frente de Pareto, por lo que inesperadamente, el foco de la tesis se concentró en mejorar este método para representar el experimento requerido.

Como primera medida, entonces, se evaluó y mejoró la herramienta de optimización multi-objetivo MOCOM-UA produciendo una adaptación en el método para mejorar su robustez. Los primeros resultados demostraron que el método, en su estado original, tiende al frente de Pareto, pero es incapaz de aportar una única solución optimizada en todas las pruebas. A partir de un conjunto de técnicas de corrección se alcanzaron en este trabajo desempeños del modelo altamente superiores a las obtenidas con el método MOCOM-UA original. En efecto, pudo notarse una clara mejora de la robustez en la estimación del frente de Pareto, así como en el consumo computacional manifestado en el menor número de iteraciones requeridas para aproximarse al óptimo.

Este esfuerzo dio como fruto una primera conclusión del trabajo: la obtención de un método MOCOM Mejorado, que representa una versión con una notoria superioridad en robustez y eficacia que el original. Con esto se responde al **tercer interrogante científico** planteado en el apartado respectivo. **Sí, fue posible lograr una metodología capaz de ofrecer una herramienta de decisión al operador de un sistema de pronóstico hidrológico, que le permita evaluar en qué posición de compromiso ubicarse a la hora de producir un pronóstico de caudal.**

Se considera que este aporte al campo de la optimización de modelos matemáticos, contribuye levemente al desempeño del modelo en su proceso de calibración, pero aporta elementos sólidos a la reducción de la incertidumbre en la estimación del juego de valores óptimos para la simulación y pronóstico de modelos hidrológicos. En consecuencia, el método MOCOM-UA "modificado" es sugerido a los fines de

obtener robustamente el frente de Pareto para condiciones de confrontación de objetivos definidos como performances de un modelo de pronóstico hidrológico frente a diversos escenarios operativos.

Un segundo estudio que se abordó en este trabajo se desarrolló sobre la base de una situación clásica en operación de pronóstico hidrológico, vinculado conceptualmente a la importancia del componente de actualización de estado en el modelo de pronóstico. En efecto, se evaluó la pérdida de performance que experimenta un modelo de pronóstico cuando la componente de actualización de estado se queda súbitamente sin la provisión a tiempo real de la(s) variable(s) que permite(n) actualizar el estado del modelo. Expresado de otro modo, se confrontaron dos escenarios: el modelo con disponibilidad de información de actualización a tiempo real, *versus* el modelo sin dicha información.

Para llevar a cabo los experimentos referidos se empleó el modelo GR4P. Pudo observarse que este modelo constituye una poderosa herramienta para la estimación del caudal con un determinado horizonte de previsión. Sin embargo, se confirmó que la exactitud de la estimación es altamente dependiente de la disponibilidad a tiempo real del caudal observado, que le permite actualizar el estado del modelo. El empleo de la optimización multi-objetivo para la calibración de la cuenca “Moros en Concarneau” entre ambos escenarios, permitió alcanzar un conjunto de soluciones con distintas performances en ambas FO.

Se observó que, para horizontes menores o iguales al horizonte empleado para la calibración del modelo, la actualización de estado produce una mejora del desempeño del modelo, sin embargo, hay una notoria tendencia en desmejorar la performance debido a la actualización de estado, a medida que aumenta el horizonte L por encima del horizonte de calibración. Esto puede ser explicado porque el procedimiento de actualización del modelo, obliga a que a cada paso de tiempo el

valor pronosticado tome exactamente el valor observado en cada Δt , quitándole de este modo grados de libertad en la selección del juego óptimo de parámetros para satisfacer todos los valores del tramo de hidrograma empleado para calibrar.

Se puso en evidencia que la bondad de un juego de parámetros no sólo depende de la performance del modelo frente al horizonte de pronóstico de interés, sino también a la sensibilidad de la performance a variaciones en el tiempo de horizonte establecido, a medida que se aleja del horizonte con el cual fue calibrado.

De las reflexiones anteriores, se concluye que la optimización multi-objetivo constituye una herramienta operativa de gran utilidad, a los fines de contar con mayor información de base frente a la decisión de adoptar uno o dos juegos de valores de parámetros capaces de adaptarse a los dos escenarios posibles durante la operación: disponibilidad o no disponibilidad de los valores de caudal observado a tiempo real. Esto es porque permite tomar decisiones sobre un conjunto de soluciones de compromiso.

Finalmente, en esta tesis se abordó un tercer trabajo, mediante el cual se estudió la aptitud de mantener niveles de performance aceptables a medida que se modifican los horizontes de pronóstico estipulados. En este sentido, se empleó la técnica de optimización multi-objetivo para confrontar FO que representen las performances del modelo para dos distintos horizontes de pronóstico.

A partir de los resultados obtenidos en las pruebas y experimentos realizados, se concluyó que existe una fuerte correlación entre las FO asociadas a los diversos horizontes de pronóstico. Esto indicó que todo juego de valores de parámetros que mejora el pronóstico del modelo frente a un horizonte de pronóstico, también lo hace en los otros horizontes. En consecuencia, aun cuando en las proximidades del óptimo deja de tener validez la regla anterior, la prueba de calidad del ajuste de los hidrogramas probó que dos individuos integrantes del frente de Pareto, distantes

entre sí, pero dentro del espacio reducido referido, presentan performances semejantes, para iguales horizontes de pronóstico. Por lo que se concluye que, la optimización mono-objetivo podría reemplazar a la optimización multi-objetivo sin sacrificar fácticamente la performance de los pronósticos realizados.

De este modo, se da respuesta al primero de los objetivos planteados en la tesis que interroga con relación a la posibilidad de controlar en tiempo real, mediante las combinaciones de valores paramétricos del modelo, su adaptabilidad a los diversos escenarios de pronóstico. La respuesta es, **en algunos casos es posible (iy necesario!) contar con un juego de valores de parámetros para cada escenario**, ya que la elección del juego de valores de los parámetros óptimo para un escenario, refleja una mala calidad para el otro, y viceversa. Es decir que **sí es posible controlar la adaptabilidad del modelo a los diversos escenarios mediante la adecuada elección de los valores paramétricos.**

Por otro lado, hay casos particulares en los cuales un solo juego de valores para los parámetros consigue buena calidad de pronóstico. Un ejemplo muy concreto lo mostramos en el “individuo 22”, explicado en el apartado 13.3 de la tesis, en el experimento de confrontación de desempeño del modelo en “modo simulación” Vs “modo pronóstico”. Pero el ejemplo más concreto se da en los experimentos del modelo de pronóstico confrontando distintos horizontes (apartado 14.3), ya que se observó que el modelo tiene una tendencia a satisfacer de igual modo ambos horizontes, lo cual da muestra de la robustez del modelo GR4P.

Con estos comentarios se da respuesta al **segundo interrogante científico**, ya que ponen de manifiesto que **sí, es posible contar con algún juego de parámetros capaz de satisfacer simultáneamente diversos escenarios con aceptable precisión**. Sin embargo, aquí hay que aclarar que **sólo para algunos casos**.

Como conclusión general, se considera que la presente tesis fue capaz de explorar la aptitud de un método de optimización multi-objetivo de aportar elementos de decisión en relación a la capacidad de un modelo de pronóstico hidrológico para adaptarse con buena performance a diversos escenarios típicos de operación, a partir de una adecuación de sus valores paramétricos. En síntesis, se ofrece una herramienta que aporta elementos novedosos al operador de un sistema de pronóstico que contribuyen a enriquecer la base de información con la cual cuenta, antes de tomar una decisión.

16. Referencias bibliográficas

- Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J. (1986a). An introduction to the European Hydrological System -- Système Hydrologique Européen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.*, 87: 45-59.
- Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J. (1986b). An introduction to the European Hydrological System -- Système Hydrologique Européen, "SHE", 2: Structure of a physically-based, distributed modelling system. *J. Hydrol.*, 87: 61-77.
- Abdulla, F.A., Lettenmaier, D.P., Liang, X. (1999). Estimation of the ARNO model baseflow parameters using daily streamflow data. *Journal of Hydrology*, Vol.222 pp.37-54.
- Abebe, A.J., Solomatine, D.P. (1998). Application of global optimization to the design of pipe networks. *Proc. 3rd International Conference on Hydroinformatics*, Copenhagen, August 1998. Balkema, Rotterdam.
- Abrahart, R., Kneale, P.E., See, L.M. (2004). *Neural Networks for Hydrological Modeling* CRC Press.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2002). *Molecular Biology of the Cell*. New York: Garland Publishing. 4th ed.
- Alonso, F.J. (2008). *Calibración Automática de un Modelo Hidrológico basada en Algoritmos Genéticos, aplicación a dos cuencas típicas argentinas*. Tesis de Maestría en Ciencias de la Ingeniería, Mención en Recursos Hídricos. Universidad Nacional de Córdoba.
- Alonso, F.J., Bertoni, J.C., García, C., García, C.M. (2015). Algoritmos genéticos como herramienta de decisión frente a diversos escenarios de pronóstico hidrológico. *Aporte Santiaguino*. 8 (1), 2015: 45-56.

- Alonso, F.J., Le Moine, N., Oudin, L., Ribstein, P., Bertoni, J.C. (2017). Adaptación de un método de optimización multi-objetivo para modelos de pronóstico de inundaciones. *Rev. Int. de Desastres Naturales, Accidentes e Infraestructura Civil*. Vol. 16(1): 48-64.
- Alonso, F.J., Le Moine, N., Oudin, L., Ribstein, P., Bertoni, J.C., Patalano, A., Ramos, M-H, Perrin, Ch., Nascimento, N. (2016). Performance analysis of real time flood forecasting models for different forecast horizons using optimization algorithms. 3rd European Conference on Flood Risk Management (Flood Risk 2016). Lyon, France.
- Andreassian, V., Bergström, S., Chahinian, N., Duan, Q., Gusev, Y.M., Littlewood, I., Mathevet, T., Michel, C., Montanari, A., Moretti, G., Moussa, R., Nasonova, O.N., O'connor, K., Paquet, E., Perrin, C., Rousseau, A., Schaake, J., Wagener, T., Xie, Z. (2006). Catalogue of the models used in MOPEX 2004/2005. Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model ParameterExperiment–MOPEX. IAHS Publ. 307, 2006.
- Arduino, G., Reggiani, P., Todini, E. (2005). Recent advances in flood forecasting and flood risk assessment. *Hydrology and Earth System Sciences*, 9(4), 280-284.
- Armién, F. (2001). Manual Sistemas de Alerta Temprana. Ministerio de Educación de Panamá. MEDUCA.
- Beard, L. (1966). Optimization Techniques for Hydrologic Engineering. Technical Paper N.2, Hydrologic Engineer Center, Davis Ca.
- Berthet, L. (2010). Prévion des crues au pas de temps horaire: pour une meilleure assimilation de l'information de débit dans un modèle hydrologique. PhD Thesis de l'Institut des Sciences et Industries du Vivant et de l'Environnement AgroParisTech.
- Berthet, L., Andreassian, V., Perrin, C., Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-

- based and continuous approaches on 178 catchments. *Hydrol. Earth Syst. Sci.*, 13, 819–831, 2009.
- Bertoni, J.C. (2004a). Programa de Capacitación Para Gestores y Tomadores de Decisión Públicos en el Campo de los Recursos Hídricos. UNESCO. Global Water Partnership, GWP-SAMTAC.
- Bertoni, J.C. (2004b). Inundaciones Urbanas en Argentina. Universidad Nacional de Córdoba.
- Bertoni, J.C. (2003). Curso de Hidrología Avanzada. Notas de clase de la Maestría en Ciencias de la Ingeniería. FCEFyN. UNC.
- Beven, K. (1987). Towards a new paradigm in hydrology. *Water for the Future: Hydrology in Perspective (Proceedings of the Rome Symposium, April 1987)*. IAHS Publ. no. 164, 1987.
- Beven, K. (2004). Infiltration excess at the Horton Hydrology Laboratory (or not?). *Journal of Hydrology* 293 (2004) 219–234.
- Beven, K. (2004). Robert E. Horton's perceptual model of infiltration processes. *Hydrol. Process.* 18, 3447–3460 (2004).
- Beven, K. (1989). Changing ideas in hydrology- the case of physically-based models. *Journal of Hydrology*, 105 (1989) 157-172.
- Beven, K., Binley, A. (1992). The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes*, Vol. 6, 279-298 (1992).
- Beven, K.J., Kirkby, M.J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences-Bulletin*, 24,1, 3/1979.
- Beveridge, G.S. and Shechter, R.S. (1970). *Optimization: theory and practice*. McGraw Hill, New York.
- Box, G. E. P. & Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

- Boyle, D.P., Gupta, H.V. and Sorooshian, S. (2000). Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research*. Vol. 36. Nº12. Pags. 3663-3674.
- Bras, R.L., Rodriguez-Iturbe, I. (1989). A review of the search for a quantitative link between hydrologic response and fluvial geomorphology. *New Directions for Surface Water Modeling (Proceedings of the Baltimore Symposium, May 1989)* IAHS Publ.no. 181,1989.
- Brath, A., Rosso, R. (1993). Adaptive calibration of a conceptual model for flash flood forecasting. *Water Resources Research*, 29(8), 2561-2572.
- Bravo, J.M., Allasia, D.G., Collischonn, W., Tassi, R., Meller, A., Tucci, C.E.M. (2006). Modelo IPH2 para Windows. Manual de conceptos Instituto de Pesquisas Hidraulicas. Puerto Alegre. Brazil.
- Brazil, L.E. (1988). Multilevel calibration strategy for complex hydrologic simulation models. PhD Dissertation Colorado State University, Fort Collins.
- Brazil, L.E., Krajewski, W.F. (1987). Optimization of Complex Hydrologic Models Using Random Search Methods Conference on Engineering Hydrology, Hydraulics Division, Am. Soc. of Civ. Eng., Williamsburg V, a., Aug. 3-7, 1987.
- Burnash, R.J.C. (1995). The NWS river forecast system catchment modelling. In: Singh, V. J. (editor) *Computer models of watershed hydrology*. Water Resources Publications, Highlands Ranch. 1130 p.
- Burnash, R.J.C., R.L. Ferral, and R.A. McGuire. (1973). A Generalized Streamflow Simulation System - Conceptual Modeling for Digital Computers. U.S. Department of Commerce, National Weather Service and State of California, Dept. of Water Resources, March.
- Bustos, N. (2014). Bases para mejorar la performance de un modelo hidrológico concentrado frente a distintos patrones de distribución areal de precipitaciones.

- Aplicación a la cuenca del río San Antonio. Tesina de grado para alcanzar el título de Ingeniero Civil. Universidad Nacional de Córdoba.
- Castillo, E., Conejo, A., Pedregal, P., García, R. y Alguacil, N. (2002). *Formulación y Resolución de Modelos de Programación Matemática en Ingeniería y Ciencia*. Universidad de Castilla-La Mancha, España.
- Cemagref (2012). *Le Moros à Concarneau Description des caractéristiques morphologiques, climatiques et hydrologiques de 4436 bassins versants français. Guide d'utilisation de la base de données hydro-climatique*.
- Chow V.T., Maidment D.R., Mays L.W. (1994). *Hidrología aplicada*. McGraw-Hill. Interamérica, S.A., Colombia, 584 pp.
- Chow, V.T. (1964). *Handbook of Applied Hydrology*. New York: McGraw-Hill Book Co.
- Ciarapica L, Todini E. (2002). TOPKAPI: a model for the representation of the rainfall–runoff process at different scales. *Hydrological Processes* 16(2): 207–229.
- Clarke, R.T. (1973). A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of Hydrology*, 19, 1-20.
- Collischonn, W., Tucci, C.E.M. (2003). Ajuste multiobjetivo dos parâmetros de um modelo hidrológico. *Revista Brasileira de Recursos Hídricos*, Porto Alegre, v. 8, n.3, p. 27-39.
- Collischonn, W., Tucci, C.E.M. (2001). Calibração automática global do modelo IPH-2. In: *Simpósio Brasileiro de Recursos Hídricos, 2001, Aracaju SE. Anais do Simpósio Brasileiro de Recursos Hídricos*.
- Cooper, V.A., Nguyen, V.T.V., Nicell, J.A. (2007). Calibration of conceptual rainfall–runoff models using global optimisation methods with hydrologic process-based parameter constraints. *Journal of Hydrology* (2007) 334, 455– 466.

- Cooper, V.A., Nguyen, V.T.V., Nicell, J.A. (1997). Evaluation of global optimization methods for conceptual rainfall–runoff model calibration. *Water Sci. Technol.*, 36 (5), 53–60.
- Corral, C. (2004). Desenvolupament d'un model hidrològic per incorporar informació del radar meteorològic. Aplicació operacional a la conca del riu Besòs. Tesis doctoral. Universidad Politècnica de Catalunya.
- Crawford, N.H. and R.K. Linsley. (1966). Digital Simulation on Hydrology: Stanford Watershed Model IV. Stanford University Technical Report No. 39, Stanford University, Palo Alto, CA.
- CRED (2012). Annual Disaster Statistical Review 2011. Université Catholique de Louvain.
- Dawdy, D.R. and O'Donnell, T. (1965). Mathematical models of catchment behavior. *Journal of the Hydraulics Division*, 91(4), pp.123-137.
- Dawson, C.W., Mount, N.J., Abrahart, R.J., Shamseldin, A.Y. (2012). Ideal point error for model assessment in data-driven river flow forecasting. *Hydrol. Earth Syst. Sci.*, 16, 3049–3060, 2012.
- DesInventar (2010). Sistema de inventario de efectos de desastres. <http://www.desinventar.net>
- Dooge, J. C. I. (1974). The development of hydrological concepts in Britain and Ireland between 1674 and 1874. *Hydrol. Sci. Bull.*, 19(3), 279–302.
- Dooge, J.C.I. (1984). The waters of the Earth. *Hydrological Sciences Journal*, 29:2, 149-176.
- Dooge, J.C.I. (1973). Linear theory of hydrologic systems. Technical Bulletin No. 1468, Agricultural Research Service, U.S. Department of Agriculture, Washington, D.C.
- Duan, Q., Gupta, V.K., Sorooshian, S. (1993). A shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, 76 (3), 501-521.

- Duan, Q., Sorooshian, S., Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* Vol. 28, Nº 4 pp. 1015-1031.
- Duan, Q., Sorooshian, S., Gupta, V. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, Vol. 158, pp. 265-284.
- Duband, D. (2000). Réflexion actuelle sur la prévision hydrométéorologique. *La Houille Blanche*. Nº 1/2000.
- Duband, D. (1999). Réflexion actuelle sur la prévision hydrométéorologique. Colloque SHF "La gestion des risques liés aux inondations", 29 et 30 septembre 1999, Paris.
- Duband, D. (2000). Les aléas extrêmes de crues et inondations. Colloque Risques et Génie Civil – Paris 8/9 novembre 2000.
- Eckhardt, K., Arnold, J. G. (2001). Automatic calibration of a distributed catchment model. *Journal of Hydrology*. Vol. 251. Pags. 103-109.
- Efstratiadis, A., Koutsoyiannis, D. (2010). One decade of multi-objective calibration approaches in hydrological modelling: a review *Hydrological Sciences Journal – Journal des Sciences Hydrologiques*, 55(1) 2010.
- Engman, E.T., and A.S. Rogowski. (1974). A partial area model for storm flow synthesis. *Water Resour. Res.* 10:464–472.
- Espinosa, J.A. (2012). Water Management in the Panama Canal. International Congress PANAMA CANAL 2012. Engineering and Infrastructure. Ciudad de Panamá, Panamá.
- Failache, N., Zamanillo, E. Irigoyen, M. (2010). Precisión de los Pronósticos Hidrológicos del Río Uruguay en Salto Grande. CAPAH, Neuquén.

- Ficchi, A. (2017). Un modèle hydrologique adaptatif à différents pas de temps: diagnostic et améliorations basés sur la cohérence des flux. PhD Thesis de l'École doctorale Géosciences, ressources naturelles et environnement (Paris) en partenariat avec Hydrosystèmes et bioprocédés (laboratoire).
- Franchini, M., Galeati, G., Berra, S. (1998). Global optimization techniques for the calibration of conceptual rainfall-runoff models. *Hydrological Sciences—Journal—des Sciences Hydrologiques*, 43(3) June.
- Freeze, R.A., Harlan, R.L. (1969). Blueprint for a Physically-Based, Digitally-Simulated Hydrologic Response Model. *Journal of Hydrology* 9 (1969) 237-258.
- Friedman, M., Savage, L. (1947). *Selected Techniques of Statistical Analysis*. MacGraw-Hill Book Co. of Water Resources Systems IFIP Amsterdam.
- Gan, T.Y., Biftu, G.F. (1996). Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure. *Water Resources Research*. Vol. 32. Nº 12. Pags. 3513-3524.
- Gan, T.Y., Burges, S.J. (1990). An Assessment of a Conceptual Rainfall-Runoff Model's Ability to Represent the Dynamics of Small Hypothetical Catchments: 2. Hydrologic Responses for Normal and Extreme Rainfall. *Water Resources Research* 26 (7): 1605-1619.
- Gan, T.Y., Burges, S.J. (1990). An Assessment of a Conceptual Rainfall-Runoff Model's Ability to Represent the Dynamics of Small Hypothetical Catchments: 1. Models, Model Properties and Experimental Design. *Water Resources Research* 26 (7): 1595-1604.
- García-Bartual, R. (1989). Estimación de la Respuesta Hidrológica de una cuenca sobre la base de la Teoría del Hidrograma Unitario Geomorfológico. *Cuaternario y Geomorfología*, vol. 3 (1-4), 1-7, 1989.
- Garrick, M., Cunnane, C., Nash, J.E. (1978). A criterion of efficiency for rainfall-runoff models. *Journal of Hydrology*, 36, 375-381.

- Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaskovicova, L., Bloschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A., Kohnova, S., Koutroulis, A., Marchi, L., Matreata, S., Medina, V., Preciso, E., Sempere-Torres, D., Stancalie, G., Szolgay, J., Tsanis, I., Velasco, D., Viglione, A. (2009). A compilation of data on European flash floods. *Journal of Hydrology* 367 (2009) 70–78.
- Georgakakos, K.P. (1985). A Flash-Flood Prediction System. *Tropical Hydrology Symposium*. San Juan, Puerto Rico.
- Georgakakos, K.P., Bras, R.L. (1984). A Hydrologically Useful Station Precipitation Model. 1. Formulation. *Water Resources Research* 20 (11): 1585-1596.
- Georgakakos, K.P., Bras, R.L. (1984). A Hydrologically Useful Station Precipitation Model. 2. Case Studies. *Water Resources Research* 20 (11): 1597-1610.
- Georgakakos, K.P., Hudlow, M.D. (1984.a.). Integrated Hydrometeorological Forecast System - Design and Tests. *Hydrologic Research Laboratory - National Weather Service, NOAA*.
- Georgakakos, K.P., Hudlow, M.D. (1984.b.). Quantitative Precipitation Forecast Techniques for Use in Hydrologic Forecasting. *Bulletin American Meteorological Society*, 65(11), pp. 1186-1200.
- Georgakakos, K.P., Hudlow, M.D. (1984). Quantitative Precipitation Forecast Techniques for Use in Hydrologic Forecasting. *Bull. Amer. Meteor. Soc.*, 65, 1186–1200.
- Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Pub. Co.
- Gomez Rocha, H. y Acosta, J., Velez-Langs, O. (2006). Utilizacion de la Computacion Evolutiva y la Teoria de Juegos para la Automatización en Sistemas de Soporte a la Negociación. Cap.3. Trabajo Final de la Carrera de Ingeniero de Sistemas. Universidad de Magdalena (Santa Marta-Colombia).

- González, L.G. (2009). Reducir los riesgos y hacer frente a la incertidumbre Seminario Desafíos frente a la vida y el bienestar en la gestión del agua. IARH.
- Goswami, M., O'Connor K.M. (2007). Comparative assessment of six automatic optimization techniques for calibration of a conceptual rainfall-runoff model. *Hydrological Sciences Journal*, 52:3, 432-449.
- Goswami, M., O'Connor K.M. (2007). Real-time flow forecasting in the absence of quantitative precipitation forecasts: A multi-model approach. *Journal of Hydrology* (2007) 334, 125– 140.
- Gourley, J.J., Vieux, B.E. (2006). A method for identifying sources of model uncertainty in rainfall-runoff simulations. *Journal of Hydrology* (2006) 327, 68– 80.
- Green, I.R.A., Stephenson, D. (1986). Criteria for comparison of single event models. *Hydrological Sciences - Journal - des Sciences Hydrologiques*, 31,3, 9/1986.
- Guinot, V., Cappelaere, B., Delenne, C., Ruelland, D. (2011). Towards improved criteria for hydrological model calibration: theoretical analysis of distance- and weak form-based functions. *Journal of Hydrology* 401 (2011) 1–13.
- Gupta, H. V., L. A. Bastidas, J. A. Vrugt, and S. Sorooshian. (2003). Multiple criteria global optimization for watershed model calibration, in *Calibration of Watershed Models*, edited, pp. 125-132, AGU, Washington, DC.
- Gupta, H., Sorooshian, S., and Yapo, P. (1999). Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *J. Hydrol. Eng.*, 10.1061/(ASCE)1084-699(1999)4:2(135), 135-143.
- Gupta, H.V., Kling, H., Yilmaz, K., Martinez, G.F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* 377 (2009) 80–91.
- Gupta, V., Sorooshian, S. and Yapo, P.O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*. Vol. 34. Nº4. Pags. 751-763.

- HEC-HMS. (2000). Hydrologic Modeling System. Technical Reference Manual. U.S.Army, Hydrologic Engineering.
- Hendrichson, J.D., Sorooshian, S., Brazil, L.E. (1988). Comparison of Newton-Type and Direct Search Algorithms for Calibration of Conceptual Rainfall-Runoff Models. Water Resources Research. Vol. 24. Nº5, Pags. 691-700.
- Heraldo.es (2012). La CHE analiza la previsión de caudales para las próximas horas. En: periódico digital "Heraldo.es", del día (19/10/2012).
- Hincapié, R.A, Ríos Porras, C.A., Gallego, R.A. (2004). Técnicas heurísticas aplicadas al problema del cartero viajante (TSP). Scientia et Technica Año X, No 24, Mayo 2004. UTP. ISSN 0122-1701.
- Holland, J.H. (1975). Adaptation in natural and artificial systems. Ann Arbor: The University of Michigan Press.
- Horton, R.E. (1945). Erosional Development of Streams and their Drainage Basins; Hydrophysical Approach to Quantitative Morphology. Bulletin of the Geological Society of America. Vol. 56, pp. 275-370.
- Hydrologic Engineering Center. (1990). HEC 1 User's Manuals. U.S. Army Corps of Engineers, Davis, California.
- Ibbitt, R.P. (1972). Effects of Random Data Errors of the Parameter Values for a Conceptual Model. Water Resources Research. Vol. 8. Nº 1. Pags. 70-78.
- Ibbitt, R.P. (1970). Systematic parameter fitting for conceptual models of catchment hydrology. Ph.D. dissertation, Imperial College of Science and Technology, University of London.
- Jakeman, A.J., Hornberger, G.M. (1993). How Much Complexity Is Warranted in a Rainfall-Runoff Model? Water Resources Research, Vol. 29, Nº8, P. 2637-2649, Aug. 1993.

- Johnston, P.R., Pilgrim, D.H. (1976). Parameter optimization for watershed model. *Water Resources Research* Vol. 12, Nº 3 pp. 477-486.
- Kauark-Leite, L.A. et Nascimento, N.O. (1993). Développement, utilisation et incertitudes des modèles conceptuels en hydrologie. Séminaire Eaux - Environnement n° 3, Modélisation du comportement des polluants dans les hydrosystèmes, Paris, 191-219.
- Kitanidis, P.K., and Bras, R.L. (1980). Real-time forecasting with a conceptual hydrologic model. 2. Applications and results. *Water Resources Research*, Vol. 16, No. 6, pp. 1034:1044.
- Kitanidis, P.K., and Bras, R.L. (1980). Real-time forecasting with a conceptual hydrologic model. 1. Analysis of Uncertainty. *Water Resources Research*, Vol. 16, No. 6, pp. 1025:1033.
- Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31(1), 13-24.
- Klemeš, V. (1986). Dilettantism in Hydrology: transition or destiny? *Water Resources Research*, 22(9): 1775-1885.
- Krauß, T., Cullmann, J., Saile, P., Schmitz, G.H. (2012). Robust multi-objective calibration strategies – possibilities for improving flood forecasting. *Hydrol. Earth Syst. Sci.*, 16, 3579–3606, 2012.
- Krzysztofowicz, R. (2001). The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249 (2-9).
- Krzysztofowicz, R., Davis, D.R. (1984). Toward improving flood forecast-response systems. *Interfaces* 14(3), 1–14.
- Kuczera, G. (1997). Efficient subspace probabilistic parameter optimization for catchment models. *Water Resources Research*. Vol. 33. Nº 1. Pags. 177-185.

- Kuczera, G., Parent, E. (1998). Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology* 211 (1998) 69–85.
- Kuri, A., Galaviz, J. (2002). *Algoritmos Genéticos*. Fondo de Cultura Económica, UNAM, IPN.
- Lamb, R. and Kay, A. (2004). Confidence intervals for a spatially generalized, continuous simulation flood frequency model for Great Britain. *Water Resour. Res.*, 40, W07501.
- Lidén, R., and Harlin, J. (2000). Analysis of conceptual rainfall-runoff modelling performance in different climates. *Journal of Hydrology*, 238(3-4), 231-247.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* 201 (1997) 272-288.
- Linsley, R.K. (1982). *Hidrología para ingenieros*. McGraw-Hill, 1982 - 508 pp.
- Liu, Y. et Gupta, H. (2007). Uncertainty in hydrologic modeling : Toward an integrated data assimilation framework. *Water Resources Research*, 43(7):W07401.
- Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P. (2012). Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrol. Earth Syst. Sci.*, 16, 3863-3887, doi:10.5194/hess-16-3863-2012, 2012.
- Liu, Z-Y, Tan, B-Q, Tao, X, Xie, Z-H. (2008). Application of a Distributed Hydrologic Model to Flood Forecasting in Catchments of Different Conditions. *J. Hydrol. Eng.* 2008.13:378-384.
- Loague, K.M. et Freeze, R.A. (1985). A comparison of rainfall-runoff modeling techniques on small upland catchments. *Water Resources Research*, 21(2), 229-248.

- Luce, C. H., and T. W. Cundy. (1994). Parameter identification for a runoff model for forest roads. *Water Resour. Res.*, 30(4), 1057–1069, 1994.
- Maddaus, W.O. & Eagleson, P.S. (1969). A Distributed Linear Representation of Surface Runoff. Hydromechanics Laboratory Report No. 115. Cambridge, MA: MIT Department of Civil Engineering.
- Madsen, H. (2000). Automatic calibration and Uncertainty Assessment in Rainfall–Runoff Modeling. Joint Conference on Water Resources Engineering and Water Resources Planning & Management, Hyatt Regency Minneapolis, USA.
- Madsen, H. (2000). Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. *Journal of Hydrology* 235 (2000) 276–288.
- Madsen, H., Wilson, G., Ammentorp, H.C. (2002). Comparison of different automated strategies for calibration of rainfall–runoff models. *J. Hydrol.*, 261 (2002), pp. 48–59.
- Manley, R.E. (1975). A hydrological model with physically realistic parameters. *IAHS Publication (115):* 154-161.
- Mathevet, T. (2005). Quels modèles pluie-debit globaux au pas de temps horaire ? Développements empiriques et comparaison de modèles sur un large échantillon de bassins versants. Thèse de Doctorat, à l'Ecole Nationale du Génie Rural, des Eaux et Forêts, 463 pp.
- Mathevet, T., Michel, C., Andréassian, V. & Perrin, C. (2006). A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. *IAHS Red Books Series n°307*, pp. 211-219.
- Mendoza, P.A., McPhee, J., Vargas, X. (2012). Uncertainty in flood forecasting: A distributed modeling approach in a sparse data catchment. *Water Resources Research*, Vol. 48, W09532, doi:10.1029/2011WR011089.
- Mengelkamp, H.-T., Pestel, J., Johnsen, K.-P., Messal, H. (2005). An Integrated Atmospheric/Hydrological Flood Forecasting System for the Odra Watershed.

International conference on innovation, advances and implementation of flood forecasting technology Tromsø, Norway.

Meon, G. (2006). Past and present challenges in flash flood forecasting. En: Proceedings of First International Workshop on Flash Flood Forecasting, San José (Costa Rica). WMO.

Michaud, J. et Sorooshian, S. (1994). Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resources Research*, 30(3), 593-605.

Michel, C. (1983). Que peut-on faire en hydrologie avec un modèle conceptuel à un seul paramètre ? *La Houille Blanche*(1), 39-44.

Michel, C. (1991). *Hydrologie Appliquée aux Petits Bassins Ruraux*. CEMAGREF, Antony.

Moles, C.G., Gutierrez, G., Alonso, A.A., Banga, J. R. (2003). Integrated Process Design And Control Via Global Optimization. A Wastewater Treatment Plant Case Study. *Institution of Chemical Engineers*. Vol 81, Part A, May 2003.

Moore, R.J., Bell, V.A. (2001). Comparison of Rainfall-Runoff Models for Flood Forecasting. Part 1: Literature review of models. R&D Technical Report W241.

Moore, R.J., Bell, V.A. (2001). Comparison of Rainfall-Runoff Models for Flood Forecasting. Part 2: Calibration and evaluation of models. R&D Technical Report W242.

Moore, R.J., Bell, V.A., Carrington, D.S. (2000). Intercomparison of rainfall-runoff models for flood forecasting. En: Lees, M., Walsh, P. *Flood forecasting: What does current research offer the practitioner?*. British Hydrological Society. Occas. Pap. 12.

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., Veith. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *American Society of Agricultural and Biological Engineers*, St. Joseph, MI. Volume 50, Number 3. Pages 885-900.

- Mouelhi, S., Michel, C., Perrin, C. et Andréassian, V. (2006a). Linking stream flow to rainfall at the annual time step: the Manabe bucket model revisited. *Journal of Hydrology* 328, 283-296.
- Mouelhi, S., Michel, C., Perrin, Ch., Andréassian, V. (2006b). Stepwise development of a two-parameter monthly water balance model. *Journal of Hydrology* 318 (2006) 200–214.
- Nalbantis, I. (1995). Use of multiple-time-step information in rainfall-runoff modelling. *Journal of Hydrology* 165 (1995) 135-159.
- Nash J.E., Sutcliffe J.V. (1970). River flow forecasting through conceptual models. Part 1 : A discussion of principles. *Journal of Hydrology*, 27(3), 282-290.
- Nash, J.E. (1957). The form of the instantaneous Unit Hydrograph. *International Association of Scientific Hydrology* 45, 3. pp 114–121.
- Nelder, J.A., Mead, R. (1965). A Simplex method for function minimization. *Computer Journal*, 7: 303-313.
- NHWC - National Hydrologic Warning Council. (2002). Use and benefits of the National Weather Service River and Flood Forecasts. Rapport technique, EASPE, Inc.
- Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E. et Anquetin, S. (2008). Flash flood warning based on rainfall thresholds and soil moisture conditions : An assessment for gauged and ungauged basins. *Journal of Hydrology*, 362(3-4):274-290.
- Oudin L. (2005). Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global. Thèse de Doctorat, l'Ecole Nationale du Génie Rural, des Eaux et Forêts.
- Oudin L., Andréassian V., Lerat J., Michel, C. (2008). Has land cover a significant impact on mean annual streamflow? An international assessment using 1508 catchments. *Journal of Hydrology*, 357 :303-313.

- Oudin L., Michel, C., Anctil F. (2005). Which potential evapotranspiration input for a lumped rainfall-runoff model?. Part 1—Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? *Journal of Hydrology* 303 (2005) 275–289.
- Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andreassian, V., Anctil, F., Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall runoff modelling. *Journal of Hydrology* 303(1-4), 290-306.
- Palacios-Velez, O., Cuevas-Renaud, B. (1992). SHIFT: a distributed runoff model using irregular triangular facets. *J. Hydrol.*, 134, 35–55.
- Palahí M., Pukkala T. (2004). Métodos de optimización Heurística para la resolución de modelos de planificación forestal. *Cuadernos de S.E.C.F.* 18:103-108. I Reunión del grupo de trabajo sobre modelización forestal. Palencia, marzo, 2004.
- Pengue, W.A. (2012). Recursos Naturales y Ciudades Energívoras en América rica Latina. I Congreso Latinoamericano de Ecología Urbana. Universidad Nacional de General Sarmiento. Bs.As.
- Perrin, C. (2000). Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. Thèse de Doctorat, INPG (Grenoble) / Cemagref (Antony), 530 pp.
- Perrin, Ch., Michel, C., Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology* 279 (2003) 275–289.
- Perrin, Ch., Michel, C., Andréassian, V. (2007). Modèles hydrologiques du Génie Rural (GR) Rapport de Cemagref, UR Hydrosystèmes et Bioprocédés.
- Quintero Duque, F. (2011). Estudio de la incertidumbre asociada a la simulación hidrológica. Tesis doctoral. Universidad Politécnica de Cataluña.

- Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B. (2013). Evolutionary Multiobjective Optimization in Water Resources: The Past, Present, and Future. *Advances in Water Resources*.
- Refsgaard, J.C. (1997). Validation and intercomparison of different updating procedures for real-time forecasting. *Nordic Hydrology*, 28, 65-84.
- Refsgaard, J.C. (1997). Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology* 198 (1997) 69–97.
- Reggiani, P., Hassanizadeh, S.M., Sivapalan, M., Gray, W.G. (1999). A unifying framework for watershed thermodynamics: constitutive relationships. *Advances in Water Resources* 23 (1999) 15-39.
- Reggiani, P., Rientjes, T.H.M. (2005). Flux parameterization in the representative elementary watershed approach: Application to a natural basin. *Water Resources Research*, 41, W04013.
- Rodríguez-Iturbe, I., Valdés, J.B. (1979). The geomorphologic structure of hydrologic response. *Water Resources Research*, Vol. 15, No. 6, P. 1409.
- Rojas-Serna, C. (2005). Quelle connaissance hydrométrique minimale pour définir les paramètres d'un modèle pluie-débit?. Thèse de Doctorat, Cemagref (Antony), ENGREF (Paris), France, 319 pp.
- Rosenbrock, H.H. (1960). An Automatic Method for Finding the Greatest or Least Value of a Function. *Computer J.* 3, 175-184.
- Seneraht, S., Ogden, F.L., Downer, C.W., Sharif, H.O. (2000). On the calibration and verification of two-dimensional, distributed, Hortonian, continuous watershed models. *Water Resources Research* 36 (6): 1495-1510.
- Shamseldin, A.Y. (1997). Application of a neural network technique to rainfall-runoff modelling. *Journal of Hydrology* 199 (1997) 272-294.

- Sherali, H.D., Smith, E.P. (1997). A Global Optimization Approach to a Water Distribution Network Design Problem. *Journal of Global Optimization*, V11(2), pp. 107-132.
- Shoemaker, C.A., Regis, R., Mugunthan, P. (2006). Function approximation global optimization algorithms for calibration of expensive simulations with applications to water resources. European Geosciences Union, *Geophysical Research Abstracts*, Vol. 8.
- Shoemaker, CH.A., Regis, R.G., Fleming, R.C. (2007). Watershed calibration using multistart local optimization and evolutionary optimization with radial basis function approximation. *Hydrological Sciences–Journal–des Sciences Hydrologiques*, 52(3) June 2007.
- Siqueira, V.A., Sorribas, M.V., Bravo, J.M., Collischonn, W., Vieira Lisboa, A.M., Villa Trinidad, G.G. (2016). Real-time updating of HEC-RAS model for streamflow forecasting using an optimization algorithm. *RBRH, Porto Alegre*, v. 21, n. 4, p. 855-870, out./dez.
- Soil Conservation Service. (1973). Computer Program for Project Formulation Hydrology. U.S. Department of Agriculture, Technical Release 20.
- Sorooshian, S. and Arfi, F. (1982). Response Surface Parameter Sensitivity Analysis Methods for Postcalibration Studies. *Water Resources Research*, 18(5): 1531-1538, 1982
- Sorooshian, S. and Dracup, J. A. (1980). Stochastic parameter estimation procedures for hydrologic rainfall-runoff models. Correlated and Heteroscedastic error cases. *Water Resour. Res.* 16(2), 430–442.
- Sorooshian, S. and Gupta, K. (1983). Automatic Calibration of Conceptual Rainfall-Runoff Models: The Question of Parameter Observability and Uniqueness. *Water Resources Research*. Vol. 19. Nº1, Pags. 260-268.

- Sorooshian, S. and Gupta, V.K. (1985). The Analysis of Structural Identifiability: Theory and Application to Conceptual Rainfall-Runoff Models. *Water Resources Research*. Vol. 21. Nº 4, Pags. 487-495.
- Sorooshian, S., Duan, Q., Gupta, V.K. (1993). Calibration of rainfall-runoff models: application of global optimization to the Sacramento soil moisture accounting model. *Water Resources Research*, 29 (4), 1185-1194.
- Sorooshian, S., Gupta, V.K. and Fulton, J.L. (1983). Evaluation of Maximum Likelihood Parameter Estimation Techniques for Conceptual Rainfall-Runoff Models: Influence of Calibration. *Water Resources Research*. Vol. 19. Nº1, Pags. 251-259.
- Strahler, A.N. (1952). Hypsometric area-altitude analysis of erosional topography. *Bulletin of the Geological Society of America*, 63, p: 1117-1142.
- Tanakamaru, H. (1995). Parameter estimation for the tank model using global optimization. *Transactions of the JSIDRE*, 178, 103-112.
- Tanakamaru, H., Burges, S.J. (1996). Application of Global Optimization to Parameter Estimation of the Tank Model. *Proceedings of the International Conference on Water Resources and Environmental Research: Towards the 21st Century*, Kyoto, Japan, 29-31 October 1996 (2) 39-46.
- Tangara, M. (2005). Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global. Thèse de Doctorat, Ecole pratique des hautes études de Paris - Sciences de la Vie et de la Terre, Paris, 374 pp.
- Tassi, R., Bravo, J.M, Meller, A., Collischonn, W. (2006). Avaliação do impacto de diferentes funções-objetivo na calibração automática do modelo hidrológico IPH II. XXII Congreso Latinoamericano de Hidráulica, Venezuela, 2006.
- Thielen, J., Bartholmes, J., Ramos, M.-H., de Roo, A. (2009). The European Flood Alert System - Part 1: Concept and development. *Hydrology and Earth System Sciences*, 13(2):125-140.

- Thyer, M., Kuczera, G., Bates, B. (1999). Probabilistic optimization for conceptual rainfall-runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research* Vol. 35, Nº 3 pp. 767-773.
- Tilford, K.A., Sene, K., Chatterton, J.B., Whitlow, C. (2003). Flood Forecasting – Real Time Modelling. R&D Technical Report W5C-013/5/TR.
- Todini, E. (1996). The ARNO rainfall-runoff model. *Journal of Hydrology* 175 (1996) 339-382.
- Troutman, B. M. (1985a). Errors and parameter estimation in precipitation runoff modeling, 1.Theory. *Water Resour. Res.*, 21, 1195–1213.
- Troutman, B. M. (1985a). Errors and parameter estimation in precipitation runoff modeling, 2.Case Study. *Water Resour. Res.*, 21, 1214–1222.
- Tucci, C.E.M. (2006). Gestión de Inundaciones Urbanas UNESCO. Global Water Partnership, GWP-SAMTAC.
- Tucci, C.E.M. (2005). Modelos Hidrológicos. ABRH Editora da UFRGS. Porto Alegre.
- UCAR - University Corporation for Atmospheric Research. (2012). Guía de Referencia para Sistemas de Alerta Temprana de Crecidas Repentinas. National Oceanic and Atmospheric Administration del Departamento de Comercio de los Estados Unidos.
- UNESCO (2003). Agua para todos, agua para la vida. Division of water sciences.
- VAISALA. (2012). Como Minimizar o Impacto de Inundações Extremas. Nota de aplicação. Hobeco Sudamericana Ltda.
- Vélez Upegui, J.J., Francés García, F. (2008). Calibración automática de las condiciones iniciales de humedad para mejorar la predicción de eventos de crecida. *Avances en Recursos Hidráulicos* - Num. 18.

- Vélez, J. I. (2001). Desarrollo de un modelo hidrológico conceptual y distribuido orientado a la simulación de crecidas. Tesis doctoral. Universidad Politécnica de Valencia.
- Vélez, J.J. (2003). Desarrollo de un modelo distribuido de predicción en tiempo real de eventos de crecidas. Tesis Doctoral. Universidad Politécnica de Valencia, Departamento de Ingeniería Hidráulica y Medio Ambiente. 266 P.
- Vélez, J.J., Francés, F. (2004). Sensibility analysis of a conceptual distributed model during flood events. Proceedings I General Assembly European Geosciences Union. Nice France. Ed. Geophysical Research Abstracts.
- Verkade, J.S., Werner, M.G.F. (2011). Estimating the benefits of single value and probability forecasting for flood warning. *Hydrol. Earth Syst. Sci.*, 15, 3751–3765.
- Viveros, E.R.C. ; Barán, B. ; Almirón, M. (1998). Optimización Basada en Colonias de Hormigas con Heurística de Inicialización. In: IX Panel de Informática, 1998, Asunción. EXPOMÁTICA'98.
- Vrugt J, Robinson B, Hyman J. (2009). Self-adaptive multimethod search for global optimization in real-parameter spaces. *IEEE Trans Evol Comput* 2009;13(2):243–59.
- Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W. and Sorooshian, S. (2003). Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resources Research*. Vol. 39. Nº 8. Pags. 5-1 a 5-19.
- Vrugt, J.A., ter Braak, C.J.F., Diks, C.G.H., Schoups, G. (2012). Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. *Adv Water Resour* (2012), <http://dx.doi.org/10.1016/j.advwatres.2012.04.002>
- Wang, Q.J. (1991). The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models. *Water Resources Research*. Vol. 27. Nº 9. Pags. 2467-2471.

- Wang, T. (2001). Global Optimization for Constrained Nonlinear Programming. PhD Thesis in Computer Science, in the Graduate College of the University of Illinois at Urbana-Champaign.
- Whitley, D. (1989). The genitor algorithm and selection pressure: why rank-based allocation of reproductive trials is best. Proceedings of the Third International Conference on Genetic Algorithms, Arlington, VA, Morgan Kaufman Publishers, San Mateo, CA. pp. 116-121.
- WHO. (2018). Prevención del aborto peligroso. Datos y cifras. Pag web. oficial de la Organización Mundial de la Salud. <https://www.who.int/es/news-room/fact-sheets/detail/preventing-unsafe-abortion>.
- WMO. (2011). Manual on Flood Forecasting and Warning. World Meteorological Organization Publications - No. 1072.
- WMO. World Meteorological Office. (2009). Parte IV: Retos de gestión: administración y gobernabilidad. UNESCO.
- WMO. (2014). Atlas of mortality and economic losses from weather, climate and water extremes (1970–2012). WMO-No. 1123. World Meteorological Organization.
- Wood, E. F., Sivapalan, M., Beven, K., & Band, L. (1988). Effects of spatial variability and scale with implications to hydrologic modeling. *Journal of Hydrology*, 102(1-4), 29-47.
- Yang, T. (2015). A Framework to Provide Optimal Management Strategies for California's Reservoirs in Achieving Sustainable Water Supply and High Hydropower Productivity. PhD Thesis. The University of California
- Yapo, P.O., Gupta, H.V. and Sorooshian, S. (1998). Multi-objective global optimization for hydrologic models. *Journal of Hydrology*. Vol. 204. Pags. 83-97.
- Yapo, P.O., Gupta, H.V. and Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *Journal of Hydrology* 181 (1996) 23-48.

Yilmaz, K.K. (2007). Towards Improved Modeling for Hydrologic Predictions in Poorly Gauged Basins. PhD Thesis. The University of Arizona

Zhao, R. J. (1992). The Xinanjiang model applied in China. *Journal of Hydrology*, 135, 371-381.