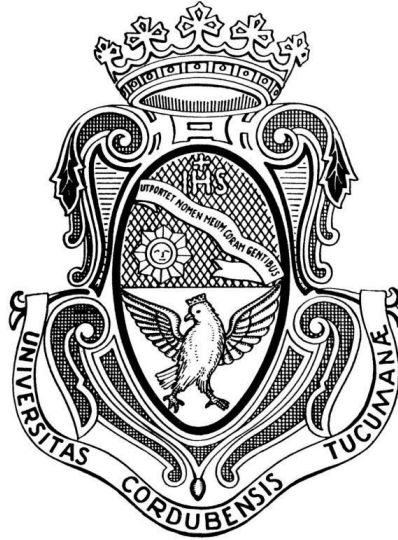


UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE CIENCIAS EXACTAS, FÍSICAS Y NATURALES
INFORME PRÁCTICA PROFESIONAL SUPERVISADA



BASES PARA EL DESARROLLO DE UN MÉTODO DE ADAPTACIÓN DE UN MODELO
HIDROLÓGICO A DIVERSOS HORIZONTES DE PRONÓSTICO

Aplicación a las cuencas de los ríos San Antonio (Córdoba) y Gualaguay (Entre Ríos)

MARÍA DE LOS ÁNGELES JUÁREZ NIHOUL

TUTOR DOCENTE: Facundo José Alonso

TUTOR EXTERNO: Tomás Santiago Vaschalde

Bases para el desarrollo de un método de adaptación de un modelo hidrológico a diversos horizontes de pronóstico

Aplicación a las cuencas de los ríos San Antonio (provincia de Córdoba) y del Gualeguay (Entre Ríos)

Resumen

Un problema recurrente en la hidrología de pronóstico de crecidas consiste en la incertidumbre asociada a la estimación de valores futuros de caudales asociados a distintos horizontes de pronóstico. Una de las fuentes del problema está relacionada con la calidad de los datos de calibración. Otra fuente está vinculada a la incapacidad de los modelos hidrológicos de responder con igual precisión frente a diversos escenarios de modelación. El presente trabajo puede ser dividido en dos grandes etapas. La primera de ellas consiste en la recopilación y procesamiento de la información hidrológica disponible sobre las cuencas de los ríos San Antonio y Gualeguay. Dentro de esta misma etapa se incluye el tratamiento llevado a cabo para depurar los datos, y obtener de esta manera, series de variables de estado confiables y con calidad apropiada para poderlas emplear en la siguiente fase del trabajo.

La segunda etapa se refiere a la calibración de un modelo hidrológico conceptual concentrado, denominado Génie Rural (GR) para distintos escenarios y la comparación entre ellas empleando como herramienta la optimización multi-objetivo. Los resultados indican que, con las herramientas empleadas no fue posible establecer un método robusto de pronóstico capaz de predecir con aceptable precisión en los horizontes practicados, pero sin embargo se infirió una cierta tendencia a converger en un único juego de parámetros capaz de alternar confiadamente entre los diversos escenarios de pronóstico con semejantes performances de las funciones objetivo obtenidas.

Basis for the development of a method of adapting of a hydrological model to various forecast horizons

Application to the San Antonio river basin (Córdoba) and Gualeguay (Entre Ríos)

Abstract

A recurring problem in hydrology of flood forecasting is the uncertainty associated with estimating future values of flow associated with different forecast horizons. One of the source of the problem is related to the quality of the calibration data. Another source is linked to the inability of hydrological models to respond with equal precision to various modeling scenarios. This essay can be divided into two main stages. The first one consist in the collection and processing of the hydrological information available of the San Antonio river basins, and Gualeguay stream. Within of this stage it is included the tratament carried out to deperate the data, and thereby obtain reliable series of state variables with appropriate quality to use in the next phase of essay.

The second stage refers to the calibration of a concentrated conceptual hydrological model, called Génie Rural (GR) for different scenarios and comparison between them using as a tool the multi-objective optimization. The results indicate that, with the tools used it was not possible to establish a robust forecasting method that can predict with reasonable accuracy the practiced horizons, but nevertheless was observed a tendency to converge in a single set of parameters able to confidently switch between the various forecast scenarios with similar performances of the obtained objective functions.

AGRADECIMIENTOS

En primer lugar quiero agradecerle a mi familia quienes son mi pilar y guía en cada camino emprendido, quienes me enseñaron con su ejemplo que todo lo que uno se propone, con esfuerzo y dedicación se alcanza. En especial por su gran confianza, apoyo y optimismo que me han brindado sobre todo en aquellos momentos de mayor cansancio y desaliento.

A mis amigas de la secundaria que siempre estuvieron presentes brindándome su amistad y apoyo.

A mis compañeros de facultad, especialmente a Consu, Sol, Agu, Augusto, Julián, Meli, Nacho A, Nacho G, Noe y Pelado, quienes se convirtieron en grandes y valiosos amigos. No solo fueron de un gran aporte en el ámbito académico donde compartimos nervios, ansiedad y alegría, sino también en lo personal mediante el apoyo y compañía en los diferentes aspectos de la vida. Además no puedo dejar de agradecer a cada una de las persona con la que me crucé durante este camino ya que entre todos nos proporcionamos ayuda y por sobre todas las cosas optimismo y alegría, haciendo que estos 5 años de carrera hayan sido una experiencia muy agradable.

A mi novio y compañero Álvaro por ser mi mayor sustento en todo momento.

A mi director de Prácticas Profesional Supervisadas, Mg. Ing. Facundo José Alonso, por su dedicación y paciencia al guiarme a lo largo de este camino.

Al Instituto Superior de Estudios Ambientales (ISEA) por brindarnos el lugar de trabajo y reuniones.

A todas aquellas personas que facilitaron el acceso a la información necesaria para el desarrollo del presente trabajo, Mg. Ing. Facundo José Alonso, Ing. Laura Colladon (CIRSA-INA), Dr. Ing. César García, Ing. Carlos Catalini, Ing. María José Martín.

INDICE

1. INTRODUCCION.....	1-1
2. OBJETIVOS.....	2-4
2.1. OBJETIVO GENERAL.....	2-4
2.2. OBJETIVOS PARTICULARES	2-4
3. REVISIÓN DE ANTECEDENTES	3-5
3.1. MODELOS HIDROLOGICOS	3-5
3.2. SISTEMAS, CONCEPTOS Y CLASIFICACION.....	3-8
3.3. EL SISTEMA HIDROLÓGICO.....	3-9
3.4. EL PROCESO DE TRASNFORMACIÓN LLUVIA-CAUDAL	3-11
3.5. MODELOS, CONCEPTOS Y CLASIFICACION	3-13
3.6. EL MODELO LLUVIA-CAUDAL	3-18
3.7. OPTIMIZACIÓN – CONCEPTOS GENERALES.....	3-20
3.7.1. METODOS DE OPTIMIZACION	3-21
3.7.2. GENERACION DE NUMEROS ALEATORIOS	3-22
3.8. ALGORITMOS GENETICOS	3-25
3.8.1. CONCEPTOS BASICO DE BIOLOGIA.....	3-26
3.8.2. LA TECNICA DE LOS ALGORITMOS GENETICOS	3-31
3.9. TÉCNICAS DE CALIBRACIÓN.....	3-35
3.9.1. CLASIFICACION DE LAS TECNICAS DE CALIBRACION	3-36
3.10. CALIBRACIÓN AUTOMATICA DE MODELOS P-Q.....	3-45
3.10.1. LA FUNCION OBJETIVO	3-45
3.10.2. ALGORITMOS DE OPTIMIZACION PARA MODELOS P-Q.....	3-58
3.10.3. CRITERIO DE PARADA	3-61
3.10.4. DATOS DE CALIBRACION	3-63
4. METODOLOGIA	4-67
4.1. DESCRIPCION DEL MODELO HIDROLOGICO GR	4-67
4.1.1. OBJETIVOS DE SU DESARROLLO	4-67
4.1.2. PRINCIPALES CARACTERÍSTICAS	4-67
4.1.3. ESTRUCTURA DEL MODELO.....	4-68

4.2.	EL ALGORITMO MOCOM-UA.....	4-77
4.2.1.	EVOLUCION COMPETITIVA COMPLEJA	4-79
4.2.2.	ALGORITMO DE NELDER Y MEAD.....	4-82
4.2.3.	PASOS DEL METODO	4-85
5.	REVISIÓN Y ACTUALIZACIÓN DE LOS DATOS DEL RIO SAN ANTONIO.....	5-91
5.1.	DESCRIPCIÓN DE LAS HERRAMIENTAS DE EXCEL EMPLEADAS EN EL PROCESAMIENTO DE LAS SERIES HIDROLÓGICAS.....	5-91
5.2.	DESCRIPCIÓN DE LA CUENCA	5-93
5.3.	PRECIPITACIÓN.....	5-98
5.3.1	INFORMACIÓN DISPONIBLE	5-98
1.14.1.....		104
5.3.2.	CONVERSION DE LA INFORMACION RECIBIDA A SERIES HORARIAS	104
5.3.3.	COMPARACIÓN DE LAS SERIES OBTENIDAS CON LAS DISPONIBLES.....	107
5.3.4.	CONFORMACION DE LA SERIE DE PRECIPTACIÓN MEDIA AREAL DE LA CUENCA... ..	109
5.4.	CURVA H-Q.....	5-114
5.4.1.	FUENTES DE INFORMACIÓN.....	5-114
5.4.2.	DESCRIPCIÓN DE LOS AFOROS	5-115
5.5.	NIVELES Y CAUDALES.....	5-116
5.5.1.	INFORMACIÓN DISPONIBLES DE NIVELES	5-116
5.5.2.	CONVERSIÓN DE LA INFORMACIÓN A SERIES DE NIVELES HORARIAS.....	5-117
5.5.3.	COMPARACIÓN DE LAS SERIES DE NIVELES	5-118
5.5.4.	CONFORMACION DE LA SERIE DE CAUDALES HORARIA	5-120
5.6.	ANALISIS P-Q	5-120
5.6.1.	CONFORMACIÓN DE SERIES P-Q.....	5-120
5.7.	EVAPOTRANSPIRACIÓN POTENCIAL.....	5-124
5.7.1.	METODO DE LUDOVIC OUDIN.....	5-125
5.7.2.	PROCESAMIENTO DE LA INFORMACIÓN	5-126
6.	RECOPILACIÓN Y REVISIÓN DE LOS DATOS DEL RÍO GUALEGUAY.....	5-129
6.1.	ANALISIS P-Q	6.1-129
6.1.1.	CONFORMACIÓN DE SERIES P-Q.....	6.1-129
7.	ANALISIS ESTADÍSTICO	7-132

7.1.	RESULTADOS CUENCA RÍO SAN ANTONIO	7-133
7.1.1.	PRECIPITACION.....	7-133
7.1.2.	CAUDAL	7-134
7.2.	RESULTADOS CUENCA GUALEGUAY.....	7-134
7.2.1.	PRECIPITACION.....	7-134
7.2.2.	CAUDAL	7-135
8.	APLICACIÓN Y RESULTADOS.....	8-136
8.1.	DESCRIPCION GENERAL- ARCHIVOS PARA LA MODELACIÓN.....	8-136
8.2.	CUENCA DEL RÍO SAN ANTONIO	8-144
8.2.1.	CALIBRACIÓN.....	8-144
8.2.2.	MODELACIÓN EN MODO PRONÓSTICO	8-148
8.3.	CUENCA DEL GUALEGUAY	8-157
8.3.1.	CALIBRACIÓN.....	8-157
8.3.2.	MODELACIÓN EN MODO PRONÓSTICO	8-161
9.	CONCLUSIÓN.....	9-168
9.1.	CONCLUSIONES SOBRE RECOPIACIÓN Y REVISIÓN DE INFORMACIÓN.....	9-168
9.2.	CONCLUSIONES SOBRE LA CALIBRACIÓN.....	9-169
10.	REFERENCIAS BIBLIOGRÁFICAS	10-171
11.	ANEXO	11-173
11.1.	ANÁLISIS P – Q.....	11-173

LISTA DE TABLAS

Tabla 4.1. Parámetros modelo GR4J, (Perrin et al. 2003).	4-75
Tabla 5.1. Pluviómetros de alta frecuencia instaladas en la cuenca del río San Antonio, denominación, emplazamiento y altitud relativa (Fuente: INA-CIRSA).	5-95
Tabla 5.2. Descripción de la Información pluviométrica faltante en el periodo de registro por año y por cada estación.	5-99
Tabla 5.3. Periodos sin información de la serie de Cesar Garcia.	5-102
Tabla 5.4. Ejemplo de presentación de la Información organizada por Bustos (2014) ..	5-103
Tabla 5.5. Ilustración del modo de empleo de la función "SI".	5-105
Tabla 5.6. Ilustración conformación serie horaria de precipitación.	5-107
Tabla 5.7. Precipitación media areal sin rellenar series.	5-113
Tabla 5.8. Precipitación media areal con relleno de series.	5-113
Tabla 5.9. Periodos de superposición de las series de niveles.	5-118
Tabla 5.10. Radiación extraterrestre.	5-126
Tabla 7.1. Errores en la aceptación de las hipótesis.	7-133

LISTA DE FIGURAS

Figura 3.1. El ciclo hidrológico (CONICYT, 2007).	3-10
Figura 3.2. Modelo hidrológico de transformación lluvia-caudal (HEC-HMS, 2000).	3-12
Figura 3.3. Etapas de la meiosis (Kuri y Galaviz, 2002).	3-30
Figura 3.4. Esquema de trabajo de la enzima ADN polimerasa (Kuti y Galaviz, 2002).	3-31
Figura 3.5. Esquema de la Regio de Pareto para dos funciones objetivo.	3-40
Figura 3.6. Proyeccion plana de la superficie de respuesta de una FO.....	3-41
Figura 4.1. Estructura del modelo GR4J (Perrin et al., 2003).	4-70
Figura 4.2. Modelo GR4J modo pronóstico.....	4-76
Figura 4.3. Zona de validez de los parámetros del modelo, (Alonso, 2008).	4-86
Figura 4.4. Puntos válidos y no válidos (Alonso, 2008).	4-86
Figura 5.1. Ubicación de la cuenca del rio San Antonio (CIRSA, 2003)	5-93
Figura 5.2. Red telemétrica de la del Río San Antonio.	5-96
Figura 5.3. Perfil de la cuenca y tributarios del Río San Antonio (CIRSA, 2003)	5-97
Figura 5.4. Información proveniente directamente de la red telemétrica del INA.	5-101
Figura 5.5. Información disponible con paso de tiempo horario.	5-102
Figura 5.6. Polígono de Thiessen N° 93.	5-111
Figura 5.7. Polígono de thiessen N° 123.....	5-111
Figura 5.8. Polígono de thiessen N°65.....	5-112
Figura 5.9. Datos de aforos realizados.	5-114
Figura 5.10. Curva H-Q confeccionada por M. J. Martín.	5-115
Figura 8.1. Carpetas y archivos necesarios para la modelación.....	8-137
Figura 8.2. Archivo de datos de la cuenca del río San Antonio.....	8-138
Figura 8.3. Archivos de salida.	8-138
Figura 8.4. Parámetros e hidrogramas simulados del individuo 30.....	8-139
Figura 8.5. Parámetros e hidrogramas pronosticados del individuo 30.	8-139
Figura 8.6. Archivo de entrada para calibración del tramo 01 del río San Antonio.....	8-140

LISTA DE GRÁFICOS

Gráfico 5.1.	5-121
Gráfico 5.2.	5-121
Gráfico 5.3.	5-122
Gráfico 5.4.	5-122
Gráfico 5.5.	5-123
Gráfico 6.1.1.	6.1-130
Gráfico 6.1.2.	6.1-130
Gráfico 6.1.3.	6.1-131
Gráfico 8.1. FO1 vs FO2. Tramo 01. Iteración 2001.	8-146
Gráfico 8.2. Región de pareto del tramo 02. Iteración 289.	8-146
Gráfico 8.3. Región de pareto del tramo 03. Iteración 603.	8-147
Gráfico 8.4. Región de pareto del tramo 04. Iteración 183.	8-147
Gráfico 8.5. Hidrograma individuo 1. Tramo 02.	8-149
Gráfico 8.6. Hidrograma del individuo 37. Tramo 02.	8-150
Gráfico 8.7. Hidrograma del individuo 1. Tramo 02.	8-150
Gráfico 8.8. Hidrograma del individuo 4. Tramo 03.	8-152
Gráfico 8.9. Hidrograma del individuo 13. Tramo 03.	8-153
Gráfico 8.10. Hidrograma del individuo 04. Tramo 03.	8-153
Gráfico 8.11. Hidrograma del individuo 13. Tramo 03.	8-154
Gráfico 8.12. Hidrograma del individuo 6. Tramo 04.	8-155
Gráfico 8.13. Hidrograma del individuo 36. Tramo 04.	8-156
Gráfico 8.14. FO1 vs FO2. Tramo 01. Iteración 15001.	8-158
Gráfico 8.15. FO1 vs FO2. Tramo 02. Iteración 30001.	8-159
Gráfico 8.16. Evolución de la población. Tramo 01.	8-160
Gráfico 8.17. Evolución de la población. Tramo 01.	8-161
Gráfico 8.18. Hidrograma del individuo 1. Tramo 01.	8-162

Gráfico 8.19. Hidrograma del individuo 4. Tramo 01.	8-163
Gráfico 8.20. Hidrograma del individuo 7. Tramo 02.	8-164
Gráfico 8.21. Hidrograma del individuo 11. Tramo 02.	8-165
Gráfico 8.22. Hidrograma del individuo 7. Tramo 02.	8-165
Gráfico 8.23. Hidrograma del individuo 11. Tamo 02.	8-166

1. INTRODUCCION

A pesar de que el agua constituye el recurso más elemental y necesario para la vida, todavía existen muchas incertidumbres en relación a los procesos que gobiernan su movimiento dentro de la hidrosfera. El sistema hidrológico global es sumamente complejo, debido a los múltiples ambientes en que el agua interacciona, pasando por los más diversos estados.

El hombre ha enfrentado este complejo problema del movimiento del agua en el planeta a través de modelos simplificados de la realidad. Los modelos matemáticos son una de las principales herramientas con que cuenta el hidrólogo para interpretar y simular el comportamiento de una cuenca hidrográfica.

La simulación hidrológica tiene aplicación en las tres principales funciones de la hidrología: la predicción, el proyecto y la planificación. En cuanto a la predicción, una de las principales aplicaciones se da en el campo del pronóstico de crecidas a tiempo real. En proyecto, los modelos se utilizan para estimar caudales de diseño de obras hidráulica. Mientras que en planificación, se utilizan para realizar una estimación de la disponibilidad del recurso hídrico a mediano y largo plazo, análisis de la variabilidad hidrológica o estudios sobre los cambios del uso de suelo, entre otras.

La bibliografía ha sido pródiga en publicaciones de modelos matemáticos aplicados a la hidrología. La filosofía de construcción de los modelos hidrológicos es variada, y su aplicabilidad está limitada a las condiciones para las cuales fue desarrollado, existiendo una amplia gama que varía en complejidad. La estructura del modelo matemático comprende un conjunto de valores constantes, denominados parámetros, que controlan los diversos procesos del modelo. Los valores de estos parámetros, varían dependiendo de la cuenca hidrográfica de aplicación.

Una de las etapas más críticas en la aplicación de modelos hidrológicos está asociada al proceso de calibración, es decir, la elección del juego de valores para los parámetros que mejor representen las condiciones locales de la cuenca hidrográfica.

Tradicionalmente, la calibración de un modelo matemático se ha venido realizando en forma manual. Pero en las últimas décadas se han desarrollado, a nivel mundial, nuevos métodos de calibración automática de modelos hidrológicos, heredados del campo de la optimización. En esencia, la calibración automática consiste en minimizar una expresión matemática, denominada función objetivo, que cuantifica el error de simulación.

Desde el advenimiento de las computadoras, los métodos de calibración automática han ido progresando. Consisten básicamente en algoritmos desarrollados para encontrar el valor óptimo de la función objetivo.

Ahora bien, el éxito de la calibración automática no sólo depende del algoritmo empleado, y de la función objetivo elegida, sino también de la performance del modelo hidrológico empleado, y de la calidad de los datos disponibles.

Los últimos desarrollos en materia de optimización han tomado prestado conceptos de la biología para incrementar la variabilidad y la capacidad de búsqueda, originando los denominados métodos heurísticos. Entre estos métodos, se destaca el método de los algoritmos genéticos (Holland, 1975) inspirados en los mecanismos de optimización de las especies, sobre la base de la selección natural y supervivencia del individuo más apto. Estos métodos han demostrado notable eficiencia y representan un considerable ahorro de consumo computacional.

En este informe se presenta, en primer lugar una recopilación y procesamiento de datos hidrológicos de la cuenca hidrográfica del río San Antonio y del Río Gualaguay con el propósito de actualizar las series de lluvia, caudal y evapotranspiración obtenidos en trabajos antecedentes.

La información recopilada fue procesada y analizada con el propósito de detectar valores anormales como producto de diversas fuentes de error, a los fines de conformar series confiables.

Posteriormente, con las series obtenidas se llevó a cabo la optimización del modelo hidrológico de transformación lluvia-caudal Génie Rural para dos horizontes de tiempo, 6

y 12 horas mediante el método de calibración global multi-objetivo, basado en algoritmos genéticos, MOCOM-UA (Yapo et al., 1998).

2. OBJETIVOS

2.1. OBJETIVO GENERAL

Estudiar la performance de un modelo hidrológico para adaptarse a diversos horizontes de pronóstico, durante la operación de una cuenca monitoreada.

2.2. OBJETIVOS PARTICULARES

- Recopilar información histórica de lluvia, temperatura y niveles de la cuenca de los ríos San Antonio y Gualeguay.
- Conformar una serie horaria actualizada de precipitación, caudales y evapotranspiración para la cuenca del río San Antonio.
- Analizar la confiabilidad de las series de datos obtenidas.
- Realizar la calibración multi-objetivo del modelo Génie Rural para la cuenca del río San Antonio y del río Gualeguay.
- Evaluar la performance de la modelación a partir de las funciones objetivo.
- Estudiar la capacidad del modelo de adaptarse a más de un horizonte de pronóstico basado en la elección del juego de parámetros adecuado.

3. REVISIÓN DE ANTECEDENTES

Durante esta práctica se llevó a cabo un importante proceso de adecuación de datos para luego efectuar la calibración automática de un modelo hidrológico empírico de transformación lluvia-caudal en dos cuencas argentinas. En consecuencia, como parte de la revisión del estado del arte, se aborda inicialmente una introducción a la teoría de modelos matemáticos.

El proceso de calibración automática de un modelo hidrológico es un problema de la optimización de modelos. Existe una gran variedad de métodos de optimización que se han ido desarrollando para distintos problemas a optimizar. La Investigación de Operaciones ha incursionado en el campo de la optimización para un gran número de casos particulares. Por esta razón, se presenta en esta tesis una introducción de los conceptos básicos que gobiernan el problema de la optimización de modelos.

Entre los métodos desarrollados para resolver problemas de optimización uno de los que mayormente se adapta a los problemas de optimización de modelos hidrológicos son los Algoritmos Genéticos (AG). Dado que se trabajó con un algoritmo de calibración automática basado en los AG, resulta necesario presentar una revisión general de los conceptos asociados a los AG.

En este capítulo se encuentran numerosos extractos de la tesis de maestría del Mg. Ing. Facundo José Alonso (2008).

3.1. MODELOS HIDROLOGICOS

Un sistema es un conjunto de elementos que interactúan entre sí. Un modelo es una representación simplificada de un sistema real. Esta representación puede hacerse en forma física (modelo físico), o bien, a través de una expresión matemática (modelo matemático) de mayor o menor complejidad. De acuerdo a la Sociedad de Investigación

de Operaciones de Gran Bretaña y de Estados Unidos (Zoraida Omaña, 2006), la esencia de la Investigación de Operaciones es el uso de modelos.

Independientemente del origen o estructura del modelo, se entiende que el mismo funciona correctamente si dado un conjunto de variables de entrada, genera la misma salida que el sistema al cual representa. El proceso por el cual un modelo imita al sistema se denomina simulación. Se dice que el modelo simula un proceso que ocurre en el sistema. Un modelo captura características selectas de un sistema, proceso o realidad, y luego las combina en una representación abstracta del original.

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. Por el contrario, las variables del sistema en general son cantidades que pueden ser medidas, al menos en forma aproximada, y además, como su nombre lo indica, varían durante el proceso. Existen tres tipos de variables: las variables de entrada, las de salida y las de estado. Estas últimas, definen el estado del sistema (Bertoni, 2003).

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).

El error cuantitativo de la simulación puede medirse pues con una función que dependa de la diferencia numérica entre el hidrograma simulado y el observado. A esta función se la denomina función objetivo (FO). De la mano de los objetivos, aparecen las restricciones, que consisten básicamente en las limitaciones asociadas a la aplicación misma del modelo, normalmente representado por un intervalo de validez de las diversas variables y parámetros que intervienen en el problema. Sin embargo, existen ciertas variables que no están sujetas a restricciones.

En síntesis, las FO son expresiones matemáticas que cuantifican el error de simulación de un modelo, como la diferencia entre los valores medidos y los simulados por el modelo. El proceso de calibración de un modelo consiste en estimar los valores de los parámetros que minimiza el valor de la función objetivo, dentro del límite definido por las restricciones. Debido al papel protagónico que desempeña la función objetivo en el

proceso de calibración automática de modelos matemáticos, más adelante se presenta una sección dedicada a las funciones objetivo.

3.2. SISTEMAS, CONCEPTOS Y CLASIFICACION

Se 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 (Bertoni, 2003).

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

y 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 $m(t)$.
- 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.

3.3. EL SISTEMA HIDROLÓGICO

De los innumerables sistemas que ofrece la naturaleza, en particular nos interesa el denominado: ciclo hidrológico (Figura 3.1.), 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 casual. 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.

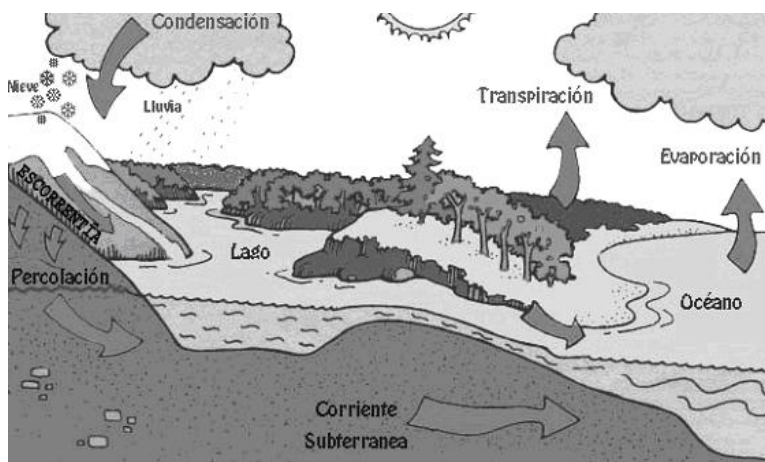


Figura 3.1. El ciclo hidrológico (CONICYT, 2007).

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. 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.

3.4. 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 la 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 3.2. ilustra el proceso.

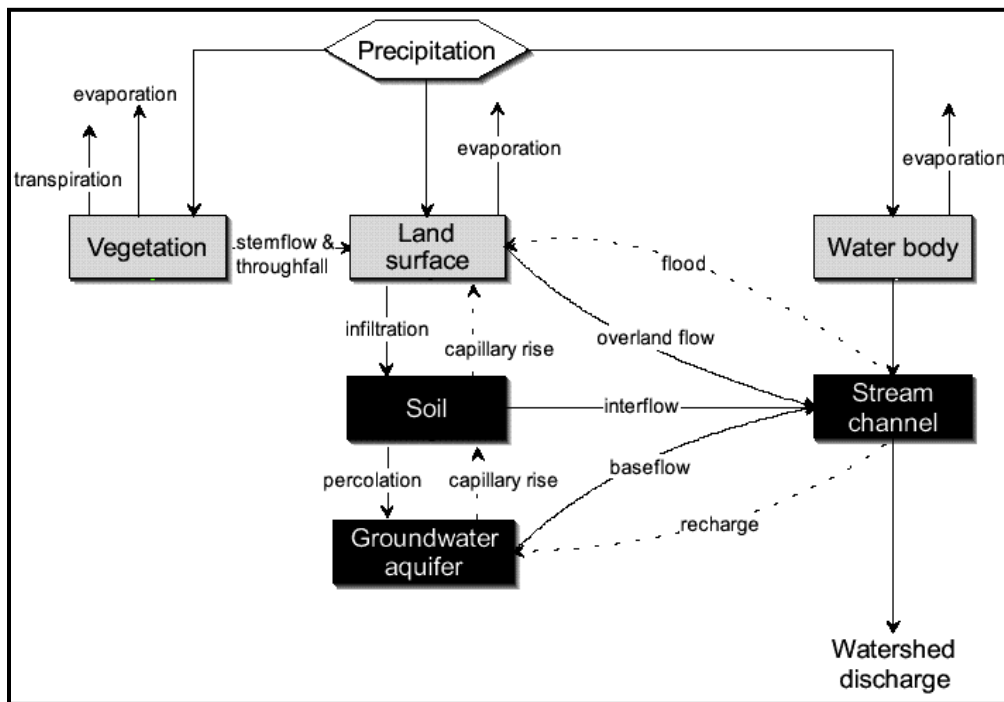


Figura 3.2. Modelo hidrológico de transformación lluvia-caudal (HEC-HMS, 2000).

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 sub-superficial; 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 sub-superficial 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 sub-superficial o subterráneo. A la inversa, el escurrimiento sub-superficial 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.

3.5. MODELOS, CONCEPTOS Y CLASIFICACION

Un modelo es una representación simplificada de un sistema real. Existen básicamente tres tipos de modelos, según su naturaleza. El *modelo físico*, que es una maqueta a escala que se construye a los fines de representar físicamente el fenómeno o proceso. El *modelo analógico*, es la representación de un sistema físico (prototipo), por un sistema físico

análogo, de modo que su comportamiento simule de forma aproximada (o exacta) el del prototipo. Un ejemplo de modelos analógico típico de la hidrología, es la simulación del flujo de agua en un cauce (o a través de un medio poroso), mediante el flujo de una corriente eléctrica en un circuito compuesto por resistencias y condensadores. Finalmente, el *modelo matemático*, consiste en un sistema de ecuaciones o formulaciones matemáticas que representan el sistema en estudio. Estos últimos, se han convertido en una poderosa herramienta de trabajo, fundamentalmente desde el vertiginoso desarrollo computacional que se comenzó a verificar en las últimas décadas.

Algunas de las ventajas principales asociadas al uso de modelos matemáticos son:

- 1) Permite estudiar y experimentar en las complejas relaciones e interacciones que existen y ocurren en el interior de un sistema real.
- 2) Permite experimentar sobre un sistema que puede no existir en la realidad (de otra manera sería imposible) y en este caso asiste a la decisión de construirlo o no.
- 3) Se ahorra tiempo y dinero, y se evitan riesgos.
- 4) Se logra independencia de la duración real del evento. Por ejemplo se puede simular procesos muy largos como muy cortos. Por ejemplo, se simula en pocos minutos una crecida ocurrida sobre una cuenca, que en la realidad podría tardar días en manifestarse.
- 5) Se puede estudiar cómo reacciona el sistema al introducir variantes, sin correr riesgos (estudio de sensibilidad).
- 6) Se puede aplicar para verificar soluciones analíticas (por ejemplo, mediante la metodología denominada de Montecarlo). Ofrece mayor facilidad para estudiar sistemas dinámicos y no determinísticos.

Con respecto a este tema, cabe aclarar que el análisis cuantitativo que resulta de la aplicación de un modelo matemático, no reemplaza, sino que complementa al cualitativo, producto del sentido común y la experiencia del modelador. Los modelos cuantitativos auxilian, pero no sustituyen a los encargados de tomar decisiones.

Un modelo matemático representa a un sistema de modo simplificado. En dicha simplificación que se hace de la realidad, se conservan algunas características propias del sistema, y evidentemente, otras, se pierden. Estas características pueden ser, por ejemplo, la linealidad, su propiedad de ser concentrado o distribuido, etc. En base a esto, podría pensarse para los modelos, una clasificación semejante a la que se realizó con los sistemas.

Sin embargo, no necesariamente un sistema de ciertas características es representado con un modelo de las mismas características. En efecto, por definición un modelo matemático es más simple que el sistema real al cual representa. La función que cumple el modelo, es convertirse en una herramienta que represente al sistema real pero del modo más simple posible. Esta simplificación, que facilita la modelación, se realiza a expensas de sacrificar algunas propiedades del sistema. Así, por ejemplo, un sistema no lineal puede representarse mediante un modelo lineal. Tal es el típico caso del hidrograma unitario. El sistema genera una transformación lluvia-caudal no lineal, sin embargo el modelo del hidrograma unitario si lo es. Otro de los muchos ejemplos que podrían citarse es el modelo de las ecuaciones de Saint Venant, que representa el flujo a superficie libre en una formulación matemática unidimensional, aunque el flujo a superficie libre, naturalmente, es tridimensional. Por último, resulta ilustrativo citar el caso de la lluvia. En el sistema real, la lluvia cae en forma continua, sin embargo, en un modelo, la lluvia es incorporada como una variable de entrada que se computa en forma discreta, ya que la misma se mide a intervalos de tiempo mediante los pluviógrafos.

De acuerdo a la función que cumple el modelo, y a la filosofía que subyace el desarrollo del mismo, el grupo EURYDICE 92 (Chocat, 1997), los clasifica del siguiente modo:

a. Según el propósito del modelo:

- Modelos Cognitivos
- Modelos Previsionales
- Modelos de toma de decisión
- Modelos normativos

b. Según la concepción planteada en su formulación:

- Modelos axiomático o heurístico
- Modelos determinístico o estadístico (o estocástico)
- Modelos reduccionista u holístico
- Modelos empírico o conceptuales
- Modelos global o distribuido

Los *Modelos Cognitivos*, tienen por objeto dar una explicación del funcionamiento o estado del sistema. Dentro de los modelos cognitivos, pueden diferenciarse dos grupos, que son los modelos *explicativos* y los *descriptivos*, de acuerdo a que permitan o no, respectivamente, comprender la influencia de las variables de entrada sobre las de salida. Por ejemplo, en la representación de las pérdidas por infiltración, el modelo de Richard, es mucho más explicativo que el modelo de Green y Ampt o que el de Horton. Pero, a su vez, estos últimos son más explicativos que el modelo del “índice Phi”. Es importante destacar que el mayor número de los modelos empleados en hidrología son de tipo Cognitivo.

Los *Modelos Provisionales*, tienen por objeto prever y simular el estado de funcionamiento de un sistema. Los *Modelos de Toma de Decisión*, se emplean para determinar las variables de entrada (o de concepción) a proveer al modelo para obtener las variables de salida conocidas y previamente impuestas. Los *Modelos Normativos*, tienen por objeto establecer las normas que relacionan la entrada y la salida. Formulan las relaciones que deben existir para que, con cierta entrada, se obtenga una salida también conocida. Se diferencian de los sistemas cognitivos en que estos modelos no analizan el proceso.

Los *modelos axiomáticos* son aquellos que se desarrollan partiendo rigurosamente de los principios y reglas que dominan el sistema, basados en hipótesis sobre el mismo. Los *modelos heurísticos* plantean principios que reflejan parcialmente el comportamiento del sistema que luego articulan en forma aproximada para obtener el modelo.

Un *modelo determinístico* es aquel que produce siempre una misma salida para una cierta entrada. Es decir que no presentan aleatoriedad sus variables (Chow, 1994). Un *modelo estadístico*, es aquel en el cual las salidas son aleatorias. Cuando en un modelo la probabilidad de ocurrencia de sus variables es tomada en cuenta, o bien, el concepto de probabilidad es introducido en el modelo mismo, se dice que el modelo es de naturaleza estadística (Clarke, 1973). Cuando el proceso evoluciona en el espacio y en el tiempo de un modo que es parcialmente predecible y parcialmente aleatorio se denomina *estocástico* (Chow, 1994), entonces el modelo que lo representa manifiesta una estructura parcialmente determinística y parcialmente aleatoria y se denomina también estocástico.

La mayoría de los procesos hidrológicos son de tipo estocásticos, y dan nacimiento a una rama importante de la hidrología denominada Hidrología Estocástica. Cuando una variable de salida de un modelo hidrológico tiene un grado de aleatoriedad, un problema de rutina para la hidrología estocástica es identificar en que grado influye su componente aleatorio y en qué grado su determinístico (Chow, 1994). Si la influencia del componente probabilístico no es significativa, es apropiado un modelo de tipo determinístico; pero si tiene importancia significativa, será necesario un modelo de tipo estocástico, pues de lo contrario, la salida real será muy distinta a la salida única que arrojaría el modelo determinístico.

Es importante destacar que, un modelo de tipo determinístico, puede tener variables de entrada o salida de tipo aleatorias, y que un modelo estocástico, puede tener variables determinísticas. Un ejemplo de esto, es un modelo determinístico que arroja para cada valor de entrada P (lluvia), su correspondiente valor de salida Q (caudal). Aún cuando P y Q son variables aleatorias, el modelo de transformación lluvia-caudal es determinístico.

El *enfoque Reduccionista* (también llamado microscópico, mecanicista o de caja blanca), incluye aquellos modelos que se desarrollan a partir de descomponer previamente el sistema real en muchos microprocesos, y luego se formula el modelo teniéndolos a todos en cuenta. En esencia se trata de un proceso de diferenciación e integración. Por ejemplo, las ecuaciones de Saint Venant. Mientras que el *enfoque Holístico* (o macroscópico, o de caja gris, o conceptual), consideran el sistema en su conjunto, y se desarrolla el modelo basado en el comportamiento global del mismo. Por ejemplo, el modelo de Muskingum.

Los *modelos Empíricos* (o de caja negra) se basan en la experimentación. Pueden no guardar relación alguna con los fenómenos físicos y conceptuales del proceso analizado. El único criterio de formulación consiste en reducir las diferencias entre las variables observadas y las calculadas. Los *modelos conceptuales* (o de enfoque teórico), en cambio, se apoyan sobre el concepto de abstracción y el de razonamiento. En este caso el modelo es construido a priori, sobre un análisis teórico, y la experimentación posterior contribuye a validar el modelo.

Los *modelos concentrados* (o globales, o puntuales, o de orden cero), no consideran la variación espacial de los parámetros y las variables. Por ejemplo el método de la Curva Número de infiltración del SCS. Los *modelos distribuidos* (o de malla, o matriciales, o de discretización espacial) consideran la heterogeneidad del cuadro físico modelado. Pero siempre es necesario realizar alguna discretización sobre el dominio para poder analizar la variación dentro del campo de estudio. Esta discretización acota su concepción de distribuidos.

3.6. EL MODELO LLUVIA-CAUDAL

Tal como ha sido indicado anteriormente, el proceso de transformación lluvia caudal, es sumamente complejo y requiere de una abstracción a través de un modelo que simplifique suficientemente el sistema como para poder simular conceptualmente sus partes.

Dentro del proceso de transformación lluvia-caudal pueden identificarse dos procesos claramente independientes desde el punto de vista conceptual, como son el proceso de producción y el de transferencia. El primero, está asociado al mecanismo por el cual sólo una parte de la precipitación llega en forma rápida a la salida de la cuenca, ya que el resto queda retenido, se evapora o se infiltra y llega a la salida de la cuenca, pero ya como flujo subterráneo. El segundo, está relacionado con el recorrido que el agua realiza desde los distintos puntos de la cuenca hasta la salida de la misma.

En síntesis, el componente de producción en un modelo de transformación lluvia-caudal se encarga de estimar el porcentaje del volumen total caído sobre la cuenca como precipitación que se convierte en caudal a la salida de la misma. Por su parte, el componente de transferencia se ocupa de establecer la forma en la cual llega repartido en el tiempo ese volumen a la salida de la cuenca, obteniendo así, el hidrograma de escurrimiento directo en la sección de salida de la cuenca.

En cuencas intermedias y grandes, para realizar el análisis del proceso de transformación lluvia caudal, normalmente, se subdividen en sub-cuencas. Entonces, se determina el hidrograma de salida de cada sub-cuenca y luego, se transita la creciente hasta las salida de la cuenca. Este proceso se denomina tránsito de creciente. Esto es, el proceso por el cual el caudal de agua a la salida de la sub-cuenca viaja hacia aguas abajo del río, alterándose el hidrograma de una sección del mismo, a la sección siguiente. Para este tipo de casos se implementan los modelos de tránsito de crecientes. Los mismos tienen por objetivo calcular el hidrograma en una sección del río, conociendo el hidrograma en otra sección, aguas arriba del mismo.

Hasta la actualidad, se ha desarrollado un gran número de modelos aplicados a estos procesos, barriendo un campo que abarca desde modelos simples de producción o transferencias a modelos más complejos que comprenden ambos procesos, más el de tránsito de crecidas. En esta práctica se utilizó el modelo Génie Rural (GR) de Charles Perrin et al,2007.

3.7. OPTIMIZACIÓN – CONCEPTOS GENERALES

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 como la mejor.

3.7.1. METODOS DE OPTIMIZACION

Las metodologías de optimización nacen de la mano de la Investigación de Operaciones, por lo cual tienen un origen común como ciencia aplicada. La mayoría de los textos introductorios a la Investigación de Operaciones describen un desarrollo histórico parecido de esta ciencia. Información más detallada se puede encontrar en literatura especializada (Alonso, 2008).

A los fines de este informe interesa señalar los avances que se han logrado desde la década de los sesenta, en cuanto a los conocidos como métodos heurísticos. Estos se aplican a aquellos problemas de optimización que 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. En los primeros tiempos se generaron métodos orientados específicamente a la resolución de cada problema, 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, SimulatedAnnealing, Algoritmos Genéticos, Redes Neuronales, Métodos basados en Trayectoria de Hormiga y Métodos basados en Inteligencia Artificial. En este informe, 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; USACH, 2007, 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 la caso de minimización o el mayor en el caso de maximización.

3.7.2. GENERACION DE NUMEROS ALEATORIOS

Los métodos heurísticos constituyen la propuesta de los nuevos paradigmas computacionales de optimización, entre los cuales se encuentran los métodos de los algoritmos genéticos. La mayoría de estos métodos tienen entre sus rutinas la generación de números aleatorios con distintas distribuciones probabilísticas (Goldberg, 1989, Duan et al., 1992, Yapo et al., 1998, Vrugt et al., 2003, Viveros, et al. 1998). Se presentarán en este informe solamente los métodos clásicos de generación de números aleatorios con distribución uniforme. Para generar números con distintas distribuciones probabilísticas, se recurre al método de Montecarlo, desarrollado principalmente por Stan Ulam y John Von Neumann, por el año 1944, cuya presentación formal puede encontrarse, entre otros, en Eckhardt (1987) y Herring y Palmore (1989).

La generación de variables aleatorias tiene un rol preponderante en el proceso de simulación. Para generar variables aleatorias es necesario contar con un generador de números aleatorios. Para que un número sea considerado aleatorio, debe tener las siguientes características (UROU, 2006): 1) Distribución Uniforme. Lo que significa que cualquier número que pertenezca al rango de interés debe tener la misma chance de resultar sorteado. 2) No exista Correlación Serial. Lo que significa que la aparición de un número no afecta de ninguna manera la chance de aparecer de otro número o de sí mismo. Existen pruebas para comprobar que un generador cumpla estas dos condiciones.

La forma más perfecta de generación de números aleatorios es a través de algún modelo físico, como por ejemplo la ruleta, que es el modelo por excelencia. También pueden conseguirse números aleatorios a través de tablas que han sido construidas con modelos físicos (en 1955 la compañía RAND publicó una tabla con un millón de números

aleatorios). Matemáticamente, se puede mencionar dos formas diferentes de generar números aleatorios:

1) *Métodos de generación aleatoria*. En base a algún circuito o característica de la máquina se generan números que son totalmente aleatorios. El método básicamente consiste en interrumpir un proceso uniforme en forma aleatoria (que es esencialmente lo que ocurre cuando la bola cae en un casillero de la ruleta).

2) *Números pseudoaleatorios*. Se generan a través de una fórmula e imitan los valores de una variable aleatoria. Se llaman pseudoaleatorios porque se obtienen realizando un conjunto de operaciones sobre el número generado antes (recurrencia) por lo que no son realmente aleatorios según lo expuesto previamente. A pesar de esto muchos de los métodos de generación de números pseudoaleatorios se comportan correctamente y pasan todos los tests como si fueran una variable aleatoria. Presenta la gran ventaja de ser un método muy veloz y barato y la mayor desventaja es que son de período finito.

En 1950, Von Neumann propuso el conocido *Método de Centros de Cuadrados*. El mismo, se basa en tomar un número, elevarlo al cuadrado y tomar los dígitos del centro como nuevo número, luego repetir el procedimiento. Por ejemplo, el número de 4 dígitos: 2061, elevado al cuadrado resulta en 4247721, del cual los cuatro números centrales son: 2477, que elevado al cuadrado resulta en 6135529, y así sucesivamente. La desventaja es que la secuencia generada por lo general es corta. El ejemplo anterior luego de 34 números degenera a 0; y si en lugar de empezar con 2061 se inicia con 2500, se repite 2500 en la primer iteración (y por lo tanto, en todas las restantes). Aún así si se toman números muy largo se puede llegar a secuencias de 100.000 números diferentes.

En 1951, Lehmer propuso el *Método Congruencial Lineal*, que se ha convertido en el generador pseudo-aleatorio estándar (Herring y Palmore, 1989), ya que es el método utilizado por excelencia, y se basa en la siguiente recurrencia (Eckhardt, 1987):

$$Z_i = (A * Z_{i-1} + C) \bmod M \quad (3-2)$$

Donde: Z_i es el i -ésimo número de la secuencia. A , C y M son parámetros que pueden ser ajustados por conveniencia y para asegurar la naturaleza pseudo-aleatoria de la secuencia. Por ejemplo, el módulo M , es normalmente tomado como la longitud de la palabra de la computadora, y el multiplicador A , se elige para producir tanto un periodo largo de la secuencia pseudo-aleatoria, como para asegurar buenas propiedades estadísticas. La expresión $y \equiv z \pmod{m}$, significa que existe un entero k tal que $mk = y - z$.

La ecuación 3-2, puede ser llevada sencillamente a la forma:

$$Z_i = A * Z_{i-1} + C - [(A * Z_{i-1} + C) / M] * M \quad (3-3)$$

Donde la notación: $[x]$ significa el mayor entero menor o igual a x , y se puede demostrar por inducción completa (UROU, 2006) que:

$$Z_n = [Z_0 * A^n + (A^n - 1) / (A - 1) * C] \pmod{M} \quad (3-4)$$

Si se quieren obtener número Uniformes $(0,1)$ se normaliza el resultado dividiendo por el módulo: $U_i = Z_i / M$.

A continuación se presentan dos ejemplos:

- a) Para los valores: $A = 3$, $C = 0$, $M = 5$ y $Z_0 = 4$; se obtiene la serie: 2,1,3,4,2,1,...
- b) Para los valores: $A = 3$, $C = 0$, $M = 9$ y $Z_0 = 4$; se obtiene la serie: 3,0,0,...

Se desprende de estos ejemplos que, en este método, si se repite un número ya se repite toda la secuencia. Presenta la ventaja de consumir poca memoria y ser tener gran velocidad de cálculo. Además, es muy fácil volver a generar la misma secuencia guardando un solo número, ya que alcanza con partir de la misma semilla Z_0 .

Es muy importante la elección de las constantes A , C , M y la semilla Z_0 ya sea para la velocidad de generación como sobre todo para el largo de la secuencia. En la práctica para aumentar la velocidad de cálculo se suele tomar $M = a^b$ donde a es el número base de

operación de la máquina, por lo que en máquinas binarias $a=2$, esto permite calcular muy fácilmente el resto y efectuar la normalización a $U(0,1)$ operando con desplazamientos y no con divisiones y multiplicaciones.

Con respecto al largo del período de la secuencia siempre será menor o igual a M y es deseable obtener el largo máximo (o sea M) aunque no siempre es posible. Cuando se obtiene largo máximo se dice que el método es de Período Completo.

Se distinguen dos casos según el valor de C .

Si $C = 0$, se llama Método Congruencial Multiplicativo Puro.

Si $C \neq 0$, se llama Método Congruencial Multiplicativo Mixto.

3.8. ALGORITMOS GENÉTICOS

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.

3.8.1. CONCEPTOS BASICO DE BIOLOGIA

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 3.3., 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 CrossingOver, 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.

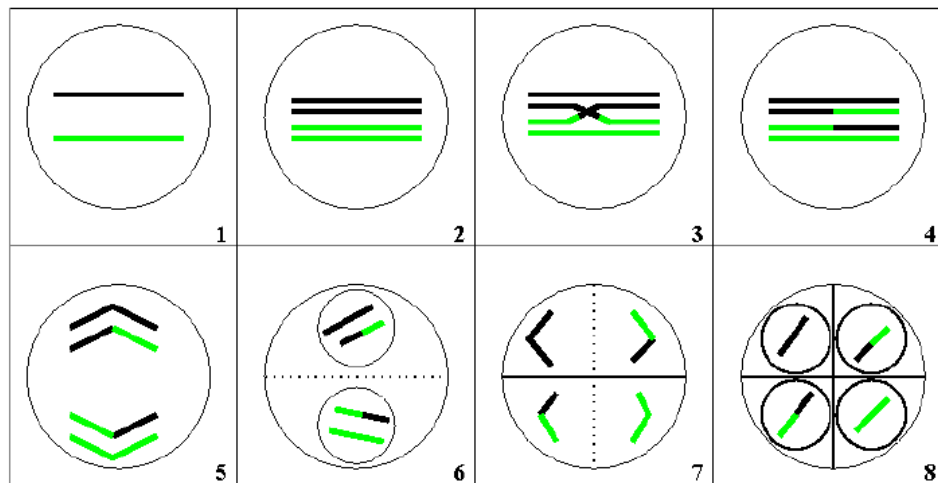


Figura 3.3. Etapas de la meiosis (Kuri y Galaviz, 2002).

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 en medio los “escalones” de esta hélice y ensambla en cada mitad los nucleótidos que debe ensamblar (Figura 3.4.).

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

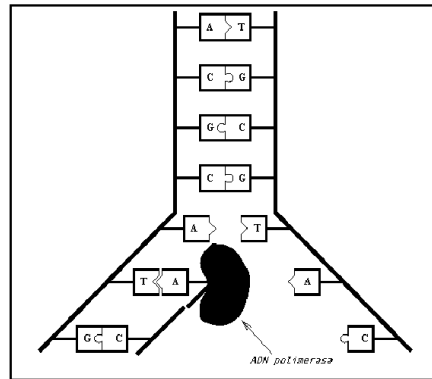


Figura 3.4. Esquema de trabajo de la enzima ADN polimerasa (Kuti y Galaviz, 2002).

3.8.2. LA TECNICA DE LOS ALGORITMOS GENETICOS

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).

3.8.2.1. *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 cumple alguna otra condición de terminación.

3.8.2.2. 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.

3.8.2.3. 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 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.

3.8.2.4. 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.

3.8.2.5. 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.

3.9. 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.

3.9.1. CLASIFICACION DE LAS TECNICAS DE CALIBRACION

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

- desde el punto de vista operativo: técnicas de calibración automática y manual.
- desde el punto de vista de la cantidad de objetivos que se propone optimizar: técnicas de calibración que utilizan funciones mono y multi-objetivo.

- desde el punto de vista del enfoque sobre el dominio analizado: técnicas de calibración local y global.

3.9.1.1. 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 han madurado al punto de sustituir enteramente a la calibración manual.

3.9.1.2. 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 3.5. 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.

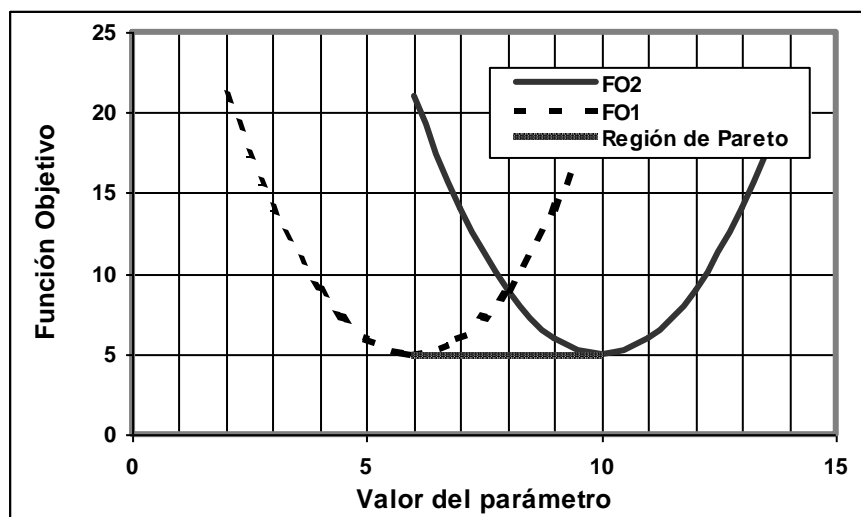


Figura 3.5. Esquema de la Región de Pareto para dos funciones objetivo.

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).

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.

3.9.1.3. *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 3.6. 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.

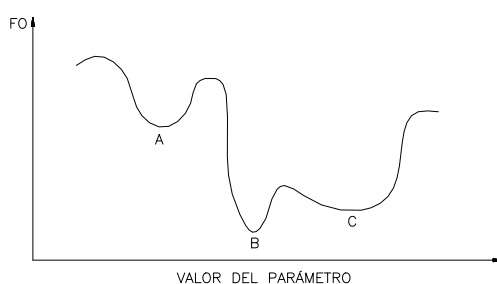


Figura 3.6. Proyección plana de la superficie de respuesta de una FO.

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 objetivos, 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.

Estos métodos iterativos de búsqueda local poseen cuatro características básicas (Tucci, 2005), que son:

- definición del punto de partida,
- dirección de la búsqueda,
- espaciamiento de cada prueba,
- y 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 DX_1 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 X_1 , en cualquier sentido se obtiene un valor superior al actual, entonces se comienza la búsqueda a lo largo de X_2 , para lo cual se elige un sentido de búsqueda y se da un paso DX_2 . 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 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 X_1 , 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 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, aún 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. En sus versiones primitivas, estos métodos consistían esencialmente en la aplicación sucesiva de la técnica de búsqueda local. Surgen los Métodos de Búsqueda Aleatoria (RandomSearch, 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 variables dadas X_1 y X_2 a calibrar en el modelo. Las mismas se van arrojando aleatoriamente según una distribución probabilística uniforme para cada variable 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 (GeneticAlgorithm, GA) que han sido tratados en sección aparte.

3.10. CALIBRACIÓN AUTOMÁTICA DE MODELOS P-Q.

Un procedimiento típico de estimación automática de los parámetros de un modelo hidrológico comprende los siguientes cuatro elementos principales (Sorooshian y Gupta, 1995):

1. Función objetivo, también llamada criterio de calibración.
2. Algoritmo de optimización.
3. Criterio de parada.
4. Datos de calibración.

3.10.1. LA FUNCION OBJETIVO

Todos los autores concuerdan con la destacada importancia de las funciones objetivos en el proceso de calibración automática de los modelos hidrológicos. Sorooshian y Gupta (1995) apuntan una sintética definición 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)". 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".

Debido a que la FO es la expresión matemática que evalúa la calidad de la calibración, debe ser elegida con cuidado, de modo que represente fielmente el error de la calibración con una estructura matemática lo más sencilla posible.

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. Las FO deben tener una estructura matemática sencilla, para poder identificar con claridad el sentido físico del número que nos arroja.

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 ecuación 3-5.

$$FO = \sum_{i=1}^n w_i (QO_i - QC_i)^2 \quad (3-5)$$

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 ecuación 3-6.

$$FO = \sum_{i=1}^n (QO_i - QC_i)^2 \quad (3-6)$$

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 una diferencia de caudales iguales, el porcentaje de error es mayor cuando los caudales son más pequeños. Por ejemplo, si se estima 18 m³/s, para un caudal observado de estiaje de 20 m³/s, y 198 m³/s, para un caudal de crecida de 200 m³/s, en ambos casos se estará cometiendo un error de 2 m³/s, sin embargo, en estiaje representa el 10% del caudal observado, mientras que en crecida, sólo el 1%. Estas diferencias todavía se marcan más, cuando consideramos que se elevan al cuadrado. Entonces, si nuestro método de ajuste asegura un error del 5%, por ejemplo, para el ajuste de los hidrogramas, en los picos es insignificante, pero en estiaje se puede estar cometiendo un error importante.

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 al cual se asumen dos hipótesis referidas 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 heteroscedasticidad (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 (Sorooshian y Gupta, 1995) ha sido una función llamada HMLE (Heteroscedastic Maximum Likelihood Estimator) 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 ecuación 3-7.

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

Donde:

$\varepsilon_i = QO_i - QC_i$, 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), propone como 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 3-8, 3-9, y 3-10. Indica Tucci (2005), al respecto de estas FO, 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.

$$FO = \sum_{i=1}^n |QO_i - QC_i| \quad (3-8)$$

$$FO = \sum_{i=1}^n \left(\frac{1}{QO_i} - \frac{1}{QC_i} \right)^2 \quad (3-9)$$

$$FO = \sum_{i=1}^n \left(\frac{QO_i - QC_i}{QO_i} \right)^2 \quad (3-10)$$

Tassi et al. (2006), presentan las tres FO anteriores, más la Función Cuadrática (ecuación 3-6), pero dividiéndolas por el número de datos n , y obteniéndole la raíz cuadrada. Con lo cual, 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, Tassi et al. (2006), agregan una nueva FO, que se muestra en la ecuación 3-11. Esta forma de FO también es elegida por Yapo et al. (1998), quienes adoptan además el criterio de HMLE para presentar y probar el desempeño del algoritmo multiobjetivo MOCOM-UA.

$$FO = \sum_{i=1}^n \left(\frac{QO_i - QC_i}{QO_i} \right) \quad (3-11)$$

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, las FO presentadas en las ecuaciones de la 3-12 a la 3-15, que asignan cierto protagonismo al peso de los errores en cada intervalo de la serie de datos, y las denomina respectivamente: error global de volumen, raíz media cuadrática global de los errores, raíz media cuadrática promedio de los errores al pico de los eventos, raíz media cuadrática promedio de los errores en los caudales bajos de los eventos.

$$FO = \left| \frac{\sum_{i=1}^n w_i (QO_i - QC_i)}{\sum_{i=1}^n w_i} \right| \quad (3-12)$$

$$FO = \left| \frac{\sum_{i=1}^n w_i^2 (QO_i - QC_i)^2}{\sum_{i=1}^n w_i^2} \right|^{1/2} \quad (3-13)$$

$$FO = \frac{1}{M_p} \sum_{j=1}^{M_p} \left| \frac{\sum_{i=1}^{n_j} w_i^2 (QO_i - QC_i)^2}{\sum_{i=1}^n w_i^2} \right|^{1/2} \quad (3-14)$$

$$FO = \frac{1}{M_l} \sum_{j=1}^{M_l} \left| \frac{\sum_{i=1}^{n_j} w_i^2 (QO_i - QC_i)^2}{\sum_{i=1}^n w_i^2} \right|^{1/2} \quad (3-15)$$

Donde

M_p es el número de eventos con caudales pico.

M_l es el número de eventos con caudales bajos.

n_j es el número de intervalos de tiempo con caudales de pico, para la ecuación 3-14, y de caudales bajos, para la ecuación 3-15.

Los eventos con caudales pico se definen como los periodos donde el caudal observado está por encima de un dado nivel prefijado. Del mismo modo, los eventos con caudales bajos se definen como los periodos donde el caudal observado está por debajo de un dado nivel prefijado.

Los pesos, siempre según Madsen (2000), representan la importancia a darle a una porción particular del hidrograma, reflejando el error en los datos y en el modelo. Es decir, se les asigna poco peso a los periodos en los cuales se presentan largas series de datos, o errores del modelo.

3.10.1.1. Coeficiente de Nash Sutcliffe

Una vez calibrado el modelo matemático, es necesario verificar la bondad del ajuste realizado. Tradicionalmente, se ha venido utilizando el famoso coeficiente de Nash y Sutcliffe (1970) para cuantificar la performance del ajuste. En la ecuación 3-16, se presenta dicho coeficiente, que representa una reinterpretación del coeficiente de determinación, empleado en regresión lineal.

$$R^2 = 1 - \frac{\sum_{i=1}^n w_i^2 (QO_i - QC_i)^2}{\sum_{i=1}^n w_i^2 (QO_i - \overline{QO})^2} \quad (3-16)$$

Donde \overline{QO} es el promedio de los caudales del hidrograma observado.

En su forma simplificada, se suprimen los pesos (w_i), y resulta la ecuación 3-17.

$$R^2 = 1 - \frac{\sum_{i=1}^n (QO_i - QC_i)^2}{\sum_{i=1}^n (QO_i - \overline{QO})^2} \quad (3-17)$$

Sin embargo, en los últimos tiempos, se emplea este coeficiente directamente para realizar la calibración, considerando claro que en este caso el objetivo es maximizar dicho coeficiente. Es decir, en vez de minimizar una FO que represente un error, se opta por maximizar una medida numérica que representa el grado de performance del modelo.

Si se observa su formulación, el criterio de Nash-Sutcliffe consiste en restar a la unidad un cociente que representa una medida cuadrática del error, ya que está dado por la relación entre dos medidas cuadráticas. En el numerador de dicha relación tenemos el desvío cuadrado de la simulación respecto de la observación, y en el denominador, el desvío cuadrado de la observación respecto de la media observada.

Dicha relación no podrá ser nunca negativa ya que tanto numerador como denominador son valores cuadráticos. En consecuencia, el coeficiente de Nash-Sutcliffe, nunca podrá ser mayor que uno. Sin embargo, 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 R sería negativo. En teoría si el error fuera inmenso R puede tender incluso a infinito.

La aptitud de este coeficiente para presentar una referencia de calidad de ajuste del modelo en modo simulación, para cualquier tipo de cuenca hidrográfica, radica en que compara el error cometido en la simulación ($QO_i - QCi$), respecto de una magnitud de referencia, dada por la amplitud promedio del hidrograma observado respecto de la media observada ($QO_i - \overline{QO}$).

En simulación, ese criterio es bueno porque del hidrograma completo observado, rescatamos un valor constante que es la media, y respecto de ése valor normalizamos. Al normalizar se puede comparar simulaciones entre sí, aunque pertenezcan a distintas cuencas, períodos, etc. Es decir, ya que el denominador mide la distancia al cuadrado entre los caudales y la media, nos da un orden de magnitud de las dimensiones de los caudales medidos. Entonces promediando esas diferencias elevadas al cuadrado, nos da la varianza de esa distribución, o sea nos da el orden de magnitud de las variaciones.

En otras palabras, es un modo de adimensionalizar el error, empleando como técnica dividir la varianza real, respecto de la varianza de las magnitudes observadas con relación a la media.

Eso es muy valioso porque, 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 calculamos 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, tenemos un orden de magnitud de los valores que alcanzan las fluctuaciones reales en ese río. Si luego dividimos la varianza de los caudales simulados respecto de los observados, tenemos un número que representa la magnitud del error. La relación entre ambos, me 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 me da una idea del error al cuadrado normalizado, es como la varianza normalizada. De modo que puedo comparar ríos con módulos del orden de los 10 m³/s con otros de 5000 m³/s.

A pesar de ser un excelente indicador de la calidad del ajuste, el valor numérico que representa puede ser un poco engañoso, ya que por ejemplo un coeficiente de 0.9 no implica que el ajuste haya sido un 10% superior a otro con un coeficiente de 0.8. En la página 118 de la tesis doctoral de Perrin (2000), se da una explicación acabada de las razones que justifican esta afirmación, pero exceden este trabajo.

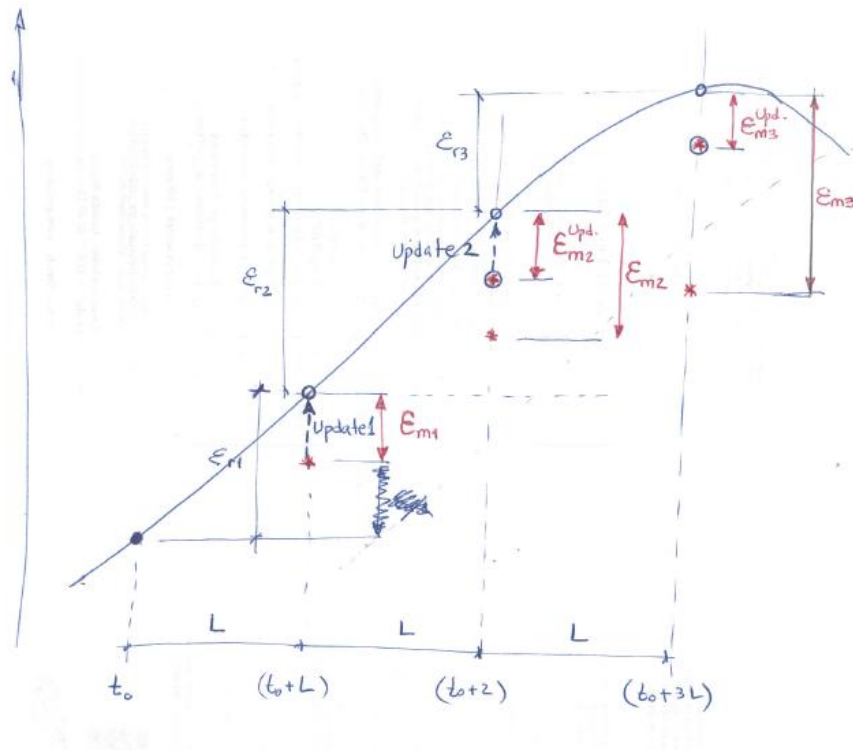
3.10.1.2. *Coefficiente de Persistencia*

El Coeficiente de Nash-Sutcliffe ha demostrado ser un buen indicador para representar de manera normalizada la calidad del ajuste de un modelo que trabaja en simulación. Sin embargo, presenta cierta deficiencia como indicador de la calidad de ajuste de un modelo aplicado al pronóstico, ya que 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 nos interesa específicamente el tramo de hidrograma en el paso en el cual estamos pronosticando. Por eso surge el concepto de “persistencia” como criterio de evaluación de la calidad de pronóstico.

La expresión del Coeficiente de Persistencia (PC, por sus siglas en inglés) tiene la forma que se presenta en la siguiente ecuación:

$$PC = 1 - \frac{\sum \epsilon_m^2}{\sum \epsilon_r^2}$$

Para comprender la notación correspondiente a la expresión del PC, es necesario interpretar el gráfico de la Figura XXXX.



En dicho gráfico se observa el tramo de un hidrograma observado en línea continua, y se indican con puntos las estimaciones que el modelo hace de los valores de caudal. Los puntos en estrella roja, corresponden al hidrograma generado en tiempo t_0 . Es decir, el dibujo indica que en tiempo t_0 , se corre el modelo de pronóstico y se obtiene para horizontes de tiempo (t_0+L) , (t_0+2L) , (t_0+3L) , etc, el hidrograma pronosticado. Cuando ha transcurrido el tiempo “L”, es decir ubicados en el instante de tiempo (t_0+L) , ya se puede constatar si el caudal pronosticado para ese instante es el correcto. De no ser así, el error cometido se denota como E_m . También en ese instante estamos en condiciones de modificar la variable de estado “R” del modelo GR, al valor necesario para obtener un pronóstico adecuado en el tiempo actual.

La dimensión E_r 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 “E”, en realidad no es un error, sino una diferencia simplemente, que indica la medida con

respecto a la cual se normalizan los errores de pronóstico. Esa dimensión es representativa en pronóstico como dimensión de referencia porque en pronóstico no nos interesa tanto cuánto ha sido el error cometido con respecto a la media (como se indicó para coeficiente de Nash-Sutcliffe), sino con respecto a una dimensión local del paso en que estamos pronosticando. Como E_r representa la diferencia de los valores de caudal entre el tiempo actual y el tiempo futuro pronosticado, me da 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-Sutcliffe, 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 le importa cuál es 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.

3.10.2. ALGORITMOS DE OPTIMIZACION 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, agregan Sorooshian y Gupta (1995), es típicamente implementado en una computadora digital para permitir una búsqueda rápida.

Una descripción detallada de la historia en cuanto al estudio de los algoritmos de optimización puede encontrarse en la tesis de maestría del Mg. Ing. Facundo José Alonso. Aca solo se abarcara lo que es de estrictamente de interés para esta práctica.

Una 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, (Ibbit, 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 mostraron no ser tan grandes como a los datos libres de errores.

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 (ShuffledComplexEvolutionmethod), 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.

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 multiobjetivo 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 multiobjetivo. 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.

Vrugt et al. (2003) propusieron una 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 ShuffledComplexEvolutionMetropolis (SCEM-UA), ya que incorpora conceptos del método de Metrópolis. Este algoritmo está diseñado para inferir la distribución posterior de los parámetros del modelo hidrológico. Proponen combinar las destrezas del exitoso algoritmo SCE-UA con la ventaja de los métodos de cadenas de Markov Monte Carlo, que han ido aumentando su popularidad debido a que estiman la distribución de probabilidades posterior de los parámetros en los modelos hidrológicos. En este trabajo Vrugt et al. (2003), presentan tres casos de estudio, en los cuales demuestran que las capacidades de adaptación del SCEM-UA reduce significativamente el número de simulaciones del modelo necesarias para inferir la distribución posterior de los parámetros cuando se comparan con el tradicional manipulador de muestras Metropoli-Hastings.

Un nuevo equipo de trabajo (Vrugt et al., 2003), encabezado por Vrugt, lanza el mismo año, la versión multiobjetivo del SCEM-UA, denominado MOSCEM-UA (MultiobjectiveShuffledComplexEvolutionMetropolis). 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 multiobjetivo 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 caso de estudio de modelaciones hidrológicas que se incrementan en complejidad.

3.10.3. CRITERIO DE PARADA

Las estrategias de optimización que se presentan en este trabajo, son todas procedimientos iterativos que buscan el juego óptimo de valores para los parámetros mediante pasos de aproximación. En consecuencia, como lo establecen Sorooshian y Gupta (1995), resulta necesario establecer un criterio que determine cuando debe detenerse la búsqueda. En principio, la solución existe en aquel punto de la superficie de respuesta en el cual la pendiente es nula y el valor de la función, un mínimo. En la práctica, es virtualmente imposible, agregan Sorooshian y Gupta (1995), conocer en qué momento exacto ese punto ha sido alcanzado, en consecuencia, estos autores proponen los criterios de terminación mayormente usados en la práctica, del modo que se transcriben a continuación.

Convergencia de la FO. Una manera sencilla de terminar la búsqueda es detenerla cuando el algoritmo es incapaz de apreciar una mejora del valor de la FO en dos iteraciones seguidas. Aunque este criterio puede indicar que se ha alcanzado el óptimo, también puede representar simplemente que se ha llegado a una zona plana de la superficie de respuesta. Si no fuese importante el valor preciso del óptimo, la convergencia de la FO puede ser muy útil como criterio de parada. Una implementación típica es detenerse cuando se cumple con la condición:

$$\frac{(FO_{i-1} - FO_i)}{FO_i} \leq \varepsilon_{FO} \quad (3-18)$$

Donde:

FO_i: valor de la FO en el paso iterativo i.

eFO : error de la FO prefijado. Por ejemplo eFO = 10⁻³.

Convergencia de los parámetros. Otra forma de terminar la búsqueda es cuando el algoritmo es incapaz de generar un cambio apreciable en el valor de los parámetros, y simultáneamente mejorar el valor de la FO. Aunque este criterio puede indicar que se ha alcanzado el óptimo, también puede representar simplemente que se ha alcanzado una región con alta interacción paramétrica en la superficie de respuesta, como por ejemplo, una región al tipo de un estrecho valle. Sorooshian y Gupta (1995) comentan que se han obtenido mejores resultados con este criterio que con el anterior. Una implementación típica de este criterio se muestra en la siguiente ecuación:

$$\frac{(\theta(j)_{i-1} - \theta(j)_i)}{(\theta(j)_{\max} - \theta(j)_{\min})} \leq \varepsilon_{\theta} \quad (3-19)$$

Donde:

$\theta(j)_i$: valor del parámetro j, en el paso iterativo i.

ε_{θ} : error paramétrico prefijado. Por ejemplo $\varepsilon_{\theta} = 10^{-3}$.

Número máximo de iteraciones. Si el tiempo de computadora es limitado, y para asegurar que el algoritmo no quedará atrapado en un ciclo infinito, se emplea como práctica normal, programar la terminación del algoritmo si se ha alcanzado un número máximo preestablecido de iteraciones, a menos que se haya cumplido previamente con alguno de los criterios anteriores. Para métodos de búsqueda aleatoria, este es el procedimiento normal de terminación. No es posible dar una regla fija respecto del número de iteraciones convenientes, ya que es dependiente, tanto del algoritmo como del problema particular tratado.

3.10.3.1. *Limitaciones*

Ninguno de estos criterios son confiables en términos de dar una indicación precisa respecto del arribo al óptimo global, excepto en el caso más trivial en el cual la FO

presenta una superficie de respuesta convexa y de comportamiento suave. Sin embargo, en la experiencia de Sorooshian y Gupta (1995), el criterio de terminación por convergencia de los parámetros es más adecuado para los estudios de calibración, ya que termina la calibración cuando el parámetro no está cambiando significativamente. El criterio de máximo número de iteraciones se emplea como respaldo para evitar consumo computacional innecesario. Si el algoritmo no concluye dentro de un número “razonable” de iteraciones, el código computacional debería ser examinado de “bugs”. Finalizan los autores citados, indicando que la mejor manera de establecer valores apropiados como criterios de terminación consiste en correr el algoritmo empleando series sintéticas de datos sobre el modelo.

3.10.4. DATOS DE CALIBRACION

Para realizar la calibración automática de un modelo hidrológico de transformación lluvia-caudal, ajustado para el funcionamiento en una cuenca específica, resulta necesario contar con una serie de datos buenos y confiables. Sin embargo, según ironizan Sorooshian y Gupta (1995), poco se conoce objetivamente respecto de aquello que constituye una “buena” serie de datos de calibración. Por esta razón, resulta necesario analizar cuántos y qué tipo de datos darán los mejores o aceptables resultados de calibración automática. Los autores citados, establecen los siguientes criterios que se transcriben a continuación.

Cantidad de datos. La práctica común consiste en utilizar la mayor cantidad de los datos disponibles para la calibración, luego de apartar una cierta cantidad de datos para la verificación de los resultados. Sin embargo, los estudios de Sorooshian et al. (1983) indicaron que, cuando el procedimiento de calibración había sido correctamente diseñado, sólo un año de datos diarios eran suficientes como para obtener estimaciones conceptualmente realistas para los parámetros del modelo de pronóstico de crecidas de la US-NationalWeatherServices (SMA-NWSRFS) bien presentado en Burnash (1995). Más aún, el uso de series de datos de mayor longitud sólo sirven para mejorar marginalmente

la estimación de los parámetros. En general, desde un punto de vista estadístico, la serie de datos debería ser al menos unas 20 veces el número de parámetros a ser estimados (por ejemplo, si hay 7 parámetros, entonces al menos se debería contar con 140 datos diarios de caudal para computar la función). Esto es, aclaran los autores, simplemente una aproximación de dedo gordo. Gupta y Sorooshian (1985) mostraron que el error estándar ($\sigma(j)$) en la estimación del parámetro (j) decrece con el tamaño n de la muestra aproximadamente de acuerdo a la fórmula de la ecuación:

$$\sigma(j) \propto \frac{1}{\sqrt{n}} \quad (3-20)$$

Como la mejora marginal en $1/n$ se convierte en pequeña después de 500 o 100 puntos de dato, esto sugiere que dos a tres años de datos de calibración deberían ser suficientes, siempre y cuando los datos sean de los “buenos”. Lo cual nos conduce a una discusión acerca de la calidad de los datos.

Calidad de datos. Desde el punto de vista de la calibración del modelo, la calidad de los datos dependen de la información (relacionada con los parámetros) contenida en los datos y el ruido (errores) en los datos. Claramente, es deseable que el contenido de información paramétrica en los datos sea tan larga como sea posible, y el ruido tan pequeño como sea posible.

Respecto de la información, cualitativamente, una serie de datos informativos son aquellos tales que contienen o representan tanta variabilidad en el comportamiento del escurrimiento, que los distintos modos de operación del proceso hidrológico son apropiadamente representados. Esto maximiza las chances de los datos de contener información suficiente como para identificar apropiadamente el valor de los parámetros del modelo (es decir, variaciones de los parámetros del modelo, resultarán en variaciones detectables en las salidas del modelo, fuera del hidrograma observado). Por ejemplo, si los datos elegidos corresponden a un año relativamente seco, ciertos procesos del modelo, no serán activados; en consecuencia, la respuesta del modelo será insensible a

algunos de los parámetros del modelo que determinan el fraccionamiento de la humedad entre las distintas capas subterráneas y el agua superficial. Sin embargo, si la serie de datos corresponde a un año tan húmedo que mantiene la capa superior del suelo saturada la mayor parte del tiempo, la respuesta del modelo puede ser insensible a variaciones en los parámetros que controlan las distintas capas subterráneas del modelo.

La mejor elección parece ser una serie de datos que contenga una gran “variabilidad hidrológica”. Por ejemplo, Gupta y Sorooshian (1985b) mostraron que la identificabilidad del parámetro correspondiente a la capa superior de humedad de un modelo de escurrimiento típico, depende no de cuán larga sea el flujo excedente en esta capa, sino de cuantas veces el flujo excede este reservorio. Por consiguiente, la mayoría de las veces en que el régimen hidrológico intercambia entre modos secos, medios y húmedos, es probable que los datos sean más informativo aún.

Con respecto al error en los datos, la presencia de errores en la medida y escritura de los datos causa un deterioro en la calidad, resultando en un desmedro de confianza en la estimación de los parámetros. El error en los datos tiene distintas fuentes, como registro incorrecto de la máquina, mal funcionamiento de la máquina, métodos de transferencia de datos a la computadora, aproximaciones en la curva de transformación, cambios en el fondo del río, aproximaciones en las fórmulas usadas para convertir registros de campo, negligencia, fraude, etc. Aunque se ha sugerido que los efectos combinados de todos esos errores deberían tender hacia una distribución normal, esto no ha sido demostrado suficientemente. En la práctica, los errores en los datos probablemente contengan sesgos, particularmente en el caso de datos de precipitación, que tienden a subestimar el valor real de lluvia. Los datos de caudal son en general, los más confiables, pero contienen típicamente un 10 % de error. En la selección de los datos para calibración del modelo, sería deseable que la serie sea cuidadosamente examinada para identificar primeramente errores groseros, por ejemplo periodos donde el caudal se eleva, pero no se observan datos de precipitación. Mientras que los errores en los datos de caudal pueden ser

manejados mediante el uso de la máxima verosimilitud y otras teorías estadísticas, todavía deben desarrollarse métodos satisfactorios para el manejo apropiado de los errores de precipitación.

4. METODOLOGIA

En este capítulo se describe el modelo hidrológico Génie Rural (Perrin et al., 2007), y el método de optimización MOCOM-UA (Yapo et al., 1998), empleado para realizar la calibración automática multi-objetivo, del modelo hidrológico.

4.1. DESCRIPCION DEL MODELO HIDROLOGICO GR

4.1.1. OBJETIVOS DE SU DESARROLLO

La simulación del caudal a 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).

4.1.2. PRINCIPALES CARACTERÍSTICAS

Como todo modelo hidrológico, los modelos GR operan una triple globalización:

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 de 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., 2006a), (ii) Versión mensual denominado GR2M (Mouelhi et al., 2006a), (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.

4.1.3. ESTRUCTURA DEL MODELO

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 se trate de un modelo orientado al paso de tiempo anual, diario u horario.

A continuación se presenta el desarrollo sólo de la versión GR4J, es decir, con paso de tiempo diario, y las pequeñas modificaciones relativas al paso de tiempo horario, ya que son las empleadas en este trabajo. La versión del modelo horario de este trabajo, no es exactamente la versión GR3H, mencionada previamente, sino muy semejante al modelo GR4J, de cuatro parámetros, con una sensible diferencia que se muestra oportunamente durante la presentación del modelo GR4J.

El modelo GR4J (Génie Rural à quatre paramètres journaliste), tiene paso de tiempo diario y dispone de 4 parámetros ajustables. Es un modelo de precipitación-caudal global. Su desarrollo se inició en el instituto “Cemagref” (hoy “Irstea”) a principios de 1980, con el

objetivo de desarrollar un modelo de simulación robusta y fiable para utilizar en problemas de lluvia-caudal aplicados a la gestión de los recursos hídricos en ingeniería (diseño, pronóstico de inundaciones y flujos bajos, manejo de reservorios, detección de impacto, etc) .

Este modelo ha sido adecuado en varias versiones en los trabajos de EdijatnoyMichel (1989), Edijatno (1991), Nascimento (1995), Edijatno et al. (1999), Perrin(2000), Perrin (2002) y Perrin et al. (2003). Estos han mejorado progresivamente el rendimiento del modelo. La versión que aquí se presenta es la de Perrin et al. (2003).

Aunque GR4J es un modelo empírico, su estructura es similar a la de los modelos conceptuales que utilizan tanques, con 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, unos hidrogramasunitarios, así como una función de apertura hacia el exterior no atmosférico (es decir, subterráneo), permitiendo simular el comportamiento global de la grancuencahidrológica, que incluye el flujo subterráneo.

4.1.3.1. DESCRIPCIÓN MATEMÁTICA DEL MODELO GR4J.

En la Figura 4.1, se muestra un diagrama de la estructura del modelo. En la misma, “Pk” es la lluvia diaria del día “k”, y “E” es la evapotranspiración potencial (EP) media del mismo día calendario.

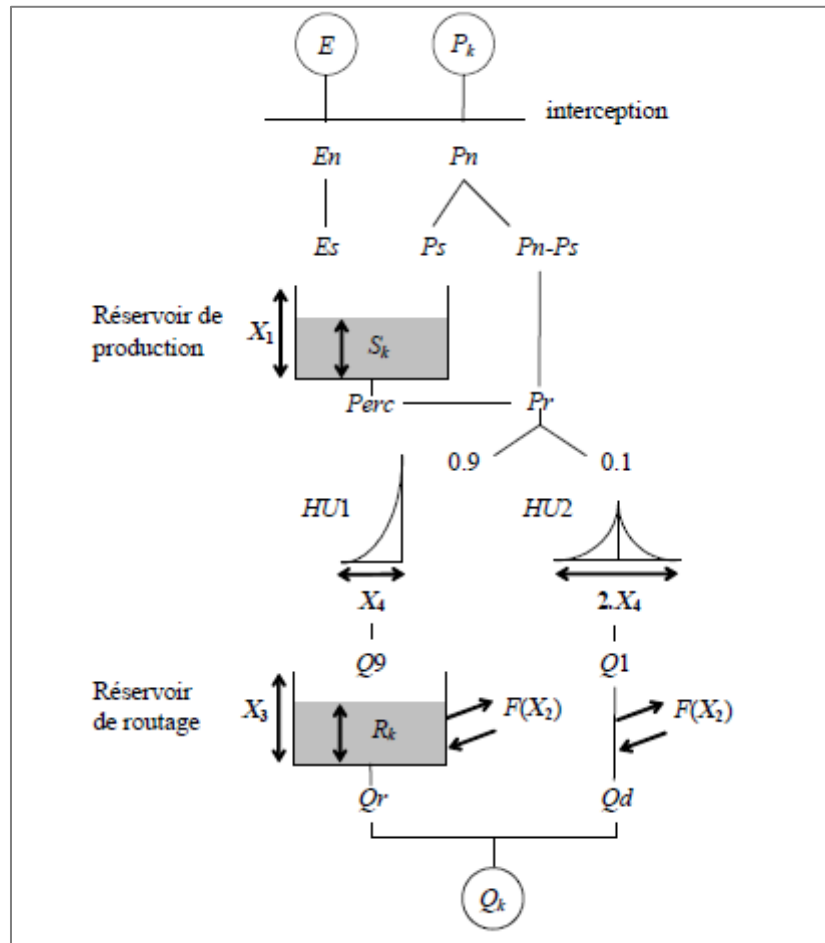


Figura 4.1. Estructura del modelo GR4J (Perrin et al., 2003).

- Neutralización

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

$$\text{Si } P_k \geq E, \text{ entonces } P_n = P_k - E \quad \text{y} \quad E_n = 0 \quad (4.1)$$

$$\text{Si } P_k < E, \text{ entonces } P_n = 0 \quad \text{y} \quad E_n = E - P_k \quad (4.2)$$

- Función de salida

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

$$P_s = \frac{X_1 \left(1 - \left(\frac{S_k}{X_1}\right)^2\right) \tanh\left(\frac{P_n}{X_1}\right)}{1 + \frac{S_k}{X_1} \tanh\left(\frac{P_n}{X_1}\right)} \quad (4.3)$$

Donde X_1 (mm) es la capacidad máxima del reservorio de producción S_k al comienzo del día "k".

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

$$E_s = \frac{S_k \left(2 - \frac{S_k}{X_1}\right) \tanh\left(\frac{E_n}{X_1}\right)}{1 + \left(1 - \frac{S_k}{X_1}\right) \tanh\left(\frac{E_n}{X_1}\right)} \quad (4.4)$$

El contenido del reservorio resultante, está dado por:

$$S' = S_k + P_s - E_s \quad (4.5)$$

- Percolación

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

$$Perc = S' * \left\{ 1 - \left[1 + \left(\frac{4}{9} * \frac{S'}{X_1} \right)^4 \right]^{\frac{1}{4}} \right\} \quad (4.6)$$

Como resultado, el contenido del tanque se convierte en:

$$S_{k+1} = S' - Perc \quad (4.7)$$

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

$$Pr = Perc + (Pn - Ps) \quad (4.8)$$

- 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 X_4 , que es el tiempo en la base de HU1 expresado en días.

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:

$$\text{Para } t \leq 0, \quad SH1(t) = 0 \quad (4.9)$$

$$\text{Para } 0 < t < X_4, \quad SH1(t) = \left(\frac{t}{X_4}\right)^{\frac{5}{2}} \quad (4.10)$$

$$\text{Para } t \geq X_4 \quad SH1(t) = 1 \quad (4.11)$$

SH2 se define de manera similar:

$$\text{Para } t \leq 0, \quad SH2(t) = 0 \quad (4.12)$$

$$\text{Para } 0 < t < X_4, \quad SH2(t) = \frac{1}{2} * \left(\frac{t}{X_4}\right)^{\frac{5}{2}} \quad (4.13)$$

$$\text{Para } 0 < t < 2 X_4, \quad SH2(t) = 1 - \frac{1}{2} * \left(2 - \frac{t}{X_4}\right)^{\frac{5}{2}} \quad (4.14)$$

$$\text{Para } t \geq X_4 \quad SH2(t) = 1 \quad (4.15)$$

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.

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

$$UH1(j) = SH1(j) - SH1(j - 1) \quad (4.16)$$

$$UH2(j) = SH2(j) - SH2(j - 1) \quad (4.17)$$

Donde j es un número entero.

En cada paso de tiempo k, los caudales Q9 y Q1 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:

$$Q9(k) = 0.9 * \sum_{j=1}^l UH1(j) * Pr(k - j + 1) \quad (4.18)$$

$$Q1(k) = 0.1 * \sum_{j=1}^m UH2(j) * Pr(k - j + 1) \quad (4.19)$$

Donde l = int(X₄) + 1 y m = int(2.X₄) + 1, int(.) denota la parte entera.

- Función de intercambio con el exterior no atmosférico

El intercambio de agua subterránea se calcula a través de:

$$F = X_2 * \left(\frac{RK}{X_3}\right)^{\frac{7}{2}} \quad (4.20)$$

Donde R_k es el nivel en el reservorio al inicio del paso de tiempo, X_3 es la capacidad del tanque a un día, 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:

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

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

$$Q_r = R' * \left\{ 1 - \left[1 + \left(\frac{R'}{X_3} \right)^4 \right]^{\frac{1}{4}} \right\} \quad (4.21)$$

Entonces el nivel del reservorio se convierte en:

$$RK + 1 = R' - Q_r \quad (4.22)$$

- Total

La salida Q_1 del hidrograma HU2 es sometida al intercambio para dar la componente de flujo Q_d :

$$Q_d = \max (0 ; Q_1(k) + F)$$

El caudal total Q resulta entonces:

$$QK = Qr + Qd \quad (4.23)$$

- Parámetros:

El modelo GR4J 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.

En una amplia muestra de las cuencas hidrográficas, los valores de parámetros que se obtuvieron se encuentran en la tabla 4.1.

Tabla 4.1. Parámetros modelo GR4J, (Perrin et al. 2003).

Paramètre	Médiane	Intervalle de confiance à 80%
X ₁ (mm)	350	100 à 1200
X ₂ (mm)	0	-5 à 3
X ₃ (mm)	90	20 à 300
X ₄ (jours)	1.7	1.1 à 2.9

- Modo pronóstico:

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 figura 4.2. se presenta un esquema del modelo en modo pronóstico.

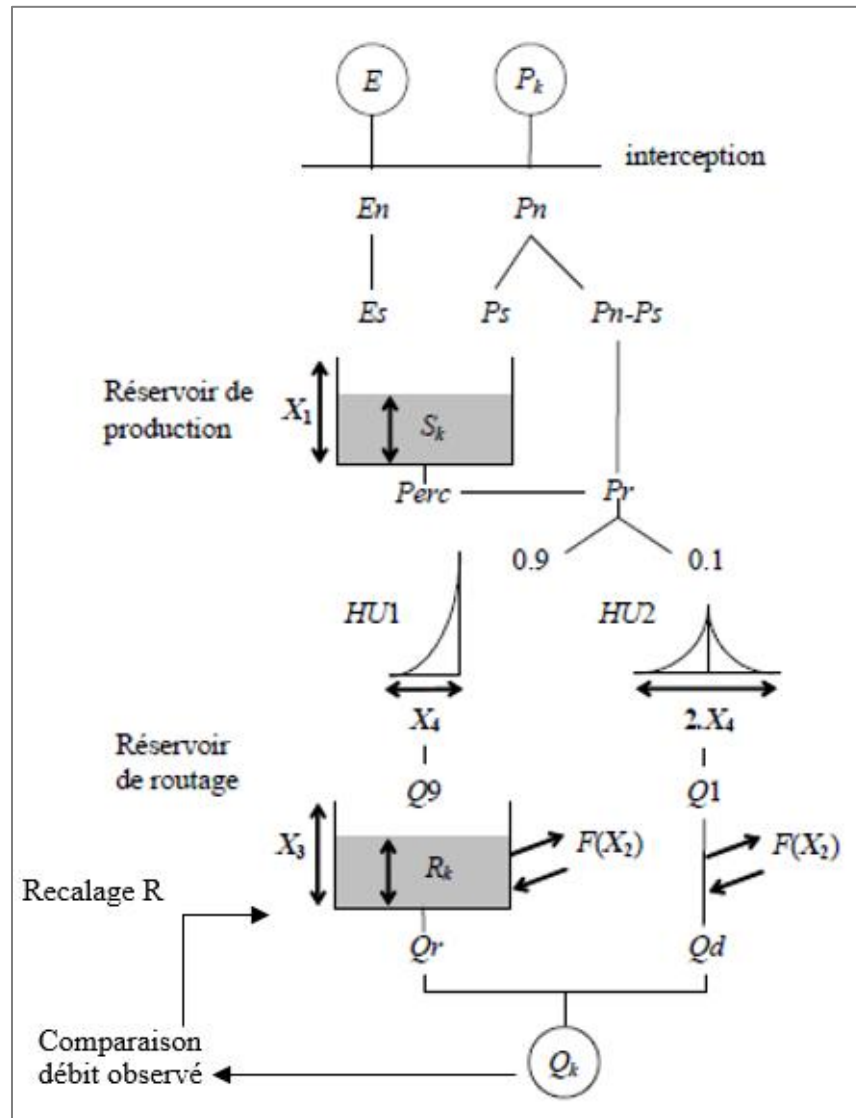


Figura 4.2. Modelo GR4J modo pronóstico.

En dicha figura 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 arriba, 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 $P(t_0)$ y $EP(t_0)$, y genera el hidrograma para t_1 (siendo $t_1 = t_0 + Dt$), es decir el valor Q_1 , y para todo valor de $t > t_0$ (es decir $Q_2, Q_3...$). Cuando transcurre un Dt , y ya estamos en t_1 , es posible verificar si el caudal generado Q_1 cuando estábamos en t_0 , coincide con el valor actual Q_1' de caudal (conocido). Si no coincide, el modelo modifica su variable de estado R_0 (para el paso de tiempo anterior), ajustando de este modo el estado del sistema para que Q_1 sea igual al observado Q_1' . 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 Q_1 , igual al medido Q_1' . Una vez que ajustó R , vuelve a correr y calcula Q_2 , con los datos de entrada P_1 y EP_1 . Y así sucesivamente con todo el periodo de tiempo simulado.

4.2. EL ALGORITMO MOCOM-UA

Como se ha presentado previamente, el algoritmo MOCOM-UA es un algoritmo de optimización multi-objetivo, ya que está diseñado para optimizar un modelo matemático empleando más de una función objetivo.

Se ha dicho que una función objetivo es una expresión matemática que cuantifica la diferencia entre la salida del modelo hidrológico (hidrograma simulado, si se trata de un

modelo de transformación lluvia-caudal), y los datos medidos de referencia (hidrograma medido). Cuanto menor sea dicha función objetivo, mejor habrá sido el ajuste.

Ahora bien, esa “diferencia” que cuantifica la FO, puede ser diferencia de caudales máximos de los hidrogramas, diferencia de volúmenes de escurrimiento directo, diferencias de los caudales de recesión, diferencia de los tiempos en que se produjeron los picos, etc. De acuerdo a la FO elegida para realizar la calibración automática, el juego de valores de parámetros que la minimice será la que mejor representa cierta parte del proceso que simula el modelo.

Por ejemplo, si se elige como FO, aquella que minimiza la diferencia entre caudal máximo simulado y el medido, entonces el juego de valores de parámetros que la minimice, será el que mejor ajusta los picos de los hidrogramas, es decir, probablemente sea el que mejor represente el proceso de transferencia. Pero si se elige la FO que cuantifica la diferencia entre volúmenes de escurrimiento directo de hidrogramas simulado y medido, el juego de parámetros que la minimice será el que mejor representa el proceso de infiltración. Sin embargo, este último juego de valores de los parámetros no tiene por qué coincidir con el anterior. En otras palabras, el conjunto de valores de parámetros que mejor representa la transferencia en la cuenca, no necesariamente sea el que mejor represente la infiltración. En la mayoría de los casos esos valores no sólo no coinciden sino que se diferencian notablemente, al menos en algunos de sus parámetros.

Como consecuencia del análisis anterior, el hidrólogo que va a realizar la calibración automática del modelo de simulación, deberá elegir un criterio de calibración, el cual dependerá de los objetivos de la simulación. Si por ejemplo, el objetivo de la simulación está asociado al proyecto del vertedero de un embalse, indudablemente que deberá elegirse una FO que conduzca a un juego de valores para los parámetros que permita una simulación adecuada de los picos (y de los volúmenes) de las crecidas. Por el contrario, si se está realizando un estudio de balance hídrico, la FO debe calibrar el modelo para adecuarse a simular con precisión los volúmenes escurridos.

A la vista de la experimentación, ha quedado ampliamente demostrado que un mismo juego de parámetros puede ser una excelente solución para un objetivo y pésima para otro, es decir que puede minimizar notablemente una FO y resultar en valores elevados para otra. En efecto, el juego de parámetros que minimiza una FO, normalmente tiene un valor elevado en la otra, y viceversa. Entre el mejor punto para una FO y el mejor para la otra, existe un conjunto de puntos intermedios que conforman una región de puntos tales que ninguno de ellos es mejor que el otro en ambas FO al mismo tiempo. Esta región se denomina Región de Pareto. Frente a este panorama de divergencias, surgen los métodos de optimización multi-objetivo, que se emplean para obtener la Región de Pareto mediante la minimización simultánea de dos o más FO.

Antes de abarcar la descripción paso a paso del método se introduce el proceso de evolución competitiva y el algoritmo de Nelder y Mead, los cuales son empleados, como se verá más adelante, para llevar a cabo la evolución del complejo.

4.2.1. EVOLUCION COMPETITIVA COMPLEJA

La Evolución Competitiva Compleja es un proceso de evolución que se lleva a cabo sobre un conjunto de individuos dados. En el contexto del Algoritmo SCE-UA, el conjunto de individuos sometido a esta evolución es el complejo. A continuación se describe su metodología.

4.2.1.1. *Atribución de pesos.*

En primer lugar, se atribuye a cada individuo del complejo un peso tal que los mejores puntos del complejo (menor valor de FO) tienen mayor peso. Ese peso podría estar dado por la expresión que proponen Duan et al. (1992):

$$\rho_i = \frac{2(m+1-i)}{m(m+1)}, \quad i = 1, \dots, m. \quad (4.24)$$

Donde:

- w_i es el valor del peso asignado al individuo i .
- m es el número de puntos en cada complejo.
- i es la posición del punto dentro del complejo.

De acuerdo a esta expresión, la suma de los pesos de los m puntos del complejo debe ser igual a 1.

4.2.1.2. Selección de subcomplejos.

Luego se elige aleatoriamente un conjunto de puntos dentro del complejo. Ese conjunto de puntos recibe el nombre de subcomplejo, y naturalmente, debe tener un número de puntos menor o igual al tamaño del complejo. Por ejemplo, si el complejo tiene 18 individuos, el tamaño del subcomplejo debe ser menor o igual a 18. Los puntos del subcomplejo son elegidos de modo tal que tengan mayor probabilidad de ser elegidos aquellos puntos del complejo con mayor peso, aun cuando cualquiera pueda serlo. O sea que el método de selección de los puntos del complejo que pasarán a formar parte del subcomplejo, debe dar mayor probabilidad de ser elegidos a los mejores, aunque los peores también podrán formar parte del Subcomplejo.

Siguiendo con nuestra analogía, la selección del Subcomplejo es análoga a la selección de los padres (o reproductores) que serán los que den origen a la siguiente generación. De todos los individuos del complejo, interesa elegir como reproductores a los mejores, por eso se le otorga mayor peso a los puntos con menor FO, para que tengan más probabilidad de formar parte del subcomplejo, o sea de reproducir. Sin embargo, los puntos malos también pueden ser elegidos, aunque tienen menor probabilidad de serlo. La única diferencia que existe con la reproducción natural es que, en la natural, los padres son sólo dos, en cambio, en el subcomplejo (los reproductores) pueden ser más de dos.

Conviene remarcar que:

El número de individuos de la población puede ser cualquiera, aunque conviene que sea grande.

El número de individuos del complejo puede ser cualquiera, siempre que sea un submúltiplo del tamaño de la población.

El número de individuos del subcomplejo puede ser cualquiera, siempre que sea menor o igual que el tamaño del Complejo.

Por ejemplo, de una población de 100 individuos pueden elegirse 5 complejos de 20 individuos cada uno. A su vez, de cada complejo pueden tomarse subcomplejos de 13 individuos.

4.2.1.3. Evolución de los Subcomplejos.

Una vez que se elige el subcomplejo, el mismo evoluciona mediante el Algoritmo de Nelder y Mead un número “a” de veces. Luego de esta evolución, el subcomplejo habrá llegado a un nuevo conjunto de puntos. El Algoritmo de Nelder y Mead es un algoritmo de búsqueda local, que se explica más adelante. El subcomplejo con sus nuevos puntos es devuelto al complejo, en donde se reúnen con los que no habían sido elegidos para formar parte del subcomplejo. Es decir que el complejo vuelve a tener la misma cantidad de puntos, pero ahora tiene algunos que han evolucionado y son distintos a los que partieron hacia el subcomplejo antes de que este fuese sometido al algoritmo de Nelder y Mead, y por lo tanto, tienen distinto valor de FO. En consecuencia los puntos están desordenados. Si por ejemplo, de un complejo de 20 individuos, se toman 13 del modo explicado para formar parte del subcomplejo, estos 13 puntos evolucionarán mediante el Algoritmo de Nelder y Mead, pero los otros 7 puntos esperarán pacientes y en su estado inicial el regreso de los 13 que partieron. Una vez que los 13 evolucionaron, se juntan con los otros 7 y vuelven a sumar 20 individuos en el complejo. Pero 13 de esos 20 individuos del complejo ya no son los de antes, por lo tanto tienen distinto valor de FO y en consecuencia es necesario volver a ordenarlos para poderlos pesar nuevamente y continuar con el proceso.

4.2.1.4. Ordenamiento de los puntos del Complejo.

Los puntos del nuevo complejo son reordenados de mejor a peor, o sea, de menor a mayor valor de la FO. Una vez reordenados, se les vuelve a asignar un peso, se vuelve a elegir un subcomplejo, y se vuelve a someter al subcomplejo al algoritmo de Nelder y Mead, y así sucesivamente un número “b” de veces. En otras palabras, se repite del paso “a” al paso “d” un número “b” de veces.

Durante la Evolución Competitiva Compleja, se han elegido y vuelto a mezclar en el complejo “b” veces subcomplejos, cada uno de los cuales fue sometido “a” veces al algoritmo de Nelder y Mead. Al cabo de la misma, el complejo ha evolucionado hacia nuevos individuos distintos a los que salieron.

4.2.2. ALGORITMO DE NELDER Y MEAD

El Algoritmo de Nelder y Mead es un método de búsqueda local que se lleva a cabo dirigiendo la evolución de un conjunto de individuos hacia el óptimo local. Una vez que el subcomplejo ha sido seleccionado del modo que se explicó previamente, será sometido a este algoritmo siguiendo los pasos que se explican a continuación.

4.2.2.1. Organización del Subcomplejo.

Los puntos del subcomplejo son organizados de mejor a peor, o sea, de menor a mayor valor de la FO. Si suponemos que nuestro subcomplejo tiene “q” individuos, entonces el primer punto será el de mejor valor de FO, y el punto q-ésimo será el peor.

4.2.2.2. Cálculo del centro de gravedad.

Se calcula el centro de gravedad de los (q-1) mejores puntos del subcomplejo, mediante la expresión:

$$g = \frac{1}{q-1} \sum_{j=1}^{q-1} u_j \quad (4.25)$$

Donde u_j es el vector posición del individuo j -ésimo, y g es el vector de coordenadas del centro de gravedad de los $q-1$ mejores puntos.

Notar que si fuesen dos parámetros del modelo, u_j sería un vector con dos coordenadas, o sea, cada punto del subcomplejo sería un punto en el plano cuya evolución podría visualizarse. Si fuesen tres parámetros, sería un subcomplejo que evoluciona en el espacio. Más de tres coordenadas no puede visualizarse, ni imaginarse, pero sí pueden realizarse proyecciones en dos dimensiones del hiperespacio, y analizar las variaciones de las funciones objetivos en estas proyecciones.

4.2.2.3. *Cálculo del punto de reflexión.*

Si trazamos un segmento que parta del peor punto del complejo, que pase por el centro de gravedad de los $(q-1)$ mejores puntos, y luego se prolongue más allá del centro de gravedad una distancia igual a la distancia que lo separa del peor punto, obtendremos un punto “ r ” equidistante al centro de gravedad con el peor punto. Ese punto así obtenido, se denomina punto de reflexión, y su vector de coordenadas de ubicación puede calcularse como:

$$r = 2g - u_q \tag{4.26}$$

4.2.2.4. *Evaluación del punto de reflexión.*

En el caso de que el punto r estuviera contenido dentro del hipercubo de validez de los parámetros, se calcula el valor de su FO y se compara con el valor de la FO del peor punto. Si es menor que aquel, el punto r reemplaza al peor, caso contrario, sigue en el paso siguiente. Ahora bien, si el punto r cayera fuera del hipercubo de validez de los parámetros, es definido el menor hipercubo que contiene todos los puntos del complejo, y se genera un nuevo punto (z) aleatoriamente en el interior de este hipercubo, el cual reemplaza al peor punto.

4.2.2.5. *Cálculo del punto de contracción.*

Como se dijo en el paso anterior, si r fuese peor que el peor punto del subcomplejo, se calculan las coordenadas de un nuevo punto “ c ” denominado punto de contracción, y sus coordenadas son calculadas de acuerdo con la ecuación:

$$c = \frac{(g + u_q)}{2} \quad (4.27)$$

Este punto está ubicado sobre el segmento que une el centro de gravedad de los $(q-1)$ mejores puntos y el peor punto, equidistante a ambos.

4.2.2.6. *Evaluación del punto de contracción.*

Se evalúa la función objetivo del punto c . Si es menor que el peor punto del subcomplejo, reemplazará a este, en caso contrario, es definido el menor hipercubo que contiene todos los puntos del complejo, y se genera aleatoriamente en su interior un nuevo punto (z).

En síntesis, el algoritmo consiste en sustituir el peor punto del subcomplejo por otro punto. Para lo cual, calcula un punto de reflexión en dirección al baricentro de los mejores puntos. Si se escapa del hipercubo, genera aleatoriamente un punto en su interior y reemplaza al peor. Si no, se analiza si el punto de reflexión es peor que el peor del subcomplejo, siendo así, genera el de contracción. Si este también es peor, define el menor hipercubo que contiene a los puntos del subcomplejo y genera aleatoriamente un punto en su interior. Pero, si el de reflexión (prioritariamente) o el de contracción, son mejores que el peor del subcomplejo, reemplaza al peor, por uno de estos.

Debe notarse que en el método descrito de búsqueda del punto óptimo, se da la oportunidad de subsistir y reproducirse a todos los individuos de la población. Por

ejemplo, en la construcción del Complejo, no se descarta ningún punto, sino que en cada complejo hay puntos buenos y malos. En la selección del subcomplejo, se da la oportunidad de salir elegido a todos los puntos del complejo, aún a los peores. De acuerdo al criterio de selección del subcomplejo, podría elegirse como reproductor al peor punto inclusive. Estos criterios, a primera vista contradictorios (ya que permiten la supervivencia de los puntos malos, cuando se buscan los buenos), tienen una propiedad muy interesante que se resume en una palabra: variabilidad. A través de este método se incursiona en todo el hiperespacio de validez de los parámetros, dando la posibilidad que el modelo se optimice localmente aún en las proximidades de los puntos malos.

4.2.3. PASOS DEL METODO

Generación de la población inicial

Se establece el rango de validez de los parámetros del modelo matemático, que definen un hipercubo de validez de los parámetros. Luego, se genera aleatoriamente una población de puntos con distribución probabilística uniforme en su interior.

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:

$$P1_{\min} \leq P1 \leq P1_{\max} \quad (4.28)$$

$$P2_{\min} \leq P2 \leq P2_{\max} \quad (4.29)$$

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 4.3.

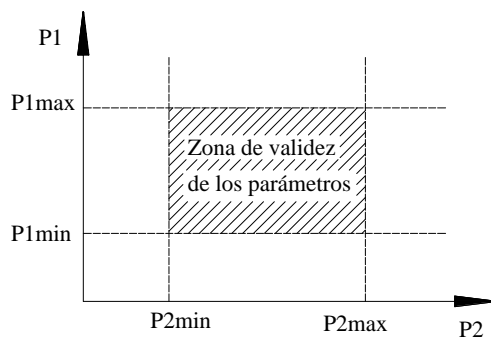


Figura 4.3. Zona de validez de los parámetros del modelo, (Alonso, 2008).

En la Figura 4.4, 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 P_1 y P_2 . Cada punto generado tiene asociado un valor de cada parámetro, así el punto B por ejemplo tiene asociado un valor de P_1 , que se ha denominado P_{1B} , y uno de P_2 , denominado P_{2B} . Como puede verse, el punto B satisface las desigualdades (4.28) y (4.29), 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 (4.28) y (4.29), por lo tanto no son puntos válidos y no pueden ser considerados en la población.

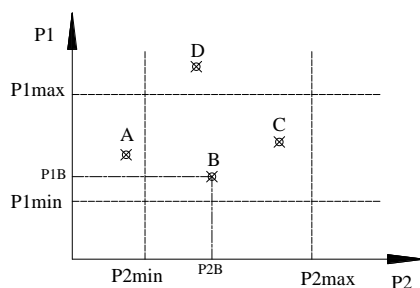


Figura 4.4. Puntos válidos y no válidos (Alonso, 2008).

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 (4.28) y (4.29), más la siguiente:

$$P3_{\min} \leq P3 \leq P3_{\max} \quad (4.30)$$

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 pueden caer los puntos aleatoriamente generados. Por ejemplo, el modelo hidrológico GR 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 GR, 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 hiperdimensional), 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 MonteCarlo.

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.

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. Luego de la evaluación resultarán puntos que tienen menor valor que otros en todas las funciones objetivos, 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, entonces no puede decirse que A es mejor que B, ni tampoco que B es mejor que A. Entonces surgen los conceptos de puntos Dominados y puntos No Dominados.

Son 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. Entonces, se eligen 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, otros 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.

De acuerdo al esquema de jerarquización antes explicado, en cada nivel habrá un conjunto de puntos. 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.

Construcción de complejos. Una vez que los puntos de la población se han jerarquizado en niveles, se les atribuye peso a cada uno de los puntos. Todos los puntos de un mismo nivel tendrán el mismo peso, pero mayor peso que los puntos de niveles inferiores. Por ejemplo, los puntos del nivel 4 tendrán mayor peso que los del nivel 5.

Luego se eligen tantos complejos como puntos tiene el último nivel. Por ejemplo, supongamos que luego de la jerarquización, se obtengan 8 niveles, y que en el nivel octavo (el último) hayan quedado 25 puntos de la población. Entonces, deben elegirse 25 complejos de la población, tales que, cada complejo debe tener uno de los puntos del último nivel y el resto de los puntos del complejo se eligen de la población aleatoriamente, pero dándole mayor probabilidad de ser elegido a los que tienen mayor peso, o sea, a los mejores. Si, por ejemplo, cada complejo debe tener 12 puntos, entonces, uno de esos será uno de los 25 correspondientes al último nivel (en nuestro ejemplo, al octavo nivel), y cada uno de los restantes 11 puntos que falta agregar, se eligen de la población en forma aleatoria, pero ponderada, dándole mayor probabilidad de ser elegido a los de mejor nivel.

Evolución independiente de cada complejo. Una vez que la muestra ha sido separada en p complejos, se da a cada complejo la oportunidad de evolucionar en forma independiente mediante la Evolución Competitiva Compleja. Con la diferencia de que no se eligen subcomplejos, sino que evolucionan los complejos mismos de acuerdo a Nelder y Mead. Se repite para cada complejo varias veces el algoritmo de Nelder y Mead. Pero como hay dos funciones objetivos el criterio de aceptación de un punto en el algoritmo de Nelder y Mead es que sea no dominado. Por ejemplo, cuando se genera el punto de reflexión, se verifica que no haya ningún punto que tenga menor valor en ambas FO; si es así, entonces se acepta el punto, puesto que es no dominado. Cuando los complejos han evolucionado independientemente, se devuelven a la población que ya está más optimizada, y se vuelve al paso 2, es decir, se vuelven a ordenar y jerarquizar, etc.

Los pasos del 1 al 4 descritos se repiten hasta que no sea posible encontrar puntos dominados. Es decir que todos los puntos de la población se encuentren en el nivel 1 del ranking de Pareto. Como consecuencia, el resultado final de este algoritmo no es un punto óptimo, como lo era el SCE, sino una población óptima de puntos que forma la región de Pareto.

5. REVISIÓN Y ACTUALIZACIÓN DE LOS DATOS DEL RIO SAN ANTONIO

En este capítulo se lleva a cabo una revisión de la información procesada con anterioridad, y se recopiló nueva información en bruto, la cual también fue procesada para obtener las series requeridas. Dado que dicha nueva información tiene una franja cronológica de superposición con la información procesada, pudo hacerse una corroboración de la calidad de estas series procesadas disponibles.

A lo largo del apartado se realiza una breve reseña de algunas herramientas del programa orientado a planillas de cálculo Excel, las cuales permitieron formar las series finales de precipitación, caudal y evapotranspiración, como así también el análisis de la serie P-Q. Seguidamente se presenta la cuenca del río San Antonio y se realiza una detallada descripción del trabajo realizado durante el procesamiento de datos de precipitación, caudal y evapotranspiración, especificando la información disponible con la que contamos para desarrollar la actualización y verificación de las series de datos.

Además se lleva a cabo un análisis de la relación causal entre precipitación y caudal, lo que nos permitirá comprobar la validez y calidad de las series conformadas.

5.1. DESCRIPCIÓN DE LAS HERRAMIENTAS DE EXCEL EMPLEADAS EN EL PROCESAMIENTO DE LAS SERIES HIDROLÓGICAS

Si bien el programa computacional Microsoft Excel es una herramienta de uso masivo, se consideró apropiado detallar en este trabajo el procedimiento abordado a los fines de converger a las series de modelación.

A continuación se mencionan algunas de las funciones más utilizadas para conformar las series de datos con paso de tiempo horario y diario.

- a. Texto en columnas: Permite dividir el texto contenido en una columna en dos o más, separándolo por el espacio presente entre palabras, por comas, barras u otro signo de ortografía.
- b. Ordenar: Como su nombre lo indica, nos posibilita ordenar la información presente en una o más columnas de acuerdo a diversos criterios como ser, de mayor a menor, de una fecha más antigua a una más reciente, etc.
- c. Función BuscarV: Es una herramienta de gran utilidad, permite buscar un valor dentro de un rango de datos, es decir, nos ayuda a obtener el valor de una tabla que coincide con el valor que estamos buscando.
- d. Función SI: Función lógica que evalúa una condición y devuelve un valor si la condición es verdadera y otro valor si la condición es falsa.
- e. Formato condicional: Formato que cambia basado en una prueba lógica, fuentes, bordes y patrones. Para una celda pueden establecerse hasta tres condiciones diferentes.
- f. Funciones de Fecha y Hora: Obtiene el número de serie que representa a la fecha y hora especificada.
- g. Pegar como valores: operación que permite copiar y pegar solo los valores de una celda, sin que se transfiera la función contenida en la misma y por lo tanto cambie el valor.
- h. Buscar: permite cambiar un valor o un signo de una celda por otro. Por ejemplo se puede cambiar un punto por una coma, una celda vacía por un guión.

5.2. DESCRIPCIÓN DE LA CUENCA

La cuenca del Río San Antonio (Figura 5.1.) se ubica a unos 30 Km al Oeste de la ciudad de Córdoba, Argentina, sobre la ladera Oriental de las Sierras Grandes (componentes del sistema de las Sierras Pampeanas de la Provincia). Geográficamente, está posicionada entre los 64° 29' 21" y 64° 51' 17" de longitud Oeste y los 31° 23' 23" y 31° 36' 18" de latitud Sur, conformando un área de 500 km² sobre el faldeo oriental de las Sierras Grandes, a una altitud promedio de 1500 m s.n.m.

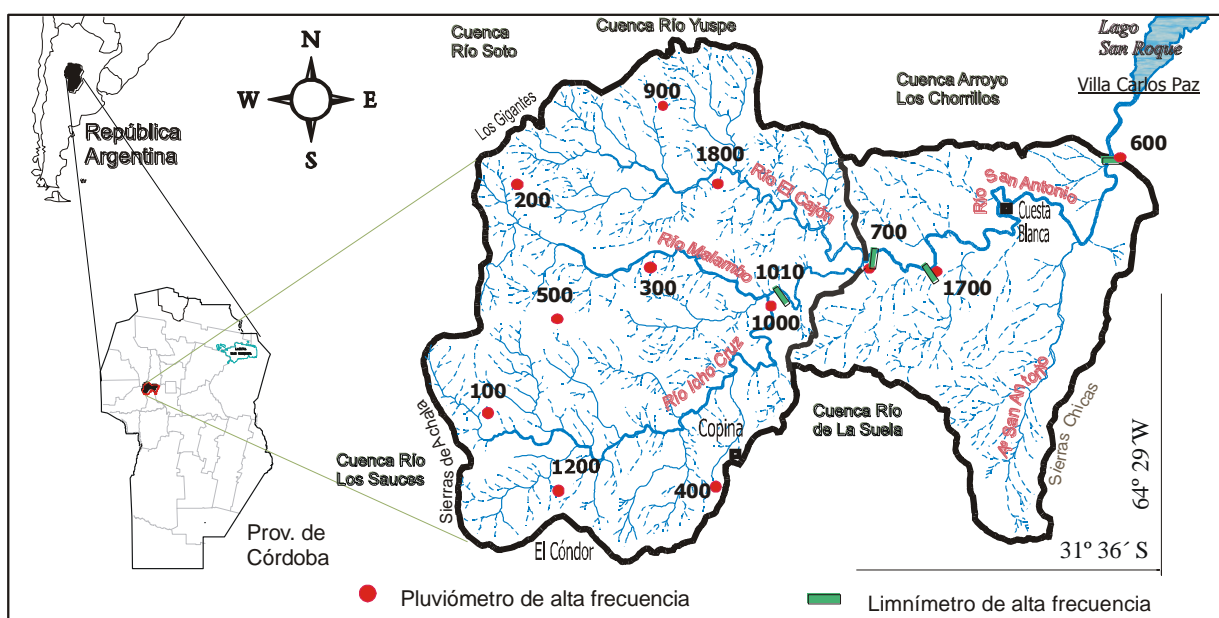


Figura 5.1. Ubicación de la cuenca del río San Antonio (CIRSA, 2003)

Constituyen sus límites naturales las Sierras Grandes al oeste, al S, por la cresta de la Sierrita o Cordón de Santiago que la separa de la cuenca del río Anizacate, elevaciones entre los cordones montañosos principales al norte y las Sierras Chicas al este. Una estación de aforo construida en Barrio El Canal (600) en la zona sur de la ciudad de Villa Carlos Paz constituye su cierre artificial, de allí el río se interna en la ciudad hasta llegar al Lago San Roque.

La Cuenca alcanza su mayor altura en el Cerro Los Gigantes de 2350 m s.n.m. en la zona NO y la salida tiene una altitud de 675 m s.n.m. Este desnivel de 1675 m se alcanza en apenas 29,7 Km.

Su principal colector, el río San Antonio, está formado por los ríos Icho Cruz, Malambo y El Cajón. El San Antonio es uno de los tributarios del Lago San Roque junto con el Río Cosquín, los Arroyos Las Mojarras y Los Chorrillos. Forma parte del Sistema del Río Suquia que nace a la salida del Embalse San Roque y desagua en el Mar de Ansenúza (Laguna de Mar Chiquita) al NE de la provincia.

La cuenca cuenta con un área de aporte de 328 Km², en la sección de aforo de la estación 700, que se encuentra en la confluencia de los ríos Icho Cruz y El Cajón, como lo muestra la Figura 5.1. En la estación 600, el área de aporte asciende a 492 Km², y finalmente a 509 Km² en su desembocadura en el lago San Roque.

La cuenca del río San Antonio, se encuentra monitoreada en forma remota, (desde inicios de la década de 1990) por una red hidrometeorológica compuesta de 13 estaciones automáticas, las cuales son capaces de medir distintas variables hidrometeorológicas a tiempo real, como ser: lámina de lluvia precipitada, niveles de agua en ríos, temperatura del aire, humedad relativa ambiente, velocidad y dirección del viento, presión barométrica, radiación solar incidente y evaporación entre otros. Esta información histórica disponible cuenta con una muy buena calidad y densidad espacial de los registros. Los pluviómetros de alta frecuencia (PAF) con los que cuentan estas estaciones (Tabla 5.1 y Figura 5.2), registran series de tiempo de los valores unitarios de milímetros precipitados en eventos lluviosos. Estos datos son transmitidos en tiempo real al centro del INA-CIRSA emplazado en la ciudad de Villa Carlos Paz.

Tabla 5.1. Pluviómetros de alta frecuencia instaladas en la cuenca del río San Antonio, denominación, emplazamiento y altitud relativa (Fuente: INA-CIRSA).

Código en el Sistema	Denominación de la Estación	Latitud [WSG1984]	Longitud [WSG1984]	Altitud [msnm]
100	Santo Tomas	-31,57	-64,83	2.250
200	Puesto Pereyra	-31,45	-64,81	2.292
300	La Casita	-31,48	-64,74	1.555
400	Cañada Larga	-31,58	-64,71	1.740
500	El Galpón	-31,51	-64,82	2.380
600	Bo, El Canal	-31,40	-64,47	675
700	Confluencia El Cajón	-31,49	-64,64	760
900	Puesto Garay	-31,42	-64,74	1.625
1000	Confluencia Malambo	-31,50	-64,68	1.340
1200	Las Ensenadas	-31,60	-64,79	2.286
1700	La Quebrada	-31,49	-64,60	918
1800	El Cajón	-31,44	-64,69	1.280
1900	Las Jarillas	-31,53	-64,55	838

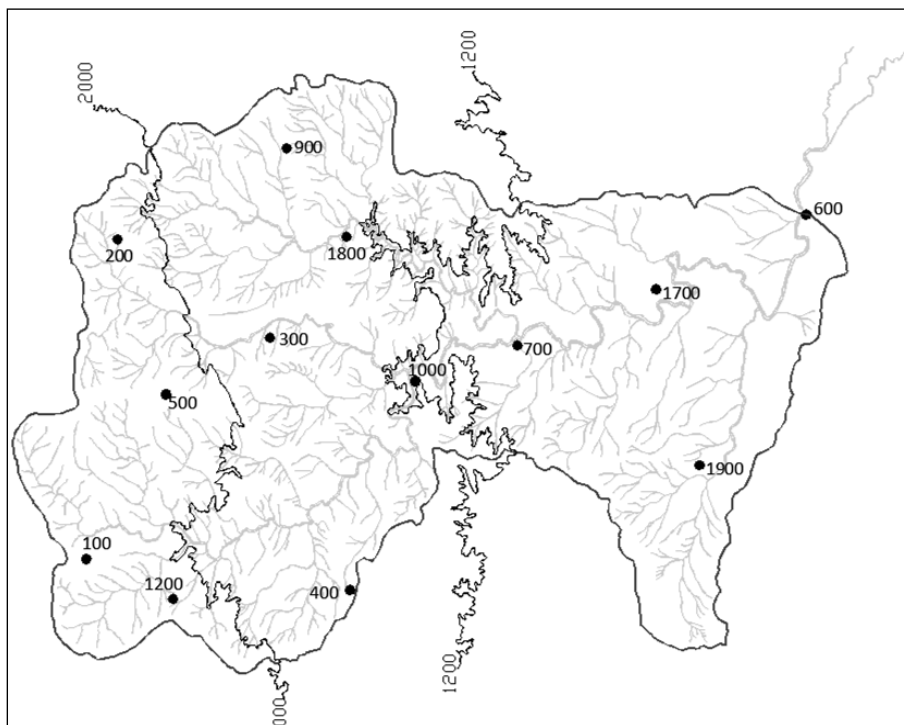


Figura 5.2. Red telemétrica de la del Río San Antonio.

Con respecto a la situación actual del pronóstico hidrológico en la cuenca del río San Antonio, el INA-CIRSA lo lleva a cabo mediante una formulación matemática polinómica muy sencilla (Colladón et al., 2007), cuyos coeficientes son ajustados estadísticamente. Mediante este modelo se da respuesta a las necesidades más imperiosas de la institución, manifestando, sin embargo, la necesidad de un método más preciso para tan delicada tarea. Este modelo está representado por dos ecuaciones: una que estima la altura pico de la crecida y otra el tiempo de llegada, ambas necesitan información de la estación 700 “Confluencia con El Cajón” para poder brindar estos datos en la estación 600 “Barrio El Canal”. Es por ello, junto con otras razones, que la densidad de estaciones es mayor en la denominada “Cuenca Alta”. Actualmente existe mayor densidad de población en el tramo que hay desde la estación 700 a la 600 y no se cuenta con ninguna herramienta para predecir la crecida a lo largo de este tramo. A esto se agrega la ocurrencia (de baja probabilidad) de precipitaciones con su núcleo temporal centrado en la cuenca baja, que al haber pocas estaciones para la toma de datos dificulta la previsión de la crecida.

El clima es del tipo semiárido. Presenta uniformidad térmica con concentración en la temporada estival de altas temperaturas y precipitaciones. La precipitación media anual oscila los 850 mm, con una concentración de lluvias torrenciales en el período de octubre a marzo (Rodríguez, 2004).

Una de las características que le otorga mayor importancia al Río San Antonio es la de constituirse en el principal tributario del Lago San Roque. Este lago es actualmente la principal fuente de agua potable para la Ciudad de Córdoba.

Otra característica atendible está representada por la peligrosa combinación de una importante afluencia turística a las costas de este río, junto con importantes, violentas y rápidas crecidas en la misma época veraniega. La cuenca del río San Antonio posee fuertes pendientes, un punto de salida bien definido, clara divisoria de aguas, baja permeabilidad y altos índices de escorrentía. La combinación de los factores anteriormente expuestos con las tormentas intensas generadas en el período estival, da como resultado la generación de hidrogramas de crecidas repentinos (tiempo de concentración pequeño), de altos valores de caudal pico y de corta duración. Estas crecidas se producen normalmente en la época del pico turístico en esta zona del valle de Punilla.

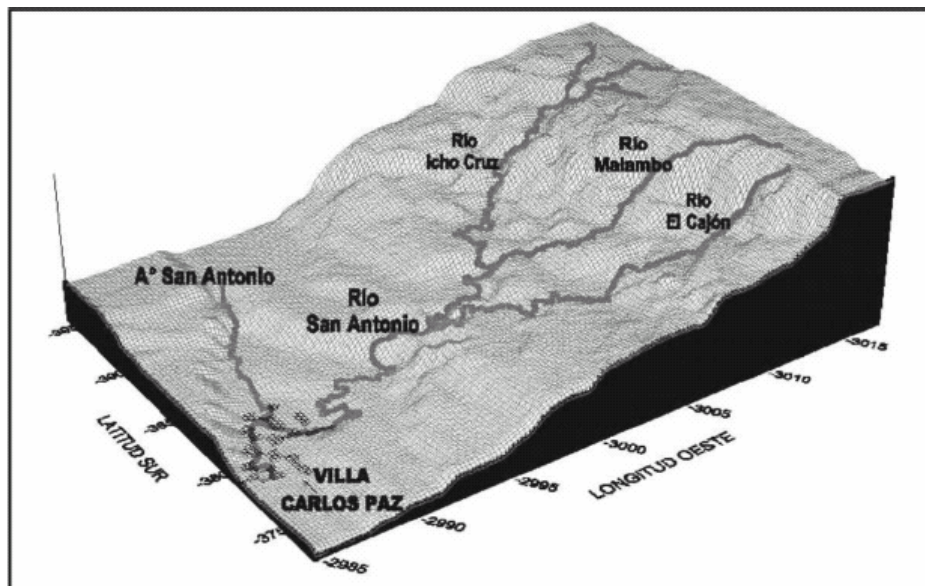


Figura 5.3. Perfil de la cuenca y tributarios del Río San Antonio (CIRSA, 2003)

Se presenta a continuación una síntesis de las características principales de la cuenca analizada.

- Superficie: 509 Km.
- Precipitación media anual de 850 mm.
- Fuertes pendientes.
- Baja permeabilidad.
- Alto índice de escorrentía.

Las últimas tres características tiene como consecuencia bajos tiempos de concentración, lo que significa una rápida respuesta de la cuenca ante las precipitaciones, desarrollándose hidrogramas de forma casi instantánea. Por tal motivo se decidió adoptar un paso de tiempo horarios para analizar la misma.

5.3. PRECIPITACIÓN

5.3.1 INFORMACIÓN DISPONIBLE

En este trabajo se contó con información proveniente de diversas fuentes, que se describen a continuación:

- a. Información original extraída directamente del sistema, es decir, la transmitida por las estaciones pluviométricas en tiempo real al centro del INA-CIRSA. Dicha información fue provista en su formato original (es decir, tal como los arroja el sistema de acopio de información que recepta las señales de las estaciones) por los ingenieros Carlos Catalini y César García, incluyendo:
 1. Archivos de precipitaciones entre los años 2006 hasta el 2014 de doce de las trece estaciones que posee la cuenca.
 2. Archivos de precipitaciones del año 1991 sólo para la estación 300.
 3. Archivos de precipitación del año 1992 al 2011 para la estación 200.

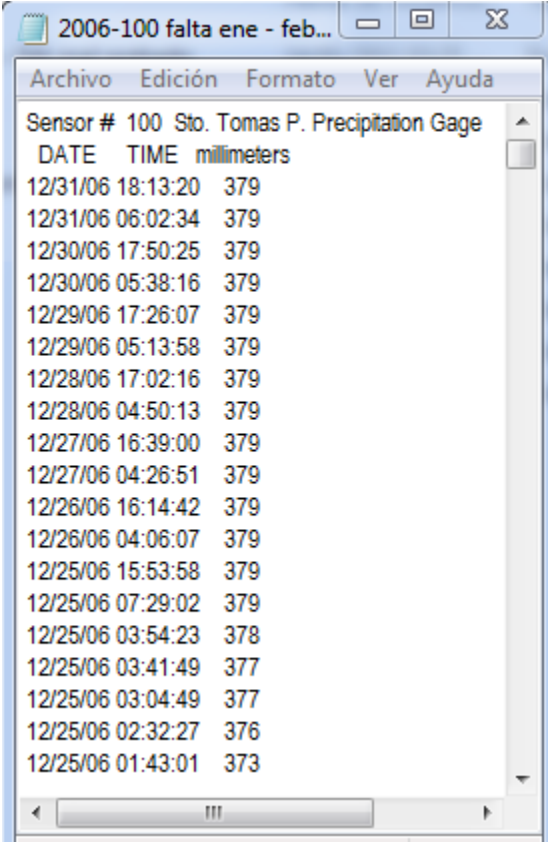
Cabe remarcar que los registros poseen notorias lagunas de información que comprenden desde algunas horas hasta varios meses. Una descripción de la principal información faltante en cada archivo de datos para cada estación se presenta en la Tabla 5.2. Adicionalmente, es importante señalar que los datos se presentan en estos archivos con pasos de tiempo variables, y con ausencia de información en varias horas del día. De modo que (por poner un ejemplo) en las 24 hs de un día cualquiera, puede haber registrado sólo 15 datos, correspondientes a horas salteadas, y a veces varios registros dentro de una misma hora. Esto puede comprenderse mejor observando la ilustración de la Figura 5.4.

También puede apreciarse en la misma que la precipitación se presenta de forma acumulada. En conclusión, a los fines de llevar esta información al esquema de paso de tiempo horario, fue necesario acumular en cada hora los valores registrados en ese paso de tiempo (en aquellos donde hay más de uno por hora), y también completar los pasos de tiempo con ausencia de información, todo lo cual se explica más adelante.

Tabla 5.2. Descripción de la Información pluviométrica faltante en el periodo de registro por año y por cada estación.

ESTACION	AÑO	DESCRIPCION
ESTACION 100	2006	No hay datos desde (1/01/2006 al 7/03/2006).
	2008	01/01 al 04/01 no hay datos.
	2009	No hay datos desde el 1/04 al 09/04 inclusive.
ESTACION 200	2006	Falta diciembre.
	2007	Falta enero.
	2008	Falta enero (01 al 30 inclusive) Y febrero (23/02 al 13/03) no funciona, reporta a partir del 14/2 (hay datos, pero son dudosos).
	2010	Falta marzo y mayo.
ESTACION 300	2006	Falta diciembre.
	2011	Falta diciembre.
	2012	Falta enero y marzo (comienza el 12 de abril); febrero incompleto (entrada errática).

ESTACION 400	2012	Falta septiembre (comienza 12/10), julio y agosto entrada errática.
ESTACION 500		-
ESTACION 600	2006	Falta noviembre.
	2010	Falta mayo, julio y agosto (recomienza el 10/09).
	2014	Pocos datos de julio y agosto.
ESTACION 700	2007	Falta abril (recomienza el 8/05) y julio.
	2008	Mayo (desde el 05/05 hasta 1/06) y agosto (desde el 1/08 al 15/08) incompletos.
	2011	Entre 02/10 y 15/11 defectos de la estación.
	2014	Falta marzo, mayo.
ESTACION 900	2007	Julio falta.
	2008	Mayo incompleto.
	2012	Agosto incompleto, septiembre falta, octubre (comienza el contador el 11/10).
ESTACION 1000	2007	Julio falta.
	2008	Faltan datos.
	2010	Julio falta.
	2012	En febrero entrada errática, comienza a funcionar a partir del 5 de marzo.
ESTACION 1200	2006	Abril falta.
	2007	Agosto falta.
	2008	Enero incompleto.
ESTACION 1700	2006	Enero, febrero, julio y noviembre faltan.
	2008	Faltan datos desde el 26 de junio al 2 de septiembre.
	2010	Faltan datos desde 21 de junio hasta 22 de septiembre. Diciembre incompleto.
	2011	Enero comienza el 27, marzo 7 a septiembre 15 falta, falta octubre, noviembre y diciembre 29.
	2012	Junio, julio y noviembre faltan.
ESTACION 1800	2006	Junio falta.
	2009	Febrero falta.
	2011	Desde 22 de junio a 16 de noviembre no hay datos.
	2012	Enero no hay datos desde el 21 al 26; febrero falta.
	2013	Desde 1 de marzo al 18 de marzo no hay datos; septiembre del 17 al 27 no hay datos.



Sensor #	100	Sto. Tomas P. Precipitation Gage
DATE	TIME	millimeters
12/31/06	18:13:20	379
12/31/06	06:02:34	379
12/30/06	17:50:25	379
12/30/06	05:38:16	379
12/29/06	17:26:07	379
12/29/06	05:13:58	379
12/28/06	17:02:16	379
12/28/06	04:50:13	379
12/27/06	16:39:00	379
12/27/06	04:26:51	379
12/26/06	16:14:42	379
12/26/06	04:06:07	379
12/25/06	15:53:58	379
12/25/06	07:29:02	379
12/25/06	03:54:23	378
12/25/06	03:41:49	377
12/25/06	03:04:49	377
12/25/06	02:32:27	376
12/25/06	01:43:01	373

Figura 5.4. Información proveniente directamente de la red telemétrica del INA-CIRSA.

- b. Series con paso de tiempo horario proporcionadas por el Dr. Ing. Cesar García, quien tomó la información original del CIRSA, y la procesó empleando la herramienta de manejo de datos on-line en la página <http://leutra.geogr.uni-jena.de/cirsaRBIS>. Las series trabajadas por Cesar García incluyen la lámina de precipitación acumulada por hora (Figura 5.5.). La información comprende un periodo de 15 años de duración, el cual inicia el 1 de enero de 1994 y finaliza el 31 de enero de 2009. Durante este tiempo se presentan periodos sin registros entre las fechas que se señalan a continuación:

Tabla 5.3. Periodos sin información de la serie de Cesar Garcia.

DESDE	HASTA
28/02/1995	06/03/1995
01/09/2005	30/11/2005
31/10/2006	01/12/2006
31/12/2007	04/01/2008

Al igual que en el caso anterior la información se recibió en formato de texto (.txt).

2006-03-08 02:00	0
2006-03-08 03:00	0
2006-03-08 04:00	0
2006-03-08 05:00	2
2006-03-08 06:00	0
2006-03-08 07:00	0
2006-03-08 08:00	0
2006-03-08 09:00	1
2006-03-08 10:00	7
2006-03-08 11:00	0
2006-03-08 12:00	0
2006-03-08 13:00	0
2006-03-08 14:00	0
2006-03-08 15:00	1

Figura 5.5. Información disponible con paso de tiempo horario.

- c. Información recopilada y procesada en el trabajo de Alonso (2012), y organizada en planillas Excel en el trabajo de Bustos (2014). La misma abarca el periodo comprendido entre los años 1994 y 2013, presentando el registro de 10 de las 13 pluviómetros. Los datos se presentan con un paso de tiempo horario y además se dispone de la precipitación media areal durante el periodo señalado. Tabla 5.4. Ejemplo de presentación de la Información organizada por Bustos (2014)

Tabla 5.4. Ejemplo de presentación de la Información organizada por Bustos (2014)

Día	Mes	Año	Hora	Est 1800	Est 1200	Est 1000	Est 900	Est 700	Est 600	Est 500	Est 400	Est 300	Est 200	Precip promedio de la cuenca
				mm acum x hora	mm acum x hora	mm acum x hora	mm acum x hora	mm acum x hora	mm acum x hora	mm acum x hora	mm acum x hora	mm acum x hora	mm acum x hora	
1	1	1994	02:00:00	-	-	-	-	-	-	-	0	-	-	0.00
1	1	1994	03:00:00	-	-	-	-	-	-	-	0	-	-	0.00
1	1	1994	04:00:00	0	-	-	-	-	-	-	0	-	-	0.00
1	1	1994	05:00:00	0	-	-	-	-	-	-	0	-	-	0.00
1	1	1994	06:00:00	0	0	-	-	0	-	-	0	0	-	0.00
1	1	1994	07:00:00	0	0	-	-	0	-	-	0	0	-	0.00
1	1	1994	08:00:00	0	0	-	-	0	-	-	0	0	-	0.00
1	1	1994	09:00:00	0	0	-	-	0	-	-	0	0	-	0.00
1	1	1994	10:00:00	0	0	-	-	0	-	-	0	0	-	0.00
1	1	1994	11:00:00	0	0	-	-	0	-	-	0	0	-	0.00
1	1	1994	12:00:00	0	0	0	0	0	-	-	0	0	-	0.00
1	1	1994	13:00:00	0	0	0	0	0	-	-	0	0	-	0.00
1	1	1994	14:00:00	0	0	0	0	0	-	-	0	0	-	0.00
1	1	1994	15:00:00	0	0	0	0	0	-	-	0	0	-	0.00
1	1	1994	16:00:00	0	0	0	0	0	-	-	0	0	-	0.00
1	1	1994	17:00:00	0	0	0	0	0	-	-	0	0	-	0.00
1	1	1994	18:00:00	0	0	0	0	0	0	-	0	0	-	0.00
1	1	1994	19:00:00	0	0	0	0	0	0	-	0	0	-	0.00
1	1	1994	20:00:00	0	0	0	0	0	0	-	0	0	-	0.00
1	1	1994	21:00:00	0	0	0	1	0	0	-	0	0	-	0.10
1	1	1994	22:00:00	0	0	0	7	0	0	-	0	0	-	0.67
1	1	1994	23:00:00	0	0	0	0	0	0	-	0	0	-	0.00
2	1	1994	00:00:00	0	0	3	0	1	0	-	1	0	-	0.56

5.3.2. CONVERSION DE LA INFORMACION RECIBIDA A SERIES HORARIAS

Como se explica en el apartado anterior, se contó con información en bruto proveniente de las estaciones, pero también con información procesada (como es la provista por García y Alonso). En consecuencia, se procedió a trabajar los registros originales, y luego emplear la información procesada disponible como contraste, a los fines de verificar que se trabajó adecuadamente.

La decisión de procesar nuevamente la información original, a pesar de disponer de documentos ya procesados, se debe a dos razones:

- i) la info procesada está incompleta, es decir, es más corta la serie que la información original que hoy en día se dispone, y
- ii) se desconoce el método empleado para obtener algunos de los tramos de las series procesadas.

A los fines entonces de llevar la información base al paso de tiempo horario se llevó adelante el siguiente procedimiento:

- 1- Como primer paso para obtener las series horarias de cada estación, se transformaron los documentos de texto (.txt) en planillas Excel (.xlsx) para poder trabajar la información en el programa Microsoft Excel, que resulta una herramienta esencial para llevar a cabo el procesamiento. Es un procedimiento muy sencillo, que consiste en copiar los datos del documento de texto en una columna de Excel, luego se divide el contenido de la misma en tres columnas, fecha, hora y precipitación acumulada haciendo uso de la función ya explicada, "texto en columna".
- 2- A su vez, las tres columnas (fecha, hora y precipitación acumulada) debieron ser tratadas ya que, por un lado la fecha estaba escrita en un formato no reconocido por Excel (primero aparece el mes, luego el día y por último el año), la precipitación acumulada se restó celda a celda para obtener la lámina caída en

cada minuto y por último a las horas se les debió retirar los minutos y segundos, ya que necesitamos un paso de tiempo horario. Para esto se volvió a utilizar la función denominada “texto en columna” y además la función “fecha”.

- 3- En una tercera instancia se acumuló en cada hora los milímetros llovidos durante la misma. Aquí fue necesario utilizar la función condicional “SI” a causa de que el texto base no posee un salto definido, es decir, no en todas las horas hay la misma cantidad de registros, sino que es variable. La herramienta fue utilizada de la siguiente manera:

Supongamos que la columna fecha se denomina F, la de hora H, la de precipitación P y la de precipitación acumulada PA, y a la filas se las designan con la letra i. Para ir acumulando las precipitaciones ocurridas en una misma hora como ya se dijo anteriormente se utilizó la función “SI” en la columna de precipitación acumulada. Si F_{i+1} es igual a F_i , y además H_{i+1} es igual a H_i quiere decir que nos encontramos en el mismo día y hora, por lo que las precipitaciones deben sumarse, por lo cual PA_{i+1} es igual a la suma de PA_i y P_{i+1} . En caso de que alguna de las dos condiciones que se plantearon no se cumplan, el valor de PA_{i+1} es igual a P_{i+1} . Se ejemplifica a continuación lo explicado.

Tabla 5.5. Ilustración del modo de empleo de la función "SI".

ORDEN	FECHA	HORA	PRECP	PREC. ACUM. POR HORA
1	07/03/2006	15	0	0
2	08/03/2006	2	0	0
3	08/03/2006	5	1	1
4	08/03/2006	5	1	2
5	08/03/2006	9	1	1
6	08/03/2006	10	1	1
7	08/03/2006	10	1	2
8	08/03/2006	10	1	3
9	08/03/2006	10	1	4
10	08/03/2006	10	1	5
11	08/03/2006	10	1	6
12	08/03/2006	10	1	7
13	08/03/2006	15	1	1

Como se puede observar en la **¡Error! No se encuentra el origen de la referencia.** la serie aun posee varias celdas del mismo día y hora, por lo que para concluir con la acumulación de la precipitación es necesario eliminar las filas que no contienen la acumulación total de la precipitación de la hora correspondiente. En el caso de la imagen se deben eliminar la fila 3 y desde la 6 a la 11 inclusive. Para esto se hizo uso de la herramienta de Excel “buscarV” la cual, como ya se explicó en un apartado anterior, nos permite encontrar un número dentro de una tabla, y si lo halla colocar el valor de la columna indicada que le corresponde a ese número. Ahora bien, nosotros lo que queremos es formar una nueva serie en el que cada día y hora tenga un único valor de precipitación y sea el acumulado total. Por ejemplo, en el caso de la imagen anterior, para el día 08/03/2006 a las 10 hs el valor de precipitación que debe figurar es 7. Como paso previo a la utilización de la función “buscarV” se conformó una lista completa con todos lo día y horas desde el 01/01/2006 al 31/12/2014, luego se debió buscar la forma de identificar con un único valor la fecha y la hora, esto se llevó a cabo cambiando el formato de la fecha a número (en Excel, cada fecha tiene asociado un número), se multiplicó dicho valor por 100 y se le sumo las horas, quedando así formado un número en el que los dos últimos dígitos corresponden a la hora y los anteriores a la fecha. En la columna siguiente se realiza la función. Antes de ello, se debió realizar la fusión de la fecha y hora de la serie que se observa en Tabla 5.5, y además ordenarla a partir de la fecha (de más antiguo a más reciente), de la hora (de menor a mayor), y de la precipitación acumulada por hora (de mayor a menor), esto es así ya que la función “buscar” selecciona el primer valor que encuentra buscando de arriba hacia abajo.

Tabla 5.6. Ilustración conformación serie horaria de precipitación.

FECHA	FECHA	HORA	FECHA + HORA	PRECIPITACION
08/03/2006	38784	2	3878402	0
08/03/2006	38784	3	3878403	-
08/03/2006	38784	4	3878404	-
08/03/2006	38784	5	3878405	2
08/03/2006	38784	6	3878406	-
08/03/2006	38784	7	3878407	-
08/03/2006	38784	8	3878408	-
08/03/2006	38784	9	3878409	1
08/03/2006	38784	10	3878410	7
08/03/2006	38784	11	3878411	-
08/03/2006	38784	12	3878412	-
08/03/2006	38784	13	3878413	-
08/03/2006	38784	14	3878414	-
08/03/2006	38784	15	3878415	1

En la tabla anterior se observan que hay horas que no se encontraron en la primera serie conformada.

Por último se procedió a realizar el relleno de la serie, completando con cero las celdas vacías solo si el periodo faltante es menor a 2 días, ya que se considera que durante aquellas horas no hay registro por falta de lluvia. Cuando el rango supera las 48 hs se las deja vacías ya que podría ocurrir que el sensor haya dejado de funcionar por algún motivo.

5.3.3. COMPARACIÓN DE LAS SERIES OBTENIDAS CON LAS DISPONIBLES

Como se expresó previamente, las series de datos de precipitación procesada disponible (de los trabajos ya citados de César García y Facundo Alonso) presentan para todas las estaciones una franja de superposición con respecto a la procesada en este trabajo que abarca desde 2006 hasta 2009.

Esto permitió para cada una de las estaciones realizar una comparación de la serie obtenida con las disponibles. En primer lugar, se pudo apreciar que las series provistas por

García y las obtenidas del trabajo de Alonso, coinciden perfectamente, de donde se deduce que puede tratarse del mismo origen de los datos.

En segundo lugar, estas series antecedentes también presentaron mucha semejanza con las obtenidas en este trabajo. Sólo en pocos casos se manifestaron diferencias. En aquellos casos en los cuales no coincidieron se pudo distinguir dos patrones recurrentes:

- a- Si bien al paso de tiempo horario que fueron conformadas las series se encontraban diferencias, el valor acumulado del día completo proporcionaba el mismo valor en ambas series.
- b- La ausencia de un dato en una serie, se correspondía con la disponibilidad de datos en la otra serie, y viceversa.

A la hora de conformar las series definitivas se procedió a adoptar el valor de una u otra serie a partir de los siguientes criterios respectivamente:

- a- Dado que el valor de precipitación diaria era el mismo, se cuenta con un balance de masa coherente, lo cual permitió emplear los valores de una serie para reproducirse en la otra. En este sentido, se optó por seleccionar los valores que presentaran una mayor distribución en el tiempo, es decir, si por ejemplo en una de las series se lee que en una hora llovió 1 mm, 11 mm en otra y en la segunda serie se observa que cayó una lámina de 3 mm en la primer hora y 9 mm en la siguiente, se decide entonces por seleccionar la segunda opción, ya que es la menos concentrada.

FECHA	HORA	PRECIPITACIÓN OBTENIDA	PRECIPITACIÓN FACUNDO ALONSO	PRECIPITACIÓN ADOPTADA
20/03/2008	14	0	3	3
20/03/2008	15	1	9	9
20/03/2008	16	11	0	0

- b- Mientras que en los casos en los que una columna contenía información y la otra se encontraba vacía, se procedió a seleccionar la serie con información, ya que resulta poco probable que dichos valores sean inventados y es más razonable que dichos datos no hayan sido registrada en la información original con la que se trabajó previamente a obtener las series.

Hubieron, adicionalmente algunos casos puntuales donde además de no coincidir los valores en paso de tiempo horario, tampoco había equidad en el valor diario global de esos eventos. En esos casos, se procedió a contar con información de precipitación con paso de tiempo diario conocido y tomarlo como referencia. Es decir, se asumió que el valor de precipitación global en el día en cuestión no tiene porqué corresponder necesariamente a ninguna de las series, sino que se toma como cierta otra información que arroja series de precipitaciones con paso de tiempo diario en la cuenca del río San Antonio. Dicha información corresponde al Informe Estadístico Anual del CIRSA, provisto por la Ing. Laura Colladón.

5.3.4. CONFORMACION DE LA SERIE DE PRECIPTACIÓN MEDIA AREAL DE LA CUENCA

Los aparatos como los pluviógrafos o pluviómetros, registran la lluvia puntual, es decir la que ocurre en el punto en que está instalado el aparato y para los cálculos ingenieriles, es necesario conocer la lluvia media en toda la cuenca.

Por tal motivo, una vez comparadas las series de lluvia se procedió a obtener la precipitación media areal de la cuenca desde 1994 al 2014. Para esto se comenzó uniendo la precipitación horaria de cada estación actualizada desde el 2006 al 2014 a la obtenida con anterioridad por el Ing. Alonso. Además se debieron crear nuevas combinaciones de estaciones para ser sumadas a las 64 ya realizadas, teniéndose actualmente 164 conjuntos de polígonos de thiessen. Obsérvese algunas de las nuevas combinaciones realizadas en las Figura 5.6, Figura 5.7 y Figura 5.8.

Se presentaron casos en los que no había registros en ninguna estación o tramos con menos de 3 estaciones con datos. Se decidió entonces hacer el relleno en dichas horas para que la precipitación media sea representativa de toda la cuenca y no solo de aquellos escasos sectores con información. Puede verse en la Tabla 5.7. la precipitación promedio de la cuenca determinada a partir de las series de precipitación sin rellenar. En la última columna, “cantidad de estaciones”, se señalan aquellas horas que presentan menos de 3 estaciones disponibles, las cuales posteriormente deberán ser rellenadas. Ver Tabla 5.8.

La manera de trabajar, entonces, fue la siguiente:

- a- Con las precipitaciones horarias de cada estación y junto a los polígonos se obtiene la precipitación media areal.
- b- En aquellas horas en la que la precipitación media areal fuera obtenida a partir de 3 o menos estaciones (o ninguna) se procede a rellenar dichos tramos con los milímetros caídos en un día dividido por la cantidad de horas en las que no hay datos. Esto fue realizado para cada una de las estaciones.
- c- Se determina la precipitación promedio de la cuenca en los tramos anteriores y se los acopla con la serie de completa.

Se realizo de esta manera ya que se consideró más conveniente rellenar los datos cuando hay pocas estaciones con registros que completar las series en su totalidad, ya que de esta forma se estaría modificando la serie con datos que no tienen suficiente confiabilidad (la precipitación diaria se obtiene a partir de registros de personas particulares que se encuentran en las cercanías de la estación) en casos en los que no se es necesario ya que hay cantidad suficiente de estaciones para calcular la P.M.A.

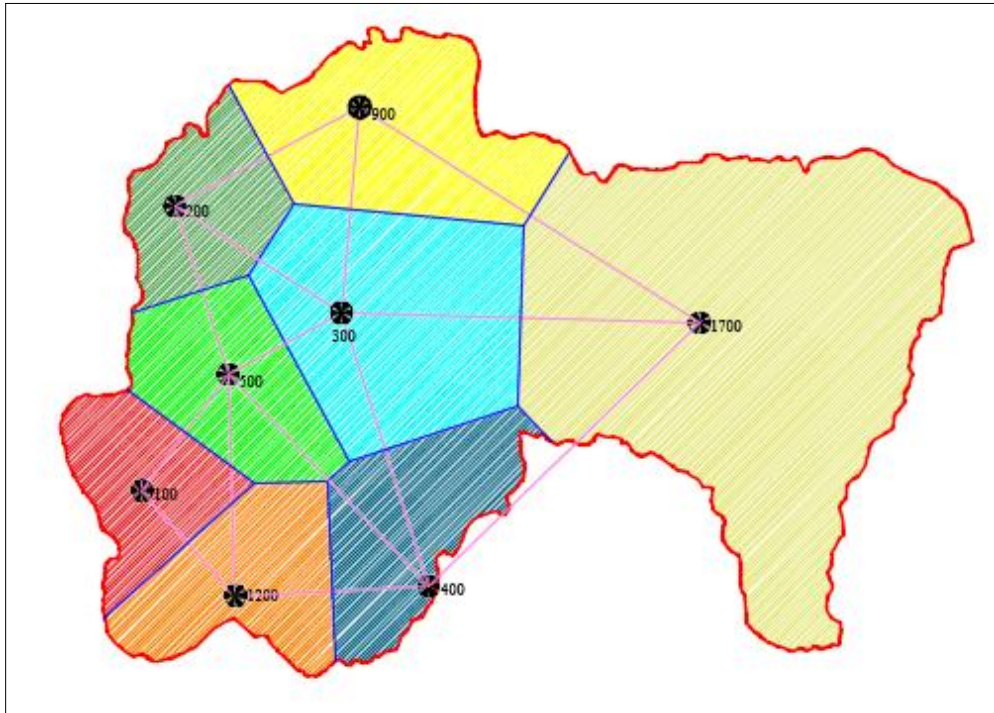


Figura 5.6. Polígono de Thiessen N° 93.

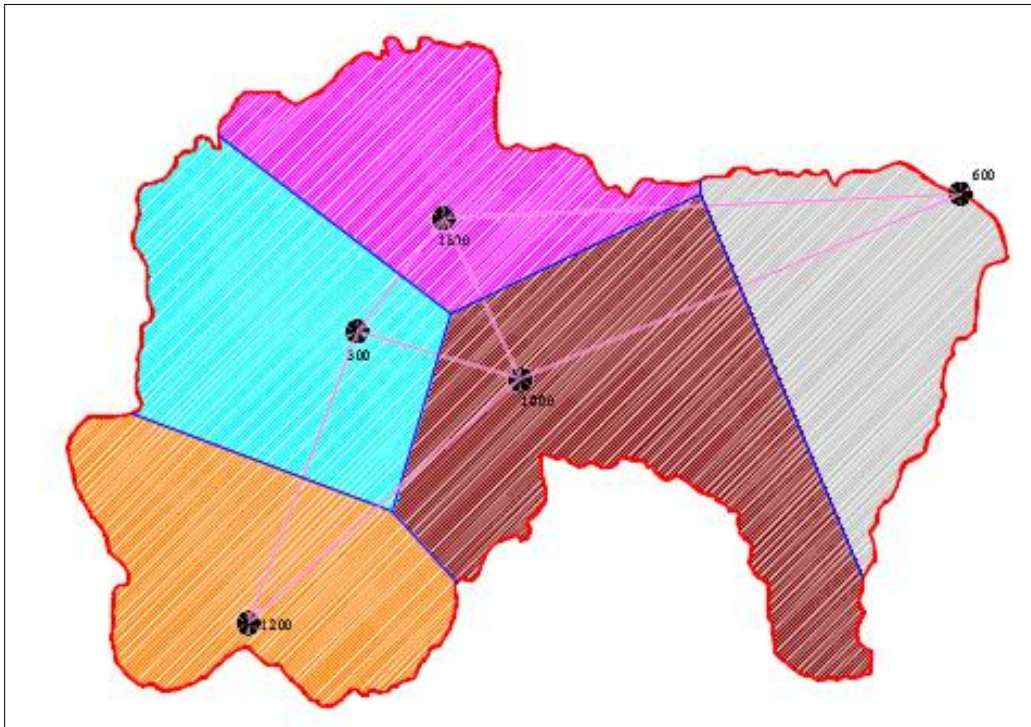


Figura 5.7. Polígono de thiessen N° 123.

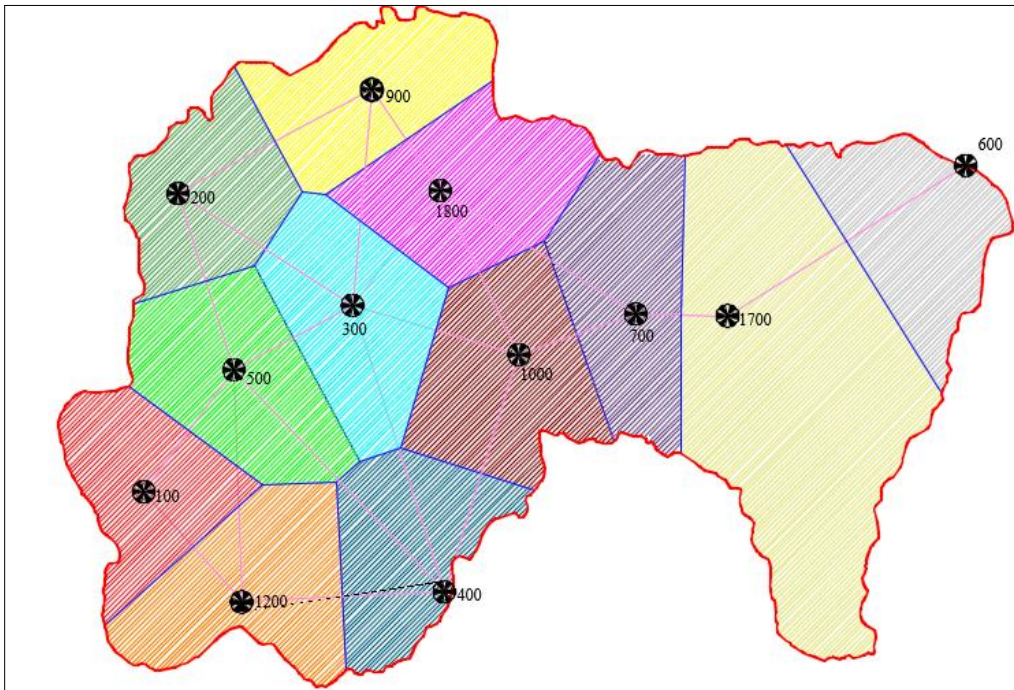


Figura 5.8. Polígono de thiessen N°65.

Tabla 5.7. Precipitación media areal sin rellenar series.

PRECIPITACION														PRECIPITACION MEDIA AREAL	CANTIDAD DE ESTACIONES	
FECHA	HORA	Est. 1800	Est. 1700	Est. 1200	Est. 1000	Est. 900	Est. 700	Est. 600	Est. 500	Est. 400	Est. 300	Est. 200	Est. 100			
30/11/2005	22:00:00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
30/11/2005	23:00:00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
01/12/2005	00:00:00	0	-	-	-	-	-	-	-	-	-	-	-	0	1	
01/12/2005	01:00:00	0	-	-	-	-	-	-	0	-	-	-	-	0	2	
01/12/2005	02:00:00	0	-	-	-	0	-	-	0	-	-	-	-	0	3	
01/12/2005	03:00:00	0	-	-	-	0	0	-	0	-	-	-	-	0	4	
01/12/2005	04:00:00	0	-	-	-	0	0	0	0	-	0	1	-	0.062001673	7	
01/12/2005	05:00:00	1	-	0	-	1	0	1	1	-	0	0	-	0.371904358	8	

Tabla 5.8. Precipitación media areal con relleno de series.

RELLENO DE CELDAS														PRECIPITA CION MEDIA AREAL	CANTID AD DE ESTACI ONES
PRECIPITACION															
FECHA	HORA	Est. 1800	Est. 1700	Est. 1200	Est. 1000	Est. 900	Est. 700	Est. 600	Est. 500	Est. 400	Est. 300	Est. 200	Est. 100		
30/11/2005	22:00	0.042	-	0.167	0.083	0.125	0.083	0.042	0	0.1667	0.125	0.167	-	0.0916591	10
30/11/2005	23:00	0.042	-	0.167	0.083	0.125	0.083	0.042	0	0.1667	0.125	0.167	-	0.0916591	10
01/12/2005	00:00	0	-	0	0.167	0	0	0	0	0.1667	0	0	-	0.0241909	10
01/12/2005	01:00	0	-	0	0.167	0	0	0	0	0.1667	0	0	-	0.0241909	10
01/12/2005	02:00	0	-	0	0.167	0	0	0	0	0.1667	0	0	-	0.0241909	10
01/12/2005	03:00	0	-	-	-	0	0	-	0	-	-	-	-	0	4
01/12/2005	04:00	0	-	-	-	0	0	0	0	-	0	1	-	0.0620167	7
01/12/2005	05:00	1	-	0	-	1	0	1	1	-	0	0	-	0.3719043	8

5.4. CURVA H-Q

Debido a que las series de caudales son poco habituales es necesario contar con una curva H-Q, la que permite obtenerlos a partir de una serie de niveles. Dicha curva se obtiene en base a niveles medidos y de algunos aforos de caudales realizados en campaña de aforos específicos en secciones estables del río.

5.4.1. FUENTES DE INFORMACIÓN

Si bien el CIRSA brindada una curva de transformación, durante el desarrollo de su PS M. J. Martín, confeccionó y analizó 4 curvas H-Q, con el fin de obtener un mejor ajuste de los caudales transformado. Para ello se utilizaron datos de 725 aforos de río San Antonio realizados entre 1969 y 1995. En el siguiente gráfico (Figura 5.9.) se representa en abscisas los niveles y en ordenadas los caudales aforados.

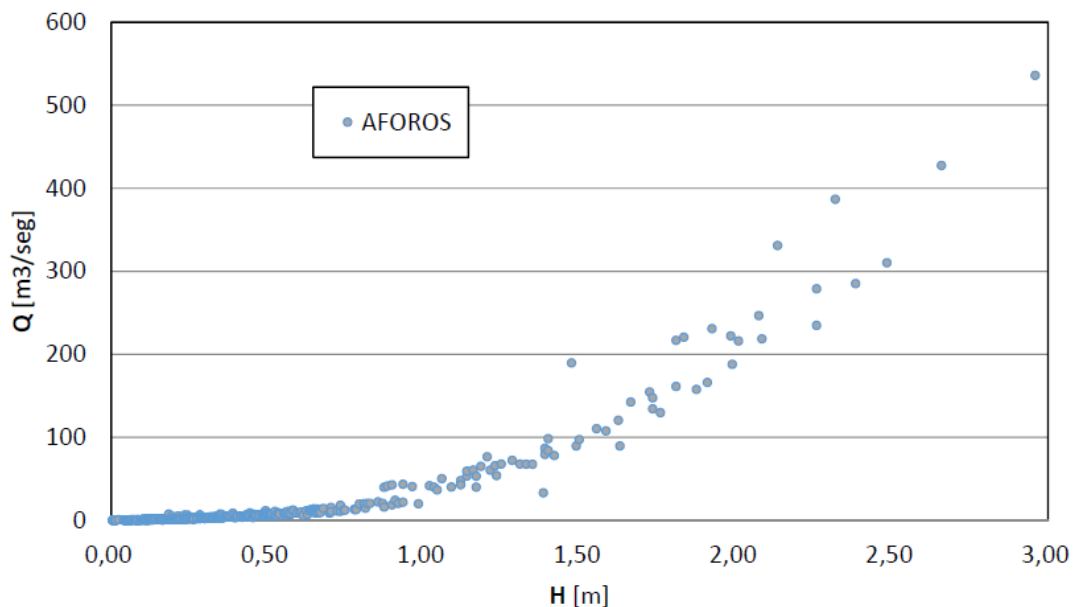


Figura 5.9. Datos de aforos realizados.

La alternativa adoptada, consiste en una curva de dos tramos con punto de quiebre en $H=0,85$ m. Las ecuaciones de los dos tramos son las siguientes:

$$\text{Para } H < 0.85 \quad Q = 19.86 * H^2 + 7.8944 * H + 0.053$$

Para $H > 0.85$ $Q = 77.9 * H^2 - 54.3206 * H + 11.0019$

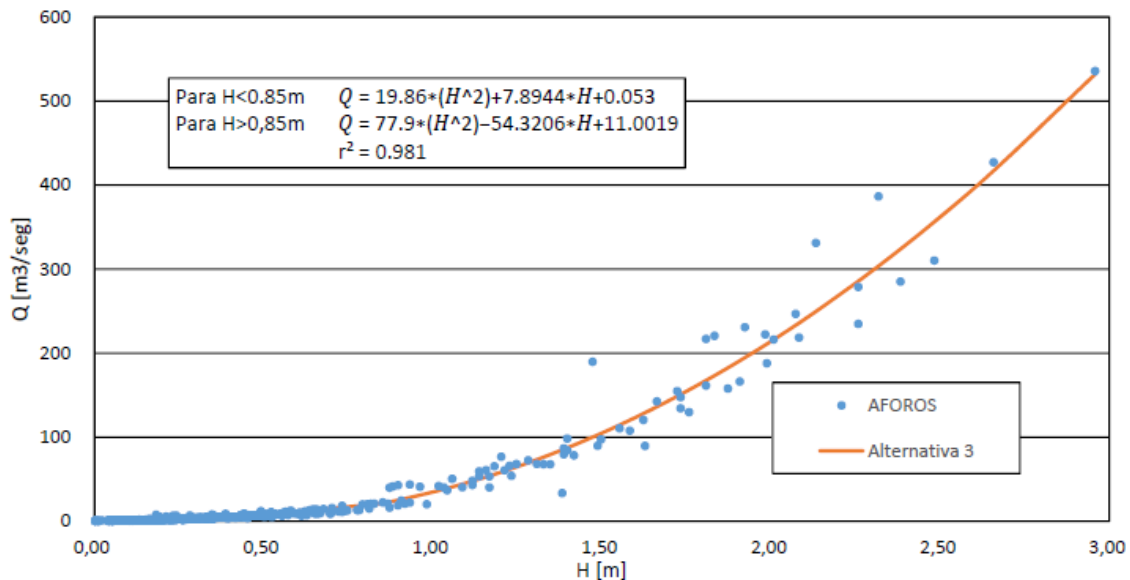


Figura 5.10. Curva H-Q confeccionada por M. J. Martín.

5.4.2. DESCRIPCIÓN DE LOS AFOROS

En el momento en que se llevaron a cabo los aforos, la lectura se hacía en escala hidrométrica con su cero ubicado por debajo del labio del vertedero de la sección de control una magnitud igual a (-0,1167 metros), mientras que los datos de nivel medidos por el sensor 604 tienen como cero el labio del vertedero. Por lo tanto, antes de estimar la curva, se le restó 0,1167 metros a los niveles leídos durante los aforos, para que de esta manera estuviesen referenciados al mismo cero que los datos base.

5.4.3. DESCRIPCIÓN DEL SENSOR 604

La serie de niveles medidas por el CIRSA (sensor 604) inicia en febr. 1993 a la actualidad. Sin embargo dentro de este periodo hay muchas lagunas, no habiendo datos desde abril de 1998 a marzo de 2002. De este tramo solo se recuperó noviembre y diciembre de 1999.

El limnógrafo de la estación 604, siempre han tenido el cero del nivel coincidente con el labio del vertedero, hasta que el 25 de enero de 2007 el cero se mueve a 30 cm por debajo del

mismo. Esto significa que si el nivel del limnógrafo marca 0,30 m, recién el agua está llegando al labio del vertedero. Por lo tanto para calcular los caudales con la curva de transformación se deben utilizar los niveles tal cual como se encuentran registrados en los archivos originales txt (Evout) del CIRSA hasta el 25 de enero del 2007, fecha a partir de la cual se le deberán restar 30 cm a los niveles originales. A modo de ejemplo, si el 10 de julio del 2001 el nivel registrado es 0,58 m, se ingresa con este valor a la curva H-Q, y se obtiene el caudal. Pero si en cambio se midió un nivel de 0,89 m el 07 de marzo del 2011, se entra a la curva con una altura igual a 0,59 m (0,85-0,30).

Además es importante señalar que cuando se reinstala el sensor en febrero de 2002, se advierte q el muro estaba roto en la parte más baja. La rotura abarca un área de 2.4 m².

La medición con el sensor 604 inicialmente tenía un rango de 0-4m, a partir del 12 de enero de 1994 se cambió por uno de rango 0-10m. El sensor de 4m es capaz de interpretar variaciones de 0,0125m (1 step), mientras que en los sensores de 10m, la apreciación mínima es del orden de los 0,04m (1step).

5.5. NIVELES Y CAUDALES

5.5.1. INFORMACIÓN DISPONIBLES DE NIVELES

A continuación se describe la información con la cual se contó al momento de realizar una serie final confiable de niveles, para posteriormente convertirla en caudales con la utilización de la curva H-Q realizada y verificada por María José Martín en el desarrollo de su Práctica Supervisada.

- a- Registro original venido directamente del sistema, como salidas de la red telemétrica. La información provista consiste en una serie de niveles de la estación 600 "El Canal" la cual presenta un paso de tiempo variable. Laura Colladon proporcionó registros en

formatos de texto que datan desde el año 1987 al 2011 y Carlos Catalini desde el 2006 al 2014.

- b- Información semi-procesada de niveles con paso de tiempo horario realizada por Cesar García. La serie de datos abarca el periodo de años que va desde el 2011 al 2015 presentados en un documento de texto.
- c- Información procesada y organizada en planillas Excel por el Ing. Alonso. La serie de niveles y caudales aquí obtenida presenta un paso de tiempo horario que comienza en el año 1994 y finaliza en el 2009.

5.5.2. CONVERSIÓN DE LA INFORMACIÓN A SERIES DE NIVELES HORARIAS.

A los fines de llevar la información original a paso de tiempo horario se llevó adelante el siguiente procedimiento:

- 1- Los documentos de texto recibidos por L. Colladon y C. Catalini debieron ser transformados, como primer paso a planillas Excel, las cuales nos permiten posteriormente llevar a cabo el procesamiento necesario. Se realiza de la misma manera que para el caso de precipitación, obteniéndose aquí tres columnas, fecha, hora y niveles.
- 2- Se debe modificar la fecha ya que la misma se encuentra en un formato no reconocido por Excel. Y a la columna de las horas se le retiran los minutos y segundos, ya que queremos conseguir una serie con paso de tiempo horario.
- 3- En una tercera instancia se ordeno de mayor a menor los niveles de cada hora de manera de elegir posteriormente el máximo valor de nivel de cada hora.
- 4- Por último, se armó una serie completa, es decir, con todos los días y horas entre los años en los que recibimos la información, ya que en los registros con los que nos encontramos había horas y días con datos faltantes. Para esto hicimos uso de la función BuscarV, la cual nos permitió colocar el nivel en la nueva serie en su posición

correcta, es decir, en la fecha y hora correspondiente. Este paso fue realizado de la misma manera que en el caso de la conformación de las series de precipitación.

En la siguiente imagen se ejemplifica el proceso llevado a cabo en la obtención de las series de niveles con paso de tiempo horario.

FECHA	HORA	FECHA + HORA	NIVEL		NIVEL		FECHA	HORA	FECHA + HORA	NIVEL
01/01/2006	17	3871817	0.31	Ordenar de mayor a menor	0.31	Serie final (Buscar V)	01/01/2006	18	3871818	0.27
01/01/2006	17	3871817	0.31		0.31		01/01/2006	19	3871819	0.31
01/01/2006	17	3871817	0.27		0.27		01/01/2006	20	3871820	0.31
01/01/2006	18	3871818	0.27		0.27		01/01/2006	21	3871821	0.31
01/01/2006	18	3871818	0.27		0.27		01/01/2006	22	3871822	0.31
01/01/2006	18	3871818	0.24		0.24		01/01/2006	23	3871823	0.27
01/01/2006	19	3871819	0.27		0.31		02/01/2006	0	3871900	0.24
01/01/2006	19	3871819	0.27		0.27		02/01/2006	1	3871901	0.2
01/01/2006	19	3871819	0.31		0.27		02/01/2006	2	3871902	0.2
01/01/2006	20	3871820	0.31		0.31		02/01/2006	3	3871903	0.2
01/01/2006	20	3871820	0.31	0.31	02/01/2006	4	3871904	0.16		

5.5.3. COMPARACIÓN DE LAS SERIES DE NIVELES

Con el fin de obtener una serie de niveles confiable se comparan las series obtenidas a partir de la información base con las provistas por García y Alonso durante los periodos de superposición. En la siguiente tabla se esquematizan los periodos abarcados por cada serie pudiéndose apreciar claramente los tramos de superposición, (Tabla 5.9).

Tabla 5.9. Periodos de superposición de las series de niveles.

AÑO	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015					
L. Colladon	■																										
C. Catalini													■														
C. García																		■									
F. Alonso	■																										

- Serie L. Colladon vs C. catalini: Al analizar ambas series se comprobó que las mismas presentan un alto porcentaje de similitud, exceptuando el valor de una de las horas y tramos en los que una de las series no presenta información y la otra si, o viceversa. Cabe destacar que la columna de Carlos Catalini posee mayor cantidad de horas con datos que la de Colladon.
- Verificación de los niveles de C. Catalini con los de C. García: En una primera instancia, al compararlas se pudo observar que había una gran desigualdad entre ellas. Se identificó que gran parte de la falta de coincidencia podía ser subsanada al sumarle a la serie de Cesar 30 cm desde su comienzo, el 18/07/2011 a las 16 Hs hasta el 31/12/2012 a las 23 Hs y adelantándolo una hora desde el 6/11/2011 hasta finalizar la serie de niveles de C. Catalini el 31/12/2014. De esta manera se logró reducir la cantidad de errores de 4030 diferencias a 489.
- Comparación de las series de Carlos Catalini y Facundo Alonso: durante los tramos de superposición se comprueba que en todas horas en las que ambas series contienen datos los valores son exactamente iguales. La única diferencia encontrada al analizar el periodo yuxtaposición se debe a la existencia de datos en una de las series mientras que en la otra hay ausencia de ellos. La columna obtenida a partir de los niveles proporcionados por Catalini posee menor cantidad de celdas vacías.

Por los motivos expuestos en los párrafos anteriores es que se descartaron las series de Laura Colladon y Facundo Alonso a la hora de conformar la serie final. Ambas columnas contienen iguales valores a la de C. Catalini y mayor cantidad de horas vacías. Es decir, que la obtención de la serie de niveles en el tramo de superposición se obtuvo a partir de la fusión de la columna de Catalini con la de Garcia. El proceso consistió básicamente en completar los huecos de la primer serie con los valores presentes en la otra y, en aquellas horas que presentaban desigualdades entre una y otra serie generalmente prevaleció la información contenida en la columna de C. Catalini.

5.5.4. CONFORMACION DE LA SERIE DE CAUDALES HORARIA

Una vez obtenida la serie de niveles se la debe corregir como se explicó en el capítulo anterior restándole 0,30 m a partir del 25 de enero del 2007 en adelante. Finalmente utilizando las ecuaciones de la curva H-Q se transforma la serie de niveles en una serie de caudales.

A diferencia de la precipitación se decidió no rellenar la serie de caudales en las horas faltantes por las siguientes razones:

- El método utilizado antiguamente, interpolación lineal entre los valores de los datos extremos a cada periodo sin dato, alberga cierto grado de inexactitud ya que entre medio de las mismas podría haber ocurrido un pico que no se tendría en cuenta.
- No es esencial contar con una serie de caudales completamente llena ya que simplemente el modelo no se calibrará en los espacios vacíos.

5.6. ANALISIS P-Q

5.6.1. CONFORMACIÓN DE SERIES P-Q

Con intención de detectar anomalías en las series de precipitación y caudal se confeccionaron gráficos donde se representan en la parte inferior los hidrógrafas y en la parte superior la precipitación. De esta manera se puede visualizar de manera sencilla inconsistencias de la serie P-Q, como ser periodos en los que la lluvia no genera caudal o viceversa.

La serie P-Q que es utilizada posteriormente en la modelación comienza el 06 de marzo de 1995 y finaliza el 31 de diciembre de 2014. Se debió eliminar el primer tramo de la serie que comienza en 1994 debido a que del 02/03 al 06/03 no hay registros de lluvia en ninguna de las estaciones de la cuenta y tampoco información de precipitación diaria para poder rellenar las horas vacías.

A continuación se presentan algunos de los gráficos realizados en Excel.

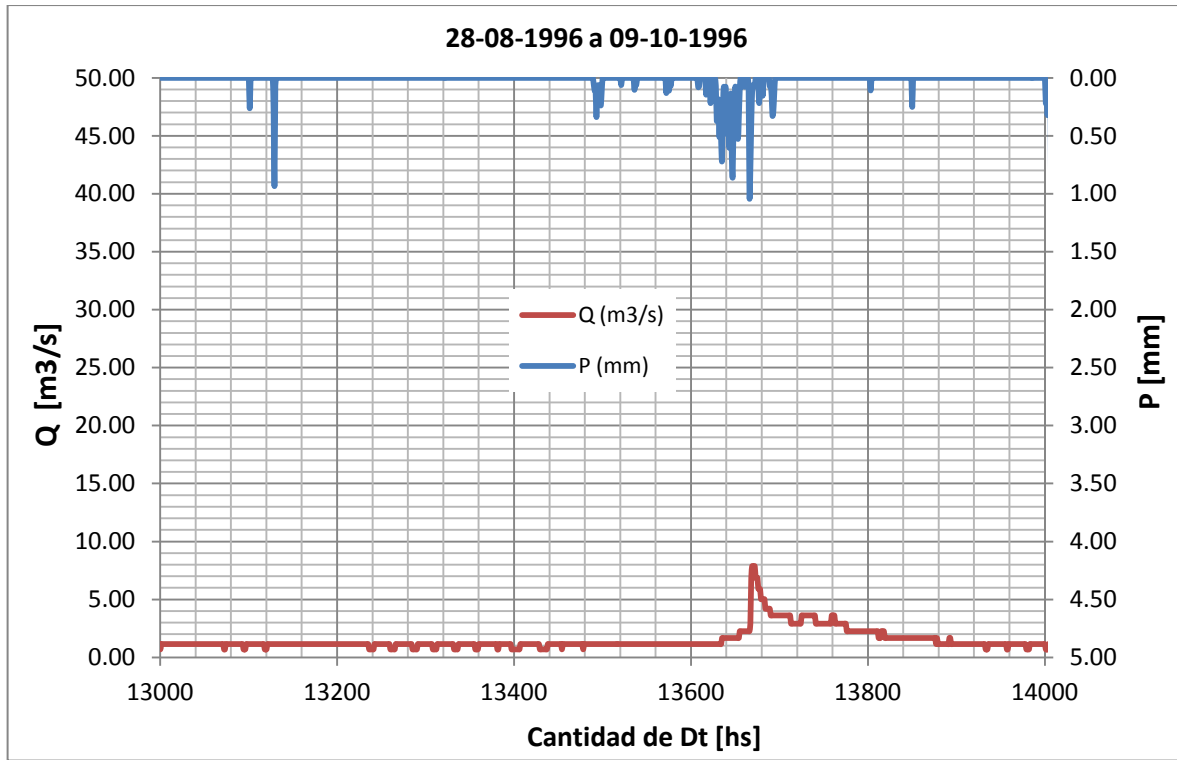


Gráfico 5.1.

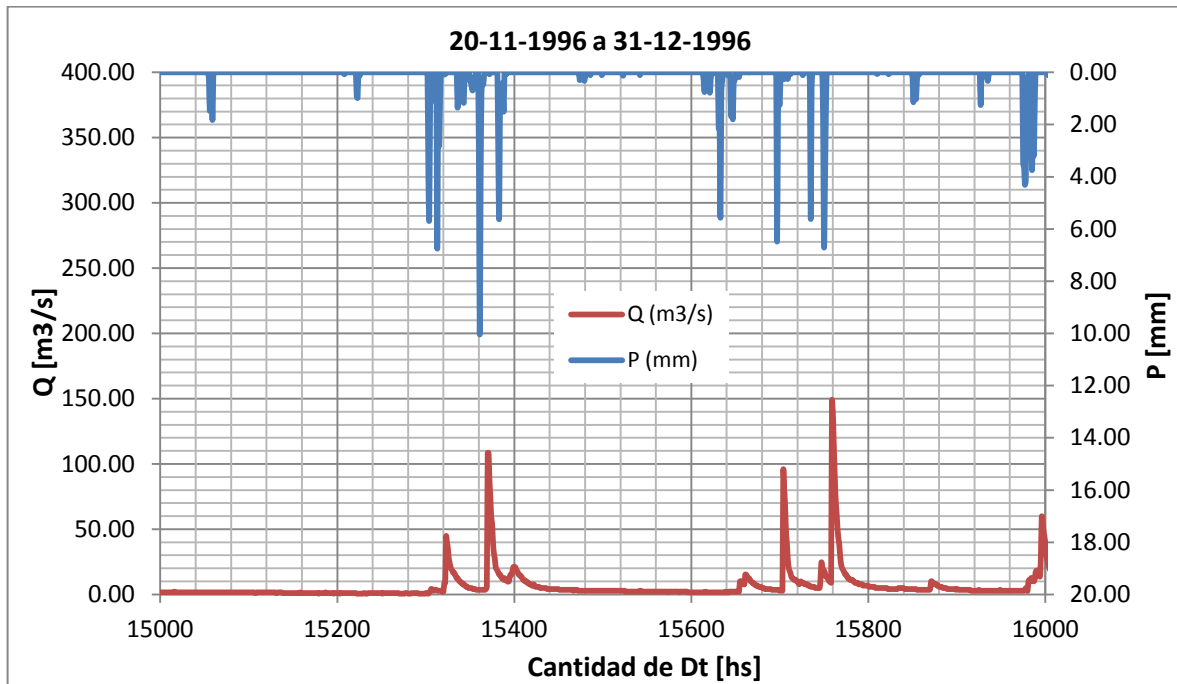


Gráfico 5.2

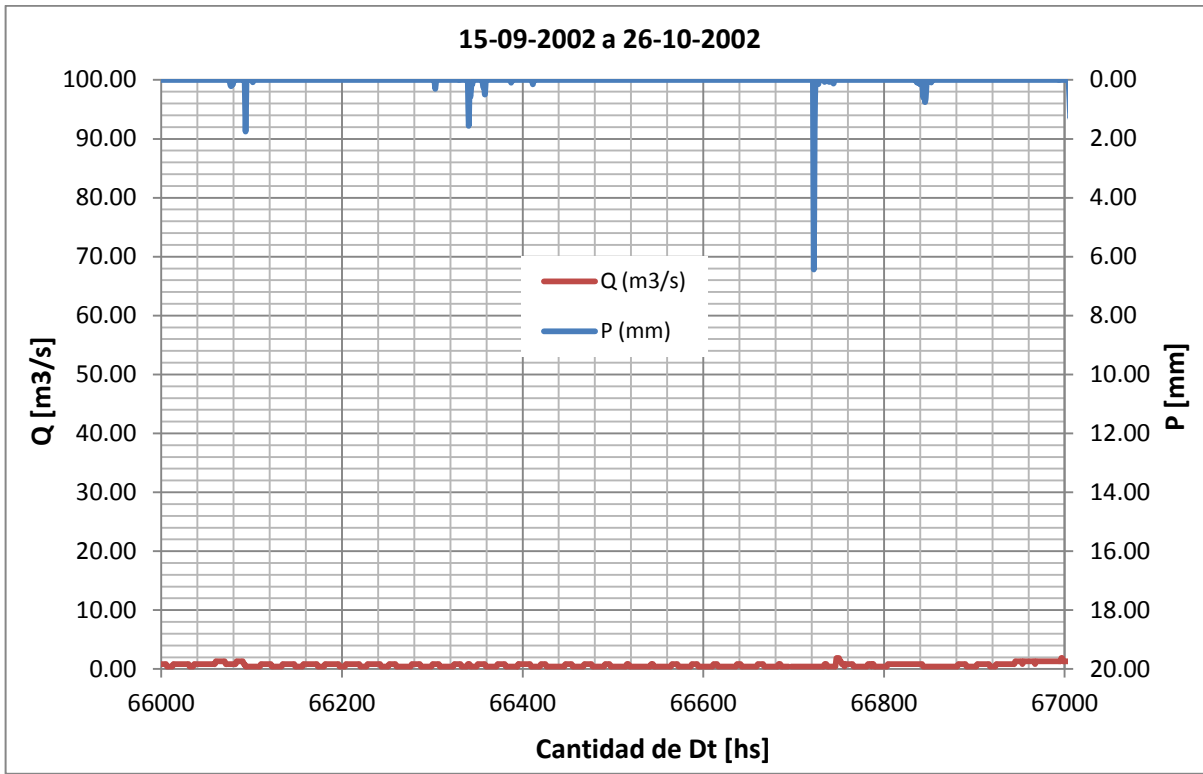


Gráfico 5.3

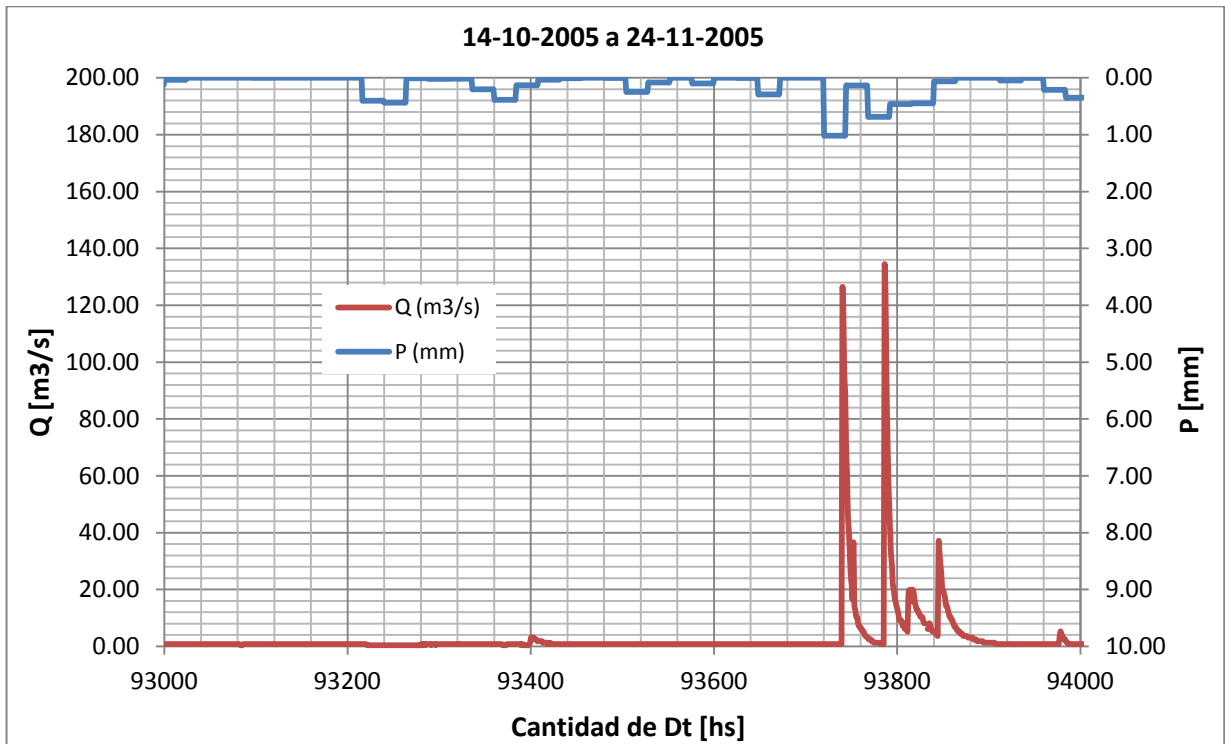


Gráfico 5.4

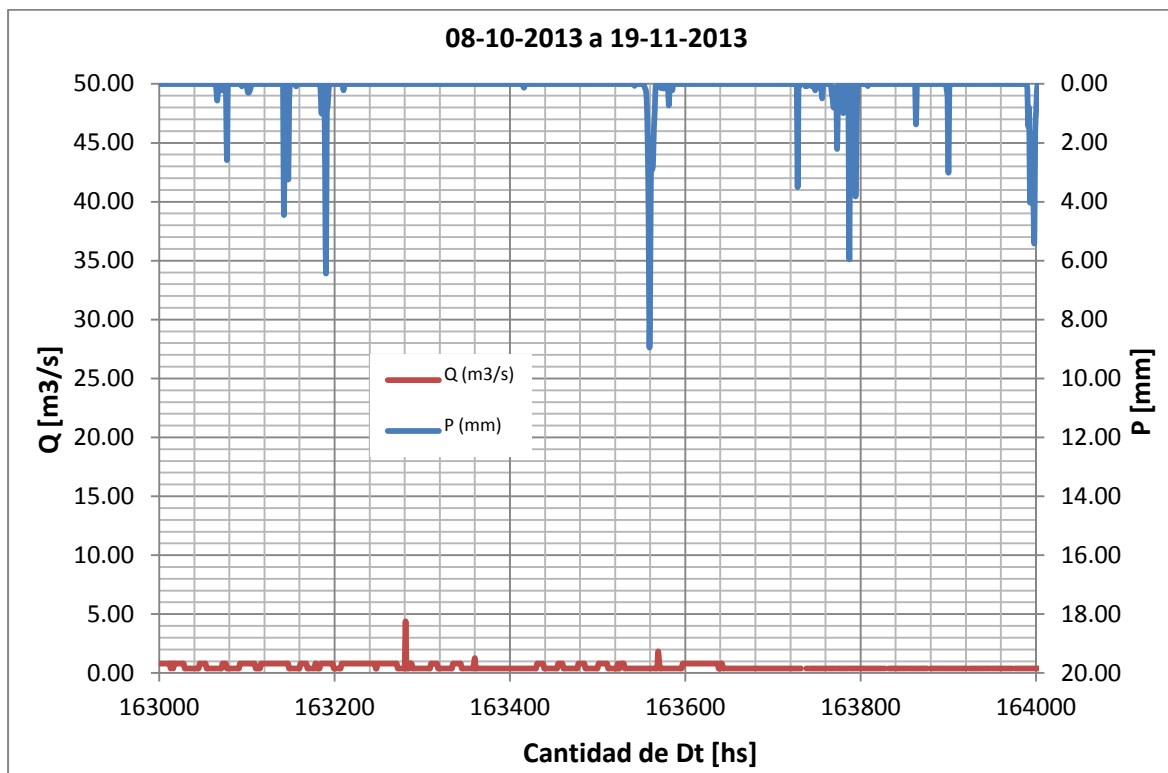


Gráfico 5.5

Se realizará una breve descripción de las gráficas destacando los sucesos de relevancia.

- Gráfico 5.1 y Gráfico 5.2: En ellos se aprecia una gran correspondencia entre la lluvia ocurrida y el hidrograma generado.
- Gráfico 5.3: Una lluvia puntual no origina un aumento de caudal debido a que la misma se infiltra.
- Gráfico 5.4: Irregularidad de la serie P-Q. Se genera un caudal considerable en comparación con la escasa precipitación.
- Gráfico 5.5: Al igual que en el caso anterior, la serie presenta cierto grado de inconsistencias, siendo en este caso bajo el caudal en comparación con la lluvia. Obsérvese el Gráfico 5.2, la precipitación es similar al caso en análisis mientras que los caudales generados son considerablemente mayores.

5.7. EVAPOTRANSPIRACIÓN POTENCIAL

Se define a la evapotranspiración potencial como la máxima cantidad de agua que puede evaporarse desde un suelo completamente cubierto de vegetación, que se desarrolla en óptimas condiciones, y en el supuesto caso de no existir limitaciones en la disponibilidad de agua. Según esta definición, la magnitud de la ETP está regulada únicamente por las condiciones climáticas del lugar de medida.

Existe una amplia variedad de autores que han desarrollado formulaciones matemáticas para su determinación, entre ellos se pueden mencionar Charles Thornthwite, Perman, Perman-Monteith, Jensen y Haisen, McGuinness y Bordne, Oudin, etc.

El método de Perman (1948) utiliza como datos de entrada las horas de insolación, la temperatura, humedad relativa y la velocidad del aire. El modelo de Perman-Monteith además de las anteriores variables introduce la radiación solar y emplea valores mínimos y máximos de temperatura. Mientras que Oudin et al. (2005) propuso un método con el objetivo de desarrollar una ecuación simple dentro de un modelo lluvia-escurrimiento, con el cual se puede calcular la ETP diaria a partir de la temperatura media del día y la radiación extraterrestre únicamente.

En general, con los métodos Penman y Penman-Monteith se obtienen valores de mayor proximidad a los observados, con la desventaja de requerir un número considerable de variables de entrada. Por otro lado el método de Oudin, presenta un desempeño similar, considerando que incluyen pocas variables.

Por los motivos expuestos y debido a que los estudios hidrológicos por lo general carecen de suficiente información climática, en el presente trabajo se utiliza la ecuación de Oudin para calcular la evapotranspiración potencial.

Además Oudin et al. (2005.a) concluyen que los modelos de EP calculada empleando series diarias de temperatura medida, presentan igual performance que cuando se calcula

empleando la temperatura diaria promedio para cada día del año. Ese concepto indica que de no disponer series largas medidas de temperatura, puede calcularse la temperatura promedio día a día, con el registro de la temperatura de ese día para una serie de años.

5.7.1. METODO DE LUDOVIC OUDIN

$$\text{Si } T_a + 5 > 0 \qquad PE = \frac{Re}{\lambda * \rho} * \frac{T_a + 5}{100}$$

$$\text{Caso contrario} \qquad PE = 0$$

Donde:

- PE es la evapotranspiración potencial [mm/día].
- Re es la radiación extraterrestre [MJ/m²día].
- λ es el calor latente del agua [MJ/kg].
- ρ es la densidad del agua [kg/m³].
- Ta es la temperatura diaria [°C].

Radiación extraterrestre (Re): para determinar un valor es necesario contar con latitudes de la cuenca y su hemisferio. Tabla 5.10. Radiación extraterrestre.

Tabla 5.10. Radiación extraterrestre.

Northern Hemisphere												Lat	Southern Hemisphere											
Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec	deg	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
0.0	2.6	10.4	23.0	36.2	42.5	36.4	26.0	14.9	4.9	0.1	0.0	70	41.4	26.6	15.8	4.9	0.2	0.0	0.0	3.2	10.7	23.5	37.3	45.3
0.1	3.7	11.7	23.9	26.3	42.0	38.9	26.6	16.1	6.0	0.7	0.0	68	41.0	26.3	16.9	6.0	0.8	0.0	0.0	3.2	11.9	24.4	37.4	44.7
0.6	4.8	12.9	24.8	36.6	41.4	38.8	29.3	17.3	7.2	1.5	0.1	66	40.9	30.0	18.1	7.2	1.5	0.1	0.5	4.2	13.1	25.4	37.6	44.1
1.4	5.9	14.1	25.8	36.9	41.2	38.8	30.0	18.4	8.5	2.4	0.6	64	41.0	30.8	19.3	8.4	2.4	0.6	1.2	5.3	14.4	26.3	38.0	43.8
2.3	7.1	15.4	26.6	36.3	41.2	39.0	30.6	19.5	9.7	3.4	1.3	62	41.2	31.5	20.4	9.6	3.4	1.2	2.0	6.4	15.5	27.2	38.3	43.9
3.3	8.3	16.6	27.5	36.6	41.2	39.2	31.3	20.6	10.9	4.4	2.2	60	41.5	32.3	21.5	10.8	4.4	2.0	2.9	7.6	16.7	28.1	38.7	43.9
4.3	9.6	17.7	28.4	37.0	41.3	39.4	32.0	21.7	12.1	5.5	3.1	58	41.7	33.0	22.6	12.0	5.5	2.9	3.9	8.7	17.9	28.9	39.1	44.0
5.4	10.8	18.9	29.2	37.4	41.4	39.6	32.6	22.7	13.3	6.7	4.2	56	42.0	33.7	23.6	13.2	6.6	3.9	4.9	9.9	19.0	29.8	39.5	44.1
6.5	12.0	20.0	30.0	37.8	41.5	39.8	33.2	23.7	14.5	7.8	5.2	54	42.2	34.3	24.6	14.4	7.7	4.9	6.0	11.1	20.1	30.6	39.9	44.3
7.7	13.2	21.1	30.8	38.2	41.6	40.1	33.8	24.7	15.7	9.0	6.4	52	42.5	35.0	25.6	15.6	8.0	6.0	7.1	12.2	21.2	31.4	40.2	44.4
8.9	14.4	22.2	31.5	38.5	41.7	40.2	34.4	25.7	16.9	10.2	7.5	50	42.7	35.6	26.6	16.7	10.0	7.1	8.2	13.4	22.2	32.1	40.6	44.5
10.1	15.7	23.3	32.2	38.8	41.8	40.4	34.9	26.8	18.1	11.4	8.7	48	42.9	36.3	27.5	17.9	11.1	8.2	9.3	14.6	23.3	32.8	40.9	44.5
11.3	16.9	24.3	32.9	39.1	41.9	40.6	35.4	27.5	19.2	12.6	9.9	46	43.0	36.7	28.4	19.0	12.3	9.3	10.4	15.7	24.3	33.5	41.1	44.6
12.5	18.0	25.3	33.5	39.3	41.9	40.7	35.9	28.4	20.3	13.9	11.1	44	43.2	37.2	29.3	20.1	13.5	10.5	11.6	16.8	25.2	34.1	41.4	44.6
13.8	19.2	26.3	34.1	39.5	41.9	40.8	36.3	29.2	21.4	15.1	12.4	42	43.3	37.7	30.1	21.2	14.6	11.8	12.8	18.0	26.2	34.7	41.6	44.6
15.0	20.4	27.2	34.7	39.7	41.9	40.8	36.7	30.0	22.5	16.3	13.6	40	43.4	38.1	30.9	22.3	15.6	12.8	13.9	19.1	27.1	35.3	41.8	44.6
16.2	21.5	28.1	35.2	39.9	41.8	40.8	37.0	30.7	23.6	17.5	14.8	38	43.4	38.6	31.7	23.3	16.9	13.9	15.1	20.2	28.0	35.8	41.9	44.5
17.5	22.6	29.0	35.7	40.0	41.7	40.8	37.4	31.5	24.6	18.7	16.1	36	43.4	38.9	32.4	24.3	18.1	15.1	16.2	21.2	28.8	36.3	42.0	44.4
18.7	23.7	29.9	36.1	40.0	41.6	40.8	37.8	32.1	25.6	19.9	17.3	34	43.4	39.2	33.0	25.3	19.2	16.2	17.4	22.3	29.6	36.7	42.0	44.3
19.9	24.8	30.7	36.5	40.0	41.4	40.7	37.9	32.8	26.6	21.1	18.5	32	43.3	39.4	33.7	26.3	20.3	17.4	18.5	23.3	30.4	37.1	42.0	44.1
21.1	25.8	31.4	36.8	40.0	41.2	40.6	38.0	33.4	27.6	22.2	19.8	30	43.1	39.6	34.3	27.2	21.4	18.5	19.6	24.3	31.1	37.5	42.0	43.9
22.3	26.8	32.2	37.1	40.0	40.9	40.4	38.2	33.9	28.5	23.3	21.0	28	43.0	39.8	34.8	28.1	22.5	19.7	20.7	25.3	31.8	37.8	41.9	43.5
23.4	27.8	32.8	37.4	39.9	40.6	40.2	38.3	34.5	29.3	24.5	22.2	26	42.8	39.9	35.3	29.0	23.5	20.8	21.8	26.3	32.5	38.0	41.8	43.3
24.6	28.8	33.5	37.6	39.7	40.5	39.9	38.3	34.9	30.2	25.5	23.3	24	42.5	40.0	35.8	29.8	24.6	21.9	22.9	27.2	33.1	38.3	41.7	43.0
25.7	29.7	34.1	37.8	39.5	40.0	39.6	38.4	35.4	31.0	26.6	24.5	22	42.2	40.1	36.2	30.5	25.6	23.0	24.0	28.1	33.7	38.4	41.4	42.6
26.8	30.6	34.7	37.9	39.3	39.5	39.3	38.3	35.8	31.8	27.7	25.6	20	41.9	40.0	36.6	31.3	26.6	24.1	25.0	28.9	34.2	38.6	41.2	42.1
27.9	31.5	35.2	38.0	39.0	39.1	38.9	38.2	36.1	32.5	28.7	26.8	18	41.5	40.0	37.0	32.1	27.5	25.1	26.0	29.8	34.7	38.7	40.9	41.7
28.9	32.3	35.7	38.1	38.7	38.6	38.5	38.1	36.4	33.2	29.6	27.9	16	41.1	39.9	37.2	32.8	28.5	26.2	27.0	30.6	35.2	38.7	40.6	41.2
29.9	33.1	36.1	38.1	38.4	38.1	38.1	38.0	36.7	33.9	30.6	28.9	14	40.6	39.7	37.5	33.4	29.4	27.2	27.9	31.3	35.8	38.7	40.2	40.6
30.9	33.8	36.5	38.0	38.0	37.6	37.6	37.8	36.9	34.5	31.5	29.0	12	40.1	39.6	37.7	34.0	30.2	28.1	28.9	32.1	36.0	38.6	39.8	40.0
31.9	34.5	36.9	37.9	37.6	37.0	37.1	37.5	37.1	35.1	32.4	31.0	10	39.5	39.1	37.8	34.6	31.1	29.1	29.8	32.8	36.3	38.5	39.3	39.4
32.8	35.2	37.2	37.8	37.1	36.3	36.5	37.2	37.2	35.6	33.3	32.0	8	38.9	39.0	37.9	35.1	31.9	30.0	30.7	33.4	36.6	38.4	38.8	38.7
33.7	35.8	37.4	37.6	36.6	35.7	35.9	36.9	37.3	36.1	34.1	32.9	6	38.3	38.7	38.0	35.6	32.7	30.9	31.5	34.0	36.8	38.2	38.2	38.0
34.6	36.4	37.6	37.4	36.0	35.0	35.3	36.5	37.3	36.6	34.9	33.9	4	37.6	38.3	38.0	36.0	33.4	31.8	32.3	34.6	37.0	38.0	37.6	37.2
35.4	37.0	37.8	37.1	35.4	34.2	34.6	36.1	37.3	37.0	35.6	34.8	2	36.9	37.9	38.0	36.4	34.1	32.6	33.1	35.2	37.1	37.7	37.0	36.4
36.2	37.5	37.9	36.8	34.8	33.4	33.9	35.7	37.2	37.4	36.3	35.6	0	36.2	37.5	37.9	36.9	34.8	33.4	33.9	35.7	37.2	37.4	36.3	35.6

5.7.2. PROCESAMIENTO DE LA INFORMACIÓN

La información con la que se contó para la obtener la serie horaria de temperatura es proporcionada por Cesar García la cual se encuentra semi procesada en formato de texto. La misma contiene registros de 6 estaciones de la cuenca del río San Antonio que abarcan el periodo desde 1994 a 2009. La serie posee un paso de tiempo diario aunque con algunas faltantes.

Una vez convertido el archivo de texto en formato de Excel se procede a calcular la evapotranspiración potencial con la expresión de Ludovic Oudin. Como ya se explicó con anterioridad para utilizar dicha fórmula es necesario contar con temperaturas diarias y con la radiación extraterrestre. Esta última se obtiene a partir de la latitud de la cuenca, 31° en nuestro caso, y del hemisferio en el que se encuentra. Entrando a la tabla de la Tabla 5.10 se obtiene un valor de acuerdo al mes en el cual se calcule la ETP. Debido a que la expresión de

Ludovic Oudin fue formulada para temperaturas diarias y nosotros contamos con una serie de paso de tiempo horario se debió modificar la ecuación dividiéndola por 24, es decir, se considera a la temperatura horaria como si fuera diaria calculándose así una ETP de todo un día, luego al dividir por 24 obtenemos la que ocurre en una hora si durante a lo largo del mismo día la temperatura no hubiera variado.

Luego de determinarse la ETP de las seis estaciones se conforma la serie de evapotranspiración media de la cuenca con los polígonos ya realizados. Los mismos resultaron ser escasos, pero al no requerirse tanta exactitud con esta serie, se tomó la decisión de modificar la información contenida en algunas de las estaciones para que de esta manera se pueda hacer uso de las combinaciones existentes. Por ejemplo, un tramo de la serie de ETP contiene información de la estación 200, 500, 600 y 900, pero esta combinación es inexistente, pero sí poseemos los coeficientes de thiesen de la combinación 200, 500, 600, 700 y 900 por lo que para subsanar el inconveniente simplemente basta con colocarle a la estación 700 los datos de la 600 por ser estaciones próximas. Otra posibilidad es que la combinación disponible sea la de las estaciones 200, 600 y 900, en este caso se elimina la estación 500 a la hora de calcular la evapotranspiración media. Sin dudas el resultado no es exacto, pero no se está introduciendo un error significativo, recordar que para calcular la evapotranspiración se utiliza una formula simplificada.

Lo descrito anteriormente fue realizado para el periodo comprendido entre los años 1994 y 2009. Con estos 16 años de información se realizó el promedio de las ETP horarias de los días homólogos formándose así una serie histórica de evapotranspiración potencial. Es decir, se obtiene la evapotranspiración para cada hora de cada día del año. Por ejemplo, la ETP del 2 de enero a las 3 de la tarde se obtiene realizando el promedio de la ETP de todos los 2 de enero a las 15 Hs de todos los años con información.

Luego, con la serie histórica puede rellenarse los tramos sin información durante el año 1994 y el 2009 y además utilizarse la misma para completar la serie hasta el 2014.

A modo de control de la serie obtenida se procedió comparar los datos de la serie histórica de evapotranspiración provista por el sitio web www.scielo.org.ar. Dicha página posee información de la estación del SNM ubicada en el aeropuerto de Córdoba la cual registra 1511 mm por año, lo que por día valdría alrededor de 4,14 mm y si lo dividimos por 24 obtenemos 0,17 mm por hora; o bien 182 mm durante el mes de enero, la que al dividirla por 31 obtenemos 5,8 mm por día, o sea 0,24 mm por hora. Es decir, analizando estos resultados podemos decir que en promedio la evapotranspiración en la ciudad de Córdoba ronda entre 0,15 y 0,30 mm por hora.

En cuanto a la serie del río San Antonio, observamos que la misma posee valores comprendidos entre 0,15 y 0,30. Además la suma de la evapotranspiración del mes de enero de la serie histórica es 131 mm, valor semejante al registrado en la estación del aeropuerto. Con esto podemos concluir que la serie obtenida del San Antonio es confiable.

6. RECOPIACIÓN Y REVISIÓN DE LOS DATOS DEL RÍO GUALEGUAY

Por motivos de que la modelación del río San Antonio no arrojó los resultados esperados, se decidió llevar a cabo la optimización de la cuenca del río Gualeguay y de esta manera comparar resultados y posteriormente sacar conclusiones al respecto.

Las series utilizadas para la programación corresponden a información ya procesada con anterioridad y organizada en planillas de Excel.

- Precipitación: comienza el 01/01/1998 y finaliza el 10/11/2010.
- Caudal: desde el 01/11/1992 al 31/08/2012.
- Evapotranspiración: del 01/08/1998 al 10/11/2010.

Es decir, el tramo de serie empleado comienza el 01/08/1998 y culmina el 10/11/2010. Las mismas fueron realizadas con un paso de tiempo diario ya que el tiempo de concentración de la cuenca es aproximadamente de 5 a 10 días.

La serie de precipitación fue rellenada con la intención de conseguir una serie ininterrumpida, mientras que la de caudales, por motivos expuestos en el capítulo anterior, en los días en los que no hay registros se los dejó vacíos para de esta manera no introducir datos erróneos sin necesidad. La serie de evapotranspiración potencial, al igual que la de precipitación es ininterrumpida.

6.1. ANALISIS P-Q

6.1.1. CONFORMACIÓN DE SERIES P-Q

Al igual que en el caso de la cuenca del río San Antonio se confeccionó una serie precipitación - caudal con la intención de detectar anomalías en las series. Se elaboraron gráficos donde se representan en la parte inferior los hidrógrafas y en la parte superior la precipitación. De esta manera se puede visualizar de manera sencilla inconsistencias entre las serie P y Q, como ser periodos en los que la lluvia no genera caudal o viceversa.

A continuación se presentan algunos de los gráficos realizados en Excel.

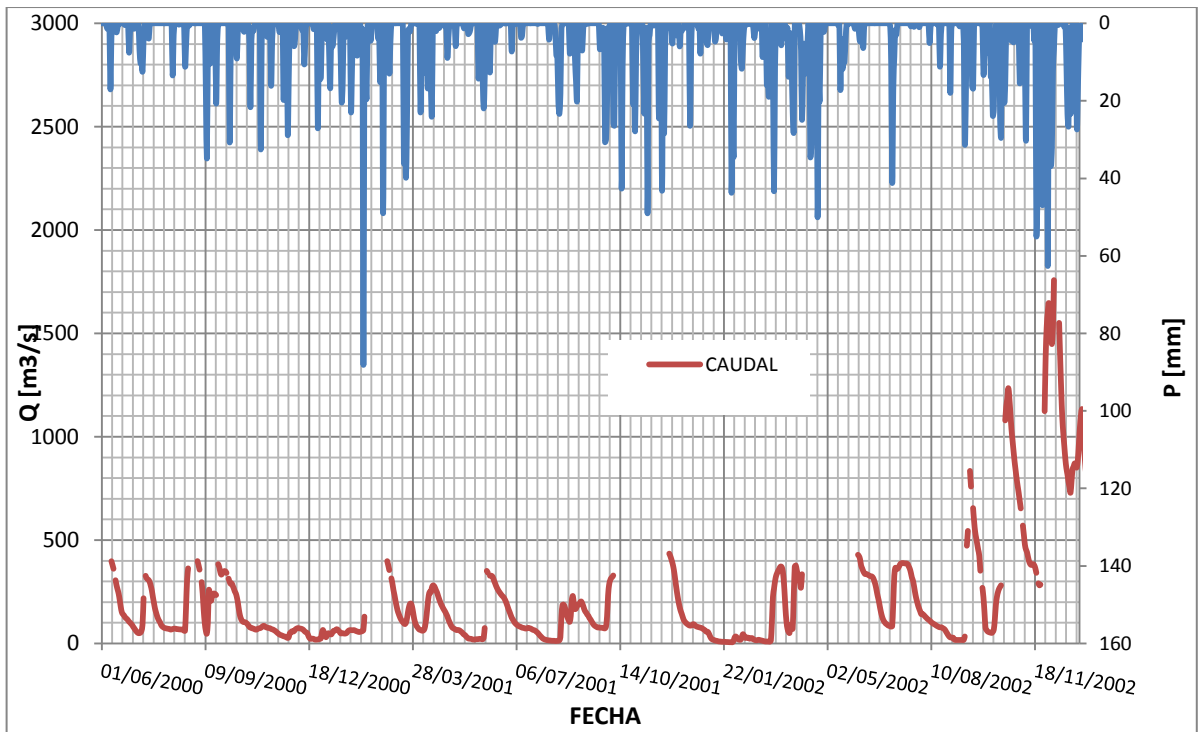


Gráfico 6.1.1.

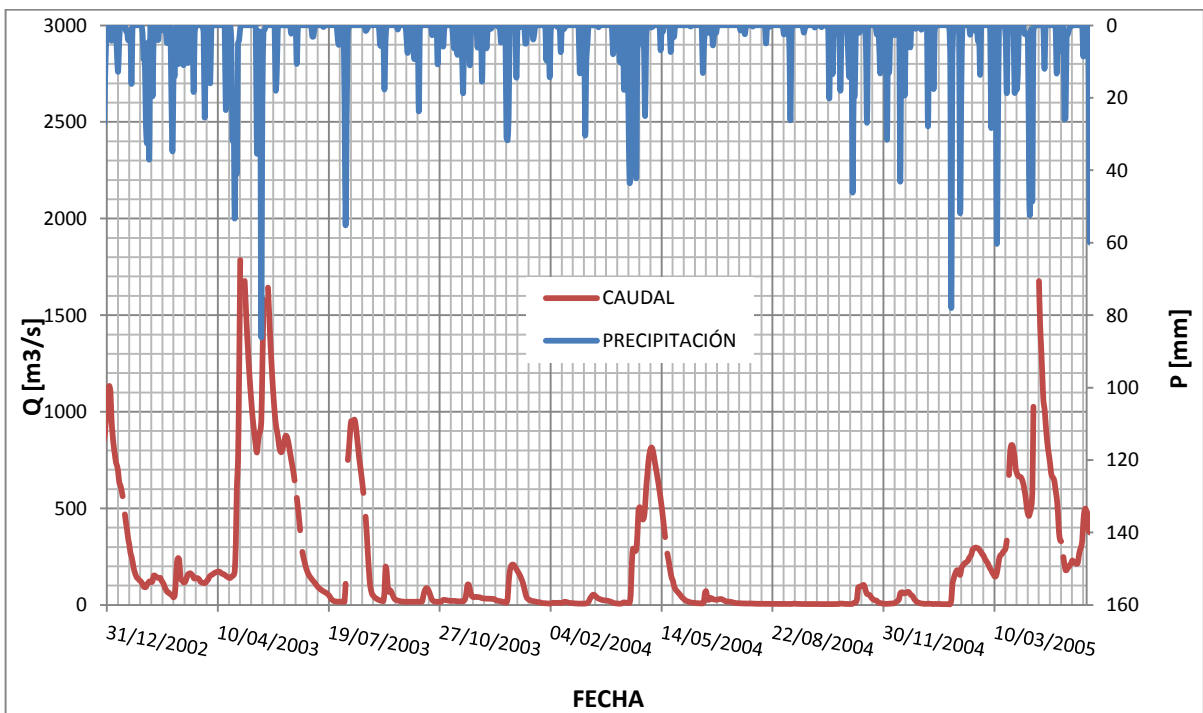


Gráfico 6.1.2.

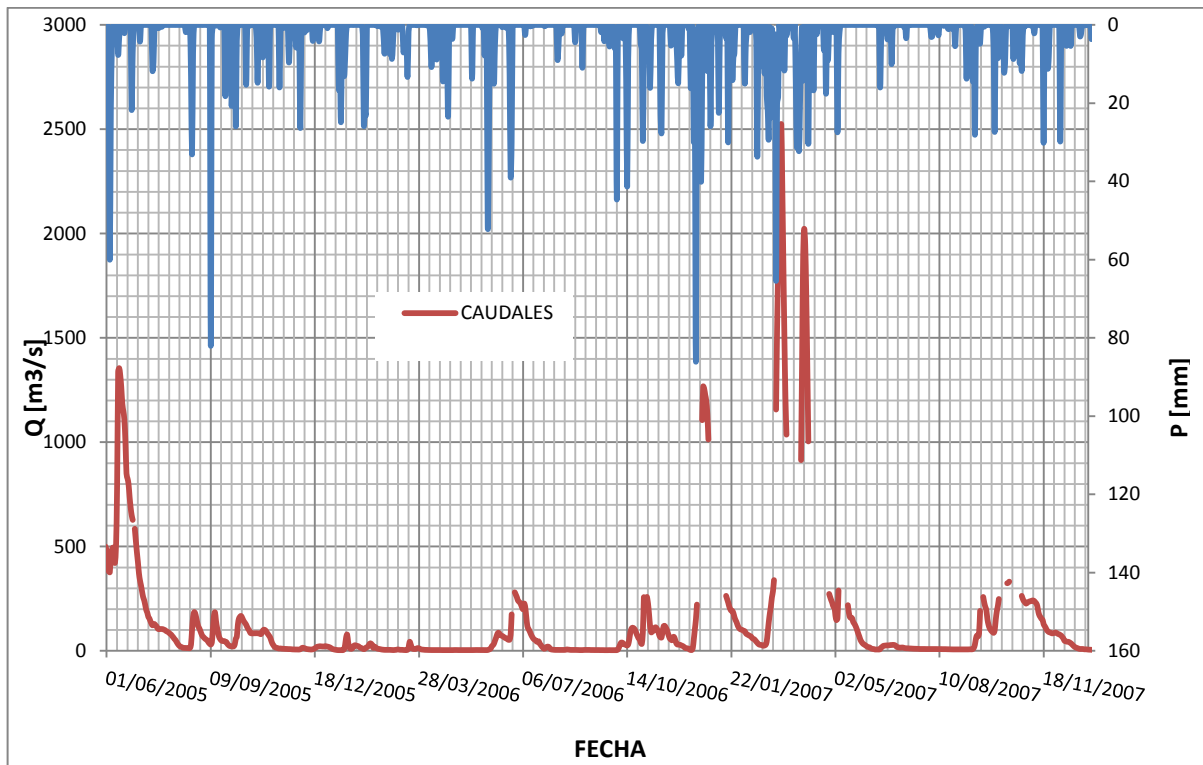


Gráfico 6.1.3.

En los gráficos anteriores se puede apreciar claramente las interrupciones en la serie de caudal.

A simple vista se puede concluir que hay una cierta coherencia y correlación entre ambas series. En el siguiente apartado, para mayor información en cuanto a la calidad de las series, se realizará una abstracción del análisis estadístico llevado a cabo en años anteriores para las series de precipitación y caudal.

7. ANALISIS ESTADÍSTICO

El análisis estadístico de las series de precipitación y caudal es de gran importancia, pues en una próxima instancia se hará uso de las mismas para llevar a cabo una modelación lluvia-caudal. Por ello, necesitamos de series confiables y que presenten ciertas condiciones para su aceptación.

Sin embargo en el presente trabajo, por motivos de tiempo y debido a que ya se han realizado con anterioridad estudios al respecto, se decidió aceptar las conclusiones adoptadas durante las Práctica Supervisada de María José Martín y Nereo Juárez en el año 2013 en cuanto a la confiabilidad de las series. A continuación se menciona la metodología utilizada con la cual se llegó a los resultados.

Las series analizadas corresponden a las de precipitación y caudal, descartándose la de evapotranspiración ya que como se señaló en apartados anteriores la misma no requiere de demasiada precisión y confiabilidad por no modificar significativamente los resultados durante la modelación.

Tanto la serie de precipitación como la de caudal fueron sometidas a dos pruebas estadísticas:

- Prueba de estacionalidad (Kendall): se busca que las series no presenten una tendencia importante.
- Prueba de homogeneidad a escala anual (Wilcoxon): permite corroborar si el promedio de un primer subconjunto de la muestra difiere significativamente del promedio del un segundo subconjunto de la misma muestra.

En las distintas pruebas se plantean dos hipótesis, una hipótesis nula y una hipótesis alternativa, esta ultima queda definida implícitamente como "H0 es falsa". Se adopta mediante niveles de significancia un rango, dentro del cual la hipótesis nula es rechazada, mientras que fuera de él, es aceptada.

Al aceptar una u otra hipótesis se puede estar incurriendo en un error como se esquematiza en la siguiente Tabla 7.1.

Tabla 7.1. Errores en la aceptación de las hipótesis.

	Ho es cierta	H1 es cierta
Se acepta H0	No hay error	Error tipo II
Se acepta H1	Error tipo I	No hay error

La probabilidad de cometer un error tipo I es conocida como nivel de significancia. En las pruebas realizadas se plantearon dos niveles de significancia, 5 % y 1%, en caso de no cumplirse para el rango de 5% se analiza con una significancia del 1%

7.1. RESULTADOS CUENCA RÍO SAN ANTONIO

7.1.1. PRECIPITACION

7.1.1.1. *Prueba de estacionalidad*

Se realizó la prueba a la serie con paso de tiempo diario.

- Hipótesis nula: se plantea que no existe tendencia entre las observaciones.

Con un nivel de significancia del 5% la hipótesis nula se debe rechazar. Sin embargo con un nivel de 1% H0 se acepta. Por lo cual se acepta que la serie no presenta ninguna tendencia.

7.1.1.2. *Prueba de homogeneidad*

Al igual que en el caso anterior se utilizó la serie con paso de tiempo diario.

- Hipótesis nula: señala que los promedios de las series son iguales.

Se obtuvieron los mismos resultados que en la prueba anterior en cuanto al rechazo de H0 para el 5% y la aceptación para el 1%. De esta manera se acepta que las dos sub-series son homogéneas y provenientes de la misma población.

7.1.2. CAUDAL

7.1.2.1. *Prueba de estacionalidad*

Al igual que en el caso de la precipitación se realizó la prueba estadística sobre una serie de caudales con paso de tiempo diario, es decir, una serie de caudales medios diarios.

- Hipótesis nula: se plantea que no existe tendencia entre las observaciones.

En este caso la H_0 se rechaza para ambos niveles de significancia, como M. J. Martín explica, esto puede deberse a que series que se desarrollan en un paso de tiempo corto y a esta escala, pueden estar afectadas por tendencias que se deban a la misma variabilidad climática. Es por ello que se aceptará la serie.

7.1.2.2. *Prueba de homogeneidad*

La serie utilizada es la misma que para la prueba anterior.

- Hipótesis nula: señala que los promedios de las series son iguales.

Como resultado se obtuvo que se debe rechazar H_0 para un nivel de significancia del 1%. Ante esto decidieron confeccionar una serie de caudales medios mensuales y se la sometió a prueba. Obteniéndose como resultado que las dos sub-series eran homogéneas para un nivel de significancia del 5% pudiendo concluir que las series provienen de una misma población.

7.2. RESULTADOS CUENCA GUALEGUAY

7.2.1. PRECIPITACION

7.2.1.1. *Prueba de estacionalidad*

A partir de las precipitaciones diarias se obtuvo una serie de precipitaciones mensuales acumuladas sobre las que se llevó a cabo la prueba estadística.

- Hipótesis nula: se plantea que no existe tendencia entre las observaciones.

Se acepta la hipótesis nula con un nivel de significancia del 5%, por lo cual se concluye que la serie no presenta ninguna tendencia.

7.2.1.2. Prueba de homogeneidad

Al igual que en el caso anterior se utilizó la serie de precipitación mensual acumulada.

- Hipótesis nula: señala que los promedios de las series son iguales.

Para un nivel de significancia del 5% se obtuvo que se debe aceptar la hipótesis nula, es decir que las dos sub-series son homogéneas y provenientes de la misma población.

7.2.2. CAUDAL

7.2.2.1. Prueba de estacionalidad

Al igual que en el caso de la precipitación se obtuvo una serie de caudales medios mensuales a partir de los caudales diarios.

- Hipótesis nula: se plantea que no existe tendencia en la serie.

En este caso la H0 se rechaza para ambos niveles de significancia, se justifica de la siguiente manera. Al no ser las series suficientemente largas pueden verse influenciadas por los ciclos multidecádicos comunes.

7.2.2.2. Prueba de homogeneidad

La serie utilizada es la misma que para la prueba anterior.

- Hipótesis nula: señala que los promedios de las series son iguales.

Como resultado se obtuvo que se debía rechazar H0 para un nivel de significancia del 1%. Al no ser las series suficientemente largas pueden verse influenciadas por los ciclos multidecádicos comunes.

8. APLICACIÓN Y RESULTADOS

En el presente capítulo se analiza los resultados obtenidos de la modelación del río San Antonio y el Gualeguay mediante del modelo hidrológico Génie Rural.

Se comienza realizando una descripción general del programa y de los archivos de entrada necesarios para hacerlo correr. Luego se explica la calibración de las cuencas de ambos ríos, y se muestran los resultados a los que hemos llegado utilizando el modelo en modo pronostico.

La serie de la cuenca del río San Antonio de 173784 datos fue dividida en 4 tramos de 47946 cada uno. Por su lado, la del río Gualeguay, con paso de tiempo diario de 4697 datos, se la decidió dividir en dos tramos.

La metodología adoptada para trabajar con cada cuenca consiste en realizar la calibración del modelo con todos los tramos en los que fueron divididas las series. Luego se seleccionaron algunos individuos para analizar sus hidrogramas pronosticados a distintos horizontes de tiempo, centrando nuestra atención a los de 6 y 12 horas.

8.1. DESCRIPCION GENERAL- ARCHIVOS PARA LA MODELACIÓN

El modelo fue programado por Alonso (2012) mediante el lenguaje de programación de alto nivel llamado fortran (Formula Translating System), el cual ha sido empleado con anterioridad en cuencas francesas. Como resultado de dicha programación se obtiene un ejecutable, el cual requiere para su funcionamiento una serie de archivos y carpetas.

En lo que sigue se realiza una ejemplificación de los archivos necesarios a partir de documentos tomados de la modelación de la cuenca del río San Antonio.

Nombre	Fecha de modifica...	Tipo	Tamaño
Control	28/03/2016 02:59 ...	Carpeta de archivos	
Datos	28/03/2016 02:52 ...	Carpeta de archivos	
Salidas	30/03/2016 11:34 a...	Carpeta de archivos	
Ent_Main.txt	28/03/2016 02:54 ...	Documento de tex...	5 KB
Libreria_Diaria.txt	14/05/2015 08:21 a...	Documento de tex...	12 KB
Libreria_Horaria.txt	14/05/2015 08:11 a...	Documento de tex...	12 KB
Main.exe	05/02/2016 05:57 ...	Aplicación	1 364 KB
Matriz.txt	04/10/2015 01:38 ...	Documento de tex...	51 KB
Poblac.txt	24/06/2015 10:46 a...	Documento de tex...	1 KB

Figura 8.1. Carpetas y archivos necesarios para la modelación.

1- Carpeta control: en la misma se guardan archivos de salida previos a la optimización, contienen los hidrogramas generados, tanto en simulación como en pronóstico, para cada juego de parámetros de la población inicial que se va a utilizar.

2- Carpeta datos: en ella se encuentran distintos archivos de textos los cuales deben poseer un formato y nombre determinado para poder ser llamados por el ejecutable. En su designación se utilizan 12 caracteres, los primeros 6 hacen referencia al nombre de la cuenca, los siguientes dos al número del tramo, a continuación se ubica un guion y luego de este, dos caracteres más para señalar a que país pertenece la cuenca ya que la forma en la que son leídos los archivos de Argentina (Ar) y los de Francia (Fr) es diferente, y el último carácter es para indicar si la serie posee un paso de tiempo horario (H) o diario (D).

Estos documentos contienen las variables de estado de la cuenca en análisis, precipitación, evapotranspiración y caudales observados y además el área de la misma. A continuación se observa el formato de los archivos mencionados para la cuenca Argentina del río San Antonio con paso de tiempo horario.

Area (km²)	Precipitación (mm)	EP (mm)	Caudal (m³/s)
509.			
0.000	0.037	-9.999	
0.000	0.029	-9.999	
0.000	0.026	-9.999	
0.000	0.024	-9.999	
0.000	0.024	-9.999	
0.000	0.024	-9.999	
0.000	0.024	-9.999	
0.000	0.025	-9.999	
0.000	0.024	-9.999	
0.000	0.025	-9.999	
0.000	0.025	-9.999	
0.000	0.025	-9.999	
0.000	0.025	-9.999	
0.000	0.026	-9.999	
0.000	0.026	-9.999	

Figura 8.2. Archivo de datos de la cuenca del río San Antonio.

En la fila uno se ubica el valor del área de la cuenca en kilómetros cuadrados seguido de un punto para que de esta manera el programa lo lea correctamente. La primera, segunda y tercer columna de la fila dos en adelante, corresponden a la precipitación en mm, la EP en mm y a los caudales observados en m³/s respectivamente. El valor -9.999 significa que en esa hora no hay datos de caudales.

3- Carpeta salidas: En ella se generan 4 tipos de archivos una vez finalizada la optimización.

- SanAnt01_ArH_030_Qp1.txt
- SanAnt01_ArH_030_Qp2.txt
- SanAnt01_ArH_Evol.txt
- SanAnt01_ArH_Sal.txt

Figura 8.3. Archivos de salida.

- Hidrogramas del individuo 30 con distintos horizontes de tiempo y el observado. Estos hidrogramas pueden ser simulados o pronosticados según

sean las opciones adoptadas en el archivo de entrada. Además se encuentra el valor de la función objetivo 1.

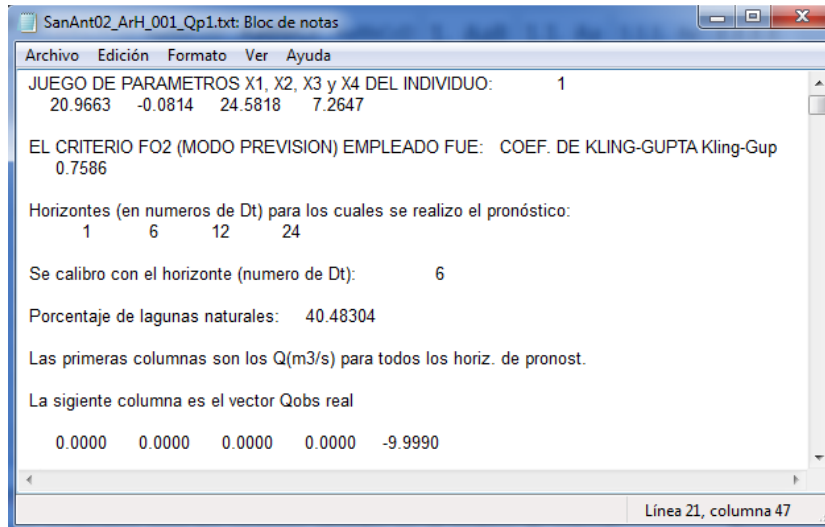


Figura 8.4. Parámetros e hidrogramas simulados del individuo 30.

- Hidrogramas del individuo 30 con distintos horizontes de tiempo y el observado. Los mismos, cualquiera sean las opciones elegidas en el archivo de entrada, serán siempre hidrogramas pronosticados, correspondientes a la función objetivo 2.

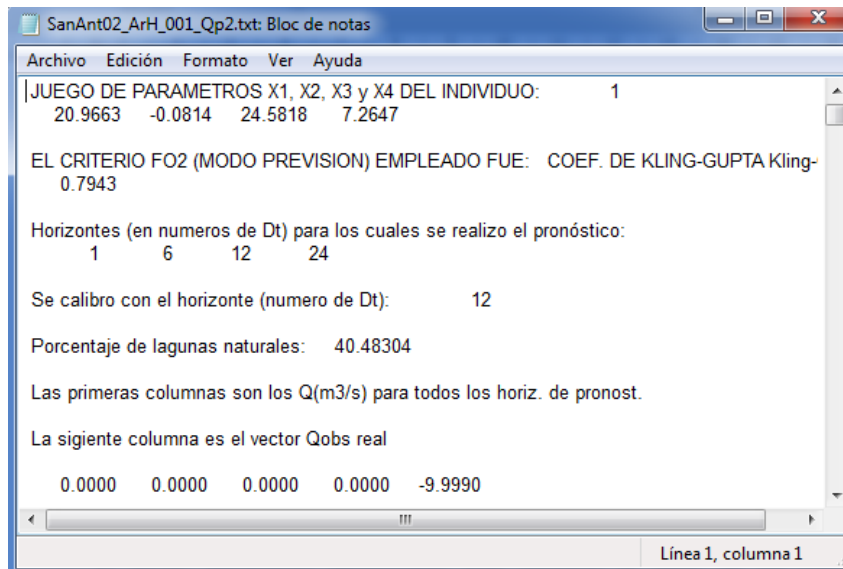


Figura 8.5. Parámetros e hidrogramas pronosticados del individuo 30.

- El tercer documento denominado Evol contiene el valor de los cuatro parámetros de cada uno de los individuos de la población y los valores de las funciones objetivo 1 y 2 para diferentes iteraciones. Es decir, cada un cierto número de interacciones, según sea lo establecido en el archivo de entrada, se registran las variables mencionadas. De esta manera se puede observar la manera en la que la población va variando hasta converger o hasta que el programa finalice.
- El archivo de salida contiene el valor de los parámetros de cada individuo con su correspondiente valor de función objetivo 1 y 2 de la última iteración.

4- Documento Ent Main: es un archivo de entrada que contiene las indicaciones necesarias para que el ejecutable lleve a cabo el proceso requerido. En la siguiente figura se observa su configuración.

Archivo	Edición	Formato	Ver	Ayuda						
	4		2		1	2	10	1		
	2		4		3	2	3			
	9000		4		4	2	478	1		
	1		6		12	24	48	72	73	74
	0.0001		0.0001		0.85	0.				90
	482.1		-0.698		444.5	5.909				
	30		10		1500	1	10	3	1	5
	20.0		-30.		1.	0.10				1
	4090.		10.		4000.	240.0				
SanAnt01_ArH			47946							
J4514010_FrH			22010							
Y5235010_FrH			53000							
J4514010_FrD			3600							
Y1345010_FrH			35000							
A3301010_FrH			15000							
A1050310_FrD			15000							
A9091060_FrD			15000							

Figura 8.6. Archivo de entrada para calibración del tramo 01 del río San Antonio.

- Primer fila: El primer valor corresponde al tipo de modelo a emplear, IPH2 (1) o GR4 (2), o el modelo a calibrar, Mocom-IPH2 (3) o Mocom-GR4 (4). La segunda columna admite dos valores, 1 si se desea optimizar entre el modelo funcionando en simulación y en previsión o 2 si se optimiza el modelo en previsión pero en dos horizontes distintos de tiempo. El siguiente valor se utiliza para señalar la cantidad de cuencas que quieren correr. La columna 4 será tenida en cuenta únicamente cuando el primer valor de la fila sea 1 o 2, es decir, cuando únicamente se haga correr el

modelo sin calibrar, permite indicar si se quiere que el modelo funcione en modo simulación (1) o en previsión (2). La penúltima columna consiste en una variable entera que indica que cada un cierto número de ciclos completos del MOCOM, se guarda en un arreglo (matriz "PobEvol") la población de ese ciclo. Esto sirve para poder luego escribir en algún archivo de salida la evolución de la población a lo largo de los ciclos completos del MOCOM. Por ejemplo si este valor es igual a 10, cada 10 ciclos completos del MOCOM, se guarda una población. Si el ultimo valor de la fila vale uno, indica que en la carpeta "Control" se guardarán dos archivos por cada individuo de la población, un hidrograma simulado y uno pronosticado si se ha elegido previamente optimizar el modelo en modo simulación y previsión, si de lo contrario se opto por optimizar en dos horizontes distintos en previsión se guardarán dos hidrogramas pronosticados iguales, pero con distinto FO. Si en cambio este carácter es distinto de uno no se guardara ningún archivo en la carpeta.

- Cuarta fila: en la misma, como se explica oportunamente, se colocan todos los horizontes que se deseen tener disponibles para poder ser empleados, con la única condición de que sean escritos de menor a mayor.
- Segunda fila: Los valores aquí presentes llaman a la fila 4. Por ejemplo si el primer valor de esta fila es 3 quiere decir que el horizonte de tiempo al cual se está refiriendo es el que se encuentre posicionado en la columna tres de la fila 4. La primera columna denominada "NTsim" corresponde al horizonte de previsión en el cual se calibrará el modelo funcionando en modo simulación, la segunda "NTprev" al horizonte de previsión que será acumulado para luego sacar en archivo de salida, la tercera "NTopt" al horizonte en el cual se calibrará el modelo funcionando en modo previsión, la cuarta "NTopt1" al horizonte de previsión de la FO1, y la ultima "NTopt2" al horizonte de previsión de la FO2. Las dos últimas columnas son tenidas en cuenta únicamente cuando se optimiza el modelo en modo previsión vs previsión.

- Tercera fila: está formada por seis valores. El primero representan la cantidad de Δt del tramo que se descarta para el cálculo de la FO (longitud de desfasaje). El segundo y tercer número son el código de FO1 y FO2, los cuales pueden tomar un valor entero de 1 a 4, Coef.Nash-Sut. (1), Coef de persits. (2), Coef.Nash-Sut RAIZ (3), Coef. de KGE. (4). En nuestro caso el coeficiente utilizado es el de KGE. El cuarto valor consiste en una variable entera que define de qué manera se va a genera la población inicial, si este vale 1 se elige de manera aleatoria de una librería de parámetros, si es 2 se extraen los individuos de una matriz de datos, en cambio si toma el valor 3 se genera una población a partir de los mejores individuos de la librería de parámetros, en caso de que el número sea distinto de los mencionados la población se creará de forma aleatoria con distribución probabilística aleatoria. En la quinta columna se indica la cantidad de datos que posee la librería de parámetros. El último, denominado fuga es una variable entera que permite salir de la optimización antes de encontrar todos los individuos no dominados. Esta variable está relacionada con TolFug1, TolFug2, y PND, que se verán más adelante. Por definición del método MOCOM, la optimización multi-objetivo concluye cuando se encuentra la región de Pareto, es decir, cuando todos los individuos de la población son no dominados. Pero hasta llegar a ese objetivo final a veces se requiere mucho tiempo, y realmente lo que nos interesa no es que todos sean no dominados, sino que sean buenos. Por eso, esta variable "Fuga" justamente permite (cuando Fuga=1) que el procedimiento de optimización se escape del ciclo y concluya antes de llegar a lograr la no-dominancia de todos los individuos, pero a condición de que el ajuste de FO1 y FO2 entre un ciclo y el siguiente sea menor que "TolFug1" y "TolFug2" respectivamente, y además que al menos un porcentaje "PND" de los individuos de la población sean no dominados.
- Quinta fila: el primer y segundo valor corresponden a "TolFug1" y "TolFug2" (tolerancia a la fuga), es decir, al mínimo porcentaje de cambio de FO1 y FO2 que se admite, si el porcentaje de cambio es menor al valor aquí indicado se termina la

calibración. En la tercer columna se ubica el valor que señala el porcentaje de individuos no dominado (PND) mínimo que permite la salida antes de que finalice por completo la optimización. El último valor de la fila es el porcentaje total de laguna que se asigna al caudal observado para simular ausencia de datos.

- Sexta fila: contiene todos los valores de los parámetros del modelo elegido ya sea IPH2 o GR. Si es IPH2.
- Séptima fila: 9 parámetros y datos que requiere el Algoritmo MOCOM. Número de individuos de la población (Nip), número de individuos del complejo (Nic), número de iteraciones en que son tomados los complejos de la población (NiterP), número de veces en que es sometido el Complejo al algoritmo de Nelder y Mead (NiterC), número de veces que el vector aleat unif se automultiplica (Nunif), número de ajustes del unto de reflexión intentando obtener un punto no dominado (Najr), número de ajustes del punto de contracción intentando obtener un punto no dominado (Najc), número de ajustes para lograr que el nuevo punto de reflexión caiga dentro del espacio de validez de parámetros (Rfint). La última columna se denomina "Simp", si vale 1 el programa elige la forma simplificada de optimizar, o sea adopta directamente un punto de contracción, si simp toma otro cualquier otro valor se evalúa si el punto de contracción es malo y se elige uno de manera aleatoria.
- Octava y novena fila: límites inferiores y superiores de los parámetros, cualquiera sea el modelo elegido (IPH2 o GR).
- Decima fila en adelante: Lista de códigos de las cuencas a leer y a su izquierda la cantidad de datos que posee el archivo.

5- Main_exe: archivo ejecutable que resulta de la programación en lenguaje Fortran.

6- Librería Horaria, Librería Diaria, Matriz, Poblac: son archivos que contienen una cierta cantidad de individuos con el valor de sus cuatro parámetros (X1, X2, X3 y X4) desde los que se extrae la población inicial.

8.2. CUENCA DEL RÍO SAN ANTONIO

8.2.1. CALIBRACIÓN

A la hora de realizar la calibración del modelo hidrológico para la cuenca del río San Antonio se realizaron diversas modificaciones en el archivo de entrada con la intención de que el programa converja rápidamente sin tanto costo de tiempo computacional. Las variables de ajuste elegidas para ir tendiendo a los resultados más adecuados fueron las siguientes:

- Número de iteraciones
- Número de individuos del complejo
- La manera en la que se genera la población inicial
- Número de ajustes del punto de reflexión intentando obtener un punto no dominado.
- Número de ajustes del punto de contracción intentando obtener un punto no dominado.
- Número de ajustes para lograr que el nuevo punto de reflexión caiga dentro del espacio de validez de parámetros.

Se llevó a cabo la optimización en los cuatro tramos en las que fue dividida la serie. En la siguiente tabla se presentan los valores de los parámetros del archivo de entrada de la corrida que arrojó los mejores resultados en cuanto a valor de las FO y a la cantidad de individuos no dominados para cada uno de los tramo.

RIO SAN ANTONIO	N° DE ITERAC.	N° DE IND.	GENERACION DE POBL. INICIAL	N° DE AJUSTES DEL PTO DE REFL.	N° DE AJUSTES DEL PTO DE CONTRA.	PTO DE REFL. DENTRO DEL ESPACIO DE VALIDEZ
TRAMO 01	1000	50	Aleatoriamente de librería de parámetros.	3	1	5
TRAMO 02	1000	50	Aleatoriamente de librería de parámetros.	3	1	5
TRAMO 03	1000	50	Aleatoriamente de librería de parámetros.	3	1	5
TRAMO 04	1000	50	Aleatoriamente de librería de parámetros.	3	1	5

Luego de cada corrida se realizó un gráfico donde se representan a los individuos de la población a partir de sus valores de función objetivo 1 y 2. Como resultado se obtuvo, que las corridas de 3 tramos finalizaron correctamente la calibración del modelo, formándose así en el gráfico, una región de Pareto. Sin embargo, los valores tanto de FO1 como FO2 son considerablemente bajos, por lo que no se logra una buena correlación entre los caudales observados y los pronosticados. Con respecto al otro tramo, el denotado 01, no pudo ser optimizado, es decir, no se halló la región de Pareto y además los valores de FO resultaron bajos con respecto a lo esperado.

Los tramos 02, 03 y 04 se optimizaron en la primer y única corrida, mientras que para la calibración de tramo 01 se realizaron numerosos intentos, ninguno con éxito.

Es necesario aclarar, que durante una corrida puede ocurrir que a partir de un momento la variación de los valores de FO de una iteración a otra sea significativamente escasa. Por lo que para lograr la calibración requeriría de demasiado tiempo. Para el tramo 01 se han realizado más de 5 corridas y las mismas tuvieron una duración de entre 3 y 4 días, y además se detectó que la posición de los individuos en el gráfico FO1 vs FO2 prácticamente no cambiaba desde la iteración 400 hasta la 2000. Por lo que seguir probando significaría mucho costo de tiempo y aun así tal vez, no se lograría obtener los resultados esperados.

A continuación se representan en gráficos de dispersión los individuos a partir de sus valores de funciones objetivos 1 y 2 para los 4 tramos.

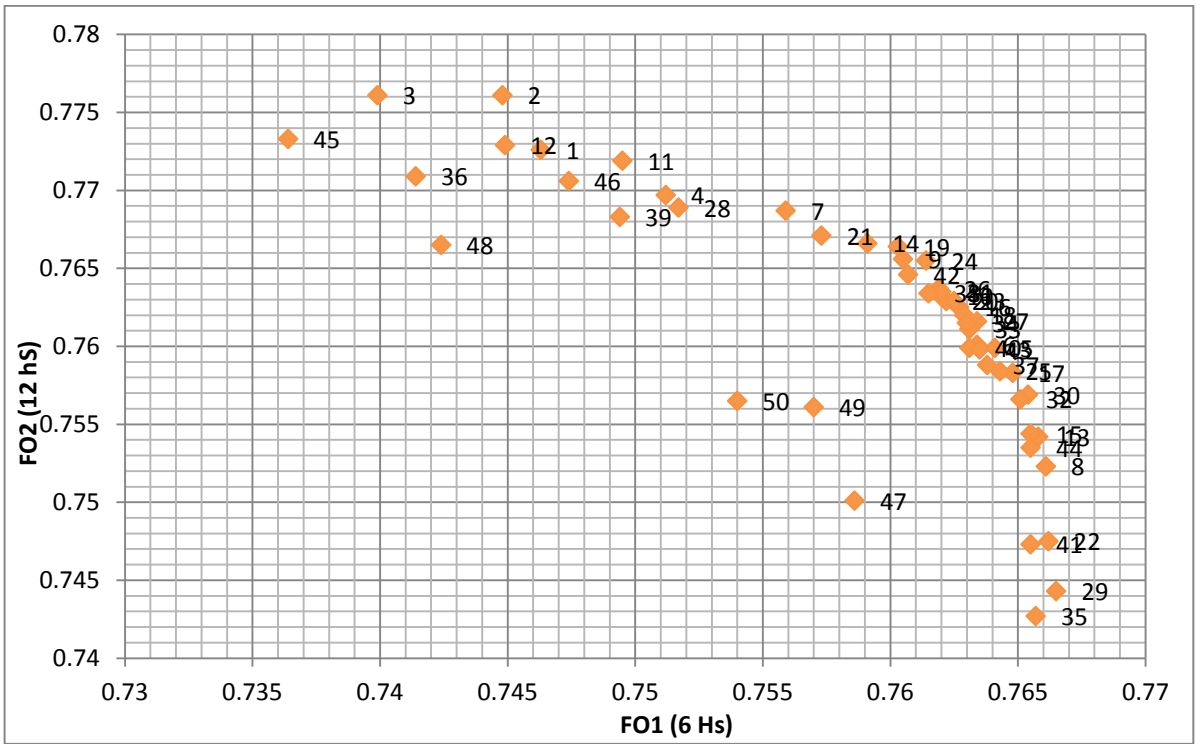


Gráfico 8.1. FO1 vs FO2. Tramo 01. Iteración 2001.

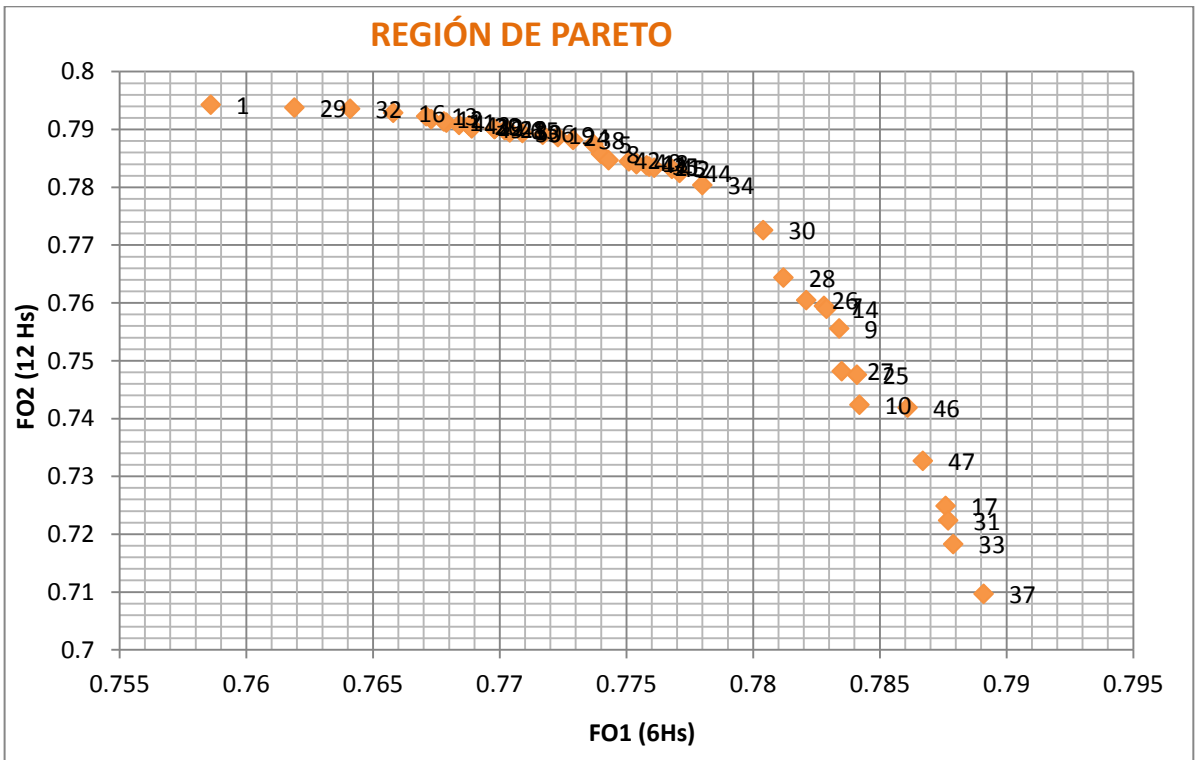


Gráfico 8.2. Región de pareto del tramo 02. Iteración 289.

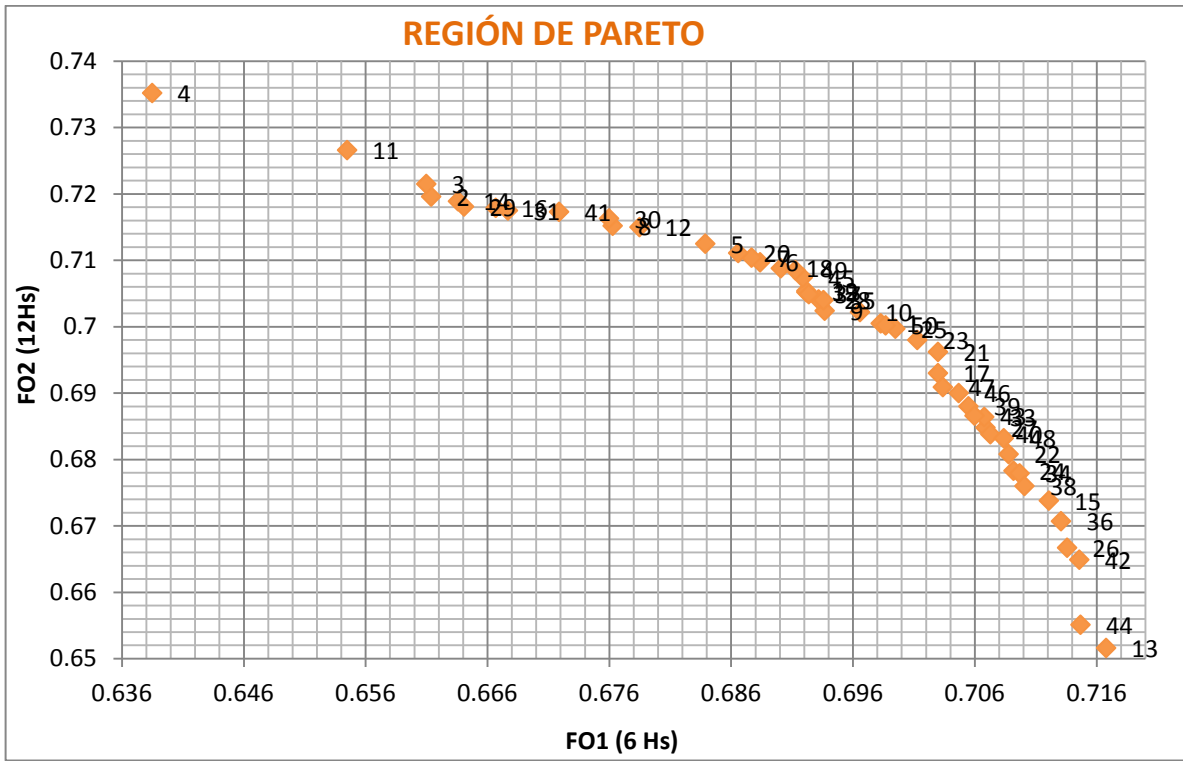


Gráfico 8.3. Región de Pareto del tramo 03. Iteración 603.

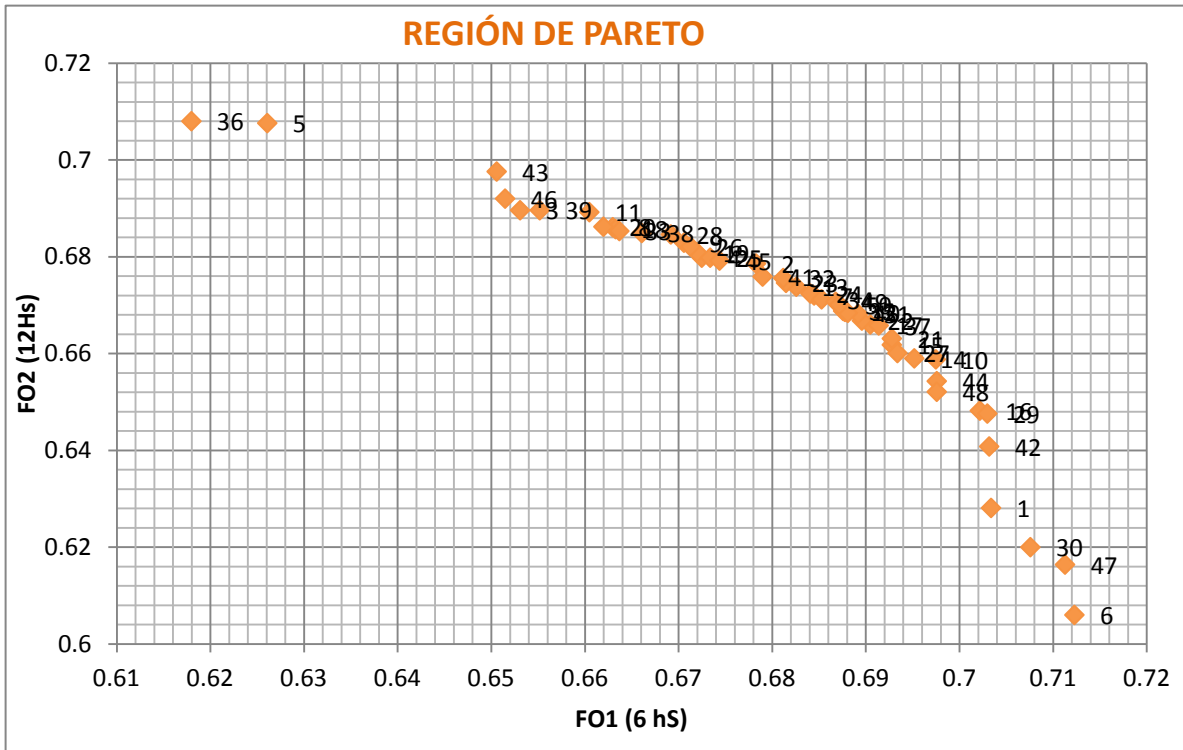


Gráfico 8.4. Región de Pareto del tramo 04. Iteración 183.

En todos los gráficos anteriores se ve que los valores de FO2 son mayores que los de FO1, excepto en el Gráfico 8.4. Esto es contrario a lo que se esperaba, de que el pronóstico a 6 horas tenga una mejor correlación. Se especula con que esa aparente falta de coherencia en los resultados puede estar relacionada a que la escala de respuesta de los hidrogramas del río San Antonio (la cual es infra-horaria) es notoriamente inferior al tiempo de horizonte pronosticado. Es decir que, desde este punto de vista, la actualización del estado del modelo hora a hora, no sólo deja de contribuir a la mejora de la performance del pronóstico, sino que por el contrario condiciona su capacidad de simulación.

8.2.2. MODELACIÓN EN MODO PRONÓSTICO

En este apartado se muestran y analizan algunos hidrogramas para 4 horizontes de tiempo (1, 6, 12 y 24 horas), obtenidos durante la calibración del modelo hidrológico en modo pronóstico. En los mismos se encuentra además, representada la precipitación correspondiente a esas horas.

8.2.2.1. Tramo 02

En la siguiente tabla se encuentran los valores de los parámetros de los individuos elegidos para graficar los hidrogramas y el de las funciones objetivos uno y dos.

INDIV.	JUEGOS DE PARAMETROS				FO1	FO2
	X1	X2	X3	X4		
IND. 1	20.9663	-0.0814	24.5818	7.2647	0.7586	0.7943
IND. 37	20.3295	0.0282	23.9763	7.0878	0.7891	0.7097

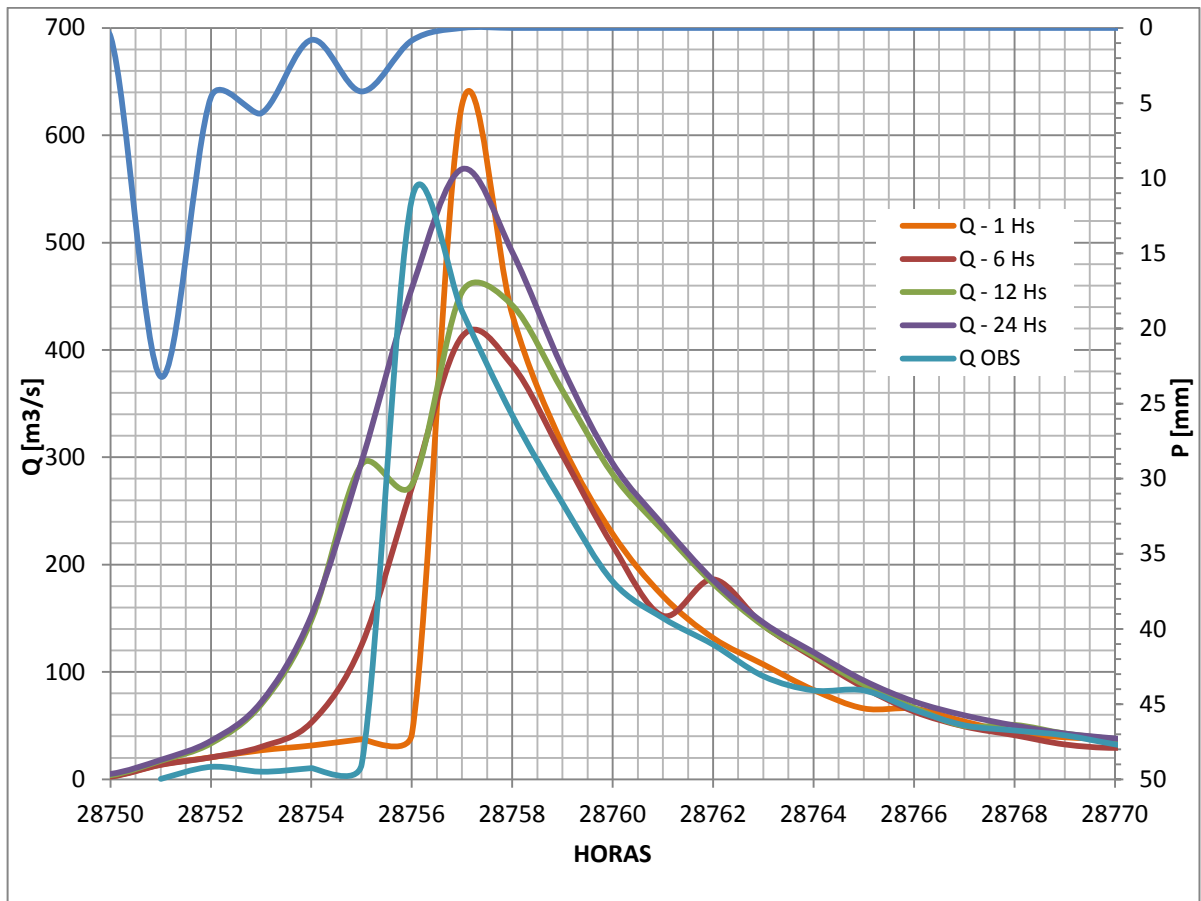


Gráfico 8.5. Hidrograma individuo 1. Tramo 02.

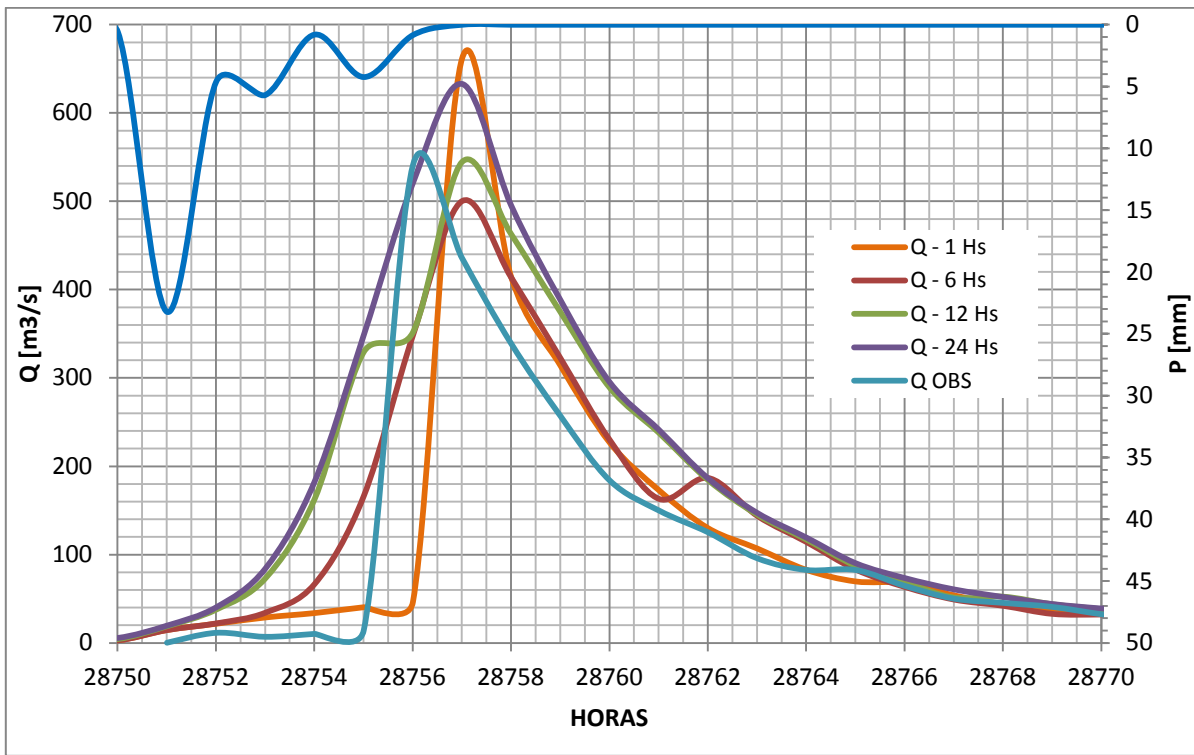


Gráfico 8.6. Hidrograma del individuo 37. Tramo 02.

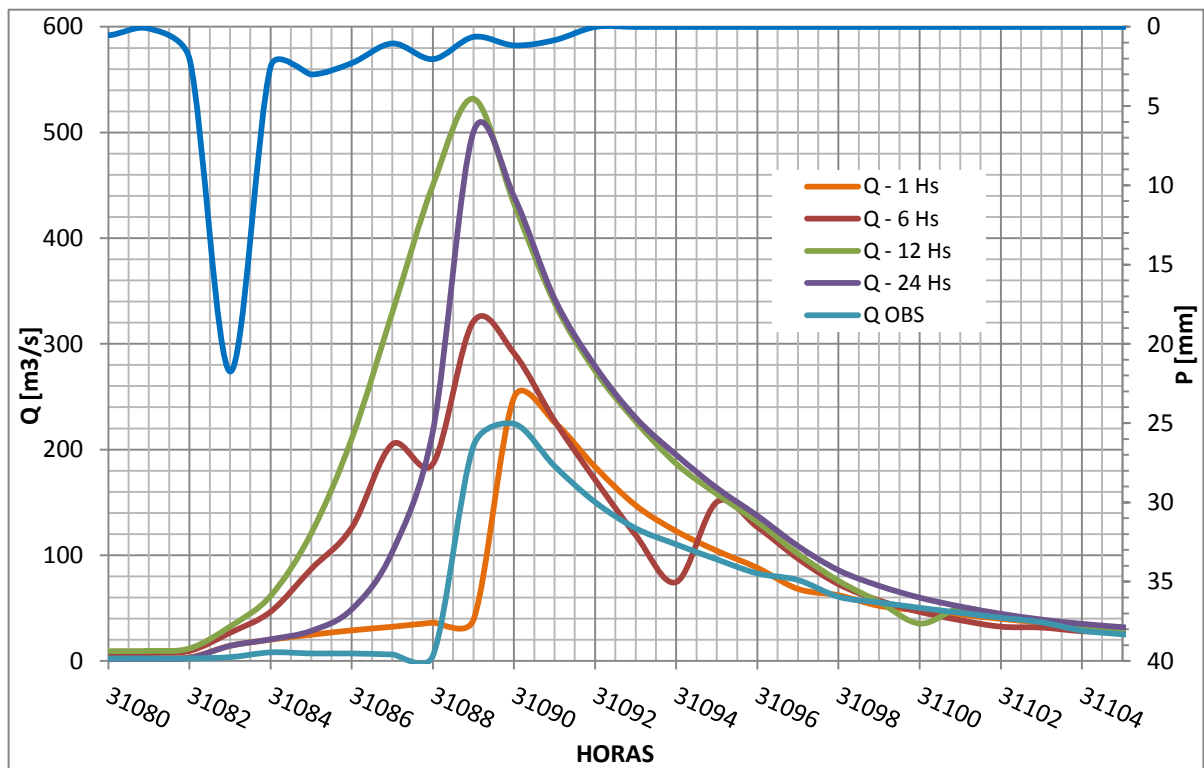


Gráfico 8.7. Hidrograma del individuo 1. Tramo 02.

Se observa en los Gráfico 8.5 y Gráfico 8.6 que la ubicación del pico del hidrograma observado no coincide con los demás, con los pronosticados a distintos horizontes de tiempo, el desfase es de aproximadamente 1 hora. Mientras que, en el Gráfico 8.7, se puede ver que los picos se encuentran aproximadamente alineados excepto el de 1 hora.

En cuanto a los valores de los caudales máximos se observa que el Gráfico 8.6, a diferencia del Gráfico 8.5, presenta menor distancia entre el máximo de los caudales observados y el de los pronosticados a 6 y 12 horas. Al analizar el Gráfico 8.7 se puede distinguir que los caudales máximos pronosticados distan significativamente del valor observado, en un 50 (pronóstico a 6 Hs) y 140 (pronóstico a 12 Hs) por ciento más. Mientras que en los dos gráficos anteriores estos valores no superan el 25 por ciento de diferencia.

En la siguiente tabla se pone de manifiesto en forma clara y sencilla lo expuesto en el párrafo anterior. Se indica el porcentaje de diferencia que existe entre los caudales pronosticados y el observado.

N° de Gráf.	Q OBS	Q Pronosticado	IND. 1	IND. 37
8.5 y 8.6	550	6 Hs	-24%	-10%
		12 Hs	-17%	-1%
8.7	220	6 Hs	47%	-----
		12 Hs	140%	-----

Se puede observar que para los Gráfico 8.5 y Gráfico 8.6 el individuo 37 es el que mejor representa en cuanto a la magnitud del pico. Además, a diferencia de lo que se esperaba, la mayor diferencia se muestra para el pronóstico a 6 horas, lo que no parece razonable ya que mientras menor sea el horizonte de pronóstico, mejor se espera que sea la correlación.

8.2.2.2. Tramo 03

En este caso, se decidió graficar los individuos que poseen el mejor valor de una de las dos funciones objetivos 4 (mejor FO2) y 13 (mejor FO1).

INDIV.	JUEGOS DE PARAMETROS				FO1	FO2
	X1	X2	X3	X4		
IND 4	29.6484	-0.0567	19.4847	9.8325	0.6385	0.7352
IND 13	29.8271	-0.0091	13.5384	10.4747	0.7168	0.6516

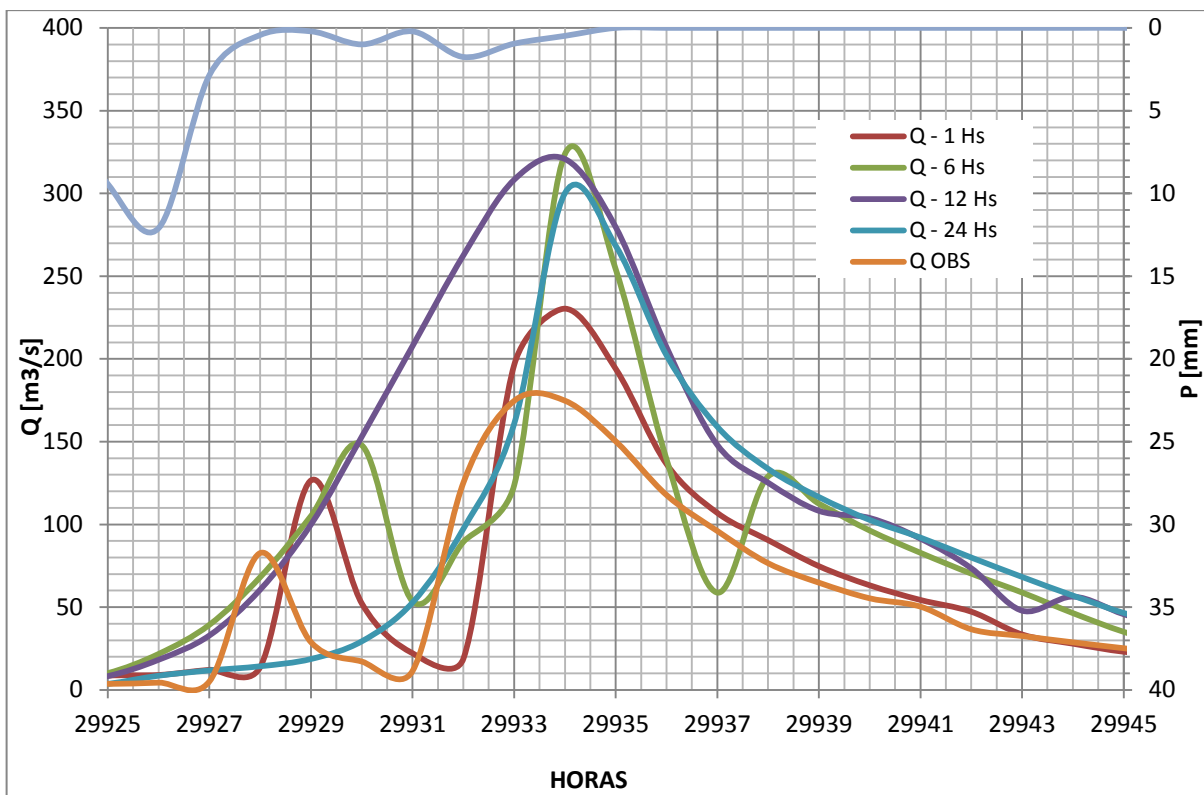


Gráfico 8.8. Hidrograma del individuo 4. Tramo 03.

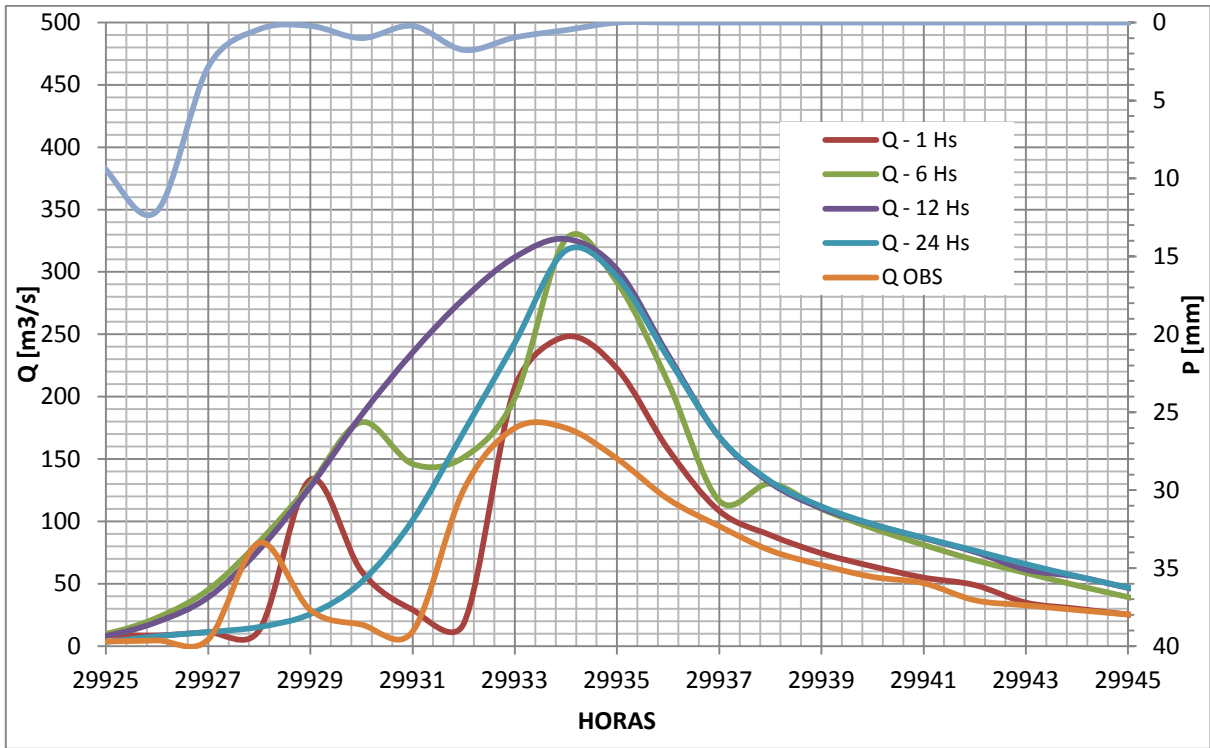


Gráfico 8.9. Hidrograma del individuo 13. Tramo 03.

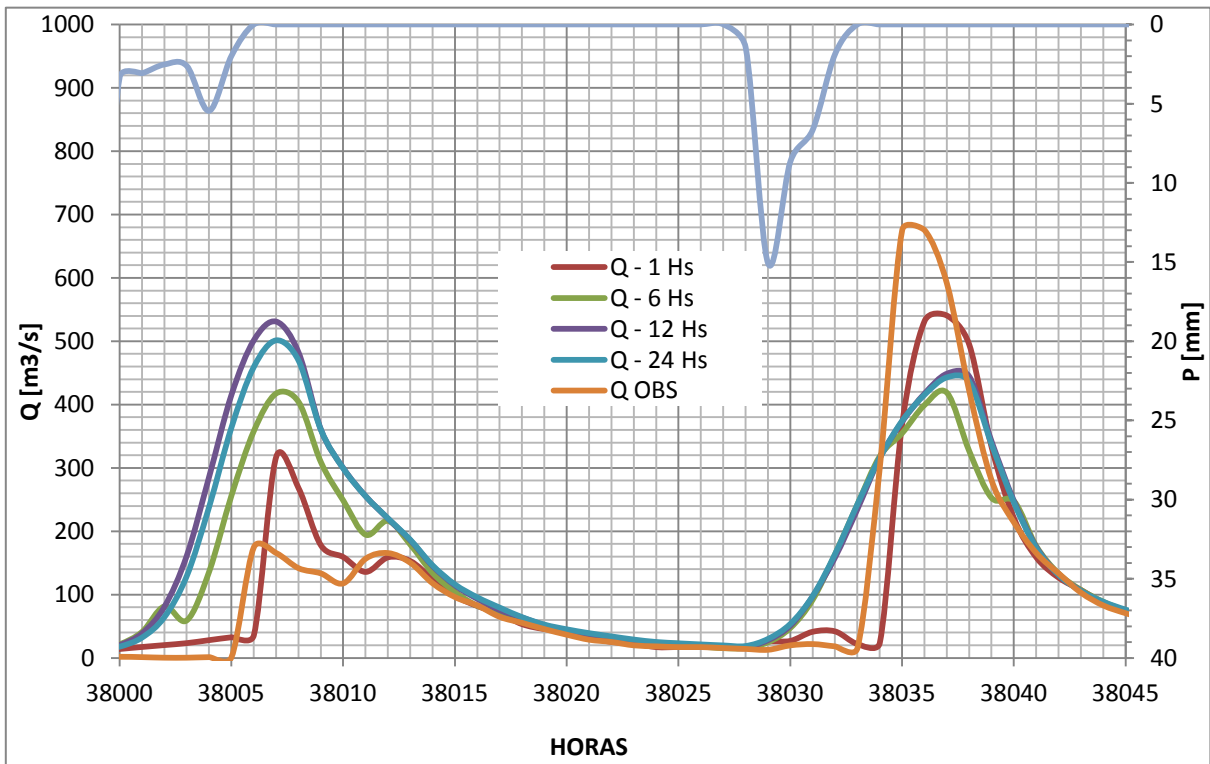


Gráfico 8.10. Hidrograma del individuo 04. Tramo 03.

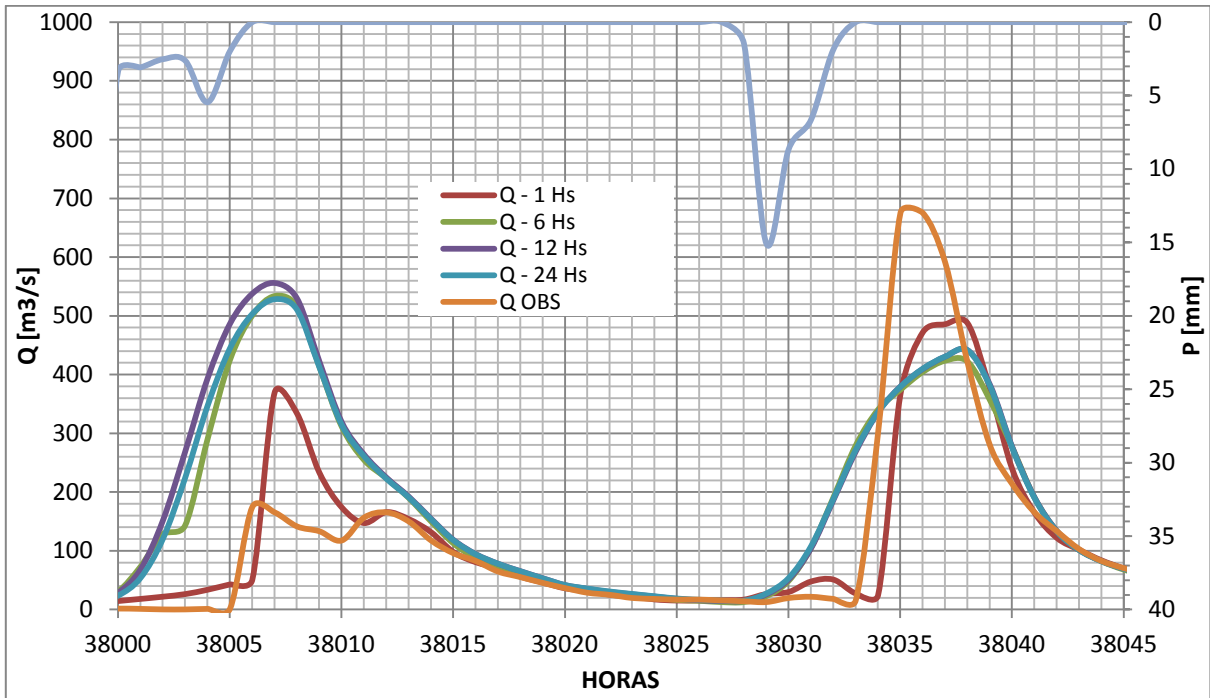


Gráfico 8.11. Hidrograma del individuo 13. Tramo 03.

En los gráficos anteriores se puede apreciar que en todos ellos los desfases del caudal observado con el pronosticado a 6 horas es de 1 hora, y para 12 horas es de 30 minutos aproximadamente. Además, es notoria la diferencia existente entre los caudales máximos observados y los pronosticados. En el siguiente gráfico se puede apreciar claramente.

N° de Gráf.	Q OBS	Q Pronosticado	IND. 4	IND. 13
8.8 y 8.9	180	6 Hs	82%	83%
		12 Hs	77%	82%
8.10 y 8.11	180	6 Hs	133%	194%
		12 Hs	194%	208%

Al ver la tabla anterior se aprecia que el individuo 4 es el que mejor correlación presentan, sin embargo también posee un alto porcentaje de diferencia entre los caudales pronosticados y el observado. En los Gráfico 8.8. Hidrograma del individuo 4. Tramo 03. y Gráfico 8.9, la mejor representación del hidrograma observado por el pronosticado a 12 horas en cuanto al valor del caudal máximo, lo cual no es lo que se esperaba como ocurrió en los gráficos del tramo 02. Mientras que en los gráficos Gráfico 8.10 y Gráfico 8.11 el

hidrograma a 6 horas, en comparación con el de 12, es el que mejor representa al observado.

8.2.2.3. Tramo 04

Como se ve en la siguiente tabla, los individuos elegidos para realizar la ejemplificación de los hidrogramas obtenidos durante la calibración son el 6 (mejor FO1) y el 36 (mejor FO2).

INDIV.	JUEGOS DE PARAMETROS				FO1	FO2
	X1	X2	X3	X4		
IND 6	20.2882	0.0391	14.5342	9.6861	0.7123	0.606
IND 36	21.3765	-0.1153	19.5438	8.6088	0.618	0.708

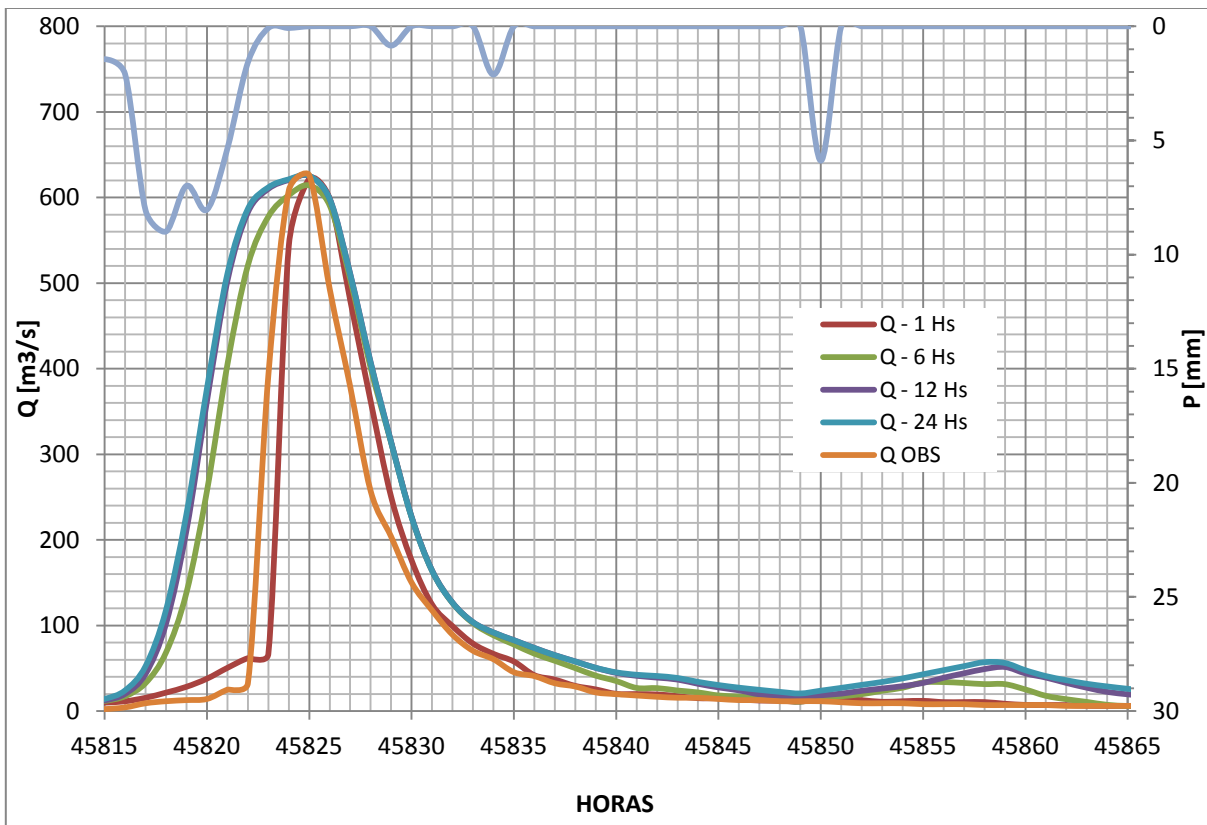


Gráfico 8.12. Hidrograma del individuo 6. Tramo 04.

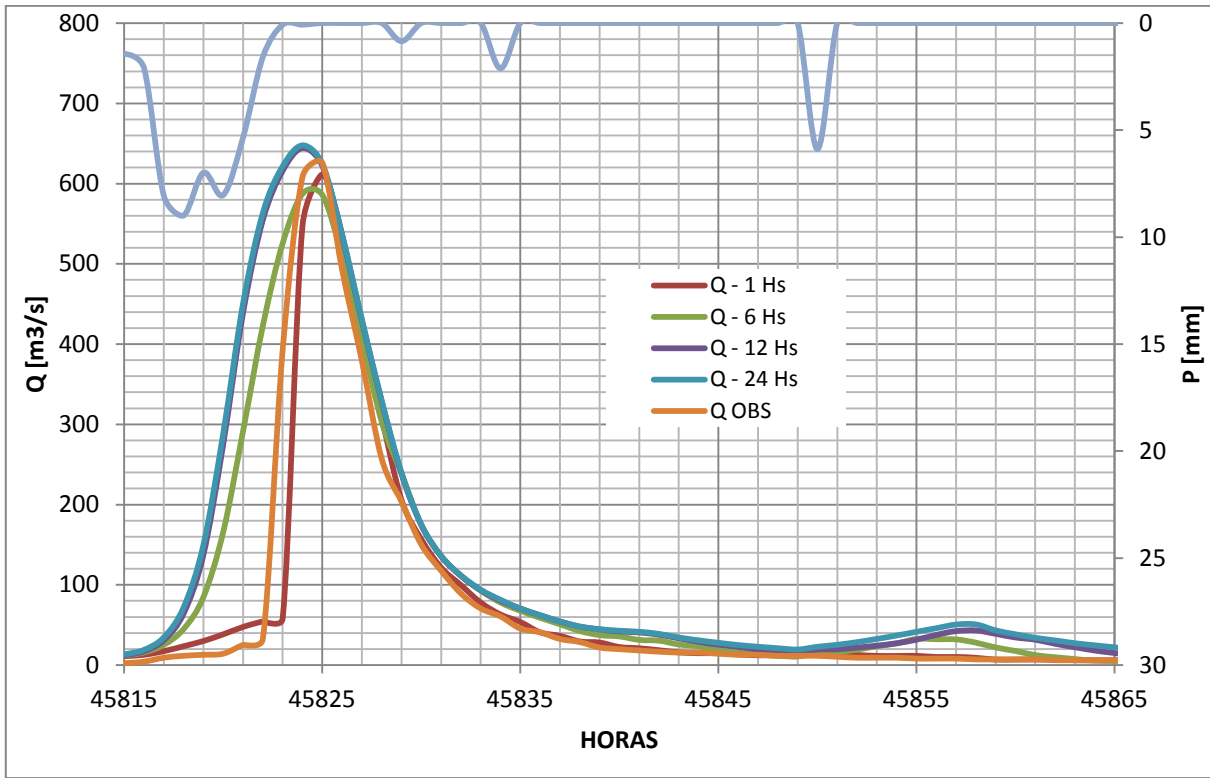


Gráfico 8.13. Hidrograma del individuo 36. Tramo 04.

A diferencia de los casos anteriores, en el tramo 04 se logró una mejor correlación, los picos de los hidrogramas se encuentran muy cerca unos de otros, tanto en el valor de los caudales como en horas. Esto puede verse en la siguiente tabla, donde los porcentajes son muy bajos.

N° de Gráf.	Q OBS	Q Pronosticado	IND. 6	IND. 36
8.12 y 8.13	631	6 Hs	-3%	-6%
		12 Hs	-1%	2%

Al igual que en la mayoría de los gráficos vistos en los tramos anteriores, los menores porcentajes se obtienen con el hidrograma pronosticada a 12 horas, y además el individuo 6 es el que mejor ajusta de los dos tanto para el pronóstico a 6 horas como a 12.

8.3. CUENCA DEL GUALEGUAY

8.3.1. CALIBRACIÓN

Al igual que en el caso anterior, para llevar a cabo la calibración del modelo hidrológico para la cuenca del río Gualeguay se realizaron diversas modificaciones en el archivo de entrada con la intención de que el programa converja rápidamente sin demasiado costo de tiempo computacional. Las variables que fueron elegidas para alterarse en las distintas corridas son las siguientes:

- Número de iteraciones
- Número de individuos del complejo
- La manera en la que se genera la población inicial
- Número de ajustes del punto de reflexión intentando obtener un punto no dominado.
- Número de ajustes del punto de contracción intentando obtener un punto no dominado.
- Número de ajustes para lograr que el nuevo punto de reflexión caiga dentro del espacio de validez de parámetros.

En la siguiente tabla se presentan los valores de los parámetros del archivo de entrada de la corrida que arrojó los mejores resultados para cada tramo.

RIO GUALEGUAY	N° DE ITERAC.	N° DE IND	GENERACION DE POBL. INICIAL	N° DE AJUSTES DEL PTO. DE REFL.	N° DE AJUSTES DEL PTO. DE CONTRA.	PTO. DE REFL. DENTRO DEL ESPACIO DE VALIDES
TRAMO 01	15000	50	Aleatoriamente de librería de parámetros.	10	10	10
TRAMO 02	30000	50	Aleatoriamente de librería de parámetros.	10	10	10

A continuación se representan en gráficos de dispersión las funciones objetivas 1 y 2 de cada uno de los individuos de la población. Cada uno de ellos se encuentra identificado con su correspondiente número.

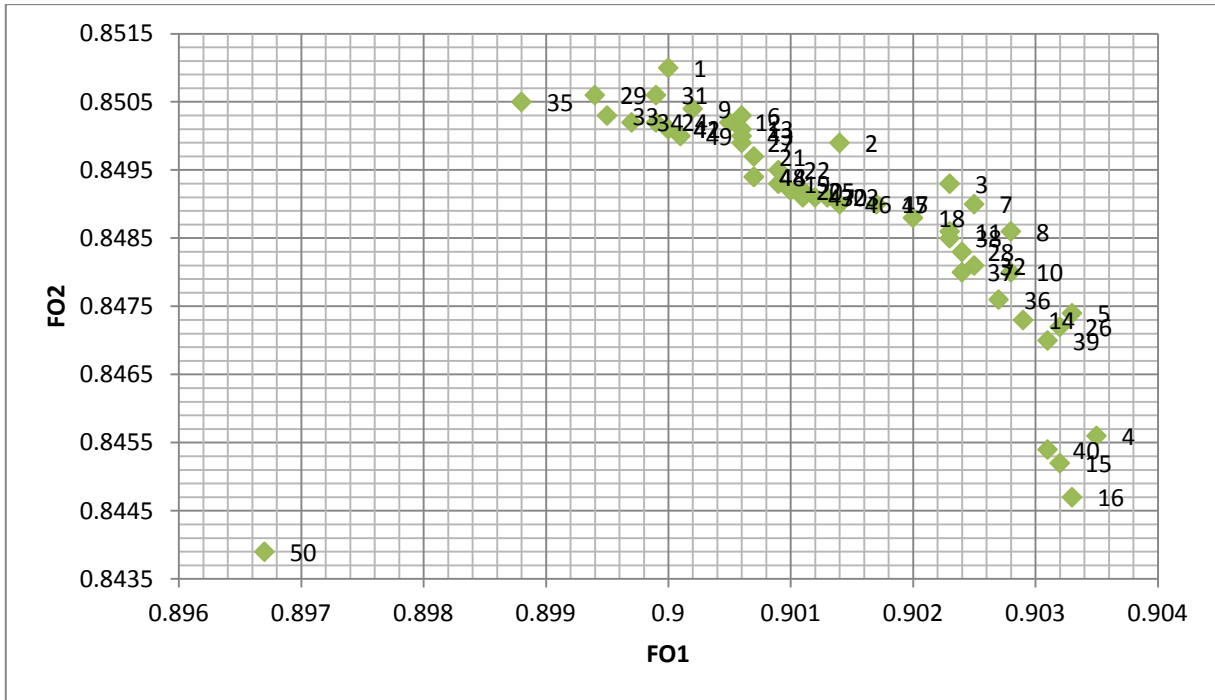


Gráfico 8.14. FO1 vs FO2. Tramo 01. Iteración 15001.

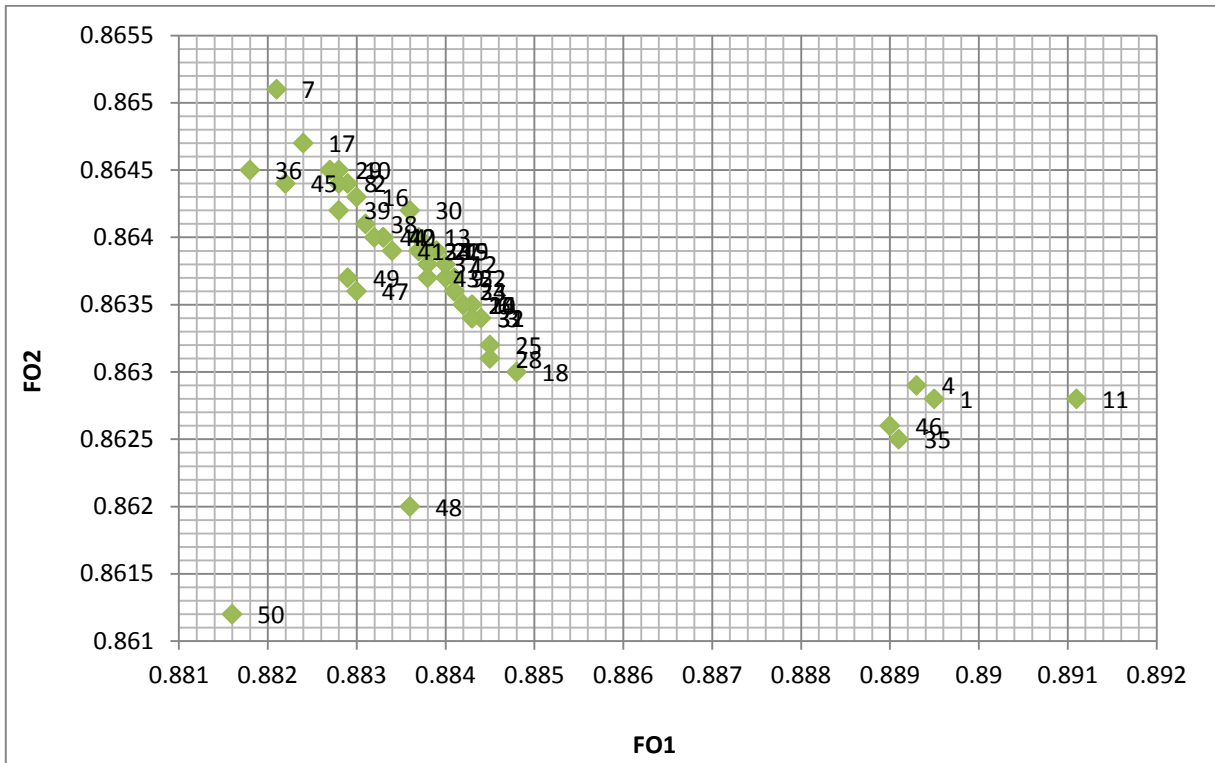


Gráfico 8.15. FO1 vs FO2. Tramo 02. Iteración 30001.

Como se puede apreciar en el Gráfico **8.14** y Gráfico **8.15** en ninguno de los dos tramos en la que fue dividida la serie se logró llegar a la optimización. Si bien se puede observar que los individuos se encuentran distribuidos de manera bastante próxima a la región de Pareto, por más que se siguiera dejando correr el programa con un mayor número de iteraciones, no se hubiera obtenido mejores resultados ya que llega un momento en el que las variaciones en cuanto al valor de las FO y cantidad de individuos no dominados desde una iteración a otra son mínimas.

Con la intención de clarificar lo explicado anteriormente se realiza un gráfico donde se esquematiza la evolución de la población del tramo 01. Al ver el Gráfico 8.16 se puede distinguir que los individuos de todas las iteraciones excepto la 10 se agrupan en un mismo sector.

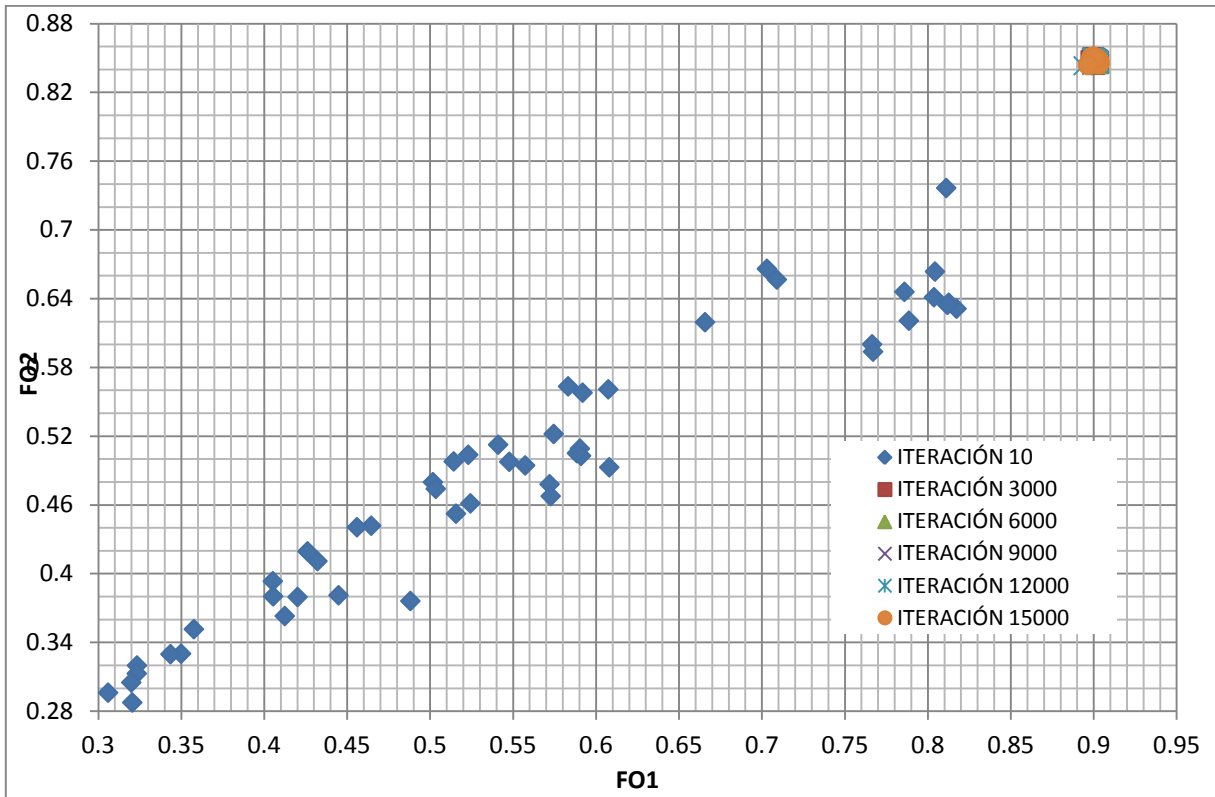


Gráfico 8.16. Evolución de la población. Tramo 01.

Con la intención de clarificar lo explicado anteriormente se realiza un gráfico donde se esquematiza la evolución de la población del tramo 01. Al ver el Gráfico 8.16 se puede distinguir que los individuos de todas las iteraciones excepto la 10 se agrupan en un mismo sector.

Si se realiza un acercamiento en el gráfico anterior en aquel sector donde se acumulan los individuos de todas las iteraciones se puede apreciar mejor los cambios que se producen al aumentar el número de iteraciones. Al ver el Gráfico 8.17 se puede advertir que los cambios en cuanto los valores de las funciones objetivas no varían significativamente desde la iteración 3000 a la 1500.

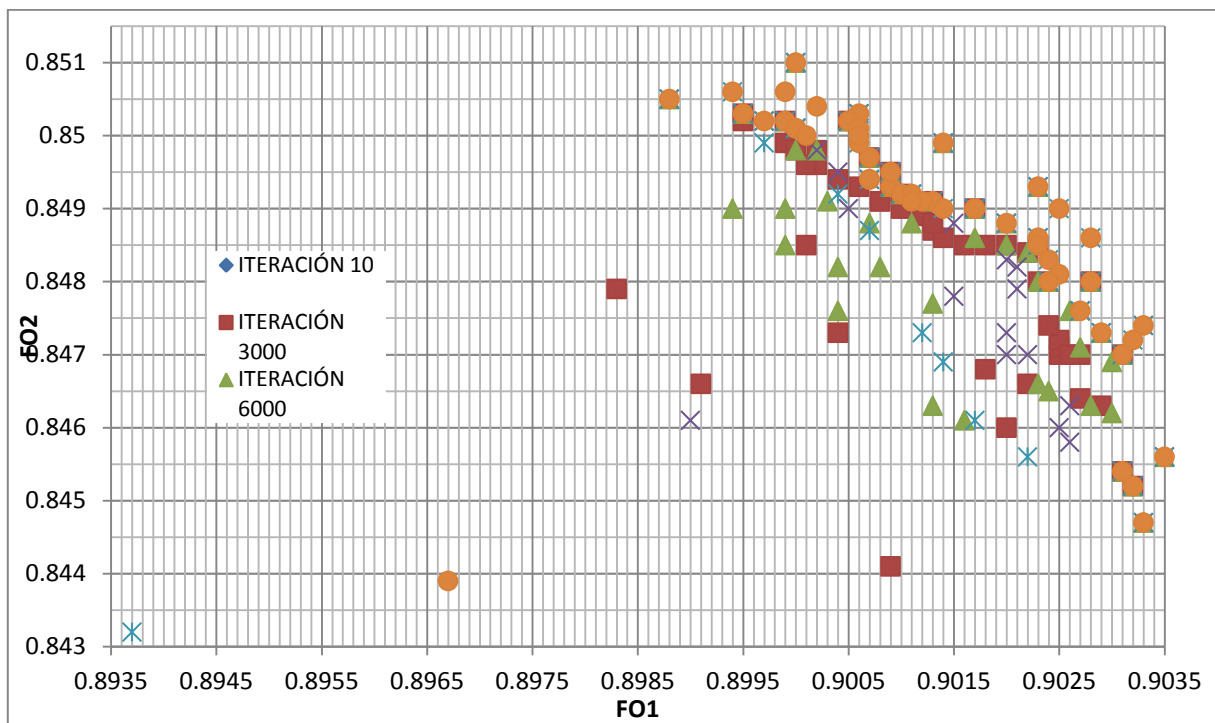


Gráfico 8.17. Evolución de la población. Tramo 01.

Además, puede verse en las gráficas Gráfico 8.14 y Gráfico 8.15 que a diferencia de la optimización del río San Antonio, los valores de la función objetivo 1 son mayores a los de FO2, lo cual es lógico y esperable ya que FO1 es la función objetivo para una previsión a 6 horas y FO2 a 12 horas.

8.3.2. MODELACIÓN EN MODO PRONÓSTICO

En lo que sigue se mostrarán y analizarán los hidrogramas para 4 horizontes de tiempo (1, 6, 12 y 24 días), obtenidos durante la calibración del modelo hidrológico en modo pronóstico.

8.3.2.1. Tramo 01

INDIV.	JUEGOS DE PARAMETROS				FO1	FO2
	X1	X2	X3	X4		
IND. 1	104.6507	3.7235	251.0602	6.2086	0.9	0.851

IND. 4	105.9008	4.9072	272.7364	5.5304	0.9035	0.8456
--------	----------	--------	----------	--------	--------	--------

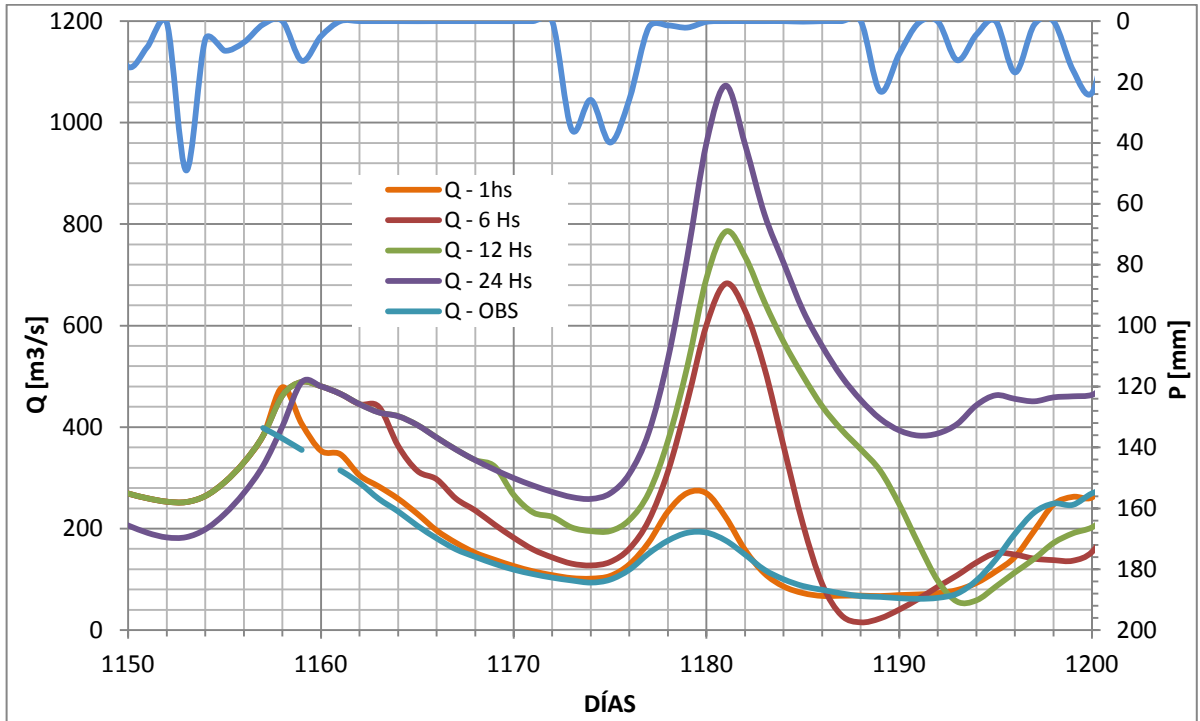


Gráfico 8.18. Hidrograma del individuo 1. Tramo 01.

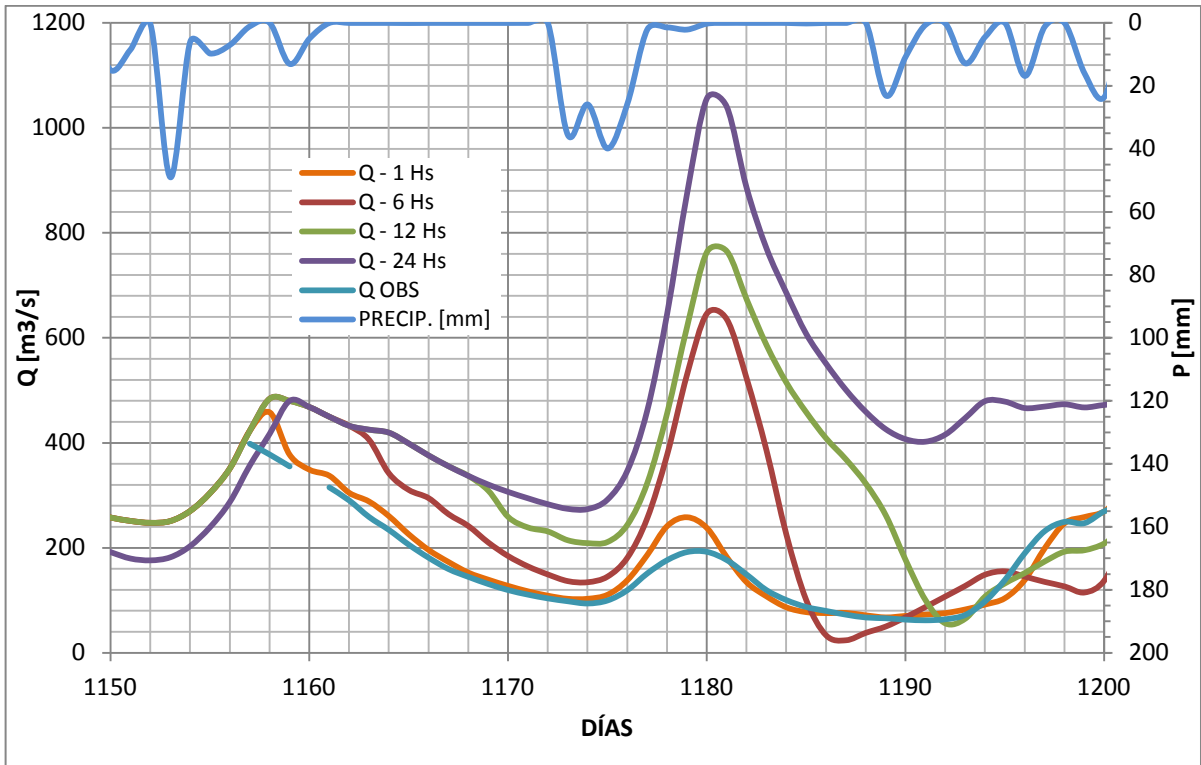


Gráfico 8.19. Hidrograma del individuo 4. Tramo 01.

Analizando los Gráfico 8.18 y Gráfico 8.19 se observa que hay una desviación en tiempo entre los picos de aproximadamente un día. Además a simple vista se puede ver que hay mucha diferencia en relación al valor de los caudales máximos, siendo, el caudal observado siempre superior por los pronosticados.

N° de Gráf.	Q OBS	Q Pronosticado	IND. 1	IND. 4
8.18 y 8.19	195	6 Hs	249%	233%
		12 Hs	305%	300%

Como era de esperarse el porcentaje es menor para la previsión a 6 días que para 12. El individuo 4 es el que posee un menor porcentaje de diferenciación para ambos días de pronóstico (6 y 12).

8.3.2.2. Tramo 02

Del Gráfico 8.15 se eligieron los individuos 7 y 11 para realizar el análisis de sus hidrogramas. En la siguiente tabla se encuentran los valores de los cuatro parámetros del modelo de ambos individuos y además el valor de las funciones objetivos 1 y 2.

INDIV.	JUEGOS DE PARAMETROS				FO1	FO2
	X1	X2	X3	X4		
IND 7	88.7095	2.5147	257.507	6.4391	0.8821	0.8651
IND 11	87.3773	2.6077	261.0729	6.4318	0.8911	0.8628

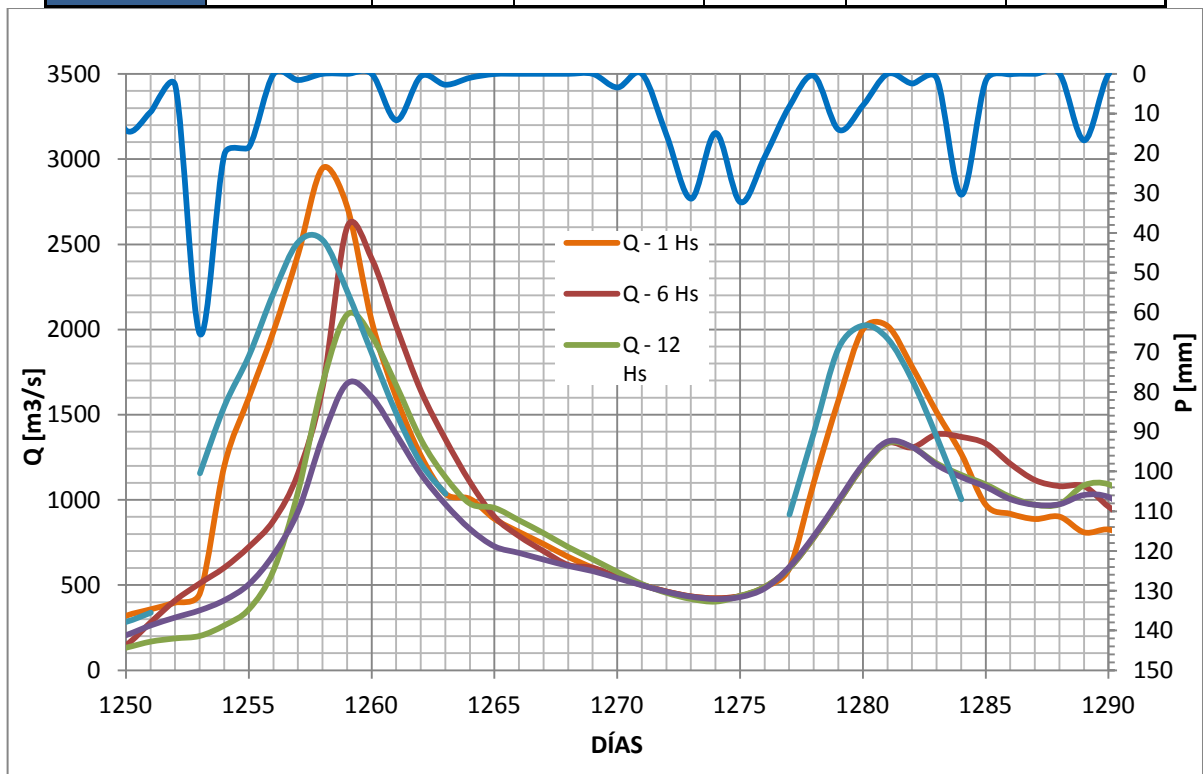


Gráfico 8.20. Hidrograma del individuo 7. Tramo 02.

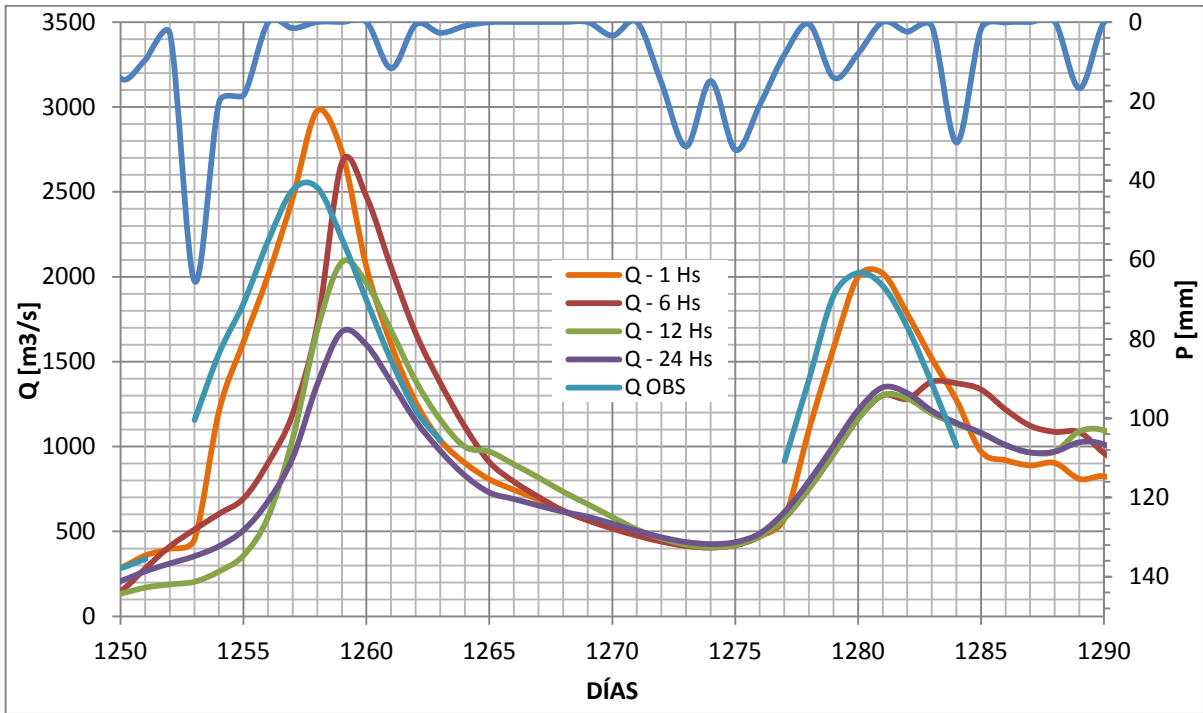


Gráfico 8.21. Hidrograma del individuo 11. Tramo 02.

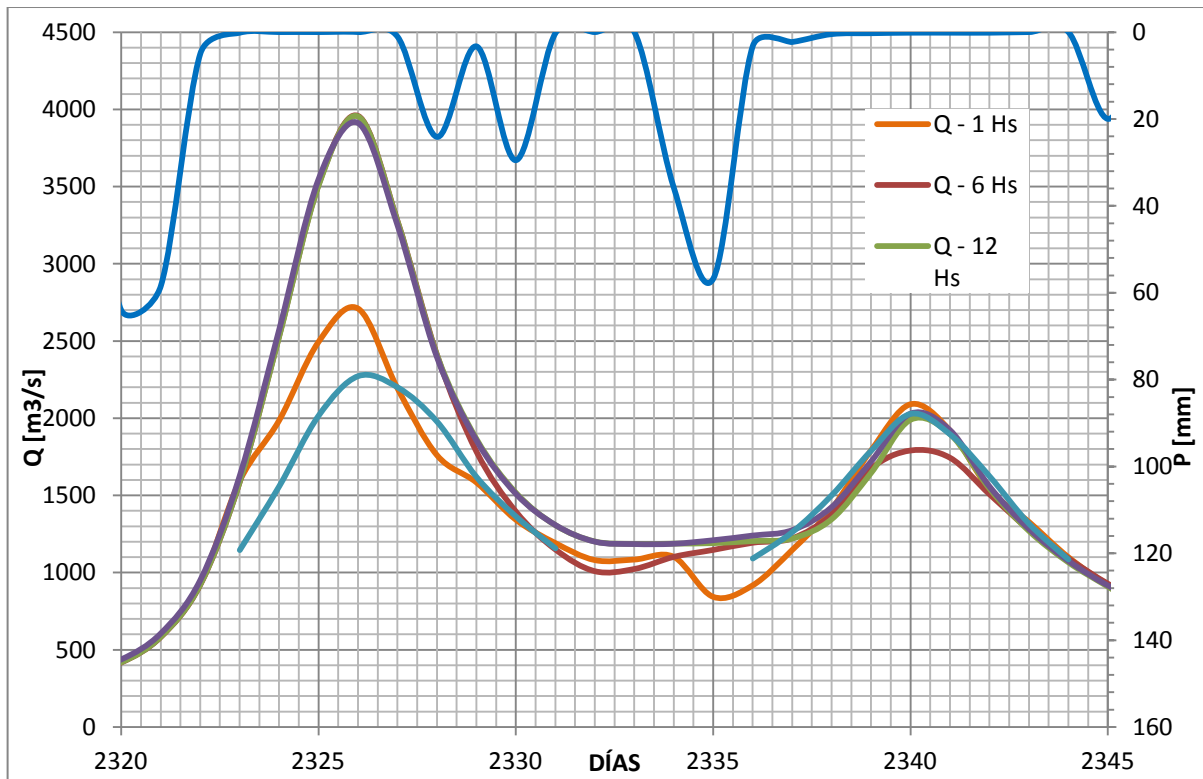


Gráfico 8.22. Hidrograma del individuo 7. Tramo 02.

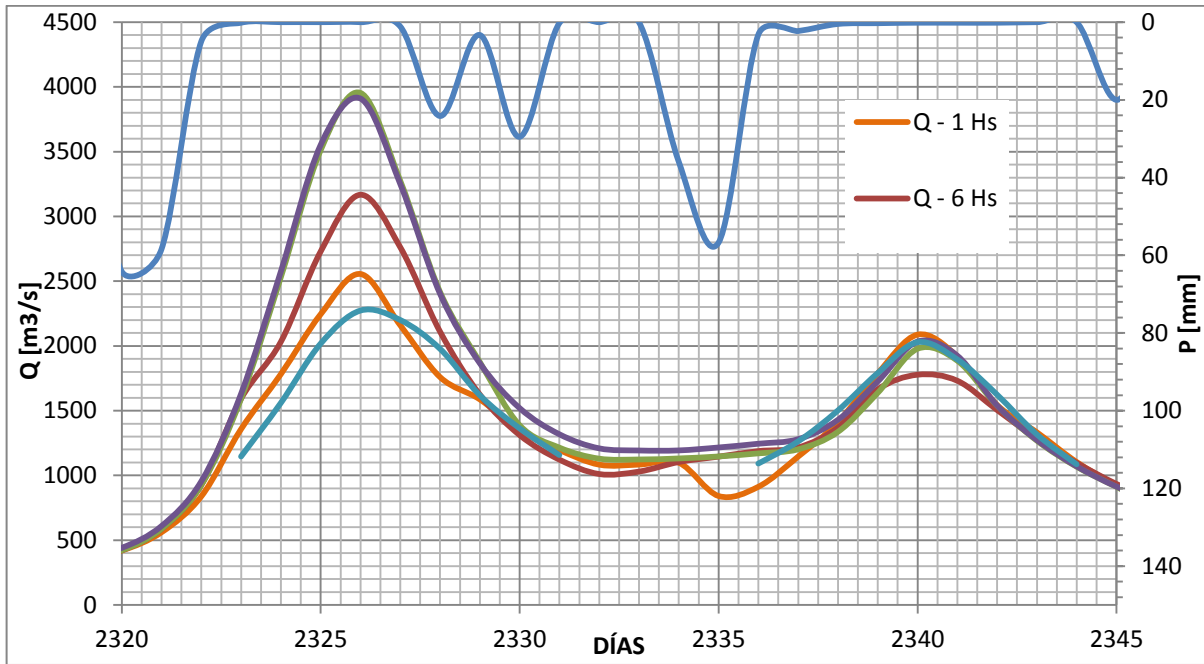


Gráfico 8.23. Hidrograma del individuo 11. Tamo 02.

Si se ven los Gráfico 8.20 y Gráfico 8.21 se puede apreciar que hay un desfase de dos días entre la ocurrencia del pico observado y los pronosticados a 6 y 12 horas. En los dos siguientes gráficos, el Gráfico 8.22 y Gráfico 8.23, no se distingue desviación entre los picos.

Para analizar la aproximación lograda con la previsión en cuanto a los caudales máximos se realiza la siguiente tabla con los dos caudales picos de los cuatro gráficos y los porcentajes de desviación entre el caudal observado y los pronosticados a 6 y 12 horas. La primera columna de cada individuo corresponde al primer caudal observado y la segunda al caudal número 2.

N° de Gráf.	Q Observado		Q Pronosticado	IND. 7		IND. 11	
	1	2		6 Hs	12 Hs	6 Hs	12 Hs
8.20 y 8.21	2250	2020	6 Hs	2%	-34%	6%	-36%
			12 Hs	-19%	-34%	-19%	-36%
8.22 y 8.23	2280	2020	6 Hs	73%	-11%	38%	-12%
			12 Hs	73%	-1%	73%	-2%

Se puede distinguir que la diferencia entre los Gráfico 8.20 y Gráfico **8.21** es mínima. En los Gráfico 8.22 y Gráfico **8.23** ocurre lo mismo, excepto con el pronóstico a 6 horas del primer caudal observado.

9. CONCLUSIÓN

El objetivo de la práctica profesional supervisada se divide en dos grandes ejes. El primero de ellos fue llevar a cabo la recopilación y procesamiento de la información de precipitación, caudal y evapotranspiración de dos cuencas argentinas: la del río San Antonio y la del Gualeguay. El segundo fue la evaluación de la performance del modelo hidrológico Génie Rural para adaptarse a distintos escenarios de pronóstico en ambas cuencas.

9.1. CONCLUSIONES SOBRE RECOPIACIÓN Y REVISIÓN DE INFORMACIÓN

Durante el presente trabajo se llevo a cabo la recopilación, actualización y depuración de los datos de la cuenca del río San Antonio. Mientras que para el río Gualeguay, únicamente se recopiló la información ya procesada y se la analizó para comprobar su confiabilidad.

Los datos recopilados del San Antonio fueron provistos por diversas fuentes, en distintos formatos y grado de procesamiento. Todas las series formadas a partir de dichos datos difieren unas de otras por lo que se debió adoptar un criterio a la hora de conformar la serie final. Es aquí donde se hace notar la importancia del criterio ingenieril y la necesidad de realizar un trabajo manual, para de esta forma poder detectar errores e inconsistencias.

Se pudo detectar, que la situación actual en Argentina, en cuanto a registros de datos es bastante deficiente, no solo por poseer largas lagunas, sino también porque contienen información dudosa y de escasa fiabilidad. Se observó a lo largo del procesamiento, la gran cantidad de errores presentes en los registros provenientes de las distintas estaciones de medición de la red telemétrica. Por lo cual se debió trabajar con suficiente minuciosidad y eliminar así, en lo posible, los datos erróneos, obteniéndose series más confiables.

Sin embargo, luego de todo el tratamiento realizado se puede observar cierta falta de calidad en los datos en ambas cuencas trabajadas.

9.2. CONCLUSIONES SOBRE LA CALIBRACIÓN

Durante esta segunda etapa se comprendió el funcionamiento del modelo hidrológico de gran difusión en Europa, llamado GRP (Génie Rural Prevision) y del método de optimización global multi-objetivo, MOCOM – UA.

Se obtuvieron los siguientes resultados durante la calibración del modelo:

- En algunos tramos no se logró completar la calibración ya que no se llegó a la región de pareto.
- En los gráficos de las regiones de pareto de la cuenca del río San Antonio se encontró que las funciones objetivos 1 son menores a las de FO2. En cambio, en los gráficos de la cuenca del río Gualaguay se obtuvo que los valores de FO1 son mayores a los de FO2.
- Por lo general, uno de los dos individuos elegidos por tramo, presenta un mejor ajuste que el otro, para ambos horizontes de tiempo, 6 y 12 pasos de tiempo.
- En cuanto a los hidrogramas pronosticados se pudo apreciar que existe una cierta diferencia con los observados. El modelo no ajusta adecuadamente el pico (en magnitud y en posición), la forma del hidrograma y tampoco la velocidad de vaciado de los reservorios.

Ante los resultados descritos anteriormente se plantearon las siguientes conclusiones:

- Se observa (a priori) una incapacidad del modelo GRP para pronosticar caudales en un horizonte de previsión igual o superior a 6 horas, en la cuenca del río San Antonio. Se presume que dicha incapacidad está relacionada a la escala temporal de respuesta (infra horaria) de las crecidas que registra esta cuenca.
- La desigual escala temporal entre el pronóstico y los tiempos de respuesta de la cuenca, convierte al modelo en una herramienta de simulación, y resta importancia al módulo de actualización de estado del modelo de pronóstico, el cual no sólo es

incapaz de mejorar la “peformance” de la estimación, sino que condiciona la capacidad de simular del modelo. Obteniendo de ese modo resultados en apariencia incoherentes, ya que manifiesta mejores aciertos cuanto mayores son los horizontes de pronóstico.

Como conclusión final, es importante destacar la necesidad de contar con un sistema de pronóstico temprano y confiable de crecidas, más aún, en todos aquellos ríos del país de respuesta rápida, para de esta manera llevar a cabo la evacuación de la zona esperada de inundación con éxito.

Si se pronostica un caudal máximo considerablemente menor al que realmente se genera, la zona inundada será mayor a la esperada, por lo que se pondría en riesgo la seguridad de personas no notificadas. Si de lo contrario, ocurriera que la crecida sea menor a la esperada, inundándose una menor superficie, generaría en la población una especie de incredulidad en el sistema de alerta, por lo que podría ser contraproducente para crecientemente venideras.

Esto demuestra la gran importancia que tiene en la práctica, la totalidad del trabajo realizado durante la elaboración de este informe, tanto el procesamiento y su análisis, como la calibración del modelo hidrológico utilizado.

10. REFERENCIAS BIBLIOGRÁFICAS

Alonso, F. (2008) “Calibración automática de un modelo hidrológico basada en algoritmos genéticos, aplicación a dos cuencas típicas argentinas”. *Tesis de maestría*.

Alonso, F. (2012). Informe Final de estadía en la Universidad Pierre et Marie Curie, Paris. Programa Movilidad de docentes a Paris. Ministerio de Educación de la Nación.

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”. *Prácticas supervisadas*.

Catalini, C., García, C., Ortega, F., Ottaviani, M., Aliaga, I., García, C. (2010) “Análisis de sensibilidad sobre parámetros de un modelo hidrológico conceptual para pronóstico de crecidas en una cuenca de montaña (río San Antonio, prov. de Córdoba)”, *XXIV congreso latinoamericano de hidráulica punta del este, uruguay, noviembre 2010*.

Colladon, L., Caamaño, N., Pazos (2007) “Modelos estadísticos para pronosticar crecidas a tiempo real en las Sierras Cordobesas”, *XXI Congreso Nacional del Agua CONAGUA 2007, Tucumán, República Argentina*.

Guillén, N., Patalano, A., Gracia, C. (2015) “Validación de la técnica PIV a gran escala para la estimación de los caudales en ríos de montaña”, *IV Simposio sobre métodos experimentales en hidráulica, La Plata, Argentina, marzo 2015*.

Juárez, N. (2013) “Calibración y modelación hidrológica preliminar con énfasis en el ordenamiento y control de datos. Aplicación a las cuencas de los ríos Salado (Santa Fe) y Gualaguay (Entre Ríos)”. *Prácticas supervisadas*.

Martín, M. (2013) “Calibración y modelación hidrológica preliminar con énfasis en el ordenamiento y control de datos. Aplicación a las cuencas de los ríos San Antonio (Córdoba) y Gualaguaychú (Entre Ríos)”. *Prácticas supervisadas*

Scielo Argentina (2015) [http://www.scielo.org.ar/scielo.php?pid=S1853-86652015000100008 & script = sci_arttext](http://www.scielo.org.ar/scielo.php?pid=S1853-86652015000100008&script=sci_arttext)

11. ANEXO

11.1. ANALISIS P – Q

En lo que sigue se presentan gráficos de P-Q de la cuenca del río San Antonio en donde se observan irregularidades.

