



Canal de  
Isabel II

## CUADERNOS DE I+D+i

# 13

Desarrollo de un sistema de validación,  
estimación y predicción de consumos  
horarios por sectores para la red de  
distribución de Canal de Isabel II







© Canal de Isabel II – 2011

**Autores:**

Alejandro Carrasco Mínguez  
Luis García Amor

**Dirección del estudio:**

Juan Carlos Ibáñez

**Agradecimientos:**

Agradecemos la colaboración en los trabajos que acompañaron la ejecución de este proyecto a Guillaume Raimbault y Bruno Rodríguez Rodríguez

Depósito Legal: M-39407-2011  
Maquetación: BLINK Comunicación S.L.  
Impresión: Estugraf Impresores S.L.



## CUADERNOS DE I+D+i

13

Desarrollo de un sistema de validación, estimación y predicción de consumos horarios por sectores para la red de distribución de Canal de Isabel II



**EXCLUSIÓN DE RESPONSABILIDAD**

Las afirmaciones recogidas en el presente documento reflejan la opinión de los autores y no necesariamente la de Canal de Isabel II.

Tanto Canal de Isabel II como los autores de este documento declinan todo tipo de responsabilidad sobrevenida por cualquier perjuicio que pueda derivarse a cualesquiera instituciones o personas que actúen confiadas en el contenido de este documento, o en las opiniones vertidas por sus autores.



## PRESENTACIÓN

Los cuadernos de I+D+i de Canal de Isabel II forman parte de la estrategia de gestión del conocimiento de la empresa y del desarrollo del Plan de Investigación, Desarrollo e Innovación.

Son elemento de difusión de proyectos e iniciativas desarrollados y auspiciados desde Canal de Isabel II para la innovación en las áreas relacionadas con el servicio de agua en el entorno urbano.

Exponen las diferentes problemáticas abordadas en cada proyecto junto con los resultados obtenidos. La intención al difundirlos mediante estas publicaciones es compartir las experiencias y conocimientos adquiridos con todo el sector de servicios de agua, con la comunidad científica y con cuantos desarrollan labores de investigación e innovación. La publicación de estos cuadernos pretende contribuir a la mejora y eficiencia de la gestión del agua y, en consecuencia, a la calidad del servicio prestado a los ciudadanos.

Los cuadernos de I+D+i ya publicados son los que figuran en la presente tabla.

Nº colección	Cuadernos Investigación, Desarrollo e Innovación publicados
1	Transferencias de derechos de agua entre demandas urbanas y agrarias. El caso de la Comunidad de Madrid
2	Identificación de rachas y tendencias hidrometeorológicas en el ámbito del sistema de Canal de Isabel II
3	Participación de Canal de Isabel II en el Proyecto Internacional de Eficiencia en la Gestión (IDMF)
4	Microcomponentes y factores explicativos del consumo doméstico de agua en la Comunidad de Madrid
5	El agua virtual y la huella hidrológica en la Comunidad de Madrid
6	Estudio de potenciales de ahorro de agua en usos residenciales de interior
7	Investigación sobre potenciales de eficiencia con el empleo de lavavajillas
8	Precisión de la medida de los consumos individuales de agua en la Comunidad de Madrid
9	Proyecto de investigación para la definición y evaluación de la aplicabilidad de un bioensayo para la determinación de la toxicidad del agua utilizando embriones de Pez Cebra
10	Eficiencia en el uso del agua en jardinería en la Comunidad de Madrid
11	Técnicas de teledetección y sistemas de información geográfica para la evaluación de la demanda de agua para usos de exterior en la Comunidad de Madrid
12	Estudio sobre la dinámica de cianotoxinas en dos embalses de abastecimiento de Canal de Isabel II



## FICHA TÉCNICA

**Título del proyecto** Desarrollo de un sistema de validación, estimación y predicción de consumos horarios por sectores para la red de distribución de Canal de Isabel II.

**Línea de investigación** Aseguramiento del equilibrio disponibilidades /demandas.

**Unidades implicadas de Canal de Isabel II** Subdirección de I+D+i.

**Participación externa** Getinsa Ingeniería S.L.

**Objeto y justificación del proyecto**

En el momento presente, Canal de Isabel II está llevando a cabo la sectorización de su red de distribución. Gracias a esta medida se puede disponer de lecturas de consumo horario a una escala notablemente más desagregada. En este contexto, surge la necesidad de analizar estas nuevas series de datos, prestando especial atención a determinadas variables del comportamiento que hasta ahora se consideraban poco trascendentes o bien habían pasado desapercibidas, de cara a obtener una serie de patrones identificables y en buena medida interpretables, en función de los hábitos y actividades de la población.

En paralelo, este cambio de escenario comporta la necesidad de rediseñar las herramientas de gestión y análisis de Canal de Isabel II para adaptarlas a este nuevo enfoque.

**Contribución al estado del arte**

Se ha avanzado en el conocimiento de nuevos patrones de consumo que, merced a la sectorización de la red de distribución llevada a cabo, han podido identificarse. La mayor finura de la división realizada ha permitido agrupar unidades de consumo, no sólo según su ubicación geográfica, sino también de acuerdo a su tipología. Los resultados obtenidos del análisis de los datos de consumo disponibles pueden considerarse novedosos.

**Resumen del desarrollo del proyecto e hitos relevantes**

La investigación ha permitido desarrollar una formulación matemática y, a partir de ella, una herramienta de modelización de consumos horarios que ha resultado ser aplicable al amplio abanico de tipologías de unidades de demanda que se pueden identificar en el sistema abastecido por Canal de Isabel II dentro de la Comunidad de Madrid.

**Resumen resultados obtenidos**

Desarrollo de modelos de estimación y predicción de consumos de agua, a nivel horario, para más de 300 unidades de consumo resultantes de la sectorización de la red de distribución.

Actualización de la aplicación informática CHYPRE para adaptar sus funcionalidades a la gestión de los datos de consumo recogidos por los sensores de control hidráulico instalados en los distintos sectores.

**Líneas de investigación abiertas para la continuación de los trabajos**

Generación de conocimiento relativo a los patrones de consumo en función de las características tipológicas de los sectores de la red. El desarrollo de la formulación ha servido para facilitar el proceso de adquisición de conocimiento. Se considera que la fase de explotación y mantenimiento de los modelos de consumo desarrollados podrá servir para profundizar en esta tarea.



## RESUMEN EJECUTIVO



En el año 2004, Canal de Isabel II puso en marcha un plan para la compartimentación de su red de distribución en sectores independientes de extensión limitada. Esta sectorización consiste básicamente en subdividir la red en zonas de tamaño reducido, aisladas entre sí, y conectadas con las arterias principales mediante un número reducido de puntos de entrada y salida. Equipando cada una de estas vías de entrada y salida con una serie de captadores capaces de suministrar información precisa y a cortos intervalos relacionada con el control hidráulico de los flujos, se está en condiciones de poder establecer un diagnóstico y control del funcionamiento de la red a pequeña escala.

Merced a esta actuación, en el sistema de Telecontrol se registran de manera continua datos de caudal entrante a escala horaria correspondientes a cada uno de los sectores implementados. A medida que se ha ido acumulando información histórica de estos registros, han quedado evidenciados los ciclos diarios, semanales y anuales a los que responde buena parte de la variabilidad temporal de los consumos. En la mayoría de los casos, estos comportamientos se corresponden claramente con lo esperable en función de las actividades, hábitos y desplazamientos de la población. Por otro lado, parte de la oscilación no explicada por las variables anteriores puede ser relacionada con factores meteorológicos, cuya acción puede ser también susceptible de interpretarse en función de mecanismos comprensibles.

Naturalmente, en función de cuál sea la tipología de demanda o la unidad doméstica predominante en un sector determinado, el patrón representativo del consumo histórico registrado puede variar de manera ostensible. Sin embargo, estos comportamientos pueden entenderse como distintas respuestas a la evolución de una misma serie de variables, en general, fácilmente anticipables.

A partir de estos antecedentes, la investigación realizada se ha basado en intentar desarrollar una formulación matemática, aplicable a todos los sectores implementados, y capaz de establecer una relación entre los datos de caudal obtenidos y una serie de predictores vinculados a unas variables sobre cuya base se pueden entender las fluctuaciones en el consumo de agua registradas en las series históricas.

Basándose en la formulación anterior, como segundo alcance de la presente investigación se ha desarrollado una metodología implementada en una aplicación informática que permite el ajuste de los parámetros que definen los modelos de consumos correspondientes a las unidades sectorizadas.

Estos modelos de demanda horaria constituyen la base del funcionamiento de un sistema de validación, estimación y predicción de consumos horarios, individualizado para cada sector de consumo, con que Canal de Isabel II cuenta en la actualidad, resultado final, tangible y práctico del trabajo de investigación llevado a cabo.

Para el establecimiento de dicho sistema, los modelos de consumo han sido desarrollados de acuerdo con unas características que se pudieran adecuar a las funcionalidades requeridas. A grandes rasgos, éstas se pueden resumir en los siguientes aspectos:

- Resolución temporal: horaria.
- Resolución espacial: cada uno de los sectores de consumo en los que quedará dividida la red de distribución.
- Capacidad predictiva.
- Posibilidad de actualización.

A continuación se resumen, a grandes rasgos, los pasos seguidos en la investigación realizada que ha permitido llegar al establecimiento de los modelos de consumo y a su aplicación práctica en el proceso de estimación, validación y predicción de caudales.

## Caracterización del comportamiento de los consumos a partir del análisis de datos históricos

El conocimiento previo del comportamiento de los consumos y, en particular, su interpretación en términos de los mecanismos o causas que lo producen ha constituido una actividad primordial. Como parte de este aprendizaje, se llevó a cabo un análisis de las series históricas de consumos individualizados por sectores. Como conclusión a este trabajo se determinó que, en términos generales, la evolución de los patrones de consumo podía interpretarse como respuesta a los siguientes factores:

- Variables de calendario (efecto de los días festivos, tanto en lo que se refiere a diferencias en la actividad urbana como a desplazamientos poblacionales asociados a los días de descanso).
- Estacionalidad del consumo (variaciones en los niveles de consumo que se producen entre unas épocas del año y otras y que se pueden atribuir tanto a los movimientos vacacionales, como a la estacionalidad de determinados usos del agua).
- Efecto de la oscilación meteorológica, en particular, precipitación y desviación de la temperatura respecto a los valores esperables en una determinada época del año.
- Fenómenos de tendencia de lenta evolución que cabe atribuir a variaciones a largo plazo en la población abastecida, en los usos de agua, hábitos de consumo etc.

## Diseño de una formulación adecuada para los modelos

El carácter periódico de buena parte de la variabilidad temporal del consumo de agua y la clara relación de dependencia que se ha encontrado con las anteriores variables, ha permitido desarrollar una formulación matemática de carácter lineal que, apoyándose en una serie de predictores asociados a las mismas permite, tanto reproducir adecuadamente los patrones de consumo registrados en las series históricas como la posibilidad de obtener, de manera sencilla, los parámetros de dicha formulación.

Así, se ha elaborado un modelo matemático de consumo de agua para cada sector de consumo existente. La finalidad última de esta formulación es poder disponer de un sistema de estimación y validación de consumos horarios individualizados para cada unidad sectorizada, de manera que puedan someterse a control los datos registrados en los sensores instalados en los puntos de entrada y salida de caudal.

Además, estos modelos disponen de capacidad predictiva, por cuanto las relaciones matemáticas que los constituyen se han establecido de modo que simulen comportamientos observados en las series que responden a causas, con muchas perspectivas de permanecer en el tiempo, a medio o largo plazo.

En la metodología desarrollada en este trabajo, se siguió el criterio de que la variabilidad intradiaria del consumo podía ser tratada de forma independiente de la variabilidad entre días. Este criterio condujo al desarrollo en paralelo de modelos en dos niveles de resolución temporal diferentes:

- Modelo horario, que contempla la forma en que el consumo diario se distribuye entre las horas del día. En este caso, la magnitud que se ha utilizado para establecer los modelos horarios ha sido el consumo horario referido al consumo medio horario del día.
- Modelo diario, que simula la evolución de los consumos diarios en escalas de tiempo iguales, o superiores a un día.

A continuación se resumen las principales características que presenta la formulación desarrollada en relación a las variables explicativas que incorpora:

- En la mayor parte de los sectores analizados, se evidencia claramente un ciclo semanal en el patrón de consumos, existiendo una notable diferencia en la magnitud y patrón de consumo de días festivos y

laborables. En la modelización, esta variabilidad ha quedado reflejada tanto a nivel de descenso de actividad, como de movilidad de la población.

- La estacionalidad de consumo ha quedado recogida a través de un factor multiplicativo en la modelización del consumo diario. Además, se ha tenido también en cuenta el componente estacional de la evolución del patrón intradiario, incorporando en el modelo horario un predictor afectado por este factor de estacionalidad.
- Para poder representar en la modelización posibles efectos derivados de fenómenos evolutivos a largo plazo, así como para utilizar las diferencias registradas entre las observaciones y las estimaciones del pasado inmediato como variable de entrada, se ha incorporado en el modelo diario un factor multiplicativo de corrección de tendencia que funciona adecuadamente en una escala temporal de trabajo de entre una y dos semanas.
- La influencia de la meteorología en el consumo es significativa, si bien comparativamente menor que la de los efectos de calendario. Su efecto se ha incorporado en la modelización del consumo diario a través de dos variables relacionadas, una con la desviación entre la temperatura observada y la esperable para la época del año, y la otra con el percentil que, en la serie histórica, ocupa la precipitación registrada. A su vez, estos predictores están afectados por un factor multiplicador que incorpora la estacionalidad atribuible al componente meteorológico, dado que se ha evidenciado que, en los meses de verano y de transición entre estaciones, el efecto meteorológico tiene una mayor repercusión en el consumo que durante el periodo invernal.

## Ajuste e implementación de los modelos

El proceso de modelización desarrollado consiste, básicamente, en la obtención de las constantes de cada uno de los modelos de regresión lineal contemplados en la formulación, y en determinar unas curvas de estacionalidad capaces de representar las variaciones de consumo observadas entre unas épocas del año y otras.

Para poder realizar estas tareas, como paso previo resulta necesario seleccionar los registros de consumo observado sobre los cuales realizar el ajuste. A partir del análisis individualizado de las series históricas de cada sector, se escogieron aquellos datos que podían considerarse representativos del patrón de consumos de la serie. Una vez seleccionado este conjunto de referencia, por aplicación del criterio de mínimos cuadrados, se establecieron las constantes de modelización mediante las cuales se lograba el mayor grado de ajuste entre estos consumos observados y las estimaciones obtenidas. Como es evidente, la bondad del ajuste depende, en buena medida, de la calidad de la serie de datos de consumo disponible en el momento de realizar la modelización del sector.

Los errores cometidos en la estimación pueden ser atribuibles tanto a limitaciones de la formulación del modelo que hagan imposible el dar respuesta a ciertos componentes de la variabilidad de los consumos, como a la influencia de fenómenos que pudieran considerarse como aleatorios o a errores cometidos en la toma de lecturas. En todo caso, dado el buen resultado obtenido en las modelizaciones efectuadas, puede concluirse que estas desviaciones tienen una consideración menor, por lo que pueden considerarse como ruido.

Estas diferencias entre los caudales medidos y los valores proporcionados por el modelo sirvieron para establecer los parámetros de validación, límite razonable para poder diagnosticar si un consumo entrante es lo suficientemente discordante respecto al valor esperado como para darlo por no representativo de la realidad que pretende medir.

La contribución de la investigación realizada a esta faceta de los trabajos, ha pasado por estimar una manera coherente para calcular los criterios y parámetros de validación en base a los que se determina el límite a partir del cual se produce el rechazo de un consumo entrante.

Para la evaluación de la precisión de cada modelo, se estableció como base de comparación el resultado obtenido en la validación realizada sobre el conjunto de referencia. Se seleccionó el coeficiente de regresión ( $r^2$ ) como representativo del nivel de ajuste y el error cuadrático medio como medida del error del modelo. En términos generales, el grado de ajuste de los modelos es satisfactorio. En conjuntos de referencia de suficiente

longitud y calidad, la regresión toma valores próximos a 0,9 mientras que el error cuadrático se mantiene en el entorno de un 10 por ciento del consumo medio diario medido, resultados que avalan claramente la validez de la formulación adoptada.

Fuera de la fase de modelización, valorar el nivel de ajuste de un modelo durante su explotación es, en general, complicado y no siempre representativo, ya que los errores obtenidos están muy condicionados por la evolución de los datos de entrada. En la medida en que estos se alejen del patrón observado en el conjunto de referencia, serán susceptibles de verse rechazados aun cuando no presenten anomalías significativas.

## **Empleo de los modelos desarrollados en los procesos de validación, estimación y predicción de consumos**

Como resultado primero del trabajo realizado, se dispuso de unos modelos de tipo autorregresivo, basados en una formulación lineal, tanto a escala diaria como en el caso de cada uno de los componentes del consumo horario.

Como ya se ha indicado, estos modelos han sido desarrollados con vistas a servir de base en los procesos de validación, estimación y predicción de datos de consumo por sectores. La estimación permite proporcionar un valor de consumo independientemente del dato medido, utilizando la herramienta de modelización, dicha herramienta, como ya se ha indicado, se apoya en un conjunto de predictores estacionales, de calendario y meteorológicos y en los consumos registrados (o estimados) de ese mismo sector en fechas precedentes.

La validación se utiliza en la fase de explotación de los modelos para diagnosticar la validez de los datos de consumo obtenidos de los sistemas de telecontrol instalados. Se entiende que una lectura es válida si responde con cierta precisión a la realidad que pretende medir. Los errores en el sistema de toma de datos o las anomalías en la realidad medida producen datos no válidos, es decir, no representativos y que, por tanto, deben ser descartados de la serie histórica. El rechazo de un dato genera una laguna de validación que se rellena con el dato estimado por el modelo para esa misma hora y día.

Por su formulación, relativamente sencilla, parsimoniosa en cuanto a su dependencia de datos de consumo previo y con un número limitado de parámetros, los modelos tienen, en general, un comportamiento estable (no producen valores fuera de rango u oscilaciones bruscas en la serie simulada). Gracias a estas características, es posible atribuirles capacidad predictiva.

La implementación de estos modelos de consumo individualizados en los más de 300 sectores analizados hasta la fecha de redacción del presente documento ha servido para avalar tanto la formulación obtenida del trabajo de investigación desarrollado, como los procesos seguidos para el cálculo de los parámetros de modelización a partir del análisis de series históricas.

Como es sabido, la población abastecida por Canal de Isabel II comprende la práctica totalidad de la Comunidad de Madrid. Como consecuencia, la heterogeneidad de las unidades de consumo abastecidas es muy notable, circunstancia que se traduce, de manera directa, en que las tipologías de sectores analizados hayan sido muy diversas. Si bien predominan las unidades de consumo de tipo doméstico, se distinguen igualmente sectores localizados en zonas de tejido industrial. Existe también una amalgama de unidades de consumo, todas ellas particulares en sí mismas, como los campos de golf, campings o centros de ocio.

A otra escala de desagregación, incluso dentro de los consumos asociados a áreas residenciales, pueden distinguirse numerosas categorías atendiendo a la representatividad del sector respecto a las variables explicativas del consumo: tipología de la vivienda (unifamiliar o plurifamiliar), uso al que se destina (primera o segunda residencia), nivel de renta de la población residente, etc.

Como ya se ha apuntado, las variables en base a las cuales se ha establecido la relación de dependencia con la evolución de los consumos son las mismas en todos los sectores analizados (tendencia, estacionalidad, calendario y meteorología). La versatilidad de la formulación adoptada se basa en que, independientemente de cuáles sean las características del sector considerado, en todos los casos se ha encontrado buena correlación entre la evolución de los consumos observados y la variación de estas variables explicativas (si bien, evidentemente, la respuesta que se obtiene puede variar en gran medida según cuál sea la tipología del sector en cuestión).

Como conclusión final, cabe destacar que el desarrollo de modelos teóricos de consumo para cada una de las unidades sectorizadas, ha servido tanto para la generación de conocimiento relativo a los patrones de consumo asociados a las distintas tipologías de demanda que se han identificado, como para profundizar en el comportamiento, a pequeña escala, de la red de distribución. La propia formulación del modelo desarrollado es capaz de contener el conocimiento generado sobre patrones de consumo y las interpretaciones que permiten explicarlos.

Por otra parte, merced a la implementación de los modelos, los contrastes entre las estimaciones de consumo elaboradas por los modelos teóricos y los caudales reales registrados, sirven para mejorar el control operativo de la red de distribución, permitiendo identificar variaciones en la pauta de consumos, así como incidencias de cualquier tipo (aparición de fugas, roturas, maniobras, anomalías en el funcionamiento de la instrumentación, etcétera).



# ÍNDICE DE CONTENIDOS



## **1. INTRODUCCIÓN página 19**

- 1.1. Planteamiento del problema
- 1.2. Plan del documento

## **2. FORMULACIÓN DE LOS MODELOS DE ESTIMACIÓN DE CONSUMO página 25**

- 2.1. Introducción
- 2.2. Análisis de las series de datos de consumo
- 2.3. Modelo diario
- 2.4. Modelo horario

## **3. TAREAS DE MODELIZACIÓN Y MANTENIMIENTO página 57**

- 3.1. Introducción
- 3.2. Desarrollo del modelo
- 3.3. Validación y estimación
- 3.4. Predicción

#### **4. COMPORTAMIENTO OBSERVADO EN EL PERIODO DE MANTENIMIENTO página 75**

- 4.1. Modificación de los parámetros de validación
- 4.2. Corrección del factor de tendencia
- 4.3. Recalibración completa del sector

#### **5. IDENTIFICACIÓN DE PATRONES DE CONSUMO página 93**

- 5.1. Introducción
- 5.2. Sectores pertenecientes a grandes núcleos urbanos del área metropolitana de Madrid
- 5.3. Sectores pertenecientes a áreas residenciales situadas al noroeste del núcleo urbano de Madrid
- 5.4. Sectores pertenecientes a municipios localizados en la Sierra Norte de Madrid
- 5.5. Consumos en sectores donde predomina la actividad industrial
- 5.6. Asignación de curvas de estacionalidad tipo

#### **6. INFLUENCIA DE LA METEOROLOGÍA EN LOS CONSUMOS POR SECTORES página 113**

- 6.1. Introducción
- 6.2. Periodo de referencia seleccionado
- 6.3. Respuesta individualizada por sectores al efecto meteorológico

#### **7. REFERENCIAS A OTROS ALGORITMOS DE MODELIZACIÓN-VALIDACIÓN DE DATOS DE CONSUMO página 129**

- 7.1. Introducción
- 7.2. Análisis de series temporales
- 7.3. Aproximación a la modelización mediante técnicas de reconocimiento de patrones
- 7.4. Metodologías basadas en redes neuronales artificiales

#### **8. CONCLUSIONES página 137**

- 8.1. Formulación de los modelos
- 8.2. Aplicación de los modelos

#### **9. ANEXOS página 143**

- Anexo 1. Glosario de términos
- Anexo 2. Referencias
- Anexo 3. Índice de figuras



1

Introducción

## 1.1. Planteamiento del problema

Canal de Isabel II dispone de herramientas y procedimientos sofisticados para la ayuda en la gestión de su red de distribución. Dichos sistemas se ven sometidos a un continuo proceso de actualización y mejora, que puede venir condicionado por diversas circunstancias.

Un claro ejemplo de esta necesidad de puesta al día, ha surgido de la decisión por parte de Canal de Isabel II de realizar una sectorización de su red de distribución. Derivado de esta actuación, se plantea la necesidad de conocer el funcionamiento de la misma a una escala diferente, lo que obliga a prestar especial atención a unas variables del comportamiento hidráulico de la red que antes no habían podido ser tenidas en consideración, y a rediseñar algunas de sus herramientas de gestión y análisis para adaptarse a este nuevo enfoque.

La sectorización consiste básicamente en subdividir la red de distribución en zonas de tamaño reducido, aisladas entre sí, que se encuentran conectadas con las arterias principales mediante un número reducido de puntos de entrada y salida a la red principal. Equipando cada una de estas vías de entrada y salida con una serie de captadores de campo que puedan suministrar información precisa y a cortos intervalos relacionada con consumos (caudales de entrada y salida), variaciones de presión, o con cualquier otra variable que se desee estudiar, se estará en condiciones de poder establecer un diagnóstico y control del funcionamiento de la red, a pequeña escala.

Sin embargo, la sectorización no debe entenderse sólo en términos de configuración de la red, sino que ha de concebirse como una nueva forma de gestión de la misma. Su principal ventaja radica en el hecho de que permite alcanzar un mejor conocimiento del patrón de consumos, lo que necesariamente se ha de traducir en un respaldo a la capacidad del gestor de la explotación para ofrecer un mejor servicio.

Con el conocimiento acumulado en los años de gestión, los patrones temporales de consumo a una escala de demanda agregada son sobradamente conocidos por los técnicos de Canal de Isabel II. Así, se sabe que buena parte de la variabilidad temporal del consumo de agua responde a ciclos diarios, semanales y anuales claramente identificables. Igualmente, se ha constatado la existencia un conjunto de variaciones que se pueden asociar a la distribución de los días festivos en el calendario.

En la mayoría de los casos, los comportamientos observados en las series de consumo de agua se corresponden claramente con lo esperable en función de las actividades, hábitos y desplazamientos de la población. Los mínimos nocturnos, los descensos de fin de semana -que se tornan aumentos en las poblaciones donde predomina la segunda residencia- los máximos estacionales de junio y julio, que van seguidos de descensos acusados en agosto -salvo en zonas de segunda residencia- son fácilmente interpretables. También, resulta significativo que el Jueves Santo suela ser el día de menor consumo de agua a escala anual, o que el día primero de enero, Año Nuevo, presente un patrón de consumo claramente distinto del resto de días del año.

Por otro lado, parte de la variabilidad no explicada por las pautas de variación anteriores puede ser relacionada con factores meteorológicos, cuya acción puede ser también susceptible de interpretarse en función de mecanismos comprensibles.

Aprovechando este conocimiento acumulado por la experiencia, en el año 2002, a partir del análisis de las series de consumo disponibles, y del desarrollo de formulaciones adecuadas para la modelización del comportamiento observado, Canal de Isabel II puso en marcha el desarrollo de un sistema de validación, estimación y predicción de consumos. El ámbito de aplicación de esta herramienta fue el de grandes zonas de demanda, dividiéndose para ello la red de abastecimiento en 48 unidades disjuntas entre sí. Los principales objetivos alcanzados en este estudio fueron los siguientes:

- Obtención de modelos teóricos de estimación y predicción de consumos horarios para cada una de las zonas en las que quedó dividida la red de distribución de Canal de Isabel II.
- Desarrollo de un conjunto de técnicas para explotar los citados modelos, de manera que fuese posible realizar funciones de validación, estimación y predicción de valores de consumo a escala horaria.
- Puesta en funcionamiento de una aplicación informática llamada CHYPRE (Consumos Horarios: validación, estimación Y PREDicción) que, utilizando los modelos citados en el párrafo anterior, facilitase la gestión de los datos de consumo, tanto históricos como los que se generasen de forma continuada, utilizando los modelos descritos para las funciones citadas en el punto anterior.

Como consecuencia de la sectorización de la red de abastecimiento, ha surgido la necesidad de llevar a cabo el desarrollo de otros modelos teóricos capaces de recoger las particularidades de cada uno de los sectores que constituyen las nuevas unidades de consumo. Este objetivo ha exigido buscar una nueva formulación matemática que permita reproducir adecuadamente dichos comportamientos y elaborar un método consistente para establecer los valores de los parámetros de dicha formulación. Por medio de un ajuste realizado a partir de los datos de consumo de la serie histórica, se ha podido disponer de un modelo matemático de consumo de agua en cada sector. Además, tales modelos presentan capacidad predictiva en la medida en que las relaciones matemáticas que los constituyen simulan comportamientos observados en la serie, que responden a causas interpretables y que se espera permanezcan en el tiempo a medio o largo plazo.

Para la modelización de sectores, además de la necesidad de modificar la formulación en la que se basan los modelos de estimación, ha sido necesario adaptar la herramienta informática CHYPRE a las condiciones particulares que caracterizan los patrones de consumo de cada una de estas unidades. A grandes rasgos, el grueso de los trabajos realizados puede dividirse en cuatro actividades principales:

- Desarrollo de la herramienta de modelización.
- Actualización de la aplicación informática.
- Modelización de sectores.
- Revisión y mantenimiento de los sectores modelizados.

El desarrollo de modelos teóricos de consumo para cada una de las unidades sectorizadas ha servido, tanto para la generación de conocimiento relativo a los patrones de consumo asociados a las distintas tipologías de demanda que se han identificado, como para profundizar en el comportamiento a pequeña escala de la red de distribución. Por otra parte, los contrastes entre las estimaciones de consumo elaboradas por los modelos teóricos y los caudales reales registrados, permiten identificar variaciones en la pauta de consumos, así como incidencias de cualquier tipo (aparición de fugas, roturas, maniobras, anomalías en el funcionamiento de la instrumentación, etc.).

## 1.2. Plan del documento

La investigación realizada se expone en los capítulos 2 a 8, según el esquema de contenido que se indica a continuación:

### Formulación de los modelos de estimación de consumo

Se describen cuales son las principales variables para las que se ha identificado una respuesta a nivel global en el patrón de consumo del conjunto de sectores contemplados. Basado en estas relaciones, se describe la formulación matemática de los modelos teóricos de consumo adoptada y la selección de los predictores sobre los que se apoya la anterior formulación para la estimación de consumos, tanto a escala horaria como diaria.

### Tareas de modelización y mantenimiento

Se expone cual es el proceso seguido para determinar las constantes de modelización del término lineal y los factores de estacionalidad que definen el modelo de consumo de un sector. Igualmente, se destaca la importancia de la determinación del conjunto de referencia de datos históricos a partir del cual se realizará el ajuste por mínimos cuadrados para fijar los parámetros del modelo.

### Comportamiento observado en el periodo de mantenimiento

En este capítulo se exponen cuales son las tareas habituales que comprende el mantenimiento de un modelo de consumo. En particular, se describen las funcionalidades principales desarrolladas para poder corregir desviaciones del modelo y de su capacidad de estimación y validación de consumos.

### Identificación de patrones de consumo

Se describen aquellos patrones de consumo, tanto a escala diaria como horaria, que son claramente representativos de una determinada tipología de sector.

### Influencia de la meteorología en los consumos por sectores

Se detalla en qué medida una desviación meteorológica acusada puede incidir en la distribución del patrón de consumo intradiario, en sectores de distintas tipologías.

### Referencia a otros algoritmos de modelización y validación de datos de consumo

De manera sucinta, se presenta una revisión del estado del arte en lo que a técnicas de estimación y validación de consumos urbanos a escala horaria se refiere.

### Conclusiones

Finalmente, se exponen las conclusiones que se extraen del análisis realizado.

Además, el documento incluye un glosario de términos utilizados en el desarrollo de las exposiciones del presente trabajo.





# 2

Formulación de los modelos de  
estimación de consumo

## 2.1. Introducción

El análisis realizado sobre las series de datos de consumo individualizadas para cada sector ha revelado las pautas de variación en el tiempo de dichos consumos. En general, se ha comprobado que buena parte de la variabilidad temporal del consumo de agua en un determinado sector presenta un marcado carácter periódico, con ciclos diarios, semanales y anuales principalmente. Asimismo, se constata la existencia de tendencias a largo plazo de diversa naturaleza. Hay también un conjunto de variaciones que se pueden asociar claramente a la distribución de los días festivos en el calendario.

En la mayoría de los casos, los comportamientos enumerados en el párrafo anterior son interpretables en términos de actividades de los usuarios.

Los mínimos nocturnos en el consumo se corresponden con el periodo de descanso de la población. Es, en estas horas de la noche, donde más fácilmente pueden evidenciarse aquellos caudales que puedan estar relacionados con fugas existentes en la red de distribución.

Por su parte, los descensos de fin de semana se pueden atribuir, en términos generales, al descanso en muchas actividades laborales y a los desplazamientos de fin de semana.

La causa de los máximos estacionales de junio y julio, en determinadas familias de sectores, se puede buscar en la estacionalidad de ciertos usos del agua por su dependencia de las condiciones climáticas: riego de jardines, uso de piscinas, etc. El bajón de agosto en el municipio de Madrid y núcleos periféricos, y su contrapunto de ascenso en otros municipios donde abundan las segundas residencias, se relaciona fácilmente con los desplazamientos vacacionales de la población.

Finalmente, parte de la variabilidad no comprendida en las pautas de variación anteriores puede ser debida a factores meteorológicos. La acción de estos factores es también susceptible de interpretarse en función de mecanismos comprensibles: disminución de riegos o baldeo de calles después de precipitaciones, aumento de ciertos consumos como riegos en días más calurosos, etc.

Se trata, entonces, de buscar una formulación matemática que permita reproducir adecuadamente dichos comportamientos; de un método consistente, para establecer los valores de los parámetros de dicha formulación. Los requisitos exigidos a los modelos de estimación de consumos para poder adecuarse a las funcionalidades requeridas fueron los siguientes:

- Resolución temporal: horaria.
- Resolución espacial: sector de consumo obtenido de la compartimentación de la red.
- Capacidad predictiva.
- Posibilidad de actualización.

Se planteó, pues, desarrollar un modelo matemático de consumo de agua que pudiera aplicarse a cada uno de los sectores creados. Estos modelos llevan atribuida una capacidad predictiva por cuanto las relaciones matemáticas que los constituyen se han establecido de modo que simulen comportamientos observados en la serie, que responden a causas que es muy verosímil su permanencia en el tiempo a medio o largo plazo.

En la metodología propuesta en este trabajo, se ha seguido el criterio de que la variabilidad intradiaria del consumo se puede tratar de forma independiente de la variabilidad entre días. Este criterio ha conducido al desarrollo en paralelo de modelos en dos niveles de resolución temporal diferentes:

- Modelo horario, que contempla la forma en que el consumo diario se distribuye entre las horas del día.
- Modelo diario, que simula la evolución de los consumos diarios en escalas de tiempo iguales o superiores a un día.

Este enfoque ha proporcionado buenos resultados, con la ventaja de simplificar el problema, al dividirlo en dos cuestiones independientes de menor complejidad.

Al margen de esto, una dificultad añadida en el proceso de modelización llevado a cabo ha venido impuesta por los problemas tanto de cantidad, como de calidad, en algunas series de datos.

La sectorización física de la red de distribución comenzó en 2004 y, a fecha de redacción del presente cuaderno, aun no está completamente finalizada. A las lagunas en la toma de datos se unen las derivadas del proceso de depuración realizado. Algunas series, ya cortas de por sí, quedan, de este modo, bastante incompletas, especialmente en los últimos meses del periodo de estudio, que por ser los más recientes, son también los más valiosos. Además, el proceso de depuración, debido a que se ha llevado a cabo con criterios conservadores, no ha eliminado ciertas irregularidades que probablemente no reflejen un comportamiento real del consumo y sí pueden perjudicar el proceso de modelización.

En los apartados siguientes se detalla la metodología seguida para el desarrollo de los modelos a nivel diario y nivel horario respectivamente.

## 2.2. Análisis de las series de datos de consumo

### 2.2.1. Consideraciones generales

Como paso previo a la consecución de los modelos de demanda, se ha realizado un análisis minucioso de todas y cada una de las series de consumo, con las dos siguientes finalidades:

- Conocer la calidad y completitud de los datos disponibles.
- Caracterizar los aspectos más relevantes del comportamiento del consumo en cada una de las unidades de demanda.

En relación con la primera, la presencia de datos fuera de rango, de lagunas o de cambios bruscos en el nivel de las series, debidos a causas ajenas a las que intervienen en el modelo, puede enmascarar las relaciones reales entre variables y, consecuentemente, dificultar seriamente la tarea de modelización. De ahí la necesidad de analizar la calidad de los datos y, en su caso, de depurar o rellenar las series.

Sin embargo, la frontera entre los datos desestimables y los datos simplemente incómodos es siempre indefinida y, en cualquier tarea de depuración de series, se corre el riesgo de eliminar datos representativos de la realidad que se desea modelizar (es más, representativos de los extremos de dicha realidad) y perder así información valiosa. En este sentido, se apostó por ser conservador en toda decisión concerniente a desestimar datos y, al mismo tiempo, se renunció a cualquier intento de relleno previo de lagunas.

El criterio para la depuración de los datos ha sido subjetivo, basado en el estudio del comportamiento general de cada serie y la coherencia con otras series. El único criterio sistemático utilizado ha sido el de descartar los datos de caudal negativos. A todos los efectos del presente trabajo, los datos negativos han sido considerados como lagunas.

En relación con la segunda de las finalidades planteadas, debe destacarse que el conocimiento previo del comportamiento de los consumos y, en particular, su interpretación en términos de los mecanismos o causas que lo producen se considera fundamental. Dicho conocimiento, expresado en términos de una caracterización del comportamiento de la serie, servirá de base en el diseño y posterior desarrollo de los modelos de consumo.

Aunque, por comodidad en la exposición, esta tarea de análisis se describe como previa a la de desarrollo del modelo, debe entenderse que ambas están relacionadas. El desarrollo de los modelos requiere la realización previa del análisis, pero, a su vez, puede producir resultados que lleven a revisar dicho análisis.

### **2.2.2. Datos diarios**

La casuística de problemas detectados en el nivel diario es variada.

De forma concisa se pueden citar los cambios bruscos en el nivel de consumo, la presencia de picos con valores fuera de lo normal, de uno o varios días de duración, y la existencia de periodos de varios días con caudal constante. La mayoría de las veces, estos comportamientos anómalos no parecen atribuibles a la evolución real del consumo de agua sino, más bien, a cambios en la operación de la red, problemas de los instrumentos de medida de caudal que no han sido correctamente instalados e incluso averías o fugas.

Al margen de estas incidencias, el análisis de las series de datos de consumo diario ha permitido identificar un conjunto de patrones de comportamiento, en términos de los que es posible describir una parte importante de la variabilidad del consumo diario de cada sector. La tendencia a largo plazo, la estacionalidad, el efecto de los fines de semana, los días festivos y los "puentes", etc. se han descrito para cada una de las unidades de demanda consideradas.

### **2.2.3. Datos horarios**

El análisis exhaustivo de las series de datos horarios plantea el problema del elevado número de los mismos.

Para facilitar el proceso, se ha utilizado un método para detectar distribuciones anómalas del consumo horario. El procedimiento consiste, en primer lugar, en determinar la distribución media diaria de los consumos horarios relativos (caudal horario/caudal medio diario) de la serie de datos, del sector considerado. Seguidamente, se eleva para cada día la distancia cuadrática de sus consumos horarios relativos a la distribución media. Los días con una distribución de caudales muy diferente de la distribución media presentan un valor mucho más alto de esta distancia cuadrática y son fácilmente detectables.

Los días anómalos detectados, según el procedimiento descrito en el párrafo anterior, se analizan uno a uno para conocer la causa de su diferencia respecto de la media.

Entre la casuística habitual de las anomalías encontradas se pueden citar picos bruscos de caudal en una o algunas horas consecutivas; descensos bruscos del consumo que se prolongan varias horas, incluso días; datos de consumo constante durante varias horas seguidas en sectores donde existe una pauta clara de variación horaria del consumo, etc. Claramente, estos comportamientos anómalos parecen consecuencia de maniobras en la operación de la red, averías, fugas o problemas de los caudalímetros más que reflejo de las variaciones reales del consumo horario.

Como parte fundamental del análisis de datos y paso previo de las tareas de modelización, se han buscado en las lecturas históricas patrones de comportamiento del consumo que se pudieran poner en relación con:

- a) El conocimiento previo existente (efecto del fin de semana, estacionalidad, etc.).
- b) La tipología de los sectores (núcleos urbanos, urbanizaciones abiertas, polígonos industriales, etc.).

A grandes rasgos, el análisis de las series de datos horarios revela una tipología de patrones de distribución del consumo horario:

- Al trabajar a un nivel de agregación pequeño, con frecuencia se han identificado patrones de consumo distintos, correspondientes a tipos más "puros": polígono industrial, urbanización de viviendas unifamiliares, núcleo urbano, etc.
- En términos generales, en la mayoría de los sectores la distribución de consumo durante el día responde a un patrón claro con mínimos nocturnos y máximos matinales. Es frecuente la presencia de una punta moderada de consumo alrededor de las 22 horas. Se trata de la distribución de consumos horarios esperables en un área urbana.
- Con todo, se observa una variación en el patrón horario que responde, de manera generalizada, a la estación del año. Esta estacionalidad se manifiesta sobre todo en las horas finales del día. Esto hace que, según la época del año y el tipo de sector considerado, la punta de consumo se atrase o adelante respecto a la referencia de las 22 horas.
- En otros sectores se aprecia una pauta distinta, a base de tramos con caudal casi constante y otros con caudal nulo, generalmente en ciclos diarios. Este patrón de consumo es el obtenido en caudalímetros instalados en la toma de entrada desde la arteria principal de depósitos de regulación y carga utilizados para la alimentación de la unidad sectorizada.
- En algunos sectores, con frecuencia se aprecian cambios anómalos de comportamiento horario a partir de un determinado día u oscilaciones erráticas difíciles de analizar. Un caso particular de esta situación lo constituyen aquellas series de datos en las que aparece un salto brusco de magnitud de consumo en algún punto, que da lugar a la aparición de dos o más subseries coherentes, ambas por separado, pero no homogéneas entre sí. Esta respuesta en el histórico de consumos es típica en el caso de haber llevado a cabo modificaciones en la compartimentación de la red (subdivisión de un sector o incorporación al mismo de nuevas unidades de consumo).
- Por último, existen sectores de reducido tamaño y constituidos fundamentalmente por segundas residencias en los que el consumo en invierno es tan reducido, en comparación al observado en el periodo estival, como para que varíe el orden de magnitud. Esta circunstancia dificulta en gran medida conseguir una modelización válida para ambos periodos estacionales.

En el apartado dedicado a las labores de modelización y mantenimiento, se da cumplida cuenta de las tareas realizadas para poder constituir un conjunto de referencia de longitud suficiente sobre el que establecer los parámetros del modelo para cada uno de los sectores.

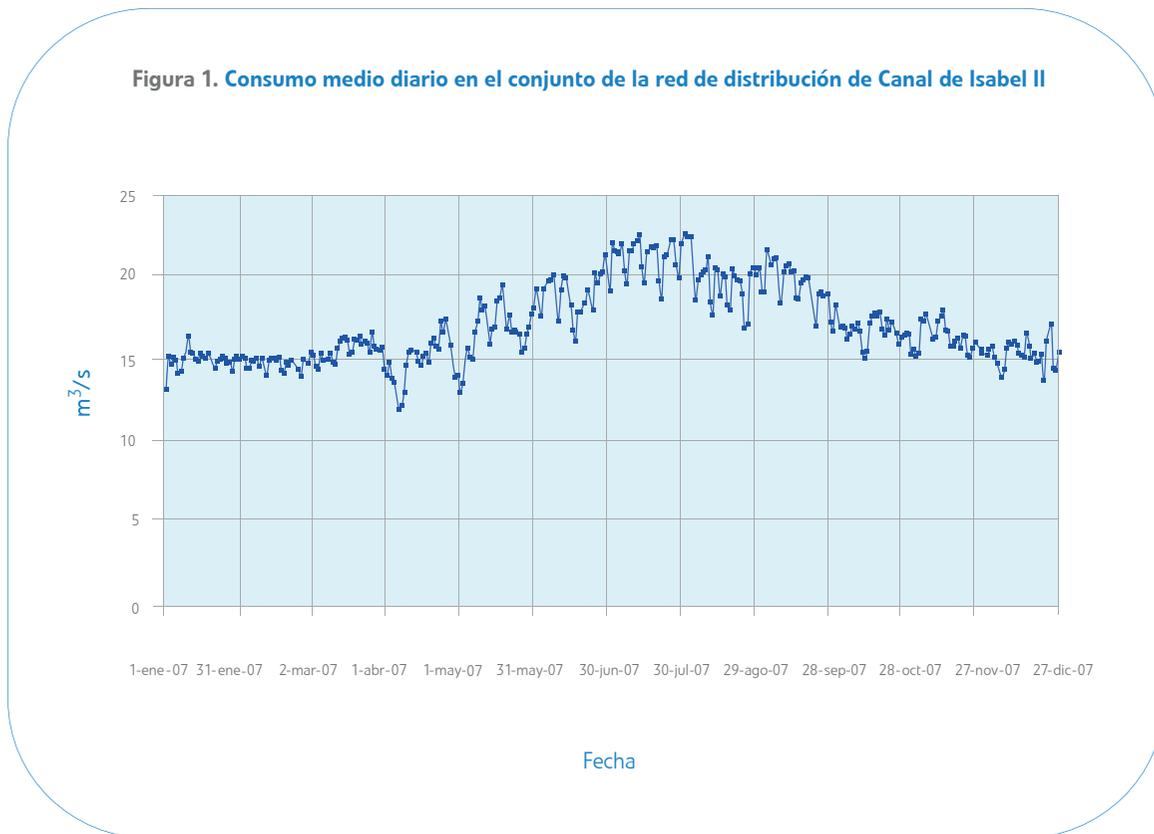
Se apunta en cualquier caso que, debido a que la sectorización se ha llevado a cabo en fechas muy recientes, con el fin de disponer de un número de datos suficientes, ha sido necesario adoptar criterios muy conservadores en la selección de datos de referencia, lo que ha desembocado en que, en la fase de explotación del modelo, se haya podido comprobar que algunas de las modelizaciones realizadas resultan excesivamente laxas. Por este y otros motivos, se destaca la importancia de las tareas de mantenimiento y revisión de los modelos, iniciadas en esta fase de los trabajos y que deberán prolongarse como parte intrínseca a la explotación.

## 2.3. Modelo diario

### 2.3.1. Componentes de la variación de consumos diarios

La figura 1 muestra la serie de consumos de agua diarios (expresado como caudal medio diario) del conjunto de la red de Canal de Isabel II, para el año 2007. En ella se aprecian claramente una serie de características de la evolución temporal de consumos diarios que se observan con carácter bastante general en la mayoría de los sectores estudiados:

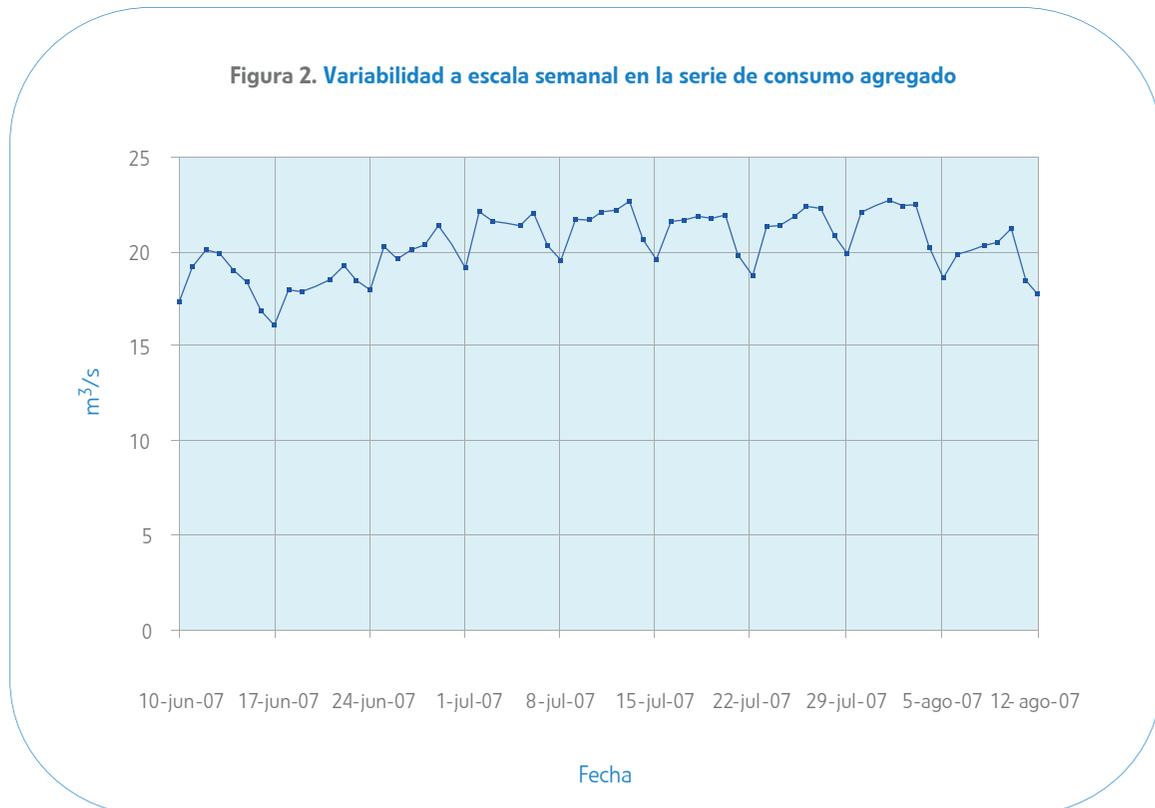
- Estacionalidad en el nivel general de consumo.
- Efecto del fin de semana.
- Estacionalidad en el efecto del fin de semana.
- Efecto de la Semana Santa y de los días festivos.
- Meteorología.



La estacionalidad del nivel general se traduce en un claro aumento del consumo de agua asociado al periodo estival: los valores máximos se dan durante los meses de junio y julio. Determinados usos estacionales del agua como el riego de jardines o el llenado y uso de piscinas y un mayor consumo doméstico inducido por el calor pueden explicar este patrón. Por otra parte, el claro descenso de consumo de agua en el mes de agosto (que se corresponde con un aumento en otros núcleos de población de uso residencial asociado a los meses estivales) puede relacionarse con los desplazamientos vacacionales de la población.

El efecto del fin de semana produce un ciclo semanal claro, con una caída del consumo en sábado y domingo, que se aprecia en la forma aserrada de la serie de consumos del gráfico de la figura 1.

En la figura 2 se muestra con mayor detalle una parte de la serie de consumos medios diarios. La escala horizontal se ha dividido en semanas y las fechas presentadas en el eje de abscisas corresponden a los domingos, de modo que el ciclo semanal puede apreciarse con absoluta claridad.

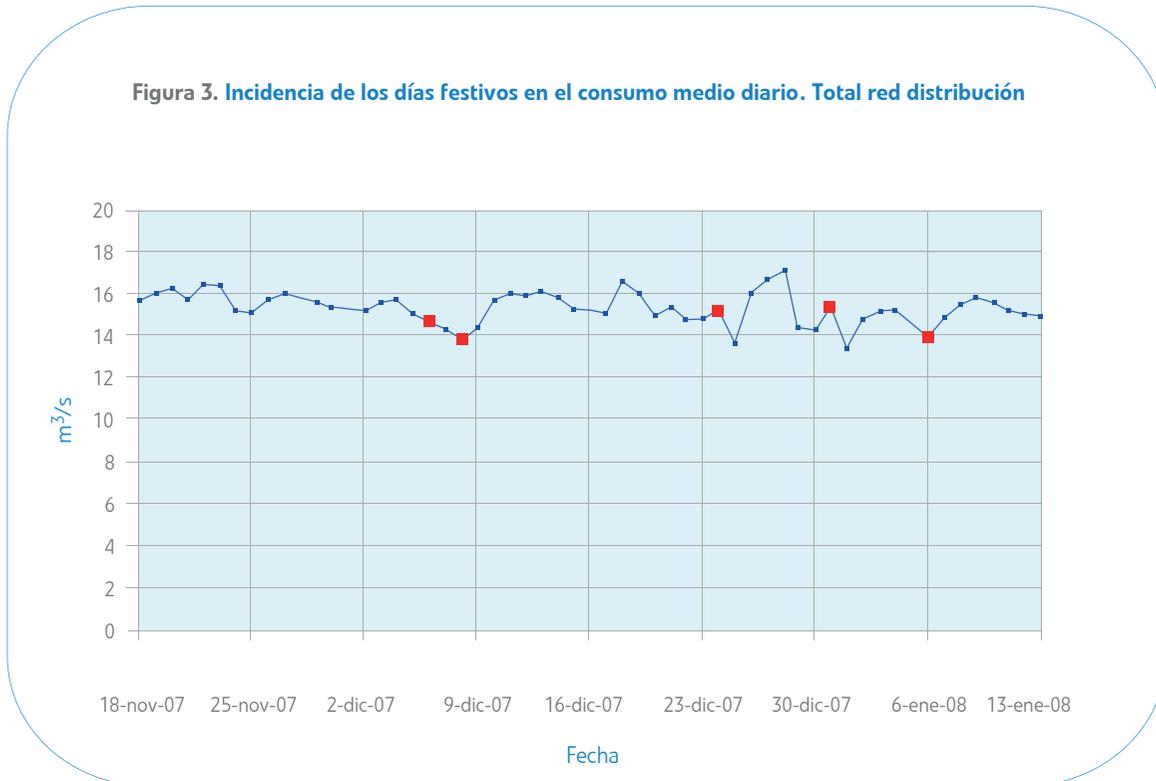


Conviene tener en cuenta que, en sectores donde predomina la segunda residencia (aquellos situados en la zona de la sierra, próximos al núcleo urbano especialmente), se observa el efecto contrario de fin de semana, es decir, máximos de consumo en sábado y domingo.

Por otra parte, el efecto de fin de semana también está sujeto a estacionalidad. En la figura 1 se comprueba que el descenso de consumo en fines de semana es muchísimo más marcado en los meses de verano. La explicación parece residir en la estacionalidad en el uso de la segunda vivienda y de las salidas de fin de semana, que se suponen más frecuentes en las épocas de buen tiempo.

Finalmente, el efecto de la Semana Santa y otros días festivos es claro. En 2007, la Semana Santa fue la comprendida entre el 2 y el 8 de abril. En la figura 1 se aprecia el claro descenso de consumo de esos días en relación con el nivel general de las semanas anteriores y posteriores. Es bastante frecuente que el Jueves Santo marque el mínimo consumo diario del año en el conjunto de consumos de Canal de Isabel II. En general, este descenso se observa todos los años y de manera generalizada en la mayoría de los sectores en los que ha quedado dividida la red de distribución (quedan fuera de este caso los sectores donde predominan las segundas residencias que, como ya se ha indicado, suelen observar un aumento notable del consumo en estas fechas).

El periodo navideño presenta un comportamiento similar. En el gráfico de la figura 3 se han señalado en rojo los días festivos de diciembre y enero de 2007 y 2008. Se comprueba su efecto general en la reducción de consumo, y cierta sinergia con los fines de semana próximos y días “de puente”.



De cara a la modelización, conviene resaltar que, para representar el patrón de consumo de fines de semana y festivos, conviene analizar al menos dos componentes. El primero se refiere a las diferencias en la actividad urbana entre días festivos y días de diario: las personas no acuden al trabajo, los comercios cierran, etc. El segundo se refiere a los movimientos de personas asociados a los días de descanso: viajes de fin de semana o puente, desplazamiento a una segunda residencia, etc.

Al margen de los factores ya reseñados y pese a que en la figura 3 no pueda percibirse, hay que mencionar que la evolución a largo plazo del consumo diario de algunos sectores cuya serie histórica tiene una longitud suficiente, presenta tendencia significativa a aumentar o a disminuir. La tendencia a largo plazo representa una variación parsimoniosa del nivel general de consumo que cabe atribuir a variaciones a largo plazo en la población abastecida, en los usos del agua, en los hábitos de consumo, etc.

Evidentemente, esta tendencia no se aprecia en una serie que no abarque más de un año, puesto que en ese nivel predominan otros efectos como la estacionalidad. En cualquier caso, sea o no perceptible esta característica con los datos disponibles, es esperable que se ponga de manifiesto con el paso del tiempo, de modo que su inclusión en el modelo es completamente necesaria.

En el presente trabajo se ha renunciado a la modelación conceptual del efecto de tendencia, ya que esta correspondería a un estudio sobre niveles de resolución temporal mayores. En principio, se asume que la tendencia actúa de forma gradual, de modo que su efecto para los próximos días se podrá estimar con precisión suficiente a partir del nivel de consumo observado en los días del pasado inmediato.

Por último, cabe reseñar la influencia de la meteorología, ya que ciertas condiciones meteorológicas tienen efectos potenciadores o inhibidores sobre ciertos usos del agua. El riego de jardines o el uso de piscinas son claros ejemplos.

Agregando todos los factores considerados en los párrafos anteriores dentro de la formulación, de forma esquemática, el modelo adoptado se articula de la siguiente manera:

$$c^* = e_c \times [b_0 + \text{TérminoLineal}] \times TEN$$

donde:

- $c^*$  es la estimación de consumo de agua diario en un sector.
- $e_c$  es un factor de estacionalidad de consumo, cuyo valor depende del día del año y del sector de consumo considerado.
- $b_0$  es un valor base constante de consumo diario.
- *TérminoLineal* es una función lineal de predictores de calendario y meteorológicos. Los predictores, variables meteorológicas o de festividad, pueden estar afectados, a su vez, por factores de estacionalidad (cuyo valor depende del día del año).
- **TEN** es el factor de tendencia, una corrección de la estimación de consumo para llevarla al nivel que marque la tendencia a largo plazo.

El término lineal puede expresarse de la siguiente manera:

$$\text{TermLineal} = \sum_{i=1}^n a_i \cdot f_i \cdot p_i$$

donde:

- $n$  es el número de variables que intervienen en el término lineal en cuestión.
- $a_i$  es el coeficiente del término  $i$ -ésimo.
- $f_i$  es la función de estacionalidad del término  $i$ -ésimo.
- $p_i$  es el valor del predictor correspondiente al término  $i$ -ésimo.

Un mismo predictor suele aparecer dos veces en el mismo término lineal, una de ellas afectado por una función de estacionalidad constante (es decir, sin estacionalidad).

La estrategia de construcción del modelo que se ha adoptado ha consistido en el estudio por etapas sucesivas de la estacionalidad, el calendario y los factores meteorológicos. Se ha considerado conveniente esta estrategia porque, tratándose de un problema tan complejo, es la que permite obtener de forma simultánea tanto resultados como conocimiento que los avale. Por otra parte, la viabilidad de la estrategia y el diseño de la misma resultan de los siguientes hechos y consideraciones:

- a) Como ya se ha indicado, no se aborda la modelización de la tendencia a largo plazo a partir de sus causas. Alternativamente, es posible eliminar su efecto en los datos originales, construir los modelos y, por último, aplicarles una corrección de tendencia basada en sus desviaciones respecto de observaciones recientes.
- b) Se ha encontrado un método que permite aislar razonablemente la estacionalidad del resto de factores que determinan el consumo. A partir del mismo, se obtiene para cada sector una representación de la estacionalidad en forma de factor multiplicativo que resulta satisfactoria.
- c) Una vez eliminada la estacionalidad y la tendencia, son los factores de calendario los que, a priori, parecen ser más determinantes. En principio, ha resultado posible determinar su influencia prescindiendo de las variables meteorológicas.
- d) El efecto de la meteorología y su interacción con otras variables se puede analizar, finalmente, sobre los residuos de un modelo que incluya los factores tratados con anterioridad.

El análisis de cada uno de los factores considerados en la simulación de consumos diarios se describe en detalle en los apartados que siguen.

### 2.3.2. Análisis de la estacionalidad

En la formulación a nivel diario, se han considerado dos niveles de estacionalidad para el desarrollo de los modelos:

- a) Estacionalidad general o estacionalidad del consumo.
- b) Estacionalidad de los efectos de calendario, o estacionalidad de la festividad.

La estacionalidad general aparece representada de forma explícita en el esquema del modelo descrito en el apartado anterior: es el factor de estacionalidad del consumo,  $e_c$ . La representación de la estacionalidad de los efectos de calendario se ha realizado también mediante un factor,  $e_f$ , que interviene en la formulación de las variables de calendario consideradas en el esquema del modelo propuesto. Ambas estacionalidades, por tanto, se tratan como factores multiplicativos.

Los valores de  $e_c$  nos permitirán, dividiendo los consumos observados por este factor, "desestacionalizar" las series de consumo. Los valores de  $e_f$  se utilizarán para modelar la interacción entre los factores de calendario y la época del año, en los casos en los que aquellos se vean influenciados por ésta.

Para el estudio de la estacionalidad se ha trabajado con datos semanales, incluidos en el conjunto de referencia y correspondientes únicamente a las semanas sin días festivos distintos del domingo. Se han escogido las semanas sin festivos porque un día festivo suele ser de consumo diferente que los restantes, lo que podría conducir a una aproximación falseada del componente de estacionalidad.

Tales semanas son unidades de tiempo prácticamente idénticas en cuanto a efectos de calendario, por lo que bastará con eliminar la tendencia a largo plazo para aislar el efecto de la estacionalidad en los consumos. Estamos suponiendo que los efectos meteorológicos y el conjunto de efectos no considerados tienen una relevancia menor en relación con la estacionalidad y que, en el nivel semanal, se pueden considerar ruido.

En el momento en el que se llevaron a cabo las tareas de modelización de sectores, las series de consumos diarios contaban con apenas cuatro años de datos, en los mejores casos. Para este análisis se ha exigido un mínimo de dos años, de cara a asegurar una representación aceptable de las variaciones estacionales.

Cada dato de consumo semanal se asocia a una posición a escala anual que se caracteriza por el día del año en que se encuentra el jueves de la semana. De esta manera, se calculan los valores de consumo acumulado de cada semana sin festivos y se sitúan el jueves de la semana correspondiente. Las posiciones en el año, por tanto, toman valores entre 1 y 366. Por ejemplo, para la semana del 1 al 7 de enero de 2001, la posición en el año es 4.

Esta ubicación en el año de los datos semanales sirve de base para establecer los valores de  $e_c$  y  $e_f$ , en función del día del año, es decir, establecer las funciones de estacionalidad de consumo y de calendario.

A la vista de esta serie de datos, se puede definir el intervalo considerado como más adecuado para establecer la función de estacionalidad. Debe ser un intervalo que contenga representadas aproximadamente por igual todas las épocas del año.

Una vez establecida la forma de definir el periodo de referencia, en los siguientes apartados se detallan por pasos los procedimientos para determinar las funciones  $e_c$  y  $e_f$ .

### 2.3.2.1. Determinación de $e_c$

#### Establecer una función de tendencia con carácter descriptivo

Como paso previo a la obtención de la curva de estacionalidad, es necesario estimar una línea base que sirva para determinar el nivel de consumo general a gran escala. Esta línea de tendencia se ha establecido a partir de la serie de consumos semanales por uno de entre dos métodos posibles: línea de tendencia de orden 1, o valor medio de la serie.

Es preceptivo seleccionar la última opción cuando los datos observados no permiten evidenciar la existencia de una evolución monótona, a nivel general, en los consumos registrados. En particular, cuando el intervalo de análisis es muy corto (menos de dos años), resulta difícil identificar un comportamiento de este tipo salvo que éste sea muy marcado. Contrariamente, se deberá optar por la línea de tendencia cuando en el intervalo de análisis se aprecie claramente la existencia de una tendencia creciente o decreciente en la magnitud de los consumos.

Nuevamente conviene advertir que no se pretende establecer un modelo con capacidad predictiva para la tendencia a largo plazo, sino simplemente obtener una representación matemática de la misma, con carácter meramente descriptivo, que se utilizará, a efectos de este análisis, para eliminar su efecto sobre los datos originales.

#### Eliminación de la tendencia de los datos de consumo semanal

Se pretende obtener una serie que represente las variaciones del consumo semanal no relacionadas con la tendencia a largo plazo. Para ello, para cada dato de consumo semanal, se calcula el ratio entre dicho valor y la línea de tendencia correspondiente, según su posición en el año.

Estos cocientes son los factores por los que hay que multiplicar los valores de la línea de tendencia para restaurar los consumos semanales. Indican en qué medida el consumo observado es mayor o menor que el que le correspondería si no existiera variación debida a la tendencia. Acumula pues, el efecto de la estacionalidad más otros factores (meteorológicos, aleatorios, etc.) entre los que, por construcción, no se encuentra el efecto de calendario. Cabe suponer, por tanto, que la variabilidad de este parámetro será debida, en buena medida, a la estacionalidad.

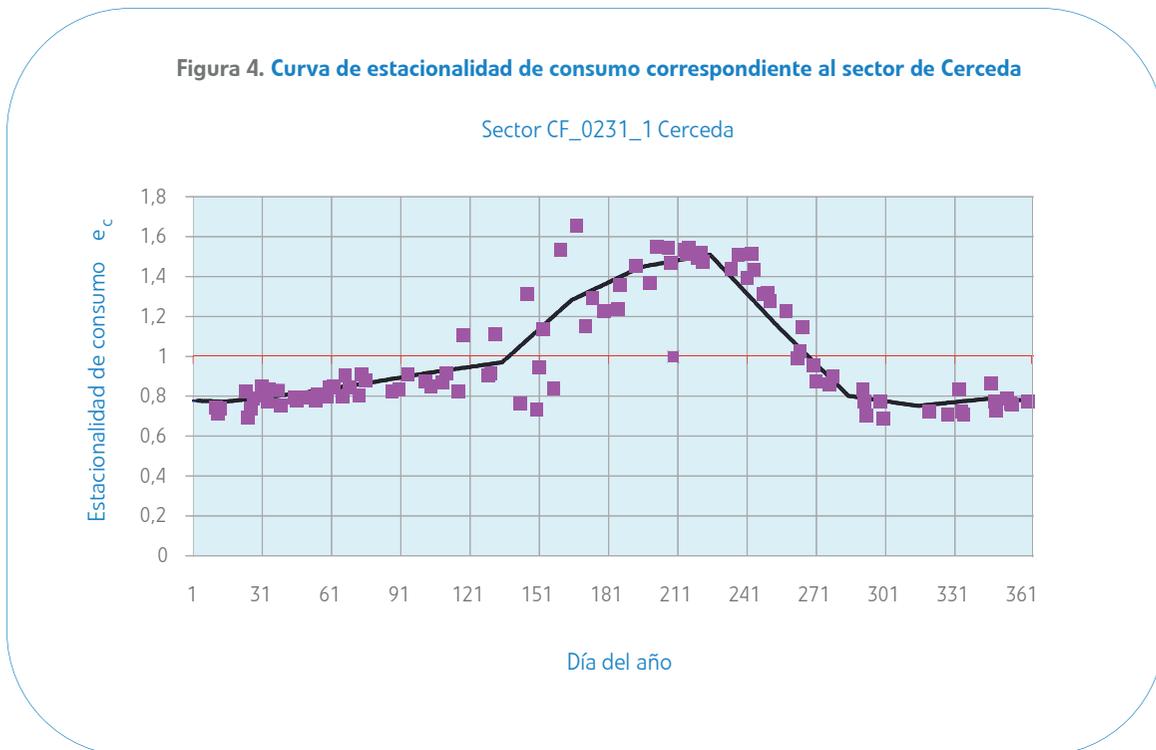
Como consecuencia de lo anterior, al representar el valor de este parámetro contra su posición en el año (día del año correspondiente al jueves de la semana) deberá quedar plasmado el patrón de estacionalidad general de la serie.

#### Determinación de la curva de estacionalidad

La curva de estacionalidad se obtendrá de ajustar una línea al conjunto de puntos definidos por cada uno de los ratios obtenidos para cada jueves del mes, pertenecientes al conjunto de referencia, una vez excluidas las semanas con festivos.

Ante la variabilidad e irregularidad de las formas que se han observado en dichas nubes de puntos, se ha descartado cualquier aproximación funcional. En su lugar, se ha optado por una función empírica construida a partir del promedio mensual de los ratios. Este promedio se ha asignado a la fecha intermedia del mes en cuestión. A partir de estos 12 valores, por medio de una interpolación lineal, se obtienen los correspondientes a los 365 días del año. De esta manera, queda definida la función de estacionalidad.

En la figura 4 se incluye, a modo de ejemplo, la curva de estacionalidad obtenida para el sector de Cerceda junto con los valores del ratio entre consumo medio semanal y tendencia correspondientes a cada jueves del periodo de referencia.



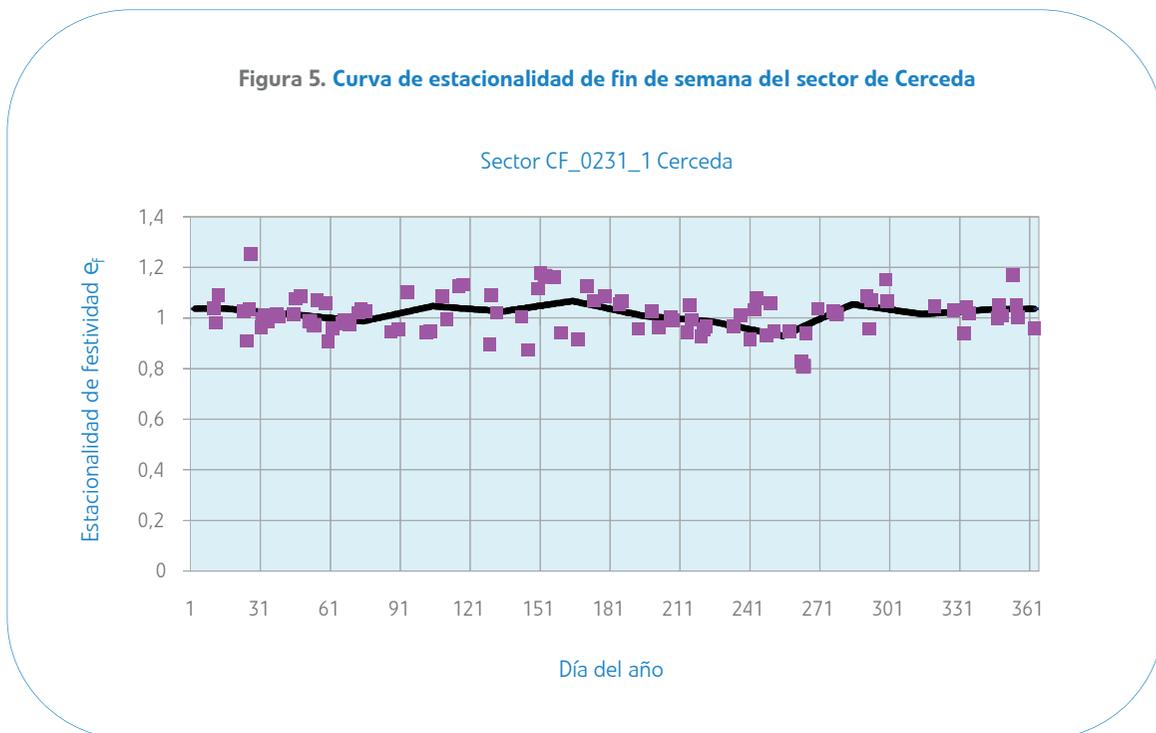
Se hace evidente el claro componente estacional que presenta este sector. El fuerte aumento de los consumos asociados a los meses de verano se justifica por la tipología del sector, compuesto fundamentalmente por segundas residencias. La dispersión existente en los meses de mayo y junio es imputable a la influencia que la meteorología suele tener en esas fechas, ya que suele condicionar el inicio de los riegos programados o el llenado de piscinas.

### 2.3.2.2. Determinación de $e_f$

#### Determinar los datos de consumo domingo/consumo medio diario

Como dato base para estudiar la estacionalidad del calendario se ha tomado la relación entre el consumo acumulado del domingo y el consumo medio diario de la semana correspondiente, evaluado en las semanas sin festivos. Se trata, por tanto, de un dato semanal cuyo valor se puede interpretar como un factor representativo del efecto de la festividad en el consumo.

A partir de estos valores, el método aplicado para obtener la curva de estacionalidad es igual que el empleado para el cálculo del otro factor, es decir, un promedio por mes de la relación antes descrita y una interpolación con un intervalo de un día. En el siguiente gráfico se representa la nube de puntos correspondiente al ratio de consumos junto con la curva de ajuste obtenida para el sector de Cerceda.



En este caso particular se comprueba cómo el efecto de la festividad es de reducción del consumo en los meses estivales (julio y agosto) donde se espera un consumo diario del mismo orden de magnitud que en fines de semana, mientras que para el resto del año la estacionalidad se mantiene por encima de 1.

Como último apunte cabe destacar que, durante las tareas de modelización, para no pocos sectores los datos de consumo de semanas sin festivos disponibles en el conjunto de referencia no han sido suficientes para completar un año. Para solventar esta carencia se han creado curvas tipo de estacionalidad del consumo ( $e_c$ ) y de estacionalidad de fin de semana ( $e_f$ ) por tipologías de sectores, a partir de aquellos que sí contaban con una serie de referencia de suficiente longitud.

Como paso previo, se ha realizado una clasificación de sectores por tipos a los que se atribuyen unas funciones de estacionalidad similares. Los tipos se identifican por una serie de características generales del sector y de

sus datos de consumo, que pueden determinarse con series de datos menos completas que las necesarias para establecer la función de estacionalidad.

Esta clasificación permite asignar una función de estacionalidad (siquiera provisional) a un sector con datos insuficientes para calcularla por el procedimiento directo.

El método de determinación de estas curvas tipo, asociado a la identificación de patrones de consumo por sectores, se explica en detalle en el apartado 5.6.

### 2.3.3. Análisis de los efectos de calendario

Los efectos de calendario que se han identificado en las series de consumos diarios, y las variables con que se han representado en la modelización desarrollada son los que se describen a continuación.

#### 2.3.3.1. Efecto de fin de semana y festividad

El efecto calendario presenta dos facetas fundamentales: las diferencias en las actividades urbanas según se trate de días festivos o no, y los desplazamientos de la población asociados a los días de descanso. Parece evidente que, en relación con estos efectos, conviene distinguir, al menos, tres categorías o tipos de días:

- a) Días laborables normales.
- b) Días laborables, con horario comercial normal, pero en los que la mayoría de la población no tiene jornada laboral (sábados, días de puente).
- c) Días festivos.

Para caracterizar estos tres tipos de días se han definido dos variables: **FFA** y **FFM** (siglas que responden a Factor de Festividad de Actividad y Movilidad, respectivamente).

- **FFA**, festividad estricta, toma valor 1 en los días festivos y valor 0 en los días laborables.
- **FFM**, más asociado a la movilidad de la población, toma valor 1 los días festivos, los sábados y los días de diario que quedan entre dos festivos o un festivo y un sábado (días de puente).

Con la variable **FFA** se pretende abarcar las diferencias en la actividad urbana entre los días festivos y días de diario: personas que acuden al trabajo, comercios que cierran, etc.

Por su parte, la variable **FFM** se refiere a los movimientos de personas asociados a los días de descanso: viajes de fin de semana o puente, desplazamiento a una segunda residencia, etc.

#### 2.3.3.2. Estacionalidad del calendario

En relación con el efecto del calendario, hay que señalar que éste presenta, a su vez, estacionalidad. En múltiples sectores se ha comprobado que el efecto de los días festivos es más acusado en la época estival que en la invernal. Para incorporar esta particularidad dentro de la modelización se han incorporado las variables **FFAE** y **FFME** que son, respectivamente, los cocientes de **FFA** y **FFM** entre  $e_f$ . Estas variables actúan como términos de interacción entre las variables de festividad y la estacionalidad.

### 2.3.3.2. Efecto de la Semana Santa

En muchas series de datos se observan bajadas y ascensos muy significativos en los consumos de los días de Semana Santa, especialmente de Jueves Santo a Domingo de Pascua. Se puede suponer que el efecto de varios días de fiesta consecutivos, vacaciones escolares y la tradición vacacional de la Semana Santa se combinan de modo que estos días resulten excepcionales en cuanto a consumo de agua.

Para incluir la Semana Santa en el modelo se ha definido la variable **SS**, que toma los siguientes valores:

- Lunes Santo a Miércoles Santo: 0,5
- Jueves Santo a Domingo de Pascua: 1
- Resto del año: 0

### 2.3.3.4. Festividad de Madrid

Las variables de festividad pueden tomar valores diferentes, según el sector considerado, debido a las fiestas locales. En este sentido, se ha considerado que la festividad del municipio de Madrid puede ser relevante en todos los sectores y no sólo en aquellos pertenecientes a la capital, ya que es evidente que muchos de los hábitos poblacionales están vinculados a la actividad de la urbe. Así pues, en los sectores exteriores al municipio de Madrid se han proporcionado al modelo las variables de festividad de Madrid, además de las propias.

### 2.3.3.5. Resultados

Con este juego de variables de calendario se ha procedido a determinar su influencia en el consumo diario, suponiendo que esta puede representarse mediante una ecuación lineal. Para ello se han seguido los pasos que se describen a continuación.

1. Determinar la serie de consumos diarios "desestacionalizados", dividiendo los datos de consumo diario por los valores correspondientes del factor de estacionalidad de consumo,  $e_c$ .
2. Ajustar a esta serie un modelo de regresión lineal con las variables de calendario. En el proceso se analiza la significación de los coeficientes asignados a cada una de las variables y la colinealidad de las mismas, de modo que se seleccionan aquellas que son relevantes. También se analiza el signo de los coeficientes, en relación con el efecto que a priori se atribuyen las variables correspondientes.
3. Una vez seleccionado un modelo de regresión adecuado, sus coeficientes se interpretan como los coeficientes del término de calendario para las variables correspondientes, en el modelo previo a la incorporación de la meteorología. Las variables que intervienen en cada modelo pasan a ser consideradas como predictores de calendario para el sector correspondiente.

#### 2.3.4. Análisis de la influencia de la meteorología

Para determinar e intentar cuantificar en qué medida el consumo de agua de la red gestionada por Canal de Isabel II podía estar influido por las condiciones meteorológicas, se partió del análisis de la serie histórica de consumos diarios del conjunto de la red de Canal de Isabel II. Se utilizaron datos de consumo medio diario correspondientes al periodo enero 1997- enero 2008.

No se descarta que el efecto de la meteorología pueda ser diferente entre sectores y, que incluso, se manifieste con más claridad en el consumo de agua de determinados sectores que en el consumo total. Con todo, se optó por el análisis de la serie agregada atendiendo a las siguientes razones:

- a) La serie de consumo total es, con diferencia, la que presenta un periodo de datos más largo y de mejor calidad. Por lo general, los periodos de datos válidos pertenecientes a sectores son más cortos y fraccionados. La heterogeneidad en los periodos de datos válidos podría ser fuente de confusión al intentar capturar la heterogeneidad del comportamiento de los sectores.
- b) Se confió en que, una vez identificado el efecto de la meteorología en la serie de consumo total, y establecida una formulación genérica adecuada para simularla, la particularización de esta en cada uno de los sectores permitiría una buena representación de su comportamiento específico.
- c) Las relaciones cuantitativas entre meteorología y consumo de agua para la serie de consumo total pueden servir de base para establecer predicciones de consumo que son interesantes desde el punto de vista de la planificación general.

Por otra parte, se optó por prescindir del análisis a escala horaria fundamentalmente por el hecho de que las fluctuaciones de la temperatura se antojaban demasiado parsimoniosas como para manifestarse de forma significativa en el consumo de agua de una hora. Para la precipitación, aunque presenta fluctuaciones más rápidas (pasa de no llover a llover, en un instante), no se espera un efecto inmediato en el consumo.

Como variables meteorológicas, se optó por seleccionar las variables temperatura y precipitación. Se descartaron otros parámetros como insolación, nubosidad, velocidad del viento etc. ya que, debido a la práctica ausencia de datos históricos, no resultaba posible establecer relaciones significativas con el consumo de agua.

Para el estudio se utilizaron las series de precipitación diaria y temperaturas media, mínima y máxima diaria de la estación de Retiro, que se puede considerar representativa de la situación meteorológica en el conjunto del territorio abastecido por Canal de Isabel II.

##### 2.3.4.1. Influencia de la temperatura

Como es obvio, el ciclo estacional de la temperatura ambiente coincide sensiblemente con las variaciones estacionales del consumo de agua (mayores consumos en la época estival). La correspondencia entre estacionalidad del consumo y estacionalidad de la temperatura indica, entre otras cosas, una relación general y positiva entre consumo de agua y temperatura. Sin embargo, esta relación ya está recogida en el término de estacionalidad del modelo y no es la que interesa en este análisis.

Más allá de esta variación, a la fluctuación estacional general de la temperatura diaria se superponen unas variaciones de frecuencia más alta, de manera que un día concreto puede ser más caluroso o más frío de lo que puede considerarse normal para la época del año. El objetivo planteado fue intentar cuantificar el reflejo de estas variaciones en el consumo de agua y plantear una formulación que permitiese reproducir este comportamiento.

Para ello resultó necesario, en primer lugar, establecer cuál es la temperatura que se considera “normal” para una época del año. De forma genérica, se adoptó el criterio de considerar temperatura normal para una época del año, el promedio de las temperaturas de los días de esa época del año, en el periodo de datos analizado.

A efectos de este estudio se estableció un valor de temperatura normal para cada uno de los 365 días del año, que se determina según se explica en los puntos siguientes:

- Para un día del año dado, se consideran “misma época del año” los 15 días que van desde el séptimo anterior, al séptimo posterior. Por ejemplo, los días de la misma época del año que el 15 de marzo son los días desde el 8 hasta el 22 de marzo.
- La temperatura normal de un día del año es el promedio de las temperaturas diarias de todos los días de la misma “época del año” del periodo de datos estudiado (enero de 1997- enero de 2008).

El cálculo de las temperaturas diarias normales se realizó para la máxima, la mínima y la media diarias. A partir de los valores de temperatura normal, se define la anomalía de temperatura para un día concreto como la diferencia entre la temperatura registrada ese día y la temperatura normal que corresponde a ese día del año. La anomalía de temperatura es una medida de cuánto se aparta la temperatura registrada en un día cualquiera respecto del valor normal o esperable.

A modo de ejemplo, en la figura 6 se muestra la serie de temperatura media diaria de 2007 superpuesta a la serie de temperatura media normal, y la serie resultante de anomalía de temperatura media (**ATM**). Se observan, por ejemplo, periodos fríos en la segunda mitad de marzo, la segunda mitad de mayo y la quincena central de junio. Los días centrales de abril, en cambio, fueron más calurosos de lo normal.

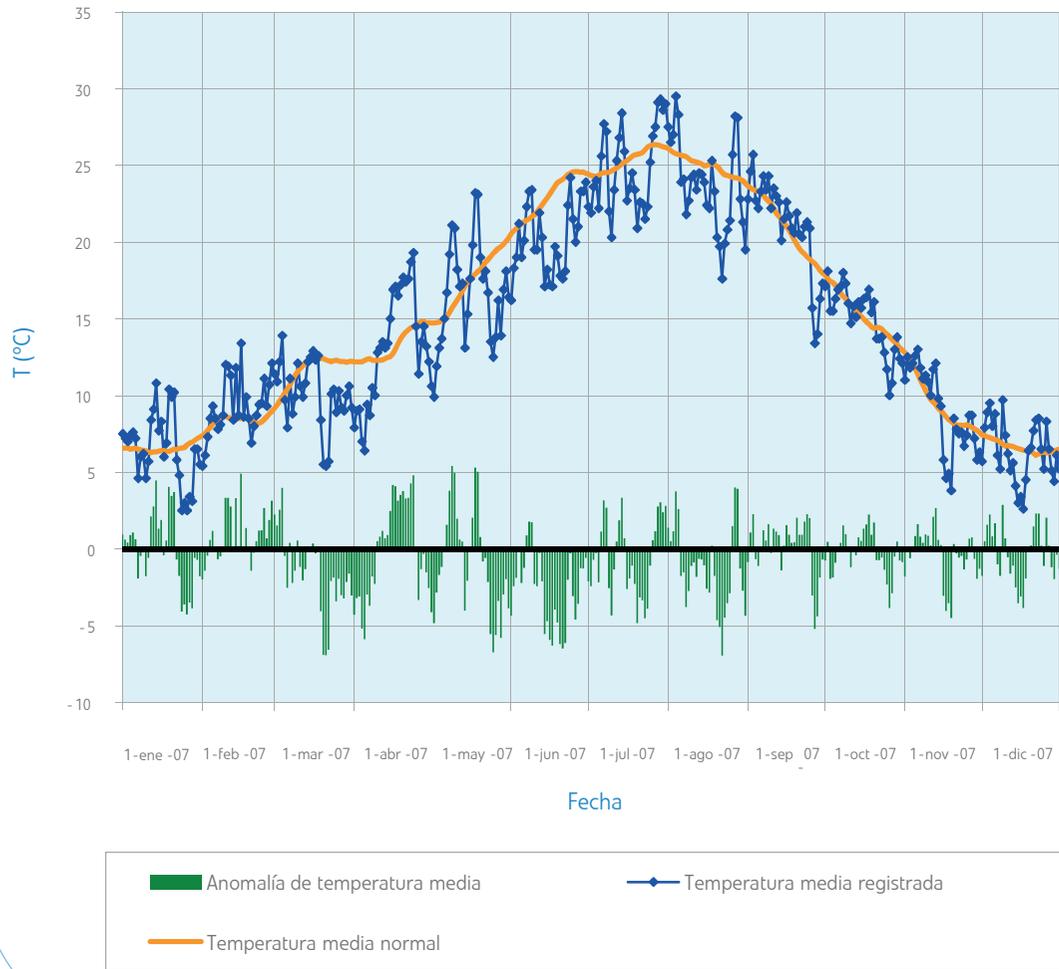
Una vez definida esta nueva variable, el siguiente paso consistió en intentar interpretar a través de ella las fluctuaciones de consumo que no podían ser explicadas por los factores de “tendencia”, “estacionalidad” y “calendario” según la formulación adoptada, es decir, los residuos obtenidos en el ajuste de la serie de consumos diarios a través de la formulación que no considera el efecto meteorológico.

Planteada la correlación entre estas dos variables, los resultados obtenidos fueron positivos, existiendo una relación claramente significativa en prácticamente todos los meses del periodo estudiado. El coeficiente de correlación obtenido fue superior a 0,4. Esto quería decir que, aproximadamente un 20 por ciento de la anomalía de consumo respecto al modelo Tendencia, Estacionalidad y Calendario (en adelante lo referiremos por las siglas de sus iniciales, modelo **TEC**, expuesto en el apartado anterior) podía ser explicada a partir de la variable de anomalía de temperatura. Obviamente, la relación es de tipo potenciadora, es decir, temperaturas por encima de lo normal potencian el consumo.

Como ejemplo del comportamiento observado, en la figura 7 se representan, para el año 2007, las series de consumo diario observado y consumo simulado **TEC**, a las que se añade la serie de anomalía de temperatura media (**ATM**). “De visu” se pueden apreciar algunas coincidencias.

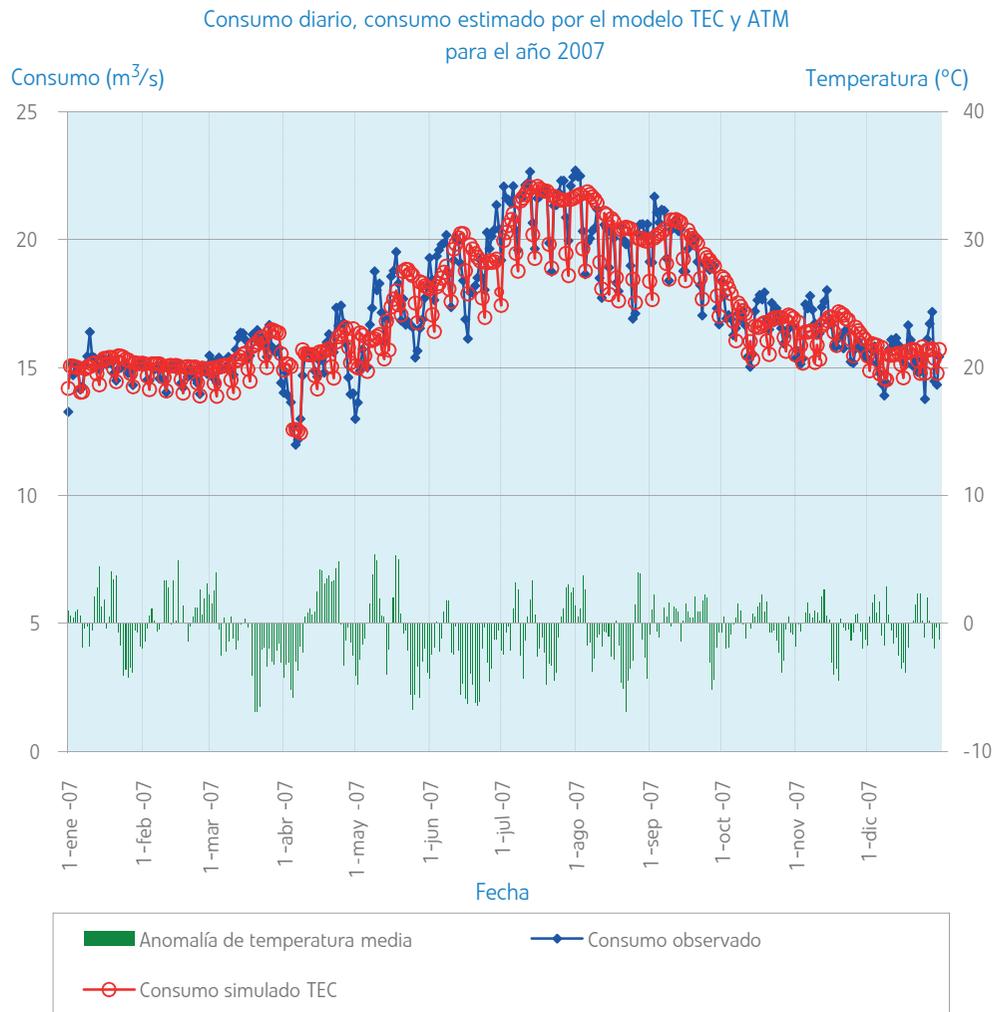
Figura 6. Evolución de las variables de temperatura en el año 2007

Temperatura media registrada y anomalía de temperatura media (ATM) para el año 2007



En los meses de mayo y junio, los periodos calurosos, ( $ATM > 0$ ) coinciden claramente con consumos por encima del consumo esperado **TEC** y, los periodos fríos, ( $ATM < 0$ ) coinciden con consumos por debajo del consumo esperado **TEC**.

**Figura 7. Desviación en la estimación del consumo en el modelo TEC y relación con el predictor ATM**



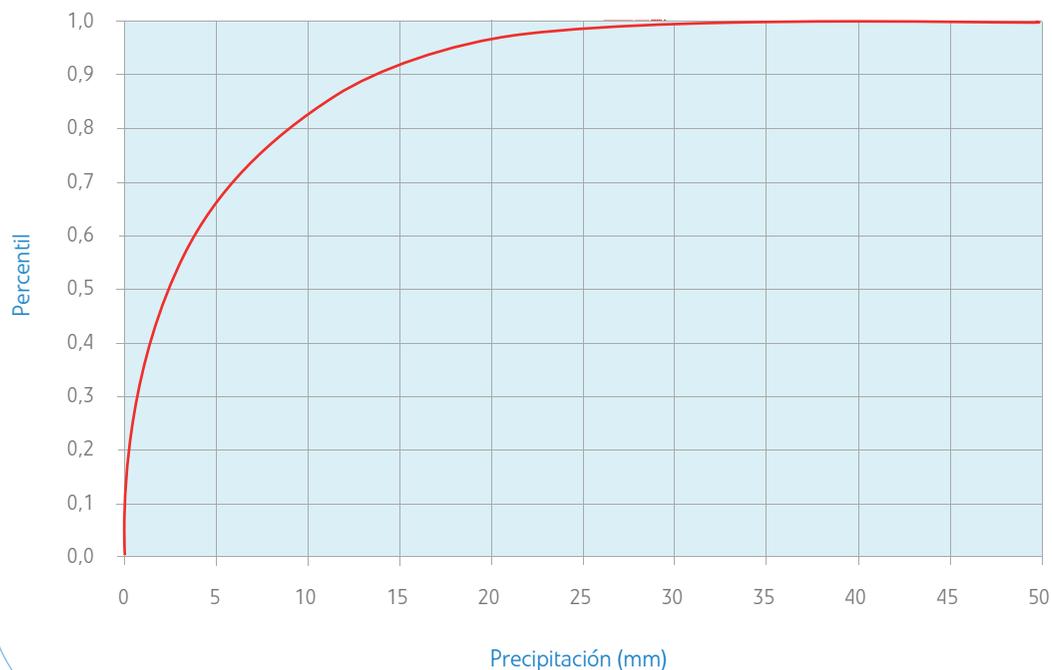
### 2.3.4.2. Influencia de la precipitación

Como punto de partida, se aceptó la hipótesis de suponer que la respuesta del consumo de agua a la precipitación no fuese lineal. En efecto, la magnitud de dicha respuesta depende en buena medida de la percepción que tengan los usuarios de su magnitud. Parece claro que entre 1 y 5 milímetros se percibe una diferencia clara de magnitud, mientras que un día con 30 milímetros de lluvia y otro con 34 milímetros se percibirán prácticamente iguales, aunque en ambos casos la diferencia sea de 4 milímetros.

Para utilizar una variable que permitiese incorporar este comportamiento del efecto de la precipitación, se optó por transformar el valor de la precipitación diaria por su percentil en el conjunto de valores de precipitación diaria  $>0$  del periodo de estudio. Para días sin precipitación, la variable vale 0 y para valores de precipitación diaria igual o mayor que 50,2 milímetros, (máximo registrado en el histórico de datos utilizado) el valor de la variable es 1. Se ha denominado **pP** (por "percentil de precipitación" a esta variable). En la figura 8 se representa la transformación de P a **pP**.

**Figura 8. Transformación de la variable de precipitación en su percentil**

Transformación del valor de la precipitación diaria por su percentil a partir de los registros existentes en la serie histórica



Por otra parte, es esperable que el efecto de la precipitación pueda presentarse diferido en el tiempo al menos algunos días. Poniendo por caso el uso de agua para riego de jardines, es concebible que un día de lluvia abundante tenga efecto inhibitor en ese día, pero también en los siguientes. Por tal motivo, se ensayaron variables que combinaran la precipitación de varios días consecutivos.

Para una amplia gama de percentiles y promedios de distinto alcance de variables de precipitación, se plantearon las relaciones existentes con los residuos del ajuste del consumo medio diario según el modelo **TEC**, de manera semejante a como se realizó en el análisis de la influencia de la temperatura (ver apartado anterior).

A partir de los resultados obtenidos se pudieron establecer las siguientes conclusiones generales:

- Hay correlación significativa entre la precipitación y el consumo de agua, si bien es menos marcada que la relación con la temperatura.
- Esta relación es de tipo inhibitor ( $r < 0$ ): la ocurrencia de precipitación reduce el consumo de agua.
- En general, el efecto de la precipitación sobre el consumo perdura varios días. La correlación con la anomalía de consumo **TEC** es mejor para las variables de precipitación que promedian la precipitación de varios días. Concretamente, las mejores correlaciones se obtienen para las variables que acumulan la precipitación de tres o cuatro días, sin apenas diferencia entre ambas.
- Las variables de precipitación transformadas por el percentil proporcionan mejores correlaciones que las variables de precipitación sin transformar. En consecuencia, se considera que las variables **pP** capturan mejor que la propia variable precipitación el efecto de esta sobre el consumo.

En función de lo anterior, se decidió adoptar **pP3** (percentil de la precipitación media de los tres últimos días) como variable representativa para el estudio del efecto de la precipitación sobre el consumo de agua.

#### 2.3.4.3. Estacionalidad de la influencia de la meteorología

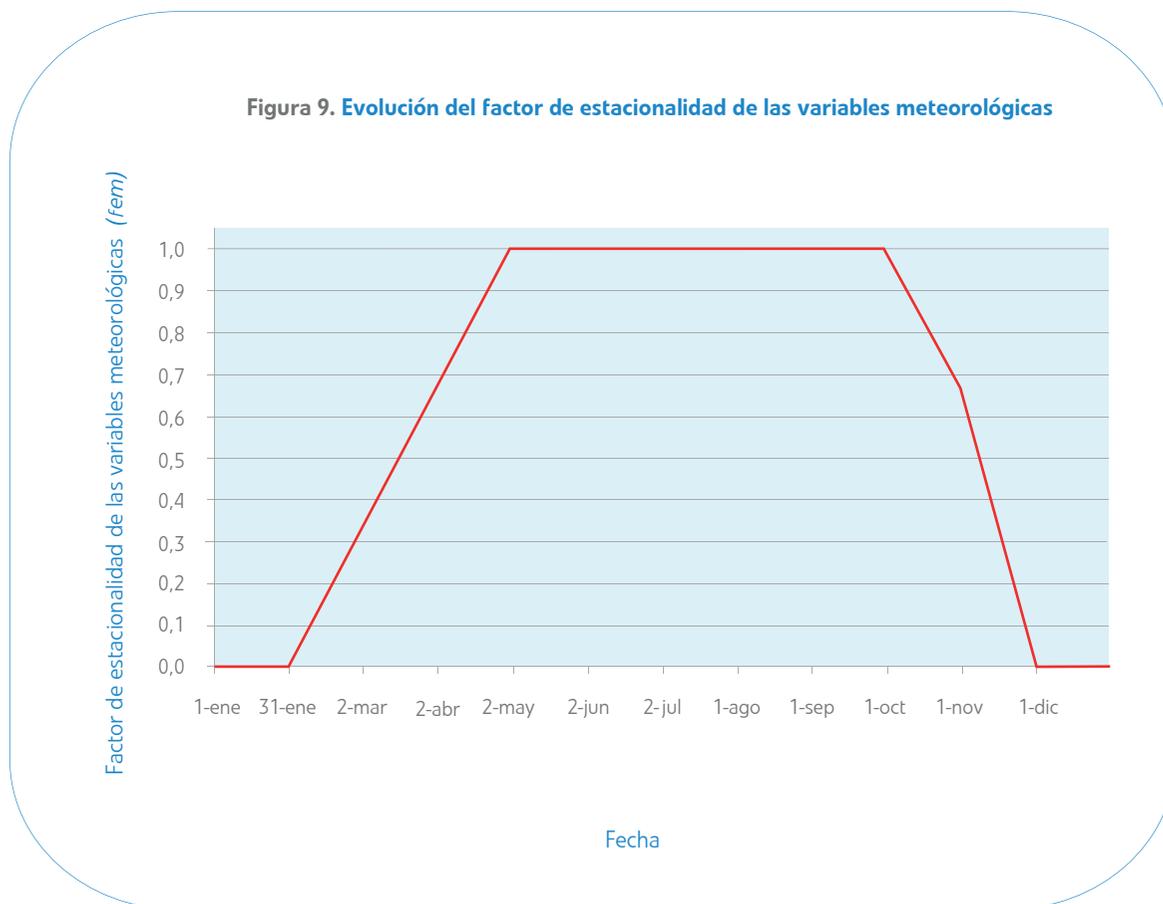
En los apartados anteriores ha quedado fundamentado el proceso de selección de las variables del modelo matemático con el que se quiso representar la influencia del factor meteorológico en el patrón de consumos.

Sin embargo, hasta ahora se ha omitido el hecho de que la influencia de las variables meteorológicas en el consumo de agua no es homogénea a lo largo del año. Como es previsible, el reflejo que en términos de consumo puede tener el hecho de que un mes de enero o febrero sea más cálido o más lluvioso de lo habitual, será mucho menor que si esa desviación respecto a la temperatura media o esperada se produce en el mes de mayo o septiembre. En los meses de invierno es difícil que la temperatura o precipitación puedan condicionar los hábitos de la población, mientras que en los meses inmediatamente anteriores o posteriores al verano una temperatura o precipitación anormal pueden traducirse en una modificación muy notable de los hábitos poblacionales, un adelanto (o retraso) en el inicio (o fin) del periodo de riegos o de llenado de piscinas, factores todos ellos de gran relevancia a efectos del consumo.

Para evaluar la estacionalidad del efecto de la meteorología se calcularon los coeficientes de determinación, **r<sup>2</sup>**, entre las variables meteorológicas **ATM** y **pP3** con la anomalía de consumo por meses para toda la serie que constituye el conjunto de referencia. Los resultados permitieron poner de manifiesto la evidente variación de la influencia de la respuesta a la meteorología en función de la época del año.

Así, la correlación entre anomalías de temperatura y anomalías de consumo **TEC** es muy baja o prácticamente nula en noviembre, diciembre y enero, y presenta un periodo de valores altos entre abril y septiembre, con máximo en mayo.

La correlación entre precipitación y anomalía de consumo **TEC** sigue un patrón más irregular, pero los valores más bajos también se sitúan en el periodo noviembre-enero. En función de estos resultados, la estacionalidad de la respuesta a las variables meteorológicas se ha incorporado afectando a los predictores meteorológicos con un factor variable en función del día del año (*fem*). Las formas de variación adoptadas han sido simples: un conjunto de meses en el centro del año con *fem*=1, otro periodo de meses invernales con *fem*=0, y transiciones lineales en los meses intermedios. En la figura 9 se recoge la curva adoptada para este factor de estacionalidad de las variables meteorológicas a lo largo del año.



La figura anterior responde a las siguientes relaciones:

- En diciembre - enero, *fem*=0.
- En mayo - septiembre, *fem*=1.
- Entre el 31 de enero y el 1 de mayo, *fem* aumenta linealmente desde 0 hasta 1.
- Entre el 30 de septiembre y el 31 de octubre, *fem* disminuye linealmente entre 1 y 2/3.
- Entre el 31 de octubre y el 1 de diciembre, *fem* disminuye linealmente entre 2/3 y 0.

### 2.3.5. Tendencia

En la formulación recogida en el apartado 2.3.1, el único término que aún no ha sido descrito es el factor de tendencia **TEN**. La tendencia a largo plazo representa una variación parsimoniosa del nivel general de consumo que se pone de manifiesto en un número significativo de sectores, que cuentan con series de datos de longitud suficiente como para que ésta se haga visible y que cabe atribuir a múltiples factores: variaciones a largo plazo en la población abastecida, en los usos del agua, en los hábitos de consumo, modificaciones en la red de distribución, etc.

Como se ha indicado anteriormente, se ha prescindido de plantear una formulación de la tendencia partiendo de las causas que la motivan, que por otra parte es esperable que sean muy diversas debido a la heterogeneidad de la tipología de patrones estudiada.

En su lugar, se ha planteado incorporarla como un factor corrector del valor de consumo inicialmente simulado por el modelo. Así, se ha asumido que la tendencia actúa de forma gradual, de modo que su efecto para los próximos días se podrá estimar con precisión suficiente a partir del nivel de consumos validados de los 14 días pasados (se elige un número múltiplo de 7 para que el periodo de referencia recoja semanas completas). Para el ajuste de los modelos se consideran exclusivamente consumos diarios del conjunto de referencia.

De esta manera, este factor actúa como corrección sobre los consumos simulados para adecuarlos a los niveles de consumo que marca la tendencia a largo plazo. Su valor se establece en función de las observaciones de consumo recientes, según la expresión:

$$T_i = \frac{\sum_{j=i-14}^{i-1} c_j}{\sum_{j=i-14}^{i-1} \frac{c_j^*}{T_j}}$$

Donde los **j** se refieren a los 14 últimos días con dato de consumo, **c<sub>j</sub>** son los valores de consumo medio diario validados para el día **j**-ésimo, **c<sub>j</sub><sup>\*</sup>** son las estimaciones de consumo realizadas por el modelo antes de incorporar el factor corrector de tendencia para el día **j**-ésimo y **T<sub>j</sub>** se corresponde con la estimación del factor de tendencia correspondiente también al día **j**-ésimo. Es decir, el cociente anterior es la relación entre la suma de consumos aceptados en los 14 días anteriores al que se estima y las 14 estimaciones del modelo para esos 14 días, en el caso de no aplicar el factor de tendencia. A este último término (estimación del modelo antes de aplicar el factor de tendencia) se le ha denominado en la terminología seguida como **cest** (consumo estimado sin tendencia).

$$cest_j = \frac{c_j^*}{T_j}$$

Por lo que la expresión que define el factor de tendencia puede escribirse de manera equivalente como:

$$T_i = \frac{\sum_{j=i-14}^{i-1} c_j}{\sum_{j=i-14}^{i-1} cest_j}$$

### 2.3.6. Formulación

Una vez definidos los planteamientos utilizados para incorporar a la formulación los efectos de los factores que se ha considerado que tienen mayor influencia en la evolución del consumo diario, puede establecerse la formulación del modelo que a continuación se recoge:

$$c_i^* = e_{ci} [a_0 + a_1 \cdot FFA_i + a_2 \cdot FFM_i + a_3 \cdot SS_i + a_4 \cdot FFAE_i + a_5 \cdot FFME_i + a_6 \cdot ATM_i + a_7 \cdot pP3_i] \cdot T_i$$

Donde:

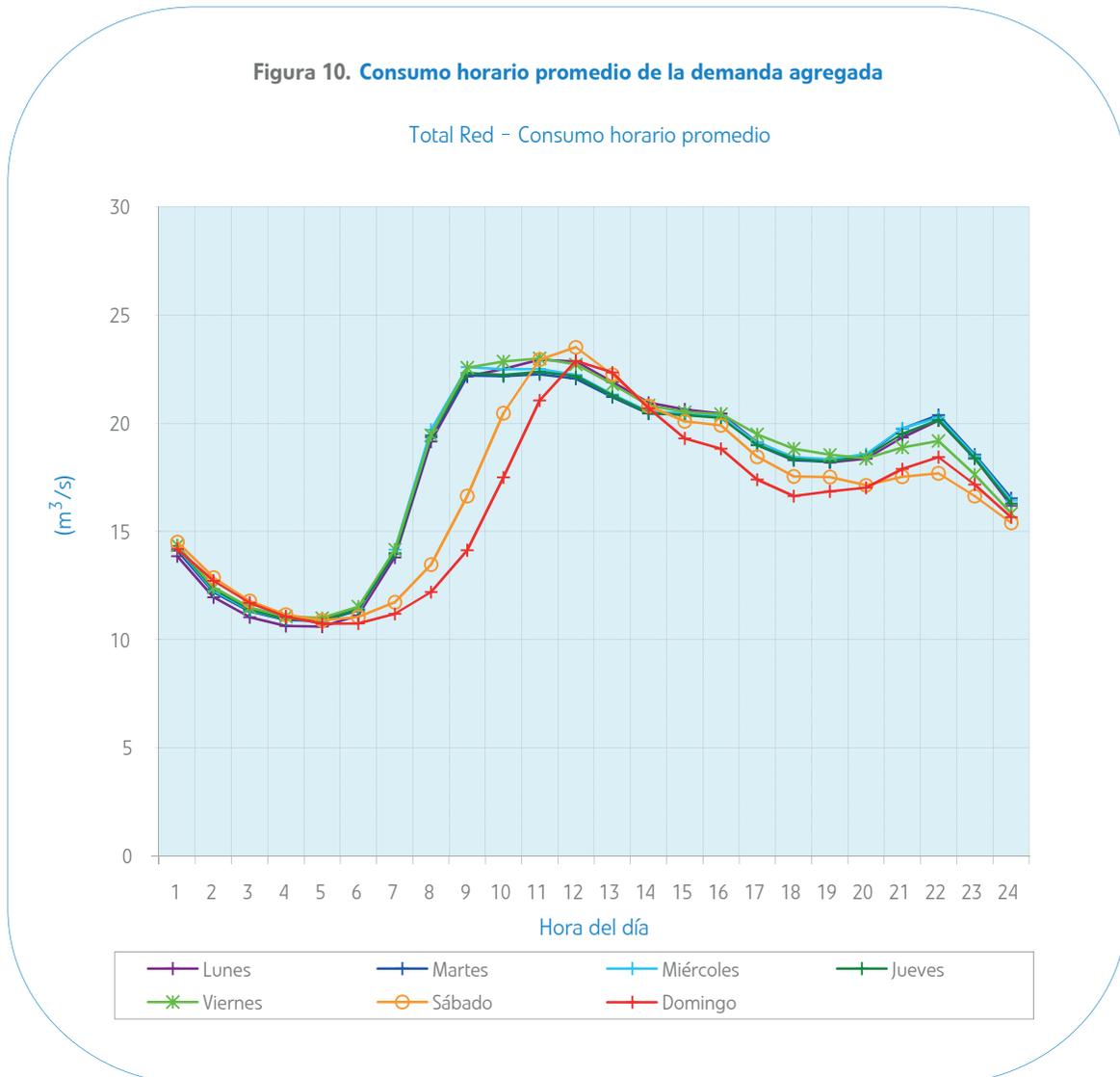
- $c_i^*$  es la estimación de consumo en el día  $i$ -ésimo.
- $e_{ci}$  es el factor de estacionalidad de consumo correspondiente a ese día.
- $a_0, a_1, \dots, a_7$ . son las constantes de modelización, propias a cada sector.
- FFA, FFM, FFAE, FFME son los predictores asociados al componente de movilidad y actividad del calendario (incluyendo, en los dos últimos, el componente estacional de estos patrones).
- SS es el predictor asociado a la Semana Santa.
- ATM y pP3 son los predictores meteorológicos relacionados con la anomalía de temperatura y precipitación respectivamente.

- $T_i$  es el factor de tendencia dónde
- $$T_i = \frac{\sum_{j=i-14}^{i-1} c_j}{\sum_{j=i-14}^{i-1} \frac{c_j^*}{T_j}}$$

## 2.4. Modelo horario

### 2.4.1. Patrones de variación intradiaria

Al igual que ocurría con el patrón de consumo diario a escala anual o semanal, a nivel horario, la evolución del consumo de agua también suele presentar regularidades observables e interpretables. Como ilustración, en la figura 10 se recogen los valores promedio de consumos totales horarios suministrados a la red de abastecimiento por Canal de Isabel II, para cada día de la semana, del periodo comprendido entre los años 1997 y 2007.



Si bien el estudio extendido al conjunto de sectores, ha permitido identificar comportamientos a escala intradiaria intrínsecos a una determinada tipología de sector de consumo (por ejemplo el patrón observado en sectores donde predomina el uso industrial es muy característico, y se aleja de manera considerable, del patrón de consumo que puede caracterizar una pequeña población donde abundan viviendas unifamiliares, también fácilmente identificable por otra parte); el análisis del conjunto de la demanda tiene la ventaja de ser generalista

y de presentar características comunes a todas las agrupaciones de sectores que han podido identificarse. En cualquier caso, el apartado 5 está dedicado íntegramente a la descripción de los patrones de consumo que han podido identificarse como asociados a una determinada tipología de sectores.

Según se observa en la figura 10, en todos los días de la semana existe un claro descenso nocturno del consumo que alcanza siempre un nivel mínimo parecido alrededor de la hora 5. En la evolución del consumo durante las horas diurnas, en cambio, cabe diferenciar dos tipos de comportamiento: el de los días de diario y el de los fines de semana.

En los días de diario (lunes a viernes no festivo), el valle nocturno concluye con una subida rápida del consumo entre las horas 7 y 8. Entre las horas 9 y 16 se mantienen los niveles de consumo más altos del día a los que sucede un valle poco pronunciado con mínimo relativo a la hora 18. En este punto se produce un repunte del consumo con máximo en la hora 22, que da paso a la caída hasta los niveles nocturnos.

En el fin de fin de semana (sábado y domingo), la reactivación matinal del consumo es más tardía y progresiva: se produce entre las horas 8 y 10. Entre las horas 11 y 12 se tiene el máximo del día y, a partir de ese punto se produce una caída progresiva hasta los niveles nocturnos. Por otra parte, los patrones de consumos horarios observados para días festivos resultan similares a los del fin de semana.

El valle nocturno de consumo tiene una interpretación evidente asociada al descanso de la población en esas horas. Las diferencias entre el patrón de consumo de los días de diario y de los días del fin de semana debe atribuirse a la diferente actividad urbana entre un tipo de día y otro.

#### **2.4.2. Variabilidad estacional del patrón horario**

Al margen de las características generales reseñadas en el apartado anterior, el análisis particularizado para cada sector ha permitido observar variaciones estacionales importantes en los patrones de consumo horario de muchas de las series estudiadas.

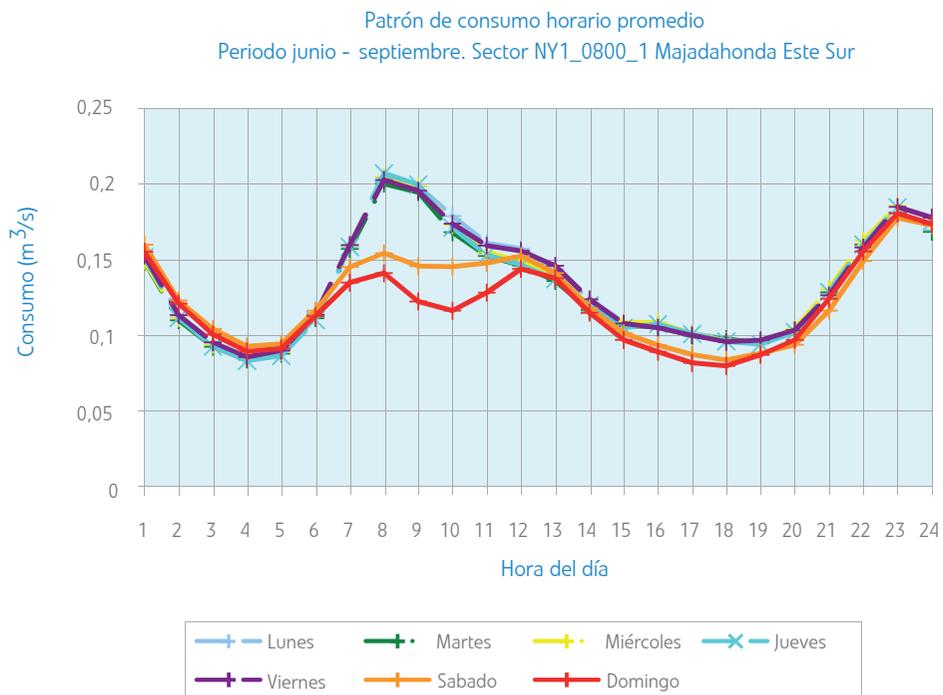
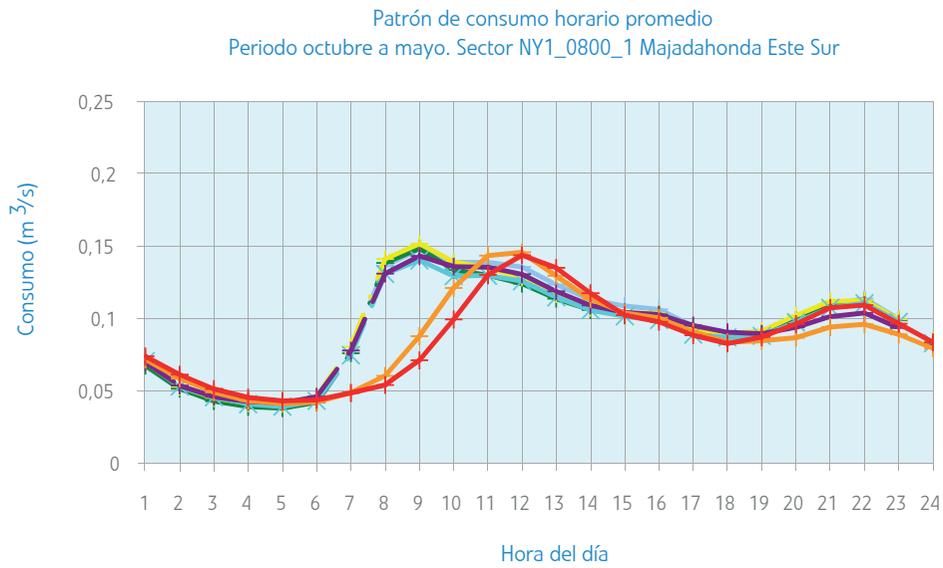
Se comprobó que existe una variabilidad estacional del patrón horario muy clara que se manifiesta sobre todo en las horas finales del día y, aunque es una característica generalizada, es especialmente llamativa en aquellos sectores donde una parte importante de las unidades de consumo se corresponden con viviendas destinadas a un uso de fin de semana o vacacional.

La pauta seguida para poner en evidencia este componente estacional ha sido la de representar separadamente los consumos correspondientes a los meses de verano (julio a septiembre) de los valores promedio pertenecientes al resto del año. Las principales diferencias se producen en el consumo durante el periodo nocturno. Así, entre las 22 y 24 horas se observa un significativo aumento de los consumos relativos en los meses de verano, llegándose a alcanzar en estas horas los máximos absolutos. Por el contrario, en los meses de invierno, estos sectores se caracterizan por presentar una punta de consumo nocturno mucho más reducida y que además se produce a una hora más temprana.

Claramente, este componente estacional va asociado a los hábitos de comportamiento propios de los meses de verano, caracterizado por un retraso en el inicio del periodo de descanso nocturno, quedando las horas finales del día destinadas al aseo personal y doméstico, y al riego de jardines.

Como ejemplo, en la figura 11 se aprecia la diferencia en el patrón horario que se produce en un sector perteneciente al municipio de Majadahonda, localidad residencial situada a unos 20 kilómetros del núcleo urbano de Madrid, exclusiva en cuanto a la renta media de la población allí residente y donde predomina la vivienda unifamiliar.

**Figura 11. Efecto de la estacionalidad en el consumo horario**



Es clara la diferencia entre el patrón de ambas curvas. El máximo nocturno correspondiente a los meses de verano se produce una hora más tarde (en la hora 23) y está al mismo nivel que el que se produce durante el mediodía. En el invierno, por el contrario, la curva ascendente correspondiente al consumo nocturno es mucho más tendida y los valores máximos son notablemente inferiores tanto en términos relativos como absolutos.

Al margen de la estacionalidad de los consumos en las horas finales del día, también resulta fácilmente apreciable la variabilidad estacional correspondiente al aumento de las demandas nocturnas en los meses de verano, que llegan a doblar los registros correspondientes a los meses de invierno en esa misma franja horaria. Esta diferencia de consumos es atribuible a los riegos programados de parques y jardines, que únicamente se producen en el periodo de estío.

### 2.4.3. Predictores seleccionados

En el análisis de las series de datos por sectores se ha comprobado que, en los casos en que las distribuciones horarias responden a una pauta interpretable en términos de comportamiento del consumo, dichas distribuciones son bastante estables. En estos casos, además, es posible establecer un conjunto de patrones horarios "tipo", correspondientes cada uno de ellos a determinadas condiciones definibles mediante variables principalmente de calendario. Es decir, se puede distinguir un patrón de consumo horario para los días laborables, otro para los domingos y festivos, etc. Al depender de variables de calendario, estos patrones son, además, anticipables.

Un conjunto de patrones tipo como el citado constituye una representación bastante completa de la variabilidad de los consumos horarios de un sector, al menos de la variabilidad que, como resultado del análisis realizado, se considera modelizable con la información disponible.

De lo anterior se desprende que los predictores principales en el nivel horario serán variables de calendario. Esto es así, en efecto, por lo que las variables de calendario **FFA** y **FFM** (Factor de Festividad de Actividad y Movilidad, respectivamente), utilizadas en el nivel diario, se han planteado también como predictores en los modelos horarios.

Sin embargo, los predictores considerados en el modelo diario presentaban el inconveniente de que no permitían representar el efecto de la estacionalidad a nivel de consumos horarios reseñada en el apartado anterior. A la hora de intentar modelizar el consumo horario exclusivamente con estas variables, se observó que los parámetros de validación horaria otorgaban una tolerancia claramente mayor a los consumos de las horas 22, 23 y 24 por presentar una menor precisión en la estimación de éstas, circunstancia debida precisamente a la variabilidad estacional.

Para responder a esta carencia se optó por definir dos nuevos predictores al modelo horario que intentaran reflejar esta variabilidad estacional del patrón horario: **FDEc** y **SEMEc**.

**FDEc** toma valor 1 dividido por el factor de estacionalidad del consumo ( $e_c$ ) cuando el **FFM** es igual a 1 y toma valor 0 cuando el **FFM** es igual a 0. **SEMEc** toma valor 1 dividido por el factor de estacionalidad del consumo ( $e_c$ ) cuando el **FFM** es igual a 0 y toma valor 0 cuando el **FFM** es igual a 1. Es decir, **FDEc** actúa en los días festivos y de fin de semana, y **SEMEc** actúa en los días de diario.

Simultáneamente, para no sobrecargar el modelo horario, es decir, para no aumentar el número de parámetros libres, se han desestimado los predictores **FFAE** y **FFME**, utilizados en la modelización a escala diaria pero que aportaban poca información en relación con **FFA** y **FFM**.

Además de las anteriores, el nivel horario incorpora una variable de calendario nueva: **AN**, que toma valor 1 los días 1 de enero, y 0 el resto del año. Esta variable se incluye para dar cabida en el modelo al hecho de que los días de Año Nuevo, aunque poco numerosos en la muestra, presentan, por lo general, un comportamiento claramente diferenciado del resto de días (con consumos nocturnos mayores y máximo matinal menos pronunciado y más tardío).

Por último, se ha renunciado a incorporar la influencia de la meteorología en el nivel horario del modelo. Esta influencia puede ser relevante (unos ejemplos del efecto meteorológico a escala horaria se han recogido en el capítulo 6), pero el nivel de resolución temporal de las predicciones actualmente disponibles no es suficiente para sacar provecho de un modelo que las contemplara.

#### 2.4.4. Formulación

La modelización de consumos en el nivel horario se ha acometido analizando las distribuciones horarias del consumo diario. En consecuencia, la magnitud que se ha utilizado para establecer los modelos horarios es el consumo horario referido a consumo medio horario del día, magnitud adimensional que hemos denominado consumo horario relativo.

$$\gamma_{ih} = \frac{c_{ih}}{\frac{1}{24} \cdot \sum_{\eta=1}^{24} c_{i\eta}}$$

Donde  $\gamma_{ih}$  es el consumo horario relativo de la hora **h** del día **i**-ésimo y  $c_{ih}$  es el consumo de la hora **h** del día **i**-ésimo.

Las estimaciones de consumo horario relativo permiten obtener estimaciones de consumo horario al combinarse con estimaciones de consumo diario.

Tal como está planteado, el modelo horario permite estimar la distribución de un consumo diario en las 24 horas del día. El modelo horario consiste, en realidad, en 24 modelos lineales que estiman, cada uno, el consumo horario relativo, es decir, la relación entre el consumo horario de una hora y el consumo horario medio del día.

Según se ha definido en el apartado anterior, los predictores principales en el nivel horario son las variables de festividad **FFA** y **FFM** y de Semana Santa, **SS**. Además, se han incluido las variables de festividad afectadas por funciones de estacionalidad con vistas a permitir un tratamiento diferenciado en función de la época del año y poder representar así la estacionalidad de consumo (predictores **FDEc** y **SEMEc**). Asimismo, el nivel horario incorpora una variable de calendario nueva: **AN**, que toma valor 1 el primer día de enero, y 0 el resto del año.

El modelo horario se formula en términos estrictamente lineales sobre los predictores enumerados en el párrafo anterior. El modelo para un sector consta de 24 ecuaciones, cada una de las cuales proporciona la estimación de consumo horario relativo para una hora del día, formuladas como se indica seguidamente.

$$\begin{aligned}
\gamma_{i1}^* &= a_{01} + a_{11} \cdot FFA_i + a_{21} \cdot FFM_i + a_{31} \cdot FDEC_i + a_{41} \cdot SEMEC_i + a_{51} \cdot AN_i \\
&\quad \dots \\
\gamma_{ih}^* &= a_{0h} + a_{1h} \cdot FFA_i + a_{2h} \cdot FFM_i + a_{3h} \cdot FDEC_i + a_{4h} \cdot SEMEC_i + a_{5h} \cdot AN_i \\
&\quad \dots \\
\gamma_{i,23}^* &= a_{0,23} + a_{1,23} \cdot FFA_i + a_{2,23} \cdot FFM_i + a_{3,23} \cdot FDEC_i + a_{4,23} \cdot SEMEC_i + a_{5,23} \cdot AN_i \\
\gamma_{i,24}^* &= 24 - \sum_{h=1}^{23} \gamma_{ih}
\end{aligned}$$

Donde \* indica valor estimado, **i** indica el día **i**–ésimo, **h** indica la hora del día y cada **a<sub>jh</sub>** es el coeficiente del modelo para el predictor **j**–ésimo en la hora **h** (**a<sub>0h</sub>** se refiere a la constante del modelo), y donde algunos **a<sub>j</sub>** pueden ser 0. Obsérvese que la selección de predictores se hace por sectores y afecta por igual a las ecuaciones de todas las horas. Es decir, cuando un **a<sub>jh</sub>** es 0 para alguna de las 23 primeras ecuaciones, lo es para todas las demás, de modo que el predictor **j**–ésimo **no interviene en el modelo de ese sector**.

Por construcción, la suma de los 24 consumos horarios relativos es 24, lo que permite la expresión compacta para la ecuación correspondiente a **h=24**. Tratándose de modelos de regresión lineal por mínimos cuadrados con las mismas variables independientes, la restricción de que deben sumar 24 determina los coeficientes de una de las ecuaciones a partir de las otras 23, de modo que no es necesario calcular el último modelo de regresión.





# 3

Tareas de modelización y  
mantenimiento

## 3.1. Introducción

Una vez definida la formulación en la que se basa la modelización del consumo, la tarea siguiente consiste en el ajuste de los parámetros que multiplican a cada uno de los predictores considerados, tanto en el modelo diario como en el horario. Este ajuste se realiza por el criterio de mínimos cuadrados, confrontando los valores estimados por el modelo con los correspondientes a los consumos horarios pertenecientes a la serie histórica que se utiliza como periodo de referencia.

Como paso previo, es necesario realizar una revisión de la serie histórica que permita descartar aquellos datos para los que existe una elevada probabilidad de que no sean representativos del consumo registrado en el sector considerado.

La tarea de desarrollo de los modelos es la antesala de las fases de explotación y validación de la propia modelización. La validación consiste en comparar los datos de consumo registrados en un determinado sector con las estimaciones de caudal que proporciona el modelo matemático de manera independiente al registro observado. Si de la comparación entre el consumo medido y el valor esperable se concluye que la diferencia no es excesiva, el dato es validado. En caso contrario se rechaza.

La base de la validación es, por tanto, la comparación del dato medido con el valor estimado por el modelo.

En los sucesivos apartados se describe, con algo más de detalle, el planteamiento teórico en el que se basa el desarrollo de cada una de estas actividades.

## 3.2. Desarrollo del modelo

### 3.2.1. Selección del conjunto de referencia

Los datos centrales del análisis estuvieron constituidos, lógicamente, por las series de consumo horario por sectores, residentes en el sistema de base de datos de Canal de Isabel II. Con todo, de manera complementaria, se utilizó además, esta otra información adicional:

- Datos de caracterización de sectores.
- Informes de sectorización.
- Datos de facturación.

El conjunto de referencia es un subconjunto del total de datos que se considera representativo del comportamiento del consumo del sector. Se obtiene por eliminación, descartando los datos que se consideran no válidos. Todas las modelizaciones se intentaron realizar contrastándolas con series históricas de, al menos un año de longitud, de cara a poder evaluar las diferencias existentes entre el consumo de verano e invierno y poder así estimar una curva de estacionalidad.

Como es evidente, el resultado de la modelización que se pueda llegar a obtener depende en buena medida de la calidad de los datos que constituyen el conjunto de referencia. La presencia de registros fuera de rango, de lagunas o de cambios bruscos, que responden a causas ajenas a las variaciones reales de consumo puede enmascarar las relaciones reales entre variables y, consecuentemente, dificultar la tarea de modelización.

Sin embargo, la frontera entre los datos desestimables y los datos simplemente incómodos es siempre indefinida. En cualquier tarea de depuración de series se corre el riesgo de eliminar datos representativos de la realidad que se desea modelizar (es más, datos representativos de los extremos de dicha realidad) y perder así información valiosa.

Por la necesidad de adoptar una posición de equilibrio entre estos argumentos contrapuestos, se consideró que la depuración de datos no podía basarse en criterios puramente objetivos. Al margen de estos razonamientos ocurría también que, en sectores de reciente incorporación, introducir unas holguras considerables en la fase de depuración era obligado si se quería tener una serie de longitud suficiente como para poder acometer la modelización, circunstancia que favorecía la adopción de una depuración manual o, al menos, no completamente automatizada.

En general, debe destacarse que el proceso de depuración se llevó a cabo siguiendo criterios conservadores, aun asumiendo que esto implicaba aceptar ciertas irregularidades en los conjuntos de referencia que muy probablemente no reflejaban un comportamiento real de las unidades de consumo y que, como consecuencia, podían derivar en una modelización de peor calidad.

Los criterios seguidos han sido suficientemente laxos como para asegurar que la probabilidad de rechazar datos válidos es mínima. En cualquier caso, es también recomendable modular estos criterios a medida que se adquiera un mayor conocimiento sobre el comportamiento de cada sector, de manera que de forma progresiva se disponga de procesos de depuración más ajustados a la naturaleza intrínseca de cada unidad de consumo.

Según el procedimiento finalmente adoptado, el proceso de depuración de datos de partida quedó dividido en dos fases:

- Depuración automática.
- Depuración manual (o subjetiva).

En los siguientes apartados se explica el alcance de cada una de estas tareas.

#### 3.2.1.1. Depuración automática

La serie de datos de consumo horario de cada sector analizado se sometió a un procedimiento de depuración automático. Esta depuración tuvo por objeto eliminar de la serie los valores que claramente estuvieran fuera del rango de validez.

La depuración de las series se aplicó exclusivamente a los días con datos completos y no todos iguales a 0. Es decir, los días en los que faltaba algún dato horario o en los que el consumo diario era igual a 0 se rechazaban de manera automática.

Como primer criterio se impuso, para los sectores correspondientes a núcleos urbanos, que ningún consumo horario ( $c_{ih}$ , consumo de la hora  $h$  correspondiente al día  $i$ ), expresado en metros cúbicos por segundo, pudiera ser superior a la suma de las viviendas del sector por un caudal de un litro por segundo.

$$C_{ih} < 0,001 * \sum \text{viviendas}$$

El segundo criterio de depuración buscaba rechazar aquellos consumos horarios claramente alejados del resto de valores del conjunto de referencia pertenecientes a la misma franja horaria. Como paso previo se calculaban, para cada hora del día, el promedio de los consumos horarios ( $\lambda_h$ ) y la desviación típica de la muestra ( $\sigma_h^2$ ).

A partir de estos valores de referencia, el rechazo se producía cuando la diferencia entre el consumo promedio ( $\lambda_h$ ) y el evaluado ( $C_{ih}$ ) superaba un número determinado de desviaciones típicas (valor umbral).

Resultaba forzosamente necesario establecer un valor holgado para este umbral ya que, además de la inevitable variabilidad aleatoria, debía considerarse también la existencia de una variabilidad semanal y estacional que, con frecuencia, se caracterizaba por ser notablemente elevada. De inicio, este umbral fue fijado en 5 de manera empírica si bien, en sectores con series de datos de reducida longitud fue necesario aumentarlo con el fin de alcanzar un periodo de referencia con un número de registros suficiente para abordar la modelización.

Planteado como fórmula, el requisito que se le imponía a cada consumo horario  $C_{ih}$  para no ser descartado del conjunto de referencia era el siguiente:

Donde:

$$\frac{(\lambda_h - c_{ih})^2}{\sigma_h^2} < 5$$

$$\lambda_h = \frac{\sum_{i=1}^{i=n} c_{ih}}{n}$$

$$\sigma_h^2 = \frac{\sum_{i=1}^{i=n} (\lambda_h - c_{ih})^2}{n}$$

con  $C_{ih} < 0,001 * \sum \text{viviendas}$

### 3.2.1.2. Depuración manual

Una vez realizada la depuración automática, los datos no descartados se sometían a un segundo proceso de depuración, esta vez de carácter subjetivo, si bien se apoyaba en indicadores objetivos. Esta tarea se realizó a dos niveles que se describen a continuación.

### Distribución diaria

El procedimiento consistió en representar los promedios de consumo diario de la serie histórica en forma de gráfico y, a la vista de este, rechazar aquellos periodos en los que, "de visu", se observaba un comportamiento anómalo por ser mucho más altos o más bajos que los inmediatamente anteriores o posteriores, por no seguir los patrones de estacionalidad observados, etc.

### Distribución horaria

La depuración horaria se realizaba una vez acometido el análisis a nivel diario. Se perseguía identificar aquellos días cuyo patrón intradiario fuese claramente descartable, aun en el caso de que el promedio diario estuviese dentro de los límites establecidos.

Como baremo de la separación de la distribución horaria respecto del patrón habitual del sector (constituido por el resto de valores aceptados en el conjunto de referencia), se utilizó la distancia cuadrática a la media de consumos horarios relativos. Es decir, como indicador se calculaba la distancia cuadrática media entre los consumos horarios de un día determinado, con respecto a los del consumo medio calculado, hora a hora, para todo el conjunto de la serie.

Tratándose de una depuración subjetiva, además de obtener este valor numérico se desarrolló una aplicación que permitiera representar, de manera automática en un gráfico, aquellos días de consumos horarios que más se alejaban de la distribución intradiaria media junto con la curva correspondiente al patrón horario medio de la serie. Para facilitar la comparación, se optó por representar dos curvas de distribución intradiaria típica, una correspondiente a días festivos y la otra a laborables. A la vista de la anomalía existente, tanto a nivel numérico como de diferencia entre la evolución horaria de las curvas, se decidía si el día en cuestión debía o no ser descartado. Obviamente, la decisión de en qué momento la diferencia entre los patrones horarios era lo suficientemente pequeña como para parar el proceso de eliminación resulta completamente subjetiva.

## 3.2.2. Ajuste del modelo

Una vez conformado el conjunto de referencia, el siguiente paso consiste en ajustar los valores de los parámetros del modelo. Para la realización de esta tarea se distinguen tres etapas independientes:

- Cálculo (o asignación) de las curvas de estacionalidad.
- Cálculo de coeficientes del modelo diario.
- Cálculo de coeficientes del modelo horario.

A continuación se desarrolla el alcance de cada uno de estos pasos. Se destaca que necesariamente han de realizarse según este orden, ya que cada fase de modelización se apoya en los resultados obtenidos en la anterior etapa.

### 3.2.2.1. Determinación de las curvas de estacionalidad

Como ya se ha puesto de manifiesto en los apartados 2.3. y 2.4, de acuerdo con la formulación planteada tanto en el modelo diario como en el horario, existen predictores que están afectados por los coeficientes de estacionalidad  $e_f$  y  $e_c$ . Por lo tanto, antes del ajuste de los parámetros es necesario tener definidas estas

curvas de estacionalidad. Esta tarea se realiza a partir de los datos de consumo medidos que constituyen el conjunto de referencia, de acuerdo con el procedimiento explicado en el apartado 2.3.2.

Alternativamente, para un sector en el que, debido a problemas en la calidad o cantidad de los datos de referencia, no pudiera determinarse una curva de estacionalidad, se contempló en la modelización la posibilidad de aprovechar una curva de "estacionalidad tipo" correspondiente a un sector de similares características y tipología.

Una vez definido el valor que toman estos factores de estacionalidad para cada día del año, es posible calcular el valor de cada uno de los restantes parámetros que intervienen en la fase de ajuste del modelo.

### 3.2.2.2. Cálculo de los coeficientes del modelo diario

Para cada sector, el modelo lineal se ajustó con el criterio de mínimos cuadrados a la serie de consumos de referencia, utilizando como variables predictoras todas las que intervienen en la formulación del modelo (tendencia, estacionalidad, calendario y meteorología). El método de ajuste seguido fue el de pasos sucesivos, por lo que siempre podía darse la circunstancia de que alguno de los predictores pudiese quedar fuera.

Al margen, se admitió la posibilidad de que manualmente se pudiera imponer o restringir la incorporación de alguno de los predictores en la modelización. De esta manera, resultaba posible valorar la incidencia exclusiva de alguna de las variables (por ejemplo la meteorología) en el grado de ajuste del modelo obtenido.

### 3.2.2.3. Cálculo de los coeficientes del modelo horario

La determinación del modelo horario de cada sector consiste, evidentemente, en la determinación de los coeficientes que multiplican a los predictores de las primeras 23 ecuaciones horarias. En el proceso de determinación de los coeficientes existe un paso previo de selección de predictores del conjunto de predictores potenciales. Ambas tareas, selección de predictores y establecimiento de los valores de los coeficientes, se han realizado simultáneamente siguiendo el proceso por aproximaciones sucesivas que se describe a continuación:

1. Se ajustan las 23 ecuaciones de regresión lineal con todos los predictores potenciales.
2. Se recopilan los valores de significación estadística de los coeficientes de cada uno de los predictores.
3. Se rechaza el predictor que no participe de forma significativa (valores de significación en el orden de 0, inferiores a 0,05) en un número elevado de ecuaciones (del orden de 8 o más).
4. Si no hay motivos para rechazar ninguno de los predictores, se da por válido el modelo y finaliza el proceso.
5. En caso contrario, se ajustan las 23 ecuaciones de regresión lineal con el nuevo conjunto, más restringido, de predictores y se vuelve al paso 2.

Al margen del nivel de significación que le pudiese corresponder a cada predictor en el proceso de ajuste, de manera manual pudo imponerse la aceptación o rechazo de una determinada variable. En este sentido, se advierte que se ha sido algo benevolente con la variable **AN** (Año Nuevo), por cuanto, en general, sólo resulta determinante en las primeras horas del día, y rara vez superaría un criterio prefijado de significación en un número elevado de horas.

### 3.3. Validación y estimación

#### 3.3.1. Planteamiento

Una vez definidos los parámetros de un modelo de consumo, éstos quedan incorporados en la aplicación informática CHYPRE. Con esta herramienta se lleva a cabo la utilización operativa de los modelos y la gestión, tanto de los datos de consumo observados, como de los estimados por el modelo.

Como tarea primera se lleva a cabo la confrontación de los datos de consumo medidos con los estimados por el modelo. A este proceso se le denomina validación de los datos. En términos generales, si la diferencia entre observación y estimación está dentro de un intervalo de aceptación, el dato se da por validado y, en caso contrario, se rechaza. Conviene destacar que, en la estimación, el valor proporcionado por el modelo de consumo es totalmente independiente del dato medido.

Como es evidente, resulta inevitable que exista divergencia entre los caudales realmente medidos y los valores proporcionados por el modelo. Sin embargo, un modelo correctamente formulado y ajustado, debería ser capaz de responder con una notable precisión a la realidad que pretende medir. De esta confrontación de resultados medidos y estimados pueden plantearse tres posibles situaciones:

- Existe suficiente coherencia entre el dato medido y la estimación realizada, por lo que se dan por válidos ambos valores.
- Las lecturas de consumo o bien no existen o son anómalas y, en cualquier caso, están muy alejadas de los valores estimados por el modelo. En esta situación sería factible la eliminación del registro medido de la serie de consumos, siendo sustituido por el caudal obtenido de la modelización.
- La serie de consumos medidos muestra un comportamiento coherente. Sin embargo, las estimaciones del modelo se alejan de este patrón de comportamiento. Este comportamiento podría obedecer a una modificación en las unidades de consumo del sector debidas a maniobras realizadas en la compartimentación de la red. La consecuencia directa de esta situación sería el rechazo del modelo que debería ser sometido de nuevo a un proceso de ajuste, ampliando en la medida de lo posible el conjunto de referencia de datos históricos.

A la hora de valorar una modelización, se entiende que son datos válidos aquellos que se acercan con cierta precisión al valor real que pretenden estimar. En este sentido, verificar que un dato medido es válido es virtualmente imposible. La validación pretende, únicamente, identificar aquellos registros para los que la probabilidad de que no sean válidos es suficientemente alta como para que resulte más conveniente desestimarlos.

Inciendo en este aspecto, la aparición de datos de consumo incoherentes y muy alejados del consumo esperable no necesariamente debe asociarse a un mal funcionamiento de los aparatos de medida; también podría ser debida a incidencias que tengan lugar en la red de distribución (fugas, maniobras, etc.). En este sentido, las modelizaciones efectuadas juegan un papel relevante a la hora de identificar estas situaciones anómalas.

El rechazo de un dato genera una laguna de validación que se rellenará con el dato estimado correspondiente. Las lagunas de observación se rellenan igualmente con datos estimados por el modelo. Así pues, parece natural que los procesos de validación vayan acompañados de los procesos de relleno de lagunas. En lo que sigue, "proceso de validación" o simplemente "validación" se referirá al proceso conjunto de validación y relleno

de lagunas, salvo que se indique lo contrario. Además, el término "laguna" se referirá, tanto a las lagunas de observación, como a las de validación.

A pesar de la sencillez del planteamiento general, el proceso de validación es bastante más complejo que el desarrollo propuesto para la obtención de los parámetros del modelo. Esto se debe a la existencia de distintos niveles temporales de variación de los datos que han de ser contrastados: intradiario, diario, semanal y estacional.

Esta circunstancia ha conducido a la utilización de un modelo estructurado en 2 niveles de resolución temporal: el nivel diario, que estima la magnitud del consumo en el día, y el nivel horario, que estima el patrón de distribución horaria del consumo diario. La forma directa de aplicación del modelo consiste en estimar primero el consumo diario y, a partir de él, los 24 consumos horarios.

Teniendo presente lo anterior, podríamos considerar los siguientes casos:

- a) En un día determinado se ha producido una laguna en un dato horario, pero los 23 datos horarios restantes existen y, además, resultan coherentes con el patrón de distribución horaria esperable.
- b) Existen los 24 datos horarios de un día determinado, pero no son coherentes con el patrón de distribución horaria esperable.

En relación con el nivel diario, el caso (a) representa una situación de laguna: la falta de un dato horario impide determinar el consumo diario. La aproximación directa consistiría en estimar primero el dato diario con el modelo diario y, a partir de él, el consumo horario que falta. Procediendo de esta forma, sin embargo, se estaría desperdiciando la información que sobre el nivel diario proporcionan los 23 datos horarios disponibles. La estimación del consumo diario que puede obtenerse a partir de 23 datos horarios puede ser mucho mejor que la que proporciona el modelo diario.

En el caso (b) puede determinarse el consumo diario a partir de los datos y este consumo puede validarse mediante el modelo diario. Sin embargo, la incoherencia de la distribución horaria debe hacernos sospechar de la validez de dicho dato diario. Parece más adecuado rechazar los 24 datos horarios y no utilizar el dato diario correspondiente.

Se pone de manifiesto, por tanto, que el proceso de validación debe ser más sofisticado que la aplicación directa del modelo. En particular, se plantea claramente la necesidad de estructurar el proceso de validación en dos etapas de acuerdo con los niveles de resolución temporal adaptados para la modelización:

- Coherencia de los datos horarios con el patrón de distribución horaria esperable.
- Coherencia del dato diario con el nivel de consumo diario esperable.

Como es evidente, el proceso de validación es totalmente dependiente de la etapa de modelización. La meta de la validación es la de funcionar automáticamente y rechazar los datos que parecen malos remplazándolos por datos modelizados. Para ello, es condición necesaria que los resultados que el modelo proporcione sean, al menos, verosímiles.

### 3.3.2. Proceso de validación

Con el planteamiento anterior, la unidad de validación debe ser el conjunto de datos horarios de consumo en un día para un sector determinado. Es decir, el proceso de validación unitario comprende el conjunto de operaciones que se aplican sobre todos los datos horarios disponibles para un día y un sector dados, y que producen como resultado 24 datos de consumo horario validados o, en su caso, estimados.

Por simplicidad de notación, supondremos que los datos de consumo horario de un día y un sector se han copiado en un vector de 24 componentes, **CH** [1...24], en el que se utiliza un valor "Nulo" para representar las eventuales lagunas. El proceso de validación se describirá en términos de las operaciones que se realizan sobre dicho vector. En todo momento será posible diferenciar qué datos son originales (observados) y cuáles no (estimados).

De acuerdo con lo referido en el apartado anterior, el proceso de validación se basa en 2 operaciones o subprocesos distintos:

- Estimación por la distribución horaria. Esta operación consiste en estimar uno o varios datos horarios a partir del resto de datos horarios del día. Las estimaciones persiguen que la distribución horaria final sea coherente con la predicha por el modelo horario. La estimación es independiente del modelo diario.
- Estimación por el modelo. Consiste en estimar los 24 datos horarios a partir del modelo en 2 etapas: estimación del consumo medio diario y, a partir de él, la estimación de su distribución horaria. A diferencia de la anterior, esta operación utiliza los dos niveles del modelo, es decir, el diario y el horario.

El proceso de validación que utiliza la aplicación CHYPRE sigue el esquema siguiente:

1. Los valores negativos de **CH** se consideran "Nulo". Esto es, cualquier dato de consumo observado negativo se considera no válido y genera una laguna de validación.
2. Si hay menos de seis datos originales, se salta directamente al paso número 6 (se eliminan todos los registros originales y se hace la estimación completa del modelo diario y horario). Se considera que 5 ó menos datos horarios no proporcionan ninguna información relevante sobre el consumo diario.
3. Si hay valores "Nulo", se remplazan por sus correspondientes estimaciones por la distribución horaria.
4. Se comprueba la coherencia de los datos originales que quedan en **CH** con la distribución horaria estimada por el modelo (horario).

Si hay alguna discrepancia no aceptable:

- a. Se devuelve **CH** a los valores que tenía antes de empezar el paso 3.
- b. Se pone a "Nulo" el valor correspondiente a la discrepancia más significativa.
- c. Se salta al paso 2.

En caso contrario (no hay discrepancias significativas en ningún dato horario) se continua en el paso 5.

5. Se comprueba la coherencia del dato de consumo medio diario, **Cmd** con el modelo diario. **Cmd** es la suma de los 24 valores de consumo (originales o estimados; nótese que en este estadio, no hay valores "Nulo" en **CH**). Si se da por válido el valor de **Cmd**, se salta a 7.
6. Se remplazan los 24 componentes de **CH** por sus valores estimados por el modelo. Para ello se realiza la estimación completa en 2 fases: modelo diario y modelo horario.
7. Final:
  - Los datos originales que quedan en **CH** son datos validados.
  - Los datos originales que había en **CH** antes de iniciar el paso 1, y que ya no están en **CH**, son datos rechazados.
  - Los datos no originales de **CH** son datos estimados, bien por rechazo de datos (laguna de validación), bien por relleno de laguna de observación.

Los pasos 4 y 5 del proceso requieren un conjunto de valores umbral para definir los rangos de aceptación de los datos. Es decir, la validación está parametrizada o gobernada por un conjunto de parámetros que deben definirse a priori.

Esta parametrización permite, por otra parte, modular la validación, haciéndola más o menos restrictiva. De este modo, el proceso de validación podrá ajustarse en función del criterio de los usuarios del modelo, a medida que vayan adquiriendo experiencia sobre su comportamiento.

En los apartados siguientes se describen con un mayor detalle las operaciones enumeradas en los puntos anteriores que se ha considerado que entrañan algo más de complejidad en su planteamiento. Son las siguientes:

- Sustitución de valores "Nulo" por estimaciones de la distribución horaria del modelo (paso 3 en la relación anterior).
- Comprobación de la coherencia de los datos originales con la distribución horaria esperable (la estimada por el modelo). Se corresponde con el paso 4 anterior.
- Comprobación de la coherencia del dato de consumo diario con el valor proporcionado por el modelo (paso 5 anterior).

### 3.3.3. Estimación por la distribución horaria

El objetivo de esta operación consiste en poder completar las lagunas de observación de consumo horario que puedan existir en los datos de consumo de un día determinado, en un sector cualquiera. Como ya se ha expuesto en el apartado anterior, esta tarea se intenta llevar a cabo en el caso de existir un mínimo de seis datos de consumo horario originales no descartables.

Con menos de 6 datos originales, se considera que la información aportada es tan escasa que está justificado eliminar estos registros y remplazarlos por la estimación completa de caudal diario y horario que proporciona el modelo.

Como ya es sabido, el modelo de distribución horaria de consumo proporciona los valores de los 24 consumos horarios relativos,  $\gamma_h$ , que son los cocientes entre el consumo horario y el consumo medio diario **Cmd**. Cada uno de estos  $\gamma_h$  se obtiene a su vez como función lineal de un conjunto de predictores. Además, por definición, la suma de los 24 consumos horarios relativos ha de ser 24. Como recordatorio de lo anterior, se incluye nuevamente la formulación recogida en el apartado 2.4.4.

$$\begin{aligned}\gamma_{i1}^* &= a_{01} + a_{11} \cdot FFA_i + a_{21} \cdot FFM_i + a_{31} \cdot FDEC_i + a_{41} \cdot SEMEC_i + a_{51} \cdot AN_i \\ &\dots \\ \gamma_{ih}^* &= a_{0h} + a_{1h} \cdot FFA_i + a_{2h} \cdot FFM_i + a_{3h} \cdot FDEC_i + a_{4h} \cdot SEMEC_i + a_{5h} \cdot AN_i \\ &\dots \\ \gamma_{i,23}^* &= a_{0,23} + a_{1,23} \cdot FFA_i + a_{2,23} \cdot FFM_i + a_{3,23} \cdot FDEC_i + a_{4,23} \cdot SEMEC_i + a_{5,23} \cdot AN_i \\ \gamma_{i,24}^* &= 24 - \sum_{h=1}^{23} \gamma_{ih}^*\end{aligned}$$

Se parte del supuesto de que en los datos de partida del consumo horario para un día determinado hay **i** componentes nulos (obviamente, con  $i < 24$ ). Por simplificar la notación y sin pérdida de generalidad se acepta que estos sean los componentes **1... i**. Estos **i** consumos horarios a estimar ( $c_1^*, c_2^*, \dots, c_i^*$ ) deberán verificar que:

$$\begin{aligned}24 \cdot Cmd^* &= c_1^* + \dots + c_i^* + c_{i+1} + \dots + c_{24} \\ &\dots \\ \gamma_h^* &= a_{0h} + a_{1h} \cdot FFA + a_{2h} \cdot FFM + a_{3h} \cdot FDEC + a_{4h} \cdot SEMEC + a_{5h} \cdot AN \quad \text{para todo } h \text{ en } \{1..i\} \\ &\dots \\ c_h^* &= \gamma_h^* \cdot Cmd^* \quad \text{para todo } h \text{ en } \{1..i\}\end{aligned}$$

De las tres ecuaciones anteriores es sencillo despejar el valor de **Cmd\***. Conocido **Cmd\*** y los valores de los predictores pueden determinarse los  $\gamma_h^*$  utilizando las ecuaciones del modelo horario.

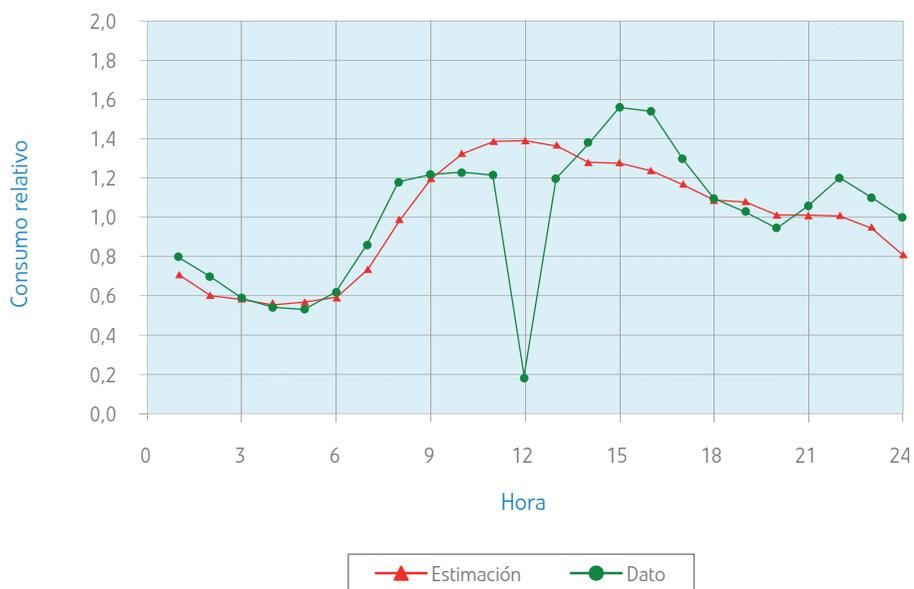
Por último, conocido los  $\gamma_h^*$  y los **Cmd\*** podrán determinarse los **c<sub>h</sub>\*** que deberán rellenar las lagunas de observación existentes.

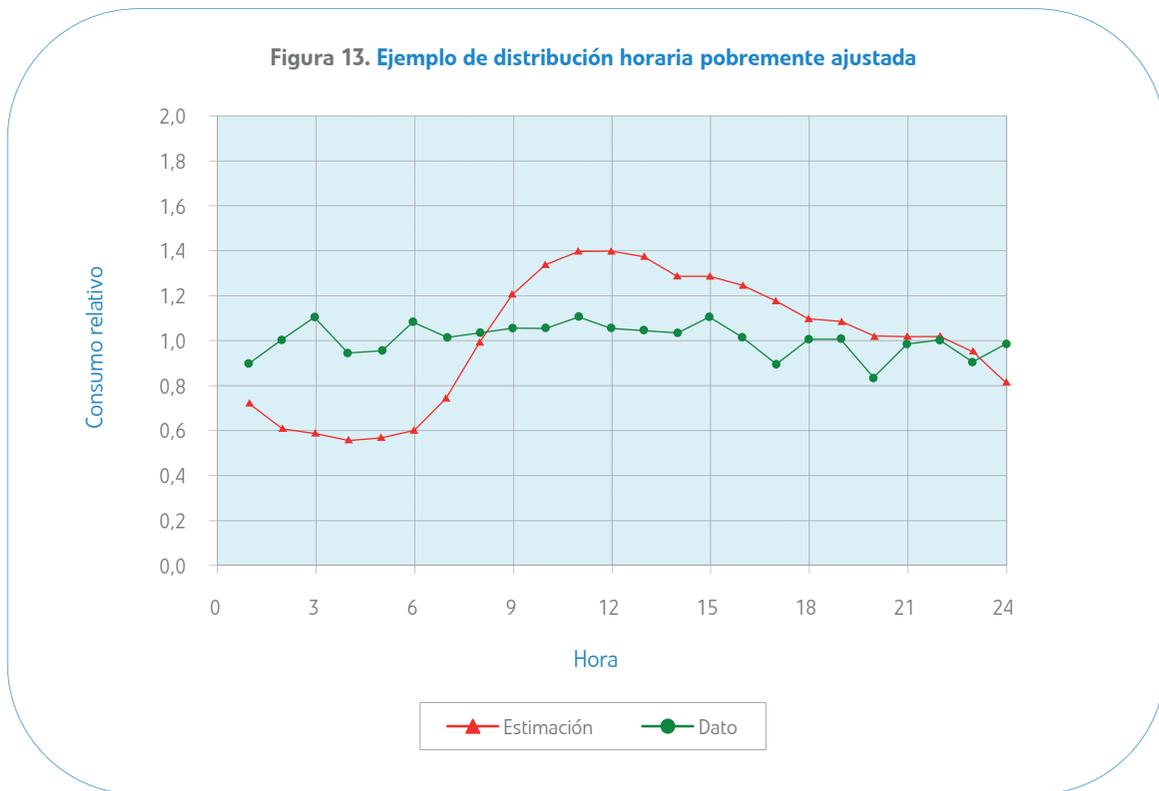
### 3.3.4. Comprobación de la coherencia de la distribución horaria

Resultaba necesario implantar un procedimiento que permitiera comprobar cómo de parecida podía resultar una distribución horaria de consumos en un día determinado en relación con la distribución horaria predicha por el modelo.

El método finalmente adoptado se basa en la aplicación de dos tipos de criterios. El primero de los criterios es una evaluación general de la distribución horaria. El segundo determina la validez de cada uno de los datos horarios individualmente. Con este criterio doble se pretende identificar tanto distribuciones horarias pobremente ajustadas al patrón esperado, pero sin ningún dato singular excepcionalmente malo, como distribuciones buenas, en general, pero con uno o pocos datos muy anómalos (el caso típico es el de una subida o una bajada brusca en una hora aislada). En las figuras 12 y 13 se presentan ejemplos de estas dos situaciones.

**Figura 12. Ejemplo de distribución horaria con un dato claramente anómalo**





El indicador que se utiliza para valorar el primer criterio es:

$$S^2 = \sum_{h=1}^{24} (\gamma_h - \gamma_h^*)^2$$

donde  $c_h^* = \gamma_h^* \cdot Cmd$

El criterio de aceptación será la no superación de un valor umbral para  $S^2$ ,  $S^2_{\max}$ , que puede determinarse en función de la particularidad de cada sector modelizado. Como criterio general, de inicio se ha establecido como límite aquel valor que deja fuera el uno por ciento de los datos de referencia, si bien en la fase de explotación de los modelos este valor puede corregirse al alza o a la baja en función del comportamiento observado.

Este parámetro de la bondad del patrón horario es el mismo que se empleó a la hora de establecer los criterios de validez de los datos que deben conformar el periodo de referencia a partir del cual realizar la modelización (ver apartado 3.2.1).

Debe destacarse que en esta fase se obvia la distinción entre los consumos horarios originales y los consumos estimados ( $C_h$  y  $C_h^*$ ).  $Cmd$  es la media de los 24 consumos horarios considerados en el día, independientemente de su origen.

Para el segundo criterio, el valor que se utiliza para definir el intervalo de aceptación de cada dato horario es:

$$\delta_h = |\gamma_h - \gamma_h^*| / \sigma_h$$

Donde  $\sigma_h$  es un parámetro cuyo valor depende del sector y de la hora de que se trata. Este parámetro es la desviación estándar de las diferencias  $\gamma_h - \gamma_h^*$  observadas en los datos de referencia.

Para  $\delta_h$  se define un valor umbral,  $\delta_{\max h1}$ . Si  $\delta_h \leq \delta_{\max h1}$ , el dato se considera válido.

El proceso de validación de la distribución horaria se realiza en dos etapas sucesivas, la segunda de las cuales, a su vez, consiste en una sucesión de pasos.

#### Etapa primera: validación general de la distribución

Consiste en aplicar los dos criterios descritos para verificar la validez de la distribución horaria en su conjunto. El proceso es el siguiente:

- a) Se calculan los 24 valores de  $\gamma_h$  y  $\gamma_h^*$  (consumo horario relativo observado y estimado, respectivamente).
- b) Se calcula el valor  $S^2$ . Si este valor supera el umbral de aceptación  $S^2_{\max}$ , es necesario reestimar el vector **CH**. En caso contrario, se continúa en (c).
- c) Se calculan los 24 valores de  $\delta_h$  y se localiza el máximo.
- d) Si el máximo  $\delta_h$  no supera el umbral de aceptación,  $\delta_{\max h1}$ , se acepta la distribución horaria y se continúa en la etapa segunda. En caso contrario, la prueba de validación no es superada y es necesario volver a estimar el vector **CH**.

#### Etapa segunda: volver a estimar el vector CH

Si no supera la prueba anterior, se deben identificar los datos no válidos y sustituirlos por estimaciones. Este proceso se realiza en etapas sucesivas, rechazando cada vez un único dato: el que presenta la mayor discrepancia. Es necesario hacerlo así porque el rechazo de un dato modifica el **Cmd** y, por tanto, los valores de  $\gamma_h$  y  $\gamma_h^*$  del resto de horas.

En esta fase se prescinde del criterio de  $S^2$ : ya no se trata de evaluar la distribución horaria en su conjunto sino de identificar los datos que pueden considerarse anómalos y sustituirlos. Además, se utiliza un umbral más restrictivo para  $\delta_h$ ,  $\delta_{\max h2}$  ( $\delta_{\max h1} > \delta_{\max h2}$ ). Se ha comprobado que produce mejores resultados aplicar un criterio más laxo para identificar distribuciones horarias anómalas y otro más estricto a la hora de depurarlas.

Expresado en forma esquemática, el proceso seguido en la validación de los consumos horarios es el siguiente:

- a) Se calculan los 24 valores de  $\gamma_h$  y  $\gamma_h^*$  (consumo horario relativo observado y estimado, respectivamente).
- b) Se calculan los 24 valores de  $\delta_h$  y se localiza el máximo.
- c) Si el máximo  $\delta_h$  no supera el umbral de aceptación,  $\delta_{\max h2}$ , se acepta la distribución horaria y se finaliza el paso 4. En caso contrario, se sigue el paso (d).
- d) Se devuelve **CH** a los valores que tenía antes de iniciar el paso 3 por última vez (es decir, con "Nulo" en las lagunas, los datos negativos y los datos rechazados en todas las comprobaciones de la distribución horaria anteriores).
- e) Se pone a "Nulo" el dato correspondiente a la discrepancia máxima, localizado en el paso (b).

Se vuelve al paso número 2 del proceso general de validación. Ahora, **CH** tendrá un valor nulo más que en la anterior llamada. La vuelta al paso 2 implica que si se acaba rechazando un número suficiente de datos, se da por "**no valida**" la totalidad de la distribución horaria debiendo ser remplazada por la estimación del modelo.

### 3.3.5. Comprobación de la coherencia del dato diario

El dato de **Cmd** debe ser coherente con el modelo diario. Esta coherencia se evalúa en términos de **cest\*** (consumo estimado antes de aplicar el factor de tendencia), mediante el indicador:

$$\delta = |Cmd/T - cest^*|/\sigma$$

donde **T** es el término de tendencia según el modelo diario (formulado en función de los valores de **cest** y de consumo vigente en los catorce días previos, tal y como se recoge en el apartado 2.3.5) y  $\sigma$  es un parámetro cuyo valor depende del sector. Este último parámetro es el valor de la desviación estándar de las diferencias (**Cmd/T – cest\***) de los datos de referencia.

Previamente se habrá definido un valor umbral para  $\delta$ ,  $\delta_{max}$ , cuya no superación supone la aceptación del dato de consumo diario como válido. Inicialmente, se adoptó el valor "3" para este umbral. Este valor es susceptible, como los anteriores, de ser modificado o ajustado según las particularidades de cada sector.

### 3.3.6. Parámetros de validación

Según ha quedado expuesto en los apartados anteriores, en los procesos de validación se utiliza un conjunto de valores umbral y otros parámetros que sirven para definir los rangos de aceptación de los datos de consumo. Es decir, la validación, tal como se ha definido, está parametrizada o gobernada por un conjunto de parámetros que deben definirse a priori.

Como ya se ha destacado, esta parametrización permite, por otra parte, modular la validación, haciéndola más o menos restrictiva. De este modo, el proceso de validación podrá ajustarse en función del criterio de los usuarios del modelo a medida que vayan adquiriendo experiencia sobre el comportamiento individual de cada sector.

A continuación se incluyen en un listado todos los parámetros de validación que se han considerado. Como ya se ha indicado, la determinación del valor asignado a cada uno de estos estadísticos depende de las particularidades de cada sector. Como norma general, los valores se han calculado a partir de las características estadísticas de los datos del periodo de referencia. En particular se ha utilizado la población de datos de los días completos (es decir, con 24 datos horarios) que no fueron desestimados durante la fase de análisis de las series de datos de consumo previa al desarrollo de los modelos.

### 3.3.6.1. Parámetros de validación del modelo horario

Se han considerado los parámetros que se describen a continuación.

**S<sup>2</sup><sub>max</sub>** Define un primer umbral de aceptación de una distribución horaria en su conjunto. Cuanto menor sea, más restrictivo es el criterio. Por defecto, se ha definido para cada sector aquel valor que deja fuera el uno por ciento de los datos de referencia. Sin embargo, durante la fase de mantenimiento, con frecuencia se ha efectuado la relajación de este criterio.

**σ<sub>h</sub>** Para cada sector y para cada hora se ha tomado la desviación típica de las diferencias  $\gamma_h - \gamma_h^*$  observadas en los datos de referencia. Se trata de asimilar, de manera informal, las diferencias  $\gamma_h - \gamma_h^*$  a errores normales, de modo que el parámetro **σ<sub>h</sub>** no hace sino tipificarlas. De este modo, se puede establecer de forma directa el valor de los umbrales **δ<sub>max</sub>**- siguiente:

**δ<sub>max h1</sub>** Inicialmente se estableció en "5" para todos los sectores y todas las horas. Este valor proporciona una horquilla amplia para los datos horarios. En términos de distribución normal (y con todas las reservas) equivale a adoptar un umbral de 5 sigmas. Se trata de identificar datos horarios individuales clara mente fuera de rango.

**δ<sub>max h2</sub>** Se ha establecido en "3" para todos los sectores y todas las horas. Este umbral se aplica a la identificación de los datos horarios que deben sustituirse por estimaciones, una vez se ha identificado la distribución horaria como anómala. Siguiendo con la referencia a la distribución normal, equivale a un umbral de 3 sigmas. Para evaluar la verdadera magnitud del umbral, téngase en cuenta que se aplica a grupos de 24 datos; sirva de referencia que un test de 24 pruebas independientes, todas ellas con un umbral de significación de 3 sigmas, excluyen a más del 6 por ciento de los casos válidos.

### 3.3.6.2. Parámetros de validación del modelo diario

Se han considerado dos parámetros:

**σ** Se ha tomado, para cada sector, el valor de la desviación estándar de las diferencias **Cmd/T – cest\*** de los datos diarios que constituyen el conjunto de referencia. De este modo, como en el caso del parámetro equivalente para la modelización horaria, se persigue tipificar dichas referencias.

**δ<sub>max d</sub>** Como criterio general se ha tomado el valor de "3" para todos los sectores. Echando mano de nuevo a la referencia a la distribución normal, se trata de un umbral equivalente a 3 sigmas.

## 3.4. Predicción

### 3.4.1. Definición

Se entiende por predicción el proceso de estimación de consumos futuros. Las predicciones no son otra cosa que estimaciones de datos de consumo futuros realizadas por el modelo. La única diferencia formal entre las predicciones y las estimaciones en sentido estricto es la información de partida que se utiliza para realizarla.

Así, las predicciones únicamente pueden utilizar como datos de partida, aparte de los predictores de estacionalidad y calendario, que son en todo momento anticipables, los consumos registrados en las fechas anteriores a aquella en la que se efectúa la predicción y los datos de predicción meteorológica disponibles, que se utilizarán para hacer una aproximación de los valores que tomarán los predictores de lluvia y temperatura para el día futuro en el que se quiere efectuar la predicción de consumo.

### 3.4.2. Proceso de cálculo

Las predicciones se calculan por aplicación directa del modelo en sus dos fases de resolución temporal. En primer lugar, se determina la predicción de consumo medio diario con el modelo diario; en segundo lugar, se determinan los consumos horarios con el dato de promedio diario estimado y el modelo horario. Sin embargo, las condiciones particulares en las que se establece la predicción obligan a realizar algunas matizaciones que guardan relación con el valor del factor corrector de tendencia descrito en el apartado 2.3.5 como parte de la formulación del modelo diario.

Al no disponer de los datos de consumo vigentes u observados para los días comprendidos entre aquel en el que se efectúa la predicción y el correspondiente al día futuro para el que se quiere aventurar el consumo, no se puede realizar el cálculo del factor de tendencia. Sirva como recordatorio que este factor de corrección recoge la discrepancia entre las estimaciones del modelo y los datos vigentes correspondientes a los 14 días anteriores al día para el que se estima. En el caso de predicciones de alcance superior a "1", no se podrá disponer de datos de consumo observado para al menos uno de los 14 días anteriores y, por lo tanto, no se podrá aplicar la fórmula anterior.

Para soslayar este problema se adopta el criterio de suponer un factor de tendencia constante e igual al correspondiente al alcance "1" para todos los alcances. Con este criterio, es necesario disponer de datos de consumo vigente y **cest** hasta el día anterior a la fecha para la que se quiere disponer de la predicción.

Como suele suceder con cualquier tipo de predicción, la calidad de la misma disminuye a medida que aumenta el alcance. En este caso, la disminución de la calidad se debe a:

- La calidad de las predicciones meteorológicas que alimentan el modelo disminuye.
- La información sobre la tendencia recogida se va degradando al referirse a datos cada vez más alejados de aquel para el que se establece la predicción.

El sistema de predicción de la aplicación CHYPRE se ha concebido para alcances de hasta 7 días.



# 4

Comportamiento observado  
en el periodo de mantenimiento

Las tareas de explotación y mantenimiento de las unidades de consumo sectorizadas tienen un doble objetivo. Por una parte, verificar que el comportamiento de cada modelo de consumo es correcto y que presenta capacidad para responder de manera adecuada a las evoluciones de la demanda de consumos que puedan producirse. Por otra, también deben servir para la evaluación de los datos de consumo registrados y, por extensión, como herramienta de apoyo para la detección de anomalías en el funcionamiento de la red de distribución y sus equipamientos (maniobras en la red, cierres de válvulas, fallos en la captación de datos por los sensores, existencia de fugas, etc.).

A falta de otra información, un porcentaje elevado de rechazos de datos de consumo durante la validación debería considerarse señal sintomática de la posible existencia de una incidencia en relación al comportamiento de la red de distribución. Sin embargo, para poder hacer esta afirmación, como condición necesaria deberá garantizarse que el modelo esté funcionando de manera correcta. En este sentido, la única forma de asegurar el comportamiento coherente de los modelos pasa por realizar un seguimiento continuado y en detalle de la respuesta que presenta a las sucesivas cargas de datos realizadas.

La utilización operativa de los modelos y las operaciones que sobre los datos de consumo horario se efectúan (estimación, rechazo, validación y predicción) así como otras funciones generales de gestión de dichos datos (carga de datos, visualización, generación de informes, etc.), se realizan a través de la herramienta informática CHYPRE. Sin embargo, pese a la ayuda que brinda en las tareas de mantenimiento de los modelos desarrollados, la intervención de usuarios familiarizados con la formulación desarrollada es insustituible, y no puede remplazarse por procedimientos automatizados, tanto a la hora de diagnosticar el mal comportamiento de un modelo como para establecer las medidas necesarias para su corrección.

En este sentido, cuando por los resultados observados en la carga y validación de datos de consumo se identifica un comportamiento anómalo en un determinado sector, a la hora de intervenir manualmente se han planteado tres posibles alternativas:

- Modificación de los parámetros de validación.
- Corrección del factor de tendencia.
- Recalibración completa del sector (nueva modelización).

En los siguientes apartados se describe el alcance de cada una de estas tareas, y se incorporan algunos ejemplos ilustrativos de los casos más sintomáticos de un mal comportamiento del modelo inicialmente desarrollado y de las medidas adoptadas para devolverlo a un nivel de ajuste adecuado.

#### **4.1. Modificación de los parámetros de validación**

Aun suponiendo que se pueda disponer de un conjunto de referencia de calidad suficiente como para elaborar una modelización ajustada, y aun aceptando que los datos de consumo entrantes durante la fase de mantenimiento no presenten anomalías evidentes, inevitablemente durante la fase de explotación existirán diferencias significativas entre las observaciones y las estimaciones.

Estas diferencias pueden deberse a varias causas:

- Incapacidad de la formulación adoptada para representar determinadas particularidades del comportamiento.
- Evolución de los consumos registrados a valores y comportamientos alejados de los que conformaron el periodo de referencia.
- Registros de variables meteorológicas extraordinarios o, al menos, muy distintos a los observados en las fechas sobre las que se definió el conjunto de referencia.

Estas diferencias son asumibles y, en la medida en que no se alejen en exceso respecto a las desviaciones que se observaron durante la fase de validación realizada sobre el conjunto de referencia, el modelo deberá aceptar estos datos como válidos y no realizar su remplazo.

Una modelización correctamente realizada y apoyada en unos parámetros de validación bien ajustados será capaz de discriminar, en condiciones normales, si un dato de consumo entrante puede ser dado por válido o debe ser rechazado por alejarse de lo esperable en una magnitud tal como para considerarlo no verosímil. En este sentido, un mal comportamiento del modelo puede manifestarse de dos maneras independientes.

- Validación generalizada por parte del modelo de datos de consumo que parecen no responder a un patrón coherente.
- Rechazo y reemplazo por estimaciones del modelo de datos de consumo de evolución aparentemente aceptables.

Las medidas adoptadas para corregir una de estas tipologías de fallo, inevitablemente, hacen el modelo más vulnerable al otro tipo de deficiencias. Resulta necesario adoptar una situación de equilibrio que, generalmente, sólo se obtiene si se efectúa un seguimiento prolongado del comportamiento según se van incorporando nuevos datos de consumo actualizados y se van modificando, en un sentido u otro, los parámetros de validación.

Tal y como se ha expuesto en el apartado anterior, por defecto, los parámetros de validación de inicio aceptados en cada modelo se determinaron exclusivamente a partir de la dispersión del conjunto de referencia y de las desviaciones que existieron entre las estimaciones del modelo y la muestra poblacional. En este sentido, es claro que si en una modelización se evidencia que la formulación no tiene capacidad para reproducir un determinado comportamiento, o si el conjunto de referencia presenta una excesiva heterogeneidad, el modelo obtenido arrastrará una tolerancia muy elevada.

De manera inversa, si la validación del modelo evidenció un muy buen grado de ajuste entre estimación y conjunto de referencia, la modelización será muy estricta y, en consecuencia, existirá más probabilidad de que se puedan llegar a rechazar datos sobre los que no pueda existir suficiente convencimiento acerca de la conveniencia de su eliminación.

Por estos motivos, como criterio general puede apuntarse que las modelizaciones de gran calidad adolecerán de ser excesivamente restrictivas mientras que en los ajustes pobres el principal fallo observado en las labores de seguimiento y mantenimiento será el acercamiento a un grado de tolerancia excesivo. Puede deducirse, por tanto, que una parte importante de las tareas de mantenimiento consistirá en la revisión de los parámetros de validación seleccionados.

#### 4.1.1. Modificación del coeficiente de validación del consumo diario

Como ya se explicó en el apartado 3.3, el proceso de validación actúa de manera independiente a dos escalas temporales: horaria y diaria. Como primera etapa, se realiza la validación sobre la distribución intradiaria a partir de los consumos horarios relativos. Una vez revisados los 24 componentes del consumo horario (mediante las correspondientes validaciones y, en el caso de producirse algún rechazo, el correspondiente reemplazo por el dato estimado), el valor del consumo diario obtenido es sometido a una nueva validación.

Al ser estos criterios independientes, una familia de datos horarios correspondientes a un día determinado puede ser rechazada en su totalidad en el caso de que el consumo medio diario se aleje lo suficiente del valor esperado aun cuando su patrón horario pueda presentar una evolución del todo coherente.

Según se expuso en el apartado 3.3.5, la validación del consumo diario va ligada a la evolución del factor corrector de tendencia. Este parámetro sirve para acercar los valores de consumo diario estimados por el modelo a los valores reales observados. Estrictamente, pretende dar respuesta a las evoluciones a largo plazo de los niveles de consumo diarios, variaciones que generalmente en la fase de modelización no llegan a hacerse perceptibles.

Sin embargo, no es infrecuente que los consumos medidos presenten oscilaciones bruscas que excedan los límites de respuesta de este factor corrector y que, como resultado, se sobrepasen los límites de tolerancia de la validación del modelo diario.

Es indudable que, en el caso de observarse un buen nivel de coherencia entre los componentes del consumo horario estimado y los datos medidos, el rechazo de todos estos valores debido a una discordancia en el valor de consumo medio diario contradice la percepción intuitiva de que las lecturas registradas pudieran ser válidas. En este sentido, en la fase de desarrollo del modelo, cabría la posibilidad de haber condicionado la eliminación de un registro diario a la bondad del ajuste del modelo horario. Sin embargo, de esta manera se restarían las ventajas que desde el punto de vista de la formulación presenta el poder hacer por separado el análisis de las dos escalas temporales.

Más allá de los criterios ligados exclusivamente a la formulación del modelo, prestando atención también a los aspectos relacionados con la explotación, un cambio de magnitud ostensible en los consumos medios diarios de un sector debe considerarse a priori como una anomalía de entidad suficiente como para que la herramienta de modelización forzosamente tenga que detectarlo. De no ser así, la aceptación implícita de esta situación podría hacer que quedaran inadvertidas para el explotador incidencias de relevancia, bien en la red de suministro o en el sistema de adquisición de datos.

Aunque la heterogeneidad de sectores es muy grande y las causas no siempre son conocidas, se pueden encontrar algunos condicionantes que pueden favorecer que se produzcan cambios de magnitud en el consumo diario sin que vayan asociados a variación alguna en el patrón de consumo horario relativo:

- Estacionalidad acentuada, no observada en la fase de modelización y sin repercusión significativa en el consumo horario.
- Incorporación/eliminación de unidades de consumo en el sector considerado (asociado generalmente a maniobras en la red de distribución).
- Aparición de fugas en algún punto del sector considerado.

En el caso de producirse alguna de estas circunstancias, sería razonable que los consumos variaran de escala sin variar su evolución horaria. En estos casos, las lecturas obtenidas serían correctas, en tanto en cuanto que ajustadas a la realidad, si bien el modelo las podría rechazar por quedar alejadas del consumo medio diario esperable.

Como ya se ha indicado, la no validación de los datos de consumo en esta situación puede ser ventajosa para el explotador, de cara a disponer de una alerta para una situación anómala. Con todo, el paso siguiente consistiría en contrastar a qué obedece el cambio de magnitud observado y, en el caso de descartar la posibilidad de que pudiera deberse a anomalías en las lecturas, proceder a corregir los parámetros de validación del modelo para que éste pueda responder de manera correcta a la nueva evolución del consumo.

Para lograr una mayor adecuación de las estimaciones del modelo a la evolución de los consumos observados, cabe la posibilidad de imponer la aceptación de un periodo de registros rechazados, de inicio, por el modelo y así lograr una corrección del factor de tendencia. Este procedimiento se describe en apartados sucesivos. Su aplicación va destinada más a sectores de modelización consolidada que han presentado un comportamiento estable a lo largo del tiempo y para los que no se considera necesario realizar modificaciones de entidad en los parámetros de validación.

Sin embargo, para unidades de consumo de reciente modelización, en las que no existe aún un conocimiento profundo de su evolución o en las que cabe esperar que la oscilación del consumo diario pueda ser recuperable, suele ser preferible revisar los parámetros de validación del consumo diario, a fin de dotar al modelo de una mayor tolerancia en relación a la misma. La evolución posterior del comportamiento del sector servirá para dictaminar sobre la bondad de esta actuación.

Como ilustración de este tipo de anomalía, en la figura 14, se recogen las lecturas de consumos horarios del sector de "Alcobendas centro" en el momento en que se produce un aumento en la magnitud de los consumos diarios, sin cambio significativo en el patrón horario. Se observa que en la noche del día 19 de octubre se produce un salto brusco del nivel de consumo leído (del orden de 20 litros por segundo) en todas las horas registradas, por lo que apenas existe alteración del patrón de consumo horario. Coincidiendo con este escalón, se produce el rechazo completo de todos los registros horarios y el reemplazo por las correspondientes estimaciones que quedan siempre en un nivel de consumo por debajo al del escalón anterior.

Ante la aparición de esta anomalía, al margen de la conveniencia de conocer el origen de este comportamiento, se revisó la modelización para determinar en qué momento se producía el rechazo del dato entrante. Repitiendo el proceso de validación manualmente, se comprobó que el conjunto de los datos horarios de consumo relativo sí era aceptado (en el gráfico puede observarse la aparente coherencia que existe entre los registros observados y estimados), produciéndose el rechazo en el momento de realizar la validación del consumo diario.

El buen comportamiento del patrón horario medido invitaba a pensar que los datos podrían ser fidedignos. Ante esta aparente coherencia de las lecturas y a falta de otros elementos de juicio, se optó por modificar el parámetro de aceptación del consumo diario, de cara a hacerlo más tolerante con este súbito aumento del consumo medio. La modificación se refleja en la figura 15.

Como resultado se obtiene una aceptación generalizada de la práctica totalidad de los registros medidos. Se observa que tan sólo se producen unos rechazos localizados en las horas finales del día 19, las inmediatas al aumento de consumo. En este caso, la eliminación de estos registros se debe al distanciamiento existente con respecto al consumo horario esperable. Recuérdese que la validación horaria se hace en términos relativos (consumo horario dividido por el consumo medio diario). Debido a que el escalón de consumo no abarca el día completo, los consumos relativos de las horas finales del día quedan descompensados y resultan ser muy elevados en comparación a los valores esperables; de ahí que se produzca el remplazo por una estimación horaria notablemente inferior.

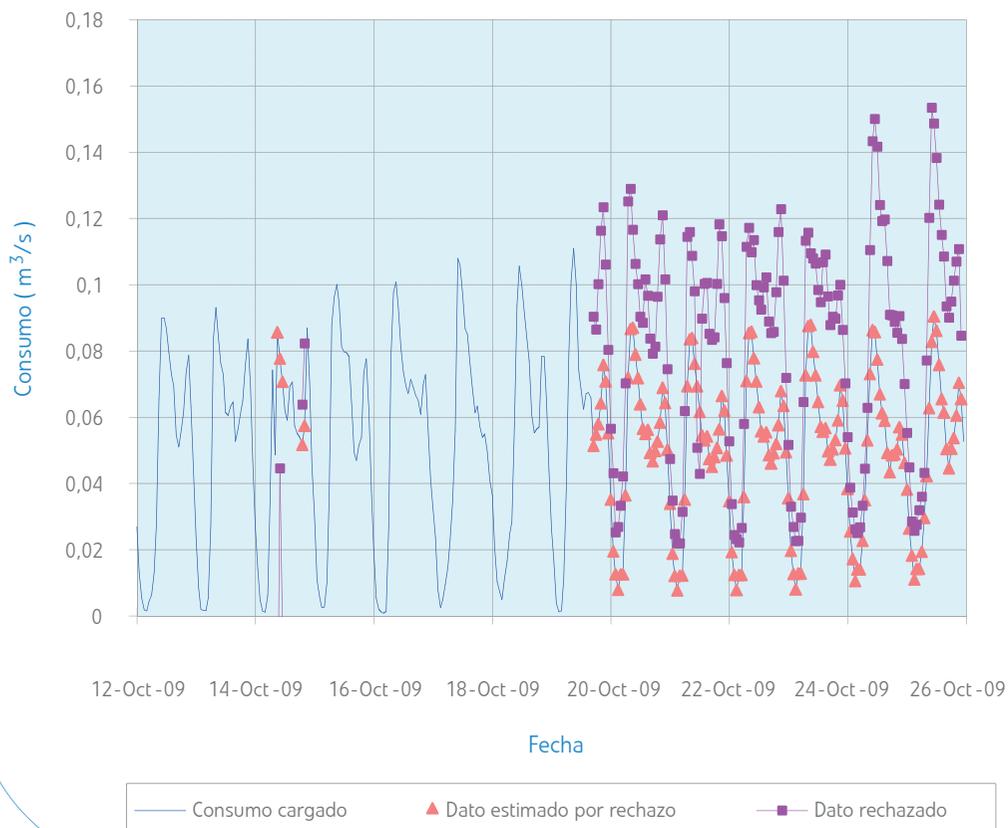
Sin embargo, a partir del día 20, al estar la totalidad de los consumos horarios igualmente mayorados, existe total coherencia con el patrón horario. Como la aceptación a nivel diario viene impuesta por el coeficiente de validación, se obtiene como resultado la aceptación generalizada de las lecturas entrantes.

Como último apunte conviene destacar que la relajación de los parámetros de validación implica aceptar el riesgo de que puedan darse por válidos registros espurios que, de no llevar a cabo esta actuación, quedarían descartados. Por este motivo, suele ser preferible actuar de esta manera únicamente en sectores nuevos, con alta probabilidad de verse sometidos a maniobras en la red, o donde los patrones de consumo no quedan claramente definidos en la modelización horaria y en los que interesa, ante todo, realizar sucesivas cargas de datos para tener una serie de longitud suficiente como para poder evaluar tanto el comportamiento del modelo como el de la evolución de los consumos del sector.

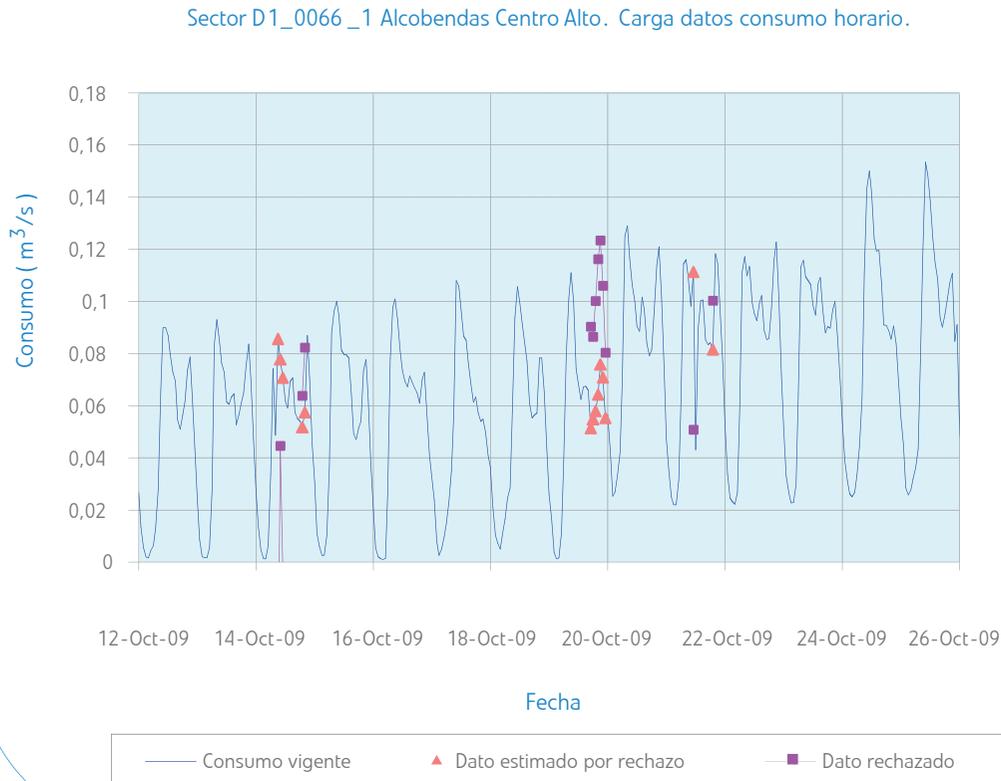
En estos sectores que aun no están "asentados", intentar reconducir el modelo actuando, exclusivamente, mediante validaciones de registros de manera manual impuestas por el explotador, puede constituir una tarea farragosa, compleja e innecesaria para una fase inicial de la modelización.

**Figura 14. Ejemplo de rechazo por modificación del consumo diario**

Sector D1\_0066\_1 Alcobendas Centro Alto. Carga datos consumo horario



**Figura 15. Resultados de la carga de datos una vez modificado al alza el coeficiente de validación del consumo diario**



#### 4.1.2. Modificación de los coeficientes de validación del modelo horario

La validación a escala horaria se basa en una formulación notablemente más compleja que la correspondiente al nivel diario. Ya se indicó en el apartado 3.3.4 que existen hasta tres etapas sucesivas para la aceptación de un registro, y estas se gobiernan por un total de tres parámetros de validación.

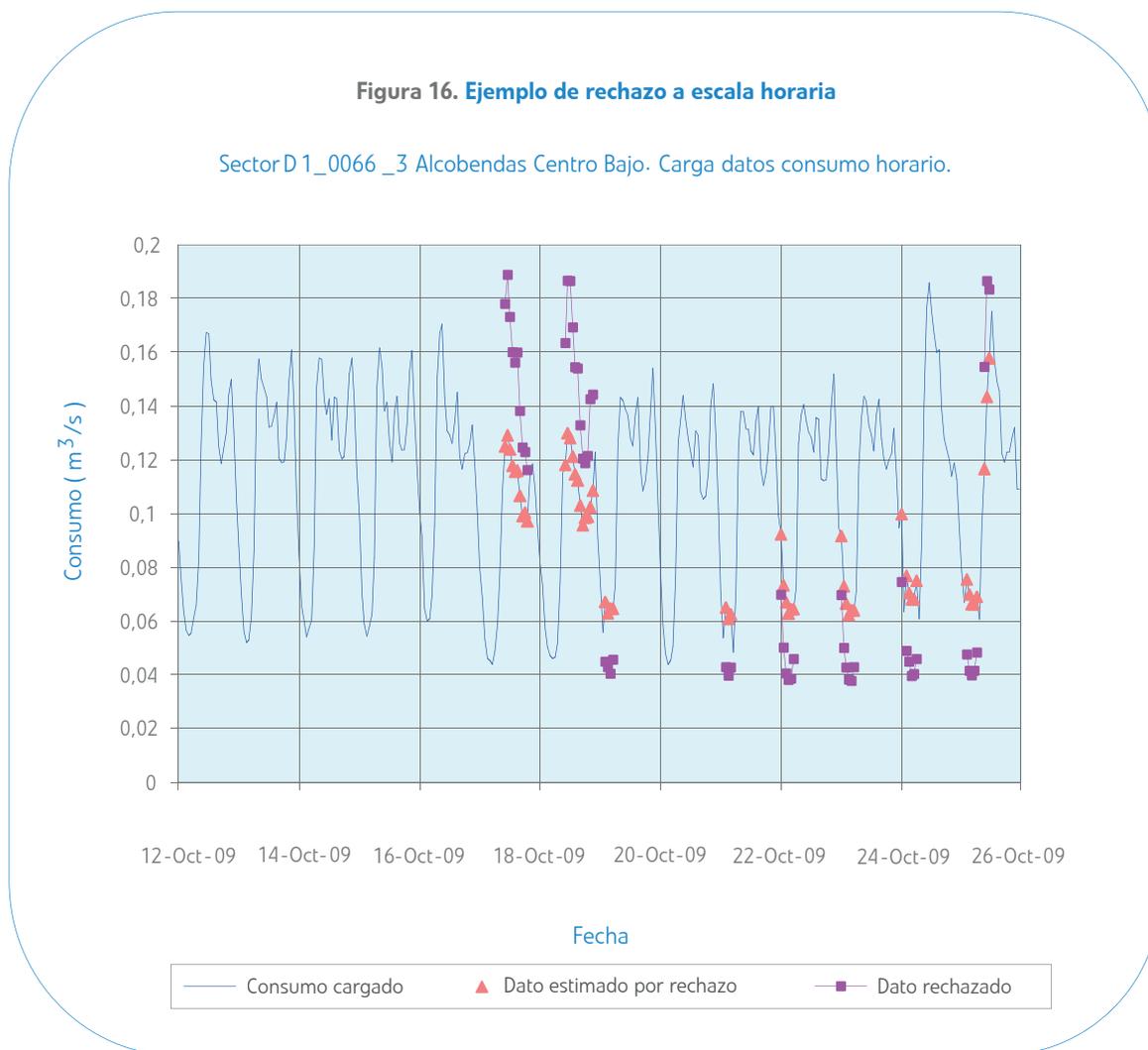
Por otra parte, si bien es relativamente frecuente que puedan existir modelos en los que se observen buenos ajustes a escala horaria pero en los que aparezca una falta de coherencia a nivel diario, en general, el comportamiento inverso no suele producirse (consumo diario aceptable y modelización horaria pobre). Además, un ajuste pobre a nivel horario suele ser argumento suficiente como para descartar la validación independientemente de la validez del consumo diario.

Este hecho es coherente con el planteamiento de modelización adoptado: en el caso de rechazarse un número suficiente de consumos horarios, se prescinde de la totalidad de los datos entrantes y se realiza una estimación completa tanto a nivel diario, como horario.

Por este motivo, las tareas de revisión a nivel horario suelen presentar una menor relevancia y se han centrado fundamentalmente en corregir los parámetros de aceptación del consumo de una franja horaria determinada más que en abordar fallos que pudiesen afectar al conjunto de datos horarios.

Como ya es sabido, la validación horaria considera como parámetros de validación la desviación existente dentro del conjunto de referencia en cada hora del día, de manera que se obtienen 24 valores distintos de desviación tolerable. Obviamente, la tolerancia del modelo es menor en las franjas de consumo horario más estables (valles nocturnos), mientras que se acepta una mayor variabilidad en las horas centrales del día o en los picos de consumo donde existirá una mayor dispersión en las lecturas del conjunto de referencia.

Esta diferencia en los criterios de tolerancia según la hora del día da lugar a un tipo de fallo relativamente común a muchos sectores, y que se hace evidente al efectuar la comparación entre consumos observados y validados. Consiste en el rechazo sistemático y puntual de los consumos correspondientes a un mismo tramo horario. La figura 16 ilustra bien esta tipología de fallo. En ella se presentan los registros de consumos horarios validados y rechazados por el modelo para el sector "Alcobendas Centro Bajo".



En primer lugar, debe destacarse el brusco descenso global (cerca de 20 litros por segundo) de los consumos de todas las horas que se inicia a lo largo de los días 17 y 18 de octubre (fin de semana) y que se mantiene hasta el final del periodo de análisis. Este descenso, sabiendo que es coincidente tanto en fechas como en magnitud con el aumento de consumo observado en el gráfico del apartado anterior, y siendo también conocido que los

gráficos se corresponden con sectores que pertenecen a un mismo núcleo de población (Alcobendas) puede asociarse, casi con total probabilidad, a una maniobra realizada en la compartimentación de la red.

En cualquier caso, más allá de las causas que pudieran justificar esta anomalía en los registros, interesa analizar la respuesta del modelo a este cambio de magnitud en el consumo diario. Al revés de lo que ocurría en el ejemplo del apartado anterior, tal y como puede observarse en la figura 16, la mayor parte de los consumos entrantes son dados por válidos. Se confirma, por tanto, que no se produce el rechazo completo derivado de una falta de coherencia a nivel diario.

En cambio, a nivel horario sí se produce el rechazo sistemático de los consumos correspondientes a los valles nocturnos. Esta respuesta obedece a la circunstancia, ya apuntada, de la poca holgura que presenta la validación de los registros nocturnos (hecho normal y esperable por otra parte).

Nótese que, además del escalón de consumo, se produce una ligera modificación del patrón intradiario. Los consumos de las horas centrales del día prácticamente quedan equiparados con los de la punta de la mañana y la del final del día. Como consecuencia, el consumo medio diario se aleja de los mínimos nocturnos. Por otra parte, de acuerdo con las desviaciones existentes en el conjunto de referencia, la modelización establece distintos límites de rechazo para cada consumo horario, parámetros que obviamente son más estrictos en los consumos que menos oscilan. En el sector considerado, la tolerancia es suficiente como para aceptar los registros en todas las horas excepto en las nocturnas.

Es importante destacar que estos rechazos puntuales son debidos exclusivamente a la modificación que se produce en el patrón intradiario y que no guardan relación alguna con el escalón de consumo medio diario pese a que se produzcan coincidentes en el tiempo con éste.

Para corregir este comportamiento inadecuado, cabría la posibilidad de modificar los parámetros de validación del consumo horario para las horas de la madrugada. Sin embargo, se advierte también que esta actuación puede tener efectos negativos. El seguimiento de los consumos mínimos nocturnos es útil para detectar fugas o consumos vegetativos que en las horas con mayor oscilación son difíciles de detectar. Aumentar la holgura de aceptación de consumos en estos periodos de valle sin duda reducirá la sensibilidad del modelo para detectar estas anomalías.

## 4.2. Corrección del factor de tendencia

### 4.2.1. Introducción

La función y formulación del factor de tendencia está recogida en el apartado 2.3.5. Esencialmente, este parámetro sirve para intentar acercar, en el análisis del consumo diario, las estimaciones del modelo a los valores observados. Para ello, se apoya en los datos de consumo correspondientes al pasado inmediato.

El objetivo perseguido al introducir este factor en la formulación fue el de permitir a los modelos responder a fenómenos evolutivos que pudieran presentarse, de respuesta lenta pero prolongada en el tiempo y, en consecuencia, difícilmente identificables en el momento de llevar a cabo las modelizaciones.

En este sentido, a la hora de establecer la formulación, no se pretendió dotar al modelo de capacidad predictiva para la estimación de tendencias a largo plazo sino que, más allá de un planteamiento conceptual, se estableció un factor corrector calculado a partir de las observaciones en el consumo de los días pasados y las estimaciones del modelo para ese conjunto de fechas. Su cálculo se realiza considerando como periodo de referencia los consumos aceptados en los 14 días anteriores según la siguiente expresión (ver apartado 2.3.5.).

$$T_i = \frac{\sum_{j=i-14}^{i-1} c_j}{\sum_{j=i-14}^{i-1} \frac{c_j}{T_j}}$$

donde  $T_i$  es el factor de tendencia para el día  $i$ , y, para los 14 días anteriores al  $i$ ,  $c_j$  es el consumo medio diario validado por el modelo para el día  $j$ ,  $c_j^*$  la estimación del consumo medio diario para ese día y  $T_j$  el factor de tendencia correspondiente también a ese día  $j$ .

El factor de tendencia correspondiente a un día y un sector determinado viene a ser, por tanto, el cociente entre los valores del consumo diario aceptados en ese sector en las dos últimas semanas, (valores vigentes tras los procesos de validación, rechazo y estimación de datos entrantes) y las estimaciones de consumo diario para cada una de esas fechas, proporcionadas por el modelo a partir del conjunto de predictores correspondientes a esos días y de las funciones que componen los términos de estacionalidad, calendario y meteorología del modelo (estimaciones por tanto no afectadas por el factor corrector de tendencia).

Si bien en la familia de consumos vigentes lo normal es esperar que la mayor parte de los registros se correspondan con datos de consumo entrantes (se acepta que puedan haberse producido rechazos en determinados componentes horarios o, incluso en algunos días completos, pero lo normal es que estos valores estimados sean minoría), hay que destacar que el término del denominador de la ecuación anterior (el valor del **cest**, de acuerdo con la terminología adoptada en el apartado 2.3.5), está formado exclusivamente por estimaciones realizadas completamente al margen del consumo que haya podido medirse.

En la medida en que los consumos aceptados/observados sean mayores que las estimaciones del modelo, el factor de tendencia tomará valores superiores a la unidad, de manera que las futuras estimaciones de consumo irán mayoradas. En cambio, si el consumo observado quedara por debajo del estimado, la tendencia iría disminuyendo de manera paulatina y por consiguiente también las estimaciones del consumo diario proporcionadas por el modelo en cuestión.

Como consecuencia de todo lo anterior, puede deducirse que el factor de tendencia juega un papel muy importante en el proceso de validación de consumos entrantes. Si bien inicialmente fue concebido para dotar a los modelos de una flexibilidad que les permitiera adaptarse a variaciones en el consumo, en general parsimoniosas y de carácter monótono, eventualmente puede servir también para permitir al modelo "reconectarse" en el caso de que se produzcan excesivas divergencias entre los consumos observados, aparentemente coherentes, y los valores estimados por el modelo.

En el apartado siguiente se describe de qué manera puede llevarse a cabo esta acción.

#### 4.2.2. Recuperación manual de datos rechazados

Como ya se ha explicado, el factor de tendencia actúa como un elemento de retroalimentación de errores. Así, el caudal estimado para una fecha concreta pasa a ser corregido en función de las diferencias que se observen entre los caudales simulados y caudales observados para los 14 días precedentes.

Como ya se ha apuntado, existen múltiples factores que pueden dar lugar a la aparición de desviaciones entre el consumo observado y el estimado por el modelo. La mayor parte de ellos (limitaciones de la formulación, error en los datos meteorológicos de entrada o diferencias debidas a errores en la lectura de caudales) supuestamente deberían tener una consideración menor y, además, un carácter aleatorio, por lo que pueden considerarse como ruido. En este sentido, no afectarían en gran medida al factor de tendencia, cuyo valor debería mantenerse estable.

Otra situación bien distinta se produce en el caso de que en una modelización comiencen a aparecer errores persistentes siempre del mismo signo, es decir, estimaciones de consumo sistemáticamente por encima (o por debajo) de los valores observados. En este caso, el factor de tendencia debería actuar corrigiendo las estimaciones de consumo a partir de la evaluación de los errores cometidos en el pasado inmediato (en particular, en los 14 días anteriores) logrando así un paulatino acercamiento entre los valores estimados y los observados.

Sin embargo, puede ocurrir que la evolución del consumo observado hacia un nuevo orden de magnitud sea tan brusca como para que el factor de tendencia, debido a presentar una inercia de 14 días, no pueda responder a ella debidamente y, como consecuencia, no pueda alcanzarse el reajuste necesario para acercar el valor de consumo estimado al dato medido. En esa situación es esperable que la distancia entre el registro de consumo y la estimación del modelo resulte excesiva y, como consecuencia, el consumo entrante original acabe siendo rechazado.

Siempre que se produce una laguna en un dato de consumo horario, bien por ausencia de lectura o por rechazo y reemplazo del dato original, la estimación de la tendencia para días sucesivos se realiza tomando como dato de consumo de partida el valor estimado (es decir, el aceptado tras el proceso de validación) en detrimento del observado. Efectivamente, no parecería lógico introducir para el cálculo del factor de tendencia datos de consumo observado que no hubieran quedado validados previamente. Este criterio tiene su relevancia, ya que, por lo general, suele ser difícil encontrar periodos de 14 días consecutivos en los que no existan lagunas de observación y/o validación. Adoptando siempre el valor vigente para el cálculo de la tendencia se consiguen varias ventajas:

- Se evita el tratamiento de lagunas.
- Todos los datos vigentes pasan a tener igual consideración e intervienen por igual en la determinación del factor de tendencia.
- En periodos largos sin datos, se consigue mantener estable el factor de tendencia en torno al correspondiente a los últimos datos observados.

Este último punto es importante: si se produce un “escalón” suficientemente pronunciado en la magnitud del consumo entrante y al que no consiga responder el factor de tendencia, en el proceso de validación podrán llegarse a rechazar estos datos y, a partir de ese momento, se mantendrán vigentes unos valores estimados en el nivel anterior al “escalón”.

Mientras que los consumos diarios observados no cambien en magnitud y retornen a valores próximos a los registrados antes de la aparición del “escalón”, el modelo seguirá generando lagunas por rechazo y realizará los pertinentes reemplazos por valores de magnitud similar a los registrados en el periodo anterior al salto, ya que el factor de tendencia apenas habrá sufrido variación (mantendrá un valor estable, semejante al obtenido antes de producirse la oscilación en el nivel de consumo). En caso de que no se produzca el retorno de las lecturas a los órdenes de magnitud anteriores al “escalón” en un máximo de 14 días, el modelo quedará a partir de entonces completamente “desconectado” de los datos observados.

Tal y como se detalló en el apartado anterior, 4.1.1, el rechazo de datos de consumo diario motivado por la aparición de un brusco salto en la magnitud de las lecturas se considera que responde a un comportamiento esperado por parte del modelo desarrollado. Es función del modelo detectar posibles anomalías en el funcionamiento de la red y los equipos de lectura y, en este caso, un ascenso (o descenso) marcado y producido en un corto lapso de tiempo puede ser sintomático de esta situación.

Es posible sin embargo que, tras el estudio de las causas a las que puede responder este cambio de magnitud en los consumos medidos, se concluya que los datos entrantes son válidos y que, por tanto, deberían ser aceptados por el modelo.

Para lograr este objetivo, como primera alternativa se expuso en el apartado 4.1.1 la posibilidad de modificar los parámetros de validación del consumo diario para hacerlos menos sensibles a estos cambios de magnitud. Como ya se indicó, esta opción se considera preferente para sectores de reciente modelización y en los que aun no se dispone de un conocimiento certero acerca de cuál será su evolución.

Sin embargo, en los casos en los que el comportamiento del sector ya está consolidado, y en los que se considera que los parámetros de validación adoptados son correctos y están avalados por el buen comportamiento observado en la modelización, puede ser preferible no modificar estos coeficientes de validación diaria y actuar exclusivamente corrigiendo el factor de tendencia.

Este tipo de actuación puede ser llevada a cabo por cualquier usuario familiarizado con la aplicación CHYPRE y que disponga de los correspondientes permisos de acceso. El objetivo perseguido es el de modificar este factor multiplicador de tendencia, de manera que el nuevo valor permita aproximar la estimación de consumo realizada por el modelo al dato observado lo suficiente como para que la discordancia existente entre estos valores sea pequeña y el consumo registrado se acepte como vigente.

Operativamente, en lugar de modificar manualmente el factor de tendencia se opta por respetar su formulación. Como alternativa, el proceso indirecto seguido para corregir el valor de este parámetro consiste en realizar una recuperación de un conjunto de datos de consumo rechazados por el modelo en primera instancia. Así, el procedimiento consiste en recuperar un conjunto de datos de consumo diario originales con fecha posterior a la aparición del "escalón" de lecturas y que de inicio quedaron rechazados por la modelización efectuada. La aplicación CHYPRE incluye esta funcionalidad dentro de sus opciones de gestión de datos. En concreto existen dos posibilidades de recuperación de datos:

- Recuperar datos e imponerlos como vigentes, de manera que quedan exentos de ser sometidos al proceso de validación (la modelización los acepta por buenos de antemano).
- Recuperar datos rechazados de inicio pero sin imponerlos como válidos, de forma que quedan como pendientes de validación (deberán someterse, por tanto, al habitual proceso de validación para determinar si son finalmente aceptados, o no).

Mediante estas opciones de recuperación, los datos rechazados en los procesos de validación pueden recuperarse como datos vigentes o como datos pendientes de validación. Esta función, por tanto, permite en cierta medida revertir el proceso de validación.

Ante la aparición de un salto en el nivel de consumo medio diario que haya provocado el rechazo masivo de los datos de consumo entrantes hasta el punto de que el modelo haya quedado "desconectado" de los posteriores registros observados, se propone proceder de la siguiente manera:

- Se recuperan y validan manualmente lecturas diarias correspondientes a 14 días consecutivos posteriores a la aparición del cambio de magnitud de las lecturas entrantes.
- Se recuperan manualmente los datos entrantes posteriores a los días validados en el paso anterior y se dejan como pendientes de validación.
- Se somete a validación la nueva serie de datos de consumo.

Al imponer la validación manual de datos observados, indirectamente se modifica el factor de tendencia (obtenido como cociente entre estos registros aceptados de antemano y los valores de **cest** estimados por el modelo y que, previsiblemente, quedarán lejos del "escalón" de consumo identificado). A partir de estos primeros 14 días recuperados y validados, la tendencia habrá evolucionado muy notablemente hasta adoptar un valor que permitirá acercar las estimaciones a las lecturas medidas supuesto que éstas se mantengan en los mismos órdenes de magnitud (es decir, que no se produzcan nuevos saltos bruscos en los niveles de consumo).

Nótese por tanto que, mientras que el modelo se esté comportando adecuadamente a nivel de respuesta del patrón diario a las variables estacionales, de calendario y meteorológicas, la existencia de un distinto orden de magnitud entre los consumos entrantes y los estimados en ausencia de información de las lecturas registradas puede quedar totalmente soslayada gracias al factor corrector de tendencia.

Debe advertirse también que la necesidad de intervenir manualmente sobre el modelo en estas circunstancias se debe exclusivamente a que la evolución del consumo observado es más rápida que la respuesta que el modelo teórico de consumo puede dar. En el caso en que, en lugar de producirse el escalón de consumo súbitamente, éste evolucionara de manera paulatina hasta alcanzar la nueva magnitud, el consumo modelizado debería responder por sí mismo adecuadamente a este cambio, aceptando todos los registros entrantes.

Tras la explicación teórica, quizás la mejor forma de entender esta tarea de mantenimiento de un modelo de consumo sea a través de un ejemplo práctico. En la figura 17 se recogen las curvas de datos aceptados, rechazados y estimados por el modelo de consumo del sector Venturada – Cotos de Monterrey. Este sector se corresponde con una urbanización perteneciente a un pequeño municipio situado a unos 50 kilómetros de la capital.

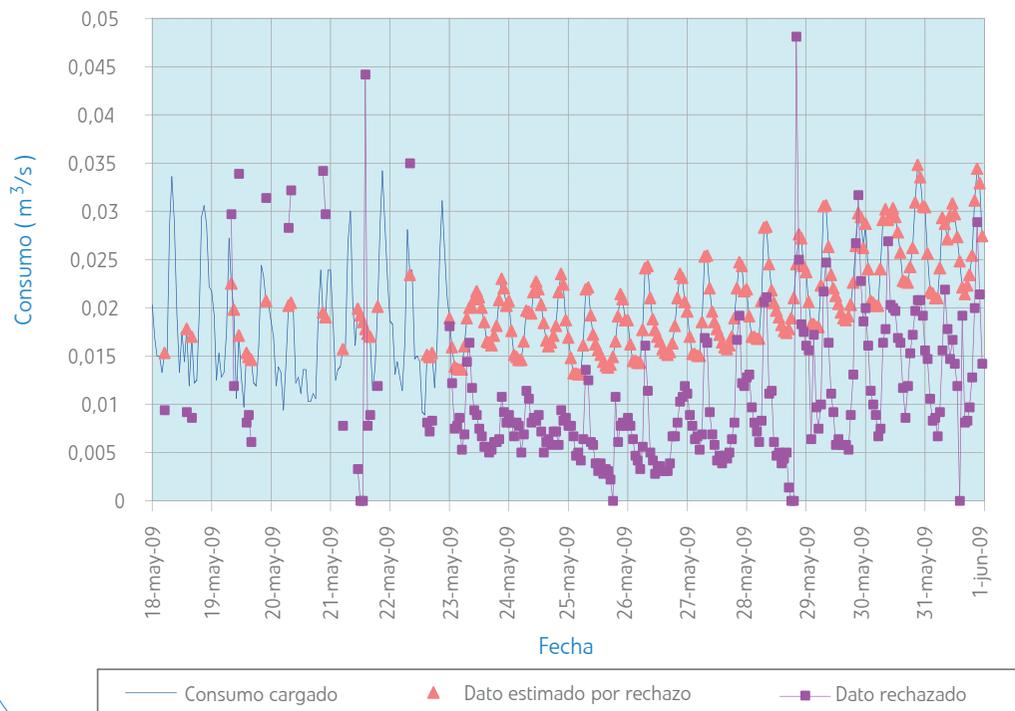
En la figura 17 puede comprobarse que, para los días anteriores al 23 de mayo, se produce la aceptación generalizada de la práctica totalidad de los datos de consumo cargados. Sin embargo, a partir de ese día, coincidiendo con un brusco descenso en los caudales medidos que afecta prácticamente por igual a todos los consumos horarios, el modelo pasa a rechazar los valores observados y los reemplaza por una serie estimada cuyo consumo medio diario es del orden de 0,02 metros cúbicos por segundo superior al consumo de la serie rechazada.

Los patrones horarios de la serie de consumos rechazada y estimada son similares, lo que permite concluir que el rechazo se está produciendo debido exclusivamente a la falta de coherencia del consumo medio diario.

Como complemento a la figura 17, conviene indicar que el **cest**, en las fechas mostradas en el gráfico oscila relativamente poco, adoptando en todo momento valores comprendidos entre 0,010 y 0,015 metros cúbicos por segundo. A la vista de la magnitud de los datos de consumo cargados en los días anteriores al 23, es obvio que el factor de tendencia para esas fechas adquiere valores notablemente elevados (alrededor de 1,6).

Figura 17. Ejemplo de rechazo a escala horaria

Sector TL\_1697\_2 Venturada - Cotos de Monterrey  
Carga de datos de consumo horario

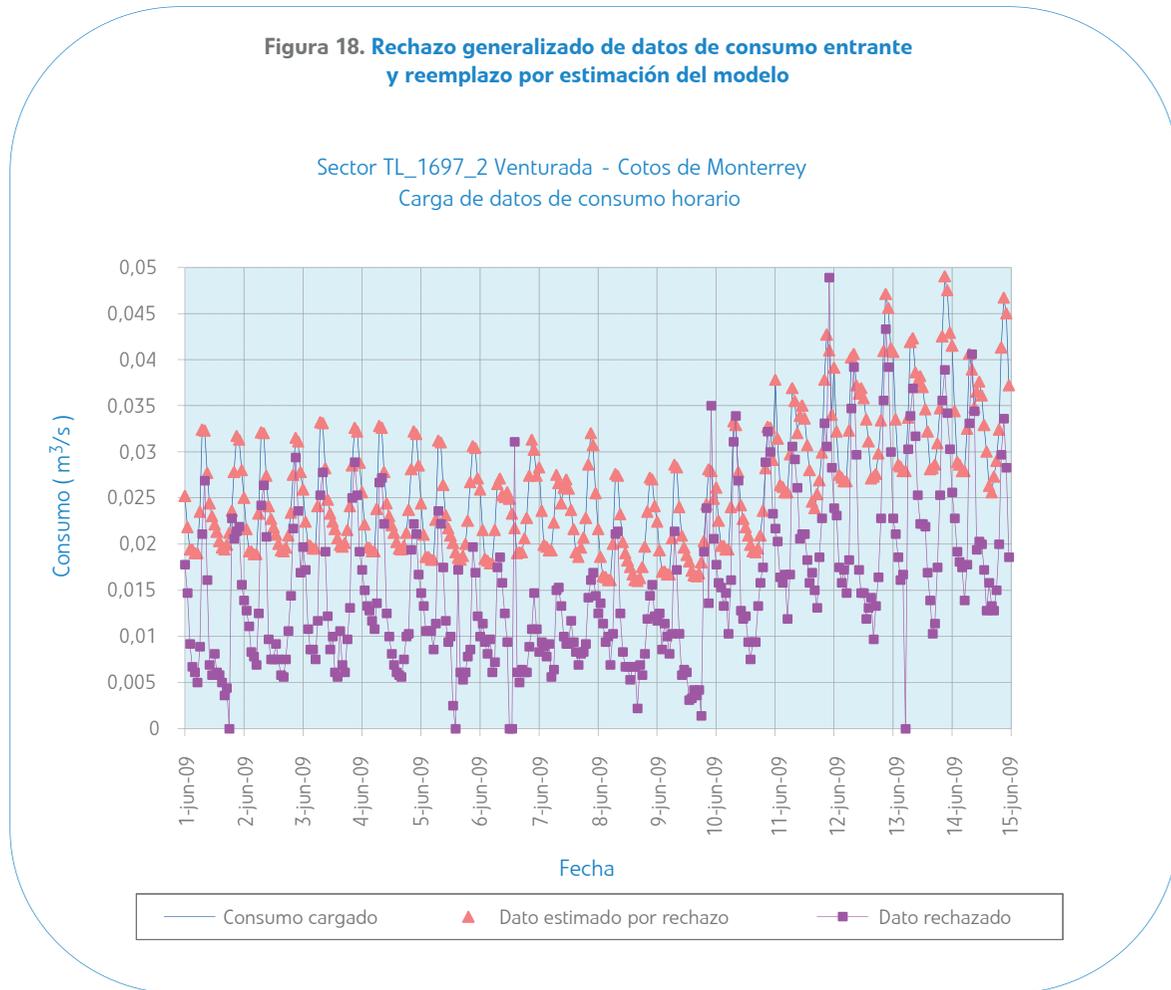


Al heredarse un valor tan alto del factor de tendencia, el consumo medio diario estimado resulta muy superior al caudal entrante una vez se ha producido el brusco descenso, motivo por el cual empiezan a producirse rechazos sistemáticos. En tanto en cuanto no se recupere el salto de caudal, los rechazos seguirán produciéndose, tal y como queda reflejado en la figura 18.

Debe destacarse que, si bien se observa una tendencia creciente en las lecturas rechazadas, ésta obedece a fenómenos de estacionalidad contemplados en la modelización. El sector de consumo pertenece a una urbanización en la que abundan las segundas residencias, por lo que es esperable que con el comienzo del verano se inicie el traslado poblacional a estos núcleos. En este sentido esta evolución es “esperable” por el modelo, de manera que las estimaciones también responden de manera semejante. En tanto en cuanto ambas curvas mantengan evoluciones “paralelas”, no será posible lograr que converjan los consumos estimados con los observados.

Debido al rechazo indiscriminado de lecturas entrantes que se produce, el cálculo del factor de tendencia se mantiene siempre en valores muy altos, ya que en todo momento el consumo vigente es notablemente superior al valor del **cest** (en definitiva, se mantienen valores de tendencia semejantes a los recogidos antes de la aparición del “escalón” de consumo).

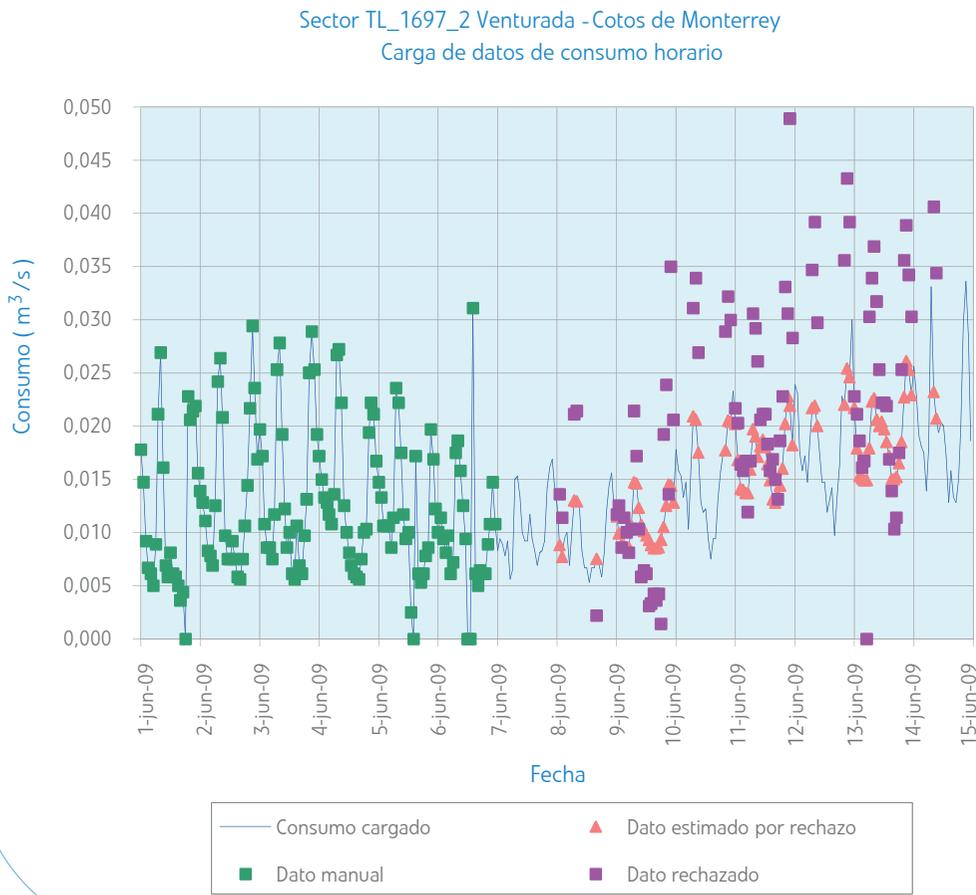
**Figura 18. Rechazo generalizado de datos de consumo entrante y reemplazo por estimación del modelo**



Este sería el comportamiento del modelo en el caso de no intervenir de manera alguna sobre él. Sin embargo, si pudiese descartarse que el escalón en la magnitud de consumo observado el día 23 de mayo fuera debido a un error de lectura, la aparente normalidad del patrón de consumos horarios invitaría a pensar que los registros entrantes deberían ser aceptados por el modelo. Para lograr la aceptación de estos registros, como ya se ha apuntado, podrían corregirse los parámetros de validación a escala diaria, pero esta medida podría hacer el modelo excesivamente tolerante. En esta ocasión se consideró preferible modificar el valor del factor de tendencia por medio de una validación manual de los registros entrantes inmediatamente posteriores al salto en el nivel de consumo.

La figura 19 ilustra este proceso. Los caudales horarios marcados en verde se corresponden con lecturas entrantes que se han aceptado y validado de forma manual, de manera que el modelo no las somete al proceso de validación. Estos datos, ya validados, pasan a participar en la determinación ulterior del factor de tendencia. Como ya se ha indicado, es conveniente recuperar un periodo de 14 días posteriores al “escalón” para que, a partir del día 15, el factor de tendencia se calcule íntegramente con datos en el nuevo nivel de consumo.

**Figura 19. Corrección del factor de tendencia a partir de la validación manual de consumos entrantes entre el 23 de mayo y el 6 de junio**



El comportamiento del modelo en este caso es muy diferente al que se observa en el gráfico de la figura 18 anterior. Como ya se ha indicado, todas las lecturas anteriores al día 7 de junio se han aceptado de manera manual al margen de los criterios de validación del modelo. En las fechas posteriores en cambio, el modelo ya funciona de manera automática. Los datos entrantes son aceptados en su mayoría, produciéndose tan sólo rechazos puntuales por pérdidas de coherencia con el patrón de distribución horaria. Además, los datos estimados por el modelo se mantienen en valores notablemente inferiores a las estimaciones de consumo que aparecen en la figura 18.

Este radical cambio de comportamiento del modelo se debe exclusivamente a la corrección del factor de tendencia que produce la aceptación manual de los datos de consumo de 14 días posteriores al escalón de lecturas registrado el 23 de mayo. El modelo pasa a calcular la tendencia a partir del día 7 de junio, tomando como valores vigentes unos registros de caudal medio diario que quedan incluso por debajo del valor que adopta el **cest**. Como consecuencia, los consumos diarios estimados en fechas posteriores son notablemente menores en magnitud. Debido a esto, los datos entrantes pasan a ser concordantes con los estimados y de ahí

la aceptación a escala diaria. Gracias a los 14 días de datos recuperados y validados, resulta posible “reconectar” el modelo, que a partir de ese momento vuelve a mostrar un comportamiento aceptable.

Conviene destacar una vez más que, en esta situación, se consigue que el modelo se comporte de manera mucho más coherente sin modificar ninguno de los coeficientes de su formulación o los parámetros de validación. Es decir, el modelo es idéntico en ambos casos. El mejor ajuste se logra modificando exclusivamente las “condiciones de contorno” (imponiendo la aceptación de las lecturas de 14 días consecutivos).

### 4.3. Recalibración completa del sector

Aunque no es frecuente, en determinadas ocasiones las modelizaciones efectuadas presentan un comportamiento que diverge, tanto a escala horaria como diaria, de las lecturas observadas, sin que existan sospechas de que los datos cargados puedan ser espurios. En estos casos se hace necesario proceder a una nueva calibración del sector. Esta circunstancia se produce especialmente en modelizaciones realizadas a partir de conjuntos de referencia pobres o de escasa longitud. Sobre todo cuando este periodo de referencia no abarca un año completo, puede ocurrir que los consumos de determinadas estaciones se alejen en gran medida, tanto en magnitud como en patrón, de los valores que en un principio se tomaron como referencia, lo que deriva en que el modelo sea incapaz de reproducir la evolución de los consumos observados en fechas distintas a aquellas en las que fue ajustado.

El patrón de consumo de un sector también puede verse alterado severamente en el caso de que se realicen maniobras de compartimentación en la red de distribución. La incorporación a la unidad sectorizada de elementos de consumo correspondientes a otras tipologías sin duda puede producir alteraciones de este tipo.

En el caso de identificarse este tipo de anomalías, será necesario comparar los distintos patrones identificados a fin de poder dilucidar si, más allá de un cambio de patrón, las alteraciones pueden obedecer a un efecto de la estacionalidad no identificado en la modelización inicial y, por tanto, no incorporado en los parámetros del modelo elaborado. En el caso de que pudiese existir una asociación entre ambos comportamientos, la recalibración del sector debería hacerse considerando como periodo de referencia toda la serie histórica de datos de consumo.

Por el contrario, si el análisis permite concluir que existe un cambio de patrón motivado por una circunstancia concreta, la recalibración del sector debería realizarse considerando como periodo de referencia exclusivamente las fechas más recientes, en las que quedará reproducido el nuevo comportamiento de la unidad de una manera más fidedigna.



# 5

Identificación de patrones  
de consumo

## 5.1. Introducción

A lo largo de los capítulos anteriores ha quedado evidenciado que la modelización de unidades de demanda de agua que se presenta está basada en representaciones matemáticas de los distintos hábitos poblacionales, obtenidas del análisis del conjunto de datos históricos de consumos horarios. Gracias a la compartimentación física de la red de distribución realizada por Canal de Isabel II, se puede disponer de información histórica correspondiente a un nivel de agregación mucho menor. Esta nueva configuración ha tenido una repercusión directa en los patrones de comportamiento que se persigue modelizar y que se puede sintetizar en dos aspectos fundamentales:

- Por una parte, la mayor escala de detalle permitió observar particularidades en los patrones de consumo asociados a unas tipologías de demanda determinadas y que, hasta ese momento, no habían podido ser identificadas.
- Por otra, al haberse multiplicado muy significativamente el número de unidades de demanda a modelizar y ser muy diferentes las evoluciones reflejadas en el patrón de consumos, resultaba mucho más difícil establecer una formulación y una utilidad de modelización que pudiese adaptarse de manera satisfactoria a toda las tipologías de sectores identificadas y, muy especialmente, a toda la casuística de anomalías, y cambios de comportamiento que los registros de consumo con frecuencia presentaban.

La tarea de análisis de patrones de comportamiento de cada unidad sectorizada se lleva a cabo sobre la serie histórica de consumos una vez realizada la depuración. Sobre el conjunto de referencia, se debe tratar de identificar patrones y regularidades asociados a factores y variables que sirven de predictores según la formulación adoptada e interpretar estas asociaciones desde un punto de vista conceptual.

Esta tarea constituye la parte del trabajo más implicada con el objetivo de generación de conocimiento. En este sentido, el disponer de unas funcionalidades implementadas en la herramienta de modelización desarrollada facilita de manera significativa la labor.

En general, la sectorización de la red se ha realizado asociando unidades de consumo de una misma tipología (por ejemplo, sectores correspondientes a polígonos industriales, municipios o comunidades urbanas que comparten hábitos poblacionales, urbanizaciones donde predomina la vivienda unifamiliar, pueblos de la sierra donde abundan las segundas residencias etc.). El análisis de las series históricas ha servido para poner en evidencia en qué manera y con qué grado de sensibilidad cada una de las variables de comportamiento consideradas en la formulación del modelo afecta al patrón temporal de consumo de una determinada unidad de demanda.

Para facilitar esta tarea, a partir del conjunto de datos históricos depurados de un sector, previo a su modelización, de manera sistemática se han obtenido una serie de gráficos de agregados de consumo y valores promedio a escala horaria, diaria y semanal, mediante los cuales resulta más sencillo obtener una composición de lugar sobre la evolución de los niveles de consumo y, en particular, sobre las siguientes características:

- Estacionalidad a escala interanual. En casi la totalidad de los casos, la estacionalidad está presente y en todos los niveles de resolución temporal. Sin embargo, según la tipología del sector, este efecto puede ser más o menos acusado o llevar signos contrarios.
- Efecto de calendario (diferencias entre consumos en días laborables y fines de semana). Suele estar presente en prácticamente la totalidad de las unidades identificadas, y afecta tanto al nivel horario, como al diario. Sin embargo, su influencia puede ser muy distinta según cuál sea la tipología de sector considerada.

- Influencia de la estacionalidad en los consumos de fin de semana (estacionalidad en la variable de calendario o diferencias entre consumos en días laborables y festivos pueden variar según la época del año considerada).
- Tendencias hiperanuales.

En los siguientes apartados se presentan distintas agrupaciones de sectores establecidas a partir de la identificación de una serie de patrones de comportamiento comunes a todos ellos. Naturalmente, la idiosincrasia de patrones de consumo existente entre los más de 500 sectores en los que quedará compartimentada la red de abastecimiento de Canal de Isabel II forzosamente será muy amplia, por lo que los conjuntos que a continuación se presentan no dejan de ser sino una aproximación muy simplificada. Con todo, se considera que esta agrupación sirve para identificar, de manera nítida, una serie de respuestas de las unidades de demanda muy diferenciadas entre sí y que sin duda pueden tener un carácter divulgativo.

## 5.2. Sectores pertenecientes a grandes núcleos urbanos del área metropolitana de Madrid

En el entorno del municipio de Madrid se asientan grandes núcleos urbanos donde predomina la edificación residencial. Estas poblaciones originalmente eran pequeños municipios de carácter rural o semirural. Sin embargo, a partir de la década de los 60 sufrieron un fuerte crecimiento demográfico, consecuencia del éxodo de la población del campo a las grandes ciudades. Concebidas inicialmente como “ciudades dormitorio”, en la actualidad disponen de una gran oferta de servicios públicos, industrias y comercios, por lo que los desplazamientos diarios de la población residente a centros de trabajo en la capital ya no son tan acentuados.

Las urbes de mayor tamaño asociadas a esta tipología son las que componen el “cinturón Sur” de Madrid: Alcorcón, Fuenlabrada, Getafe, Leganés y Móstoles. Todas ellas superan los 150.000 habitantes. Además, dentro de la red de abastecimiento de Canal de Isabel II, existen otras localidades en el corredor del Henares que presentan similares características poblacionales y de ordenación demográfica. Es el caso de núcleos como Torrejón de Ardoz, Coslada o San Fernando de Henares.

Pese a que es frecuente que estos núcleos cuenten con una importante presencia de tejido industrial, gracias a que generalmente estas industrias se concentran en polígonos a la hora de llevar a cabo la sectorización de la red de distribución ha sido posible segregar las dotaciones de manera que, salvo excepciones, cada sector se puede asociar a un uso preponderante.

Ciñéndonos a los consumos residenciales, debe destacarse que las poblaciones de estos núcleos se asientan predominantemente en viviendas plurifamiliares. Salvo en los barrios de más reciente creación, no es frecuente que existan piscinas comunitarias dentro de las urbanizaciones. Originalmente, se asentaron en estos núcleos población rural emigrante con niveles de renta bajos. Sin embargo, en la última década se ha producido un notable ascenso en el nivel salarial de la población residente.

Debido al tipo de edificación predominante, la densidad poblacional es elevada, por lo que los consumos medios por sector obtenidos son, en general, altos (alrededor de 0,1 metros cúbicos por segundo de consumo medio diario). Sirva como referencia, que los consumos medios de las residencias situadas en estos núcleos de población se estiman en el entorno de los 300 litros por vivienda y día.

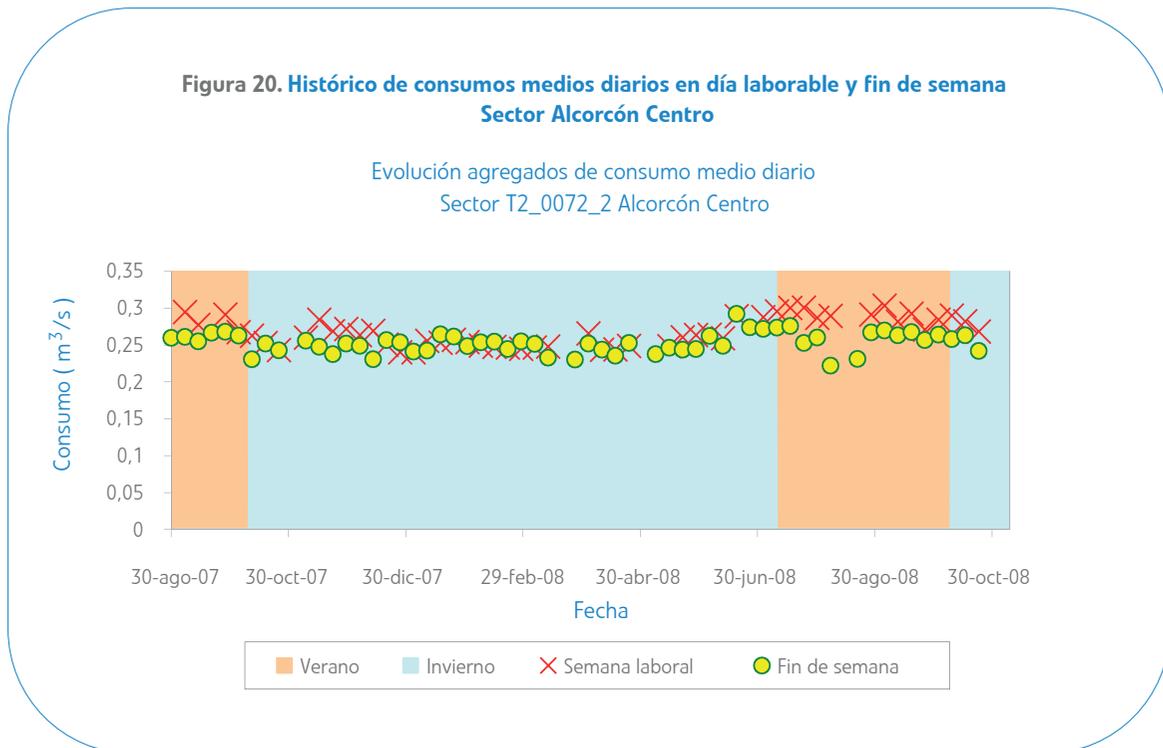
### 5.2.1. Análisis a escala diaria

Una característica común a los sectores pertenecientes a estos municipios es que presentan, en general, una estacionalidad de consumo poco acentuada. Esta reducida influencia se puede justificar sobre la tipología de demanda y los hábitos poblacionales. Las demandas típicamente asociadas a los meses de estío son reducidas en estos núcleos de población: apenas existen viviendas unifamiliares y piscinas, y las zonas ajardinadas son también escasas. Tampoco se producen grandes oscilaciones de demanda por desplazamientos poblacionales a zonas de veraneo. Esta movilidad poblacional tiende a concentrarse fundamentalmente en los fines de semana.

Como resultado, se distingue un ligero aumento de los consumos medios diarios durante los meses de verano en lo que se refiere a los días laborables. En los fines de semana, la tendencia es contraria, durante el verano disminuyen ligeramente los consumos respecto a los registros observados fuera del periodo vacacional.

El efecto de las variables de calendario es notable a lo largo de todo el año y siempre en el sentido de descenso del consumo de fin de semana respecto del correspondiente a días laborables. Se achaca esta diferencia a los desplazamientos de fin de semana. De acuerdo con lo expuesto en el párrafo anterior, en los meses de verano se agudizan estas diferencias.

Para ilustrar esta tipología de comportamiento, en la figura 20 se presentan los consumos medios diarios de cada una de las semanas que componen el conjunto de referencia de datos históricos a partir de los cuales se estableció la modelización del sector "Alcorcón Centro". Para valorar el efecto del calendario y su correspondiente estacionalidad, para cada semana del conjunto de referencia se recogen dos valores: uno correspondiente al consumo promedio de los días laborables, y otro correspondiente al consumo medio diario del fin de semana.



La evolución de este histórico de consumos puede considerarse representativa de la familia de sectores a la que pertenece, si bien en este caso, el efecto de la estacionalidad puede considerarse especialmente atenuado. Se observa que a lo largo de todo el año el consumo medio en fin de semana queda sistemáticamente por debajo del consumo promedio correspondiente a los días no festivos. En el verano esta diferencia queda algo más marcada, ya que mientras que el consumo en día laborable aumenta ligeramente por efecto de la estacionalidad, en fin de semana los consumos de verano llegan a descender respecto a los registros medidos fuera de la época estival.

### 5.2.2. Análisis a escala horaria

Para ilustrar lo que podría ser una evolución del consumo intradiario representativa de los sectores pertenecientes a los grandes núcleos urbanos situados en el extrarradio, a continuación se recogen los gráficos de consumo medio horario para cada día de la semana del sector "Alcorcón Centro". Para apreciar el efecto de la estacionalidad a nivel horario, en la figura 21 se presentan dos gráficos; uno, correspondiente a los consumos promedio del periodo octubre–mayo (invierno) y otro para el periodo estival (de junio a septiembre).

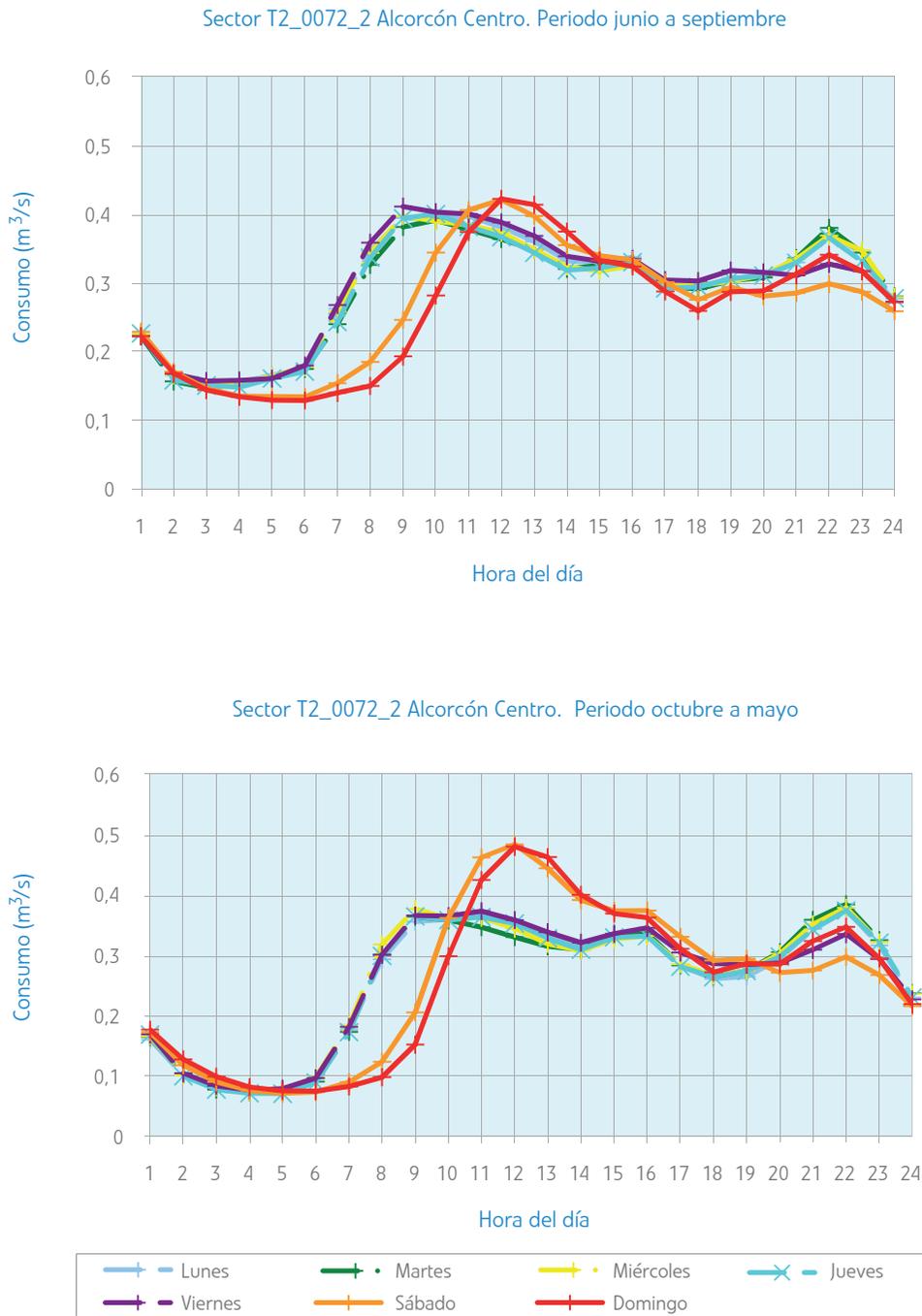
Las curvas que recogen la figura 21 representan patrones de consumo que responden al comportamiento general expuesto en el apartado 2.4. Aun así, existen una serie de particularidades que merece la pena destacar:

- Los niveles de consumo más altos se alcanzan, en todos los casos, al final del ascenso matinal (en torno a las 8 de la mañana en los días laborables y hacia las 12 del mediodía en los fines de semana). A partir de esta hora, el consumo se mantiene bastante estable durante el resto del día, si bien existe un pequeño valle relativo en la hora 18. El repunte nocturno es moderado en todos los casos.
- En los meses de verano, la diferencia entre el patrón de día laborable y el correspondiente a festivo se limita al retraso esperable en la reactivación matinal del consumo. En comparación, en el periodo octubre–mayo se hace evidente otra diferencia notable: la punta matinal es muy superior en fin de semana al valor correspondiente a un día laborable.
- El anterior es el único efecto de la estacionalidad identificable a nivel horario. El patrón de consumo en día laborable es prácticamente idéntico en las dos estaciones consideradas. No hay rastro de variabilidad estacional en los niveles de consumo en las horas finales del día.

Una vez expuestas las principales singularidades de los patrones de consumo, interesa justificar éstas en función de los comportamientos poblacionales esperables. El reducido efecto de la estacionalidad debe adjudicarse a la marginalidad de las demandas que mayor incidencia tienen en este sentido (uso y mantenimiento de piscinas y riegos). También debe interpretarse este comportamiento entendiendo que los hábitos poblacionales no varían significativamente de una estación del año a otra.

La elevada punta de consumo en fin de semana en el periodo octubre–mayo puede asociarse a la concentración de tareas domésticas y de higiene personal durante este periodo. El hecho de que este comportamiento no se repita en los meses de verano puede entenderse si se supone que, al disponer de más horas de luz, estas tareas se pueden realizar de una manera más paulatina. Complementariamente, este distinto comportamiento también podría obedecer a desplazamientos a segundas residencias.

**Figura 21. Patrón de consumo intradiario, según época del año. Sector Alcorcón Centro**



Como colofón, interesa hacer hincapié en que la práctica totalidad de los sectores integrados dentro la tipología a la que está dedicada el presente capítulo presentan un comportamiento a escala intradiaria muy semejante al observado en los gráficos de la figura 21.

Es decir, las curvas anteriores son claramente identificables como propias de un sector perteneciente a una urbe del extrarradio. Si bien a escala diaria los sectores estudiados pueden tener particularidades de comportamiento, sobre todo en relación a la influencia del factor estacional, a escala horaria, prácticamente, no existe desviación respecto al patrón representado en la figura 21.

### **5.3. Sectores pertenecientes a áreas residenciales situadas al noroeste del núcleo urbano de Madrid**

Dentro de esta clasificación se incluyen una serie de municipios situados en los alrededores de la capital, preferentemente en la zona noroeste, y que presentan unas características muy similares en cuanto a edificación predominante y perfil poblacional. Se trata de poblaciones en las que predomina la urbanización residencial unifamiliar y en la que la población residente dispone en general de una elevada renta per cápita.

Originalmente, estos municipios fueron pequeños pueblos cuyos habitantes se dedicaban al pastoreo y tareas agrícolas. Sin embargo, en las tres últimas décadas el crecimiento demográfico ha sido muy elevado, gracias al desarrollo de nuevos planes de ordenación urbana que dieron lugar al nacimiento de grandes áreas residenciales de viviendas unifamiliares. Se trasladaron a estas zonas familias de profesionales liberales que seguían manteniendo sus puestos de trabajo en la capital. Esta inercia ha cambiado en los últimos tiempos, ya que recientemente se han localizado en los alrededores de estos municipios algunos parques empresariales de considerable tamaño, por lo que los flujos diarios para acudir a los puestos de trabajo tienden a reducirse y acortarse.

En estos núcleos de población existen, además, grandes centros comerciales y buenos usos dotacionales, por lo que las necesidades tanto de ocio como de servicios están cubiertas sin necesidad de tener que realizarse desplazamientos a la capital.

Los principales municipios englobados dentro de esta tipología son Majadahonda, Pozuelo de Alarcón, Las Rozas, Villaviciosa de Odón y algunos sectores de Alcobendas (aquellos pertenecientes a La Moraleja). Sus poblaciones están comprendidas en el entorno de los 50.000 habitantes y, debido fundamentalmente a la tipología de vivienda predominante, se caracterizan por abarcar una gran extensión. Como es de esperar abundan las zonas verdes y ajardinadas. El número de piscinas, tanto de uso individual como comunitario, es también significativo.

Pese a que la densidad poblacional es reducida debido al modelo de urbanización que caracteriza a estos municipios, los consumos de estos sectores suelen ser elevados, sobre todo en los meses de verano, por efecto de los riegos y uso de piscinas. Los consumos medios diarios oscilan entre 0,05 y 0,1 metros cúbicos por segundo. Un valor esperable del consumo medio registrado en una vivienda unifamiliar de este tipo de sectores residenciales oscila en el entorno de los 600 litros por día.

Como último apunte conviene destacar que no han sido incluidos en esta clasificación municipios de semejantes características a las expuestas, como pueden ser Torreldones, Galapagar, Alpedrete o Collado Villalba. Pese a que, efectivamente, comparten todas las propiedades reseñadas en el apartado anterior, existe una particularidad que hace que los patrones de consumo de estos sectores presenten diferentes pautas de comportamiento: se trata del hecho de que en estas poblaciones coexisten primeras y segundas residencias. En cambio, en los núcleos citados en el párrafo anterior, el porcentaje de segundas residencias es prácticamente insignificante, por lo que no se producen aumentos poblacionales en los meses de verano o fin de semana. Esta diferencia redundará en una alteración significativa del patrón de consumo en los meses de estío, circunstancia que justifica que su inclusión en otra tipología de consumo.

### 5.3.1. Análisis a escala diaria

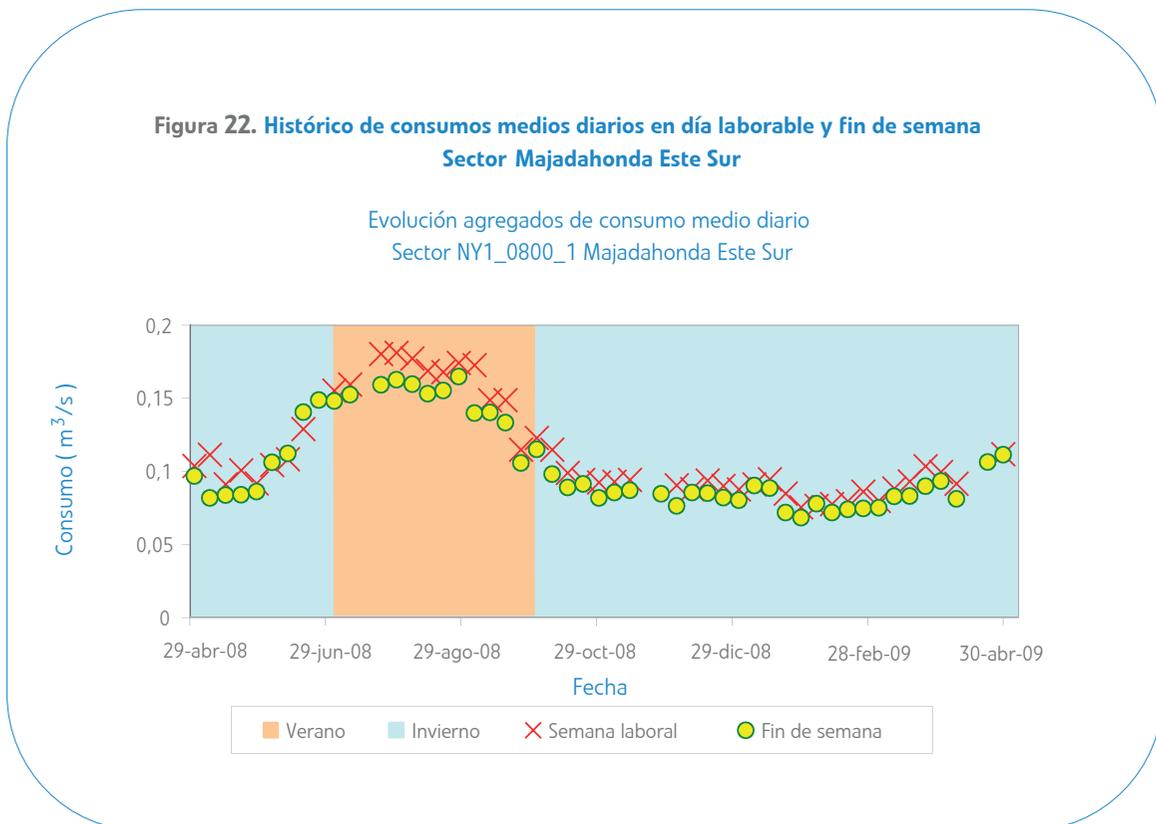
El análisis del consumo medio diario pone de manifiesto la fuerte estacionalidad existente: los consumos de los meses de verano son notablemente superiores a los del resto del año (del orden del doble).

Por su parte, las diferencias entre consumos en festivos y consumos en días laborables son considerablemente menos acusadas. Aun así, se cumple en todo momento que los consumos en periodo laborable son superiores a los correspondientes a fines de semana. En los meses de verano esta diferencia se hace más notable.

Estas pautas de comportamiento son coherentes con los hábitos de consumo esperables en esta tipología de sectores. La menor magnitud de los consumos en fin de semana apunta a salidas de la población residente a segundas residencias, de ahí que tenga una mayor incidencia en los meses de verano.

En este periodo estival, pese a que son esperables descensos en los consumos domésticos debidos a los cambios de residencia de la población a destinos vacacionales, los consumos de riegos de jardines y el mantenimiento de piscinas contrarrestan ese efecto y, como resultado final, se obtiene el aumento de los consumos de verano antes expuesto.

Como ejemplo representativo de la evolución de los consumos diarios la figura 22 recoge un gráfico de valores promedio registrados en días laborables y fines de semana correspondiente al sector "Majadahonda Este Sur".



El gráfico recoge datos de consumo promedio semanal para un total de 12 meses. Resulta muy representativo del comportamiento expuesto en los párrafos anteriores. Queda marcada la estacionalidad que se traduce en un fuerte aumento en los meses de verano a la vez que se hace evidente también la pequeña diferencia

existente entre consumos de fin de semana y de días laborables que se hace un tanto más notoria únicamente en los meses de verano.

### 5.3.2. Análisis a escala horaria

La estacionalidad observada en el análisis del consumo medio diario también presenta un claro reflejo a escala horaria. Se observan notables diferencias entre los patrones de los meses estivales y los correspondientes al resto del año.

Fuera de los meses de verano, el valle del descenso nocturno es muy pronunciado, y la reactivación matinal del consumo tiene fuerte pendiente y presenta un considerable recorrido hasta alcanzar un máximo hacia las 10 de la mañana. A continuación, el descenso es mucho más progresivo y menos acusado. El mínimo relativo se produce en la hora 19. Le sigue una suave curva ascendente que conduce al máximo nocturno (hora 22). Desde esta hora el consumo decrece hasta volver al mínimo en la madrugada.

En el verano, el descenso del consumo que se produce a partir de la medianoche no resulta tan pronunciado, de ahí que la curva ascendente de la mañana no tenga un recorrido tan grande. A cambio, comparativamente, sí se produce un descenso más acusado, hacia la media tarde. El mínimo de este descenso es en la hora 17 y el valor que se alcanza es comparable al valle de la madrugada. Desde esa hora se produce un repunte muy marcado, alcanzándose un nuevo máximo de magnitud semejante a la punta matinal hacia la hora 22. Por lo tanto, la estacionalidad al nivel horario es acusada en casi todas las franjas horarias, si bien resulta más evidente en las horas finales del día y en el mínimo nocturno.

Adicionalmente, existe otra pauta de variabilidad estacional en los patrones de consumo asociados a esta tipología de sectores. Responde a una particularidad del comportamiento del consumo durante los fines de semana.

Durante los meses de invierno, las diferencias entre consumo en día laboral y festivo se limitan al retraso en el inicio de la curva ascendente de la mañana. Los máximos de las horas del mediodía, si bien diferidos en el tiempo, alcanzan valores muy parecidos. Por el contrario, en el verano, si bien no se observa un retraso en el inicio de la reactivación matinal, se hace muy evidente una reducción en la punta de consumo alcanzada en la mañana para los días festivos en comparación con el registro alcanzado en días laborables.

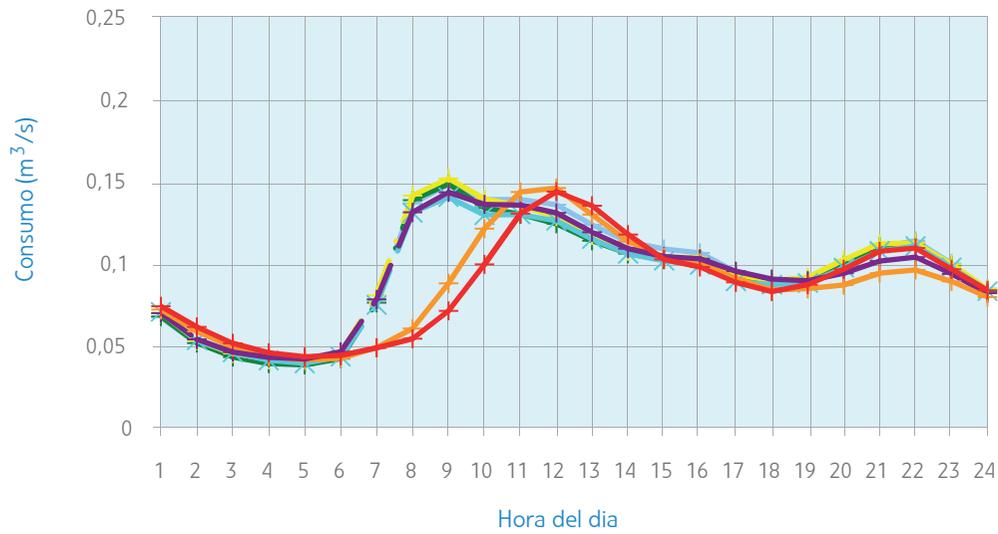
La figura 23 refleja gráficamente los patrones descritos en los párrafos anteriores. Se corresponden con las curvas de consumo medio horario por días y por periodo estacional del sector "Majadahonda Este Sur", cuyo comportamiento puede considerarse representativo del de los sectores pertenecientes a estas áreas residenciales.

Resultan fácilmente identificables la estacionalidad de las horas finales del día, la distinta magnitud de los descensos en las horas de descanso, la atenuación del valle de media tarde en los meses de invierno, respecto al consumo estival, y las particularidades del comportamiento en fin de semana en cada uno de los periodos estacionales considerados.

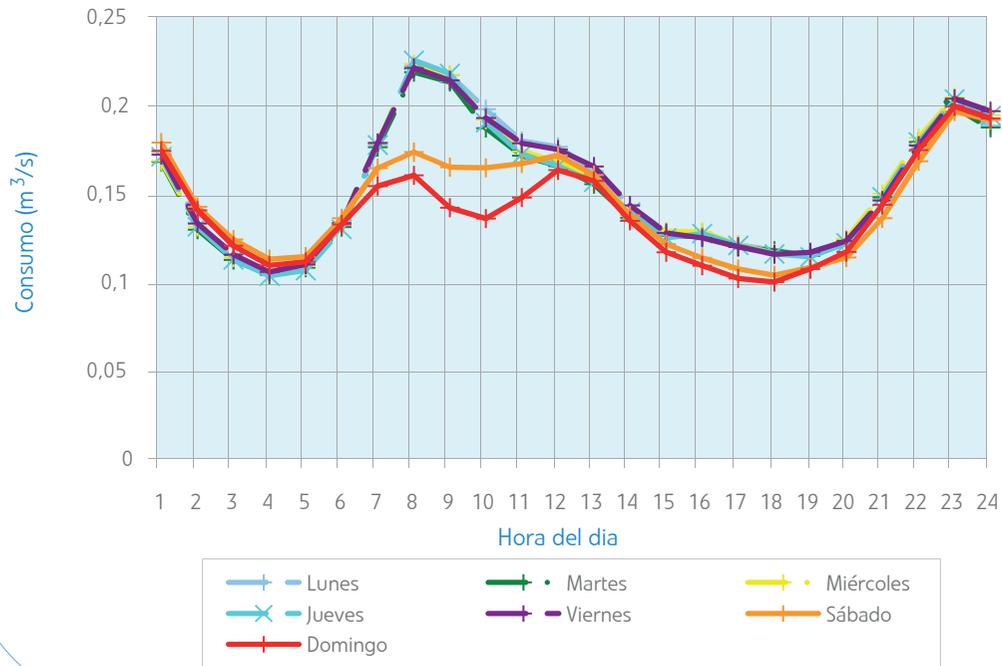
Esta variación del patrón intradiario debe interpretarse a partir de las pautas de comportamiento poblacional ya expuestas. Centrándonos en los meses de verano, el aplanamiento de la punta de la mañana en fin de semana debe interpretarse como una reducción del consumo doméstico debida a desplazamientos generalizados de la población a segundas residencias. Complementariamente, también puede interpretarse como que se produce un cambio en los hábitos de consumo de la población no desplazada asumiendo que se produce un retraso en el aseo personal, destinándose a este cometido las horas finales del día en lugar del inicio de la mañana.

**Figura 23. Patrones de consumo intradiario, según la época del año  
Sector NY1\_0800\_1 Majadahonda Este Sur**

Patrón de consumo horario promedio. Periodo octubre - mayo



Patrón de consumo horario promedio. Periodo junio - septiembre



La atenuación del valle del consumo mínimo nocturno se entiende que está motivada por las tareas de baldeo de calles y riego de parques y jardines que suelen desarrollarse en las horas de la madrugada y únicamente durante los meses de verano. Esta tipología de consumo probablemente también tenga su repercusión en los elevados consumos registrados en las horas finales del día.

Como último apunte, puede destacarse que la similar magnitud observada en las puntas matinales de los días laborables y festivos no pertenecientes al periodo estival parece ser reflejo de que las tareas domésticas no quedan concentradas en el fin de semana, sino que se desarrollan indistintamente en días laborables, quizás debido a la presencia de personal empleado en el servicio doméstico.

#### 5.4. Sectores pertenecientes a municipios localizados en la Sierra Norte

Dentro de esta tipología se engloban sectores pertenecientes a pueblos localizados en la sierra madrileña, concretamente en los valles de los ríos Guadarrama, Guadalix, Manzanares y Lozoya. Las características que la mayoría de estos municipios comparten son las siguientes:

- Tejido residencial conformado por población establecida en esa localidad de manera permanente y segundas viviendas que se ocupan en fines de semana y periodos vacacionales. Los porcentajes varían para cada núcleo, pero generalmente es predominante la segunda residencia.
- La población censada y residente durante la mayor parte del año en el municipio oscila entre los 1.000 y los 8.000 habitantes.
- Ausencia de tejido industrial. Las actividades laborales son, en general, escasas y se concentran fundamentalmente en el sector servicios. Debido a la afluencia de población estacional, los días de descanso de los comercios no se producen durante el fin de semana (preferentemente el día de cierre suele ser el lunes).
- Debido a la reducida oferta laboral de estos municipios, la ciudad de Madrid sigue constituyendo el foco principal de afluencia de la población trabajadora.
- La edificación residencial predominante es la vivienda unifamiliar, generalmente vinculada con una zona ajardinada y piscina de uso individual.
- Ocasionalmente, pueden existir explotaciones agropecuarias en las proximidades de los cascos urbanos.

Estos núcleos de población en su origen fueron municipios de carácter rural. A partir de la década de los sesenta, empezaron a desarrollarse urbanizaciones con viviendas destinadas a usos vacacionales y de descanso de población con residencia habitual en la ciudad de Madrid. Posteriormente, a partir de la década de los ochenta, estos municipios sufrieron un nuevo crecimiento motivado por el cambio de residencia de población proveniente de la capital.

De acuerdo con esta evolución, la distribución urbanística predominante en estos municipios se basa en un casco urbano donde se concentra la mayor parte de la primera residencia, así como el sector terciario y los servicios públicos. En torno a este núcleo se establecen las urbanizaciones de segundas residencias, siendo la mayor parte de ellas viviendas unifamiliares. Generalmente, a la hora de establecer la sectorización de la red de distribución, se ha tenido en cuenta esta división territorial, por lo que es habitual que un sector abarque la extensión del "pueblo" propiamente dicho, existiendo después tantos sectores individuales como urbanizaciones de entidad suficiente se cuenten.

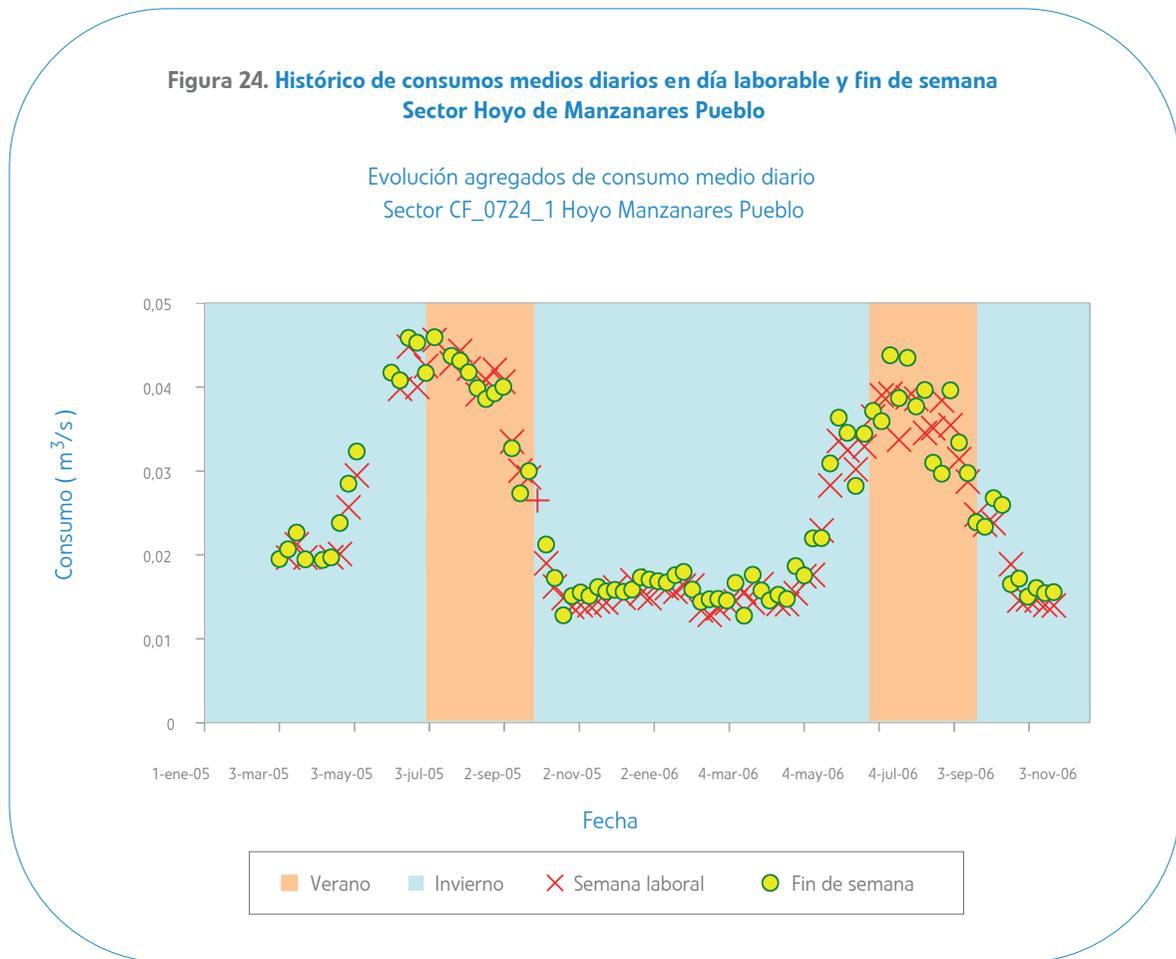
Al concentrarse en el casco urbano la mayor parte de la población residente en días de diario y las actividades que tienen lugar en esas fechas, la estacionalidad de los consumos de estos sectores es más reducida que la

correspondiente a las urbanizaciones donde, fuera de los meses de verano, la demanda en días laborables es casi despreciable en muchos casos.

### 5.4.1. Análisis a escala diaria

Como ya se ha apuntado, debido por una parte al elevado porcentaje de segundas residencias existente y por otra, a la tipología de demandas propia de estos sectores, la estacionalidad de los consumos diarios es muy marcada. En el mes de mayo comienza un progresivo ascenso del consumo medio diario que alcanza su valor máximo entre los meses de julio y agosto. A partir de septiembre se inicia de nuevo el descenso, que finaliza en el mes de noviembre cuando vuelven a registrarse los consumos promedio mínimos. El ratio entre el consumo promedio máximo alcanzado en verano y el mínimo de invierno puede llegar a ser holgadamente superior, del orden de 4 a 1.

Según se ha hecho en los capítulos anteriores para ilustrar el patrón de comportamiento de las demandas de agua en estos municipios, en la figura 24 se presenta un histórico de consumos medios diarios de un sector que se ha considerado representativo de la tipología a la que corresponde el presente apartado. En este caso, se ha seleccionado la unidad sectorizada denominada "Hoyo de Manzanares Pueblo".



El efecto de las variables de calendario es muy notable fuera de la época de veraneo, ya que, en los puentes y fines de semana, se producen desplazamientos de población de la capital hacia las segundas residencias que abundan en estas localidades. En el verano este efecto del calendario se difumina, ya que la movilidad poblacional no está comprometida a los días festivos. En general, al inicio del periodo de vacaciones escolares las familias trasladan sus residencias a estos municipios por lo que ya no se concentran los desplazamientos en el fin de semana. Alternativamente, se produce el traslado diario del fragmento de población que tiene que atender a puestos de trabajo ubicados en la capital.

Como es habitual, se presentan los promedios diarios de cada una de las semanas que componen el conjunto de referencia de datos históricos a partir de los cuales se estableció la modelización del sector. Se distingue entre consumos promedio de fin de semana y de día laborable para poner en evidencia las variaciones producidas por efecto del calendario.

Se observa que la estacionalidad del consumo es muy fuerte, los máximos de verano son del orden de 4 veces superiores a los correspondientes a los meses de invierno. En general, cuanto mayor es el porcentaje de segundas residencias en un sector, más marcada es la estacionalidad. En relación al efecto de calendario, éste es evidente en los meses de invierno, donde sistemáticamente el consumo en día laborable queda por debajo del promedio correspondiente a días festivos. Por el contrario, en los meses de verano, esta tendencia desaparece y el efecto de calendario queda muy difuminado.

#### 5.4.2. Análisis a escala horaria

La fuerte estacionalidad observada a escala diaria tiene también reflejo en el consumo intradiario. Para apreciar este efecto, tal y como se ha hecho en los casos anteriores, se presentan dos gráficos que pueden servir de ejemplo, uno correspondiente a los consumos promedio horario del periodo octubre a mayo y otro para el periodo estival (meses de junio a septiembre). Pertenecen al sector denominado "Hoyo de Manzanares Pueblo".

Es muy notable el aumento del consumo en fines de semana respecto a los días de diario durante los meses de invierno y primavera y durante las horas centrales del día. Fuera de estas fechas, los consumos de días laborables y festivos quedan prácticamente equiparados.

En el periodo octubre-mayo en fin de semana, además del mayor consumo horario se produce el habitual retraso en el inicio de la curva ascendente de la mañana con respecto a los días de diario. En estos meses, tras el máximo de la mañana se produce un descenso atenuado de los consumos. A partir de media tarde las demandas quedan estabilizadas con la salvedad de un pequeño máximo relativo en las horas finales del día previo al descenso del descanso nocturno.

Comparativamente, dejando al margen los meses de verano, el patrón de consumos intradiarios de estos sectores es semejante al que se observó en los pertenecientes a los núcleos urbanos del cinturón sur de Madrid (ver figura 21). La similitud en los consumos de días laborables es esperable, ya que el sector de Hoyo de Manzanares considerado se corresponde con el núcleo urbano del municipio, donde se concentra la mayor parte de la población que tiene residencia fija en esta localidad, así como los servicios públicos y comercios. Como además, fuera de los meses de verano la influencia en el consumo de jardines y piscinas es muy poco significativa, los hábitos de consumo resultan comparables a los de cualquier sector residencial donde prime la actividad urbana.

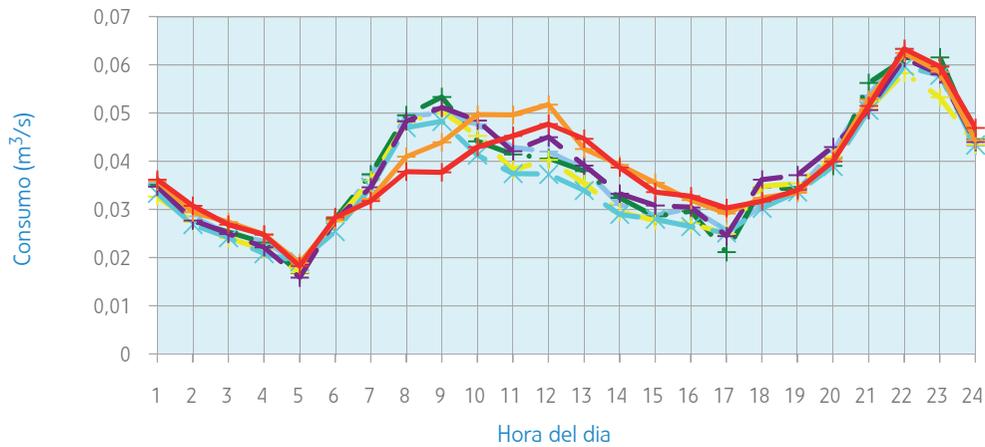
En fines de semana, en los municipios de la sierra de Madrid los consumos registrados son los de la población habitualmente residente más la parte correspondiente a la población desplazada. Mientras que en los sectores del cinturón sur el aumento de consumo en fin de semana se achacaba a la realización de tareas domésticas que no son llevadas a cabo en días laborables, en el caso de los municipios de la sierra el aumento de consumo en festivo es achacable fundamentalmente al aumento de población que se produce.

**Figura 25. Patrones de consumo intradiario, según la época del año  
Sector Hoyo de Manzanares Pueblo**

Patrón de consumo horario promedio. Periodo octubre - mayo  
Sector CF\_0724\_1 Hoyo Manzanares Pueblo



Patrón de consumo horario promedio. Periodo junio - septiembre  
Sector CF\_0724\_1 Hoyo Manzanares Pueblo



Por otra parte, en los meses de verano se observan dos particularidades claras: por un lado la homogeneización de los consumos de día laborable y festivo. Y, por otro, el fuerte aumento de los consumos registrados en las horas finales del día. La atenuación de las diferencias derivadas de la variable de calendario obedece claramente al asentamiento de la población eventual a lo largo de todo el verano. Con todo, sigue siendo perceptible un retraso y atenuación de la punta matinal en fin de semana. Esto se debe a que un porcentaje significativo de población residente sigue teniendo que obedecer a sus compromisos laborales.

Por su parte, la punta de consumo de las horas finales del día obedece a la relajación de los hábitos poblacionales nocturnos propia del verano (retraso en las horas de descanso) y a las demandas de riego de jardines, actividad que suele desarrollarse a partir del ocaso.

## 5.5. Consumos en sectores donde predomina la actividad industrial

Existen diversos polígonos industriales localizados en el extrarradio de la capital. Generalmente pertenecen a algunos de los municipios limítrofes a los que se hizo referencia en el apartado 5.2. Por la singularidad de sus consumos, la compartimentación de la red de distribución se ha llevado a cabo intentando aislarlos como sectores independientes. Por lo tanto, estas unidades de consumo se caracterizan por presentar una gran homogeneidad interna, de manera que el patrón de consumo que se obtiene es muy nítido y claramente identificable.

### 5.5.1. Análisis a escala diaria

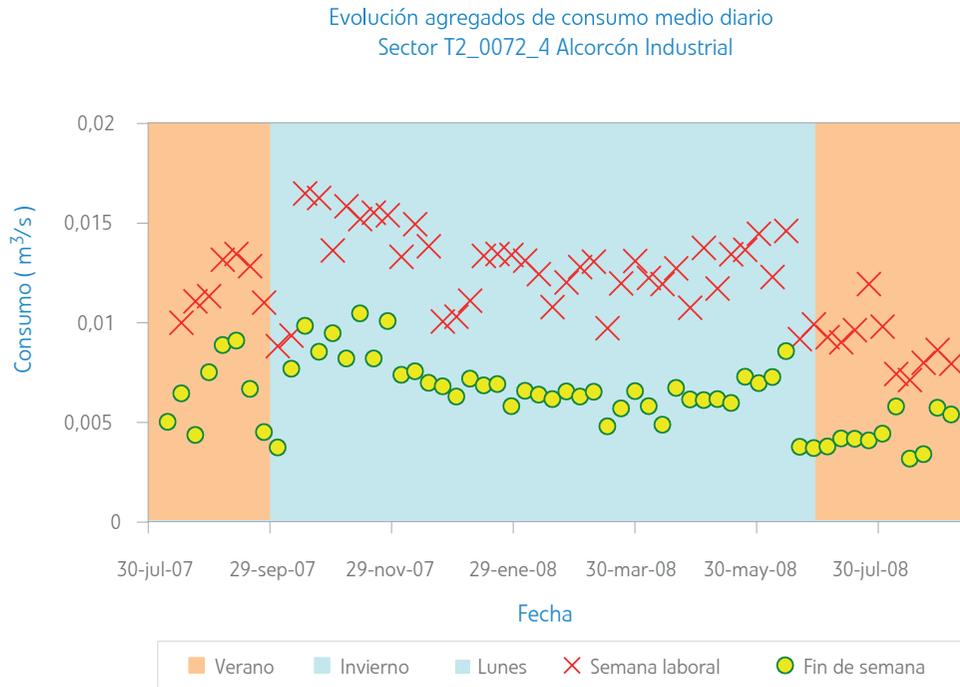
Como es esperable, la característica más evidente de esta tipología de consumos se basa en la acentuada influencia de la variable de calendario. El consumo en los fines de semana queda sistemáticamente por debajo del de días laborables y la diferencia entre ambos registros es siempre muy notable.

Aparte del efecto de calendario, los históricos de consumos diarios de los sectores de esta tipología analizados no reflejan la existencia de un patrón estacional común a todos ellos. En algunos sectores se aprecia un crecimiento sensible del consumo medio diario durante los meses de verano mientras que en otros, el efecto es el contrario. Alternativamente, en otro conjunto de sectores la estacionalidad apenas tiene influencia en los niveles de consumo.

Probablemente, la respuesta estacional de este tipo de unidades de consumo venga muy condicionada por el tipo de actividad industrial que se desarrolle en ellas; de ahí la variabilidad observada. A priori, es esperable que en consumos de agua relacionados con procesos de refrigeración sean mayores las demandas en los meses de verano. Sin embargo, también hay que tener en cuenta que es razonable esperar que, coincidiendo con el periodo vacacional, los niveles de producción puedan descender, lo que podría invertir el comportamiento antes apuntado.

Como ejemplo de la evolución estacional del consumo de un sector industrial se ilustra la figura 26, correspondiente al polígono industrial de Alcorcón. Destaca, por encima de todo, la diferencia de consumos entre días laborables y festivos. Además, este sector se caracteriza por presentar una elevada estacionalidad. En los meses de mayo y junio se atisba un crecimiento de los consumos promedio tanto en fin de semana como en día laborable. Parece razonable ligar este crecimiento con el aumento de temperatura correspondiente al cambio de estación. Sin embargo, a partir del mes de julio esta tendencia creciente se trunca ya que se produce un brusco escalón descendente en los niveles de consumo. Este descenso es achacable al inicio del periodo vacacional, que se entiende que redundará en un descenso de los niveles de producción y, como consecuencia, de los consumos de agua.

**Figura 26. Histórico de consumos medios diarios en día laborable y fin de semana Sector Alcorcón Industrial**



Aparte de la variabilidad estacional ya comentada, en aquellas semanas que cuentan con días festivos se evidencia también un descenso significativo en los consumos promedios en el periodo laborable, consecuencia de la menor presencia de mano de obra en esas fechas por puentes vacacionales. Obviamente, en fines de semana esta variabilidad apenas se acusa; los consumos se mantienen en valores notablemente estables.

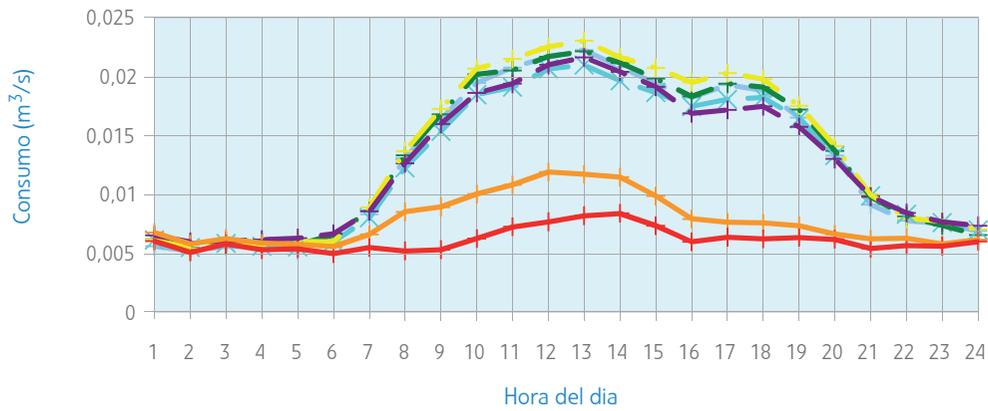
### 5.5.2. Análisis a escala horaria

Este tipo de sectores se caracteriza por presentar un patrón intradiario muy representativo y claramente identificable. En días laborables, el consumo empieza a crecer a primera hora de la mañana (hora 6 ó 7). Tras un tramo de fuerte crecimiento, a partir de la hora 9 se produce una atenuación de la curva ascendente. Los consumos se mantienen en general estables hasta la hora 16, momento en el que comienza el tramo descendente. Este descenso es continuado hasta la hora 23. A partir de este momento, los consumos se estabilizan y se mantienen casi constantes a lo largo de toda la noche. En días laborables por tanto, existe tan sólo un tramo ascendente y otro descendente.

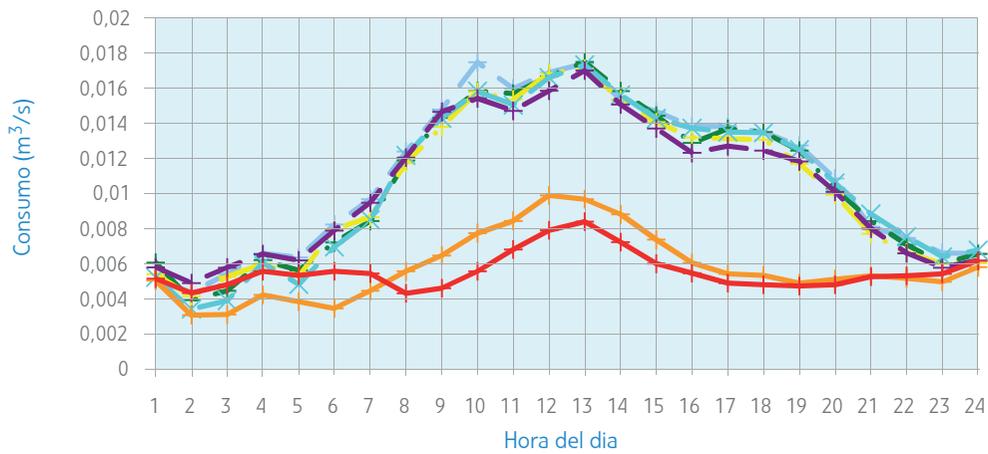
La figura 27 ejemplifica estos patrones de comportamiento particularizados para el sector "Alcorcón Industrial".

**Figura 27. Patrones de consumo intradiario, según la época del año  
Sector Alcorcón Industrial**

Patrón de consumo horario promedio. Periodo octubre a mayo  
Sector T2\_0072\_4 Alcorcón Industrial



Patrón de consumo horario promedio. Periodo junio a septiembre  
Sector T2\_0072\_4 Alcorcón Industrial



En fines de semana, la modulación es mucho más aplanada. Pese a esta atenuación de la diferencia entre máximos y mínimos, el patrón intradiario es semejante al observado en días laborables: ascenso matinal, estabilización en las horas centrales del día y descenso a media tarde para alcanzar de nuevo un mínimo, casi constante, a lo largo de toda la noche. Sin embargo, con respecto a los días laborables se produce un retraso en el inicio de la curva ascendente matinal así como un adelanto en el comienzo del tramo descendente de media tarde. Se distingue también un mayor nivel de consumo los sábados, con respecto a lo registrado en domingos, para las horas centrales del día. Por la noche, estos valores se estabilizan.

Estos gráficos permiten evidenciar la práctica ausencia de estacionalidad, característica común a este tipo de sectores. Únicamente se evidencia un adelanto de la curva de ascenso matinal en los meses de verano, que arranca hacia la hora 5. Para el resto de las horas del día el patrón es casi coincidente.

Como rasgo particular del consumo intradiario, que si bien no afecta a todos los sectores de tipología industrial es una característica común a varios de ellos, se identifica un escalón de consumo en el entorno de la hora 16. Se produce un descenso respecto al máximo de la mañana para, a continuación, estabilizarse. A partir de la hora 20 se retoma el patrón descendente.

Esta evolución escalonada se asocia a los cambios de turno y paradas de producción durante las horas de la comida. Obviamente, las curvas de ascenso y descenso deben coincidir con el inicio y finalización de las jornadas de trabajo. Por último, los consumos nocturnos se asocian a procesos industriales que se desarrollan de manera continua, aunque no es descartable que una parte no despreciable de este consumo vegetativo sea debida también a la existencia de fugas, que podrán estar localizadas tanto en la red de distribución como en válvulas o llaves de paso de las distintas instalaciones industriales.

## 5.6. Asignación de curvas de estacionalidad tipo

Las tipologías de usos de demanda analizadas en los apartados anteriores constituyen un claro ejemplo de la influencia de la estacionalidad en el patrón de consumo obtenido tanto a nivel diario como horario. Esta estacionalidad es evidente incluso en sectores que no pueden asociarse de manera concreta a un tipo específico de demanda. Por lo tanto, resulta evidente que prescindir de ella en la formulación de la modelización hubiera supuesto tener que admitir una tolerancia muy elevada de cara a que en todas las horas del día y en todos los días del año pudiese admitirse la variación debida al componente estacional.

Considerar la estacionalidad en la fase de modelización se convertía, pues, en una necesidad insalvable. Surgía, sin embargo, un inconveniente: en el arranque de las tareas de elaboración de modelos existían sectores para los que no se disponía de una serie de referencia de suficiente longitud o calidad como para definir de manera completa su curva de estacionalidad a la hora de llevar a cabo su modelización.

Como subterfugio, se planteó la posibilidad de que esta curva pudiese establecerse de antemano no a partir de los registros históricos, sino en función de la tipología del sector. Esta afirmación se fundamentaba a partir del conocimiento que paulatinamente se iba adquiriendo sobre los patrones de consumo. Tal como ha quedado comprobado, sectores de similar tipología y usos presentan curvas de consumo que, tanto a escala horaria como a nivel intranual, resultan notablemente similares, por lo que podría ser asumible el "adoptar" una curva tipo.

Las ventajas de esta forma de actuación resultaron incuestionables. Inevitablemente, imponer una curva de estacionalidad implicaba introducir en la modelización una ajena al comportamiento del periodo de referencia del sector en cuestión pero, en cualquier caso, el resultado era notablemente más ventajoso que si se renunciaba de manera completa a considerar la estacionalidad.

Según este criterio, a partir del análisis de las series históricas de los sectores de más antigüedad se intentaron realizar distintas agrupaciones por unidades de consumo que presentaran una respuesta similar en su patrón de consumo a la variable estacional. Como resultado se establecieron una serie de familias de sectores más o menos "homogéneos". Para cada familia se obtuvo una curva de estacionalidad de calendario y de festividad tipo, como valor promedio de las de cada uno de los sectores que integraban ese conjunto.

El siguiente paso consistió en establecer una forma objetiva mediante la cual, durante la fase de modelización de un sector con una serie histórica de poca calidad o longitud, se le pudiese asignar una de estas estacionalidades "tipo".

El método adoptado consistió en establecer un árbol de decisiones basado en el análisis de las características que pudiese presentar el conjunto de referencia del sector en cuestión. De esta manera, respondiendo a una serie de preguntas sobre el comportamiento de los datos históricos, se acababa obteniendo cual era la familia que teóricamente mejor se ajustaba.



# 6

## Influencia de la meteorología en los consumos por sectores

## 6.1. Introducción

Dentro del capítulo dedicado a la formulación del modelo diario, en el apartado 2.3.4 quedó recogida la descripción de los predictores meteorológicos y la justificación de su adopción. A modo de recordatorio, interesa destacar en este punto que su selección se realizó a partir de las relaciones de correlación que pudieron identificarse entre variables asociadas a los registros diarios de precipitación y temperatura acumulados y los correspondientes históricos de consumo agregado en el conjunto de la red de distribución de Canal de Isabel II.

La experiencia adquirida en la fase de mantenimiento de sectores modelizados ha servido para atestiguar que, a nivel global, los resultados que con esta formulación genérica adoptada se han obtenido son buenos. Comparando los niveles de ajuste alcanzados en las modelizaciones de sectores realizadas prescindiendo de los predictores meteorológicos y los correspondientes al modelo completo, queda evidenciado este hecho.

Pese al buen comportamiento observado en términos generales, la bondad del ajuste conseguido analizado individualmente para cada sector es muy variable. Es evidente que, según la tipología de unidad de consumo que se considere, la respuesta a los efectos meteorológicos será muy dispar: la decisión de activar o no un sistema de riego puede estar muy condicionada por el hecho de que se haya producido precipitación en fechas recientes, mientras que la repercusión de esta lluvia en un consumo industrial previsiblemente será reducida. Este distinto efecto debería estar contemplado en la modelización a través de los coeficientes del término lineal del modelo diario correspondientes a las variables meteorológicas. Según cuál sea la sensibilidad del consumo a la oscilación de temperatura y precipitación identificada en el conjunto de referencia, estos parámetros adoptan valores que dan un mayor o menor peso al componente meteorológico en relación a los restantes factores (tendencia, estacionalidad y calendario).

Con todo, es evidente que la dificultad en poder identificar el componente meteorológico de un patrón de consumos es muy superior si se realiza para un sector en particular que si se considera a escala agregada. Ya se ha apuntado que la parte fundamental del efecto meteorológico queda considerada al abordar la estacionalidad, ya que la temperatura ambiente claramente está sujeta al ciclo estacional. El análisis de la influencia de la variabilidad meteorológica constituye un paso más, ya que intenta traducir en términos de consumo, las desviaciones respecto a los valores esperados (en qué medida afecta al consumo que un día sea más caluroso o más frío de lo esperable en esas fechas). Dado que las series históricas de datos de consumo a partir de las cuales se realizaron las modelizaciones de sectores en pocas ocasiones superaban los dos años de longitud, la dificultad para identificar estas variaciones que se superponen a las propias de la estacionalidad resultaba muy notable.

El objetivo perseguido se basa, pues, en encontrar la respuesta a registros climatológicos “anómalos” o bien alejados de la media. A un nivel agregado, esta tarea puede ser más sencilla de acometer, debido a que la persistencia es mayor e influyen menos las desviaciones que puedan achacarse a fenómenos no contemplados en la modelización o a errores de lectura. Además, se cuenta con una serie de registros de mayor longitud (por encima de los 10 años). Sin embargo, al nivel de detalle de un sector particular, para conocer de manera adecuada la respuesta a la variabilidad meteorológica y poder discriminarla de las perturbaciones debidas a la calidad de los registros o a oscilaciones no imputables a ella, se requiere tener registrados consumos históricos coincidentes con una amplia variedad de niveles de precipitación y temperatura, condición que sólo en contadas ocasiones se ha cumplido, dado el poco tiempo transcurrido desde el inicio de la sectorización de la red de distribución.

Además de la ya apuntada, una limitación adicional que la selección de predictores meteorológicos considerada y la formulación desarrollada presentan es que no permiten realizar una aproximación de la variabilidad del consumo horario en respuesta a los fenómenos meteorológicos. De la misma manera que la estacionalidad de consumo tiene reflejo a nivel diario, pero también a escala intradiaria, sin duda la variabilidad meteorológica ha

de tener un efecto sobre la forma en que el consumo diario se distribuye entre las horas del día. Es evidente que determinados usos que pueden verse más afectados por los efectos meteorológicos, se concentran habitualmente en una franja horaria en particular (es el caso, por ejemplo, de los riegos de jardines que suelen realizarse al final de la tarde o primera hora de la mañana).

Abordar este estudio habría planteado una complejidad muy superior a la que se encontró en el análisis del componente estacional del consumo horario. Debe destacarse que, a priori, la estacionalidad afecta de una manera muy similar a todas las épocas del año, es decir, el modelo asume que el efecto estacional será el mismo en todos los veranos de la serie analizada, cosa que no ocurre en el análisis de la oscilación meteorológica.

Además, cuanto menor es la escala temporal considerada, mayores resultan las dificultades en encontrar la parte de la variabilidad de consumo achacable al efecto meteorológico. La oscilación a escala horaria es, en general, mucho mayor que la del consumo medio diario por lo que resulta más difícil distinguir la influencia de la meteorología aislada del resto de efectos no contemplados en la modelización realizada y que, por tanto, deberían considerarse como ruido en esta aproximación.

Pese a estas dos limitaciones evidentes, el análisis del efecto de la meteorología a escala de consumo global sirvió para evidenciar que, incluso para los periodos más extremos (años meteorológicos más inhibidores y más favorecedores del consumo de agua), los resultados obtenidos, si bien tomados estos con las cautelas correspondientes a la naturaleza y limitaciones del modelo utilizado, permitieron concluir que el impacto de desviaciones en el régimen de precipitación y temperatura tenían un efecto sin duda apreciable pero, en cualquier caso, limitado en los niveles de consumo (desde luego muy por debajo del asociado a las variables de estacionalidad y calendario). Este argumento se consideró razón de peso suficiente como para que no resultase necesario tener que intentar abordar una formulación que permitiera afinar más la escala temporal y la precisión del ajuste en el análisis de la influencia meteorológica.

Pese a todo lo dicho, y aun aceptando que los resultados obtenidos son buenos y que la capacidad del modelo para representar la variabilidad meteorológica es más que aceptable, se ha considerado oportuno realizar un análisis en detalle de los consumos registrados en aquellos sectores que a priori podrían resultar más susceptibles de presentar una respuesta más acusada, incluso a nivel intradiario, al componente meteorológico. De esta manera, se ha querido avanzar en el conocimiento del comportamiento observado en los patrones de consumo más allá del nivel alcanzado en la fase de modelización. En los siguientes apartados se presentan los resultados obtenidos del estudio realizado, junto con las limitaciones del análisis que han podido identificarse.

## 6.2. Periodo de referencia seleccionado

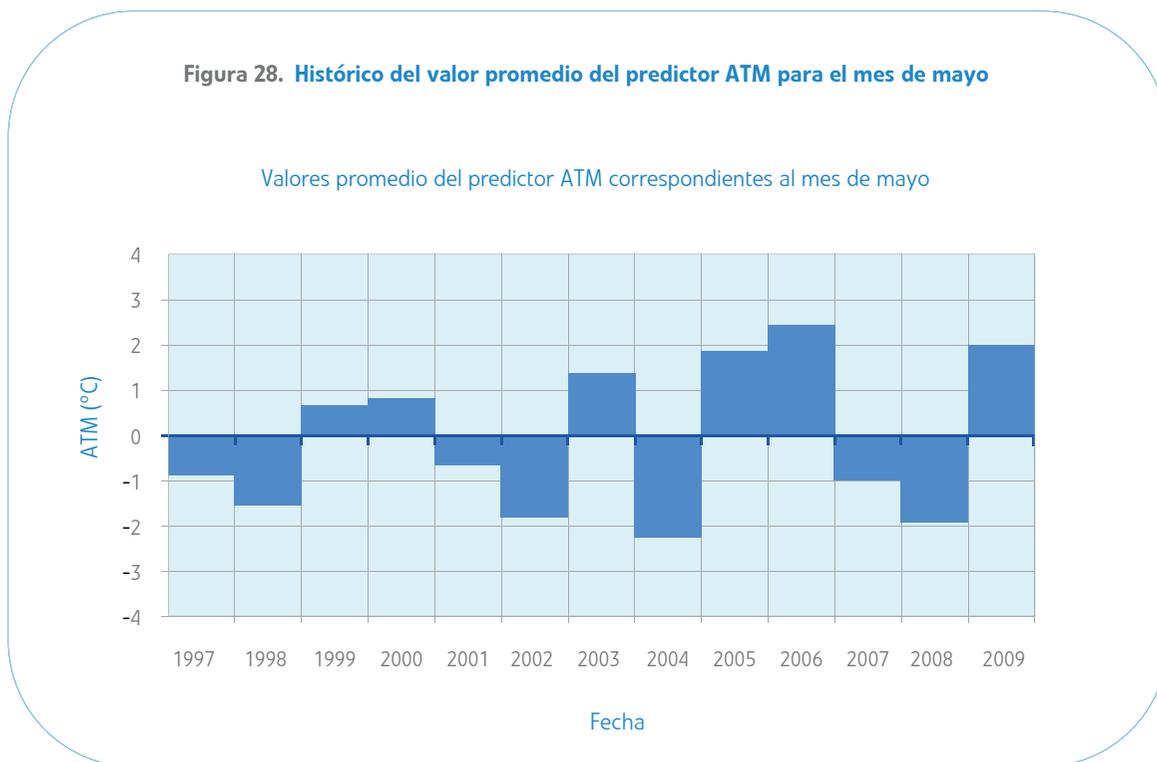
Como ya se apuntó en el apartado 2.3.4, la influencia de las variables meteorológicas en el consumo de agua no es homogénea a lo largo del año. En los meses inmediatamente anteriores o posteriores al verano la respuesta a la oscilación térmica es mucho más acentuada, mientras que en los meses de invierno apenas existe correlación. Este comportamiento es coherente por dos motivos independientes:

- Por una parte, las mayores desviaciones de temperatura se producen generalmente en los meses donde tiene lugar la transición entre las estaciones fría y seca. Debido a la continentalidad de la climatología de la Comunidad de Madrid, los inviernos y veranos suelen ser muy marcados, siendo tan sólo los meses de abril, mayo, septiembre y octubre los que pueden considerarse como de transición.
- Por otra parte, los mayores escalones de consumo se producen al inicio y a la finalización de la temporada de verano debido a la fuerte estacionalidad que presentan los usos de mayor demanda.

Se ha considerado que la mejor forma de plasmar el efecto de la meteorología consiste en comparar dos periodos correspondientes a la época de mayor "sensibilidad" meteorológica, y en los que la anomalía de precipitación y temperatura adopte signos contrarios. Preferentemente, sería deseable que estos dos periodos de comparación fuesen próximos en el tiempo y relativamente recientes, ya que la disponibilidad de datos de consumo en sectores es limitada y se reduce a unos cuantos años (no más de 5 en el mejor de los casos).

De acuerdo a los anteriores razonamientos, se ha hecho el ejercicio de comparar la variabilidad meteorológica que se ha registrado en los meses de mayo (el mes donde más peso toma el efecto meteorológico) de los últimos años. Interesaba identificar aquellos meses que presentaran mayor disparidad entre sí, entendiendo que serían los periodos en los que más evidentes se harían las anomalías de consumo imputables a la meteorología.

Dado que, según lo expuesto en el apartado 2.3.4, la influencia de la precipitación en la demanda de agua es menos marcada que la correspondiente a la temperatura para identificar los meses con meteorología más y menos favorecedora del consumo, se centró el análisis en la comparación del valor medio mensual del predictor de temperatura **ATM**. Así, en la figura 28, se presentan los resultados obtenidos del valor promedio de esta variable para los meses de mayo comprendidos entre el año 1997 y 2009.



Cada mes adopta un valor positivo o negativo según si, en su conjunto, la temperatura media registrada en los días correspondientes es superior o inferior al valor que sería esperable de acuerdo con los datos observados en la serie histórica.

Como ya se ha apuntado, para valorar la incidencia meteorológica en el consumo interesan sobre todo los años con anomalías de temperatura acusadas y de signo contrario, ya que serán aquellos donde previsiblemente las relaciones entre ambas variables se hagan más evidentes. Obviamente, estos años deben estar dentro del periodo abarcado por las series históricas de datos de consumos de sectores con modelo validado.

A la vista de los resultados del gráfico de la figura 28, los meses de mayo de 2008 y 2009 resultan del todo propicios para llevar a cabo el análisis. Mayo de 2009 se identifica como el más caluroso de los que componen la serie mientras que el mismo mes del año 2008 fue notablemente más frío que la media histórica. La ventaja de que se consideren años consecutivos sirve también para minimizar efectos de evolución del consumo imputables a tendencias de largo recorrido.

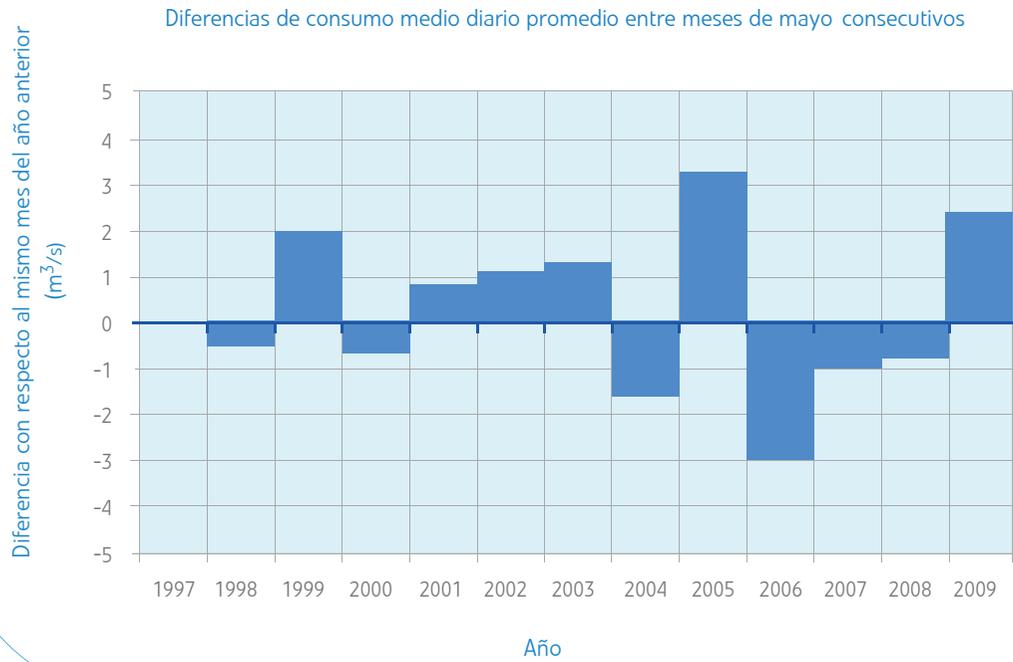
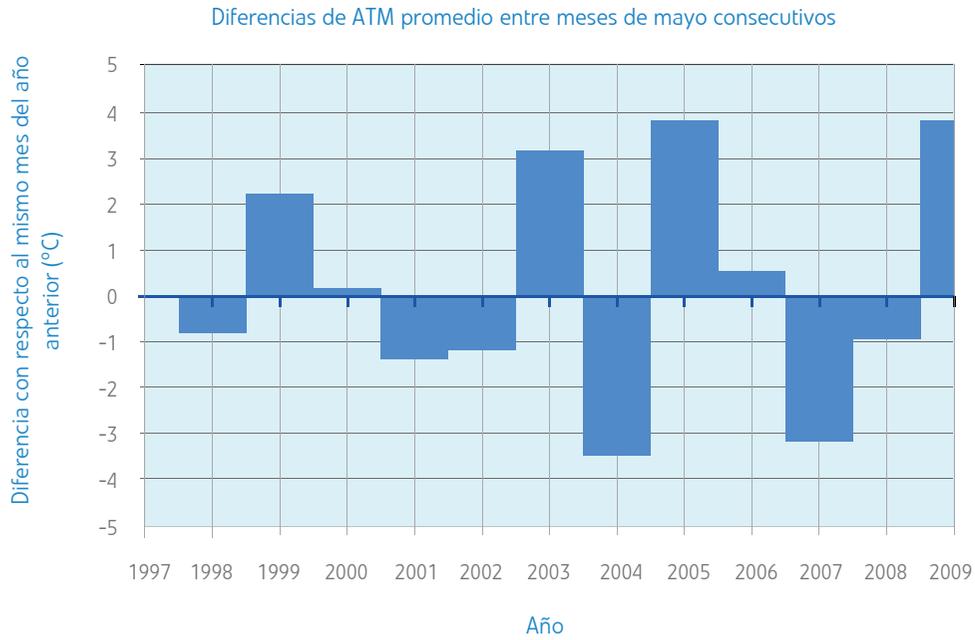
Necesariamente, estas opuestas condiciones meteorológicas han de tener su traducción en términos de consumo medido. En la figura 29 se presentan las diferencias entre los consumos promedio del conjunto de la red para cada uno de los meses de mayo de la serie histórica 1998-2009, con respecto al mismo mes del año anterior. Para evidenciar la relación de correlación positiva existente, en la otra figura que lo acompaña se recogen de manera análoga las diferencias del valor promedio del predictor **ATM**, correspondientes a meses de mayo de años consecutivos.

Las mayores diferencias meteorológicas van asociadas a variaciones de consumo igualmente destacadas. La única excepción se produce en el año 2006, donde el consumo medido desciende muy notablemente respecto al mes de mayo del año anterior, sin que exista una oscilación de temperatura consonante. Esta mala correlación se achaca a la distorsión producida por efecto de las restricciones y medidas de ahorro establecidas en la Comunidad de Madrid por efecto de la sequía acaecida entre los años 2003 y 2006. El resultado del descenso de consumo en 2006 se achaca a la entrada en vigor de estas medidas.

Entre los meses de mayo de 2008 y 2009 se observa una clara evolución positiva del consumo que, al menos en parte, deberá estar justificada por la notable diferencia meteorológica identificada. Por tanto, el análisis detallado por sectores se ha centrado en la comparación de la evolución de los consumos correspondientes a estos dos meses.

En los apartados siguientes se recogen las comparativas de los patrones de consumo horarios recogidas en los meses de mayo de 2008 y mayo de 2009 de un conjunto de sectores que representan cada una de las tipologías descritas en el capítulo 5. Como puede observarse, la manera en la que el efecto meteorológico afecta a la modulación horaria es muy distinta para cada una de las unidades de consumo que se presentan.

**Figura 29. Sintonía entre variación del consumo promedio y diferencia de anomalía de temperatura media (ATM) correspondientes al mes de mayo**



### 6.3. Respuesta individualizada por sectores al efecto meteorológico

A la hora de evaluar la manera en la que la variabilidad meteorológica registrada en los meses de mayo de 2008 y 2009 afectó al consumo de agua, se seleccionaron para el análisis exclusivamente sectores que cumplieran los siguientes requisitos:

- Serie histórica de longitud suficiente como para avalar un comportamiento homogéneo del sector. En particular se descartaron aquellas unidades de consumo en las que, entre el año 2008 y 2009, en su histórico de lecturas apareciesen escalones bruscos o tendencias marcadas.
- Coeficientes elevados del término lineal correspondientes a las variables meteorológicas. En la medida en que se acepta como válida la formulación planteada, los modelos de los sectores en los que el componente meteorológico tenga un efecto notable en el nivel de consumo se caracterizarán por asignar un peso elevado a los predictores meteorológicos en el término lineal de la modelización del consumo diario.
- Porcentaje de rechazos reducido en las validaciones efectuadas por el modelo, correspondientes a los dos meses de mayo analizados. Se le ha concedido a la modelización el "beneficio de la duda", en el sentido de que se le supone capaz de representar adecuadamente la variabilidad meteorológica. De esta manera, un porcentaje de rechazos elevado se ha interpretado como imputable a una anomalía en las lecturas, de ahí que se haya optado por descartar el análisis en estos casos. Este criterio puede calificarse de conservador, ya que es perfectamente asumible que, en un sector donde en las lecturas se observe una nítida respuesta al efecto meteorológico, el modelo pueda presentar limitaciones a la hora de realizar la validación de registros muy afectados por esta variable. Estas limitaciones pueden ser de dos tipos:
  - Si el efecto meteorológico tiene una incidencia muy marcada en los consumos de determinadas horas del día, como resultado puede producirse un rechazo de esos consumos horarios en la validación a escala intradiaria. Alternativamente, si la meteorología tuviese efecto en todas las horas del día excepto en una franja horaria determinada, entonces el rechazo podría producirse en aquellas horas no afectadas por esa variable. Estos efectos son consecuencia de la limitación de la formulación planteada, que no contempla la influencia de la meteorología en el consumo intradiario. En el caso en que por efecto de ésta se alterase el patrón horario considerado en el conjunto de referencia en magnitud suficiente, la validación horaria rechazaría aquellas lecturas que presentasen mayor desviación. Para soslayar esta posible deficiencia, los parámetros de validación deben tener la holgura suficiente como para aceptar estas posibles desviaciones.
  - A nivel diario, el rechazo de un consumo fuertemente afectado por el efecto meteorológico se producirá únicamente si los predictores considerados y sus correspondientes coeficientes obtenidos en la modelización no tienen alcance suficiente como para poder interpretar este comportamiento.

Se ha hecho, por tanto, un análisis pormenorizado de sectores con modelo validado restringido a aquellos que cumplieran las anteriores condiciones con la intención de evidenciar las oscilaciones en el patrón de consumo intradiario que pudieran ser imputables a la gran anomalía de temperatura registrada en los meses de mayo de los años 2008 y 2009.

La mejor forma de establecer esta comparación ha consistido en calcular los consumos horarios promediados entre todos los registros acumulados en cada mes. Se han determinado, por tanto, las distribuciones medias de consumo horario, distinguiendo entre días festivos y laborables para mayo de 2008 y 2009, y se han representado en un mismo gráfico con la intención de hacer más evidentes los cambios en el patrón horario producidos.

Las desviaciones en las modulaciones horarias identificadas son llamativas y varían de manera ostensible según la tipología de sector considerada. Incluso en sectores de semejantes características (uso predominante, perfil poblacional, tamaño, tipo de edificación etc.), los comportamientos observados son en ocasiones diferentes entre sí. Esta heterogeneidad puede obedecer a distintas causas. Entre las esperables a priori cabe destacar las siguientes:

- Debido a la contraria evolución meteorológica observada en los meses analizados, se espera que el efecto meteorológico tenga una repercusión importante en la demanda de agua. Sin embargo, esto no quita para que los consumos registrados puedan haber quedado alterados también por efecto de otras variables o factores no conocidos y no contemplados en la modelización. Tal sería el caso de la aparición o corrección de fugas, maniobras en la red, desplazamientos de población, incorporación al sector de nuevas unidades de consumo etc. Debe destacarse que se están comparando registros de más de un año de diferencia entre sí. En este lapso de tiempo es previsible que la configuración de determinados sectores haya sido modificada de manera significativa.
- Por otra parte, aun descartando la posibilidad de que la influencia de otras variables tuviese una repercusión relevante en el nivel de consumo en comparación con la meteorológica, también es factible que sectores aparentemente semejantes respondan de distinta manera a este componente atmosférico. Por ejemplo, el inicio del riego de jardines públicos puede retrasarse o adelantarse en relación a la variable meteorológica en función de la decisión que a este respecto adopte cada ayuntamiento. La propia meteorología puede afectar de manera distinta a cada sector (por ejemplo en el caso de tormentas en la época estival que pueden llegar a ser muy localizadas).

Más allá de estos argumentos, hay que destacar que el estudio realizado se ha basado en el análisis de las series históricas de datos relevantes (consumo de agua y variables meteorológicas) y sus relaciones estadísticas. No se han analizado, por tanto, los mecanismos por los que se materializa la influencia de la meteorología en el consumo de agua, cometido que se escapa del alcance del presente trabajo.

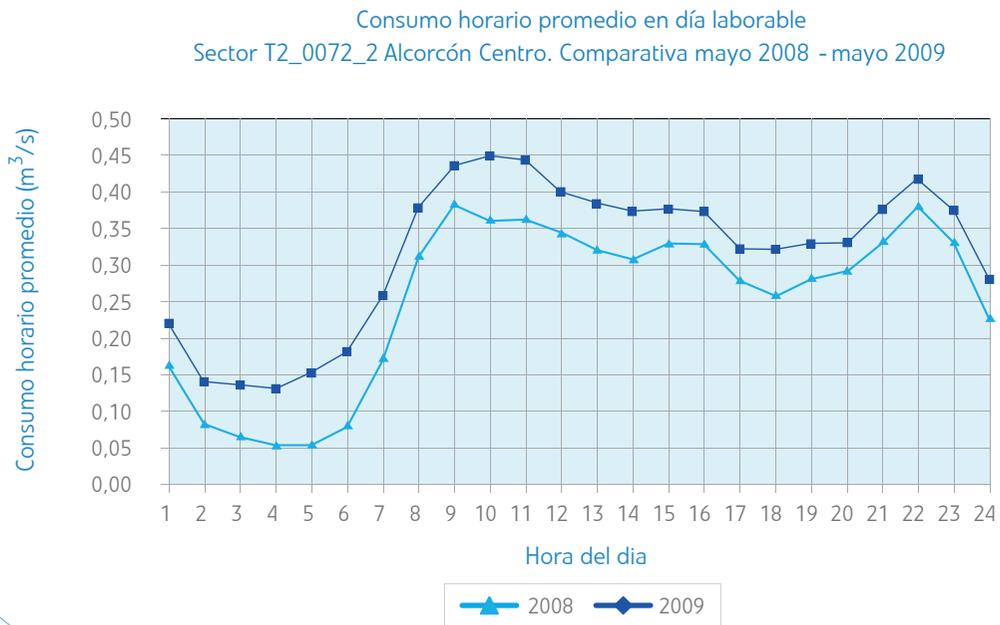
Así, la investigación ha permitido constatar y cuantificar el aumento de consumo de agua que se produce en días más calurosos de lo habitual, pero no proporciona información sobre si tal aumento se debe a un mayor consumo doméstico, a un incremento del riego de jardines o, en general, a tales o cuales hábitos de consumo que se modifican en función de la temperatura.

Es cierto que este hecho no impide que, para orientar el análisis, se hayan manejado razonamientos genéricos sobre el modo en el que la meteorología puede influir en el consumo de agua. Sobre esos mismos razonamientos se han intentado interpretar las oscilaciones en el patrón de demanda intradiaria que en la comparativa establecida han quedado patentes. Sin embargo, no teniendo un conocimiento profundo de los efectos concretos de la meteorología sobre los diferentes usos, resulta más difícil identificar en función de qué podrían estar justificados los distintos comportamientos observados en sectores de una misma tipología.

### **6.3.1. Efecto meteorológico en grandes núcleos urbanos del área metropolitana de Madrid**

La incidencia del efecto meteorológico en esta tipología de sectores es moderada. Esta atenuada respuesta es imputable a la poca presencia de demandas de agua sensibles a la evolución de la temperatura (piscinas y jardines privados). Para ilustrar este comportamiento, la figura 30 recoge dos gráficos donde se pueden comparar las curvas de consumo medio horario correspondientes a los meses de mayo de 2008 y 2009 en el sector llamado "Alcorcón Centro". Se han incluido dos gráficos independientes para reflejar el distinto comportamiento observado entre los consumos en días festivos y laborables.

**Figura 30. Efecto de la anomalía de temperatura en el sector Alcorcón Centro, comparación entre día festivo y día laborable**



En los consumos correspondientes a días festivos, llama la atención que la única desviación apreciable entre las dos curvas de consumo se produzca en las horas de la madrugada. Esta variación se imputa a baldeos de calles y riegos de parques públicos.

A priori debería entenderse, por tanto, que el resto de demandas de fin de semana no van vinculadas al componente meteorológico. Aparentemente, esto debería aplicar también a los desplazamientos de fin de semana (poco significativos en cualquier caso, como ya se apuntó en el apartado 5.2). Dada la similitud de ambas curvas, cabría pensar que éstos no se ven condicionados por la coincidencia con buen tiempo atmosférico.

Sin embargo, se comprueba que en el año 2009 el consumo promedio en días laborables es muy superior al correspondiente a días festivos, mientras que en 2008 apenas existe diferencia entre uno y otro. La disminución del consumo medio diario en festivos de 2009 en relación al registrado en días laborables invita a pensar que, efectivamente, sí se produce un mayor porcentaje de desplazamientos en puentes y fines de semana cuando las temperaturas registradas son más elevadas (así se observa en la serie histórica, la estacionalidad de la variable de calendario es patente en la figura 20 del apartado 5.2.1).

Como consecuencia de lo anterior, la conclusión final que puede hacerse en relación a los consumos en fin de semana y festivos es doble:

- Por una parte, en épocas cálidas se produce un mayor éxodo poblacional a segundas residencias.
- Por otra, todo apunta a que el consumo de la población no desplazada aumenta en una magnitud tal como para contrarrestar el descenso de demanda que sería esperable como consecuencia de los desplazamientos de fin de semana.

En días laborables se observa un aumento generalizado de los consumos en todas las horas del día, si bien éste se acentúa algo más en la franja de la madrugada. La variabilidad intradiaria derivada de la meteorología es reducida en cualquier caso.

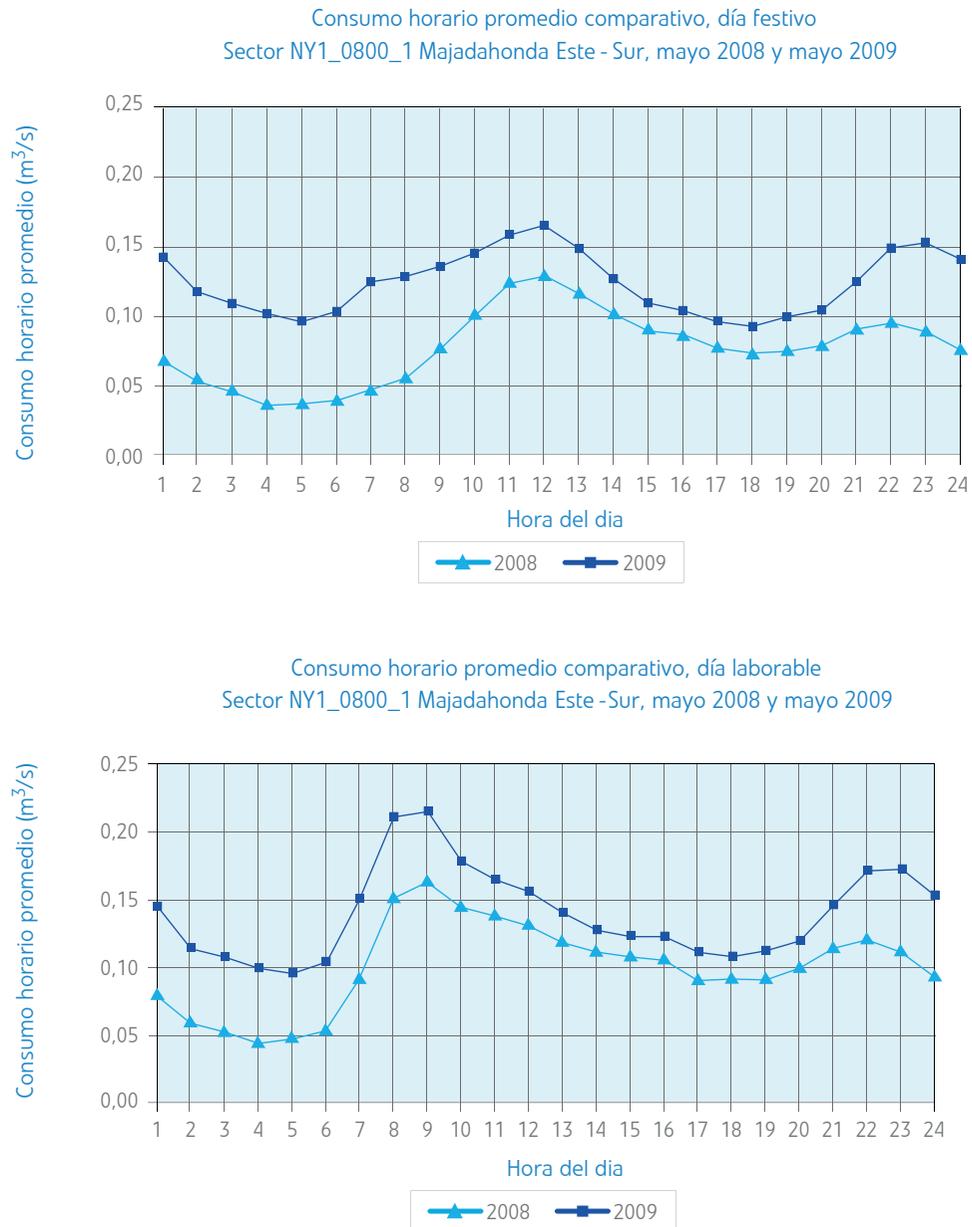
Los usos de demanda preponderantes son los de tipo doméstico junto con los correspondientes al sector terciario, por lo que no cabría esperar que la incidencia meteorológica fuese mayor en un tramo horario determinado. La uniformidad en el ascenso de los consumos avala la teoría apuntada para los días festivos, acerca de la compensación de efectos contrarios entre el aumento de consumo de la población que no viaja y la reducción de demanda derivada del éxodo de la población desplazada.

### **6.3.2. Efecto meteorológico en áreas residenciales situadas al noroeste del núcleo de Madrid**

De la misma manera que se ha hecho en el apartado anterior, la figura 31 muestra dos gráficos donde aparecen las diferencias de consumos en valores promedio entre los meses de mayo de 2008 y 2009 de un sector representativo de la tipología. Se ha diferenciado entre consumos en días laborables y en festivos, ya que la influencia de la interacción entre la variable meteorológica y la de calendario es significativa.

En los días festivos, la modulación horaria en mayo de 2009 se adapta al patrón típico de verano, es decir, aumentan los consumos en la franja de mínimos (horas de la madrugada y media tarde) y, a cambio, se reduce la magnitud del máximo de la mañana por lo que, en términos globales, la oscilación intradiaria disminuye notablemente. La atenuación de la punta matinal se atribuye al desplazamiento de la población residente, ya que se interpreta como una reducción de los consumos domésticos. Por su parte, el aumento del consumo vegetativo de los valles nocturnos y de media tarde se asocia a riegos programados para esas horas del día y al inicio del llenado y mantenimiento de piscinas. La diferencia con respecto al patrón observado en 2008, donde los máximos y mínimos quedan mucho más marcados, es notable. Más allá de la variación en el patrón, también es ostensible el aumento en términos de consumo promedio, que crece por encima del 50 por ciento en mayo de 2009 con respecto al mismo mes de 2008.

**Figura 31. Efecto de la anomalía de temperatura en el sector Majadahonda Este Sur, comparación entre día festivo y día laborable**



En días laborables, también se produce un aumento notable del consumo en el mes de mayo de 2009 con respecto al de 2008, de cerca del 40 por ciento. Aunque es generalizado, este tiene una mayor incidencia en la

punta matinal y vespertina y en las horas de la madrugada. El aumento de consumo en las horas finales del día es característica típica del comportamiento estival, igual que la atenuación del valle de la madrugada.

Como conclusión general, puede afirmarse que este tipo de sectores son muy sensibles a la variabilidad meteorológica. El aumento de temperatura registrado en el mes de mayo de 2009 altera los patrones de consumo hasta aproximarlos a los asociados a los meses de verano.

### **6.3.3. Efecto meteorológico en municipios localizados en la Sierra Norte**

En esta tipología de sectores, el efecto meteorológico presenta una incidencia que se manifiesta con mayor significación en los consumos de determinadas horas del día, lo que deriva en una alteración significativa de la modulación horaria obtenida. Se ilustra este comportamiento en la figura 32 con los consumos del sector llamado "Colmenarejo Pueblo", representativo de este conjunto.

En los días laborables, el consumo crece muy especialmente en la punta de la mañana (horas 8 y 9) y en la del final del día (horas 21 y 22). En este sentido, la respuesta al componente meteorológico recuerda a la variabilidad estacional de los patrones de consumo horario que también caracteriza a esta tipología de sectores. En estas franjas horarias, el aumento de consumo observado en mayo de 2009 es de casi el cien por ciento del medido en mayo de 2008. En comparación, es mucho menor el aumento de caudal registrado en las horas de la madrugada, lo que invita a pensar que las tareas de riego y limpieza de calles se concentran en las franjas horarias donde mayor es la demanda.

En días festivos se observa que, en mayo de 2009 el consumo crece respecto al de 2008 pero de una manera mucho más homogénea a lo largo del día. Con todo, sigue siendo perceptible un repunte en el máximo vespertino mientras que en la madrugada ambas curvas se aproximan.

El hecho de que el efecto intradiario de la meteorología no sea tan acusado, se asocia al aumento de población producido por el desplazamiento a las segundas viviendas localizadas en este municipio. En general, cuanto mayor sea la población residente en un núcleo, más facilidad habrá para que el efecto de la meteorología se haga perceptible en la demanda de consumo a nivel global. Este es el motivo por el que la demanda aumenta de manera generalizada a lo largo de todas las horas del día en fines de semana y festivos. En cambio, en días laborables, se produce un descenso de población en las horas centrales del día, ya que la oferta de trabajo en estos municipios es reducida. Como consecuencia, la demanda baja y por tanto menor es la susceptibilidad a la variable meteorológica.

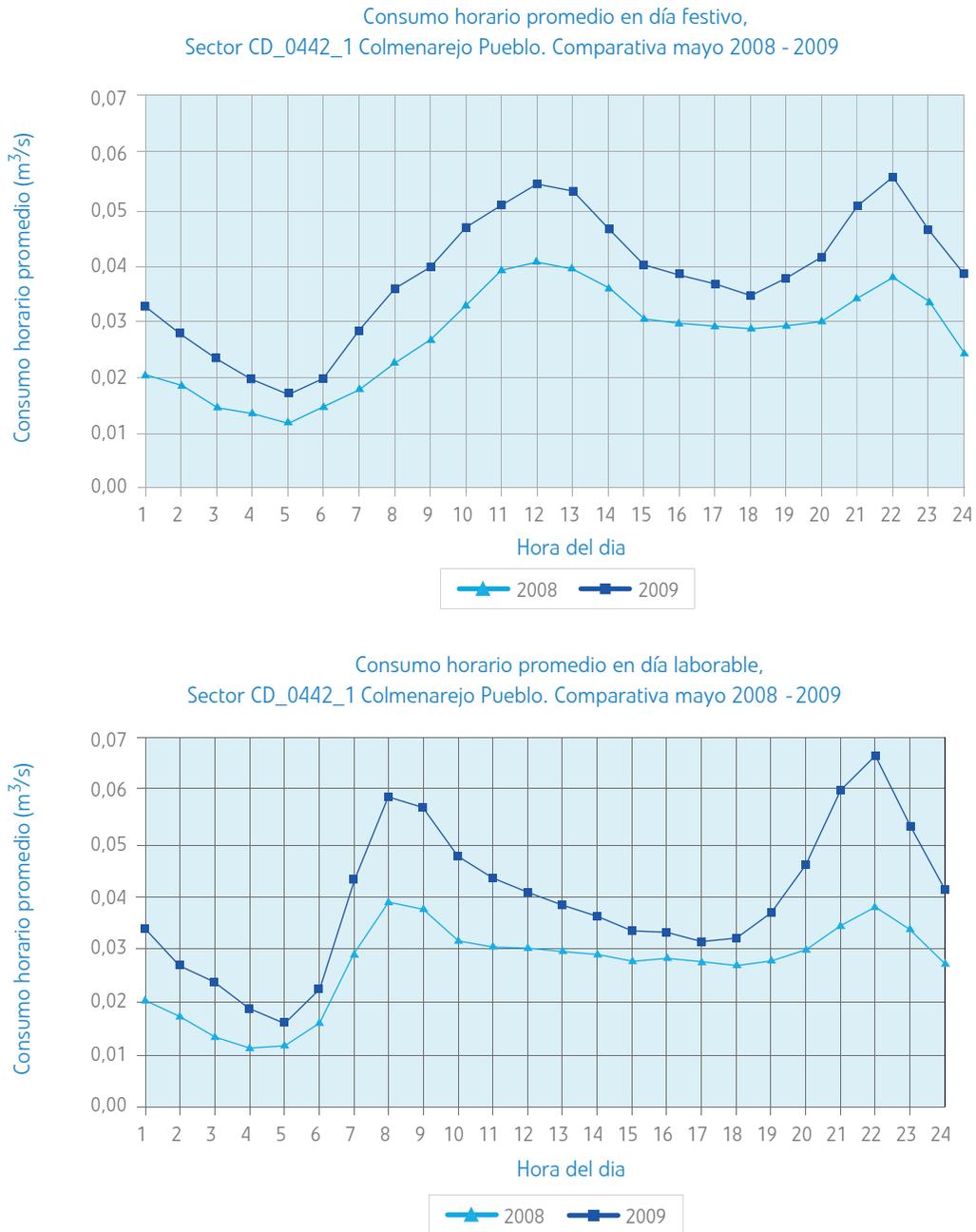
### **6.3.4. Efecto meteorológico en sectores donde predomina la actividad industrial**

Los consumos de los sectores en los que predomina la actividad industrial también muestran una respuesta clara a la influencia meteorológica. Ésta se hace evidente en los gráficos de la figura 33, correspondientes al sector llamado "Torrejón Oeste y San Fernando Industrial".

El crecimiento del consumo es incuestionable en todas las franjas horarias, si bien se intensifica en las horas centrales, coincidiendo con la jornada laboral. En estos días crece más rápido la punta de la mañana; el máximo que se alcanza en 2009 es notablemente superior al de 2008. Este aumento se imputa a mayores demandas en los procesos industriales de fabricación que allí se desarrollan. Indudablemente, las demandas de refrigeración de equipos aumentan en tiempo caluroso. Marginalmente, también aumentarán las demandas de riegos de zonas verdes que puedan localizarse y las correspondientes al personal empleado en las industrias localizadas en ese sector.

En días festivos también se observa un aumento del consumo intradiario, si bien este es más generalizado y no se intensifica en las horas centrales del día. El consumo vegetativo crece, pero de una manera más moderada.

**Figura 32. Efecto de la anomalía de temperatura en el sector Colmenarejo Pueblo, comparación entre día festivo y día laborable**



**Figura 33. Efecto de la anomalía de temperatura en el sector Torrejón Oeste y San Fernando Industrial, comparación entre día festivo y día laborable**

Consumo horario promedio en día festivo. Sector BC3\_1489\_1 Torrejón Oeste y San Fernando Industrial. Comparativa mayo 2008 - 2009



Consumo horario promedio día laborable. Sector BC3\_1489\_1 Torrejón Oeste y San Fernando Industrial. Comparativa mayo 2008 - 2009







# 7

Referencias a otros algoritmos de estimación de datos de consumo

## 7.1. Introducción

Existe una notable variedad de aproximaciones al análisis de la simulación de consumos de agua potable en núcleos urbanos. De manera general, estas pueden clasificarse de acuerdo al nivel de agregación de demanda sobre el que trabajan así como en función de la escala temporal considerada en el análisis.

Así, el estudio de demandas por sectores contemplado en la investigación presentada en este cuaderno de divulgación no es más que un ejemplo representativo de este tipo de modelizaciones. En este caso se cumple que la unidad de consumo es de un tamaño suficientemente grande como para suavizar la aleatoriedad asociada a los consumos domésticos individuales, si bien no tanto como para impedir que pueda apreciarse una diferenciación clara de patrones de acuerdo a la tipología de consumo preponderante.

Como nivel de agregación mayor, tanto a nivel de unidad de consumo como de escala temporal, se identifican los modelos de estimación de consumos de agua urbana orientados a la prognosis de demandas a largo plazo a escala municipal. Como es sabido, la utilidad de este tipo de herramientas se basa en el dimensionamiento de grandes elementos de regulación y distribución. En este tipo de modelos, el punto de partida pasa por estimar una dotación a futuro por habitante y una prognosis de la evolución de la población.

En el otro extremo, el caso de máxima desagregación puede venir representado por la caracterización de microcomponentes del consumo doméstico mediante una monitorización en continuo de consumos reales (lecturas por segundos) e individualizadas por unidad doméstica de consumo (domicilio). Mediante un adecuado tratamiento de los datos obtenidos, las mediciones permiten identificar los distintos usos domésticos (cisterna, ducha, lavavajillas, lavadora, etc.) según la forma característica de los pulsos de consumo registrados.

Más allá de una segregación por agrupaciones de demanda y escalas temporales, los modelos de estimación de demandas pueden clasificarse de acuerdo a la técnica empleada en la modelización.

Un modelo, entendido éste como una representación matemática de la realidad, se apoya en una serie de variables predictoras para hacer una estimación de cuál será el comportamiento de la variable problema, aquella cuya evolución queremos ser capaces de predecir. En el caso de que la relación establecida entre la variable dependiente y las correspondientes variables explicativas intente aproximarse a unas relaciones teóricas establecidas de antemano, de acuerdo a un sentido físico, la modelización será conceptual.

Sin embargo, debe tenerse en cuenta que la naturaleza del problema que se aborda en la modelización de demandas urbanas no tiene por qué responder a un proceso físico, o al menos no de manera exclusiva. Las necesidades de consumo de una determinada población pueden verse afectadas por factores de muy diversa índole (evolución de los niveles de renta, crecimiento demográfico, repercusión de medidas de ahorro o variaciones en la tarificación) y para los que puede ser difícil establecer los mecanismos por los que se materializa su influencia en el consumo doméstico. Esto se aplica incluso a fenómenos para los que, al menos teóricamente, debería ser más fácil establecer una base física que permitiera explicar las relaciones de interdependencia con la evolución del consumo, como es el caso de la variabilidad meteorológica. En los casos en los que no se plantee el aspirar a entender los fundamentos de la interdependencia observada, el modelo obtenido será puramente estadístico, es decir, el modelo será capaz de proporcionar una descripción del comportamiento del consumo de agua únicamente en términos de sus relaciones estadísticas con otras series de datos relevantes.

Con todo, si bien un modelo obtenido exclusivamente a partir de un análisis estadístico puede ser completamente válido para conseguir una modelización adecuada, el disponer de un conocimiento previo sobre las interpretaciones que permiten explicar el comportamiento de la variable dependiente, siempre resulta beneficioso que, además de permitir adquirir una mayor certeza sobre su capacidad para representar la realidad, puede servir para agilizar el proceso de selección de variables explicativas así como para avalar relaciones teóricas que pudieran haber sido establecidas de antemano.

La modelización y predicción de consumos de agua ha sido abordada tradicionalmente dentro del análisis de series temporales y, por lo tanto, siguiendo la metodología clásica de Box y Jenkins. El estudio desarrollado en el presente cuaderno puede integrarse dentro de esta metodología. En el apartado siguiente se realiza una revisión somera de esta técnica y de sus posibilidades de aplicación.

De manera alternativa, existen otras aproximaciones de análisis, de corte más vanguardista, que parecen haber adquirido un mayor auge en la última década. En este capítulo se ha dedicado espacio a la descripción de dos de estas herramientas: la aproximación mediante el reconocimiento de patrones de demanda y las aplicaciones realizadas a partir de redes neuronales artificiales.

## 7.2. Análisis de series temporales

El enfoque del análisis de series temporales aplicado a los modelos de consumos en núcleos urbanos se basa en considerar, de manera independiente, cada uno de los componentes que afectan a la evolución de los registros medidos. Se asume que el comportamiento global observado es suma de distintos factores cuya modelización puede abordarse por separado. De esta manera, a partir de la serie original pueden obtenerse nuevas series sintéticas en cada una de las cuales tendrá reflejo únicamente la parte de variabilidad explicada por la variable dependiente que en cada etapa se considere.

Generalmente, los componentes explicativos que siempre aparecen en las series históricas de demanda de agua son la estacionalidad, la tendencia y una oscilación cíclica que dependerá de la escala temporal considerada (consumos diarios, semanales o mensuales). A estos factores estacionarios se le añade un componente errático o, al menos, no sistemático. El aspecto final de la serie original estará formado por la suma de todos estos componentes parciales.

La tendencia suele ser reflejo de variables que tienen una influencia parsimoniosa sobre el consumo y sólo apreciable en un horizonte a largo plazo. Es el caso de la modificación de hábitos de población, variación en los niveles de renta, efecto de campañas de concienciación y ahorro, políticas tarifarias etc.

A su vez, la estacionalidad representa el patrón cíclico de la variación en el consumo de agua durante el año.

Los parámetros de las funciones que describen la tendencia y la estacionalidad se considera que, una vez estimados, pueden operar independientemente de los valores que tome la demanda de agua en un año concreto. Por este motivo, a estos dos factores se les denomina como componentes de memoria larga.

La parte de la variabilidad de la serie original no explicada a partir de estos parámetros define el componente de memoria corta. En general, este componente no sistemático, de carácter aparentemente aleatorio, debe ser analizado de manera independiente para tratar de identificar algún patrón de interés en su evolución que ayude a entender la progresión de la serie completa. Para la obtención de formulaciones en base a las que explicar el comportamiento de este componente de memoria corta, suele recurrirse a la modelización ARIMA así como al establecimiento de correlación con distintas variables climáticas.

Como puede deducirse de la lectura de los capítulos anteriores, la formulación desarrollada en la presente investigación toma como base de partida el análisis contemplado en esta metodología. Es esta una característica común de la mayor parte de las aproximaciones a este estudio realizadas hasta la aparición de las técnicas basadas en la inteligencia artificial y el auge de las redes neuronales.

Las particularidades de cada metodología se basan en la forma de acometer la partición de la serie original a partir de sus componentes estacionarios (eliminación de tendencia y desestacionalización) y en el desarrollo del posterior análisis del componente filtrado de memoria corta que ayude a entender la progresión de la serie completa.

Más allá de los grados de ajuste alcanzados, la principal virtud de esta metodología se basa en la capacidad para ir eliminando, y por tanto estimando, el porcentaje de la variabilidad de la demanda atribuible a cada uno de los factores explicativos.

Mención especial merece el análisis del componente no sistemático, de carácter aparentemente aleatorio, para tratar de identificar algún patrón de interés en su evolución que ayude a entender la progresión de la serie completa. En el ejemplo desarrollado en la presente investigación, el comportamiento de este componente del consumo se interpretó a partir de una oscilación atribuible a los efectos del calendario y a un componente relacionado con el efecto meteorológico.

Alternativamente, otras modelizaciones se basan en utilizar un modelo autorregresivo para el componente no explicado por la estacionalidad y la tendencia y luego aplicar una regresión entre los residuos obtenidos y las variables de precipitación y temperatura.

La consideración de la variable meteorológica puede implementarse en la modelización a partir de distintas aproximaciones. La relación de causalidad entre meteorología y consumo es significativa en prácticamente todas las escalas temporales y a casi cualquier nivel de agregación de demanda. Sin embargo, la caracterización de esta relación es compleja, debido a la variabilidad que puede presentar la respuesta al efecto meteorológico a lo largo del tiempo y a su carácter no lineal. Otro rasgo característico que dificulta la modelización de este componente es que el efecto de una perturbación en la meteorología puede empezar a notarse en el consumo con cierto retardo y/o prolongarse a lo largo del tiempo respecto al instante en que tiene lugar.

Debido, por tanto, al carácter variable que prevalece en esta relación, son numerosos los intentos de modelizar este fenómeno a través de herramientas de programación dinámica. Así, se han utilizado para este fin modelos de función de transferencia, buscando representar la estacionalidad del componente meteorológico, la no linealidad, y el efecto de memoria que produce en la demanda de días sucesivos.

Siguiendo esta línea de actuación, las funciones de transferencia también se han utilizado para representar los efectos de factores que tienen una incidencia muy ocasional, que se producen de manera puntual y cuyo efecto también puede entenderse como un proceso dinámico. Ejemplos de esta variedad de fenómenos pueden ser las restricciones y campañas de ahorro promulgadas en situaciones de sequía por las entidades responsables del suministro. De ahí que, en ocasiones, se haya planteado hacer uso de estas formulaciones para poder tener una estimación de la incidencia de posibles medidas a adoptar de cara a disponer de una información que pueda servir de ayuda en la toma de decisiones.

### **7.3. Aproximación a la modelización mediante técnicas de reconocimiento de patrones**

El componente cíclico que habitualmente caracteriza las curvas de demanda hace posible su análisis a través de técnicas de reconocimiento de patrones. Como ya se ha expuesto en otros apartados de este documento, son varios los factores que determinan la existencia de oscilaciones periódicas en la evolución de las series de consumo. El componente estacional que afecta a la mayor parte de los usos domésticos de agua repercute en la existencia de un patrón cíclico a escala anual. Similarmente, el efecto de calendario, entendido éste como la variabilidad de consumo que se produce entre días laborables y festivos deriva en la existencia de un componente cíclico de los consumos diarios con periodicidad semanal.

Por último, si consideramos los consumos a escala horaria, el componente cíclico del patrón de consumo intradiario se hace todavía más palpable. Si bien en el trabajo presentado ha quedado evidenciado que el patrón horario está afectado tanto por el efecto estacional como por la variable de calendario, es indudable que en todos los días del año pueden distinguirse claramente las curvas ascendentes y descendentes asociadas

a la periodicidad que caracteriza la mayor parte de los hábitos poblacionales: consumos bajos en las horas nocturnas, reactivación al inicio de la mañana, tramo oscilante en las horas centrales del día, repunte hacia el ocaso y descenso hasta alcanzar los mínimos nocturnos.

La existencia de estos patrones definidos favorece que, de manera repetitiva, se produzcan tramos diferenciados de subida y bajada en el consumo. Esta circunstancia hace factible modelizar la demanda como una sucesión de "estados" que se suceden con arreglo a una ley de probabilidad condicional de transición de uno a otro. Para analizar el comportamiento estocástico de estos sucesos, la herramienta estadística habitualmente utilizada es la de procesos en cadena de Markov.

Mekorot, la compañía de abastecimiento de agua israelí, desarrolló una aproximación en este sentido que pasa por subdividir la curva de la demanda horaria en tramos individualizados, de acuerdo a cual sea el diagnóstico sobre la evolución del consumo, en cada intervalo de tiempo considerado. Se consideraron tres posibles "estados" para representar el comportamiento de la serie histórica: "creciente", "decreciente" u "oscilante". Alternativamente, se añadió una tipología adicional, pasando a distinguirse entre "rápidamente creciente" y "moderadamente creciente", de cara a poder identificar los distintos ritmos observados entre la reactivación matinal de la demanda en días festivos y laborables.

A la par que la caracterización de posibles estados, es necesario definir las probabilidades de variación de uno a otro para cada paso del modelo (cada una de las horas del día en el caso de análisis de la demanda intradiaria). La forma de determinar esta probabilidad suele basarse en la sucesión de estados que haya tomado la variable en los intervalos de cálculo anteriores.

En el caso de que el paso a un estado futuro dependa exclusivamente del valor que tome la variable en el momento presente, el proceso se denomina de **orden 1**. Inevitablemente, la situación habitual es que el cambio hacia un estado u otro no dependa exclusivamente de la situación inmediatamente anterior sino de la trayectoria observada en una sucesión de etapas. La forma de modelar esta transición queda definida a través de un proceso de Markov, que permite transformar sucesos de orden superior a cadenas de orden 1. De la misma manera, mediante esta formulación es posible establecer una modelización capaz de predecir el estado que tomará el sistema para un alcance cualquiera, simplificándolo en forma de probabilidades de transición de paso unidad.

A partir de esta formulación, el periodo de análisis queda subdividido en segmentos donde se mantiene un mismo estado (tramos de demanda "creciente", "decreciente" u "oscilante"), debiendo establecerse en la modelización los intervalos de transición donde se produce el cambio de un estado a otro. Este proceso de determinación de estados es estocástico y para su obtención se ha asimilado su comportamiento al de una cadena de Markov. La longitud resultante de cada uno de los segmentos asociados a un determinado estado es aleatoria. En cualquier caso, un diagnóstico de la bondad del ajuste es el verificar que todas las transiciones de un estado a otro se produzcan de manera suave. De esta manera, no son admisibles oscilaciones erráticas entre estados en cortos intervalos de tiempo.

Con posterioridad a la definición de los tramos correspondientes a un mismo estado y sus puntos de transición, el siguiente paso consiste en determinar los parámetros de un modelo autorregresivo para cada uno de los posibles estados considerados. Es decir, en el estudio realizado por Mekorot se definió una función de ajuste para los puntos identificados como en zona "creciente", otra ecuación para los puntos pertenecientes a los segmentos "oscilantes" y una tercera para los tramos "descendientes".

Para la determinación de las probabilidades de transición de un estado a otro y el ajuste de los modelos autorregresivos, resulta necesario fijar un conjunto de referencia dentro de la serie histórica de demandas a partir del cual definir la modelización. Sobre esta serie de datos se establecen, de antemano y de forma manual en base a la observación, los segmentos correspondientes a cada posible estado y sus puntos de transición.

A partir de los datos registrados y de la división realizada, mediante la herramienta de modelización se calculan las probabilidades de transición de un estado a otro y los parámetros de los modelos autorregresivos definidos para cada uno de ellos.

La capacidad predictiva del modelo se basa en las probabilidades de evolución de un estado a otro para distintos alcances y en los resultados proporcionados por los modelos autorregresivos definidos para cada estado. A su vez, la distribución de probabilidades a futuro para cada uno de los posibles estados, de carácter claramente cíclico, también está sujeta a su ajuste a través de un modelo ARIMA. El modelo desarrollado por Mekorot se apoyaba en datos de demanda horaria de los 15 días pasados para establecer predicciones a un horizonte de 24 horas.

Esta modelización se concibió para permitir la operación sobre el sistema en tiempo real. Por lo tanto, el comportamiento de la demanda era monitorizado de manera continua. Las desviaciones observadas entre las predicciones del modelo y los consumos observados se corregían utilizando para futuras estimaciones información actualizada de consumos registrados.

#### **7.4. Metodologías basadas en redes neuronales artificiales**

El empleo de redes neuronales artificiales también tiene su aplicación al campo de la estimación y predicción de consumos de demanda urbana. La gran ventaja de su empleo radica en el hecho de que su mayor rendimiento se alcanza en aquellas situaciones donde mayores limitaciones presenta la modelización estadística a partir del análisis de series temporales. Es el caso, por ejemplo, de la modelización de fenómenos no lineales, donde los modelos de regresión presentan limitaciones que sólo pueden soslayarse parcialmente introduciendo términos de interacción. Análogamente, las redes neuronales permiten prescindir de las hipótesis de normalidad y no necesitan de un algoritmo definido para definir la relación causal entre variables.

Los principales desempeños de las redes neuronales para la estimación de consumos urbanos han venido relacionados con las predicciones a corto alcance de demanda diaria, escala temporal donde más difícil resulta encontrar comportamientos lineales o, al menos, asimilables a un proceso lineal. Las variables de entrada que generalmente se han utilizado en este tipo de modelizaciones suelen ser algún tipo de indicador de días festivos, la demanda registrada en los días inmediatamente anteriores y distintos conjuntos de variables meteorológicas. Los resultados obtenidos, en términos de calidad del ajuste, han sido en general destacables.

El empleo de redes neuronales puede contemplarse como una herramienta complementaria de los modelos de análisis de series temporales, o bien como un método alternativo. Visto como un complemento, existen referencias a su empleo para intentar establecer predicciones de los residuos obtenidos en una modelización de carácter lineal tipo ARIMA realizada con anterioridad: De esta manera, no resulta necesario renunciar a las ventajas que confiere la modelización estocástica y, a la vez, se aprovecha la versatilidad del aprendizaje de las redes neuronales en entornos aparentemente aleatorios.

Otra ventaja adicional de la aplicación conjunta de ambos métodos es que, al modelizarse a través de técnicas de inteligencia artificial únicamente los residuos del modelo estocástico, se acotan los márgenes de error que pueden producirse por las dificultades para garantizar un buen comportamiento de las redes neuronales en regiones no cubiertas por los datos a partir de los que se constituyó el modelo.

El uso complementario de ambas técnicas permite, además, aprovechar el conocimiento del comportamiento físico del proceso adquirido a través de la modelización estocástica para la selección de variables a emplear en el proceso de análisis por redes neuronales. En general, se obtienen mejores resultados cuando existe un buen entendimiento del proceso a modelar.

El empleo alternativo de una u otra técnica está sujeto a debate y no siempre es fácil encontrar posturas conciliadoras. Si la comparación se basa en el grado de ajuste, los resultados obtenidos a partir del análisis mediante redes neuronales pueden llegar a superar los obtenidos en procesos de series temporales, principalmente si nos encontramos en situaciones donde resulta especialmente difícil encontrar soluciones algorítmicas.

Sin embargo, tan importante o más que la bondad en el ajuste puede ser tener conocimiento del grado de incertidumbre asociado a una predicción, así como el fondo conceptual del modelo y cada una de las variables implicadas. En ese sentido, las ventajas de las técnicas basadas en series temporales son claras.

Por último, de manera inevitable un elemento que juega a favor del empleo de técnicas tradicionales es el hecho de que, aunque pueda resultar una postura contraria al avance científico, parece más seguro acudir a estos métodos avalados por la práctica antes que a estas técnicas novedosas aun en fase de desarrollo.



# 8

## Conclusiones

## 8.1. Formulación de los modelos

El punto de partida de la investigación realizada puede resumirse en los siguientes enunciados:

### Enunciados

- a) La evolución temporal del consumo de agua en la red de Canal de Isabel II presenta una serie de patrones identificables y en buena medida interpretables en función de los hábitos y actividades de la población. Se observan ciclos diarios y semanales, estacionalidad y relaciones claras con variables de calendario. También se observan relaciones, más débiles, con variables meteorológicas.
- b) Gracias a la compartimentación de la red de distribución en sectores emprendida por Canal de Isabel II, es posible disponer de un conocimiento de este comportamiento a mayor escala de detalle y, en determinadas ocasiones, individualizada por tipologías de consumo.
- c) En el nivel diario, la estacionalidad del consumo (variaciones que se producen entre unas épocas del año y otras) es evidente, si bien según la tipología del sector, ésta puede adoptar comportamientos muy diversos. El efecto de calendario es también muy claro. Se aprecia, en general, un ciclo semanal nítido, con variación en el nivel de consumo, entre días laborables y festivos. A su vez, existe interacción entre los efectos de la estacionalidad y los de calendario. La influencia de la meteorología es también clara en el nivel diario, aunque comparativamente menor que la de los efectos de calendario. El factor meteorológico también queda afectado por la estacionalidad. Por el contrario, no se ha podido evidenciar interrelación entre la meteorología y la variable de calendario. Por último, algunas de las series históricas de consumo analizadas presentan alguna tendencia a largo plazo apreciable.
- d) En el nivel horario, la mayoría de los sectores presentan también un patrón anticipable, caracterizado generalmente por un máximo matinal, mínimos nocturnos y una punta moderada en las horas finales del día. Con todo, la pauta de consumo varía de manera significativa en función de la tipología del sector considerada. Este patrón de consumo está claramente afectado por el efecto de calendario, circunstancia que se traduce en retrasos en el inicio de la curva ascendente de la mañana y en un mayor apuntamiento del consumo de las horas finales del día. El análisis ha evidenciado también la existencia de diferencias estacionales del patrón de consumo horario. Por el contrario, la variabilidad horaria asociada al efecto meteorológico, si bien existe, tiene una repercusión marginal en relación a los otros factores, por lo que se ha optado por no considerarla en la formulación del modelo adoptada.

### Resultados

A partir de los presupuestos enunciados se han obtenido diversos resultados que se sintetizan seguidamente.

- a) El desarrollo de modelos de consumo por sectores ha pasado por concebir una formulación matemática que permita describir las regularidades observadas en las series temporales de consumos y las relaciones con otras variables ya apuntadas.
- b) Como resultado general de los trabajos referidos se dispone de unos modelos de tipo autorregresivo, basados en una formulación lineal, tanto a escala diaria como en el caso de cada uno de los componentes del consumo horario, capaces de estimar o predecir el consumo horario de manera individualizada para cada sector. Estos modelos se utilizan para realizar estimaciones o predicciones de consumo a partir de una serie de predictores desarrollados de manera específica para la formulación adoptada y que se apoyan en las siguientes variables:
  - Variables de calendario
  - Variables meteorológicas
  - Consumos recientes en el sector o, en su defecto, estimaciones de los mismos

- c) En la metodología aplicada se ha seguido el criterio de que la variabilidad intradiaria del consumo se puede tratar de forma independiente de la variabilidad entre días. Este criterio ha conducido al desarrollo en paralelo de modelos en dos niveles de resolución temporal diferentes:
- Modelo horario, que simula la forma en que el consumo diario se distribuye entre las horas del día.
  - Modelo diario, que simula la evolución de los consumos diarios en escalas de tiempo iguales o superiores a un día.
- d) Todas las variables anteriores están disponibles, tanto en un escenario de validación de consumos pasados como de predicción, con hasta siete días de alcance. Por otra parte, la capacidad predictiva de los modelos viene derivada por cuanto las relaciones matemáticas que los constituyen se han establecido de modo que simulen comportamientos observados en la serie que responden a causas con muchas trazas de permanecer en el tiempo a medio o largo plazo.
- e) Además, la formulación del modelo desarrollada, es capaz de contener el conocimiento generado sobre patrones de consumo y las interpretaciones que permiten explicarlos.

## 8.2. Aplicación de los modelos

Los modelos desarrollados se diseñaron con vistas a servir de base en los procesos de validación, estimación y predicción de datos de consumo. La estimación permite proporcionar un valor de consumo independientemente del dato medido, utilizando como herramienta el modelo de consumo el cual, como ya se ha indicado, se apoya en un conjunto de predictores estacionales, de calendario y meteorológicos, y en los consumos de ese mismo sector registrados (o estimados) en fechas precedentes.

La validación se utiliza en la fase de explotación de los modelos para diagnosticar la validez de los datos de consumo obtenidos de los sistemas de medida localizados en la red de distribución. Se entiende que una lectura es válida si responde con cierta precisión a la realidad que pretende medir. Los errores en el sistema de medida o las anomalías en la realidad medida producen datos no válidos, es decir, no representativos y que por tanto deben ser descartados de la serie histórica. El rechazo de un dato genera una laguna de validación que se rellena con el dato estimado por el modelo para esa misma hora y día.

Por su formulación relativamente sencilla, parsimoniosa en cuanto a su dependencia de datos de consumo previos y con un número limitado de parámetros, los modelos tienen, en general, un comportamiento estable (no producen valores fuera de rango u oscilaciones bruscas en la serie simulada). Gracias a estas características, es posible atribuirles capacidad predictiva.

La formulación de los modelos desarrollada en la presente investigación ha resultado ser válida para representar el comportamiento de los más de 300 sectores analizados hasta la fecha de redacción del presente documento. Debe destacarse que, si bien las variables en base a las cuales se ha establecido la relación de dependencia con la evolución de los consumos son las mismas en todos los casos (tendencia, estacionalidad, calendario y meteorología), la respuesta de cada sector a una variación de estas variables explicativas puede variar en gran medida según cuál sea su tipología. En este sentido, el buen comportamiento observado en términos generales avala la versatilidad de la formulación adoptada.

El grado de precisión de estos modelos es variable, pero satisfactorio en la mayoría de los casos. En cualquier caso, la bondad del ajuste depende, en buena medida, de la calidad de la serie de datos de consumo disponible en cada caso en el momento de realizar la modelización del sector. La evaluación de la precisión de cada modelo se ha medido a partir de los resultados de la validación realizada sobre el conjunto de referencia utilizado para el ajuste de los parámetros, del término lineal y el cálculo de las funciones de estacionalidad.

Se ha seleccionado el coeficiente de regresión  $r^2$  como representativo del nivel de ajuste del modelo y el error cuadrático medio como medida del error del modelo. En conjuntos de referencia de suficiente longitud y calidad, la regresión toma valores próximos a 0,9 mientras que el error cuadrático se mantiene en el entorno de un 10 por ciento del consumo medio medido.

Valorar el nivel de ajuste de un modelo en su fase de explotación es, en general, complicado y no siempre representativo de su comportamiento, ya que está muy condicionado a la evolución de los datos de entrada. Las desviaciones observadas respecto al comportamiento observado en el conjunto de referencia pueden ser asumibles por el modelo siempre que no excedan del límite establecido en los parámetros de validación y de la rapidez de respuesta al cambio que admite el factor corrector de tendencia.

Como conclusión final, cabe destacar las ventajas derivadas de la sectorización llevada a cabo y del respaldo que para la misma constituye el desarrollo de modelos de consumo horarios obtenidos de la presente investigación:

- Detección de fugas no sintomáticas que puedan evidenciarse desde la certidumbre de consumos anormales a partir de desviaciones en las lecturas, respecto a las estimaciones del modelo.
- Establecimiento de prioridades de actuación en cuanto a tareas de mantenimiento, mejora o ampliación.
- Mejor control de la presión. Puede realizarse la regulación de presión en sectores determinados con la consiguiente disminución del volumen suministrado y del número de averías.
- Caracterización de los patrones de consumo. Profundización en el conocimiento de la demanda.
- Capacidad de prognosis, estimación de horizontes de demanda para evoluciones hipotéticas de la meteorología o de desarrollo de nuevas actuaciones urbanísticas.





Anexos

## Anexo 1. Glosario de términos

En el presente apartado se definen y explican algunos términos que se han empleado en la descripción funcional de la aplicación de validación, estimación y predicción de consumos horarios desarrollada. En particular, se detalla el modo concreto en que deben ser interpretados en el contexto del presente trabajo.

### Alcance (de una estimación o de una predicción de consumo)

Se denomina "alcance" a la distancia temporal (en días) entre la fecha del día problema (el día para el cual queremos obtener alguna información mediante el modelo) y la fecha del día en el que nos situamos para realizar las predicciones o estimaciones de caudal para los días siguientes (día 0). Viene a ser, por tanto, la anticipación en relación al momento en que se producirá el dato predicho. En CHYPRE, los alcances se expresan en días con el convenio siguiente: las predicciones para datos de un día, disponibles la mañana de ese mismo día, son de alcance 1 (predicciones de "hoy para hoy"), las disponibles para la mañana del día anterior son de alcance 2 (predicciones de "hoy para mañana"), y así sucesivamente.

A la hora de realizar una predicción, se suponen conocidos los datos de consumo y los valores de los predictores meteorológicos para el día 0 y todas las fechas anteriores. En cambio, son (o se suponen) desconocidos los datos correspondientes a los días sucesivos. El valor de los predictores meteorológicos para el día problema se obtiene a partir de las predicciones meteorológicas disponibles para esa fecha.

El sistema concebido en la aplicación CHYPRE permite realizar predicciones con un alcance máximo de siete días.

### Anomalía de temperatura

A partir de los valores de temperatura normal se define la anomalía de temperatura, para un día concreto, como la diferencia entre la temperatura registrada ese día y la temperatura normal referida al día del año correspondiente. La anomalía de temperatura es una medida de cuánto se aparta la temperatura registrada en un día cualquiera, respecto del valor normal o esperable.

### Consumo estimado o estimación de consumo

Es el consumo simulado por el modelo, para un día determinado, suponiendo conocidos los siguientes valores:

- Datos meteorológicos correspondientes al día para el que se quiere establecer la estimación y todos los anteriores.
- Consumos medios diarios vigentes correspondientes a los días previos a aquel para el que se quiere realizar la estimación.

La estimación se apoya en los consumos de los días anteriores a través del factor de tendencia, pero el valor de consumo estimado por el modelo, desarrollado para el día en cuestión, se obtiene de manera independiente al dato medido. Las estimaciones se utilizan para el relleno de lagunas y para la validación. La técnica de validación propuesta se basa en la comparación de los datos observados y sus estimaciones correspondientes.

### Consumo predicho o predicción de consumo con alcance x

Las predicciones son estimaciones de datos de consumo futuros realizadas por el modelo, esto es, datos que aún no se han medido. La única diferencia formal entre las predicciones y las estimaciones, en sentido estricto, es la información de partida que se utiliza para realizarla. La predicción se realiza para un momento futuro y, como consecuencia, la información de que se dispone en el momento de llevarla a cabo es limitada.

A la hora de establecer una predicción, se suponen desconocidos tanto los consumos vigentes como los datos meteorológicos correspondientes al periodo comprendido entre la fecha en la que nos situamos para realizar la predicción y la fecha del día problema (el día para el que queremos realizar esa predicción). La información de partida que se utiliza para esta tarea consiste en:

- Datos de consumo vigentes de los días anteriores al de partida (aquel día en el que se realiza la predicción).
- Predictores meteorológicos obtenidos a partir de predicciones meteorológicas correspondientes a los días sucesivos a aquel en el que se realiza la predicción y hasta el día para el que se quiere realizar la predicción.

### Consumo horario relativo

Se trata de una magnitud adimensional definida como el consumo medio de una hora determinada de un día, dividido por el consumo medio horario del día en cuestión. Este es el valor incógnita (predictando) empleado en la modelización de consumos en el nivel horario. Las estimaciones (predicciones) de consumo horario relativo permiten obtener estimaciones (predicciones) de consumo horario al combinarse con estimaciones (predicciones) de consumo diario.

### Estacionalidad de consumo

La estacionalidad de consumo se refiere a las variaciones en los niveles de consumo que se producen entre unas épocas del año y otras. Se puede atribuir a la estacionalidad de determinados usos del agua (riegos, piscinas, etc.) y a la estacionalidad de la población debida, principalmente, a movimientos vacacionales. Para cada sector, la curva de estacionalidad se obtiene a partir del análisis de los datos históricos. Se mide con referencia a una línea base que marca el nivel de consumo general una vez eliminado el efecto de la tendencia. Como resultado, se obtiene una función de estacionalidad del consumo ( $e_c$ ) que proporciona los distintos valores de este coeficiente para cada día del año.

### Estacionalidad de fin de semana (o estacionalidad de la festividad)

El efecto de la variable de calendario en el consumo presenta a su vez estacionalidad. De manera habitual, el efecto de los días festivos en el consumo se hace más acusado en la época estival que en la invernal. La función de estacionalidad de fin de semana sirve para modelar la interacción entre los factores de calendario y la época del año, en los casos en los que aquellos se vean influenciados por ésta.

Como dato base para estudiar la estacionalidad del calendario, se ha tomado la relación entre el consumo acumulado del domingo y el consumo medio diario de la semana correspondiente, evaluado en las semanas sin festivos. Igual que en el caso anterior, la estacionalidad de la festividad toma distintos valores para cada día del año, según cuál sea la respuesta del consumo observada en el periodo de referencia.

### Estado de validación

Se entiende por estado de validación a la clasificación asignada a cada dato de consumo entrante en relación al diagnóstico que sobre su validez determina la aplicación CHYPRE. Una de las funciones primordiales de esta herramienta consiste en la obtención de unas series de datos de consumo completas y depuradas, esto es, sin lagunas y con todos sus datos validados. Como consecuencia, las series de datos resultantes contienen datos de diverso origen: datos observados y validados, datos estimados por el modelo, consecuencia de una laguna de observación y datos estimados por el modelo, consecuencia de una laguna de validación (rechazo).

Aunque los datos de las series completas y validadas se utilizan generalmente sin atender a su origen, se considera necesario mantener en la base de datos de CHYPRE información sobre el mismo. Esta información

queda almacenada a través de códigos numéricos que van asociados a cada uno de los posibles estados de validación contemplados.

### Factor de tendencia

El factor de tendencia es un parámetro que interviene en el nivel diario de la modelización y que pretende reflejar variaciones a largo plazo del nivel de consumo, las cuales no se han contemplado de forma conceptual en el modelo.

Para adecuar los consumos simulados a los niveles observados que puedan estar afectados por el efecto de una tendencia a largo plazo, se introdujo este parámetro en la formulación. Se calcula a partir de la relación entre las estimaciones de consumo medio diario, efectuadas por el modelo para los 14 días pasados, y los correspondientes valores de consumo validado en esas fechas. A partir de esta relación el factor de tendencia toma valores distintos para cada sector modelizado y para cada día del año.

Indirectamente, el factor de tendencia sirve también para corregir desviaciones en la modelización que puedan ser imputables a otros factores distintos de las desviaciones asociadas a la variación parsimoniosa del nivel de consumo con que se identifica la tendencia.

### Predictor

Los predictores constituyen el conjunto de variables en función de los cuales se pretende representar el comportamiento de la evolución del consumo horario y diario. Se asume que los cambios en los valores de estas variables tienen una respuesta directa en los datos de consumo, que constituyen la variable dependiente. Los modelos desarrollados estiman el consumo a partir de los siguientes factores:

- Consumos de agua de fechas pasadas
- Variables de calendario
- Variables meteorológicas

Las dos últimas clases de variables (de calendario y meteorológicas) deben suministrarse a la aplicación desde fuentes externas al sistema de medida de datos de consumo. Se ha dado a ambos tipos de variables el nombre genérico de "predictores". Los valores de estas variables se han definido intentando dar representatividad a los efectos que producen sobre los niveles de consumo.

### Sector de consumo

Es la unidad de consumo resultado de la subdivisión de la red de distribución en zonas de tamaño reducido y aisladas entre sí, obtenidas mediante el control hidráulico de los flujos de entrada y salida a la red principal. Permite disponer de medidas precisas de los balances de suministros y consumos, y de los niveles de eficiencia conseguidos en áreas reducidas, de más fácil control. Como resultado se consigue, no sólo modificar la configuración de la red, sino plantear una nueva forma de gestión de la misma.

### Tablas de datos

Es una categoría del conjunto de tablas de la base de datos sobre la que trabaja la aplicación CHYPRE. Almacenan series temporales de datos organizados por registros horarios y diarios. Su tabla principal, llamada "CONSUMOS", contiene los datos de consumos horarios por zonas. Además, otra tabla recoge los valores que adoptan todas las variables predictoras consideradas en la formulación.

### Tablas de modelos

Es otra de las categorías del conjunto de tablas de la base de datos sobre la que trabaja la aplicación CHYPRE. Contienen los datos necesarios para la definición completa de los modelos de estimación de consumos: coeficientes de los modelos, parámetros de validación y funciones de estacionalidad.

### Temperatura normal

Es el valor normal de la variable de temperatura para la época del año correspondiente a la fecha. Para el cálculo del predictor de anomalía de temperatura se ha definido como "temperatura normal" de un día del año, como el promedio de las temperaturas diarias de todos los días de la misma "época del año" pertenecientes al periodo de referencia considerado (enero de 1997, a enero de 2008). Se consideran de la misma "época del año" los 15 días que van desde el séptimo anterior al séptimo posterior.

Por lo tanto, la temperatura normal correspondiente a la fecha dd/mm/aaaa es el promedio de los datos de temperatura para el conjunto de todos los días dd/mm del periodo de referencia, más aquellos que no distan más de 7 días de uno de ellos. El valor "7" como número máximo de días de separación es arbitrario.

### Validación

Se entiende por validación de un dato de consumo el proceso por el cual se diagnostica si el dato es aceptable, en el sentido de que no es demasiado discordante con el valor de consumo que sería esperable. La base de la validación es la comparación del dato con el valor estimado por el modelo. Como resultado de la validación, un dato puede ser validado (si se acepta), o rechazado (si no se acepta).

## Anexo 2. Referencias

### CANAL DE ISABEL II.

Trabajos de mejora de los procesos de validación, estimación y predicción de los consumos horarios. No publicado, 2002

### Cubillo González, F.; Moreno Rueda, T.; Ortega Les, S.

Microcomponentes y factores explicativos del consumo doméstico de agua en la Comunidad de Madrid. Cuaderno de I+D+i nº 4. Canal de Isabel II, 2008.

### Díaz Arévalo J. L.

Utilización de técnicas avanzadas en el tratamiento y manejo de datos. Aplicación a la gestión de sistemas de abastecimiento de agua. Departamento de Ingeniería Hidráulica y Medio Ambiente. Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos. Universidad Politécnica de Valencia, 2010.

### Díaz Guzmán, I.; Flores Cabeza, J.

Precisión de la medida de los consumos individuales de agua en la Comunidad de Madrid. Cuaderno de I+D+i nº 8. Canal de Isabel II, 2010.

### Espinoza Rodríguez C. A.

Predicción a corto plazo de la demanda de agua en áreas densamente pobladas. Departamento de Ingeniería Hidráulica y Medio Ambiente. Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos. Universidad Politécnica de Valencia, 2010.

### Fernández, M.; Sainctavit, L.

Estudio de potenciales de ahorro de agua en usos residenciales de interior. Cuaderno de I+D+i nº 6. Canal de Isabel II, 2009.

### Ibáñez Carranza J. C., I.; Martínez Gimeno, V.; Pérez Bueno D.

Investigación sobre potenciales de eficiencia con el empleo de lavavajillas. Cuaderno de I+D+i nº 7. Canal de Isabel II, 2009.

### Shvartser, L.; Shamir, U.; Feldman, M.

Forecasting Hourly Water Demands by Pattern Recognition Approach. Journal of Water Resources Planning and Management, 1993.

### Anexo 3. Índice de figuras

- Figura 1. Consumo medio diario en el conjunto de la red de distribución de Canal de Isabel II - Pág. 30
- Figura 2. Variabilidad a escala semanal en la serie de consumo agregado - Pág. 31
- Figura 3. Incidencia de los días festivos en el consumo medio diario. Total red distribución - Pág. 32
- Figura 4. Curva de estacionalidad de consumo correspondiente al sector de Cerceda - Pág. 36
- Figura 5. Curva de estacionalidad de fin de semana del sector de Cerceda - Pág. 37
- Figura 6. Evolución de las variables de temperatura en el año 2007 - Pág. 42
- Figura 7. Desviación en la estimación del consumo en el modelo TEC y relación con el predictor ATM - Pág. 43
- Figura 8. Transformación de la variable de precipitación en su percentil - Pág. 44
- Figura 9. Evolución del factor de estacionalidad de las variables meteorológicas - Pág. 46
- Figura 10. Consumo horario promedio de la demanda agregada - Pág. 49
- Figura 11. Efecto de la estacionalidad en el consumo horario - Pág. 51
- Figura 12. Ejemplo de distribución horaria con un dato claramente anómalo - Pág. 68
- Figura 13. Ejemplo de distribución horaria pobremente ajustada - Pág. 69
- Figura 14. Ejemplo de rechazo por modificación del consumo diario - Pág. 80
- Figura 15. Resultados de la carga de datos una vez modificado al alza el coeficiente de validación del consumo diario - Pág. 81
- Figura 16. Ejemplo de rechazo a escala horaria - Pág. 82
- Figura 17. Ejemplo de rechazo a escala horaria - Pág. 88
- Figura 18. Rechazo generalizado de datos de consumo entrante y reemplazo por estimación del modelo - Pág. 89
- Figura 19. Corrección del factor de tendencia a partir de la validación manual de consumos entrantes entre el 23 de mayo y el 6 de junio - Pág. 90
- Figura 20. Histórico de consumos medios diarios en día laborable y fin de semana. Sector Alcorcón Centro - Pág. 96
- Figura 21. Patrón de consumo intradiario, según época del año. Sector Alcorcón Centro - Pág. 98
- Figura 22. Histórico de consumos medios diarios en día laborable y fin de semana. Sector Majadahonda Este Sur - Pág. 100

[Figura 23](#). Patrones de consumo intradiario, según la época del año.  
Sector NY1\_0800\_1 Majadahonda Este Sur - Pág. 102

[Figura 24](#). Histórico de consumos medios diarios en día laborable y fin de semana.  
Sector Hoyo de Manzanares Pueblo - Pág. 104

[Figura 25](#). Patrones de consumo intradiario, según la época del año. Sector Hoyo de Manzanares Pueblo - Pág. 106

[Figura 26](#). Histórico de consumos medios diarios en día laborable y fin de semana.  
Sector Alcorcón Industrial - Pág. 108

[Figura 27](#). Patrones de consumo intradiario, según la época del año. Sector Alcorcón Industrial - Pág. 109

[Figura 28](#). Histórico del valor promedio del predictor ATM para el mes de mayo - Pág. 116

[Figura 29](#). Sintonía entre variación del consumo promedio y diferencia de anomalía de temperatura media (ATM) correspondientes al mes de mayo - Pág. 118

[Figura 30](#). Efecto de la anomalía de temperatura en el sector Alcorcón Centro, comparación entre día festivo y día laborable - Pág. 121

[Figura 31](#). Efecto de la anomalía de temperatura en el sector Majadahonda Este Sur, comparación entre día festivo y día laborable - Pág. 123

[Figura 32](#). Efecto de la anomalía de temperatura en el sector Colmenarejo Pueblo, comparación entre día festivo y día laborable - Pág. 125

[Figura 33](#). Efecto de la anomalía de temperatura en el sector Torrejón Oeste y San Fernando Industrial, comparación entre día festivo y día laborable - Pág. 126



Canal de   Isabel II

Canal de  Isabel II

[www.cyii.es](http://www.cyii.es)

Canal de Isabel II  
Santa Engracia, 125. 28003 Madrid