



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola Superior d'Enginyeries Industrial,
Aeroespacial i Audiovisual de Terrassa

TRABAJO FIN DE MÁSTER

ESTUDIO DE VALIDACIÓN, CORRECCIÓN Y
ALMACENAMIENTO DE DATOS DE LOS
CAUDALÍMETROS DE LA RED DE DISTRIBUCIÓN DE
AGUA DE TERRASSA

Autor: Ivan Campaña Soler
Director: Joseba-Jokin Quevedo Casin
Alejandro Bachiller Matarranz
Titulación: Máster Universitario en Ingeniería de Sistemas Automáticos y Electrónica
Convocatoria: Primavera 2022



Agradecer en primera instancia a Joseba, Alejandro y Joan por su ayuda y guía en el proyecto, y a la empresa TAIGUA por proporcionar la información necesaria para llevarlo a cabo.

Asimismo, agradecer a Toni, por el largo camino juntos y a familiares y pareja por el apoyo incondicional.

Resumen

En este presente documento se presenta el análisis de validación y reconstrucción de los datos recogidos por los caudalímetros de la red de distribución de Terrassa, concretamente del piso de Sulleva.

Debido a la complejidad de la red de distribución así como sus dimensiones, cobra de gran importancia realizar un estudio de fiabilidad de los datos con tal de poder encontrar minimizar las pérdidas de agua o detectar un funcionamiento erróneo de alguno de los sensores. Es por ello que TAIGUA ha proporcionado datos de 2019 para poder realizar un estudio exhaustivo de los datos.

En este proyecto se presenta el estudio de los diferentes métodos de validación dependiendo de su procedencia. Asimismo se analizan y se examinan las casuísticas de cada uno de los valores mediante los tests de validación establecidos para decretar su validez.

Con los valores inválidos se procederá a la reconstrucción, donde dependiendo de la causa de error se corregirá mediante su método correspondiente. Estos métodos serán determinados mediante un estudio previo comparando las diferentes metodologías posibles.

Por último, se añadirán los valores a una base de datos con el fin de almacenarlos de un modo más accesible mediante las consultas SQL.

Resum

En aquest document es presenta l'anàlisi de validació i reconstrucció de les dades recollides pels cabalímetres de la xarxa de distribució de Terrassa, concretament del pis de Sulleva.

A causa de la complexitat de la xarxa de distribució així com les seves dimensions, cobra de gran importància fer un estudi de fiabilitat de les dades per tal de trobar minimitzar les pèrdues d'aigua o detectar un funcionament erroni d'algun dels sensors. És per això que TAIGUA ha proporcionat dades del 2019 per poder realitzar un estudi exhaustiu de les dades.

En aquest projecte es presenta l'estudi dels diferents mètodes de validació depenent de la seva procedència. Així mateix, s'analitzen i s'examinen les casuístiques de cadascun dels valors mitjançant els tests de validació establerts per decretar-ne la validesa.

Amb els valors invàlids es procedirà a la reconstrucció, on depenent de la causa de l'error es corregirà mitjançant el seu mètode corresponent. Aquests mètodes seran determinats mitjançant un estudi previ comparant les diferents metodologies possibles.

Finalment, s'afegiran els valors a una base de dades per emmagatzemar-los de manera més accessible mitjançant les consultes SQL.



Abstract

This document presents the analysis of validation and reconstruction of the data collected by the flow meters of the distribution network of Terrassa, specifically of the Sulleva floor.

Due to the complexity of the distribution network and its dimensions, it is of great importance to carry out a study of the reliability of the data in order to minimise water losses or detect the malfunctioning of any of the sensors. For this reason, TAIGUA has provided data for 2019 in order to carry out an exhaustive study of the data.

This project presents the study of the different validation methods depending on their origin. Furthermore, the casuistics of each of the values are analysed and examined by established validation tests in order to decree their validity.

Invalid values are reconstructed and, depending on the cause of the error, corrected by the corresponding method. These methods will be determined by means of a previous study comparing the different possible methodologies.

Finally, the values will be added to a database in order to store them in a more accessible way by means of SQL queries.

Índice

RESUMEN	II
RESUM	II
ABSTRACT	III
ÍNDICE	IV
ÍNDICE DE TABLAS	VI
ÍNDICE DE FIGURAS	VI
LISTA DE ABREVIATURAS	VIII
1. INTRODUCCIÓN	1
1.1 DESCRIPCIÓN DEL PROYECTO	1
1.2 JUSTIFICACIÓN DEL PROYECTO	2
1.3 OBJETO	2
1.4 OBJETIVOS DEL PROYECTO	3
1.5 ALCANCE	3
2 DESCRIPCIÓN DE LA PROBLEMÁTICA	4
2.1 TAIGUA	4
2.2 RED DE DISTRIBUCIÓN	5
2.3 PISO DE SULLEVA	6
3 TRATAMIENTO E IMPORTACIÓN DE LOS DATOS	10
3.1 ESTUDIO DE LOS DATOS	12
3.2 SOLUCIÓN ADOPTADA	13
4 ESTUDIO Y ELECCIÓN DE METODOLOGÍAS DE VALIDACIÓN	15
4.1 METODOLOGÍAS DE VALIDACIÓN	15
4.2 CC0201 – BOMBEO AL DEPÓSITO	17
4.3 NV_SU1 Y NV_SU2 – NIVELES DE DEPÓSITO	20
4.4 CC0202 – SALIDA DEL DEPÓSITO	22
4.5 CC0210 Y CC0211 – SECTOR S0201	26
4.6 CC0213 Y CC0216 – SECTOR S0203	30
5 RECONSTRUCCIÓN DE LOS DATOS	34
5.1 METODOLOGÍAS DE RECONSTRUCCIÓN	34
5.2 CC0201 – BOMBEO AL DEPÓSITO	35
5.3 NV_SU1 Y NV_SU2 – NIVELES DE DEPÓSITO	38
5.4 CC0202 – SALIDA DEL DEPÓSITO	39
5.5 CC0210 Y CC0211 – SECTOR S0201	40
5.6 CC0213 Y CC0216 – SECTOR S0203	44
6 ALMACENAMIENTO DE DATOS	48
6.1 MODELO ENTIDAD-RELACIÓN	48
6.2 MODELO RELACIONAL	50
6.3 CONEXIÓN DESDE MATLAB	51
6.4 CREACIÓN Y CARGA DE LA BBDD	52
7 RESULTADOS	56
7.1 CASO 1: 01/01/2019 – 14/01/2019	56
7.2 CASO 2: 13/04/2019 – 26/04/2019	63
7.3 CASO 3: 14/07/2019 – 27/07/2019	67
7.4 CASO 4: 27/09/2019 – 10/10/2019	71
7.5 CASO 5: 05/12/2019 – 17/12/2019	73



8	CONCLUSIONES	76
8.1	POSIBLES MEJORAS.....	77
9	REFERENCIAS.....	80
ANEXOS	82	
	ANEXO I: TRATADO, VALIDACIÓN Y RECONSTRUCCIÓN	82
	ANEXO II: ESTUDIO CONSUMO DIARIO	93
	ANEXO III: CREACIÓN E INSERT EN BBDD	96

Índice de tablas

TABLA 1. INFORMACIÓN DE LOS CAUDALÍMETROS EN EL PISO DE SULLEVA.	10
TABLA 2. REGISTROS SIN TRATAR EN SULLEVA.	11
TABLA 3. DESCRIPCIÓN DE LA VALIDACIÓN CC0201.	19
TABLA 4. DESCRIPCIÓN DE LA VALIDACIÓN DE NIVEL DEL DEPÓSITO.	21
TABLA 5. DESCRIPCIÓN DE LA VALIDACIÓN DE NIVEL DEL DEPÓSITO.	25
TABLA 6. DESVIACIONES ESTÁNDAR POR TRAMOS EN S0202.	28
TABLA 7. DESCRIPCIÓN DE LA VALIDACIÓN PARA CC0210 Y CC0211.	29
TABLA 8. DESVIACIONES ESTÁNDAR POR TRAMOS EN S0204.	32
TABLA 9. DESCRIPCIÓN DE LA VALIDACIÓN PARA CC0213 Y CC0216.	33
TABLA 10. DESCRIPCIÓN DE LOS RESULTADOS DE LA VALIDACIÓN.	56

Índice de figuras

FIGURA 1. DESCRIPCIÓN DEL PROYECTO.	1
FIGURA 2. MEDIDA ANUAL EN 2019 DEL CAUDALÍMETRO CC0202.	2
FIGURA 3. DISTRIBUCIÓN DE LOS PISOS DE LA RED DE DISTRIBUCIÓN EN TERRASSA.	5
FIGURA 4. RED DE DISTRIBUCIÓN DE ENTRADA DE AGUA A SULLEVA.	6
FIGURA 5. ESQUEMÁTICO DEL PISO DE SULLEVA.	7
FIGURA 6. SECTOR S0201.	7
FIGURA 7. SECTOR S0202.	8
FIGURA 8. SECTOR S0203.	8
FIGURA 9. SECTOR S0204.	9
FIGURA 10. TRATAMIENTO DE DATOS BIMINUTAL.	13
FIGURA 11. EXCEPCIÓN EN EL TRATAMIENTO DE DATOS.	14
FIGURA 12. RESULTADO DE TRATAMIENTO DE DATOS QUINCEMINUTAL.	14
FIGURA 13. PROCESO DE LA VALIDACIÓN DE DATOS.	16
FIGURA 14. DATOS ANUALES DEL CAUDALÍMETRO CC0201.	17
FIGURA 15. PRIMERA PRUEBA DE VALIDACIÓN CC0201.	18
FIGURA 16. VALIDACIÓN DE SENSOR CC0201.	20
FIGURA 17. REGISTROS DE LOS NIVELES DEL DEPÓSITO.	20
FIGURA 18. VALIDACIÓN DE SENSOR DE NIVEL.	22
FIGURA 19. DATOS DEL CAUDALÍMETRO CC0202.	22
FIGURA 20. SECTOR S0202.	23
FIGURA 21. CÁLCULO SECTOR S0202.	23
FIGURA 22. VALORES REGISTRADOS Y CALCULADOS (ARRIBA) Y DIFERENCIAS PORCENTUALES EN VALOR ABSOLUTO DEL CC0202 (ABAJO).	24
FIGURA 23. VALIDACIÓN SENSOR CC0202.	25
FIGURA 24. SECTOR S0201.	26
FIGURA 25. DATOS DE CONSUMO S0201.	26
FIGURA 26. REGLA DE 68-95-99.7.	27
FIGURA 27. MEDIA ARITMÉTICA, LÍMITE SUPERIOR Y LÍMITE INFERIOR DEL S0201.	28
FIGURA 28. ENTRENAMIENTO MÉTODO DE VALIDACIÓN S0201.	29
FIGURA 29. VALIDACIÓN S0201.	30
FIGURA 30. SECTORES S0203 Y S0204.	30
FIGURA 31. CONSUMO EN EL SECTOR S0204.	31
FIGURA 32. CONSUMO EN EL SECTOR S0203.	31
FIGURA 33. MEDIA ARITMÉTICA, LÍMITE SUPERIOR Y LÍMITE INFERIOR EN S0203.	32
FIGURA 34. ENTRENAMIENTO MÉTODO DE VALIDACIÓN S0203.	33
FIGURA 35. VALIDACIÓN CC0213 Y CC0214.	33
FIGURA 36. PROCESO DE RECONSTRUCCIÓN.	35
FIGURA 37. PRIMERA RECONSTRUCCIÓN DEL SENSOR CC0201.	37
FIGURA 38. SEGUNDA RECONSTRUCCIÓN DEL SENSOR CC0201.	38
FIGURA 39. RECONSTRUCCIÓN DE SENSORES DE NIVEL DEL DEPÓSITO.	39
FIGURA 40. RECONSTRUCCIÓN DEL SENSOR CC0202.	40



FIGURA 41. DATOS REGISTRADOS EN S0201 SIN RECONSTRUIR.	41
FIGURA 42. RECONSTRUCCIÓN S0201 CON AR.	42
FIGURA 43. RELACIÓN ENTRE CC0210 Y CC0211.	43
FIGURA 44. PROCESO DE RECONSTRUCCIÓN S0201.	44
FIGURA 45. RECONSTRUCCIÓN CC0210 Y CC0211.	44
FIGURA 46. DATOS EN S0203 SIN RECONSTRUIR.	45
FIGURA 47. RECONSTRUCCIÓN DEL SECTOR S0203 CON AR.	46
FIGURA 48. PROCESO RECONSTRUCCIÓN S0203.	46
FIGURA 49. RECONSTRUCCIÓN DE CC0213 Y CC0216.	47
FIGURA 50. PROCESO DE CREACIÓN BBDD.	48
FIGURA 51. DIAGRAMA DE MODELO ENTIDAD-RELACIÓN.	50
FIGURA 52. DIAGRAMA DE MODELO RELACIONAL.	51
FIGURA 53. ODBC DATA SOURCE.	52
FIGURA 54. RELACIONES DE LAS TABLAS EN ACCESS.	53
FIGURA 55. TABLA DE PISOS.	53
FIGURA 56. TABLA DE SENSOR.	54
FIGURA 57. TABLA DE DATOS.	54
FIGURA 58. EJEMPLO DE CONSULTA EN LA BBDD.	55
FIGURA 59. CASO 1 VALIDACIÓN CC0201.	57
FIGURA 60. CASO 1 RECONSTRUCCIÓN CC0201.	57
FIGURA 61. CASO 1 VALIDACIÓN NV2.	58
FIGURA 62. CASO 1 RECONSTRUCCIÓN NV2.	58
FIGURA 63. CASO 1 VALIDACIÓN CC0210.	59
FIGURA 64. CASO 1 RECONSTRUCCIÓN CC0210.	59
FIGURA 65. CASO 1 VALIDACIÓN CC0211.	60
FIGURA 66. CASO 1 RECONSTRUCCIÓN CC0211.	60
FIGURA 67. CASO 1 VALIDACIÓN CC0213.	61
FIGURA 68. CASO 1 RECONSTRUCCIÓN CC0213.	61
FIGURA 69. CASO 1 VALIDACIÓN CC0216.	62
FIGURA 70. CASO 1 RECONSTRUCCIÓN CC0216.	62
FIGURA 71. CASO 2 VALIDACIÓN CC0201.	63
FIGURA 72. CASO 2 RECONSTRUCCIÓN CC0201.	63
FIGURA 73. CASO 2 VALIDACIÓN NV1.	64
FIGURA 74. CASO 2 RECONSTRUCCIÓN NV1.	64
FIGURA 75. CASO 2 VALIDACIÓN CC0202.	65
FIGURA 76. CASO 2 RECONSTRUCCIÓN CC0202.	65
FIGURA 77. CASO 2 VALIDACIÓN CC0216.	66
FIGURA 78. CASO 2 RECONSTRUCCIÓN CC0216.	66
FIGURA 79. CASO 3 VALIDACIÓN NV2.	67
FIGURA 81. CASO 3 RECONSTRUCCIÓN NV2.	67
FIGURA 81. CASO 3 VALIDACIÓN CC0202.	68
FIGURA 82. CASO 3 RECONSTRUCCIÓN CC0202.	68
FIGURA 83. CASO 3 VALIDACIÓN CC0210.	69
FIGURA 84. CASO 3 RECONSTRUCCIÓN CC0210.	69
FIGURA 85. CASO 3 VALIDACIÓN CC0211.	70
FIGURA 86. CASO 3 RECONSTRUCCIÓN CC0211.	70
FIGURA 87. CASO 4 VALIDACIÓN CC0216.	71
FIGURA 88. CASO 4 RECONSTRUCCIÓN CC0216.	71
FIGURA 89. CASO 4 VALIDACIÓN S0203.	72
FIGURA 90. CASO 5 RECONSTRUCCIÓN S0203.	72
FIGURA 91. CASO 5 VALIDACIÓN CC0201.	73
FIGURA 92. CASO 5 RECONSTRUCCIÓN CC0201.	73
FIGURA 93. CASO 5 VALIDACIÓN CC0210.	74
FIGURA 94. CASO 5 RECONSTRUCCIÓN CC0210.	74
FIGURA 95. CASO 5 VALIDACIÓN CC0211.	75
FIGURA 96. CASO 5 RECONSTRUCCIÓN CC0211.	75
FIGURA 98. PISO DE CAN COLOMER.	78

Lista de abreviaturas

Abreviatura	Descripción
TAIGUA	Aigua Municipal de Terrassa
BBDD	Base de datos
EPEL	Entidad pública empresarial local
OAT	Observatorio del Agua de Terrassa
CSV	Comma Separated Values
AR	Modelo autorregresivo
PK	Primary Key
FK	Foreign Key
ODBC	Open DataBase Connectivity
JDBC	Java DataBase Connectivity
SQL	Structured Query Language

1. Introducción

El presente documento recoge el Trabajo Final de Máster en Ingeniería de Sistemas Automáticos y Electrónica Industrial, especializado en Tecnologías de la Producción y Automatización Avanzadas, realizado por Ivan Campaña Soler, siendo los codirectores Joseba-Jokin Quevedo Casin y Alejandro Bachiller Matarranz en el periodo que comprende desde febrero hasta junio del presente año 2022.

El proyecto se centra en el estudio de validación, corrección y almacenamiento de datos de caudalímetros de la red de distribución de agua de Terrassa.

1.1 Descripción del proyecto

En este proyecto se realiza un estudio de los datos obtenidos por los caudalímetros de la compañía TAIGUA, encargada de la red de distribución de agua de Terrassa.

Con el objetivo de poder estudiar el sistema, TAIGUA ha proporcionado los datos del año 2019 de sensores de la red de distribución de agua de Terrassa, especialmente del piso de Sulleva.

Partiendo de estos datos, se pretende crear un modelo mediante Matlab que valide los datos, según el estado del sensor o el historial de consumo en los sectores del piso de Sulleva, para posteriormente reconstruir o corregir dicho dato erróneo.

Por último, se crea una base de datos con el fin de almacenar y tener una mayor accesibilidad a los datos obtenidos antes y después de los procesos, tanto de validación como de reconstrucción.

Finalmente, se concluye el proyecto mostrando los resultados del modelo aplicado, verificando el correcto funcionamiento y proporcionando posibles mejoras respecto, así como los anexos donde se incluirá el código implementado.

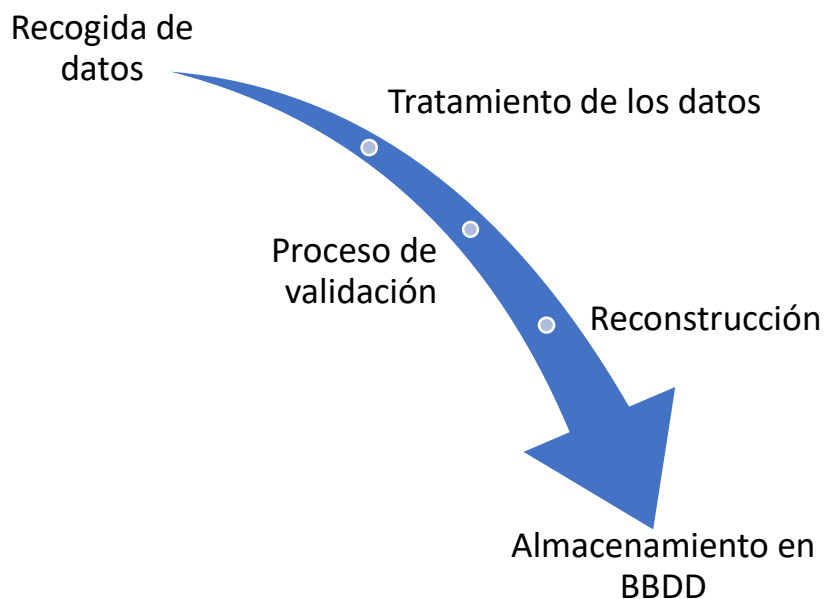


Figura 1. Descripción del proyecto.

1.2 Justificación del proyecto

TAIGUA opera con la red de distribución de agua que abastece a Terrassa y otras pequeñas ciudades de los alrededores. Se trata de una red compleja debido a su gran escala y su distribución geográfica, y los datos proporcionados por los sensores pueden verse afectados por el mal funcionamiento del propio sensor, errores de comunicación, ruido, malas calibraciones o simplemente lecturas erróneas. Es por ello que se necesita de un control a tiempo real que garantice un alto nivel de fiabilidad.

Es de gran importancia tener una gran fiabilidad en la información procedente de los caudalímetros con el objetivo de optimizar el consumo de energía y reducir las pérdidas de agua, de esta manera, aprovechando la información proporcionada por los diferentes sensores se puede conocer la cantidad de agua que se distribuye por cada zona de la red y conocer el estado del sistema al momento. Toda esta información es esencial para garantizar un suministro adecuado a los consumidores.

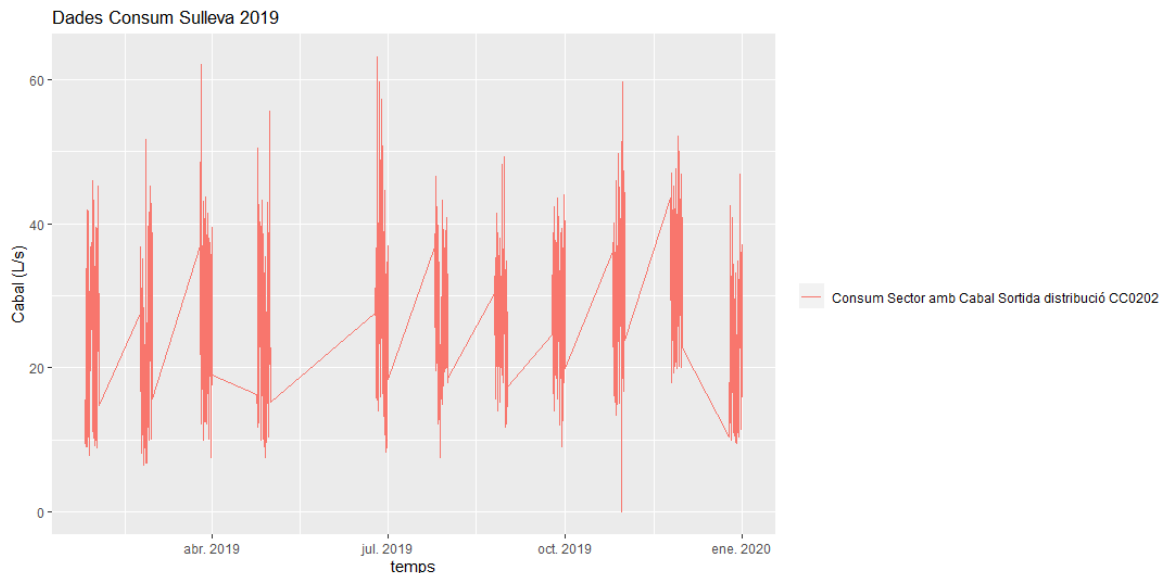


Figura 2. Medida anual en 2019 del caudalímetro CC0202.

En la imagen anterior se muestra los datos recogidos en 2019 por TAIGUA en el caudalímetro CC0202 del piso de Sulleva. Tal y como se observa, hay una importante falta de datos, ya que únicamente existen datos en 11 intervalos de tiempo durante todo el año. Esa gran carencia de datos pone en riesgo a todo el sistema, ya que en esos tramos vacíos puede haber algún fallo en la red de distribución y dificulta que se detecte.

Por ello, es necesario encontrar un sistema que sea capaz de determinar la validez los datos procedentes de los diferentes sensores o caudalímetros. Así, utilizando diferentes técnicas de reconstrucción, poder estimar el dato erróneo o no existente con la finalidad de conseguir una base de datos fiable.

1.3 Objeto

El objeto del proyecto es encontrar las técnicas de validación óptimas para validar los datos proporcionados por la red de distribución de TAIGUA y conseguir un procedimiento de validación fiable.

Por otro lado, se deberá realizar el estudio independiente de cada sensor o depósito para conocer sus características propias. Con ello, mediante la utilización de diferentes modelos se podrá reconstruir o corregir el dato inválido correctamente y almacenarlo de forma estructurada y adecuada.

1.4 Objetivos del proyecto

Los objetivos del proyecto son los siguientes:

Objetivo general:

- Encontrar un procedimiento de validación de la red de caudalímetros que pueda ser utilizado para garantizar la fiabilidad de los datos y reconstruir los datos inválidos y almacenarlo en una base de datos.

Objetivos específicos:

- Tratamiento de datos brutos.
- Estudio de los diferentes métodos de validación a aplicar.
- Implementación de la validación de datos mediante Matlab.
- Estudio de la reconstrucción de los datos.
- Implementación de la corrección y reconstrucción de datos mediante Matlab.
- Desarrollo e implementación de una BBDD.

1.5 Alcance

En este proyecto, se pretende realizar un análisis de la fiabilidad de los datos registrados por los caudalímetros de una parte de la red de distribución de agua de Terrassa para conocer qué datos pueden ser inválidos.

Primeramente, se debe establecer la manera eficiente de tratar los datos. Los datos entre sí, pueden ser de diferente periodicidad y magnitud, por lo que se debe encontrar cuál es el modo más cómodo para trabajar.

Posteriormente, se realiza el estudio exhaustivo de los datos en cuestión para conocer su comportamiento durante los diferentes intervalos del día y la relación entre los diferentes sensores. Con ello, se podrán establecer los diferentes métodos de validación según de qué sensor se trate.

Después, con tal de conseguir el mayor número de datos fiables, siempre que sea posible se reconstruirán los datos inexistentes o invalidados en la validación. Se estudiará cuáles son los mejores métodos de reconstrucción según el caudalímetro.

Por último, se creará una BBDD que sea intuitiva y permita acceder a los datos iniciales y finales, conociendo su valor, ubicación y estado de validez, reconstruido o corregido.

2 Descripción de la problemática

En este apartado se pretende mostrar un contexto del proyecto explicando la procedencia de los datos, conociendo la empresa que los gestiona. En él se detallarán los diferentes sensores que contiene la red de distribución de agua de Terrassa, así como los pisos y sectores en los que está dividido.

2.1 TAIGUA

TAIGUA [1], es la entidad pública empresarial local (EPEL) que se encarga de gestionar la red de distribución de agua de Terrassa, teniendo como objetivo ofrecer un servicio eficiente y de calidad, velando por los intereses de la ciudadanía egarenca. En los Estatutos de TAIGUA [2] se encuentra la siguiente definición respecto a sus objetivos y funciones:

“Constituye el objetivo de la Entidad y, para ello su competencia, la gestión y prestación del servicio público de abastecimiento de agua potable a la ciudad de Terrassa, que comprende: la regulación en el marco de la normativa básica que pueda establecer y aprobar el Ayuntamiento de Terrassa; la captación, el tratamiento, la potabilización, la conducción, el almacenamiento y la distribución y abastecimiento de agua potable, así como todas aquellas operaciones y actividades necesarias, conexas y consecuentes con las anteriores, relacionadas con el ciclo del agua.”

Anteriormente, el abastecimiento de agua había sido un servicio público gestionado por la empresa Mina Pública de Aguas de Terrassa, gracias a una concesión del Ayuntamiento de Terrassa que realizó en 1941. Tras 75 años, en 2016, al finalizar la concesión y con ello su servicio, el Ayuntamiento de Terrassa estudió las diferentes alternativas de gestión de servicio y concluyó que la gestión pública directa era más eficiente y sostenible económicamente que la gestión de una empresa privada, además de ofrecer más transparencia y participación.

Así, a mediados de 2018, se aprobó la gestión mediante una entidad pública empresarial local TAIGUA. Por otro lado, también se creó el OAT (Observatorio del Agua de Terrassa) con el objetivo de mejorar la participación, colaboración y concertación con la ciudadanía, el cual tiene la capacidad de alcanzar acuerdos que deberán ser considerados por el Gobierno Municipal.

En cuanto a los servicios proporcionados por TAIGUA, destacan los siguientes:

- Red en alta: Consiste en la captación, tratamiento y transporte de agua hasta los depósitos.
- Red en baja: Consiste en la regulación de los depósitos y la distribución a los domicilios.
- Realización de todas las obras de infraestructura e instalaciones que posibiliten el servicio.
- Todas las actividades de atención, averías, reparaciones y rehabilitaciones.
- Todas las actividades administrativas y comerciales.

Este proyecto se centra principalmente en los segmentos en alta y baja, ya que tiene en cuenta los datos recogidos por los diferentes caudalímetros a lo largo de la red de distribución y sensores de nivel de los depósitos, de manera que pueda facilitar la detección

de malos funcionamientos de los sensores o averías que provoquen grandes pérdidas de agua.

2.2 Red de distribución

La red de distribución de aguas de Terrassa está dividida en 8 pisos, los cuáles contienen varios sectores de distribución de agua. Los pisos son los siguientes:

- Sulleva
- Can Colomer
- Zona Alta
- Can Boada
- Can Poal
- Can Parellada Residencial
- Can Parellada Industrial
- Can Palet de Vista Alegre

La infraestructura, hasta la fecha, cuenta con 98 caudalímetros [6], entre los cuales se encuentra los de impulsión a los depósitos y los de distribución, y 13 sensores de nivel de los diferentes depósitos repartidos por todos los pisos comentados anteriormente.

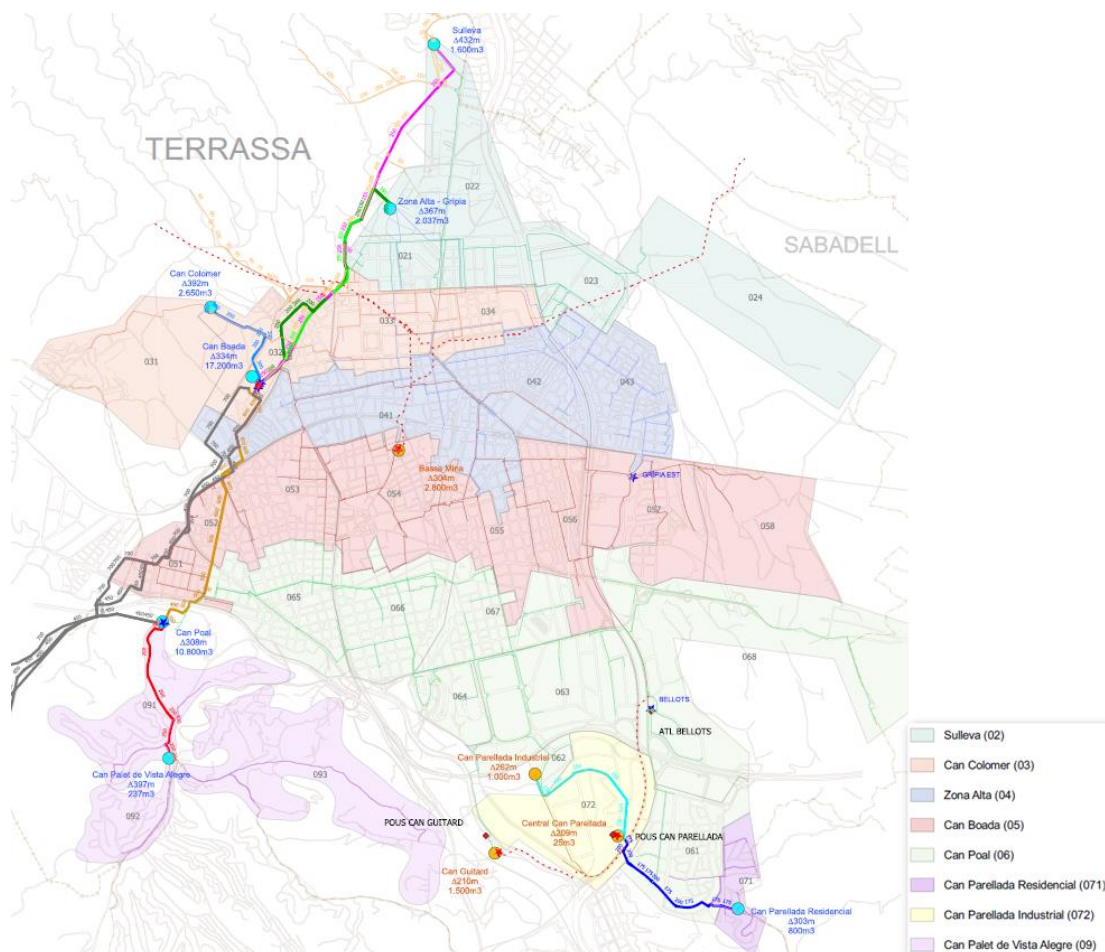


Figura 3. Distribución de los pisos de la red de distribución en Terrassa.

Este proyecto se centra en el estudio del piso de Sulleva, con el cual se pretende crear un modelo de estudio que puede servir de base para en un futuro estudiar los diferentes pisos de Terrassa.

2.3 Piso de Sulleva

El piso de Sulleva tiene una infraestructura óptima para realizar un primer modelo, ya que cuenta con diversos elementos que proporcionan información de la red de distribución, como son los sensores de nivel del depósito o los diversos caudalímetros.

En la imagen anterior, donde se encuentra la distribución de los diferentes pisos se puede observar en la parte superior el piso de Sulleva, que se encuentra en el norte de la ciudad. Este piso, recibe el agua impulsada desde el depósito de Can Boada que registra los datos en el caudalímetro CC0201, y se almacena en el depósito. Posteriormente, el piso de presión de Sulleva es alimentado desde el depósito por gravedad y el flujo de salida del depósito queda registrado por el caudalímetro CC0202.

A continuación se muestra una parte de la red de distribución donde aparece la infraestructura hasta el depósito de Sulleva, se muestran los caudalímetros de CC0201, el depósito, el caudalímetro CC0202 y el piso de Sulleva:

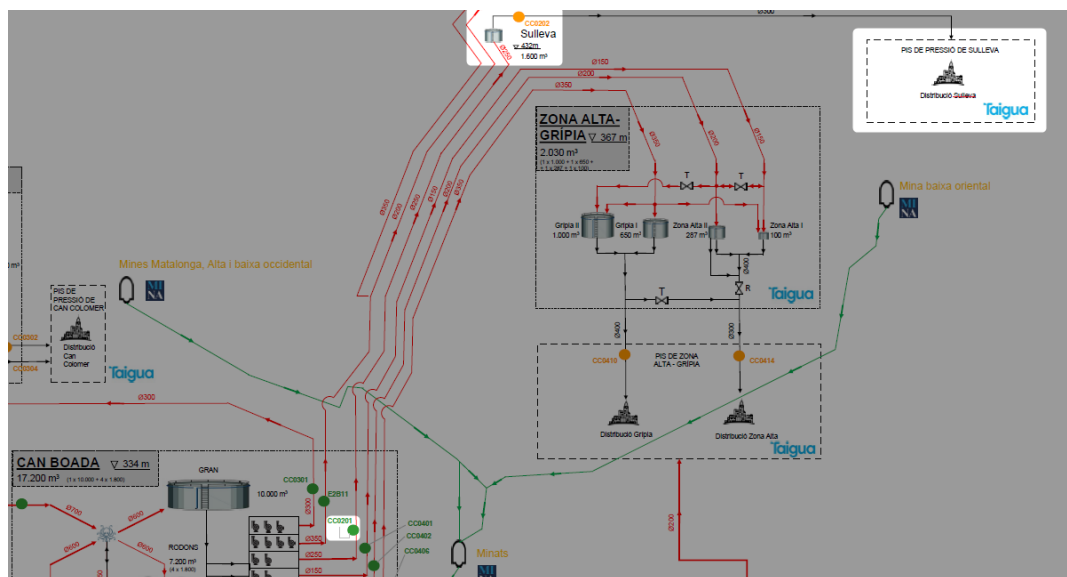


Figura 4. Red de distribución de entrada de agua a Sulleva.

La instalación de Sulleva tiene como principal componente el depósito. Este, se encuentra a 432 metros sobre el nivel del mar y tiene una capacidad de 1600 m³. Según los datos catastrales, el terreno tiene una superficie total construida de 432 m².

Se trata de un depósito rectangular de hormigón de 4,5 metros de altura y con una altura de lámina de agua de 3,5 m, con una renovación constante de agua y un tiempo de vacío de 14 horas. El depósito cuenta con dos cámaras, interconectadas entre sí, y cada una de ellas tiene un sensor de nivel de agua.

Cabe destacar, que además de la entrada de agua procedente de Can Boada (caudalímetro CC0201), el depósito tiene una entrada exterior que no está telecontrolada, (aportación externa mediante un camión cisterna) por lo que en la validación, se deberá tener en cuenta que puede haber momentos en los que el nivel suba, pero la bomba de Can Boada esté apagada.

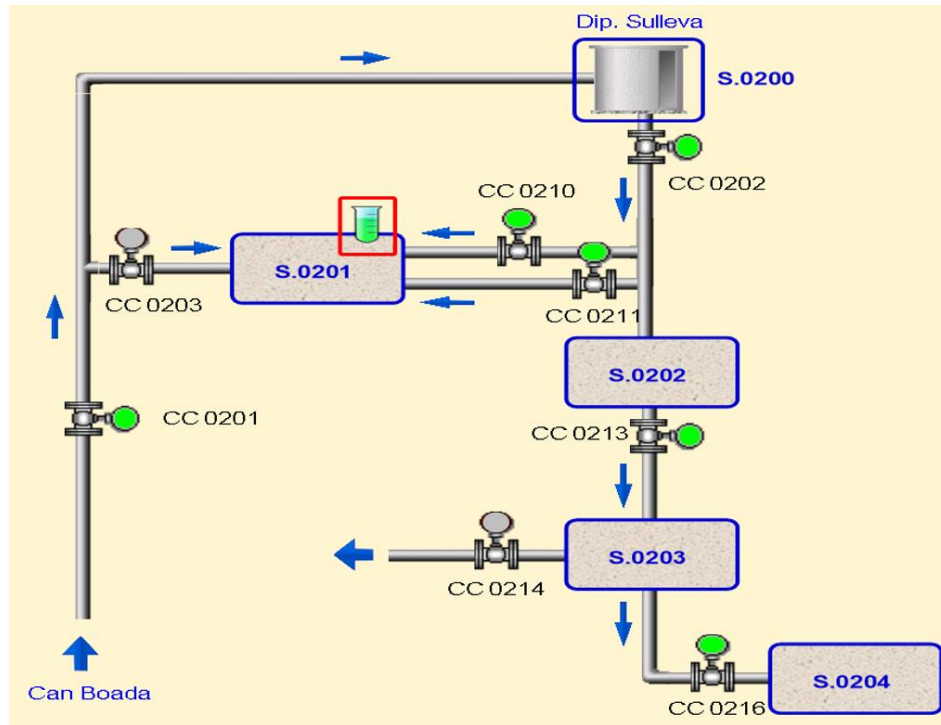


Figura 5. Esquemático del piso de Sulleva.

Tras el depósito y pasar por el caudalímetro CC0202, llega el agua a los diferentes sectores del piso de Sulleva:

- El sector S0201, cuyo consumo viene dado por los dos caudalímetros CC0210 y CC0211, el caudalímetro CC0203 no tiene datos, ya que actualmente ese paso está cerrado. Se trata de un sector residencial, por lo que su consumo puede ser constante a lo largo de los días. Las ubicaciones de los caudalímetros son las siguientes:

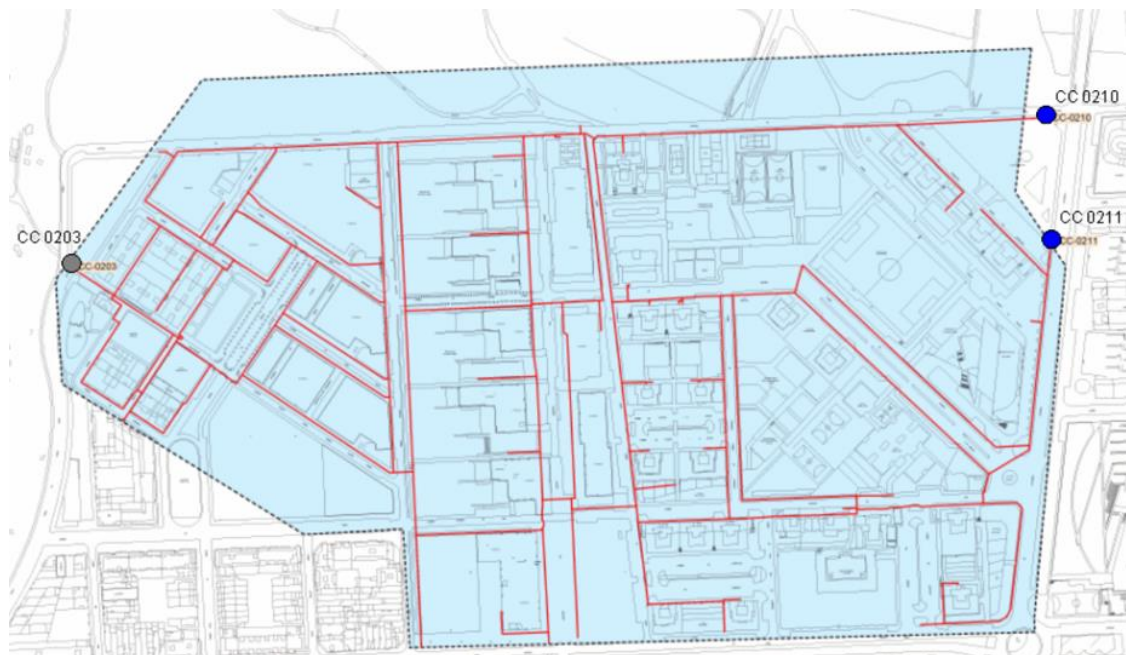


Figura 6. Sector S0201.

- El sector S0202 es un sector industrial y su consumo viene condicionado por diferentes caudalímetros, ya que será la diferencia entre lo que pasa del caudalímetro CC0202 y los caudalímetros CC0210, CC0211 y CC0213:

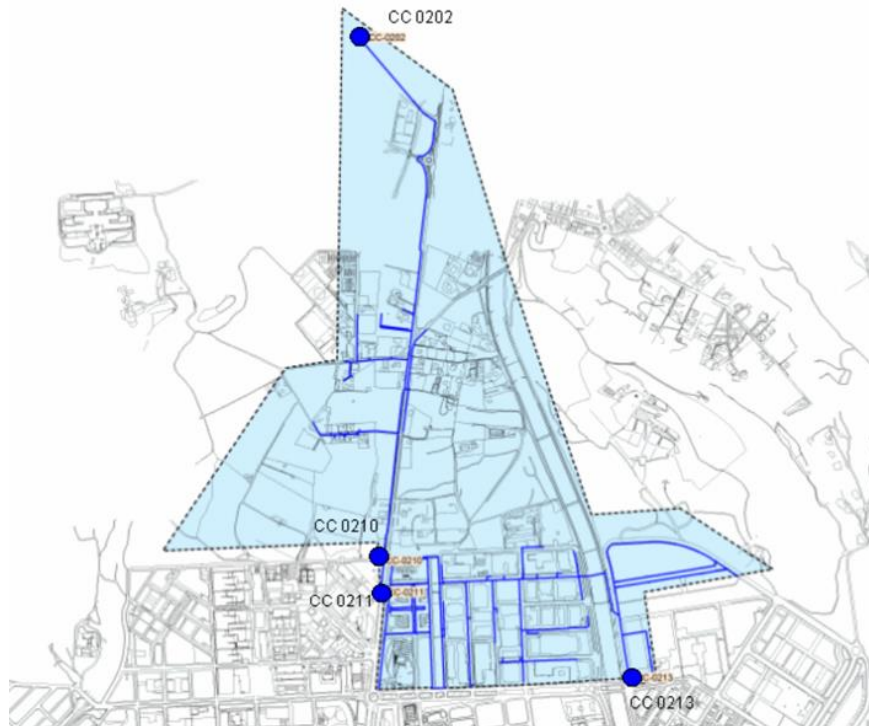


Figura 7. Sector S0202.

- El sector S0203, cuyo consumo viene dado de la diferencia de los caudalímetros CC0213 y CC0216. El caudalímetro CC0214 no recibe datos, ya que no pasa agua por esa vía. También se trata de un sector residencial:

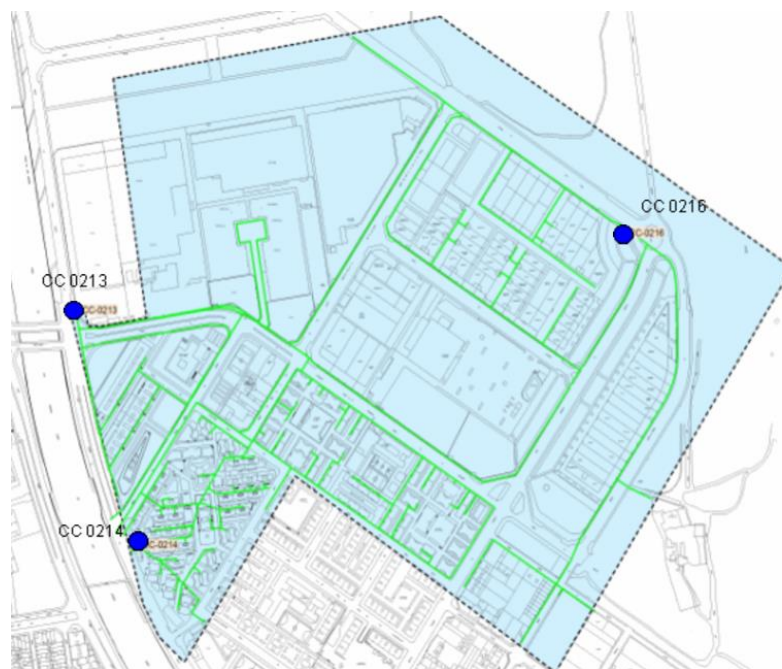


Figura 8. Sector S0203.

- Por último, el sector S0204 el cual únicamente tiene como entrada el caudalímetro CC0216. En este sector cobra importancia la existencia de un hotel que rellena esporádicamente su propio depósito provocando picos de consumo:

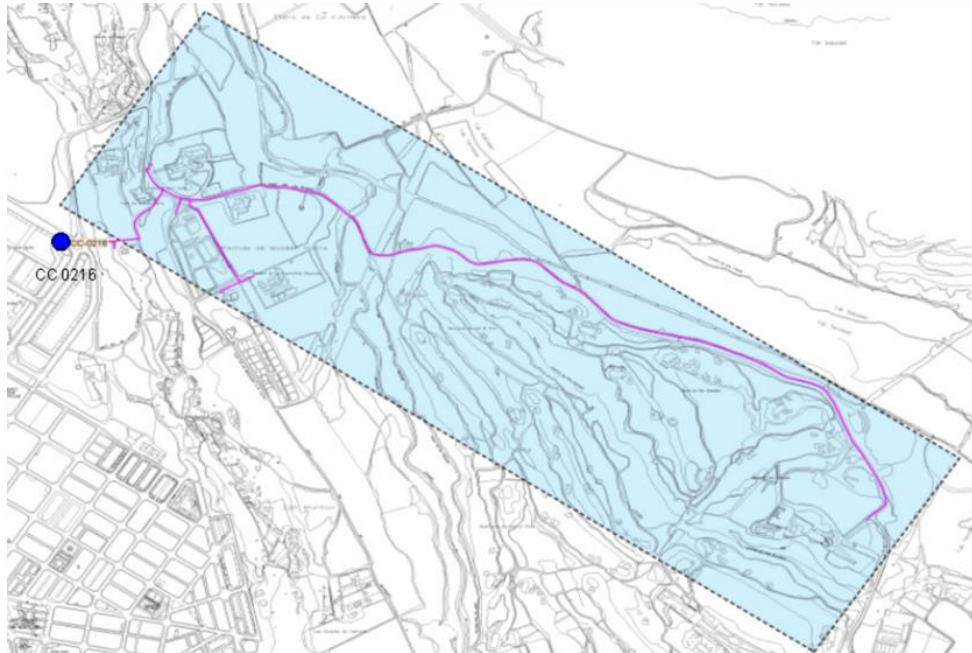


Figura 9. Sector S0204.

3 Tratamiento e importación de los datos

En este apartado se detallarán los primeros pasos del tratamiento de datos. En él se explicará el procedimiento seguido para poder llevar a cabo un tratamiento útil y eficaz para que posteriormente facilite el estudio y validación de estos datos.

Tal y como se ha comentado anteriormente, la empresa TAIGUA ha proporcionado un archivo CSV con los datos obtenidos de los sensores distribuidos por toda la red de distribución de agua de Terrassa durante el año 2019. En él, se encuentra una gran diversidad de datos debido a la variedad de sensores. Se ha de tener en cuenta la procedencia de los datos, para poder conocer la frecuencia de recogida de estos, así como el modo de funcionamiento, sin pasar por alto el objetivo principal del tratamiento que es utilizarlos con facilidad en Matlab.

Debido a los diferentes sensores, existe un gran número de datos que a ciertas horas, no tienen valor. Esto puede darse por el mal funcionamiento del sensor, como por ejemplo que esté durante dos horas sin registrar datos. En ese caso, se deberá tratar de manera que la información llegue correctamente a Matlab para que se pueda validar sabiendo que el valor es nulo.

Otro punto a tener en cuenta, es que existe una diversidad de frecuencia en la recogida de datos. Hay datos recogidos cada 2 minutos, otros cada 10 y otros cada 15 minutos. De esta manera, hay una gran cantidad de sensores de los cuales se disponen, por ejemplo, 30 registros por hora, en el caso de los biminutales, 6 registros por hora en el caso de los diezminutales y 4 en el caso de los quinceminutales.

Poniendo el foco en el piso de Sulleva, obtenemos datos de 6 caudalímetros de la red y 2 sensores de nivel del depósito:

Tabla 1. Información de los caudalímetros en el piso de Sulleva.

Pis de pressió	Codi	Descripció o adreça	Marca	Model	Nº serie	DN	Telecontrolat	Bateria	Data instal·lació
1	CC0201	Impulsió Can Boada a Sulleva					si	no	
2	CC0202	Sortida dipòsit Sulleva	Siemens	MAG 5100W			si	no	
3	CC0203	Ronda Ponent amb C. Palautordera	Woltex			100	no	no	
4	CC02D	Av. Lacetània - Ctra. Matadepera	Siemens	MAG 8000	7ME68 02YC31AA 1-Z	100	si	si	nov-08
5	CC02I	Ctra. Matadepera - C. Granollers	Siemens	MAG 8000	7ME68 02YC31AA 1-Z	100	si	si	nov-08
6	CC02B	Av. del Vallès - C. Castelltassa	Siemens	MAG 8000		100	si		
7	CC02H	Av. del Vallès - C. Canal del Llor	Siemens	MAG 8000	7ME68 02YC31AA 1-Z	100	no	si	nov-08
8	CC02E	Ctra. Castellàr - C. Font Sagüé	Siemens	MAG 8000	7ME68 02YC31AA 1-Z	50	si	si	nov-08

En la tabla anterior se puede observar como los caudalímetros CC0203 y CC0214 no están telecontrolados, pero es debido a que actualmente el paso de agua está deshabilitado y no se está usando esa vía. Por lo que el tratamiento de los datos se centrará en los demás caudalímetros.

Tabla 2. Registros sin tratar en Sulleva.

	DATETIME	DATE	TIME	CC0201	CC0202	CC0210	CC0211	CC0213	CC0214	CC0216	NV_SU_1	NV_SU_2
1	01/01/2019 0:02	01/01/2019	0:02:00	0	NA	NA	NA	NA	NA	NA	NA	NA
2	01/01/2019 0:04	01/01/2019	0:04:00	0	NA	NA	NA	NA	NA	NA	NA	NA
3	01/01/2019 0:06	01/01/2019	0:06:00	0	NA	NA	NA	NA	NA	NA	NA	NA
4	01/01/2019 0:08	01/01/2019	0:08:00	42,939816	NA	NA	NA	NA	NA	NA	NA	NA
5	01/01/2019 0:10	01/01/2019	0:10:00	41,359955	NA	NA	NA	NA	NA	NA	NA	NA
6	01/01/2019 0:12	01/01/2019	0:12:00	39,739582	NA	NA	NA	NA	NA	NA	NA	NA
7	01/01/2019 0:14	01/01/2019	0:14:00	39,739582	NA	NA	NA	NA	NA	NA	NA	NA
8	01/01/2019 0:15	01/01/2019	0:15:00	NA	NA	1,888889	1,7777779	10,666667	NA	3,3777778	2,2212863	2,2490273
9	01/01/2019 0:16	01/01/2019	0:16:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
10	01/01/2019 0:18	01/01/2019	0:18:00	39,739582	NA	NA	NA	NA	NA	NA	NA	NA
11	01/01/2019 0:20	01/01/2019	0:20:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
12	01/01/2019 0:22	01/01/2019	0:22:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
13	01/01/2019 0:24	01/01/2019	0:24:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
14	01/01/2019 0:26	01/01/2019	0:26:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
15	01/01/2019 0:28	01/01/2019	0:28:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
16	01/01/2019 0:30	01/01/2019	0:30:00	39,78009	NA	2,2222223	1,888889	11,000001	NA	3,2333333	2,2730143	2,2896774
17	01/01/2019 0:32	01/01/2019	0:32:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
18	01/01/2019 0:34	01/01/2019	0:34:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
19	01/01/2019 0:36	01/01/2019	0:36:00	39,739582	NA	NA	NA	NA	NA	NA	NA	NA
20	01/01/2019 0:38	01/01/2019	0:38:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
21	01/01/2019 0:40	01/01/2019	0:40:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
22	01/01/2019 0:42	01/01/2019	0:42:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
23	01/01/2019 0:44	01/01/2019	0:44:00	39,901619	NA	NA	NA	NA	NA	NA	NA	NA
24	01/01/2019 0:45	01/01/2019	0:45:00	NA	NA	2,3333335	2,0000002	8,5555563	NA	0,7555557	2,321172	2,3396659
25	01/01/2019 0:46	01/01/2019	0:46:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
26	01/01/2019 0:48	01/01/2019	0:48:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
27	01/01/2019 0:50	01/01/2019	0:50:00	39,78009	NA	NA	NA	NA	NA	NA	NA	NA
28	01/01/2019 0:52	01/01/2019	0:52:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
29	01/01/2019 0:54	01/01/2019	0:54:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
30	01/01/2019 0:56	01/01/2019	0:56:00	39,901619	NA	NA	NA	NA	NA	NA	NA	NA
31	01/01/2019 0:58	01/01/2019	0:58:00	39,861111	NA	NA	NA	NA	NA	NA	NA	NA
32	01/01/2019 1:00	01/01/2019	1:00:00	NA	NA	2,3333335	2,1111112	8,2222223	NA	0,76666671	2,3692379	2,3895628

En la tabla 2 se puede ver el ejemplo en una hora, concretamente del día 1 de enero de 2019, de los registros del piso de Sulleva. En la columna CC0201, se observa que se trata de un sensor que proporciona los datos de manera biminutal, por lo que en todos los minutos pares tiene un dato registrado, y en los cuartos de hora de 15 y 45 minutos no tiene datos, apareciendo así "NA". Como excepción, justo a la hora de los datos, también aparece como valor nulo.

Siguiendo con el ejemplo de la tabla anterior, hay dos columnas que no aparecen datos registrados, en este caso se trata de dos situaciones diferentes. La columna CC0202, debería tener registros quinceminutales, ya que es un sensor que sí que está activo, pero en este momento había fallado el funcionamiento. Por otro lado, la columna CC0214 se trata de un sensor que no está en uso, por lo que todos los datos serán nulos.

Por último, referido a los demás sensores, se tratan de valores quinceminutales, por lo que registran datos en los minutos 00, 15, 30 y 45. De esta manera, en el resto de minutos quedan registrados como nulos.

Este trabajo centra su estudio de validación y reconstrucción de datos en el piso de Sulleva, por lo que en este caso no hay datos diezminutales con los que trabajar.

Para poder tratar los datos correctamente es necesario tener valores de forma regular en un tiempo definido, es decir, en el caso de los datos originales, al tener frecuencias diferentes no se estaría trabajando con unos tiempos regulares, sino que la escala de tiempo variaría de periodos de dos minutos a periodos de quince minutos. Es por ello que no se debe únicamente rellenar los vacíos con aproximaciones, sino que se debe encontrar la manera que se consiga limpiar los datos para tener los registros ordenados.

3.1 Estudio de los datos

Al tener dos tipos de sensores, lo más lógico sería trabajar adaptarse a uno de los sensores para tener más facilidades, por lo que se podría trabajar con datos biminutales o quinceminutales (también se podría con intervalos mayores como 30 o 60 minutos). Por lo que se obtendrían dos opciones diferentes: seguir la frecuencia biminutal del sensor CC0201 y adaptar el resto de sensores consiguiendo los valores biminutales, o seguir la frecuencia quinceminutal del resto de sensores y adaptar los valores del sensor CC0201.

Para conocer la situación de cada uno de los sensores se ha solicitado a TAIGUA más información sobre la recogida de datos de estos en el piso de Sulleva:

- El sensor CC0201 se trata de un caudalímetro de entrada al depósito de Sulleva que proporciona los datos instantáneos en l/s cada dos minutos.
- El sensor CC0202 es el caudalímetro de salida del depósito. Su funcionamiento es muy irregular ya que se han detectado numerosos fallos de funcionamiento. Proporciona los datos instantáneos cada 15 minutos en l/s.
- Los sensores CC0210, CC0211, CC0213 y CC0216 son caudalímetros que se encuentran entre los diferentes sectores de Sulleva y proporciona el promedio del caudal en l/s² que ha pasado en los últimos 15 minutos.
- Los sensores NV_SU_1 y NV_SU_1 son los de nivel del depósito. Proporcionan el valor instantáneo de la altura del depósito en metros cada 15 minutos.
- El sensor CC0214 no está en funcionamiento.

Con esta información, se procede a realizar el tratamiento de los datos. Como se ha comentado, está la opción de tratarlos con el objetivo de tener registros cada dos minutos. En este caso, se deberían interpolar el resto de los valores de los sensores quinceminutales para adaptarlos a la frecuencia del sensor CC0201 que no haría falta modificarlo. Los demás sensores se deberían interpolar por intervalos, es decir, conociendo el valor previo y el posterior interpolar según el tiempo que haya transcurrido o dejar el mismo valor cada 15 minutos, ya que la información que nos proporcionan los sensores es el promedio de los 15 minutos previos. De esta manera, sí que se conseguiría una mayor cantidad de registros, pero los valores obtenidos tendrían menor exactitud.

En la siguiente imagen se muestra un ejemplo de lo que se pretende con este tratamiento de datos biminutal: en la columna CC0210 para rellenar los valores nulos entre las 00:15 y las 00:30 se interpolarían entre los dos valores, mientras que en la columna CC0213 se dejaría el valor de las 00:15 para todos los valores previos, ya que el propio sensor proporciona el valor medio de los últimos 15 minutos.

DATETIME	CC0210	CC0211	CC0213	CC0214	CC0216
01/01/2019 0:02	NA	NA	NA	NA	NA
01/01/2019 0:04	NA	NA	NA	NA	NA
01/01/2019 0:06	NA	NA	NA	NA	NA
01/01/2019 0:08	NA	NA	NA	NA	NA
01/01/2019 0:10	NA	NA	NA	NA	NA
01/01/2019 0:12	NA	NA	NA	NA	NA
01/01/2019 0:14	NA	NA	NA	NA	NA
01/01/2019 0:15	1,888889	1,7777779	10,666667	NA	3,3777778
01/01/2019 0:16	NA	NA	NA	NA	NA
01/01/2019 0:18	NA	NA	NA	NA	NA
01/01/2019 0:20	NA	NA	NA	NA	NA
01/01/2019 0:22	NA	NA	NA	NA	NA
01/01/2019 0:24	NA	NA	NA	NA	NA
01/01/2019 0:26	NA	NA	NA	NA	NA
01/01/2019 0:28	NA	NA	NA	NA	NA
01/01/2019 0:30	2,2222223	1,888889	11,000001	NA	3,2333333

Figura 10. Tratamiento de datos biminutal.

En cuanto al tratamiento quinceminutal, se trataría de algo más sencillo. Todos los sensores proporcionan valores cada quince minutos a excepción del caudalímetro de la entrada del depósito (CC0201), por lo que únicamente se debería modificar este último. En cuanto a la veracidad de los datos, a excepción del caudalímetro CC0202 que proporciona los datos instantáneos quinceminutalmente, los demás ya proporcionan la media de esos 15 minutos, por lo que se puede realizar el mismo proceso para el CC0201.

El mayor inconveniente que tiene el tratamiento quinceminutal es que se ha de calcular la media de los valores cada 15 minutos, pero conociendo 7 valores instantáneos, se puede hacer una mayor aproximación que en tratamiento biminutal. Es por ello que se ha elegido el tratamiento quinceminutal.

3.2 Solución adoptada

Una vez establecido el tratamiento que se va a realizar, el procesado de los datos se ha realizado mediante el programa Matlab. El primer paso a realizar es la importación de los datos.

En primer lugar, al tratarse de una gran cantidad de datos, se va a importar como tabla ya que facilita el acceso a los datos, ya que cada columna está diferenciada por el nombre del sensor o variable a estudiar. Como detalle importante, se ha especificado que todos aquellos valores que sean leídos como "NA" que los trate como vacíos "TreatAsEmpty", de esta manera, entiende las columnas de los sensores como valores numéricos en cambio a caracteres.

Al importarse la tabla es necesario encontrar en qué filas se encuentran los cuartos de hora, ya que todo se centrará en ellos. Es por eso que se busca en la columna de tiempos todos aquellos instantes en que los minutos coincidan con 00, 15, 30 o 45. De esta manera se conocerá la posición en la que hay datos quinceminutales.

Cabe destacar alguna excepción debido a la diferencia de caudalímetros. Se han encontrado situaciones en la que en minuto 00 con 00 segundos no hay datos, pero en el minuto 00 y 01 segundos tiene registrado un dato. Un ejemplo sería el siguiente caso, donde a las 3:00:00 no tiene registros (y debería haber al tratarse de minuto par), y sin embargo a las 3:00:01 sí tiene valor:

TIME	CC0201	CC0202	CC0210
2:38:00	39,861111	NA	NA
2:40:00	39,861111	NA	NA
2:42:00	39,901619	NA	NA
2:44:00	39,942131	NA	NA
2:45:00	NA	NA	2,222223
2:46:00	39,901619	NA	NA
2:48:00	39,901619	NA	NA
2:50:00	39,861111	NA	NA
2:52:00	39,942131	NA	NA
2:54:00	39,901619	NA	NA
2:56:00	39,942131	NA	NA
2:58:00	39,942131	NA	NA
3:00:00	NA	NA	2,111112
3:00:01	39,901619	NA	NA
3:02:00	39,78009	NA	NA
3:04:00	39,78009	NA	NA
3:06:00	39,861111	NA	NA
3:08:00	39,861111	NA	NA
3:10:00	39,901619	NA	NA
3:12:00	39,901619	NA	NA
3:14:00	39,861111	NA	NA
3:15:00	NA	NA	1,888889

Figura 11. Excepción en el tratamiento de datos.

Con el objetivo de no registrar horas duplicadas, se ha especificado que únicamente se coja los valores con los segundos exactos en 00. De esta manera, el siguiente valor se tendría en cuenta para los siguientes 15 minutos.

Una vez conseguidas las filas en las que se encuentran los cuartos de hora se deben guardar esos valores. Así, se recorrerá la tabla y cada vez que pase por el número de fila correspondiente a un cuarto de hora, cogerá todos los valores quinceminutales en todos aquellos sensores que tengan esa frecuencia, mientras que para el caudalímetro CC0201 se realiza la media de los valores de los anteriores 15 minutos.

La tabla resultado de la primera hora sería la siguiente:

date	Var2	Var3	Var4	Var5	Var6	Var7
01-Jan-2019 00:15:00	23.397	NaN	1.8889	1.7778	10.667	3.3778
01-Jan-2019 00:30:00	39.795	NaN	2.2222	1.8889	11	3.2333
01-Jan-2019 00:45:00	39.81	NaN	2.3333	2	8.5556	0.75556
01-Jan-2019 01:00:00	39.844	NaN	2.3333	2.1111	8.2222	0.76667
01-Jan-2019 01:15:00	39.78	NaN	2.7778	2.2222	8.4444	0.75556

Figura 12. Resultado de tratamiento de datos quinceminutal.

En la figura anterior, Var2 corresponde al sensor CC0201, Var3 al sensor CC0202, Var4 al sensor CC0210, Var5 al sensor CC0211, Var6 al sensor CC0213 y Var7 al sensor CC0216. Comparándola con la imagen mostrada al principio del apartado, se ve como corresponden todos los datos quinceminutales, mientras que los correspondientes al caudalímetro CC0201 se realiza la media.

Con este método, se ha conseguido una menor cantidad de registros con los que trabajar pero de una manera más limpia sin tanto ruido en las aproximaciones, facilitando la posterior validación.

4 Estudio y elección de metodologías de validación

Tal y como se ha comentado con anterioridad, la validación de los datos es la parte central del proyecto. Cobra de gran importancia encontrar un modelo de validación fiable en el que poder comparar los datos y conocer si son válidos o no.

En este presente apartado se profundizará sobre el estudio de las diferentes técnicas de validación que aplicar a los datos tratados. Se detallará el procedimiento seguido sobre los tests para cada caudalímetro dependiendo de su funcionamiento, como puede ser caudalímetro de bombeo o que forme parte de un sector. Teniendo en cuenta los diferentes factores se definirá un método para cada sensor.

Con el fin de conseguir una validación eficaz, se va a realizar por partes, es decir, al tener sensores de diferentes tipos y cada uno puede estar relacionado con otros, se separarán en los siguientes grupos:

- El sensor CC0201 de bombeo al depósito, se deberá tener en cuenta el nivel del depósito.
- Los sensores NV_SU_1 y NV_SU_1 son los de nivel del depósito, en los que al ser una medida continua se ha de tener en cuenta tanto su último valor como el del otro sensor.
- El sensor CC0202, el de salida del depósito, se deberá tener en cuenta tanto el nivel del depósito, como la participación en el sector S0202.
- Los sensores CC0210 y CC0211, que son la entrada al sector S0201 donde se podrá conocer el consumo de la zona.
- Los sensores CC0213 y CC0216, que participan en el sector S0203 y al conocer el consumo se podrá realizar la validación. El segundo sensor también tiene participación en el S0204.

4.1 Metodologías de validación

En una primera instancia, se define el conjunto de tests que se aplicarán para realizar la validación de los datos. Estos test no se utilizarán en su totalidad en cada uno de ellos sino, que se escogerá la mejor opción según el tipo de caudalímetro que sea o de la información de la que se disponga.

- **Valor nulo:** Se examinará si el valor en cuestión existe o por contrario aparece como inexistente ('NaN' en Matlab).
- **Valor fuera de rango:** Se revisará el histórico de los datos y se establecerán un máximo y mínimo de los datos. El mínimo absoluto permitido para todos los casos será 0, ya que no puede haber valores negativos (significaría que el agua está yendo en dirección contraria).
- **Test de coherencia local:** Se examinará la coherencia de los datos comparándolo con datos de otros elementos relacionados. El ejemplo más claro será el depósito, con la variación del nivel de agua se podrá comprobar el correcto funcionamiento del sensor de entrada y de salida.

- **Valores consecutivos:** Para señales continuas se revisará que la variación entre un valor y su consecutivo se encuentre dentro de sus posibilidades. Por ejemplo, el nivel de agua no deberá variar más de lo permitido.
- **Patrones diarios:** Se analizará el consumo diario en todos los sectores realizando un balance de masas de los sensores que intervenga, teniendo varios días de estudio y otros días para testearlo. Se establecerá un intervalo de confianza que permita identificar los valores válidos.

El procedimiento de la validación vendrá diferenciada por el tipo de sensor a validar. Para todos ellos se validará que no se trate de un valor nulo o fuera de los intervalos permitidos, una vez realizada la validación común, según qué tipo de información se disponga se realizarán los diferentes tests. Por ejemplo, en el caso de los sensores relacionados con el consumo del sector, se realizarán los tests comunes y el de patrones diarios.

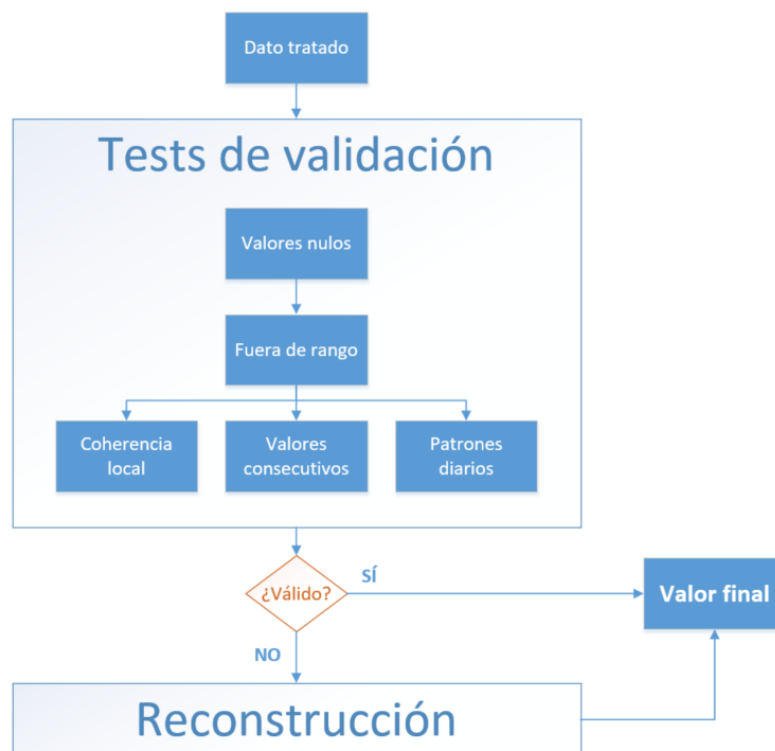


Figura 13. Proceso de la validación de datos.

Una vez pasado por los tests, para poder conocer de qué tipo de validación se ha encontrado el error, se establecerá un código de validación con números del 0 al 6. Estos números serán genéricos para todos los casos, por lo que si se encuentra un valor nulo, marcará el mismo número si se trata del llenado del depósito o de la entrada del sector S0203. Es una práctica que facilitará la reconstrucción, ya que al conocer el origen del fallo se podrá corregir el dato de la manera más conveniente.

En el caso que el valor haya resultado válido, se confirmará como valor final.

4.2 CC0201 – Bombeo al depósito

El primer caudalímetro que validar, es el CC0201 que corresponde al de la entrada al piso de Sulleva. En él básicamente se recoge el caudal que es impulsado por la bomba para llenar el depósito.

La bomba tiene únicamente dos modos de funcionamiento, cuando está parada no hay caudal, por lo que en los datos debería ser 0 l/s y cuando está en funcionamiento tiene un caudal de aproximadamente 40 l/s. También TAIGUA, ha hecho hincapié que en algunos datos, al encender la bomba, registra un sobrepico significativamente mayor a 40 l/s. En este caso, al realizar el tratamiento de los datos y tratarlos quinceminutalmente, ese error será disminuido.

En la figura 15 se muestran los datos registrados en el año 2019 donde se pueden diferenciar claramente los dos estados (en marcha o paro). También se detectan sobre picos muy significativos.

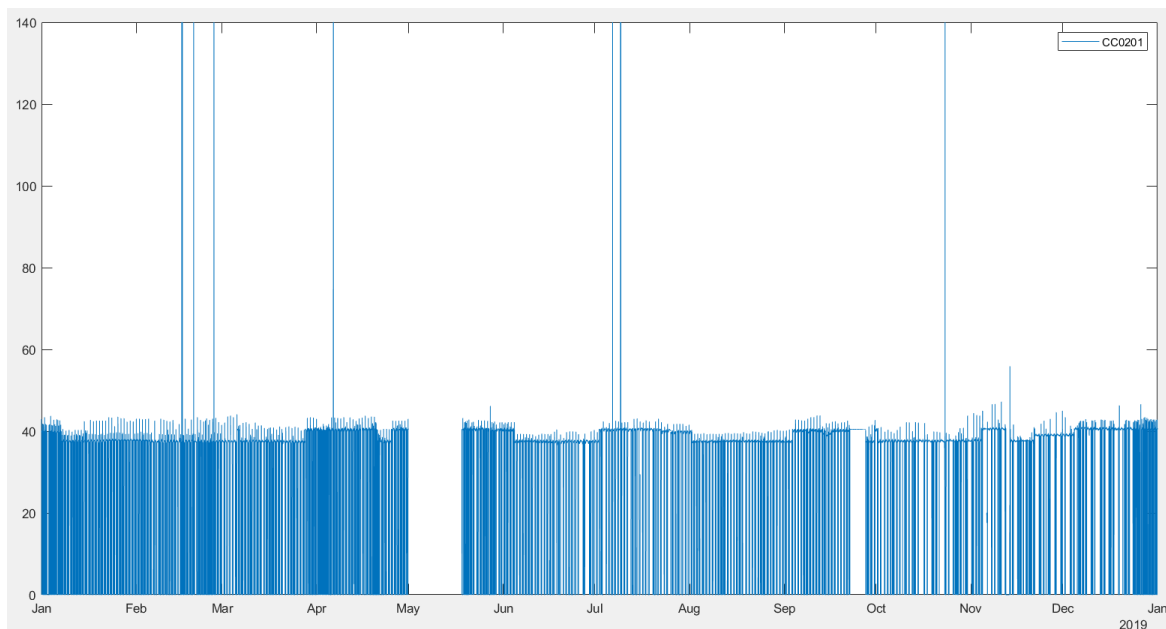


Figura 14. Datos anuales del caudalímetro CC0201.

Conociendo esta primera premisa, se conoce el primer test de validación:

Ecuación 1. Primer test de validación de CC0201.

$$0 \text{ l/s} \leq CC0201 \leq 49 \text{ l/s}^2$$

El valor nunca puede ser negativo y por otro lado, revisando los datos anuales se observa que extrayendo los sobrepicos iniciales de cada ciclo el valor máximo anual es de 41 l/s², se añade un margen del 20% para aceptar los sobrepico del inicio del encendido de la bomba.

Al no tener más información de la bomba, el siguiente elemento del que podemos extraer información es del nivel del depósito, y con ello del sensor CC0202, realizando una relación de entrada y salida del depósito conociendo la variación del nivel, aunque este segundo caso queda descartado ya que la información que proporciona el sensor CC0202 es escasa

además de poco fiable en cuanto a la variación del tiempo, ya que proporciona datos quinceminutales instantáneos.

En cuanto al nivel del depósito, sí que puede proporcionar información interesante. Por la salida del depósito siempre debe estar pasando agua, ya que es la entrada a los diferentes sectores del piso de Sulleva, por lo que cuando la bomba no esté impulsando agua, el nivel de agua debería disminuir. Y cuando la bomba esté encendida con casi total seguridad, el nivel del depósito aumente.

En este caso, se trata de un test de validación que tiene en cuenta la diferencia de nivel entre un instante y los quince minutos siguientes, se considerará como válido en las siguientes circunstancias:

- Si la bomba está activa y el nivel aumenta en los últimos 15 minutos.
- Si la bomba está desactiva y el nivel disminuye en los últimos 15 minutos.

Con estos dos primeros tests, se realizan las primeras pruebas para conocer la efectividad de la validación. En este caso, se cogerán los datos de los 3 primeros días de enero de 2019.

date	CC0201	NV1	validacionCC0201
02-Jan-2019 20:45:00	0	3.0573	0
02-Jan-2019 21:00:00	0	3.011	1
02-Jan-2019 21:15:00	0	NaN	1
02-Jan-2019 21:30:00	0	2.9167	1
02-Jan-2019 21:45:00	0	2.8649	1
02-Jan-2019 22:00:00	0	2.8187	1
02-Jan-2019 22:15:00	0	2.7779	1
02-Jan-2019 22:30:00	0	2.7373	1
02-Jan-2019 22:45:00	0	2.7021	1
02-Jan-2019 23:00:00	0	2.667	1
02-Jan-2019 23:15:00	0	2.6337	1
02-Jan-2019 23:30:00	0	2.5986	1
02-Jan-2019 23:45:00	0	2.5708	1
03-Jan-2019 00:00:00	0	2.5393	1
03-Jan-2019 00:15:00	0	2.5301	1
03-Jan-2019 00:30:00	0	2.5893	0
03-Jan-2019 00:45:00	0	2.6448	0
03-Jan-2019 01:00:00	0	2.6985	0
03-Jan-2019 01:15:00	0	2.7539	0
03-Jan-2019 01:30:00	0	2.8113	0
03-Jan-2019 01:45:00	0	2.8723	0
03-Jan-2019 02:00:00	0	2.9278	0
03-Jan-2019 02:15:00	0	2.987	0
03-Jan-2019 02:30:00	0	3.048	0
03-Jan-2019 02:45:00	0	3.1072	0
03-Jan-2019 03:00:00	0	3.1682	0
03-Jan-2019 03:15:00	0	3.2274	0

Figura 15. Primera prueba de validación CC0201.

En esta primera validación, se observa el buen funcionamiento tanto en los dos estados de la bomba, pero se encuentra un caso extraño mostrado en la tabla de resultados anterior: observamos como el nivel del depósito (NV1) aumenta, mientras que por el caudalímetro

CC0201 no hay registros que haya habido un paso agua, entonces la señal de validación aparece como 0 dando así erróneo.

Al tener descartada la opción de comparar con la salida, se solicita a TAIGUA más información, ya que en este caso, parece claro que está fallando el sensor de entrada. TAIGUA nos indica que hay momentos en los que el depósito es llenado externamente por otra vía en la cual no se registran datos, pero con el objetivo de clarificar la situación, ha proporcionado información al respecto de la bomba. En ella están los datos de corriente eléctrica de la bomba y presión de la tubería para conocer si ha estado en funcionamiento o no.

Realizando una comparación con los datos válidos, previos a los mostrados en la tabla anterior, se puede extraer la siguiente información:

- La bomba en funcionamiento consume alrededor de 110 A y provoca una presión de 114 mca.
- La bomba en paro no consume intensidad y la presión está entorno a los 100 mca.

Conociendo estos datos, se procede a volver a verificar el proceso de validación añadiendo más datos. Al tener 3 sensores diferentes, en caso que todos ellos concuerden se dará como válida la información. De esta manera quedan las siguientes validaciones:

- Si el sensor CC0201 es mayor a 20 l/s y la intensidad de la bomba es mayor a 80 A se dará por válido.
- Si el sensor CC0201 es mayor a 20 l/s y la presión de la bomba es mayor a 110 mca se dará por válido.

Como excepción, están los momentos de arranque y paro de la bomba. Los datos de intensidad y presión son biminutales, mientras que los que se validan es una media de los últimos 15 minutos, por lo que asume que si el caudal está por debajo del 90% de su valor nominal se considerará que ha parado en los 15 minutos previos, mientras que si el valor del caudal está por encima del 10% se asumirá que la bomba se ha encendido.

De esta manera, se registrarán los diferentes estados de la validación de la siguiente manera:

Tabla 3. Descripción de la validación CC0201.

Nº	Descripción
0	Valor nulo o fuera de rango
1	Válido
2	Intensidad de la bomba y CC0201 no concuerda
3	Presión de la bomba y CC0201 no concuerda
4	Llenado externo

A continuación se muestra el mismo rango de tiempo que en la tabla anterior. En él se observa como en los momentos en los que sube el nivel y el caudalímetro no registra valores, es debido a que efectivamente la bomba estaba apagada. Tanto los datos de intensidad como de presión confirman que no pasaba agua al depósito por esa vía, pero lo más probable es que se estuviese llenando por otra vía:

date	CC0201	NV1	valCC0201
02-Jan-2019 20:45:00	0	3.0573	0
02-Jan-2019 21:00:00	0	3.011	1
02-Jan-2019 21:15:00	0	NaN	1
02-Jan-2019 21:30:00	0	2.9167	1
02-Jan-2019 21:45:00	0	2.8649	1
02-Jan-2019 22:00:00	0	2.8187	1
02-Jan-2019 22:15:00	0	2.7779	1
02-Jan-2019 22:30:00	0	2.7373	1
02-Jan-2019 22:45:00	0	2.7021	1
02-Jan-2019 23:00:00	0	2.667	1
02-Jan-2019 23:15:00	0	2.6337	1
02-Jan-2019 23:30:00	0	2.5986	1
02-Jan-2019 23:45:00	0	2.5708	1
03-Jan-2019 00:00:00	0	2.5393	1
03-Jan-2019 00:15:00	0	2.5301	1
03-Jan-2019 00:30:00	0	2.5893	4
03-Jan-2019 00:45:00	0	2.6448	4
03-Jan-2019 01:00:00	0	2.6985	4
03-Jan-2019 01:15:00	0	2.7539	4
03-Jan-2019 01:30:00	0	2.8113	4
03-Jan-2019 01:45:00	0	2.8723	4
03-Jan-2019 02:00:00	0	2.9278	4
03-Jan-2019 02:15:00	0	2.987	4
03-Jan-2019 02:30:00	0	3.048	4
03-Jan-2019 02:45:00	0	3.1072	4
03-Jan-2019 03:00:00	0	3.1682	4
03-Jan-2019 03:15:00	0	3.2274	4

Figura 16. Validación de sensor CC0201.

4.3 NV_SU1 y NV_SU2 – Niveles de depósito

Se trata de dos sensores de nivel que miden dos cámaras diferentes del depósito, pero interconectadas entre sí, por lo que registrarán prácticamente los mismos datos.

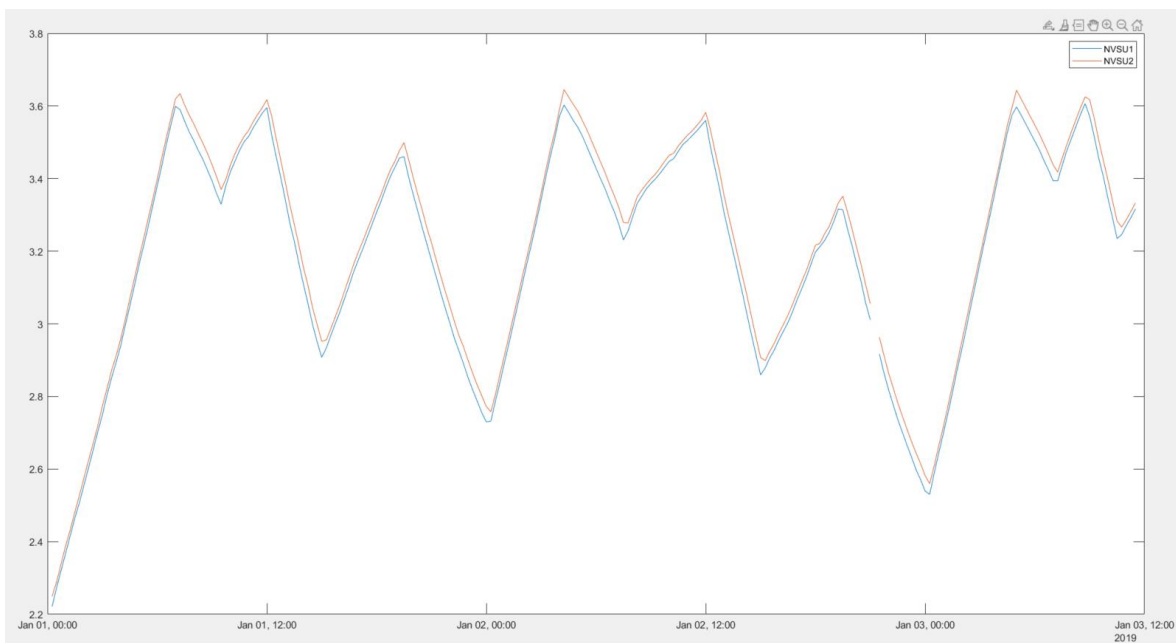


Figura 17. Registros de los niveles del depósito.

En este caso, al tratarse de una señal continua, se puede evaluar respecto a su último estado, ya que debe tener un incremento o decremento lógico según la entrada o salida del depósito.

La primera de las validaciones, es que debe ser mayor o igual 0. En este, como en el primer test del caudalímetro anterior, también se evalúa si tiene datos. En el caso que falle, se asignará un cero a la validación.

La segunda validación que se realizará es conocer el máximo incremento o decremento que puede tener en un intervalo de tiempo de 15 minutos. De esta manera se podrá establecer un rango de posibles registros entre un dato y otro. Ya se conoce el área del depósito, por lo que conociendo la entrada y la salida, se puede conocer la diferencia de nivel:

Ecuación 2. Validación de sensores de nivel.

$$\frac{(entrada - salida) * \Delta t}{\Delta nivel} = A$$

$$\frac{(CC0201 - CC0202) * 0.9}{\Delta nivel} = 432$$

$$\Delta nivel = \frac{(45 - 0) * 0.9}{432} = 0.09375 m$$

En la ecuación anterior se considera como la mayor diferencia entre la entrada y la salida de 45 l/s, un caso extremo, ya que en la salida siempre va a pasar agua hacia los diferentes sectores. De esta manera se considerará un incremento o decremento mayor a 0,1 como inválido.

La descripción de la validación quedará de la siguiente manera:

Tabla 4. Descripción de la validación de nivel del depósito.

Nº	Descripción
0	Valor nulo o fuera de rango
1	Válido
5	Fallo con valores previos

A continuación se muestra un ejemplo de la validación, en él se muestra como en el primer datos no puede compararse con el dato anterior (validación en 5), y en los otros hay un dato inexistente (validación en 0) y el siguiente no puede compararse (validación en 5).

date	NV1	valNV1	NV2	valNV2
02-Jan-2019 20:45:00	3.0573	5	3.1072	5
02-Jan-2019 21:00:00	3.011	1	3.0554	1
02-Jan-2019 21:15:00	NaN	0	NaN	0
02-Jan-2019 21:30:00	2.9167	5	2.963	5
02-Jan-2019 21:45:00	2.8649	1	2.913	1
02-Jan-2019 22:00:00	2.8187	1	2.8649	1
02-Jan-2019 22:15:00	2.7779	1	2.8242	1
02-Jan-2019 22:30:00	2.7373	1	2.7835	1
02-Jan-2019 22:45:00	2.7021	1	2.7465	1
02-Jan-2019 23:00:00	2.667	1	2.7114	1
02-Jan-2019 23:15:00	2.6337	1	2.6762	1
02-Jan-2019 23:30:00	2.5986	1	2.6448	1
02-Jan-2019 23:45:00	2.5708	1	2.6152	1
03-Jan-2019 00:00:00	2.5393	1	2.5819	1
03-Jan-2019 00:15:00	2.5301	1	2.5598	1
03-Jan-2019 00:30:00	2.5893	1	2.6097	1
03-Jan-2019 00:45:00	2.6448	1	2.6651	1
03-Jan-2019 01:00:00	2.6985	1	2.717	1
03-Jan-2019 01:15:00	2.7539	1	2.7743	1
03-Jan-2019 01:30:00	2.8113	1	2.8316	1

Figura 18. Validación de sensor de nivel.

4.4 CC0202 – Salida del depósito

Este caudalímetro es el más conflictivo del piso de Sulleva, se trata de un sensor que funciona en escasas ocasiones durante un periodo de pocos días, por lo que el resto de los datos va a estar sin valor.

Otro detalle a tener en cuenta es que este valor registra los datos quinceminutalmente y de manera instantánea, por lo que no se puede conocer los datos instantáneos del caudal, ni tampoco se puede conocer el caudal acumulado entre. Este punto, cobra una gran importancia para comparar la diferencia de niveles del depósito o con datos históricos.

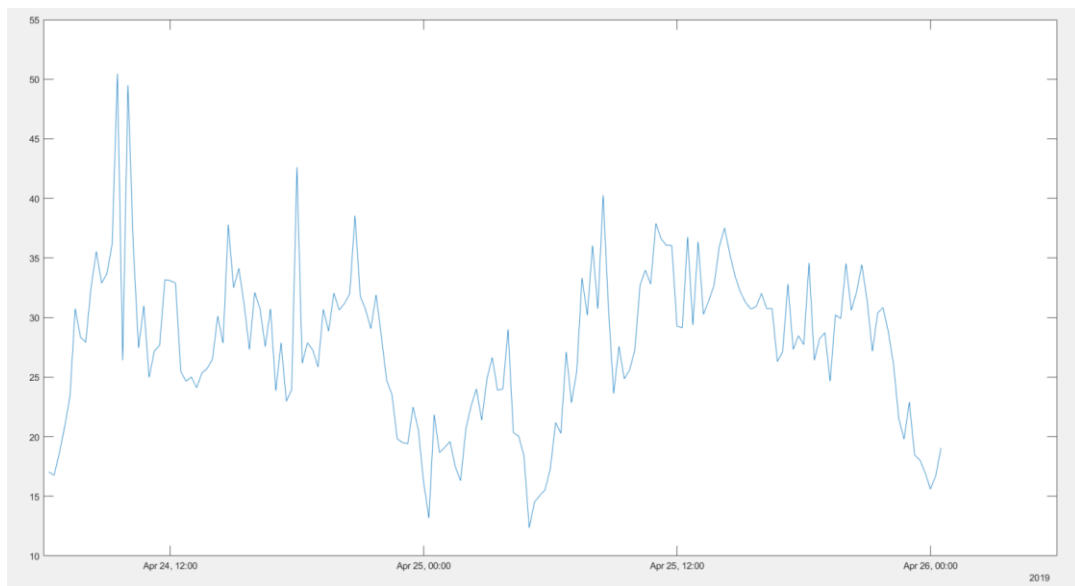


Figura 19. Datos del caudalímetro CC0202.

El primer test de validación, como en los anteriores, es que el valor debe ser mayor que 0 y no puede ser nulo. Al ser la salida del depósito y tras él se reparte el agua por todos los sectores, siempre ha de tener un valor.

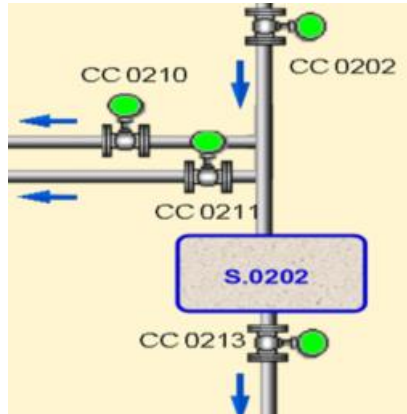


Figura 20. Sector S0202

El sensor CC0202, es la entrada para todos los sectores, pero para conocer el consumo del S0202 se ha de tener en cuenta directamente, ya que será la diferencia respecto a los que van para el sector S0201 y el saliente de este propio sector. Por lo que el cálculo será el siguiente:

Ecuación 3. Cálculo consumo S0202.

$$S0202 = CC0202 - (CC0210 + CC0211) - CC0213$$

Haciendo el cálculo para 23 días con registro, se obtienen los siguientes resultados diarios:

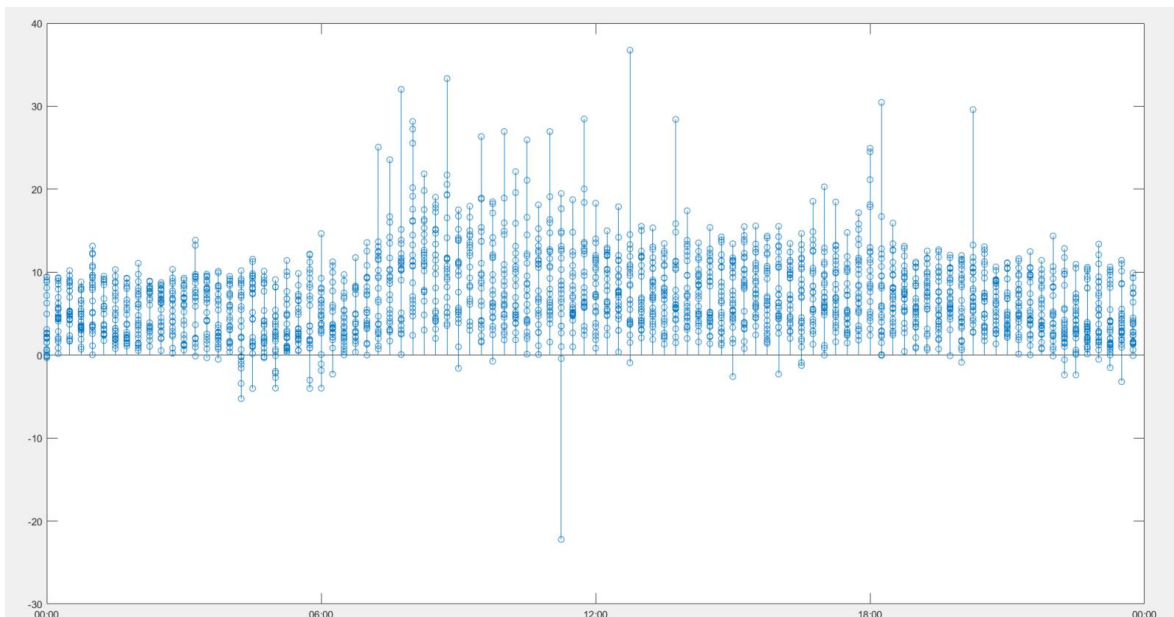


Figura 21. Cálculo sector S0202.

En la anterior figura se muestran los datos del sector con intervalos de 15 minutos a lo largo del día. Tal y como se puede observar hay valores muy dispares y sería muy difícil extraer un modelo en condiciones. Por otro lado, TAIGUA informó que a parte de recoger pocos datos, es un sensor que no pueden asegurar su rendimiento, por lo que tampoco tendría sentido realizar un modelo teniendo en cuenta un comportamiento dudoso. Un ejemplo de esto último sería los muchos valores negativos que serían imposibles, ya que representan el consumo de un sector de población. Es por ello que este método queda inválido.

Descartando el sector, la otra fuente de información que tenemos es el depósito. Ya se conocen la validez de los datos del sensor de entrada y el del nivel del agua, por lo que pueden conocer qué valor debería obtenerse con los datos de los últimos 15 minutos. En este caso, se realiza el mismo cálculo que en la ecuación 2, conociendo el área, la entrada y la diferencia de nivel, se aproxima el caudal que debería haber salido en los anteriores 15 minutos.

En esta figura se muestra los datos reales del sensor CC0202 y los datos estimados con la ecuación 2:



Figura 22. Valores registrados y calculados (arriba) y diferencias porcentuales en valor absoluto del CC0202 (abajo).

Tal y como se observa en la imagen inferior, se muestra las diferencias porcentuales entre el valor aproximado y el registrado por el sensor. Encontramos que excepto algunos puntos, la diferencia no sobrepasa el 25%. En este caso, al tratarse de una aproximación y comparación con el sensor instantáneo cada 15 minutos, se aproximará al 35%.

Tabla 5. Descripción de la validación de nivel del depósito.

Nº	Descripción
0	Valor nulo o fuera de rango
1	Válido
5	Fallo con valores previos

En este caso, la tabla de validación sería la misma que en el sensor anterior.

date	CC0202	calculado	valCC0202
24-Apr-2019 06:15:00	17.049	0	5
24-Apr-2019 06:30:00	16.767	20.066	1
24-Apr-2019 06:45:00	18.616	19.137	1
24-Apr-2019 07:00:00	20.824	20.827	1
24-Apr-2019 07:15:00	23.341	24.398	1
24-Apr-2019 07:30:00	30.736	27.872	1
24-Apr-2019 07:45:00	28.348	27.813	1
24-Apr-2019 08:00:00	27.937	28.727	1
24-Apr-2019 08:15:00	32.508	30.492	1
24-Apr-2019 08:30:00	35.538	33.962	1
24-Apr-2019 08:45:00	32.893	33.043	1
24-Apr-2019 09:00:00	33.663	32.174	1
24-Apr-2019 09:15:00	36.153	33.027	1
24-Apr-2019 09:30:00	50.457	35.639	1
24-Apr-2019 09:45:00	26.422	44.545	5
24-Apr-2019 10:00:00	49.506	33.861	5
24-Apr-2019 10:15:00	36.052	41.803	1
24-Apr-2019 10:30:00	27.449	30.355	1
24-Apr-2019 10:45:00	30.993	29.448	1
24-Apr-2019 11:00:00	24.984	36.868	5
24-Apr-2019 11:15:00	27.192	26.631	1
24-Apr-2019 11:30:00	27.68	26.631	1
24-Apr-2019 11:45:00	33.175	19.374	5
24-Apr-2019 12:00:00	33.098	32.148	1
24-Apr-2019 12:15:00	32.893	30.306	1

Figura 23. Validación sensor CC0202

Se ha escogido un tramo en los que sí que existían datos del sensor, por lo que no aparece ningún error de tipo 0, que sería sin dato. Los datos con validación de tipo 5 son aquellos que el valor calculado está fuera de lo permitido.

4.5 CC0210 y CC0211 – Sector S0201

Los caudalímetros CC0210 y CC0211 son los encargados de medir la parte del caudal que sale del depósito que entra al S0201. Este sector, también tiene el caudalímetro CC0203, pero como se ha comentado anteriormente, esa vía se encuentra cerrada y el sensor no está disponible, por lo que el consumo del sector viene dado por los sensores CC0210 y CC0211.

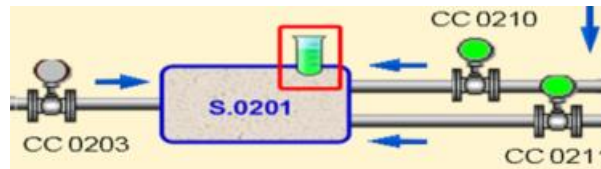


Figura 24. Sector S0201

De esta manera, el consumo del sector vendría dado por la siguiente ecuación:

Ecuación 4. Cálculo consumo S0201.

$$S0201 = CC0210 + CC0211$$

Así, se ha calculado el consumo del sector con los datos de 23 días durante diferentes épocas del año con el objetivo de conocer su comportamiento:

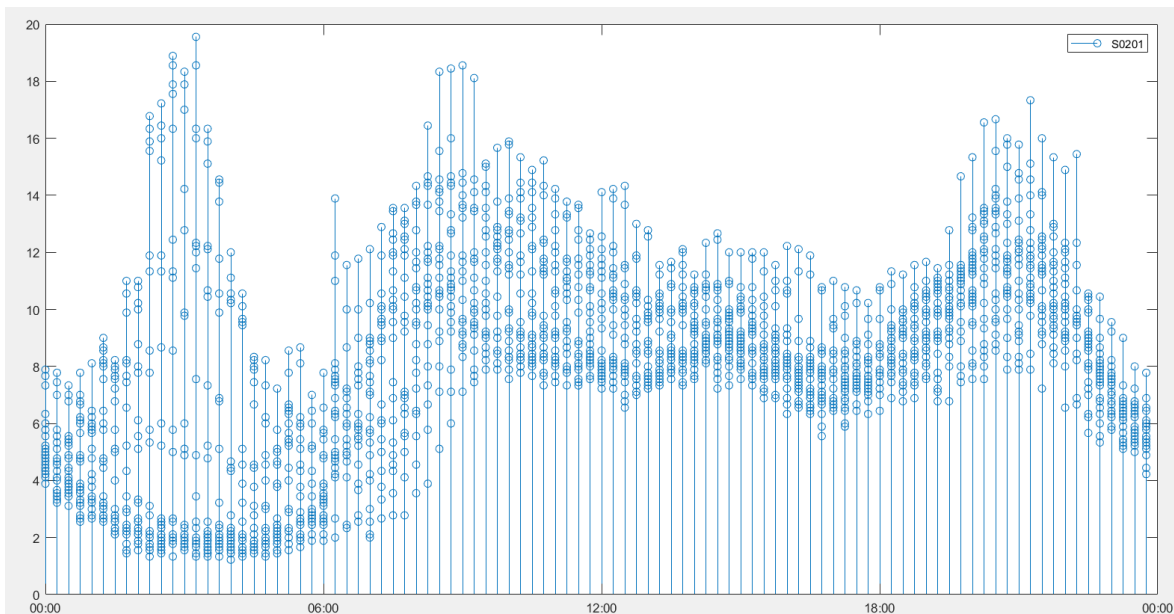


Figura 25. Datos de consumo S0201.

Tal y como se observa en la imagen anterior, se observa cómo parece haber un patrón diario, diferenciando las horas nocturnas, con las de la mañana y la tarde. Es por ello que la validación de estos sensores pasará por encontrar un intervalo de aceptación según la franja horaria en la que se encuentre y así poder validar estos datos según los históricos. Las franjas que se han tomado como referencia son las siguientes:

- Madrugada de 0h a 6h
- Mañana de 6h a 12h
- Tarde de 12h a 20h
- Noche de 20h a 00h

Con ello, teniendo los datos quinceminutales del sector durante 23 días, se pretende crear un entorno en el que una parte de esos 23 datos sea para estudiar y encontrar los modelos y la otra parte sea para entrenarlo, de manera que se ha establecido que 15 datos serán para el estudio, y los 8 restantes serán para testarlo.

Para encontrar el intervalo de validación, se utilizará la desviación estándar (σ) [8]. La desviación estándar es una medida utilizada para conocer la dispersión de un conjunto de datos, por ejemplo, en la figura anterior, en el tramo de la madrugada, habría una mayor desviación estándar que en el tramo de la tarde.

El dato de la desviación estándar es utilizado en estadística para estudiar la distribución normal de un conjunto de datos mediante la regla empírica [10] del "68-95-99.7", en la que se indica que el 68% de los valores se encontrarán en el intervalo entre la media de los valores más la desviación estándar y la media menos la desviación estándar, el 95% se encontrará entre la media más y menos 2 veces la desviación estándar y el 99,7% entre la media más y menos 3 veces la desviación estándar:

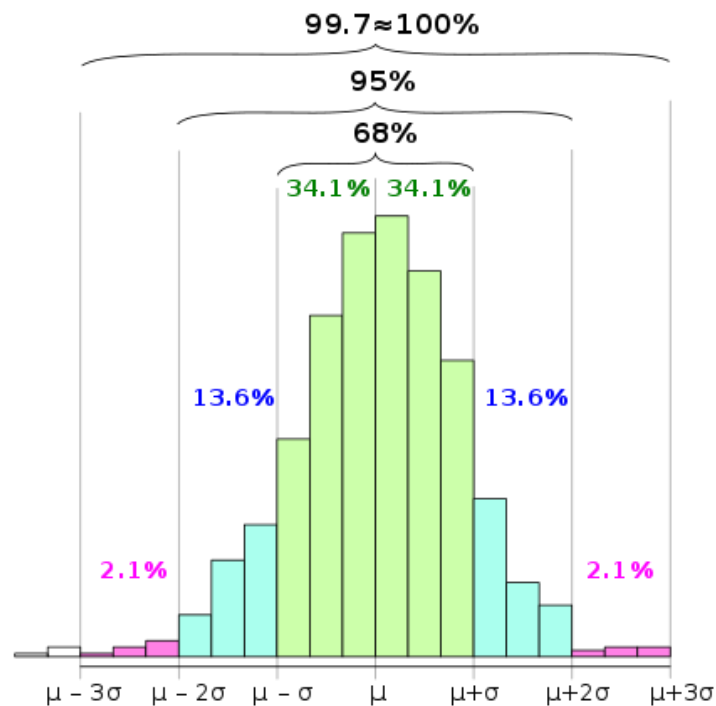


Figura 26. Regla de 68-95-99.7

Con ello, se realiza el cálculo de las desviaciones estándar en los diferentes tramos del día obteniendo los siguientes resultados:

Tabla 6. Desviaciones estándar por tramos en S0202.

Tramo	Desviación estándar
0h – 6h	4,3019
6h – 12h	3,8339
12h – 20h	1,5903
20h – 00h	2,8979

Así se realiza la media de los 15 datos quinceminutales y posteriormente se le sumará 3 veces la desviación estándar de su tramo, consiguiendo así el 99,7% de fiabilidad. En la siguiente figura se muestra la media de los valores en azul, el límite superior en rojo y el inferior en amarillo (se ha establecido la consigna que el límite inferior no puede ser menor a 0):

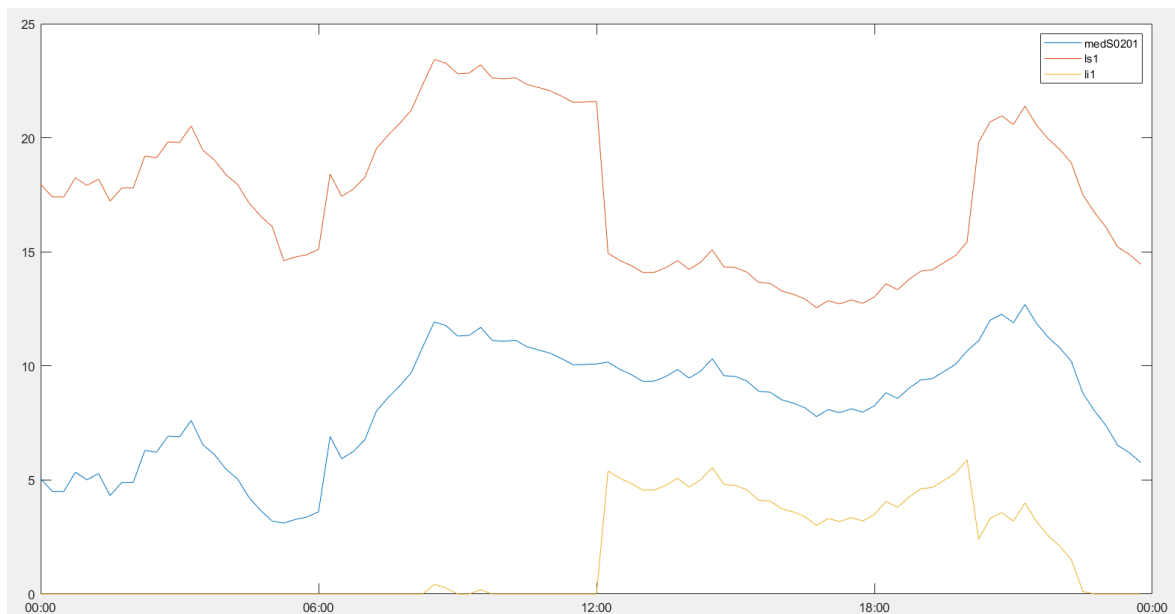


Figura 27. Media aritmética, límite superior y límite inferior del S0201.

Finalmente, se verifica el funcionamiento del método de validación con los 8 días restantes reservados para el entreno. En este caso se muestra como no hay ningún valor fuera de los límites, por lo que todos habrían pasado el test de validación:

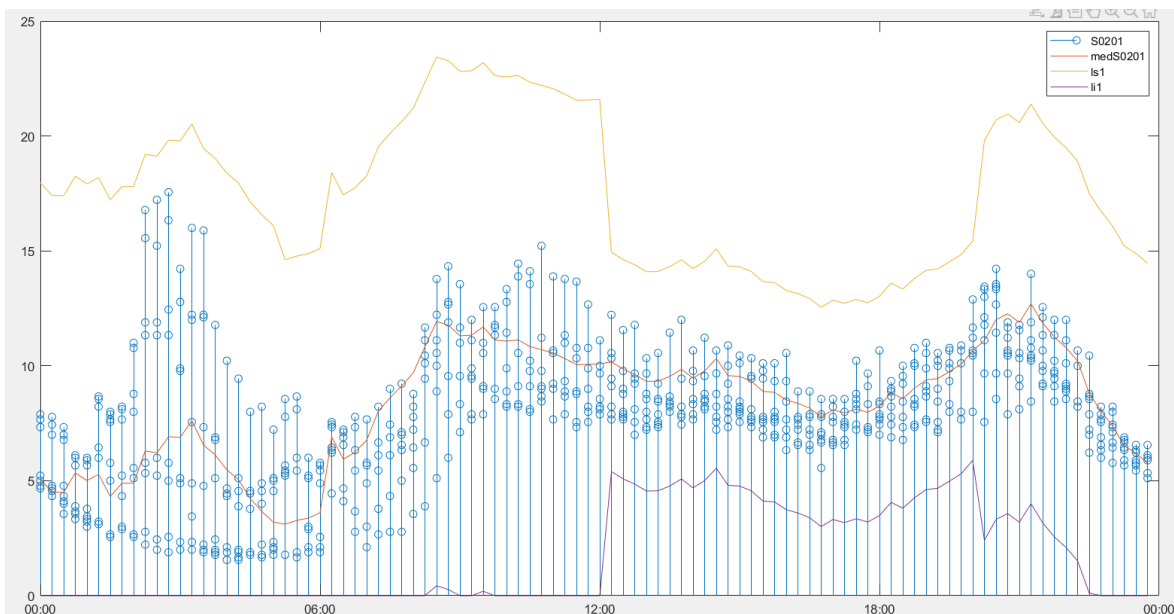


Figura 28. Entrenamiento método de validación S0201.

Una vez finalizado el test de validación con los límites superiores e inferiores, se realizaría el test del dato con valor nulo o fuera de rango. Como en todos los casos, pasa a ser el test de mayor importancia, ya que si falla este fallarán los demás con toda seguridad.

Tabla 7. Descripción de la validación para CC0210 y CC0211.

Nº	Descripción
0	Valor nulo o fuera de rango
1	Válido
6	Fallo en el sector.

Si se da el caso que con el consumo se valida que el dato no es correcto, pasaría a revisar los datos de ambos sensores, y en el caso que uno o los dos dé nulo o fuera de rango, se especificaría con un 0 mientras el otro seguiría con un 6.

En el caso que sea válido, se añadirá un 1. En la siguiente tabla se puede observar la validación realizada en un periodo de tiempo del 24 de abril de 2019:

S0201	lsp1	lip1	vals0201	CC0210	valCC0210	CC0211	valCC0211
4.1111	17.401	0	1	2.2222	1	1.8889	1
4.3333	18.248	0	1	2.3333	1	2	1
4.4444	17.914	0	1	2.3333	1	2.1111	1
5	18.188	0	1	2.7778	1	2.2222	1
4.7778	17.222	0	1	2.5556	1	2.2222	1
5	17.795	0	1	2.6667	1	2.3333	1
4.4444	17.803	0	1	2.4444	1	2	1
4.2222	19.205	0	1	2.2222	1	2	1
4.4444	19.119	0	1	2.4444	1	2	1
4.1111	19.82	0	1	2.2222	1	1.8889	1
3.8889	19.795	0	1	2.1111	1	1.7778	1
3.5556	20.517	0	1	1.8889	1	1.6667	1
3.1111	19.461	0	1	1.6667	1	1.4444	1
3.2222	19.025	0	1	1.6667	1	1.5556	1
3.2222	18.384	0	1	1.7778	1	1.4444	1
2.8889	17.957	0	1	1.5556	1	1.3333	1
2.3333	17.136	0	1	1.2222	1	1.1111	1
2.3333	16.572	0	1	1.2222	1	1.1111	1
2.2222	16.102	0	1	1.2222	1	1	1
3	14.613	0	1	1.5556	1	1.4444	1
3	14.775	0	1	1.5556	1	1.4444	1
2.5556	14.878	0	1	1.3333	1	1.2222	1
2.1111	15.108	0	1	1.1111	1	1	1

Figura 29. Validación S0201

4.6 CC0213 y CC0216 – Sector S0203

Los sensores CC0213 y CC0216 tienen participación en dos sectores directamente, en el sector S0203 en el que el primero es la entrada y el segundo la salida, por lo que la diferencia será el consumo (por el sensor CC0214 no pasa caudal), y en el sector S0204 en el que únicamente tiene una entrada la cual es el CC0216:

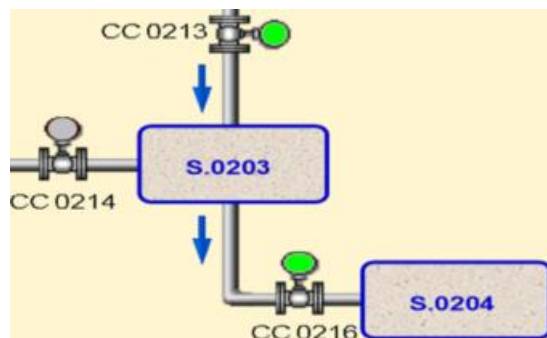


Figura 30. Sectores S0203 y S0204.

En esta validación se ha optado por seguir los pasos al sector S0201, realizando el test de validación con la desviación estándar por tramos y conociendo si estos datos son nulos o se encuentran fuera de rango.

Como se ha comentado con anterioridad en el sector S0204 hay un hotel que llena su propio depósito cuando lo necesitan, no hay periodicidad ninguna, por lo que puede ser llenado a cualquier hora del día. Esto se ve reflejado en el consumo del sector, que es directamente el caudal que pasa por el sensor CC0216, ya que como se puede observar

en la siguiente imagen, no se puede establecer un modelo claro de comportamiento, además de tener una gran desviación estándar en todos sus valores, por lo que tampoco cobraría sentido aplicar tratarla como distribución normal:

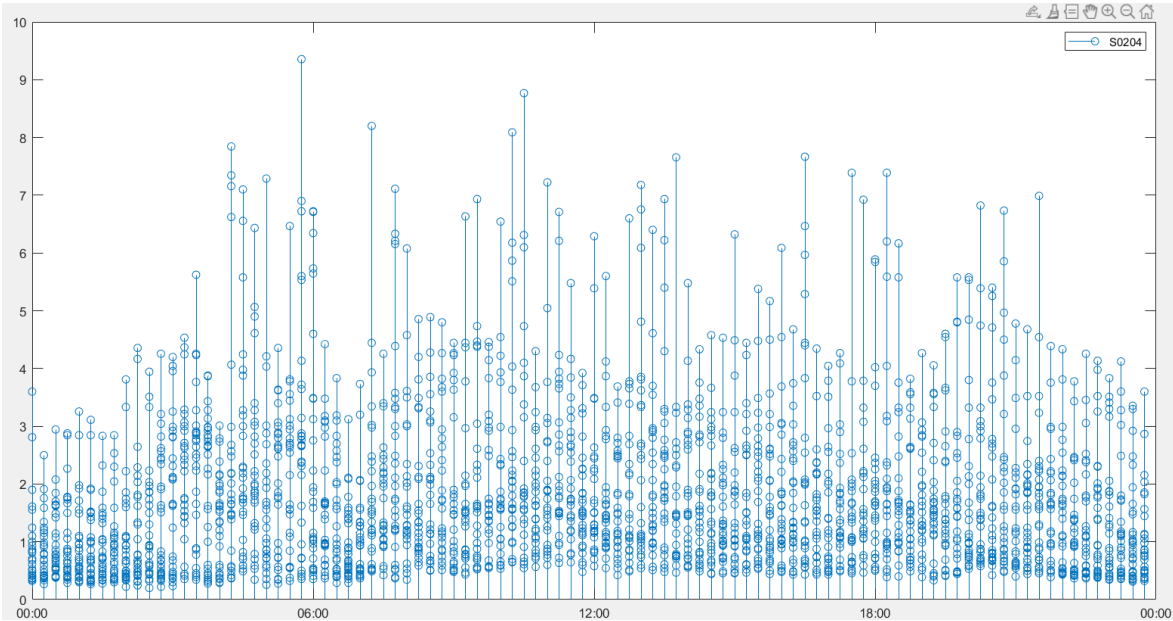


Figura 31. Consumo en el sector S0204.

Es por ello que para realizar la validación de los dos sensores, se utilizará el sector S0203. En este caso, también tiene participación el sensor CC0216, pero los picos de consumo del hotel se ven contrarrestados con el sensor CC0213, ya que también contiene esos picos:

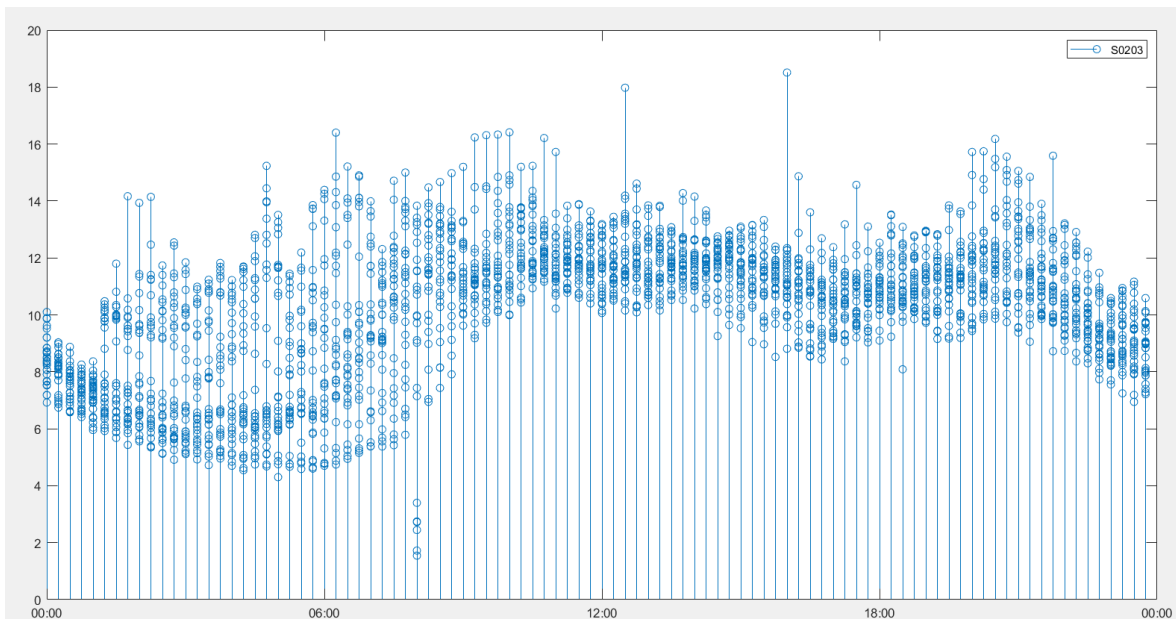


Figura 32. Consumo en el sector S0203.

En el caso del S0203, sí que se ve un modelo diferenciado, donde se muestra una mayor desviación estándar durante la madrugada y la mañana, y a partir de las 12h se recogen unos datos más cercanos entre sí.

Realizando la regla del “68-95-99.7” por tramos durante 15 días, tal y como se ha realizado en el subapartado anterior, recogemos los siguientes resultados:

Tabla 8. Desviaciones estándar por tramos en S0204.

Tramo	Desviación estándar
0h – 6h	2,1466
6h – 12h	2,5721
12h – 20h	1,0993
20h – 00h	1,8526

En la siguiente imagen se muestra la media aritmética quinceminutalmente en azul, el límite superior aceptado en rojo y el límite inferior en amarillo:

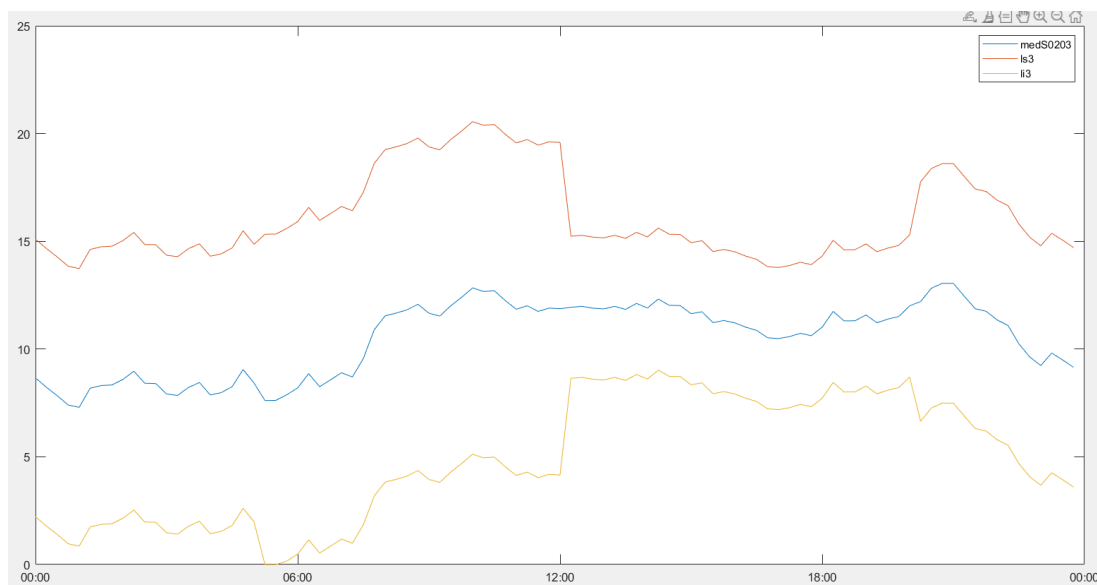


Figura 33. Media aritmética, límite superior y límite inferior en S0203.

Como en el caso anterior, se testea el funcionamiento con los datos de los 8 días restantes. En él se puede observar como hay algún dato que no se encuentra dentro de los límites establecidos, sobretodo en la tarde, ya que al tener la desviación estándar menor, los datos abruptos destacan más.

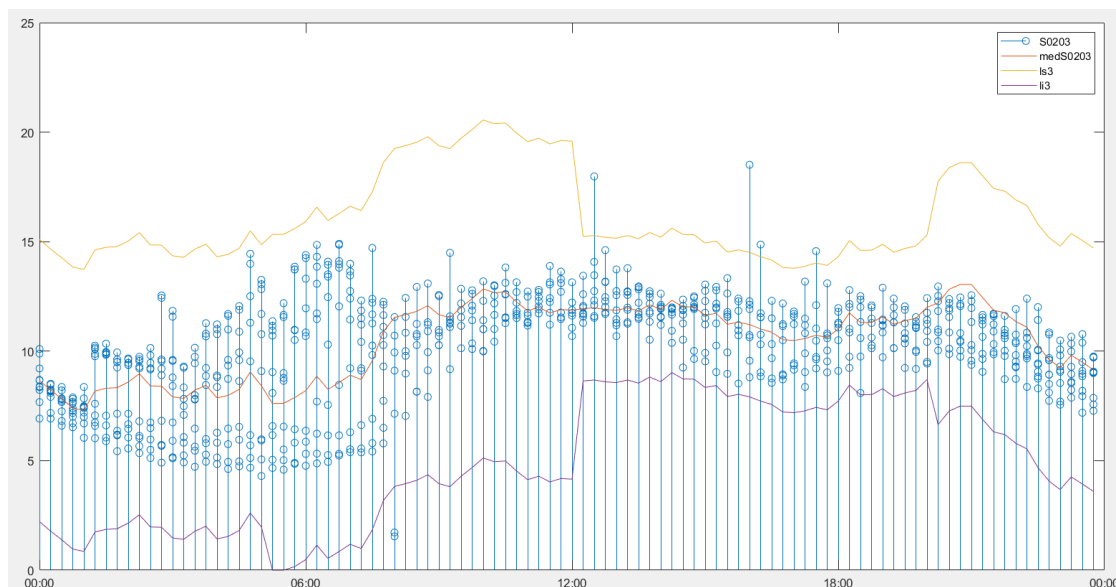


Figura 34. Entrenamiento método de validación S0203.

Para estos sensores, la categorización de la validación tiene los mismos factores que en el caso anterior:

Tabla 9. Descripción de la validación para CC0213 y CC0216.

Nº	Descripción
0	Valor nulo o fuera de rango
1	Válido
6	Fallo en el sector.

A continuación se muestra la validación realizada durante el 24 de abril de 2019:

S0203	lsp3	lip3	vals0203	CC0213	valCC0213	CC0216	valCC0216
7.8444	15.969	0.53631	1	8.5556	1	0.71111	1
8.4889	16.285	0.85261	1	9.1111	1	0.62222	1
9.1778	16.616	1.1837	1	9.6667	1	0.48889	1
10.011	16.419	0.98669	1	10.556	1	0.54444	1
11.244	17.275	1.8422	1	11.778	1	0.53333	1
13.711	18.613	3.18	1	14.222	1	0.51111	1
11.933	19.258	3.8252	1	12.556	1	0.62222	1
12.656	19.384	3.9519	1	14	1	1.3444	1
13.122	19.536	4.103	1	14.333	1	1.2111	1
13.044	19.794	4.3615	1	14.222	1	1.1778	1
12.322	19.383	3.9504	1	13.333	1	1.0111	1
10.556	19.244	3.8119	1	14.333	1	3.7778	1
10.933	19.721	4.2889	1	15.667	1	4.7333	1
10.967	20.119	4.6859	1	12.333	1	1.3667	1
11.067	20.556	5.123	1	15	1	3.9333	1
11.3	20.387	4.9541	1	15.333	1	4.0333	1
11.889	20.42	4.9874	1	13.667	1	1.7778	1
11.711	19.962	4.5296	1	13.111	1	1.4	1

Figura 35. Validación CC0213 y CC0214.

5 Reconstrucción de los datos

Tras la validación de los datos, se procede a la reconstrucción de los registros que han quedado invalidados en cada uno de los sensores del piso de Sulleva. Para realizar un modelo de reconstrucción, se debe tener en cuenta el tipo de sensor, para conocer la relación con otros componentes del sistema, así como los datos históricos y los datos previos a la fecha actual para reconstruir con un dato coherente a ese momento en específico.

En este apartado se presentarán las diferentes metodologías de reconstrucción estudiadas para este piso, así como la aplicación en cada uno de los sensores teniendo en cuenta qué fallo de validación se ha obtenido para enfocar correctamente la reconstrucción.

5.1 Metodologías de reconstrucción

Los tipos de reconstrucción de datos vendrán dados por la casuística del sensor que se haya definido para invalidar ese registro en concreto. De esta manera, la reconstrucción está muy relacionada al proceso de validación, ya que el mismo procedimiento para determinar la validez del dato, puede construir un camino para reconstruir o estimar el dato correcto.

A continuación, se especificaran las metodologías de reconstrucción que se han implementado en los sensores:

- **Coherencia local.** Se determina el valor estimado según los datos obtenidos de componentes relacionados del propio caudalímetro. El ejemplo más claro es el caudalímetro de entrada, ya que TAIGUA también ha proporcionado los valores de presión e intensidad de la bomba de impulsión.
- **Interpolación lineal.** En el caso que haya un único valor erróneo, se determina que el valor erróneo se estimará mediante interpolación lineal.
- **Serie temporal.** Se realiza un estudio de los últimos valores previos al dato erróneo para estimar dicho valor.
- **Reconstrucción semanal.** Se estimará el valor con el registro de una semana previa, ya que se considera que el consumo es constante semanalmente pero variable entre los días, ya que el consumo de los domingos no es el mismo que el de los lunes.
- **Datos históricos.** Con el estudio de validación se obtuvieron los datos medios diarios de consumo por cada 15 minutos, por lo que puede estimar el valor erróneo.
- **Modelo espacial.** Se realiza una relación de caudal entre diferentes sensores para conocer un valor estimado.

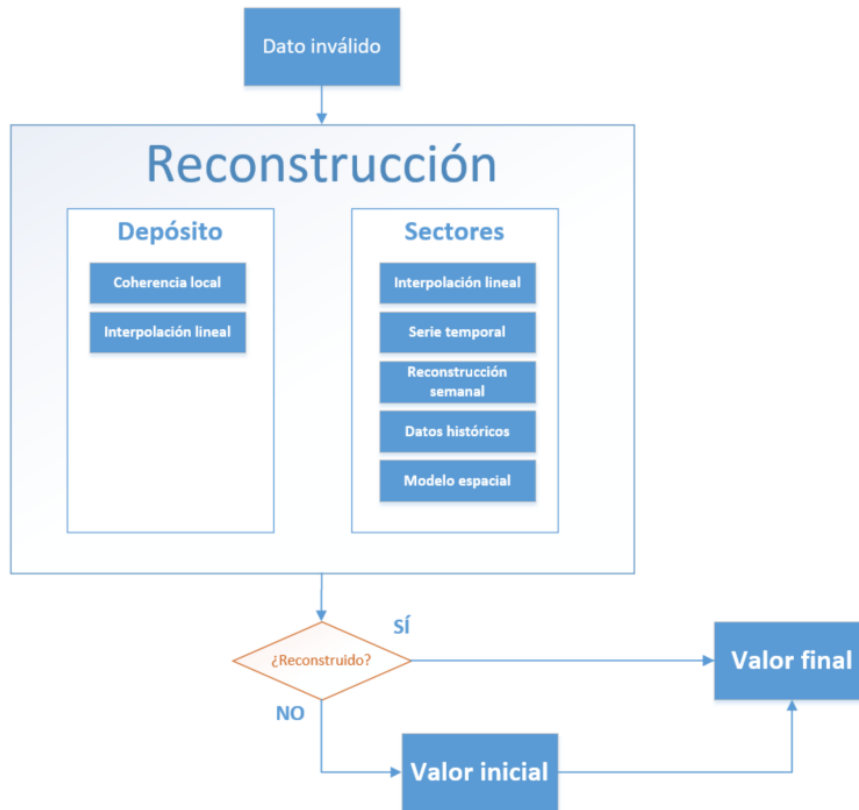


Figura 36. Proceso de reconstrucción

Tal y como muestra la figura anterior, según de donde proceda el dato inválido se usarán diferentes métodos de reconstrucción:

- Depósito: Se encuentran los caudalímetros de entrada (CC0201) y salida del depósito (CC0202), así como los sensores de nivel de agua (NIV1 y NIV2).
- Sectores: Se tienen en cuenta los dos sectores con los que se puede trabajar, como son el sector S0201, con los caudalímetros CC0210 y CC0211, y el sector S0203 con los caudalímetros CC0213 y CC0216.

Una vez realizado el proceso de reconstrucción se determinará el valor final que se ha obtenido en el modelo, por ejemplo en los casos válidos, el valor final será el mismo que el registrado por el sensor. En los valores que hayan pasado por el proceso de reconstrucción, se especificará el valor final que se ha estimado, o en el caso que no haya sido posible, se dejará el mismo que se registró.

Por último, se especificará mediante un 1 si ese dato ha sido reconstruido o un 0 si no ha sido necesaria la reconstrucción o no ha sido posible.

5.2 CC0201 – Bombeo al depósito

El primer caudalímetro a tratar será el caudalímetro CC0201. Como ya se ha comentado con anterioridad, este sensor aporta la información del caudal en l/s que entra al depósito de Sulleva.

En el apartado de validación, debido a la falta de información, sobretodo en la salida del depósito, era difícil conocer la validez de los datos en el caudalímetro de entrada. Además también se da la casuística que el depósito puede ser llenado por otra entrada de la cual no hay registros, por lo que TAIGUA proporcionó datos de la bomba de impulsión al depósito. En estos datos, se recogen los datos de la intensidad de la bomba y de la presión de la vía.

El proceso de reconstrucción del caudalímetro CC0201 se ha enfocado por coherencia local, de manera que teniendo los tres sensores será suficiente para estimar el valor. Esto es debido principalmente a que la bomba únicamente tiene 2 estados (encendida y apagada), por lo que facilita la estimación.

La reconstrucción se ha realizado de manera que:

- Si el estado de validación es un 4 es que se ha llenado el depósito mediante otra entrada. Se asume que el valor es correcto ya que la causa de que suba el nivel es el llenado externo.
- Si el estado de validación es un 3 significa que ha habido un fallo de coherencia con el sensor de presión. En este caso se compara el valor con el sensor de intensidad y si concuerda con el caudalímetro, se reconstruye con el mismo valor previo.
- Si el estado de validación es un 2 es que el sensor de intensidad no concuerda con el caudalímetro. Se compara con la presión y si equivale el estado, se reconstruye con el valor registrado.
- Si el estado de validación es un 0, significa que es un valor fuera de rango o inexistente. En este caso se revisan los sensores de la bomba y se determina el estado (encendido o apagado) y se estima con 0 l/s o 40 l/s.

A continuación se muestran algunos ejemplos de la reconstrucción de los datos. En estos ejemplos aparecen 5 columnas, la fecha (daet), el dato registrado (CC0201), el estado de la validación (valCC0201), el valor tras haber pasado la reconstrucción (R_CC0201) y si ha pasado por la reconstrucción (recCC0201).

En este primer ejemplo se observa como todos los valores han sido válidos, por lo que no ha sido necesario cambiar el valor a excepción del dato de las 4:45 en el que había discordancia con el amperímetro de la bomba. Tras revisarlo, con el sensor de presión se confirma que el valor era correcto, por lo que se deja el mismo valor en el campo R_CC0201 y se añade un 0 en el campo recCC0201 conforme en la reconstrucción no ha cambiado su valor.

date	CC0201	valCC0201	R_CC0201	recCC0201
01-Jan-2019 23:15:00	0	1	0	0
01-Jan-2019 23:30:00	0	1	0	0
01-Jan-2019 23:45:00	0	1	0	0
02-Jan-2019 00:00:00	0	1	0	0
02-Jan-2019 00:15:00	20.65	1	20.65	0
02-Jan-2019 00:30:00	40.099	1	40.099	0
02-Jan-2019 00:45:00	40.094	1	40.094	0
02-Jan-2019 01:00:00	40.129	1	40.129	0
02-Jan-2019 01:15:00	40.048	1	40.048	0
02-Jan-2019 01:30:00	40.079	1	40.079	0
02-Jan-2019 01:45:00	40.069	1	40.069	0
02-Jan-2019 02:00:00	40.119	1	40.119	0
02-Jan-2019 02:15:00	40.099	1	40.099	0
02-Jan-2019 02:30:00	40.135	1	40.135	0
02-Jan-2019 02:45:00	40.14	1	40.14	0
02-Jan-2019 03:00:00	40.114	1	40.114	0
02-Jan-2019 03:15:00	40.119	1	40.119	0
02-Jan-2019 03:30:00	40.124	1	40.124	0
02-Jan-2019 03:45:00	40.114	1	40.114	0
02-Jan-2019 04:00:00	40.15	1	40.15	0
02-Jan-2019 04:15:00	35.096	1	35.096	0
02-Jan-2019 04:30:00	0	1	0	0
02-Jan-2019 04:45:00	0	2	0	0
02-Jan-2019 05:00:00	0	1	0	0
02-Jan-2019 05:15:00	0	1	0	0
02-Jan-2019 05:30:00	0	1	0	0
02-Jan-2019 05:45:00	0	1	0	0
02-Jan-2019 06:00:00	0	1	0	0
02-Jan-2019 06:15:00	0	1	0	0

Figura 37. Primera reconstrucción del sensor CC0201.

En el siguiente y último ejemplo se observan dos casuísticas más. Entre las 7:30 y las 8:30 con el estado de validación en 4, se verifica el estado de la bomba con los otros sensores y en el caso de confirmarlo, se deja el mismo valor registrado del caudalímetro y se indica con un 0 en la reconstrucción ya que no se ha modificado el valor.

Por otro lado, hay un caso de valor perdido a las 8:45. Al no tener valor del caudalímetro se comprueba el estado de la bomba con los otros dos sensores. En ese caso, se certifica que estaba parada por lo que se cambia el valor a 0 l/s y se da por reconstruido.

date	CC0201	valCC0201	R_CC0201	recCC0201
03-Jan-2019 06:30:00	0	1	0	0
03-Jan-2019 06:45:00	0	1	0	0
03-Jan-2019 07:00:00	0	1	0	0
03-Jan-2019 07:15:00	0	1	0	0
03-Jan-2019 07:30:00	0	4	0	0
03-Jan-2019 07:45:00	0	4	0	0
03-Jan-2019 08:00:00	0	4	0	0
03-Jan-2019 08:15:00	0	4	0	0
03-Jan-2019 08:30:00	0	4	0	0
03-Jan-2019 08:45:00	NaN	0	0	1
03-Jan-2019 09:00:00	17.101	1	17.101	0
03-Jan-2019 09:15:00	0	1	0	0
03-Jan-2019 09:30:00	0	1	0	0
03-Jan-2019 09:45:00	0	1	0	0
03-Jan-2019 10:00:00	0	1	0	0
03-Jan-2019 10:15:00	0	1	0	0
03-Jan-2019 10:30:00	0	1	0	0
03-Jan-2019 10:45:00	30.179	1	30.179	0
03-Jan-2019 11:00:00	39.952	1	39.952	0
03-Jan-2019 11:15:00	39.912	1	39.912	0
03-Jan-2019 11:30:00	39.927	1	39.927	0
03-Jan-2019 11:45:00	39.912	1	39.912	0
03-Jan-2019 12:00:00	39.913	1	39.913	0
03-Jan-2019 12:15:00	4.9826	1	4.9826	0
03-Jan-2019 12:30:00	0	1	0	0
03-Jan-2019 12:45:00	0	1	0	0
03-Jan-2019 13:00:00	0	1	0	0

Figura 38. Segunda reconstrucción del sensor CC0201.

5.3 NV_SU1 y NV_SU2 – Niveles de depósito

Los siguientes caudalímetros a reconstruir son los dos sensores de nivel de agua del depósito. Como se ha comentado con anterioridad, se trata de dos sensores que deben proporcionar un valor muy aproximado entre sí, ya que están situados en dos cámaras diferentes pero que están interconectadas entre ellas, por lo que el nivel debe ser el mismo en ambas.

Para estos sensores la validación tenía tres estados diferentes: 0 si no había valor o estaba fuera de rango, 1 si era válido o 5 si había fallo con los valores previos.

En el primero de los casos, se verifica que únicamente falla un dato (representado por un 0). Si es así, se realiza una interpolación lineal entre el anterior y el siguiente, conociendo el valor intermedio entre ambos. Se podría realizar para más valores, pero se considera que el valor en más de media hora puede cambiar sustancialmente.

Si fallan más datos, se recurre a la ecuación 2 (mostrada en el apartado 5.4), donde si el sensor de entrada (CC0201) y de salida del depósito (CC0202) están validados, se estima el nivel del depósito.

Por último, caso que por ejemplo el sensor de nivel 1 haya dado erróneo debido a la ausencia de dato, pero el de nivel 2 haya dado correcto, se estima que es el mismo valor para ambos sensores.

En el caso que no se haya podido reconstruir, se copia el valor recogido y se informa que no se ha podido reconstruir.

En el siguiente ejemplo se muestran datos del 2 de enero de 2019 de ambos sensores de nivel:

date	NV1	valNV1	R_NV1	recNV1	NV2	valNV2	R_NV2	recNV2
02-Jan-2019 04:00:00	3.5714	1	3.5714	0	3.5899	1	3.5899	0
02-Jan-2019 04:15:00	3.6028	1	3.6028	0	3.6454	1	3.6454	0
02-Jan-2019 04:30:00	3.5825	1	3.5825	0	3.6251	1	3.6251	0
02-Jan-2019 04:45:00	3.5604	1	3.5604	0	3.6048	1	3.6048	0
02-Jan-2019 05:00:00	3.5419	1	3.5419	0	3.5863	1	3.5863	0
02-Jan-2019 05:15:00	3.5178	1	3.5178	0	3.5604	1	3.5604	0
02-Jan-2019 05:30:00	3.4882	1	3.4882	0	3.5344	1	3.5344	0
02-Jan-2019 05:45:00	3.4586	1	3.4586	0	3.5049	1	3.5049	0
02-Jan-2019 06:00:00	3.429	1	3.429	0	3.4752	1	3.4752	0
02-Jan-2019 06:15:00	3.3994	1	3.3994	0	3.4456	1	3.4456	0
02-Jan-2019 06:30:00	3.3717	1	3.3717	0	3.4161	1	3.4161	0
02-Jan-2019 06:45:00	3.3384	1	3.3384	0	3.3827	1	3.3827	0
02-Jan-2019 07:00:00	3.3088	1	3.3088	0	3.3532	1	3.3532	0
02-Jan-2019 07:15:00	3.2737	1	3.2737	0	3.3199	1	3.3199	0
02-Jan-2019 07:30:00	3.2311	1	3.2311	0	3.2792	1	3.2792	0
02-Jan-2019 07:45:00	3.2551	1	3.2551	0	3.2774	1	3.2774	0
02-Jan-2019 08:00:00	NaN	0	3.2931	1	NaN	0	3.3134	1
02-Jan-2019 08:15:00	3.331	5	3.331	1	3.3495	5	3.3495	1
02-Jan-2019 08:30:00	3.3513	1	3.3513	0	3.368	1	3.368	0
02-Jan-2019 08:45:00	3.3717	1	3.3717	0	3.3847	1	3.3847	0
02-Jan-2019 09:00:00	3.3865	1	3.3865	0	3.3994	1	3.3994	0
02-Jan-2019 09:15:00	3.3994	1	3.3994	0	3.4124	1	3.4124	0
02-Jan-2019 09:30:00	3.4142	1	3.4142	0	3.429	1	3.429	0
02-Jan-2019 09:45:00	3.4309	1	3.4309	0	3.4475	1	3.4475	0
02-Jan-2019 10:00:00	3.4475	1	3.4475	0	3.4641	1	3.4641	0
02-Jan-2019 10:15:00	3.4549	1	3.4549	0	3.4697	1	3.4697	0
02-Jan-2019 10:30:00	3.4752	1	3.4752	0	3.49	1	3.49	0

Figura 39. Reconstrucción de sensores de nivel del depósito.

En la figura anterior se observan dos fallos consecutivos en la columna valNV1 y valNV2 a las 8:00 y 8:15, al no existir el primer valor, el segundo no se puede confirmar por lo que también es inválido. En este caso, se realiza interpolación lineal para estimar el valor perdido (en las columnas R_NV1 y R_NV2) y se verifica que el siguiente valor también es correcto, por lo que no se modifica el valor.

5.4 CC0202 – Salida del depósito

La salida del depósito viene representada por el sensor CC0202 que a su vez es la entrada del piso de Sullea. Este sensor como se ha comentado anteriormente, tiene una gran ausencia de datos por lo que será, con diferencia, el que más se tenga que reconstruir.

Para este caudalímetro, sea cuál sea el estado de la validación que determine la causa del fallo del dato, se procederá con el mismo procedimiento de reconstrucción, la ecuación 2. En esta ecuación se tiene en cuenta los valores del caudalímetro de entrada del depósito (CC0201) y los dos sensores de nivel, que conociendo el área del depósito se extrae una relación de caudal entrada-salida.

A pesar de que sea el único método de reconstrucción, se calcula con tres de los sensores con más fiabilidad de todo el sistema, ya que como se ha comentado, se puede contrastar la información desde varios puntos.

date	CC0202	valCC0202	R_CC0202	recCC0202
25-Jul-2019 11:00:00	NaN	0	37.645	1
25-Jul-2019 11:15:00	NaN	0	32.314	1
25-Jul-2019 11:30:00	36.847	1	36.847	0
25-Jul-2019 11:45:00	32.7	1	32.7	0
25-Jul-2019 12:00:00	28.117	1	28.117	0
25-Jul-2019 12:15:00	31.121	1	31.121	0
25-Jul-2019 12:30:00	35.936	1	35.936	0
25-Jul-2019 12:45:00	34.01	1	34.01	0
25-Jul-2019 13:00:00	34.652	1	34.652	0
25-Jul-2019 13:15:00	34.177	1	34.177	0
25-Jul-2019 13:30:00	29.94	1	29.94	0
25-Jul-2019 13:45:00	27.898	1	27.898	0
25-Jul-2019 14:00:00	27.256	1	27.256	0
25-Jul-2019 14:15:00	31.866	1	31.866	0
25-Jul-2019 14:30:00	31.584	1	31.584	0
25-Jul-2019 14:45:00	27.45	1	27.45	0
25-Jul-2019 15:00:00	28.682	1	28.682	0
25-Jul-2019 15:15:00	31.814	1	31.814	0
25-Jul-2019 15:30:00	30.466	1	30.466	0
25-Jul-2019 15:45:00	27.796	1	27.796	0
25-Jul-2019 16:00:00	26.602	1	26.602	0
25-Jul-2019 16:15:00	28.194	1	28.194	0
25-Jul-2019 16:30:00	30.171	1	30.171	0
25-Jul-2019 16:45:00	31.943	1	31.943	0
25-Jul-2019 17:00:00	30.71	1	30.71	0
25-Jul-2019 17:15:00	25.755	1	25.755	0
25-Jul-2019 17:30:00	24.856	1	24.856	0
25-Jul-2019 17:45:00	24.804	5	41.936	1
25-Jul-2019 18:00:00	30.145	1	30.145	0
25-Jul-2019 18:15:00	36.975	1	36.975	0

Figura 40. Reconstrucción del sensor CC0202.

En la figura anterior se muestran dos casos diferentes de reconstrucción, pero afrontados de la misma manera:

- Cuando el valor registrado no existe, en el caso de las 11:00 y las 11:15, se estima mediante la ecuación 2 con los valores de los diferentes componentes del depósito.
- Valor fuera de lo permitido, como sucede a las 17:45 del ejemplo anterior y también se estima con la ecuación 2.

5.5 CC0210 y CC0211 – Sector S0201

En este subapartado se reconstruirán los datos de los sensores que influyen directamente en el sector S0201, que son los caudalímetros CC0210 y CC0211. Cabe recordar que, como se ha comentado en la ecuación 4, el consumo del sector es la suma de ambos caudalímetros.

Como en el caso de la validación, se ha enfocado de manera que para poder estimar los valores de los caudalímetros, se partirá del consumo de sector para establecer coherencia en el sistema.

Como primera metodología a aplicar, se revisará si es un único valor el que falla o que está ausente. De esta manera se realizará una interpolación lineal con el valor anterior y el

siguiente. Al tratarse de un sector, se revisará de qué sensor es el valor que ha sido erróneo para sustituirlo, y calcular el consumo estimado del sector.

En otras ocasiones, puede haber un gran conjunto de datos perdidos que pueden llegar a ser días e incluso semanas. Por ejemplo la siguiente figura extraída de los datos de las dos primeras semanas de enero de 2019. En ella se observa que una vez pasada la primera semana, deja de haber datos del caudalímetro CC0211, por lo que el sector S0201 a partir del 7 de enero tiene el mismo valor que el CC0210.

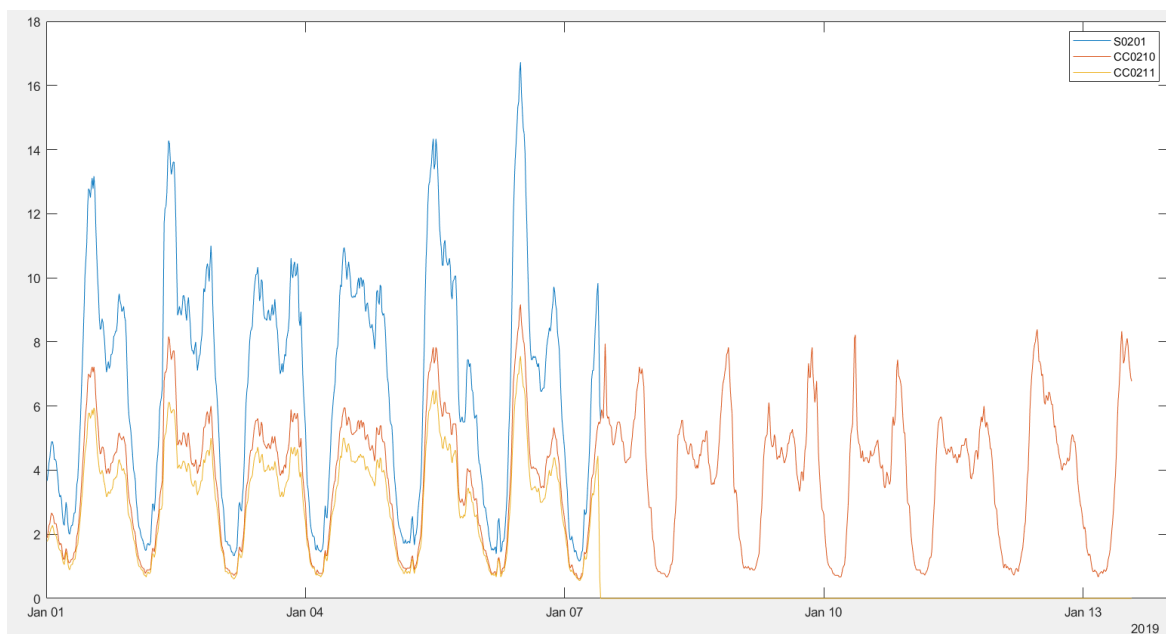


Figura 41. Datos registrados en S0201 sin reconstruir.

Para estudiar estas situaciones se debe hallar un modelo que sea capaz de determinar qué valor correspondería a esa situación. Para ello se ha de tener en cuenta diferentes factores que influyen directamente sobre ese valor, ya que según la hora del día o el día de la semana en que se encuentre puede haber una importante variación en él.

Otro factor importante sobre el que enfocar la reconstrucción de los datos son los valores previos. Teniendo una cantidad de valores válidos previos al dato erróneo se puede estimar qué valor debería tener. En este caso se estaría hablando de un modelo autorregresivo (AR) [12], el cual la variable de salida depende linealmente de sus valores anteriores.

El polinomio característico de los modelos autorregresivos es el siguiente:

Ecuación 5. Modelo autorregresivo.

$$Y_t = c + \phi_1 * Y_{t-1} + \dots + \phi_p * Y_{t-p}$$

Donde 'Y_t' es el valor estimado, 'c' es una constante determinada por el modelo, 'φ_x' son los parámetros del modelo y por último, 'Y_{t-1}' ... 'Y_{t-p}' son los valores previos.

Para este caso, al tener una variación quinceminutal de los datos, donde su mayor similitud serían los días previos a la misma hora, carece de sentido realizar un modelo autorregresivo con todos los datos quinceminutales. Es por ello que se plantea un modelo que estime el consumo del día actual y no del minuto. Para ello se ha realizado la suma de

los consumo de los 5 días previos y se ha realizado un modelo de autorregresión con esos 5 datos.

Una vez obtenido el consumo estimado del día actual, se calcula el porcentaje usual de consumo de ese minuto en específico respecto al consumo total de un día. Este cálculo se ha realizado con los datos de la validación.

Tras calcular el dato estimado del sector, se resta el caudalímetro CC0210 para conocer el valor del CC0211.

Realizando un único filtro posterior al modelo autorregresivo de coherencia, como es que los valores del caudalímetro CC0211 y consecuentemente del sector S0201 no pueden ser negativos, se extrae el siguiente resultado:

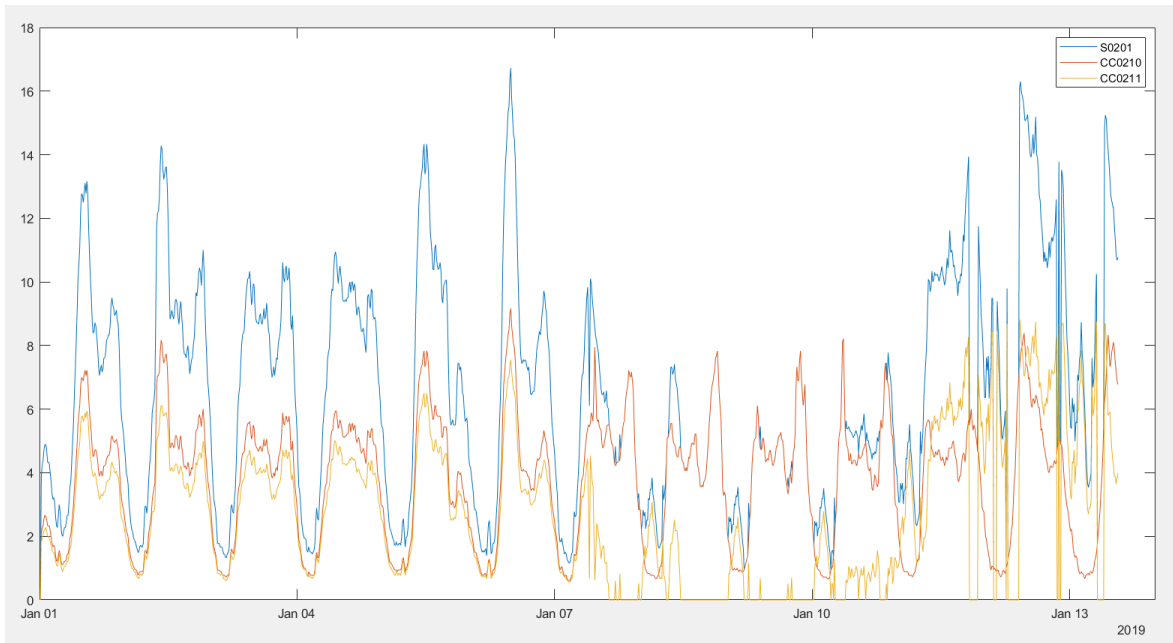


Figura 42. Reconstrucción S0201 con AR.

En este caso, en todos los valores negativos se ha vuelto a establecer el mismo valor recogido por el propio sensor. Tal y como se observa en la anterior figura, el modelo autorregresivo no ha sido excesivamente eficaz, ya que en ninguno de los días ha estimado correctamente el consumo del sector excepto el algún momento del principio de la reconstrucción que ha sido más coherente.

Partiendo de este resultado, se procede añadir más herramientas de reconstrucción de datos. En un primer momento, se determinará si el valor ha sido reconstruido correctamente, para así entrar en otra reconstrucción. Para ello, el primer paso será determinar los valores máximos y mínimos de los diferentes sensores según los valores históricos y añadiendo un margen de seguridad de un 10%:

- CC0210 se debe encontrar entre 0.6 l/s – 10 l/s.
- CC0211 se debe encontrar entre 0.5 l/s – 9 l/s.

Otro punto importante, es que el caudalímetro CC0211 siempre es menor que el CC0210, pero se debe determinar cuánto. Revisando los datos históricos, se determina que la media de la relación entre ambos caudalímetros es de 0.845 unidades, pero se han encontrado

relaciones menores al 0.60, por lo que se considerará que un valor del CC0211 siendo el 50% del caudalímetro CC0210 se considerará bien reconstruido.

A continuación se muestra un ejemplo de varios días de la relación entre los caudalímetros CC0211 y CC0210:

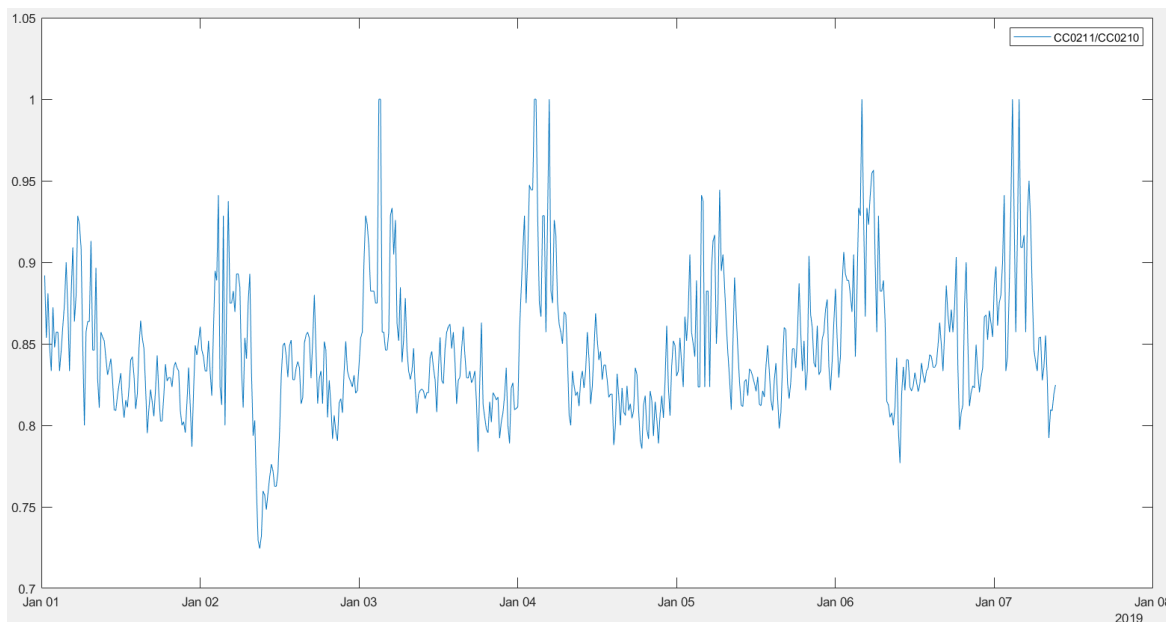


Figura 43. Relación entre CC0210 y CC0211.

En el caso que la reconstrucción por el modelo autorregresivo no haya pasado por lo anteriores filtros, se recurrirá a revisar los datos del mismo día de la semana, pero de la semana anterior. En ese caso el consumo del sector sería el mismo, pero el caudal del sensor que ha registrado el fallo con casi total seguridad no, ya que por ejemplo, si tenemos valores del sensor CC0210 y se estima el consumo del sector con el dato de una semana atrás, el valor reconstruido del sensor CC0211 se determinará la diferencia de ambos, dando así un nuevo valor.

En el caso que tampoco se validen estos datos reconstruidos, se recurrirá a los datos medios históricos. En el proceso de validación se extrajo con la información de varios días un día "modelo" que puede ser utilizado en caso que fallen las anteriores técnicas de reconstrucción o en el caso que no haya datos anteriores a los que recurrir.

Como última opción, si los datos medios de consumo no son aceptados por los filtros, se calculará el valor erróneo mediante la relación entre ambos caudalímetros. Teniendo la media de relación CC0211/CC0210 siendo un 0.845 se estimará el valor restante.



Figura 44. Proceso de reconstrucción S0201.

A continuación se muestra el ejemplo de reconstrucción completa del sector S0201 utilizando las metodologías mencionadas:

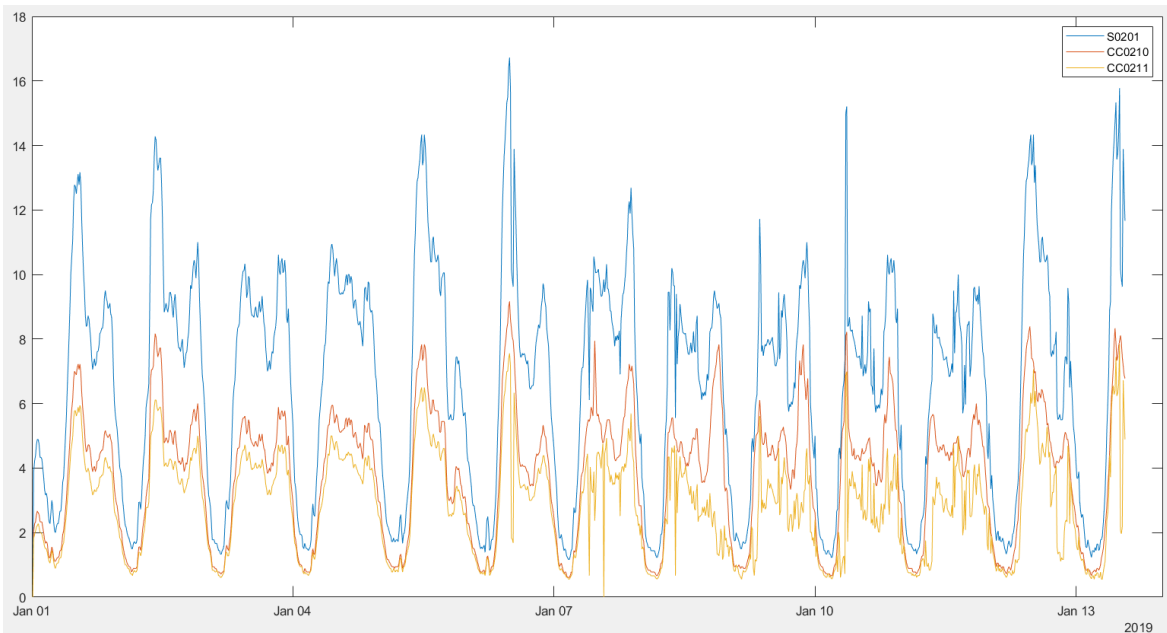


Figura 45. Reconstrucción CC0210 y CC0211.

En la figura anterior se observa como los valores de consumo y del caudalímetro CC0211 reconstruidos (a partir del 7 de enero) se asemejan mucho más a los valores reales. Pese a ello, sigue habiendo algún valor que no ha sido posible reconstruir.

5.6 CC0213 y CC0216 – Sector S0203

En este apartado se realizará la reconstrucción del consumo del sector S0203, que es calculado por la diferencia entre los caudalímetros CC0213 y CC0216.

Para este sector se ha aplicado el mismo desarrollo de la reconstrucción que en el sector anterior, empezando por la interpolación lineal. En caso que haya un valor cuyos datos previos y posteriores sean válidos, se estimará el valor erróneo mediante interpolación lineal. En este caso como en todos los de la reconstrucción del sector, se realizará la estimación del sector y posteriormente se calculará mediante el balance de masas de sector el valor erróneo.

Para mostrar la reconstrucción, se partirá del siguiente ejemplo que compone las dos primeras semanas de enero de 2019.

Tal y como se observa en la siguiente figura, hay una ausencia de datos del caudalímetro CC0216 a partir del día 11 de enero. Los valores recogidos por los sensores CC0213 y CC0216 y el consumo del sector son los siguientes:

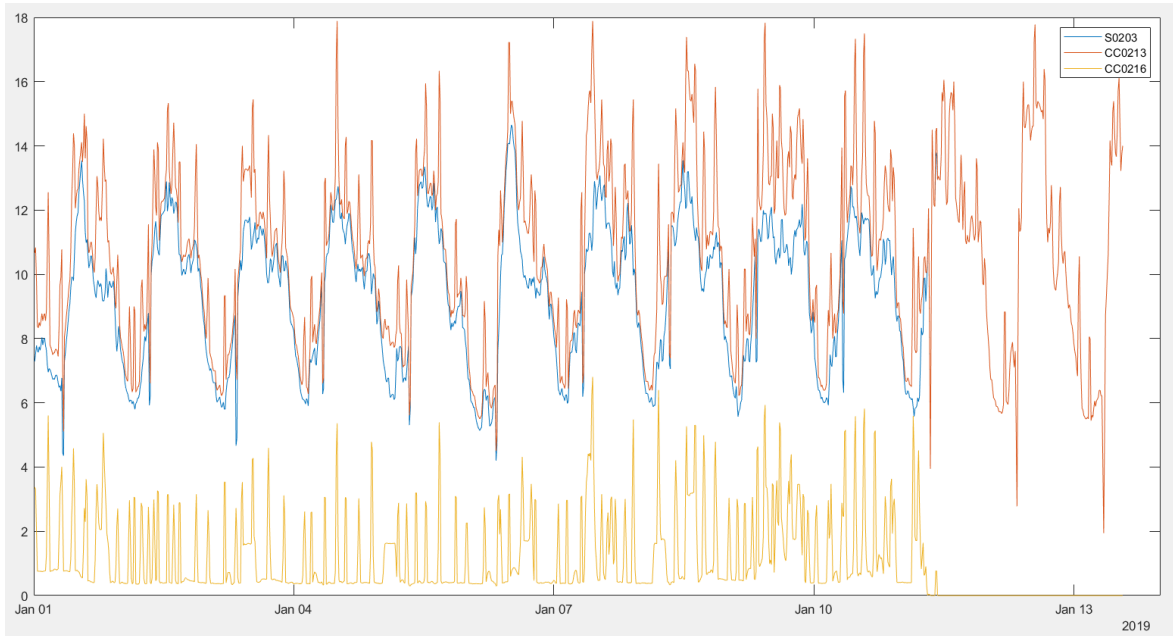


Figura 46. Datos en S0203 sin reconstruir.

Para este sector también se ha aplicado el modelo autorregresivo (ecuación 5) para estimar cuando fallen una serie de valores consecutivos. Se recogerán los consumos totales de los 5 días previos y se estimará el consumo del día actual. Con ello conociendo la proporción de consumo de cada minuto del día se estimará el consumo quinceminutal del sector en ese instante. Esta proporción se ha extraído de los datos históricos que se han estudiado en la validación.

A continuación se muestra la reconstrucción mediante el modelo AR, filtrando únicamente que los valores reconstruidos deben ser positivos:

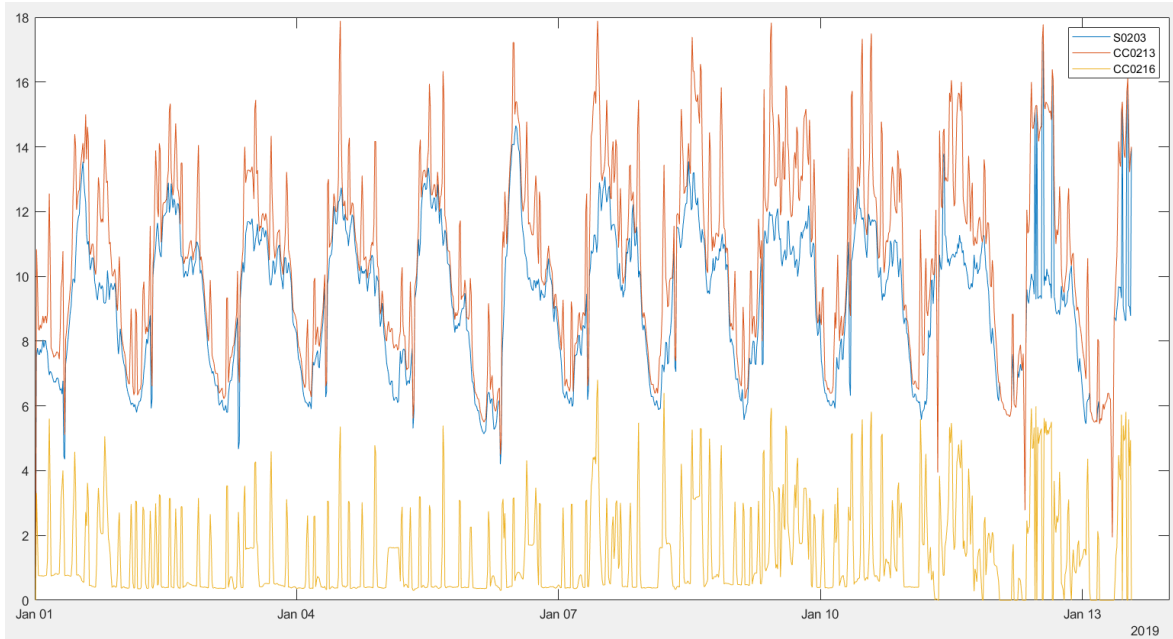


Figura 47. Reconstrucción del sector S0203 con AR.

Tal y como se observa en la figura anterior, se puede determinar que la reconstrucción mediante el modelo AR ha resultado más satisfactoria que en el sector anterior. Se observa que el comportamiento del consumo diario es más constante, pese a que hay ciertos puntos que existen ciertos picos.

Con tal de mejorar la reconstrucción se realizarán una serie de filtros. En el primer caso, se impondrán unos valores máximos y mínimos revisando el total de los datos de 2019. Aplicando un factor de seguridad del 10% los máximos y mínimos de los caudalímetros son los siguientes:

- CC0213 se debe encontrar entre 4 l/s – 20 l/s.
- CC0216 se debe encontrar entre 0.25 l/s – 8 l/s.

Por otro lado, el consumo del sector será limitado por los intervalos de confianza determinados en la validación del sector S0203 (visible en la figura 32).

Una vez realizado el filtro, en el caso que no haya sido aceptado, se revisará el valor de una semana atrás y se cogerá el mismo consumo que entonces. El consumo será aceptado, pero se volverá a filtrar los datos de los caudalímetros por los intervalos marcados anteriormente.



Figura 48. Proceso reconstrucción S0203.

En el caso que no haya sido posible reconstruir mediante el dato de la semana previa, se reconstruirá con el valor medio diario calculado en la validación. Como en el caso anterior, el consumo del sector será correcto, pero al calcular el valor del caudalímetro erróneo, puede dar fallo en alguno de los fatos.

Como última opción, en caso que el valor del caudalímetro CC0216 (que es el más frecuente en fallo) dé un valor negativo, se sustituirá por el valor mínimo que es 0.25 l/s.

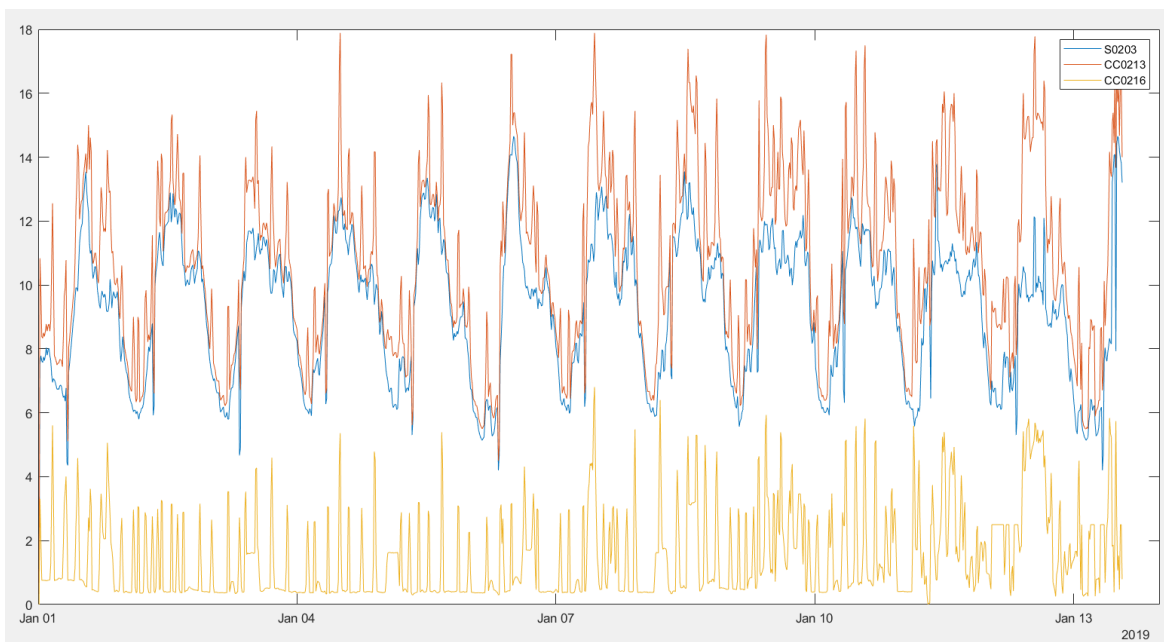


Figura 49. Reconstrucción de CC0213 y CC0216.

En la anterior figura se muestra la reconstrucción del sector S0203. Se observa una clara mejora respecto a la reconstrucción únicamente por modelo AR, dando una menor cantidad de sobrepicos y siendo capaz de reconstruir un alto porcentaje de los datos nulos.

6 Almacenamiento de datos

Una vez realizado el proceso de validación y reconstrucción de los datos, se procede a almacenarlos en una BBDD con el objetivo de conseguir mayor accesibilidad a estos y encontrar la información deseada más fácilmente.

En este apartado describirá el proceso realizado para implementar la BBDD y se presentará la conexión desde Matlab con la BBDD de Access. Para realizar la definición de la BBDD, se ha comenzado con el desarrollo del diagrama entidad-relación y posteriormente con el modelo relacional. Una vez creada la BBDD se procede a añadir la información de los sensores. En la Figura 51 se muestra de manera esquemática el proceso seguido.

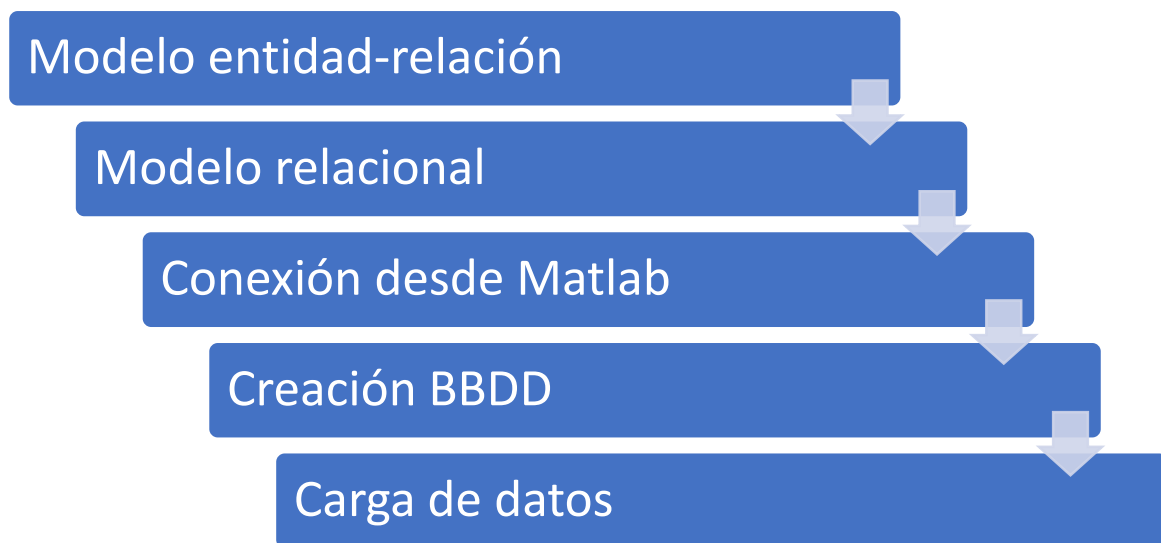


Figura 50. Proceso de creación BBDD.

Una vez creada la BBDD únicamente será necesario cargar los datos siempre que se indique.

6.1 Modelo entidad-relación

Para crear toda BBDD el primer paso es diseñar un modelo entidad-relación para tener una primera visión de qué información se quiere reflejar en la propia BBDD y de qué manera organizar los datos.

Los modelos entidad-relación tienen 3 componentes claves [15] a la hora de diseñarlos que se deben tener en cuenta:

- Las entidades: es el objeto del cual se quiere almacenar información, pueden tratarse de entidades débiles si dependen de otra entidad o por el contrario, de entidades fuertes. Serán representadas como rectángulos si se trata de una entidad fuerte, y como doble rectángulo si es débil.
- Los atributos: Son los parámetros o características de las entidades, en ellos se destacan la PK (primary key) el cuál es el atributo único que define la entidad, la FK (foreign key) atributo único de una entidad relacionada. También se puede catalogar los atributos según si es libre o debe tener una respuesta acotada entre unas opciones. Los atributos se representan mediante óvalos en el modelo.

- La cardinalidad o relación entre entidades: Define la participación que hay entre las entidades. Las relaciones son representadas como rombos.

En la BBDD implementada se han identificado 3 entidades diferenciadas:

- Pisos: se trata de entidad fuerte debido a que no necesita de más entidades para existir. Será la entidad en la que se vean reflejados los diferentes pisos que hay en la red de distribución. Contiene los siguientes atributos:
 - ID_pisos: El cuál representará el número de identificación del piso en cuestión.
 - Nombre: Atributo donde se detallará el nombre del piso, será la PK de la entidad.
- Sensor: Es una entidad débil, ya que cada sensor está localizado en un piso, por lo que necesita de la existencia de la entidad Pisos. En esta entidad se pretende definir los sensores que se encuentran en cada Piso así como detallar qué tipo de sensor es. Los atributos de la entidad serán los siguientes:
 - ID_sensor: Es la PK de la entidad que define el código identificador de cada sensor.
 - Tipo: Se definirá el tipo de sensor que es, por ejemplo, si es de nivel, o un caudalímetro.
 - Unidades: Se define en qué unidades vendrá representado el valor de cada sensor.
 - Estado: atributo con respuesta cerrada, ya que se definirá si está disponible o no.
- Datos: Se trata de una entidad débil, ya que depende de la entidad Sensores debido a que cada dato proviene de un sensor. En esta entidad se representarán la información de los datos registrados por los sensores, y posteriormente la información de la validación y reconstrucción. Los atributos que contiene esta entidad son los siguientes:
 - ID_medida: Se define un número identificador para cada dato. Actuará como PK.
 - Fecha: La fecha de cuando fue registrado el dato.
 - Valor: El valor del dato registrado:
 - Valor_final: El valor del dato una vez pasado el proceso de validación y reconstrucción.
 - Validado: Representará con un número del 0 al 6 el estado de la validación, por lo que se trata de respuesta cerrada.
 - Reconstruido: Representará con un 1 si se ha modificado el valor en el proceso de reconstrucción, o por el contrario, con un 0.

De esta manera, el modelo entidad-relación diseñado [17] para la BBDD de este proyecto es el siguiente:

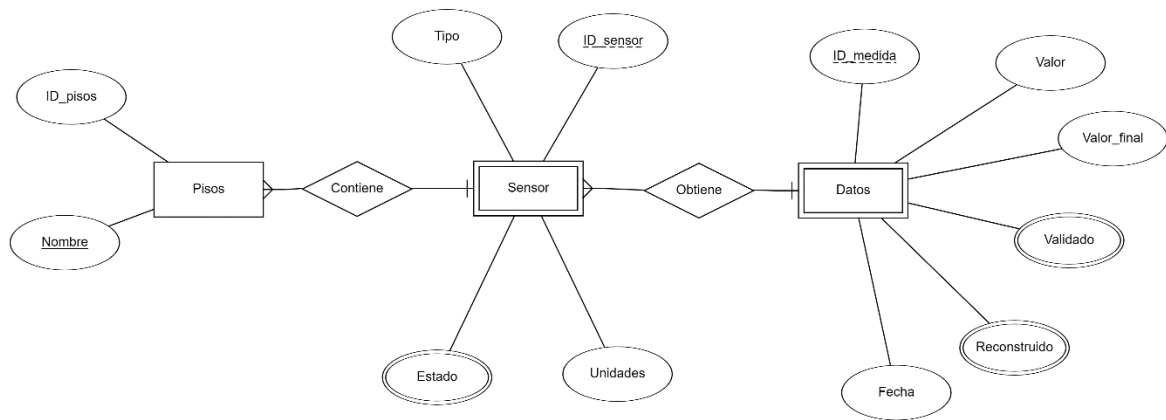


Figura 51. Diagrama de modelo entidad-relación

Siguiendo la figura anterior, se puede identificar la relación o cardinalidad entre cada una de las entidades:

- Relación entre Pisos y Sensor: Se trata de una cardinalidad de un piso a muchos sensores, ya que en un piso pueden encontrarse varios sensores, pero por el contrario, cada sensor pertenece a un piso.
- Relación entre Sensor y Datos: Se trata de la misma cardinalidad que en el anterior caso, cada sensor puede contener infinitos datos, pero cada dato únicamente puede pertenecer a un sensor en específico.

Una vez finalizado con el modelo entidad-relación, se procede a realizar el traspaso a modelo relacional, que será el paso previo a crear la BBDD desde Matlab, ya que definirá las tablas a realizar y su relación entre ellas.

6.2 Modelo relacional

El modelo relacional tiene como objetivo crear, a partir del modelo anterior de entidad-relación, las tablas que se usarán en la BBDD. Una vez conocido el modelo relacional, se podrá acceder a crear la BBDD desde Matlab para posteriormente cargarla.

La base principal para realizar el modelo relacional pasa por tratar las entidades como tablas, donde cada atributo será un parámetro que se deberá tener en cuenta para las tablas. Por otro lado, las relaciones vendrán dadas por las PK y FK de cada una de las tablas.

El modelo relacional es el siguiente:

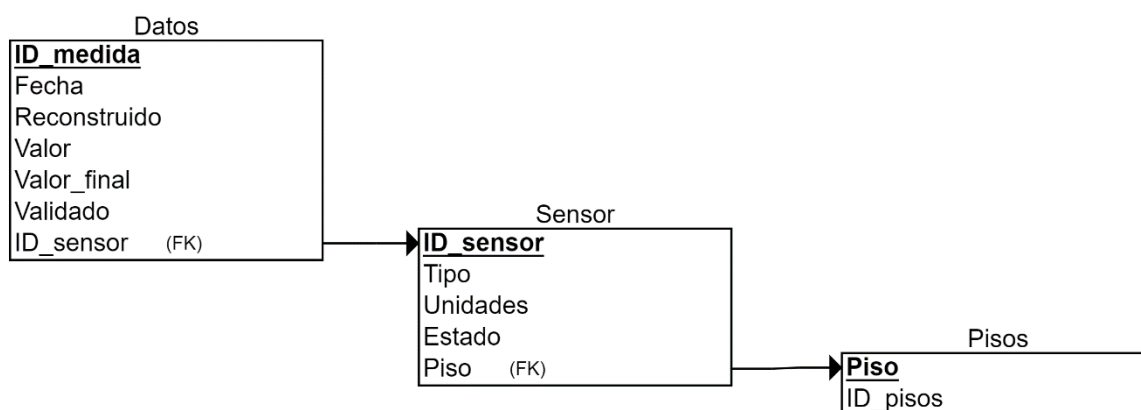


Figura 52. Diagrama de modelo relacional.

Al tratarse de la entidad fuerte del esquema de entidad-relación, todo se inicia con la tabla de Pisos, donde se indica que la columna Piso será PK, también tendrá otra columna de ID_pisos.

Posteriormente la tabla de Sensor donde se identifica ID_sensor como PK, también contra con las columnas Tipo, Unidades y Estado y por último tendrá como FK el identificador del piso (ID_pisos).

Por último, la tabla de datos que tendrá una estructura similar a la de Sensor, ya que tendrá el PK con ID_medida, las columnas Fecha, Reconstruido, Valor, Valor_final y Validado, y el identificador ID_sensor que actuará como FK.

6.3 Conexión desde Matlab

La BBDD será gestionada desde el programa Microsoft Access, pero los datos se introducirán desde Matlab. Para ello es necesario crear una conexión con la BBDD mediante la conexión ODBC.

ODBC (Open DataBase Connectivity) [18] es el tipo de conexión abierta que es utilizada para las BBDD. Con esta conexión se consigue que sea cuál sea la aplicación desde la cual se estén cargando los datos, se consiga acceder a los datos.

Para configurar la conexión, en un primer caso se ha de crear la BBDD en Access y guardarla sin llegar a hacer ninguna acción. Posteriormente, se debe acceder a “ODBC Data Sources” de Windows y en la pestaña de “DSN de usuario” crear una conexión de tipo “Microsoft Access Driver” y seleccionar la BBDD creada:

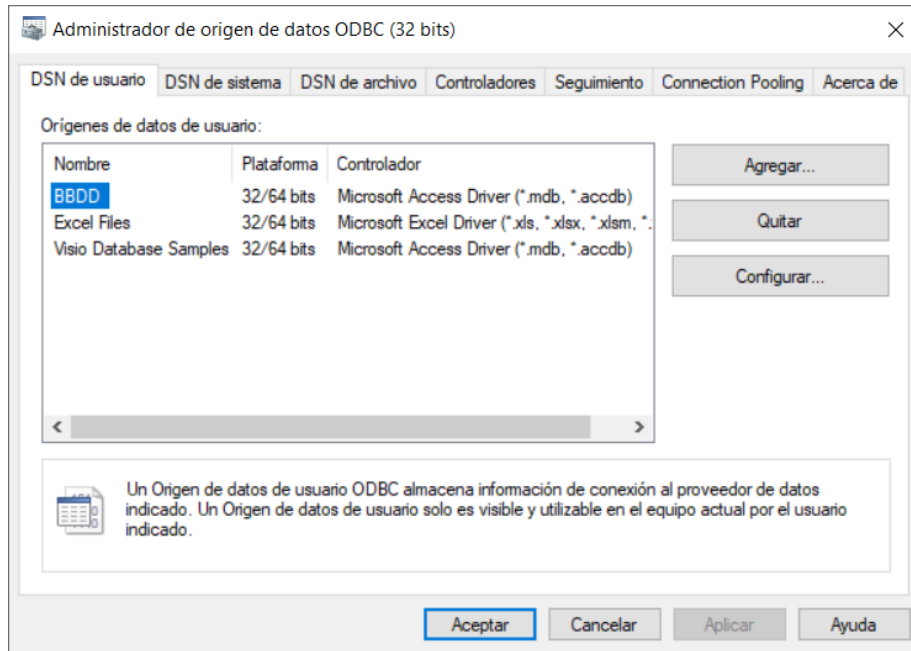


Figura 53. ODBC Data Source

En este caso, la base de datos se llamará “BBDD”.

Una vez configurada la conexión para Access, se deberá acceder desde Matlab. En Matlab es necesario tener instalada la “Database Toolbox” que es la que contiene los drivers de comunicación ODBC y JDBC (para conectarse a BBDD de Java).

Finalmente, con la función ‘database(datasource,username,password)’ se realiza la conexión desde Matlab.

6.4 Creación y carga de la BBDD

Finalmente, se accede a la creación de la BBDD desde Matlab partiendo del modelo relacional diseñado.

Las tablas son creadas mediante la función CREATE de SQL y se especifican los siguientes tipos de variable para cada una de las columnas:

- En la tabla de Pisos, ambos atributos son referidos como Varchar, de esta manera se pueden añadir tanto números como caracteres.
- En la tabla Sensor, también se definen todas las columnas como Varchar.
- En la tabla Datos, se encuentran diferentes tipos:
 - ID_medida, Validado y Reconstruido son definidos como reales, ya que todos ellos únicamente se identificarán números.
 - Fecha como datetime, ya que especificará año, mes, día, hora, minuto y segundo del dato.
 - ID_sensor como varchar, debido a que debe estar definido igual que en la tabla Sensor.

- Valor y Valor_final como float, ya que se introducirán valores con decimales.

Tras crear la BBDD desde Matlab, desde Access se puede confirmar que se ha realizado correctamente las relaciones entre sí:

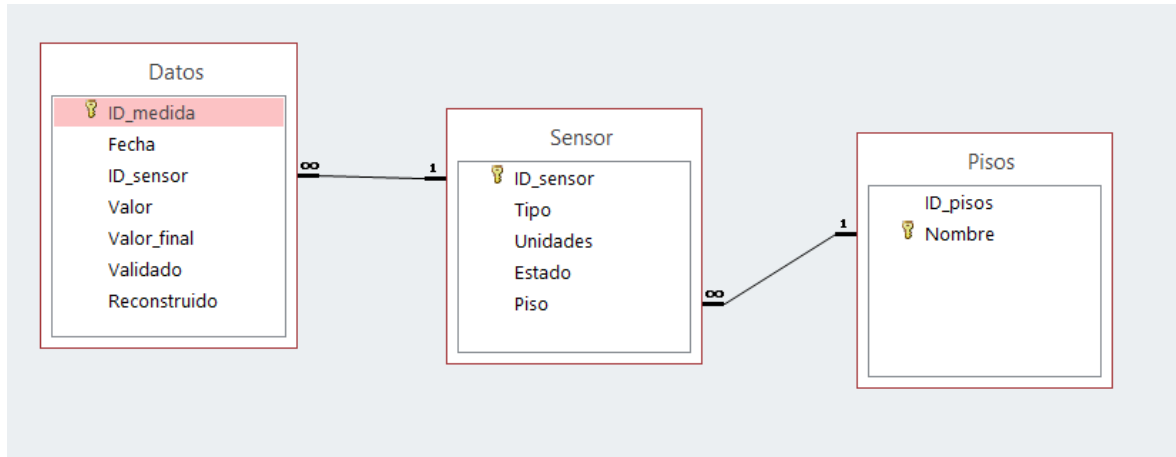


Figura 54. Relaciones de las tablas en Access

Un detalle a tener en cuenta es que las columnas de las tablas, pueden tener un nombre diferente a su relación en otra tabla, por ejemplo en la tabla sensor, se encuentra la columna Piso que actúa como FK ya que hace referencia a la columna Nombre (que actúa como PK) de la tabla Pisos.

Una vez creada las tablas ya se pueden cargar los datos en cada una de ellas. Para este proyecto hay una clara diferencia entre las tablas, ya que sobre las tablas de Pisos y Sensor únicamente habrá que cargar una vez los datos por lo que serán tablas estáticas, mientras que en la de Datos se irán añadiendo datos en cuanto acabe el proceso de validación, por lo que se tratará de una tabla dinámica.

En cualquier caso, se ha decidido añadir los datos a las tablas mediante la función INSERT de SQL. Otra opción para sobrescribir datos sería realizar un UPDATE, pero en este caso no se quiere cambiar el dato ya cargado, sino añadir otro nuevo.

A continuación se muestra las tablas cargadas con los datos:

- Tabla de Pisos:

ID_pisos	Nombre
2	Sulleva
3	Can Colomer
4	Zona Alta
5	Can Boada
6	Can Poal
7	Can Parellada Residencial
8	Can Parellada Industrial
9	Cam Palet de Vista Alegre

Figura 55. Tabla de Pisos.

- Tabla de Sensor:

	ID_sensor	Tipo	Unidades	Estado	Piso
+	CC0201	Caudal	l/s	Disponible	Sulleva
+	CC0202	Caudal	l/s	Disponible	Sulleva
+	CC0210	Caudal	l/s	Disponible	Sulleva
+	CC0211	Caudal	l/s	Disponible	Sulleva
+	CC0213	Caudal	l/s	Disponible	Sulleva
+	CC0214	Caudal	l/s	No Disponible	Sulleva
+	CC0216	Caudal	l/s	Disponible	Sulleva
+	SU_NV1	Nivel	m	Disponible	Sulleva
+	SU_NV2	Nivel	m	Disponible	Sulleva

Figura 56. Tabla de Sensor.

- Tabla de Datos:

ID_medida	Fecha	ID_sensor	Valor	Valor_final	Validado	Reconstruido
1670	01/01/2019 1:00:00	CC0202		16,7721	0	1
480	01/01/2019 1:00:00	SU_NV2	2,38956	2,38956	1	0
242	01/01/2019 1:00:00	SU_NV1	2,36924	2,36924	1	0
1432	01/01/2019 1:00:00	SU_NV2	2,38956	2,38956	1	0
1194	01/01/2019 1:00:00	SU_NV1	2,36924	2,36924	1	0
4	01/01/2019 1:00:00	CC0201	39,8437	39,8437	1	0
956	01/01/2019 1:00:00	CC0201	39,8437	39,8437	1	0
957	01/01/2019 1:15:00	CC0201	39,7801	39,7801	1	0
5	01/01/2019 1:15:00	CC0201	39,7801	39,7801	1	0
1433	01/01/2019 1:15:00	SU_NV2	2,43397	2,43397	1	0
243	01/01/2019 1:15:00	SU_NV1	2,4173	2,4173	1	0
719	01/01/2019 1:15:00	CC0202		16,7085	0	1
481	01/01/2019 1:15:00	SU_NV2	2,43397	2,43397	1	0
1671	01/01/2019 1:15:00	CC0202		16,7085	0	1
1195	01/01/2019 1:15:00	SU_NV1	2,4173	2,4173	1	0
244	01/01/2019 1:30:00	SU_NV1	2,46537	2,46537	1	0
958	01/01/2019 1:30:00	CC0201	39,7902	39,7902	1	0
1434	01/01/2019 1:30:00	SU_NV2	2,48203	2,48203	1	0
720	01/01/2019 1:30:00	CC0202		16,7186	0	1
1196	01/01/2019 1:30:00	SU_NV1	2,46537	2,46537	1	0
1672	01/01/2019 1:30:00	CC0202		16,7186	0	1
6	01/01/2019 1:30:00	CC0201	39,7902	39,7902	1	0
482	01/01/2019 1:30:00	SU_NV2	2,48203	2,48203	1	0
721	01/01/2019 1:45:00	CC0202		19,3857	0	1
1673	01/01/2019 1:45:00	CC0202		19,3857	0	1
1197	01/01/2019 1:45:00	SU_NV1	2,50794	2,50794	1	0
7	01/01/2019 1:45:00	CC0201	39,8206	39,8206	1	0
1435	01/01/2019 1:45:00	SU_NV2	2,52827	2,52827	1	0
483	01/01/2019 1:45:00	SU_NV2	2,52827	2,52827	1	0
959	01/01/2019 1:45:00	CC0201	39,8206	39,8206	1	0

Figura 57. Tabla de Datos.

Para no sobrescribir sobre los mismos datos, antes de volver a cargar registros en la tabla Datos, desde Matlab se ejecutará una consulta mediante la función SELECT que devuelva el último valor registrado en el campo ID_medida, así empezará a registrar desde ese último valor y no habrá duplicidad de valores.

Una vez cargados los datos se pueden realizar consultas desde Access, como por ejemplo revisar los datos de una hora en específico o incluso de un sensor en concreto. A continuación, se muestran los valores de una consulta realizada con el objetivo de mostrar únicamente los valores del caudalímetro CC0201 ordenando mediante fecha ascendentemente:

ID_medida	Fecha	ID_sensor	Valor	Valor_final	Validado	Reconstruido
1	01/01/2019 0:15:00	CC0201	23,397	23,397	1	0
2	01/01/2019 0:30:00	CC0201	39,7953	39,7953	1	0
3	01/01/2019 0:45:00	CC0201	39,8105	39,8105	1	0
4	01/01/2019 1:00:00	CC0201	39,8437	39,8437	1	0
5	01/01/2019 1:15:00	CC0201	39,7801	39,7801	1	0
6	01/01/2019 1:30:00	CC0201	39,7902	39,7902	1	0
7	01/01/2019 1:45:00	CC0201	39,8206	39,8206	1	0
8	01/01/2019 2:00:00	CC0201	39,8763	39,8763	1	0
9	01/01/2019 2:15:00	CC0201	39,856	39,856	1	0
10	01/01/2019 2:30:00	CC0201	39,8662	39,8662	1	0
11	01/01/2019 2:45:00	CC0201	39,8814	39,8814	1	0
12	01/01/2019 3:00:00	CC0201	39,9132	39,9132	1	0
13	01/01/2019 3:15:00	CC0201	39,856	39,856	1	0
14	01/01/2019 3:30:00	CC0201	39,9269	39,9269	1	0
15	01/01/2019 3:45:00	CC0201	39,9269	39,9269	1	0
16	01/01/2019 4:00:00	CC0201	39,9371	39,9371	1	0
17	01/01/2019 4:15:00	CC0201	39,9168	39,9168	1	0
18	01/01/2019 4:30:00	CC0201	39,9421	39,9421	1	0
19	01/01/2019 4:45:00	CC0201	39,9674	39,9674	1	0
20	01/01/2019 5:00:00	CC0201	39,9978	39,9978	1	0
21	01/01/2019 5:15:00	CC0201	40,0333	40,0333	1	0
22	01/01/2019 5:30:00	CC0201	40,0637	40,0637	1	0
23	01/01/2019 5:45:00	CC0201	40,0637	40,0637	1	0
24	01/01/2019 6:00:00	CC0201	40,0535	40,0535	1	0
25	01/01/2019 6:15:00	CC0201	40,0434	40,0434	1	0
26	01/01/2019 6:30:00	CC0201	40,013	40,013	1	0
27	01/01/2019 6:45:00	CC0201	39,9978	39,9978	1	0
28	01/01/2019 7:00:00	CC0201	40,0463	40,0463	1	0
29	01/01/2019 7:15:00	CC0201	20,042	20,042	1	0
30	01/01/2019 7:30:00	CC0201	0	0	1	0
31	01/01/2019 7:45:00	CC0201	0	0	1	0

Figura 58. Ejemplo de consulta en la BBDD.

7 Resultados

En este presente apartado se mostrarán los diferentes resultados de 5 casos a lo largo del año. Para realizar estas pruebas se han recogido 10000 datos recogidos por los sensores de diferentes épocas para contrastar el modelo.

En estos resultados se podrán comprobar el correcto tratamiento de los datos, ya que de los 10000 datos variados entre quinceminutales y biminutales, se mostrarán el equivalente quinceminutal. Por lo tanto, el número de datos disminuirá significativamente.

Por otro lado, se pretende mostrar el correcto funcionamiento de la validación mediante representaciones gráficas de Matlab. Se mostrará los datos registrados en color azul siguiendo el eje 'y' izquierdo de la gráfica (en las unidades del sensor) y en color naranja el estado de la validación con un valor del 0 al 6 tal y como se ha muestra en la tabla 10, siguiendo el eje 'y' derecho.

Tabla 10. Descripción de los resultados de la validación.

Nº	Descripción
0	Valor nulo o fuera de rango
1	Válido
2	Intensidad de la bomba y CC0201 no concuerda
3	Presión de la bomba y CC0201 no concuerda
4	Llenado externo
5	Fallo con valores previos
6	Fallo en el sector.

Y por último, para la comprobación de la reconstrucción se seguirá el mismo patrón que en la validación. En azul se mostrarán los valores finales tras el proceso, sea reconstruido o no, siguiendo el eje 'y' izquierdo, mientras que en naranja se mostrará el estado de la reconstrucción, siendo 0 si no ha pasado por el proceso y 1 si ha sido reconstruido.

Cabe destacar que se mostrarán los casos en los que se dé una circunstancia diferente a las demás. En el caso de ser correcto, no se mostrará más de una por sensor.

7.1 Caso 1: 01/01/2019 – 14/01/2019

El primer caso es de los primeros 14 días de enero del 2019. Obtenemos los siguientes resultados:

CC0201:

El primer fallo es detectado por discordancia con el sensor de intensidad de la bomba. El segundo fallo detecta que se ha llenado el depósito externamente por una vía no controlada. Además posteriormente, aparece un 0 indicando que no hay dato.

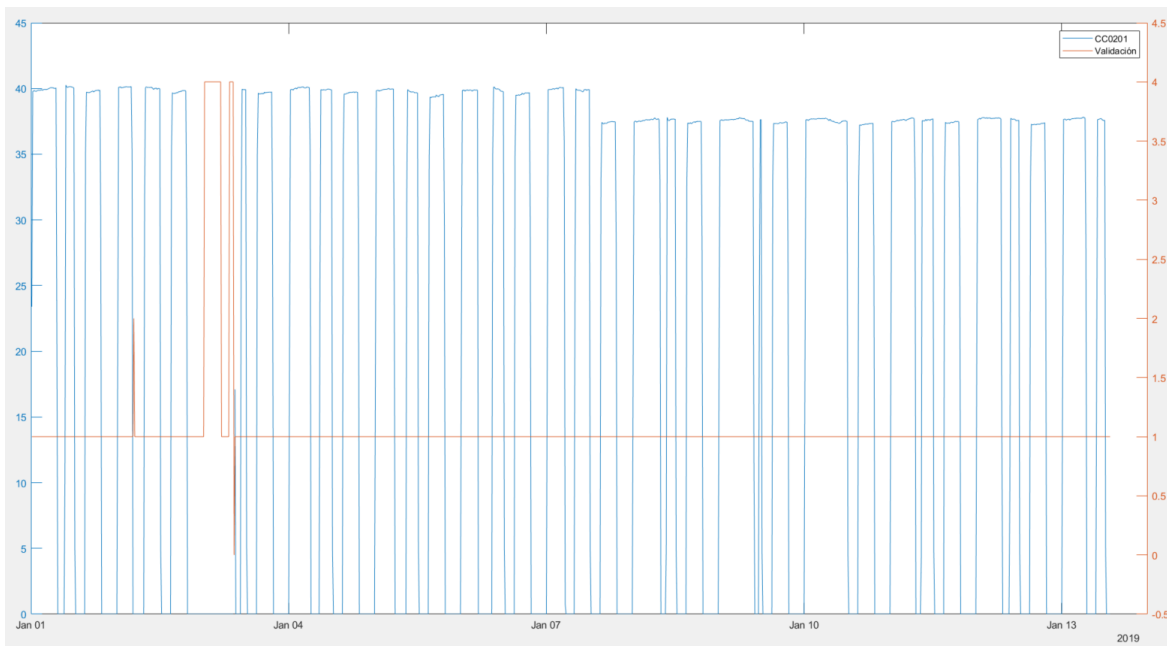


Figura 59. Caso 1 validación CC0201

Para los fallos de discordancia con el sensor de intensidad y el llenado externo, se observa como no se ha reconstruido el valor, por lo que es correcto. El único valor que reconstruido es el que no había dato.

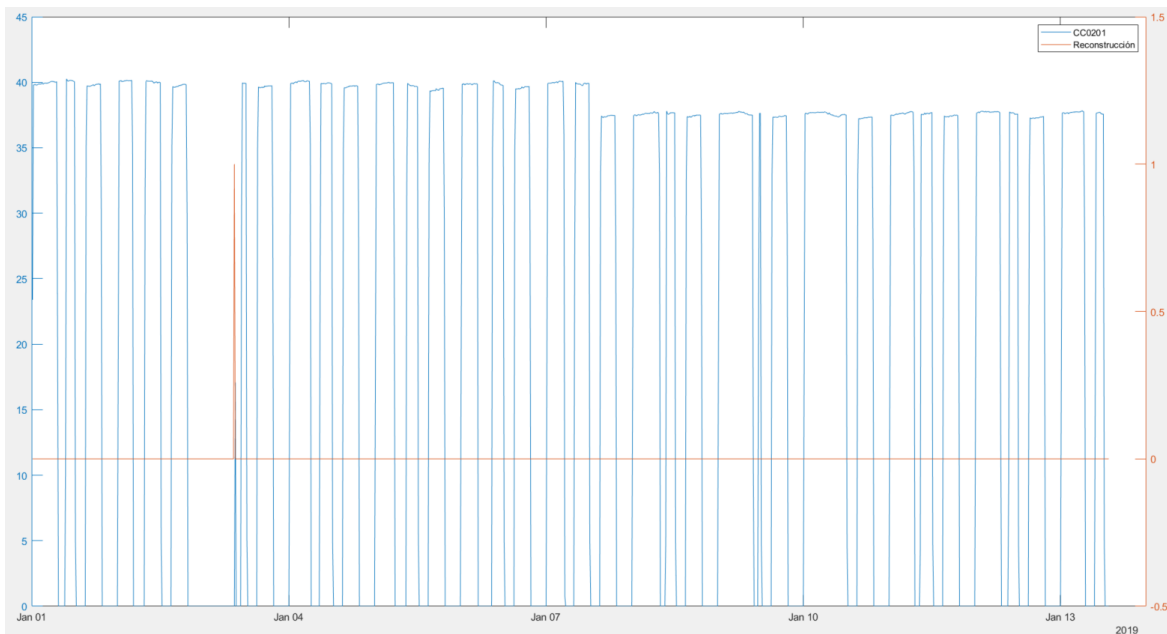


Figura 60. Caso 1 reconstrucción CC0201

Sensor de nivel:

Se detectan dos fallos consecutivos, el primer caso marca un 0 por lo que ese valor no existe, mientras que el segundo no se puede comprobar por la ausencia del primero.

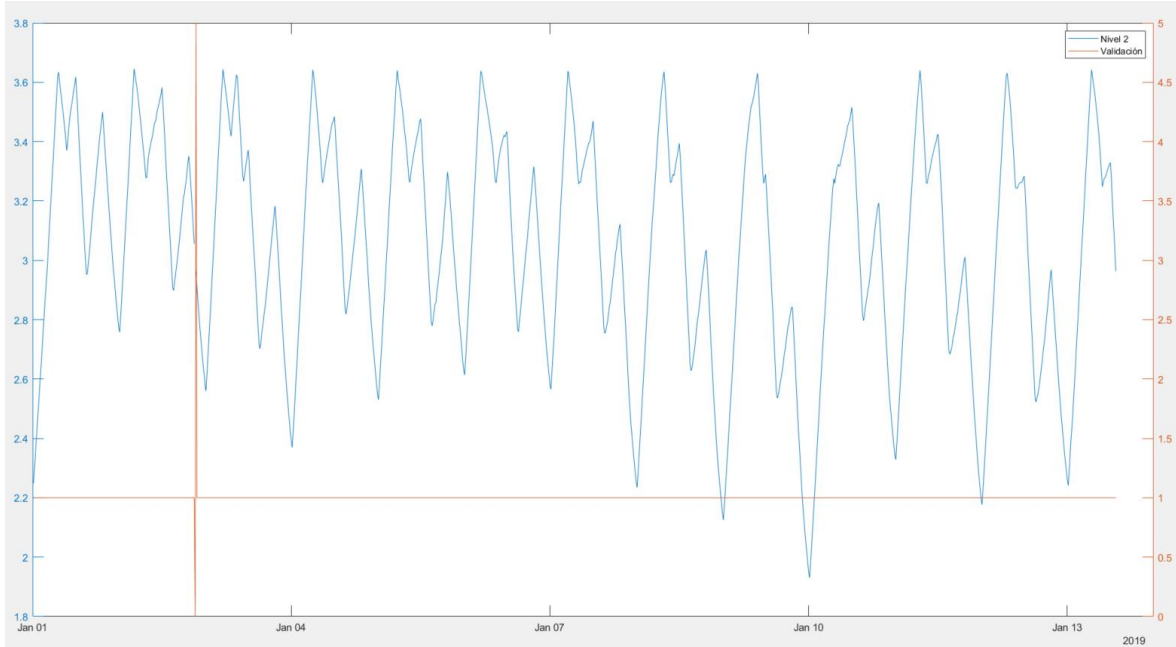


Figura 61. Caso 1 validación NV2

Se reconstruye el primer valor mediante interpolación lineal y se confirma el segundo valor correctamente.

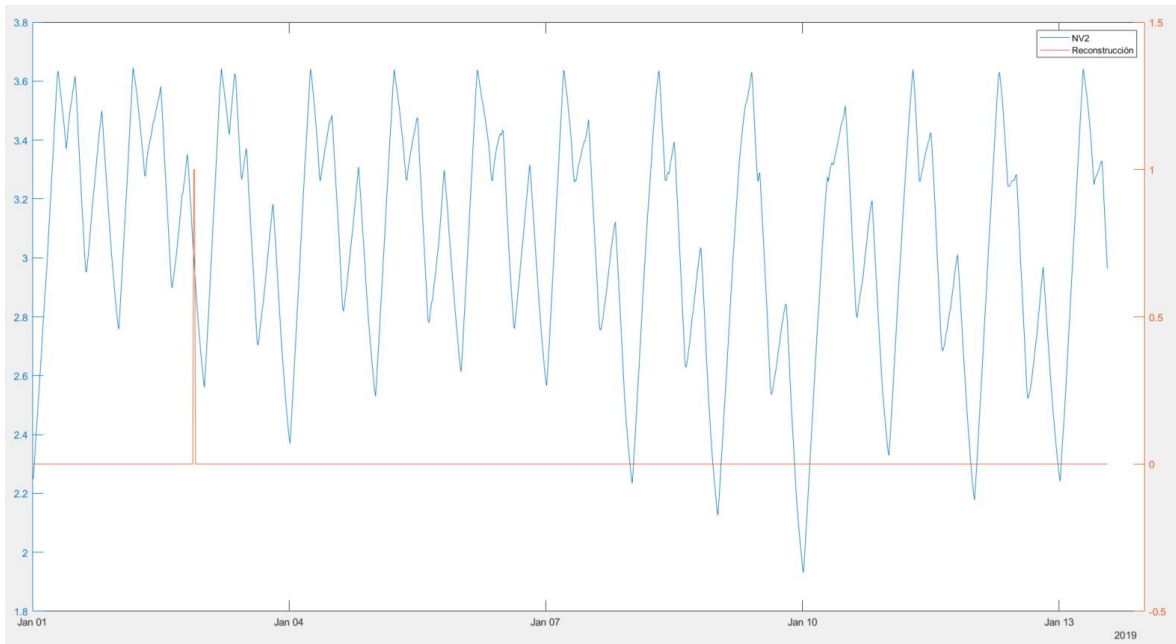


Figura 62. Caso 1 reconstrucción NV2

CC0210:

En este caso, todos los valores son válidos.

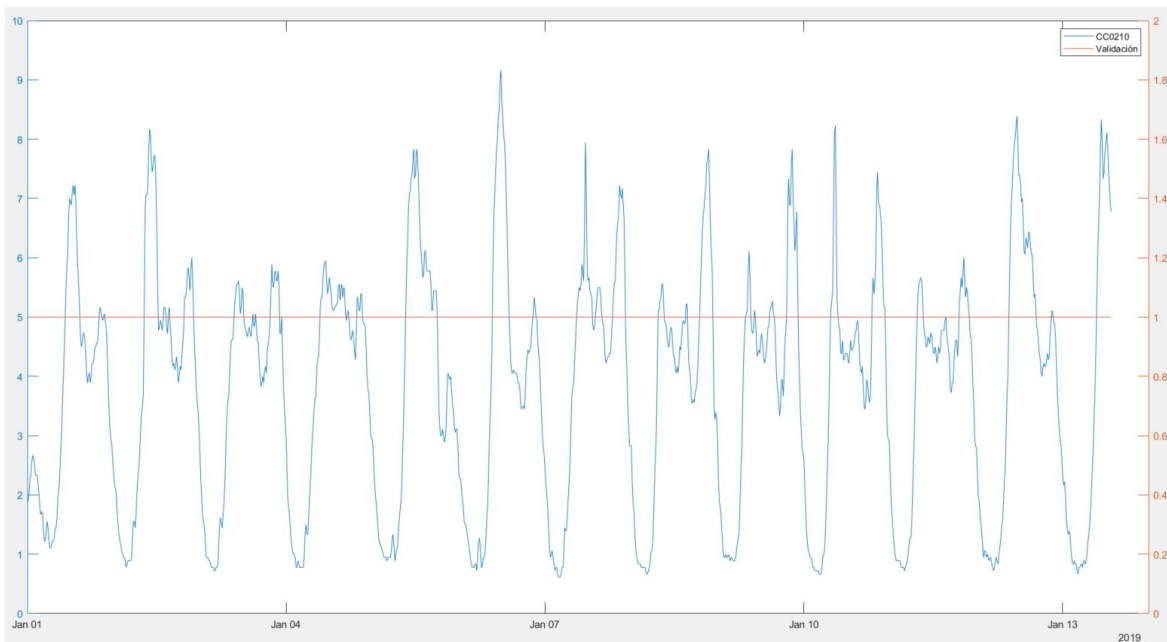


Figura 63. Caso 1 validación CC0210

Al ser todos válidos, no entra en el proceso de reconstrucción.

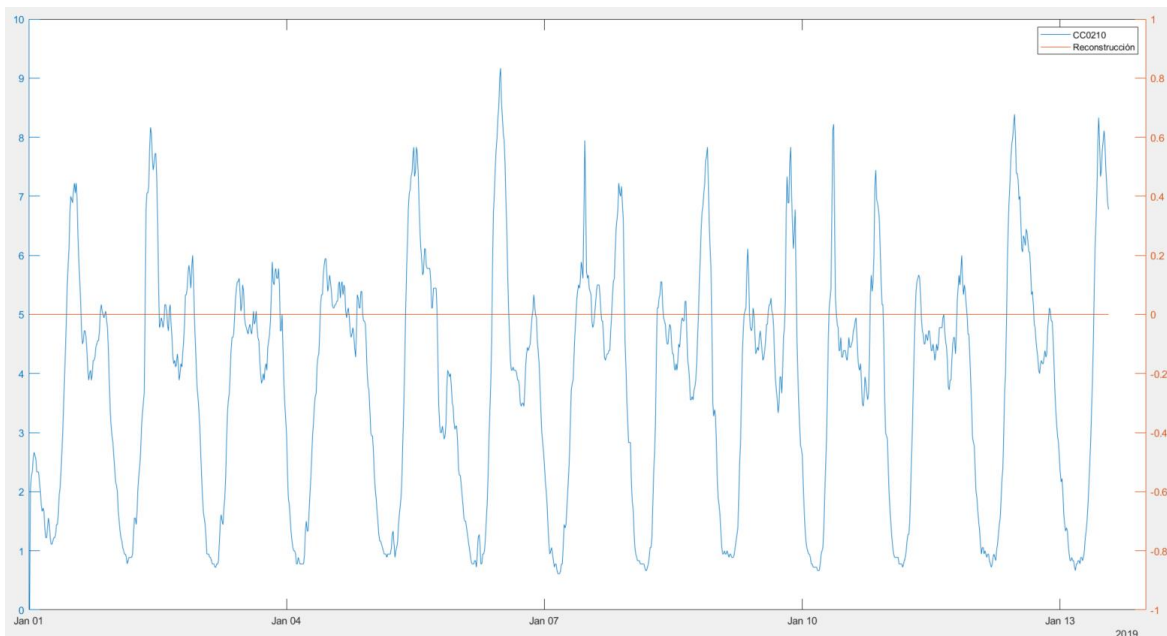


Figura 64. Caso 1 reconstrucción CC0210

CC0211:

Para este sensor, una vez pasado el 7 de enero, deja de haber valores, por lo que la validación pasa a 0.

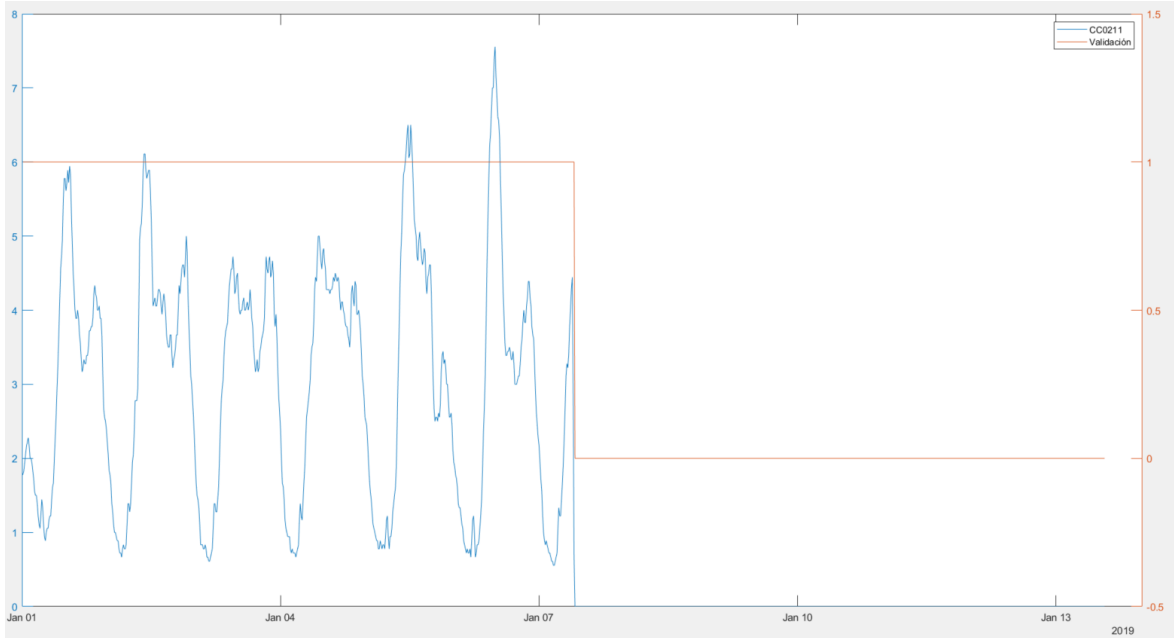


Figura 65. Caso 1 validación CC0211

No entra en la reconstrucción en la primera mitad, y en la segunda realiza la reconstrucción mediante el modelo autorregresivo y los datos históricos.

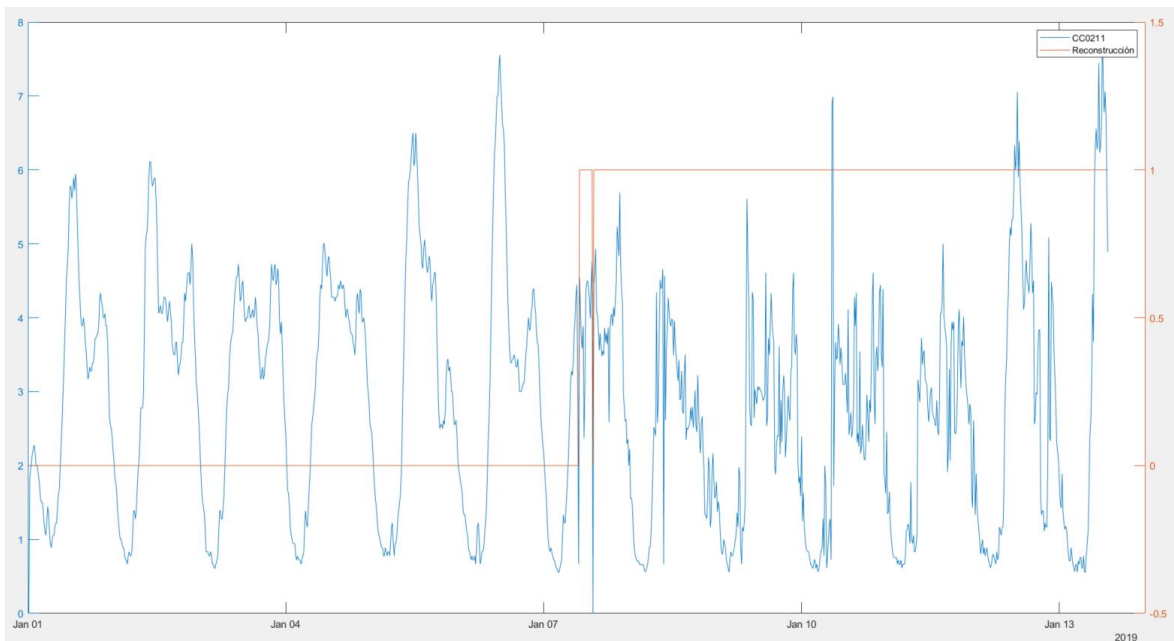


Figura 66. Caso 1 reconstrucción CC0211

CC0213:

Para este caudalímetro detecta varios puntos de validación 6, que significa que hay fallo en el sector S0203.

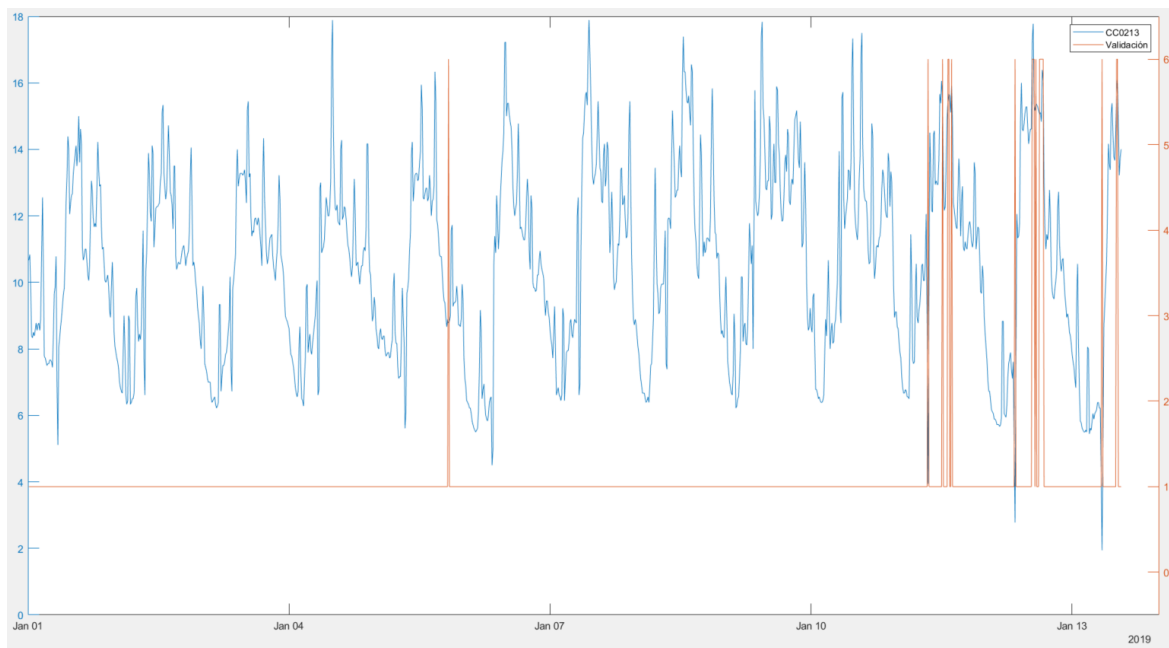


Figura 67. Caso 1 validación CC0213

Al realizar la reconstrucción del sector, indica que no se ha tenido que modificar el dato de este sensor, tal y como se muestra con una reconstrucción total en 0.

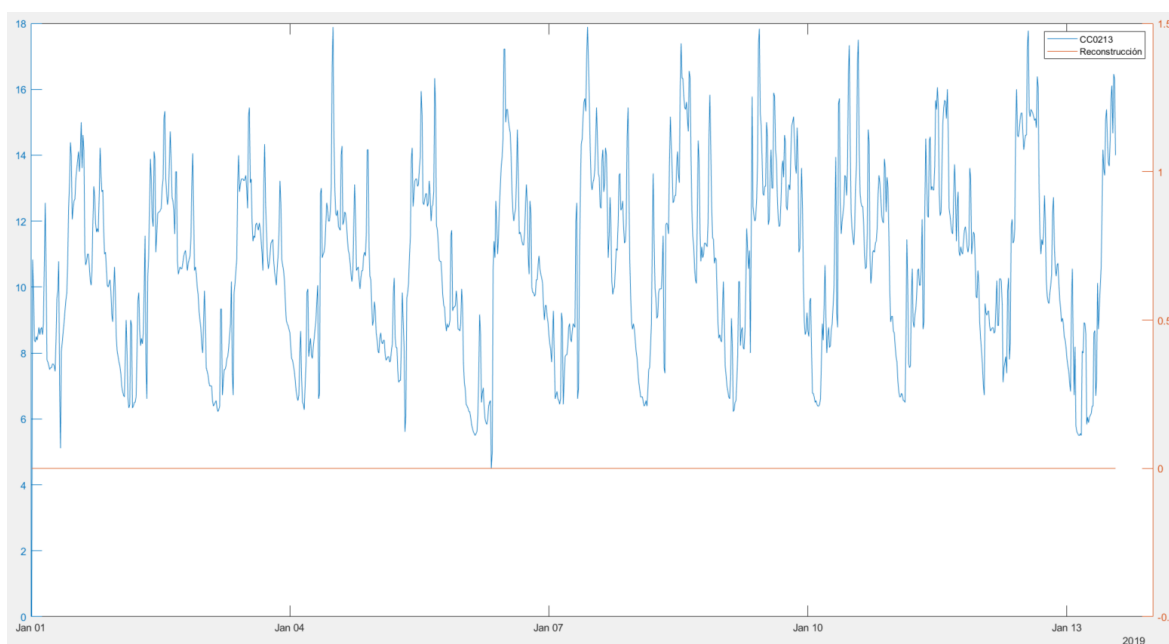


Figura 68. Caso 1 reconstrucción CC0213

CC0216:

En este caso detecta el mismo fallo que en el caudalímetro anterior y posteriormente se queda sin datos.

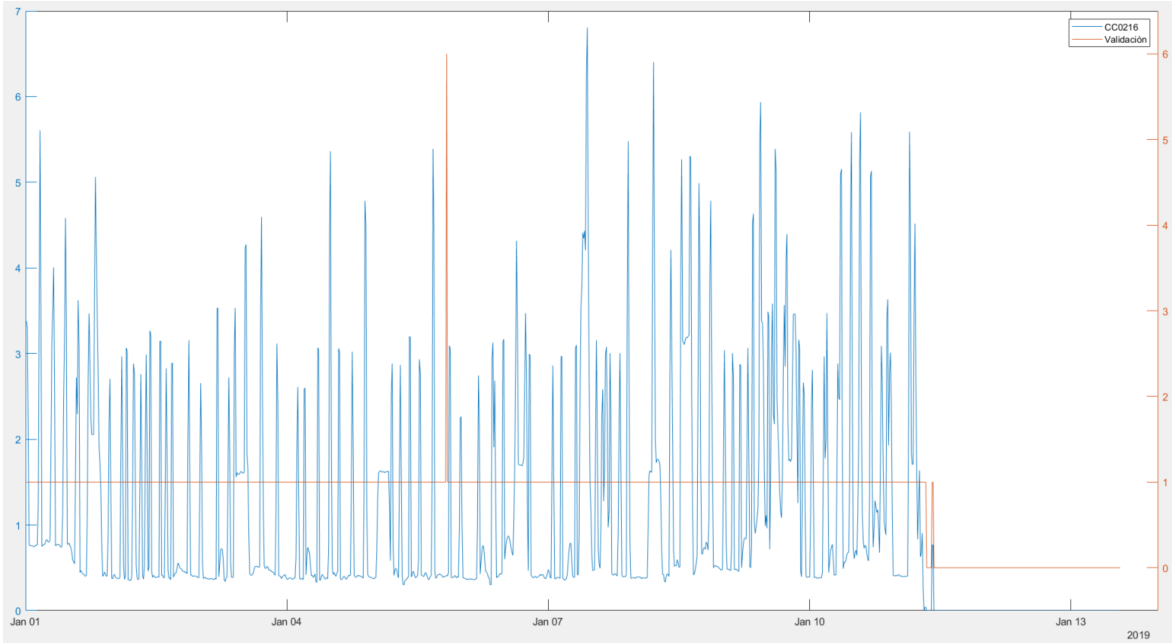


Figura 69. Caso 1 validación CC0216

En la parte final sí ha realizado la reconstrucción mediante el modelo AR y los datos históricos, mientras que el fallo puntual, no ha sido posible reconstruirlo. Esto puede ser debido a que ambos caudalímetros tengan coherencia con sus valores consecutivos.

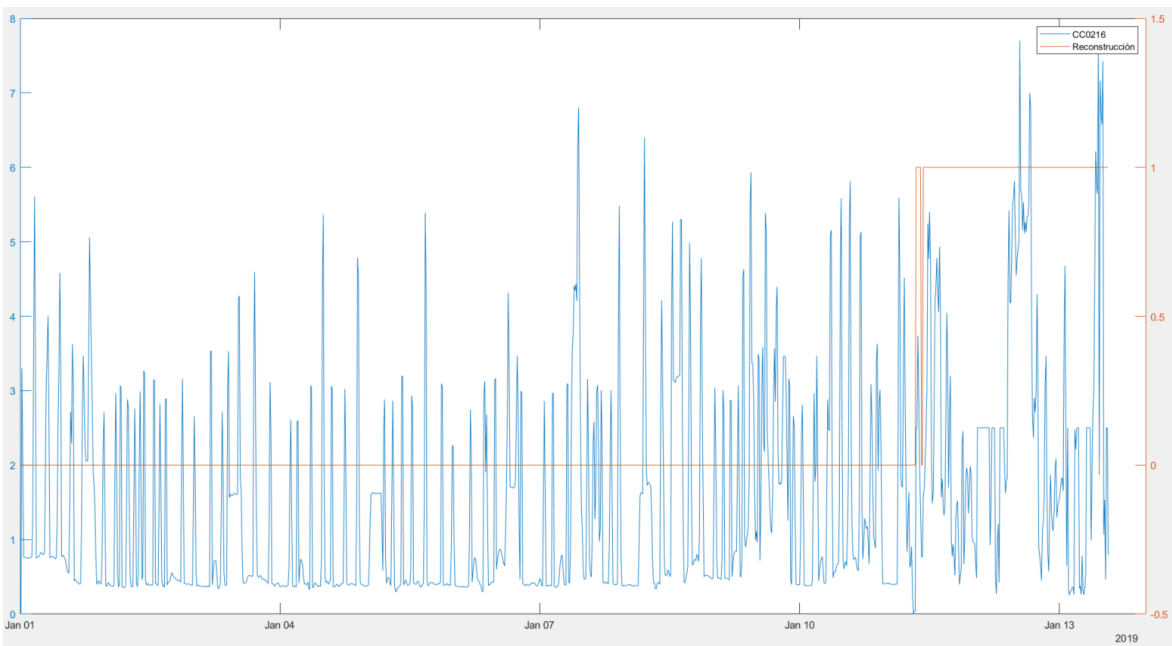


Figura 70. Caso 1 reconstrucción CC0216

7.2 Caso 2: 13/04/2019 – 26/04/2019

CC0201:

Se observan fallos por falta de datos (0) y por discordancia con el sensor de presión (3).

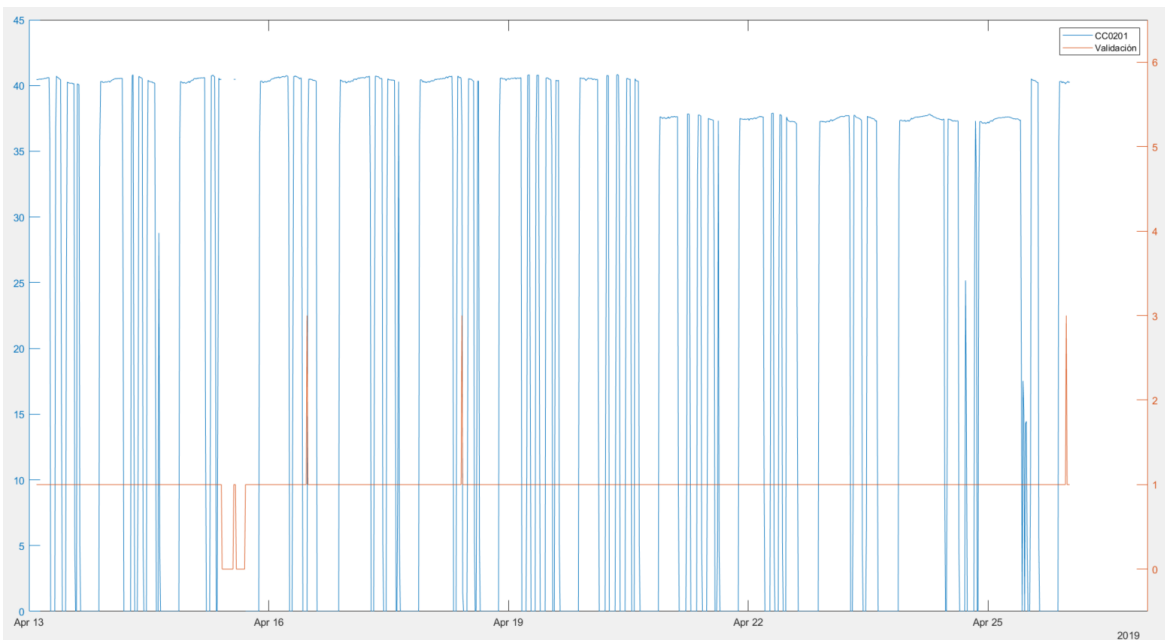


Figura 71. Caso 2 validación CC0201

En el primer fallo, se observa cómo se reconstruye a un valor indicando que la bomba está en funcionamiento. En las demás alarmas se confirma que el valor era correcto.

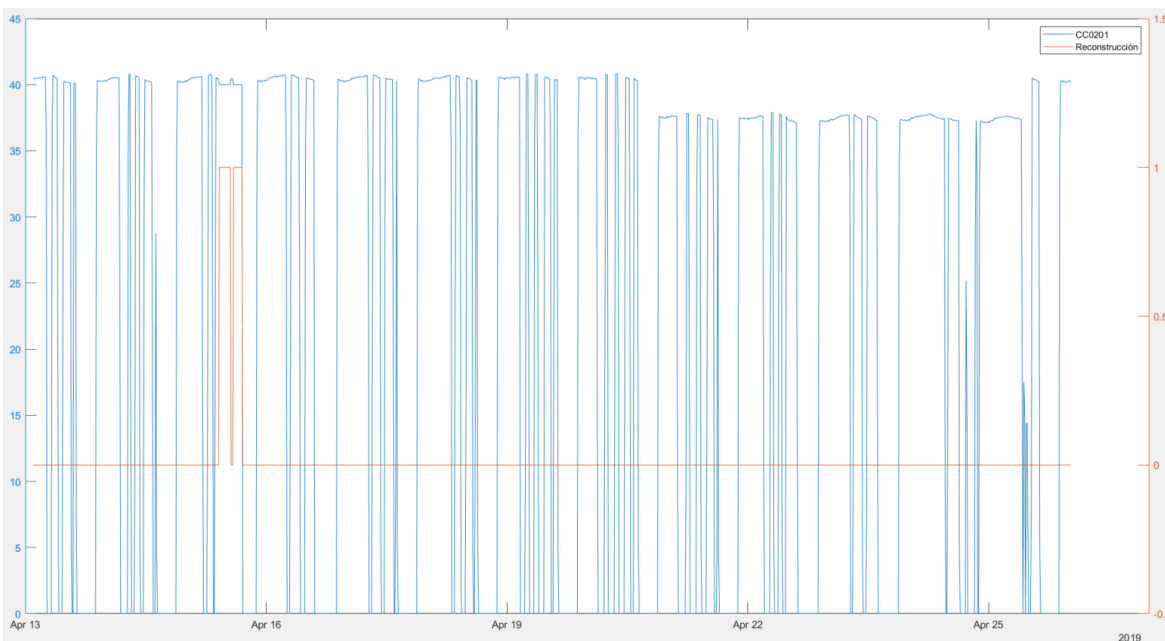


Figura 72. Caso 2 reconstrucción CC0201.

Sensor de nivel:

Se detecta una ausencia de datos del sensor de nivel y posteriormente un fallo único.

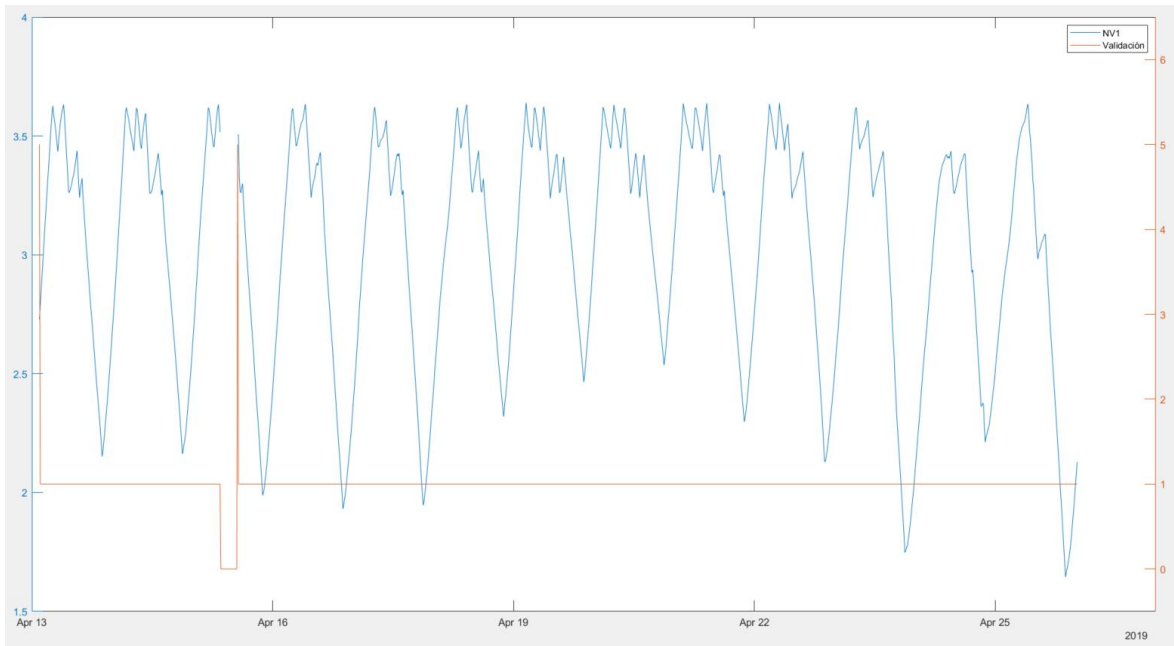


Figura 73. Caso 2 validación NV1.

La ausencia no ha podido reconstruirse debido a que el sensor de salida del caudalímetro CC0202 también estaba en fallo. Posteriormente el fallo único es reconstruido.

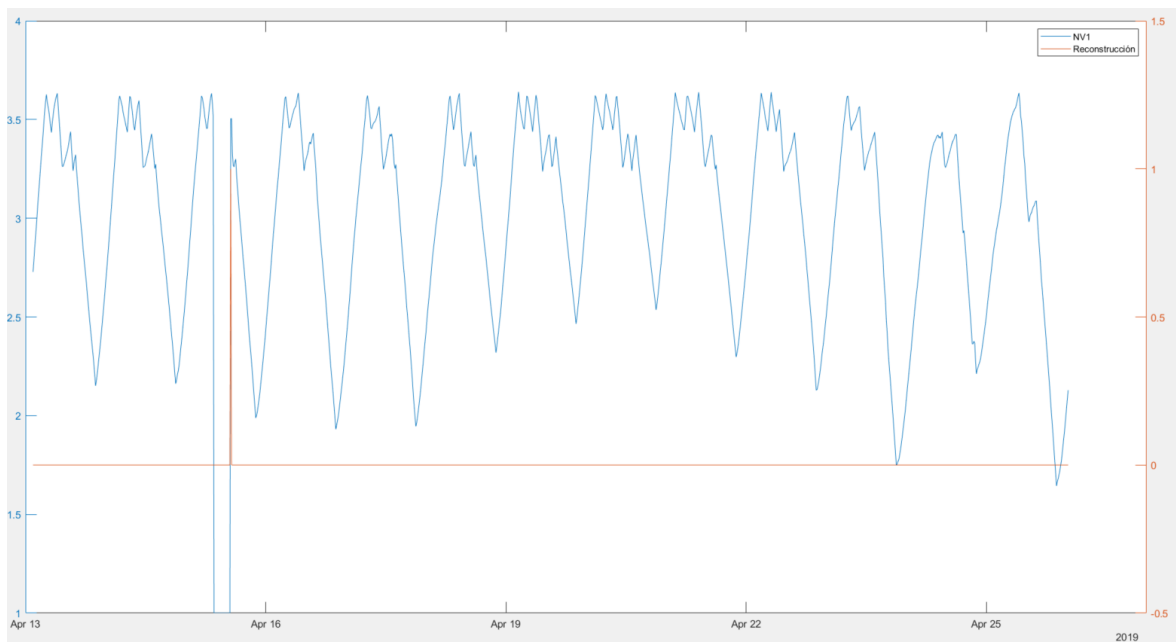


Figura 74. Caso 2 reconstrucción NV1

CC0202:

Hay una gran ausencia de los datos hasta llegar al 24 de abril, donde se valida y se encuentran ciertos errores.

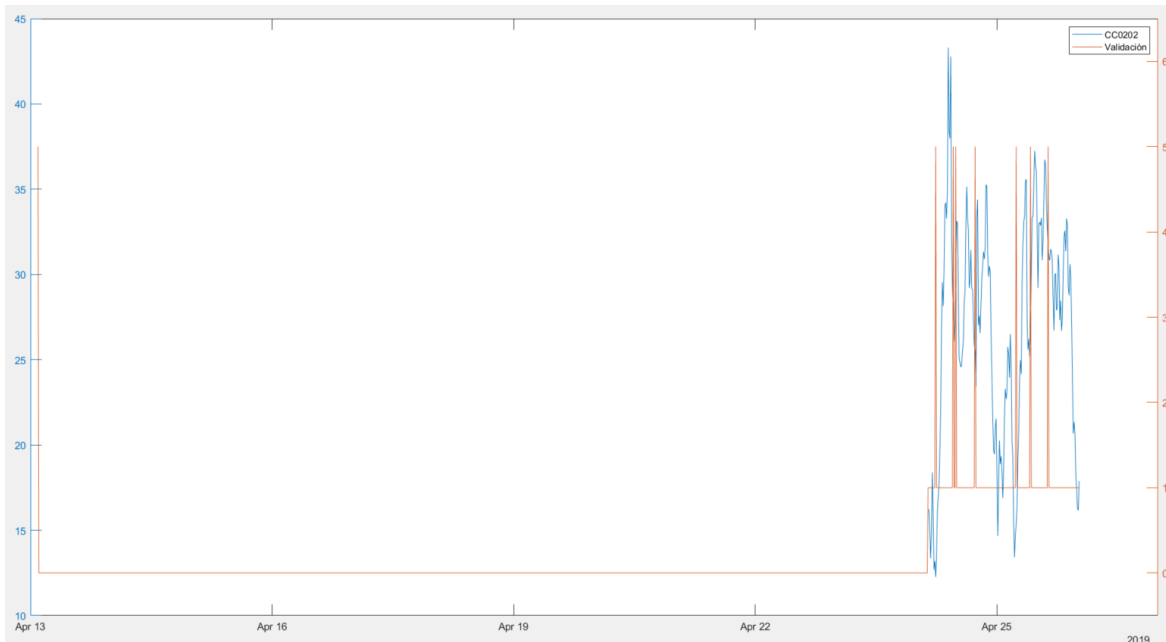


Figura 75. Caso 2 validación CC0202.

Se reconstruyen con los valores del CC0201 y los de nivel los valores ausentes y los erróneos. Los valores iniciales que no han sido posible reconstruir es debido a la ausencia de datos en el nivel también.

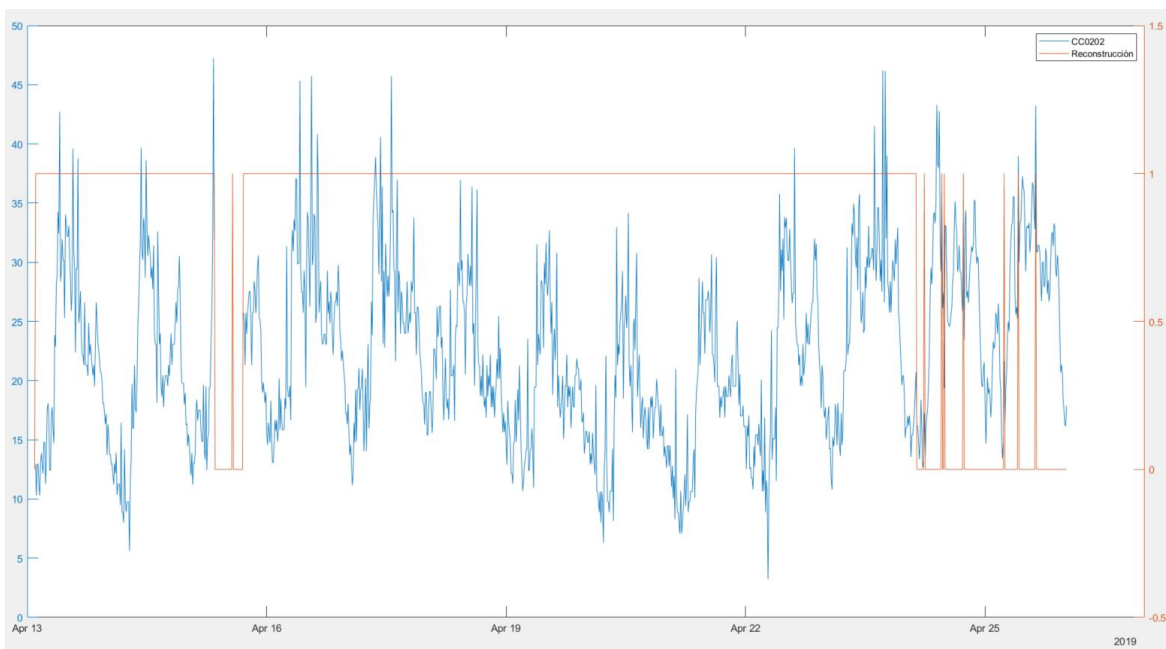


Figura 76. Caso 2 reconstrucción CC0202

CC0216:

Se detecta un fallo puntual en el sector S0203.

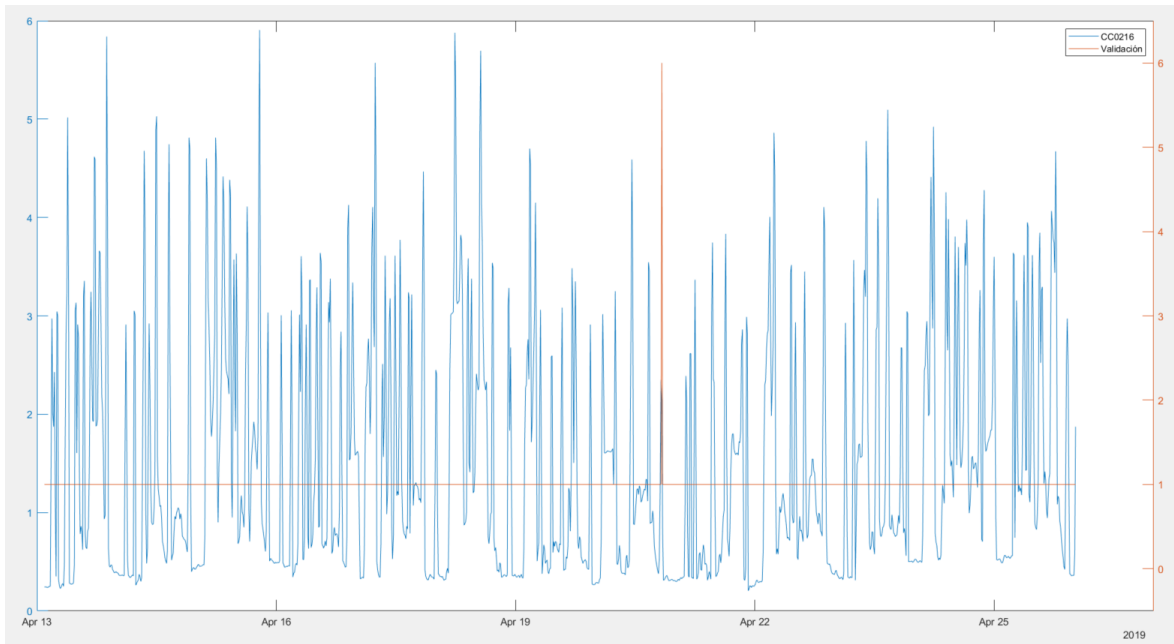


Figura 77. Caso 2 validación CC0216

Se reconstruye el valor usando interpolación lineal para entrar en los límites de consumo del sector.

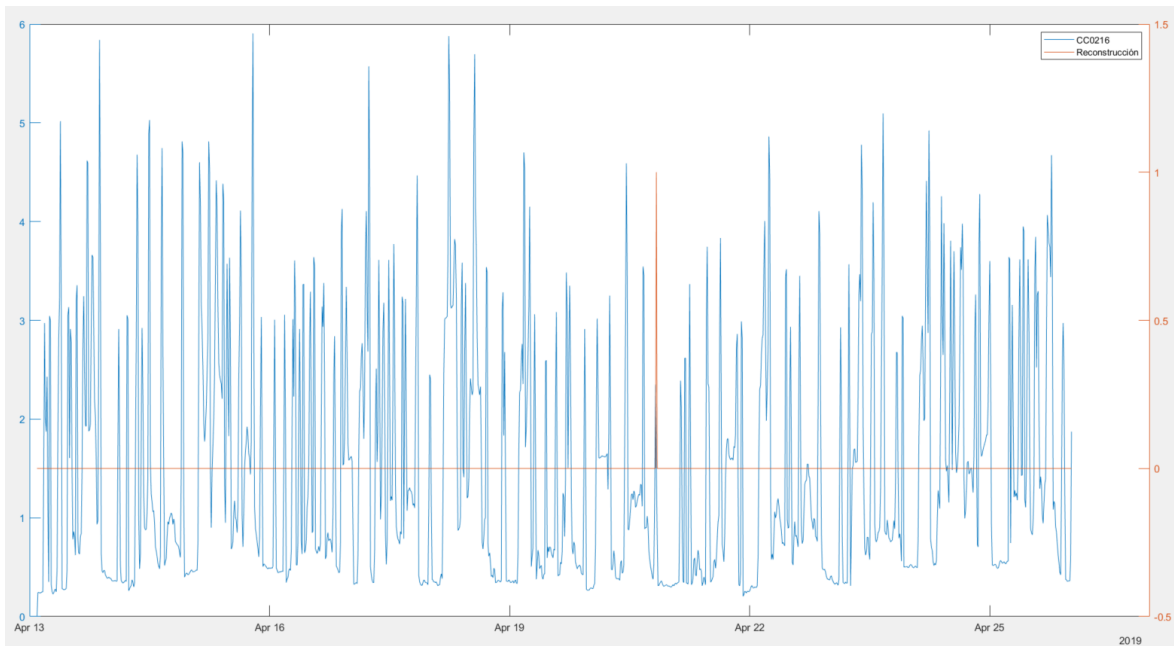


Figura 78. Caso 2 reconstrucción CC0216

7.3 Caso 3: 14/07/2019 – 27/07/2019

Sensor de nivel:

Todos los datos son correctos.

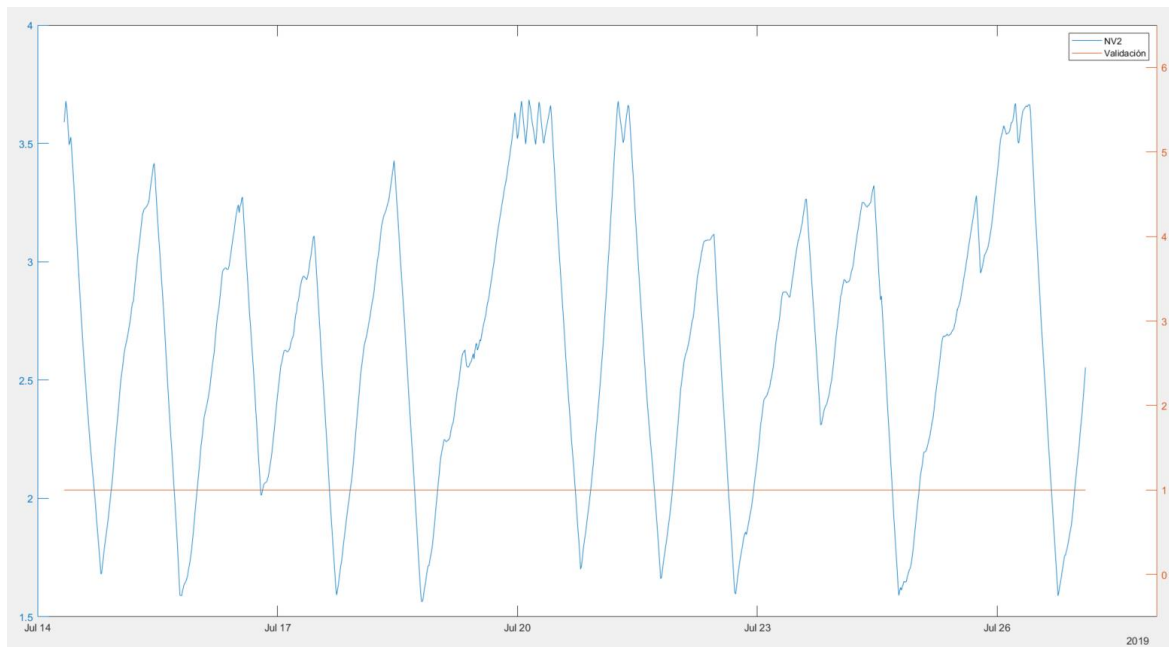


Figura 79. Caso 3 validación NV2

No es necesario entrar a reconstruir.

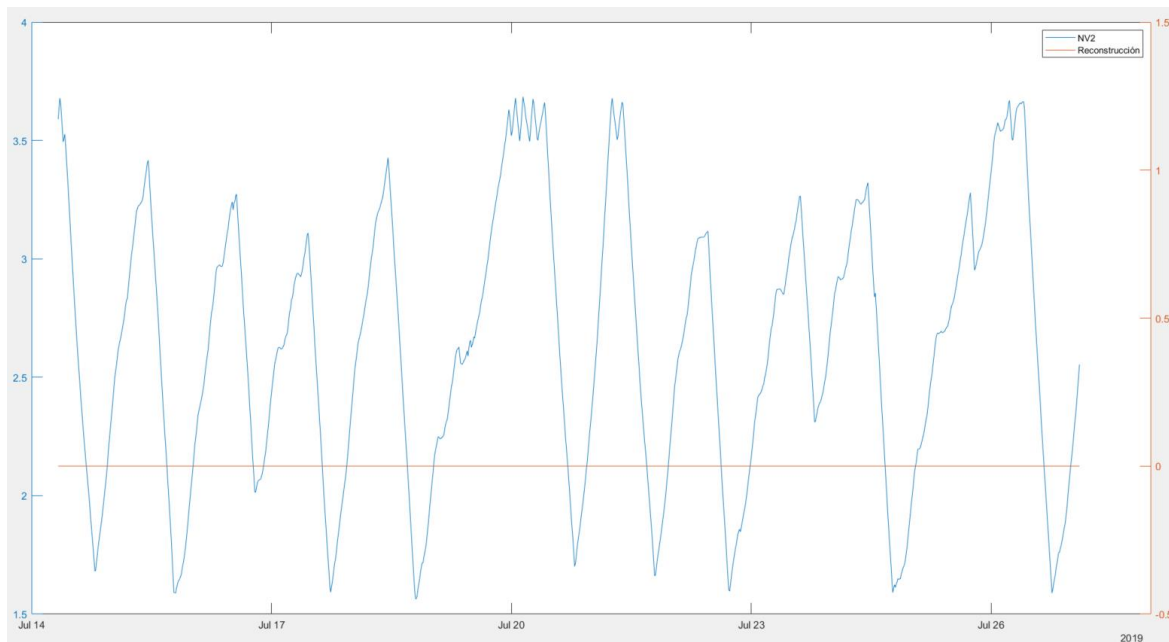


Figura 80. Caso 3 reconstrucción NV2

CC0202:

Se encuentra una gran ausencia de datos y posteriormente se valida mediante coherencia local.

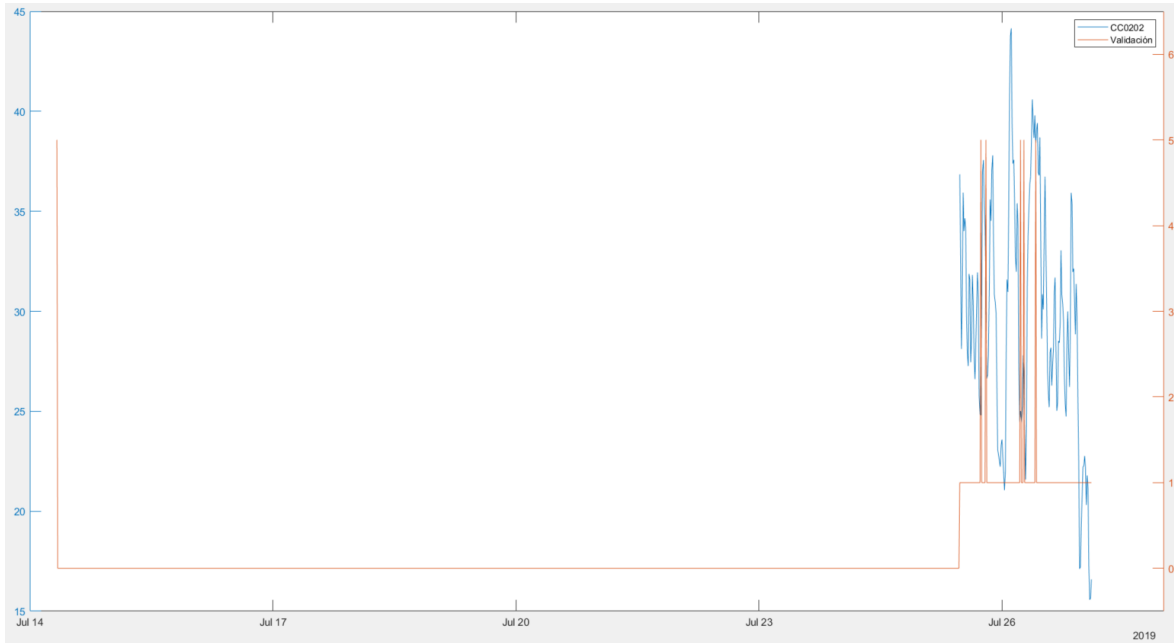


Figura 81. Caso 3 validación CC0202.

Se reconstruye mediante los valores del depósito, a excepción de algún dato que no ha sido posible.

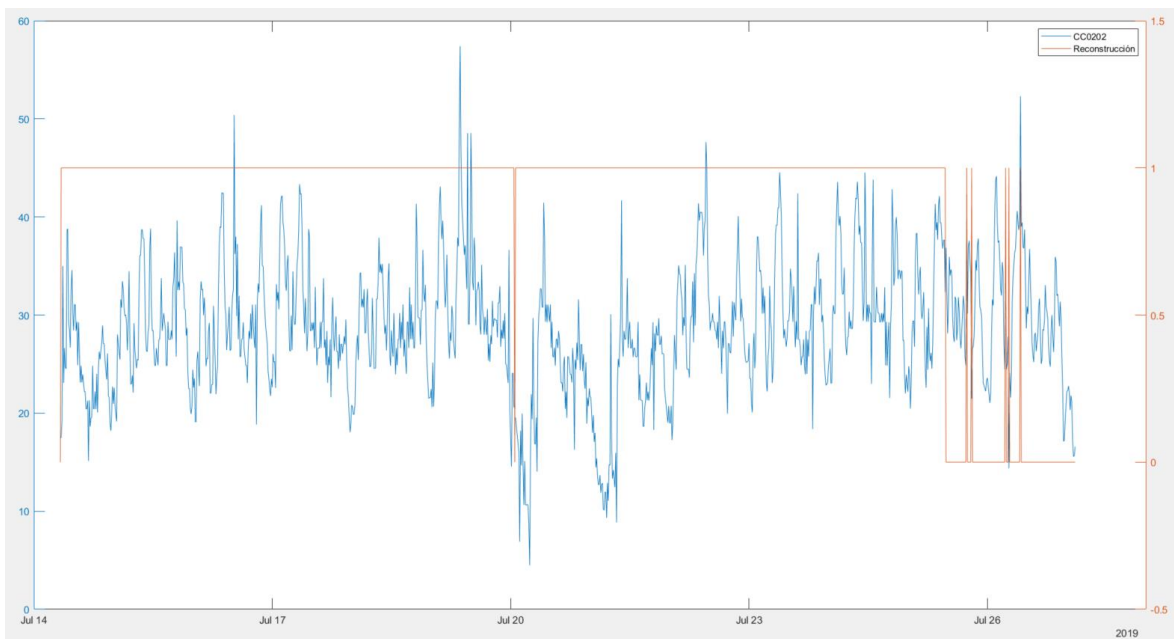


Figura 82. Caso 3 reconstrucción CC0202

CC0210:

Los datos son correctos.

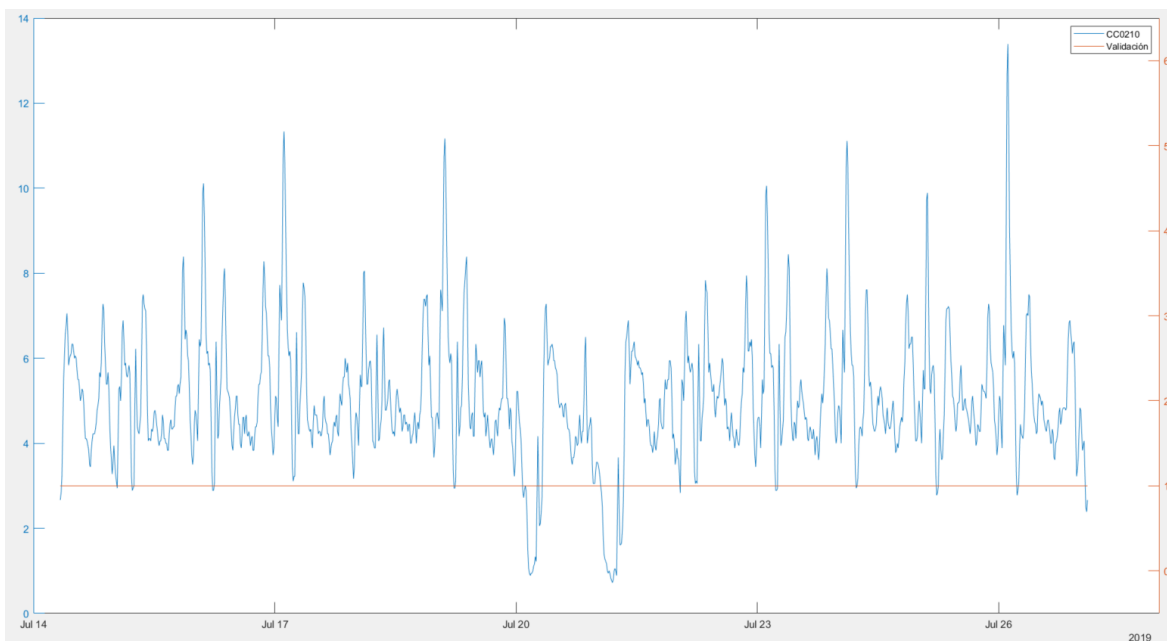


Figura 83. Caso 3 validación CC0210

No entra a reconstruir porque son correctos.

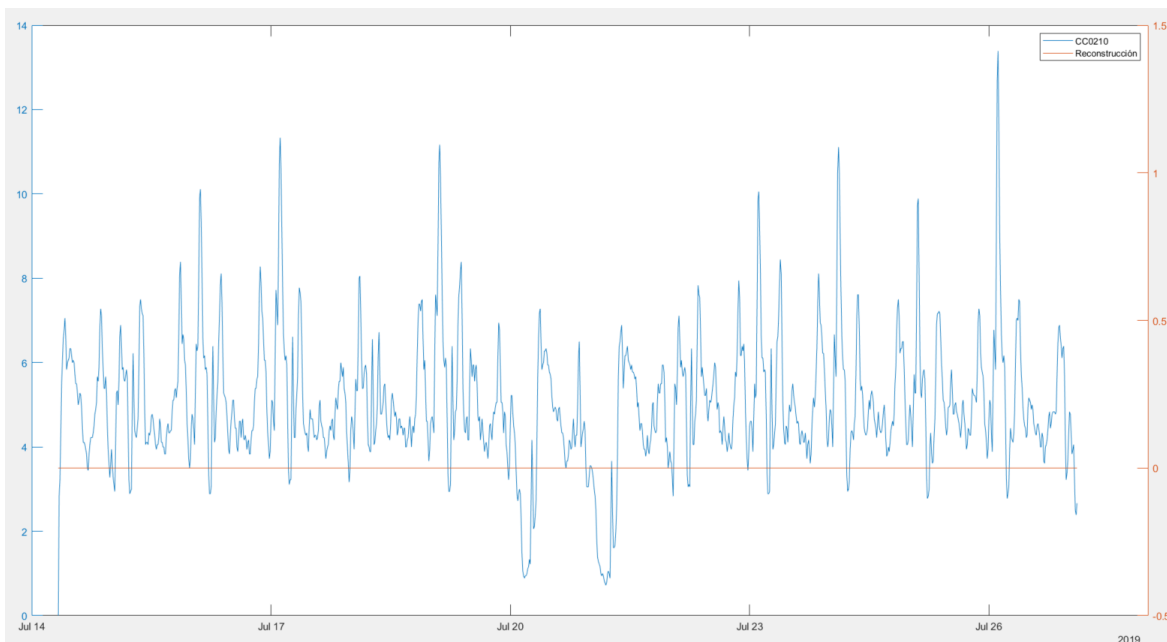


Figura 84. Caso 3 reconstrucción CC0210

CC0211:

Se validan correctamente los datos hasta llegar al día 25 de julio, donde empieza una total ausencia de datos.

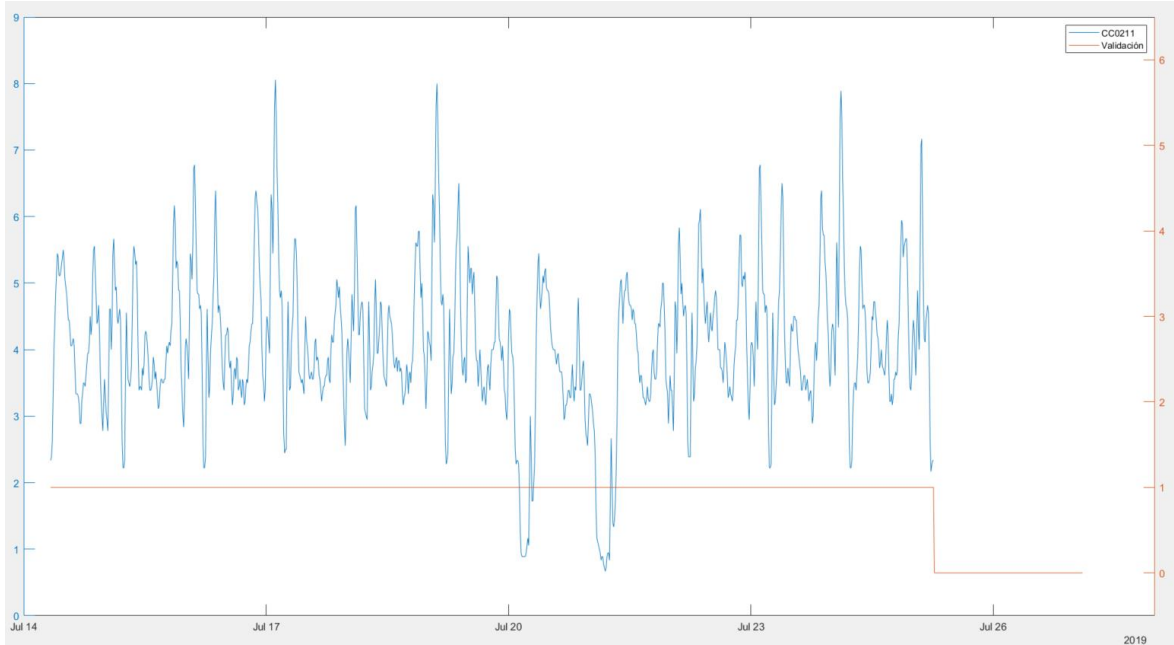


Figura 85. Caso 3 validación CC0211

Reconstruye mediante el modelo AR y los datos históricos el último tramo. En algún punto no ha sido posible realizar la reconstrucción ya que excedería los límites establecidos.

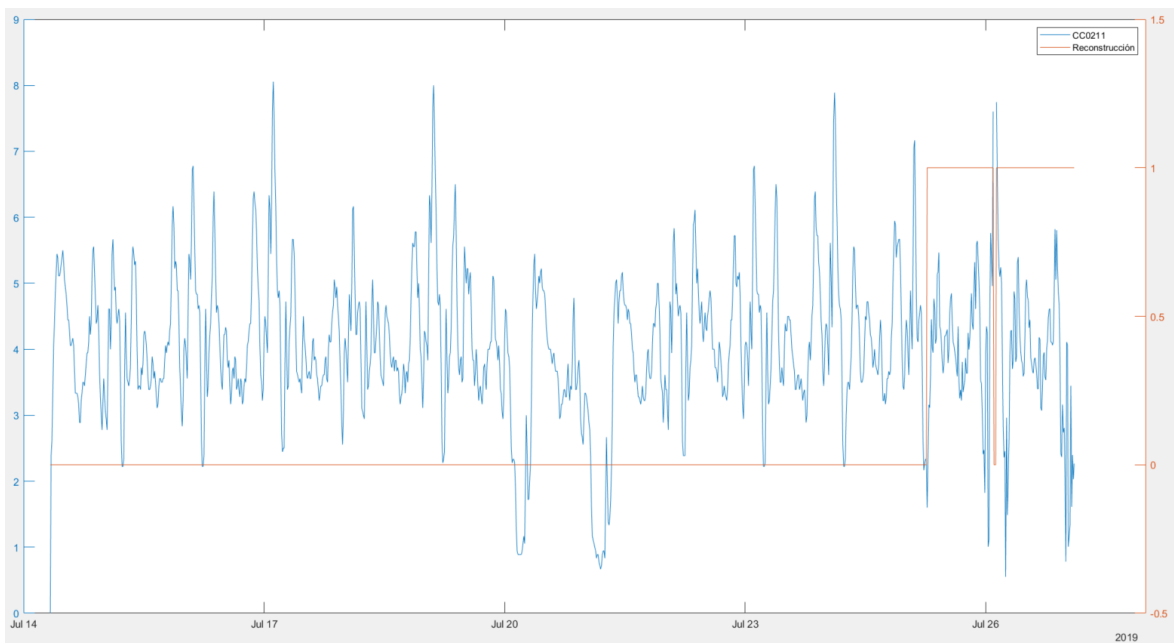


Figura 86. Caso 3 reconstrucción CC0211

7.4 Caso 4: 27/09/2019 – 10/10/2019

CC0216:

Se detecta un fallo puntual en el sector S0203.

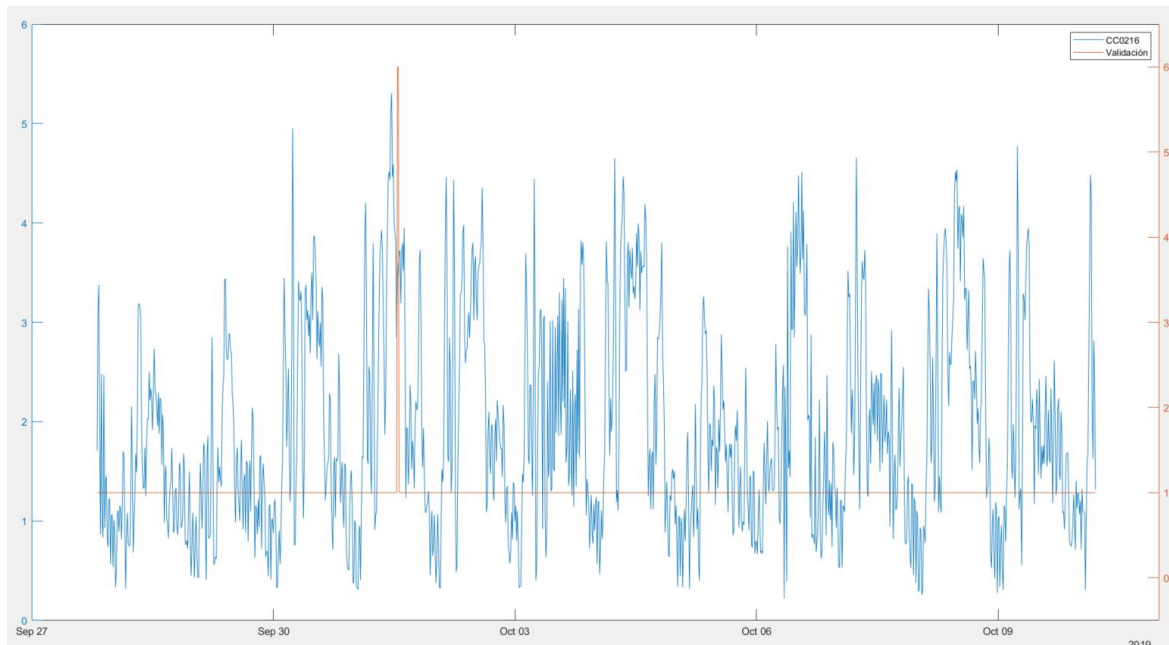


Figura 87. Caso 4 validación CC0216

Mediante la reconstrucción se cambia el valor del sensor CC0216 y se confirma que el consumo del sector es correcto.

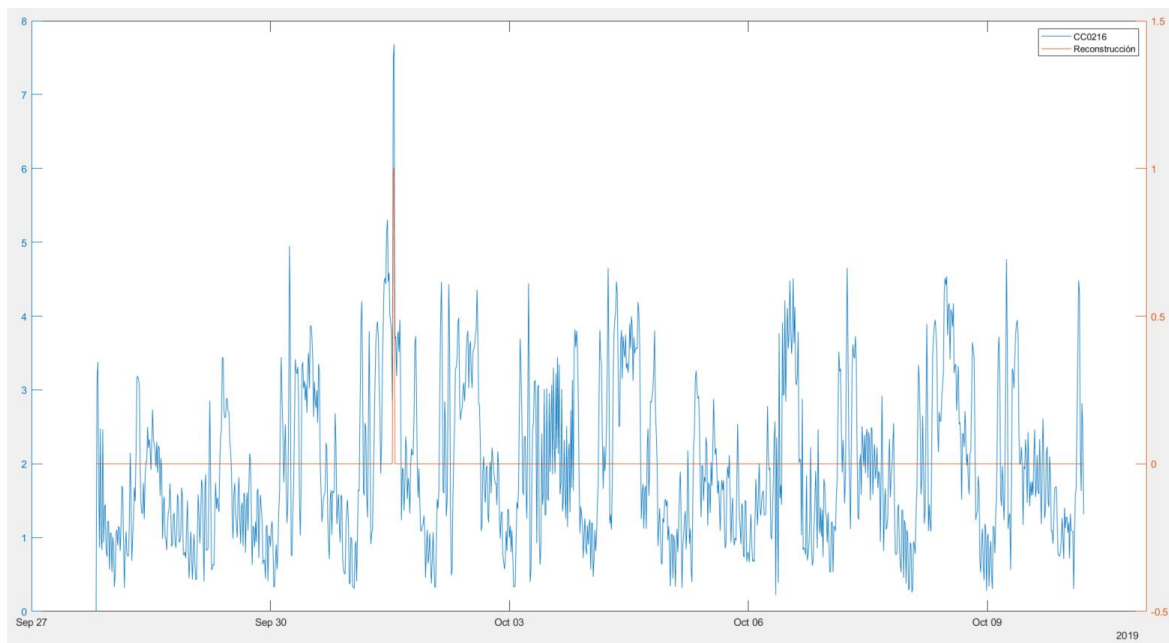


Figura 88. Caso 4 reconstrucción CC0216

Entrando en detalle del sector S0203 para este caso, se observa como en un momento puntual, excede el límite superior establecido.

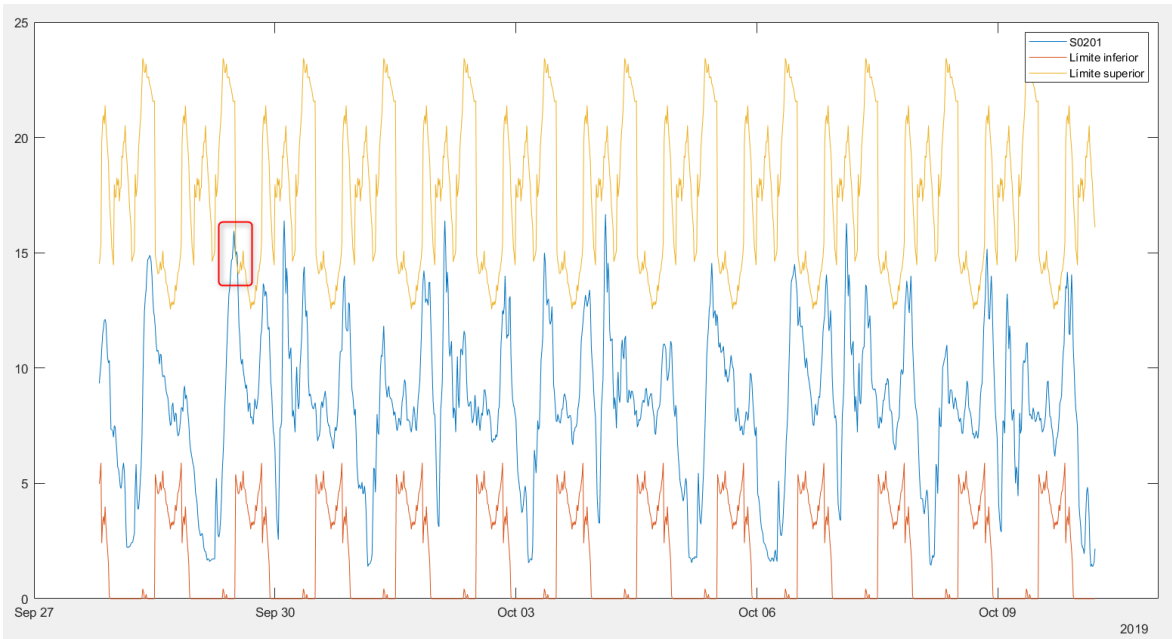


Figura 89. Caso 4 validación S0203

Tras el cambio en el sensor CC0216 se corrige el dato y se verifica que entra dentro del intervalo de confianza.

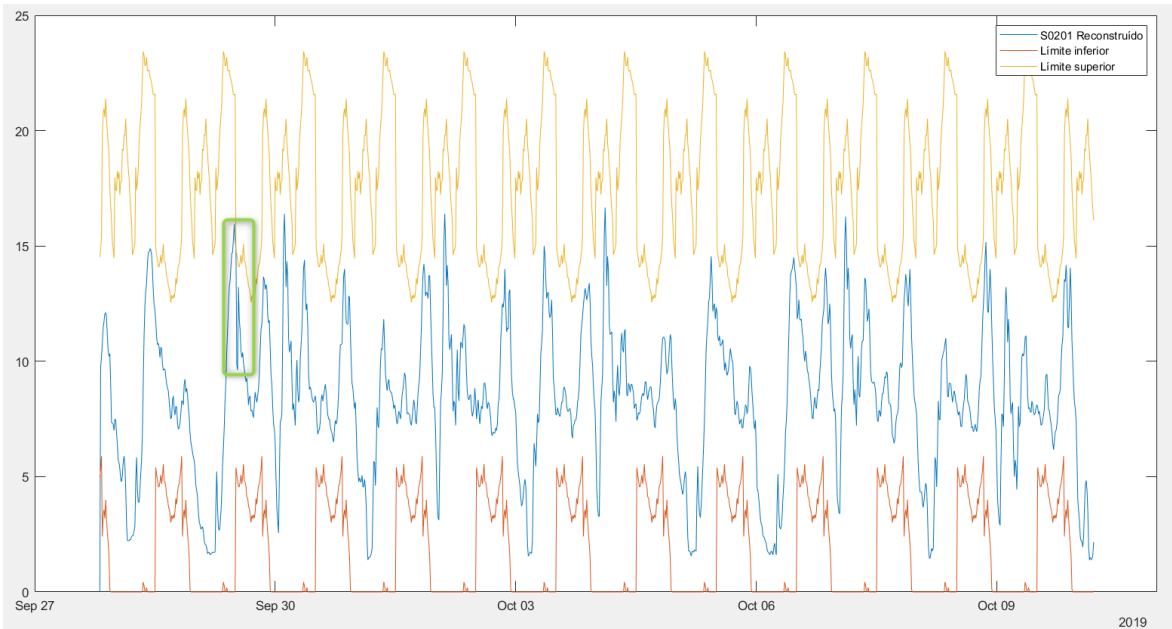


Figura 90. Caso 5 reconstrucción S0203

7.5 Caso 5: 05/12/2019 – 17/12/2019

CC0201:

Se detectan tres fallos en su validación, por discordancia con el sensor de presión.

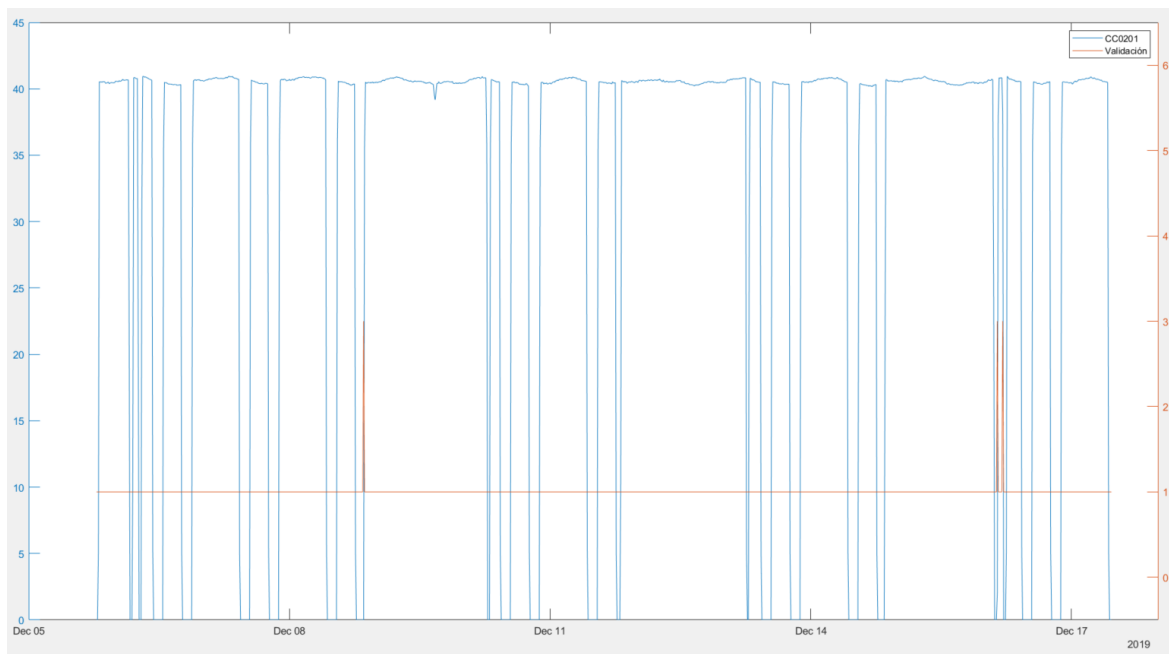


Figura 91. Caso 5 validación CC0201

Tras realizar la reconstrucción verifica el estado de la bomba en los dos primeros errores, mientras que en el tercero no reconstruye debido a que la bomba está apagada, pero el caudal en los 15 minutos no es menor al 10% del caudal normal.

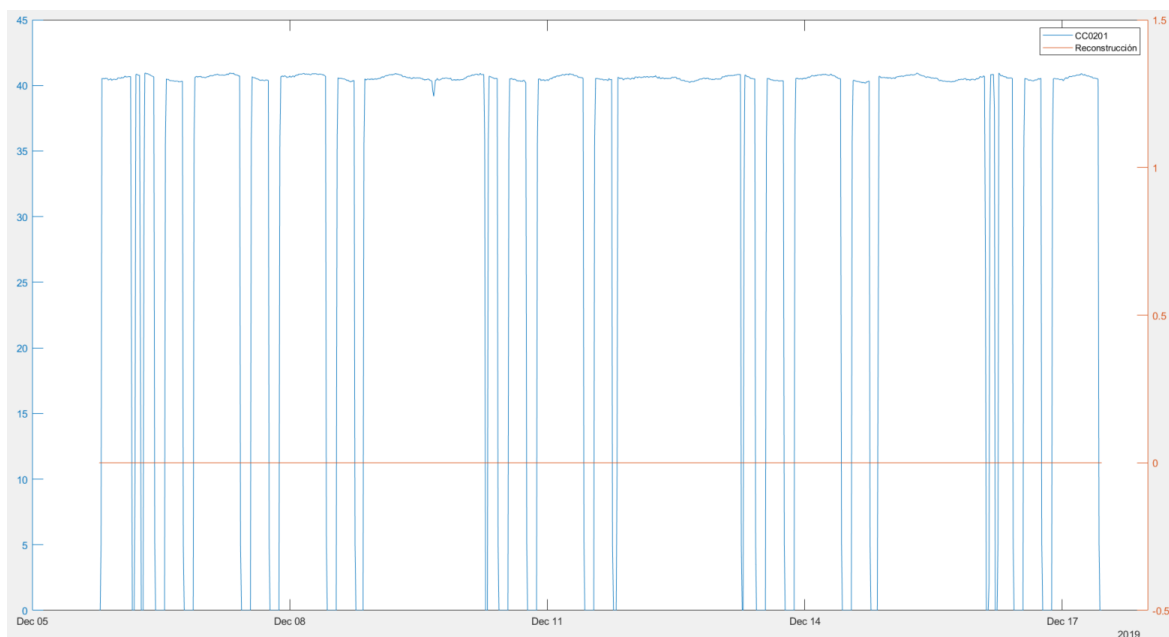


Figura 92. Caso 5 reconstrucción CC0201

CC0210:

Todos los valores son correctos.

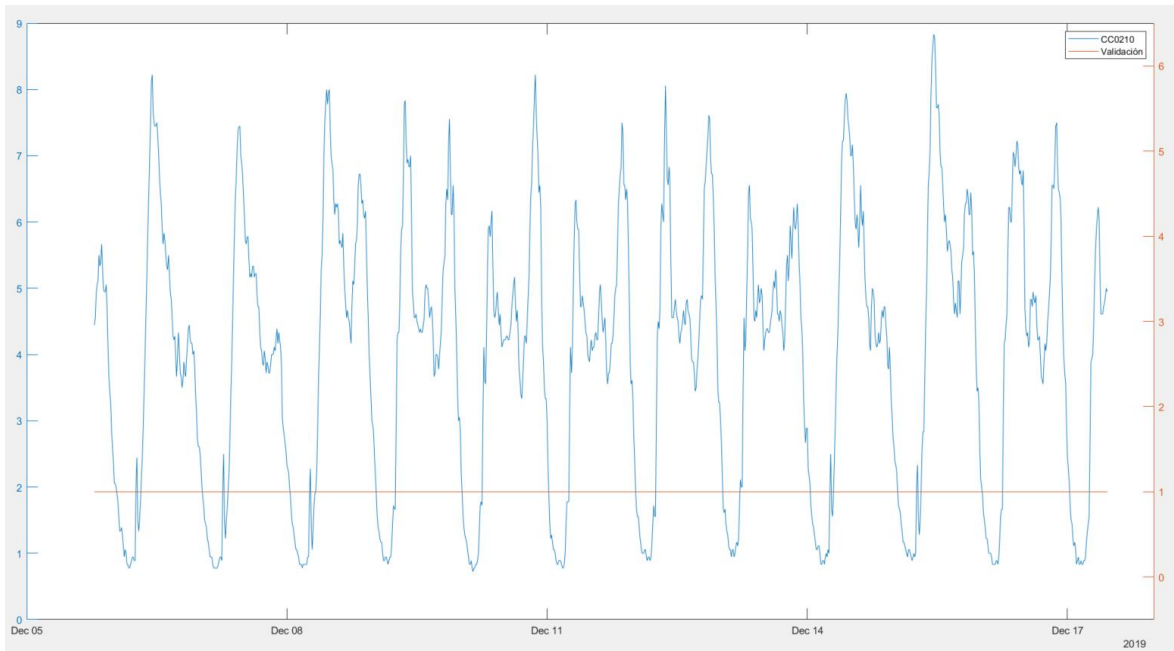


Figura 93. Caso 5 validación CC0210

Al ser los valores correctos no reconstruye.

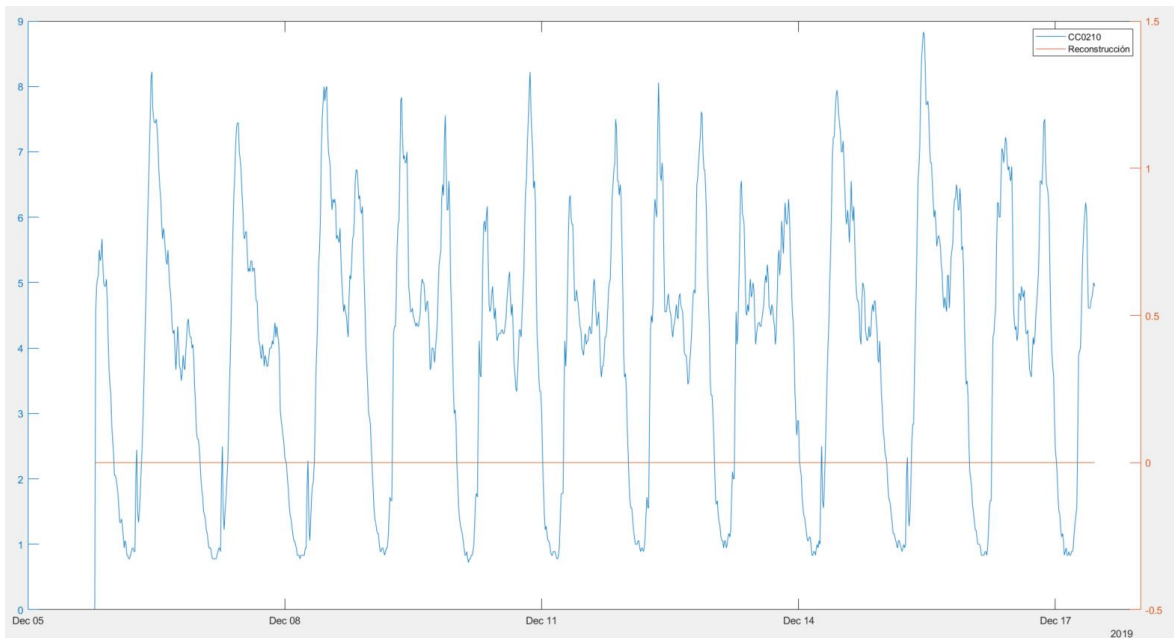


Figura 94. Caso 5 reconstrucción CC0210

CC0211:

Todos los valores son correctos.

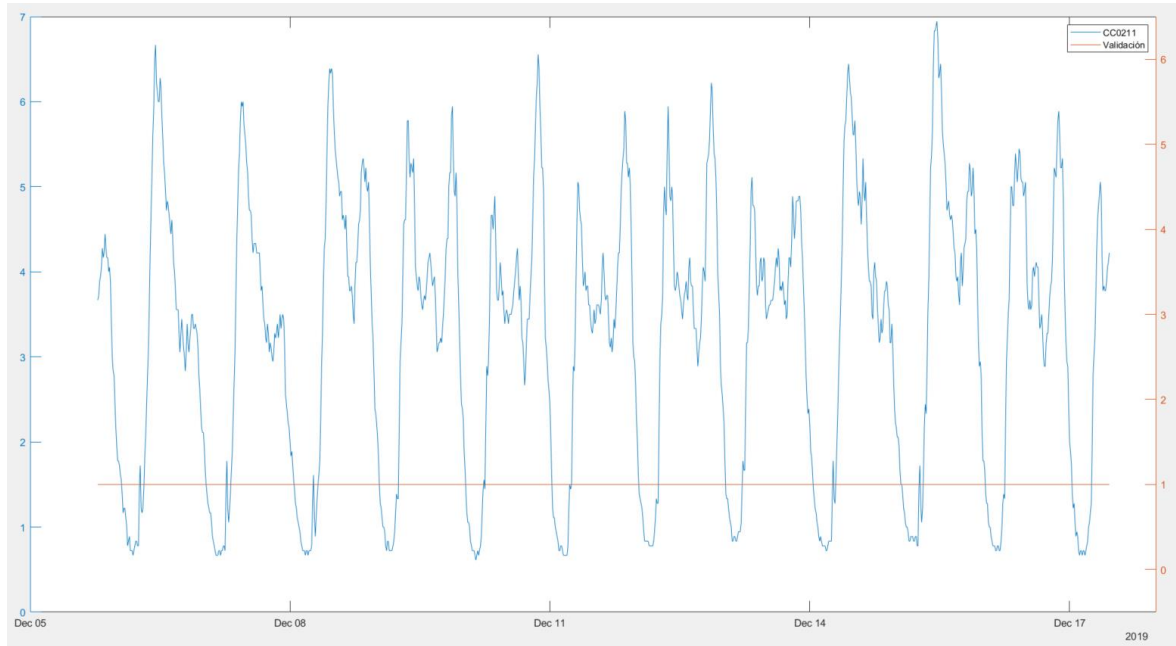


Figura 95. Caso 5 validación CC0211

No hace falta reconstruir ya que los valores son correctos.

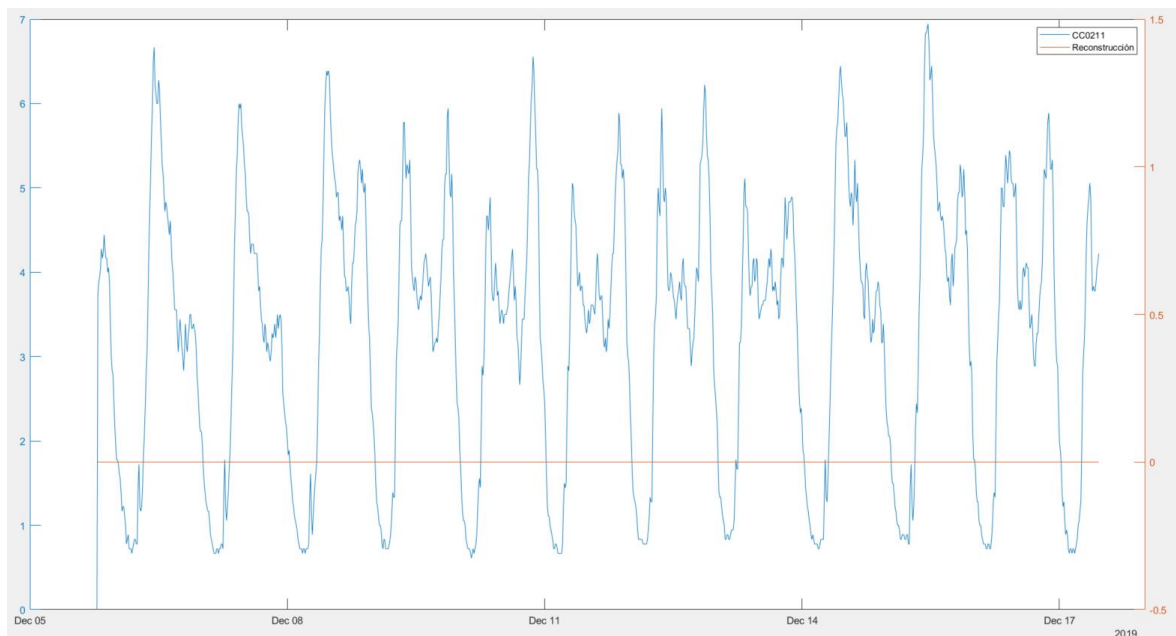


Figura 96. Caso 5 reconstrucción CC0211

8 Conclusiones

El presente proyecto se ha centrado en la validación y reconstrucción de los datos telecontrol de la red de distribución de aguas de Terrassa, concretamente los procedentes de los sensores del piso de Sulleva. Este procedimiento es necesario debido a que es imprescindible llevar un correcto control de los consumos de agua, así como la detección y anticipación de las averías.

Para llevar a cabo el proyecto, en una primera instancia ha sido necesario unificar la periodicidad de los datos para facilitar su validación. En el piso de Sulleva, se han encontrado algunos sensores que proporcionan datos biminutalmente y otros quinceminutalmente. Con el objetivo de tratar con una información más clara y precisa, se decide que la mejor opción es utilizar una frecuencia de datos de 15 minutos. De esta manera, para modificar la frecuencia de los datos biminutales se realiza la media cada 15 minutos consiguiendo un único valor, mientras que los quinceminutales siguen tal y como estaban.

Respecto a la validación, se ha realizado un estudio de las diferentes metodologías para aplicar a cada uno de los casos, dependiendo del sensor en cuestión o de las diferentes circunstancias. Con los métodos elegidos, se ha podido comprobar la veracidad de los datos de los diferentes sensores y conocer cuáles son más propensos al fallo. Posteriormente, se ha seguido un patrón similar estudiando los casos de reconstrucción para encontrar los métodos más adecuados según el contexto.

Partiendo por el caudalímetro de bombeo de agua hacia el depósito (CC0201), se puede afirmar que es un sensor fiable. En un primer momento, debido a la dependencia de otros sensores del piso, existían varios momentos de incertidumbre, es por ello que se solicitó a TAIGUA más información al respecto del funcionamiento de la bomba. Con estos datos extra, se ha asegurado una validación más fiable y ha facilitado la reconstrucción, ya que se dispone de 3 fuentes de información para confirmar qué valor debía ser.

Siguiendo con los componentes del depósito, el caudalímetro de salida (CC0202) es el que más errores se han detectado, principalmente porque tiene una gran ausencia de datos. Además, el hecho que informe los datos instantáneos quinceminutalmente ha dificultado encontrar un modelo diario que sea fiable. La parte positiva de este sensor ha sido la reconstrucción, gracias a la fiabilidad del sensor de entrada y de los sensores de nivel, se ha podido reconstruir eficazmente con la variación del volumen de agua en el depósito y conocer el consumo estimado de todo el piso.

Los sensores de nivel se han mostrados mayormente fiables. La validación se ha podido realizar sin ningún inconveniente, ya que se puede confirmar la información con el otro sensor además de tener una coherencia entre valores consecutivos. En cuanto a la reconstrucción, la falta de valores en el caudalímetro de salida del depósito ha imposibilitado su reconstrucción.

Entrando en los diferentes sectores, el primero a comentar es el S0201, cuya participación directa viene dada por los caudalímetros CC0210 y CC0211. Cuando ambos sensores han funcionado, el sector se ha mantenido dentro del intervalo de confianza determinado en la validación. En cuanto a los caudalímetros, el primer sensor tiene un comportamiento regularmente bueno, sin embargo, el CC0211 se han detectado ciertos momentos de ausencia de una gran cantidad de datos equivalente a días e incluso semanas. Mediante la reconstrucción por el modelo AR y con los datos históricos del consumo del sector, se ha podido reconstruir el valor coherentemente. Cabe decir que ha habido casos que no ha sido posible, ya que los diferentes métodos resultaban erróneos.

El segundo sector que se ha podido estudiar ha sido el S0203, cuyo consumo es la diferencia entre el caudalímetro CC0213 y CC0216. Este sector es algo similar al

comportamiento del anterior, el primer sensor tiene un comportamiento bueno y con poco fallo, mientras que el segundo falla algo más. La validación con los intervalos de confianza ha permitido la validación del sector dando un alto porcentaje de acierto cuando ambos funcionaban. Cabe recordar que el CC0216 es equivalente directamente al consumo del sector S0204, donde se encuentra un hotel que rellena su depósito interno sin ningún tipo de planificación horaria, es decir, se llena cuando es necesario. Esto causa una gran varianza de los datos del sensor, ya que durante un día pueden encontrarse varios picos que destacan sobre el caudal normal, por lo que no se ha podido establecer un modelo de consumo diario de este sector. Con ello, la reconstrucción del sector S0203 principalmente realizada por el modelo AR, ha podido reconstruir gran cantidad de datos ausentes del sensor CC0216, a excepción de algunos casos que ha resultado un valor incongruente.

Por último, el sector S0202, no ha sido utilizado para la validación debido a que tiene dependencia directa del sensor CC0202 y la falta de datos ha imposibilitado determinar un modelo de comportamiento diario.

En cuanto a la BBDD, una vez realizado el modelo entidad-relación y el modelo relacional, se ha creado desde Matlab mediante la conexión ODBC para utilizar Microsoft Access. Tras la validación y la reconstrucción de los datos, se ha añadido correctamente la información en las diferentes tablas.

Como conclusión este proyecto cumple con el objeto y los objetivos previamente establecidos, ya que se han encontrado modelos fiables de validación y reconstrucción, y posteriormente la carga en la BBDD ha sido correcta. Por otro, es de gran importancia haber creado una base donde se pueda seguir construyendo, por ello se han propuesto una serie de posibles mejoras a implementar.

8.1 Posibles mejoras

Ampliación a otros pisos

Como primera mejora, este proyecto se ha centrado en el piso de Sulleva, ya que dentro de los pisos, se encuentra en un nivel medio-bajo de complejidad. En él se encuentran diferentes componentes que pese a que no haya una gran cantidad de sensores, hacen que el estudio de cada uno de ellos sea diferente.

Partiendo de este trabajo, se propone implementar las metodologías aplicadas de validación y reconstrucción en este piso para estudiar otro. Un ejemplo de piso que podría adaptarse es el de Can Colomer.

El piso de Can Colomer cuenta con un depósito y cuatro sectores, al igual que el piso de Sulleva, aunque cabe destacar que tiene más caudalímetros. Por otro lado, en el estudio de los sectores podrían implementarse las mismas metodologías para cada uno de ellos ya que prácticamente tienen la misma estructura en cuanto a entradas y salidas.

El esquema del Piso de Can Colomer es el siguiente:

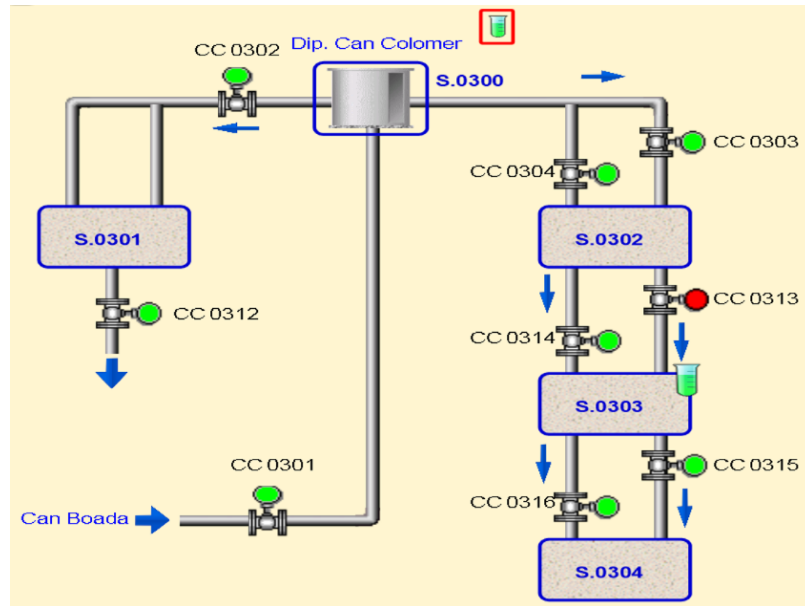


Figura 97 . Piso de Can Colomer

Querys en la BBDD

En este proyecto se ha realizado el diseño y creación de la BBDD así como el almacenamiento de los datos. Sin embargo, una parte fundamental es la accesibilidad de estos gracias a las consultas mediante SQL.

Un detalle importante a mejorar podría ser el realizar unas consultas predeterminadas para que el usuario únicamente tuviera que elegir qué sensor o de qué días quiere ver la información del piso. De esta manera, se realizaría el desarrollo de una aplicación de usuario "front-end" para que personal sin conocimientos técnicos pudieran realizar consultas de forma intuitiva y obtener los datos deseados.

Estudio temporal

Para la validación de los datos, se ha implementado una metodología que consistía en la revisión de varios días de consumo para determinar un consumo medio y aplicando la regla del "68-95-99.7", establecer un intervalo de confianza para determinar la validez de un cierto dato.

Respecto a este punto, se podrían establecer otros métodos que revisaran los datos previos además de los históricos, de esta manera se podría entrar más al detalle de la coherencia de los datos en el momento del año en el que se dan.

Por otro lado, se podría estudiar de establecer un modelo semanal en cambio de diario, así, se tendría en cuenta la variación para los fines de semana.

Red de distribución

Con este proyecto se han podido detectar qué sensores proporcionan información más fiable, y cuáles menos. Con esta información se podría proponer a TAIGUA sustituir los sensores menos fiables o defectuosos.



Otro punto a tener en cuenta respecto a los sensores es el tipo de recogida de datos. Se podría indicar que sustituyendo los sensores instantáneos por sensores acumulativos, aumentaría la fiabilidad de los datos.

9 Referencias

- [1] Taigua - Aigua Municipal de Terrassa [Internet]. [Consultado en marzo de 2022]. Disponible en:
<https://taigua.cat/es/>.
- [2] Taigua - Aigua Municipal de Terrassa [Internet]. Estatutos de Taigua; 20 de julio de 2020 [consultado en marzo de 2022]. Disponible en:
<https://taigua.cat/wp-content/uploads/2022/04/ESTATUTS-TAIGUA.pdf>
- [3] Taigua - Aigua Municipal de Terrassa [Internet]. [Internet]. MEMÒRIA JUSTIFICATIVA PER LA DETERMINACIÓ DE LA FORMA DE GESTIÓ PER A LA PRESTACIÓ DEL SERVEI PÚBLIC D'ABASTAMENT D'AIGUA A TERRASSA; 17 de julio de 2017 [consultado en marzo de 2022]. Disponible en:
https://www.terrassa.cat/documents/12006/15032606/Memoria_gestio_aigua_c.pdf
- [4] Quevedo J, Blanch J, Puig V, Saludes J, Roquet J, Espín S. Fiabilizar datos en el telecontrol de redes de transporte de agua; Febrero de 2010 [Consultado en marzo de 2022].
- [5] Pesántez JL. STUDY OF A FAULT DIAGNOSIS SYSTEM FOR A SENSOR NETWORK; [consultado en abril de 2022].
- [6] Taigua - Aigua Municipal de Terrassa, Ajuntament de Terrassa. Pla Director del Servei d'Abastament d'Aigua Potable de Terrassa, Annex 1: Infraestructures existents [consultado en marzo de 2022].
- [7] Quevedo J, Garcia D, Puig V, Saludes J. Metodología para validación y reconstrucción de datos y su aplicación en redes de agua potable [Internet]; 2 de junio de 2011 [consultado en abril de 2022]. Disponible en:
https://intranet.ceautomatica.es/sites/default/files/upload/955/files/VII_Simposio_de_Control_Inteligente-2011/comunicacion7_Quevedo.pdf
- [8] Colaboradores de los proyectos Wikimedia. Wikipedia, la enciclopedia libre [Internet]. Desviación típica - Wikipedia, la enciclopedia libre; 23 de septiembre de 2003 [consultado en abril de 2022]. Disponible en:
https://es.wikipedia.org/wiki/Desviación_típica
- [9] Colaboradores de los proyectos Wikimedia. Wikipedia, la enciclopedia libre [Internet]. Regla 68-95-99.7 - Wikipedia, la enciclopedia libre; 22 de abril de 2019 [consultado en abril de 2022]. Disponible en:
https://es.wikipedia.org/wiki/Regla_68-95-99.7
- [10] DATA SCIENCE [Internet]. Regla empírica — Matemática y Estadística — DATA SCIENCE; 13 de abril de 2020 [consultado en abril de 2022]. Disponible en:
<https://datascience.eu/es/matematica-y-estadistica/regla-empirica-que-es/>.
- [11] MathWorks - Creadores de MATLAB y Simulink - MATLAB y Simulink - MATLAB & Simulink [Internet]. Linear Prediction and Autoregressive Modeling- MATLAB & Simulink Example- MathWorks España; [consultado en mayo de 2022]. Disponible en:
<https://es.mathworks.com/help/signal/ug/linear-prediction-and-autoregressive-modeling.html>
- [12] Colaboradores de los proyectos Wikimedia. Wikipedia, la enciclopedia libre [Internet]. Modelo autorregresivo - Wikipedia, la enciclopedia libre; 24 de julio de 2013 [consultado en mayo de 2022]. Disponible en:

https://es.wikipedia.org/wiki/Modelo_autorregresivo

[13] Departamento de Estadística - UC3M [Internet]. Procesos Autorregresivos; 17 de septiembre de 2008 [consultado en mayo de 2022]. Disponible en:

<http://halweb.uc3m.es/esp/Personal/personas/imolina/MiDocencia/SeriesTemporales/Tema3SeriesEstud.pdf>

[14] MathWorks - Creadores de MATLAB y Simulink - MATLAB y Simulink - MATLAB & Simulink [Internet]. Estimate parameters when identifying AR model or ARI model for scalar timeseries - MATLAB ar- MathWorks España; [consultado en mayo de 2022]. Disponible en:

https://es.mathworks.com/help/ident/ref/ar.html?searchHighlight=ar&s_tid=srchtitle_ar_1

[15] Blog de ILERNA Online [Internet]. El modelo Entidad-Relación: el esquema de una base de datos; 12 de noviembre de 2019 [consultado en mayo de 2022]. Disponible en:

<https://www.ilerna.es/blog/informatica-comunicacion/modelo-entidad-relacion-base-de-datos/>

[16] Home | Bookdown [Internet]. Capítulo 7 El Modelo Relacional | Base de Datos; 31 de marzo de 2021 [consultado en junio de 2022]. Disponible en:

<https://bookdown.org/paranedagarcia/database/el-modelo-relacional.html>

[17] ERDPlus [Internet]; [consultado en junio de 2022]. Disponible en:

<https://erdplus.com/>

[18] Colaboradores de los proyectos Wikimedia. Wikipedia, la enciclopedia libre [Internet]. Open Database Connectivity - Wikipedia, la enciclopedia libre; 25 de noviembre de 2003 [consultado en junio de 2022]. Disponible en:

[https://es.wikipedia.org/wiki/Open_Database_Connectivity#:~:text=Open%20DataBase%20Connectivity%20\(ODBC\)%20es,\(DBMS\)%20almacene%20los%20datos.](https://es.wikipedia.org/wiki/Open_Database_Connectivity#:~:text=Open%20DataBase%20Connectivity%20(ODBC)%20es,(DBMS)%20almacene%20los%20datos.)

ANEXOS

ANEXO I: Tratado, validación y reconstrucción

```

data = readtable('Dades_Telecontrol_2019.csv', 'Range', 'A2:DI230000', 'TreatAsEm-
pty', {'NA'});
bomb = readtable('Datos_2019_Bombeo.csv', 'Range', 'A1:DI230000', 'TreatAsEm-
pty', {'NA'});
conn = database('BBDD', '', '');

Hbomb = height(bomb);
L = height(data);
load('limites.mat')
load('tiempo.mat')
load('partic.mat')
load('medias.mat')
SU = zeros(); %Matriz para rellenar con valores para reconstruir
date = datetetime(); %Vector con las fechas cada 15 min
datebomb = datetetime();
reg = datetetime();
datebomb = bomb.DATETIME;
AREA = 432;
%Tabla de datos
[Y,M,D,H,MN,S] = datevec(data.Var2);
index=find((MN==00) & (S==0) | (MN==15) & (S==0) | (MN==30) & (S==0) | (MN==45) & (S==0));
Hindex = height(index);

selectquery = 'SELECT TOP 2 ID_medida FROM Datos order by ID_medida desc';
query = select(conn,selectquery);
num = 0;
%num = query.ID_medida(1,1);

%for para rellenar la matriz SU con los valores de los sensores cada 15 min
for i = 1:L
    if i == index(s)
        date(s,1) = data.Var2(i);
        CC0201(s,1) = nanmean(data.Var5(prev:index(s))); %Cálculo de la media con
los valores de los 15 min previos
        CC0202(s,1) = nanmean(data.Var6(prev:index(s)));
        CC0210(s,1) = nanmean(data.Var7(prev:index(s)));
        CC0211(s,1) = nanmean(data.Var8(prev:index(s)));
        CC0213(s,1) = nanmean(data.Var9(prev:index(s)));
        CC0216(s,1) = nanmean(data.Var11(prev:index(s)));
        NV1(s,1) = data.Var112(i);
        NV2(s,1) = data.Var113(i);
        prev = index(s);
        s = s+1;
        if s == Hindex
            s = 1;
        end
    end
end

SU; %Tabla de valores cada 15 min
date; %fechas cada 15 min

for v1 = 2:(Hindex-1)

    %validación CC0201

    q=1;
    for i = 2:Hbomb
        if date(v1,1) < datebomb(i)

```



```
        if ((bomb.Intensidad_Grupo_Bombeo_1(i-1)<= 80 || bomb.Inten-
sidad_Grupo_Bombeo_2(i-1)<= 80) && CC0201(v1,1) < 36) || ((bomb.Inten-
sidad_Grupo_Bombeo_1(i-1) > 80 || bomb.Intensidad_Grupo_Bombeo_2(i-1) > 80) &&
CC0201(v1,1) >= 4)
            val2CC0201(v1,1) = 1;
        else
            val2CC0201(v1,1) = 0;
            valCC0201(v1,1) = 2;
        end

        if (bomb.Presion_Bombeo(i-1) < 110 && CC0201(v1,1) < 36) ||
(bomb.Presion_Bombeo(i-1) >= 110 && CC0201(v1,1) >= 4)
            val3CC0201(v1,1) = 1;
        else
            val3CC0201(v1,1) = 0;
            valCC0201(v1,1) = 3;
        end

        if ((bomb.Intensidad_Grupo_Bombeo_1(i-1)<= 80 || bomb.Inten-
sidad_Grupo_Bombeo_2(i-1)<= 80) && CC0201(1,1) < 36) || ((bomb.Inten-
sidad_Grupo_Bombeo_1(i-1) > 80 || bomb.Intensidad_Grupo_Bombeo_2(i-1) > 80) &&
CC0201(1,1) >= 4)
            val2CC0201(1,1) = 1;
        else
            val2CC0201(1,1) = 0;
            valCC0201(1,1) = 2;
        end

        if (bomb.Presion_Bombeo(i-1) < 110 && CC0201(1,1) < 36) ||
(bomb.Presion_Bombeo(i-1) >= 110 && CC0201(1,1) >= 4)
            val3CC0201(1,1) = 1;
        else
            val3CC0201(1,1) = 0;
            valCC0201(1,1) = 3;
        end

        if bomb.Intensidad_Grupo_Bombeo_1(i-1)<= 80 && bomb.Presion_Bom-
beo(i-1) < 110 || bomb.Intensidad_Grupo_Bombeo_2(i-1)<= 80 && bomb.Presion_Bom-
beo(i-1) < 110
            estCC0201(v1,1) = 0;
        elseif bomb.Intensidad_Grupo_Bombeo_1(i-1)> 80 && bomb.Pre-
sion_Bombeo(i-1) >= 110 || bomb.Intensidad_Grupo_Bombeo_2(i-1)> 80 && bomb.Pre-
sion_Bombeo(i-1) >= 110
            estCC0201(v1,1) = 1;
        else
            estCC0201(v1,1) = NaN;
        end
        reg(q,1) = datebomb(i);
        q=q+1;
        break
    end
end

if (CC0201(v1,1) <= 0 && CC0201(v1,1) > 49) || isnan(CC0201(v1,1))
    val1CC0201(v1,1) = 0;
    valCC0201(v1,1) = 0;
else
    val1CC0201(v1,1) = 1;
end

if (CC0201(1,1) <= 0 && CC0201(1,1) > 49) || isnan(CC0201(1,1))
    val1CC0201(1,1) = 0;
    valCC0201(1,1) = 0;
else
    val1CC0201(1,1) = 1;
end

if val1CC0201(1,1)*val2CC0201(1,1)*val3CC0201(1,1) == 1
```

```

        valCC0201(1,1) = 1;
    end

    if val1CC0201(v1,1)*val2CC0201(v1,1)*val3CC0201(v1,1) == 1
        valCC0201(v1,1) = 1;
    end

    if not (CC0201(v1,1) == 0 && NV1(v1,1) > NV1(v1-1,1))
        wnCC0201(v1,1) = 1;
    else
        wnCC0201(v1,1) = 0;
        valCC0201(v1,1) = 4;
    end
    idCC0201(v1,1) = num+v1;
    idCC0201(1,1) = num+1;
    sensCC0201(v1,1) = "CC0201";
    sensCC0201(1,1) = "CC0201";
    recCC0201(v1,1) = 0;
%-----
%   %validación sensores de nivel
%   %Sensor1
%   %aproximadamente máximo inc/dec de nivel es 0.1
%
    if (NV1(v1,1)-0.1 <= NV1(v1-1,1)) || (NV1(v1,1)+0.1 >= NV1(v1-1,1))
        valNV1(v1,1) = 1;
    else
        valNV1(v1,1) = 5;
    end

    if NV1(v1,1) < 0 || isnan(NV1(v1,1))
        valNV1(v1,1) = 0;
    end
    if valNV1(2,1) == 1
        if (NV1(2,1)-0.1 <= NV1(2,1)) || (NV1(1,1)+0.1 >= NV1(2,1))
            valNV1(1,1) = 1;
        else
            valNV1(1,1) = 5;
        end
    else
        valNV1(1,1) = 5;
    end
    valNV1(1,1) = 5;
    idNV1(v1,1) = num+Hindex+v1-1;
    idNV1(1,1) = num+Hindex;
    sensNV1(v1,1) = "SU_NV1";
    sensNV1(1,1) = "SU_NV1";
    recNV1(v1,1) = 0;

    %Sensor 2
    %aproximadamente máximo inc/dec de nivel es 0.1

    if (NV2(v1,1)-0.1 <= NV2(v1-1,1)) || (NV2(v1,1)+0.1 >= NV2(v1-1,1))
        valNV2(v1,1) = 1;
    else
        valNV2(v1,1) = 5;
    end

    if NV2(v1,1) < 0 || isnan(NV2(v1,1))
        valNV2(v1,1) = 0;
    end

    if valNV2(2,1) == 1
        if (NV2(1,1)-0.1 <= NV2(2,1)) || (NV2(1,1)+0.1 >= NV2(2,1))
            valNV2(1,1) = 1;
        else
            valNV2(1,1) = 5;
        end
    else
        valNV2(1,1) = 5;
    end
end

```



```
    valNV2(1,1) = 5;
end
idNV2(v1,1) = num+(2*Hindex)+v1-2;
idNV2(1,1) = num+(2*Hindex)-1;
sensNV2(v1,1) = "SU_NV2";
sensNV2(1,1) = "SU_NV2";
recNV2(v1,1) = 0;

%-----
%validación CC0202
calculado(v1,1) = CC0201(v1,1) - ((AREA*(NV1(v1,1)-NV1(v1-1,1)))/0.9);

if (calculado(v1,1) >= (CC0202(v1,1)*0.7)) && (calculado(v1,1) <=
(CC0202(v1,1)*1.30))
    valCC0202(v1,1) = 1;
else
    valCC0202(v1,1) = 5;
end

if CC0202(v1,1) <= 0 || isnan(CC0202(v1,1))
    valCC0202(v1,1) = 0;
end
valCC0202(1,1) = 5;
DIFCC0202(v1,1) = abs((calculado(v1,1) - CC0202(v1,1))/calculado(v1,1));
idCC0202(v1,1) = num+(3*Hindex)+v1-3;
idCC0202(1,1) = num+(3*Hindex)-2;
sensCC0202(v1,1) = "CC0202";
sensCC0202(1,1) = "CC0202";
recCC0202(v1,1) = 0;

%-----

[Y,M,D,H,MN,S] = datevec(date(v1,1));
[Y1,M1,D1,H1,MN1,S1] = datevec(date(1,1));
%validación S0201 (CC0210 y CC0211)
S0201(v1,1) = CC0210(v1,1) + CC0211(v1,1);
S0201(1,1) = CC0210(1,1) + CC0211(1,1);

for i = 1:(height(tiempo))
    [Yt,Mt,Dt,Ht,MNt,St] = datevec(tiempo(i,1));
    if H == Ht && MN == MNt
        lsp1(v1,1) = ls1(i,1);
        lip1(v1,1) = li1(i,1);
        if (S0201(v1,1) >= li1(i,1)) && (S0201(v1,1) <= ls1(i,1))
            valS0201(v1,1) = 1;
            valCC0210(v1,1) = 1;
            valCC0211(v1,1) = 1;
        else
            valS0201(v1,1) = 0;
            valCC0210(v1,1) = 6;
            valCC0211(v1,1) = 6;
        end
    elseif H1 == Ht && MN1 == MNt
        lsp1(1,1) = ls1(i,1);
        lip1(1,1) = li1(i,1);
        if (S0201(1,1) >= li1(i,1)) && (S0201(1,1) <= ls1(i,1))
            valS0201(1,1) = 1;
            valCC0210(1,1) = 1;
            valCC0211(1,1) = 1;
        else
            valS0201(1,1) = 0;
            valCC0210(1,1) = 6;
            valCC0211(1,1) = 6;
        end
    end
end
end
```

```

if (CC0210(v1,1) <= 0 || isnan(CC0210(v1,1)))
    valCC0210(v1,1) = 0;
    valS0201(v1,1) = 0;
else
    valCC0210(v1,1) = 1;
end
idCC0210(v1,1) = (3*Hindex)+v1-3;
idCC0210(1,1) = (3*Hindex)-2;
sensCC0210(v1,1) = "CC0210";
sensCC0210(1,1) = "CC0210";
recCC0210(v1,1) = 0;

if CC0211(v1,1) <= 0 || isnan(CC0211(v1,1))
    valCC0211(v1,1) = 0;
    valS0201(v1,1) = 0;
else
    valCC0211(v1,1) = 1;
end
idCC0211(v1,1) = (4*Hindex)+v1-4;
idCC0211(1,1) = (4*Hindex)-3;
sensCC0211(v1,1) = "CC0211";
sensCC0211(1,1) = "CC0211";
recCC0211(v1,1) = 0;
corrCC0211(v1,1) = 0;
%-----

%validación S0203 (CC0213 y CC0216)
S0203(v1,1) = CC0213(v1,1) - CC0216(v1,1);
S0203(1,1) = CC0213(1,1) - CC0216(1,1);

for i = 1:(height(tiempo))
    [Yt,Mt,Dt,Ht,MNt,St] = datevec(tiempo(i,1));
    if H == Ht && MN == MNt
        lsp3(v1,1) = ls3(i,1);
        lip3(v1,1) = li3(i,1);
        if (S0203(v1,1) >= li3(i,1)) && (S0203(v1,1) <= ls3(i,1))
            valS0203(v1,1) = 1;
            valCC0213(v1,1) = 1;
            valCC0216(v1,1) = 1;
        else
            valS0203(v1,1) = 0;
            valCC0213(v1,1) = 6;
            valCC0216(v1,1) = 6;
        end
    elseif H1 == Ht && MN1 == MNt
        lsp3(1,1) = ls3(i,1);
        lip3(1,1) = li3(i,1);
        if (S0203(1,1) >= li3(i,1)) && (S0203(1,1) <= ls3(i,1))
            valS0203(1,1) = 1;
            valCC0213(1,1) = 1;
            valCC0216(1,1) = 1;
        else
            valS0203(1,1) = 0;
            valCC0213(1,1) = 6;
            valCC0216(1,1) = 6;
        end
    end
end

if (CC0213(v1,1) <= 0 || isnan(CC0213(v1,1)))
    valCC0213(v1,1) = 0;
    valS0203(v1,1) = 0;
end

if (CC0216(v1,1) <= 0 || isnan(CC0216(v1,1)))
    valCC0216(v1,1) = 0;
end

```




```
        valS0203(v1,1) = 0;
    end
    recCC0213(v1,1) = 0;
    recCC0216(v1,1) = 0;
end
valCC02011 = ta-
ble(date,CC0201,val1CC0201,val2CC0201,val3CC0201,valCC0201,wnCC0201);
valCC02021 = table(date,CC0202,calculado, valCC0202);
valNIV = table(date,NV1,valNV1,NV2,valNV2);

valS02011 = table(S0201,lsp1,lip1,valS0201,CC0210,valCC0210, CC0211,valCC0211);
valS02031 = table(S0203,lsp3,lip3,valS0203,CC0213,valCC0213, CC0216,valCC0216);

%% Reconstrucción
for r = 2:(Hindex-1)
    %Reconstrucción CC0201
    if valCC0201(r,1) == 4
        R_CC0201(r,1) = CC0201(r,1);
        recCC0201(r,1) = 0;
    elseif valCC0201(r,1) == 3
        if val1CC0201(r,1) == 1 && val2CC0201(r,1) == 1
            R_CC0201(r,1) = CC0201(r,1);
            recCC0201(r,1) = 0;
        else
            R_CC0201(r,1) = CC0201(r,1);
            recCC0201(r,1) = 0;
        end
    elseif valCC0201(r,1) == 2
        if val1CC0201(r,1) == 1 && val3CC0201(r,1) == 1
            R_CC0201(r,1) = CC0201(r,1);
            recCC0201(r,1) = 0;
        else
            R_CC0201(r,1) = CC0201(r,1);
            recCC0201(r,1) = 0;
        end
    elseif valCC0201(r,1) == 0
        if val2CC0201(r,1) == 1 && val3CC0201(r,1) == 1
            R_CC0201(r,1) = CC0201(r,1);
            recCC0201(r,1) = 0;
        elseif estCC0201(r,1) == 0
            R_CC0201(r,1) = 0;
            recCC0201(r,1) = 1;
        elseif estCC0201(r,1) == 1
            R_CC0201(r,1) = 40;
            recCC0201(r,1) = 1;
        else
            R_CC0201(r,1) = CC0201(r,1);
            recCC0201(r,1) = 0;
        end
    else
        R_CC0201(r,1) = CC0201(r,1);
        recCC0201(r,1) = 0;
    end
end
R_CC0201(1,1) = CC0201(1,1);

%Reconstrucción NV1
R_NV1(r,1) = NV1(r,1);
if (valNV1(r,1) == 0)
    if (valNV1(r-1,1) == 1) && (valNV1(r+1,1) == 5)
        R_NV1(r,1) = (NV1(r-1,1)+NV1(r+1,1))/2;
        recNV1(r,1) = 1;
    elseif (valCC0201(r,1) == 1) && (valCC0202(r,1) == 1)
        R_NV1(r,1) = ((CC0202(r,1) - CC0201(r,1))*0.9)/AREA - R_NV1(r-1,1);
    end
end
```

```

elseif valNV2(r,1) == 1
    R_NV1(r,1) = valNV2(r,1);
else
    R_NV1(r,1) = ((calculado(r,1) - CC0201(r,1))*0.9)/AREA - R_NV1(r-
1,1);
end
elseif (valNV1(r,1) == 5)
    if (NV1(r,1)-0.1 <= R_NV1(r-1,1)) || (NV1(r,1)+0.1 >= R_NV1(r-1,1))
        R_NV1(r,1) = NV1(r,1);
        recNV1(r,1) = 1;
    end
else
    R_NV1(r,1) = NV1(r,1);
    recNV1(r,1) = 0;
end
R_NV1(1,1) = NV1(1,1);
recNV1(1,1) = 0;

%Reconstrucción NV2
R_NV2(r,1) = NV2(r,1);
if (valNV2(r,1) == 0)
    if (valNV2(r-1,1) == 1) && (valNV2(r+1,1) == 5)
        R_NV2(r,1) = (NV2(r-1,1)+NV2(r+1,1))/2;
        recNV2(r,1) = 1;

    elseif (valCC0201(r,1) == 1) && (valCC0202(r,1) == 1)
        R_NV2(r,1) = ((CC0202(r,1) - CC0201(r,1))*0.9)/AREA - R_NV2(r-1,1);

    elseif valNV1(r,1) == 1
        R_NV2(r,1) = NV1(r,1);
    end
elseif (valNV2(r,1) == 5)
    if (NV2(r,1)-0.1 <= R_NV2(r-1,1)) || (NV2(r,1)+0.1 >= R_NV2(r-1,1))
        R_NV2(r,1) = NV2(r,1);
        recNV2(r,1) = 1;
    end
else
    R_NV2(r,1) = NV2(r,1);
    recNV2(r,1) = 0;
end
R_NV2(1,1) = NV2(1,1);
recNV2(1,1) = 0;

%Reconstrucción CC0202
if (valCC0202(r,1) ~= 1)
    if (valCC0201(r,1) == 1) && (valNV1(r,1) == 1)
        if calculado(r,1) > 0 && calculado(r,1) < 60
            R_CC0202(r,1) = CC0201(r,1) - ((AREA*(NV1(r,1)-NV1(r-1,1)))/0.9);
            recCC0202(r,1) = 1;
        else
            R_CC0202(r,1) = CC0202(r,1);
            recCC0202(r,1) = 0;
        end

    elseif ((valCC0201(r,1) == 1) && (valNV2(r,1) == 1))
        if calculado(r,1) > 0 && calculado(r,1) < 60
            R_CC0202(r,1) = CC0201(r,1) - ((AREA*(NV2(r,1)-NV2(r-1,1)))/0.9);
            recCC0202(r,1) = 1;
        else
            R_CC0202(r,1) = CC0202(r,1);
            recCC0202(r,1) = 0;
        end
    elseif (recCC0201(r,1) == 1)
        if calculado(r,1) > 0 && calculado(r,1) < 60
            R_CC0202(r,1) = R_CC0201(r,1) - ((AREA*(R_NV1(r,1)-R_NV1(r-
1,1)))/0.9);
            recCC0202(r,1) = 1;
        end
    end
end

```

```

else
    R_CC0202(r,1) = CC0202(r,1);
    recCC0202(r,1) = 0;
end
else
    R_CC0202(r,1) = CC0202(r,1);
    recCC0202(r,1) = 0;

end
else
    R_CC0202(r,1) = CC0202(r,1);
    recCC0202(r,1) = 0;
end
R_CC0202(1,1) = CC0202(1,1);

%Reconstrucción S0201
R_S0201(r,1) = CC0210(r,1) + CC0211(r,1);
p_CC0210 = CC0210;
p_CC0211 = CC0211;
R_CC0210(r,1) = CC0210(r,1);
R_CC0211(r,1) = CC0211(r,1);

if (valS0201(r,1) ~= 1)
    if (valS0201(r-1,1) == 1) && (valS0201(r+1,1) == 1)
        p_CC0210(r,1) = (p_CC0210(r+1,1)+p_CC0210(r-1,1))/2;
        R_S0201(r,1) = p_CC0210(r,1) + CC0211(r,1);
        if (R_S0201(r,1) >= lip1(r,1)) && (R_S0201(r,1) <= lsp1(r,1))
            recCC0210(r,1) = 1;
            R_CC0210(r,1) = p_CC0210(r,1);
            R_CC0211(r,1) = CC0211(r,1);
            recCC0211(r,1) = 1;
        else
            p_CC0213(r,1) = (p_CC0213(r+1,1)+p_CC0213(r-1,1))/2;
            R_S0201(r,1) = CC0210(r,1) + p_CC0211(r,1);
            if (R_S0201(r,1) >= lip1(r,1)) && (R_S0201(r,1) <= lsp1(r,1))
                recCC0210(r,1) = 1;
                recCC0211(r,1) = 1;
                R_CC0210(r,1) = CC0210(r,1);
                R_CC0211(r,1) = p_CC0211(r,1);
            end
        end
    end
else
    consum1 = [0; 0; 0; 0; 0];
    if r > 480 && ~(isnan(S0201(r,1)))
        for i=1:5
            for c=1:96
                if i == 1
                    consum1(5,1) = consum1(5,1) + R_S0201(r-c,1);
                elseif i == 2
                    consum1(4,1) = consum1(4,1) + R_S0201(r-96-c,1);
                elseif i == 3
                    consum1(3,1) = consum1(3,1) + R_S0201(r-2*96-c,1);
                elseif i == 4
                    consum1(2,1) = consum1(2,1) + R_S0201(r-3*96-c,1);
                elseif i == 5
                    consum1(1,1) = consum1(1,1) + R_S0201(r-4*96-c,1);
                end
            end
        end
        end
        AR1 = ar(consum1,2);
        consdia1 = (AR1.A(1,1) + consum1(5,1)*AR1.A(1,2) + con-
sum1(4,1)*AR1.A(1,3))*-1;
        for i = 1:(height(tiempo))
            [Yt,Mt,Dt,Ht,MNt,St] = datevec(tiempo(i,1));
            [Y,M,D,H,MN,S] = datevec(date(r,1));
            if H == Ht && MN == MNt

```

```

R_S0201(r,1) = consdial*partS0201(i,1);
R_CC0211(r,1) = R_S0201(r,1) - R_CC0210(r,1);
if R_CC0211(r,1) > 0.5 && R_CC0211(r,1) < 9 &&
R_CC0210(r,1) > R_CC0211(r,1) && R_CC0210(r,1)*0.5 < R_CC0211(r,1)
    recCC0211(r,1) = 1;
    col(r,1) = consdial;
elseif r > 672
    R_S0201(r,1) = R_S0201(r-672,1);
    R_CC0211(r,1) = R_S0201(r,1) - R_CC0210(r,1);
    recCC0211(r,1) = 1;
    if R_CC0211(r,1) < 0.5 || R_CC0211(r,1) > 9 ||
R_CC0210(r,1) < R_CC0211(r,1)
        R_CC0211(r,1) = R_CC0210(r,1)*0.85;
        R_S0201(r,1) = R_CC0210(r,1) + R_CC0211(r,1);
    elseif R_CC0211(r,1) < 0.5 || R_CC0211(r,1) > 9 ||
R_CC0210(r,1) < R_CC0211(r,1)
        R_S0201(r,1) = medS0201(i,1);
        R_CC0211(r,1) = R_S0201(r,1) - R_CC0210(r,1);
        recCC0211(r,1) = 1;
        if R_CC0211(r,1) < 0.5 || R_CC0211(r,1) > 9 ||
R_CC0210(r,1) < R_CC0211(r,1)
            recCC0211(r,1) = 0;
            R_CC0211(r,1) = CC0211(r,1);
        end
    end
end
else
    R_S0201(r,1) = medS0201(i,1);
    R_CC0211(r,1) = R_S0201(r,1) - R_CC0210(r,1);
    recCC0211(r,1) = 1;
    if R_CC0211(r,1) < 0.5 || R_CC0211(r,1) > 9 ||
R_CC0210(r,1) < R_CC0211(r,1)
        recCC0211(r,1) = 0;
        R_CC0211(r,1) = CC0211(r,1);
    end
end
end
end
else
    for i = 1:(height(tiempo))
        [Yt,Mt,Dt,Ht,MNt,St] = datevec(tiempo(i,1));
        [Y,M,D,H,MN,S] = datevec(date(r,1));
        if H == Ht && MN == MNt
            R_S0201(r,1) = medS0201(i,1);
            R_CC0211(r,1) = R_S0201(r,1) - R_CC0210(r,1);
            recCC0211(r,1) = 1;
            if R_CC0211(r,1) < 0.5 || R_CC0211(r,1) > 9 ||
R_CC0210(r,1) < R_CC0211(r,1)
                R_CC0211(r,1) = R_CC0210(r,1)*0.85;
                R_S0201(r,1) = R_CC0210(r,1) + R_CC0211(r,1);
                if R_CC0211(r,1) < 0.5 || R_CC0211(r,1) > 9
                    recCC0211(r,1) = 0;
                    R_CC0211(r,1) = CC0211(r,1);
                end
            end
        end
    end
end
end
end
else
    R_CC0210(r,1) = CC0210(r,1);
    R_CC0211(r,1) = CC0211(r,1);
    recCC0210(r,1) = 0;
    recCC0211(r,1) = 0;
end
end

```

```

%Reconstrucción S0203
R_S0203(r,1) = CC0213(r,1) - CC0216(r,1);
p_CC0213 = CC0213;
p_CC0216 = CC0216;
R_CC0213(r,1) = CC0213(r,1);
R_CC0216(r,1) = CC0216(r,1);

if (valS0203(r,1) ~= 1)
    if (valS0203(r-1,1) == 1) && (valS0203(r+1,1) == 1)
        p_CC0213(r,1) = (p_CC0213(r+1,1)+p_CC0213(r-1,1))/2;
        R_S0203(r,1) = p_CC0213(r,1) - CC0216(r,1);
        if (R_S0203(r,1) >= lip3(r,1)) && (R_S0203(r,1) <= lsp3(r,1))
            recCC0213(r,1) = 1;
            R_CC0213(r,1) = p_CC0213(r,1);
            R_CC0216(r,1) = CC0216(r,1);
            recCC0216(r,1) = 1;
        else
            p_CC0216(r,1) = (p_CC0216(r+1,1)+p_CC0216(r-1,1))/2;
            R_S0203(r,1) = CC0213(r,1) - p_CC0216(r,1);
            if (R_S0203(r,1) >= lip3(r,1)) && (R_S0203(r,1) <= lsp3(r,1))
                recCC0213(r,1) = 1;
                recCC0216(r,1) = 1;
                R_CC0213(r,1) = CC0213(r,1);
                R_CC0216(r,1) = p_CC0216(r,1);
            else
                recCC0213(r,1) = 0;
                recCC0216(r,1) = 0;
                R_CC0213(r,1) = CC0213(r,1);
                R_CC0216(r,1) = CC0216(r,1);
            end
        end
    end
else
    consum = [0; 0; 0; 0; 0];
    if r > 480
        for i=1:5
            for c=1:96
                if i == 1
                    consum(5,1) = consum(5,1) + R_S0203(r-c,1);
                elseif i == 2
                    consum(4,1) = consum(4,1) + R_S0203(r-96-c,1);
                elseif i == 3
                    consum(3,1) = consum(3,1) + R_S0203(r-2*96-c,1);
                elseif i == 4
                    consum(2,1) = consum(2,1) + R_S0203(r-3*96-c,1);
                elseif i == 5
                    consum(1,1) = consum(1,1) + R_S0203(r-4*96-c,1);
                end
            end
        end
        AR = ar(consum,2);
        consdia = (AR.A(1,1) + consum(5,1)*AR.A(1,2) + con-
sum(4,1)*AR.A(1,3))*-1;
        for i = 1:(height(tiempo))
            [Yt,Mt,Dt,Ht,MNt,St] = datevec(tiempo(i,1));
            [Y,M,D,H,MN,S] = datevec(date(r,1));
            if H == Ht && MN == MNt
                R_S0203(r,1) = consdia*partS0203(i,1);
                R_CC0216(r,1) = R_CC0213(r,1) - R_S0203(r,1);
                if R_CC0216(r,1) > 0.25 && R_CC0216(r,1) < 8 &&
lsp3(r,1) >= R_S0203(r,1) && lip3(r,1) <= R_S0203(r,1)
                    recCC0216(r,1) = 1;
                    recCC0213(r,1) = 0;
                elseif r > 672
                    R_S0203(r,1) = R_S0203(r-672,1);
                    R_CC0216(r,1) = R_CC0213(r,1) - R_S0203(r,1);
                    recCC0216(r,1) = 1;
                    if R_CC0216(r,1) < 0.25 || R_CC0216(r,1) > 8

```

```

        R_CC0216(r,1) = 2.5;
        R_CC0213(r,1) = R_CC0216(r,1) + R_S0203(r,1);
        elseif R_CC0216(r,1) < 0.25 || R_CC0216(r,1) > 8 ||
lsp3(r,1) < R_S0203(r,1) || lip3(r,1) > R_S0203(r,1)
            R_S0203(r,1) = medS0203(i,1);
            R_CC0216(r,1) = R_CC0213(r,1) - R_S0203(r,1);
            recCC0216(r,1) = 1;
            if R_CC0216(r,1) < 0.25 || R_CC0216(r,1) > 8
                recCC0216(r,1) = 0;
                R_CC0216(r,1) = CC0216(r,1);
            end
        end
    else
        R_S0203(r,1) = medS0203(i,1);
        R_CC0216(r,1) = R_CC0213(r,1) - R_S0203(r,1);
        recCC0216(r,1) = 1;
        if R_CC0216(r,1) < 0.25 || R_CC0216(r,1) > 8
            recCC0216(r,1) = 0;
            R_CC0216(r,1) = CC0216(r,1);
        end
    end
end
end
else
    for i = 1:(height(tiempo))
        [Yt,Mt,Dt,Ht,MNt,St] = datevec(tiempo(i,1));
        [Y,M,D,H,MN,S] = datevec(date(r,1));
        if H == Ht && MN == MNt
            R_S0203(r,1) = medS0203(i,1);
            R_CC0216(r,1) = R_CC0213(r,1) - R_S0203(r,1);
            recCC0216(r,1) = 1;
            if R_CC0216(r,1) < 0.25
                recCC0216(r,1) = 1;
                R_CC0216(r,1) = 0.25;
            elseif R_CC0216(r,1) >= R_CC0213(r,1)
                recCC0216(r,1) = 0;
                R_CC0216(r,1) = CC0216(r,1);
            end
        end
    end
end
end
end
else
    R_CC0213(r,1) = CC0213(r,1);
    R_CC0216(r,1) = CC0216(r,1);
    recCC0213(r,1) = 0;
    recCC0216(r,1) = 0;

end
cont(r,1)=r;
eee(r,1) = R_CC0211(r,1)/R_CC0210(r,1);

end
valCC02011 = ta-
ble(date,CC0201,val1CC0201,val2CC0201,val3CC0201,valCC0201,R_CC0201,recCC0201);
valCC02021 = table(date,CC0202,calculado, valCC0202, R_CC0202, recCC0202);
valNIV = table(date,NV1,valNV1,R_NV1, recNV1,NV2,valNV2,R_NV2, recNV2);
valS02011 = table(S0201,lsp1,lip1,vals0201,CC0210,valCC0210,R_CC0210,recCC0210,
CC0211,valCC0211,R_CC0211,recCC0211);
valS02031 = table(S0203,lsp3,lip3,vals0203,CC0213,valCC0213,R_CC0213,recCC0213,
CC0216,valCC0216,R_CC0216,recCC0216);

```

ANEXO II: Estudio consumo diario

```
[Y,M,D,H,MN,S] = datevec(DadesTelecontrol2019.DATETIME);
index=find((MN==00) & (S==0)) | (MN==15) & (S==0) | (MN==30) & (S==0) | (MN==45) & (S==0));
L = height(S);
Hindex = height(index);
hora = datetime();
s=1;
for i = 1:L
    if i == index(s)
        hora(s,1) = DadesTelecontrol2019.TIME(i);
        CC202(s,1) = DadesTelecontrol2019.CC0202(i);
        CC210(s,1) = DadesTelecontrol2019.CC0210(i);
        CC211(s,1) = DadesTelecontrol2019.CC0211(i);
        CC213(s,1) = DadesTelecontrol2019.CC0213(i);
        CC216(s,1) = DadesTelecontrol2019.CC0216(i);

        s = s+1;
        if s == Hindex
            s = 1;
        end
    end
end

horaheigth = height(hora);
for y = 1:horaheigth
    S0201(y,1) = CC210(y,1) + CC211(y,1);
    S0202(y,1) = CC202(y,1) - (CC210(y,1) + CC211(y,1)) - CC213(y,1);
    S0203(y,1) = CC213(y,1) - CC216(y,1);
    S0204(y,1) = CC216(y,1);
end
PR1 = S0203(1440:2410,1);
PR2 = hora(1440:2410,1);
tiempo = unique(hora);
th = height(tiempo);
[Y,M,D,H,MN,S] = datevec(tiempo);
ma =find((H==5) & (MN==00) & (S==0));
me =find((H==12) & (MN==00) & (S==0));
ta =find((H==20) & (MN==00) & (S==0));
no =find((H==23) & (MN==45) & (S==0));
t = 1;
calc1 = zeros(); %para calcular las medias
calc2 = zeros();
calc3 = zeros();
calc4 = zeros();
medS0201 = zeros(); %medias de sectores
medS0202 = zeros();
medS0203 = zeros();
medS0204 = zeros();
ls1 = zeros(); %límite superior
ls2 = zeros();
ls3 = zeros();
ls4 = zeros();
li1 = zeros(); %límite inferior
li2 = zeros();
li3 = zeros();
li4 = zeros();
for j =1:th

    r = 1;
    for t=1:horaheigth
        if tiempo(j) == hora(t) && r < 16
            calc1(j,r) = S0201(t);
            calc2(j,r) = S0202(t);
            calc3(j,r) = S0203(t);
            calc4(j,r) = S0204(t);
```

```

        r = r+1;
    end
end
end

dsS0201ma = nanstd(reshape(calc1(1:ma,1:15),ma*15,1));
dsS0201me = nanstd(reshape(calc1(me:me,1:15),(me-ma+1)*15,1));
dsS0201ta = nanstd(reshape(calc1(me:ta,1:15),(ta-me+1)*15,1));
dsS0201no = nanstd(reshape(calc1(ta:no,1:15),(no-ta+1)*15,1));
dsS0202ma = nanstd(reshape(calc2(1:ma,1:15),ma*15,1));
dsS0202me = nanstd(reshape(calc2(me:me,1:15),(me-ma+1)*15,1));
dsS0202ta = nanstd(reshape(calc2(me:ta,1:15),(ta-me+1)*15,1));
dsS0202no = nanstd(reshape(calc2(ta:no,1:15),(no-ta+1)*15,1));
dsS0203ma = nanstd(reshape(calc3(1:ma,1:15),ma*15,1));
dsS0203me = nanstd(reshape(calc3(me:me,1:15),(me-ma+1)*15,1));
dsS0203ta = nanstd(reshape(calc3(me:ta,1:15),(ta-me+1)*15,1));
dsS0203no = nanstd(reshape(calc3(ta:no,1:15),(no-ta+1)*15,1));
dsS0204ma = nanstd(reshape(calc3(1:ma,1:15),ma*15,1));
dsS0204me = nanstd(reshape(calc3(me:me,1:15),(me-ma+1)*15,1));
dsS0204ta = nanstd(reshape(calc3(me:ta,1:15),(ta-me+1)*15,1));
dsS0204no = nanstd(reshape(calc3(ta:no,1:15),(no-ta+1)*15,1));

for j =1:th

    %cálculo medias y valores umbrales
    f = 3;
    if j <= ma
        ds1 = dsS0201ma;
        ds2 = dsS0202ma;
        ds3 = dsS0203ma;
        ds4 = dsS0204ma;
    elseif (ma < j) && (j <= me)
        ds1 = dsS0201me;
        ds2 = dsS0202me;
        ds3 = dsS0203me;
        ds4 = dsS0204me;
    elseif (me < j) && (j <= ta)
        ds1 = dsS0201ta;
        ds2 = dsS0202ta;
        ds3 = dsS0203ta;
        ds4 = dsS0204ta;
    elseif (ta < j) && (j <= no)
        ds1 = dsS0201no;
        ds2 = dsS0202no;
        ds3 = dsS0203no;
        ds4 = dsS0204no;
    end
    medS0201(j,1) = nanmean(calc1(j,:));
    ls1(j,1) = medS0201(j,1) + f*ds1;
    li1(j,1) = medS0201(j,1) - f*ds1;
    if li1(j,1) < 0
        li1(j,1) = 0;
    end
    medS0202(j,1) = nanmean(calc2(j,:));
    ls2(j,1) = medS0202(j,1) + f*ds2;
    li2(j,1) = medS0202(j,1) - f*ds2;
    if li2(j,1) < 0
        li2(j,1) = 0;
    end
    medS0203(j,1) = nanmean(calc3(j,:));
    ls3(j,1) = medS0203(j,1) + f*ds3;
    li3(j,1) = medS0203(j,1) - f*ds3;
    if li3(j,1) < 0
        li3(j,1) = 0;
    end
    medS0204(j,1) = nanmean(calc4(j,:));
    ls4(j,1) = medS0204(j,1) + f*ds4;
end
end
end

```




```
    li4(j,1) = medS0204(j,1) - f*ds4;
    if li4(j,1) < 0
        li4(j,1) = 0;
    end
end
tab = table((1:96)',medS0201,ls1,li1);

figure(1)
stem(hora,S0201,'o-')
hold on
plot(tiempo,medS0201,tiempo,ls1,tiempo,li1)
legend('S0201','medS0201','ls1','li1')
hold off

figure(2)
stem(hora,S0202,'o-')
hold on
plot(tiempo,medS0202,tiempo,ls2,tiempo,li2)
legend('S0202','medS0202','ls2','li2')
hold off

figure(3)
stem(hora,S0203,'o-')
hold on
plot(tiempo,medS0203,tiempo,ls3,tiempo,li3)
legend('S0203','medS0203','ls3','li3')
hold off

figure(4)
stem(hora,S0204,'o-')
hold on
plot(tiempo,medS0204,tiempo,ls4,tiempo,li4)
legend('S0204','medS0204','ls4','li4')
hold off

%% PARTICIPACIÓN QUINCEMINUTAL
for j =1:th
    partS0201(j,1) = medS0201(j,1)/sum(medS0201);
    partS0203(j,1) = medS0203(j,1)/sum(medS0203);
end
```

ANEXO III: Creación e INSERT en BBDD

```
conn = database('BBDD', '', '');
```

```
sqlquery = ['CREATE TABLE Pisos(ID_pisos varchar, Nombre varchar PRIMARY KEY)'];
execute(conn,sqlquery)
```

```
sqlquery = ['CREATE TABLE Sensor(ID_sensor varchar PRIMARY KEY, Tipo varchar,
Unidades varchar, Estado varchar, Piso varchar, FOREIGN KEY (Piso) REFERENCES Pi-
sos(Nombre))'];
execute(conn,sqlquery)
```

```
sqlquery = ['CREATE TABLE Datos(ID_medida real PRIMARY KEY, Fecha datetime,
ID_sensor varchar, Valor float, Valor_final float, Validado real, Reconstruido
real, FOREIGN KEY (ID_sensor) REFERENCES Sensor(ID_sensor))'];
execute(conn,sqlquery)
```

```
tablename = 'Pisos';
data = table(["2";"3";"4";"5";"6";"7";"8";"9"], ["Sulleva";"Can Colomer";"Zona
Alta";"Can Boada";"Can Poal";"Can Parellada Residencial";"Can Parellada Indus-
trial";"Cam Palet de Vista Alegre"], 'VariableNames', {'ID_pisos', 'Nombre'});
sqlwrite(conn,tablename,data)
```

```
tablename = 'Sensor';
data = ta-
ble(["SU_NV1";"SU_NV2";"CC0201";"CC0202";"CC0210";"CC0211";"CC0213";"CC0214";"CC0
216"], ["Nivel";"Nivel";"Caudal";"Caudal";"Caudal";"Caudal";"Caudal";"Cau-
dal";"Caudal"], ["m";"m";"l/s";"l/s";"l/s";"l/s";"l/s";"l/s";"l/s"], ["Disponi-
ble";"Disponible";"Disponible";"Disponible";"Disponible";"Disponible";"Disponi-
ble";"No Disponible";"Disponible"], ["Sulleva";"Sulleva";"Sulleva";"Sulleva";"Su-
lleva";"Sulleva";"Sulleva";"Sulleva";"Sulleva"], 'VariableNames', {'ID_sensor',
'Tipo', 'Unidades', 'Estado', 'Piso'});
sqlwrite(conn,tablename,data)
```

```
tablename = 'Datos';
data = ta-
ble([idCC0201;idNV1;idNV2;idCC0202;idCC0210;idCC0211;idCC0213;idCC0216], [date;dat
e;date;date;date;date;date;date], [sensCC0201;sensNV1;sensNV2;sensCC0202;sen-
sCC0210;sensCC0211;sensCC0213;sen-
sCC0216], [CC0201;NV1;NV2;CC0202;CC0210;CC0211;CC0213;CC0216], [R_CC0201;R_NV
2;R_CC0202;R_CC0210;R_CC0211;R_CC0213;R_CC0216],
[valCC0201;valNV1;valNV2;valCC0202;valCC0210;valCC0211;valCC0213;valCC0216], [recC
C0201;recNV1;recNV2;recCC0202;recCC0210;recCC0211 recCC0213;recCC0216], 'Variable-
Names', {'ID_medida', 'Fecha', 'ID_sensor', 'Valor', 'Valor_final', 'Validado',
'Reconstruido'});
sqlwrite(conn,tablename,data)
```

