

R.LTWB – SECTION 03

Descarga, procesamiento y análisis de datos
hidroclimatológicos

Completado y extendido de series -
Imputación

<https://github.com/jlgingcivil/R.LTWB.CS2120>

JORGE LUIS GONZALEZ CASTRO
CC: 1032395475

TABLA DE CONTENIDO

1.	Introducción	3
2.	Objetivo General	3
3.	Actividad 1: Procesamiento en software	3
4.	ACTIVIDAD 2: SENSIBILIDAD PARÁMETROS ESTADÍSTICOS	49
5.	ACTIVIDAD 3: ANÁLISIS ESTACIÓN PRECIPITACIÓN 35070110	51
6.	Conclusiones.....	54
7.	Referencias Bibliográficas	55

ÍNDICE DE ILUSTRACIONES

Ilustración 3-1. Descarga Impute.py	3
Ilustración 3-2. Script Outlier.py	5
Ilustración 3-3. Ejecución inicial Impute.py	1
Ilustración 3-4. Resultados Precipitación script Impute.py	6
Ilustración 3-5. Resultados Impute Precipitación	7
Ilustración 3-6. Resultados Impute Precipitación (2)	11
Ilustración 3-7. Resultados Impute Caudal.....	15
Ilustración 3-8. Resultados Impute Caudal (2)	19
Ilustración 3-9. Resultados Impute Evaporación.....	23
Ilustración 3-10. Resultados Impute Evaporación (2)	27
Ilustración 3-11. Registro de datos filtrados Temperatura Máxima	31
Ilustración 3-12. Registro de datos originales Temperatura Máxima.....	32
Ilustración 3-13. Resultados Impute Temperatura Máxima	32
Ilustración 3-14. Resultados Impute Temperatura Máxima (2)	37
Ilustración 3-15. Resultados Impute Temperatura Mínima	41
Ilustración 3-16. Resultados Impute Temperatura Mínima (2)	45
Ilustración 4-1. Resultados Precipitación	49
Ilustración 4-2. Estaciones Parciales Precipitación No Consistentes	50
Ilustración 4-3. Resultados Caudal.....	50
Ilustración 4-4. Estaciones Caudal No Consistentes	50
Ilustración 4-5. Resultados Evaporación.....	50
Ilustración 4-6. Resultados Temperatura Mínima	51
Ilustración 4-7. Resultados Temperatura Máxima	51
Ilustración 5-1. Ejecución script Impute.py Estación 35070110	52
Ilustración 5-2. Resultados método KNNN y MICE Estación 35070110.....	53
Ilustración 5-3. Comparación Resultados Estación 35070110.....	54

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

1. INTRODUCCIÓN

Se continua con curso Balance hidrológico de largo plazo para estimación de caudales medios usando SIG – LWTB con el desarrollo de la sección 3 Descarga, procesamiento y análisis de datos hidroclimatológicos. A continuación, se presenta en cada numeral las actividades realizadas de acuerdo con cada capítulo de la sección de estudio, incluyendo el resumen de actividades, logros alcanzados y capturas de pantalla de los ejercicios realizados. Se ha creado el repositorio <https://github.com/jlgingcivil/R.LTWB.CS2021> para la inclusión de los archivos y documentos de las actividades desarrolladas.

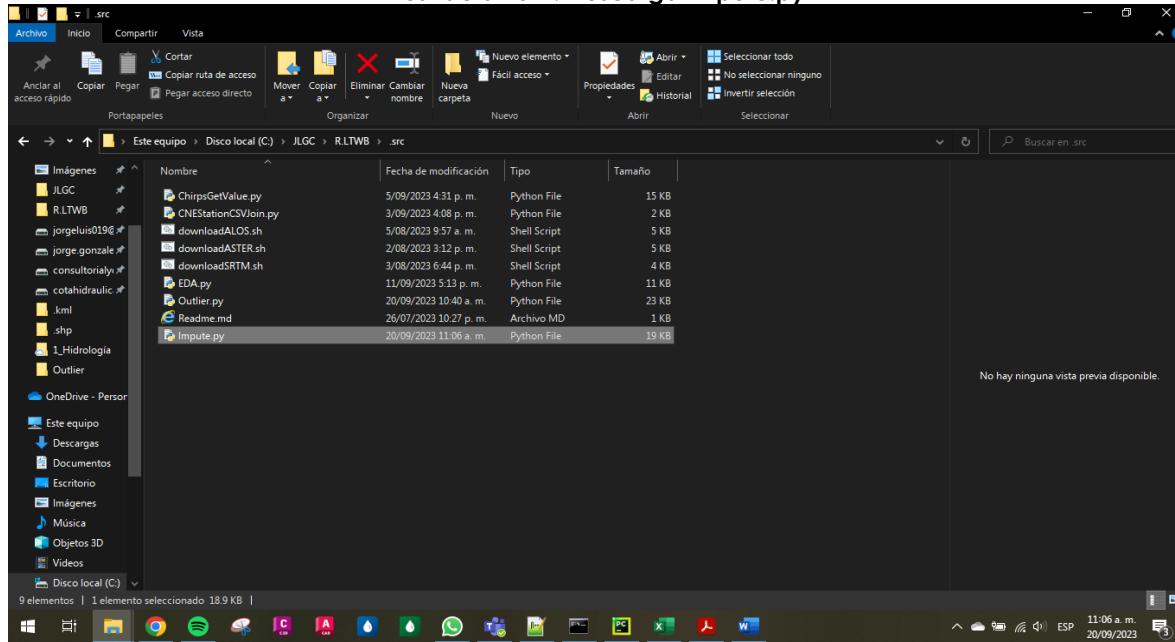
2. OBJETIVO GENERAL

El objetivo general en esta sección es realizar el complemento de datos a partir de 8 métodos estadísticos de las series no excluidas de datos descargadas desde el portal del IDEAM definidas a partir de la actividad de análisis de datos atípicos.

3. ACTIVIDAD 1: PROCESAMIENTO EN SOFTWARE

En primera medida se realiza la descarga del script `Impute.py` y la creación de la carpeta para almacenamiento de archivos.

Ilustración 3-1. Descarga `Impute.py`



Fuente: Elaboración Propia, 2023.

Luego se ajusta el script para que lea la ruta de almacenamiento de archivos, así como la definición de los parámetros tomando el ejemplo de clase para el archivo pivot de precipitación y se inicia con la ejecución de la herramienta; de igual manera se modifica en el script el código de las estaciones a ser excluidas del análisis.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Las estaciones excluidas corresponden al filtro realizado en las actividades anteriores, sin embargo, en los ejercicios preliminares de la presente actividad se evidenció que algunas estaciones con muy pocos datos generan una tendencia lineal en el llenado de datos, por lo que es necesario eliminarlas del conjunto de datos; como criterio se eliminan las estaciones que tengan una cantidad de datos menor equivalente a 10 años, es decir que tengan menos de 120 registros para el caso de la precipitación y de temperatura equivalente a 5 años es decir 1825 datos.

En el caso de evaporación se evidenció que la estación 35035130 presenta una tendencia lineal, pero esto corresponde a que se encuentra en una zona de páramo; se elimina una estación de este parámetro por falta de datos.

Es así que el listado completo de 123 estaciones a excluir es el siguiente:

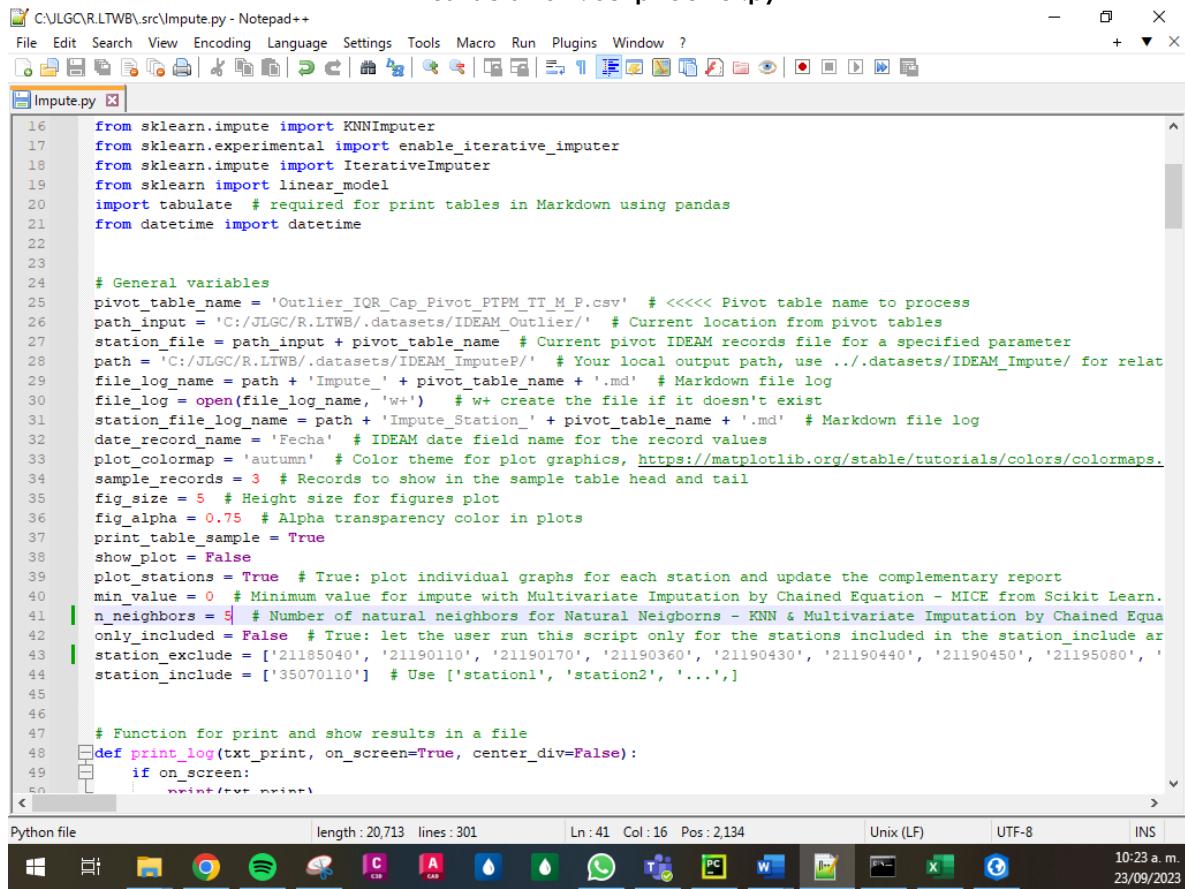
- 21185040
- 21190110
- 21190170
- 21190360
- 21190430
- 21190440
- 21190450
- 21195080
- 21200040
- 21200170
- 21200390
- 21200440
- 21200500
- 21200580
- 21200590
- 21200610
- 21200650
- 21200660
- 21200700
- 21200710
- 21200720
- 21200800
- 21200830
- 21201090
- 21201130
- 21201150
- 21201290
- 21201380
- 21201620
- 21201670
- 21201680
- 21201690
- 21201700
- 21201720
- 21201730
- 21201740
- 21201750
- 21201760
- 21201780
- 21201790
- 21201810
- 21201820
- 21201830
- 21201840
- 21201870
- 21202100
- 21205012
- 21205090
- 21205300
- 21205340
- 21205360
- 21205370
- 21205450
- 21205470
- 21205540
- 21205550
- 21205580
- 21205670
- 21205700
- 21205750
- 21205770
- 21205910
- 21205970
- 21206070
- 21206100
- 21206160
- 21206190
- 21206200
- 21206230
- 21206280
- 21206310
- 21206350
- 21206390
- 21206410
- 21206460
- 21206500
- 21206510
- 21206550
- 21206570
- 21206610
- 21206620
- 21206630
- 21206640
- 21206660
- 21206670
- 21206680
- 21206700

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

- 21206970
- 21230080
- 21255080
- 21255160
- 23010140
- 23060040
- 23060070
- 23060080
- 23060130
- 23060210
- 23060220
- 23060250
- 23060300
- 23060310
- 23060320
- 23065120
- 23065140
- 23065150
- 23065200
- 23125170
- 24010330
- 24010380
- 24010490
- 24015380
- 35020080
- 35020090
- 35020370
- 35025060
- 35035030
- 35035040
- 35035050
- 35060010
- 35060280
- 35065010
- 35067050
- 35070160

Ilustración 3-2. Script Outlier.py



The screenshot shows a Notepad++ window titled "C:\JLGC\R.LTWB\src\Impute.py - Notepad++". The code in the editor is as follows:

```
16 from sklearn.impute import KNNImputer
17 from sklearn.experimental import enable_iterative_imputer
18 from sklearn.impute import IterativeImputer
19 from sklearn import linear_model
20 import tabulate # required for print tables in Markdown using pandas
21 from datetime import datetime
22
23
24 # General variables
25 pivot_table_name = 'Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv' # <<<< Pivot table name to process
26 path_input = 'C:/JLGC/R.LTWB/.datasets/IDEAM_Outlier/' # Current location from pivot tables
27 station_file = path_input + pivot_table_name # Current pivot IDEAM records file for a specified parameter
28 path = 'C:/JLGC/R.LTWB/.datasets/IDEAM_ImputeP/' # Your local output path, use ../../datasets/IDEAM_Impute/ for relat
29 file_log_name = path + 'Impute.' + pivot_table_name + '.md' # Markdown file log
30 file_log = open(file_log_name, 'w+') # w+ create the file if it doesn't exist
31 station_file_log_name = path + 'Impute_Station.' + pivot_table_name + '.md' # Markdown file log
32 date_record_name = 'Fecha' # IDEAM date field name for the record values
33 plot_colormap = 'autumn' # Color theme for plot graphics, https://matplotlib.org/stable/tutorials/colors/colormaps.
34 sample_records = 3 # Records to show in the sample table head and tail
35 fig_size = 5 # Height size for figures plot
36 fig_alpha = 0.75 # Alpha transparency color in plots
37 print_table_sample = True
38 show_plot = False
39 pict_stations = True # True: plot individual graphs for each station and update the complementary report
40 min_value = 0 # Minimum value for impute with Multivariate Imputation by Chained Equation - MICE from Scikit Learn.
41 n_neighbors = 5 # Number of natural neighbors for Natural Neighbors - KNN & Multivariate Imputation by Chained Equa
42 only_included = False # True: let the user run this script only for the stations included in the station_include ar
43 station_exclude = ['21185040', '21190110', '21190170', '21190360', '21190430', '21190440', '21190450', '21195080', '2
44 station_include = ['35070110'] # Use ['station1', 'station2', '....']
45
46
47 # Function for print and show results in a file
48 def print_log(txt_print, on_screen=True, center_div=False):
49     if on_screen:
50         print(txt_print)
```

The status bar at the bottom of the Notepad++ window shows: Python file, length : 20,713 lines : 301, Ln:41 Col:16 Pos:2,134, Unix (LF), UTF-8, INS, 10:23 a.m., 23/09/2023.

Fuente: Elaboración Propia, 2023.

SECTION 03 DESCARGA, PROCESAMIENTO Y ANÁLISIS DE DATOS HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO

CC: 1032395475

CS2021

Ilustración 3-3. Ejecución inicial Impute.py

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO

CC: 1032395475

CS2021

```
Selección Simbolo del sistema
## Method 2 - Imputing with median values
According to this technique, the missing values are imputed using the median value in each feature and the serie has been completed filled.

Imputed file: [Impute_Median_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv](Impute_Median_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)

[R.LTWB]Missingno.Impute_Median_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png

General statistics table - Imputed file
| count | mean | std | min | 25% | 50% | 75% | max | |
|---|---|---|---|---|---|---|---|---|
| 21170020 | 516 | 98.825 | 73.1159 | 0 | 46.675 | 81.95 | 137.25 | 446 |
| 21190030 | 516 | 95.1084 | 60.9247 | 0.4 | 52.05 | 87.2 | 118.225 | 390.356 |
| 21190060 | 516 | 132.145 | 49.4843 | 0 | 132.5 | 132.5 | 132.5 | 342 |
| 21190210 | 516 | 128.939 | 98.9186 | 0 | 44 | 101.9 | 175.025 | 593.066 |
| 21190300 | 516 | 163.839 | 131.813 | 0 | 89.25 | 135.4 | 200.475 | 764 |
| 21190310 | 516 | 142.188 | 87.9782 | 6.6 | 73.875 | 124 | 194.2 | 479 |
| 21190460 | 516 | 92.4956 | 43.9161 | 0 | 75.9 | 85.5 | 98 | 351.402 |
| 21190470 | 516 | 73.1138 | 26.4815 | 0 | 71.75 | 71.75 | 71.75 | 205.3 |
| 21195030 | 516 | 82.6931 | 34.6833 | 5 | 88 | 88 | 80 | 283.9 |
| 21195120 | 516 | 115.022 | 67.8597 | 1.9 | 64.875 | 105.65 | 153.25 | 352.2 |
| 21195140 | 516 | 44.8624 | 36.5643 | 0 | 38.95 | 38.95 | 38.95 | 277.2 |
| 21195190 | 516 | 74.4764 | 46.0321 | 0 | 39.275 | 64.25 | 104.025 | 262.6 |
| 21200000 | 516 | 103.876 | 50.5886 | 4.6 | 78.475 | 97.65 | 117.075 | 378.3 |
| 21200110 | 516 | 63.0475 | 33.779 | 1.7 | 45.475 | 57.7 | 68.4 | 261.7 |
| 21200130 | 516 | 193.683 | 50.7401 | 1 | 78.95 | 99.5 | 119.675 | 367.186 |
| 21200160 | 516 | 93.0062 | 58.5718 | 0 | 53.15 | 85.7 | 123.5 | 369.5 |
| 21200190 | 516 | 116.384 | 57.9018 | 0 | 91.2 | 111.75 | 127.625 | 345.2 |
| 21200200 | 516 | 67.0661 | 35.2478 | 0.1 | 48.775 | 62.8 | 78.25 | 230.1 |
| 21200230 | 516 | 83.1942 | 51.6653 | 1.4 | 61.875 | 73.5 | 87.35 | 345.2 |
| 21200240 | 516 | 188.624 | 51.0053 | 0 | 88.875 | 104.2 | 126.225 | 372.6 |
| 21200260 | 516 | 88.3255 | 44.4351 | 0 | 58.9 | 73.1 | 89.45 | 327.571 |
| 21200270 | 516 | 59.9486 | 35.8596 | 0 | 45.225 | 54.1 | 65.4 | 253 |
| 21200310 | 516 | 73.9953 | 44.6937 | 0.9 | 49.075 | 66.2 | 88.525 | 277.6 |
| 21200320 | 516 | 93.3962 | 46.5612 | 8.3 | 69 | 85.6 | 108.15 | 338.66 |
| 21200340 | 516 | 82.6566 | 41.4669 | 0.9 | 66.7 | 88.7 | 98.775 | 235.1 |
| 21200380 | 516 | 6.53837 | 25.7897 | 0 | 0 | 0 | 0 | 171.2 |
| 21200400 | 516 | 88.5479 | 50.6891 | 5.4 | 66.5 | 80.7 | 100.525 | 341.6 |
| 21200430 | 516 | 53.2843 | 27.7304 | 0 | 44.875 | 53.3 | 55.325 | 264.1 |
| 21200510 | 516 | 50.2847 | 30.6995 | 0 | 34.15 | 47.3 | 53.425 | 197 |
| 21200520 | 516 | 54.7731 | 34.4258 | 0 | 37.5 | 46.6 | 63.175 | 227.3 |
| 21200550 | 516 | 69.6566 | 44.745 | 0 | 45.875 | 63.2 | 88.125 | 297 |
| 21200600 | 516 | 79.314 | 50.7283 | 0 | 58.375 | 69.6 | 96.875 | 359 |
| 21200620 | 516 | 107.691 | 85.3389 | 0 | 48.075 | 89.8 | 139.6 | 454.6 |
```

```
Selección Simbolo del sistema
## Method 3 - Imputing with Last Observation Carried Forward (LOCF) values
According to this technique, the missing values are imputed using the immediate values before it in the time series and the missing values at the start are not filled but the series are completed filled to the end.

Imputed file: [Impute_LOCF_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv](Impute_LOCF_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)

[R.LTWB]Missingno.Impute_LOCF_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png

General statistics table - Imputed file
| count | mean | std | min | 25% | 50% | 75% | max | |
|---|---|---|---|---|---|---|---|---|
| 21170020 | 462 | 101.331 | 77.157 | 0 | 40.6 | 82.55 | 148.825 | 446 |
| 21190030 | 433 | 96.7692 | 66.4604 | 0.4 | 44.2 | 87.2 | 129.8 | 390.356 |
| 21190060 | 492 | 59.3961 | 79.5348 | 0 | 26 | 26 | 44.25 | 342 |
| 21190210 | 516 | 123.246 | 100.822 | 0 | 44 | 102.3 | 180 | 593.066 |
| 21190300 | 516 | 212.025 | 144.871 | 0 | 89.25 | 208.3 | 307 | 764 |
| 21190310 | 509 | 143.846 | 89.5832 | 6.6 | 71.7 | 125.7 | 196.5 | 479 |
| 21190460 | 300 | 97.4952 | 57.1753 | 0 | 57.525 | 85.3 | 132.875 | 351.402 |
| 21190470 | 300 | 39.6557 | 51.5454 | 0 | 0 | 0 | 73.35 | 205.3 |
| 21195030 | 516 | 108.68 | 39.7813 | 5 | 91.85 | 122.2 | 122.2 | 283.9 |
| 21195120 | 513 | 114.25 | 71.1152 | 1.9 | 60.1 | 103.3 | 158.6 | 352.2 |
| 21195140 | 318 | 38.996 | 53.415 | 0 | 0 | 0 | 46.3 | 277.2 |
| 21195190 | 516 | 74.8731 | 46.2783 | 0 | 39.275 | 64.65 | 104.125 | 262.6 |
| 21200000 | 516 | 161.335 | 108.561 | 4.6 | 72.275 | 123.35 | 317.1 | 378.3 |
| 21200110 | 516 | 84.4448 | 47.6087 | 1.7 | 42.65 | 72.15 | 138.1 | 261.7 |
| 21200130 | 516 | 137.616 | 75.4277 | 1 | 76.1 | 122.1 | 228.3 | 367.186 |
| 21200160 | 453 | 94.1951 | 62.7549 | 0 | 46.3 | 85.8 | 130.7 | 369.5 |
| 21200190 | 516 | 108.717 | 64.3636 | 0 | 73 | 78.7 | 140.575 | 345.2 |
| 21200200 | 516 | 88.9929 | 40.4933 | 0.1 | 48.775 | 80.8 | 113.9 | 230.1 |
| 21200230 | 516 | 122.731 | 70.3191 | 1.4 | 59.4 | 126.65 | 195.7 | 345.2 |
| 21200240 | 516 | 136.262 | 66.5759 | 0 | 80.875 | 134.2 | 206.8 | 372.6 |
| 21200260 | 516 | 91.3499 | 49.1885 | 0 | 55.6 | 102.55 | 106.4 | 327.571 |
| 21200270 | 515 | 62.7561 | 37.3428 | 0 | 36.95 | 67.4 | 67.4 | 253 |
| 21200310 | 516 | 91.4266 | 51.2795 | 0.9 | 46.375 | 91.55 | 132.4 | 277.6 |
| 21200320 | 516 | 126.456 | 66.6687 | 8.3 | 68.95 | 118.85 | 204.5 | 338.66 |
| 21200340 | 516 | 96.7233 | 46.768 | 0.9 | 60.7 | 103.5 | 132.2 | 235.1 |
| 21200380 | 516 | 40.331 | 29.4714 | 0 | 0 | 53 | 53 | 171.2 |
| 21200400 | 516 | 101.402 | 52.9634 | 5.4 | 60.45 | 104.8 | 127.5 | 341.6 |
| 21200430 | 516 | 46.4734 | 30.9029 | 0 | 30.575 | 42.6 | 56.15 | 204.1 |
| 21200510 | 516 | 35.7976 | 38.9155 | 0 | 0 | 25.45 | 57.125 | 197 |
| 21200520 | 516 | 85.0769 | 57.107 | 0 | 37.075 | 68.35 | 159.8 | 227.3 |
| 21200550 | 516 | 85.438 | 47.8372 | 0 | 45.875 | 106.25 | 118 | 297 |
| 21200600 | 516 | 60.3965 | 63.2398 | 0 | 0 | 45.9 | 99 | 359 |
```

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO

CC: 1032395475

CS2021

```
## Method 4 - Imputing with Next Observation Carried Backward (NOCB) values
According to this technique, the missing values are imputed using the immediate values after it in the time series and the missing values at the end are not filled but the series are completed fillet to the start.

Imputed file: [Impute_NOCB_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv](Impute_NOCB_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)

!R.LTWB](Impute_NOCB_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

![R.LTWB](Missingno_Impute_NOCB_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)
General statistics table - Imputed file
```

	count	mean	std	min	25%	50%	75%	max
21170020	516	91.6349	77.6988	0	25	72.75	138.25	446
21190030	516	89.4854	62.9717	0.4	51.875	71.3	118.225	390.356
21190060	180	147.016	93.0233	0	69.375	152.5	232	342
21190210	516	120.595	100.476	0	42	101.1	176.25	593.066
21190300	388	173.639	151.2	0	66.2	130.5	236	764
21190310	516	142.883	93.1608	6.6	68.8	123.4	196.6	479
21190460	516	65.3344	58.4929	0	19.9	43.75	98.475	351.402
21190470	372	31.9804	48.8612	0	0	0	56.9	205.3
21195030	228	105.554	74.3875	5	49.625	90.15	142.325	283.9
21195120	516	116.466	70.1289	1.9	61.15	106.5	159.7	352.2
21195140	372	74.589	45.2677	0	46.075	90.5	96.5	277.2
21195190	516	74.9725	46.3653	0	38.85	64.55	105.975	262.6
21200080	376	105.009	60.4283	4.6	62.825	94.55	133.825	378.3
21200110	377	67.1279	42.6686	1.7	35.6	57.8	88.2	261.7
21200130	377	105.686	59.4306	1	62.9	100.7	135.2	367.186
21200160	516	89.4545	60.5758	0	51.3	69.4	124.475	369.5
21200190	374	116.827	72.7211	0	61.6	110.8	157.275	345.2
21200200	377	68.6448	41.1392	0.1	39.6	62.8	89.8	230.1
21200230	375	76.1739	67.2061	1.4	25.3	59.9	106.3	345.2
21200240	377	110.255	59.6162	0	69.7	104.2	145.4	372.6
21200260	384	79.5887	54.9381	0	38.5	69.7	110.675	327.571
21200270	345	59.9901	45.5673	0	21.8	49.3	88	253
21200310	377	76.7684	52.279	0.9	35.6	65.3	186.4	277.6
21200320	377	98.4941	55.5566	8.3	57.6	89.8	129.2	338.66
21200340	375	83.384	48.566	0.9	45.15	88.7	113.25	235.1
21200370	187	18.0047	46.1079	0	0	0	0	100.2
21200400	377	89.1798	68.686	5.4	44.1	76.8	115.8	341.6
21200430	345	53.4441	36.45	0	25.3	51.3	72.5	284.1
21200510	372	58.9258	36.8057	0	23.35	48.45	71.425	197
21200520	376	57.1189	46.4284	0	27.875	48.5	79.225	227.3
21200550	342	72.9415	54.6957	0	33.35	63.2	104	297

```
## Method 5 - Impute missing values with Linear Interpolation values
According to this technique, the missing values are imputed using the linear interpolation between knowing pair values in the time series and the missing values at the start are not filled but the series are completed fillet to the end.

Imputed file: [Impute_InterpolateLinear_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv](Impute_InterpolateLinear_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)

!R.LTWB](Missingno_Impute_InterpolateLinear_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

![R.LTWB](General_statistics_table - Imputed file)
General statistics table - Imputed file
```

	count	mean	std	min	25%	50%	75%	max
21170020	462	100.962	77.0481	0	40.6	82.1	148	446
21190030	433	96.6241	66.4127	0.4	44.2	87.2	128.4	390.356
21190060	492	59.3961	79.5348	0	26	26	44.25	342
21190210	516	121.92	99.4334	0	44	102.1	177.4	593.066
21190300	516	209.373	143.202	0	89.25	207	307	764
21190310	509	144.189	99.3353	6.6	71.7	124.9	196.9	479
21190460	300	97.7262	57.1634	0	57.975	85.6	133.5	351.402
21190470	300	39.6557	51.5454	0	0	0	73.35	205.3
21195030	516	111.762	41.7737	5	103.825	122.2	122.2	283.9
21195120	513	115.619	69.6248	1.9	62.1	105.6	156.7	352.2
21195140	318	30.996	53.415	0	0	0	46.3	277.2
21195190	516	74.9228	46.1358	0	38.9	64.65	104.375	262.6
21200080	516	161.944	107.869	4.6	72.375	122.7	217.1	378.3
21200110	516	85.3455	46.984	1.7	43.975	74.45	138.1	261.7
21200130	516	138.166	74.7355	1	78.375	122.1	228.3	367.186
21200160	453	94.1408	62.6447	0	47.1	85.7	130.7	369.5
21200190	516	106.742	63.4289	0	73	75.6	138.975	345.2
21200200	516	80.8693	40.4849	0.1	48.775	80.8	113.9	230.1
21200230	516	115.783	72.4663	1.4	50.65	100.692	195.7	345.2
21200240	516	136.262	66.5759	0	80.875	134.2	206.8	372.6
21200260	516	88.9897	46.7293	0	53.9417	96.75	106.4	327.571
21200270	515	62.6237	37.2879	0	36.95	67.4	67.4	253
21200310	516	91.5095	51.0653	0.9	47.2	91.55	132.4	277.6
21200320	516	126.753	66.6989	8.3	68.65	118.85	204.5	338.66
21200340	516	96.7233	46.768	0.9	60.7	103.5	132.2	235.1
21200380	516	40.331	29.4714	0	0	53	53	171.2
21200400	516	100.452	53.5094	5.4	57.3962	104.8	127.5	341.6
21200430	516	48.1619	29.7126	0	35.675	42.6	59.5	204.1
21200510	516	36.2108	38.5456	0	0	27.75	54.9406	197
21200520	516	85.0275	57.0834	0	37.5	67.4	159.8	227.3

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO

CC: 1032395475

CS2021

```
## Method 6 - Impute missing values with Exponential (Weighted) Moving Average - EWM = 3
According to this technique, the missing values are imputed using the moving average values in the time series and the missing values at the start are not filled but the series are completed fillet to the end.

Imputed file: [Impute_MeanEWM_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv](Impute_MeanEWM_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)

!R.LTWB](Impute_MeanEWM_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

!R.LTWB](Missingno.Impute_MeanEWM_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

General statistics table - Imputed file
```

	count	mean	std	min	25%	50%	75%	max
21170020	462	100.93	77.856	0	40.6	81.95	148	446
21190030	433	96.6499	66.4112	0.4	44.2	87.2	128.4	390.356
21190060	492	133.142	50.6914	0	133.985	133.985	133.985	342
21190210	516	122.972	99.6736	0	44	102.75	180	593.066
21190300	516	196.337	136.515	0	89.25	208.3	245.577	764
21190310	509	143.042	88.6675	6.6	72.9	124.7	196.4	479
21190460	300	93.5232	57.1058	0	57.975	85.475	132.875	351.402
21190470	300	63.6309	37.0655	0	49.9484	49.9484	73.35	205.3
21195030	516	101.222	36.7228	5	88.343	112.161	112.161	283.9
21195120	513	115.443	68.3819	1.9	63.1	105.6	154.6	352.2
21195140	318	51.7649	45.693	0	33.4	45.9311	46.3	277.2
21195190	516	74.7529	46.0234	0	39.275	65	104.025	262.6
21200080	516	131.692	66.4323	4.6	74.375	122.15	201.176	378.3
21200110	516	83.4978	45.1586	1.7	45.475	72.15	132.851	261.7
21200130	516	124.361	60.1721	1	76.1	122.1	177.366	367.186
21200160	453	94.1847	62.4623	0	48	85.9	130.1	369.5
21200190	516	114.415	58.6881	0	87.2141	101.872	133.2	345.2
21200200	516	82.7152	42.6997	0.1	48.775	86.6	120.814	230.1
21200230	516	114.25	67.5482	1.4	60.05	93.4	185.442	345.2
21200240	516	135.814	66.1022	0	80.875	134.2	205.135	372.6
21200260	516	98.6632	56.314	0	58.9	101.55	138.997	327.571
21200270	516	66.8835	36.9577	0	38.8763	77.4822	77.4822	253
21200310	516	95.8349	54.6985	0.9	47.2	91.55	147.927	277.6
21200320	516	122.838	63.6202	8.3	68.95	101.9	194.7	388.66
21200340	516	85.9197	41.5306	0	60.7	92.6638	98.775	238.0
21200380	516	38.1205	28.5558	0	0	49.5472	49.5472	171.2
21200400	516	114.365	62.7337	5.4	60.45	104.8	175.704	341.6
21200430	516	54.9037	29.5435	0	32.323	63.2774	63.2774	204.1
21200510	516	45.9881	31.8453	0	32.9901	54.5972	197	

```
## Method 7 - Impute missing values with Natural Neighbors - KNN = 5 Imputer from Scikit Learn
According to this technique, the missing values are imputed using the natural neighbors values and the serie has been completed filled. More information in https://scikit-learn.org/stable/modules/generated/sklearn.impute.KNNImputer.html

Imputer = KNNImputer(n_neighbors=n_neighbors, weights='uniform', metric='euclidean')

Imputed file: [Impute_KNN_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv](Impute_KNN_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)

!R.LTWB](Impute_KNN_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

!R.LTWB](Missingno.Impute_KNN_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

General statistics table - Imputed file
```

	count	mean	std	min	25%	50%	75%	max
21170020	516	100.325	75.1493	0	41	83.1	148	446
21190030	516	95.5934	63.6489	0.4	45.4	86.85	127.625	390.356
21190060	516	128.899	68.2985	0	73.55	124.5	176.65	342
21190210	516	121.328	99.773	0	43.225	101.1	177.3	593.066
21190300	516	171.429	141.722	0	69.3	139.45	233.4	764
21190310	516	142.512	89.0214	6.6	72.6	123.75	196.3	479
21190460	516	91.2105	50.6585	0	57.35	80.53	118.212	351.402
21190470	516	70.5093	37.9981	0	41.905	64.17	93.6	205.3
21195030	516	85.1333	50.7221	5	42.115	78.68	122.21	283.9
21195120	516	113.963	69.3007	1.9	60.1	104.3	155.725	352.2
21195140	516	55.7023	46.0254	0	21.12	43.42	83.345	277.2
21195190	516	74.2368	46.2252	0	38.625	63.6	104.025	262.6
21200080	516	107.098	55.8372	4.6	66	101.38	135.3	378.3
21200110	516	63.8069	37.2142	1.7	38.17	56.84	82.25	261.7
21200130	516	104.885	56.2799	1	64.51	101.2	133.655	367.186
21200160	516	91.4979	61.4419	0	45.675	82.2	126.725	369.5
21200190	516	119.793	69.484	0	68.245	110.8	156.825	345.2
21200200	516	68.5922	39.6082	0.1	40.875	63.65	91.42	230.1
21200230	516	85.3989	58.1572	1.4	42.63	69.37	113.625	345.2
21200240	516	110.479	56.5468	0	70.38	107.12	146.375	372.6
21200260	516	82.4448	50.1714	0	48.725	72.85	108.675	327.571
21200270	516	64.0074	42.7226	0	32.48	55.25	91.06	253
21200310	516	75.9345	49.6152	0.9	38.11	65.59	104.195	277.6
21200320	516	95.9833	51.2799	8.3	59.7	87.06	125.3	338.66
21200340	516	83.5812	47.2881	0.9	46.175	80.57	113.185	235.1
21200380	516	15.5349	27.8828	0	0	0	19.32	171.2

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO

CC: 1032395475

CS2021

```
## Seleccionar Símbolo del sistema
warnings.warn()

## Method 8 - Impute missing values with Multivariate Imputation by Chained Equation - MICE from Scikit Learn
According to this technique, the missing values are imputed using MICE values and the serie has been completed filled. More information in https://scikit-learn.org/stable/modules/generated/sklearn.impute.IterativeImputer.html

Imputer = IterativeImputer(estimator=BayesianRidge(), min_value=0, n_nearest_features=5)

Imputed file: [Impute_MICE_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv](Impute_MICE_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)
[R.LTMB](Impute_MICE_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)
![R.LTMB](Missingno_Impute_MICE_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

General statistics table - Imputed file
   count | mean | std | min | 25% | 50% | 75% | max |
-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:
  21170020 | 516 | 100.39 | 76.1207 | 0 | 41 | 81.45 | 147.85 | 446
  21190030 | 516 | 95.412 | 63.0788 | 0.4 | 47.0715 | 85 | 124.225 | 390.356
  21190060 | 516 | 137.495 | 60.6393 | 0 | 104.551 | 134.671 | 173.126 | 342
  21190210 | 516 | 121.729 | 69.3728 | 0 | 44 | 102.05 | 177 | 593.066
  21190300 | 516 | 174.345 | 137.55 | 0 | 81.8391 | 142.5 | 232 | 764
  21190310 | 516 | 142.421 | 88.6113 | 6.6 | 72.6 | 123.45 | 196.3 | 479
  21190460 | 516 | 93.5065 | 49.8363 | 0 | 58.9958 | 84 | 119.747 | 351.402
  21190470 | 516 | 72.6005 | 36.9518 | 0 | 46.8994 | 63.9185 | 93.5002 | 208.907
  21195030 | 516 | 91.3499 | 45.3852 | 5 | 61.1883 | 85.2385 | 117.057 | 283.9
  21195120 | 516 | 114.631 | 68.9886 | 1.9 | 62.2024 | 105.05 | 155.725 | 352.2
  21195140 | 516 | 54.6626 | 39.3168 | 0 | 33.7933 | 49.6947 | 68.7387 | 277.2
  21195190 | 516 | 74.3688 | 46.118 | 0 | 39.275 | 63.6 | 104.025 | 262.6
  21200080 | 516 | 106.832 | 52.7712 | 4.6 | 74.775 | 98.7066 | 138.05 | 378.3
  21200110 | 516 | 66.1288 | 35.9659 | 1.7 | 41.6181 | 68.6 | 83.75 | 261.7
  21200130 | 516 | 187.645 | 55.579 | 1 | 69 | 102.5 | 136.151 | 367.186
  21200160 | 516 | 91.5286 | 61.1882 | 0 | 45.5546 | 83.6675 | 126.926 | 369.5
  21200190 | 516 | 127.695 | 72.1703 | 0 | 72.8428 | 118.3 | 171.304 | 345.2
  21200200 | 516 | 68.6192 | 36.1131 | 0.1 | 48.2 | 64.2901 | 86.0003 | 230.1
  21200230 | 516 | 92.598 | 57.3154 | 1.4 | 55.2381 | 81.9 | 128.3 | 359.493
  21200240 | 516 | 111.112 | 53.8435 | 0 | 76.4 | 105.7 | 143.9 | 372.6
  21200260 | 516 | 83.4369 | 49.2965 | 0 | 50.1927 | 73.55 | 109.723 | 327.571
  21200270 | 516 | 64.0810 | 41.1015 | 0 | 33.3113 | 59.4 | 90.0551 | 259.6
  21200310 | 516 | 76.8786 | 48.1848 | 0.9 | 42.0114 | 68.405 | 102.05 | 277.6
  21200320 | 516 | 97.3295 | 49.2507 | 8.3 | 63.4969 | 91.4536 | 123.696 | 338.66
  21200340 | 516 | 83.0897 | 45.4621 | 0.9 | 47.4521 | 80.2216 | 112.346 | 235.1
  21200380 | 516 | 18.4138 | 25.157 | 0 | 0 | 14.0411 | 22.6452 | 171.2
  21200400 | 516 | 92.6533 | 53.5494 | 5.4 | 55.775 | 83.5234 | 115.761 | 341.6
```

Windows Taskbar: Seleccionar Símbolo del sistema, warnings.warn(), R.LTMB, Missingno_Impute_MICE_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png, ESP 10:38 a.m. 23/09/2023


```
## Seleccionar Símbolo del sistema
warnings.warn()

## Method 8 - Impute missing values with Multivariate Imputation by Chained Equation - MICE from Scikit Learn
According to this technique, the missing values are imputed using MICE values and the serie has been completed filled. More information in https://scikit-learn.org/stable/modules/generated/sklearn.impute.IterativeImputer.html

Imputer = IterativeImputer(estimator=BayesianRidge(), min_value=0, n_nearest_features=5)

Imputed file: [Impute_Station_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.md](Impute_Station_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv)
[R.LTMB](Impute_Station_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)
![R.LTMB](Missingno_Impute_Station_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png)

General statistics table - Imputed file
   count | mean | std | min | 25% | 50% | 75% | max |
-----:-----:-----:-----:-----:-----:-----:-----:-----:-----:
  35020420 | 516 | 82.1483 | 48.8884 | 0 | 45.7834 | 75.1 | 112.475 | 305.5
  35020500 | 516 | 109.745 | 66.3028 | 0 | 57.15 | 103.08 | 150.725 | 341.3
  35020510 | 516 | 220.193 | 143.37 | 0 | 97.9926 | 208.217 | 306.273 | 737.5
  35030020 | 516 | 58.52 | 20.3981 | 0 | 29.75 | 500 | 699.25 | 1091
  35030080 | 516 | 77.0154 | 49.6907 | 0 | 34.45 | 74.8334 | 104.99 | 245.9
  35030110 | 516 | 155.81 | 94.8749 | 0 | 85.1991 | 114.184 | 212.173 | 522.2
  35030120 | 516 | 151.146 | 93.8199 | 0 | 70.5456 | 143.73 | 204.125 | 566.9
  35030140 | 516 | 145.839 | 86.7801 | 0 | 82.8428 | 133.639 | 194.8 | 553.4
  35030160 | 516 | 124.788 | 76.6748 | 0.8 | 64.95 | 112.575 | 174.483 | 448.1
  35030170 | 516 | 99.517 | 60.3841 | 0 | 41.3304 | 80.0822 | 131.023 | 317.8
  35030180 | 516 | 130.488 | 85.6541 | 0 | 66.7 | 120.4 | 181.369 | 585.4
  35030210 | 516 | 148.022 | 96.9875 | 0 | 72.8852 | 133.435 | 208.678 | 512.4
  35030230 | 516 | 185.084 | 99.1923 | 2.5 | 105.806 | 176.8 | 249.911 | 522.2
  35030250 | 516 | 179.016 | 98.755 | 0 | 132.532 | 169.948 | 215.821 | 599.9
  35030260 | 516 | 88.7136 | 56.073 | 0 | 50.1663 | 77.0084 | 114.445 | 341.6
  35040010 | 516 | 552.369 | 324.344 | 0 | 309.5 | 549.085 | 755.475 | 1726
  35060020 | 516 | 113.769 | 81.8992 | 0 | 55.1 | 98.25 | 158.625 | 477.7
  35060050 | 516 | 183.361 | 69.5564 | 0 | 48.875 | 93.25 | 144.175 | 378
  35060090 | 516 | 185.446 | 115.136 | 0 | 94.775 | 172.95 | 261.425 | 698.557
  35060100 | 516 | 177.235 | 108.052 | 0 | 92.225 | 169 | 249.9 | 514.3
  35060120 | 516 | 245.679 | 146.295 | 0.7 | 135 | 231.65 | 343.66 | 919
  35060130 | 516 | 268.444 | 167.843 | 0 | 139.5 | 255.1 | 380.525 | 974
  35060150 | 516 | 277.559 | 176.38 | 0 | 144.769 | 259.35 | 388.225 | 1078.65
  35060160 | 516 | 145.163 | 84.9871 | 2 | 78.75 | 136.3 | 194.775 | 421.8
  35060170 | 516 | 143.682 | 96.9798 | 0 | 64.525 | 136.85 | 213.1 | 534
  35060180 | 516 | 167.156 | 107.637 | 0 | 83.85 | 154.7 | 233.76 | 573
  35060200 | 516 | 149.797 | 93.2987 | 0.8 | 83.2 | 133.55 | 203.7 | 493.6
  35060210 | 516 | 161.098 | 98.5496 | 0.1 | 85.4827 | 149.7 | 228.425 | 545
  35060230 | 516 | 179.252 | 114.601 | 0 | 89.175 | 168.6 | 252 | 667
  35060240 | 516 | 106.012 | 75.4851 | 0 | 45.8306 | 95.55 | 149.469 | 435
  35070110 | 516 | 97.5742 | 60.6355 | 0 | 48.3769 | 93.2707 | 144.966 | 283.9
  35070230 | 516 | 94.4347 | 60.5767 | 0 | 45.4 | 87.8 | 134.15 | 310.4
  35070480 | 516 | 95.0005 | 59.9649 | 0.3 | 46.0825 | 91.4 | 134.75 | 280.731
  35070490 | 516 | 189.68 | 120.207 | 0 | 93.625 | 175.5 | 264.791 | 603.7
```

Windows Taskbar: Seleccionar Símbolo del sistema, warnings.warn(), R.LTMB, Missingno_Impute_Station_Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv.png, ESP 10:38 a.m. 23/09/2023

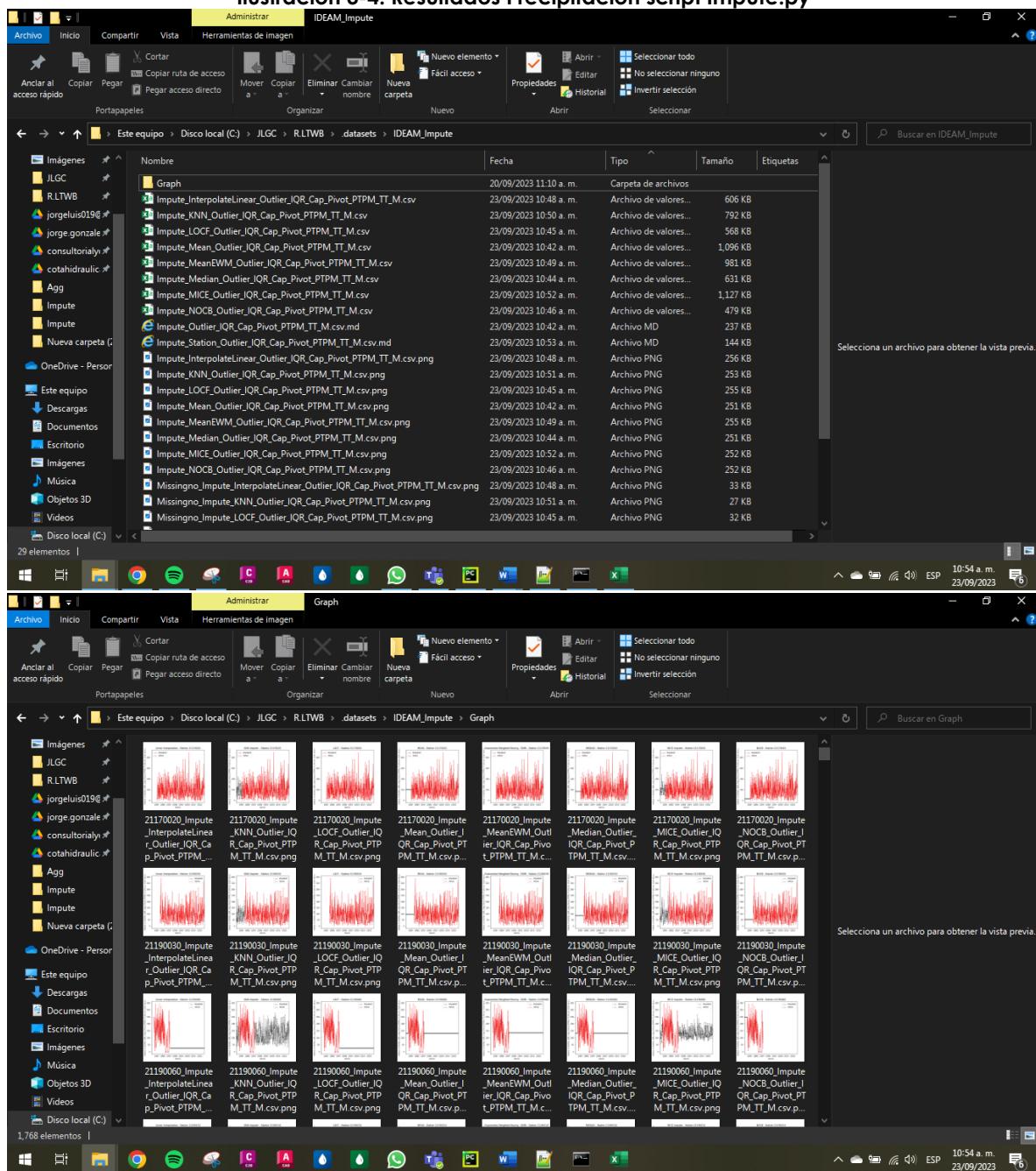
Fuente: Elaboración Propia, 2023.

Se verificó que en la carpeta ./datasets/IDEAM_Impute se almacenaron los resultados del script en cuanto a gráficas, tablas y archivo de visualización en formato Markdown.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
 CC: 1032395475
 CS2021

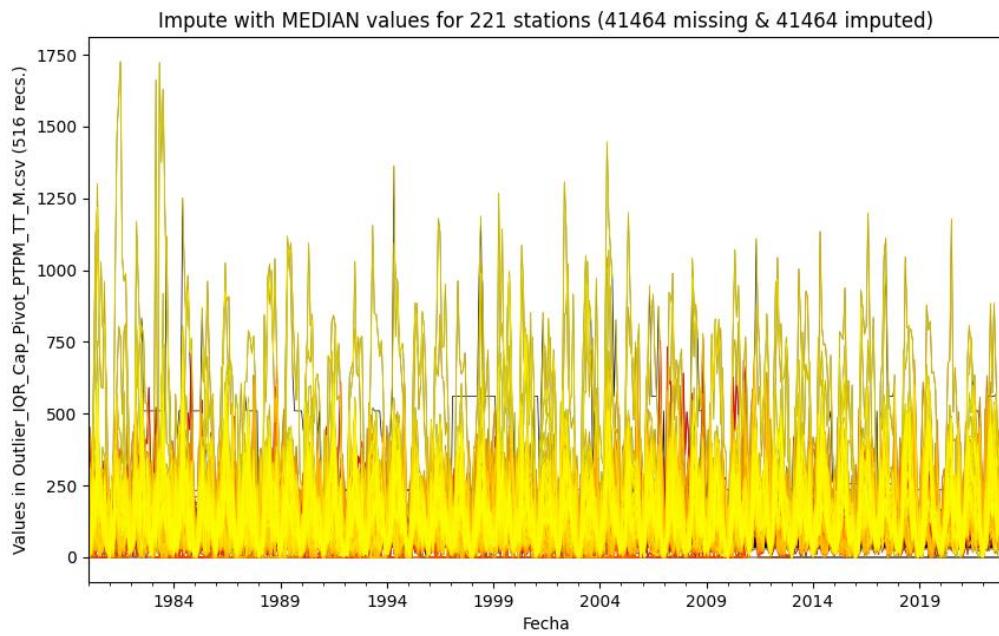
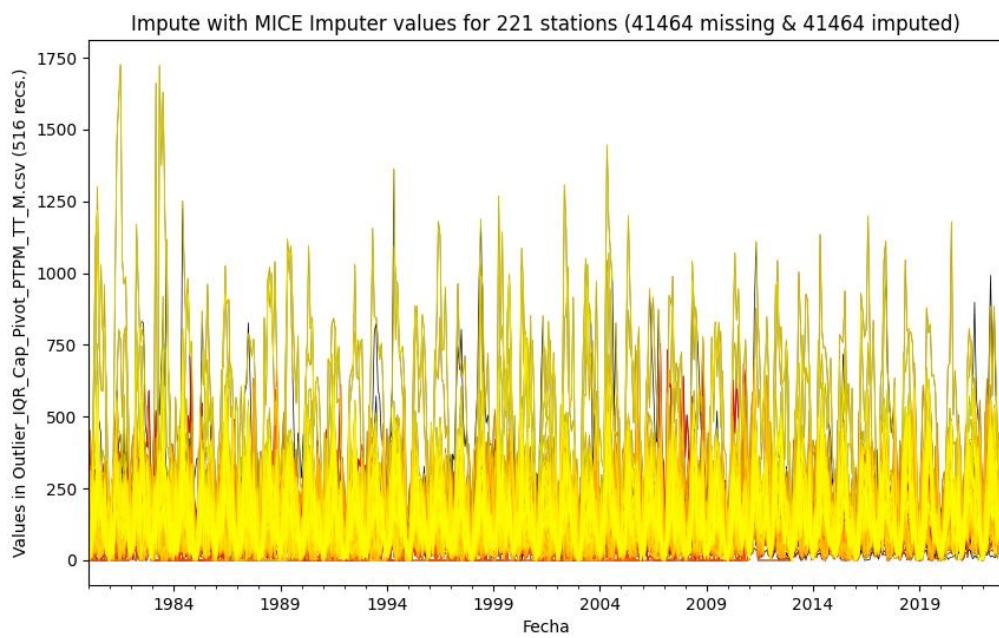
Ilustración 3-4. Resultados Precipitación script Impute.py



Fuente: Elaboración Propia, 2023.

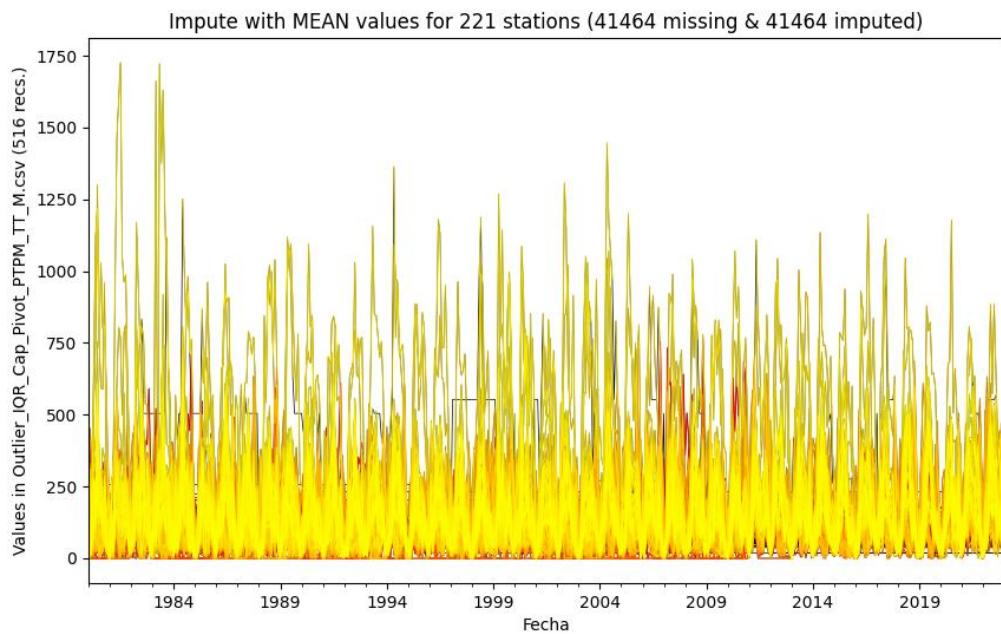
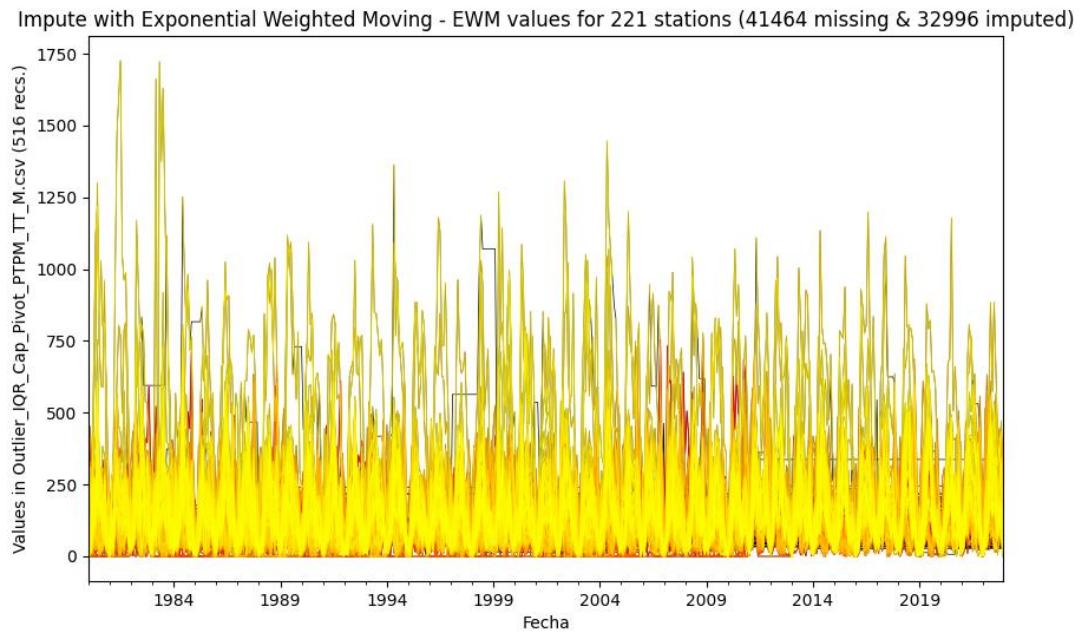
Las gráficas de resultados por cada método se presentan a continuación.

Ilustración 3-5. Resultados Impute Precipitación



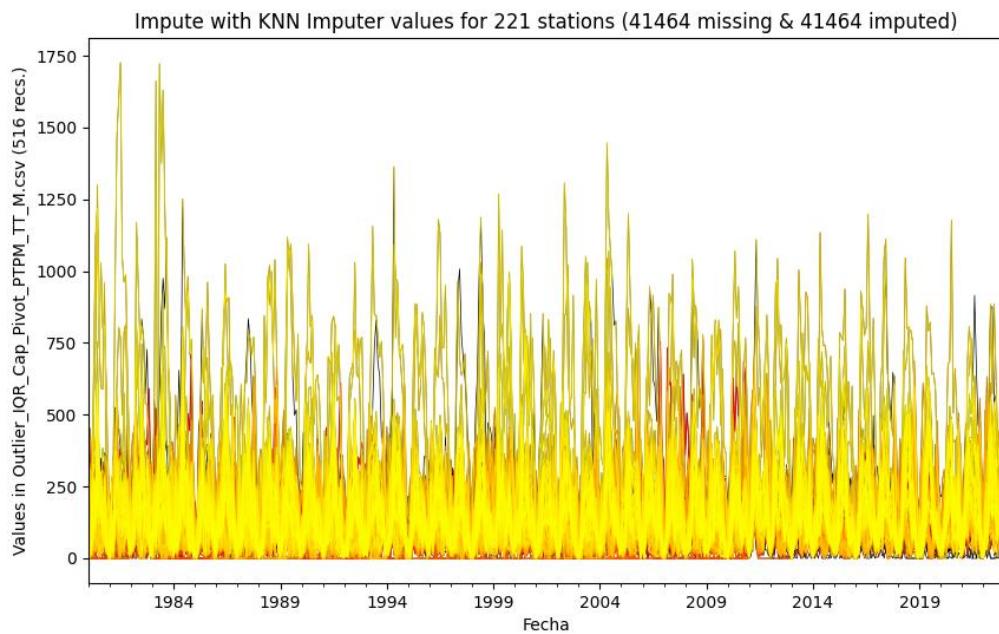
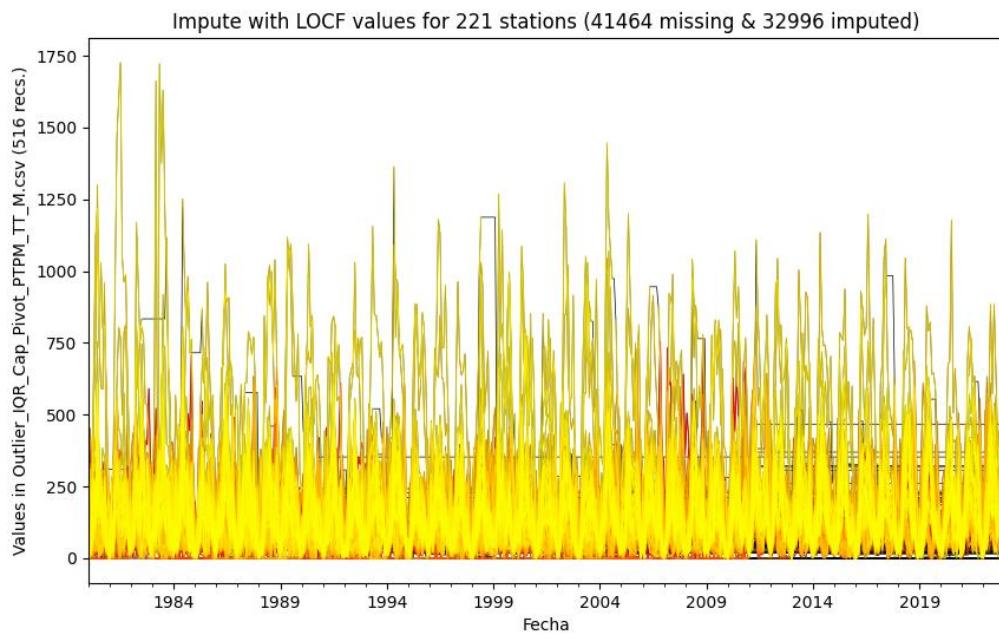
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



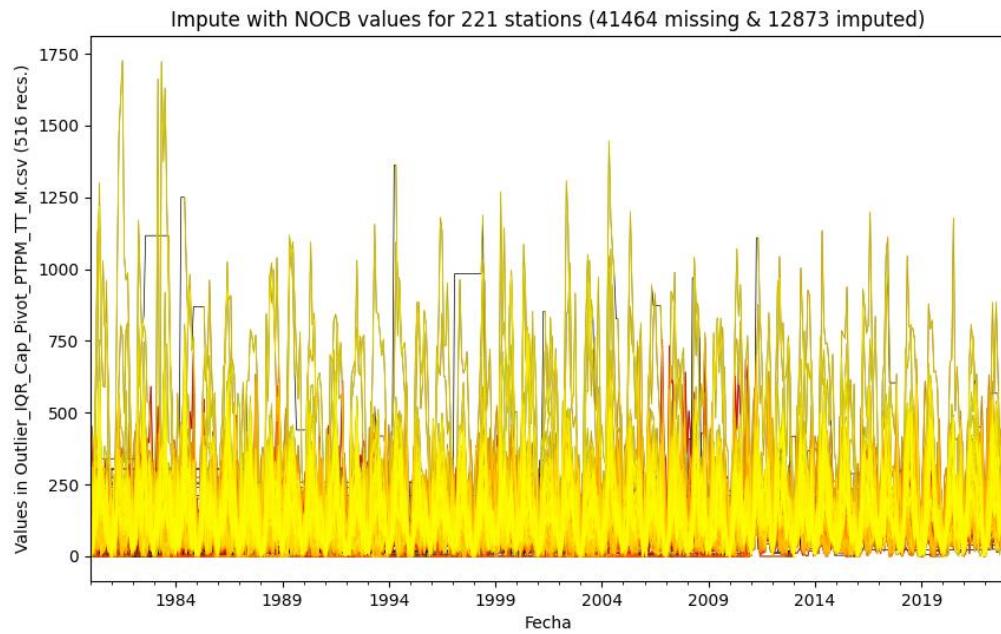
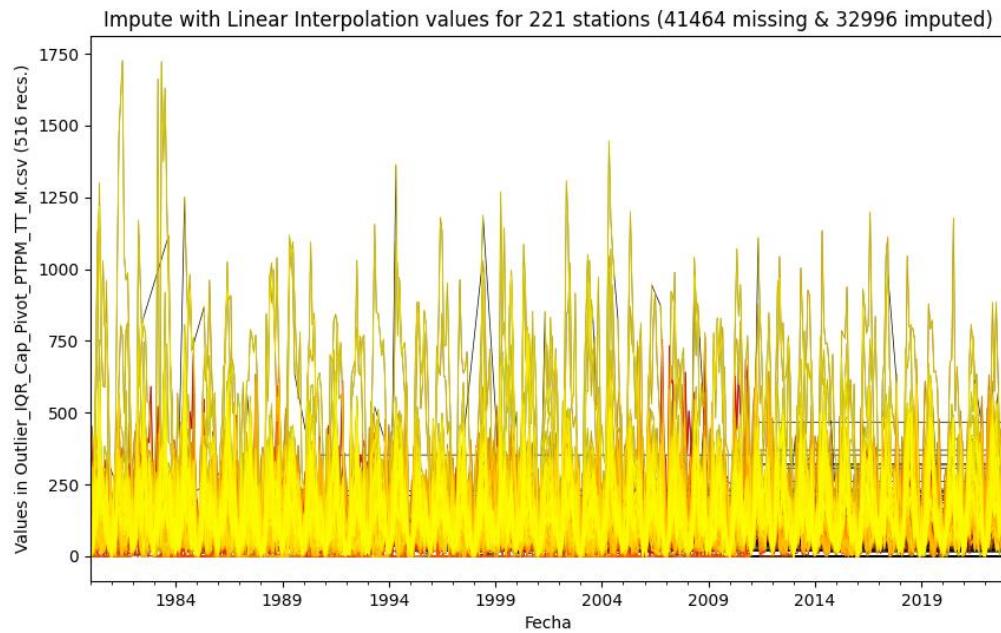
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



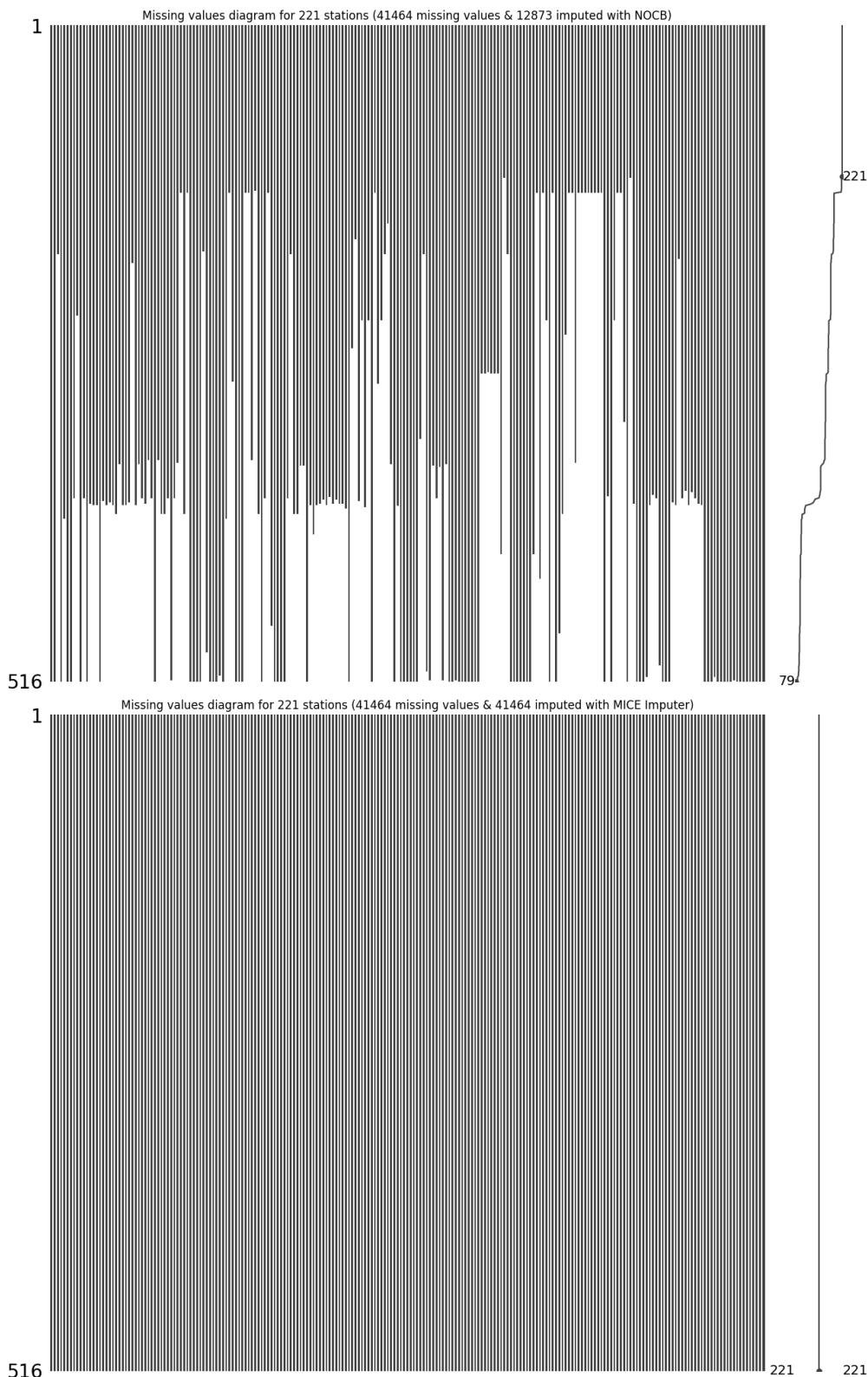
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



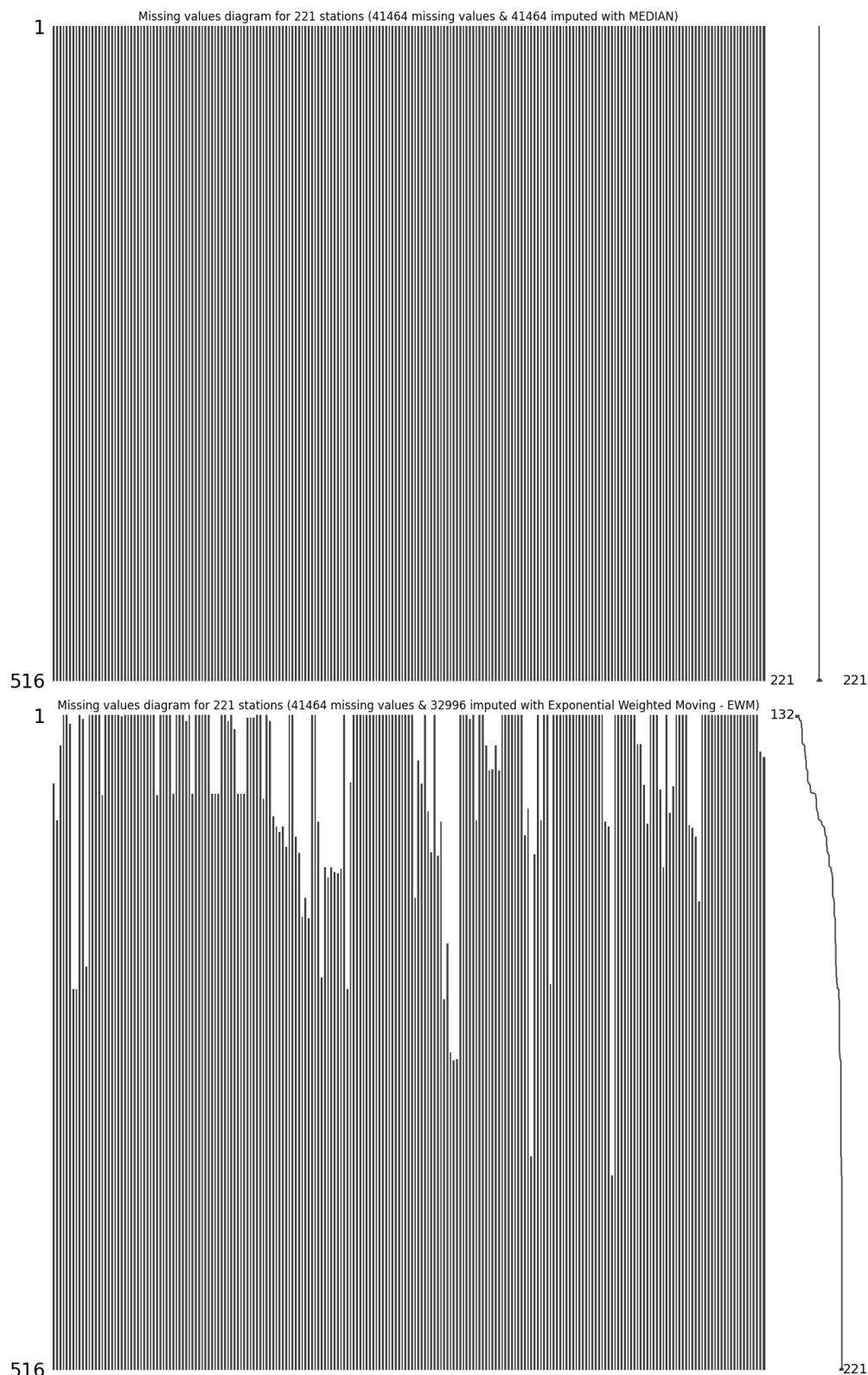
Fuente: Elaboración Propia, 2023.

Ilustración 3-6. Resultados Impute Precipitación (2)



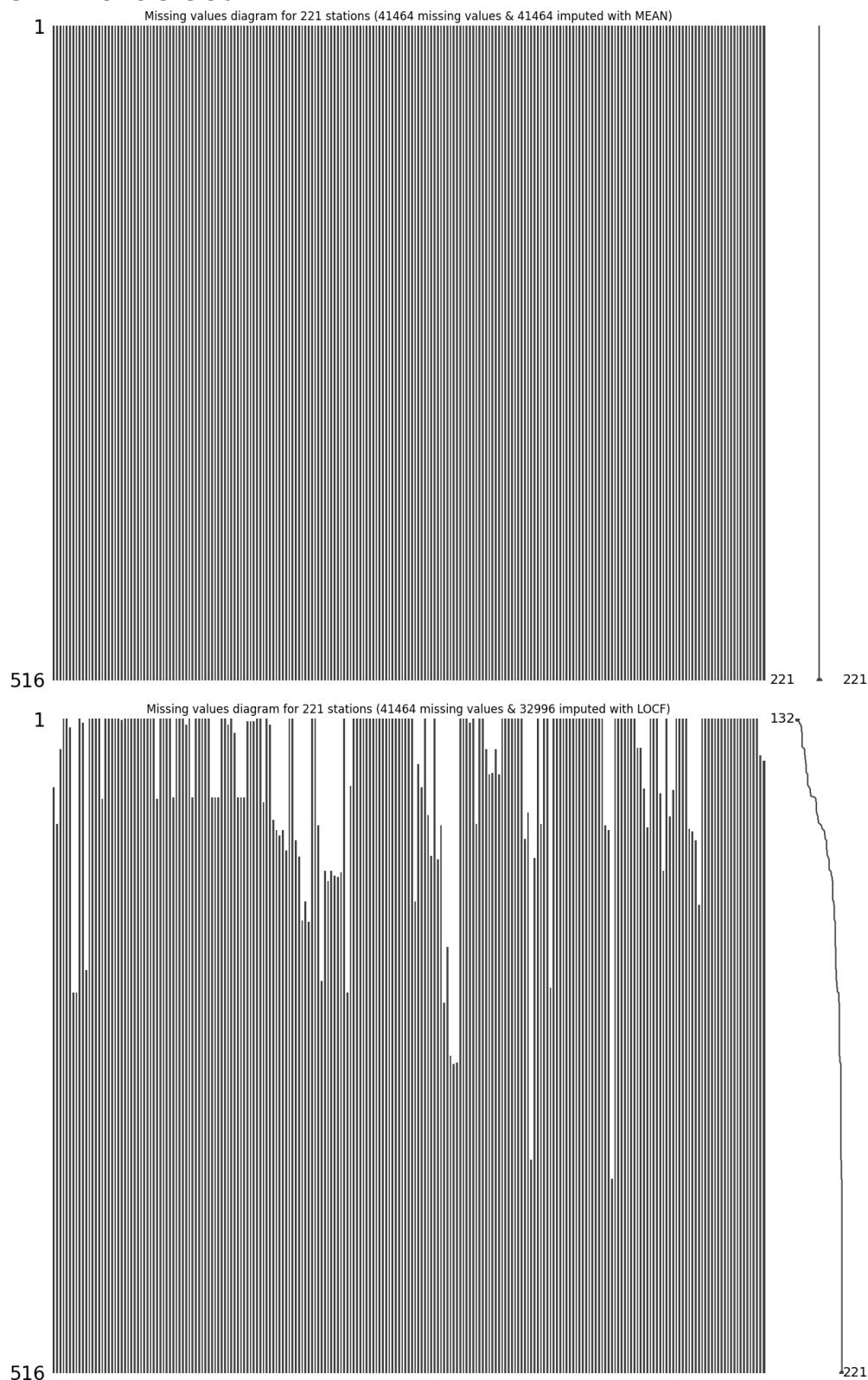
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



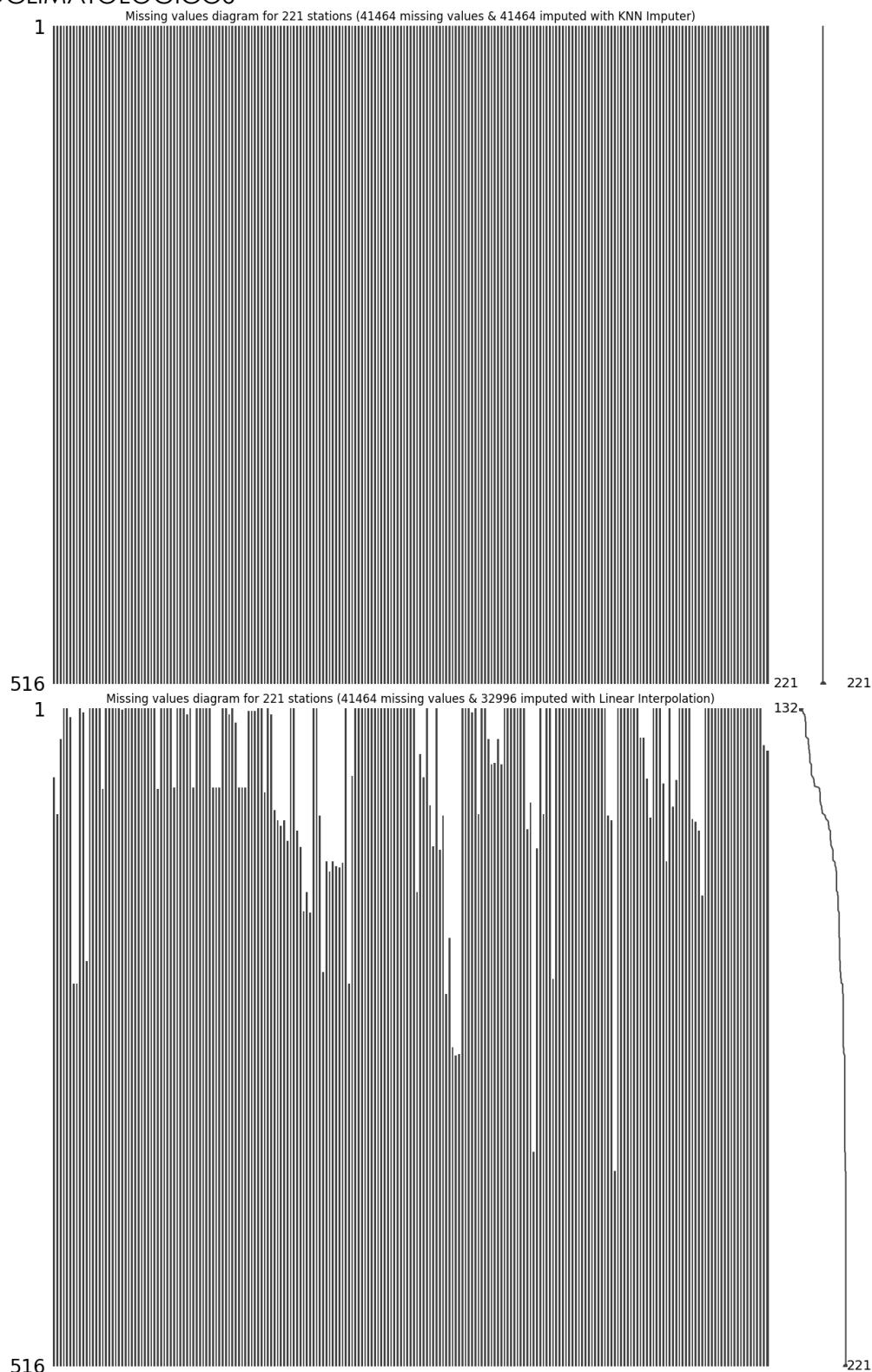
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



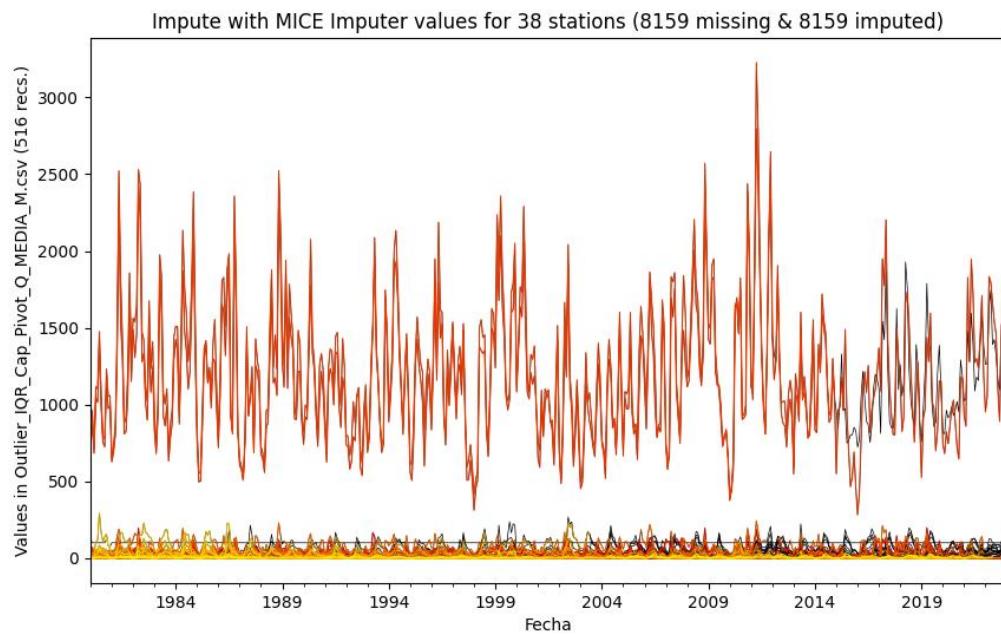
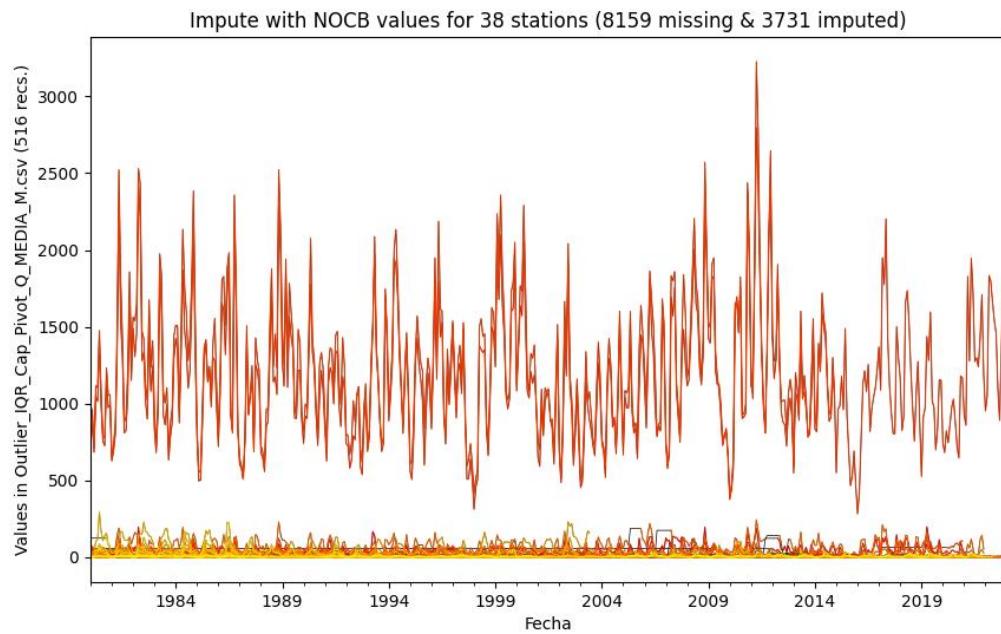
Fuente: Elaboración Propia, 2023.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

Se realizó el mismo ejercicio para los parámetros de caudal.

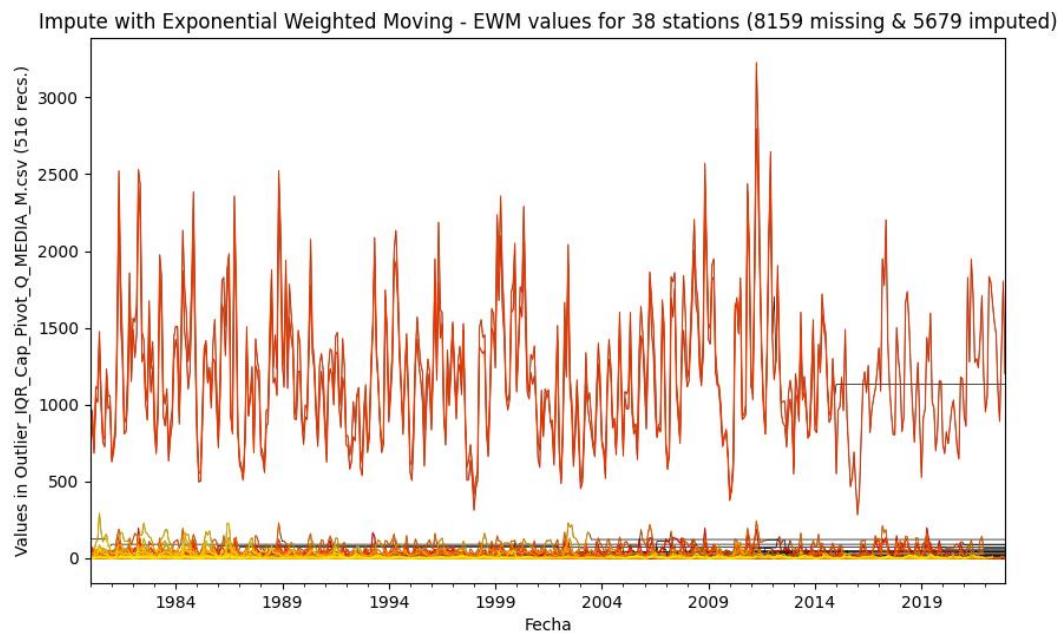
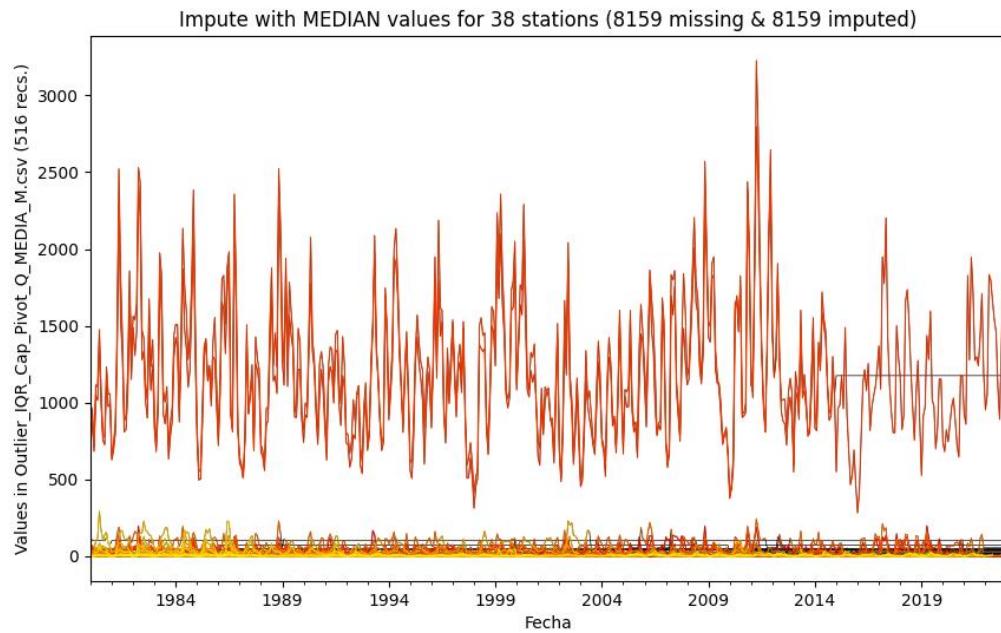
JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Ilustración 3-7. Resultados Impute Caudal



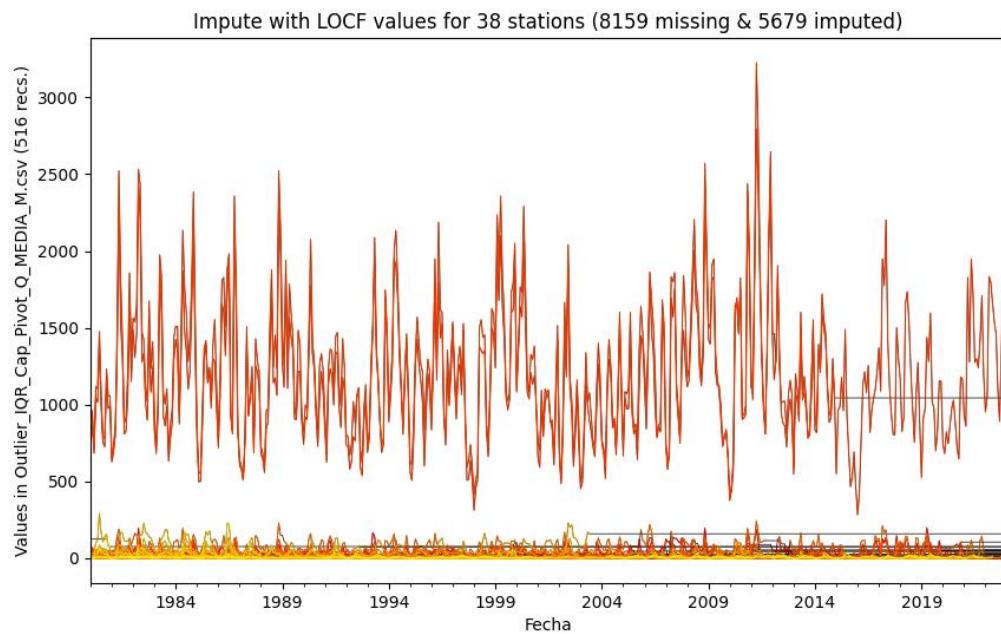
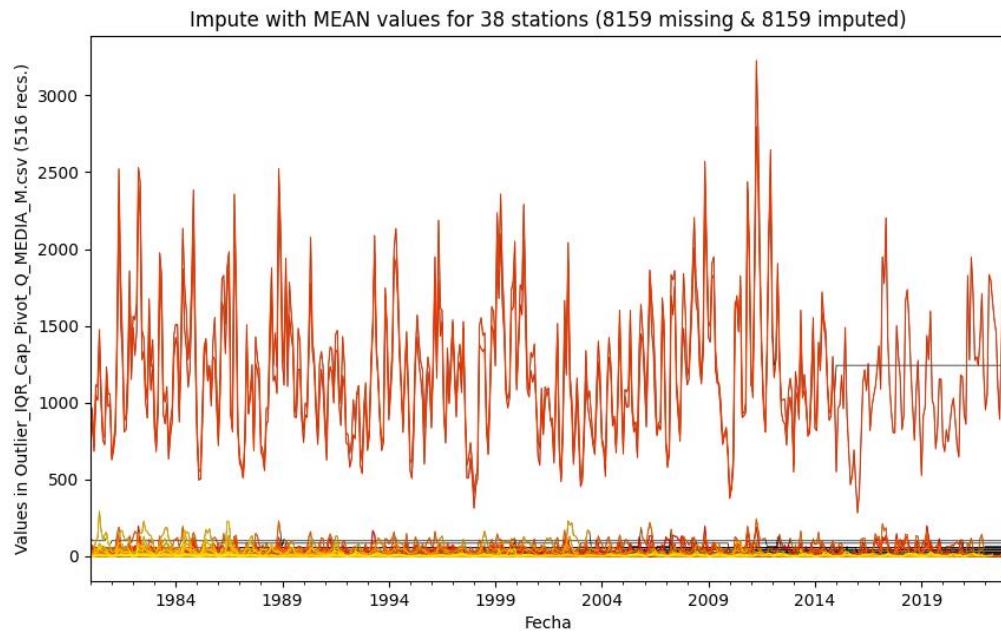
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



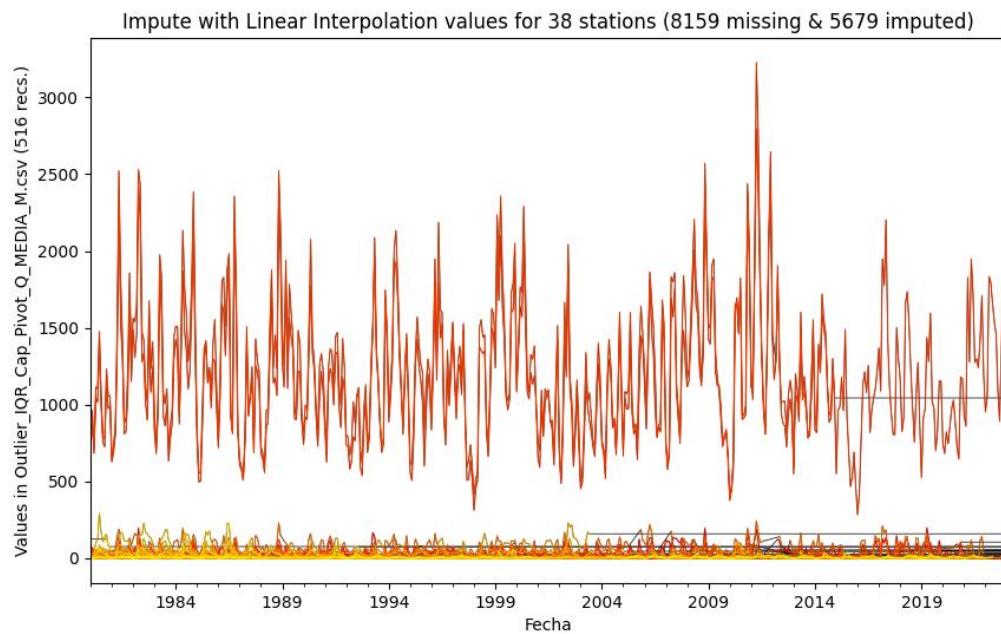
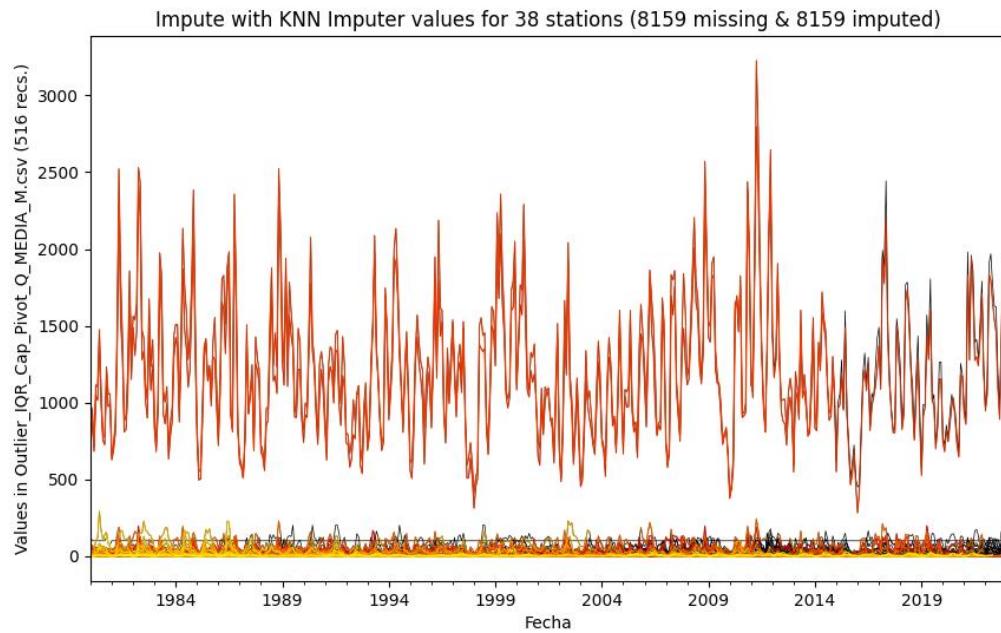
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



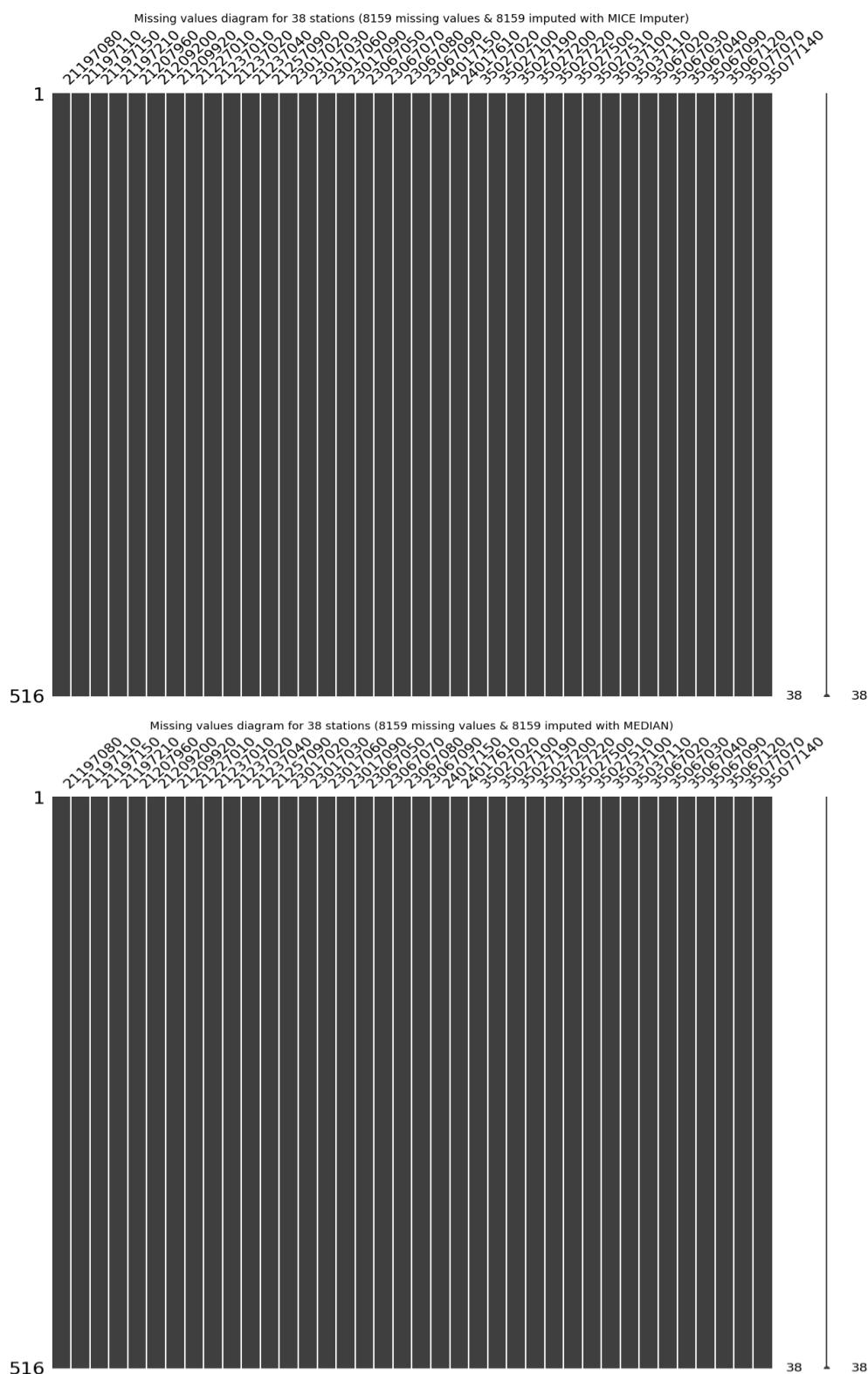
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



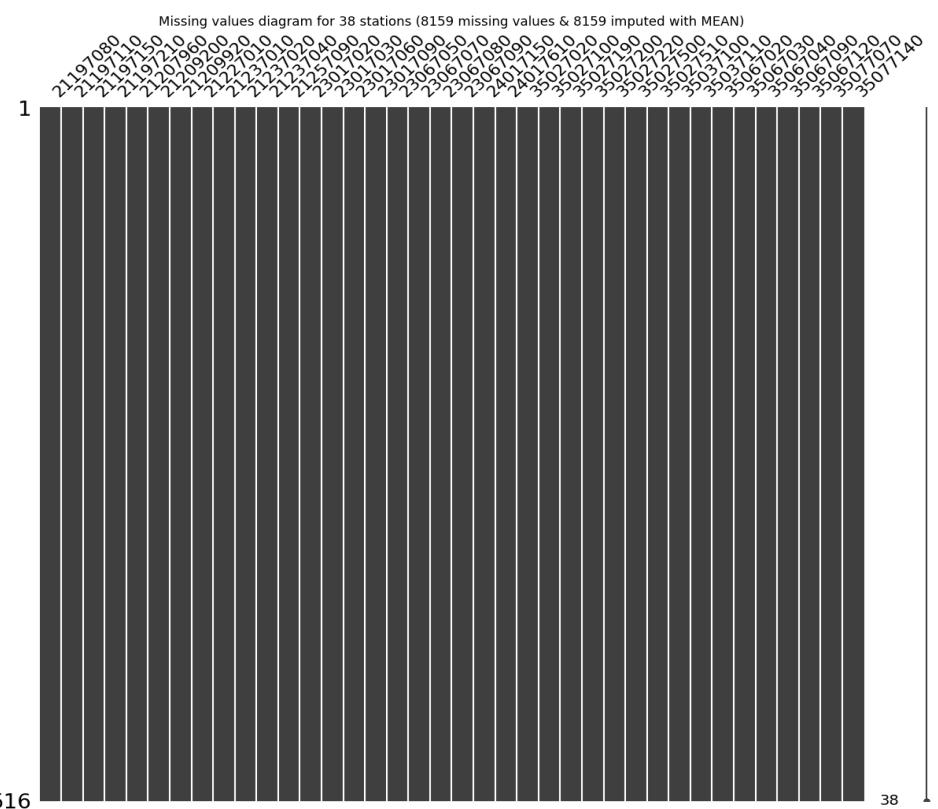
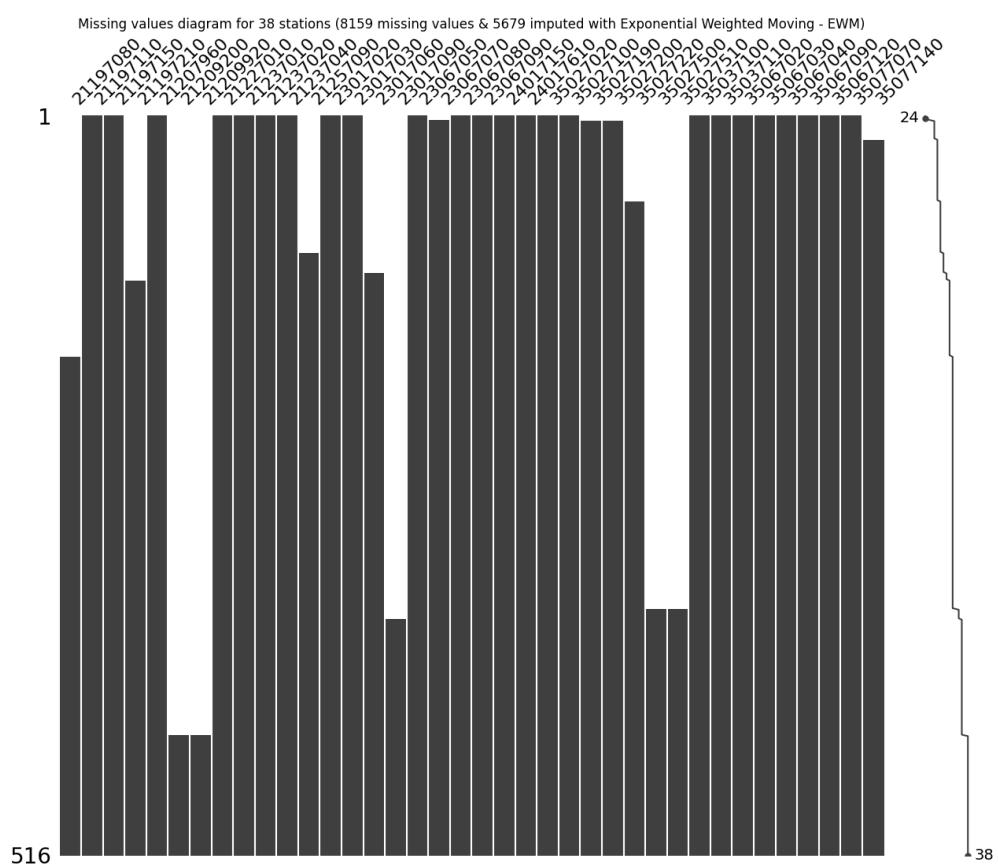
Fuente: Elaboración Propia, 2023.

Ilustración 3-8. Resultados Impute Caudal (2)



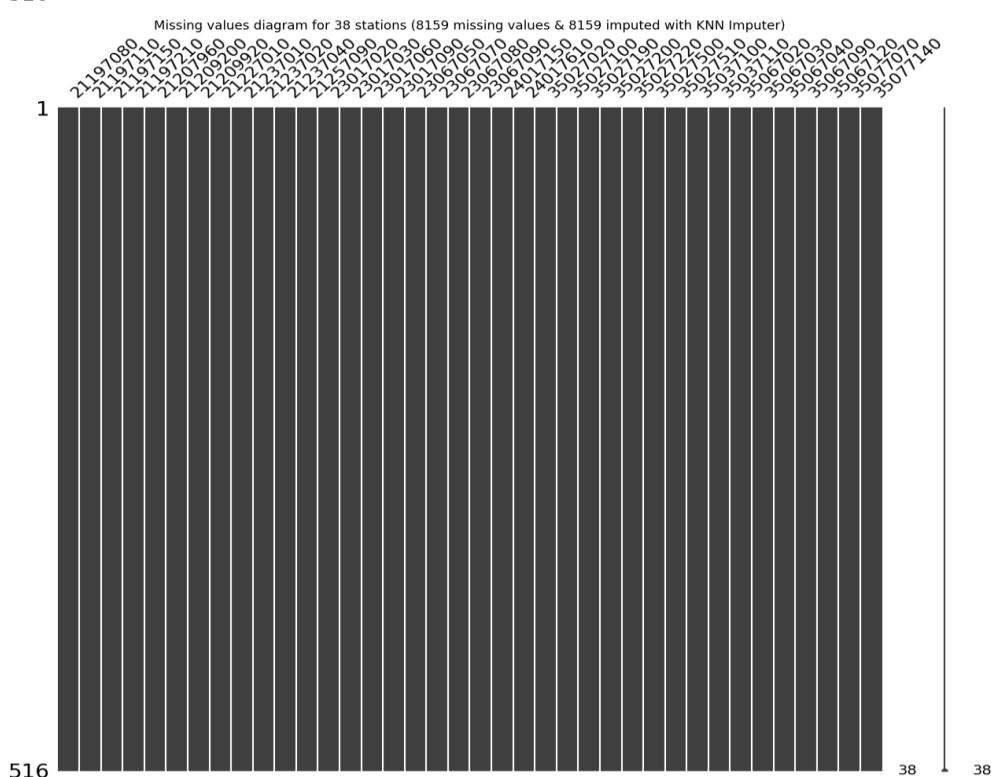
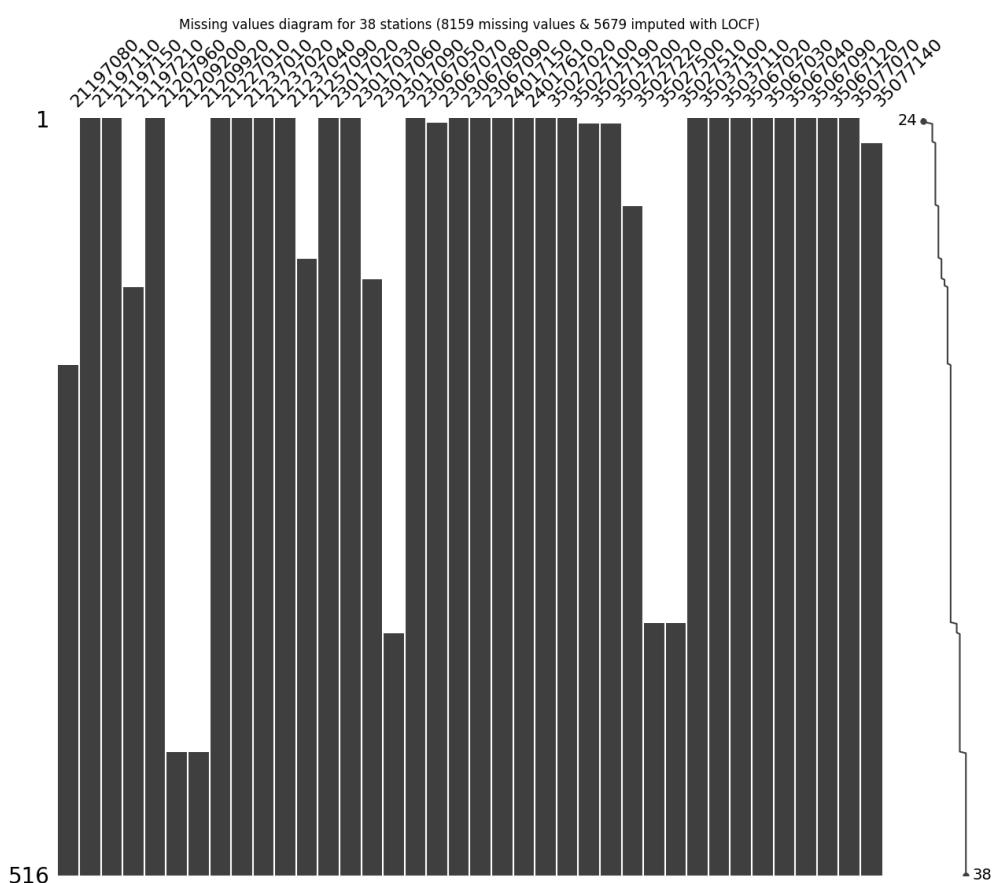
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



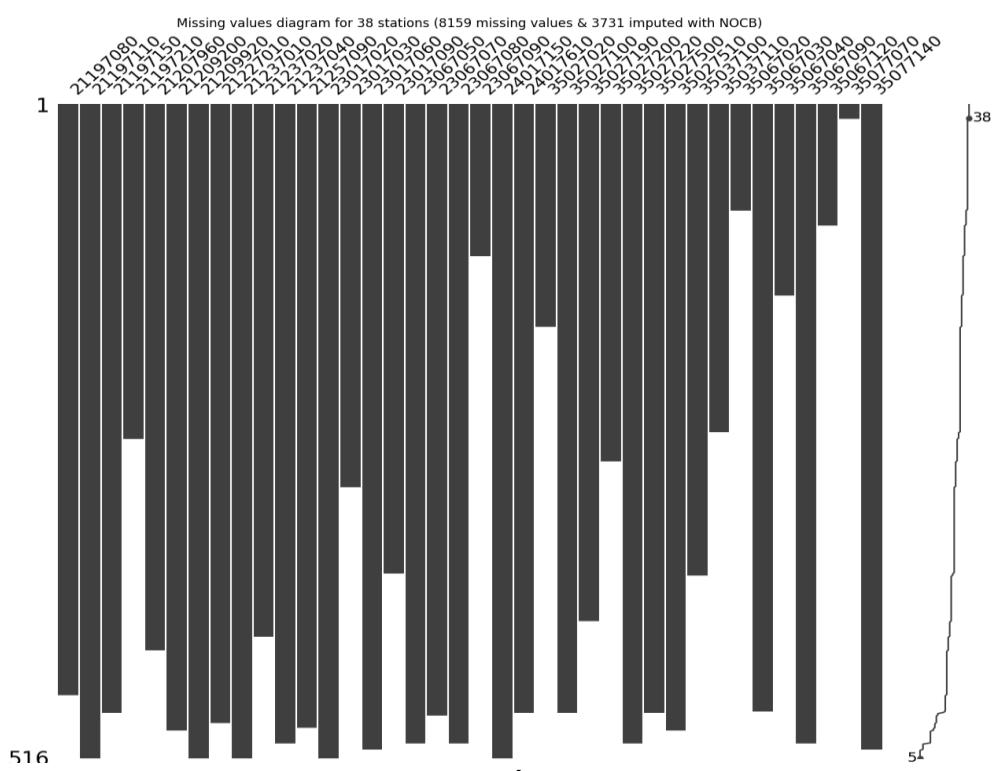
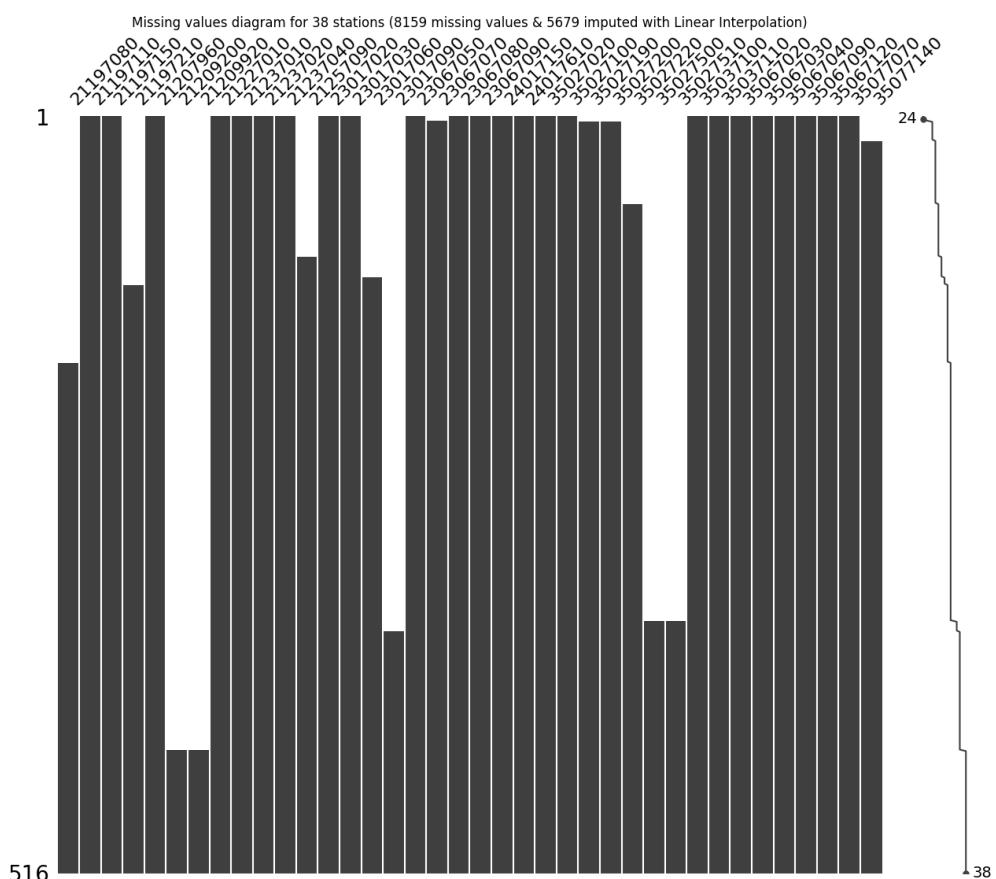
SECTION 03 DESCARGA, PROCESAMIENTO Y ANÁLISIS DE DATOS HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03 DESCARGA, PROCESAMIENTO Y ANÁLISIS DE DATOS HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



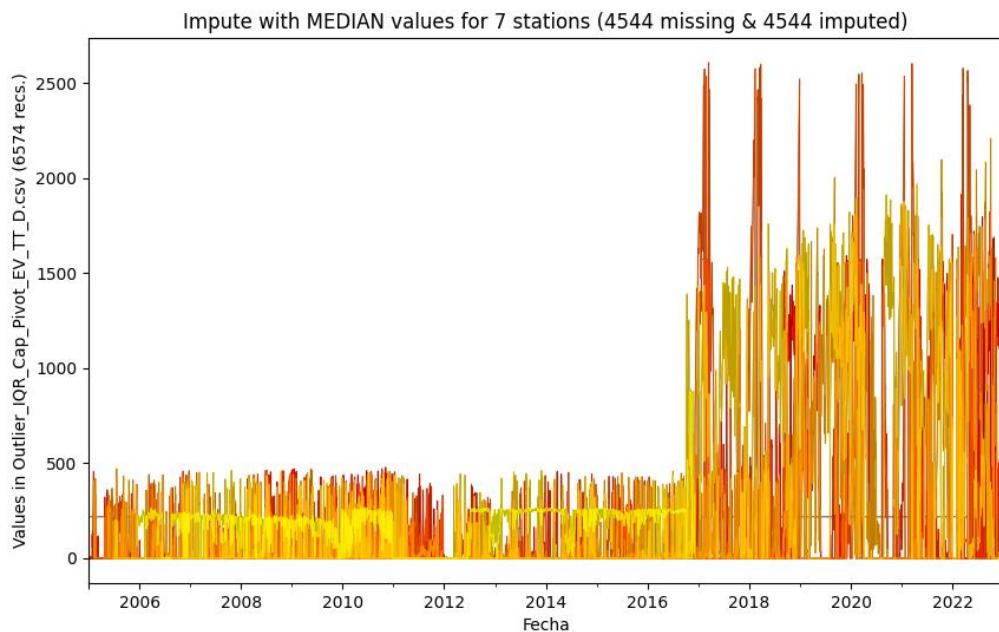
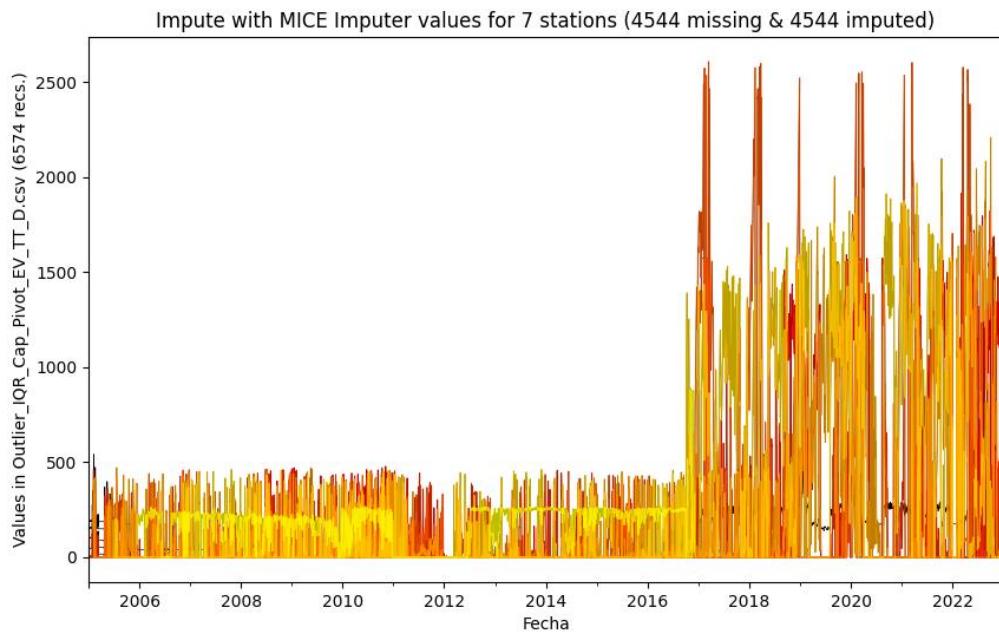
Fuente: Elaboración Propia, 2023.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

Se realizó el mismo ejercicio para los parámetros de evaporación.

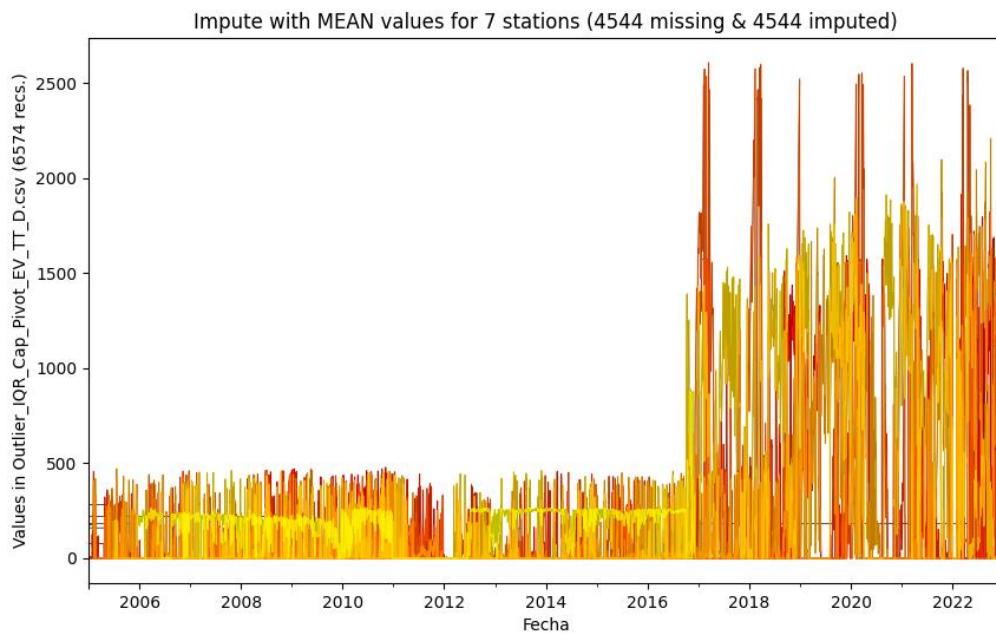
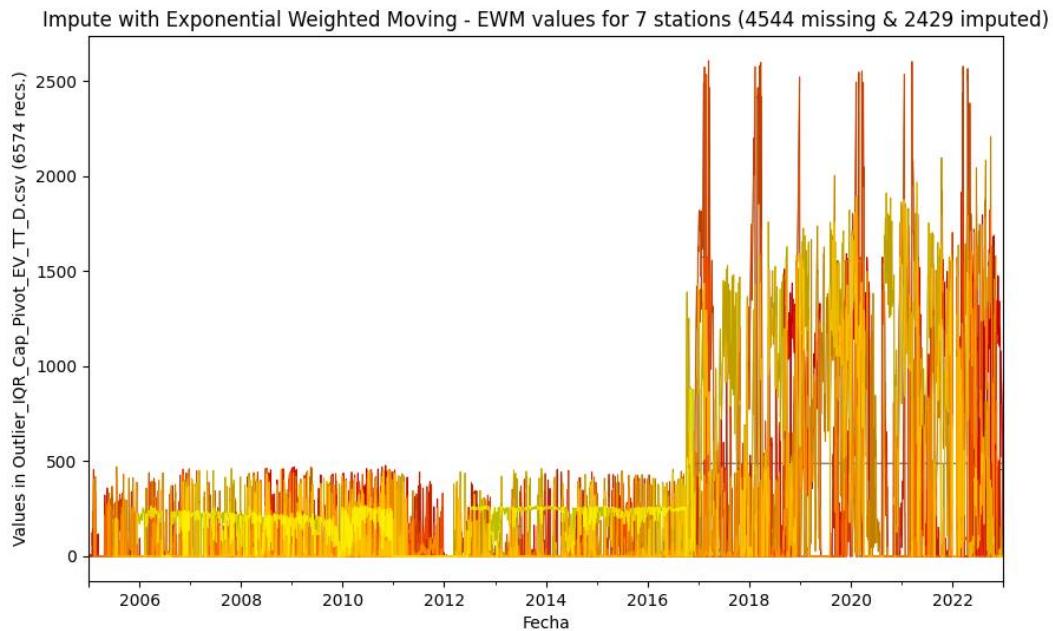
JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Ilustración 3-9. Resultados Impute Evaporación



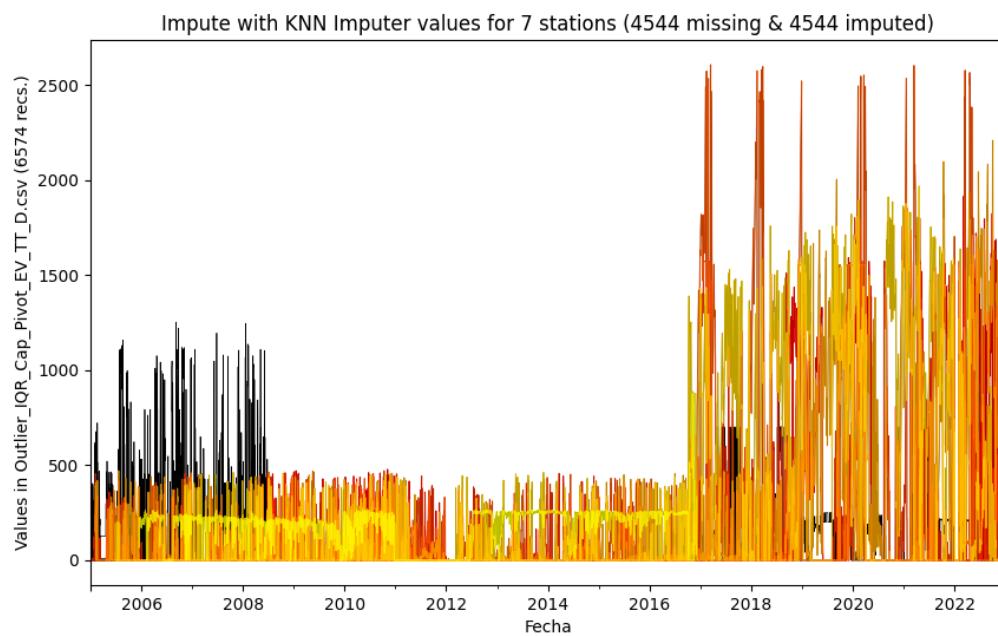
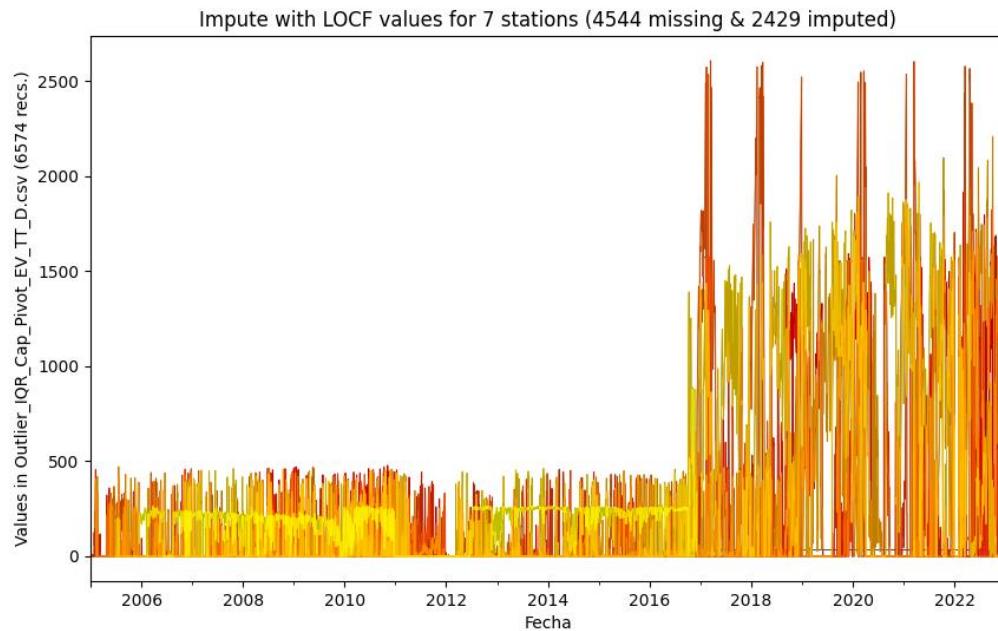
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



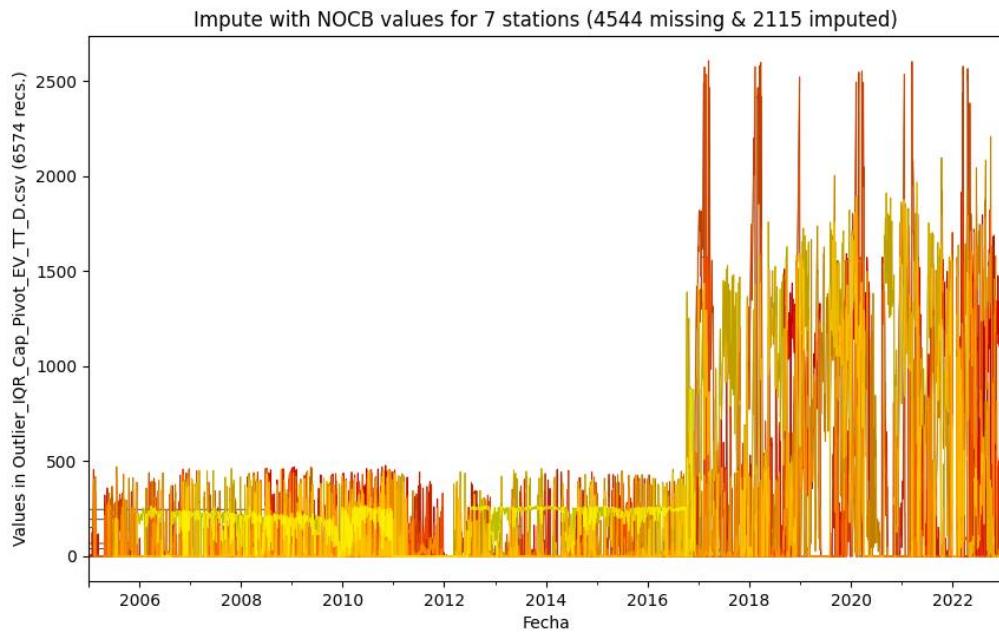
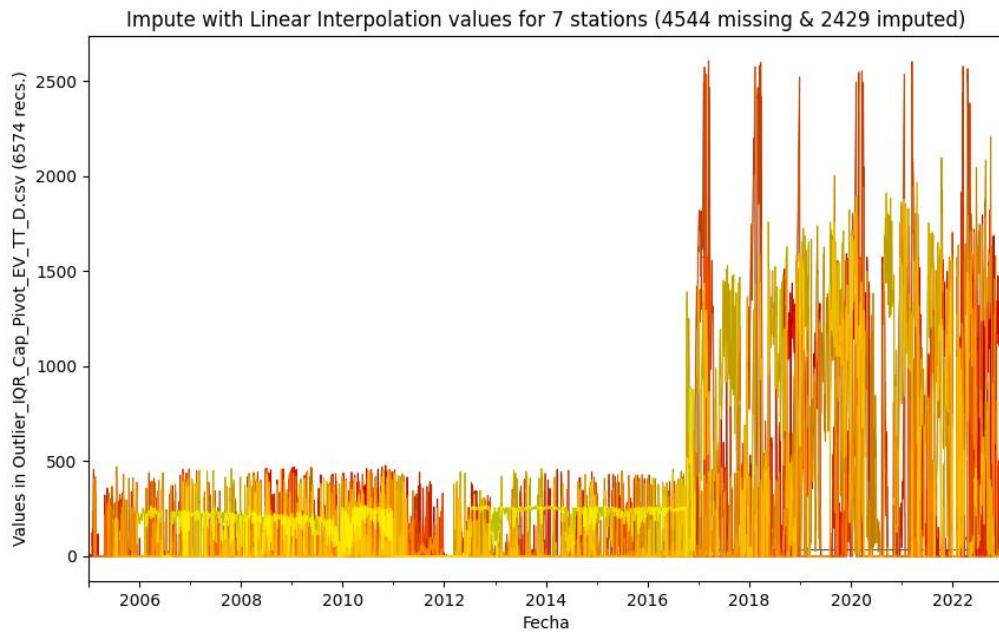
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

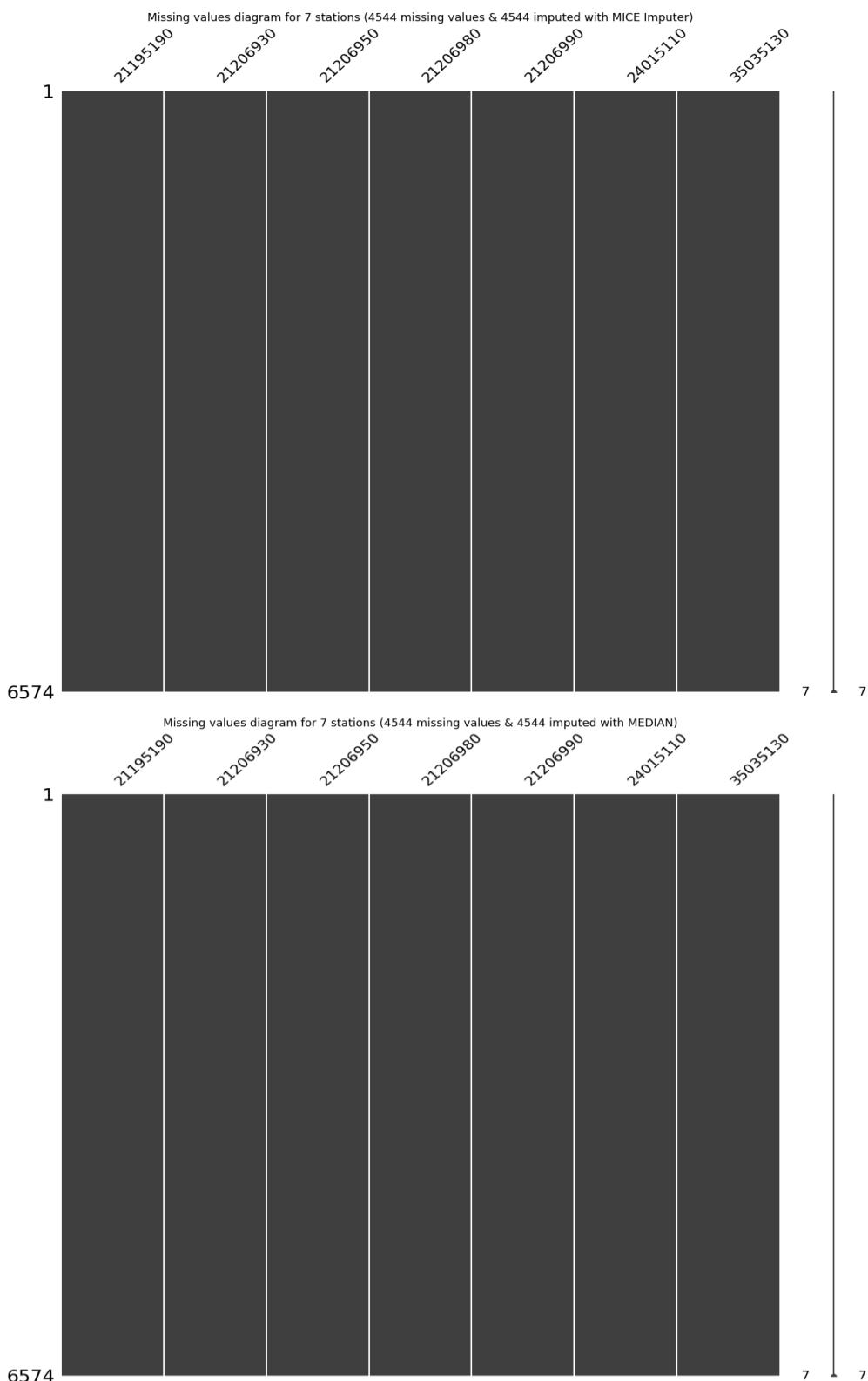


Fuente: Elaboración Propia, 2023.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

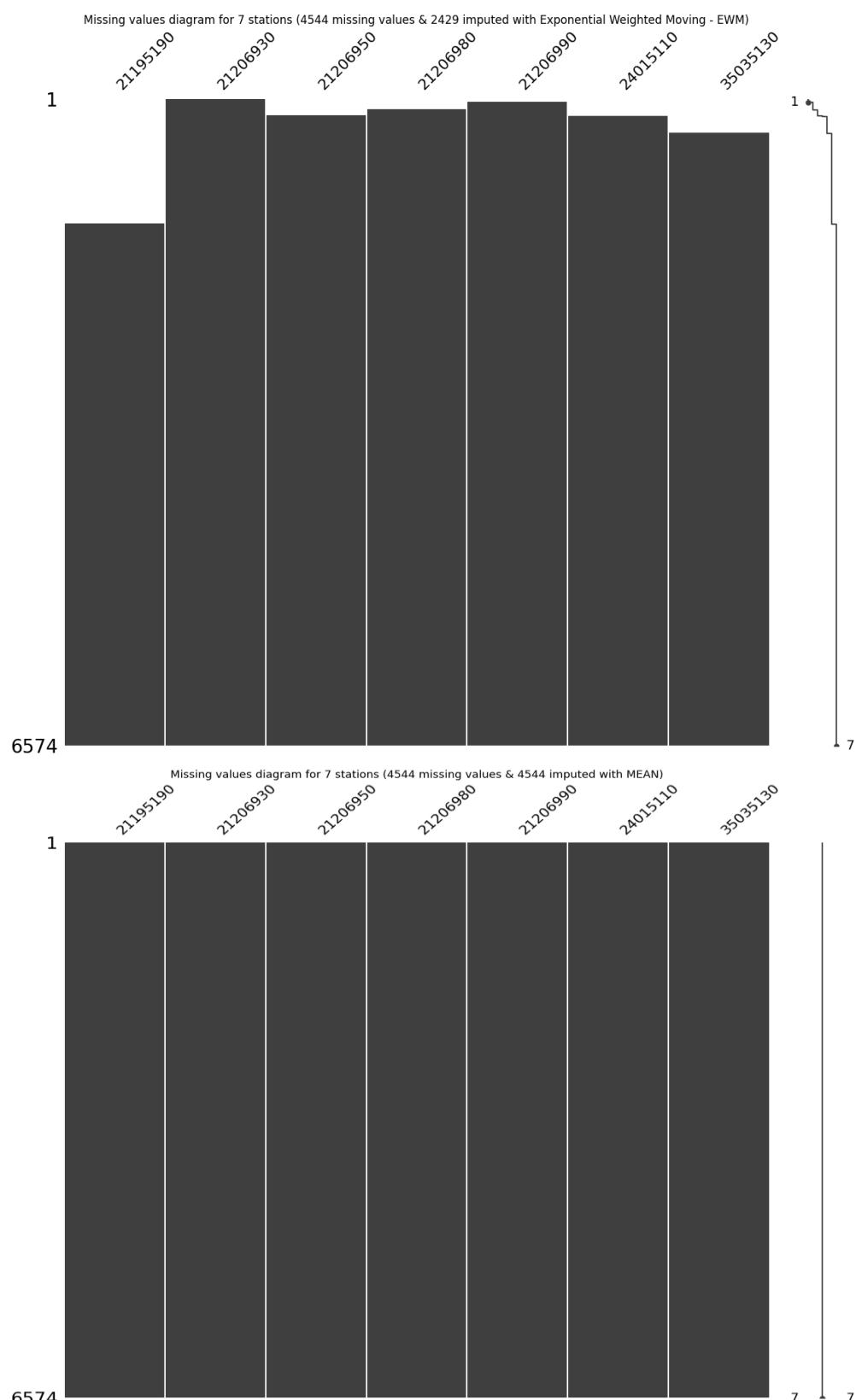
JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Ilustración 3-10. Resultados Impute Evaporación (2)



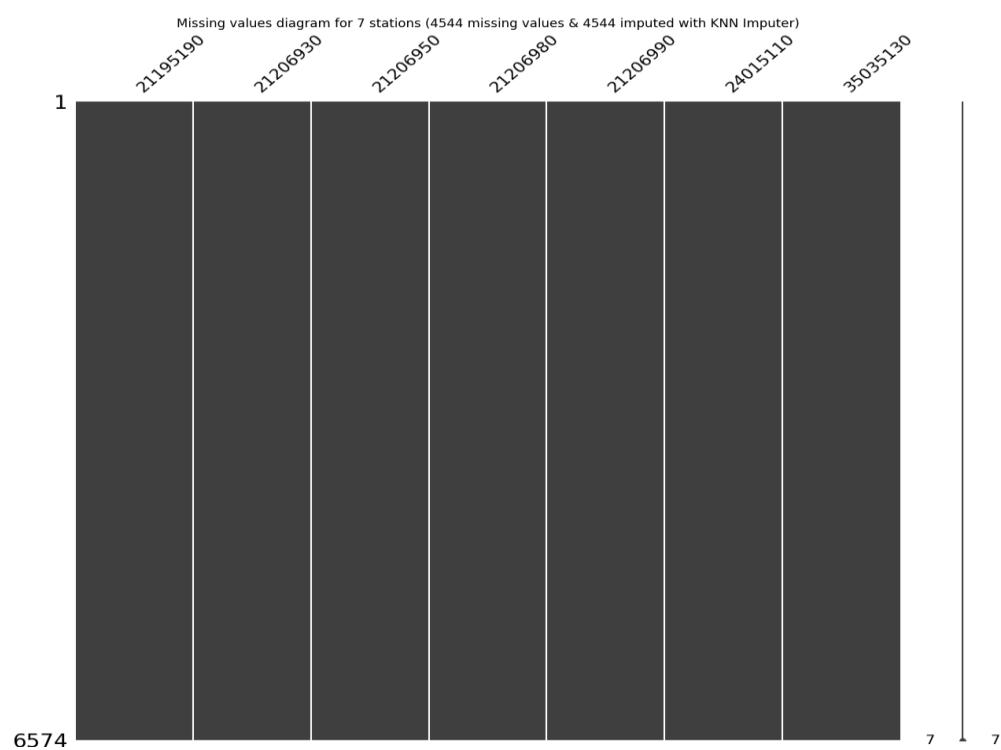
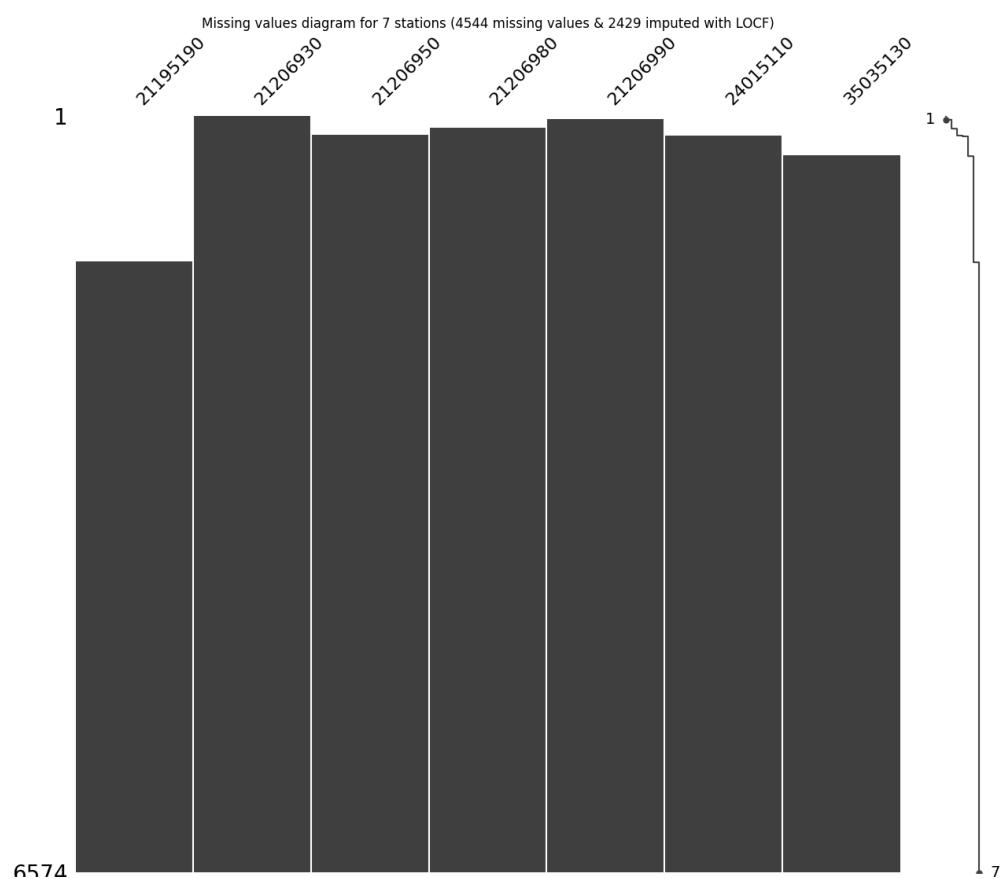
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



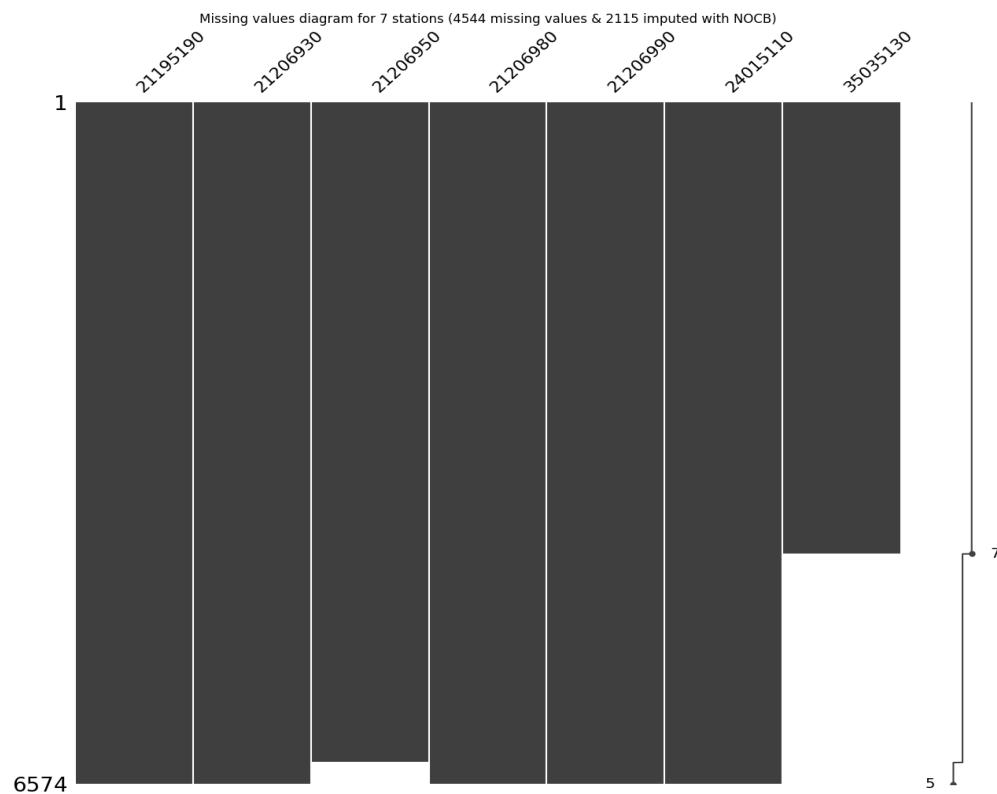
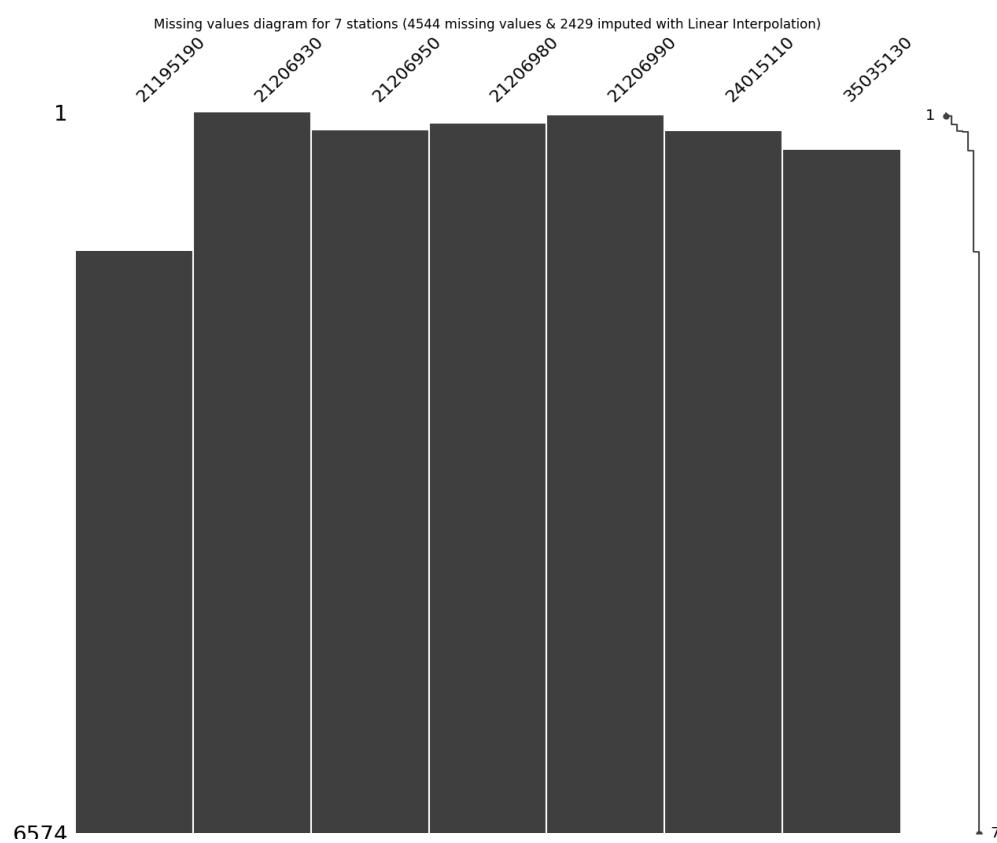
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



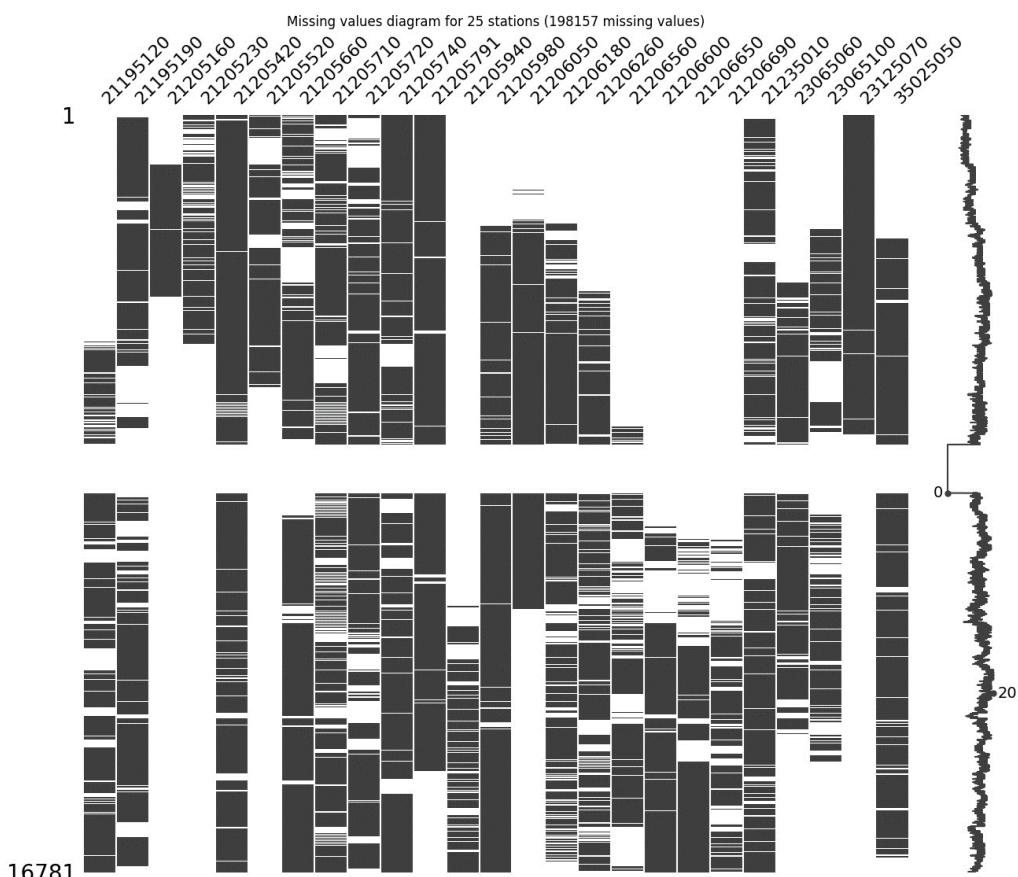
Fuente: Elaboración Propia, 2023.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Se realizó el mismo ejercicio para los parámetros de temperatura máxima. Para el caso de este parámetro se observa en la grafica de datos faltantes que el periodo del año 1995 no tiene información con las estaciones filtradas de acuerdo con el criterio mencionado al inicio de este capítulo.

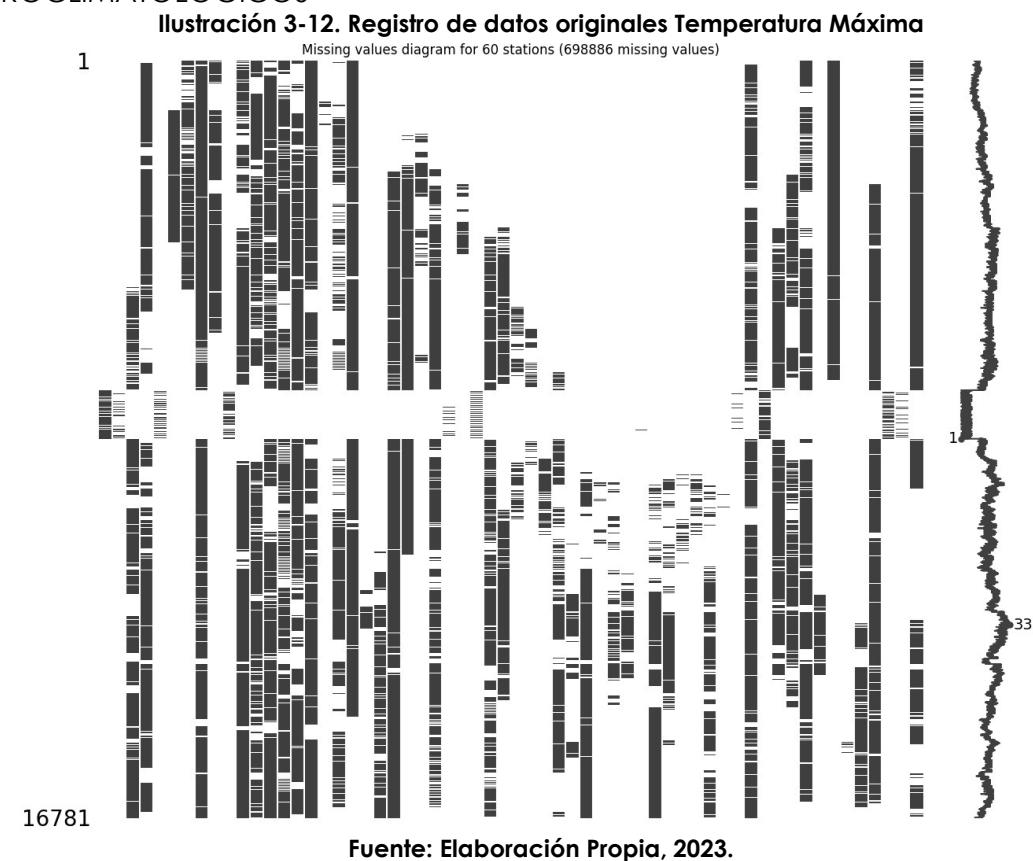
Ilustración 3-11. Registro de datos filtrados Temperatura Máxima



Fuente: Elaboración Propia, 2023.

Sin embargo, si se compara con la serie de datos sin eliminar estaciones, se observa que se evidencia un caso especial en este año ya que algunas estaciones tienen información solo en ese periodo, como se observa a continuación.

En este caso, se opta por utilizar las series de las estaciones filtradas ya que se considera en mejor medida completar los datos de un año a completar el resto del periodo analizado para varias estaciones.

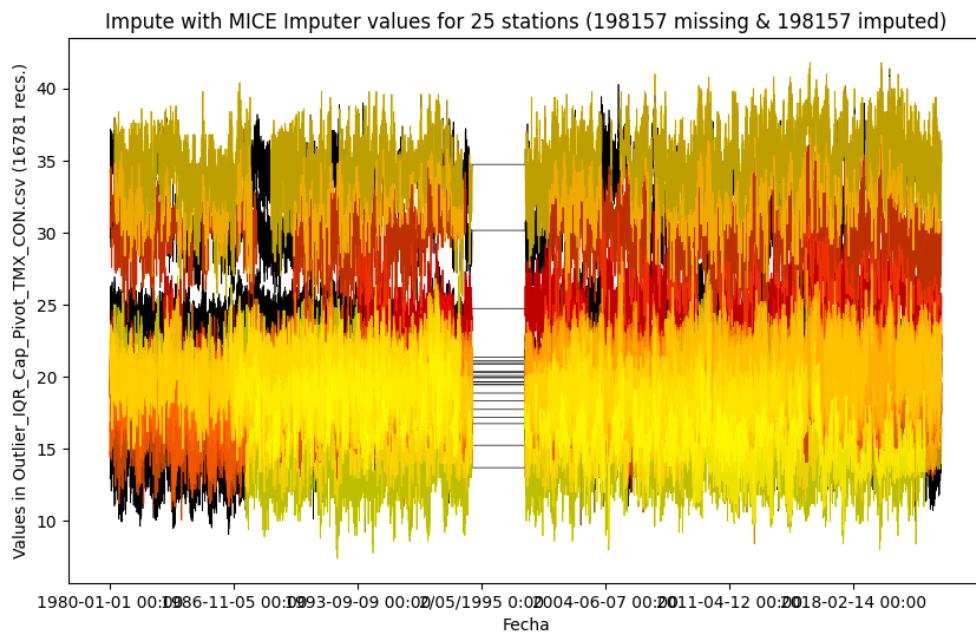
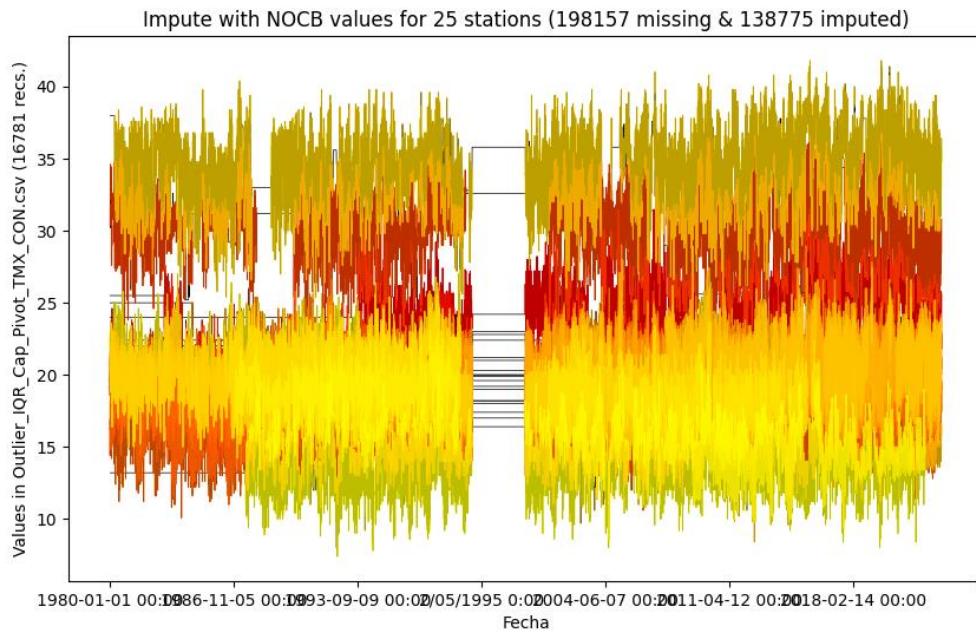


A continuación, los resultados de las estaciones de temperatura máxima completada.

Ilustración 3-13. Resultados Impute Temperatura Máxima

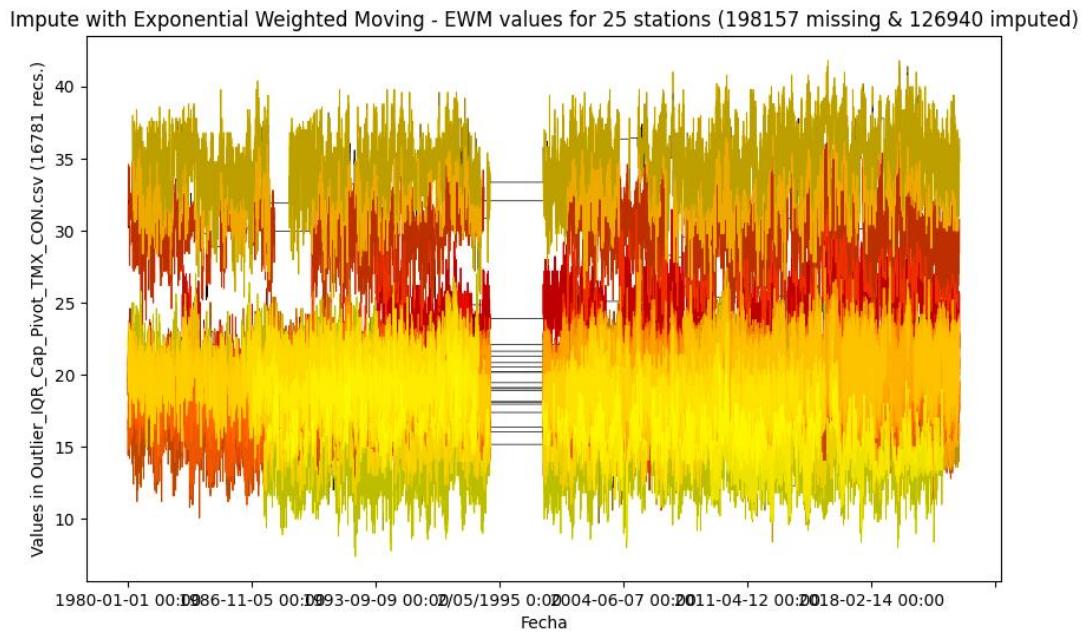
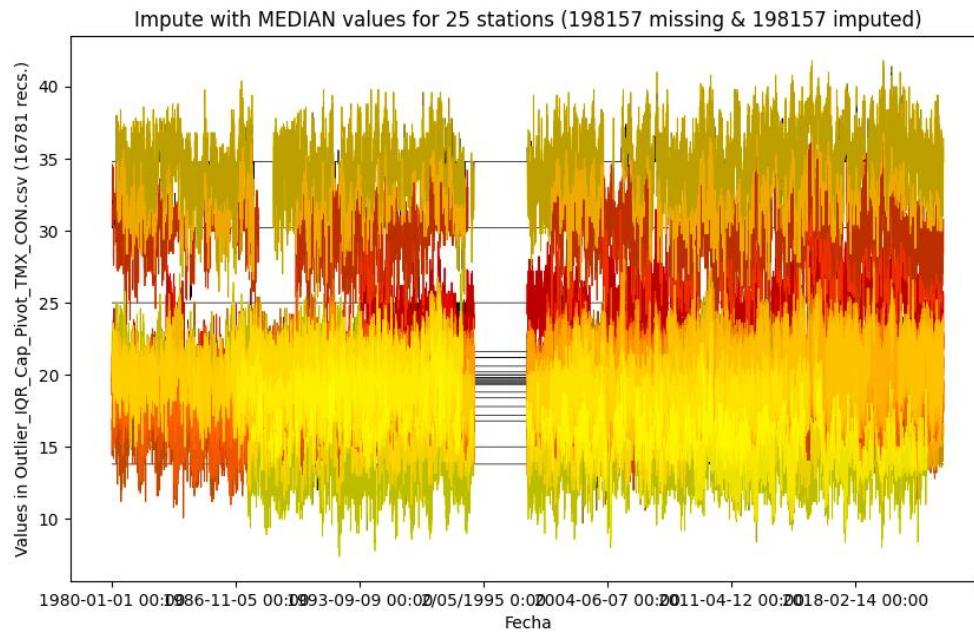
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



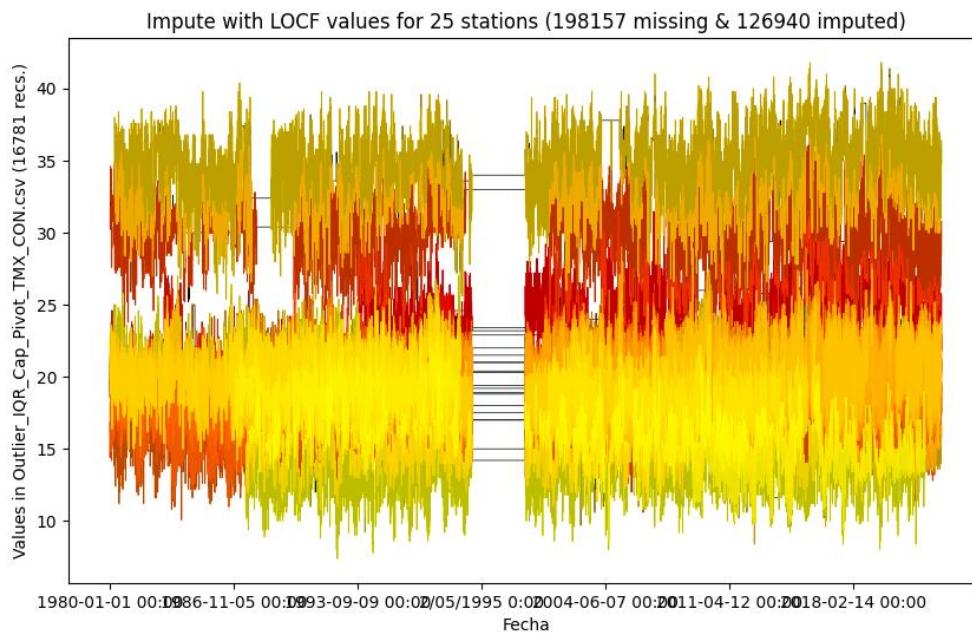
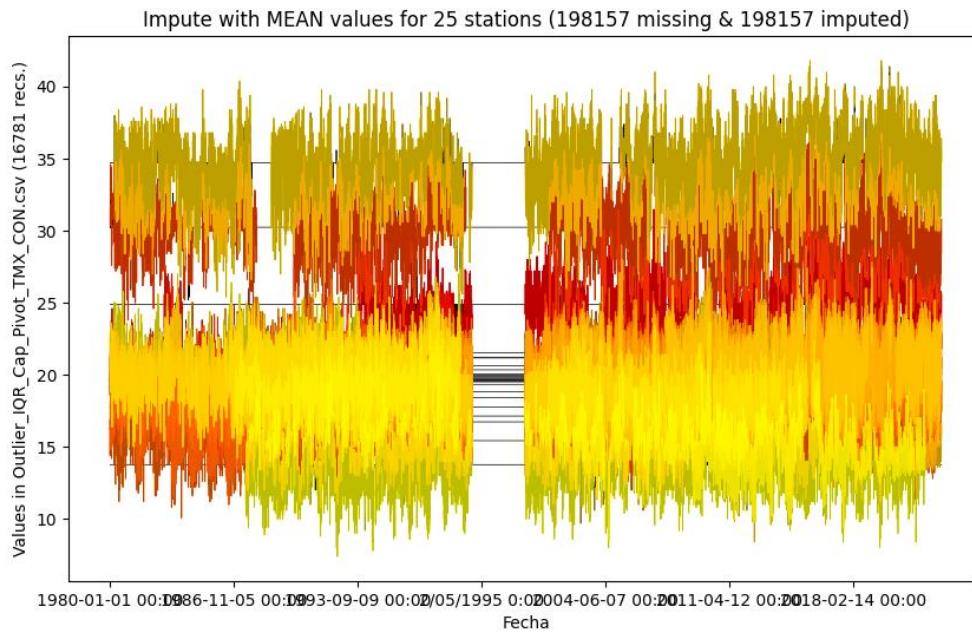
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



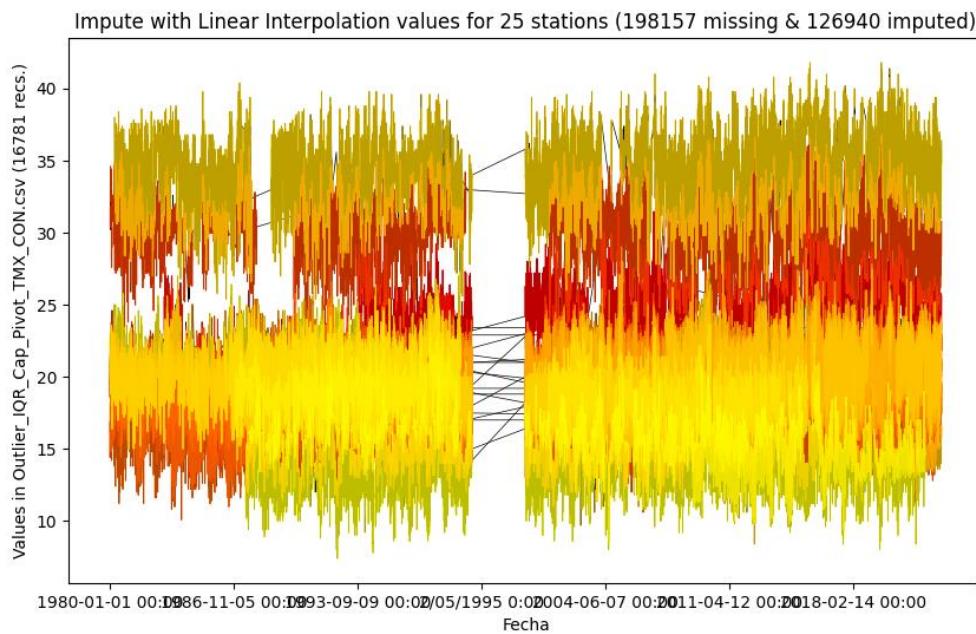
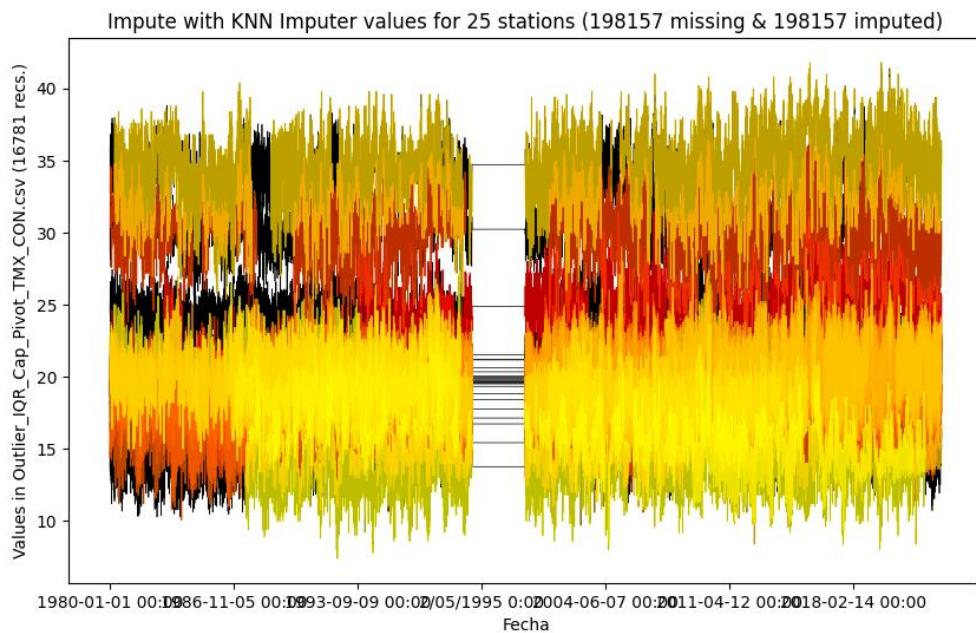
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

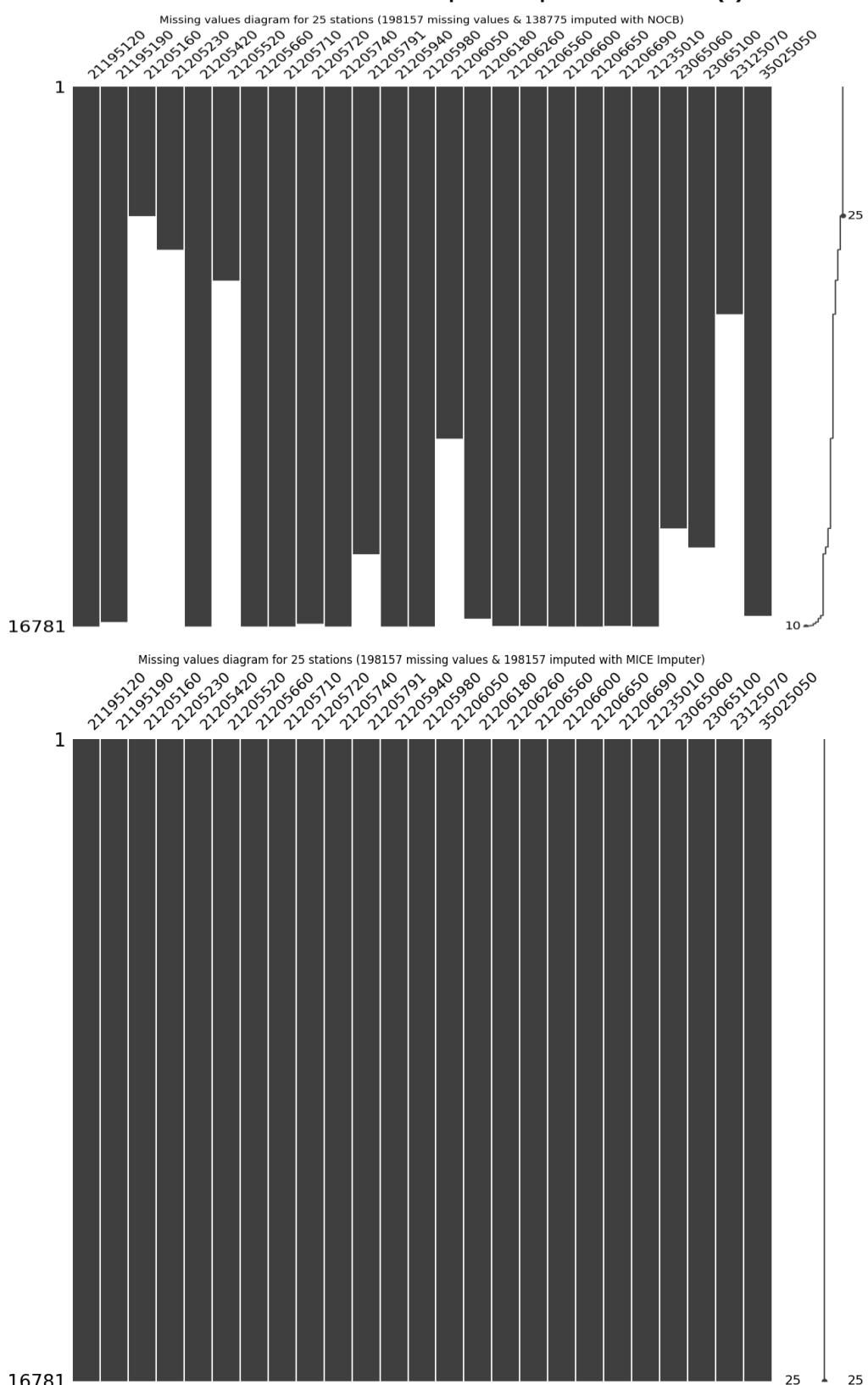


Fuente: Elaboración Propia, 2023.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

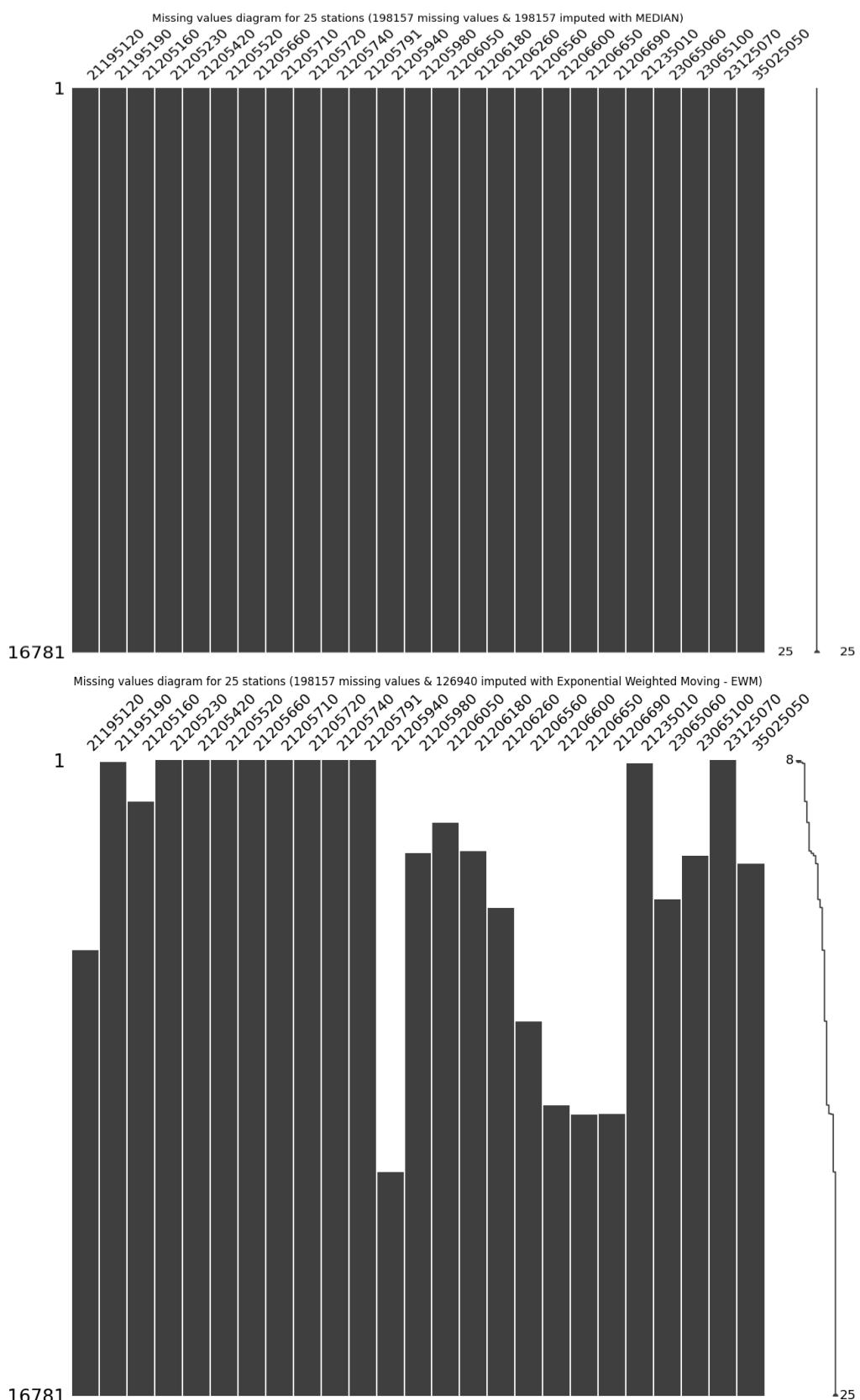
JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Ilustración 3-14. Resultados Impute Temperatura Máxima (2)



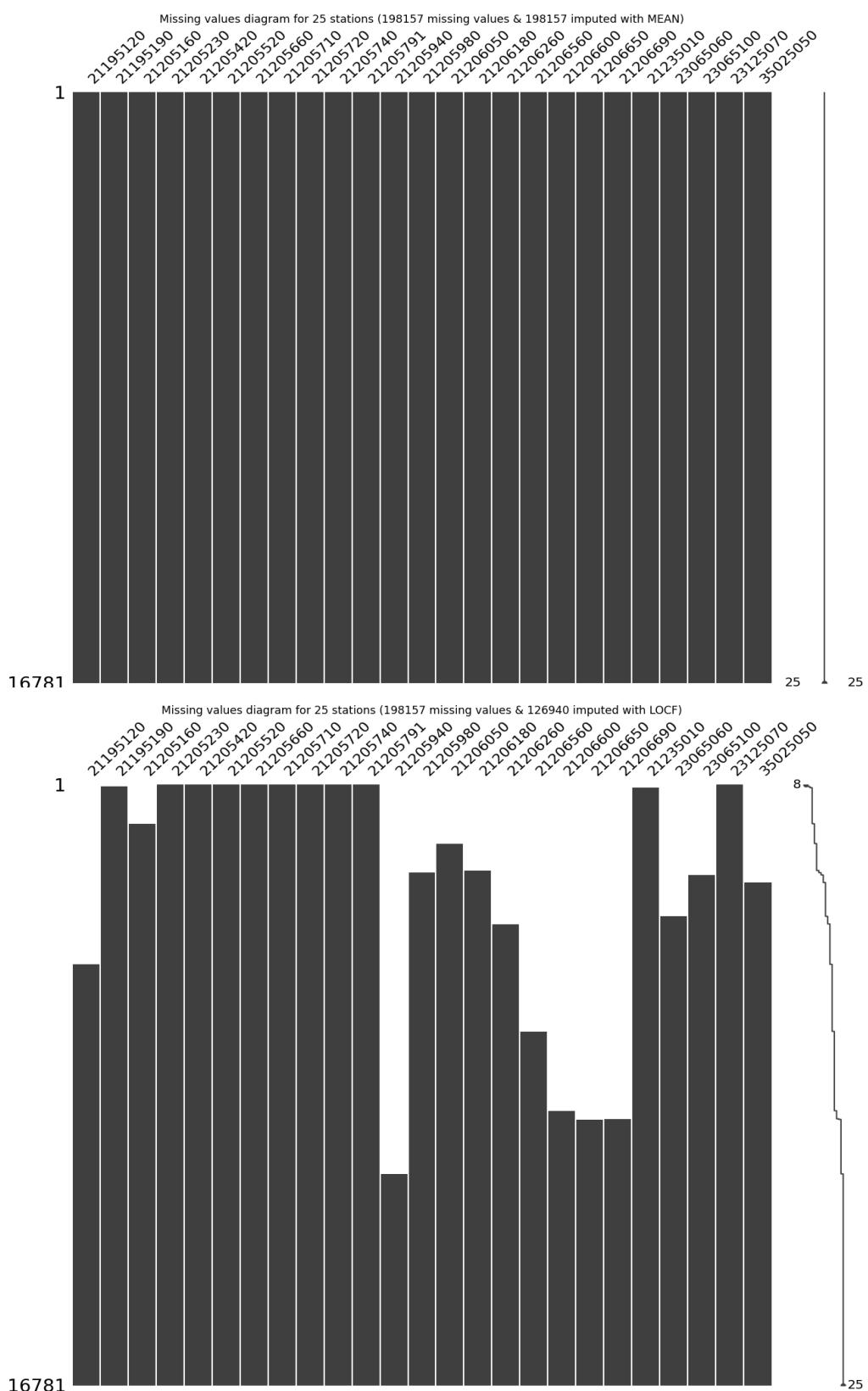
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



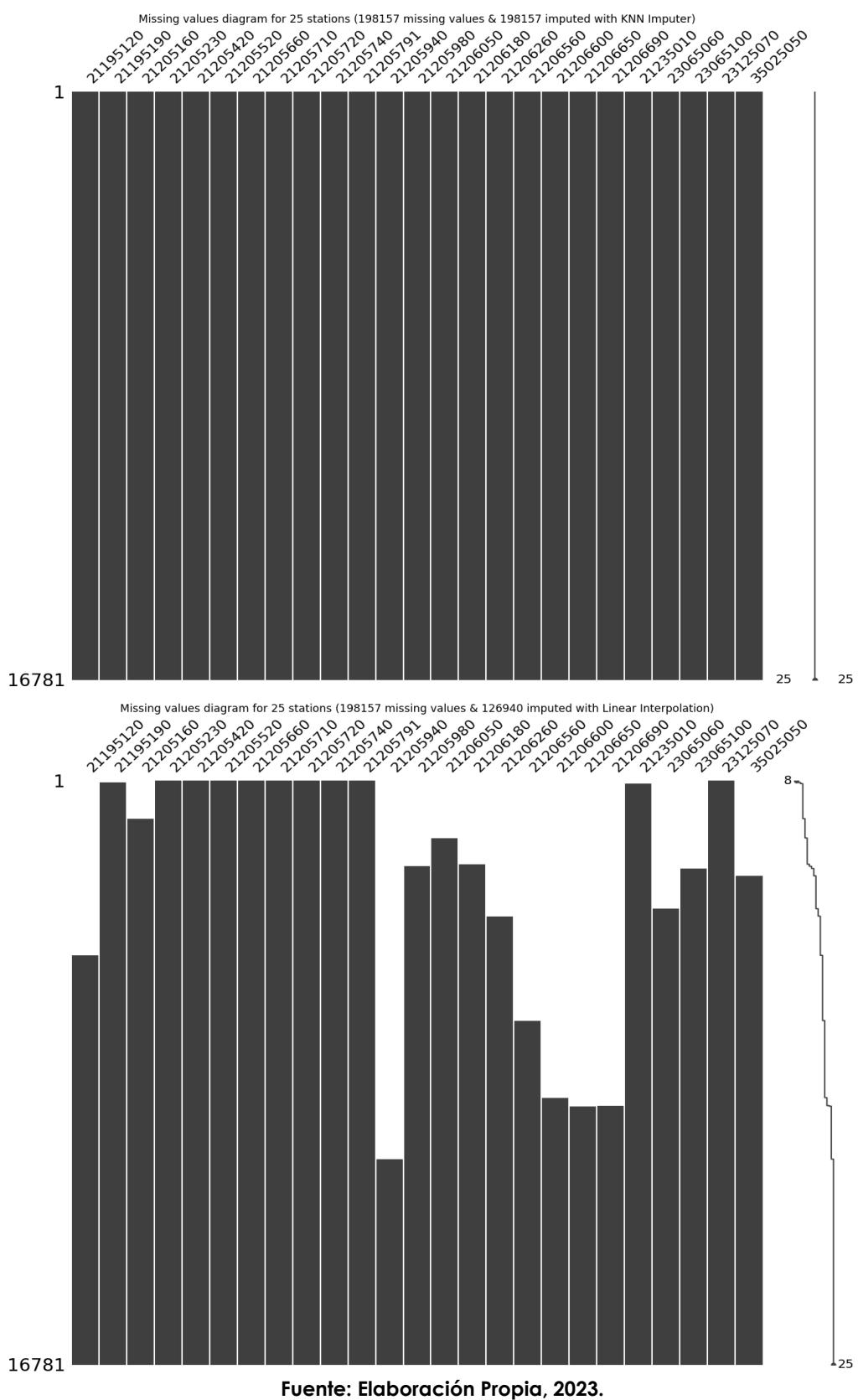
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

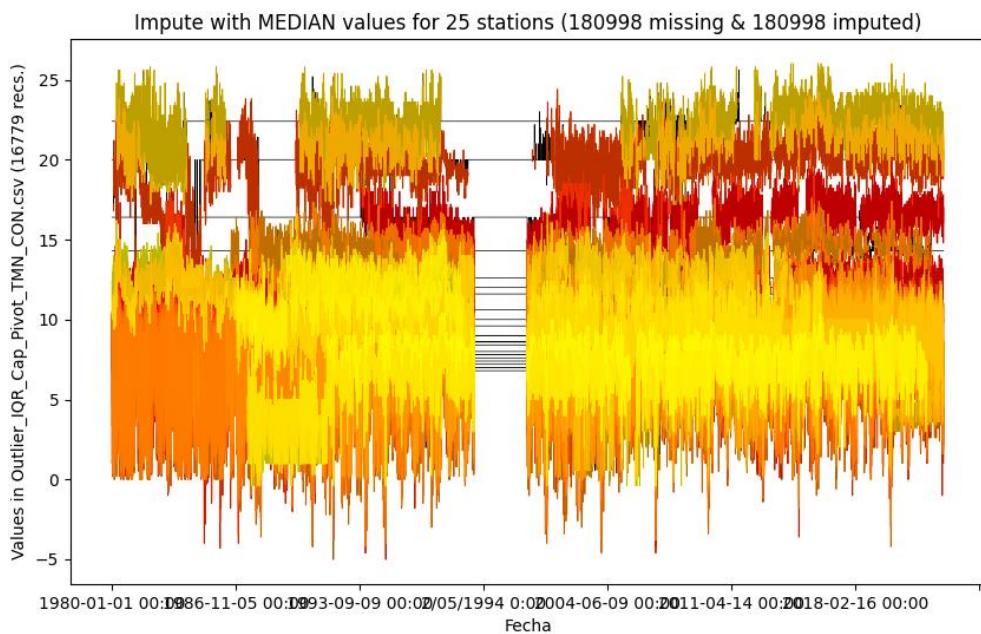
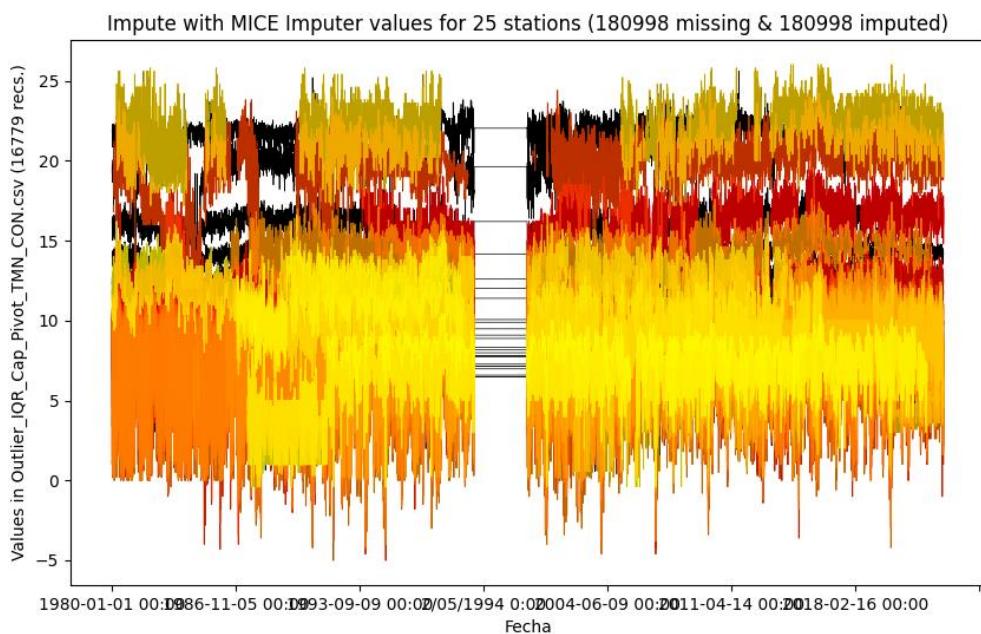


SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

Se realizó el mismo ejercicio para los parámetros de temperatura mínima. En este caso aplica el mismo concepto de la temperatura máxima para el periodo 1995.

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

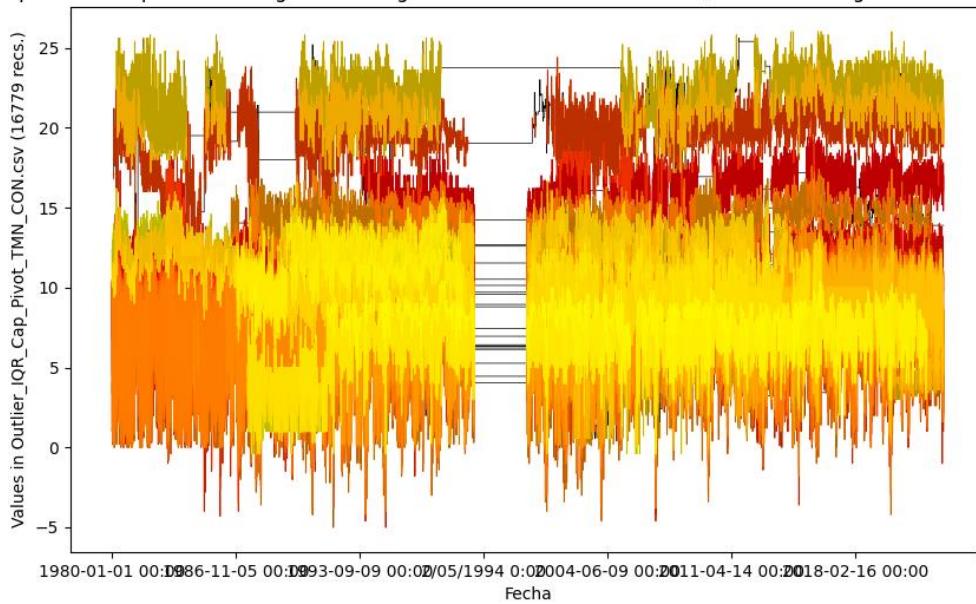
Ilustración 3-15. Resultados Impute Temperatura Mínima



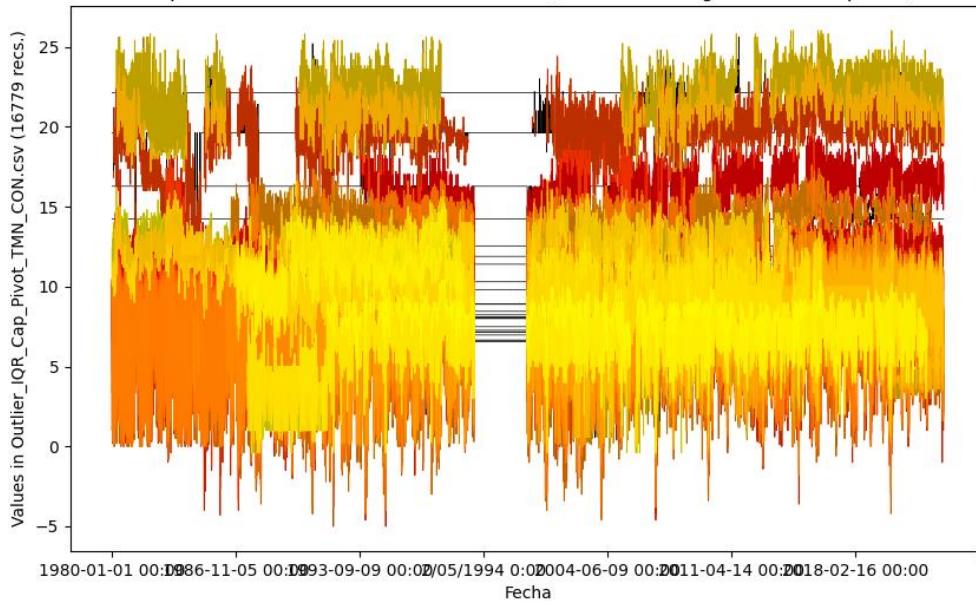
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Impute with Exponential Weighted Moving - EWM values for 25 stations (180998 missing & 123550 imputed)

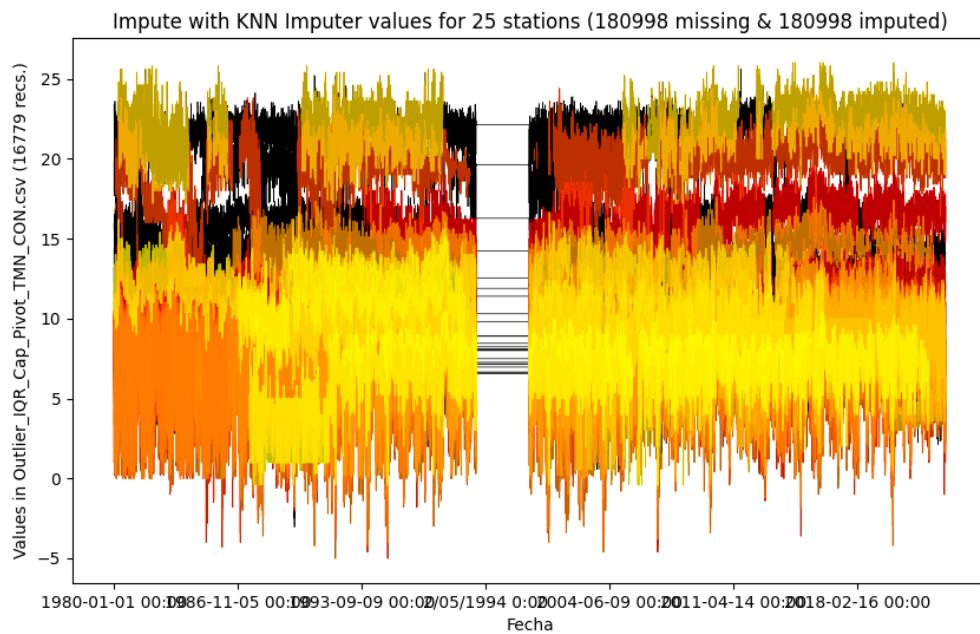
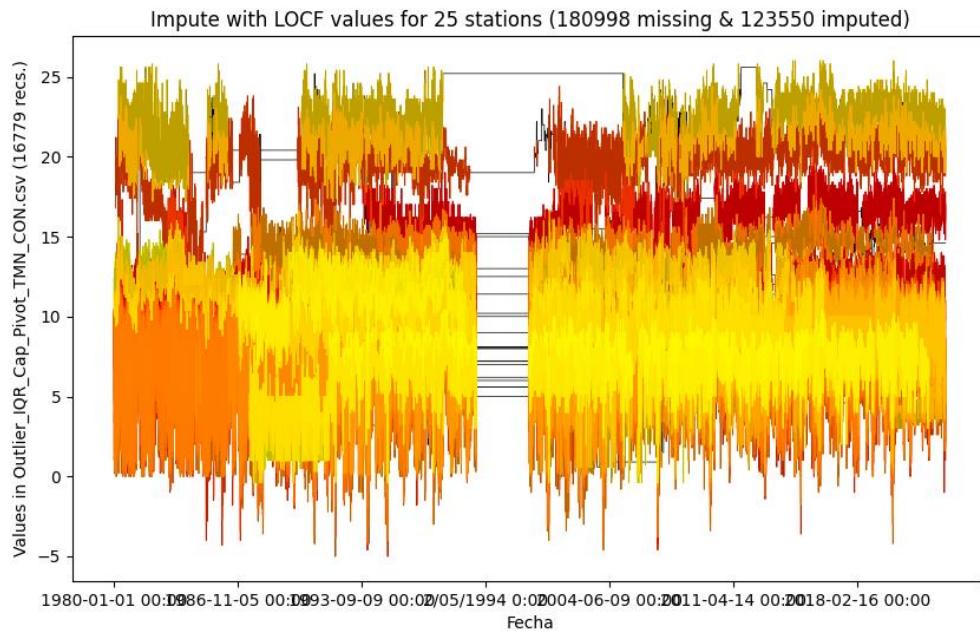


Impute with MEAN values for 25 stations (180998 missing & 180998 imputed)



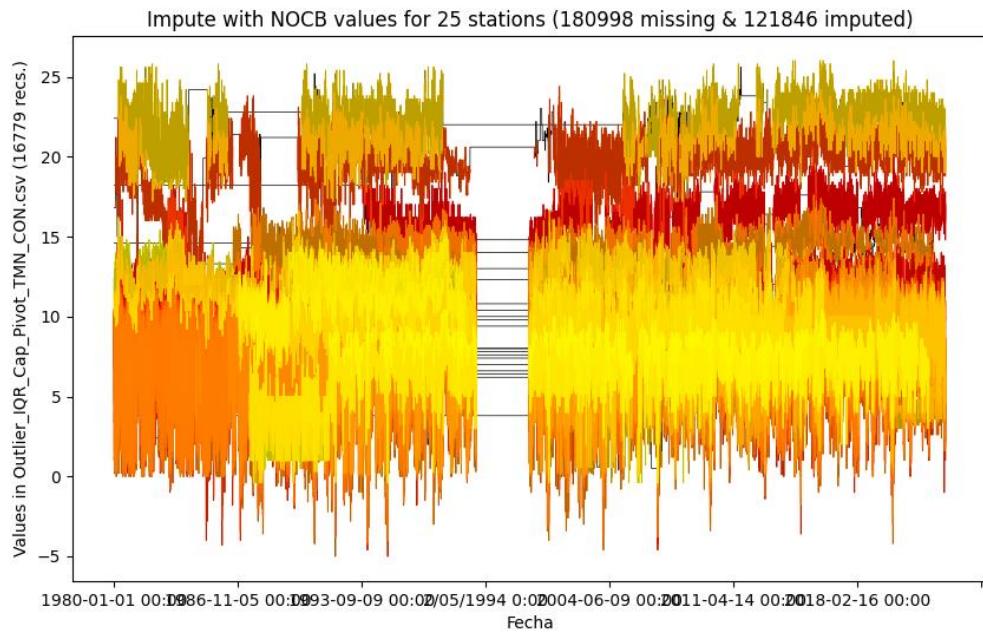
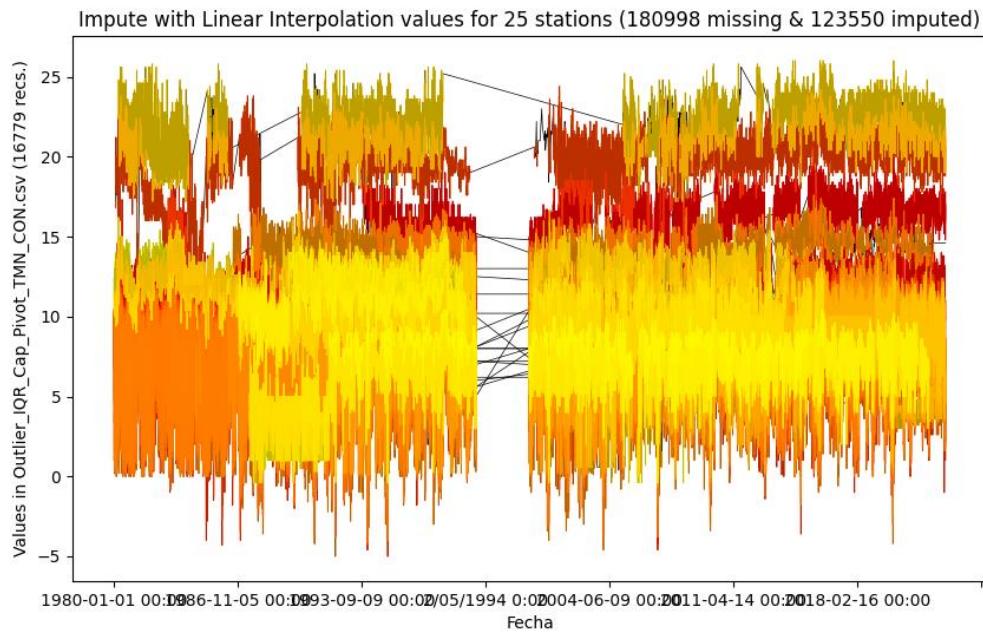
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

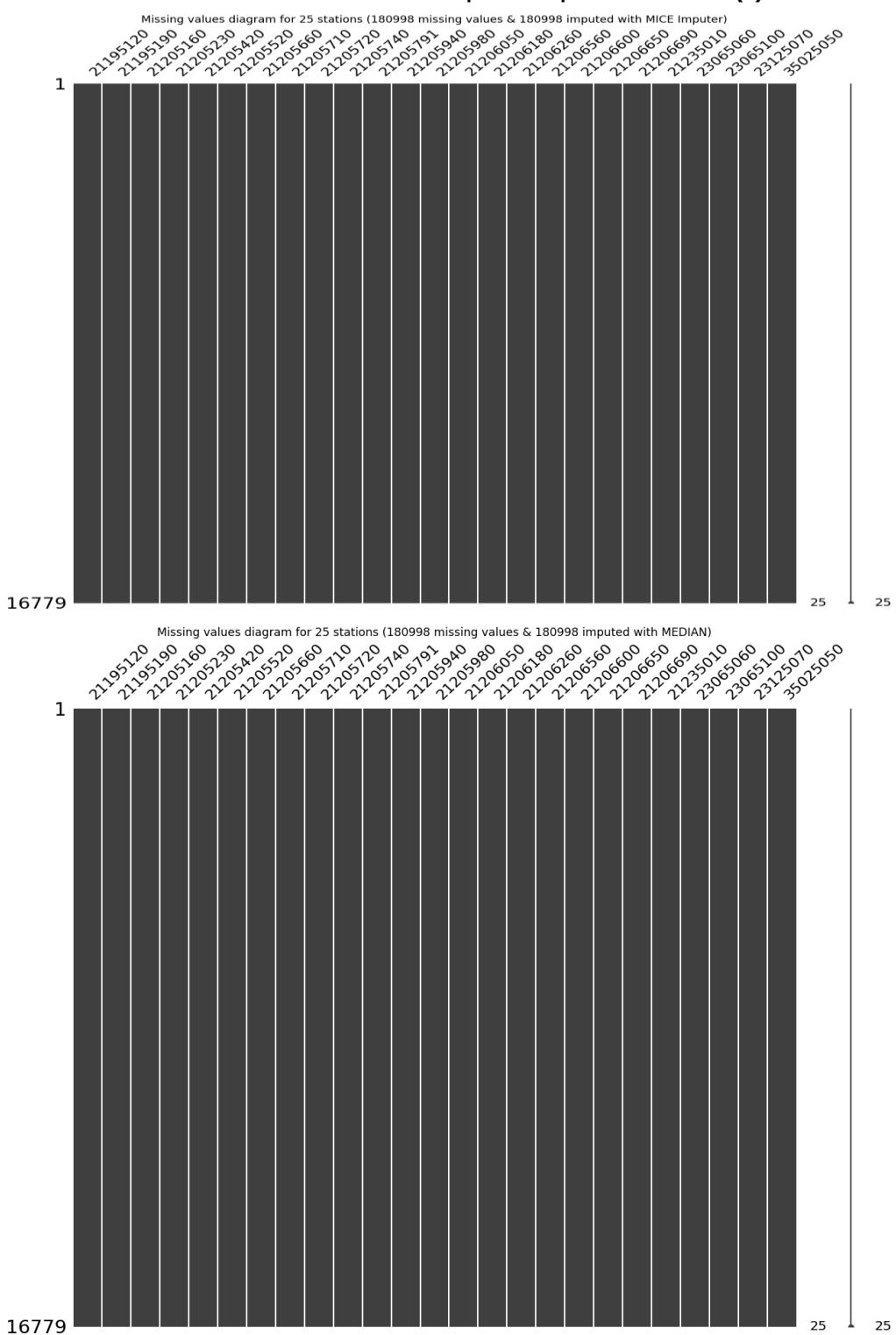


Fuente: Elaboración Propia, 2023.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

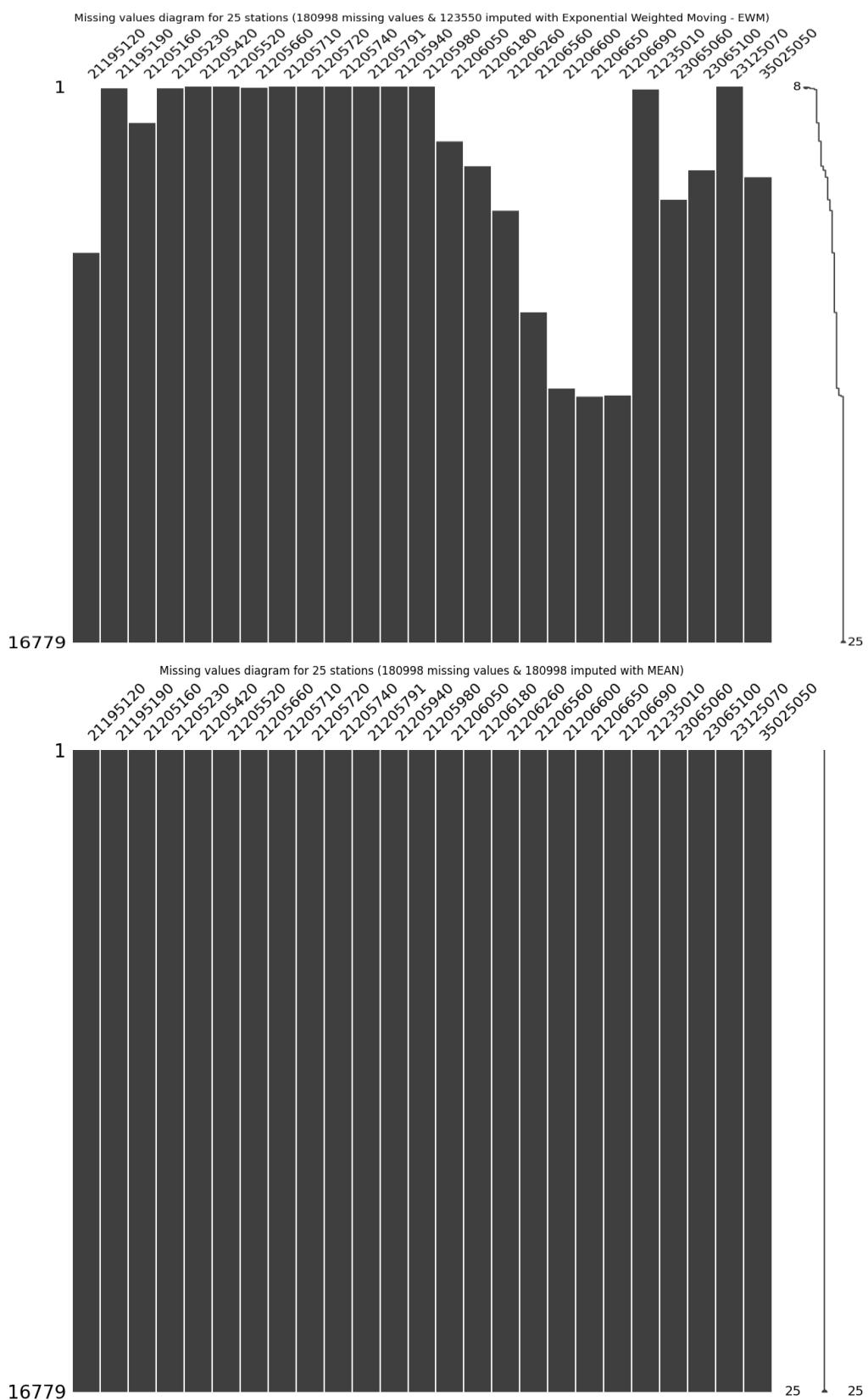
JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Ilustración 3-16. Resultados Impute Temperatura Mínima (2)



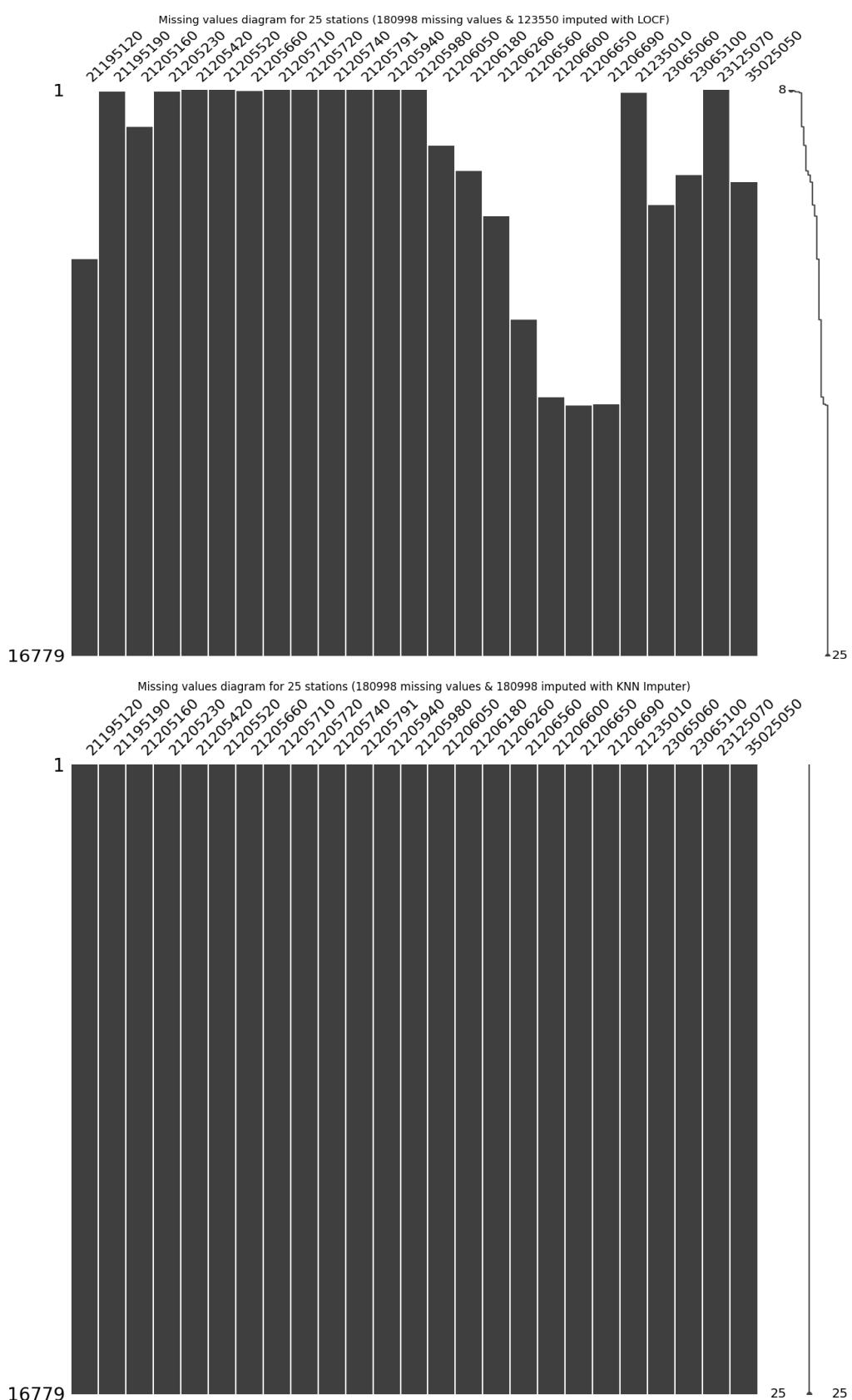
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



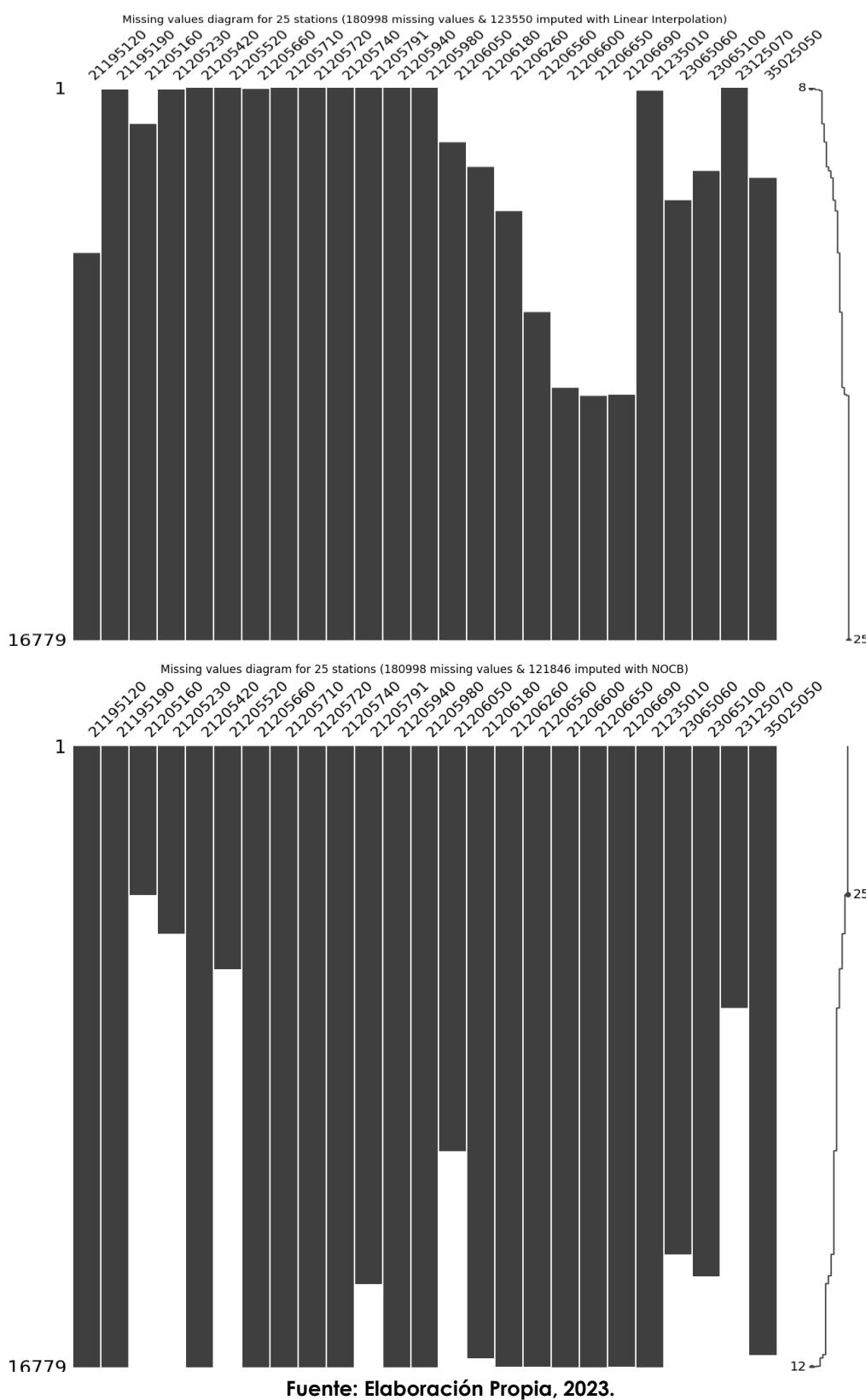
SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021



SECTION 03
 DESCARGA, PROCESAMIENTO
 Y ANÁLISIS DE DATOS
 HIDROCLIMATOLÓGICOS
 En resumen, se tiene que:

JORGE LUIS GONZÁLEZ CASTRO
 CC: 1032395475
 CS2021

Tabla 3-1. Resultados imputación de datos

Parámetro	Precipitación	Caudal	Evaporación	T. Mínima	T. Máxima
Estaciones	221	38	7	25	25
Registros	516	516	6574	16779	16781
Faltantes	41464	8159	4544	180998	198157
M1	41464	8159	4544	180998	198157
M2	41464	8159	4544	180998	198157
M3	32996	5679	2429	123550	126940
M4	12873	3731	2115	121846	138775
M5	32996	5679	2429	123550	126940
M6	32996	5679	2429	123550	126940
M7	41464	8159	4544	180998	198157
M8	41464	8159	4544	180998	198157

Fuente: Elaboración Propia, 2023.

4. ACTIVIDAD 2: SENSIBILIDAD PARÁMETROS ESTADÍSTICOS

En las siguientes ilustraciones se observa el resultado de las medias y desviaciones estándar de los resultados del ejercicio del numeral anterior, donde se compara en cada estación las series con datos imputados por cada uno de los métodos usados. A estos valores estadísticos se les calcula la desviación estándar para conocer la dispersión entre sí y valorar cuales estaciones no fueron consistentes en su llenado.

Se esperaría que la desviación estándar de los resultados entre los métodos sea similar o baja para determinar si se puede considerar consistente el llenado.

Ilustración 4-1. Resultados Precipitación

Estación	Serie Original		M1		M2		M3		M4		M5		M6		M7		M8		Análisis Desviación Estándar	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
21170020	100.962	77.354	100.962	72.868	98.825	73.116	101.331	77.157	91.635	77.699	100.962	77.048	100.930	77.056	100.316	75.134	100.748	74.161	3.255775958	1.96135051
21190030	96.646	66.488	96.646	60.825	95.108	60.925	96.769	66.460	89.485	62.972	96.624	66.413	96.650	66.411	95.674	63.560	95.710	64.260	2.440470139	2.34655082
21190060	131.326	90.194	131.326	49.481	132.145	49.484	59.396	70.535	147.016	93.023	59.396	70.535	133.142	50.691	128.772	68.246	132.418	57.147	35.02990095	15.0729075
21190210	121.431	100.143	121.431	98.871	120.939	98.919	123.246	100.822	120.595	100.476	121.920	99.433	122.972	99.674	121.328	99.773	122.184	99.723	0.939062517	0.67984391
21190300	174.428	153.124	174.428	130.664	163.839	131.813	212.025	144.871	173.639	151.200	209.373	143.202	196.337	136.515	171.490	141.868	178.223	137.608	18.40009477	6.91092032
21190310	142.807	89.402	142.807	87.914	142.188	87.978	143.846	89.583	142.883	93.161	144.189	90.335	143.042	88.668	142.495	89.028	142.941	88.916	0.663419591	1.69756646
21190460	97.573	57.197	97.573	43.509	92.496	43.916	97.405	57.175	65.334	58.493	97.726	57.163	97.532	57.106	91.248	50.600	93.981	51.508	10.94494131	6.09232562
21190470	76.261	48.122	76.261	26.400	73.114	26.482	39.656	51.545	31.980	48.861	39.656	51.545	63.631	37.066	70.593	37.893	74.805	35.801	18.44134179	10.2892037
21195030	87.462	58.519	87.462	34.500	82.603	34.683	108.680	39.781	105.554	74.388	111.762	41.774	101.222	36.723	85.059	50.641	88.628	48.093	11.6686505	13.2438999
21195120	115.898	70.895	115.896	67.799	115.022	67.860	114.250	71.115	116.466	70.129	115.619	69.625	115.443	68.382	113.930	69.308	114.890	68.956	0.840194185	1.14564981
21195140	56.483	61.441	56.483	35.611	44.862	36.564	30.906	53.415	45.589	45.268	30.906	53.415	51.705	45.693	55.504	45.974	58.545	41.385	14.67212671	6.70415223
21195190	74.679	46.463	74.679	46.010	74.476	46.032	74.873	46.278	74.973	46.365	74.923	46.136	74.753	46.023	74.232	46.231	74.442	46.074	0.263363903	0.13346994
21200080	106.725	60.891	106.725	50.413	103.876	50.589	161.335	108.561	105.009	60.420	161.944	107.869	131.092	66.432	107.131	55.824	108.491	55.003	25.26426598	24.5207709
21200110	65.473	40.510	65.473	33.586	63.048	33.779	84.445	47.609	67.128	42.669	85.346	46.984	83.498	45.159	63.870	37.196	65.300	36.202	10.15664209	5.87076883
21200130	105.287	59.623	105.287	50.674	103.683	50.740	137.616	75.428	105.686	59.431	138.166	74.736	124.361	60.172	104.950	56.290	106.220	55.841	15.20610811	9.69474783
21200160	94.210	63.143	94.210	58.497	93.006	58.572	94.105	62.755	89.455	60.576	94.141	62.645	94.105	61.465	91.490	62.274	61.826	1.707188737	1.73557428	
21200190	118.996	72.317	118.996	57.797	116.384	57.902	108.717	64.364	116.827	72.721	106.742	63.429	114.415	58.608	119.706	69.468	122.408	68.134	5.399198696	5.70982552

Fuente: Elaboración Propia.

Se tienen 38 estaciones de precipitación donde la variación entre la estadística indica que no fueron consistentes entre los métodos el llenado de datos.

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
CC: 1032395475
CS2021

Ilustración 4-2. Estaciones Parciales Precipitación No Consistentes

Estación	Serie Original		M1		M2		M3		M4		M5		M6		M7		M8		Análisis Desviación Estándar	
	meai	std	meai	std	meai	std	meai	std	meai	std	meai	std	meai	std	meai	std	meai	std	mean	std
21190060	131.326	90.194	131.326	49.481	132.145	49.484	59.396	70.535	147.016	93.023	59.396	70.535	133.142	50.691	128.772	68.246	132.418	57.147	35.02990095	15.0729075
21190470	76.261	48.122	76.261	26.400	73.114	26.482	39.656	51.545	31.980	48.861	39.656	51.545	63.631	37.065	70.593	37.893	74.805	35.801	18.44134179	10.2892037
21195038	87.462	58.519	87.462	34.500	82.603	34.683	108.680	39.781	105.554	74.388	111.762	41.774	101.222	36.723	85.059	50.641	88.628	48.093	11.6686505	13.2438999
21200080	106.725	60.891	106.725	50.413	103.876	50.589	161.335	108.561	105.009	60.420	161.944	107.869	131.092	66.432	107.131	55.824	108.491	55.003	25.26426598	24.5207709
21201110	100.327	77.277	100.327	43.873	86.579	44.894	70.071	49.335	96.240	70.140	76.866	49.020	73.699	47.691	92.674	56.193	100.578	47.376	13.22046296	10.7456117
21201250	62.662	38.203	62.662	31.674	60.777	31.798	100.183	76.122	61.010	37.840	100.465	75.716	76.119	41.183	61.353	34.764	62.821	33.732	17.49363423	19.1256966
21201260	141.252	96.316	141.252	46.299	134.727	46.438	152.868	52.427	136.438	93.061	150.295	47.953	134.249	46.642	136.518	77.438	153.385	85.812	8.360791982	19.9415937
21201330	77.704	52.216	77.704	25.414	72.370	25.589	57.543	28.342	73.888	52.627	57.188	28.028	72.101	25.948	71.146	41.627	81.122	46.837	8.682038031	10.9945138
21201374	128.578	96.087	128.578	48.090	125.902	48.115	296.460	108.734	128.578	96.087	296.460	108.734	128.991	48.091	125.971	73.139	129.430	61.060	78.03479339	26.9432284
21201770	18.281	49.128	18.281	36.675	10.204	37.784	53.739	66.354	33.136	49.722	53.739	66.354	55.648	68.607	17.708	39.296	18.522	37.979	19.09293639	14.4720109
21202000	81.356	58.976	81.356	33.684	76.140	33.881	119.434	60.441	77.086	39.663	119.434	60.441	77.939	43.718	78.428	49.683	79.185	47.608	19.07706271	10.535884
21202124	59.068	45.446	59.068	38.780	56.908	38.942	96.844	73.103	59.068	45.446	96.844	73.103	81.070	52.978	59.266	42.896	59.639	41.388	17.67777365	14.4635104
21205230	81.394	56.090	81.394	31.940	73.136	32.453	82.814	31.956	81.394	56.090	82.814	31.956	82.042	31.944	79.352	47.093	87.932	51.289	4.132667108	10.3469042
21206080	127.452	87.617	127.452	48.067	122.776	48.166	58.199	71.304	128.358	82.685	57.934	71.155	119.294	48.507	123.475	74.022	136.697	63.309	32.01654044	13.6083567
21206180	86.179	60.547	86.179	52.824	83.371	53.063	86.805	66.003	114.239	87.909	87.752	64.193	88.830	61.067	83.090	57.884	83.951	57.006	10.30161568	11.3198882
21250570	110.108	82.560	110.108	54.933	99.814	55.699	106.706	57.537	129.399	87.507	106.706	57.537	99.801	58.308	106.139	70.673	110.466	70.076	9.3058956	11.3457937
23060050	232.225	138.577	232.225	66.613	234.355	66.624	215.936	67.215	232.225	138.577	215.936	67.215	221.677	66.866	220.180	113.1061	226.768	96.514	7.490777131	27.8497947

Fuente: Elaboración Propia.

Ilustración 4-3. Resultados Caudal

Estación	Serie Original		M1		M2		M3		M4		M5		M6		M7		M8		Análisis Desviación Estándar	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
21197080	1.855	1.567	1.855	1.147	1.670	1.164	1.854	1.462	1.951	1.226	1.891	1.436	1.823	1.412	1.778	1.328	1.785	1.300	0.084439704	0.12228844
21197110	3.168	2.272	3.168	2.252	3.159	2.253	3.134	2.267	3.151	2.261	3.142	2.261	3.147	2.258	3.137	2.266	3.138	2.265	0.011590076	0.00564785
21197150	46.722	33.494	46.722	31.483	45.632	31.626	46.712	32.054	47.514	34.407	47.154	32.100	48.094	31.846	46.940	32.742	47.738	33.276	0.759150025	0.89867264
21197210	0.120	0.262	0.120	0.113	0.085	0.115	0.134	0.149	0.073	0.124	0.121	0.136	0.105	0.314	0.130	0.188	0.136	0.2083257	0.216058284	
21207960	41.549	22.437	41.549	19.525	40.476	19.617	36.845	24.304	40.650	22.322	36.244	23.669	36.727	23.118	41.462	21.470	42.214	21.266	2.479132963	1.77310054
21209200	54.858	29.852	54.858	10.274	46.602	10.718	49.474	27.159	35.050	15.992	49.474	27.159	51.797	26.108	57.290	27.812	59.325	25.445	11.32053073	7.6817891
21209920	5.186	1.815	5.186	0.711	4.926	0.719	5.221	1.778	3.904	0.898	5.167	1.776	5.226	1.777	5.161	1.119	5.077	1.451	0.446246945	0.47513747
21227010	5.047	4.132	5.047	3.789	4.832	3.821	5.090	4.410	4.776	4.034	4.883	4.062	4.897	3.894	5.002	3.941	5.082	3.889	0.12010614	0.19754725
21237010	1148.510	410.094	1148.510	410.094	1148.510	410.094	1148.510	410.094	1148.510	410.094	1148.510	410.094	1148.510	410.094	1148.510	410.094	1148.510	410.094	0	0
21237020	1242.840	452.952	1242.840	408.072	1230.270	408.908	1206.270	415.509	1242.930	452.415	1206.090	415.418	1223.330	410.844	1234.160	445.371	1253.440	434.090	17.22994232	17.5905905
35027200	55.253	37.405	55.253	27.284	50.740	27.707	102.794	59.167	54.865	37.103	102.814	59.137	85.957	43.389	55.181	31.236	57.158	30.571	22.69665492	13.2534428
35067020	56.361	42.223	56.361	16.951	51.843	17.067	32.165	20.034	56.361	42.223	20.034	69.351	17.894	53.934	27.015	68.165	41.974	14.08491143	10.7905019	
35077070	101.498	27.046	101.497	2.064	102.125	2.065	77.667	6.636	117.166	18.259	77.669	6.647	91.860	4.711	101.497	2.064	79.877	47.528	14.40096322	15.6002062

Fuente: Elaboración Propia.

Ilustración 4-5. Resultados Evaporación

Estación	Serie Original		M1		M2		M3		M4		M5		M6		M7		M8		Análisis Desviación Estándar	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
21195190	221.429	446.829	221.429	401.271	178.585	410.696	221.429	446.829	226.377	401.399	221.429	446.829	221.429	446.829	221.429	446.829	221.429	446.829	195.798	405.756
21206930	130.541	287.193	130.541	287.193	130.541	287.193	130.541	287.193	130.541	287.193	130.541	287.193	130.541	287.193	130.541	287.193	130.541	287.193	3.03841E-14	0
21206950	160.363	440.671	160.363	427.743	151.094	429.377	157.055	433.792	157.103	435.153	157.055	433.792	170.042	436.609	158.914	433.071	157.184	428.178	5.3383535385	3.33587476
21206980	76.514	275.364	76.514	273.029	76.514	275.364	76.514	275.364	76.514	275.364	76.514	275.364	75.832							

SECTION 03
DESCARGA, PROCESAMIENTO
Y ANÁLISIS DE DATOS
HIDROCLIMATOLÓGICOS

JORGE LUIS GONZÁLEZ CASTRO
 CC: 1032395475
 CS2021

Se tienen 2 estaciones de evaporación donde la variación entre la estadística indica que no fueron consistentes entre los métodos el llenado de datos; una de ellas corresponde a zona de páramo.

Ilustración 4-6. Resultados Temperatura Mínima

Estación	Serie Original		M1		M2		M3		M4		M5		M6		M7		M8		Análisis Desviación Estándar	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
21195120	16.272	1.239	16.272	0.901	16.333	0.903	16.172	1.188	16.771	1.513	16.166	1.269	16.090	1.265	16.126	1.047	16.195	1.001	0.218227764	0.21218857
21195190	11.394	1.487	11.394	1.347	11.431	1.350	11.456	1.854	11.508	1.680	11.483	1.742	11.257	1.574	11.395	1.379	11.392	1.367	0.07708346	0.2032225
21205160	8.227	3.455	8.227	1.434	8.039	1.437	8.036	1.495	8.777	3.088	8.037	1.493	7.188	1.565	8.285	2.422	9.485	2.498	0.661847266	0.64454526
21205230	8.891	2.353	8.891	1.113	9.441	1.152	9.699	1.608	9.028	2.407	9.774	1.403	9.946	1.562	9.331	1.669	9.089	1.471	0.383142519	0.40169141
21205420	6.967	2.795	6.967	2.642	7.013	2.646	6.943	2.728	7.031	2.701	6.987	2.687	6.818	2.764	6.989	2.692	6.982	2.694	0.065661294	0.03990515
21205520	7.262	2.769	7.262	1.562	7.629	1.582	7.206	1.671	6.987	2.862	7.165	1.639	8.303	1.815	7.855	2.421	7.777	2.302	0.44393714	0.48523174
21205660	19.625	2.109	19.625	1.876	19.703	1.883	19.581	2.001	19.805	1.997	19.696	1.964	19.506	2.004	19.537	1.947	19.629	1.906	0.097384994	0.052653
21205710	8.902	1.933	8.902	1.606	8.933	1.606	8.779	2.023	9.032	1.959	8.905	1.909	8.899	1.756	8.848	1.784	8.882	1.734	0.072002631	0.15541025
21205720	7.141	1.765	7.141	1.545	7.155	1.545	6.914	1.745	7.175	1.801	7.045	1.702	6.926	1.753	7.129	1.618	7.143	1.584	0.104967295	0.10113225
21205740	7.999	2.375	7.999	2.138	8.113	2.151	8.251	2.285	7.882	2.417	8.067	2.262	8.173	2.214	7.980	2.230	7.988	2.179	0.118915101	0.08976523
21205791	8.049	2.513	8.049	2.222	8.169	2.234	7.568	2.532	8.148	2.432	7.620	2.540	8.024	2.279	8.179	2.382	8.187	2.394	0.25368101	0.12419576
21205940	6.654	2.712	6.654	2.408	6.727	2.413	6.695	2.553	6.612	2.777	6.654	2.542	6.590	2.471	6.639	2.540	6.587	2.515	0.049331953	0.11584992
21205980	7.166	2.956	7.166	2.828	7.203	2.831	7.037	2.913	7.130	2.858	7.083	2.871	6.982	2.918	7.159	2.861	7.158	2.860	0.074686907	0.03304548
21206050	8.047	2.299	8.047	1.541	8.352	1.565	7.268	1.825	7.664	2.449	7.367	1.871	7.524	1.744	8.227	2.050	8.327	2.053	0.442970741	0.29715117
21206180	14.234	0.980	14.234	0.780	14.258	0.781	14.125	1.009	14.002	1.022	14.128	1.020	14.122	0.966	14.180	0.850	14.166	0.842	0.078999212	0.10648747
21206260	7.487	2.313	7.487	1.767	7.618	1.774	7.539	2.232	7.407	2.039	7.357	2.158	7.498	2.111	7.376	2.051	7.274	2.115	0.111005607	0.17156235
21206560	10.798	1.946	10.798	1.140	10.487	1.194	10.235	1.968	10.588	1.942	9.996	1.889	10.084	1.955	10.106	1.567	10.063	1.477	0.213210026	0.32474

Fuente: Elaboración Propia.

La mayor diferencia entre las desviaciones de temperatura mínima es 0.69, lo que se considera mínimo y por ende una buena consistencia entre los métodos para todas las estaciones.

Ilustración 4-7. Resultados Temperatura Máxima

Estación	Serie Original		M1		M2		M3		M4		M5		M6		M7		M8		Análisis Desviación Estándar	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
21195120	24.884	1.708	24.884	1.208	24.942	1.210	24.693	1.708	24.631	1.542	24.797	1.657	24.789	1.539	24.830	1.433	24.741	1.449	0.100527054	0.18503887
21195190	19.626	1.564	19.626	1.314	19.618	1.314	19.459	1.697	19.461	1.473	19.467	1.539	19.549	1.668	19.564	1.392	19.646	1.390	0.078088017	0.14964939
21205160	19.878	2.320	19.878	0.963	19.813	0.963	22.336	1.540	21.407	3.191	22.336	1.545	21.688	1.322	20.543	1.857	20.303	1.551	0.10851338	0.170614518
21205230	19.973	1.591	19.973	0.758	19.994	0.759	20.692	0.996	19.946	1.737	20.687	0.979	20.119	0.814	20.268	1.344	20.090	1.336	0.306375029	0.3505812
21205420	19.625	1.470	19.625	1.377	19.622	1.377	19.682	1.426	19.631	1.433	19.656	1.411	19.635	1.394	19.622	1.404	19.637	1.407	0.020849593	0.02057307
21205520	19.517	1.418	19.517	0.764	19.505	0.764	18.942	1.179	19.736	1.412	19.039	0.978	19.156	0.973	19.789	1.263	19.661	1.298	0.328155603	0.24554109
21205660	30.241	2.016	30.241	1.758	30.231	1.758	30.430	2.093	30.557	1.995	30.493	1.999	30.370	1.886	30.178	1.837	30.165	1.855	0.150083714	0.12108487
21205710	20.347	1.510	20.347	1.261	20.302	1.262	20.606	1.624	20.436	1.545	20.521	1.479	20.494	1.398	20.309	1.353	20.365	1.341	0.110022042	0.13145368
21205720	16.733	1.728	16.733	1.469	16.752	1.470	16.809	1.622	16.691	1.570	16.757	1.577	16.716	1.573	16.675	1.572	16.748	1.542	0.042447327	0.0537048
21205740	17.140	1.951	17.140	1.755	17.152	1.755	16.951	1.973	17.408	2.065	17.179	1.914	17.147	1.815	17.149	1.828	17.181	1.841	0.123272531	0.10968678
21205791	19.336	1.474	19.336	1.303	19.328	1.303	19.599	1.415	19.383	1.419	19.586	1.408	19.393	1.330	19.446	1.401	19.444	1.393	0.103755258	0.05062182
21205940	18.406	1.319	18.406	0.668	18.402	0.668	18.244	1.303	18.918	0.868	18.505	1.252	18.355	1.179	18.308	0.904	18.336	0.796	0.210119064	0.25631954
21205980	19.731	1.648	19.731	1.435	19.747	1.436	19.660	1.597	19.950	1.733	19.628	1.607	19.599	1.626	19.609	1.582	19.656	1.518	0.115022432	0.10065482
21206050	19.789	1.874	19.789	1.246	19.740	1.247	20.245	1.480	20.825	2.444	20.283	1.442	19.638	1.362	20.015	1.688	19.946	1.542	0.38522185	0.38791136
21206180	21.172	1.350	21.172	1.058	21.183	1.058	21.395	1.335	21.322	1.343	21.155	1.271	21.160	1.198	21.121	1.161	21.103	1.153	0.070990844	0.1197319
21206260	19.575	1.525	19.575	1.139	19.498	1.143	19.655	1.537	19.519	1.293	19.638	1.456	19.620	1.408	19.516	1.294	19.487	1.310	0.067211031	0.14073172
21206560	20.064	1.601	20.064	0.895	20.070	0.895	20.420	1.466	20.367	1.471	20.373	1.495	20.475	1.444	20.444	1.078	20.66629419	0.21679753		

Fuente: Elaboración Propia.

La mayor diferencia entre las desviaciones de temperatura máxima es 0.70, lo que se considera mínimo y por ende una buena consistencia entre los métodos para todas las estaciones.

5. ACTIVIDAD 3: ANÁLISIS ESTACIÓN PRECIPITACIÓN 35070110

Para la estación de precipitación con código 35070110 que cuenta con 506 registros, se eliminaron desde el año 2002 los datos para ser generados sintéticamente y compararlos con los métodos de llenado, como se muestra a

SECTION 03
 DESCARGA, PROCESAMIENTO
 Y ANÁLISIS DE DATOS
 HIDROCLIMATOLÓGICOS
 continuación. Para el caso del método KNN se asigna el número de vecinos naturales a 10.

JORGE LUIS GONZÁLEZ CASTRO

CC: 1032395475

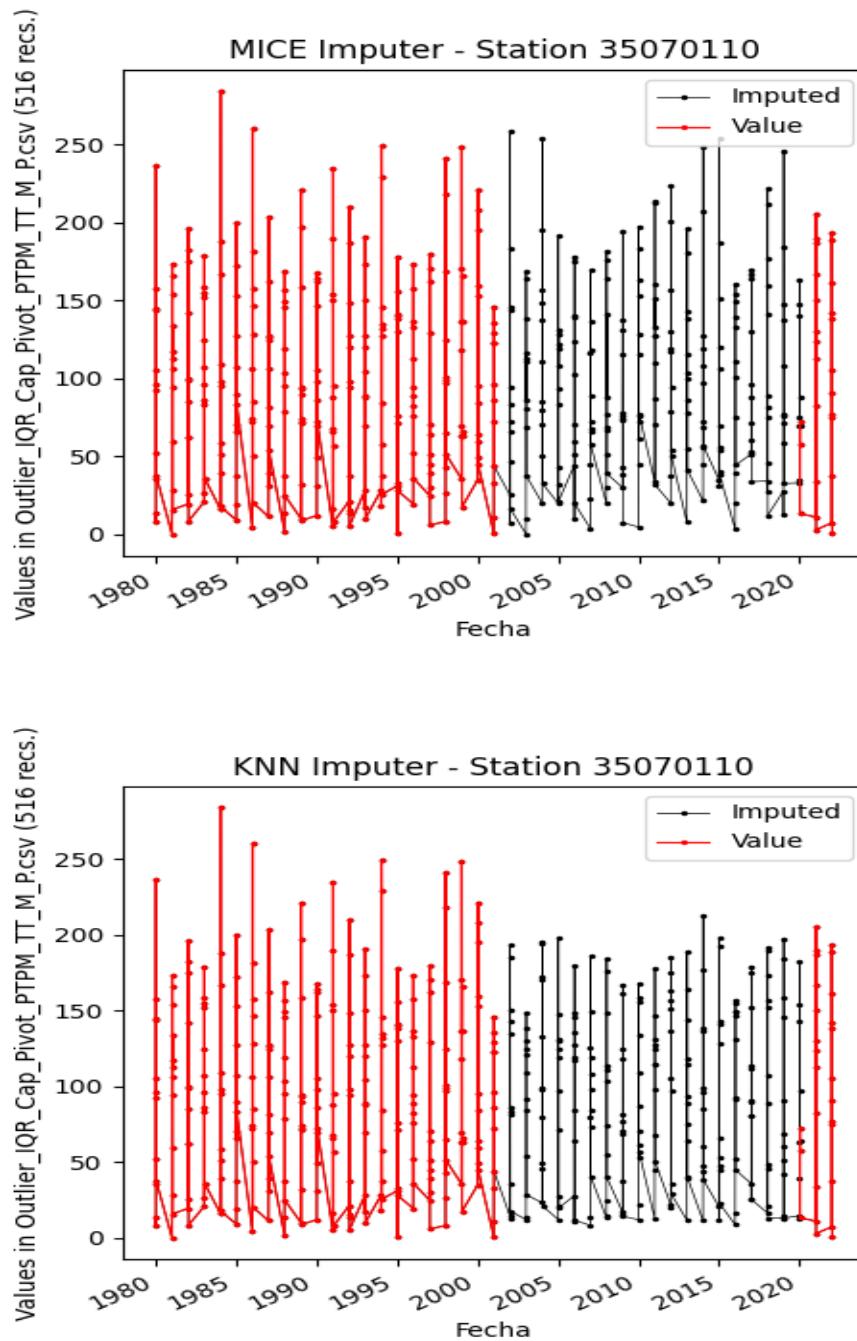
CS2021

Ilustración 5-1. Ejecución script Impute.py Estación 35070110

```

C:\JLGC\R.LTWB\src\Impute.py - Notepad++
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
Agg.py Impute.py
16     from sklearn.impute import KNNImputer
17     from sklearn.experimental import enable_iterative_imputer
18     from sklearn.impute import IterativeImputer
19     from sklearn import linear_model
20     import tabulate # required for print tables in Markdown using pandas
21     from datetime import datetime
22
23
24     # General variables
25     pivot_table_name = 'Outlier_IQR_Cap_Pivot_PTPM_TT_M_P.csv' # <<<< Pivot table name to process
26     path_input = 'C:/JLGC/R.LTWB/.datasets/IDEAM_Outlier/' # Current location from pivot tables
27     station_file = path_input + pivot_table_name # Current pivot IDEAM records file for a specified parameter
28     path = 'C:/JLGC/R.LTWB/.datasets/IDEAM_ImputeP/' # Your local output path, use ../../datasets/IDEAM_Impute/ for relative
29     file_log_name = path + ' /Impute_ ' + pivot_table_name + '.md' # Markdown file log
30     file_log = open(file_log_name, 'w+') # w+ create the file if it doesn't exist
31     station_file_log_name = path + ' /Impute_Station_ ' + pivot_table_name + '.md' # Markdown file log
32     date_record_name = 'Fecha' # IDEAM date field name for the record values
33     plot_colormap = 'autumn' # Color theme for plot graphics, https://matplotlib.org/stable/tutorials/colors/colormaps.
34     sample_records = 3 # Records to show in the sample table head and tail
35     fig_size = 5 # Height size for figures plot
36     fig_alpha = 0.75 # Alpha transparency color in plots
37     print_table_sample = True
38     show_plot = False
39     plot_stations = True # True: plot individual graphs for each station and update the complementary report
40     min_value = 0 # Minimum value for impute with Multivariate Imputation by Chained Equation - MICE from Scikit Learn.
41     n_neighbors = 10 # Number of natural neighbors for Natural Neighbors - KNN & Multivariate Imputation by Chained Equ
42     only_included = False # True: let the user run this script only for the stations included in the station_include ar
43     station_exclude = ['21200710', '21201090', '21201130', '21201380', '21201620', '21201720', '21201750', '21201760', '21201770', '21201780', '21201790', '21201800', '21201840', '21206070', '21206080', '21206090', '21206100', '21206110', '21206120', '21206130', '21206140', '21206150', '21206160', '21206170', '21206180', '21206190', '21206200', '21206210', '21206220', '21206230', '21206240', '21206250', '21206260', '21206270', '21206280', '21206290', '21206300', '21206310', '21206320', '21206330', '21206340', '21206350', '21206360', '21206370', '21206380', '21206390', '21206400', '21206410', '21206420', '21206430', '21206440', '21206450', '21206460', '21206470', '21206480', '21206490', '21206500', '21206510', '21206520', '21206530', '21206540', '21206550', '21206560', '21206570', '21206580', '21206590', '21206600', '21206610', '21206620', '21206630', '21206640', '21206650', '21206660', '21206670', '21206680', '21206690', '21206700', '21206710', '21206720', '21206730', '21206740', '21206750', '21206760', '21206770', '21206780', '21206790', '21206800', '21206810', '21206820', '21206830', '21206840', '21206850', '21206860', '21206870', '21206880', '21206890', '21206900', '21206910', '21206920', '21206930', '21206940', '21206950', '21206960', '21206970', '21206980', '21206990', '21207000', '21207010', '21207020', '21207030', '21207040', '21207050', '21207060', '21207070', '21207080', '21207090', '21207100', '21207110', '21207120', '21207130', '21207140', '21207150', '21207160', '21207170', '21207180', '21207190', '21207200', '21207210', '21207220', '21207230', '21207240', '21207250', '21207260', '21207270', '21207280', '21207290', '21207300', '21207310', '21207320', '21207330', '21207340', '21207350', '21207360', '21207370', '21207380', '21207390', '21207400', '21207410', '21207420', '21207430', '21207440', '21207450', '21207460', '21207470', '21207480', '21207490', '21207500', '21207510', '21207520', '21207530', '21207540', '21207550', '21207560', '21207570', '21207580', '21207590', '21207600', '21207610', '21207620', '21207630', '21207640', '21207650', '21207660', '21207670', '21207680', '21207690', '21207700', '21207710', '21207720', '21207730', '21207740', '21207750', '21207760', '21207770', '21207780', '21207790', '21207800', '21207810', '21207820', '21207830', '21207840', '21207850', '21207860', '21207870', '21207880', '21207890', '21207900', '21207910', '21207920', '21207930', '21207940', '21207950', '21207960', '21207970', '21207980', '21207990', '21208000', '21208010', '21208020', '21208030', '21208040', '21208050', '21208060', '21208070', '21208080', '21208090', '21208100', '21208110', '21208120', '21208130', '21208140', '21208150', '21208160', '21208170', '21208180', '21208190', '21208200', '21208210', '21208220', '21208230', '21208240', '21208250', '21208260', '21208270', '21208280', '21208290', '21208300', '21208310', '21208320', '21208330', '21208340', '21208350', '21208360', '21208370', '21208380', '21208390', '21208400', '21208410', '21208420', '21208430', '21208440', '21208450', '21208460', '21208470', '21208480', '21208490', '21208500', '21208510', '21208520', '21208530', '21208540', '21208550', '21208560', '21208570', '21208580', '21208590', '21208600', '21208610', '21208620', '21208630', '21208640', '21208650', '21208660', '21208670', '21208680', '21208690', '21208700', '21208710', '21208720', '21208730', '21208740', '21208750', '21208760', '21208770', '21208780', '21208790', '21208800', '21208810', '21208820', '21208830', '21208840', '21208850', '21208860', '21208870', '21208880', '21208890', '21208900', '21208910', '21208920', '21208930', '21208940', '21208950', '21208960', '21208970', '21208980', '21208990', '21209000', '21209010', '21209020', '21209030', '21209040', '21209050', '21209060', '21209070', '21209080', '21209090', '21209100', '21209110', '21209120', '21209130', '21209140', '21209150', '21209160', '21209170', '21209180', '21209190', '21209200', '21209210', '21209220', '21209230', '21209240', '21209250', '21209260', '21209270', '21209280', '21209290', '21209300', '21209310', '21209320', '21209330', '21209340', '21209350', '21209360', '21209370', '21209380', '21209390', '21209400', '21209410', '21209420', '21209430', '21209440', '21209450', '21209460', '21209470', '21209480', '21209490', '21209500', '21209510', '21209520', '21209530', '21209540', '21209550', '21209560', '21209570', '21209580', '21209590', '21209600', '21209610', '21209620', '21209630', '21209640', '21209650', '21209660', '21209670', '21209680', '21209690', '21209700', '21209710', '21209720', '21209730', '21209740', '21209750', '21209760', '21209770', '21209780', '21209790', '21209800', '21209810', '21209820', '21209830', '21209840', '21209850', '21209860', '21209870', '21209880', '21209890', '21209900', '21209910', '21209920', '21209930', '21209940', '21209950', '21209960', '21209970', '21209980', '21209990', '21209999', '21210000', '21210001', '21210002', '21210003', '21210004', '21210005', '21210006', '21210007', '21210008', '21210009', '21210010', '21210011', '21210012', '21210013', '21210014', '21210015', '21210016', '21210017', '21210018', '21210019', '21210020', '21210021', '21210022', '21210023', '21210024', '21210025', '21210026', '21210027', '21210028', '21210029', '21210030', '21210031', '21210032', '21210033', '21210034', '21210035', '21210036', '21210037', '21210038', '21210039', '21210040', '21210041', '21210042', '21210043', '21210044', '21210045', '21210046', '21210047', '21210048', '21210049', '21210050', '21210051', '21210052', '21210053', '21210054', '21210055', '21210056', '21210057', '21210058', '21210059', '21210060', '21210061', '21210062', '21210063', '21210064', '21210065', '21210066', '21210067', '21210068', '21210069', '21210070', '21210071', '21210072', '21210073', '21210074', '21210075', '21210076', '21210077', '21210078', '21210079', '21210080', '21210081', '21210082', '21210083', '21210084', '21210085', '21210086', '21210087', '21210088', '21210089', '21210090', '21210091', '21210092', '21210093', '21210094', '21210095', '21210096', '21210097', '21210098', '21210099', '212100999', '212101000', '212101001', '212101002', '212101003', '212101004', '212101005', '212101006', '212101007', '212101008', '212101009', '212101010', '212101011', '212101012', '212101013', '212101014', '212101015', '212101016', '212101017', '212101018', '212101019', '212101020', '212101021', '212101022', '212101023', '212101024', '212101025', '212101026', '212101027', '212101028', '212101029', '212101030', '212101031', '212101032', '212101033', '212101034', '212101035', '212101036', '212101037', '212101038', '212101039', '212101040', '212101041', '212101042', '212101043', '212101044', '212101045', '212101046', '212101047', '212101048', '212101049', '212101050', '212101051', '212101052', '212101053', '212101054', '212101055', '212101056', '212101057', '212101058', '212101059', '212101060', '212101061', '212101062', '212101063', '212101064', '212101065', '212101066', '212101067', '212101068', '212101069', '212101070', '212101071', '212101072', '212101073', '212101074', '212101075', '212101076', '212101077', '212101078', '212101079', '212101080', '212101081', '212101082', '212101083', '212101084', '212101085', '212101086', '212101087', '212101088', '212101089', '212101090', '212101091', '212101092', '212101093', '212101094', '212101095', '212101096', '212101097', '212101098', '212101099', '2121010999', '2121011000', '2121011001', '2121011002', '2121011003', '2121011004', '2121011005', '2121011006', '2121011007', '2121011008', '2121011009', '2121011010', '2121011011', '2121011012', '2121011013', '2121011014', '2121011015', '2121011016', '2121011017', '2121011018', '2121011019', '2121011020', '2121011021', '2121011022', '2121011023', '2121011024', '2121011025', '2121011026', '2121011027', '2121011028', '2121011029', '2121011030', '2121011031', '2121011032', '2121011033', '2121011034', '2121011035', '2121011036', '2121011037', '2121011038', '2121011039', '2121011040', '2121011041', '2121011042', '2121011043', '2121011044', '2121011045', '2121011046', '2121011047', '2121011048', '2121011049', '2121011050', '2121011051', '2121011052', '2121011053', '2121011054', '2121011055', '2121011056', '2121011057', '2121011058', '2121011059', '2121011060', '2121011061', '2121011062', '2121011063', '2121011064', '2121011065', '2121011066', '2121011067', '2121011068', '2121011069', '2121011070', '2121011071', '2121011072', '2121011073', '2121011074', '2121011075', '2121011076', '2121011077', '2121011078', '2121011079', '2121011080', '2121011081', '2121011082', '2121011083', '2121011084', '2121011085', '2121011086', '2121011087', '2121011088', '2121011089', '2121011090', '2121011091', '2121011092', '2121011093', '2121011094', '2121011095', '2121011096', '2121011097', '2121011098', '2121011099', '21210110999', '21210111000', '21210111001', '21210111002', '21210111003', '21210111004', '21210111005', '21210111006', '21210111007', '21210111008', '21210111009', '21210111010', '21210111011', '21210111012', '21210111013', '21210111014', '21210111015', '21210111016', '21210111017', '21210111018', '21210111019', '21210111020', '21210111021', '21210111022', '21210111023', '21210111024', '21210111025', '21210111026', '21210111027', '21210111028', '21210111029', '21210111030', '21210111031', '21210111032', '21210111033', '21210111034', '21210111035', '21210111036', '21210111037', '21210111038', '21210111039', '21210111040', '21210111041', '21210111042', '21210111043', '21210111044', '21210111045', '21210111046', '21210111047', '21210111048', '21210111049', '21210111050', '21210111051', '21210111052', '21210111053', '21210111054', '21210111055', '21210111056', '21210111057', '21210111058', '21210111059', '21210111060', '21210111061', '21210111062', '21210111063', '21210111064', '21210111065', '21210111066', '21210111067', '21210111068', '21210111069', '21210111070', '21210111071', '21210111072', '21210111073', '21210111074', '21210111075', '21210111076', '21210111077', '21210111078', '21210111079', '21210111080', '21210111081', '21210111082', '21210111083', '21210111084', '21210111085', '21210111086', '21210111087', '21210111088', '21210111089', '21210111090', '21210111091', '21210111092', '21210111093', '21210111094', '21210111095', '21210111096', '21210111097', '21210111098', '21210111099', '212101110999', '212101111000', '212101111001', '212101111002', '212101111003', '212101111004', '212101111005', '212101111006', '212101111007', '212101111008', '212101111009', '212101111010', '212101111011', '212101111012', '212101111013', '212101111014', '212101111015', '212101111016', '212101111017', '212101111018', '212101111019', '212101111020', '212101111021', '212101111022', '212101111023', '212101111024', '212101111025', '212101111026', '212101111027', '212101111028', '212101111029', '212101111030', '212101111031', '212101111032', '212101111033', '212101111034', '212101111035', '212101111036', '212101111037', '212101111038', '212101111039', '212101111040', '212101111041', '212101111042', '212101111043', '212101111044', '212101111045', '212101111046', '212101111047', '212101111048', '212101111049', '212101111050', '212101111051', '212101111052', '212101111053', '212101111054', '212101111055', '212101111056', '212101111057', '212101111058', '212101111059', '212101111060', '212101111061', '212101111062', '212101111063', '212101111064', '212101111065', '212101111066', '212101111067', '212101111068', '212101111069', '212101111070', '212101111071', '212101111072', '212101111073', '212101111074', '212101111075', '212101111076', '212101111077', '212101111078', '212101111079', '212101111080', '212101111081', '212101111082', '212101111083', '212101111084', '212101111085', '212101111086', '212101111087', '212101111088', '212101111089', '212101111090', '212101111091', '212101111092', '212101111093', '212101111094', '212101111095', '212101111096', '212101111097', '212101111098', '212101111099', '2121011110999', '2121011111000', '2121011111001', '2121011111002', '2121011111003', '2121011111004', '2121011111005', '2121011111006', '2121011111007', '2121011111008', '2121011111009', '2121011111010', '2121011111011', '2121011111012', '2121011111013', '2121011111014', '212
```

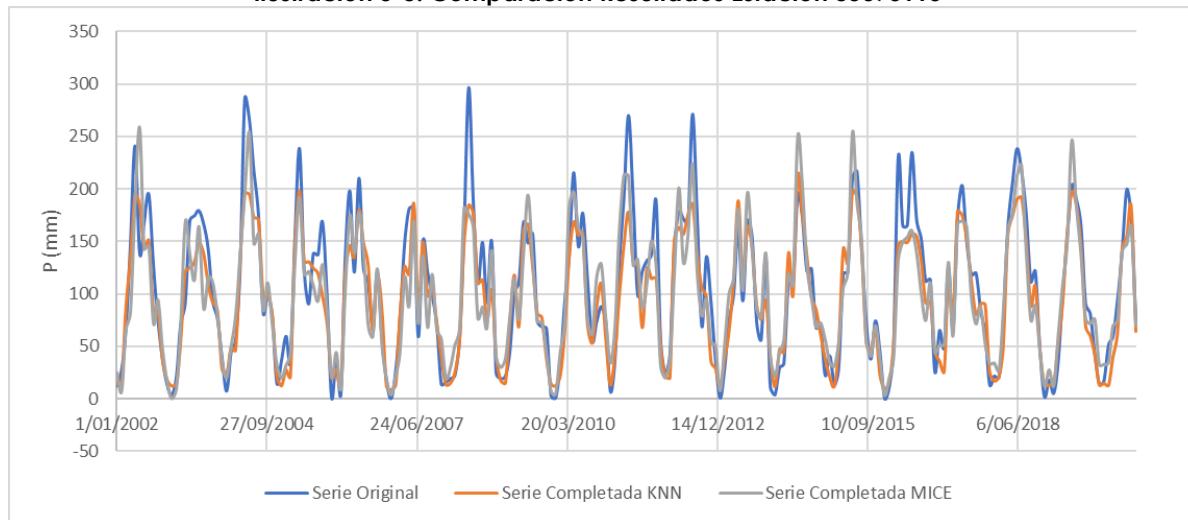
Ilustración 5-2. Resultados método KNNN y MICE Estación 35070110



Fuente: Elaboración Propia.

En la siguiente gráfica se observa que para los valores altos de la serie original no coinciden en de manera exacta con las series sintéticas, mientras que para valores medios y bajos si se tiene una tendencia similar.

Ilustración 5-3. Comparación Resultados Estación 35070110



Fuente: Elaboración Propia.

Comparándose la estadística se tiene que:

Tabla 5-1. Comparación estadística estación 35070110

	Serie Original	Serie Completada KNN	Serie Completada MICE
Media	103.90	93.78	97.76
Desviación Estándar	69.07	57.58	59.49

Fuente: Elaboración Propia.

Se observa que respecto a la serie original los valores de la media coinciden más con el método MICE, mientras que la desviación estándar es similar en los dos métodos usados, pero difiere bastante de la serie original.

Lo anterior indica que en casos extremos donde una estación no tenga datos registrados en períodos de tiempo seguidos y extensos puede usarse uno de estos métodos para completarla, pero con una alta incertidumbre especialmente en datos altos.

6. CONCLUSIONES

- Se realizó el complemento de datos con los 8 métodos del script Impute.py.
- Se eliminaron las series de estaciones con muy pocos datos.
- Se realizó el análisis de los resultados de cada parámetro donde se observa que en no todos los casos los datos sintéticos entre métodos son consistentes entre sí.
- Se realizó el análisis de una estación de precipitación con datos completos suponiendo que la serie no lo estaba en un periodo seguido de 10 años para comprar los datos, donde se evidencio que los métodos MICE y KNN pueden ser utilizados en estos casos.

7. REFERENCIAS BIBLIOGRÁFICAS

- RCFDTOOLS, 2023. Balance hidrológico de largo plazo para estimación de caudales medios usando SIG. Contenido del curso: <https://github.com/rcfdtools/R.LTWB/blob/main/Section03/Impute>.