

Informe de control de calidad de datos climáticos - Parcela 3

Ing. Miguel Silva

2025-12-06

Table of contents

1	Introducción	1
2	Descripción general de la base de datos	1
3	Análisis exploratorio	5
4	Control de calidad	6
5	Resultados	9
6	Visualización gráfica de las variables	13

1 Introducción

En este informe se presenta el control de calidad de la base de datos climáticos registrada en la Parcela 3, a partir del archivo **Datos_Parcela_3.xlsx**.

El objetivo principal es identificar y cuantificar:

- Datos faltantes por variable.
- Valores fuera de rangos físicos razonables.
- Registros atípicos que puedan comprometer análisis posteriores (balance de energía, estimación de ET, etc.).
- Visualizar el comportamiento temporal de las variables.

2 Descripción general de la base de datos

La base de datos esta conformada por los registros horarios de las siguientes variables:

Variables meteorológicas:

- Temperatura (°C)
- Humedad relativa (%)

- Velocidad del viento (m/s)
- Dirección del viento (°)
- Presión atmosférica (mbar)
- Radiación global (Wm-2)
- Precipitación (mm)

Variables de flujos de energía:

- Temperatura superficial (°C)
- Flujo de energía en el suelo (Wm-2)
- Radiación Neta (Wm-2)
- Temperatura termocupla 1 (°C)
- Temperatura termocupla 2 (°C)
- Humedad del suelo a los 5 cm (m³/m³)
- Humedad del suelo a los 40 cm (m³/m³)

Para su procesamiento preliminar se aplica lo siguiente:

```
# Cargar paquetes
library(readxl)
library(dplyr)
```

Adjuntando el paquete: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(lubridate)
```

Adjuntando el paquete: 'lubridate'

The following objects are masked from 'package:base':

date, intersect, setdiff, union

```
library(ggplot2)
library(tidyr)
library(purrr)
library(readr)
library(ggpmisc)
```

Cargando paquete requerido: ggpp

```
Registered S3 methods overwritten by 'ggpp':  
  method           from  
heightDetails.titleGrob  ggplot2  
widthDetails.titleGrob   ggplot2
```

Adjuntando el paquete: 'ggpp'

The following object is masked from 'package:ggplot2':

```
annotate  
  
library(openair)  
  
# Lectura del archivo original  
datos_brutos <- readxl::read_excel("Datos_Parcela_3.xlsx")  
  
# Renombrar variables según indicaciones  
datos_brutos <- datos_brutos %>%  
  rename(  
    Radiacion_global_Wm2      = SlrFD_W_Avg,  
    Precipitacion_mm          = Rain_mm_Tot,  
    Rayos_conteo              = Strikes_Tot,  
    Velocidad_viento_ms       = WS_ms_Avg,  
    Direccion_viento_grados  = WindDir,  
    Rachas_ms                 = MaxWS_ms_Avg,  
    Temperatura_aire_C        = AirT_C_Avg,  
    Presion_atmosferica_mbar = BP_mbar_Avg,  
    Humedad_relativa_porc    = RH,  
    Radiacion_global_MJm2     = SlrTF_MJ_Tot,  
    Temperatura_superficie_C = SBT_C_Avg,  
    Flujo_energia_suelo_Wm2  = SHF_Avg,  
    Radiacion_neta_Wm2        = NR_Wm2_Avg,  
    Temperatura_termocupla1_C = Temp_TC1_C_Avg,  
    Temperatura_termocupla2_C = Temp_TC2_C_Avg,  
    Humedad_suelo_5cm_m3m3   = SWC_5CM_Avg,  
    Humedad_suelo_40cm_m3m3  = SWC_40CM_Avg  
) %>%  
  mutate(  
    fecha_hora = ymd_hms(fecha_hora),  
    fecha      = as_date(fecha)  
) %>%  
# Forzar variables climáticas a numérico  
  mutate(  
    Radiacion_global_Wm2      = as.numeric(Radiacion_global_Wm2),  
    Precipitacion_mm          = as.numeric(Precipitacion_mm),  
    Rayos_conteo              = as.numeric(Rayos_conteo),  
    Velocidad_viento_ms       = as.numeric(Velocidad_viento_ms),  
    Direccion_viento_grados  = as.numeric(Direccion_viento_grados),  
    Rachas_ms                 = as.numeric(Rachas_ms),  
    Temperatura_aire_C        = as.numeric(Temperatura_aire_C),  
    Presion_atmosferica_mbar = as.numeric(Presion_atmosferica_mbar),  
    Humedad_relativa_porc    = as.numeric(Humedad_relativa_porc),  
    Radiacion_global_MJm2     = as.numeric(Radiacion_global_MJm2),  
    Temperatura_superficie_C = as.numeric(Temperatura_superficie_C),
```

```

Flujo_energia_suelo_Wm2      = as.numeric(Flujo_energia_suelo_Wm2),
Radiacion_neta_Wm2           = as.numeric(Radiacion_neta_Wm2),
Temperatura_termocupla1_C    = as.numeric(Temperatura_termocupla1_C),
Temperatura_termocupla2_C    = as.numeric(Temperatura_termocupla2_C),
Humedad_suelo_5cm_m3m3       = as.numeric(Humedad_suelo_5cm_m3m3),
Humedad_suelo_40cm_m3m3      = as.numeric(Humedad_suelo_40cm_m3m3)
) %>%
arrange(fecha_hora)

Warning: There was 1 warning in `mutate()` .
i In argument: `fecha_hora = ymd_hms(fecha_hora)` .
Caused by warning:
! 105 failed to parse.

#| label: resumen-temporal
#| echo: true

datos_brutos %>%
summarise(
fecha_inicio = min(fecha_hora, na.rm = TRUE),
fecha_fin = max(fecha_hora, na.rm = TRUE),
n_registros = n()
)

# A tibble: 1 x 3
  fecha_inicio     fecha_fin   n_registros
  <dttm>          <dttm>        <int>
1 2025-07-21 14:00:00 2025-11-28 15:00:00      2438

#| label: resumen-completitud
#| echo: true

datos_brutos %>%
summarise(
n_Temperatura_aire_C = sum(!is.na(Temperatura_aire_C)),
n_Humedad_relativa_porc = sum(!is.na(Humedad_relativa_porc)),
n_Precipitacion_mm = sum(!is.na(Precipitacion_mm)),
n_Velocidad_viento_ms = sum(!is.na(Velocidad_viento_ms)),
n_Radiacion_neta_Wm2 = sum(!is.na(Radiacion_neta_Wm2)),
n_Radiacion_global_Wm2 = sum(!is.na(Radiacion_global_Wm2)),
n_Temperatura_superficie_C = sum(!is.na(Temperatura_superficie_C)),
n_Humedad_suelo_5cm_m3m3 = sum(!is.na(Humedad_suelo_5cm_m3m3)),
n_Humedad_suelo_40cm_m3m3 = sum(!is.na(Humedad_suelo_40cm_m3m3))
)

# A tibble: 1 x 9
  n_Temperatura_aire_C n_Humedad_relativa_porc n_Precipitacion_mm
  <int>                  <int>                  <int>
1          2434                  2434                  2434
# i 6 more variables: n_Velocidad_viento_ms <int>, n_Radiacion_neta_Wm2 <int>,
#   n_Radiacion_global_Wm2 <int>, n_Temperatura_superficie_C <int>,
#   n_Humedad_suelo_5cm_m3m3 <int>, n_Humedad_suelo_40cm_m3m3 <int>

```

3 Análisis exploratorio

Este análisis primero identifica el período temporal cubierto, el número de registros y variables, y determina qué columnas contienen información numérica. Luego calcula estadísticas descriptivas básicas; media, mediana, desviación estándar, mínimos, máximos y cantidad de datos faltantes para todas las variables, permitiendo detectar rangos típicos, dispersión y posibles anomalías. Finalmente, cuantifica los valores ausentes por variable, lo cual ayuda a evaluar la completitud del conjunto de datos antes de aplicar controles de calidad más estrictos.

```
analisis_exploratorio <- function(datos) {
  resultados_ae <- list()

  #Información básica (proteger si fecha_hora es todo NA)

  fecha_min <- ifelse(all(is.na(datos$fecha_hora)), NA, min(datos$fecha_hora, na.rm = TRUE))
  fecha_max <- ifelse(all(is.na(datos$fecha_hora)), NA, max(datos$fecha_hora, na.rm = TRUE))
  duracion <- ifelse(
    is.na(fecha_min) | is.na(fecha_max),
    NA,
    as.numeric(difftime(fecha_max, fecha_min, units = "days"))
  )

  resultados_ae$info_basica <- list(
    n_filas = nrow(datos),
    n_columnas = ncol(datos),
    fecha_inicio = fecha_min,
    fecha_fin = fecha_max,
    duracion_dias = duracion
  )

  #Variables numéricas disponibles

  vars_numericas <- datos %>% dplyr::select(where(is.numeric)) %>% colnames()
  resultados_ae$variables_numericas <- vars_numericas

  #Variables principales

  vars_principales <- c(
    "Radiacion_global_Wm2",
    "Precipitacion_mm",
    "Rayos_conteo",
    "Velocidad_viento_ms",
    "Direccion_viento_grados",
    "Rachas_ms",
    "Temperatura_aire_C",
    "Presion_atmosferica_mbar",
    "Humedad_relativa_porc",
    "Radiacion_global_MJm2",
    "Temperatura_superficie_C",
    "Flujo_energia_suelo_Wm2",
    "Radiacion_neta_Wm2",
    "Temperatura_termocupla1_C",
    "Temperatura_termocupla2_C",
    "Humedad_suelo_5cm_m3m3",
    "Humedad_suelo_40cm_m3m3"
  )

  #Estadísticas descriptivas
```

```

estadisticas <- list()
for (var in vars_principales) {
  if (var %in% colnames(datos)) {
    estadisticas[[var]] <- list(
      media = mean(datos[[var]], na.rm = TRUE),
      mediana = median(datos[[var]], na.rm = TRUE),
      desviacion = sd(datos[[var]], na.rm = TRUE),
      min = min(datos[[var]], na.rm = TRUE),
      max = max(datos[[var]], na.rm = TRUE),
      nas = sum(is.na(datos[[var]])))
    )
  }
}
resultados_ae$estadisticas <- estadisticas

#Conteo de valores faltantes

na_count <- colSums(is.na(datos))
resultados_ae$valores_faltantes <- na_count[na_count > 0]

return(resultados_ae)
}

```

4 Control de calidad

Las reglas de control de calidad aplicadas en este informe son:

Continuidad temporal

Verificación de que la serie está ordenada y sin duplicados en fecha_hora.

Rangos físicos razonables (pueden ajustarse a tus criterios locales):

Temperatura_aire_C: -10 a 45 °C

Humedad_relativa_porc: 0 a 100 %

Precipitacion_mm: 0 mm

Velocidad_viento_ms: 0 m/s

Radiacion_neta_Wm2: -200 a 1000 W/m²

Tratamiento de datos

Los valores fuera de rango se marcan mediante flags y, en la base validada, se reemplazan por NA para evitar su uso directo en análisis.

```

control_calidad <- function(datos) {
  resultados_cc <- list()

  #Límites físicos

  limites <- list(
    Temperatura_aire_C = c(-10, 45),
    Humedad_relativa_porc = c(0, 100),
    Velocidad_viento_ms = c(0, 50),
    Radiacion_global_Wm2 = c(0, 1500),
    Precipitacion_mm = c(0, 100),
    Humedad_suelo_5cm_m3m3 = c(0, 0.6),
    Humedad_suelo_40cm_m3m3 = c(0, 0.6)
  )

```

```

)

#Fuera de rango

fuera_rango <- list()
for (var in names(lmites)) {
  if (var %in% colnames(datos)) {
    idx <- which(datos[[var]] < lmites[[var]][1] | datos[[var]] > lmites[[var]][2])
    fuera_rango[[var]] <- list(
      n_fuera_rango = length(idx),
      porcentaje = round(length(idx) / nrow(datos) * 100, 2)
    )
  }
}
resultados_cc$fuera_rango <- fuera_rango

#Atípicos por IQR

atipicos_iqr <- list()
variables_verificar <- c(
  "Radiacion_global_Wm2",
  "Precipitacion_mm",
  "Rayos_conteo",
  "Velocidad_viento_ms",
  "Direccion_viento_grados",
  "Rachas_ms",
  "Temperatura_aire_C",
  "Presion_atmosferica_mbar",
  "Humedad_relativa_porc",
  "Radiacion_global_MJm2",
  "Temperatura_superficie_C",
  "Flujo_energia_suelo_Wm2",
  "Radiacion_neta_Wm2",
  "Temperatura_termocupla1_C",
  "Temperatura_termocupla2_C",
  "Humedad_suelo_5cm_m3m3",
  "Humedad_suelo_40cm_m3m3"
)

for (var in variables_verificar) {
  if (var %in% colnames(datos)) {
    Q1 <- quantile(datos[[var]], 0.25, na.rm = TRUE)
    Q3 <- quantile(datos[[var]], 0.75, na.rm = TRUE)
    IQR_val <- Q3 - Q1
    lower <- Q1 - 1.5 * IQR_val
    upper <- Q3 + 1.5 * IQR_val
    idx_atip <- which(datos[[var]] < lower | datos[[var]] > upper)
    atipicos_iqr[[var]] <- list(
      n_atipicos = length(idx_atip),
      porcentaje = round(length(idx_atip) / nrow(datos) * 100, 2)
    )
  }
}
resultados_cc$atipicos_iqr <- atipicos_iqr

#Completitud de variables esenciales

vars_esenciales <- c(

```

```

"Radiacion_global_Wm2",
"Precipitacion_mm",
"Rayos_conteo",
"Velocidad_viento_ms",
"Direccion_viento_grados",
"Rachas_ms",
"Temperatura_aire_C",
"Presion_atmosferica_mbar",
"Humedad_relativa_porc",
"Radiacion_global_MJm2",
"Temperatura_superficie_C",
"Flujo_energia_suelo_Wm2",
"Radiacion_neta_Wm2",
"Temperatura_termocupla1_C",
"Temperatura_termocupla2_C",
"Humedad_suelo_5cm_m3m3",
"Humedad_suelo_40cm_m3m3"
)

completitud <- sapply(vars_esenciales, function(x) {
  if (x %in% colnames(datos)) {
    round(sum(!is.na(datos[[x]])) / nrow(datos) * 100, 2)
  } else {
    NA
  }
})
resultados_cc$completitud <- completitud

#Saltos temporales distintos a 1 hora

datos2 <- datos %>%
  arrange(fecha_hora) %>%
  mutate(diff_tiempo_h = as.numeric(difftime(fecha_hora, lag(fecha_hora), units = "hours")))

saltos_tiempo <- which(datos2$diff_tiempo_h != 1 & !is.na(datos2$diff_tiempo_h))
resultados_cc$saltos_temporales <- length(saltos_tiempo)

return(resultados_cc)
}

#| label: crear-datos-qc
#| echo: true

#Rangos para flags (coherentes con la metodología)

rango_temp <- c(-10, 45)
rango_hr <- c(0, 100)
rango_pp <- c(0, Inf)
rango_viento <- c(0, Inf)
rango_rn <- c(-200, 1000)

datos_qc <- datos_brutos %>%
  mutate(
    flag_temp_rango = case_when(
      is.na(Temperatura_aire_C) ~ "na",
      Temperatura_aire_C < rango_temp[1] | Temperatura_aire_C > rango_temp[2] ~ "fuera_rango",
      TRUE ~ "ok"
    ),

```

```

flag_hr_rango = case_when(
  is.na(Humedad_relativa_porc) ~ "na",
  Humedad_relativa_porc < rango_hr[1] | Humedad_relativa_porc > rango_hr[2] ~ "fuera_rango",
  TRUE ~ "ok"
),
flag_pp_rango = case_when(
  is.na(Precipitacion_mm) ~ "na",
  Precipitacion_mm < rango_pp[1] ~ "negativa",
  TRUE ~ "ok"
),
flag_ws_rango = case_when(
  is.na(Velocidad_viento_ms) ~ "na",
  Velocidad_viento_ms < rango_viento[1] ~ "negativa",
  TRUE ~ "ok"
),
flag_rn_rango = case_when(
  is.na(Radiacion_neta_Wm2) ~ "na",
  Radiacion_neta_Wm2 < rango_rn[1] | Radiacion_neta_Wm2 > rango_rn[2] ~ "fuera_rango",
  TRUE ~ "ok"
)
)
)

```

5 Resultados

Resultados del análisis exploratorio

```

resultados_ae <- analisis_exploratorio(datos_brutos)

#Info básica en tabla

info_basica <- as.data.frame(resultados_ae$info_basica)
info_basica

  n_filas n_columnas fecha_inicio fecha_fin duracion_dias
1    2438           32   1753106400 1764342000      130.0417

#| label: estadisticas-principales
#| echo: true

estadisticas_df <- purrr::map_dfr(
  resultados_ae$estadisticas,
  ~ as.data.frame(.x),
  .id = "variable"
)

estadisticas_df

```

	variable	media	mediana	desviacion	min
1	Radiacion_global_Wm2	1.891088e+02	10.8350	2.617007e+02	0.000
2	Precipitacion_mm	2.940427e-02	0.0000	2.730409e-01	0.000
3	Rayos_conteo	2.875924e-03	0.0000	8.354031e-02	0.000
4	Velocidad_viento_ms	5.160140e-01	0.4580	3.449411e-01	0.048
5	Direccion_viento_grados	1.761297e+02	163.9000	7.951649e+01	0.900
6	Rachas_ms	1.511889e+00	1.3470	1.043812e+00	0.123
7	Temperatura_aire_C	1.559380e+01	14.2250	6.835956e+00	2.317

```

8 Presion_atmosferica_mbar 1.016620e+03 1016.3310 3.036572e+00 1009.648
9     Humedad_relativa_porc 6.761639e+01   70.6500 2.316700e+01    15.700
10    Radiacion_global_MJm2 6.816402e-01    0.0366 9.424293e-01    0.000
11 Temperatura_superficie_C 1.596208e+01   13.7500 8.175708e+00    1.126
12 Flujo_energia_suelo_Wm2 -9.835949e-02  -1.0150 4.760450e+00   -10.820
13     Radiacion_neta_Wm2  5.264392e+02  317.9000 7.687920e+02  -1642.000
14 Temperatura_termocupla1_C 1.658387e+01   14.1500 8.853681e+00    0.719
15 Temperatura_termocupla2_C 1.658374e+01   14.1500 8.853524e+00    0.721
16     Humedad_suelo_5cm_m3m3 3.570538e-01   0.3650 3.624735e-02    0.214
17     Humedad_suelo_40cm_m3m3 4.897650e-01   0.4910 5.753877e-03    0.476

      max nas
1 873.0000 0
2 6.0350 4
3 3.0000 4
4 2.1030 4
5 359.7000 4
6 6.3280 4
7 32.9200 4
8 1027.0980 4
9 100.0000 4
10 3.1428 4
11 36.3800 4
12 12.4600 4
13 3004.0000 4
14 38.4300 4
15 38.4300 4
16 0.4500 4
17 0.5010 4

```

```

#| label: valores-faltantes
#| echo: true

valores_faltantes_df <- data.frame(
  variable = names(resultados_ae$valores_faltantes),
  n_na = as.numeric(resultados_ae$valores_faltantes)
)

valores_faltantes_df

```

	variable	n_na
1	fecha_hora	105
2	Precipitacion_mm	4
3	Rayos_conteo	4
4	Dist_km_Avg	4
5	Velocidad_viento_ms	4
6	Direccion_viento_grados	4
7	Rachas_ms	4
8	Temperatura_aire_C	4
9	VP_mbar_Avg	4
10	Presion_atmosferica_mbar	4
11	ETos	4
12	Rso	4
13	Humedad_relativa_porc	4
14	RHT_C	4
15	TiltNS_deg_Avg	4
16	TiltWE_deg_Avg	4
17	Radiacion_global_MJm2	4
18	CVMeta	14

```

19           Invalid_Wind      4
20             TT_C_Avg       4
21 Temperatura_superficie_C      4
22 Flujo_energia_suelo_Wm2      4
23     Radiacion_neta_Wm2      4
24             CNR_Wm2_Avg     4
25 Temperatura_termocupla1_C      4
26 Temperatura_termocupla2_C      4
27   Humedad_suelo_5cm_m3m3      4
28   Humedad_suelo_40cm_m3m3      4

```

Resumen de flags por variable

```

#| label: resumen-flags
#| echo: true

resumen_flags <- datos_qc %>%
summarise(
temp_ok = sum(flag_temp_rango == "ok", na.rm = TRUE),
temp_fuera = sum(flag_temp_rango == "fuera_rango", na.rm = TRUE),
temp_na = sum(flag_temp_rango == "na", na.rm = TRUE),

hr_ok      = sum(flag_hr_rango == "ok",           na.rm = TRUE),
hr_fuera   = sum(flag_hr_rango == "fuera_rango",  na.rm = TRUE),
hr_na      = sum(flag_hr_rango == "na",            na.rm = TRUE),

pp_ok      = sum(flag_pp_rango == "ok",           na.rm = TRUE),
pp_negativa = sum(flag_pp_rango == "negativa",    na.rm = TRUE),
pp_na      = sum(flag_pp_rango == "na",            na.rm = TRUE),

ws_ok      = sum(flag_ws_rango == "ok",           na.rm = TRUE),
ws_negativa = sum(flag_ws_rango == "negativa",    na.rm = TRUE),
ws_na      = sum(flag_ws_rango == "na",            na.rm = TRUE),

rn_ok      = sum(flag_rn_rango == "ok",           na.rm = TRUE),
rn_fuera   = sum(flag_rn_rango == "fuera_rango",  na.rm = TRUE),
rn_na      = sum(flag_rn_rango == "na",            na.rm = TRUE)

)

resumen_flags

# A tibble: 1 x 15
  temp_ok temp_fuera temp_na hr_ok hr_fuera hr_na pp_ok pp_negativa pp_na ws_ok
  <int>     <int>    <int> <int>    <int> <int>    <int>     <int> <int> <int>
1     2434         0        4   2434        0        4   2434        0        4   2434
# i 5 more variables: ws_negativa <int>, ws_na <int>, rn_ok <int>,
#   rn_fuera <int>, rn_na <int>

#| label: resumen-flags-porcentaje
#| echo: true

n_total <- nrow(datos_qc)

resumen_flags_porcentaje <- resumen_flags %>%
mutate(across(everything(), ~ .x / n_total * 100))

resumen_flags_porcentaje

```

```

# A tibble: 1 x 15
  temp_ok temp_fuera temp_na hr_ok hr_fuera hr_na pp_ok pp_negativa pp_na ws_ok
  <dbl>      <dbl>    <dbl> <dbl>    <dbl> <dbl>    <dbl> <dbl>    <dbl> <dbl>
1    99.8        0   0.164   99.8       0  0.164   99.8       0  0.164   99.8
# i 5 more variables: ws_negativa <dbl>, ws_na <dbl>, rn_ok <dbl>,
#   rn_fuera <dbl>, rn_na <dbl>

#| label: resumen-flags-porcentaje
#| echo: true

n_total <- nrow(datos_qc)

resumen_flags_porcentaje <- resumen_flags %>%
  mutate(across(everything(), ~ .x / n_total * 100))

#resumen_flags_porcentaje

```

Resultados del control de calidad

```

resultados_cc <- control_calidad(datos_brutos)

fuera_rango_df <- purrr::map_dfr(
  resultados_cc$fuera_rango,
  ~ as.data.frame(.x),
  .id = "variable"
)

atipicos_iqr_df <- purrr::map_dfr(
  resultados_cc$atipicos_iqr,
  ~ as.data.frame(.x),
  .id = "variable"
)

completitud_df <- data.frame(
  variable = names(resultados_cc$completitud),
  completitud = as.numeric(resultados_cc$completitud)
)

list(
  fuera_rango = fuera_rango_df,
  atipicos_iqr = atipicos_iqr_df,
  completitud = completitud_df,
  saltos_tiempo = resultados_cc$saltos_temporales
)

```

	variable	n_fuera_rango	porcentaje
1	Temperatura_aire_C	0	0
2	Humedad_relativa_porc	0	0
3	Velocidad_viento_ms	0	0
4	Radiacion_global_Wm2	0	0
5	Precipitacion_mm	0	0
6	Humedad_suelo_5cm_m3m3	0	0
7	Humedad_suelo_40cm_m3m3	0	0

	variable	n_atipicos	porcentaje
--	----------	------------	------------

```

1      Radiacion_global_Wm2      0      0.00
2      Precipitacion_mm      179      7.34
3      Rayos_conteo      3      0.12
4      Velocidad_viento_ms      10      0.41
5      Direccion_viento_grados      0      0.00
6      Rachas_ms      4      0.16
7      Temperatura_aire_C      0      0.00
8      Presion_atmosferica_mbar      28      1.15
9      Humedad_relativa_porc      0      0.00
10     Radiacion_global_MJm2      0      0.00
11     Temperatura_superficie_C      0      0.00
12     Flujo_energia_suelo_Wm2      0      0.00
13     Radiacion_neta_Wm2      66      2.71
14     Temperatura_termocupla1_C      0      0.00
15     Temperatura_termocupla2_C      0      0.00
16     Humedad_suelo_5cm_m3m3      190      7.79
17     Humedad_suelo_40cm_m3m3      0      0.00

```

\$completitud

	variable	completitud
1	Radiacion_global_Wm2	100.00
2	Precipitacion_mm	99.84
3	Rayos_conteo	99.84
4	Velocidad_viento_ms	99.84
5	Direccion_viento_grados	99.84
6	Rachas_ms	99.84
7	Temperatura_aire_C	99.84
8	Presion_atmosferica_mbar	99.84
9	Humedad_relativa_porc	99.84
10	Radiacion_global_MJm2	99.84
11	Temperatura_superficie_C	99.84
12	Flujo_energia_suelo_Wm2	99.84
13	Radiacion_neta_Wm2	99.84
14	Temperatura_termocupla1_C	99.84
15	Temperatura_termocupla2_C	99.84
16	Humedad_suelo_5cm_m3m3	99.84
17	Humedad_suelo_40cm_m3m3	99.84

\$saltos_tiempo

[1] 103

6 Visualización gráfica de las variables

Series temporales facetadas de variables principales

```

#| label: graf-series-temporales-facetadas
#| fig-cap: "Series temporales de variables meteorológicas (facetadas por variable)."
#| echo: false

datos_largo <- datos_brutos %>%
  dplyr::select(
    fecha_hora,
    Radiacion_global_Wm2,
    Temperatura_aire_C,
    Humedad_relativa_porc,
    Precipitacion_mm,
    Velocidad_viento_ms,

```

```

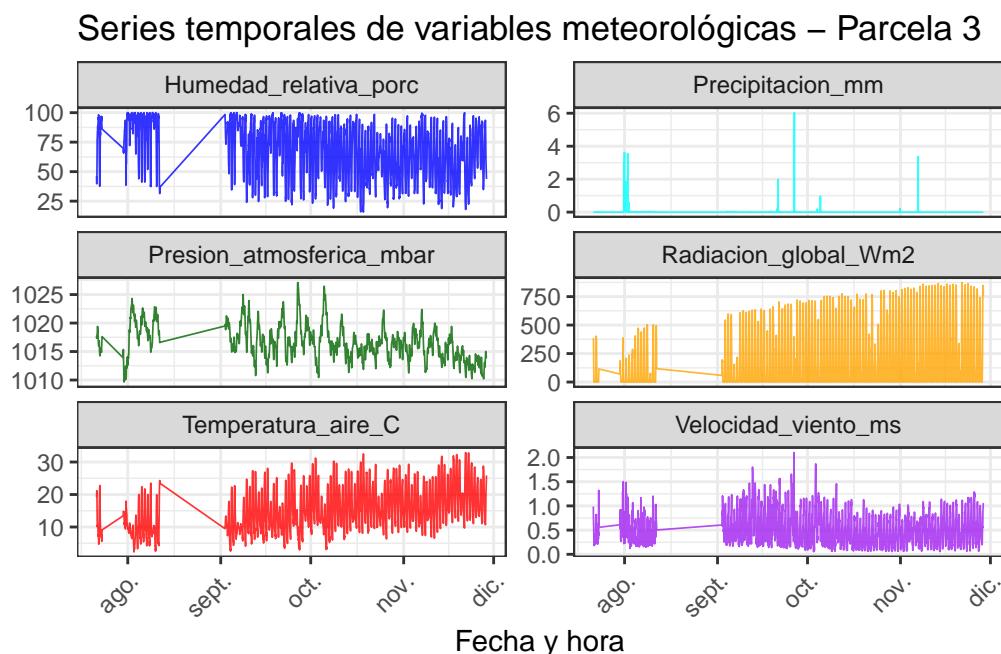
Presion_atmosferica_mbar
) %>%
tidy::pivot_longer(
cols = -fecha_hora,
names_to = "Variable",
values_to = "Valor"
)

series_temporales <- ggplot(datos_largo, aes(x = fecha_hora, y = Valor, color = Variable)) +
geom_line(linewidth = 0.3, alpha = 0.8) +
facet_wrap(~ Variable, scales = "free_y", ncol = 2) +
scale_color_manual(values = c(
"Humedad_relativa_porc" = "blue",
"Presion_atmosferica_mbar" = "darkgreen",
"Temperatura_aire_C" = "red",
"Precipitacion_mm" = "cyan",
"Radiacion_global_Wm2" = "orange",
"Velocidad_viento_ms" = "purple"
)) +
labs(
title = "Series temporales de variables meteorológicas – Parcela 3",
x = "Fecha y hora",
y = ""
) +
theme_bw() +
theme(
axis.text.x = element_text(angle = 45, hjust = 1),
legend.position = "none"
)

series_temporales

```

Warning: Removed 630 rows containing missing values or values outside the scale range (`geom_line()`).



Humedad del suelo en el tiempo (2 profundidades)

```

#| label: graf-humedad-suelo-tiempo
#| fig-cap: "Humedad del suelo a 5 y 40 cm en el tiempo."
#| echo: false

datos_suelo <- datos_brutos %>%
  dplyr::select(
    fecha_hora,
    Humedad_suelo_5cm_m3m3,
    Humedad_suelo_40cm_m3m3
  ) %>%
  tidyr::pivot_longer(
    cols = -fecha_hora,
    names_to = "Profundidad",
    values_to = "Humedad"
  )

Humedad_Suelo <- ggplot(datos_suelo, aes(x = fecha_hora, y = Humedad, color = Profundidad)) +
  geom_line(lineWidth = 0.5) +
  labs(
    title = "Humedad del suelo a diferentes profundidades",
    x = "Fecha y hora",
    y = "Humedad (m³/m³)",
    color = "Profundidad"
  ) +
  theme_bw() +
  theme(legend.position = "top")

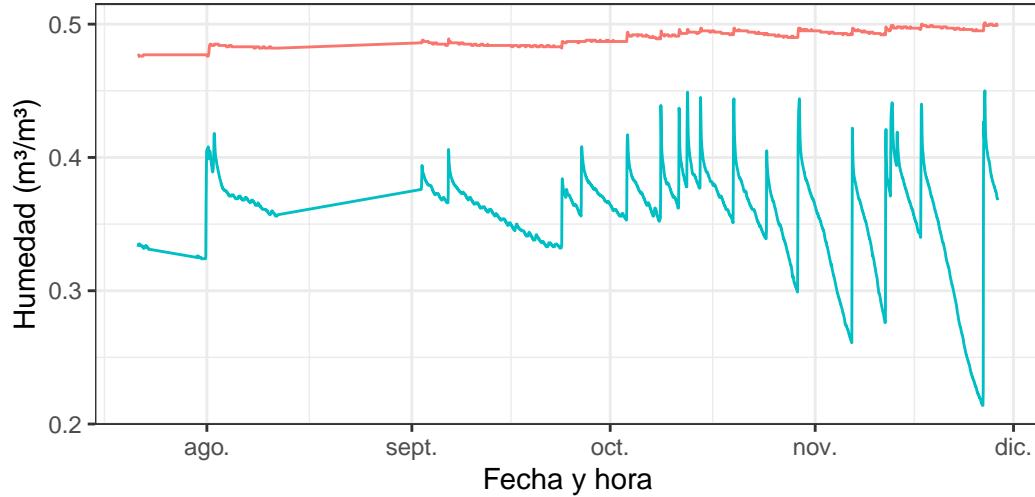
Humedad_Suelo

```

Warning: Removed 210 rows containing missing values or values outside the scale range (`geom_line()`).

Humedad del suelo a diferentes profundidades

Profundidad — Humedad_suelo_40cm_m3m3 — Humedad_suelo_5cm_m3m3



Precipitación diaria acumulada

Warning: Removed 1 row containing missing values or values outside the scale range (`geom_col()`).

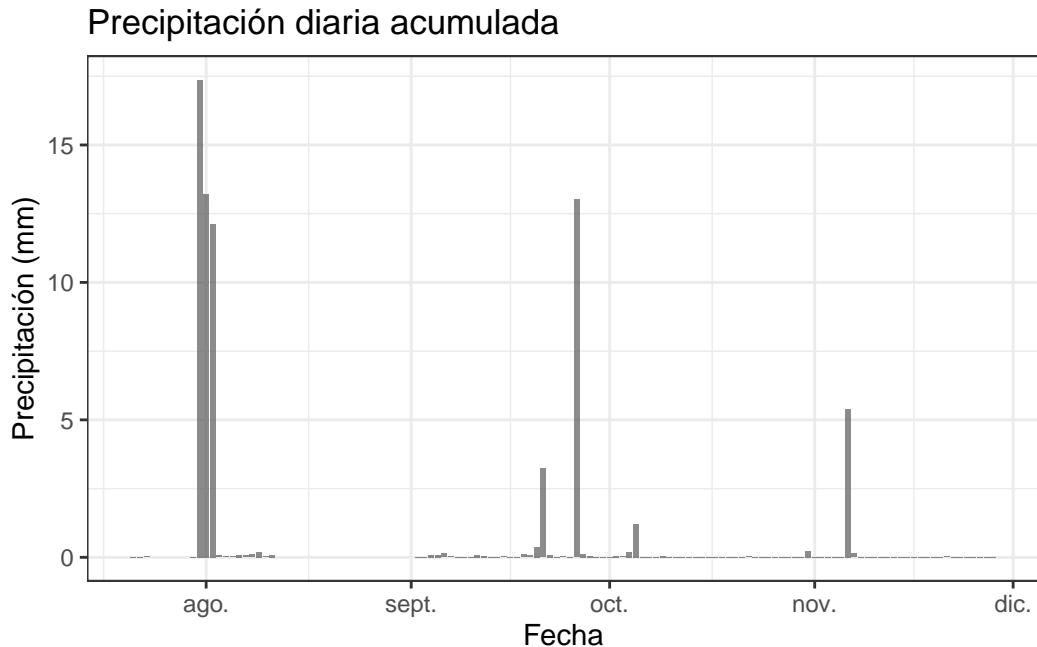


Figure 1: Precipitación diaria acumulada.

Rosa de vientos

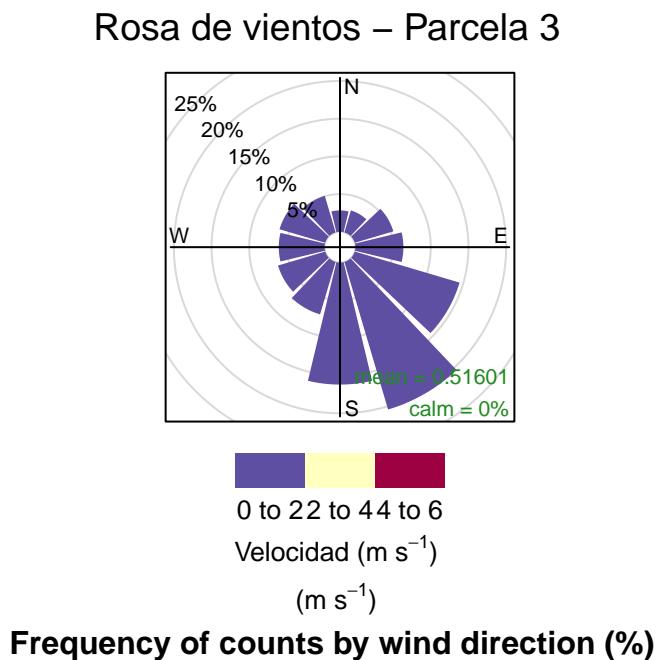


Figure 2: Rosa de vientos (dirección y velocidad del viento).

Distribución de temperaturas

```
Warning: Removed 16 rows containing non-finite outside the scale range
(`stat_boxplot()`).
```

```
Warning: Removed 16 rows containing non-finite outside the scale range
(`stat_summary()`).
```

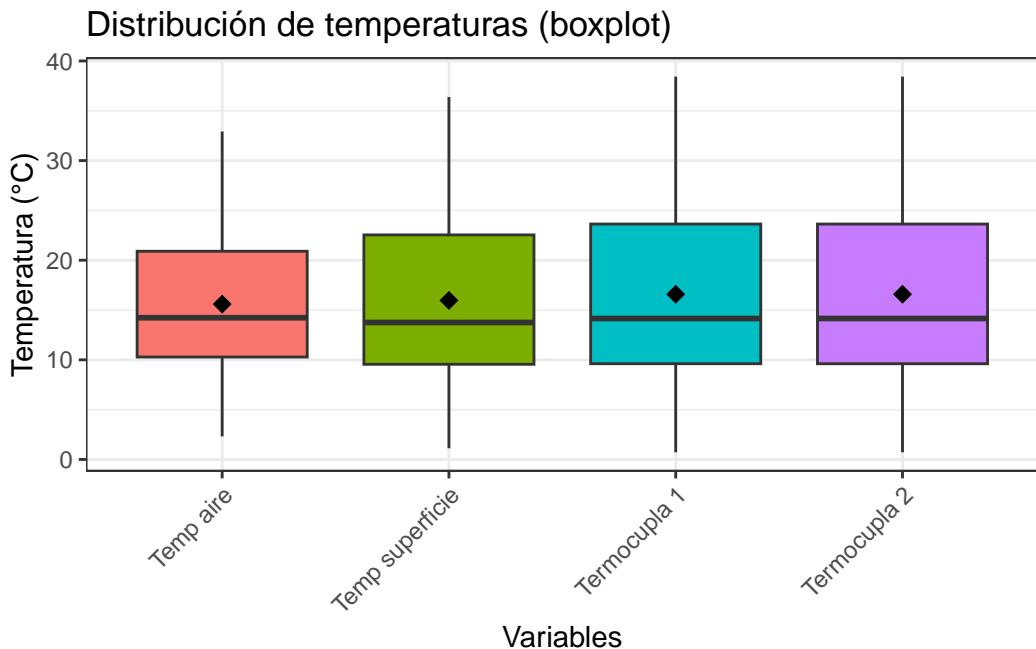


Figure 3: Distribución de temperaturas del aire, superficie y termocuplas.

Generación de base de datos validada

```

datos_validados <- datos_qc %>%
  mutate(
    Temperatura_aire_C = if_else(flag_temp_rango == "ok", Temperatura_aire_C, NA_real_),
    Humedad_relativa_porc = if_else(flag_hr_rango == "ok", Humedad_relativa_porc, NA_real_),
    Precipitacion_mm = if_else(flag_pp_rango == "ok", Precipitacion_mm, NA_real_),
    Velocidad_viento_ms = if_else(flag_ws_rango == "ok", Velocidad_viento_ms, NA_real_),
    Radiacion_neta_Wm2 = if_else(flag_rn_rango == "ok", Radiacion_neta_Wm2, NA_real_)
  )

readr::write_csv(datos_validados, "Datos_Parcela_3_validado.csv")

datos_validados %>%
  summarise(
    n_temp_validadas = sum(!is.na(Temperatura_aire_C)),
    n_hr_validadas = sum(!is.na(Humedad_relativa_porc)),
    n_pp_validadas = sum(!is.na(Precipitacion_mm)),
    n_ws_validadas = sum(!is.na(Velocidad_viento_ms)),
    n_rn_validadas = sum(!is.na(Radiacion_neta_Wm2))
  )

# A tibble: 1 x 5
#>   n_temp_validadas n_hr_validadas n_pp_validadas n_ws_validadas n_rn_validadas
#>   <int>           <int>           <int>           <int>           <int>
#> 1     2434            2434            2434            2434            1821

```