

Graficos

Diego Vega Viquez

Resultados

```
base <- read_excel("data/respaldo_base_limpia.xlsx")

# Extra: volver a poner las categóricas como factores y la fecha como Date por que
# al pasarse a excel pierde esta característica.

base$sitio <- as.factor(base$sitio)
base$cuerpo <- as.factor(base$cuerpo)
base$ubi_muestra <- as.factor(base$ubi_muestra)
base <- base %>% mutate(across(
  c(sitio, ubi_muestra, cuerpo),
  as.factor))
base$latitud <- as.factor(base$latitud)
base$longitud <- as.factor(base$longitud)
base$horarecolectaj <- as_hms(base$horarecolectaj)

base <- base[!is.na(base$dqoj),]

base <- base %>%
  select(-all_of(
    base %>%
      summarise(across(everything(), ~sum(is.na(.)))) %>%
      pivot_longer(everything(), names_to = "columna", values_to = "cantidad_na") %>%
      mutate(
        total = nrow(base),
        porcentaje_na = (cantidad_na / total) * 100
      ) %>%
      filter(porcentaje_na > 25) %>%
      pull(columna)
```

```
))
str(base)
```

```
tibble [33 x 34] (S3: tbl_df/tbl/data.frame)
 $ sitio          : Factor w/ 10 levels "Agujitas","Claro",...: 1 1 1 1 1 1 2 2 2 3 ...
 $ ubi_muestra    : Factor w/ 2 levels "Fondo","Superficie": 1 2 1 2 1 2 1 1 2 2 ...
 $ cuerpo         : Factor w/ 2 levels "Dulce","Salado": 1 1 1 1 1 1 1 1 1 1 ...
 $ fecharecolectaf: POSIXct[1:33], format: "2019-02-15" "2019-02-17" ...
 $ fecharecolectaj: POSIXct[1:33], format: "2019-07-29" "2019-07-29" ...
 $ horarecolectaf : chr [1:33] "11:30:00" "08:22:00" "08:09:00" "11:04:00" ...
 $ horarecolectaj : 'hms' num [1:33] 13:43:00 13:43:00 13:53:00 13:53:00 ...
 ..- attr(*, "units")= chr "secs"
 $ salinidad      : num [1:33] 31 30 30 29 33 30 29 31 31 29 ...
 $ oxigeno        : num [1:33] 5.61 6.41 5.81 6.15 4.76 5.94 93.6 2.87 6.09 6.15 ...
 $ sat_oxigen     : num [1:33] 95.9 110.3 99.6 106.1 82.9 ...
 $ precipitacionf : num [1:33] 0 0 0 0 0.01 0 0.01 0.02 0.02 0 ...
 $ tempairef      : num [1:33] 29.2 28.7 28.6 28.6 29.4 ...
 $ precipitacionj : num [1:33] 0.74 0.74 0.88 0.88 0.72 0.72 0.75 0.78 0.78 0.81 ...
 $ tempairej      : num [1:33] 31.2 31.2 31.1 31.1 30.1 30.1 29.8 33 33 29.9 ...
 $ colifecalj     : num [1:33] 1 1 1 1 1 1 1 1 1 1 ...
 $ ecolij         : num [1:33] 1 1 1 1 1 1 1 1 1 1 ...
 $ enterococoj    : num [1:33] 1 1 1 1 1 1 1 1 1 1 ...
 $ ph             : num [1:33] 8.05 8.06 8.13 8.08 8.14 8.12 7.99 7.87 8.24 8.08 ...
 $ fosfatos       : num [1:33] 0.35 0.32 0.9 0.41 0.33 2.65 0.76 1.12 0.00001 0.41 ...
 $ silicatos      : num [1:33] 8.2 12.93 0.00001 0.00001 0.00001 ...
 $ amonio         : num [1:33] 3.83 3.61 4.02 3.67 4.48 3.88 3.73 4.59 6.31 3.75 ...
 $ nitritos       : num [1:33] 6.72 6.17 7 5.89 8.4 7.7 7.84 8.96 7.84 8.4 ...
 $ nitratos       : num [1:33] 6.07 5.31 6.66 5.4 7.29 7.71 8.95 8.99 8.95 9.31 ...
 $ chla_agua      : num [1:33] 0.75 0.2 0.17 0.25 0.24 0.13 0.76 0.51 0.1 1.25 ...
 $ faopigmentos   : num [1:33] 0.1 0.35 -0.06 -0.17 -0.47 -0.31 0.21 0.16 -0.4 -2.19 ...
 $ matsuspension  : num [1:33] 1.26 1.32 0.95 1.01 1.14 0.88 1.55 1.37 1.62 1.03 ...
 $ alcali_total   : num [1:33] 254 252 250 254 251 ...
 $ dureza         : num [1:33] 8018 8205 8018 7877 8064 ...
 $ dbof           : num [1:33] 7.1 7.1 3.2 3.2 12.6 12.6 3.8 3.2 3.2 0 ...
 $ dboj           : num [1:33] 6 6 1.4 1.4 11.1 11.1 14 182 182 20.4 ...
 $ dqof           : num [1:33] 1443 1443 1457 1457 1507 ...
 $ dqoj           : num [1:33] 1110 1110 1400 1400 1070 1070 1000 1010 1010 1090 ...
 $ latitud        : Factor w/ 27 levels "8.54564","8.56092",...: 14 14 20 20 22 22 10 15 15 20 ...
 $ longitud       : Factor w/ 27 levels "-83.88581","-83.86688",...: 19 19 13 13 11 11 8 4 4 4 ...
```

```
table(base$sitio)
```

Agujitas	Claro	Drake	Isla del Caño	Llorona
6	3	3	0	6
Río Agujitas	Río Claro	Río Drake	Río Llorona	Río Sierpe
3	1	2	3	6

```
table(base$ubi_muestra)
```

Fondo	Superficie
12	21

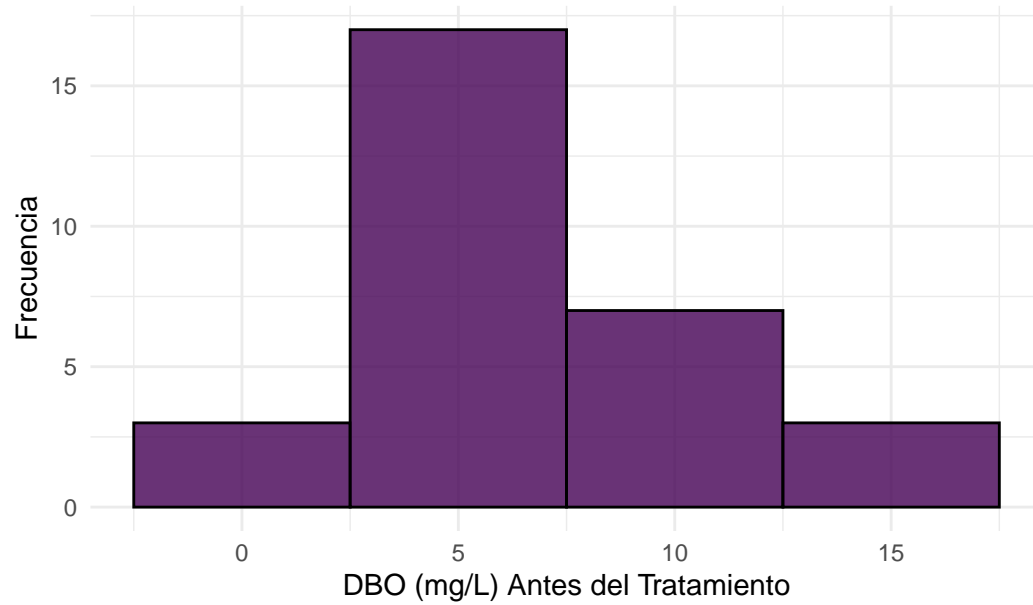
```
table(base$cuerpo)
```

Dulce	Salado
18	15

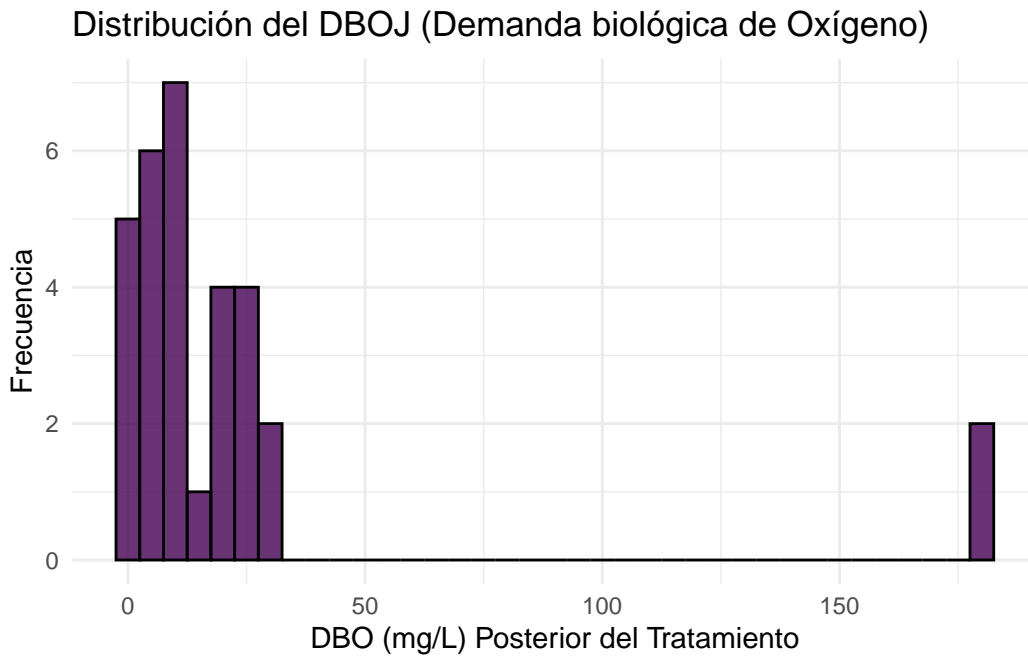
Gráficos

Distribución de la demanda biológica de oxígeno antes del tratamiento

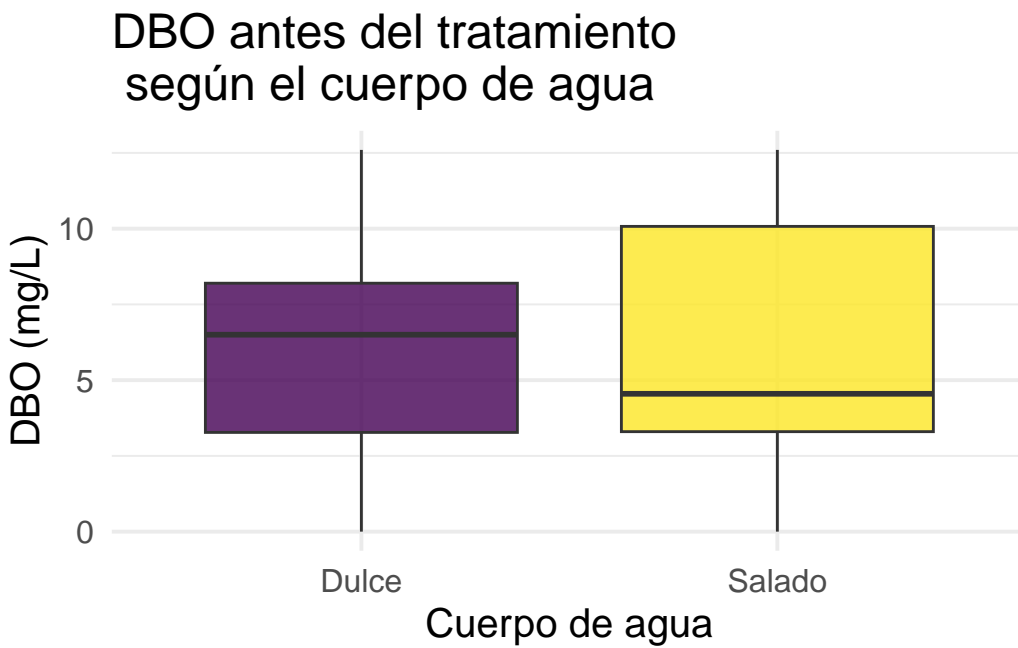
Distribución del DBOF (Demanda biológica de Oxígeno)



Distribución de la demanda biológica de oxígeno posterior al tratamiento

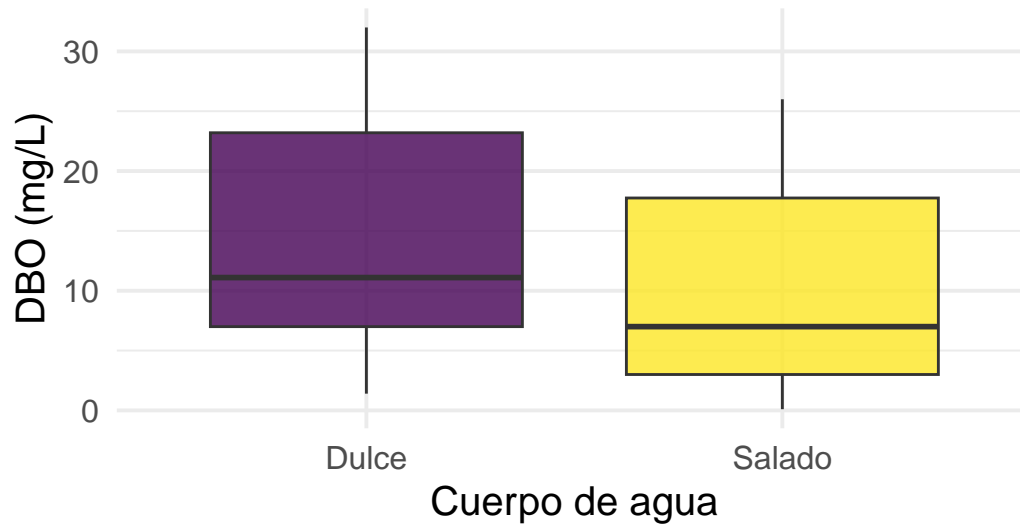


Boxplot de la demanda biológica de oxígeno antes al tratamiento



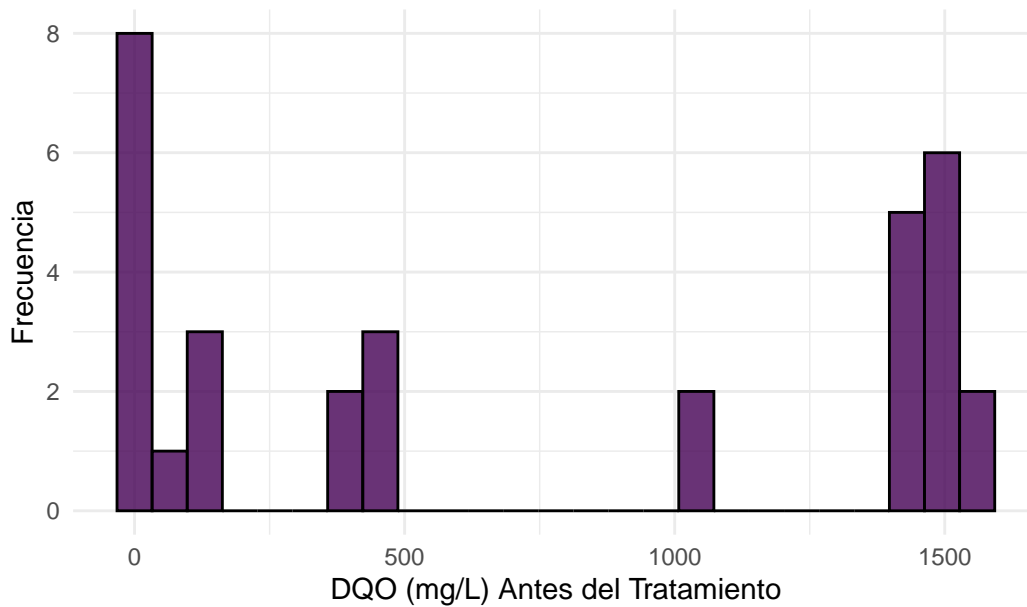
Boxplot de la demanda biológica de oxígeno posterior al tratamiento

DBO posterior del tratamiento
según el cuerpo de agua



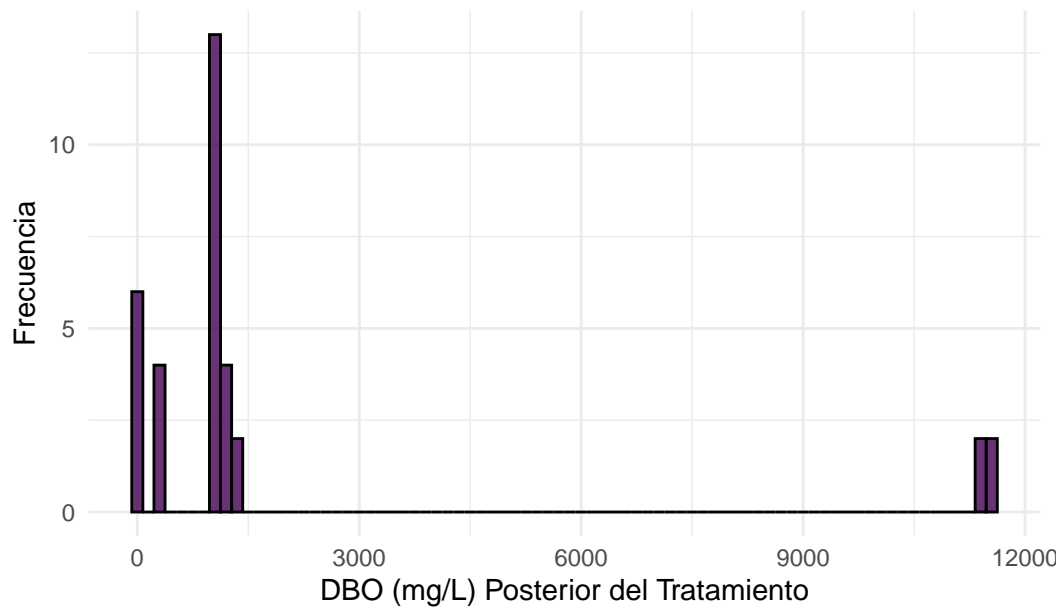
Distribución de la demanda química de oxígeno antes del tratamiento

Distribución del DQOF (Demanda química de Oxígeno)



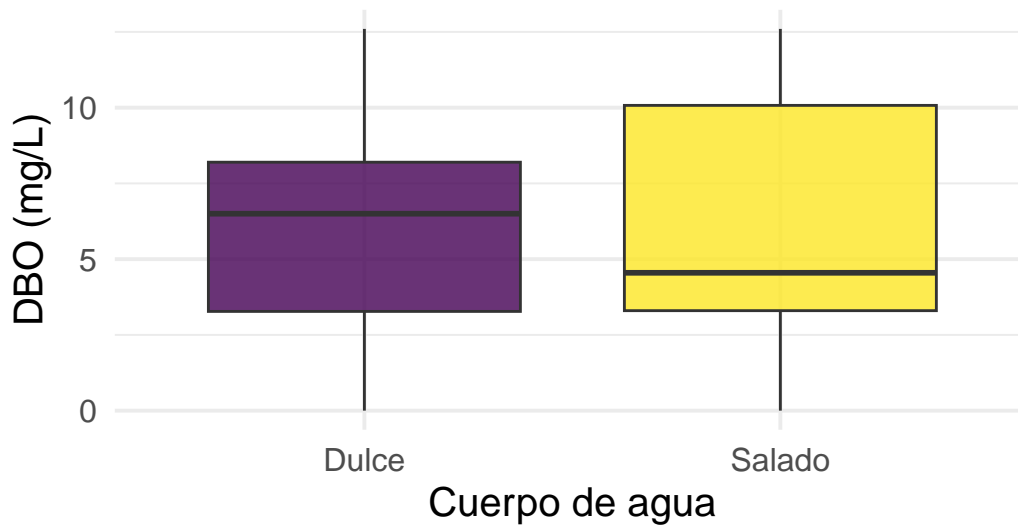
Distribución de la demanda biológica de oxígeno posterior al tratamiento

Distribución del DQOJ (Demanda química de Oxígeno)



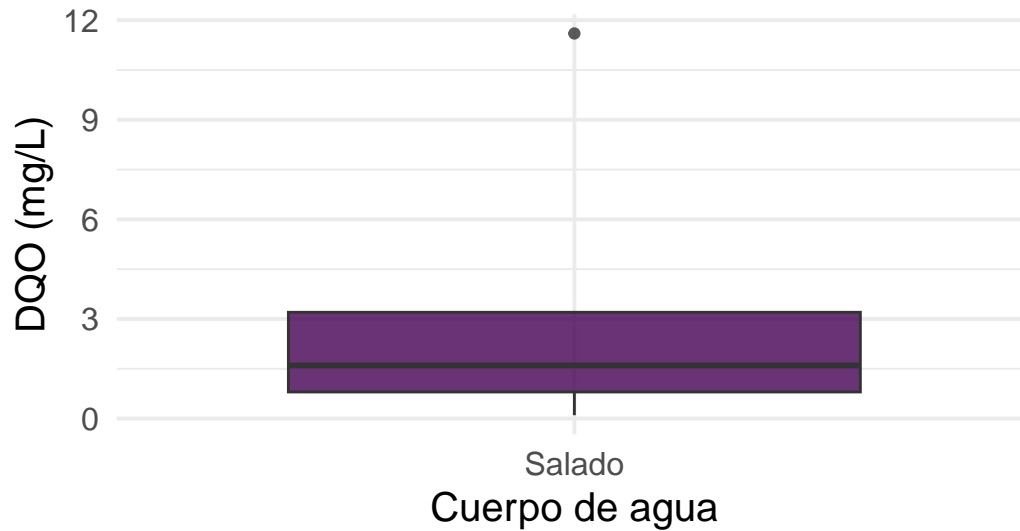
Boxplot de la demanda biológica de oxígeno antes al tratamiento

DBO antes del tratamiento
según el cuerpo de agua



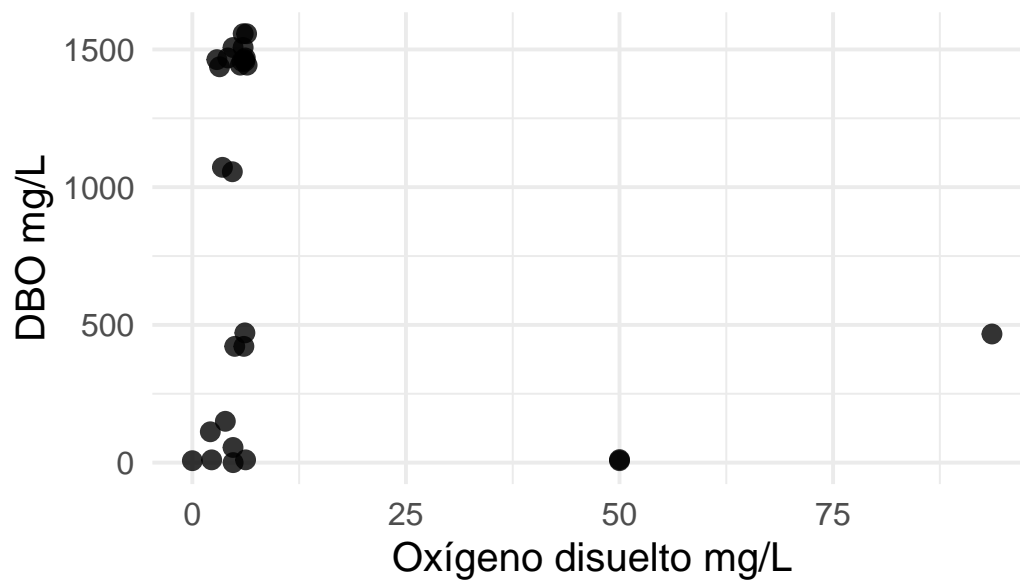
Boxplot de la demanda biológica de oxígeno posterior al tratamiento

DQO posterior del tratamiento
según el cuerpo de agua

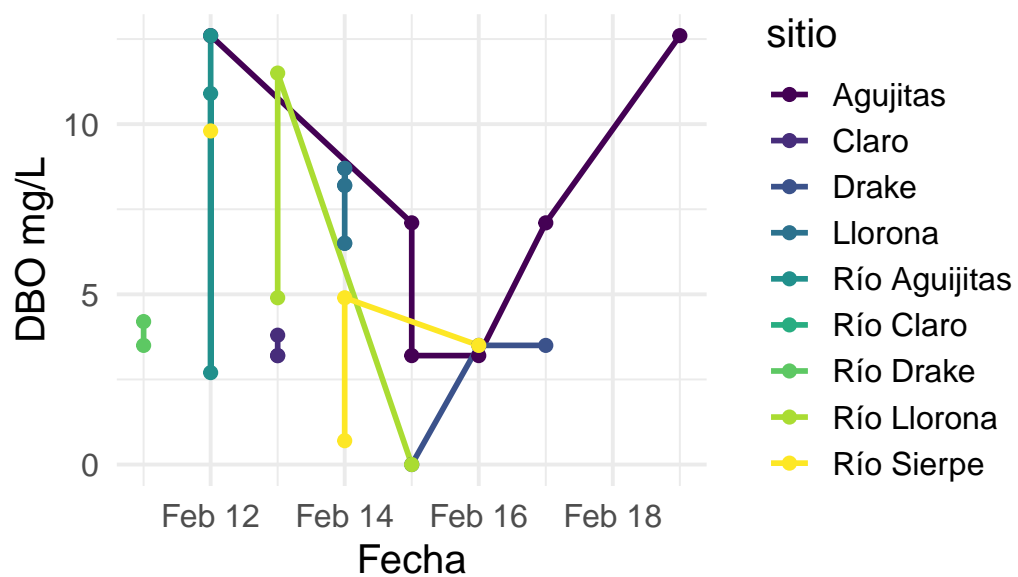


Boxplot de DBO según la calidad del agua

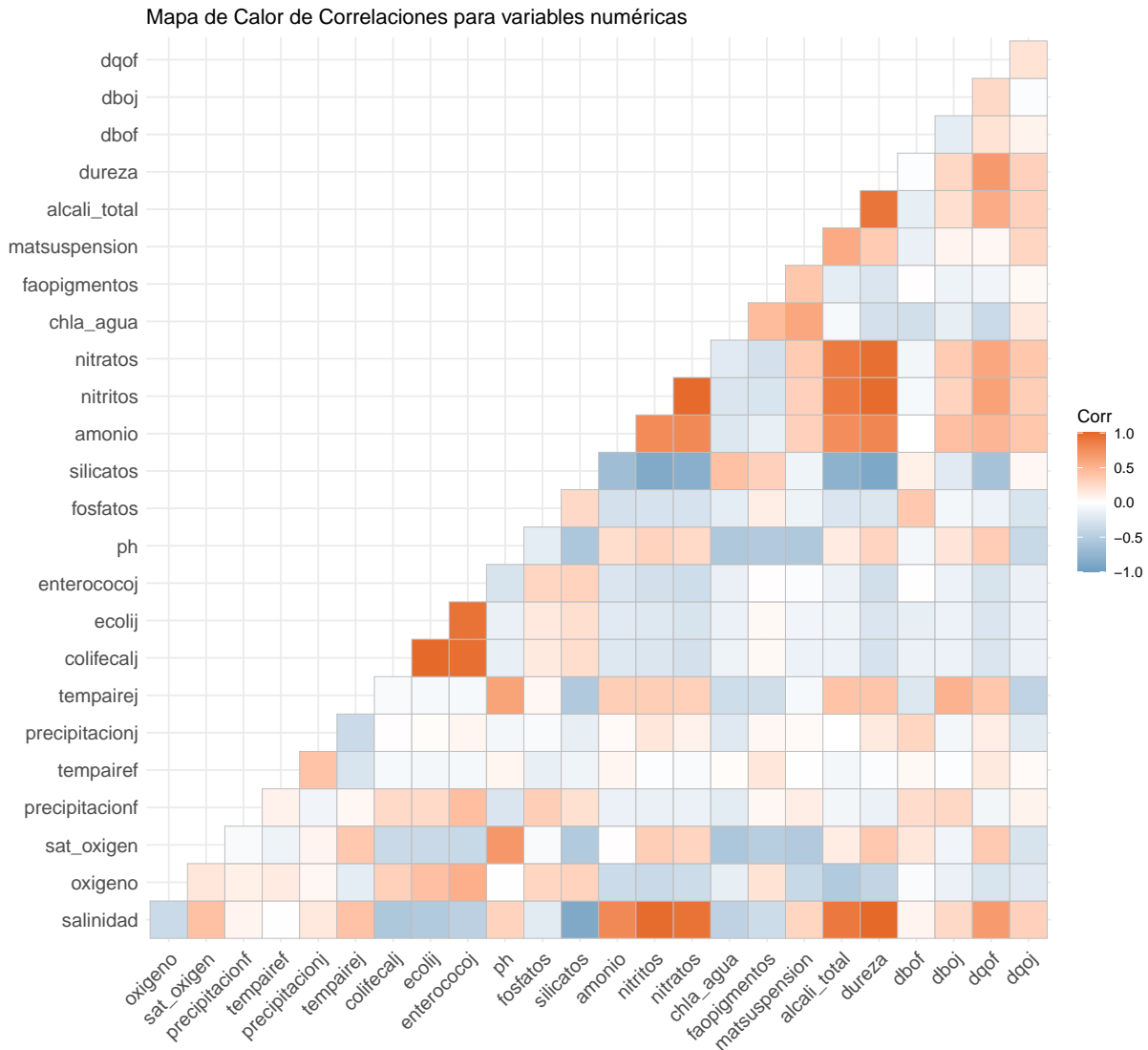
Oxígeno disuelto vs DBO



Evolución del DBO en el tiempo por sitio



Heatmap de Correlación



Este gráfico permite visualizar la fuerza y dirección de las relaciones lineales entre pares de variables mediante la correlación de Pearson, donde los valores cercanos a +1 indican una correlación positiva fuerte, los cercanos a -1 una correlación negativa fuerte, y los cercanos a 0 indican ausencia de relación lineal.

En primer lugar, se observa una correlación positiva alta entre la **salinidad** y varias variables clave, como **amonio**, **nitritos**, **nitratos**, **alcalinidad total** y **dureza**. Este patrón es esperable en cuerpos de agua que presentan influencia marina, intrusión salina o descargas domésticas, ya que la salinidad refleja la concentración de iones disueltos como sodio, cloruro, calcio y magnesio. En estos contextos, también es común que se incremente la presencia de

compuestos nitrogenados, ya sea por procesos de mineralización de materia orgánica o por aportes externos, como aguas residuales o escorrentía agrícola.

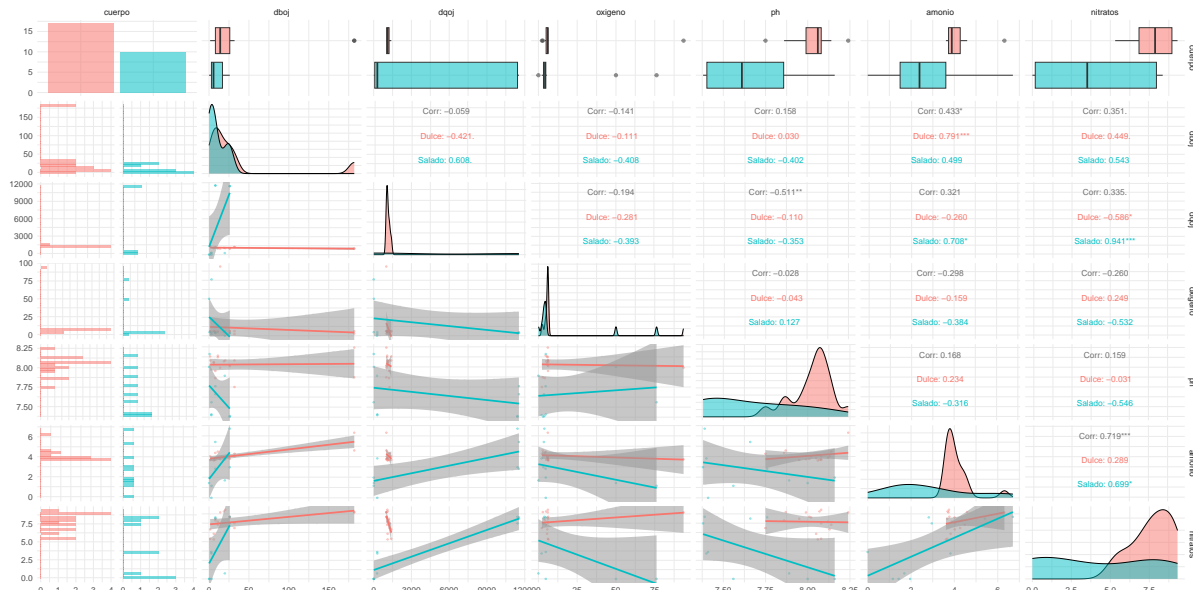
Asimismo, se identifica una correlación positiva fuerte entre la **alcalinidad total** y la **dureza** con las formas inorgánicas del nitrógeno (**amonio**, **nitritos** y **nitratos**). Este hallazgo es consistente con la idea de que dichas variables comparten fuentes comunes, como la disolución de minerales carbonatados en aguas subterráneas, así como la presencia de residuos orgánicos o fertilizantes nitrogenados en cuerpos de agua superficiales. En zonas con alta actividad agrícola o urbana, es frecuente que estos componentes aumenten simultáneamente, reflejando una combinación de procesos naturales y antropogénicos.

Por otro lado, se reportan correlaciones negativas marcadas entre algunas variables, particularmente entre dureza y silicatos, nitritos y silicatos, y salinidad y silicatos. Estos resultados también pueden interpretarse como parte de una dinámica hidrogeoquímica esperable. Los silicatos provienen usualmente de la meteorización de rocas silíceas y tienden a estar más presentes en aguas dulces o de baja mineralización, características de ambientes naturales con baja intervención humana. En contraste, la dureza, la alcalinidad y la salinidad aumentan típicamente en aguas más mineralizadas, influenciadas por rocas calcáreas, intrusión marina o actividades antrópicas. Por tanto, la correlación negativa entre estas variables podría reflejar un gradiente entre zonas de origen geológico silíceo (más naturales) y zonas con mayores cargas minerales y nutrientes (más intervenidas).

Finalmente, la correlación negativa entre **silicatos** y **nitritos** puede interpretarse como una señal indirecta de presión antrópica: en áreas donde predominan las cargas nitrogenadas —producto de descargas residuales o fertilización—, suele observarse una menor concentración de elementos como el silicio, probablemente debido a diferencias en el origen de los aportes o al desplazamiento de fuentes naturales por actividades humanas.

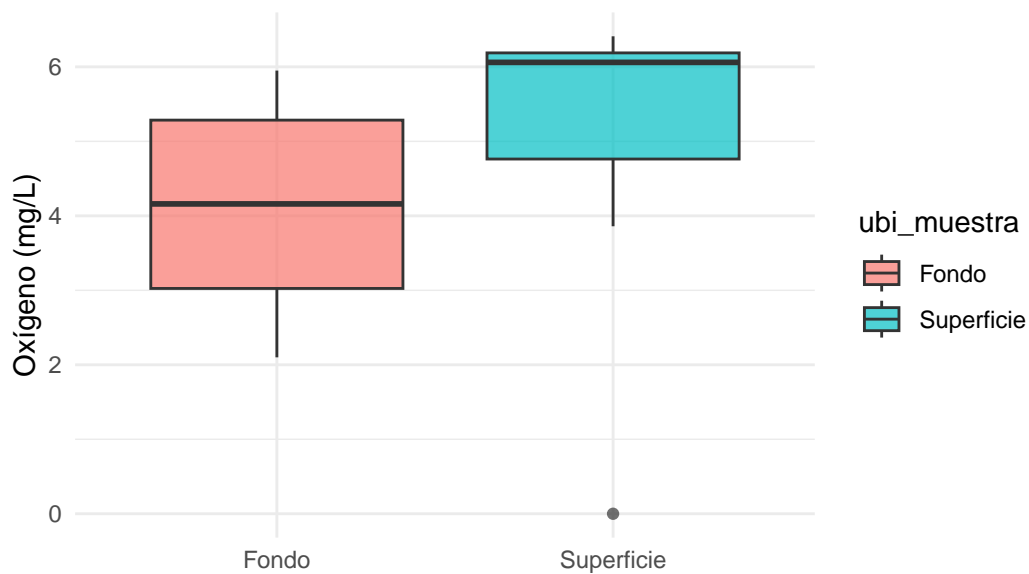
En conjunto, estos patrones de correlación no solo son coherentes con el conocimiento hidrogeoquímico existente, sino que también aportan evidencia sobre la coexistencia de procesos naturales y presiones antrópicas en los cuerpos de agua estudiados.

Matriz de correlaciones visuales y distribuciones para variables fisicoquímicas



Ubicación

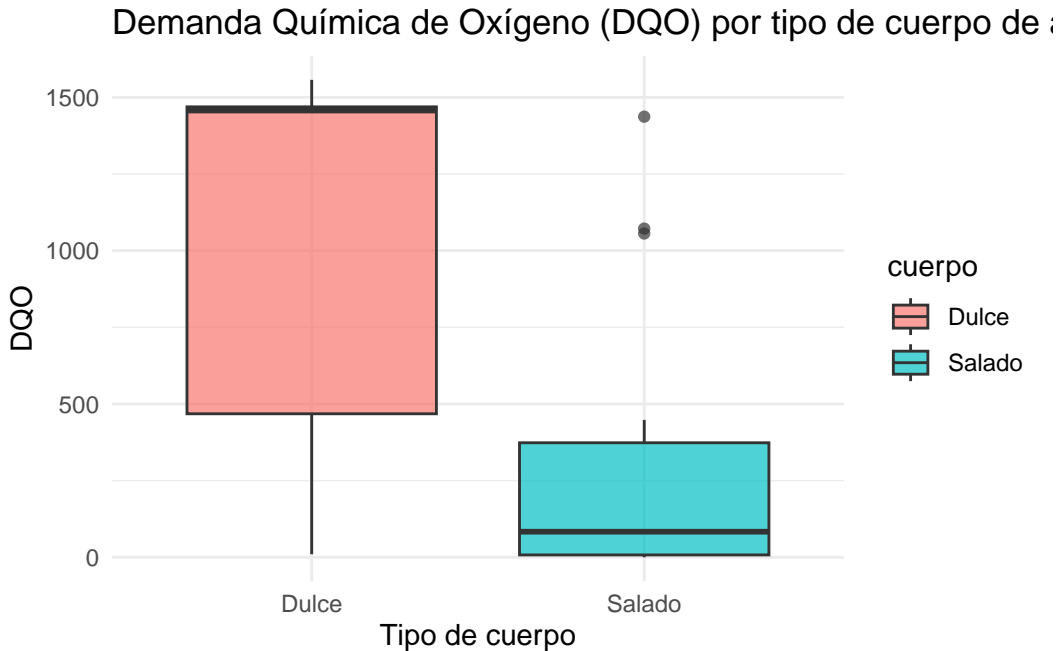
Oxígeno disuelto según ubicación (valores < 20 mg/L)



El análisis de oxígeno disuelto según la ubicación de muestreo revela un patrón claro: las muestras tomadas en el fondo presentan concentraciones significativamente más bajas que aquellas recolectadas en la superficie. Mientras que en la superficie los niveles de oxígeno

oscilan alrededor de los 6 mg/L, en el fondo se observan valores medianos en torno a los 4 mg/L, con algunos registros cercanos a 2 mg/L que indican posibles condiciones de hipoxia. Este comportamiento es coherente con la dinámica de cuerpos de agua estratificados, donde el intercambio gaseoso, la fotosíntesis y la mezcla turbulenta favorecen la oxigenación superficial, mientras que la acumulación de materia orgánica y la falta de renovación vertical provocan un empobrecimiento del oxígeno en las capas profundas.

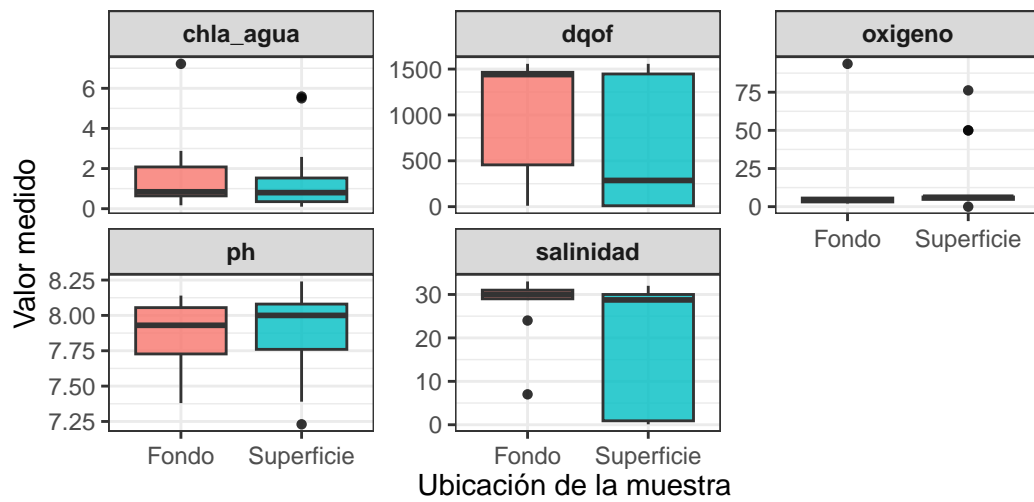
DQO por tipo de Cuerpo



La comparación de la Demanda Química de Oxígeno (DQO) entre cuerpos de agua dulce y salada muestra diferencias marcadas en los niveles de materia orgánica presente. En los cuerpos de agua dulce, la DQO alcanza valores considerablemente más altos, con una mediana cercana a 1200 mg/L, lo cual indica una elevada carga contaminante posiblemente asociada a descargas de origen doméstico, agrícola o industrial. Por el contrario, los cuerpos de agua salada presentan valores de DQO mucho más bajos y estables, lo que podría explicarse por su mayor capacidad de dilución, influencia de corrientes marinas y menor exposición a fuentes contaminantes directas. Estos resultados sugieren un mayor grado de deterioro en los cuerpos de agua dulce, lo cual refuerza la necesidad de monitoreo y control más estricto en estos ecosistemas.

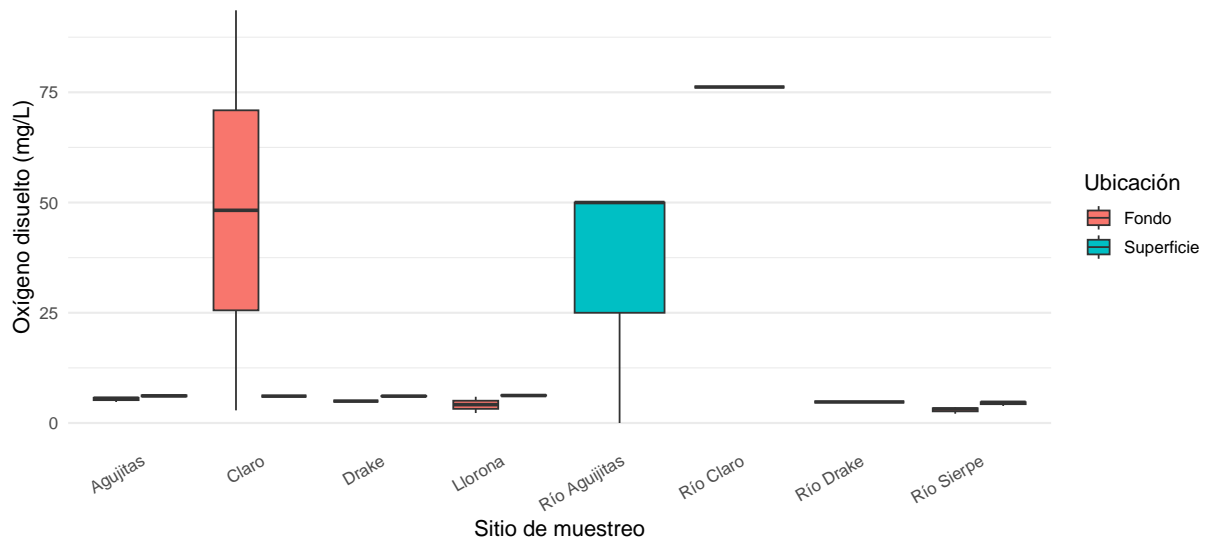
Distribución conjunta (facet) para varias variables clave

Distribuciones de variables seleccionadas según ubicación ver

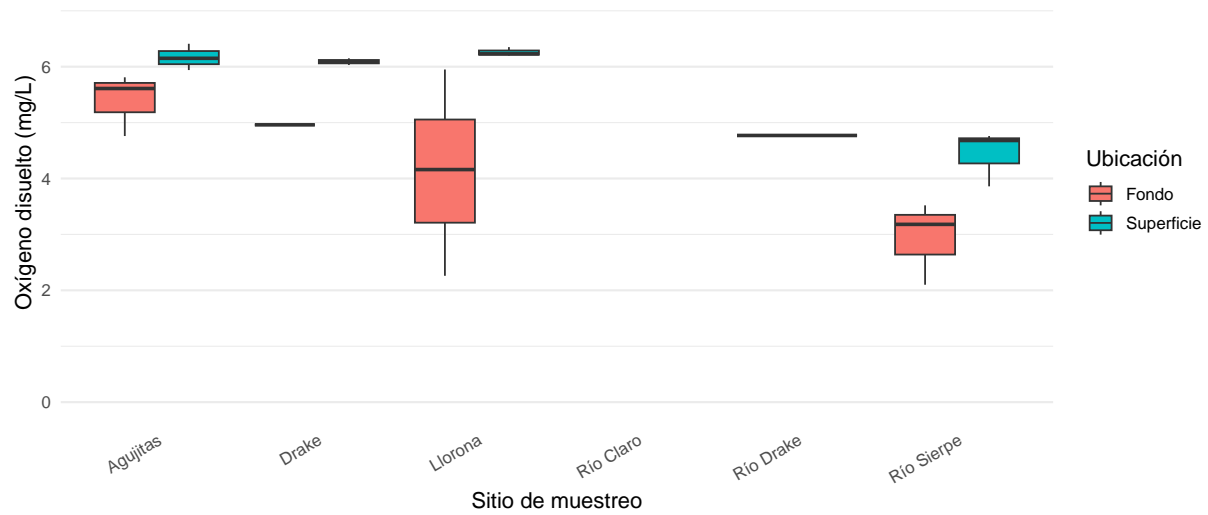


ubi_muestra Fondo Superficie

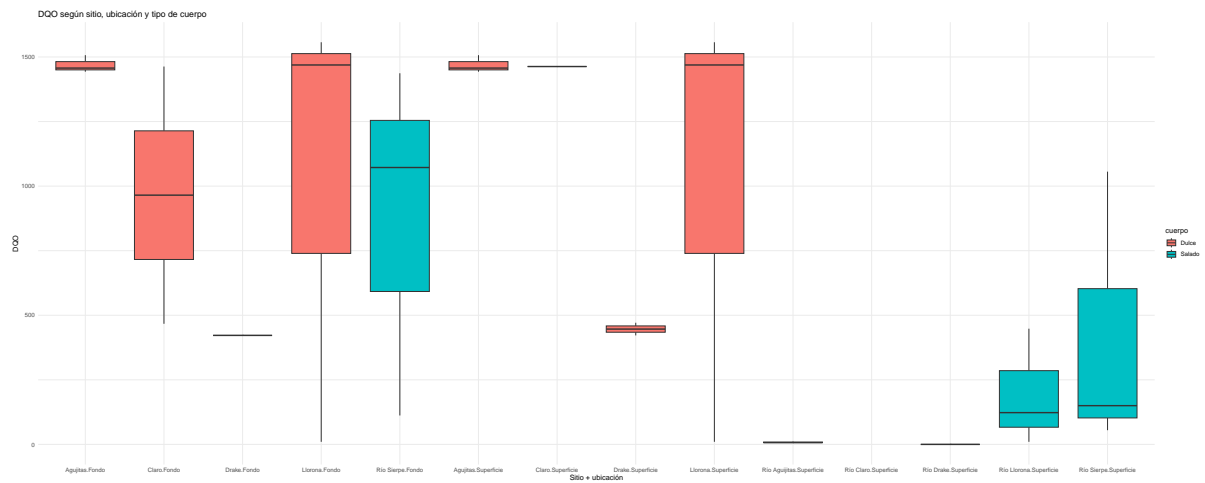
Oxígeno disuelto por sitio y profundidad



Oxígeno disuelto por sitio y profundidad sin Claro ni Río Agujitas



sitio	Dulce	Salado
Agujitas	6	0
Claro	3	0
Drake	3	0
Llorona	6	0
Río Aguijitas	0	3
Río Claro	0	1
Río Drake	0	2
Río Llorona	0	3
Río Sierpe	0	6



```
base %>%
  group_by(sitio, cuerpo, ubi_muestra) %>%
  summarise(
    n = n(),
    dqof_mediana = median(dqof, na.rm = TRUE),
    ph_promedio = mean(ph, na.rm = TRUE),
    salinidad_sd = sd(salinidad, na.rm = TRUE)
  ) %>%
  arrange(desc(dqof_mediana)) %>% kableExtra::kable()
```

`summarise()` has grouped output by 'sitio', 'cuerpo'. You can override using the `.groups` argument.

sitio	cuerpo	ubi_muestra	n	dqof_mediana	ph_promedio	salinidad_sd
Llorona	Dulce	Fondo	3	1469.0	7.876667	0.7637626
Llorona	Dulce	Superficie	3	1469.0	8.013333	1.7559423
Claro	Dulce	Superficie	1	1463.0	8.240000	NA
Agujitas	Dulce	Fondo	3	1457.0	8.106667	1.5275252
Agujitas	Dulce	Superficie	3	1457.0	8.086667	0.5773503
Río Sierpe	Salado	Fondo	3	1072.0	7.480000	11.5325626
Claro	Dulce	Fondo	2	965.0	7.930000	1.4142136
Drake	Dulce	Superficie	2	446.5	8.065000	0.0000000
Drake	Dulce	Fondo	1	422.0	8.070000	NA
Río Sierpe	Salado	Superficie	3	150.0	7.520000	14.0000000
Río Llorona	Salado	Superficie	3	123.0	7.636667	NA
Río Aguijitas	Salado	Superficie	3	7.0	7.966667	0.4163332
Río Drake	Salado	Superficie	2	0.0	8.000000	0.0707107
Río Claro	Salado	Superficie	1	NA	7.560000	NA