Guía de uso del paquete de R *climatol* (versión 4.3-1)

José A. Guijarro (jaguijarro21@gmail.com)

2025-10-28*

Índice

Preámbulo **2**

1. Homogeneización de series **2**
1.1. Introducción . . . 2
1.2. Metodología . . . 3
1.3. Procedimiento de homogeneización . . . 4
1.3.1. Preparación de los ficheros de entrada . . . 4
1.3.2. Homogeneización de las precipitaciones diarias . . . 6
1.3.3. Homogeneización de las temperaturas mensuales . . . 8
1.3.4. Otros parámetros de la función `homogen` . . . 12
1.3.5. Ficheros de resultados . . . 13
1.4. Obtención de productos con los datos homogeneizados . . . 14
1.4.1. Resúmenes estadísticos y series homogeneizadas . . . 14
1.4.2. Series de rejillas homogeneizadas . . . 14
1.5. Dudas frecuentes sobre la función `homogen` . . . 15
1.5.1. Cómo guardar los resultados de diferentes pruebas . . . 15
1.5.2. Cómo cambiar el nivel de corte en el análisis de agrupamiento . . . 15
1.5.3. Cómo usar series de reanálisis como referencias . . . 15
1.5.4. Qué series homogeneizadas debería usar . . . 16
1.5.5. El proceso está durando demasiado tiempo . . . 16
1.5.6. ¿Puede usarse *climatol* para homogeneizar series de caudales? . . . 16
1.6. Bibliografía . . . 16

2. Otras funciones **17**
2.1. Utilidades . . . 17
2.2. Productos gráficos . . . 18
2.2.1. dens2Dplot: Diagrama de dispersión bidimensional . . . 18
2.2.2. diagwl: Diagrama de Walter y Lieth . . . 19
2.2.3. IDFcurves: Diagrama de intensidad-duración-frecuencia a partir de datos subdiarios de precipitación . . . 19
2.2.4. meteogram: Meteograma de 1 día . . . 20
2.2.5. MHisopleths: Isopletas en un diagrama Meses-Horas . . . 20
2.2.6. runtnd: Diagramas de tendencias móviles . . . 21
2.2.7. windrose: Rosa de los vientos a partir de datos de dirección y velocidad del viento . . . 21

Preámbulo

Esta guía es un complemento al manual estándar de R incluido en el propio paquete *climatol*, que es donde se encuentran detallados todos los aspectos de cada función (descripción de los parámetros, información complementaria y ejemplos). Aquí se explicará cómo utilizar esas funciones para el control de calidad, homogeneización y relleno de datos ausentes de un conjunto de series climáticas, y cómo obtener productos derivados de las mismas. También se mostrarán ejemplos de funciones que generan diversos gráficos de utilidad en climatología, pero la información sobre su uso ha de buscarse en el manual estándar.

El historial de cambios y novedades incorporadas a las distintas versiones se encuentra (en inglés) en NEWS.

Las actualizaciones menores (4.3-*) que se vayan produciendo se podrán encontrar, junto con esta guía de usuario y algunos vídeos de ayuda, en https://climatol.eu

Esta guía se estructura en dos capítulos. El primero y principal está dedicado a la homogeneización de series e interpretación de los resultados generados, mientras que en el segundo se muestran ejemplos de otras funciones.

1. Homogeneización de series

1.1. Introducción

Las series de observaciones meteorológicas son de capital importancia para el estudio de la variabilidad climática. Sin embargo, estas series se ven frecuentemente contaminadas por eventos ajenos a dicha variabilidad: errores en la toma de medidas o en su transmisión, cambios en el instrumental utilizado, en la ubicación del observatorio o en su entorno. Estos últimos pueden ser cambios bruscos, como el incendio de un bosque colindante, o graduales, como la posterior recuperación de la vegetación. Estas alteraciones de las series, denominadas inhomogeneidades[1], enmascaran los verdaderos cambios del clima y hacen que el estudio de las series conduzca a conclusiones erróneas.

Para abordar este problema se han desarrollado desde hace muchos años metodologías de homogeneización para eliminar o reducir en lo posible estas alteraciones indeseadas. Inicialmente consistían en comparar una serie problema con otra supuestamente homogénea, pero como esta suposición es muy arriesgada, se pasó a construir series de referencia a partir del promedio de otras seleccionadas por su proximidad o elevada correlación, diluyendo así sus posibles inhomogeneidades. Otros métodos proceden a comparar todas las series disponibles por parejas, de modo que la repetida detección de un cambio en la media permita identificar las series erróneas. Para mayor información pueden consultarse trabajos como los de Peterson *et al.* (1998), Aguilar *et al.* (2003) y Venema *et al.* (2012), así como las directrices sobre homogeneización de la OMM (2020), que pasan revista a estas metodologías.

Existen varios paquetes de programación que implementan estos métodos para que puedan ser usados por la comunidad climatológica. La Acción COST ES0601 *(Advances in homogenisation methods of climate series: an integrated approach, "HOME")* financió un esfuerzo internacional de comparación de los mismos (Venema *et al.*, 2012). Posteriormente el proyecto MULTITEST (Guijarro *et al.*, 2023) realizó otra comparación de los métodos actualizados que pudieran ejecutarse en modo totalmente automático. Hasta ese momento la atención estuvo centrada en la homogeneización de series mensuales, principalmente de temperatura y precipitación, pero el estudio de la variabilidad de los fenómenos extremos suscitó un interés creciente en la homogeneización de series diarias, y en el proyecto INDECIS se homogeneizaron los datos diarios de ocho variables climáticas esenciales de la base de datos ECA&D.

Los resultados de las comparaciones realizadas en el marco del citado proyecto MULTITEST (Guijarro *et al.*, 2023) y en una tesis doctoral sobre homogeneización de temperaturas diarias (Killick, 2016) mostraron que usando *climatol* la calidad de las homogeneizaciones está entre las mejores que se pueden obtener por otros métodos. Además, mientras que algunos otros programas son poco tolerantes a la ausencia de datos,

[1] En realidad el término correcto es *heterogeneidades*, pero *inhomogeneidades* se ha popularizado más en la comunidad climatológica.

climatol se diseñó para poder utilizar series muy cortas o fragmentadas, aprovechando así toda la información climática disponible en la zona de estudio.

1.2. Metodología

Inicialmente *climatol* se programó para rellenar los datos ausentes mediante estimaciones calculadas a partir de las series más próximas. Para ello se adaptó el método de Paulhus y Kohler (1952) para rellenar precipitaciones diarias mediante promedios de valores medidos en las proximidades, normalizados mediante división por sus respectivas precipitaciones medias. Este método se escogió por su simplicidad y por permitir el uso de series vecinas aunque no dispongan de un periodo común de observación con la serie problema, lo que impediría el cálculo de correlaciones y el ajuste de modelos de regresión.

Además de normalizar los datos mediante división por sus valores medios, *climatol* ofrece también hacerlo restando las medias o mediante una estandarización completa. Así, denominando m_X y s_X a la media y desviación típica de una serie X, tenemos estas opciones para su normalización:

1. Restar la media: $x = X - m_X$
2. Dividir por la media: $x = X/m_X$
3. Estandarizar: $x = (X - m_X)/s_X$

El principal problema de esta metodología es que las medias (y desviaciones típicas en el tercer caso) de las series en el periodo de estudio no se conocen si las series no están completas, que es lo más frecuente en las bases de datos reales. Entonces *climatol* calcula primero estos parámetros con los datos disponibles en cada serie, rellena los datos ausentes usando estas medias y desviaciones típicas provisionales, y vuelve a calcularlas con las series rellenadas. Después se vuelven a calcular los datos inicialmente ausentes usando los nuevos parámetros, lo que dará lugar a nuevas medias y desviaciones típicas, repitiendo el proceso hasta que ninguna media cambie al redondearla con la precisión inicial de los datos.

Una vez estabilizadas las medias, se normalizan todos los datos y se procede a estimarlos (tanto si existen como si no, en todas las series) mediante la sencilla expresión:

$$\hat{y} = \frac{\sum_{j=1}^{j=n} w_j x_j}{\sum_{j=1}^{j=n} w_j}$$

en la que $\hat{y}$ es un dato estimado mediante los correspondientes n datos x_j más próximos disponibles en cada paso temporal, y w_j es el peso asignado a cada uno de ellos.

Estadísticamente, $\hat{y_i} = x_i$ es un modelo de regresión lineal denominado *Eje Mayor Reducido* o *Regresión Ortogonal*, en el que la recta se ajusta minimizando las distancias de los puntos medidas en dirección perpendicular a la misma (regresión modelo II) en lugar de en dirección vertical (regresión modelo I) como se hace generalmente (figura 1), cuya formulación (con series normalizadas) es $\hat{y_i} = r \cdot x_i$, siendo r el coeficiente de correlación entre las series x e y. Nótese que este tipo de ajuste se basa en la presunción de que la variable independiente x se mide sin error (Sokal y Rohlf, 1969), presunción que no se sostiene cuando ambas son series climáticas.

Las series estimadas a partir de las demás sirven como referencias para sus correspondientes series observadas, de forma que el siguiente paso es obtener series de anomalías (espaciales) restando los valores estimados a los observados (siempre en forma normalizada). Estas series de anomalías van a permitir:

- Controlar la calidad de las series y eliminar los datos cuyas anomalías espaciales superen un umbral prefijado `dz.max`.
- Comprobar su homogeneidad mediante la aplicación del *Standard Normal Homogeneity Test* (SNHT; Alexandersson, 1986). Alternativamente se puede elegir el test de Cucconi (1968), pero solo si las series de referencia son completas o casi completas.

Cuando los máximos valores obtenidos al aplicar el test a las series son mayores que un umbral `inht` *(INHomogeneity Threshold)* predefinido, la serie se divide por el punto de máximo valor, pasando todos los

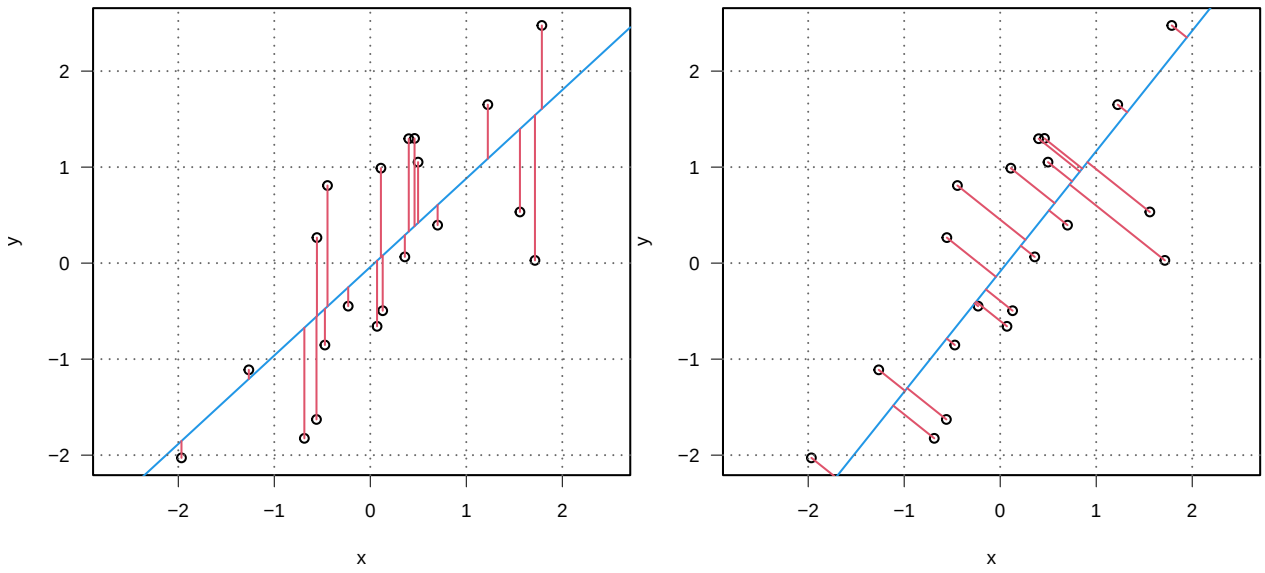


Figura 1: En rojo, desviaciones de la recta de regresión lineal (azul) minimizadas por mínimos cuadrados en los modelos I (izquierda) y II (derecha).

datos anteriores a una nueva serie que se suma a las demás con las mismas coordenadas pero añadiendo un sufijo numérico al código y al nombre de la estación. Este procedimiento se realiza de forma iterativa, cortando solo las series con mayores valores SNHT en cada ciclo, hasta que no se encuentren más inhomogeneidades. Además, como SNHT es una prueba originalmente ideada para encontrar un solo punto de ruptura en una serie, la existencia de dos o más saltos en la media de un tamaño similar podría enmascarar sus resultados. Para minimizar este problema, en una primera pasada se aplica SNHT sobre ventanas temporales solapadas, y después en una segunda pasada se aplica SNHT a las series completas, que es cuando la prueba tiene más poder de detección. Finalmente, una tercera pasada se dedica a rellenar todos los datos ausentes en todas las series y subseries homogéneas con el mismo procedimiento de estimación de datos explicado anteriormente. Por lo tanto, aunque la metodología subyacente del programa es muy simple, su operación se complica a través de una serie de procesos iterativos anidados, como se muestra en el diagrama de flujo de la figura 2.

Aunque se han publicado umbrales de SNHT para distintas longitudes de serie y niveles de significación estadística, la experiencia demuestra que esta prueba puede arrojar valores muy diferentes según la variable climática estudiada, el grado de correlación entre las series y su frecuencia temporal. *Climatol* adopta por defecto el valor umbral `inht=25`, apropiado para valores mensuales de temperatura, aunque un poco conservador, tratando de no detectar falsos saltos en la media a costa de dejar pasar los de menor importancia. Para otras variables puede ser necesario ajustar ese umbral, para lo que pueden utilizarse los histogramas de los valores del test que aparecen en el fichero de gráficos. Si se quisieran homogeneizar series diarias directamente el umbral debería ser alrededor de 10 veces mayor, pero lo aconsejable es realizar la detección de cambios en la media sobre las series mensuales, y luego usar los puntos de corte (opcionalmente ajustados a los metadatos disponibles) para obtener las series diarias homogeneizadas.

1.3. Procedimiento de homogeneización

Tras haber expuesto la metodología seguida por el paquete *climatol*, esta sección se dedicará a ilustrar su aplicación práctica a través de algunos ejemplos.

1.3.1. Preparación de los ficheros de entrada

Climatol solo necesita dos ficheros de entrada, uno con la lista de coordenadas, códigos y nombres de las estaciones, y otro con todos los datos, estación por estación, en el mismo orden en que figuran en el fichero de estaciones. Ninguno de estos ficheros lleva líneas de cabecera ni números de fila, y sus datos se separan

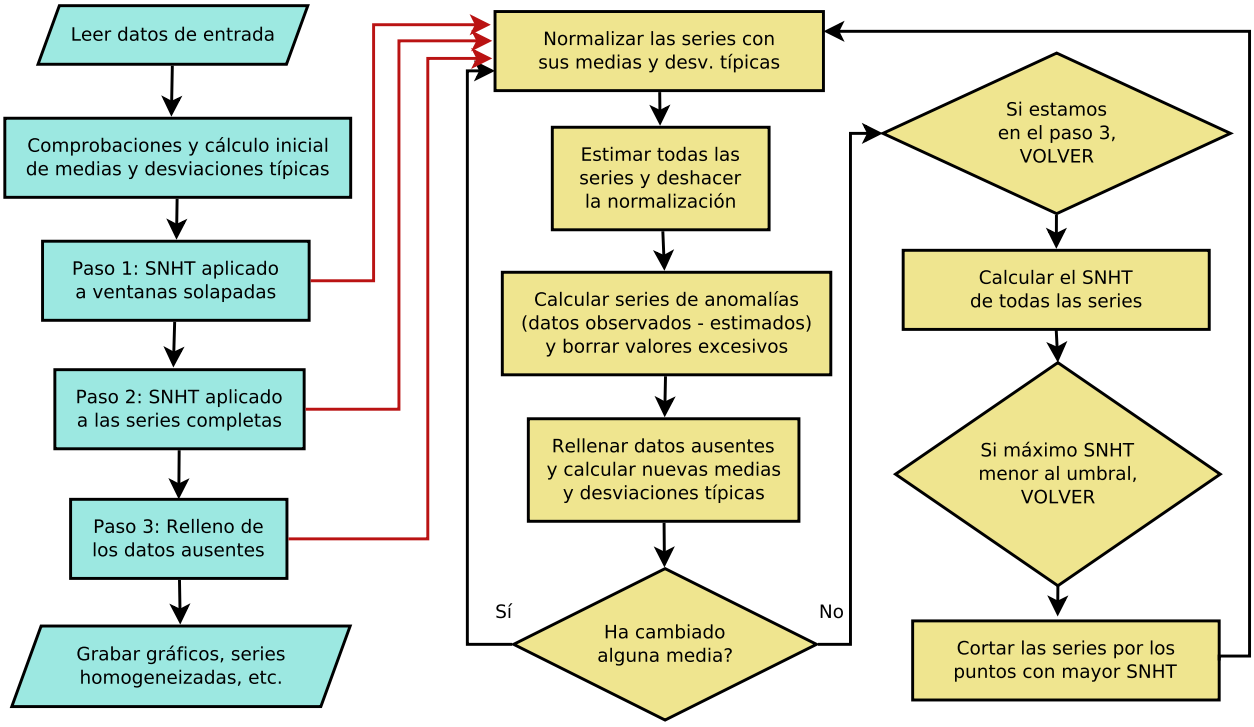


Figura 2: Diagrama de flujo de la función *homogen*, mostrando sus procesos iterativos.

por espacios. (Si los nombres de estación están compuestos por más de una palabra hay que ponerlos entre comillas). En el fichero de datos todas las series han de estar completas, representando los datos ausentes con `NA` u otro código distintivo, pero las líneas pueden tener cualquier número de datos, puesto que se leerán secuencialmente. Además, para que las funciones de postproceso que se comentarán más adelante funcionen sin problemas, el periodo de estudio debería abarcar años completos, comenzando en enero (el día 1 si son datos diarios) del año inicial y terminando en diciembre (el día 31 en el caso de datos diarios) del año final, aunque esto no es estrictamente necesario.

Ambos archivos comparten el mismo nombre básico `VRB_aaaa-AAAA` (donde `VRB` es una abreviatura de la variable estudiada, `aaaa` el primer año y `AAAA` el último abarcado por las series), pero tienen distintas extensiones: `dat` para los datos y `est` para las estaciones. Estas extensiones no son reconocidas directamente por Windows, por lo que los usuarios de este sistema operativo tendrán que indicar que deben abrirse con el bloc de notas u otro editor de texto plano, evitando el uso de procesadores de texto como MS-Word.

Como muestra podemos recurrir a los ejemplos de la documentación estándar de *climatol*. (Un requerimiento del repositorio CRAN es que los ejemplos se ejecuten en unos pocos segundos, por lo que los datos de los ejemplos son mucho más reducidos que los que se usarán normalmente en aplicaciones reales):

```
library(climatol) #cargar las funciones en memoria
data(climatol_data) #cargar los datos de ejemplo
#mostrar un fragmento del fichero de datos:
write(Temp.dat[41:80,4],stdout(),ncolumns=10)
```

```
20.4 23 25.8 26.1 24.8 18.9 14.9 11.7 10.6 8.5
12.6 14.2 19.3 22.4 26.4 26.1 NA NA 15.9 12.5
13.2 NA 12.6 17 NA NA NA 25.5 23.1 18.2
12.5 10.1 10.7 11.2 13.5 NA 18.1 19.5 NA 24.9
```

```
#ver el fichero de estaciones:
write.table(Temp.est,stdout(),row.names=FALSE,col.names=FALSE)
```

```
-2.5059 39.0583 210 "st01" "Station 1"
-2.7028 38.9808 112 "st02" "Station 2"
-2.63 38.8773 111 "st03" "Station 3"
-2.5699 38.9205 112 "st04" "Station 4"
-2.4663 38.9885 125 "st05" "Station 5"
```

Se puede observar que cada línea contiene las coordenadas X (longitud, °), Y (latitud, °), Z (altitud, m), código y nombre de la estación, todo ello separado por espacios. Las coordenadas se expresan en grados con decimales (no en grados, minutos y segundos) y con el signo adecuado para indicar Oeste, Este, Norte o Sur.

Vamos a guardar ficheros de entrada de temperaturas mensuales y precipitaciones diarias para realizar un par de ejemplos de homogeneización (elegir primero un directorio de trabajo en la sesión de R):

```
#temperaturas mensuales de 5 estaciones en 1961-2005:
write.table(Temp.est,'Temp_1961-2005.est',row.names=FALSE,col.names=FALSE)
write(Temp.dat,'Temp_1961-2005.dat')
#precipitaciones diarias de 3 estaciones en 1981-1995:
write.table(SIstations,'Prec_1981-1995.est',row.names=FALSE,col.names=FALSE)
dat <- as.matrix(RR3st[,2:4])
#las series de precipitación están completas, pero vamos a borrar algunos datos:
dat[1:300,1] <- dat[c(1000:1200,2000:2015),2] <- dat[5000:5478,3] <- NA
#y también introducir unos pocos errores:
dat[500:509,1] <- 9.9; dat[600,2] <- -9.9; dat[3000,3] <- 999
#ahora las escribimos en el fichero de datos:
write(dat,'Prec_1981-1995.dat')
```

Para ayudar en la preparación de los ficheros de entrada con este formato, *climatol* provee algunas funciones útiles (ver la documentación estándar de *climatol* para más detalles sobre su uso):

- **db2dat** genera los ficheros extrayendo las series de una base de datos a través de una conexión ODBC.
- **daily2climatol** sirve para cuando cada estación tiene los datos diarios almacenados en archivos individuales.
- **rclimdex2climatol** es útil si los ficheros están en formato RClimDex.
- **sef2climatol** convierte los datos de los ficheros SEF[2].
- **xls2csv** extrae los datos de ficheros **xls** o **xlsx** y los vuelca a un único fichero con formato CSV.
- **csv2climatol** lee los datos de un único fichero CSV y genera los ficheros para *climatol*.

1.3.2. Homogeneización de las precipitaciones diarias

Las precipitaciones diarias procedentes de lecturas manuales en pluviómetros tipo Hellmann tienen la particularidad de que los datos reportados por los observadores suelen dejar en blanco los días sin lluvia, originando una ambigüedad por no especificar si no llovió o si no se pudo hacer la observación. Cuando esto sucede en uno o varios días, la siguiente lectura corresponde a la lluvia acumulada durante los días en que no se realizó la observación. En estos casos el observador debe comunicar la incidencia señalando qué días no pudo realizar su tarea, y entonces a esos días se puede asignar un código especial, como por ejemplo `-1`.

Pero otras veces puede ser que haya una falta sistemática de observación en los fines de semana (o en algunos otros días de la semana) que no se haya comunicado convenientemente. Si se sospecha que esto puede haber sucedido en alguna de nuestras series de datos diarios de precipitación, podemos aplicar la función **weekendaccum** para detectar en qué años la frecuencia de ceros en entre 1 y 3 días consecutivos es anormalmente alta y está seguida por una frecuencia demasiado baja. Si se detecta esta anomalía con un

[2] SEF (*Station Exchange Format*) es el formato usado en los proyectos de rescate de datos del *Copernicus Climate Change Service*.

nivel de significación prefijado (0,01 % por defecto) en uno o más años de la serie, se substituirán los ceros de esos días por el código de acumulación especificado (`-1` por defecto, pero debe coincidir con el código usado normalmente). Ejemplo con los ficheros de precipitación grabados anteriormente:

```
weekendaccum('Prec', 1981, 1995)
```

Esta orden no encuentra ceros sospechosos en esas series diarias de precipitación. Si los hubiera detectado, por defecto no habría aplicado ningún cambio a las series, pero volviendo a ejecutar la misma orden con el parámetro `expl=FALSE` cambiaría los ceros sospechosos por el código de acumulación por defecto `cumc=-1`, renombraría el fichero de datos original a `Prec-wkn_1981-1995.dat` por si se quiere recuperar para repetir el proceso, y las series modificadas se grabarían con el nombre original `Prec_1981-1995.dat`.

Tanto si las series originales ya contenían datos acumulados como si se han detectado acumulaciones con la función `weekendaccum`, el siguiente paso es distribuir las precipitaciones acumuladas entre los días codificados con `cumc`, para lo que usaremos la función `homogen` del siguiente modo:

```
homogen('Prec', 1981, 1995, cumc=-1) #poner el código de acumulación adecuado
```

Además de desagregar las precipitaciones diarias acumuladas, esta aplicación de la función `homogen` también realiza un control de calidad inicial a las series, que en este ejemplo ha detectado y borrado un dato inferior a cero (distinto de los codificados con `cumc`) y un dato excesivamente alto, así como demasiados datos idénticos consecutivos en la serie `p064`. Si no hubieramos aplicado `homogen` para desagregar los datos acumulados, habría convenido realizar este control de calidad inicial mediante la orden `homogen('Prec', 1981, 1995, onlyQC=TRUE)`, que además de informar sobre el borrado de datos demasiado anómalos también generaría un fichero `Prec-QC_1981-1995.pdf`, cuyos tres primeros gráficos mostrarían, para cada serie (figura 3):

1. Diagramas de caja de los datos, señalando el dato anómalo borrado con un punto rojo.
2. Diagramas de caja de las segundas diferencias de la serie, para detectar datos anómalos aislados.
3. Longitudes de las secuencias de datos idénticos, marcando en rojo una secuencia borrada por contener 10 datos iguales.

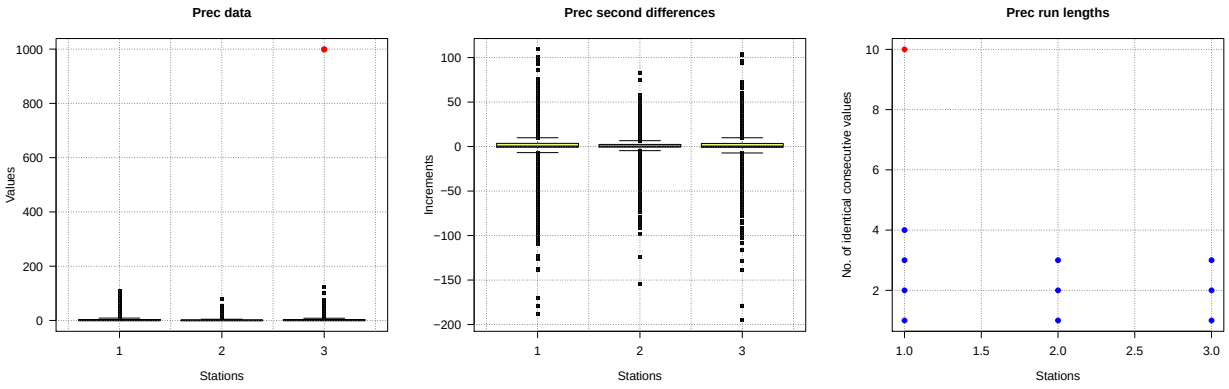


Figura 3: Diagramas de caja y longitudes de secuencias constantes para el control de calidad inicial.

Como la precipitación diaria tiene una distribución de frecuencia fuertemente sesgada, se excluye el borrado de elevados datos aislados porque pueden producirse lluvias fuertes entre dos días con poca o nula precipitación. Con esta variable también se excluyen automáticamente los ceros para no incluir en el análisis de secuencias de datos idénticos las de días con precipitación nula. De este modo, se ha detectado una secuencia de 10 días con el mismo dato en la serie 1, que se ha eliminado.

Los umbrales para este control de calidad inicial se pueden modificar mediante el parámetro `niqd`, que por defecto está fijado en `c(4,6,1)` rangos intercuartílicos para los tres tipos de control respectivamente.

Al haber encontrado claros errores en los datos de entrada, los ficheros originales y los resultados del control de calidad se guardan añadiendo el sufijo `-QC` al nombre de la variable, y las series libres de dichos errores se graban con los nombres originales de los ficheros de entrada.

Ahora ya podemos abordar la homogeneización de las series. La función que realiza esta tarea sigue siendo `homogen`, que inicialmente solo requiere especificar los tres primeros parámetros usados anteriormente: la abreviatura de la variable y los años inicial y final del periodo que abarcan los datos. Pero como la elevada variabilidad de los datos diarios (o subdiarios) dificulta mucho la detección de cambios en la media, la homogeneización directa de dichos datos no está recomendada, y se generaría un error aconsejando homogeneizar primero sus datos mensuales, lo que podemos hacer del siguiente modo:

```
# valm=1 es para calcular totales mensuales en lugar de valores medios:
dd2m('Prec', 1981, 1995, valm=1) #genera las series mensuales 'Prec-m'
# especificamos 'std=2' (normalización recomendada para precipitaciones)
# porque no se asignará automáticamente para series mensuales.
# annual='total' hace que en los últimos gráficos se muestren valores anuales
# totales en lugar de medios:
homogen('Prec-m', 1981, 1995, std=2, annual='total')
```

Tanto el fichero de gráficos `Prec-m_1981-1995.pdf` como el que contiene la lista de cortes efectuados `Prec-m_1981-1995_brk.csv` (vacío en este caso) indican que no se ha cortado ninguna serie, pudiéndose dar por homogéneas. Si se hubieran detectado saltos en la media convendría editar las fechas de corte para ajustarlas a eventos del historial de las estaciones (metadatos), si se dispone de ellos, que justifiquen los cambios detectados. Para finalizar, obtendremos las series diarias homogeneizadas mediante:

```
homogen('Prec', 1981, 1995, annual='total', metad=TRUE)
```

Nuevamente conviene revisar los resultados para comprobar si ha habido algún problema. Especial cuidado hay que tener con la lista de datos anómalos que aparece en el fichero `Prec_1981-1995_out.csv`. Vemos que hay muchos datos sospechosos que no se han borrado (`Deleted=0`), y los mensajes de la consola (guardados en `Prec_1981-1995.txt`) indican que algunos datos anómalos se hubiesen borrado si hubieran tenido más de una referencia disponible. El histograma de anomalías estandarizadas del fichero de gráficos `Prec_1981-1995.pdf` (figura 4) muestra que dos de los tres datos eliminados son bastante anómalos respecto de la distribución de frecuencias de todos ellos. Sin embargo la lista del fichero `Prec_1981-1995_out.csv` nos indica que ambos datos corresponden al mismo día (1993-10-06) en dos series diferentes con anomalías de signo opuesto debido a su discrepancia. Estas anomalías recíprocas se han corregido automáticamente marcando `Deleted=-1` en el fichero `Prec_1981-1995_out.csv` y aplicando la función `datrestore`, que restituye los datos originales en las series homogeneizadas almacenadas en el fichero binario de resultados `Prec_1981-1995.rda`. Pero si alguno de estos datos fuera realmente erróneo, habría que eliminarlo del fichero de entrada `Prec_1981-1995.dat` y repetir el ajuste de las series.

De todos modos hay que tener en cuenta que un dato muy anómalo, aunque sea correcto por un fenómeno meteorológico local, es mejor eliminarlo antes de la homogeneización y restituirlo después en el fichero de series homogeneizadas, puesto que de lo contrario esa anomalía local puede provocar cambios indeseados en las series vecinas. (La función `datrestore` automatiza la restitución de los valores anómalos a los que el usuario haya cambiado manualmente el signo a negativo en la columna `Deleted` del fichero `*_out.csv`).

1.3.3. Homogeneización de las temperaturas mensuales

Veamos ahora otro ejemplo de homogeneización con las temperaturas mensuales que habíamos grabado:

```
homogen('Temp', 1961, 2005)
```

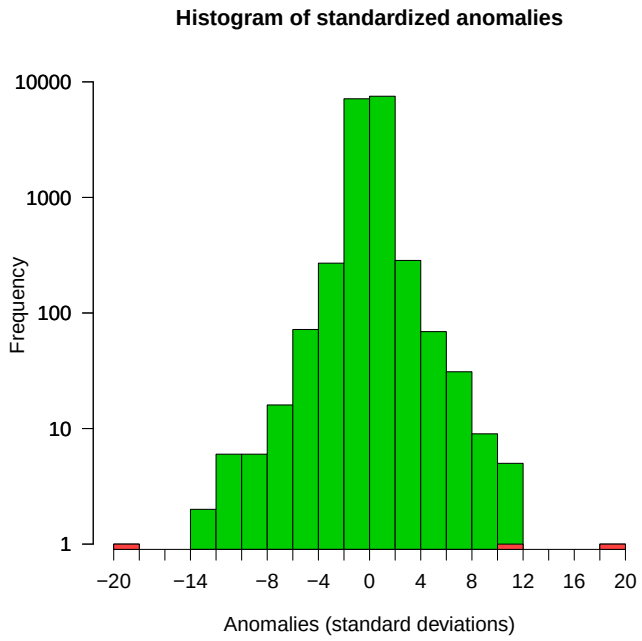


Figura 4: Histograma de anomalías espaciales mostrando en rojo los datos desechados.

El control de calidad inicial ha eliminado un dato muy anómalo y una serie de 9 datos consecutivos idénticos (puntos rojos en la figura 5). El fichero `Temp_1961-2005_out.csv` nos informa de que esta serie de estaba compuesta por 9 datos seguidos con valor 11 que comenzaba el día 1988-02-01.

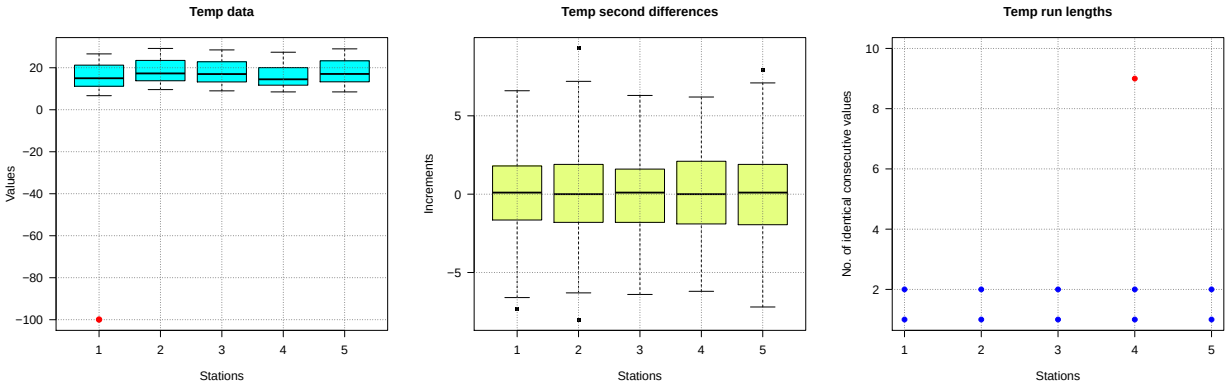


Figura 5: Diagramas de caja y longitudes de secuencias constantes del control de calidad inicial.

Después de los tres primeros gráficos del control de calidad inicial podemos ver la disponibilidad de datos, tanto estación por estación como en total (figura 6). Vemos que la serie 3 es la única que está completa, mientras que la 1 es bastante corta y la 4 está muy fragmentada, con pocos datos pero repartidos a lo largo del periodo de estudio. (Otros métodos de homogeneización no podrían funcionar con tantos datos ausentes). La figura 6-derecha muestra cuántos datos existen en cada paso temporal. Las líneas horizontales a trazos de color verde y rojo indican respectivamente cuál es el mínimo deseable (5 datos) y el mínimo necesario (3 datos) para poder detectar datos sospechosos, pues si en un paso de tiempo solo hay dos datos y resultan discrepantes entre sí, no podremos aventurar cuál será el erróneo.

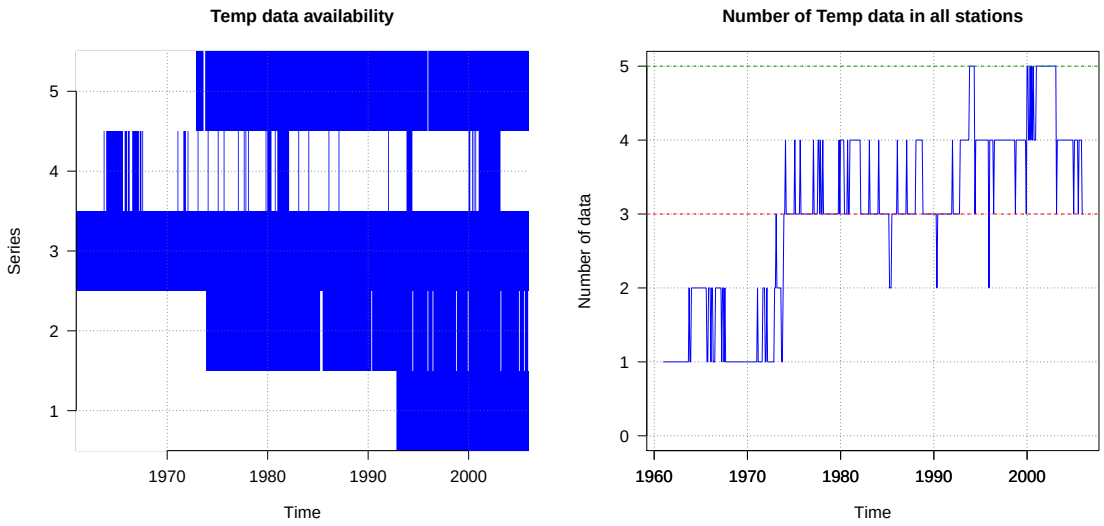


Figura 6: Disponibilidad de datos a lo largo del periodo estudiado.

El mínimo absoluto para que climatol funcione es que haya al menos un dato, pues el objetivo final es rellenar todas la ausencias mediante interpolación espacial, y para ello precisa de la existencia de al menos un dato de referencia. Si en uno o más pasos temporales no existe ningún dato en las series el proceso se detendrá con un mensaje de error, y si esto no es debido a que se hayan borrado demasiados datos anómalos (lo que se puede corregir disminuyendo el umbral `dz.max`) habrá que añadir nuevas series que contengan datos en los periodos críticos o acortar el periodo de estudio para evitar el problema.

Los gráficos que siguen se centran en las correlaciones entre las series y su clasificación en grupos con variabilidad similar, que luego se representan en un mapa. Las correlaciones tienden a disminuir al aumentar la distancia entre las estaciones. Cuanto más altas sean las correlaciones, mayor será la fiabilidad de la homogeneización y el relleno de datos ausentes. En particular, las correlaciones deben ser siempre positivas, al menos dentro de un rango de distancias razonable. De lo contrario, probablemente haya discontinuidades geográficas que produzcan diferencias climáticas. (Por ejemplo, una cresta montañosa puede producir regímenes de precipitación opuestos a ambos lados de la misma). Esto puede confirmarse con el mapa de estaciones, en el que los grupos de variabilidad similar se ubicarían en distintas zonas, en cuyo caso habría que homogeneizar sus series por separado[3].

En áreas de topografía compleja y/o baja densidad de estaciones, las correlaciones pueden estar lejos de ser óptimas. En esta situación, los datos rellenados se verán afectados individualmente por errores importantes, pero es de esperar que sus parámetros estadísticos sean aceptables.

Para evitar el procesamiento de matrices de correlación demasiado grandes, el número de series utilizado para este análisis de conglomerados está limitado por defecto a 300, y se utiliza una muestra aleatoria de este tamaño cuando el número de series es mayor, pero el usuario puede modificar este número mediante el parámetro `nclust`.

Después de estos gráficos iniciales dedicados a verificar los datos, las siguientes páginas del documento muestran gráficos de anomalías (espaciales) estandarizadas para cada una de las tres etapas:

[3]La función `datsubset` permite obtener ficheros *climatol* de un subconjunto de estaciones.

1. Detección en ventanas escalonadas superpuestas
2. Detección en las series completas
3. Anomalías finales de las series homogeneizadas

Las gráficas de las dos primeras etapas muestran las series de anomalías en las que se han detectado cambios en la media, marcando los puntos de ruptura por donde son cortadas mediante una línea vertical de trazos rojos y rotulando el valor del test de homogeneidad en su extremo superior. Las anomalías finales (en la tercera etapa) sirven para comprobar si han quedado evidentes cambios en la media sin corregir, en cuyo caso habría que volver a homogeneizarlas fijando un umbral de inhomogeneidad `inht` más bajo que el valor 25 usado por defecto. Si, por el contrario, se considerara que los últimos cortes en las series no están justificados en las series de anomalías, lo que haríamos es aumentar dicho umbral.

En nuestro ejemplo se han detectado 5 saltos en la media en cuatro series. En una de ellas ha detectado dos saltos que delimitaban un corto periodo de datos anómalos, que ha sido eliminado (figura 7).

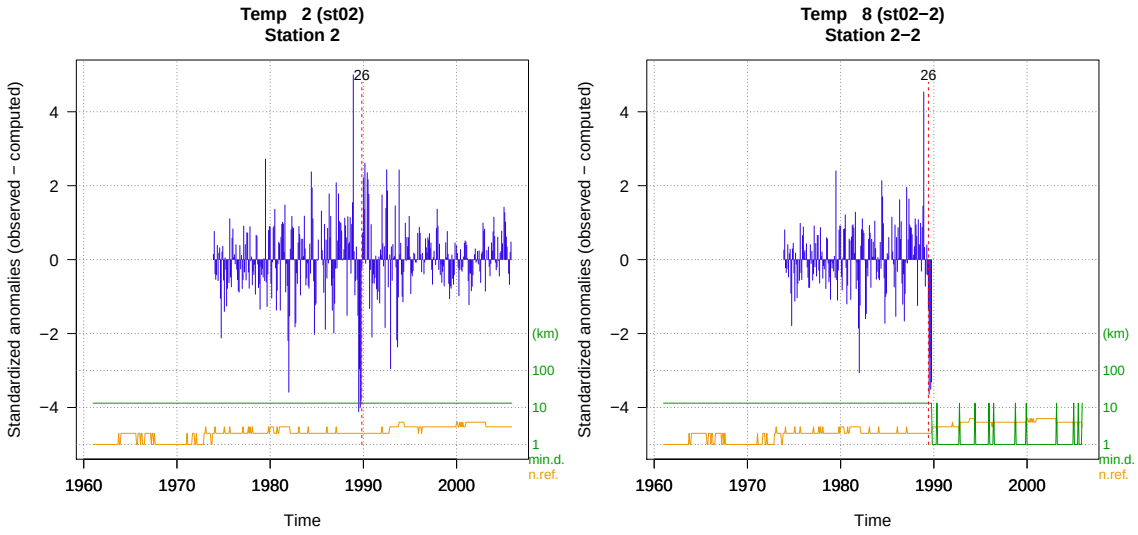


Figura 7: Detección de un breve periodo anómalo.

Tras las series de anomalías podemos ver los gráficos de las series reconstruidas y las correcciones aplicadas. La figura 8 muestra un ejemplo, con las anomalías de la serie 1 a la izquierda y la reconstrucción de series completas a partir de los dos fragmentos homogéneos a la derecha. El gráfico de anomalías presenta dos líneas adicionales en la parte inferior que informan sobre la distancia mínima de los datos vecinos (en verde) y el número de datos de referencia utilizados (en naranja), usando ambos la escala logarítmica del eje derecho. En el gráfico de series reconstruidas se representan sus medias[4] anuales móviles, salvo que las series sean muy cortas (hasta 120 términos), en cuyo caso se dibujan todos los valores. Las series originales se trazan en color negro[5], y en colores las reconstruidas.

Tras cada fase del proceso de homogeneización se presentan también histogramas de los valores residuales del test de homogeneidad empleado (SNHT por defecto) que pueden ayudar a variar el umbral `inht` si se desea repetir el proceso. Pero cuando trabajemos con pocas series (como en nuestro ejemplo) será difícil discriminar qué valor separa mejor las series homogéneas de las que no lo son. En ese caso convendrá revisar los gráficos de anomalías, como se ha comentado. Como por defecto se usa un máximo de 4 datos de referencia en la última fase (en lugar del máximo de 10 usado en las fases de detección), las últimas series de anomalías pueden presentar valores del test superiores al umbral utilizado, debido a su mayor variabilidad por reducirse el número de referencias.

[4]Totales si se da el parámetro `annual='total'`, que es lo recomendado para precipitación.

[5]Pero los valores anuales no se podrán calcular cuando falte algún dato, y entonces aparecerán con el color del fragmento reconstruido al que pertenezcan.

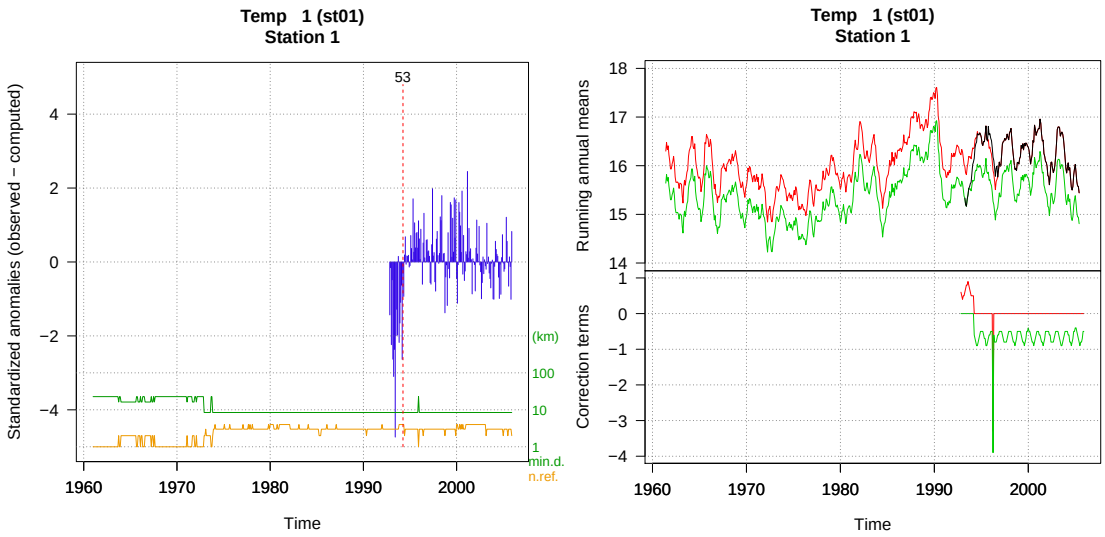


Figura 8: Ejemplo de detección de salto en la media (izquierda) y reconstrucción de las series a partir de cada fragmento homogéneo (derecha).

El histograma de anomalías que aparece casi al final del documento ya se ha comentado anteriormente al tratar de la homogeneización de las precipitaciones diarias `Prec`. Finalmente, la última página del fichero de gráficos contiene una figura que indica su calidad o singularidad, en la que las estaciones se sitúan según sus errores típicos (RMSE por sus siglas en inglés) finales y los valores del test de homogeneidad. Los RMSE se calculan al comparar los datos estimados y los observados en cada serie. Un valor alto puede indicar una mala calidad, pero también podría deberse a que la estación se encuentra en un sitio peculiar con un microclima distinto. De todos modos, las series homogéneas de estaciones que comparten el clima común de la región tenderán a agruparse en la parte inferior izquierda del gráfico.

1.3.4. Otros parámetros de la función `homogen`

Esta función tiene una gran cantidad de parámetros, como se puede ver en la documentación estándar de la misma. En general no hace falta especificarlos, pues tienen valores por defecto que suelen ser apropiados, pero según la variable estudiada o los primeros resultados de la homogeneización puede convenir cambiar sus valores, especialmente en los siguientes:

- `dz.max` fija los umbrales de rechazo de datos anómalos o de advertencia de datos sospechosos. Ejemplo: `dz.max=c(7,9)` eliminará los datos cuya anomalía sea mayor a 9 desviaciones típicas, y listará como sospechosos los que tengan anomalías entre 7 y 9 desviaciones típicas. Si se desea fijar umbrales diferentes en la cola izquierda de la distribución de anomalías se pueden dar valores al parámetro `dz.min`.
- `inht` es el umbral de inhomogeneidad, es decir, el valor del test de homogeneidad por encima del cual la serie se partirá en dos. La revisión de los histogramas del test y los gráficos de anomalías pueden sugerir variar el valor por defecto, que es 25. Si se quiere forzar la homogeneización directa de series diarias habrá que aumentar un orden de magnitud este valor (usando, por ejemplo, '`inht=250, force=TRUE`').
- `std` es el tipo de normalización que se aplica a los datos. Si se detecta que la variable está fuertemente sesgada y limitada por cero, por defecto se usará `std=2` (los datos se dividirán por su valor medio). Como esta es la normalización recomendada para precipitación y velocidad del viento, aunque en las series diarias se asignará automáticamente, no sucederá lo mismo con los valores mensuales, para los que convendrá especificar `std=2`. La normalización por defecto es `std=3` (restar la media y dividir por la desviación típica), válida para otras variables como temperatura, humedad relativa, presión atmosférica,

etc. Un tercer tipo de normalización que el usuario puede especificar es `std=1`, que únicamente centra los datos al restarles su valor medio.

- `vmin` y `vmax` sirven para limitar los valores posibles que pueden adoptar los datos. `vmin` se fija automáticamente a 0 si la normalización es `std=2`, pero por ejemplo para la humedad relativa convendrá especificar '`vmin=0, vmax=100`'.
- `nref` es el número máximo de datos próximos a usar para estimar los de la serie problema. Por defecto se utilizarán hasta 10 (si existen en cada paso de tiempo considerado) en las dos primeras etapas, y hasta 4 en la última, pero a veces puede convenir cambiar estos valores. Por ejemplo, en la homogeneización de precipitaciones diarias, usar 4 datos de referencia suavizará los datos estimados, con lo que aumentará el número de días de lluvia y disminuirá los valores máximos. Eso se evitará fijando `nref=1`, aunque un valor muy alto de la serie más próxima puede producir uno demasiado elevado en la serie problema si la precipitación media es mucho mayor de la de la serie vecina.
- `wd` especifica la distancia a la que el peso de los datos vecinos vale la mitad. Por defecto se establece `wd=c(0,0,100)`, con lo que no se asignarán pesos a los datos en las dos primeras etapas de detección de saltos en la media y anomalías espaciales, y en la tercera etapa (relleno final de todos los datos ausentes) los datos perderán peso con la distancia, tal como muestra la figura 9, de forma que a 100 km pesarán la mitad.

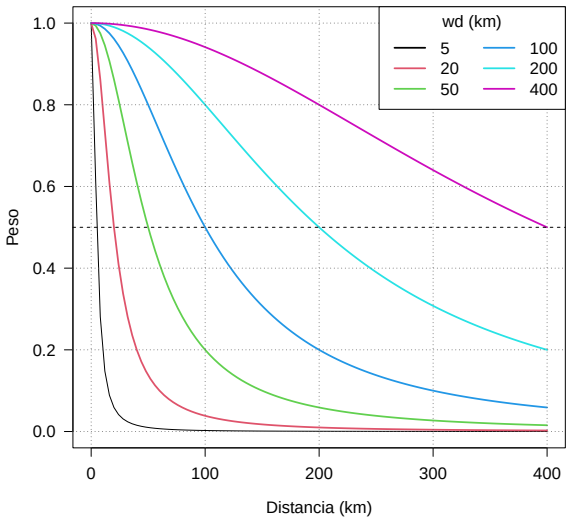


Figura 9: Variación de los pesos para distintos valores de wd.

1.3.5. Ficheros de resultados

Los únicos ficheros no comentados hasta ahora que también genera la función `homogen` son los que guardan los resultados en un archivo binario de R. Tienen el mismo nombre base que los demás, pero con extensión `rda`. En la documentación estándar de la función se dan detalles sobre su contenido. El usuario puede cargar los resultados de uno de los ejemplos anteriores en la memoria de trabajo de R para su manipulación mediante la orden:

```
load('Temp_1961-2005.rda')
```

Pero *climatol* proporciona funciones de postproceso para facilitar la obtención de productos a partir de las series homogeneizadas, por lo que en la mayoría de los casos no es necesario usar los ficheros `*.rda` directamente, como veremos a continuación.

1.4. Obtención de productos con los datos homogeneizados

Aunque el usuario puede cargar los resultados de la homogeneización como se indicó anteriormente, *climatol* provee las funciones de postproceso `dahstat` y `dahgrid` para facilitar la obtención de productos de uso corriente a partir de las series homogeneizadas.

1.4.1. Resúmenes estadísticos y series homogeneizadas

Las series homogeneizadas pueden volcarse a dos ficheros de texto CSV mediante la función `dahstat` especificando el parámetro `stat='series'`. En los ejemplos anteriores `Temp` eran temperaturas mensuales y `Prec` eran precipitaciones diarias, de las cuales se obtuvieron los valores mensuales grabados con el nombre de `Prec-m`. Por tanto podemos obtener las series homogeneizadas (ajustadas desde el último fragmento homogéneo hacia atrás) mediante:

```
dahstat('Temp', 1961, 2005, stat='series')   #temperaturas mensuales
dahstat('Prec', 1981, 1995, stat='series')   #precipitaciones diarias
```

Cada una de estas órdenes genera dos ficheros CSV. Los llamados `*_series.csv` contienen las series homogeneizadas, mientras que en los `*_flags.csv` aparecen códigos que indican si los datos son observados (0), rellenados (1, ausentes originalmente) o corregidos (2, por inhomogeneidades o por excesiva anomalía). Con una orden similar aplicada a `Prec-m` se podrían obtener las series de la homogeneización de las precipitaciones mensuales, pero es mejor calcular estas series a partir de las diarias, pues la ausencia de datos diarios en el momento del cálculo de los agregados mensuales hará que existan algunas diferencias. Para ello puede usarse la función `dahstat` con la opción `stat='mseries'`.

Los resúmenes estadísticos se crean con la misma función. Aquí se presentan algunos ejemplos (más información en la documentación de R de la función `dahstat`):

```
dahstat('Prec',1981,1995) #medias mensuales (el estadístico por defecto)
dahstat('Prec',1981,1995,stat='tnd') #tendencias y p-valores
dahstat('Prec',1981,1995,stat='q',prob=.2) #primer quintil
```

Esta función incluye parámetros para escoger un subconjunto de las series, bien dando una lista con los códigos deseados, como por ejemplo con `cod=c('p064','p084')`, o especificando que queremos las series reconstruidas desde el subperiodo más largo (`long=TRUE`). También podemos pedir los estadísticos para todas las series (reconstruidas desde todos los fragmentos homogéneos) mediante el parámetro `all=TRUE`.

1.4.2. Series de rejillas homogeneizadas

La otra función de postproceso, `dahgrid`, genera rejillas calculadas a partir de las series homogeneizadas (sin usar datos rellenados). Pero antes de aplicar esta función, el usuario debe definir los límites y la resolución de la rejilla, como en este ejemplo que usa los resultados de la homogeneización de `Temp` realizada anteriormente:

```
grd <- expand.grid(x=seq(-2.7,-2.5,.025),y=seq(38.8,39.1,.025)) #rejilla deseada
#la siguiente orden precisa que el paquete sp esté instalado:
sp::coordinates(grd) <- ~x+y #convertir la rejilla en un objeto espacial
```

La función de R `expand.grid` se usa para definir las secuencias de coordenadas X e Y, y luego se aplica `coordinates` (del paquete `sp`) para convertir la rejilla, guardada con el nombre `grd` (se podría haber usado cualquier nombre), en un objeto de clase espacial.

Ahora las rejillas homogeneizadas pueden generarse (en formato NetCDF) con:

```
dahgrid('Temp', 1961, 2005, grid=grd) #rejillas mensuales
```

Estas rejillas se han construido con valores normalizados adimensionales. Se pueden obtener nuevas rejillas con las unidades originales (°C en el ejemplo) por medio de herramientas externas, como las *Climate Data Operators (CDO)*, aprovechando que `dahgrid` también ha guardado rejillas con las medias (`*_m.nc`) y las desviaciones típicas (`*_s.nc`). Así, si las `CDO` están instaladas en nuestro sistema, podemos invocarlas desde R con:

```
orden <- paste('cdo add -mul Temp_1961-2005.nc Temp_1961-2005_s.nc',
  'Temp_1961-2005_m.nc Temp-u_1961-2005.nc')
system(orden)
```

Pero las nuevas rejillas contenidas en `Temp-u_1961-2005.nc` (podríamos haber dado cualquier nombre al fichero de salida, respetando la extensión `nc`) solo estarán basadas en interpolaciones geométricas, de modo que si hay montañas desprovistas de datos las variaciones climáticas esperables no se verán reflejadas en las rejillas. Para conseguir una mejor representación del clima de la zona estudiada deberían obtenerse mejores rejillas de medias `Temp_1961-2005_m.nc` y desviaciones típicas `Temp_1961-2005_s.nc` mediante métodos geoestadísticos antes de utilizarlas para obtener las rejillas de valores con sus unidades.

1.5. Dudas frecuentes sobre la función `homogen`

Los ejemplos anteriores muestran y comentan las aplicaciones más usuales de las funciones de homogeneización de `climatol`. No obstante, pueden surgir dudas relativas a cómo proceder cuando tratamos con otras variables climáticas o resoluciones temporales. Esta sección está dedicada a resolver las dudas más frecuentes.

1.5.1. Cómo guardar los resultados de diferentes pruebas

Si se ejecuta `homogen` con diferentes parámetros para explorar cuáles dan mejores resultados, puede evitarse sobrescribir las salidas anteriores renombrándolas con la ayuda de la función `outrename`. Por ejemplo, la siguiente orden renombrará todos los ficheros de salida `Temp_1961-2005*` a `Temp-old_1961-2005*`:

```
outrename('Temp', 1961, 2005, 'old')
```

1.5.2. Cómo cambiar el nivel de corte en el análisis de agrupamiento

En el análisis de agrupamiento que realiza *climatol* en su comprobación inicial de los datos, el número de grupos se determina automáticamente. Mirando el dendrograma (en los primeros gráficos del documento de salida en PDF) se puede elegir un nivel de corte diferente mediante el parámetro `cutlev`. Esto no tendrá ningún efecto sobre los resultados de la homogeneización, puesto que únicamente sirve para optimizar la división de las estaciones en grupos de régimen climático más similar en caso de que el usuario considere conveniente no homogeneizar todas las series conjuntamente. (La composición de los grupos aparecerá en la salida de texto y en el objeto `ct` guardado en el fichero binario `*.rda`).

1.5.3. Cómo usar series de reanálisis como referencias

Cuando las series están muy fragmentadas y algunos pasos temporales de nuestro periodo de estudio no tienen datos en ninguna de ellas, o bien cuando tratamos de homogeneizar una serie aislada, una solución es usar series procedentes de productos de reanálisis para servir como referencias que provean datos en esos huecos críticos.

Aunque la aparición de nuevos sistemas de observación (como los satélites) introduce inhomogeneidades en la cantidad de datos disponibles para su asimilación por los modelos, podemos considerar que los productos de reanálisis son más homogéneos en general que las series observadas. Para usar estos productos como referencias, las series de uno o más puntos de rejilla localizados en el dominio de estudio deben añadirse al fichero de datos `*.dat`, y las coordenadas de esos puntos añadirse al fichero de estaciones `*.est`. Sus códigos deben empezar por un asterisco (ejemplo: `*R43`) para que los controles de calidad y homogeneidad se salten esas series más confiables.

Como un estudio mostró que las series de reanálisis son peores referencias que las constituidas por observaciones, por defecto a las de reanálisis se les añaden 1000 km de distancia para que pesen menos que las observadas al usarlas para calcular datos por interpolación. Este valor por defecto se puede modificar mediante el parámetro `raway`.

1.5.4. Qué series homogeneizadas debería usar

La mayoría de métodos de homogeneización devuelven las series ajustadas desde el último sub-periodo homogéneo, pero *climatol* genera reconstrucciones completas a partir de cada subperiodo (salvo que sea demasiado corto para que dicha reconstrucción sea fiable). Por tanto, el usuario puede preguntarse cuál utilizar en su estudio climático. La respuesta depende del objetivo de la investigación. Para obtener valores normales con los que calcular las anomalías de nuevos datos entrantes para monitorización climática, usará las series ajustadas a partir del último sub-periodo homogéneo (la opción por defecto de `dahstat`). Pero si el objetivo es hacer mapas, deberían considerarse todas las series (añadiendo `all=TRUE` a los parámetros de `dahstat`), escogiendo luego las que se ajusten mejor a la variabilidad espacial de la escala del mapa e ignorando las que se sospeche que están afectadas por microclimas locales.

1.5.5. El proceso está durando demasiado tiempo

Esto puede suceder si estamos procesando muchas series, muy largas y con muchos saltos en la media. Ejemplo: 400 series termométricas diarias de 1951-2020. La homogeneización directa de estas largas series diarias puede durar días, sobre todo si no se han dado valores suficientemente altos a los parámetros `inht` y `dz.max`, porque entonces las series estarían sufriendo un elevado número de cortes y rechazando demasiados datos que luego se tendrán que rellenar.

Si se sigue el procedimiento recomendado de homogeneizar primero las series mensuales obtenidas con la función `dd2m`, el tiempo de proceso será mucho más corto. Pero de todos modos conviene considerar si es necesario homogeneizar todas las series a la vez, porque probablemente sea mejor dividir nuestros datos en conjuntos más reducidos, agrupando las series según subregiones climáticamente más uniformes. (Se puede usar la función `datsubset` para generar los ficheros de un grupo seleccionado de estaciones, que puede estar basado en el análisis de agrupamiento generado por `homogen`, adaptado por el usuario según su conocimiento del clima y la fisiografía de la zona).

1.5.6. ¿Puede usarse *climatol* para homogeneizar series de caudales?

Para poderlo usar debe haber cierto grado de correlación (positiva) entre las series. Se puede hacer una prueba con `onlyQC=TRUE` y ver el correlograma. Entre caudales de diferentes puntos de una misma cuenca es posible que exista la necesaria correlación (aunque probablemente con cierto desfase temporal). Otra posibilidad sería usar como serie de referencia caudales generados con un modelo hidrológico.

1.6. Bibliografía

Aguilar E, Auer I, Brunet M, Peterson TC, Wieringa J (2003): *Guidelines on climate metadata and homogenization*. WCDMP-No. 53, WMO-TD No. 1186. World Meteorological Organization, Geneve. LINK

Alexandersson H (1986): A homogeneity test applied to precipitation data. *Jour. of Climatol.*, 6:661-675.

Cucconi O (1968): Un nuovo test non parametrico per il confronto tra due gruppi campionari. *Giornale degli Economisti*, 27:225-248.

Guijarro JA, López JA, Aguilar E, Domonkos P, Venema VKC, Sigró J, Brunet M (2023): Homogenization of monthly series of temperature and precipitation: Benchmarking results of the MULTITEST project. Int. J. Climatol., 19 pp, DOI 10.1002/joc.8069 LINK

Khaliq MN, Ouarda TBMJ (2007): On the critical values of the standard normal homogeneity test (SNHT). *Int. J. Climatol.*, 27:681687.

Killick RE (2016): Benchmarking the Performance of Homogenisation Algorithms on Daily Temperature Data. PhD Thesis, University of Exeter, 249 pp. LINK

OMM (2020): Directrices sobre homogeneización. OMM Nº 1245, 61 pp., Ginebra, Suiza, ISBN 978-92-63-31245-8. LINK

Paulhus JLH, Kohler MA (1952): Interpolation of missing precipitation records. *Month. Weath. Rev.*, 80:129-133.

Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Böhm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Førland E, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D (1998): Homogeneity Adjustments of 'In Situ' Atmospheric Climate Data: A Review. *Int. J. Climatol.*, 18:1493-1518.

Sokal RR, Rohlf PJ (1969): *Introduction to Biostatistics*. 2^{nd} edition, 363 pp, W.H. Freeman, New York.

Venema V, Mestre O, Aguilar E, Auer I, Guijarro JA, Domonkos P, Vertacnik G, Szentimrey T, Stepanek P, Zahradnicek P, Viarre J, Müller-Westermeier G, Lakatos M, Williams CN, Menne M, Lindau R, Rasol D, Rustemeier E, Kolokythas K, Marinova T, Andresen L, Acquaotta F, Fratianni S, Cheval S, Klancar M, Brunetti M, Gruber C, Prohom Duran M, Likso T, Esteban P and Brandsma T (2012): Benchmarking homogenization algorithms for monthly data. *Clim. Past*, 8:89-115.

2. Otras funciones

Además de las funciones de homogeneización explicadas hasta ahora, *climatol* también proporciona algunas utilidades y productos gráficos que se mostrarán brevemente a continuación (ver el manual estándar para conocer todos los detalles sobre su uso).

2.1. Utilidades

- **`fix.sunshine`** sirve para recortar los excesos de horas de sol que puedan haberse producido al ajustar las series diarias (ver ejemplo en la ayuda de la función).
- **`QCthresholds`** permite obtener, para cada serie diaria (o subdiaria), cuantiles mensuales de valores extremos, de incrementos entre valores consecutivos y de secuencias de valores idénticos. Estos cuantiles se pueden utilizar para implementar alertas de control de calidad en los sistemas de gestión de datos climáticos. Ejemplo para las series `Prec` homogeneizadas anteriormente (se establece un valor mínimo para evitar contabilizar las series largas de ceros consecutivos):

```
QCthresholds('Prec_1981-1995.rda',minval=0.1)
```

```
========== thr1: Monthly quantiles of the data
--------- Station p064
         1    2    3    4    5    6    7     8    9    10   11   12
0      0.0  0.0  0.0  0.0  0.0  0.0  0.0   0.0  0.0   0.0  0.0  0.0
0.001  0.0  0.0  0.0  0.0  0.0  0.0  0.0   0.0  0.0   0.0  0.0  0.0
0.01   0.0  0.0  0.0  0.0  0.0  0.0  0.0   0.0  0.0   0.0  0.0  0.0
0.99  35.1 40.9 44.5 33.1 31.6 31.0 31.9  49.5 46.7  60.2 63.9 42.9
0.999 54.1 50.8 56.5 54.2 69.4 44.6 48.3  91.2 71.7 120.2 91.9 56.1
1     54.3 53.8 57.1 62.7 92.8 49.0 54.2 110.1 73.1 140.0 93.8 57.1
--------- Station p084
         1    2    3    4    5    6    7    8     9    10   11   12
0      0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0   0.0   0.0  0.0  0.0
0.001  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0   0.0   0.0  0.0  0.0
0.01   0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0   0.0   0.0  0.0  0.0
0.99  24.1 31.3 29.0 27.9 30.0 29.7 30.9 36.6  38.0  51.5 49.2 30.1
0.999 39.8 42.2 40.7 46.3 50.1 46.5 40.4 67.5  74.6  87.8 67.2 39.8
1     42.2 48.2 44.2 52.8 51.1 51.6 42.0 76.7 100.4 116.5 71.3 41.1
```

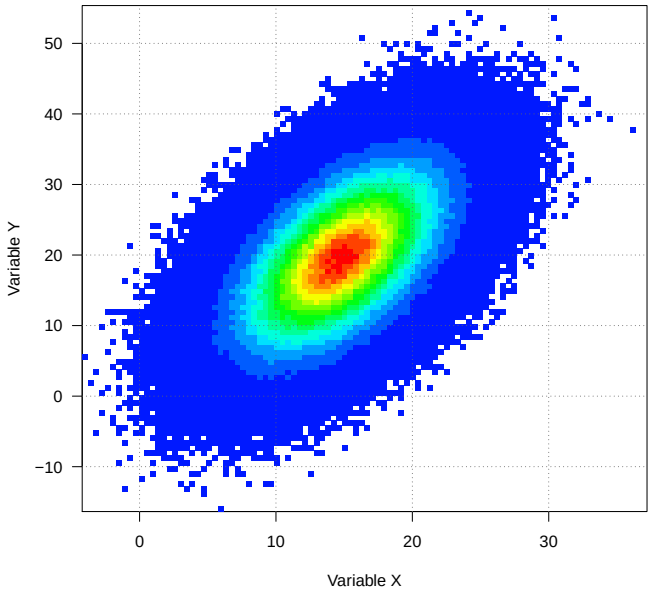
```
--------- Station p082
         1    2    3    4    5    6    7    8    9   10   11   12
0      0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
0.001  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
0.01   0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
0.99  17.9 24.2 21.4 23.1 21.0 26.4 28.2 33.2 29.3 38.2 31.8 19.7
0.999 24.3 32.7 29.9 31.5 34.4 34.0 47.1 48.0 61.4 63.4 48.0 24.5
1     26.9 37.6 30.4 33.8 36.8 34.0 47.9 52.0 78.7 78.0 51.2 26.3

========== thr2: Quantiles of the first differences
     0.99 0.999     1
p064 46.1  84.0 129.4
p084 37.8  66.9 103.6
p082 28.8  49.2  78.5

========== thr3: Quantiles of run lengths of constant values >= 0.1
     0.99 0.999 1
p064    2     3 4
p084    2     3 3
p082    2     3 3

Thresholds thr1,thr2,thr3 saved into QCthresholds.Rdat
(Rename this file to avoid overwriting it in the next run.)
```

Con `load('QCthresholds.Rdat')` tendríamos estos resultados en la memoria de la sesión de R y podríamos escribirlos en el formato adecuado para importarlos en el sistema de gestión de datos climáticos y poder implementar las alertas de valores sospechosos.

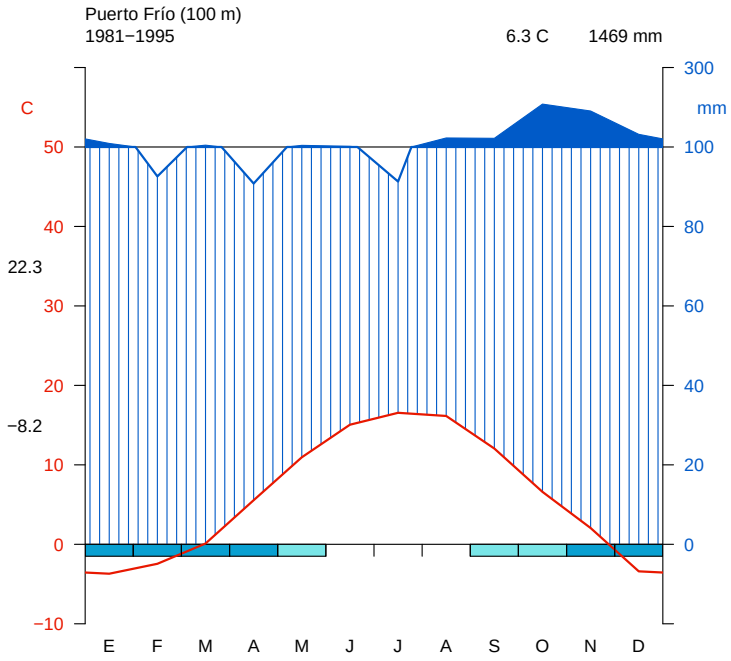
2.2. Productos gráficos

En este apartado únicamente se mostrarán ejemplos de las funciones que producen gráficos de uso climatológico. El manual estándar de *climatol* muestra los detalles de cada una de ellas.

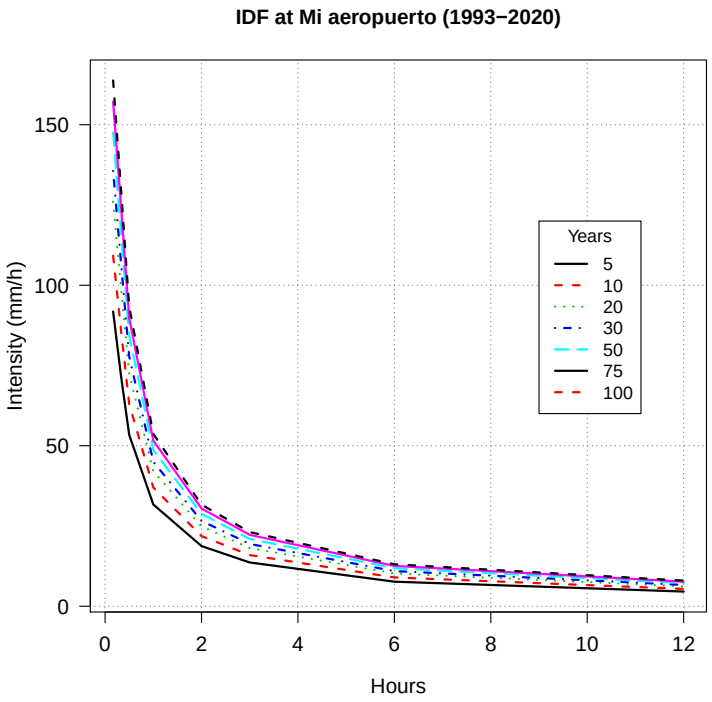
2.2.1. dens2Dplot: Diagrama de dispersión bidimensional



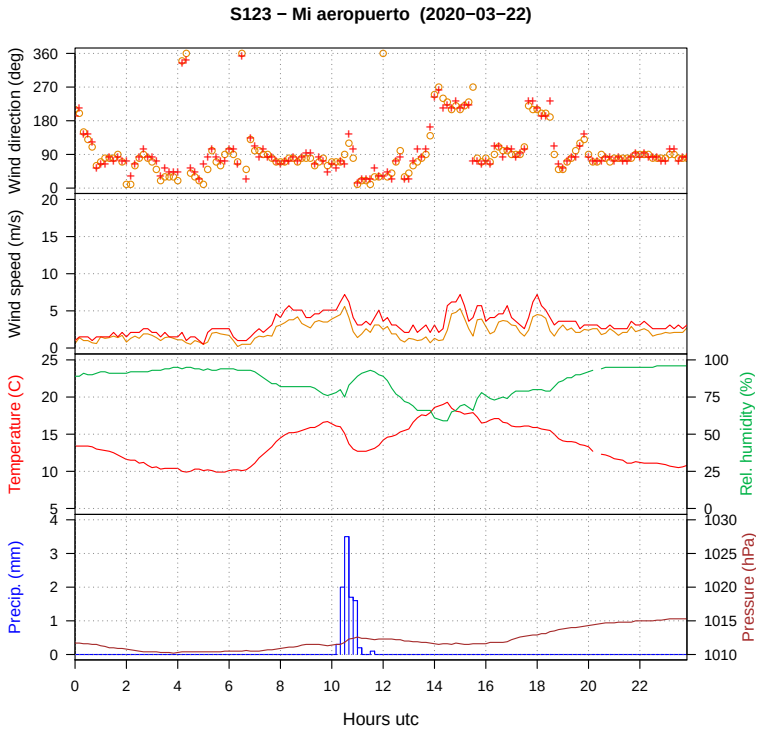
2.2.2. diagwl: Diagrama de Walter y Lieth



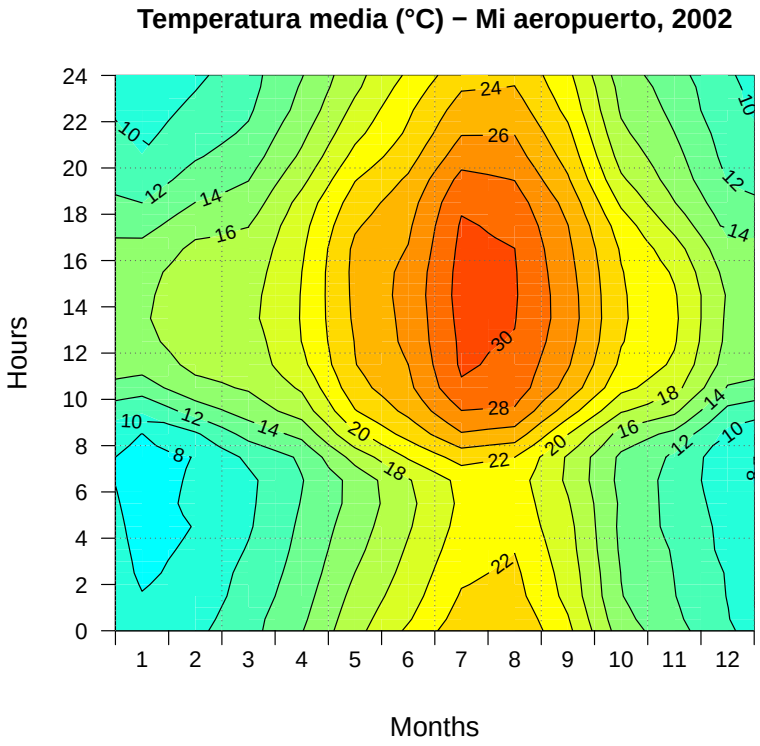
2.2.3. IDFcurves: Diagrama de intensidad-duración-frecuencia a partir de datos subdiarios de precipitación



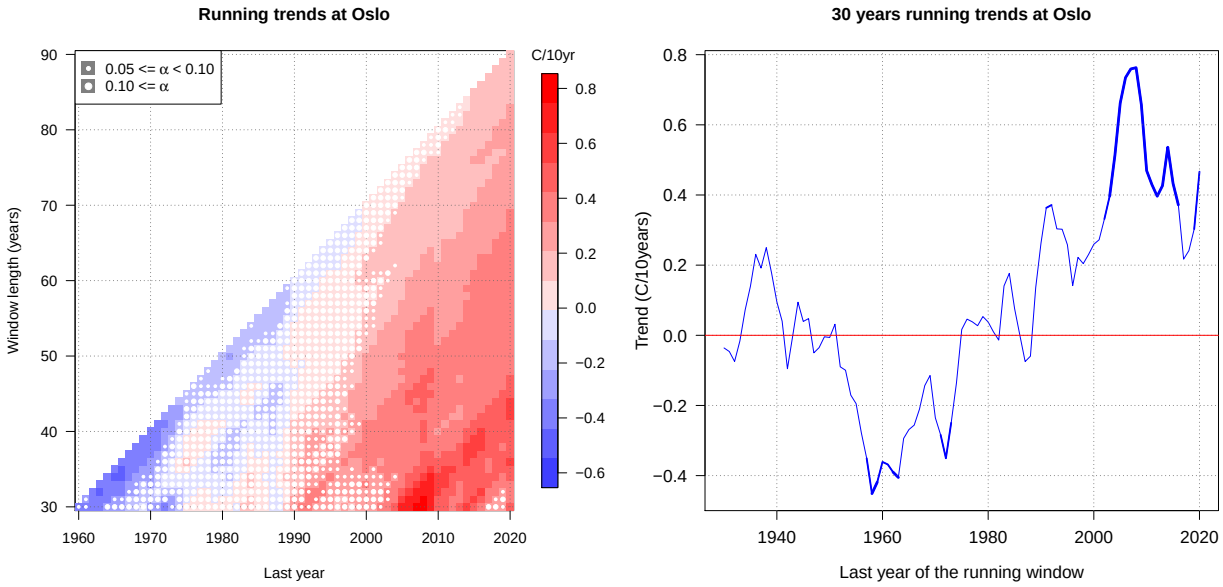
2.2.4. meteogram: Meteograma de 1 día



2.2.5. MHisopleths: Isopletas en un diagrama Meses-Horas



2.2.6. runtnd: Diagramas de tendencias móviles



2.2.7. windrose: Rosa de los vientos a partir de datos de dirección y velocidad del viento