

Manual para el cálculo de índices de sequía a partir de series sintéticas

Alessio Bocco

19 April, 2021

Contents

1	Introducción	2
1.1	Índice de Precipitación Estandarizado (SPI)	2
1.2	Índice de Precipitación – Evapotranspiración Estandarizado (SPEI)	2
2	Metodología	3
2.1	Cálculo en distintas escalas temporales	3
2.2	Período de referencia para el cálculo de índices	3
2.3	Actualización frecuente de índices basada en péntadas	4
3	Pasos necesarios para el cálculo de índices de sequía	5
3.1	Definición de las configuraciones para cálculo de índices	6
3.2	Cálculo de valores agregados o estadísticos con ventana móvil	7
3.3	Cálculo de índices	7
4	Ejemplo de aplicación	10
4.1	Generación de configuraciones	10
4.2	Resultados	15
	Bibliografía	19

1 Introducción

En el SISSA se utilizan datos climáticos observados in situ para calcular cuatro índices de sequía en función de las necesidades de los miembros y usuarios del sistema. Éstos son: (i) Índice de Precipitación Estandarizado (SPI) (McKee *et al.* (1993)); (ii) Índice de Precipitación – Evapotranspiración (SPEI) (Vicente-Serrano *et al.* (2010)); (iii) Deciles de Precipitación; y (iv) Porcentaje de Precipitación Normal (PPN). Sin embargo, con las series sintéticas solamente se calcularán dos índices – SPI y SPEI – para la caracterización de la amenaza de sequía, el primer paso hacia el análisis probabilista del riesgo de sequías.

1.1 Índice de Precipitación Estandarizado (SPI)

El Índice de Precipitación Estandarizado (IPE, o SPI por sus siglas en inglés) cuantifica las condiciones de déficit o exceso de precipitación en un lugar y para una escala determinada de tiempo (World Meteorological Organization 2012). El SPI fue desarrollado por McKee *et al.* (1993) con la finalidad de mejorar la detección del inicio y el monitoreo de la evolución de las sequías meteorológicas (definidas únicamente en función de la precipitación). La principal ventaja de este índice es que su cálculo requiere de una única variable climática: la precipitación. Utilizar sólo precipitación requiere de dos asunciones: (i) la variabilidad de la precipitación es mucho mayor que la de otras variables climáticas como la temperatura o evapotranspiración; y (ii) otras variables son estacionarias, es decir, sus niveles promedio no cambian en el tiempo. En 2009 la Organización Meteorológica Mundial recomendó el SPI como el principal índice que los países deben utilizar para el monitoreo de las sequías (Hayes *et al.* (2011)).

1.2 Índice de Precipitación – Evapotranspiración Estandarizado (SPEI)

El Índice de Precipitación – Evapotranspiración Estandarizado (IPEE o SPEI por sus siglas en inglés), es un índice cuyo cálculo es similar al del SPI, pero incorpora el efecto de la evapotranspiración (i.e., la demanda atmosférica de agua) que influye en las condiciones de sequía. El SPEI fue desarrollado por Vicente-Serrano *et al.* (2010) y ya ha sido ampliamente utilizado para analizar distintas características de la sequía, como ser su variabilidad, impactos y mecanismos atmosféricos que la producen (Beguiría *et al.* (2013); Vicente-Serrano *et al.*, (2015)).

El SPEI utiliza como valor de entrada al balance hídrico (es decir, la diferencia entre precipitación y evapotranspiración potencial o PET por sus siglas en inglés). El cálculo de la evapotranspiración potencial es complicado debido a que involucra muchas variables meteorológicas (temperatura, humedad del aire, viento y radiación, entre otros). En este proyecto se ha utilizado la ecuación de Hargreaves y Samani ((1982)) modificada por Droogers and Allen (2002), que es eficiente en el cálculo de la evapotranspiración potencial utilizando sólo medias mensuales de temperatura máxima y mínima, y radiación solar extraterrestre.

Dado que el balance entre precipitación y evapotranspiración puede tomar valores negativos, el ajuste del mismo se realiza con la distribución teórica Log-Logística, que acepta valores nulos y negativos.

2 Metodología

2.1 Cálculo en distintas escalas temporales

Los efectos de las sequías se manifiestan en diferentes escalas temporales, ya que las respuestas de diferentes sistemas hidrológicos y biológicos a las anomalías de precipitación varían mucho (Ji and Peters, 2003). Es decir, puede haber grandes diferencias en la duración de los déficits hídricos necesarios para causar impactos negativos en diferentes sistemas. Como la sequía es un fenómeno multiescalar, es necesario el uso de indicadores que puedan capturar adecuadamente las escalas temporales relevantes para detectar impactos negativos sobre los diferentes sistemas de interés. Por ejemplo, la sequía agrícola – definida como la escasez de agua para satisfacer las necesidades de un cultivo – puede ser bien representada por las escalas de 2 y 3 meses, mientras que déficits en los caudales de río o arroyos se reflejan mejor por medio de las escalas de 3 a 6 meses. Asimismo, se han encontrado asociaciones entre la variación del nivel de la napa freática y los valores de los índices con escalas de 6 a 24 meses. Todos los índices de sequía que se han descrito en secciones anteriores están siendo calculados para múltiples escalas temporales: 1, 2, 3, 6, 9, 12, 18, 24, 36 y 48 meses. Es decir, los índices se basan en series de precipitaciones acumuladas para cada escala. Por ejemplo, para calcular un índice con una escala de 6 meses para junio de 2003, se considera la suma de los valores mensuales de precipitación desde enero hasta junio de 2003; este valor se pone en el contexto de los totales de lluvia para enero-junio en el registro histórico. De la misma forma, se puede realizar el cálculo para otras escalas temporales. Para el cálculo del SPEI, que requiere temperaturas máximas y mínimas para estimar la evapotranspiración, se usa el promedio de estas temperaturas para cada escala temporal. Siguiendo con el ejemplo anterior, el cálculo de SPEI para una escala de 6 meses para junio de 2003 utilizará los promedios mensuales de temperaturas máximas y mínimas desde enero hasta junio de 2003.

2.2 Período de referencia para el cálculo de índices

Para el cálculo de los diferentes índices de sequía, generalmente es necesario estimar parámetros o cuantiles de la distribución empírica o teórica de precipitación. En el caso de la estimación de índices a partir de series sintéticas, se debe definir un período común para todas las realizaciones para que los valores de los índices obtenidos sean comparables. Para ello se determina un período de referencia fijo para el cual se estiman los parámetros necesarios. Es deseable que el período de referencia contenga la mayor parte de la variabilidad de las series climáticas. Para las series sintéticas, la importancia de la longitud del período de referencia escogido dependerá del tipo de series producidas por el generador estocástico. Si las series sintéticas son estacionarias, la variabilidad de los datos será aproximadamente la misma para

toda la serie y para todas las realizaciones. Esto quiere decir que la definición del período de referencia no será una decisión crítica. El ajuste obtenido usando un período de referencia comprendido entre 1971-2010 (que se utiliza para el cálculo de índices a partir de observaciones in situ) sería igual al obtenido usando el período 1961-2000 o cualquier otro rango de fechas ya que los valores generados para todos los años tienen la misma probabilidad.

En el caso de series pseudohistóricas o condicionadas por modelos de cambio climático, el período de referencia debe ser lo suficientemente largo para capturar oscilaciones de baja frecuencia (épocas secas y húmedas asociadas con la variabilidad climática multidecadal). Para este tipo de series el período de referencia es muy importante, ya que los valores de los índices calculados a partir de distintos ajustes pueden cambiar de manera importante. Por ejemplo, si se utiliza un período en donde las precipitaciones fueron relativamente bajas, es posible que se subestimen los eventos secos cuando las precipitaciones sean un poco mayores. Lo contrario ocurriría con los eventos húmedos que serían sobreestimados a partir de pequeños aumentos en la precipitación. Por lo tanto, para los índices calculados a partir de series pseudohistóricas, se definió el período 1971-2010 (40 años) como período de referencia. Aunque los años en series sintéticas no corresponden a los años reales, como las series pseudohistóricas están condicionadas por valores de covariables (ej., totales trimestrales de lluvia) calculados usando observaciones históricas para esos años, se usa la convención de nombrar las series sintéticas con los años de la serie histórica usada para el ajuste de parámetros para el generador estocástico.

2.3 Actualización frecuente de índices basada en péntadas

Como se mencionó en el documento correspondiente al cálculo de Estadísticas Móviles, la demanda de una mayor frecuencia de actualización de los índices motivó el cambio del método de cálculo para aumentar la frecuencia temporal de valores actualizados de índices de sequía. Para más detalles del cambio se sugiere ver la página 4 de dicho reporte. El cálculo frecuente de índices se basa en el concepto de péntadas (o períodos de aproximadamente 5 días). Se utiliza aquí una definición flexible: (a) todos los meses tienen 6 péntadas y (b) las cinco primeras péntadas dentro de un mes tienen una duración de 5 días, en tanto que la sexta y última péntada de un mes puede tener 3, 4, 5 o 6 días – para meses con 28, 29, 30 o 31 días, respectivamente (Funk et al., 2015). Por ejemplo, la primera péntada de un mes incluye los días 1 a 5 de ese mes, la segunda péntada los días 6 a 10, y así sucesivamente. Una ventaja de esta definición es que las péntadas no cruzan límites de meses, lo que hace sencilla la agregación de datos a diferentes escalas. Por ejemplo, los totales mensuales de precipitación se calculan como la suma de totales para las 6 péntadas en cada mes/año. Siguiendo esta lógica, los índices de sequía se calculan para cada péntada del año. Sin embargo, se debe enfatizar que, aunque los valores de los índices se actualizan cada 5 días aproximadamente, esos valores se calculan usando valores agregados para cada escala temporal especificada. La escala temporal de un índice (o sea, el período para el cual se acumula la precipitación o se promedian las temperaturas) se define en función de péntadas, y no de meses. Sin embargo, para facilitar la comprensión de los resultados, las escalas de los índices se expresan en meses (la cantidad de péntadas dividida por seis). Por ejemplo, un SPI-3 es un índice que acumula precipitación por 18 péntadas, o aproximadamente 3

meses. Una de las principales ventajas del cálculo por péntadas es la precisión con la que es posible definir un evento seco. Anteriormente, la mínima resolución de detección era un mes completo, lo que ciertamente es un problema para cuantificar los impactos sobre algunos sistemas. Por ejemplo, en los sistemas agrícolas los momentos críticos de máxima sensibilidad son acotados y no necesariamente coinciden con meses calendario.

3 Pasos necesarios para el cálculo de índices de sequía

En esta sección se describe brevemente el flujo de cálculo de índices de sequía a partir de series climáticas sintéticas. La Figura 1 muestra un diagrama de flujo del proceso. Los pasos de este flujo se describen brevemente aquí. Para simplificar la discusión, los pasos involucrados se describen para una realización de series sintéticas. Si la simulación incluye múltiples realizaciones, la mayoría de los pasos son idénticos para cada realización; se notarán los pasos que utilicen todas las realizaciones simultáneamente. El flujo ilustrado en la Figura 1 es igual para ambos índices calculados (SPI y SPEI); sólo varían (a) los estadísticos necesarios para cada índice y (b) las funciones de probabilidad usadas en los ajustes de la climatología.

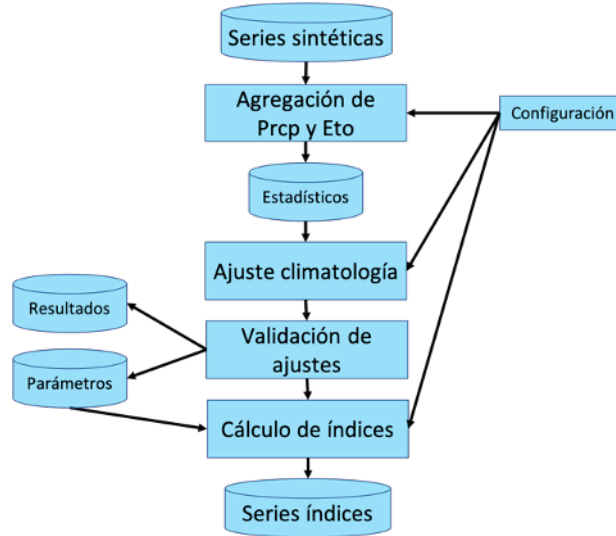


Figure 1: Flujo del proceso de cálculo de índices de sequía a partir de series sintéticas

El cálculo comienza con las series sintéticas (precipitación diaria y temperaturas máxima y mínima) simuladas previamente por el generador estocástico (Se sugiere leer la documentación sobre el mismo en <https://github.com/CRC-SAS/weather-generator>). Las series sintéticas están almacenadas en un archivo en formato de texto separados por comas (.csv). Las series pueden simularse (i) para una ubicación donde existen estaciones meteorológicas, (ii) para puntos geográficos determinados donde no hay estaciones meteorológicas, o (iii) para una serie de puntos que forman una grilla regular en el espacio. Asimismo, las series sintéticas pueden ser estacionarias o pseudohistóricas (estos términos se definen más abajo) ya que el cálculo de índices de sequías es el mismo para ambos tipos.

En primer lugar, las variables climáticas simuladas a paso diario para cada estación meteorológica o punto geográfico son agregadas a las escalas temporales especificadas para los índices de sequía; los datos agregados se guardan en la base de datos para uso posterior. Los datos agregados incluyen, por ejemplo, los totales de precipitación para cada escala temporal deseada; estas sumas se calculan para cada p ntada en la serie de datos sint ticos. Para las temperaturas m ximas y m nimas diarias utilizadas en el SPEI, la agregaci n implica el c lculo de valores promedio por p ntada para cada escala temporal. En este documento, los valores agregados tambi n se denominan estad sticos. A continuaci n, se seleccionan los datos agregados comprendidos dentro del per odo de referencia especificado para construir la climatolog a de cada estaci n meteorol gica (ver Secci n 2.3 sobre per odos de referencia). El per odo de referencia es un segmento de la serie (que puede incluir toda una serie sint tica) que es considerado representativo del clima local. Por ejemplo, si se deseara calcular el SPI-3 para el 31 de enero de 2019 a partir de series pseudohist ricas, se puede utilizar el per odo de referencia 1971–2010 (el mismo utilizado para las series in situ). Este per odo de referencia contiene la precipitaci n sint tica acumulada para los trimestres (o 18 p ntadas) de noviembre a enero de cada a o entre 1971 y 2010. Recordamos al lector que los a os mencionados no se refieren a fechas reales, sino a a os que se asignan a los valores simulados con covariables calculadas a partir de datos hist ricos para esos a os.

Con los valores agregados para el per odo de referencia, se estiman los par metros correspondientes a la distribuci n te rica utilizada para un  ndice (ej., la distribuci n gamma para el SPI). Los par metros estimados se usan luego para transformar la precipitaci n acumulada para el per odo entre el 1 de noviembre de 2018 y el 31 de enero de 2019 en los valores del SPI-3 para el 31 de enero de 2019. A continuaci n, se ajusta una funci n de probabilidad a todos los valores dentro del per odo de referencia para cada p ntada. Luego, se comprueba la bondad de ajuste de las distribuciones para asegurar que los valores del  ndice calculados a posteriori sean confiables. Si los ajustes son v lidos, los par metros de cada distribuci n ser n utilizados a continuaci n para calcular los  ndices de sequ a a partir de los valores agregados o estad sticos. En las subsecciones siguientes se dan m s detalles sobre cada paso para el c lculo de  ndices de sequ a.

3.1 Definici n de las configuraciones para c lculo de  ndices

Al igual que para series clim ticas observadas, es necesario definir o especificar la configuraci n deseada para el c lculo de  ndices de sequ a a partir de series sint ticas. La configuraci n de un  ndice se define como la combinaci n  nica de (i) el  ndice a calcular, (ii) la escala temporal deseada, (iii) el per odo de referencia y (iv) la metodolog a de ajuste. La cantidad de configuraciones posibles es definida por el usuario en funci n de sus necesidades. En este caso, se definen 40 configuraciones posibles, que resultan de la combinaci n de 2  ndices (SPI y SPEI), 10 escalas temporales (1, 2, 3, 6, 9, 12, 18, 24, 36 y 48 meses), 2 m todos de ajuste (por m xima verosimilitud o no param trico), y un  nico per odo de referencia para series pseudohist ricas (1971–2010). Este set de 40 configuraciones es v lido para series pseudohist ricas, ya que  stas requieren de un per odo de referencia determinado. Para series estacionarias, como el per odo de referencia es flexible,  ste no necesariamente debe ser 1971–2010. Todas las posibles configuraciones especificadas se guardan en la base de

datos local. Esta tabla se denominará de ahora en más tabla de configuraciones de cálculo. La Tabla 2 muestra sólo las primeras filas de esta tabla y es un ejemplo del resultado de esta etapa. Cada configuración posee un índice unívoco (id) que permite su identificación durante el proceso de cálculo de índices. La variable escala expresa la escala de agregación del índice medida en meses. Como se mencionó más arriba, internamente el script maneja la escala de agregación en péntadas, pero para facilitar la interpretación se decidió expresarla en meses completos.

3.2 Cálculo de valores agregados o estadísticos con ventana móvil

El insumo primario principal para el cálculo de índices de sequía son los valores de variables climáticas (precipitación, temperaturas máxima y mínima) simulados a paso diario. Estos valores, también llamados estadísticos en esta discusión, se agregan para las distintas escalas temporales definidas previamente. La agregación involucra (a) la suma de precipitaciones diarias y (b) los promedios de temperaturas máximas y mínimas diarias, para la escala temporal seleccionada. Los valores agregados o estadísticos para cada escala temporal se calculan para toda la serie sintética, repitiéndose el cálculo para cada péntada. Para el cálculo de valores agregados se implementó un script en lenguaje R (R Development Core Team, 2008) que procesa los datos diarios simulados de las tres variables definidas para todas las localidades (o puntos geográficos) y realizaciones generadas. Este script se ejecuta cada vez que un nuevo set de datos es simulado utilizando el generador estocástico de datos meteorológicos. Si se realiza una nueva simulación con otra configuración del generador, ya sea para otras fechas o localidades, las estadísticas también deben recalcularse porque las nuevas series generadas no serán necesariamente iguales a las originales dado el carácter estocástico del generador. Una vez que ha finalizado el proceso de agregación, los datos obtenidos, de aquí en más llamados estadísticos, se guardan en una base de datos local. Dada una fecha de fin de una péntada (días 5, 10, 15, 20, 25 o fin del mes), el proceso se encarga de generar ventanas del ancho correspondiente a cada escala temporal y de computar los estadísticos previamente enumerados con los valores dentro de esa ventana. Luego, el proceso se mueve una péntada hacia adelante y se vuelve a realizar el mismo cómputo. Este proceso iterativo finaliza una vez que se completa la última péntada de la serie simulada. Además de los estadísticos, se guarda un registro de todas las tareas realizadas por el script. En este registro se detalla el nombre y origen de las series sintéticas para así asegurar la trazabilidad de los datos de inicio y evitar confusiones en el caso que se realicen múltiples simulaciones mediante el generador. Por último, notamos que el uso de series sintéticas tiene la ventaja que éstas no poseen datos faltantes por lo que no es necesario un procesamiento especial para el cálculo de los estadísticos.

3.3 Cálculo de índices

En esta subsección se describirá específicamente el cálculo de los índices partiendo de los estadísticos o valores agregados y las configuraciones especificadas por el usuario. El proceso comienza al leerse los estadísticos para una estación meteorológica o punto de grilla,

incluyendo todas las realizaciones. Es decir, se agrupan todas las realizaciones para un mismo punto en una sola tabla. Dado que cada estación/realización es independiente de las demás, el proceso se paraleliza para optimizar los tiempos de cálculo. Una vez que se tienen los datos estadísticos de una estación, el paso siguiente es la estimación de la evapotranspiración potencial (ETP) utilizando el método de Hargreaves-Samani (Hargreaves and Samani, 1985). El conjunto de datos resultante es, a su vez, manipulado y transformado con el fin de optimizar los procesos posteriores. Llamaremos a este conjunto de datos estadísticas preprocesadas. Considerando los datos de cada una de las 40 configuraciones de cálculo posibles, se seleccionan los datos de las estadísticas preprocesadas cuyo ancho de ventana coincida con la escala temporal del índice a calcular. Es decir, de la tabla que contiene los estadísticos para todas las escalas temporales posibles se filtran aquellos estadísticos que fueron agregados en la escala detallada en la configuración. También en esta etapa se definen las fechas para las cuales se realizará el cálculo de los índices de sequía (fechas procesables). Las fechas procesables serán iguales a las fechas simulados, a menos que el usuario quisiera calcular los índices para un período acotado de tiempo. Por ejemplo, si se simularon 1000 años con el generador y sólo se necesitan los valores de los índices para los primeros 500 años.

En el caso de las series sintéticas no se guarda una climatología en la base de datos como se hace para las series in situ. Esto se debe a que cada generación de series será diferente de la anterior. Es decir, existen diferentes tipos de series – estacionarias, pseudohistóricas, condicionadas por un modelo de cambio climático – que produce el generador estocástico presentado en el Reporte 1. Por este motivo, una misma estación meteorológica puede tener distintos valores simulados de precipitación y temperaturas máxima y mínima. Además, si bien se puede especificar una semilla para que los resultados de una misma configuración del generador se repitan, para evitar errores siempre se calcula la climatología para cada set de datos sintéticos. El ajuste de la climatología se realiza utilizando el método indicado en la configuración y las estadísticas preprocesadas para el período de referencia (también definido en la configuración). El flujo de este proceso termina con el cálculo del índice para una estación o ubicación geográfica (o punto de grilla) y realización a partir de los datos agregados de cada fecha procesable y los ajustes de la climatología de esa estación. Además del valor del índice se incorpora el percentil correspondiente a ese valor en función de la climatología y el valor del estadístico para esa fecha (por ejemplo, para el caso de un SPI-3 será un dato de precipitación acumulada de 3 meses). Esta tabla luego es guardada en la base de datos. Este flujo es iterativo y se repite para todas las estaciones/realizaciones y fechas procesables, así como para todas las configuraciones definidas por el usuario. A continuación, se describirán con más detalles cada uno de los pasos mencionados en esta sección.

3.3.1 Métodos de ajuste de climatología

El ajuste de la climatología se realiza utilizando dos metodologías: i) ajuste paramétrico por máxima verosimilitud (EMV) y ii) ajuste no paramétrico (logsplines). Ambos métodos son los mismos que se utilizan para series in situ, por lo que se sugiere ver el reporte de Róvere para más detalles. La estimación de los parámetros por máxima verosimilitud para cada péntada de la climatología se realiza utilizando el paquete `fitdistrplus` (<https://>

[//cran.r-project.org/package=fitdistrplus](https://cran.r-project.org/package=fitdistrplus)) del lenguaje R (Delignette-Muller and Dutang, 2015). Este paquete permite la estimación de los parámetros para una función Gamma – utilizada para el cálculo del SPI – y para una función Log-logística – para el caso del SPEI. La otra variante para el ajuste de la climatología es a través de un método no paramétrico. Al tratarse de una metodología no paramétrica no se obtiene una función matemática explícita. El ajuste en cambio se realiza a través de logsplines, cuyo objetivo es ajustar una función logarítmica a los datos observados a partir de una serie de funciones cúbicas (Perperoglou et al., 2019). La implementación en R fue realizada con el paquete logspline (<https://cran.r-project.org/package=logspline>). Sin embargo, se han realizado modificaciones a la implementación desarrollada para series in situ, de modo de poder manejar más de una serie temporal por cada estación meteorológica. Esta modificación consiste en agrupar a todas las realizaciones en una única tabla para poder ajustar una sola climatología para todas ellas. Esto permite que las realizaciones sean comparables entre sí. De otra manera, cada realización tendría un set de parámetros diferente. Por ejemplo, si se han generado 25 realizaciones para la estación de Junín y se está ajustando la climatología de la primera péntada del año considerando el período de referencia 1971-2010. Se obtendrían 40 valores para esa péntada por realización, y dado que se generaron 25 realizaciones el total de valores para ajustar la climatología sería de 1000 valores. Una muestra de este tamaño permite obtener ajustes muy precisos y robustos. Al finalizar el proceso de ajuste de la climatología se produce una lista de parámetros (en caso de ser un ajuste paramétrico) o un objeto más complejo (en caso de ser un ajuste no paramétrico). Estos resultados se guardan temporalmente en la base de datos local y luego son utilizados para validar los ajustes a las distintas funciones de probabilidad. Si el ajuste fue satisfactorio, los parámetros estimados junto con el resultado de los tests de bondad de ajuste se almacenan en la base de datos. Si las pruebas son rechazadas, es decir, no se encuentra un ajuste apropiado de la distribución a los datos sintéticos, el índice no se puede calcular porque se incurriría en un error que podría generar valores del índice poco confiables.

3.3.2 Pruebas de bondad de ajuste

Las pruebas de bondad de ajuste son tests estadísticos que permiten verificar si una distribución de probabilidades supuesta es congruente con un conjunto de datos dado (en este caso, totales de precipitación calculados a partir de series sintéticas para un lugar, escala temporal deseada y péntada del año). Para que los tests pueden llevarse a cabo es necesaria la definición de una función de probabilidad de referencia, por lo que estas pruebas no pueden ser aplicadas a ajustes no paramétricos. Los tests de bondad de ajuste se dividen en dos grupos: i) probabilísticos; y ii) basados en cuantiles. Entre los tests probabilísticos se encuentran las pruebas de: Kolmogorov-Smirnov, Anderson-Darling y Cramér-von Mises (Wilks, 2011). Los tres tienen como objetivo determinar si los datos observados son consistentes con una distribución determinada. La ejecución de las pruebas devuelve el valor del estadístico correspondiente y un p-valor asociado al valor del estadístico. Debe recordarse que el p-valor de un test de hipótesis, indica la probabilidad de obtener un valor del estadístico igual o más extremo que el observado bajo el supuesto de la hipótesis nula. Si el valor observado del estadístico es muy extremo, entonces resulta muy improbable que bajo el supuesto de

la hipótesis nula puedan ocurrir observaciones como la de la muestra. El p-valor obtenido es comparado con un umbral de la prueba ($\alpha = 0.05$). El ajuste será rechazado cuando al menos un p-valor resultante de los tests sea menor al umbral. La razón de utilizar tres tests radica en que cada uno tiene fortalezas y debilidades, por lo que se complementan muy bien. Los tests basados en cuantiles son complementarios a los anteriores, y sólo se utilizan si los ajustes pasaron las tres pruebas probabilísticas. Los tests han sido desarrollados específicamente para poder comparar las distintas metodologías de ajuste, ya que permiten seleccionar el método que produce los mejores resultados. La primera prueba es una variante de la raíz del error cuadrático medio entre los cuantiles observados y los teóricos. Mientras más pequeño sea el error, mejor será el ajuste (Wilks, 2011). El segundo test es una prueba de correlación basada en la concordancia de los cuantiles (Lawrence and Lin, 1989). La interpretación es idéntica a la de un coeficiente de concordancia ordinario, los valores cercanos a 1 indican la máxima concordancia mientras que valores próximos a 0 indican lo contrario. La última prueba de este tipo consiste en una comparación entre los cuantiles teóricos para una distribución determinada y los cuantiles estimados para los datos observados (Wilcox et al., 2014). Se comparan sólo 10 cuantiles que son considerados representativos de la función de probabilidad; a través de una prueba de Monte-Carlo se determina si los cuantiles teóricos son iguales a los ajustados. Para cada uno de los 10 cuantiles probados se obtiene un p-valor que debe ser superior al α de la prueba (0.05) para ser considerado exitoso. Si el p-valor es inferior a 0.05, se considera que los cuantiles observados y los teóricos son diferentes. Este test tiene como objetivo determinar si los cuantiles extremos de la distribución se encuentran bien ajustados ya que son los que definirán las probabilidades de ocurrencia de valores extremos de los índices. Finalmente, si el ajuste es exitoso, los resultados de los dos tipos de pruebas de bondad son almacenados en la base de datos, al igual que los parámetros estimados. Esto permite la evaluación posterior y comparación de los distintos métodos de ajuste.

4 Ejemplo de aplicación

Como se mencionó anteriormente, luego del cálculo de las estadísticas móviles se procede al cálculo de los índices de sequía. Para ello se deben correr los siguientes scripts respetando el orden `01_generador_configuraciones.R`, `02_calculador_indices.R` y, de manera opcional, `03_diagnosticos.R`. No es necesario modificar nada en el código ya que éste es válido para calcular cualquier tipo y configuración de índice de sequía. Solo se deben modificar los archivos de configuración y parámetros correspondientes a cada uno de los scripts.

4.1 Generación de configuraciones

La generación de configuraciones hace referencia a los índices que desean calcularse por lo que se debe definir la escala de agregación, el período de referencia, función de distribución a utilizar y método de ajuste. Antes de correr el script `01_generador_configuraciones.R` es necesario modificar los archivos `configuracion_generador_configuraciones.yml` y `parametros_generador_configuraciones.yml`. Estos archivos se pueden editar con

cualquier editor de texto, incluso con el propio RStudio. A continuación se muestra la estructura de los archivos de texto y como modificarlos según cada usuario. El primero que modificaremos es `configuracion_generador_configuraciones.yml`. Este archivo define las rutas a las distintas carpetas donde se almacenarán los resultados y la cantidad de máxima de procesos que pueden realizarse en paralelo.

```
# -----  
# Directorios de trabajo  
# -----  
dir:  
  base: /Users/alessiobocco/Documents/Documentos/SISSA/Devel/indices-eventos/IndicesSequ  
  data: /Users/alessiobocco/Documents/Documentos/SISSA/Devel/indices-eventos/data/  
# -----
```

El primer bloque corresponde a las rutas. Se deben definir cuatro rutas:

- **base:** corresponde a la ruta donde se encuentra el proyecto.
- **data:** corresponde a la ruta donde se encuentran los datos diarios de entrada para el cálculo de las estadísticas móviles. Esta carpeta no se encuentra junto al proyecto sino un nivel más arriba dentro de `indices-eventos`.

En el segundo bloque se deben definir la cantidad de procesos en paralelo que pueden llevarse adelante. La mayoría de las computadoras personales hoy poseen más de un núcleo lo que aumenta en gran medida el poder de cómputo. Se recomienda elegir una cantidad de núcleos tal que permita seguir utilizando la computadora con fluidez, en el caso que se trate de una computadora personal. Si se desconoce la cantidad de núcleos puede escribir lo siguiente en la consola.

```
require(parallel)  
parallel::detectCores()
```

```
## [1] 8
```

La función `detectCores()` del paquete `parallel` permite identificar la cantidad de núcleos disponibles. En este caso son 8 pero se recomendaría usar 6 para así dejar dos disponibles para otras tareas.

Cabe mencionar que el usuario debe crear una carpeta

El segundo archivo de texto que debe modificarse es `parametros_generador_configuraciones.yml`. Este archivo define las características que tendrá el proceso de cálculo de índices

```

# -----
# Definicion de configuraciones de indices
# -----
indices:
  SPI:
    escala: [ 1, 3]
    periodo.referencia: [ 1971-2010 ]
    distribucion: [ 'Gamma' ]
    metodo.ajuste: [ 'NoParametrico', 'ML-SinRemuestreo' ] # 'L-Momentos-PP-PWM', 'L-Mo
  SPEI:
    escala: [ 1, 3]
    periodo.referencia: [ 1971-2010 ]
    distribucion: [ 'Log-Logistica' ]
    metodo.ajuste: [ 'NoParametrico', 'ML-SinRemuestreo' ] # 'L-Momentos-PP-PWM', 'L-Mo

# -----

```

En el ejemplo se muestran dos bloques, uno correspondiente al SPI y otro al SPEI. Si sólo se desea calcular uno de ellos basta con borrar las líneas del bloque no deseado. Sin embargo, la opción más segura sería comentarlas añadiendo un `#` al comienzo de cada línea. En cada uno de los bloques se define la escala en meses, que debe coincidir con las estadísticas móviles agregadas. De otra forma no se podrá calcular el índice y arrojará un error ya que los datos no existen. En los métodos de ajuste pueden seleccionarse más de uno si el usuario tiene interés de compararlos. Sin embargo, debe contemplarse que la cantidad de configuraciones pueden crecer rápidamente lo que multiplica el tiempo total que insume el proceso. `##` Cálculo de índices

Luego de haber definido las configuraciones se procede con el cálculo de los índices a partir de las estadísticas móviles. Antes de correr el script `02_calculador_indices.R` es necesario modificar los archivos `configuracion_calculador_indices.ym` y `parametros_calculador_indices.ym`.

El primero que modificaremos es `configuracion_calculador_indices.ym`. Este archivo define las rutas a las distintas carpetas donde se encuentran los archivos de entrada y se almacenarán los resultados y la cantidad de máxima de procesos que pueden realizarse en paralelo.

```

# -----
# Directorios de trabajo
# -----
dir:
  base: /Users/alessiobocco/Documents/Documentos/SISSA/Devel/indices-eventos/IndicesSequ
  run: /Users/alessiobocco/Documents/Documentos/SISSA/Devel/indices-eventos/IndicesSequi
  estadisticas:
    run: /Users/alessiobocco/Documents/Documentos/SISSA/Devel/indices-eventos/Estadistic

```

```

lib: /Users/alessiobocco/Documents/Documentos/SISSA/Devel/indices-eventos/lib/R/
data: /Users/alessiobocco/Documents/Documentos/SISSA/Devel/indices-eventos/data/
# -----

# -----
# Cantidad maxima de procesos paralelos
# -----
max.procesos: 8
# -----

```

- **base:** corresponde a la ruta donde se encuentra el proyecto.
- **run:** corresponde a la ruta donde se guardarán los logs de cada una de las tareas realizadas. Puede crearla manualmente como una subcarpeta dentro del directorio base o solo especificarla y el script la creará automáticamente. Siempre debe ser una subcarpeta dentro de **base**.
- **estadisticas > run:** corresponde a la ruta de ejecución del código para la agregación de las estadísticas móviles.
- **lib:** corresponde a la ruta donde se encuentran los scripts desarrollados por el SISSA. Esta carpeta no se encuentra junto al proyecto sino un nivel más arriba dentro de **indices-eventos**.
- **data:** corresponde a la ruta donde se encuentran los datos diarios de entrada para el cálculo de las índices. Esta carpeta no se encuentra junto al proyecto sino un nivel más arriba dentro de **indices-eventos**.

El segundo que archivo que se debe modificar es `parametros_calculador_indices.yml`.

```

# -----
# Escalas para las cuales serán calculados los índices.
# --- Tener en cuenta que hay una relación entre los anchos de ventana utilizados
# --- al calcular las estadísticas móviles y las escalas para las cuales se podrán
# --- calcular los índices. Para poder calcular índices para una escala x, deben
# --- haberse calculado estadísticas móviles para el ancho de ventana medido en
# --- pentadas igual a x o un ancho de ventana medido en dias igual a x*6 !!!
# -----
escalas: [1, 3]
# -----

# -----
# Maxima proporcion de faltantes para periodo de referencia
# -----
min.proporcion.disponibles.referencia: 0.75
# -----

```

```

# -----
# Definicion de funciones de ajuste y sus parametros generales
# -----
ajuste.general:
  NoParametrico: { funcion: 'AjustarNoParametrico', parametros: { } }
  #ML-ConRemuestreo: { funcion: 'AjustarMaximaVerosimilitud', parametros: { numero.muest
  ML-SinRemuestreo: { funcion: 'AjustarMaximaVerosimilitud', parametros: { min.tasa.valo
# -----

# -----
# Definicion de funciones de ajuste y sus parametros particulares por indice
# -----
ajuste.particular:
  NoParametrico: { SPI: { lbound: 0, min.tasa.valores.positivos: 0.95 } }
# -----

# -----
# Definicion de parametros para tests de bondad de ajuste
# -----
tests:
  umbral.p.valor: 0.05
  umbral.ccc: 0.9
  qcomhd.muestras: 100
  probs:
    ccc: { from: 0.01, to: 0.99, by: 0.01 }
    qcomhd: { from: 0.1, to: 0.9, by: 0.1 }
# -----

# -----
# Definicion de funciones de calculo de indices
# -----
calculo:
  SPI: { funcion: 'CalcularSPI', parametro.lista: 'parametros.gamma', parametro.logsplin
  SPEI: { funcion: 'CalcularSPEI', parametro.lista: 'parametros.log.logistica', paramet
# -----

# -----
# Sistema de referencia espacial de los datos
# -----
projections:
# Proyectado
  gk: '+proj=tmerc +lat_0=-90 +lon_0=-60 +k=1 +x_0=5500000 +y_0=0 +ellps=intl +towgs84=-
# Geográfico
  latlon: '+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0'

```

El segundo grupo de archivos a modificar corresponden a la configuración del cálculo propiamente dicho de los índices de sequía. En esta sección se definen los parámetros que condicionarán el proceso de calculo de índices.

- **escalas:** deve coincidir con las configuraciones seleccionadas anteriormente.
- **min.proporcion.disponibles.referencia:** cantidad mínima de datos en el período de referencia para que una pentada puede ser ajustada a la distribución de probabilidad seleccionadas.
- **ajuste general:** define las funciones que serán utilizadas para el ajuste de cada péntada del periodo de referencia. Se sugiere no modificar esta opción.
- **ajuste particular** define las funciones para el ajuste no paramétrico. Se sugiere no modificar esta opción.
- **test:**
 - **umbral.p.valor:** valores de alpha para cada uno de los test de bondad de ajuste. El estándar para ellos es 5% pero el usuario puede relajarlo si así lo desea.
- **calculo:** nombre de las funciones que ajustan los indices. Esta opción no debe modificarse.
- **projection:** sistema de referencia espacial para los datos de entrada. Por defecto, el generador devuelve un objeto que tiene una proyección planar que debe ser la misma a la indicada aquí.

4.2 Resultados

Una vez completado los archivos de configuración y parámetros, sólo queda abrir el script y ejecutarlo. A modo de ejemplo se mostrarán a continuación algunos resultados. En la subcarpeta **data** dentro del directorio raíz **indices-eventos** se encuentra una carpeta llamada **output**. Dentro de la misma se guardarán los resultados del cálculo de índices. En esa carpeta también se encuentran los resultados intermedios de los parámetros de las funciones de distribución de probabilidad elegida y los resultados de los test de bondad de ajuste.

A continuación se muestran los resultados para la estación de Villa Reynolds, San Luis, Argentina.

```
indices_sequia <- vroom::vroom('/Users/alessiobocco/Documents/Documentos/SISSA/Devel/ind

## Rows: 1,699,200
## Columns: 15
## Delimiter: ","
## chr   [ 3]: indice, distribucion, metodo_ajuste
## dbl   [10]: station_id, realizacion, pentada_fin, ano, metodo_imputacion_id, conf_id,
## date  [ 2]: referencia_comienzo, referencia_fin
##
## Use 'spec()' to retrieve the guessed column specification
## Pass a specification to the 'col_types' argument to quiet this message
```

En la Tabla siguiente se muestran las primeras 75 filas de la tabla con los resultados del proceso de agregación. Se muestran las primeras 75 porque al calcular los índices a una escala de agregación de un mes, el primer se pierde y es reemplazado por NA. Por ello como el dataset usado tiene 50 realizaciones, las primeras 50 filas serán NA. Mientras que a partir de la 51, se muestra la segunda péntada de la primera realización y así sucesivamente. Las variables que posee la Tabla son las siguientes:

- **station_id** : Código unívoco que identifica a la estación meteorológica
- **point_id** : Código unívoco que representa a cada punto de la simulación. Esta variable cobra sentido cuando se simula sobre un grilla o en locaciones arbitrarias donde no hay un identificador previo para el punto. Si los datos a agregar solo corresponden a estaciones meteorológicas, esta columna podría descartarse.
- **pentada_fin** : Valor de la última péntada de la ventana de agregación.
- **ano** : Año de la péntada correspondiente.
- **metodo_imputacion_id** : Código unívoco de identificación del método de imputación.
***conf_id** : Código de identificación de la configuración generada por el primer script. Indica la combinación de índices de sequía, escala de agregación, período de referencia y método de ajuste.
- **indice** : Índice de sequía utilizado. Puede ser SPI o SPEI.
- **escala** : Escala de agregación seleccionada.
- **distribución** : Función de distribución de probabilidad usada para el cálculo. Puede ser Gamma o Log-Logística.
- **método_ajuste** : Tipo de método de ajuste elegido.
- **referencia_comienzo** : Fecha de comienzo del período de referencia.
- **referencia_fin** : Fecha de finalización del período de referencia. ***valor_dato** : Valor de la variable agregada a la escala seleccionada.
- **valor_indice** : Valor del índice seleccionado.
- **percentil_dato** : Percentil correspondiente al **valor_dato** comparado con el período de referencia.

station_id	realizacion	pentada_fin	ano	metodo_imputacion_id	conf_id	indice	escala	distribucion	metodo_ajuste	referencia_comienzo	referencia_fin	valor_datos	valor_indice	percentil_datos
87448	1	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	2	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	3	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	4	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	5	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	6	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	7	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	8	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	9	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	10	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	11	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	12	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	13	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	14	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	15	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	16	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	17	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	18	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	19	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	20	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	21	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	22	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	23	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	24	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	25	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	26	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	27	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	28	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	29	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	30	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	31	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	32	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	33	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	34	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	35	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	36	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	37	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	38	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	39	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	40	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	41	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	42	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	43	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	44	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	45	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	46	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	47	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	48	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	49	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	50	1	1961	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	NA	NA	NA
87448	1	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	45.97134	-1.2399500	10.749695
87448	2	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	41.64498	-1.3687553	8.553789
87448	3	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	83.86446	-0.2904626	38.573118
87448	4	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	100.88314	0.0367147	51.464377
87448	5	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	109.17459	0.1758507	56.979438
87448	6	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	60.34732	-0.8427620	19.968078
87448	7	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	194.92667	1.1679458	87.858569
87448	8	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	172.44004	0.9568559	83.068001
87448	9	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	54.45310	-1.0000264	15.864886
87448	10	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	92.65052	-0.1142641	45.451422
87448	11	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	132.31832	0.5071552	69.397705
87448	12	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	109.31606	0.1781188	57.068516
87448	13	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	92.81652	-0.1110873	45.577357
87448	14	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	74.53060	-0.4954132	31.015425
87448	15	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	46.62861	-1.2207697	11.108661
87448	16	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	48.01200	-1.1807283	11.885535
87448	17	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	37.89711	-1.4840674	6.889552
87448	18	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	93.43698	-0.0992627	46.046486
87448	19	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	150.66914	0.7273964	76.650843
87448	20	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	54.80676	-0.9903741	16.099566
87448	21	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	87.13012	-0.2231073	41.172600
87448	22	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	113.24083	0.2397021	59.471941
87448	23	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	122.37727	0.3737817	64.571663
87448	24	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	77.34256	-0.4317124	33.297524
87448	25	1	1962	0	1	SP1	1	Gamma	NoParametrico	1971-01-01	2010-12-31	116.36739	0.2869821	61.293700

La tabla resultante se encuentra en formato largo, es decir, cada realizacion/configuración de los índices se encuentra una debajo de la otra. Para visualizar los resultados se transforma la tabla de formato largo a ancho. Se filtrará solo una realización para los índices SPI y SPEI agregados a una escala de 1 mes.

Filtrar valor de los índices SPI y SPEI para las escalas de 1 mes:

```
indices_sequia_escala_uno <- indices_sequia %>%
  dplyr::filter(conf_id %in% c(1, 5), realizacion == 1) %>%
```

```

dplyr::select(station_id, realizacion, ano, pentada_fin,
              indice, valor_indice) %>%
tidyr::spread(indice, valor_indice) %>%
dplyr::mutate(date = pentada.ano.a.fecha.inicio(pentada_fin, ano))

# Estructura de la tabla transformada
indices_sequia_escalado %>%
  dplyr::slice(1:10) %>%
  knitr::kable() %>%
  kableExtra::kable_styling() %>%
  kableExtra::scroll_box(width = "100%", height = "200px")

```

station_id	realizacion	ano	pentada_fin	SPEI	SPI	date
87448	1	1961	1	NA	NA	1961-01-01
87448	1	1961	2	NA	NA	1961-01-06
87448	1	1961	3	NA	NA	1961-01-11
87448	1	1961	4	NA	NA	1961-01-16
87448	1	1961	5	NA	NA	1961-01-21
87448	1	1961	6	-0.3451669	-0.1901227	1961-01-26
87448	1	1961	7	-0.2626959	-0.0929402	1961-02-01
87448	1	1961	8	0.1529654	0.1390311	1961-02-06
87448	1	1961	9	-0.6694804	-0.8846628	1961-02-11
87448	1	1961	10	-0.7091847	-1.2694952	1961-02-16

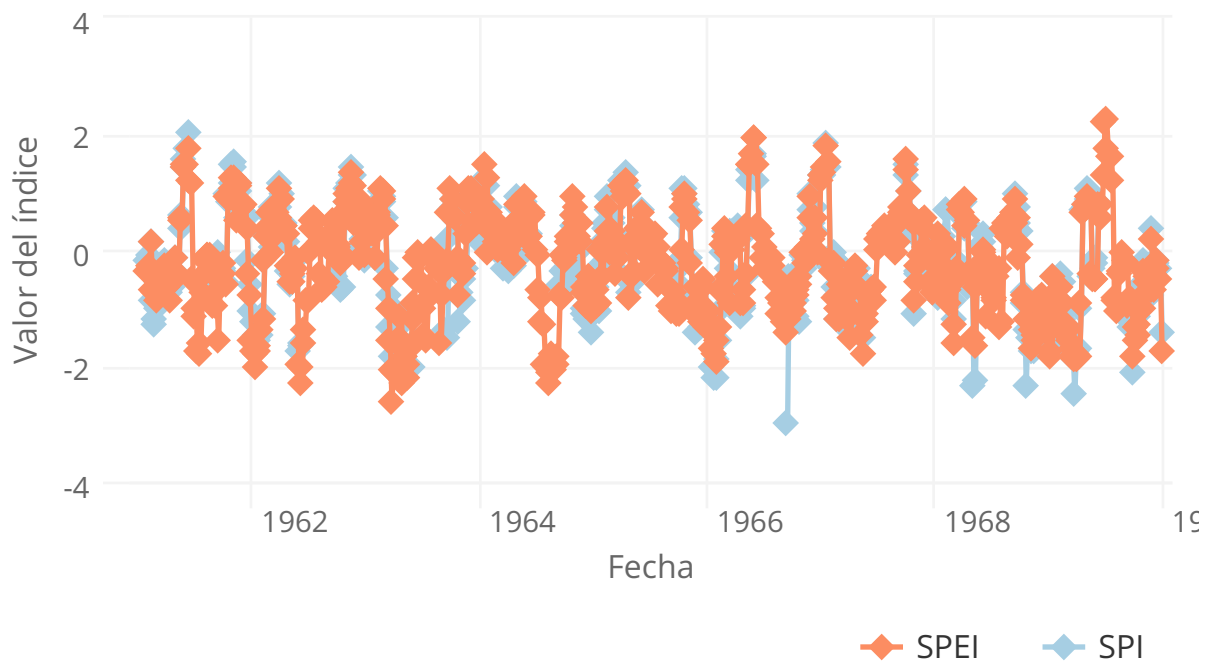
Para visualizar los efectos de la agregación en alta frecuencia se muestra un fragmento de 10 años de la tabla anterior. Como se mencionó, este dataset posee 50 realizaciones pero aquí solo se filtra una para optimizar la visualización. Las líneas naranjas de la Figura corresponden a la serie temporal de SPEI mientras que la celeste a la de SPI.

```

## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo

```

Valores de SPE y SPI de Villa Reynolds (San Luis, Argentina)



Bibliografía

- Beguería, Santiago, Sergio M. Vicente-Serrano, Fergus Reig, and Borja Latorre. 2013. "Standardized Precipitation Evapotranspiration Index (Spei) Revisited: Parameter Fitting, Evapotranspiration Models, Tools, Datasets and Drought Monitoring." Journal Article. *International Journal of Climatology* 34: 3001–23.
- Hargreaves, GH, and ZA Samani. 1982. "Estimating Potential Evapotranspiration." Journal Article. *Journal of the Irrigation and Drainage*.
- Hayes, Michael, Mark Svoboda, Nicole Wall, and Melissa Widhalm. 2011. "The Lincoln Declaration on Drought Indices: Universal Meteorological Drought Index Recommended." Journal Article. *Bulletin of the American Meteorological Society* 92 (4): 485–88. <https://doi.org/10.1175/2010bams3103.1>.
- McKee, Thomas B, Nolan J Doesken, and John Kleist. 1993. "The Relationship of Drought Frequency and Duration to Time Scales." Journal Article. *AMS 8th Conference on Applied Climatology*, 179–84. <https://doi.org/citeulike-article-id:10490403>.
- Vicente-Serrano, Sergio, Santiago Beguería, and Juan López-Moreno. 2010. "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index." Journal Article. *Journal of Climate* 23: 1696–1718.

Vicente-Serrano, S. M., O. Chura, J. I. López-Moreno, C. Azorin-Molina, A. Sanchez-Lorenzo, E. Aguilar, E. Moran-Tejeda, F. Trujillo, R. Martínez, and J. J. Nieto. 2015. “Spatio-Temporal Variability of Droughts in Bolivia: 1955–2012.” Journal Article. *International Journal of Climatology* 35 (10): 3024–40. <https://doi.org/10.1002/joc.4190>.