

Universidad Nacional de Córdoba
Facultad de Ciencias Exactas Físicas y Naturales



Practica Supervisada. Ingeniería Civil

Informe técnico final

Título: Elaboración del programa SOS

Año: 2013

Alumno: Guillermo Marraco

Matrícula: 199120327

Plan: 2005

Tutor: Gerardo Hillman

Supervisor Externo: Andrés Ravelo

Expediente interno N°:

Índice

Contenido

Índice.....	2
Tabla de ilustraciones.....	3
Introducción.....	4
Objetivos y criterios de la metodología usada.....	4
Desarrollo, cálculo y diagramas.....	5
Datos.....	5
Formato Excel de archivo de datos de precipitación.....	5
Ejecución.....	5
Salida.....	5
Cálculo y diagramas.....	5
Resultados obtenidos en la práctica profesional	7
Conclusiones.....	8
Bibliografía.....	9
Informes mensuales del avance de la Práctica Supervisada.....	10
Informe N°1.....	10
Informe N°2.....	10
Anexos.....	11
Manual del programa SOS.....	11

Tabla de ilustraciones

Introducción

La práctica supervisada consistió en la creación de software para automatizar el procesamiento de datos hidrológicos, para calcular índices en base a la precipitación medida.

El servicio meteorológico provee al CREAN de datos de precipitación diarios para numerosas estaciones meteorológicas argentinas.

Los datos son recibidos en formato de texto, y el resultado buscado son diversos índices hidrológicos, y mapas que los representen gráficamente.

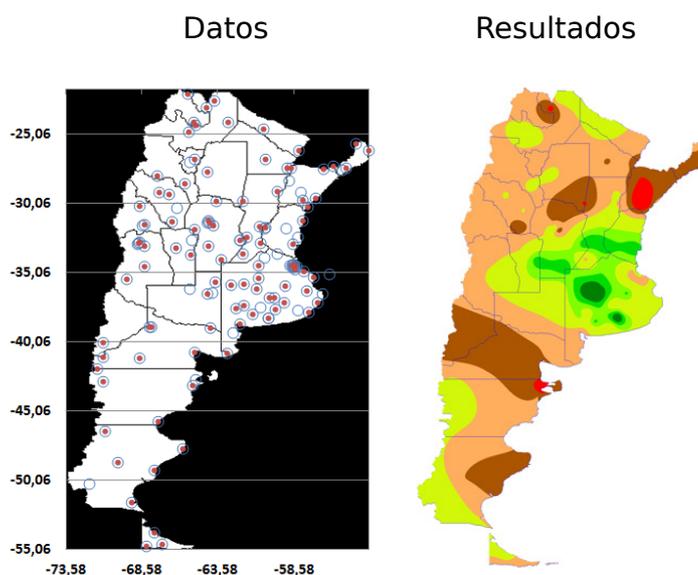


Figura : Estaciones meteorológicas y mapa de índices

Objetivos y criterios de la metodología usada.

El **objetivo** es automatizar los pasos, que previamente eran ejecutados manualmente, para el procesamiento de datos de precipitación obtenidos del Servicio Meteorológico Nacional, con el fin de elaborar diversos índices de sequía. El trabajo consistió en programar el software necesario para esa tarea.

Se requiere cierta flexibilidad, para seleccionar y modificar las estaciones, y las zonas geográficas en las que se elaboren los mapas.

El **criterio** es obtener resultados que sean indistinguibles del procedimiento manual anterior, y alcanzar esos resultados rápidamente, aprovechando el poder de cómputo de una PC de escritorio.

Desarrollo, cálculo y diagramas

El programa fue desarrollado en Visual Basic .NET, un lenguaje completamente orientado a objetos, que generalmente comparte todas las características de Visual C++ .NET (excepto por la múltiple herencia), y es traducible recíprocamente.

El lenguaje fue adoptado por la facilidad con que interacciona con otros programas, en particular MS Excel e IDRISI.

Sin embargo, no requiere de IDRISI para su funcionamiento, aunque éste programa puede ser un útil auxiliar, debido a que varios de los formatos de archivos usados son compatibles.

En particular, el programa usa como dato de entrada un mapa de máscara en formato Raster IDRISI, y puede entregar sus resultados en el mismo formato.

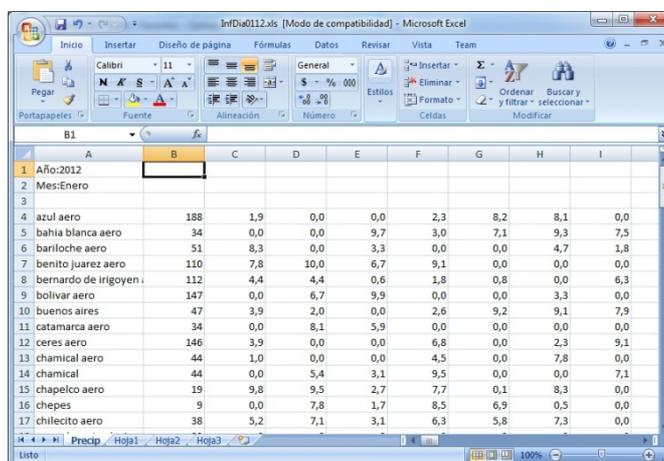
Se requiere Excel si se usan sus formatos de planillas como datos de entrada, lo cual es recomendado.

Datos

Los datos se recibieron en varios formatos de texto, que fueron cambiando a medida que progresó el trabajo, por lo que se decidió estandarizar el formato de entrada, adoptando el de una planilla Excel.

De esta manera se facilita la preparación de los datos, y se garantiza cierta compatibilidad futura.

El programa admite ingreso mediante algunos otros formatos de datos, que se detallan en el manual.



	A	B	C	D	E	F	G	H	I
1	Año:2012								
2	Mes:Enero								
3									
4	azul aero	188	1,9	0,0	0,0	2,3	8,2	8,1	0,0
5	bahia blanca aero	34	0,0	0,0	9,7	3,0	7,1	9,3	7,5
6	bariloche aero	51	8,3	0,0	3,3	0,0	0,0	4,7	1,8
7	benito juarez aero	110	7,8	10,0	6,7	9,1	0,0	0,0	0,0
8	bernardo de irigoyen	112	4,4	4,4	0,6	1,8	0,8	0,0	6,3
9	bolivar aero	147	0,0	6,7	9,9	0,0	0,0	3,3	0,0
10	buenos aires	47	3,9	2,0	0,0	2,6	9,2	9,1	7,9
11	catamarca aero	34	0,0	8,1	5,9	0,0	0,0	0,0	0,0
12	ceres aero	146	3,9	0,0	0,0	6,8	0,0	2,3	9,1
13	chemical aero	44	1,0	0,0	0,0	4,5	0,0	7,8	0,0
14	chemical	44	0,0	5,4	3,1	9,5	0,0	0,0	7,1
15	chapelco aero	19	9,8	9,5	2,7	7,7	0,1	8,3	0,0
16	chepes	9	0,0	7,8	1,7	8,5	6,9	0,5	0,0
17	chilecto aero	38	5,2	7,1	3,1	6,3	5,8	7,3	0,0

Figura : Planilla de datos

Formato Excel de archivo de datos de precipitación

Los datos de precipitación se leen desde una planilla en formato Excel, y se toman de la hoja titulada “precip”, o, de no existir, de la primera hoja con el texto “precip”.

Por ejemplo, si la primer hoja se titula “Datos de Precipitación”, y la segunda se titula “Precip”, se toman los datos de la segunda hoja. Pero si la segunda se titula “Precipitación”, se toman los datos de la primer hoja, ya que es la primera cuyo nombre contiene el texto “precip”.

El uso de mayúsculas o minúsculas es indistinto.

Cómo indicar la fecha a la que corresponden los datos

Existen dos métodos; indicarlo en el nombre de archivo, o en el interior de la planilla.

- Si la planilla tiene como nombre de archivo el siguiente formato:
infdiaMMAA.xls

En donde dice **MM** se indica el mes al que corresponde la planilla (dos caracteres numéricos), y en lugar de **AA**, (dos caracteres numéricos), se indica el año al que corresponde la planilla. Deben utilizarse siempre dos cifras. Si el número es menor que diez, debe anteponersele un cero.

Por ejemplo **infdia0108.xls**, es el nombre que deben tener los datos de enero de 2008 (01/08).

Dado que sólo se consideran dos cifras para indicar el año, sólo se admiten planillas que contengan datos entre los años 1931 y 2030.
- El segundo método admite cualquier nombre de archivo, pero la planilla con los datos de precipitación debe contener dos celdas con los siguientes textos:

Texto: “**Año:XXXX**”. Debe iniciar con la palabra “**año**”. En lugar de **XXXX** debe indicarse el año al que corresponden los datos, utilizando hasta 4 cifras numéricas. Ejemplo: una celda conteniendo el texto “**año:1912**”, o “**Año, 12**” (que se interpreta como año 12, y no 2012).

Texto: “**Mes:YYYY**”. En Lugar de **YYYY** debe indicarse el mes al que corresponden los datos. Se pueden utilizar una o dos cifras numéricas, o bien el nombre del mes.

Cómo indicar los datos de precipitación

Se reconocen como datos de precipitación, todas las filas, en las que la primera columna contenga el nombre de una estación, y las siguientes columnas contengan datos de lluvia, en mm.

Debe haber exactamente un dato por cada día. Si no se registraron datos, puede utilizarse los textos “-”, (signo de negación), o “INAP” (inapreciable); éstos valores serán considerados de precipitación nula. Si alguna de las celdas no corresponde a un número, o alguno de los textos opcionales, se descartará toda la fila de datos.

La fila debe contener un dato por cada día de ese mes específico; 28 días para febrero (29 en años bisiestos), ó 31 días para Diciembre.

No se puede dejar una celda en blanco dentro de las que corresponden a ese mes.

Debe prestarse especial atención al nombre de la estación. Éste debe ser siempre el mismo nombre en las diferentes planillas. No se admiten abreviaciones, ni nombres repetidos en diferentes estaciones. En caso de conflicto, se recomienda añadir, al final del nombre, el número internacional de la estación. El uso de mayúsculas o minúsculas es indistinto.

El programa requiere otros archivos de entrada, que se detallan en el manual.

Ejecución

Cuando el programa lee la planilla de datos, verifica la corrección del formato, y realiza algunos chequeos, como la presencia de valores negativos, no numéricos, o la falta de datos.

Antes de procesar los datos, se pueden revisar las estaciones a procesar, y los datos que les corresponden, como su ubicación, (latitud, y longitud), si se procesarán, y cuáles archivos mantienen las bases de datos de cada estación, para cada índice.

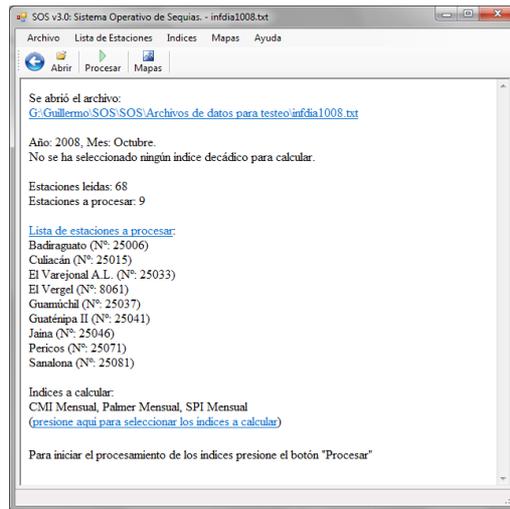


Figura : Apertura de datos

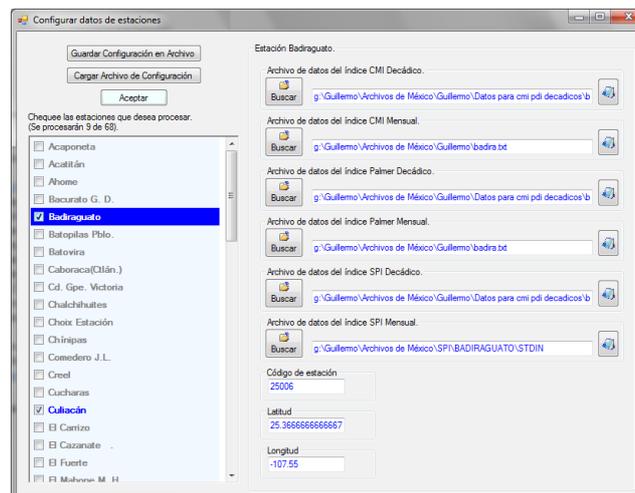


Figura : Configuración de estaciones

También se pueden elegir los índices a calcular, que son 6: CMI (mensual y decádico), Palmer (mensual y decádico), y SPI (mensual y decádico).

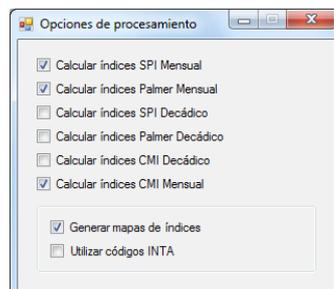


Figura : Selección de índices

El procesamiento de los datos inicia al presionar el botón “Procesar”, que posee un ícono triangular verde.

Salida

Si el programa no encuentra errores durante la ejecución, (en cuyo caso describe esos errores al usuario, junto a una sugerencia para corregirlos, si es posible), lista como resultado, los índices calculados, y grafica los mapas en pantalla.

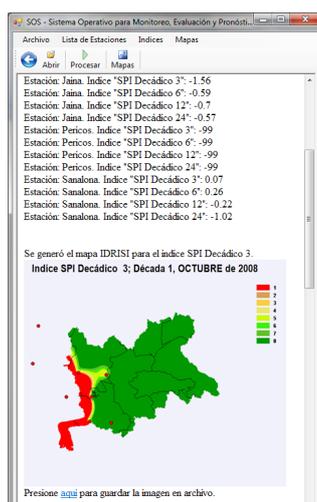


Figura : Resultados

El formato en que se muestran los resultados, es un formato de hipertexto, con vínculos. El usuario puede guardar los mapas generados usando los vínculos debajo de cada mapa, y optar por un formato gráfico (BMP), o formato Raster IDRISI.

(El formato Raster IDRISI se guarda sin reclasificar ni enmascarar).

Los resultados numéricos quedan almacenados en formato de texto, para cada estación, en archivos a los que se puede acceder utilizando los hipervínculos.

En general, los resultados numéricos sólo se usan para investigar y solucionar problemas imprevistos, o en casos especiales, en los que se necesiten sus valores.

La necesidad principal del CREAN era la obtención de los mapas, y con ese objetivo está pensada la interface.

Cálculo y diagramas

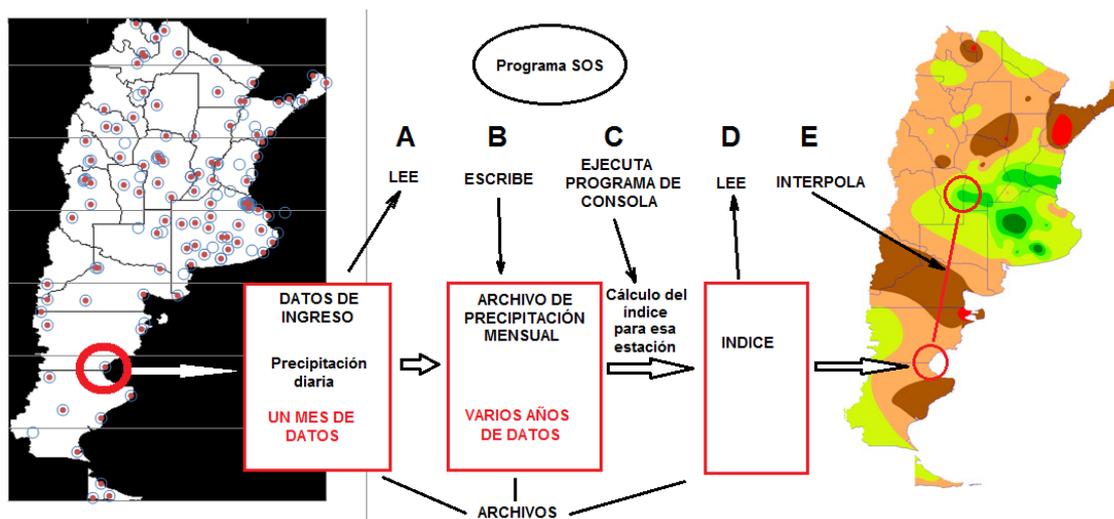


Figura : Esquema del procedimiento

Procedimiento principal

- A. De los archivos de datos, el software lee datos diarios de precipitación, correspondientes a un mes completo, de todas las estaciones meteorológicas.

Para cada estación, el primer cálculo es la suma de la precipitación acumulada en períodos de 10 días y/o mensuales, dependiendo del índice al calcular.

- B. Esa suma se guarda en un archivo de texto, de la misma estación.
- C. El archivo anterior se usa como entrada de datos de un programa de consola, que se ejecuta, y da como salida el valor de los índices.
- D. Se lee el índice resultante para esa estación, y se repite todo el procedimiento con todas las estaciones.

- E. Una vez que se obtienen los índices de todas las estaciones se los interpola espacialmente, para calcular su valor en todo el mapa.

El procedimiento de interpolación, en base al criterio de que deben obtenerse exactamente los mismos resultados que con el procedimiento manual, es el método de interpolación Kriging, (también conocido como Kolmogorov Wiener), con variograma lineal. (Krige, 1951)

Interpolación. Pasos de cálculo

El primer paso es la estimación de una función llamada variograma. Este paso fue omitido, como se explicó, y se adoptó un modelo de variograma lineal, o sea $S_{(d)}=d$, en donde d es la distancia correspondiente entre el pixel a calcular, y las diferentes estaciones.

El variograma representa la varianza S , de la diferencia entre los valores de los índices, pertenecientes a cada par de estaciones, en función de la separación espacial "d" de esas estaciones.

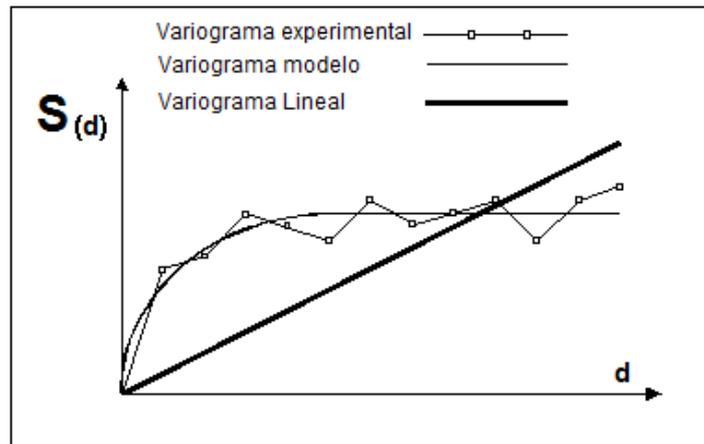


Figura : Variograma

El valor F de la interpolación en un punto (x,y) del mapa se obtiene como un promedio ponderado de los n puntos conocidos, en cada uno de los cuales hay un valor f_i a interpolar. Los pesos w_i son función de la distancia, afectada por el variograma.

$$F_{(x,y)} = \sum_{i=1}^n w_i f_i$$

Ecuación

Los pesos w_i son la solución del sistema de ecuaciones siguiente:

$$w_1 \cdot S(d_{11}) + w_2 \cdot S(d_{12}) + \dots + w_n \cdot S(d_{1n}) = S(d_{1p})$$

$$w_1 \cdot S(d_{21}) + w_2 \cdot S(d_{22}) + \dots + w_n \cdot S(d_{2n}) = S(d_{2p})$$

:

$$w_1 \cdot S(d_{n1}) + w_2 \cdot S(d_{n2}) + \dots + w_n \cdot S(d_{nn}) = S(d_{np})$$

$$w_1 + w_2 + \dots + w_n = 1$$

Ecuación

En donde $S(d_{ij})$ es el valor del variograma a la distancia d_{ij} entre los puntos i y j .

d_{ip} es la distancia correspondiente entre el punto i y pixel p .

La última ecuación es la que garantiza que los pesos sumen 1.

La Ecuación puede resumirse así:

$$W.$$

$$[S]=S_p$$

Ecuación

En donde [S] es una matriz constante, invertible, Sp un vector a calcular para cada pixel, y W el vector incógnita para cada pixel.

Como [S] es constante, puede invertirse una sola vez y ser reutilizado calculando los pesos W como:

$$W=S_p.$$

$$[S]^{-1}$$

Ecuación

Debido a que este paso debe repetirse para cada pixel del mapa resultante, y aún mapas pequeños tienen millones de píxeles, éste es el paso más costoso de todo el proceso, y mucho esfuerzo fue dedicado a optimizarlo, pero no se llegó a modificarlo para provechar el paralelismo de los procesadores por falta de tiempo.

Índices SPI, Palmer, y CMI

El índice SPI es la desviación estándar, según una distribución gamma, de la precipitación acumulada en 3, 6, 12 ó 24 meses. (McKee, Doesken, & Kleist, Enero 1993).

El índice Palmer es similar.

Utilizando un modelo estandarizado, que considera la evapotranspiración potencial, la capacidad de campo, y diferentes capas de suelo, estima un balance de agua, y usa el resultado para calcular un índice de sequía. (Palmer, 1965).

El índice CMI (Crop Moisture Index, o Índice de Humedad de Cultivo), es derivado del Índice Palmer. (Palmer, 1968)

Inicialmente no fue requerido programar el cálculo de los índices, ya que el CREAN proveyó programas de consola que lo hacían, por lo que el cálculo, desde el punto de vista del alumno, fue tratado como una caja negra, en la que se proveían los archivos de texto, se ejecutaban los programas para cada estación, y se leían los resultados en los archivos de salida.

Sin embargo, ello no fue posible con el índice SPI, ya que, por ser un programa de consola creado con un compilador antiguo, presentaba limitaciones, incluyendo el error del año 2000, y límites de memoria que no le permitían procesar largas series de datos. Asimismo fallaba frecuentemente cuando las series de datos presentaban lagunas temporales, aún cuando ello no fuera una dificultad lógica para el cálculo del índice.

Por esa razón, creé un programa de consola independiente para calcular el índice SPI.

Cálculo del índice SPI

El ingreso de datos consiste en un archivo de texto, conteniendo las series de precipitación acumuladas mensualmente.

Pasos de cálculo:

1. Calcular la precipitación acumulada, para un tiempo t , y un período n de 3, 6, 12 ó 24 meses. Se obtienen hasta doce valores A_t , uno para cada mes.

$$n=3, 6, 12, 24$$

$$A_{t_n} = \sum_{k=0}^n P_{t-k}$$

2. Para cada período n :

- a. Se toman todos los valores de precipitación acumulada A_{t_n} en el mes al que corresponde t . Para cada período n , y mes t , se calcula el índice $SPI_{n,t}$, del siguiente modo:

- i. Calcular $C_t = (\text{cantidad de } A_{t_n} = 0)$. Valores nulos.

- ii. $D_t = (N^\circ \text{ de } A_{t_n} \neq 0)$. Valores no nulos.

- iii. $N_t = C_t + D_t$. Cantidad de valores.

- iv. $\hat{x}_t = \text{promedio de los valores } A_{t_n} \text{ no nulos.}$

- v. Frecuencia de ceros $q_t = C_t/N_t$

- vi. Parámetro $A_t = \ln(\hat{x}_t) - \text{promedio}(\ln(\hat{x}_t))$

- vii.
$$\alpha_t = \frac{1 + \sqrt{1 + \frac{4}{3} A_t}}{4 A_t}$$

- viii.
$$\beta_t = \frac{\alpha}{\hat{x}_t}$$

- ix. $\gamma_t = \text{Probabilidad de distribución gama} = f(\hat{x}_t, \alpha_t, \beta_t)$

El cálculo de $f(\hat{x}_t, \alpha_t, \beta_t)$ es complejo, y propenso a errores, por lo que se utilizó una solución robusta, de la librería Math.NET Iridium. (Rüegg, Vermorel, & Kitchin, 2008)

La corrección de los cálculos fue verificada numéricamente, manualmente, utilizando planillas Excel en donde

$$f(\hat{x}_t, \alpha_t, \beta_t) = \text{Distr. Gamma}(\hat{x}_t, \alpha_t, \beta_t, \text{verdadero})$$

La planilla está disponible para el cálculo de $SPI_{3\text{meses}}$.

- La ecuación para la función de densidad

de probabilidad gamma es:

$$f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

La función de densidad de probabilidad gamma estándar es:

$$f(x; \alpha) = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)}$$

- Cuando el argumento $\alpha_t = 1$,

$f(x_t, \alpha_t, \beta_t)$, devuelve la distribución

exponencial con:

$$\lambda = \frac{1}{\beta}$$

- Para un entero positivo n , cuando los argumentos $\alpha_t = n/2$, $\beta_t = 2$, la función

$f(x_t, \alpha_t, \beta_t)$ devuelve (1 -

DISTR.CHI(x)) con n grados de libertad.

x. $H_t = q_t + (1 - q_t) \gamma_t$

1. Si $0 < H_t \leq 0,5$

$$t_t = \sqrt{\ln(H_t^{-2})}$$

$$Z_t = -(t_t - (cc_0 + cc_1 \cdot t_t + cc_2 \cdot t_t^2)) / (1 + d_1 \cdot t_t + d_2 \cdot t_t^2 + d_3 \cdot t_t^3)$$

2. Si $0,5 < H_t \leq 1$

$$\frac{1 - H_t}{(\ln t)^{-2}}$$

$$t_t = \frac{\ln \frac{1 - H_t}{(\ln t)^{-2}}}{\sqrt{t}}$$

$$Z_t = (t_t - (cc_0 + cc_1 \cdot t_t + cc_2 \cdot t_t^2)) / (1 + d_1 \cdot t_t + d_2 \cdot t_t^2 + d_3 \cdot t_t^3)$$

xi. Valor de índice $SPI_{n,t} = Z_t$

Resultados obtenidos en la práctica profesional

El estudiante ganó experiencia aplicando conocimientos de hidrología aplicados a datos reales, así como aprendió que hay importantes diferencias, entre el empleo de herramientas informáticas para uso propio, o la creación de herramientas para terceros.

Los mapas producidos fueron comparados con los obtenidos con el procedimiento anterior, resultando idénticos para las mismas fechas, al usar los mismos datos de entrada.

Originalmente sólo se elaboraban los mapas para la región pampeana, y se tardaban varios días en hacerlos.

Con el nuevo programa se pueden calcular los mapas para todo el territorio nacional, usando más estaciones, y con mayor agilidad. Eso permite concentrar los esfuerzos en el análisis de los resultados, y la detección de anomalías en vez de la confección de los mapas.

En un mismo día fue posible elaborar todos los mapas de índices SPI y Palmer correspondientes a una década, totalizando unos 600 mapas.

Los resultados son publicados mensualmente en una página web de la institución. (CREAN)

CREAN | Centro de Relevamiento y Evaluación de Recursos Agrícolas y Naturales

Google™ Búsqueda personalizada

PORTADA

INSTITUCIONAL

PERSONAL

INVESTIGACION

MONITOREO DE SEQUIAS

RECURSOS IDRISI

PUBLICACIONES

FORMACIÓN

CONTACTO

 UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE CIENCIAS AGROPECUARIAS

Indice de Palmer (PDI)

Setiembre de 2013

Enero	Febrero	Marzo	Abril
Mayo	Junio	Julio	Agosto
Setiembre	Octubre	Noviembre	Diciembre

The figure displays a series of 12 maps of Argentina, one for each month of the year. The maps use a color scale to represent the Palmer Index (PDI) values. The largest map shows the current state for September 2013, with significant red and brown areas in the northern and central regions, indicating low PDI values (drier conditions). The smaller maps show the seasonal progression of these values throughout the year.

Figura : Publicación de los mapas. CREAN, UNC.

Conclusiones

Se obtuvo una experiencia laboral valiosa.

En el aspecto técnico se adquirió experiencia en la manipulación de datos de precipitación reales.

Los datos oficiales frecuentemente necesitan control personal, y no pueden ser dejados en manos del software.

Es normal que en una serie de datos haya valores faltantes y varios tipos de errores e incorrecciones.

En ciertos casos es importante usar series de datos de igual duración, aunque sea necesario descartar datos excedentes. Por ejemplo, no se pueden mezclar datos dos grupos de estaciones cercanas si un grupo posee una serie más larga. Por caso, si uno solo de los grupos registró un período de sequía, los valores medios de ambas series serán diferentes aunque las series estén correlacionadas, y eso distorsiona los cálculos que derivan de ambas.

No puede asumirse automáticamente que si una estación posee una serie más larga de datos, incluirlos mejorará la precisión de los resultados.

Hay una escasez importante de estaciones meteorológicas en el norte argentino, así como en la Patagonia.

La programación para su uso por otras personas, consume la mayor parte del tiempo empleado en elaborar el código, y obliga a hacerlo con extremo cuidado y atención al detalle, no obstante lo cual siempre aparecen problemas imprevistos.

Nunca debe asumirse que el proyecto avanzará según lo planeado, y deben tomarse medidas redundantes para anticiparse a eventos fortuitos. Las medidas redundantes también pueden originar nuevos problemas.

Bibliografía

1. CREAN. (s.f.). *Monitoreo de sequías*. Obtenido de http://www.crean.unc.edu.ar/secciones/monitoreo/monitoreo_intro.html
2. Krige, D. G. (1951). *A statistical approach to some mine valuations and allied problems*. University of Witwatersrand: Master's thesis .
3. McKee, T., Doesken, N., & Kleist, J. (Enero 1993). The relationship of drought frequency and duration to time scale. *Proceedings of the Eighth Conference on Applied Climatology. Boston, American Meteorological Society* , 179-184.
4. Palmer, W. C. (1965). "*Meteorological Drought*". *Research paper N° 45, U.S. Department of Commerce Weather Bureau, February 1965 (58 pgs)*. Obtenido de NOAA National Climatic Data Center: <http://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf>
5. Palmer, W. C. (1968). *Keeping track of crop moisture conditions, nationwide: the new Crop Moisture Index, Weatherwise, 21* , 156-161.
6. Rüegg, C., Vermorel, J., & Kitchin, M. (Abril de 2008). *Math.NET Iridium, v2008.4.14.425*. Obtenido de <http://mathnetiridium.codeplex.com/releases/view/13948>

Informes mensuales del avance de la Práctica Supervisada

Informe N°1

Tareas realizadas:

- Me explicaron el procedimiento usado para hacer los mapas, y me entregaron varios archivos de ejemplo, de datos, y auxiliares, para que me familiarice con ellos.
- Los datos de precipitación corresponden a 25 estaciones meteorológicas de la zona pampeana. Son datos diarios de precipitación.
- Repetí los procedimientos para generar los mapas, y busqué la información necesaria para implementar los algoritmos matemáticos.
- Programé la estructura principal del software, logrando producir resultados comparables a los buscados para el índice Palmer.

Informe N°2

Tareas realizadas:

- Finalicé el código necesario para calcular el índice Palmer y generar los mapas.
- Adapté ese código para procesar los índices SPI y CMI.
- Comparé los mapas elaborados con los originales, y coinciden cuando se usan los mismos datos. Difieren ligeramente cuando se usan series más largas de datos, como es esperable.
- Enseñé a una de las personas del instituto CREAN a operar el programa, y encontramos varias correcciones a realizar.
- Redacté el manual de operación.
- Corregí varios errores en el código, e implementé algunas mejoras.

Anexos

Manual del programa SOS