

1. INTRODUCCIÓN

Frecuentemente el tratamiento de las series de temperatura y precipitación pasa por alto algunos supuestos de los análisis estadísticos tradicionales. Si la técnica de análisis viola estos supuestos, la validez de los resultados se verá afectada. En otras palabras, si la información es asimétrica, heterogénea y/o con tendencia, será indispensable usar estadísticos y técnicas distintas a las tradicionales. Hoffman et al. (1997) demostraron la influencia de causas artificiales en la marcha de las lluvias y temperaturas de varias estaciones argentinas. Resta conocer en que medida las series disponibles reúnen requisitos de simetría, integridad, homogeneidad y ausencia de tendencia.

Los objetivos de este estudio son:

1. Caracterizar los supuestos de normalidad, homogeneidad y falta de tendencia de las series mensuales de temperatura y lluvias.
2. Identificar la presencia de valores ajenos al conjunto de datos ("outliers") en las series.
3. Sintetizar los atributos de calidad de las series con un índice de utilidad para el control e interpretación de los datos.

2. MATERIAL Y MÉTODOS

Se obtuvieron series mensuales de temperaturas medias y precipitaciones de 44 estaciones del Servicio Meteorológico Nacional (S.M.N.) y del Instituto Nacional de Tecnología Agropecuaria (I.N.T.A.) de la región pampeana (Figura 1). El período de análisis fue 1961-1990, excepto en 8 estaciones con registros entre 1971 y 2000.

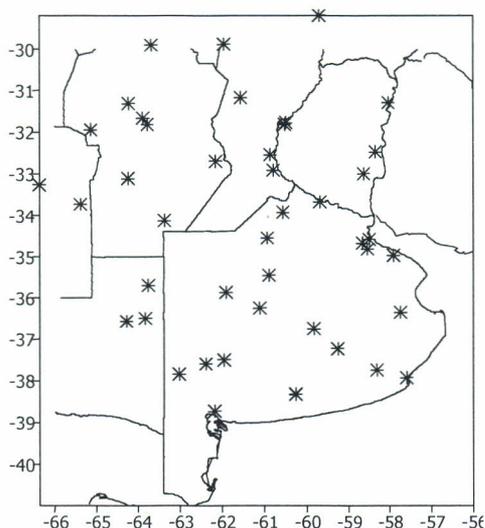


Figura 1. Ubicación de las estaciones del estudio

Grado de simetría: se empleó una modificación de la prueba de Wilcoxon-Mann-Whitney, que determina si las medianas de dos conjuntos son iguales. La mediana *M* separa la serie en dos: *XG* con los valores mayores que *M* y *XL* con los datos menores a *M*. Un tercer grupo *Y* se obtiene como:

Instituto de Clima y Agua, I.N.T.A., Castelar, Argentina

$$Y = M + (M - XL)$$

Un test *z* al 95% de confianza decide la normalidad o asimetría de la distribución.

Grado de homogeneidad: supone que una discontinuidad provocará un cambio significativo en la mediana de una serie de *n* valores (Siegel y Castellan, 1988). Con el valor de orden *R* de cada valor, se calculan *SRi* y *SAi*:

$$SRi = \sum_{j=1}^i R_j$$

$$SAi = |(2SRi) - i(n+1)|$$

El máximo de *SAi* identifica el punto *n1* donde la discontinuidad es máxima. El valor de *SRi* en *n1* sirve para calcular el estadístico *z*, cuya significancia al 95% decidirá si la serie es homogénea o no.

Grado de asociación: El coeficiente de correlación de Pearson es válido siempre y cuando la serie no tenga "outliers" y la distribución sea normal. El coeficiente de correlación de Spearman es un método no paramétrico que evita estos problemas porque las dos variables se ordenan y la correlación se calcula entre los valores de posición (Smith, 1995). El coeficiente *rs* de Spearman es:

$$rs = 1 - [(6 SRD) / (n^2 - n)]$$

donde *SRD* se obtiene igual que *SRi*.

El valor del estadístico *t* se calcula como:

$$t = rs[(n - 2) / (1 - rs^2)]^{0.5}$$

Se rechaza la hipótesis de tendencia en la distribución si el valor absoluto de *t* > *tstat*.

Grado de consistencia interna: se identificaron los valores extraños al conjunto de datos ("outliers") empleando un intervalo de confianza que depende de la media y el desvío estandar si la serie es simétrica o, caso contrario, de la media *Xbi* y el desvío bponderado *Sbi*. *MAD* es la mediana de los desvíos absolutos de los valores *Xi* con la mediana y el factor de peso *ui* se calculó con la ecuación:

$$ui = \frac{Xi - M}{0,75 \times MAD}$$

$$Xbi = M + \frac{\left[\sum_{i=1}^n (Xi - M)(1 - ui^2)^2 \right]}{\sum_{i=1}^n (1 - ui^2)^2}$$

$$Sbi = \frac{\left[n \sum_{i=1}^n (Xi - M)(1 - ui^2)^4 \right]}{\sum_{i=1}^n (1 - ui^2)(1 - 5ui^2)}$$

El límite inferior (superior) del intervalo se fijó sumando (restando) al valor de posición el doble del valor de dispersión. El resultado expresa la proporción de datos comprendidos en el intervalo en el total de 360 valores (30 años x 12 meses).

Un índice *I* da una medida del grado en que las series cumplen con los requerimientos de simetría (*S*),

homogeneidad (H) y sin tendencia (T) y reflejan la consistencia (C) interna de los datos:

$$I = 0.25 \times (S + H + T + C)$$

S, H, T, y C en %. Un valor de 100% indica que la serie cumple con los supuestos de un análisis estadístico tradicional y no tiene "outliers".

3. RESULTADOS

Las series pluviométricas presentan menos casos de asimetría y tendencia que las de temperaturas (Cuadro 1). En parte uno podría atribuirlo a diferencias en datos faltantes: 2,5% de los datos de temperatura y 0,7% de lluvia. También influye el efecto de la urbanización, que afecta más al régimen térmico que al de lluvias (Hoffman et al, 1997). Conspira contra la normalidad de una distribución que la base de cálculo sea desuniforme (no todos los promedios se calculan sobre el mismo número de días). El cambio de ubicación de estaciones afecta la homogeneidad, efecto al que no escapan las lluvias.

Cuadro 1. Frecuencia relativa (% de N = 528 = 12 meses * 44 estaciones) de meses con series normales, homogéneas y sin tendencias

% de meses	Temperatura	Precipitación
Normalidad	74	91
Homogeneidad	69	75
Sin tendencia	85	95

Una hipotética estación promedio tendría series de temperaturas con distribución normal en 9 meses del año y en 11 meses para las de lluvia. En Bolívar y Dolores, sólo tres series mensuales de temperatura fueron normales. La mitad de las series mensuales de lluvia de Anguil fueron asimétricas. Cuando una serie no es simétrica, no es posible detectar "outliers" con el promedio y la variancia.

Las series mensuales de temperatura de 8 estaciones no presentaron tendencia lineal. Lo mismo ocurrió con las lluvias de 25 estaciones. La asociación positiva fue más fuerte en las series térmicas (348 casos sobre 528), siendo Octubre el mes con mayor número de estaciones (20) con tendencia significativa.

La consistencia interna no mostró diferencias entre variables y estaciones; la proporción de "outliers" varió entre 4 y 11% (Cuadro 2).

Cuadro 2. Proporción de casos con datos mensuales de la variable incluidos en el intervalo de confianza. N = 360 (12 meses * 30 años)

% de meses	Temperatura	Precipitación
Medio	94	94
Mínimo	89	90
Máximo	96	96

En general la calidad de las series es buena. El índice promedio de las series de temperaturas fue 85% y para las de lluvias 90%, reflejando las diferencias del Cuadro 1. Bolívar y Dolores tuvieron series de temperaturas con índices 67 y 68, en contraste con un valor de 96 en Tres Arroyos (Figura 2). La confiabilidad de los datos de lluvia

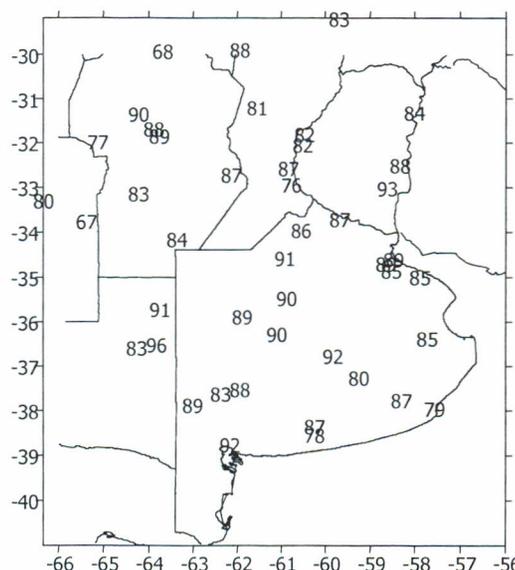


Figura 2. Índice de calidad de las series de temperaturas de 44 estaciones meteorológicas

quedó reflejada en un rango desde 83 en Pigüé hasta 97 en La Plata (datos no mostrados).

4. CONCLUSIONES

Existen diferencias de calidad entre las series mensuales de 30 años de lluvias y temperaturas. Estas muestran mayor frecuencia de casos en que los datos son homogéneos, asimétricos y/o con tendencia. El procedimiento permitió separar estaciones en forma objetiva hasta el punto de identificar los datos de temperaturas de las estaciones Dolores y Bolívar como poco confiables. El estudio describe una metodología de detección de "outliers" en base a medidas de posición y dispersión diferentes a las tradicionales.

5. REFERENCIAS

- Hoffman, J.A.J., Núñez, S.E., Vargas, W.M. 1997. Temperature, humidity and precipitation variations in Argentina and the adjacent sub-antarctic during the present century. *Meteorol. Z.*, 6:3-11.
- Lanzante, J. 1996. Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *Intl. J. Climatol.* 16:1197-1226.
- Siegel, S., Castellan, N. 1988. *Nonparametric statistics for the behavioural sciences*, McGraw-Hill, New York, 399 pp.
- Smith, K. 1995. Precipitation over Scotland, 1757-1992: some aspects of temporal variability. *Intl. J. Climatol.* 15: 543-556.

Agradecimientos

Este estudio fue financiado parcialmente con fondos del Proyecto de Investigación Científica y Tecnológica BID 802/OC-AR-PICT N°02108/97 de la Agencia Nacional de Promoción Científica y Tecnológica. A la Sra. Graciela Cazenave (Inst. de Clima y Agua, INTA) por los datos suministrados.