TRATAMIENTO ESTADISTICO DE DATOS METEOROLOGICOS DEL NOA PARA SU USO EN SIMULACION: ESTADO DE AVANCE

O. Avila y A. Passaro
Departamento de Matemáticas, UNSa.

G. Lesino*
INENCO#, Universidad Nacional de Salta
Buenos Aires 177 - 4400 Salta

RESUMEN

Para realizar una evaluación del comportamiento térmico de un edificio es necesario contar con datos climáticos (temperatura, radiación, velocidad del viento, humedad, en valores medios y de series reales, diarios, horarios) para su simulación computacional. Interesa realizar dicha evaluación en situaciones promedio, en casos desfavorables significativos desde el punto de vista estadístico y en casos extremos. Estas situaciones se dan tanto en el acondicionamiento térmico de verano como en el de invierno. Las variables meteorológicas de interés y la "situación desfavorable" será función del sistema empleado y del recurso climático que utilice.

Los objetivos generales del tratamiento estadístico son:
a) lograr un estudio seriado completo de datos climáticos por medios analíticos y gráficos,
b) diseñar un conjunto de datos estándar para su posterior utilización en simulación.

Para comenzar el estudio, se ha tratado la variable tensión de vapor en Salta. Se continuará con otras variables de interés y en otras localidades de las que existen datos. Se ha hecho un análisis de gráficos y un ajuste empírico de modelos de tipo autorregresivo y de promedio móviles de diferentes órdenes.

1. INTRODUCCION

En Europa y los Estados Unidos se han desarrollado métodos de evaluación del comportamiento térmico de edificios bioclimáticos. El más conocido quizás sea el desarrollado en Los Alamos por D. Balcomb. Todos ellos realizan la hipótesis de que existe un termostato y un calefactor auxiliar que en caso necesario, mantiene las condiciones de confort y evalúa la calidad del acondicionamiento térmico solar a través de la energía auxiliar necesaria para mantener al edificio en la zona de confort.

En los casos de los países en desarrollo y en el de la Argentina en particular, este método no resulta adecuado, ya que no se ajusta a la realidad. Según el nivel económico del usuario, la respuesta de éste a la falta de confort será generalmente acondicionar una zona de la vivienda o en el peor de los casos, soportar el problema. Salvo en los niveles de ingreso altos, la respuesta a una vivienda fría no es calefaccionarla totalmente a 21°C durante todo el día, como suponen los métodos de tipo fracción de ahorro solar, sino calefaccionar algún lo-

* Investigador del CONICET
\# Instituto UNSa.- CONICET

cal durante algunas horas. Esta apreciación es cualitativa. Se están realizando estudios a nivel nacional por la Dirección Nacional de Conservación y Nuevas Fuentes y Grupos de investigación sobre el comportamiento en este sentido de los usuarios.Como en muchos otros temas de tipo social, el problema debe plantearse en términos de calidad de vida y no de combustible o su equivalente en australes ahorrados.

Es necesario planter entonces métodos alternativos de evaluación. Entendemos que una primer idea puede consistir en establecer porcentajes de tiempo que un edificio está dentro la zona de confort, en una zona extendida de "falta de confort tolerable" o francamente fuera de ella. Dado el carácter bioclimático de los edificios a que, en primera instancia, nos referimos, será fundamental el valor que tomen las variables climáticas en los períodos considerados. En el caso de desear realizar simulaciones para prever la calidad de un cierto diseño térmico será de fundamental importancia el conjunto de datos meteorológicos que se emplee en el cálculo. Serán necesarios datos promedio, fluctuaciones, casos desfavorables de baja frecuencia y eventualmente "catástrofes". En los casos de edificios de gran inercia, importarán no solamente los valores diarios sino las secuencias de valores diarios y su frecuencia.

Por lo mencionado anteriormente es que se ha iniciado un análisis de datos climáticos de series temporales, de hasta treinta años, en algunos casos, para la región NOA.

En lo que se refiere al tratamiento estadístico, los objetivos generales son:

a) lograr un estudio seriado completo de datos climáticos por medios analíticos y gráficos.
b) diseñar un conjunto de datos estándar para su posterior utilización en simulación.

Los pasos a seguir en esta etapa son:

a) análisis de series de datos meteorológicos correspondientes a un período de treinta años.
b) selección de variables relevantes por medio de un estudio de regresión múltiple
c) obtención de medidas de posición y dispersión para las series con variables relevantes seleccionadas en b)
d) homogeneización de estas series por métodos estadísticos usuales
e) análisis de existencia de autorregresión, determinación del orden
f) construcción de gráficos diversos: seriales, correlogramas, etc.

2. MODELOS CONSIDERADOS: NOTACION Y RESUMEN TEORICO

2.1. Modelos estacionarios

Dada la serie cronológica (y_t) con observaciones correspondientes a períodos equiespaciados, consideremos como posibles modelos teóricos a los siguientes:

(2.1) $$y_t + b_1 Y_{t-1} + \dots + b_p Y_{t-p} = u_t$$

que es el modelo autorregresivo de orden p o AR(p).

2.2) $$y_t = u_t + a_1 u_{t-1} + \dots + a_q u_{t-q}$$

modelo de promedios móviles de orden q o MA(q). y

2.3) $$y_t + b_1 y_{t-1} + \dots + b_p y_{t-p} = u_t + a_1 u_{t-1} + \dots + a_q u_{t-q}$$

conocido como modelo mixto de órdenes p y q o ARMA (p,q).
En todos los casos, los y_t son los términos de la serie observada y los u_t se consideran como variables aleatorias independientes, con esperanza matemática cero y varianza constante (condición de homocedasticidad) simbólicamente:

$E(u_t) = 0$ para todo t

$E(u_t, u_s) = 0$ para todo $t \neq s$. (independencia)

$= \sigma^2$ para todo $t = s$ (homocedasticidad)

Los elementos a_i, b_i, σ^2 son parámetros; por lo tanto, el modelo AR(p) tiene p+ 1 parámetros, el MA(q) tiene q+1 parámetros y el mixto ARMA (p,q) tiene p+q+1 parámetros.

Una de las notaciones más difundidas es la que permite definir el operador B como:

(2.4) $$By_t = y_{t-1}$$

consecuentemente: $B^r y_t = y_{t-r}$ con $r = 1,2,3,\dots$

con lo cual el modelo AR(p), por ejemplo, puede expresarse mediante:

(2.5) $$y_t + b_1 By_t + b_2 B^2 y_t + \dots + b_p B^p y_t =$$
$$= (1 + b_1 B + b_2 B^2 + \dots + b_p B^p) y_t$$
$$= b_p (B) y_t = u_t$$

Expresiones similares se encuentran para los otros modelos.

2.2. MODELOS ESTACIONARIOS ESTACIONALES

Para las situaciones en las cuales se emplean datos mensuales, además de los elementos comunes como tendencias, ciclos, variaciones estocásticas, etc., deben considerarse posibles movimientos estacionales intraanuales.

En el caso de componentes estacionales podemos suponer un modelo:

(2.6) $$y_t + b_1^{(12)} y_{t-12} + \dots + b_p^{(12)} y_{t-12p} = u_t$$

Con este modelo relacionamos todos los datos de Enero entre sí, los de agosto entre sí, etc.

Empleando el operador B podemos escribir la ecuación anterior como

(2.7) $$b_p (B^{12}) y_t = u_t$$

donde $b_p (B12) = 1 + b_1^{(12)} B^{12} + \dots + b_p^{(12)} B^{12p}$

utilizando posturas análogas podemos escribir los modelos de promedios móviles y mixto estacional como:

(2.8) $$y_t \doteq a_Q (B^{12}) u_t$$

y (2.9) $$b_p (B^{12}) y_t = a_Q (B12) u_t$$

Suponiendo que las series con datos mensuales no tienen sólo la componente estacional sino que además presentan tendencias, ciclos, etc., resulta entonces que una forma factible de combinar estos dos efectos distintos, intraanuales y otros, es por medio de un esquema de tipo multiplicativo. Ante un caso netamente autorregresivo tomaríamos un modelo como:

(2.10) $$b_p (B)\, b_p (B^{12}) y_t = u_t$$

Luego, sobre esta idea, los modelos posibles son:
(2.11) y los siguientes:

(2.12) $$y_t = a_q (B)\, a_Q (B^{12}) u_t$$

(2.13) $$b_p (B)\, b_p(B^{12}) y_t = a_q(B)\, a_Q(B^{12}) u_t$$

La designación de estos modelos son AR (p) x $(P)_{12}$, MA(q) x $(Q)_{12}$ y ARMA (p,q) x $(P,Q)_{12}$ respectivamente.

2.3. LOS MODELOS NO ESTACIONARIOS

En el estudio de series cronológicas puede suceder a menudo que los datos no satisfagan las condiciones siguientes:
1) Las componentes de tipo autorregresivo (estacionales o no estacionales) tienen asociados polinomios cuyas raíces son iguales a 1 o muy próximas a este valor con lo que el proceso aleatorio pierde su carácter estacionario y se convierte en no invertible.
2) Las componentes del tipo de promedios móviles tienen vinculados polinomios con todas sus raíces iguales a la unidad o próximas a ésta, como consecuencia de lo cual, el modelo pierde su invertibilidad. Para tales situaciones los métodos propuestos no son aplicables y consideramos el tratamiento de diferencias sucesivas, trabajando con los operadores.

2.14) $$y_t - y_{t-1} = (1-B)\, y_t = \nabla y_t$$
$$y_t - y_{t-12} = (1-B^{12})\, y_t = \nabla_{12}\, y_t$$

Por lo tanto denotamos a las diferencias sucesivas como:

$$(1-B)^d\, y_t = \nabla^d\, y_t \qquad \text{y} \qquad (1-B^{12})^D\, y_t = \nabla_{12}^D\, y_t$$

y aplicando ambas:

$$(1-B)^d\,(1-B^{12})^D\, y_t = \nabla^d\, \nabla_{12}^D\, y_t$$

A partir de aquí por medio de diferencias sucesivas de forma $(1-B)^d$ o bien $(1-B^{12})^D$, se desea lograr una serie a la que se puedan aplicar métodos estadísticos que estén desarrollados bajo hipótesis de estacionalidad e invertibilidad.

Se consigue una mejora sustancial si se transforma los datos por medio de un función sencilla,

$$y_t^* = \begin{cases} \dfrac{y_t - 1}{\lambda} & \text{si} \quad \lambda \neq 0 \\ \log y_t & \text{si} \quad \lambda = 0 \end{cases}$$

la que permite la continuidad de esta función de λ.
Se pude tomar una versión aún más sencillas de esta función

$$y_t^* = \begin{cases} y_t^{\lambda} \ ; \ \lambda \neq 0 \\ \log y_t \ ; \ \lambda = 0 \end{cases}$$

que es el modelo que se adopta en este trabajo. Así un conjunto de modelos aleatorios adecuados es:

2.16) $$b_p\,(B)\, b_P\,(B^{12})\, \nabla^d \nabla_{12}^D\, y_t = a_q\,(B)\, a_Q\,(B^{12})\, u_t$$

conocido como el modelo ARIMA (p,d,q) x $(P,D,Q)_{12}$, donde ARIMA significa modelo autorregresivo, integrado y con promedíos móviles.

Algunos modelos interesantes son: 1) ARIMA (p,d,O) x $(P,D,O)_{12}$ y

2) $b_{p*}\,(B)\, \nabla^d\, \nabla_{12}^D\, y_t^{\lambda} = u_t$

donde p* es tal que tiene en cuenta la posible presencia de efectos debido a la estacionalidad o sea, cuando p* pueda ser mayor que 12,24 36,.....

Por lo tanto contamos con el esquema:

$$x_t = \nabla \cdot \nabla_{12} \log y_t = (1-B)(1-B^{12}) \log y_t$$
$$= \log y_t - \log y_{t-1} - \log y_{t-12} + \log y_{t-13}$$

Además, las correlaciones parciales observadas, tienden a ser pequeñas frente a las bandas aproximadas $\pm 2/\sqrt{T}$, bajo las hipótesis de independencia.

Se procede a continuación, a ajustar el modelo tomado recientemente. La idea principal es que se debe ajustar, en forma recursiva o secuencial, modelos autorregresivos de orden p menor que 50 empleando las ecuaciones de Yule y Walker y con las estimaciones del modelo de orden superior, proceder a ajustar modelos de promedios móviles con órdenes q hasta 15. Para realizar la determinación del orden óptimo para el modelo autorregresivo tomado, pueden utilizarse los métodos:

- procedimiento de Parzen, el que consisten en elegir el valor de p para el cual el estadístico CAT se hace mínimo.
- procedimiento de Akaike, que consisten en elegir el valor de p de modo que el estadístico FPE(p) tome un mínimo.
- procedimiento de Anderson, que consiste en realizar en forma sucesiva, tests de hipótesis de nulidad del último coeficiente del modelo ajustado, empezando por el orden superior y disminuyendo en forma sucesiva hasta rechazar la hipótesis o hasta aceptar el orden cero.

Del estudio de estos casos se concluye en este ejemplo que: es aconsejable elegir en forma preliminar un modelo autorregresivo de orden 3 puesto que los estadísticos CAT y FPE son mínimos para p=3 mientras que t y t* para la hipótesis de nulidad son máximos también para este valor de p. Luego el modelo aconsejado es (AR (3)) para modelar:

$x_t = \nabla \cdot \nabla_{12} \log y_t$ con lo cual tenemos:

$$(1 + b_1 B + b_2 B^2 + b_3 B^3)(1-B)(1-B^{12}) \log y_t = u_t$$

Para este modelo se han obtenido además las estimaciones de los parámetros y éstas son:

$$b_1 = 0.0266 \qquad b_2 = -0.1348 \qquad b_3 = -0.3367$$

Considerando ahora la información para el ajuste de los modelos de promedios móviles resulta que el estadístico t para testar el nulidad del último coeficiente del modelo es grande para el valor de q=12. Si tomamos esto como significativo, nos llevaría a decir que un componente de tipo autorregresivo significativo estaría presente en los términos de orden 12, 24, 36, 48, etc., en un modelo del tipo AR(p*) con orden grande. Por ej. si p = 27, los valores de t o t* son grandes por lo recomendamos finalmente que se tome un modelo AR(3) o bien uno AR(30) o AR(36) como modelos iniciales y mediante un procedimiento de regresión "paso a paso", buscar un modelo que tenga pocos parámetros estimados. Finalmente, la serie de la variable Tensión de vapor puede ajustarse según:

$$(1 + 0.0266 B - 0.134 B^2 - 0.3367 B^3)(1-B)(1-B^{12}) \log y_t = u_t$$

Con procedimientos análogos se puede tratar las series correspondientes a las restantes variables: temperatura máxima, mínima, heliofanía relativa etc.

Desde luego, es posible que no puedan ajustarse por medio de modelos autorregresivos sino por uno de promedios móviles o en último caso, por medio de un modelo mixto.

Como última acotación cabe mencionar en todos los procedimientos efectuados se observa un coeficiente de correlación múltiple R^2 bajo.

4. TAREAS FUTURAS.

- Generalizar el tipo de estudio realizado para otras variables de importancia en el proyecto, teniendo en cuenta también los datos provenientes de otras estaciones importantes tales como Tucumán, Santiago del Estero, Jujuy y otras.
- Estudiar las posibles correlaciones múltiples entre variables de interés.
- Simular modelos de comportamiento aleatorio aplicados al estudio físico en particular.

5. REFERENCIAS BIBLIOGRAFICAS

- Estimación en los modelos autorregresivos y de promedio móviles. R. P. Mentz. Cuaderno Nº 23 (1982). I.N.I.E. UNT.

- Gráficos estadísticos: Un conjunto de programas de computación. C Martínez, N. Jarma de Cortés. Nota Nº 29 (1982). INIE . UNT.

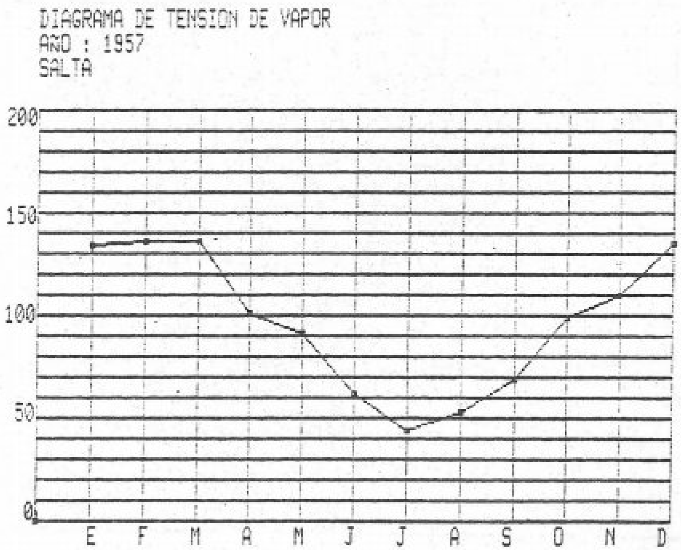


Fig. 1.

Tensión de vapor d=1 D=1 Lamb=0

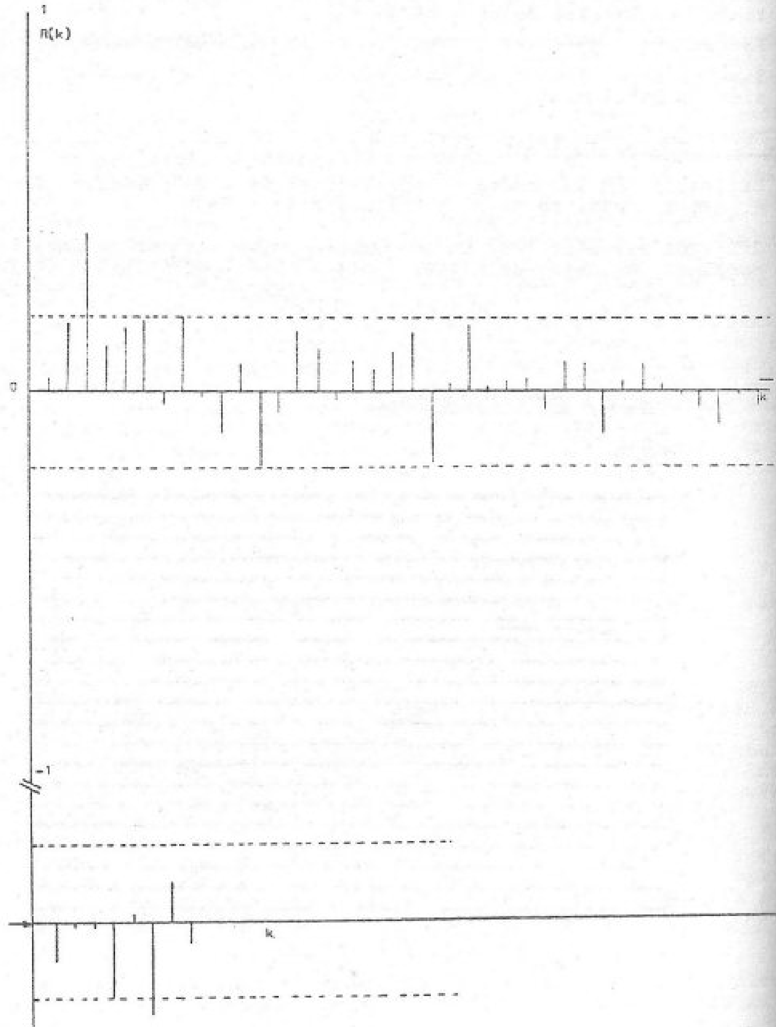
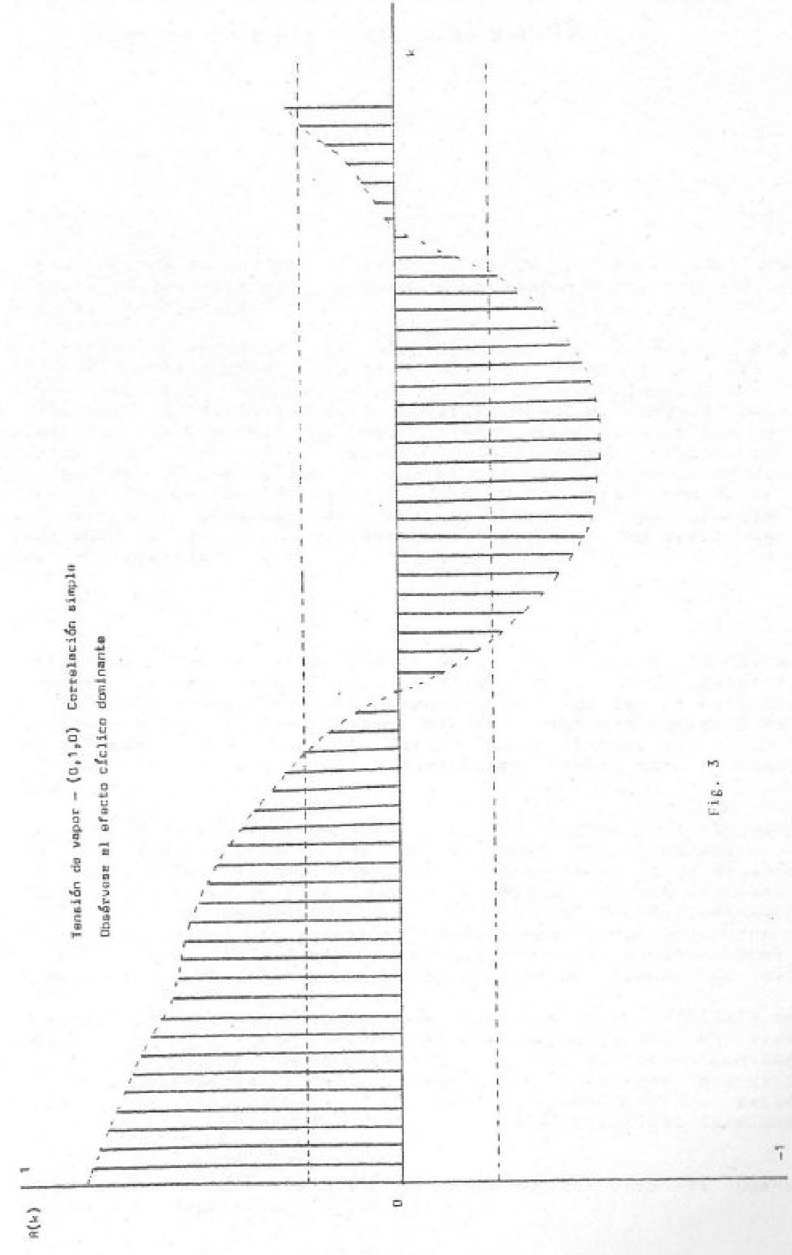


Fig. 2



Tensión de vapor - (0,1,0) Correlación simple

Obsérvese el efecto cíclico dominante

Fig. 3