

## **EMISIVIDAD NOCTURNA: MODELO DE PREDICCIÓN ESTADÍSTICA CON VARIANZA ESTOCÁSTICA**

**E. Frigerio<sup>1</sup>, O. Avila Blas<sup>2</sup>**  
Av. Bolivia 5150 - 4400 Salta  
Tel. : (0387) 4255424, (0387) 4255385,  
Email : [frigerio@unsa.edu.ar](mailto:frigerio@unsa.edu.ar), [oblas@unsa.edu.ar](mailto:oblas@unsa.edu.ar)

**RESUMEN:** Se buscó mejorar modelos estadísticos previos de la emisividad de cielo para la zona, usando el análisis estadístico estructural para el tratamiento de series de tiempo. Se usó para el modelado una serie de emisividad nocturna formada por promedios mensuales obtenida a partir de datos de Salta del Servicio Meteorológico Nacional. Se presenta un modelo que permite tratar series de tiempo no gaussianas mediante la introducción de una varianza de tipo estocástica. La ventaja de este modelo es que permite realizar pronósticos con una alta confiabilidad y a mayor plazo.

**Palabras clave:** Radiación, Atmósfera, Emisividad, Estadística, Análisis Estructural.

### **INTRODUCCION**

El establecer un modelo estadístico apropiado de la emisividad para la región de Salta es de importancia para poder predecir los valores esperados en el corto plazo.

Para poder realizar este estudio se han usado los datos de la estación Salta proporcionados por el Servicio Meteorológico Nacional que abarcan el período junio/1965-diciembre/1984. Al promedio mensual de éstos se les aplicó una correlación que da la emisividad en función de la temperatura de rocío obtenida de ajustar medidas de emisividad en Salta (Frigerio, 2000). La serie de datos obtenida sirvió de punto de partida para el análisis.

En un trabajo previo (Frigerio y Avila Blas, 1999), se hizo el modelado de una serie más restringida de datos, abarcando desde julio/1971 hasta diciembre/1984. El modelo óptimo obtenido en esa oportunidad presentaba la desventaja de mostrar un corrimiento de la serie pronosticada respecto de la serie original. El modelo encontrado en un trabajo posterior (Frigerio y Avila Blas, 2000) no presenta tal dificultad, aumentado la confiabilidad en las predicciones. En estos dos modelos se supuso que la emisividad estaba compuesta por un término promedio fijo al cuál se le agregaba la estacionalidad correspondiente y un término aleatorio que contenía todas las variables no controlables en el proceso de medición. Se partió del supuesto básico de que el término aleatorio asociado a la distribución de los datos tenía una distribución normal simple en el primero y resultante de una combinación de normales en el segundo. En ambos casos se supuso que la varianza asociada a dicho término permanecía constante en el tiempo.

### **FORMULACIÓN DEL MODELO**

Al igual que en los dos trabajos previos, se ha empleado el tratamiento de series de tiempo mediante modelos estructurales, escribiendo un modelo en forma de espacio de estado que tiene la particularidad de suponer que la varianza asociada al término estocástico es una variable aleatoria. Esta suposición fue introducida por observación en la serie de datos de valores extremos pronunciados y de la inspección del correlograma asociado, el que presenta una notable divergencia.

Otros supuestos adicionales son:

- la componente estocástica ( $\epsilon_t$ ) presenta autocorrelación nula con lo que puede suponerse independencia de valores entre pares asociados a tiempos distintos;
- la distribución de esta componente no tiene que ser necesariamente normal pero sí con media cero y varianza unitaria. Esta condición es relevante ya que elimina el condicionamiento que daba cierta rigidez en los modelos previos.

La serie original tratada aquí consta de  $T=223$  valores mensuales. La representación gráfica de esta serie junto con el correspondiente correlograma se presentan en las figuras 1 y 2 .

---

<sup>1</sup> INENCO, Instituto UNSa - CONICET

<sup>2</sup> Cátedra de Probabilidades y Estadística, Fac. Cs. Exactas - UNSa

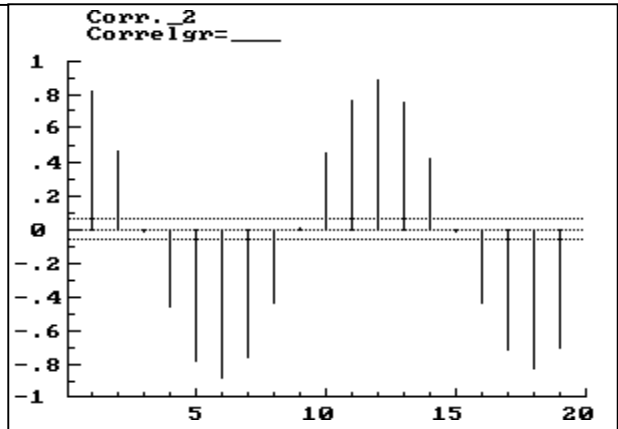
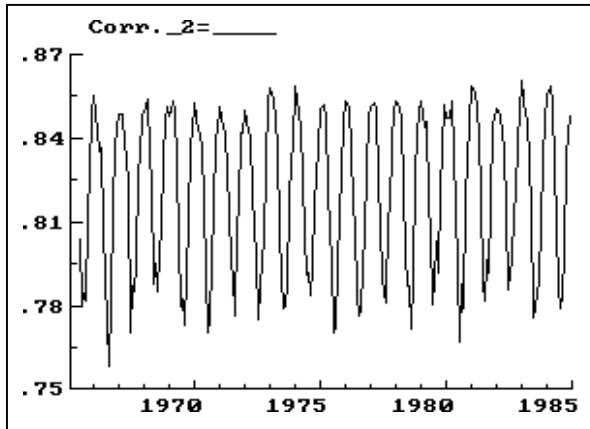


Figura 1 : Serie de emisividad mensual (Corr.\_2)

Figura 2 : correlograma de la serie de emisividad.

Antes de llegar al modelo final, se han intentado diferentes modelos como ser : tomar la variable original y considerarla como una combinación lineal de componente de nivel fijo, estacionalidad fija y componente irregular estocástica; ó bien, estacionalidad dummy manteniendo los otros componentes sin modificación. En todos ellos el modelado no resultó con las propiedades deseadas, dado que los valores de los estadísticos asociados a la bondad de ajuste : BS (de normalidad), Q (para medir autocorrelación serial parcial), DW (para medir autocorrelación serial total) y especialmente el coeficiente de determinación  $R_s^2$ , no presentaron una mejora sustancial respecto a los modelos presentados anteriormente.

La transformación final consiste en proponer un **modelo de varianza** o **volatilidad estocástica** (SV), el que puede formularse como sigue :

$$\text{Emisividad}_t = y_t = \sigma_t \varepsilon_t = \sigma \varepsilon_t \exp(h_t/2), \quad \varepsilon_t \sim \text{IID}(0, 1), \quad t=1, \dots, T \quad (1)$$

$$h_t = \phi h_{t-1} + \eta_t, \quad \eta_t \sim \text{NID}(0, \sigma_\eta^2), \quad |\phi| \leq 1 \quad (2)$$

El término  $\sigma^2$  es el denominado factor de escala,  $\phi$  es un parámetro a estimar y  $\eta_t$  es un término de disturbio que se considera no correlacionado con  $\varepsilon_t$ . La condición  $|\phi| \leq 1$  es la que garantiza la convergencia de las iteraciones en las distintas etapas del proceso de estimación.

La ecuación (2) representa una caminata aleatoria de primer orden que permite definir la función  $h_t$ , conocida como componente de suavizado, actuando sobre los datos de alta variabilidad y sobre los datos extremos.

En este tipo de modelo no podemos realizar directamente la maximización de la función de verosimilitud a fin de aplicar el Filtro de Kalman como en los modelos básicos, debido a que el modelo no es aditivo. Sin embargo, podemos aplicar un método de cuasi-máxima verosimilitud (QML), fácil de aplicar y razonablemente eficiente. El mismo está basado en la siguiente aproximación en serie de Taylor :

$$y_t^* = \log y_t^2 \cong \log (y_t^2 + 0.02 s_y^2) - 0.02 s_y^2 / (y_t^2 + 0.02 s_y^2), \quad t=1, \dots, T \quad (3)$$

donde  $s_y^2$  es la varianza muestral estimada a partir de los datos. Para poder realizar el modelado, se le aplica dicha transformación a la serie original; la serie resultante ( $y_t^*$ ) y su correspondiente correlograma se presentan en las figuras 3 y 4.

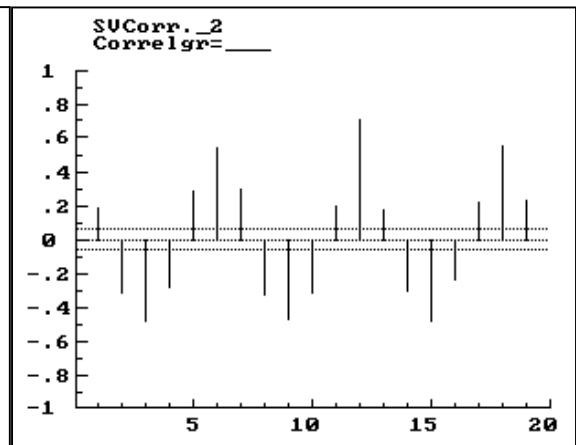
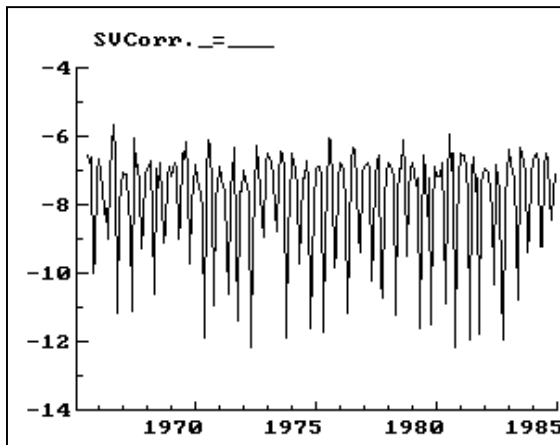


Figura 3: serie  $y_t^*$  (ecuación 3)

Figura 4 : correlograma de la serie  $y_t^*$

Posteriormente se ajusta  $y_t^*$  a un modelo con componentes : nivel estocástico + AR(1) + irregular + estacionalidad fija.

Algunos de los resultados computacionales usando el soft STAMP 5.0 (Koopman *et al.*, 1995) se muestran a continuación :

Resumen de estadísticos para la variable $y_t^*$		Ec. 3 : Devianciones estándar estimadas de los disturbios.	
Error est.	0.83619	Componente	$y_t^*$ (razón q)
Normalidad	20.87	Irr	0.000160103 ( 0.0002)
H( 74)	1.193	Nivel	5.2795e-006 ( 0.0000)
r( 1)	0.0059533	Ar1	0.859257 ( 1.0000)
r(13)	-0.093247	Ec. 3: Coeficiente autorregresivo estimado.	
DW	1.978	El coeficiente rho del AR(1) es $\phi = - 0.1091878$ .	
Q(13,10)	18.40		
Rs <sup>2</sup>	0.55367		

Tabla 1: Resumen de estadísticos más destacados

El estadístico de normalidad junto con los demás indicados en la tabla 1, especialmente el coeficiente de determinación  $Rs^2$ , indica que el modelo transformado es aproximadamente Gaussiano; además, la estimación de la componente irregular  $\varepsilon_t$  es del orden de  $10^{-4}$  lo que justifica considerar una autocorrelación nula, es decir independencia entre estos términos. El valor bajo del coeficiente indica una mejora en el correlograma de la figura 4 respecto al mostrado en la figura 2 en el sentido de asegurar la convergencia de esta función.

Las consideraciones previas, nos permiten realizar la estimación suavizada del proceso de volatilidad  $h_t$ . por medio de la cuál se consiguió una bondad de ajuste muy buena.

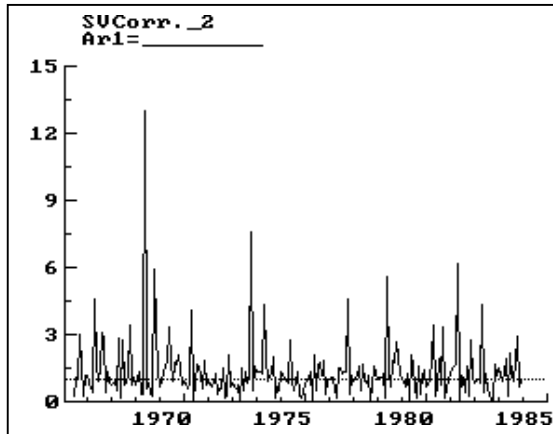


Figura 6 : estimación de la componente de suavizado  $h_t$

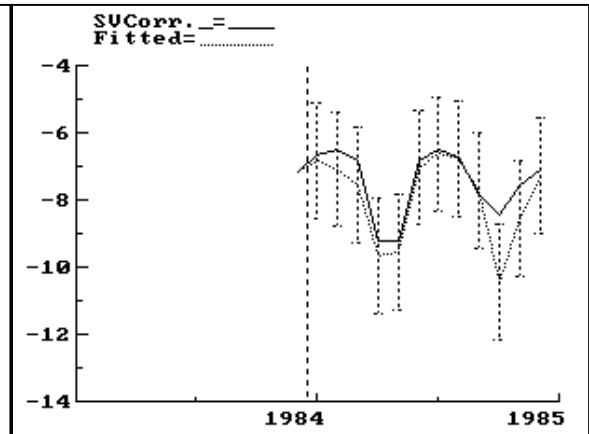


Figura 7 : Bondad de ajuste de la serie  $y_t^*$  vs. la ajustada.

Para estimar el parámetro  $\sigma^2$ , se forma la serie de observaciones corregidas por heterocedasticidad  $\tilde{y}_t$  y a ella se le computan los estadísticos descriptivos usando la transformación inversa *Anti-log* ( $y_t^*$ ):

$$\text{Anti-log}(y_t^*) = y_t \exp(-h_t/2) \tag{4}$$

La gráfica de esta transformación y su correspondiente correlograma se muestran en las figuras 8 y 9. En esta última se aprecia la convergencia final del proceso de estimación.

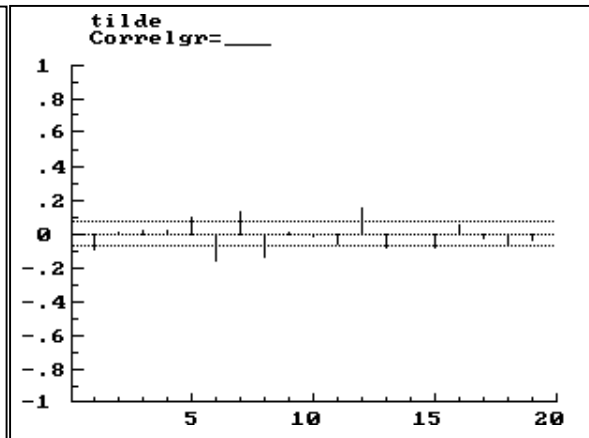
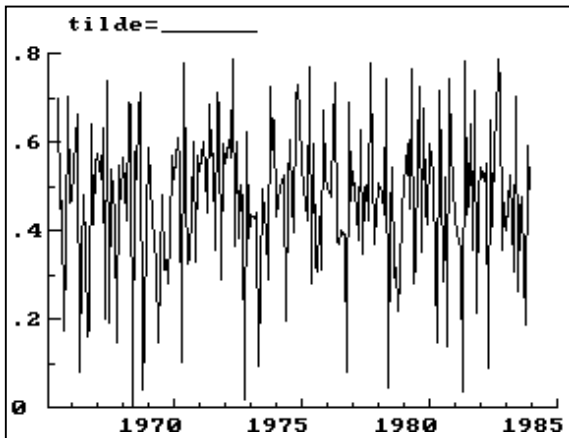


Figura 8: Serie tilde = Anti-log ( $y_t^*$ )

Figura 9: Correlograma de la serie de la figura 8.

Finalmente, el modelo que permite ajustar nuestros datos se puede escribir como :

$$y_t = \sigma_t \varepsilon_t = 0.029 \varepsilon_t \exp(h_t/2), \quad \varepsilon_t \sim \text{IID}(0, 1), \quad t=1, \dots, 223; \quad \text{donde } h_t = -0.109 \cdot h_{t-1} + \eta_t, \quad \eta_t \sim \text{NID}(0, 0.17) \quad (5)$$

En la figura 10 se muestra un dibujo comparativo entre la predicción correspondientes a 12 meses hacia adelante empleando la serie modelada y los valores originales correspondientes. En la tabla 2 se da numéricamente esta comparación conjuntamente con el intervalo de confianza para cada predicción.

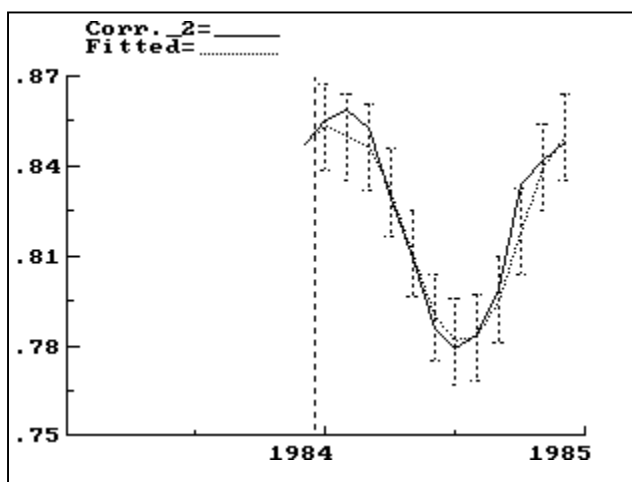


Figura 10: comparación entre los 12 últimos datos originales y los predichos por la serie ajustada (línea débil)

Pronósticos para Anti-log ( $y_t^*$ )					
Período	Serie orig.	Pronóst.	R.m.s.e.	- Rmse	+ Rmse
1984. 1	0.855	0.852745	0.00719019	0.845555	0.859935
1984. 2	0.858	0.849144	0.00730177	0.841843	0.856446
1984. 3	0.852	0.844186	0.00733571	0.836850	0.851522
1984. 4	0.829	0.828916	0.00734756	0.821568	0.836263
1984. 5	0.809	0.809709	0.00735256	0.802357	0.817062
1984. 6	0.785	0.789075	0.00735520	0.781720	0.796430
1984. 7	0.779	0.781855	0.00735693	0.774498	0.789212
1984. 8	0.783	0.783478	0.00735827	0.776120	0.790836
1984. 9	0.799	0.795314	0.00735943	0.787954	0.802673
1984.10	0.834	0.817333	0.00736051	0.809972	0.824693
1984.11	0.842	0.836033	0.00736154	0.828671	0.843395
1984.12	0.848	0.847224	0.00736255	0.839862	0.854587

Tabla 2: valores pronosticados por el modelo usado.

## CONCLUSIÓN.

El modelo usado permitió un pronóstico confiable de la serie tratada del orden del 95 %.

El nuevo modelo presenta ventajas respecto a los usados anteriormente desde el punto de vista estadístico debido a que se consiguió:

- levantar la condición de normalidad en la distribución de los datos;
- disminuir considerablemente el número de iteraciones necesarias para minimizar numéricamente la función de máxima verosimilitud asociada al proceso de estimación;
- se logró aumentar el coeficiente de determinación  $R_s^2$  de 0,41 a 0,55;
- aumentar el número de períodos de predicción “hacia adelante” (hasta 22 meses).

## REFERENCIAS.

- Abril (1999). *Análisis de Series de Tiempo Basado en Modelos de Espacio de Estado*. Apuntes de curso. Tucumán.
- Frigerio, E. (2000) *Emisividad en el N.O.A. Dependencia con la altura*. Avances en Energías Renovables y Medio Ambiente, Vol. 4, N° 2, 11.07-11.12.
- Frigerio, E. y Avila Blas, O. (2000) *Estudio de emisividad: predicción estadística*. Avances en Energías Renovables y Medio Ambiente, Vol. 4, N° 2, 11.19-11.24.
- Frigerio, E. y Avila Blas, O.,(1999). *Análisis estadístico estructural de emisividad*. Avances en Energías Renovables y Medio

Ambiente, Vol. 3, N° 1, 03.29-03.32.  
Harvey A. C. y Shepard N. (1993). *Estructural Time Series Models*. En Handbook of Statistics, Elsevier Science Publishers B.V (Editores), Vol. 11, pp. 261-302.  
Koopman S. J., Harvey A. C., Doornik J. A. y Shepard N. (1995). *STAMP 5.0, Structural Time Series Analyser, Modeller and Predictor*. 1a. edición. Chapman and Hall, London.

#### **ABSTRACT**

The aim was to improve previous statistical models for the sky emissivity of the zone by using structural time series model statistical analysis. A monthly averages nocturnal emissivity series obtained from data of the Servicio Meteorológico Nacional was used for modelling. A model that allows treatment of non-gaussian time series by introducing a random variance. The advantage of this model comes from the fact that it permits to make forecasting with high confidence and for longer periods.

**Keywords:** Radiation, Atmosphere, Emissivity, Statistics, Structural Analysis.