

# Los perfiles meteorológicos a las 1200 GMT y la lluvia en Camagüey.

**Autores:** IBIS R. RIVERO, DANIEL MARTÍNEZ, REINALDO BÁEZ Y ELOIDA HECHAVARRÍA

Centro de Física de la Atmósfera, Instituto de Meteorología. E-mail: finubes@met.inf.cu

## **Resumen.**

*Se investiga la relación con la lluvia de algunas variables de superficie y aire superior. Se utilizaron 1738 observaciones aerológicas realizadas en Camagüey a las 1200 GMT (1982-1991) y los datos de lluvia en 24 horas de la Red Básica del Instituto Nacional de Recursos Hidráulicos (INRH). Como estimadores de volumen de lluvia, se utilizan la lluvia media de todos los pluviómetros con datos válidos (PMED) y la lluvia máxima reportada en una de las estaciones (PMAX). La fracción de área cubierta por lluvia (FRAREA) se estimó a partir del cociente entre el número de estaciones con lluvia y el número total de estaciones con datos válidos. El parámetro que mejor correlaciona con el área cubierta para la muestra es la humedad relativa media en la capa de 0 a 6 km. Al analizar la muestra completa, los diagramas de dispersión muestran dependencia lineal débil entre la mayor parte de los índices y las variables de lluvia, en sentido de aumento de la lluvia con la inestabilidad, aunque dicha dependencia prácticamente desaparece al analizar las muestras para los periodos lluvioso y poco lluvioso por separado. La aplicación del análisis de regresión lineal multivariada a las variables de lluvia en función de las del sondeo da como resultado varianzas totales descritas por el modelo de regresión menores del 33% de la total. Se utilizó el análisis discriminante lineal para determinar categorías de lluvia relacionadas con las variables definidas anteriormente en función de las variables del sondeo, obteniéndose porcentajes de clasificación correcta cercanos al 70%. Para todos los índices en forma individual, se calcularon el valor medio, la desviación estándar y el error estándar de estimación de la media para cada periodo, lo que permitió definir los intervalos de variación de cada índice para cada grupo y los umbrales para los cuales se puede diagnosticar la variable de lluvia utilizando sólo el índice dado, con un nivel de significación superior al 85%.*

## **Introducción**

La influencia de los perfiles meteorológicos obtenidos a través de sondeos atmosféricos en las condiciones del tiempo a escala local se ha relacionado tradicionalmente con los llamados índices termodinámicos, o índices de estabilidad, que han sido utilizados, tanto en la investigación como en la práctica operativa del pronóstico del tiempo. En el trabajo de Pepler y Lamb, (1989) se resumen las definiciones de muchos de estos índices y se discute su efectividad como estimadores de lluvia.

También han encontrado aplicación los estimadores de la cizalladura vertical del viento, como la cizalladura entre dos niveles estándar y la cizalladura burda del viento (Lucas et al., 1994; Westcott et al., 1993; Moller et al., 1994)

Más recientemente, Rasmussen y Straka (1998) estudiaron los perfiles termodinámicos característicos de tres tipos diferentes de tormentas de supercelda definidas a partir de la precipitación que

aportan, obteniendo que las tormentas con precipitación muy alta o muy baja se caracterizan por perfiles termodinámicos ambientales con menores valores de la energía potencial convectiva disponible que aquellas con precipitación normal, observándose una clara relación no lineal en la dependencia lluvia-inestabilidad.

Los trabajos de investigación encaminados a evaluar la efectividad de los índices termodinámicos u otros parámetros energéticos como predictores de fenómenos convectivos en Cuba son muy escasos. Entre ellos se puede citar el de Sosa (1981), para la provincia La Habana, quien señaló el alto grado de información de las magnitudes relacionadas con la humedad.

Alfonso (1985), encontró una ecuación de predicción para el pronóstico a corto plazo de las tormentas eléctricas en la provincia Ciudad de La Habana en función de magnitudes derivadas del sondeo aerológico. De 14 predictores potenciales, se

seleccionaron al final el déficit del punto de rocío a 700 hPa, la componente meridional de la velocidad del viento a 500 hPa y el índice de Showalter.

En los trabajos de Martínez et al., (2000, 2001) sobre la influencia de los perfiles meteorológicos a escala local sobre la lluvia en Camagüey, se utilizaron sondeos cercanos a las 1800 GMT y se utilizó la red básica del INRH. Se analizaron varios índices termodinámicos tradicionales y una serie de parámetros derivados tanto del modelo pseudoadiabático de ascenso de la parcela como de un modelo de nube lagrangiano, estacionario y no adiabático. La variable más informativa fue la humedad media entre 0 y 6 km y también resultaron informativos el índice de George, la altura del tope de la nube predicha por el modelo no adiabático, el exceso máximo de temperatura de la parcela respecto al medio en el modelo pseudoadiabático, la cizalladura del viento en la tropósfera media y en la capa subnubosa.

El objetivo del presente trabajo es determinar la relación de algunas variables meteorológicas de superficie y aire superior determinadas a partir de los sondeos de las 1200 GMT con la lluvia diaria en Camagüey. Para esto se utilizaron los datos de lluvia en 24 horas de la Red Básica del INRH para dicha provincia. Se aplican los métodos de análisis de correlación, regresión lineal múltiple y el análisis discriminante lineal, según el método desarrollado por Martínez et al. (2000; 2001) añadiendo variables que no fueron analizadas en dichos trabajos, con la ventaja de que el uso de los sondeos de las 1200 GMT (0700 hora local) permite trabajar con una muestra mucho mayor y que las conclusiones del trabajo puedan resultar más directamente aplicables al pronóstico de lluvia.

## Materiales y métodos

### Índices termodinámicos y otras variables.

Los índices termodinámicos son magnitudes sencillas, que aportan criterios cuantitativos a partir de los cuales se juzga la posibilidad de ocurrencia de fenómenos relacionados con las nubes convectivas, los cuales han sido utilizadas tradicionalmente como estimadores de inestabilidad o humedad. Pueden calcularse directamente a partir de diagramas termodinámicos, contruidos con los datos correspondientes a los niveles estándar de presión atmosférica contenidos en la información sinóptica de intercambio internacional. Algunos de ellos presuponen la aplicación del modelo pseudoadiabático de ascenso de una parcela y otros utilizan solamente los perfiles de los parámetros del sondeo. En la tabla 1 se describen brevemente los índices utilizados en el presente trabajo.

Un resumen descriptivo de las definiciones de los índices termodinámicos, incluyendo una detallada lista de referencias puede encontrarse en el artículo de Peppler y Lamb (1989).

Además de los índices termodinámicos, en el presente trabajo se han utilizado otras variables, algunas de ellas de superficie, medidas en el punto de sondeo y otras relacionadas con los propios perfiles meteorológicos. Estas son la presión, temperatura y humedad relativa superficiales, la humedad relativa media entre la superficie y 6 km de altura, la cizalladura burda del viento, definida como:

$$B_{\text{shear}} = ((\Delta U)^2 + (\Delta V)^2)^{0.5} \quad (\text{m s}^{-1})$$

$$\text{donde } \Delta U = \frac{1}{\sum_{\text{sup}} \rho_i} \sum_{\text{sup}} \rho_i u_i - \frac{1}{\sum_{\text{sup}} \rho_i} \sum_{\text{sup}} \rho_i u_i, \text{ y análogamente para } \Delta V$$

y la cizalladura superficial del viento, utilizada también por Martínez et al., (2001), es el módulo del vector cizalladura del viento entre los niveles de superficie y 300 m y está relacionado con la vorticidad en la parte baja de la capa subnubosa según Houze, (1993). Valores altos de esta cizalladura van acompañados de intensidades altas de turbulencia en esta capa. Se define a través de la expresión:

$$S_{\text{ushear}} = ((u_{\text{sup}} - u_{300\text{m}})^2 + (v_{\text{sup}} - v_{300\text{m}})^2)^{0.5}$$

donde el subíndice "sup" representa que la variable es de superficie. u y v representan las componentes zonal y meridional de la velocidad del viento. El subíndice "300 m" se refiere a la altura en metros.

Para estudiar la influencia del viento medio en la vertical, se introdujeron en el análisis los valores medios de ambas componentes del viento en las capas 1-6 km y 6-12 km (U1-6, V1-6; U6-12, V6-12).

Tabla 1. Índices termodinámicos calculados para los sondeos de las 1200 GMT en la estación aerológica de Camagüey.

Símbolo de la variable	Denominación	Fórmula o descripción y unidades
SI	Índice de Showalter	$T_{850} - T_{500}$ [°C]. Definiendo para la capa 850-500
LI	Índice de ascenso	$T_{850} - T_{500}$ (ascenso desde la superficie) [°C].
K	Índice de George	$T_{850} + T_{500} - 2T_{550}$ [°C]
TT	Índice total de totales	$T_{850} + T_{500} - 2T_{550}$ [°C]
Jeff	Jefferson	$1.6 \theta_{850} - T_{500} - 0.5(T_{850} - T_{500}) - 8$ [°C]
CT	Cross Totals	$T_{850} - T_{500}$ [°C]
SICP	Showalter modificado.	$T_{850} - T_{500}$ [°C] (niveles de la parcela se inicia en 850hPa, pero se le asigna el contenido de humedad media en la capa 850-500)
K mod	K modificado	$(T - T_{500}) + T_e (T_{500} - T_{300})$ [°C] T y $T_e$ promedios para la capa Sup-850
Sweat	Índice Sweat	$12 T_{500} + 20(T - 49) + 2T_{500} + T_{500} + 25(S + 0.2)$ (-velocidad del viento, s - sea (d <sub>sup</sub> -d <sub>850</sub> )

## Descripción de la muestra.

La muestra empleada consta de 1738 observaciones de la estación aerológica de Camagüey correspondientes al horario de las 1200 GMT del periodo 1982-1991. Se consideraron los sondeos que alcanzaron al menos 6 km de altura y para los cuales, al menos el registro de temperatura resultó válido.

La distribución de los sondeos por meses fue bastante homogénea, como se muestra en la tabla 2.

Tabla 2. Cantidad de datos (N) por meses en la muestra de sondeos de la estación de Camagüey de las 1200 GMT.

MES	E	F	M	A	M	J	J	A	S	O	N	D	TOTAL
N	117	111	165	121	170	137	166	169	180	141	134	127	1738

Los datos de lluvia utilizados para la estratificación de la muestra se calcularon a partir de las observaciones de la red pluviométrica básica del Instituto Nacional de Recursos Hidráulicos (INRH) para la provincia de Camagüey.

Los criterios de estratificación de la muestra y las definiciones de las variables de lluvia, se tomaron del trabajo de Martínez et al. (2001). Como estimadores de volumen de lluvia, se utilizan la lluvia media de todos los pluviómetros con datos válidos

(PMED) y la lluvia máxima reportada en una de las estaciones (PMAX). Para el cálculo de la media no se utilizó la ponderación por área en aras de la simplicidad, y sobre la base de la relativa homogeneidad y adecuada cobertura de la red. La fracción de área cubierta por lluvia (FRAREA) se estimó también de manera simple y aproximada como el cociente entre el número de estaciones con lluvia no nula y el número total de estaciones con datos válidos.

## Resultados y discusión.

### Análisis de correlación.

Para los 1738 casos de la muestra total se calculó la matriz cuadrada de correlación que se muestra en la tabla 3, la que incluye los 1061 casos donde ninguna de las variables resultó faltante. Los valores significativos al 95% se denotan con un "\*" a la derecha del número de las variables estudiadas.

Los índices Jeff, Kmod y K presentan coeficientes de correlación lineal entre sí superiores a 0.93 y lo mismo ocurre con TT, CT y SI. Por este motivo se decidió escoger, de este grupo de índices, solamente a K y SI para el análisis ulterior teniendo en cuenta que han sido utilizados tradicionalmente en Cuba.

La mayoría de las variables estudiadas para la muestra total de datos presentan correlación significativa con los tres parámetros de la lluvia, especialmente con el estimador de área cubierta (Tabla 4).

Tabla 3. Matriz de correlación lineal para los 1061 casos de la muestra total sin variables faltantes a las 1200 GMT en la Estación Camagüey. Los asteriscos a la derecha de las cifras indican que el nivel de significación estadística del coeficiente de correlación es superior al 95%. 1200Z.

VARIABLES	PSUP	TSUP	HS	SI	SICP	JEFF	TT	CT	LJ	K	KMOD	SWEAT	HR6KM	BSHEAR	SUSHEAR	UI 6	VI 6	U6 12	V6 12
PSUP	1	-0.57*	-0.16*	0.34*	0.40*	-0.38*	-0.28*	-0.27*	0.31*	-0.39*	-0.38*	-0.27*	-0.35*	0.01	0.29*	-0.26*	-0.18*	0.05	-0.09*
TSUP	-0.37*	1.00	0.01	-0.62*	-0.65*	0.59*	0.60*	0.49*	-0.88*	0.57*	0.64*	0.54*	0.37*	-0.24*	-0.16*	-0.27*	0.30*	-0.49*	0.07*
HS	-0.16*	0.01	1.00	-0.25*	-0.24*	0.27*	0.21*	0.20*	-0.22*	0.30*	0.18*	0.31*	-0.19*	-0.10*	-0.11*	0.15*	0.15*	-0.24*	0.08*
SI	0.34*	-0.62*	-0.25*	1.00	0.87*	0.92*	0.97*	-0.95*	0.71*	-0.80*	-0.81*	-0.85*	-0.69*	0.13*	0.10*	0.06	-0.30*	0.28*	-0.13*
SICP	0.40*	-0.65*	-0.24*	0.87*	1.00	0.92*	0.87*	-0.76*	0.74*	-0.91*	0.90*	-0.74*	-0.78*	0.11*	0.15*	-0.01	-0.36*	0.27*	-0.18*
JEFF	0.38*	0.59*	0.27*	0.92*	0.92*	1.00	0.88*	0.88*	-0.66*	0.99*	0.95*	0.80*	0.83*	-0.11*	-0.14*	0.01	0.31*	-0.24*	0.15*
TT	-0.28*	0.60*	0.21*	-0.97*	-0.87*	0.88*	1.00	0.93*	-0.77*	0.81*	0.78*	0.84*	0.61*	-0.11*	-0.06	-0.07*	0.28*	-0.27*	0.11*
CT	-0.27*	0.49*	0.20*	-0.95*	-0.76*	0.88*	0.93*	1.00	-0.62*	0.82*	0.75*	0.83*	0.71*	-0.07*	-0.06	-0.02	0.24*	-0.18*	0.10*
LJ	0.31*	-0.88*	-0.22*	0.71*	0.74*	-0.66*	-0.77*	-0.62*	1.00	-0.62*	-0.70*	-0.62*	-0.42*	0.21*	0.11*	0.26*	-0.29*	0.43*	-0.06
K	-0.39*	0.57*	0.27*	-0.86*	-0.91*	0.99*	0.81*	0.82*	-0.62*	1.00	0.97*	0.75*	0.85*	-0.10*	-0.16*	0.03	0.31*	-0.23*	0.15*
KMOD	-0.38*	0.64*	0.30*	-0.81*	-0.90*	0.95*	0.78*	0.75*	-0.70*	0.97*	1.00	0.70*	0.83*	-0.15*	-0.19*	-0.03	0.33*	-0.28*	0.15*
SWEAT	-0.27*	0.54*	0.18*	-0.69*	-0.74*	0.80*	0.84*	0.83*	-0.62*	0.75*	0.70*	1.00	0.60*	-0.08*	-0.05	-0.08*	0.28*	-0.24*	0.10*
HR6KM	-0.35*	0.37*	0.31*	-0.69*	-0.78*	0.83*	0.61*	0.71*	-0.42*	0.85*	0.83*	0.60*	1.00	-0.06	-0.18*	0.11*	0.29*	-0.10*	0.18*
BSHEAR	0.01	-0.24*	-0.19*	0.13*	0.11*	-0.11*	-0.11*	-0.07*	0.21*	-0.10*	-0.15*	-0.08*	-0.06	1.00	0.15*	0.37*	0.17*	0.37*	0.16*
SUSHEAR	0.29*	0.16*	-0.10*	0.10*	0.15*	-0.14*	-0.06	-0.06	0.11*	0.16*	-0.19*	-0.05	-0.18*	0.15*	1.00	-0.16*	-0.02	0.03	-0.08*
UI 6	-0.26*	-0.27*	-0.11*	0.06	-0.01	0.01	-0.07*	-0.02	0.26*	0.03	-0.03	-0.08*	0.11*	0.37*	-0.16*	1.00	0.03	0.68*	0.13*
VI 6	-0.18*	0.30*	0.13*	-0.30*	-0.36*	0.31*	0.28*	0.24*	-0.29*	0.31*	0.33*	0.28*	0.29*	0.17*	-0.02	0.03	1.00	-0.18*	0.45*
U6 12	0.05*	-0.49*	-0.24*	0.28*	0.27*	-0.24*	-0.27*	-0.18*	0.43*	0.23*	-0.28*	-0.24*	-0.10*	0.37*	0.03	0.68*	-0.18*	1.00	0.06
V6 12	-0.09	0.07*	0.08*	-0.13*	-0.18*	0.15*	0.11*	0.10*	-0.06	0.15*	0.15*	0.10*	0.18*	0.16*	-0.08*	0.15*	0.43*	0.06	1.00

Tabla 4: Matriz de correlación entre las variables del sondeo y las relacionadas con la lluvia. Los valores significativos al 95% se denotan con un asterisco a la derecha del número.

VARIABLES DEL SONDEO	VARIABLES DE LLUVIA		
	FRAREA	PMED	PMAX
PSUP	-0,40*	-0,37*	-0,37*
TSUP	0,32*	0,27*	0,41*
HS	0,13*	0,11*	0,13*
SI	-0,39*	-0,33*	-0,39*
SICP	-0,48*	-0,41*	-0,46*
LI	-0,34*	-0,28*	-0,41*
K	0,46*	0,38*	0,43*
SWEAT	0,33*	0,27*	0,32*
HR6KM	0,52*	0,41*	0,41*
BSHEAR	-0,03	0,01	-0,07*
SUSHEAR	-0,22*	-0,20*	-0,17*
U1_6	0,10*	0,16*	0,09*
V1_6	0,30*	0,28*	0,27*
U6_12	-0,08*	0,00	-0,11*
V6_12	0,15*	0,12*	0,14*

Puede observarse que existe una correlación lineal negativa entre los valores de SI, SICP y LI con las tres variables de lluvia seleccionadas, siendo SICP el que mayor índice de correlación muestra con todas ellas. Esto implica que, generalmente, para la muestra total, a mayor inestabilidad puede esperarse mayor cantidad de lluvia. El índice K, también presenta correlación alta con la lluvia, mientras que el SWEAT algo más baja, pero también significativa. La alta correlación del índice SICP con los parámetros de la lluvia puede deberse a que en el cálculo del mismo se le asigna el contenido medio de la humedad de la capa entre 850 y 500 hPa, de manera que está relacionado tanto con la alimentación de las nubes con aire húmedo a través de sus bases como con el intercambio lateral entre las fronteras de la nube y el medio ambiente en la troposfera media. Debe notarse que el parámetro que mejor correlaciona con el área cubierta no es un índice de estabilidad, sino la HR6KM y que la presión y temperatura de superficie poseen coeficientes de correlación del mismo orden que los índices.

En el caso del índice K y el SWEAT se puede apreciar que la dependencia con los parámetros de lluvia es creciente (Fig.1), lo cual implica que el índice es mayor a mayor inestabilidad y humedad. La dependencia de los otros dos índices, que son puramente de estabilidad, es decreciente, denotando también que, para la muestra total, a mayor inestabilidad corresponde más lluvia.

Los diagramas de dispersión muestran que cuando existe estabilidad alta, los valores de los estimadores de lluvia no son elevados y tienen una dispersión relativamente baja para cada valor del

índice dado. Sin embargo, en los casos de alta inestabilidad (LI y SICP muy negativos, o K y SWEAT muy altos) la lluvia presenta una gran variabilidad, existiendo una gran dispersión para los puntos arriba y debajo de la línea de regresión.

En la muestra total están incluidos los datos correspondientes tanto al período lluvioso (PLL) como al período poco lluvioso (PPLL), por lo que la correlación entre las variables del sondeo y las relativas a la lluvia, que se evidencia en los diagramas de dispersión, se debe en cierta medida a la estacionalidad, ya que las diferencias de estabilidad y de precipitación entre los dos períodos pudieran tener un mayor peso en los coeficientes de correlación que las mismas diferencias de inestabilidad entre sondeos correspondientes del mismo período del año.

Para estudiar este fenómeno, en cada punto de la Fig. 1 se situó una etiqueta que simboliza el período a que corresponde cada dato (0 - PPLL; 1- PLL). La estacionalidad se refleja en el hecho de que los casos con mayor inestabilidad se concentran preferentemente en el PLL (a la cual corresponden los mayores valores de FRAREA y PMAX) y los estables, en la poco lluviosa. Por otra parte, la gran dispersión presente en la parte más inestable de los diagramas (especialmente en los de LI y SICP) refleja que la inestabilidad alta no es condición suficiente para la formación de grandes volúmenes de lluvia. Los diagramas del índice K de George muestran un grado similar de dispersión, aunque este índice está relacionado no sólo con la inestabilidad, sino también con la humedad a 850 hPa, de manera que los valores altos de K implican alta inestabilidad y alta humedad.

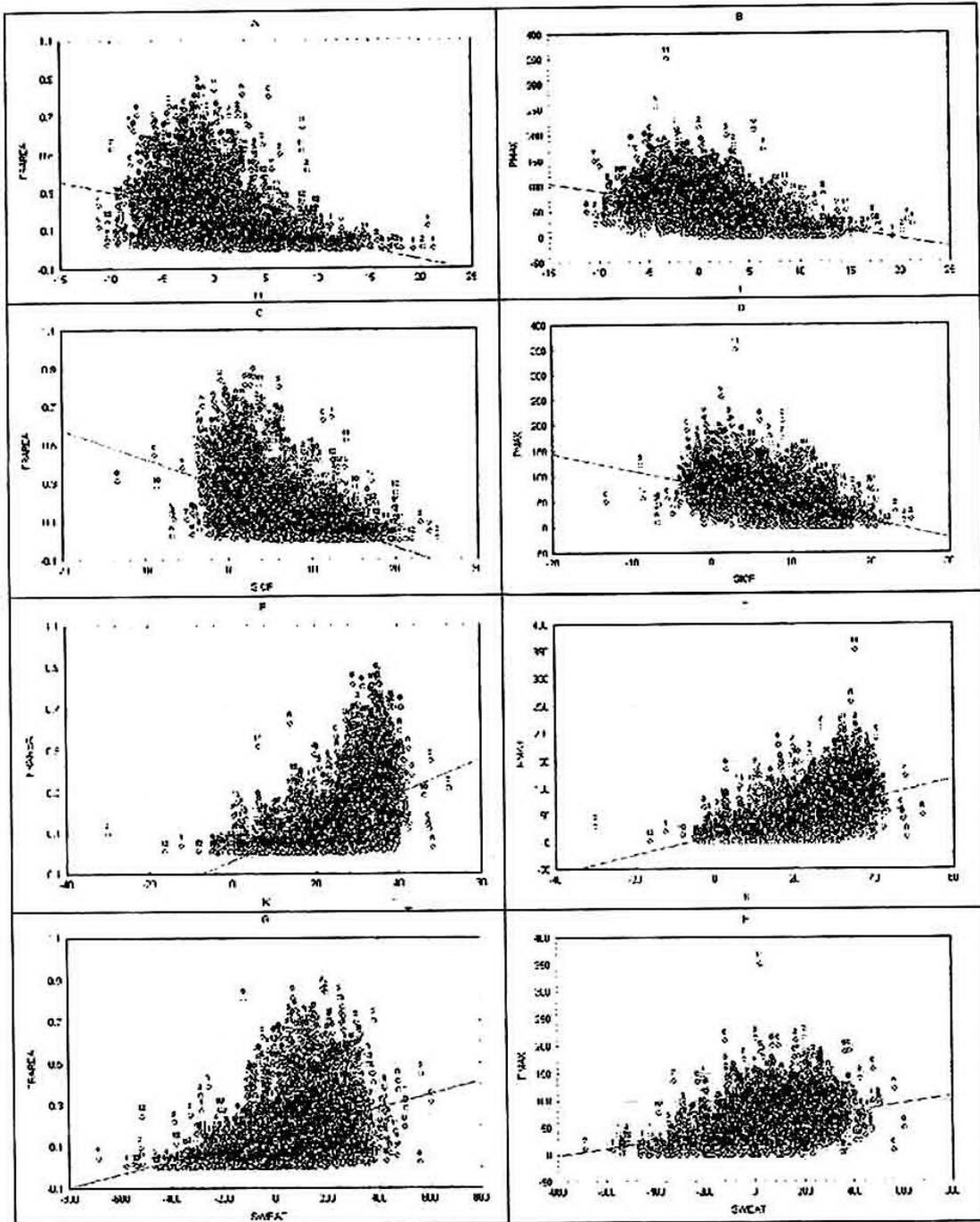


Fig. 1. Diagramas de dispersión de FRAREA y PMAX en función de LI, SICP, Ky SWEAT, para toda la muestra, incluyendo las líneas de regresión para las 1200 GMT en la Estación Camagüey. La etiqueta encima de los puntos corresponde al mes.

La alta diferenciación estacional del comportamiento de los índices que se evidencia de las Figs 2 y 3 sugiere la conveniencia de analizar por separado los datos para los dos períodos, siguiendo la metodología empleada por Martínez et al (2001).

La tabla 4 muestra las matrices de correlación para ambos períodos, la cual evidencia que los coeficientes son, en general, menores para cada período por separado, que cuando se considera la muestra completa. La dependencia de la lluvia con la inestabilidad es más marcada en la estación poco lluviosa que en la lluviosa.

Tabla 4 Matriz de correlación de las variables derivadas del sondeo estudiadas con las variables de lluvia para el período lluvioso (PLL) y el poco lluvioso (PPLL), para la muestra completa a las 1200 GMT en la Estación Camagüey.

VARIABLES DERIVADAS DEL SONDEO	VARIABLES DE LLUVIA					
	(Lluvioso)			(Poco lluvioso)		
	FRAREA	PMED	PMAx	FRAREA	PMED	PMAx
PSUP	-0,37*	-0,34*	-0,23*	-0,20*	-0,18*	-0,21*
TSUP	-0,07	-0,07	0,03	0,21*	0,10*	0,19*
HS	0,10*	0,08	0,05	0,12*	0,08	0,13*
SI	-0,18*	-0,12*	-0,10*	-0,32*	-0,24*	-0,28*
SICP	-0,29*	-0,21*	-0,13*	-0,40*	-0,31*	-0,35*
LI	0,01	0,02	-0,05	-0,27*	-0,16*	-0,24*
K	0,33*	0,25*	0,16*	0,38*	0,27*	0,32*
SWEAT	0,13*	0,08	0,06	0,28*	0,21*	0,23*
HR6KM	0,44*	0,34*	0,22*	0,44*	0,31*	0,35*
BSHEAR	0,06	0,06	0,05	0,09	0,18*	0,06
SUSHEAR	-0,21*	-0,20*	-0,10*	-0,12*	-0,10*	-0,10*
U1_6	0,32*	0,35*	0,26*	0,03	0,13*	0,12*
V1_6	0,20*	0,19*	0,11*	0,23*	0,21*	0,19*
U6_12	0,22*	0,26*	0,15*	0,05	0,17*	0,15*
V6_12	0,11*	0,07	0,02	0,15*	0,12*	0,19*

Se puede apreciar que los valores del coeficiente de correlación de los índices SICP y K para ambos períodos son superiores a los de LI y SWEAT, condicionado esto a la definición de los mismos, los cuales tienen en cuenta además de la inestabilidad, el valor de la humedad en la capa 850 - 500 hPa.

En el comportamiento LI y SWEAT, existen apenas correlación con las variables de lluvia en el PLL, lo que está relacionado con la gran variabilidad que presenta dicho índice, mostrado en el diagrama de dispersión con las variables de FRAREA y Pmax en ese mismo período (Fig 2). Para valores extremos, tanto positivos como negativos de LI y SWEAT resulta poco probable encontrar valores altos de FRAREA y Pmax.

Para el PPLL LI se desplaza hacia los valores mayores (más positivos), que corresponden a mayor estabilidad, aunque siguen existiendo algunos casos con valores negativos, cercanos a los valores extremos de la estación lluviosa. Aunque la variabilidad de FRAREA y PMAx sigue siendo considerable,

los valores de los coeficientes de correlación son superiores a los del PLL, evidenciándose la tendencia a aumentar los estimadores de lluvia para las inestabilidades mayores. Se nota que para valores del índice superiores a 10, FRAREA se encuentra por debajo del 15%.

### 3.2 Análisis por regresión múltiple. Porcentaje de varianza descrita.

Del análisis de los diagramas de dispersión se puede concluir que no es posible establecer relaciones funcionales univariadas entre los parámetros del sondeo y las variables relacionadas con la lluvia, que pudieran conducir a ecuaciones de regresión con un grado aceptable de significación, especialmente para cada período por separado, que es lo que tendría mayor aplicación práctica. Una alternativa a esta situación es el empleo del análisis multivariado, que permite considerar el carácter complejo de las interacciones que se deben tomar en cuenta. La alternativa más simple a considerar es la regresión li-

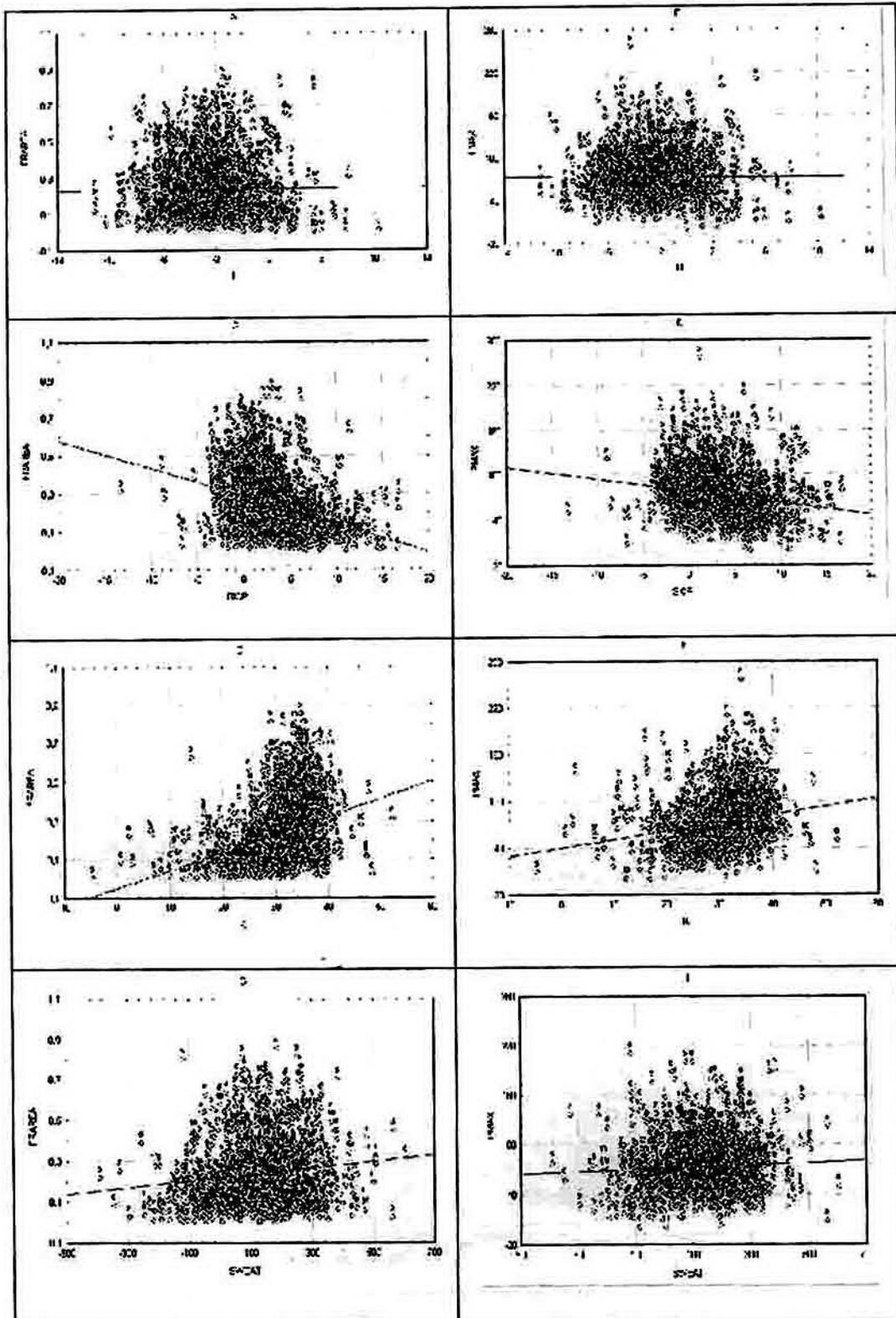


Fig.2. Diagrama de dispersión de FRAREA y Pmax en función de LI, SICP, K y SWEAT, para el PLL, incluyendo las líneas de regresión a las 1200 GMT en la Estación Camagüey. La etiqueta encima de los puntos corresponde al mes.

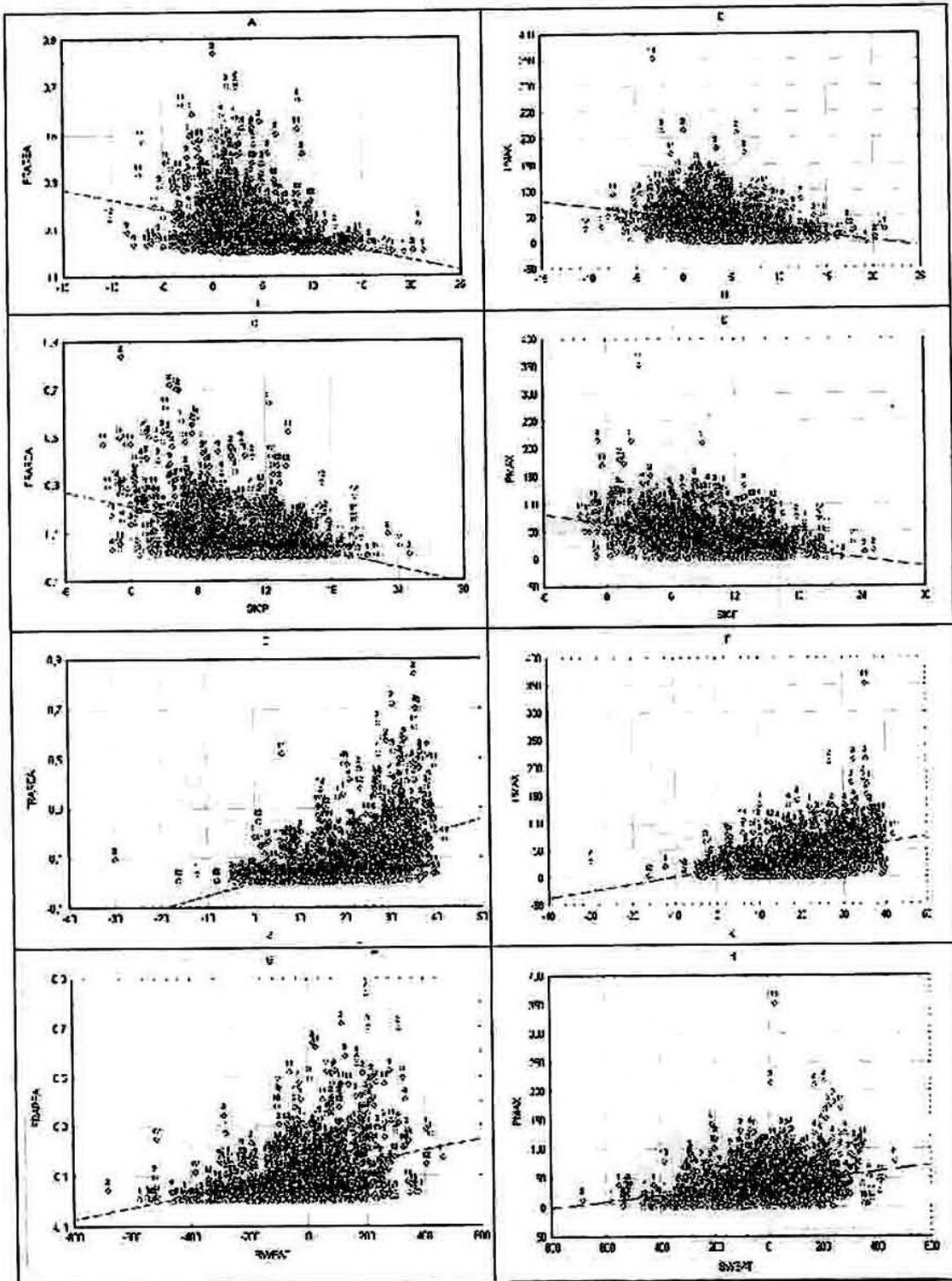


Fig.3. Diagrama de dispersión de FRAREA y Pmax en función de LI, SICP, K y SWEAT, para el PPLL, incluyendo las líneas de regresión a las 1200 GMT en la Estación Camagüey. La etiqueta encima de los puntos corresponde al mes.

neal múltiple, que pudiera permitir obtener ecuaciones para las variables de lluvia a partir de las del sondeo. La aplicación de este método demostró que aún considerando todas las variables del sondeo analizadas anteriormente en conjunto, no es posible estimar las variables de lluvia con una varianza del error suficientemente baja. No obstante, este análisis permitió evaluar la parte de la varianza de la regresión descrita por cada una de las varia-

bles del sondeo con coeficiente de determinación significativo. Esto se hizo empleando el método de regresión lineal progresiva por paso y evaluando en cada paso, el incremento aportado por cada variable a la varianza de la regresión. La tabla 5 muestra las variables que resultaron significativas en la regresión múltiple para los estimadores de lluvia y el porcentaje de la varianza de la regresión descrita por cada una de ellas.

Tabla 5. Porcentaje de la varianza de las variables de lluvia (dependientes), descrito por las variables del sondeo (independientes), por el método de la regresión lineal progresiva por pasos, para la muestra completa a las 1200 GMT en la Estación Camagüey.

Variables de lluvia	Variables del perfil termodinámico							
	HR6KM	Psup	LI	Bshear	Sushear	u1-6	v1-6	u6-12
	Periodo lluvioso							
FRAREA	21.1	6.4				1.7	2.0	
PMED	13.1	6.9				3.5	1.1	
PMAX	1.1	1.7	2.5			6.6		2.1
	Periodo poco lluvioso							
FRAREA	20.5	1.0				6.6	2.0	2.2
PMED	8.0			3.0	1.6	1.0	1.4	2.1
PMAX	11.0				1.0		1.7	2.2

Según los resultados mostrados en la tabla 5, donde se relacionan todas las variables cuya varianza descrita resultó al menos del 1%, la variable de lluvia que presenta mayor relación con los parámetros del sondeo es FRAREA, pero aún en el PLL, cuando el porcentaje de su varianza que puede ser descrito por las variables del modelo es mayor, éste no llega al 30 %, por lo que la utilización de ecuaciones de regresión múltiple no permitiría un diagnóstico preciso del valor de esta variable a partir de la información del sondeo. Por otra parte, se evidencia que la humedad relativa en la tropósfera media y baja es la variable más importante para dos de las variables de lluvia en el PLL y para las tres variables en la PPLL. Un resultado similar encontró Martínez et al. (2001) para sondeos de las 1800 GMT. La única variable para la cual HR6km no resultó entre las variables más informativas fue Pmax para la estación lluviosa, siendo ocupado este lugar por la componente zonal del viento en la tropósfera baja. La fracción de varianza descrita por las variables relacionadas directamente con la inestabilidad, como LI es, en general, muy pequeña en comparación con el aporte de la humedad y la presión superficial. En el PPLL, las variables relacionadas con la cizalladura del viento, contribuyen modestamente a la varianza en el caso de Pmed, mientras que para FRAREA su contribución es algo superior a la correspondiente al PLL. Las componentes del viento tienen un valor significativo en la determinación de la varianza para todas

las variables de lluvia. Estas variables tienen gran importancia, pues reflejan indirectamente la situación sinóptica, tanto de superficie como de aire superior.

### 3.3. Estratificación de la muestra por grupos de casos según las variables de lluvia y aplicación del análisis discriminante lineal

La imposibilidad de aplicar un modelo de regresión, sea univariado o multivariado, que permita establecer relaciones funcionales con un error aceptable entre las variables relacionadas con la lluvia y las derivadas del sondeo, conduce a la aplicación de técnicas que permitan clasificar los días de acuerdo a la lluvia diaria en grupos de casos definidos por diferentes intervalos de variación de las variables de lluvia. El más sencillo de estos métodos es el análisis discriminante lineal, que permite definir con un cierto grado de arbitrariedad la cantidad de grupos que se pretende discriminar y los umbrales que delimitan dichos grupos. En el presente trabajo se emplearon clasificaciones dicotómicas para cada una de las tres variables de lluvia definidas anteriormente, y se emplearon los umbrales definidos por Martínez et al., (2001), tratando de dividir la muestra en grupos de tamaño comparable y de maximizar en lo posible el poder discriminante del método.

El problema consistió en encontrar las variables más informativas y con la menor correlación posible entre sí, que permitieran clasificar los días con la mayor resolución posible.

Tabla 6. Coeficientes de la función discriminante de las variables del perfil termodinámico del sondeo para cada una de las variables de clasificación en el período lluvioso (PLL) y poco lluvioso (PPLL) a las 1200 GMT en la Estación Camagüey.

Variables de clasificación	Variables del perfil termodinámico									
	HR6km	Psup	LI	K	Sushear	Bshear	u1-6	v1-6	u6-12	const
<b>Período lluvioso</b>										
CLAREA	.10	-.20	-.11	-.05		-.14	.16	.05	-.04	198.5
CLPMED	.08	-.21	-.12			-.15	.17	.06	-.05	192.2
CLPMAX	.034	-.25	-.055		10.2		.105			251.6
<b>Período poco lluvioso</b>										
CLAREI	.07	-.07	-.03		-8.54		-.03	.047		67.6
CLPMEDI	.08		-.05			-.09		.15	.04	-5.77
CLPMAXI		-.05		.05		-.09	.03	.08		44.6

Tabla 7. Matrices de clasificación para cada una de las variables y para ambos períodos a las 1200 GMT en la Estación Camagüey.

Grupo observado	Porcentaje de clasificación correcta	Grupo clasificado	
		Gr. 1	Gr. 2
<b>CLAREA</b>			
Gr. 1	73	177	66
Gr. 2	72	119	310
Total	73	296	376
<b>CLPMAX</b>			
Gr. 1	66	333	174
Gr. 2	63	167	282
Total	64	500	456
<b>CLPMED</b>			
Gr. 1	72	148	59
Gr. 2	73	128	338
Total	72	276	397
<b>CLAREI</b>			
Gr. 1	67	211	105
Gr. 2	68	59	124
Total	68	270	229
<b>CLPMAXI</b>			
Gr. 1	66	229	119
Gr. 2	72	93	249
Total	69	322	368
<b>CLPMEDI</b>			
Gr. 1	73	311	113
Gr. 2	73	41	109
Total	73	352	222

El problema consistió en encontrar las variables más informativas y con la menor correlación posible entre sí, que permitieran clasificar los días con la mayor resolución posible.

Las variables de clasificación definidas en ambos periodos fueron las siguientes:

**Período lluvioso**

CLAREA = 1 si FRAREA ≤ 0.15    CLAREA = 2 si FRAREA > 0.15

CLPMED = 1 si PMED ≤ 2 mm    CLPMED = 2 si PMED > 2 mm

CLPMAX = 1 si PMAX ≤ 70 mm    CLPMAX = 2 si PMAX > 70 mm

**Período poco lluvioso**

CLAREI = 1 si FRAREA ≤ 0.1    CLAREI = 2 si FRAREA > 0.1

CLPMEDI = 1 si PMED ≤ 1.5 mm    CLPMEDI = 2 si PMED > 1.5 mm

CLPMAXI = 1 si PMAX ≤ 30 mm    CLPMAXI = 2 si PMAX > 30 mm

La aplicación del análisis discriminante lineal permitió la obtención de ecuaciones discriminantes para cada una de las variables de clasificación, del tipo:

$$Y = \sum a_i x_i + c$$

donde  $x_i$  son las variables del sondeo que resultaron informativas para predecir el valor de la variable de clasificación,  $a_i$  son los coeficientes de la ecuación para la variable y  $c$  es una constante. Si  $Y > 0$ , la variable de clasificación se asigna al grupo 2 y en caso contrario, al grupo 1. La tabla 6 muestra los coeficientes obtenidos para dichas ecuaciones, aplicando el método de regresión lineal progresiva y en la tabla 7 se observan las matrices de clasificación correspondientes en la cual los porcentajes de clasificación correctos están próximos al 70 % para todos los grupos, siendo los más altos los correspondientes a las variables CLAREA y CLPMED. La variable CLPMAX es más difícil de predecir debido a su mayor aleatoriedad, ya que está definida por la medición de un solo pluviómetro y puede depender de efectos puramente locales como interacciones entre sistemas convectivos y entre estos y los vientos locales. Estos valores son algo menores que los reportados por Martínez et al. (2001) para una muestra mucho menor de sondeos de las 1800 Z. Esto se debe, sobre todo al mayor intervalo de tiempo entre el sondeo y la ocurrencia de la lluvia, que en la mayoría de los casos tiene lugar en horas de la tarde o la noche.

### 3.4. Evaluación de los índices y otros parámetros en forma individual.

La utilización tradicional de los índices de estabilidad como predictores de diversos fenómenos relacionados con la lluvia confiere especial interés a la evaluación del comportamiento de sus valores para las diferentes categorías de lluvia.

Con este objetivo se evaluaron los valores medios, desviaciones estándar y errores estándar de estimación de la media para cada uno de los índices por separado para cada grupo, tomando a CLAREA y CLAREI como variable de clasificación.

El método empleado permitió obtener los valores medios de los índices y umbrales a partir de los cuales se puede diagnosticar la variable de lluvia.

La Fig. 4 muestra, como ejemplo, los diagramas de caja y bigote correspondientes a cada grupo para el índice SICP, que permiten visualizar la variación de dicho parámetro y definir valores umbrales del mismo relacionados con cada grupo de lluvia.

Las diferencias de las medias para ambos grupos son significativas, al no existir solapamiento entre los intervalos de valores definidos por la altura de las cajas, que representan los intervalos de confianza. No obstante, la gran variabilidad de los valores individuales de SICP, reflejada por la desviación estándar, limita el intervalo de valores de SICP que permiten discriminar el grupo, de forma tal que:

Para el PLL si  $SICP \geq 5.8$ , entonces el día corresponderá con mayor probabilidad al grupo 1 ( $FRAREA \leq 0.15$ ); si  $SICP \leq 0.25$ , entonces el día corresponderá al grupo 2 ( $FRAREA > 0.15$ ). Los casos intermedios podrán ser clasificados en cualquiera de los dos grupos.

Para el PPLL, si  $SICP \geq 12.2$ , entonces el día corresponderá con mayor probabilidad al grupo 1 ( $FRAREA \leq 0.1$ ); si  $SICP \leq 6.2$ , entonces el día corresponderá al grupo 2 ( $FRAREA > 0.1$ ). Los casos intermedios se podrán clasificar en cualquiera de los dos grupos.

En los diagramas de caja y bigotes analizados para las diferentes variables del sondeo seleccionadas, se aprecian los valores medios y los umbrales que se presentan en la Tabla 8.

Los valores umbrales de las variables de sondeo, que definen el grupo para las variables de clasificación CLAREA y CLAREI, se determinaron con un nivel de significación del 85%, ya que la amplitud de los "bigotes" se escogió como igual a la desviación estándar. Todas las medias son significativamente diferentes para ambos grupos a más del 95%, excepto la de LI para el PLL.

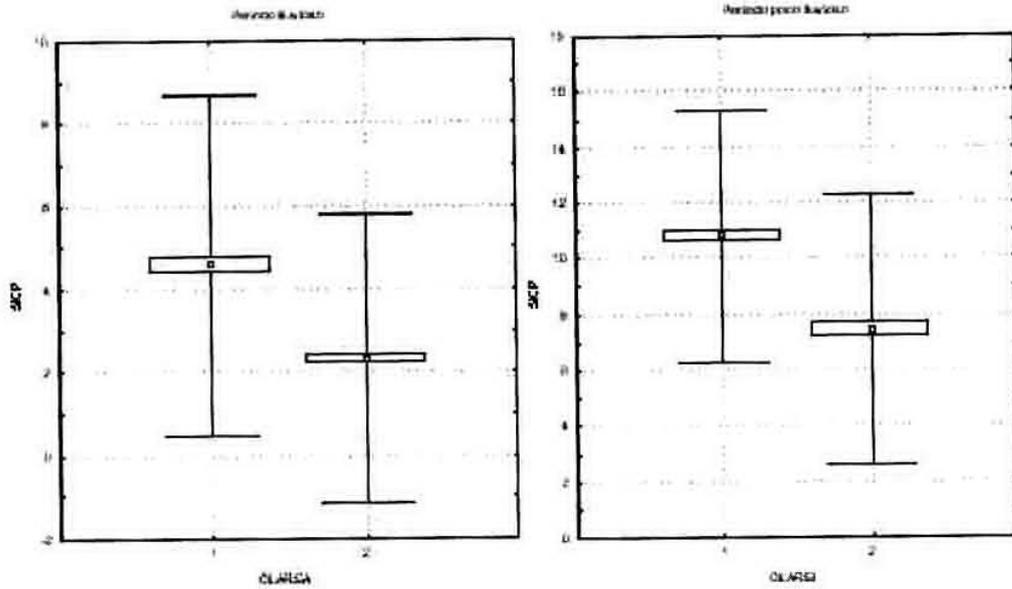


Fig4. Diagrama de caja y bigote correspondiente a las variables de clasificación CLAREA y CLAREI para el índice SACP a las 1200 GMT en la Estación Camagüey. El punto central del diagrama representa el valor de la media, mientras que el rectángulo o "caja" representa el error de estimación de dicha media. Las líneas verticales acotadas, o "bigotes", representan la desviación estándar.

Tabla 8. Medias y valores umbrales de las variables derivadas del sondeo, para las variables de clasificación CLAREA y CLAREI, para el período lluvioso (PLL) y poco lluvioso (PPLL) a las 1200 GMT en la Estación Camagüey.

Variable de clasificación	Variables derivadas sondeo	- Grupo 1		Grupo 2		Indeterminado
		Media	Umbral	Media	Umbral	
CLAREA (PLL)	SICP	4.6	$\geq 5.8$	2.2	$\leq 0.25$	$0.25 < SICP < 5.8$
	K	28.5	$\leq 27$	32.5	$\geq 35.5$	$27 < K < 35.5$
	SI	3.5	$\geq 3.2$	0.4	$\leq -2.1$	$-2.1 < SI < 3.2$
	LI	-2.5	$\geq -0.9$	-2.8	$\leq -5.3$	$-5.3 < LI < -0.9$
	SWEAT	120	$\leq 20$	150	$\geq 270$	$20 < SWEAT < 270$
	HR6Km	69.9	$\leq 68$	77	$\geq 79$	$68 < HR6km < 79$
CLAREI (PPLL)	SICP	10.8	$\geq 12.2$	6.5	$\leq 6.2$	$6.2 < SICP < 12.2$
	K	18.5	$\leq 17.5$	26.5	$\geq 28.5$	$17.5 < K < 28.5$
	SI	5.8	$\geq 7$	3.4	$\leq 1.5$	$1.5 < SI < 7$
	LI	4.5	$\geq 5.8$	2	$\leq 0$	$0 < LI < 5.8$
	SWEAT	-60	$\leq -100$	50	$\geq 140$	$-100 < SWEAT < 140$
	HR6Km	63	$\leq 61.5$	71.5	$\geq 74$	$61.5 < HR6km < 74$

## Conclusiones

Se observa que para la muestra total de datos, la mayoría de las variables estudiadas presentan correlación significativa con los tres parámetros de la lluvia, especialmente con el estimador de área cubierta. El parámetro que mejor correlaciona con el área cubierta no es un índice de estabilidad, sino la HR6KM y que la presión y temperatura de superficie poseen coeficientes de correlación del mismo orden que los índices. Los índices que mejor correlacionan para la muestra total son el SICP y K.

Los diagramas de dispersión para la muestra total presentan dependencia lineal débil entre la mayor parte de los índices y las variables de lluvia, en sentido de aumento de la lluvia con la inestabilidad. Al estudiar las matrices de correlación y diagramas de dispersión por periodos, se concluye que las dependencias son en general, del mismo signo que para toda la muestra, con coeficientes de correlación inferiores. Esto se explica a partir de que una componente importante de la variabilidad en la muestra completa es la dependencia estacional. Para el período lluvioso (PLL), los coeficientes de correlación son más bajos que para el poco lluvioso, aunque no llegan a cambiar de signo.

La aplicación del análisis de regresión lineal multivariada a cada una de las variables de lluvia en función de las derivadas del sondeo, produce valores de la varianza descrita por el modelo de regresión menores del 33% de la total, lo que implica que no sea recomendable la obtención de ecuaciones de regresión múltiple. La variable del sondeo que describe el mayor porcentaje de la varianza de la de lluvia es HR6km, excepto en el caso de Pmax en el período lluvioso (PLL), para el cual la de mayor varianza descrita es u1-6.

Se definieron tres variables de clasificación dicotómicas para cada período, una para cada variable de lluvia. La aplicación del análisis discriminante lineal al diagnóstico de estas variables permitió la clasificación correcta de cerca del 70 % de los casos para cada una de las variables a partir de las ecuaciones discriminantes.

Considerando el valor medio, la desviación estándar y el error estándar de estimación de la media de todos los índices de forma individual, para cada período se definieron los intervalos de variación de cada índice en cada grupo y los umbrales a partir de los cuales se puede diagnosticar la variable de lluvia utilizando sólo el índice dado, con un nivel de significación superior al 85%. Los valores medios de los índices por grupos se diferenciaron con más de un 95% de significación, excepto en el caso del LI para el período lluvioso (PLL), que no presentó diferencia significativa.

## Reconocimientos

Los autores agradecen las valiosas críticas y sugerencias de Luciano Amaro, Dulce Pérez y Mario Carnesoltas. El trabajo se realizó con apoyo financiero del Proyecto 014-1999 de la Agencia de Medio Ambiente del CITMA de Cuba, del Proyecto IAI/PES-CA "Aplicación de modelos numéricos de alta resolución en la simulación de la circulación atmosférica a escala local y regional en áreas del Caribe", adjunto al Proyecto IAI/CRN-073 y del Proyecto de Colaboración Cubano-Mexicano CITMA-CONACyT 01-2000 "Modelación matemática de nubes convectivas y de su impacto en la evolución de los contaminantes de la atmósfera". Se agradece además la colaboración del INRH, que proporcionó los datos de lluvia.

## Referencias

- Alfonso A., 1985:** Pronóstico a corto plazo de turbonadas en la Provincia Ciudad de La Habana. *Ciencias de la Tierra y el Espacio*, No. 10, 115-18.
- Houze R. A. Jr., 1993:** Cloud dynamics. *International Geophysics Series*. Vol. 53, 573 pp. Academic Press.
- Lucas C., Zipser E. J. and LeMone M. A., 1994:** Vertical velocity in oceanic convection off tropical Australia. *J. Atmos. Sci.*, 51: 3183-3193.
- Martínez D., L. Alfonso, R. Báez and I. Jo, 2000:** The influence of Meteorological profiles at a local scale in rainfall over Camagüey, Cuba. *Proceedings of 13th International Conference on Clouds and Precipitation*. Reno, Nevada, USA. Vol. 1, 440-443.
- Martínez D., Alfonso L., Báez R., and Jo I, 2001:** La influencia de los perfiles meteorológicos a escala local en la lluvia sobre Camagüey, Cuba. *Rev. Bras. Meteor.* 16(1), 67 - 79
- Moller A.R., Doswell C. A., Foster M. P., and Woodall G. R., 1994:** The operational recognition of supercell thunderstorm environments and storm structures. *Wea Forecasting*, 9: 327-347.
- Peppler R. A. and Lamb. P. J., 1989:** Tropospheric stability and Central North American growing season rainfall. *Mon. Wea. Rev.*, 117: 1156-1180.
- Rasmussen E. N. and Straka J. M., 1998:** Variations in supercell morphology. Part I: Observations of the role of upper-level storm-relative flow. *Mon. Wea. Rev.*, 126: 2406-2421.
- Sosa J.A., 1981:** Pronóstico objetivo de turbonadas para las provincias La Habana y Ciudad de La Habana. *Ciencias de la Tierra y el Espacio*, No. 3: 119-127.

**Westcott N.E., S. A. Changnon, R. R. Czys, R. W. Scott and M. S. Patersen, 1993: Results of the 1989 exploratory cloud seeding experiment in Illinois based on synoptic weather conditions. J. Wea. Mod., 25, 1, 26-49.**

**Abstract**

*The relation between rainfall and some surface and upper air variables. Data from 1738 1200 GMT soundings of the Camaguey radiosond station and 24 hour daily rainfall data of the National Institute for Hydraulic Resources (INRH) Basic Network were used. As estimators of rainfall volume were chosen the the average rainfall measured in all raingauges with valid data, and the maximum rainfall reported by one of the raingauges (P<sub>MAX</sub>). The fractional area covered by rain (FRAREA) was estimated as the ratio between the number of raingauge stations with rainfall and the total number of station with valid data. The best correlating parameter with rain covered area is the relative humidity in the 0-6 km layer. The scatterplots*

*of the whole sample show weak linear dependence between most of the indices and the rainfall variables, in the sense of increasing rainfall with instability, though this dependence practically disappears when the samples for the rainy and less rainy periods are analyzed independently. The application of multivariate regression analysis to the rainfall variables as function of the sounding related ones produces total variances described by the regression model less than 33 % of the total variance. Linear discriminant analysis was used to find rainfall categories defined by the rainfall related variables to be related with the sounding variables. The percentages of correct classification were near to 70 %. Mean value, standard deviation, and standard error of determination of the mean were found for all the indices for each of the two periods. This allowed to define the ranges of variation of each index in each category and the threshold values for which the rainfall variable can be classified using only the given index, with a significance level higher than 85 %.*

**Palabras clave:**

perfiles meteorológicos, inestabilidad, lluvia.