

PREDICCIÓN ESTADÍSTICA GEM. APLICACIÓN A LOS JUEGOS OLÍMPICOS DE BARCELONA

Guillermo Balester Valor
Oficina meteorológica
de la Base Aérea de Armilla (Granada), INM

INTRODUCCION

El objetivo de este artículo es doble, por una parte, la descripción breve del modelo de predicción estadística GEM de forma que éste sea conocido entre los predictores y demás personal del INM, por otra, la exposición del estado actual de desarrollo de este modelo por el INM y el resultado de las primeras verificaciones en el Area Olímpica.

La aplicación a los Juegos Olímpicos de Barcelona de 1992 ha dictado la máxima prioridad en la adaptación de este método. En la actualidad, se está desarrollando un esquema de predicción inspirada en GEM para esta zona. Existen varias versiones adaptadas por el autor para el INM en sistemas basados en PC's y desarrollados íntegramente en Fortran pero que utilizan las ecuaciones de validez general desarrolladas en EEUU.

Como se indica en este artículo, la generalidad de estas ecuaciones es la máxima limitación del método. Las conclusiones son muy claras: no sirven para mucho en la zona estudiada. Hay que incluir de alguna manera los condicionantes locales. A la vista de estos resultados se está desarrollando una predicción que, muy similar al GEM, utiliza ecuaciones desarrolladas con datos del Area Olímpica.

DESCRIPCION DE GEM

Aquí sólo se exponen las características más significativas del GEM. Para un estudio exhaustivo véase (1). El nombre de este método proviene de su denominación en inglés (Generalized Equivalent Markovian).

Es General porque unas mismas ecuaciones son aplicadas para cualquier observatorio en cualquier época del año (aunque dentro de los predictores se incluye, por ejemplo, el mes y la estación del año).

Es casi Equivalente a un proceso de Markov de orden 1. Se trata de una aproximación lineal, en este caso la tabla de probabilidades de transición se sustituye por un sistema de ecuaciones de regresión lineal múltiple. Los predictandos son el valor de los predictores una hora más tarde. Si la salida de estas ecuaciones son introducidas de nuevo en el esquema, se va obteniendo la evolución de las variables meteorológicas a intervalos de una hora, y por eso recuerda a un proceso markoviano.

Este tipo de ecuaciones tiene similares propiedades de convergencia que las markovianas puras: tras determinado número de pasos la predicción se aproxima a la climatología, de forma que ésta deja de tener validez más allá de 8 o 10 horas (o iteraciones).

Otra de las características más notable del GEM, y la más sorprendente, es que sólo se incluyen como predictores datos de superficie y nubosidad de una sola estación en un determinado instante. Con estos datos, cualquiera que tenga algún contacto con la predicción intuirá que esto sólo puede ir bien cuando el tiempo atmosférico evolucione de acuerdo a lo normal en cada zona, es mucho pedir que detecte cambios bruscos.

En GEM, las variables numéricas (como pueden ser la temperatura, presión, altura de capa nubosa, etc...) son divididas en clases, de forma que el valor de cada predictor (y pre-

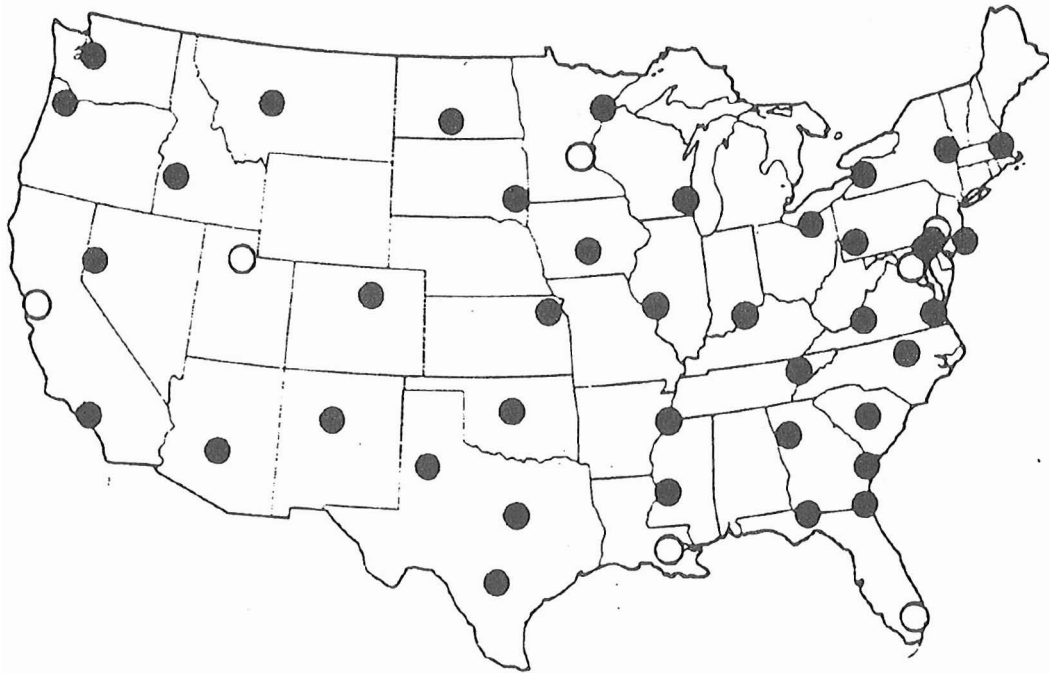


Figura 1: Localización de las estaciones para la construcción de GEM. Los círculos oscuros denotan aquellos observatorios con los que se han construido las ecuaciones de predicción. Los círculos claros indican las estaciones seleccionadas para verificación.

```

      XXX XXXXX X X
      X XX X XX XX
      X XX X XX X X X
      X X X X X X
      XXX XXXXX X X
  
```

TECHNIQUES DEVELOPMENT LABORATORY

FOR STATION: DCA

VALID FOR 12 HOURS AFTER MAR 21, 1980 7 LOCAL

	7	8	9	10	11	12	13	14	15	16	17	18	19
TEMPERATURE (F)	60	61	64	66	67	67	68	68	67	66	64	62	59
DEW POINT (F)	59	60	61	60	60	59	58	57	56	55	54	53	53
VSRY (1000MS SM)	0600	0600	0600	0600	0600	0600	0600	0500	0500	0400	0400	0400	0400
FOG, ICE FOG	F	F	F	F	F	F	F	F	F	F	F	F	F
GROUND FOG													
SMOKE, HAZE													
FLOWING													
DRIZZLE													
RAIN													
RAIN SHOWER	R-	R-	R-	R-	R-	R-	R-	R-	R-	R-	R-	R-	R-
SNOW, IC													
SNOW SHOWER, SLP													
FREEZE DRIZZLE													
FREEZE RAIN													
THUNDERSTORM													
THUNDERSTORM*													
WIND (DREF)	151X	1710	1719	1820	1921	2021	2121	2221	2321	2321	2421	2419	2410
SLP (1013 HPa)	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
CLOUD COVER #1	000	000	000	000	000	000	000	000	000	000	000	000	000
CLOUD HEIGHT #1	7	7	7	7	7	7	7	7	7	7	7	7	7
CLOUD COVER #2	000	000	000	000	000	000	000	000	000	000	000	000	000
CLOUD HEIGHT #2	10	10	10	10	10	10	10	10	10	10	10	10	10
LOT CLOUD COVER	000	000	000	000	000	000	000	000	000	000	000	000	000
CEILING 1000 FT	7	7	7	7	7	7	7	7	7	7	7	7	7

Figura 2: Ejemplo de salida de GEM.

horas sucesivas se utilizan distintas matrices de coeficientes, en vez de remultiplicar cada salida por la matriz de coeficientes a una hora, como sugiere un proceso markoviano. La razón estriba en que la multiplicación se haría mucho más lenta a partir de la primera hora porque la entrada ya no se trataría de ceros y unos. La contrapartida es que se utiliza mucho más espacio en disco para los coeficientes, aproximadamente 70 Kbytes por cada hora que se avance en la predicción.

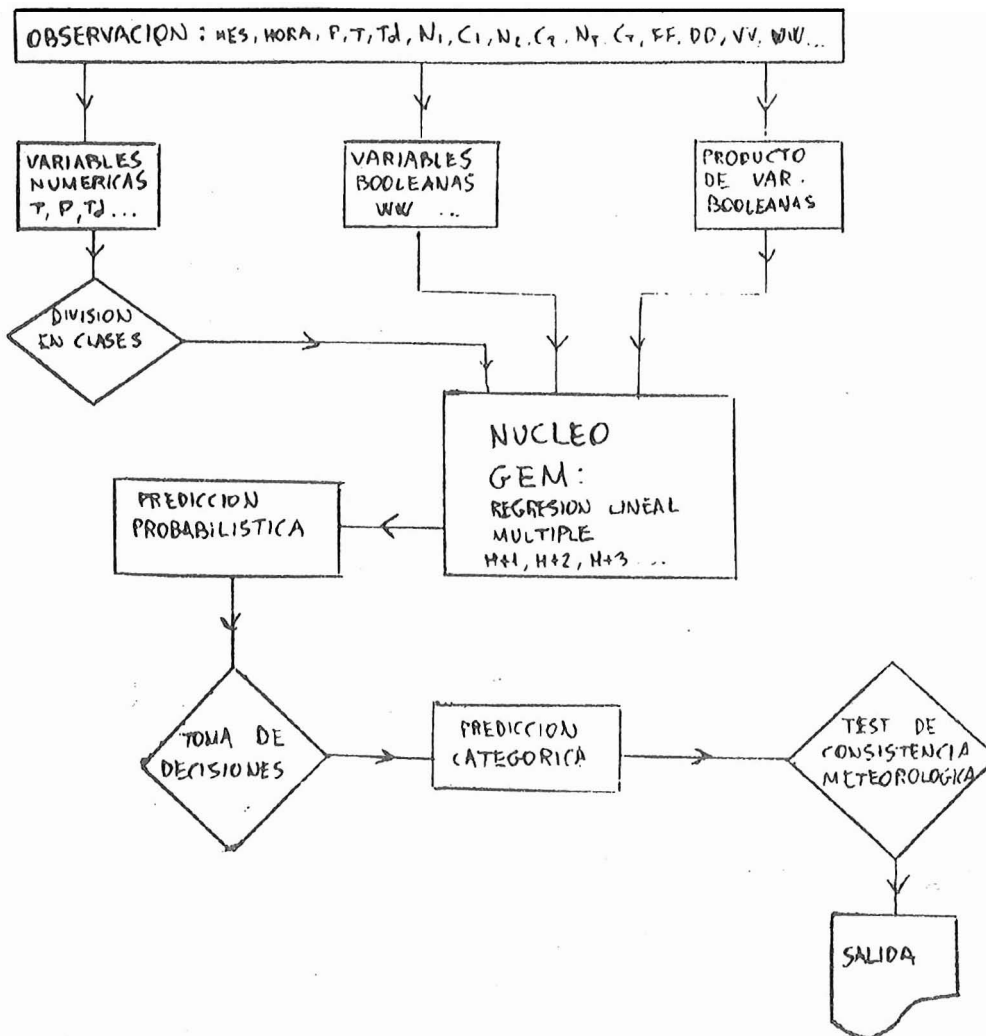


Figura 3: Esquema general de la predicción GEM.

El resultado de cada iteración es una predicción estadística, en la que cada clase tiene una probabilidad dada. Esto no tiene mucha utilidad para el predictor, ya que es difícilmente interpretable. La predicción la recoge entonces un módulo de toma de decisiones y pasa a ser categórica, en la que se pronostica el valor numérico de cada variable y la existencia o no de determinados fenómenos. Pero como "la estadística no sabe meteorología" pueden darse predicciones sin sentido (como nevadas intensas con buena visibilidad, nubosidad total inferior a la de algunas capas, etc...) de forma que un módulo de consistencia meteorológica ha de chequear y modificar en su caso la salida antes de ofrecerse al usuario.

En el INM, el autor ha adaptado el programa original a los datos y tipos de observaciones utilizados en España. Las observaciones (en este caso tipo METAR), pueden introducirse manualmente o leerse directamente desde ficheros. Un ejemplo de la salida puede verse en la figura 4. El tiempo de cómputo en un PC no supera en todo caso los 10 segundos por estación.

En este punto, cabe hacerse la pregunta que todo predictor estará pensando: ¿funciona?, es decir, ¿ofrece alguna ventaja o da alguna información útil? En (3) y (4) se hacen algunas comparaciones respecto otras técnicas de predicción estadísticas.

Los resultados para algunas variables de interés aeronáutico, como por ejemplo el techo de nubes y la presencia de nieblas, es satisfactorio.

No obstante, el test realizado en el INM sobre algunas zonas, fundamentalmente el área de Barcelona, GEM no sirve de gran cosa. Para confirmar lo que en principio era una sospecha se realizó una verificación automática y objetiva de las predicciones de GEM para los meses de Julio y Agosto de 1990 en Barcelona.

II

```

II  NN  MM  GGG  EEEEE  M  M
II N  NNM  MMM  G  E  MM  MM  INSTITUTO NACIONAL
II N  NN  M  MM  ==  G  GGG  EEE  M  M  M
II  NNN  MM  G  G  E  M  M  DE METEOROLOGIA
II  NN  MM  GGG  EEEEE  M  M
  
```

Estacion: LEMD Val. 8 h. a partir de 8 Loc. del 17-HAY-90

HORA	VIENTO	VIS	FENOM.	SIG.	N1	H1	N2	H2	N	H	T /TR	PRE.
800	CALMA	10.0			SCD	30	SCD	60	SCD	150	12/ 11	1018
900	CALMA	>13.			SCD	30	SCD	60	SCD	150	14/ 11	1018
1000	CALMA	>13.			SCD	30	SCD	76	SCD	150	16/ 9	1017
1100	CALMA	>13.			SCD	30	SCD	150	BKN	150	18/ 9	1017
1200	WSW-W 10-19	>13.			SCD	30	SCD	158	BKN	150	19/ 9	1016
1300	WSW-W 10-19	>13.			SCD	30	SCD	150	BKN	150	20/ 9	1016
1400	WSW-W 10-19	>13.			SCD	40	BKN	150	BKN	150	21/ 8	1015
1500	WSW-W 10-19	>13.			SCD	40	BKN	151	BKN	150	21/ 8	1015
1600	WNW-NW 10-19	>13.			SCD	40	BKN	100	BKN	100	21/ 8	1015

Figura 4: Ejemplo de salida de GEM adaptado por el INM. Aquí, la entrada es un parte METAR: "LEMD 0000KT 9999 2CUSCO30 2AC070 12/11 1018 NOSIG" del día 17 de Mayo de 1990 a las 8 h.

En la figura 5 se muestra la diferencia entre el valor previsto y observado de la temperatura para algunos plazos de predicción (H+3, H+6 y H+8). Como puede verse, los errores promedios alcanzaron los 4°C por la tarde. GEM parece dar una mayor oscilación diurna que la observada. Esta mayor oscilación diurna también apareció en la depresión del punto de rocío y en la presión, lo cual hizo pensar que podía funcionar mejor en zonas continentales.

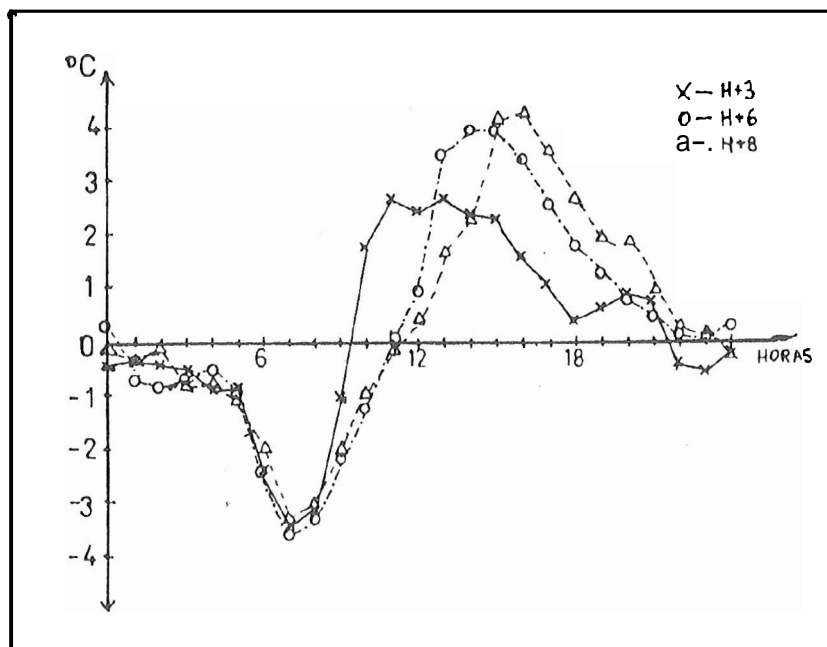


Figura 5: Diferencia entre la temperatura predicha por GEM y el valor observado para distintos plazos de predicción en Barcelona (Aeropuerto del Prat), en Julio de 1990.

Se comprobó también la posible eficiencia de GEM frente a la técnica más utilizada por bastantes predictores a primera hora de la mañana, cuando apenas tienen datos, a la hora de hacer una primera aproximación, por ejemplo, de la evolución de la temperatura: Mirar lo que ocurrió el día anterior. Si llamamos a esto persistencia, su comparación con GEM puede verse en la tabla 2. Tanto el error medio como el error absoluto medio es por lo general peor en el GEM.

LEBL Julio 1990

	09 Z	12 Z	15 Z	18 Z	21 Z	00 Z	03 Z	06 Z
GEM EAM:	2.1	1.5	1.8	1.5	1.3	1.6	1.8	2.0
EM:	-1.3	1.1	1.8	1.4	0.6	-0.1	0.0	-1.4
PER EAM:	1.4	1.4	1.5	1.2	1.6	1.4	1.8	1.6
EM:	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0

LEBL Agosto 1990

	09 Z	12 Z	15 Z	18 Z	21 Z	00 Z	03 Z	06 Z
GEM EAM:	2.9	1.6	1.9	1.2	1.5	1.8	1.9	2.4
EM:	-2.2	-0.2	0.8	0.6	-0.6	-1.4	-1.6	-2.2
PER EAM:	1.9	1.2	1.8	1.3	1.4	1.3	1.4	1.2
EM:	0.2	0.2	0.3	0.1	0.1	0.2	0.2	0.2

Tabla 2. Verificación de la predicción GEM de temperatura para el Aeropuerto del Prat durante julio y agosto de 1990. Se comparan los resultados con los obtenidos con la persistencia, entendiéndose esta como la repetición de lo que ocurrió el día anterior. Para GEM y para la persistencia (PER) se calcula el error absoluto medio (EAM) y el error medio (EM) entre los valores predichos y observados a las horas indicadas. En el caso del GEM, la predicción se realiza con la observación de las 06 Z.

Por comprobar si en una zona más continental funcionaba mejor el GEM, esta misma prueba se hizo para Sevilla (Aeropuerto de San Pablo) y la cosa fue mucho peor, como puede verse en la tabla 3.

Comparación de GEM con persistencia. Temperatura.

LEZL Julio 1990

	09 Z	12 Z	15 Z	18 Z	21 Z	00 Z	03 Z	06 Z
GEM EAM:	4.6	6.3	7.5	7.5	4.2	3.3	1.8	2.0
EM:	-4.6	-6.3	-7.5	-7.4	-4.1	-3.1	-1.3	-1.4
PER EAM:	2.0	2.7	2.1	2.5	1.0	1.3	1.2	1.4
EM:	0.0	-0.4	-0.1	0.0	-0.2	-0.2	-0.2	-0.2

LEZL Agosto 1990

	09 Z	12 Z	15 Z	18 Z	21 Z	00 Z	03 Z	06 Z
GEM EAM:	4.8	5.8	7.8	7.5	4.6	3.4	2.6	2.2
EM:	-4.5	-5.8	-7.8	-7.6	-4.6	-3.3	-2.6	-1.9
PER EAM:	2.6	3.1	2.5	2.3	1.9	1.9	1.2	0.9
EM:	0.4	0.1	0.2	0.4	0.4	0.2	0.1	0.2

Tabla 3. Verificación de la predicción GEM de temperatura para el Aeropuerto del San Pablo (Sevilla) durante julio y agosto de 1990. Se comparan los resultados con los obtenidos con la persistencia, entendiéndose ésta como la repetición de lo que ocurrió el día anterior. Para GEM y para la persistencia (PER) se calcula el error absoluto medio (EAM) y el error medio (EM) entre los valores predichos y observados a las horas indicadas. En el caso del GEM, la predicción se realiza con la observación de las 06 Z.

En cuanto a variables de otro tipo, como el pronóstico de nieblas o tormentas, se calculó el Brier Score y el Heide Skill Score (véanse (2) y (3)) de GEM y de persistencia, entendiendo esta como que los fenómenos significativos persisten inalterados desde el momento de la observación (tabla 4). A pesar de que este tipo de persistencia no es muy lógico, los resultados sólo son algo mejores para GEM.

Brier Score de GEH y persistencia en la predicción de nieblas.

		LEBL Julio 1990							
		09 Z	12 Z	15 Z	18 Z	21 Z	00 Z	03 Z	06 Z
BS	GEM:	.0345	.0345	.0000	.0000	.0357	.2143	.1429	.0370
	PER:	.0690	.0690	.0357	.0370	.0714	.2500	.1786	.0714
HSS	GEM:	.000	.000	---	---	.000	.000	.000	.000
	PER:	-.036	-.036	.	0	-.037	-.065	-.061	-.038

Tabla 4. Brier Score (BS) y Heide Skill Score (HSS) calculados para el pronóstico de niebla en el Aeropuerto de Barcelona en julio de 1990. El Brier score es tanto peor cuanto más alto. Por el contrario, el Heide Skill Score es peor cuanto más bajo. Aquí se comparan la efectividad del pronóstico GEM basado en la observación de las 06 Z y la persistencia, entendida como la continuación de la situación existente en la hora de observación.

Todas estas pruebas demuestran que, al menos para los Juegos Olímpicos de Barcelona, el GEM debe incluir algo que represente la climatología de la zona. A este efecto, TDL ha desarrollado otras ecuaciones basadas en las anomalías de las observaciones respecto de la media climatológica de cada observatorio, lo cual se traduce en unos términos independientes diferentes para cada observatorio y unas ecuaciones que ciertamente no difieren mucho de las generales. Los resultados que ha dado este ajuste climatológico en algunas estaciones de EEUU puede verse en (1) y, desde luego, las mejoras son mínimas en algunas variables.

La otra solución, más drástica, es desarrollar unas ecuaciones de validez local basadas en observaciones locales. Con este método se pierde la generalidad pretendida en GEM. Los resultados de ensayos realizados por TDL tampoco ofrecen frutos demasiado satisfactorios (ver (1)).

Una de las razones en la que estriba la relativa poca eficiencia de GEM es, creo, la poca representatividad que se le da al ciclo estacional. Como se puede ver en la tabla 1, el ciclo estacional tan sólo queda representado con la inclusión del mes y de la variable invierno/verano. Quizás fuese más adecuado que las ecuaciones fueran distintas para cada periodo definido del año.

ADAPTACION DE GEM A BARCELONA

Siguiendo la idea de desarrollar ecuaciones para una determinada época del año y para que tengan una fuerte componente local, se ha optado por obtener la matriz de predicción con datos del Aeropuerto del Prat durante los meses de Julio y Agosto.

Los datos utilizados provienen de los partes metar, algunos de ellos completados con synops, a partir de 1979 y durante los meses referidos anteriormente.

El esquema adoptado debe tener la facilidad de extrapolación de las predicciones a otras sedes olímpicas a partir de la predicción base del Prat. Para que las observaciones tengan más generalidad, se puede proceder previamente a la adimensionalización de las variables numéricas, de forma que en vez de trabajar con cada variable se trate con valores percentiles acumulados en la distribución de probabilidades de cada variable en cada sede. Este procedimiento, no obstante, incluye una cantidad de cálculo considerable y además puede hacer perder algunas significaciones meteorológicas implícitas en los valores absolutos de las variables (por ejemplo, la temperatura de 0°C).

En la figura 6 se muestra un esquema del proceso. Los datos pueden provenir, bien de la estación principal (El Prat) o de estaciones remotas (como por ejemplo de las estaciones

automáticas de cada sede olímpica). Para poder utilizar las observaciones remotas es necesario el establecimiento de funciones de transformación que relacionen los datos de estas estaciones con las de la principal, y también sus transformaciones inversas de forma que una predicción para el Prat se pueda extrapolar a cada sede. El núcleo de la predicción no difiere del GEM salvo en las ecuaciones que son mucho más restringidas a un lugar y época determinada.

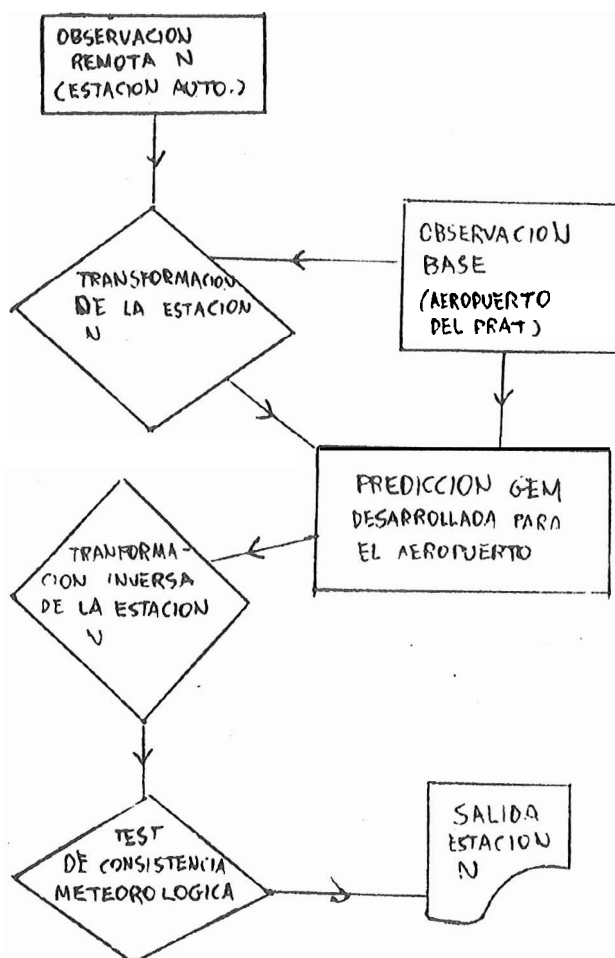


Figura 6: Esquema general de la predicción tipo GEM adaptada a Barcelona.

Las hipótesis en las que se basa este esquema de predicción son las siguientes:

a) Las ecuaciones de predicción, tras la transformación para cada sede, deben ser muy similares en todas las sedes cercanas, es decir, el tiempo evoluciona de igual forma en todas ellas salvo pequeñas diferencias las locales tenidas en cuenta en las funciones de transformación.

b) Asimismo, la existencia de funciones de transformación directas e inversas requiere que dos tipos de flujos de viento y tiempo característico esencialmente distintos no corresponda a un mismo tipo de tiempo en el otro observatorio.

c) Las funciones de transformación deben representar aceptablemente bien las diferencias locales, de forma que las ecuaciones de regresión ofrezcan coeficientes de correlación suficientemente representativos.

Todas estas condiciones se verifican razonablemente si la distancia que separan ambas sedes son pequeñas y la orografía y condiciones locales no imponen transiciones microclimáticas drásticas.

En los próximos meses se espera tener los primeros resultados de este nuevo tipo de predicción, que proponemos se denomine MEM (Modelo Equivalente Markoviano). Se tienen esperanzas de que, al menos en la evolución normal de las variables en días no muy problemáticos, se obtengan resultados más satisfactorios.

CONCLUSIONES

Si bien este artículo es esencialmente divulgativo, hay que resaltar algunas ideas respecto el GEM a modo de conclusión.

1) GEM es una predicción estadística en la que sólo intervienen datos de superficie y nubosidad de un observatorio en un instante dado. Las mismas ecuaciones sirven para cualquier observatorio en cualquier época del año.

2) La generalidad y simplicidad que supone el punto anterior es su gran limitación. Se ha comprobado que en distintos puntos de la Península Ibérica GEM no sirve de mucho si no se plasma de alguna forma la climatología de cada lugar.

3) A pesar de todo, es un método rápido y accesible de obtener una primera aproximación de la evolución meteorológica en cualquier lugar de la zona templada del Hemisferio Norte.

REFERENCIAS BIBLIOGRAFICAS

GEM. A statistical weather forecasting procedure. Robert G. Miller. Short and Medium Range Weather Prediction Research OMM. Publication series N.º 10. 1981.

Survey of common verification methods in meteorology. Henry R. Stanski, Laurence J. Wilson, William R. Burrows. Technical report nº 8. OMM N.º 358. 1989.

Comparative verification between GEM and the official aviation terminal forecast (FT). R.G. Miller and T.J. Perrone. TDL office note 85-8, National Weather Service, NOAA, U.S. Department of Commerce. 1985.

A comparative verification of GEM and MOS. T.J. Perrone and R.G. Miller. TDL office note 83-10, National Weather Service, NOAA, U.S. Department of Commerce. 1983.