

## ESTIMACION DE LA COSECHA DE TRIGO DE LA PROVINCIA DE VALLADOLID

por Juan Maynar Ferrero

### Resumen

El presente trabajo desarrolla una estimación del rendimiento potencial del cultivo de trigo de secano para la provincia de Valladolid, en un período que va del año 1968 al 1982. Dicha estimación se realiza a partir de variables meteorológicas a determinar, y mediante un modelo uniecuacional de regresión múltiple, Mínimos Cuadrados Ordinarios; este modelo nos permitirá, además, predecir el rendimiento potencial para el año 1983.

### Introducción

En España la producción anual de trigo constituye un 5,2 % de la producción total agraria (6), siendo la más alta de la de todos los productos agrícolas. Esto hace que el conocimiento del rendimiento potencial de este cultivo sea de vital importancia tanto a nivel macroeconómico como microeconómico; podría ser, además, útil en la creación de una apropiada planificación, tanto en el transporte, como en el uso de la maquinaria agrícola, almacenamiento, análisis de las necesidades de mano de obra, política de precios, etc.

La tendencia al aumento en el rendimiento observada en los últimos tiempos se debe a la mejora habida en la tecnología agrícola: maquinaria, abonos químicos, nuevas variedades de semilla, etc. Ahora bien, año a año, el rendimiento de un cultivo depende, principalmente, de las condiciones climáticas.

La estimación del rendimiento potencial de un cultivo de secano, en una región determinada, en función de la climatología, puede hacerse conociendo las producciones reales, y la superficie sembrada a lo largo de un período de tiempo, además de los datos meteorológicos pertinentes.

En un estudio realizado por Micevic (9), para un cultivo de trigo, bajo unas condiciones determinadas en cuanto a período de siembra, abonado, variedad de semilla, etc., se observó que la influencia de la climatología en el rendimiento del trigo era muy fuerte, presentando un coeficiente de correlación  $R$  alrededor de 0,90. Esto es cierto para las regiones áridas y semiáridas de la Tierra, donde la humedad es el factor limitante más importante de la producción de una cosecha de secano, pero a medida que aumenta la humedad disponible del suelo, se van haciendo necesarios métodos y técnicas más complejas para relacionar las variables meteorológicas con el rendimiento de un cultivo.

El presente trabajo consiste en la estimación del rendimiento potencial del cultivo de trigo, de la provincia de Valladolid, para los años del período 1968/1982; el método de cálculo es el de Mínimos Cuadrados ordinarios.

Un modelo rendimiento-clima, debe incluir las variables meteorológicas básicas que afectan al crecimiento y desarrollo del cultivo; dichas variables deben estar disponibles por las diferentes estaciones.

Las variables utilizadas en el presente estudio son: precipitación, número de días de heladas y temperatura máxima absoluta.

La precipitación es el elemento climatológico más importante para la cosecha. El máximo beneficio de la precipitación se obtiene cuando la cantidad es la adecuada, tanto el defecto como el exceso incide negativamente en el cultivo, así como cuando su distribución a lo largo del año es la correcta, de tal forma que no falte en las estaciones cruciales para la planta.

Cada tipo de cultivo tiene una serie de épocas críticas durante las cuales la planta presenta una mayor sensibilidad a la falta de agua. Estas épocas críticas se dan durante la nascencia de la planta (octubre), el encañado (marzo y abril) y la floración y fecundación (mayo).

Para obtener unos buenos rendimientos, es necesario que, durante estos períodos críticos, la evaporación real esté próxima a la evaporación potencial. Por el contrario, en otras épocas, pueden existir diferencias notables entre ambos tipos de evapotranspiración sin que afecte sensiblemente a la planta.

Williams (3) en sus modelos ha empleado, además de la precipitación (P), la evapotranspiración potencial (E); la evapotranspiración potencial es un índice de evaporatividad. El índice más conocido es el de Thornthwaite, que expresa los efectos combinados de la evaporación del suelo y de la transpiración de las plantas. Se define como la cantidad de agua que perderá una superficie completamente cubierta de vegetación en crecimiento activo, si en todo momento existe en el suelo humedad suficiente para su utilización máxima por las plantas (2). Una combinación de P y E, como  $P-E$  o  $P/E$ , puede ser un buen indicador del balance hídrico de la planta. Las ecuaciones en las que aparece la E además de la P, han dado mejores resultados que aquellas basadas únicamente en la P (3); esta variable, es decir E, ha sido finalmente desestimada en nuestro modelo ya que, tratándose la provincia de Valladolid de una región semiárida, la variable P resulta suficiente (4).

El número de días de helada considera aquellos en los que la temperatura mínima ha sido igual o inferior a cero grados centígrados (1). Un dato de

importancia en agricultura es el de las heladas tardías, entendiéndose por tales a aquellas que ocurren después del 1 de abril, ya que éstas ocurren durante el encañado de los cereales, incidiendo negativamente en la cosecha, constituyendo, por tanto, un factor limitante.

La temperatura máxima absoluta es la tercera variable considerada en nuestro estudio. Una entrada de aire caliente y seco, procedente del continente africano, o un tiempo soleado y seco prolongado, si se produce durante la fecundación (mayo) o durante el cuajado del fruto (junio, julio y agosto), produce en la cosecha una merma considerable y constituye, por lo tanto, otro factor limitante.

Es de esperar que la temperatura máxima absoluta en ese período sea una variable adecuada a tener en cuenta. Aunque hubiera sido preferible utilizar la humedad relativa en ese período, esto no ha sido posible al no haber podido obtener dicha variable; hay que tener en cuenta, no obstante, que para la provincia de Valladolid se ha obtenido un coeficiente de correlación entre la temperatura máxima absoluta media mensual, y la humedad relativa media mensual del período 1941/1960 de 0,905, ya que dicha provincia está dentro de una región semiárida.

### El clima de Valladolid

La provincia de Valladolid está situada en la Meseta Norte, a unos 700 metros sobre el nivel del mar, y a una latitud, para la capital de  $41^{\circ}39'N$ .

Font Tullot en su "Climatología de España y Portugal" (1), clasifica a la provincia de Valladolid dentro de la Subregión Continental Extremada, caracterizada por un frío y largo invierno, y un verano cálido. Las heladas son muy frecuentes y pueden presentarse desde finales de octubre hasta primeros de mayo. Hay relativamente pocos días de escarcha y rocío en el año, consecuencia de la insuficiente humedad del aire.

Debido al acusado aislamiento orográfico de la Meseta Norte, las precipitaciones totales anuales

no alcanzan los 400 mm. La estación de máxima pluviosidad es la primavera. La Meseta Norte está considerada como una zona semiárida. Las zonas semiáridas de la Península Ibérica superan más de la mitad de la superficie peninsular.

Comparando las medias mensuales de precipitación, temperatura máxima absoluta, y temperatura mínima absoluta del período 1931/1960 (\*) con las del período de nuestro estudio 1968/1982, se observan las siguientes características:

1) La precipitación media anual para el período 68/82 fue de 460 mm., mientras que para el período 31/60 fue de 362,3 mm. (gráfica 1). La correlación entre estos dos períodos es de 0,695.

PRECIPITACION MEDIA MENSUAL

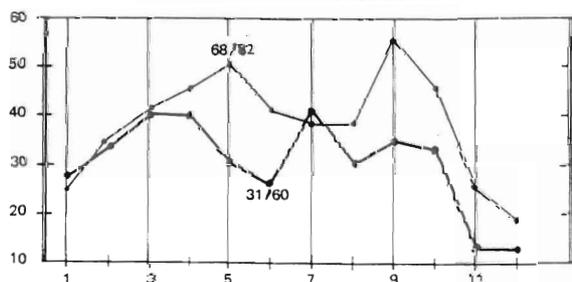


Gráfico 1.

2) La temperatura máxima absoluta anual para el período 68/82 fue de 13,5°, 3,4° inferior a la que se dio en el período 31/60 (gráfica 2). El coeficiente de correlación entre los dos períodos es de 0,973.

TEMPERATURA MAX. ABS. MEDIA MENSUAL

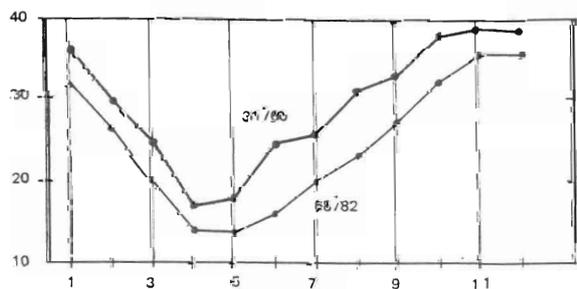


Gráfico 2.

(\*) Los meteorólogos consideran que un período de 30 años es suficiente para definir el clima de una región en las latitudes templadas (1), aunque últimamente esta cuestión ha sido puesta en duda.

3) La temperatura mínima absoluta anual para el período 68/82 fue de -0,74°, 1,3° superior a la correspondiente al período 31/60, que fue de -2,06°. El mínimo invernal para el período 68/82 se produce en diciembre, mientras que en el período 31/60 se producía en febrero. El coeficiente de correlación entre los dos períodos ha sido de 0,926 (gráfica 3).

TEMPERATURA MIN. ABS. MEDIA MENSUAL

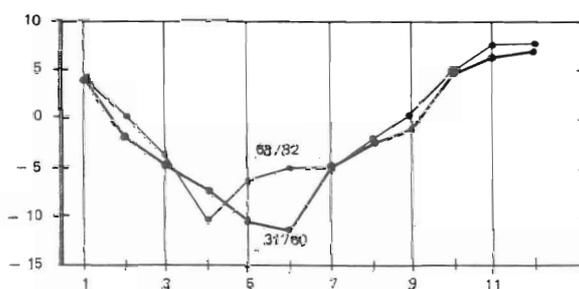


Gráfico 3.

### Resultados

Para cada una de las variables han sido utilizados los valores que presentan en los meses considerados como cruciales en el desarrollo del cultivo. La elección de los meses más representativos para cada variable, se ha llevado a cabo mediante el cálculo de su correlación con el rendimiento, eligiéndose los meses que presentaron las mayores correlaciones, teniendo en cuenta si el signo estaba de acuerdo con el considerado por nosotros, y si estaba dentro del período crucial para el cultivo.

Los resultado para cada variable son los siguientes:

#### Correlación entre rendimiento y precipitación (P)

	Rend.
P (oct.)	0,628
P (feb.+mar.+abr.+may.)	0,413
P (oct.+feb.+mar.+abr.+may.)	0,624

#### Correlación entre rendimiento y nº de días de heladas (H)

	Rend.
H (nov.)	-0,264
H (abr.)	-0,384
H (nov.+abr.)	-0,373

Correlación entre rendimiento y temp. máx. abs.  
(T)

	Rend.
T (may.)	-0,641
T (jun.)	-0,520
T (jul.)	-0,567
T (may.+jun.+jul.)	-0,798

El mejor modelo de regresión múltiple (MCO), ha sido aquel en el que al rendimiento (R) se le ha hecho depender de las siguientes variables:

$$R = f(P_O, P_{FMAM}, H_{NA}, T_{MJJ})$$

Donde  $P_O$  es la precipitación en octubre (gráf. 6),  $P_{FMAM}$  es la suma de la precipitación en febrero, marzo, abril y mayo (gráf. 7),  $H_{NA}$  es la suma del número de días de helada en noviembre y abril (gráf. 8),  $T_{MJJ}$  es la suma de la temperatura máxima absoluta en mayo, junio y julio (gráf. 9).

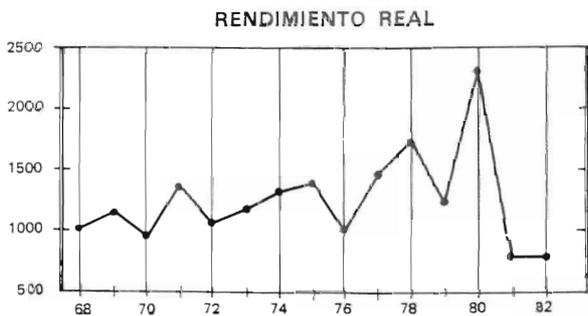


Gráfico 4.



Gráfico 5.

La matriz de correlación entre estas variables y el rendimiento es la siguiente:

	R	$P_O$	$P_{FMAM}$	$H_{NA}$	$T_{MJJ}$
R	1	0,628	0,413	-0,373	-0,798
$P_O$		1	0,119	-0,201	-0,313
$P_{FMAM}$			1	0,368	-0,245
$H_{MA}$				1	0,222
$T_{MJJ}$					1

PRECIPITACION EN OCTUBRE

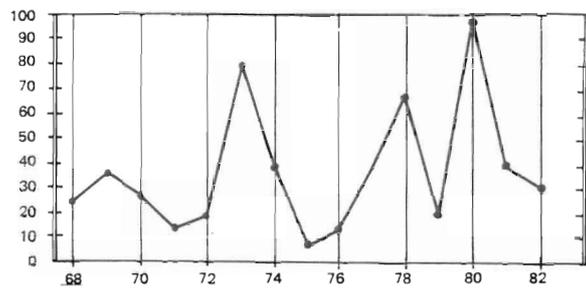


Gráfico 6.

PRECIPITACION EN FEB., MAR., ABR., MAY.

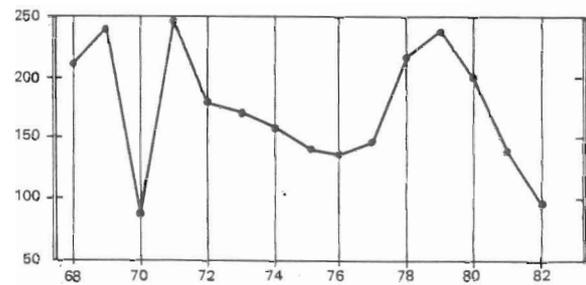


Gráfico 7.

Principales resultados empíricos (ver gráficos 5 y 10)

$$R = 4320,16 + 5,394 P_O + 2,761 P_{FMAM} - (901,85) (1,524) (0,842) - 16,645 H_{NA} - 41,712 T_{MJJ} (5,767) (8,38)$$

Coefficiente de correlación múltiple	0,954
Coefficiente de correlación múltiple corregido	0,936
Suma de residuos al cuadrado	194596
Estadístico F (4,10)	25,822
Estadístico de Durbin Watson	1,471

DIAS DE HELADAS EN NOV. Y ABR.

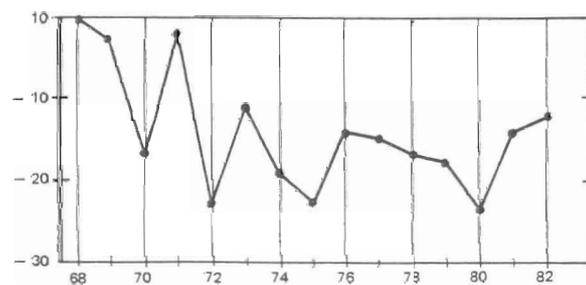


Gráfico 8.

TEMP. MAX. ABS. EN MAY., JUN., JUL.

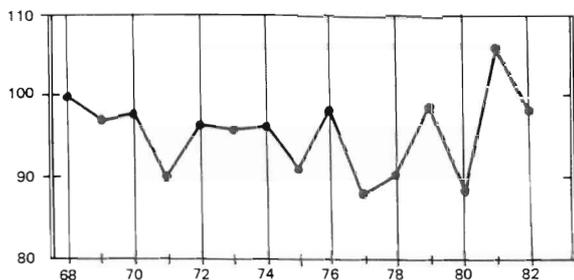


Gráfico 9.

RENDIMIENTO REAL Y RESIDUOS

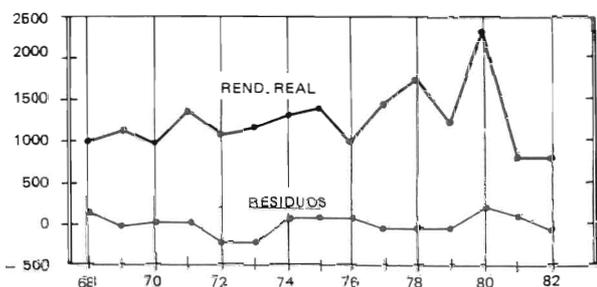


Gráfico 10.

a) Un parámetro podrá considerarse como estadísticamente significativo, a niveles de confianza del 95 %, si el estimador tiene un valor que supere dos veces su desviación típica. En nuestro caso, se pueden aceptar los cuatro parámetros a un nivel de confianza del 95 %:

Variable	Desv. Tip.	Param.
$P_O$	$1,524 \times 2 = 3,048 <$	$5,394$
$P_{FMAM}$	$0,842 \times 2 = 1,684 <$	$2,761$
$N_{NA}$	$5,767 \times 2 = 11,534 <$	$16,645$
$T_{MJJ}$	$8,380 \times 2 = 16,760 <$	$41,712$

b) Cuando contrastamos la hipótesis nula con la F, lo que estamos contrastando es la hipótesis de nulidad de todos los parámetros, y su rechazo no implica que todos los parámetros sean estadísticamente significativos, sino que al menos un parámetro no puede considerarse nulo al nivel de significación elegido. En nuestro caso, se puede rechazar la hipótesis nula al nivel de confianza del 95 %:

$$F_{0,95}(4, 10) = 3,48 < 25,82 = F(4, 10)$$

c) Los coeficientes de determinación parcial se interpretan como la proporción de la variable independiente no especificada por el resto de las

variables del modelo, que explica directamente la variable dependiente X, para la cual se ha calculado dicho coeficiente de correlación parcial. En nuestro caso los resultados obtenidos son los siguientes:

$r_{R,P_O}$	$= 0,525$
$r_{R,P_{FMAM}}$	$= 0,491$
$r_{R,N_{NA}}$	$= 0,429$
$r_{R,T_{MJJ}}$	$= 0,673$

d) El contraste F parcial puede interpretarse como la diferencia relativa entre los residuos de la regresión con  $4-m$  variables explicativas, y con la totalidad de las variables  $K = 4$ . Si la diferencia es grande, significa que las  $m$  variables juegan un importante papel en el modelo, ya que disminuyen significativamente los residuos, con lo que la hipótesis de nulidad debe ser rechazada. Los resultados obtenidos para la F-parcial son los siguientes:

$F(1, 10) P_O$	$= 12,53$
$F(1, 10) P_{FMAM}$	$= 10,75$
$F(2, 10) P_O P_{FMAM}$	$= 26,55$
$F(1, 10) N_{NA}$	$= 8,33$
$F(1, 10) T_{MJJ}$	$= 24,77$
	$(F_{0,95}(1, 10) = 4,96)$
	$(F_{0,95}(2, 10) = 4,10)$

e) El modelo presenta únicamente un error de tipo I, es decir, un cambio falsamente predicho para el año 73/74 (ver gráfico 5).

f) Un estudio visual del gráfico de los residuos (gráfico 11), nos indica la inexistencia de he-

RESIDUOS

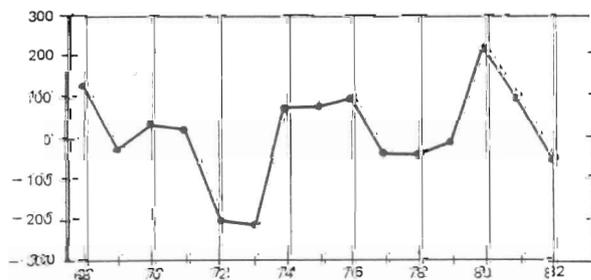


Gráfico 11.

terocedasticidad, ya que muestra un comportamiento sin cambio aparente en su dispersión a lo largo del período muestral. Sin embargo, en el gráfico 4, se observa que las fluctuaciones de la se-

rie son de mayor amplitud según avanzamos en el tiempo, hecho que nos indica la presencia de heterocedasticidad. Estas fluctuaciones de mayor amplitud no son debidas enteramente a la variabilidad de las variables meteorológicas, sino más bien al uso de fertilizantes químicos, que en los años climatológicamente buenos hacen que se alcancen rendimientos muy altos.

g) El contraste de Shapiro-Wilk, contrasta la hipótesis de normalidad en la distribución de residuos. Para nuestro modelo  $W = 0,889$ , que es mayor que 0,881 para un nivel de confianza del 95 %, con lo que puede mantenerse la hipótesis de normalidad de los residuos.

h) Theil, establece el coeficiente de la desigualdad  $U$ , que nos permite analizar la bondad de la predicción. Si hubiera una predicción perfecta  $U = 0$ , y en caso de máxima desigualdad  $U = 1$ . Para nuestro modelo  $U = 0,177$ .

Theil descompone  $U$  en tres sumandos  $U_m$ ,  $U_s$  y  $U_c$  con  $U^2 = U_m^2 + U_s^2 + U_c^2$  donde  $U_m$  es el error sistemático: Si  $0 < U_m < 1$  entonces hay supervaloración de los cambios, si  $-1 < U_m < 0$ , hay infravaloración de los cambios, y si  $U_m = 0$  el error sistemático es nulo.  $U_s$  es el error de dispersión, y se anula si la dispersión en la predicción y en los valores reales es la misma.  $U_c$  representa el error de correlación, y nos indica la mayor o menor correlación entre los valores de predicción y los reales. Para un error de correlación nulo  $U_c = 0$  ( $r = 1$ ).

Para nuestro modelo tenemos los siguientes valores:

$$\begin{aligned} U_m &= 0,0285 \\ U_s &= 0,0014 \\ U_c &= 0,1808 \\ (R &= 0,954) \end{aligned}$$

i) Si observamos el gráfico de los residuos (gráfico 11), se observa un comportamiento sistemático cada tres años para los que se mantiene un error del mismo signo. Parece fácil admitir la existencia de autocorrelación a lo largo del período de estimación del modelo, ya que sólo una total alea-

toriedad en la distribución de los errores corresponde a la idea de ausencia de autocorrelación. El resultado parece confirmarse con el estadístico de Durbin-Watson, cuyo valor calculado es 1,471 que está situado en la zona de duda para un nivel de confianza del 95 %. Para un nivel de confianza del 99 % estaríamos dentro de la zona de ausencia de autocorrelación de primer orden.

### Discusión

Como se observa en el gráfico 5, la curva del rendimiento estimado se ajusta bien a la curva del rendimiento real, salvo para los años 72 y 73, que dan un rendimiento estimado mayor que el observado en un 20,1 % para el año 72, y un 19,3 % para el año 73. Esto puede ser debido a que se ha dejado fuera del modelo alguna variable explicativa. También se observa en el gráfico 11 de los residuos, que para los años 72 y 73, toman valores altos consecutivos, que además son los mayores del modelo.

En un modelo MCO con un período muestral de 11 años, 1972 a 1982, y en el que se utilizaron las variables meteorológicas  $P_{\text{oct.}+\text{nov.}+\text{mar.}+\text{abr.}+\text{may.}}$ ,  $H_{\text{nov.}}$ ,  $T_{\text{dic.}+\text{ene.}}$ ,  $T_{\text{may.}+\text{jun.}+\text{jul.}}$  y  $T_{\text{min. mar.}+\text{abr.}}$ , la curva del rendimiento estimado se ajustaba muy bien a la curva del rendimiento real, incluso en los años 72 y 73. Los principales resultados empíricos fueron:

Un coeficiente de correlación múltiple corregido de 0,959. Un estadístico F (5,5) de 24,240. Una suma de residuos al cuadrado de 78797. Un estadístico de Durbin-Watson de 2,683.

Cuando graficamos el rendimiento real desde el año 61 hasta el año 82, se observa una tendencia suave positiva (7 %). Además, el rendimiento a medida que avanzamos en el tiempo fluctúa con una amplitud mayor. Esta fluctuación, cada vez mayor en amplitud, creemos que es debida a las nuevas técnicas agrícolas introducidas en la agricultura a partir de finales de la década de los 60, principalmente al uso de fertilizantes químicos, ya que pensamos que no es debido a las fluctuaciones de las variables climatológicas utilizadas en el modelo (ver gráfs. 6, 7, 8 y 9).

Una aplicación interesante de un modelo como el estudiado, sería la evaluación de los efectos de las plagas y epidemias en el cultivo, al poder contrastar el rendimiento real con el estimado. Williams (3) al emplear su modelo de estimación del rendimiento de la cosecha de trigo para una región de Canadá para el año 1954, en la que hubo una importante epidemia de roya, observó que las pérdidas debidas a esta epidemia se elevaban al 60 % de la cosecha estimada por su modelo.

Otra aplicación sería la del análisis de la producción potencial de un cultivo de secano si las condiciones climáticas a largo plazo se modificasen. La simulación nos daría los efectos del cambio climático en la producción del cultivo.

Con un modelo general para cada cultivo de secano, podría, también evaluarse la totalidad de la producción agrícola de secano, y de ella estimar la producción total agrícola.

## Conclusiones

Como ya se ha indicado, sería deseable el realizar un estudio más general, que comprendiese todas las provincias, y que abarcase la totalidad de los productos agrícolas de secano.

Un proyecto como éste, tendrá que tener en cuenta el cambio que se da en la agricultura española a partir de los años 60 hasta la fecha: la mecanización y el uso de fertilizantes químicos. Este proceso no es gradual y varía de una región a otra. Tendrían que estudiarse, además, cuáles son los períodos cruciales para cada producto, de esta for-

ma se podría determinar qué variables meteorológicas afectan a cada tipo de cultivo.

## Bibliografía

- (1) FONT TULLOT, I. (1983): Climatología de España y Portugal. Inst. Nac. de Meteorología (1983).
- (2) CASTILLO, E., del GIMÉNEZ ORTIZ (1965): Evapotranspiración potencial y balances de agua en España. Ministerio de Agricultura (1965).
- (3) WILLIAMS, G. D. V. (1973): Estimates of prairie provincial wheat yields based on precipitation and potential evapotranspiration. Canada Journal Plant Science, 53: 17-33.
- (4) GEO, W., ROBERTSON, A. (1974): A factorial yield-weather model for wheat. The Canadian Wheat Board, Winnipeg, Canada (1974).
- (5) PULIDO, E. (1983): Modelos Econométricos. Ed. Pirámide (1983).
- (6) RODRÍGUEZ, J., SANZ, R.: Trimestalización del P.I.B. por ramas de actividad. Servicio de Estudios del Banco de España.
- (7) JOHNSTON, J. (1975): Métodos de Econometría. Ed. Vicens Universidad (1975).
- (8) BAIER, W., WILLIAMS, G. D. V. (1973): Regional wheat yield predictions weather data in Canada. Agrometeorology Research Inst. Agriculture Canada, Ottawa.
- (9) MICEVIC, Z. S. (1973): On early prognosis of wheat yield in the production region of Bosnia. Republic Hidrometeorological Inst. Sarajevo, Yugoslavia, 1973.

Datos recogidos de:

Anuario de Estadística Agraria. Ministerio de Agricultura, 1960-1982.

Calendarios Fenológicos y Meteorológicos. I.N.M. 1960-1983.

Programas utilizados: VISICALC, VISITREND, VISIPILOT.

Ordenador: APPLE II/e, 64K.