

SISTEMA INFORMÁTICO PARA LA GENERACIÓN DE DATOS CLIMÁTICOS Y FITOCLIMÁTICOS

¹E. Manrique Menéndez y ²A. Fernández Cancio

¹ E.U.I.T. Forestal, U.P.M., Ciudad Universitaria s/n. 28040 Madrid.
e-mail: emilio.manrique@upm.es Tfno: 91 3367657

² CIFOR-INIA. Ctra. de la Coruña s/n. km 7. 28040 Madrid.
e-mail: fernand@inia.es Tfno: 91 3476778

Resumen

En este trabajo se presenta un sistema informático para la generación de datos climáticos de cualquier punto del país, conocidas sus coordenadas y su altitud. Para ello, se parte de una base de datos de estaciones meteorológicas reales y se halla la que está más próxima al punto problema, que se convierte en el centro de círculos de radio creciente. Se consideran además en cada círculo cuatro cuadrantes definidos por las líneas perpendiculares WE y SN. El sistema genera los datos climáticos de la estación real, mes a mes, a partir de las estaciones que están dentro de cada círculo y de cada cuadrante. Los valores se calculan de dos formas: mediante modelos de regresión, relacionando la variable climática con la altitud o a partir de valores medios de las estaciones reales. Dado que los datos de la estación próxima son conocidos, la mejor aproximación entre el dato generado y el dato real sirve para realizar la reconstrucción de las variables climáticas del punto problema, por el mismo método. Los programas calculan también diversas variables fitoclimáticas, que relacionan el clima y los tipos de vegetación existentes en la zona.

Palabras clave: Temperaturas, precipitaciones, variables fitoclimáticas, modelos de regresión.

INTRODUCCIÓN

Los programas GENPT (FERNÁNDEZ-CANCIO y MANRIQUE, 2001) sirven para la generación de datos climáticos de cualquier lugar del país, conocidas sus coordenadas y su altitud. Para ello, se parte de una base de datos de estaciones meteorológicas reales, térmicas y pluviométricas, de las que se conoce su localización y su altitud, así como los datos mensuales de precipitaciones y temperaturas medias, máximas y mínimas absolutas y medias de las máximas y de las mínimas.

El proceso comienza hallando la estación real más próxima al punto problema. Ésta se convierte en el centro de una serie de círculos de radio creciente, desde aproximadamente 20 km hasta 100 km, aunque se puede ampliar este límite en zonas con pocas estaciones. Se consideran además en cada círculo cuatro cuadrantes, los definidos por las líneas perpendiculares WE y SN, que se cortan en la estación de referencia. Con todas estas posibilidades se realiza la reconstrucción climática de la estación real, mes a mes, a partir de las estaciones que están dentro de cada círculo y de cada cuadrante. La reconstrucción se realiza de dos formas:

- Mediante un modelo de regresión, relacionando la variable climática con la altitud.
- Hallando el valor medio de los datos conocidos de todas las estaciones.

Dado que los datos de la estación próxima son conocidos, la mejor aproximación entre el dato reconstruido y el real, que corresponde a un cuadrante, a un radio y a uno de los dos métodos, sirve para realizar la reconstrucción del dato desconocido del punto problema. Puesto que se han considerado diversas posibilidades de distancia y orientación para mejorar la elección, el error cometido suele ser bastante aceptable, inferior al 5%.

Para reducir el efecto negativo que pudieran tener estaciones con muy pocos datos, o que sean claramente inhomogéneas con su entorno, se realiza un análisis estadístico de la variabilidad, a fin de no considerar los valores que pudieran distorsionar los resultados. Con este mismo objetivo se puede realizar *a posteriori* un análisis de valores fuera de rango, e incluso realizar el corte de los datos que se aparten excesivamente del valor medio.

Los programas no se limitan al cálculo de valores mensuales de precipitaciones y temperaturas. También calculan las variables fitoclimáticas, que relacionan el clima y los tipos de

vegetación existentes en la zona, como la aridez de Gaussen y diversos parámetros ombrotérmicos relacionados con los diagramas de Walter-Lieth (ALLUÉ-ANDRADE, 1990). En todos los casos, dado que se dispone de los datos reales y estimados de la estación más próxima al punto problema, se realiza la correspondiente estimación del error.

En determinadas zonas del país, la escasez de datos reales disponibles da lugar a la generación de archivos climáticos con importantes lagunas en algunos períodos. Los subprogramas COMPLET se encargan reducir o eliminar estos huecos, utilizando criterios de regresión cuando es posible o medias cuando no lo es, obteniéndose así series de datos muy completas.

GENERACIÓN DE DATOS TÉRMICOS Y PLUVIOMÉTRICOS

Determinación de estaciones en el entorno del punto problema

El proceso comienza generando un archivo que contiene la información más relevante sobre las estaciones meteorológicas que se van a utilizar para la generación de los datos climáticos del punto problema, caracterizado por sus coordenadas geográficas y su altitud.

La estación que denominamos de referencia es la más próxima al punto, siempre que cumpla unos mínimos de calidad en sus datos (12 años completos para precipitaciones y 8 para temperaturas). El programa genera un archivo en el que aparecen los datos del punto (coordenadas y altitud), los de la estación de referencia y los de entornos circulares de ésta, ordenadas por distancias. La distancia máxima considerada puede ser la que se quiera, aunque no conviene alejarse excesivamente para mantener la idea de aproximación local.

La estación de referencia se sitúa en el centro de un círculo, respecto al cual se pueden considerar 4 cuadrantes, los definidos por las líneas perpendiculares WE y SN, que se cortan en la estación y servirán para situar las demás estaciones (figura 1). Existe una variante en la aplicación del programa, en el que se fuerza la elección de la estación de referencia a una concreta cuando ello se considere oportuno.

Como se comentará posteriormente, se van a utilizar dos métodos de generación, por regresión y por medias. En este último caso, debido a que los datos deben ser homogéneos, se impone la condición de que la diferencia de altitud entre las estaciones del entorno y el punto no supere los 100 m.

Análisis de la variabilidad de los datos

Es posible que no todas las estaciones del entorno sean válidas para su utilización, por presentar datos anómalos en alguna variable o intervalo temporal. El programa trata de detectar esta circunstancia, calculando las desviaciones típicas en cada estación, para cada variable y cada mes. Se obtiene así una matriz en la que las filas corresponden a las estaciones y las columnas a las variables mensuales, es decir 12 columnas para las precipitaciones y 12 x 5 columnas para las temperaturas (máximas/mínimas, media de máximas/mínimas y media mensual), siendo los elementos de la matriz las correspondientes desviaciones típicas. Para cada columna se halla el valor medio y el error cuadrático medio de las desviaciones típicas. Sólo se consideran válidas, para la reconstrucción de cada variable mensual, las estaciones cuya desviación típica esté comprendida en un intervalo de confianza del 95%. Para su manejo posterior, se asigna a cada estación y a cada variable mensual un dígito, 1 si es válida y 0 si no lo es.

Generación de datos climáticos

El proceso continúa con la lectura de la información almacenada en varios archivos por los programas anteriores. Así, se reconoce la estación de referencia, que es la primera de la lista para la que existe dato de la variable, el mes y el año, y con varianza homogénea.

Existen dos formas de tratar las temperaturas extremas C y F: se puede optar por que sean los de la estación de referencia, modificados por la altitud (± 0.6 °C cada 100 m) o que se estimen como el resto de las variables. Esto es debido a que, por sus peculiaridades, estos datos frecuentemente no se pueden estimar correctamente mediante modelos lineales.

El programa calcula los datos climáticos de dos formas:

- A partir de regresiones, en las que la variable independiente es la altitud.
- Tomando el valor medio de los datos climáticos del entorno.

Se elige el mejor sistema en función de la mejor aproximación en la reconstrucción de la estación de referencia.

El programa desarrolla varios bucles anidados, en función de que se esté considerando:

- Medias o Regresiones: No se consideran estimaciones si no hay un número mínimo de datos.
- Cuadrantes: Entre 0 y 4. El cuadrante 0 incluye los 4 cuadrantes restantes.
- Distancias a la estación de referencia: Se consideran diversos radios: $0,2^\circ = 27$ km; $0,3^\circ = 40$ km $0,4^\circ = 54$ km $0,5^\circ = 68$ km $0,6^\circ = 81$ km, etc.
- Años extremos del intervalo temporal.
- Meses, entre 1 y 12.
- Variables (1 para precipitaciones y 5 para temperaturas).

Para cada variable, cada mes y cada año, realiza $2 \times 5 \times N$ estimaciones (Media/Regresión $\times 5$ cuadrantes $\times N$ distancias). Sólo se conserva en los archivos la estimación que proviene de la mejor distancia, es decir $2 \times 5 \times 1$ estimaciones. En una primera etapa, se reconstruye sólo la estación de referencia, a partir de las de su entorno. Como de ella se conocen los datos reales y cada una de las $2 \times 5 \times 1$ estimaciones, se determina la mejor, que corresponde al valor mínimo de de la diferencia entre el valor real y el valor estimado, en valor absoluto. Una vez conocido el mejor método, el mejor cuadrante y la mejor distancia, se elige y se aplica ese método de generación para el punto cuyos datos climáticos se desconocen.

Cuando no se alcanzan los valores límite del número de datos en el proceso de calibración, no se realiza la estimación del dato, pues no se podría conocer el error cometido. Al no disponer de una esta información, esencial, se interrumpe el proceso.

Errores cometidos al generar los datos

Para el cálculo del error se considera una nueva variable, la diferencia entre el valor real y el valor estimado, a partir de los datos de la estación de referencia. Se determinan las medias y los errores cuadráticos medios, mes a mes de esta nueva variable. Esto se hace para las 3 variables térmicas media de las mínimas, media de las máximas y temperatura media mensual (se excluyen las temperaturas extremas) y para las precipitaciones. Además se generan también las variables anuales T media y P total; con estas variables se procede del mismo modo, pero año a año.

Todas estas variables se distribuirán en torno a 0 (son diferencias entre valores reales y estimados), con valores medios positivos (subestimación) o negativos (sobrestimación), y con un error cuadrático medio que se toma con el 95% de significación. Al no ser la media necesariamente cero, los intervalos serán asimétricos. En la información que queda grabada en los archivos aparece siempre la media y el error: el signo de la media da información sobre la subestimación (+) o sobrestimación (-); el error puede tomarse en el caso más desfavorable, sumando los valores absolutos de la media y el error.

Los errores de TMF y TMC (temperaturas medias del mes más frío/cálido), TMMF y TMMC (temperaturas medias de las mínimas/máximas del mes más frío/cálido) se calculan como se haría para variables compendiales. Dado que C y F son las temperaturas extremas y por tanto corresponden a un único dato del archivo, el error es el la diferencia entre el valor real y el estimado en ese punto concreto. Toda esta información queda almacenada en los archivos generados.

Ejemplo de aplicación

Como ejemplo de aplicación de todo lo anterior, se muestra la generación de las temperaturas del año 1988 en un punto situado a 1500 m de altitud y con coordenadas 0459W 3642. Se han determinado 2 estaciones de referencia distintas, la 5555O para las medias y la 6138 para las regresiones. Este hecho es debido a la exigencia en el primer caso de que la diferencia de altitud sea menor de 100 m, lo que ha obligado a tomar una estación que no es la más próxima. La reconstrucción se ha hecho para las variables térmicas media de las mínimas, media de las máximas y temperatura media mensual. Los valores obtenidos se muestran en las tablas 1 y 2.

ESTUDIOS COMPLEMENTARIOS

Una vez generados los datos termopluviométricos de un punto con GENPT, una adaptación del programa CLIMOAL (MANRIQUE, 1992; MANRIQUE, 1993; MANRIQUE *et al.*, 1995) permite calcular los datos taxonómicos del sistema de ALLUÉ-ANDRADE (1990), es decir: K, A, P, PE, TMF, T, TMC, TMMF, TMMC, OSC, C, F, HS y HP. Los programas GENPT determinan el

error en la estimación de todas estas variables, que en algunos casos como K y A, dependen de las temperaturas y de las precipitaciones.

Los programas hacen también un estudio de los valores que se alejan de la media en más de un número determinado de veces la desviación típica, por meses y por variables. A continuación, si se desea, se pueden cortar estos valores fuera de rango, colocándolos dentro del intervalo considerado.

RELLENADO DE HUECOS

Proceso de completar los archivos con datos faltantes

La escasez de datos reales disponibles es un hecho relativamente frecuente y da lugar a la generación de archivos climáticos con importantes huecos en algunos intervalos. Los subprogramas COMPLET se encargan de reducir o eliminar estos huecos, utilizando criterios de regresión cuando es posible o medias cuando no lo es, obteniéndose así series de datos más completas.

Inicialmente se hallan las 50 estaciones más próximas a la estación problema, sin importar su altitud. Luego, las 10 más próximas que alcancen un número mínimo de años (45). Se consideran distancias máximas de aproximadamente 100 km y se trabaja con un nivel de significación estadística del 99%.

Los programas leen los archivos climáticos de la estación a completar y los de las estaciones auxiliares. Se generan dos tipos de archivos:

- a) Archivos de medias: se toma la media de 10 valores, desde -5 hasta +5 de cada dato del archivo original.
- b) Archivos de regresiones: las regresiones son siempre entre cada estación auxiliar y la estación problema, entre los datos climáticos de la variable, considerando separadamente cada uno de los 12 meses del año. Por tanto, no interviene la altitud. Una vez obtenidas las rectas de regresión, se hallan los valores estimados para la estación problema y se hace una media ponderada por los coeficientes de correlación r , si los valores son significativos al nivel indicado (99%, siempre que $r > 0$).

A continuación, se genera el archivo completado, tomando los datos en el siguiente orden: 1. Original (O); 2. Regresión (R); 3. Media móvil (M). Este es el orden que se sigue para ir incorporando los datos al archivo definitivo; si falta un escalón, se recurre al siguiente. Estas claves se añaden después de los datos, para tener constancia de su origen. Se añaden también otras informaciones valiosas; por ejemplo, en el caso de las regresiones se indica si estamos dentro del rango de los datos originales de la estación o fuera del mismo.

Por último, se realiza un último intento de perfeccionar los archivos reduciendo las lagunas persistentes. Primero se hace una interpolación sólo para temperaturas: si falta un mes y están los meses adyacentes del mismo año, asigna la media al faltante. A continuación rellena los huecos que todavía persistan, en P ó T, con la media de 10 datos, 5 a un lado y 5 al otro, siempre del mismo mes y distintos años. En los extremos, como es lógico, el intervalo de 10 es asimétrico.

Errores cometidos al completar los archivos

Los errores también son calculados en el proceso de completar datos. Para ello, se calcula la diferencia entre el valor real y el valor estimado, por regresiones y por medias. Dado que son datos pareados, se hallan las diferencias y se determina la distribución por meses, con su media y error cuadrático medio. No se puede trabajar con las estimaciones de datos faltantes, dado que en este caso no se pueden hallar las mencionadas diferencias. Se ha optado por trabajar con distribuciones genéricas por meses, para una variable de precipitación ó cinco térmicas. Aunque es discutible asignar distribuciones a C ó F, su utilización permite considerar un error genérico a estas temperaturas extremas. Para las variables Precipitación total y Temperatura media, se hallan los valores anuales, y a estas nuevas variables se les aplica el mismo criterio que ya se mencionó. Las variables TMF, TMC, TMMF y TMMC se calculan hallando a nivel compendial el mes más frío y el más cálido y asignando la media y el error (diferencias) que ya fueron determinadas antes.

CONCLUSIONES

El sistema de generación de variables climáticas que aquí se ha expuesto tiene la ventaja de basarse en datos locales, próximos al punto problema, a los que se aplican modelos lineales de reconstrucción. Esto permite una correcta aproximación al mesoclima de la zona estudiada. Los

programas permiten generar datos en puntos de elevada altitud, a pesar de que disminuye el número de estaciones que se pueden utilizar, suavizando las condiciones exigidas. Las aplicaciones de los programas se extienden también a las variables fitoclimáticas. Los datos faltantes en los archivos climáticos se pueden completar por otros procedimientos, hasta obtener series climáticas de gran calidad.

BIBLIOGRAFÍA

ALLUÉ ANDRADE, J.L.; 1990. *Atlas fitoclimático de España. Taxonomías*. Instituto Nacional de Investigaciones Agrarias (MAPA). Madrid

FERNÁNDEZ CANCIO, A. y MANRIQUE MENÉNDEZ, E.; 2001. *Programas GenPT y Complet. Instrucciones de utilización*. Registro de la propiedad intelectual. N.º solicitud: 106649. Madrid.

MANRIQUE MENÉNDEZ, E.; 1992. Informatización de los Diagramas de Walter (Gausson) y la Taxonomía Fitoclimática de Allué Andrade 1990. *Montes*. 1º trimestre de 1992: 62–64.

MANRIQUE MENÉNDEZ, E.; 1993. *Informatizaciones Climoal*. Fundación Conde del Valle de Salazar. Madrid

MANRIQUE MENÉNDEZ, E.; FERNÁNDEZ MELÉNDEZ, J.A. y GRAU CORBÍ, J.M.; 1995. *Informatizaciones Climoal. Instrucciones de utilización de la versión de 1995*. E.U.I.T. Forestal. Universidad Politécnica de Madrid.

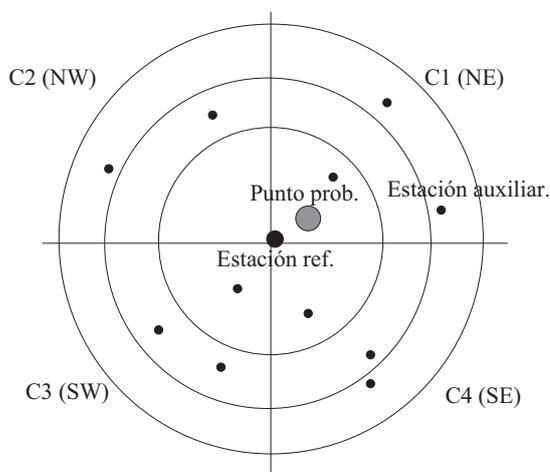


Figura 1: Punto problema (círculo grande), estación de referencia (centro de coordenadas) y estaciones auxiliares. Se indican los cuadrantes y varios círculos de distintos radios.

Tabla 1: Generación del año 1988 en el punto de altitud 1500 m y coordenadas 0459W 3642

MES	C	F	TMMC	TMMF	T
1	10.7	-6.9	5.4	-1.9	2.4
2	15.1	-4.9	4.7	1.1	3.0
3	21.3	-4.9	10.9	3.3	5.0
4	20.1	-2.9	12.3	3.2	6.2
5	17.5	5.1	13.0	6.7	12.6

6	27.1	7.6	15.2	8.7	14.9
7	30.6	10.6	24.9	14.9	22.0
8	34.3	14.1	28.7	17.0	22.1
9	32.7	9.1	26.5	16.0	19.3
10	25.1	5.6	15.6	6.7	14.1
11	19.1	2.6	9.0	5.2	7.6
12	12.1	0.1	8.8	2.5	4.7

Tabla 2: Generación del año 1988 en el punto de altitud 1500 m y coordenadas 0459W 3642. Se indica la estación de referencia y un grupo de dígitos que indica: n.º de estaciones auxiliares, coef. de correlación x 100 (r; si es significativo, R), distancia d en grados x 10 (hasta 0.12°, C hexadecimal ≈ 160 km) y cuadrante c.

Meses	TMMC		TMMF		T	
	Estac Ref.	Varios	Estac Ref.	Varios	Est. Ref.	Varios
1	6143	13R-93d7c0	6031	16R-88dCc0	6118A	14R-93d6c0
2	6139	15R-94d7c0	6139	19R-53dCc0	6139	12R-92d7c0
3	6139	12R-94d5c0	6139	11r-48d6c0	6139	11R-93d6c0
4	6139	13R-77dCc1	6120	11R-65d7c1	6118A	11R-94d4c0
5	6139	11R-88dCc1	6139	13r-47d8c0	6139	12R-79d8c0
6	6119I	12R-79d8c0	6031	13R-93d7c0	6143	12R-67d6c0
7	6139	11R-82d6c1	6031	12R-67d6c0	6139	11R-61dCc1
8	6118A	16R-61dCc0	6139	12r 13dCc0	6139	14R-60d7c0
9	6118A	11r-30dCc0	6139	14r-14d8c0	6139	11R-76dCc1
10	6139	11R-90d8c1	6139	19R-84d8c0	6118A	15R-66dCc0
11	6118A	11R-97dCc0	6139	12R-64dCc1	6118A	13R-84d8c0
12	6119I	12R-85dCc1	6143	17r-47dCc0	6032	12R-71dCc0