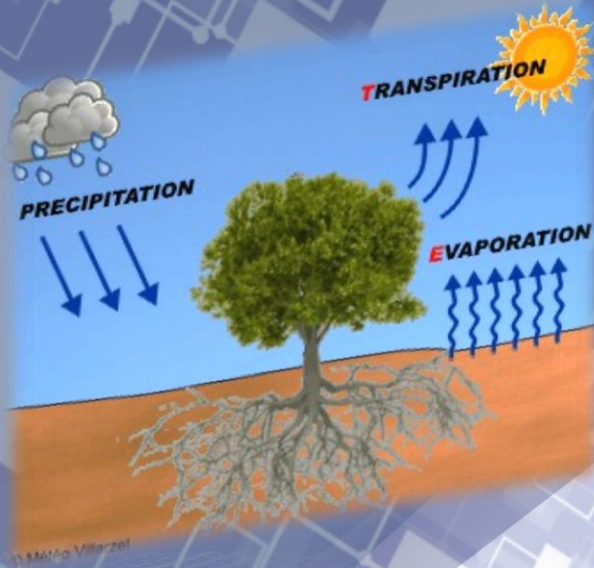




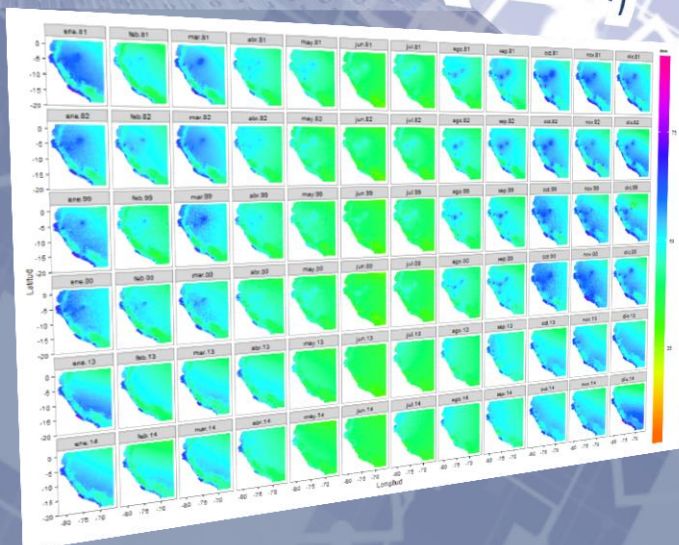
PERÚ

Ministerio del Ambiente

Servicio Nacional de Meteorología e Hidrología del Perú - SENAMHI



# GENERACIÓN DE UNA BASE DE DATOS DE EVAPOTRANSPIRACIÓN GRILLADA Y DE ALTA RESOLUCIÓN A NIVEL NACIONAL (1981-2014)



2015





PERSONAL DIRECTIVO

**Ing. Amelia Díaz Pabló**  
*Presidente Ejecutivo del SENAMHI*

**Ing. Esequiel Villegas Paredes**  
*Director Científico del SENAMHI*

**Ing. Oscar G. Felipe Ovando**  
*Director de Hidrología y Recursos Hídricos del SENAMHI*

**PhD. Waldo Lavado Casimiro**  
*Director de Hidrología Aplicada del SENAMHI*

REVISOR DEL ESTUDIO

**PhD. Waldo Lavado Casimiro**  
*Director de Hidrología Aplicada del SENAMHI*

RESPONSABLE DEL ESTUDIO

**Ing. Eber Risco Sence**  
*Especialista en Hidrología del SENAMHI*

Lima-Perú  
Noviembre 2015

---

## ÍNDICE

RESUMEN .....	4
ABSTRACT .....	5
I. ASPECTOS GENERALES .....	6
1.1 INTRODUCCIÓN .....	6
1.2 OBJETIVOS .....	7
II. ZONA DE ESTUDIO .....	8
2.1 UBICACIÓN .....	8
2.2 DATOS DISPONIBLES .....	9
III. MÉTODOS .....	10
3.1 DATOS GRILLADOS DE TEMPERATURA (1981-2014) .....	10
3.2 DATOS GRILLADOS DE EVAPOTRANSPIRACIÓN POTENCIAL (1981-2014) .....	13
3.2.1 Método de Hargreaves y Samani .....	13
3.3 EFICIENCIA DE LAS REGRESIONES MULTIVARIANTES .....	14
3.3.1 Validación cruzada .....	14
3.3.2 Error cuadrático medio .....	14
3.3.3 Coeficiente de determinación múltiple .....	15
IV. RESULTADOS Y DISCUSIÓN .....	16
4.1 EXPLORATORIO DE DATOS .....	16
4.2 VALIDACIÓN CRUZADA .....	18
4.3 COEFICIENTE DE DETERMINACIÓN MÚLTIPLE .....	19
4.4 DATOS GRILLADOS DE TEMPERATURA (1981-2014) .....	20
4.5 DATOS GRILLADOS DE EVAPOTRANSPIRACIÓN POTENCIAL (1981-2014) .....	25
V. CONCLUSIONES .....	29
VI. BIBLIOGRAFÍA .....	30

## LISTA DE FIGURAS

<b>Figura 2.2.</b> Localización del área de estudio y ubicación espacial de las estaciones meteorológicas utilizadas (1981-2014).	8
<b>Figura 2.3.</b> Evolución temporal mensual de la cantidad de datos utilizados para generar los mapas de temperatura mínima y máxima.	9
<b>Figura 2.4.</b> Variables explicativas utilizadas en la regresión multivariante, la longitud y latitud están expresadas en $^{\circ}/1000$ , la elevación en km, el seno y coseno representan razones de cambio en el terreno (m/m).	10
<b>Figura 3.1.</b> Componentes principales de las Variables explicativas utilizadas en la regresión multivariante y porcentaje de varianza explicada por cada componente principal.	11
<b>Figura 3.2.</b> Diagrama de flujo de la metodología utilizada para la obtención de los mapas de temperatura.	12
<b>Figura 3.3.</b> Variación latitudinal de los valores de la Radiación solar extraterrestre en MJ m <sup>-2</sup> d <sup>-1</sup> para cada mes del año.	15
<b>Figura 4.1.</b> Variación temporal de las temperaturas en las estaciones Tarata y Bambamarca. Se puede apreciar carencia de información en ciertos periodos como saltos en la media.	16
<b>Figura 4.2.</b> a) Distancia de Mahalanobis para los datos de las variables originales b) Distancia de Mahalanobis para los datos de los componentes principales.	17
<b>Figura 4.3.</b> Variación temporal del error cuadrático medio de las temperaturas mínimas y máximas.	18
<b>Figura 4.4.</b> Variación temporal del coeficiente de determinación múltiple de las regresiones multivariantes generadas para la temperatura mínima y máxima.	19
<b>Figura 4.5.</b> Variación espacial de la temperatura mínima ( $^{\circ}\text{C}$ ) obtenida del modelo de regresión multivariante para los meses del año 2013.	20
<b>Figura 4.6.</b> Variación espacial de los errores interpolados mediante la técnica IDW de la temperatura mínima para los meses del año 2013.	22
<b>Figura 4.7.</b> Variación espacial de la temperatura mínima ( $^{\circ}\text{C}$ ) para el año 2013.	23
<b>Figura 4.8.</b> Variación espacial de la temperatura máxima ( $^{\circ}\text{C}$ ) para el año 2013.	24
<b>Figura 4.9.</b> Datos grillados de temperatura mínima ( $^{\circ}\text{C}$ ) a nivel nacional (1981-2014).	25
<b>Figura 4.10.</b> Variación espacial de la radiación solar incidente (mm/día) para el año 2013.	26
<b>Figura 4.11.</b> Variación evapotranspiración potencial (mm/mes) para el año 2013.	27
<b>Figura 4.12.</b> Datos grillados de evapotranspiración potencial (mm/mes) a nivel nacional (1981-2014).	27

## LISTA DE TABLAS

<b>Tabla 3.1:</b> Tabla de Radiación solar extraterrestre en MJ m <sup>-2</sup> d <sup>-1</sup> (Allen et al., 1998).	14
---	----

## RESUMEN

El presente estudio tiene como principal objetivo la generación de una base de datos de Evapotranspiración grillada y de alta resolución a nivel nacional para el periodo 1981-2014. Para el estudio se ha utilizado la información de temperatura máxima y mínima de entre 70 y 200 estaciones meteorológicas dentro del ámbito del territorio peruano.

En la metodología se describe la aplicación del método para la determinación de las variables topográficas expresadas en los ángulos seno y coseno de la orientación de la pendiente, además de la ubicación espacial (latitud, longitud y altitud) como variables explicativas para la estimación de la temperatura máxima y mínima, mediante una regresión multivariante, para obtener variables incorrelacionadas se ha aplicado componentes principales a las variables mencionadas, con estas nuevas variables se procedió a realizar la regresión multivariante, finalmente se le adiciona la interpolación de los errores de la estimación del modelo, para tal fin se ha utilizado la interpolación de la distancia inversa ponderada, con lo cual se han obtenido mapas de temperatura máxima y mínima a una resolución espacial de  $0.05^\circ$  y temporal de un mes.

Los modelos de regresiones multivariantes se presenta robusto en el proceso de validación (realizada mediante validación cruzada), los coeficiente de determinación nos muestran que los modelos explican de mejor manera la temperatura mínima, donde el  $R^2$  se encuentra en promedio para la serie 1981-2014 por encima de 0.5, mientras para la temperatura máxima este indicador estadístico disminuye a 0.4. Debido a la disponibilidad de información la Evapotranspiración Potencial (EP) fue obtenida mediante un modelo de balance de energía como el de Hargreaves y Samani. No se cuenta con información observada para validar los mapas de evapotranspiración potencial generados, por lo cual se recomienda para futuros trabajos la validación de estos mapas con otras metodologías de obtención de la EP, además de información satelital, como por ejemplo con el producto MOD 16 - Evapotranspiration del sensor MODIS.

## ABSTRACT

The present study aims to create a gridded and high resolution data set of evapotranspiration estimates over the territory of the Peru for the period 1981-2014. For the study we used the maximum and minimum temperature of between 70 and 200 weather stations.

Methodology describes the method for determining the topographical variables such as sin and cosine angle of slope, in addition to the spatial location (latitude, longitude and altitude) as explanatory variables for estimation of maximum and minimum temperature, using a multivariate regression, for uncorrelated variables was applied principal components to these variables, with these new variables we proceeded to perform multivariate regression, finally, interpolation errors was added of the model estimation, we used the inverse distance weighted interpolation, which were obtained maps of maximum and minimum temperature for a 0.05 ° resolution gridded data set of monthly.

Multivariate regression models were robust in the validation process (performed using cross-validation), the coefficient of determination show that the models better explain the minimum temperature, where average R<sup>2</sup> for the series 1981-2014 is above 0.5, while the maximum temperature for this statistical indicator is 0.4. Due to the availability of Potential Evapotranspiration (PE) data, it was obtained using a model of energy balance such as Hargreaves and Samani method. The lack of EP observed data prevents validate generated maps, thus, we recommended for future study to validate these kind of maps with other methods of EP, satellite information such as the MOD 16 product - Evapotranspiration MODIS.

## I. ASPECTOS GENERALES

### 1.1 Introducción

La evapotranspiración (ET) es un elemento central en el ciclo hidrológico, que rige la transferencia de humedad a la atmósfera y su influencia en las propiedades fundamentales de los ecosistemas terrestres, como la escorrentía, la humedad del suelo y crecimiento de la planta. Sin embargo, las interacciones básicas entre la atmósfera, suelo y las plantas aún se encuentran en investigación (por ejemplo, Brutsaert y Parlange, 1998; Golubev et al., 2001; Ohmura y Wild, 2002; Peterson et al., 1995; Roderick y Farquhar, 2002) propiedades básicas tales como totales anuales y la variabilidad de ET no son obtenidos con precisión en muchas regiones del mundo. En el Perú, la representación de la variabilidad espacio temporal de la evapotranspiración es considerado uno de los mayores inconvenientes al momento de realizar estudios hidrológicos y ambientales.

Como la ET no se puede medir directamente sin un considerable esfuerzo, es necesario realizar estimaciones a partir de la meteorología de datos y/o el uso de métodos hidrológicos. El método de Hargreaves y Samani, representa una ecuación de balance de energía para estimar la ET potencial (ETP) a partir de datos de temperatura mínima y máxima.

Existen estimaciones ET grilladas del modelo climático global (GCM por sus siglas en inglés) cuyos resultados carecen de una buena resolución requeridos para representar la variabilidad espacial. Esto implica que datos de ET observados sean necesarios para calibrar los resultados.

En general existen mayor cantidad de estaciones que miden la temperatura del aire que estaciones que midan la insolación o la velocidad del viento, razón por lo cual se ha optado por utilizar un método de balance de energía para estimar la ETP. Sin embargo, para ello, previamente se deberá obtener las variables de temperatura máxima y mínima de manera espacial.

A pesar de una intensa investigación durante varias décadas (OMM, 1972) la interpolación de datos climáticos, en particular en las regiones montañosas, para obtener datos espacialmente distribuidos ha seguido siendo un problema. Incluso si se utiliza técnicas multivariante o métodos de interpolación geoestadística (por ejemplo, Hutchinson, 1995a; Martínez-Cob, 1996), la interpolación directa sin la inclusión de los datos topográficos no puede capturar la variabilidad espacial a menos. Incluyendo información de elevación de un modelo digital de elevación (DEM) en el procedimiento de regionalización (por ejemplo, Phillips et al., 1992; Hudson y Wackernagel, 1994; Hutchinson, 1995b; Fleming et al., 2000; Goovaerts, 2000; New et al., 2002) conduce a mejores resultados.

En la mayoría de los casos, los procesos climatológicos, como el levantamiento orográfico de las masas de aire, están influenciados por aspectos morfológicos y diferencias de elevación relativos de la topografía local, más que por la altitud.

Para tener en cuenta la influencia de las variables topográficas, tales como la exposición, la pendiente o la distancia y la diferencia de elevación a la más alta barrera topográfica se han utilizado (ver Prudhomme y Reed (1998)), en particular en la regionalización de los datos climatológicos.

En este trabajo se describe la aplicación de un método objetivo para la determinación de las variables topográficas expresadas en los ángulos seno y coseno de la orientación de la pendiente, además de la ubicación espacial (latitud, longitud y altitud) como variables explicativas para la estimación de la temperatura máxima y mínima, mediante una regresión multivariante, para obtener variables incorrelacionadas se ha aplicado componentes principales a las variables mencionadas, con estas nuevas variables se procedió a realizar la regresión multivariante, finalmente se le adiciona la interpolación de los errores de la estimación del modelo, para tal fin se ha utilizado la interpolación de la distancia ponderada.

El servicio nacional de meteorología e hidrología del Perú (SENAMHI) viene realizando trabajos para explicar la variabilidad espacial y temporal de las variables meteorológicas en el territorio peruano, a partir de información observada y datos de satélite basado en técnicas estadísticas ha generado mapas a una resolución espacial de 5 km y una resolución temporal de un mes. En tal sentido, el producto PISCO (Peruvian Interpolation of the SENAMHI's Climatological and Hydrological data observed) temperatura y evapotranspiración estima las variaciones espaciales y temporales de la temperatura mínima y máxima (mediante regresiones múltiples), evapotranspiración potencial (por el método de Hargreaves y Samani) para el periodo comprendido entre 1981-2014.

En vista de importancia atribuida por la comunidad científica en los datos climáticos espaciales, mapas de temperatura máxima, mínima y de evapotranspiración potencial se han obtenido para el territorio peruano.

## **1.2 Objetivos**

### **1.2.1 Objetivo general**

Generar una base de datos de Evapotranspiración grillada y de alta resolución a nivel nacional para el periodo 1981-2014.

### **1.2.2 Objetivos específicos**

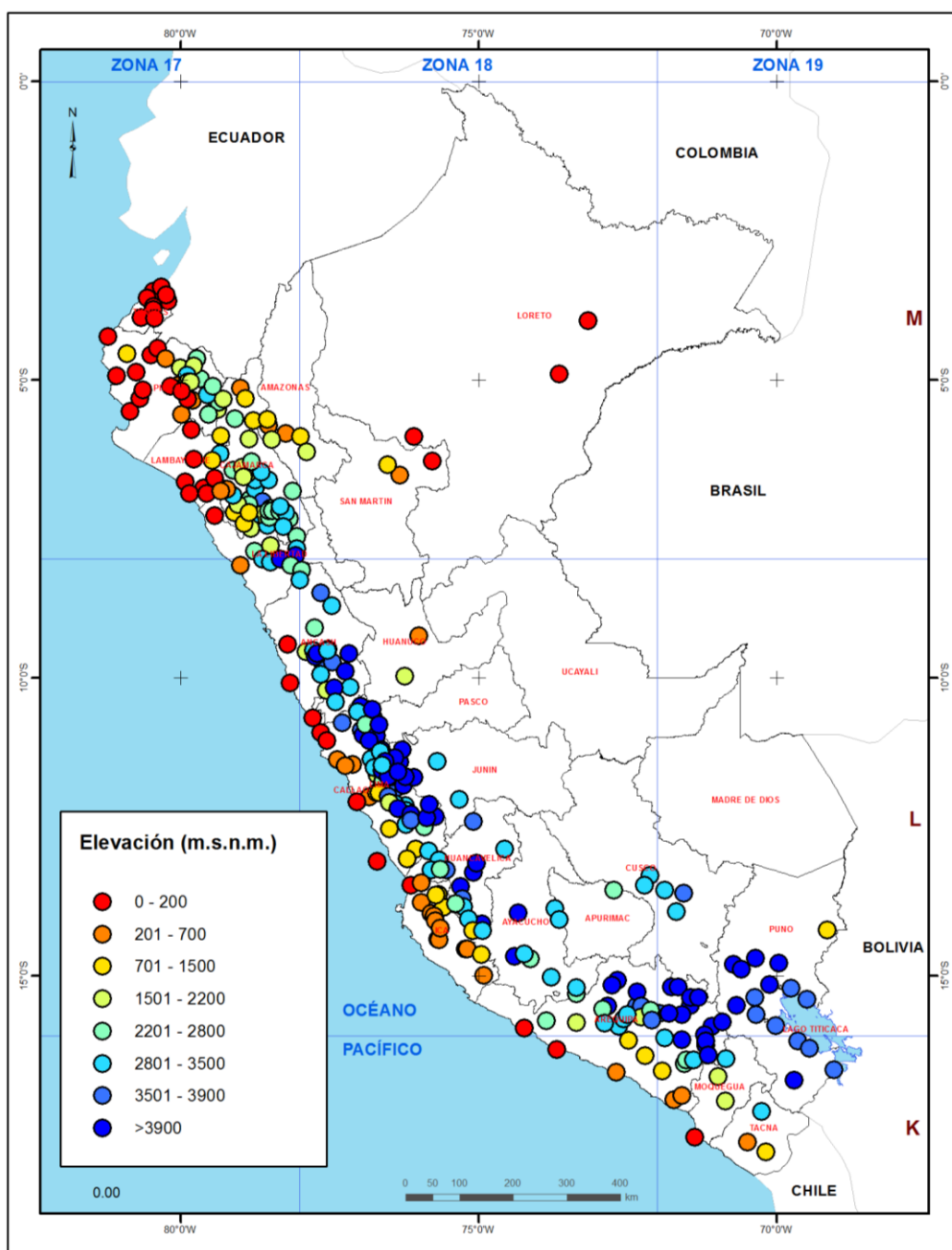
- Realizar el análisis de los datos de temperatura máxima y mínima de temperatura a nivel nacional.
- Implementar modelos de regresiones multivariantes que expliquen la variabilidad espacial y temporal de la temperatura máxima y mínima usando variables auxiliares de ubicación espacial y características topográficas.
- Generar mapas de temperatura máxima y mínima a una resolución espacial de 0.05° y temporal de un mes.
- Implementar el método de balance de energía de Hargreaves y Samani para generar mapas de evapotranspiración potencial.



## II. ZONA DE ESTUDIO

### 2.1 Ubicación

El área de estudio está representada por el territorio peruano, se encuentra situado en la parte central y occidental de América del Sur, ubicado entre las latitudes  $0^{\circ}01' - 18^{\circ}20' S$  y las longitudes  $68^{\circ}39' - 81^{\circ}19' W$ . Está conformado por un territorio de una superficie continental de 1.285.215,60 km<sup>2</sup> de superficie, lo que representa el 0.87% del planeta, que se distribuyen en región costera 136.232,85 km<sup>2</sup> (10,6%), región andina 404.842,91 km<sup>2</sup> (31,5%) y región amazónica 754.139,84 km<sup>2</sup> (57,9%).



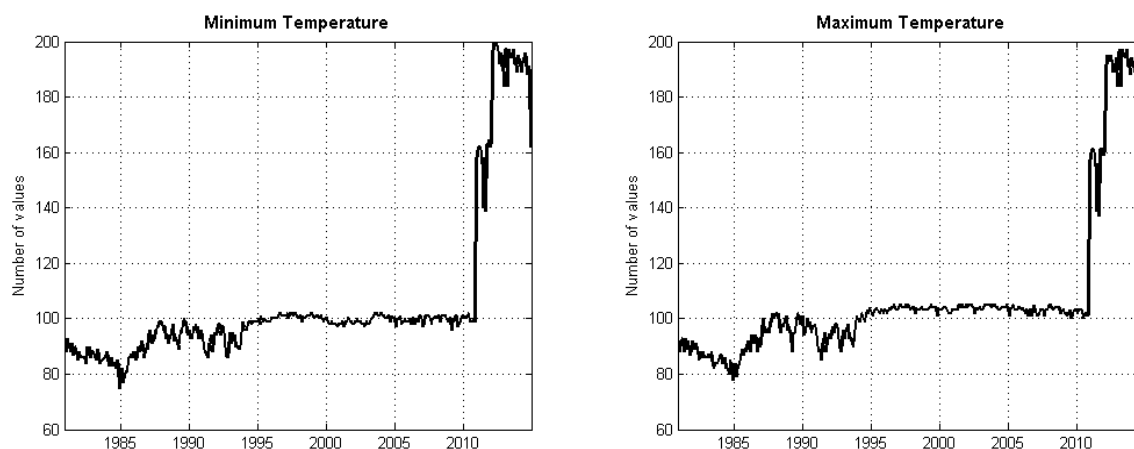
**Figura 2.1.** Localización del área de estudio y ubicación espacial de las estaciones meteorológicas utilizadas (1981-2014).

Los Andes atraviesan el Perú de norte a sur, condicionando con su imponente presencia el clima y la orografía del país. Aunque el abra de Porculla marque, a 2137 msnm, su punto más bajo, la Cordillera Blanca y la Cordillera de Huayhuash albergan en el sector norte las cumbres más altas del Perú. A partir del nudo de Pasco, los Andes centrales se ensanchan y presenta mesetas entre las cordilleras y cimas como las del Coropuna, el Ampato o el Salcantay. La meseta del Collao, a 3600 msnm y la cordillera Volcánica, con los nevados Misti, Pichu Pichu, y Ubinas, compone el sector meridional de los Andes peruanos a partir del Nudo de Vilcanota.

El hecho de estar en el Perú cerca de la línea ecuatorial indicaría que su clima debería ser eminentemente tropical, sin embargo dos factores alteran notablemente el clima. En primer lugar la existencia de la elevada Cordillera de los Andes paralela en América del Sur al Océano Pacífico y, en segundo lugar, la fría Corriente Peruana o de Humboldt que se manifiesta de sur a norte hasta la latitud 5° y que choca con la Corriente del Niño en las costas de Piura y Tumbes hasta la latitud 3.2°, al sur de la línea ecuatorial. Estos accidentes, más el anticiclón del Pacífico sur en esta parte del continente, originan una disminución de las temperaturas promedio anuales de unos diez grados centígrados en la costa y una gran variedad de climas simultáneos en todo el país el cual ha situado al Perú como uno de los países con mayor variedad de climas en el mundo.

## 2.2 Datos disponibles

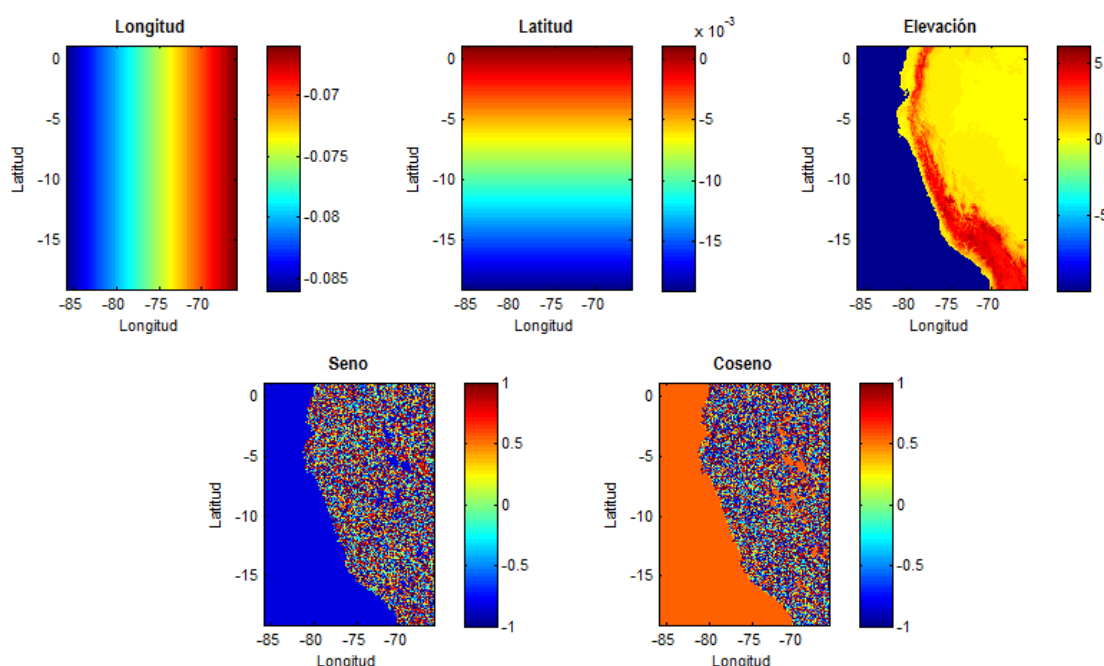
Para obtener una base de datos de Evapotranspiración grillada y de alta resolución a nivel nacional (1981-2014) se ha partido desde la obtención de mapas de temperatura máxima y mínima, los cuales son estimados a partir de datos de estaciones meteorológicas del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). Estos datos fueron sometidos a un proceso de control de calidad visual enfocada en reportar valores atípicos y quiebres temporales de temperatura. El periodo de análisis es de 1981-2014 con un número de estaciones variables desde 75 hasta 200, cuya condición era que tengan al menos 10 años de registro (un año con al menos 9 meses de datos, ver Figura 2.2).



**Figura 2.2.** Evolución temporal mensual de la cantidad de datos utilizados para generar los mapas de temperatura mínima y máxima.

Los procesos climatológicos están influenciados por aspectos morfológicos y diferencias de elevación relativos de la topografía local, más que por la altitud, por lo cual para estimar las temperaturas, se han utilizado variables que expliquen la ubicación espacial como la longitud, latitud y elevación, para explicar condiciones de relieve se han utilizado el seno y coseno de la orientación del terreno, estos datos fueron obtenidos de manera espacial mediante interpolación lineal (latitud, longitud), del modelo digital de elevaciones MDE de la NASA se obtuvieron las altitudes, a partir del MDE se han obtenido las orientaciones del terreno, y posteriormente el seno y coseno de las inclinaciones, las cuales se muestran en la Figura 2.3.

Las variables están expresadas en  $^{\circ}/1000$  para la latitud y longitud, en km para la elevación, los valores de seno y coseno representan las razones de cambio en el terreno con respecto a la orientación del terreno, expresadas en m/m.



**Figura 2.3.** Variables explicativas utilizadas en la regresión multivariante, la longitud y latitud están expresadas en  $^{\circ}/1000$ , la elevación en km, el seno y coseno representan razones de cambio en el terreno (m/m).

### III. MÉTODOS

#### 3.1 Datos grillados de temperatura (1981-2014)

La metodología para obtener los datos grillados de temperatura consiste en predecir la temperatura a partir de la relación entre las temperaturas mínima y máxima con variables auxiliares como la elevación, la latitud, longitud, el seno y coseno del aspecto del terreno mediante una regresión multivariante, los residuos del modelo de regresión se interpola mediante IDW y finalmente la suma de los valores estimados tanto por regresión e interpolación de los residuales, resultan ser los mapas de temperaturas.

El modelo de regresión multivariante se expresa de la siguiente manera:

$$Y_i = \beta_0 + \beta_1 \cdot X_{i1} + \dots + \beta_p \cdot X_{ip} + \mu_i$$

Donde:

$Y_i$ = temperatura máxima, mínima

$\beta_i$ = coeficientes del modelo.

$X_i$ =longitud, latitud, elevación, seno y coseno del aspecto del terreno.

$\mu_i$ = error de estimación.

Sin embargo, la regresión multivariante debe cumplir algunos supuestos, por lo cual se realizaron análisis previos antes de efectuar las regresiones múltiples, como el análisis exploratorio de datos multivariantes (el cual fue realizado mediante la distancia de Mahalanobis).

Distribución normal multivariante (medida pruebas de asimetría y curtosis), prueba de linealidad (Coeficiente de dependencia efectiva), para tener variables incorrelacionadas se utilizaron los componentes principales de las variables iniciales, con lo cual el modelo de regresión multivariante finalmente utilizado es de la forma.

$$Y_i = \beta_0 + \beta_1 \cdot CP_{i1} + \dots + \beta_p \cdot CP_{ip} + \mu_i$$

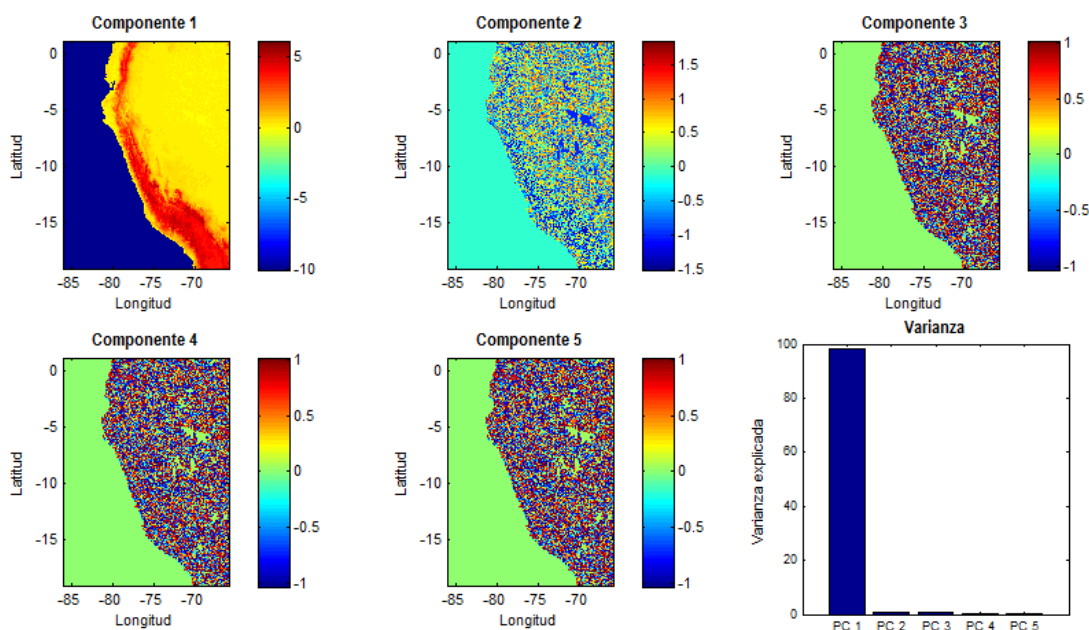
Donde:

$Y_i$ = temperatura máxima, mínima

$\beta_i$ = coeficientes del modelo.

$CP_i$ =componente principal, con  $i=1,2,\dots,5$ .

$\mu_i$ = error de estimación.



**Figura 3.1.** Componentes principales de las Variables explicativas utilizadas en la regresión multivariante y porcentaje de varianza explicada por cada componente principal.

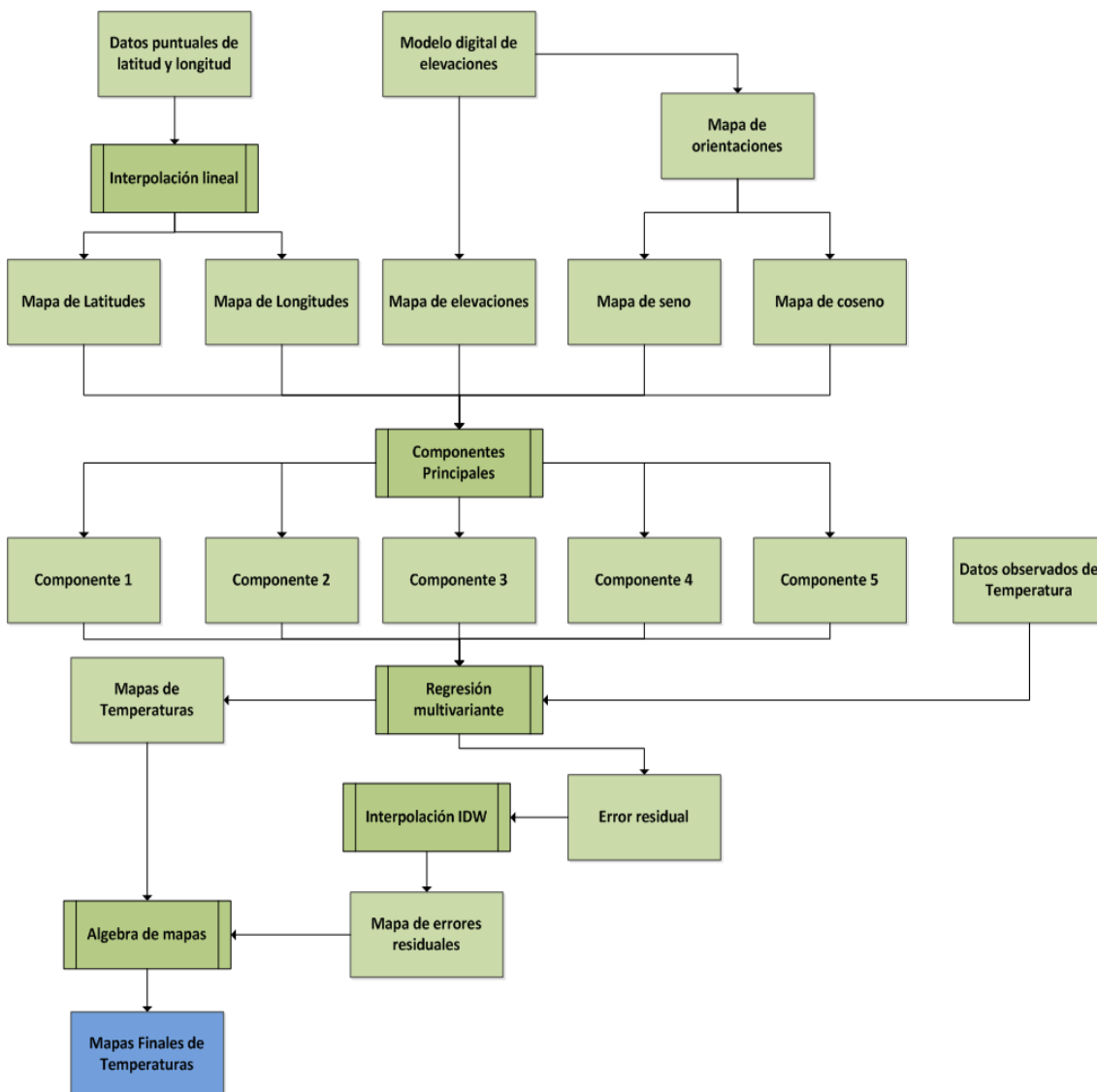
Los residuos fueron interpolados mediante interpolación de la inversa ponderada de la distancia, definido como:

$$\lambda_i = \frac{1}{|x_i - x_{0i}|^\beta}$$

Con  $\beta > 0$

El resultado es que los datos de los puntos cercanos son ponderados por el inverso de la distancia para obtener el valor en el punto objetivo. La opción más usada es cuando  $\beta=2$ .

Los mapas de temperaturas fueron obtenidos a 0.05° de resolución espacial y de un mes de resolución temporal.



**Figura 3.2.** Diagrama de flujo de la metodología utilizada para la obtención de los mapas de temperatura.

### 3.2 Datos Grillados de evapotranspiración potencial (1981-2014)

El método usado para obtener la evapotranspiración potencial fue el propuesto por Hargreaves y Samani, el cual se detalla a continuación.

#### 3.2.1 Método de Hargreaves y Samani

La fórmula de Hargreaves (Hargreaves y Samani, 1985) para evaluar la evapotranspiración potencial necesita los datos de temperatura y de radiación solar. La expresión general es la siguiente:

$$ET_o = 0.0135(t_{med} + 17.78)R_s$$

Donde:

ET<sub>o</sub>=evapotranspiración potencial diaria, mm/día.

t<sub>med</sub>= temperatura media °C.

R<sub>s</sub>= radiación solar incidente, convertida en mm/día.

La radiación solar incidente, R<sub>s</sub> se evalúa a partir de la radiación solar extraterrestre (la que llega a la parte exterior de la atmosfera, que sería la que llegaría al suelo si no existiera atmosfera); esta última aparece según los autores como R<sub>o</sub> ó R<sub>a</sub>, y la leemos en tablas en función de la latitud del lugar y del mes. En este documento nos referimos a ella como R<sub>o</sub>.

#### Obtención de la radiación solar incidente (R<sub>s</sub>)

Samani (2000) propone la siguiente formula:

$$R_s = R_o * K_T * (t_{max} - t_{min})^{0.5}$$

Donde:

R<sub>s</sub>= Radiación solar incidente

R<sub>o</sub>=Radiación solar extraterrestre (tabulada)

K<sub>T</sub>=coeficiente

t<sub>max</sub>= temperatura máxima

t<sub>min</sub>=temperatura mínima

Puesto que los valores de R<sub>o</sub> están tabulados y las temperaturas máximas y mínimas son datos empíricos relativamente fáciles de obtener, la dificultad para aplicar esta sencilla expresión la encontramos en el coeficiente K<sub>T</sub>.

Para evaluar la Radiación Solar Extraterrestre (R<sub>o</sub>) existen varias tablas, todas ellas en función de la latitud y del mes. Esta tabla está en MJulio/m<sup>2</sup>/día, para pasar a mm./día (de agua evaporada) multiplicar por 0.4082.

El coeficiente K<sub>T</sub> de la expresión anterior es un coeficiente empírico que se puede calcular a partir de datos de presión atmosférica, pero Hargreaves (citado en Samani, 2000) recomienda K<sub>T</sub> = 0.162 para regiones del interior y K<sub>T</sub> = 0.19 para regiones costeras. Para el presente estudio se utilizó el valor de K<sub>T</sub> medio de 0.17.

Al ser los valores de  $R_o$  tabulados, mediante interpolación lineal se obtuvieron los mapas de  $R_o$  para cada mes, a una resolución espacial de  $0.05^\circ$ , los cuales se muestran en la Figura 3.3.

**Tabla 3.1:** Tabla de Radiación solar extraterrestre en MJ m<sup>-2</sup> d<sup>-1</sup> (Allen et al., 1998).

Latitud	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2	35.4	37.0	37.8	37.1	35.4	34.2	34.6	36.1	37.3	37.0	35.6	34.8
0	36.2	37.5	37.9	36.8	34.8	33.4	33.9	35.7	37.2	37.4	36.3	35.6
-2	36.9	37.9	38.0	36.4	34.1	32.6	33.1	35.2	37.1	37.7	37.0	36.4
-4	37.6	38.3	38.0	36.0	33.4	31.8	32.3	34.6	37.0	38.0	37.6	37.2
-6	38.3	38.7	38.0	35.6	32.7	30.9	31.5	34.0	36.8	38.2	38.2	38.0
-8	38.9	39.0	37.9	35.1	31.9	30.0	30.7	33.4	36.6	38.4	38.8	38.7
-10	39.5	39.3	37.8	34.6	31.1	29.1	29.8	32.8	36.3	38.5	39.3	39.4
-12	40.1	39.6	37.7	34.0	30.2	28.1	28.9	32.1	36.0	38.6	39.8	40.0
-14	40.6	39.7	37.5	33.4	29.4	27.2	27.9	31.3	35.6	38.7	40.2	40.6
-16	41.1	39.9	37.2	32.8	28.5	26.2	27.0	30.6	35.2	38.7	40.6	41.2
-18	41.5	40.0	37.0	32.1	27.5	25.1	26.0	29.8	34.7	38.7	40.9	41.7
-20	41.9	40.0	36.6	31.3	26.6	24.1	25.0	28.9	34.2	38.6	41.2	42.1

### 3.3 Eficiencia de las regresiones multivariantes

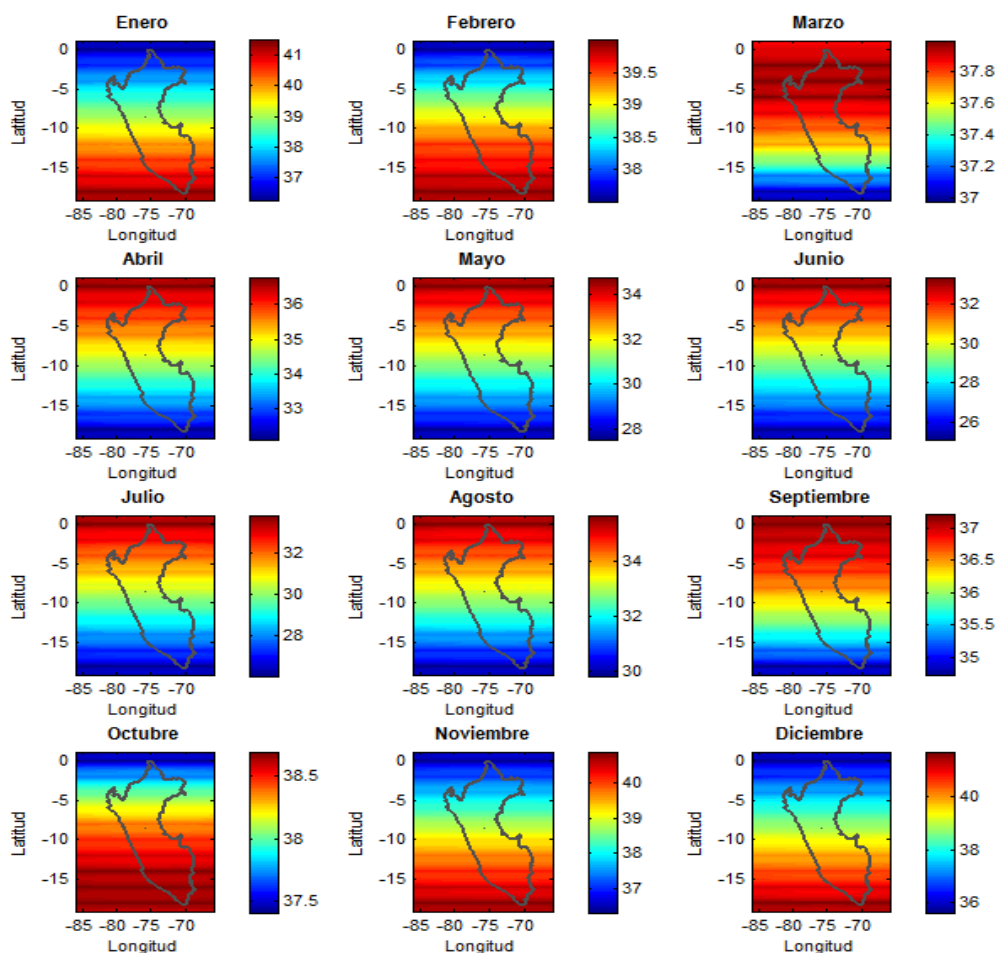
#### 3.3.1 Validación cruzada

La validación cruzada o cross-validation es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba. Consiste en repetir y calcular la media aritmética obtenida de las medidas de evaluación sobre diferentes particiones. Se utiliza en entornos donde el objetivo principal es la predicción y se quiere estimar cómo de preciso es un modelo que se llevará a cabo a la práctica. Es una técnica muy utilizada en proyectos de inteligencia artificial para validar modelos generados.

El estadístico para evaluar la eficiencia de las regresiones multivariantes se utilizó el error medio cuadrático.

#### 3.3.2 Error cuadrático medio

En estadística, el error cuadrático medio (ECM) de un estimador mide el promedio de los errores al cuadrado, es decir, la diferencia entre el estimador y lo que se estima. El ECM es una función de riesgo, correspondiente al valor esperado de la pérdida del error al cuadrado o pérdida cuadrática. La diferencia se produce debido a la aleatoriedad o porque el estimador no tiene en cuenta la información que podría producir una estimación más precisa.



**Figura 3.3.** Variación latitudinal de los valores de la Radiación solar extraterrestre en MJ m-2 d-1 para cada mes del año.

El ECM es el segundo momento (sobre el origen) del error, y por lo tanto incorpora tanto la varianza del estimador así como su sesgo. Para un estimador insesgado, el ECM es la varianza del estimador. Al igual que la varianza, el EMC tiene las mismas unidades de medida que el cuadrado de la cantidad que se estima. En una analogía con la desviación estándar, tomando la raíz cuadrada del EMC produce el error de la raíz cuadrada de la media o la desviación de la raíz cuadrada media (RMSE o RMSD), que tiene las mismas unidades que la cantidad que se estima; para un estimador insesgado, el RMSE es la raíz cuadrada de la varianza, conocida como la desviación estándar.

### 3.3.3 Coeficiente de determinación múltiple

El coeficiente de determinación ( $R^2$ ) o coeficiente de correlación múltiple al cuadrado, es una medida descriptiva que sirve para evaluar la bondad de ajuste del modelo a los datos, ya que mide la capacidad predictiva del modelo ajustado. Se define como el cociente entre la variabilidad explicada por la regresión y la variabilidad total, esto es:

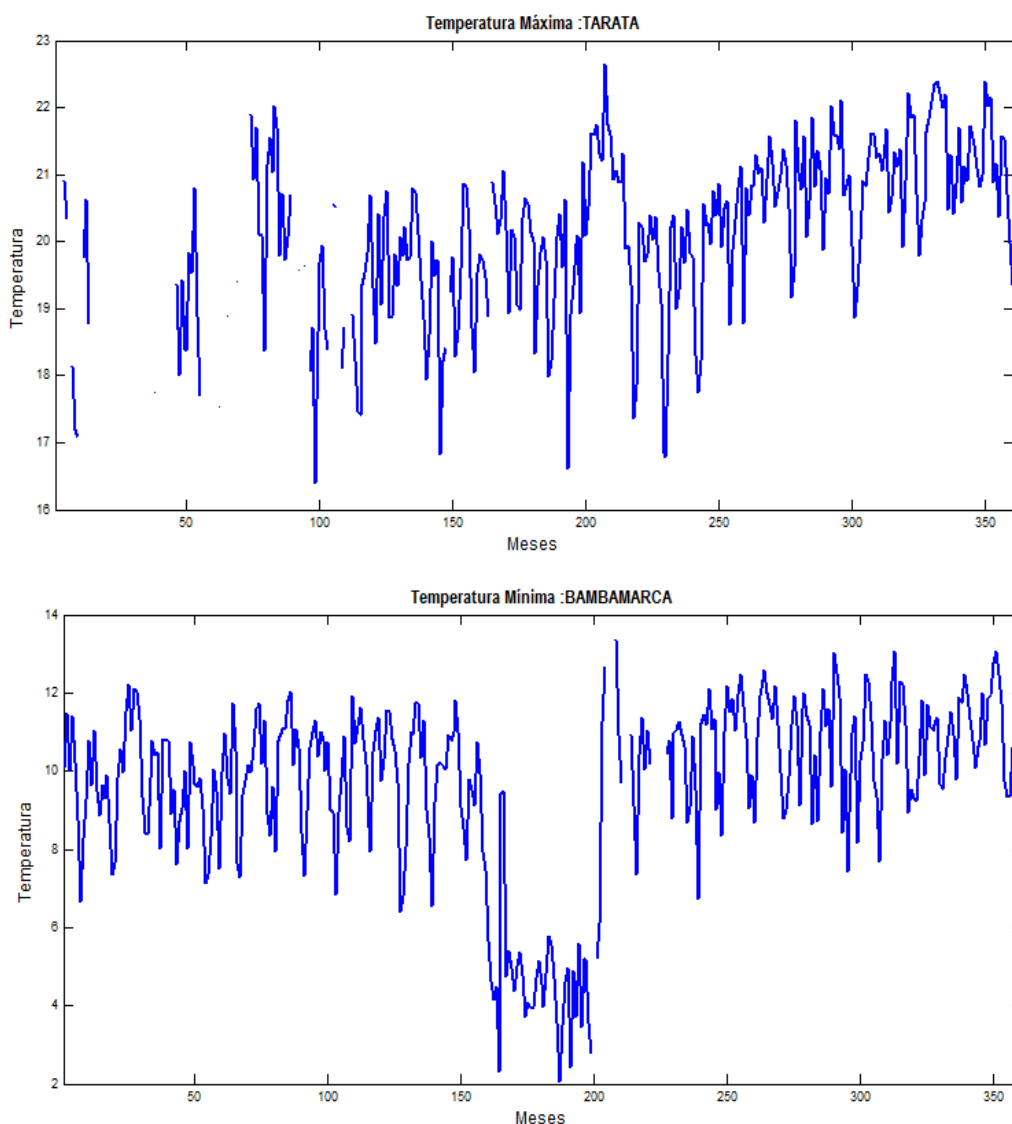
$$R^2 = \frac{\text{Varianza explicada}}{\text{Varianza total}} = 1 - \frac{\text{Varianza no explicada}}{\text{Varianza total}}$$



## IV. RESULTADOS Y DISCUSIÓN

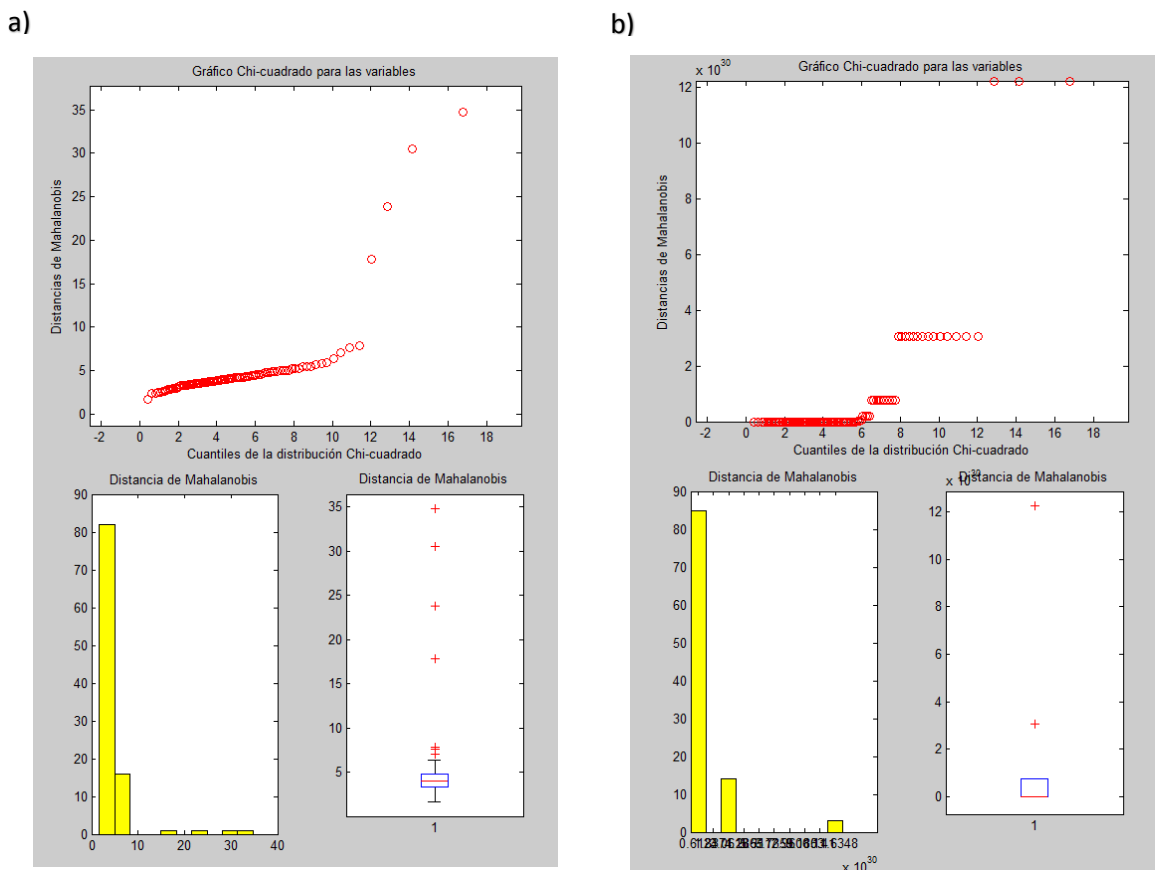
### 4.1 Exploratorio de datos

El exploratorio de los datos observados se realizó de manera visual, de tal manera de poder identificar estaciones con información no coherente, que pudieran resultar considerables en la modelación de la temperatura espacial, en ese sentido se han encontrado estaciones que presentan información con saltos como se puede apreciar en la estación Bambamarca, estos periodos atípicos fueron retirados, también se observaron estaciones con carencia de información como se aprecia en la estación Tarata, no se ha realizado completación de datos, ante la carencia de información la estación no era considerada en el modelo de regresión multivariante.



**Figura 4.1.** Variación temporal de las temperaturas en las estaciones Tarata y Bambamarca. Se puede apreciar carencia de información en ciertos periodos como saltos en la media.

A partir de la ubicación espacial de las estaciones se extraen los valores de las variables explicativas iniciales (latitud, longitud, altitud, seno, coseno) y se realizó el análisis exploratorio multivariante para identificar valores atípicos. Además, se realizó la prueba de dependencia efectiva para observar si los datos están correlacionados, obteniéndose un valor de 0.46, lo cual indica que los datos se encuentran correlacionados, para convertir los datos en variables incorrelacionadas se utilizaron los componentes principales.



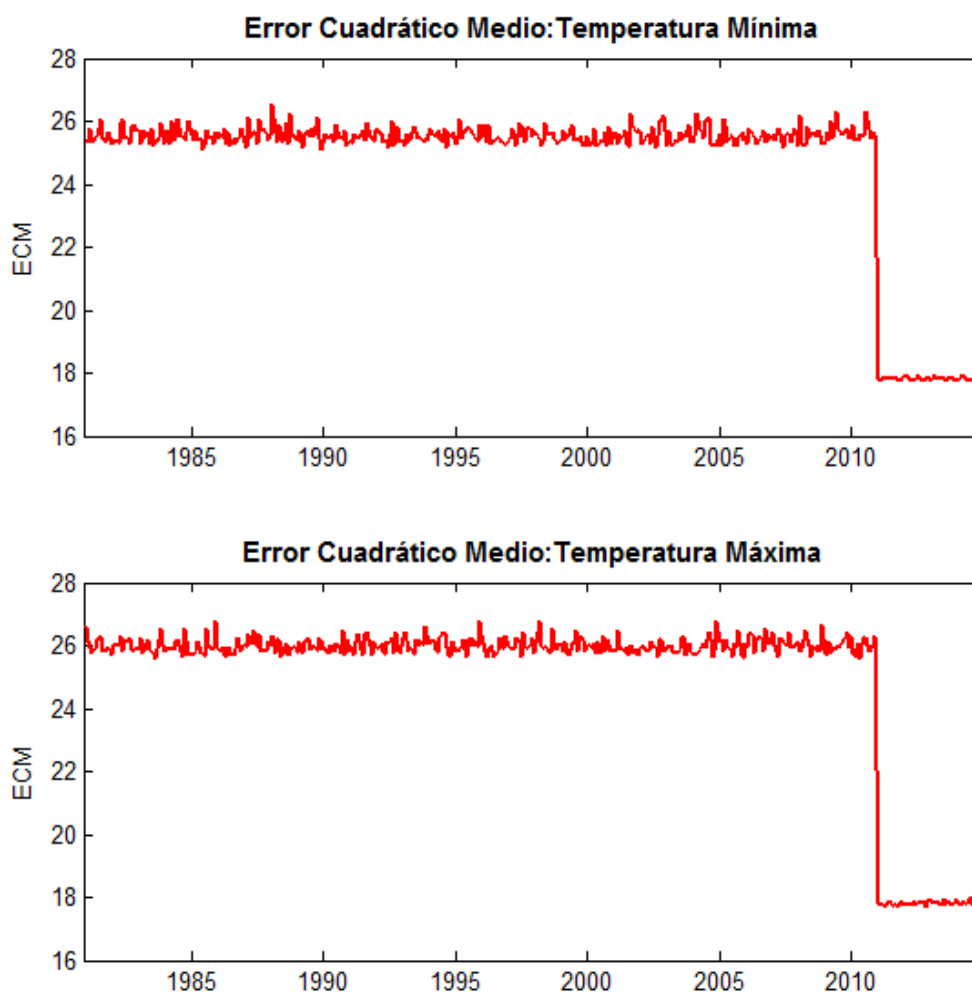
**Figura 4.2.** a) Distancia de Mahalanobis para los datos de las variables originales b) Distancia de Mahalanobis para los datos de los componentes principales.

De acuerdo a la Figura 4.2 para los datos de las variables originales se aprecian cuatro valores atípicos, correspondientes a las estaciones Tamshiyacu, Genaro Herrera, Punta Ático, Cañete. Estas estaciones no han sido retiradas del análisis, las estaciones mencionadas representan comportamientos distintos a las demás estaciones por lo cual presentan el comportamiento atípico en relación a las demás estaciones, así se tiene a las estaciones Tamshiyacu y Genaro Herrera, ambas ubicadas en la región de Loreto, además de ser las únicas en dicha región, por lo cual no se puede colegir que el comportamiento en esas estaciones en cuanto a la variabilidad de la temperatura es distinto al resto de las estaciones, por otro lado las estaciones de Punta Ático se encuentra ubicada en la zona muy costera, al igual que la estación Cañete, por lo cual el comportamiento distinto al resto de las estaciones. Con los datos de componentes principales la estación atípica sería la de la Mallares,

ubicada en la región de Piura, sin embargo, la otra estación debajo de ella en cuanto a la distancia de Mahalonabis es la estación de Ayabaca también ubicada en la misma región.

#### 4.2 Validación cruzada

La validación cruzada fue realizada dejando una estación fuera del análisis para generar el modelo de regresión multivariante y fue evaluada la eficiencia mediante el error medio cuadrático, los cuales se muestran en la Figura 4.3.



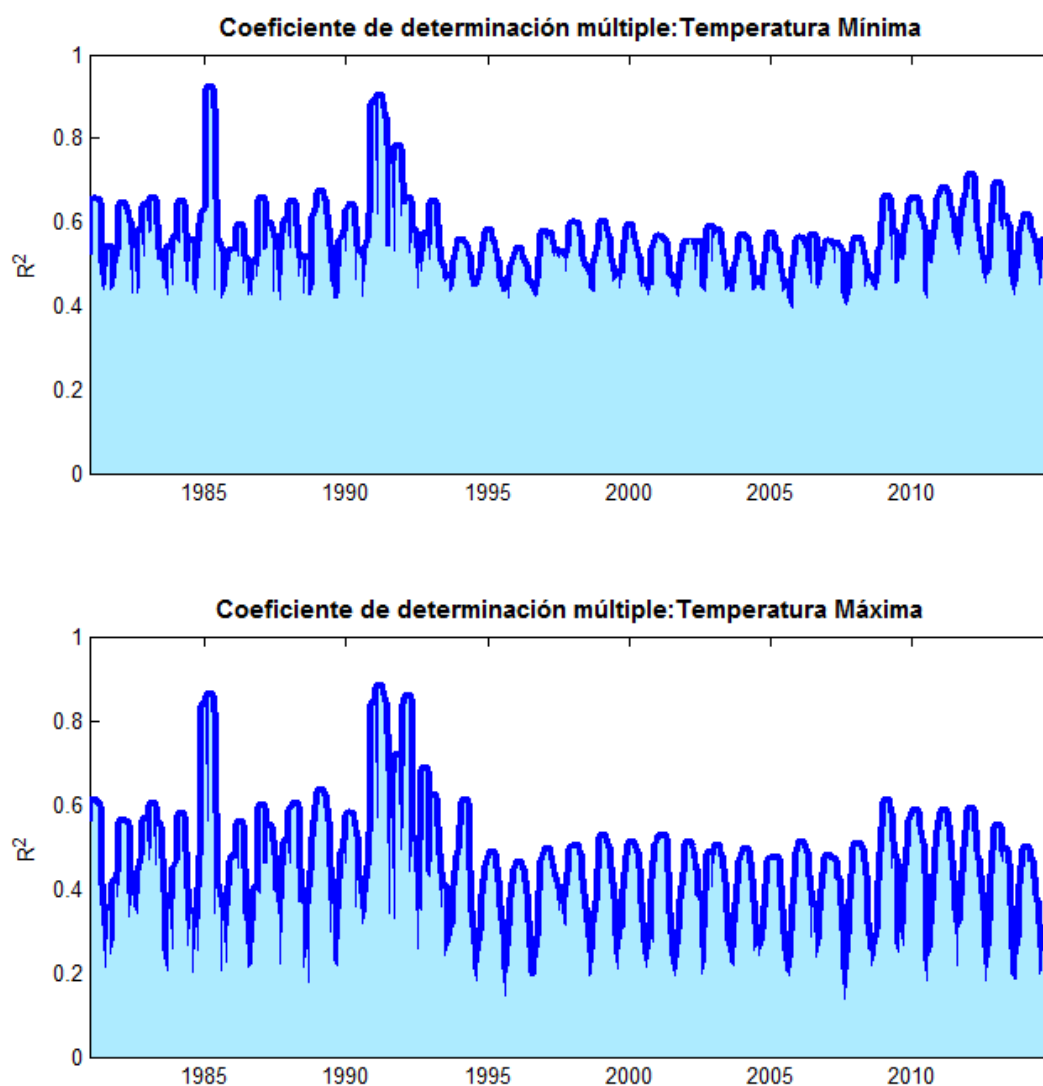
**Figura 4.3.** Variación temporal del error cuadrático medio de las temperaturas mínimas y máximas.

Como se puede apreciar en la Figura 4.3 existe un decaimiento del error cuadrático medio para ambas variables analizadas (temperatura mínima y máxima), desde 1981 al 2010 se aprecia un ECM alrededor del 26, a partir de enero del 2011 hasta el 2014 el valor fluctúa en 18, esto se debe a la cantidad de estaciones que se utilizaron en estos dos periodos mencionados (ver Figura 2.2), a partir del 2011 la cantidad de estaciones utilizadas se incrementaron considerablemente desde alrededor de 100 a

más de 200 estaciones. Por lo que se aprecia que a mayor cantidad de estaciones el método de regresión multivariante disminuye el error medio cuadrático.

### 4.3 Coeficiente de determinación múltiple

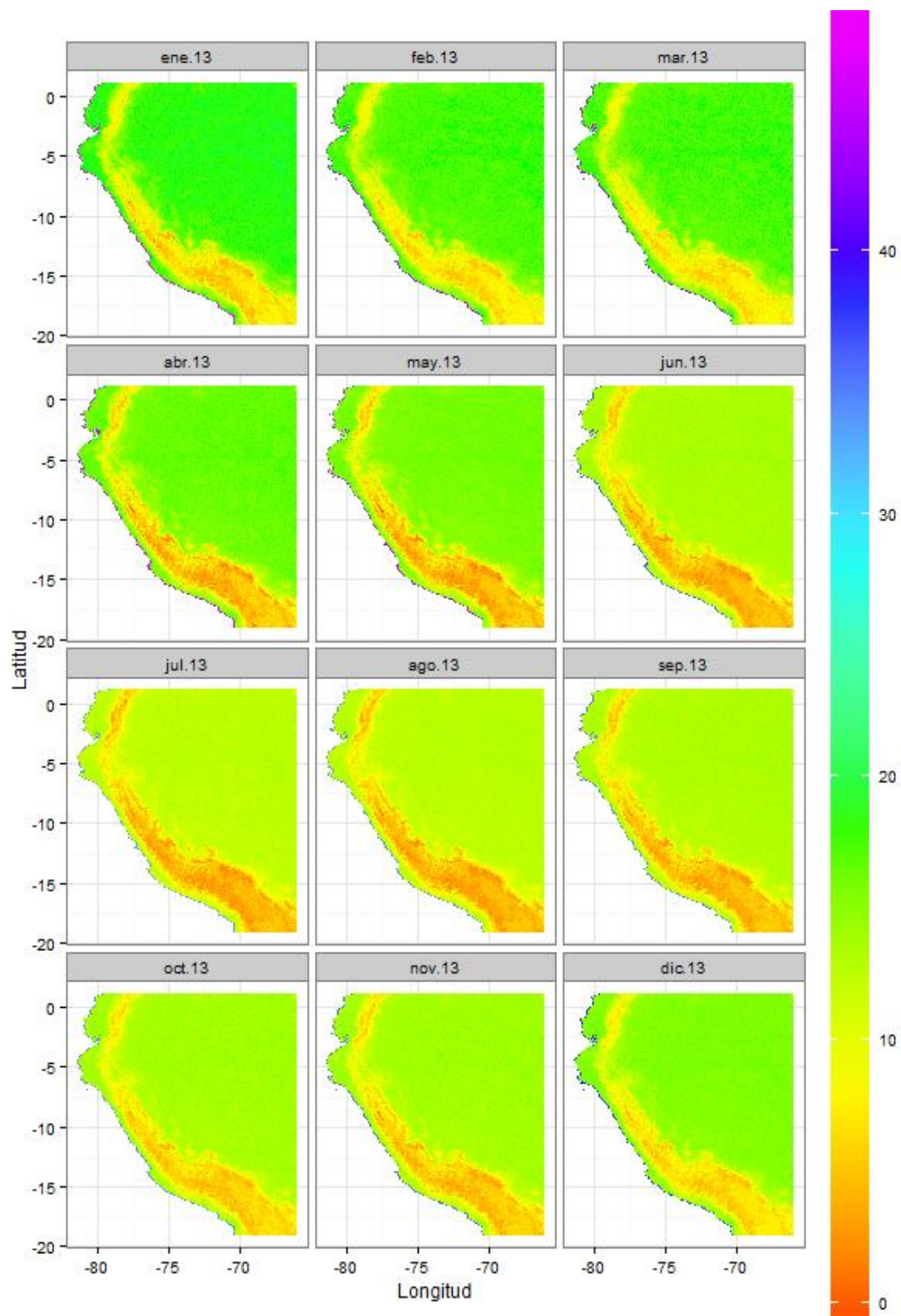
El coeficiente de determinación múltiple fue obtenido para cada mes y año de la serie 1981-2014 de acuerdo a las regresiones multivariantes generadas, las cuales se muestran en la Figura 4.4. Se puede apreciar que los modelos de regresiones presentan mejores resultados para explicar la temperatura mínima, ya que el valor del coeficiente de determinación ( $R^2$ ) se encuentra por encima de 0.5 en promedio, para la temperatura máxima se aprecia que el  $R^2$  en promedio es de 0.4.



**Figura 4.4.** Variación temporal del coeficiente de determinación múltiple de las regresiones multivariantes generadas para la temperatura mínima y máxima.

#### 4.4 Datos grillados de temperatura (1981-2014)

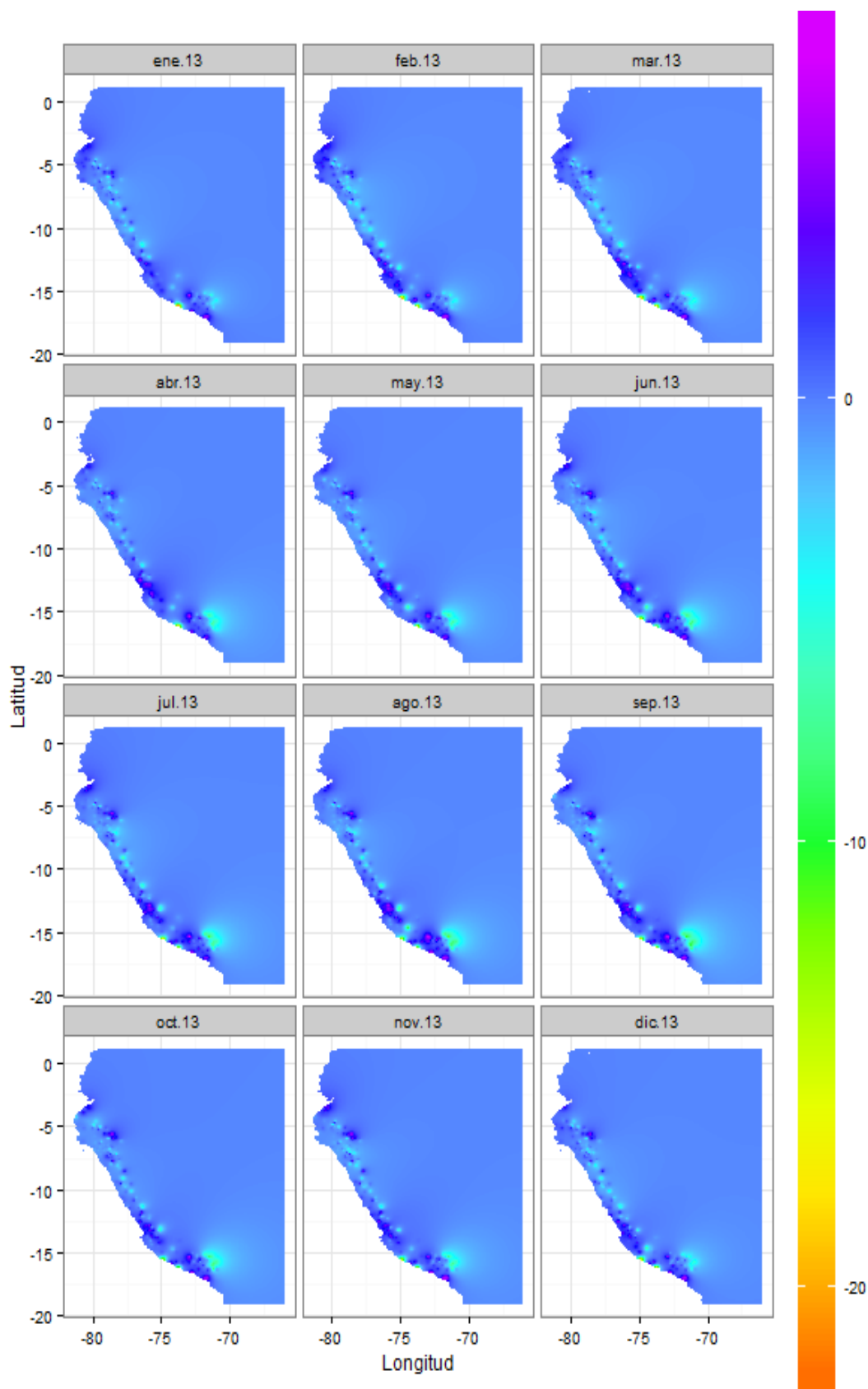
Con los modelos de regresión multivariante obtenidos para cada mes y año se procede a obtener los valores de temperaturas en cada una de las celdas del ámbito de estudio (ver Figura 2.1).



**Figura 4.5.** Variación espacial de la temperatura mínima (°C) obtenida del modelo de regresión multivariante para los meses del año 2013.

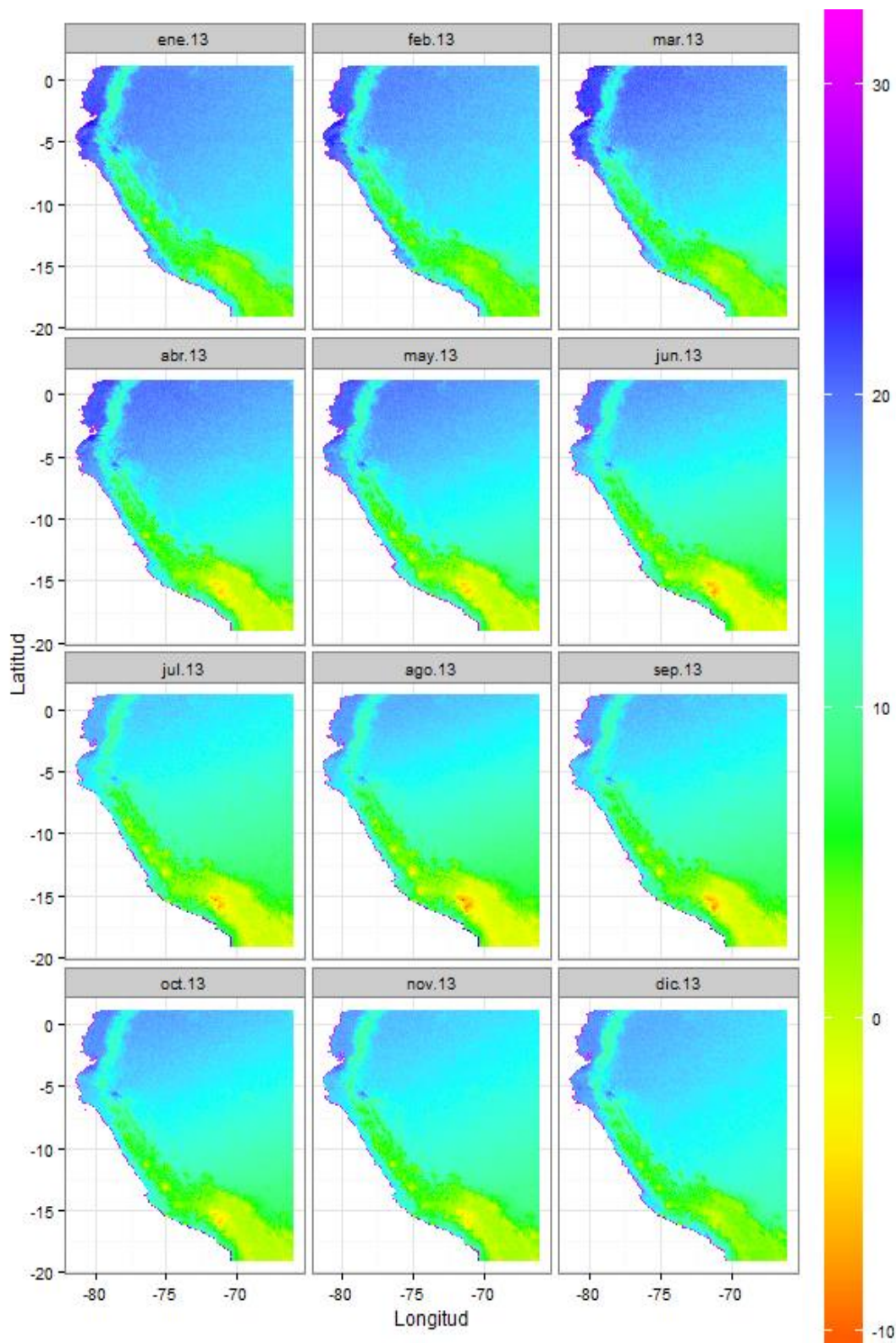
En la Figura 4.5 a manera de ejemplo se muestra la variación espacial de la temperatura mínima para los meses del año 2013, obtenida a partir de los modelos de regresión multivariante para los meses del 2013. Se puede apreciar temperaturas cálidas en la zona costera, como también en la zona de la selva, las menores temperaturas se aprecian en la región andina, los mapas se aprecian bien influenciados por la variación altitudinal del relieve del terreno, de acuerdo a la Figura 3.1 se aprecia que el primer componente principal se encuentra bastante influenciado por el relieve del terreno, además que la primera componente explica más del 90 % de la varianza total, por lo cual los resultados de la regresión multivariante es de esperarse que este bastante marcada por el componente de elevación.

Una vez obtenido los valores espacializados de las regresiones multivariantes, se obtuvieron los errores de estimación en las ubicaciones de las estaciones (las cuales cuentan con información), estas fueron espacializados mediante interpolación IDW, en la Figura 4.6 se muestran los errores espacializados del año 2013. Se puede apreciar que los modelos de regresiones multivariantes sobreestiman, ya que los errores obtenidos son por lo general negativos, la interpolación se ve bastante influenciada de acuerdo a la ubicación de las estaciones (donde existe información), donde se ven las mayores variaciones en relación a distancias, en la región de la selva la interpolación se suaviza, pero es debido a la carencia de estaciones en dicha zona.



**Figura 4.6.** Variación espacial de los errores interpolados mediante la técnica IDW de la temperatura mínima para los meses del año 2013.

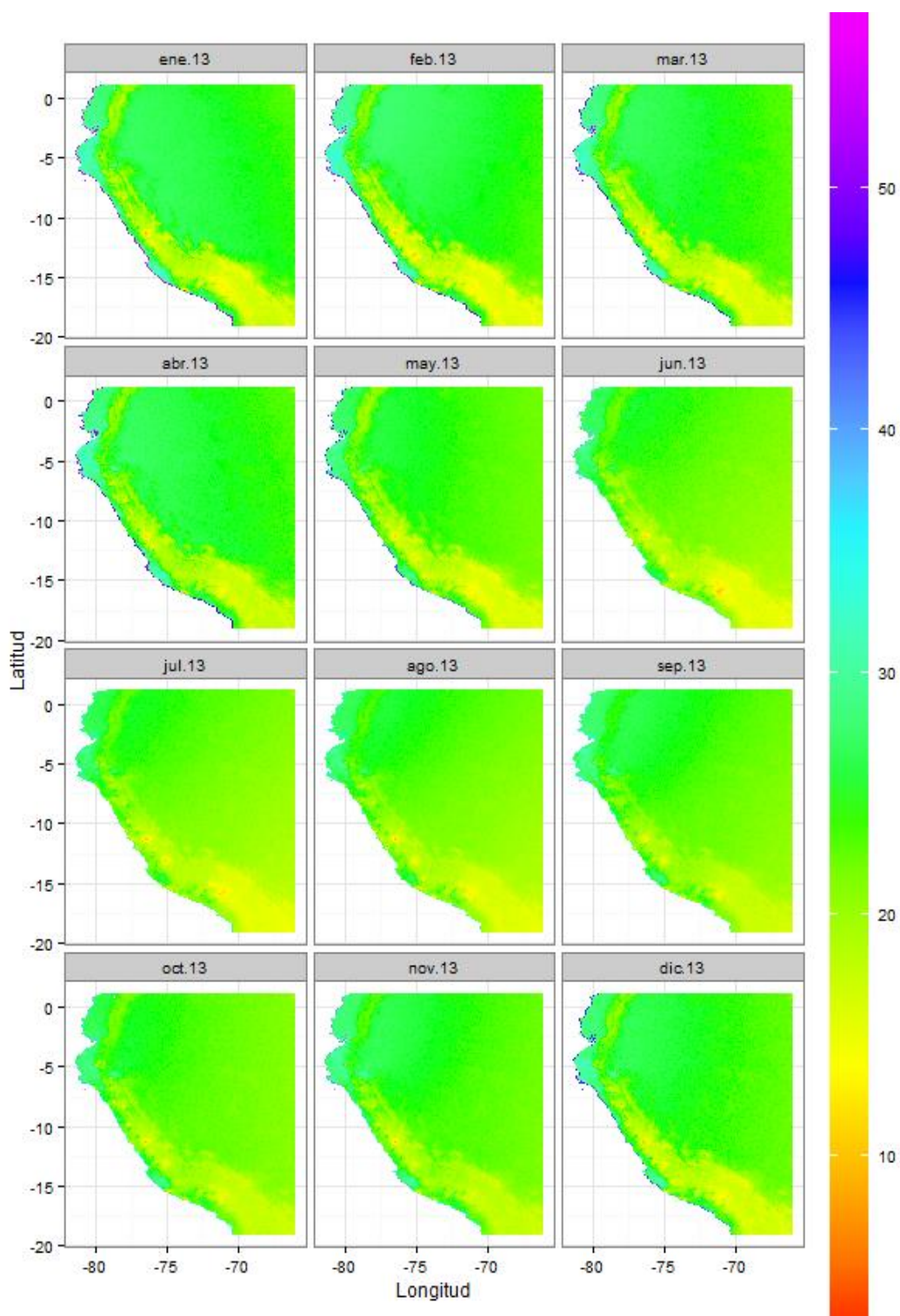
Con los mapas de temperaturas obtenidos mediante las regresiones y los mapas de errores espacialmente distribuidos, los mapas de temperaturas máximas y mínimas son obtenidos mediante una suma algebraica, en la Figura 4.7 se aprecia la variabilidad de la temperatura mínima para el año 2013 y en la Figura 4.8 la variabilidad de la temperatura máxima para el año 2013.



**Figura 4.7.** Variación espacial de la temperatura mínima (°C) para el año 2013.



En la Figura 4.7 a manera de ejemplo se muestra la variación espacial de la temperatura mínima para los meses del año 2013, obtenida de la suma algebraica del modelo de regresión multivariante y los errores de estimación. Se puede apreciar temperaturas cálidas en la zona costera, como también en la zona de la selva, las menores temperaturas se aprecian en la región andina, los mapas ya no se aprecian tan influenciados por la elevación del terreno, presentándose los menores valores en la zona del altiplano.



**Figura 4.8.** Variación espacial de la temperatura máxima (°C) para el año 2013.

En la Figura 4.8 a manera de ejemplo se muestra la variación espacial de la temperatura máxima para los meses del año 2013, obtenida de la suma algebraica del modelo de regresión multivariante y los errores de estimación. Se aprecia una distribución similar a la de la temperatura mínima, valores demasiado altos en el mapa (40°-50°) se deben a la zona del mar del pacífico (en esta zona no existen datos por lo cual estas estimaciones no son consideradas).

Los mismos procedimientos fueron realizados para toda la serie de tiempo analizada, obteniéndose así mapas de temperatura mínima y máxima para cada mes del periodo 1981-2014. En la Figura 4.9 se aprecia información generada para los años 1981, 1982, 1999, 2000, 2013, 2014.

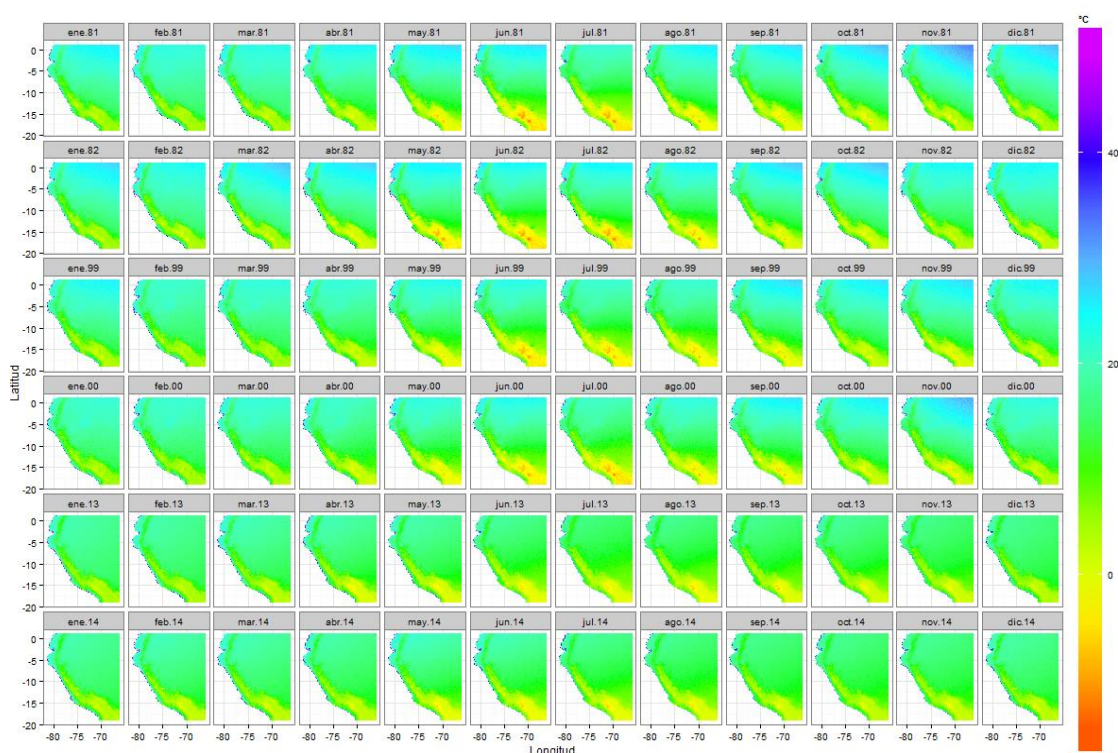


Figura 4.9. Datos grillados de temperatura mínima (°C) a nivel nacional (1981-2014).

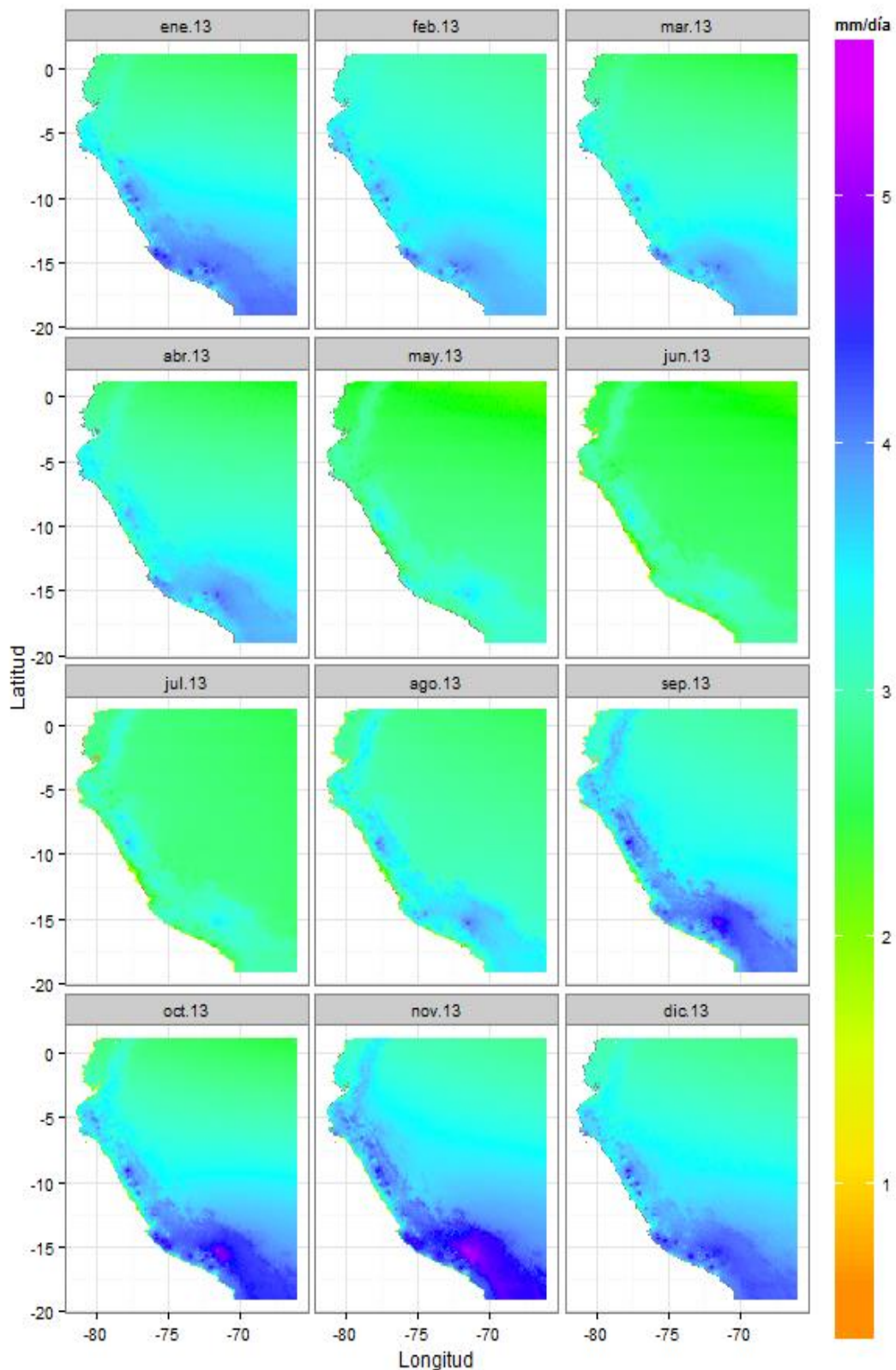
#### 4.5 Datos grillados de evapotranspiración potencial (1981-2014)

Una vez obtenidos los mapas de temperatura mínima y máxima se obtuvieron los mapas de evapotranspiración potencial usando el método de Hargreaves y Samani para cada mes para el periodo 1981-2014. En la Figura 4.10 se pueden apreciar los mapas para el año 2013. Previamente fueron obtenidos los mapas de Rs (radiación solar incidente), los cuales se muestran en la Figura 4.11 para el año 2013.

Los mapas de Rs nos muestran un comportamiento estacional, presentando los mayores valores en los meses de Septiembre a Enero y en las zonas de la sierra sur del territorio peruano y la región del altiplano.

Los mapas de evapotranspiración potencial (ETP) nos muestran un comportamiento estacional, presentando las mayores demandas en los meses de verano (Enero a Marzo). Se puede apreciar una zona bien diferenciada en la región de Ica, donde se

presentan las mayores demandas de ETP, al igual, que en la zona costera de la región de Piura.



**Figura 4.10.** Variación espacial de la radiación solar incidente (mm/día) para el año 2013.

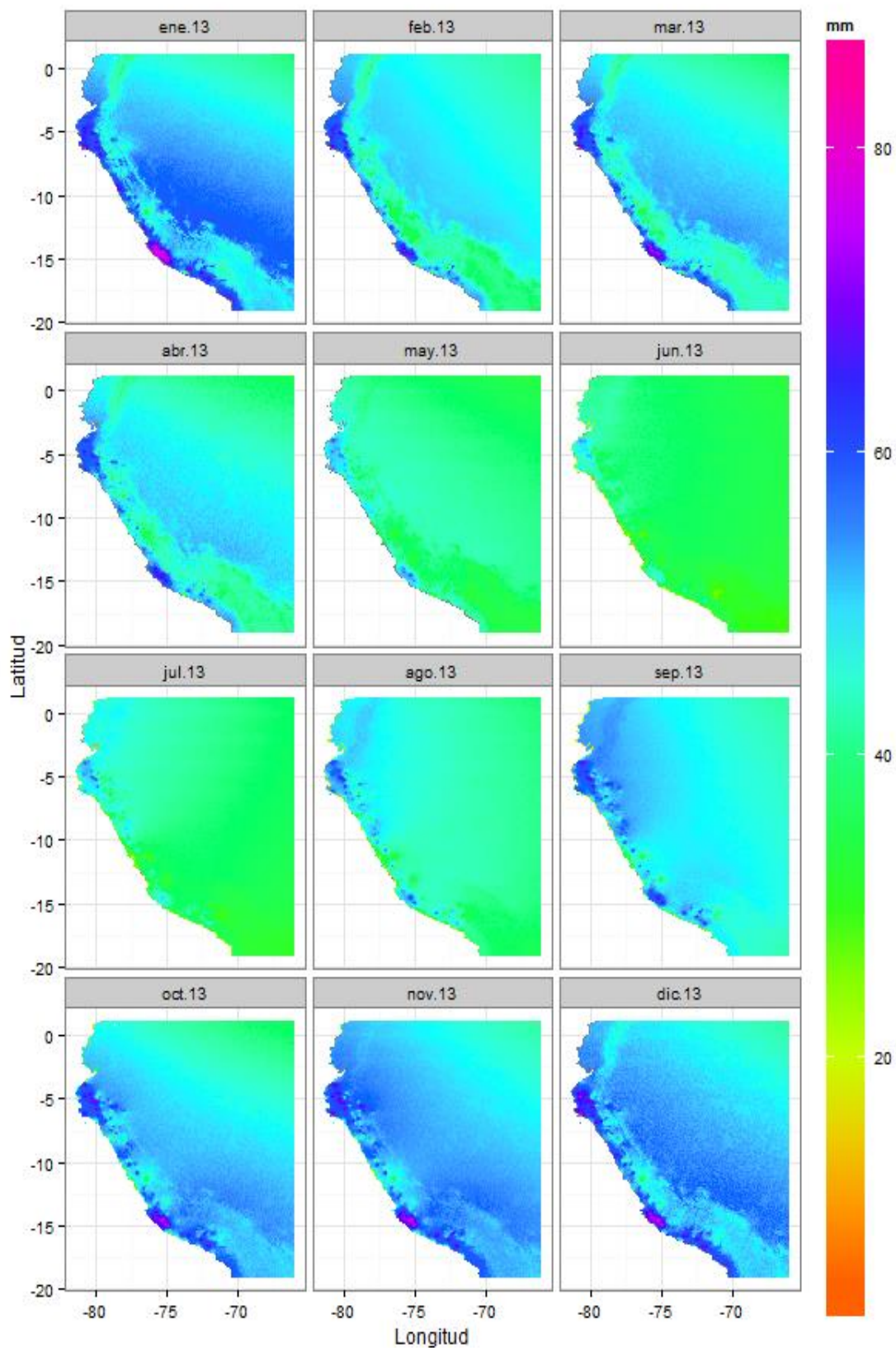


Figura 4.11. Variación evapotranspiración potencial (mm/mes) para el año 2013.

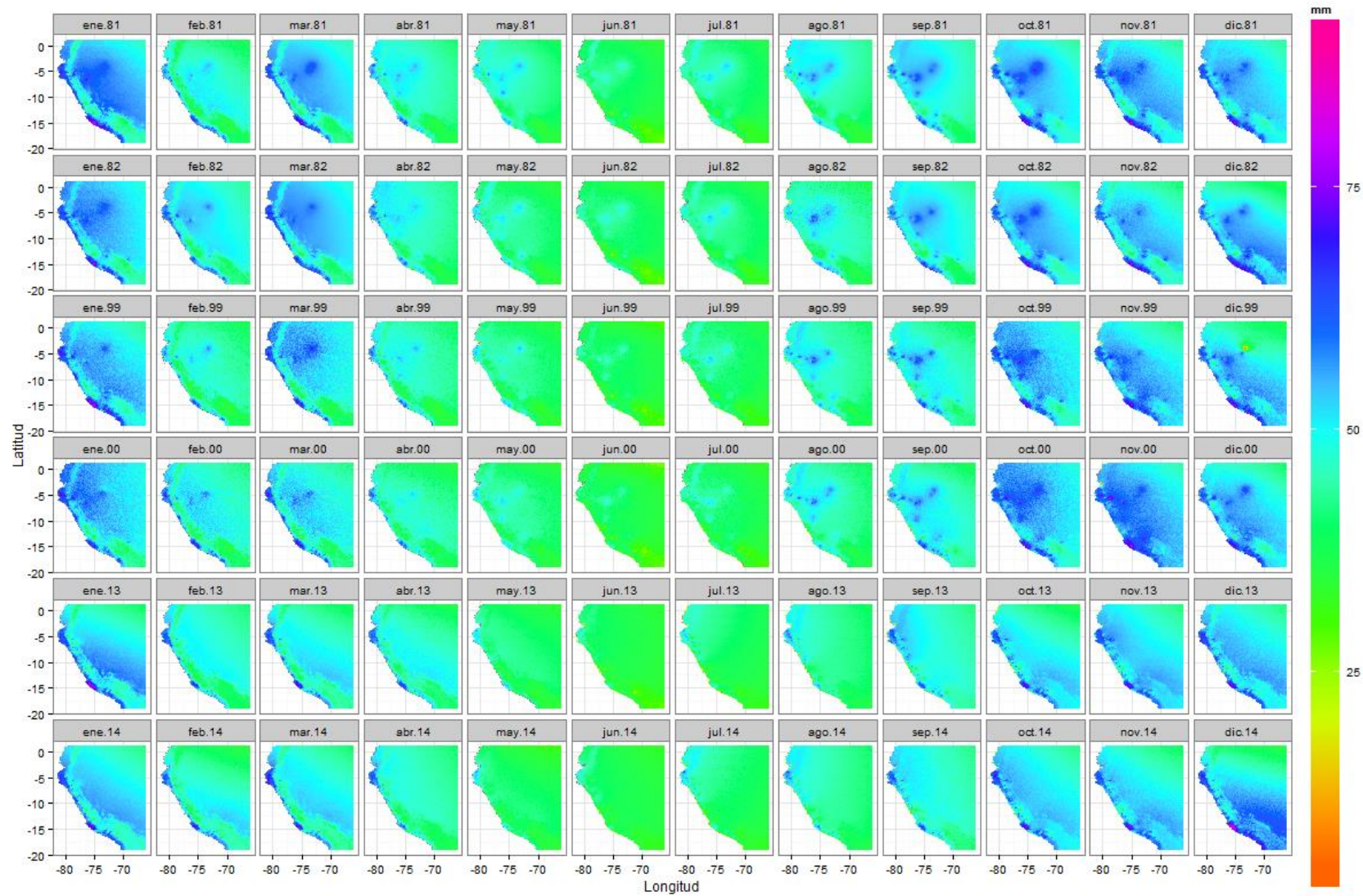


Figura 4.12. Datos grillados de evapotranspiración potencial (mm/mes) a nivel nacional (1981-2014)

Los mismos procedimientos descritos anteriormente fueron realizados para toda la serie de tiempo analizada, obteniéndose así mapas de evapotranspiración potencial para cada mes del periodo 1981-2014. En la Figura 4.12 se aprecia información generada para los años 1981, 1982, 1999, 2000, 2013, 2014.

Los datos grillados generados han sido almacenados en formato GEOTIFF, de tal manera que puedan ser leídos y editados mediante programas de sistemas de información geográfica.

## V. CONCLUSIONES

- Se han implementado los procesos para la generación de datos grillados de temperatura mínima, máxima y evapotranspiración potencial a nivel nacional, estas variables nos permiten explicar la variabilidad espacio-temporal de las variables mencionadas.
- Se ha adoptado el uso de regresiones multivariantes basadas en variables explicativas de condiciones de ubicación (latitud, longitud, elevación) y de relieve (seno y coseno de la inclinación del terreno) para explicar la variabilidad de las temperaturas, se ha utilizado la técnica de análisis de componentes principales para obtener variables incorrelacionadas.
- Los errores obtenidos de las regresiones han sido interpolados mediante IDW, los mapas finales de temperaturas corresponden a la suma algebraica de las regresiones y los errores. La representación de los mapas es a una resolución espacial de  $0.05^\circ$  y resolución temporal mensual. Como en todo modelo, la calidad de los resultados en una zona determinada está influenciada de manera importante a la calidad y cantidad de información observada disponible.
- Los modelos de regresiones multivariantes se presenta robusto en el proceso de validación (realizada mediante validación cruzada), los coeficiente de determinación nos muestran que los modelos explican de mejor manera la temperatura mínima, donde el  $R^2$  se encuentra en promedio para la serie 1981-2014 por encima de 0.5, mientras para la temperatura máxima este indicador estadístico disminuye a 0.4.
- Debido a la disponibilidad de información la evapotranspiración fue obtenida mediante un modelo de balance de energía como el de Hargreaves y Samani.
- No se cuenta con información observada para validar los mapas de evapotranspiración potencial generados, por lo cual se recomienda para futuros trabajos la validación de estos mapas con otras metodologías de obtención de la EP, además de información satelital, como por ejemplo con el producto MOD 16 - Evapotranspiration del sensor MODIS, estas comparaciones van a permitir identificar patrones de comportamientos de la variable ETP e identificar si los mapas generados estiman de manera adecuada la variabilidad de dicha variable.

## VI. BIBLIOGRAFÍA

- Allen, R.G.; L. S. Pereira y D. Raes (1998).- Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56 Disponible en Internet en: <http://www.fao.org/docrep/X0490E/X0490E00.htm#Contents>.
- Brutsaert, W., Parlange, M.B., 1998. Hydrological cycle explains the evaporation paradox. *Nature* 396, 30.
- Fleming MD, Chapin FS, Cramer WP, Hufford GL, Serreze MC. 2000. Geographic patterns and dynamics of Alaskan climate interpolated from a sparse station record. *Global Change Biology* 6: 49–58.
- Goovaerts P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology* 228: 113–129.
- Golubev, V.S., Lawrimore, J.H., Groisman, P.Y., Speranskaya, N.A., Zhuravin, S.A., Menne, M.J., Peterson, T.C., Malone, R.W., 2001. Evaporation changes over the contiguous United States and the former USSR: a reassessment. *Geophysical Research Letters* 28, 2665–2668.
- Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. *Applied Eng. in Agric.*, 1(2): 96-99.
- Hudson G, Wackernagel H. 1994. Mapping temperature using kriging with external drift: theory and an example from Scotland. *International Journal of Climatology* 14: 77–91.
- Hutchinson MF. 1995a. Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems* 9: 385–403.
- Hutchinson MF. 1995b. A new objective method for spatial interpolation of meteorological variables from irregular networks applied to the estimation of monthly mean solar radiation, temperature, precipitation and windrun. CSIRO Division of Water Resources Technical Memo 89/5; 95–104.
- Martinez-Cob A. 1996. Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *Journal of Hydrology* 174: 19–35.
- New MG, Lister D, Hulme M, Makin I. 2002. A high-resolution data set of surface climate over global land areas. *Climate Research* 21: 1–25.
- Ohmura, A., Wild, M., 2002. Climate change: is the hydrological cycle accelerating? *Science* 298, 1345–1346.
- Peterson, T.C., Golubev, V.S., Groisman, P.Ya., 1995. Evaporation losing its strength. *Nature* 377, 687–688.
- Phillips DL, Dolph J, Marks D. 1992. A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agricultural and Forest Meteorology* 58: 119–141.
- Roderick, M.L., Farquhar, G.D., 2002. The cause of decreased pan evaporation over the past 50 years. *Science* 298, 1410–1411.