

**UNIVERSIDAD NACIONAL AGRARIA  
LA MOLINA**

**FACULTAD DE CIENCIAS**



**“MODELO EMPÍRICOS PARA PRONÓSTICO DE  
TEMPERATURA MÍNIMAS EN PERIODO DE INICIO  
Y FIN DE CAMPAÑA AGRÍCOLA”**

Presentada por:

**Bremilda Andrea Sutizal Sánchez**

Tesis para Optar el Título Profesional de:

**INGENIERO METEORÓLOGO**

Lima – Perú

**2017**

**UNIVERSIDAD NACIONAL AGRARIA**

**LA MOLINA**

**FACULTAD DE CIENCIAS**

**“MODELO EMPÍRICOS PARA PRONÓSTICO DE  
TEMPERATURA MÍNIMAS EN PERIODO DE INICIO  
Y FIN DE CAMPAÑA AGRÍCOLA”**

Presentada por:

**Bremilda Andrea Sutizal Sánchez**

Tesis para Optar el Título Profesional de:

**INGENIERO METEORÓLOGO**

Sustentada y aprobada por el siguiente jurado:

---

Mg. Sc. Ever Menacho Casimiro  
PRESIDENTE

---

Ing. Franklin Unsihuay Tovar  
MIEMBRO

---

Mg. Sc. Vidal Villagomez Castillo  
MIEMBRO

---

Mestre. Jerónimo García Villanueva  
ASESOR

## **AGRADECIMIENTOS**

El presente trabajo de tesis se desarrolló gracias a la contribución de conocimiento y tiempo de mi asesor Mg. Sc. Jerónimo García Villanueva, a quien expreso un especial agradecimiento. Así mismo, quiero manifestar mi agradecimiento infinito a mis padres, por todo su apoyo. También, quiero agradecer a mis compañeras de la carrera de Meteorología, quienes de algún modo me ayudaron a realizar esta investigación y por estar pendientes de mí, a pesar de la distancia.

# ÍNDICE GENERAL

<b>I. INTRODUCCIÓN .....</b>	<b>1</b>
<b>II. REVISIÓN DE LITERATURA .....</b>	<b>3</b>
2.1. DESCRIPCIÓN DEL VALLE DEL MANTARO .....	3
2.1.1. VARIACIÓN MENSUAL DE LA TEMPERATURA MÍNIMA .....	4
2.1.2. ANÁLISIS DE HELADAS.....	5
2.1.3. CLIMA Y AGRICULTURA .....	6
2.2. TEMPERATURA MÍNIMA .....	7
2.2.1. HELADAS .....	8
2.3. PRONÓSTICO METEOROLÓGICO .....	11
2.3.1. CONFIABILIDAD DE UN PRONÓSTICO NUMÉRICO DEL TIEMPO .....	12
2.3.2. PREDICCIÓN NUMÉRICA DEL TIEMPO (NWP) .....	12
2.4. BRAZILIAN REGIONAL ATMOSPHERIC MODELING SYSTEM (BRAMS) .	18
2.4.1. CARACTERÍSTICAS DEL MODELO BRAMS.....	19
2.4.2. CONSIDERACIONES PARA LA EJECUCIÓN DEL MODELO BRAMS ....	23
2.5. ANÁLISIS EXPLORATORIO DE DATOS (A.E.D.) .....	24
2.5.1. DATOS ATÍPICOS (OUTLIERS) .....	25
2.5.2. DISTRIBUCIÓN NORMAL DE DATOS .....	26
2.6. MODELO DE REGRESIÓN LINEAL .....	28
2.6.1. ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO.....	29
2.6.2. EL COEFICIENTE DE CORRELACIÓN LINEAL Y EL COEFICIENTE ....	30
DE DETERMINACIÓN .....	30
2.6.3. ANÁLISIS DE RESIDUOS.....	32
2.7. ANTECEDENTES .....	34
<b>III. MATERIALES Y MÉTODOS .....</b>	<b>36</b>
3.1. ÁREA DE ESTUDIO .....	36

3.2. DATOS Y SOFTWARES .....	37
3.3. METODOLOGÍA .....	37
3.3.1. PARA LA DETERMINACIÓN DE LOS MODELOS EMPÍRICOS.....	37
3.3.2. PARA LA DETERMINACIÓN DE LA DISTRIBUCIÓN ESPACIAL .....	49
<b>IV. RESULTADOS Y DISCUSIONES.....</b>	<b>50</b>
4.1. DETERMINACIÓN DE LOS MODELOS EMPÍRICOS DE PRONÓSTICO DE TEMPERATURAS MÍNIMAS .....	50
4.1.1. ANÁLISIS EXPLORATORIO DE LOS DATOS.....	50
4.1.2. ANÁLISIS DE NORMALIDAD Y LINEALIDAD DE LOS DATOS UTILIZADOS .....	56
4.1.3. MODELOS EMPÍRICOS DE PRONÓSTICO.....	66
4.2.3. VALIDACIÓN DEL MODELO EMPÍRICO Y ANÁLISIS DE RESIDUOS..	69
4.2. DETERMINACIÓN DE LA DISTRIBUCIÓN ESPACIAL .....	73
<b>V. CONCLUSIONES .....</b>	<b>74</b>
<b>VI. RECOMENDACIONES.....</b>	<b>75</b>
<b>VII. REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>77</b>
<b>VIII. ANEXO .....</b>	<b>80</b>

## ÍNDICE DE TABLAS

<b>Tabla 1:</b> Ecuaciones modelos empíricos para pronóstico de temperaturas mínimas en inicio de campaña agrícola para el tiempo de pronóstico dos (T2). .....	66
<b>Tabla 2:</b> Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico dos (T2). .....	66
<b>Tabla 3:</b> Ecuaciones de modelos empíricos para pronóstico de temperatura mínima en inicio de campaña agrícola para el tiempo de pronóstico tres (T3). .....	67
<b>Tabla 4:</b> Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico tres (T3). .....	67
<b>Tabla 5:</b> Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en inicio de campaña agrícola para el tiempo de pronóstico cuatro (T4). .....	67
<b>Tabla 6:</b> Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico cuatro (T4). .....	67
<b>Tabla 7:</b> Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en inicio de campaña agrícola para el tiempo de pronóstico cinco (T5). .....	68
<b>Tabla 8:</b> Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico cinco (T5). .....	68
<b>Tabla 9:</b> Resultados de la validación de los modelos empíricos para los tiempos de pronóstico (T2, T3 y T4) en ambas campañas agrícolas. ....	69
<b>Tabla 10:</b> Resultados del análisis cuantitativo de los residuos para los tiempos de pronóstico (T2 y T4) en ambas campañas agrícolas. ....	73

## ÍNDICE DE FIGURAS

<b>Figura 1:</b> Ubicación del valle del Mantaro.....	3
<b>Figura 2:</b> Panorama económico de la producción de papa en Perú.....	4
<b>Figura 3:</b> Variabilidad mensual de las temperaturas mínimas en el valle del Mantaro. Periodo común: enero 1992 a diciembre 2009, excepto para Viques de enero 1998 a diciembre 2009 .....	5
<b>Figura 4:</b> Tipos de heladas .....	9
<b>Figura 5:</b> Símbolos de las ecuaciones que usa el modelo BRAMS .....	21
<b>Figura 6:</b> Ejemplo del anidamiento de rejillas en el modelo BRAMS.....	22
<b>Figura 7:</b> Flujo de ejecución de BRAMS .....	24
<b>Figura 8:</b> Gráficos para análisis de residuos .....	33
<b>Figura 9:</b> Mapa del Valle del Mantaro y ubicación de estaciones .....	36
<b>Figura 10:</b> Flujo grama básico del funcionamiento del Modelo BRAMS. ....	39
<b>Figura 11:</b> Configuración en tres dominios del modelo BRAMS.....	39
<b>Figura 12:</b> Representación de las características de un Box-Plot .....	41
<b>Figura 13:</b> Tipos de distribución según Diagrama de Cuantiles .....	42
<b>Figura 14:</b> Representación del análisis gráfico de residuos que cumple con los supuestos para un modelo de regresión lineal.....	47
<b>Figura 15:</b> Representación del análisis gráfico de residuos que incumple con el supuesto de linealidad.....	48
<b>Figura 16:</b> Representación del análisis gráfico de residuos que incumple con el supuesto de homocedasticidad.....	48
<b>Figura 17:</b> Diagrama de cajas de temperatura mínima de las cuatro estaciones de estudio para el mes de septiembre.....	51
<b>Figura 18:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de septiembre.....	52
<b>Figura 19:</b> Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de febrero.....	53
<b>Figura 20:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de febrero.....	55

<b>Figura 21:</b> Análisis de normalidad, de temperaturas mínimas de estación, para las cuatro estaciones de estudio en inicio de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2). .....	57
<b>Figura 22:</b> Análisis de normalidad, de temperaturas mínimas pronosticadas por BRAMS, para las cuatro estaciones de estudio en inicio de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2).....	58
<b>Figura 23:</b> Análisis de normalidad, de temperaturas mínimas de estación, para las cuatro estaciones de estudio en fin de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2). .....	60
<b>Figura 24:</b> Análisis de normalidad, de temperaturas mínimas pronosticadas por BRAMS, para las cuatro estaciones de estudio en fin de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2).....	61
<b>Figura 25:</b> Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico dos (T2) y en todas las estaciones de estudio. ....	62
<b>Figura 26:</b> Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico tres (T3) y en todas las estaciones de estudio. ....	63
<b>Figura 27:</b> Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico cuatro (T4) y en todas las estaciones de estudio. ....	64
<b>Figura 28:</b> Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico cinco (T5) y en todas las estaciones de estudio. ....	65
<b>Figura 29:</b> Análisis de residuos para el tiempo dos de pronóstico (T2), para las estaciones que pasaron satisfactoriamente la prueba Fisher y la validación. ....	71
<b>Figura 30:</b> Análisis de residuos para el tiempo cuatro de pronóstico (T4), para las estaciones que pasaron satisfactoriamente la prueba Fisher y la validación. ....	72
<b>Figura 31:</b> Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de octubre. ....	80
<b>Figura 32:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de octubre. ....	81
<b>Figura 33:</b> Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de noviembre. ....	81

<b>Figura 34:</b> Continuación de la figura 32.....	82
<b>Figura 35:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de noviembre. ....	82
<b>Figura 36:</b> Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de diciembre. ....	83
<b>Figura 37:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de diciembre. ....	83
<b>Figura 38:</b> Continuación de la figura 36.....	84
<b>Figura 39:</b> Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de marzo. ....	84
<b>Figura 40:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de marzo. ....	85
<b>Figura 41:</b> Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de abril.....	85
<b>Figura 42:</b> Continuación de la figura 40.....	86
<b>Figura 43:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de abril.....	86
<b>Figura 44:</b> Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de mayo. ....	87
<b>Figura 45:</b> Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de mayo. ....	87
<b>Figura 46:</b> Continuación de la figura 44.....	88
<b>Figura 47:</b> Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin_P) para el tiempo de pronóstico tres (T3), para inicio de campaña y en las cuatro estaciones de estudio. ....	89
<b>Figura 48:</b> Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin_P) para el tiempo de pronóstico cuatro (T4), para inicio de campaña y en las cuatro estaciones de estudio. ....	90
<b>Figura 49:</b> Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin_P) para el tiempo de pronóstico cinco (T5), para inicio de campaña y en las cuatro estaciones de estudio. ....	91
<b>Figura 50:</b> Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin_P) para el tiempo de pronóstico tres (T3), para fin de campaña y en las cuatro estaciones de estudio.....	92

**Figura 51:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico cuatro (T4), para fin de campaña y en las cuatro estaciones de estudio..... 93

**Figura 52:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico cinco (T5), para fin de campaña y en las cuatro estaciones de estudio. .... 94

## ÍNDICE DE ANEXOS

<b>ANEXO 1. UBICACIÓN DE LAS ESTACIONES METEOROLÓGICAS QUE SE TOMARON EN CUENTA EN LA INVESTIGACIÓN.....</b>	<b>78</b>
<b>ANEXO 2. ANÁLISIS EXPLORATORIO DE DATOS, DIAGRAMAS DE CAJAS Y SERIES DE TIEMPO.....</b>	<b>78</b>
<b>ANEXO 3. ANÁLISIS DE NORMALIDAD DE LOS DATOS QUE SE TOMARON EN CUENTA PARA LA DETERMINACIÓN DE LOS MODELOS EMPÍRICOS.....</b>	<b>86</b>
<b>ANEXO 4. ESTADÍSTICAS Y ANÁLISIS DE VARIANZA DE LOS MODELOS DE REGRESIÓN LINEAL.....</b>	<b>93</b>

## RESUMEN

Este trabajo de investigación tuvo como objetivo determinar modelos empíricos para pronóstico de temperaturas mínimas en el Valle del Mantaro, de hasta con cinco días de anticipación, para inicio (setiembre a diciembre) y fin (febrero a mayo) de campaña agrícola. Para lo cual, se ha usado datos (del período de 2012 a 2016) de pronóstico del modelo regional BRAMS 4.2 y datos de temperatura mínima de cuatro estaciones (Viques, Huayao, Ingenio y Jauja). Para generar los datos pronosticados el modelo fue configurado en tres dominios (72 km, 12km y 3km), en el estudio se usó el de 3km. Se realizaron pronósticos horarios - semanales y en la determinación de los modelos empíricos se tomó en cuenta los pronósticos del día dos al cinco (T2, T3, T4 y T5). Se verificó la asociación lineal entre las temperaturas mínimas del modelo BRAMS (variable independiente) y de estación (variable dependiente), se evaluaron las significancias de los modelos lineales de regresión simple mediante el análisis de varianza y el estadístico de Fisher al 95% de confianza, además se verificaron los modelos mediante la determinación del error cuadrático medio (RMSE), el Bias, la correlación de Pearson y el análisis de residuos. Se encontraron modelos significativos para los días de pronóstico dos, tres y cuatro. La validación mostró RMSE entre 1,5 y 2,7°C, y Bias entre  $\pm 2.0^{\circ}\text{C}$ . Así mismo las correlaciones fueron significativas hasta en 0,844 (especialmente en fin de campaña). Con el análisis de residuos se identificó heterocedasticidad y distribución de datos no normales en algunos modelos, y se concluyó que, los modelos empíricos de pronóstico de temperaturas mínimas, con dos días de anticipación para Huayao (en inicio de campaña agrícola) y de cuatro días de anticipación para Ingenio (en fin de campaña agrícola) pueden ser aplicados como modelos operativos.

**Palabras clave:** Modelo empírico, pronóstico, temperaturas mínimas, modelo BRAMS, regresión lineal.

## ABSTRACT

This research was performed in order to determine empirical models for forecasting minimum temperatures in the Valle del Mantaro, up to five days in advance, for the beginning (September to December) and the end (February to May) of the agricultural season. For this purpose, data from the regional model BRAMS 4.2 and minimum temperature data from Viques, Huayao, Ingenio and Jauja stations have been used (from 2012 to 2016). In order to generate the predicted data the model was configured in three domains (72 km, 12km and 3km), the study was performed with 3km. Weekly forecasts were made and in the determination of the empirical models, the forecasts were taken into account Two to five days (T2, T3, T4 and T5). The linear association between the minimum temperatures of the BRAMS (independent variable) and stations (dependent variable) was verified, the significance of linear simple regression models was evaluated by analysis of variance and the Fisher statistic test at 95% confidence level, in addition the models were verified by the determination of the mean square error (RMSE), the Bias, the Pearson correlation and the residue analysis. Significant models were found at the two, three and four forecast days. Validation showed RMSE between 1.5 and 2.7 ° C, and Bias between  $\pm 2.0$  ° C. Correlations were also significant up to 0.844 (especially in end of the agricultural season). With the residue analysis we identified heteroscedasticity and distribution of non-normal data in some models, finally concluded that empirical models of minimum temperature forecast, two days in advance for Huayao (at the beginning of the agricultural season) and four days in advance for Ingenio (at the end of the agricultural season) can be applied as operational models.

**Key words:** Empirical model, forecast, minimum temperatures, BRAMS model, linear regression.

## I. INTRODUCCIÓN

Desde la década de los ochenta los modelos de predicción numérica han ido evolucionando en complejidad y su desempeño cada vez es mejor, sin embargo en los últimos años su evolución no ha sido homogénea pues las previsiones presentan incertidumbres. Éstas pueden deberse a las condiciones iniciales, resolución del modelo, parametrización deficiente de los procesos físicos o de variables límite como la orografía, condiciones de superficie (albedo, vegetación, etc.) y condiciones de contorno. En las últimas décadas, debido al avance en los medios computacionales, modelos de mesoescala tales como el Weather Research and Forecasting (WRF) del National Centers for Environmental Prediction/ National Centers for Environmental Prediction (NCEP/NCAR), el ETA del NCEP y el Brazilian Regional Atmospheric Modeling System (BRAMS) se utilizan con el fin de generar pronósticos regionales en tiempo real. Los modelos numéricos constituyen hoy en día una herramienta indispensable para la comunidad meteorológica, tanto para fines operativos como de investigación. A la vez resulta indispensable realizar una verificación de los pronósticos numéricos con el fin de brindar a los usuarios una medida de su precisión y detectar las deficiencias principales de los mismos para así poder realizar las modificaciones que sean necesarias para mejorar la calidad del pronóstico.

Uno de los problemas que afronta el Perú es la gran demanda alimenticia, problema que se ve agravado por la escasez de suelos con aptitud agrícola, heterogeneidad fisiográfica, falta de tecnología y riesgos naturales y climáticos que limitan el crecimiento y desarrollo de los cultivos y reducen sus rendimientos. Uno de los elementos del clima vinculado con los riesgos climáticos, es la temperatura mínima extrema, la cual origina heladas. La helada es un factor meteorológico, que dependiendo de su intensidad, duración y la fase fenológica en que se encuentre el cultivo, puede provocarle graves daños, además de, mortalidad en la ganadería, mayor número de enfermedades respiratorias y pérdidas económicas. Son consideradas como uno de los fenómenos extremos que mayor impacto tiene en los andes peruanos. En el valle del Mantaro es el segundo factor meteorológico que causa preocupación en la población dedicada a la agricultura (IGP, 2005), debido a que, en esta zona la agricultura es el sector económico más importante, constituyéndose como una de las

principales fuentes para la seguridad alimentaria, para la región Junín y la ciudad de Lima (IGP, 2010).

Es entonces importante poder predecir cuándo cae la temperatura hasta un valor crítico, ya que, proporciona una oportunidad a los agricultores y la población en general de prepararse y evitar daños y pérdidas. En el Valle del Mantaro, se han determinado modelos en base a procesos físicos relacionados a conservación de energía como el de Cellier y modelos empíricos basados en mediciones de humedad relativa, temperaturas del bulbo seco y húmedo como el de García y Villegas (1993), para pronósticos con 18 horas de anticipación, pero ningún modelo empírico se ha determinado en base a un modelo numérico tal como el BRAMS o el WRF.

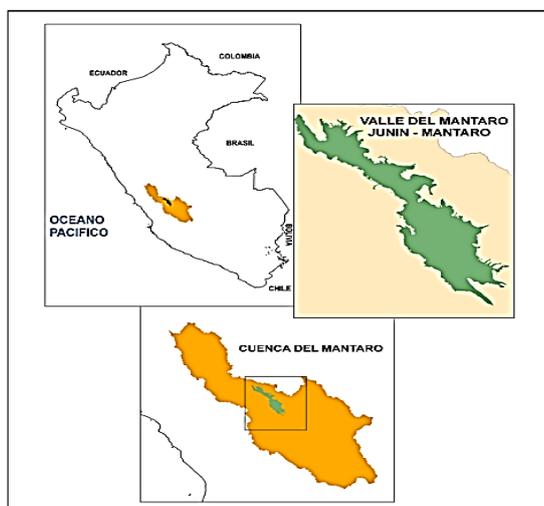
Por lo tanto nuestro interés de estudio fue establecer modelos empíricos para pronosticar temperaturas mínimas con cinco días de anticipación, para el Valle del Mantaro, en periodos de inicio y fin de campaña agrícola (setiembre-diciembre y febrero-mayo) de 2012-2016, utilizando datos registrados de 4 estaciones meteorológicas de SENAMHI y temperaturas mínimas pronosticadas por el modelo BRAMS 4.2. Para ello, se plantearon los siguientes objetivos específicos: 1) Determinar modelos de pronóstico de las temperaturas mínimas y 2) Determinar la distribución espacial de las temperaturas mínimas pronosticadas.

## II. REVISIÓN DE LITERATURA

### 2.1. DESCRIPCIÓN DEL VALLE DEL MANTARO

El valle del Mantaro, formado por el río del mismo nombre, se encuentra entre los paralelos  $11^{\circ} 40'$  y  $12^{\circ} 10'$  sur y los meridianos  $75^{\circ} 10'$  y  $75^{\circ} 15'$  oeste en la sierra central del Perú. Tiene una longitud de 70 km desde los alrededores de la ciudad de Jauja (al norte) hasta las localidades de Pucará y Huacrapuquio al extremo sur. Su ancho es variable con rangos entre 2 y 8 km, presenta altitudes entre 3100 y 3400 msnm, aproximadamente, y tiene una superficie total aproximada de 70 000 ha (Vega, 2005).

Se encuentra en el departamento de Junín (figura 01), ocupa hasta 60 capitales de distritos como parte de las provincias de Chupaca, Concepción, Huancayo y Jauja. El sector económico más importante es sin duda la agricultura. La sierra de Junín es uno de los más importantes productores de papa, también cultiva una proporción importante de maíz amiláceo, quinua, cebada, trigo, habas, entre otros. Tiene como grandes núcleos a las confluencias de los ríos Cunas, Achamayo, Shullcas y Chanchas, todos los cuales desembocan en el río Mantaro. El valle es una de las principales fuentes para la seguridad alimentaria, no solo para la región, sino también para la ciudad de Lima (IGP, 2010).

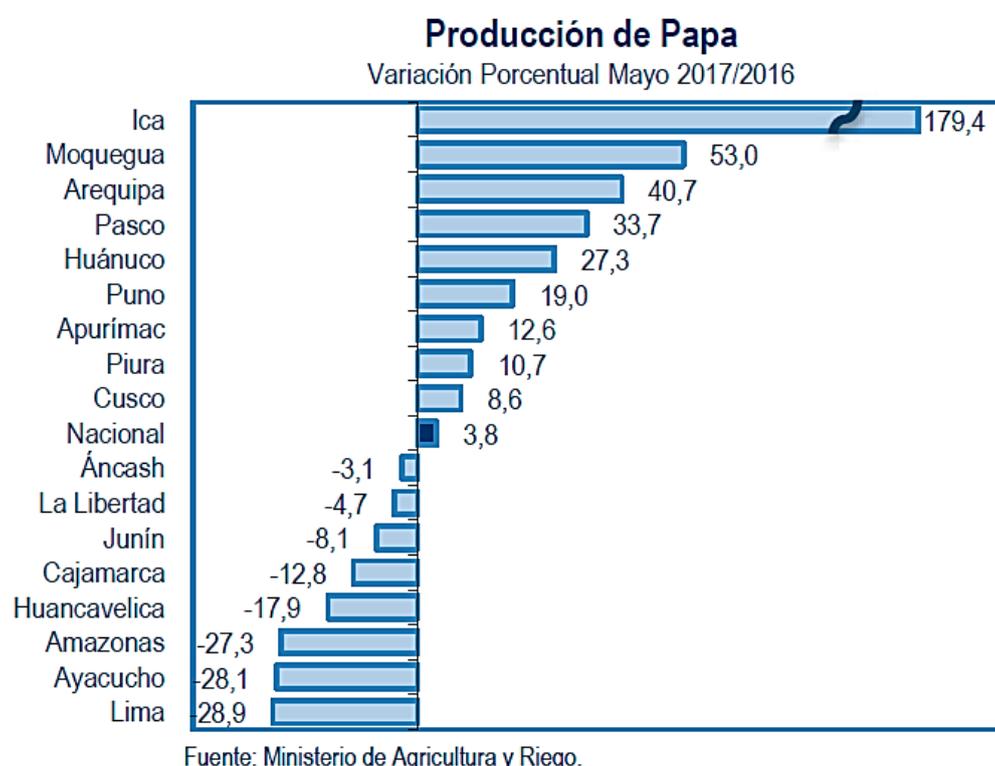


**Figura 1:** Ubicación del valle del Mantaro

FUENTE: IGP, 2010

## Panorama económico departamental para la producción de papa en Perú

La producción de papa, en mayo 2017, totalizó 1 millón 22 mil 148 toneladas y se incrementó en 3,8% respecto a igual mes de 2016, explicado por las temperaturas normales y lluvias ligeras que favorecieron la mayor producción en los departamentos de Huánuco (27,3%), Puno (19,0%), Apurímac (12,6%) y Cusco (8,6%), los cuales concentraron el 61,1% del total nacional. Igualmente, aumentó en Ica (179,4%), Moquegua (53,0%), Arequipa (40,7%), Pasco (33,7%) y Piura (10,7%). No obstante, se redujo en Lima (-28,9%), Ayacucho (-28,1%), Amazonas (-27,3%), Huancavelica (-17,9%), Cajamarca (-12,8%), Junín (-8,1%), La Libertad (-4,7%) y Áncash (-3,1%). (INEI, 2017).



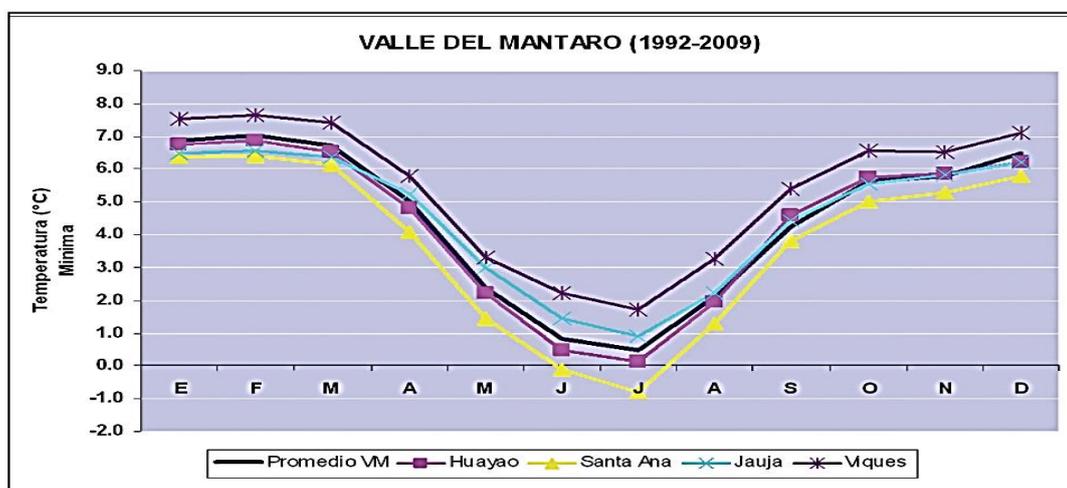
**Figura 2:** Panorama económico de la producción de papa en Perú

FUENTE: INEI, 2017

### 2.1.1. VARIACIÓN MENSUAL DE LA TEMPERATURA MÍNIMA

La variación mensual de la temperatura mínima en el valle del Mantaro está fuertemente asociada a la estacionalidad (Figura 02). Se observan dos periodos bien definidos durante el año: uno de mínimos valores, centrado en la estación de invierno (junio y julio), siendo el valor más bajo de 0,5°C en julio; y otro de valores máximos en verano (entre enero y marzo),

con un valor promedio alrededor de 7,0 °C (IGP, 2010). La mayor variabilidad de las temperaturas mínimas, en las estaciones de estudio, se da en los meses de invierno (junio-agosto), época donde son más frecuentes las heladas en la región (IGP, 2005), y la menor variabilidad se da durante el verano, de enero a marzo.



**Figura 3:** Variabilidad mensual de las temperaturas mínimas en el valle del Mantaro. Periodo común: enero 1992 a diciembre 2009, excepto para Viques de enero 1998 a diciembre 2009

FUENTE: IGP, 2010

### 2.1.2. ANÁLISIS DE HELADAS

Estudios realizados en el Mantaro señalan que las heladas se incrementan con la altura en ciertas zonas en el Valle, y que los eventos de heladas durante el periodo de siembra se está incrementando en un promedio de ocho días por cada diez años (Trasmonte et al. 2008). Sin embargo las heladas en el valle tienen mayor ocurrencia durante los meses entre junio y agosto, asociados a poca cobertura nubosa y aire muy seco (IGP, 2005 y Villegas, 1991).

Villegas y Rosas (1989) mediante el método de la topoclimatología determinaron que las diferencias locales de las temperaturas mínimas en el Valle del Mantaro se originan a partir de las características topográficas y fisiográficas propias del valle. Asimismo encontraron la peligrosidad de ocurrencia de heladas de tipo radiativas para cuatro niveles de intensidad (-4; -3; 0 y -2) y para los tres períodos: siembra, desarrollo vegetativo y cosecha. Las zonas más frías del valle se ubican en el norte, en Jauja, así como en Sicaya y las Pampas de

Chichipampa; en tanto que las zonas más cálidas se ubican al sur, en Viques y en Aza Palián. En relación al tipo de heladas, el 83% son por irradiación o heladas de enfriamiento nocturno, éstas se dan muy frecuentemente en el invierno (junio-julio-agosto), iniciándose su período de ocurrencia a mediados de otoño (abril) y finalizando generalmente en setiembre; sin embargo esporádicamente se pueden presentar en noviembre o en diciembre, que es cuando son más peligrosas y dañinos para los cultivos ya establecidos.

### **2.1.3. CLIMA Y AGRICULTURA**

La papa es un cultivo oriundo y representativo del Perú; es el patrimonio más grande de nuestra cultura agraria, y uno de los mayores problemas que afecta a su producción es el factor climático, el cual es un factor no controlado por el productor y que puede impactar positiva o negativamente en el rendimiento y calidad del producto cosechado. De todos los factores ambientales que inciden sobre el cultivo de papa, tal vez la menos estudiada y comprendida, a pesar de su prioritaria importancia, sean las condiciones atmosféricas en el que está inmerso el cultivo y probablemente la temperatura del aire, la precipitación y el fotoperiodo son las variables ambientales que más impactan en este cultivo. Es necesaria la cultura de usar la información climática, la cual es muy pobre entre los agricultores del valle del Mantaro y ahora más que antes dado que, en los últimos años, hay un cambio significativo en el clima y debemos adaptarnos. (IGP, 2010).

El maíz amiláceo es otro de los principales cultivos y constituye uno de los productos básicos para el consumo humano en nuestro País. Es la sierra central la que tiene mayor área dedicada a este cultivo. El valle del Mantaro es la mayor productora de maíz de grano seco y abastece al principal mercado de Lima, aportando con aproximadamente el 20% de la producción nacional. Sin embargo, estando ubicado en la sierra altoandina del Perú, presenta ambientes bastante variados, con insuficientes e irregulares precipitaciones que provocan periodos de escasez de agua y variaciones de temperatura que van desde los 0,7 °C a 21 °C aunque alcanza muchas veces valores bajo cero, lo que afecta a los cultivos. Aún con estas condiciones negativas, el maíz amiláceo es el principal cultivo; su siembra está limitada al uso de variedades muy tradicionales y complejas, como la variedad Blanco Urubamba y San Gerónimo. Sus rendimientos son bajos, dado a que es cultivado mayormente por pequeños y medianos agricultores, quienes lo producen para autoconsumo o su pequeña

comercialización, ya sea en el estado de choclo o grano. Constituye además una base fundamental de su alimentación y economía. Las variedades más comerciales son aquellas que muestran buena característica de mazorca con hileras bien conformadas y granos medianos o grandes, como el Blanco Urubamba, que es la preferida en el mercado de maíz amiláceo. Sin embargo, tal variedad es de un periodo vegetativo largo, lo que la hace susceptible al frío, por tanto, tiene mayor requerimiento de agua y necesita un manejo agronómico adecuado. Estos factores hacen que aumenten los riesgos y reduzca notoriamente la rentabilidad a causa de las variaciones en el clima que se presentan con bastante frecuencia en los últimos años. (IGP, 2010).

El cultivo de quinua es una actividad de gran importancia en la sierra peruana desde tiempos ancestrales hasta la fecha. Se cultiva en zonas áridas y semiáridas de los Andes. Tiene una gran adaptabilidad, tanto en latitud como en altitud, encontrándose en el Perú desde Tacna hasta Piura, y desde el nivel del mar hasta los 4 000 metros de altura. La quinua constituye uno de los elementos más importantes de la dieta alimenticia de la población humana, principalmente en los Andes peruanos, debido a sus propiedades nutricionales. En el valle del río Mantaro, la quinua es uno de los cultivos más populares, aunque allí están expuestos a las variaciones del clima, las cuales son bastante extremas, como son las sequías o heladas, que algunos autores sugiere el mejoramiento genético como alternativa para la mejor adaptación de la quinua a climas hostiles como los del valle del río Mantaro. Actualmente, en el valle del Mantaro, se vienen observando variaciones en el clima como sequías prolongadas y presencia de lluvias en fechas que antes no se presentaban; estas variaciones, en muchos casos, son la causa del bajo rendimiento de los cultivos de la zona. (IGP, 2010).

## **2.2. TEMPERATURA MÍNIMA**

Es la temperatura más baja registrada en un día, mes o año, este valor puede ser negativo si su registro se da por debajo de cero. Es una de las variables críticas que impulsa todos los sistemas biológicos. (Norabuena, 2014).

### 2.2.1. HELADAS

La helada es un fenómeno atmosférico que se presenta cuando la temperatura del aire, existente en las cercanías del suelo, desciende por debajo de cero grados. De otro lado, existen especies tropicales, como el caucho y el cacao, que sufren del frío a temperaturas marcadamente superiores a 0°C. Generalmente la helada se presenta en la madrugada o cuando está saliendo el Sol.

Técnicamente, la palabra “helada” se refiere a la formación de cristales de hielo sobre las superficies, tanto por congelación del rocío como por un cambio de fase de vapor de agua a hielo; no obstante, la palabra es ampliamente utilizada por el público para describir un evento meteorológico cuando los cultivos y otras plantas experimentan daño por congelación. Los agricultores a menudo utilizan los términos “helada” y “congelación” de forma indistinta, con la definición vaga de “una temperatura del aire inferior o igual a 0 °C”.

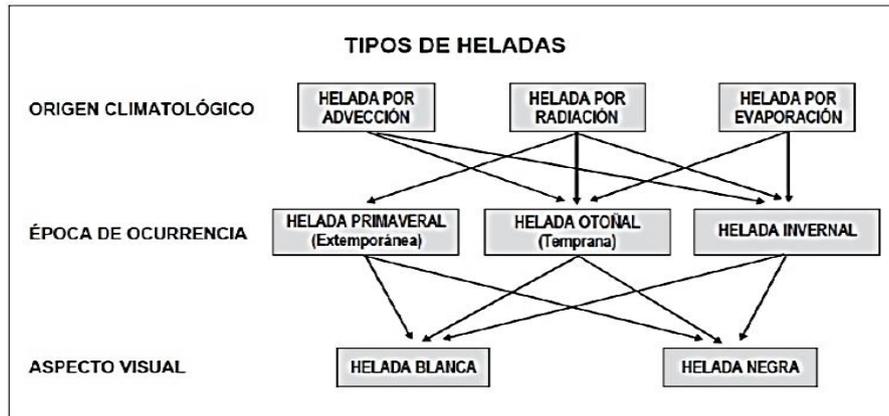
Ejemplos de definiciones de helada en la literatura incluyen: La ocurrencia de una temperatura inferior o igual a 0 °C medida en una garita “tipo Stevenson” a una altura entre 1,25 y 2,0 m. La ocurrencia de una temperatura inferior a 0 °C, sin definición del tipo de garita y de la altura. Cuando la temperatura de la superficie cae por debajo de 0 °C; y la existencia de una temperatura del aire baja que causa el daño o la muerte de las plantas, sin mencionar la formación de hielo.

Así mismo se utilizan los siguientes términos:

**Helada Meteorológica.-** Se define como helada meteorológica a la ocurrencia de una temperatura mínima diaria que no supere los 0 °C en abrigo meteorológico (medida a 1.5 m del suelo). **Helada Agronómica.-** En forma general, se define como el descenso de la temperatura del aire a niveles críticos para los cultivos, sin llegar necesariamente a 0°C.

### CLASIFICACIÓN

Las heladas se pueden agrupar desde los puntos de vista de origen climatológico, época de ocurrencia o aspecto visual. Algunas de las categorías se relacionan entre sí, por ejemplo una helada por radiación puede ocurrir en la estación primaveral, otoñal o invernal, etc. (Ver figura 3).



**Figura 4:** Tipos de heladas

FUENTE: Matías Ramírez, 2001

## **POR SU GÉNESIS:**

### **a) Helada de Advección de Viento frío**

Son aquellas provocadas por la ocurrencia de vientos fuertes, constantes, con temperaturas muy bajas por muchas horas seguidas. El aire frío reseca el follaje causando su muerte. Por lo tanto, los vientos son los causantes de los daños a las plantas. La helada de advección se puede presentar en cualquier hora del día, con independencia del estado del cielo. Tiene su origen en una invasión de aire frío, con una temperatura inferior al punto de congelación. Suelen afectar a amplias zonas y por sus características los métodos de lucha contra este tipo de helada acostumbran a ser eficaces. En muchos casos, una serie de noches bajo cero empiezan como una helada de advección y cambiarán más tarde a noches con heladas de radiación.

### **b) Helada de Radiación**

Ocurre cuando hay enfriamiento intenso de la superficie, que pierde energía durante las noches con cielo despejado, sin viento y con dominio de un anticiclón estacionario de alta presión (masa polar de aire frío), con baja concentración de vapor de agua (seca). Una pérdida radiativa de la superficie hace que el aire adyacente a ella también se enfríe. Luego, el agente causador es la pérdida radiativa intensa. Esa situación ocurre frecuentemente en regiones de clima árido, en que la falta de vapor de agua atmosférico reduce el efecto de estufa local.

Durante el día, la temperatura en la superficie se mantiene encima del punto de congelamiento. Por eso, y por el suelo, durante la noche, la pérdida de energía de la superficie por emisión de radiación de onda larga (ley de Stefan-Boltzmann) se acentúa, provocando una caída rápida de la temperatura del aire próximo a la superficie, resultando en lo que se denomina inversión térmica, o sea, la temperatura aumenta con la altura, en los primeros metros, en lugar de disminuir (situación normal). Las heladas de radiación son acontecimientos usuales. Se caracterizan por un cielo despejado, en calma o con poco viento, inversión de temperatura, temperaturas del punto de rocío bajas y temperaturas del aire que normalmente caen por debajo de 0 °C durante la noche pero que están por encima de 0 °C durante el día. La temperatura del punto de rocío es la temperatura que se alcanza cuando el aire se enfría hasta que alcanza una humedad relativa del 100%, y es una medida directa del contenido de vapor de agua del aire. Durante la noche con cielos despejados, se pierde más calor por radiación que la que se recibe y ello conduce a una caída de la temperatura. La temperatura cae más rápido cerca de la superficie que está radiando provocando la formación de una inversión de temperatura (la temperatura aumenta con la altura por encima del suelo).

### **c) Helada de Evaporación**

Se produce al evaporarse el agua depositada sobre las plantas, con el consiguiente enfriamiento al ser absorbido del aire el calor latente necesario para la evaporación. Si después de una precipitación desciende la humedad relativa del aire, lo que es frecuente después del paso de un frente frío, el agua que recubre los vegetales se evapora rápidamente. La intensidad de estas heladas depende de la cantidad de agua que se evapora, de la temperatura del aire y de la humedad relativa.

### **POR LA ÉPOCA EN LA QUE OCURREN:**

De acuerdo con la estación del año en que se presentan, se tienen tres clases de heladas:

**Heladas primaverales.**- Este tipo de helada afecta principalmente a los cultivos de ciclo anual (como el maíz) cuando se encuentran en la etapa de brotación de ramazo con pocos días de crecimiento. Se presentan cuando en el ambiente se genera un descenso de temperatura.

**Heladas otoñales.**- También llamadas heladas tempranas, son perjudiciales para los cultivos porque pueden interrumpir bruscamente el proceso de formación de botones de las flores y la maduración de frutos. A estas heladas se le atribuye la reducción de la producción agrícola

de una región. Se forman por la llegada de las primeras masas de aire frío de origen polar sobre el país durante los meses de marzo y abril.

**Heladas Invernales.**- Se forman durante el invierno si la temperatura ambiente disminuye notablemente. Estas heladas afectan principalmente a los árboles perennes con frutos y especies forestales, especialmente cuando encuentran un periodo de reposo, lapso en el que las plantas disponen de mayores posibilidades de soportar bajas temperaturas.

#### **POR SU ASPECTO VISUAL:**

Atendiendo a la apariencia de los cultivos expuestos a las bajas temperaturas del aire se tienen dos tipos de heladas: la blanca y la negra. El contenido de humedad en las masas de aire determina estos tipos de heladas.

**Helada Negra.**- Ocurre cuando una atmósfera tiene baja concentración de vapor de agua y una pérdida radiativa intensa, causando enfriamiento acentuado de la vegetación, llegando a la temperatura letal. En función de baja lectura de humedad de aire, no hay deposición de hielo por falta de agua. Este tipo de helada es el más severo, pues una baja humedad del aire permite la ocurrencia de temperaturas bastante menores.

**Helada Blanca.**- Ocurre cuando un intenso enfriamiento nocturno produce condensación de vapor de agua y su congelamiento sobre la plantas. En ese caso, una concentración de vapor de agua en la atmósfera adyacente a la superficie es más elevada que una helada negra. Cuando se tiene más humedad en el aire, primero ocurre una condensación con liberación de calor latente, hecho que ayuda a reducir la caída de la temperatura. Por lo tanto, una helada blanca es menos severa que una negra.

### **2.3. PRONÓSTICO METEOROLÓGICO**

Actualmente existen diferentes métodos para realizar un pronóstico. El método que un pronosticador utilice depende básicamente de su experiencia, la cantidad de información disponible, del nivel de dificultad que presenta la situación y del grado de exactitud o confianza necesaria en el pronóstico.

### **2.3.1. CONFIABILIDAD DE UN PRONÓSTICO NUMÉRICO DEL TIEMPO**

Las ecuaciones matemáticas de un modelo numérico del tiempo pueden ser resueltas en una región limitada o en el globo entero. Cuando es en una región limitada se llama MODELO REGIONAL, cuando es en el globo entero se le llama MODELO GLOBAL. La capacidad del hombre para predecir el tiempo con anticipación de 24, 48, 72 horas o más, es muy limitada. Se podría decir que es casi imposible de predecir el tiempo con varios días de anticipación. Sin embargo a través de un MODELO GLOBAL se puede tener pronósticos confiables hasta con una semana de anticipación. La confiabilidad es mayor para periodos de tiempo menor que 5 días, y en regiones de latitudes medias, tales como Estados Unidos, Argentina, Sur de Brasil, Europa, Rusia, etc. La habilidad (skill) de un Modelo Global para hacer buenas predicciones del tiempo depende de varios factores. Por ejemplo: las condiciones iniciales; condiciones de superficie; tipos de parametrizaciones de los procesos de superficie, radiación y la convección; la resolución del modelo; el tipo de método numérico usado; etc.

En general un modelo Global complejo que incluye todos los procesos físicos solamente puede ser corrido en una supercomputadora como la CRAY, IBM, SX-3 o SX-4. De esta forma, en las latitudes medias podemos tener pronósticos confiables hasta con una semana de anticipación. En las regiones tropicales como Perú, Colombia, parte Norte y Central de Brasil, Cuba, India, etc, la confiabilidad de estos modelos disminuye.

Un modelo numérico regional es muy útil para el pronóstico del tiempo con alta resolución (desde 500-600 m hasta 60-100 Km) y con antecendencia de 48 horas. Estos modelos son más confiables que los modelos globales, y no necesariamente requieren de un supercomputador. La confiabilidad de un modelo numérico depende también de la cantidad y calidad de datos que son proporcionados al modelo como condición inicial. Esto implica que para el Perú existe la necesidad de tener la mayor cantidad de estaciones meteorológicas de superficie y de altura.

### **2.3.2. PREDICCIÓN NUMÉRICA DEL TIEMPO (NWP)**

La predicción numérica del tiempo, usa complejos programas de cómputo, conocidos como modelos numéricos de pronóstico, que procesan (“corren”) datos en supercomputadoras y

proporcionan predicciones de variables. Un modelo numérico es un conjunto de ecuaciones matemáticas cuya solución requiere de métodos numéricos. Las ecuaciones básicas de un modelo numérico del tiempo son aquellas que rigen el movimiento del aire (horizontal y vertical), conservación de la masa y la energía, las transformaciones termodinámicas, los procesos de formación y desarrollo de las nubes, etc. Los métodos numéricos más comunes usados para resolver el sistema de ecuaciones diferenciales en derivadas parciales (modelo numérico del tiempo) son: diferencias finitas, métodos espectrales y elementos finitos.

### **MODELAMIENTO NUMÉRICO DE LA DINÁMICA ATMOSFÉRICA**

La predicción meteorológica consiste en determinar anticipadamente los valores de una serie de variables meteorológicas, como temperatura, humedad, presión, dirección e intensidad del viento, precipitación, etc., que pueden llegar a afectar a determinada región. Actualmente las predicciones se realizan utilizando modelos de predicción numérica que emplean las ecuaciones de las leyes físicas que describen el comportamiento de la atmósfera. La solución de estas ecuaciones es algo compleja, debido a que no existe una solución matemática exacta que permita determinar los valores futuros de las variables, por tal razón es necesario recurrir a las técnicas de modelamiento numérico. (Jiménez, 2014).

Formalmente, los modelos de predicción meteorológica se basan en resolver de forma numérica estas ecuaciones, para ello recurren a realizar aproximaciones a un conjunto de ecuaciones diferenciales referidas a la conservación de la masa, la energía y el momento lineal en la atmósfera. Partiendo de condiciones iniciales obtenidas mediante un proceso de asimilación de datos, a partir de un modelo de circulación general y un conjunto de observaciones en un determinado momento de tiempo ( $t$ ), el modelo resuelve las ecuaciones para encontrar valores a las variables pronosticadas en un tiempo futuro ( $t + \Delta t$ ). Para ello es necesario hacer uso de las ecuaciones en derivadas parciales, que traducen las leyes generales de la física que rigen la dinámica de la atmósfera terrestre. (Jiménez, 2014).

La atmósfera es un fluido en movimiento sometido a procesos de intercambio de energía. En consecuencia, se necesita describir adecuadamente el movimiento de las partículas de aire sometidas a un conjunto de fuerzas e introducir el balance energético una vez que se tiene en cuenta los intercambios que tienen lugar en el medio atmosférico, tanto en el medio como entre la atmósfera y la superficie terrestre. Para describir su comportamiento se utilizan las ecuaciones generales de la mecánica de fluidos y de la termodinámica. Además, el

movimiento de las partículas de fluidos atmosféricos debe ser estudiado teniendo en cuenta que nuestro planeta es un sistema de referencia no inercial. Por lo tanto, para que se verifiquen las leyes de la mecánica clásica, es necesario introducir algunos términos denominados fuerzas aparentes. (Jiménez, 2014).

El movimiento de una partícula de fluido atmosférico viene dada por:

$$\frac{d\vec{V}}{dt} = \sum_i \frac{\vec{F}_i}{m} = -\frac{1}{\rho} \vec{\nabla} p + \vec{g} - 2\vec{\Omega} * \vec{V} + \vec{F}_r \quad (..)$$

donde,

$\left(\frac{d\vec{V}}{dt}\right)$  → Aceleración por unidad de masa

$\left(\sum_i \frac{\vec{F}_i}{m}\right)$  → Fuerza por unidad de masa

$\left(-\frac{1}{\rho} \vec{\nabla} p\right)$  → Gradiente de presión

$\vec{g}$  → Gravedad

$(-2\vec{\Omega} * \vec{V})$  → Fuerza de Coriolis

$\vec{F}_r$  → Fuerza de rozamiento

El término de la izquierda representa la fuerza por unidad de masa que actúa sobre una burbuja de aire, medida desde un sistema de referencia no inercial (la Tierra). Los términos de la derecha señalan dicha burbuja: la fuerza debido al gradiente de presión, resultante de la fuerza gravitatoria y céntrica, la fuerza de Coriolis, que junto a la anterior, son los dos términos ficticios de la ecuación, y el término asociado a la fuerza de rozamiento. Para resolver esta ecuación es necesario introducir una serie de simplificaciones en función de las escalas espacio-temporal utilizado en cada caso. (Jiménez, 2014).

Desde el punto de vista termodinámico hay que tener en cuenta los intercambios de calor asociados a los cambios de fases del agua (vapor, líquido y estado sólido) que afectan a la formación y desarrollo de las nubes y a los intercambio de energía con la superficie de la tierra, los océanos y la troposfera. Otro factor a tener en cuenta es el hecho de que la atmósfera es un fluido gaseoso que sufre variaciones de densidad como consecuencia de la dinámica a la que se encuentra sometida. (Jiménez, 2014).

Las ecuaciones matemáticas que describen todos estos procesos no son lineales y, por tanto, las soluciones no se pueden obtener de forma analítica. Para resolverlas es preciso recurrir al cálculo numérico, que proporciona una solución aproximada. Para obtener estas soluciones aproximadas, el primer paso consiste en discretizar el área geográfica objeto de estudio. Para ello se define una rejilla (grid) tridimensional que la abarque. La distancia entre dos puntos de la rejilla se conoce como resolución del modelo. El siguiente paso consiste en resolver las ecuaciones, utilizando cálculo numérico, en todos los puntos de la grilla. (Jiménez, 2014).

## **CLASIFICACIÓN DE LOS MODELOS DE PREDICCIÓN DEL TIEMPO**

Los modelos numéricos de predicción del tiempo se pueden clasificar en:

- Modelos Globales: describen la atmósfera a gran escala, es decir, sobre el planeta como un todo. Utilizan sólo las condiciones de contorno verticales. Identifican fenómenos de escala sinóptica; Y
- Modelos de Mesoescala: Posee espaciamiento de rejilla reducida en relación a los modelos globales, lo que hace que represente fenómenos típicos de mesoescala, que van desde cientos de metros a decenas de kilómetros de extensión. Estos modelos generalmente se aplican a estudios o previsiones de tiempo a escala regional o local. Los modelos de mesoescala son capaces de permitir la representación de fenómenos que difícilmente se simulan en modelos globales, como la circulación local, la influencia de la topografía, entre otros. Los modelos de mesoescala utilizan resolución en torno a algunas decenas de kilómetros, es decir, su resolución es mayor en comparación con los globales. (Pereira, 2011).

### **MODELOS MESOESCALARES:**

Los modelos mesoescalares son aquellos que permiten estudiar y pronosticar fenómenos meteorológicos, incluyendo los mesoescalares, es decir, aquellos que ocurren a una escala espacial relativamente pequeña. Entre los primeros modelos se encuentran los hidrostáticos pues consideraban despreciables los movimientos en la escala vertical frente a los movimientos en la horizontal. Esto ocurre con modelos que simulan a escala sinóptica, válidos para establecer el comportamiento de las masas de aire, frentes polares, etc. En estos casos se mantiene la condición hidrostática y la presión está completamente determinada por

las masas de aire extendidas, quedando inhibidos los movimientos del aire en la vertical en las hipótesis iniciales. Sin embargo, cuando la escala de la rejilla del modelo se reduce, la resolución espacial aumenta y es posible incorporar movimientos en la vertical, y por tanto, la hipótesis hidrostática, deja verificarse. Se puede decir que esto permite incorporar algunos factores meteorológicos como la convección, ondas, etc. que son de tipo mesoescalar. (Jiménez, 2014).

Entre tanto, hay dos grandes tipos de modelos numéricos: hidrostáticos, como el High Resolution Limited Area Model (HIRLAM), y no hidrostáticos, como el Mesoscale Atmospheric Simulation System (MASS), Regional Atmospheric Modeling System (RAMS), PennState-NCAR (MM5), o la versión avanzada de este último denominada Weather Research and Forecasting (WRF). Actualmente, los modelos mesoescalares de área limitada permiten realizar simulaciones de la atmósfera bajo condiciones no hidrostáticas, y resultan muy útiles para el estudio de fenómenos de precipitaciones intensas y severas. (Jiménez, 2014).

Todos los modelos de simulación mesoescalares hacen uso de distintos esquemas de parametrizaciones para resolver aquellos fenómenos cuya escala espacial es inferior a la resolución del modelo, entre ellos, parametrización de cúmulos, microfísica, capa límite, etc. (Jiménez, 2014).

## **CONSIDERACIONES GENERALES DE LOS MODELOS DE PREDICCIÓN**

Los modelos numéricos se basan en ecuaciones matemáticas, que representan las leyes de la física que rigen los movimientos de la atmósfera y las interacciones con la superficie. A continuación se enumeran los principios físicos; Como se tratan las coordenadas verticales y horizontales y las condiciones de contorno. (Pereira, 2011).

**Ecuaciones básicas.-** Los modelos de mesoescala utilizan los siguientes principios físicos: conservación de masa; conservación del calor; conservación del movimiento; conservación del agua; conservación de los gases. En Pielke (2002) se desarrollan las ecuaciones de conservación citadas anteriormente y también se presentan sus simplificaciones.

**Coordenadas horizontales.-** Cuando se trabaja con modelos atmosféricos es muy común no utilizar coordenadas cartesianas para representar el terreno. Una alternativa a esta representación es el sistema de coordenada horizontal polar estereográfica. Esta proyección

tiene en cuenta las diferencias entre el plano de la rejilla y la superficie de la esfera, teniendo en cuenta que el dominio de un modelo mesoesalar puede ser extenso. La deformación asociada a la representación de la esfera sobre la rejilla del modelo puede ser dada por el factor de escala de proyección ( $m$ ) se define como:  $m = Dp / Dr$ , donde  $Dp$  es la distancia proyectada en la cuadrícula del modelo y  $Dr$  es la distancia real sobre la superficie terrestre.

**Coordenadas verticales.-** La estructura vertical del modelo es tan importante en la definición del comportamiento del modelo como la configuración horizontal. La adecuada representación de la estructura vertical de la atmósfera requiere la selección apropiada de la coordenada vertical. A diferencia de las coordenadas horizontales, donde se puede elegir entre una representación discreta (puntos de rejilla) o continua (espectral), la representación vertical es discreta (COMET, 2017).

- a) Altura ( $z$ ): tiene la ventaja de ser intuitiva y fácil de construir ecuaciones. La desventaja es la dificultad de representar la superficie de la tierra porque diferentes lugares poseen diferentes alturas.
- b) Presión ( $p$ ): sus ventajas son la facilidad de representar la parte superior de la atmósfera ( $p = 0$ ) y de incorporar datos de radiosonda. La desventaja es la dificultad de representar la superficie de la tierra pues la presión varía de un punto a otro.
- c) Sigma ( $\sigma = p / p_{superficie}$ ): es fácil de representar la parte superior y la base de la atmósfera, en la parte superior  $\sigma = 0$  y en la superficie de la tierra  $\sigma = 1$ . Lo que tiene como desventaja la posibilidad de errores de cálculo del gradiente de presión horizontal en áreas con pendientes acentuadas.
- d) Eta ( $\eta$ ): fue creado en la década de 1980, en un intento de reducir los errores que a veces ocurren cuando la fuerza de gradiente de presión horizontal fue calculada usando sigma. Su ventaja es mejorar el cálculo del gradiente de presión horizontal. Y tiene la desventaja de no representar la realidad de la topografía con exactitud.
- e) Isentrópica ( $\theta$ ): utiliza la temperatura potencial  $\theta$  como coordenada vertical. Cuando la atmósfera está estable e insaturada  $\theta$  aumenta con la altura.

**Condiciones de contorno.-** Los modelos pueden contener dos tipos de condiciones de contorno. **Vertical:** especifica la relación utilizada para definir la magnitud de las variables en la parte superior y en la base del modelo. **Lateral:** se refiere a la relación que se utiliza para especificar la magnitud de las variables en los vértices horizontales del dominio del modelo.

**Operación.-** La inicialización del modelo es el proceso de preparar los datos para la entrada del modelo. Los datos proceden de diferentes fuentes y pueden contener errores. El proceso de arranque debe juntar los datos en un conjunto coherente que representa con precisión el estado de la atmósfera. Las ecuaciones se resuelven utilizando la aproximación numérica para truncar los resultados, esto produce los errores de truncamiento. Los modelos numéricos generaron archivos que contienen las previsiones del modelo. El post-procesamiento de las previsiones se realiza para hacer ajustes en las salidas y para generar productos gráficos de las previsiones. Las previsiones del modelo se utilizan a veces como entrada para métodos estadísticos, lo que puede generar previsiones para ubicaciones específicas.

**Post-procesamiento.-** El post-procesamiento transforma las salidas de los modelos en la información que se utiliza en la previsión del tiempo. Algunas variables previstas por los modelos ya poseen alto grado de precisión, mientras que otras, como la precipitación no presentan resultados plenamente satisfactorios. Comúnmente los modelos producen resultados que no son utilizados por los meteorólogos, como por ejemplo: resultados en las capas sigma y no en los niveles verticales estándares. Durante el post procesamiento, los campos nativos generados por el modelo se transforman en un sistema de coordenadas que tienen significado para el predicador. En esta transformación se generan los grid points o incluso la información para puntos específicos y varios niveles verticales. Las salidas de las variables patrones (presión, temperatura, humedad específica, componentes del viento, etc.) así como las derivadas (presión media a nivel del mar, humedad relativa, punto de rocío, altura geopotencial, CAPE, CINE, entre otros). A partir de estos resultados otros productos pueden ser generados en el post-procesamiento. Hay varias técnicas para generar los campos derivados. La mayoría se calcula a partir de fórmulas y no son específicas de un determinado modelo (COMET, 2017).

#### **2.4. BRAZILIAN REGIONAL ATMOSPHERIC MODELING SYSTEM (BRAMS)**

Brazilian Regional Atmospheric Modeling System - BRAMS, es una versión adaptada del Regional Atmospheric Modeling System - RAMS desarrollado en la Universidad de Colorado (CSU) (PIELKE et al., 1992; WALKO, TREMBACK, 2005) es el resultado de un proyecto financiado por la Financiadora de Estudios y Proyectos (FINEP) involucrando al Centro de Previsión de Tiempo y Estudios Climáticos del Instituto Nacional de Investigaciones Espaciales (CPTEC / INPE), Instituto de Matemática y Estadística e

Instituto de Astronomía, Geofísica y Ciencias De la Universidad de São Paulo (IME / USP), (IAG / USP) y Atmospheric, Environmental and Environmental Technologies (ATMET). El RAMS surgió de esfuerzos realizados separadamente, desde los años 70, comandados por William R. Cotton en la parte de modelado de sistemas dinámicos de microescala y procesos microfísicos, y Roger A. Pielke en la parte de modelado de sistemas de mesoescala y en la influencia de la superficie De la tierra en las características de la atmósfera; Sin embargo, sólo en 1986 esos esfuerzos se sumaron, con el fin de generar un modelo más completo. Este modelo atmosférico fue construido con base en las ecuaciones de la dinámica de la atmósfera, complementadas con parametrizaciones de difusión turbulenta, radiación solar y terrestre, procesos húmedos incluyendo la formación y la interacción de nubes y agua líquida precipitante y hielo, calor sensible y latente, suelo, vegetación y superficie de agua, los efectos cinemáticos del terreno y convección cumulus. Todas estas parametrizaciones y características del modelo pueden ser alteradas para que mejor representen las condiciones específicas de determinado local, o las condiciones ideadas para simulaciones de situaciones, lo que lo convierte en una excelente herramienta para investigaciones meteorológicas. El BRAMS tiene mejoras en el código, dispone de otras opciones de parametrización de la convección profunda y, además, una parametrización para convección rasa (SOUZA, 1999). (Pereira, 2011).

#### **2.4.1. CARACTERÍSTICAS DEL MODELO BRAMS**

El sistema de ecuaciones que utiliza el modelo es el sistema de ecuaciones de Navier-Stokes considerando un estado de referencia hidrostático (representado con el subíndice cero en las variables). Las ecuaciones generales del modelo se describen a continuación y el significado de los símbolos se presenta en la Tabla 1. (García, 2007).

Ecuaciones de movimiento:

$$\frac{\partial u}{\partial t} = -u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} - w \frac{\partial u}{\partial z} - \theta \frac{\partial \pi'}{\partial x} + fv + TURB(u)$$

$$\frac{\partial v}{\partial t} = -u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} - w \frac{\partial v}{\partial z} - \theta \frac{\partial \pi'}{\partial y} - fu + TURB(v)$$

$$\frac{\partial w}{\partial t} = -u \frac{\partial w}{\partial x} - v \frac{\partial w}{\partial y} - w \frac{\partial w}{\partial z} - \theta \frac{\partial \pi'}{\partial z} - \frac{g\theta'_v}{\theta_0} + TURB(w)$$

Ecuación termodinámica:

$$\frac{\partial \theta_{it}}{\partial t} = -u \frac{\partial \theta_{it}}{\partial x} - v \frac{\partial \theta_{it}}{\partial y} - w \frac{\partial \theta_{it}}{\partial z} + \frac{\partial}{\partial x} \left( K_h \frac{\partial \theta_{it}}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_h \frac{\partial \theta_{it}}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_h \frac{\partial \theta_{it}}{\partial z} \right) + \left( \frac{\partial \theta_{it}}{\partial t} \right)_{rad}$$

Ecuación de continuidad:

El modelo utiliza una ecuación de continuidad compresible que escrita como ecuación de pronóstico de la perturbación de  $\pi$  en la mesoescala retiene la naturaleza predictiva de la ecuación a través de retener el flujo de calor de escala sinóptica, siguiendo la aproximación realizada por Klemp y Wilhelmson, 1978. (García, 2007).

$$\frac{\partial \pi'}{\partial t} = - \frac{R\pi_0}{c_v \rho_0 \theta_0} \left( \frac{\partial \rho_0 \theta_0 u}{\partial x} + \frac{\partial \rho_0 \theta_0 v}{\partial y} + \frac{\partial \rho_0 \theta_0 w}{\partial z} \right)$$

donde:

$$\pi = C_p \left( \frac{p}{1000} \right)^{\frac{R}{C_p}}$$

$$\pi = \pi_0(z) + \pi'(x, y, z, t)$$

Símbolo	Significado
u	Componente zonal del viento
v	Componente meridional del viento
w	Componente vertical del viento
f	Parámetro de Coriolis
$\Pi$	Función de Exner
$\pi'$	Perturbación de la función de Exner
$\Theta$	Temperatura potencial
$\rho$	Densidad
$\theta_v$	Temperatura potencial virtual
g	Gravedad
TURB	Término turbulento
R	Constante de los gases
Cp	Calor específico del aire a presión constante
Cv	Calor específico del aire a volumen constante
$\theta_{il}$	Temperatura potencial del hielo y del agua líquida
$K_h$	Coefficiente de viscosidad para el calor y la humedad
Subíndice 0	Indica que corresponde al estado de referencia

**Figura 5:** Símbolos de las ecuaciones que usa el modelo BRAMS

FUENTE: García, 2007

Los términos advectivos se configuran en el modelo en forma de flujo con el fin de conservar la masa y la cantidad de movimiento. Considerando, por ejemplo, la dirección x el término advectivo para una variable  $\Phi$  se expresa de la siguiente forma:

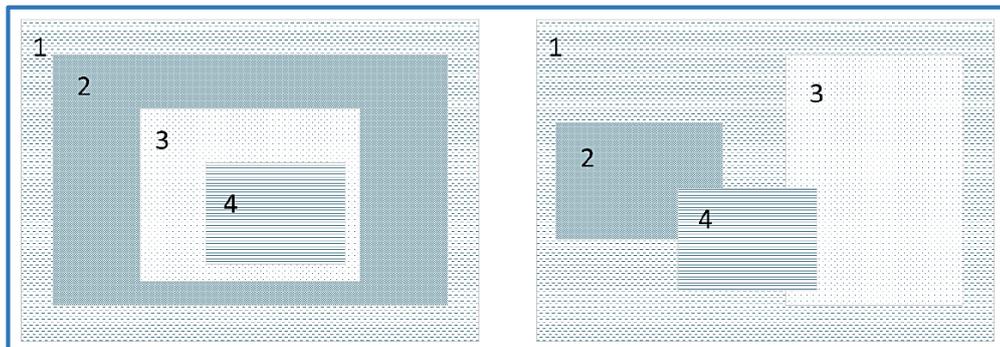
$$-u \frac{\partial \phi}{\partial x} = -\frac{1}{\rho_0} \left[ \frac{\partial(\rho_0 u \phi)}{\partial x} - \phi \frac{\partial(\rho_0 u)}{\partial x} \right]$$

La asimilación de datos se realiza alimentando a los campos generados por el modelo con los análisis producidos por el esquema RAMS/ISAN a medida que el modelo evoluciona en el tiempo. Utiliza un esquema simple de nudging donde se agrega a las ecuaciones de pronóstico del modelo un término que puede ser escrito de la siguiente forma:

$$\frac{\partial \phi}{\partial t} = \frac{(\phi_{observada} - \phi_{modelo})}{\tau}$$

donde  $\Phi$  es una variable del modelo y  $\tau$  es el tiempo que controla la intensidad del *nudging*. En el modelo la escala de tiempo que regula el *nudging* se divide en tres, bordes laterales, borde superior e interior del dominio. El  $\tau$  para cada punto de retícula va a ser el mínimo de las tres escalas computadas, es decir el que va a producir la máxima intensidad del *nudging*. (García, 2007).

De acuerdo con Treichrieb (2008) la estructura de la rejilla es del tipo C de Arakawa. La proyección horizontal de las coordenadas es estereográfica con el polo de proyección en el área central del dominio. En vertical, se utiliza el sistema sigma Z ( $\sigma_z$ ), que rodea la topografía del terreno. Para las simulaciones de previsión de tiempo, el BRAMS necesita datos de pronósticos para las fronteras laterales (condición de contorno) provenientes de un modelo global y de datos de análisis para las condiciones iniciales. Los datos que se van a utilizar proceden del RAMS / ISAN - ISentropic Analysis package (TREMBACK, 1990). Este paquete convierte el formato del análisis proveniente de un modelo global para el formato del BRAMS, o combina y procesa datos observacionales de superficie y de aire superior. En cada paso de tiempo, el BRAMS calcula la previsión de tiempo forzando los datos de su frontera con los datos de la media ponderada del modelo global, a través de una técnica de asimilación de datos (nudging). Esta técnica se utiliza cuando es deseable que los resultados de las ecuaciones del modelo sean forzados a tender a otro resultado ya conocido, por ejemplo, en los análisis de los datos del modelo global, comúnmente utilizados como condiciones de contorno (WALKO, TREMBACK, 2001). Una de las más importantes capacidades del modelo es el de anidamiento de rejillas, pues permite la creación de una cuadrícula detallada (mayor resolución) a partir de una cuadrícula de menor resolución (es posible anidar más de una cuadrícula dentro de una cuadrícula de menor resolución). Los ejemplos de anidamiento se pueden ver en la Figura 5. (Pereira, 2011).



**Figura 6:** Ejemplo del anidamiento de rejillas en el modelo BRAMS

FUENTE: Pereira, 2011

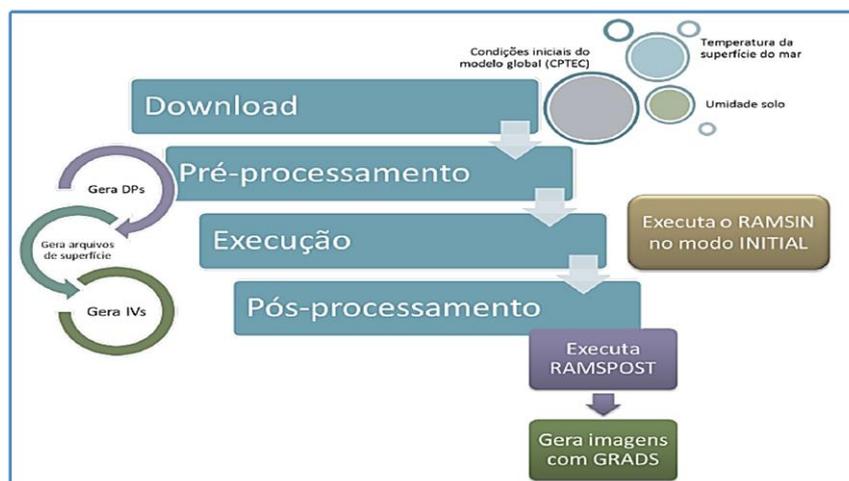
El BRAMS es un modelo meteorológico proyectado para ejecución en clústeres, conjunto de ordenadores interconectados por una red de comunicación, por lo que puede ejecutarse paralelamente (ALMEIDA, 2007). Pero también es posible ejecutarlo en un equipo con un

solo procesador. La elección del hardware se dará de acuerdo con la aplicación del modelo. (Pereira, 2011).

#### **2.4.2. CONSIDERACIONES PARA LA EJECUCIÓN DEL MODELO BRAMS**

**Condiciones iniciales y de contorno:** son datos provenientes de un modelo atmosférico con rejilla mayor que la primera en ejecución en el BRAMS y que posee las variables viento zonal y meridional, geopotencial, temperatura y humedad relativa en varios niveles de presión, por lo menos 12. Estos datos también deben tener una buena frecuencia temporal, preferentemente de 6 en 6 horas. Ejemplo: CPTEC (T126 o T213), GFS, reanálisis del NCEP, ETA20, etc; **datos observados:** pueden ser de superficie o sondeo; **temperatura de la superficie del mar (SST):** se puede utilizar la serie climatológica o semanal; **topografía:** se utilizan los datos del Servicio Geológico de los Estados Unidos (USGS) con resoluciones de 10 o 1 km; **suelo:** los archivos son actualizados por el INPE a partir de la base de datos de la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO); **humedad del suelo:** estos datos se pueden actualizar diariamente; **índice de vegetación por diferencia normalizada (NDVI):** para américa meridional es mejor utilizar los datos provenientes del Moderate Resolution Imaging Spectroradiometer (MODIS) con resolución de 1 Km. (Pereira, 2011).

En la Figura 6 se presenta un esquema simplificado de la ejecución de BRAMS. En (Download) sólo se representa la información que necesita actualizarse diariamente o semanalmente. La etapa de preprocesamiento es cuando los datos de entrada se transforman en los formatos de archivo que utiliza la plantilla. En la ejecución, se crean los archivos de análisis. En el post-procesamiento, se generan las figuras, otras variables no calculadas por el modelo y también se pueden hacer las correcciones estadísticas.



**Figura 7:** Flujo de ejecución de BRAMS

FUENTE: Pereira, 2011

## 2.5. ANÁLISIS EXPLORATORIO DE DATOS (A.E.D.)

El Análisis Exploratorio de Datos (A.E.D.) es un conjunto de técnicas estadísticas cuya finalidad es conseguir un entendimiento básico de los datos y de las relaciones existentes entre las variables analizadas. Para conseguir este objetivo el A.E.D. proporciona métodos sistemáticos sencillos para organizar y preparar los datos, detectar fallos en el diseño y recogida de los mismos, tratamiento y evaluación de datos ausentes (missing), identificación de casos atípicos (outliers) y comprobación de los supuestos subyacentes en la mayor parte de las técnicas multivariantes (normalidad, linealidad, homocedasticidad). El examen previo de los datos es un paso necesario, que lleva tiempo, y que habitualmente se descuida por parte de los analistas de datos. Las tareas implícitas en dicho examen pueden parecer insignificantes y sin consecuencias a primera vista, pero son una parte esencial de cualquier análisis estadístico. (Salvador y Gargallo, 2003).

Antes de proceder a cualquier análisis, es necesario garantizar que la información prevista por instrumentos hidrometeorológicos posee un adecuado control de calidad tal como recomienda la Organización Meteorológica Mundial (OMM). El objetivo es garantizar que las series de tiempo de las distintas estaciones están en realidad mostrando señales asociadas a los sistemas meteorológicos que se presentaron, y no señales que tengan involucrados errores que ocurren durante la adquisición, manipulación, formato, transmisión y archivo de los datos. (Jiménez, 2014).

### **2.5.1. DATOS ATÍPICOS (OUTLIERS)**

Los casos atípicos son observaciones con características diferentes de las demás. Este tipo de casos no pueden ser caracterizados categóricamente como benéficos o problemáticos sino que deben ser contemplados en el contexto del análisis y debe evaluarse el tipo de información que pueden proporcionar. Su principal problema radica en que son elementos que pueden no ser representativos de la población pudiendo distorsionar seriamente el comportamiento de los contrastes estadísticos. Por otra parte, aunque diferentes a la mayor parte de la muestra, pueden ser indicativos de las características de un segmento válido de la población y, por consiguiente, una señal de la falta de representatividad de la muestra. (Salvador y Gargallo, 2003).

#### **TIPOS DE OUTLIERS**

Los casos atípicos pueden clasificarse en 4 categorías. (Salvador y Gargallo, 2003).

La primera categoría contiene aquellos casos atípicos que surgen de un error de procedimiento, tales como la entrada de datos o un error de codificación. Estos casos atípicos deberían subsanarse en el filtrado de los datos, y si no se puede, deberían eliminarse del análisis o recodificarse como datos ausentes.

La segunda clase es la observación que ocurre como consecuencia de un acontecimiento extraordinario. En este caso, el outlier no representa ningún segmento válido de la población y puede ser eliminado del análisis.

La tercera clase contiene las observaciones cuyos valores caen dentro del rango de las variables observadas pero que son únicas en la combinación de los valores de dichas variables. Estas observaciones deberían ser retenidas en el análisis pero estudiando qué influencia ejercen en los procesos de estimación de los modelos considerados.

La cuarta y última clase comprende las observaciones extraordinarias para las que el investigador no tiene explicación. En estos casos lo mejor que se puede hacer es replicar el análisis con y sin dichas observaciones con el fin de analizar su influencia sobre los resultados. Si dichas observaciones son influyentes el analista debería reportarlo en sus conclusiones y debería averiguar el porqué de dichas observaciones.

## IDENTIFICACIÓN DE OUTLIERS

Los casos atípicos pueden identificarse desde una perspectiva univariante o multivariante. La *perspectiva univariante* examina la distribución de observaciones para cada variable, seleccionando como casos atípicos aquellos casos cuyos valores caigan fuera de los rangos de la distribución. La cuestión principal consiste en el establecimiento de un umbral para la designación de caso atípico. Esto se puede hacer gráficamente mediante histogramas o diagramas de caja o bien numéricamente, mediante el cálculo de puntuaciones tipificadas. (Salvador y Gargallo, 2003).

### 2.5.2. DISTRIBUCIÓN NORMAL DE DATOS

Muchos métodos estadísticos se basan en la hipótesis de normalidad de la variable objeto de estudio. De hecho, si la falta de normalidad de la variable es suficientemente fuerte, muchos de los contrastes utilizados en los análisis estadístico-inferenciales no son válidos. Incluso aunque las muestras grandes tiendan a disminuir los efectos perniciosos de la no normalidad, el investigador debería evaluar la normalidad de todas las variables incluidas en el análisis. Existen varios métodos para evaluar la normalidad de un conjunto de datos que pueden dividirse en dos grupos: los *métodos gráficos* y los *contrastos de hipótesis*. (Salvador y Gargallo, 2003).

### MÉTODOS GRÁFICOS PARA EVALUAR LA NORMALIDAD

El método gráfico univariante más simple para diagnosticar la normalidad es una comprobación visual del *histograma* que compare los valores de los datos observados con una distribución normal. Aunque atractivo por su simplicidad, este método es problemático para muestras pequeñas, donde la construcción del histograma puede distorsionar la representación visual de tal forma que el análisis sea poco fiable. Otras posibilidades, también basadas en información gráfica, consisten en realizar *diagramas de cuantiles* (Q-Q plots). Los diagramas de cuantiles comparan en un sistema de coordenadas cartesianas, los cuantiles muestrales (eje X) con los cuantiles esperados bajo la hipótesis normalidad. Si la distribución de partida es normal dichos diagramas tenderán a ser rectas que pasan por el

origen. Cuanto más se desvíen de una recta, menos normales serán los datos. (Salvador y Gargallo, 2003).

## CONTRASTES DE HIPÓTESIS PARA EVALUAR LA NORMALIDAD

La segunda de las formas para comprobar la normalidad de una distribución se efectúa a través de un *contraste de hipótesis*. No existe un contraste óptimo para probar la hipótesis de normalidad. La razón es que la potencia relativa depende del tamaño muestral y de la verdadera distribución que genera los datos. Desde un punto de vista poco riguroso, el contraste de Shapiro y Wilks es, en términos generales, el más conveniente en muestras pequeñas ( $n < 30$ ), mientras que el contraste de Kolmogorov-Smirnov, en la versión modificada de Lilliefors es adecuado para muestras grandes. (Salvador y Gargallo, 2003).

En el *test de Kolmogorov-Smirnov* la hipótesis nula que se pone a prueba es que los datos proceden de una población con distribución normal frente a una alternativa de que no es así. Este contraste calcula la distancia máxima entre la función de distribución empírica de la muestra y la teórica. Si la distancia calculada es mayor que la encontrada en las tablas, fijado un nivel de significación, se rechaza el modelo normal.

El *contraste de Shapiro y Wilks* se utiliza para muestras pequeñas ( $n < 30$ ) y utiliza el hecho de que si  $x(1) \leq x(2) \leq \dots \leq x(n)$  es una muestra ordenada de la  $N(\mu, \sigma)$  entonces:

$$E\left[\frac{x(i) - \mu}{\sigma}\right] = C_{i,n} \quad \text{donde } C_{i,n} = \phi^{-1}\left(\frac{i-3/8}{n+1/4}\right)$$

por lo que  $E[x(i)] = \mu + \sigma C_{i,n}$  y el gráfico de  $x(i)$  frente a  $C_{i,n}$  será una recta.

Dado que  $\phi(C_{i,n}) + \phi(C_{n+1-i,n}) = 1$   $i=1, \dots, n/2$  se tiene que  $C_{i,n} = -C_{n+1-i,n}$  por lo que  $C_{1,n} + \dots + C_{n,n} = 0$ . El test de Shapiro-Wilks se basa en calcular el coeficiente de correlación entre  $x(i)$  y  $C_{i,n}$  y cuanto más cerca de 1 esté, mayor será el grado de normalidad de la distribución y viene dado por la expresión:

$$r^2 = \frac{\left(\sum_{i=1}^n x(i) C_{i,n}\right)^2}{ns^2 \left(\sum_{i=1}^n C_{i,n}^2\right)}$$

Shapiro y Wilks evalúan la distribución del estadístico  $r^2$  bajo hipótesis de normalidad y proporcionan un test que rechaza dicha normalidad cuando el ajuste es bajo, es decir, cuando el estadístico toma valores pequeños.

Otros contrastes muy utilizados son los **tests de asimetría y curtosis** cuyos estadísticos muestrales vienen dados por:

$$z_{\text{asimetría}} = \frac{\text{asimetría}}{\sqrt{\frac{6}{N}}} \text{ y } z_{\text{curtosis}} = \frac{\text{curtosis}}{\sqrt{\frac{24}{N}}}$$

donde  $N$  es el tamaño muestral. Si es cierta la hipótesis de normalidad ambos se distribuyen asintóticamente según una  $N(0,1)$ .

## 2.6. MODELO DE REGRESIÓN LINEAL

La estructura del modelo de regresión lineal es la siguiente:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

En esta expresión se admite que todos los factores o causas que influyen en la variable respuesta  $Y$  pueden dividirse en dos grupos: el primero contiene a una variable explicativa  $X$  y el segundo incluye un conjunto amplio de factores no controlados que englobaremos bajo el nombre de perturbación o error aleatorio ( $\varepsilon$ ) que provoca que la dependencia entre las variables dependiente e independiente no sea perfecta, sino que esté sujeta a incertidumbre. Por ejemplo, en el consumo de gasolina de un vehículo ( $Y$ ) influyen la velocidad ( $X$ ) y una serie de factores como el efecto conductor, el tipo de carretera, las condiciones ambientales, etc, que quedarían englobados en el error. (Regresión lineal simple, en línea).

Lo que en primer lugar sería deseable en un modelo de regresión es que estos errores aleatorios sean en media cero para cualquier valor  $x$  de  $X$ , es decir,  $E[\varepsilon/X = x] = E[\varepsilon] = 0$ , y por lo tanto:

$$E[Y / X = x] = \beta_0 + \beta_1 x + E[\varepsilon/X = x] = \beta_0 + \beta_1 x$$

En dicha expresión se observa que:

- La media de Y, para un valor fijo x, varía linealmente con x.
- Para un valor x se predice un valor en Y dado por  $\hat{y} = E[Y / X = x] = \beta_0 + \beta_1 x$ , por lo que el modelo de predicción puede expresarse también como  $\hat{Y} = \beta_0 + \beta_1 X$ .
- El parámetro  $\beta_0$  es la ordenada al origen del modelo (punto de corte con el eje Y) y  $\beta_1$  la pendiente, que puede interpretarse como el incremento de la variable dependiente por cada incremento en una unidad de la variable independiente. Estos parámetros son desconocidos y se deben de estimar para poder realizar predicciones.

Además de la hipótesis establecida sobre los errores de que en media han de ser cero, se establecen las siguientes hipótesis:

- ii) La varianza de  $\varepsilon$  es constante para cualquier valor de x, es decir,

$$\text{Var}(\varepsilon / X = x) = \sigma^2$$

- iii) La distribución de  $\varepsilon$  es normal, de media 0 y desviación  $\sigma$ .

- iv) Los errores asociados a los valores de Y son independientes unos de otros.

En consecuencia, la distribución de Y para x fijo es normal, con varianza constante  $\sigma^2$ , y media que varía linealmente con x, dada por  $\beta_0 + \beta_1 x$ . Además los valores de Y son independientes entre sí.

### 2.6.1. ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO

Se parte de una muestra de valores de X e Y medidos sobre n individuos:  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , y quiere estimar valores en Y según el modelo  $\hat{Y} = \beta_0 + \beta_1 X$ , donde  $\beta_0$  y  $\beta_1$  son desconocidos. Se debe de encontrar entonces la recta que mejor se ajuste a los datos observados, es decir, aquellos valores de  $\beta_0$  y  $\beta_1$  que hagan mínimos los errores de estimación. Para un valor  $x_i$ , el modelo estima un valor en Y igual a  $\hat{y}_i = \beta_0 + \beta_1 x_i$  y el valor observado en Y es igual a  $y_i$ , con lo cual el error de estimación en ese caso vendría dado por  $e_i = y_i - \hat{y}_i = y_i - (\beta_0 + \beta_1 x_i)$ . Entonces como estimaciones de  $\beta_0$  y  $\beta_1$ , que se denotan por  $\hat{\beta}_0$  y  $\hat{\beta}_1$ , para aquellos valores que hagan mínima la suma de los errores al cuadrado, que viene dada por:

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2$$

De ahí que al método de estimación se le llame método de mínimos cuadrados. La solución se obtiene por el mecanismo habitual, derivando SSE con respecto a  $\beta_0$  y  $\beta_1$  e igualando a cero. Los estimadores resultan:

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Siendo:

$$SS_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n \bar{x} \bar{y},$$

$$SS_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n \bar{x}^2 = n \sigma_x^2$$

Donde ( $x_i$  e  $y_i$ ) representan cada par de variable independiente y dependiente dentro del modelo de regresión lineal.

A la recta resultante  $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$  se le llama recta de regresión lineal de Y sobre X.

Un último parámetro a estimar en el modelo es la varianza de los errores ( $\sigma^2$ ). A su estimador se le denomina varianza residual y viene dada por:

$$\hat{\sigma}_R^2 = \frac{SSE}{n-2} = \frac{\sum_{i=1}^n e_i^2}{n-2} = \frac{SS_{yy} - \beta_1 SS_{xy}}{n-2}$$

## 2.6.2. EL COEFICIENTE DE CORRELACIÓN LINEAL Y EL COEFICIENTE DE DETERMINACIÓN

**El coeficiente de correlación lineal.**- Como solución al inconveniente planteado, para medir la asociación lineal entre dos variables X e Y se utiliza una medida adimensional denominada coeficiente de correlación lineal, dado por:

$$r = \frac{E[(X - E(X))(Y - E(Y))]}{\sqrt{VAR(X)VAR(Y)}} = \frac{\sqrt{VAR(X)}}{\sqrt{VAR(Y)}}\beta_1$$

y su estimación a partir de datos de una muestra resulta:

$$\hat{r} = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}} = \frac{\sqrt{SS_{xx}}}{\sqrt{SS_{yy}}}\hat{\beta}_1$$

El coeficiente de correlación lineal toma valores entre -1 y 1 y su interpretación es la siguiente:

- Un valor cercano o igual a cero indica respectivamente poca o ninguna relación lineal entre las variables.
- Cuanto más se acerque en valor absoluto a 1 mayor será el grado de asociación lineal entre las variables. Un coeficiente igual a 1 en valor absoluto indica una dependencia lineal exacta entre las variables.
- Un coeficiente positivo indica asociación lineal positiva, es decir, tienden a variar en el mismo sentido.
- Un coeficiente negativo indica asociación lineal negativa, es decir, tienden a variar en sentido opuesto.

Nótese que si  $\beta_1 = 0$  entonces  $r = 0$ , en cuyo caso hay ausencia de linealidad. Por lo tanto, contrastar si el coeficiente de correlación lineal es significativamente distinto de cero sería equivalente a contrastar si  $\beta_1$  es significativamente distinto de cero.

**El coeficiente de determinación.-** El coeficiente de correlación lineal puede interpretarse como una medida de la bondad del ajuste del modelo lineal, concretamente, un valor del coeficiente igual a 1 o -1 indica dependencia lineal exacta, en cuyo caso el ajuste es perfecto. No obstante, para cuantificar la bondad del ajuste de un modelo, lineal o no, se utiliza una medida que se denomina coeficiente de determinación lineal ( $R^2$ ), que es la proporción de variabilidad de la variable Y que queda explicada por el modelo de entre toda la presente, y cuya expresión es:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SSE}{SS_{yy}},$$

que en un modelo de regresión lineal coincide con el cuadrado del coeficiente de correlación lineal:  $R^2 = r^2$

El coeficiente de determinación toma valores entre 0 y 1, y cuanto más se aproxime a 1 mejor será el ajuste y por lo tanto mayor la fiabilidad de las predicciones que con él realicemos. Nótese que si el coeficiente de correlación lineal  $r$  es igual a 1 o -1 entonces  $R^2 = 1$  y por lo tanto el ajuste lineal es perfecto.

### 2.6.3. ANÁLISIS DE RESIDUOS

Si las observaciones provienen de distribuciones normales todas con la misma varianza  $\sigma$ , entonces los residuos deberían mostrar una distribución aproximadamente normal. Para verificarlo se puede usar métodos gráficos o el test de Shapiro-Wilk. Métodos gráficos: Histograma / box-plot (box) / gráfico de probabilidad normal (normal probability plot). La mayoría de los supuestos puede verificarse usando los residuos  $(Y_i - \hat{Y}_i)$ , ellos representan la distancia de cada observación a la recta ajustada. Si se detectan observaciones outliers, se deberían controlar si son correctas e investigar si son influyentes. (Orellana, 2008).

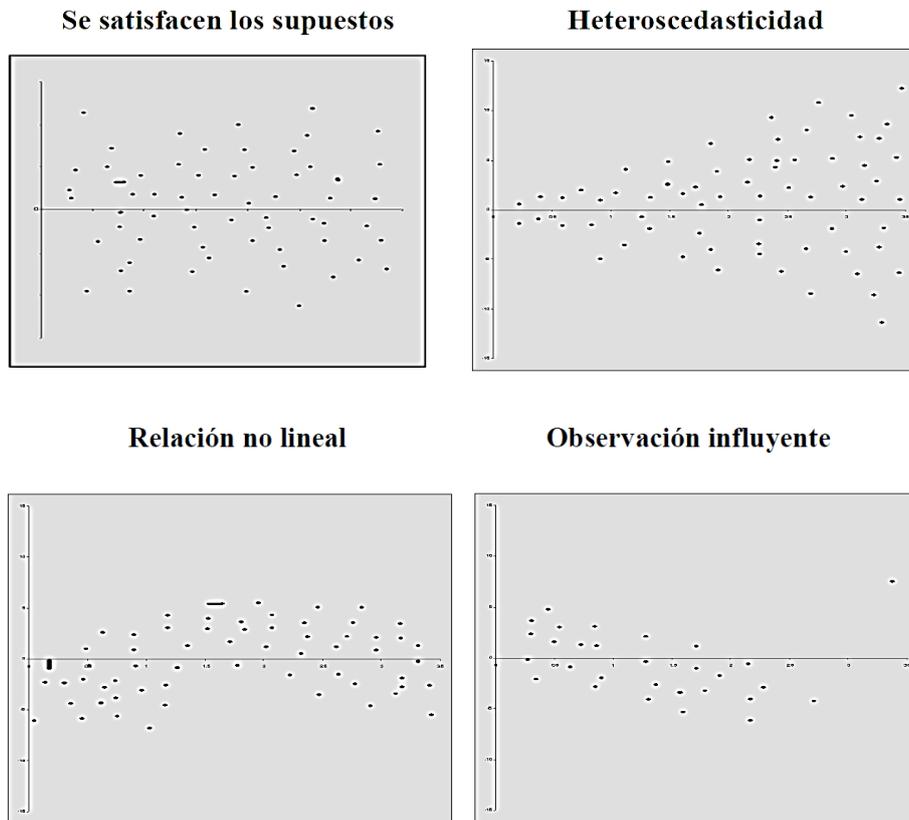
#### **Linealidad y homoscedasticidad**

Para verificar que el modelo lineal es una buena aproximación a la verdadera relación entre las variables y para constatar que el supuesto de homogeneidad de varianzas, se usa el gráfico de residuos versus valores predichos.

Si el gráfico muestra una nube de puntos alrededor de cero sin evidencia de estructura, tendencia o cambio de la dispersión, entonces no hay sospecha de que se violen ninguno de estos dos supuestos.

En general, los datos influyentes no son fácilmente detectables como outliers en gráficos de residuos vs predichos, porque ejercen “palanca” sobre la recta forzándola a pasar relativamente cerca de ellos. En consecuencia, pueden presentar residuos que no llaman la atención por su magnitud.

En la práctica los gráficos de residuos no son tan claros como estos y aun cuando los datos satisficieran todos los supuestos, la variabilidad muestral podría hacer que el gráfico tuviera pequeños apartamientos de la imagen ideal.



**Figura 8:** Gráficos para análisis de residuos

FUENTE: Orellana, 2008

### **Independencia**

El hecho de haber tomado una muestra aleatoria de sujetos desde alguna población asegura que, en principio, tendremos observaciones independientes. Algunas situaciones en las que este supuesto puede fallar se dan en estudios en los que los datos se recolectan secuencialmente, las observaciones consecutivas pueden no ser independientes, determinaciones de laboratorio hechas secuencialmente en el tiempo pueden mostrar un cierto patrón, dependiendo de cómo funcionan los equipos, los observadores, etc.

## 2.7. ANTECEDENTES

En la actualidad, en el Perú existen pocos estudios de modelos empíricos para pronóstico de temperaturas mínimas en campañas agrícolas, en el Valle del Mantaro se presentan las siguientes investigaciones como las más importantes.

García y Villegas (1993), presentan en su estudio pronóstico de temperaturas mínimas, a nivel de caseta meteorológica (1,60 m), para previsión de heladas con 18 horas de anticipación para la localidad de Huayao, mediante un modelo empírico, el conjunto de las variables están relacionados con los factores controladores del balance de energía y del efecto invernadero en el proceso de enfriamiento nocturno, además considera una variable de mayor influencia en el enfriamiento nocturno que es la temperatura mínima de un día anterior al que se va a pronosticar ( $T_{\min y-1}$ ), el cual va indicar la energía base o energía inicial disponible en el ambiente. Para la determinación o evaluación de las constantes ha utilizado correlación serial múltiple, en las que los modelos son adoptados como válidos si la prueba estadística de las constantes diferentes de cero resultan significativos para  $(1-\alpha) \times 100$  mayor de 90%. En lo que respecta a la verificación o validación, ella consistió en la aplicación del modelo determinado para los días de los meses de abril, mayo y junio (por coincidir estos meses con las fases fenológicas más críticas de la mayoría de los cultivos de la localidad) de 1986, además, en la verificación fijan como límites ciertos márgenes de errores, ello consiste en determinar la probabilidad empírica del error de las temperaturas pronosticadas que superan a  $-1.5\text{ }^{\circ}\text{C}$  y a  $-2.0\text{ }^{\circ}\text{C}$ . El modelo encontrado permite pronosticar temperaturas mínimas con 18 horas de anticipación y con la probabilidad de 75% en el grado de error de  $\pm 2\text{ }^{\circ}\text{C}$ .

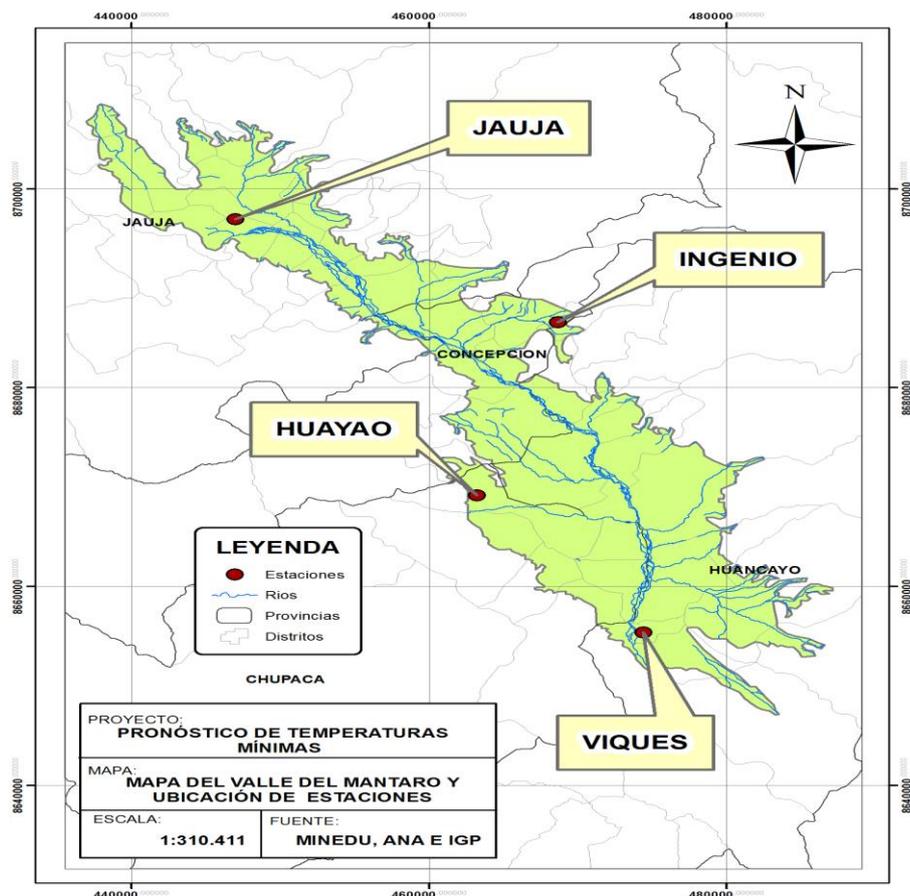
Sanabria (2009), presenta una tesis con el título "Calibración y validación de modelos de pronóstico de heladas en el valle del Mantaro"; aquí Sanabria hace uso los modelos de Lhomme y Cellier basados en procesos físicos relacionados a conservación de energía, también usa los modelos empíricos de Allen, García y Young basados en mediciones de humedad relativa, temperaturas del bulbo seco y húmedo, en tres localidades del valle (Huayao, Jauja y Santa Ana). De estos modelos, sólo los de Cellier y García fueron validados, además mostraron ser útiles sólo en determinadas zonas del valle, los de Cellier fueron calibrados y validados a nivel de 10 cm de altura para la localidad de Santa Ana y el de García se determinó y validó a nivel de caseta (1.60m) para las tres localidades. Sanabria

resalta que una de las principales fuentes de error de las salidas de los modelos Cellier y Lhomme fue la estimación de la cantidad de radiación infrarroja atmosférica.

### III. MATERIALES Y MÉTODOS

#### 3.1. ÁREA DE ESTUDIO

El estudio se realizó para el Valle del Mantaro (ver Figura 9), el cual está ubicada en el departamento de Junín, entre los paralelos 11°40' y 12°10' sur y los meridianos 75°10' y 75°15' oeste en la sierra central del Perú. Para la determinación de los modelos empíricos se tomó en cuenta 4 puntos, en los cuales se encuentra ubicada una estación meteorológica (Viques, Huayao, Ingenio y Jauja), estas estaciones fueron seleccionadas debido a que cuentan con datos disponibles hasta el 2016 y están ubicados dentro del valle.



**Figura 9:** Mapa del Valle del Mantaro y ubicación de estaciones

FUENTE: Elaboración propia

## **3.2. DATOS Y SOFTWARES**

Se hizo uso de datos de pronósticos del modelo global de CPTEC (<http://brams.cptec.inpe.br/> ) y datos registrados de temperatura mínima, de 4 estaciones meteorológicas de SENAMHI (ver Anexo 01) ubicadas en el valle del Mantaro (departamento de Junín, Perú) al inicio y final de la campaña agrícola (setiembre-diciembre y febrero-mayo) para el período de 2012-2016.

La UNALM viene implementando una base de datos de pronóstico de salidas del modelo global del CPTEC, desde hace cinco años atrás, esta información tiene una resolución temporal de 6 horas y espacial de 110km.

Los datos de entrada al modelo son altura geopotencial (hgt), temperatura (T), humedad relativa (HR), viento zonal (u) y viento meridional (v).

Los softwares usados en el estudio fueron: modelo BRAMS 4.2, R Studio, Office, GrADS, QGIS y Google Earth.

## **3.3. METODOLOGÍA**

### **3.3.1. PARA LA DETERMINACIÓN DE LOS MODELOS EMPÍRICOS**

#### **a. EJECUCIÓN DEL MODELO BRAMS Y DETERMINACIÓN DE TEMPERATURAS PRONOSTICADAS**

Los datos del modelo global del CPTEC primeramente fueron convertidos (creación de los dp's), para poder ser utilizados en el modelo regional BRAMS; después se procedió a realizar el pre procesamiento de los datos, corrida o procesamiento y el post procesamiento de las salidas (BRAMS, 2007). (Ver Figura 10).

**Conversión de archivos globales CPTEC.-** Se crearon dp's a partir de los datos del modelo global del CPTEC para poder ser utilizados en el modelo regional BRAMS. CPTEC produce archivos en formato grib que contienen el estado de la atmósfera en una rejilla gaussiana (13 niveles verticales, coordenadas sigma) que abarca América del Sur. BRAMS no puede leer este formato de archivo - espera un archivo ASCII (dpYYYY-MM-DD-HHHH) en una cuadrícula regular que contiene el estado de la atmósfera.

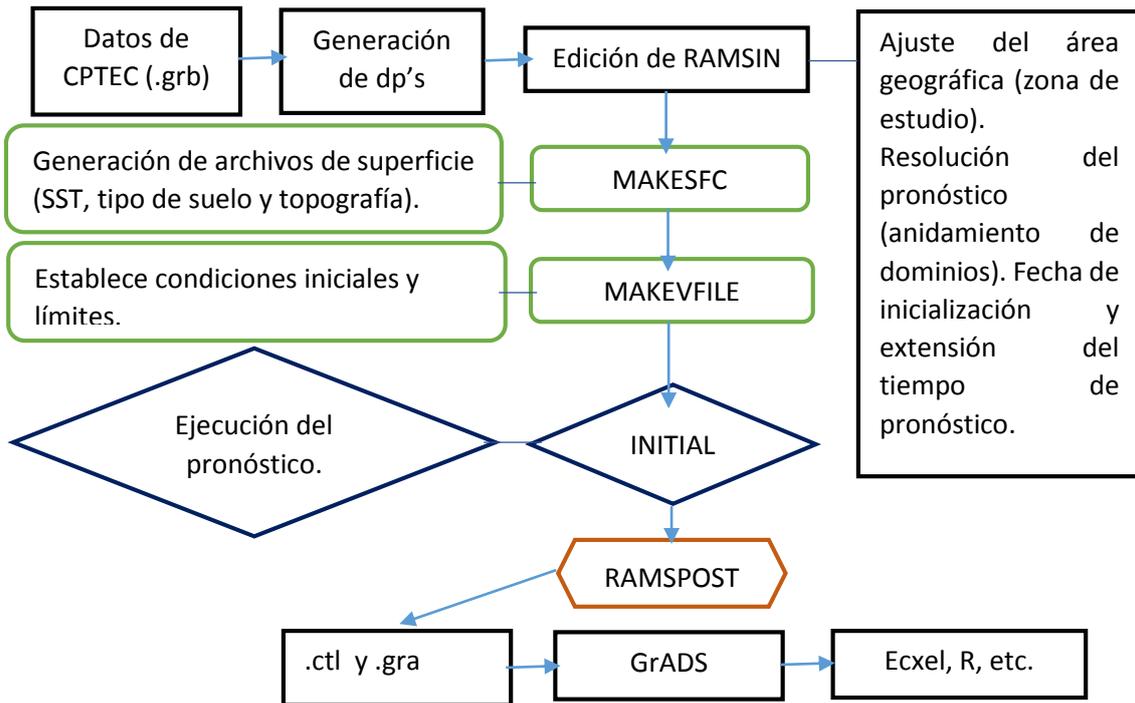
**Construcción de archivos de superficie para el área limitada deseada.-** Se ejecutó BRAMS con RUNTYPE establecido en MAKESFC (en RAMSIN). Con esta ejecución, los archivos de datos globales con temperatura de la superficie del mar, tipo de suelo y topografía se convirtieron en archivos que sólo cubren el área limitada (y cuadrícula) como se define en la lista de nombres \$ MODEL\_GRIDS (en RAMSIN).

**Establecimiento de condiciones iniciales y límites para el área limitada deseada.-** Se ejecutó BRAMS con RUNTYPE establecido en MAKEVFILE. Esta ejecución produjo archivos con condiciones iniciales y límites para el área de interés. Esto se realizó cortando los archivos dp (salida del primer paso) a las áreas geográficas, las cuadrículas y el intervalo de tiempo deseados, como se define en la lista de nombres \$ MODEL\_GRIDS (en RAMSIN). Los archivos de salida se conocen como VARFILES (es por eso que RUNTYPE se establece en MAKEVFILE). Cada VARFILE contiene el viento horizontal, la presión, la temperatura potencial y los campos de mezcla de vapor interpolados a la (s) rejilla (s) del modelo y listo para la inicialización.

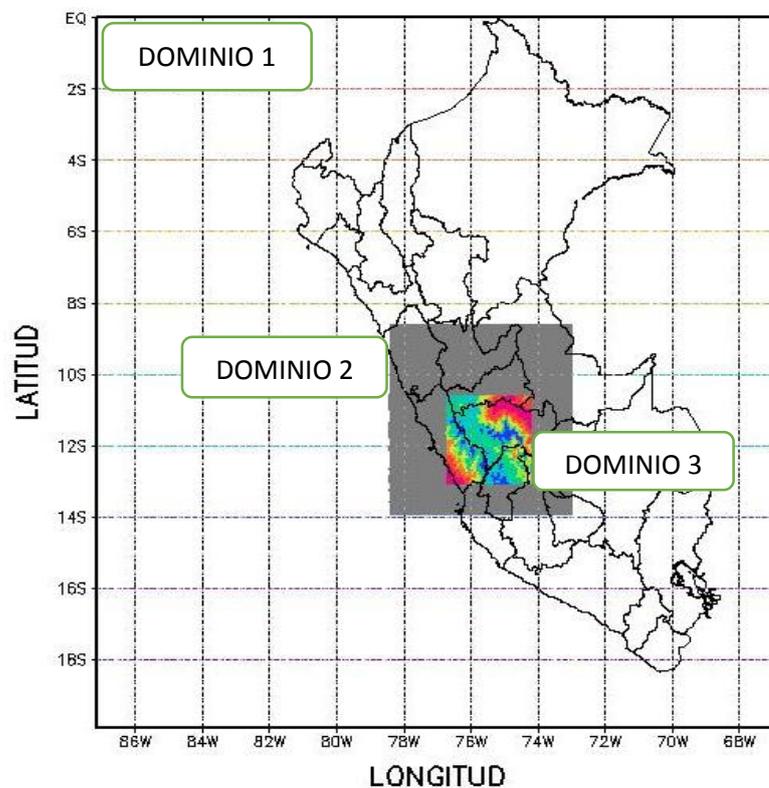
**Ejecución del Pronóstico.-** En este paso, se realizó la previsión con BRAMS estableciendo RUNTYPE en INITIAL. Esta configuración significa que una simulación comenzará desde el tiempo cero y se ejecutará hasta el tiempo especificado por la variable TIMEMAX. Esta ejecución supone que todas las variables atmosféricas y del suelo, que se van a pronosticar, se inicializan a partir de un varfile preparado anteriormente.

**Post procesamiento.-** Se realizó el post-procesamiento en el directorio RAMSPOST. Se utilizó el programa ramspost60.x, en el directorio RAMSPOST, para post procesar los archivos de pronóstico BRAMS. Toma como entrada el archivo de lista de nombres ramspost.inp en el mismo directorio y produce los archivos de salida result\_g1.ctl y result\_g1.gra, en el mismo directorio. El archivo result\_g1.ctl describe el contenido de result\_g1.gra en un formato adecuado para grads.

Para el procesamiento respectivo, el modelo se configuró en tres dominios (ver Figura 11), el primero de 72km, el segundo de 12km y el tercero de 3km y para el estudio se tomó en cuenta el tercer dominio. Así mismo, el tiempo de pronóstico se determinó para 7 días y con frecuencia horaria de las salidas.



**Figura 10:** Flujo grama básico del funcionamiento del Modelo BRAMS.



**Figura 11:** Configuración en tres dominios del modelo BRAMS.

Para el análisis de las salidas del BRAMS se tuvo en cuenta 4 estaciones meteorológicas de SENAMHI (Ver Anexo 1), ubicadas en el Valle del Mantaro (ver Figura 8), considerando que estas estaciones cuentan con datos de temperatura mínima durante el periodo de investigación.

Por último se prepararon scripts en GrADS con la finalidad de exportar los datos de temperatura mínima pronosticadas (se usó la variable *Tempc2m*, que representa la temperatura pronosticada a 2m), para los puntos seleccionados en el paso anterior. Los datos se exportaron a una hoja de cálculo de Excel donde fueron ordenados por inicio y fin de campaña agrícola, la temperatura mínima pronosticada se obtuvo extrayendo la menor temperatura entre las 00hrs y 12hrs locales, considerando que la temperatura mínima se presenta alrededor de las 7hrs. Este análisis se realizó semanalmente durante las campañas agrícolas comprendidas entre el 2012 y 2016.

## **b. ANÁLISIS MENSUAL DE DATOS OBSERVADOS ( $T_{min}$ ) EN EL PERIODO DE ESTUDIO**

Este análisis se realizó con los datos de las estaciones meteorológicas (ver Anexo 1) para los meses de febrero, marzo, abril, mayo, setiembre, octubre, noviembre y diciembre en el periodo de 2012 a 2016, con la finalidad de constatar la calidad de los datos que se tomaron en cuenta en el estudio. Para lo cual se realizaron los siguientes análisis gráficos.

### **Gráficos de series de tiempo**

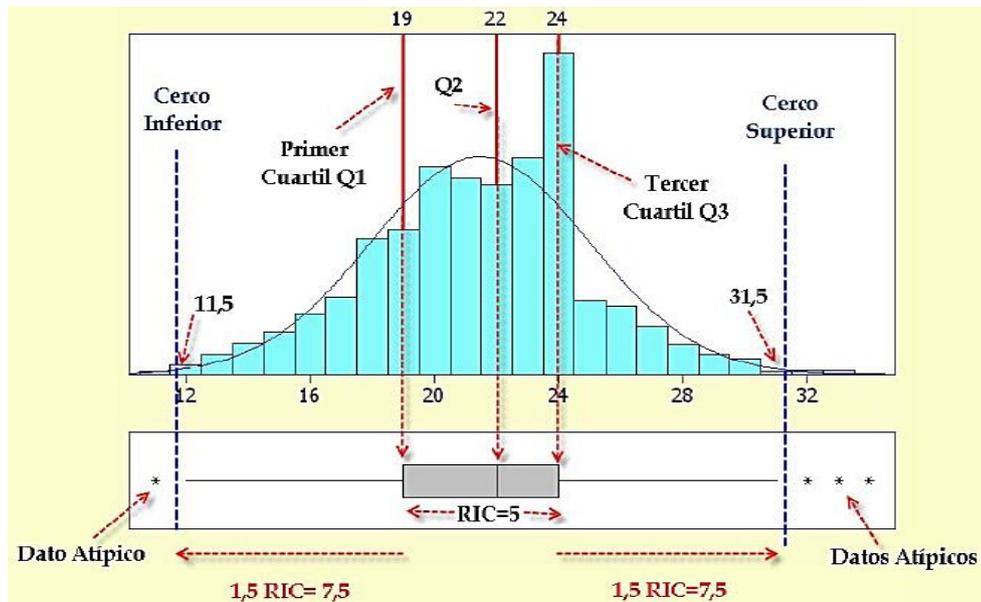
Este gráfico nos permitió evaluar el comportamiento de la variable durante todo el periodo de estudio y así poder identificar alguna anomalía en los datos de las estaciones, ya que, de este universo de datos provienen los datos observados que intervienen en los modelos empíricos.

### **Diagrama de cajas (Box-Plot)**

Este diagrama nos permitió resumir la complejidad de los datos, además nos proporcionó una muy buena idea de su distribución y especialmente nos ayudó en la identificación de datos atípicos (ver Figura 12).

El Box-Plot es un gráfico muy usado que involucra los tres cuartiles (dividen la muestra en cuatro partes iguales) junto con el mínimo y el máximo de los datos, la caja está delimitada por los cuartiles  $Q_1$  (percentil 25) y  $Q_3$  (percentil 75), en el interior de la caja aparece una

línea divisoria que corresponde al segundo cuartil Q2 (percentil 50, la mediana). La longitud de la caja está determinada por el “RIC” (*Rango Inter Cuartílico*), el cual representa la diferencia entre los cuartiles extremos, los cercos se encuentran a una distancia de 1,5 veces el RIC, medida a partir de los cuartiles extremos, los datos menores que el cerco inferior y mayores que el cerco superior se consideraron como atípicos.



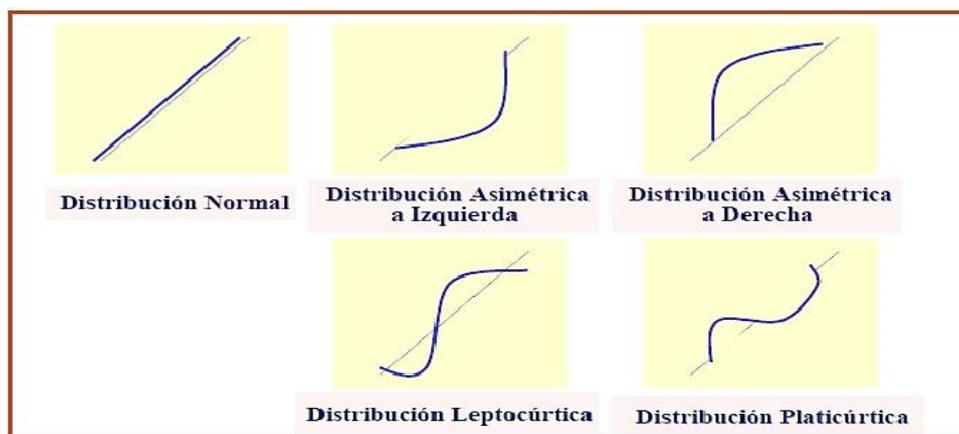
**Figura 12:** Representación de las características de un Box-Plot

FUENTE: (Behar, 2008)

### c. ANÁLISIS DE DATOS QUE SE TOMARON EN CUENTA PARA LA DETERMINACIÓN DE LOS MODELOS EMPÍRICOS

El análisis exploratorio se realizó con la finalidad de determinar el ajuste de los datos (tanto de las temperaturas mínimas pronosticadas por el modelo BRAMS y las observadas en las estaciones meteorológicas) a la distribución normal, lo cual facilitó el uso de las técnicas de análisis estadísticos, para ello se hizo uso del software R Studio y del método gráfico, en este caso se trabajó con el Diagrama de Cuantiles (Q-Q plots). De igual modo se realizaron gráficos de dispersión entre las temperaturas mínimas pronosticadas por el modelo BRAMS y las temperaturas mínimas de las estaciones, estos graficos nos permitieron identificar el tipo de relación que presentan nuestros datos. Ambos análisis se realizaron para cada una de las estaciones y en los tiempos de pronóstico dos, tres, cuatro y cinco (T2, T3, T4 y T5) de cada campaña agrícola (inicio y fin).

Los diagramas de cuantiles comparan en un sistema de coordenadas cartesianas, los cuantiles muestrales (eje X) con los cuantiles esperados bajo la hipótesis normalidad. Si la distribución de partida es normal dichos diagramas tenderán a ser rectas que pasan por el origen. Cuanto más se desvíen de una recta, menos normales serán los datos (ver figura 13).



**Figura 13:** Tipos de distribución según Diagrama de Cuantiles

FUENTE: Salvador, 2003

Una vez determinado el análisis de normalidad y el tipo de relación de los datos se procedió a determinar los modelos empíricos.

#### **d. DETERMINACIÓN DE LOS MODELOS EMPÍRICOS DE PRONÓSTICO DE TEMPERATURAS MÍNIMAS**

Una vez detectada una relación significativa entre las dos variables ( $T_{min}$  y  $T_{min\_P}$ ), el siguiente paso fue determinar una fórmula matemática que formalice esa relación y que permita calcular pronósticos de una variable a partir de la otra.

La determinación de los Modelos empíricos se realizó para los tiempos de pronóstico ( $T_2$ ,  $T_3$ ,  $T_4$  y  $T_5$ ) del modelo BRAMS, teniendo en cuenta que el  $T_2$  representa el pronóstico para el día 2 y así respectivamente.

Se determinaron modelos de regresión lineal simple de las temperaturas mínimas. El modelo utilizó como variables predictoras las temperaturas mínimas pronosticadas por el modelo BRAMS ( $T_{min\_P}$ ). Este modelo pronostica la temperatura mínima del aire a nivel de caseta

meteorológica. Los coeficientes del modelo empírico de la temperatura mínima fueron calculados mediante el ajuste de los datos a un modelo de regresión lineal simple.

Para el periodo de inicio de campaña, se consideró la información de los meses de setiembre a diciembre de cada año y los modelos están expresados de la siguiente manera:

$$T_{\min\_IC} = b_1 + b_2 * T_{\min\_P}$$

Para el periodo de fin de campaña, se consideró la información de los meses de febrero a mayo de cada año y los modelos están expresados de la siguiente manera:

$$T_{\min\_FC} = b_1 + b_2 * T_{\min\_P}$$

$T_{\min\_IC}$ : Temperaturas mínimas de inicio de campaña.

$T_{\min\_FC}$ : Temperaturas mínimas de fin de campaña.

$T_{\min\_P}$ : Temperaturas mínimas pronosticadas por el modelo BRAMS.

$b_n$ : Son las constantes determinadas como factor de aporte de cada una de las estaciones meteorológicas.

### **PRUEBA ESTADÍSTICA DE FISHER A LOS MODELO EMPÍRICO:**

La prueba estadística de Fisher se utilizó principalmente para probar la igualdad entre dos varianzas y por lo tanto nos permitió determinar cambios importantes en el comportamiento de la variable en estudio.

La distribución F se utilizó para demostrar la hipótesis nula

$H_0: b_1 = b_2 = 0,$

En caso contrario la hipótesis planteada será diferente de cero.

$H_a: b_1 \neq b_2 \neq 0,$

Como parte de la prueba de hipótesis se calculará el  $F_c$  (F calculado) y  $F_t$  (F tabular), con un nivel de significancia de  $\alpha = 0.05$ ; para  $F_c > F_t$ , se rechazó la hipótesis nula y se aceptó la hipótesis alternante, con lo cual se concluyó que las ecuaciones obtenidas son correctas.

El estadístico  $F_c$  y el  $F_t$  se calcularon con:

$$F_c = \frac{R^2(N - 1 - P)}{P(1 - R^2)}$$

$$F_t = [(1 - \alpha); P, (N - 1 - P)]$$

Donde:

N= número de datos (tamaño de la muestra).

P= número de parámetros de variables predictoras (b2).

R<sup>2</sup>= coeficiente de determinación.

## **VALIDACIÓN DE LOS MODELOS EMPÍRICOS CON RMSE, BIAS Y EL COEFICIENTE DE DETERMINACIÓN Y ANÁLISIS DE RESIDUOS:**

Para el proceso de validación de los modelos empíricos, se comparó las temperaturas mínimas observadas con las determinadas por los modelos empíricos, y estos se evaluaron a través de una serie de indicadores estadísticos (error cuadrático medio, bias y el coeficiente de determinación). Un modelo fue considerado válido cuando los indicadores de error (RMSE y Bias) tuvieron valores bajos y los coeficientes de determinación valores altos. Estos indicadores se explican a continuación.

### **a. Error cuadrático medio (RMSE) Y Bias (B)**

El RMSE, mide el error total, incluye los componentes sistemáticos y al azar. El RMSE es una medida de aproximación entre valores observados y simulados, pone una mayor influencia sobre los errores grandes más que los errores pequeños, esto puede ser especialmente perjudicial en la presencia de posibles valores atípicos en los conjuntos de datos. La desventaja es que no indica la dirección de las desviaciones. Su alcance es de cero al infinito, con una puntuación perfecta de RMSE = 0. (Jiménez, 2014).

El Bias, representa el error medio o sistemático (sesgo). Debido a que el Bias carece de términos en valor absoluto o al cuadrado, se permite la cancelación de errores individuales positivos y negativos, poniendo en evidencia solamente el exceso, es decir el error sistemático. El rango del Bias va desde menos infinito hasta infinito, y una puntuación

perfecta es, Bias = 0. Sin embargo, es posible llegar a una puntuación perfecta para un conjunto de datos con errores grandes, si hay errores de compensación de un signo inverso. El Bias no es una medida de precisión, ya que no proporciona información sobre la magnitud de los errores de pronóstico. Una medida sencilla para compensar los posibles errores positivos y negativos que puede generar el Bias es mediante el cálculo del RMSE. (Jiménez, 2014).

Estos se calcularon con:

$$RMSE = \sqrt{\sum_{i=1}^n (s_i - m_i)^2 / n} \quad B = \sum_{i=1}^n (s_i - m_i) / n$$

Donde:

$s_i$ : valor estimado del modelo empírico

$m_i$ : valor observado correspondiente

$n$ : número de datos

### **b. Coeficiente de determinación**

El coeficiente de determinación ( $R^2$ ) es una medida que nos permitió inferir qué tan bien se ajusta la recta de regresión muestral a los datos, es decir, fue una medida útil que nos mostró la proporción de varianza (variación) de una variable que es predecible a partir de la otra variable. El  $R^2$  es equivalente al cuadrado del coeficiente de correlación ( $r$ ).

El coeficiente de determinación varía entre  $0 < R^2 < 1$ , y denota el grado de la asociación lineal entre  $s_i$ = valor simulado y el  $m_i$ = valor observado.

El cual se calculó con:

$$R^2 = \left\{ \frac{\sum (s_i - S)(m_i - M)}{\sqrt{\sum (s_i - S)^2} \sqrt{\sum (m_i - M)^2}} \right\}^2$$

Donde:

$s_i$ = valor simulado de temperatura mínima

$m_i$ = valor observado de temperatura mínima

S y M= Valores promedio

Por ejemplo, si  $r = 0.922$ , entonces  $R^2 = 0.850$ , lo que significa que el 85% de la variación total en los valores simulados se pueden explicar por la relación lineal que existe entre los valores observados y simulados y que el otro 15% de la variación total en los valores simulados, permanece sin explicación.

Para evaluar la validez de los modelos se establecieron criterios limitantes a estos indicadores estadísticos: El *Bias* debe de estar en el intervalo  $\pm 2.0^\circ C$  y el diagrama de dispersión entre el valor simulado (si) y el valor observado (mi) deben de analizarse a un nivel de significancia de 95%.

### **c. Análisis de residuos**

Que la recta se ajuste a los datos no significa que el modelo sea correcto, depende del uso que se le quiera dar. Si sólo se pretende hallar la relación entre dos variables, con calcular la recta de mínimos cuadrados es suficiente, esa recta describe la relación entre las variables, otra cosa es que los datos tengan una buena relación lineal. Podría ser que los datos tengan muy mala relación lineal y la recta seguiría existiendo. En cambio si pretendemos describir la estructura general de los datos, o inferir/predecir con la recta de regresión se debe de comprobar que se verifican unas reglas ya establecidas y aceptadas que aseguran que el modelo es bueno. (Ferre, 2014).

Los modelos lineales de regresión simple, tienen asociado un error de predicción (a los que se denomina residuos), ellos representan la distancia de cada observación a la recta ajustada, reflejan la ausencia de dependencia perfecta entre las variables y la incertidumbre en su relación. Así mismo el diagnóstico de los residuos permite evaluar las características de los datos usados en la determinación de los modelos, tales como, normalidad, homogeneidad de varianzas, linealidad, etc. (Ferre, 2014).

Las hipótesis básicas del modelo de regresión lineal simple son: Incorrelación o independencia, varianza constante y normalidad de los residuos.

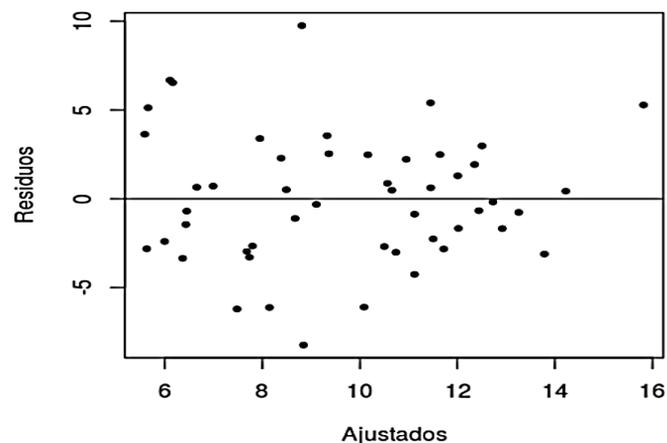
Una vez que se obtuvo el modelo ajustado se procedió con su diagnóstico, el cual se realizó a través del análisis de los residuos.

#### **- Análisis gráfico**

Las hipótesis de linealidad, homocedasticidad e independencia se contrastaron a través de un análisis gráfico que enfrenta los valores de los residuos con los valores ajustados. Estos análisis se realizaron usando la librería “Rcmdr” dentro del software R Studio, en el que se consideraron los datos del modelo BRAMS y los de estaciones, con los cuales se generó modelos de regresión lineal, a partir de las cuales se obtuvieron los gráficos de dispersión entre los residuos y los valores pronosticados y los gráficos qqplot de los residuos para determinar la normalidad de los datos.

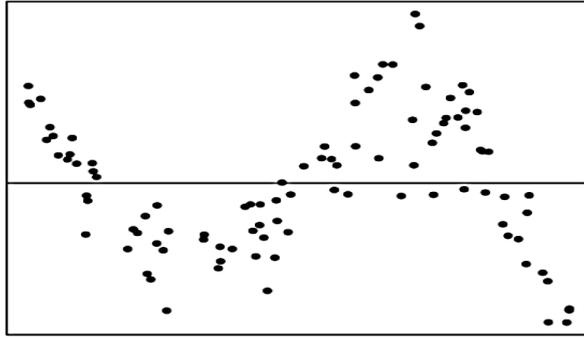
En la Figura 14, se muestra un ejemplo de cuando se satisfacen las hipótesis básicas de un modelo de regresión lineal simple (mencionados anteriormente), en este caso los residuos formaron una nube de puntos sin estructura y con, aproximadamente, la misma variabilidad por todas las zonas.

Los residuos de la Figura 15 muestran una estructura que sugiere una relación no lineal entre las variables, y la Figura 16 sugiere la ausencia de homocedasticidad.



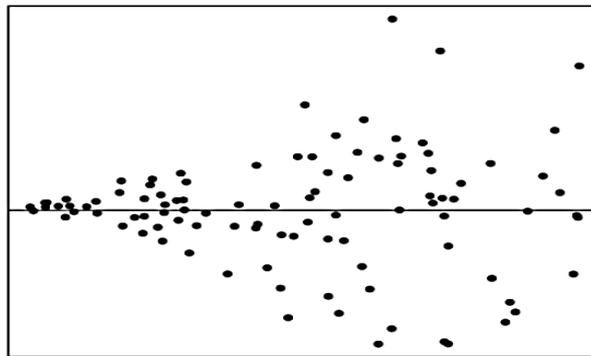
**Figura 14:** Representación del análisis gráfico de residuos que cumple con los supuestos para un modelo de regresión lineal

FUENTE: Ferre, 2014



**Figura 15:** Representación del análisis gráfico de residuos que incumple con el supuesto de linealidad

FUENTE: Ferre, 2014



**Figura 16:** Representación del análisis gráfico de residuos que incumple con el supuesto de homocedasticidad

FUENTE: Ferre, 2014

**-Análisis analítico**

De igual modo las hipótesis de linealidad, varianza constante, incorrelación y normalidad se comprobaron analíticamente. Estos análisis se llevaron a cabo haciendo uso de la librería “Rcmdr” del software R Studio, se trabajó con los datos de Tmin y Tmin\_P que se tomaron en cuenta en la determinación de los modelos empíricos, con los cuales se crearon modelos de regresión lineal y se aplicaron los siguientes tests.

**Test de Normalidad.-** Se realizó el Shapiro-Wilk normality test, donde el p-valor obtenido fue comparado con el nivel de significancia de ( $\alpha = 0.05$ ) con el que se trabajó, se consideró que los datos presentan distribución normal cuando el p-valor fue mayor que el nivel de significancia.

**Homogeneidad de varianzas.-** Se realizó el studentized Breusch-Pagan test, donde se obtuvo un p-valor y el cual fue comparado con el nivel de significancia, se consideró homogeneidad de varianzas cuando el p-valor fue mayor al nivel de significancia.

**Autocorrelación.-** Se utilizó el test de Durbin-Watson, en el contraste de autocorrelación se consideró que no existe correlación entre los residuos cuando se obtuvo un p-valor superior al nivel de significancia.

**Linealidad.-** Se aplicó el RESET test, se consideró que existe relación lineal cuando el p-valor, obtenido en el test, fue mayor al nivel de significancia.

### **3.3.2. PARA LA DETERMINACIÓN DE LA DISTRIBUCIÓN ESPACIAL**

Para la determinación de la distribución espacial de las temperaturas mínimas pronosticadas se tomó en cuenta los pronósticos generados por el modelo empírico para el inicio y fin de la campaña agrícola del año 2015. Se generarán mapas de temperaturas mínimas pronosticadas, haciendo uso del software QGIS, considerando los 4 puntos tomados en cuenta en el área de estudio y especificados en el Anexo 1. De tal forma que se podrá conocer la intensidad de temperaturas mínimas y por consiguiente, las zonas de heladas agronómicas y/o meteorológicas.

## **IV. RESULTADOS Y DISCUSIONES**

### **4.1. DETERMINACIÓN DE LOS MODELOS EMPÍRICOS DE PRONÓSTICO DE TEMPERATURAS MÍNIMAS**

#### **4.1.1. ANÁLISIS EXPLORATORIO DE LOS DATOS**

En esta parte se realizaron gráficos mensuales, de temperaturas mínimas de estación, mediante Box-Plot y series de tiempo, para las cuatro estaciones (Viques, Huayao, Ingenio y Jauja).

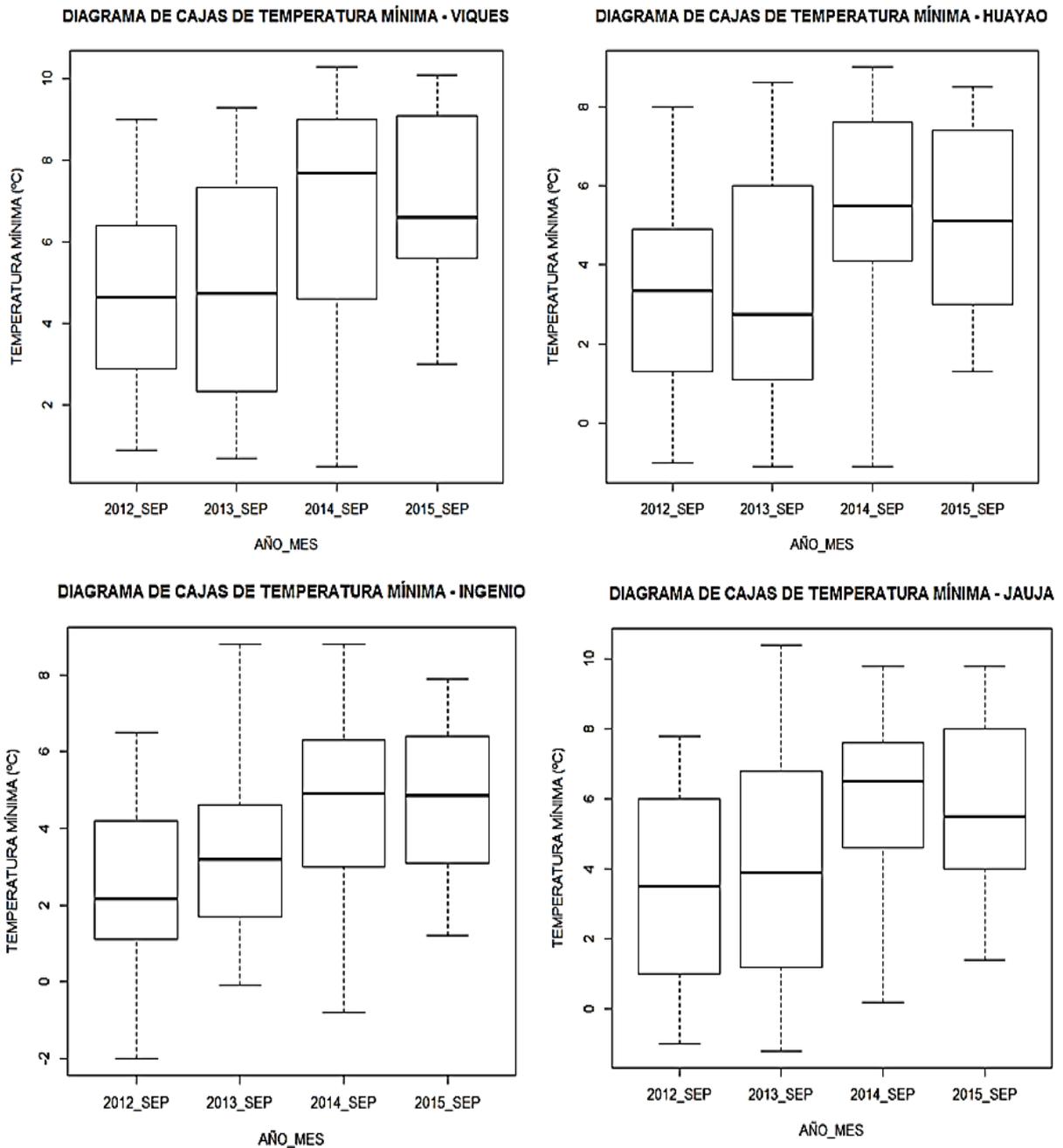
Referente a inicio de campaña, según los gráficos Box-Plot de temperatura mínima, en general presentaron datos más homogéneos (la amplitud y mediana de los datos fueron más parecidas en estos meses) que en los meses de fin de campaña. Se encontró pocos datos atípicos (Uno por mes) especialmente en las estaciones de Ingenio y Jauja, estos se verificaron y se determinaron como datos correctos. La amplitud de datos (diferencia entre la temperatura mínima más baja y la más alta) fue mayor en el mes de septiembre y menor en diciembre, las cuatro estaciones presentan características similares de sus datos en los diferentes meses, sin embargo, Viques fue la que presentó datos de temperatura mínima más altos y Jauja e Ingenio los más bajos, así también la variación mensual de sus medianas fue similar a la determinada en una investigación del IGP (IGP, 2010).

En las Figuras 17 y 18, se muestran los gráficos Box-Plot y las series de tiempo para el mes de septiembre y para las cuatro estaciones de estudio, se encontró la temperatura mínima más alta en Viques (10 °C) y la más baja en Ingenio (-2 °C), la amplitud varió aproximadamente en 10°C y la mediana mensual aproximada fue 4 °C en todas las estaciones.

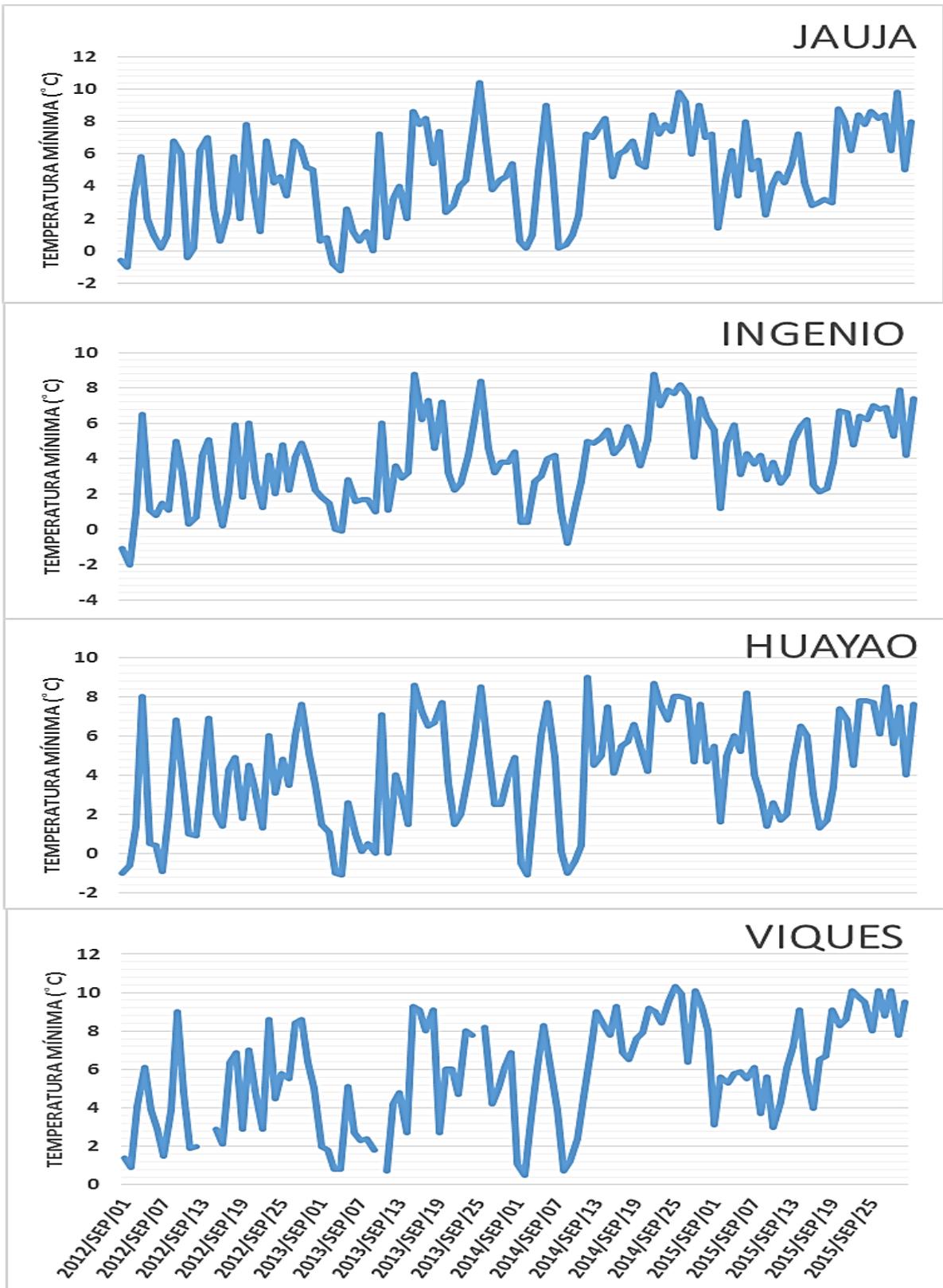
Para los meses de octubre, noviembre y diciembre (ver Anexo 2): octubre, presentó un dato máximo de 11 °C en Jauja y un mínimo de -1 °C en Huayao, su amplitud de datos varió en aproximadamente 8,5 °C y su mediana fue aproximadamente 6,5 °C en todas las estaciones; noviembre, presentó un máximo de 12 °C en Viques y un mínimo de 0 °C en Huayao, su amplitud de datos varió en aproximadamente 9 °C y su mediana fue 6 °C aproximadamente

en todas las estaciones; diciembre, presentó un máximo de 12 °C en Viques y un mínimo de 2 °C en Huayao, su amplitud de datos varió en 7 °C aproximadamente y su mediana fue 7,5 °C aproximadamente en todas las estaciones.

**a. Gráficos Box-Plot y series de tiempo para los meses de Inicio de Campaña.**

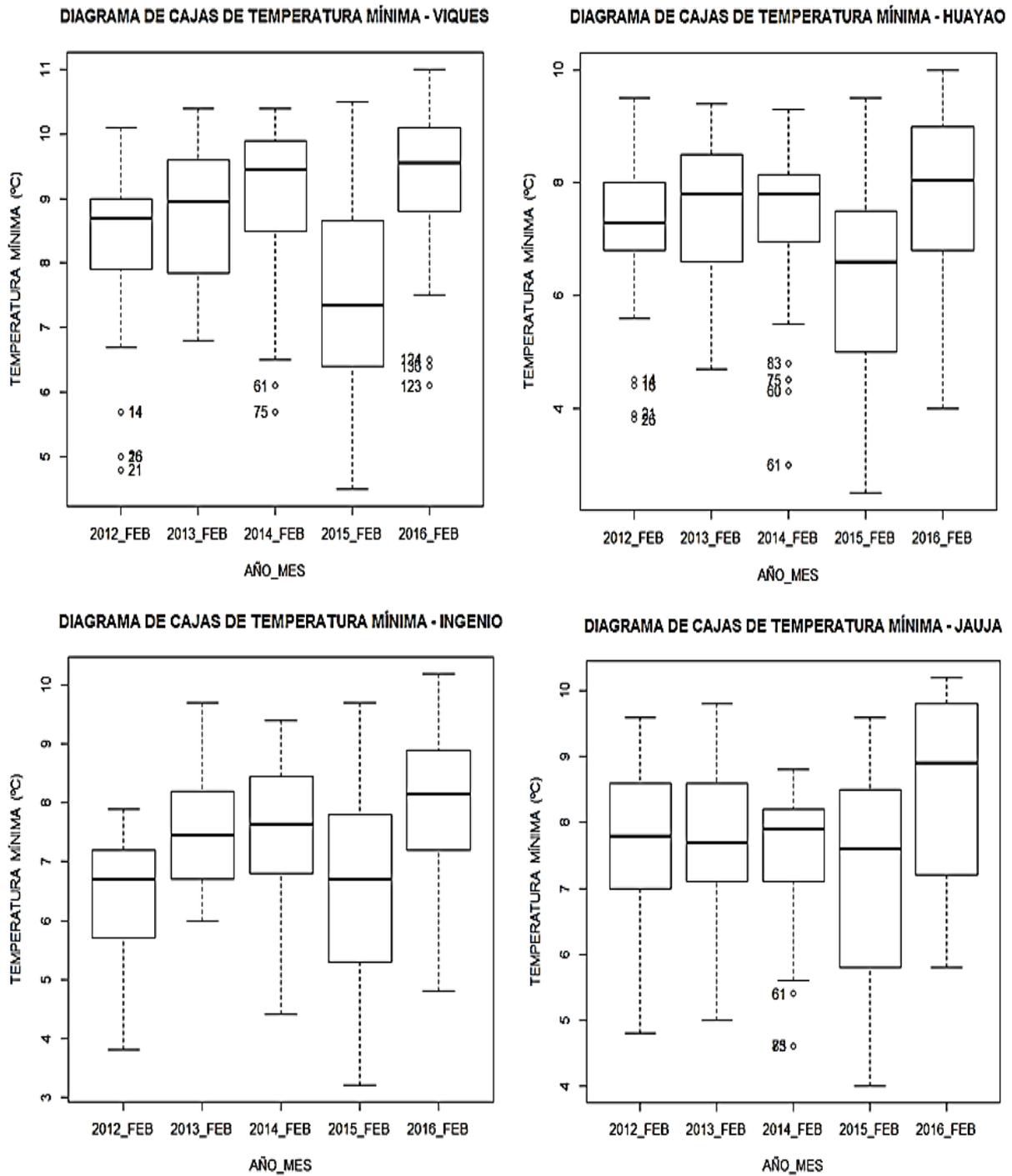


**Figura 17:** Diagrama de cajas de temperatura mínima de las cuatro estaciones de estudio para el mes de septiembre.



**Figura 18:** Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de septiembre.

**b. Gráficos Box-Plot y series de tiempo para los meses de Fin de Campaña.**



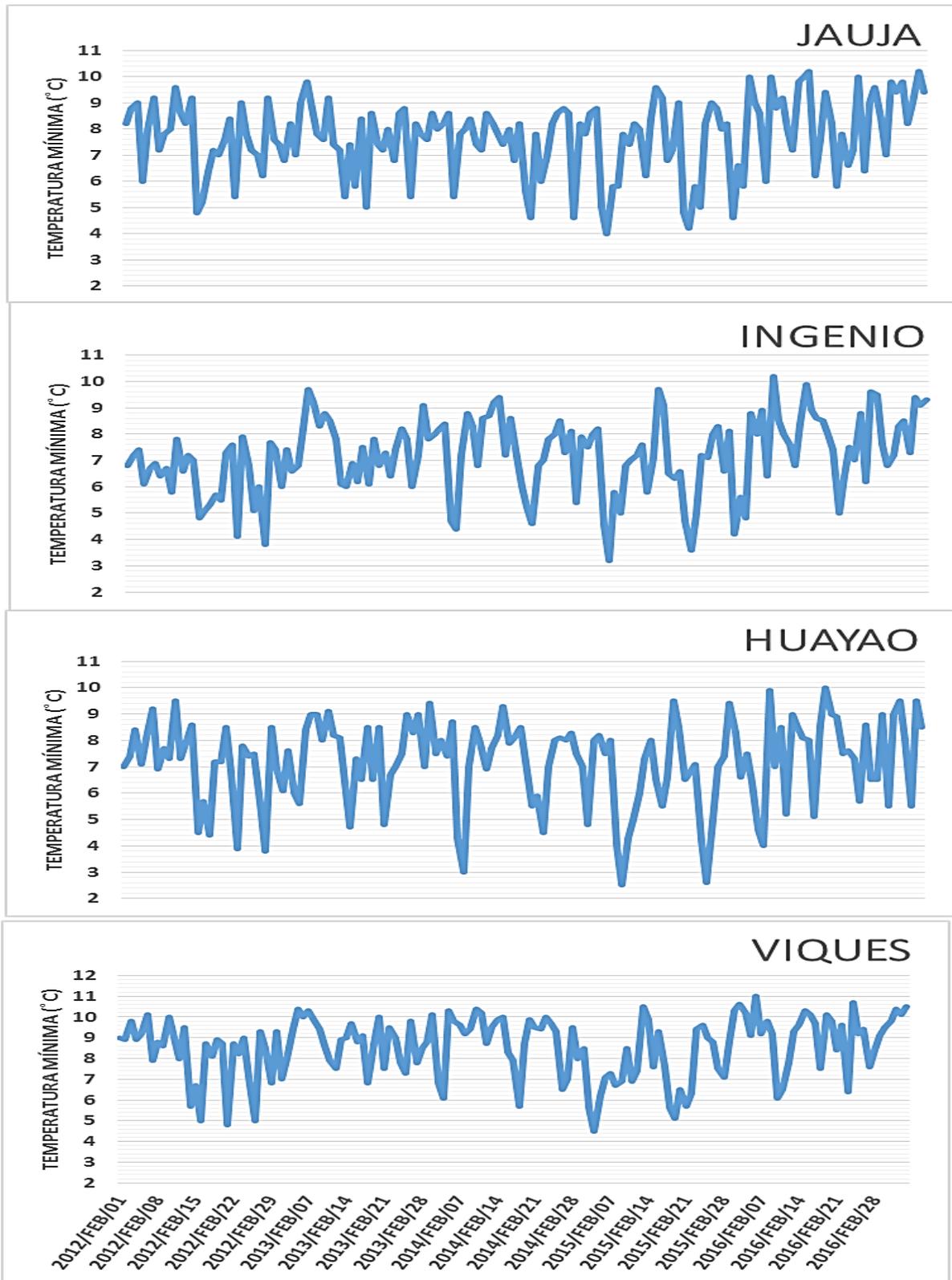
**Figura 19:** Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de febrero.

En los meses de fin de campaña, según los gráficos Box-Plot y las series de tiempo de temperatura mínima, se encontraron más datos atípicos (hasta cuatro por mes) especialmente en las estaciones de Viques y Huayao, estos se verificaron y se determinaron como datos correctos. La amplitud de datos (diferencia entre la temperatura mínima más baja y la más alta) fue mayor en el mes de mayo y menor en febrero, las cuatro estaciones presentan características similares de sus datos en los diferentes meses, sin embargo, Viques siguió siendo el que presentó datos de temperatura mínima más altos y Jauja e Ingenio los más bajos, así también la variación mensual de sus medianas fue similar a la encontrada en una investigación del IGP (IGP, 2010).

En las Figuras 19 y 20, se muestran los Box-Plot y las series de tiempo para el mes de febrero y para las cuatro estaciones de estudio, se encontró la temperatura mínima más alta en Viques de 11 °C y la más baja en Huayao de 2 °C, la amplitud varió aproximadamente en 6,5 °C y la mediana mensual fue aproximadamente 7,5 °C en todas las estaciones.

Para los meses de marzo, abril y mayo (ver Anexo 2): marzo, presentó un valor máximo de 11 °C en Viques y un valor mínimo de 1 °C en Huayao, su amplitud de datos varió en aproximadamente 7,3 °C y su mediana fue aproximadamente 7 °C en todas las estaciones; abril, presentó un máximo de 10 °C en Viques y un mínimo de 0 °C en Jauja, su amplitud de datos varió en aproximadamente 9 °C y su mediana fue 6 °C aproximadamente en todas las estaciones; mayo, presentó un máximo de 10 °C en Viques y un mínimo de -4 °C en Ingenio, su amplitud de datos varió en 10 °C aproximadamente y su mediana fue 4 °C aproximadamente en todas las estaciones.

Teniendo en cuenta los resultados del análisis mensual de las temperaturas mínimas de cada una de las estaciones, tanto para inicio y fin de campaña agrícola, se determinó que la variación mensual de las temperaturas mínimas está fuertemente asociada a la estacionalidad (IGP, 2010), es así que, se encontró mayor variabilidad de los datos de temperatura mínima (tanto en cada una de las estaciones como entre las estaciones) en los meses de mayo y septiembre (especialmente en mayo) debido a que estos meses se encuentran cercanos a la temporada de invierno, y por ello pueden presentar temperaturas cercanas a cero o bajo cero.



**Figura 20:** Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de febrero.

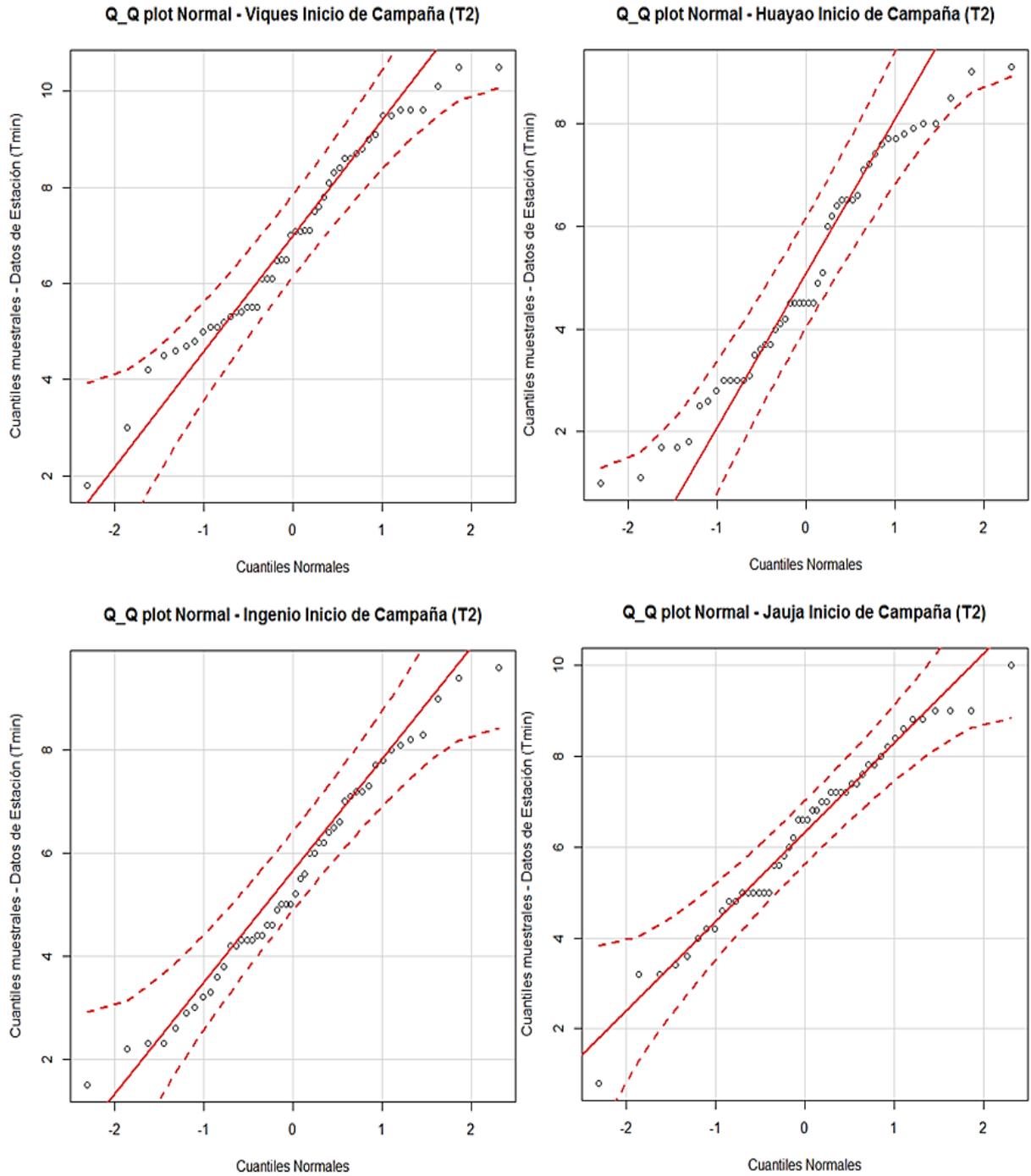
#### **4.1.2. ANÁLISIS DE NORMALIDAD Y LINEALIDAD DE LOS DATOS UTILIZADOS**

##### **a. Para Inicio de Campaña agrícola.**

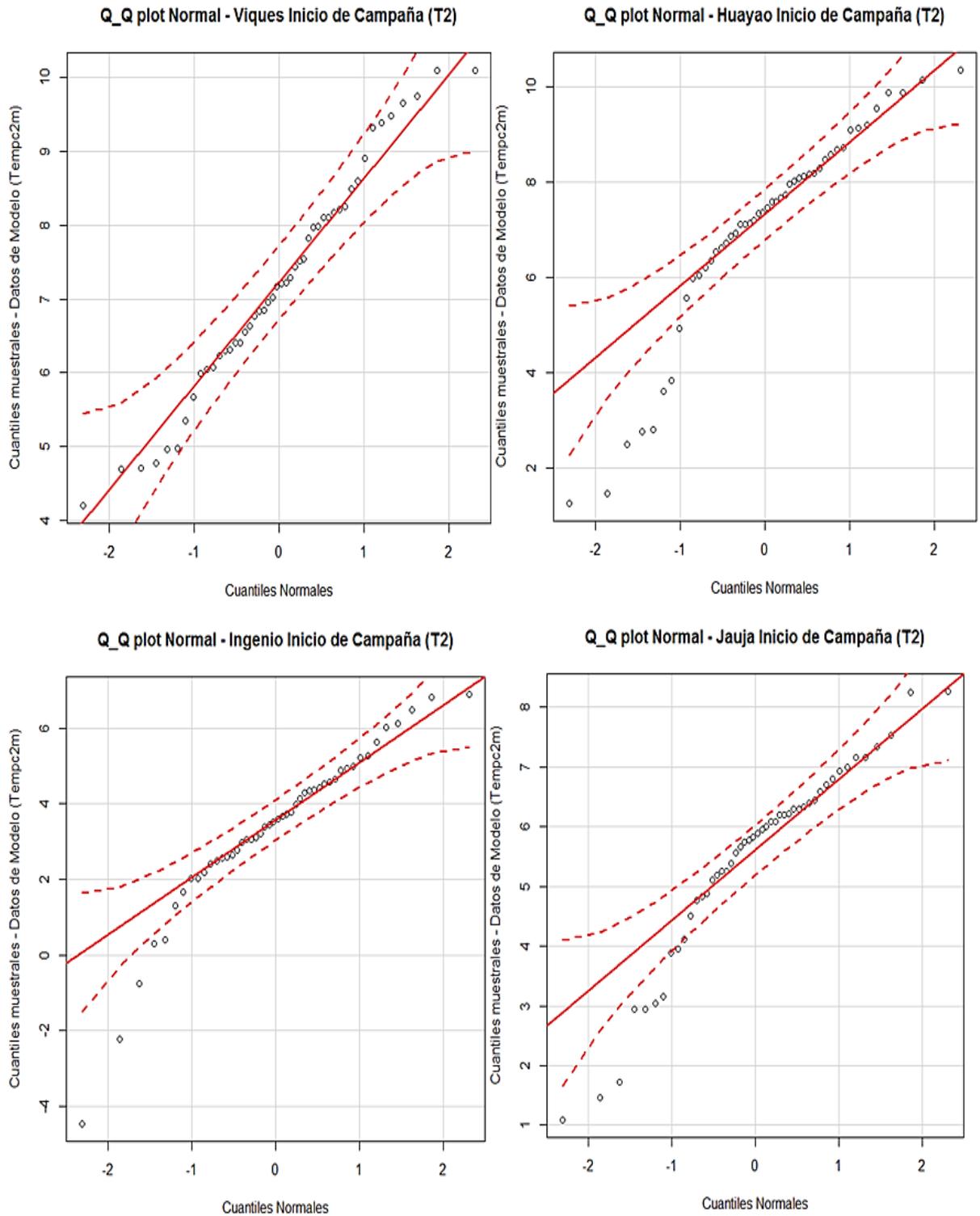
En este periodo los gráficos qqplots, mostraron una distribución más normal para los datos de temperatura mínima de las estaciones (Tmin) que para los datos de temperatura mínima pronosticada por BRAMS (Tempc2m o Tmin\_P). Teniendo en cuenta la temperatura mínima de las estaciones, la estación que presentó una distribución de datos más normal fue Ingenio, el cual presento esta característica en todos los tiempos de pronóstico (T2, T3, T4 y T5), así también, Viques fue la estación que presentó una distribución de datos menos normal en los tiempos de pronóstico (T3, T4 y T5). En cuanto a los datos de temperatura del BRAMS, las estaciones que presentaron una mejor distribución normal de sus datos fueron Ingenio en los tiempos de pronóstico (T3 y T5) y Viques en los tiempos de pronóstico (T2 y T4), además estas temperaturas mínimas presentaron en general algunos datos que se alejaron de la línea normal de su qqplot, casi en todas las estaciones y tiempos de pronóstico.

En el tiempo de pronóstico dos (T2) y para Tmin en inicio de campaña (ver Figura 21), se puede apreciar en las gráficas qqplots que las cuatro estaciones presentan datos con muy buena distribución normal, debido a que, todos los datos se ajustan a la línea normal y así mismo están dentro de los límites establecidos. Es sin duda Ingenio la estación que presenta un mejor ajuste. En los tiempos de pronóstico tres, cuatro y cinco (T3, T4 y T5) para Tmin en inicio de campaña (ver Anexo 3a): las estaciones de Ingenio y Huayao presentaron un mejor ajuste a la distribución normal en el tiempo tres y cuatro y en el tiempo cinco solamente Ingenio.

En el tiempo de pronóstico dos (T2) y para Tmin\_P en inicio de campaña (ver Figura 22), se puede apreciar en las gráficas qqplots que de las cuatro estaciones, Viques es la única que presenta datos con buena distribución normal (todos los datos se ajustan a la línea normal y así mismo están dentro de los límites establecidos), las otras estaciones presentan datos que se alejan de los límites de normalidad, lo cual nos sugiere la presencia de datos atípicos que influye en su normalidad. En los tiempos de pronóstico tres, cuatro y cinco (T3, T4 y T5) para Tmin\_P en inicio de campaña (ver Anexo 3a): la estacione de Ingenio presentó un mejor ajuste a la distribución normal en el tiempo tres, Viques y Huayao en el tiempo cuatro, e Ingenio y en el tiempo cinco.



**Figura 21:** Análisis de normalidad, de temperaturas mínimas de estación, para las cuatro estaciones de estudio en inicio de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2).



**Figura 22:** Análisis de normalidad, de temperaturas mínimas pronosticadas por BRAMS, para las cuatro estaciones de estudio en inicio de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2).

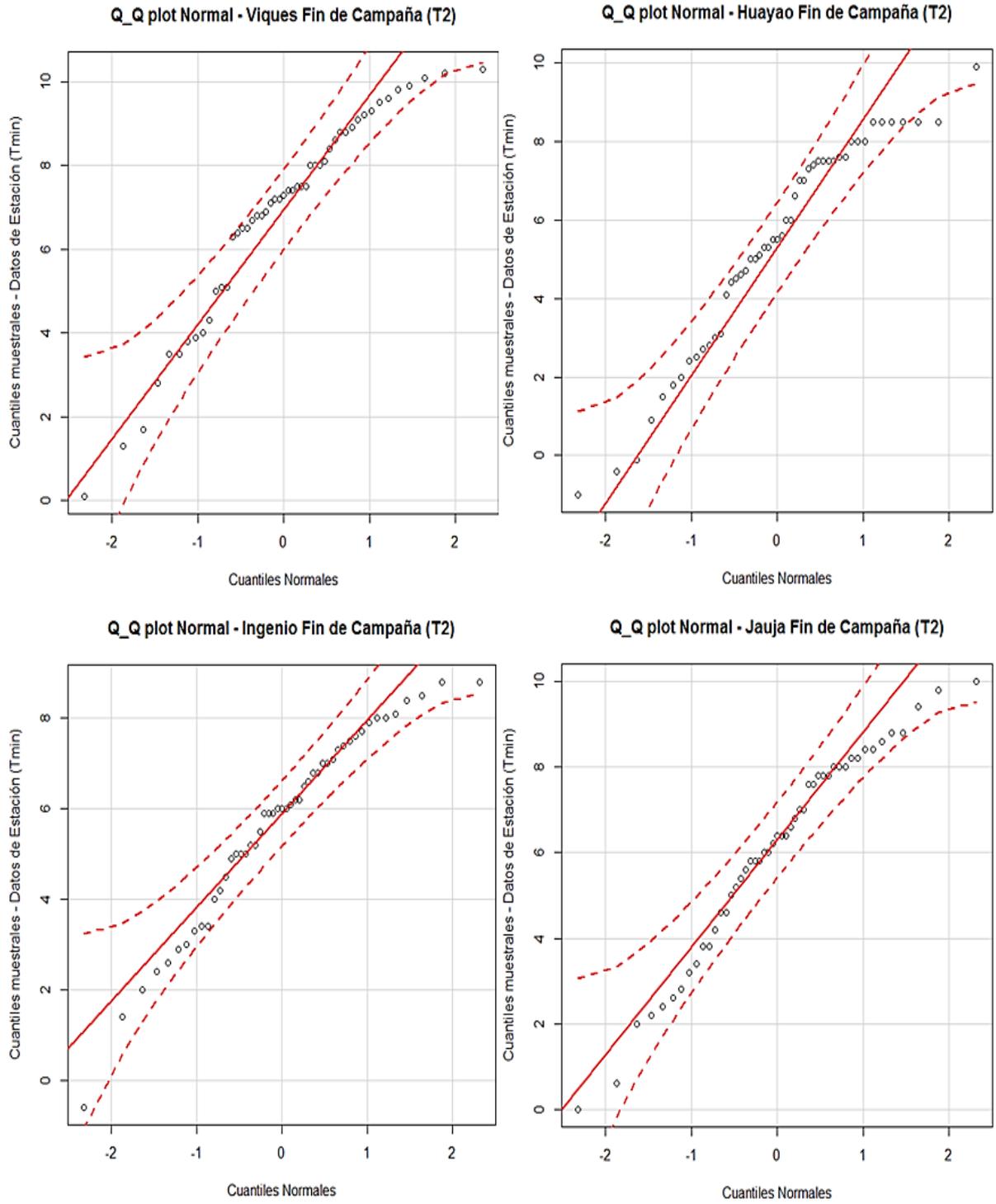
## **b. Para Fin de Campaña agrícola.**

En este periodo los gráficos qqplots de los datos de temperatura mínima de las estaciones (Tmin) también mostraron una distribución con mejor ajuste a la normalidad que los datos de temperatura mínima pronosticada por BRAMS (Tempc2m o Tmin\_P). Teniendo en cuenta la temperatura mínima de las estaciones, la estación que presentó una distribución de datos más normal fue Ingenio (a excepción del T3), Viques y Huayao fueron las estaciones que presentaron una distribución de datos menos normal en la mayoría de los tiempos de pronóstico. En cuanto a los datos de temperatura del BRAMS, Ingenio fue la única estación que presentó una distribución de sus datos con mejor ajuste a la normalidad (en el tiempo de pronóstico cuatro (T4)), en general se encontraron datos que se alejaron de la línea normal de su qqplot en todos los tiempos y para todas las estaciones (en mayor cantidad que en inicio de campaña y con un mayor alejamiento).

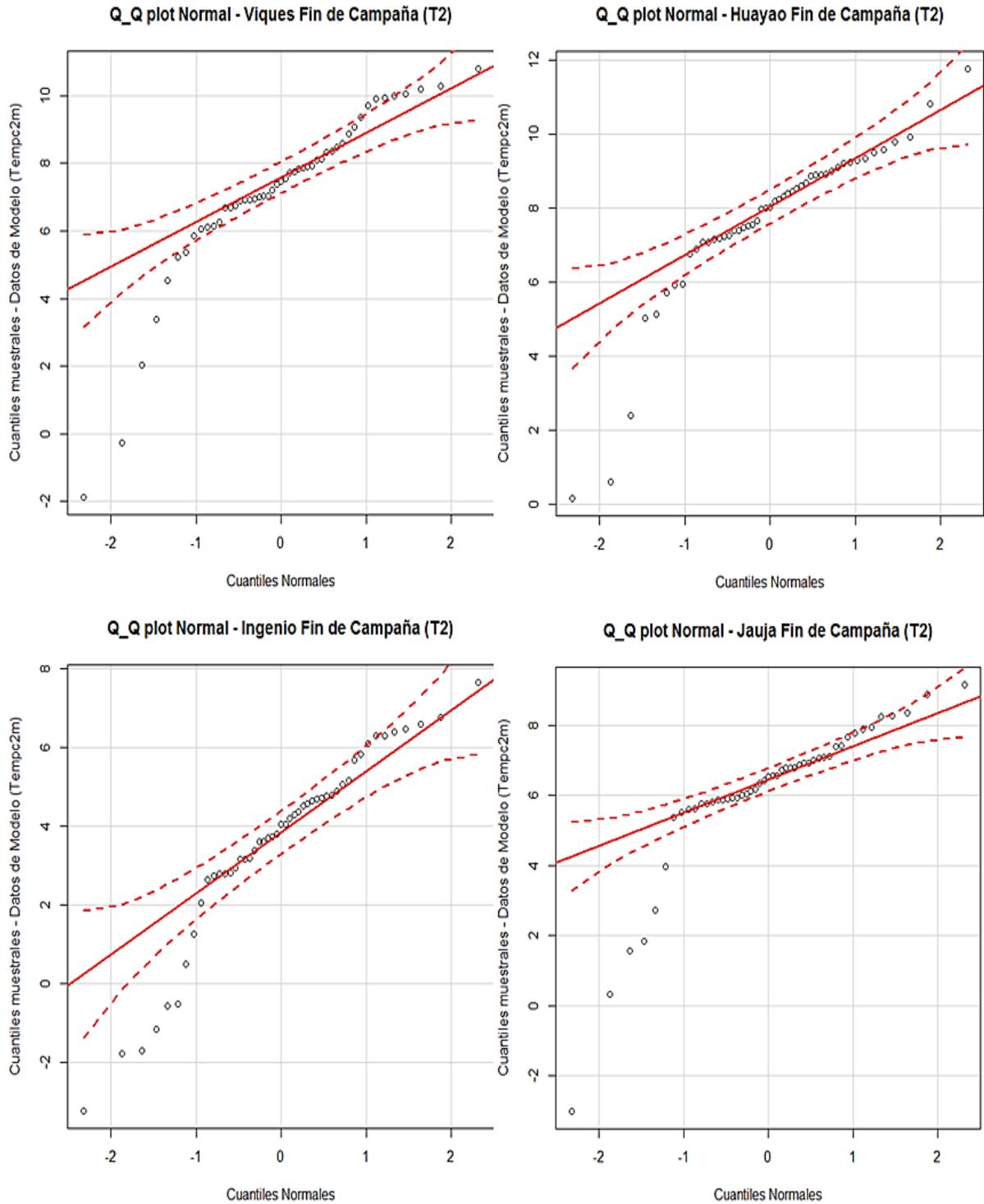
En el tiempo de pronóstico dos (T2) y para Tmin en fin de campaña (ver Figura 23), se puede apreciar en las gráficas qqplots que los datos de Viques, Jauja e Ingenio se ajustan a la recta normal y permanecen dentro del límite permitido; en los tiempos de pronóstico tres (T3), cuatro (T4) y cinco (T5) (ver Anexo 3b), la estación de Ingenio presentó un mejor ajuste a la distribución normal en el tiempo cuatro, en los tiempos tres y cinco todas las estaciones mostraron un menor ajuste.

En el tiempo de pronóstico dos (T2) y para Tmin\_P en fin de campaña (ver Figura 24), se puede apreciar en las gráficas qqplots que las cuatro estaciones presentan datos que se alejan de los límites de normalidad, lo cual nos sigue sugiriendo la existencia de datos atípicos, de igual modo en los tiempos de pronóstico tres (T3), cuatro (T4) y cinco (T5) (ver Anexo 3b).

En vista de la presencia de datos atípicos dentro de los datos del modelo BRAMS, estos se analizaron y se identificó que corresponden a temperaturas mínimas de estación cercanas a cero (especialmente las negativas), debido a que el modelo es sensible a pronosticar este tipo de datos y más aún en periodos en los cuales la mayoría de los datos son generalmente positivos ( el cual es el caso de nuestro periodo de estudio), es así que se determinó tomarlos en cuenta para la determinación de los modelos empíricos.

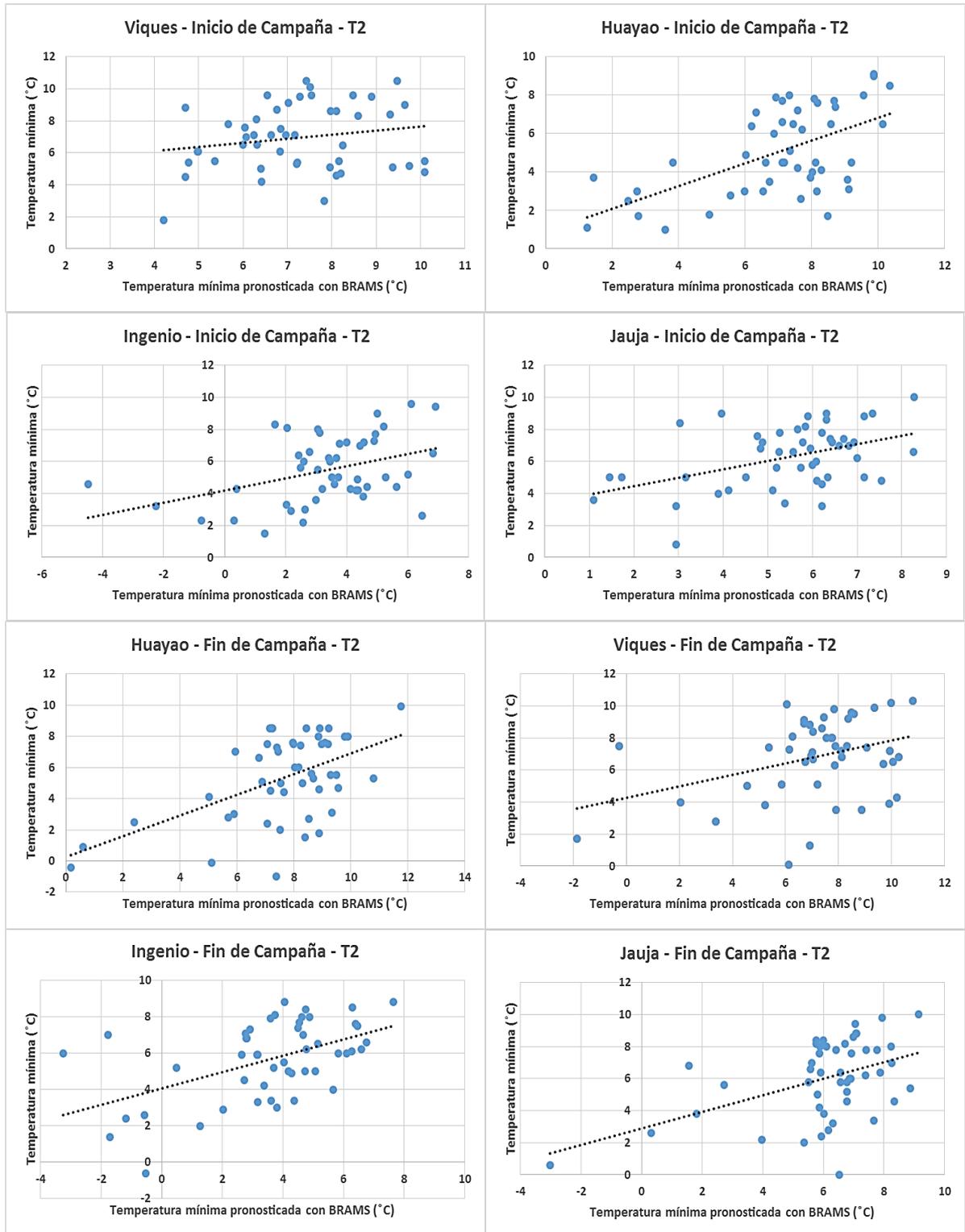


**Figura 23:** Análisis de normalidad, de temperaturas mínimas de estación, para las cuatro estaciones de estudio en fin de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2).

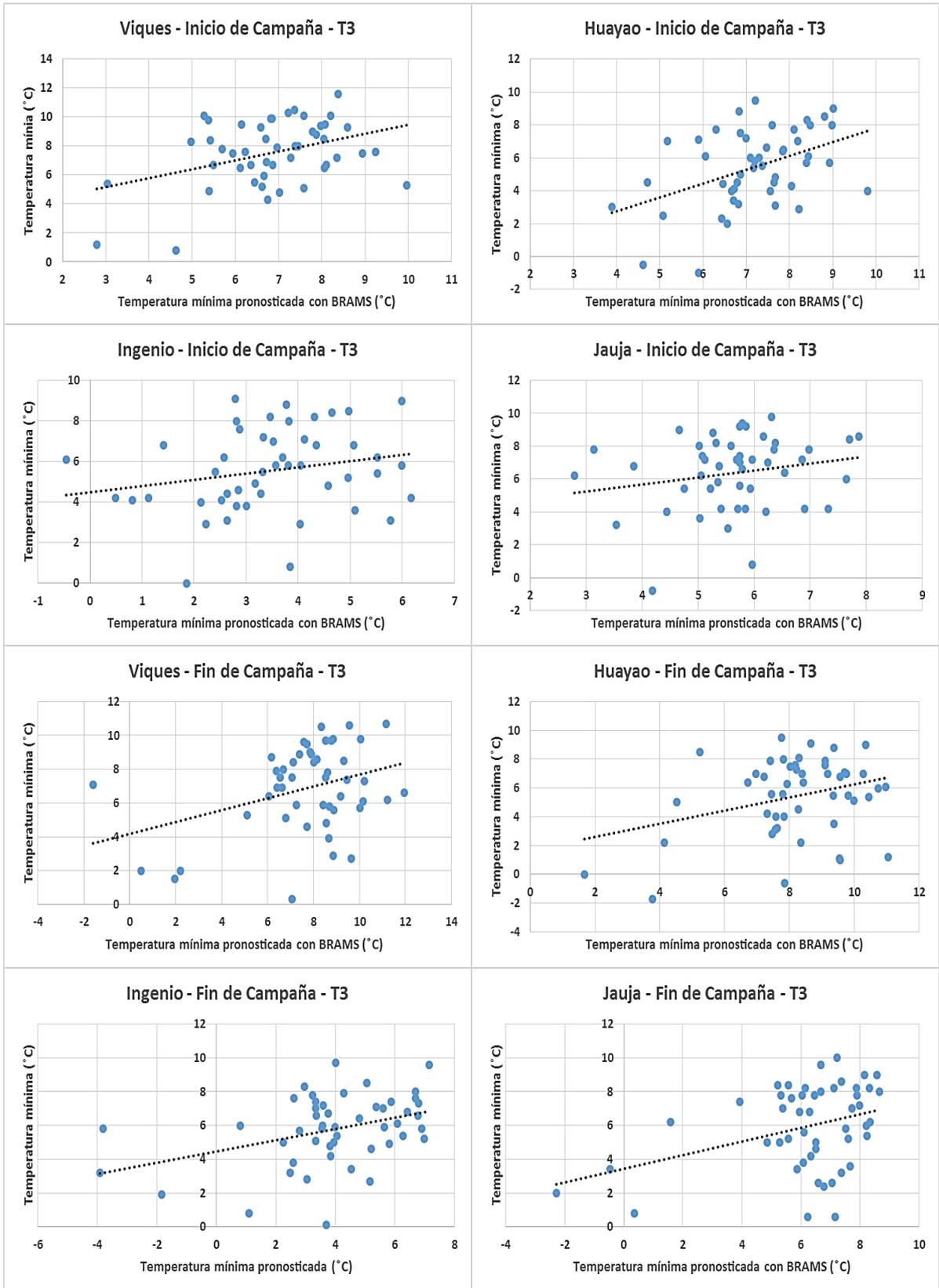


**Figura 24:** Análisis de normalidad, de temperaturas mínimas pronosticadas por BRAMS, para las cuatro estaciones de estudio en fin de campaña agrícola y correspondientes al tiempo dos de pronóstico (T2).

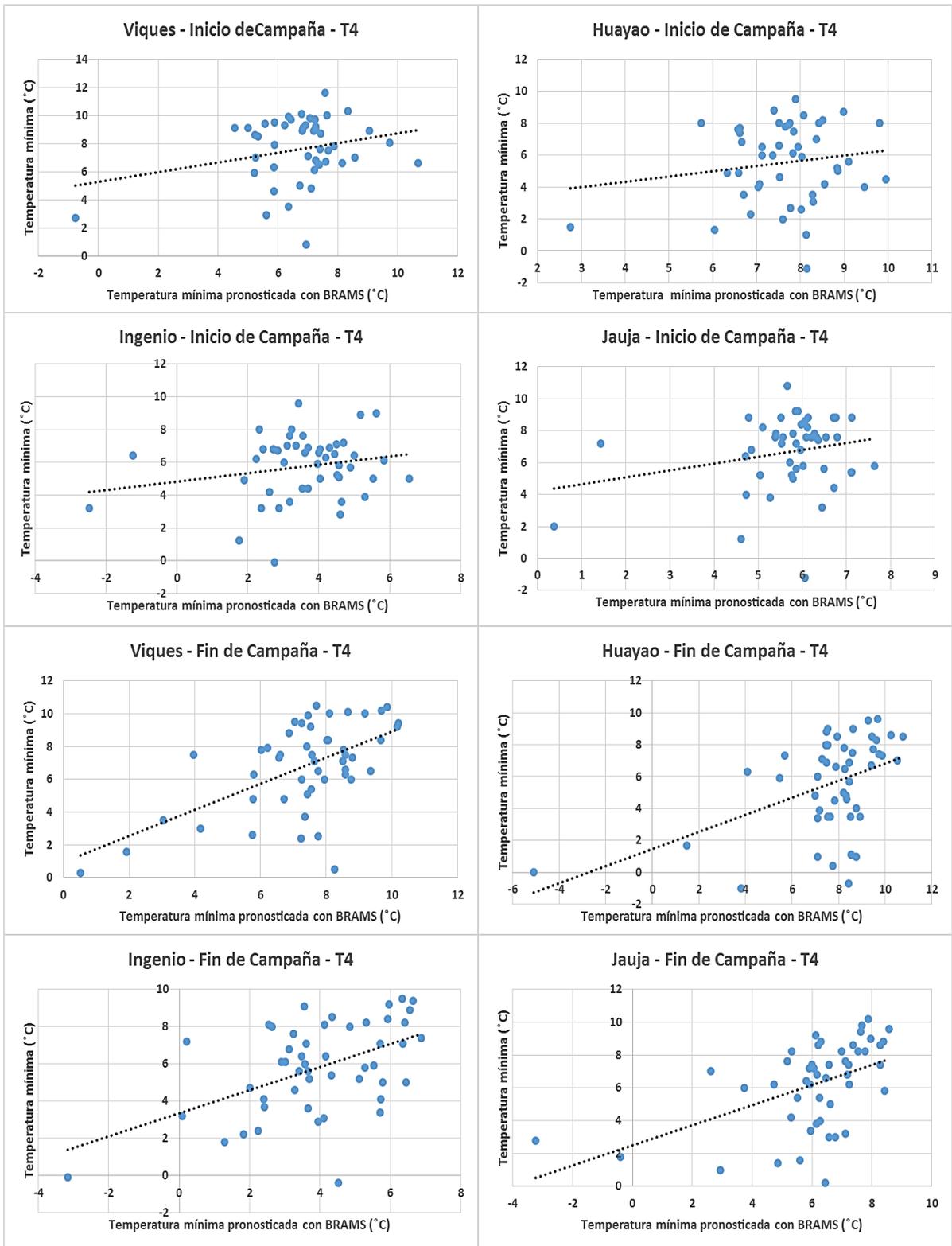
**c. Análisis de linealidad entre los datos que se tomaron en cuenta para la determinación de los modelos empíricos.**



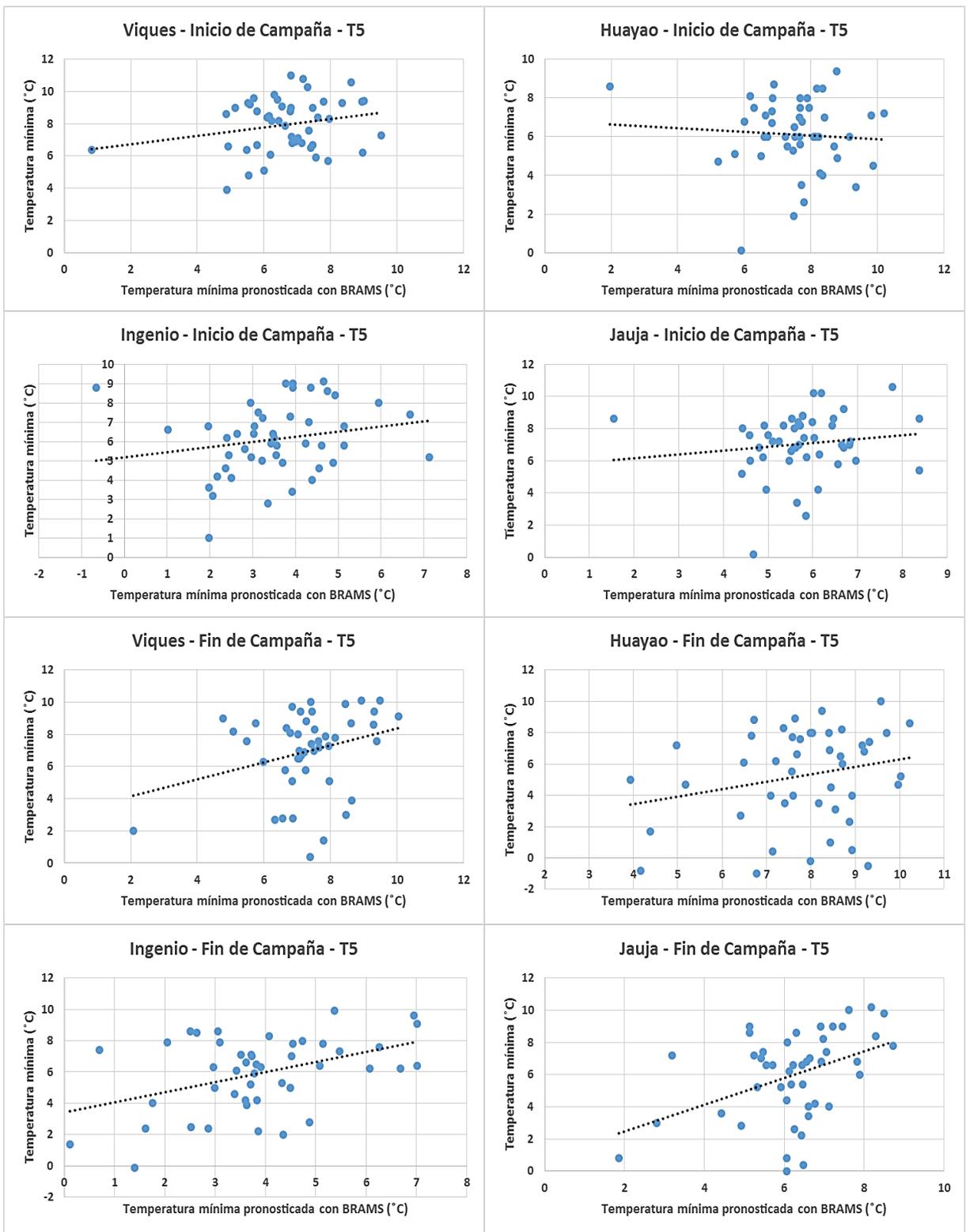
**Figura 25:** Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico dos (T2) y en todas las estaciones de estudio.



**Figura 26:** Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico tres (T3) y en todas las estaciones de estudio.



**Figura 27:** Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico cuatro (T4) y en todas las estaciones de estudio.



**Figura 28:** Gráficas de dispersión entre la temperatura mínima de estación y la temperatura mínima del BRAMS para inicio y fin de campaña agrícola del tiempo de pronóstico cinco (T5) y en todas las estaciones de estudio.

En las figuras 25, 26, 27 y 28, para las cuatro estaciones, en el T2 y T3 se evidenciaron relación lineal entre la variable independiente ( $T_{min\_P}$ ) y dependiente ( $T_{min}$ ) tanto para inicio y fin de campaña agrícola, en los tiempos T4 y T5, se percibe una mejor linealidad para fin de campaña.

#### 4.1.3. MODELOS EMPÍRICOS DE PRONÓSTICO

A continuación, se presentarán las ecuaciones de regresión lineal simple, entre la temperatura mínima pronosticada por el modelo BRAMS y la temperatura mínima de cada estación, en base al análisis de varianza ANOVA (ver Anexo 4) para cada uno de los tiempos de pronóstico (T2, T3, T4 y T5) y en las 4 estaciones de estudio, con un nivel de confianza del 95% ( $\alpha = 0.05$ ).

**Tabla 1:** Ecuaciones modelos empíricos para pronóstico de temperaturas mínimas en inicio de campaña agrícola para el tiempo de pronóstico dos (T2).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{min\_IC} = 5,4823 + 0,1745 * T_{min\_P}$	0,639	4,162	No significativo
HUAYAO	$T_{min\_IC} = 0,3377 + 0,6591 * T_{min\_P}$	16,062	4,162	Significativo
INGENIO	$T_{min\_IC} = 4,3070 + 0,3162 * T_{min\_P}$	4,408	4,162	Significativo
JAUJA	$T_{min\_IC} = 3,2004 + 0,5569 * T_{min\_P}$	7,352	4,162	Significativo

**Tabla 2:** Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico dos (T2).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{min\_FC} = 4,807 + 0,3328 * T_{min\_P}$	2,883	4,162	No significativo
HUAYAO	$T_{min\_FC} = 0,4078 + 0,6976 * T_{min\_P}$	8,533	4,111	Significativo
INGENIO	$T_{min\_FC} = 4,5131 + 0,3994 * T_{min\_P}$	8,690	4,111	Significativo
JAUJA	$T_{min\_FC} = 3,8119 + 0,38 * T_{min\_P}$	1,718	4,162	No significativo

**Tabla 3:** Ecuaciones de modelos empíricos para pronóstico de temperatura mínima en inicio de campaña agrícola para el tiempo de pronóstico tres (T3).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{\min\_IC} = 1,3944 + 0,9151 * T_{\min\_P}$	<b>12,020</b>	<b>4,162</b>	<b>Significativo</b>
HUAYAO	$T_{\min\_IC} = -2,4363 + 1,1231 * T_{\min\_P}$	<b>11,437</b>	<b>4,162</b>	<b>Significativo</b>
INGENIO	$T_{\min\_IC} = 4,6058 + 0,2891 * T_{\min\_P}$	<b>1,108</b>	<b>4,162</b>	No significativo
JAUJA	$T_{\min\_IC} = 4,5031 + 0,3762 * T_{\min\_P}$	<b>0,792</b>	<b>4,162</b>	No significativo

**Tabla 4:** Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico tres (T3).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{\min\_FC} = 5,6410 + 0,1822 * T_{\min\_P}$	<b>0,844</b>	<b>4,162</b>	No significativo
HUAYAO	$T_{\min\_FC} = 3,0209 + 0,3188 * T_{\min\_P}$	<b>0,877</b>	<b>4,162</b>	No significativo
INGENIO	$T_{\min\_FC} = 5,0965 + 0,1900 * T_{\min\_P}$	<b>1,508</b>	<b>4,162</b>	No significativo
JAUJA	$T_{\min\_FC} = 4,8120 + 0,1910 * T_{\min\_P}$	<b>0,549</b>	<b>4,162</b>	No significativo

**Tabla 5:** Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en inicio de campaña agrícola para el tiempo de pronóstico cuatro (T4).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{\min\_IC} = 4,4073 + 0,4451 * T_{\min\_P}$	<b>3,077</b>	<b>4,162</b>	No significativo
HUAYAO	$T_{\min\_IC} = 3,8120 + 0,1769 * T_{\min\_P}$	<b>0,241</b>	<b>4,162</b>	No significativo
INGENIO	$T_{\min\_IC} = 4,8703 + 0,1817 * T_{\min\_P}$	<b>0,769</b>	<b>4,162</b>	No significativo
JAUJA	$T_{\min\_IC} = 4,2197 + 0,3831 * T_{\min\_P}$	<b>1,508</b>	<b>4,162</b>	No significativo

**Tabla 6:** Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico cuatro (T4).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{\min\_FC} = 0,6512 + 0,8241 * T_{\min\_P}$	<b>8,974</b>	<b>4,162</b>	<b>Significativo</b>
HUAYAO	$T_{\min\_FC} = 1,3183 + 0,5078 * T_{\min\_P}$	<b>1,714</b>	<b>4,162</b>	No significativo
INGENIO	$T_{\min\_FC} = 3,4007 + 0,5875 * T_{\min\_P}$	<b>9,106</b>	<b>4,111</b>	<b>Significativo</b>
JAUJA	$T_{\min\_FC} = 1,8718 + 0,7077 * T_{\min\_P}$	<b>9,360</b>	<b>4,111</b>	<b>Significativo</b>

**Tabla 7:** Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en inicio de campaña agrícola para el tiempo de pronóstico cinco (T5).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{\min\_IC} = 5,1365 + 0,4263 * T_{\min\_P}$	<b>3,785</b>	<b>4,162</b>	No significativo
HUAYAO	$T_{\min\_IC} = 7,0417 - 0,1265 * T_{\min\_P}$	<b>0,247</b>	<b>4,162</b>	No significativo
INGENIO	$T_{\min\_IC} = 4,9679 + 0,3032 * T_{\min\_P}$	<b>1,028</b>	<b>4,162</b>	No significativo
JAUJA	$T_{\min\_IC} = 6,45348 + 0,06997 * T_{\min\_P}$	<b>0,040</b>	<b>4,162</b>	No significativo

**Tabla 8:** Ecuaciones de modelos empíricos para pronóstico de temperaturas mínimas en fin de campaña agrícola para el tiempo de pronóstico cinco (T5).

Estación	Modelo Empírico	Fc	Ft	SIGNIFICANCIA
VIQUES	$T_{\min\_FC} = 5,8514 + 0,1443 * T_{\min\_P}$	<b>0,090</b>	<b>4,162</b>	No significativo
HUAYAO	$T_{\min\_FC} = 2,7327 + 0,3203 * T_{\min\_P}$	<b>0,642</b>	<b>4,162</b>	No significativo
INGENIO	$T_{\min\_FC} = 3,6607 + 0,5372 * T_{\min\_P}$	<b>2,743</b>	<b>4,162</b>	No significativo
JAUJA	$T_{\min\_FC} = 2,7673 + 0,5119 * T_{\min\_P}$	<b>1,141</b>	<b>4,162</b>	No significativo

En las Tablas 01 y 02, se presentan las ecuaciones de regresión lineal para el tiempo de pronóstico dos (T2), es decir modelos con capacidad de pronóstico de dos días, en inicio y fin de campaña respectivamente. La prueba estadística Fisher resultó significativa al 95% de confianza ( $\alpha = 0,05$ ), para las estaciones de Huayao, Ingenio y Jauja en inicio de campaña agrícola y para Huayao e Ingenio en fin de campaña, esto quiere decir que las variables  $T_{\min\_P}$  (independiente) y  $T_{\min}$  (dependiente) cumplen con la asociación lineal, es decir, se cumplió el rechazo de la hipótesis nula ( $H_0$ ), lo cual significa que las pendientes e interceptos de las ecuaciones de regresión lineal son diferentes de cero.

En las Tablas 03 y 04, se presentan las ecuaciones de regresión lineal para el tiempo de pronóstico tres (T3), es decir modelos con capacidad de pronóstico de tres días, en inicio y fin de campaña respectivamente. La prueba estadística Fisher resultó significativa al 95% de confianza ( $\alpha = 0,05$ ), para las estaciones de Viques y Huayao en inicio de campaña agrícola.

En las Tablas 05 y 06, se presentan las ecuaciones de regresión lineal para el tiempo de pronóstico cuatro (T4), es decir modelos con capacidad de pronóstico de cuatro días, en inicio y fin de campaña respectivamente. La prueba estadística Fisher resultó significativa

al 95% de confianza ( $\alpha = 0,05$ ), para las estaciones de Viques, Ingenio y Jauja en fin de campaña agrícola.

En la tabla 07 y 08, se presentan las ecuaciones de regresión lineal para el tiempo de pronóstico cinco (T5), es decir modelos con capacidad de pronóstico de cinco días, en inicio y fin de campaña respectivamente. La prueba estadística Fisher no resultó significativa al 95% de confianza ( $\alpha = 0,05$ ), para ninguna de las estaciones en ninguna de las campañas agrícolas.

#### 4.2.3. VALIDACIÓN DEL MODELO EMPÍRICO Y ANÁLISIS DE RESIDUOS

**Tabla 9:** Resultados de la validación de los modelos empíricos para los tiempos de pronóstico (T2, T3 y T4) en ambas campañas agrícolas.

Estación	RMSE (°C)	Bias (°C)	Coefficiente de determinación	Coefficiente de correlación	Valor crítico de Pearson ( $\alpha = 0,05$ )
Huayao_IC_T2	2,088	-0,250	0,325	0,570	0,514
Ingenio_IC_T2	2,172	-0,288	0,184	<b>0,429</b>	0,514
Jauja_IC_T2	1,829	-0,156	0,199	<b>0,446</b>	0,514
Huayao_FC_T2	2,689	1,900	0,485	0,696	0,632
Ingenio_FC_T2	1,595	1,284	0,713	0,844	0,632
Viques_IC_T3	2,347	0,408	0,012	<b>0,108</b>	0,514
Huayao_IC_T3	2,354	0,679	0,138	<b>0,372</b>	0,514
Viques_FC_T4	1,565	-0,287	0,609	0,780	0,497
Ingenio_FC_T4	1,945	-0,379	0,462	0,680	0,632
Jauja_FC_T4	2,578	-0,225	0,363	<b>0,602</b>	0,632

En la Tabla 9, se muestran los estadísticos de validación de los modelos empíricos de pronóstico de las temperaturas mínimas pronosticadas con dos (T2), tres (T3) y cuatro (T4) días de anticipación, tanto para inicio y fin de campaña agrícola. A través del RMSE se verifica que las temperaturas mínimas pronosticadas con dos, tres y cuatro días de anticipación varían entre 1,5 a 2,7°C (en todas las estaciones), con respecto a los valores de temperaturas mínimas registradas en las estaciones meteorológicas, tanto para inicio y fin de las campañas agrícolas entre los años 2012 a 2016. Según la prueba estadística Bias, se observa que los modelos de pronóstico de dos días de anticipación subestiman las temperaturas mínimas en las estaciones de Huayao, Ingenio y Jauja, para inicio de campaña agrícola, y sobrestiman en la estación de Huayao para fin de campaña agrícola; los modelo

de pronóstico de tres días de anticipación sobreestiman las temperaturas mínimas tanto en Viques y Huayao en inicio de campaña agrícola; y los modelos de pronóstico con cuatro días de anticipación subestiman las temperaturas mínimas en las estaciones de Viques, Ingenio y Jauja en fin de campaña agrícola. La prueba estadística de Pearson resultó significativa con un nivel de confianza de 95% para los modelos de pronóstico de dos días de anticipación en Huayao (inicio y fin de campaña agrícola) e Ingenio (fin de campaña agrícola), y para los modelos de cuatro días de anticipación en Viques e Ingenio (fin de campaña agrícola). Es así que, respecto a los estadísticos de validación las estaciones mencionadas anteriormente son las que pasan la validación.

### **ANÁLISIS DE RESIDUOS:**

Este análisis se realizó para verificar los supuestos de linealidad, homocedasticidad, normalidad e independencia, que se debe de cumplir para que la ecuación lineal simple tenga capacidad de predicción válida. En este sentido solo se realizó este análisis en las ecuaciones que resultaron significativas a la prueba estadística Fisher y a los que pasaron la validación. Se realizó el análisis para un nivel de significancia de ( $\alpha = 0,05$ ).

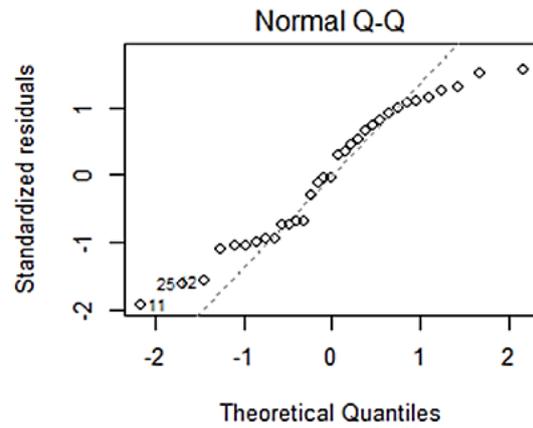
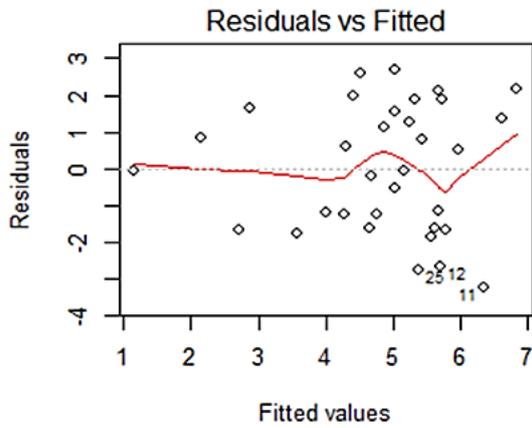
#### **a. Análisis gráfico**

En las Figuras 29 y 30 no se puede determinar exactamente si los residuos cumplen con los supuestos que se indicó, es evidente que los datos no cumplen con la gráfica ideal de residuos vs valores pronosticados (ver Figura 13), es por ello, que se realizó el análisis cuantitativo o analítico.

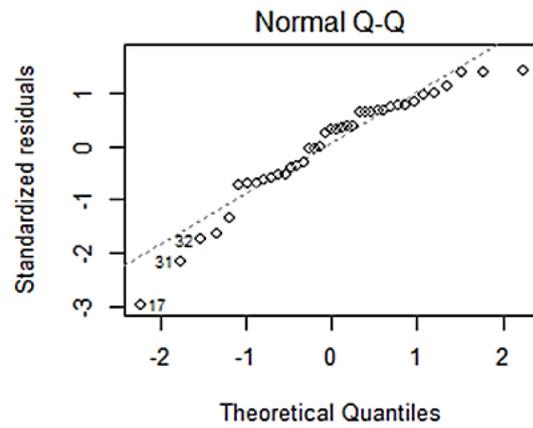
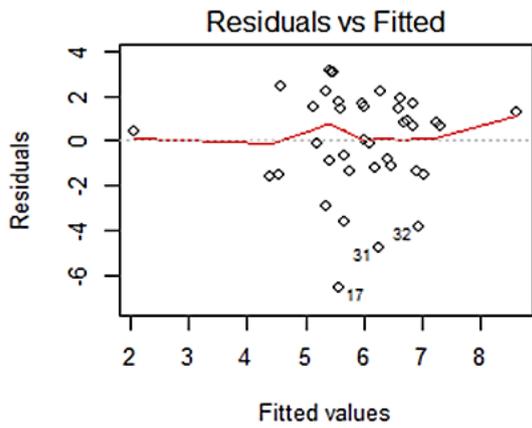
#### **b. Análisis cuantitativo**

En la Tabla 10, presenta los resultados de los tests estadísticos para verificar las hipótesis de los supuestos que se indicó. Se encontró que, en fin de campaña agrícola, tanto Huayao para el tiempo de pronóstico dos (T2) y Viques para el tiempo de pronóstico cuatro (T4), no pasó la prueba estadística de normalidad, debido a que su p-valor obtenido fue menor al nivel de significancia ( $\alpha = 0,05$ ); en cuanto a la homocedasticidad la estación de Ingenio para el tiempo de pronóstico dos (T2) en fin de campaña agrícola, no pasó esta prueba, debido a que su p-valor obtenido también fue menor que el nivel de significancia; por último todas las estaciones pasaron la prueba de linealidad.

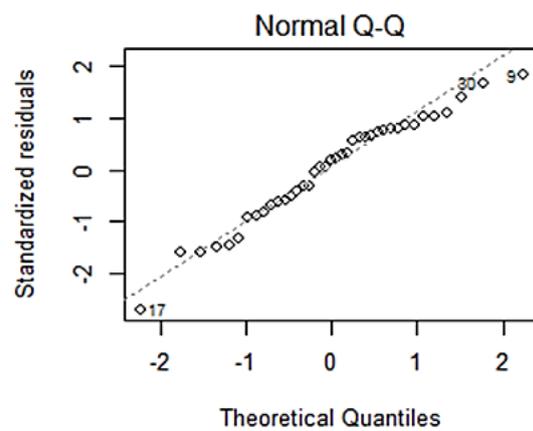
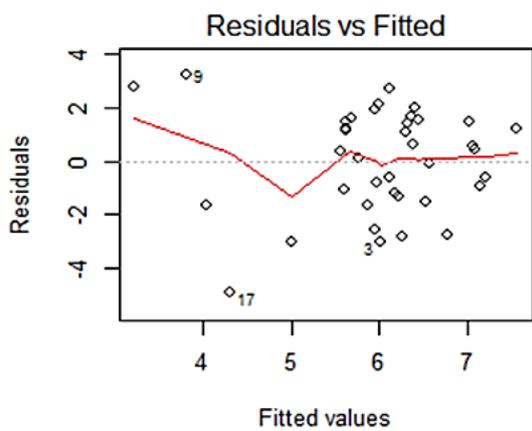
$\text{lm}(\text{Huayao\_IC\_T2\_Tmin} \sim \text{Huayao\_IC\_T2\_TminP})$



$\text{lm}(\text{Huayao\_FC\_T2\_Tmin} \sim \text{Huayao\_FC\_T2\_TminP})$

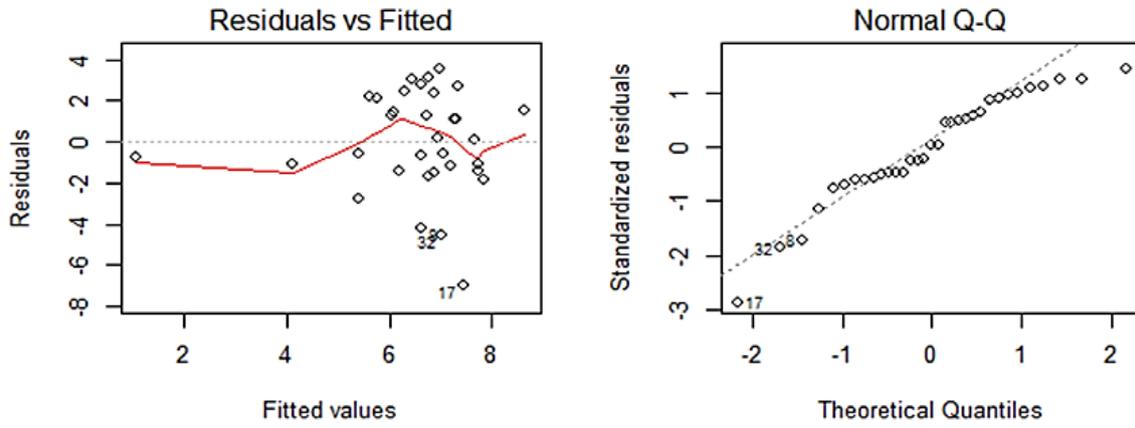


$\text{lm}(\text{Ingenio\_FC\_T2\_Tmin} \sim \text{Ingenio\_FC\_T2\_TminP})$

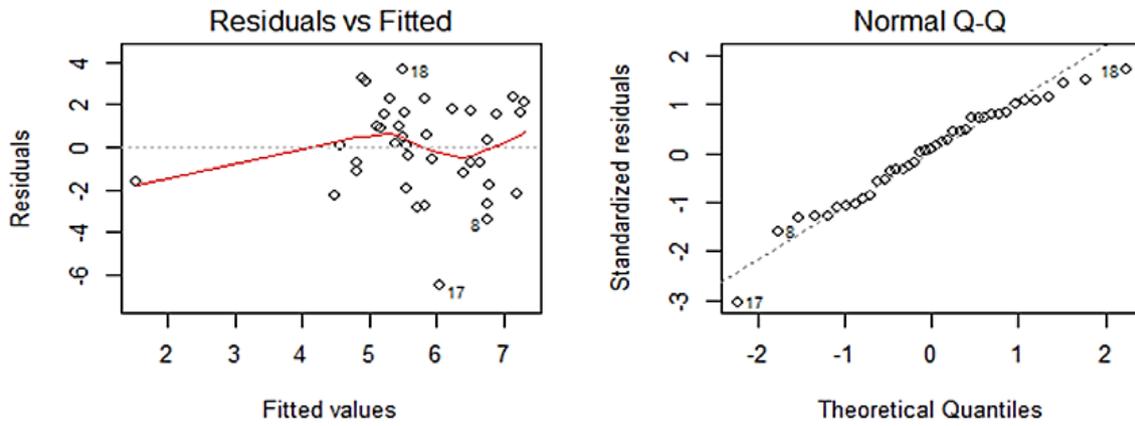


**Figura 29:** Análisis de residuos para el tiempo dos de pronóstico (T2), para las estaciones que pasaron satisfactoriamente la prueba Fisher y la validación.

$\text{lm}(\text{Viques\_FC\_T4\_Tmin} \sim \text{Viques\_FC\_T4\_TminP})$



$\text{lm}(\text{Ingenio\_FC\_T4\_Tmin} \sim \text{Ingenio\_FC\_T4\_TminP})$



**Figura 30:** Análisis de residuos para el tiempo cuatro de pronóstico (T4), para las estaciones que pasaron satisfactoriamente la prueba Fisher y la validación.

**Tabla 10:** Resultados del análisis cuantitativo de los residuos para los tiempos de pronóstico (T2 y T4) en ambas campañas agrícolas.

( $\alpha = 0,05$ )	Normalidad	Homocedasticidad	Autocorrelación	Linealidad
	Shapiro-Wilk test	Breusch-Pagan test	Durbin-Watson test	RESET test
Huayao_IC_T2	0,076	0,056	0,832	0,918
Huayao_FC_T2	<b>0,018</b>	0,662	0,857	0,823
Ingenio_FC_T2	0,220	<b>0,002</b>	0,596	0,121
Viques_FC_T4	<b>0,027</b>	0,382	0,726	0,784
Ingenio_FC_T4	0,248	0,674	0,805	0,291

#### 4.2. DETERMINACIÓN DE LA DISTRIBUCIÓN ESPACIAL

Debido a que solo dos estaciones pasaron la validación y la prueba de residuos, y que estas corresponden a diferentes tiempos de pronóstico y a la vez a diferentes campañas, no se dio la posibilidad de cumplir con este objetivo.

## V. CONCLUSIONES

- 1) Los modelos empíricos de pronóstico de temperaturas mínimas, de dos días de anticipación para la estación de Huayao en inicio de campaña agrícola y de cuatro días de anticipación para la estación de Ingenio en fin de campaña agrícola, se ajustaron a ecuaciones de regresión lineal simple y resultaron estadísticamente significativos, pasando la etapa de validación y la de análisis de residuos, por tanto pueden ser aplicados como modelos operativos. Sin embargo, los modelos para la mayoría de las estaciones, tanto para inicio y fin de campaña agrícola, en los tiempos de pronóstico dos, tres y cuatro, no pasaron todas las pruebas estadísticas respectivas. Así también los modelos para el tiempo cinco no presentaron un buen ajuste al modelo de regresión lineal simple.

## VI. RECOMENDACIONES

En base a los resultados obtenidos en el presente trabajo de investigación se hacen las siguientes recomendaciones:

- Se recomienda volver a realizar el estudio para el valle del Mantaro, pero con una mayor cantidad de datos, especialmente para los tiempos de pronóstico dos, tres y cuatro. Dado que se obtuvieron resultados positivos, sería interesante realizar un estudio similar en otra zona de interés.
- Para los modelos empíricos que pasaron la etapa de validación (aun así no hayan pasado satisfactoriamente el análisis de residuos) se recomienda, evaluar su capacidad de pronóstico, dado que el análisis de residuos se realizó teniendo en cuenta las condiciones estadísticas necesarias para determinar modelos con capacidad predictiva confiable, sin embargo, los estadísticos usados en la validación también nos indican que los modelos empíricos determinados poseen capacidad predictiva, por ello la necesidad de evaluar sus alcances y limitaciones.
- Se podrían determinar modelos empíricos de regresión lineal múltiple teniendo en cuenta las temperaturas pronosticadas del BRAMS, a nivel de caseta meteorológica (2 metros) y las del primer nivel z (aproximadamente 28 metros), debido a que, las temperaturas del BRAMS que se tomaron en cuenta para la realización de la tesis (Tempc2m) tienen la característica de exagerar los valores muy bajos y altos, a diferencia de la otra temperatura (Tempc) el cual presenta una variación más homogénea, usar estas dos temperaturas quizá podría generar modelos con un mejor ajuste a los datos de estación.
- Se debe de tener cuidado al momento de configurar el RAMSIN del modelo BRAMS, especialmente al ubicar la zona de estudio y realizar el anidamiento de sus tres dominios, debido a que, existe la posibilidad de crear un anidamiento inestable, en ese caso la ejecución del modelo presentará fallas, especialmente, el truncamiento de esta en un tiempo anterior al establecido.

- El modelo BRAMS posee la capacidad de realizar pronósticos con muy buena resolución (hasta 2 km), sería interesante hacer uso de este recurso para poder estudiar otras variables meteorológicas tales como la precipitación, viento, etc.
- El uso del modelo BRAMS no es muy complejo, sin embargo, si se quiere realizar una investigación por un tiempo muy largo es recomendable tener una computadora con buena capacidad, si es más de uno mejor, ya que, debido a su gran capacidad de resolución demanda un mayor tiempo de ejecución.
- Existe la posibilidad de usar el modelo BRAMS tanto en serie como en paralelo, la segunda forma de instalación es la más recomendada cuando se quiere trabajar con muy buena resolución, debido a que este tipo de instalación permite al modelo usar todos los núcleos de una computadora y es así como reduce el tiempo de ejecución considerablemente.
- Debido a las características del modelo BRAMS (especialmente su portabilidad e instalación), se hace factible su uso como herramienta de aprendizaje sobre modelos a nivel universitario.

## VII. REFERENCIAS BIBLIOGRÁFICAS

- Behar Gutiérrez, Roberto. 2008. Introducción al Análisis Exploratorio de Datos. Edición Intrauniversitaria, Universidad del Valle. Versión 2.0, Cali, Colombia, 2009.
- Carlos Gómez, G. 2010. Monografía “Las heladas en el Perú y el mundo”. Huancayo, PE. (42p). Consultado noviembre 2015. [http://www.academia.edu/1770180/Las\\_heladas\\_en\\_el\\_Per%C3%BA\\_y\\_el\\_mundo](http://www.academia.edu/1770180/Las_heladas_en_el_Per%C3%BA_y_el_mundo)
- COMET. Intelligent Use of Model-Derived Products -version 2. Agosto 2017. [http://www.meted.ucar.edu/nwp/model\\_derivedproducts/](http://www.meted.ucar.edu/nwp/model_derivedproducts/)
- CPTEC. 2007. First Time User’s Guide (BRAMS Version 4.0). Version 5. Consultado Junio 2016. [http://downloads.cptec.inpe.br/~rdown/upload/Brams\\_First\\_Time\\_User\\_Guide\\_Ver4.pdf](http://downloads.cptec.inpe.br/~rdown/upload/Brams_First_Time_User_Guide_Ver4.pdf)
- Ferre Jaén, María Elvira. 2014. Modelos de regresión. Curso 2014-15. Consultado Julio 2017. <http://www.um.es/ae/FEIR/40/>
- Ferreira Pereira, Simone. 2011. Estudio comparativo do pós-processamento estatístico aplicado ao modelo BRAMS. Disertación (Maestría en Sensoramento Remoto). Porto Alegre – Brasil, Universidade Federal do Rio Grande do Sul. 5-12, 132p.
- García J, Villegas E. Pronóstico de temperaturas mínimas para previsión de heladas con 18 horas de anticipación. Anales Científicos de la Universidad Nacional Agraria La Molina. [Aprobado para su publicación]. Ediagraria, Lima, 2006.
- García Skabar, Yanina. 2007. Análisis enriquecidos, evaluación de su impacto en el pronóstico y su aplicación para describir circulaciones de mesoescala en el centro y norte de Argentina. Tesis de doctorado en Ciencias de la Atmósfera y los Océanos. Buenos Aires - Universidad de Buenos Aires. 150-154, 177p.
- Guerrero Higuera, Á. M., García Ortega, E., Matellán Olivera, V., & Sánchez Gómez, J. L. (2013). Procesamiento paralelo de los pronósticos meteorológicos del modelo WRF mediante NCL.

- IGP (Instituto Geofísico del Perú). 2005. Diagnóstico de la cuenca del Mantaro bajo la visión de Cambio Climático. CONAM, Instituto Geofísico del Perú. 94 p.
- IGP (Instituto Geofísico del Perú). 2010. Memoria del Subproyecto “Pronóstico estacional de lluvias y temperaturas en la cuenca del río Mantaro para su aplicación en la agricultura”. Yamina Silva Vidal. Lima, PE. Biblioteca Nacional del Perú N° 2010-08456. p.13-14, 42, 82-99.
- INEI (Instituto Nacional de Estadística e Informática). 2017. Perú: Panorama económico departamental. Informe Técnico N° 07 – Julio 2017.
- Jiménez García, Mauricio. 2014. Validación de la capacidad del modelo WRF “Weather Research and Forecasting” para pronosticar lluvia intensa, usando el método orientado a objetos y tablas de contingencia. Tesis Magister en Ciencias Meteorológica. Bogotá – Colombia, Universidad Nacional de Colombia. 20-23 y 31-33, 144p.
- LMEIDA, E. S. 2007. Climatología de mesoescala em grade computacional. Tesis (Doctorado en Computación Aplicada) - Instituto Nacional de Investigaciones Espaciales, São José dos Campos.
- Norabuena E., A. 2014. Modelamiento empírico para pronóstico de temperaturas mínimas en períodos de inicio y fin de campaña agrícola, Universidad Nacional Agraria la Molina. Ingeniera meteoróloga.
- Orellana, Liliana. 2008. Regresión lineal simple. Consultado Julio 2017. 38-41, 49p. [http://www.dm.uba.ar/materias/estadistica\\_Q/2011/1/clase%20regresion%20simple.pdf](http://www.dm.uba.ar/materias/estadistica_Q/2011/1/clase%20regresion%20simple.pdf)
- Pielke, R. A. 2002. Mesoscale Meteorological Modeling (Academic Press, New York, 1984); 2nd ed.
- Pielke, R. A.; Cotton, W. R.; Walko, R. L.; Trembaek, C. J.; Lyons, W. A.; Grasso, L. D.; Nieholls, M. E.; Moran, M. D.; Wesley, D. A.; Lee, T. J.; Copeland, J. H. 1992. A comprehensive meteorological modeling system - rams. Meteorology and atmospheric physics. Meteorology and Atmospheric Physics, v. 49, p. 69–91.
- Regresión lineal simple. Capítulo 9. Consultado Agosto 2017. Disponible en: <<http://www4.ujaen.es/~dmontoro/Metodos/Tema%209.pdf>>
- Salvador Figueras, M; Gargallo, P. (2003). "Análisis Exploratorio de Datos". Consultado Julio 2017. 5, 26-29 y 53-54, 68p. Disponible en: 5campus.com, Estadística <<http://www.5campus.com/leccion/aed>>

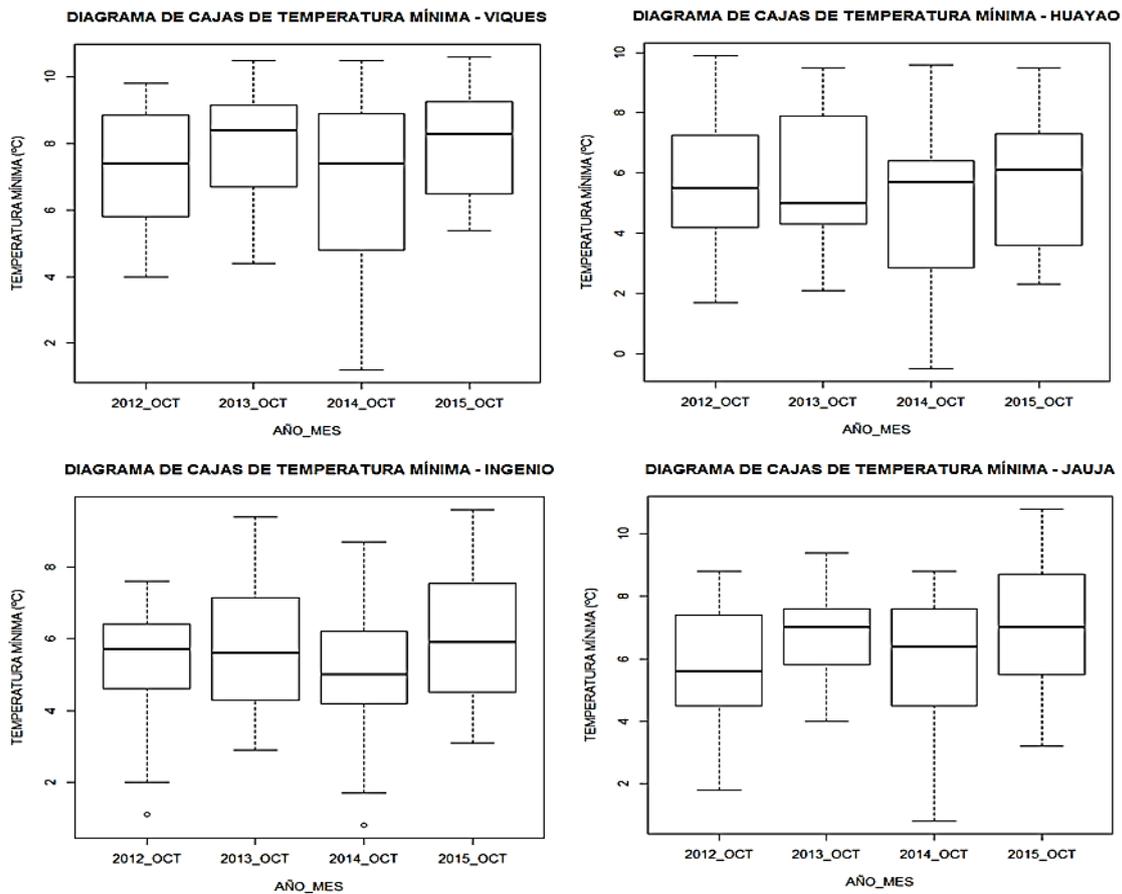
- Sanábria, J. 2005, Calibración y Validación de Modelos de Pronóstico de Heladas en el Valle del Mantaro, Univesidad Nacional Agraria La Molina. Tesis de Maestría.
- SENAMHI (Servicio Nacional de Meteorología e Hidrología, PE). Aprendiendo/Meteorología/Guía básica de Meteorología General/Capítulo 13 (en línea). Consultado 05 nov. 2015. [file:///C:/Users/pc1/Downloads/cap13%20\(1\).pdf](file:///C:/Users/pc1/Downloads/cap13%20(1).pdf)
- Souza, E. P. 1999. Estudo Teórico e Numérico da Relação entre Convecção e Superfícies Heterogêneas na Região Amazônica. Tesis de Doctorado, São Paulo – SP, DCA/IAG/USP.
- Trasmonte, G.; Chavez, R.; Segura, B. and Rosales, J. L. 2008, Frost risks in the Mantaro river basin, *Advances in Geosciences* 14, 265-27.
- Treichrieb, C. A. 2008. Sensibilidade do BRAMS para previsão de chuva e temperatura, no nordeste do Rio Grande do Sul, para diferentes resoluções espaciais. Disertación (Maestría) - Universidad Federal de Santa María.
- Tremback, C. J. 1990. Numerical simulation of a mesoscale convective complex model development and numerical results. Tesis (Doctorado) - Department of Atmospheric Science, Colorado State University.
- Vega L., G. 2005. Diagnóstico de la Cadena Productiva de la Alcachofa. Dirección Regional Agraria Junín, 20 pp.
- Villegas, E. 1991. Zonificación del Valle del Mantaro según la Intensidad y Riesgo de Ocurrencia de Heladas Radiacionales, Tesis para optar el título de Ingeniero Meteorólogo, Universidad Nacional Agraria La Molina, 134 p.
- Walko, R.; Tremback, J. 2001. RAMS The Regional Atmospheric Modeling System. [S.l.]. Draft.
- Walko, R.; Tremback, J. 2005. Regional atmospheric modeling system. Version 6, User's Guide - Introduction. [S.l.].
- WMO. 2011. Guide to climatological practices. WMO-No 100. [http://www.wmo.int/pages/prog/wcp/ccl/guide/documents/wmo\\_100\\_es.pdf](http://www.wmo.int/pages/prog/wcp/ccl/guide/documents/wmo_100_es.pdf)

## VIII. ANEXO

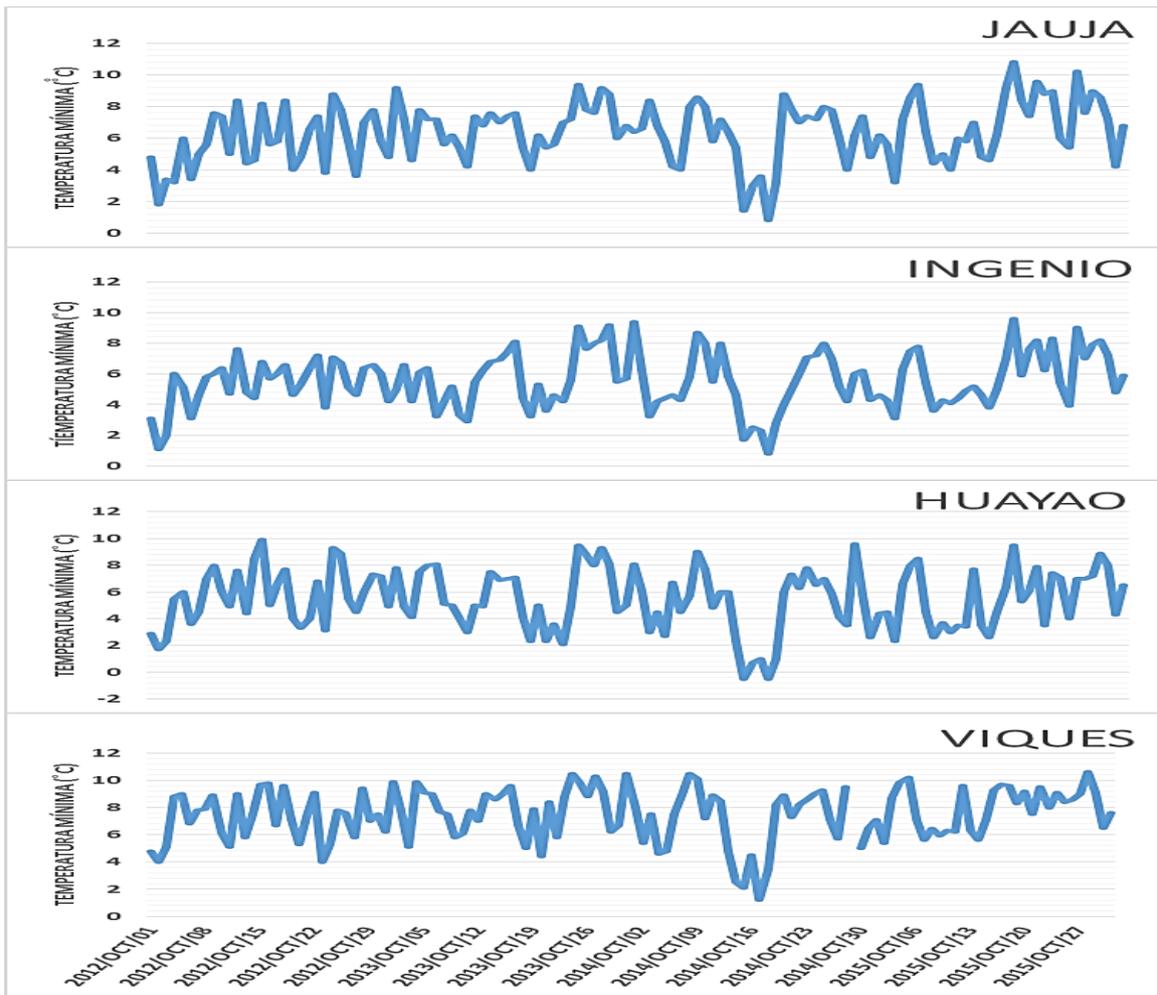
### ANEXO 1. UBICACIÓN DE LAS ESTACIONES METEOROLÓGICAS QUE SE TOMARON EN CUENTA EN LA INVESTIGACIÓN.

Estación Meteorológica	Latitud ( S°)	Longitud ( W°)	Altitud (m)
Viques	12,16 °	75,23 °	3186
Huayao	12,03 °	75,34 °	3360
Ingenio	11,88 °	75,29 °	3422
Jauja	11,78 °	75,48 °	3360

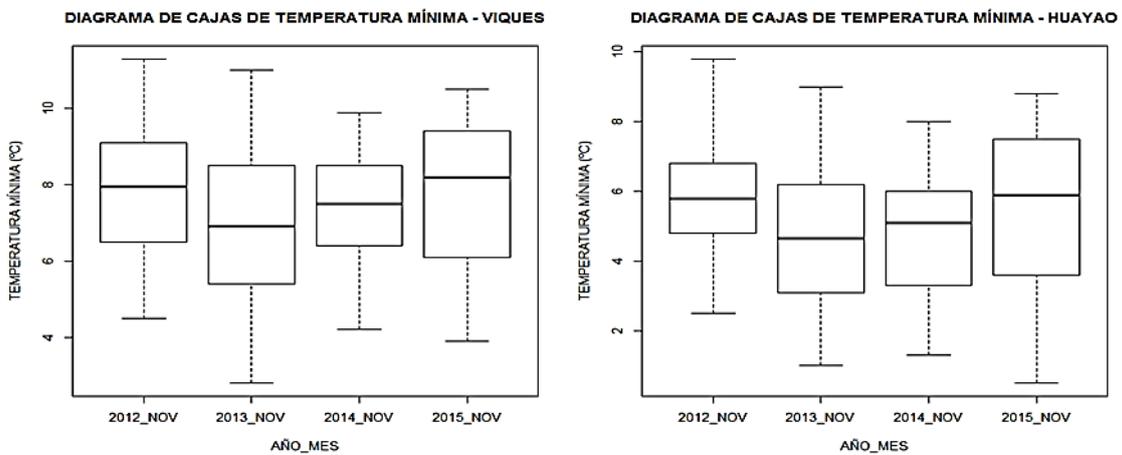
### ANEXO 2. ANÁLISIS EXPLORATORIO DE DATOS, DIAGRAMAS DE CAJAS Y SERIES DE TIEMPO.



**Figura 31:** Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de octubre.



**Figura 32:** Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de octubre.



**Figura 33:** Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de noviembre.

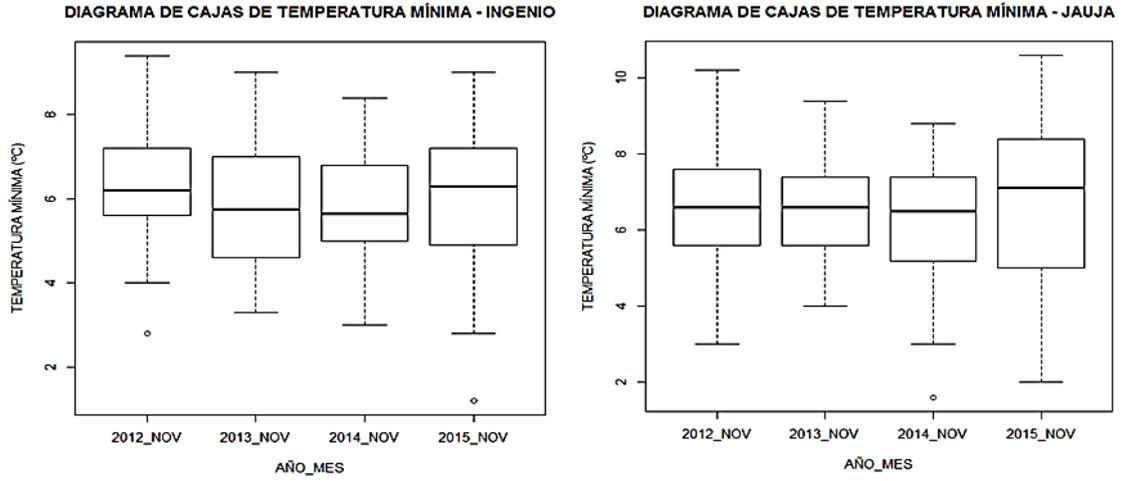


Figura 34: Continuación de la figura 32.

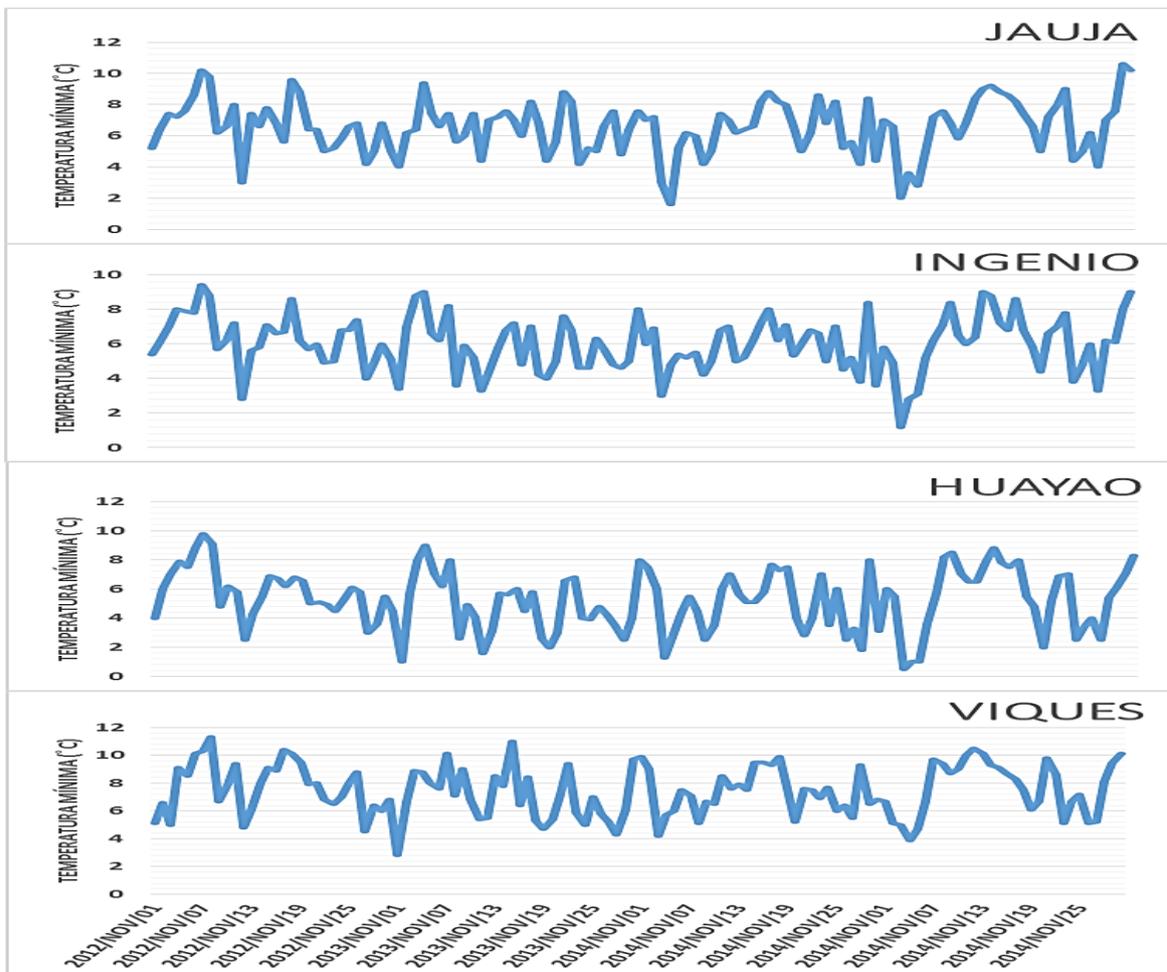
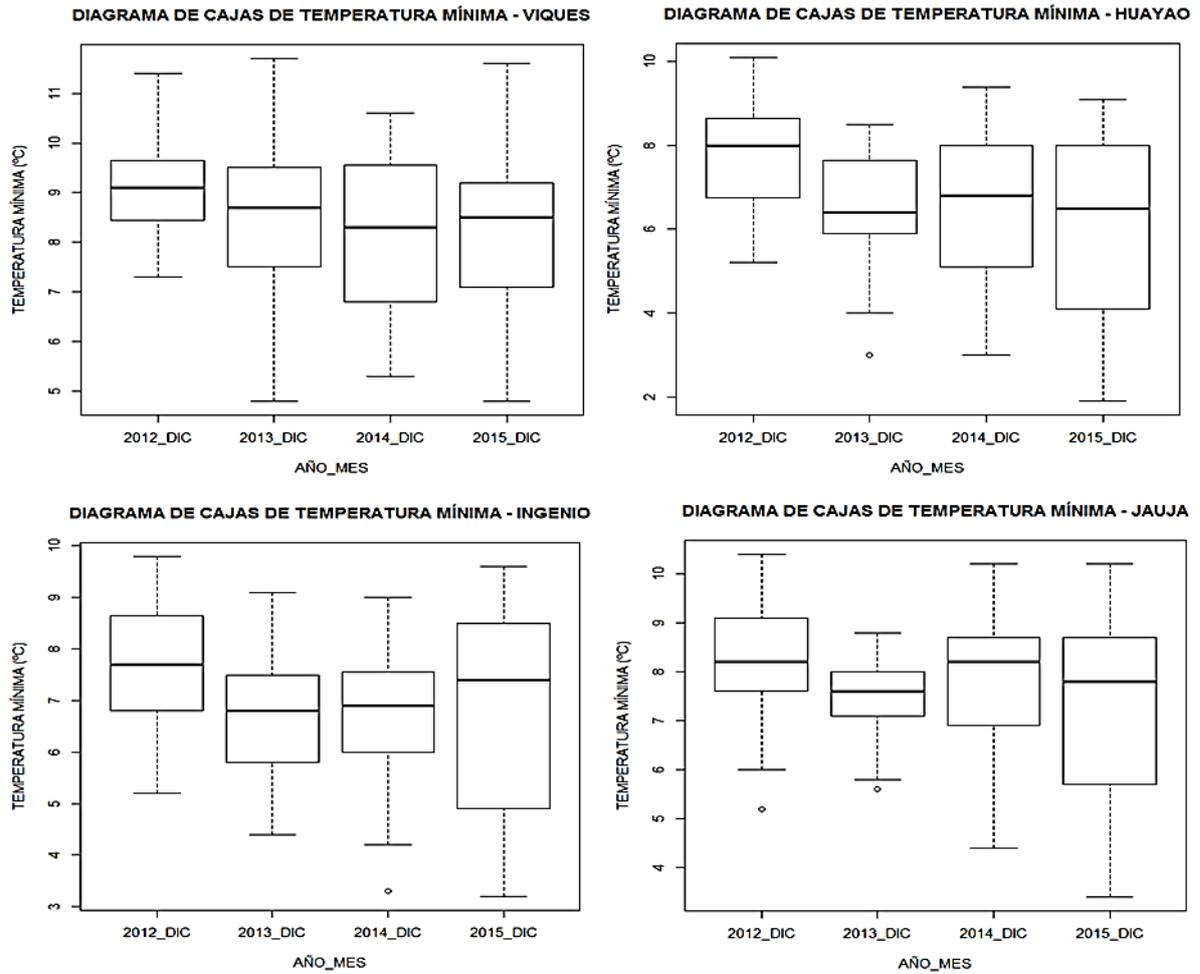
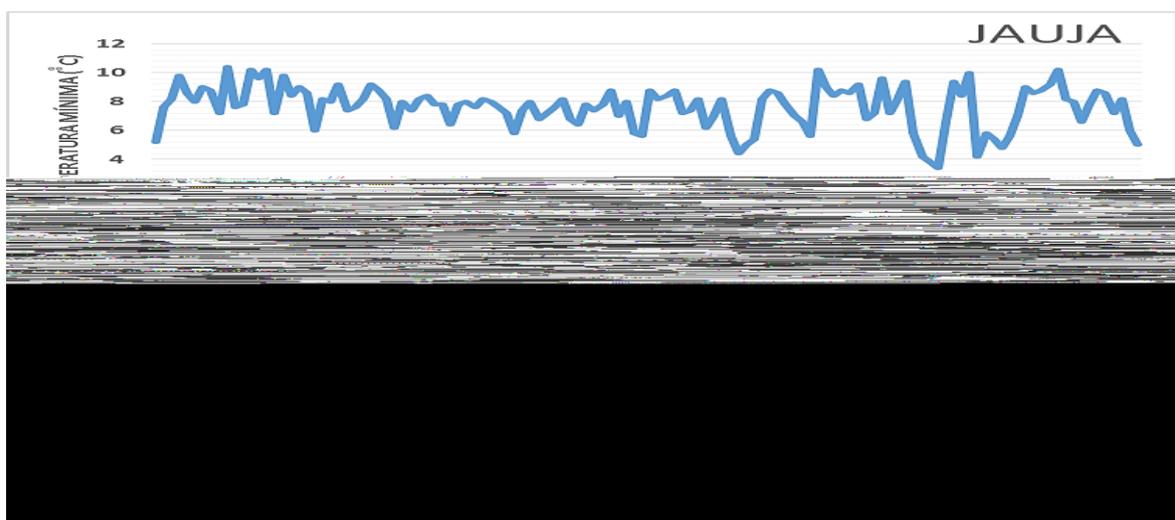


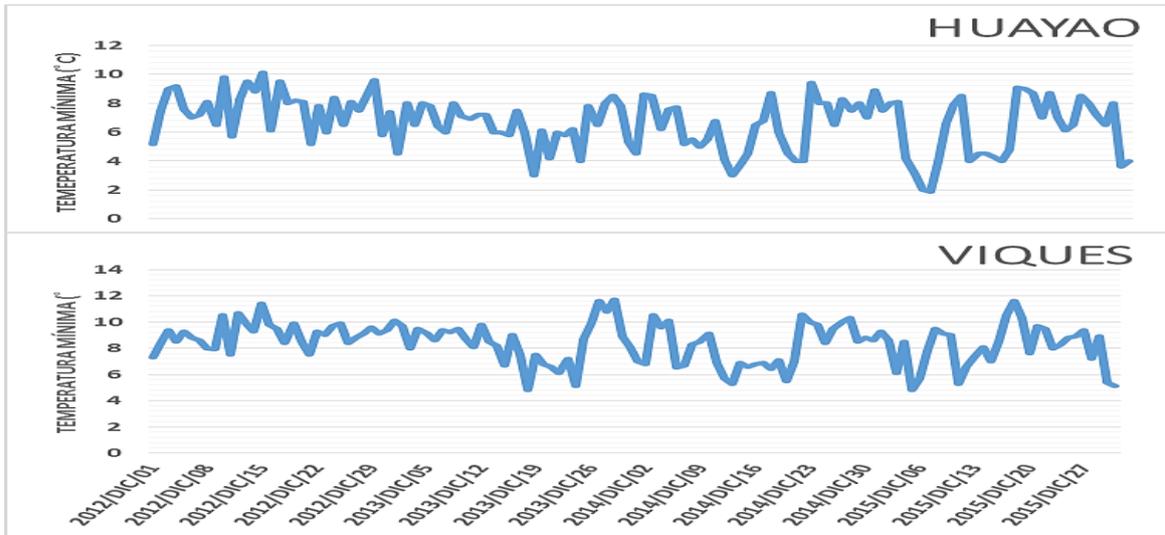
Figura 35: Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de noviembre.



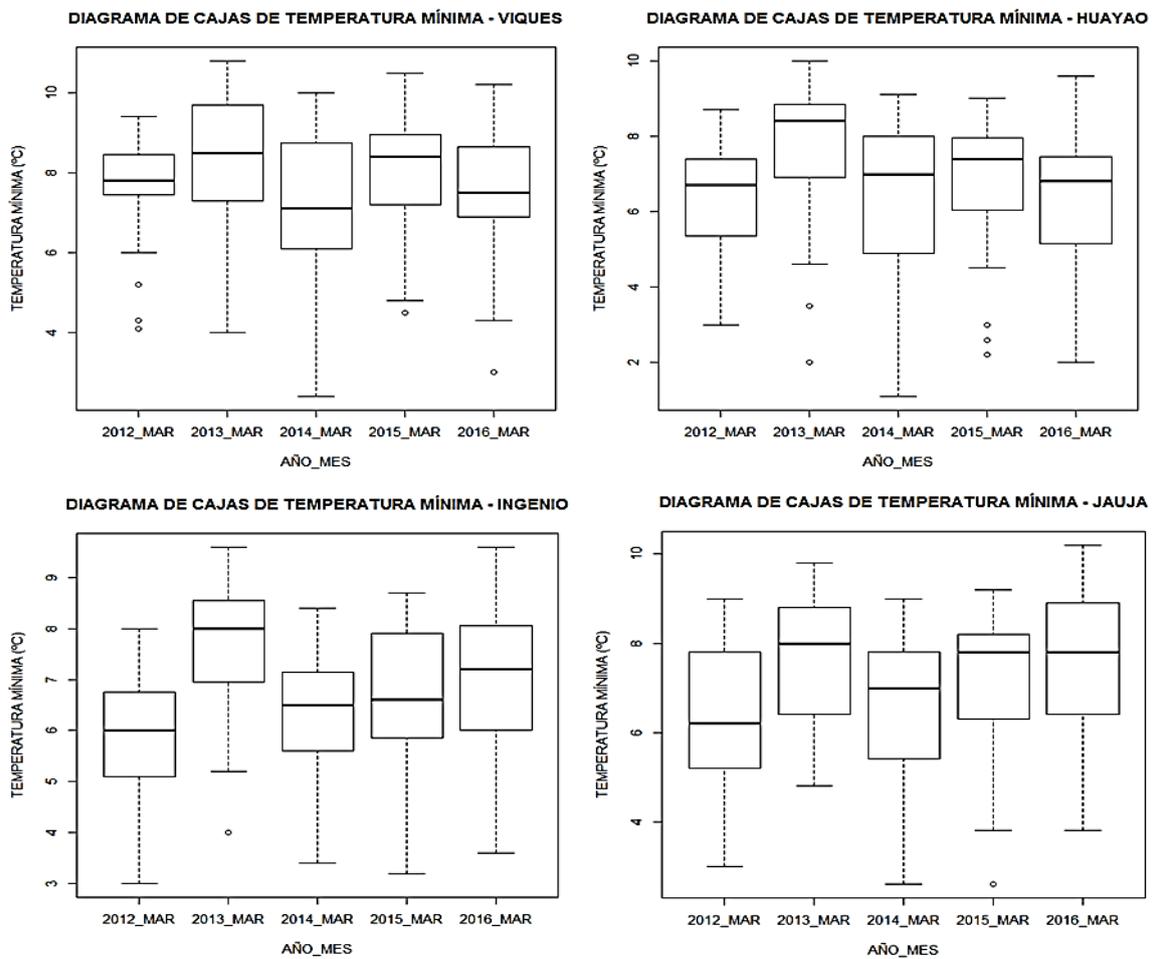
**Figura 36:** Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de diciembre.



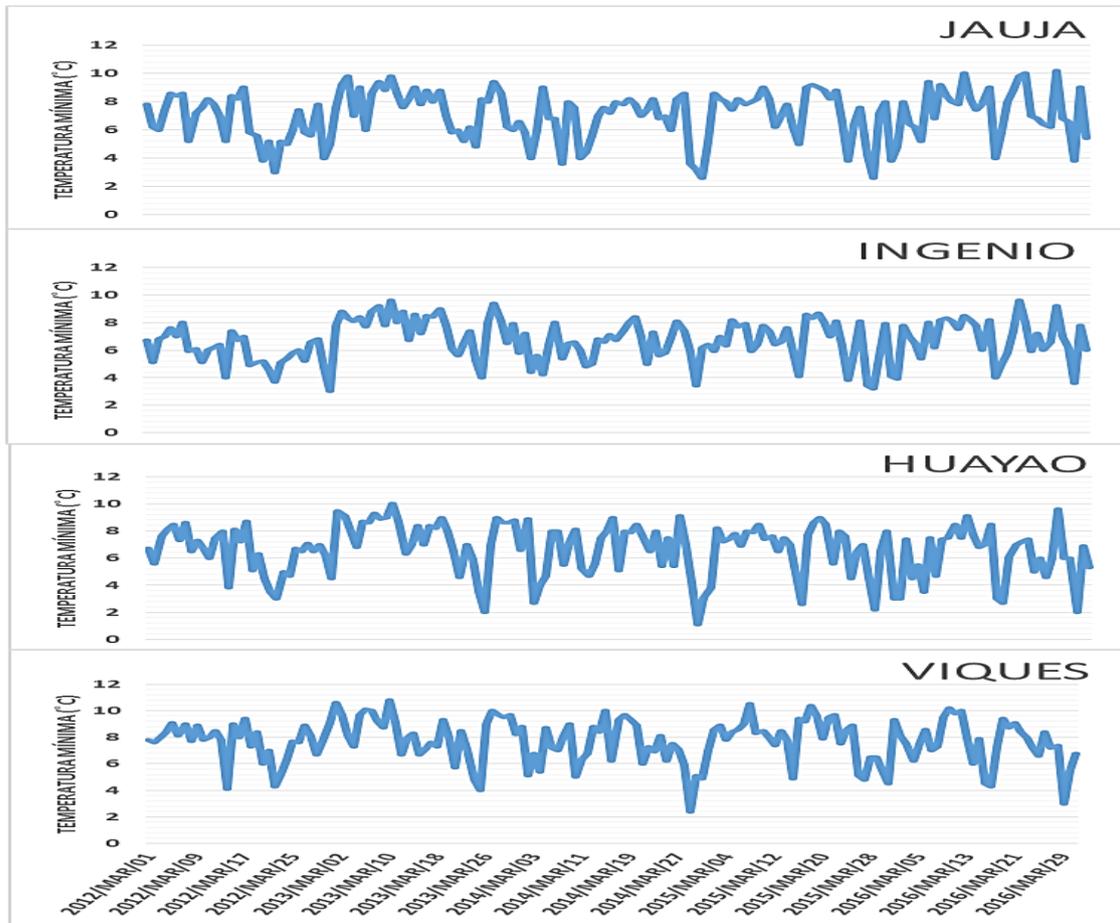
**Figura 37:** Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de diciembre.



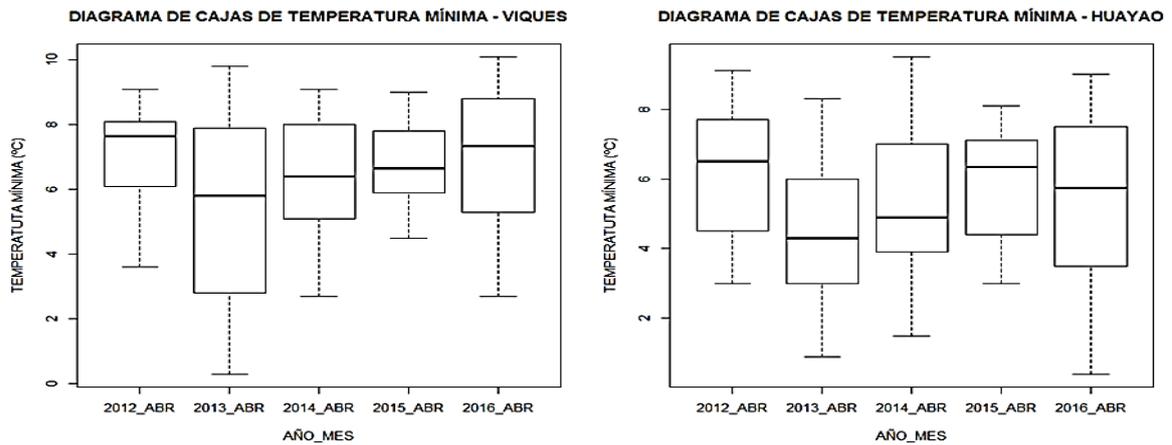
**Figura 38:** Continuación de la figura 36.



**Figura 39:** Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de marzo.



**Figura 40:** Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de marzo.



**Figura 41:** Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de abril.

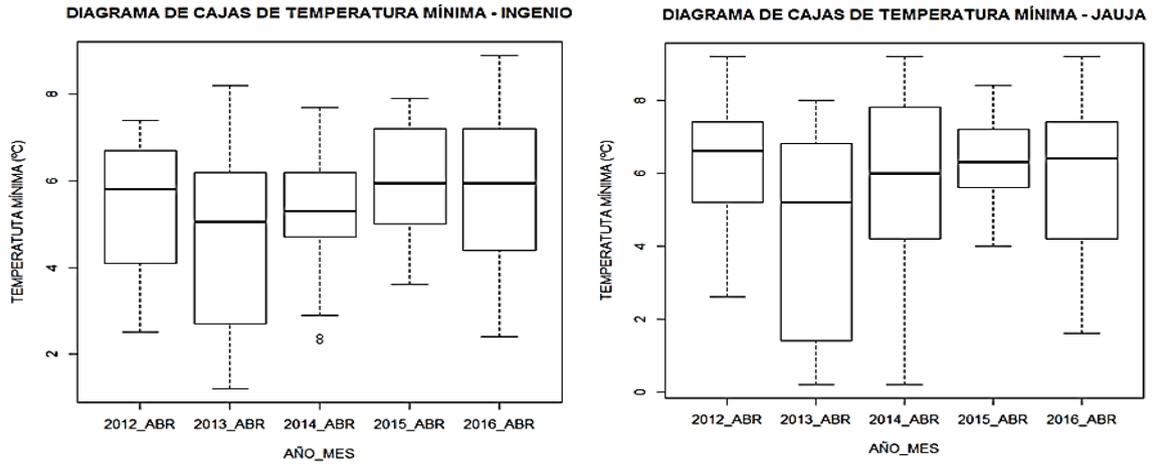


Figura 42: Continuación de la figura 40.

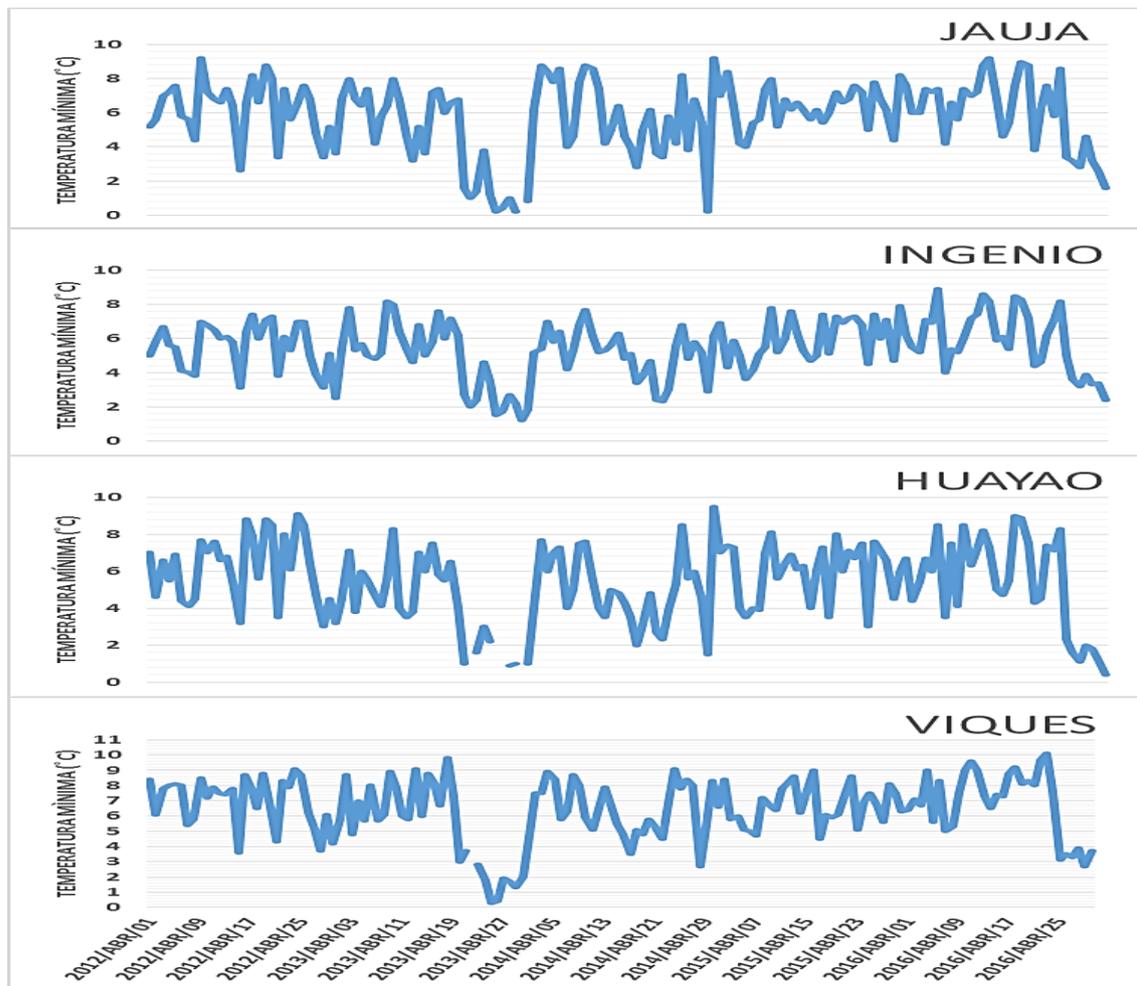
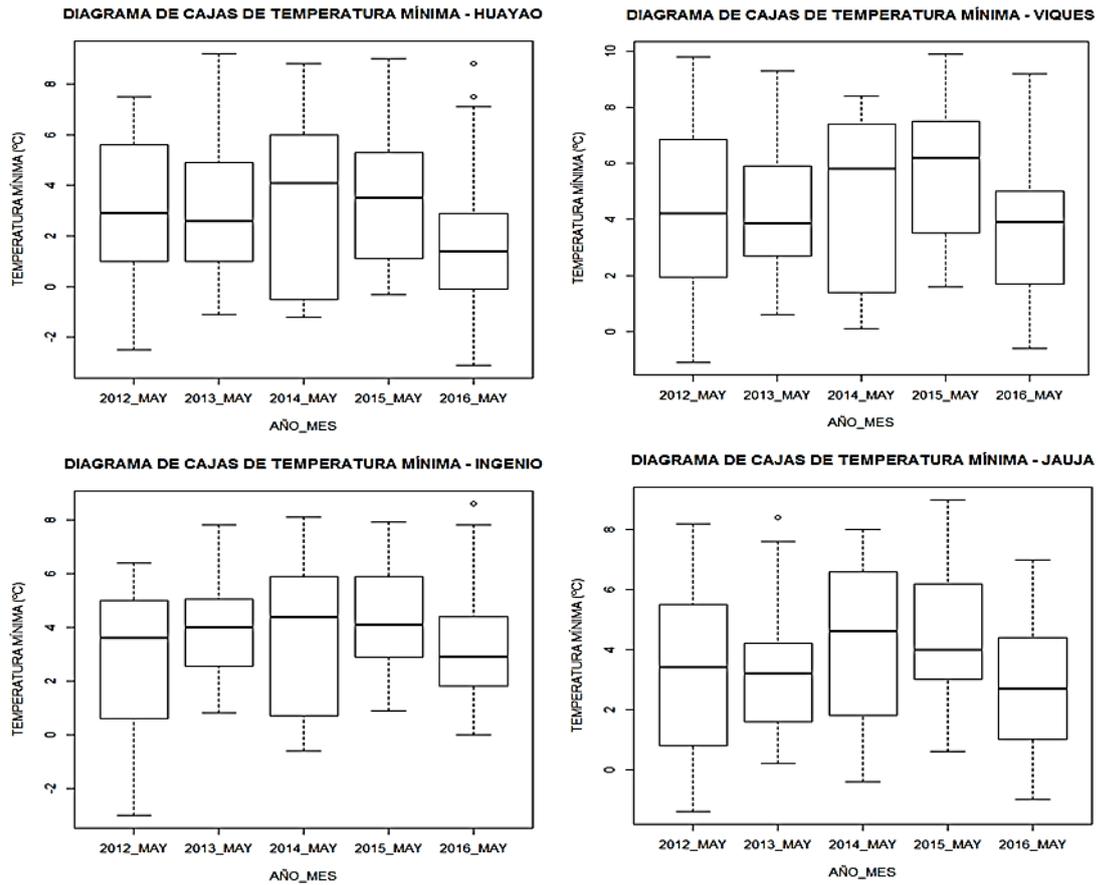
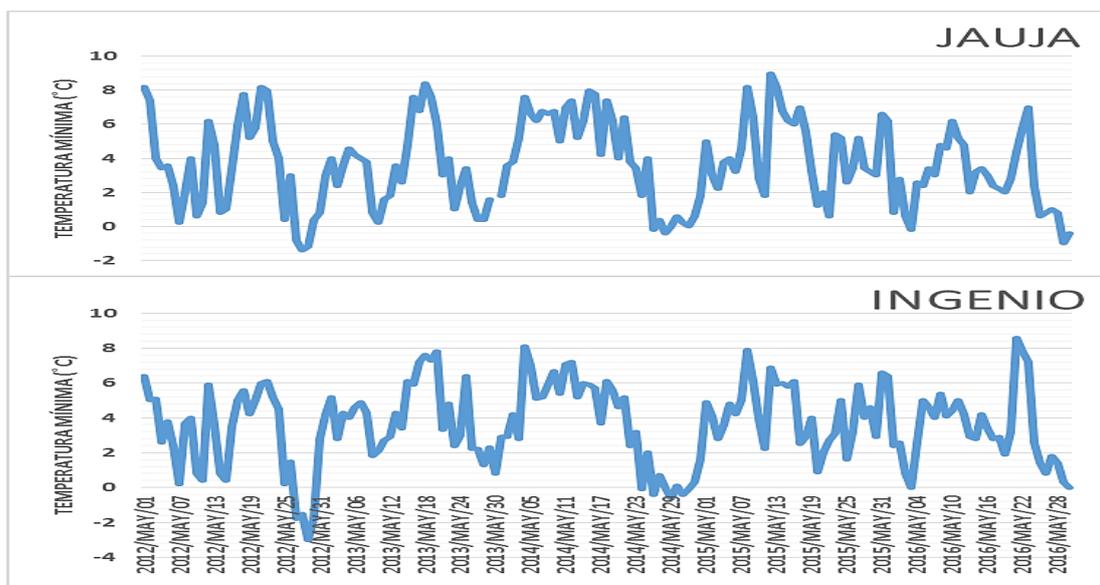


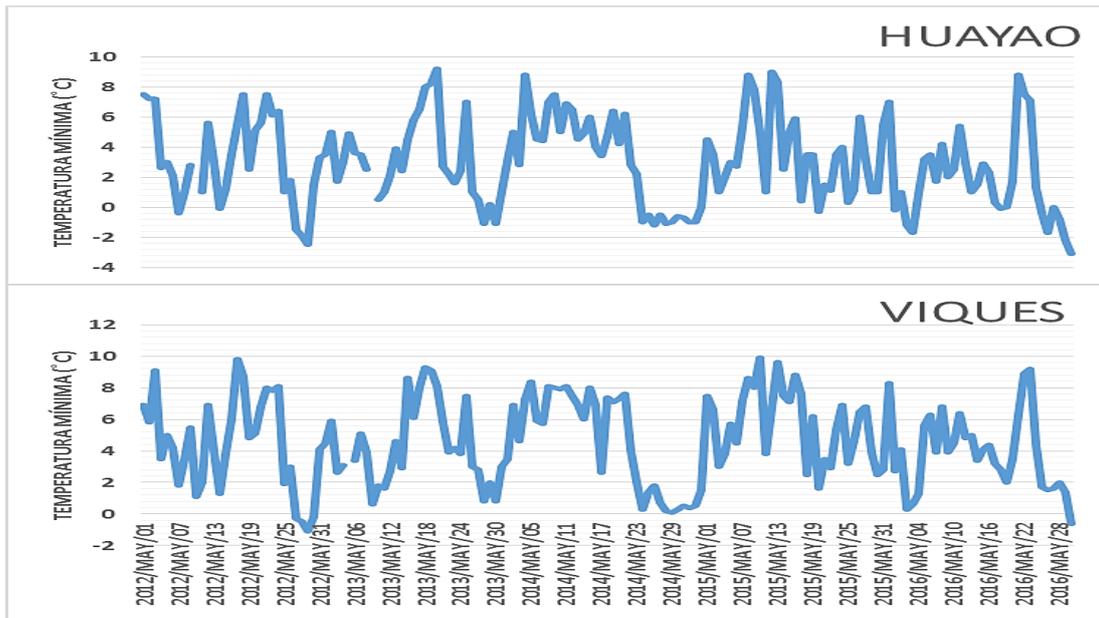
Figura 43: Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de abril.



**Figura 44:** Diagrama de cajas de temperatura mínima, de las cuatro estaciones de estudio para el mes de mayo.



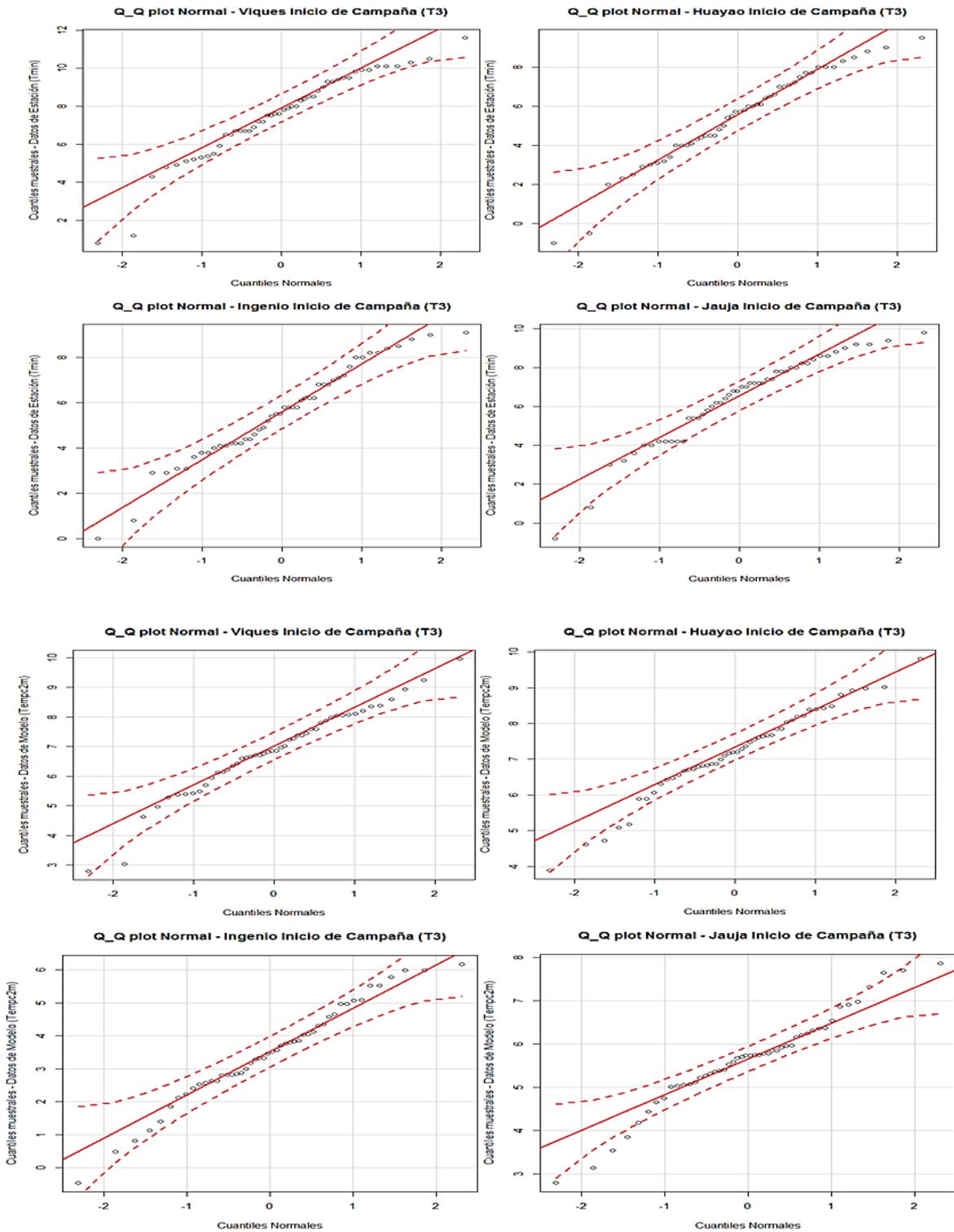
**Figura 45:** Series de tiempo de temperatura mínima de las cuatro estaciones de estudio para el mes de mayo.



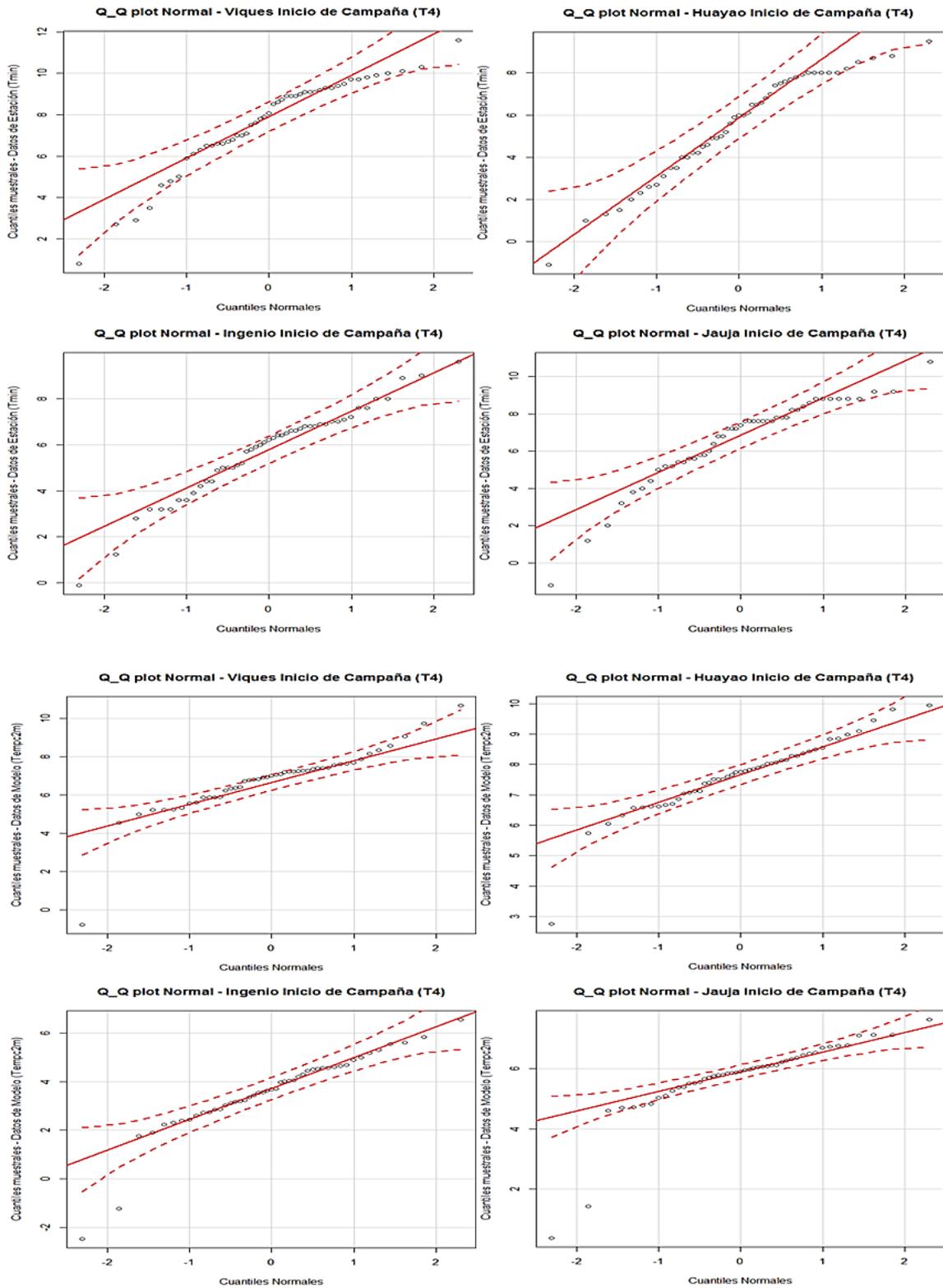
**Figura 46:** Continuación de la figura 44.

**ANEXO 3. ANÁLISIS DE NORMALIDAD DE LOS DATOS QUE SE TOMARON EN CUENTA PARA LA DETERMINACIÓN DE LOS MODELOS EMPÍRICOS**

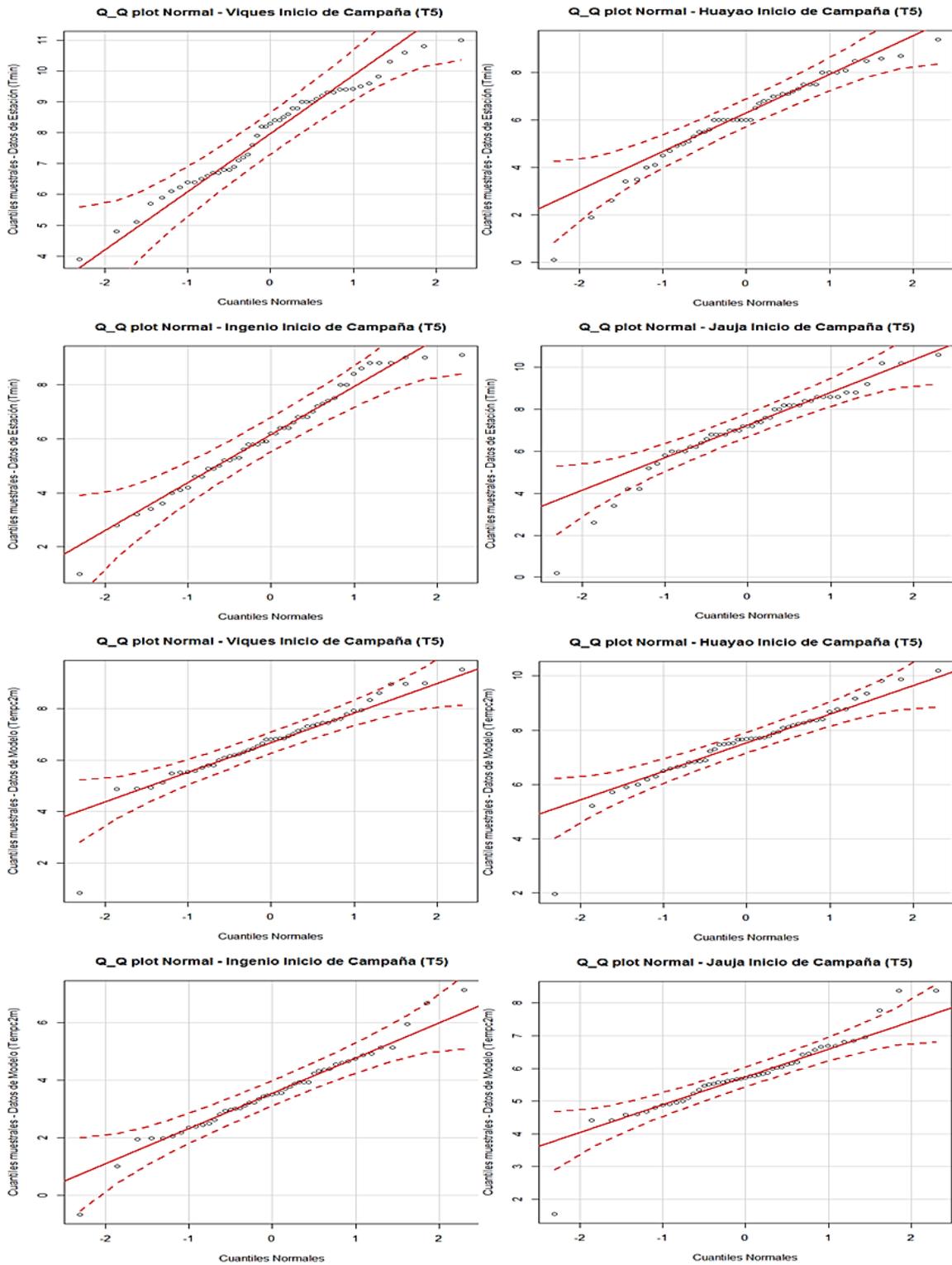
**a. Para inicio de campaña, en las cuatro estaciones de estudio, y para los tiempos de pronóstico (T3, T4 y T5).**



**Figura 47:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico tres (T3), para inicio de campaña y en las cuatro estaciones de estudio.

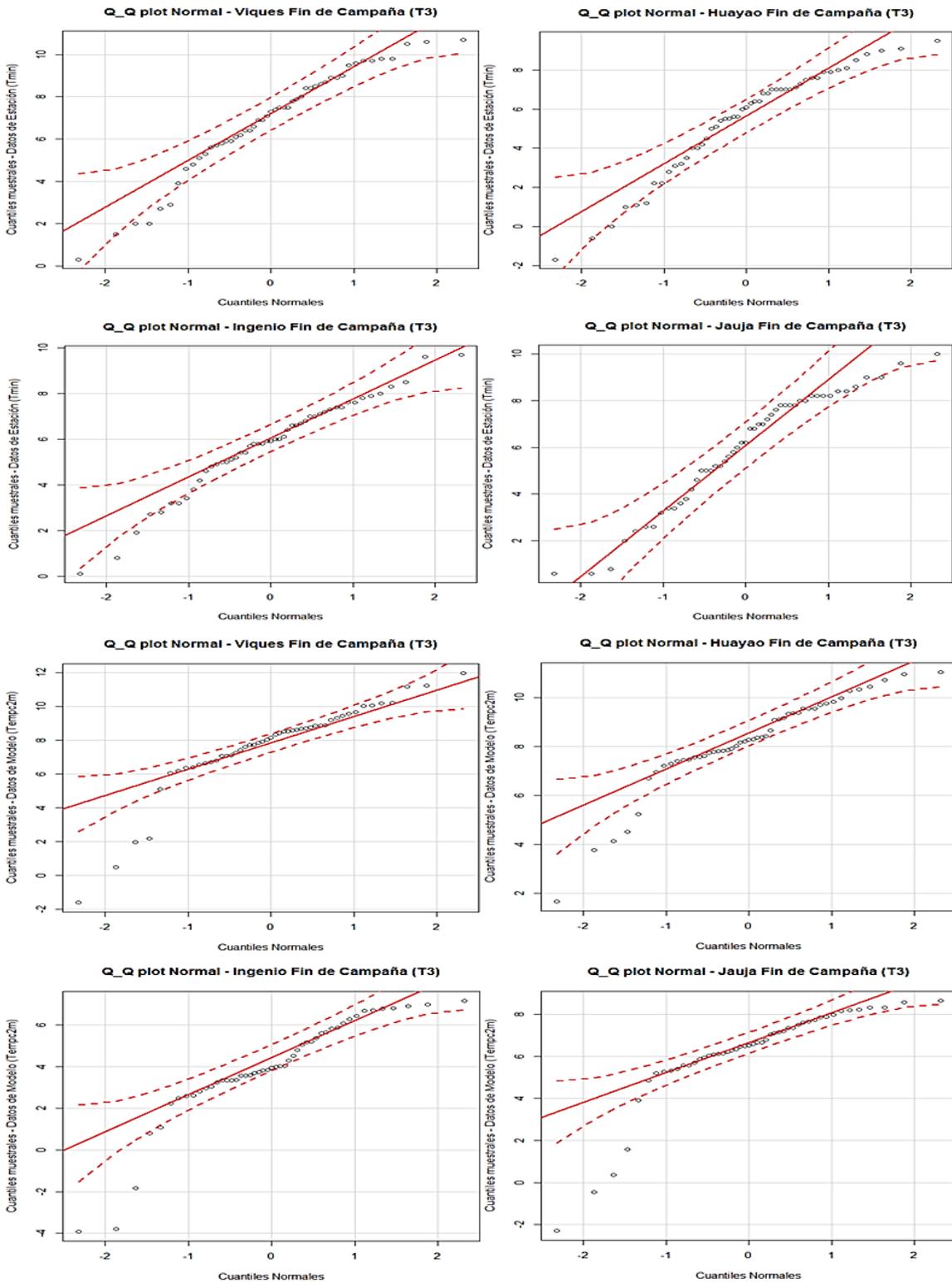


**Figura 48:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico cuatro (T4), para inicio de campaña y en las cuatro estaciones de estudio.

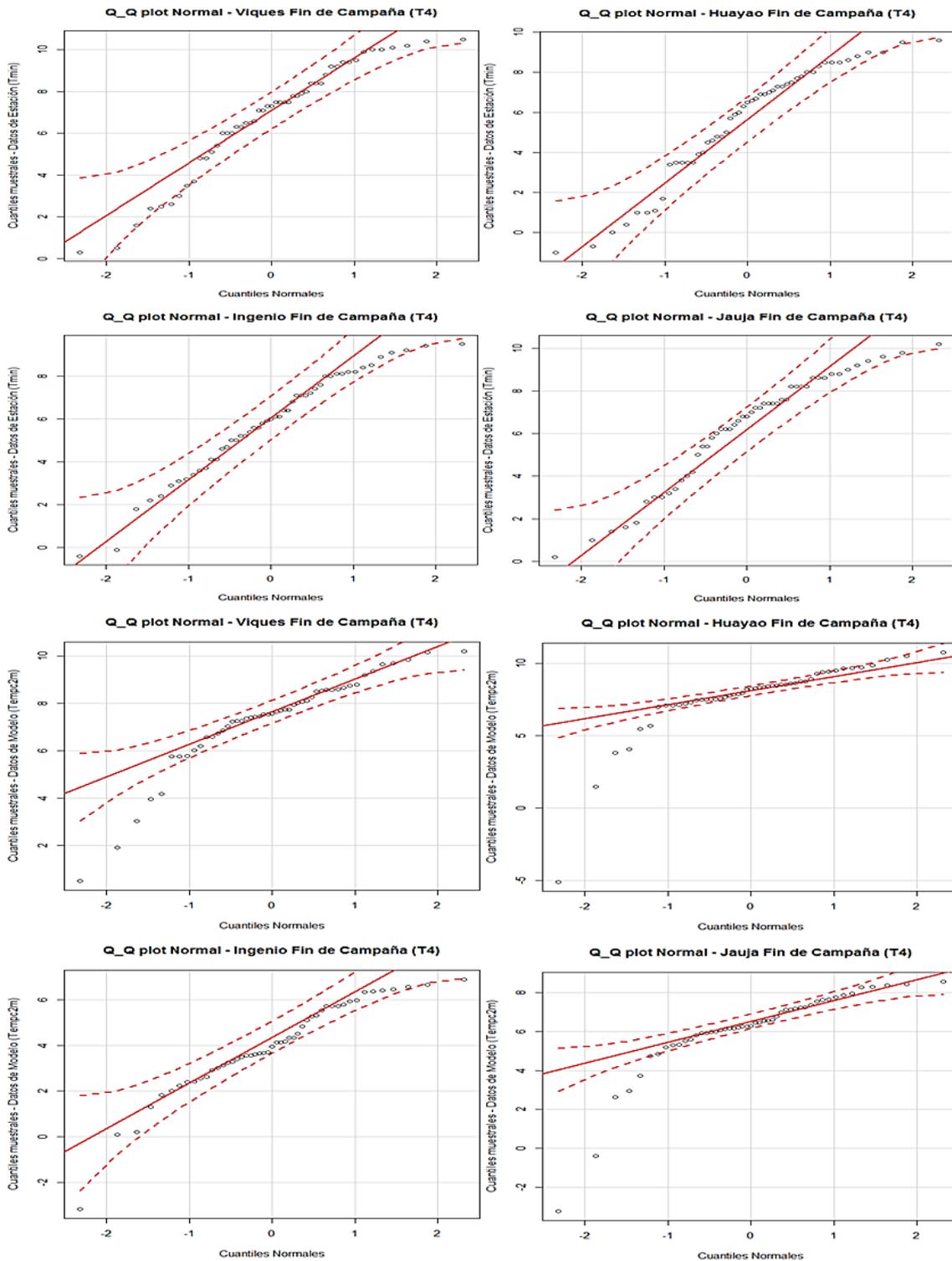


**Figura 49:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico cinco (T5), para inicio de campaña y en las cuatro estaciones de estudio.

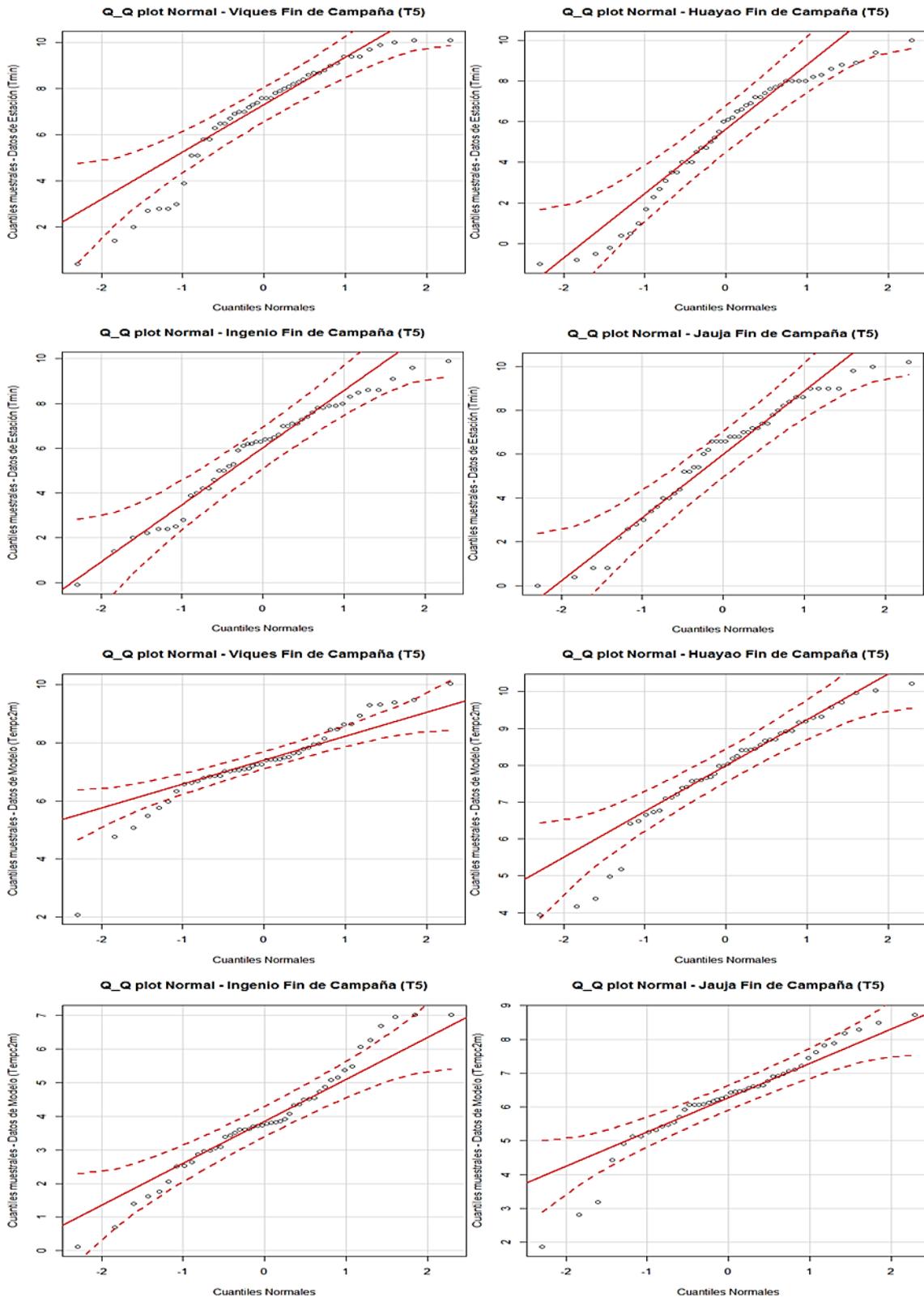
a. Para fin de campaña, en las cuatro estaciones de estudio, y para los tiempos de pronóstico (T3, T4 y T5).



**Figura 50:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico tres (T3), para fin de campaña y en las cuatro estaciones de estudio.



**Figura 51:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico cuatro (T4), para fin de campaña y en las cuatro estaciones de estudio.



**Figura 52:** Análisis de normalidad de temperaturas mínimas de estación (Tmin) y de temperaturas mínimas de pronóstico del BRAMS (Tmin\_P) para el tiempo de pronóstico cinco (T5), para fin de campaña y en las cuatro estaciones de estudio.

## ANEXO 4. ESTADÍSTICAS Y ANÁLISIS DE VARIANZA DE LOS MODELOS DE REGRESIÓN LINEAL

- Para el tiempo de pronóstico dos (T2) en inicio de campaña agrícola.

ESTADÍSTICAS DE LA REGRESIÓN				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,7640	0,5840	0,3530	0,4378
Coefficiente de Determinación R <sup>2</sup>	0,0202	0,3413	0,1245	0,1917
R <sup>2</sup> Ajustado	-0,0114	0,3200	0,0963	0,1656
Error Típico	1,9600	1,7580	1,7740	1,7110
Observaciones	33	33	33	33

ANÁLISIS DE VARIANZA					
VIQUES	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	2,460	2,460	0,640	0,4298
Residuos	31	119,110	38,424		
Total	32	121,57			
HUAYAO	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	49,605	49,605	16,059	0,00036
Residuos	31	95,756	3,089		
Total	32	145,361			
INGENIO	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	13,883	13,883	44,098	0,04397
Residuos	31	97,598	31,483		
Total	32	111,481			
JAUJA	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	21,513	21,513	73,505	0,01083
Residuos	31	90,729	29,268		
Total	32	112,242			

- Para el tiempo de pronóstico dos (T2) en fin de campaña agrícola.

<b>ESTADÍSTICAS DE LA REGRESIÓN</b>				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,2900	0,4329	0,4361	0,2300
Coefficiente de Determinación R <sup>2</sup>	0,0851	0,1874	0,1902	0,0525
R <sup>2</sup> Ajustado	0,0555	0,1655	0,1683	0,0220
Error Típico	2,1820	2,2460	1,9220	2,3300
Observaciones	33	39	39	33

<b>ANÁLISIS DE VARIANZA</b>					
<b>VIQUES</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	13,717	13,717	2,882	0,0996
Residuos	31	147,552	47,597		
Total	32	161,269			
<b>HUAYAO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	43,047	43,047	85,342	0,0059
Residuos	37	186,630	5,044		
Total	38	299,677			
<b>INGENIO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	32,118	32,118	86,909	0,0055
Residuos	37	136,738	3,696		
Total	38	168,856			
<b>JAUJA</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	9,326	9,326	17,182	0,1995
Residuos	31	168,255	54,276		
Total	32	177,581			

- Para el tiempo de pronóstico tres (T3) en inicio de campaña agrícola.

<b>ESTADÍSTICAS DE LA REGRESIÓN</b>				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,5286	0,5191	0,1857	0,1578
Coefficiente de Determinación R <sup>2</sup>	0,2794	0,2695	0,0345	0,0249
R <sup>2</sup> Ajustado	0,2562	0,2460	0,0034	-0,0066
Error Típico	2,0320	2,0450	2,1150	2,2710
Observaciones	33	33	33	33

<b>ANÁLISIS DE VARIANZA</b>					
<b>VIQUES</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	49,627	49,627	12,02	0,0016
Residuos	31	127,986	4,129		
Total	32	177,613			
<b>HUAYAO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	47,862	47,862	11,44	0,00196
Residuos	31	129,701	4,184		
Total	32	177,563			
<b>INGENIO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	4,958	4,958	1,1083	0,3006
Residuos	31	138,671	4,4733		
Total	32	143,629			
<b>JAUJA</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	4,078	4,078	0,7906	0,3808
Residuos	31	159,895	5,1579		
Total	32	163,973			

- Para el tiempo de pronóstico tres (T3) en fin de campaña agrícola.

<b>ESTADÍSTICAS DE LA REGRESIÓN</b>				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,1628	0,1657	0,2154	0,1319
Coefficiente de Determinación R <sup>2</sup>	0,0265	0,0275	0,0464	0,0174
R <sup>2</sup> Ajustado	-0,0049	-0,0039	0,0156	-0,0143
Error Típico	2,3820	2,4950	1,9700	2,4120
Observaciones	33	33	33	33

<b>ANÁLISIS DE VARIANZA</b>					
<b>VIQUES</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	4,780	4,780	0,8424	0,3658
Residuos	31	175,910	5,675		
Total	32	180,69			
<b>HUAYAO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	5,448	5,448	0,8754	0,3567
Residuos	31	192,913	6,223		
Total	32	198,361			
<b>INGENIO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	5,853	5,853	1,5089	0,2285
Residuos	31	120,256	3,8792		
Total	32	126,109			
<b>JAUJA</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	3,195	3,195	0,5491	0,4642
Residuos	31	180,361	5,8181		
Total	32	183,556			

- Para el tiempo de pronóstico cuatro (T4) en inicio de campaña agrícola.

<b>ESTADÍSTICAS DE LA REGRESIÓN</b>				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,3005	0,0877	0,1556	0,2154
Coefficiente de Determinación R <sup>2</sup>	0,0903	0,0077	0,0242	0,0464
R <sup>2</sup> Ajustado	0,0609	-0,0243	-0,0073	0,0157
Error Típico	2,339	2,512	1,921	2,389
Observaciones	33	33	33	33

<b>ANÁLISIS DE VARIANZA</b>					
<b>VIQUES</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	16,839	16,839	3,0774	0,0893
Residuos	31	169,628	5,4719		
Total	32	186,467			
<b>HUAYAO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	1,526	1,526	0,2419	0,6263
Residuos	31	195,599	6,3097		
Total	32	197,125			
<b>INGENIO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	2,833	2,833	0,7677	0,3877
Residuos	31	114,386	3,6899		
Total	32	117,219			
<b>JAUJA</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	8,611	8,611	1,509	0,2285
Residuos	31	176,895	5,7063		
Total	32	185,506			

- Para el tiempo de pronóstico cuatro (T4) en fin de campaña agrícola.

<b>ESTADÍSTICAS DE LA REGRESIÓN</b>				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,4738	0,2289	0,4444	0,4493
Coefficiente de Determinación R <sup>2</sup>	0,2245	0,0524	0,1975	0,2019
R <sup>2</sup> Ajustado	0,1995	0,0218	0,1758	0,1803
Error Típico	2,474	2,749	2,158	2,177
Observaciones	33	33	39	39

<b>ANÁLISIS DE VARIANZA</b>					
<b>VIQUES</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	54,919	54,919	8,9738	0,0053
Residuos	31	189,719	6,12		
Total	32	244,638			
<b>HUAYAO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	12,954	12,954	1,714	0,2001
Residuos	31	234,293	7,5578		
Total	32	247,247			
<b>INGENIO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	42,382	42,382	9,1041	0,0046
Residuos	37	172,247	4,655		
Total	38	214,629			
<b>JAUJA</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	44,35	44,35	9,3595	0,0041
Residuos	37	175,33	4,739		
Total	38	219,68			

- Para el tiempo de pronóstico cinco (T5) en inicio de campaña agrícola.

<b>ESTADÍSTICAS DE LA REGRESIÓN</b>				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,3298	0,0889	0,1792	0,0361
Coefficiente de Determinación R <sup>2</sup>	0,1088	0,0079	0,0321	0,0013
R <sup>2</sup> Ajustado	0,0801	-0,0241	0,0009	-0,0309
Error Típico	1,693	1,941	1,919	1,941
Observaciones	33	33	33	33

<b>ANÁLISIS DE VARIANZA</b>					
<b>VIQUES</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	10,846	10,846	3,7849	0,0608
Residuos	31	88,835	2,8656		
Total	32	99,681			
<b>HUAYAO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	0,935	0,935	0,2481	0,6219
Residuos	31	116,771	3,7668		
Total	32	117,706			
<b>INGENIO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	3,786	3,786	1,0277	0,3185
Residuos	31	114,184	3,6834		
Total	32	117,97			
<b>JAUJA</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	0,151	0,151	0,04	0,8428
Residuos	31	116,716	3,7666		
Total	32	116,867			

- Para el tiempo de pronóstico cinco (T5) en fin de campaña agrícola.

<b>ESTADÍSTICAS DE LA REGRESIÓN</b>				
	VIQUES	HUAYAO	INGENIO	JAUJA
Coefficiente de Correlación	0,0539	0,1425	0,2851	0,1884
Coefficiente de Determinación R <sup>2</sup>	0,0029	0,0203	0,0813	0,0355
R <sup>2</sup> Ajustado	-0,0292	-0,0113	0,0517	0,0043
Error Típico	2,528	2,941	2.205	2,486
Observaciones	33	33	33	33

<b>ANÁLISIS DE VARIANZA</b>					
<b>VIQUES</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	0,585	0,585	0,0915	0,7643
Residuos	31	198,045	6,3885		
Total	32	198,63			
<b>HUAYAO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	5,562	5,562	0,6429	0,4288
Residuos	31	268,16	8,6503		
Total	32	273,722			
<b>INGENIO</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	13,34	13,34	2,7447	0,1077
Residuos	31	150,67	4,8603		
Total	32	164,01			
<b>JAUJA</b>	Grados de Libertad	Suma de Cuadrados	Promedios de los Cuadrados	F	Valor Crítico de F
Regresión	1	7,048	7,048	1,14	0,2939
Residuos	31	191,653	6,1823		
Total	32	198,701			