




PERÚ

Ministerio
del Ambiente

Servicio Nacional de
Meteorología e Hidrología
del Perú - SENAMHI



CORRECCIÓN DE BIAS PARA PRONÓSTICOS DE PRECIPITACIÓN DEL MODELO ETA/SENAMHI

**DIRECCIÓN DE METEOROLOGÍA Y
EVALUACIÓN AMBIENTAL
ATMOSFÉRICA**

**SERVICIO NACIONAL DE METEOROLOGÍA E HIDROLOGÍA DEL PERÚ
– SENAMHI –**

**NOVIEMBRE, 2016
LIMA, PERÚ**



PRESIDENTA EJECUTIVA DEL SENAMHI
Ing. AMELIA DIAZ PABLÓ

**DIRECTOR GENERAL DE METEOROLOGÍA Y EVALUACIÓN
AMBIENTAL ATMOSFÉRICA**
Lic. LUIS ALFARO LOZANO

**SUB-DIRECTORA DE MODELAMIENTO NUMÉRICO DE LA
ATMOSFERA**
Ing. GABRIELA ROSAS BENANCIO

AUTOR
Met. GUSTAVO DE LA CRUZ MONTALVO

2016

TABLA DE CONTENIDO

I.	INTRODUCCIÓN.....	1
II.	OBJETIVOS	1
III.	DATOS Y METODOLOGÍA.....	2
3.1.	Datos.....	2
3.1.1.	Datos observados de precipitación	2
3.1.2.	Información del modelo ETA/SENAMHI	3
3.2.	Metodología.....	3
3.2.1.	Metodologías de corrección de bias	4
3.2.2.	Análisis de la eficiencia de las distintas técnicas de corrección de bias. 7	
IV.	RESULTADOS.....	10
4.1.	Resultados de la corrección de bias.	10
4.2.	Análisis de la eficiencia por cada método de corrección	14
4.3.	Análisis de extremos de precipitación.....	18
V.	CONCLUSIONES.....	22
VI.	RECOMENDACIONES	23
VII.	BIBLIOGRAFÍA.....	24
VIII.	ANEXOS	25
8.1.	Correlación de spearman y root mean squared error (RMSE) del modelo ETA/SENAMHI.....	25
8.2.	Series diarias de precipitación.....	28

I. INTRODUCCIÓN

Actualmente, los modelos numéricos son una herramienta bien importante para el pronóstico del tiempo e investigación de los diversos procesos que ocurren en la atmósfera. Debido a la fuerte influencia que tienen los factores meteorológicos sobre las diferentes actividades humanas, es de gran requerimiento tener pronósticos mucho más precisos y aterrizados, especialmente de variables tan importante como lo es la precipitación. Sin embargo, los modelos no siempre presentan simulaciones que sean consistente con lo observado y generalmente tiene errores sistemáticos y aleatorios que deben ser analizados. Es por ello que un paso importante en la generación de pronósticos numéricos es el pos-procesamiento de las salidas del modelo, que se realiza a través de técnicas de ajustes de bias usando modelos estadísticos calibrados con series históricas de datos observados y del modelo, proceso comúnmente conocido como *Model Output Statistic* (MOS). Existen una gran variedad de técnicas estadísticas que se requiere desarrollar, calibrar y validar con el fin de eliminar los errores sistemáticos de los modelos. En ese contexto y teniendo la alta complejidad de pronosticar la precipitación debido a su alta dependencia a la geografía y de procesos a diferentes escalas, el presente trabajo tiene como objetivo calibrar y validar cuatro metodologías de corrección de bias para la precipitación pronosticada del modelo ETA/SENAMHI con la finalidad de conocer la habilidad y limitaciones de estos.

II. OBJETIVOS

- Calcular series de precipitación corregida usando métodos de corrección de bias para pronósticos de precipitación del modelo ETA-SENAMHI (22 km de resolución).
- Analizar la eficiencia de las distintas metodologías estadísticas de corrección de bias en el pronóstico cuantitativo de precipitación y sus extremos.

III. DATOS Y METODOLOGÍA

3.1. Datos

3.1.1. Datos observados de precipitación

En este estudio se utilizó información pluviométrica acumulada en 24 horas de 365 estaciones distribuidos en todo el Perú y las cuales todas pertenecen a la red de estaciones del SENAMHI. El periodo al cual corresponde esta información es de Enero del 2010 a Diciembre del 2015. La distribución espacial de las estaciones se puede apreciar en la figura 1.

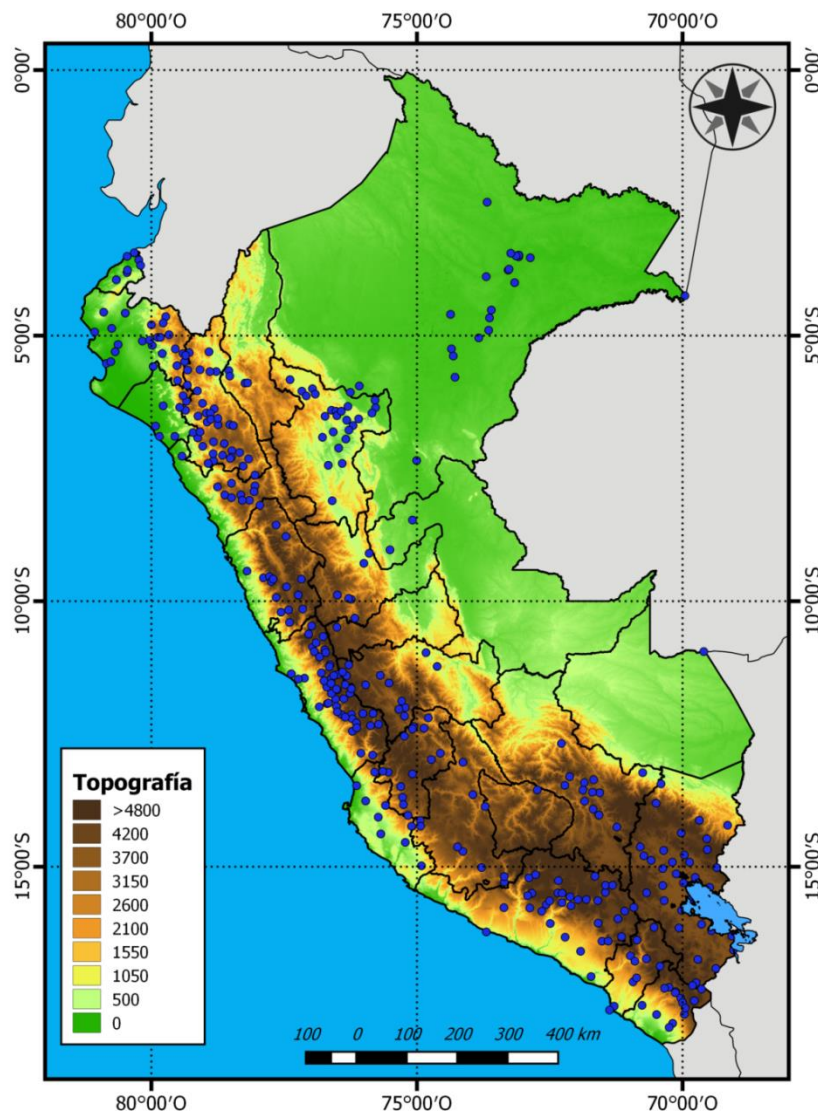


Fig. 1. Mapa de topografía y localización de estaciones utilizadas en el estudio.

3.1.2. Información del modelo ETA/SENAMHI

El modelo ETA es un modelo atmosférico regional usado para propósitos de investigación y operacional. Actualmente en el SENAMHI se viene corriendo el modelo ETA en forma operacional en dos dominios distintos que cubren Sudamérica y Perú. El dominio de Sudamérica posee una resolución horizontal de 32 Km y 50 niveles en la vertical, mientras que el segundo cuenta con 22 Km de resolución horizontal y 38 niveles en la vertical. Las previsiones son realizadas con intervalos de 6 horas a las 00Z y 12Z, extendiéndose hasta 120 horas para el dominio de Sudamérica y 72 horas para el dominio de Perú. Como condiciones iniciales este modelo utiliza datos del modelo global océano-atmosférico GFS (Global Forecast System) el cual es un modelo global de pronóstico producido por NCEP (National Centers for Environmental Prediction).

La información utilizada para este estudio corresponde a la precipitación acumulada en 24 horas para dos días, es decir, es la precipitación acumulada entre las 12 a 36 horas (Día 1) y de 36 a 60 horas (Día 2) de previsión. Dicha información se extrajo de la grilla más cercana a las coordenadas de las estaciones meteorológicas y para el mismo periodo de tiempo de estudio, 2010-2015. La parametrización de convección que usa esta configuración del modelo es la de Kain – Fritsch el cual está basado en un consumo completo de la energía potencial convectiva disponible (CAPE) en condiciones de diferentes mezclas por un efecto de arrastre para varios niveles de presión. El ciclo convectivo en esta parametrización es de 30 a 60 minutos de duración y tiende a dejar una capa profundamente saturada (COMET, 2012).

3.2. Metodología

La metodología planteada en el presente estudio consiste en el uso de cuatro técnicas de corrección de bias de la precipitación del modelo ETA/SENAMHI y su posterior validación. Una de estas técnicas utiliza un factor de corrección multiplicativo mientras que las otras tres son variantes del Quantile mapping el cual es un método de corrección basado en el ajuste de cuantiles de los datos de precipitación del modelo. Todos estos métodos fueron aplicados para los

datos de puntos de grilla correspondientes a cada una de las estaciones, y para los pronósticos de Día 1 y Día 2.

3.2.1. Metodologías de corrección de bias

a) Método Scaling (SCAL)

Este método consiste en ajustar la precipitación modelada multiplicándolo con un factor calculado en base a series largas de precipitación observada y modelada (Weterball et. al, 2012) tal y como se muestra en la siguiente formula:

$$\widehat{PP}_{mod} = PP_{mod} \times \frac{\overline{PP}_{obs}}{\overline{PP}_{mod}} \dots\dots\dots(1)$$

Donde:

\widehat{PP}_{mod} = Precipitación del modelo ajustado estadísticamente

PP_{mod} = Precipitación del modelo sin ajuste.

\overline{PP}_{obs} = Promedio de la precipitación observada sobre el periodo de calibración.

\overline{PP}_{mod} = Promedio de la precipitación del modelo sobre el periodo de calibración.

Según Lenderink et al. (2007) este método remueve bias en los promedios pero no el coeficiente de varianza de la precipitación modelada.

b) Empirical Quantil Mapping (eQM)

Este método es ampliamente usado en corrección de bias de distintas variables meteorológicas. El eQM está basado en una función no paramétrica que compensa promedio, variabilidad y tamaño de los errores de las funciones de distribución acumulada (CDF's) de la precipitación (Amengual et. al, 2011). Utiliza las distribuciones empíricas acumuladas de la precipitación observada y modelada y calcula los cambios, cuantil por cuantil entre dichas distribuciones. Posteriormente, dichos cambios son añadidos en cada cuantil a la distribución de precipitación modelada, para obtener una nueva serie calibrada.

$$\widehat{PP}_{mod} = F_{obs}^{-1}(F_{sim}(PP_{mod})) \dots\dots\dots (2)$$

Donde:

F_{obs}^{-1} = Función inversa de la distribución empirica acumulada observada.

F_{sim} = Función de la distribución empirica acumulada del modelo.

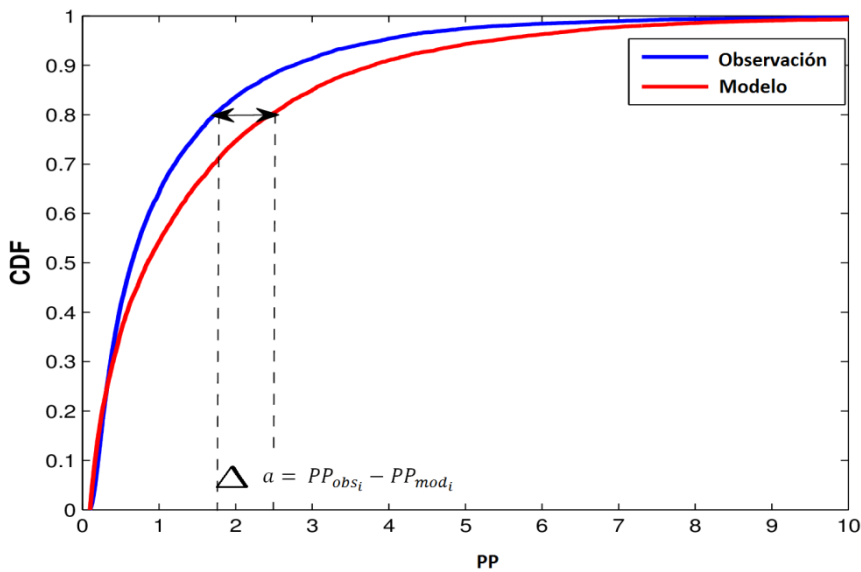


Fig 2. Ilustración del método eQM (modificado de Henning, 2016)

La figura 2 muestra de manera gráfica como el método eQM trabaja, siendo Δa el cambio del dato del modelo comparado con el observado en el percentil i . Estos cambios se obtuvieron para los 99 percentiles de cada serie y también para los valores entre dos percentiles en base a una interpolación lineal.

c) Gamma Quantil Mapping (gQM)

Este método reemplaza la empírica función CDF con una distribución paramétrica y se basa en la suposición de que tanto como las distribuciones de intensidad observada y simulada son bien representados por la distribución Gamma la cual depende de dos parámetros y es definido como:

$$pdf(pp) = \frac{\left(\frac{pp}{\beta}\right)^{\alpha-1} e^{\left(\frac{-pp}{\beta}\right)}}{\beta\Gamma(pp)} \quad pp, \alpha, \beta > 0 \dots\dots\dots (3)$$

Donde:

β = Parámetro de escala

α = Parámetro de forma

Γ =Función gamma

Los parámetros de la distribución son estimados mediante método de máxima verosimilitud. La distribución gamma no es definido para $pp=0\text{mm/día}$; por lo tanto, el proceso de corrección se realiza en dos pasos (Piani et al., 2010b). Primero, el número de días secos es corregido usando un valor umbral de precipitación (0.1 mm en este trabajo) de tal forma que valores de precipitación del modelo por debajo de este umbral son iguales a 0. Segundo, las distribuciones gammas son integradas para obtener sus respectivos CDF's y posteriormente ajustadas para las series de precipitación modelada y observada diferente de cero y reemplazar las empíricas CDF's.

d) Gamma-Pareto Quantile Mapping (gpQM)

Este método fue propuesto por Gutjahr & Heinemann (2013) y al igual que el método anterior este también es un método paramétrico de quantile mapping el cual utiliza una combinación de distribución gamma y de Pareto Generalizado (GPD). Como primer paso, se realiza el mismo ajuste de días secos en base al valor umbral de 0.1 mm. Segundo, se realiza el ajuste a las distribuciones paramétricas considerando un valor umbral de percentil 95 (propuesto por Yang et. al, 2010 y usado también por Gutjahr & Heinemann, 2013); de tal forma que valores por debajo del percentil 95 son ajustados a una distribución gamma; mientras que valores mayores o iguales se ajustan a una distribución Pareto.

$$\widehat{pp}_{mod} \begin{cases} F_{obs, gamma}^{-1}(F_{mod, gamma}), & \text{si } pp < \text{percentil } 95 \\ F_{obs, GPD}^{-1}(F_{mod, GPD}), & \text{si } pp \geq \text{percentil } 95 \end{cases} \dots\dots(4)$$

3.2.2. Análisis de la eficiencia de las distintas técnicas de corrección de bias.

a) Método de validación cruzada

Para el análisis de los valores de precipitación Las cuatro metodologías de corrección estadísticas mencionadas en el anterior capítulo fueron aplicados usando un periodo de calibración y otro para la validación. Un método de selección de estos periodos comúnmente usados es el de usar un 80% de todo el periodo para calibración y 20% para la validación; sin embargo, este método es limitado ya que no se puede evaluar al modelo en distintas condiciones que no pueden estar presentes en los periodos de validación y además, el periodo de validación no puede ser lo suficientemente largo para generar series independientes y obtener una buena comprensión de la performance del modelo. En ese sentido, se utilizó un método de validación cruzada (conocido en inglés como k-fold cross validation) para evitar los sobreajustes, el cual se basa en dividir la cantidad de datos disponibles en k subconjuntos que no se solapan, de tal forma que uno de ellos es usado para la validación, mientras que los $k-1$ subconjuntos son usados conjuntamente para la calibración. Dicho proceso es repetido k veces utilizando diferentes subconjuntos de validación en cada paso o iteración, con el fin de obtener una serie total de validación con la misma cantidad de elementos de la serie observada.

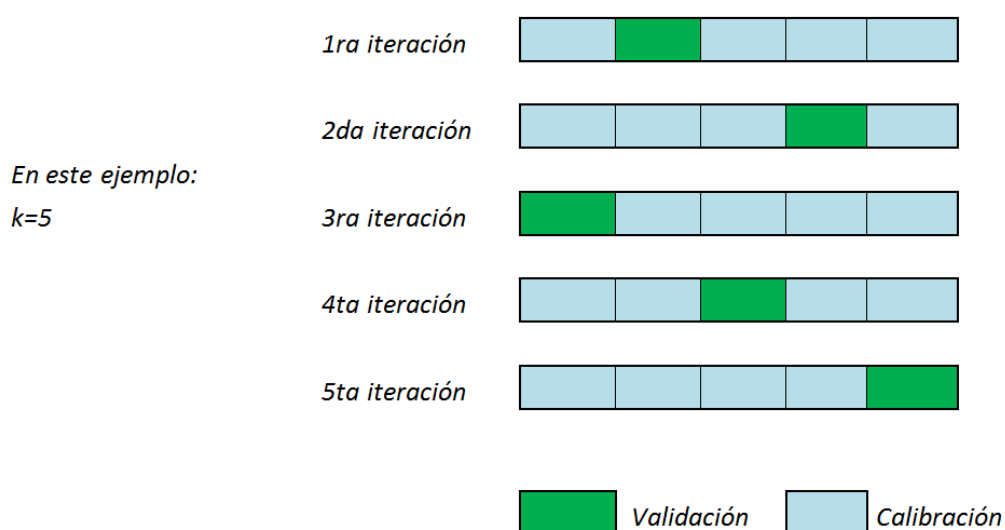


Fig. 3. Ejemplo de validación cruzada.

Para este estudio se aplicó un esquema de validación cruzada con $k=5$ subconjuntos a nivel trimestral (Diciembre-Enero-Febrero, Marzo-Abril-Mayo, Junio-Julio-Agosto, Septiembre-Octubre-Noviembre) y para cada metodología de corrección de bias. Ya que la precipitación en el Perú tiene un comportamiento estacional bien marcado, es conveniente obtener parámetros de calibración y realizar el análisis por cada trimestre del año.

b) Verificación de extremos de precipitación

Para el análisis de extremos de precipitación se analizaron los percentiles 90, 95 y 99 para los datos observados, ETA/SENAMHI y de cada uno de los métodos de corrección. El cálculo de los percentiles se ha basado en la metodología establecida por Alfaro (2014), considerando solamente los valores de precipitación mayores 0.1 mm y excluyendo los valores más altos de cada serie. Es importante mencionar que el cálculo de los percentiles para todas las series se realizó solamente para el periodo 2010-2015, ya que es el periodo con el que se cuenta información. La comparación de estos valores se realizó a través de gráficos de dispersión, ploteando los datos observados vs cada uno de los métodos para todas las estaciones agrupados convenientemente en sectores tal y como se muestra en la figura 5.

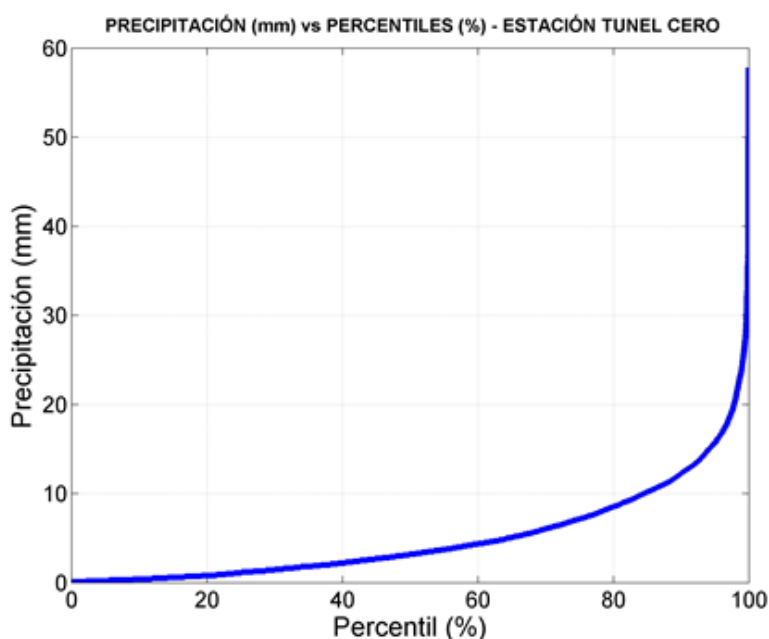


Fig. 4. Ejemplo de estimación de percentiles con datos observados

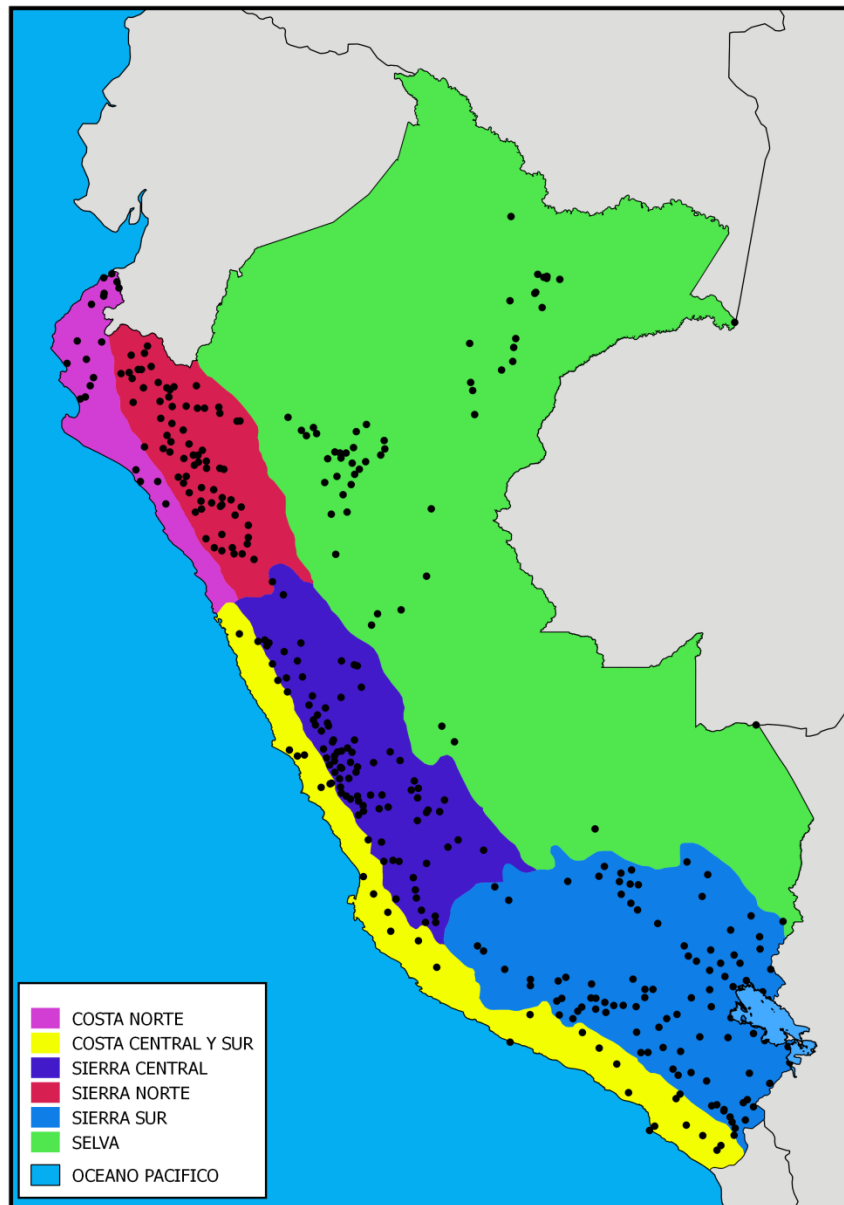


Figura 5. Agrupación de estaciones por sectores para la validación de extremos de precipitación

IV. RESULTADOS

4.1. Resultados de la corrección de bias.

La figura 6 muestra los promedios de los acumulados trimestrales de precipitación en el periodo 2010-2015 para cada serie de datos: observado, ETA/SENAMHI (Día 1) sin corregir y de las series corregidas por cada uno de las metodologías planteadas. Se aprecia una fuerte sobreestimación del modelo ETA/SENAMHI sin corregir en la vertiente oriental de los andes y en la selva, principalmente en el trimestre Diciembre-Enero-Febrero, que corresponde a los meses con mayor humedad y calentamiento. Los valores de precipitación del modelo pueden alcanzar los 1000 mm acumulados, en estaciones de Junín y Huancavelica, donde según los datos observados, llegan en realidad a 400 mm. En el trimestre JJA la sobreestimación es menor; sin embargo, este es más notorio en estaciones cercanas a Iquitos. Esta fuerte sobreestimación del modelo ETA/SENAMHI en dichas regiones puede deberse al esquema de convección de Kain-Fritsch que deja la capa excesivamente saturada y además, produce fuertes mezclas de la nube con su entorno, dando una capa más húmeda e inestable con alto potencial de generar precipitaciones intensas.

Estas fuertes precipitaciones dado por el modelo ETA/SENAMHI es reducido por los cuatro métodos de corrección. Particularmente, las altas cantidades de precipitación en la sierra oriental son reducidas notablemente y los valores son bien cercanos a los datos observados. Aparentemente, los resultados de precipitación trimestral promedio de los cuatro métodos son bien similares entre sí; sin embargo, se observa una ligera sobreestimación de la selva norte con el método gpQM siendo más marcado en el trimestre Diciembre-Enero-Febrero.

La figura 7 muestra los mismos resultados pero usando los pronósticos correspondientes al Día 2. Se observa similitud en la sobreestimación del modelo sin corregir para la vertiente oriental de la cordillera y en la selva usando el modelo ETA/SENAMHI, los cuales son posteriormente ajustados y más cercanos a los observados usando los método de corrección.

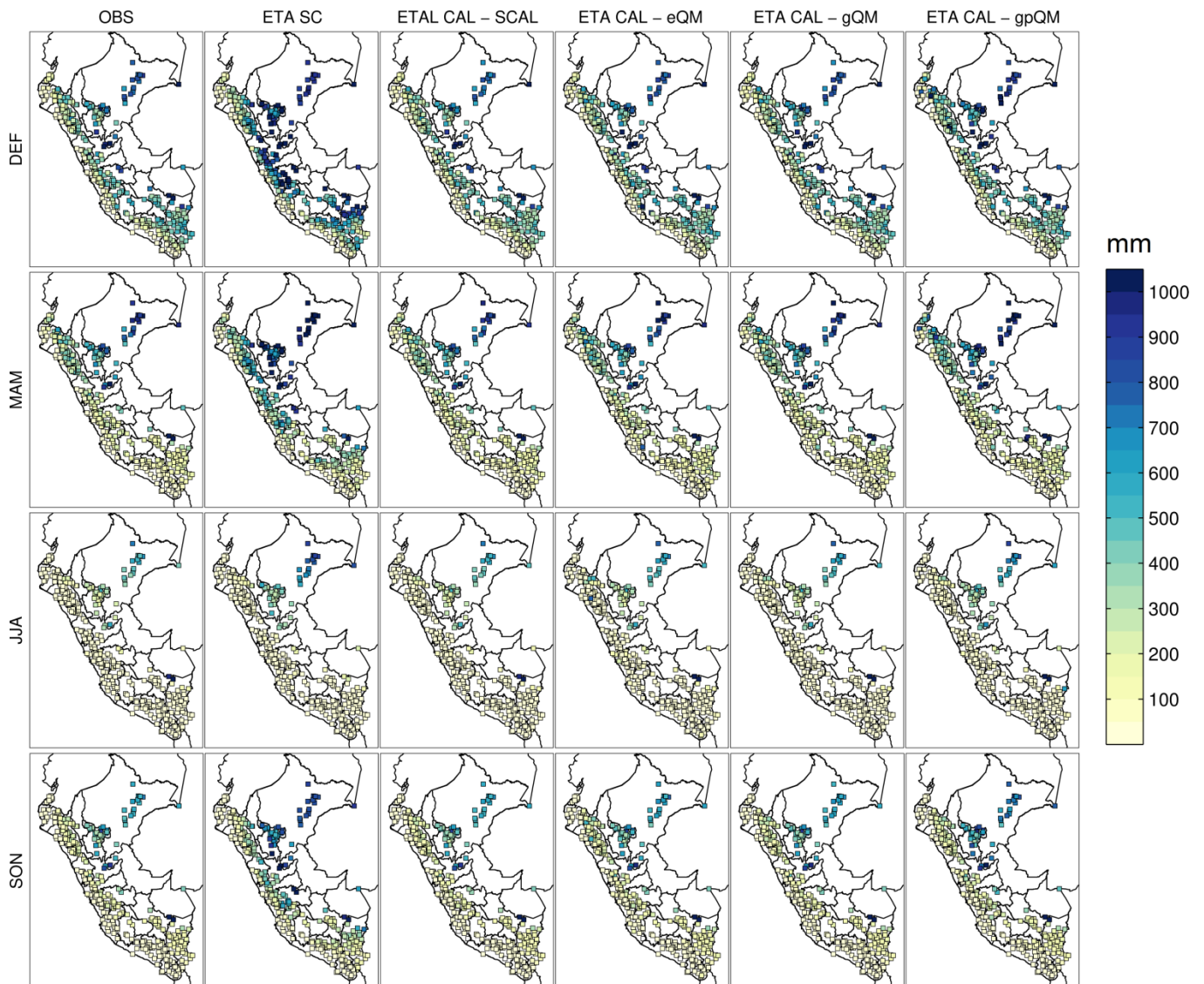


Figura 6. Promedio de la precipitación acumulada trimestral para el periodo 2010-2015 de los datos observados (obs), datos del eta/senamhi sin calibrar (ETA SC) y de los métodos de corrección: SCAL, eQM, gQM, gpQM, correspondientes al día 1 de pronósticos

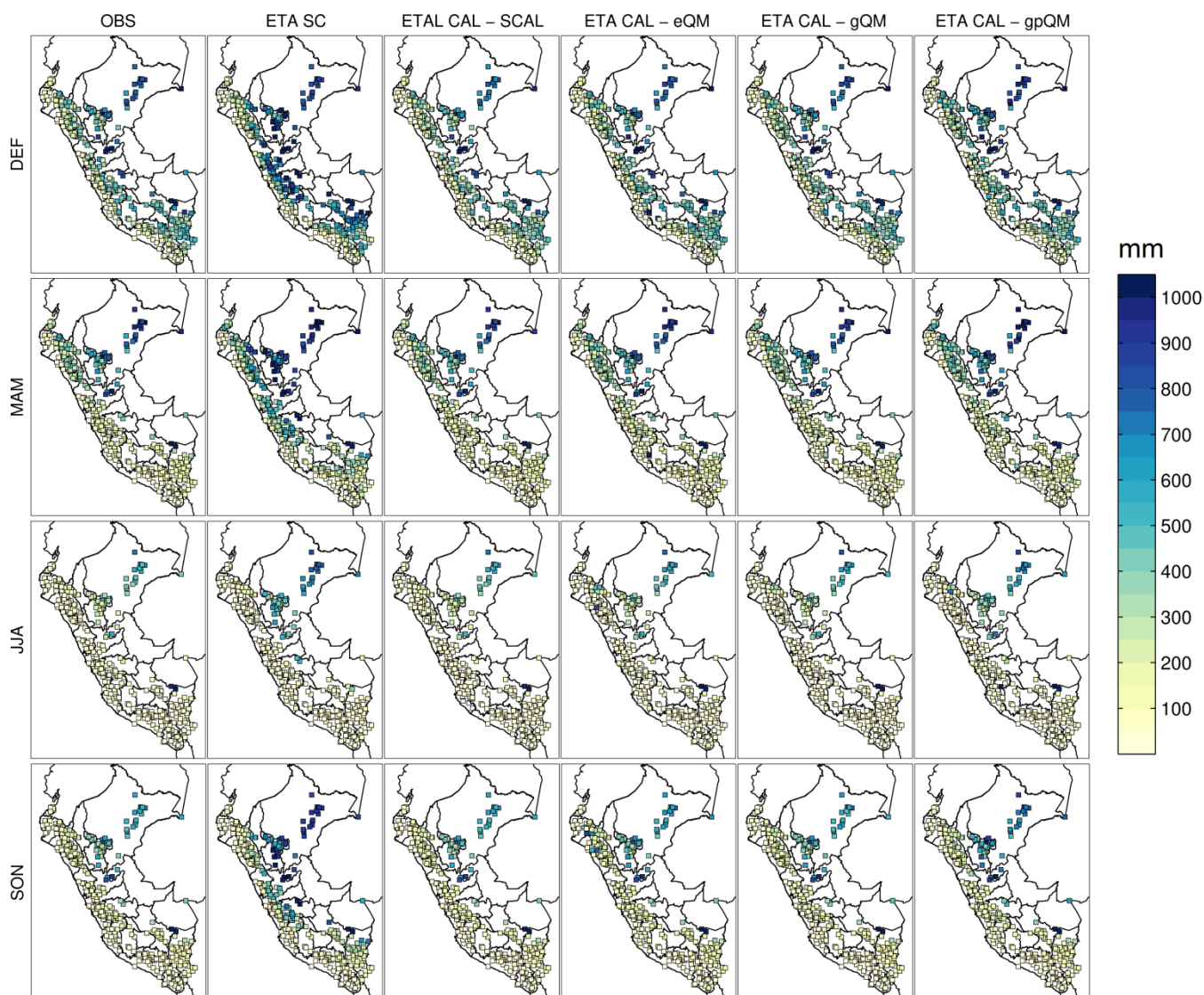


Figura 7. Igual que figura 3 para el día 2 de pronóstico del modelo ETA/SENAMHI y de las series corregidas.

Para el análisis a nivel temporal se seleccionaron un grupo de nueve estaciones distribuidos en diferentes regiones, con el fin de comparar sus series temporales observadas con los del ETA/SENAMHI sin corregir y corregidos con el método Scaling. Cabe recordar que las series correspondientes al Scaling son las series construidos juntando los cinco periodos de validación que se seleccionaron para el método de validación cruzada. En todas las series, se observa sobreestimación por parte del ETA/SENAMHI sin corrección, a excepción de la estación El Salto (figura 8a), donde por el contrario, existe una subestimación. Para los dos días de pronóstico se aprecia que las series del método Scaling se aproximan mejor a los datos observados. Este ajuste se resalta más en la estación de Punta Ático el cual está ubicado en la costa

sur, donde los acumulados mensuales generalmente no superan 1 mm; sin embargo, ETA/SENAMHI presenta acumulados de 5 mm a 6 mm, mientras que las series corregidas están más cercano a 0 siendo estas más coherentes. Es preciso señalar que a pesar de que las series ajustados son más cercanas a las observadas aún siguen habiendo ciertos valores puntuales sobreestimados considerablemente tal y como sucede en la estación de El Salto. Esto último puede ser a causa de un sobreajuste de los coeficientes estimados de la estación por el hecho de que hayan varios días sin lluvias y ciertos días con lluvias aisladas de fuerte intensidad, lo cual influye en que el coeficiente sea mucho mayor, y genere resultados exagerados cuando el modelo presente pequeñas lluvias.

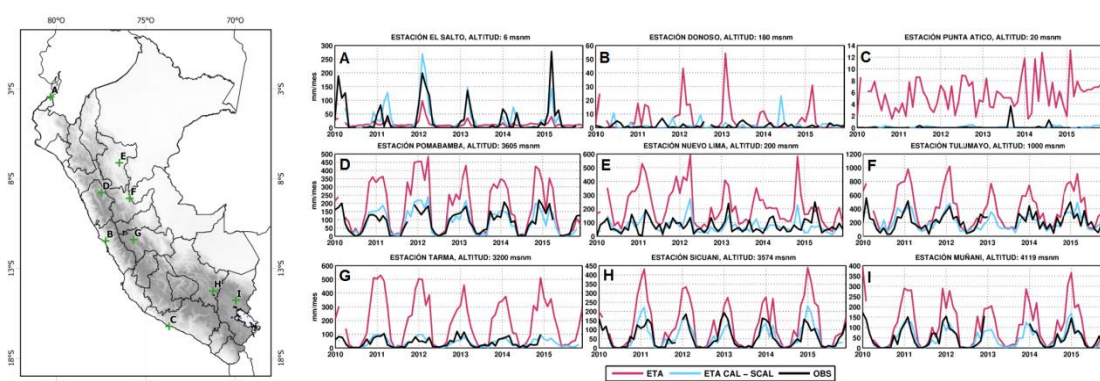


Figura 8. Series de precipitación acumulados mensual para el periodo 2010-2015 de datos observados (línea negra) datos del modelo ETA/SENAMHI Día 1 (línea roja) y corregido con el método SCAL (línea celeste).

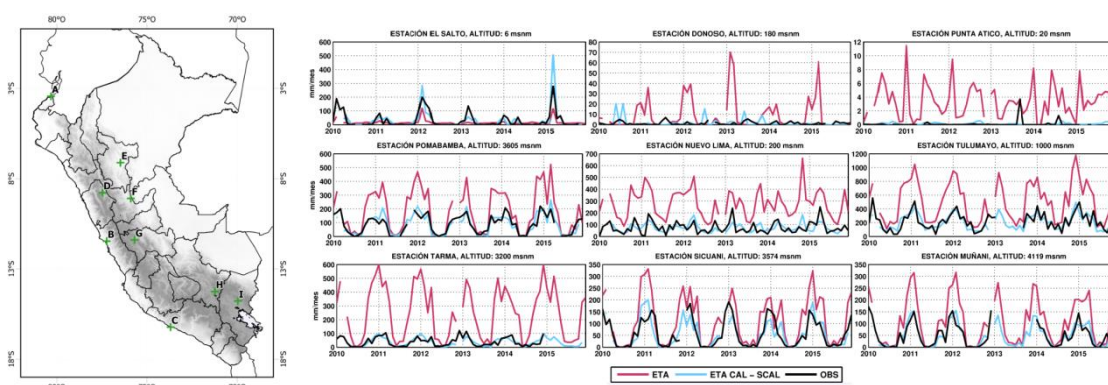


Figura 9. Igual que figura 5 para el día 2 de pronóstico del modelo ETA/SENAMHI y de las series corregidas.

4.2. Análisis de la eficiencia por cada método de corrección

La distribución espacial de los bias para los datos del modelo ETA/SENAMHI sin corregir y de los métodos de corrección son mostrados en las figuras 10 y 11 para los pronósticos correspondientes al día 1 y día 2 respectivamente. Se aprecia que el modelo ETA/SENAMHI sin corregir presenta un error sistemático positivo para toda la parte oriental de la cordillera y en la selva, lo cual guarda coherencia con los resultados de precipitación trimestral promedio. Además, notamos que existe un grupo de estaciones pertenecientes a la sierra norte y altiplano con bias negativos, tanto en el día 1 como día 2 de pronóstico. Mayores bias fueron encontrados en el trimestre Diciembre-Enero-Febrero, en el cual se tiene el valor más alto que es de 14 mm, mientras que el más bajo se obtiene en el trimestre Junio-Julio-Agosto con -10 mm.

Los resultados con los métodos corregidos evidencian una fuerte reducción de bias, siendo esta más notable para el método Scaling en el cual los valores varían de -0.26 a 0.43. De igual forma, el método eQM y gQM confirman fuerte reducción de bias a excepción de ciertas estaciones cuya sobreestimación todavía prosiguen. Diferencias importantes se puede observar del método gpQM con los demás métodos, ya que este indica bias positivos principalmente en estaciones de la selva y sierra norte, y para los cuatro trimestres del año. Hay que considerar que el qpQM utiliza un ajuste a la distribución GPD para valores de precipitación por encima del percentil 95, lo cual puede desencadenar una exageración en los valores extremos de precipitación y por ende mayores valores de bias en los resultados.

Otra forma de apreciar dichos resultados, es mediante la figura 12 donde se confirma el bias cercano a cero que presenta el método Scaling mientras que el método gpQM presenta valores mayores a 2mm para gran cantidad de estaciones y principalmente en los meses Diciembre-Enero-Febrero y Marzo-Abril-Mayo. En menor medida, los métodos eQM y gQM presentan algunos bias ligeramente positivos (menores a 4 mm en su mayoría).

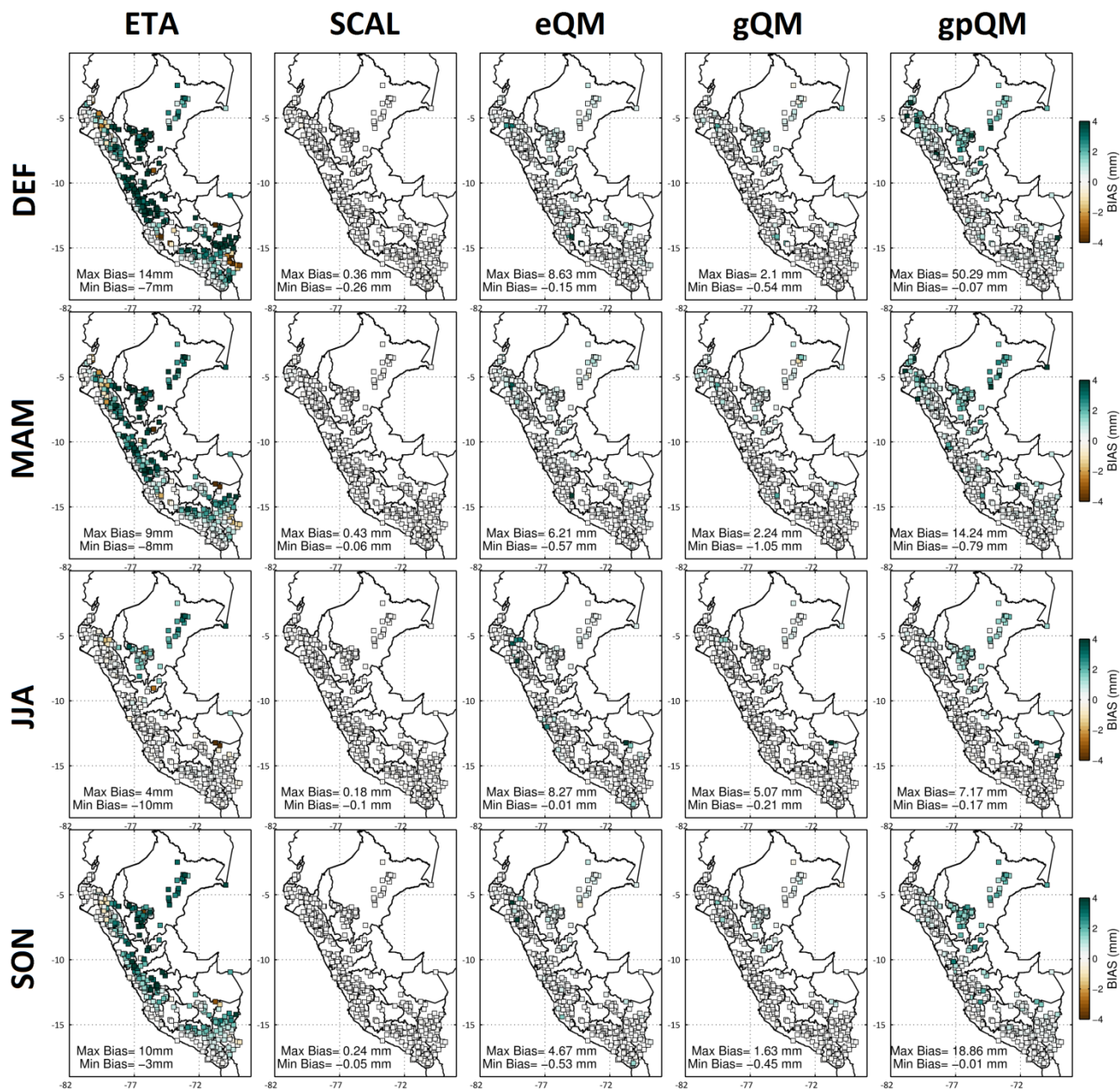


Figura 10. Bias trimestrales para datos del modelo ETA/SENAMHI correspondientes al día 1 y datos corregidos con distintas metodologías.

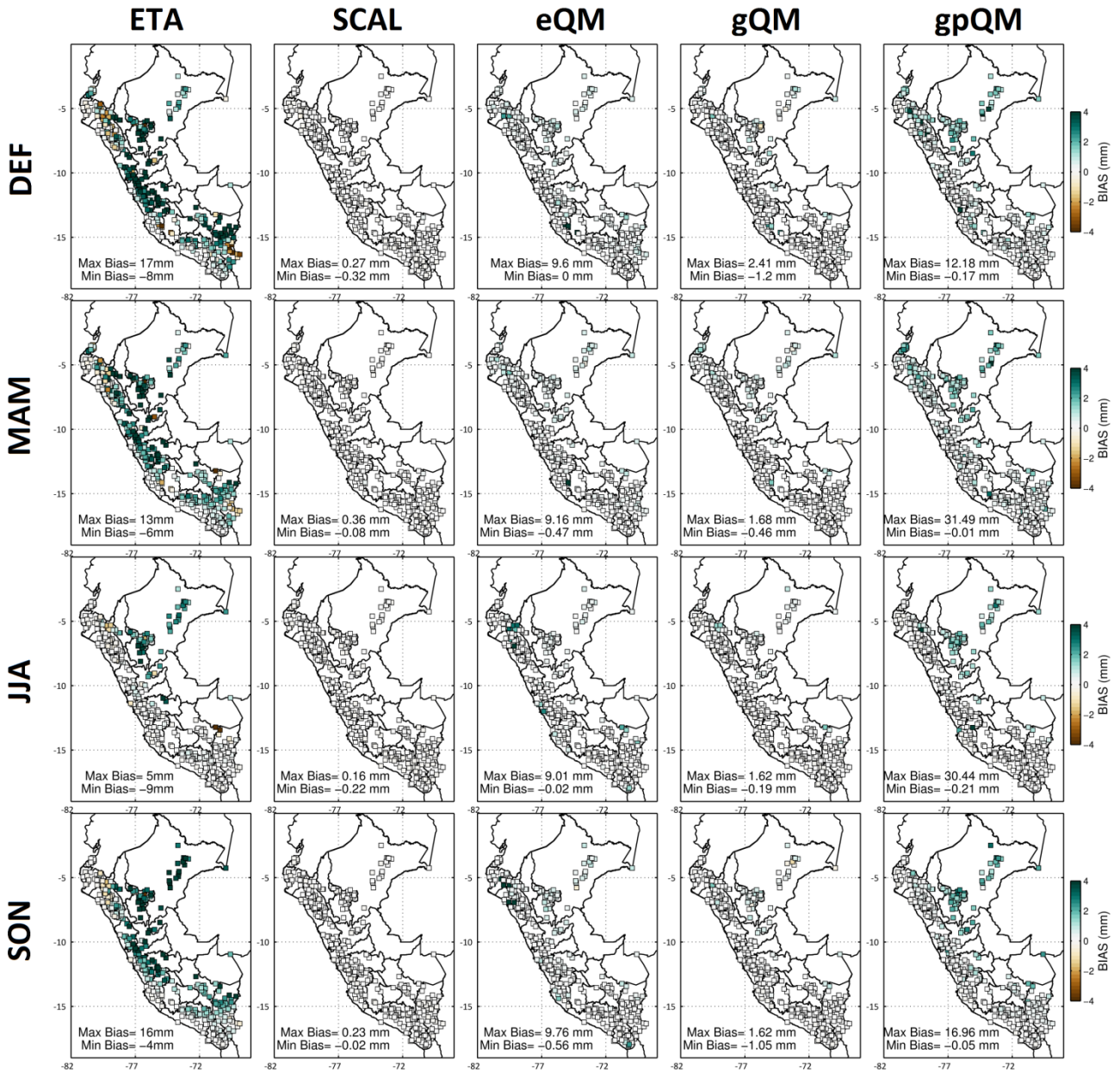


Figura 11. Igual que figura 7 pero con datos del modelo ETA/SENAMHI correspondientes al día 2.

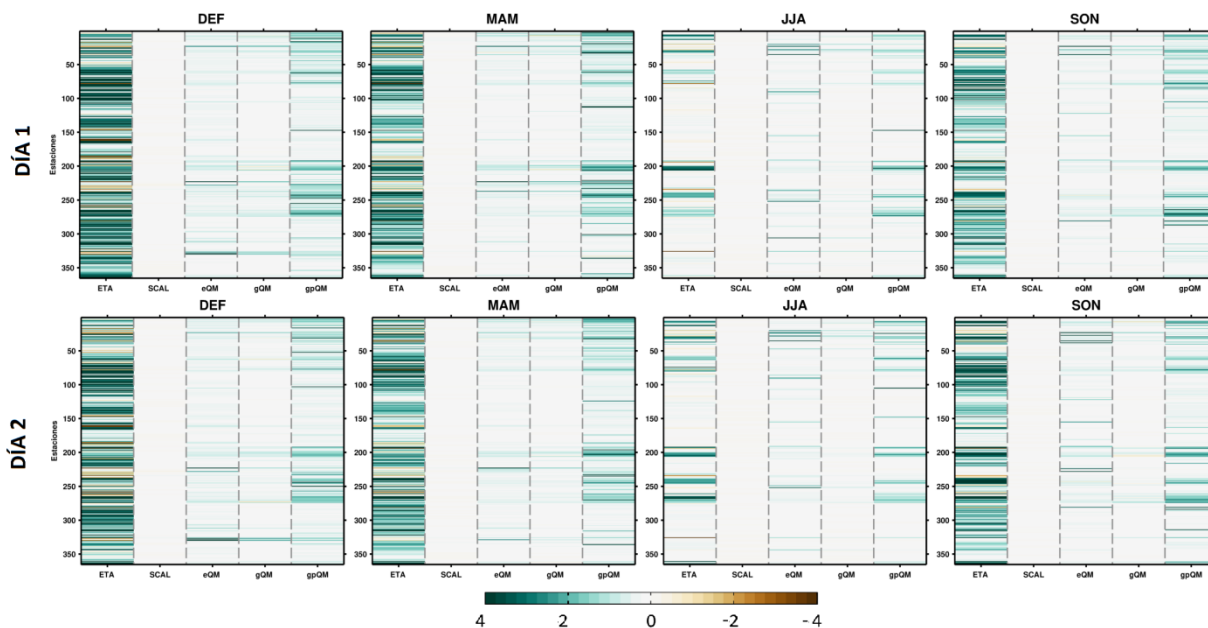


Figura 12. Bias trimestrales correspondientes al día 1 (arriba) y día 2 (abajo) para las 365 estaciones (filas) y los datos del ETA/SENAMHI sin corregir y corregidos con todos los métodos (columnas).

Los bias calculados para todas las estaciones fueron graficados a través de boxplots en la figura 13. De todos los métodos de corrección el qpQM presenta mayores cantidades de outliers los cuales corresponden a las estaciones de la selva y sierra norte tal y como se observó en los mapas de las figuras 8 y 9. Estos valores anómalos de bias son mayores a 10 e inclusive llega a más de 50 para una estación en el Día 1, indicando una sobreestimación en demasía por parte de este método. Dichos outliers ponen en evidencia la baja eficiencia de esta técnica al ajustar las distribuciones extremas de precipitación a la distribución GPD. Resultados más óptimos se obtuvieron con el eQM y gQM, teniendo esta última menores cantidades de outliers. Los mejores resultados de reducción de bias los presenta el método Scaling ya que sus valores son bien cercanos a cero.

Se puede evidenciar una diferencia entre los pronósticos del día 1 y día 2 para cada método. Con respecto a los datos del ETA/SENAMHI, se aprecia un ligero incremento de los bias en el pronóstico del día 2. En el método Scaling y eQM, el día 2 presenta mayores cantidades de outliers y con incrementos a comparación del día 1. En el método gQM no se presenta una diferencia

marcada entre los días de pronóstico. Por último, en el método gpQM la concentración de la mayor parte de los outliers se encuentra en el rango de 0-20 mm en el día 1, y entre 10-35 mm para el día 2, resaltando de esta forma un incremento en los bias para este pronóstico.

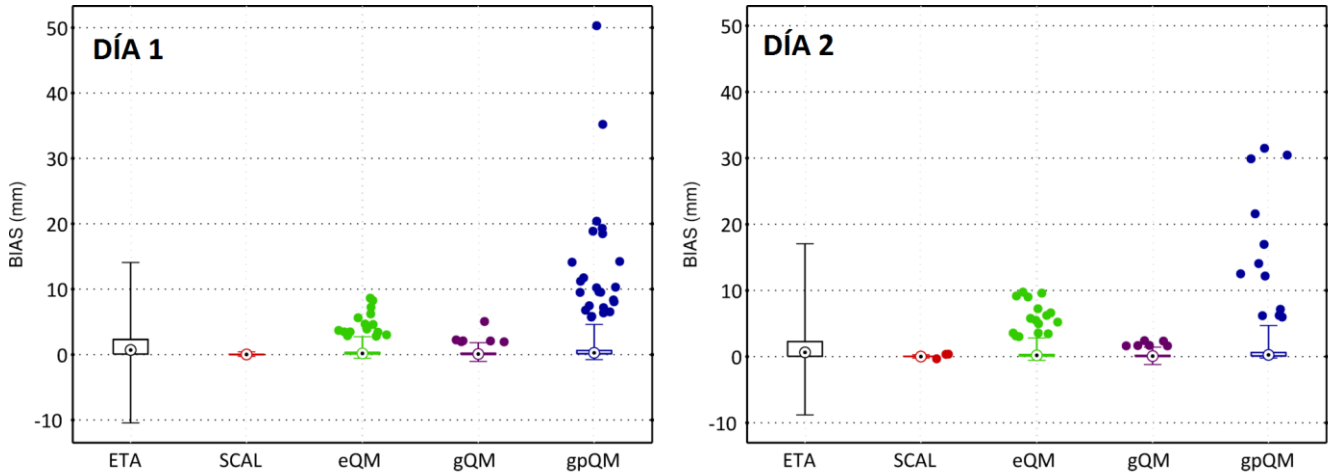


Figura 13. Boxplots de bias del ETA/SENAMHI y para los métodos de corrección, correspondientes al día 1 (izquierda) y día 2 (derecha) de pronóstico.

4.3. Análisis de extremos de precipitación

Una mayor verificación de las metodologías de corrección consiste en determinar no solo si el bias ha sido reducido, sino también como son simulados los extremos de precipitación después de la corrección. Esto se realizó mediante gráficos de dispersión agrupando las estaciones por regiones (figura 5) con datos observados vs. los pronóstico del ETA/SENAMHI y de cada uno de los métodos de corrección.

Los resultados correspondientes al ETA/SENAMHI muestran nubes de puntos con correlaciones que van de 0 a 0.75, presentándose los valores más bajos en la sierra centro y selva, y para los dos días de pronóstico. Cabe resaltar que los extremos de precipitación no son sobreestimados en la selva a diferencia de sus bias, lo cual sugiere que el modelo presenta una sobreestimación de la precipitación como consecuencia posiblemente de una mayor frecuencia o con precipitaciones de intensidades medias.

Con respecto a los métodos de corrección, se aprecia que en el método Scaling existe mejora en las correlaciones a comparación del ETA/SENAMHI para todas las regiones, a excepción de la costa centro-sur y sierra norte donde el desempeño es mucho menor. Cabe resaltar que a pesar de que el bias del modelo se reduce notablemente con esta técnica de corrección, las correlaciones de los extremos de precipitación no presentan mejora importante por el hecho de que el método Scaling no cambia la variabilidad de los datos siendo una desventaja.

Mejores resultados se presentan con eQM, en el cual las correlaciones mejoran notablemente, teniendo valores por encima de 0.7 y 0.99 como el valor más alto. La razón de que en este método se obtengan correlaciones bien altas, se debe a la naturaleza de esta técnica en ajustar los cuantiles de los datos del modelo a los de los observados y resultando de esta forma que los extremos de precipitación sean mejores representados. Es importante mencionar que a pesar de que eQM no sea el mejor de los métodos en reducir el error sistemático, tiene una muy buena performance en representar los extremos.

Valores de correlación ligeramente menores al eQM se obtuvieron con el método gQM; sin embargo, siguen siendo óptimas ya que los estos son mayores a 0.73, a excepción de la costa centro-sur que presenta una correlación de 0.22 para el percentil 90 en el día 1; sin embargo, esta mejora para el día 2. El método gQM se caracteriza por ajustar las distribuciones del modelo a una distribución paramétrica, en este caso a la distribución Gamma. Esto da a entender que hay una buena representación usando esta distribución. Sin embargo; los valores de percentil 99 para la selva son sobreestimados llegando inclusive a 200 mm en algunos casos.

En el método gpQM los valores de percentiles para la costa norte no guardan coherencia con lo observado, habiendo algunos valores que inclusive sobrepasan los 200 mm diarios. Las correlaciones en general, son más bajas a comparación de las otras variantes de quantile mapping, pero tiene mejor representación de extremos para la selva.

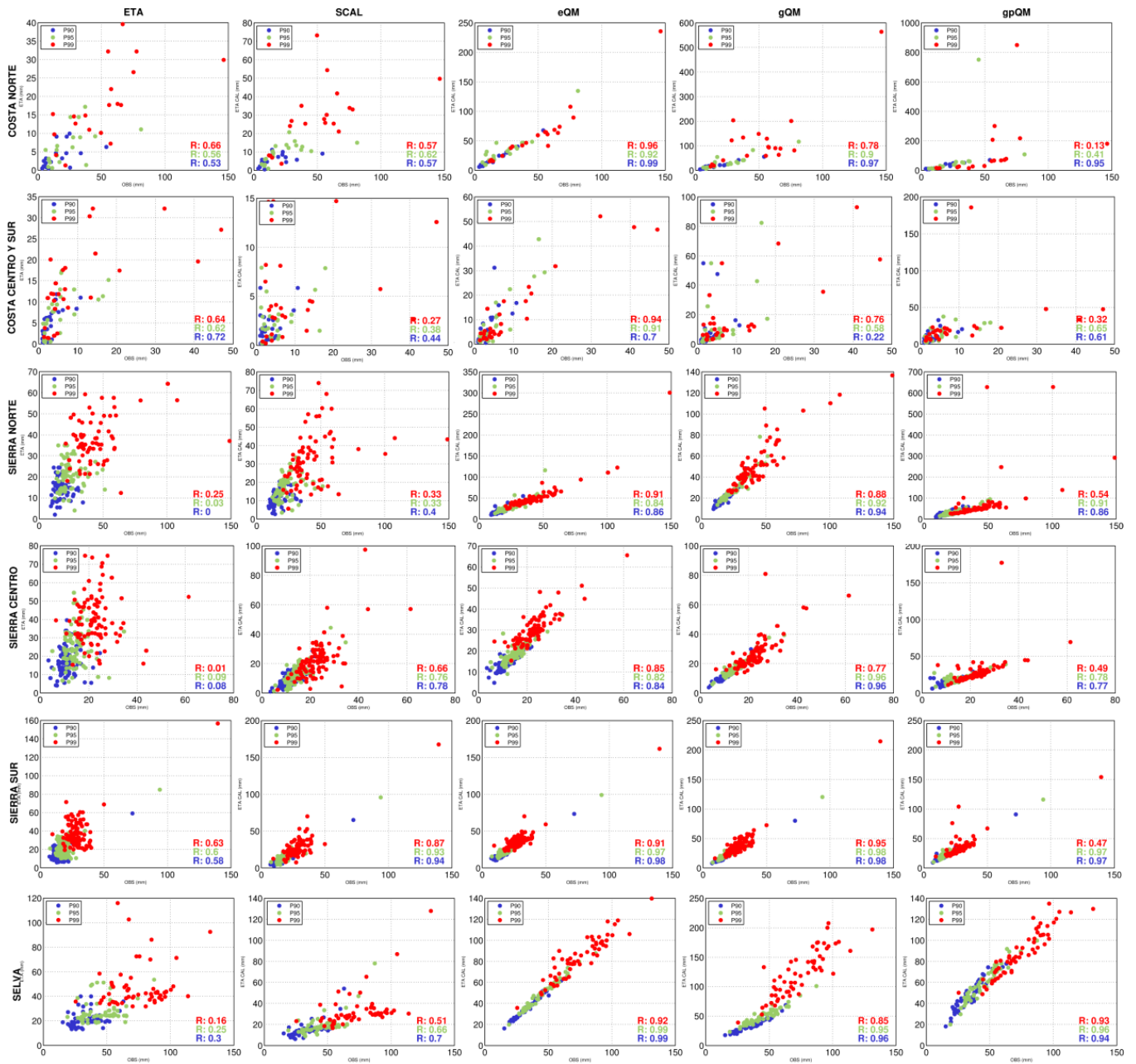


Figura 14. Gráficos de dispersión y correlaciones de la data observada vs datos del día 1 del ETA/SENAMHI, SCAL, eQM, gQM, gpQM (columnas) y para cada una de los sectores (filas).

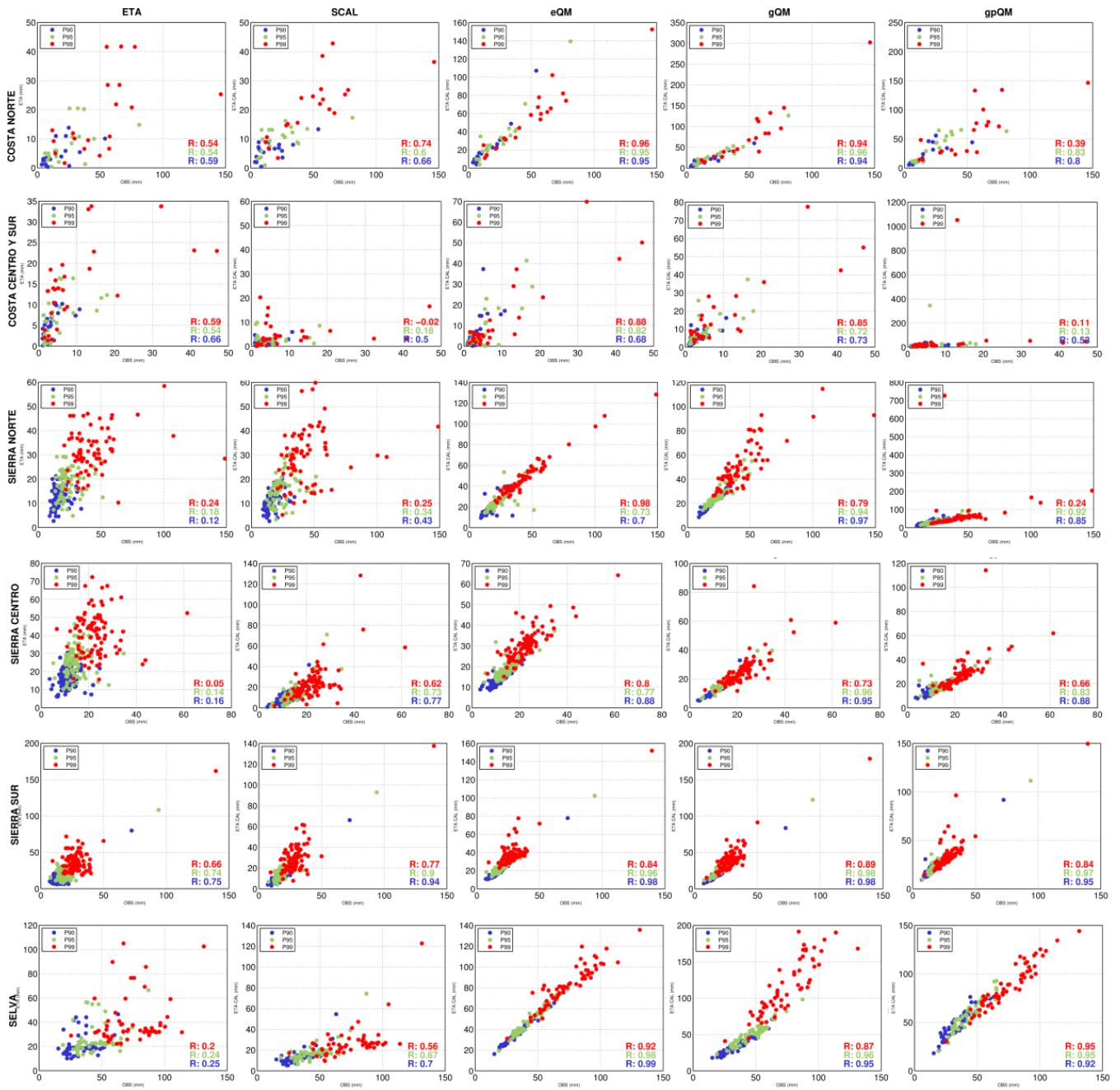


Figura 14. Igual que figura 13 para el día 2 de pronóstico del modelo ETA/SENAMHI y de las series corregidas.

V. CONCLUSIONES

- Existe una variación estacional del error sistemático del modelo ETA/SENAMHI que estaría asociado a un bajo desempeño del modelo en simular procesos de convección y humedad. Mayores valores de bias se encontraron principalmente en la vertiente oriental de los andes y en la selva para los meses de Diciembre-Enero-Febrero.
- El método Scaling fue el que mejor rendimiento presentó reduciendo el bias; sin embargo, tuvo menor habilidad para simular los extremos a consecuencia de que no modifica la variabilidad. Por otro lado, el método eQM pudo representar con más eficiencia los extremos debido al ajuste de cuantiles en el que está basado, pero no corrigió adecuadamente el bias para ciertas estaciones. Resultados óptimos se también obtuvieron con el método gQM el cual redujo el bias para la mayor parte de las estaciones y tuvo buena performance al representar los extremos, pero con correlaciones ligeramente menores a las del eQM. Por otro lado, el método qpQM es el que menor desempeño presentó ya que no pudo reducir el bias para estaciones de la selva y parte de la sierra norte.
- Comparando los pronóstico tanto para el día 1 como para el día 2 se concluye que ambos presentaron resultados similares en la reducción de bias; sin embargo, el pronóstico del día 1 presento ligeramente un mejor rendimiento que el del día 2. En cuanto a la representación de los extremos, el día 2 resulto más eficientes para algunas regiones y algunos métodos de corrección, por lo que no se puede dar una conclusión definitiva de cuál es más eficiente.

VI. RECOMENDACIONES

- Dado los resultados obtenidos en base a los métodos de calibración usados en este estudio, se debe tener en cuenta que estos se han obtenido usando información de cinco años debido a la disponibilidad de información del modelo ETA/SENAMHI. Sin embargo, es recomendable en lo posible trabajar con series más largas de datos para tener una mayor precisión de los resultados. Por lo tanto, se debe aplicar constantemente estas metodologías y validando a medida que se tenga mayores años de información.
- Dado que una técnica puede ser eficiente en un aspecto tal como en la reducción de bias, pero menos eficiente en otro como en el pronóstico de los extremos y además dependiente también de las regiones de estudio; se recomienda no solo aplicar una técnica, sino dos o más a la vez, considerando las fortalezas que presenta cada metodología de acuerdo a las necesidades que se tengan.
- Se recomienda seguir probando otras metodologías de corrección así como también otros modelos numéricos con el fin de seguir generando resultados muchos más robustos.

VII. BIBLIOGRAFÍA

- Gutjahr and Heinemann (2013). Comparing precipitation bias correction methods for high-resolution regional climate simulations using COSMO-CLM – Effects on extreme values and climate change signal.
- Henning W. Rust, Andy Richling, Edmund Meredith, Madlen Fischer, Christos Vagenas, Christopher Kadow and Uwe Ulbrich (2016). Freie Universität Berlin. A plug-in for data extraction and conversion developed within and for BINGO
- Kain, J. S. y J.M. Fritsch (1990). A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784-2802.
- Lenderink, G., Buishand, A., and van Deursen, W (2007). Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, *Hydrol. Earth Syst. Sci.*, **11**, 1145–1159, doi:10.5194/hess-11-1145-2007.
- Piani C, Weedon GP, Best M, Gomes SM, Viterbo P, Hagemann S, Haerter JO (2010b) Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J Hydrol* 395:199–215. doi:10.1016/j.jhydrol. 2010.10.024.
- Alfaro, L, (2014). Estimación de umbrales de precipitaciones extremas para la emisión de avisos meteorológicos. Nota Técnica 001-SENAMHI-DGM-2014. Servicio Nacional de Meteorología e Hidrología del Perú.
- THE COMET PROGRAM (2010). ¿Cómo producen los modelos la precipitación en las nubes?. Disponible en: https://www.meted.ucar.edu/nwp/model_precipandclouds_es/print.htm
- Wetterhall, F., Pappenberger, F., He, Y., Freer, J., and Cloke, H. L. (2012) Conditioning model output statistics of regional climate model precipitation on circulation patterns, *Nonlin. Processes Geophys.*, **19**, 623-633, doi:10.5194/npg-19-623-2012, 2012.

VIII. ANEXOS

8.1. Correlación de spearman y root mean squared error (RMSE) del modelo ETA/SENAMHI

Las correlaciones de los datos observados vs el pronóstico del modelo ETA/SENAMHI (figura 15) indican que estos son mayores en la zona andina y en algunos puntos superan 0.5. Estacionalmente se observan mayores correlaciones en el trimestre de Marzo-Abril-Mayo. En las estaciones de costa los valores son cercanos a cero corroborando la baja eficiencia que presenta este modelo en simular las precipitaciones para esta región. Así mismo, los correlaciones del pronóstico del día 2 son menores a comparación del día 1, lo cual puede ser también verificado en los boxplots de la figura 16a.

Con respecto al RMSE se aprecia que los errores son mayores en la parte oriental de la region andina y en la selva, lo cual guarda coherencia con el bias mostrado anteriormente. Los valores más altos de RMSE superan los 15 mm principalmente en el trimestre Diciembre-Enero-Febrero, debido a la fuerte sobreestimación de humedad en la temporada lluviosa; mientras que los valores más bajos se dan en el trimestre Junio-Julio-Agosto. Resaltan algunas estaciones que superan inclusive los 40 mm de error medio tal y como se puede en el boxplot de la figura 16b.

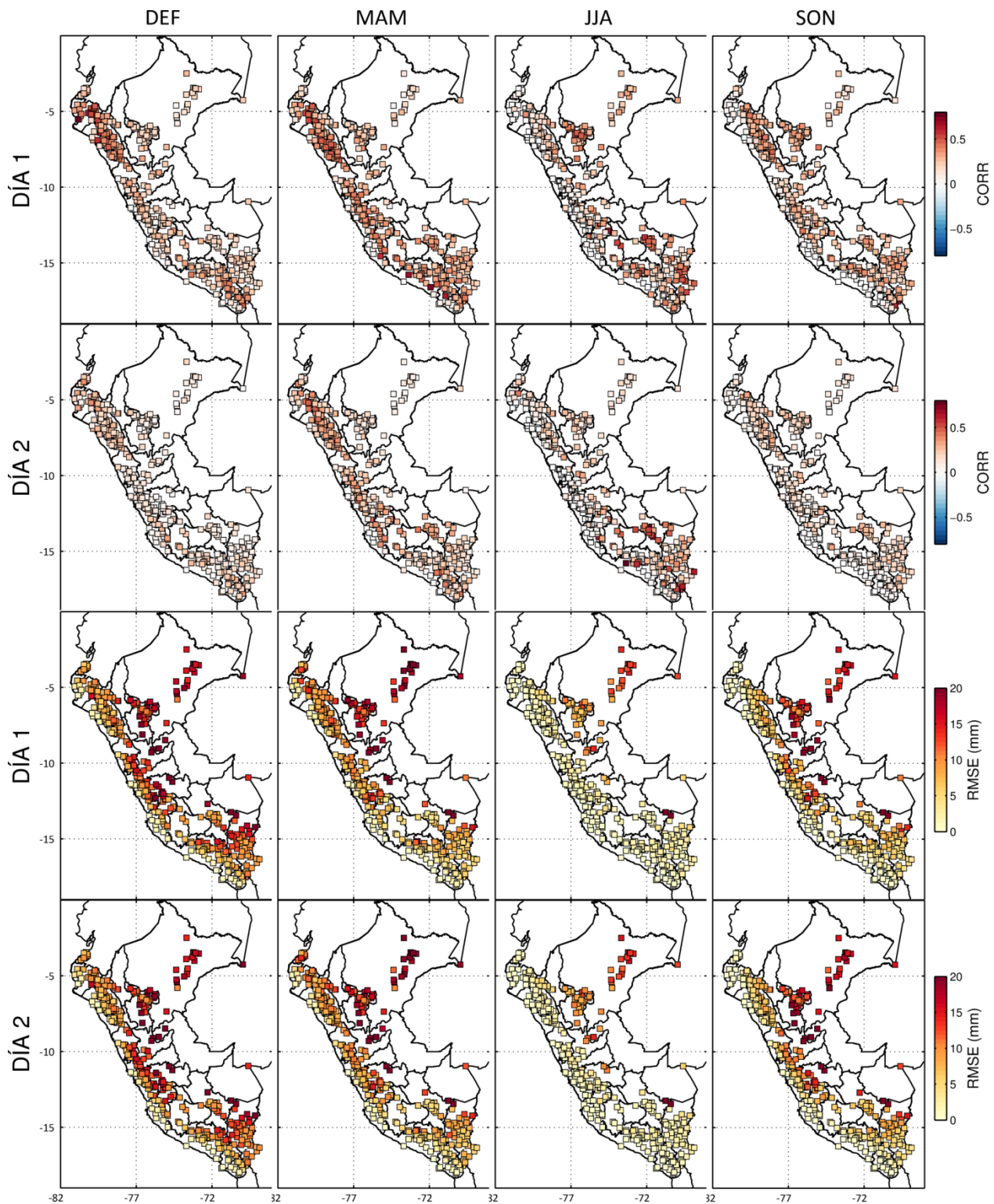


Figura 15. Las dos filas superiores representan la correlación de la data observada vs los pronósticos del modelo ETA/SENAMHI, y las dos filas inferiores representan el RMSE.

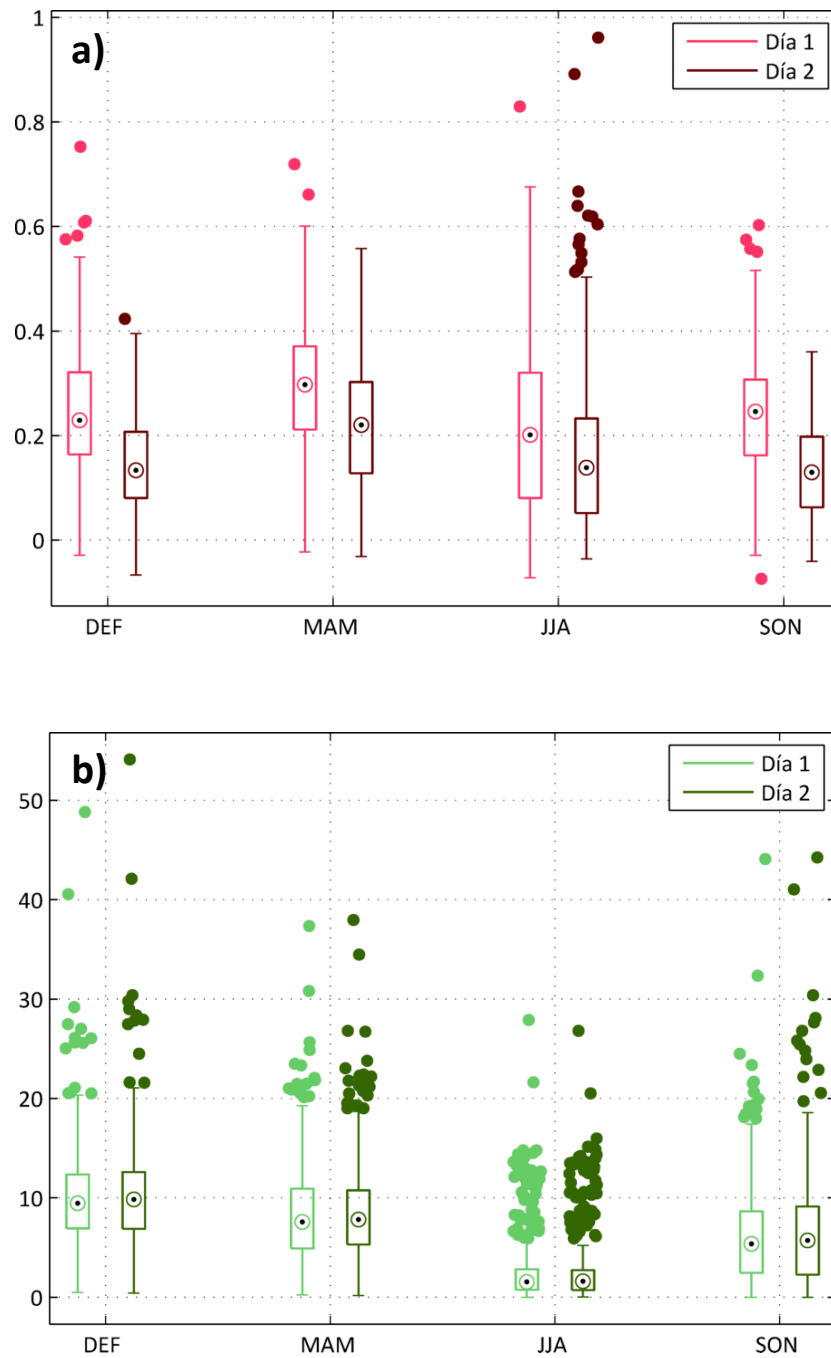


Figura 16. (a) boxplot de correlaciones de datos observados vs pronósticos
ETA/SENAMHI (b) boxplot de rmse del modelo ETA/SENAMHI

8.2. Series diarias de precipitación

Las series diarias de precipitación de la figura 17 y 18 indican una mejora notable de los acumulados diarios que genera el modelo ETA/SENAMHI usando el método de corrección Scaling, reproduciendo parcialmente la tendencia de la variación diaria para la mayoría de las estaciones. Sin embargo, en Nuevo Lima y Tulumayo es menos eficiente en representar las frecuencias y mayores cantidades de precipitaciones, lo cual es consistente con los resultados mostrados en el análisis de pronósticos de extremos de precipitación.

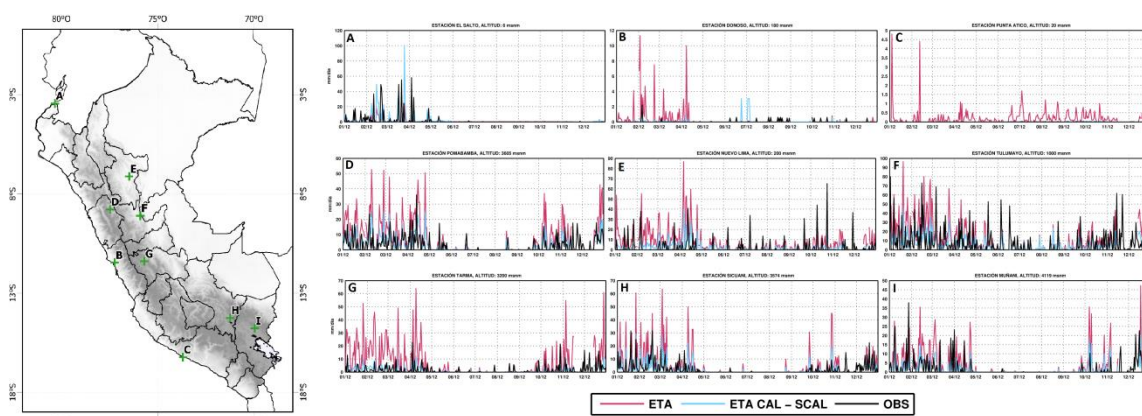


Figura 17. Series de precipitación diaria para el año 2012 de datos observados (línea negra) datos del modelo ETA/SENAMHI Día 1 (línea roja) y corregido con el método SCAL (línea celeste).

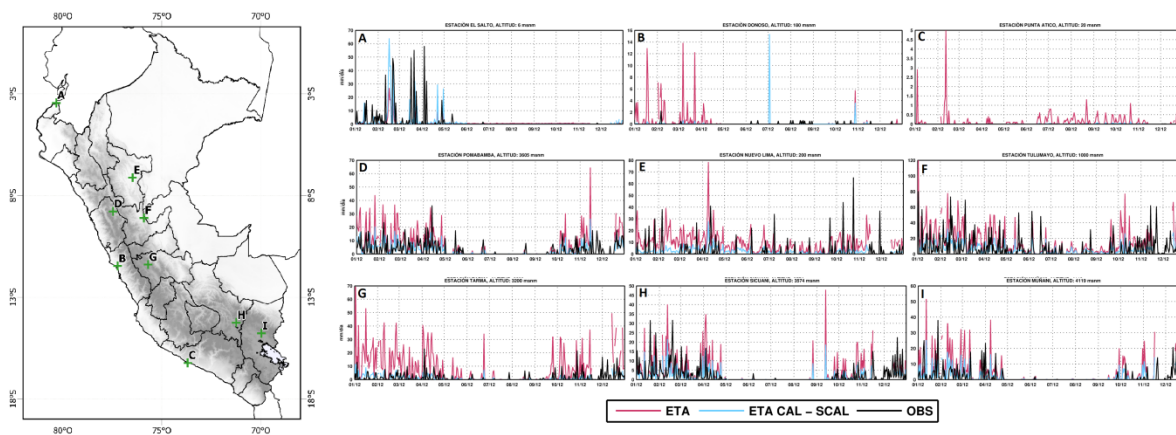


Figura 18. Igual que la figura 17 pero correspondiente al pronóstico del día 2 del modelo ETA/SENAMHI.

Informe preparado por:

Met. Gustavo De la Cruz Montalvo
Especialista de la Subdirección de Modelamiento Numérico
SENAMHI-PERÚ