

UNIVERSIDAD NACIONAL AGRARIA

**LA MOLINA
FACULTAD DE CIENCIAS**



**“PRONÓSTICO CUANTITATIVO DE PRECIPITACIÓN EN
TÉRMINOS DE PROBABILIDADES”**

Presentado por:

Gustavo Adolfo De la Cruz Montalvo

Tesis para Optar el Título Profesional de:

INGENIERO METEORÓLOGO

Lima-Perú

2016

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA**

FACULTAD DE CIENCIAS

**“PRONÓSTICO CUANTITATIVO DE PRECIPITACIÓN EN
TÉRMINOS DE PROBABILIDADES”**

Presentado por:

Gustavo Adolfo De la Cruz Montalvo

Tesis para Optar por el Título Profesional de:

INGENIERO METEORÓLOGO

Sustentado y aprobado por el siguiente Jurado:

**Mg. Sc. Victoria Calle Montes
PRESIDENTE**

**Mestre. Eng. Esaúl Obregón Párraga
MIEMBRO**

**Ing. Franklin Unsihuay Tovar
MIEMBRO**

**Mg. Sc. Ever Menacho Casimiro
PATROCINADOR**

**Ing. Clara Oria Rojas
CO-PATROCINADORA**

DEDICATORIA

A mis padres Ena Montalvo y Alberto De la Cruz y a mi hermana Aida Villanueva por todo el apoyo incondicional que me brindaron a lo largo de todos mis estudios y ser mi motivación cada día.

AGRADECIMIENTOS

Muchas gracias a la Ing. Clara Oria por haberme impulsado en realizar la presente tesis y su constante apoyo y asesoría a lo largo de este proceso hasta el final. Gracias por estar pendiente del avance de esta investigación.

Muchas gracias al profesor Ever Menacho por la asesoría brindada y su constante ánimo. Sin sus consejos y apoyo hubiese sido difícil terminar esta tesis.

Miles de agradecimientos al Servicio Nacional de Meteorología e Hidrología por su colaboración con respecto a la información meteorológica. De igual forma, un agradecimiento especial a mis colegas y amigos de la Dirección de Meteorología Aplicada por sus valiosos consejos, en especial para el Fis. Gerardo Jácome y la Ing. Gabriela Rosas.

Quisiera agradecer también a todos mis amigos de la carrera que me motivaron constantemente a seguir avanzando con este trabajo y me brindaron bastante apoyo emocional. Un agradecimiento especial a Isabel Ramos por todo su apoyo y por estar pendiente de que terminara esta tesis.

Finalmente, quisiera agradecer también al jurado calificador de esta tesis por sus valiosos consejos y sugerencias que hicieron que esta tesis vaya mejorando con cada revisión.

ÍNDICE GENERAL

ÍNDICE DE ANEXOS	iv
ÍNDICE DE CUADROS	v
ÍNDICE DE FIGURAS	vi
RESUMEN	ix
ABSTRACT	x
I. INTRODUCCIÓN	1
II. REVISIÓN DE LA LITERATURA.....	3
2.1. VARIABILIDAD ESPACIO – TEMPORAL DE LA PRECIPITACIÓN.	3
2.1.1. Variabilidad interanual	7
2.2. PATRONES DE CIRCULACIÓN ATMOSFÉRICAS ASOCIADOS A LA PRECIPITACIÓN EN EL PERÚ	7
A. Sistemas atmosféricos predominantes durante el verano austral	8
B. Sistemas atmosféricos predominantes durante el invierno austral	11
2.3. MODELOS DE PREDICCIÓN NUMÉRICA DEL TIEMPO	13
2.4. MODELO REGIONAL ETA	16
2.5. POST- PROCESAMIENTO ESTADÍSTICO MOS Y SUS APLICACIONES EN EL PRONÓSTICO DEL TIEMPO.....	22
2.6. MODELOS ESTADÍSTICOS EMPLEADOS EN EL AJUSTE ESTADÍSTICO MOS. 23	
2.6.1. REGRESIÓN LINEAL MULTIPLE.....	23

2.6.2.	REGRESIÓN LOGÍSTICA.....	24
2.7.	APLICACIONES DE LA REGRESIÓN LOGÍSTICA EN EL PRONÓSTICO PROBABILÍSTICO DE LA PRECIPITACIÓN	26
III.	MATERIALES Y MÉTODOS.....	29
3.1.	ÁREA DE ESTUDIO	29
3.2.	MATERIALES REQUERIDOS PARA LA INVESTIGACIÓN.....	31
3.2.1.	INFORMACIÓN DE PRECIPITACIÓN OBSERVADA	31
3.2.2.	INFORMACIÓN GENERADA POR EL MODELO ETA-SENAMHI.....	34
3.2.3.	DATOS DE REANALYSIS ERA-INTERIM.....	36
3.2.4.	IMÁGENES DE SATÉLITE.....	37
3.2.5.	SOFTWARES	37
3.3.	ENFOQUE METODOLÓGICO.....	38
3.3.1.	METODOLOGÍA PARA EL PRONÓSTICO CUANTITATIVO DE LA PRECIPITACIÓN EN TÉRMINO DE PROBABILIDADES.	39
3.3.2.	ANÁLISIS DE LA EFICIENCIA DEL PRONÓSTICO PROBABILÍSTICO DE PRECIPITACIÓN.....	50
3.3.3.	APLICACIÓN DEL PRONÓSTICO PROBABILÍSTICO DIARIO DE PRECIPITACION EN UN CASO DE ESTUDIO.....	58
IV.	RESULTADOS Y DISCUSIONES.....	59
4.1.	RESULTADOS DE LA APLICACIÓN DE LA METODOLOGÍA PARA EL PRONÓSTICO PROBALÍSTICO DE LA PRECIPITACIÓN.....	59
4.1.1.	RESULTADOS DE SELECCIÓN DEL NÚMERO DE PREDICTORES ...	59
4.1.2.	PREDICTORES SELECCIONADOS PARA CADA ESTACIÓN.....	65
4.2.	ANÁLISIS DE LA EFICIENCIA DEL PRONÓSTICO PROBABILÍSTICO DIARIO DE PRECIPITACIÓN.	68
4.2.1.	RESULTADOS DEL ATRIBUTO DE FIABILIDAD	68

4.2.2.	RESULTADOS DEL ATRIBUTO DE PRECISIÓN	73
4.2.3.	RESULTADOS DEL ATRIBUTO DE HABILIDAD.....	76
4.2.4.	RESULTADOS DEL ATRIBUTO DE RESOLUCIÓN.	80
4.3.	APLICACIÓN DEL PRONÓSTICO PROBABILÍSTICO DE LA PRECIPITACIÓN EN UN EVENTO DE FUERTES PRECIPITACIONES.	82
4.3.1.	ANÁLISIS DE ENTORNO SINÓPTICO DEL EVENTO	82
4.3.2.	RESULTADOS DE LA APLICACIÓN DEL PRONÓSTICO PROBABILÍSTICO DE LA PRECIPITACIÓN PARA EL EVENTO DE LLUVIAS FUERTES.....	89
V.	CONCLUSIONES	97
VI.	RECOMENDACIONES	100
VII.	REFERENCIAS BIBLIOGRÁFICAS	101
VIII.	ANEXOS	113

ÍNDICE DE ANEXOS

ANEXO 1: Diagramas de fiabilidad para cada estación.....	113
ANEXO 2: Componente Fiabilidad del Brier Score dad del Brier Score	117
ANEXO 3: Componente Resolución del Brier Score.....	118
ANEXO 4: Resultados del Brier Score por estación	119
ANEXO 5: Resultados del Brier Skill Score por estación.....	120
ANEXO 6: Resultados del Brier Score a nivel estacional.	121
ANEXO 7: Resultados del Brier Skill Score a nivel estacional.	122
ANEXO 8: Precipitación diaria en 6 estaciones para el periodo 01-15 de Febrero del 2013	123
ANEXO 9: Imágenes satelitales del GOES 13 correspondiente al canal visible para el periodo 02-09 de Febrero del 2013.....	124
ANEXO 10: Relación de mezcla y vientos en 500 hPa para el periodo 02-09 de Febrero del 2013	124
ANEXO 11: Pronóstico de probabilidades de precipitación a 24 horas y ocurrencia observada de precipitación para el periodo 02-09 de Febrero del 2013.	126

ÍNDICE DE CUADROS

Cuadro 1: Características del modelo ETA SENAMHI	19
Cuadro 2: Productos de pronóstico del modelo ETA-SENAMHI (32 km) disponibles en http://www.senamhi.gob.pe/?p=0260	21
Cuadro 3: Productos de pronóstico del modelo ETA-SENAMHI (22 km) disponibles en http://www.senamhi.gob.pe/?p=0260	22
Cuadro 4: Estaciones usados para el pronóstico de probabilidad de precipitación	32
Cuadro 5: Conjunto de datos extraídos del modelo ETA - SENAMHI.....	35
Cuadro 6: Categorías de pronóstico probabilístico	41
Cuadro 7: Atributos analizados e indicadores estadísticos del desempeño del modelo	54
Cuadro 8: Tabla de contingencia para pronósticos binarios.	55
Cuadro 9: Nivel de presión y correlación de Spearman seleccionados para cada variable.....	63
Cuadro 10: Conjunto de predictores seleccionados para cada estación.....	67
Cuadro 11: Umbrales de probabilidad para la discriminación de eventos.....	82
Cuadro 12: Razón de aciertos y falsas alarmas correspondientes al periodo del 02-09 de Febrero del 2013.....	90

ÍNDICE DE FIGURAS

Figura 1: Precipitación acumulada durante el periodo lluvioso normal Setiembre – Mayo.	6
Figura 2: Climatología para enero de: (a) de vientos en 200 Hpa y (b) Presión atmosférica a nivel medio del mar.	10
Figura 3: Climatología para julio de (a) de vientos en 200 Hpa y (b) Presión atmosférica a nivel medio del mar.	12
Figura 4: Representación del mundo por modelos numéricos.	14
Figura 5: Representación de un modelo regional anidado a un modelo global.	15
Figura 6: Representación gráfica de la coordenada vertical ETA.	17
Figura 7: Dominios espaciales del modelo ETA-SENAMHI	20
Figura 8: Gráfico de un modelo de regresión logística simple.	25
Figura 9: Comparación de pronósticos de probabilidad de precipitación para Ithaca, usando regresión lineal (línea discontinua) y regresión logística (línea solida).	27
Figura 10: Mapa de ecorregiones del Perú.	30
Figura 11: Mapa con ubicación de estaciones meteorológicas usadas en el trabajo.	33
Figura 12. Esquema de extracción de variables predictores para el modelo ETA-SENAMHI	35
Figura 13: Esquema de la metodología de la presente investigación.	38
Figura 14: Ejemplo de diagrama de fiabilidad	51
Figura 15: Diagramas de fiabilidad de acuerdo a la calidad del pronóstico probabilístico	52
Figura 16: Dos ejemplos de curvas ROC para diferentes eventos.	56
Figura 17: Tipos de curva ROC según la resolución del pronóstico.	57
Figura 18: Correlación de Spearman Precipitación vs humedad relativa (Periodo DEFMA)	61
Figura 19: Correlación de Spearman de Precipitación vs viento zonal (periodo DEFMA)	61
Figura 20: Correlación de Spearman de Precipitación vs viento meridional para el periodo DEFMA	62

Figura 21: Correlación de Spearman de Precipitación vs temperatura para el periodo DEFMA	62
Figura 22: Frecuencia de predictores significativos al 95% de confianza para (a) estaciones de costa (b) estaciones de sierra y (c) estaciones de selva.	65
Figura 23: Diagramas de fiabilidad correspondientes a todo el periodo de estudio	70
Figura 24: Diagramas de fiabilidad correspondientes al periodo Diciembre-Enero-Febrero	71
Figura 25: Diagramas de fiabilidad correspondiente al periodo Junio- Julio- Agosto.....	72
Figura 26: Componente fiabilidad del Brier Score.	73
Figura 27: Resultados de Brier Score (atributo de precisión) para estaciones de (a) costa (b) sierra y (c) selva.....	74
Figura 28: Resultados de Brier Skill Score (atributo de habilidad) para estaciones de (a) costa (b) sierra y (c) selva.	77
Figura 29: Comparación Resultados del Brier Skill Score (atributo de habilidad) del modelo ETA-SENAMHI (izquierda) y MOS (derecha).....	79
Figura 30: Resultados de curvas ROC (atributo de resolución).....	81
Figura 31: Líneas de corriente en 200 para los días (a) 02 de febrero (b) 05 de febrero (c) 07 de febrero y (d) 09 de febrero del 2013	84
Figura 32: Líneas de corriente en 500 hPa para los días (a) 02 de febrero (b) 05 de febrero (c) 07 de febrero y (d) 09 de febrero del 2013	85
Figura 33: Flujo de razón de mezcla en 600 hPa para los días (a) 02 de febrero (b) 05 de febrero (c) 08 de febrero y (d) 09 de febrero del 2013	86
Figura 34: Imágenes de satélite del GOES 13 para el día de 04 de febrero del 2013 22: 30 UTC. (a) Imagen vapor de agua (b) imagen infrarrojo (c) imagen visible. Fuente: CPTEC-INPE	88
Figura 35: Parte superior, pronóstico probabilístico MOS de la categoría 1 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 0.01mm.....	91
Figura 36: Parte superior, pronóstico probabilístico MOS de la categoría 2 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 1.00 mm.....	92

Figura 37: Parte superior, pronóstico probabilístico MOS de la categoría 3 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 2.50 mm..... 93

Figura 38: Parte superior, pronóstico probabilístico MOS de la categoría 4 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 5.00mm..... 94

Figura 39. Ruta operativa para la generación de pronósticos probabilísticos 96

RESUMEN

En la presente tesis se utilizó la técnica estadística de ajuste estadístico MOS (Model Output Statistic) el cual se basa en desarrollar modelos estadísticos que expresen la relación entre las salidas de los modelos numéricos (predictores) y la variable a pronosticar (predictando). Esta técnica fue aplicada en la generación de pronósticos de precipitación a nivel diario expresados en probabilidades, para acumulaciones diarias que superen los umbrales de 0.01, 1.00, 2.50 y 5.00 mm en 40 estaciones ubicadas en todo el Perú. Los predictores seleccionados provienen de diferentes variables de pronóstico generados por el modelo ETA-SENAMHI durante el periodo de Enero del 2010 a Enero del 2013. Al ser los predictandos variables del tipo probabilística, se desarrollaron modelos de regresión logística, los cuales fueron calibrados en base a la información de precipitación observada y los predictores del modelo ETA-SENAMHI. El análisis de la eficiencia de los pronósticos se hizo midiendo los diferentes atributos de las variables probabilísticas tales como la fiabilidad, precisión y habilidad; sin embargo, se ha encontrado que a pesar de ser diferentes atributos, estas medidas están relacionados directamente, dando resultados muy similares entre ellos. Las estaciones ubicadas en la sierra presentan la mejor calidad de pronósticos probabilísticos ya que poseen los mejores resultados de fiabilidad, habilidad y resolución; mientras que las estaciones de la costa son las que presentan los pronósticos probabilísticos más deficientes, y que puede estar asociado al bajo rendimiento del modelo ETA-SENAMHI en representar las condiciones atmosféricas en dicha región. A nivel estacional, se observa mayor rendimiento de los pronósticos en el trimestre de diciembre-enero-febrero, mientras que en el trimestre de junio-julio-agosto los pronósticos probabilísticos presentaron para todas las estaciones un bajo desempeño. A nivel estacional, los resultados más óptimos de pronósticos fueron para umbrales de 0.01 mm; mientras que para mayores umbrales de precipitación, el rendimiento del pronóstico disminuye.

Key words: Precipitación, pronóstico probabilístico, MOS, regresión logística, ETA-SENAMHI

ABSTRACT

The statistical post-processing technique MOS (Model Output Statistic), which is used to develop statistical models that express the relationship between the outputs of numerical models (predictor) and predicting variable, was applied in the present study. This technique was used to make twenty – four – hour probabilistic quantitative precipitation forecasts for accumulations exceeding thresholds of 0.01, 1.00, 2.50 and 5.00 mm in 40 meteorological stations over Perú. The selected predictors are results of different variables predicted generated by the ETA-SENAMHI model during the period January 2010 to January 2013. Because the predictands are probabilistic variables, logistic regression models were used and calibrated using observed daily rainfall information and predictors from ETA-SENAMHI model. The analysis to know the efficiency of the forecast was made considering the attributes of probabilistic variables such as reliability, precision and skill. The results show that the stations located in the mountains have the best quality of probabilistic forecasts because they have good results for reliability, precision and skill. Stations located in the rainforest exhibit fair results because they have good reliability but poor precision and skill. Stations located in the coast have the worst probabilistic forecasts, which may be associated with the poor performance of ETA-SENAMHI model in representing the weather conditions in this region. According to the seasonal analysis, summer months (December to February) have the best probabilistic forecasts, while in winter months (June to August) probabilistic forecasts for all stations showed poor performance. Finally, it was found that the best results of probabilistic forecasts were for threshold of 0.01mm, at higher thresholds of accumulated precipitation the performance decreased.

Key words: Precipitation, Probabilistic forecast, MOS, logistic regression, ETA-SENAMHI

I. INTRODUCCIÓN

La constante mejora de los modelos de predicción numérica del tiempo ha dado como resultado un gran avance en los pronósticos durante las últimas décadas. A pesar de este avance, la precipitación es una variable que sigue siendo complicada de pronosticar ya que es altamente dependiente de procesos a pequeña escala y a la geografía local, los cuales no pueden ser simulados por los modelos debido a su gruesa resolución. Por estos motivos, el pronóstico de cantidades precipitación todavía no es tan preciso como los campos a escala sinóptica de presión, humedad y velocidad de viento (Applequist *et al.*, 2002). En el Perú se puede observar este efecto ya que debido a la accidentada topografía y factores que condiciona el estado de la atmosfera, hay una gran complejidad al realizar pronósticos de precipitación a nivel puntual.

El pronóstico numérico trae asociada una determinada incertidumbre (Lorenz, 1963). Los modelos meteorológicos dinámicos presentan claras dificultades a la hora de predecir correctamente variables altamente dependientes de las parametrizaciones o de procesos a escala local y regional, como son la precipitación y la nieve (Subias *et al.* 2012). Existe la necesidad de desarrollar, calibrar y validar diferentes modelos estadísticos o técnicas que mejoren los resultados de los modelos de predicción numérica del tiempo, también conocidos como técnicas de pos-procesamiento.

Una de las técnicas estadísticas de post-procesamiento muy usada es conocida como MOS (Model Output Statistic) (Glahn y Lowry, 1972) que consiste en desarrollar ecuaciones que expresen la relación entre las salidas de los modelos numéricos (predictores) y la variable a pronosticar (predictando).

Una importante ventaja de los métodos de pronósticos estadísticos como el MOS sobre los métodos dinámicos es la capacidad de producir pronósticos de probabilidad. La inclusión de elementos de probabilidad dentro del formato de pronósticos es ventajoso ya que provee una expresión explícita de la inherente incertidumbre o estado del conocimiento sobre el tiempo futuro (Wilks, 2006).

La ventaja del pronóstico probabilístico frente al determinístico es que no sólo brinda la información acerca de la ocurrencia/no-ocurrencia de precipitación sino que indica la probabilidad con la que puede superar cierto umbral (Matsudo *et al.*, 2012). Un pronóstico de tipo probabilístico permite tener en cuenta las diferentes fuentes de incertidumbres asociadas a cualquier pronóstico (Medeiro, 2007). Con el fin de cuantificar la incertidumbre, los pronósticos probabilísticos van adquiriendo cada vez mayor importancia y suponen un avance respecto a los del tipo determinístico (Medeiro, 2007) facilitando la comprensión en los usuarios y dando un pronóstico mucho más confiable.

El objetivo general de esta tesis es proponer una metodología para el pronóstico cuantitativo de la precipitación en términos de probabilidades. Esta metodología se basa en la técnica de ajuste estadístico MOS y la cual necesita ser analizada para su aplicación; por consiguiente se plantearon los siguientes objetivos específicos:

- Determinar la eficiencia del pronóstico probabilístico diario de la precipitación.
- Aplicar los resultados de los pronósticos probabilísticos diario de la precipitación en un evento de fuertes precipitaciones.

II. REVISIÓN DE LA LITERATURA

2.1. VARIABILIDAD ESPACIO – TEMPORAL DE LA PRECIPITACIÓN.

Perú se caracteriza por poseer una gran diversidad de climas debido a su compleja orografía. Según Peñaherrera (1969) las características climatológicas de nuestro país se reflejan desde lo tropical, cálido y húmedo hasta el glaciar frío y seco, pasando por el árido y templado. Por su localización geográfica, dentro de la franja tropical, al Perú le corresponde un clima cálido, húmedo y lluvioso; sin embargo, la presencia de la Cordillera de los Andes, la existencia de la Corriente de Humboldt y la Circulación Anticiclónica del Pacífico Sur, han modificado esta condición climática teórica y lo cual a su vez influye que se presente una gran variabilidad con respecto a la precipitación.

En la costa del Perú se observa normalmente la ocurrencia de lluvias en los meses de verano para la región norte, mientras que en la costa central y sur las cantidades de precipitación son muy bajas debido a la presencia de una capa de inversión en el Pacífico oriental que impide el desarrollo de células convectivas y forma nubes tipo estratos por debajo de la base de inversión térmica. En los meses de invierno, esta inversión térmica es más intensa y ocasiona lloviznas, mientras que en el verano la inversión térmica es débil, lo que produce ingreso de masas de aire del este provocando lluvias esporádicas con gotas más gruesas. En la figura 1 se muestran los valores de precipitación acumulada durante los meses de septiembre a mayo siendo la costa la región con los valores más bajos de precipitación acumulada.

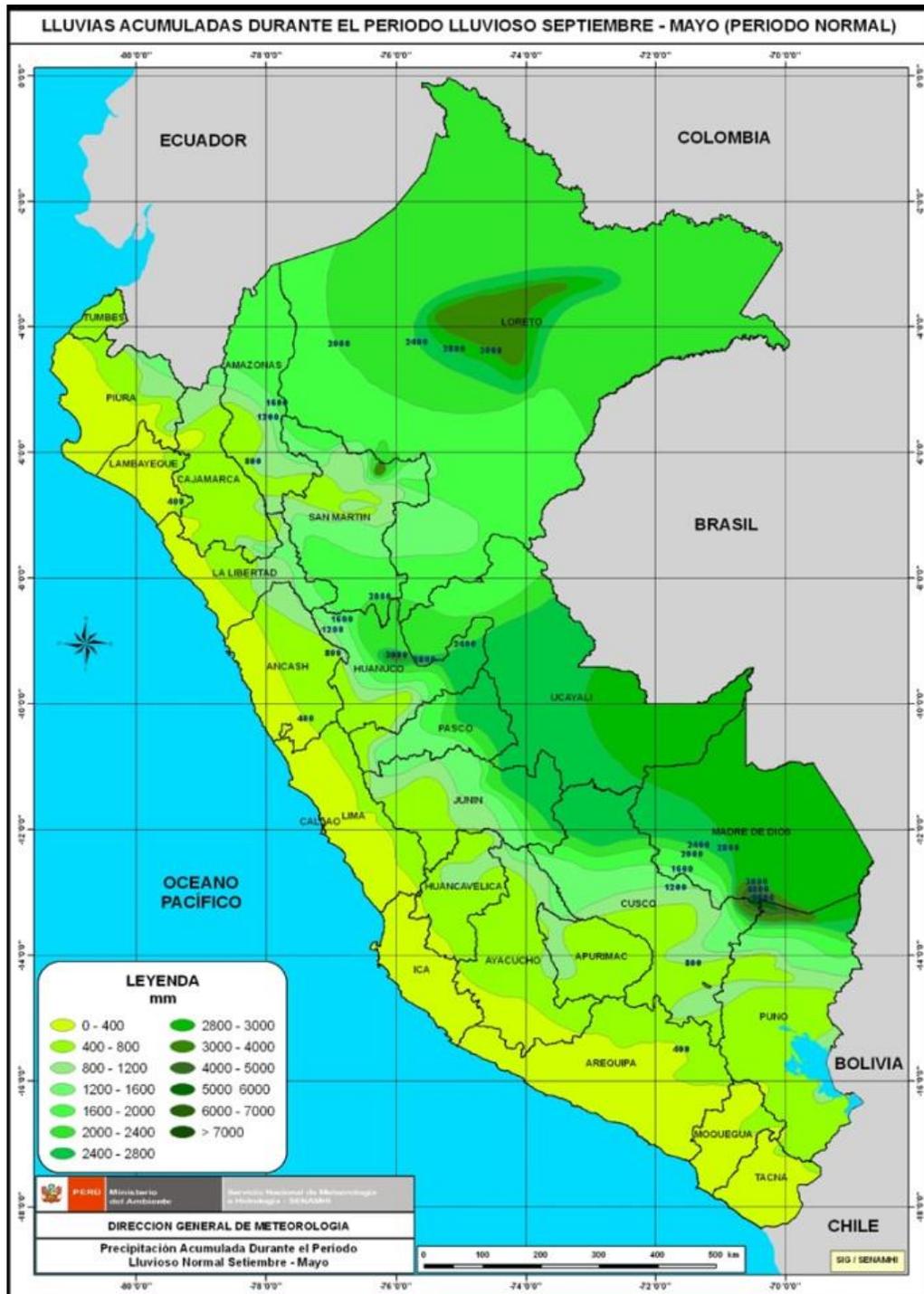
Respecto a los andes, Lagos *et al.* (2008) mencionan que la precipitación empieza en Septiembre y acaba en Mayo, con valores máximos en los meses de Enero-Marzo, y una rápida disminución en Abril. En el invierno austral se presenta un periodo seco que es más intenso en el oeste e interior de los Andes. El comportamiento climatológico de la

precipitación en los andes se ve influenciada por el desplazamiento estacional norte-sur y este-oeste de los sistemas anticiclónicos del Pacífico Sur y Atlántico Sur, que permiten la formación de una vaguada, durante el verano, en el continente. Asimismo, el desplazamiento estacional norte-sur de la Zona de Convergencia Intertropical (ZCIT), la convección y el transporte de humedad desde la Amazonía y la formación de un núcleo anticiclónico a una altura aproximada de 12 Km, conocido como la Alta de Bolivia, son fuertes factores que modulan la precipitación en los andes (Virji, 1981, Silva Días *et al.*, 1983, citados por Lagos, 2007). En la región de la sierra central, Takahashi (2004) encontró que hay una buena correlación entre la intensificación de los vientos del este y el aumento de las lluvias en el valle del río Mantaro durante la estación lluviosa, lo cual sugiere que la incursión de aire del este tiene mayor importancia en esta región. En el altiplano, Garreaud *et al.* (2003) demostraron que el mecanismo de lluvias se debe fundamentalmente al transporte de humedad desde la cuenca amazónica, mientras que Lenters y Cook (1995) demostraron que el flujo de convergencia de humedad en la vertical debido a la orografía local es el principal mecanismo por la que se produce la precipitación sobre el Altiplano.

Asimismo, la altitud también influye notablemente en las características de la precipitación en los andes. A poca altura, las precipitaciones abundantes se relacionan con el aire cálido húmedo y la liberación de gran cantidad de vapor de agua sobre las primeras laderas orientales de los Andes (Espinoza *et al.*, 2009). Sin embargo, puede haber zonas que estén protegidas por las montañas, de los vientos alisios húmedos con dirección predominantemente del este, lo cual se manifiesta en bajos niveles de precipitación lo cual se traduce en una fuerte variabilidad espacial de las lluvias por debajo de los 2000 msnm. (Espinoza *et al.*, 2009). Además, la altitud de la Cordillera es distinta en diferentes latitudes. Al respecto, Hernán (1951), señala que las cordilleras son más bajas y las temperaturas más elevadas en el norte del Perú, los vientos alisios se despojan totalmente de su vapor acuoso y una proporción de ella transmonta la cordillera. Por el contrario, en el sur y centro, la cordillera alcanza grandes alturas, los vientos van despojándose paulatinamente de su humedad a medida que ascienden y que al llegar a las cumbres están casi desprovistas de vapor acuoso y no reciben lluvias, pero en verano, al incrementarse la velocidad de los vientos, la precipitación aumenta en la parte alta.

Lo que respecta a la selva, Marengo, J. (1983), menciona que el inicio de las lluvias corresponde al mes de setiembre, con máximas entre febrero-abril y mínimos en julio-agosto. Además, existe un segundo mínimo de la precipitación, denominada localmente “veranillo” que ocurre entre diciembre y marzo. En la figura 1, los mayores valores de precipitación acumulada durante el periodo lluvioso normal se encuentran particularmente en la región centro y sur, mientras que en la región norte se aprecia un núcleo de máxima precipitación en Loreto. Los altos valores de precipitación debido al incremento de la concentración de humedad permite que esta región sea una de las zonas con mayor convección y lluvias a nivel de Sudamérica (CONAM, 2005). Durante los meses de menor precipitación en la baja atmosfera se puede distinguir determinados factores que originan convección para la selva peruana como el fenómeno “Friaje” el cual presenta características climáticas y sinópticas propias de una perturbación frontal (Marengo, 1983). Quispe (2005) describió los friajes como la incursión de un sistema frontal en la región sur y centro del Perú y sobre la selva norte, generando gran inestabilidad y produciendo fuertes precipitaciones. Estas invasiones frontales frías hacia latitudes tropicales que se presentan en los meses de invierno con mayor intensidad y frecuencias.

Figura 1: Precipitación acumulada durante el periodo lluvioso normal Setiembre – Mayo.



Fuente: SENAMHI, 2002 citado por INDECI, 2003

2.1.1. Variabilidad interanual

El comienzo y la intensidad de las temporadas de precipitación puede variar año a año y en algunos años la variabilidad puede tomar forma de un evento extremo, ya sea como exceso o déficit de precipitación, lo cual pueden convertirse en desastres naturales.

La variabilidad interanual de la precipitación en el Perú están en gran medida determinadas por la presencia del El Niño-Oscilación Sur (ENOS), el cual es una perturbación del sistema Océano Atmósfera en el Pacífico ecuatorial. El ENOS posee una fuerte y directa influencia sobre gran parte de Sudamérica tropical por medio de las teleconexiones¹ (SENAMHI, 2007). Eventos de precipitaciones extremas en los Andes ha sido asociado por investigadores locales con eventos del ENOS, a pesar de que la relación entre estos eventos climáticos y El Niño no ha sido establecido todavía (Lagos *et al.*, 2008); sin embargo, autores como Garreaud y Aceituno (2007) afirman que los principales efectos del ENOS con respecto a las precipitaciones es el aumento de las lluvias sobre lo normal en el norte de Perú y Ecuador y lluvias debajo de su normal en el Altiplano (Peruano – Boliviano).

2.2.PATRONES DE CIRCULACIÓN ATMOSFÉRICAS ASOCIADOS A LA PRECIPITACIÓN EN EL PERÚ

La distribución espacio-temporal de la precipitación depende en gran medida de las circulaciones que pueda haber en niveles altos y medios de la tropósfera² así tenemos que la precipitación están estrechamente relacionadas con la circulación en niveles altos, con un flujo zonal del este el cual favorece condiciones húmedas y un flujo del oeste causando condiciones secas (Condom *et al.* 2010). Para comprender mejor la distribución espacial y temporal de la precipitación es necesario tener un conocimiento cabal de los mecanismos de la circulación de la atmósfera a nivel de Sudamérica.

¹ Cambios en la circulación atmosférica global inducida por anomalías océano-atmosférica del Pacífico Ecuatorial.

² Llámese niveles altos a las capas de la atmósfera en donde la presión es de 300-200 hPa (12000 msnm aproximadamente), mientras que niveles medios y bajos la presión es de 500 hPa y 1000-850 hPa respectivamente (5500 msnm y 100-1500 msnm)

A. Sistemas atmosféricos predominantes durante el verano austral

Los máximos de precipitaciones en los meses de verano están asociados a un sistema anticiclónico en niveles altos, llamado también Alta de Bolivia, el cual se ubica al sur este de los Andes centrales (Garreaud, 2000). Este sistema fue estudiado por primera vez por Schwerdtfeger (1961) encontrando que es típico del verano austral, y desaparece durante el invierno austral.

La Alta de Bolivia es caracterizado por un núcleo cálido ubicado aproximadamente en 200 hPa y es generado a través de una combinación de calor sobre el altiplano Boliviano y el calor latente de condensación liberado por la intensa convección al oeste de la cuenca amazónica. Se forma en Diciembre y permanece hasta Marzo, perdiendo sus principales características al comienzo de Abril, cuando la vaguada ecuatorial se mueve hacia el norte (Satyamurty, Nobre y Silva Días, 1998). La principal función de la Alta de Bolivia es el de hacer converger, sobre Perú y otros países de Sudamérica, la inestabilidad y humedad necesaria para la generación de lluvias, en particular sobre las zonas andinas (IGP, 2010).

Existen numerosos estudios en el que han notado una relación entre la variación estacional en las lluvias sobre Sudamérica y la posición e intensidad del Alta de Bolivia (Kousky y Kagano 1981; Nishizawa y Tanaka 1983) ya que durante el verano este centro anticiclónico se hace más meridional, aumenta el contenido de humedad sobre la Amazonía y se establece una convergencia de los vientos al norte del Perú favoreciendo el flujo de humedad desde el este. La convergencia de vapor de agua en bajos niveles y la liberación de calor latente en la troposfera media ayuda a sostener el movimiento ascendente y mantiene a la región relativamente caliente en la troposfera alta.

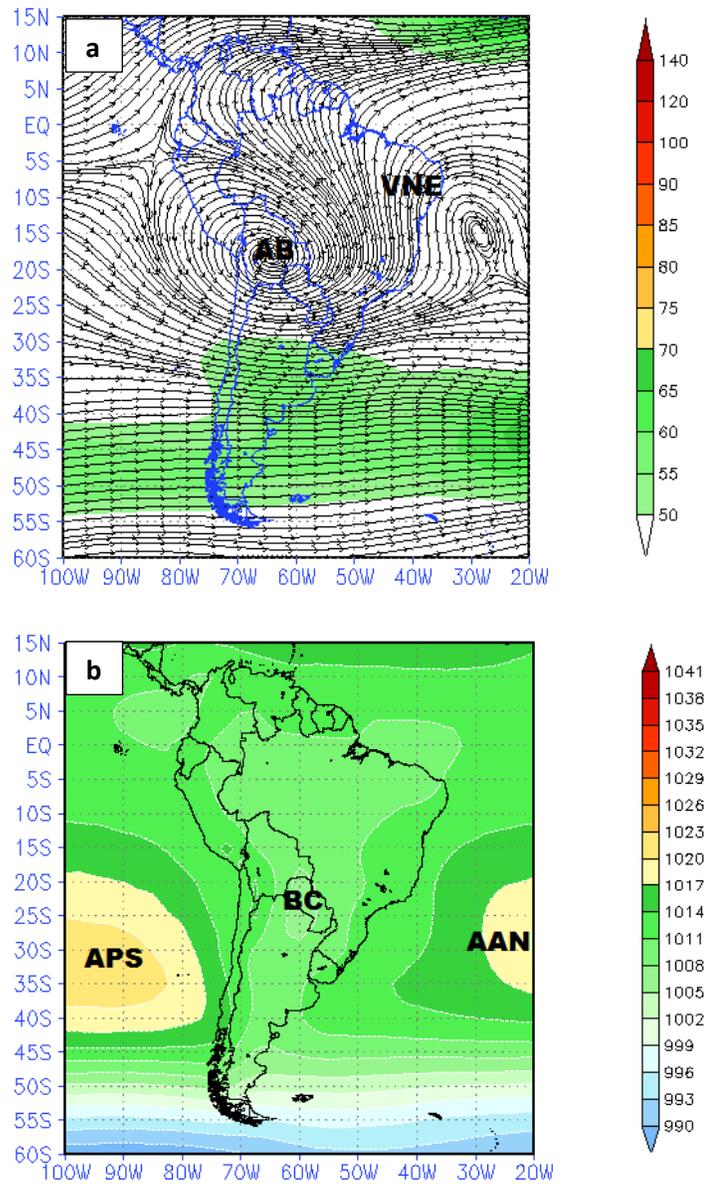
La figura 2a muestra mediante líneas de corriente en 200 hPa que el centro de la Alta de Bolivia se encuentra ubicado en 18°S, 68°W para el mes de Enero. Asimismo, al noreste de Brasil se presenta una profunda vaguada fría, la cual está conectada dinámicamente con la Alta de Bolivia y forma muy a menudo vórtices ciclónicos de núcleo frío (Satyamurty, Nobre y Silva Dias., 1998) el cual puede trasladarse hacia nuestro país cuando la alta de Bolivia se

debilita, originando inestabilidad atmosférica. Esta vaguada es llamada comúnmente Vaguada del Noreste (VNE).

Cerca de la superficie, en la capa baja de la troposfera a nivel de 850 hPa hay presencia de vientos que ingresan por el norte de Sudamérica procedente del Atlántico norte, y fluyen hacia el Oeste sobre el continente donde se encuentra con los Andes entre Colombia y Perú, que bloquea su avance en los meses de verano para luego circular al sur y sureste. Esto último se ve reflejado en la dirección de los vientos promedios, los cuales son del este y norte, favoreciendo la presencia constante de flujos cálidos y húmedos, y alcanzan gran parte del territorio con mayor intensidad y extensión en comparación a otras estaciones del año (CONAM, 2005).

A nivel de superficie (figura 2b) para el mes de enero, se forma un centro de baja presión al noroeste de Argentina y Paraguay conocido como Baja de Chaco (BC) y que favorece la inestabilidad atmosférica y el desarrollo de convección y lluvias en la región de ceja de selva, andes centrales y sur (CONAM, 2005).. Al este de los Andes la circulación en la baja atmosfera es controlado por esta baja térmica y la alta humedad prevalece en la atmosfera media y baja (Garreaud, 1998). Asimismo, se observan la presencia de dos centros de alta presión en el pacifico sur y atlántico sur conocidos como Anticiclón del Pacifico Sur (APS) y Anticiclón del Atlántico Sur (AAS) respectivamente, los cuales se desplazan más hacia el sur en estos meses. Por otro lado, la Zona de Convergencia Intertropical (ZCIT), que es la región donde confluyen los vientos alisios del hemisferio norte y del sur, puede favorecer el desarrollo de lluvias, ya que para esta estación del año se encuentra más posicionado al sur.

Figura 2: Climatología para enero de: (a) de vientos en 200 Hpa y (b) Presión atmosférica a nivel medio del mar.



Fuente: Solis et. al, 2012.

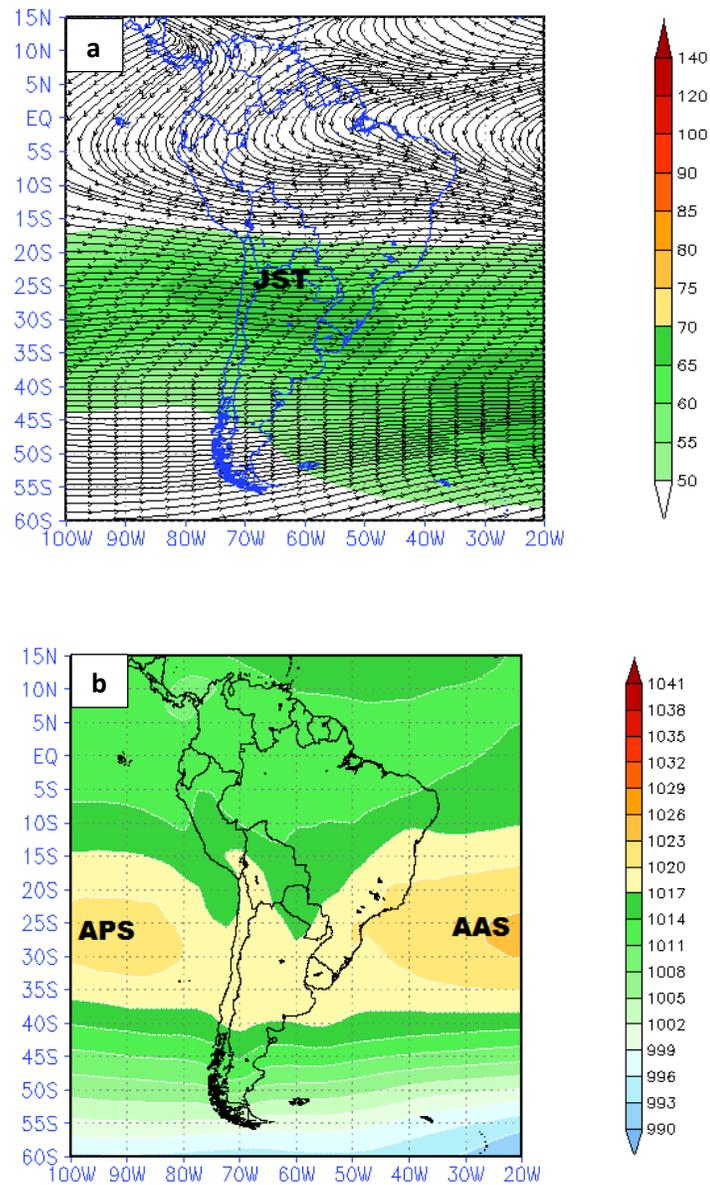
B. Sistemas atmosféricos predominantes durante el invierno austral

En el invierno austral, la configuración de la alta de Bolivia desaparece. Como se aprecia en la figura 3a, los vientos en 200 hPa predominantes sobre el centro y sur de Sudamérica (incluyendo el centro y sur del Perú) se desplazan de Oeste a Este y se caracterizan por ser flujos de aire seco, y asociados al enfriamiento generalizado en la región, contribuyendo a la marcada estabilidad atmosférica sobre el territorio y al decremento de las lluvias, característico de esta época del año.

En niveles cerca de superficie, los vientos del norte y este se debilitan en general para Perú, aun cuando en promedio se presentan ligeros vientos del Este en el sector norte (principalmente en zonas bajas del oriente). Se fortalecen los sistemas de alta presión del pacífico sur y del atlántico sur (con intensidades en promedio de 1023 hPa y 1024 hPa como se aprecia en la figura 3b) desplazándose más al norte y cerca al continente junto con la ZCIT (que ocupa su posición norte más extrema, alrededor de 10°N en el Pacífico y 5°N en el Atlántico aproximadamente), provocando que en el centro y sur del continente, se observan altas presiones y estabilidad atmosférica favoreciendo la inversión térmica y ocasionando lloviznas en la región costera del Perú.

La humedad es la más baja en estos meses del año en varios sectores del centro y sur del continente. En el Perú, la poca humedad que existe se concentra en el norte, particularmente en el sector nororiental, mientras que en el sector central y sur se consolida el período de estiaje (CONAM, 2005).

Figura 3: Climatología para julio de (a) de vientos en 200 Hpa y (b) Presión atmosférica a nivel medio del mar.



Fuente: Solis et. al, 2012.

2.3. MODELOS DE PREDICCIÓN NUMÉRICA DEL TIEMPO

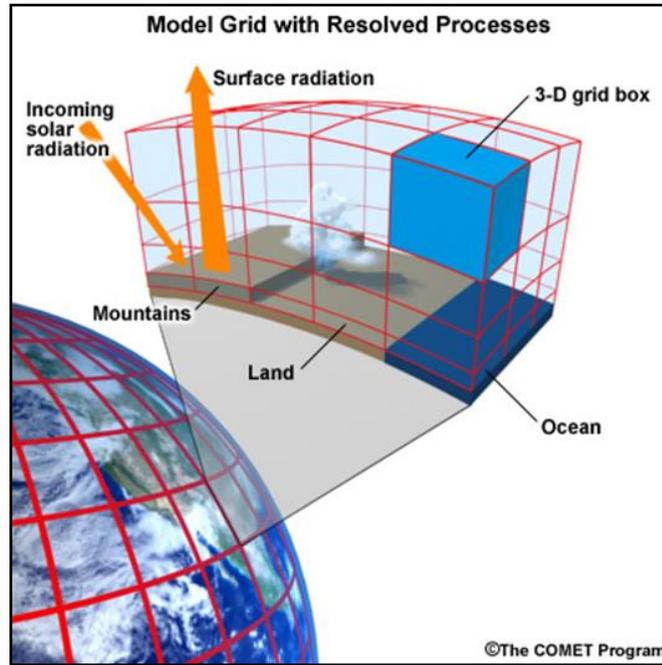
La predicción numérica del tiempo es una metodología de pronóstico del tiempo que emplea modelos matemáticos formulados por ecuaciones en derivadas parciales, las cuales traducen las leyes generales de la física que rigen la atmósfera terrestre y representan la evolución temporal y espacial de los flujos atmosféricos. Las ecuaciones matemáticas obtenidas son no lineales por lo que no pueden ser resueltas de forma analítica y se debe apelar al cálculo numérico, de tal forma que se reemplazan las ecuaciones en variables continuas por ecuaciones en las que las variables son discretas (Illanes, 2010). El estado futuro de la atmósfera se calcula usando aproximaciones numéricas de las ecuaciones dinámicas.

De acuerdo con lo anterior, se puede definir un modelo numérico de predicción del tiempo como un conjunto de instrucciones ejecutadas en un equipo de cómputo (programa), el cual resuelve las ecuaciones de pronóstico mediante técnicas numéricas para generar una aproximación del estado futuro de la atmósfera (Lopez, 2012).

Las diferentes ecuaciones que gobiernan el comportamiento de los flujos son las de cantidad de movimiento, ecuaciones de continuidad, conservación de la energía y humedad, conocidas también bajo el nombre de ecuaciones primitivas y están basadas en los principios de conservación de masa, cantidad de movimiento y energía (Pielke, 2002). Las soluciones que se obtienen se integran en un área denominada dominio.

Los modelos numéricos representan los flujos atmosféricos dividiendo la atmósfera en varias capas discretas horizontales y cada una de ellas a su vez es dividida en un número de celdas, donde las variables son evaluadas en el centro de cada celda (ver figura 4). De forma similar, el tiempo es llevado a un número finito de pasos discretos, siendo la evolución continua de las variables aproximadas por pasos discretos (Lighezzolo, 2014). Para los procesos físicos en meteorología que ocurren a escalas espaciales menores al tamaño de una celda de grilla, los modelos necesitan recurrir a valores paramétricos, los cuales representan efectos físicos totales mediante un único número que represente a la celda, se concibe a la parametrización en términos de modelar los efectos de un proceso (emulación) en lugar de modelar el proceso en sí (simulación) (METED, 2010).

Figura 4: Representación del mundo por modelos numéricos.



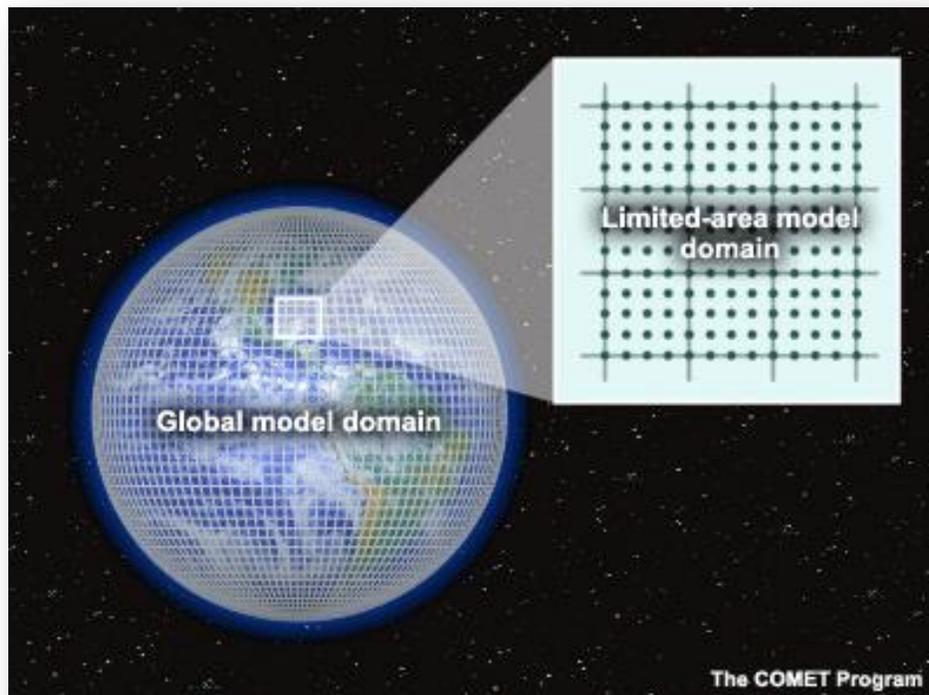
Fuente: MetEd, 2010.

Basado en la escala temporal y espacial de los fenómenos a estudiar los modelos numéricos usados en Ciencias Atmosféricas pueden clasificarse como modelos de circulación general (MCG) y modelos regionales.

Los **modelos globales** son usados para obtener aproximaciones de procesos importantes que ocurren en la atmósfera a nivel global y a escalas temporales diversas con resoluciones espaciales que van de decenas a cientos de kilómetros. Estos modelos consideran los procesos de mayor influencia en la circulación del planeta como el flujo de energía térmica, cantidad de energía recibida y emitida en forma de radiación de onda larga y corta, etc (Lopez, 2012). Un ejemplo de estos modelos es el Global Forecast System (GFS) el cual es desarrollado por National Centers for Environmental Prediction (NCEP) USA, utilizado en la predicción numérica del tiempo y que corre cuatro veces al día produciendo pronósticos de hasta 16 días.

Los **modelos regionales**, los cuales se conocen también como modelos de área limitada, simulan los procesos dinámicos de la atmósfera en resoluciones espaciales muy altas (unos cuantos kilómetros) ya que poseen condiciones de frontera bien definidas en dirección horizontal y vertical. La información meteorológica que proporcionan estos modelos es más detallada para regiones específicas con fines distintos a la investigación (agricultura, aeronáutica, protección civil, etc). Consideran parametrizaciones para describir procesos como formación de nubes o la dinámica de la capa planetaria (Lopez, 2012). Ejemplos de estos modelos son el Workstation ETA model, Weather Research and Forecasting Model (WRF), PSU/NCAR Mesoscale Model versión 5 (MM5). Debido a que estos modelos corren en un área limitada requieren información en el contorno de sus dominios, siendo esta información proveniente de modelos globales por lo que generalmente son anidados dentro de los modelos globales (ver figura 5).

Figura 5: Representación de un modelo regional anidado a un modelo global.



Fuente: MetEd, 2010

2.4. MODELO REGIONAL ETA

El modelo ETA es un modelo atmosférico regional usado para propósitos de investigación y operacional. Su principal ventaja es la utilización de coordenada vertical ETA. Mesinger (1984) definió la coordenada vertical ETA con el objetivo de remover, en gran medida, los errores de los modelos con coordenadas sigma que se producen al calcular la fuerza de gradiente de presión, así como la advección y difusión horizontal, a lo largo de una pendiente muy pronunciada.

La coordenada vertical hidrostática ETA (η) (Mesinger, 1984) está dada por la siguiente expresión:

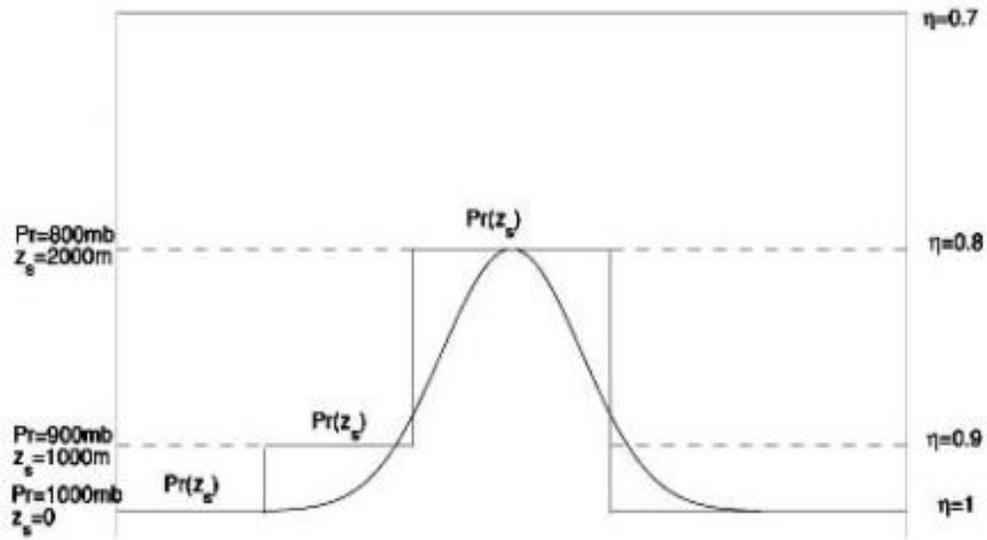
$$\eta = \frac{p - p_t}{p_s - p_t} \eta_s$$

Dónde:

$$\eta_s = \frac{p_{rf}(z_s) - p_t}{p_{rf}(0) - p_t}$$

Los sufijos s y t representan la superficie y el tope de la atmosfera del modelo. $p_{rf}(z)$ Es un nivel de presión de referencia expresado como una función de la altura. El valor de z_i indica una altura definida de referencia del terreno el cual puede tomar solo valores discretos. Con la coordenada vertical ETA, las montañas son creadas como cajas de grillas en tres dimensiones.

Figura 6: Representación gráfica de la coordenada vertical ETA.



Fuente: Popovic, 2006.

La topografía del modelo ETA es construida en pasos discretos tal y como se muestra en la figura 6, con topes localizados en las interfaces del modelo. Las coordenadas por pasos reducen los errores que ocurren cuando se calcula derivadas horizontales en la proximidad de terrenos en pendiente. Por lo tanto, las superficies de nivel resultan prácticamente horizontales, permitiendo mejorar el cálculo de magnitudes horizontales en presencia de topografía abrupta (Garreaud & Seluchi, 2003).

El hecho de que las superficies del ETA tiende a ser cuasi-horizontal y permita que las características del terreno, como montañas, puedan ser representados como bloques de tierra sólida en varios niveles del modelo tiene las siguientes ventajas (METED, 2010):

- Con las coordenadas ETA no es necesario realizar una interpolación vertical necesaria para calcular la fuerza de gradiente de presión, lo cual reduce el error en los cálculos y mejora el pronóstico del viento, temperatura y cambios de humedad en zonas con fuerte pendiente del terreno.

- Mediante las coordenadas ETA, la convergencia en bajos niveles para terrenos empinados es mejor representado, esto a su vez conlleva a mejores pronósticos de precipitación en estas áreas.
- Comparado con modelos que usan coordenadas sigma, los modelos con coordenadas ETA puede mejorar los pronósticos de brotes de aire frío y ciclogénesis a sotavento.

La resolución vertical del modelo es mayor en capas más bajas, donde la primera capa esta aproximadamente a 20 metros de profundidad, y en las capas cerca a la tropopausa (Chou & da Silva, 1999). Por estas razones, este modelo resulta especialmente apto para el análisis de fenómenos con fuerte componente orográfica, como la circulación en la región del altiplano (Garreaud & Seluchi, 2002).

Además, el modelo ETA presenta diferentes esquemas de parametrización de los siguientes procesos (Sharan & Raman, 2007):

- Procesos en superficie.
- Mezcla en turbulencia vertical.
- Convección de cúmulos.
- Difusión lateral.
- Radiación.

En nuestro país, la Dirección de Meteorología Aplicada del Servicio Nacional de Meteorología e Hidrología SENAMHI, viene corriendo el modelo ETA en forma operacional desde julio del año 2000 (considerando el esquema de convección Betts-Miller hasta el 2006, y con el esquema Kain-Fritsch, a partir del 2007) (Avalos, 2009). Actualmente el modelo corre dos dominios diferentes que cubren Sudamérica y Perú (figura 7), teniendo el primero una resolución de 32 Km y 50 niveles en la vertical, mientras que el segundo cuenta con 22 Km de resolución 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

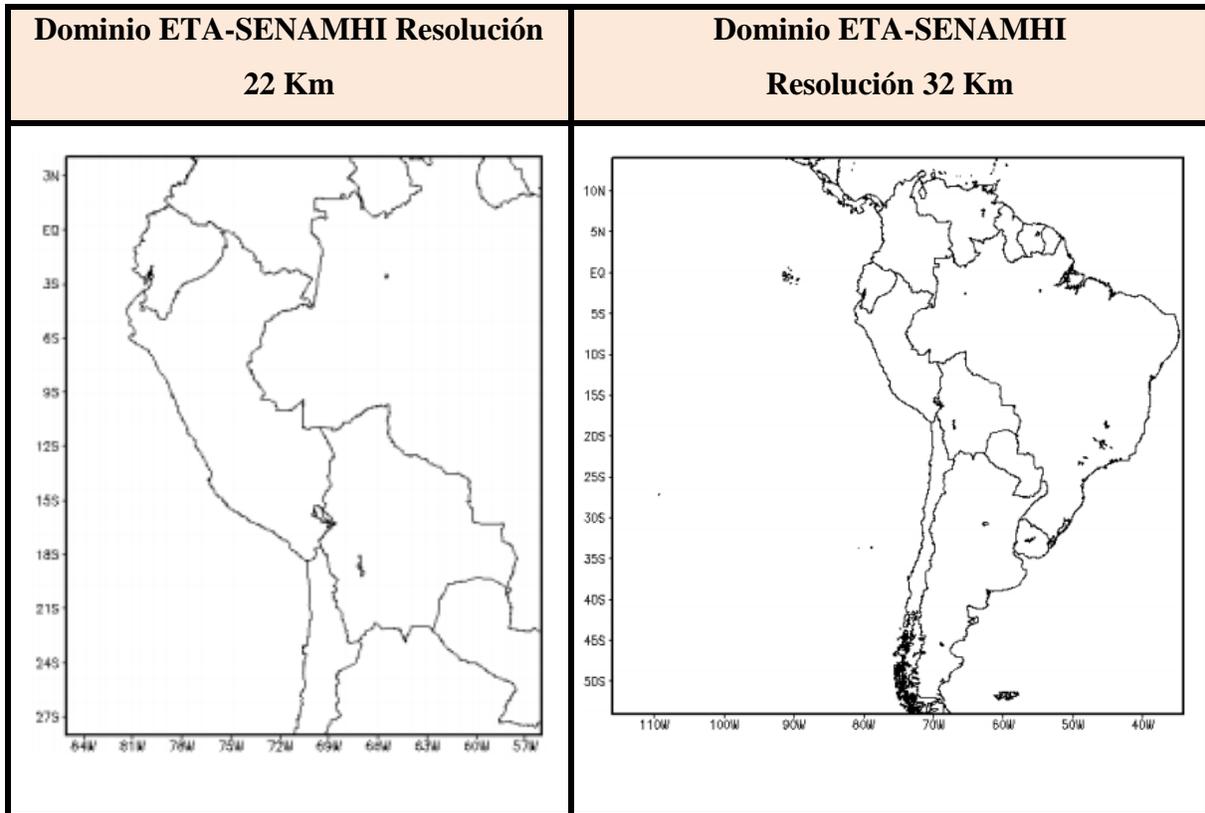
para el dominio de Perú. Los dominios de previsión se encuentran operativos en un servidor SUN para Sudamérica y en un servidor DELL para Perú. Como condiciones iniciales y de borde los datos del modelo global GFS.

Cuadro 1: Características del modelo ETA SENAMHI.

	Perú	Sudamérica
Área	longitud : 85° W 56° W latitud : 28° S 4° N	longitud : 116° W 34° W latitud : 54° S 13° N
Resolución espacial	22 km x 22 km ,38	32 km x 32 km , 50
Resolución temporal	Cada 6 horas Dos veces al día (00Z y 12Z)	
Condiciones iniciales	Atmosfera y Océano : G F S (Global Forecast System)	
Parametrizaciones	<ul style="list-style-type: none"> • Convección: Kain-Fritsch • Capa limite: Mellor – Yamada 2.5 • Radiación: GFDL • Modelo dinámica : Hidrostático 	
Tiempo de pronóstico	72 horas	120 horas
Periodo disponible	Enero 2007 - actualidad	Enero 2009 - actualidad

Fuente: Llacza, 2013

Figura 7: Dominios espaciales del modelo ETA-SENAMHI



Fuente: Llacza, 2013

Los productos del modelo ETA-SENAMHI para los dominios de resolución 32 km y de 22 km son listados en los cuadros 2 y cuadros 3 respectivamente y se encuentran disponibles actualmente de forma gratuita. Los mapas pueden ser descargados en el siguiente enlace web: <http://www.senamhi.gob.pe/?p=0260>.

Cuadro 2: Principales productos de pronóstico del modelo ETA-SENAMHI (32 km) disponibles en <http://www.senamhi.gob.pe/?p=0260>.

Pronóstico (Horas) Variables	0	6	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108	114	120	
Temperatura y viento (1000 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Temperatura y viento (950 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Temperatura y viento (850 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Temperatura y viento (500 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Jet Stream	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adveccion de Vorticidad	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Vorticidad potencial (300hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Isentropico 320°K	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Vorticidad Relativa (500hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Presión nm/espesor (500/1000hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Omega/H. Relativa (500hPa)		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Líneas de Corriente (200hPa,500 hPa, 850 hPa, 700hPa, 100 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adveccion de temp. (850 hPa, 925 hPa, 950 hPa, 975 hPa, 1000 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Humedad relativa (850 hPa, 700 hPa, 600 hPa, 300 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Temperatura de Rocío (925 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Flujo de humedad (600hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Convergencia/Divergencia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Convergencia de Hum. (850 hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Jet bajos niveles Costero (950hPa)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Índice K	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Índice Sweat	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Índice Total de Totales	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Índice Cape	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Precipitación SA 06h		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Precipitación PE 06h		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Precipitación PE 24h				✓		✓		✓		✓		✓		✓		✓		✓		✓		✓
Nieve SA 06h		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Corte sur 5°, 10°, 15°, 20°, 25°		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Cuadro 3: Productos de pronóstico del modelo ETA-SENAMHI (22 km) disponibles en <http://www.senamhi.gob.pe/?p=0260>

Tiempo (Horas) Variables	Análisis	Día 0.5 12 horas	Día 1 24 horas	Día 1.5 36 horas	Día 2 48 horas	Día 2.5 60 horas	Día 3 72 horas
Vorticidad y Altura Geopotencial 500 hPa	✓	✓	✓	✓	✓	✓	✓
Presión a nivel medio del mar y espesor 1000-500 hPa	✓	✓	✓	✓	✓	✓	✓
Velocidad vertical en 700 hPa y Precipitación	✓	✓	✓	✓	✓	✓	✓
Temperatura, humedad relativa y vientos en 850 hPa	✓	✓	✓	✓	✓	✓	✓
Isotacas y líneas de corrientes en 200 hPa	✓	✓	✓	✓	✓	✓	✓
Agua precipitable e índice CAPE	✓	✓	✓	✓	✓	✓	✓

2.5. POST- PROCESAMIENTO ESTADÍSTICO MOS Y SUS APLICACIONES EN EL PRONÓSTICO DEL TIEMPO

La técnica de ajuste estadístico MOS (Model Output Statistics), nombrada así por Glahn y Lowry (1972) es un método de pronóstico el cual usa una muestra de datos de pronósticos en un corto periodo de tiempo producidos por modelos numéricos los cuales son ajustados con las observaciones del tiempo local con el fin de mejorar objetivamente las correlaciones entre las simulaciones y las observaciones (Rincón, 2013). Es una técnica que tiene la capacidad de predecir el error sistemático en sistemas numéricos y se aplica a las salidas de los modelos de predicción numérica del tiempo. Una característica peculiar del MOS es que esta puede incluir diferentes tipos de predictores tales como velocidad vertical, temperatura, viento, humedad, etc.

Dado el gran avance alcanzado con respecto al pronóstico numérico, la técnica de post procesamiento MOS³ se está aplicando en el pronóstico de diversas variables tales como: temperatura mínima y máxima, probabilidad de precipitación, temperatura de punto de rocío, dirección y velocidad de viento, ocurrencia de tormentas severas, visibilidad, entre otras y está

³ Llámese post-procesamiento al conjunto de procedimientos que se hacen a las salidas crudas de los modelos numéricos con el fin de optimizar su manejo y rendimiento.

siendo implementado continuamente operacionalmente en varios centros meteorológicos incluyendo estos en Países Bajos (Lemcke y Kruizinga, 1988), Gran Bretaña (Francis *et al.*, 1982), Italia (Conte *et al.*, 1980), China (Lu, 1991), España (Azcarraga y Ballester, 1991), Canadá (Brunet *et al.*, 1988), y Estados Unidos (Carter *et al.*, 1989). Además, es especialmente útil en regiones con complejas orografía, donde pueden ser encontradas largas diferencias ente la elevación de los modelos numéricos del tiempo y la verdadera orografía (Eccel *et al.*, 2007).

La deficiencia principal de este método estadístico es que se necesita de un elevado número de datos disponibles de salidas de una misma versión de un modelo numérico para obtener un mejor performance de las ecuaciones de previsión.

2.6. MODELOS ESTADÍSTICOS EMPLEADOS EN EL AJUSTE ESTADÍSTICO MOS.

Existen distintos métodos estadísticos que son usados para desarrollar relaciones entre las variables de las salidas de los modelos numéricos del tiempo y las variables locales. El método más común usado es el de regresión lineal múltiple; sin embargo, existe también otros metodologías que son igualmente viables. A continuación, se explicaran algunos de estos métodos.

2.6.1. REGRESIÓN LINEAL MULTIPLE

La regresión lineal múltiple relaciona una variable Y llamada variable dependiente o predictando, con otras variables X , llamadas variable independientes o predictores. El resultado es una ecuación para usado para estimar el predictando como una combinación lineal de los predictores.

$$Y = b_0 + b_1X_1 + b_2X_3 + \dots + b_kX_k \dots \dots \dots (1)$$

De donde Y es el predictando, b_k son las constantes de regresión o coeficientes, X_k los predictores y k el número total de predictores. Nótese que cada predictor X_k tiene su

respectiva constante b_k . Estos coeficientes de regresión son también llamados parámetros de la regresión, y pueden ser estimados con métodos de minimización de sumatorias de cuadrados residuales. (Wilks, 2006).

2.6.2. REGRESIÓN LOGÍSTICA

La técnica estadística que es más usada para variables probabilísticas es modelo de regresión logística el cual es método de regresión no lineal. Se basa en usar los valores de los predictores para formar una ecuación que expresa la probabilidad de algún evento categórico. Es similar a un modelo de regresión lineal pero está adaptado para modelos en los que la variable dependiente es dicotómica.

La ecuación de regresión logística posee la siguiente forma:

$$P_i = \frac{\exp(b_0 + b_1x_1 + \dots + b_Kx_K)}{1 + \exp(b_0 + b_1x_1 + \dots + b_Kx_K)} \dots\dots\dots (2)$$

De donde:

p_i = Probabilidad de la ocurrencia de algún evento determinado para un tiempo i .

$x_1, x_2, x_1, \dots, x_k$ = Predictandos los cuales son salidas del modelo numérico

$b_0, b_1, b_2, \dots, b_k$ = Parámetros de la regresión

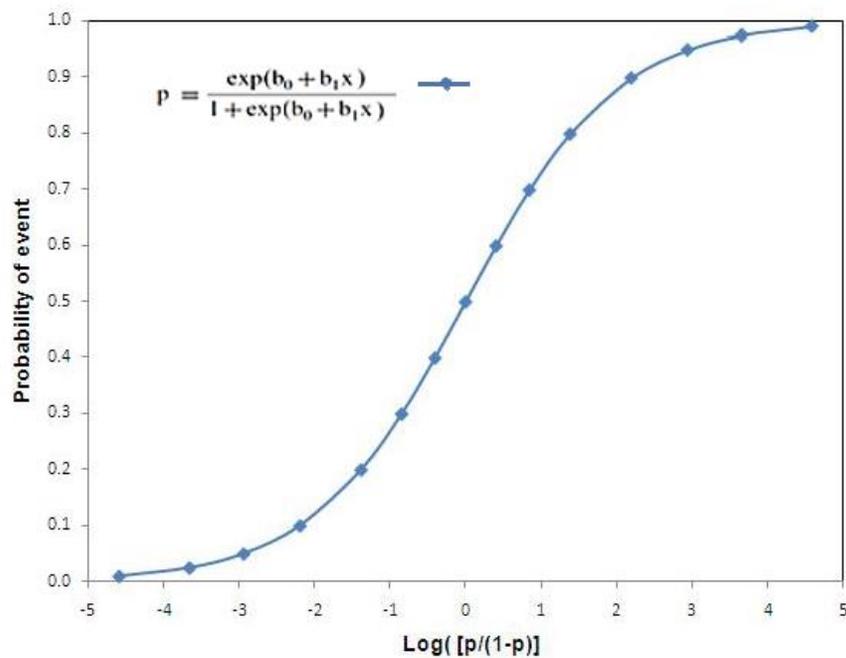
La ecuación 2 se puede expresar también de la siguiente forma:

$$\ln \left(\frac{P_i}{1 - P_i} \right) = b_0 + b_1x_1 + \dots + b_Kx_K. \dots\dots\dots(3)$$

La ecuación 3 es del tipo lineal y similar a la regresión múltiple; sin embargo, debemos tener en cuenta que p_i es un valor de probabilidad, por lo tanto este puede variar entre 0 y 1. Los métodos estándar de regresión lineal no pueden ser aplicados para estimar los parámetros de

regresión en la ecuación 3, ya que los valores observados del predictando son binarios (0 y 1). Para estimar los parámetros o coeficientes se puede utilizar procedimientos iterativos tales como el de máxima verosimilitud (maximum likelihood) (Wilks, 2006a) el cual se basa en encontrar una función que maximice la habilidad para pronosticar la probabilidad basado en lo que se sabe de los predictores. Otro método de estimación de parámetros es el de descenso de gradiente el cual comienza con un conjunto aleatorio de pesos iniciales (b_k) y, a través, de una serie de iteraciones, encuentra nuevos pesos, los cuales minimizan el error cuadrado del predictando (Applequist, 2002).

Figura 8: Gráfico de un modelo de regresión logística simple.



Fuente: Modificado de Strand *et. al*, 2011.

Según Burns (2008) la regresión logística debe cumplir los siguientes supuestos:

- La regresión logística no asume una relación lineal entre las variables dependientes e independientes.
- La variable dependiente debe ser dicotómica (2 categorías).

- Las variables independientes no tienen que estar en intervalo ni linealmente relacionados, ni tampoco poseer igual varianza dentro de cada grupo.
- Se necesitan muestras más grandes a comparación de la regresión lineal, porque los coeficientes de máxima verosimilitud son grandes estimaciones de la muestra.

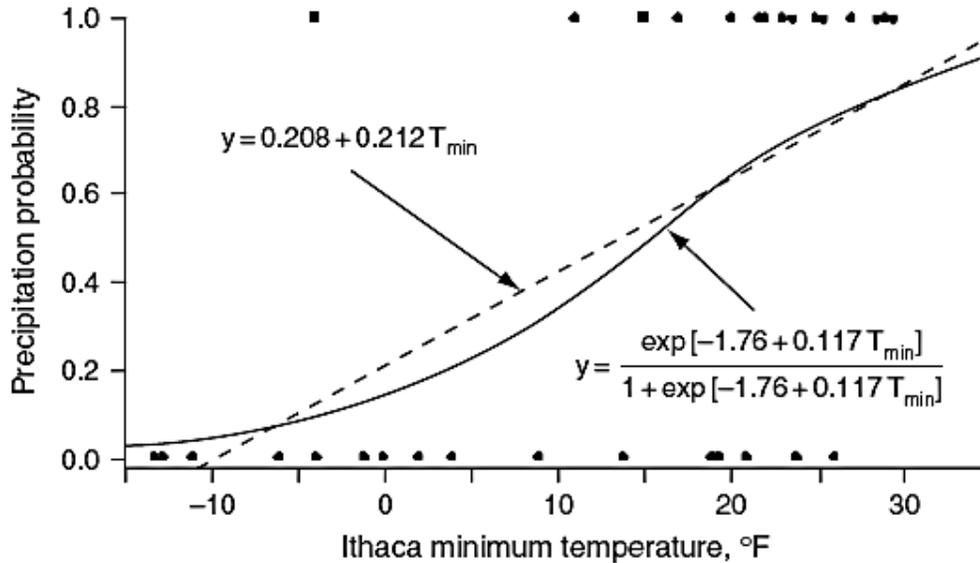
2.7. APLICACIONES DE LA REGRESIÓN LOGÍSTICA EN EL PRONÓSTICO PROBABILÍSTICO DE LA PRECIPITACIÓN

Se ha dado importantes usos de la regresión logística en post-procesamiento estadísticos para el pronóstico de temperatura y precipitación (e.g. Hamill *et al.*, 2004; Hamill *et al.*, 2008; Wilks y Hamill, 2007 citados por Wilks, 2006).

En el caso de las precipitación, las aplicaciones se dieron a través del pronóstico de la probabilidad de precipitación la cual es una variable definida como la probabilidad de que al menos 0.01 mm de precipitación caerá en una locación determinada dentro de un intervalo de tiempo dado (6, 12 o 24 horas). Cuando la cantidad de precipitación pronosticada es mayor (por ejemplo 0.1 mm, 0.25 mm, 1 mm, etc.) la variable se denomina probabilidad de la precipitación cuantitativa. Ejemplos de la aplicación de la regresión logística para los pronósticos probabilísticos de la precipitación se detalla a continuación:

- Wilks (2006) comparó los resultados de pronóstico probabilístico de precipitación de la ciudad de Ithaca (EEUU) calculados a partir de dos métodos de regresión: regresión lineal y regresión logística usando como único predictor la temperatura mínima del aire (figura 9), encontrando que el método de regresión logística produce una curva más ajusta al pronóstico probabilístico ya que los valores varían entre 0 y 1; en cambio con la regresión lineal las probabilidades salen mayores a 1 con temperaturas mínimas mayores 37°F, menores 0 con temperaturas mínimas menores a -10°F

Figura 9: Comparación de pronósticos de probabilidad de precipitación para Ithaca, usando regresión lineal (línea discontinua) y regresión logística (línea sólida).



Fuente: Wilks, 2006.

- Applequist *et al.*(2001) realizaron una evaluación de diferentes metodologías para producir pronósticos probabilísticos MOS de precipitación en la parte este y centro de Estados Unidos, encontrando que la regresión logística funciona mejor a comparación de otras metodologías tales como regresión lineal múltiple, redes neuronales y análisis discriminante para el pronóstico probabilístico de precipitación.
- Crochet (2003) desarrollo un modelo estadístico basado de ecuaciones de regresión logística para pronosticar la probabilidad de precipitación en locaciones específicas de Islandia. Como predictores utilizó salidas de pronósticos del modelo ECMWF de hasta con 5 días de pronóstico, dando como resultados predicciones confiables y útiles.
- Wilks (2009) aplicó la regresión logística para obtener pronósticos MOS de probabilidad con distribución continua y consistente expresados en cuantiles en algunas ciudades de Estados Unidos. Este procedimiento fue realizado usando salidas del modelo GFS.

- Chourio (2009) utilizó índices de estabilidad atmosférica en ecuaciones de regresión logística para el pronóstico de precipitación en la región centro norte costera de Venezuela, encontrando que este modelo cumple con las condiciones para aplicarse como método de pronóstico en esta región.

- Correia *et al* (2013) realizaron un diagnóstico de los extremos de precipitación sobre el noreste de Brasil basado en la regresión logística para obtener asociaciones entre extremos de precipitación y variables meteorológicas. Los resultados evidenciaron que la radiación en onda larga saliente fue la variable clave en la detección de extremos de precipitación en tres subregiones mientras que al sur del área de estudio la variable clave fue la humedad relativa. Así mismo la temperatura máxima, el viento zonal, evaporación y humedad específica también tuvieron influencias en estos extremos.

III. MATERIALES Y MÉTODOS

3.1. ÁREA DE ESTUDIO

Perú se encuentra ubicado en la costa occidental de América del Sur, por debajo de la línea ecuatorial. Tiene como coordenadas a $0^{\circ}01'48''$ S en el extremo norte hasta los $18^{\circ}21'08''$ S en el extremo sur y $68^{\circ}39'27''$ W en el extremo este hasta los $81^{\circ}21'35''$ W en el extremo este. Posee una extensión de $1\ 285\ 215.6\ \text{km}^2$ siendo el tercer país más grande de América del Sur y uno de los 20 más extensos del mundo. Presenta una gran diversidad geográfica que tiene una gran influencia en el desarrollo socioeconómico del país.

Debido a que Perú se encuentra próximo a la línea ecuatorial le debería corresponder un clima del tipo tropical; no obstante, la influencia de factores geográficos como la Cordillera de los Andes y las corrientes marinas determinan su diversidad climática las cuales varían desde un clima tropical en la montaña hasta un clima de nieve o gélido en los Andes. En el territorio peruano se identifican tres regiones que ha sido la forma tradicional de dividirla según sus características geográficas: Costa, Sierra y Selva.

La costa está conformada por una estrecha región en donde podemos encontrar valles y desiertos, situado entre el lado occidental de la Cordillera de los Andes y el océano Pacífico. El clima que posee es generalmente templado cálido con elevada humedad, y durante el invierno se produce densas neblinas lo cual provoca que la sensación térmica sea menor. Las precipitaciones son mayores en el norte mientras que en el sur son escasas, siendo esta región extremadamente árida. Estas condiciones se deben principalmente al afloramiento costero que es originada por la corriente de Humboldt e impide la formación de nubes con gran desarrollo vertical.

La sierra se caracteriza por la presencia de la Cordillera de los Andes y que provoca una accidentada topografía en esta región. Climáticamente cuenta con dos periodos: un periodo

cálido y con abundantes lluvias que comprende los meses de septiembre a abril y un periodo frío y con pocas lluvias que está entre los meses de mayo de agosto. Debido a su geografía esta región presenta una gran variedad de climas los cuales varían entre gélido, lluvioso, semiárido y climas con poca humedad de acuerdo a sus diferentes pisos altitudinales.

Por último, la región selva es la más extensa del territorio peruano. Las principales características es que posee un gran tamaño territorial con una variedad de recursos los cuales no son explotados totalmente. Sus climas cálidos-húmedos es un factor importante para que esta región posea una vegetación muy tupida y variada. Durante los meses de mayo y agosto se producen los fenómenos meteorológicos conocidos como “frijes”, los cuales ocurren por el ingreso de masas de aire frío proveniente del extremo sur del continente provocando que las temperaturas desciendan hasta 12 °C.

Figura 10: Mapa de ecorregiones del Perú.



Fuente: MINAM, 2011.

3.2. MATERIALES

3.2.1. INFORMACIÓN DE PRECIPITACIÓN OBSERVADA

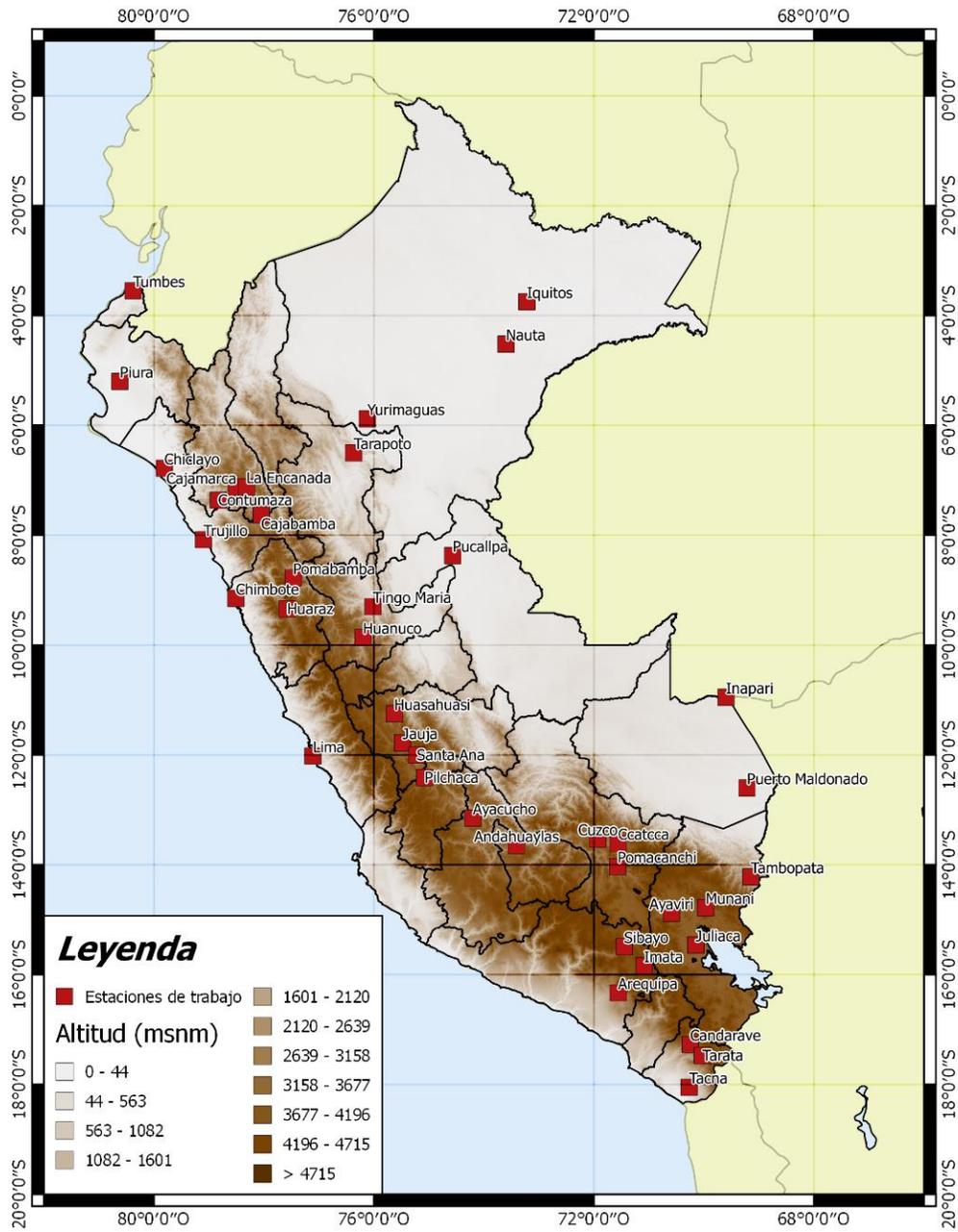
Para el presente trabajo se recopiló información pluviométrica diaria de 40 estaciones que están distribuidas por todo el Perú y los cuales se encuentran listados en el cuadro 4. La mitad de estas estaciones pertenecen a la red de estaciones meteorológicas aeronáuticas de la Corporación Peruana de Aeropuertos y Aviación Comercial (CORPAC); y la otra mitad, a la red de estaciones del Servicio Nacional de Meteorología e Hidrología (SENAMHI). Estos datos se encuentran comprendidos en el periodo de Enero de 2010 a Febrero de 2013. Según la Organización Meteorológica Mundial (OMM, 1981), la precipitación diaria es definida como la precipitación acumulada en 24 horas; en base a esta regla SENAMHI considera la precipitación acumulada de un día cualquiera como la suma de la precipitación registrada a las 19:00 h (hora local) de este día, más la precipitación que se registra a las 07:00 h. (hora local) del día siguiente.

Debido a las condiciones áridas que se presentan en la costa del Perú, las precipitaciones no son abundantes en esta región por lo que se seleccionaron poca cantidad de estaciones siendo un total de siete, las cuales se encuentran ubicadas principalmente en la costa centro y norte, en donde se presentan las mayores frecuencias de precipitación. La región con mayor cantidad de estaciones, siendo un total de 24, es la sierra y se encuentran distribuidas equitativamente a lo largo de la cordillera. Para la región de la selva se cuenta con un total de 9 estaciones. La distribución espacial de estaciones se puede ver en la figura 11.

Cuadro 4: Estaciones usados para el pronóstico de probabilidad de precipitación

REGIÓN	NOMBRE DE LA ESTACIÓN	LATITUD (°) S	LONGITUD (°) W	ALTITUD (msnm)
COSTA	CHICLAYO	6.78	79.81	29
	CHIMBOTE	9.15	78.51	20
	TUMBES	3.55	80.38	25
	TRUJILLO	8.08	79.10	26
	PIURA	5.20	80.62	49
	LIMA	12.02	77.11	12
	TACNA	18.05	70.27	468
SIERRA	CUZCO	13.53	71.93	3248
	ANDAHUAYLAS	13.65	73.41	3444
	CAJAMARCA	7.13	78.50	2620
	AYACUCHO	13.15	74.20	2740
	HUARAZ	9.34	77.58	3080
	HUÁNUCO	9.86	76.20	1859
	JULIACA	15.46	70.15	3826
	AREQUIPA	16.33	71.56	2538
	CONTUMAZA	7.36	78.83	2610
	CAJABAMBA	7.63	78.05	2612
	POMABAMBA	8.78	77.47	3605
	SANTA ANA	12.00	75.22	3302
	JAUIJA	11.78	75.48	3360
	HUASAHUASI	11.25	75.63	3359
	PILCHACA	12.41	75.08	3650
	CCATCCA	13.61	71.56	3693
	SIBAYO	15.49	71.45	3827
	IMATA	15.84	71.09	4445
	AYAVIRI	14.88	70.59	3920
	MUÑANI	14.78	69.97	4119
POMACANCHI	14.03	71.57	3686	
CANDARAVE	17.27	70.25	3435	
TARATA	17.47	70.03	3050	
LA ENCAÑADA	7.12	78.32	2950	
SELVA	IQUITOS	3.75	73.22	125
	PUERTO MALDONADO	12.60	69.22	265
	YURIMAGUAS	5.88	76.12	179
	PUCALLPA	8.37	74.57	148
	TARAPOTO	6.50	76.37	281
	TINGO MARÍA	9.30	76.02	664
	NAUTA	4.52	73.60	112
	IÑAPARI	10.95	69.60	250
TAMBOPATA	14.22	69.15	1264	

Figura 11: Mapa con ubicación de estaciones meteorológicas usadas en el trabajo.



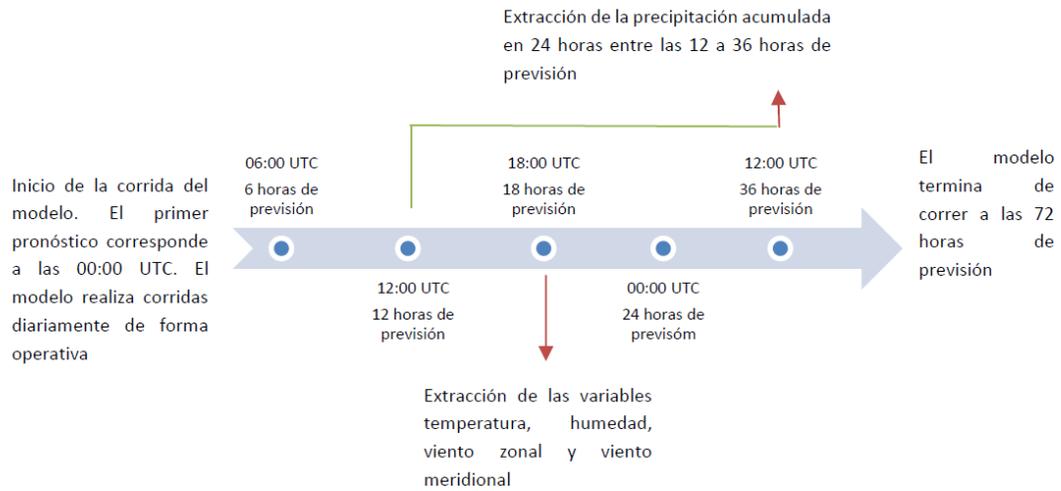
Fuente: Elaboración propia

3.2.2. INFORMACIÓN GENERADA POR EL MODELO ETA-SENAMHI

Los datos de modelo numérico que se utilizaron en el presente trabajo son resultados del modelo ETA-SENAMH de pronósticos correspondientes a las 18 horas de previsión para un análisis de las 00:00 UTC; es decir las variables extraídas corresponden al pronóstico de las 18:00 UTC. Esta información fue proporcionada por la Dirección de Meteorología Aplicada (DMA) del SENAMHI. Para extraer esta información se utilizó el dominio que cubre al Perú cuya área se encuentra ubicada entre los 85°W-56°W de longitud y 28°S-4°N y además posee una resolución espacial de 22 km en la horizontal y 38 niveles de presión. 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). Dicha información se puede descargar de manera gratuita en el siguiente link: <http://nomads.ncep.noaa.gov/>

Las variables que fueron generadas por modelo son: temperatura del aire, humedad relativa, viento meridional, viento zonal. Toda esta información se extrajo para los principales niveles de presión (ver Cuadro 5). Adicionalmente, se generó la precipitación acumulada en 24 horas, la cual corresponde a la precipitación acumulada entre las 12 UTC (7:00 am hora local) de un día hasta las 12 UTC (7:00 am hora local) del día siguiente, con el fin de que esta información se asemeje a los datos de estaciones; es decir, es la precipitación acumulada entre las 12 a 36 horas de previsión. La información de estas variables se extrajo a partir de la grilla más cercana a las coordenadas de las estaciones mencionadas en la sección anterior y para el mismo periodo de tiempo de estudio (Enero 2010 – Febrero 2013).

Figura 12. Esquema de extracción de variables predictoras para el modelo ETA-SENAMHI



Cuadro 5: Conjunto de datos extraídos del modelo ETA - SENAMHI.

Variables	Niveles de presión
Precipitación diaria acumulada	Superficie
Temperatura del aire	950, 925, 850, 700, 600, 500, 400, 300, 200
Humedad relativa	950, 925, 850, 700, 600, 500, 400, 300
Viento meridional	950, 925, 850, 700, 600, 500, 400, 300, 200
Viento zonal	950, 925, 850, 700, 600, 500, 400, 300, 200

Adicionalmente, en base a estos datos del modelo ETA-SENAMHI, se generaron datos de tres índices de estabilidad, definidos como indicadores de los movimientos convectivos verticales predominantes de una masa de aire. Los índices generados fueron:

- **Índice K:** Índice desarrollado por Whiting el cual es una medida del potencial de tormenta de una masa de aire, y se basa en el gradiente vertical de la temperatura y en

la cantidad y extensión vertical del contenido de humedad en la atmósfera inferior (Chourio, 2009). Se calcula mediante la siguiente expresión.

$$K = (T850-T500) + Td850-(T700-Td700)$$

- **Índice TT:** Introducido por Miller (1972), es utilizado para identificar potenciales áreas de desarrollo de tormentas en base a la estabilidad estática o el gradiente entre los niveles de 850 y 500, y la humedad a 850 hPa (Chourio, 2009).. Su valor se obtiene mediante la siguiente expresión:

$$TT = T850 + Td850 - 2 * T500$$

- **Índice W:** El valor del índice de humedad W representa la humedad contenida en los niveles de 850, 700 y 500 mb y se trabaja con las diferencias entre la temperatura y el punto de rocío en los 3 niveles antes mencionados (Chourio, 2009). Está dado por la siguiente expresión:

$$W = (T850-Td850) + (T700-Td700) + (T500-Td500)$$

3.2.3. DATOS DE REANALYSIS ERA-INTERIM

Se recolectó información correspondiente al periodo del 02 al 09 de febrero del 2013 del Reanálisis ERA-INTERIM el cual es un tipo de reanálisis global de observaciones climáticas registradas en los últimos 35 años y producido por el ECMWF (European Centre for Medium-Range Weather Forecasts). Esta información se encuentra disponible libremente en la siguiente página web: http://data-portal.ecmwf.int/data/d/interim_daily/

Estos datos, según sus características, cuentan con una resolución temporal de 4 horas en UTC por día (00 UTC, 06 UTC, 12 UTC y 18 UTC), una resolución espacial de 1.5° x 1.5° en formato global y además posee 37 niveles de presión para la mayoría de las variables. El área

que comprendió la descarga de los datos está comprendida desde 60° S a 30° N de latitud y desde los 120° W a 2° E de longitud, abarcando, de esta manera, Sudamérica.

Las variables que se descargaron fueron: temperatura del aire, viento meridional, viento zonal, humedad relativa y humedad específica para los niveles isobáricos de 1000, 925, 850, 700, 500, 300 y 200 hPa

3.2.4. IMÁGENES DE SATÉLITE.

Adicionalmente a la información del reanálisis ERA-INTERIM, se utilizaron también imágenes de satélite del GOES 13 de Sudamérica para las 00:00 Z, 12:00 Z y 18:00 Z, del periodo del 02-09 de Febrero del 2013, y de los siguientes canales:

- Canal 1: Imagen Visible (0.52 - 0.71 μm).
- Canal 4: Imagen Infrarrojo (10.2 - 11.2 μm).
- Canal 3: Imagen Vapor de agua (6.47-7.02 μm).

Estas imágenes se encuentran disponibles de manera gratuita en el Banco de Imágenes del Centro de Previsión, Tiempo y Estudios Climáticos CPTEC. (<http://satellite.cptec.inpe.br/home/novoSite/index.jsp>).

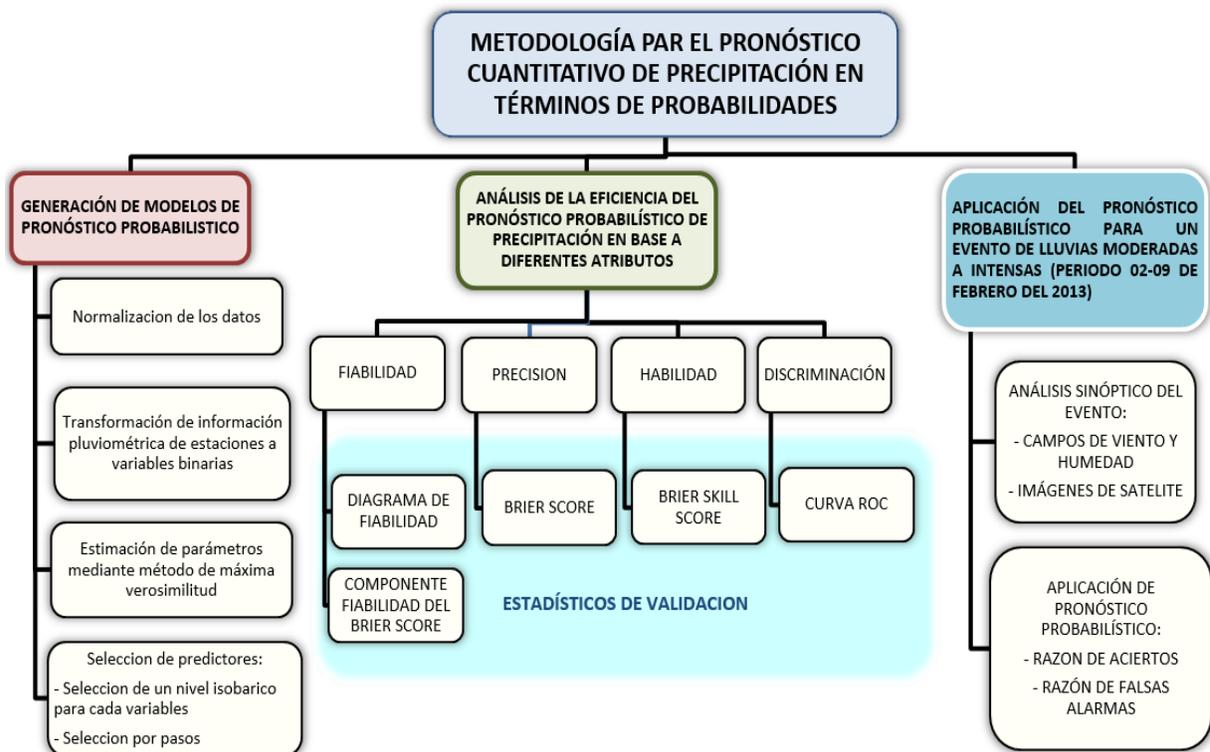
3.2.5. SOFTWARES

- La extracción y simulaciones del modelo ETA-SENAMHI se realizó en entorno Linux Centos mientras que el análisis de los resultados se hizo en Windows.
- Microsoft Office 2010
- Fortran 90
- GrADS 2.0.2
- Octave
- R x64 versión 3.1.1.
- Rstudio

3.3. ENFOQUE METODOLÓGICO.

En la presente sección se explica las diferentes técnicas y procedimientos que se utilizaron para cumplir con los objetivos de la presente investigación. El método estadístico usado para establecer la relación entre predictores y predictando es el de regresión logística ya que, acorde con Applequist et al. (2002) es el mejor método de regresión en pronóstico probabilístico y para la calibración de estos modelos se utilizó los datos en el periodo de enero del 2010 a enero del 2013. Con las ecuaciones de regresión establecidas, se procedió a evaluar los pronósticos probabilístico en base a los de diferentes atributos que tienen este tipo de pronósticos: confiabilidad, habilidad, precisión y resolución. Por último, se aplicaron los pronósticos probabilístico diarios en un evento de intensas lluvias ocurrido durante el periodo del 02 al 09 de febrero del año 2013. En la figura 13 se presenta el esquema detallado de la metodología del trabajo:

Figura 13: Esquema de la metodología de la presente investigación



3.3.1. METODOLOGÍA PARA EL PRONÓSTICO CUANTITATIVO DE LA PRECIPITACIÓN EN TÉRMINO DE PROBABILIDADES.

Los modelos de regresión logística generados para el pronóstico probabilístico diarios de precipitación poseen la siguiente representación matemática:

$$\Pr\{y\} = \frac{1}{1+e^{-(\beta_0+\beta_1x_1+\beta_2x_2+\dots+\beta_kx_k)}} \dots\dots\dots (4)$$

Donde:

$\Pr\{y\}$ = Predictando, siendo esta la probabilidad de que ocurra una precipitación mayor a 0.01 mm, 1 mm, 2.5 mm ó 5 mm acumulada en el intervalo de 7:00 am de un día a 7:00 am del día siguiente.

x_n = Predictores, los cuales son variables de pronóstico generados por el modelo ETA-SENAMHI.

β_k = Parámetros de la regresión.

k = Número de predictores

La creación de los modelos de regresión logística tiene como componentes fundamentales la estimación de los parámetros β_k de la regresión y la obtención del conjunto de predictores adecuados para cada modelo de regresión. Previamente, se tuvo que realizar un análisis de control de calidad de la información de estaciones el cual consistió principalmente en la detección de datos anómalos que pudieran afectar en la calibración del modelo. Posteriormente se crearon los modelos de regresión a nivel mensual para cada estación, siguiendo los pasos que se detallan a continuación:

a) Normalización de los datos

El primer paso para la creación de los modelos de regresión es la normalización de los datos de los predictores. Este proceso es importante ya que los predictores pueden tener diferentes magnitudes físicas y pueden oscilar sobre diferentes órdenes de magnitud. (Applequist *et al.*,

2002). Cada potencial predictor, que proviene de las simulaciones generadas por el modelo ETA-SENAMHI, fueron normalizados usando la siguiente expresión:

$$X_k(\textit{normalizado}) = \frac{X_k - \bar{X}_k}{S_d} \dots\dots\dots (5)$$

De donde:

x_k = Valor k-ésimo predictor potencial

\bar{x}_k = Promedio mensual de k - ésimo predictor potencial

S_d = Desviación estándar para el k-ésimo predictor potencial

b) Transformación de información pluviométrica de estaciones a variables binarias.

La regresión logística es especialmente útil en particular cuando solo hay dos posibles respuestas; es decir, cuando la variable de respuesta es dicotómica, que es el caso más común. Las variables del tipo continua, como la precipitación acumulada diaria, puede transformarse a variable dicotómica en base a la definición de diferentes umbrales, de tal forma que si la precipitación supera cierto umbral se puede asignar valores de 1 y 0 para los que estén debajo del umbral. Por lo tanto, el propósito de los modelos de regresión logística consiste en determinar las probabilidades en las que ocurra una precipitación acumulada diaria mayor a un umbral.

Con el fin de evaluar la eficiencia del pronóstico probabilístico de diferentes cantidades de precipitación, se establecieron cuatro categorías de pronósticos probabilísticos de precipitación a partir de cuatro diferentes umbrales, los cuales son mostrados en el cuadro seis. Los datos mayores o iguales a cada uno de estos umbrales fueron recodificados como 1 y los menores como 0 para cada una de las cuatro categorías.

La categoría uno (precipitación acumulada diaria mayor a umbral de 0.01 mm), la cual es la más simple de todas, indica únicamente la ocurrencia o no ocurrencia de precipitación, sin tomar en cuenta la cantidad. Esta categoría es parte del pronóstico de muchos servicios meteorológicos y es conocida con el nombre de PoP (Probability of Precipitation). Las

siguientes categorías que contienen umbrales mayores de precipitación fueron escogidas con el fin de obtener una estimación de la cantidad de precipitación acumulada diaria. Los pronósticos para estas categorías indican una estimación de la cantidad, se usa en distintos servicios meteorológicos bajo el nombre de PQPF (Probability Quantitative of Precipitation Forecast); sin embargo, estos umbrales pueden diferir de acuerdo a las características climáticas de cada región.

En este trabajo, se establecieron los umbrales mostrados en el cuadro seis, para obtener una primera aproximación del rendimiento de los modelo de regresión logística en el pronóstico probabilísticos de cantidades de precipitación. Se podría trabajar con umbrales mayores de precipitación; sin embargo, diversos estudios (Applequist, 2002; Matsudo, 2013; Gutierrez, 2004) hacen de cuenta que los pronósticos probabilísticos para mayores umbrales de precipitación conducen a una mayor incertidumbre y error del pronóstico.

Es importante resaltar que no todas las categorías se aplicaron para todas las estaciones del Perú. Para las estaciones que se encuentran ubicadas en la región de la costa, se elaboraron modelos de regresión únicamente para la categoría 1, debido a que sus condiciones climáticas dificulta la ocurrencia de precipitaciones acumuladas diarias mayores a 1 mm. En contraste con las estaciones de costa, las condiciones son favorables para que puedan ocurrir eventos de precipitaciones abundantes de precipitación en la sierra y selva.

Cuadro 6: Categorías de pronóstico probabilístico

Categorías	Definición
1	Precipitación acumulada en 24 horas ≥ 0.01 mm
2	Precipitación acumulada en 24 horas ≥ 1 mm
3	Precipitación acumulada en 24 horas ≥ 2.5 mm
4	Precipitación acumulada en 24 horas ≥ 5 mm

c) Estimación de parámetros mediante método de máxima verosimilitud

Para la estimación de los parámetros de los modelos de regresión logística se utilizó el método de máxima verosimilitud usando la información tanto del modelo ETA-SENAMHI y de las estaciones meteorológicas en el periodo de enero del 2010 a enero del 2013. Este método propuesto por Wilks (2006), consiste en encontrar valores de la distribución de parámetros que maximice la función de verosimilitud. El proceso se inicia construyendo una función de verosimilitud de la ecuación de regresión logística, la cual se debe maximizar y está representado por la siguiente forma:

$$L(\beta) = \prod_{i=1}^N \frac{y_i \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k) + (1 - y_i)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)} \dots (6)$$

De donde:

y_i = Datos pluviométricos de estación en representación binaria, siendo 0 cuando no llueve y 1 cuando sí llueve.

x_1, x_2, \dots, x_k = Predictores

n = Tamaño de la muestra.

k = Número de variables predictoras.

$\beta_0, \beta_1, \dots, \beta_k$ = Parámetros de la regresión

Dado que la variable y_i es binaria, cada factor en la ecuación (6) para lo cual $y_i = 1$ es igual a la probabilidad p_i , y los factores para lo cual $y_i = 0$ son iguales a $1 - p_i$. Es más conveniente la estimación de los parámetros a partir del logaritmo de la función de verosimilitud:

$$LL(\beta) = Ln(L(\beta)) = \sum_{i=1}^n \{y_i (\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k) - \ln[1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)]\} \dots (7)$$

El siguiente paso es estimar aquellos valores de β que hacen máxima la función de verosimilitud L (o su logaritmo LL). El cálculo de los parámetros se puede realizar mediante

el método iterativo de Newton-Raphson teniendo en cuenta que si se iguala a cero la derivada parcial de una función respecto a un parámetro, el resultado va a ser un valor de dicho parámetro que conducen a la función a un valor máximo o mínimo, es decir, un punto de inflexión de la curva. El cálculo de la segunda derivada de la función con respecto a dicho parámetro nos va a confirmar si el punto de inflexión se trata de un máximo y no de un mínimo, debiendo ser este menor a 0:

La primera derivada de $LL(\beta)$ respecto de β en su forma matricial se expresa de la siguiente forma:

$$\frac{\partial LL(\beta)}{\partial \beta} = X'(Y - p) \dots \dots \dots (8)$$

De donde:

p = Matriz de n filas y una columna que contiene las probabilidades para cada dato de que tengan su correspondiente evento p_i

X = Matriz que contiene a los predictores x_1, x_2, \dots, x_k

β = Matriz que contiene los parámetros de regresión $\beta_0, \beta_1, \dots, \beta_n$

La segunda derivada, en su forma matricial se expresa de la siguiente forma:

$$\frac{\partial^2 LL(\beta)}{\partial \beta \partial \beta} = -X'WX \dots \dots \dots (9)$$

De donde:

W = Matriz diagonal ⁴ de n filas y n columnas en la que los elementos de su diagonal están dado por $p_i(1 - p_i)$, de tal forma que se expresa de la siguiente manera:

⁴ Matriz cuadrada en la que todos los valores de sus elementos son 0 excepto su diagonal.

$$W = \begin{pmatrix} p_1(1-p_1) & 0 & 0 & 0 \\ 0 & p_2(1-p_2) & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n(1-p_n) \end{pmatrix} \dots (10)$$

Y para cada fila su p_i es:

$$p_i = \frac{1}{1+e^{-(\beta_0+\beta_1x_1+\beta_2x_2+\dots+\beta_kx_k)}} \dots\dots\dots(11)$$

Una vez se dispone del cálculo de las derivadas, se aplicó el método iterativo para la determinación de los parámetros de la regresión. Para ello, se asignó un valor inicial empírico a los parámetros de regresión, como por ejemplo cero. En cada iteración t la nueva matriz de parámetros experimentales se obtiene de sumar matricialmente un gradiente a la matriz de coeficientes experimentales del paso anterior. El gradiente resulta del cociente entre la primera derivada (ecuación 8) y la segunda derivada (ecuación 9) de la función de verosimilitud de la ecuación de regresión.

$$\hat{\beta}_t = \hat{\beta}_{t-1} + (X'W_{t-1}X)'X'(Y - p_{t-1}) \dots\dots (12)$$

Las iteraciones se repite tantas veces como sea necesario hasta que la diferencia entre la matriz de coeficientes de regresión de la iteración t y la matriz de la iteración $t-1$, sea cero o muy cercano a cero. De esta forma, los resultados de las iteraciones convergen a un vector formado por los parámetros de la regresión que se ajusten mejor en el pronóstico probabilístico.

La cantidad de parámetro a estimar va a depender del número de predictores a usar; así por ejemplo si se trabaja con dos predictores, la ecuación (12) se puede expresar como en (13), en la cual se debe calcular un total de tres parámetros:

$$\begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \beta_2^* \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} - \begin{bmatrix} \sum p_i^2 - p_i & \sum x_1(p_i^2 - p_i) & \sum x_2(p_i^2 - p_i) \\ \sum x_1(p_i^2 - p_i) & \sum x_1^2(p_i^2 - p_i) & \sum x_1x_2(p_i^2 - p_i) \\ \sum x_2(p_i^2 - p_i) & \sum x_2x_1(p_i^2 - p_i) & \sum x_2^2(p_i^2 - p_i) \end{bmatrix}^{-1} \begin{bmatrix} y - p_i \\ x_1(y - p_i) \\ x_2(y - p_i) \end{bmatrix} \dots(13)$$

Donde:

$\beta_0, \beta_1, \beta_2$ = Parámetro genérico.

$\beta_0^*, \beta_1^*, \beta_2^*$ = Verdaderos valores de parámetros para ser aproximado.

p_i = Probabilidad en un tiempo i.

x_1 = Valor de la primera variable predictora

x_2 = Valor de la segunda variable predictora

y = Datos pluviométricos de estación en representación binaria, cero cuando no precipita y uno cuando sí precipita.

La estimación de los parámetros de los modelos de regresión logística a partir del método de máxima verosimilitud, se realizó mediante un paquete de librería abierto en lenguaje R (versión 3.1.1) llamado *Bias reduction in binomial-response generalized linear models* (brglm) ⁵ cuya principal función es el ajuste de modelos de regresión y la estimación de parámetros de regresión.

Los cálculos de los parámetros se realizaron de tal manera que hubiera un conjunto de parámetros para cada mes y para cada estación; es decir, doce modelos de regresión (mensuales) para cada estación, sumando un total de 480 modelos de regresión logística (40 estaciones x 12 meses) para cada categoría de pronóstico. Para ello se agrupó la data diaria de cada mes para todos los años y se calculó los modelos de regresión a nivel mensual. Por ejemplo, para la calibración de un modelo de regresión correspondiente al mes de enero, se agrupó la información de enero del 2010, 2011, 2012 y 2013 en una sola serie por variable esto para los datos de estación (predictando) como también para los datos de cada variable del modelo ETA-SENAMHI (predictores), y con estas series ordenadas se aplicó el método iterativo para la estimación de los parámetros de regresión.

⁵ Disponible en: <https://cran.r-project.org/web/packages/brglm/index.html>

d) Método de selección de predictores.

Existe aproximadamente un total de 40 potenciales predictores para cada modelo de regresión logística. La cantidad de predictores o variables independientes que puede tener el modelo dependen de la cantidad de observaciones que tengamos. Según Freeman (1987) para un total de k predictores, debemos tener como mínimo $10*(k+1)$ observaciones para satisfacer los supuestos de regresión logística; es decir, al menos 10 casos por cada variable que intervienen en el modelo, lo cual nos da la posibilidad de tener un total de ocho predictores para cada modelo⁶. Sin embargo; Schmeits *et al* (2008) recomienda usar un máximo de cinco predictores para cada modelo de regresión ya que más predictores puede ocasionar sobreajustes. Bajo este enfoque se estableció que cada modelo de regresión logística tenga un máximo de cinco predictores.

La selección de los predictores se realizó en dos fases. La primera selección de predictores consistió en escoger un nivel isobárico para las variables de temperatura, humedad relativa, viento zonal y viento meridional; mientras que la segunda fase consistió en la aplicación un método de selección por pasos.

- **Primera fase de selección de predictores: selección de un nivel isobárico para cada variable predictor.**

Las variables humedad relativa, temperatura del aire, viento meridional y viento zonal se encuentran en diferentes niveles de presión de acuerdo a la altitud en la capa atmosférica. El comportamiento de estas variables no es igual en cada nivel isobárico y puede ejercer un rol diferente en la precipitación. Así por ejemplo tenemos que el viento zonal en 500 hPa influye en el transporte de humedad, mientras que en 200 hPa puede influir mayormente en la inestabilidad.

⁶ Considerando que las ecuaciones son a nivel mensual para tres años (enero 2010-enero 2013) tenemos una cantidad de 90 observaciones aproximadamente para cada modelo de regresión.

En esta fase se escogió un solo nivel isobárico para cada una de las cuatro variables mencionadas anteriormente. Esta elección se realizó mediante la correlación de Spearman entre los datos de precipitación observada y los predictores de humedad relativa, temperatura del aire, viento meridional y viento zonal. A su vez se comprobó que las variables que estén más correlacionadas tengan coherencia física con la ocurrencia de precipitación

Correlación de Spearman (rho de Spearman)

Según Pita & Pértega (1997) este coeficiente del tipo no paramétrico es una medida de asociación lineal que utiliza los rangos, números de orden, de cada grupo de sujetos y compara dichos rangos. Los valores de este coeficiente varían de -1 a 1 de tal forma que valores próximos a 1; indican una correlación fuerte y positiva. Valores próximos a -1 indican una correlación fuerte y negativa. Valores próximos a cero indican que no hay correlación lineal. La expresión matemática para el cálculo es la siguiente:

$$\rho = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2-1)} \dots\dots\dots (14)$$

De donde:

D=Diferencia entre los rangos (Xi-Yi)

n= Número de observaciones

La ventaja de este coeficiente de correlación es que la distribución de ambas variables a analizar, no necesariamente debe asemejarse a la curva normal (Badii et. al, 2014); es decir, no hay alguna restricción con respecto al tipo de distribución que posean las variables a analizar. Esto es favorable a la hora de estimar la correlación de precipitación con otras variables atmosféricas, ya que no posee una distribución del tipo normal. Por lo tanto, para cada variable se escogió el nivel de presión en el que posea el mayor valor absoluto de la correlación de Spearman con la precipitación observada, teniendo en cuenta además de la influencia física que tengan estos en la ocurrencia de precipitación.

En base a esta primera selección la cantidad de posible predictores totales para cada modelo de regresión logística se redujo a un total de 9 predictores: precipitación acumulada (PP_{acum}), humedad relativa (RH), temperatura (T), viento zonal (U) y viento meridional (V), además los índices de estabilidad Ki, TT y W, tal y como se muestra en el siguiente modelo:

$$\Pr\{y\} = \frac{1}{1+e^{-(\beta_0+\beta_1RH_{n1}+\beta_2T_{n2}+\beta_3U_{n3}+\beta_4V_{n4}+\beta_5PP_{acum}+\beta_6Ki+\beta_7TT+\beta_8WW)}} \dots(15)$$

De donde:

$\Pr\{y\}$ = Probabilidad de que ocurra una precipitación mayor a 0.01 mm, 1 mm, 2.5 mm ó 5 mm.

$n1, n2, n3$ y $n4$ = Niveles de presión de los predictores RH, T, U y V respectivamente.

$\beta_0, \beta_1, \dots, \beta_8$ = Parámetros de la regresión.

- Segunda fase de selección de predictores: método de selección por pasos

Posteriormente de haber escogido un nivel isobárico para cada variable, el siguiente procedimiento usado para seleccionar las variables independientes fue el de selección por pasos. Este método consiste en construir sucesivos modelos de tal manera que cada uno difiera del precedente en una sola variable y en cada paso se compara los resultados de cada versión con los de la anterior en función de su significancia estadística en la explicación de la variable respuesta. El criterio para la selección de predictores fue mediante el cálculo de las desviación, el cual se define a continuación:

Desviación

Es considerada como una medida de ajuste del modelo y de los datos. Supongamos que tenemos un modelo llamado $m1$ con p parámetros y otro modelo $m2$ con q parámetros (para n datos y con la misma respuesta) y supongamos además que $p < q$. Entonces la prueba consiste

en comparar las funciones de verosimilitud de estos dos modelos para decidir cuál es el mejor. Su expresión matemática está dada por:

$$D_{m1} = -2 (\log(L(\beta_{m1})) - \log(L(\beta_s))) = -2 \log\left(\frac{L(\beta_{m1})}{L(\beta_s)}\right) \dots\dots(16)$$

$$D_{m2} = -2 (\log(L(\beta_{m2})) - \log(L(\beta_s))) = -2 \log\left(\frac{L(\beta_{m2})}{L(\beta_s)}\right) \dots\dots(17)$$

De donde:

$L(\beta_{m1})$: Función de verosimilitud del modelo $m1$ con p variables predictoras

$L(\beta_{m2})$: Función de verosimilitud del modelo $m2$ con q variables predictoras

$L(\beta_s)$: Función de verosimilitud del modelo saturado; es decir, el modelo con máximo número de parámetros que puede ser estimado.

En general, a menores valores de desviación, el modelo se ajusta mejor a los datos. Este estadístico no es usualmente interpretado directamente, por lo que es preferible comparar desviaciones de otros modelos ajustados a la misma data. La comparación se procede mediante la diferencia de estas tal y como se muestra a continuación:

$$D_{m1} - D_{m2} = -2 \log\left(\frac{L(\beta_{m1})}{L(\beta_s)}\right) + 2 \log\left(\frac{L(\beta_{m2})}{L(\beta_s)}\right)$$

$$D_{m1} - D_{m2} = -2 \log(L(\beta_{m1})) + 2 \log(L(\beta_s)) + 2 \log(L(\beta_{m2})) - 2 \log(L(\beta_s))$$

$$D_{m1} - D_{m2} = -2 \log(L(\beta_{m1})) + 2 \log(L(\beta_{m2}))$$

$$D_{m1} - D_{m2} = -2 \log\left(\frac{L(\beta_{m1})}{L(\beta_{m2})}\right) \dots\dots(18)$$

Esta diferencia de desviaciones se puede comparar con una distribución chi-cuadrado de muestra grande con grados de libertad igual a la diferencia del número de parámetros

estimados en el modelo $m1$ y el modelo $m2$, con un nivel deseado de significancia (Hogg y Craig, 1994).

$$D = D_{m1} - D_{m2} \sim \chi^2_{p-q, \alpha} \dots (19)$$

La hipótesis nula de esta prueba es que no existe mejora estadísticamente significativa en el modelo ajustado después de añadir un predictor potencial. En adición, un p-valor resultante menor o igual a 0.05 es requerido para ser considerado estadísticamente significativo (en un nivel de 95% de confianza) y rechazar la hipótesis nula. De esta forma, cada predictor fue añadido a los modelos de regresión hasta que ninguna de las variables restantes sea significativa. Por lo tanto, este método se cubre comparando modelos simples (el más simple de los cuales es el modelo que contiene solo el parámetro de intercepción) con modelos más complejos (que incluyen otros parámetros además de la intercepción) y decidiendo con el criterio de desviaciones si un modelo más complejo mejoran significativamente el ajuste producido por modelos más simple (Arnau, 1996).

En base a la selección por pasos se obtuvieron las variables significativas que se presentaron con mayor frecuencia en los modelos de regresión logística y que fueron escogidos para los modelos de regresión logística finales. Debido a que los factores que influyen en la ocurrencia de precipitación son diferentes para cada región de nuestro país se agruparon estos resultados en costa, sierra y selva, de tal forma que se escogieran un máximo de cinco predictores que tuvieran las más altas frecuencia de significancia estadística al 95% de confianza. Este proceso se realizó también teniendo como criterio la influencia física que tengan los predictores en la ocurrencia de precipitación de cada lugar de acuerdo a los patrones de circulación.

3.3.2. ANÁLISIS DE LA EFICIENCIA DEL PRONÓSTICO PROBABILÍSTICO DE PRECIPITACIÓN

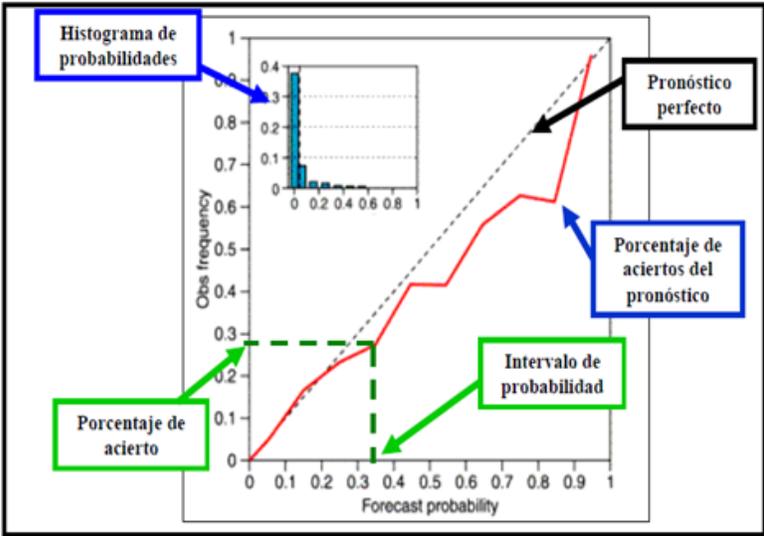
La verificación de los pronósticos probabilísticos se fundamenta en la evaluación de diferentes atributos que influyen en la calidad de un pronóstico tales como por ejemplo la fiabilidad, precisión, habilidad y resolución; ya que no basta con medir solamente el error. Los sistemas

de predicción probabilística exigen la verificación de sus diferentes atributos mediante técnicas estadísticas que son diferentes a las de las predicciones deterministas. Por este motivo, no existe una única medida que sea capaz de brindar una información completa de la calidad de estos pronósticos. A continuación se presenta la elección de las medidas y gráficos que analizan la correspondencia entre las probabilidades pronosticadas y los datos observados.

a) Análisis de la fiabilidad

El atributo de la fiabilidad es una medida de que tan cercano los pronósticos probabilísticos se corresponden a las frecuencias condicionales de los eventos. Este atributo se mide a partir del “Diagrama de fiabilidad” y se grafica dibujando la curva de probabilidad prevista frente a la probabilidad observada para intervalos de 5% o 10%. Este diagrama nos va a indicar que tan bien la probabilidad pronosticada de un evento se corresponde a su frecuencia observada. Se fundamenta en la idea de que, por ejemplo, una probabilidad de pronóstico del 0.3 se tendría que cumplir en el 30% de los casos pronosticados con dicha probabilidad. El ‘eje x’ representa los intervalos de probabilidad de pronóstico y en el ‘eje y’ la frecuencia relativa de aciertos de los pronósticos incluidos en cada uno de los intervalos de probabilidad.

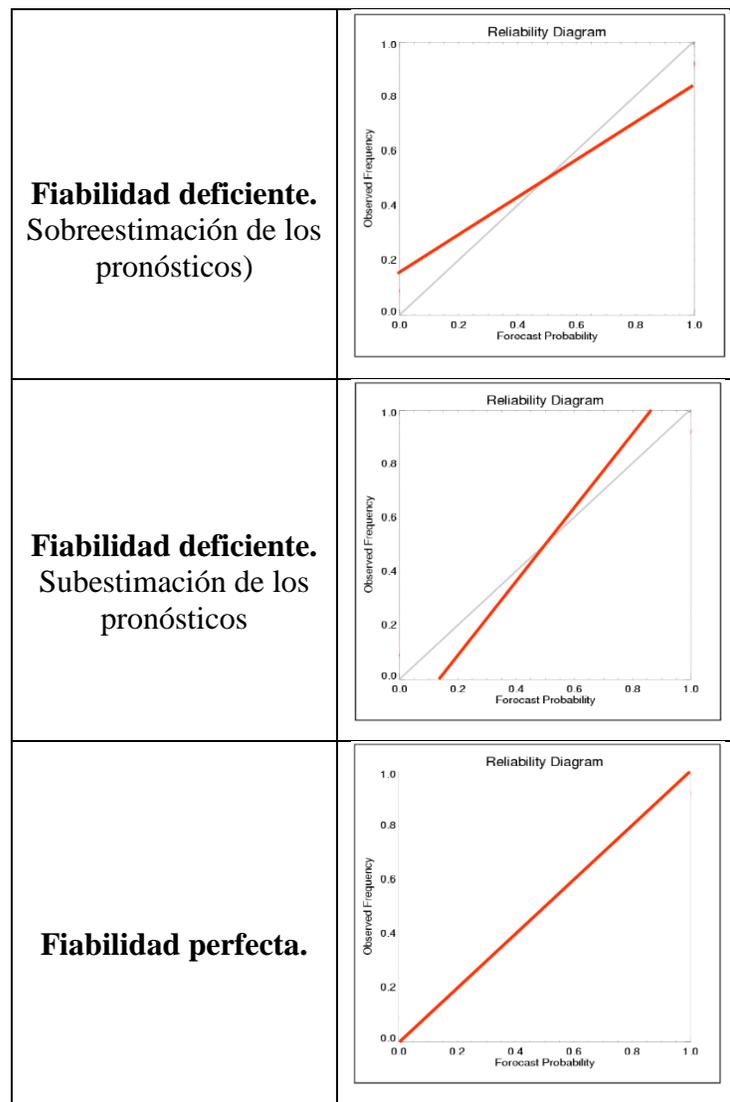
Figura 14: Ejemplo de diagrama de fiabilidad



Fuente: Medeiro (2007)

Para una fiabilidad perfecta el pronóstico de probabilidad y la frecuencia de ocurrencia deben ser iguales, y los puntos de la curva deben estar en la línea diagonal, por lo tanto, la fiabilidad del pronóstico probabilístico será mejor cuanto más se acerque a la gráfica de la fiabilidad a la diagonal principal. Si el diagrama de fiabilidad queda por debajo de la diagonal, las probabilidades de pronóstico son sobreestimadas, mientras que si queda por encima de la diagonal las probabilidades de pronóstico serían subestimadas (véase figura 15).

Figura 15: Diagramas de fiabilidad de acuerdo a la calidad del pronóstico probabilístico



Fuente: Hagerdorn, 2007.

Con el objetivo de tener una mejor representación de la fiabilidad del pronóstico probabilístico para todas las estaciones, se elaboró un diagrama de fiabilidad correspondiente a cada región del país y para cada categorías de pronóstico, construidos en base a los datos observados de precipitación y las probabilidad pronosticadas para todo el periodo de estudio. Así mismo, se elaboraron diagramas de fiabilidad correspondiente a los meses de Diciembre-Enero-Febrero y Junio-Julio-Agosto con el fin de conocer en qué estación del año los resultados son más fiables.

b) Análisis cuantitativo de la fiabilidad, precisión y habilidad.

El cuadro 7 resume los estadísticos utilizados para la evaluación numérica de los atributos de fiabilidad, precisión y habilidad; así como sus principales características.

La **precisión** es una medida de la concordancia entre el valor pronosticado y el observado, promediado sobre una muestra de parejas individuales de pronósticos y observaciones. Para pronósticos probabilístico, la medición de la precisión se realiza mediante el Brier Score, de tal forma que valores de este índice más cercanos a 0 indican una mayor precisión de los resultados

La **habilidad** es una medida de la precisión media de los pronósticos respecto de los pronósticos dados por un método de referencia, los cuales pueden ser métodos basados en persistencia y climatología y que pueden utilizarse como métodos de predicción probabilística. El índice utilizado para la medición de habilidad es el Brier Skill Score y se interpreta de tal forma que un valor positivo (negativo) indica que el sistema es más (menos) hábil que el de referencia.

Cuadro 7: Atributos analizados e indicadores estadísticos del desempeño del modelo

Atributo	Descripción	Estadístico	Formula
FIABILIDAD	Diferencia entre el pronóstico y el promedio de observaciones asociados con ese valor de pronóstico	Componente fiabilidad del Brier Score	$Fiabilidad = \frac{1}{N} \sum_{k=0}^M N_k (f_k - \bar{o}_k)^2$ <p>De donde: f_k = Probabilidad pronosticada o_k = Dato observado binario (1 cuando el evento ocurrió y 0 si no ocurrió) n = Número total de parejas de datos Pronóstico perfecto: Fiabilidad = 0</p>
PRECISIÓN	Correspondencia media entre las parejas individuales de pronóstico y datos observados	Brier Score	$BS = \frac{1}{n} \sum_{k=1}^n (f_k - o_k)^2$ <p>De donde: f_k = Probabilidad pronosticada o_k = Dato observado binario (1 cuando el evento ocurrió y 0 si no ocurrió) n = Número total de parejas de datos. Pronóstico perfecto: BS = 0 Pronóstico deficiente: BS = 1</p>
HABILIDAD	Precisión de los pronósticos de interés con respecto a un pronóstico de referencia tales como la climatología o la persistencia	Brier Skill Score	$BSS = 1 - \frac{BS}{BS_{cli}}$ <p>De donde: BS = Brier Score BScli = Brier Score de la climatología Pronóstico perfecto: BSS = 1 Pronóstico deficiente: BS < 0</p>

Medidas de validación de predicciones binarias.

Otra forma de realizar la verificación de los pronósticos probabilísticos, es transformando estos valores de pronóstico a variables dicotómicas (si/no). Dada la predicción probabilística para una variable binaria, se puede obtener una predicción determinista considerando un umbral u_p para la probabilidad. Cuando la probabilidad prevista supera ese umbral, se predice la ocurrencia del evento y en caso contrario se predice la no ocurrencia. Los pronósticos probabilísticos fueron transformados a valores binarios para conseguir un pronóstico de ocurrencia o no-ocurrencia del evento y son verificadas mediante tablas de contingencia, que se construyen combinando todas las posibilidades entre categoría prevista y observada.

Cuadro 8: Tabla de contingencia para pronósticos binarios.

		PRONÓSTICO	
		SI	NO
OBSERVACIÓN	SI	O1	NO1
	NO	O2	NO2

Donde evidentemente O1, O2, NO1, NO2 dependen de u_p . A partir de los valores de las tablas de contingencia, se pueden obtener las siguientes medidas de la calidad del pronóstico:

– **Razón de aciertos (Hit Rate).**

Se interpreta como la fracción del número de ocurrencia de eventos pronosticados correctamente con respecto al número total de casos en que se dio el evento. Mide la capacidad de predecir eventos que realmente se han observado. Una mejor razón de acierto va a ser cercano a 1 mientras que cercano a 0 va a ser peor.

$$H = \frac{O1}{O1 + O2}$$

– **Razón de falsas alarmas (False Alarm Rate)**

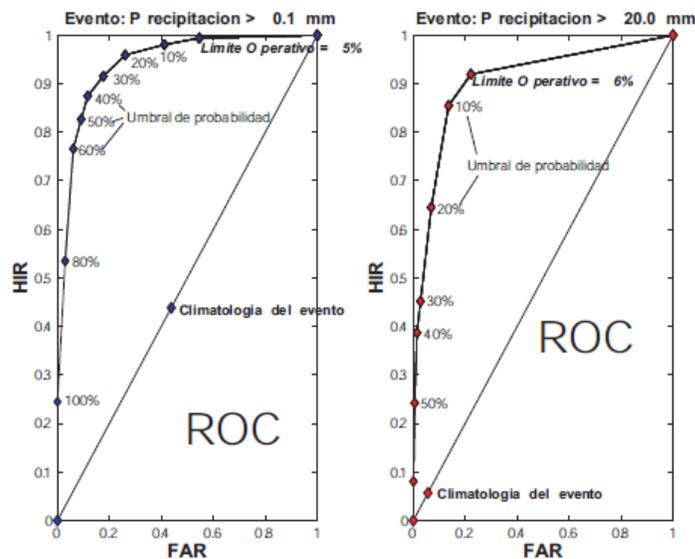
Mide la proporción de fallos cuando se predice positivamente un evento. Es la frecuencia relativa condicional de un pronóstico errado dado que el evento no ocurre. Valores cercanos a 0 nos indica buenos pronósticos y cercanos a 1 son malos.

$$F = \frac{NO1}{NO1 + NO2}$$

c) Análisis de la resolución: Curva ROC.

La resolución es la habilidad del pronóstico probabilístico para discriminar entre eventos y no-eventos mediante y una forma de medirla es mediante la curva ROC (Receiver Operating Characteristic). Esta curva se obtiene representando la razón de falsas alarmas (FAR) frente la razón de aciertos (HIT) para diferentes umbrales de probabilidad entre 0 y 1 (por ejemplo $u_p = 0, 0.1, 0.2, \dots, 1$).

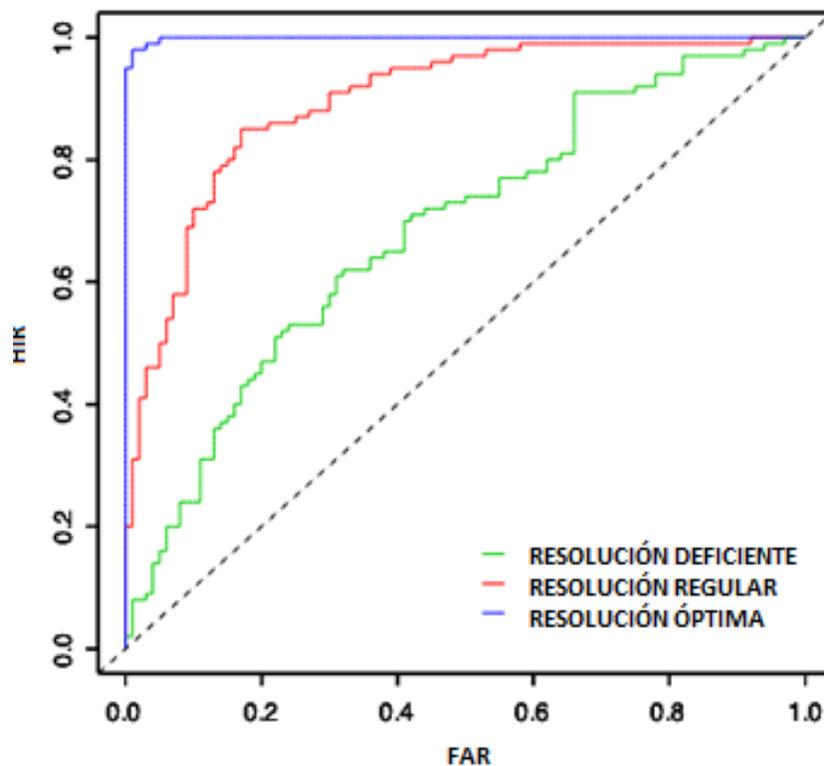
Figura 16: Dos ejemplos de curvas ROC para diferentes eventos.



Fuente: Ancell (2009)

Un buen modelo de pronóstico de probabilidad daría los valores de la curva más cercana a la esquina superior izquierda (alta tasa de aciertos y baja tasa de falsas alarmas). El modelo de pronóstico va a tener resolución deficiente mientras la curva ROC se acerque más a la diagonal. Esta gráfica es importante además, para encontrar el umbral u_p que mejor discrimina entre ocurrencia y no ocurrencia del evento para cada estación ya que el mejor sistema será aquel cuya tasa de aciertos sea más alta para un mismo nivel de tasas de falsas alarmas.

Figura 17: Tipos de curva ROC según la resolución del pronóstico.



Fuente: Modificado de Weiss, 2006.

3.3.3. APLICACIÓN DEL PRONÓSTICO PROBABILÍSTICO DIARIO DE PRECIPITACION EN UN CASO DE ESTUDIO.

El desempeño del pronóstico probabilístico de precipitación fue evaluado durante un periodo de días lluvias intensas que se encuentra comprendida del 02 al 09 de febrero del 2013, en la que se tienen registros de fuertes precipitaciones (ver ANEXO 8).

El análisis de los factores físicos de entorno sinóptico asociados a estos eventos de precipitación es de vital importancia con el fin de verificar si los resultados de los pronósticos probabilísticos representan con altas (bajas) probabilidades las regiones donde las condiciones fueron más (menos) favorables para la ocurrencia de lluvias.

El análisis sinóptico se realizó mediante información del Reanalysis ERA-INTERIM, elaborando campos de líneas de corriente y flujo de razón de mezcla para Sudamérica y Perú, principalmente en niveles medios y altos, mediante el software Grads. Además, se complementó con la información satelital del GOES-13 para identificar las zonas con alta actividad convectiva.

Posteriormente, los resultados de probabilidad de precipitación y de los datos observados fueron visualizados en forma puntual. Para ellos se elaboraron mapas de pronóstico probabilístico diario de precipitación, en el software Grads, para las cuatros categorías de pronóstico (umbrales de precipitación de 0.01 mm, 1.0 mm, 2.5 mm y 5 mm) y que fueron contrastados con los mapa de datos observados binarios de la ocurrencia de precipitación. Para la medición en forma cuantitativa de la eficiencia en el pronóstico, se calcularon la cantidad de aciertos y falsas alarmas a partir de las tablas de contingencia elaboradas para ese evento y agrupados por regiones costa, sierra y selva.

IV. RESULTADOS Y DISCUSIONES

4.1. RESULTADOS DE LA APLICACIÓN DE LA METODOLOGÍA PARA EL PRONÓSTICO PROBABILÍSTICO DE LA PRECIPITACIÓN.

El procedimiento para la elaboración del pronóstico probabilístico de la precipitación en base al ajuste estadístico MOS (corrección estadística de las salidas de los modelos numéricos para mejorar la eficiencia del pronóstico), dependió en gran medida de la selección de predictores óptimos que se usaría en los modelos de regresión logística a nivel mensual para cada punto de estación. A continuación se describen los resultados obtenidos de la selección de predictores óptimos para los modelos de regresión logística de acuerdo a las dos fases de la metodología planteada.

4.1.1. RESULTADOS DE LA SELECCIÓN DEL NÚMERO DE PREDICTORES

a) Resultados de la primera fase de selección: selección de un nivel isobárico para cada variable predictor

Las figuras del 18 al 21 muestran las correlaciones de Spearman entre la precipitación observada y las variables del modelo ETA-SENAMHI: humedad relativa (figura 18), viento zonal (figura 19), viento meridional (figura 20) y temperatura del aire (figura 21), para distintos niveles de presión de la columna atmosférica y para el periodo de Diciembre a Abril (periodo con mayores frecuencias de precipitación). Se escogió un solo nivel de presión por cada variable predictor de tal forma que tuviera la mayor magnitud de correlación con la precipitación observada y que guarde relación física con la actividad convectiva en la estación.

De manera general, la humedad relativa (figura 18) presenta correlaciones positivas con la precipitación en todas las 40 estaciones, siendo estas de mayor magnitud (mayores a 0.6) en

niveles medios (700 hPa-400 hPa) para estaciones de sierra, lo cual podría estar asociado al ingreso de humedad en estos niveles proveniente de la amazonia.

Las correlaciones de viento zonal (figura 19) con la precipitación para las estaciones de costa indican fuertes magnitudes negativas en los niveles de 500-400 hPa y que estaría asociado al ingreso de humedad proveniente de la sierra que logró traspasar la cordillera. En las estaciones de sierra se muestra una fuerte correlación negativa para niveles de 200 y 300 hPa asociado a la circulación de la Alta de Bolivia que causa divergencia en estos niveles de presión, mientras que para estaciones de selva no existen niveles en común de fuertes correlaciones para todas las estaciones por lo tanto las mayores magnitudes depende de cada estación.

En el caso de viento meridional (figura 20) las correlaciones con la precipitación observada son más cercanas a cero, a excepción de algunas estaciones de sierra (ejemplo Sibayo e Imata) que muestra valores de -0.4 de correlación aproximadamente en niveles de 500-300hPa posiblemente relacionado al movimiento anticiclónico con componente meridional negativo que favorece el transporte de humedad en estos niveles.

Por último, las correlaciones entre precipitación y temperatura (figura 21) son casi nulas para estaciones de costa, mientras que para estaciones de sierra existen correlaciones positivas en 400-300 hPa y en niveles medios (700 hPa-500 hPa) son negativo. En la selva la temperatura entre los niveles bajos a medios (950-500 hPa) se encuentra correlacionado negativamente.

Figura 18: Correlación de Spearman Precipitación vs humedad relativa (Periodo DEFMA)

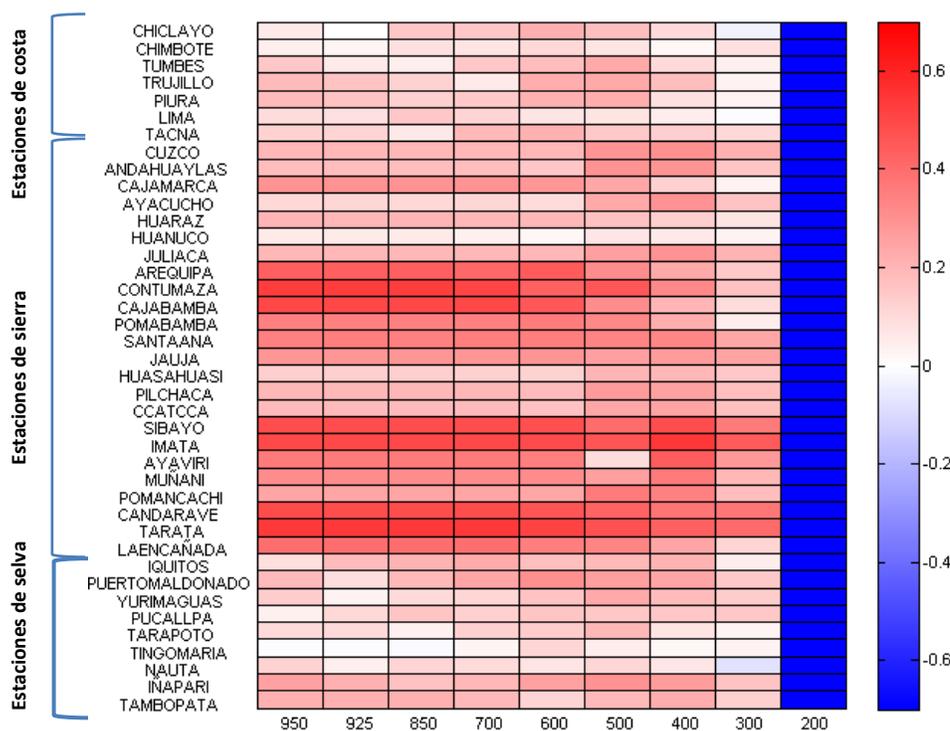


Figura 19: Correlación de Spearman de Precipitación vs viento zonal (periodo DEFMA)

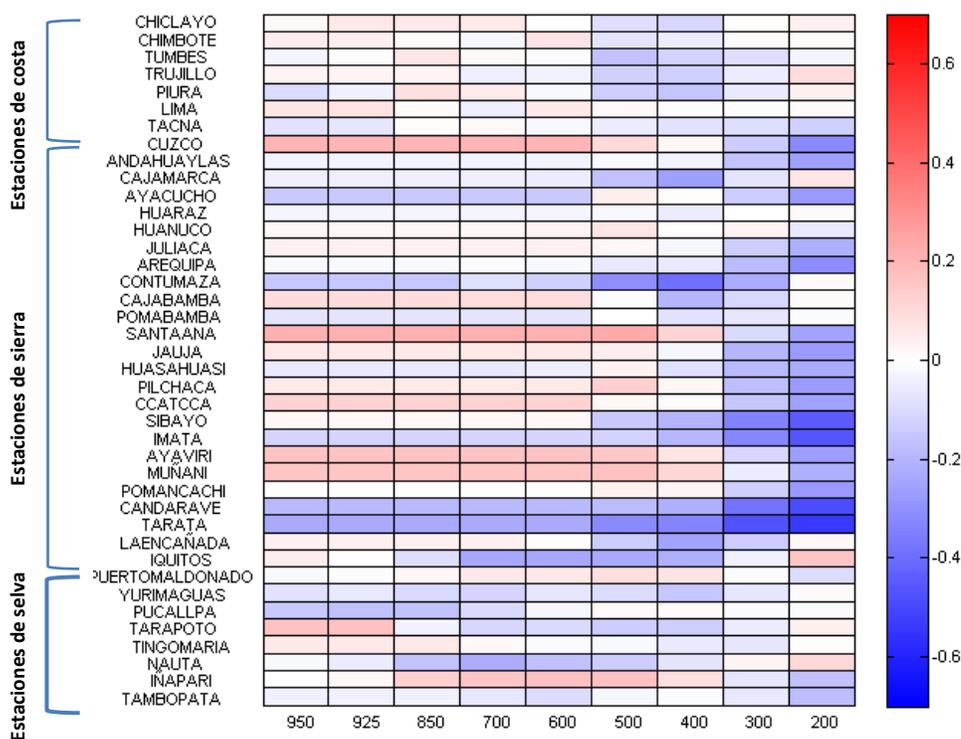


Figura 20: Correlación de Spearman de Precipitación vs viento meridional para el periodo DEFMA

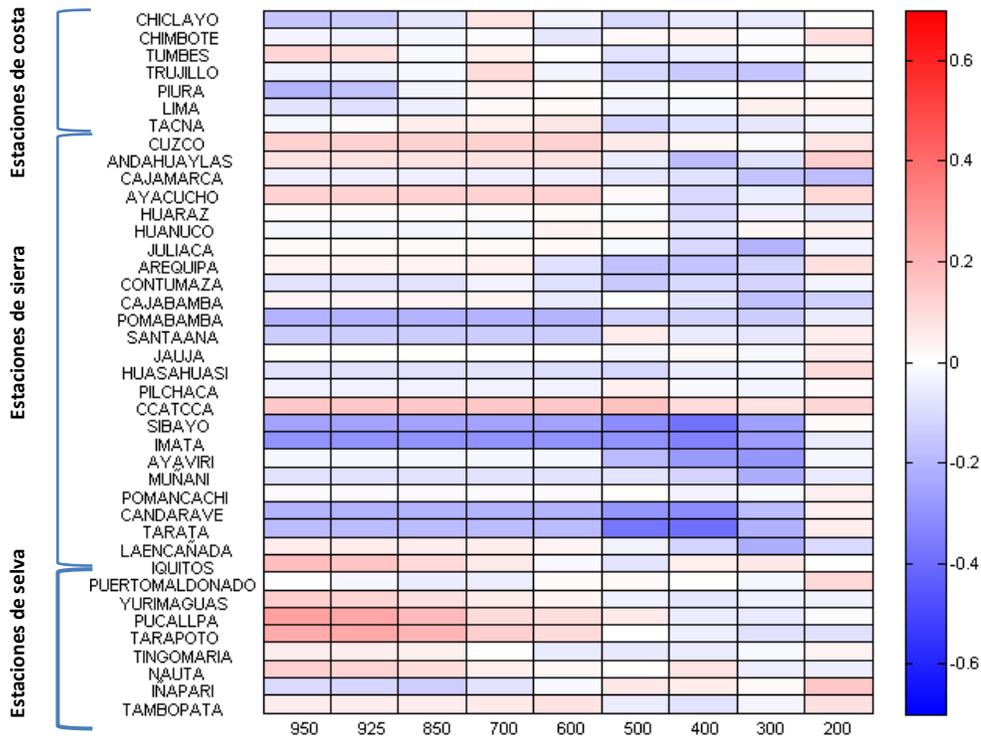
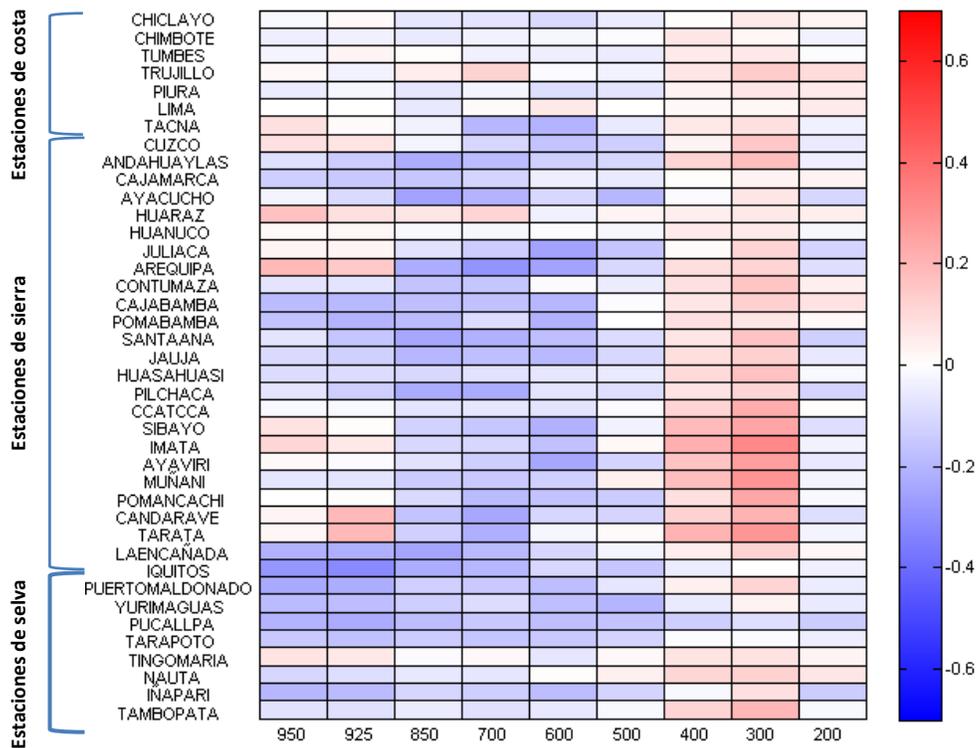


Figura 21: Correlación de Spearman de Precipitación vs temperatura para el periodo DEFMA



Como resultado de esta primera fase para escoger los predictores que ingresarían a los modelos de regresión, los niveles isobáricos de los cuatro variables que resultaron seleccionados para cada estación, se muestran en el cuadro 9 junto con sus respectivos valores de correlación de Spearman.

Cuadro 9: Nivel de presión y correlación de Spearman seleccionados para cada variable.

		Humedad relativa (RH)		Viento zonal (U)		Viento meridional (V)		Temperatura	
		Nivel de presión (hPa)	Correlación de Spearman	Nivel de presión (hPa)	Correlación de Spearman	Nivel de presión (hPa)	Correlación de Spearman	Nivel de presión (hPa)	Correlación de Spearman
Estaciones de costa	Chiclayo	600	0.21	700	-0.11	950	-0.15	925	-0.10
	Chimbote	600	0.10	500	0.08	950	0.09	850	0.07
	Tumbes	600	0.23	500	-0.17	500	0.11	925	0.06
	Trujillo	850	0.24	300	-0.13	400	-0.16	500	0.14
	Piura	600	0.22	700	-0.15	950	-0.20	600	-0.09
	Lima	925	0.15	300	0.08	700	-0.08	925	0.06
	Tacna	925	0.21	700	-0.13	200	-0.12	600	-0.20
Estaciones de sierra	Cuzco	400	0.30	200	-0.32	200	0.13	300	-0.16
	Andahuaylas	400	0.30	200	-0.26	300	-0.18	300	-0.22
	Cajamarca	925	0.29	200	-0.26	300	-0.18	850	-0.15
	Ayacucho	400	0.30	200	-0.28	300	0.12	300	-0.25
	Huaraz	600	0.20	200	-0.05	600	-0.10	300	0.17
	Huanuco	500	0.07	200	0.07	400	-0.07	300	0.06
	Juliaca	400	0.30	200	-0.22	300	-0.21	300	-0.25
	Arequipa	600	0.45	200	-0.31	700	-0.17	700	-0.29
	Contumaza	925	0.53	300	-0.39	300	-0.15	700	-0.16
	Cajabamba	925	0.50	700	-0.20	850	-0.17	300	-0.20
	Pomabamba	600	0.36	200	-0.08	300	-0.20	850	-0.21
	Santa Ana	500	0.35	200	-0.25	300	-0.14	300	-0.24
	Jauja	500	0.29	200	-0.27	850	0.05	300	-0.19
	Huasahuasi	500	0.20	200	-0.23	300	-0.10	300	0.17
	Pilchaca	400	0.28	200	-0.27	200	0.04	850	-0.23
	Ccatcca	400	0.25	200	-0.26	500	0.16	300	0.23
	Sibayo	600	0.49	200	-0.45	300	-0.39	300	0.24
	Imata	600	0.54	200	-0.46	300	-0.34	300	0.32
	Ayaviri	400	0.44	200	-0.27	300	-0.28	300	0.26
	Muñani	700	0.36	200	-0.22	300	-0.22	300	0.29
Pomancachi	400	0.36	200	-0.28	200	0.05	300	0.24	
Candarave	700	0.49	200	-0.49	500	-0.31	300	-0.24	
Tarata	600	0.54	200	-0.54	500	-0.39	300	0.29	
LaEncañada	500	0.40	700	-0.25	700	-0.23	300	-0.25	
Estaciones de selva	Iquitos	400	0.23	850	-0.24	700	0.18	950	-0.32
	Puerto	500	0.30	200	-0.09	850	0.10	950	-0.23
	Yurimaguas	500	0.23	200	-0.15	700	0.14	950	-0.20
	Pucallpa	500	0.16	200	-0.17	950	0.26	925	-0.23
	Tarapoto	500	0.19	950	0.17	925	0.23	950	-0.17
	Tingo	600	0.11	200	-0.07	700	-0.06	850	0.08
	Nauta	950	0.12	850	-0.22	700	0.13	950	0.12
	Iñapari	950	0.30	200	0.17	850	0.15	300	-0.20
	Tambopata	500	0.22	200	-0.18	950	0.08	300	0.19

b) Resultados de la segunda fase de selección: selección por pasos

Para realizar la segunda fase mediante el método de selección por pasos se consideraron los predictores que se muestran en el cuadro 9 además de la precipitación acumulada del modelo ETA-SENAMHI y los tres índices de estabilidad, sumando un total de 8 predictores para cada modelo de regresión. La figura 22 muestra las frecuencias de predictores que resultaron significativas estadísticamente a un nivel de 95% de confianza a partir del método de selección basado en la comparación de desviaciones. Los predictores que seleccionaron son los que tienen mayor frecuencia de significancia estadística para cada región.

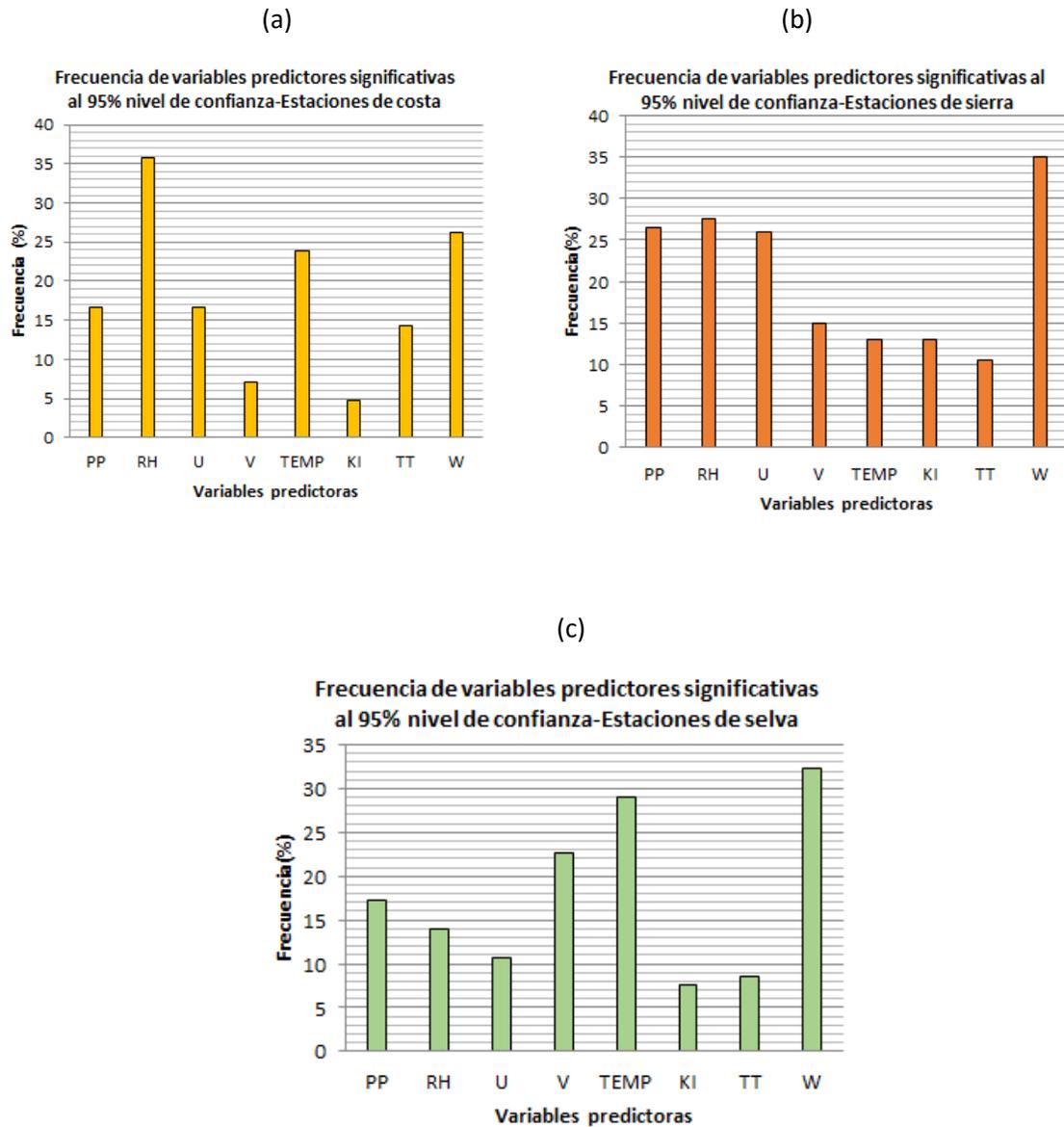
Se aprecia que para las estaciones ubicadas en costa (figura 22a) las variables que resultaron ser más significativas fueron la humedad relativa, temperatura y el índice de estabilidad W. En menor medida también resultaron frecuentes la precipitación acumulada y el viento zonal.

Para el grupo de estaciones de la sierra (figura 22b) las variables que tuvieron mayor frecuencia de significatividad estadística fueron la precipitación, humedad relativa, temperatura, viento zonal y el índice de estabilidad W.

Para el grupo de estaciones de la selva (figura 22c), las variables con más alta frecuencia de significatividad estadística fueron la precipitación, viento meridional, temperatura y el índice de estabilidad W.

Es importante resaltar que el índice de estabilidad W es la variable predictora que se muestra en mayor frecuencia para las estaciones de sierra y selva lo cual nos hace posible afirmar que es un buen índice que podría adaptarse para estas regiones. Los índices K y TT no mostraron gran aporte en los modelos de regresión y fueron los predictores con menores frecuencias de significancia estadística para las tres regiones.

Figura 22: Frecuencia de predictores significativos al 95% de confianza para (a) estaciones de costa (b) estaciones de sierra y (c) estaciones de selva.



4.1.2. PREDICTORES SELECCIONADOS PARA CADA ESTACIÓN

Como resultado de haber aplicado los dos pasos de selección de predictores se obtuvieron los predictores óptimos finales que se usarían para cada estación y los cuales son mostrados en el cuadro 10, teniendo en cuenta que cada modelo de regresión logística debe poseer como máximo 5 variables independientes. En todas las estaciones fueron considerados los predictores de precipitación acumulada y el índice de estabilidad W ya que resultaron

significativas al 95% de confianza con mayor frecuencia. En cuanto a los predictores por niveles de presión, en la región costa se encontró que estos predictores se encontraban mayormente entre niveles bajos a medios (950-500 hPa) a causa de la concentración de humedad en niveles bajos y también por el pase poco frecuente de la nubosidad proveniente de la cordillera.

En estaciones de la región sierra los predictores de humedad relativa se encuentran mayormente en niveles medios (entre 700 hPa a 400 hPa), ya que la mayor advección de humedad ocurre en estos niveles debido al desplazamiento de las masas de aire en dirección este a oeste y el ascenso orográfico e influyen en la ocurrencia e intensidad de la precipitación. Con respecto a vientos, las componentes del viento en niveles medios y altos (entre 400 hPa a 200 hPa) fueron seleccionadas, debido al transporte de humedad en niveles medios y al aporte de inestabilidad en niveles altos producto de la circulación de la alta de Bolivia.

Por último, para las estaciones de la selva, los predictores de viento meridional y temperatura que fueron seleccionados se encuentran mayormente en niveles bajos y medios (entre 950 hPa a 500 hPa) y que se relaciona con la mayor concentración de humedad en estos niveles producto de los flujos anticiclón del atlántico sur y también del aporte de la evapotranspiración de la extensa selva amazónica.

Estos mismos predictores se usaron para las cuatro categorías de pronóstico en cada estación pero con diferentes parámetros de regresión, con el objetivo de obtener resultados más consistentes y comparables.

Cuadro 10: Conjunto de predictores seleccionados para cada estación.

REGIÓN	ESTACIÓN	PREDICTORES SELECCIONADOS
COSTA	CHICLAYO	PP acum 24 horas, RH850, U700, TEMP925, W
	CHIMBOTE	PP acum 24 horas, RH600, U500, TEMP850, W
	TUMBES	PP acum 24 horas, RH600 U500 TEMP925, W
	TRUJILLO	PP acum 24 horas, RH850, U500, TEMP500, W
	PIURA	PP acum 24 horas, RH600, U700, TEMP600, W
	LIMA	PP acum 24 horas, RH925, U850, TEMP925, W
	TACNA	PP acum 24 horas, RH925, U700, TEMP600, W
SIERRA	CUZCO	PP acum 24 horas, RH400, U200, W
	ANDAHUAYLAS	PP acum 24 horas, RH400, U200, W
	CAJAMARCA	PP acum 24 horas, RH700, U200, W
	AYACUCHO	PP acum 24 horas, RH400, U200, W
	HUARAZ	PP acum 24 horas, RH600, U200, W
	HUANUCO	PP acum 24 horas, RH500, U200, W
	JULIACA	PP acum 24 horas, RH400, U200, W
	AREQUIPA	PP acum 24 horas, RH600, U200, W
	CONTUMAZA	PP acum 24 horas, RH700, U300, W
	CAJABAMBA	PP acum 24 horas, RH700, U700, W
	POMABAMBA	PP acum 24 horas, RH600, U200, W
	SANTAANA	PP acum 24 horas, RH700, U200, W
	JAUJA	PP acum 24 horas, RH500, U200, W
	HUASAHUASI	PP acum 24 horas, RH500, U200, W
	PILCHACA	PP acum 24 horas, RH400, U200, W
	CCATCCA	PP acum 24 horas, RH400, U200, W
	SIBAYO	PP acum 24 horas, RH700, U200, W
	IMATA	PP acum 24 horas, RH700, U200, W
	AYAVIRI	PP acum 24 horas, RH400, U200, W
	MUÑANI	PP acum 24 horas, RH700, U200, W
	POMANCACHI	PP acum 24 horas, RH400, U200, W
	CANDARAVE	PP acum 24 horas, RH700, U200, W
TARATA	PP acum 24 horas, RH700, U200, W	
LA ENCAÑADA	PP acum 24 horas, RH500, U850, W	
SELVA	IQUITOS	PP acum 24 horas, V700, TEMP950, W
	PUERTO MALDONADO	PP acum 24 horas, V850, TEMP950, W
	YURIMAGUAS	PP acum 24 horas, V700, TEMP950, W
	PUCALLPA	PP acum 24 horas, V950, TEMP925, W
	TARAPOTO	PP acum 24 horas, V925, TEMP950, W
	TINGO MARIA	PP acum 24 horas, V700, TEMP850, W
	NAUTA	PP acum 24 horas, V700, TEMP950, W
	IÑAPARI	PP acum 24 horas, V850, TEMP300, W
	TAMBOPATA	PP acum 24 horas, V950, TEMP300, W

4.2. ANÁLISIS DE LA EFICIENCIA DEL PRONÓSTICO PROBABILÍSTICO DIARIO DE PRECIPITACIÓN.

El análisis de la eficiencia del pronóstico probabilístico de precipitación se realizó enfocándose en los diferentes atributos de este tipo de pronóstico: fiabilidad, precisión, habilidad y resolución. A consecuencia de que estos atributos miden aspectos diferentes del pronóstico se realizó un análisis para cada uno de ellos; sin embargo, se ha encontrado que la mayoría de los atributos presentan resultados muy similares de la eficiencia de los modelos. Estos atributos están relacionados directamente, a pesar de que miden parámetros muy diferentes pero presentan un comportamiento similar. Por ejemplo, mientras más cercanos este el diagrama de fiabilidad a la diagonal contribuye a tener un valor más alto de Brier Skill Score; y a su vez la resolución se vuelve más eficiente ya que mejora la discriminación de eventos. Se ha encontrado que el atributo de fiabilidad presenta gran influencia sobre los atributos de habilidad y resolución; mientras que el atributo de precisión no es adecuado para el análisis de la eficiencia del pronóstico probabilístico debido a la gran sensibilidad que presenta ante eventos de precipitación poco frecuente.

A continuación se muestran tablas y gráficos de los resultados para cada atributo y con las estaciones agrupadas en región costa, sierra y selva; así como la discusión y análisis estadístico de cada uno de ellos.

4.2.1. RESULTADOS DEL ATRIBUTO DE FIABILIDAD

La figura 23 muestra los diagramas de fiabilidad que corresponden a los resultados de probabilidades para todo el periodo de trabajo (Enero 2010 - Enero 2013). El diagrama de fiabilidad para el pronóstico probabilístico de la región costa muestra las mayores desviaciones respecto de la línea diagonal, lo cual presenta una deficiente fiabilidad y que estaría asociado a una inadecuada simulación de las variables meteorológicas por parte del modelo ETA-SENAMHI en esta región. Betts y Miller (1986) señalan que el modelo ETA no tiene un control explícito de la capa de inversión el cual es el principal factor que controla la precipitación en la costa centro y sur. Para las estaciones de la sierra y selva el diagrama se encuentran más cercanos a la recta diagonal con pocas diferencias entre las categorías de

pronóstico siendo la fiabilidad muy buena para ambas regiones; sin embargo, existen estaciones en donde los resultados son pésimos tales como Huaraz, Huánuco y Tingo María (véase Anexo 1).

La figura 24 muestra los diagramas de fiabilidad para el trimestre de Diciembre a Febrero, periodo de abundantes precipitaciones. La costa es la que muestra la menor fiabilidad. El diagrama para la sierra y selva es coincidente con la diagonal, principalmente en la categoría uno pero se separan ligeramente según vayamos aumentando la categoría de pronóstico lo cual sugiere una pérdida de la fiabilidad cuando se realiza pronóstico probabilístico para mayores cantidades de precipitación.

Los diagramas de fiabilidad para el periodo de Junio a Agosto (figura 25), periodo en el cual las precipitación son escasas, indican que la fiabilidad es sumamente baja para esta época bajo la condición de que los diagramas se separan notablemente de la diagonal en las tres regiones. Los resultados más óptimos se presentan en la región selva; sin embargo, a medida que se aumenta la categoría de pronóstico la separación va siendo mayor, mientras que en la sierra la fiabilidad es deficiente para las cuatro categorías. Estos resultados son debido a que en este trimestre del año periodo seco el transporte zonal de humedad desde la cuenca amazónica es mucho menor y los eventos de precipitación disminuyen considerablemente en la sierra y en la selva, lo cual se ve reflejado en los diagramas de fiabilidad. Según Ancel (2009), el estudio de la fiabilidad de los sistemas para diferentes eventos es bien importante, ya que los eventos raros casi nunca se predicen con probabilidades altas, manifestando una fuerte pérdida de fiabilidad. Situación similar de pérdida de la fiabilidad ocurre con las estaciones de costa en donde los eventos de precipitación son poco frecuentes durante todo el año.

Figura 23: Diagramas de fiabilidad correspondientes a todo el periodo de estudio

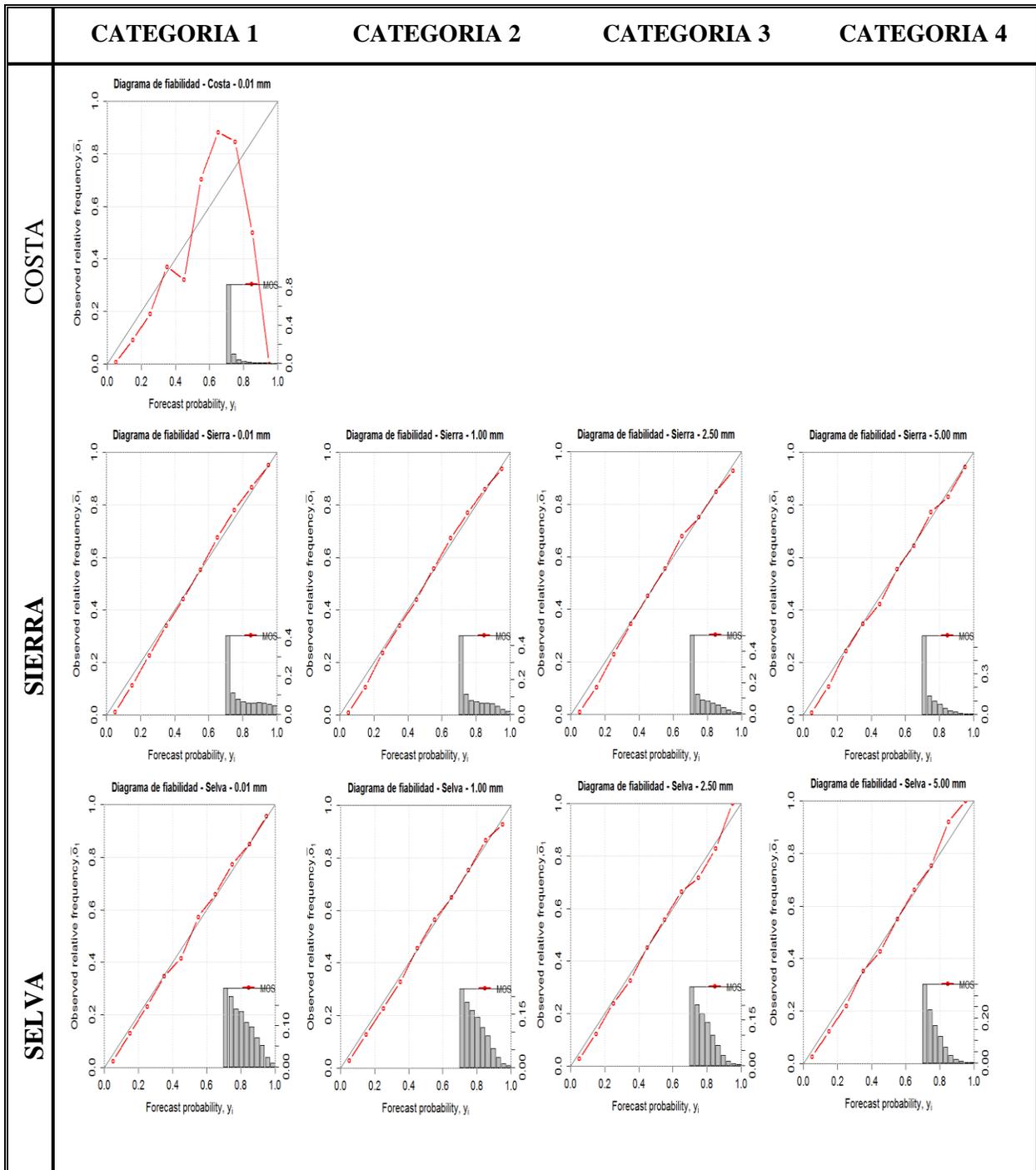


Figura 24: Diagramas de fiabilidad correspondientes al periodo Diciembre-Enero-Febrero

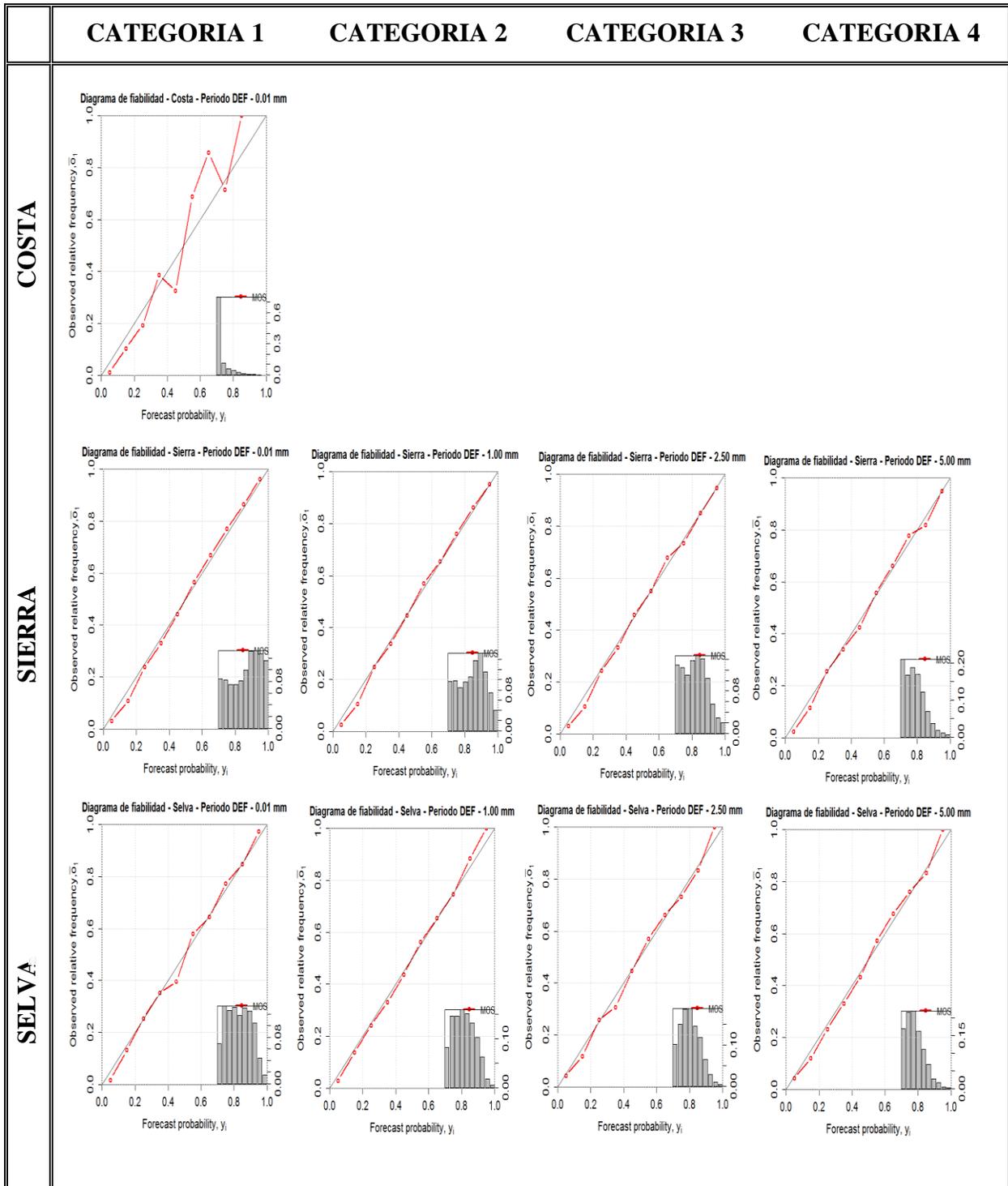
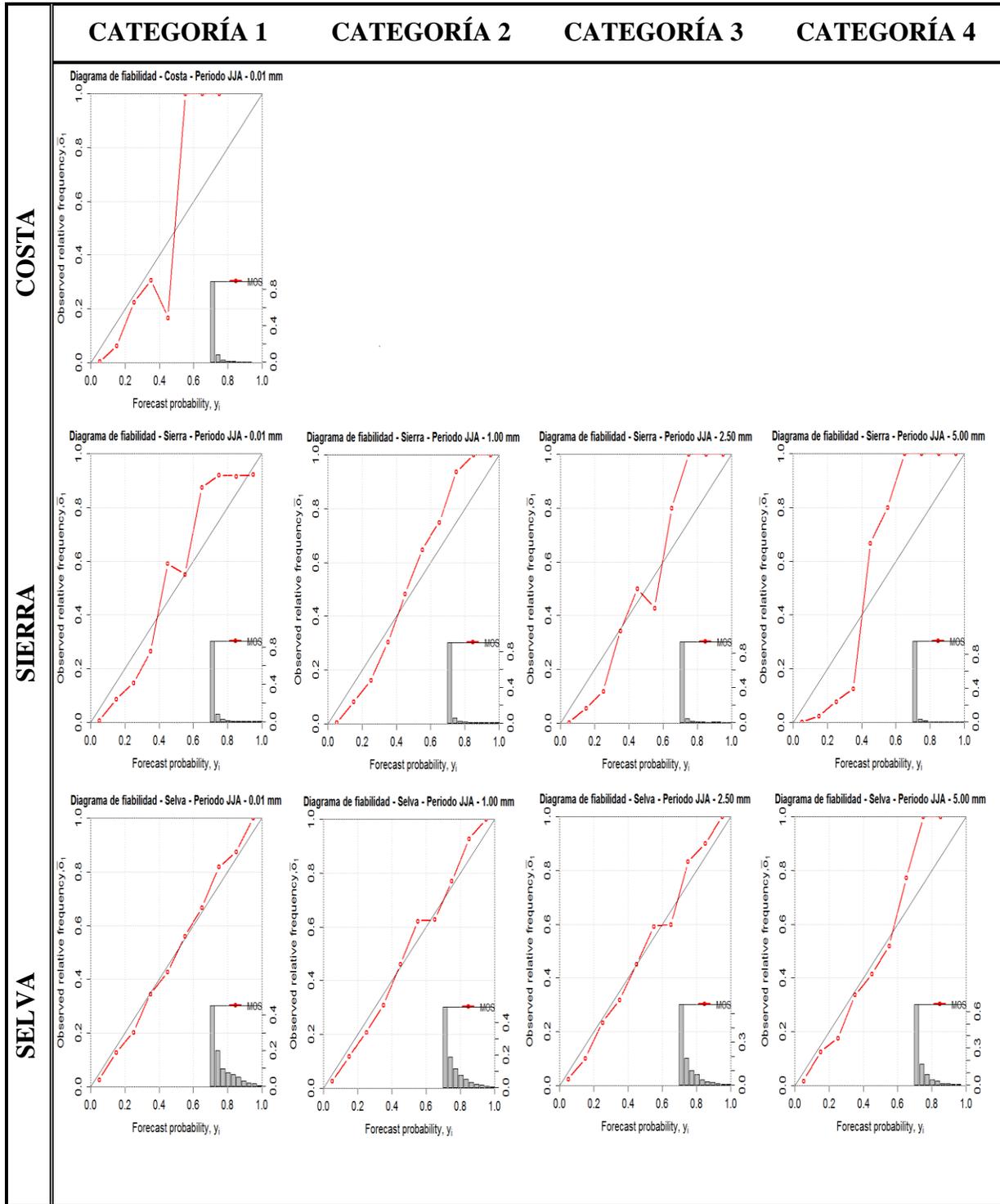


Figura 25: Diagramas de fiabilidad correspondiente al periodo Junio- Julio- Agosto.



Complementando el análisis de la fiabilidad, los diagramas mostrados anteriormente fueron cuantificados mediante la descomposición del Brier Score, obteniéndose valores numéricos de la calidad del atributo de fiabilidad mediante este coeficiente. Los resultados (figura 26) son similares al del diagrama de fiabilidad encontrándose que las estaciones de la sierra y selva presentan eficiente fiabilidad ya que sus valores son más cercanos a cero, a comparación de las estaciones de la costa y en el periodo de verano es cuando se presentan los mejores valores de fiabilidad.

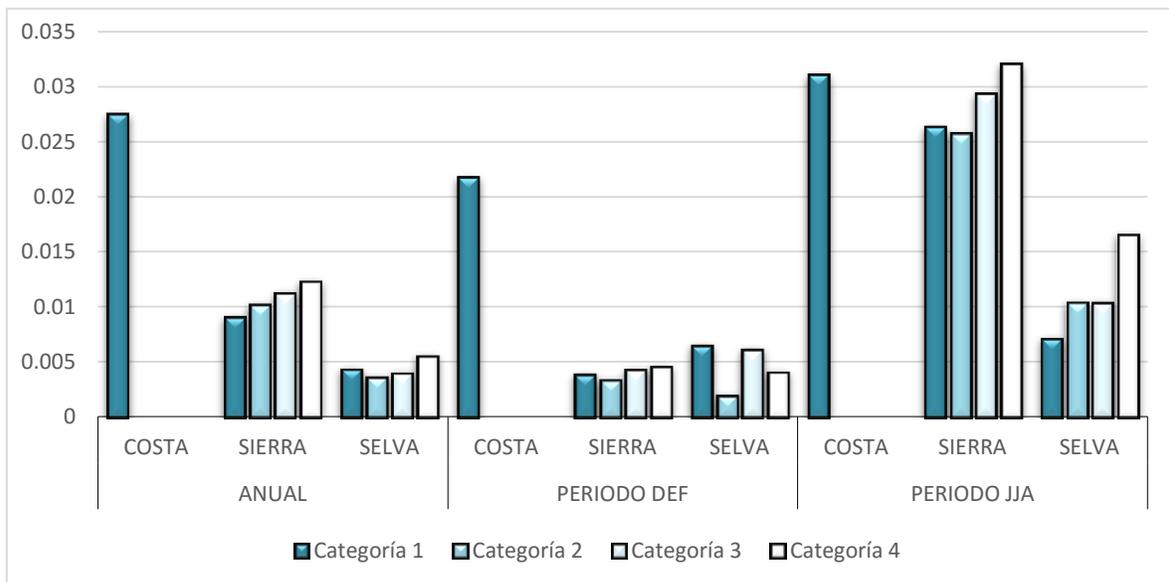


Figura 26: Componente fiabilidad del Brier Score.

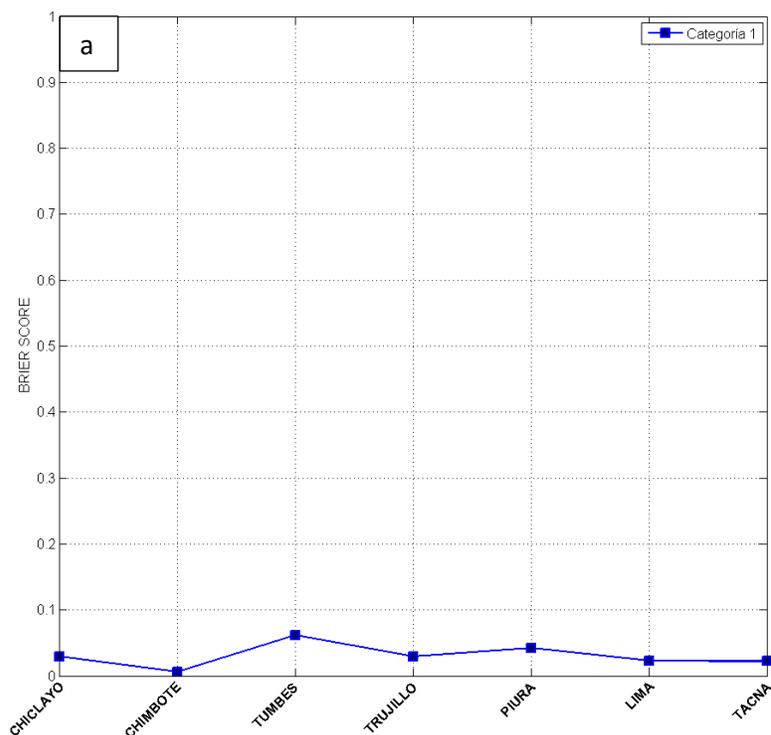
4.2.2. RESULTADOS DEL ATRIBUTO DE PRECISIÓN

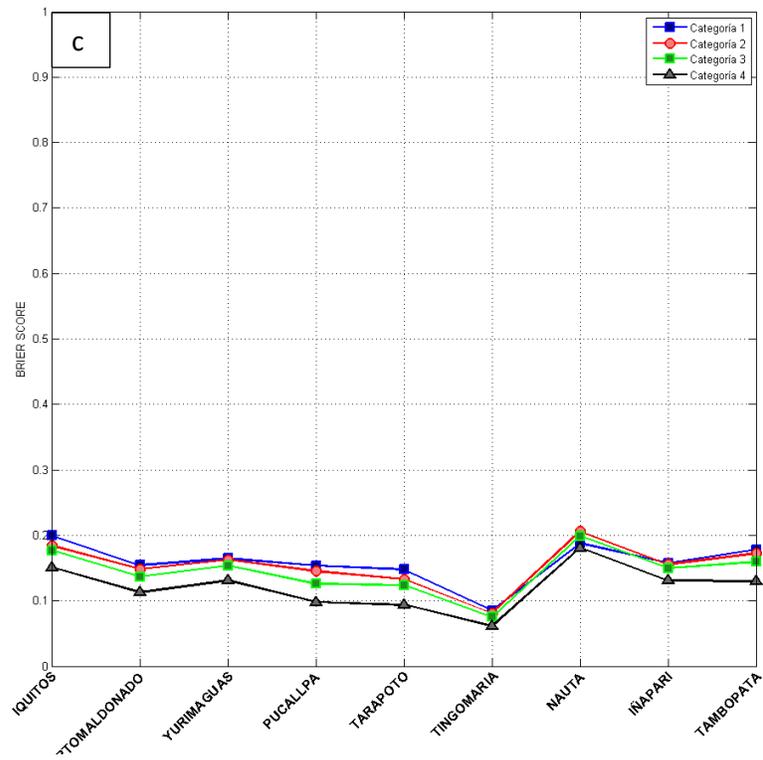
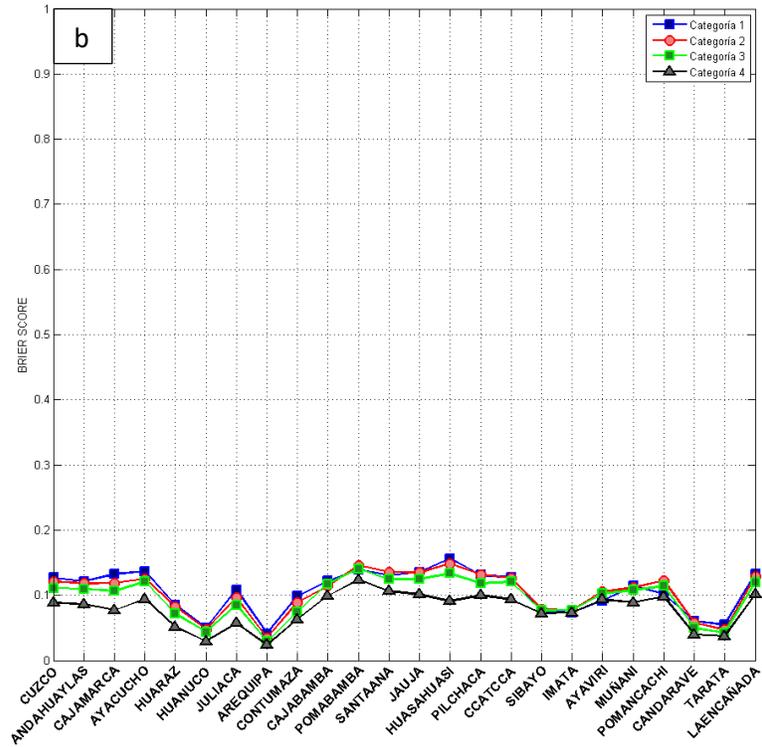
En la figura 27 se muestran los resultados del Brier Score para el pronóstico probabilístico MOS en los cuales se observan que las cuatro categorías y las tres regiones presentan buena precisión (valores de Brier Score cercanos a cero) especialmente en estaciones de costa y menor en estaciones de selva. A pesar de que la precisión es óptima para las estaciones de costa (figura 27a), se deben tomar estos resultados con mucha precaución ya que este estadístico es muy sensible a la gran cantidad de eventos nulos. Tal y como manifiesta Ancel (2009) si el evento es muy raro (como sucede con la precipitación en la costa del Perú) un

sistema sin utilidad que siempre diga que no va a llover (como la climatología o la persistencia) tendría un Brier Score casi nulo por lo que sería casi imposible de superar por un sistema de mayor utilidad. De la misma forma Gutierrez *et al.*, (2004) menciona que el Brier Score es un índice poco representativo cuando está descompensada la frecuencia de ocurrencia y de no ocurrencia del evento. Por lo tanto, la gran cantidad de pronósticos probabilísticos cercanos a cero y la poca cantidad de días lluviosos hacen que el Brier Score resulte muy cercano a cero sugiriendo un pronóstico bien preciso, de esta forma este estadístico solo es representativo del evento más frecuente.

En las estaciones de sierra (figura 27b) y selva (figura 27c), se observan que los resultados del Brier Score del pronóstico probabilístico MOS presentan valores promedios de aproximadamente 0.1 y 0.15 respectivamente. Además se aprecia una ligera mejora de la precisión a medida que aumenta la categoría, lo cual se debe a que la frecuencia de estos eventos disminuye a medida que aumentamos los umbrales de cantidades de precipitación.

Figura 27: Resultados de Brier Score (atributo de precisión) para estaciones de (a) costa (b) sierra y (c) selva.





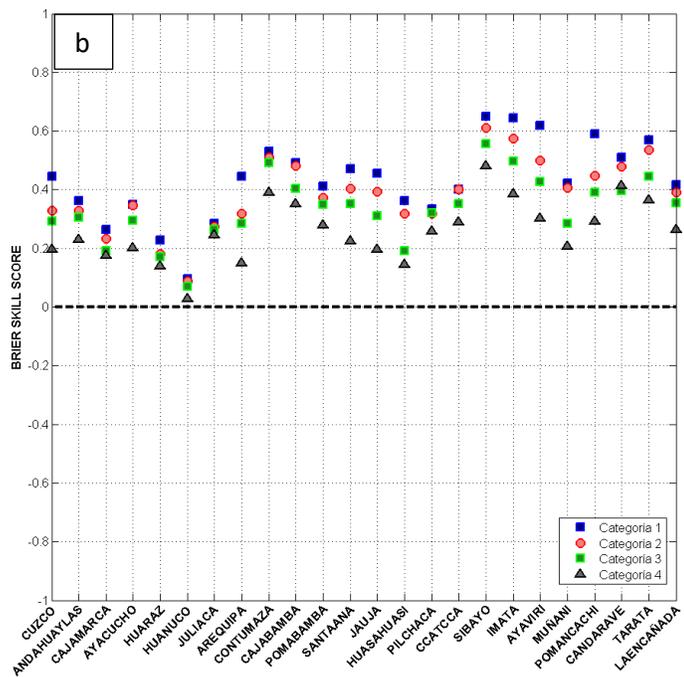
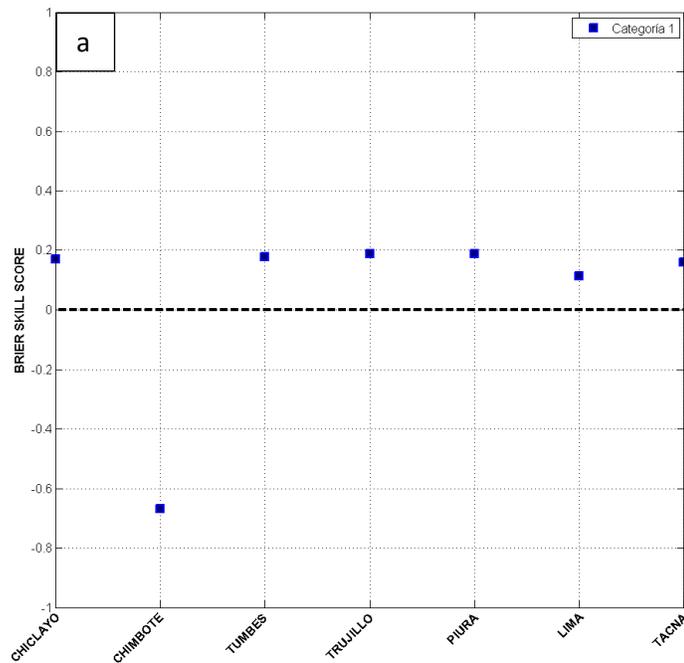
4.2.3. RESULTADOS DEL ATRIBUTO DE HABILIDAD

La figura 28 muestra los resultados del BSS del pronóstico probabilístico MOS. Este estadístico sirve para el análisis de la habilidad del modelo comparando el Brier Score del pronóstico contra el Brier Score de la climatología.

A consecuencia de que los resultados del diagrama de fiabilidad contribuyen también en cierta medida a los resultados del Brier Skill Score (BSS) los resultados que se muestran son muy similares. Así tenemos que los BSS son superiores a cero tanto como para las tres regiones y para las cuatro categorías de pronóstico, siendo los valores más altos los de la sierra (figura 28b) los cuales sobrepasan 0.6 (estaciones como Sibayo, Imata, Ayaviri y Tarata). Estos valores positivos indican una mejora del pronóstico MOS por sobre la climatología y una buena habilidad de este modelo. Los resultados de BSS para la región selva (figura 28c) son ligeramente menores a los de la sierra, teniendo valores de 0.1 a 0.3 pero siguen siendo positivos. Los más bajos valores los presenta la costa (figura 28a) con un BSS promedio son de 0.1 y además la estación de Chimbote es negativo, sugiriendo una pobre habilidad del pronóstico MOS en esta región. Se observa, además, que los valores de BSS del pronóstico MOS disminuyen a medida que aumenta la categoría, lo cual nos sugiere que la habilidad del pronóstico probabilístico se va perdiendo cuando se usa mayores umbrales de cantidades de precipitación.

Adicionalmente, los gráficos de caja de la figura 29 muestran los resultados del BSS tanto del modelo ETA-SENAMHI como del pronóstico probabilístico MOS. Se aprecia que los pronósticos del modelo ETA-SENAMHI presentan deficiente habilidad ya que la gran mayoría de los valores de BSS son negativos, lo cual significa que no se pudo agregar información más allá del conocimiento de la frecuencia climatológica del evento. Por otro lado, los valores de BSS con el pronóstico MOS son todos positivos lo cual que nos permite afirmar que la habilidad del pronóstico probabilístico MOS de precipitación es generalmente mejor que los del modelo ETA-SENAMHI.

Figura 28: Resultados de Brier Skill Score (atributo de habilidad) para estaciones de (a) costa (b) sierra y (c) selva.



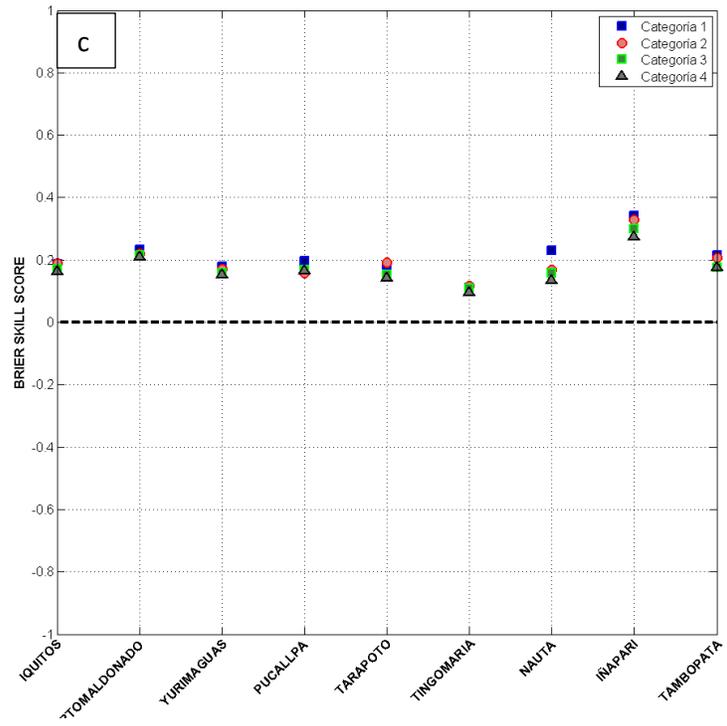
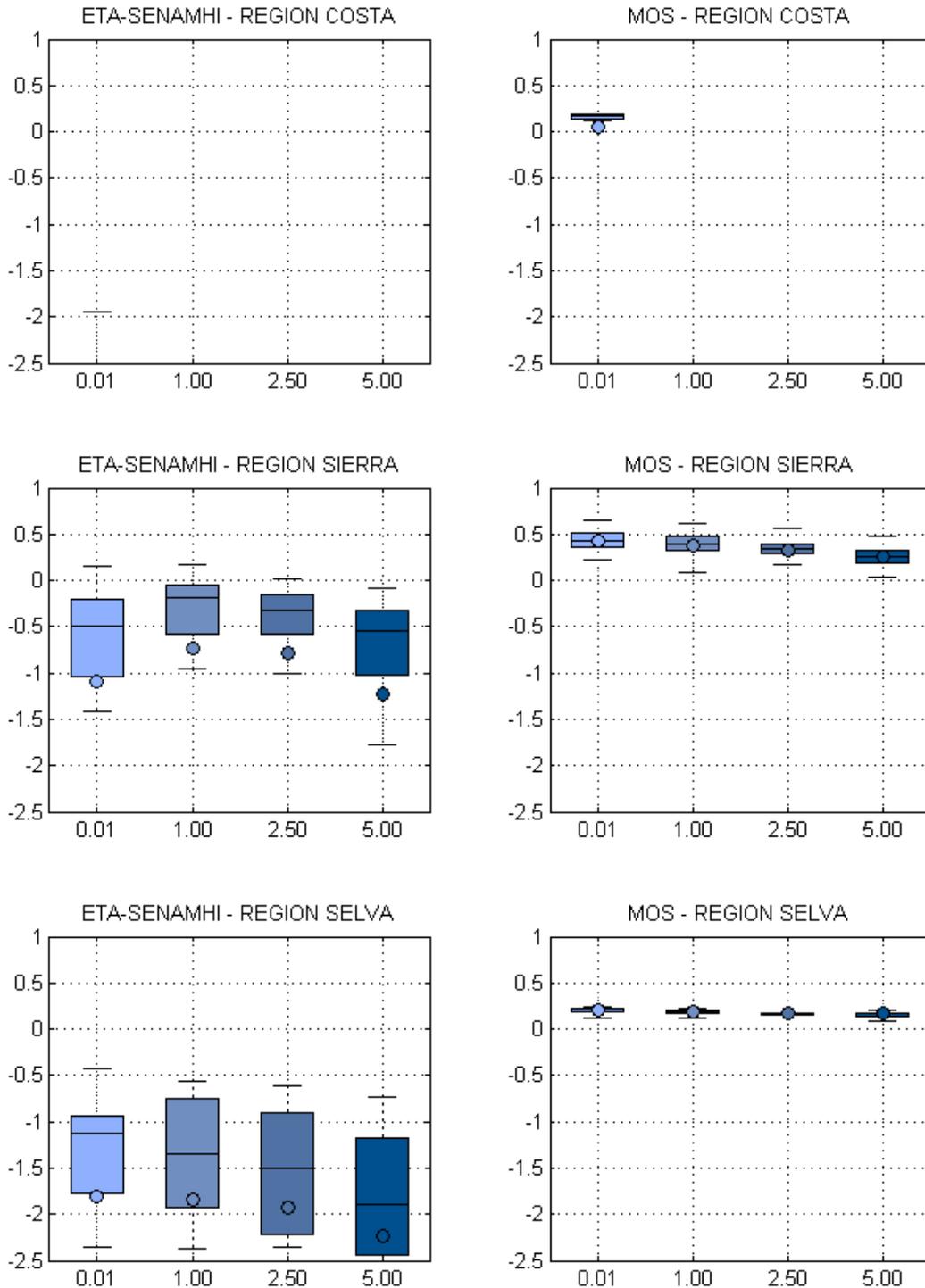


Figura 29: Comparación Resultados del Brier Skill Score (atributo de habilidad) del modelo ETA-SENAMHI (izquierda) y MOS (derecha).

BRIER SKILL SCORE



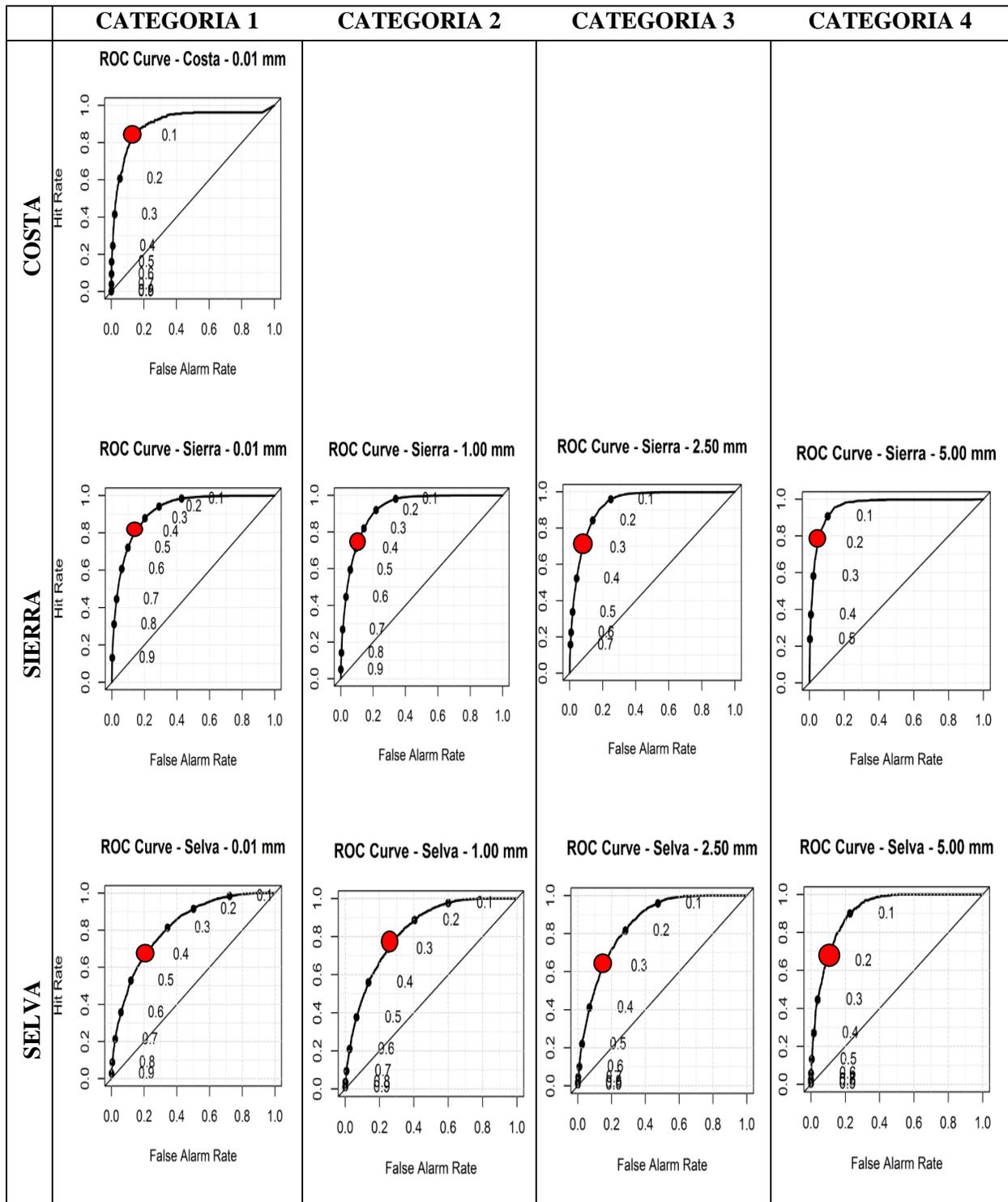
4.2.4. RESULTADOS DEL ATRIBUTO DE RESOLUCIÓN.

La figura 30 muestra las curvas ROC correspondientes a cada región y a cada categoría de pronóstico. Estas curvas fueron construidas a partir de los valores de razón de aciertos y falsas alarmas calculados con diferentes umbrales de probabilidad, por lo tanto es útil para conocer la resolución de los modelos; es decir, la capacidad de discriminar entre sucesos y no-sucesos.

Los resultados más deficientes de resolución se muestran para las estaciones región selva ya que la curva ROC se encuentra más cercana a la diagonal a comparación de las curvas para la sierra y costa, los cuales indican una menor capacidad de discriminar eventos. A esto puede complementarse los resultados de la resolución calculados a partir de la descomposición del Brier Score (ANEXO 3), en donde se aprecia que la selva posee menores coeficientes de resolución a comparación de la sierra y costa y lo cual se puede traducir también en valores de razón de aciertos más bajos y de razón de falsas alarmas más altos. A medida que aumentamos la categoría de pronóstico las falsas alarmas disminuyen para todos los umbrales de probabilidad, dado que la frecuencia climatológica del evento disminuye y la probabilidad de predecir el evento es también baja.

La curva ROC nos permite conocer además cuál es el umbral óptimo de probabilidad que discrimine la ocurrencia y no ocurrencia del evento (probabilidad en la que se tenga mayores tasas de aciertos y menor cantidad de falsas alarmas). Estos resultados se resumen en el cuadro 11. A pesar de que la costa presenta buena resolución, su umbral óptimo de probabilidad es de 0.1 siendo este un valor bien bajo debido al efecto de pocos eventos de ocurrencia de la precipitación y que influyen en que los modelos de regresión logística no sean eficientes para esta región. Con respecto a las estaciones de la sierra y selva, los umbrales probabilísticos óptimos para la discriminación de eventos se encuentran entre 0.2 a 0.4 dependiendo de la categoría de pronóstico de tal forma que a mayor categoría, el umbral va a ser menor. Resultados similares fueron encontrados en España por Gutierrez *et al* (2004) encontrando que, a medida que el evento es más extremo, las falsas alarmas disminuyen para todos los umbrales de probabilidad dado que la frecuencia climatológica del evento disminuye y por tanto la posibilidad de predecir el evento es también baja.

Figura 30: Resultados de curvas ROC (atributo de resolución).



Cuadro 11: Umbrales de probabilidad para la discriminación de eventos

	CATEGORÍA 1	CATEGORÍA 2	CATEGORÍA 3	CATEGORÍA 4
COSTA	0.1			
SIERRA	0.4	0.4	0.3	0.2
SELVA	0.4	0.3	0.3	0.2

4.3. APLICACIÓN DEL PRONÓSTICO PROBABILÍSTICO DE LA PRECIPITACIÓN EN UN EVENTO DE FUERTES PRECIPITACIONES.

4.3.1. ANÁLISIS DE ENTORNO SINÓPTICO DEL EVENTO

La circulación promedio climática en verano sobre Sudamérica, presenta un patrón característico de circulación anticiclónica conocido como Alta de Bolivia (AB), asociado a la liberación de calor latente y calor sensible del Altiplano y región subtropical (norte de Argentina y Paraguay) (Trebejo et. al, 2011). Dicho sistema favorece el desarrollo de la precipitación; sin embargo, no está siempre estable y su configuración puede ser modificada debido a la interacción con otros sistemas de escala sinóptica. Esto se encontró en las circulaciones en niveles altos para el periodo del 02 al 09 de Febrero del 2013 cuyo posicionamiento de la Alta de Bolivia al sur del Perú debido a la interacción con una baja segregada ubicada al oeste, ocasionó que se tuviera precipitaciones principalmente en la parte centro y sur de la sierra y selva (Ver Anexo 8).

a) Análisis de la circulación en niveles altos (200 hPa)

En la figura 31 los campos de líneas de corriente en el nivel de 200 hPa (tropósfera alta) indican que durante los días 02 y 03 de febrero, cerca de la costa peruana hay un sistema de cuña vaguada de onda larga asociada a una circulación ciclónica con núcleo en 25°S/88°W y a un jet streak que cruza la región centro y norte de Chile. Este sistema provocó que se amplificara una dorsal posicionada en continente con eje paralelo a Chile y con vientos de más de 30 m/s, favoreciendo a que la Alta de Bolivia se desplazaría más hacia el sur de su posición climatológica, ubicándose su núcleo en 20°S/62°W para el día 03 de febrero.

Para el día 04 de febrero la perturbación meridional de la vaguada se comienza a debilitar y desplazarse hacia el suroeste favoreciendo a que la circulación anticiclónica de la alta de Bolivia se desplace ligeramente al noroeste ubicándose su núcleo al norte de Chile y permitiendo el paso de flujos zonales a lo largo de la cordillera peruana.

Durante los días del 05 al 07 de febrero la circulación anticiclónica de la Alta de Bolivia posicionada en el sur del Perú se volvió más zonal, al norte de su núcleo, permitiendo el aporte de humedad y con flujos que presentaron divergencia por velocidad, lo cual dio condiciones óptimas para la continua precipitación, especialmente en la parte centro y sur del país.

Durante los días del 08 al 09 de febrero, la circulación ciclónica en altura desaparece completamente quedando la dorsal que se encontraba en continente cuyo desplazamiento hacia el este (aumentando su amplitud) favoreció a que una vaguada se forme delante y desplace la Alta de Bolivia hacia el Océano Pacífico, perdiendo sus características.

a) Análisis de la circulación en niveles medios (600 - 500 hPa)

En niveles medios (figura 32), durante los días 02 y 03 de febrero, se presenta una circulación anticiclónica con núcleo en 25°S/74°W, el cual es reflejo del movimiento de la posición de la alta de Bolivia en niveles altos, permitiendo el ingreso de flujos del sureste hacia el Perú. Para el día 05 de febrero la circulación se hace más zonal con vientos del este, siendo trascendental en la advección de humedad desde la Hoya Amazónica y Atlántico hacia la parte sur y centro del territorio peruano. Para los días posteriores la circulación ciclónica se va debilitando y desplazándose hacia el suroeste producto de los flujos zonales del este en continente que continúan adveccionando aire cálido y húmedo (razón de mezcla) hacia la ladera oriental de los Andes.

La figura 33 muestra los flujos de razón de mezcla en el nivel de 600 hPa. Durante los días 04 y 05 de febrero la razón de mezcla alcanza valores mayores de 5g/kg en el sur y centro del Perú debido a la advección de humedad en estos niveles producto del posicionamiento de la Alta de Bolivia en el sur del Perú. Según Garreaud (2001) es necesario alcanzar un mínimo de 5-6g/kg de razón de mezcla para que ocurra la precipitación en el altiplano. Este aporte de humedad continúa durante los próximos días para luego disminuir paulatinamente durante los días 08 y 09 de Febrero.

Figura 31: Líneas de corriente en 200 para los días (a) 02 de febrero (b) 05 de febrero (c) 07 de febrero y (d) 09 de febrero del 2013

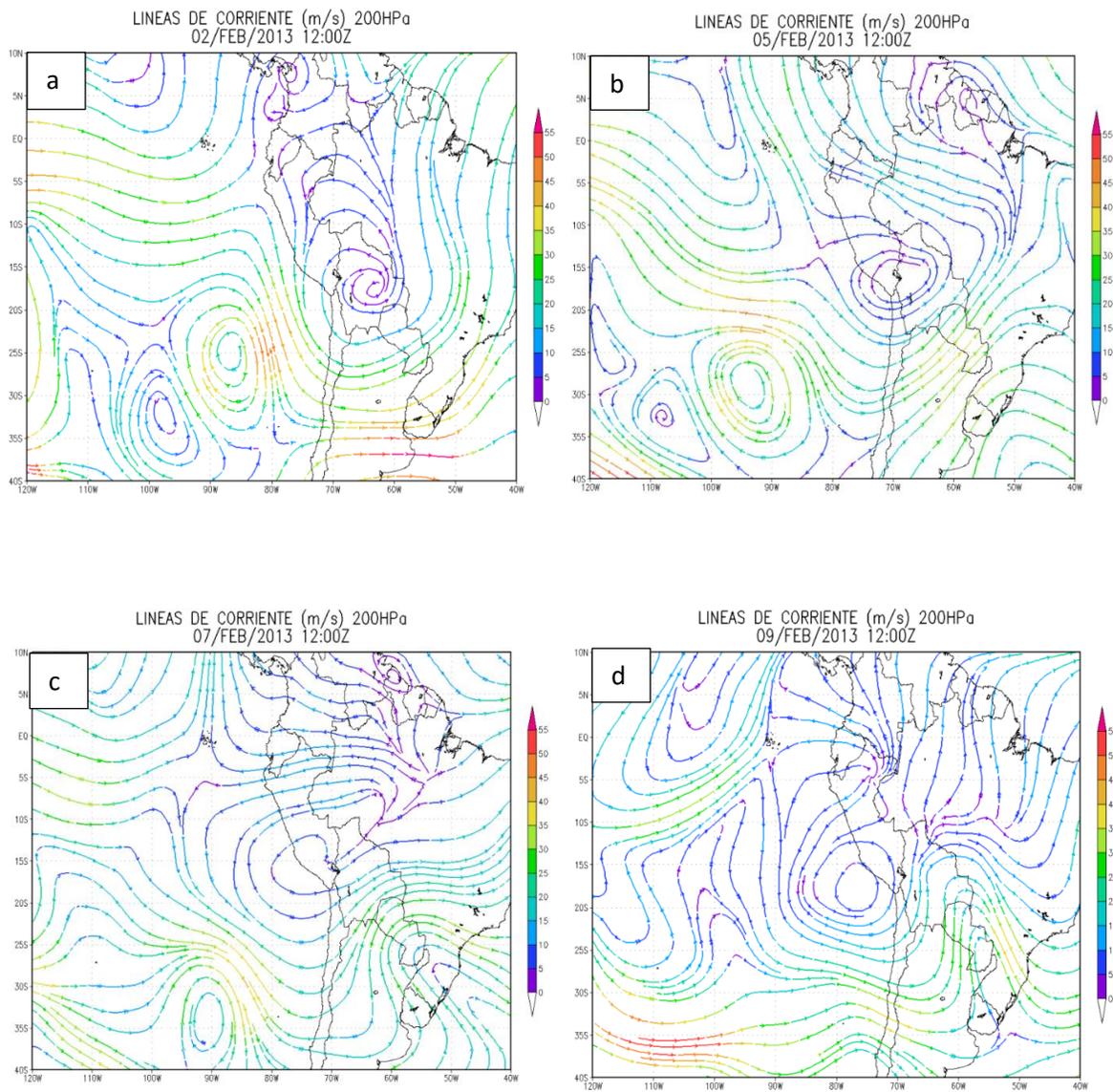


Figura 32: Líneas de corriente en 500 hPa para los días (a) 02 de febrero (b) 05 de febrero (c) 07 de febrero y (d) 09 de febrero del 2013

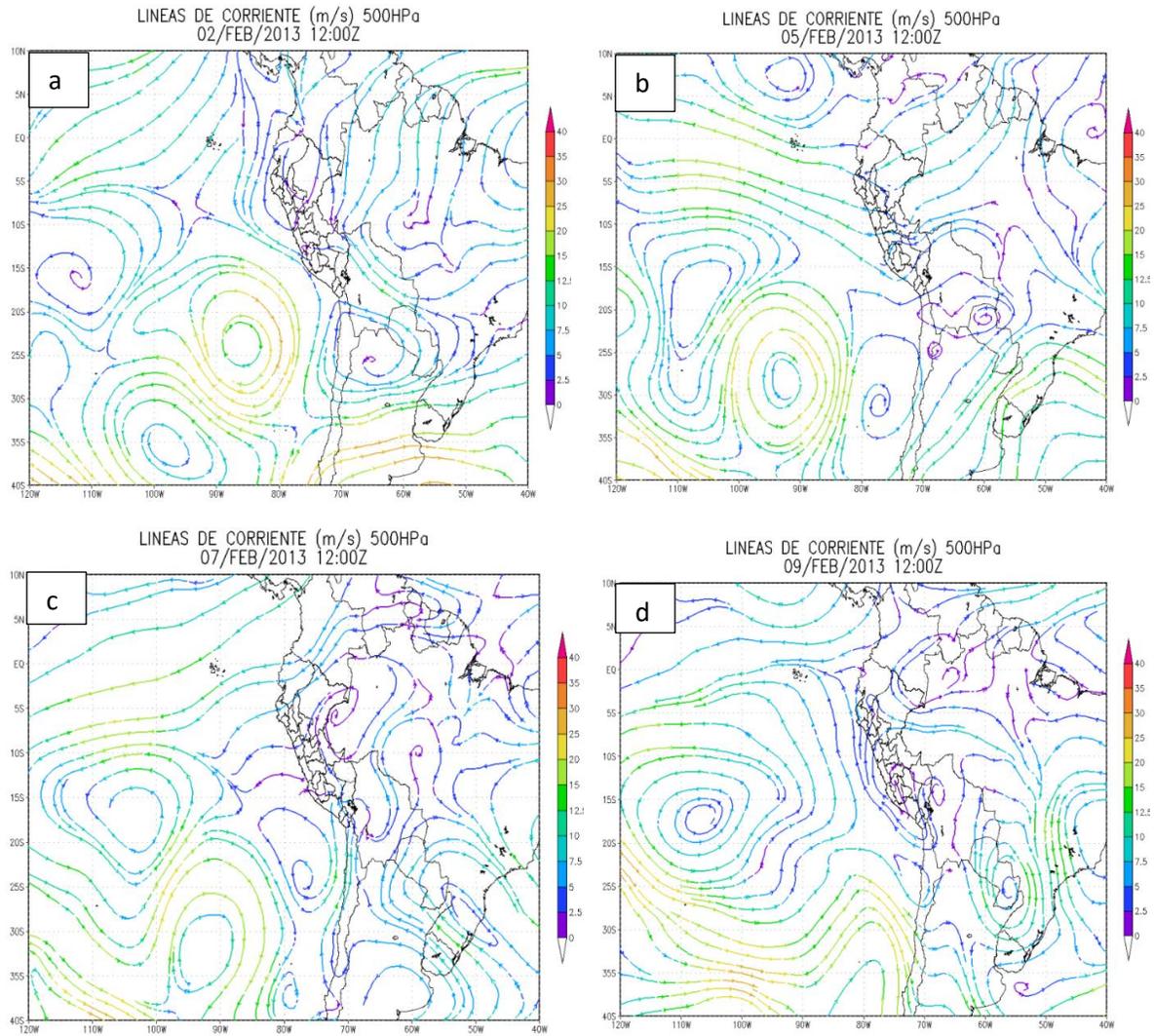
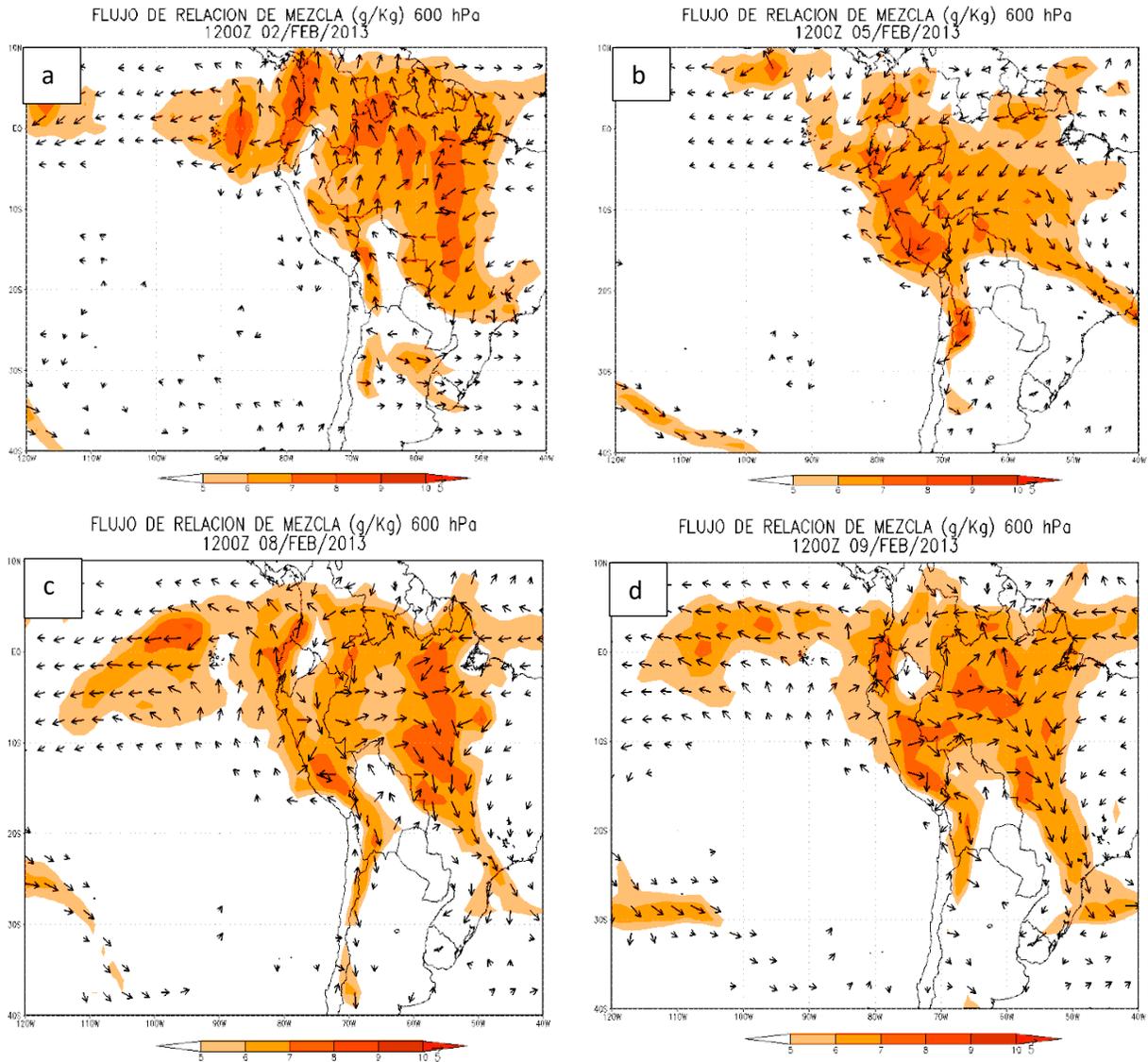


Figura 33: Flujo de razón de mezcla en 600 hPa para los días (a) 02 de febrero (b) 05 de febrero (c) 08 de febrero y (d) 09 de febrero del 2013

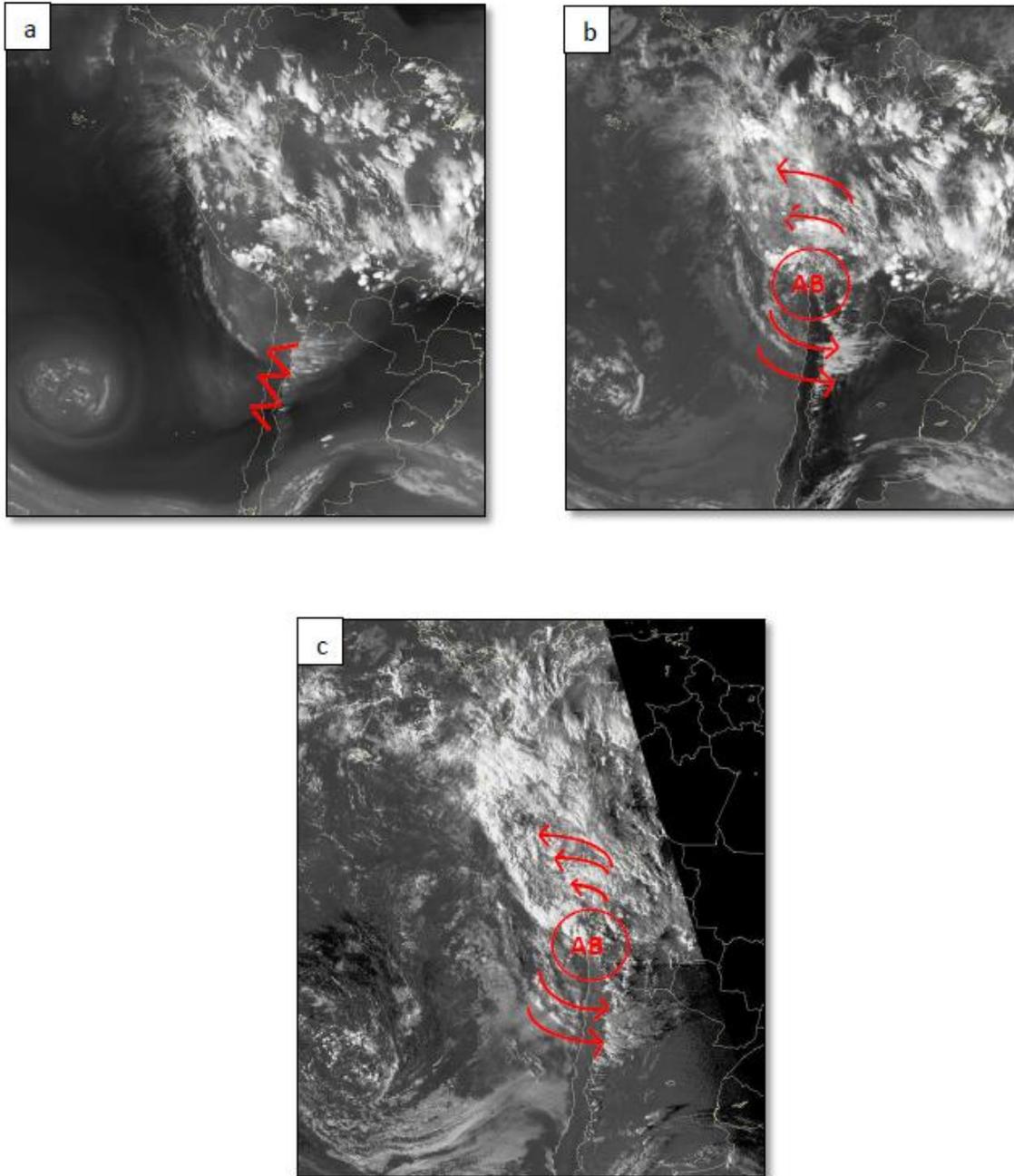


Para complementar el análisis sinóptico, la figura 34 muestra las imágenes del satélite GOES 13 para los canales de vapor de agua, infrarrojo y visible para el 04 de Febrero del 2013 a las 22: 30 UTC. La imagen de vapor de agua (figura 34a) nos proporciona información acerca de las condiciones de humedad en niveles altos y medios, además en ella podemos identificar la baja segregada asociada con la vaguada de gran amplitud de onda que se ubica en el pacífico oriental y que contribuyeron a la inestabilidad atmosférica sobre nuestro país. Hacia el sur, se pude distinguir la dorsal de gran amplitud cuyo eje se encuentra en el océano pacífico paralelo a Chile. Los colores más claros y grises en la parte sur y ciertas zonas del centro y norte del Perú representan el ingreso de humedad proveniente de la Amazonía, y la formación de sistemas convectivos.

En la imagen infrarrojo (figura 34b) se muestra la nubosidad en la región centro y sur del Perú con topes de nubes de colores bien claros, los cuales significan nubes de gran desarrollo vertical. Además, en la misma figura se observan los rastros de nubes que deja la circulación anticiclónica de la Alta de Bolivia cuyo centro se encuentra ubicado en cerca al norte de Chile.

En la imagen visible (figura 34c) se aprecia la extensión de nubosidad cubriendo a Perú y concentrada mayormente al sur y en la parte occidental de la cordillera y que está relacionado a la intensa convectividad produciéndose en esas zonas.

Figura 34: Imágenes de satélite del GOES 13 para el día de 04 de febrero del 2013 22: 30 UTC. (a) Imagen vapor de agua (b) imagen infrarrojo (c) imagen visible. Fuente: CPTEC-INPE



4.3.2. RESULTADOS DE LA APLICACIÓN DEL PRONÓSTICO PROBABILÍSTICO DE LA PRECIPITACIÓN PARA EL EVENTO DE LLUVIAS FUERTES.

Las figuras 35 al 38 muestran los pronósticos probabilístico correspondiente a cada estación graficados en el mapa para los días 02, 04, 07 y 09 de Febrero del 2013. En la parte inferior a dichos gráficos se aprecian los datos observados de precipitación en forma binaria correspondiente a cada categoría de pronóstico.

En la figura 35, según los resultados del pronóstico probabilístico para la categoría 1 vemos que el día 02 de febrero, las mayores valores de probabilidad se concentraron en la sierra sur del Perú con probabilidades superiores al 70% para estaciones ubicadas en los departamentos de Puno, Arequipa y sur de Cuzco lo cual es coincidente con los registros de ocurrencia de precipitación para ese día (gráfico inferior). En los días 04 y 07 de febrero las probabilidades de precipitación aumentaron considerablemente, llegando a valores superiores de más de 90% en estaciones de la sierra central, producto de la inestabilidad producida por el posicionamiento del núcleo de la Alta de Bolivia en el sur de Perú y norte de Chile y su interacción con la circulación ciclónica ubicado en el pacifico oriental. Según los gráficos de ocurrencia de precipitación para dichos días hubo precipitaciones en gran parte de la cordillera peruana. En el día 09 de febrero las probabilidades disminuyen, concentrándose las mayores probabilidades en estaciones ubicadas en la sierra central.

Los resultados de la categoría 2 (figura 36) son muy similares al de la categoría 1, con la diferencia que en este caso las probabilidades son ligeramente menores a comparación de la categoría 1, debido a que la incertidumbre aumenta al pronosticar mayores cantidades de precipitaciones. Las mayores probabilidades se concentran en estaciones de la sierra sur, sierra centro y parte de sierra norte, en donde ocurrieron precipitaciones mayores a 1.00 mm, para los días 04 y 07 de febrero.

Para la categoría 3 (figura 37) las probabilidades disminuyen estando las más altas entre 50% y 70%. Debido a que hay mayor incertidumbre al pronosticar cantidades más altas de precipitación es importante el uso de los umbrales de probabilidades que mejor discrimine la ocurrencia del evento, los cuales fueron calculados a partir de la curva ROC (cuadro 11)

siendo este valor 30% para las regiones de la sierra y selva en la categoría 2. Bajo este enfoque, se asume que probabilidades pronosticadas de 50% son altas para el pronóstico de precipitaciones mayores a 2.5 mm.

Finalmente, en la categoría 4 (figura 38) las probabilidades son más bajas a comparación de las demás categorías, llegando los valores más altos a 55%. El umbral óptimo de probabilidad para la discriminación de los eventos es de 20% tanto como para la sierra y la selva.

Los umbrales de probabilidad estimados se utilizaron para cuantificar los valores de razón de acierto y de falsas alarmas en porcentajes para el periodo del 02 al 09 de Febrero del 2013. Estos resultados son mostrados en el cuadro 12. Los mejores resultados lo obtuvieron las estaciones de la sierra, con 81% de razón de aciertos para la categoría 1. Las estaciones de costa muestran los resultados más deficientes, con 25% de razón de aciertos a pesar de que tiene 0% de razón de falsas alarmas debido a que la precipitación es poco frecuente en dicha región y por ello es difícil pronosticar la ocurrencia de estos eventos. A mayor categoría de pronóstico tanto como la razón de aciertos y de falsas alarmas va disminuyendo debido a la dificultad de pronosticar probabilidades para mayores cantidades de precipitación.

Cuadro 12: Razón de aciertos y falsas alarmas correspondientes al periodo del 02-09 de Febrero del 2013

		CATEGORÍA 1	CATEGORÍA 2	CATEGORÍA 3	CATEGORÍA 4
COSTA	Razón de aciertos (HIT)	25%			
	Razon de falsas alarmas (FAR)	0%			
SIERRA	Razón de aciertos (HIT)	81%	73%	65%	49%
	Razon de falsas alarmas (FAR)	36%	46%	38%	22%
SELVA	Razón de aciertos (HIT)	64%	58%	51%	51%
	Razon de falsas alarmas (FAR)	39%	45%	41%	19%

Figura 35: Parte superior, pronóstico probabilístico MOS de la categoría 1 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 0.01mm.

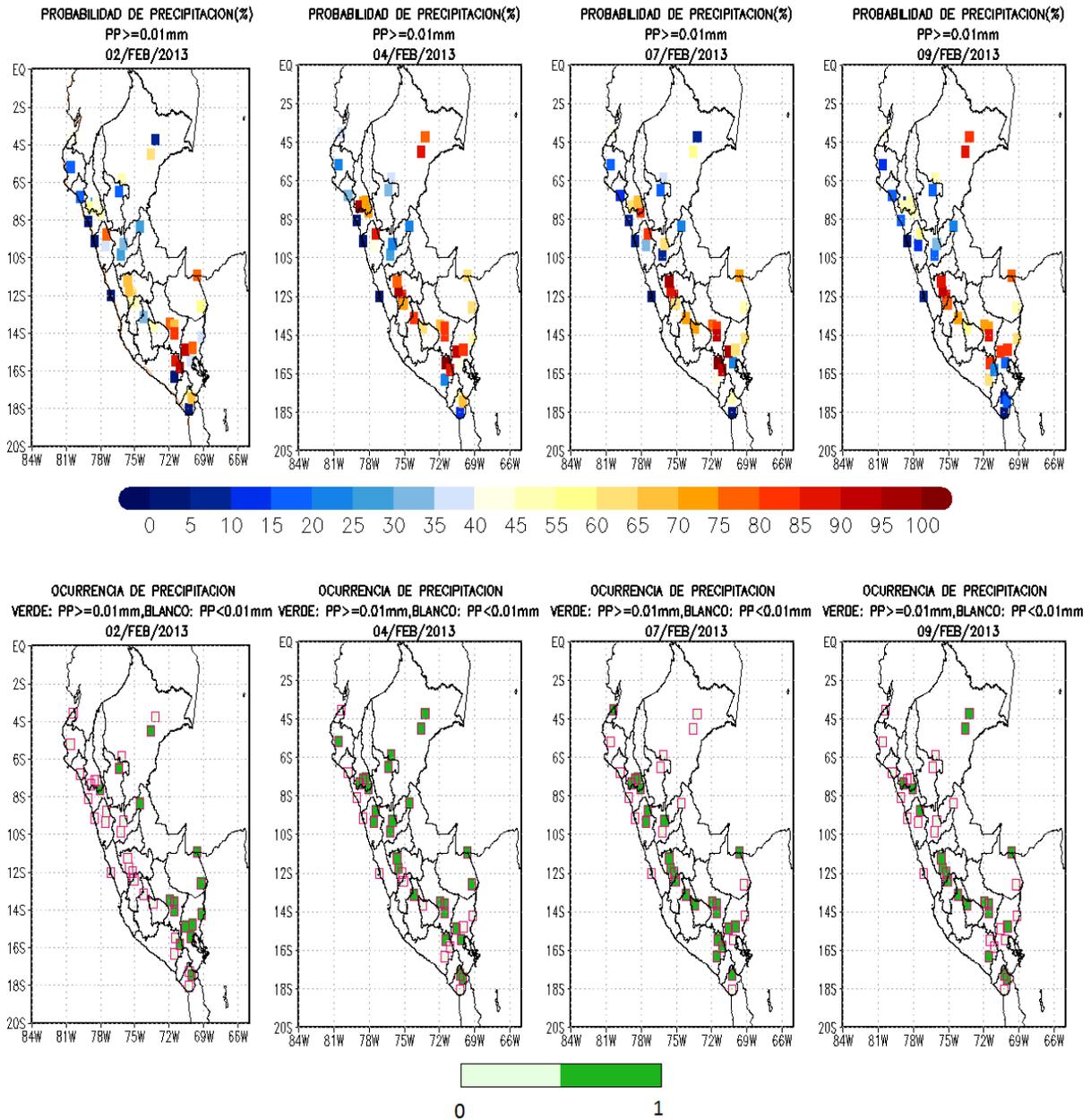


Figura 36: Parte superior, pronóstico probabilístico MOS de la categoría 2 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 1.00 mm.

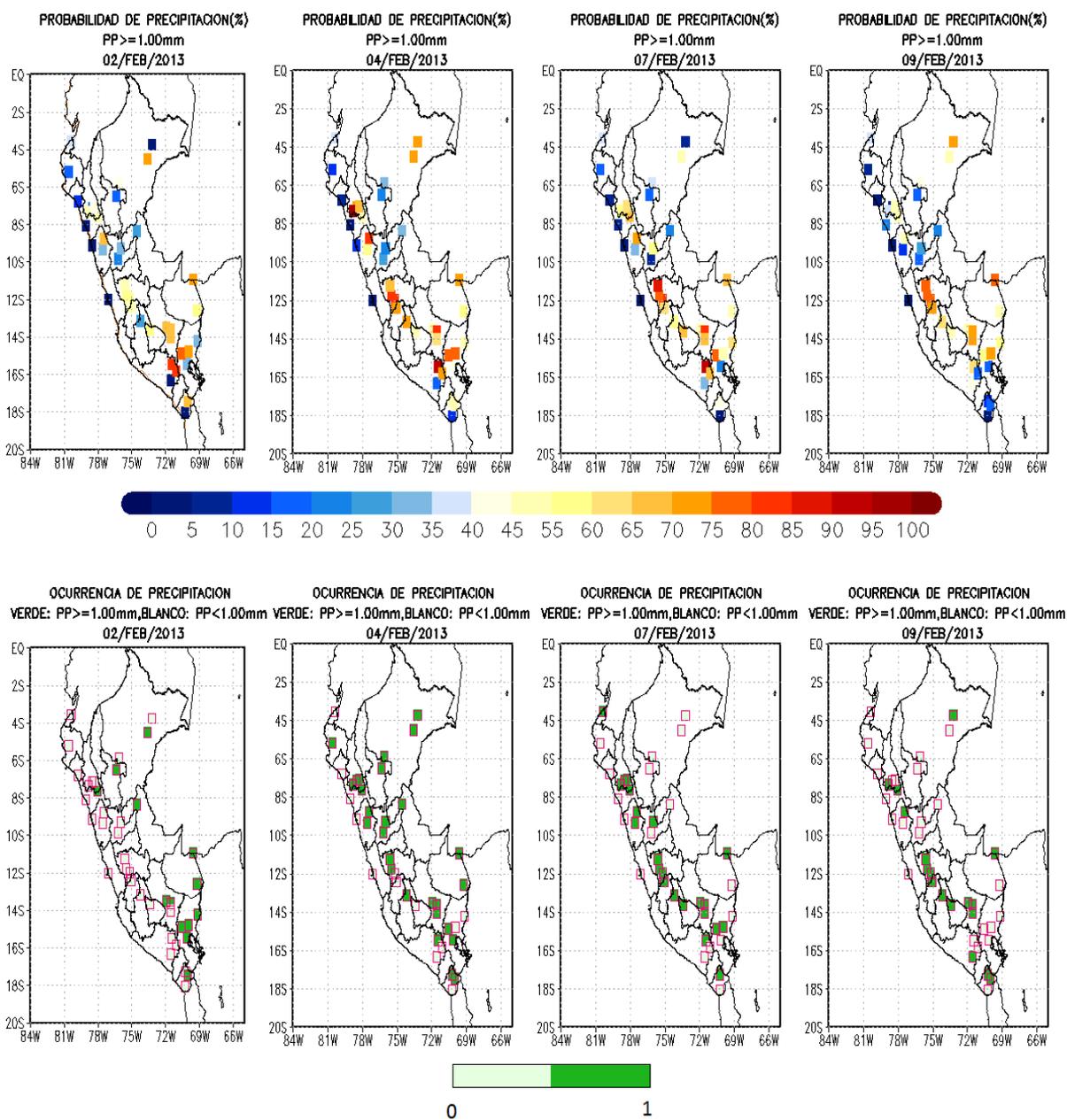


Figura 37: Parte superior, pronóstico probabilístico MOS de la categoría 3 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 2.50 mm.

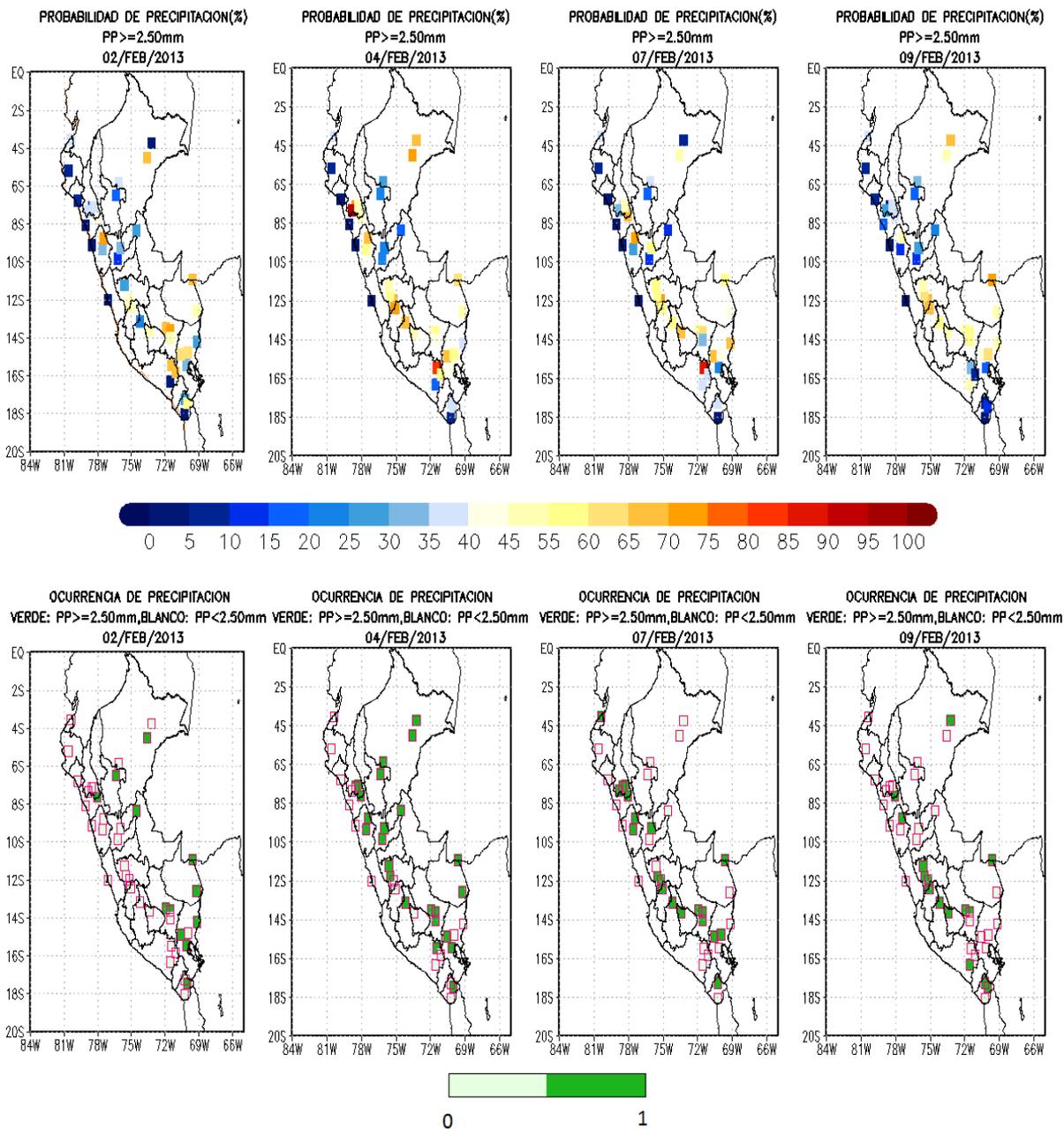
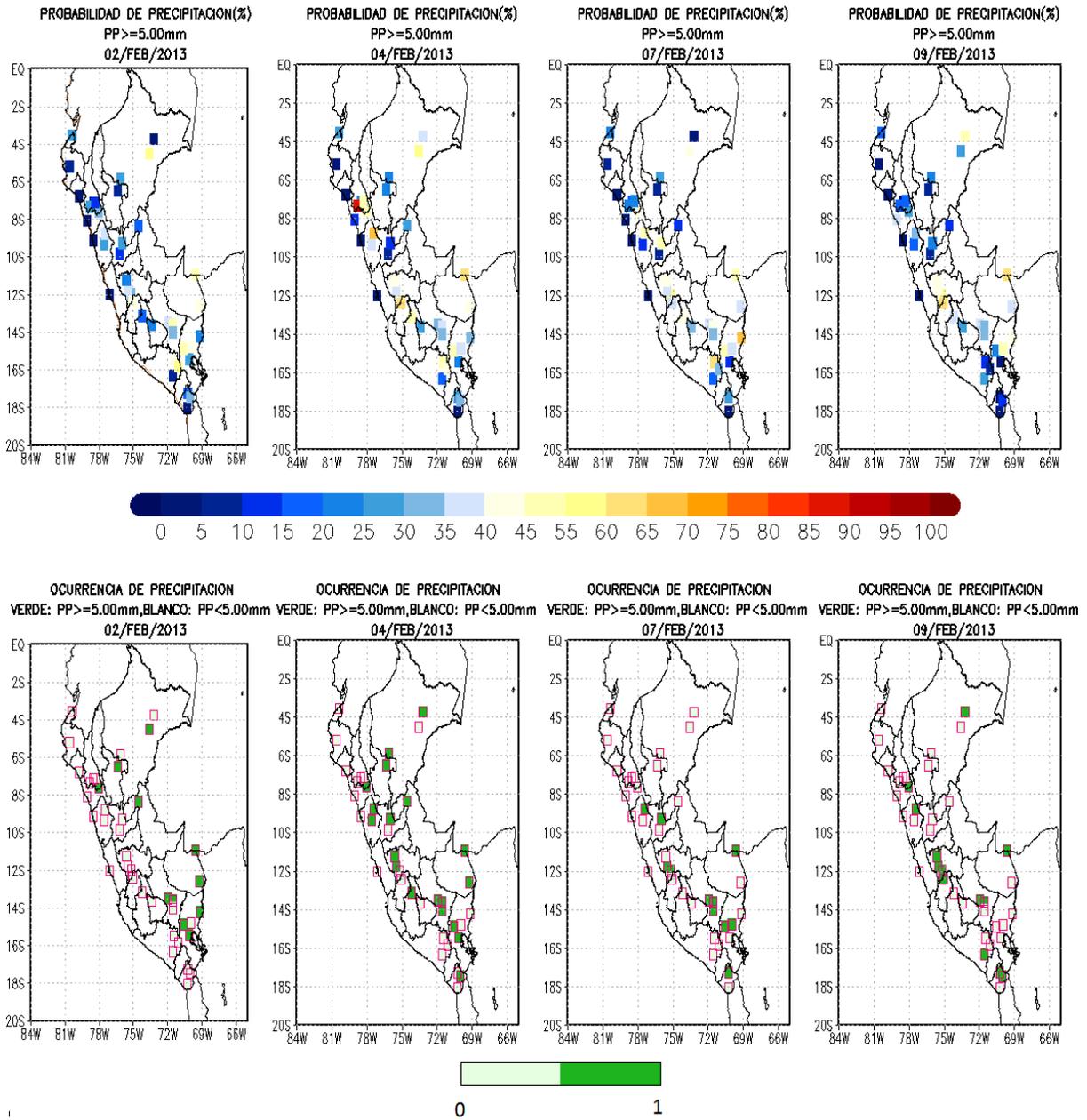


Figura 38: Parte superior, pronóstico probabilístico MOS de la categoría 4 para los días 02, 04, 07 y 09 de Febrero del 2013. Parte inferior, ocurrencia de precipitación observada para un umbral de 5.00mm.

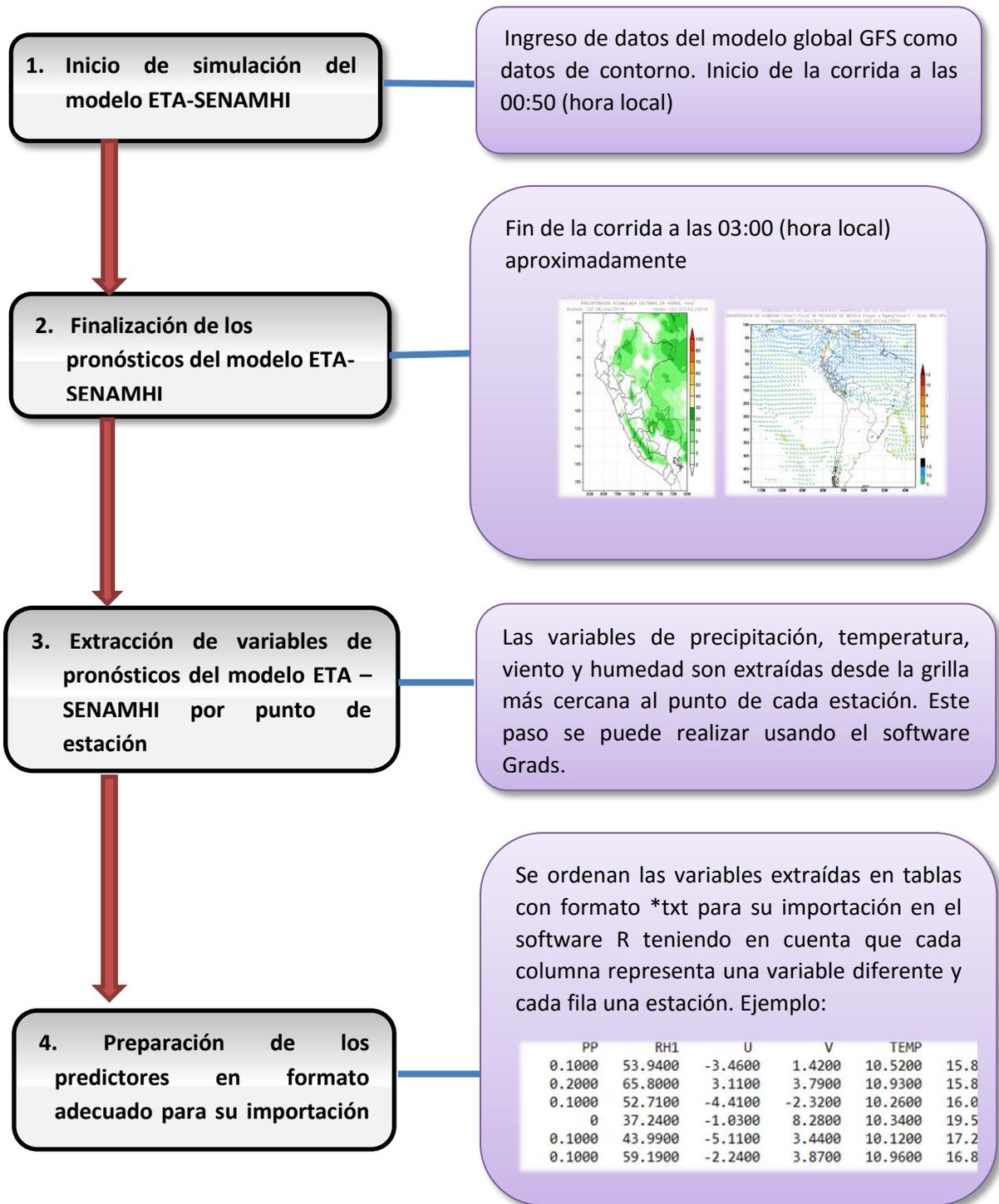


Finalmente, en la figura 39 se expone la secuencia de pasos para la implementación del sistema de pronóstico probabilístico en forma diaria basándose en la metodología que se ha desarrollado en la presente tesis. Esta secuencia de pasos está diseñada para que se pueda implementar y generar de forma automática los pronósticos probabilísticos cada día, iniciándose el proceso a las 00:50 (hora local) y finalizando a las 04:00 (hora local).

El proceso inicia en la corrida del modelo numérico ETA-SENAMHI con datos de entrada de pronósticos del modelo global GFS que deben ser descargados previamente, y para un análisis de las 00:00 UTC. El tiempo de corrida del modelo ETA-SENAMHI es de aproximadamente 2 horas y 10 minutos. Una vez finalizado, se extraen los datos de los predictores correspondientes a las grillas más cercanas de cada estación y se deben preparar estos en formato de tablas para su correcta importación en el software R; este formato consiste en tener los datos ordenados en columnas por cada variable. El siguiente paso es generar los pronósticos probabilísticos usando los datos de predictores ordenados y los parámetros de regresión a nivel mensual que se han calculado en la presente tesis. La automatización de la ruta operativa se realiza mediante la preparación de scripts en lenguaje de programación en Shell y en R.

Además de la automatización de la generación de los pronósticos probabilísticos diarios, es recomendable generar también nuevos parámetros de regresión al menos cada año para que estos tengan una mayor precisión basándose en una muestra de datos más largos; y de igual forma, la validación también se debe realizar de forma continua con los nuevos parámetros de regresión y los nuevos datos de precipitación observada.

Figura 39. Ruta operativa para la generación de pronósticos probabilísticos



(Continuación de la figura 39)

5. Ingreso de variables pronosticadas a modelos de regresión logística y generación de pronósticos

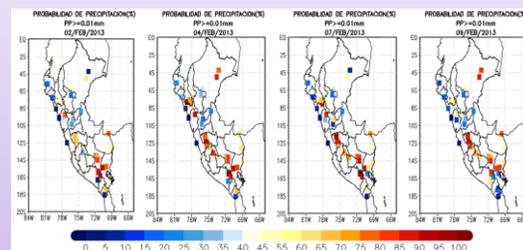
- Importación de datos de predictores en formatos *.txt al software R.
- Para el cálculo de las probabilidades se debe usar la librería en R *Bias reduction in binomial-response generalized linear models* (brglm).
- Con los parámetros obtenidos del proceso de calibración y los datos de predictores del modelo numérico ETA-SENAMHI se estima la probabilidad pronosticada en el software R en base a modelos de regresión logística. Como un ejemplo se muestra el modelo para estimar la probabilidad de que ocurra una precipitación mayor a 1 mm en la ciudad de Cusco para el mes de Enero:

$$Pr\{y\} = \frac{1}{1 + e^{-(-2.57 + 0.11 * RH_{n1} - 0.13 U_{n3} - 0.01 * PP_{acum} - 0.21 WW)}}$$

- Finalmente, el software R exporta los resultados del pronóstico probabilístico en formato *.txt. Fin del proceso a las 04:00 (hora local) aproximadamente

6. Presentación de resultados de pronóstico probabilístico probabilísticos

Elaboración de mapas y tablas para la difusión de resultados



V. CONCLUSIONES

1. Se generaron modelos de regresión logística a nivel mensual usados para el pronóstico probabilístico diario de la precipitación. Un paso importante fue la selección adecuada de los predictores los cuales están relacionados a los mecanismos físicos que originan la precipitación para cada región del país. Se encontró que para la costa los predictores óptimos son las variables de humedad y temperatura en niveles bajos; mientras que en la sierra, la influencia es mayor con las variables de humedad y viento zonal en niveles medios y altos; y para la selva resultaron más significativas las variables de viento meridional y temperatura en niveles bajos a medios. Asimismo, el índice de estabilidad W y la precipitación diaria acumulada del modelo ETA-SENAMHI fueron las variables que presentaron una alta frecuencia de significancia estadística en los modelos de regresión para las tres regiones.
2. Considerando los diferentes atributos del pronóstico probabilístico (fiabilidad, precisión, habilidad y resolución) para el análisis de la eficiencia, se concluye que, los atributos de fiabilidad, habilidad y resolución presentan resultados muy similares y tienen una relación directa a pesar de medir aspectos diferentes. El atributo de precisión no es una medida recomendable debido a que es sensible cuando la frecuencia climatológica del evento es muy baja.
3. Se concluye que las estaciones ubicadas en la costa presentan los pronósticos probabilísticos más deficientes debido principalmente a una inadecuada simulación por parte del modelo ETA-SENAMHI. Esto se debe principalmente a que el modelo no tiene una adecuada simulación de la capa de inversión tal y como manifiestan Betts y Miller (1986). Las estaciones de la sierra presentan los resultados más óptimos, evidenciándose en la eficiente fiabilidad lo cual influyó también en que se obtuviera la mayor habilidad (0.5 de Brier Skill Score como promedio para la categoría 1).

4. Los resultados para las estaciones de la selva presentaron eficiente fiabilidad y menor habilidad (valores de 0.2 de Brier Skill Score como promedio) comparando con los resultados de las estaciones de la sierra. Además, poseen la menor resolución lo cual puede deberse a que los predictores seleccionados no fueron los más adecuados.
5. A nivel estacional se encontraron resultados más eficientes en el periodo Diciembre-Enero-Febrero ya que en este periodo la atmósfera es más barotrópica y los patrones de circulación a escala sinóptica que favorecen la precipitación son representados adecuadamente por el modelo ETA-SENAMH. El periodo Junio-Julio-Agosto se asocia a una atmósfera menos barotrópica y con menor advección de humedad de la hoya amazónica, lo cual conduce que las frecuencias climatológicas de precipitación diaria disminuyen y los pronósticos probabilísticos fueron deficientes.
6. Respecto a las categorías de pronóstico, se concluye que los resultados más eficientes de probabilidades se presentan en la categoría 1 (umbral de 0.01 mm). A medida que aumenta de categoría (umbrales 1, 2.5 y 5 mm) la eficiencia del pronóstico probabilístico disminuye debido al aumento de incertidumbre del pronóstico a cantidades mayores de precipitación, especialmente en estaciones donde son menos frecuente la ocurrencia de cantidades más altas de precipitación diaria.
7. Las lluvias producidas durante el periodo del 02 al 09 de Febrero del 2013 se produjeron por un posicionamiento del núcleo de la Alta de Bolivia al norte de Chile y un sistema ciclónico que bloqueaba su paso hacia el oeste, favoreciendo al transporte de humedad e inestabilidad en los sectores sur y centro de la sierra y selva. Estos días lluviosos fueron representados adecuadamente por el pronóstico probabilístico MOS para las estaciones de la sierra con valores de 81% de razón de aciertos en la categoría 1, mientras que en la costa se obtuvieron solamente 25% de razón de acierto en la categoría 1. Además, la razón de aciertos disminuye cuando la categoría es mayor lo cual nos sugiere que se debe usar con reserva los pronósticos de probabilidades para mayores cantidades de precipitaciones.

VI. RECOMENDACIONES

1. Dado que la fiabilidad no fue posible encontrar en estaciones de costa, se recomienda seguir realizando más pruebas con otros tipos de predictores que sirvan de inputs para los modelos de regresión con el fin de mejorar el pronóstico probabilístico.
2. Se recomienda utilizar salidas de otros modelos de predicción numérica del tiempo, como por ejemplo el modelo WRF (Weather Research and Forecasting), ya que podrían dar resultados más eficientes a comparación a las que se hicieron con salidas del modelo ETA-SENAMHI. Previo a ello, previamente se debe hacer un análisis de la simulación atmosféricas sobre nuestro país por parte de estos modelos, de lo contrario, se necesitaría hacer ajustes ya sea de las condiciones iniciales y/o de las parametrizaciones para obtener las representación más óptima.
3. Se recomienda hacer análisis del pronóstico cuantitativo de la precipitación expresados en probabilidades para un plazo de previsión mayor de previsión. Además, en estaciones cuya climatología nos indica que la precipitación diaria acumulada puede sobrepasar fácilmente los 5mm, se recomienda trabajar con pronósticos probabilísticos para mayores umbrales de precipitación (por ejemplo, 10 mm o 20mm) siendo esta información mucho más útil para los usuarios que se encuentren en las localidades donde estén ubicadas dichas estaciones para la reducción de riesgo ante posibles desastres.
4. Se recomienda en el futuro realizar investigaciones con otras metodologías para generar pronósticos probabilísticos tales como las regresiones lineales múltiples, redes bayesianas, redes neuronales y los pronósticos por conjuntos (ensembles) con el fin de obtener los modelos más óptimos.

VII. REFERENCIAS BIBLIOGRÁFICAS

- Aldeco, L. S. (2011). Aplicación de la técnica de análogos a la generación de pronósticos probabilísticos de precipitación sobre algunas estaciones de la Argentina. Tesis de Licenciatura. Universidad de Buenos Aires, Argentina.
- Ancell, R. (2009). Aplicaciones de las redes bayesianas en Meteorología. Predicción probabilística de precipitación. Tesis doctoral. Universidad de Cantabria. Santander. 228 pp.
- Applequist, S.; Gahrs G. E.; Pfeffer R. L. y Niu X. (2002). Comparison of Methodologies for Probabilistic Quantitative Precipitation Forecasting (en línea). *Wea. Forecasting*, **17**, 783–799. Consultado 13 nov. 2013. Disponible en: <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0434%282002%29017%3C0783%3ACOMFPQ%3E2.0.CO%3B2>
- Avalos G. J. (2009). Evaluación del modelo ETA/SENAMHI en la previsión de la precipitación durante los periodos lluviosos 2002 y 2003. *Revista peruana geo atmosférica* (1), 108 – 124.
- Arnau, J. (1996). Métodos y técnicas avanzadas de análisis de datos en ciencias del comportamiento. Universidad de Barcelona
- Azcarraga, R. y Ballester, G. (1991). Statistical system for forecasting in Spain. *World Meteorological Organization (WMO)*. No 421, XX23-25.
- Badii, M.H.; Guillen A.; Serrato L. y Aguilar J. (2014). Correlación No-Paramétrica y su Aplicación en la Investigaciones Científica. *Non-Parametric Correlation and Its*

- Application in Scientific Research. *International Journal of Good Conscience*. 9(2)31-40. Revisado el 15 abr. 2014. Disponible en [http://www.spentamexico.org/v9-n2/A5.9 \(2\)31-40.pdf](http://www.spentamexico.org/v9-n2/A5.9%20(2)31-40.pdf)
- Betts, A y Miller, M. (1986). A new convective adjustment scheme. Part II: Single column test using GATE wave, BOMEX, and arctic air-masses data sets. *Meteor. Soc.*, 112, pp. 1306-1335
- Brier, G.W. (1950). Verification of forecasts expressed in terms of probability (en línea). *Mon. Wea. Rev.*, Vol. 78, 1-3. Consultado 10 nov. 2013. Disponible en: <http://docs.lib.noaa.gov/rescue/mwr/078/mwr-078-01-0001.pdf>.
- Brocker, J. y Smith, L.A. (2007). Increasing the reliability of reliability diagrams. *Weather Forecast*. 22, 651–661.
- Brunet, N.; Verret R. y Yacowar, N. (1988). An objective comparison of model output statistics and “perfect prog” systems in producing numerical weather element forecasts. *Weather and Forecasting*, 3, 273–283.
- Burns, R. (2008). Business Research Methods and Statistics Using SPSS (en línea). *Sage Publications Inc, London*. Consultado 02 abr. 2014. Disponible en: <http://books.google.com.pe/books?id=bPvCRzBou3gC&printsec=frontcover&hl=es#v=onepage&q&f=false>
- Carter, G.M., Dallavalle, J.P. y Glahn, H.R. (1989). Statistical forecasts based on the National Meteorological Center’s numerical weather prediction system. *Weather and Forecasting*, 4, 401–412.
- Chou, S. C., y Justi Da Silva, M. G. A. (1999). Objective evaluation of Eta model precipitation forecasts over South America (en línea). *Climanálise*, 14(1), 235-239. Consultado 11 abr. 2014. Disponible en:

http://www.researchgate.net/publication/37680112_Objective_evaluation_of_ETA_Model_precipitation_forecasts_over_South_America/file/60b7d517ae325ce852.pdf.

Chourio, M. C. (2009). Aplicación de índices de estabilidad atmosférica al pronóstico meteorológico de la región centro norte costera de Venezuela (en línea). Trabajo especial de grado. Universidad Central de Venezuela. 89p. Consultado 14 abr. 2014. Disponible en: http://www.inameh.gob.ve/documentos/indices_estabilid_venezuela.pdf.

Consejo Nacional del Ambiente (CONAM). (2005). Atlas Climático de precipitación y temperatura del aire de la Cuenca del Río Mantaro. *Fondo Editorial del CONAM*. Consultado 03 abr. 2014. Disponible en: http://www.met.igp.gob.pe/publicaciones/2000_2007/Atlas_Climatico.pdf

Condom, T.; Rau, P.; y Espinoza, J. C. (2010). Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007 (en línea). Consultado 01 abr. 2014. *Hydrological Processes*, 25(12), 1924–1933. Disponible en: <http://www.lamolina.edu.pe/postgrado/doctorh/2012/Correction.pdf>

Conte, M., C. Desimone, y C. Finizio, (1980). Post-processing of numerical models: forecasting the maximum temperature at milano linete. *Rev. Meteor. Aeronautica*, **40**, 247–265.

Crochet, P. (2003). A statistical model for predicting the probability of precipitation in Iceland (en línea). Vedurstofa Islands. *Report 03028*. Consultado 03 abr. 2014. Disponible en: <http://www.vedur.is/media/vedurstofan/utgafa/greinargerdir/2003/03028.pdf>.

Eccel, E., Ghielmi, L., Granitto, P., Barbiero, R., Grazzini, F., And Cesari, D. (2007). Prediction of minimum temperatures in an alpine region by linear and non-linear post-processing of meteorological models. *Nonlinear Processes in Geophysics*, 14 (3), pp. 211–222.

Espinoza V. J. C., Ronchail J., Lavado W., Carranza J., Cochonneau G., De Oliveira E., Pombosa R., Vauchel P. y Guyot J.L. (2009). Variabilidad espacio-temporal de las lluvias en la cuenca amazónica y su relación con la variabilidad hidrológica regional. Un enfoque particular sobre la region andina. Spatio-temporal rainfall variability in the amazon basin and its relation with regional hydrological variability.a particular focus in the andean region. *Revista Peruana Geo-Atmosférica*, 1: 66-89.

Figuroa, M.R. (2010). Modelos matemáticos de apoyo a la gestión para la prevención de incendios forestales. Memoria para optar el título de Ingeniero Civil Matemático. Consultado 02 mayo. 2014. Universidad de Concepción. 131 pp.

Figuroa, S.N.; Satyamurty, P.; y Silva Dias, P.L. (1995). Simulation of the summer circulation over the South American region with an Eta coordinate model (en línea). *Journal of the atmospheric sciences*, 52 (10), 1573-1584. Consultado 03 abr. 2014. Disponible en: [http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493\(1999\)127%3C0901%3AMAOTSP%3E2.0.CO%3B2](http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493(1999)127%3C0901%3AMAOTSP%3E2.0.CO%3B2).

Francis, P. E.; Day, A. y Davis G. (1982). Automated temperature forecasting, an application of model output statistics to the Meteorological Office numerical weather prediction model. *Meteor. Mag.* 111, 73-87.

Freeman, D.H., 1987. Applied Categorical Data Analysis. *Marcel Decker*, 318 pp.

Garreaud, R. D. (1998). Multiscale analysis of the summertime precipitation over the central Andes (en línea). *Monthly Weather Review-Usa*, 127, 901-921. Consultado 04 abr. 2014. Disponible en: <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%281999%29127%3C0901%3AMAOTSP%3E2.0.CO%3B2>.

Garreaud, R. D. (2000a). Intraseasonal variability of moisture and rainfall over the South American Altiplano (en línea). *Monthly Weather Review* 128:3337-3346. Consultado 02

abril de 2014. Disponible en:
http://www.dgf.uchile.cl/rene/PUBS/Altiplano_intra_MWR.pdf

Garreaud, R. D. (2000b). Cold air incursions over subtropical South America: mean structure and dynamics (en línea). *Monthly Weather Review* 128:2544-2559. Consultado 11 abr. 2014. Disponible en: http://www.dgf.uchile.cl/rene/PUBS/cusrges_clima_MWR.pdf

Garreaud, R. D y Aceituno, P. (2007). Atmospheric circulation over South America: Mean features and variability. Chapter 2 in *The Physical Geography of South America*. T. Veblen, K. Young and A. Orme, Eds. Oxford University Press. Consultado 01 abr. 2014. Disponible en: http://dgf.uchile.cl/rene/PUBS/book_PHGSA.pdf

Garreaud, R. y Seluchi, M. (2003). Pronóstico de la convección en el altiplano sudamericano: empleo del modelo regional ETA/Cptec (en línea). *Meteorológica* 26, N°1 y 2 25 – 38.

Garreaud, R.; Vuille, M. y Clement, A. (2003). The climate of the Altiplano : Observed current conditions and mechanisms of past changes (en línea). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 194, 5-22. Consultado 07 abr. 2014. Disponible en: http://www.atmos.albany.edu/facstaff/mathias/pubs/Garraud_et_al_2003.pdf.

Glahn, H R. y Lowry D.A. (1972). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.

Gutierrez, J.M.; Cano, R; Cofiño A. y Sordo C. (2004). Redes probabilísticas y neuronales en las ciencias atmosféricas. Series Monográficas del Instituto Nacional de Meteorología. Universidad de Cantabria. Madrid.

Hagedorn, R. y Smith L. (2009). Communicating the value of probabilistic forecasts with weather roulette, *Met. Appl.*, 16, 143-155.

Hogg, R. y Craig, A. (1994). Introduction to mathematical statistics. 5th ed. Upper Saddle River, NJ: Prentice Hall.

Instituto Geofísico Del Perú (IGP). (2010). Memoria del Subproyecto “Pronóstico estacional de lluvias y temperaturas en la cuenca del río Mantaro para su aplicación en la agricultura” (en línea). Consultado 08 abr. 2014. Disponible en: <http://www.met.igp.gob.pe/proyectos/incagro/datos/libromemoria.pdf>

Instituto Nacional De Defensa Civil (INDECI). (2003). Atlas de peligros naturales del Perú (en línea). Disponible en: <http://www.bvcooperacion.pe/biblioteca/bitstream/123456789/5473/1/BVCI0005224.pdf>

Illanes, L. P. (2010). Sensibilidad a diferentes parametrizaciones de capa límite atmosférica y suelo con el modelo wrf, para el llano de chajnantor–chile. Postulación al título de Meteorólogo. Universidad de Valparaíso. Chile. 98 pp

Jacks, E., Bower, J. B., Dagostaro, V. J., Dallavalle, J. P., Erickson, M. C. y Su, J. C. (1990). New NGM-based MOS guidance for maximum/minimum temperature, probability of precipitation, cloud amount, and surface wind. *Weather and Forecasting*, 5(1), 128-138

Karl, T.R.; Schlesinger, M.E. y Wang, W. C. (1989). A method of relating general circulation model simulated climate to the observed local climate. Part I: Central tendencies and dispersion (en línea). Preprints, Sixth Conference on Applied Climatology. *American Meteorological Society*, 188–196.

Kousky, V. E. y Kagano. M. T. (1981). A climatological study of the tropospheric circulation over the Amazon region. *AcraAmazonico* 11, 743-758.

Kousky, V.E. y M. A. Gan, (1981). Upper tropospheric cyclonic vortex in the tropical South Atlantic (en línea). *Tellus*, 33, 538-551. Consultado 09 abr. 2014. Disponible en: <http://onlinelibrary.wiley.com/doi/10.1111/j.2153-3490.1981.tb01780.x/pdf>

- Lagos, P.; Silva, y.; Nickl, E. y Mosquera. K. (2008). El Nino – related precipitation variability in Peru (en línea). *Advances in Geosciences*, 14, 231-237. Consultado 11 nov. 2013. Disponible en: <http://hal.archives-ouvertes.fr/docs/00/29/71/03/PDF/adgeo-14-231-2008.pdf>
- Lemcke C, y Kruizinga S. (1988). Model Output Statistics forecasts: Three years of operational experience in the Netherlands. *Mon. Weather Rev.* 116: 1077–1090.
- Lenters, J. D. y Cook, K. H. (1995). Simulation and Diagnosis of the Regional Summertime Precipitation Climatology of South America (en línea). *Journal Climate*, 8, 2988–3005. Consultado 10 abr. 2014. Disponible en: <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0442%281995%29008%3C2988%3ASADOTR%3E2.0.CO%3B2>.
- Lighezzolo, R. (2014). Integración de modelos numéricos de predicción meteorológica en sistema de alerta temprana a emergencias. Tesis de maestria. Universidad Nacional de Cordoba. Consultado 11 abr. 2014. Disponible en : <http://www.famaf.unc.edu.ar/wp-content/uploads/2014/05/21-Gulich-Lighezzolo.pdf>
- Llacza, A. (2013). Modelo ETA- SENAMHI en el. Pronóstico operativo (diapositivas). IV WORKETA – 03 al 08 marzo 2013.
- Lopez, L. C. (2012). Evaluación de la calidad del pronóstico numérico del tiempo en la ciudad de Mexico. Tesis para obtener el grado de Ingeniero Geofísico. *Universidad Nacional Autónoma de México*. 73 pp. Consultado el 10 de mar. 2014. Disponible en: http://132.248.9.195/ptd2012/junio/0681178/0681178_A1.pdf
- Lorenz, E.N. (1963). Deterministic non periodic flow (en línea). *Journal of the Atmospheric Sciences*, **20**, 130–141. Consultado 17 abr. 2014. Disponible en: <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0469%281963%29020%3C0130%3ADNF%3E2.0.CO%3B2>

- Lu, R. (1991). The application of NWP products and progress of interpretation techniques in China. In: H.R. Glahn, A.H. Murphy, L.J. Wilson, and J.S. Jensenius, Jr., eds., *Programme on Short-and Medium-Range Weather Prediction Research*. World Meteorological Organization WM/TD No. 421. XX19–22.
- Maloney, J. C. (2002). Eta-based MOS probability of precipitation (PoP) and quantitative precipitation forecast (QPF) guidance for the continental United States (en línea). *NWS Technical Procedures Bulletin* No 487. Silver Spring, MD: National Oceanic and Atmospheric Administration. Consultado 12 abr. 2014. Disponible en: <http://www.nws.noaa.gov/mdl/synop/tpb/487.pdf>
- Marengo, J.A. (1983). Estudio agroclimático en la zona de Jenaro Herrera (Requena-Loreto) y climático en la Selva Baja Norte del Perú. Tesis para optar el título de Ingeniero Meteorólogo. Universidad Nacional Agraria La Molina. 441p.
- Martínez, R.; Ruiz, L.; Andrade, M.; Biacutt, L.; Pabón, D.; Jaimes, E.; León, G.; Viiiicís, M.; Quintana, J.; Montealegre, E.; y Euscátegui, C. (2001). Síntesis del Clima de los Andes Tropicales (en línea). Capítulo seis del libro *Cambio Climático y Biodiversidad en los Andes Tropicales*. Consultado 13 abr. 2014. Disponible en: http://www.iai.int/files/communications/publications/scientific/Climate_Change_and_Biodiversity_in_the_Tropical_Andes_SP/capitulo6.pdf.
- Matsudo, C.; García, Y; Ferreira, L.; Ruiz, J.; Salio, P.; Vidal, L. y Nicolini, M. (2013). Pronósticos determinístico y probabilístico en alta resolución con el modelo BRAMS (en línea). Consultado 13 abr. 2014. Disponible en: http://www.ina.gov.ar/pdf/ifrrhh/03_037_Matsudo.pdf.
- Medeiro, L. (2007). Pronóstico probabilístico de caudales de avenidas mediante redes bayesianas aplicadas sobre un modelo hidrológico distribuido. *Universidad de Cantabria*. 490 pp.

- Mesinger, F. (1984). A blocking technique for representation of mountains in atmospheric models. *Riv. Meteor. Aeronaut.*, 44, 195-20.
- METED. (2010). Funcionamiento de los modelos de mesoescala (en línea). COMET. Consultado 03 marzo 2014. Disponible en: www.meted.ucar.edu/mesoprim/models_es/
- METOFFICE. (2013). Reliability and sharpness diagrams (en línea). Consultado 12 abr. 2014. Disponible en: <http://www.metoffice.gov.uk/research/areas/seasonal-to-decadal/gpc-outlooks/user-guide/interpret-reliability>
- Ministerio Nacional Del Ambiente. (2011). Mapa de ecorregiones del Perú Escala 1/2000000. Viceministerio de Desarrollo Estratégico de Recursos Naturales. Disponible en:http://geoservidor.minam.gob.pe/geoservidor/Archivos/Mapa/N53_ECORREGIONES.pdf
- Nishizawa, T., y M. Tanaka. (1983). The annual change in the tropospheric circulation and the rainfall in South America. *Arch. Meteor. Geophys. Bioclimatol.*, 33B, 107–116.
- Peñaherrera, C. (1969). Geografía general del Perú. *Ausonia*. 313 pp.
- Pita S., Pértega S. (1997). Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario Juan Canalejo. A Coruña. *Cad Aten Primaria*; 4: 141-144.
- Pielke R. (2002). Mesoscale Meteorological Modeling. Academic Press, 693 pp.
- Popovic, J. (2006). ETA model in Weather Forecast. Master's Thesis in Numerical Analysis. Royal Institute of Technology. 56 pp.
- Quispe N. (2004). Vórtice Ciclónico En Bajos Niveles Caso Estudio: 09 De Febrero De 2004 (en línea). Consultado 03 abr. 2014. Disponible en: http://www.senamhi.gob.pe/pdf/estudios/vcn_ram.pdf

- Quispe, N. (2005). Condiciones Sinópticas de la Incursión de una Masa de Aire Frío en Sudamérica y su Impacto en los Cultivos de la Selva Caso específico: verano de 1996 e invierno de 2000. Tesis para optar el Título de Ingeniero Meteorólogo. *Universidad Nacional Agraria La Molina*.122p.
- Raghavendra, B. y Srivatsa, S. (2011).Evaluation of Logistic Regression and Neural Network Model with Sensitivity Analysis on Medical Datasets (en línea). *International Journal of Computer Science and Security (IJCSS)*, Vol. 5, Issue 5, pp. 503-511. Consultado 08 abr. 2014.Disponible en:
<http://www.cscjournals.org/csc/manuscript/Journals/IJCSS/volume5/Issue5/IJCSS-568.pdf>
- Rincón, A. (2013). Sistema de pronóstico de radiación solar a corto plazo a partir de un modelo meteorológico y técnicas de post-proceso para España. Tesis doctoral. *Universidad Nacional de Asunción*. 279 pp.
- Roballino, A. (2010). Automatización de un Modelo numérico de predicción metereologicaWeatherResearch And ForecastModel – Wrf (en línea). *Universidad de Huelva – España*. Consultado 09 abr. 2014. Disponible en: http://uhu.es/noticieros/master-ta/files/ROBALINO_LOPEZ_RESUMEN.pdf
- Ruiz, J. (2009). Pronóstico probabilístico ¿por qué? y ¿para quién? (diapositivas). CONGREGMET X, CLIMET XIII. Departamento de Cs. de la Atmósfera y los Océanos Centro de Investigaciones del Mar y la Atmósfera. Consultado 11 abr. 2014. Disponible en: <http://www.at.fcen.uba.ar/docs/presentacionRuiz.pdf>
- Satyamurty P., Nobre A., y Silva Dias P. (1998). South America. In *Meteorology of the Southern Hemisphere*. In: *Meteorological Monographs*, vol. **27**, No. 49, KarolyDJ, VincentDG . *American Meteorological Society*: Boston, Massachusetts. pp. 119–139.

Schmeits M. J.; Kok K., Vogelesang D. y Van Westrhenen R. (2008). Probabilistic Forecasts of (Severe) Thunderstorms for the Purpose of Issuing a Weather Alarm in the Netherlands. *Wea. Forecasting*, **23**, 1253–1267.

Schwerdtfeger, W. (1961) Flujo y campo de temperatura de la atmósfera libre sobre los Andes. *Meteor. Rundsch.*, 14, 1-6. Consultado 13 marzo 2014.

Servicio Nacional de Meteorología e Hidrología. (2007). Escenarios de cambio climático en la Cuenca del río Urubamba para el año 2100 (en línea). Consultado 12 marzo 2014. Disponible en:
http://www.senamhi.gob.pe/main_down.php?ub=cmn&id=PRAA_resu_tec_escenarios_2100_URUBAMBA

Sharan, M. y Raman, S. (2007). Atmospheric and Oceanic Mesoscale Processes. Birkhäuser (en línea). Consultado 01 abr. 2014. Disponible en:
<http://books.google.com.pe/books?id=bsDvIXtAvxMC&printsec=frontcover#v=onepage&q&f=false>

Solis, O.; Quispe, N. y Ferrari, C. (2012). Software Climatología Sinóptica Clisyn v1.0. Disponible en:
http://www.senamhi.gob.pe/main_down.php?ub=zip&id=INSTALADOR_CLISYN

Strand, S.; Cadwallader, S. y Firth, D. (2011). Using statistical regression methods in education research. *ESRC National Center for Research Methods*. Disponible en
<http://www.restore.ac.uk/srme/www/fac/soc/wie/research-new/srme/index.html>

Subías, A.; Pascual, A.; Martín, M.; Valero, F. y Luna, M. (2012). Validación de un modelo de análogos para la predicción de precipitación y nieve en la sierra de Guadarrama (en línea). Consultado 08 abr. 2014. Disponible en:
<http://fundacion.usal.es/conaec/pendrive/ficheros/ponencias/ponencias1/17-Modelos.pdf>

Takahashi, K. 2004. Statistical downscaling of rainfall in the Mantaro basin. *Scientific report to PROCLIM*. 17 pp.

The Centre for Australian Weather and Climate Research. (Cawcr). Methods and scores used for verifying ensemble forecasts (en línea). Consultado 12 abr. 2014. Disponible en: <http://www.cawcr.gov.au/projects/EPsverif/scores/scores.html>

Virji, H. (1981). A Preliminary Study of Summertime Tropospheric Circulation Patterns over South America Estimated from Cloud Winds. *Mon. Wea. Rev.*, **109**, 599–610. Consultado 07 abr. 2014. Disponible en: <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%281981%29109%3C0599%3AAPSOST%3E2.0.CO%3B2>.

Weiss, J. (2006). *Statistical Analysis in Ecology and Evolution*. University of North Carolina. Disponible en <https://www.unc.edu/courses/2006spring/ecol/145/001/docs/lectures.htm>

Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences*. 2d ed. *International Geophysics Series, Vol. 91*, Academic Press, 627 pp.

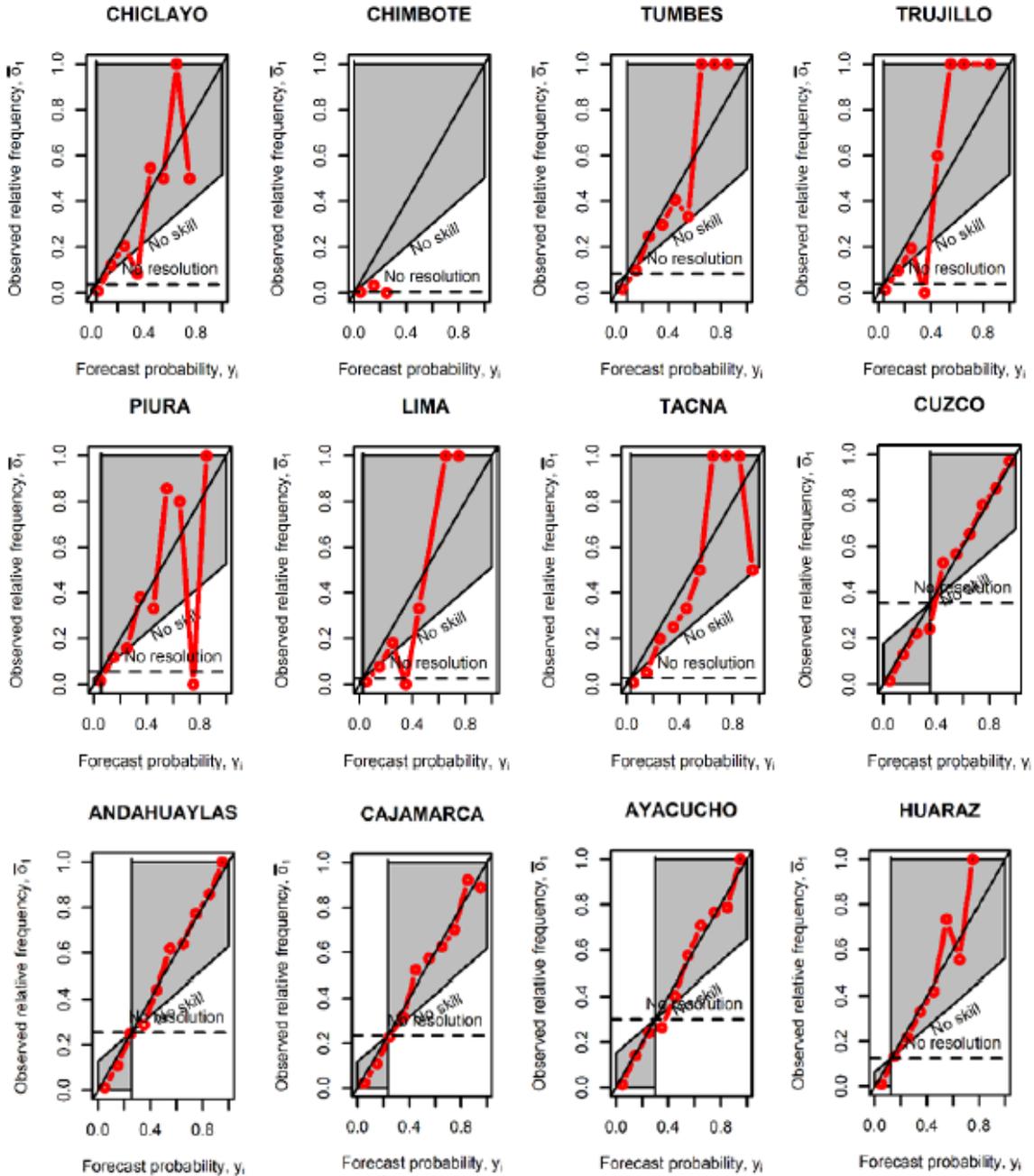
Woodcock, F. (1984). Australian Experimental Model Output Statistics Forecasts of Daily Maximum and Minimum Temperature (en línea). *Mon. Wea. Rev.*, 112, 2112–2121. Consultado 08 abr. 2014. Disponible en: <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%281984%29112%3C2112%3AAEMOSF%3E2.0.CO%3B2>

World Meteorological Organization (WMO). (2003). Documentation and software for verification scores: Reliability diagrams (en línea). Consultado 06 abr. 2014. Disponible en: <http://www.bom.gov.au/wmo/lrfvs/reliability.shtml>.

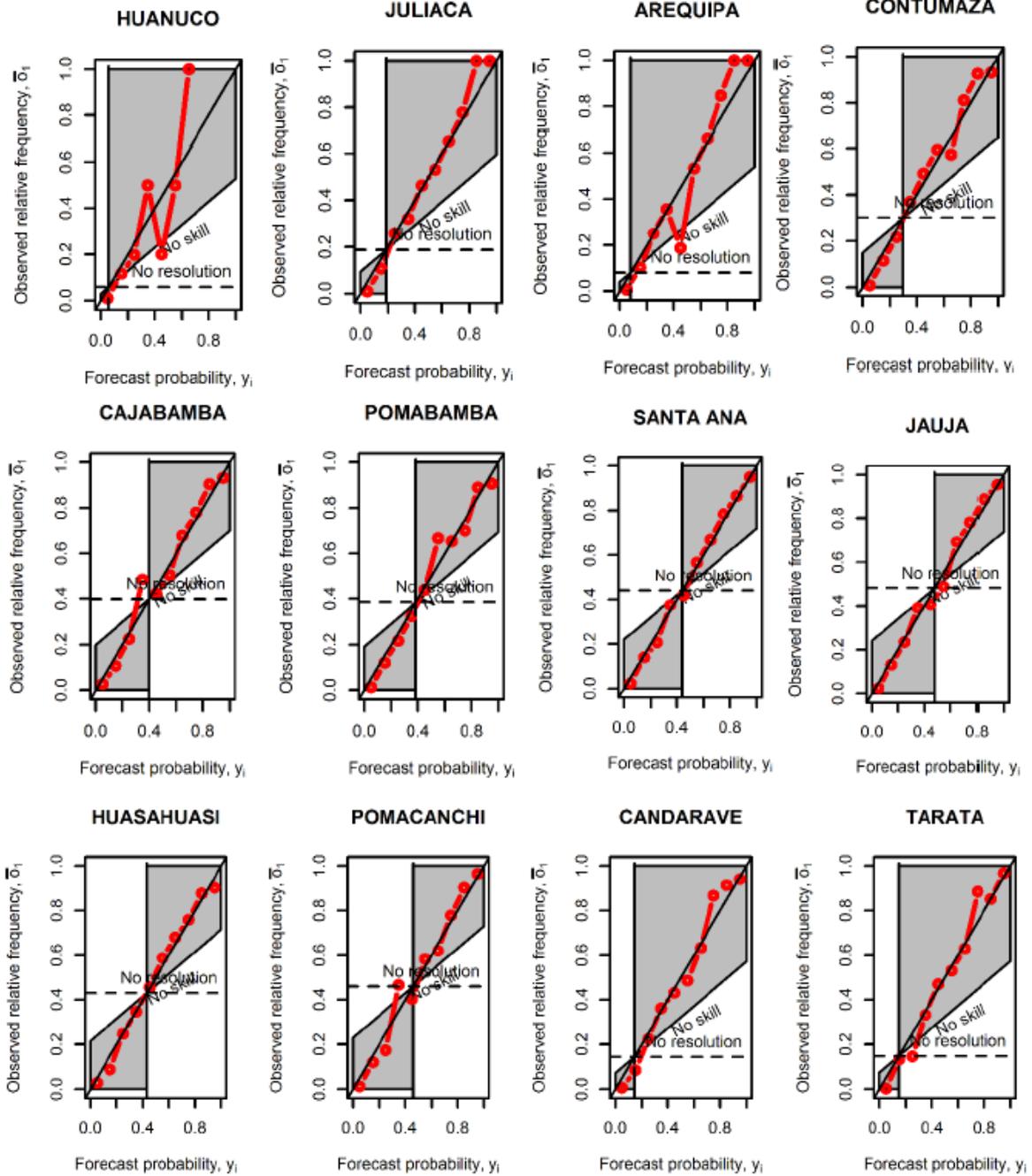
World Meteorological Organization (WMO). (1981). *Manual on the Global Observing System*. WMO-No. 544. World Meteorological Organization, Geneva. Consultado 03 de marzo del 2014. Disponible en: <http://www.bom.gov.au/wmo/lrfvs/reliability.shtml>.

VIII. ANEXOS

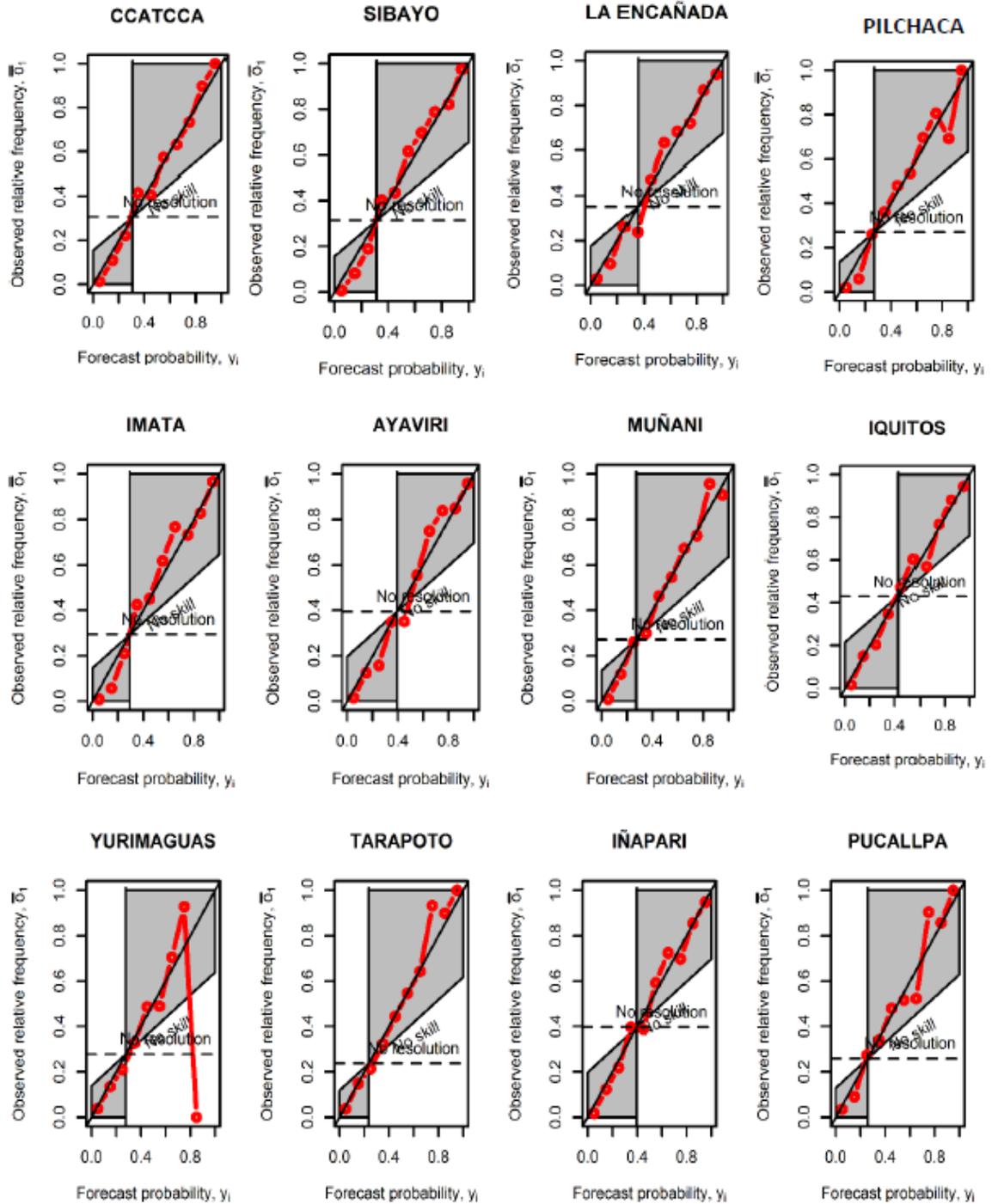
ANEXO 1: Diagramas de fiabilidad para cada estación para categoría 1



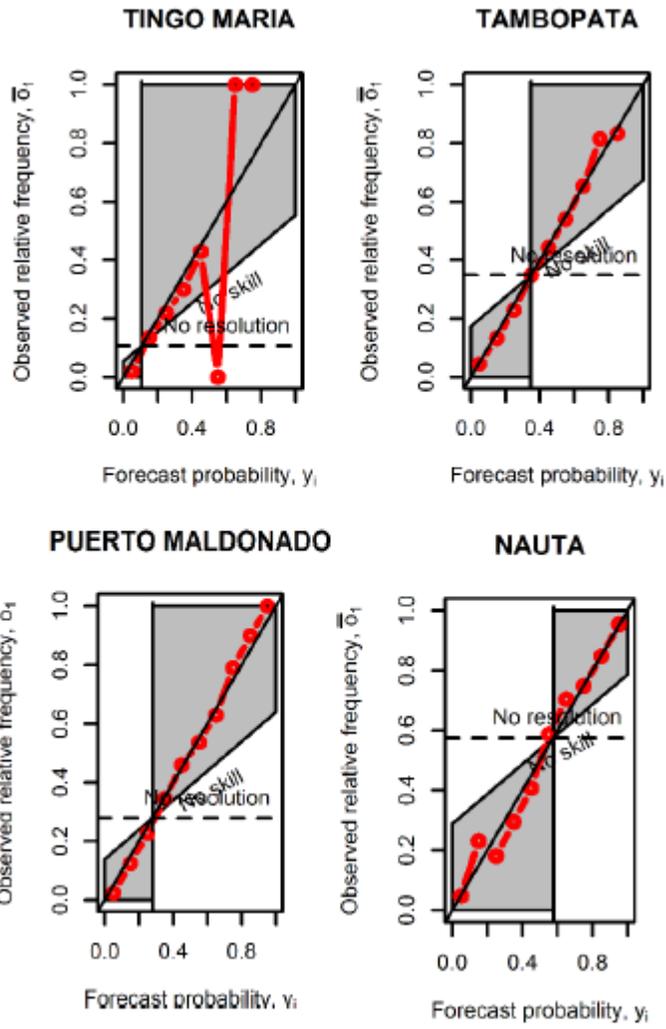
Continuación del Anexo 1



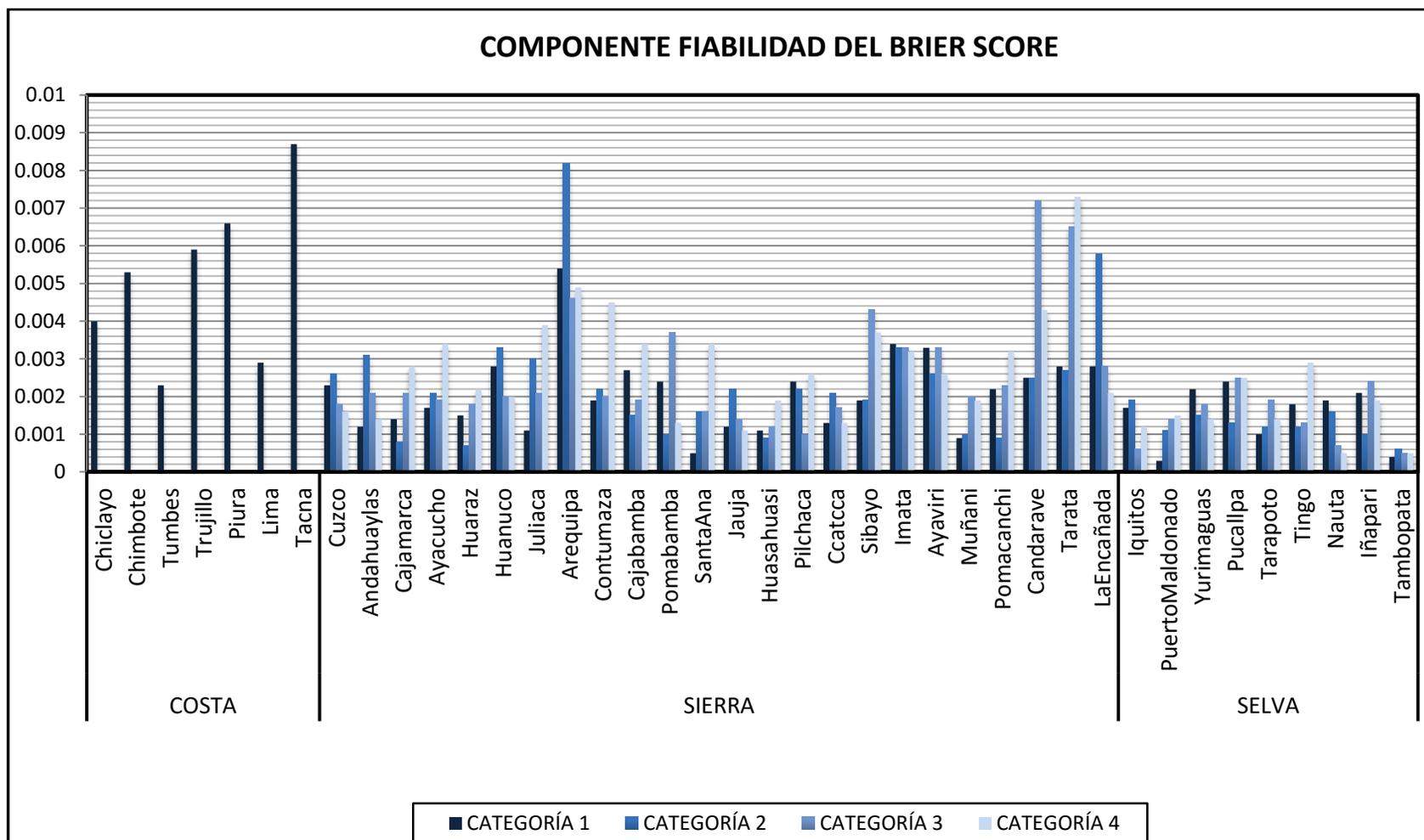
Continuación del Anexo 1



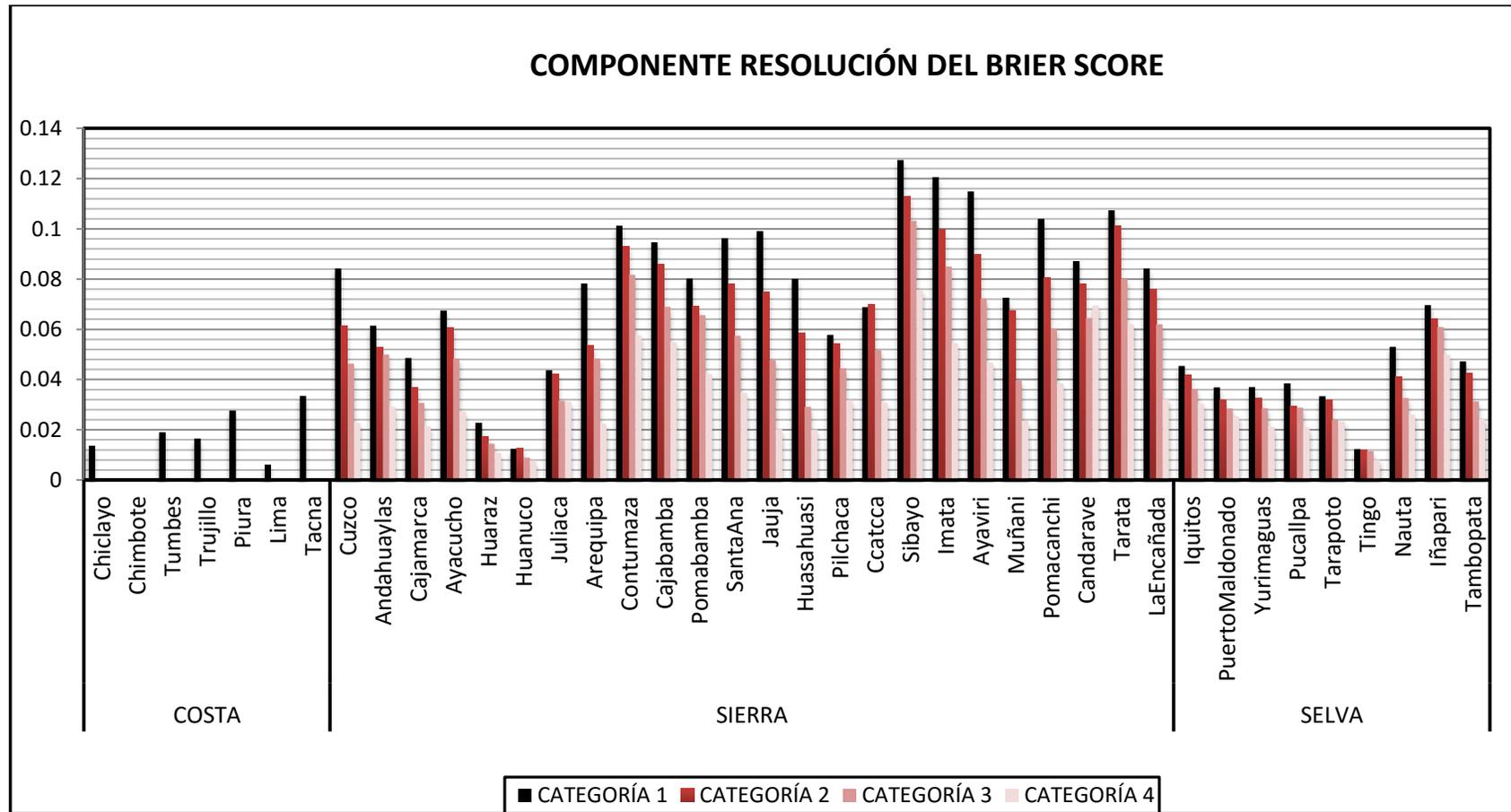
Continuación del Anexo 1



ANEXO2: Componente Fiabilidad del Brier Score dad del Brier Score



ANEXO 3: Componente Resolución del Brier Score



ANEXO 4: Resultados del Brier Score por estación

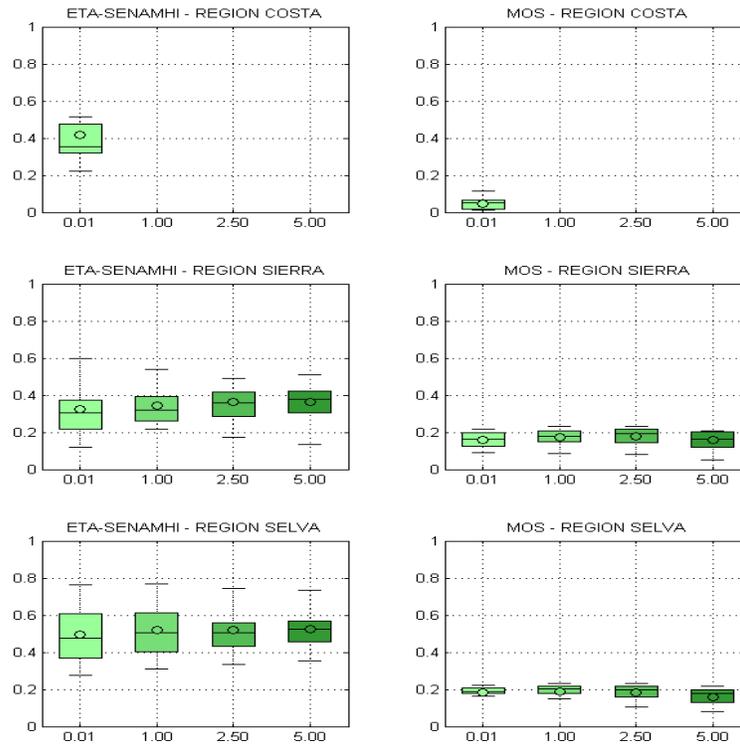
ESTACIÓN	PRONÓSTICO ETA -SENAMHI				PRONÓSTICO MOS				
	CATEGORÍA 1	CATEGORÍA 2	CATEGORÍA 3	CATEGORÍA 4	CATEGORÍA 1	CATEGORÍA 2	CATEGORÍA 3	CATEGORÍA 4	
COSTA	Chiclayo	0.54	-	-	-	0	-	-	-
	Chimbote	0.92	-	-	-	0	-	-	-
	Tumbes	0.55	-	-	-	0	-	-	-
	Trujillo	0.19	-	-	-	0	-	-	-
	Piura	0.15	-	-	-	0	-	-	-
	Lima	0.6	-	-	-	0	-	-	-
	Tacna	0.21	-	-	-	0	-	-	-
SIERRA	Cuzco	0.46	0.34	0.25	0.14	0.13	0.09	0.06	0.01
	Andahuaylas	0.57	0.39	0.28	0.21	0.09	0.08	0.05	0.01
	Cajamarca	0.49	0.36	0.3	0.23	0.07	0.04	0.02	0
	Ayacucho	0.53	0.3	0.21	0.15	0.12	0.1	0.06	0.02
	Huaraz	0.57	0.44	0.37	0.32	0.01	0.01	0.01	0
	Huánuco	0.81	0.63	0.51	0.44	0	0	0	0
	Juliaca	0.47	0.28	0.21	0.13	0.04	0.03	0.02	0.01
	Arequipa	0.14	0.07	0.05	0.02	0.01	0.01	0	0
	Contumaza	0.38	0.22	0.17	0.13	0.09	0.07	0.04	0.02
	Cajabamba	0.59	0.3	0.24	0.18	0.16	0.11	0.09	0.05
	Pomabamba	0.54	0.38	0.33	0.28	0.16	0.16	0.14	0.05
	Santa Ana	0.53	0.32	0.26	0.22	0.18	0.13	0.1	0.03
	Jauja	0.47	0.35	0.33	0.29	0.2	0.14	0.09	0.01
	Huasahuasi	0.6	0.4	0.39	0.35	0.2	0.15	0.04	0.01
	Pilchaca	0.62	0.36	0.26	0.2	0.1	0.09	0.07	0.03
	Ccatcca	0.54	0.31	0.21	0.14	0.13	0.11	0.11	0.03
	Sibayo	0.36	0.2	0.17	0.14	0.08	0.09	0.06	0.04
	Imata	0.3	0.16	0.13	0.1	0.08	0.08	0.06	0.04
	Ayaviri	0.38	0.27	0.21	0.21	0.12	0.11	0.08	0.04
	Muñani	0.45	0.29	0.25	0.2	0.11	0.11	0.05	0.02
Pomacanchi	0.4	0.27	0.22	0.15	0.17	0.13	0.07	0.03	
Candarave	0.3	0.13	0.1	0.06	0.04	0.03	0.02	0.01	
Tarata	0.33	0.11	0.08	0.05	0.03	0.02	0.02	0.01	
La Encañada	0.44	0.29	0.24	0.2	0.13	0.1	0.06	0.03	
SELVA	Iquitos	0.92	0.87	0.82	0.7	0.21	0.11	0.07	0.03
	Puerto Maldonado	0.61	0.55	0.5	0.45	0.08	0.07	0.05	0.03
	Yurimaguas	0.83	0.75	0.67	0.58	0.07	0.06	0.03	0.02
	Pucallpa	0.72	0.66	0.6	0.51	0.07	0.04	0.02	0.01
	Tarapoto	0.86	0.73	0.6	0.41	0.04	0.02	0.01	0.01
	Tingo María	0.88	0.79	0.7	0.58	0	0	0	0
	Nauta	0.86	0.82	0.77	0.66	0.39	0.26	0.15	0.06
	Iñapari	0.65	0.58	0.5	0.39	0.19	0.17	0.11	0.06
Tambopata	0.71	0.55	0.47	0.38	0.14	0.12	0.08	0.03	

ANEXO 5: Resultados del Brier Skill Score por estación

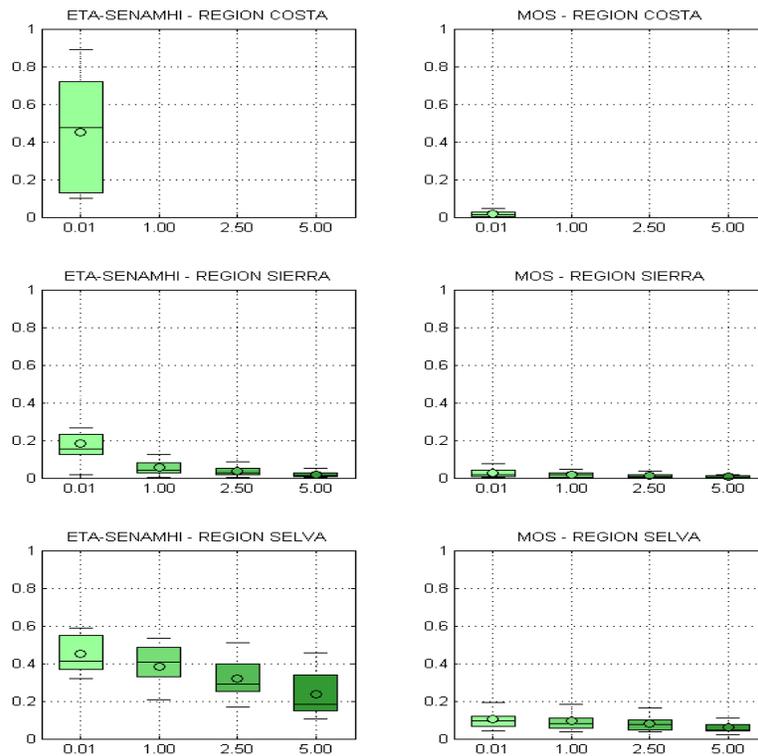
	ESTACIÓN	PRONÓSTICO ETA -SENAMHI				PRONÓSTICO MOS			
		CATEGORÍA 1	CATEGORÍA 2	CATEGORÍA 3	CATEGORÍA 4	CATEGORÍA 1	CATEGORÍA 2	CATEGORÍA 3	CATEGORÍA 4
COSTA	Chiclayo	0.12	-	-	-	-12.76	-	-	-
	Chimbote	-0.75	-	-	-	-999.00	-	-	-
	Tumbes	0.18	-	-	-	-5.36	-	-	-
	Trujillo	0.11	-	-	-	-4.00	-	-	-
	Piura	0.19	-	-	-	-1.94	-	-	-
	Lima	0.03	-	-	-	-19.91	-	-	-
	Tacna	0.19	-	-	-	-6.56	-	-	-
SIERRA	Cuzco	0.44	0.36	0.31	0.20	-0.29	0.03	-0.54	-0.61
	Andahuaylas	0.35	0.31	0.31	0.24	-1.06	-0.51	-0.62	-0.99
	Cajamarca	0.29	0.25	0.22	0.18	-1.03	-0.59	-0.87	-1.42
	Ayacucho	0.36	0.34	0.29	0.20	-0.72	-0.23	-0.35	-0.56
	Huaraz	0.20	0.18	0.15	0.12	-3.19	-2.30	-2.67	-3.86
	Huánuco	0.12	0.12	0.09	0.07	-3.19	-2.30	-2.67	-3.86
	Juliaca	0.28	0.28	0.24	0.25	-1.41	-0.69	-0.92	-1.06
	Arequipa	0.45	0.35	0.34	0.17	-0.87	-0.14	-0.54	-0.52
	Contumaza	0.51	0.50	0.47	0.38	-0.27	0.10	-0.11	-0.31
	Cajabamba	0.46	0.44	0.38	0.35	-0.41	0.13	-0.07	-0.24
	Pomabamba	0.40	0.36	0.34	0.26	-0.32	-0.02	-0.16	-0.42
	Santa Ana	0.44	0.37	0.35	0.24	-0.17	0.15	-0.17	-0.60
	Jauja	0.46	0.40	0.30	0.18	0.00	0.23	-0.51	-1.11
	Huasahuasi	0.37	0.32	0.21	0.18	-0.33	0.08	-0.92	-1.78
	Pilchaca	0.31	0.30	0.28	0.23	-1.14	-0.38	-0.31	-0.56
	Ccatcca	0.39	0.39	0.33	0.25	-0.63	-0.12	-0.20	-0.45
	Sibayo	0.64	0.59	0.54	0.45	-0.13	0.22	0.00	-0.15
	Imata	0.63	0.56	0.50	0.37	0.01	0.35	0.01	-0.08
	Ayaviri	0.59	0.49	0.42	0.32	0.07	0.30	-0.09	-0.54
	Muñani	0.40	0.38	0.28	0.21	-0.57	-0.09	-0.47	-0.85
Pomacanchi	0.56	0.43	0.36	0.27	0.16	0.31	-0.14	-0.33	
Candarave	0.51	0.47	0.40	0.44	-0.96	-0.08	-0.39	-0.28	
Tarata	0.58	0.56	0.48	0.40	-1.12	0.16	-0.17	-0.27	
La Encañada	0.40	0.37	0.34	0.23	-0.24	0.03	-0.21	-0.46	
SELVA	Iquitos	0.19	0.19	0.18	0.17	-1.00	-0.94	-1.51	-1.89
	Puerto Maldonado	0.22	0.20	0.19	0.18	-1.13	-0.99	-1.30	-1.72
	Yurimaguas	0.18	0.17	0.16	0.14	-1.78	-1.59	-1.74	-2.05
	Pucallpa	0.21	0.19	0.19	0.16	-1.69	-1.56	-2.10	-2.74
	Tarapoto	0.20	0.20	0.16	0.18	-2.35	-1.96	-2.36	-2.45
	Tingo María	0.11	0.12	0.12	0.08	-6.44	-5.73	-5.89	-6.32
	Nauta	0.22	0.17	0.14	0.13	-0.42	-0.39	-0.90	-1.18
	Iñapari	0.34	0.34	0.33	0.29	-0.56	-0.43	-0.62	-0.73
Tambopata	0.22	0.21	0.17	0.17	-0.93	-0.64	-0.83	-1.07	

ANEXO 6: Resultados del Brier Score a nivel estacional.

BRIER SCORE - PERIODO DEF

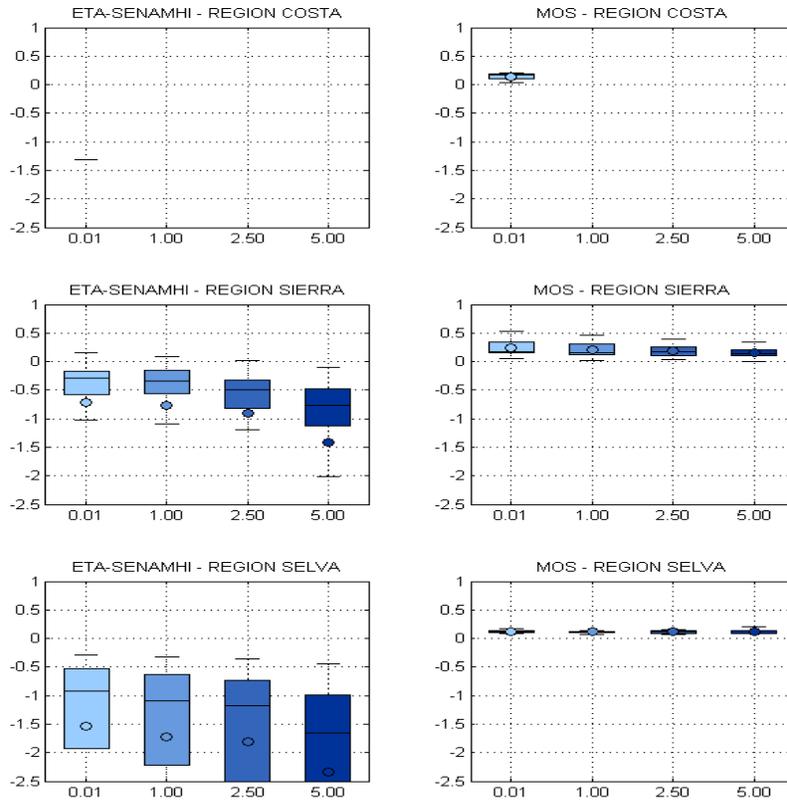


BRIER SCORE - PERIODO JJA

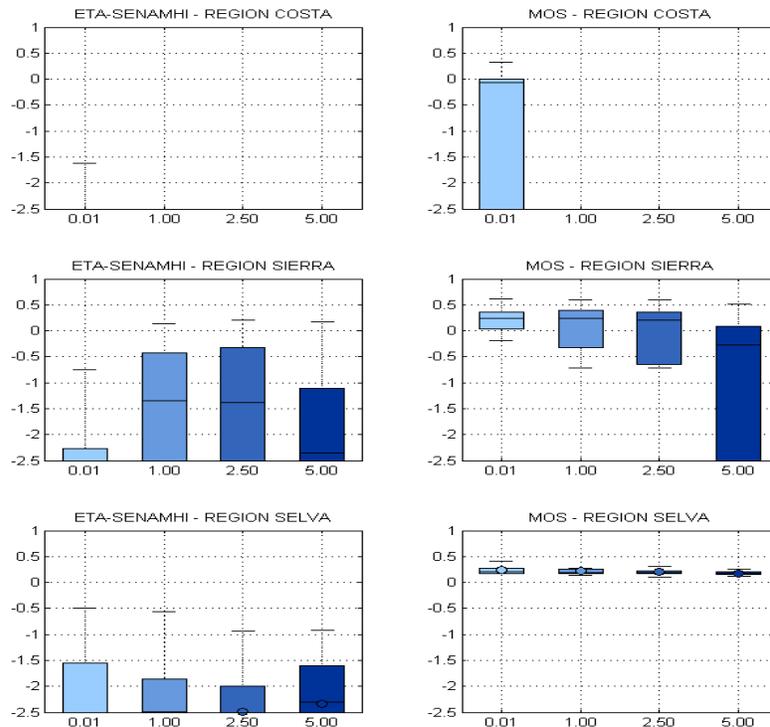


ANEXO 7: Resultados del Brier Skill Score a nivel estacional.

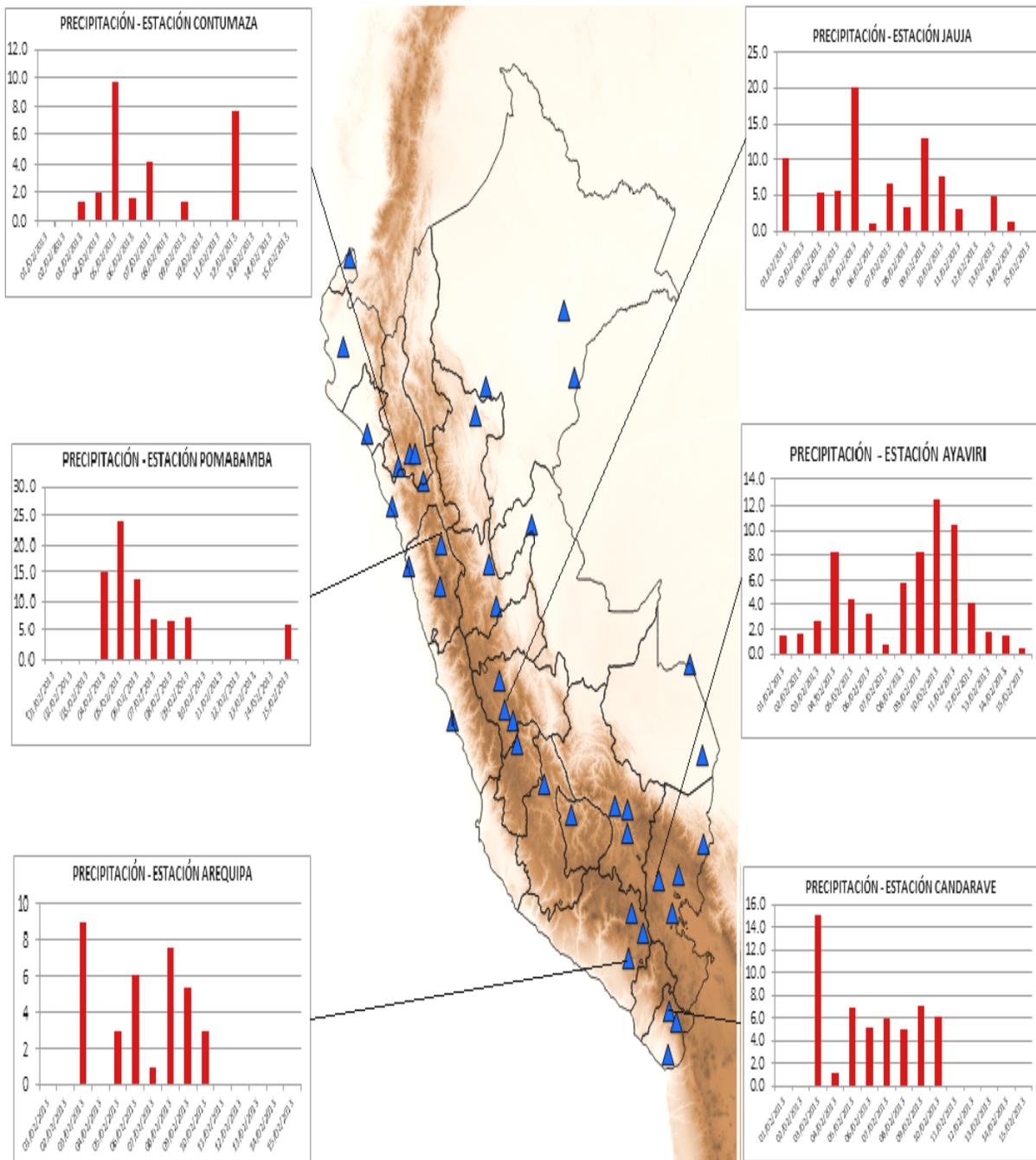
BRIER SKILL SCORE - PERIODO DEF



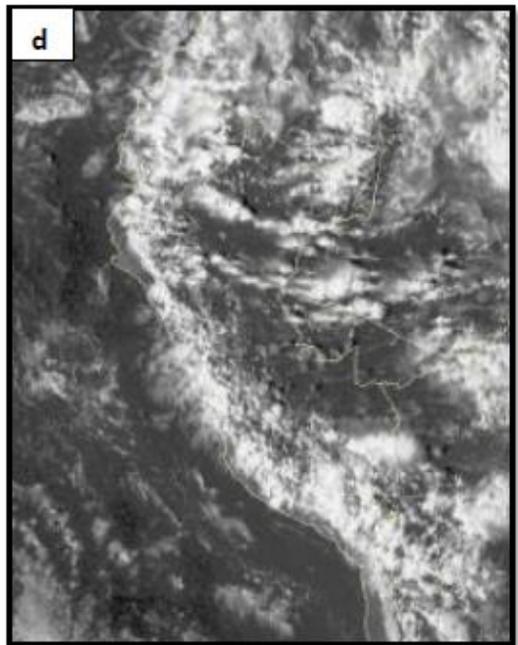
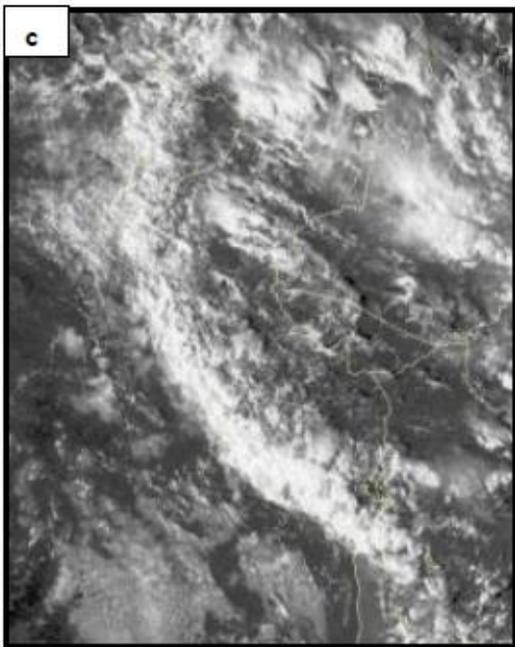
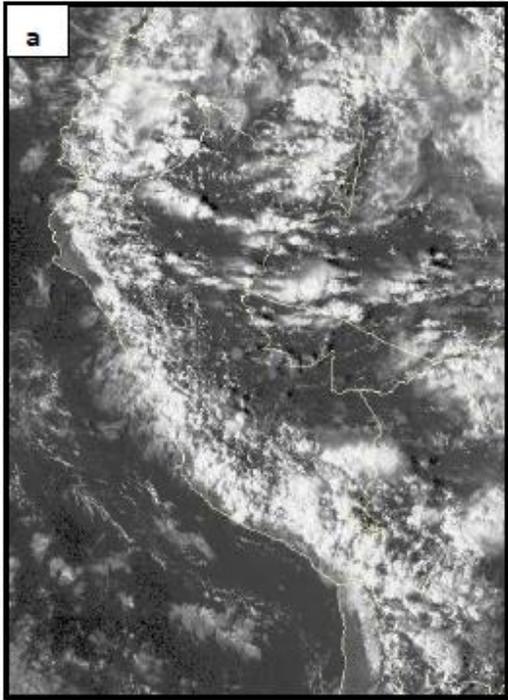
BRIER SKILL SCORE - PERIODO JJA



ANEXO 8: Precipitación diaria en 6 estaciones para el periodo 01-15 de Febrero del 2013

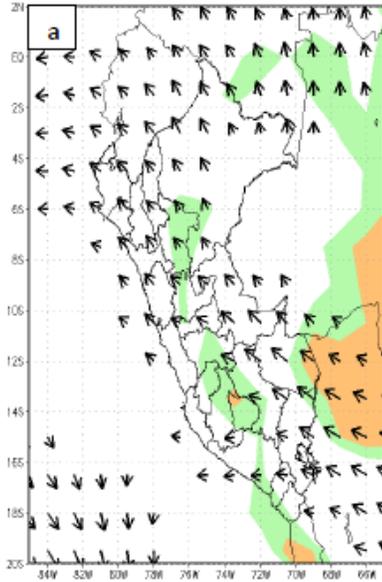


ANEXO 9: Imágenes satelitales del GOES 13 correspondiente al canal visible para los días (a) 02 de febrero (b) 04 de febrero (c) 05 de febrero y (d) 06 de febrero del 2013

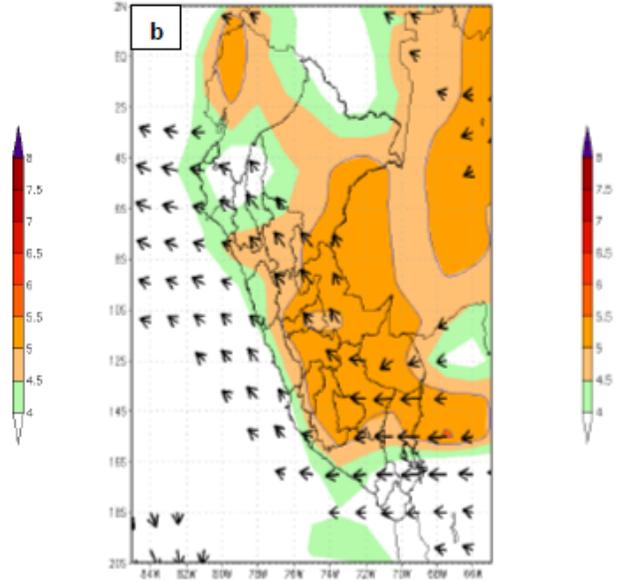


ANEXO 10: Relación de mezcla y vientos en 500 hPa para los días (a) 03 de febrero (b) 04 de febrero (c) 05 de febrero y (d) 07 de febrero del 2013

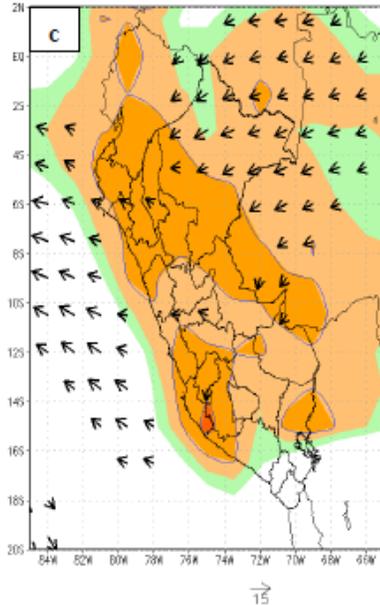
RELACION DE MEZCLA (g/Kg) VIENTOS (m/s) 500HPa
03/FEB/2013 12:00Z



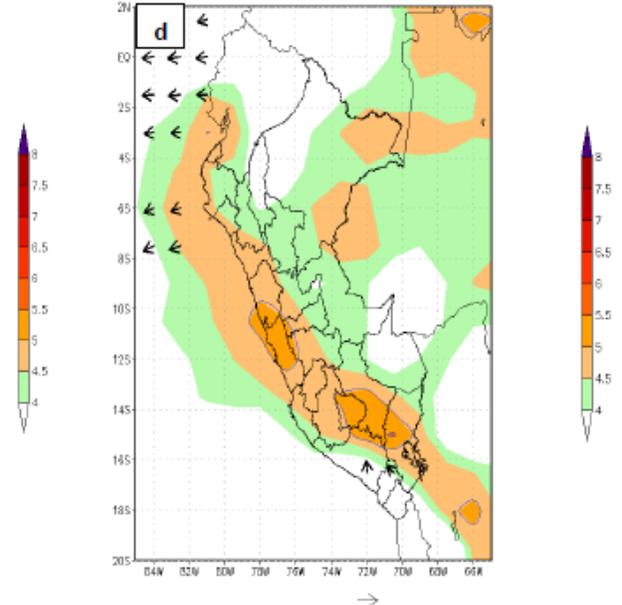
RELACION DE MEZCLA (g/Kg) VIENTOS (m/s) 500HPa
04/FEB/2013 12:00Z



RELACION DE MEZCLA (g/Kg) VIENTOS (m/s) 500HPa
05/FEB/2013 12:00Z



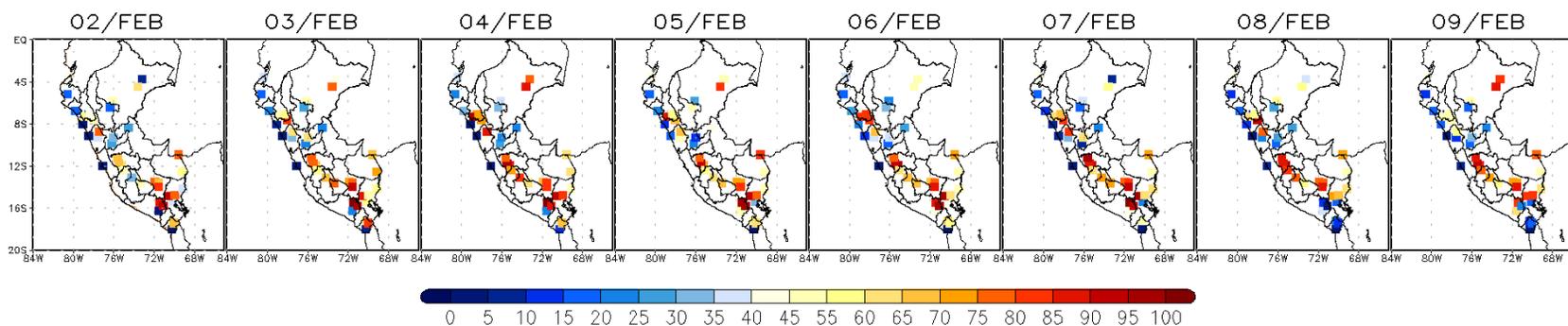
RELACION DE MEZCLA (g/Kg) VIENTOS (m/s) 500HPa
07/FEB/2013 12:00Z



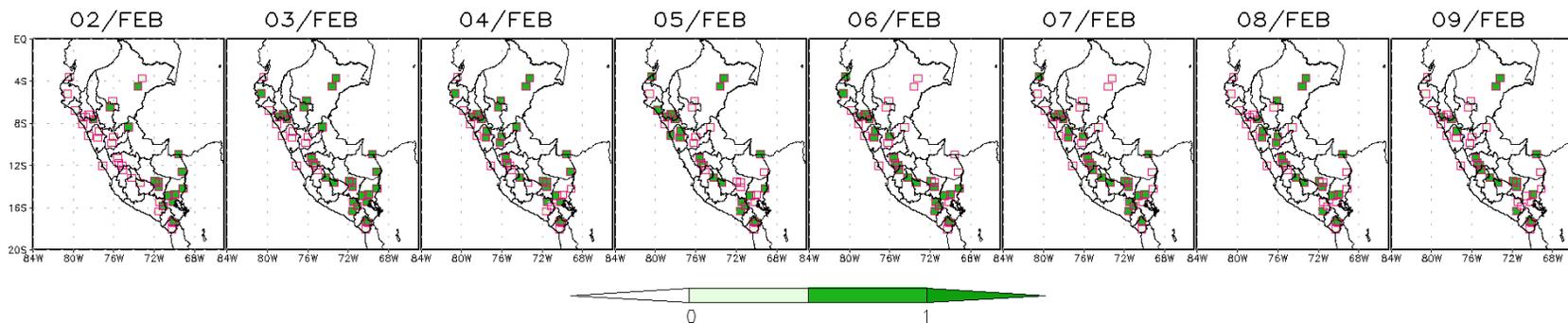
ANEXO 11: Pronóstico de probabilidades de precipitación a 24 horas y ocurrencia observada de precipitación para el periodo 02-09 de Febrero del 2013.

1. Pronósticos probabilísticos y datos observados para categoría 1

PROBABILIDAD DE PRECIPITACION(%)
PP >= 0.01mm

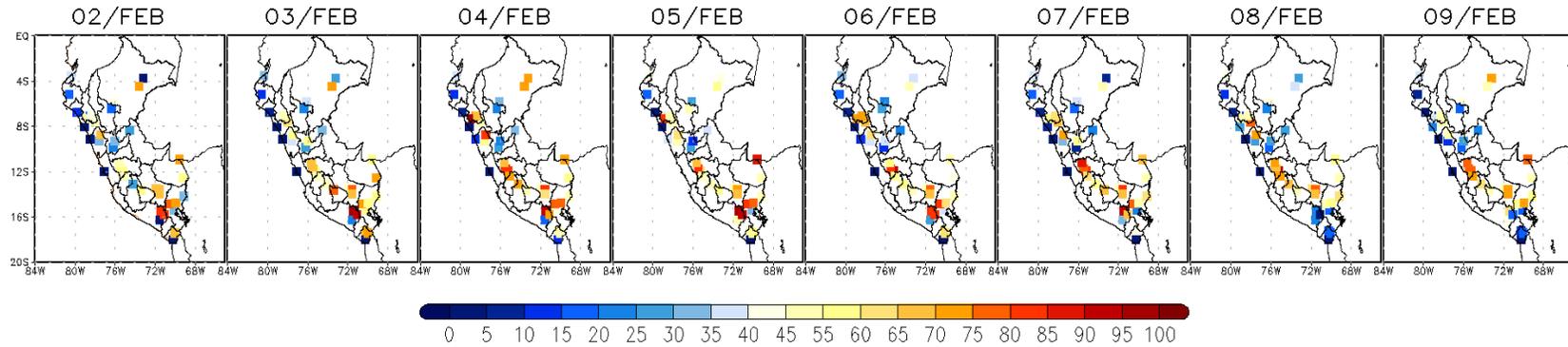


OCURRENCIA DE PRECIPITACION
VERDE: PP >= 0.01mm, BLANCO: PP < 0.01mm

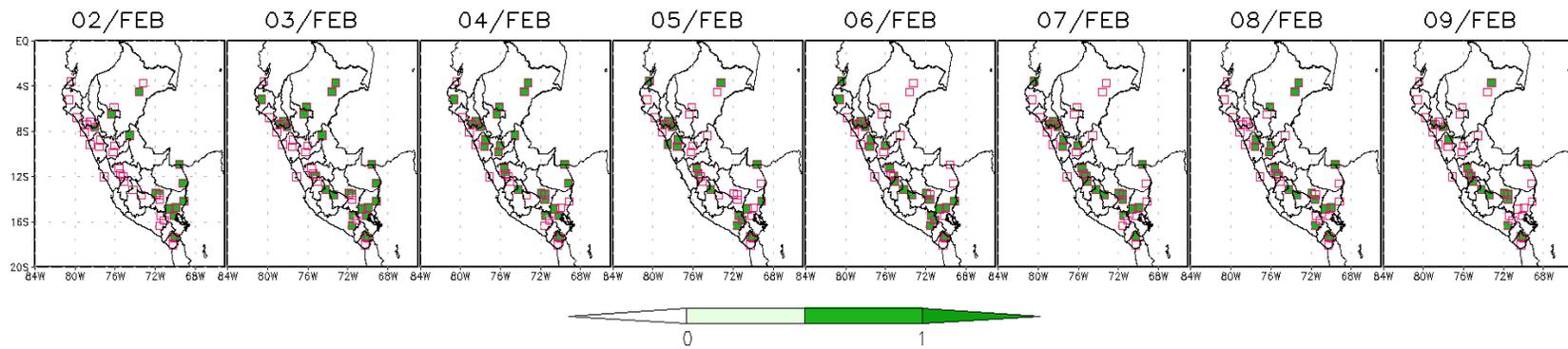


2. Pronósticos probabilísticos y datos observados para categoría 2

PROBABILIDAD DE PRECIPITACION(%)
PP >= 1.00mm

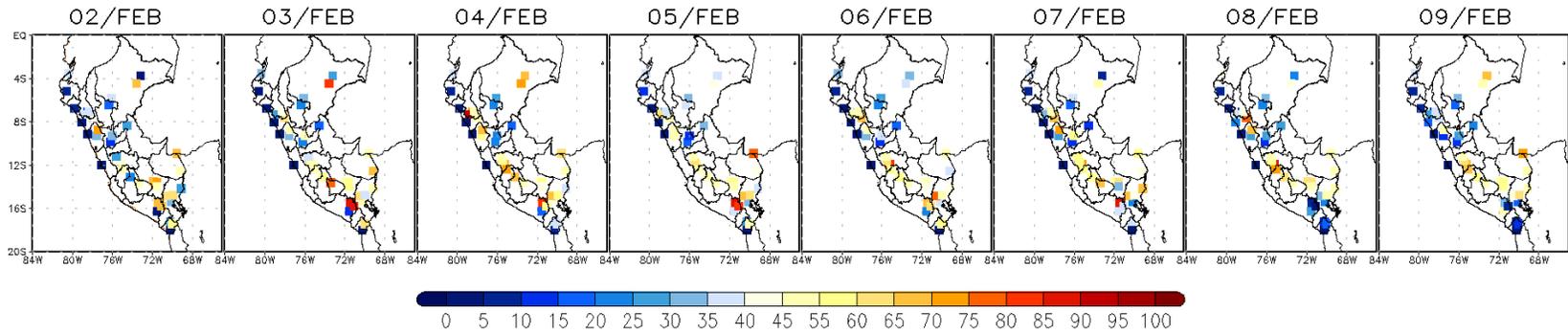


OCURRENCIA DE PRECIPITACION
VERDE: PP >= 1.00mm, BLANCO: PP < 1.00mm

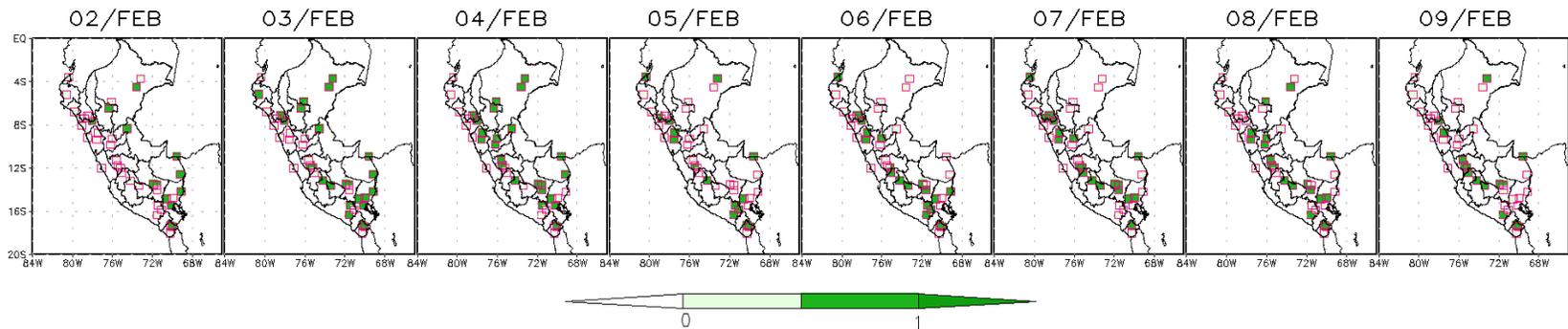


3. Pronósticos probabilísticos y datos observados para la categoría 3

PROBABILIDAD DE PRECIPITACION(%)
PP >= 2.50mm

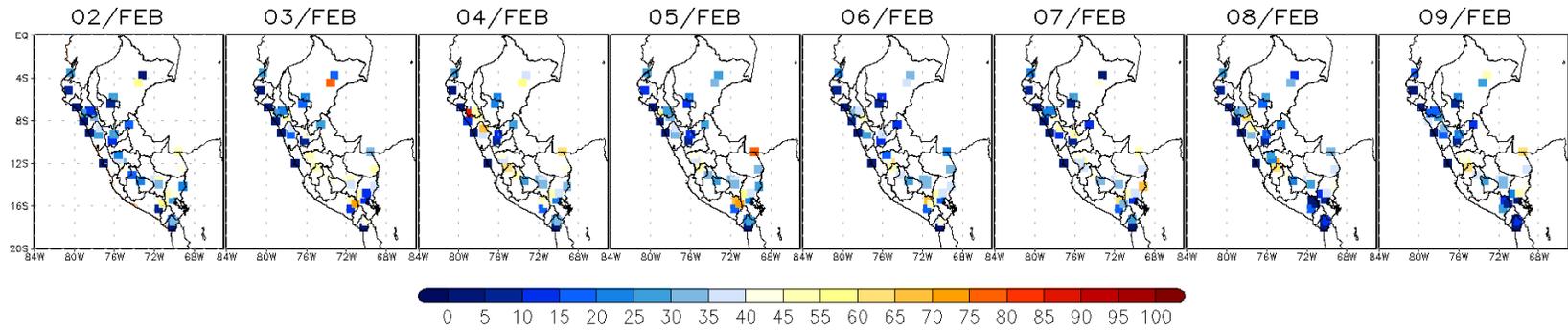


OCURRENCIA DE PRECIPITACION
VERDE: PP >= 2.50mm, BLANCO: PP < 2.50mm



4. Pronósticos probabilísticos y datos observados para la categoría 4

PROBABILIDAD DE PRECIPITACION(%)
PP >= 5.00mm



OCURENCIA DE PRECIPITACION
VERDE: PP >= 5.00mm, BLANCO: PP < 5.00mm

