

AVANCE DE INVESTIGACIÓN

MODELAMIENTO ESTADÍSTICO PARA EL PRONÓSTICO MENSUAL DE PRECIPITACIONES EN PERÚ

Aldo S. Moya-Álvarez¹, Gerardo Rivera¹ y Kobi Mosquera¹

RESUMEN

Se construye un modelo estadístico de pronóstico mensual de precipitaciones para Perú, basado en la relación de varios índices climáticos oceánicos y atmosféricos, así como de los índices de la oscilación de Madden-Julian, con la variabilidad estacional de la precipitación en el país, a inicializar en el mes de junio de cada año. Se aplicó la técnica de regresión lineal entre los índices y la precipitación grillada de la fuente de datos de PISCO para el período diciembre-marzo, meses que corresponden a la estación lluviosa en Perú. El modelo se construyó con datos correspondientes al período 1981-2016 y se verificó con los registros de precipitación correspondientes a los años 2017-2022 para 15 zonas dentro de la costa, sierra y selva. El valor del error medio absoluto (MAE, por sus siglas en inglés) fue mejor para la selva, con excepción de los pronósticos para enero en la selva norte. El mejor rendimiento se obtuvo para marzo en la selva sur y central. En cuanto a la sierra, el mejor rendimiento del modelo se obtuvo para la sierra sur y central, en ambos casos para la vertiente Atlántica. Para la selva, el mejor rendimiento se obtuvo en la parte alta. En 8 de las 12 regiones correspondientes a la sierra y la selva, el modelo realizó “pronósticos correctos” en más del 60 % de los casos, y en uno, selva sur baja, muy cercano al 60 %.

Palabras clave: pronóstico de lluvias, modelamiento estadístico, regresión lineal

AFILIACIÓN:

1. Instituto Geofísico del Perú, Lima, Perú.

Citar como: Moya, A., Rivera G. y Mosquera K. (2022). Modelamiento estadístico para el pronóstico mensual de precipitaciones en Perú. *Boletín científico El Niño*, Instituto Geofísico del Perú, Vol. 9 n.º 06. págs. 14-20.

INTRODUCCIÓN

Las precipitaciones juegan un papel fundamental en la sociedad y la economía de cada país dada su utilidad como recurso indispensable, tanto para el consumo humano como para su uso en la industria y, fundamentalmente, en el desarrollo agrícola de cualquier nación. Sin embargo, éstas también pueden constituir un peligro cuando exceden su comportamiento normal, que producen eventos lluviosos que frecuentemente dan lugar a inundaciones y al desborde de los ríos, que por lo general provocan la pérdida de cuantiosos recursos económicos y, en ocasiones, hasta de vidas humanas. A veces, estos eventos están relacionados con la fase cálida o fría de El Niño-Oscilación del Sur (Takahashi, 2004; Lagos et al., 2008; Lavado-Casimiro y Espinoza, 2011; Sulca et al., 2017). Por ello, se requiere generar herramientas que apoyen el monitoreo y pronóstico de eventos de lluvias. La mayoría de las lluvias en Perú se concentran en el período comprendido entre septiembre y abril (Silva et al, 2008), definiendo una marcada estacionalidad, con una estación seca entre mayo y agosto (Aceituno, 1989; Vuille et al, 2008; Garreaud, 2009).

En el presente trabajo se construye un modelo estadístico de pronóstico de precipitaciones mensuales para todo el Perú, basado en la relación entre la variabilidad estacional de la lluvia en el Perú y varios índices climáticos oceánicos y atmosféricos, así como de los índices RMM (Real-Time Multivariate MJO). El modelo se construyó aplicando la técnica de regresión lineal entre los mencionados índices y una base de datos grillada de precipitación para el período diciembre-marzo (período de lluvias). El modelo se construyó con datos correspondientes al período 1981-2016, y se inicializó en junio de cada año.

DATOS Y METODOLOGÍA

Los datos de precipitación (“predictando”) fueron tomados para diciembre, enero, febrero y marzo, de la base de datos interpolados de las observaciones climatológicas e hidrológicas del SENAMHI (PISCO; Lavado et al., 2018) del período 1981-2016. Los datos de PISCO constituyen una grilla de 26 334 puntos que abarcan todo el Perú, de los cuales 19 617 se ubican sobre tierra. Los predictores (ver Tabla 1) se tomaron también para el período 1981-2016. Se tomó el comportamiento de los predictores durante los 12 meses previos al mes de inicialización del modelo, en este caso, junio.

Foto: Jorge Concha



Tabla 1. Resumen de las teorías de interacción entre las ondas Kelvin y TIWs.

ONI	Índice Oceánico Niño
ICEN	Índice costero El Niño
NAO	Oscilación del Atlántico Norte
AO	Oscilación del océano Ártico
EP/NP	Oscilación del Pacífico Este / Pacífico Norte
WP	Índice del Pacífico Occidental
SOI	Índice de Oscilación del Sur
TSA	Índice del Atlántico Sur Tropical
TNA	Índice del Atlántico Norte Tropical
C y E	Nuevos índices no correlacionados, basados en los EOF principales de la anomalía de la TSM en el océano Pacífico, que dan cuenta respectivamente de los eventos cálidos extremos en el este y los eventos cálidos fríos / moderados en el Pacífico ecuatorial central, correspondientes a regímenes con diferente evolución.
RMM1 y RMM2	Índices Multivariados en tiempo real de la oscilación de Madden-Julian

MÉTODO

Se aplicó la técnica de regresión lineal múltiple en su modalidad de Perfect-Prog (Glahn y Lowry, 1972). Sobre esta base, se construyó una ecuación de regresión múltiple para cada punto de la grilla de PISCO, de manera que el pronóstico de precipitaciones comprende todo el país. Visto así, se construirían 19 617 ecuaciones de pronóstico (modelos).

VERIFICACIÓN

La verificación de los resultados se realizó para 15 regiones, promediando para cada una los resultados de los pixeles correspondientes a cada región, tanto del pronóstico como de la precipitación registrada:

- Coste norte, costa central y costa sur
- Sierra norte occidental y oriental
- Sierra central occidental y oriental
- Sierra sur occidental y oriental
- Selva norte alta y baja
- Selva central alta y baja
- Selva sur alta y baja

Los estadísticos empleados fueron el sesgo (bias) y el error absoluto medio (MAE), los que se calcularon con los resultados obtenidos para la muestra independiente 2017-2022. También se calculó el porcentaje de “pronósticos correctos”, entendiéndose el porcentaje de ocasiones en que el modelo pronosticó por encima/debajo de la media histórica y se cumplió. También se incluyeron como pronósticos correctos a aquellos casos en que el módulo de las anomalías, tanto pronosticada como observada, no rebasó el 10 %.

RESULTADOS

La Figura 1 muestra la desviación absoluta media del pronóstico respecto a la precipitación registrada para diciembre, enero, febrero y marzo para las regiones consideradas. De forma general, se puede apreciar que el modelo muestra el mejor rendimiento para la selva, donde el valor del MAE estuvo por debajo o cercano al 25 %, con excepción de los pronósticos para enero en la selva norte. En este sentido, hay que resaltar que los mejores resultados se obtuvieron para marzo, tanto en la selva sur como en la selva central.

En relación con la sierra, los mejores rendimientos se obtuvieron para la sierra sur y central, en ambos casos en la vertiente oriental para diciembre, enero y marzo. Los pronósticos para la vertiente del Pacífico de la sierra mostraron bajo rendimiento.

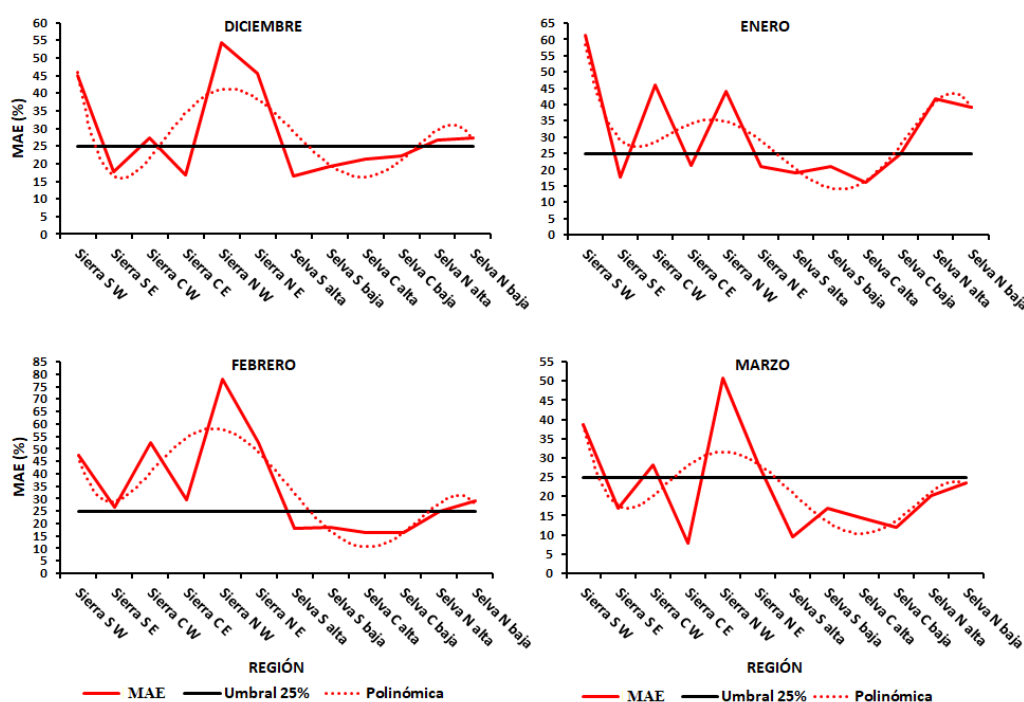


Figura 1. Desviación absoluta media del pronóstico respecto a la precipitación registrada para diciembre, enero, febrero y marzo para todas las regiones consideradas. La curva polinómica es la tendencia de quinto orden aplicada al rendimiento del modelo a través de las regiones mostradas en la gráfica.

En cuanto a la efectividad general del modelo, la Figura 2 muestra el porcentaje general de los pronósticos correctos emitidos para todas las regiones consideradas. En 8 de las 12 regiones correspondientes a sierra y selva el modelo realizó “pronósticos correctos” en más del 60 % de los casos, y en uno, selva sur baja, muy cercano al 60 %. En este sentido, el modelo mostró un rendimiento bajo solo en la sierra norte y en la selva norte baja, con efectividad por debajo del 45 %. Para la selva central baja mostró una efectividad cercana al 80 %.

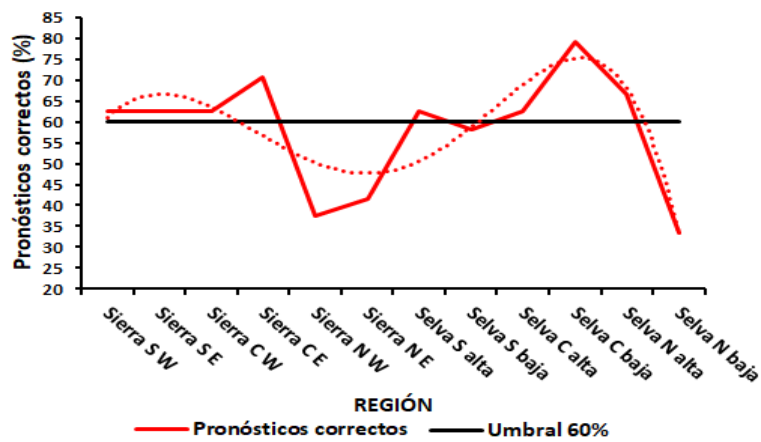


Figura 2. Porcentaje general de pronósticos correctos emitidos para todas las regiones consideradas. La curva punteada es del tipo polinómica.

La Figura 3 resume el comportamiento de la desviación absoluta media del pronóstico respecto a la precipitación registrada para las vertientes del Pacífico y Atlántico de la sierra y para la selva alta y baja. En la figura se observa que el modelo mostró el mejor rendimiento para la vertiente Atlántica de la sierra y para la selva alta. El rendimiento del modelo para la zona costera (no se muestra en el artículo) resultó muy bajo en cuanto al porcentaje de error. De igual forma, el modelo no detectó los eventos lluviosos asociados a los eventos de El Niño 1982-1983 y 1997-1998. Los predictores más informativos fueron los índices WP, NAO, E, AO, TSA y la MJO.

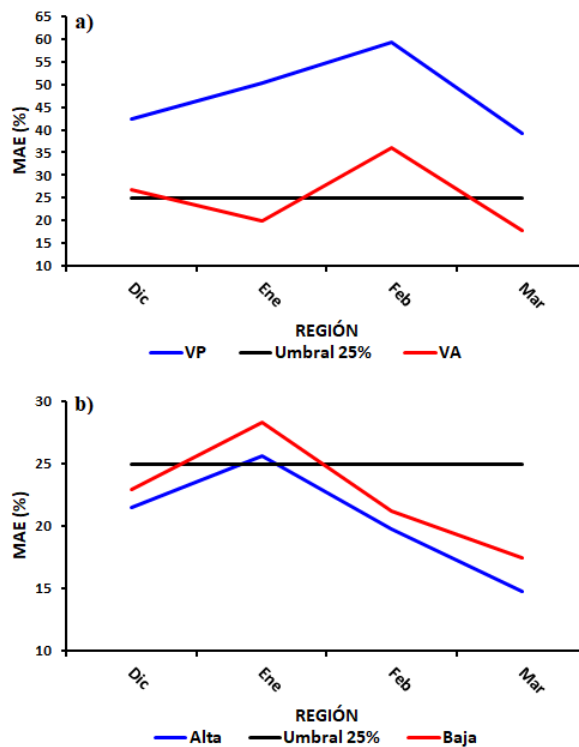


Figura 3. Desviación absoluta media del pronóstico respecto a la precipitación registrada para las vertientes del Pacífico y Atlántico de la sierra (a) y para la selva alta y baja (b).

CONCLUSIONES

Se construyó un modelo estadístico de pronóstico estacional de precipitaciones, basado en la técnica de regresión lineal para 15 zonas de Perú, que cubren costa, sierra y selva. Los resultados fueron los siguientes:

- El modelo muestra el mejor rendimiento para la región de la selva, donde el MAE se comportó por debajo o cerca al 25 %, con excepción de los pronósticos para enero en la selva Norte.
- El mejor rendimiento del modelo se obtuvo para el mes de marzo en la selva sur y selva central.
- Para la sierra, el mejor rendimiento del modelo se obtuvo para la sierra sur y central, en ambos casos para la vertiente Atlántica.
- En 8 de las 12 regiones correspondientes a sierra y selva, el modelo realizó “pronósticos correctos” en más del 60 % de los casos, y en uno, selva sur baja, muy cercano al 60 %.
- Para la sierra el modelo mostró mejor rendimiento en su vertiente atlántica, mientras para la selva, el mejor rendimiento se obtuvo en la parte alta.
- Los predictores más informativos fueron los índices WP, NAO, E, AO, TSA y la MJO.

REFERENCIAS

Aceituno, P. (1989). On the functioning of the southern oscillation in the South American, sector. Part II. Upper-air circulation, *J. Clim.*, 2(4), 341-355, [https://doi.org/10.1175/1520-0493\(1988\)116<0505:OTFOTS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1988)116<0505:OTFOTS>2.0.CO;2)

Garreaud, R. (2009). The Andes climate and weather. *Adv. in Geosc.*, 22, 3-11. <https://doi.org/10.5194/adgeo-22-3-2009>

Glahn, H.R. y Lowry, D.A. (1972). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *J. Appl. Meteor. and Clim.*, 11 (8), 1203-1211. [https://doi.org/10.1175/1520-0450\(1972\)011<1203:tuomos>2.0.co;2](https://doi.org/10.1175/1520-0450(1972)011<1203:tuomos>2.0.co;2)

Lagos, P., Silva, Y., Nickl, E., y Mosquera, K. (2008). El Niño–related precipitation variability in Perú. *Adv. Geosci.*, 14, 231-237. <https://doi.org/10.5194/adgeo-14-231-2008>

Lavado-Casimiro, W. y Espinoza, J. C. (2014). Impactos de El Niño y La Niña en las Lluvias del Perú (1965-2007). *Rev. bras. meteorol.*, 29(2), 171-182. <https://doi.org/10.1590/S0102-77862014000200003>

Lavado, W., Fernandez, C., Vega, F., Caycho, T., Endara, S., Huerta, A. y Obando, O. F. (2016). PISCO: Peruvian interpolated data of the SENAMHI's climatological and hydrological observations. *Precipitación v1.0*. Servicio Nacional de Meteorología e Hidrología, pp. 5.

Silva, Y., Takahashi, K. y Chávez, R. (2008). Dry and wet rainy seasons in the Mantaro river basin (Central Peruvian Andes). *Adv. Geosci.*, 14, 261-264. <https://doi.org/10.5194/adgeo-14-261-2008>

Sulca, J., Takahashi, K., Espinoza, J.C., Vuille, M. y Lavado-Casimiro, W. (2017). Impacts of different ENSO flavors and tropical Pacific convection variability (ITCZ, SPCZ) on austral summer rainfall in South America, with a focus on Peru. *Int. J. Climatol.*, 38(1), 420-435. <https://doi.org/10.1002/joc.5185>

Takahashi, K. (2004). The atmospheric circulation associated with extreme rainfall events in Piura, Peru, during the 1997—1998 and 2002 El Niño events. *Ann. Geophys.*, 22, 3917–3926. <https://doi.org/10.5194/angeo-22-3917-2004>

Vuille, M., Kaser, G. y Juen, I. (2008). Glacier mass balance variability in the Cordillera Blanca, Peru and its relationship with climate and the large-scale circulation. *Glob. Planet. Chang.*, 62(1-2), 14-28. <https://doi.org/10.1016/j.gloplacha.2007.11.003>