

METODOLOGÍA DE VERIFICACIÓN DEL PRONÓSTICO ESTACIONAL POR CONSENSO DEL SERVICIO METEOROLÓGICO NACIONAL

María Mercedes Poggi¹, María de los Milagros Skansi¹, Ramón de Elía¹
mpoggi@smn.gob.ar

¹Servicio Meteorológico Nacional (SMN)

Palabras clave: Evaluación, verificación, predicciones climáticas, métricas.

1) INTRODUCCIÓN

En los últimos años el Servicio Meteorológico Nacional de Argentina (SMN) ha puesto una creciente atención en la necesidad de establecer verificaciones objetivas de todos los pronósticos que produce. Entre ellos, se encuentra el pronóstico estacional, el cual es utilizado por una gran variedad de usuarios y resulta de relevancia para aquellos sectores que necesitan tener algún tipo de predictibilidad para planificar sus acciones en la escala de mediano plazo, como es el caso de las actividades agrícolas, hídricas, el sector energético, entre otros.

El pronóstico estacional del SMN consiste en un pronóstico probabilístico por consenso elaborado desde la Dirección Central del Monitoreo del Clima en conjunto con la Dirección Nacional de Servicios Sectoriales. También participan el Departamento de Ciencias de la Atmósfera y los Océanos, y el Centro de Investigaciones del Mar y la Atmósfera. Dicho pronóstico se realiza mensualmente para todo el país para las variables temperatura media y precipitación acumulada, y corresponde a un campo o mapa con un sistema de probabilidades de tres categorías (“superior a lo normal”, “normal” e “inferior a lo normal”). Tiene la particularidad de contener una componente subjetiva: su elaboración se basa en el análisis que realizan los expertos no sólo de las condiciones atmosféricas-oceánicas prevalentes, sino también de las salidas disponibles de diversos modelos climáticos. Si bien todos los expertos utilizan como guía o referencia la misma información objetiva, sus valoraciones son únicas y difieren entre sí producto de su experiencia y percepción.

El proceso de verificación tiene como principal objetivo la evaluación del pronóstico por consenso. A su vez, comprende la comparación entre el desempeño de dicho pronóstico con el de los pronósticos objetivos provenientes de otros centros, de modelos climáticos, o del propio SMN. El propósito de éste es determinar si la contribución subjetiva de los pronosticadores podría mejorar las predicciones derivadas de modelos y algoritmos. Al momento, los pronósticos probabilísticos considerados para la comparación son: North American Multi-model Ensemble (NMME) (Kirtman y otros, 2014), CLIMAX (calibración de modelos que integran el NMME - Osman y otros, 2021), IRI (calibración de modelos que integran el NMME), SMN-CPT (modelo estadístico) y SMN-CPT.NMME (combinación de modelo estadístico con calibración del ensemble del NMME).

2) METODOLOGÍA

Para la verificación se utilizaron diversas métricas sugeridas por la “Guía de Verificación de los Pronósticos Climáticos Estacionales Operativos” (Mason, 2018), de la Organización Meteorológica Mundial, en donde se diferencian los procedimientos a seguir según si se tiene una serie de pronósticos o un mapa de pronóstico individual para un trimestre específico. En este trabajo, se consideraron ambos enfoques. Las métricas empleadas para el primer enfoque

fueron los diagramas y área ROC, y los diagrama de confiabilidad; mientras que para el segundo, el hit score, el ignorance score y average interest rate. Cabe destacar que, si bien un sistema de pronóstico sólo puede ser evaluado adecuadamente sobre la base de una larga serie, analizar el desempeño por trimestre también resulta útil y valioso, en particular para ser conscientes del valor de la información brindada al público en un determinado período.

Las métricas fueron calculadas en puntos de estación y sintetizadas a nivel nacional y regional. Para ello, fue necesario interpolar los datos grillados de los modelos NMME, IRI y CLIMAX a las coordenadas de las estaciones, y remover aquellos pronósticos no informativos (aquellos que no se diferencian del uso de la climatología o los que carecen de interés por coincidir con la estación seca) previo a cualquier cálculo. Lo segundo obligó a añadir, junto con los resultados de las métricas, la información del porcentaje de predicciones intervinientes en el trimestre a la hora de evaluar un mapa de pronóstico individual, para lograr una correcta interpretación de las verificaciones al comparar. Por último, para el análisis de series de pronóstico, se computaron los intervalos de confianza de cada métrica mediante la técnica de remuestreo o bootstrap (Efron, 1979) con reposición, con el objetivo de subsanar las incertidumbres acarreadas debido que se cuenta con una muestra pequeña de pronósticos del consenso (desde 2018 hasta la fecha).

Los procedimientos que se llevaron a cabo se enmarcan dentro del “Plan transversal de verificación de pronósticos del SMN” (de Elia, 2021), e intentan ser coherentes con los adoptados en las verificaciones de los otros pronósticos de la institución. Su descripción detallada se encuentra en Poggi y otros (2021).

3) RESULTADOS

En las Tablas 1 y 2 se presentan los valores de las áreas ROC para la temperatura y la precipitación, respectivamente, por categoría de pronósticos. Esta métrica permite evaluar el desempeño de los modelos considerando las series totales de predicciones, que abarcan desde enero-febrero-marzo (EFM) de 2018 hasta octubre-noviembre-diciembre (OND) de 2021.

Área ROC	Modelos	Categorías		
		Superior a lo normal	Normal	Inferior a lo normal
Temperatura	Consenso	0.60	0.57	0.54
	SMN-CPT	0.56	0.61	0.53
	SMN-CPT.NMME	0.56	0.61	0.55
	CLIMAX	0.55	0.54	0.50
	IRI	0.54	0.50	0.53
	NMME	0.58	0.52	0.53

Tabla I: Áreas ROC de los modelos considerados, para la temperatura y las tres categorías de pronóstico, desde EFM de 2018 a OND de 2021. Valores mayores que 0.5 indican skill.

Área ROC	Modelos	Categorías		
		Superior a lo normal	Normal	Inferior a lo normal
Precipitación	Consenso	0.60	0.50	0.56
	SMN-CPT	0.56	0.58	0.56
	SMN-CPT.NMME	0.51	0.57	0.51
	CLIMAX	0.61	0.52	0.57
	IRI	0.63	0.49	0.60
	NMME	0.57	0.52	0.50

Tabla II: Ídem Tabla I pero para la variable precipitación.

Para la temperatura, se observa que el consenso es el que presenta el mayor valor en la categoría “superior a lo normal”, exhibiendo la mejor habilidad para distinguir un evento cálido de un no evento. El SMN-CPT y el SMN-CPT.NMME muestran el mejor desempeño en discriminar las condiciones normales, no así el IRI. Éste último tiene la menor cantidad de pronósticos informativos de temperatura de entre todos los modelos considerados. La máxima performance para discriminar las condiciones frías se advierte en el SMN-CPT.NMME seguido por el consenso, mientras que la mínima, en el CLIMAX.

Para la variable precipitación, el IRI presenta los máximos valores de área ROC en las categorías “superior a lo normal” e “inferior a lo normal”. Como en el caso de la temperatura, estos valores están condicionados por la poca cantidad de pronósticos informativos que brinda el modelo. Al momento, el IRI tiene la mayor capacidad para discriminar las condiciones húmedas y secas, pero no así las normales. Junto con el consenso, muestra la menor habilidad en distinguir los eventos de lluvia dentro del rango normal. Para esta categoría, nuevamente, el SMN-CPT y SMN-CPT.NMME exhiben los mejores resultados.

La Tabla III muestra los límites superiores e inferiores de los intervalos de confianza de los valores de área ROC, únicamente para el consenso. El conocimiento de los intervalos de confianza permite hacer inferencias, con un cierto nivel de confianza, sobre el verdadero resultado de las métricas. Esto cobra relevancia al tener muestras pequeñas de datos.

Categorías	Temperatura		Precipitación	
	Límite inferior	Límite superior	Límite inferior	Límite superior
Superior a lo normal	0.59	0.61	0.58	0.62
Normal	0.56	0.58	0.49	0.51
Inferior a lo normal	0.52	0.55	0.55	0.58

Tabla III: Límites inferior y superior de los intervalos de confianza para los valores de área ROC, calculados a partir de la técnica de bootstrap con reposición, en función del período EFM de 2018 a OND de 2021. Se consideró un remuestreo con 20 muestras y un nivel de confianza del 90%.

REFERENCIAS

Efron, B., 1979: Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1–26.

de Elía, R., 2021: Plan transversal de verificación de pronósticos del SMN. Nota Técnica SMN 2021-84.

Kirtman, B.P. and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. *Bull. Amer. Meteor. Soc.*, 95, 585–601, <https://doi.org/10.1175/BAMS-D-12-00050.1>.

Mason, S.J., 2018: Guidance on Verification of Operational Seasonal Climate Forecasts. World Meteorological Organization, Commission for Climatology XIV Technical Report.

Osman, M., Coelho, C. and Vera, C., 2021: Calibration and combination of seasonal precipitation forecasts over South America using Ensemble Regression. *Climate Dynamics*. doi:10.1007/s00382-021-05845-2.

Poggi, M.M., Skansi, M.M. y de Elía, R., 2021: Metodología de verificación del pronóstico estacional por consenso. Nota Técnica SMN 2021-110.