

PREDICTABILIDAD ESTACIONAL DEL ÍNDICE DÍAS CÁLIDOS DE VERANO EN ARGENTINA: APLICACIÓN DE UN MODELO ESTADÍSTICO

Soledad M. COLLAZO ^{1,2}, Mariana G. BARRUCAND ^{1,2}, Matilde M. RUSTICUCCI ^{1,2}
scollazo@at.fcen.uba.ar

¹Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET-UBA)

²Departamento de Ciencias de la Atmósfera y los Océanos (FCEyN, UBA)

RESUMEN

La habilidad de predecir extremos de temperatura permitirá tomar medidas de precaución para evitar o reducir sus riesgos. El objetivo de este estudio es ajustar un modelo estadístico que permita predecir los días cálidos de verano en escala estacional en Argentina al norte de 40 °S. Para alimentar el modelo, se tuvieron en cuenta gran cantidad de índices representativos de distintos forzantes atmosféricos, oceánicos y de interacción suelo-atmósfera en el mes previo al trimestre de verano. Debido a la colinealidad de los predictores, se aplicó el modelo estadístico Regresión de Componentes Principales para simular los días cálidos de verano en ocho localidades representativas de distintas regiones del país. Para evaluar el desempeño del modelo, se calcularon distintas métricas estadísticas entre los valores observados y simulados sobre una muestra de testeo. Se encontraron correlaciones significativas entre ambos (superiores a 0.5), excepto en la región noroeste, y sesgos fríos en todas las estaciones.

ABSTRACT

The ability to predict temperature extremes will allow to take precautionary measures to avoid or reduce risks. The aim of this study is to adjust a statistical model to predict warm summer days on a seasonal scale in Argentina north of 40 °S. To feed the model, a large number of representative indices of different atmospheric, oceanic and soil-atmosphere interaction drivers were taken into account in the previous month to the summer season. Due to the colinearity of the predictors, the statistical model of Principal Components Regression was applied to simulate warm summer days in eight representative locations of different regions of the country. To evaluate the performance of the model, different statistical metrics were calculated between the observed and simulated values on a test sample. Significant correlations were found between both (higher than 0.5), except in the northwest region, and cold biases in all stations.

Palabras clave: predicción estacional, extremos de temperatura, regresión de componentes principales

1) INTRODUCCIÓN

La predicción de fenómenos meteorológicos es muy compleja, esto se debe a la naturaleza caótica de la atmósfera. Como una alternativa a los sistemas de predicción dinámica, los enfoques empíricos apuntan a describir una relación física conocida entre una variable respuesta y los fenómenos climáticos precedentes (los predictores). Una técnica estadística útil para predecir series temporales cuando hay una gran cantidad de variables, es el análisis de componentes principales seguido de regresión lineal múltiple (Çelik, 2018). Esta combinación se llama regresión del componente principales (RCP). Este método reduce la complejidad del sistema ya que las variables se combinan linealmente para obtener las componentes principales y solo se conservan aquellas que maximizan la varianza explicada (Zhang et al., 2013). El objetivo de este trabajo es ajustar el modelo estadístico RCP para predecir los extremos cálidos de la temperatura máxima de verano (TX90 DEF) en Argentina al norte de 40°S y evaluar su desempeño sobre una muestra de testeo.

2) DATOS Y METODOLOGÍA

Para construir el índice TX90 DEF, se utilizaron datos diarios de temperatura máxima de verano (Diciembre-Enero-Febrero) de 8 estaciones representativas de distintas regiones ubicadas al norte de 40°S en el periodo 1970-2015, provistos por el SMN y el INTA. El índice TX90 representa el porcentaje de días en que la temperatura máxima excede el percentil 90, estimado en el periodo base 1981-2010. Además, se consideraron los siguientes forzantes: índices climáticos (como el ENOS), índices de circulación regional (anomalías de altura geopotencial, viento zonal en niveles altos, índices de bloqueo, etc), índices de precipitación estandarizada, temperatura superficial del mar del Atlántico, entre otros. Para detectar la colinealidad, analizamos las correlaciones de Pearson simultáneamente entre los índices, después de haber

filtrado las tendencias. Por ultimo, se ajustó el modelo de RCP para predecir TX90 DEF. Para ello se dividieron las series en una muestra de entrenamiento (1970-1999) y una muestra de testeo (2000-2015) sobre la que se evaluó el desempeño del modelo mediante el cálculo de diferentes métricas: bias, error cuadrático medio (RMSE), error absoluto medio (MAE) y coeficiente de correlación de Pearson. Para la regresión, solo se consideraron las componentes principales necesarias para explicar el 90% de la varianza acumulada.

3) RESULTADOS

El análisis de la colinealidad reveló que el El Niño-Oscilación del Sur (ENOS) está significativamente correlacionados con el dipolo del Indico, jet subtropical, Oscilación Decadal del Pacífico, Índice de Precipitación Estandarizada, entre otros; que enfatizó cuán relevante es ENOS para la predicción estacional de la temperatura extrema en Argentina.

Se ajustó el modelo RCP a ocho estaciones representativas de distintas regiones empleando todos los índices para la construcción de las componentes principales. La Tabla 1 presenta distintas métricas estadísticas para evaluar el desempeño del modelo sobre la muestra de testeo. Exceptuando Famailla, las correlaciones entre el TX90 DEF observado y el modelado fueron significativas al 5%, lo que indica que el modelo es capaz de representar la variabilidad interanual. En todas las estaciones analizadas se encontró un bias frío del modelo. En particular, TX90 DEF del año 2013 fue claramente subestimado en las ocho estaciones analizadas, probablemente debido a la ocurrencia de una inusual ola de calor extrema en Diciembre 2013 que dificultó que el modelo representara adecuadamente a TX90.

4) CONCLUSIONES

En el presente trabajo, se presenta un enfoque novedoso al ajustar el modelo estadístico RCP para predecir el índice de extremos de temperatura TX90 DEF en Argentina al norte de 40°S. Las métricas estadísticas obtenidas al evaluar el modelo en la muestra de testeo revelaron buenos resultados para predecir la temperatura máxima extrema en verano, con excepción de la región noroeste. El modelo presentó bias fríos en todas las estaciones, y en particular el año 2013 fue claramente subestimado. En el futuro, se busca extender este análisis a otros trimestres e índices de extremos, así como también aplicar otros modelos estadísticos.

AGRADECIMIENTOS

Esta investigación fue financiada por CONICET PIP 0137-Res 4248/16. Queremos agradecer al SMN y al INTA por proporcionar los datos para este estudio.

REFERENCIAS

- Çelik, K., 2018. Predicting chlorophyll-a concentrations in two temperate reservoirs with different trophic states using Principal Component Regression (PCR). *Oceanol. and Hydrobiol. Stud.* 47(1), pp. 1-9. doi:10.1515/ohs-2018-0001
- Zhang, W., Lou, IC., Kong, Y., Ung, WK., Mok, M., 2013. Eutrophication analysis and principal component regression for two subtropical storage reservoirs in Macau. *Desalin. Water. Treat.* 51: 7331–7340. DOI: 10.1080/19443994.2013.793921.

| | Bias | RMSE | MAE | Correlación |
|--------------|-------|-------|------|-------------|
| Posadas | -2,85 | 6,88 | 5,56 | 0,71 |
| Famailla | -6,03 | 10,11 | 8,17 | 0,46 |
| San Luis | -4,20 | 6,99 | 5,02 | 0,61 |
| Pilar Obs | -1,97 | 7,58 | 5,96 | 0,53 |
| Paraná Aero | -2,89 | 6,16 | 4,56 | 0,73 |
| Junín | -5,26 | 6,60 | 5,48 | 0,86 |
| Tandil | -3,60 | 6,63 | 4,94 | 0,64 |
| General Pico | -2,85 | 7,40 | 6,28 | 0,60 |

Tabla 1: Métricas estadísticas (bias, RMSE, MAE, coeficiente de correlación) calculadas entre TX90 DEF observado y modelado usando RCP en la submuestra de testeo. Bias, RMSE y MAE en [% de días]

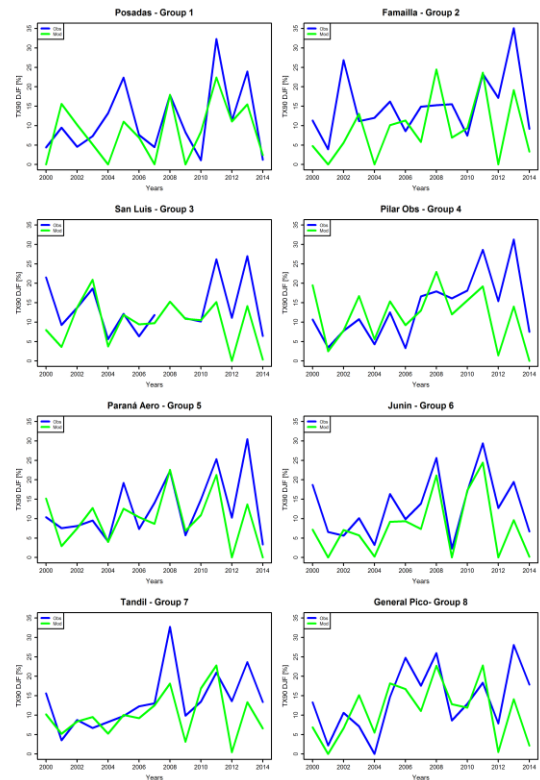


Figura 1: TX90 DEF observado (azul) y modelado (verde) usando RCP en la submuestra de testeo para ocho estaciones.