Comparison between the Operational and Statistical Daily Maximum and Minimum Temperature Forecasts on the Central Coast of Peru

VANNIA ALIAGA-NESTARES,[a] GUSTAVO DE LA CRUZ,[a] AND KEN TAKAHASHI[a,b]

[a] *Servicio Nacional de Meteorología e Hidrología del Perú—SENAMHI, Lima, Peru*
[b] *Instituto Geofísico del Perú, Lima, Peru*

(Manuscript received 10 June 2021, in final form 21 November 2022)

ABSTRACT: Multiple linear regression models were developed for 1–3-day lead forecasts of maximum and minimum temperature for two locations in the city of Lima, on the central coast of Peru (12°S), and contrasted with the operational forecasts issued by the National Meteorological and Hydrological Service—SENAMHI and the output of a regional numerical atmospheric model. We developed empirical models, fitted to data from the 2000–13 period, and verified their skill for the 2014–19 period. Since El Niño produces a strong low-frequency signal, the models focus on the high-frequency weather and subseasonal variability (60-day cutoff). The empirical models outperformed the operational forecasts and the numerical model. For instance, the high-frequency annual correlation coefficient and root-mean-square error (RMSE) for the 1-day lead forecasts were 0.37°–0.53° and 0.74°–1.76°C for the empirical model, respectively, but from around −0.05° to 0.24° and 0.88°–4.21°C in the operational case. Only three predictors were considered for the models, including persistence and large-scale atmospheric indices. Contrary to our belief, the model skill was lowest for the austral winter (June–August), when the extratropical influence is largest, suggesting an enhanced role of local effects. Including local specific humidity as a predictor for minimum temperature at the inland location substantially increased the skill and reduced its seasonality. There were cases in which both the operational and empirical forecast had similar strong errors and we suggest mesoscale circulations, such as the low-level cyclonic vortex over the ocean, as the culprit. Incorporating such information could be valuable for improving the forecasts.

SIGNIFICANCE STATEMENT: We wanted to compare the temperature of the operational forecast of the Meteorological and Hydrological Service, an atmospheric model, and persistence with the observed temperatures on the Peruvian central coast. In addition, we generated an empirical forecast model considering both atmospheric and local predictors. We got better results with this empirical model, considering the highest Pearson correlations and the lowest RMSE values. These results will allow us to use this empirical model as the main tool to automate the forecast on the central coast of Peru. Future work should be aimed at testing the skill of this model for forecasting in other cities of Peru.

KEYWORDS: Synoptic climatology; Synoptic-scale processes; Regression analysis; Forecast verification/skill; Numerical weather prediction/forecasting

1. Introduction

Daily weather forecasting is one of the most basic activities in National Meteorological and Hydrological Services (NMHSs). At the *Servicio Nacional de Meteorología e Hidrología* of Peru (SENAMHI), forecasters do so by analyzing the evolution of weather systems, aided by numerical model output. However, human-modified forecast information has been lately found to not necessarily outperform model systems, particularly as the model skill continues to increase (Bosart 2003; Doswell 2004; Novak et al. 2014). Thus, it is important for NMHSs to investigate and identify the conditions and applications for which forecasters can provide the most added value.

Peru is a tropical country, and its weather predictability is partly associated with tropical systems, such as the Bolivian high, the summertime upper-level anticyclonic circulation associated with moisture transport from the Amazon to the Andes (Garreaud 1999) and equatorial convectively coupled waves (Takahashi 2004; Huaman et al. 2020). Furthermore, El Niño makes a substantial contribution to the temperature variability on interannual time scales on the central coast (Takahashi and Martínez 2017; Cai et al. 2020), while oceanic Kelvin waves modulate it on subseasonal time scales (Illig et al. 2014). On the other hand, extratropical phenomena have a large influence on the Peruvian weather, e.g., cold surges in the Amazon region (Espinoza et al. 2013), snowfall in the Andes associated with cutoff lows (COLs) (Quispe Vega 2017), and the influence on the coastal trade winds of the variability in the South Pacific subtropical anticyclone (Dewitte et al. 2011; Illig et al. 2014). Weather variability along the arid central coast of Peru is mild, without severe weather events, and although it is influenced by

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/WAF-D-21-0094.s1.

Corresponding author: Vannia Aliaga Nestares, valiaga@senamhi.gob.pe

DOI: 10.1175/WAF-D-21-0094.1

large-scale conditions, the experience of the forecasters indicates that local mesoscale circulations also play a role. Furthermore, since a third of the population of Peru lives in this region, primarily in its capital Lima and its adjacent port province Callao (hereafter jointly referred to as city of Lima, ~12°S), the forecasts for the city of Lima have high visibility to the public. These conditions make weather forecasting for this region a challenge to SENAMHI, particularly with respect to the day-to-day temperature and cloudiness variability, even for lead times of one day.

In this study we document the skill of operational daily maximum and minimum temperature forecasts from SENAMHI in 10 locations of the central coast of Peru, including the city of Lima, and propose empirical models to improve said forecasts. Furthermore, we address two specific research goals: (i) to develop empirical models based on multiple linear regression for the daily prediction of maximum and minimum temperature; and (ii) to assess and compare four different approaches to forecast: SENAMHI's operational forecast, regional numerical model forecast, persistence forecast (based on the previous day's temperature), and the empirical model.

a. *Climate on the central coast of Peru*

The arid central coast of Peru is under the influence of the cool eastern boundary coastal upwelling in the Pacific Ocean, capped by a persistent low-level thermal inversion (Prohaska 1973; Enfield 1981; Arellano Rojas 2013) and isolated from the Amazonian warm and moist conditions by the Andes mountains (Takahashi and Battisti 2007). Thus, despite its low latitude, this region has a mild climate and an annual cycle modulated by the ocean (Takahashi 2004). The winters feature low stratiform cloudiness, high relative humidity, and drizzle that peaks in June–August, while summer is generally sunnier, warmer, and has higher specific humidity, with the highest temperatures in January–March, but also features high cloudiness and some rain episodes associated with the rainy season in the Andes (Prohaska 1973).

The main large-scale atmospheric circulation features that affect the climate of the central coast are the southeast Pacific anticyclone (SEPA) and the Bolivian high (BH). The SEPA corresponds to the subtropical subsidence branch of the Hadley circulation, affected by the presence of the Andes, with contribution from the summertime monsoon in the Amazon (Rodwell and Hoskins 2001; Takahashi and Battisti 2007). The associated southeasterly trade winds along the coast of Peru are responsible for the coastal ocean upwelling that maintains the typically low sea surface temperature (SST), which is also contributed by dry subsidence and the positive feedback from stratiform clouds (Klein and Hartmann 1993; Takahashi and Battisti 2007). The SEPA is more intense in austral spring, and reaches its southernmost position (centered at 35°S) during summer and its northernmost position (center at 27°S) during fall and winter (Febre Pérez 2018). The winds of the SEPA flow over cold water as they approach the coast, thus cooling and increasing the relative humidity. These conditions are favorable to the formation of stratiform cloudiness and fog on the subtropical west coast of South America, mainly in austral winter (Pinche Laurre 1986). The BH is a summertime upper-level anticyclonic circulation centered above the altiplano that favors the transport of moisture from the Amazon to the Andes (Garreaud 1999) associated with a mixing ratio above 9–10 g kg^{-1} at 600 hPa and over 4–5 g kg^{-1} at 500 hPa (Rivas Quispe 2019). The BH or, more generally, mid- and upper-level easterly flow, may generate spillover from the Andes toward the coast producing rain in the latter region (Rollenbeck and Bendix 2011).

b. *Subseasonal and interannual variability along the coast*

Interannual climatic variability of the coast of Peru is dominated by El Niño and La Niña (Cai et al. 2020; Capotondi et al. 2020) through the positive effect of local SST anomalies on air temperature and, during the warmest months, the occurrence of heavy precipitation and flooding that is also favored by cool conditions in the equatorial central Pacific (Takahashi and Martínez 2017). Extreme El Niño events, such as the 1983, 1998, and 2016 events, are much stronger than La Niña in the eastern Pacific (e.g., Takahashi et al. 2011), although the 2016 event was associated with warming but not heavy precipitation (L'Heureux et al. 2017). Although the large-scale ENSO variability occurs on the interannual scale, the "coastal El Niño" occurs on subseasonal to seasonal time scales associated with air–sea interactions in the far-eastern Pacific, such as in 1891, 1925 and 2017 (Takahashi and Martínez 2017; Garreaud 2018; Takahashi et al. 2018; Echevin et al. 2018; Peng et al. 2019).

On subseasonal time scales, the Madden–Julian oscillation modulates rainfall in northern Peru through its effect on the Walker circulation. Furthermore, equatorial ocean Kelvin waves and extratropical atmospheric variability also modulate the coastal SST on subseasonal time scales through their effect on local oceanic thermal advection and surface heat fluxes, with weaker trade winds and downwelling Kelvin waves leading to positive SST anomalies (Dewitte et al. 2011; Illig et al. 2014).

At higher frequencies, the main synoptic systems that determine the meteorological conditions on the central coast of Peru are the low-level trade winds (850–1000 hPa) associated with the SEPA and modulated by eastward subtropical transients (Dewitte et al. 2011; Rahn and Garreaud 2013), which have more predominance in winter; and the zonal moisture transport at midlevels (500–600 hPa) and short-wave troughs at high levels (300–200 hPa), associated with BH variability, predominantly in summer (Rivas Quispe 2019). The analysis of such synoptic systems is the backbone of the conceptual models considered for weather forecasting on the central coast. Thus, in 2019, in a meeting with 10 weather forecasters of SENAMHI, it was determined that the main large-scale factors that regulate the temperatures on the central Peruvian coast are the SST, the position of the APSO, the interaction of winds at low levels off the Peruvian coast, i.e., northerly winds at 850 hPa and southerly winds at lower levels (950–1000 hPa), precipitable water on the coast, winds and moisture advection from the east at medium levels (500–600 hPa), the BH location, and the presence of high-level systems, such as a cutoff low (COL) (Quispe-Gutiérrez et al. 2021) (Fig. 1).

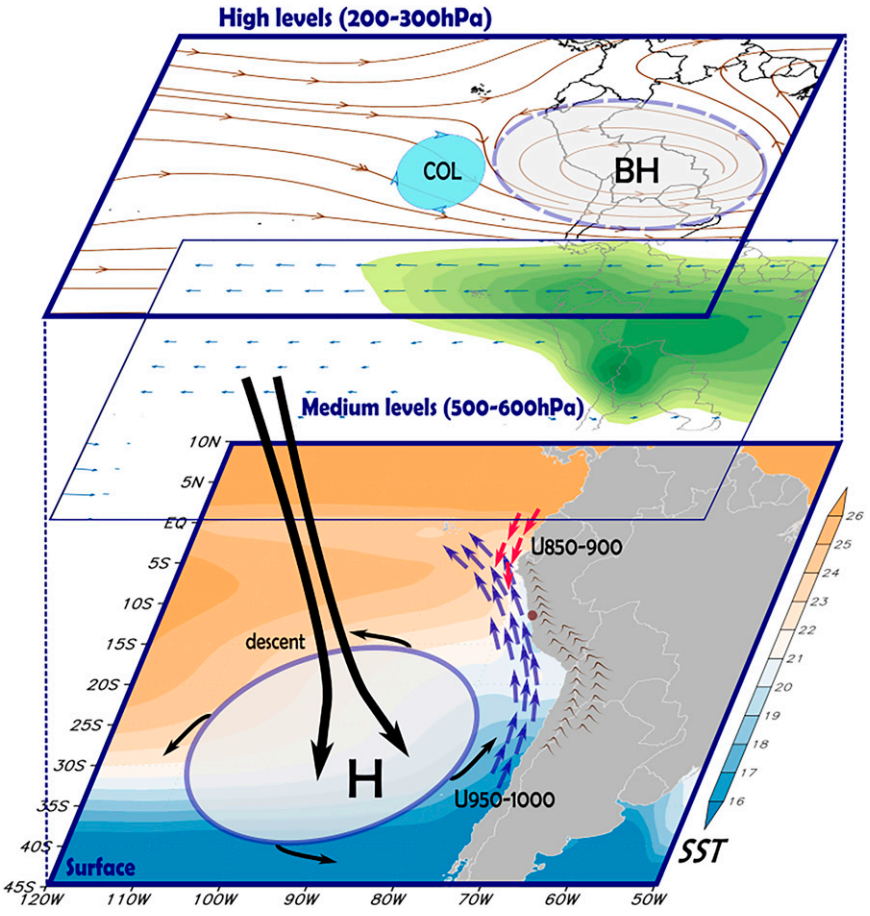

FIG. 1. Main large-scale variables considered by the SENAMHI meteorologists for the temperature forecast on the central Peruvian coast. Sea surface temperature (SST; °C; shaded), the SEPA denoted by the "H" at the surface, wind at 850–900 hPa (U850–900), wind at 950–1000 hPa (U950–1000), the Bolivian high (BH), and cutoff low (CL).

c. *Climate of the city of Lima*

Climatically, the city of Lima can be divided into a western and eastern zone. The western zone is near the sea and the eastern zone is more continental. This condition affects the temperature due to the thermoregulatory effect of water. On the other side, the main type of precipitation in Lima is drizzle, which occurs from April to October, with more intensity and frequency in winter (June–August), occurring during more than 90% of days with drizzle in this period. The second type of precipitation in Lima is rain, which occurs between November and March, with more frequency in summer (January–March); however, it is less frequent than drizzle, raining only between 5% and 10% of the days of this period.

This behavior is evident when reviewing meteorological data. According to the Callao meteorological station, which represents the western zone (Fig. 2), the monthly mean maximum and minimum temperature are 27.4°C (February) and 14.6°C (July), respectively (Fig. 3); the drizzliest month is July with a mean monthly amount of 1.8 mm of total precipitation, and the rainiest month is January with 0.9 mm. On the other hand, the eastern zone, where the La Molina meteorological station is located (Fig. 2), the monthly mean maximum and minimum temperature are 28.7°C (February) and 13.8°C (August), respectively, the drizzliest month is August with 2.3 mm month^{-1} of total precipitation, and the rainiest month is November with a monthly mean of 1.1 mm month^{-1} (Fig. 3).

About the types of clouds in Lima, stratiform clouds are the main type of cloudiness in Lima and are responsible for the drizzle. They present in a saturated layer whose position varies along the year. This layer is limited below by the level of condensation and above by the base of the subsidence inversion (Pinche Laurre 1986), which is higher in the winter than in the summer Prohaska (1973). According to Albrecht (1981) and Arellano Rojas (2013), the inversion is strongest and most persistent during the austral winter with an average base height of ~700 m MSL and a bit weaker and shallower during the fall, with a base pressure of around 350 m MSL (Fig. 4).

On the other hand, rain in Lima is generated by the horizontal advection of moisture and cloudiness from the eastern Cordillera to the medium and high atmosphere of the occidental highlands. Additionally, the convergence with the mountains

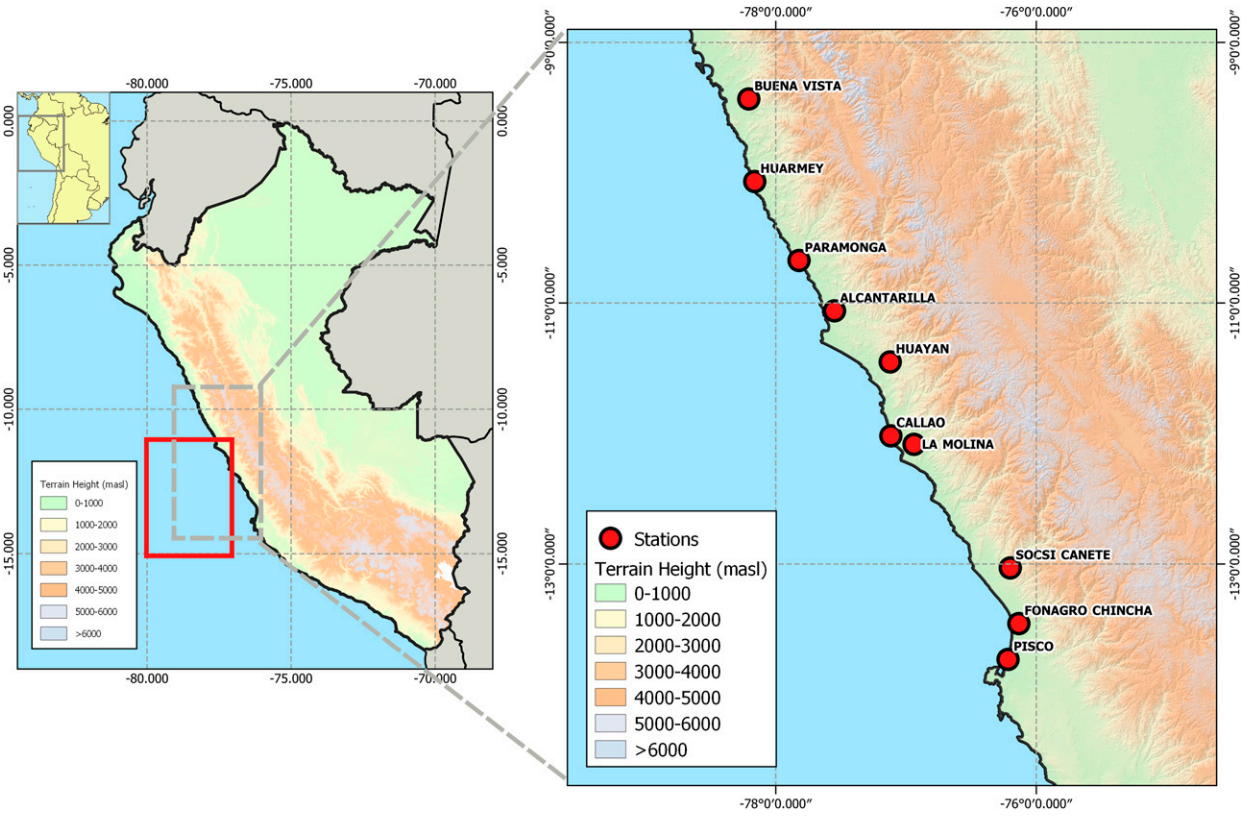

FIG. 2. Spatial locations of 10 stations on the central Peruvian coast (red points) and selected area for large-scale predictors (red box).

of the low-level westerlies may help the triggering of convection by orographic lifting over the western slope of the Andes (Takahashi 2004). Many times, these westerlies are associated with a sea breeze, that is stronger in austral summer (Enfield 1981), or to synoptic winds. This concept changes a bit from what was said by Prohaska (1973), that the occasional light rainfalls in summer over Lima are mostly not of Pacific origin but are a spillover of the intense precipitation in the Andes and the Amazonian watershed.

d. *Temperature variations in Lima*

The daily temperature variation in Lima is normally slight (around 1°C), which makes the use of the previous day's temperature (persistence) a useful method in most

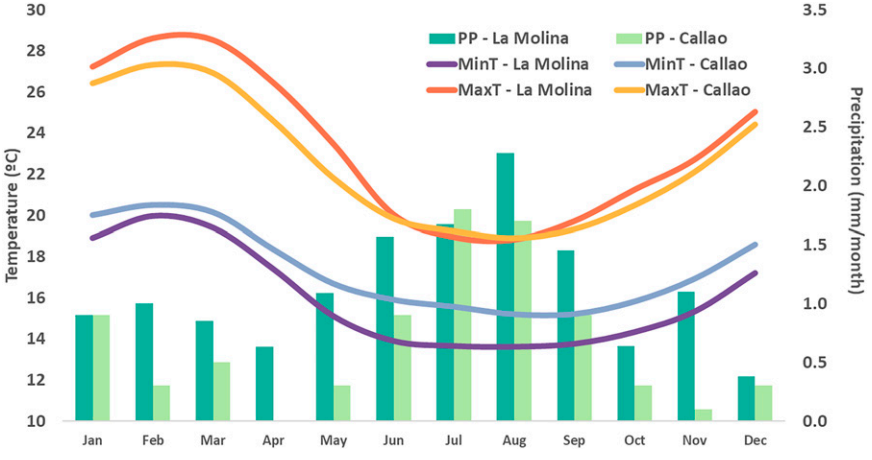

FIG. 3. Monthly climatology of temperature (lines) and precipitation (bars) at the La Molina and Callao stations. The precipitation between April and October represents drizzle, and between November and March, it represents rain.

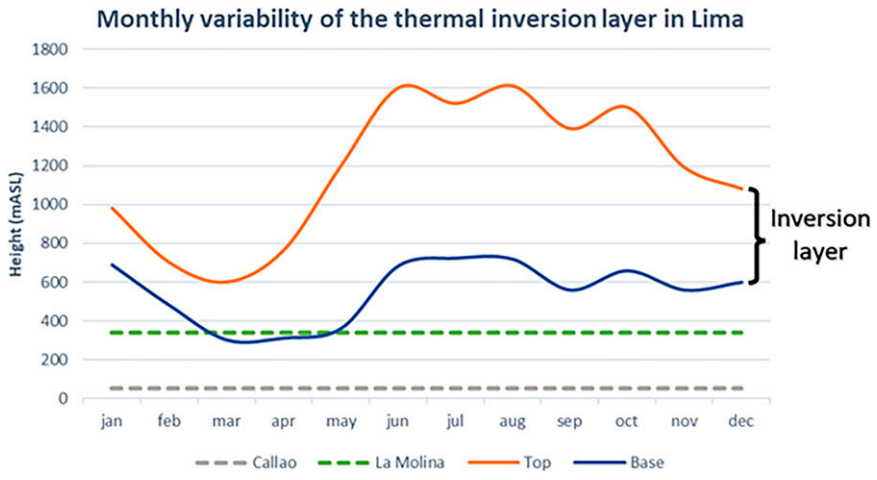

FIG. 4. Monthly variability of the thermal inversion layer in Lima. Gray and green lines represent the altitude of the Callao and La Molina stations, respectively. The top and the base of the inversion layer are shown in orange and blue, respectively. Adapted from Arellano Rojas (2013).

daily forecasts; however, there are some situations in which there are very marked anomalies of the synoptic systems or the presence of mesoscale systems that generate a sudden variation in temperature from one day to the next. One of these mesoscale systems is the low-level cyclonic vortex (VCNB, for its acronym in Spanish) which we will describe later.

The objective of this article is to detect, through an automated empirical forecast, these sudden changes in temperature from one day to the next.

2. Data and methodology

a. Data

1) OPERATIONAL FORECAST DATA

In this study, in order to evaluate the performance of the operational forecast, we use the daily forecast data made by SENAMHI for 10 localities: Callao, Pisco, Huarmey, Paramonga, Fonagro Chincha, and Alcantarilla for the west zone, and La Molina, Buena Vista, Socsi Cañete, and Huayan for the east zone, for 1–3-day leads. This forecast is submitted and published on the SENAMHI web page (https://www.senamhi.gob.pe/?p=pronostico-meteorologico) every day at 1900 local time (LT). It is based on the expert analysis by the forecaster, which considers the atmospheric patterns and their behavior predicted by numerical models, station data, soundings, and satellite imagery.

2) ETA—SENAMHI MODEL

The ETA model has been widely used as a weather forecasting tool at NCEP and at other national services and research institutions (Janjić 1984; Mesinger 1984; Mesinger et al. 1988; Black 1994). At SENAMHI, it has been run operationally with the current configuration since 2009, in support of weather forecasts. It has two domains with grid spacings of 32 km (ETA32) and 22 km (ETA22), and both of them provide forecasts at 6-h temporal resolution. The 32-km domain covers the South America region and has 50 vertical levels, and the 22-km is a domain that covers Peru and has 38 pressure levels. The physical parameterization schemes used in these domains include the Kain–Fritsch cumulus parameterization scheme, the Mellor–Yamada 2.5 planetary boundary layer (PBL) scheme, and the Geophysical Fluid Dynamics Laboratory (GFDL) radiation scheme. The model is initialized daily at 0000 UTC and the data from the Global Forecast System (GFS) model are used as lateral boundary conditions. In this study, we used temperature at 2 m from the ETA22 model data (1200 and 1800 UTC).

3) OBSERVATIONAL DATA

Daily maximum and minimum temperatures registered from January 2000 to January 2019 at the representative meteorological stations for the 10 forecast points of the central coast of Peru (Fig. 2) were selected and used. We also considered specific humidity from all stations except in Callao and Pisco (according to availability).

For the period from 2000 to 2019, daily temperature series from 10 stations along the central coast of Peru were carefully selected and quality controlled. A quality control software developed at SENAMHI was employed in combination with a visual inspection to identify and remove measurement errors. Additionally, these time series were homogenous and had sufficient data, with less than 30% of daily missing values.

4) REANALYSIS DATA

We employed daily SST and mean sea level pressure, wind, temperature, and relative humidity at different pressure levels for 0000, 0600, 1200, and 1800 UTC as potential predictors (Table 1) from ERA5 reanalysis. This dataset has global gridded information on 37 pressure levels (from 1000 to 1 hPa) and the surface with $0.25° \times 0.25°$ spatial resolution, hourly, and from 1979 to the present.

TABLE 1. Predictors identified after correlation matrix analysis and forward selection. PersTmax: maximum temperature value of the previous day; PersTmin: minimum temperature value of the previous day; PersQ: specific humidity value of the previous night; V1000_00z: meridional wind at the 1000-hPa isobaric level at 0000 UTC; T850_00z: temperature at the 850-hPa isobaric level at 0000 UTC; T925_18z: temperature at the 925-hPa isobaric level at 1800 UTC; Q1000_00z: specific humidity at the 1000-hPa isobaric level at 0000 UTC.

Predictand	Selected predictors	Calibration period	Validation period
Tmax-Callao	PersTmax, V1000_00z, T850_00z	2000–13	2014–19
Tmin-Callao	PersTmin, Q1000_00z, V1000_00z	2000–13	2014–19
Tmax-La Molina	PersTmax, T925_18z, V1000_00z	2000–13	2014–19
Tmin-La Molina (EMP)	PersTmin, Q1000_00z	2000–13	2014–19
Tmin-La Molina (EMPq)	PersTmin, PersQ	2010–13	2014–19

b. Methodology

1) Data filtering

Since our main purpose is to evaluate and improve the daily forecast, it is important to improve the accuracy in forecasting the high-frequency variability. To do this, an eighth-order Butterworth low-pass filter using a cutoff period of 60 days was first applied to remove the seasonal and lower-frequency variability of all datasets.

2) Identification of potential predictors

The procedure to identify and select the predictors from the large-scale atmospheric variables is as follows (Fig. 5). We first calculated the daily correlation between the high-pass-filtered series of predictands (T'^{pred}) and the potential large-scale predictors (X') (Table 1). Then we constructed potential predictor indices from the spatial average of the atmospheric fields over the areas that had the highest correlations and that were physically related to the atmospheric circulation of the Peruvian central coast.

Additionally, we considered the previous day (persistence) high-pass filtered station temperatures (T'^{obs}_{t-1}) in the predictor suite.

For Tmin La Molina we also considered the station specific humidity from the previous night (q'^{obs}_{t-1}) as a potential predictor, which was motivated by a prior exploratory statistical analysis that showed a high positive correlation (Fig. S1).

3) Selection of predictors

To select the best sets of predictors, we established a Pearson correlation matrix between the potential predictor indices, and selected only those with a correlation less than 0.5 with the rest to avoid multicollinearity. With this, we removed highly correlated potential predictors.

Then, to select the best model of predictors for each predictand, we used the cross-validation method (Wilks 2006). For that, we divided the data in subsets to develop the empirical models, considering this formulation:

$$T'^{\text{pred}}_t = \beta_o T'^{\text{obs}}_{t-1} + \sum_{i=1}^{N} \beta_i X'_{i,t}, \tag{1}$$

where T_t is the temperature for the time t, $X_{i,t}$ is the large-scale predictor, N is the total number of large-scale predictors, and β are the parameters for the regression model. The "pred" and "obs" terms represent the predicted and observed temperature, respectively, in Eq. (1). It is important to mention that the prime symbol ($'$) in the variables means the high-frequency component.

After this, we divided the observations randomly into k groups of approximately the same size. Next, an iterative process began with k cycles in which

- The best model for each size (1 predictor, 2 predictors, …, n predictors) was identified using as a training set| the observations of all k groups except one and evaluating them according to the lowest residual sum of squares (RSS).
- We estimated the test error and stored it for each of the selected models using the observations of the group that was excluded in the previous step.
- We repeated the process k times, excluding a different group from the training set in each iteration.
- We calculated the average of the k test errors [mean squared error (MSE)] for each model size. This average value is known as the mean cross-validation test error estimate (lower values indicate better predictive accuracy).
- We identified the model size that achieves the smallest mean cross-validation test error. But, according to the principle of parsimony, if several models have similar test cross-validation error values, the simplest of them (fewer predictors) should be chosen.

4) Design of the empirical forecast model

The empirical forecast model (EMP) is based on multiple linear regression. The model was fitted for every season separately (DJF, MAM, JJA, SON). Based on the available observed data, we considered 2000–13 as the calibration period, and 2014–19 as the validation period, for all cases.

For minimum temperature, we considered an alternative formulation of the model (EMPq) that includes the specific humidity observed during the previous night (1900 LT) as a predictor. Although most of the selected stations had humidity data in the period 2000–19, some key stations such as Callao and Pisco did not have this data, so for consistency we considered 2010–13 for the calibration of the EMPq models.

We used the same equation and predictors of La Molina for the other points located in the eastern zone (Buena Vista, Socsi Cañete, and Huayan), and the same equation and predictors of Callao for the points located in the western zone (Pisco, Huarmey,

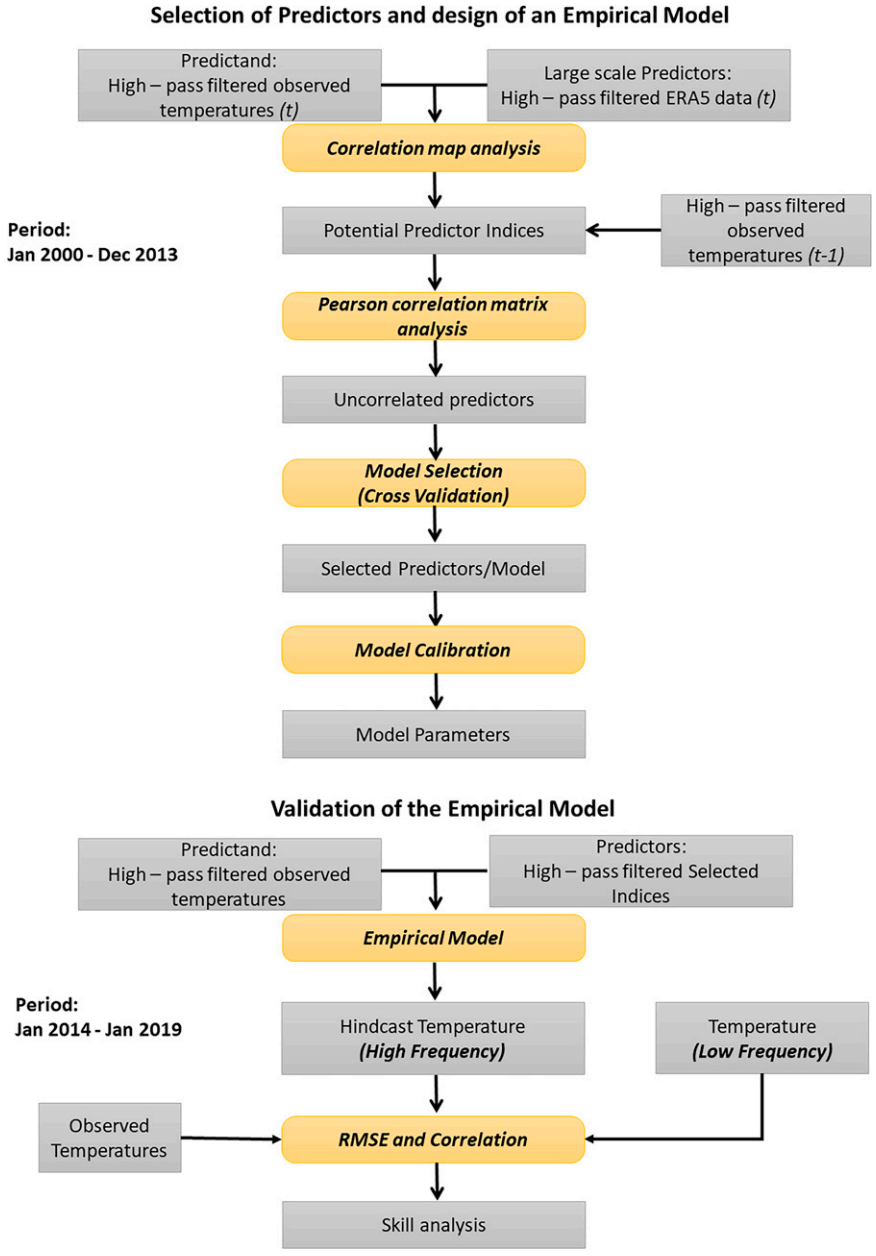

FIG. 5. Diagram of the methodology to selection of predictors and design of the empirical temperature forecast model.

Paramonga, Fonagro Chincha, and Alcantarilla). According to our experience in forecasting, the effect of synoptic variables on the Peruvian central coast is similar across the board; however, there is a marked difference between the west coast and the east coast, due to the location of the thermal inversion layer, which has a greater effect in the west and less in the east.

5) VERIFICATION OF FORECASTS

In total, we evaluated five forecast sets of Tmin and Tmax in Callao and La Molina: 1) the operational forecast of SENAMHI (operational), 2) numerical weather prediction using the operational ETA22 model (ETA22), 3) persistence of the previous day (Pers1d), 4) persistence using the mean of the previous 3 days (Pers3d), and 5) the empirical model (EM).

The skill scores used were the Pearson correlation and root-mean-squared error (RMSE).

3. Results

a. Identification of potential predictors

The annual mean of the daily spatial correlation maps between the potential predictor atmospheric variable fields

(Table S1 in the online supplemental material) and the station temperatures with the highest correlations are shown in Fig. 6. In general, Tmax has a negative correlation with meridional wind at 1000 hPa (V1000) and temperature at 850 hPa (T850) off the Peruvian coast, and with mean sea level pressure (MSLP) and meridional wind at 600 hPa (V600) over the Pacific off the Chilean coast; it has a positive correlation with temperature at 925 hPa (T925), zonal wind at 925 hPa (U925), and with specific humidity at 850 hPa (Q850) off the Peruvian coast. On the other hand, Tmin has lower values of correlation than Tmax, showing a positive correlation with specific humidity at 1000 hPa (Q1000), temperature at 1000 hPa (T1000), U925, and specific humidity at 600 hPa (Q600); it has a negative correlation with V1000 and T850 off the Peruvian coast, and meridional wind at 600 hPa (V600) and 200 hPa (V200) over the Pacific off the Chilean coast.

These relationships show why higher maximum temperatures in Lima are associated with high specific humidity and high temperature at low levels near the coast (Q850 and T925), as well as with weak southerly trade winds along the coast of Peru (V1000), which is related to a weakened SEPA (MSLP). The positive correlation between U925 and Tmax reflects the anticlockwise circulation over the Pacific associated with the SEPA, which is observed in the negative correlation with V1000. That is, if the easterly wind in the Pacific at low levels intensifies, the southern winds at low levels also intensify, and therefore the Tmax decreases.

On the other hand, higher minimum temperatures are associated with high specific humidity at medium and low levels (Q600 and Q1000), and high temperature at low levels (T1000) close to the coast. This happens because higher midlevel humidity triggers more midlevel clouds in the east zone of Lima, and so, the heat stored during the day is not lost at night and the minimum temperature increases. Furthermore, higher humidity and temperature in low levels mean increased Tmin.

The indirect relationship between the Tmin and T850 can be explained by the dynamically heated layer above the marine layer (Prohaska 1973), which is warmer if the subsidence is larger and, for that, the inversion is more intense and the temperature over the surface is cooler. On the other hand, the meridional wind at medium (V600) and high levels (V200) near the Chilean coast would reflect the formation of troughs or COLs with cyclonic circulation and that would generate an increase of both Tmin and Tmax.

The annual mean of the daily spatial correlation maps show that the lowest correlation corresponds to austral winter (JJA) and the highest correlation are in summer (DJF) and fall (MAM) in all the cases (Figs. S2–S5).

After a spatial correlation, we added local variables of a previous day, which were called persistence. Thus, we had in total five predictors for Tmin in Callao, six for Tmax in Callao, eight for Tmin in La Molina, and seven predictors for Tmax in La Molina (Table S2).

b. Selection of predictors

After the Pearson correlation matrix, with which we removed highly correlated potential predictors (Fig. S5); we applied the cross-validation method, with which we identified the smallest model size that has the smallest mean cross-validation test error. In Fig. 7 we show the better model size selected in each case (red point), and the predictors in each size were selected according to the lowest RSS considering all observations. The results of selected predictors for each predictand are in Table 1.

c. Empirical forecast model

The proposed empirical forecast models are defined as follows:

$$T_t^{\text{pred}} = \beta_o + \beta_1 T'^{\text{obs}}_{t-1} + \sum_{i=1}^{N} \beta_{i+1} X'_{i,t} + \overline{T}_t^{\text{obs}}, \tag{2}$$

$$T_t^{\text{pred}} = \beta_o + \beta_1 T'^{\text{obs}}_{t-1} + \beta_2 q'^{\text{obs}}_{t-1} + \sum_{i=1}^{N} \beta_{i+2} X'_{i,t} + \overline{T}_t^{\text{obs}}, \tag{3}$$

where T_t is the temperature for the time t, X_t is the large-scale predictor i, N is the total number of large-scale predictors, and β are the parameters for the regression model. The overbar in Eqs. (2) and (3) indicates the low-frequency signal of the variable and the prime ($'$) indicates the high-frequency signal. The second regression model in Eq. (3) has the predictor q_{t-1}, which is the specific humidity of the previous night and was used only for the minimum temperature of La Molina.

To have empirical models for 2- and 3-day lead forecasts, additional calibrations were performed considering the same predictors for day $t - 1$ but using the observed temperatures for days $t + 1$ and $t + 2$ as predictands.

d. Forecast verification

Results of the empirical model, as well as of the other types of forecasts, were contrasted against the observed data between 2014 and 2019. Figures 8–11 show Pearson correlation and RMSE results between the high-frequency of all forecasts and the observed data at the annual and seasonal scale.

In most cases, the forecasts based on the empirical model present the highest correlations followed by the Pers1d forecast and the lowest correlations are found for the ETA22 forecast. Therefore, it is possible that persistence performs better than the SENAMHI forecast. At the seasonal level, correlations in austral autumn are higher for all cases and the lowest correlation results are found in austral winter.

Correlations between the empirical model and the observed data are acceptable for both Tmax and Tmin (Figs. 8 and 9). The correlations between the observed and predicted high-frequency series of Tmax are in the range of 0.38–0.53 for the empirical model, while the correlations with the operational forecast show values between 0.02 and 0.21. In the case of Tmin, correlations are in the range of 0.28–0.52 for the empirical model. However, the performance improves notably with EMPq in the majority of the stations. A possible explanation for this improvement is that some stations such as La Molina are further from the sea and surrounded by hills. As a consequence, the daily Tmin could be more associated with local-scale factors, such as geography and specific humidity.

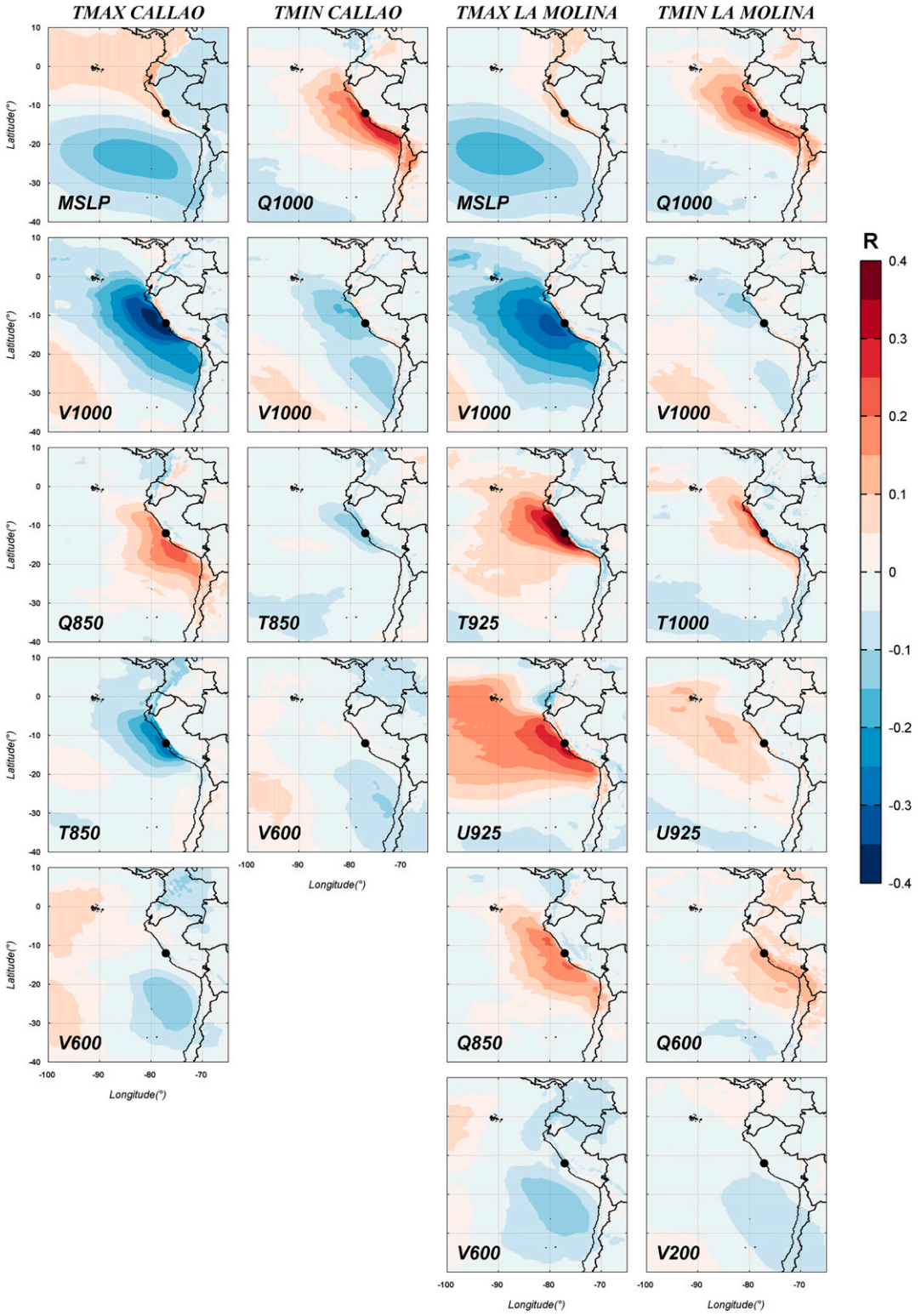

FIG. 6. Annual correlation maps between the station (dot) temperature and the large-scale variables. Seasonal and lower-frequency variabilities were removed in all datasets.

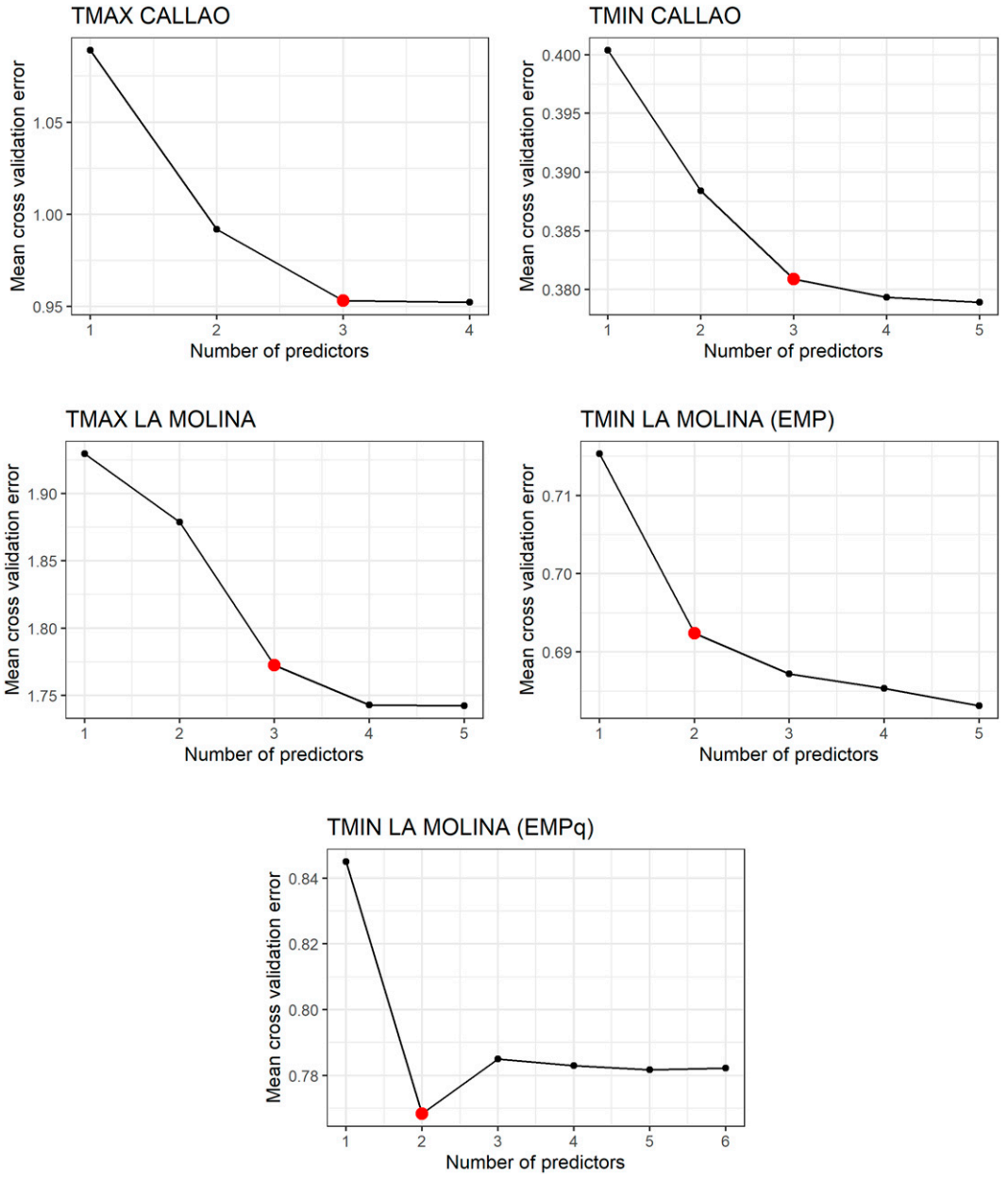

FIG. 7. Mean cross-validation error for different model sizes (number of predictors) for each predictand, using data of the calibration period. Selected model sizes are in red.

An analysis of correlation between the Tmin^{obs} and the Q^{obs} at 1900 LT of the previous day in La Molina (Fig. S1) suggests that lower Tmin values are associated with drier local conditions. This kind of analysis might be relevant to select local-scale predictors which could contribute to the skill of empirical models.

Results of RMSE show that EMP has smaller errors than the ETA22. In general, forecasts based on empirical models for La Molina present larger errors compared to Callao. Meanwhile Tmax presents the largest errors of any variable in both stations (values of RMSE greater than 0.76°C), with La Molina having the most remarkable errors for EMP (Fig. 10). In contrast, Tmin shows a slight improvement in La Molina for EMPq. In the case of Callao, the Tmin has the smallest errors (Fig. 11).

Figure 12 presents results of correlation and RMSE for the forecasts of days t, $t + 1$, and $t + 2$. For the empirical model, correlations decline for days $t + 1$ and $t + 2$ but these are still higher than the other forecasts. The ETA22 model has the lowest correlations for day t but maintains its performance for days $t + 1$ and $t + 2$, being superior to the persistence and operational forecast on those lead times. Regarding the RMSE, results indicate increases for days $t + 1$ and $t + 2$ in most of the forecasts. The empirical model presents the lowest RMSE

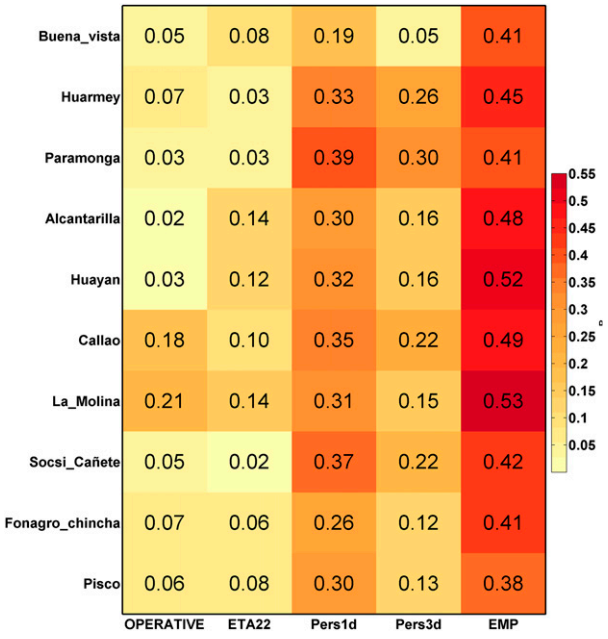

FIG. 8. Annual Pearson correlations between observed maximum temperatures and the results of forecast methods (x axis) for 10 stations (y axis) using data of validation period. Seasonal and lower-frequency variabilities were removed in all datasets.

values in all cases, while the ETA22 model has the highest values in most cases. The largest errors are seen for Tmax-La Molina which are above 1.4°C, while Tmin-Callao presents the lowest values which are in the range of 0.6°–1.0°C.

To evaluate the improvement of the empirical model against the SENAMHI operational forecast of daily temperatures in La

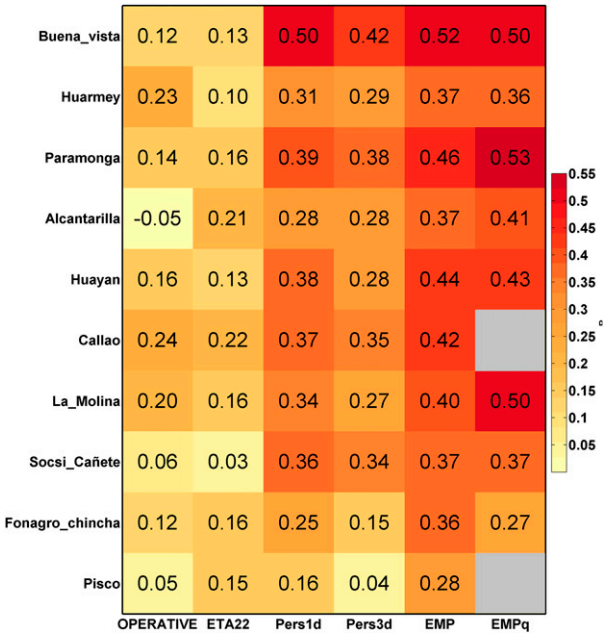

FIG. 9. As in Fig. 8, but for minimum temperature.

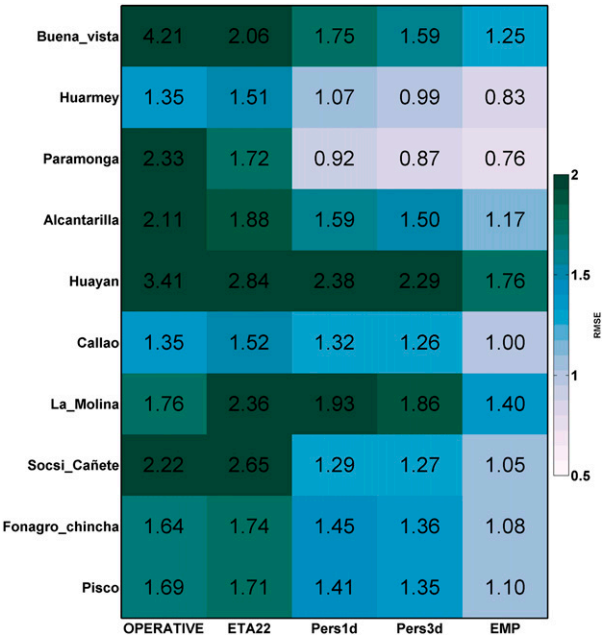

FIG. 10. As in Fig. 8, but for RMSE.

Molina and Callao, Fig. 13 clearly shows the largest errors in the operational forecast.

4. Discussion

Empirical models were developed for 1–3-day lead forecasts of maximum and minimum temperature for two locations in the city of Lima (central coast of Peru at 12°S), and the results

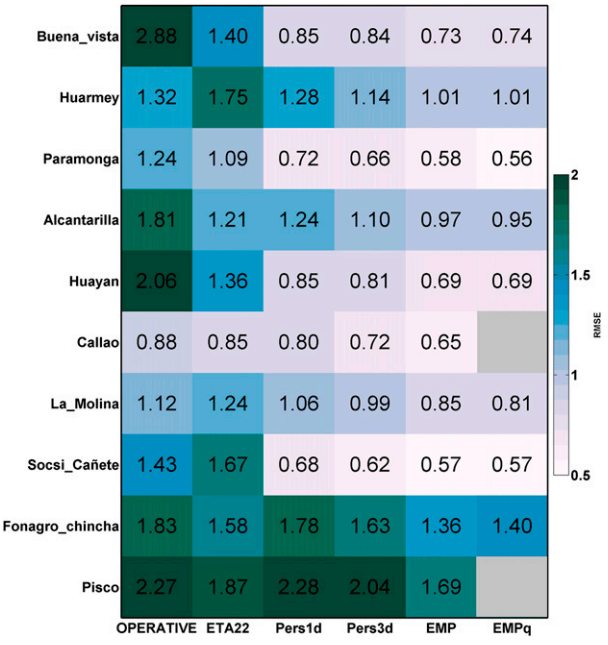

FIG. 11. As in Fig. 9, but for RMSE.

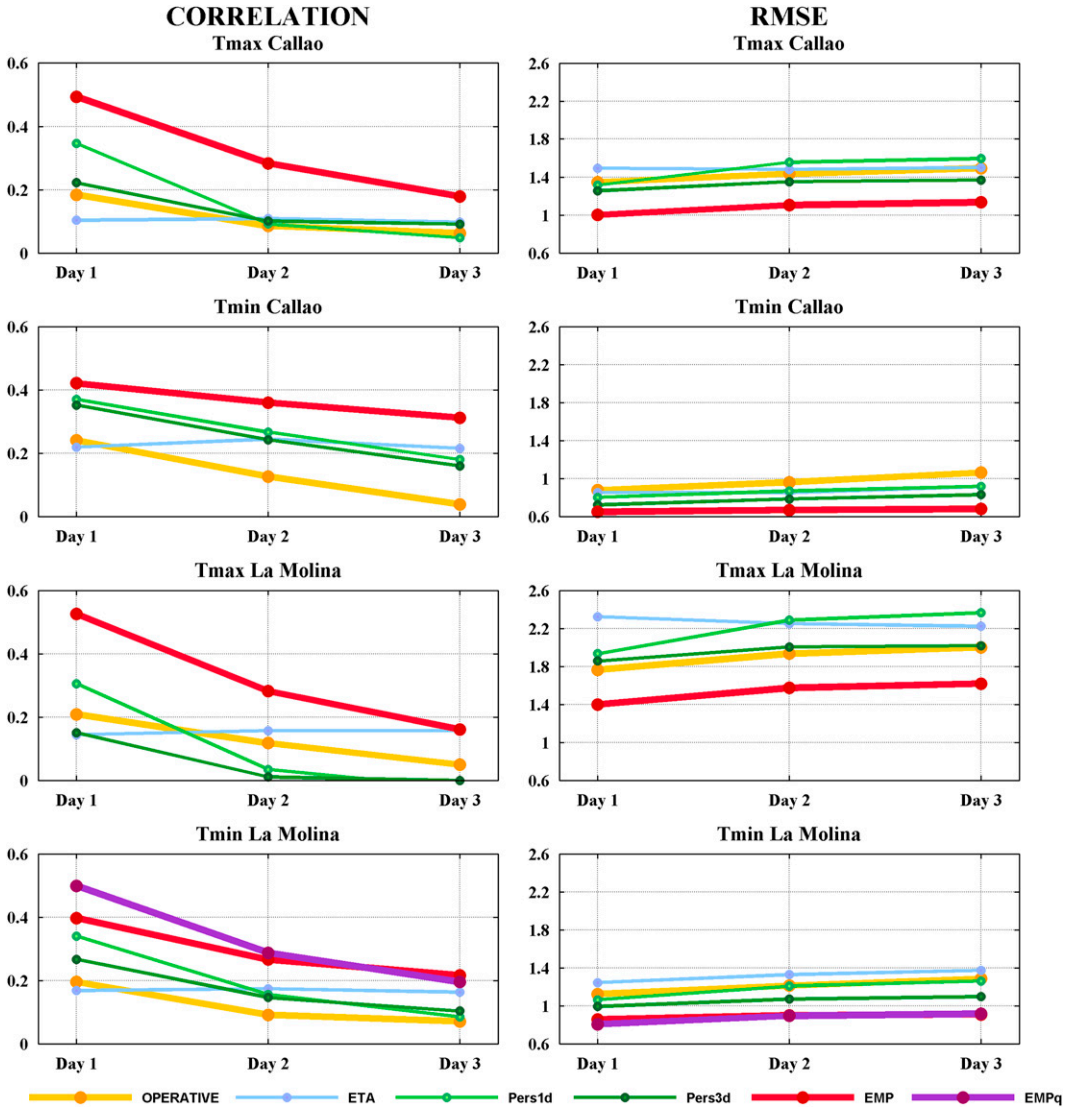

FIG. 12. (left) Correlation and (right) RMSE between observed temperature and the five forecast methods for 1–3-day lead forecast.

were compared with the operational forecasts issued by the National Meteorological and Hydrological Service—SENAMHI and the output of the operational regional numerical atmospheric model (ETA22) of SENAMHI.

The results of the empirical model have better results with the highest Pearson correlations and the lowest RMSE values for all predictands: Tmax-Callao, Tmin-Callao, Tmax-La Molina, and Tmin-La Molina (EMP and EMPq), showing a correct selection of both atmospheric and local predictors.

Results of high-frequency correlation coefficients and root-mean-square error reveal the empirical models have the best performances compared to the operational forecast and numerical model. The highest annual correlations for the 1-lead day operational forecast were obtained by the empirical model with values of around 0.38–0.53 for Tmax and 0.28–0.52 for Tmin. Annual values of RMSE are smaller with the empirical models, around 0.76°–1.76°C and 0.57°–1.69°C for Tmax and Tmin, respectively. The skill of the model is lower in the austral winter, which would indicate the low contribution of large-scale predictors during this period. Therefore, it is necessary to consider more local-scale predictors that could better contribute to the forecast. For example, the observed specific humidity was considered as a local scale predictor. As a result of this, the performance of the empirical model improved substantially in all seasons and especially in austral winter. We observed that the empirical forecast, on average for the 10 forecast points, considering the specific humidity as a predictor, improves by 47% compared to the operational forecast of SENAMHI, and by 5%, compared to the empirical forecast without considering the specific humidity.

On the other hand, the performance of the empirical models declines as lead time increases; however, it is still more skillful than the other forecasts. Although the decreases in correlations and increases

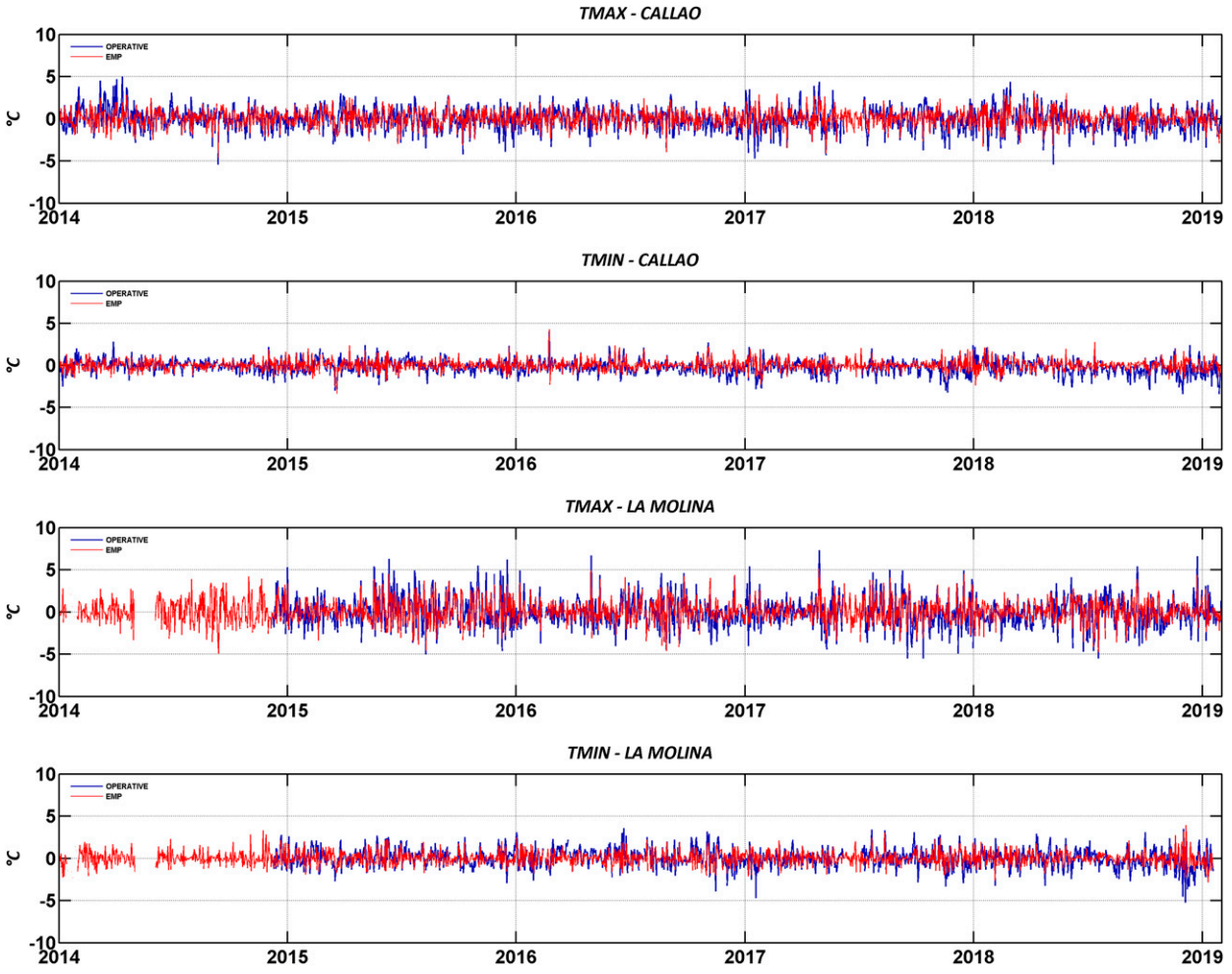

FIG. 13. Daily variation of the errors of SENAMHI operational forecast (blue) and of the empirical model (red) from January 2014 to January 2019.

in RMSE for the 2- and 3-day forecast suggest a decline in performance of the empirical model, it is still more skillful than the other forecasts. Furthermore, it should be noted that large-scale predictors for day 1 were maintained, but these results could be improved if numerical forecasts for 2–3-day lead, for example from the NOAA Global Forecast System (GFS), were used for the large-scale predictors in the empirical model, i.e., applying the model output statistics (MOS; Glahn and Lowry 1972) approach, in which large-scale predictors from models such as the GFS will be tested as predictors in order to improve the accuracy of these forecasts.

Apart from the analyzed predictors, the weather forecasters state that cloud cover affects the behavior of temperatures, mainly in the eastern part of the Peruvian coast. A cloudy day means lower maximum temperature, due to less shortwave radiation reaching the surface; in contrast, a clear night means lower minimum temperature, due to the rapid loss of heat gained during the day.

Just as there are cases in which the empirical model is much better than the operational forecast, there are situations in which neither of them performs well, generating an error of more than 3°C in both cases (Fig. 14). In 33% of these cases, a low-level cyclonic vortex (VCNB) (SENAMHI 2021) occurs, such as the one that formed on 24 December 2018 (Fig. 15), which generated cloudy coverage throughout the day and generated an error in the forecast of Tmax of 6.6° and 4.4°C, for the operational and empirical model, respectively, in La Molina.

As long as the VCNB generates a layer of clouds in the coastal strip of central Peru, regardless of where its core is located, forecast temperatures will be overestimated, such as what happened on 24 December 2018 (Fig. 15). On the other hand, there are isolated situations in which the formation of the vortex makes it possible to clear the central coastal strip and generates an underestimation of the forecast, such as what happened on 17 September 2016 and 25 June 2018 (Fig. 14).

This vortex is a mesoscale system formed by the low-level confluence of south and north winds over the coast of Lima, and it is possible that the topography of the central coast of Peru is responsible for the formation of this vortex; it is typically found between May and December. We recommend

	Callao						La Molina					
	Date	Observed (O)	SENAMHI Forecast (S)	Empiric Model (EM)	S-O	EM-O	Date	Observed (O)	SENAMHI Forecast (S)	Empiric Model (EM)	S-O	EM-O
Tmax	30/03/2011	20.6	25	24.6	4.4	4.0	01/01/2015	22.2	27.5	25.6	5.3	3.4
	26/07/2011	22.1	19	18.7	-3.6	-3.4	28/04/2015	22.8	26.5	25.9	3.7	3.1
	25/03/2012	23.6	27	27.3	3.4	3.7	19/05/2015	21.6	27	25.5	5.4	3.9
	12/09/2014	24.4	19	19.8	-5.4	-4.6	22/05/2015	21.5	25	24.5	3.5	3.0
	08/10/2015	25.2	21	22.1	-4.2	-3.1	02/06/2015	20.4	25	23.7	4.6	3.3
	28/08/2016	22.9	19.5	19.0	-3.4	-3.9	12/06/2015	19.2	25.5	23.7	6.3	4.5
	09/03/2017	33.4	30	30.0	-3.4	-3.4	20/07/2015	23.1	19.5	19.9	-3.6	-3.2
	10/05/2017	30.3	26	26.3	-4.3	-4.0	29/07/2015	24.3	21	20.5	-3.3	-3.8
	07/04/2018	23	26	26.3	3.0	3.3	07/08/2015	17.2	20.5	20.9	3.3	3.7
	08/05/2018	26.4	21	23.3	-5.4	-3.1	10/08/2015	23.5	18.5	18.9	-5.0	-4.6
							27/08/2015	23.7	20	20.5	-3.7	-3.2
							28/08/2015	25.6	22.5	22.4	-3.1	-3.2
							16/09/2015	18.9	22.5	22.6	3.6	3.7
							31/10/2015	19.5	22.5	22.6	3.0	3.1
							27/11/2015	20.5	25	23.7	4.5	3.2
							07/12/2015	21.1	26	24.3	4.9	3.2
							09/12/2015	29.6	26	26.5	-3.6	-3.1
							18/12/2015	21.8	28	25.1	6.2	3.3
							07/01/2016	24.6	29.5	28.0	4.9	3.4
							30/04/2016	22.3	29	27.2	6.7	4.9
							14/05/2016	21.6	26	25.5	4.4	3.9
							23/06/2016	17	21	21.1	4.0	4.1
							29/06/2016	17.5	21	20.6	3.5	3.1
							25/08/2016	24	20.5	20.4	-3.5	-3.6
							28/08/2016	24.4	20	19.8	-4.4	-4.6
							30/08/2016	16.5	20	19.5	3.5	3.0
							17/09/2016	24.7	21.5	20.6	-3.2	-4.1
							18/09/2016	25.7	22	22.2	-3.7	-3.5
							25/09/2016	17.9	22.5	22.2	4.6	4.3
							15/12/2016	23.8	28	28.2	4.2	4.4
							29/04/2017	20.7	28	25.8	7.3	5.1
							10/05/2017	30.5	27	27.3	-3.5	-3.2
							27/05/2017	24.8	21	21.4	-3.8	-3.4
							19/07/2017	16.8	20	20.6	3.2	3.8
							24/07/2017	17.3	22	21.1	4.7	3.8
							10/08/2017	16.9	20	20.3	3.1	3.4
							20/08/2017	15	20	19.0	5.0	4.0
							05/09/2017	22.3	19	18.8	-3.3	-3.5
							02/10/2017	24.5	21	21.2	-3.5	-3.3
							12/10/2017	25.5	20	22.4	-5.5	-3.1
							04/11/2017	18.9	22	22.1	3.1	3.2
							05/12/2017	20.6	24	24.0	3.4	3.4
							16/12/2017	20.1	25	24.6	4.9	4.5
							29/12/2017	22.1	26	25.3	3.9	3.2
							25/06/2018	21.7	17	18.1	-4.7	-3.6
							11/07/2018	22.6	18	18.5	-4.6	-4.1
							13/07/2018	23.4	20	20.3	-3.4	-3.1
							19/07/2018	22.5	17	17.8	-5.5	-4.7
							01/09/2018	22.6	19	19.1	-3.6	-3.5
							19/09/2018	16.6	22	20.5	5.4	3.9
							03/10/2018	26	23	22.6	-3.0	-3.4
							24/12/2018	21.4	28	25.8	6.6	4.4
Tmin	23/02/2016	19	23	23.3	4.0	4.3	12/08/2017	10.7	14	13.8	3.3	3.1
							02/12/2018	13.5	17	16.8	3.5	3.3
							06/12/2018	12.8	16	17.2	3.2	4.4

FIG. 14. Dates with errors of more than 3°C, both in the empirical model and the operational forecast for the validation period. Orange colors indicate underestimation of the observed temperature; blue colors indicate overestimation of the observed temperature. Dates marked in purple represent days with identified VCNBs.

deepening the studies in this regard, considering the automatic identification of VCNB.

The ETA-SENAMHI model is capable of simulating the occurrence of low-level cyclonic vortex events. One perspective emerging from this research is to perform numerical model experiments at finer resolutions in order to improve pattern simulation accuracy.

On the other hand, Fig. 14 shows that in 67% of cases where the error is greater than 3°C in both cases (operational forecast and empirical model) do not correspond to the presence of the VCNB, and instead, we found that the anomalies of 1000-hPa wind parallel to the Peruvian coast in the most extreme cases, where the absolute error is greater 5°C, were more intense from the north (underestimation) or south (overestimation). The meaning of this finding could be the objective of a future study.

5. Conclusions

Empirical models were developed for 1–3-day lead forecasts of maximum and minimum temperature for 10 locations in the city of Lima (central coast of Peru at 12°S), and the results were

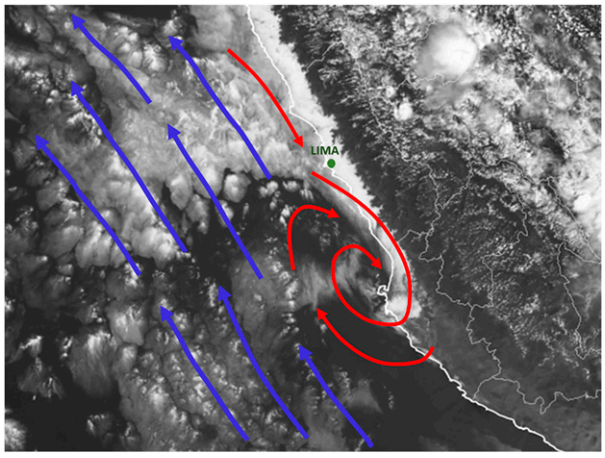

FIG. 15. Low-level cyclonic vortex (VCNB, for its acronym in Spanish) at 1100 LT 24 Dec 2018. Red arrows represent the mesoscale flux responsible for the formation of the VCNB. Blue arrows represent the synoptic southerlies in the Pacific near the Peruvian central coast. The green point represents the location of Lima.

compared with the operational forecasts issued by the National Meteorological and Hydrological Service—SENAMHI and the output of the operational regional numerical atmospheric model (ETA22) of SENAMHI.

The empirical model has better results with the highest Pearson correlations and the lowest RMSE values for all predictands, showing a correct selection of both atmospheric and local predictors.

Acknowledgments. We extend our gratitude to the SENAMHI's team of forecasters, since they helped with the identification of the main synoptic systems for weather forecasting in Lima, as well as participating in the identification of the VCNB during 2011–18; and the support from the Meteo-Huascarán project (Prociencia 036-2021), that allows the development of science in Peru.

Data availability statement. Datasets analyzed during the current study are available at this link: https://github.com/GustavoDelaCruz/TemperatureForecastLima.

REFERENCES

Albrecht, J. A., 1981: The twenty-year average atmospheric structure at Lima, Peru. M.S. thesis, Dept. of Meteorology, Florida State University, 94 pp.

Arellano Rojas, C. S., 2013: Condições meteorológicas e níveis de poluição na região metropolitana de Lima—Perú. M.S. in Meteorology thesis, Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo, 148 pp.

Black, T. L., 1994: The new NMC Mesoscale Eta Model: Description and forecast examples. *Wea. Forecasting*, **9**, 265–278, https://doi.org/10.1175/1520-0434(1994)009<0265:TNNMEM>2.0.CO;2.

Bosart, L. F., 2003: Whither the weather analysis and forecasting process? *Wea. Forecasting*, **18**, 520–529, https://doi.org/10.1175/1520-0434(2003)18<520:WTWAAF>2.0.CO;2.

Cai, W., and Coauthors, 2020: Climate impacts of the El Niño–Southern Oscillation on South America. *Nat. Rev. Earth Environ.*, **1**, 215–231, https://doi.org/10.1038/s43017-020-0040-3.

Capotondi, A., A. T. Wittenberg, J.-S. Kug, K. Takahashi, and M. J. McPhaden, 2020: ENSO diversity. *El Niño Southern Oscillation in a Changing Climate*, *Geophys. Monogr.*, Vol. 253, Amer. Geophys. Union, 65–86.

Dewitte, B., S. Illig, L. Renault, K. Goubanova, K. Takahashi, D. Gushchina, K. Mosquera, and S. Purca, 2011: Modes of covariability between sea surface temperature and wind stress intraseasonal anomalies along the coast of Peru from satellite observations (2000–2008). *J. Geophys. Res.*, **116**, C04028, https://doi.org/10.1029/2010JC006495.

Doswell, C. A., 2004: Weather forecasting by humans—Heuristics and decision making. *Wea. Forecasting*, **19**, 1115–1126, https://doi.org/10.1175/WAF-821.1.

Echevin, V., F. Colas, D. Espinoza-Morriberon, L. Vasquez, T. Anculle, and D. Gutierrez, 2018: Forcings and evolution of the 2017 coastal El Niño off northern Peru and Ecuador. *Front. Mar. Sci.*, **5**, 367, https://doi.org/10.3389/fmars.2018.00367.

Enfield, D. B., 1981: Thermally driven wind variability in the planetary boundary layer above Lima, Peru. *J. Geophys. Res.*, **86**, 2005–2016, https://doi.org/10.1029/JC086iC03p02005.

Espinoza, J. C., J. Ronchail, M. Lengaigne, N. Quispe, Y. Silva, M. L. Bettolli, G. Avalos, and A. Llacza, 2013: Revisiting wintertime cold air intrusions at the east of the Andes: Propagating features from subtropical Argentina to Peruvian Amazon and relationship with large-scale circulation patterns. *Climate Dyn.*, **41**, 1983–2002, https://doi.org/10.1007/s00382-012-1639-y.

Febre Pérez, C. J., 2018: Estudio de la dinámica y patrones de circulación media del jet costero de bajos niveles en la costa central y sur del Perú. Meteorological Engineering thesis, Facultad de Ciencias, Universidad Nacional Agraria La Molina, 282 pp.

Garreaud, R., 1999: Multiscale analysis of the summertime precipitation over the central Andes. *Mon. Wea. Rev.*, **127**, 901–921, https://doi.org/10.1175/1520-0493(1999)127<0901:MAOTSP>2.0.CO;2.

——, 2018: A plausible atmospheric trigger for the 2017coastal El Niño. *Int. J. Climatol.*, **38**, e1296–e1302, https://doi.org/10.1002/joc.5426.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211, https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Huaman, L., C. Schumacher, and G. Kiladis, 2020: Eastward-propagating disturbances in the tropical Pacific. *Mon. Wea. Rev.*, **148**, 3713–3728, https://doi.org/10.1175/MWR-D-20-0029.1.

Illig, S., and Coauthors, 2014: Forcing mechanisms of intraseasonal SST variability off central Peru in 2000–2008. *J. Geophys. Res. Oceans*, **119**, 3548–3573, https://doi.org/10.1002/2013JC009779.

Janjić, Z. I., 1984: Nonlinear advection schemes and energy cascade on semi-staggered grids. *Mon. Wea. Rev.*, **112**, 1234–1245, https://doi.org/10.1175/1520-0493(1984)112<1234:NASAEC>2.0.CO;2.

Klein, S. A., and D. L. Hartmann, 1993: The seasonal cycle of low stratiform clouds. *J. Climate*, **6**, 1587–1606, https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2.

L'Heureux, M. L., and Coauthors, 2017: Observing and predicting the 2015/16 El Niño. *Bull. Amer. Meteor. Soc.*, **98**, 1363–1382, https://doi.org/10.1175/BAMS-D-16-0009.1.

Mesinger, F., 1984: A blocking technique for representation of mountains in atmospheric models. *Riv. Meteor. Aeronaut.*, **44**, 195–202.

——, Z. I. Janjic, S. Nickovic, D. Gavrilov, and D. G. Deaven, 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of an Appalachian redevelopment. *Mon. Wea. Rev.*, **116**, 1493–1518, https://doi.org/10.1175/1520-0493(1988)116<1493:TSMCMD>2.0.CO;2.

Novak, D. R., C. Bailey, K. F. Brill, P. Burke, W. A. Hogsett, R. Rausch, and M. Schichtel, 2014: Precipitation and temperature forecast performance at the Weather Prediction Center. *Wea. Forecasting*, **29**, 489–504, https://doi.org/10.1175/WAF-D-13-00066.1.

Peng, Q., S. P. Xie, D. Wang, X. T. Zheng, and H. Zhang, 2019: Coupled ocean-atmosphere dynamics of the 2017 extreme coastal El Niño. *Nat. Commun.*, **10**, 298, https://doi.org/10.1038/s41467-018-08258-8.

Pinche Laurre, C., 1986: Estudio de las condiciones climáticas y de la niebla en la costa norte de Lima. Meteorological Engineering thesis, Facultad de Ciencias, Universidad Nacional Agraria La Molina, 166 pp.

Prohaska, F. J., 1973: New evidence on the climatic controls along the Peruvian coast. *Coastal Deserts: Their Natural and Human Environments*, D. H. K. Ameran and A. Wilson, Eds., University of Arizona Press, 91–107.

Quispe-Gutiérrez, N., and Coauthors, 2021: Cutoff low over the southeastern Pacific Ocean: A case study. *J. South. Hemisphere Earth Syst. Sci.*, **71**, 17–29, https://doi.org/10.1071/ES19051.

Quispe Vega, K. R., 2017: Patrones sinópticos de una DANA asociados a la ocurrencia de nevadas en la sierra central y sur del Perú. Meteorological Engineering thesis, Facultad de Ciencias, Universidad Nacional Agraria La Molina, 105 pp.

Rahn, D., and R. Garreaud, 2013: A synoptic climatology of the near-surface wind along the west coast of South America. *Int. J. Climatol.*, **34**, 780–792, https://doi.org/10.1002/joc.3724.

Rivas Quispe, P. R., 2019: Identificación de patrones de circulación característicos en la formación de precipitación en Lima Metropolitana y Callao durante 1980–1995. Meteorological Engineering thesis, Facultad de Ciencias, Universidad Nacional Agraria La Molina, 187 pp.

Rodwell, M. J., and B. J. Hoskins, 2001: Subtropical anticyclones and summer monsoons. *J. Climate*, **14**, 3192–3211, https://doi.org/10.1175/1520-0442(2001)014<3192:SAASM>2.0.CO;2.

Rollenbeck, R., and J. Bendix, 2011: Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. *Atmos. Res.*, **99**, 277–289, https://doi.org/10.1016/j.atmosres.2010.10.018.

SENAMHI, 2021: Análisis del Vórtice Ciclónico de Niveles Bajos (VCNB) en la costa central Peruana. Informe Técnico, 14 pp., https://hdl.handle.net/20.500.12542/794.

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.

——, and D. S. Battisti, 2007: Processes controlling the mean tropical Pacific precipitation pattern. Part I: The Andes and the eastern Pacific ITCZ. *J. Climate*, **20**, 3434–3451, https://doi.org/10.1175/JCLI4198.1.

——, and A. G. Martínez, 2017: The very strong coastal El Niño in 1925 in the far-eastern Pacific. *Climate Dyn.*, **52**, 7389–7415, https://doi.org/10.1007/s00382-017-3702-1.

——, A. Montecinos, K. Goubanova, and B. Dewitte, 2011: ENSO regimes: Reinterpreting the canonical and Modoki El Niño. *Geophys. Res. Lett.*, **38**, L10704, https://doi.org/10.1029/2011GL047364.

——, and Coauthors, 2018: The 2017 coastal El Niño [in "State of the Climate in 2017"]. *Bull. Amer. Meteor. Soc.*, **99** (8), S210–S211, https://doi.org/10.1175/2018BAMSStateoftheClimate.1.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. International Geophysics Series, Vol. 100, Academic Press, 648 pp.