Predictability of December–April Rainfall in Coastal and Andean Ecuador

G. CRISTINA RECALDE-CORONEL

National Institute of Meteorology and Hydrology, and Observatorio Latinoamericano de Eventos Extraordinarios, and Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador

ANTHONY G. BARNSTON

International Research Institute for Climate and Society, Palisades, New York

ÁNGEL G. MUÑOZ

International Research Institute for Climate and Society, Palisades, New York, and Centro de Modelado Científico, Universidad del Zulia, and Observatorio Latinoamericano de Eventos Extraordinarios, Maracaibo, Venezuela

(Manuscript received 15 April 2013, in final form 3 February 2014)

ABSTRACT

In Ecuador, forecasts of seasonal total rainfall could mitigate both flooding and drought disasters through warning systems if issued at useful lead time. In Ecuador, rainfall from December to April contributes most of the annual total, and it is crucial to agricultural and water management. This study examines the predictive skill for February–April and December–February seasonal rainfall totals using statistical and dynamical approaches. Fields of preceding observed sea surface temperature (SST) are used as predictors for a purely statistical prediction, and predictions of an atmospheric general circulation model (AGCM) are used as predictors with a model output statistics correction design using canonical correlation analysis. For both periods, results indicate considerable predictive skill in some, but not all, portions of the Andean and especially coastal regions. The skill of SST and AGCM predictors comes mainly through skillful rainfall anomaly forecasts during significant ENSO events. Atlantic Ocean SST plays a weaker predictive role. For the simultaneous diagnostic highest skill is obtained using the eastern Pacific Ocean domain, and for time-lagged forecasts highest scores are found using the global tropical ocean domain. This finding suggests that, while eastern Pacific SST is what matters most to Ecuadorian rainfall, at sufficient lead time these local SSTs become most effectively predicted using basinwide ENSO predictors. In Ecuador's coastal region, and in some parts of the Andean highlands, skill levels are sufficient for warning systems to reduce economic losses associated with flood and drought. Accordingly, the Instituto Nacional Meteorologia e Hidrologia of Ecuador issues forecasts each month using methods described here—also implemented by countries of the Latin American Observatory partnership, among other South American organizations.

1. Introduction

a. Ecuador climate

Ecuador is located in northwestern South America, between Peru and Colombia, within 1.5°N–3.4°S, 75.2°–81.0°W. The Galapagos Islands, approximately 1000 km to the west, are part of Ecuador but are not included in this study. Intersected by the equator and by the Andes Mountains, Ecuador has a complex topography and a variety of regional climates and subregional microclimates. Four natural regions are identified: the coast, the highlands (Andes), the Oriente (Amazon), and the Galapagos Islands. Each region has its own climatological rainy season: For the coast and the Galapagos it occurs from late December through May, in the highlands it runs from September through April or May, and in the Amazon it rains throughout the year (Cañadas 1983) with the wettest (driest) months being April–July (September and October). Here, we examine the levels and sources of predictive skill both for the February–April (FMA) season, because it represents a major portion of the main rainy seasons in all Ecuadorian regions, and for the December–February (DJF) season,

Corresponding author address: Anthony Barnston, Climate Group, International Research Institute for Climate and Society, Monell Bldg., 61 Rte. 9W, Palisades, NY 10964.
E-mail: tonyb@iri.columbia.edu

DOI: 10.1175/JAMC-D-13-0133.1

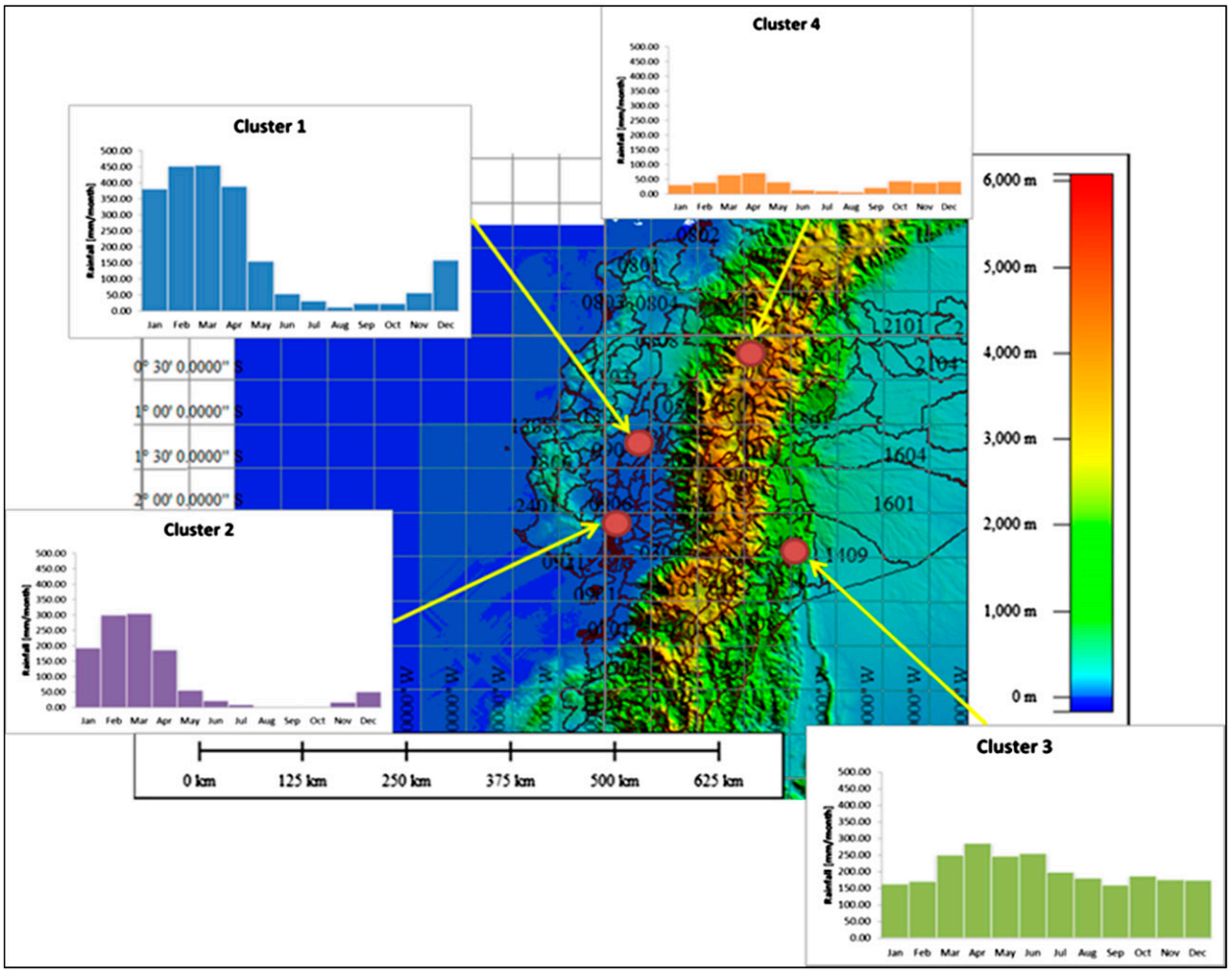

FIG. 1. Representation of the climatological monthly rainfall distribution for each of four clusters using the stations Pichilingue (1.10°S, 79.46°W; cluster 1 for coast; wetter), José Joaquín de Olmedo International Airport in Guayaquil (2.15°S, 79.88°W; cluster 2 for coast; less wet), Edmundo Carvajal Airport in Macas (2.29°S, 78.11°W; cluster 3 for Amazon), and El Quinche-Pichincha (0.10°S, 78.30°W; cluster 4 for highlands).

because it also brings significant rainfall in the economically critical coastal regions. In addition, the DJF season is better timed to take advantage of the predictable effects of El Niño–Southern Oscillation (ENSO), because ENSO episodes are often near or just past their peak in DJF but may have significantly dissipated by FMA.

In an unpublished cluster-analysis study conducted previously by the lead author, using monthly precipitation data for 121 stations covering Ecuador for the 1971–2010 period, four clusters were found to capture best the seasonality of Ecuador's rainfall. Selecting one station to represent the typical rainfall for each cluster (Fig. 1), clusters 1 and 2 (both in the coastal region) are seen to have unimodal rainfall patterns with the same wet season, with cluster 1 receiving higher rainfall amounts. Cluster 3 (in the Amazon) receives more uniform precipitation throughout the year, and cluster 4 (highlands) has a bimodal rainfall pattern, with less total annual rainfall than the other clusters.

b. Interannual variability: ENSO and seasonal forecasting

The influence of the ENSO cycle (Diaz and Markgraf 1992) on rainfall has been studied in numerous regions of the world, and Ecuador is among those countries that are most strongly affected by ENSO (Rossel et al. 1999). ENSO most strongly modulates precipitation and temperature patterns in the Ecuadorian coastal region. During El Niño (La Niña), there are higher-than-average (lower than average) sea surface temperatures (SST) in the central and eastern equatorial Pacific Ocean and above-average (below average) rainfall in most of

Ecuador west of the highlands. The flooding in the Ecuadorian coastal region associated with El Niño can cause considerable material loss and deaths, whereas La Niña can result in drought, which has slower-acting but also potentially disastrous consequences. Long-term preventive measures, such as irrigation systems, could be implemented in preparation for multiyear La Niña episodes.

The equatorial Andes tend to experience below-normal precipitation during El Niño episodes because of an anomalous Hadley cell that inhibits convection over the high terrain (Vuille et al. 2000a,b; Francou et al. 2004). The Andean signal is opposite to that observed just a few hundred kilometers to the west, where El Niño brings excessive rain over the normally arid coastal area (e.g., Horel and Cornejo-Garrido 1986). The tropical Andes also experience significant warming (cooling) of up to 1°C during El Niño (La Niña) years (Francou et al. 2004). Rainfall variability over the eastern side of the Andes is not appreciably governed by ENSO but has been associated with a north–south dipole-like SST pattern over the tropical Atlantic Ocean (Vuille et al. 2000a,b).

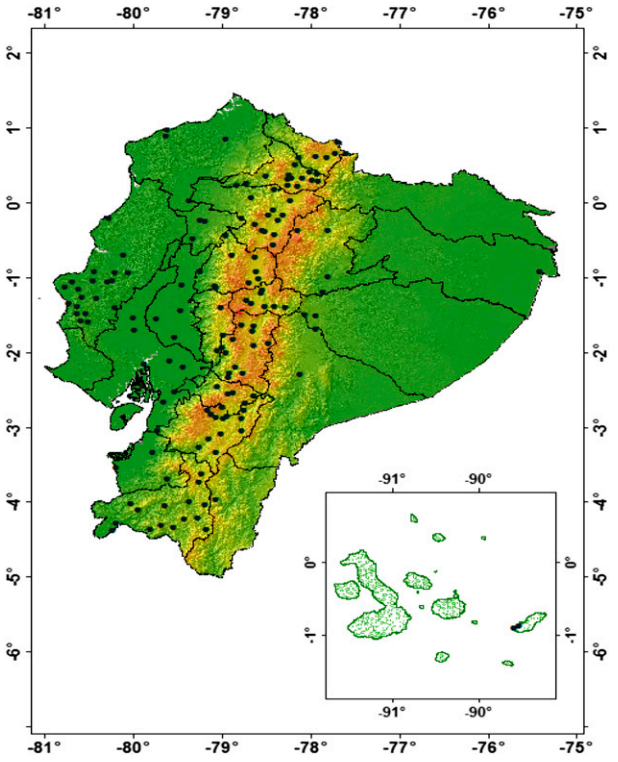

FIG. 2. Distribution of rainfall stations used in this study, following QC and homogenization. The Galapagos stations are omitted, resulting in 149 stations. The topography map comes from the International Maize and Wheat Improvement Center at the Universidad del Azuay (http://www.uazuay.edu.ec/promsa/ecuador.htm).

2. Data and methods

The study is conducted in steps. First, the monthly precipitation data for Ecuador are quality controlled, checking for incorrect (or suspicious) readings, and the data are subjected to a homogenization process in which artifacts related to instrument changes and/or station relocations are identified and the data may be adjusted. An objective analysis is then applied to derive a gridded rainfall dataset. This gridded rainfall is the predictand data to which canonical correlation analysis (CCA) is applied to examine the strongest rainfall predictor patterns, based on either observed SSTs or atmospheric general circulation model (AGCM) predictions.

a. Data sources

Monthly gauge rainfall data at Ecuadorian stations are provided by the National Institute of Meteorology and Hydrology (INAMHI) of Ecuador (''Meteorological Historical Data''; accessed in 2012). After quality control and testing for homogeneity, the original set of about 160 stations is narrowed to 151 stations for the 46-yr period of 1965–2010 that are eligible to be used here (Fig. 2). We omit the two Galapagos Island stations because of their geographical inconvenience, resulting in 149 stations for the study. The study uses the monthly extended reconstructed SST (ERSSTv3) data (Smith et al. 2008) extracted from the U.S. National Oceanic and Atmospheric Administration/National Climatic Data Center (NOAA/NCDC) archives. This SST dataset is a blend from ships and buoys on a 2° × 2° grid and does not include satellite data.

The ''ECHAM4.5'' European AGCM has hybrid sigma-pressure vertical coordinates, T42 spectral horizontal resolution (an approximately 2.8° grid), 19 vertical levels, and its top at 10 hPa (Roeckner et al. 1996). Here, we use only the mean of an ensemble of 24 members, integrated at the International Research Institute for Climate and Society (IRI) and stored in its online data library (http://iri.columbia.edu/resources/data-library/), to form a 3-month total rainfall output. As a predictor field, outputs from ECHAM4.5 for a selected domain for the FMA or DJF rainfall seasons are the result of a two-tiered process in which SST is initially prescribed and used as a lower boundary condition for the ECHAM4.5 integrations. The SST can be actually observed, so that the ECHAM4.5 output is a diagnostic simulation, or the SST can be predicted, so that the ECHAM4.5 output is a true forecast. Both types of SST prescription are used here.

b. *Quality control, homogenization, and gridding of rainfall data*

The monthly total precipitation data are subjected to quality control (QC), a temporal homogeneity assessment, and an objective analysis to create a gridded rainfall dataset. The QC check is first performed to identify outliers that may indicate erroneous data. Here, outliers are defined as being at least 3 standard deviation units away from the mean (Aguilar et al. 2009). Some outliers are caused by human error in digitizing the data (Zhang et al. 2005), and others are deemed correct—as in, for example, cases of extremely high local precipitation associated with a strong El Niño episode.[1] To help to determine the likelihood of an outlier being real, values at nearby stations for the same month may be examined. Entries that are considered to be uncorrectable errors are labeled as missing.

To test for lack of homogeneity in a station time series of monthly total precipitation, we use a software package called RHtestsV3 that was developed by the Climate Research Division of Environment Canada (Wang and Feng 2013) and is available online (http://etccdi.pacificclimate.org/software.shtml). This software can detect and adjust for one or more changepoints (shifts) in a data series, given the overall lag-1 autocorrelation. It is based on the penalized maximal t test (Wang et al. 2007) and F test (Wang 2008b), which are embedded in a recursive testing algorithm (Wang 2008a) assuming the lag-1 autocorrelation (Wang and Feng 2013). Additional details are found in Wang (2003). Here, we follow the procedures detailed in the RHtestsV3 user manual that accompanies the software package.

The homogeneity test procedures identify some problems in our data. Because of the lack of metadata records accompanying the rainfall data that are needed to apply the homogenization, possible cases of inhomogeneity (e.g., step changes in mean or variance) identified by the algorithm must be manually examined and corrected using one or more of 1) the climate information of a neighboring station or stations, 2) topographical information, 3) consideration of the ENSO state, and/or 4) the expert judgment of the authors and/or local INAMHI experts. Several stations are eliminated because of their moderate or severe inhomogeneity. Seventeen stations are retained but are modified from the raw data using the homogenization process. Figure 3 (bottom) illustrates a step change and its correction, for the Cañar station in the Andean highlands. Following the basic QC and then homogenization, and after a decision to exclude the two stations in the Galapagos Islands, the resulting dataset includes 149 stations in Ecuador for the predictability experiments (Fig. 2). Nearly all of the stations are located in the coastal or Andean regions of the country.

The final data preparation step is application of a Cressman objective analysis (Cressman 1959) to create regular gridded rainfall data from the station rainfalls. It is an iterative process in which gridpoint rainfall estimates are progressively refined with each of several passes. The analysis is also used to fill in some missing data at stations having otherwise mostly complete data, so that stations that originally are missing up to 25% of their data subsequently have as little as 10% missing. The gridding of the rainfall effectively smooths the patterns created from the station data. This smoothing tends to make a greater difference in the Andean highlands, where the station patterns are ''noisy'' because of the mountainous terrain, than it does in the coastal areas where the terrain is relatively more uniform. This smoothing, which carries some advantages (e.g., noise filtering) and some disadvantages (loss of station-level specificity), applies to all of the patterns of predictive diagnostics and skills to be presented in the study. The Cressman gridding process leads to a set of roughly 150 active grid squares at 30-km resolution. Details of the Cressman analysis, including the formulas used, are provided in the appendix.

With the QC, homogenization, and filling-in processes, the 149 stations and gridded data are believed to have better homogeneity, fewer missing cases, and generally improved quality.

c. *Methods*

In 2008, INAMHI implemented a local statistical forecast system using CCA to produce probabilistic forecasts of tercile-based precipitation categories (below, near, and above normal), using fields of SST and/or atmospheric variables as predictors. CCA is a multivariate statistical method that has been widely described in the literature (e.g., Hotelling 1936; Glahn 1968) and has been applied to short-term seasonal climate forecasting (Barnett and Preisendorfer 1987; Barnston 1994; Barnston and Smith 1996; Thiaw et al. 1999; Korecha and Barnston 2007). CCA calculates linear combinations of a set of predictors that maximizes relationships, in a least-squared-error sense, to similarly calculated linear combinations of a set of predictands. It is a multivariate regression, involving patterns on both the predictor and predictand sides. One combination of predictor coefficients and its corresponding combination of predictand coefficients

[1] Very high rainfall in response to El Niño may create a positively skewed distribution, in which the threshold of 3 standard deviations would not represent as unusual of a case as it would in a Gaussian distribution.

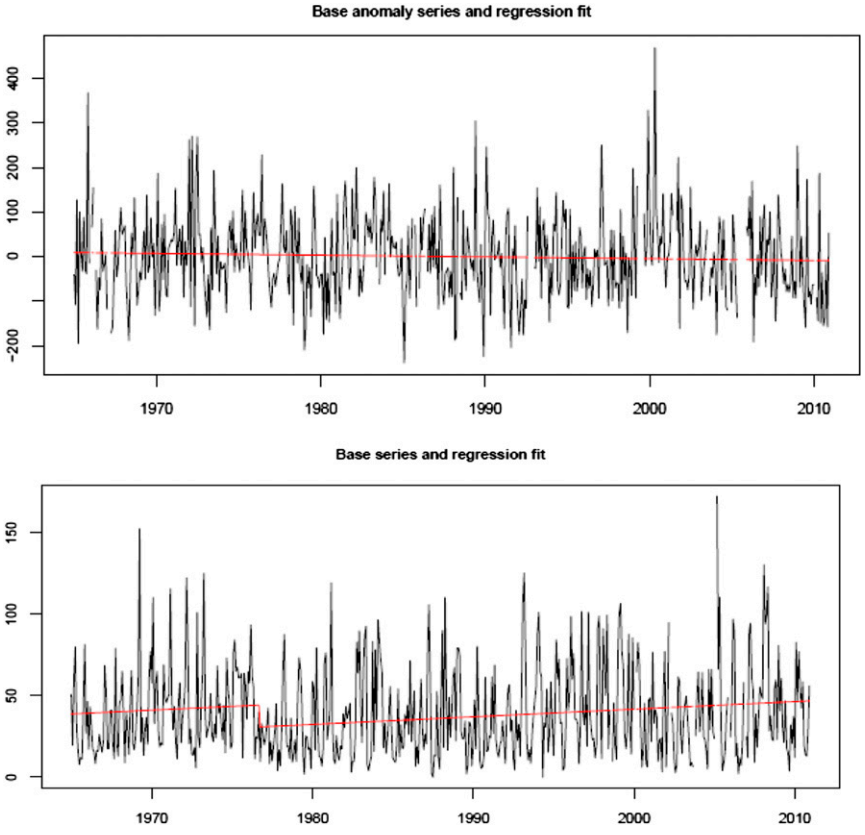

FIG. 3. (top) Homogeneity test for monthly total precipitation (mm) for Sangay. No significant step change in mean or slope is detected, and therefore the time series is considered to be homogeneous. (bottom) As in the top panel, but for Cañar. A statistically significant discontinuity detected near 1977 is verified manually against a nearby station.

constitutes one canonical mode, describing the preferred *coupled* spatial patterns relating predictor and predictand. In an initial step in running the CCA, the predictor and predictand data are separately prefiltered using empirical orthogonal function analysis, thereby preserving the most coherent, and presumably most physically meaningful, modes of variability. Here, CCA is conducted using a software package called the Climate Predictability Tool (CPT; available online from the IRI at http://iri.columbia.edu/our-expertise/climate/tools/cpt/), providing data that describe the cross-validated forecast skills and diagnostics underlying the coupled patterns and making possible forecasts both within the historical training period and for independent cases (e.g., for a future season).

Although use of the CCA is Ecuador's principal forecast method at INAMHI, dynamical input is also examined using the approach followed by the Latin American Observatory partnership (Muñoz et al. 2010, 2012), including the Weather Research and Forecasting (WRF) model (Kuo et al. 2004) with three ensemble members and two sets of parameterizations and the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; Anthes and Warner 1978) with two members and a single set of parameterizations. Together, these tools are used to develop seasonal probabilistic rainfall forecasts that provide climate services for the general public and decision makers in sectors such as agriculture, water management, energy, and health (Vaughan and Dessai 2014). Regular use of the statistical CCA represents an advance in climate-forecasting methods in Ecuador as compared with the previous use of more subjective methods. Here, we document the accuracy and other performance attributes of the forecasts produced by the CCA, helping to fill a gap in knowledge among both the scientific community and the general public.

The CPT software package builds a CCA-based prediction model that optimizes the relationships between the patterns in the predictor and those in Ecuador rainfall and then verifies the goodness, or skill, of the resulting predictions using cross validation or retroactive

designs to minimize artificial skill. In this study we use cross validation with five running consecutive cases (years for the selected season) held out, with the middle year predicted (and later verified) as a simulated independent case outside of the training sample (Michaelsen 1987; Barnston and van den Dool 1993). This process is repeated so that each year in the dataset is forecast with the climatological data redefined each time a new set of 5 years is withheld. The predictor variables may be observed fields from earlier than the forecast-targeted time, or they may be the predictions of a global dynamical model for the desired target time. Here, we use both approaches.

In the first case, when using observed predictor data, there are two timing arrangements of the predictor and predictand fields: 1) they both can be at the same time (e.g., the same season, year by year), identifying diagnostic relationships between predictor and predictand fields, or 2) they can be time lagged, representing a predictive relationship in which the predictands occur later than the predictors (e.g., January for the predictors and FMA for the predictands). The diagnostic arrangement is intended to maximize the strength of predictor–predictand relationships to represent an upper limit of predictability if the predictors were perfectly known at an earlier time, whereas the predictive arrangement represents what would be possible in a practical prediction situation. The timing arrangements used here in predicting either the FMA or DJF season include a simultaneous (diagnostic) design and two time-lagged (predictive) designs. For example, in predicting FMA the predictor is for FMA in the diagnostic design and for January or November for the short- or long-lead predictive design, respectively.

Meanwhile, in the second case, when the predictor field is a dynamical model prediction for a given time (usually, but not necessarily, the same variable as the predictand), the model forecast is for the same time as the desired forecast season (e.g., both are for FMA). The CCA statistically corrects systematic biases in the model predictions as identified over the historical period of the predictions (i.e., hindcasts) against the corresponding observations. Because CCA is a multivariate linear regression in which the model forecasts are treated as predictors and the corresponding observations as predictands, it identifies patterns in the model predictions that tend to be associated with patterns in the verifying observations over the hindcast period. For example, if the model predictions tend to place rainfall maxima over northern Peru that should appear over southern Ecuador, CCA identifies this systematic positional error and "corrects" the forecasts to become more like the past observed patterns in such cases. The dynamical model may be a fully coupled (one tiered) ocean–atmosphere model or an atmosphere-only model (an AGCM) run in a two-tiered design, such as is done here for ECHAM4.5, where the AGCM is forced using prescribed SST as the lower boundary condition. In this two-tiered case, the SST used to force the model could be observed SST, in which case the model climate "forecast" is a simulation or a perfect prognosis, or the SST could be actually predicted, in which case the model forecast is a truly time-lagged forecast.

The two cases of using the AGCM output as the predictor are examples of model output statistics (MOS) designs (e.g., Tippett et al. 2003, 2005), in which the advantages of a dynamical prediction (or simulation), such as its ability to capture nonlinear relationships, and its statistical correction can sometimes deliver better predictions than either an uncorrected dynamical prediction or a purely statistical observational prediction alone. Both cases also constitute downscaling, because gridded predictor information over a large spatial scale is used to produce more local predictions (e.g., at individual stations or at smaller grid squares than those in the model). Such downscaling makes detailed calibrations on the basis of the local predictand data history.

For the experiments using observed predictors, four different SST predictor domains are tested: the wide global tropics (30°N–30°S), the narrow global tropics (15°N–15°S), the eastern tropical Pacific (5°N–15°S, 80°–136°W), and a substantial part of the tropical Atlantic (20°N–25°S, 19°–69°W). These domains are designed to isolate potential diagnostic and predictive signals coming from each or several of these regions and to explore their relative importance while excluding information from SST regions believed to be physically irrelevant for Ecuador (e.g., the extratropics). Large amounts of such extraneous information may contribute to artificial skill resulting from fitting some of the accidental variability, even with cross validation. Within the generally more relevant predictor domains, however, inclusion of some unnecessary information usually does not hinder the CCA from identifying the predictively critical portions, the patterns for which are provided in the diagnostic output. For the AGCM predictor experiments, the predictor domain is 16°N–16°S, 64°–90°W—much larger than the domain of Ecuador—to allow for correction of the AGCM's spatial biases (to be discussed further in section 3).

The main validation measures used here are 1) the spatial distribution of the correlation between forecasts (or simultaneous simulations) and observations and the mean correlation[2] and 2) the spatial structure and temporal

[2] The overall correlation, reflecting mean skill, is computed through averaging the Fisher Z equivalents of the gridpoint correlations and then computing the anti-Fisher Z of the average.

TABLE 1. Skill results of the diagnostic and predictive experiments for Ecuador rainfall during FMA and DJF. In diagnostic experiments (top row in each set of timing designs), the predictor X and predictand Y are concurrent (depicting simultaneous relationships); in predictive experiments, the lead time is defined as the number of months between the end of the predictor month and the beginning of the forecast season. The selected statistical model contains the numbers of X, Y, and CCA modes that maximize the overall correlation skill (third column from right). Abbreviated interpretations of the CCA modes are indicated with the canonical correlation coefficient (CCC).

	Timing design	No. of modes X/Y	No. of modes CCA	Mode-1 CCC	Mode-2 CCC	Mode-3 CCC	Overall correlation	Guayaquil correlation/rank correlation	Quito correlation/ rank correlation
				FMA rainfall					
Narrow tropical oceans (15°N–15°S)	FMA–FMA	3/3	3	0.71 ENSO	0.51 trend	0.35 Andes	0.346	0.79/0.46	0.25/0.19
	Jan–FMA lead 0	4/2	1	0.61 ENSO			0.229	0.57/0.35	−0.30/−0.27
	Nov–FMA lead 2	4/3	3	0.63 trend	0.46 ENSO	0.32 Andes	0.145	0.13/0.07	0.25/0.25
Wide tropical oceans (30°N–30°S)	FMA–FMA	3/3	3	0.66 ENSO	0.51 trend	0.33 Andes	0.287	0.71/0.32	0.18/0.11
	Jan–FMA lead 0	4/2	1	0.56 ENSO			0.194	0.41/0.27	−0.20/−0.18
Eastern tropical Pacific	FMA–FMA	4/3	3	0.77 ENSO	0.53 trend	0.24 Andes	0.332	0.76/0.58	0.06/0.13
	Jan–FMA lead 0	4/3	3	0.61 ENSO	0.47 Andes, trend	0.30 unknown	0.183	0.49/0.22	0.09/0.08
Southwestern tropical Atlantic	FMA–FMA	4/2	1	0.53 monopole			0.114	0.34/0.37	−0.47/−0.46
	Jan–FMA lead 0	4/3	2	0.62 trend	0.56 monopole		0.155	0.21/0.42	−0.14/−0.05
ECHAM4.5 with MOS correction; model forced by obs or predicted SST	FMA–FMA	4/3	3	0.80 ENSO	0.60 Andes	0.47 unknown	0.387	0.72/0.70	−0.13/0.06
	Jan–FMA lead 0	3/2	2	0.50 ENSO	0.46 "ENSO"		0.190	0.49/0.12	−0.28/−0.39
	Nov–FMA lead 2	3/3	1	0.57 trend			−0.004	−0.35/−0.44	−0.07/0.02
Westward extended domain	Nov–FMA lead 2	4/3	2	0.60 trend	0.57 ENSO		0.112	0.48/0.01	0.04/0.04
				DJF rainfall					
Narrow tropical oceans (15°N–15°S)	DJF–DJF	3/2	2	0.73 ENSO	0.48 Andes		0.352	0.69/0.32	−0.02/−0.09
	Nov–DJF lead 0	3/4	3	0.67 trend	0.54 ENSO	0.38 Andes	0.238	0.41/0.02	0.09/0.05
	Sep–DJF lead 2	3/2	2	0.56 ENSO	0.45 Andes		0.232	0.39/0.13	−0.15/−0.21
ECHAM4.5 with MOS correction; model forced by obs or predicted SST	DJF–DJF	2/2	2	0.78 ENSO	0.42 Andes		0.391	0.77/0.46	−0.10/−0.09
	Nov–DJF lead 0	3/3	3	0.68 ENSO	0.45 Andes	0.43 trend	0.285	0.70/0.21	−0.32/−0.35
	Sep–DJF lead 2	2/3	2	0.58 "ENSO"	0.45 trend		0.113	0.21/0.10	−0.22/−0.22

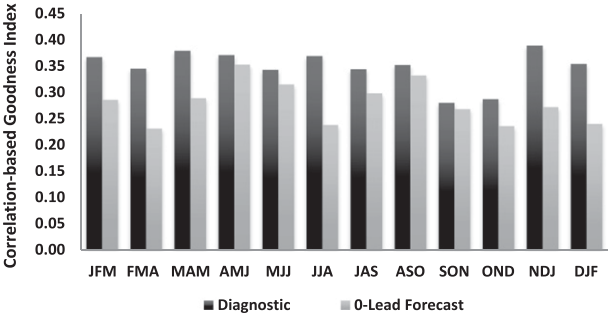

FIG. 4. Overall correlation skill over Ecuador of simultaneous diagnostic (black bars) and 0-lead forecast (gray bars), using the narrow global ocean SST predictor domain (15°N–15°S).

information revealed in the CCA loading patterns. The second measure includes maps of the coupled predictor–predictand patterns (the CCA modes) and the canonical correlation coefficients showing the strength of the linear association between the time coefficients of each.

3. Results

Results are shown first for predicting rainfall for the FMA season and then are shown for DJF. For each season, we first summarized results using observed predictors in the CCA, followed by results using dynamical model simulations or predictions from the ECHAM4.5 model. In all cases, seasonal total rainfall at the 30-km grid squares over Ecuador serves as the predictand. The CCA graphics are shown to illustrate the spatial and temporal patterns of the coupled relationships between the predictor (either SST observations or ECHAM4.5 model output) and the predictand rainfall for the leading CCA modes for the diagnostic (simultaneous) or the forecast (time lagged) time arrangement. A summary of the key results, including overall skills, for the experiments performed in this study is provided in Table 1.

a. Seasonal variation of skill for observed SST predictors

The seasonal cycle of overall correlation between CCA-generated rainfall predictions and the corresponding observations, averaged over Ecuador's gridded network, is shown in Fig. 4. The narrow global tropical SST predictor domain is selected because it is found to deliver the highest skill among the four domains in the largest number of cases. Skills are shown for cases of simultaneous diagnostics (predictor and predictand both for the same season) and for predictions at 0-month lead time, in which SST is for the month immediately preceding the predicted 3-month season.

Overall correlations average in the middle 0.30s for the simultaneous diagnostics, and in the upper 0.20s for the 0-lead predictions. Skills remain relatively constant throughout the annual cycle, including some relatively dry seasons in the third quarter of the year. For the FMA and DJF seasons to be examined here, the difference in skill between the diagnostic and the prediction is relatively large. As will be discussed below, this loss of skill with lead time may be related to uncertainty in the persistence of the state of ENSO between November and April, especially in terms of the SST anomaly in the far eastern tropical Pacific.

b. Skill for FMA target period using observed SST predictors

1) DIAGNOSTIC DESIGN

Experiments are done to determine which of the four SST predictor domains tends to result in the greatest cross-validated hindcast skill and shows the strongest SST-versus-rainfall relationships. The CCA graphics (Fig. 5) show the spatial and temporal patterns of these coupled relationships for the leading CCA mode for experiments using each of the four SST predictor domains for the diagnostic (simultaneous) time arrangement for the FMA target period.

Of the four predictor domains, the highest overall correlation (0.35) is produced using the narrow global tropical SST domain (see Table 1), and the eastern tropical Pacific predictor domain results in nearly as high a correlation (0.33). The wide global tropical domain produces slightly lower skill (0.29), and the tropical Atlantic yields a much lower skill (0.11). Of the four predictor domains, the highest canonical correlation coefficient for the leading CCA mode (i.e., the correlation between the predictor and predictand time series shown in the middle column of Fig. 5) is 0.77, obtained using the eastern tropical Pacific domain. The narrow and wide tropical ocean domains result in slightly lower coefficients, at 0.71 and 0.66, respectively. The tropical Atlantic's leading coefficient is 0.53. Except for the tropical Atlantic, the spatial patterns of the leading CCA modes describe the effects of ENSO-related SST variability in the far eastern tropical Pacific on rainfall in Ecuador's coastal region.[3] Meanwhile, consistent with experience, a group of stations in the northeastern Andes shows an opposing rainfall response (e.g., positive anomalies with negative eastern tropical Pacific SST anomalies). Although Pacific SSTs are not included in

[3] Because CCA is linear, the choice of polarity (i.e., whether El Niño or La Niña is shown, and the sign of the temporal series) is arbitrary, and the meaning of either polarity is identical.

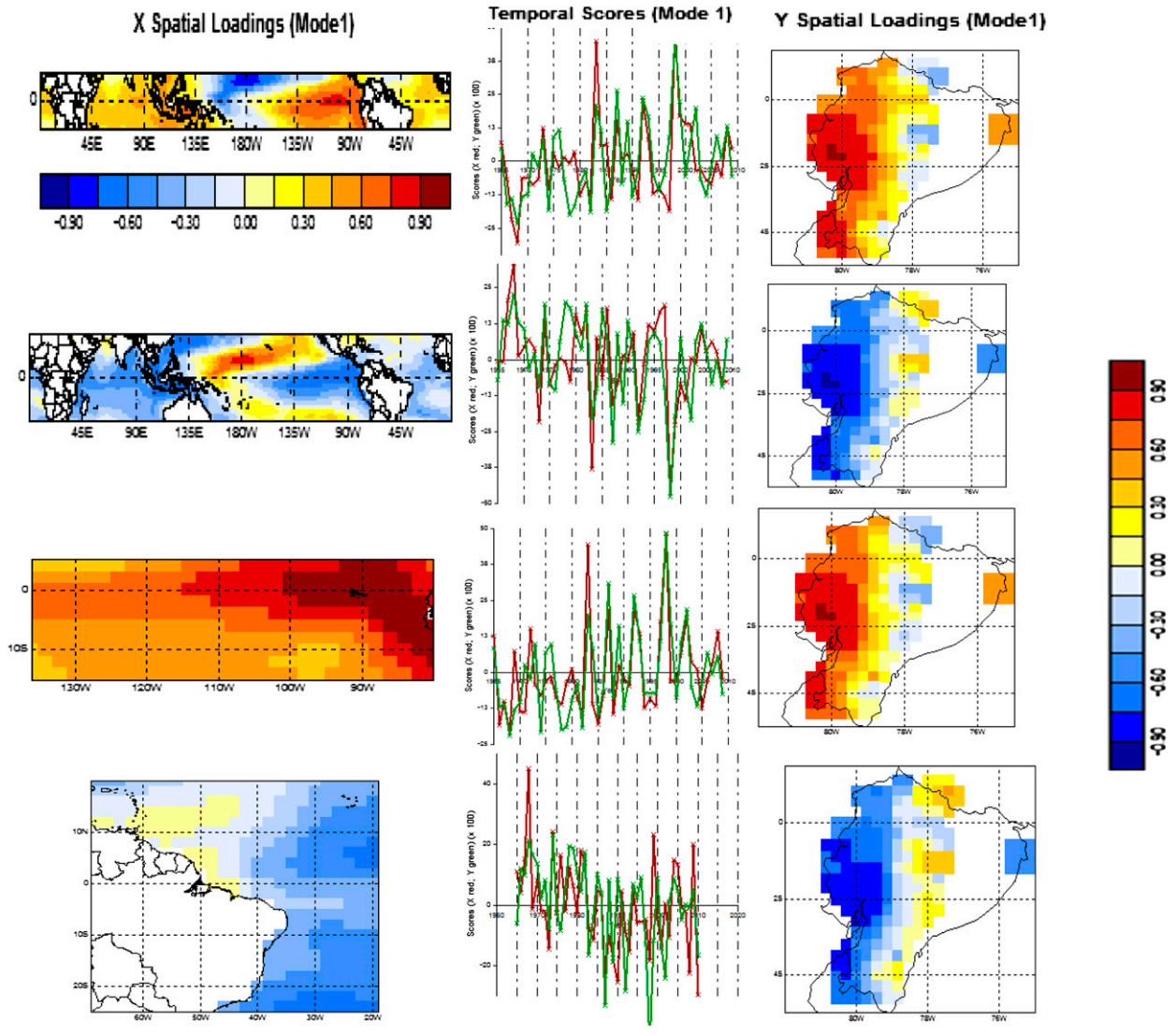

FIG. 5. CCA leading-mode loading patterns for (left) SST predictor for FMA and (right) Ecuador rainfall for FMA, and (center) the associated time series of each (red line for SST; green line for rainfall). Shown are results for the (top) narrow tropics, (top middle) wide tropics, (bottom middle) eastern tropical Pacific, and (bottom) tropical Atlantic SST predictor domains. The canonical correlation coefficients for the four rows are 0.71, 0.66, 0.77, and 0.53, respectively.

the tropical Atlantic predictor domain, the rainfall pattern over Ecuador described by the leading Atlantic mode (Fig. 5, bottom row) is similar to that for the other SST predictor domains, suggesting that the near-monopole pattern in the Atlantic SST may be a delayed effect of ENSO in which the SST anomalies in the Atlantic mirror those in the eastern Pacific.

The temporal scores (Fig. 5, middle column) show that the FMA SST patterns of greatest relevance for Ecuador's rainfall are most extreme during some of the known El Niño years (e.g., 1983, 1987, 1992, 1998, and 2002) but during only a few known La Niña–like years (1982 and 1985). La Niña often results in suppressed rainfall in the coastal region, but some ENSO-neutral years also have below-average rainfall, and as a result the La Niña years are not well differentiated. The warm and cold ENSO years do not match the traditionally known ENSO years closely because the far eastern tropical Pacific SST anomalies do not always follow the SST anomalies in the east-central basin (e.g., the Niño-3.4 region) that reflect the basinwide ENSO phenomenon.

The second and third CCA modes for FMA for the diagnostic relationship using the narrow global tropical SST domain are shown in Fig. 6. Mode 2 represents a general secular trend toward warmer SST in most of the tropics except for the central and eastern Pacific,

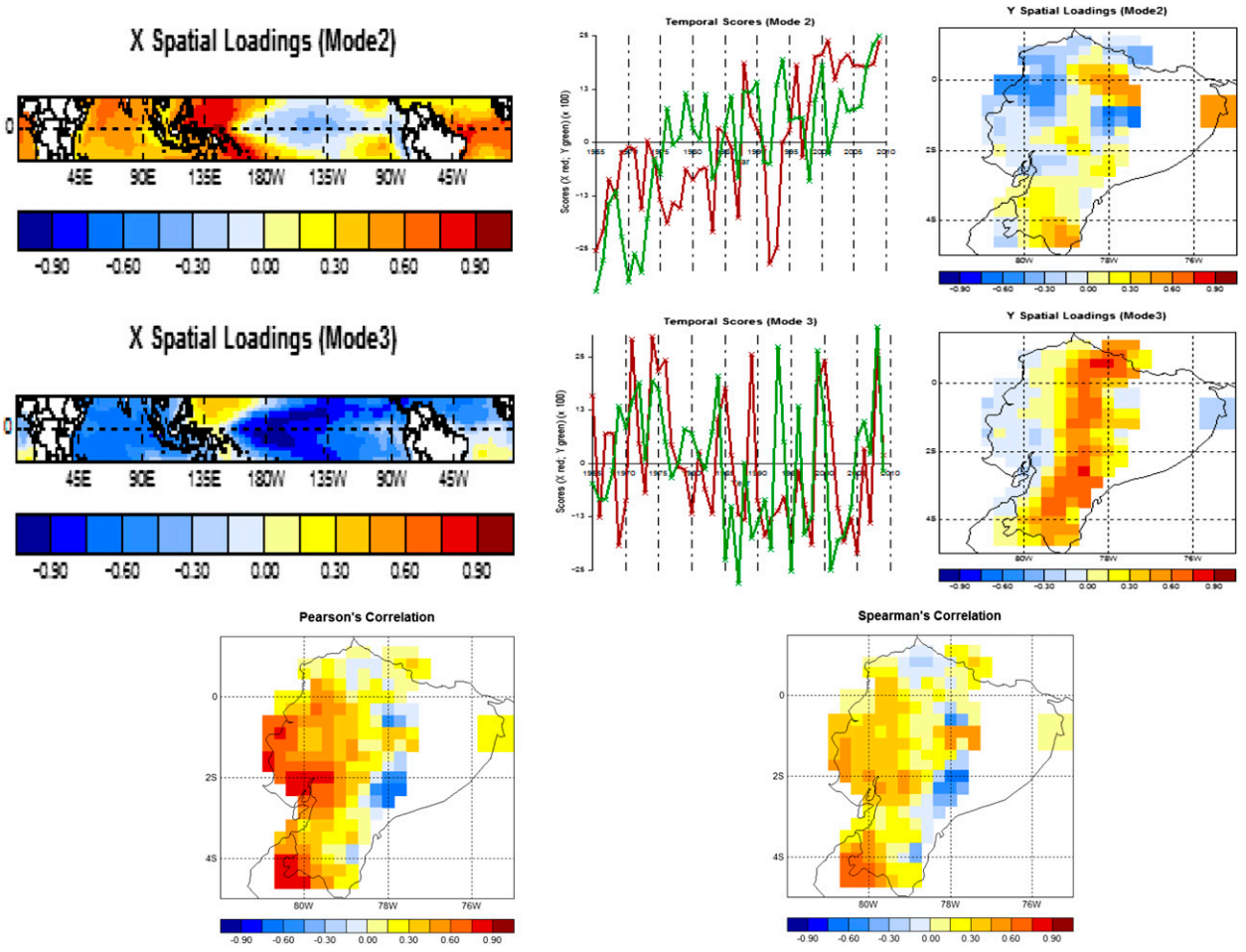

FIG. 6. SST and rainfall loading patterns and their time series for CCA (top) mode 2 and (middle) mode 3 for the FMA rainfall prediction experiment using simultaneous SST predictors in the narrow tropical ocean domain. (bottom) The spatial distribution of cross-validated (left) standard and (right) rank correlation skill for the experiment. The canonical correlation coefficients for modes 2 and 3 are 0.51 and 0.35, respectively.

with a mixed associated rainfall trend pattern in Ecuador (e.g., a drying tendency along the northern coast). This trend mode will appear in other experiments here. Mode 3 is a near-monopole SST anomaly pattern associated with a rainfall anomaly of opposite sign over the Andes and therefore will be called the Andes mode from here forward. The spatial distribution of cross-validated skill of this three-mode diagnostic model over Ecuador is shown in Fig. 6 (bottom) in terms of the standard (Pearson) correlation and the Spearman rank correlation. Skill is highest over the coastal area in the west—the region with strongest ENSO involvement—and weakens eastward into the Andes and the Amazon.

The results using only the eastern tropical Pacific SST (Fig. 5, third row) are generally similar to those of the narrow global tropical SST in both skill sources and final skill level (Table 1), except that the strength of the SST–rainfall relationship in the ENSO mode is greater for the eastern Pacific as shown by the higher canonical correlation coefficient (0.77 vs 0.71). The SST loading pattern for the ENSO mode shows that SST in the far eastern Pacific (east of 100°W) is most important for Ecuador's coastal-region rainfall. The fact that the eastern tropical Pacific SST domain produces nearly the same overall skill level as the narrow global tropical SST domain suggests that much of the key SST information affecting Ecuador rainfall is in the local waters. The question then arises as to whether prediction of ENSO itself is necessary to predict Ecuador's coastal regional rainfall, or to what degree the far eastern SST is dependent on the basinwide ENSO phenomenon.

Diagnostic results using the wide global tropical domain have features that are similar to those for the narrow SST domain (Fig. 5, row 2), but final skills are

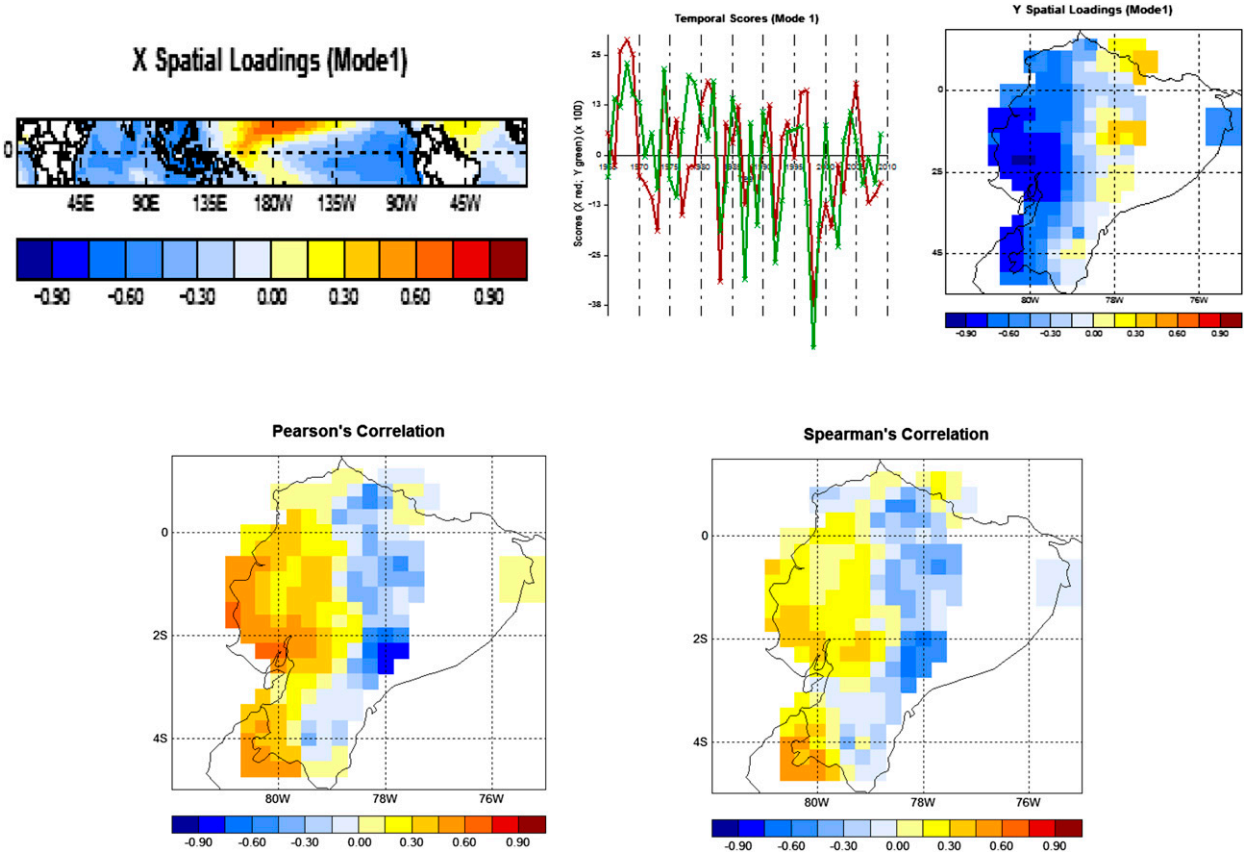

FIG. 7. (top) SST and rainfall loading patterns and their time series for CCA mode 1 for the FMA rainfall prediction experiment using January SST predictors in the narrow tropical ocean domain. (bottom) The spatial distribution of cross-validated (left) standard and (right) rank correlation skill for the experiment. The canonical correlation coefficient for mode 1 is 0.61.

slightly lower. This finding suggests that SST variability outside of the 15°N–15°S zone does not provide valuable additional information for Ecuador rainfall.

Use of the tropical Atlantic SST domain results in a single CCA mode and markedly lower skill (Table 1). As mentioned above, that mode (Fig. 5, bottom row) indicates a near-monopole SST anomaly with a relative weakness near the equator and a lack of a signal in the Caribbean Sea, in association with rainfall anomalies of the same sign over the coastal region of Ecuador. The known effect of a tropical Atlantic dipole on northeastern Brazil (Moura and Shukla 1981; Vuille et al. 2000a) does not appear in Ecuador's Amazon region; the dearth of stations in eastern Ecuador makes for difficult recognition of an Atlantic influence there, however.

2) FORECAST DESIGN

For the time-lagged forecast design using January observed SSTs as the predictor (a 0-month lead prediction), relationships between predictor and predictand patterns, and skills, are somewhat weaker than those for the simultaneous variables (Table 1; Fig. 4). The mean correlation skill and the canonical correlation coefficient for the leading CCA mode are highest when the narrow global tropical ocean SST domain is used (0.23 and 0.61, respectively). The leading canonical mode for that domain shows an ENSO-related SST pattern and the familiar rainfall pattern emphasizing Ecuador's coastal region (Fig. 7, top). These patterns, and the associated temporal scores, are similar to those seen in the simultaneous diagnostic. Although a second and third mode help in the diagnostic model, they do not enhance the prediction model and are not included. The spatial distribution of skill (Fig. 7, bottom) is a weakened version of that found for the simultaneous diagnostic (Fig. 6, bottom).

The experiment using the eastern tropical Pacific domain yields the same canonical correlation coefficient (0.61) as that of the narrow global tropical domain but yields a slightly lower skill of 0.18. The leading mode is the ENSO mode, as found in the simultaneous diagnostic, but the spatial focus is now less strongly limited to the far eastern basin (Fig. 8). One interpretation of this result is

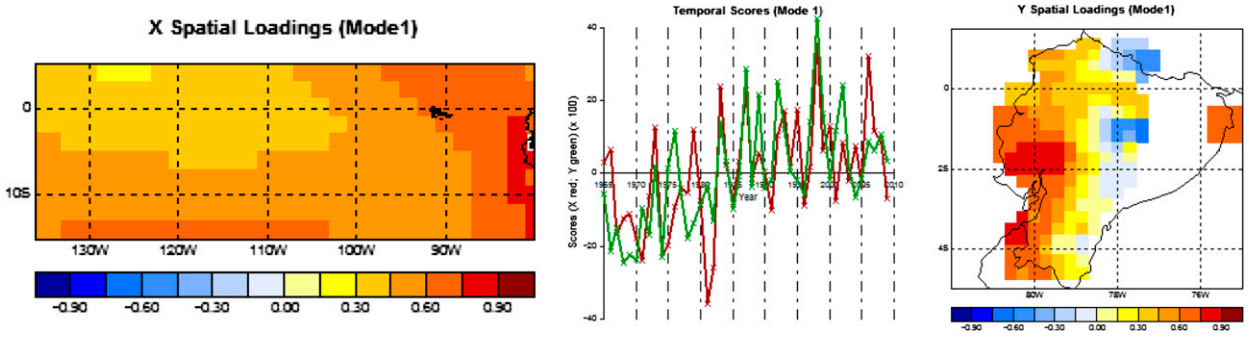

FIG. 8. SST and rainfall loading patterns and their time series for CCA mode 1 for the FMA rainfall prediction experiment using January SST predictors in the eastern tropical Pacific domain. The canonical correlation coefficient for mode 1 is 0.61.

that the far eastern SST is what ultimately matters for Ecuador's rainfall, but the general (basinwide) ENSO state becomes more important at an earlier time because it implies what SST anomaly is most likely farther east at the forecast target time. The canonical correlation coefficient of the leading CCA mode of the eastern Pacific SST domain model, which was higher than that for the narrow global tropical SST mode model for the simultaneous diagnostic relationship, is no longer so for the January predictor time (Table 1). In addition, the skill difference between the two models, very slightly favoring the narrow global tropical SST domain for the diagnostic relationship, is increased for the January predictor time. These findings support the idea that the basinwide ENSO state assumes greater importance for Ecuador's rainfall relative to the far eastern Pacific SST anomaly alone for time-lagged predictions, despite the likelihood that the latter may be the most important diagnostic factor.

To help to determine when the general ENSO state may be a better time-lagged predictor of the far eastern tropical Pacific SST than is the eastern SST itself, lag correlations are computed for the central Niño-3.4 SST (5°N–5°S, 120°–170°W), for the eastern Niño-1+2 region (0°–10°S, 90°–80°W), and between the two regions, as based on 1965–2010 monthly SST data. The correlation results are shown in Table 2 for lags ending in March (representative of the FMA season) and January (for the DJF season). Lag correlations of Niño-3.4 are far greater than those of Niño-1+2. The simultaneous correlation between the two regions is only moderate and drops off as Niño-3.4 precedes Niño-1+2 by increasing lag time. When the lag time increases sufficiently, however, there comes a point beyond which the cross-region lag correlation exceeds the Niño-1+2 lag correlation. For forecasts beyond that lag time, Niño-3.4 (representing the basinwide ENSO) becomes a better predictor of Niño-1+2 (and of Ecuador's rainfall) than does Niño-1+2 itself. Note, however, that the expected skill of forecasts of Niño-1+2 at lead times sufficient for Niño-3.4 to be a better predictor than Niño-1+2 is low for the predictions of March.

The tropical Atlantic SST predictor domain provides weaker predictive information than do the other domains (in part due to the lack of stations in the Amazon region) but delivers slightly more predictive information for the 0-lead forecast than simultaneously (overall correlation of 0.15 vs 0.11), has a stronger leading-mode canonical correlation (0.62 vs 0.53), and carries two CCA modes instead of 1. The coupled patterns described by the two modes are shown in Fig. 9. The leading mode shows a secular trend and a rainfall anomaly pattern similar to that found for the trend mode in other experiments (e.g., Fig. 7, top row) and has an Atlantic SST dipole pattern with a strong (weak) anomaly to the north (south) of about 10°S. The second CCA mode is similar to that found for the simultaneous diagnostic (Fig. 5, bottom row), with a monopole in much of the Atlantic domain. As mentioned earlier, the second mode may be an indirect consequence of ENSO variability in the tropical Pacific. The pattern of resulting predictive skill (not shown) resembles that of the simultaneous diagnosis, with highest skill over Ecuador's coastal region.

Given that the relationships described in the simultaneous diagnostics are also found at lag time from January, one might ask if they would appear also at a longer lag time, implying potential for longer-lead forecasts. We repeat the narrow global tropical SST domain experiment using November SST as predictor for the FMA target season (a 2-month lead time). The result is a lower level of skill (overall correlation of 0.145) and now three CCA modes contributing to the forecasts. The leading mode is the general secular trend, the second mode is ENSO related, and the third is the Andes rainfall mode discussed earlier. The ENSO mode (mode 2; Fig. 10) shows the familiar ENSO versus Ecuador rainfall relationship. Because a typical ENSO episode is near peak strength by November, the potential for skillful forecasts for FMA exists if the eastern Pacific SST remains

TABLE 2. Lag correlations for the Niño-3.4 and Niño-1+2 SST index regions, and cross correlations between them, for the March target month (representing the FMA target season examined here; top half) and for the January target month (representing the DJF target season; bottom half). Correlations corresponding to the specific timing designs used in this study are shown in boldface type. An asterisk in the "Nino3.4 vs Nino1+2" column indicates that the cross-region lag correlation exceeds the lag correlation for Nino1+2. SST data spanning 1965–2010 are used.

Lag correlation	Niño-3.4 vs itself	Niño-1+2 vs itself	Niño-3.4 vs Niño-1+2	Timing design in this study
Mar–Mar	**1.00**	**1.00**	**0.57**	**Diagnostic**
Feb–Mar	0.97	0.83	0.53	Partial diagnostic
Jan–Mar	**0.93**	**0.63**	**0.46**	**0-month lead**
Dec–Mar	0.90	0.51	0.42	1-month lead
Nov–Mar	**0.87**	**0.47**	**0.40**	**2-month lead**
Oct–Mar	0.84	0.42	0.38	3-month lead
Sep–Mar	0.81	0.49	0.38	4-month lead
Aug–Mar	0.78	0.33	0.35*	5-month lead
Jan–Jan	**1.00**	**1.00**	**0.76**	**Diagnostic**
Dec–Jan	0.98	0.92	0.76	Partial diagnostic
Nov–Jan	**0.96**	**0.86**	**0.76**	**0-month lead**
Oct–Jan	0.94	0.82	0.77	1-month lead
Sep–Jan	**0.92**	**0.77**	**0.76**	**2-month lead**
Aug–Jan	0.88	0.65	0.70*	3-month lead
Jul–Jan	0.84	0.57	0.60*	4-month lead
Jun–Jan	0.75	0.46	0.55*	5-month lead

anomalous over the intervening period. The result found here indicates that such SST anomaly endurance happens in some cases but that it is not guaranteed; therefore, in contrast to the shorter-lead forecast and the simultaneous diagnostic, the ENSO mode shows a weaker canonical correlation than the trend mode.

c. *Skill for FMA target period using ECHAM4.5 model output predictor*

For the dynamical model correction design we define a predictor domain that is larger than that of the Ecuador predictand, because the CCA would then have the opportunity to adjust for locational biases. The enlarged domain would enable detection and spatial correction of features predicted both inside and outside Ecuador—in some cases, moving them inside the country. For the experiments here, we use the predictor domain of 16°N–16°S, 64°–90°W (see Fig. 11, top-left panel).

In the ECHAM4.5 model diagnostic correction MOS design, the model is forced using prescribed *observed* SST for the target period of FMA. The model rainfall "forecast" is then often called a simulation, or a perfect prognosis. In the predictive case, the model is forced using SST that was actually predicted from an earlier time—in this case, by the constructed-analog statistical method (van den Dool 1994, 2007).

1) MODEL SIMULATIONS USING CONCURRENT OBSERVED SST

In the case of the diagnostic analysis, the corrected dynamical model simulation delivers an overall correlation value of 0.39 and a canonical correlation coefficient of 0.80 for the leading canonical mode. In the leading mode, a spatial pattern correction over Ecuador is not apparent (Fig. 11, top-right panel vs top-left panel), but local calibrations (i.e., downscaling) are noted in the details of the rainfall prediction at grid squares that are much smaller than those of the AGCM, of which only about three lie over Ecuador. The large-scale rainfall predictor pattern is clearly that of ENSO. Accordingly, the temporal scores (Fig. 11, top-center panel) show that the simulation reproduces the anomalous rainfall during the strongest El Niño years in Ecuador fairly well. Because the model simulation patterns of CCA modes 2 and 3 (not shown) depend on the peculiarities of the ECHAM4.5 model within the limited domain, the meanings of these modes are better identified through the rainfall patterns. The second CCA mode appears to be the Andes mode described earlier, while the identity or physical meaning of the third CCA mode is unknown.

The coherency of the leading CCA mode and the resulting overall correlation skill (Fig. 11, bottom) are stronger than those found in the diagnostic observational experiments (Table 1). This superiority of the model simulations may be due to its successful reproduction of some of the nonlinear dynamics relating SST to precipitation—dynamics that the linear CCA cannot model.

2) MODEL PREDICTIONS USING PREDICTED SST

In the predictive analysis for FMA from the January start time, the model is forced by actually predicted SST

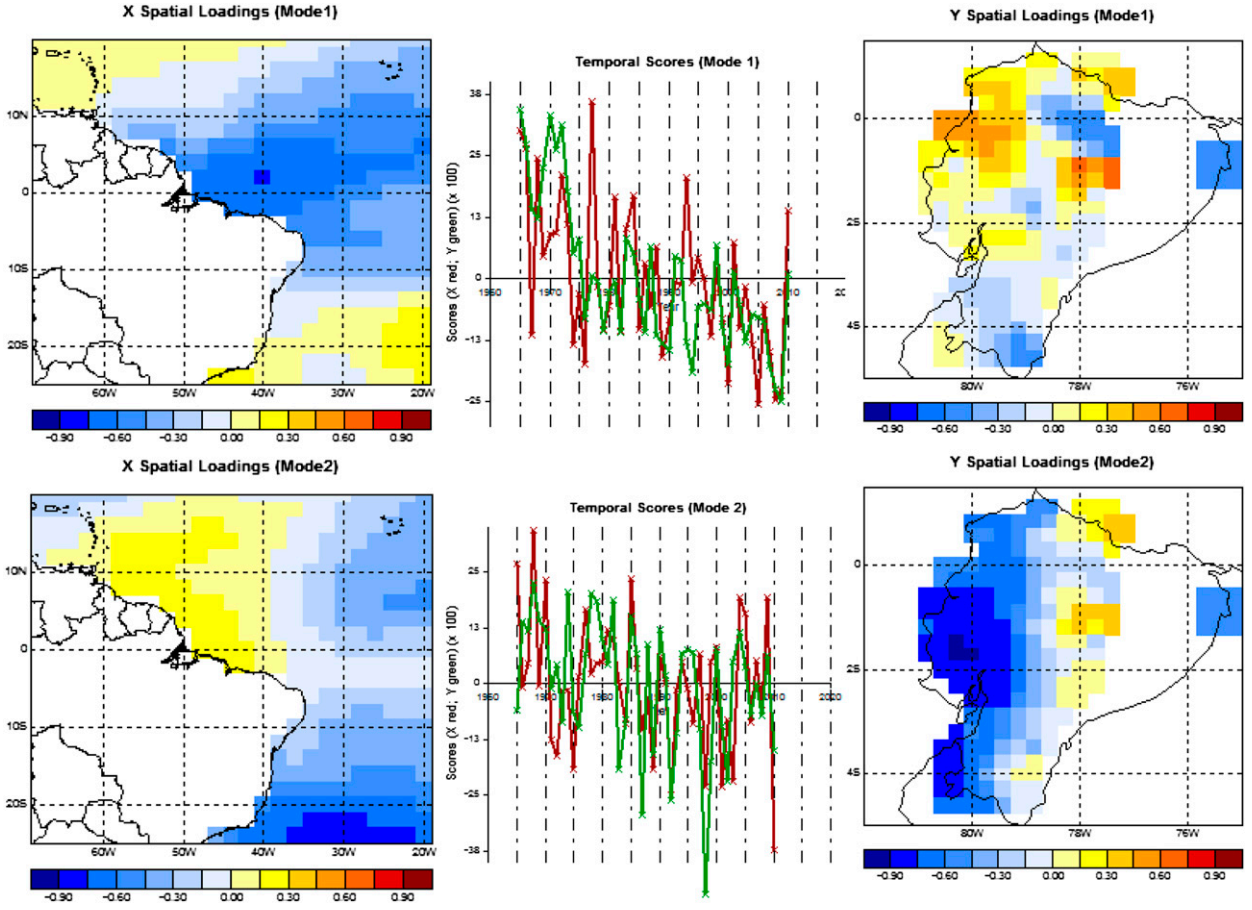

FIG. 9. SST and rainfall loading patterns and their time series for CCA mode (top) 1 and (bottom) mode 2 for the FMA rainfall prediction experiment using January SST predictors in the tropical Atlantic domain. The canonical correlation coefficients for modes 1 and 2 are 0.62 and 0.56, respectively.

using the constructed-analog method. The corrected dynamical model forecast produces an overall correlation of 0.19 and a canonical correlation coefficient of 0.50 for the leading CCA mode. These are both considerably lower than were found in the corrected model simulation. The spatial patterns shown in the model predictor and predictand maps for the two included CCA modes (Fig. 12) suggest two ''flavors'' of ENSO relationship,[4] which together incorporate the general ENSO influence on rainfall in Ecuador's coastal region. A spatial pattern correction between predictor and predictand maps (over Ecuador) consists mainly of local calibrations (Fig. 12, left vs right). The temporal scores show that both of the leading modes help to capture the seasonal rainfall anomalies observed during the strongest El Niño years in Ecuador. The superiority of the dynamical model over the observational predictor approach seen in the simultaneous diagnostic is not present in the 0-lead predictions, however. This loss of relative skill may be due either to the quality of the ECHAM4.5 model or to the SST forecast used to drive it, or to both. Below, we will show evidence that the second factor is likely.

When the lead time for the predictive analysis using the dynamical model is increased such that the SST forecast is from November, the predictive skill decreases to zero and the sole CCA mode resembles the trend mode from the earlier experiments. Because many ENSO episodes dissipate between December and March and may dissipate sooner in the far eastern tropical Pacific than farther toward the central part of the basin, one possibility is that the constructed-analog forecasts of SST tend do not maintain an ENSO-related SST anomaly near the South American coast long enough for the AGCM to respond with anomalous

[4] The appearance of two ENSO flavors may be physically based but could also be a mathematical artifact of the CCA, given the finite data sample.

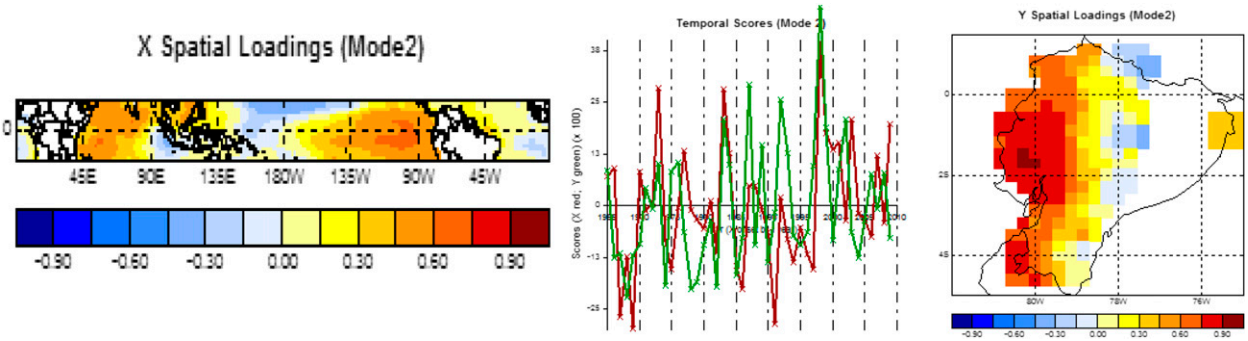

FIG. 10. SST and rainfall loading patterns and their time series for CCA mode 2 for the FMA rainfall prediction experiment using November SST predictors in the narrow tropical ocean domain. The canonical correlation coefficient for mode 2 is 0.46.

rainfall in Ecuador. The constructed analog has had reasonably competitive forecasts for the Niño-3.4 region (e.g., Barnston et al. 2012), but that region is better predicted than the SST along the South American coast.

To explore the possibility of this premature ending of the predicted eastern SST anomalies in ENSO events, the longer-lead experiment is repeated but with the model forecast domain extended westward to 175°W. The result (Fig. 13) is an increased overall correlation

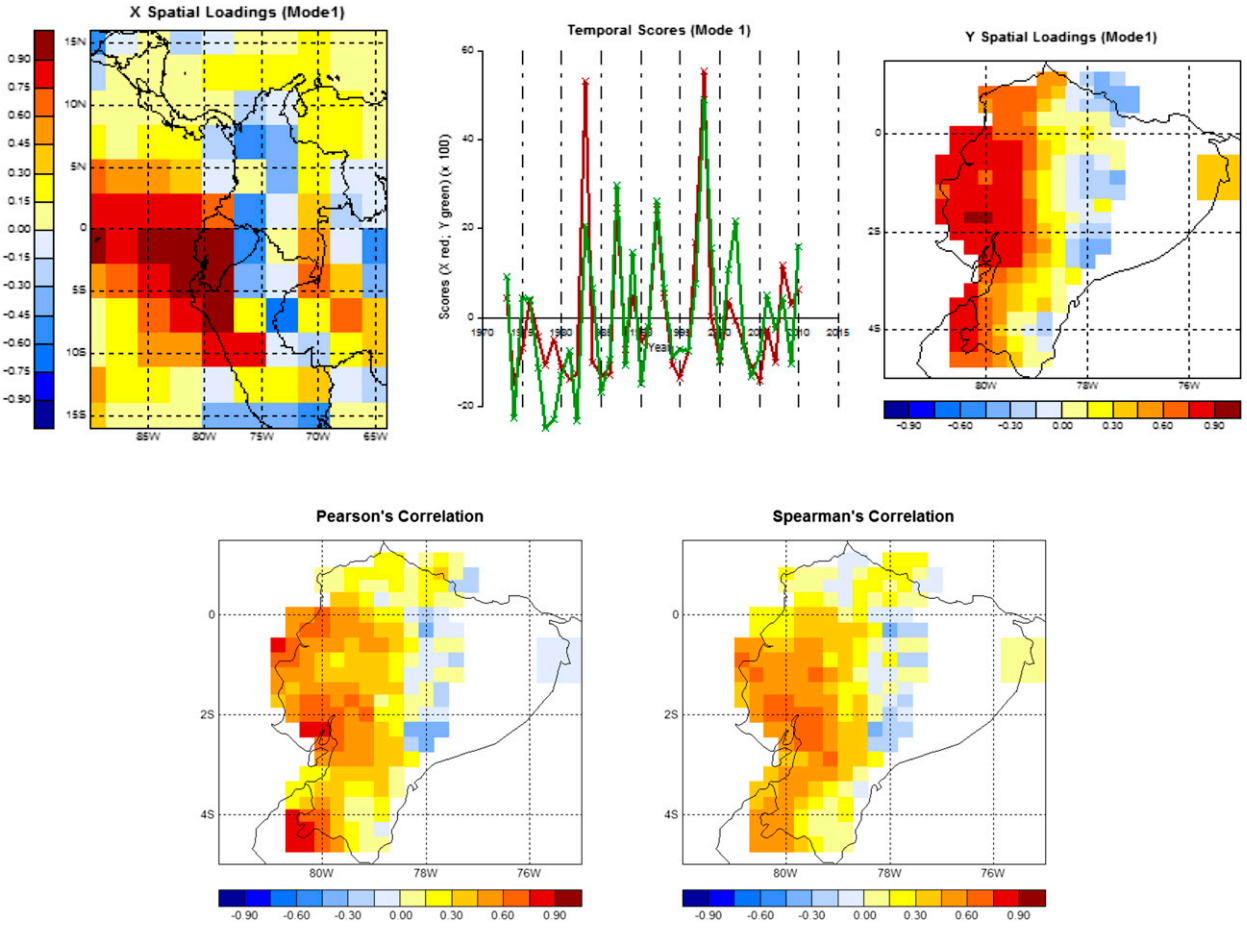

FIG. 11. (top) Model-predicted and observed rainfall loading patterns and their time series for CCA mode 1 for the FMA rainfall prediction model using ECHAM4.5 simulations as the predictor, forced by observed FMA SST. (bottom) The spatial distribution of cross-validated (left) standard and (right) rank correlation skill for the experiment. The canonical correlation coefficient for mode 1 is 0.80.

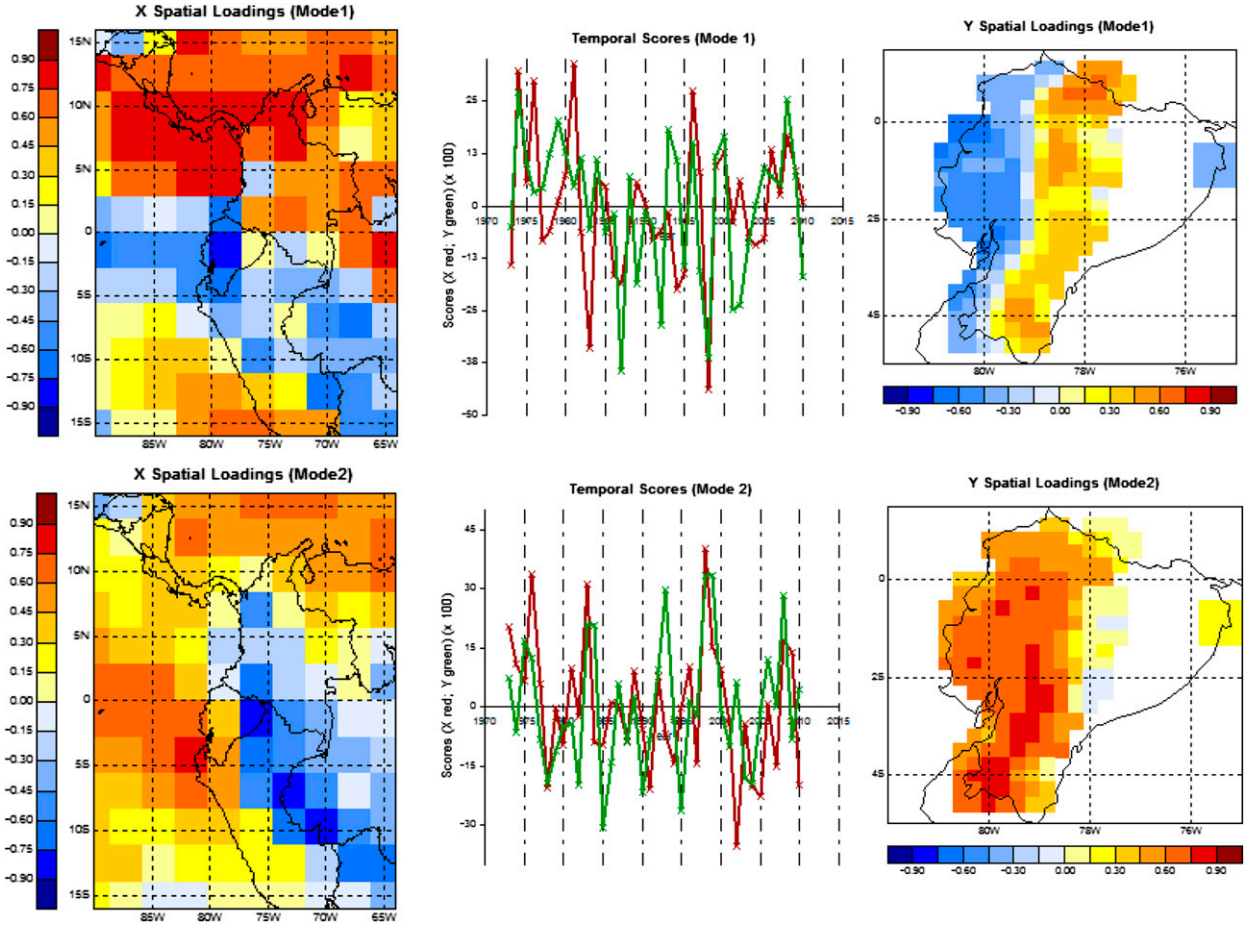

FIG. 12. Model-predicted and observed rainfall loading patterns and their time series for CCA modes (top) 1 and (bottom) 2 for the FMA rainfall prediction experiment using SST predicted from a January start time to force the ECHAM4.5 model. The canonical correlation coefficients for modes 1 and 2 are 0.50 and 0.46, respectively.

skill of 0.11 and an ENSO-related second CCA mode (with canonical correlation of 0.57). (Mode 1 is the trend mode, with 0.60.) The predictor pattern of mode 2 shows anomalous model-predicted rainfall over the east-central tropical Pacific that does not extend to near the South American coast, suggesting that the SST prediction provided insufficient anomalies in the far eastern tropical Pacific basin. This result indicates that a farther western-extended ECHAM4.5 model prediction domain might have helped the skills for some of the other seasonal forecast timings in this study, such as the 0-month lead forecast. Although CCA-based corrections for dynamical model predictions have used predictor locations that are remote from the predictand location before (e.g., Tippett et al. 2003, 2005) and can be effective, they make the rainfall prediction more of a statistical exercise. We chose not to use such remote model indicators here with the hope that there would be model predictor signals close to the region being predicted.

d. Skill for DJF target period using observed SST predictors

We explore the predictive skill potential for the DJF season using the narrow global tropical SST predictor domain for the observational design, because of its highest relative skill in the FMA experiments, and the ECHAM4.5 model correction design for simultaneous simulations and time-lag predictions.

1) DIAGNOSTIC DESIGN

Overall correlation skill for DJF rainfall over Ecuador using DJF observed SST is 0.35, a usefully high level equaling that of the comparable experiment for the FMA season (Table 1). The leading CCA mode is ENSO related, shown in Fig. 14 (top), with a canonical correlation of 0.73. The relationship looks similar to that noted above for FMA. A second CCA mode (not shown) is the one identified in the FMA experiments

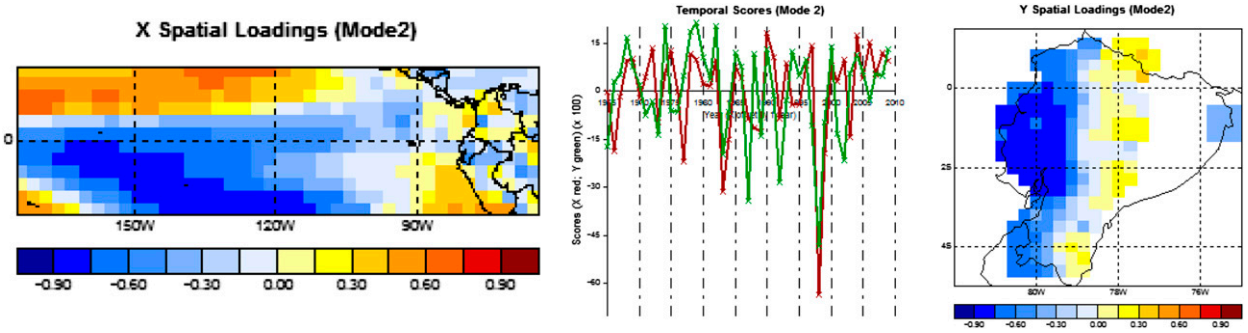

FIG. 13. Model-predicted rainfall in a westward-expanded domain and observed rainfall loading patterns and their time series for CCA mode 2 for the FMA rainfall prediction experiment using SST predicted from a November start time to force the ECHAM4.5 model. The canonical correlation coefficient for mode 2 is 0.57.

affecting precipitation in the Andes. The spatial distribution of correlation skill (Fig. 14, bottom) features maximum skill in the coastal region. In comparison with the FMA result, DJF has a more concentrated influence from ENSO and less influence from trend.

2) FORECAST DESIGN

With the SST predictor in November, skill levels drop to 0.24, just incrementally higher than the January predictor for the FMA season, and three CCA modes are used: the trend, ENSO, and the Andes rainfall modes, with canonical correlations of 0.67, 0.54, and 0.38, respectively. The leading two modes and the spatial distribution of correlation skill are shown in Fig. 15. In comparing the skill map with those for FMA forecasts from January, it is seen that skill is higher along the coast for the FMA forecasts than it is for the DJF forecasts, but some positive skill is noted farther eastward into the Andes in DJF than in FMA. The higher coastal-region skill during FMA may be related to the greater sensitivity of rainfall to ENSO when the climatological SST is closer to convective thresholds, even if ENSO predictability is slightly lower for FMA than for DJF because of uncertainty in the ENSO dissipation time.

The 2-month-lead forecast experiment using September SST indicates a remarkable retention of overall skill when compared with the November start time for the FMA season, with correlation of 0.23 for DJF as compared with 0.145 for FMA. The leading CCA mode is the ENSO mode (not shown) with a canonical correlation of 0.56, and the second is the Andes rainfall mode noted earlier. The spatial distribution of skill (Fig. 16) is similar to that found for the shorter-lead forecast, but with smaller magnitude. A reason for better forecasts for DJF from September SST than for FMA forecasts from November SST may be the typically greater endurance of an ENSO episode over the former time span than the latter one, at the end of which some ENSO episodes are largely completed—especially in the far eastern Pacific SST that matter most for Ecuadorian rainfall anomalies. The appearance of this usable level of skill 2 months in advance of the target period implies the possibility for useful early warning well ahead of the rainy season in Ecuador's coastal region.

e. Skill for DJF target period using ECHAM4.5 model output predictor

The model correction MOS design is repeated for the DJF season, first for simulations in which the ECHAM4.5 model is driven by observed SST simultaneous with the target period and then for forecasts in which SST is predicted from an earlier time using the statistical constructed-analog method.

1) MODEL SIMULATIONS USING CONCURRENT OBSERVED SST

The diagnostic analysis with the statistical correction of ECHAM4.5 simulations results in overall correlation skill of 0.39 for the DJF season (Table 1), nearly identical to the comparable experiment for FMA. The leading CCA mode represents ENSO (Fig. 17), with a canonical correlation of 0.78. The spatial distribution of correlation skill (not shown) indicates high skill over Ecuador's coastal region and western portions of the Andes. We ask whether the ECHAM4.5 model can deliver greater forecast skill at lag times than is found for FMA, which was seen to drop off sharply when lags were introduced.

2) MODEL PREDICTIONS USING PREDICTED SST

For the predictive analysis with the ECHAM4.5 model forced by predicted SST from a November start, results appear similar in structure to those found for the diagnostic analysis but with a weaker ENSO signal and an overall correlation skill of 0.29 rather than 0.39. The leading CCA mode (Fig. 18, top) is the ENSO mode with canonical correlation of 0.68. Two additional CCA

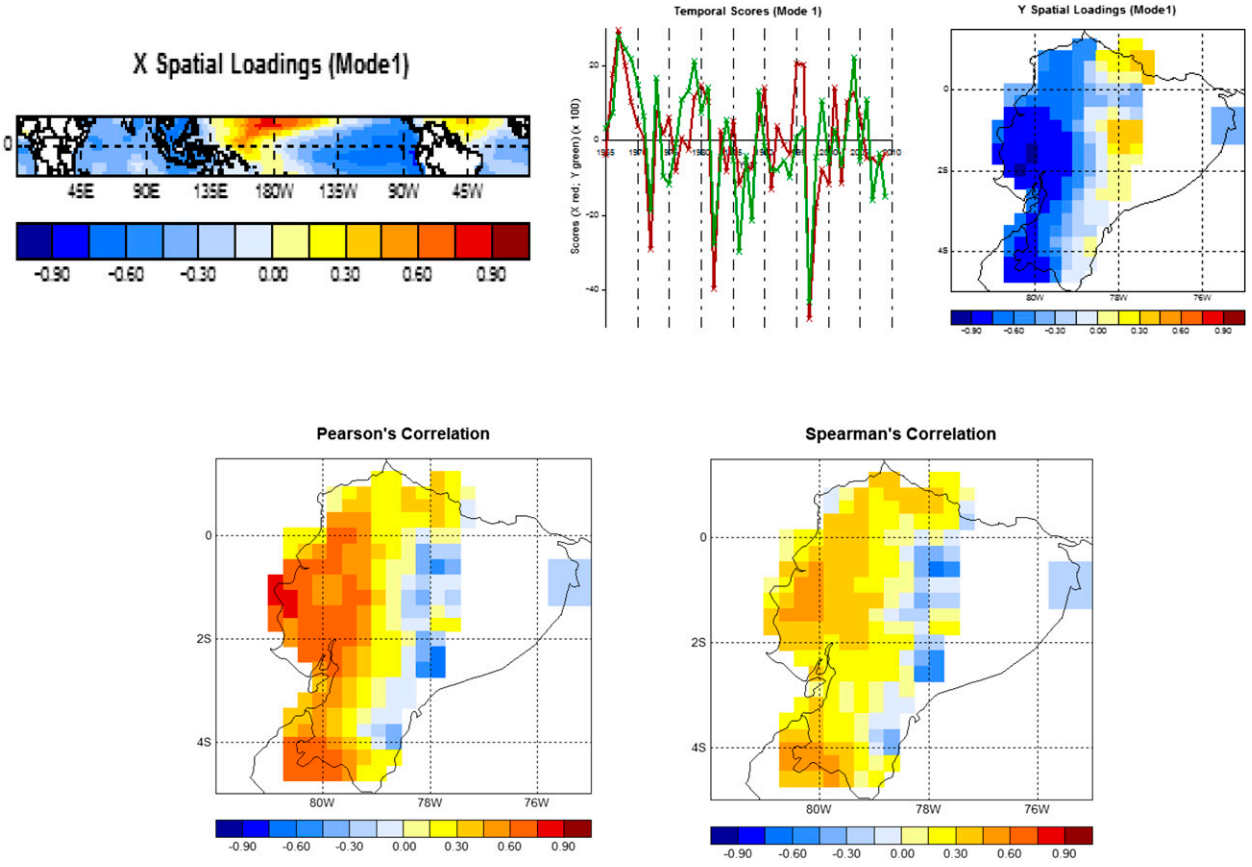

FIG. 14. (top) SST and rainfall loading patterns and their time series for CCA mode 1 for the DJF rainfall prediction experiment using simultaneous SST predictors in the narrow tropical ocean domain. (bottom) The spatial distribution of cross-validated (left) standard and (right) rank correlation skill levels for the experiment. The canonical correlation coefficient for mode 1 is 0.73.

modes (the Andes rainfall mode and the trend mode; not shown) add to the total predictive skill, whose spatial pattern is shown in Fig. 18 (bottom). The overall skill is slightly higher for the ECHAM4.5-based model than for the comparable experiment using the observational model, whereas for the FMA season the ranking of the two methods is the opposite. For FMA the 0-month-lead ECHAM4.5 forecast may have suffered from an SST forecast that had inadequate SST anomaly amplitude over the eastern tropical Pacific, as seen unequivocally above in the 2-month-lead ECHAM4.5 forecast.

With a longer-lead ECHAM4.5 forecast using SST predicted from September, some overall skill remains, but only at the 0.11 level. An ENSO-like CCA mode and a trend mode appear (not shown), but with weak canonical correlations (0.58 for the ENSO-like mode) and, although the spatial distribution of skill emphasizes the coastal region, levels are modest, with only a few squares attaining 0.45. With this longer-lead forecast, the observational approach yields a more usable level of skill in Ecuador. As shown above for the ECHAM4.5 forecast for FMA, besides possible errors in the ECHAM4.5 model, the constructed-analog SST forecast may have questionable skill in maintaining ENSO conditions in the far eastern part of the tropical Pacific even if it maintains them farther west toward the central Pacific.

4. Discussion and conclusions

Statistical and partially dynamical diagnostic and forecast experiments are performed for seasonal total rainfall anomalies in the coastal and Andean regions of Ecuador, with a focus on the FMA season because it is the rainy period common to these portions of the country. The DJF season is also examined because it is more closely linked to ENSO while being economically sensitive to rainfall performance. The purpose of the study is to estimate the levels and sources of predictive skill for Ecuador's interannual rainfall variability, using historical data on a 30-km grid covering 1965–2010 (46 yr).

The experiments include several predictor domain choices, for a simultaneous diagnostic and for the

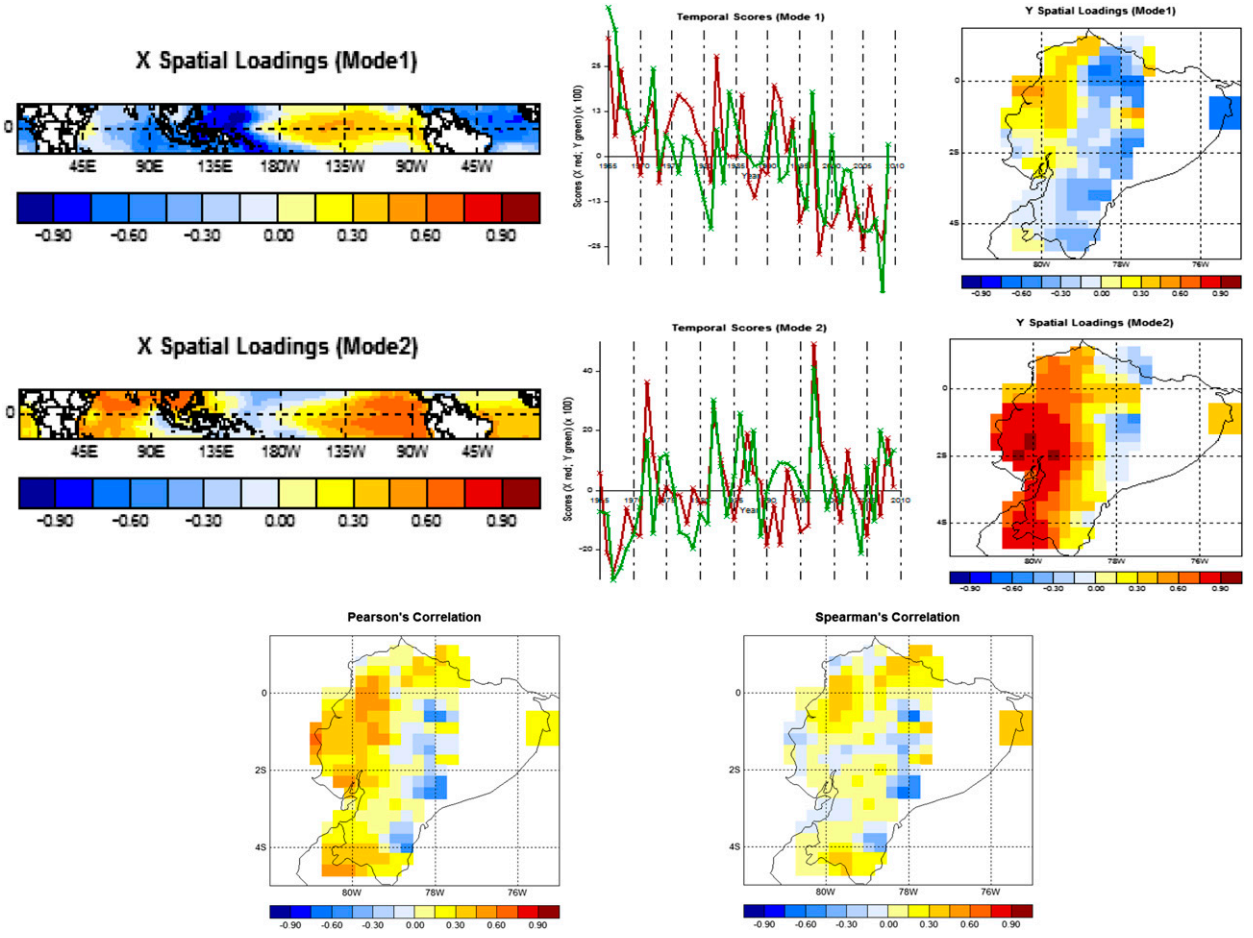

FIG. 15. SST and rainfall loading patterns and their time series for CCA modes (top) 1 and (middle) 2 for the DJF rainfall prediction experiment using November SST predictors in the narrow tropical ocean domain. (bottom) The spatial distribution of cross-validated (left) standard and (right) rank correlation skill levels for the experiment. The canonical correlation coefficients for modes 1 and 2 are 0.67 and 0.54, respectively.

time-lagged forecasts. Experiments are conducted 1) using observed SST predictors and 2) for the ECHAM4.5 dynamical model predictor in a MOS correction design. For the diagnostic and predictive analyses using observed SST predictors, the highest skill for the FMA season is obtained using SST in a narrow (15°N–15°S) global tropical ocean domain, and this domain is then used for most of the subsequent experiments.

The findings of the study, summarized in Table 1, show the spatial average of cross-validated rainfall forecast correlation skill, and the principal sources of predictive skill as revealed by the leading modes of the CCA. For the ECHAM4.5 model MOS correction design, the simulation analysis (in which observed SST is prescribed to force the ECHAM4.5 model) produces the highest spatially averaged correlation skill of any of the tests performed in this study—including those from the simultaneous (diagnostic) SST predictor design. On the other hand, the ECHAM4.5 forecasts (with SST forcing predicted from earlier start times) showed lower rainfall prediction skill than did forecasts using observed SST from the same start time. This skill difference may be due in part to imperfections in the ECHAM4.5 model and/or in the constructed-analog statistical model used to predict the SST to force ECHAM4.5. Evidence of this last possibility is seen in 2-month-lead ECHAM4.5 forecasts for the FMA season, in which the SST forecast fails to maintain anomalous SST in the far eastern tropical Pacific during ENSO events.

In all designs, the ENSO phenomenon plays a dominating role in creating skillful predictions (or simultaneous diagnostics) of Ecuador rainfall—particularly in the coastal region. The central Andes and Amazon regions have comparatively lower associations with the predictor fields, and the complex Andes topography adds additional ''noise'' to the relationships with governing SST. Although

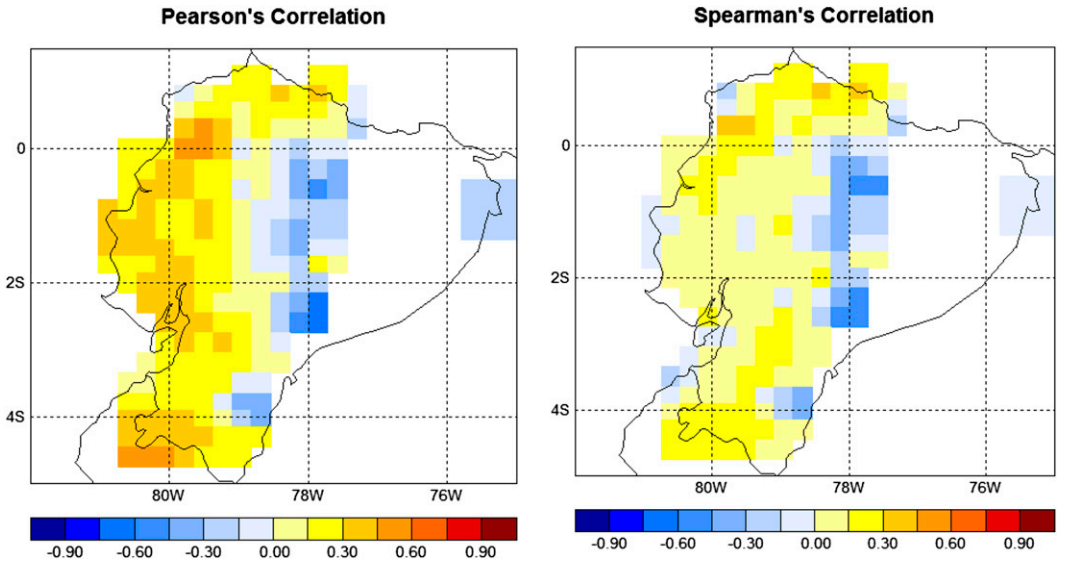

FIG. 16. Spatial distribution of cross-validated (left) standard and (right) rank correlation skill for the experiment predicting the DJF rainfall using September SST predictors in the narrow tropical ocean domain.

the ENSO phenomenon is instrumental in enabling diagnostic and predictive skill in Ecuador from the western Andes westward, the relationship between the basinwide ENSO phenomenon and rainfall is not direct, because the basinwide ENSO may not necessarily include SST anomalies in the far eastern tropical Pacific (e.g., east of 105°W), where the SST anomalies are essential in governing rainfall anomalies in western Ecuador. The importance of the far eastern SSTs is confirmed by strong relationships in the simultaneous diagnostic experiments using eastern tropical Pacific SST predictors alone, versus stronger relationships with the basinwide ENSO SST pattern in time-lagged predictive mode. Thus, with increasing forecast lead time, the far eastern Pacific SSTs become increasingly predictable through prediction of the basinwide ENSO rather than by prediction of the eastern Pacific SST alone.

Modes less important than ENSO for rainfall in much of Ecuador include 1) a secular trend in much of the global SST (but not the ENSO-related central and eastern Pacific) and an associated complex pattern of trend in Ecuador rainfall and 2) a near-monopole tropical ocean SST anomaly associated with a rainfall anomaly of opposite sign over the Andes. These two additional modes provide incremental cross-validated

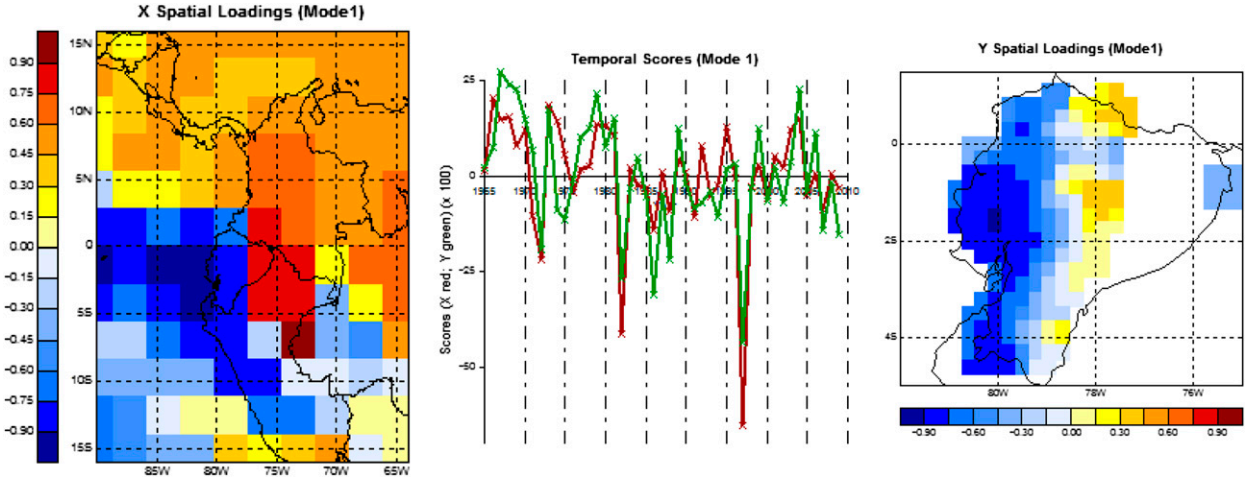

FIG. 17. Model-predicted and observed rainfall loading patterns and their time series for CCA mode 1 for the DJF rainfall prediction model using ECHAM4.5 simulations as the predictor, forced by observed DJF SST. The canonical correlation coefficient for mode 1 is 0.78.

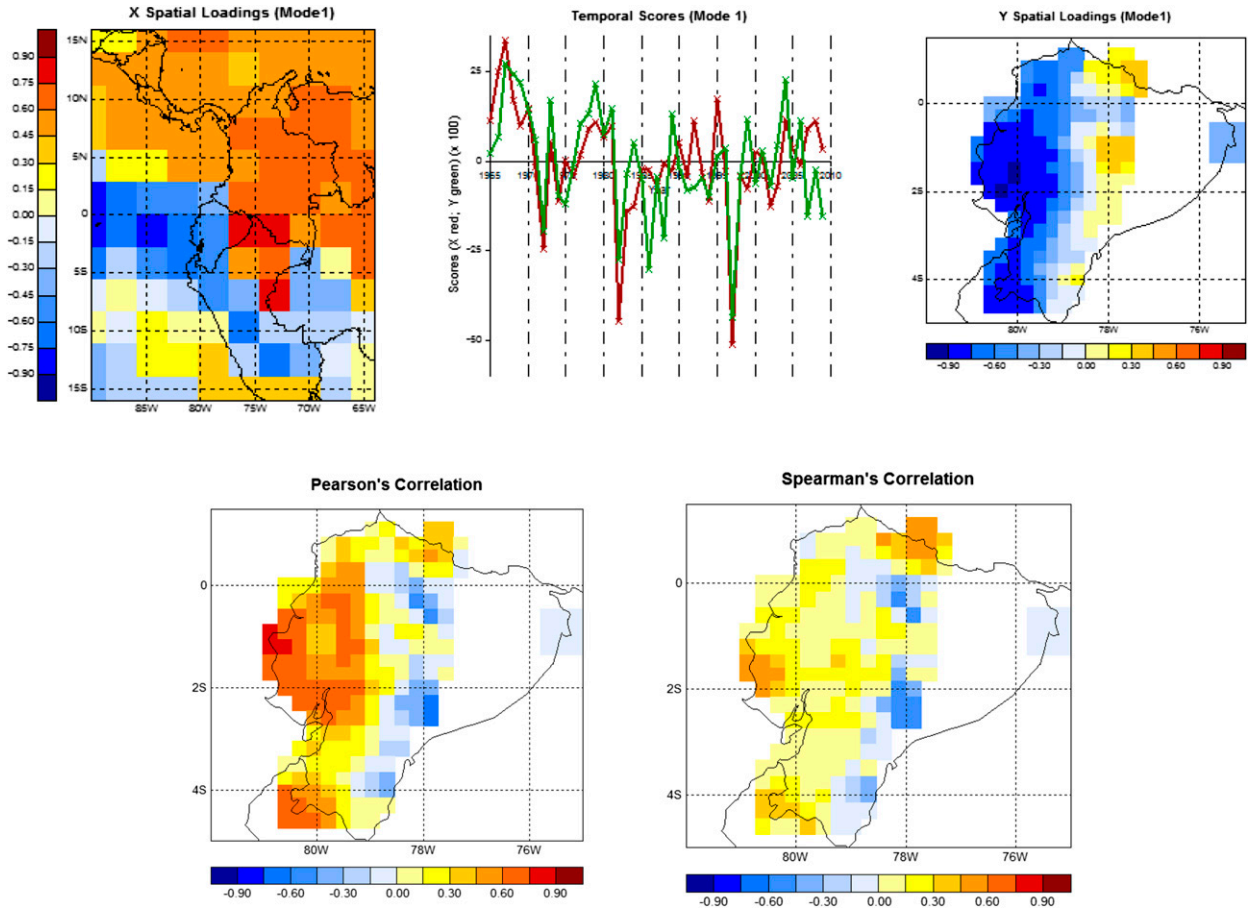

FIG. 18. Model-predicted and observed rainfall loading patterns and their time series for CCA mode 1 for the DJF rainfall prediction experiment using SST predicted from a November start time to force the ECHAM4.5 model. The canonical correlation coefficient for mode 1 is 0.68.

forecast skill over portions of Ecuador and are most important in regions of the country that lack a significant influence from ENSO.

SST in the tropical Atlantic Ocean does not contribute materially to prediction skill in this study, despite its role in governing Amazon rainfall variability, as seen in earlier studies (e.g., Moura and Shukla 1981; Vuille et al. 2000a,b). The dearth of rainfall predictands in eastern Ecuador undoubtedly suppresses the appearance of an Atlantic relationship, however. A mode relating an anomaly in much of the tropical Atlantic SST to a like-signed anomaly in mainly the coastal region of Ecuador may be linked to the ENSO mode, which involves a similar rainfall pattern in Ecuador, given that ENSO can affect the tropical Atlantic in its later stages (e.g., from DJF through March–May).

Results of the experiments for DJF are generally consistent with those for FMA and even slightly stronger because ENSO episodes often remain strong through DJF, but not necessarily FMA. Ecuadorian rainfall tends to be greater in FMA because of the climatological maximum in SST at that time in the eastern tropical Pacific, however. Greater ENSO strength does not guarantee greater SST anomaly strength in the far eastern tropical Pacific—especially in FMA (Table 2)—which ultimately governs Ecuador's rainfall west of the Andes.

Overall findings suggest that in most of Ecuador's coastal region seasonal rainfall predictions are skillful enough for climate services and, in particular, for warning systems to be able to substantially reduce economic losses associated with flood and drought. Such favorable predictive skills are found to a lesser extent in the Andean highlands, where there is a mixture of high and low skills over small distances (indicated clearly in individual station skills; not shown) because of the irregular terrain. In both coastal and Andean regions, much of the seasonal climate predictability arises as a result of the influence of ENSO, and thus years with the strongest ENSO conditions are expected to produce the strongest rainfall anomaly forecasts. Within the

subset of years without strong ENSO conditions, the signal-to-noise ratio is lower and the predictive skill indicated in a study such as this one would be lower. The large influence of the strongest El Niño years in the correlation skill is evidenced by greater values of correlation than of Spearman rank correlation, as noted particularly in Figs. 6, 14, 16, and 18.

Ecuador's INAMHI issues forecasts each month using the methods described here and also considers dynamical tools following the Latin American Observatory partnership method (Muñoz et al. 2010, 2012). Work in progress involves a similar study but using high-resolution dynamical model outputs (WRF and MM5) for both temperature and precipitation.

Acknowledgments. We are grateful for the helpful suggestions of the three anonymous reviewers. Research of the lead author was supported by a master fellowship by the National Secretary of Higher Education, Science, Technology and Innovation of Ecuador. This work was funded by a grant/cooperative agreement from the National Oceanic and Atmospheric Administration (NA05OAR4311004). The views expressed are those of the authors and do not necessarily reflect the views of NOAA or its subagencies.

APPENDIX

Cressman Objective Analysis to Create Gridded Data

The final data preparation step is application of an objective analysis to create a gridded rainfall dataset and to fill in missing data at stations having otherwise mainly complete data. Missing data at the 149 retained stations having no more than 25% of years missing for a given month were filled, where possible, by using a Cressman objective analysis (Cressman 1959). The analysis derives gridbox rainfalls by conducting several passes of a weighted averaging of the standardized precipitation anomalies at neighboring stations within a progressively decreasing radius of the grid box.

Although the objective-analysis algorithm generally follows the iterative scheme suggested by Cressman (1959), it includes a few modifications developed for the Observatorio Latinoamericano de Eventos Extraordinarios (OLE^2; Muñoz et al. 2010, 2012). The algorithm first computes standardized anomalies for each station with respect to its own observed climatological values. Then, three radii of influence are defined in terms of the maximum radius from a grid box to a station, from which a weight is assigned to the station's observation to estimate the gridbox value; stations beyond the radius of influence have zero weight. In this study, 15, 9, and 3 km were chosen as the successive radii of influence. The program then interpolates the station's standardized anomalies to a 30-km latitude–longitude grid for Ecuador, and the three passes are made through the grid at the consecutively smaller radii R to increase precision, using the correcting factor $C(i, j)$ for each iteration and grid point, given by

$$C(i,j) = \left(\frac{\sum_{s=1}^{n} W_s Q_s}{\sum_{s=1}^{n} W_s} \right),$$

where the weights are set to zero if the station is outside the radius of influence; otherwise, they are computed as

$$W_s = \frac{(R^2 - d_s^2)}{(R^2 + d_s^2)},$$

where d_s is the distance between each station and grid-point center. The correction factors are applied to all grid points before the next pass is made. Some grid boxes are not sufficiently close to stations having ample data, and their final rainfalls remain missing. Most grid boxes in the sparsely populated Amazon are in this missing category. It was decided not to include the single Galapagos Island grid square in the experiments.

The resulting 30-km grid has roughly 150 nonmissing squares over Ecuador. (The exact number varies slightly by month.) After the gridding process, stations originally having up to 25% missing have only up to 10% missing. In the CPT software used, grid points with up to 10% missing values are filled according to a user-selected algorithm. Here, we selected the option of a multiple regression that is based on neighboring grid points having nonmissing data.

REFERENCES

Aguilar, E., J. Sigró, and M. Brunet, 2009: RClimdex con funcionalidades extras de control de calidad: Manual de uso, versión 1.0 (RClimdex with extra quality control functionality: User manual, version 1.0). Universitat Rovira i Virgili Center for Climate Change (C3) Tech. Doc., 12 pp. [Available online at http://cmc.org.ve/descargas/Cursos/CRRH/Manual_rclimdex_extraQC.r.pdf.]

Anthes, R. A., and T. T. Warner, 1978: Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. *Mon. Wea. Rev.,* **106,** 1045–1078, doi:10.1175/1520-0493(1978)106<1045:DOHMSF>2.0.CO;2.

Barnett, T. P., and R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for U.S. surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.,* **115,** 1825–1850, doi:10.1175/1520-0493(1987)115<1825:OALOMA>2.0.CO;2.

Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J. Climate,* **7,** 1513–1564, doi:10.1175/1520-0442(1994)007<1513:LSSTCP>2.0.CO;2.

——, and H. M. van den Dool, 1993: A degeneracy in cross-validated skill in regression-based forecasts. *J. Climate,* **6,** 963–977, doi:10.1175/1520-0442(1993)006<0963:ADICVS>2.0.CO;2.

——, and T. M. Smith, 1996: Specification and prediction of global surface temperature and precipitation from global SST using CCA. *J. Climate,* **9,** 2660–2697, doi:10.1175/1520-0442(1996)009<2660:SAPOGS>2.0.CO;2.

——, M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11. Is our capability increasing? *Bull. Amer. Meteor. Soc.,* **93,** 631–651, doi:10.1175/BAMS-D-11-00111.1.

Cañadas, L., 1983: *El Mapa Bioclimático y Ecológico del Ecuador* (*Bioclimatic and Ecological Map of Ecuador*). Banco Central del Ecuador, 210 pp.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87,** 367–374, doi:10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Diaz, H. F., and V. Markgraf, Eds., 1992: *El Niño: Historical and Paleoclimatic Aspects of the Southern Oscillation.* Cambridge University Press, 476 pp.

Francou, B., M. Vuille, V. Favier, and B. Cáceres, 2004: New evidence for an ENSO impact on low latitude glaciers: Antizana 15, Andes of Ecuador, 0°28'S. *J. Geophys. Res.,* **109,** D18106, doi:10.1029/2003JD004484.

Glahn, H. R., 1968: Canonical correlation and its relationship to discriminant analysis and multiple regression. *J. Atmos. Sci.,* **25,** 23–31, doi:10.1175/1520-0469(1968)025<0023:CCAIRT>2.0.CO;2.

Horel, J., and A. Cornejo-Garrido, 1986: Convection along the coast of northern Peru during 1983: Spatial and temporal variation of clouds and rainfall. *Mon. Wea. Rev.,* **114,** 2091–2105, doi:10.1175/1520-0493(1986)114<2091:CATCON>2.0.CO;2.

Hotelling, H., 1936: Relations between two sets of variates. *Biometrika,* **28,** 321–377.

Korecha, D., and A. Barnston, 2007: Predictability of June–September rainfall in Ethiopia. *Mon. Wea. Rev.,* **135,** 628–650, doi:10.1175/MWR3304.1.

Kuo, Y.-H., J. B. Klemp, and J. Michalakes, 2004: Mesoscale numerical weather prediction with the WRF model. Presentations, *Symp. on the 50th Anniversary of Operational Numerical Weather Prediction,* College Park, Maryland, NCEP. [Available online at http://www.ncep.noaa.gov/nwp50/Presentations/Thu_06_17_04/Session_9/Kuo_50th_NWP/Kuo_50th_NWP.ppt.]

Michaelsen, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.,* **26,** 1589–1600, doi:10.1175/1520-0450(1987)026<1589:CVISCF>2.0.CO;2.

Moura, A. D., and J. Shukla, 1981: On the dynamics of droughts in northeast Brazil: Observations, theory and numerical experiments with a general circulation model. *J. Atmos. Sci.,* **38,** 2653–2675, doi:10.1175/1520-0469(1981)038<2653:OTDODI>2.0.CO;2.

Muñoz, Á. G., and Coauthors, 2010: An environmental watch system for the Andean countries: El Observatorio Andino. *Bull. Amer. Meteor. Soc.,* **91,** 1645–1652, doi:10.1175/2010BAMS2958.1.

——, and Coauthors, 2012: Risk management at the Latin American Observatory. *Risk Management—Current Issues and Challenges,* N. Banaitiene, Ed., InTech, 533–556. [Available online at http://cdn.intechopen.com/pdfs-wm/38982.pdf.]

Roeckner, E., K. and Coauthors, 1996: The atmospheric general circulation model ECHAM-4: Model description and simulation of present day climate. Max-Planck Institute for Meteorology Tech. Rep. 218, 90 pp. [Available online at http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/MPI-Report_218.pdf.]

Rossel, F., P. Goulven, and E. Cadier, 1999: Areal distribution of the influence of ENSO on the annual rainfall in Ecuador. *Rev. Sci. Eau,* **12,** 183–200.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006). *J. Climate,* **21,** 2283–2296, doi:10.1175/2007JCLI2100.1.

Thiaw, W. M., A. G. Barnston, and V. Kumar, 1999: Predictions of African rainfall on the seasonal timescale. *J. Geophys. Res.,* **104,** 31 589–31 597, doi:10.1029/1999JD900906.

Tippett, M. K., M. Barlow, and B. Lyon, 2003: Statistical correction of central southwest Asia winter precipitation simulations. *Int. J. Climatol.,* **23,** 1421–1433, doi:10.1002/joc.947.

——, L. Goddard, and A. G. Barnston, 2005: Statistical–dynamical seasonal forecasts of central-southwest Asian winter precipitation. *J. Climate,* **18,** 1831–1843, doi:10.1175/JCLI3371.1.

van den Dool, H. M., 1994: Searching for analogues, how long must we wait? *Tellus,* **46A,** 314–324, doi:10.1034/j.1600-0870.1994.t01-2-00006.x.

——, 2007: *Empirical Methods in Short-Term Climate Prediction.* Oxford University Press, 215 pp.

Vaughan, C., and S. Dessai, 2014: Climate services for society: Origins, institutional arrangements, and design elements for an evaluation framework. *Wiley Interdiscip. Rev.: Climate Change,* doi:10.1002/wcc.290, in press.

Vuille, M., R. S. Bradley, and F. Keimig, 2000a: Climate variability in the Andes of Ecuador and its relation to tropical Pacific and Atlantic sea surface temperature anomalies. *J. Climate,* **13,** 2520–2535, doi:10.1175/1520-0442(2000)013<2520:CVITAO>2.0.CO;2.

——, ——, and ——, 2000b: Interannual climate variability in the central Andes and its relation to tropical Pacific and Atlantic forcing. *J. Geophys. Res.,* **105,** 12 447–12 460, doi:10.1029/2000JD900134.

Wang, X. L., 2003: Comments on ''Detection of undocumented changepoints: A revision of the two-phase regression model.'' *J. Climate,* **16,** 3383–3385, doi:10.1175/1520-0442(2003)016<3383:CODOUC>2.0.CO;2.

——, 2008a: Accounting for autocorrelation in detecting mean shifts in climate data series using the penalized maximal *t* or *F* test. *J. Appl. Meteor. Climatol.,* **47,** 2423–2444, doi:10.1175/2008JAMC1741.1.

——, 2008b: Penalized maximal *F* test for detecting undocumented mean shifts without trend change. *J. Atmos. Oceanic Technol.,* **25,** 368–384, doi:10.1175/2007JTECHA982.1.

——, and Y. Feng, 2013: RHtestsV4 user manual. Environment Canada Science and Technology Branch Atmospheric Science and Technology Directorate Climate Research Division Res. Rep., 28 pp. [Available online at http://etccdi.pacificclimate.org/RHtest/RHtestsV4_UserManual_20July2013.pdf.]

——, Q. H. Wen, and Y. Wu, 2007: Penalized maximal *t* test for detecting undocumented mean change in climate data series. *J. Appl. Meteor. Climatol.,* **46,** 916–931, doi:10.1175/JAM2504.1.

Zhang, X., and Coauthors, 2005: Trends in Middle East climate extreme indices from 1950 to 2003. *J. Geophys. Res.,* **110,** D22104, doi:10.1029/2005JD006181.