RESEARCH ARTICLE

Preconditioning of the precipitation interannual variability in southern Mexico and Central America by oceanic and atmospheric anomalies

Alejandra Straffon[1,2] | Jorge Zavala-Hidalgo[1] | Francisco Estrada[1,3]

[1]Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de, Mexico City, Mexico

[2]Posgrado en Ciencias de la Tierra, Universidad Nacional Autónoma de, Mexico City, Mexico

[3]Institute for Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands

Correspondence
Alejandra Straffon, 1Centro de Ciencias de la Atm´osfera, Universidad Nacional Aut´onoma de M´exico, Mexico, Mexico.
Email: alejandra.straffon@atmosfera.unam.mx

Abstract

The interannual variability of precipitation in southern Mexico, northwestern Central America, and part of the northeastern tropical Pacific and its relation with previous oceanic and atmospheric anomalies are studied using the National Centers for Environmental Prediction Climate Forecast System Reanalysis data for 1979–2015. The rainy season over the study region was divided into three intraseasonal phases: before, during, and after the mid-summer drought. The interannual precipitation anomalies between the three phases are significantly correlated with each other, although there are years in which the precipitation anomalies for different phases differ in their sign. The interannual variability of the precipitation in this region is closely related with previous oceanic and atmospheric anomalies that occurred in the 60–31 days before the beginning of each phase. Multilinear regression models, between precipitation anomalies and previous anomalies of oceanic and atmospheric variables, explain 49, 60, and 55% of the variance of the precipitation inter-annual variability for the respective phase. Within a phase, in more than 76% of the cases, the estimated precipitation anomalies had the correct sign but in most cases their amplitude, irrespective of sign, were underestimated, as is expected from a regression model which aims to estimate the conditional mean. The more relevant processes associated with the interannual variability of precipitation in the study area are the westward displacement and strength-ening of the North Atlantic Subtropical High, the meridional displacement of the Intertropical Convergence Zone and the strength of the Caribbean Low-Level Jet. In general, the mechanisms that explain interannual variability of precipitation are the same through the phases. However, the magnitude of the anomalies and the size of the regions evolve in time in a particular way for each variable.

KEYWORDS

intraseasonal forecast, midsummer drought, precipitation interannual variability

1 | INTRODUCTION

Most of the agriculture in southern Mexico and Central America is strongly dependent on rainfall. Reliable advance warning about precipitation anomalies can benefit potable water supply, hydroelectrical power production, and flood risk management in the region.

The annual precipitation cycle over southern Mexico and Central America (SMCA) is characterized by a summer-dominant season that starts in early May and ends in late October or the beginning of November (Magaña et al., 1999). There is a climatological mid-summer drought (MSD), with rainfall ≈40% less in late July and early August than the relative maxima in June and September (Curtis and Gamble, 2008). The MSD over SMCA is the most extreme in the world (Curtis, 2002). Over the Pacific coast its spatial scale, coherence, and intensity (Curtis, 2002; Curtis and Gamble, 2008; Karnauskas et al., 2013) are greater than in adjacent land regions.

The bimodal distribution of the precipitation was considered to be caused by a local mechanism that involves lagged feedbacks between solar radiation, sea surface temperature (SST), and intraseasonal variations of the low-level convergent flow (MSD, Magaña et al. (1999)). In contrast, many studies attribute it to the dominant role of the westward displacement and strengthening of the North Atlantic Subtropical High (NASH) during July–August through its associated easterly trade winds (Romero-Centeno et al., 2007; Muñoz et al., 2008). Regionally, the changing patterns associated with the MSD are an intensification of the low-level easterlies, in particular, those crossing trough the Tehuantepec, Papagayo, and Panama mountain gaps (Romero-Centeno et al., 2007). The Caribbean Low-Level Jet (CLLJ, Amador (1998)) consists of a region with strong zonal winds at around 925 hPa extending from the Lesser Antilles to the Western Caribbean (Hidalgo et al., 2015). These wind patterns cool the SSTs in the Caribbean Sea (Mestas-Nuñez et al., 2007) and in the Eastern Pacific Warm Pool (Xie et al., 2005; Herrera et al., 2015). The wind patterns also displace low-level wind convergence areas away from the Pacific coasts (Small et al., 2007; Wang et al., 2007, 2008; Muñoz et al., 2008; Xie et al., 2008), prevent the southerly trade winds over the Eastern Pacific Warm Pool, and convection from reaching the SMCA coast (Romero-Centeno et al., 2003, 2007). Anomaly patterns of lower (higher)-than-normal SST (sea level pressure, SLP) over the Atlantic and higher (lower)-than-normal SST (SLP) over the Pacific drive an intensification of the CLLJ, these patterns are characteristic of droughts in the Central America Dry Corridor (Hidalgo et al., 2019). A positive correlation/anticorrelation between SST anomalies over the Tropical Atlantic, Eastern Pacific, and Western Pacific, and the summer precipitation in southern Mexico, was reported by Fuentes-Franco et al. (2018); they also evaluated the forecast of the seasonal precipitation via a multiple regression model.

The Western Hemisphere warm pool (WHWP) and its variability are related with precipitation through an increase in SST anomalies, a decrease in atmospheric SLP anomalies and an anomalous increase in atmospheric convection and cloudiness which reinforces SST anomalies (Wang and Enfield, 2001, 2003). During the boreal spring until fall, the WHWP is a source for the seasonal heating of the Walker and Hadley circulations in the Western Hemisphere which affects global and regional climate variability (Wang and Enfield, 2001, 2003).

The storm activity in the Caribbean and over the far-eastern Pacific shows a bimodal distribution that is in phase with the MSD (Curtis, 2002; Inoue et al., 2002). However, Amador et al. (2016a) found the opposite for the Caribbean Sea, where the tropical storm peak activity was identified during August. During the last period of the summer rainy season (late August, September, and October) the hurricane season in the Tropical North Atlantic and in the Eastern Tropical Pacific are more active, increasing the precipitation over the entire region (Amador et al., 2006).

The presence of the WHWP on both sides of the study region provide favourable conditions for tropical cyclone development. However, tropical storms over the Caribbean Sea behave different that those over the Eastern Tropical Pacific, important differences related with low-frequency modes of variability and the regional impacts of climate change may explain the observations (Amador et al., 2016a, 2016b).

The interannual variability of the rainy season over SMCA is determined by several modes of variability. In the boreal summer, during warm (cold) El Niño Southern Oscillation (ENSO) conditions the CLLJ is stronger (weaker) than normal (Amador, 2008). Also, during stronger (weaker) CLLJ the Intertropical Convergence Zone (ITCZ) is shifted southward (northward) from its average position, leading to a reduction (increase) in the precipitation, and thus affecting the MSD (Hidalgo et al., 2015). The Madden–Julian Oscillation (MJO) also modulates the summer rainfall over Mexico and Central America, with the positive phase of the local MJO tending to increase precipitation (Barlow and Salstein, 2006).

This study focuses on the precipitation over an area that includes part of southern Mexico, northwestern Central America and the northeastern tropical Pacific, see black rectangle (100–85°W, 10–18°N) in Figure 1. This area is characterized by a strong MSD signal that is
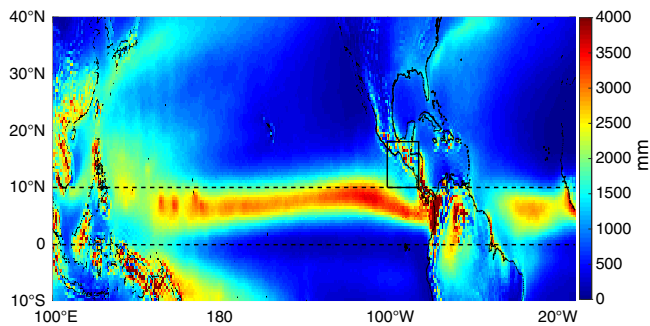
FIGURE 1 Mean accumulated precipitation from May to October CFSR data, 1979–2015. Rectangle indicates the study region (Southern Mexico, Central America, and a part of the northeastern tropical Pacific), segmented lines are the equator and the 10 N parallel. The figure limits correspond to the study domain [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 NCEP-CFSR data used in this study

Variable	Notation	Units
Precipitation rate	P	$kg/(m^2 s)$
Pressure reduced to MSL	PRMSL	Pa
Zonal component of wind at 10 m	U10	$m \cdot s^{-1}$
Meridional component of wind at 10 m	V10	$m \cdot s^{-1}$
Zonal component of wind at 850 mb	U850	$m \cdot s^{-1}$
Meridional component of wind at 850 mb	V850	$m \cdot s^{-1}$
Zonal component of wind at 200 mb	U200	$m \cdot s^{-1}$
Meridional component of wind at 200 mb	V200	$m \cdot s^{-1}$
Sea surface temperature	SST	K
Ocean heat content	OCNHEAT	J

highly correlated with the zonal wind component averaged within the central region of the northeastern tropical Pacific (Romero-Centeno *et al*., 2007), and is also associated with the strengthening of the Caribbean trade winds (indexed by the CLLJ) (Hidalgo *et al*., 2019). The method separates the summer rainy season into three phases, based on the climatic periods corresponding to before, during and after the MSD. The aim of the study is to use information of previous anomalies of oceanic and atmospheric variables to estimate, 1 month before the beginning of each phase, the interannual variability of precipitation anomalies over the region of interest. For each phase, an ensemble of multilinear regression models was performed, the sensitivity to data removal was assessed, and the ensemble's average was compared with the reanalysis data. Dynamical processes involved with wet and dry years over the study region are discussed for each phase.

2 | DATA AND METHODS

2.1 | Climate Forecast System Reanalysis

The analysis is based on data (Table 1) from the National Centers for Environmental Prediction Climate Forecast System Reanalysis (NCEP-CFSR), which was the first quasi-global, fully coupled atmosphere–ocean–land model used at NCEP for seasonal prediction (Saha *et al*., 2010, 2014). CFSR represents a clear improvement over the earlier reanalysis products, although biases remain, for example, mass balance of the atmospheric water component (Saha *et al*., 2010). Studies have shown improvements in CFSR daily precipitation statistics (Higgins *et al*., 2010) and surface heat fluxes and wind stress over

the ocean have smaller biases compared with previous NCEP-reanalysis (Xue *et al*., 2011).

The analysis covers the 1979–2015 period. CFSR ds093.1 covers the period 1979–2010 and CFSv2 ds094.1 the 2011–2015 period. The data have a $0.5° \times 0.5°$ spatial resolution for all variables excluding *precipitation rate* which has higher resolution ($0.3125° \times 0.3125°$ for CFSR and $0.2045° \times 0.2045°$ for CFSv2). For this study, precipitation rate was interpolated to the $0.5°$ resolution by linear interpolation and the 6-hr CFSR data were averaged to obtain daily mean values.

2.2 | Methods outline

A methodology to quantify, and also anticipate, the interannual variability of precipitation anomalies was developed for this study. This section briefly introduces the main steps followed, a detailed explanation of them is in their specific section. Steps 1–4 are presented as part of Section 2 and steps 5–6 are described in Section 3 of the manuscript.

1. Precipitation season over the study region was divided into three phases that correspond to the climatic periods: before, during, and after the MSD (see Section 2.3).
2. Correlation maps between precipitation anomalies and previous anomalies of atmospheric and oceanic variables were calculated. From these maps regions (boxes) of high interannual correlation/anticorrelation between anomalies of each variable were determined, considering reanalysis data of 60–31 days before the beginning of each phase. The 30 days mean was

computed to filter high-frequency variability. Annual time series of each variable were calculated within each region (box) for the three phases (see Section 2.4).

3. Multilinear regression models were developed considering all possible combinations between boxes (up to three boxes/variables in each model), but only those models that satisfy all statistical misspecification tests, and shown statistical significance at the 5% level were considered. Each phase has specific variables that participate more in models than the others, that is, those variables in specific box(es) appear more frequently in models that satisfy all the statistical requirements, so from now on, they are called *the most frequent variables*. The explained variance of the mean of the models with same number of variables/boxes was compared, in order to find out the optimal number of variables/boxes (see Section 2.5).

4. The robustness of the multilinear regression models to the input data was assessed considering the whole time series, removing 1 year (out-of-sample forecast) and dividing into halves (see Section 2.6).

5. Precipitation anomalies were predicted for all cases mentioned previously, and compared with reanalysis data (see Section 3.1).

6. The anomaly patterns and processes involved before and during each phase of the summer rainy season are discussed (see Section 3.2).

2.3 | Intraseasonal and interannual precipitation variability

In order to better understand the intraseasonal variability of precipitation over the study region, the rainy season was divided into three phases: the May–June precipitation, the MSD, and the September–October precipitation. Daily climatology of precipitation is used for the division in phases as follows: the date in which the mean value between the first (second) maximum and the relative minimum was considered to mark the beginning (end) of phase 2. Then, the dates of the beginning of phase 1, and the end of phase 3, were defined as those in which the 10-day low-pass filter reaches the climatological value of the seasonal minimum associated with the MSD (see Figure 2 and its legend).

Maps of precipitation anomalies for each phase of the rainy season relative to the mean of the period—24 May (onset of phase 1) to 22 October (end of phase 3)—are illustrated in Figure 3. Differences between the three phases are large in the equatorial and tropical Pacific, the central tropical Atlantic, the Gulf of Mexico, north western Mexico, and SMCA.
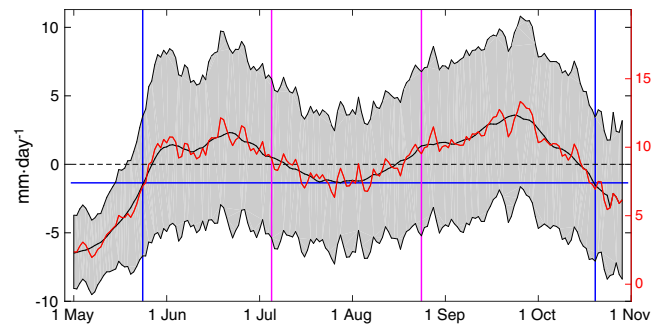

FIGURE 2 Daily precipitation (CFSR) and its anomalies relative to the period mean over the study region during 1979–2015 (*red*), 10-day low-pass filter (*black*) and one standard deviation of the anomalies signal (*grey shading*). Left vertical axis: anomaly; right vertical axis: absolute variation. Vertical blue lines: beginning (end) of phase 1 (phase 3). Horizontal blue line: relative minimum value. Magenta lines: beginning and end of phase 2 [Colour figure can be viewed at wileyonlinelibrary.com]
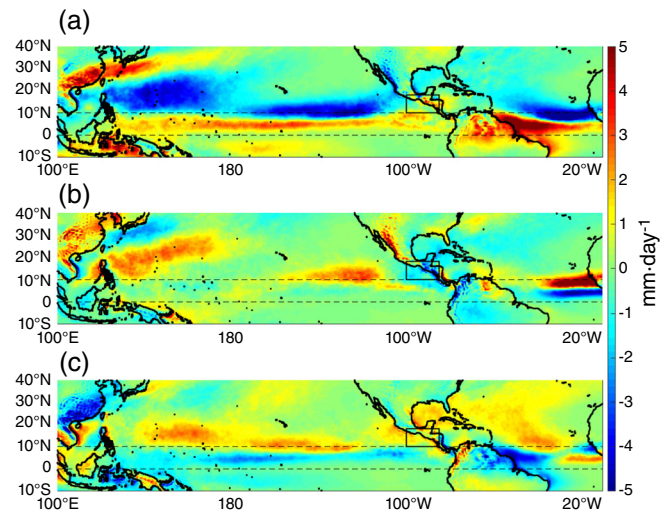

FIGURE 3 Precipitation anomalies relative to the precipitation mean of the period 24 May to 22 October for the period 1979–2015 for (a) phase 1 (24 May–6 July), (b) phase 2 (7 July–24 August) and (c) phase 3 (25 August–22 October). Black rectangle: Study region. Dotted lines: Equator and the 10 N parallel [Colour figure can be viewed at wileyonlinelibrary.com]

The meridional displacement of the ITCZ is identified through a positive precipitation anomaly between 0 and 10 N over the equatorial Pacific during phase 1; this becomes near mean during phase 2 and negative in phase 3 (Figure 3).

During phase 2, the precipitation areas associated with low-level wind convergence over the northeastern tropical Pacific are displaced away from the Pacific coasts, due to an intensification of winds that blows through the Isthmus of Tehuantepec and the Gulf of Papagayo (Romero-Centeno *et al.*, 2007). During phase

3, there are positive precipitation anomalies in both the Gulf of Mexico and the northeastern tropical Pacific.

The differences in precipitation patterns at global scale of the three phases suggest that the precipitation in the study region (black rectangle in Figure 3) during each phase is caused by different large-scale processes. Moreover, precipitation anomaly maps during periods of maximum precipitation (phases 1 and 3) are noticeably different each other, which supports the idea of separating the corresponding periods for precipitation studies.

Furthermore, the precipitation internannual variability in the study region was assessed by taking the precipitation anomaly for each phase of every year (Figure 4). The anomalies for each phase are significantly correlated between each other. Correlation decreases between phases 1 and 3, which implies the short-term nature of the memory of the system (see Table 2). However, there are years when the sign of the anomalies of the different phases differ.

2.4 | Correlation maps

Maps were made to illustrate the correlation of the precipitation anomaly time series for each phase over the
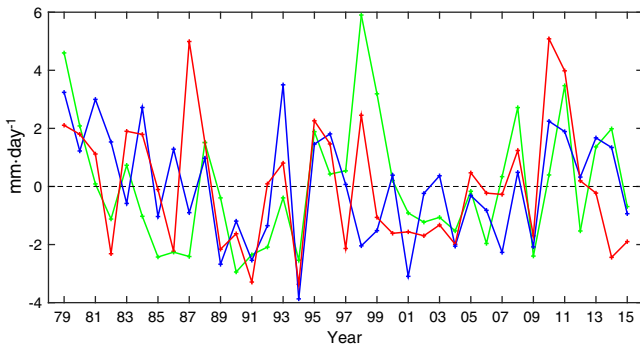
study region with the anomaly time series of all variables listed in Table 1. A period of 60–31 days before the beginning of each precipitation phase was considered for the analysis (Figures 5–7). Maps of variables with at least one box involved in a multilinear regression model (see next section) are shown in descending order, from the map of the most frequent variable in the multilinear regression (MLR) models to that with least frequent, for each phase. A large correlation (anticorrelation) means increased (decreased) rainfall in the study region. For each variable, a rectangular box indicates where correlations or anticorrelations are large.
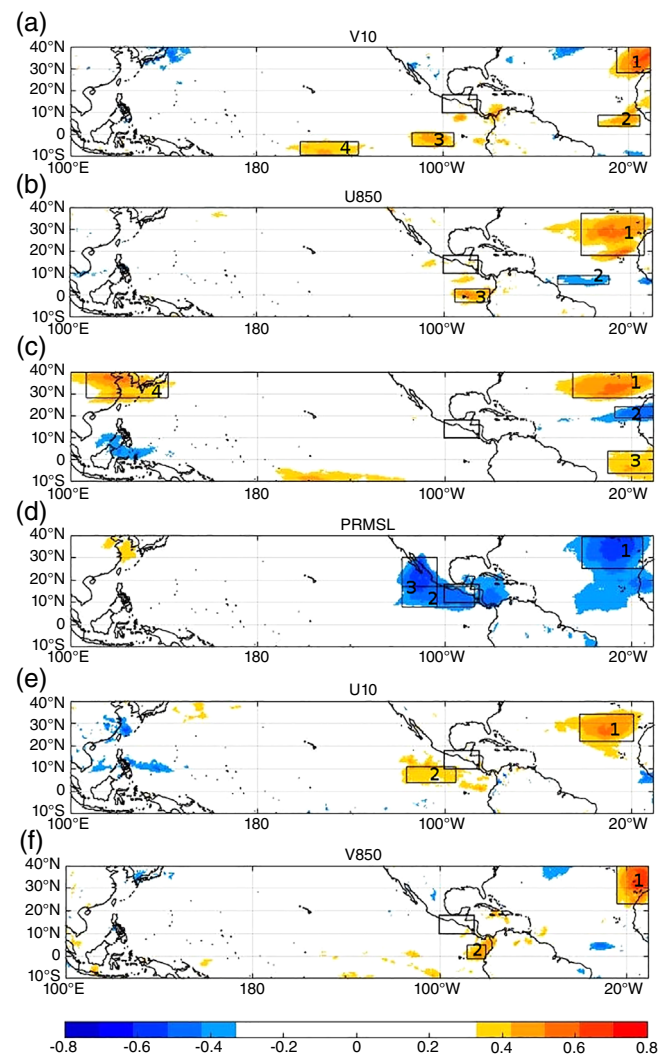

FIGURE 4 Annual precipitation anomalies for each phase in the study region: Phase 1 *(blue)*, phase 2 *(red)* and phase 3 *(green)* [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Coefficients of determination (R^2), Pearson and Spearman correlations between time series of precipitation anomalies

	R^2	Corr. (Pearson)	Corr. (Spearman)
Phase 1–2	0.23	0.48*	0.47*
Phase 2–3	0.19	0.44*	0.48*
Phase 1–3	0.12	0.34*	0.43*

*Statistical significance at the 5% level.

FIGURE 5 Correlation between precipitation anomalies for phase 1 in study region and the variables with at least one box involved in an MLR model (using information from the preceding 60–31 days). Each box indicates the region with the largest correlation, *(red)* or anticorrelation *(blue)*. Study region is indicated as a black rectangle without number [Colour figure can be viewed at wileyonlinelibrary.com]
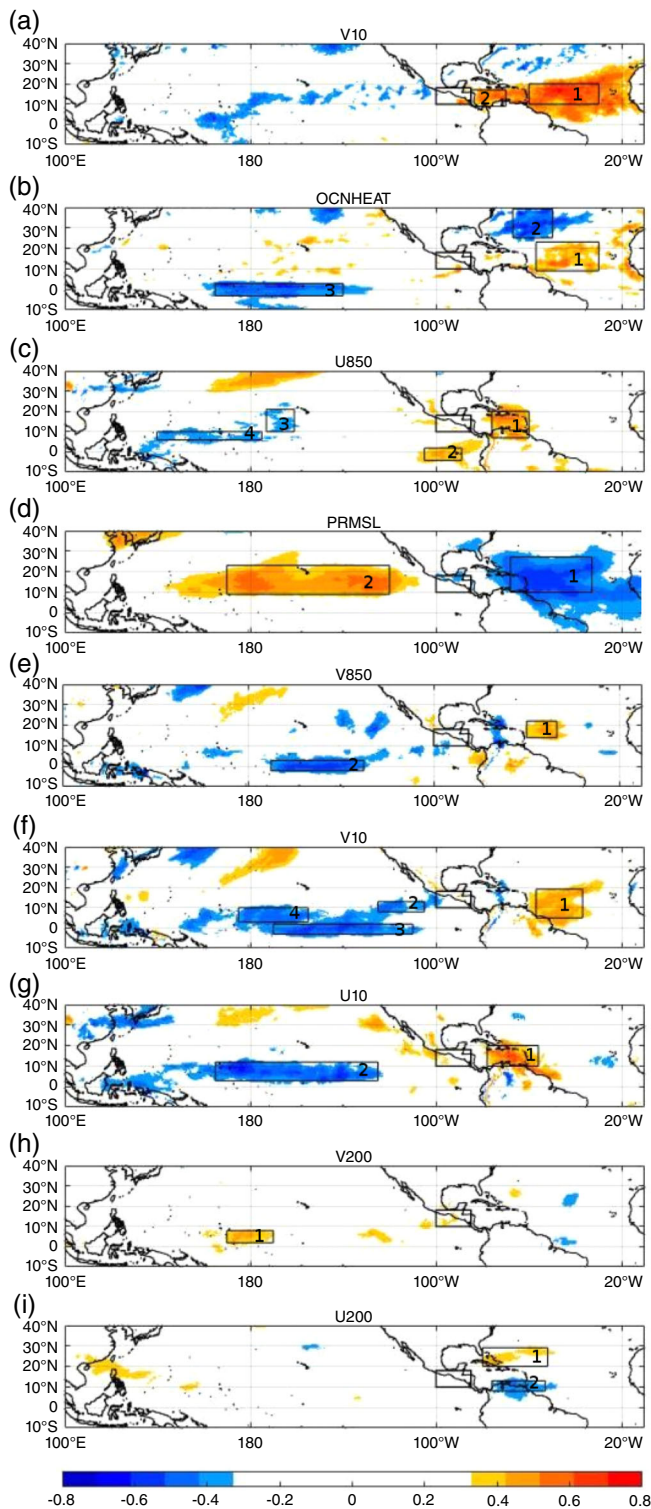
FIGURE 6   Same as in Figure 5 for phase 2 [Colour figure can be viewed at wileyonlinelibrary.com]

The variables and regions (boxes) that are the most frequent in each phase are different. In addition, for phase 2, all studied variables have at least one box
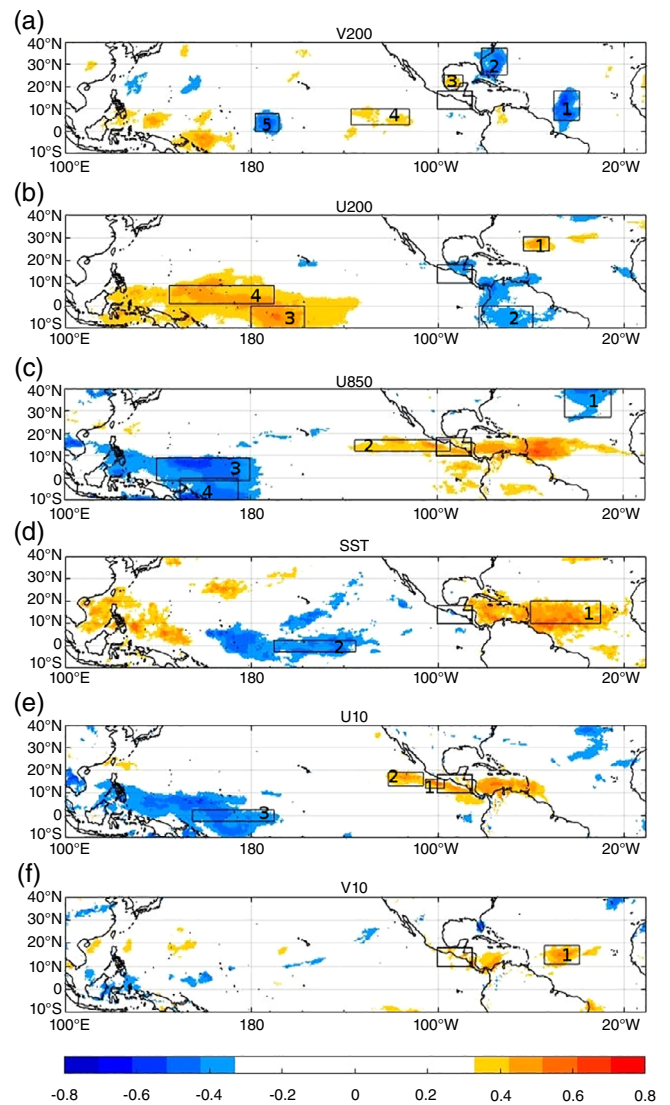

FIGURE 7   Same as in Figure 5 for phase 3 [Colour figure can be viewed at wileyonlinelibrary.com]

involved in a MLR model, whereas for phases 1 and 3 only six variables are involved.

2.5 | Multilinear regression models

For each phase of the summer rainy season, MLR models were developed for the precipitation anomalies within the study region. This study considers a set of atmospheric and oceanic variables averaged over the boxes with the larger correlation/anticorrelation. Boxes of each phase are shown in Figures 5–7, for each box of each variable was assigned an identification number. To develop an MLR model, all-time series of anomalies were normalized. For a particular year and phase, precipitation anomalies over the study region (P_t) are represented by

TABLE 3 Comparison of the number of MLR models and the coefficient of determination (R^2) of the mean time series considering two, three and two and three variables

Phase	No. (2 var)	R^2	No. (3 var)	R^2	No. (2 and 3 var)	R^2
1	44	0.57	53	**0.60**	97	0.60
2	91	0.65	61	**0.71**	152	0.68
3	56	0.56	38	**0.59**	94	0.58

Note: Highest R^2 values are indicated with bold.

$$P_t = a_0 + a_1(Var_1^x) + a_2(Var_2^y) + \cdots + a_n(Var_n^z) + e, \quad (1)$$

where $Var_i, \in \{1, ..., n\}$, are atmospheric and oceanic normalized variables averaged over the boxes x, y, and z (more than one box of the same variable is allowed). Coefficients a_i are estimated by ordinary least squares (OLS) and e is a Gaussian noise component with zero mean and constant variance.

All possible combinations considering two and three variables were explored. The inclusion of more than three variables in a model decreased the adjusted coefficient of determination (R_{adj}^2). Therefore, an algorithm was applied that ensures statistical significance at the 5% level for all variables involved in each linear regression. Finally, six statistical misspecification tests were used to assess the adequacy of the linear regression model: functional form (Ramsey RESET test), no autocorrelation (Durbin–Watson test), normality (Jarque–Bera test), homoscedasticity (White test), structural permanence (Quandt–Andrews test), and no multicollinearity (variance inflation factor). Only linear regressions that satisfied all tests, and also all variables that showed statistical significance, were considered for the following calculations.

In Table 3, the number of MLR models and the coefficient of determination (R^2) of the mean time series for the groups with different number of variables are compared. The group of MLR models with 2 and 3 variables has the largest number of MLR models. However, the explained variance of these models is equal or less than the group with three variables. So, a larger number of models does not imply higher explained variance (e.g., phase 2).

This study considers MLR models of three variables because the mean of this group has higher R^2 values than those for the 2 and 2 and 3 variable groups. Therefore, Equation 1 takes the form

$$P_t = a_0 + a_1(Var_1^x) + a_2(Var_2^y) + a_3(Var_3^z) + e. \quad (2)$$

The average of all MLR models of three variables explains between 59 and 71% of the interannual variability in precipitation anomalies over the study region when the complete time series of data (37-year) is used for the estimation.

2.6 | Sensitivity analysis

The sensitivity of the model to the amount of data was evaluated. For each year of the period, precipitation anomalies were reconstructed using the linear regression coefficients obtained when: (a) using the whole 37-year time series, (b) using 36 years of data to estimate the other year (out-of-sample forecast), and (c) using the first half (19 years, from 1979 to 1997) of the time series to estimate the second half (18 years, from 1998 to 2015) and vice versa.

The scatter plots between the coefficients from the 37-year and those from computations using 36-year indicate a robust model with small changes of the coefficients (Figure 8). Mean standard deviations are 0.0216, 0.0244, and 0.0242, for phases 1, 2, and 3, respectively. Coefficients are not evenly distributed over the axes, but all phases have coefficients a_0 around zero because anomalies are normalized.

When halves of time series are considered, the scatter plots show a different slope for each half (not shown). For the three phases, the linear regression of the first half (period 1979–1997) has a slope above or equal to one. In contrast, the second-half coefficients (period 1998–2015) have a slope below one.

The sensitivity analysis indicates that the reconstruction of each MLR model performs well even when only a part of the dataset is used; however, the best estimates are obtained for the longest time series.

3 | RESULTS

3.1 | Reconstruction of precipitation anomalies

For the three cases studied, that is, including data from all years, removing 1 year (out-of-sample forecast), and

with half of the years, results are similar. In almost all cases the right sign of the precipitation anomaly was obtained even though the mean of the MLR models does not reproduce the amplitude of the precipitation variability of the CFSR data, as would be expected from a
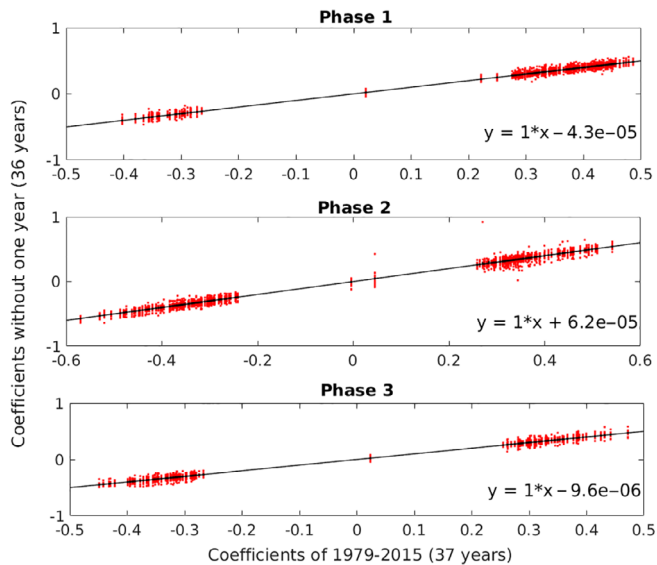

FIGURE 8 Scatter plot of MLR coefficients of the complete time series (37-year) versus the corresponding coefficients from time series without 1 year of all MLR models of each phase (*red dots*). Interannual mean of each coefficient (*black dots*) and its linear regression (*black line*) equation is shown for each phase [Colour figure can be viewed at wileyonlinelibrary.com]

regression model which aims to estimate the conditional mean (Figure 9a, c and e). This behaviour was similar for the three phases. Table 4 shows the standard deviation (*SD*) values as a measure of the data dispersion and the explained variance (measured by R^2) of the reconstructed time series. *SD* values do not depend on the amount of information as R^2 does, for example, phase 2 has maximum *SD* using halves of the time series.

The percentage of years with a correct anomaly sign is greater than 75% for phase 3 and much higher for the other two phases. Also, the number of years with precipitation absolute errors beyond 1 *SD* of the CFSR data time series for phase 3 is almost twice the number for either of the other phases (Table 5). For the more anomalous years the right sign is obtained, although larger errors are present (Figure 9).

3.2 | Anomaly patterns and processes

For each phase, the large-scale precipitation patterns are different (see Figure 3), which suggest that in each case different mechanisms modulate the precipitation. Moreover, the most frequent variables in MLR models differ between phases. Also, the high correlation/anticorrelation regions (boxes) change for a given variable. In this section, the anomaly patterns of each variable, before and during each phase, are analysed based on the composite of the anomalies for the 6 years with higher and lower precipitation over the study region.
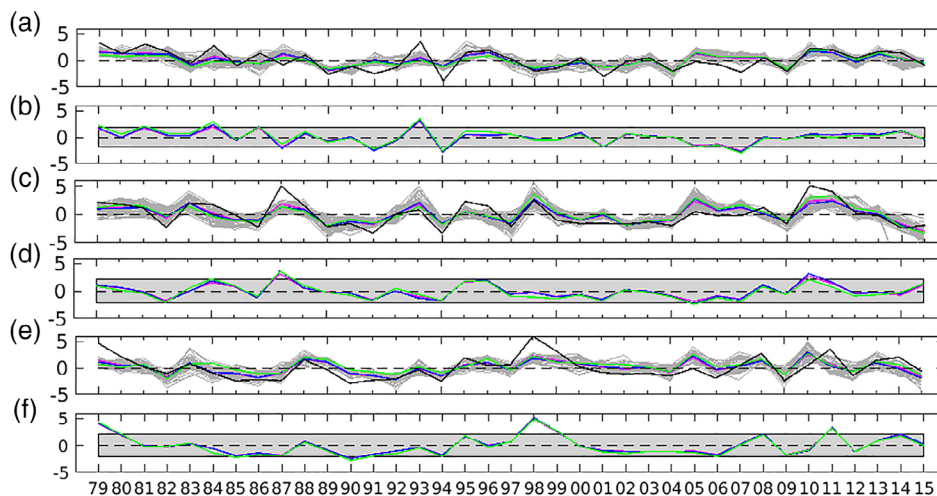

FIGURE 9 Precipitation anomalies (mm·day^{-1}) comparison between CFSR data (*black*) and the three cases of the MLR model: Mean using 37-year time series for the computation (*magenta*), mean using 36-year time series (*blue*) and mean using data of the first half to estimate the second half and vice versa (*green*) for phase 1, 2, and 3 (a, c, and e), MLRs (*grey shading*). Absolute errors comparison between CFSR data and the mean of the three cases of the MLR model for phase 1, 2, and 3 (b, d, and f), 1 *SD* of the CFSR anomalies time series (*grey shading*). The MLR series were obtained as the mean of all combinations of the three variables [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 4 Standard deviation (SD) and coefficients of determination (R^2) respect to CFSR data of the precipitation anomalies time series

	CFSR data	Using 37-year time series		Without 1 year		Using half time series	
	SD	SD	R^2	SD	R^2	SD	R^2
Phase 1	1.88	0.60	0.60	0.57	0.49	0.55	0.45
Phase 2	2.16	0.67	0.71	0.67	0.60	0.74	0.60
Phase 3	2.12	0.57	0.59	0.58	0.55	0.56	0.35

	Years with the correct sign		Years with errors >1 SD	
	N^o	%	N^o	%
Phase 1	28	76	8	22
Phase 2	33	89	5	14
Phase 3	31	84	5	14

TABLE 5 Years with the correct sign and years with errors >1 *SD* in the precipitation anomaly, when a 36-year time series is used to estimate the other year

3.2.1 | Phase 1

For phase 1, which contains the first relative maximum in precipitation, the key variables that have relevant information from previous periods are the low-level winds (V10 and U850), zonal wind at 200 mb, and PRMSL. Moreover, most of the highly correlated regions are located over the Atlantic Ocean (Figure 5).

Before phase 1, precipitation spatial patterns appear similar between the composite of years with precipitation above and below the mean (Figure 10a and f). For the composite of wet years (left side in Figure 10), previous reduction of the anticyclonic circulation over the North Atlantic through a weakening of Atlantic westerlies at mid-latitudes and easterlies in the tropics (Figure 10c) is noticeable. PRMSL weakening over the North Atlantic (Figure 10e), in particular, over the Gulf of Mexico and the Caribbean Sea, reduce the pressure gradient with respect to the Pacific basin, and thus the funnelling of the trade winds through the Isthmus of Tehuantepec and the Papagayo gap weaken. This signal is evident in the positive low-level wind anomaly over the Pacific coast in front of the Isthmus of Tehuantepec and Papagayo gap in Figure 10b and (c). While zonal wind at 200 mb (U200) increase its speed over the tropical Pacific, particularly over the northeastern tropical Pacific (Figure 10d). The opposite patterns occur for the composite of dry years.

Moreover, for the composite of wet years, during the previous period, positive anomalies of V10 over the equatorial eastern Pacific (Figure 10b) associated with a northward shift of the ITCZ, which (during the phase) bring intense convection close to the study region. The opposite occurs for the composite of dry years, when anomalous weak meridional low level winds (V10) over

the northeastern tropical Pacific (Figure 10(g)) shift the ITCZ position further south than normal (Figure 11a and f).

During phase 1, the large-scale patterns observed 60–31 days before the phase remain, but anomalies are smaller. Regions over North America and adjacent oceans with PRMSL anomalies reduce its area and magnitude (Figure 11e and j). However, anomalous westerlies (easterlies) in the composite of wet years (dry years) (Figure 11c and h), associated with the interannual variability of the CLLJ, intensify over the Caribbean Sea and extends towards the northeastern tropical Pacific.

In summary, negative PRMSL anomalies over the North Atlantic are associated with weaker anticyclonic circulation at low-level. Those conditions maintain the convergence areas and precipitation closer to the eastern Pacific coasts during the phase (Figure 11a) in comparison with the composite of dry years (Figure 11f).

3.2.2 | Phase 2

For phase 2, the oceanic variables SST and OCNHEAT, as well as the meteorological variables U850 and PRMSL, are crucial in anticipating variability in precipitation (Figure 6). Notice that the precipitation composites of wet and dry years are similar before the beginning of this phase (Figure 12a and f), although very different during the phase (Figure 13a and f).

The composite maps based on the 6 years with the highest (lowest) precipitation in the study region shows previous warm (cold) SST anomalies (more than 1 K from 1979–2015 mean) in the tropical North Atlantic, except for the Gulf of Mexico and north to 20 N. Also,
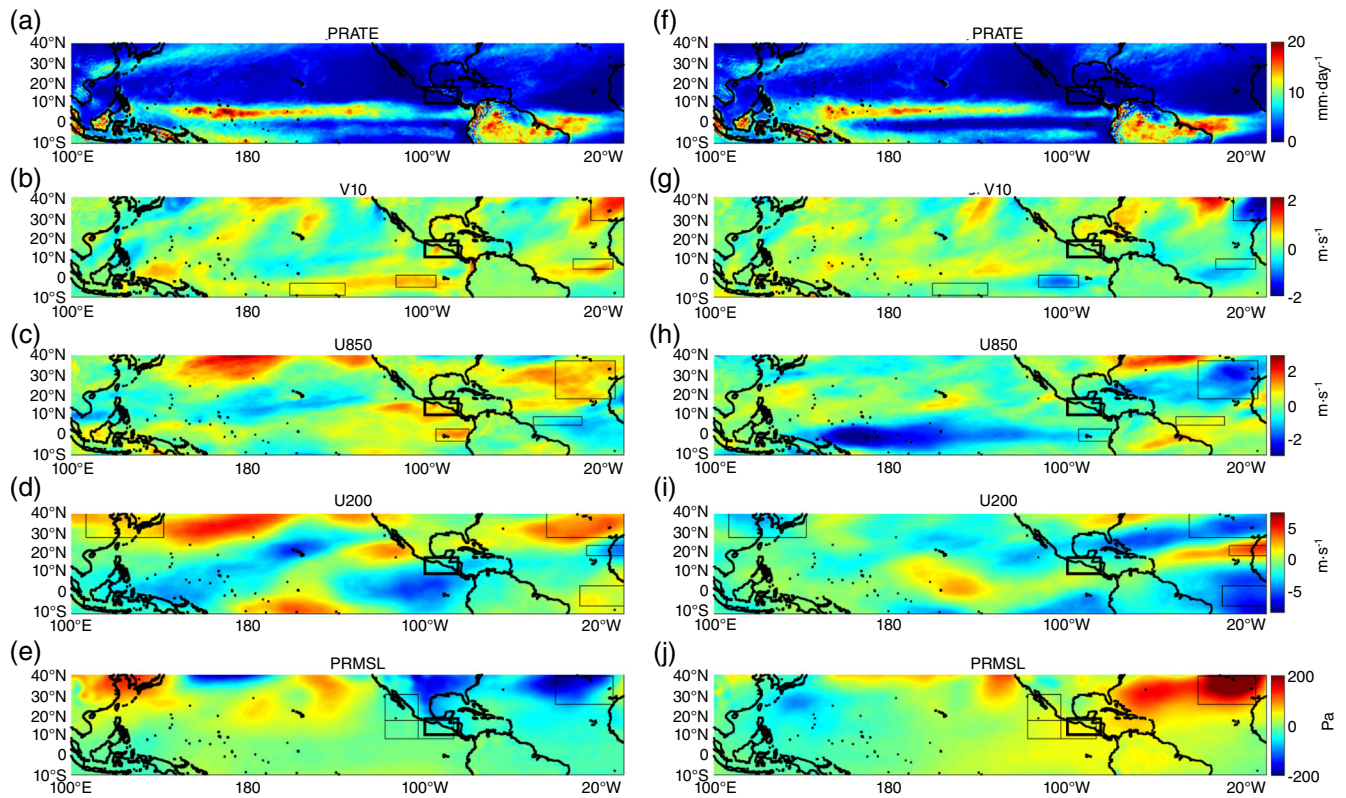
FIGURE 10 Composite maps of precipitation (*first row*) and composite anomaly maps (*second to fifth row*) of the 6 years with the highest positive (*left*) and negative (*right*) precipitation anomaly in study region (*black bold rectangle*) for the most frequent variables (*boxes indicate regions with the largest interannual correlation*) in the MLR model during 60–31 days before the start of phase 1 [Colour figure can be viewed at wileyonlinelibrary.com]
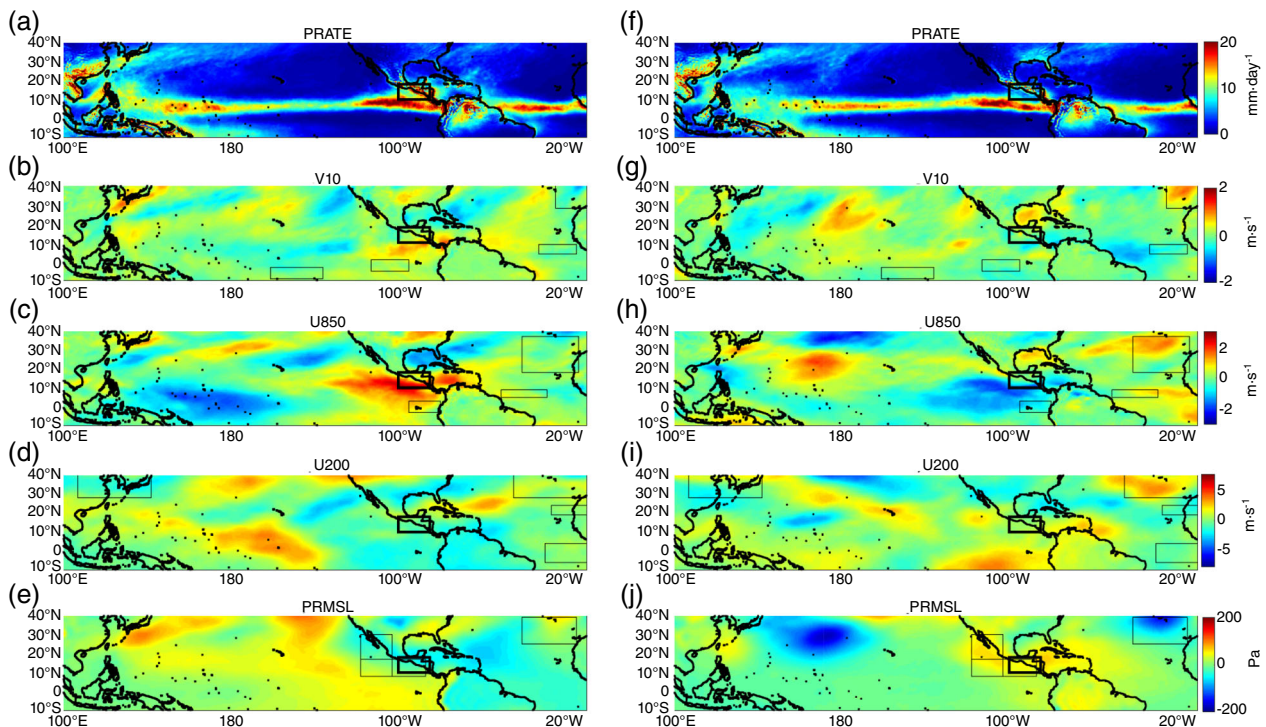

FIGURE 11 Composite maps of precipitation (*first row*) and composite anomaly maps (*second to fifth row*) of the 6 years with the highest positive (*left*) and negative (*right*) precipitation anomaly in study region (*black bold rectangle*) for the most frequent variables (*boxes indicate regions with the largest interannual correlation*) in the MLR model during phase 1 [Colour figure can be viewed at wileyonlinelibrary.com]
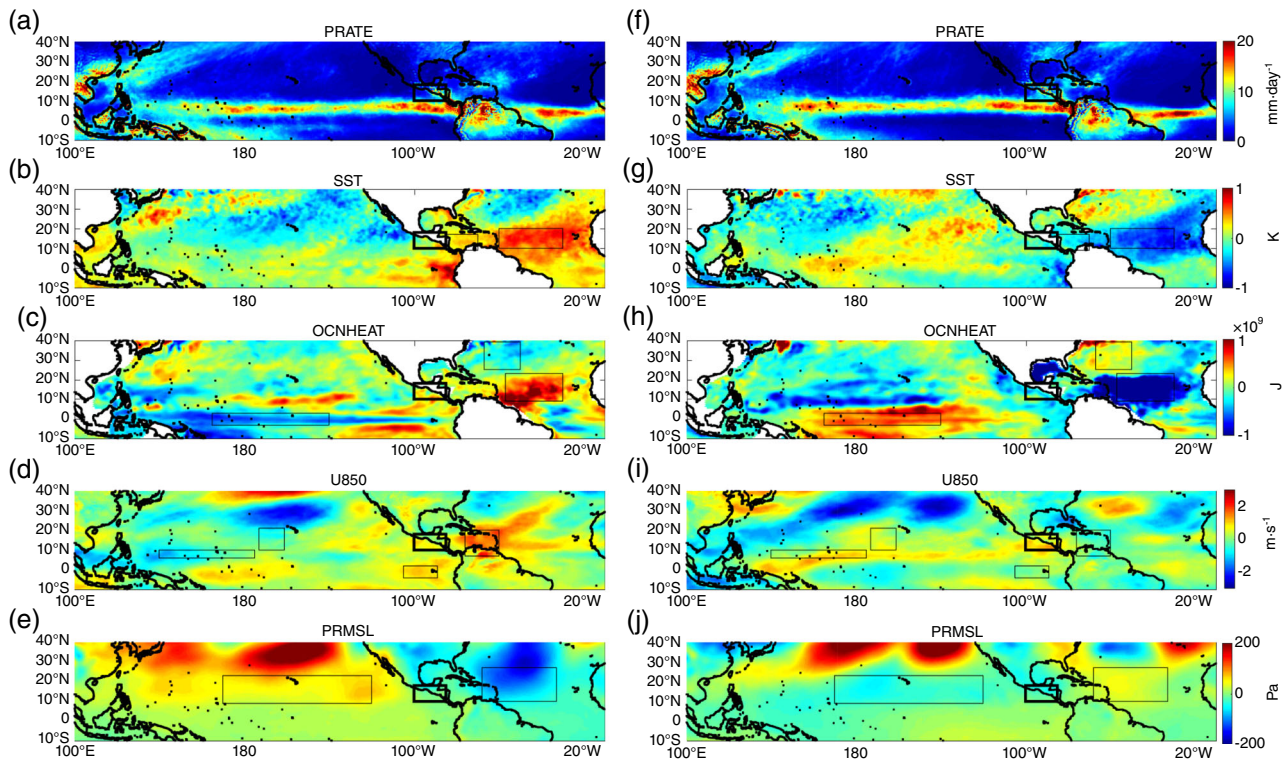
FIGURE 12 Composite maps of precipitation (*first row*) and composite anomaly maps (*second to fifth row*) of the 6 years with the highest positive (*left*) and negative (*right*) precipitation anomaly in study region (*black bold rectangle*) for the most frequent variables (*boxes indicate regions with the largest interannual correlation*) in the MLR model during 60–31 days before the start of phase 2 [Colour figure can be viewed at wileyonlinelibrary.com]
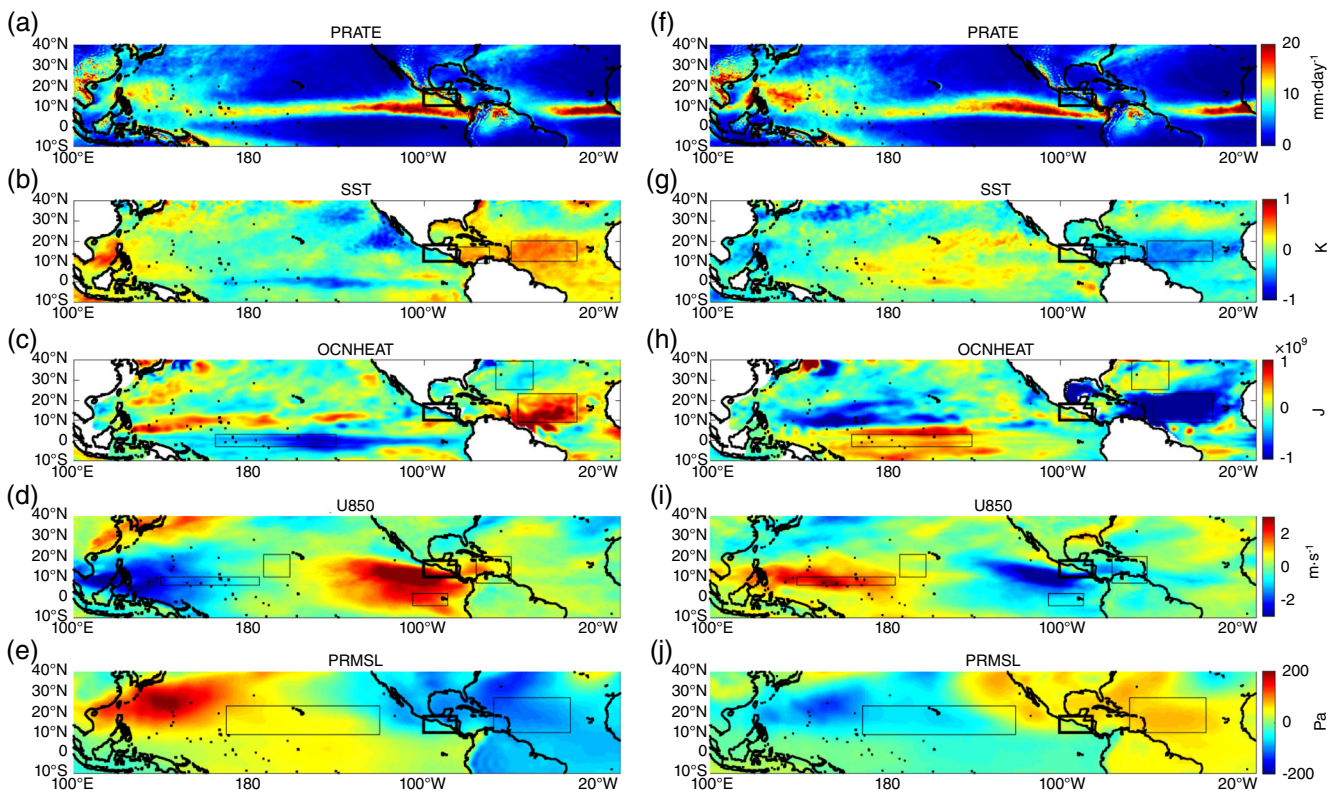

FIGURE 13 Composite maps of precipitation (*first row*) and composite anomaly maps (*second to fifth row*) of the 6 years with the highest positive (*left*) and negative (*right*) precipitation anomaly in study region (*black bold rectangle*) for the most frequent variables (*boxes indicate regions with the largest interannual correlation*) in the MLR model during phase 2 [Colour figure can be viewed at wileyonlinelibrary.com]
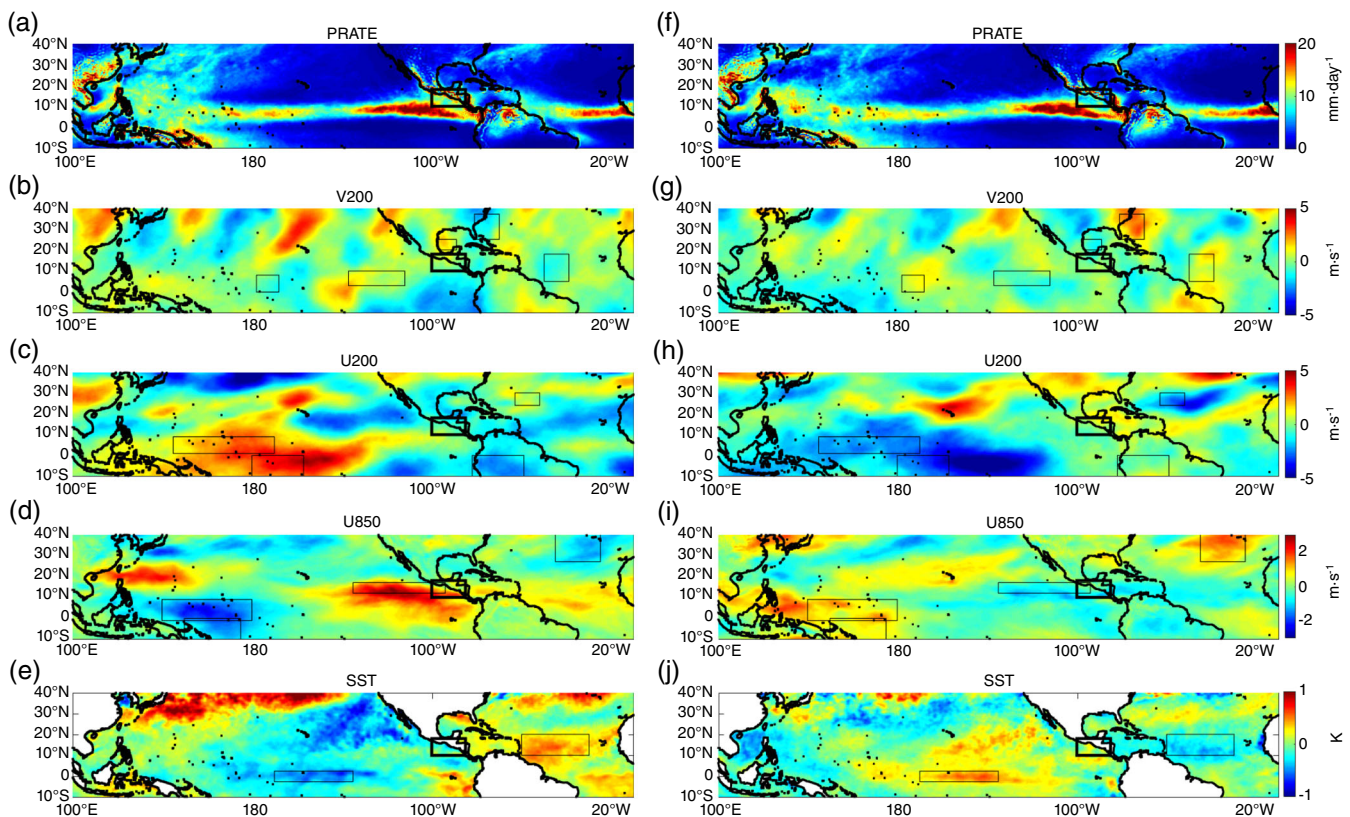
FIGURE 14 Composite maps of precipitation (*first row*) and composite anomaly maps (*second to fifth row*) of the 6 years with the highest positive (*left*) and negative (*right*) precipitation anomaly in study region (*black bold rectangle*) for the most frequent variables (*boxes indicate regions with the largest interannual correlation*) in the MLR model during 60–31 days before the start of phase 3 [Colour figure can be viewed at wileyonlinelibrary.com]

positive SST anomalies are found in the equatorial Pacific off the coasts of South America (Figure 12b). Spatial patterns of OCNHEAT (Figure 12c and h) are similar to those associated with SST; however, a in the OCNHEAT anomaly map cold tongue in front of the coast of equatorial South America appears in the composite of wet years and extends over the entire Pacific ocean.

The composite of wet years shows a weakening of easterlies (U850) over the Caribbean Sea and northern Venezuela due to a weakening in PRMSL over the same region (Figure 12d and e). These zonal low-level wind anomalies reduce wind speed over Caribbean Sea and a weakening of the CLLJ. The configuration of atmospheric and oceanic anomalies is consistent with those observed during phase 1.

During phase 2 (Figure 13), oceanic variables and PRMSL maintain almost the same anomaly patterns than those of the previous 60–31 days. However, PRMSL anomalies have a westward shift and an increase in area as a result of the displacement of the NASH. All continental America between 10 S and 40 N, almost the whole Atlantic and along the Pacific coasts of North America, there is an increase in PRMSL as a result of the

displacement of the NASH. In the composites (Figure 13e and j) those large regions of anomalies are organized as a dipole; for the composite of wet years, the Atlantic Ocean, the Americas, and northeastern tropical Pacific have negative anomalies, while the northwestern and Central Pacific have positive anomalies. The opposite pattern is observed in the composite of years with precipitation below mean.

As in phase 1, during phase 2, strong low-level positive (negative) wind anomalies (U850) are found over the northeastern tropical Pacific, and thus a diminish (increase) in the displacement of convergence areas away from the eastern Pacific coasts (Figure 13d and i).

3.2.3 | Phase 3

Upper winds (U200 and V200), U850, and SST were the most useful variables for anticipating the anomalies of phase 3, which includes the second relative maximum in precipitation (Figure 7).

For the composite of wet years in SMCA, progressive strengthening of the Walker circulation over the tropical
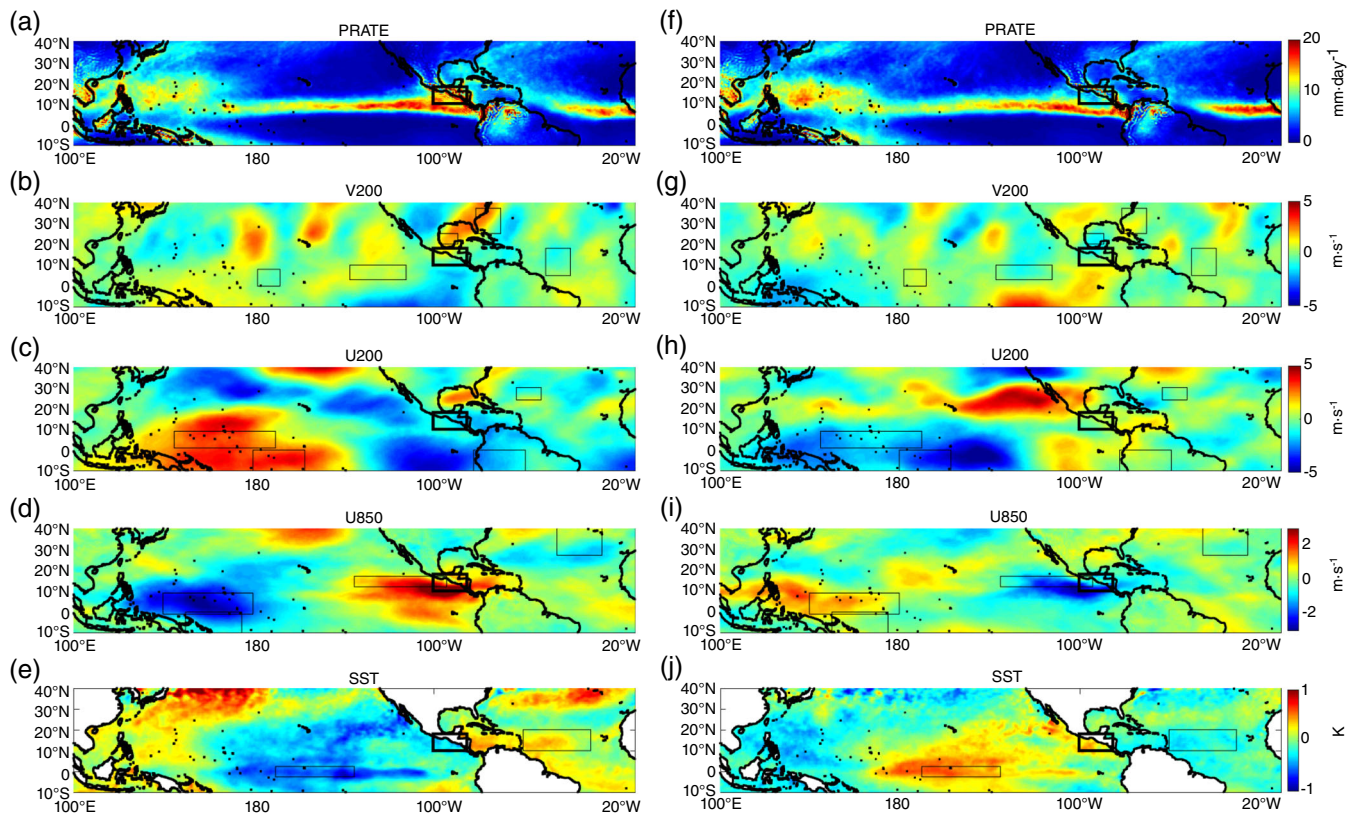
FIGURE 15 Composite maps of precipitation (*first row*) and composite anomaly maps (*second to fifth row*) of the 6 years with the highest positive (*left*) and negative (*right*) precipitation anomaly in study region (*black bold rectangle*) for the most frequent variables (*boxes indicate regions with the largest interannual correlation*) in the MLR model during phase 3 [Colour figure can be viewed at wileyonlinelibrary.com]

Western Pacific is deduced by anomalous U200 and U850 patterns. Anomalous low-level easterlies at around 150E (Figure 14d), associated with anomalous westerlies at 200 mb (Figure 14c) produce an intensification over the western side of the Walker cell. Also, an increase of zonal wind divergence at 850 mb over the central tropical Pacific, between 170 and 140 W, and the corresponding convergence at 200 mb, is associated with low SSTs in that region for wet years, see Figure 14c–e, respectively. The opposite patterns are observed in the composite of dry years.

In the North Atlantic basin, for the composite of wet years, warmer SSTs are also observed; however, their magnitude decreases with respect to previous phases (Figure 14e).

During phase 3 (Figure 15), for the composite of wet years, spatial patterns associated with an intensification of the western side of the Walker cell intensify over the tropical Pacific Ocean (Figure 15c,d). A large region over the tropical Pacific (170 E–100 W) shows lower SSTs (Figure 15e).

As in previous phases, for the composite of wet years, during this phase low-level westerly anomalies increase

over the northeastern tropical Pacific, the Caribbean Sea and continental masses between them (Figure 15d). This low-level wind configuration tends to maintain convergence areas close to the study region. The opposite is shown for the composite of dry years.

4 | DISCUSSION AND CONCLUSIONS

Most of the previous studies about the bimodal distribution of precipitation over the study region and its surroundings have focused on the whole summer season or in the MSD. To the best of our knowledge, the separation into phases (according to the interannual variability) is a novelty in this study.

The physical processes associated with the phenomena observed during wet years are as follows: the ITCZ meridional displacement is northward than normal (Hidalgo *et al.*, 2015) which is observable before the beginning of phase 1 and is a result of a cold tongue in the equatorial Pacific (characteristic of La Niña pattern). Hidalgo *et al.* (2019) showed a strong correlation between

ENSO, CLLJ, and precipitation in northwestern Central America that is consistent with previous studies and the results of this study. Wang (2005) has explained positive SST anomalies over the tropical North Atlantic and the WHWP as a result of decreased subsidence over the subtropical North Atlantic and the weaker north-east trades off Africa, which are associated with a progressive strengthening of the Walker circulation over the tropical Western Pacific. The opposite occurs for the composite of dry years.

Results presented in this study indicate the relative importance of particular variables and regions to anticipate precipitation anomalies over the study region for each phase of the summer rainy season. Interannual variability of the NASH and its implications appear to be a key determinant of precipitation anomalies during phases 1 and 2. Positive PRMSL anomalies over the tropical and mid-latitude North Atlantic are associated with an intensified anticyclonic circulation at low-level. An intensification in low-level wind speed tends to increase evaporation and reduce SSTs over the Atlantic, and this also increases the pressure gradient with respect to the Pacific basin. Such configuration is associated with the westward displacement of the convergence areas, away from the Pacific coasts, thereby, reducing precipitation over the study region. Hidalgo *et al.* (2019) found anomaly patterns of SST and SLP for dry episodes in the Pacific slope of Central America, with anomalous intense easterly flow in the Pacific coast of Central America and stronger CLLJ, that are consistent with the results presented in this study. During phase 3, SSTs show a similar pattern than that observed in phase 2, but the anomalies are smaller. During all phases, the composite of years with precipitation above (below) normal implies a large-scale configuration that is associated with weak (strong) easterlies over SMCA and adjacent oceans.

A computation of the number of hurricanes over the northeastern tropical Pacific and the Caribbean Sea (considering tropical storms and hurricanes of all intensities) in the area of study and its vicinity (5° around) was performed, results show that the total number of hurricanes (normalized by the number of days in each phase) during the period of interest follow a bimodal distribution in both oceans, which means a relative reduction in hurricane activity during the MSD. In relation to the interannual variability, when considering the number of hurricanes (normalized) in the composites of wet and dry years over the Caribbean Sea, there were more hurricanes during wet years (8, 4, 4) than in dry years (3, 3, 2), for the three phases, over the northeastern tropical Pacific the number of events was too small to make any conclusion. The relative contribution of cyclones to the interannual variability of precipitation in the study

region should be addressed by other studies since it was not conclusive base only on the number of events.

It is remarkable that different subsets of information, that is, 37 and 36 years and the halves of the time series, reproduce closely the time series of precipitation anomalies for each phase of the rainy season, although the explained variance (measured by R^2) decreases when the amount of input information is reduced (see Table 4).

When considering an out-of-sample forecast, the average of the group of linear regression models obtained for each phase explains 49, 60, and 55% of the variance for the three respective phases. However, differences between the coefficients of the first and second halves of the time series are probably related with multidecadal variability and/or climate change. A further comparison with other reanalyses, seasonal forecast and observational data sets using the proposed methodology will be desirable.

In 76, 89 and 84% of all cases for the respective phases (when considering an out-of-sample forecast), the estimates had the correct sign but in most cases the amplitude of the anomalies, irrespective of the sign, were underestimated, as would be expected from a regression model which aims to estimate the conditional mean. The MSD period (phase 2) had the greatest number of years with the correct sign and the lowest number of years with errors larger than one standard deviation.

Previous to phase 1, north Atlantic atmospheric conditions were more relevant in the statistical analysis than the anomalies on the Pacific basin. This was observed for both cases, when the precipitation was extremely high or low. In contrast, for phase 3 the previous anomalies in the Pacific Ocean, mainly in the equatorial Pacific, have a larger influence in the statistical analysis, with higher (lower) precipitation in the study area when La Niña (El Niño) patterns are observed. However, the SST anomalies in the tropical Atlantic remain important for this phase. Previous anomalies for phase 2 are distributed between the Pacific and the Atlantic basins. SST and OCNHEAT anomalies in the equatorial Atlantic resulted very important, as well as equatorial Pacific OCNHEAT. Zonal winds in the Caribbean Sea and in both the western (5 –10 N) and eastern tropical Pacific (2 S–4 N) have also influence for phase 2.

In general, we found that the mechanisms that explain precipitation anomalies are the same through the phases. For example, for wet years, low-level positive wind anomalies over the northeastern tropical Pacific, the Caribbean Sea, and continental masses between them tend to maintain convergence areas close to the study region, which is consistent with negative anomalies of PRMSL. However, the magnitude of the anomalies and the size of the regions evolve in time (phases) in a particular way for each variable.

The proposed methodology could help predict precipitation anomalies 1 month ahead, or more, on the basis of data that are available in almost real time. Our results indicate that the reconstruction of those anomalies performs well even in the absence of some data; however, best estimates are obtained using the longest time series.

Due to the socioeconomic benefits of obtaining advanced information about precipitation variability different efforts seek to generate forecasts on global or regional scale and there are several studies to better understand the physical phenomena and patterns on the ocean and the atmosphere associated with the interannual variability of precipitation (Fuentes-Franco et al., 2018). This work contributes to this collective effort through the suggestion of separating the rainy season in SMCA into three phases, a methodology based on multiple linear models, and the consideration of including meteorological and ocean variables averaged over a period to filter the high-frequency variability. In this case, the CFSR reanalysis was used but the methodology could be used with other reanalysis such as the ECMWF and with observations. It could also be combined with other methods and evaluate whether combining them can improve the seasonal precipitation forecast.

In particular, these results help to understand the interannual variability of the precipitation of southern Mexico and northwestern Central America and may help in the development of seasonal forecast of precipitation for this region. This methodology can also be applied to other regions of the world. The processes proposed here for each phase are based on the statistical analysis of the data, further research is needed in order to increase the understanding on the physical processes underlying the phenomena.

ORCID
Jorge Zavala-Hidalgo 🄳 https://orcid.org/0000-0002-2737-434X

REFERENCES

Amador, J.A. (1998) A climatic feature of the tropical Americas: the trade wind easterly jet. *Topicos Meteorologicosy Oceangraficos*, 5 (2), 91–102.

Amador, J.A. (2008) The intra-Americas sea low-level jet overview and future research. *Annals of the New York Academy of Sciences*, 69(1146), 153–188.

Amador, J.A., Alfaro, E.J., Lizano, O.G. and Magaña, V.O. (2006) Atmospheric forcing of the eastern tropical Pacific: a review. *Progress in Oceanography*, 69, 101–142.

Amador, J.A., Durán-Quesada, A.M., Rivera, E.R., Mora, G., Sáenz, F., Calderón, B. and Mora, N. (2016a) The easternmost tropical Pacific. Part II: seasonal and intraseasonal modes of atmospheric variability. *Revista de Biología Tropical*, 64(Suppl. 1), S23–S57.

Amador, J.A., Rivera, E.R., Durán-Quesada, A.M., Mora, G., Sáenz, F., Calderón, B. and Mora, N. (2016b) The easternmost tropical Pacific. Part I: a climate review. *Revista de Biología Tropical*, 64(Suppl. 1), S1–S22.

Barlow, M. and Salstein, D. (2006) Summertime influence of the madden-Julian oscillation on daily rainfall over Mexico and Central America. *Geophysical Research Letters*, 33, L21708.

Curtis, S. (2002) Interannaul variability of the bimodal distribution of summertime rainfall over Central America and tropical storm activity in the far-eastern Pacific. *Climate Research*, 22, 141–146.

Curtis, S. and Gamble, D.W. (2008) Regional variations of the Caribbean midsummer drought. *Theoretical and Applied Climatology*, 94, 25–34.

Fuentes-Franco, R., Giorgi, F., Pavia, E.G., Graef, F. and Coppola, E. (2018) Seasonal precipitation forecast over Mexico based on a hybrid statistical-dynamical approach. *International Journal of Climatology*, 38, 4051–4065.

Herrera, E., Magaña, V. and Caetano, E. (2015) Air-sea interactions and dynamical processes associated with the midsummer drought. *International Journal of Climatology*, 35, 1569–1578.

Hidalgo, H.G., Alfaro, E.J., Amador, J.A. and Bastidas, A. (2019) Precursors of quasi-decadal dry-spells in the Central America dry corridor. *Climate Dynamics*, 53, 1307–1322.

Hidalgo, H.G., Durán-Quesada, A.M., Amador, J.A. and Alfaro, E.J. (2015) The Caribbean low-level jet, the inter-tropical Convergence zone and precipitation patterns in the intra-Americas sea: a proposed dynamical mechanism. *Geografiska Annaler: Series A, Physical Geography Geografiska Annaler: Series A, Physical Geography*, 97, 41–59.

Higgins, R.W., Kousky, V.E., Silva, V.B.S., Becker, E. and Xie, P. (2010) Intercomparison of daily precipitation statistics over the United States in observations and in NCEP reanalysis products. *Journal of Climate*, 23, 4637–4650.

Inoue, M., Handoh, I.C. and Bigg, G.R. (2002) Bimodal distribution of tropical cyclogenesis in the Caribbean: characteristics and environmental factors. *J ournal of Climate*, 15, 2897–2905.

Karnauskas, K. B., Seager, R., Giannini, A., and A.J. Busalacchi, A. (2013). A simple mechanism for the climatological midsummer drought along the Pacific coast of Central America. *Atmósfera*, 26(25):261–281.

Magaña, V., Amador, A. and Medina, S. (1999) The midsummer drought over Mexico and Central America. *Journal of Climate*, 12, 1577–1588.

Mestas-Nuñez, A.M., Enfield, D.B. and Zhang, C. (2007) Water vapor fluxes over the intra-Americas sea: seasonal and interannual variability and associations with rainfall. *Journal of Climate*, 20, 1910–1922.

Muñoz, E., Busalacchi, A.J., Nigam, S. and Ruiz-Barradas, A. (2008) Winter and summer structure of the Caribbean low-level jet. *Journal of Climate*, 21, 1260–1276.

Romero-Centeno, R., Zavala-Hidalgo, J., Gallegos, A. and O'Brien, J.J. (2003) Isthmus of Tehuantepec wind climatology and ENSO signal. *Journal of Climate*, 16, 2628–2639.

Romero-Centeno, R., Zavala-Hidalgo, J. and Raga, G.B. (2007) Mid-summer gap winds and low- level circulation over the eastern tropical Pacific. *Journal of Climate*, 20, 3768–3784.

Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Golberg, M., et al. (2010) The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91, 1015–1057.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Becker, E., et al. (2014) The NCEP climate forecast system version 2. *Journal of Climate*, 27, 2185–2208.

Small, R.J.O., Szoeke, S.P.D. and Xie, S.P. (2007) The central American midsummer drought: regional aspects and large-scale forcing. *Journal of Climate*, 20, 4853–4873.

Wang, C. (2005) ENSO, Atlantic climate variability, and the Walker and Hadley circulations. In: Diaz, H.F. and Bradley, R.S. (Eds.) *The Hadley Circulation: Present, Past and Future, Volume 21 of Advances in Global Change Research*. Dordrecht: Springer, pp. 173–202.

Wang, C. and Enfield, D.B. (2001) The tropical Western hemisphere warm pool. *Geophysical Research Letters*, 28(8), 1635–1638.

Wang, C. and Enfield, D.B. (2003) Further study of the tropical Western hemisphere warm Pool. *Journal of Climate*, 16, 1476–1493.

Wang, C., Lee, S.K. and Enfield, D.B. (2007) Impact of the Atlantic warm pool on the summer climate of the Western hemisphere. *Journal of Climate*, 20, 5021–5040.

Wang, C., Lee, S.K. and Enfield, D.B. (2008) Climate response to anomalously large and small Atlantic warm pools during the summer. *Journal of Climate*, 21, 2437–2450.

Xie, S.P., Okumura, Y., Miyama, T. and Timmermann, A. (2008) Influences of Atlantic climate change on the tropical Pacific via the central American isthmus. *Journal of Climate*, 21, 3914–3928.

Xie, S.P., Xu, H., Kessler, W.S. and Nonaka, M. (2005) Air-sea interaction over the eastern Pacific warm pool: gap winds, thermocline dome, and atmospheric convection. *Journal of Climate*, 18, 5–20.

Xue, Y., Huang, B., Hu, Z., Kumar, A., Wen, C., Behringer, D. and Nadiga, S. (2011) An assessment of oceanic variability in the NCEP climate forecast system reanalysis. *Climate Dynamics*, 37, 2511–2539.