ORIGINAL ARTICLE

Climate Resilience and Sustainability
Open Access
Interdisciplinary Approaches Towards Solutions for Climate Change
RMetS

Forecasting annual maximum water level for the Negro River at Manaus

Amulya Chevuturi[1,2,3] | **Nicholas P. Klingaman**[1,2] | **Conrado M. Rudorff**[4] |
Caio A. S. Coelho[5] | **Jochen Schöngart**[6]

[1] National Centre for Atmospheric Science, University of Reading, Reading, UK

[2] Department of Meteorology, University of Reading, Reading, UK

[3] UK Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford, UK

[4] National Center for Monitoring and Early Warning of Natural Disasters (CEMADEN), SP, Brazil

[5] Centre for Weather Forecast and Climate Studies (CPTEC), National Institute for Space Research (INPE), Cachoeira Paulista, SP, Brazil

[6] Coordenação de Dinâmica Ambiental (CODAM), Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Brazil

Correspondence
Amulya Chevuturi, UK Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford OX10 8BB, UK.
Email: amuche@ceh.ac.uk

Abstract

More frequent and stronger flood hazards in the last two decades have caused considerable environmental and socio-economic losses in many regions of the Amazon basin. It is therefore critical to advance predictions for flood levels, with adequate lead times, to provide more effective and earlier warnings to safeguard lives and livelihoods. Water-level variations in large, low-lying, free-flowing river systems in the Amazon basin, such as the Negro River, follow large-scale precipitation anomalies. This offers an opportunity to predict maximum water levels using observed antecedent rainfall. This study aims to investigate possible improvements in the performance and extension of the lead time of existing operational statistical forecasts for annual maximum water level of the Negro River at Manaus, occurring between May and July. We develop forecast models using multiple linear regression methods, to produce forecasts that can be issued in March, February and January. Potential predictors include antecedent catchment rainfall and water levels, large-scale modes of climate variability and the long-term linear trend in water levels. Our statistical models gain one month of lead time against existing models for same skill level, but are only moderately better than existing models at similar lead times. All models lose performance at longer lead times, as expected. However, our forecast models can issue skilful operational forecasts in March or earlier. We show the forecasts for the Negro River maximum water level at Manaus for 2020 and 2021.

1 | INTRODUCTION

The Amazon is the largest river basin in the world, draining about one-sixth of global freshwater to the ocean (Callède et al., 2010). It is also one of the few remaining networks of free-flowing large rivers on Earth (Grill et al., 2019). The Amazonian floodplains have been long settled and used by indigenous populations, as the plains provide essential ecosystem services and natural resources for agriculture, fishing, livestock rearing and forest management for subsidence and trade (Junk et al., 2014). Most settlements and cities in the Amazon basin have been built along the large rivers, which provide water for domestic use and irrigation, an important network for transport and recently are increasingly used for hydro-power production. Water evaporated from the Amazon Basin provides freshwater to other parts of South America through moisture transport and eventual precipitation (Drumond et al., 2014).

Variations in the annual hydrological cycle over the Amazon result from the complex dynamics of the climate system, at interannual (e.g. Xavier et al., 2010) or decadal (e.g. Marengo, 2004) time scales. Variations in rainfall patterns, and thus river discharge, are strongly influenced by sea surface temperature (SST) anomalies in either Equatorial Pacific (e.g. Jiménez-Muñoz et al., 2016) or the tropical Atlantic (e.g. Espinoza et al., 2014), or in both (e.g. Barichivich et al., 2018). The detailed processes that influence hydrological oscillations are summarized in Betts et al. (2009), Costa et al. (2009), Marengo et al. (2009) and Nobre et al. (2009). Extreme droughts and floods in the Amazon are caused by atmospheric circulation anomalies that drive deficient or excessive rainfall and late or early onsets of the rainy seasons (Espinoza et al., 2014; Marengo et al., 2012, 2013). Over the last two decades, the hydrograph amplitude of some Amazonian rivers, between the wet and dry seasons, has increased significantly (Gloor et al., 2013), related to changes in climate (Barichivich et al., 2018) and catchment land use (Dias et al., 2015). Future projections suggest further increases in mean and maximum river discharge for large river basins in western Amazonia (Sorribas et al., 2016), however, with large uncertainties related to regional variability and land-use changes.

Natural disasters have gained attention as one of the main challenges global society faces. The 2030 Agenda for Sustainable Development recognized the urgent need for action and set objectives and targets aimed at disaster risk reduction. Forecasting and early warning systems (e.g. CPRM's Negro River Flood Warning system for Manaus; Maciel et al., 2020) are essential to generate and communicate information to effectively prepare for natural disasters, especially for managing high-risk regions (Pappenberger et al., 2015). Flood pulse is a naturally occurring phenomenon for the Amazon rivers, which is important for regional ecology (Junk et al., 2011) and supports the local population's socio-economic activities such as fisheries, forestry and transport (Langill and Abizaid, 2020). However, extreme flood events in the Amazonian rivers endanger human lives, properties and infrastructure (e.g. transport and energy systems; Marengo and Espinoza, 2016). For example, the 2012 extreme flood event damaged neighbourhoods, businesses, roads, bridges, ports etc., and affected more than 29,000 people in Manaus (Marengo et al., 2013). Floods also cost billions in damages to local communities and the Brazilian government (Dolman et al., 2018), lead to loss of lives and livelihoods and force people to evacuate their residences (Schöngart and Junk, 2007). In urban areas, floods increase disease risk, as people suffer prolonged exposure to water contaminated by sewage, industrial and domestic waste due to the lack of basic sanitation and sewage treatment. In rural regions, severe flood events cause losses of crops and livestock.

Recent extreme floods (2009, 2012, 2013, 2014, 2015, 2019 and 2021) have affected hundreds of thousands of people in urban and rural regions of the Amazon Basin, who mostly live along large rivers or in the várzea floodplains (Junk et al., 2011), resulting in major social problems, economic losses and impacts on regional ecosystems (Gloor et al., 2015). Under-developed rural regions (Marengo et al., 2013), increased rural–urban migration (Guedes et al., 2009) and unplanned growth of urban areas has led to deficient infrastructure, inadequate social and health services, pollution, developments in flood hazard areas (Santos et al., 2017; Dolman et al., 2018; Mansur et al., 2018), which significantly exacerbate flood risk. Studies have devised methods for flood probability and risk mapping for the Amazon region (Lima et al., 2015; de Andrade et al., 2017). However, most towns in the Amazon Basin still lack integrated flood risk management plans and effective actions to assess the implications of the changes in risk and develop disaster preparation, response, recovery and mitigation strategies (Dolman et al., 2018). The severe impacts of recent extremes have revealed the need for timely, economically sustainable and socially acceptable actions, such

CHEVUTURI ET AL.

Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change   Open Access   RMetS

3 of 17

as risk assessments, management of human settlements in flood-prone areas and the introduction of early warning systems. A skilful flood forecasting system, at seasonal lead times, is an important component of the flood risk management actions to improve preparedness to extreme events.

Large Amazonian rivers tend to have annual, predictable high-amplitude flood pulses (Junk et al., 2011), due to the seasonality of precipitation in their large catchments (Salati and Marques, 1984). Most Amazonian floodplains become inundated from January–March (Papa et al., 2010), from which the water drains into the river systems slowly over the following months (de Paiva et al., 2013). Most of the rivers in the Amazon network reach flood levels during May–July and low levels during September–November, with variations in the average timing of peak/low discharge across the Amazon basin (Junk et al., 2011). The large drainage basins integrate precipitation variability; the river valley topography and wetlands attenuate and delay the flood wave. In the Central Amazon region, the maximum water level coincides with the beginning of the dry season, as the water needs 2–3 months to flow the several hundred kilometres from the headwaters to this region (Schöngart and Junk, 2007). Thus, interannual variability in the maximum water levels, in these free-flowing rivers, results from rainfall variability over the catchment regions in the months prior to the peak water level. Such regularity and temporal predictability enables statistical seasonal forecast models to predict the magnitude of hydrological peak water levels, which have high interannual variation (Schöngart and Junk, 2007).

Rainfall over the Amazon region is modulated by the South Atlantic Convergence Zone (e.g. Carvalho et al., 2011), which in turn is modulated seasonally by large-scale climate variability in the Atlantic (e.g. Yoon and Zeng, 2010; Gloor et al., 2013) and Pacific (e.g. Fu et al., 2001; Liebmann and Marengo, 2001). These teleconnections between large-scale modes of climate variability and Amazonian rainfall ultimately influence Amazon river discharge (Richey et al., 1989), and thus offer additional potential predictability for Amazon basin terrestrial water (De Linage et al., 2014). Studies such as Towner et al. (2020) show that the floods in Amazon can be attributed to large-scale coupled ocean–atmosphere modes of variability. Thus, statistical seasonal forecast models using state-of-the-art gridded rainfall datasets, in addition to other predictors such as the large-scale modes of variability, may improve flood forecasts compared to existing forecasts.

Currently, the Brazilian Geological Survey (CPRM) uses simple statistical models to forecast maximum water level for the Negro River at Manaus (Figure 1a) at seasonal lead times (earliest at end of March and latest at end of May; Maciel et al., 2020). More complex dynamical hydrological models, such as the Global Flood Awareness System
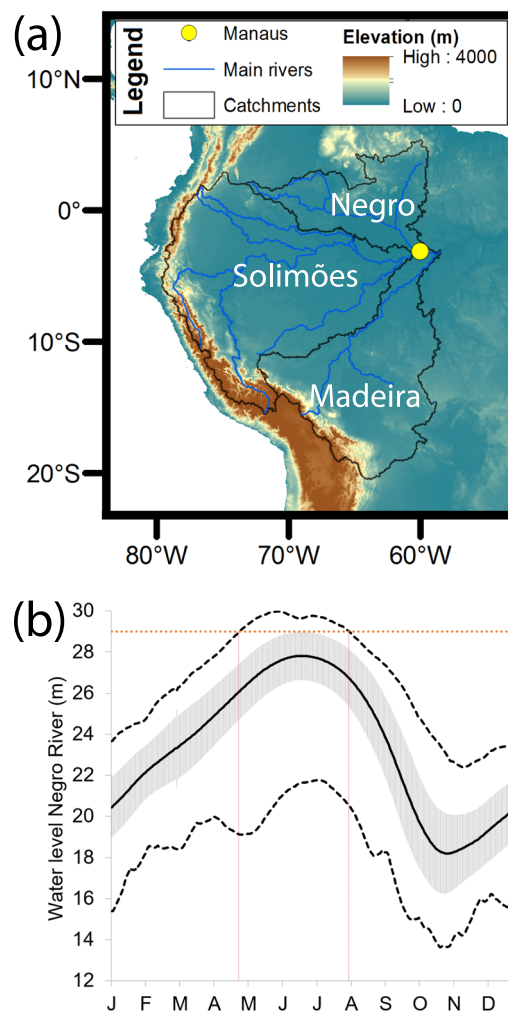

FIGURE 1 (a) The Negro River and surrounding rivers (blue lines) and catchment areas (regions bounded by black lines) contributing to the river water level at Manaus (yellow circle; 3.14°S, 60.03°W), along with the elevation (shaded; m) of the region. As demarcated in the figure: the Negro River basin is the northernmost basin; the Solimões River basin is the central basin; and the Madeira River basin is the southernmost basin. (b) Hydrograph climatology (over 1903–2019 period) with daily mean (black solid line) and standard deviation (hatched area), and maximum and minimum (dashed lines) water levels (m) for the Negro River at Manaus over the annual cycle along with the 29 m flood emergency threshold (red dotted line)

(Emerton et al., 2018), provide daily river streamflow forecasts, but they are usually skilful at the subseasonal and regional scales (Towner et al., 2019). Annual maximum flood level is a more flood impact related variable than streamflow, as flood events are extreme cases of maximum water level; this is forecasted operationally by CPRM for Manaus. Linear regression approaches have been used to develop reliable operational forecasts of the maximum water level for Manaus issued in March (Maciel et al., 2020; Schöngart and Junk, 2007, 2020), further improved by
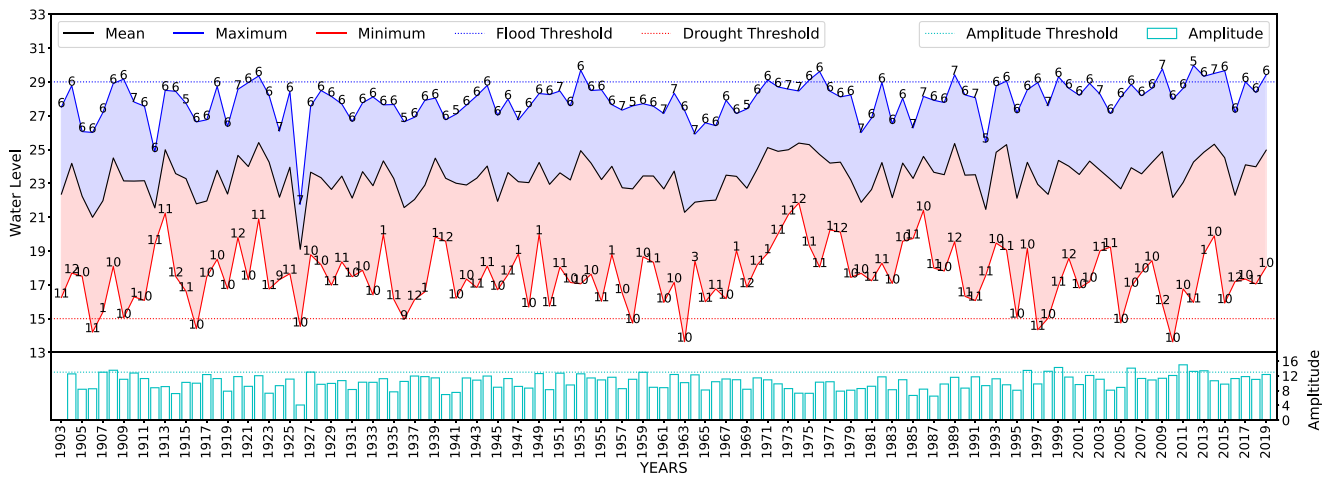
Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change — Open Access

RMetS
Royal Meteorological Society

CHEVUTURI ET AL.

FIGURE 2 Observed annual mean (black solid line), maxima (blue solid line), minima (red solid line) and amplitude (cyan bars) of Negro River water level (m) for 1903–2019. Annual maximum, mean and minimum water levels use the left-hand axis; amplitude uses the right-hand axis. Months (numbers) of occurrence for each year's maximum and minimum are shown next to their respective peaks and troughs. Amplitude does not have a value for 1903, as it is calculated as current year's maximum minus the previous year's minimum. The thresholds for extreme flood (≥29 m; blue dotted line), drought (≤15 m; red dotted line) and amplitude (≥13 m; cyan dotted line) are shown as horizontal lines

integrating teleconnections from large-scale climate parameters (Schöngart and Junk, 2020). However, existing statistical models do not use antecedent catchment rainfall as a predictor.

The main objective of this study is to develop and evaluate statistical seasonal forecast models for annual maximum water level for Negro River at Manaus, using a multiple linear regression approach. We investigate the possibility of improving existing statistical forecasts, using observed antecedent rainfall as a predictor and leveraging the teleconnections from large-scale modes of variability represented by climate indices, to develop a model that can be implemented operationally at the current earliest forecast date (March; which is three months prior to the average timing of the maximum water level around June; Figure 1b). If this method is successful for Manaus, it should be further tested to develop statistical forecast models for other location in the Amazon basin. We also examine the potential to extend the lead time by constructing models for forecasts issued in February or January. Longer lead times (i.e. forecasts issued before March) would allow additional preparation time for high-impact floods that are likely to affect large regions (Lima et al., 2015). Using the forecast models developed in this study, an operational prediction service can be implemented for Manaus by the local hydrological services.

The article is organized as follows: The data used are introduced in Section 2; the relationships between annual maximum water levels and the potential set of predictors are in Section 3; the approach for the statistical

model development is in Section 4; the model validation is described in Section 5; and Section 6 concludes this study.

2 | DATA

Capitania dos Portos (Port Authority) has measured the daily Negro River water level (stage) at the Manaus harbour station (ID: 14990000) since September 1902 (Maciel et al., 2020; Schöngart and Junk, 2007). The location of Manaus (3.14°S, 60.03°W) and the associated river network and catchments, along with the climatology of the annual hydrological cycle of Negro River at Manaus is shown in Figure 1. Figure 2 shows the interannual variability of maximum, mean, minimum water levels, and the yearly amplitude (current year maximum minus previous year minimum); along with the years exceeding the thresholds for severe flood (≥29 m), extreme hydrological drought (≤15 m) and high annual amplitudes (≥13 m). An emergency situation is declared by the government for water levels at Manaus ≥29 m. The annual maximum (peak) water level is during May–July (MJJ; Figure 1b), with most occurrences in June (Figure 2). In this study, we calculate the historical (1903–2020) annual maximum water level, from the daily data, which is our predictand.

The observed antecedent rainfall in the catchment upstream of Manaus is an important potential source of predictability for water levels. We use two rainfall datasets to determine the observed lagged relationships between catchment rainfall and the maximum water levels in

CHEVUTURI ET AL.

Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change   Open Access   RMetS   5 of 17

Manaus: CHIRPS (Climate Hazards Group InfraRed Precipitation with Station) version 2.0 (Funk et al., 2015) at 0.05° resolution, available from 1981 to present; and GPCC (Global Precipitation Climatology Centre) version 2018, available from 1891 to 2016, and version 6 available from 2017 to 2019 (Schneider et al., 2017) at 1.0° resolution. Both rainfall datasets have similar rainfall climatology and inter-annual variability for South America (not shown).

We consider the three basins of the Negro, Solimões and Madeira Rivers (Figure 1a) as the upstream catchment for the Negro River. The Negro and Solimões Rivers show similar annual cycles with average timing of annual maximum water level in June (Schöngart and Junk, 2020) because the water level in the Negro River at Manaus is strongly influenced by the backwater effect from the Solimões River (Meade et al., 1991), as the confluence of two rivers is very close to Manaus. Although the hydrograph of the Madeira River shows a maximum water level earlier than that for the Negro River (Junk et al., 2011), it still influences the water levels measured in Manaus; strong floods of the Madeira River in 2014 caused severe floods in Manaus due to the backwater effect (Schöngart and Junk, 2020). The 2014 flood at Manaus was not caused by the backwater effect from the Solimões River, as the cities upstream of Manaus on the banks of the Solimões River were not affected by flood in that year.

Rainfall patterns in the Amazon region are substantially influenced by SST anomalies in the Equatorial Pacific and the tropical Atlantic (see Section 1). These teleconnections offer additional sources of predictability for Amazon Basin water levels. As the water-level data are available from 1903 onward, we use three indices of large-scale climate variability, available over the same period. The Southern Oscillation Index (SOI) is a meteorological index based on the difference in atmospheric pressure between Darwin (Australia) and Tahiti in the Pacific Ocean, available from 1876 (Ropelewski and Jones, 1987). The SOI is one of the indices which measures the large-scale fluctuations over the tropical Pacific, and thus the conditions associated with the El Niño–Southern Oscillation (ENSO). The Interdecadal Pacific Oscillation (IPO) index can be identified using the tripole index (Henley et al., 2015), which is based on the difference between the SST anomalies over the central equatorial, the Northwest and the Southwest Pacific. As the name suggests, the IPO describes the multi-decadal climate variability over the Pacific Basin; its phases may persist for 20–30 years. The IPO index is calculated with SSTs from ERSST version 5 and is available from 1854. The Atlantic Multidecadal Oscillation (AMO) index is calculated from the detrended weighted average of the North Atlantic SSTs (Enfield et al., 2001). The AMO index describes the natural variability of the North Atlantic Ocean and has a period of 60–80 years. The AMO dataset is calculated from Kaplan SST dataset and is available from 1856 onwards.

CPRM has issued forecasts for Manaus in March (CPRM-March), April (CPRM-April) and May (CPRM-May) since 1989 (Maciel et al., 2020). The CPRM models are based on the simple linear regression of historical annual maximum water level against the water levels for the last day of the month in which the forecast is issued (Level_31Mar, Level_30Apr, Level_31May, respectively). Another seasonal forecasting model for Manaus was developed collaboratively by Instituto Nacional de Pesquisas da Amazonia (INPA) and Max-Planck Society (Schöngart and Junk, 2007). This model uses multiple linear regressions of maximum water level against historic mean water level in February and the SOI of February to forecast annual maximum water level issued in March. This model was updated (INPA-March) to forecast annual maximum water levels using Niño3.4 SST anomalies for December–February (Niño3.4_DJF), the SOI for November–January (SOI_NDJ), the Pacific decadal oscillation for February (PDO_Feb), the previous year's minimum water level (Pmin) and the 7 March water level (Level_7Mar) at Manaus (Schöngart and Junk, 2020). It is important to note that the CPRM forecasts are issued at the end of the month of the forecast (March, April and May), whereas the INPA forecasts are issued by the first week of March. We evaluate the forecasts from the statistical models developed in this study against the existing CPRM (March, April and May) and INPA (March) forecasts. This allows us to not only compare the performance of the models, we develop for Manaus, against the operational forecasts, but it also allows us to evaluate our method. We also evaluated our models against benchmark persistence and climatological forecasts, calculated as the previous years' observed maximum water level and climatological mean of maximum water level from 1903 up to the previous year, respectively.

3 | POTENTIAL PREDICTORS

In this section, we discuss the lagged relationship between annual maximum water level at Manaus and a set of potential predictors, including antecedent catchment rainfall, antecedent water level at Manaus and large-scale modes of climate variability. This discussion will identify the predictors that can be used to forecast maximum water level at Manaus using the multiple linear regression method.

First, we determine the observed lagged relationships between catchment rainfall and the maximum water level in Manaus, using gridded rainfall datasets, which will establish the typical delay between rainfall and peak water-level season. To identify the most useful areas of the
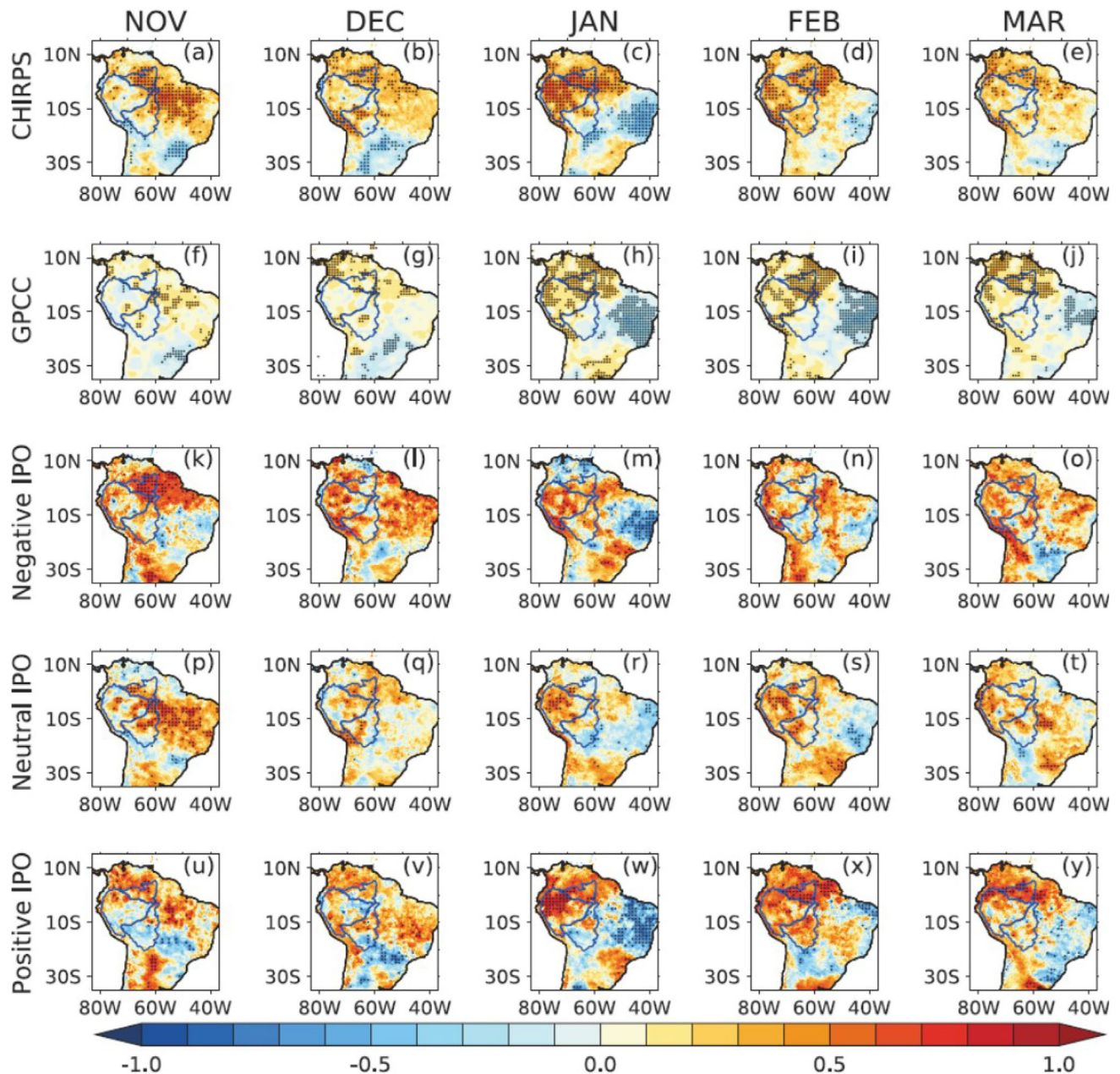
FIGURE 3 Correlation coefficient (shaded) between annual maximum water level and antecedent CHIRPS rainfall at each grid point from the previous year's (a) November and (b) December and the current year's (c) January, (d) February and (e) March over 1982–2019. Stippling shows regions with correlations significant at the 5% level. Similar to (a–e), (f–j) is with antecedent GPCC rainfall over 1904–2019, (k–o) is with CHIRPS rainfall for negative IPO phases only, (p–t) is with CHIRPS rainfall for neutral IPO phases only and (u–y) is with CHIRPS rainfall for positive IPO phases only. The regions bounded by blue lines shows the basins for the Negro, Solimões and Madeira Rivers

catchment for water-level prediction, we correlate monthly rainfall (from CHIRPS and GPCC) at each grid point for the months preceding the peak water-level season with the maximum water level (Figure 3a–j). The regions with positive and statistically significant correlations (at the 5% level) include the upstream basin areas of the three river basins (Negro, Solimões and Madeira). Although the Manaus station measures the water level for Negro River, the

water level itself is influenced by the water level from nearby rivers (pre-dominantly Solimões, and to a lesser extent Madeira) in the network (see Section 2).

We define rainfall masks to create area-mean rainfall based on the correlation maps (Figure 3). The rainfall masks are grid points within the three basins at which the correlation between rainfall and maximum water level is positive and statistically significant at the 5% level
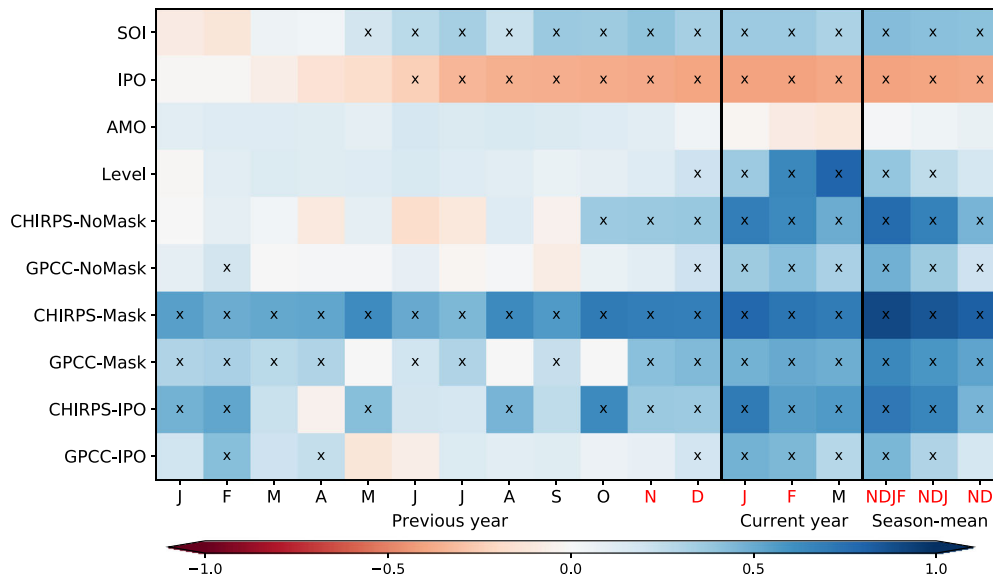
CHEVUTURI ET AL.

Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change · Open Access · RMetS

7 of 17

FIGURE 4 Correlation coefficient (shaded) between the annual maximum water level at Manaus and indices from large-scale models of climate variability (SOI, IPO, AMO); monthly mean water level at Manaus (Level); monthly mean rainfall (CHIRPS-NoMask, GPCC-NoMask, CHIRPS-Mask, CHIRPS-IPO, GPCC-IPO) for all the months in the previous year and January to March in the current year, and for the NDJF, NDJ and ND season-mean (ND from the previous year and JF from the current year; labels are shown in red). CHIRPS-NoMask and GPCC-NoMask are rainfall averaged over all three catchment regions from Figure 1a; CHIRPS-Mask and GPCC-Mask are rainfall averaged over the 'unconditional' masks from Figure 3a–j; CHIRPS-IPO and GPCC-IPO are rainfall averaged over the IPO-conditioned masks from Figure 3k–y. The correlations are calculated over the period of 1904–2019 for all indices, except for the CHIRPS indices, which are calculated over 1982–2019. Correlations statistically significant at the 5% level are designated with a 'x' sign

(stippled regions within the basin boundaries in Figure 3). The rainfall masks are computed for each month of the preceding year and from January to March for the current year (15 months; only five months shown in Figure 3). Correlations between regional-mean CHIRPS or GPCC rainfall over the masked regions (CHIRPS-Mask and GPCC-Mask) against the maximum water level, for each of the 15 months are higher than correlations for the 'non-masked' area-mean rainfall of the three catchment regions (CHIRPS-NoMask and GPCC-NoMask), as expected (Figure 4). The rainfall, from both CHIRPS and GPCC, in November–February (NDJF), preceding the flood season (MJJ), is strongly related to the maximum water level. Thus we use the masked rainfall over NDJF (Rain_NDJF) as one of the input predictors. We obtained similar masks and correlations when we used a 10% significance level.

Historical water levels at Manaus are input variables for the existing INPA and CPRM forecast models (Schöngart and Junk, 2007), with CPRM models only using water levels in their simple regression models. Thus, antecedent water level can be an important predictor for any statistical model forecasting maximum water level. As with antecedent rainfall, we compare the monthly-mean water level (Level) at Manaus against maximum water level for the same 15 months (Figure 4). The monthly (December–March) and seasonal-mean water levels for

NDJF (Level_NDJF) have statistically significant correlations with the maximum water level, thus we retain Level_NDJF as a potential predictor for the forecast model. Although the previous year's minimum water level (Pmin) at Manaus is not significantly correlated with the maximum water level ($r = -0.05$ over 1904–2019), Schöngart and Junk (2020) found that this variable is an important predictor for their model, and thus we include Pmin as a potential predictor.

The annual average and maximum water level at Manaus show increasing linear trends (Figure 2). Marengo and Espinoza (2016) discuss in detail the seasonal floods and droughts in the Amazon region. The significant increase in very severe floods over the Amazon region is characterized by the significant increase of maximum water levels over recent decades (e.g. Barichivich et al., 2018; Schöngart and Junk, 2020), which has been linked to an intensified hydrological cycle (e.g. Gloor et al., 2013) due to a stronger Walker circulation associated with warming of the tropical Atlantic and cooling over the east tropical Pacific (e.g. Barichivich et al., 2018; Wang et al., 2018). The increasing trend in the extreme flood events suggests that the linear trend of maximum water levels may also play an important role in forecasting maximum water level at Manaus. Thus, we also consider the year of the forecast (Year) as a potential predictor for our forecast models.

8 of 17

Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change Open Access RMetS

CHEVUTURI ET AL.

Pacific Ocean indices have successfully been used for forecasting maximum water levels at Manaus by INPA (Schöngart and Junk, 2007, 2020). The two Pacific Ocean indices (SOI and IPO) show strong correlations against the annual maximum water level from July of the previous year to April of the current year, but with opposite signs (Figure 4). The SOI index correlation with maximum water levels indicates the well-known relationship of severe floods during La Niña years and lower maximum water level in El Niño years (Marengo and Espinoza, 2016). We use the seasonal-mean (NDJF) indices for SOI and IPO as input predictors as they show strong correlations with maximum water level. The Atlantic Ocean index (AMO) does not show a strong relationship with the maximum water level at seasonal timescale (NDJF). A case study of 2014 flood event at Manaus, influenced by the backwater effect from the Madeira River, has been attributed to the anomalous moisture transport over the Amazon basin from the Atlantic, with variability in Atlantic SST gradients as the main driver (Espinoza et al., 2014). Further, Atlantic SSTs have strong impacts on the annual minimum water level in Manaus (Schöngart and Junk, 2020). Thus, we use the AMO index as a potential predictor.

We use season-mean (NDJF) predictors of antecedent rainfall, water level and large-scale indices, for developing a model that can issue forecasts in March. To extend the forecast lead time, by constructing models that could operationally issue forecasts in February or January, we can only use seasonal-mean predictors for November–January (NDJ) or November–December (ND), respectively. The sign of the correlations between maximum water level and different seasonal-mean predictors (NDJF, NDJ and ND) remain the same, but the correlations weaken from NDJF to NDJ to ND in all predictors (Figure 4).

4 | MODEL DEVELOPMENT

To forecast maximum water level at Manaus, we use a multiple linear regression approach, through testing combination of potential predictors, as in Schöngart and Junk (2007, 2020). Figure 5 shows a schematic for the model development process. We first identify a subset of the predictors described in Section 3 by different screening regression approaches, to find the best fit for the training period, using the ordinary least-squares method. Using this method, we derive the multiple linear regression equation (or the forecast model) for a specific subset of predictors. We then validate our forecast models against the existing operational forecasts.

We use the multiple linear regression method to develop the equation for the predictand (or dependent variable, i.e. annual maximum water level) using the predictors
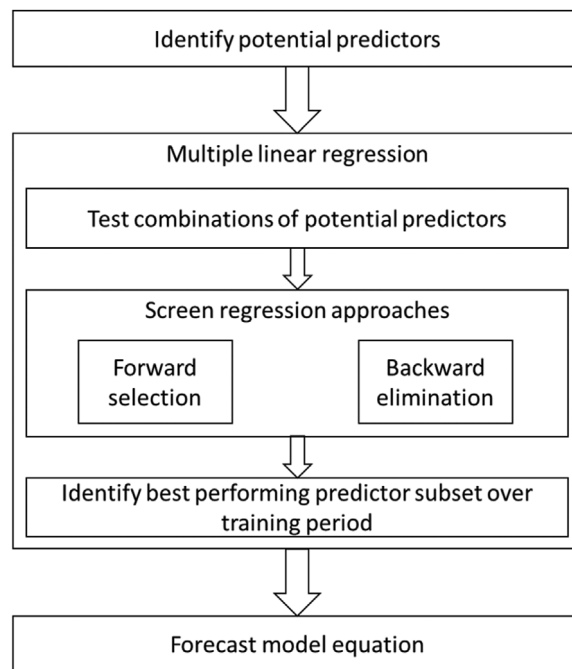

FIGURE 5 Schematic showing the model development process

(or independent variables). From the initial seven potential predictors, we filter a subset of predictors which provide the best fit (through ordinary least squares method) with the observed data during the training period, using screening regression approaches (forward selection and backward elimination). The forward selection chooses the potential predictors based on the strength of their linear relationships to the predictand. For each iteration, we test adding new potential predictors, and select the one that provides the best possible fit over the training period. Backward elimination starts with the best possible fit using all potential predictors, then removes one potential predictor each iteration to find the best combination of the fewest possible potential predictors. The best fit in both approaches is chosen on the basis of the highest adjusted coefficient of determination (Adj-R^2) and the lowest mean squared error (MSE) over the training period. We consider Adj-R^2 rather than coefficient of determination (R^2), as R^2 always increases as more predictors are added, but Adj-R^2 only increases if a new predictor improves the model by more than that would be expected by random chance. The MSE is the sum of the squared errors divided by the residual degrees of freedom (sample size minus number of parameters minus 1). For assessing the models' performance over the validation period (2005–2019), we use the Pearson correlation coefficient (CC) and root mean square error (RMSE) metrics.

Antecedent catchment rainfall is one of the most important predictors for maximum water level at Manaus, which

has not been used directly in existing operational forecasts. CHIRPS and GPCC masked rainfall for NDJF account for approximately 81% and 76% variability in the annual maximum water levels, respectively, for the common period of 1981–2019. The GPCC dataset is available from 1891 onwards, which provides a long training period for our model, but as the dataset is often updated with a lag of a month or more, it cannot be used for real-time forecasting. However, CHIRPS is available in near-real-time (with approximately two-week lag), but is available only from 1981 onwards, which provides a much shorter training period. As CHIRPS and GPCC rainfall are highly correlated, we train our models with GPCC masked rainfall for the training period of 1904–2004, then validated our models using CHIRPS masked rainfall for 2005–2019. We use standardized seasonal-mean masked rainfall for both datasets, to minimize the differences between datasets. Our tests show that the standardization of input rainfall does not change the CC between forecast and observed, but reduces the RMSE of our forecast by ≈0.20 m.

While developing the rainfall based statistical models, we performed sensitivity tests on the potential set of predictors and on our methods, to find the optimal approach. For brevity, we briefly discuss, but do not show, the different approaches we tested before finding the method discussed in this section. The performance of our forecast models was slightly poorer (but not significantly different) when considering rainfall masks calculated at 10% significance level, or when using masks computed for seasonal rainfall instead of for monthly rainfall. We also tested standardizing rainfall at grid point level, instead of standardizing the area-mean rainfall. We found that although the CC does not change, over the validation period, the grid point standardizing approach had substantially higher errors.

4.1 | Forecasts issued in March

Operationally, the earliest current forecast for annual maximum flood levels at Manaus is issued in March. Thus, we first focus on developing a multiple regression model that can also be issued in March. Figure 4 shows that most of the NDJF seasonal-mean indices show slightly stronger correlations against maximum water level than individual monthly-mean indices (see Section 3). Thus, we use the seasonal-mean (NDJF) indices as input predictors for our March forecast models, which thus, can provide real-time forecasts by mid-March. For our initial pool of potential predictors we consider the standardized NDJF-mean antecedent masked rainfall (Rain_NDJF, GPCC in the training period), water level (Level_NDJF) and large-scale teleconnection indices (SOI_NDJF, IPO_NDJF and AMO_NDJF), along with previous year's minimum water

level (Pmin) and the linear trend (Year). Please see Section 3 for the details about the potential predictor indices.

Rain_NDJF is the most important predictor for the maximum water level at Manaus. This is because in free-flowing river systems, such as the Negro River, the water levels are a consequence of the integrated precipitation anomalies over the large-scale catchment area (Junk et al., 2011). The best fit model for the training period (1904–2004) uses four predictors: GPCC_NDJF, Pmin, Year and AMO_NDJF and has Adj-R^2 of 0.53 (Figure 6a). The equation generated using these predictors is our first March forecast model, named Rain-March. During the early years of training period (up to 1930) there is underfitting of the model, but after 1930, the calibration fit of the model improves, which may be related to the sparse observations. We test the same model for forecasting maximum water level for the validation period (2005–2019), but now using standardized NDJF-mean CHIRPS masked rainfall instead of GPCC (Figure 7). Forecasting using CHIRPS masked rainfall shows improved forecasts (CC = 0.93, RMSE = 0.36) against those using GPCC masked rainfall (CC = 0.87, RMSE = 0.56) over the validation period.

Including the linear trend (Year) does not influence the CC of the forecast during the validation period, but significantly reduces RMSE, which is associated with increasing maximum water levels in the Negro River in recent years (Barichivich et al., 2018). The AMO is included as a predictor, despite having a weak correlation against flood levels, but not SOI or IPO, which have stronger correlations (Figure 4). SOI and IPO also show stronger correlations than AMO with rainfall over the Amazon region (Flantua et al., 2016). The updated INPA model, which does not use rainfall as input, includes Pacific SST indices as input predictors (Schöngart and Junk, 2020). This may suggest that the influence of SOI and IPO in the current model is included through the antecedent rainfall over the masked regions itself; further including SOI and IPO as a further predictor does not add additional value to the forecast model.

We examine the spread of the residuals from the training period against the standardized predictors (GPCC_NDJF, AMO and Pmin). We find that the forecast errors, during the training period, are less sensitive (smaller standard deviation) to predictors when the predictors' standardized values are close to 0 (not shown). This suggests that positive or negative extremes of the predictors have large influences on the predictand (maximum water level). Forecast uncertainty of the model is calculated based on the empirical distribution of residuals in the training period (as in Lee et al., 2015); the error bars are calculated as the mean of the residuals over the validation period (as in Schöngart and Junk, 2020). Residuals for the Rain-March model are normally distributed, except for in 1912 and 1926 (Figure 6b), which were years of severe drought and low
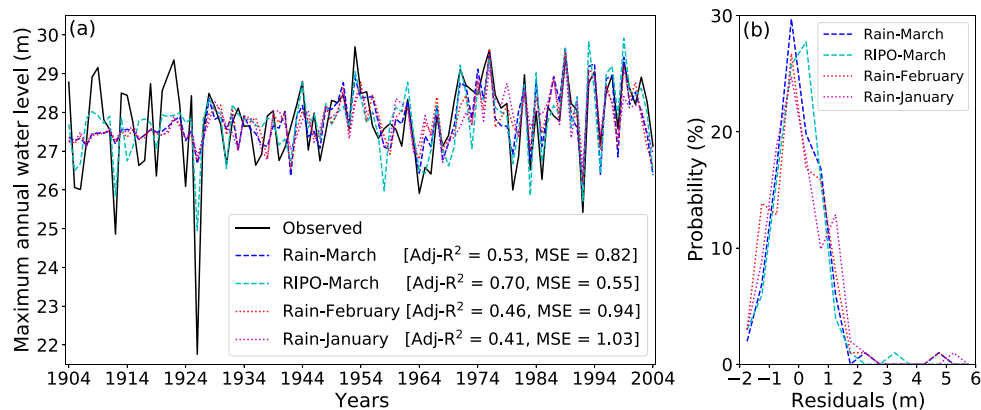
10 of 17 | **Climate Resilience and Sustainability**
Interdisciplinary Approaches Towards Solutions for Climate Change Open Access RMetS Royal Meteorological Society

CHEVUTURI ET AL.

FIGURE 6 (a) Observed (black solid line) and the best fit or calibration (coloured lines) of annual maximum water level (m) for Manaus over the training period (1904–2004) for our four models: Rain-March (blue dashed line), RIPO-March (cyan dashed line), Rain-February (red dotted line) and Rain-January (magenta dotted line). The adjusted-R^2 (Adj-R^2) and mean square error (MSE) over the training period for each model is provided in the legend. (b) Probability (%) distribution of the residuals (m; model minus observed) for the same best fit models (coloured lines) as in (a) over the training period (1904–2004)
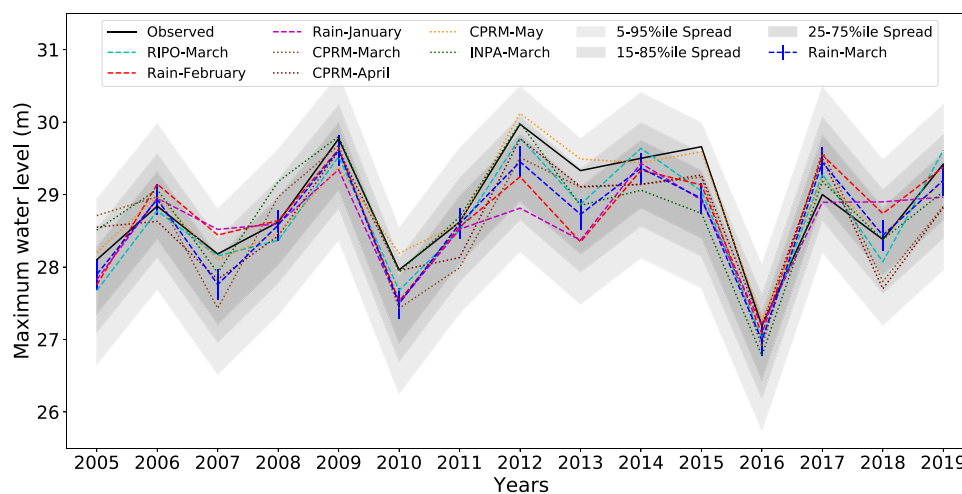
FIGURE 7 Observed (black solid line) and statistical forecasts (coloured lines) of annual maximum water level (m) for Manaus over the validation period of 2005–2019. Statistical forecasts are the CHIRPS rainfall-based March forecast (Rain-March; blue dashed line), IPO-conditioned CHIRPS rainfall-based March forecast (RIPO-March; cyan dashed line), CHIRPS rainfall-based February forecast (Rain-February; red dashed line), CHIRPS rainfall-based January forecast (Rain-January; magenta dashed line), the operational CPRM March forecast (CPRM-March; brown dotted line), CPRM April forecast (CPRM-April; maroon dotted line), CPRM May forecast (CPRM-May; orange dotted line) and the INPA March forecast (INPA-March; green dotted line). Rain-March forecast mean error over the validation period are shown as error bars (blue). The grey shading shows the forecast uncertainty of the Rain-March model; 5th–95th percentile spread (light grey shading), 15th–85th percentile spread (grey shading) and 25th–75th percentile spread (dark grey shading)

flood levels (Figure 6a) associated with strong El Niño events (Marengo and Espinoza, 2016) and to some extent by Atlantic forcing (Williams et al., 2005). The original INPA model also failed to predict the maximum water level of Negro River for 1926 (Schöngart and Junk, 2007). Using the residuals from all years including 1912 and 1926 generates an error spread (5th–95th percentile) of almost ±2 m above and below the forecast for the validation period (not shown), which is not representative of forecast performance in most years. Thus, we remove the years 1912 and 1926 and then calculate the empirical distribution of resid-

uals to show the forecast uncertainty of the Rain-March model for the validation period (Figure 7).

4.2 | Using conditional rainfall masks

As discussed in Section 1, the rainfall over the Amazon basin and thus Amazon river water levels are influenced by large-scale coupled ocean–atmosphere modes of climate variability. Seasonal forecasts show improved skill during certain phases of large-scale modes of climate variability

(ENSO) over tropical regions of South America (Coelho et al., 2006). Based on this, we evaluate if rainfall–water level correlations for different phases of the large-scale climate indices (AMO, SOI and IPO), can improve forecasts of maximum water level at the same lead-time (March). In Figure 3k–y, we show lagged correlations between CHIRPS rainfall and maximum water level at Manaus for different IPO phases. The IPO phases are identified by quartiles: the upper quartile as positive IPO, the lower quartile as negative IPO and the middle two quartiles as neutral IPO. Our results are not sensitive to defining the phases of IPO by terciles instead of quartiles (not shown). Using these lagged correlations, we create rainfall masks for each phase of IPO (Figure 3k–y) and use those to analyse the relationship between 'IPO-conditioned' masked regional-mean rainfall (CHIRPS-IPO and GPCC-IPO) and annual maximum water levels (Figure 4). Although the IPO-conditional masked seasonal-mean rainfall has high correlation with water levels, some months have few or no grid points with statistically significant correlations at the 5% level, which leads to a very small masked area (Figure 3k–y). Due to the small extent of the IPO-conditioned mask, there may be concerns about the stability of our forecast model. Despite this, we present our results of using the IPO-conditioned regional-mean rainfall as an alternate method to forecast water level at Manaus, as it seems to perform better than the model developed using unconditional rainfall masks. We discuss the implication of the IPO-conditioned forecasts in Section 5.

To develop the model with IPO-conditioned regional-mean rainfall as input, we use the same method as for the Rain-March model, but instead of using 'unconditional' rainfall masks (Figure 3a–j), we use 'IPO phase-conditioned' rainfall masks (Figure 3k–y) to calculate the seasonal-mean rainfall (RIPO_NDJF) as one of the potential predictors. The best fit model (named RIPO-March) over the training period uses RIPO_NDJF, Year and IPO_NDJF as predictors. The largest contribution to prediction of variability comes from the IPO phase-conditioned standardized rainfall (RIPO_NDJF). Compared to Rain-March, RIPO-March model has improved fit for the training period, especially for the extreme drought years of 1912 and 1926 (Figure 6a). We also developed statistical models based on SOI- and AMO-conditioned rainfall masks. The best of these 'conditional' (SOI and AMO) models are not significantly better than our original 'unconditional' (Rain-March) model.

4.3 | Extending forecast lead times

To increase the lead time of the forecasts, we develop two statistical models for issuing forecasts of annual max-

imum water level in February (Rain-February) and January (Rain-January). We use the same method as for the Rain-March model, but we use November–January mean potential predictors (Rain_NDJ, Level_NDJ, SOI_NDJ, IPO_NDJ and AMO_NDJ) for the Rain-February model, and November–December mean (Rain_ND, Level_ND, SOI_ND, IPO_ND and AMO_ND) for the Rain-January model. We also include Pmin and Year as other potential predictors for both models. Rainfall, mean water level and large-scale indices have statistically significant correlations with maximum water level for NDJ-mean and ND-mean, but with lower magnitude than for NDJF-mean (Figure 4, Section 3). We do not use months before November for the potential predictors for forecasts issued in February and January, as the forecasts do not gain any additional performance. We find that both models have a good fit with the same subset of predictors as Rain-March model (February: Adj-R^2 = 0.44, MSE = 0.97; January: Adj-R^2 = 0.39, MSE = 1.05). But Rain-February has the best fit with the combination of Rain_NDJ, AMO_NDJ and Year; Rain-January has the best fit with Rain_ND, AMO_ND and Pmin (Figure 6a). As expected, the statistical models with longer lead times have a poorer fit than models at shorter lead times (Figure 6a). We discuss forecast model performance further in Section 5.

5 | MODEL VALIDATION

In this section we validate the forecasts from the four statistical forecast models (Rain-March, RIPO-March, Rain-February and Rain-January) developed in this study (see Section 4) and compare them against existing statistical forecasts (CPRM-March, CPRM-April, CPRM-May and INPA-March). As discussed before, the CPRM forecasts are issued at the end of the month, the INPA forecasts by the first week of March; our rainfall-based forecasts can be issued by the middle of the month (see Sections 2 and 3).

For the validation period, all of the observed values fall within the 5th–95th percentile uncertainty bounds of Rain-March forecasts (Figure 7). The forecast uncertainty of the Rain-March model mostly covers the observed flood levels within the 25th–75th percentile uncertainty bounds, except for 2012 and 2015. For 2015, CPRM-March model has lower error. The INPA-March model exactly forecasts the 2012 observed value. Error bars for Rain-March forecasts are narrower than the uncertainty bounds, as they are calculated over the validation period, but still cover the observed values for most years. The IPO-conditioned forecast model (RIPO-March) outperforms all the other models for most years (Figure 7). For March forecasts, both the statistical models using rainfall (Rain-March and RIPO-March) have higher CC and lower RMSE than the existing

TABLE 1 Comparison of statistical forecasts for annual maximum water level at Manaus over the validation period of 2005–2019

Forecast model	Model predictors	CC	CC uncertainty	RMSE	RMSE uncertainty	Forecast 2020 (Bias)	Forecast 2021 (Bias)
CPRM-May	Level_31May	0.97	0.95–0.99	0.17	0.12–0.22	28.60 (+0.08)	30.00 (−0.02)
CPRM-April	Level_30Apr	0.91	0.80–0.97	0.38	0.30–0.45	28.25 (−0.27)	30.00 (−0.02)
CPRM-March	Level_31Mar	0.87	0.71–0.95	0.48	0.41–0.55	28.30 (−0.22)	29.45 (−0.57)
INPA-March	Niño3.4_DJF, SOI_NDJ, PDO_Feb, Pmin, Level_7Mar	0.88	0.69–0.96	0.39	0.27–0.50	28.48 (−0.04)	29.51 (−0.51)
Rain-March	Rain_NDJF, Year, Pmin, AMO_NDJF	0.93	0.83–0.97	0.36	0.26–0.44	28.45 (−0.07)	29.38 (−0.64)
RIPO-March	Rain_NDJF, Year, IPO_NDJF	0.96	0.90–0.98	0.29	0.22–0.36	28.52 (0.00)	29.11 (−0.91)
Rain-February	Rain_NDJ, Year, AMO_NDJ	0.86	0.66–0.94	0.43	0.29–0.55	28.84 (+0.32)	29.45 (−0.57)
Rain-January	Rain_ND, Year, Pmin	0.81	0.56–0.93	0.51	0.33–0.66	29.37 (+0.85)	29.06 (−0.96)

Note: The forecasts compared are CPRM forecasts issued in May (CPRM-May), April (CPRM-April) and March (CPRM-March), INPA forecast issued in March (INPA-March) and rainfall-based forecasts issued in March (Rain-March), February (Rain-February) and January (Rain-January) and IPO-conditioned rainfall-based forecast issued in March (RIPO-March). The forecasts are compared using correlation coefficient (CC) and its 5th–95th percentile (CC uncertainty range) and root mean square error (RMSE) and its 5th–95th percentile (RMSE uncertainty range). The models' operational forecasts and biases for the years 2020 and 2021 are also shown. Observed maximum water level at Manaus for 2020 was 28.52 m and for 2021 was 30.02 m.

CPRM-March and INPA-March forecasts (Table 1). Forecasts and bias for the year 2020 also paint a similar picture (Table 1), where RIPO-March has the exactly the correct forecast, whereas INPA-March and Rain-March also show lower biases than CPRM-March. Although the RIPO-March model offers the highest performance, the small number of grid points included in the rainfall mask (Figure 3k–y) raises concerns about the stability of the model. This may be seen in the year 2021, which was an extreme flood year, for which the RIPO-March model has a relatively high bias. The small number of grid points in the IPO-conditioned mask may cause the rainfall input for RIPO-March model to poorly represent catchment rainfall or have artificially high variability, which may lead to very poor forecasts in some years. Due to this uncertainty, the model based on IPO-conditioned rainfall cannot be recommended for operational use without further scientific analysis.

The statistical models with longer lead times (Rain-February and Rain-January) have lower performance and higher errors (Table 1). We also see a similar lead-time dependent loss in performance in CPRM forecasts (CPRM-May to CPRM-March). The Rain-March and RIPO-March models perform better than the CPRM forecast in April, but the difference is small. We also compared our models against benchmark persistence and climatological forecasts (see Section 2) by calculating their performance metrics (CC and RMSE) over the validation period (2005–2019). Our models perform significantly better than the persistence (CC = −0.2, RMSE = 1.16) and climatological (CC = 0.09, RMSE = 1.26) forecasts.

To compare the models' performance fairly, we calculate distributions of model performance metrics (CC and RMSE), using a bootstrapping approach, by randomly generating 10,000 samples, from the validation years, with replacement. We show the 5th–95th percentiles of the distributions generated from bootstrapping (Table 1). The distributions of the CC are positively skewed, with overlaps in the confidence intervals from different forecasts. The performance of the rainfall-based models in March (Rain-March and RIPO-March) is similar to that of the CPRM-April model; the performance of Rain-February model is similar to that of the INPA-March and CPRM-March models. We say similar performance, with high confidence, only when at least 95% of the CC and RMSE distributions of one model falls within the 5th–95th percentiles ranges of another model's CC and RMSE. Thus, we can say with high confidence that our rainfall-based models gain a month of lead-time against the existing models, for similar forecast performance. There is also an overlap in the confidence intervals for CC for Rain-March against existing March models (CPRM-March and INPA-March) at the same lead-time, but only part (less than 75%) of the distributions overlap. This provides us with only moderate confidence that the rainfall-based models improve on the existing models at the same lead time. It should be noted that even at same lead-time there is a couple of weeks of difference in the actual forecasting dates between the CPRM and

CHEVUTURI ET AL.

Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change | Open Access | RMetS | 13 of 17

rainfall-based models. The overlap between the distributions for the uncertainty range in RMSE is relatively lower than that in CC, which suggests that there is higher confidence for the result that the rainfall-based models have lower errors than the existing models at similar lead times.

The CPRM-May model outperforms all other models, with only less than 40% of the distribution overlapping between the CC uncertainty range of Rain-March. RIPO-March has larger overlap (greater than 75%) with CC uncertainty range of CPRM-May. It should be noted that CPRM-May forecasts are issued on 31 May, as they use water level for 31 May (Level_31May) as the predictor. As the CPRM-May forecasts are issued very close to the time of the actual flood, they are thus of less value for large-scale anticipatory action and disaster preparation.

The existing operational models and the models developed in this study are relatively simple, but they can suffer from errors related to uncertainty of the linear relationships and predictors they rely on (Coe et al., 2002; Paiva et al., 2012). There is uncertainty associated with the large-scale climate indices, for example the SOI index correlation with ENSO is lower before 1950 than after, which may be due errors in datasets and/or changes in their relationship (Bunge and Clarke, 2009). There is also considerable observational uncertainty in Amazonian rainfall (e.g. Costa and Foley, 1998), but recent satellite-based products have improved the representation of hydrological cycle over the Amazon region (e.g. Chen et al., 2009). Dynamical hydrological forecasting models have shown that meteorological input controls the accuracy of maximum river flow (Towner et al., 2019) and using rainfall as input provides skilful runoff forecasts (e.g. Zubieta et al., 2015). Thus, antecedent rainfall offer an opportunity to develop skilful seasonal statistical forecasts for water levels of free-flowing rivers. Further, the models developed in this study benefit from the long training period and use of a different dataset during the validation period, which avoids overfitting of the model and makes it more robust. However, non-stationarity of relationships between predictors and maximum water level strongly influences the statistical model forecasts. This non-stationarity may be caused by changes in the climate, land cover, land use or even the large-scale ocean–atmosphere teleconnections (e.g. Lima et al., 2015).

6 | SUMMARY AND CONCLUSIONS

Extreme overbank flooding of the Amazon River and its tributaries can cause considerable socio-economic losses in tropical Brazil, which is relatively less developed and highly vulnerable to floods. It is therefore critical to advance the prediction of high water levels in the Ama-

zon, to provide more effective and earlier warnings of impending disasters to safeguard lives and livelihoods. The Amazon basin hydrology and topography, with its extensive river network and wetlands, provides a natural delay between the seasonal rainfall and river water levels. Thus, the Amazonian rivers present a regular hydrological annual cycle. The hydrographs of low-lying large rivers in the Amazon basin present a single annual flood event that lasts for weeks to months. This regular monomodal flood-pulse offers an opportunity to predict the peak water level of the free-flowing river systems, months in advance, using the antecedent catchment rainfall.

This study develops and validates statistical forecast models for the annual maximum water level of the Negro River at Manaus, with improved performance compared to existing models, at the current earliest operational lead time (March) or longer (February and January). A set of potential predictors are identified that strongly influence the water levels of the Negro River, including catchment rainfall, large-scale teleconnection indices, the long-term linear trend of water level and antecedent water levels. We use observed gridded rainfall datasets (GPCC and CHIRPS) to identify catchment rainfall based on 'unconditional' and 'conditional' masks. Using multiple linear regression methods and screening regression approaches, we identify the best fit model over the training period (1904–2004) and validate and compare the models against existing forecasts over 2005–2019.

Using the method summarized above, we develop four statistical forecast models that have antecedent rainfall as the most important predictor. The forecast models provide real-time forecasts, based on masked rainfall calculated from regions with high correlations with the maximum water level at Manaus. We develop three models, increasing in lead time from March to February to January, which use the 'unconditional' rainfall masks, and another model for forecasts issued in March, which uses the IPO-conditioned rainfall masks. As expected, the models lose performance with increasing lead times. Our results show, with high confidence (at the 95% level), that the rainfall-based models gain a month of lead-time against the existing models, for similar forecast performance. Our March models also improve against existing models at the same lead time, but with moderate confidence (at the 75% level). None of the rainfall based statistical models were better than the operational forecasts (CPRM) issued in May. But CPRM-May model has the benefit of two additional months of input data and is of limited use for preparing for floods in May–July. This is especially true for the years exceeding the critical threshold of 29 m (at which state of emergency is declared), as by May this threshold is usually already reached. One of the benefits of our statistical models is that forecasts can be issued by the middle of the

14 of 17

Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change · Open Access · RMetS

CHEVUTURI ET AL.

month, whereas the CPRM operational forecasts are issued at the end of the month.

Large-scale changes to the land cover and land use around the river network (e.g. Spracklen and Garcia-Carreras, 2015) and any river management systems such as dams (e.g. Timpe and Kaplan, 2017) affect the relationship between rainfall and river flow over the region. Our method is not applicable to such rivers. However, the method developed in this study could be applied to develop statistical models for flood forecasts over other free-flowing rivers in the Amazon basin with intact catchments. This method could also be applied to develop forecast models for hydrological droughts over the Amazon basin, which also have severe impacts on the ecosystems and socio-economy of the local regions. The forecast models developed in this study can be used for real-time forecast application, and have been tested retrospectively for the years 2020 and 2021. For 2020, our March forecast shows adequate performance. For 2021, all of the forecasts underestimated the extreme event.

However, we should note that the efficacy of any skilful flood forecast, for the local community, may be limited by quality of the integrated flood management systems implemented, which include flood alert dissemination, flood preparedness and impact mitigation measures. Other flood characteristics (e.g. duration and timing), beyond flood levels, also influence the impacts of the Amazonian floods (Langill and Abizaid, 2020). There is a strong correlation between the maximum water level and flood duration for the severe flood years that exceed the 29 m threshold (17 years on record). Further research may improve our methods, not only for flood levels, but also for its timing and duration, by exploring additional potential indicators (e.g. upstream water levels) and other statistical regression methods (e.g. advanced machine learning techniques).

ORCID

Amulya Chevuturi https://orcid.org/0000-0003-2815-7221
Nicholas P. Klingaman https://orcid.org/0000-0002-2927-9303
Conrado M. Rudorff https://orcid.org/0000-0001-8453-1367
Caio A. S. Coelho https://orcid.org/0000-0002-9695-5113
Jochen Schöngart https://orcid.org/0000-0002-7696-9657

REFERENCES

Barichivich, J., Gloor, E., Peylin, P., Brienen, R.J., Schöngart, J., Espinoza, J.C.& Pattnayak, K.C. (2018) Recent intensification of Amazon flooding extremes driven by strengthened Walker circulation. *Science Advances*, 4(9), eaat8785. https://doi.org/10.1126/sciadv.aat8785

Betts, A., Fisch, G., Von Randow, C., Silva Dias, M., Cohen, J., Da Silva, R. & Fitzjarrald, D. (2009) The Amazonian boundary layer and mesoscale circulations. *Amazonia and global change*. Geophysical Monograph Series. Washington, DC: AGU. https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2008GM000720

Bunge, L.& Clarke, A.J. (2009) A verified estimation of the El Niño index Niño-3.4 since 1877. *Journal of Climate*, 22(14), 3979–3992. https://doi.org/10.1175/2009JCLI2724.1

Callède, J., Cochonneau, G., Vieira Alves, F., Guyot, J.-L., Santos Guimaraes, V.& De Oliveira, E. (2010) The River Amazon water contribution to the Atlantic Ocean. *Revue des sciences de l'eau*, 23(3), 247–273. https://doi.org/10.7202/044688ar

Carvalho, L.M., Silva, A.E., Jones, C., Liebmann, B., Dias, P.L.S.& Rocha, H.R. (2011) Moisture transport and intraseasonal variability in the South America monsoon system. *Climate Dynamics*, 36(9–10), 1865–1880. https://doi.org/10.1007/s00382-010-0806-2

Chen, J., Wilson, C., Tapley, B., Yang, Z. & Niu, G.-Y. (2009) 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models. *Journal of Geophysical Research: Solid Earth*, 114(B5). Available from: https://doi.org/10.1029/2008JB006056

CHEVUTURI ET AL.

Climate Resilience and Sustainability
Interdisciplinary Approaches Towards Solutions for Climate Change   Open Access   RMetS | 15 of 17

Coe, M.T., Costa, M.H., Botta, A. & Birkett, C. (2002) Long-term simulations of discharge and floods in the Amazon basin. *Journal of Geophysical Research: Atmospheres*, 107(D20), LBA–11. https://doi.org/10.1029/2001JD000740

Coelho, C., Stephenson, D., Balmaseda, M., Doblas-Reyes, F.& Van Oldenborgh, G. (2006) Toward an integrated seasonal forecasting system for South America. *Journal of Climate*, 19(15), 3704–3721. https://doi.org/10.1175/JCLI3801.1

Costa, M.H., Coe, M.T. & Guyot, J.L. (2009) Effects of climatic variability and deforestation on surface water regimes. In: Keller, M., Bustamante, M., Gash, J. & Dias, P.S. (Eds.) *Amazonia and global change*, Geophysical Monograph Series. Vol. 186. Washington, DC: AGU, pp. 543–553. https://doi.org/10.1029/2008GM000721

Costa, M.H.& Foley, J.A. (1998) A comparison of precipitation datasets for the Amazon basin. *Geophysical Research Letters*, 25(2), 155–158. https://doi.org/10.1029/97GL03502

de Andrade, M.M.N., Bandeira, I.C.N., Fonseca, D.D.F., Bezerra, P.E.S., de Souza Andrade, Á. & de Oliveira, R.S. (2017) Flood risk mapping in the Amazon. In: Hromadka, T. & Rao, P. (Eds.) *Flood risk management*, Vol. 41. InTech. Available from: https://doi.org/10.5772/intechopen.68912

De Linage, C., Famiglietti, J.& Randerson, J. (2014) Statistical prediction of terrestrial water storage changes in the Amazon basin using tropical Pacific and North Atlantic sea surface temperature anomalies. *Hydrology and Earth System Sciences*, 18(6), 2089–2102. https://doi.org/10.5194/hess-18-2089-2014

de Paiva, R.C.D., Buarque, D.C., Collischonn, W., Bonnet, M.-P., Frappart, F., Calmant, S.& Bulhões Mendes, C.A. (2013) Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research*, 49(3), 1226–1243. https://doi.org/10.1002/wrcr.20067

Dias, L.C.P., Macedo, M.N., Costa, M.H., Coe, M.T.& Neill, C. (2015) Effects of land cover change on evapotranspiration and streamflow of small catchments in the upper Xingu River basin, Central Brazil. *Journal of Hydrology: Regional Studies*, 4, 108–122. https://doi.org/10.1016/j.ejrh.2015.05.010

Dolman, D.I., Brown, I.F., Anderson, L.O., Warner, J.F., Marchezini, V.& Santos, G.L.P. (2018) Re-thinking socio-economic impact assessments of disasters: the 2015 flood in Rio Branco, Brazilian Amazon. *International Journal of Disaster Risk Reduction*, 31, 212–219. https://doi.org/10.1016/j.ijdrr.2018.04.024

Drumond, A., Marengo, J., Ambrizzi, T., Nieto, R., Moreira, L.& Gimeno, L. (2014) The role of the Amazon basin moisture in the atmospheric branch of the hydrological cycle: a Lagrangian analysis. *Hydrology and Earth System Sciences*, 18(7), 2577. https://doi.org/10.5194/hess-18-2577-2014

Emerton, R., Zsoter, E., Arnal, L., Cloke, H.L., Muraro, D., Prud-homme, C., Stephens, E.M., Salamon, P.& Pappenberger, F. (2018) Developing a global operational seasonal hydro-meteorological forecasting system: GloFAS-Seasonal v1.0. *Geoscientific Model Development*, 11(8), 3327–3346. https://doi.org/10.5194/gmd-11-3327-2018

Enfield, D.B., Mestas-Nuñez, A.M.& Trimble, P.J. (2001) The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophysical Research Letters*, 28(10), 2077–2080. https://doi.org/10.1029/2000GL012745

Espinoza, J.C., Marengo, J.A., Ronchail, J., Carpio, J.M., Flores, L.N.& Guyot, J.L. (2014) The extreme 2014 flood in south-western Amazon basin: the role of tropical-subtropical South Atlantic SST gra-dient. *Environmental Research Letters*, 9(12), 124007. https://doi.org/10.1088/1748-9326/9/12/124007

Flantua, S.G., Hooghiemstra, H., Vuille, M., Behling, H., Carson, J.F., Gosling, W., et al. (2016) Climate variability and human impact in South America during the last 2000 years: synthesis and perspectives from pollen records. *Climate of the Past*, 12, 483–523. https://doi.org/10.5194/cp-12-483-2016

Fu, R., Dickinson, R.E., Chen, M. & Wang, H. (2001) How do tropical sea surface temperatures influence the seasonal distribution of precipitation in the equatorial Amazon? *Journal of Climate*, 14(20), 4003–4026. https://doi.org/10.1175/1520-0442(2001)014%3C4003:HDTSST%3E2.0.CO;2

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. & et al. (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2(1), 1–21. https://doi.org/10.1038/sdata.2015.66

Gloor, M., Barichivich, J., Ziv, G., Brienen, R., Schöngart, J., Peylin, P., Cintra, B.B.L., Feldpausch, T., Phillips, O.& Baker, J. (2015) Recent Amazon climate as background for possible ongoing and future changes of Amazon humid forests. *Global Biogeochemical Cycles*, 29(9), 1384–1399. https://doi.org/10.1002/2014GB005080

Gloor, M., Brienen, R.J., Galbraith, D., Feldpausch, T.R., Schöngart, J., Guyot, J.-L., Espinoza, J.C., Lloyd, J.& Phillips, O.L. (2013) Intensification of the Amazon hydrological cycle over the last two decades. *Geophysical Research Letters*, 40(9), 1729–1733. https://doi.org/10.1002/grl.50377

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., et al. (2019) Mapping the world's free-flowing rivers. *Nature*, 569(7755), 215–221. https://doi.org/10.1038/s41586-019-1111-9

Guedes, G., Costa, S. & Brondizio, E. (2009) Revisiting the hierarchy of urban areas in the Brazilian Amazon: a multilevel approach. *Population and environment*, 30(4–5), 159–192. https://doi.org/10.1007/s11111-009-0083-3

Henley, B.J., Gergis, J., Karoly, D.J., Power, S., Kennedy, J. & Folland, C.K. (2015) A tripole index for the interdecadal Pacific oscillation. *Climate Dynamics*, 45(11–12), 3077–3090. https://doi.org/10.1007/s00382-015-2525-1

Jiménez-Muñoz, J.C., Mattar, C., Barichivich, J., Santamaría-Artigas, A., Takahashi, K., Malhi, Y., Sobrino, J.A.& Van Der Schrier, G. (2016) Record-breaking warming and extreme drought in the Amazon rainforest during the course of El Niño 2015–2016. *Scientific Reports*, 6, 33130. https://doi.org/10.1038/srep33130

Junk, W.J., Piedade, M.T.F., Lourival, R., Wittmann, F., Kandus, P., Lacerda, L.D., et al. (2014) Brazilian wetlands: their definition, delineation, and classification for research, sustainable management, and protection. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24(1), 5–22. https://doi.org/10.1002/aqc.2386

Junk, W.J., Piedade, M.T.F., Schöngart, J., Cohn-Haft, M., Adeney, J.M.& Wittmann, F. (2011) A classification of major naturally-occurring Amazonian lowland wetlands. *Wetlands*, 31(4), 623–640. https://doi.org/10.1007/s13157-011-0190-7

Langill, J.C.& Abizaid, C. (2020) What is a bad flood? Local perspectives of extreme floods in the Peruvian Amazon. *Ambio*, 49(8), 1423–1436. https://doi.org/10.1007/s13280-019-01278-8

Lee, C.-Y., Tippett, M.K., Camargo, S.J.& Sobel, A.H. (2015) Probabilistic multiple linear regression modeling for tropical cyclone intensity. *Monthly Weather Review*, 143(3), 933–954. https://doi.org/10.1175/MWR-D-14-00171.1

Liebmann, B.& Marengo, J. (2001) Interannual variability of the rainy season and rainfall in the Brazilian Amazon basin. *Journal of Climate*, 14(22), 4308–4318. https://doi.org/10.1175/1520-0442(2001)014⟨4308:IVOTRS⟩2.0.CO;2

Lima, C.H., Lall, U., Troy, T.J.& Devineni, N. (2015) A climate informed model for nonstationary flood risk prediction: application to Negro River at Manaus, Amazonia. *Journal of Hydrology*, 522, 594–602. https://doi.org/10.1016/j.jhydrol.2015.01.009

Maciel, J.S.C., Alves, L.G.S., Corrrâ, B.G.D.S., De Carvalho, I.M.R.& Oliveira, M.A. (2020) Flood forecast in Manaus, Amazonas, Brazil. *WIT Transactions on the Built Environment: Urban Water Systems & Floods III*, 194, 63–72. https://doi.org/10.2495/FRIAR200061

Mansur, A.V., Brondizio, E.S., Roy, S., Soares, P.P.d.M.A.& Newton, A. (2018) Adapting to urban challenges in the Amazon: flood risk and infrastructure deficiencies in Belém, Brazil. *Regional Environmental Change*, 18(5), 1411–1426. https://doi.org/10.1007/s10113-017-1269-3

Marengo, J.A. (2004) Interdecadal variability and trends of rainfall across the Amazon basin. *Theoretical and Applied Climatology*, 78(1–3), 79–96. https://doi.org/10.1007/s00704-004-0045-8

Marengo, J.A., Borma, L.S., Rodríguez, D.A., Pinho, P., Soares, W.R.& Alves, L.M. (2013) Recent extremes of drought and flooding in Amazonia: vulnerabilities and human adaptation. *American Journal of Climate Change*, 2, 87–96. https://doi.org/10.4236/ajcc.2013.22009

Marengo, J.A.& Espinoza, J.C. (2016) Extreme seasonal droughts and floods in Amazonia: causes, trends and impacts. *International Journal of Climatology*, 36(3), 1033–1050. https://doi.org/10.1002/joc.4420

Marengo, J., Nobre, C.A., Betts, R.A., Cox, P.M., Sampaio, G.& Salazar, L. (2009) Global warming and climate change in Amazonia: climate-vegetation feedback and impacts on water resources. *Amazonia and Global Change*, Geophysical Monograph Series. Vol. 186, 273–292. https://doi.org/10.1029/2008GM000744

Marengo, J.A., Tomasella, J., Soares, W.R., Alves, L.M.& Nobre, C.A. (2012) Extreme climatic events in the Amazon basin. *Theoretical and Applied Climatology*, 107(1–2), 73–85. https://doi.org/10.1007/s00704-011-0465-1

Meade, R.H., Rayol, J.M., Da Conceicão, S.C.& Natividade, J.R. (1991) Backwater effects in the Amazon River basin of Brazil. *Environmental Geology and Water Sciences*, 18(2), 105–114. https://doi.org/10.1007/BF01704664

Nobre, C.A., Marengo, J.A.& Artaxo, P. (2009) Understanding the climate of Amazonia: progress from LBA. *Amazonia and Global Change*, Geophysical Monograph Series. Vol. 186, 145–147. https://doi.org/10.1029/2008GM000716

Paiva, R., Collischonn, W., Bonnet, M.-P.& De Gonçalves, L. (2012) On the sources of hydrological prediction uncertainty in the Amazon. *Hydrology & Earth System Sciences*, 16(9), 3127–3137. https://doi.org/10.5194/hess-16-3127-2012

Papa, F., Prigent, C., Aires, F., Jimenez, C., Rossow, W. & Matthews, E. (2010) Interannual variability of surface water extent at the global scale, 1993–2004. *Journal of Geophysical Research: Atmospheres*, 115(D12). Available from: https://doi.org/10.1029/2009JD012674

Pappenberger, F., Cloke, H.L., Parker, D.J., Wetterhall, F., Richardson, D.S.& Thielen, J. (2015) The monetary benefit of early flood warnings in Europe. *Environmental Science & Policy*, 51, 278–291. https://doi.org/10.1016/j.envsci.2015.04.016

Richey, J.E., Nobre, C.& Deser, C. (1989) Amazon River discharge and climate variability: 1903 to 1985. *Science*, 246(4926), 101–103. https://doi.org/10.1126/science.246.4926.101

Ropelewski, C.F.& Jones, P.D. (1987) An extension of the Tahiti–Darwin southern oscillation index. *Monthly Weather Review*, 115(9), 2161–2165. https://doi.org/10.1175/1520-0493(1987)115⟨2161:AEOTTS⟩2.0.CO;2

Salati, E. & Marques, J. (1984) Climatology of the Amazon region. In: Sioli, H. (Eds.) *The Amazon*. Monographiae Biologicae. Springer, Dordrecht, pp. 85–126. https://doi.org/10.1007/978-94-009-6542-3_4

Santos, L.B., Carvalho, T., Anderson, L.O., Rudorff, C.M., Marchezini, V., Londe, L.R.& Saito, S.M. (2017) An RS-GIS-based comprehensive impact assessment of floods—a case study in Madeira River, Western Brazilian Amazon. *IEEE Geoscience and Remote Sensing Letters*, 14(9), 1614–1617. https://doi.org/10.1109/LGRS.2017.2726524

Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M.& Becker, A. (2017) Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, 8(3), 52. https://doi.org/10.3390/atmos8030052

Schöngart, J.& Junk, W.J. (2007) Forecasting the flood-pulse in Central Amazonia by ENSO-indices. *Journal of Hydrology*, 335(1–2), 124–132. https://doi.org/10.1016/j.jhydrol.2006.11.005

Schöngart, J. & Junk, W.J. (2020) Clima e hidrologia nas várzeas da Amazônia Central. In: Várzeas Amazônicas: Desafios para um Manejo Sustentável, Editora INPA, Manaus. https://repositorio.inpa.gov.br/bitstream/1/36480/3/manejo_sustent%C3%A1vel_das_varzeas.pdf

Sorribas, M.V., Paiva, R.C., Melack, J.M., Bravo, J.M., Jones, C., Carvalho, L., Beighley, E., Forsberg, B.& Costa, M.H. (2016) Projections of climate change effects on discharge and inundation in the Amazon basin. *Climatic Change*, 136(3–4), 555–570. https://doi.org/10.1007/s10584-016-1640-2

Spracklen, D.& Garcia-Carreras, L. (2015) The impact of Amazonian deforestation on Amazon basin rainfall. *Geophysical Research Letters*, 42(21), 9546–9552. https://doi.org/10.1002/2015GL066063

Timpe, K.& Kaplan, D. (2017) The changing hydrology of a dammed Amazon. *Science Advances*, 3(11), e1700611. https://doi.org/10.1126/sciadv.1700611

Towner, J., Cloke, H.L., Lavado, W., Santini, W., Bazo, J., Coughlan de Perez, E.& Stephens, E.M. (2020) Attribution of Amazon floods to modes of climate variability: a review. *Meteorological Applications*, 27(5), e1949. https://doi.org/10.1002/met.1949

Towner, J., Cloke, H.L., Zsoter, E., Flamig, Z., Hoch, J.M., Bazo, J., Coughlan de Perez, E.& Stephens, E.M. (2019) Assessing the performance of global hydrological models for capturing peak river flows in the Amazon basin. *Hydrology and Earth System Sciences*, 23(7), 3057–3080. https://doi.org/10.5194/hess-23-3057-2019

Wang, X.-Y., Li, X., Zhu, J.& Tanajura, C.A.S. (2018) The strengthening of Amazonian precipitation during the wet season driven by tropical sea surface temperature forcing. *Environmental Research Letters*, 13(9), 094015. https://doi.org/10.1088/1748-9326/aadbb9

Williams, E., Dall'Antonia, A., Dall'Antonia, V., Almeida, J.M.d., Suarez, F., Liebmann, B.& Malhado, A.C.M. (2005) The drought of the century in the Amazon Basin: an analysis of the regional variation of rainfall in South America in 1926. *Acta Amazonica*, 35(2), 231–238. https://doi.org/10.1590/S0044-59672005000200013

Xavier, L., Becker, M., Cazenave, A., Longuevergne, L., Llovel, W.& Rotunno Filho, O.C. (2010) Interannual variability in water storage over 2003–2008 in the Amazon Basin from GRACE space gravimetry, in situ river level and precipitation data. *Remote Sensing of Environment*, 114(8), 1629–1637. https://doi.org/10.1016/j.rse.2010.02.005

Yoon, J.-H.& Zeng, N. (2010) An Atlantic influence on Amazon rainfall. *Climate Dynamics*, 34(2–3), 249–264. https://doi.org/10.1007/s00382-009-0551-6

Zubieta, R., Getirana, A., Espinoza, J.C.& Lavado, W. (2015) Impacts of satellite-based precipitation datasets on rainfall–runoff modeling of the Western Amazon basin of Peru and Ecuador. *Journal of Hydrology*, 528, 599–612. https://doi.org/10.1016/j.jhydrol.2015.06.064