AGU100 ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2018WR022984

Key Points:
- A detailed methodology to define and select skillful SST indices to predict seasonal rainfall over Morocco is described
- Simple linear models still perform best for short data records in this region
- About 35-40% of interannual variability can be explained with a 1-month lead time

Supporting Information:
- Supporting Information S1

Correspondence to:
A. Tuel,
atuel@mit.edu

Citation:
Tuel, A., & Eltahir, E. A. (2018). Seasonal precipitation forecast over Morocco. *Water Resources Research*, *54*. https://doi.org/10.1029/2018WR022984

Received 20 MAR 2018
Accepted 23 OCT 2018
Accepted article online 29 OCT 2018

Seasonal Precipitation Forecast Over Morocco

A. Tuel[1] **and E. A. B. Eltahir**[1]

[1] Ralph M. Parsons Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract Interannual variability of precipitation is a challenge for water management in semiarid Morocco, where droughts are common and take a heavy toll on agricultural production. The question of forecast potential for winter and spring rainfall is therefore crucial for the country. As an addition to dynamical seasonal predictions, taking the case of the Oum Er Rbia watershed, the present study builds simple statistical forecasting models for Morocco's winter and spring precipitation, making use of oceanic forcing of atmospheric circulation. Previous results on the teleconnections between North Atlantic wintertime circulation and lagged ocean surface temperatures are confirmed. Based on a careful identification and selection of predictors, about 35–40% of interannual variability of precipitation may be robustly predictable with a lead time of 1 month.

1. Introduction

A large fraction of Morocco's arable land (80%) is located in arid and semiarid regions, mostly in the northeast (Moulouya Valley), central west (Casablanca, Beni Mellal, and Marrakech), and southwest (Souss) of the country. These regions are characterized by limited and erratic precipitation, with a high risk of drought. In parallel, they account for 55% of the national cereal production and support about half of the country's livestock (El Mourid & Karroui, 1996). Agriculture represents 15% to 20% of the country's gross domestic product and employs 40% of its workforce (Balaghi et al., 2012). The high interannual variability in precipitation has been a constant obstacle to efficient water management and agricultural planning over the years. As most of the cropland is rainfed and because interannual water storage is limited, crop water demand must usually be satisfied by the same year's precipitation. Precipitation therefore strongly impacts agricultural production and crop loss can be significant in times of drought (see Figure 1). As an example, during the 1994–1995 cropping season, drought was so severe that the precipitation deficit reached −60% in the north and −80% in the south. Cereal production for that season totaled a mere 1.73 million tonnes (average yield: 479 kg/ha), compared to 9.8 million tonnes the following season (1,730 kg/ha), when precipitation was in excess (Balaghi et al., 2012).

Seasonal forecasts based on dynamical models have thus far displayed very limited skill, which makes them unreliable in practice (see Method S1 and Figure S1 in the supporting information for a more quantitative discussion). Therefore, in the current decision-making system, planners base their decisions on climatological probabilities for precipitation. This compels them to adopt very conservative risk levels in order to avoid water shortages should severe droughts occur, leading to high inefficiency since it is not optimal to hedge every year against a drought. As a result, wet years are not taken advantage of as much as they could. This underlines the substantial need for precipitation forecasts in Morocco. Here we focus on one particular region, the Oum Er Rbia watershed, in central Morocco, an area of high economic importance for the country. It concentrates 30% of the country's irrigated land and half of its hydroelectric capacity and supplies water to over seven million people (including two of Morocco's largest cities, Casablanca and Marrakech). Currently, 90% of all water is used for irrigation, and 95% of available water resources are mobilized each year.

Traditionally, seasonal climate forecasting has either relied on numerical weather models (NWMs) or on statistical models. Thanks to improvements in computational efficiency, NWMs have become dominant for short- to medium-term forecasts, from 1 to about 10 days. Statistical forecasts, on the other hand, still find use in long-term forecasts, also known as season-to-season forecasts (S2S) or simply seasonal forecasts, with lead times considerably longer than the limit of deterministic atmospheric predictability (Wilks, 2005). Statistical forecasting algorithms are plenty, but all require a series of predictor variables or covariates. Forecasting a variable of interest thus requires first to identify such covariates and then to implement a forecast model.

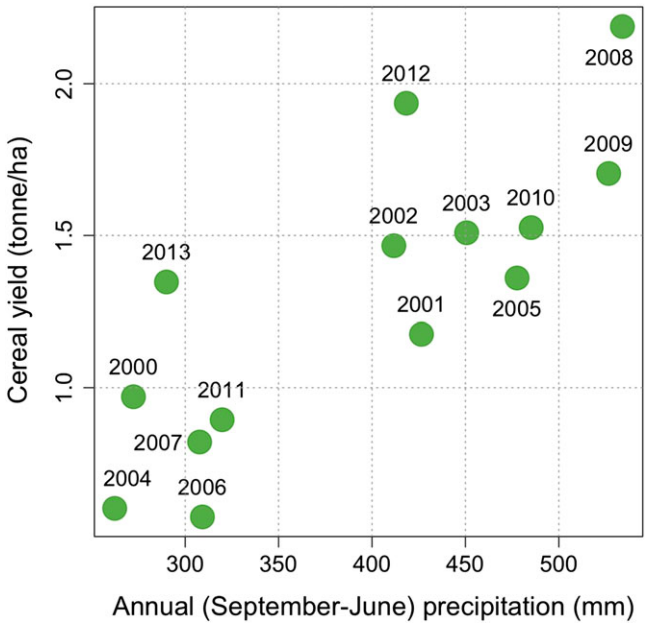

Figure 1. Average cereal yield (t/ha) as a function of total annual precipitation for the Oum Er Rbia catchment area, 2000–2013 seasons (data from M. El Gharous personal communication, August 8, 2017). Cereals are the main rainfed crop in the region; only more high-value crops, such as citrus fruits, sugar beet, or vegetables, tend to be irrigated.

Our methodology falls therefore into two parts (Figure 2): covariate definition on the one hand and covariate selection with model fit on the other.

Previous studies identified relevant areas of sea surface temperatures (SSTs) teleconnected to wintertime weather over our region of interest (Czaja & Frankignoul, 2002; Peng et al., 2003). Building on these results, we implement a series of statistical methods to identify, within those regions, the areas most connected to Moroccan precipitation (section 5). Then, in a second step, we fit linear models and identify small groups of robust predictors to get the best forecasts of three-month precipitation totals, as compared to observations.

2. Regional Precipitation Dynamics

In the northwestern quadrant of Morocco, delimited by the Atlas and Rif ranges, which contains the Oum Er Rbia watershed, the climate is mediterranean — mild and wet in winter and hot and dry in summer. Average annual precipitation varies significantly over this region, generally increasing with both latitude and elevation. Thus, the low point is around Marrakech to the south, with only 200 mm of annual precipitation, while mountainous regions to the north receive well over 1,000 mm. The wet season extends from November until April, when moist air is advected from the Atlantic ocean, located to the west and northwest. It accounts for about 90% of annual precipitation. Precipitation events during this season tend to be short and irregular, occurring when midlatitude weather systems, that formed over the Atlantic ocean, drift toward the coast, under the influence of prevailing winds. Typically, in the northwestern Maghreb, a rain season consists of at most 10–14 single precipitation events (Born et al., 2008).

A closer look at the annual precipitation cycle in the region reveals the presence of two modes in the distribution, with one peak in the early part of the wet season (around December) and a second in the later part, around March (see Figure 3). Precipitation thus appears to arrive in two independent waves, one in winter, the other in spring, with no apparent link between the two. For instance, Spearman's correlation coefficient between November-to-January and February-to-April precipitation is only 0.11 for the whole 1901–2015 period, and 0.17 for 1950–2015, both of which are insignificant at the 5% level of confidence. This suggests that precipitation during the wet season may have different physical origins depending on the time of year and that teleconnections for each "wave" will be different. This motivates us to divide the wet season into several subseasons and develop forecast models for each. We focus in this study on three precipitation series: November to January (NDJ), December to February (DJF), and February to April (FMA). A distinction is made between NDJ and DJF because, though the teleconnection of precipitation to the North Atlantic Oscillation (NAO) usually steps in during November, submonthly NAO variability may cause it to be delayed until early December, thus impacting predictability of precipitation. Separating NDJ and DJF also helps determine how predictable early-wet season precipitation is.

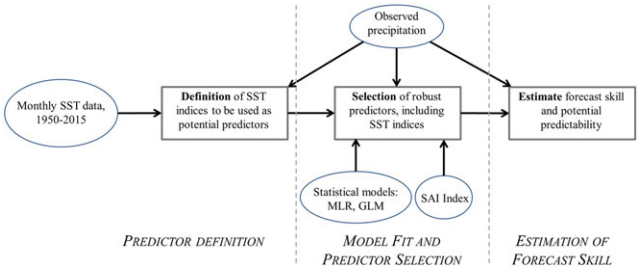

Figure 2. Summary of methodology followed in this paper. The first step is to define potential precipitation predictors from monthly SST data and observed seasonal precipitation. Then statistical models are fitted, and the most robust, or informative, predictors are selected. Finally, a traditional assessment of forecast skill is implemented. SST = sea surface temperature; MLR = multiple linear regression; GLM = generalized linear model; SAI = Snow Advance Index.

Coupled to its spatial variability, precipitation is also highly variable in time, both intraseasonally and interseasonally. The annual coefficient of variation ranges from an average of 30% in the northwest to over 100% in the most arid areas. Local climate is characterized by long periods of droughts that tend to affect wide areas. Short-term seasonal lack of moisture is also common, either at the beginning (November–December) or at the end (March–April) of the growing season.

3. Identified Teleconnections

A large body of work exists on the teleconnections to wintertime western European and northwestern African weather. In late fall and winter

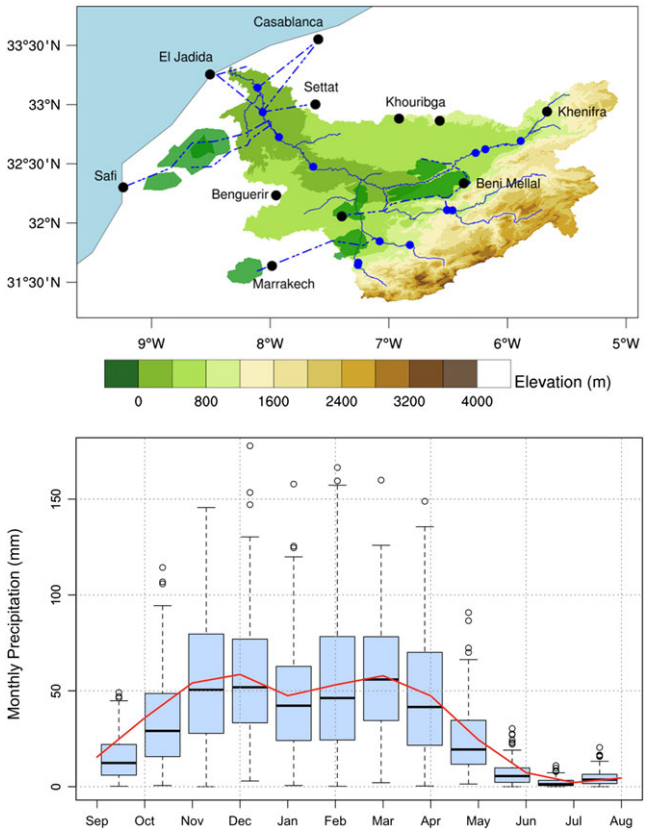

Figure 3. (top) Map of the Oum Er Rbia watershed, showing elevation (filled contours), main cities (black dots), dams (blue dots), channels (dot-dashed blue line), and irrigated perimeters (dark green). (bottom) Annual precipitation cycle over the Oum Er Rbia basin, from September to August.

(November–February), precipitation over northwestern Morocco is strongly related to the NAO, a large-scale pressure dipole characterized by a large region of relatively high pressure centered over the Azores islands (Azores High) and one of low pressure centered over Iceland (Icelandic Low; Hurrell, 1995). This bipolar mode is linked to patterns of winter temperature, precipitation, and storminess over the whole North Atlantic sector (Hurrell, 1995), because it creates powerful jets blowing from west to east, which direct storms toward Western Europe and northwestern Africa. As the relative strength of the two pressure poles evolves, due to atmospheric mass being alternately redistributed between the Arctic and subtropical Atlantic areas, so does the intensity and position of the storm track. These dynamics are characterized by the NAO index, which can be defined as the difference of renormalized sea level pressure (SLP) values between two stations, one in Iceland (Reykjavik) and one to the south (Gibraltar, Lisbon, or Ponta Delgada in the Azores; Hurrell et al., 2013).

In the spring, however, as the SLP dipole drifts north, the connection to the NAO is gradually lost. For that period, studies have suggested a stronger link to Pacific SSTs, most notably El Niño-Southern Oscillation (Mariotti et al., 2002; Shaman & Tziperman, 2010). The physical teleconnection may be due to the development of a Rossby wave train that propagated to the North Atlantic, creating a dipole of anomalous vorticity off the coasts of Europe and northwestern Africa and stimulating moisture advection from the Atlantic (Shaman & Tziperman, 2010).

The teleconnection of winter precipitation to NAO is very strong, but the latter has proven remarkably challenging to predict. Many studies have suggested that a large fraction of its dynamics arose from internal, nonlinear phenomena intrinsic to the atmosphere (Feldstein, 2000). Nevertheless, while subseasonal variability of the NAO might arise from climate noise, external forcings likely play a significant role on longer timescales (Smith et al., 2016), most notably the ocean (Bojariu & Gimeno, 2003). Given its large thermal inertia, ocean temperatures force the lower atmosphere over the course of months, making them very useful predictors for a wide range of climatic phenomena around the globe. Studies by Rodwell et al. (1999) and Saunders and Qian (2002) uncovered links between SST anomalies in the months leading to November and winter NAO. In particular, emerging in late spring and persisting until early fall is a noticeable North Atlantic horseshoe pattern (Czaja & Frankignoul, 2002), with warm anomalies southeast of Newfoundland and cold anomalies to the northeast and southeast, which precedes positive phases of the NAO and could explain up to 15% of its variability. For the North Atlantic domain, SSTs likely play a role in determining the position, extent, and strength of the meridional temperature gradient that in turn drives baroclinic instability and the storm track in the region. There is still debate about which specific region of the Atlantic is most strongly linked to winter NAO, from tropical and subtropical regions (Czaja et al., 2013) to areas off the UK or Newfoundland (Saunders & Qian, 2002). Connections to other major ocean basins have been proposed as well, such as the southern Atlantic (Robertson et al., 1999) or the tropical Indian and Pacific oceans (Sutton & Hodson, 2003). Models and regression analyses suggest however that the oceanic influence on North Atlantic climate is nonstationary and appears to have varied, sometimes significantly, over the last two centuries (Hoerling et al., 2001).

Other NAO forcings have been suggested in the literature; a recent study by Cohen and Jones (2011) pointed to the October growth rate of Eurasian snow cover as a likely NAO driver, acting through land-atmosphere forcing. They summarized this forcing in the Snow Advance Index (SAI), which we will include as a potential covariate in our models.

Finally, in recent years, attention given to climate and weather models has also grown considerably. The improvement of computational capacities made it possible to simulate winter weather of the North Atlantic

in much better detail, and some of the latest forecasts from NWMs, though they underestimate the variability of the NAO, show good skill at predicting its phase (Dunstone et al., 2016).

4. Data

Precipitation data are taken from the CRU TS4.00 database (Harris & Jones, 2017), which extends from January 1901 to December 2015. The GPCC data set (Schneider et al., 2011) is also used for comparison and to get estimates of precipitation (for 2017, which is missing from CRU TS4). The most skillful numerical NAO hindcasts at a 1-month lead are the ones published in Dunstone et al. (2016), covering the winters 1981–1982 to 2015–2016. To be consistent with that study, observed NAO values are computed from the HadSLP2 database (Allan & Ansell, 2012) as the difference in monthly SLP between the Azores (28–20°W, 36–40°N) and Iceland (25–16°W, 63–70°N). In addition, we use as predictor fields the SST data from the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) data set (Rayner et al., 2003) at 1° resolution, covering the period 1870 to present. Because SST data before the 1940s has been shown to be less precise and to have generally larger uncertainty (Kennedy, 2014) and because station data for precipitation display several missing years throughout the 1920s and 1930s, we choose to focus on the period 1950–2015. We restrict our analysis to the North Atlantic (0°N to 65°N; 80°W to 6°W), Pacific (50°S to 65°N; 120°E to 65°W) and Mediterranean (30°N to 46°N; 6°E to 37°E) domains, as the literature suggests that they may be important (Czaja & Frankignoul, 2002; Czaja et al., 2013; Rodwell et al., 1999; Saunders & Qian, 2002; Sutton & Hodson, 2003). Finally, the 1972–2015 SAI is taken from Cohen and Jones (2011).

Though other climate variables, like sea-ice concentrations, are believed to force winter atmospheric circulation (Totz et al., 2017), the focus here is mainly on SSTs, which have a long and reliable record going back to at least 1950 (Kennedy, 2014). Doing so excludes the potential additional information included in other variables but allows us to construct our forecast models on much longer time periods.

5. Defining Potential Predictors

5.1. Methods

SST fields are complex and highly multidimensional. To extract meaningful indices from those thousands of potential covariates at various different lead times, we display maps of several measures of statistical relationships between precipitation series and lagged SSTs. We go back as far as 9 months in advance (May of the preceding hydrological year), as Czaja and Frankignoul (2002) suggested that North Atlantic spring SSTs may force wintertime atmospheric circulation.

We first identify links between SST fields and precipitation series using well-known statistical methods, on a point-by-point basis, that is, by considering the SST at each point in space as a potential predictor. In addition to correlation maps, we look at SST anomalies preceding a dry, average, or wet season (composite maps). We also implement an analysis of variance procedure and calculate maps of mutual information between SSTs and precipitation time series, both of which measure how well the variability of a predictor explains that of the predictand. To take into account the strong spatial covariance of SST fields, we also rely on canonical correlation analysis, a statistical technique similar to multilinear regression that identifies maximally correlated patterns in the predictand and the predictor. Overall, these varied methods can highlight a wide spectrum of potential relationships. Their comparison helps separate spurious predictors from those with high potential skill. Using different methods also avoids being constrained by the assumptions behind a single one, nor by any p-value thresholds. Further details of the statistical methods are given in Methods S2.

For all methods, significance is based on bootstrap confidence intervals constructed by generating 5,000 random maps for each predictand. Potential SST predictors are then defined by averaging SST data encapsulated in the areas of highest significance (at a 95% confidence level), selected among those extending at least 15° in each dimension. This threshold reflects the length scale of SST variability in the selected ocean basins and is high enough to avoid defining too many potential predictors, which would inevitably lead to overfitting. We end up with lists of about two dozen potential indices for each season, at many different lead times.

5.2. Results

Figure 4 (left) highlights the most important SST predictors for each precipitation series, as selected during model fit (see section 6 for selection details). Both the NDJ and DJF precipitation series appear to be most influenced by extratropical North Atlantic SSTs, from late spring until late fall. This is in keeping with previous

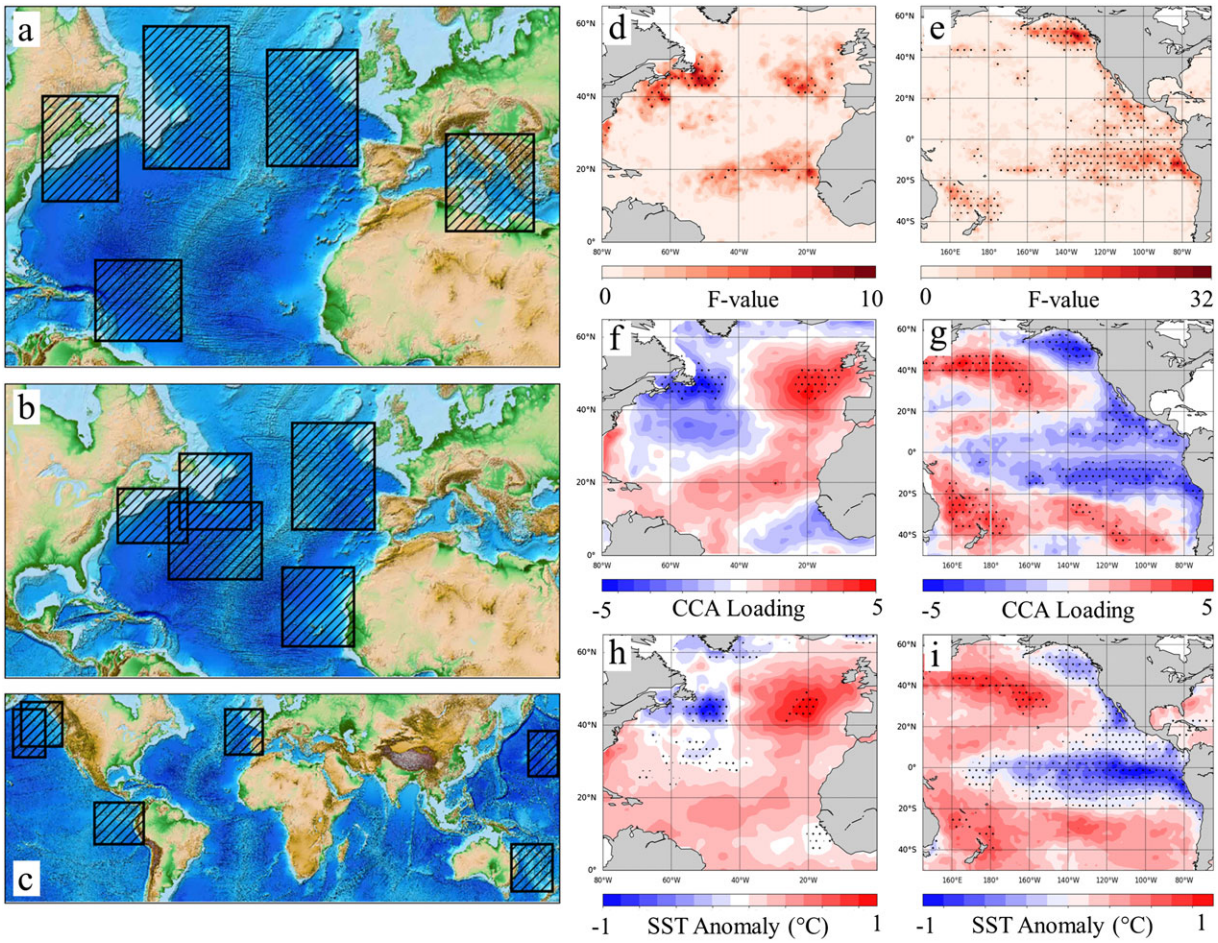

Figure 4. (left) Maps of each season's robust precipitation SST predictors: NDJ (a), DJF (b), and FMA (c). Robust predictors are defined as the ones selected most often during the cross-validation process (see section 6). The predictors shown on the above maps were selected in the models at least 90% of the time. (right) Example of results from three predictor identification methods ((d and e) analysis of variance, (f and g) CCA, and (h and i) composite maps) between the DJF precipitation series and October North Atlantic SSTs (d, f, and h) and the FMA precipitation series and December Pacific SSTs (e, g, and i). Stipling corresponds to the significance of the results with 95% confidence, as explained in the body of the article. Potential predictors were selected from such maps, drawn for every possible SST domain/lag time/precipitation series combination, and keeping areas found significant by all methods and extending at least 5° in each dimension. CCA = canonical correlation analysis; SST = sea surface temperature; NDJ = November to January; DJF = December to February; FMA = February to April.

results highlighting specific SST patterns in the same area linked to wintertime NAO (Hoerling et al., 2001; Rodwell et al., 1999; Saunders & Qian, 2002; see section 3).

Regarding the FMA series, the connection to the North Atlantic is overall weaker than for the other two series, though a horseshoe-like structure does appear in spring SSTs as well. The concurrent NAO still impacts spring precipitation, albeit much less so than in winter. The pattern of statistical links with winter (November to January) SSTs resembles that of the winter NAO itself, hinting at some potential feedback mechanism between winter NAO, winter SSTs, and spring atmospheric circulation. A wide region of positive correlations stretches from the eastern seaboard of the United States to northwestern Europe, lined with two areas of negative correlation values. FMA precipitation is most significantly linked, however, to Pacific SSTs, from early summer to December. The pattern found by all methods is very similar to that of the Pacific Decadal Oscillation, one of the major modes of SST variability in the Pacific Ocean (Trenberth & Fasullo, 2013).

A comparison of some results from three methods is shown on Figure 4 (right). Overall, the different approaches to predictor definition highlighted above agree quite well on what areas are significantly linked to precipitation over the Oum Er Rbia, though the magnitude of the link and exact geographical delimitations vary somewhat between methods.

6. Forecast Model and Predictor Selection

6.1. Methods

6.1.1. Statistical Models

Empirical forecast models are built for each precipitation series (NDJ, DJF, and FMA) by relying on the information contained in the SST indices defined in section 5. Two main models are considered: multiple linear regression (MLR) and the generalized linear model (GLM; Dobson, 1990) with a logarithmic link function. The latter is considered because it brings a slight improvement in fit when regressing winter precipitation against the concurrent NAO index (see section 7). In MLR, the expected value of seasonal precipitation y is modeled as a linear response of given covariates X: $\mathbb{E}[y] = X\beta$, while the log-link GLM model assumes $\ln(\mathbb{E}[y]) = X\beta$. In addition to the exact value of precipitation, we are interested in forecasting discrete events, like droughts. To that end, we define precipitation categories (below , near, and above average) based on the terciles $Y_{33\%percnt;}$ and $Y_{67\%percnt;}$ of each series. Thus, years for which the seasonal precipitation was below its 33% percentile ($Y_{33\%percnt;}$) were categorized as "dry." For each year t, by sampling regression coefficients and residuals, an empirical distribution $p(y, t)$ of the forecasted precipitation is estimated, from which we determine confidence intervals and probabilistic categorical forecasts, the probability of a dry year being equal to $\int_0^{Y_{33\%percnt;}} p(y,t)dy$.

The predictive potential of the best NAO forecasts (Dunstone et al., 2016) is estimated via MLR using two cross-validation methods. In the first, the model is trained on observed NAO and precipitation from 1901 to 1980 and validated using NAO forecasts over 1981–2015, rescaled to have the same mean and variance as observations. The second only relies on 1981–2015 hindcasts and a leave-one-out approach is selected.

6.1.2. Predictor Selection

The large number of potential predictors defined in section 5 must be narrowed down to avoid overfitting, given the limited length of the time series. We do this in two steps: The first, based on LASSO regression using the R package `glmnet` (Friedman et al., 2010; R Core Team, 2013), helps select a small subset of robust potential predictors (typically 12–15) out of the many initial ones. LASSO is a well-known, useful technique that performs feature selection by adding to the error a penalty term to favor a sparse vector of regression coefficients (details and references can be found in Friedman et al., 2010). In a second step, the most informative predictors are selected by a combined forward and backward variable selection process, also known as stepwise regression (see Method S3 for details).

In addition, a cross-validation procedure must also be implemented to assess model skill in the most objective way possible. Because feature selection implies looking at the data to extract the most relevant predictors, in order to avoid any bias in skill, both predictor screening and model building must be cross-validated simultaneously (Michaelsen, 1987).

We adopt a k-folds cross-validation, by dividing the sample into subsamples of size k. Each of those is withheld in turn for validation, with all remaining years serving as training set. Thus, if $k = 10$, forecasts for the years 2006–2015 are issued by models trained on 1950–2005. The value of k is varied from 5 to 25, to examine the robustness of the models. After separating the data set into training and validation sets, the two feature selection steps (LASSO and stepwise regression) are conducted on the training set only, and model skill is assessed for the left-out validation data. This cross-validation methodology ensures that the data against which model skill is evaluated is never used at any step of the model building process.

Our predictor definition methodology described in the previous section did involve using all data points to compute analysis of variance maps, canonical correlation analysis coefficients, and so forth. Still, because this process is time-consuming and involves a lot of selection by hand, it could not be repeated independently at every step of the cross-validation. A few trials have however shown no significant differences when selecting different subsets of the observed data. Most relationships appear strong and stable enough so that similar predictors would have been selected in any case.

Based on this procedure, different optimal combinations of predictors may be — and, in fact, are — selected to build the models, depending on the training set. This variation should not be seen as an attempt to maximize skill at the expense of physical interpretability; on the contrary, it helps avoid underfitting or overfitting of the models for given years and accounts for the uncertainty both in the importance of each predictor and in the number of optimal predictors. This approach can be considered more realistic and should better reflect the actual skill of the models, since forecasted years are clearly separated from the model-building process.

Table 1
Regression (r^2; RMSE in Millimeters) and Classification (FI, HS, and ROC) Skill Scores for the Three Precipitation Series at Different Time Leads, for the MLR and GLM Models, Over the 1950–2015 Period

	MLR					GLM				
	r^2	RMSE	FI	HS	ROC	r^2	RMSE	FI	HS	ROC
NDJ	0.12	58.2 (0.95)	0.44	0.51	0.67	0.14	55.5 (0.93)	0.44	0.49	0.64
DJF, 1 Nov	0.34	62.1 (0.83)	0.45	0.51	0.68	0.37	57.1 (0.76)	0.46	0.53	0.76
DJF, 1 Dec	0.37	60.5 (0.81)	0.46	0.56	0.73	0.41	55 (0.72)	0.47	0.54	0.80
FMA, 1 Jan	0.33	52 (0.81)	0.50	0.59	0.71	0.36	51 (0.8)	0.48	0.58	0.74

Note. We use a *k*-folds cross-validation strategy with *k*=10. The number in brackets in the RMSE column corresponds to the percentage of the standard deviation of the corresponding time series that the forecast RMSE is equal to. RMSE = root-mean-square error; FI = forecast index; HS = hit score; ROC = relative operating characteristic; MLR = multiple linear regression; GLM = generalized linear model; NDJ = November to January; DJF = December to February; FMA = February to April.

6.1.3. Skill Scores

For continuous forecasts, we use the coefficient of determination r^2 and root-mean-square error to evaluate the accuracy of the models. For categorical forecasts, a simple skill measure is the hit score HS, equal to the percentage of correct hits (i.e., the right category is predicted). We also consider the forecast index FI, defined as the average probability forecasted for the realized category (see Method S4 for details). A perfect model would have both FI = HS = 1 (all categories are forecasted correctly, with a probability of 1), and a no-skill model would have FI = HS = 1/3 (random category is forecasted). To assess how well the models warn against specific events, for instance, droughts, we also look at the relative operating characteristic (ROC) score, based on the comparison of hit and false-alarm rates for different warning thresholds. Models with ROC > 0.5 display skill (hit rate is higher than false-alarm rate), while a no-skill forecast will have ROC = 0.5 (hit rate and false-alarm rate are equal).

6.2. Results

The approach based on NWM NAO hindcasts (Dunstone et al., 2016) shows little skill, with the classifier trained on hindcasts only yielding FI = 0.36 and HS = 0.40. The model trained on 1901–1980 observations and validated on 1981–2015 hindcasts does not do as well as the one validated using NAO observations, as expected, but still shows a little bit of forecast potential ($r = 0.37$ against $r = 0.56$, $FI = 0.35$ against $FI = 0.38$, and $HS = 0.37$ against $HS = 0.41$). However, the calculation of the NAO as in Dunstone et al. (2016) may not be optimal to forecast winter weather in northwestern Africa; studies have shown that the southern pole of the NAO between the Azores and Gibraltar had significantly more influence on that region than SLP conditions over Iceland (Knippertz et al., 2003). Instead of summarizing the whole SLP field in the NAO index, it may be better to look at more specific SLP domains.

Models are built and fitted for all three seasons; for the DJF and FMA season in particular, we issue forecasts both on the first day of the season and one full month before the start of the season. As an example, for DJF, a model is built using all information up to 1 November, and another using all information up to 1 December. The NDJ and DJF seasons are also analysed independently over the 1972–2015 period to estimate the improvement in skill brought by the SAI. It must be noted, though, that because of the time required for processing the data, the full SST analysis fields are only available more than a month after the month of interest, that is, July SST data are available in early September. However, that delay is mainly due to the sea-ice homogenization process, and a preliminary SST field is made available on the 10th of the month after the month of interest (10 August for July SSTs). For all models, 5–7 predictors are retained after feature selection. Figure 4 (left) shows the robust SST predictors that were selected most often by each model.

As we can attest from Table 1, forecast skill is rather low for NDJ but improves significantly for the other two series. Predictability is highest for the DJF series, with r^2 close to 0.4. By comparison, dynamical models hardly reach $r^2 \sim 0.1$ at comparable time leads (see Method S1). Model skill can safely be said to be significantly above that of climatology. DJF is more predictable than NDJ, possibly because the NAO pattern during November is less stabilized, and thus less predictable, or because the beginning of the wet season in Morocco may shift by several weeks, thus adding noise to the NDJ data. DJF predictability is also more robust than that of FMA, since, when increasing *k* in the cross-validation, r^2 remains close to 0.40 for DJF but drops to 0.27 for FMA

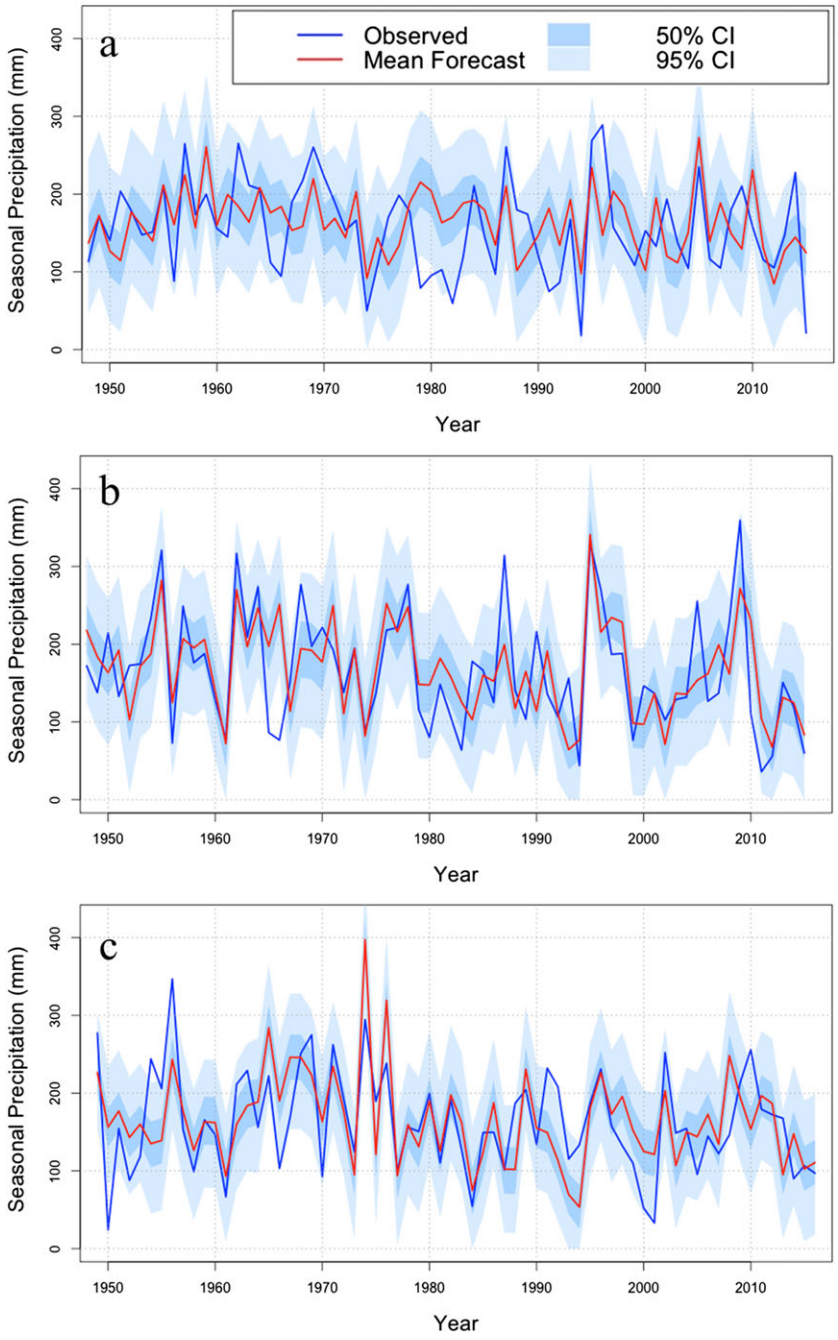

Figure 5. Results of generalized linear model precipitation forecasts for the November-to-January series, issued on 1 November (a); the December-to-February series, issued on 1 December (b); and the February-to-April series, issued on 1 January (c). The mean forecast is shown by the red line, observed precipitation by the blue line, alongside 50% (shaded in blue) and 95% (shaded in light blue) confidence intervals. CI = confidence interval.

(see Figure S2). Figure 5 shows the mean NDJ, DJF, and FMA seasonal GLM forecasts for our study period, alongside 50% and 95% confidence intervals estimated by sampling from the multilinear regression parameters. The GLM model performs better than the MLR model for all seasons. Both the increase in r^2 and the reduction in root-mean-square error are significant, though the skill is more mixed when it comes to categorical forecasts (Table 1). This may be due to the asymmetry in the confidence intervals caused by the log-transformation. As a further test for the models, we also make forecasts for the NDJ, DJF, and FMA seasons of hydrological years 2016 and 2017 (see Table S1 for details).

Table 2
GLM Skill Scores for the 1972–2015 Period, With and Without the Inclusion of the SAI

	GLM				
	r^2	RMSE (mm)	FI	HS	ROC
NDJ	0.12	60.2 (1.0)	0.42	0.46	0.58
NDJ + SAI	0.28	56.4 (0.94)	0.43	0.51	0.69
DJF, 1 Nov	0.22	70.3 (0.96)	0.39	0.46	0.62
DJF + SAI, 1 Nov	0.25	68.1 (0.92)	0.43	0.50	0.67
DJF, 1 Dec	0.25	67.6 (0.91)	0.41	0.44	0.65
DJF + SAI, 1 Dec	0.27	64.2 (0.86)	0.43	0.51	0.71

Note. We use a *k*-folds cross-validation with $k = 10$. The number in brackets in the RMSE column corresponds to the percentage of the standard deviation of the corresponding time series that the forecast RMSE is equal to. GLM = generalized linear model; SAI = Snow Advance Index; RMSE = root-mean-square error; NDJ = November to January; DJF = December to February; FI = forecast index; HS = hit score; ROC = relative operating characteristic.

Figure 4 (left) shows that for both the NDJ and DJF series, forecast skill mainly relies on North Atlantic SST data, essentially corresponding to the horseshoe pattern between May and November. The central Mediterranean basin is also selected for NDJ, although given the weak predictability of this series, the interpretation is subject to caution. The FMA series, however, is much more strongly connected to the Pacific basin, with much of the skill coming from the three poles of the Pacific Decadal Oscillation-like correlation structure. Robust predictors are detailed in Table S1.

Skill improvement between the first DJF forecast (issued 1 November) and the second (issued 1 December) is small, though detectable (Tables 1 and 2). The additional November SST information may reflect the ocean-atmosphere forcing feedback, with November SSTs giving information about the dominant, or "likely" NAO pattern that gets established for the winter. In the case of FMA, however, adding January SST information brings no predictability gain, in keeping with our observation that the teleconnection to the Pacific Ocean peaked in December.

The impact of the SAI (see section 3), detailed in Table 2, is on the whole positive, although very limited, and higher for NDJ than for DJF, even if the shortness of the record is a cause for caution. There is a decrease in all skill scores except for NDJ compared to the full record, as expected from the fact that we are fitting the models on a shorter time period. The small improvement due to the SAI suggests that the series brings independent additional information. Eurasian snow cover may indeed play a complementary role to SSTs in forcing wintertime atmospheric circulation: Whereas SSTs impact the strength and extent of the meridional temperature, the SAI would reflect a land-atmosphere forcing process over Siberia that also influences the state of the winter NAO (Cohen & Jones, 2011). In recent years, however, the link between the SAI and NAO has degraded. It is a question whether climate-change induces trends in the snow cover might disrupt this link, a question that requires more analysis and is beyond the scope of this paper.

Examination of the ROC score shows that models show on average a good ability to predict droughts, with values of 0.75-0.8 achieved for DJF and FMA. The ROC score can also be computed on each precipitation category taken separately, to estimate how well the model is able to predict it against the other two (see Figure S3 for the ROC curves obtained with the GLM model for the DJF season). The near-normal category systematically has a lower ROC score, which is due to the fact that its probability is much less sensitive to the potential predictability of the underlying phenomenon (Kharin & Zwiers, 2003). A ROC score of about 0.8 for the DJF season corresponds to a potential predictability (r^2) of about 0.4, in keeping with the findings of Kharin and Zwiers (2003) in the case of a Gaussian setting. One must be careful while interpreting those results, though, as the ROC score is relatively independent of forecast calibration. This means that two forecasts of different biases will have equal ROC scores (Kharin & Zwiers, 2003). In consequence, the ROC tends to overestimate the skill that can be achieved in practice.

7. Discussion

7.1. Physical Mechanisms and Forecast Skill

Though the models display a good ability to predict precipitation with a limited number of robust predictors, it remains to determine whether those predictors actually represent the underlying physical processes associated with seasonal variability over Morocco.

Precipitation in Morocco almost systematically comes from the northwest (see section 2); the determining factor is the direction of the winds (Knippertz et al., 2003). Mechanisms connecting SST patterns and large-scale atmospheric circulation over the North Atlantic have already been discussed in previous studies, whether for the winter NAO (Czaja & Frankignoul, 2002; Peng et al., 2003) or for a fall teleconnection to the Pacific Ocean (Shaman & Tziperman, 2010). North Atlantic SSTs, in particular the horseshoe pattern, have long been believed to force the winter NAO, via direct diabatic heating or transient eddy-mean flow interactions. Peng et al. (2003) suggest that an anomalous flow is associated with heating from SST anomalies, in turn creating an anomalous eddy forcing which drives an anomalous eddy-forced flow. In other words, North Atlantic SST anomalies, especially off of Newfoundland, impact heat release in areas of maximum eddy-growth for the North Atlantic stormtrack (Gastineau et al., 2016). This may weaken baroclinicity and in turn impact the strength and position of the stromtrack, which determines the NAO index.

For DJF, composite precipitation and circulation anomalies, associated with high (above the 67% percentile) and low (below the 33% percentile) values of each robust SST predictor from Figure 4, display significant changes in the westerlies around 30°N and 50°N (see Figures 6a–6f). Strengthened (respectively weakened) westerlies to the south are linked to weakened (respectively strengthened) westerlies to the north. Precipitation anomaly fields are consistent with this picture, with positive precipitation anomalies linked to strengthened westerlies (and vice versa). Overall, this signature is very similar to that of the NAO, albeit of smaller magnitude, and the analysis corroborates the relevance of the selected predictors as forcing variables for winter atmospheric circulation.

Regarding teleconnections to FMA precipitation, current knowledge is less specific. Though a number of studies have identified teleconnections between tropical Pacific SSTs and European precipitation (see Shaman & Tziperman, 2010, for a comprehensive review), physical explanations are yet to be identified. Positive precipitation anomalies are still linked to strengthened westerly low-level winds, but there is no large dipolar structure over the North Atlantic. A physical explanation during the fall, developed by Shaman and Tziperman (2010), consists in a Rossby wave train emanating from the Eastern Pacific, which causes vorticity anomalies off of the Iberian Peninsula.

In keeping with their methodology, we display 500-hPa vorticity anomalies associated with high/low values of the SST predictands, over the Eastern Pacific and North Atlantic sector (Figures 6g–6l). Pacific SST indices are associated with large-scale alternating vorticity anomalies extending from the Pacific to the North Atlantic. A patch of vorticity anomaly consistently appears west of the Iberian Peninsula (its sign depending on the selected SST index), accompanied by a change in westerlies around 40°N that impacts precipitation over the region. These results are similar with those of Shaman and Tziperman (2010) and consistent with an impact due to planetary waves originating from the tropical Pacific. The time lag between predictors and predictands here is too long for the teleconnection to be due to direct Rossby wave trains (Shaman & Tziperman, 2010). The link may however arise from the persistence of Pacific SSTs over the course of several months. As to the single North Atlantic SST predictor (Figure 6l), it comes with a comparable vorticity anomaly field but limited to the North Atlantic only.

7.2. Some Limitations

A criticism that can be made is that the whole record is used to identify potential predictors, which, in spite of the double cross-validation strategy, likely introduces artificial skill into the models. However, doing this procedure for every step of the cross-validation is impossible in practice and would probably have yielded similar results. This we checked for several SST-precipitation combinations and found on the whole little variability, even if only 75% of the record was used. We hope that selecting a wide range of potential predictors (see Table S1) helps deal with probable nonstationarities in the ocean-atmosphere feedback. The cross-validation process we implement is also intended as a way to prevent overfitting.

Yet the main drawbacks of hand-engineered covariates is that they are time-consuming and not scalable in practice. We attempted to define predictors based on the whole information contained in each monthly SST

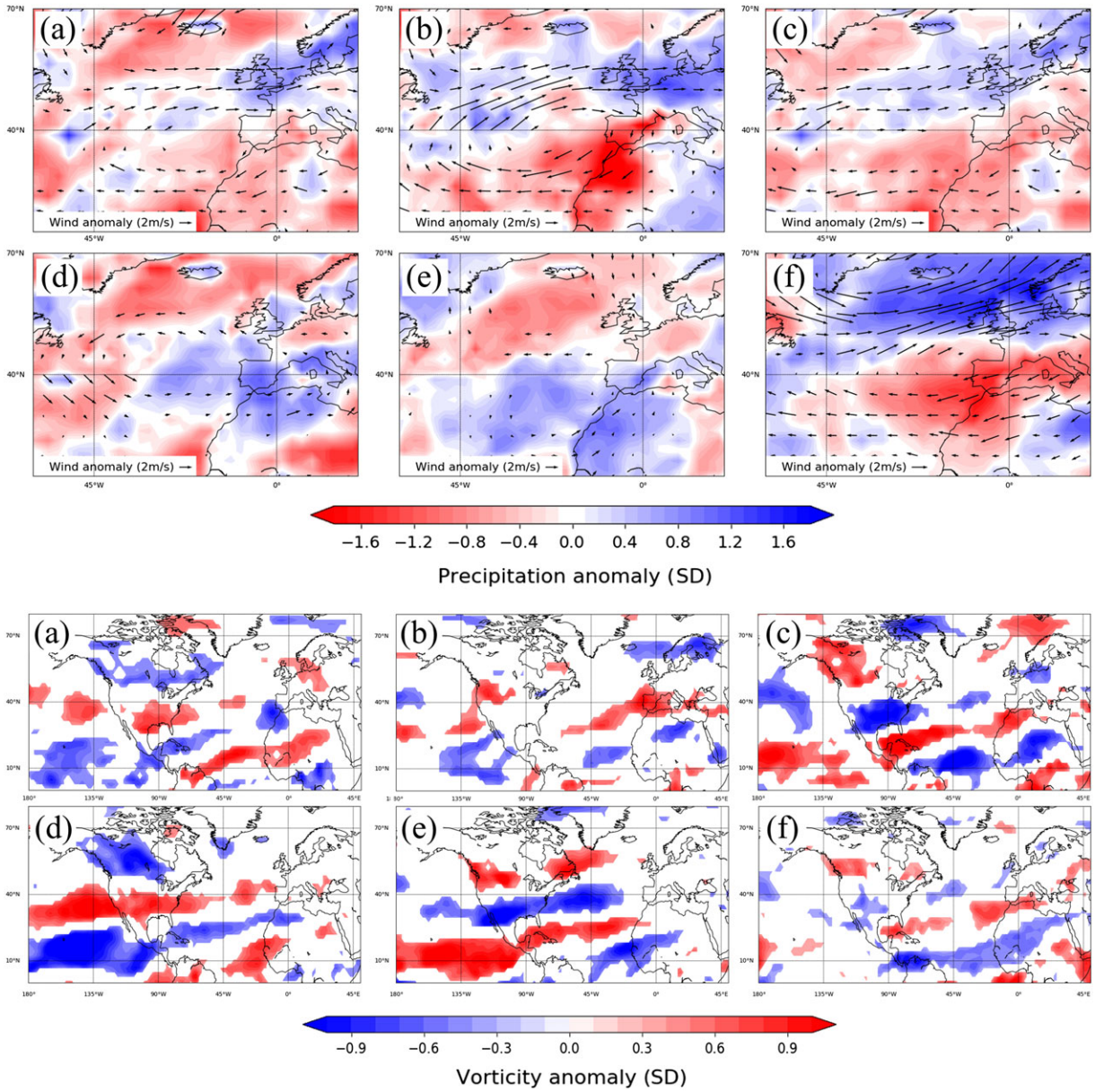

Figure 6. (a–f) Composites of 1983–2016 December-to-February PERSIANN-CDR precipitation (Sorooshian et al., 2014; filled contours) and NCEP-NCAR reanalysis 850-hPa winds, based on (a–e) the five sea surface temperature predictors defined in Figure 4 (middle row) and in the December-to-February NAO index (f). (g–l) Composites of 1950–2016 February-to-April 500-hPa vorticity from the NCEP-NCAR reanalysis, based on the six sea surface temperature predictors defined in Figure 4 (bottom row). Units for precipitation and vorticity are percent of grid point standard deviation; units for circulation are m/s. Only areas significant at the 90% level are shown, based on a bootstrap estimation of confidence intervals. SD = standard deviation.

field, by decomposing the latter into EOFs and performing an SVD decomposition on them, along with the precipitation series. This algorithm should select on its own the most relevant SST variables and patterns without any need to define by hand specific predictors. However, this approach turned out to be highly unstable and displayed little skill in practice. This may be due to our trying to predict a single, noisy time series and not a more stable multivariate field, as for instance in Totz et al. (2017), where the authors tried to predict European winter precipitation anomaly maps.

Considering a single time series (precipitation over the Oum Er Rbia) is indeed one of the limitations of our model. We know that, in winter, this series is strongly anticorrelated to precipitation over Scotland or western Scandinavia, which respond antisymmetrically to the same forcing by the NAO. Thus, looking for predictors that show skill in predicting both of these series at the same time (in a multivariate setting) may bring robust-

ness to the forecasts—by smoothing some of the noise in each of the series—and also help understand how differently, if at all, the northern pole of the NAO is affected by regional circulation. The continent-wide approach taken in Totz et al. (2017) may not yield very precise results for areas as small as ours, but it can help shed some light on the dominant NAO state to be expected for the winter, an additional piece of information to include in the forecasts.

Finally, although we rely here on linear-based forecast models, the relationship between precipitation and large-scale flow is known to be nonlinear (Pires & Perdigõ, 2007). For instance, in winter, with Morocco being very close to the southern pole of the NAO, as long as the Azores High is at least somewhat strong, storms will be deflected. When it becomes weak enough, then suddenly the door opens for storms to land in Morocco, resulting in an amplified precipitation response.

Including a GLM model in this study was motivated by the fact that it went beyond pure linearity, and a log-link fit did show some improvement for the NAO-precipitation relationship (with an Akaike information criterion slightly smaller than that of a linear regression model, $\Delta AIC = 3.7$). It must be made clear, however, that there is no physical basis for that choice, only statistical considerations. With such a short reliable record, and lacking advanced predictor selection methods, implementing nonparametric or nonlinear forecasting algorithms to represent physical connections better remains a challenge.

8. Conclusions

Forecasting seasonal precipitation from NWMs will remain quite challenging for the foreseeable future. However, this study shows that a statistical approach, based on a careful identification and selection of SST predictors, allows for a skillful prediction of the interannual variability of seasonal precipitation. Early-season precipitation is much less predictable, however, even with no lead time. The accuracy suggested by the cross-validation is still higher than what teleconnections signals predicted from the most up-to-date NWMs imply. Despite the short length of the data record, 35–40% of interannual variability can be captured by a small set of robustly selected SST indices. The identified predictors are consistent with results from previous studies, and the analysis of composite circulation fields linked to these further supports the validity of the physical connections to precipitation. At the end of the day, however, the accuracy of any forecast must be measured with respect to the potential, practical benefits it can bring to users. To paraphrase the saying, forecast skill is in the eyes of the beholder, in this case, Morocco's water and agricultural managers. Our seasonal forecasts are made to be used by them. Further work is needed to translate the forecasts to make them easily interpretable within the current framework of water management.

Acknowledgments
This work was made possible through funding from the Office Chérifien des Phosphates (OCP) through Université Mohamed VI Polytechnique. The authors would like to thank Mohamed El Gharous for his help in accessing agricultural data, as well as Judah Cohen for his advice regarding the Snow Advance Index. We are grateful to Nick Dunstone for sharing the dynamical NAO forecasts. The HadSLP2 and GPCC Precipitation data sets are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, available on their website at https://www.esrl.noaa.gov/psd/. The CRU TS4 precipitation data set is provided by the University of East Anglia's Climatic Research Unit and is made available from their website at https://crudata.uea.ac.uk/cru/data/hrg/. HadSST data are made available by the Met Office at www.metoffice.gov.uk/hadobs. The R code used for the analysis can be found at https://eltahir.mit.edu.

References

Allan, R., & Ansell, T. (2012). A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850-2004. *Journal of Climate*, *19*, 5816–5842. https://doi.org/10.1175/JCLI3937.1

Balaghi, R., Jlibene, M., Tychon, B., & Heerens, H. (2012). *Agrometeorological forecasting of cereal crop yields in Morocco*. Morocco: National Institute for Agronomic Research (INRA).

Bojariu, R., & Gimeno, L. (2003). Predictability and numerical modelling of the North Atlantic Oscillation. *Earth-Science Reviews*, *63*, 145–168. https://doi.org/10.1016/S0012-8252(03)00036-9

Born, K., Fink, A. H., & Paeth, H. (2008). Dry and wet periods in the northwestern Maghreb for present day and future climate conditions. *Meteorologische Zeitschrift*, *17*(5), 533–551. https://doi.org/10.1127/0941-2948/2008/0313

Cohen, J., & Jones, J. (2011). A new index for more accurate winter predictions. *Geophysical Research Letters*, *38*, L21701. https://doi.org/10.1029/2011GL049626

Czaja, A., & Frankignoul, C. (2002). Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation. *Journal of Climate*, *15*, 606–623. https://doi.org/10.1175/1520-0442(2002)015<0606:OIOASA>2.0.CO;2

Czaja, A., Robertson, A. W., & Huck, T. (2013). The role of Atlantic Ocean-Atmosphere coupling in affecting North Atlantic Oscillation variability. In W. Hurrell, Y. Kushnir, G. Ottersen, & M. Visbeck (Eds.), *The North Atlantic Oscillation: Climatic significance and environmental impact*. Washington, DC: American Geophysical Union. https://doi.org/10.1029/134GM07

Dobson, A. J. (1990). *An introduction to generalized linear models*. London: Chapman and Hall. https://doi.org/10.1007/978-1-4899-7252-1

Dunstone, N., Smith, D., Scaife, A. A., Hermanson, L., Eade, R., Robinson, N., et al. (2016). Skilful predictions of the winter North Atlantic Oscillation one year ahead. *Nature Geoscience*, *9*, 809–814. https://doi.org/10.1038/ngeo2824

El Mourid, M., & Karroui, M. (1996). Agriculture in arid and semi-arid regions of Morocco: Challenges and prospects. *Al Awamia*, *92*, 83–91. Retrieved from http://www.inra.org.ma/docs/awamia/article/09207.pdf

Feldstein, S. B. (2000). The timescale, power spectra, and climate noise properties of teleconnection patterns. *Journal of Climate*, *13*, 4430–4440. https://doi.org/10.1175/1520-0442(2000)013<4430:TTPSAC>2.0.CO;2

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22.

Gastineau, G., L'Heveder, B., Codron, F., & Frankignoul, C. (2016). Mechanisms determining the winter atmospheric response to the Atlantic overturning circulation. *Journal of Climate*, *29*, 3767–3785. https://doi.org/10.1175/JCLI-D-15-0326.1

Hoerling, M. P., Hurrell, J. W., & Xu, T. (2001). Tropical origins for recent North Atlantic climate change. *Science*, *292*(5514), 90–92. https://doi.org/10.1126/science.1058582

Hurrell, J. W. (1995). Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, *269*(5224), 676–679. https://doi.org/10.1126/science.269.5224.676

Hurrell, J. W., Kushnir, Y., Ottersen, G., & Visbeck, M. (2013). An overview of the North Atlantic Oscillation. In J. W. Hurrell, Y. Kushnir, G. Ottersen, & M. Visbeck (Eds.), *The North Atlantic Oscillation: Climatic significance and environmental impact*. https://doi.org/10.1029/134GM01

Kennedy, J. J. (2014). A review of uncertainty in in situ measurements and data sets of sea surface temperature. *Reviews of Geophysics*, *52*, 1–32. https://doi.org/10.1002/2013RG000434

Kharin, V. V., & Zwiers, F. W. (2003). On the ROC score of probability forecasts. *Journal of Climate*, *16*, 4145–4150. https://doi.org/10.1175/1520-0442(2003)016<4145:OTRSOP>2.0.CO;2

Knippertz, P., Christoph, M., & Speth, P. (2003). Long-term precipitation variability in Morocco and the link to the large-scale circulation in recent and future climates. *Meteorology and Atmospheric Physics*, *83*, 67–88. https://doi.org/10.1007/s00703-002-0561-y

Mariotti, A., Zeng, N., & Lau, K.-M. (2002). Euro-Mediterranean rainfall and ENSO? A seasonally varying relationship. *Geophysical Research Letters*, *29*(12), 1621. https://doi.org/10.1029/2001GL014248

Michaelsen, J. (1987). Cross-validation in statistical climate models. *Journal of Climate and Applied Meteorology*, *26*, 1589–1600. https://doi.org/10.1175/1520-0450(1987)026<1589:CVISCF>2.0.CO;2

Peng, S., Robinson, W. A., & Li, S. (2003). Mechanisms for the NAO responses to the North Atlantic SST tripole. *Journal of Climate*, *16*, 1987–2004. https://doi.org/10.1175/1520-0442(2003)016<1987:MFTNRT>2.0.CO;2

Pires, C. A., & Perdigõ, R. A. (2007). Non-Gaussianity and asymmetry of the winter monthly precipitation estimation from the NAO. *Monthly Weather Review*, *135*, 430–448. https://doi.org/10.1175/MWR3407.1

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*, D144407. https://doi.org/10.1029/2002JD002670

Robertson, A. W., Mechoso, C. R., & Kim, Y.-J. (1999). Predictability and numerical modelling of the North Atlantic Oscillation. *Journal of Climate*, *13*, 122–138. https://doi.org/110.1175/1520-0442(2000)013<0122:TIOASS>2.0.CO;2

Rodwell, M. J., Rowell, D. P., & Folland, C. K. (1999). Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, *398*, 320–323. https://doi.org/10.1038/18648

Saunders, M. A., & Qian, B. (2002). Seasonal predictability of the winter NAO from North Atlantic sea surface temperatures. *Geophysical Research Letters*, *29*(22), 2049. https://doi.org/10.1029/2002GL014952

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., & Ziese, M. (2011). GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. https://doi.org/10.5676/DWD_GPCC/FD_M_V7_050

Shaman, J., & Tziperman, E. (2010). An atmospheric teleconnection linking ENSO and Southwestern European precipitation. *Journal of Climate*, *24*, 124–139. https://doi.org/10.1175/2010JCLI3590.1

Smith, D. M., Scaife, A. A., Eade, R., & Knight, J. R. (2016). Seasonal to decadal prediction of the winter North Atlantic Oscillation: Emerging capability and future prospects. *Quarterly Journal of the Royal Meteorological Society*, *142*, 611–617. https://doi.org/10.1002/qj.2479

Sorooshian, S., Hsu, K., Braithwaite, D. K., Ashouri, H., & NOAA CDR Program (2014). NOAA Climate Data Record (CDR) of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN-CDR), version 1 revision 1., NOAA National Centers for Environmental Information, accessed November 15, 2017. https://doi.org/10.7289/V51V5BWQ

Sutton, R. T., & Hodson, D. L. R. (2003). Influence of the Ocean on North Atlantic climate variability 1871-1999. *Journal of Climate*, *16*, 3296–3313. https://doi.org/10.1175/1520-0442(2003)016<3296:IOTOON>2.0.CO;2

Totz, S., Tziperman, E., Coumou, D., Pfeiffer, K., & Cohen, J. (2017). Winter precipitation forecast in the European and Mediterranean regions using cluster analysis. *Geophysical Research Letters*, *44*, 12,418–12,426. https://doi.org/10.1002/2017GL075674

Trenberth, K. E., & Fasullo, J. T. (2013). An apparent hiatus in global warming? *Earth's Future*, *1*, 19–32. https://doi.org/10.1002/2013EF000165

University of East Anglia Climatic Research Unit, Harris, I. C., & Jones, P. D. (2017). CRU TS4.00: Climatic Research Unit (CRU) Time-Series (TS) version 4.00 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2015), Centre for Environmental Data Analysis, 25 August 2017. https://doi.org/10.5285/edf8febfdaad48abb2cbaf7d7e846a86

Wilks, D. S. (2005). *Statistical methods in the atmospheric sciences* (Vol. 2, 704 pp.). Amsterdam, Boston: Academic Press.