Centro Euro-Mediterraneo sui Cambiamenti Climatici

Research Papers
Issue RP0269
April 2016

Fondazione CMCC CSP - Climate Simulation and Prediction Division

A New Approach to Extract Monthly High Resolution Information for Precipitation from GCM Scenarios and Predictions

*By **M. Neil Ward***
mneilward@gmail.com

Stefano Materia
Fondazione CMCC – Centro Euro-Mediterraneo sui Cambiamenti Climatici
stefano.materia@cmcc.it

Silvio Gualdi
silvio.gualdi@cmcc.it

Antonio Navarra
antonio.navarra@cmcc.it

The research leading to these results has received funding from the Italian Ministry of Education, University and Research and the Italian Ministry of Environment, Land and Sea under the GEMINA project.

SUMMARY

Statistical downscaling to 0.25° spatial resolution TRMM monthly precipitation explores the potential and limitations of using the relatively short (here 15 years) but spatially extensive and complete precipitation archive, that is satellite-based and merged with station information. Downscaling models relate reanalysis circulation to TRMM, utilizing principal component regression (PCR) and canonical correlation analysis (CCA). Results are demonstrated for two contrasting regions: Northeastern Brazil (NEB, tropical, distinct wet and dry season) and Central Italy (CIT, mid-latitude/Mediterranean, complex terrain). Models are constructed for individual months (M1) and by pooling three months (M3) to increase model-training sample size when downscaling to the central month of the three. Cross-validated skill with M1 is promising and is noticeably more consistent and higher with M3, e.g., mean skill at the grid-box scale rises from r=0.44 to r=0.59 for CIT (averaged over all months), and from r=0.58 to r=0.68 for the NEB wet season months.

Spatial structure of the downscaling models (as revealed by CCA modes) supports a clear expression of orography in the precipitation anomaly fields. Application to a global coupled model climate change scenario (2012-2050) generates plausible downscaled time-series and fields. For CIT, results (skill, spatial structure) are consistent with those produced using a station-only gridded (0.25°) dataset for the extended period 1979-2012. The overall impression gained is that TRMM data enable estimation of skilful downscaling relationships, at least for some locations. Developments drawing on longer datasets to adjust the downscaled fields will likely further increase the utility of a record like TRMM.

Keywords rainfall, regression model, downscaling, correlation, principal component, orography, wind.

1. INTRODUCTION

The goal of extracting higher resolution climate information from General Circulation Model (GCM) output has been widely pursued, motivated by the potentially enhanced utility of such information. Two families of approaches have emerged: statistical downscaling and dynamical downscaling. Each has been shown to have strengths and weaknesses, which can render them both having roles to play (Hewitson and Crane 1996; Wilby and Wigley 1997; Murphy 1999; von Storch 2005; Fowler et al. 2007; Maraun et al. 2010). One constraint to the generic application of statistical downscaling has been the lack of high resolution climate datasets with complete geographic coverage and a sufficient temporal record for construction of downscaling relationships. This was noted in Wilby and Dawson (2013), who explicitly mention the emerging potential of satellite-based products (including those of the Tropical Rainfall Measuring Mission (TRMM), Huffman et al. 2007) to address this constraint.

Using the contrasting case study regions of Central Italy (CIT) and Northeastern Brazil (NEB), the objective of this paper is to make an initial assessment of the potential of using a spatially complete 0.25° resolution (~25km) monthly precipitation TRMM product to construct statistical downscaling models. In other statistical downscaling studies of monthly precipitation, regression-based models (as categorized in Fowler et al. 2007) have been widely applied (Wigley et al. 1990; von Storch et al. 1993; Zorita and von Storch 1999; Widmann et al. 2003; Vimont et

al. 2010; Nicholas and Battisti 2012). Therefore, regression approaches are considered sufficient for the initial assessment here. Perfect prognosis downscaling models are constructed by relating monthly reanalysis fields to the TRMM precipitation data. The regression-based methods of principal component regression (PCR) and canonical correlation analysis (CCA) have been explored. For this initial study, the predictor reanalysis fields are confined to circulation-based ones, which have been shown to be quite robust for statistical downscaling, given these features tend to show good consistency across reanalysis and GCMs (Fowler et al. 2007; Hewitson and Crane 2006). Their application is considered sufficient for the feasibility assessment purpose here.

One challenge in the application of regression-based methods to precipitation across all months is the evolution of downscaling relationships with the annual cycle. Since the goal is to assess feasibility of a method that can be applied as universally as possible, the approach includes from the outset the desire to allow models to evolve across the annual cycle. In other studies, this issue has been deemed secondary for initial analysis of the locations under study (e.g., Nicholas and Battisti 2012), has been treated through an approximate separation into two contrasting parts of the year (Timm and Diaz 2009) or through direct establishment of monthly-specific relationships (e.g., Sachindra et al. 2014). The approach taken here begins with monthly-specific models and then explores a generalization of the Timm and Diaz (2009) approach by pooling overlapping three-month sets together to increase sample size for models that are then used to downscale precipitation to

the central month of each three month set. The importance of allowing relationships to vary with the annual cycle is clear for the two case study regions where circulation-precipitation relationships are shown to be distinct at different times of the year.

Conclusions about the potential of the methodology introduced here will be supported by four types of information. First, the cross-validation (CV) skill of the perfect prognosis models for each month will be assessed. Clearly, this is a first prerequisite: if models are to be potentially applicable, they should be able to skilfully specify precipitation in the period of historical TRMM data (here 1998-2012). Second, the perfect prognosis models are applied to a GCM scenario for 2021-2050 and the plausibility of the generated high resolution scenarios is discussed. One of the challenges for the downscaling is the relatively short record of TRMM data (here, 15 years). While this situation is expected to incrementally improve as the record becomes longer, nonetheless, it is important to gain as many insights as possible on the expected performance now when the record is approximately 15-years in length. The third and fourth pieces of information are designed to partly address this issue. The third area considers skill enhancements when three months are pooled together to construct models. This recognizes that circulation-precipitation relationships in adjacent months will usually be quite similar. Constructing models using 45 cases (the 3 pooled months times 15 years) instead of 15 cases can be expected to lead to more robust models and improved CV skill scores, and to lead to less variability in skill levels month-to-month. The

success of such an analysis can be considered to confirm that skilful relationships are emerging through application of the TRMM data, and indeed, may provide strong evidence for potential to mine the most skilful monthly downscaling models by pooling months together.

The fourth area of evidence reported takes advantage of the existence of a gridded dataset at the same resolution as TRMM, but constructed using the high density station network that exists in the broad European and nearby region (E-OBS, Haylock et al. 2008). These data cover a longer period, here analyzed for 1979-2012. This permits the comparison of downscaling results with TRMM for the CIT case study region. The similarity of month-to-month CV skill and spatial relationship (as seen through CCA) is addressed using the full period 1979-2012 and the shorter period 1998-2012. The paper's philosophy is to consider the extent to which the shortness of the TRMM record may be a limitation that makes downscaling less exact than could be achieved using longer records. However, if approximate but still skilful relationships can be extracted, the limitation can be counter-balanced by the advantage of the data being high resolution and spatially complete, thereby at least permitting plausible scenarios to be generated, with appropriate caveats. Such scenarios may, at the least, be useful for sensitivity testing and adaptation option appraisals (Wilby and Dawson 2013).

Section 2 presents the data and methods applied, and includes discussion of the basic wind-precipitation coupling in the target regions. Section 3 validates the perfect prognosis downscaling models, including ones constructed using the 3-

month pooling approach and ones using E-OBS. It also considers dynamic feasibility of some of the spatial characteristics of the CCA modes, including comparison of modes achieved using TRMM and E-OBS for CIT. Section 4 diagnoses the downscaled precipitation generated by applying the fitted perfect prognosis models to a global change scenario for 2021-2050. Section 5 contains conclusions and discussion of some further analyses that are prompted by the results.

2. DATA, METHODS AND METHODOLOGY

a. Predictand Precipitation

The primary predictand precipitation data explored in this study are from the Version 7 TRMM Multi-Satellite Precipitation Analysis, dataset 3B43, a monthly dataset at 0.25° spatial resolution. It is a merge of the multi-satellite analysis with a single consistent station analysis from GPCC (Huffman et al. 2007; Huffman and Bolvin 2013). The data extend from 50°N to 50°S and are continuous since January 1998. The data therefore represent a source that is potentially applicable to downscaling across a large domain of the earth's surface, and represent a record that is continually expanding forward in time permitting downscaling models to be updated and revised as time progresses. For this paper, data are analyzed for January 1998 to December 2012, and for brevity, these data hereon are referred to as TRMM.

For the case study region of CIT, downscaling results achieved with TRMM are compared to those achieved using a station-only precipitation dataset (E-OBS, Haylock et al. 2008) that is available for the European domain at the same spatial grid resolution, yet covering an extended period (here analyzed for 1979 to 2012). The E-OBS data are served as daily values, and therefore for analysis here, needed to be processed into monthly values. The monthly value was calculated by averaging the daily mm totals (yielding the month's precipitation in mm/day). If a grid-box contained less than 85% of daily values in a given month, that month was set to missing. If more than one month was missing in a given analysis, then that grid-box was excluded from the analysis. Results were not sensitive to small adjustments to these thresholds.

b. Predictor Reanalysis Data

To form the predictor data for the perfect prognosis downscaling models, the European Centre for Medium-Range Weather Forecasting (ECMWF) ERA-Interim reanalysis (Dee et al. 2011) monthly means of daily means were processed at 1.5° spatial resolution. Circulation and geopotential height at low-levels and mid/upper-levels were explored. In the context of downscaling climate change scenarios, the preference was to focus on circulation, since geopotential heights can be expected to be especially vulnerable to systematic impacts from general warming. It turns out that for the case study regions explored here (CIT and NEB), geopotential height adds no substantial perfect prognosis skill to that achieved by circulation. Further, mid/upper-level circulation predictors were found to add very little to the skill

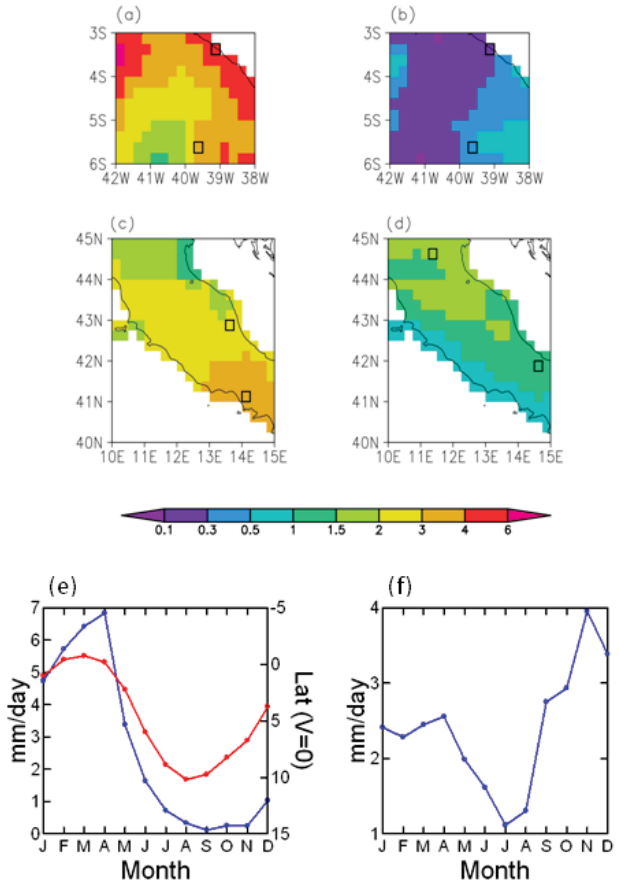
achieved with low-level circulation alone. Both 850mb and 925mb circulation prove effective for both regions. Slightly more consistent and higher skill results are found with 925mb wind for NEB, while for CIT, 850mb wind yields slightly better results, and is favored in principle since the predictor domain (see further discussion below) includes the mountainous Apennine region, such that the reanalysis 925mb level may in places represent less-consistent interpolated values.

c. Predictand Regions and Basic Circulation Linkages

Domains across NEB and CIT were selected for the case study downscaling evaluations, providing contrasting climate settings (Fig. 1). NEB is tropical with a sharp wet season (approximately January to May, through proximity to the Inter-Tropical Convergence Zone, ITCZ) and strong dry season in the remainder of the year (Fig. 1e). The mean CIT annual cycle (Fig. 1f) displays an early winter precipitation maximum, related to mid-latitude storms, and Mediterranean summer minimum (though not as precipitation deficient as the NEB dry season). The TRMM grid-boxes included in the downscaling analysis (Fig. 1) were selected after consulting NASA mask information, treating as ocean every box that is designated as greater than 90% covered with water. The ocean areas were masked because the primary concern is with downscaling to land areas; inclusion of ocean areas in the model fitting process would be sub-optimal by allowing ocean areas to influence the choice of best model. Both domains contain approximately 200 TRMM grid-boxes. The intention was to keep the target domain size modest for clarity of visualizing spatial results in this illustration, and to have a similar number of grid-

boxes in each domain. General climatic homogeneity of the target region is a consideration in domain selection (Fowler et al. 2007), and these two domains both show good first-order homogeneity, with the leading precipitation principal component (PC) in each month having the same sign of weights across the whole domain, and explaining a large fraction of total precipitation variance (for NEB, typically 80% for the wet season and 65% for the remainder of the year; for CIT, typically 65% for the winter semester and 60% for the summer semester). Therefore, to see the basic scales and nature of precipitation-wind association, wind correlations with precipitation PC1 (examples in Fig. 2) provide an introduction to the relationships that could be drawn upon for downscaling purposes.

FIG. 1. Predictand domains and precipitation climatology summary, using grid and data from TRMM (1998-2012). Illustrative monthly precipitation climatology for (a) NEB in May, (b) NEB in August, (c) CIT in January, (d) CIT in August. Units are mm/day. Also shown is the mean precipitation annual cycle for (e) NEB domain (blue line) and (f) CIT domain (blue line). In (e) the red line shows the latitude of the mean low-level meridional wind confluence averaged over the nearby tropical Atlantic (30-35°W, here shown for 925mb). Note y-axis for red line is reversed, so rainy season is accompanied by the southernmost excursion of wind confluence. The small rectangles in (a-d) locate the case study grid-boxes used in Fig. 9.



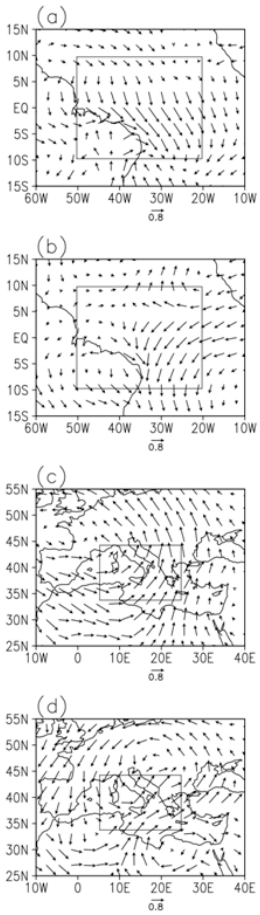
For NEB (Fig. 2a-b), there is a marked contrast in regional 925mb circulation anomalies associated with precipitation anomalies, depending on location in the annual cycle. Variability during the wet season is strongly tied to nearby equatorial Atlantic fluctuations in the trade wind – ITCZ complex (widely studied, e.g., Hastenrath and Heller 1977). This is illustrated for May (Fig. 2a) with anomalously wet conditions accompanied by anomalous northwesterly flow focused around 3°N – 6°S, indicative of a southward displaced trade wind – ITCZ complex. Broadly similar associations are seen for each month January-May. During boreal summer, even though mean precipitation is typically <0.5 mm/day, the year-to-year

fluctuations of precipitation still contain strong expression in monthly mean circulation, but markedly distinct from those in the wet season. Now, anomalously wet conditions are accompanied by (i) anomalous northerly winds in the western tropical Atlantic from 3°N southward to beyond 10°S and (ii) these northerly winds link up with reduced monsoon flow into West Africa (illustrated for August, Fig. 2b; similar results for July). It is desirable for downscaling models to reflect these evolving precipitation-circulation linkages, and this is implemented in section 3 through fitting models for each month (or set of overlapping three months). After boreal summer (especially October-November), linkages of circulation to precipitation are much weaker (not shown). This appears to represent a different phase of the dry season when year-to-year variations in precipitation have much less monthly circulation expression, and will be problematic for the downscaling approach here.

For CIT, correlation patterns between precipitation PC1 and 850mb winds contain a strong signal in all months, but again with a clear evolution in structure through the annual cycle. This is illustrated in Fig. 2c-d, with larger-scale and somewhat stronger linkages in winter compared to summer. Generally, enhanced precipitation is associated with a stronger push of cyclonic southerly anomalies over Italy (and extending eastward) during winter. In summer, enhanced precipitation is associated with a cyclonic circulation anomaly centered over the north Tyrrhenian Sea and another wider one centered near Hungary, the two forming a “glass-shaped” low pressure. As an example of the resulting circulation differences, circulation

anomalies in Figs. 2c-d over Sardinia display a southerly component in winter, and a northerly component in summer.

FIG. 2. Correlation between predictand precipitation PC1 and low-level reanalysis wind for selected months. (a) May for 925mb wind with NEB, (b) same as (a) but for August, (c) January for 850mb wind with CIT, (d) same as (c) but for August. Precipitation PC1 always displays dominantly positive weights across the precipitation domain. Correlation with u-wind and v-wind is plotted as a vector, effectively forming a standardized wind anomaly, with sign such that these wind fields correspond to positive precipitation anomaly. The predictor domains used in the PCR and CCA downscaling models are shown by the solid rectangles.



d. Downscaling Models

The methods of PCR and CCA have been widely applied in climate analysis generally (Barnett and Preisendorfer 1987; Barnston and Ropelewski 1992; von Storch and Zwiers 1999; Feddersen et al. 1999) and explicitly as regression-based methods for statistical downscaling (von Storch et al. 1993; Vimont et al. 2010; Nicholas and Battisti 2012). Details of their implementation for results in this paper are provided in Appendix A, while a summary of essential aspects for interpretation of results is included in this section.

If a model draws on the nth PC, then it is required to contain information from all PCs from PC1 to PCn (i.e., for PCR, n predictors). Similar constraints were applied for CCA models. The maximum number of PCs and CCA modes were held constant for all downscaling experiments in this paper (as described in Appendix A). For a given predictor domain and variable set, the actual choices for the best number of PCs and CCA modes (between 1 and the maximum permitted) were made based on applying a leave-one-out CV procedure, where the downscaled prediction for year i is made by a model that is fitted excluding all information about year i, including recalculation of all statistics in the predictor and predictand data (mean, standard deviation) as well as the PCs, CCA modes and prediction coefficients that are needed to specify the model. The best choice of PCs (and modes in the case of CCA) is made based on the mean (area average) of the CV correlation coefficient for each predictand grid-box, i.e., the average of the correlation skill at each grid-box in the domain. It is recognized that other skill

scores may more closely match potential utility in any given application, but it is generally the case that if good correlation skill is achieved, the potential exists to extract more tailored information, such as through variance adjustment (von Storch 1999) or other transformations / statistical model choice. Such assessments are beyond the scope of this paper, and use of correlation skill is considered sufficient for the purpose here.

The domain over which to calculate predictor PCs represents a further choice. Based on inspection of the correlation scales (Fig. 2) and through a degree of trial and error, the domains shown on Fig. 2 are used throughout the results in this paper. Results were not sensitive to minor changes in the domain selection. The choice of shifting the CIT predictor domain southward to asymmetrically straddle the target precipitation region reflects the preference to avoid the high altitudes of the Alps, where 850mb winds may be treated differently in reanalysis and the GCM. As discussed in section 2b, the preferred low-level circulation is at 850mb for CIT and 925mb for NEB. When using both zonal wind (u-wind) and meridional wind (v-wind) as predictors in combination, the u-wind and v-wind fields are stacked, such that the number of variables entering the PC analysis is doubled. For simplicity, all results presented adopt this approach (u-wind and v-wind in combination), since it provided the generally most consistent set of results.

Natural decadal to multidecadal (DMD) climate variability can be a challenge to statistical downscaling in a number of ways (further discussed in Appendix B). One major concern is the sensitivity of results to possible changes in the relationships

between predictor and predictand. The comparison of results with E-OBS provides an initial indication of whether the impact of DMD variability is overwhelming in the CIT case study example, at least across the period 1979-2012. A further major concern is that low-frequency variability can lead to strong year-to-year persistence in precipitation series which can substantially reduce the degrees of freedom in model estimation (e.g., Estrada et al. 2013) and could substantially compromise the leave-one-out CV procedure. For the demonstration domains here, lag-one serial correlation from one year to the next, is not systematically positive. Such lag-one correlation values were calculated for each grid-box in each month; averaged over all grid-boxes and months, r=-0.09 for NEB and r=-0.11 for CIT, with values randomly moving from slightly positive to slightly negative across the annual cycle of months for each location. Based on this, it appears reasonable to assume full degrees of freedom when analyzing the 15-year series of monthly precipitation values in the domains here.

e. Adjacent Three-month Pooling

A methodological innovation introduced in this paper is the construction of regression models using pooled 3-months of data, to be applied for downscaling the central month (hereafter referred to as method M3). This potentially offers a route to substantially increase the degrees of freedom for fitting models, compared to the approach in section 2d, which used a single month alone (hereafter referred to as method M1). The potential gains could be compromised if there is strong month-to-month persistence of precipitation anomalies, such that there is only limited

increase in the effective sample size. However, monthly precipitation anomalies in many locations have low persistence, and this is generally confirmed for the two case study regions here (Table 1). Across all boxes, the mean persistence correlation from one month to the next is 0.19 for NEB and 0.15 for CIT. The effective sample size for a three month set is estimated (following Yevjevich 1972; Katz and Glantz 1986):

$$n' = n \Big/ 1 + \bar{r}(n-1) \qquad (1)$$

where $\bar{r}$ is the mean correlation between the constituent three monthly time-series (i.e., average of the three correlations, see row "Mean of 3" in Table 1), n is the number of series (three) and n' is the effective number of independent series. Using the average persistence correlations, this yields values for n' around 2.3 for NEB and 2.5 for CIT (Table 1), so assuming (as discussed in section 2d) that year-to-year persistence is effectively zero, the total number of independent cases can be estimated as typically 2.3*15=34.5 for NEB and 2.5*15=37.5 for CIT. The average NEB values (over all months) match closely the values averaged across just the wet season 3-month combinations straddling January-May (Table 1). Persistence does increase as the wet season proceeds, such that persistence from April to May is highest (r=0.56). The effective sample size for March-May is noticeably the most compromised, reducing from 3 to 1.6. However, gains in skill for the M3 approach (section 3c) do not show dependence on the degree of 3-month inter-month correlation in Table 1. Therefore, it seems reasonable to conclude that, at least for

these case study regions, persistence in precipitation is not greatly compromising degrees of freedom in the 3-month sets.

TABLE 1. Persistence precipitation correlation (*100) across each 3-month set and the implied effective number of independent monthly series. Values shown are averaged over all grid-boxes for the given region (NEB or CIT). For example, for month 1 (January), correlation of (m-1) v (m) is December versus January, (m-1) v (m+1) is December versus February, Mean of 3 is the average correlation: December vs. January, January vs. February and December vs. February. Effective N is the effective number of independent series in the 3-month set, estimated using Eq. (1).

		MONTH												
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**AV**
NEB	(m-1) v (m)	20	-7	-14	29	56	18	23	39	30	-8	26	18	19
	(m-1) v (m+1)	12	10	10	43	18	33	1	25	-2	0	12	-5	13
	Mean of 3	8	-4	8	43	31	25	21	31	7	6	19	11	17
	Effective N	2.6	3.0	2.6	1.6	1.9	2.0	2.1	1.8	2.6	2.7	2.2	2.5	2.3
CIT	(m-1) v (m)	33	11	0	-8	21	-10	8	19	43	-6	31	37	15
	(m-1) v (m+1)	2	25	6	-18	-6	-8	-10	2	11	-6	29	12	3
	Mean of 3	15	12	-1	-2	2	-3	6	21	16	6	32	27	11
	Effective N	2.3	2.4	3.0	3.1	2.9	3.0	2.7	2.1	2.3	2.7	1.8	1.9	2.5

Before pooling (both precipitation and circulation), data for month n are standardized according to the mean and standard deviation of month n over the 15 years. This stops a model from fitting to any evolution of mean background state over the three months, focusing only on the departure from normal from the appropriate point in the annual cycle. The three sets of 15-year sequences are then stacked vertically to form a single model fitting dataset that contains 45 values for each precipitation and circulation grid-box. The model-fitting and CV procedure is

then implemented exactly as before: when making a perfect prognosis prediction for case i in the sequence of the 45 cases, all data for case i are withheld in the model fitting process. The CV predictions are available for all 45 cases, but here, just the central cases 16 to 30 are extracted, corresponding to the predictions for the target (central) month.

f. Application of Downscaling Models to a GCM Scenario

To generate illustrative downscaled fields, a present day (baseline) GCM experiment (1998-2012) and a GCM scenario experiment (here analyzed for 2021-2050) are subjected to the downscaling procedure. The selected PCR downscaling models (see section 2d) are applied to the monthly GCM fields. The GCM data are from archived runs of the global coupled GCM developed at the Centro Euro-Mediterraneo sui Cambiamenti Climatici (Scoccimarro et al. 2011; Bellucci et al. 2013). The present day experiment (1998-2012) has been conducted using observational radiative forcing (greenhouse gases, aerosols, solar irradiance and sulfates) from the Coupled Model Intercomparison Project phase 5 (CMIP5, Meehl et al. 2009) up to 2005, and a RCP4.5 scenario (van Vuuren et al. 2011) for the additional seven years. The same experiment (following the RCP4.5 pathway) provided the 2021-2050 scenario data.

To calculate the PC predictor time-series for the model runs, the GCM fields (for the month being analyzed, e.g., all Januaries) must first be interpolated onto the same grid as used to calculate the predictor PCs (i.e., ECMWF reanalysis at 1.5°

resolution). Then the projection of the monthly GCM fields onto the reanalysis PC spatial patterns is calculated:

$$a(m)_{jt} = \sum_{p=1..q} x'(m)_{pt} e_{pj} \qquad (2)$$

where the $a(m)_j$ form the GCM (baseline or scenario) PC time-series that will be applied to the (previously fitted) PCR equations to give the downscaled precipitation values, e_{pj} is the previously calculated (Appendix A) reanalysis PC spatial weight for grid-box p in PC j, summation is over the q grid-boxes that were included in the PC analysis, and $x'(m)_{pt}$ is, for grid-box p and month t, the scenario anomaly expressed relative to the GCM's baseline mean and standard deviation, i.e.,:

$$x'(m)_{pt} = \left(x(m)_{pt} - \overline{x(b)_p}\right) \Big/ \sigma(b)_p \qquad (3)$$

where $\overline{x(b)_p}$ and $\sigma(b)_p$ are the GCM's mean and standard deviation for grid-box p during the baseline simulation for 1998-2012. The PC time-series for the model runs (Eq. (2)) are then inserted into the downscaling equations (see Appendix A, Eq. (A1)) which were fit over the model training period using the PCR model selected by the CV procedure (the training period is always 1998-2012 for TRMM-based models).

Since the GCM scenario anomalies are calculated relative to the GCM baseline climatology (Eq. (3)), any trends in GCM circulation (scenario relative to baseline) can be expressed in the scenario PC time-series (Eq. (2)) and can therefore find expression in the downscaled model precipitation (Eq. (A1)). More sophisticated treatment of the reference level is beyond the scope of this paper. Consideration of possible adjustments based on assumptions about DMD variability is discussed in Appendix B. Also, even in the absence of DMD variability, 15 years is shorter than typically used for estimation of climatological means. To address this, in future formulations with TRMM, the downscaled scenarios may be systematically rescaled, drawing on a coarser resolution but longer time-period climatology to provide fields of adjustments based on differences in the TRMM versus longer-period climatology.

3. SKILL AND CONSISTENCY OF THE PERFECT PROGNOSIS MODELS

a. Monthly Models (M1 Approach)

For the relatively small and spatially coherent target domains tested here, the average CV correlation skill of PCR and CCA models is very similar (typically to within r=0.02), so downscaled results from the simpler technique of PCR are generally presented (while CCA modes are drawn upon in sections 3b and 3d to succinctly summarize spatial structure in the downscaling relationships). For NEB

(Table 2), the wet season skill is encouraging (averaged over the 5 months January-May, r=0.58). Substantial skill levels are also found for the dry season months of July and August (average r=0.51) when the teleconnection results (Fig. 2a-b) illustrated the strong switch in circulation-precipitation relationships that occurs after the wet season. The transition month for the switch (June) has very weak skill, while the rest of the dry season (September-December) has no skill taking the approach here. For the models here with clear skill (January-May and July-August), typically 4-6 PCs are selected as yielding maximum skill.

TABLE 2. Performance of PCR in predicting NEB using 10°N-10°S, 20-50°W, u-wind and v-wind at 925mb. N PC is the number of PCs giving the maximum cross-validation skill and R Skill is the average cross-validation correlation skill (*100) in each month (i.e., the average of the skill for each 0.25° box). Two additional summaries are provided: AV1 is mean skill across January-May and AV2 is the mean skill for July and August. The M1 approach fits models to a single month. The M3 approach fits models to overlapping three months, and verifies for the single month shown.

	Month	1	2	3	4	5	6	7	8	9	10	11	12	AV1	AV2
M1 Approach	N PC	2	6	4	1	5	6	6	5	6	2	6	4		
	R Skill	44	56	45	72	72	13	46	57	8	0	-8	2	58	51
M3 Approach	N PC	2	4	6	5	6	5	4	4	6	6	6	4		
	R Skill	56	61	65	79	79	34	63	63	8	16	28	32	68	63

For CIT (Table 3), all months except June have a mean correlation skill >0.3. There is generally a tendency for better skill in the winter semester, consistent with other studies of monthly downscaling in Europe. Mean levels of skill, while clearly positive and representative of at least some downscaling ability, are generally a little less

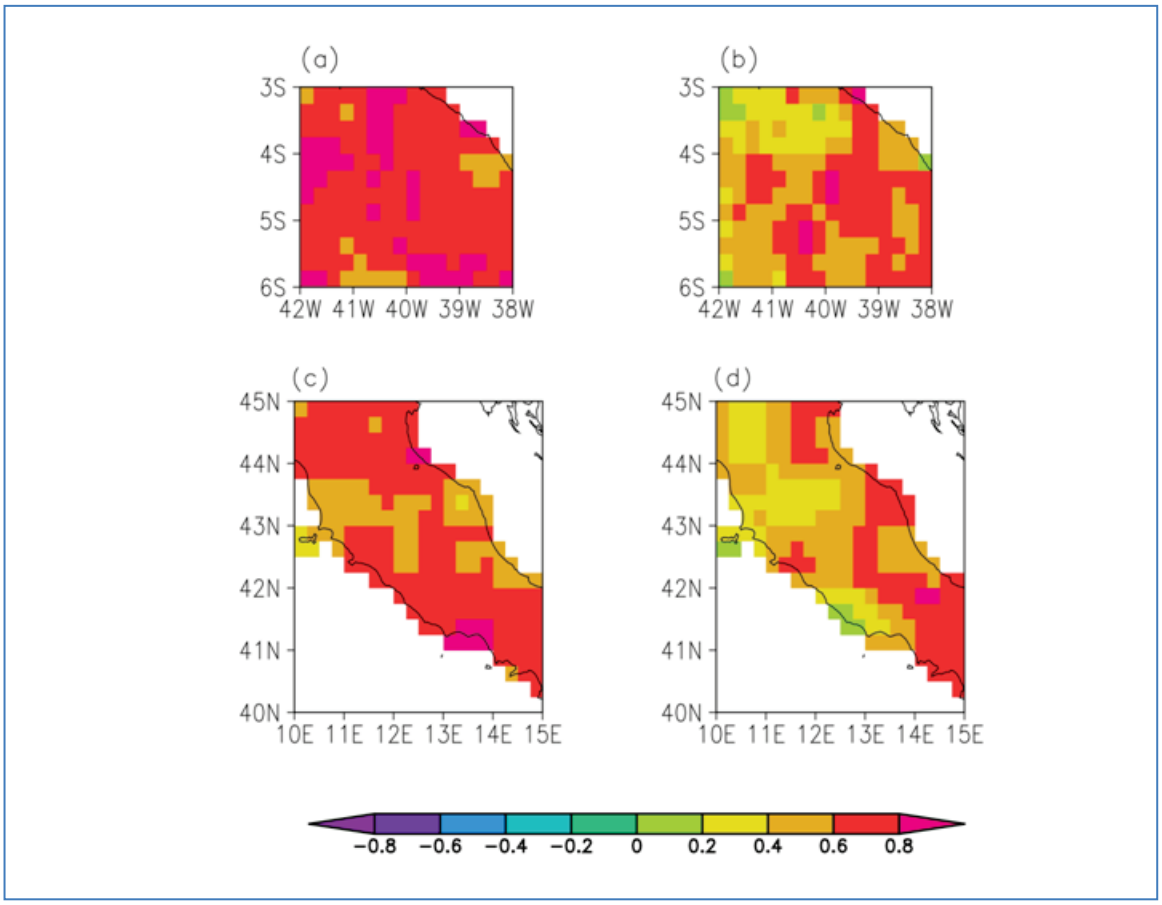
than typically achieved for European domains (albeit, with typically coarser grids). The results can be modestly improved by using u-wind alone in some summer months and v-wind alone in some winter months, consistent with Fig. 2c-d. However, these gains (not shown) are small compared to the more systematic gains reported in section 3c with the M3 approach, that moves skill levels more in line with other Europe studies.

TABLE 3. Performance of PCR in predicting CIT using 34-44°N, 5°E-25°W, u-wind and v-wind at 850mb. N PC and R Skill as defined in Table 2. AV is the mean skill across all months. Results are shown using TRMM precipitation for the M1 and M3 approaches (comparable to Table 2). In addition, results are shown using E-OBS precipitation data for 1998-2012 (comparable to all TRMM results) and for 1979-2012.

	Month	1	2	3	4	5	6	7	8	9	10	11	12	AV
TRMM M1 Approach	N PC	3	2	1	3	3	6	6	2	6	3	2	4	
	R Skill	64	40	51	54	53	14	32	51	30	36	43	54	44
TRMM M3 Approach	N PC	3	2	3	3	3	3	2	4	5	3	3	3	
	R Skill	71	55	61	69	68	45	53	63	43	57	60	66	59
E-OBS M1 1979-2012	N PC	3	4	4	3	3	4	4	2	4	3	6	3	
	R Skill	66	61	55	48	65	37	40	54	57	59	53	49	54
E-EOBS M1 1998-2012	N PC	5	3	3	4	6	4	3	1	6	3	5	4	
	R Skill	51	36	48	43	36	14	41	41	26	19	27	41	35

A relatively smaller number of PCs are needed for CIT models in winter (Table 3). This is consistent with the fact that fewer PCs are needed to explain a large fraction of the predictor variance (the leading 3 PCs explain ~85% of predictor variance in winter CIT experiments, versus ~75% for summer CIT experiments and ~70% for NEB experiments).

FIG. 3. Cross-validated skill (correlation) using PCR, for (a) NEB in May, (b) NEB in August, (c) CIT in January and (d) CIT in August. One-sided 5% significance is +0.44. The percent of significant boxes is (a) 100%, (b) 65.4%, (c) 96.1% and (d) 67.5%.

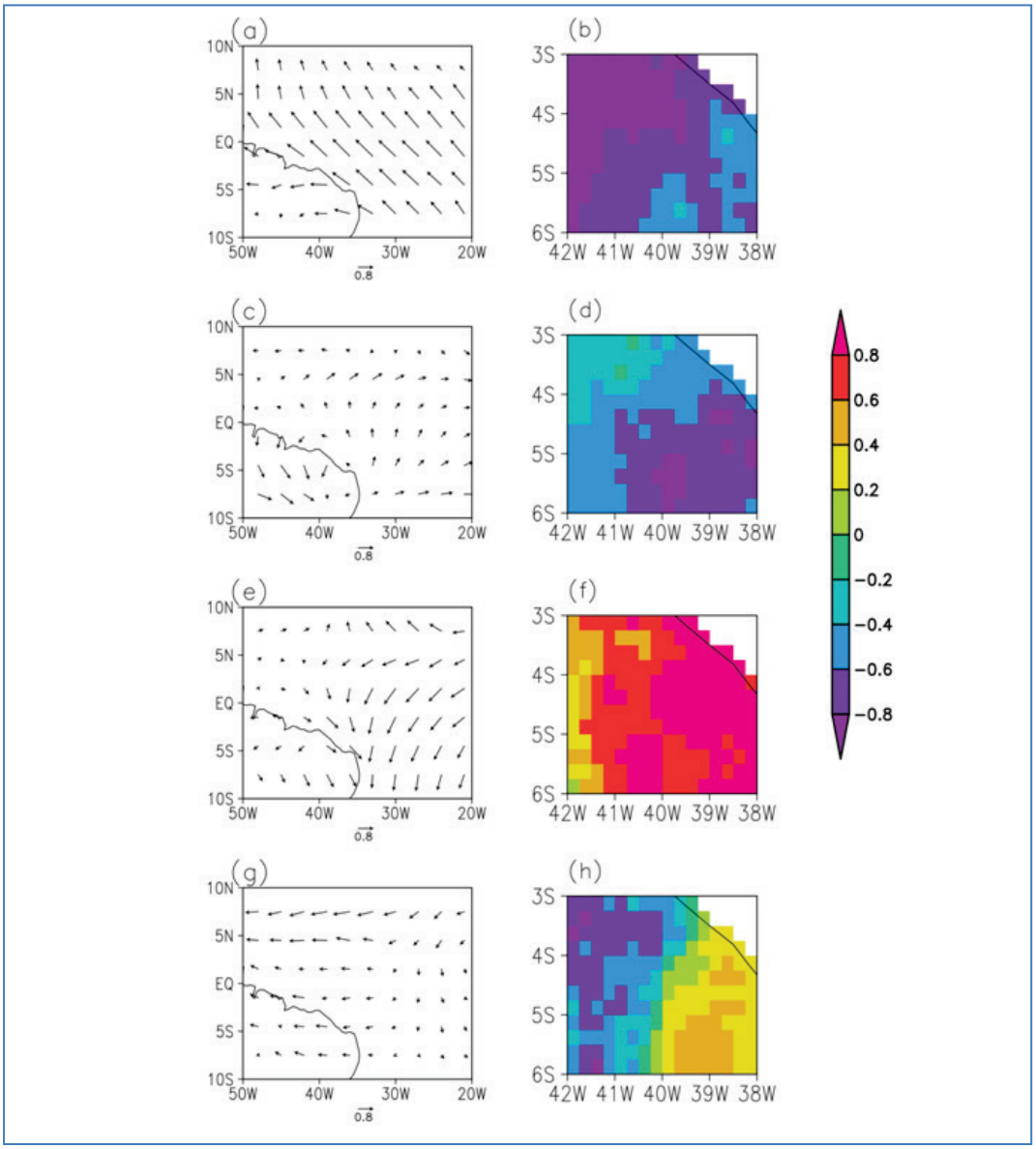


There are occasional outliers in the number of PCs selected as best model for a given month, compared to surrounding months. Examples include April for NEB and March for CIT. In both these cases, 1 PC yields maximum skill. A secondary maximum is found (in each case at 3 PCs) which yields almost identical skill. In each case, a model with 3 PCs better matches the ones selected for surrounding months. In such circumstances, the model with secondary maximum may be

selected (effectively allowing some smoothing from information in surrounding months). Indeed, both of the model selection outliers referred to above disappear when taking the M3 approach (section 3c).

While many of the months show good area-average CV skill (Tables 2 and 3), it is also relevant to assess the spatial variation of skill. Figure 3 shows the CV skill for some of the months with encouraging area-average skill (May and August for NEB, January and August for CIT). Models have good skill in almost all boxes in May for NEB (100% significant at the 5% level, assuming full degrees of freedom, as discussed in section 2d) and January for CIT (96.1% are significant); indeed most boxes have r>0.6. In August for CIT and NEB (both providing examples of performance in a climatologically relatively dry month), skill levels are generally more variable, but the majority of the domain continues to have skill levels r>0.44 (5% significance threshold) for both CIT and NEB.

FIG. 4. Leading CCA modes between 925mb wind and NEB precipitation in May (a-d) and August (e-h). (a-b) wind and precipitation mode 1 in May, (c-d) mode 2 in May, (e-f) mode 1 in August, (g-h) mode 2 in August. The CCA mode weight for u-wind and v-wind is plotted as a vector at each location to effectively form the standardized wind anomaly field that couples with the accompanying precipitation anomaly mode.



b. Diagnosis of Spatial Structure of Models using CCA

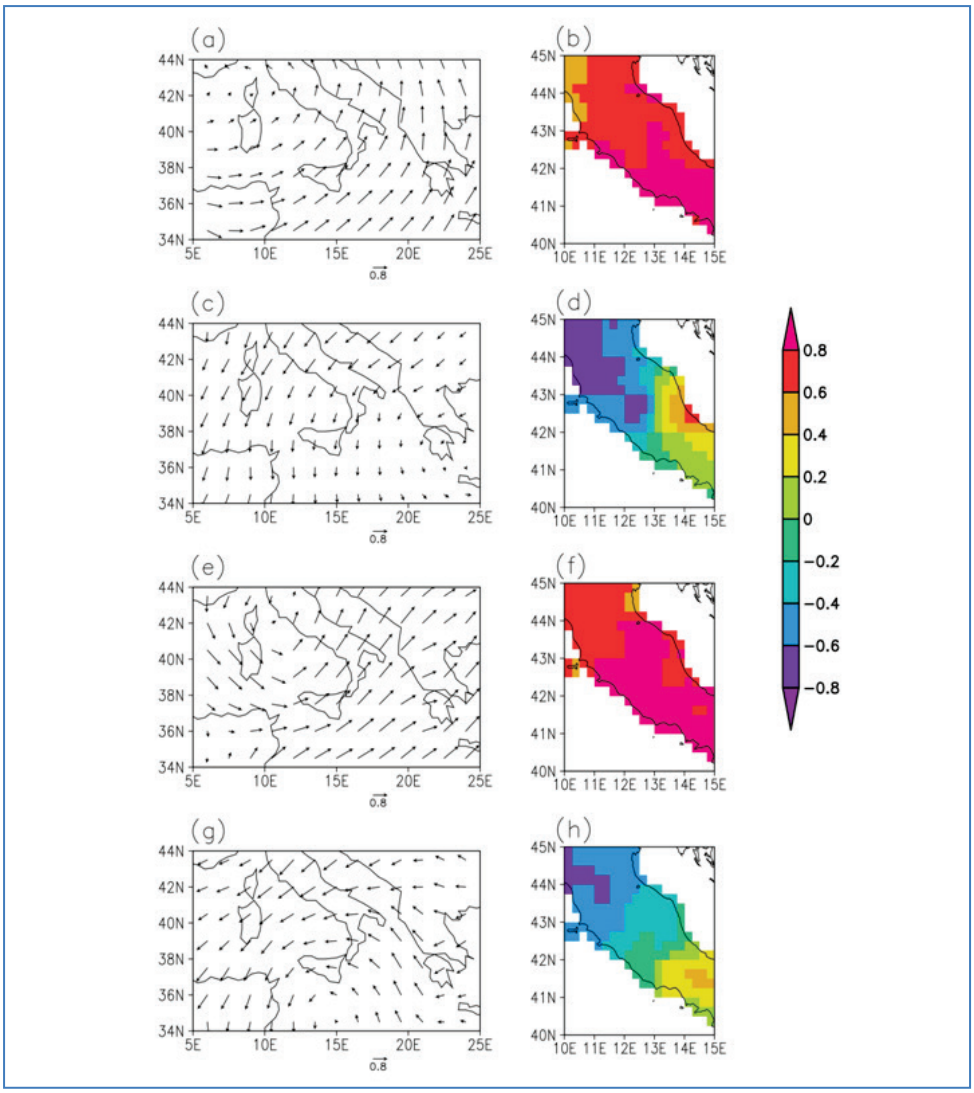
For each of the illustration months selected (May and August for NEB, January and August for CIT), 2 CCA modes achieve the best CV skill. These modes provide a succinct summary of the downscaling relationships (Fig. 4-5). For NEB, circulation

CCA mode 1 (CCA1) in May and August (Figs. 4a,e) mirrors many of the regional circulation features seen in the correlation with precipitation PC1 (Figs. 2a-b). The corresponding precipitation modes (Figs. 4b,f) describe the detailed pattern of accompanying TRMM precipitation anomalies, with strongest weights in central and northwestern parts in May, and central and eastern parts in August. In each month CCA2 is generally smaller-scale but with physically plausible wind-precipitation coupling. In May, local divergent wind anomalies near 4-8°S connect to off-shore anomalous westerly flow (Fig. 4c) and this is coupled to precipitation deficits mostly concentrated in the southeastern part of the domain (Fig. 4d). For August, anomalous easterly onshore flow (Fig. 4g) accompanies enhanced precipitation in the eastern part of the domain, and reduced precipitation in the western part (Fig. 4h).

For CIT, where there is a more complex orographic setting, the CCA modes have opportunity to more clearly identify a physiographic as well as a synoptic expression in the precipitation anomaly fields. Circulation mode 1 for January (Fig. 5a) and August (Fig. 5e) contains primarily the synoptic circulation seen in the correlation maps (Figs. 2c-d). The corresponding precipitation anomalies are of the same sign across the complete domain, but are amplified along the southwestern facing coast in both cases (Figs. 5b,f). Indeed, in the wet synoptic phase (as shown), the cyclonic southwesterly particularly enhances precipitation along the southwestern facing coast, and some amplification generally extends inland to the southwestern facing sides of the Apennine mountain chain. In the dry synoptic phase (reverse of

the phase shown), the anticyclonic northeasterly is least dry on the eastern side of Italy, while dryness is enhanced on the western side, in the lee of the Apennines.

FIG. 5. Same as Fig. 4, but for CIT precipitation versus 850mb wind, for (a-d) January, and (e-h) August.



However, superimposed on these physiographic impacts, is the dominant synoptic control, leading precipitation to nonetheless have the same sign anomaly over the complete domain (Figs. 5b,f), for a given phase of the leading wind CCA mode. This synoptic control makes statistical downscaling based solely on the projection of wind anomalies on the orography, extremely difficult to implement. For example, a southwesterly wind blowing down the mountain slope on the eastern side of Italy is actually correlated with positive precipitation anomalies. To have a direct precipitation specification simply from wind projection on orography, would require this synoptic control to be first removed from the dataset. However, the regression approaches of PCR and CCA automatically combine both the synoptic and orographic controls in the downscaling model. This is at the expense of needing high resolution local precipitation data (such as TRMM) to always specify the wind-to-precipitation relationship at each individual location.

In both January and August, CCA2 identifies a second synoptic circulation feature whose precipitation expression again appears strongly modified by orography. In January, a cyclonic centre in the southeast of the predictor domain feeds anomalous cyclonic northeasterly into southern Italy (Fig. 5c) bringing enhanced precipitation to the southeastern coastal area (Fig. 5d), but reduced precipitation to the lee side of Italy. In August, anomalous low pressure is centered south of Sicily (Fig. 5g), feeding anomalous cyclonic southeasterlies into southern areas of the predictand domain, leading to enhanced precipitation there (Fig. 5h), while in northern and especially northwestern parts in the lee of the Apennine Mountains,

precipitation anomalies associated with this mode are negative. In both months, a further contribution to the generally negative precipitation anomalies in northern parts is that here, the winds lose their anomalous cyclonic component and also become northeasterly, a direction that is typically a relatively low moisture source for northern Italy. Overall, the physical consistency and plausible nature of the CCA modes (Figs. 4-5) lends further weight to the suggestion that the TRMM data are able to identify relationships appropriate for underpinning statistical downscaling models.

c. Models constructed by pooling three months together (M3 Approach)

For NEB, taking the M3 approach, average wet season skill rises from 0.58 to 0.68, and July-August skill rises from 0.51 to 0.63 (Table 2). The increase is therefore substantial, and especially so for the dry-season months of July-August. It is hypothesized that increasing sample size in a dry season may be especially advantageous since there tends to be a relatively small frequency of high precipitation months available for establishing the circulation anomalies that are associated with high precipitation. However, this benefit can only operate if there are distinct circulation-precipitation relationships in the given climatic regime. The skill in Table 2 for September-November is still very low (even if at least averaging positively over the three months now) and suggests at this time of year in NEB, circulation alone may be insufficient for downscaling.

A further positive development with the M3 results (Table 2) is the increase in skill for June (0.13 to 0.34). This month represents a transition in precipitation-circulation relationship, from the one generally prevailing for the wet season (illustrated in Fig. 2a) to the one generally prevailing for July-August (illustrated in Fig. 2b). In such a situation, pooling three months together to construct the model is hypothesized to potentially offer enhanced gains, by allowing a clearer sampling of the relationships in the two regimes, each of which may operate in June depending on given year (e.g., whether the rainy season becomes extended in length, or ends early).

For CIT (Table 3), the benefits of the M3 approach appear even greater than those found in the tropical setting of NEB. Skill averaged across all months rises from 0.44 to 0.59. Summer semester increases are generally the greatest, and the strong dip in June and July is greatly ameliorated. Gains are also strong in the Autumn months centered on October and November.

The M3 approach is encouraging not just for the skill increases, but also because skill now shows a smoother and more plausible evolution across the annual cycle. For example, for NEB, results now strengthen gradually as the wet season develops, with strongest skill in March-May, more consistent with the months focused upon in diagnostic work relating precipitation to circulation (Hastenrath and Heller 1977; Hastenrath and Greischer 1993). Another positive development is that the number of PCs chosen for each model is now substantially more consistent, with 2-4 PCs in all but one month for CIT, and generally 4-6 PCs for the skilful NEB months.

The largest skill improvements are not focused on the 3-month sets with largest month-to-month persistence (Table 1). This appears to remove the possibility of a substantial impact of persistence compromising the CV assessment. For NEB there is almost no association between gain in skill and the N' value (r=0.03), while for CIT, there is actually a positive correlation (r=0.44) suggesting some tendency for most gain to be present when there is least inter-month correlation. Furthermore, implementation of a leave-three-out CV (not shown) was found to introduce only minor systematic skill reductions at levels that would not alter the general interpretation of results presented here.

It is also of practical relevance to assess the spatial expression of the increases in skill. The new illustrative skill maps (Figs. 6a-d) show a reduction in some of the seemingly random spatial variation in the M1 maps (Fig. 3), and a markedly reduced number of boxes with very low skill. Both these features are considered positive signs, suggesting that the M1 approach was extracting genuine relationships but with considerable noise contamination, while the M3 approach is reducing the noise substantially, crystallizing the ability of TRMM to extract clear downscaling relationships. For January in CIT (Fig. 6c), the areas with $r<0.6$ are now much reduced, and contracted into zones where topographic setting makes some reduction physically plausible, while an enhancement of skill on the southwestern facing coast now also shows coherently (southwestern zone with $r>0.8$). This is physically plausible given the primary synoptic association in this

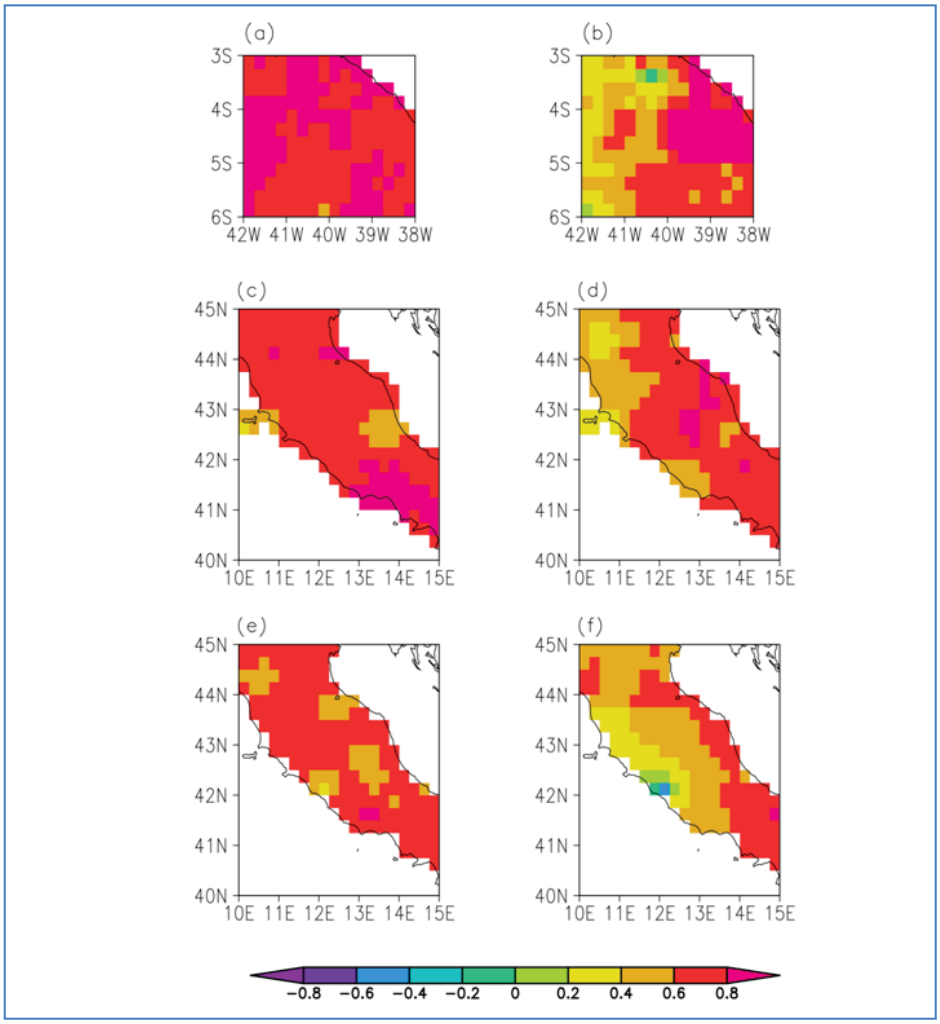
region of onshore (offshore) winds bringing enhanced (reduced) precipitation, an association that can be amplified by the prevailing coastline/mountainous terrain.

For NEB in August (Fig. 6b), the M3 results give a much clearer and consistent expression of coherent strong skill except for the western parts of the domain. Results from the M3 approach provide evidence in a number of ways that supports the existence of circulation-precipitation relationships suitable to underpin statistical downscaling schemes. The annual cycle progression and mean levels of perfect prognosis skill are also now more consistent with other methodologically similar downscaling studies (using precipitation stations or somewhat coarser precipitation grids). For example, the CIT results are comparable with a range of results for European domains (including Murphy 1999; Schmidli et al. 2007; Nicholas and Battisti 2012). For NEB, results are consistent with knowledge of the region's circulation-precipitation coupling (Hastenrath and Heller 1977; Hastenrath and Greischar 1993) and with perfect prognosis skill achieved applying similar downscaling methodologies in tropical settings in the vicinity of the ITCZ (e.g., Indonesia, Vimont et al. 2010).

d. Comparison of E-OBS and TRMM results for the CIT domain

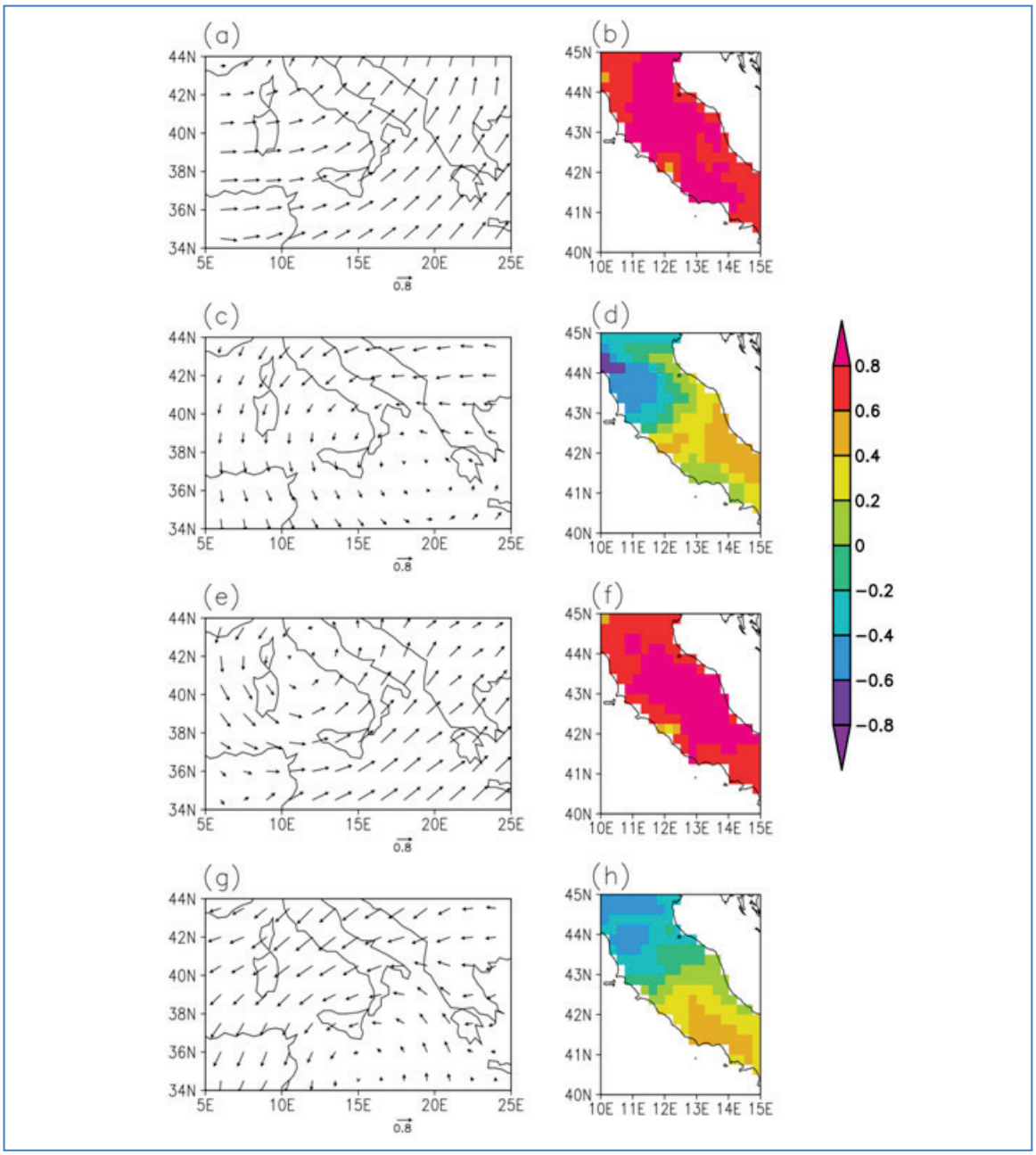
The average skill for the extended1979-2012 period that is achieved with E-OBS (Table 3, r=0.54) is comparable to that achieved with TRMM for 1998-2012. The E-OBS skill is actually a little higher than that achieved with the M1 approach (r=0.44) and a little lower than that achieved with the M3 approach (r=0.59).

FIG. 6. (a-d) are the same as Figs. 3a-d, but using the 3-month pooling method. (e-f) are the same as Figs. 3c-d, but using the E-OBS precipitation data, and for the extended period 1979-2012. The percent of boxes with r>0.44 is (a) 100%, (b) 80.9%, (c) 99.0%, (d) 89.3%, (e) 98.2% and (f) 80.2%. The percentage of boxes with r>0.30 is (a) 100%, (b) 91.2%, (c) 100%, (d) 97.6%, (e) 100% and (f) 92.4%.



Results with E-OBS also reproduce some of the same month-to-month variation in skill, such as the peak in January, and the weakest three-month average for boreal summer (June-August). Compared to the TRMM M1 results, the TRMM M3 results contain a month-to-month pattern that is closer to that achieved by E-OBS, especially noting the improvements in boreal autumn. Results with E-OBS for 1998-2012 (Table 3) do also contain some of the pattern found with TRMM (e.g., maximum in January and minimum in June), but generally the E-OBS 1998-2012 results are somewhat weaker than those achieved with TRMM. This may reflect reduced station data for E-OBS in the more recent epoch, but this would need further investigation and is not the purpose of analysis here. Rather, the aim is to assess how results with E-OBS shed light on the ability of TRMM to establish downscaling models. In this regard, it is encouraging that the TRMM results compare well with the average monthly skill achieved by E-OBS over the extended period 1979-2012. Spatial variation of skill (Figs. 6e-f) also contains broadly similar structure to that achieved with TRMM. In terms of number of CIT boxes with very low skill in Fig. 6, TRMM and E-OBS have 0 grid-boxes with skill <0.3 in January, while in August, E-OBS has 7.6% of boxes and TRMM has 2.4% of boxes <0.3.

FIG. 7. Same as Fig. 5, but using the E-OBS precipitation data, and for the extended period 1979-2012.

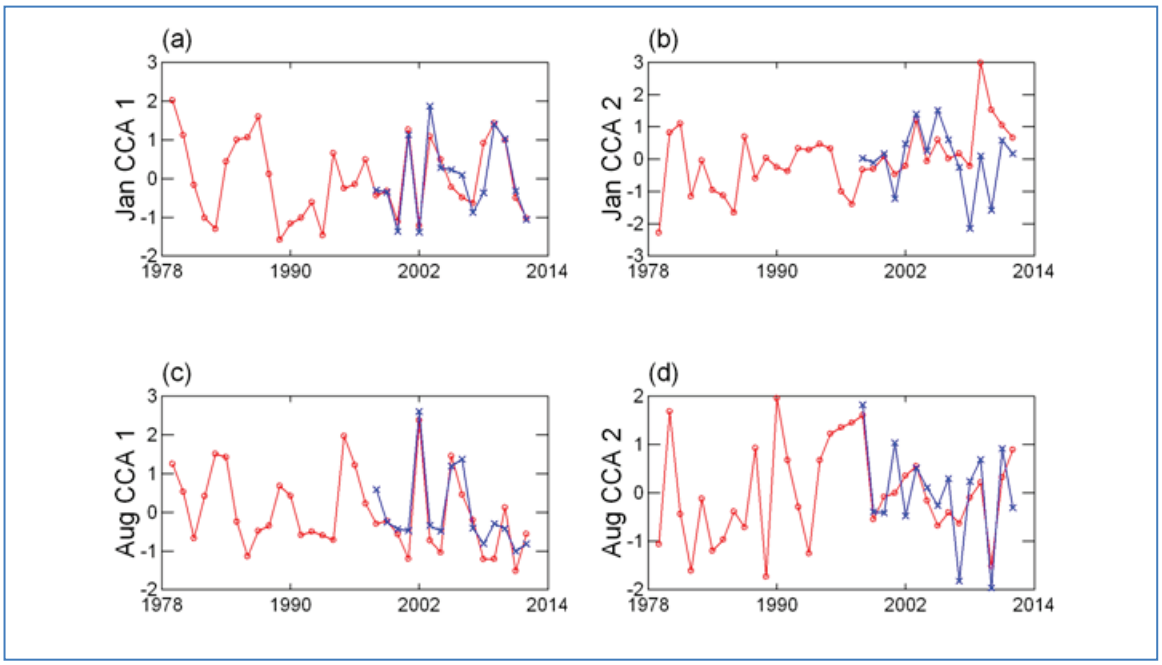


The spatial structure of the downscaling relationships is also found to be similar, illustrated as before by the leading 2 CCA modes for January and August (E-OBS modes are in Fig. 7, with counterpart TRMM modes in Fig. 5). Comparing almost all aspects, any differences tend to be of secondary, though locally important, features. For example, the January precipitation mode 1 with TRMM (Fig. 5b) appears to be amplified a little more coherently on the southwestern facing coast, when compared to the precipitation mode 1 with E-OBS (Fig. 7b). The orographic coupling of CCA precipitation-wind modes for August also appear at least as coherent in TRMM as they do in E-OBS. Overall, the spatial structures of precipitation CCA modes support the contention that TRMM contains the type of enhanced spatial structure expected of a satellite-based product, while benefiting from grounding through a station merge.

The precipitation CCA modes can also be used to efficiently summarize temporal evolution of precipitation in the two datasets (Fig. 8). For both January and August, mode 1 from TRMM and E-OBS look very consistent with each other, and over the period of common data (1998-2012), they correlate very highly (r=0.88 in January and r=0.90 in August). For mode 2, results for August with TRMM and E-OBS compare very well (Fig. 8d, r=0.74). For mode 2 in January (Fig. 8b), some discrepancies emerge around 2007, with three contrasting outliers in both E-OBS and TRMM. This result serves as a useful caution, and comparison with more than one dataset can represent a good practice in principle to highlight where potential data uncertainties lie. In section 4, downscaled scenarios for January based on E-

OBS and TRMM are compared, and it emerges that the discrepancy identified by CCA2 does not lead to major divergence in the downscaled scenarios in this case.

FIG. 8. Precipitation CCA mode time-series for the CIT domain, using TRMM precipitation (blue lines, see Fig. 5 for corresponding spatial modes) and E-OBS precipitation (red lines, see Fig. 7 for corresponding spatial modes). (a) CCA1 in January, (b) CCA 2 in January, (c) CCA1 in August, (d) CCA2 in August.



4. ASSESSING DOWNSCALED PRECIPITATION FROM A GCM SCENARIO

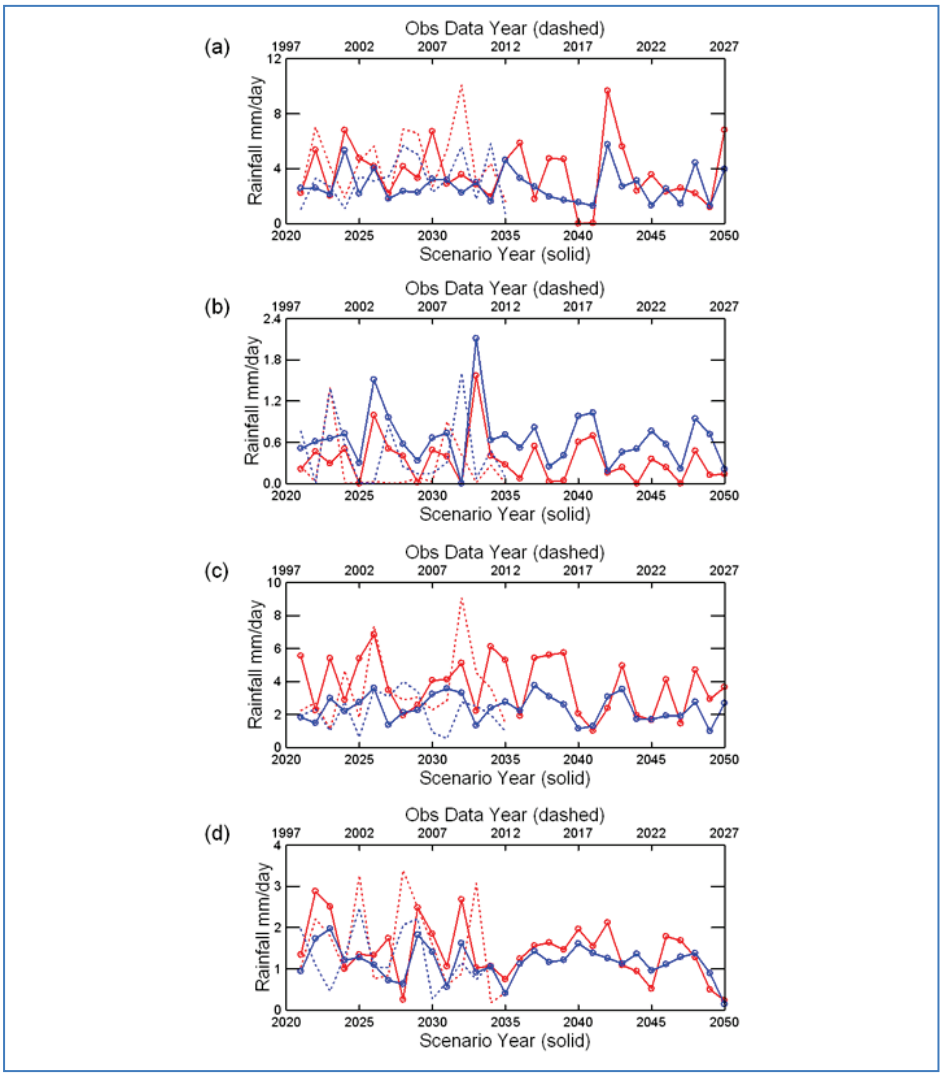
Some of the statistical downscaling models (section 3) are now applied to a GCM scenario for 2021-2050. The purpose is to illustrate whether the statistically

downscaled results appear plausible, from the perspective of being a reasonable extension of the historical record's variability and from the perspective of being dynamically consistent with the driving large-scale circulation fields. For the case study months used previously, and adopting the M1 methodology, temporal variability is illustrated in Fig. 9 while spatial anomaly precipitation fields and associated wind fields are illustrated in Fig. 10.

Downscaled precipitation anomalies for May in NEB are generally quite spatially coherent, reflecting the leading wind-precipitation regime (Figs. 4a-b). However, some spatial variability is also present, especially in the south where CCA2 (Figs. 4c-d) identified more sub-regional wind-precipitation coupling. Thus while the time-series for the coastal box (red line, Fig. 9a) is quite representative of much of the region, the time-series for the southern location (blue line, Fig. 9a) may vary out of phase. This is seen in the historical data (dashed lines on Fig. 9a) and in the scenario (solid lines on Fig. 9a). The southern box is climatologically drier, but for example, in 2044 the southern box is downscaled to be wetter than the coastal box. The spatial expression of the precipitation anomalies for this month (Fig. 10b) shows generally substantially below normal precipitation except for the area in the southwestern part of the domain (mirroring well the area identified in CCA2, Fig. 4d). The precipitation anomalies are plausible given the driving wind anomalies (Fig. 10a), with large-scale structure in the equatorial and tropical North Atlantic (easterly with a southerly component) consistent with a northward shifted ITCZ (CCA1, Fig. 4a), while in the southern part of the domain, anomalous onshore convergent wind anomalies (reverse phase of CCA2, Fig.

4c) are consistent with the less-negative (near-zero) precipitation anomalies in the southwestern part of Fig. 10b.

FIG. 9. Comparison of PCR downscaled scenarios (solid lines) and the observed data used to train the downscaling models (dashed lines). Grid-boxes used are shown on Figs. 1a-d. (a) NEB May, for coastal location (red), and southern location (blue); (b) same as (a) but for August; (c) CIT January, for southwestern location (red) and eastern location (blue); (d) same as (c) but for August and for southeastern location (blue) and northern location (red).



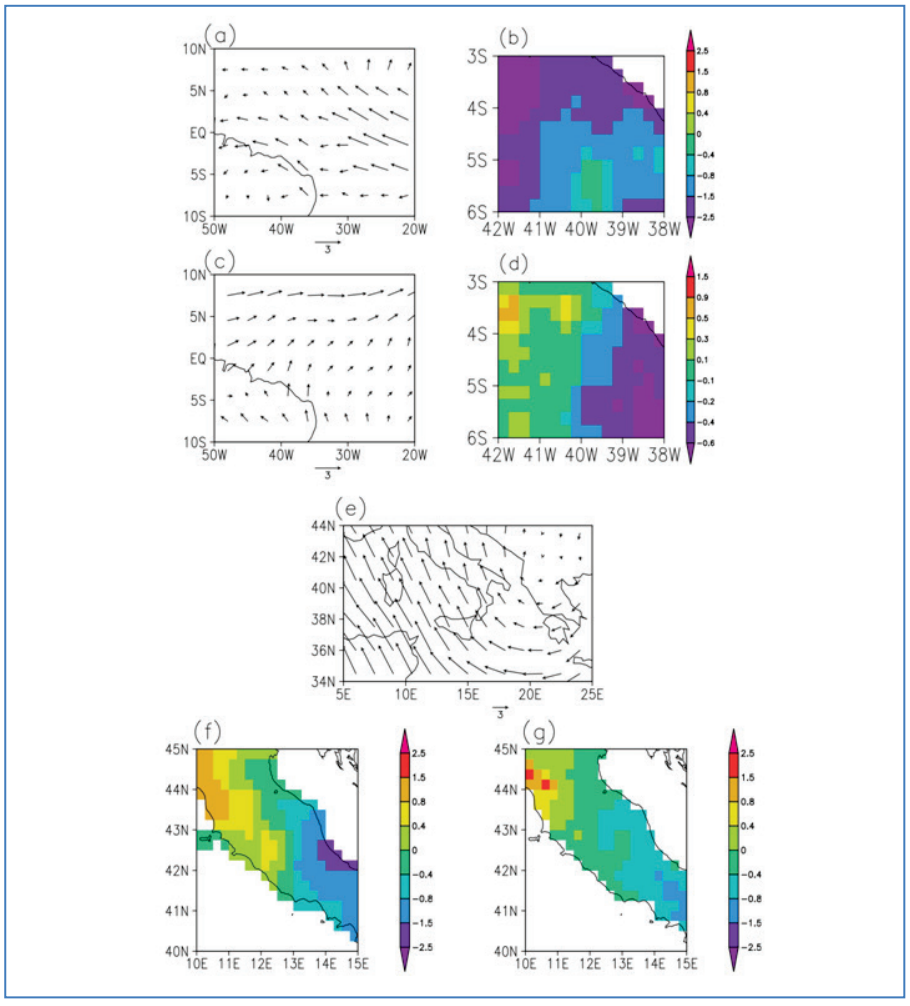
Statistically downscaled results for NEB in August (Fig. 9b) provide an illustration of the performance of the downscaling system during a climatologically dry month that, nonetheless, has strong expression of precipitation anomalies in the regional circulation. Both of the illustrative boxes show how a substantial fraction of months have near-zero or zero precipitation. Clearly, the normal distribution assumptions of linear regression are compromised in such a situation. However, the existence here of strong, approximately linear relationships with interannual variations in circulation leads to models with good explanatory power and scenarios with substantial and plausible variability. The simple treatment of zero precipitation as occurring whenever the regression prediction is less than zero also leads, in this case, to quite realistic frequencies of zero monthly precipitation. This can be seen in Fig. 9b, and is found when counting across all box-months (15 months * 204 boxes in the domain), with 12.0% of box-months being zero in the TRMM data 1998-2012, 15.0% of box-months being zero in the statistically downscaled baseline GCM simulation for 1998-2012, and 9.6% of boxes being zero in the statistically downscaled GCM scenario (consistent with a small overall increase in NEB area-average precipitation from 0.34 mm/day during the 1998-2012 baseline to 0.42 mm/day during the 2021-2050 scenario). This should not disguise the need for better statistical treatment in dry season months (as returned to in section 5). However, the time-series (Fig. 9b) and spatial anomaly patterns (Fig. 10c-d) are considered already very promising and suggest the general approach can be feasible for downscaling even in some settings with climatologically dry months. For example, some large-swings in the historical data (such as very dry 1999 to very wet

2000, Fig. 9b dashed lines) are also reproduced in the scenario (Fig. 9b solid lines), such as from 2032 (dry) to 2033 (very wet). The spatial expression of the extremely dry August of 2032 (Fig. 10d) shows that the negative precipitation anomalies are widespread across the western side of the domain, as seen in (the negative phase of) CCA1 and CCA2 (Figs. 4f,h). The driving wind anomalies in August 2032 (Fig. 10c) are mostly consistent with the negative phase of CCA1 (Fig. 4e), with anomalous low-level southerly winds extending from the tropical South Atlantic, across the equator and curving into southwesterly anomalies that continue beyond the domain shown into the West African monsoon.

For CIT in January, the contrasting downscaled precipitation features in the south of the domain are illustrated for a southwestern coast and east coast location (Fig. 9c). General character in the historical observations (dashed lines) is reproduced in the downscaled scenario (solid lines), with the east coast generally drier, but several years when the precipitation anomalies in the two locations are contrasting depending on the orientation of wind anomalies to the coastline and orography. A good example is in 2049, when the eastern location is downscaled to have its driest month in the scenario, while the western location is close to normal. The driving wind field (Fig. 10e), projects on the negative phase of CCA2 (Fig. 5c) and displays anomalous anticyclonic southerlies over the southern part of the domain, consistent with the amplification of the extreme dry conditions on the lee (northeastern-facing) side of Italy. The amplification is clearly seen in the precipitation anomaly map for 2049 near 42°N (Fig. 10f), and is seen in CCA2 (Fig. 5d). The driving wind circulation loses its anticyclonic

curvature over the northern part of the domain, giving rise to positive precipitation anomalies, especially in western parts, where the wind anomalies have generally an onshore and orographic up-slope component.

FIG. 10. Examples of monthly wind anomaly scenario fields and associated downscaled precipitation anomaly fields (based on TRMM except for (g)). (a) May 2044 925mb wind, (b) May 2044 NEB precipitation, (c) August 2032 925mb wind, (d) August 2032 NEB precipitation, (e) January 2049 850mb wind, (f) January 2049 CIT precipitation, (g) Same as (f) but using downscaling relationships constructed using E-OBS precipitation data over 1979-2012.



A further assessment is possible for CIT by comparing the downscaled scenarios based on the TRMM data, with those achieved using E-OBS. There are 178 grid-boxes where both E-OBS and TRMM downscaled scenarios are created. For a given month, the correlation between the two scenarios for grid-box i is calculated (2021-2050), and then the average of the correlations over the 178 grid-boxes is made (see Table 4). This reveals a generally very good match between the variations of the two scenarios, despite the models being based on the different precipitation datasets, and fitted over different years (1979-2012 for E-OBS, 1998-2012 for TRMM). The average correlation for the case study months of January and August (0.94 in both cases) is only slightly higher than the average over all months (0.87). This is considered further evidence that models based on TRMM can be considered to generate plausible downscaled scenarios.

TABLE 4. Cross-correlation (*100) of the downscaled CIT scenarios from PCR models trained on (a) TRMM over 1998-2012, and (b) E-OBS over 1979-2012. The value shown for each month is the average correlation between the downscaled scenario time-series for 2021-2050.

Month	1	2	3	4	5	6	7	8	9	10	11	12	AV
R	94	91	87	97	93	90	74	94	57	92	81	95	87

In addition, the spatial anomaly fields in the scenarios are also found to be generally consistent, as illustrated for the January 2049 case study (compare Figs. 10f-g). The broad pattern of anomalies is similar, but there are smaller-scale discrepancies as well, such as the weaker amplification of drought conditions on the east facing lee side of the Apennine Mountains (consistent with the difference also seen in CCA2 for TRMM and E-OBS, compare Figs. 5d, 7d).

5. DISCUSSION AND CONCLUSIONS

A generic and consistent methodology has been developed to statistically downscale GCM output to monthly precipitation at ~25km spatial resolution. An initial demonstration of the methodology has been undertaken for two contrasting locations (NEB and CIT). By establishing the potential of using TRMM monthly precipitation data for statistical downscaling, there emerges the possibility for downscaling to domains selected within the 50°S-50°N extent of the dataset, representing a large fraction of the earth's surface. The results presented here support this assertion for at least some locations and times of year, and further studies are now proposed to more clearly assess the extent of the capability. The potential has been established through a range of diagnostic assessments for the NEB and CIT domains. The magnitude and annual cycle patterns of perfect prognosis skill for NEB and CIT are consistent with levels found in previous studies taking PCR or related approaches. Diagnostic CCA results show clear coupled modes for NEB and especially CIT, where the orographic effect is very clear, driving home the sense that there is genuine high resolution spatial anomaly information in the TRMM data that can be incorporated into statistical downscaling relationships. For CIT, tests with a station-only gridded dataset (E-OBS) reproduce the basic levels of PCR skill and spatial structures seen in the TRMM CCA modes, while the downscaled scenarios based on TRMM and E-OBS have substantial common

variance, and appear dynamically consistent with the driving GCM wind fields, all suggesting that scenarios based on TRMM can be considered a plausible alternate to ones based on station data alone. One aspect for subsequent investigation (both in the context of TRMM and also more generally in statistical downscaling) is the consistency in capturing trends, for which an ensemble of scenarios with known and clear trend behavior may be explicitly studied. A further area for exploration is the extent to which prefect prognosis skill is maintained in regions with less rainfall stations contributing to the merged TRMM product; the results in this paper begin to provide a benchmark against which to make such comparisons.

The situation is likely to substantially improve further in the future for a number of reasons. Firstly, as the data set grows in length, estimation of relationships will become less prone to noise, and less variability in the levels of skill in the models can be expected. Indeed, many of the lower grid-box CV correlation scores for the TRMM results reported here most likely are attributable to sample-size rather than fundamental lack of predictability (e.g., see Fig. 10 of Schmidli et al. 2007). One improvement can be achieved through sensible pooling of months together in order to increase degrees of freedom in the estimation of model parameters. This is also considered an effective innovation to allow smooth incorporation of the evolving downscaling relationships through the annual cycle. Initial testing of pooling adjacent three months (to specify the central month of the three) has been very encouraging, with ~35% increase in mean CV correlation skill for CIT, and generally more stable and continuous results through the annual cycle. Such an outcome

also supports the assertion that TRMM data are capturing genuine circulation-precipitation relationships that can be used for downscaling. Future research can be considered to assess optimum choice of window for model calculation, including the option of weighted regression, with greatest weight given to the target month, and reduced weight to surrounding months. One further natural technical advance, that has not been introduced in this paper, is the transformation of precipitation predictand closer to normal, or application of techniques (such as skew normal regression) which explicitly adapt to the skewed nature of the predictand. These advances can also be expected to improve perfect prognosis skill and, perhaps more importantly, achieve better statistical character of precipitation scenarios, especially in climatologically dry months.

Skill levels reported here are also considered conservative estimates, given the use of just low-level circulation. There is clearly potential to enhance the perfect prognosis skill of models through additional variables, as well as to tune them to best capture global change, through explicitly incorporating key effects that may not be fully represented with circulation alone, such as effects related to humidity changes. One aspect that is very difficult to incorporate, is land surface feedback. In this regard, the results reported here for NEB are noteworthy in that persistence increases as wet season progresses (Table 1), which may be a feature of land-surface feedback, and may need special treatment through downscaling predictors such as soil moisture or antecedent precipitation anomalies. A further aspect to consider for a precipitation regime like NEB that is strongly influenced by sea

surface temperature (SST) anomalies (Ward and Folland 1991; Hastenrath et al. 2009) is whether inclusion of SST predictors may improve downscaling models, or at least provide an additional plausible alternate scenario.

Such consideration of SST as additional predictors may also help in the application of the methodology developed here for statistical downscaling of seasonal forecasts. Model Output Statistics (MOS) is the more usual approach to statistical downscaling of seasonal forecasts, where GCM seasonal forecasts are related directly to the observed high resolution predictand (e.g., Indeje et al. 2006). However, perfect prognosis models, possibly including observed SST as well as reanalysis circulation, can be applied to seasonal forecasts to at least provide a feasible high resolution precipitation outcome given the large-scale monthly GCM forecast of the coupled SST-circulation features, though this would be achieved at the expense of less effective systematic correction of model spatial anomaly modes (Feddersen et al. 1999).

Given the promising downscaling relationships identified in this paper, a family of ideas are triggered that would seek to enhance the utility of the relatively short TRMM record through approaches that merge results from TRMM with those of extended time, but lower resolution precipitation (LRP) datasets. For example, downscaling may be implemented as a two-step process, moving from circulation to LRP (long period), then LRP to TRMM (short period, but tightly coupled). Combining information in LRPs and TRMM can also provide routes to a longer climatological base-period for establishing baseline means against which to

compare downscaled scenarios, while such approaches could also be utilized to develop methodologies for assessing and minimizing sensitivities to DMD variability. Such analyses are prompted by the positive assessments in this paper, indicating that the methodology adopted is able to extract basic synoptic and physiographic features in high resolution (0.25°) monthly precipitation, and successfully incorporate these in statistical downscaling models.

APPENDIX A

Further Details of the Implemented Statistical Downscaling Models

Both PCR and CCA exploit the efficient decomposition of the predictor variable space-time matrix **X** into orthogonal time-series ($\boldsymbol{a}_j$, the PCs). The leading $\boldsymbol{a}_j$ time-series explain the maximum possible variance in **X**, for *j* orthogonal series. In PCR, the $\boldsymbol{a}_j$ immediately form the regression predictors for each predictand precipitation grid-box, with the fitted PCR best-estimate prediction model taking the form:

$$r_{pt} = c_o + \sum_{j=1..q} c_j a_{jt} \qquad \text{(A1)}$$

where r_{pt} is precipitation (mm/day) for grid-box *p* and month *t*, c_n are the regression model coefficients, a_{jt} is the predictor value for PC *j* and month *t*, and summation is over the *q* included PCs. The choice of *q* was determined through application of a leave-one-out CV procedure (see section 2d). However, it is desirable to set an upper limit (MAX) on the maximum number of PCs permitted, to avoid the CV routine cycling through low order PCs that are merely representing noise. The choice of MAX can be assessed through the presence of shelves in PC eigenvalue plots and/or through retaining a desired fraction of variance in the predictor dataset. Ideally, a shelf is seen through a drop in eigenvalue from MAX to MAX+1, with subsequent eigenvalues lying

on a gently sloping or near-flat surface (representing non-distinct eigenvalues, whereas MAX is statistically significantly different from MAX+1). Based on consultation of eigenvalue plots as well as the general patterns of percent of variance retained in each month, it was decided to set MAX to 6 for all results in the paper. The general pattern of results was not sensitive to this choice. However, the situation can vary by region, domain size and time of year, and in other settings, allowing MAX to vary may be necessary. For the domains used here, setting MAX to 6 ensured at least 90% of predictor variance was always retained; this leans towards a relatively high fraction of retained variance, since the specification relationships may potentially reflect a very tight fit between circulation and precipitation, with low order PCs potentially having some signal in precipitation, therefore warranting retention of more predictor variance than might be retained in, for example, a methodology for extracting seasonal prediction signals from predictor data fields (where a typical retention strategy might be closer to 80% of predictor variance).

In the conventional application of CCA to meteorological fields (Barnett and Preisendorfer 1987; von Storch and Zwiers 1999) the normalized predictor PC time-series are subject to a transformation to yield a new set of predictor time-series:

$$u_{gt} = \sum_{j=1..k} a_{jt} r_{gj} \qquad \text{(A2)}$$

where $\boldsymbol{u}_g$ forms the *g*th predictor CCA mode time-series, and $\boldsymbol{r}_g$ is a unit vector that determines, for the *g*th CCA mode, the weighting given to each of the *k* retained PCs (each PC standardized to unit variance). The PCs of the predictand precipitation data (**Y**) are similarly transformed by unit vectors ($\boldsymbol{s}_g$) to yield a set of predictand CCA mode time-series ($\boldsymbol{v}_g$). For CCA1, the transformation weights (i.e., the weights in $\boldsymbol{r}_1$ and $\boldsymbol{s}_1$) are selected to yield maximum correlation between $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$. This defines the first pair of coupled CCA mode time-series. Subsequent modes are defined to have

maximum cross-correlation, subject to orthogonality internal to the u_g time-series and v_g time-series. The *g*th spatial CCA modes are given by the correlation between u_g and **X** (defining the predictor CCA modes) and v_g and **Y** (defining the predictand CCA modes). These spatial modes serve as efficient diagnostics of the spatial structure of cross-correlation between **X** and **Y**. Barnett and Preisendorfer (1987) show that the value at each location in the *g*th **Y** CCA mode is proportional to the standardized least squares regression estimate of **Y** using u_g. To implement CCA in this paper, the maximum number of predictor PCs was set to 6 (consistent with the choice deployed for PCR), maximum predictand PCs was set to 4 (corresponding to retention of about 90% of variance) and maximum CCA modes was set to 3. Increasing these values rarely led to a change in results, and was considered to be drawing dominantly on noise in the system. For the different possible permutations of PCs and CCA modes, the best model was selected using the same CV procedure used for PCR.

The correlation between u_g and v_g gives the canonical correlation coefficient for CCA mode *g*. This can be very high even when there is no real CV skill present (and the canonical correlation is inevitably amplified as larger numbers of PCs are retained). The real explanatory power of a CCA mode is only really seen when CV is applied.

APPENDIX B

Natural Decadal to Multidecadal (DMD) Variability

Natural DMD variability exists in the climate system (e.g., see Meehl et al. 2009) and may be a problem for statistical downscaling in a number of ways. DMD variability in the precipitation (and circulation) during the historical fitting period may make models less reliable due to reduced degrees of freedom, and may require longer CV windows. In places where this is the case, possible actions include detrending / low-pass filtering

of data prior to model fitting. However, this may discard valuable information. Degrees of freedom reduction should, nonetheless, be considered in statistical significance estimates. For the time-period worked with here, neither domain contains a strong year-to-year persistence.

A further problem may be that the nature of the relationships between circulation and local precipitation may vary over time due to the DMD changes in the background state. This is recognized generally as a source of uncertainty in statistical downscaling. There are two potential ameliorating factors in the methodology (section 2) used here: (i) the relationships used are primarily direct circulation-precipitation ones, defined over domains that are relatively modest expansions of the target precipitation domains. Such local direct relationships can generally be expected to be more stable than ones involving more remote teleconnection relationships, such as through the use of large-scale teleconnection atmospheric indices, and further investigation is proposed to better quantify this; (ii) the formation of predictors into orthogonal series can be expected to lead to less sensitivity in model prediction coefficients, compared to the application of correlated predictors, when co-linearity and amplified standard errors prevail. The analyses presented comparing results with E-OBS (1979-2012) and TRMM (1998-2012) at least confirm for CIT, that DMD changes in relationships are not overwhelming in this case. However, more work is needed to highlight the general magnitude of this problem for the methodology used here, combined with approaches that may further ameliorate sensitivity, such as through combined analyses that include longer precipitation datasets (as discussed in section 5).

DMD variability also has the potential to systematically bias the downscaling models in terms of the mean downscaled value for a climate change scenario:

(i) It is possible that the predictand climate is biased towards a particular DMD mode during the model-training period (for TRMM models here, 1998-2012). Consider:

$$R' = \overline{R(obs)}_{98-12} - \overline{R(true)}_{98-12} \qquad \text{(B1)}$$

where $\overline{\boldsymbol{R}(\boldsymbol{true})_{98-12}}$ is considered the mean climate that would result from an ensemble of outcomes through 1998-2012 with observed radiative forcing, sampling the range of DMD variability possible and averaging to cancel the stochastic DMD signals; $\overline{\boldsymbol{R}(\boldsymbol{obs})_{98-12}}$ is the actual observed mean. If ***R'*** turns out to be substantial for the actual observed realization, the effect will not introduce any artificial trends into the GCM downscaled results (base relative to scenario), but when the downscaled values are expressed in absolute units (mm/day), the overall mean precipitation specified will be systematically altered by *R'* (i.e., the amount by which a natural DMD phase perturbed the mean climate during 1998-2012). To the extent *R'* for 1998-2012 can be estimated from knowledge of a region's climate, this could be corrected for.

(ii) It is also possible that the GCM, during the base period, happened to be in a "wet" or "dry" DMD mode for the predictand region and its associated predictors. Consider that the model is in "wet" phase. In this case, there will tend to be a bias toward relatively "dry" downscaled values in the scenario, thereby introducing an artificial downward trend from the baseline into the future scenario. This occurs because the predictors will all be referenced to mean values that correspond to relatively wet conditions in the model (see Eq. (3)). A possible solution is to use an ensemble of coupled model runs for the base period, thereby sampling across different possible DMD phases in the coupled model. In other words, an estimate of the model's *true mean* is achieved (for a given level of radiative forcing).

REFERENCES

Barnett T.P., and R.W. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States air temperature determined by canonical correlation analysis. Mon. Wea. Rev., 115: 1825–1850, doi: 10.1175/1520-0493.

Barnston A.G., and C.F. Ropelewski, 1992: Prediction of ENSO episodes using canonical correlation analysis. J. Climate, 5: 1316-1345, doi: 10.1175/1520-0442(1992) 005<1316: POEEUC>2.0.CO;2.

Bellucci A., S. Gualdi, S. Masina, A. Storto, E. Scoccimarro, C. Cagnazzo, P. Fogli, E. Manzini, A., and Navarra, 2013: Decadal cimate predictions with a coupled OAGCM initialized with oceanic reanalyses. Climate Dyn., 40: 1483-1497, doi: 10.1007/s00382-012-1468-z.

Dee D.P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quart. J. Roy. Meteor. Soc., 137: 553–597, doi: 10.1002/qj.828.

Estrada F., V.M. Guerrero, C. Gay-García, and B. Martínez-López, 2013: A cautionary note on automated statistical downscaling methods for climate change. Clim. Change, 120: 263–276, doi: 10.1007/s10584-013-0791-7.

Feddersen H., A. Navarra, and M.N. Ward, 1999: Reduction of model systematic error by statistical correction for dynamical seasonal predictions. J. Climate, 12: 1974–1989, doi: 10.1175/1520-0442(1999)012<1974:ROMSEB>2.0.CO;2.

Fowler H.J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modeling to impacts studies: recent advances in downscaling techniques for hydrological modeling. Int. J. Climatol., 27: 1547–1578, doi: 10.1002/joc.1556.

Hastenrath S., and L. Heller, 1977: Dynamics of climatic hazards in northeast Brazil. Quart. J. Roy. Meteor. Soc., 103: 77–92, doi: 10.1002/qj.49710343505.

Hastenrath S., and L. Greischar, 1993: Circulation mechanisms related to northeast Brazil precipitation anomalies. J. Geophys. Res., 98(D3): 5093–5102, doi: 10.1029/92JD02646.

Hastenrath S., L. Sun, and A.D. Moura, 2009: Climate prediction for Brazil's Northeast by empirical and numerical modeling methods. Int. J. Climatol., 29: 921–926, doi: 10.1002/joc.1770.

Haylock M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, and M. New, 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation. J. Geophys. Res. (Atmospheres), 113: D20119, doi: 10.1029/2008JD10201.

Hewitson B.C., and R.G. Crane, 1996: Climate downscaling: techniques and application. Climate Res., 7: 85-95.

Hewitson B.C., and R.G. Crane, 2006: Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. Int. J. Climatol., 26: 1315–1337, doi: 10.1002/joc.1314.

Huffman G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F. Stocker, and D.B. Wolff, 2007: The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. J. of Hydrometeor., 8: 38–55, doi: 10.1175/JHM560.1.

Huffman G.J., and D.T. Bolvin, 2013: TRMM and Other Data Set Documentation. Available at ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf, pp40.

Indeje M., M.N. Ward, L.J. Ogallo, G. Davies, M. Dilley, and A. Anyamba, 2006: Predictability of the Normalized Difference Vegetation Index in Kenya and Potential Applications as an Indicator of Rift Valley Fever Outbreaks in the Greater Horn of Africa. J. Climate, 19: 1673–1687, doi: 10.1175/JCLI3708.1.

Katz R.W., and H.M. Glantz, 1986: Anatomy of a precipitation index. Mon. Wea. Rev., 114: 764-771, doi: 10.1175/1520-0493(1986)114<0764:AOARI>2.0.CO;2.

Maraun D., and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. Rev. Geophys., 48: RG3003, doi: 10.1029/2009RG000314.

Meehl G.A., and Coauthors, 2009: Decadal Prediction: Can it be skillful? Bull. Amer. Meteor. Soc., 90: 1467–1485, doi: 10.1175/2009BAMS2778.1.

Murphy J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. J. Climate, 12: 2256–2284, doi: 10.1175/1520-0442(1999)012<2256: AEOSAD>2.0.CO;2.

Nicholas R.E., and D.S. Battisti, 2012: Empirical Downscaling of High-Resolution Regional Precipitation from Large-Scale Reanalysis Fields. J. Appl. Meteor. Climatol., 51: 100-114, doi: 10.1175/JAMC-D-11-04.1.

Sachindra D.A., F. Huang, A. Barton, and B.J.C. Perera, 2014: Statistical downscaling of general circulation model outputs to precipitation – part 1: calibration and validation. Int. J. Climatol.,: DOI: 10.1002/joc.3914.

Schmidli J., C.M. Goodess, C. Frei, M.R. Haylock, Y. Hundecha, J. Ribalaygua, and T. Schmith, 2007: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. J. Geophys. Res., 112: D04105, doi: 10.1029/2005JD007026.

Scoccimarro E., S. Gualdi, A. Bellucci, A. Sanna, P.G. Fogli, E. Manzini, M. Vichi, P. Oddo, and A. Navarra, 2011: Effects of tropical cyclones on ocean heat transport in a high resolution coupled general circulation model. J. Climate, 24: 4368–4384, doi: 10.1175/2011JCLI4104.1.

Timm O., and H. Diaz, 2009: Synoptic-Statistical Approach to Regional Downscaling of IPCC Twenty-First-Century Climate Projections: Seasonal Precipitation over the Hawaiian Islands. J. Climate, 22: 4261-4280, doi: 10.1175/2009JCLI2833.1.

van Vuuren D.P., and Coauthors, 2011: The representative concentration pathways: an overview. Clim. Change, 109: 5-31, doi: 10.1007/s10584-011-0148-z.

Vimont D.J., D.S. Battisti, and R.L. Naylor, 2010: Downscaling Indonesian precipitation using large-scale meteorological fields. Int. J. Climatol., 30: 1706–1722, doi: 10.1002/joc.2010.

von Storch H., E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. J. Climate, 6: 1161–1171, doi: 10.1175/1520-0442(1999)012<3505:OTUOII>2.0.CO;2.

von Storch H., 1999: On the use of “inflation” in statistical downscaling. J. Climate, 12: 3505–3506, doi: 10.1175/1520-0442(1999)012<3505:OTUOII>2.0.CO;2.

von Storch H., and F.W. Zwiers, 1999: Statistical Analysis in Climate Research. Cambridge University Press, 484pp.

von Storch H., 2005: Models of global and regional climate. MG. Anderson (Ed): Encyclopedia of Hydrological Sciences, Part 3. Meteorology and Climatology, 478-490, doi: 10.1002/0470848944.hsa035.

Ward M.N., and C.K. Folland, 1991: Prediction of Seasonal Precipitation in the North Nordeste of Brazil Using Eigenvectors of Sea Surface Temperature. Int. J. Climatol., 11: 711-743, doi: 10.1002/joc.3370110703.

Widmann M., C.S. Bretherton, and E.P. Salathé, 2003: Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. J. Climate, 16: 799–816, doi: 10.1175/1520-0442(2003)016<0799:SPDOTN>2.0.CO;2.

Wigley T.M.L., P.D. Jones, K.R. Briffa, and G. Smith, 1990: Obtaining subgrid scale information from coarse-resolution general circulation model output. J. Geophys. Res. (Atmospheres), 95(D2): 1943–1953, doi: 10.1029/JD095iD02p01943.

Wilby R.L., and T.M.L. Wigley, 1997: Downscaling general circulation model output: A review of methods and limitations. Prog. Phys. Geogr., 21: 530–548, doi: 10.1177/030913339702100403.

Wilby R.L., and C.W. Dawson, 2013: The Statistical DownScaling Model: insights from one decade of application. Int. J. Climatol., 33: 1707–1719, doi: 10.1002/joc.3544.

Yevjevich V., 1972: Probability and statistics in hydrology. Water Resources Publications: Fort Collins, CO, pp302.

Zorita E., and H. von Storch, 1999: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. J. Climate, 12: 2474–2489, doi: 10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2