Bridging Past and Future Climate Across Paleoclimatic Reconstructions, Observations, and Models: A Hydroclimate Case Study

JASON E. SMERDON*

Lamont-Doherty Earth Observatory of Columbia University, New York, USA

BENJAMIN I. COOK

NASA Goddard Institute for Space Studies, New York, USA

Lamont-Doherty Earth Observatory of Columbia University, New York, USA

EDWARD R. COOK

Lamont-Doherty Earth Observatory of Columbia University, New York, USA

RICHARD SEAGER

Lamont-Doherty Earth Observatory of Columbia University, New York, USA

**Corresponding author address:* Jason E. Smerdon, Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W, P.O. Box 1000, Palisades, NY 10964.

E-mail: jsmerdon@ldeo.columbia.edu

Abstract. Potential biases in tree-ring reconstructed Palmer Drought Severity Index (PDSI) are evaluated herein using Thornthwaite (TH), Penman-Montieth (PM), and Self-Calibrating Penman-Montieth (SC) PDSI in three diverse regions of the United States and tree-ring chronologies from the North American Drought Atlas (NADA). Minimal differences are found between the three PDSI reconstructions and all compare favorably to independently reconstructed Thornthwaite-based PDSI from the NADA. Reconstructions are bridged with model-derived PDSI_TH and PDSI_PM, both of which closely track modeled soil moisture (near-surface and full column) during the 20th century. Differences between modeled moisture-balance metrics only emerge in 21st-century projections. These differences confirm the tendency of PDSI_TH to overestimate drying when temperatures exceed the range of the normalization interval; the more physical accounting of PDSI_PM compares well with modeled soil moisture in the projection interval. Remaining regional differences in the secular behavior of projected soil moisture and PDSI_PM are interpreted in terms of underlying physical processes and temporal sampling. Results demonstrate the continued utility of PDSI as a metric of surface moisture balance, while additionally providing two recommendations for future work: 1) PDSI_PM (or similar moisture-balance metrics) compare well to modeled soil moisture and are an appropriate means of representing soil-moisture balance in model simulations; and 2) although PDSI_PM is more physically appropriate than PDSI_TH, the latter metric does not bias tree-ring reconstructions of past hydroclimate variability, and as such, reconstructions targeting PDSI_TH can be used with confidence in data-model comparisons. These recommendations and the collective results of our study thus provide a framework for comparing hydroclimate variability within paleoclimatic, observational, and modeled data.

1. Introduction.

The increasing availability of forced-transient simulations over the last millennium (e.g., *Fernández-Donado et al.* 2012; *Masson-Delmotte et al.* 2013) from fully-coupled General Circulation Models (GCMs) has drastically improved our ability to compare paleoclimate reconstructions with model output and to investigate multidecadal to centennial-scale climate dynamics (*Schmidt et al.* 2014). Among the currently available collection of simulations, the Coupled and Paleo Modeling Intercomparison Projects Phases 5 and 3 (CMIP5/PMIP3) have for the first time produced multiple last-millennium, historical, and future simulations using the same model configurations and resolutions (*Taylor et al.* 2012). This important development makes comparisons between paleoclimatic data and last-millennium simulations directly applicable to historical simulations and future projections. Coincident with these advances in modeling efforts, the number of gridded or regional mean proxy reconstructions of multiple climatic variables is also increasing (e.g., *Mann et al.* 2009; *Cook et al.* 2010; *PAGES2k* 2013; *Wahl and Smerdon* 2012; *Neukom et al.* 2010a,b; *Trouet et al.* 2009), as is our understanding of the methods used to perform these reconstructions (e.g. *Jones et al.* 2009; *Smerdon* 2012; *Tingley et al.* 2012), expanding the detail and accuracy with which the actual climate of the last several millennia is characterized. This collective progress opens the possibility that comparisons between paleoclimatic reconstructions and model simulations can be used to improve understanding of decadal to centennial climate dynamics and to constrain model projections of 21st-century climate change in truly direct and quantitative ways (e.g. *Schmidt* 2010; *Ault et al.* 2014).

Many paleoclimate data-model comparison studies are emerging that both interpret comparison results and work to refine the methods by which the comparisons are made (e.g.

47 *Phipps et al.* 2013; *Schmidt et al.* 2014; *Hind et al.* 2012, 2013; *Coats et al.* 2013a,b, 2014;
48 *Anchukaitis et al.* 2010; *Seager et al.* 2008; *Fernández-Donado et al.* 2012; *Ault et al.* 2013a,b;
49 *Lehner et al.* 2012; *Goosse et al.* 2010, 2012; *Sundberg et al.* 2012). This collection of studies,
50 *inter alia*, has highlighted numerous considerations for how comparisons should accommodate
51 the unique strengths, weaknesses and uncertainties of paleoclimatic reconstructions and model
52 simulations. Each new comparison requires attention to the type of proxies used, the climatic
53 variables considered, the means by which different models and model experiments are
54 incorporated, and ultimately how to statistically characterize an ensemble of comparison results.
55 Among the collection of studies, paleoclimate data-model comparisons of hydroclimatic
56 variability over the last millennium (e.g. *Seager et al.* 2008; *Coats et al.* 2013, 2014; *Ault et al.*
57 2013b; *Tierney et al.* 2013) is one emerging and important area of focus, given the critical social
58 and ecological implications of hydroclimate variability and change (e.g. *Allen and Ingram* 2002;
59 *Hoerling et al.* 2012a, 2014; *Ding et al.* 2011; *Headey* 2011; *Li et al.* 2011; *Lobell et al.* 2011;
60 *Peng et al.* 2011; *Seager et al.* 2013). Robust, high-resolution, and gridded drought atlas
61 reconstructions that span most, if not all, of the last millennium and part of the first millennium
62 CE are one critical tool for efforts on these time scales (*Cook et al.* 2007; *Cook et al.* 2010; *Cook*
63 *et al.* 2014). Comparisons between models and drought atlases are nevertheless complicated by
64 the fact that the reconstructions have targeted the Palmer Drought Severity Index (PDSI; *Palmer*
65 1965), an integrated estimate of soil moisture balance that is not a simulated state variable in
66 model integrations. In addition to the fact that model data must therefore be used in offline
67 calculations that estimate PDSI for comparison to hydroclimate reconstructions (emerging
68 capabilities using process-based tree-growth models (e.g. *Anchukaitis et al.* 2006; *Evans et al.*
69 2006) may provide an alternative approach in which model data are used to directly estimate tree

70 growth chronologies that are not calibrated on PDSI or other climate variables), a growing
71 debate has emerged around the efficacy of PDSI as a metric of soil moisture balance in
72 observations and model simulations (e.g. *Guttman* 1998; *Vicente-Serrano et al.* 2010; *Burke et*
73 *al.* 2006, 2008, 2011; *Dai* 2011a,b, 2013; *van der Schrier* et al. 2011, 2013; *Seneviratne* 2012;
74 *Hoerling et al.* 2012b; *Sheffield et al.* 2012; *Trenberth et al.* 2014; *Cook et al.* 2014).

75 Much of the discussion and criticism of PDSI has hinged on its different formulations and
76 more specifically on the means by which potential evapotranspiration (PET) is estimated within
77 the PDSI calculation. A common method for estimating PET, the Thornthwaite formulation
78 (*Thornthwaite* 1948), scales PET as a function of temperature and latitude only, and associated
79 Thornthwaite-based PDSI estimates (hereinafter PDSI_TH) can consequently overestimate
80 drying when temperatures exceed the range of variability spanned by the PDSI normalization
81 interval (*Hoerling et al.* 2012b; *Sheffield et al.* 2012; *Dai* 2013). This effect has been shown to
82 be significant for PDSI_TH estimates at the end of the 20th century in observational records
83 (*Sheffield et al.* 2012) and in the use of PDSI_TH as a moisture balance metric in 21st-century
84 model projections (*Hoerling et al.* 2012b; *Dai* 2013; *Schmidt et al.* 2013). PDSI estimates that
85 use the Penman-Monteith formulation (hereinafter PDSI_PM) for PET alternatively have been
86 proposed as more physically appropriate (e.g. *van der Schrier* et al. 2011), along with additional
87 modifications using self-calibrating PDSI (hereinafter PDSI_SC) that employ both the Penman-
88 Monteith PET formulation and regionally estimated soil and vegetation properties (*Wells and*
89 *Goddard* 2004, *van der Schrier et al.* 2013). Additional metrics, such as the Standardized
90 Precipitation-Evaporation Index, which can incorporate Penman-Monteith estimated PET, also
91 have been developed (e.g. *Vincent-Serrano et al.* 2009; *Hernandez and Uddameri* 2014) and
92 compare well with PDSI_PM in model projections (*Cook et al.* 2014).

Despite the above discussions, PDSI remains a useful metric of soil moisture balance for several reasons. While soil moisture is ultimately the applicable state variable for evaluating model-simulated changes in hydroclimate and drought, the land-surface models in coupled GCMs vary widely in their sophistication (e.g., soil depth, number of layers, etc.), tunings, and parameterizations (e.g., soil texture, rooting depths, vegetation types, etc.), thus complicating the meaningful comparison of direct soil moisture variables across models. Offline metrics like PDSI therefore serve to homogenize accounting of soil moisture balance and rely principally on atmospheric variables in their computation. Well-distributed records of observed soil moisture also are not widely available over many decades, making soil moisture a difficult quantity on which to validate simulated hydroclimate variability over the 20th century. Recent work additionally has demonstrated the benefit of separating PET or PDSI into the constituent influences on these variables (*Scheff and Frierson* 2013; *Cook et al.* 2014). The separation of such influences in coupled GCMs is not easily accomplished for model-simulated soil moisture in which variables such as temperature or precipitation cannot be held constant independent of other coupled variables. With regard to paleoclimatic data-model comparisons specifically, the currently available collection of drought atlases also have used PDSI as the reconstructed target variable, making it necessary to use model-derived PDSI for comparisons between simulations and reconstructions.

Concerns about paleoclimatic drought atlases nevertheless have been raised because tree-ring derived products have traditionally targeted PDSI_TH (e.g. *Cook et al.* 2007, 2010). *Sheffield et al.* (2012) most recently note that: "*Palaeoclimate reconstructions of drought may be particularly susceptible because they are often developed by scaling tree-ring data to match the calculated [PDSI] for their overlap period*." The authors go on to surmise that such

reconstructions may overestimate past changes, while underestimating "*recent trends in the context of the past*." Such concerns are indeed serious, given the degree to which tree-ring derived drought atlases have been used to characterize previous droughts and pluvials (e.g. *Cook et al.* 2007; *Cook et al.* 2010) and the extent to which such information has helped define potential multidecadal risk factors associated with hydroclimatic variability and change.

Herein we evaluate the dependence of regional reconstructions targeting different PDSI formulations using the same dendroclimatic chronologies used in the North American Drought Atlas (NADA) and observationally estimated PDSI_TH, PDSI_PM and PDSI_SC. We derive new reconstructions in three diverse regions of the United States for each of these PDSI formulations, analyze their respective similarities and differences, and compare them to previous NADA estimates of PDSI_TH over the same regions, the latter of which are derived using a different reconstruction method, different predictor processing steps and a spatially-explicit PDSI_TH target on a 0.5°x0.5° latitude-longitude grid (in contrast to the regional indices that are targeted herein). We subsequently turn to bridging paleoclimatic estimates of hydroclimate variability with modeled 20th-century climate simulations and 21st-century climate projections. This challenge is addressed by analyzing model-derived estimates of PDSI_TH, PDSI_PM and two soil moisture estimates from the CanESM2 and CCSM4 GCMs, both of which are available through the CMIP5 data archive. In addition to comparing simulations and reconstructions, we characterize where and how the simulated PDSI and soil moisture estimates agree within each model simulation. It is critically noted that the parameterized impacts of simulated CO_2 fertilization have implications for secular soil moisture trends and may be an important source of uncertainty in comparisons between multiple PDSI and soil moisture variables in 21st-century climate projections. Our investigation culminates with guidance on how to interpret our results

139 as a framework for comparing hydroclimate variability across overlapping observational and
140 modeling intervals and to ultimately use this framework to place future hydroclimate projections
141 into a longer paleoclimatic context.
142
143 **2. Data and Methods.**
144 *a. PDSI Calculations*
145 We focus on three areas of the United States that are representative of diverse hydroclimatic and
146 vegetation regimes: the Four Corners (4C), North Plains (NP) and Southeast (SE) regions as
147 indicated in Figure 1. These areas also contain well-documented and abundant tree-ring
148 chronologies that will be used to derive regional reconstructions. Three different formulations of
149 observational PDSI over the three regions are used: PDSI_TH, PDSI_PM and PDSI_SC. The
150 first two have been calculated using standard formulations (*Thornthwaite* 1948; *Penman* 1948;
151 *Xu and Singh* 2002), while the third is taken from the self-calibrating PDSI dataset of *van der*
152 *Schrier et al.* (2013). PET in PDSI_TH has the advantage of only requiring temperature and
153 latitude data, but is effectively a rescaling of these variables as an estimate of PET (see, for
154 example, the discussion in *van der Schrier et al.* (2011)):

155 $$PET = 16\left(\frac{10T}{I}\right)^{a}, \tag{1}$$

156 where PET is in mm month^{-1}, T is the monthly mean of daily averaged temperature (°C), I is the
157 heat index and a is a third-order polynomial of the heat index (*Thornthwaite* 1948). This
158 expression is scaled in the Thorntwaite calculation of PDSI to account for latitudinal variations
159 in the length of months and days as,

160 $$PET_{TH} = PET(\theta/30)(h/12), \tag{2}$$

161 where θ is the length of the month (in days) and h is the duration of daylight (in hours) on the

fifteenth day of the month (e.g. *van der Schrier et al.* 2011; *Willmott et al.* 1985). The dependence of PET on T in Equation 1 is thus the origin of the excessive drying estimated by PDSI_TH when temperatures are significantly outside the range of variability defined by the baseline normalization period.

The Penman-Monteith PDSI formulation (*Penman* 1948; *Xu and Singh* 2002) has been suggested as a physically appropriate alternative method for calculating PET in 21st-century projections (*Dai* 2013; *Hoerling et al.* 2012; *van der Schrier et al.* 2013; *Sheffield et al.* 2012; *Cook et al.* 2014). The Penman-Monteith formulation is based on surface moisture and energy balance considerations (*Xu and Singh* 2002) and a commonly used formulation is defined by the Food and Agricultural Organization (FAO) of the United Nations (*Allen et al.* 1998). In the FAO formulation PET in mm day^{-1} is given by:

$$PET = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_a + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)}, \quad (3)$$

where Δ is the slope of the vapor pressure curve (kPa °C^{-1}), R_n is the surface net radiation (MJ m^{-2} day^{-1}), G is the soil heat flux density (MJ m^{-2} day^{-1}), γ is the psychometric constant (kPa °C^{-1}), T_a is the air temperature at 2 m (°C), u_2 is the wind speed at 2 m (m s^{-1}), e_s is the saturation vapor pressure (kPa), and e_a is the actual vapor pressure (kPa). Although there have been some suggestions that PDSI_PM also may be susceptible to overestimated drying trends (*Hoerling et al.* 2012b), this has not been widely observed in multiple studies (e.g. *Sheffield et al.* 2012; *van der Schrier et al.* 2011; *Cook et al.* 2014) and reported differences in observationally-based PDSI_PM estimates are more likely related to differences in the precipitation datasets used as inputs for various observational PDSI_PM calculations (*Trenberth et al.* 2014).

For both PDSI_TH and PDSI_PM calculations presented herein, the normalization

interval is 1901-2012 CE. Soil moisture capacities for the top and bottom soil layers are set to the standard values of 25.4 mm (1 in.) and 127 mm (5 in.). Our observational PSDI_PM calculations use the estimate of PET from the updated version (3.21) of the latest CRU TS dataset, which has been derived using the FAO formulation of PET (*Harris et al.* 2013) given in Equation 3. The *Harris et al.* (2013) CRU TS dataset is similarly the source for all other input variables used to calculate PDSI_TH and PDSI_PM. The PDSI_SC dataset of *van der Schrier et al.* (2012) also uses the FAO PET formulation and differs from our calculated version of PDSI_PM only in its incorporation of regionally specific soil and vegetation properties and a snow model (*van der Schrier et al.* 2012); similar to PDSI_PM, the PDSI_SC estimates use climate data from the updated version (3.21) of the latest CRU TS dataset for observational forcing variables (*Harris et al.* 2013).

Area-weighted regional indices were calculated from the 0.5°x0.5° gridded datasets of PDSI_TH, PDSI_PM and PDSI_SC during the summer season (JJA) over the 4C, NP and SE regions; all of the regional PDSI series begin in the year 1901 CE and extend to 2012 CE. These regional series thus comprise the JJA PDSI calibration targets for the reconstructions discussed in the following subsections.

For calculations of PDSI_TH and PDSI_PM from the model simulations over the historical and projection intervals we use the same conventions described above, with several additional applied choices. Modeled PET in the Penman-Monteith formulation is again calculated using the FAO version. Relative to changes in energy availability and the vapor pressure deficit, Penman-Monteith PET is relatively insensitive to near-surface wind speed (*Scheff and Frierson* 2013; *Cook et al.* 2014), which we set equal to a constant 1 m s^{-1} in the PDSI_PM model calculations. Ground heat fluxes similarly are only a small fraction of the total

surface energy budget, typically totaling about 10-15% (*Betts et al.* 1996; *Sellers et al.* 1997). We therefore set the ground heat flux equal to 0 W m^{-2} in the model-derived estimates of PDSI_PM, a choice that has been found to yield little impact (*Cook et al.* 2014).

We do not calculate model-derived estimates of PDSI_SC. Although the PDSI_SC calculation can be done in principle, it additionally includes a snow model and regionally varying vegetation and soil parameters. These considerations further remove the PDSI calculation from model-derived quantities (in the case of the snow model) and involve poorly constrained or unavailable model fields (vegetation and soil properties). In light of these factors and the good agreement that is later shown for PDSI_PM and PDSI_SC, we do not additionally calculate PDSI_SC fields for the models.

b. Tree-Ring Chronologies and the NADA

We use tree-ring width chronologies from the regional collection of states shown in Figure 1. These chronologies are taken from the same network used in the NADA database to derive version 2a of the drought atlas (*Cook et al.* 2007). A total of 283, 36, and 26 chronologies are employed in the 4C, NP and SE regions, respectively. The start and end dates vary across the collection of chronologies in each region, several of which extend to the first century of the first millennium CE, while many only begin in the 18th or 19th centuries. For our purposes herein, we only employ chronologies back to the year 1000 CE; all chronologies extend to at least 1979 CE. These constraints define calibration interval selections and the nesting procedure for the reconstruction method detailed in Section 2c.

The NADA version 2a is also used for comparison to the new regional reconstructions. Version 2a of the NADA is a gridded reconstruction on a 0.5°x0.5° latitude-longitude grid of JJA average PDSI_TH values over much of North America. The grid has been reconstructed using a

point-by-point regression scheme that employs a principal component ordinary least-squares regression method, although multiple steps of predictor processing and ensemble estimation have been included (see *Cook et al.* (2007) and references therein for further details).

c. Reconstruction Method

The new regional reconstructions are derived using a nested 'composite-plus-scale' (CPS) method (e.g. *Jones et al.* 2009) with an ensemble correlation-weighting scheme (*PAGES2k* 2013). 50-year nests were used in each regional reconstruction that employed all chronologies available back to the beginning of each nest period. A CPS reconstruction was computed for each nest by standardizing (normalizing and centering) the available tree-ring chronologies over the calibration interval, and subsequently calculating a weighted composite in which the relative weight of each chronology was determined by the strength of the correlation with the target index. Each composite was finally centered and scaled to have the same mean and variance as the target index during the calibration interval.

The CPS methodology was implemented using a resampling scheme for validation and calibration that uses 50 and 29-year blocks for calibration and validation, respectively (the last year of uniformly available predictor series is 1979 CE, providing 79 years of overlap with the target indices that all start in 1901). The initial calibration period extends from 1901-1950 CE and was incremented by one year until reaching the final period of 1930-1979 CE, yielding a total of 29 reconstructions for each nest. Within each calibration step, the 29 years excluded from calibration were used for cross validation. For each nest, the final CPS reconstruction was taken as the median reconstructed value in each year within the 29-member reconstruction ensemble. Uncertainties were estimated from the mean variance of the residuals across all of the validation intervals; 1.96 times the square root of the estimated mean variance was added to the

253 median ensemble values in each year to derive the 95% confidence intervals. The final nested
254 reconstruction was combined by splicing the median reconstruction and estimated uncertainties
255 of each nest such that every reconstructed year was derived from the nest with the maximum
256 number of chronologies.

257 *d. Model Data*

258 We use publicly available GCM output from the CMIP5 archive, the suite of model experiments
259 organized and contributed in support of Assessment Report Five (AR5) of the Intergovernmental
260 Panel on Climate Change (IPCC). Output from the historical and RCP8.5 21st-century projection
261 experiments is used, the latter of which is the high-emission, business-as-usual scenario that has
262 been justified in many studies by the current lack of international action to limit greenhouse gas
263 emissions. Our analyses are nevertheless not dependent on the employed emissions scenario
264 beyond the fact that the RCP8.5 simulations represent the most extreme warming scenario and
265 therefore embody the maximum impact of temperature changes on PDSI comparisons.
266 Historical CMIP5 experiments are run for the years 1850-2005 CE and are forced with
267 observations of transient climate forcings (e.g. solar variability, land use change, greenhouse gas
268 (GHG) concentrations, etc.). These experiments are initialized in 1850 CE using output from
269 long, unforced control runs with fixed pre-industrial boundary conditions. The RCP8.5
270 projection scenario is one of a suite of future GHG forcing scenarios spanning the 2006-2099 CE
271 period; RCP8.5 is designed so that the top of the atmosphere radiative imbalance will equal
272 approximately +8.5 Wm^{-2} by the end of the 21st century, relative to pre-industrial conditions. The
273 RCP8.5 projections are initialized using the end of the historical runs. Our analysis is restricted
274 to two models (CanESM2 and CCSM4) that have available five continuous ensemble members
275 spanning the historical through projection time intervals. These models and associated ensemble

members also have been selected based on the availability of their layered soil moisture output which extends to ~4 m and ~43 m over 3 and 15 layers in the CanESM2 and CCSM4 models, respectively. For both models, we employ a near surface JJA soil moisture estimate that is taken from approximately the first 30 cm in each model and full-column soil moisture taken over the total depth of each modeled soil column. For comparison to PDSI, both soil moisture variables spanning the historical to projection interval are centered and scaled to match the corresponding PDSI_PM variance from 1901-2005 CE, the same normalization window used to calculate the PDSI variables.

3. Analyses of Target Series and Reconstructions

a. Observational Estimates of PDSI

The area-weighted regional time series for the three observational PDSI estimates are shown for the 4C, NP and SE regions in Figure 1; all time series are centered to have a mean of zero over the reconstruction calibration/validation window (1901-1979 CE). The square of Pearson's correlation coefficient calculated between the series in each region is 0.83 or higher. The largest shared variances (r^2) occur over the SE region (0.96, 0.86, 0.90), while the NP (0.88, 0.85, 0.90) and 4C (0.85, 0.83, 0.92) regions rank in descending order for parings of PDSI_TH vs. PDSI_PM, PDSI_TH vs. PDSI_SC and PDSI_PM vs. PDSI_SC, respectively. These relative comparisons are qualitatively consistent with the correlations between the Thornthwaite and Penman-Monteith versions of PDSI_SC reported by *van der Schrier et al.* (2011) for similar regions in North America.

The standard deviations of the observational PDSI series indicate regionally dependent differences in the variability of the PDSI estimates in the 4C (TH: 1.65, PM: 1.13, SC: 1.07), NP

(TH: 1.84, PM: 1.73, SC: 1.27) and SE (TH: 1.38, PM: 1.50, SC: 1.27) regions. These estimates are consistent with previously reported comparisons between Thornthwaite and Penman-Monteith derived PDSI estimates (e.g. *Sheffield et al.* 2012, *van der Schrier et al.* 2011) in which PDSI_TH displays larger variance relative to the PDSI_PM formulation because of its greater sensitivity to temperature variations. Nevertheless, the results also indicate that in some regions, such as the SE, PDSI_PM has a larger variance than PDSI_TH and therefore all areas do not follow the generally described behavior.

b. Reconstructed Estimates of PDSI

The median PDSI reconstructions and associated uncertainties for each region are compared to the corresponding observational target series in Figure 2. Each reconstruction compares well with the targets, although there are some differences among the collection of validation statistics. The PDSI_TH reconstructions are generally the most skillful across all regions and nests (Figure 3), while the PDSI_SC reconstruction is the second-most skillful in the 4C region and the PDSI_PM reconstruction is second or first most skillful in the NP and SE regions. In all cases, the reconstructions yield validation-interval coefficients of determination over all nests that are skillful above the 99% ($p<0.01$) significance level. Assuming the traditional threshold of zero for the reduction of error (RE) coefficient, all regional reconstructions are skillful across all nests, except for the NP reconstructions, for which the PDSI_TH, PDSI_PM, and PDSI_SC reconstructions are only skillful back to 1500, 1500 and 1600 CE, respectively. By the more stringent coefficient of efficiency (CE) statistic, which also uses a skill threshold of zero, the 4C and SE reconstructions are skillful across all nests, except the PDSI_SC reconstruction in the SE region during several nests in the early part of the millennium (Figure 3). In the NP, all of the reconstructions yield negative CE values prior to 1600 CE and the PDSI_SC yields an additional

322 negative nest from 1650-1700 CE. Although not all nests in the NP region are skillful back to
323 1500 CE across all of the validation metrics, the reconstructed time series are truncated at that
324 year to reflect the bulk validation performance described above.

325 Regional comparisons between each of the three different PDSI reconstructions yield
326 consistent and very similar results, despite the small differences in the validation statistics
327 discussed above. All regional pairings of reconstructions indicate shared variances of 0.99 or
328 higher during the reconstruction intervals (1000-1900 CE for the 4C and SE regions and 1500-
329 1900 CE for the NP). The only notable, though modest, differences between the reconstructions
330 are their standard deviations: 4C (TH: 1.45, PM: 1.05, SC: 1.07), NP (TH: 1.64, PM: 1.39, SC:
331 1.10) and SE (TH: 1.26, PM: 1.39, SC: 1.32). With the exception of the SE, a region in which
332 the three reconstructions are virtually identical, the PDSI_TH reconstructions exceed the
333 standard deviations of the other two reconstructions by 20-50%. These variance differences are
334 further reflected in the right-hand panels of Figure 3, in which the linear relationship between
335 each combination of the reconstructions deviate from the one-to-one line by varying amounts in
336 each region. Given the stated differences between the variances of the target indices, these
337 differences in the standard deviations are expected based on the CPS methodology that matches
338 variances between the composite reconstruction and target indices during the calibration interval.
339 In other words, the standard deviations of each reconstruction reflect the relative differences
340 between the variance of the target time series (although not perfectly so). Importantly, these
341 differences in variance are merely a scaling of the overall reconstructions and do not reveal
342 temporal variance dependencies that would bias interpretations of hydroclimate variability
343 during the 20th century relative to the past.

344 *c. Comparisons between NADA and the CPS Reconstructions*

345 It is worth investigating whether the presented CPS reconstructions are representative of the
346 NADA product, thus making our findings more applicable to interpretations of the latter. In
347 Figure 4 we compare the derived CPS PDSI_TH reconstructions to area-weighted regional
348 averages from the NADA that have been sampled from gridded regions approximating the CPS
349 reconstruction areas shown in Figure 1 (see Figures 11 and 12 for the specific sample regions
350 extracted from the NADA grid). We limit the comparisons in Figure 4 to PDSI_TH, given that
351 the NADA also has targeted PDSI_TH. The shared variances between the NADA time series
352 and the CPS reconstructions of PDSI_PM (4C: 0.94, NP: 0.65, SE: 0.68) and PDSI_SC (4C:
353 0.94, NP: 0.65, SE: 0.68) are nevertheless virtually identical to the shared variances reported for
354 PDSI_TH in Figure 4.

355 Comparisons between the NADA regional averages and the CPS PDSI_TH
356 reconstructions are consistent, despite the different methods by which the reconstructions were
357 derived. The CPS reconstructions target a single regional average and use a sliding
358 calibration/validation scheme to weight chronologies based on correlation. The NADA index is
359 the regional average of a gridded product that uses principal component regression to target
360 PDSI values at each grid point on a 0.5°x0.5° spatial grid, uses a static calibration/validation
361 interval, and involves various predictor processing steps not included in the CPS methodology
362 used herein. The reconstructions therefore have targeted different characteristics of the
363 instrumental PDSI fields and used contrasting methods to derive the respective reconstructions.
364 Despite these dissimilarities, the CPS PDSI_TH and NADA reconstructions compare extremely
365 well (Figure 4); the reported shared variances for PDSI_PM and PDSI_SC similarly support a
366 strong agreement between the NADA and the two additional sets of regional reconstructions.
367 These results not only support the idea that analyses of the new CPS reconstructions are relevant

368 to interpretations of the NADA, but they more generally support the robustness and consistency
369 of the employed methods and derived reconstructions.
370
371 **4. Bridging Paleoclimatic, Observational and Model Data**
372 We now turn specifically to the challenge of bridging PDSI reconstructions with historical model
373 simulations and future projections. We do not take up the related and important endeavor of
374 comparing last-millennium PDSI reconstructions with forced-transient model simulations
375 targeting the same interval. Such comparisons involve additional considerations regarding the
376 interpretation of internal and forced variability and the realism with which the reconstructed
377 exogenous forcings are estimated and employed in the last-millennium simulations. Detailed
378 examples of such work can be found in *Coats et al.* (2013a, 2014, 2015), who specifically
379 perform paleoclimate data-model comparisons focused on hydroclimate in North America.
380 Alternative to these specific comparisons of data and models over the last millennium, we
381 endeavor to address how to bridge PDSI reconstructions with historical and future climate
382 projections in the context of the noted differences between multiple PDSI formulations. We
383 analyze four estimates of modeled soil moisture balance: PDSI_TH, PDSI_PM, soil moisture in
384 approximately the first 30 cm of the soil column (30cmSM), and full-column soil moisture
385 (FCSM), the latter two of which are centered and scaled to match the PDSI_PM variance during
386 the 1901-2012 CE interval for comparison.
387 *a. Historical Interval*
388 The four characterizations of modeled soil moisture balance in each region over the historical
389 interval (1901-2005 CE) are shown in Figure 5 for the CanESM2 and CCSM4 models.
390 Consistent with the observation-based estimates, these four measures of soil moisture balance

391 yield internally consistent results over the 20th century, and both formulations of PDSI reproduce
392 modeled soil moisture with high fidelity. Agreement between the two PDSI calculations is
393 expected given that they are both calibrated over the 1901-2012 CE period and therefore any
394 unrealistic temperature-driven differences in the PDSI_TH calculation are minimized. PDSI_TH
395 and PDSI_PM in fact reveal a large amount of shared variance over the historical interval,
396 matching or exceeding r^2 values of 0.9 in all regions over all ensemble members in both models
397 (Figures 6 and 7).

398 Comparisons between the two PDSI estimates and modeled soil moisture during the
399 historical interval indicate weaker, but still high levels of shared variance (Figures 6 and 7). In
400 most cases, PDSI_TH and PDSI_PM compare best with 30cmSM. The shared variance is
401 strongest in CCSM4 for the 4C and SE regions, where r^2 values exceed 0.7 for comparisons
402 between the two PDSI variables and 30cmSM. These numbers reduce slightly for the NP where
403 some ensemble members yield values between 0.5 and 0.6. Comparisons weaken in CanESM2,
404 in which values of shared variance between the two PDSI variables and 30cmSM range between
405 about 0.5 and 0.75 across all regions. Comparisons between the two PDSI variables and FCSM
406 generally indicate less shared variance than with 30cmSM because FCSM incorporates longer-
407 scale variations and time lags that exceed the timescales that PDSI and near-surface soil moisture
408 more strongly sample. The depth of sampling is not, however, the only factor, as indicated by
409 the fact that shared variances between PDSI estimates and FCSM are larger than for 30cmSM in
410 the SE region in the CanESM2 model and for some variable pairings and ensemble members in
411 both models in the NP. Moreover, FCSM spans a much greater depth in the CCSM4 model than
412 in CanESM2, but the r^2 values between the PDSI variables and FCSM are comparable or larger
413 in CCSM4 than in CanESM2. Depth effects are therefore not the only determining factor in the

414 comparison.

415 Comparisons between the two direct soil moisture estimates (30cmSM and FCSM) are
416 similar or worse within each model than the comparisons between the two PDSI estimates and
417 either of the soil moisture variables, the one exception being in the CanESM2 model in the SE.
418 This highlights the fact that it is not straightforward to determine which metric is most
419 appropriate as a measure of modeled soil moisture variability or even which soil moisture target
420 is the most appropriate analogue to compare against PDSI. Even within models, the agreement
421 between direct soil moisture metrics is variable and depends strongly on the selected regions.
422 The differences among these variables are likely associated with the specific tunings,
423 parameterizations, and other modeling choices in the land-surface components of the GCMs, and
424 further highlights the challenges of even direct comparisons between soil moisture variables in
425 the current generation of coupled GCMs.

426 We finally note that we avoid direct comparisons between the regional PDSI_TH
427 reconstructions and the associated model variables beyond the collective plotting of these
428 variables in Figures 8-10. We have normalized and centered all of the moisture-balance
429 variables over the common reference interval from 1901-2005 CE, leaving any consistencies in
430 means and variances of the reconstructed and modeled variables over the reference interval a
431 product of construction. Furthermore, shared variance between the reconstructed and modeled
432 moisture-balance metrics should not be expected for the CMIP5 historical model runs, in which
433 the simulations are initialized from pre-industrial control runs. These initializations do not
434 constrain modes of internal variability that impact hydroclimate over North America, such as the
435 El Niño-Southern Oscillation or the Pacific Decadal Oscillation, to be in phase with those that
436 have dictated the actual climate states represented in the reconstructions. Unless all of the

437 hydroclimate variability in the analyzed regions were forced over the 20th century, which is
438 certainly not the case, the interannual to decadal variance in the reconstructions and the models
439 will not share consistently common features. These considerations therefore must inform
440 attempts to compare reconstructions and model simulations directly over the interval of overlap
441 and require more detailed and specific analyses that are beyond the scope of this investigation.
442 The bridging approach that we have outlined herein nevertheless demonstrates the use of the 20th
443 century as a common interval for referencing reconstruction and model data in order to compare
444 earlier paleoclimatic intervals and model projections of the 21st century.

445 *b. Projection Interval*

446 Shared variance between the four soil moisture variables during the projection interval (2005-
447 2099 CE) are given in Figures 6 and 7, while Figures 8-10 provide comparisons between the four
448 soil moisture metrics for the first ensemble member from both models over the historical through
449 projection intervals. Shared variances for detrended time series over the projection interval, in
450 addition to the original time series, are also shown in Figures 6 and 7. The shared variances
451 among all of the detrended soil moisture metrics in the projection interval in all regions are
452 generally comparable to those of the historical interval (with some exceptions particularly in the
453 CCSM4 output). Larger differences are observed between the original time series, in which
454 differences in 21st-century trends reduce the r^2 values among the different soil moisture metrics
455 (in cases where trends are comparable among variables, the shared variances in original and
456 detrended time series are similar). Secular drying in PDSI_TH exceeds that of all other modeled
457 soil moisture metrics over the 21st century in all three regions, a clear demonstration of the
458 tendency for PDSI_TH to overestimate drying during the projection interval. There are,
459 however, regional differences in the relative comparisons among variables, as discussed below.

460 PDSI_TH includes secular trends in the 4C region that in some cases are larger in
461 magnitude or opposite in sign relative to the three other metrics in both models (Figure 8),
462 despite the fact that some of the other metrics compare favorably to PSDI_TH over the full
463 century or during specific intervals of time. PDSI_PM and 30cmSM compare well throughout
464 the projection interval in the CCSM4 model, while 30cmSM projects enhanced drying, relative
465 to PDSI_PM, in the CanESM2 model. FCSM again compares favorably to PDSI_PM and
466 30cmSM in the CCSM4 simulation, while it suggests a wetting trend in the CanESM2 simulation
467 that is not reflected in any of the other estimates. This may again be representative of the
468 deeper-column soil moisture dynamics in the model (although no such trend is observed for the
469 CCSM4 model with a much deeper soil column).
470 In the NP, the behavior of PDSI_TH is again consistent with previously discussed
471 expectations (Figure 9). The variable estimates drier 21st-century mean conditions than any of
472 the other soil moisture metrics and projects secular drying trends in both models that are either
473 larger or not present in the other variables. Similar to the 4C region, PDSI_PM compares
474 favorably to 30cmSM throughout the projection interval in the CCSM4 model, while 30cmSM
475 projects slightly enhanced drying, relative to PDSI_PM, in the CanESM2 model. The FCSM
476 estimate in CanESM2 contains a pronounced wetting trend and multidecadal variability in the
477 NP that, in addition to contrasting with PDSI_PM, is also not present in the 30cmSM output.
478 The FCSM is similarly wetter in the CCSM4 projection for the NP, but not by nearly as much as
479 the CanESM2 projection. In the SE, the behavior of PDSI_TH is the same as discussed for the
480 other two regions in both models (Figure 10). Comparisons between 30cmSM and FCSM
481 variables are more consistent in both models in the SE, both of which project wetter conditions
482 or a wetting trend (CCSM4) in the SE that is not present in the PDSI_PM estimate.

483 Given the three regional analyses presented above, it is evident that all of the soil
484 moisture metrics compare well over the 20th century in both models over all three regions.
485 Regional differences are evident, however, in comparisons between the variables over the 21st-
486 century, including some large disparities between the two modeled soil moisture metrics
487 themselves. This is particularly true of the secular trends estimated by the various metrics.
488 While PDSI_TH projects varying degrees of exaggerated drying over all regions and models,
489 PDSI_PM either matches well 30cmSM, falls between 30cmSM and FCSM, or projects
490 relatively constant moisture balance conditions when 30cmSM and FCSM indicate wetting
491 trends. These regional observations are given broader context in Figures 11 and 12 that plot the
492 projected mean values of PDSI_PM, 30cmSM and FCSM for the last two decades (2080-2099
493 CE) of the 21st-century using the first ensemble member of the CanESM2 and CCSM4
494 simulation; the agreement between the three variables in the direction of either wetting or drying
495 is also plotted. Inspections of Figures 11 and 12 indicate that it is difficult to draw consistent
496 characterizations of how the variables compare across models or regions. In the CanESM2
497 simulation, for instance, PDSI_PM and FCSM project wetting throughout much of western
498 North America, while 30cmSM projects drying over those regions (note that this wetting
499 projection is anomalous over the CMIP5 ensemble (*Cook et al.* 2014), but the analyzed models
500 have been selected based on the availability of their layered soil moisture output and ensemble
501 members that are continuous across the historical to projection intervals). In contrast to
502 CanESM2, the PDSI_PM projection from CCSM4 is almost uniformly toward drying, while a
503 heterogeneous pattern of wetting emerges in 30cmSM and is further enhanced in FCSM. A full
504 diagnosis of the physical underpinnings of these differences is beyond the scope of this paper,
505 but we discuss various explanations for the regional differences in the following section and

suggest that parsing the differences between these soil moisture metrics is an important area of further research.

5. Discussion

This study has been motivated by two considerations: 1) the growing interest in performing paleoclimatic model-data comparisons; and 2) the ongoing debate about how to properly represent hydroclimatic variability and change from a host of possible soil moisture metrics. Our results demonstrate the robustness of PDSI as a metric of near-surface moisture variability in observations, reconstructions, and 20th-century model simulations. Reconstructions derived for three diverse regions of the United States compare favorably, regardless of the reconstruction target (PDSI-TH, PDSI-PM, or PDSI-SC) or technique (CPS vs. the NADA-based approach). The variances of the derived reconstructions are the principal, though modest, differences among them, and are consistent with the expected character of PDSI formulations estimated from observations over the historical interval. Importantly, these results indicate that previous concerns about biases in tree-ring reconstructions due to the use of PDSI_TH as a calibration target (*Sheffield et al.* 2012) are unfounded over North America. Similarly, model estimates of soil moisture and PDSI are all consistent during the historical interval: PDSI_TH, PDSI_PM and two normalized soil moisture estimates all compare well in the CanESM2 and CCSM4 historical model simulations. Together, these results support the continued use of PDSI as a valuable tool for empirical and model-based investigations of drought and hydroclimate.

Principal differences emerge only in model-derived estimates of PDSI and soil moisture during the 21st-century projection interval. Specifically, the secular behavior of the various metrics diverges in the projections, while the interannual variability remains relatively consistent

529 across all variables throughout the 21st-century projection interval. In the case of PDSI_TH,
530 enhanced secular drying is now a well-understood consequence of unrealistically scaling PET as
531 a function of temperature when values exceed the range defined by the normalization interval. It
532 is less clear, however, how and why modeled near-surface and full-column soil moisture would
533 compare well with the more physically-based PDSI_PM during the 20th century and in the 21st-
534 century on interannual timescales, while their 21st-century secular trends would diverge.

535 Several possibilities may explain some of the different secular behavior in the model-
536 based metrics. The differences may arise, in part, from the difficulty in identifying the most
537 appropriate modeled soil moisture variable for comparison with the model-derived PDSI. We
538 have demonstrated various situations in which PDSI most closely reflects near surface soil
539 moisture, others where PDSI was a better indicator of full column soil moisture, or places where
540 PDSI resolved both equally well. Importantly, we also illustrated ample instances where even
541 the two soil moisture metrics diverge, sometimes to quite extreme effect (e.g., the Northern
542 Plains in CanESM2). Additional differences between PDSI and modeled soil moisture also are
543 likely to arise through the typically more sophisticated treatment of processes (e.g., vegetation,
544 soil physics, etc.) in the land-surface components of GCMs. These parameterizations and
545 tunings vary across models, meaning that any concerns in comparing PDSI and soil moisture
546 within models must also be extended to comparing soil moisture trends across models.

547 One major issue that is often discussed in the context of future drought projections is the
548 CO_2 fertilization effect. The direct physiological effect of elevated atmospheric CO_2
549 concentrations (eCO2) is to reduce water loss during photosynthesis by lowering stomatal
550 conductance. This effect is typically incorporated into the physics of GCM land-surface and
551 vegetation models, but not in standard PDSI calculations, including those employed herein. Two

552 expectations are therefore associated with eCO2: 1) drought stress on model vegetation will be
553 reduced, translating into lower rates of evapotranspiration and increases in soil moisture and
554 runoff; and 2) standard PDSI projections, which do not include the impact of eCO2, will not
555 reflect any associated wetting tied to the effect. This may explain areas in the presented
556 projections where PDSI indicates drying while model soil moisture indicates wetting or little
557 change.

558 While the physiological effect of eCO2 is well understood at the leaf level (*Ainsworth*
559 *and Rogers* 2007), there are large uncertainties associated with the scaling of this effect on
560 hydrology at larger scales that may lead to an overestimate of the eCO2 effect in model
561 projections. For example, many experiments in which plants are exposed to elevated levels of
562 CO_2 show only modest and often insignificant changes in transpiration and soil moisture (e.g.,
563 *Domec et al.* 2009; *Hussain et al.* 2013; *Inauen et al.* 2013; *Naudts et al.* 2013; *Stocker et al.*
564 1997). Additionally, few eCO_2 studies simultaneously incorporate ambient warming into their
565 experiments, an accompanying element of climate change that is expected to increase ET and
566 potentially counteract any moisture gains from eCO2. Despite this ambiguity in the empirical
567 evidence, however, most land surface and vegetation models, including those models typically
568 incorporated into GCMs, substantially reduce transpiration when exposed to eCO_2 and, in some
569 cases, dramatically overestimate the transpiration response relative to observations (*De Kauwe et*
570 *al.* 2013). This suggests that the modeled response to eCO_2 may be oversimplified and the eCO_2
571 effect on hydrology overestimated. In fact, this host of uncertainties in both the modeled and
572 empirical eCO2 responses led Working Group II of the IPCC to conclude in AR5 (*Scholes et al*.
573 2014) that the net effect of eCO_2 on runoff and transpiration is still "*poorly constrained*", and
574 that "*precipitation and temperature effects are likely to remain the prime influence on global*

575 *runoff*." Notably, eCO_2 does not appear to have a strong influence on soil moisture in other
576 CMIP5 simulations before the rapid increases in the RCP scenarios, as evidenced by the tight
577 coupling between PDSI and modeled soil moisture in the historical simulations presented herein.
578 Clearly, there remain significant uncertainties in the interpretation of modeled soil
579 moisture and PDSI responses in 21st-century GCM projections. Further characterizing and
580 resolving the impact of these modeling choices will be important as a means of resolving
581 differences between PDSI_PM and the soil moisture estimates discussed herein, but even more
582 importantly to constrain the range of realistic soil moisture conditions into the future. In the
583 context of paleoclimate data-model comparisons, however, it appears that projections using
584 PDSI_PM or normalized near-surface soil moisture are the most appropriate variables for
585 characterizing future projections and for comparing projections to modern and paleoclimatic
586 observations (e.g. *Dai* 2013; *Cook et al.* 2014).
587
588 **6. Conclusions**
589 Given our principal observations, a framework emerges for comparisons between PDSI
590 reconstructions and model simulations. It is first of all evident that tree-ring based PDSI
591 reconstructions are robust to the selection of PDSI target variables and the current generation of
592 PDSI_TH drought atlases can be confidently used as estimates of hydroclimatic variability over
593 the last millennium. With regard to comparing these drought atlases to last-millennium
594 simulations, our results suggests that model-derived PDSI or soil moisture should all estimate
595 similar hydroclimate histories given that modeled temperature variations over the last
596 millennium are typically within the range of modeled 20th-century climate (e.g. *Fernández-*
597 *Donado et al.* 2012; *Masson-Delmotte et al.* 2013). Indeed, the similarity between multiple

598 hydroclimate metrics over the last millennium has been demonstrated specifically in the
599 American Southwest using multiple simulations in the context of assessing the occurrence of
600 multidecadal drought events in the region (*Coats et al.* 2013, 2014). Similarly, we have shown
601 that comparisons among observational or modeled soil moisture metrics during the 20th century
602 are largely insensitive to the choice of metric. While it is most consistent to compare the same
603 metric across reconstructions, observations, and models, such consistency does not appear
604 essential, and is sometimes not possible, over the last millennium and through the 20th century.
605 In lieu of homogenized products, it therefore is suggested that comparisons simply reference and
606 scale means and variances of all variables over a common period of overlap.

607 Overall, the current collection of tools can provide robust characterizations of
608 hydroclimate variability and change during the last millennium, which in turn can be
609 meaningfully compared to observations and model simulations of 20th-century hydroclimate and
610 model projections of 21st-century change. Within these comparisons the most important interval
611 of consideration is the 21st century, in which model projections should be characterized using
612 either PDSI_PM (or similarly a Penman-Montieth version of PDSI_SC or Standardized
613 Precipitation-Evaporation Index) or a near-surface normalized soil moisture output. In the case
614 of the latter, however, these results must be carefully interpreted in terms of their applicability
615 across a collection of model simulations and in terms of the processes that they include. The
616 impact of CO_2 fertilization on 21st-century hydroclimate projections is a particularly important
617 process to evaluate across models, given the poorly constrained nature of the process on large-
618 scale vegetation response and because it appears to have potentially important implications for
619 projected hydroclimatic trends. Although these details complicate comparisons and leave open
620 important research questions, future model-data comparisons that span the last-millennium

621 through the 21st-century projection interval will be vital for assessing and characterizing future
622 risks associated with hydroclimate variability and change. As we have demonstrated, these
623 comparisons are possible and appropriate given the current collection of data and tools, all of
624 which should be used for comparisons within the framework outlined herein.
625
626 **Acknowledgments.** We acknowledge the World Climate Research Programme's Working
627 Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate
628 modeling groups for producing and making available their model output. For CMIP, the U.S.
629 Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides
630 coordinating support and leads development of software infrastructure in partnership with the
631 Global Organization for Earth System Science Portals. Haibo Liu and Naomi Henderson
632 provided computational and data processing support at LDEO. RS and JES were supported in
633 part by the NOAA award Global Decadal Hydroclimate Variability and Change
634 (NA10OAR4310137). RS was also supported by DOE award DE-SC0005107. Further support
635 came from NSF award AGS-1243204. BIC was supported by NASA. LDEO publication
636 number #XXXX.

References

Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**(6903):224-232, DOI http://dx.doi.org/10.1038/nature01092.

Allen, R. G., L. S. Pereira, D. Raes, M. Smith, 1998: *Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements*. FAO Irrigation and drainage paper 56. Rome, Italy: Food and Agriculture Organization of the United Nations. ISBN 92-5-104219-5.

Anchukaitis, K. J., B. M. Buckley, E. R. Cook, B. I. Cook, R. D. D'Arrigo, and C. M. Ammann, 2010: Influence of volcanic eruptions on the climate of the Asian monsoon region. *Geophys. Res. Lett.*, **37**, L22703, doi:10.1029/2010GL044843.

Anchukaitis, K.J., M.N. Evans, A. Kaplan, E.A. Vaganov, H.D. Grissino-Mayer, M.K. Hughes, and M.A. Cane, 2006: Forward modeling of regional-scale tree-ring patterns in the southeastern United States and the recent influence of summer drought, *Geophys. Res. Lett.*, **33**, L04705, doi:10.1029/2005GL025050.

Ainsworth, E. A., A. Rogers, 2007: The response of photosynthesis and stomatal conductance to rising [CO2]: mechanisms and environmental interactions. *Plant, Cell and Environment*, **30**, 258-270.

Ault, T. R., C. Deser, M. Newman, and J. Emile-Geay, 2013a: Characterizing decadal to centennial variability in the equatorial Pacific during the last millennium, *Geophys. Res. Lett.*, **40**, 3450–3456, doi:10.1002/grl.50647.

Ault, T. R., J. E. Cole, J. T. Overpeck, G. T. Pederson, S. St. George, B. Otto-Bliesner, C. A. Woodhouse, C. Deser, 2013b: The Continuum of Hydroclimate Variability in Western North America during the Last Millennium. *J. Climate*, **26**, 5863-5878. doi:

http://dx.doi.org/10.1175/JCLI-D-11-00732.1

Ault, T., J. Cole, J. Overpeck, G. Pederson, and D. Meko, 2014: Assessing the risk of persistent drought using climate model simulations and paleoclimate data, *J. Climate*. doi:10.1175/JCLI-D-12-00282.1, in press.

Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. A. Viterbo, 1996: The land surface-atmosphere interaction: A review based on observational and global modeling perspectives. *J. Geophys. Res.*, **101**(D3):7209-7225, DOI 10.1029/95JD02135.

Burke, E. J., 2011: Understanding the Sensitivity of Different Drought Metrics to the Drivers of Drought under Increased Atmospheric CO2. *J. Hydromet.*, **12**(6):1378-1394, DOI 10.1175/2011JHM1386.1

Burke, E. J., and S. J. Brown, 2008: Evaluating Uncertainties in the Projection of Future Drought. *J. Hydromet.,* **9**(2):292-299, DOI 10.1175/2007JHM929.1

Burke, E. J., S. J. Brown, and N. Christidis, 2006: Modeling the recent evolution of global drought and projections for the twenty-first century with the Hadley Centre climate model. *J. Hydromet.,* **7**(5):1113-1125, DOI http://dx.doi.org/10.1175/JHM544.1

Cao, L., G. Bala, K. Caldeira, R. Nemani, G. Ban-Weiss, 2010: Importance of carbon dioxide physiological forcing to future climate change. *Proc. Nat. Acad. Sci.,* **107**, 9513-9518, DOI 10.1073/pnas.0913000107

Coats, S., J. E. Smerdon, R. Seager, B. I. Cook, and J. F. González-Rouco, 2013a: Megadroughts in Southwestern North America in ECHO-G Millennial Simulations and Their Comparison to Proxy Drought Reconstructions. *J. Clim.*, **26**, 7635-7649, DOI doi: 10.1175/JCLI-D-12-00603.1

Coats, S., J.E. Smerdon, B.I. Cook, and R. Seager, 2013b: Stationarity of the Tropical Pacific

Teleconnection to North America in CMIP5/PMIP3 Model Simulations. *Geophys. Res. Lett.*, **40**, doil:10.1002/grl.50938.

Coats, S., J.E. Smerdon, B.I. Cook and R. Seager, 2014: Are Simulated Megadroughts in the North American Southwest Forced? *J. Clim.*, in press.

Coats, S., B.I. Cook, J.E. Smerdon and R. Seager, 2015: North American Pan-Continental Droughts in Model Simulations of the Last Millennium *J. Clim.*, in review.

Cook, E. R., R. Seager, M. A. Cane, and D. W. Stahle, 2007: North American drought: reconstructions, causes and consequences. *Earth Sci. Rev.*, **81**, 93-134.

Cook, E. R., R. Seager, R. R. Heim Jr., R. S. Vose, C. Herweijer, and C. Woodhouse, 2010: Megadroughts in North America: placing IPCC projections of hydroclimatic change in a long-term palaeoclimate context. *J. Quat. Sci.*, **25**(1), 48-61, DOI 10.1002/jqs.1303.

Cook, B.I., J.E. Smerdon, R. Seager and E.R. Cook, 2014: Pan-continental droughts in North America over the last millennium, *J. Clim.*, **27**, 383-397, doi: http://dx.doi.org/10.1175/JCLI-D-13-00100.1.

Cook, B.I., J.E. Smerdon, R. Seager, and S. Coats, 2014: Global warming and 21st century drying, *Clim. Dyn.*, doi: 10.1007/s00382-014-2075-y.

Dai, A., 2011a: Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900-2008. *J. Geophys. Res.*, **116**, D12115, DOI 10.1029/2010JD015541

Dai, A., 2011b: Drought under global warming: a review. *Wiley Interdisciplinary Rev.: Clim. Ch.*, **2**(1), 45-65, DOI 10.1002/wcc.81

Dai, A., 2013: Increasing drought under global warming in observations and models. *Nature Clim. Ch.,* **3**(1), 52-58, DOI 10.1038/nclimate1633.

De Kauwe, M. G., B. E. Medlyn, S. Zaehle, A. P. Walker, M. C. Dietze, T. Hickler, A. K. Jain,

Y. Luo, W. J. Parton, I. C. Prentice, B. Smith, P. E. Thornton, S. Wang, Y. -P. Wang, D. Wårlind, E. Weng, K. Y. Crous, D. S. Ellsworth, P. J. Hanson, H.-. Seok Kim, J. M. Warren, R. Oren, and R. J. Norby, 2013: Forest water use and water use efficiency at elevated CO2: a model-data intercomparison at two contrasting temperate forest FACE sites. *Glob. Ch. Biol.*, **19**, 1759-1779. doi: 10.1111/gcb.12164

Ding, Y., M. J. Hayes, and M. Widhalm, 2011: Measuring economic impacts of drought: a review and discussion. *Disaster Prev. Man.* **20**(4), 434-446, DOI 10.1108/09653561111161752

Domec, J. C., S. Palmroth, E. Ward, C. A. Maier, M. Therezien, and R. Oren, 2009: Acclimation of leaf hydraulic conductance and stomatal conductance of Pinus taeda (loblolly pine) to long-term growth in elevated CO2 (free-air CO2 enrichment) and N-fertilization. *Plant Cell Env.*, 32(11):1500-1512, DOI 10.1111/j.1365-3040.2009.02014.x

Evans, M.N., B.K. Reichert, A. Kaplan, K.J. Anchukaitis, E.A. Vaganov, M.K. Hughes and M.A. Cane, 2006: A forward modeling approach to paleoclimatic interpretation of tree-ring data, *J. Geophys. Res.*, **111**, G03008, doi:10.1029/2006JG000166.

Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou, and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium, *Clim. Past*, **9**, 393-421, doi:10.5194/cp-9-393-2013.

Gedney, N., P. M. Cox, R. A. Betts, O. Boucher, C. Huntingford, and P. A. Stott, 2006: Detection of a direct carbon dioxide effect in continental river runoff records. *Nature*, **439**, 835-838.

Goosse, H., E. Crespin, A. de Montety, M. E. Mann, H. Renssen, and A. Timmermann, 2010: Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation. *J. Geophys. Res.*, **115**, D09108, doi:10.1029/2009JD012737.

Goosse, H., E. Crespin, S. Dubinkina, M.-F. Loutre, M. E. Mann, H. Renssen, Y. Sallaz-Damaz, and D. T. Shindell, 2012: The role of forcing and internal dynamics in explaining the "Medieval Climate Anomaly". *Clim. Dyn.*, **39**, 2847-2866.

Guttman, N. B., 1998: Comparing the Palmer Drought Index and the Standardized Precipitation Index. *J. Am. Water Res. Assoc.*, **34**, 113-121, DOI 10.1111/j.1752-1688.1998.tb05964.x

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2013: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Clim.*, **34**, 623-642

Headey, D., 2011: Rethinking the global food crisis: The role of trade shocks. *Food Policy,* **36**(2), 136-146, DOI http://dx.doi.org/10.1016/j.foodpol.2010.10.003

Hernandez, E., and V. Uddameri, 2014: Standardized precipitation evaporation index (SPEI)-based drought assessment in semi-arid south Texas. *Env. Earth Sci.*, **71**(6), 2491-2501, DOI 10.1007/s12665-013-2897-7

Hind, A., A. Moberg, and R. Sundberg, 2012: Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium – Part 2: A pseudo-proxy study addressing the amplitude of solar forcing. *Clim. Past*, **8**, 1355-1365, doi:10.5194/cp-8-1355-2012

Hind, A., A. Moberg, 2013: Past millennial solar forcing magnitude. A statistical hemispheric-scale climate model versus proxy data comparison. *Clim. Dyn.*, **41**, 2527-2537, doi:10.1007/s00382-012-1526-6

Hoerling, M., A. Kumar, R. Dole, J. W. Nielsen-Gammon, J. Eischeid, J. Perlwitz, X. W. Quan, T. Zhang, P. Pegion, and M. Chen, 2012a: Anatomy of an Extreme Event. *J. Clim.*, **26**(9), 2811-2832, DOI 10.1175/JCLI-D-12-00270.1

Hoerling, M. P., J. K. Eischeid, X. W. Quan, H. F. Diaz, R. S. Webb, R. M. Dole, and D. R. Easterling, 2012b: Is a Transition to Semipermanent Drought Conditions Imminent in the U.S. Great Plains? *J. Clim.*, **25**(24), 8380-8386, DOI 10.1175/JCLID-12-00449.1

Hoerling, M., J. Eischeid, A. Kumar, R. Leung, A. Mariotti, K. Mo, S. Schubert, and R. Seager, 2014: Causes and Predictability of the 2012 Great Plains Drought. *Bull. Am. Met. Soc.*, in press

Hussain, M. Z., A. VanLoocke, M. H. Siebers, U. M. Ruiz-Vera, R. J. Cody Markelz, A. D. B. Leakey, D. R. Ort, C. J. Bernacchi, 2013: Future carbon dioxide concentration decreases canopy evapotranspiration and soil water depletion by field-grown maize. *Glob. Ch. Biol.*, **19**(5), 1572-1584, DOI 10.1111/gcb.12155

Inauen, N., C. Körner, and E. Hiltbrunner, 2013: Hydrological consequences of declining land use and elevated CO2 in alpine grassland. *J. Ecology*, **101**, 86-96. doi: 10.1111/1365-2745.12029

Jones, P. D. et al. High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects. *Holocene,* **19**, 3-49 (2009).

Lehner, F., C. C. Raible, and T. F. Stocker, 2012: Testing the robustness of a precipitation proxy-based North Atlantic Oscillation reconstruction. *Quat. Sci. Rev.*, 45, 85-94

Li, X., S. R. Waddington, J. Dixon, A. K. Joshi, and M. C. Vicente, 2011: The relative importance of drought and other water-related constraints for major food crops in South Asian farming systems. *Food Security*, **3**(1), 19-33, DOI 10.1007/s12571-011-0111-x

Lobell, D. B., W. Schlenker, and J. Costa-Roberts, 2011: Climate Trends and Global Crop Production Since 1980. *Science,* **333**(6042), 616-620, DOI 10.1126/science.1204531

Mann, M. E., Z. Zhang, S. Rutherford, R. S. Bradley, M. K. Hughes, D. Shindell, C. Ammann, G. Faluvegi, and F. Ni, 2009: Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly, *Science*, **326**, 1256-1260, doi:10.1126/science.1177303.

Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, A. Ganopolski, J.F. González Rouco, E. Jansen, K. Lambeck, J. Luterbacher, T. Naish, T. Osborn, B. Otto-Bliesner, T. Quinn, R. Ramesh, M. Rojas, X. Shao and A. Timmermann, 2013: Information from Paleoclimate Archives. In: *Climate Change 2013: The Physical Science Basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Naudts, K., J. Berge, I. Janssens, I. Nijs, and R. Ceulemans, 2013: Combined effects of warming and elevated CO2 on the impact of drought in grassland species. *Plant and Soil*, **369**(1-2), 497-507, DOI 10.1007/s11104-013-1595-2.

Neukom, R., J. Luterbacher, R. Villalba, M. Küttel, D. Frank, P. D. Jones, M. Grosjean, J. Esper, L. Lopez, and H. Wanner, 2010a: Multi-centennial summer and winter precipitation variability in southern South America, *Geophys. Res. Lett.*, **37**, L14708, doi:10.1029/2010GL043680.

Neukom, R., J. Luterbacher, R. Villalba, M. Küttel, D. Frank, P. D. Jones, M. Grosjean, H. Wanner, J.-C. Aravena, D. E. Black, D. A. Christie, R. D'Arrigo, A. Lara, M. Morales, C. Soliz-Gamboa, A. Srur, R. Urrutia, L. von Gunten, 2010b: Multiproxy summer and winter

surface air temperature field reconstructions for southern South America covering the past centuries, *Clim. Dyn.*, **37**(1-2), 35-51 doi:10.1007/s00382-010-0793-3.

PAGES 2k Consortium, 2013: Continental-scale temperature variability over the last two millennia, *Nature Geo.*, **6**, 339-346, doi:10.1038/ngeo1797.

Palmer, W.C., 1965: Meteorological drought. Research Paper No. 45. U.S. Weather Bureau, 1-58

Peng, S., Q. Tang, and Y. Zou, 2009: Current Status and Challenges of Rice Production in China. *Plant Prod. Sci.*, **12**(1), 3-8, DOI http://dx.doi.org/10.1626/pps.12.3

Penman, H.L., 1948: Natural Evaporation from Open Water, Bare Soil and Grass. Proceedings of the Royal Society of London Series A Mathematical and Physical Sciences, **193**(1032), 120-145.

Phipps, S. J., and Coauthors, 2013: Paleoclimate Data-Model Comparison and the Role of Climate Forcings over the Past 1500 Years. *J. Climate*, **26**, 6915-6936.

Scheff, J., and D. M. W. Frierson, 2013: Scaling Potential Evapotranspiration with Greenhouse Warming. *J. Clim.*, **27**, 1539-1558, DOI 10.1175/JCLI-D-13-00233.1

Schmidt, G. A., J. D. Annan, P. J. Bartlein, B. I. Cook, E. Guilyardi, J. C. Hargreaves, S. P. Harrison, M. Kageyama, A. N. LeGrande, B. Konecky, S. Lovejoy, M. E. Mann, V. Masson-Delmotte, C. Risi, D. Thompson, A. Timmermann, L.-B. Tremblay, and P. Yiou, 2014: Using palaeo-climate comparisons to constrain future projections in CMIP5, *Clim. Past*, **10**, 221-250, doi:10.5194/cp-10-221-2014.

Scholes, R., J. Settele, R. Betts, S. Bunn, P. Leadley, D. Nepstad, J. Overpeck, M. A. Taboada, 2014: Terrestrial and Inland Water Systems. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge,

United Kingdom and New York, NY, USA.

Smerdon, J.E., 2012: Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments, *WIREs Clim. Ch.*, **3**, 63-77, doi:10.1002/wcc.149.

Sundberg, R., A. Moberg, and A. Hind, 2012: Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium – Part 1: Theory. *Clim. Past*, **8**, 1339-1353, doi:10.5194/cp-8-1339-2012

Seager, R., R. Burgman, Y. Kushnir, A. Clement, E. Cook, N. Naik, and J. Miller, 2008: Tropical Pacific forcing of North American Medieval Megadroughts: Testing the Concept with an Atmosphere Model Forced by Coral-Reconstructed SSTs. *J. Clim.*, **21**(23), 6175-6190, DOI http://dx.doi.org/10.1175/2008JCLI2170.1

Seager, R., M. Ting, C. Li, N. Naik, B. Cook, J. Nakamura, and H. Liu, 2013: Projections of declining surface-water availability for the southwestern United States. *Nature Clim. Ch.*, **3**, 482-486, DOI 10.1038/nclimate1787

Sellers, P. J., R. E. Dickinson, D. A. Randall, A. K. Betts, F. G. Hall, J. A. Berry, G. J. Collatz, A. S. Denning, H. A. Mooney, C. A. Nobre, N. Sato, C. B. Field, and A. Henderson-Sellers, 1997: Modeling the Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere. *Science,* **275**(5299), 502-509, DOI 10.1126/science.275.5299.502.

Seneviratne, S. I., 2012: Climate science: Historical drought trends revisited. *Nature*, **491**(7424), 338-339, URL http://dx.doi.org/10.1038/491338a

Sheffield, J., E. F. Wood, and M. L. Roderick, 2012: Little change in global drought over the past 60 years. *Nature,* **491**(7424), 435-438, DOI:10.1038/nature11575.

Stocker, R., P. W. Leadley, and C. Körner, 1997: Carbon and water fluxes in a calcareous grassland under elevated CO_2. *Functional Ecology,* **11**(2), 222-230, DOI:10.1046/j.1365-

2435.1997.00071.x

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An Overview Of CMIP5 And The Experiment Design. *Bull. Am. Met. Soc.*, **93**(4), 485-498, DOI: http://dx.doi.org/10.1175/BAMS-D-11-00094.1

Tierney, J.E., J.E. Smerdon, K.J. Anchukaitis, and R. Seager, 2013: Multidecadal variability in East African hydroclimate controlled by the Indian Ocean. *Nature*, **493**, 389-392.

Tingley, M. P., P. F. Craigmile, M. Haran, B. Li, E. Mannshardt, B. Rajaratnam, 2012: Piecing together the past: statistical insights into paleoclimatic reconstructions, Quat. Sci. Rev., **35**, 1-22.

Thornthwaite, C., 1948: An approach toward a rational classification of climate. *Geograph. Rev.,* **38**(1), 55-94.

Trenberth, K. E., A. Dai, G. van der Schrier, P. D. Jones, J. Barichivich, K. R. Briffa, and J. Sheffield, 2014: Global warming and changes in drought. *Nature Clim. Ch.*, **4**, 17-22, doi:10.1038/NCLIMATE2067.

Trouet, V., J. Esper, N. E. Graham, A. Baker, D. C. Frank, J. D. Scourse, 2009: Persistent positive North Atlantic Oscillation mode dominated the Medieval Climate Anomaly. *Science,* **324**, 78-80, DOI: 10.1126/science.1166349

van der Schrier, G., P. D. Jones, and K. R. Briffa, 2011: The sensitivity of the PDSI to the Thornthwaite and Penman-Monteith parameterizations for potential evapotranspiration, *J. Geophys. Res. Atmos.*, **116**, D03106, doi:10.1029/2010JD015001.

van der Schrier, G., J. Barichivich, K. R. Briffa, and P. D. Jones, 2013: A scPDSI-based global data set of dry and wet spells for 1901-2009. *J. Geophys. Res.*, DOI 10.1002/jgrd.50355

Vicente-Serrano, S. M., S. Beguera, and J. I. Lopez-Moreno, 2009: A Multiscalar Drought Index

Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *J. Clim.*, **23**(7), 1696-1718, DOI 10.1175/2009JCLI2909.1

Vicente-Serrano, S. M., S. Beguera, J. I. Lopez-Moreno, M. Angulo, and A. El Kenawy, 2010: A new global 0.5 gridded dataset (1901-2006) of a multiscalar drought index: comparison with current drought index datasets based on the Palmer Drought Severity Index. *J. Hydromet.*, **11**(4), 1033-1043, DOI http://dx.doi.org/10.1175/2010JHM1224.1.

Wahl, E.R., and J.E. Smerdon, 2012: Comparative performance of paleoclimatic field and index reconstructions derived from climate proxies and noise-only predictors, *Geophys. Res. Lett.*, **39**, L06703, doi:10.1029/2012GL051086.

Wells, N., S. Goddard, and M. J. Hayes, 2004: A self-calibrating palmer drought severity index, *J. Climate*, **17**, 2335-2351.

Wiltshire, A., J. Gornall, B. Booth, E. Dennis, P. Falloon, G. Kay, D. McNeall, C. McSweeney, and R. Betts, 2013: The importance of population, climate change and fCO2g plant physiological forcing in determining future global water stress. *Glob. Environ. Ch.*, **23**(5), 1083-1097, DOI http://dx.doi.org/10.1016/j.gloenvcha.2013.06.005

Xu, C. Y., and V. P. Singh, 2002: Cross Comparison of Empirical Equations for Calculating Potential Evapotranspiration with Data from Switzerland. *Water Res. Man.*, **16**(3), 197-219, DOI 10.1023/A:1020282515975

Figure Captions

Figure 1. Map of the contiguous United States and those states chosen to represent the Four Corners (4C), North Plains (NP) and Southeast (SE) regions (designated with grey shading). The tree-ring sites used in the regional reconstructions are shown as red dots. Plotted time series are the area-weighted mean PDSI estimates for each of the three regions estimated from observational data and three PDSI formulations: PDSI Thornthwaite (PDSI_TH), PDSI Penman-Monteith (PDSI_PM) and self-calibrating PDSI with Penman-Monteith (PDSI_SC).

Figure 2. Targeted regional PDSI time series calculated from observational data for the PDSI_TH, PDSI_PM and PDSI_SC formulations and the associated CPS reconstructions and 95% confidence intervals. Reconstructions and observationally-based estimates are shown during their common period of overlap from 1901-1979 CE, which also comprises the calibration/validation interval for the reconstructions.

Figure 3. PDSI reconstructions using the CPS method and three regional PDSI target series using the PDSI_TH, PDSI_PM and PDSI_SC formulations (a, c, and e). The 10-yr lowpass time series for each PDSI reconstruction (filtered using a 10-point butterworth filter) are also shown in panels a, c and e, while each panel also plots the annual PDSI_TH and associated 95% confidence intervals for reference. Panels b, d and f plot resolved variance, RE and CE cross-validation statistics for each of the regional PDSI reconstructions as a function of each 50-yr nest. Figure legends for panels a-f are given in panels c and d. Comparisons between the annually-reconstructed PDSI values are shown in the scatter plots in panels g-i, in which each value of reconstructed PDSI is plotted against the other. Scatter plots do not include the

calibration/validation interval from 1901-1979 and plot a dashed 1:1 line for reference.

Figure 4. Comparisons of reconstructed PDSI_TH from this study using CPS and the regional averages of the gridded NADA product over the targeted CPS reconstruction domains. Time series plot both annual values and the 10-yr lowpass filtered time series using a 10-point butterworth filter. The shared variance (r^2) between the annual (filtered) CPS and NADA time series over their periods of overlap are shown within each plot in the left column. Scatter plots for the annual CPS and NADA values are shown in the right column of the figure with one-to-one lines shown for reference.

Figure 5: Regional model estimates of PDSI_TH, PDSI_PM, normalized 30cmSM and normalized FCSM during the historical interval (1901-2005 CE) in the CanESM2 and CCSM4 models. The first ensemble member of each model is plotted (interannual r^2 estimates across all five ensemble members are shown in Figures 6 and 7). In all cases, PDSI or soil moisture normalizations used the 1901-2005 CE intervals as a baseline, but time series are recentered from 1901-1979 CE to match the calibration/validation interval of the PDSI reconstructions.

Figure 6: Regional shared variances (r^2) among the collection of soil moisture metrics in the CanESM2 model simulations during the historical (1901-2005 CE) and projection (2006-2099 CE) intervals. Vertical diamond triplets correspond to the maximum, minimum, and median shared variance across the five members of the ensemble. Detrended results have been computed for the projection interval after removing a linear trend over the same period.

Figure 7: Same as in Figure 6, but for the CCSM4 model simulations. Maximum, minimum and median results were similarly computed from an ensemble of five simulations.

Figure 8: A comparison of reconstructed PDSI_TH in the 4C region (see Figure 1) with collocated model estimates of PDSI_TH, PDSI_PM, normalized 30cmSM and normalized FCSM during the historical and projection intervals. Results are shown for the first ensemble member of the CanESM2 and CCSM4 model simulations to allow direct visual comparisons between the reconstructed and model variance. All time series have been centered over the 1901-1979 interval, the time period of common overlap. Reconstructions extend prior to 1800, but the 1800-2100 CE interval is chosen for ease of visual comparison.

Figure 9: Same as in Figure 8, but for the North Plains region (see Figure 1).

Figure 10: Same as in Figure 8, but for the Southeast region (see Figure 1).

Figure 11: (top row) Mean PDSI_PM, normalized 30cmSM and normalized FCSM for the last two decades (2080-2099 CE) of the 21st-century projection interval (ensemble member 1) from the CanESM2 model. (bottom row) Agreement between the sign of the wetting (blue) or drying (brown) as projected by the three variables in the 2080-2099 CE interval. The total percentage of grid cells that agree on the wetting or drying trends are provided in the lower left hand part of the comparison maps. The boxes defined by dashed lines indicate the regions extracted from the 0.5°x0.5° NADA, observation and model grids as representative of the 4C, NP and SE regions.

Figure 12: Same as in Figure 11, but for the CCSM4 model.

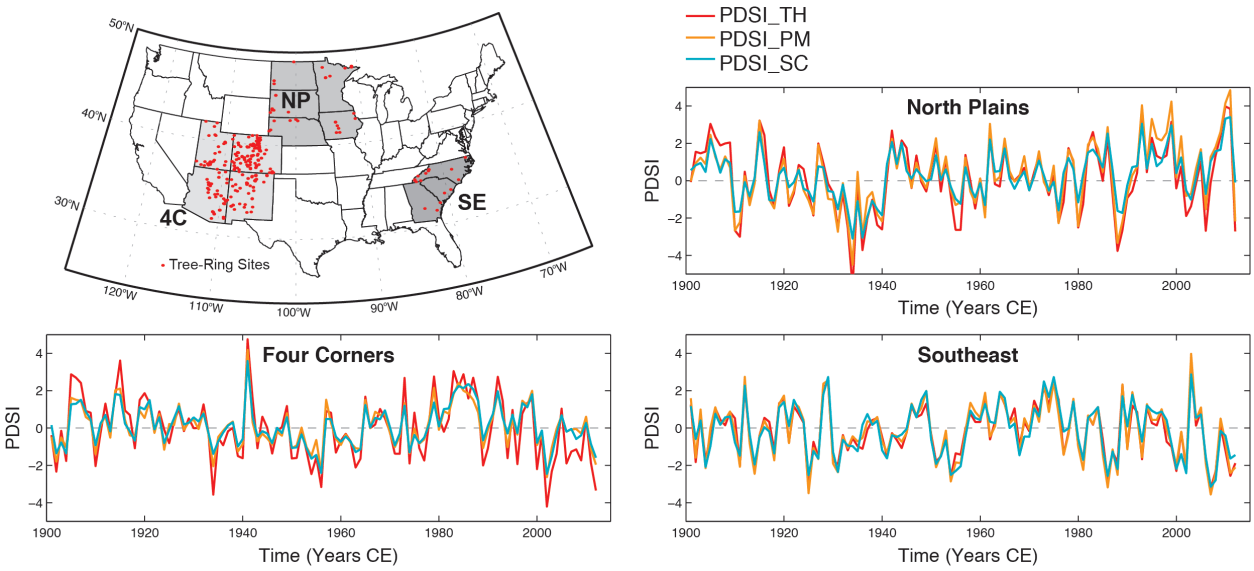

Figure 1. Map of the contiguous United States and those states chosen to represent the Four Corners (4C), North Plains (NP) and Southeast (SE) regions (designated with grey shading). The tree-ring sites used in the regional reconstructions are shown as red dots. Plotted time series are the area-weighted mean PDSI estimates for each of the three regions estimated from observational data and three PDSI formulations: PDSI Thornthwaite (PDSI_TH), PDSI Penman-Monteith (PDSI_PM) and self-calibrating PDSI with Penman-Monteith (PDSI_SC).

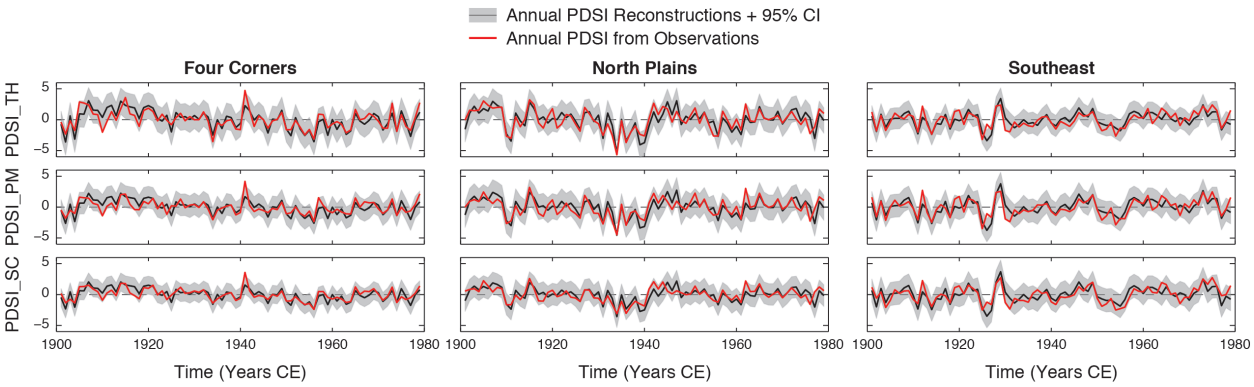

Figure 2. Targeted regional PDSI time series calculated from observational data for the PDSI_TH, PDSI_PM and PDSI_SC formulations and the associated CPS reconstructions and 95% confidence intervals. Reconstructions and observationally-based estimates are shown during their common period of overlap from 1901-1979 CE, which also comprises the calibration/validation interval for the reconstructions.

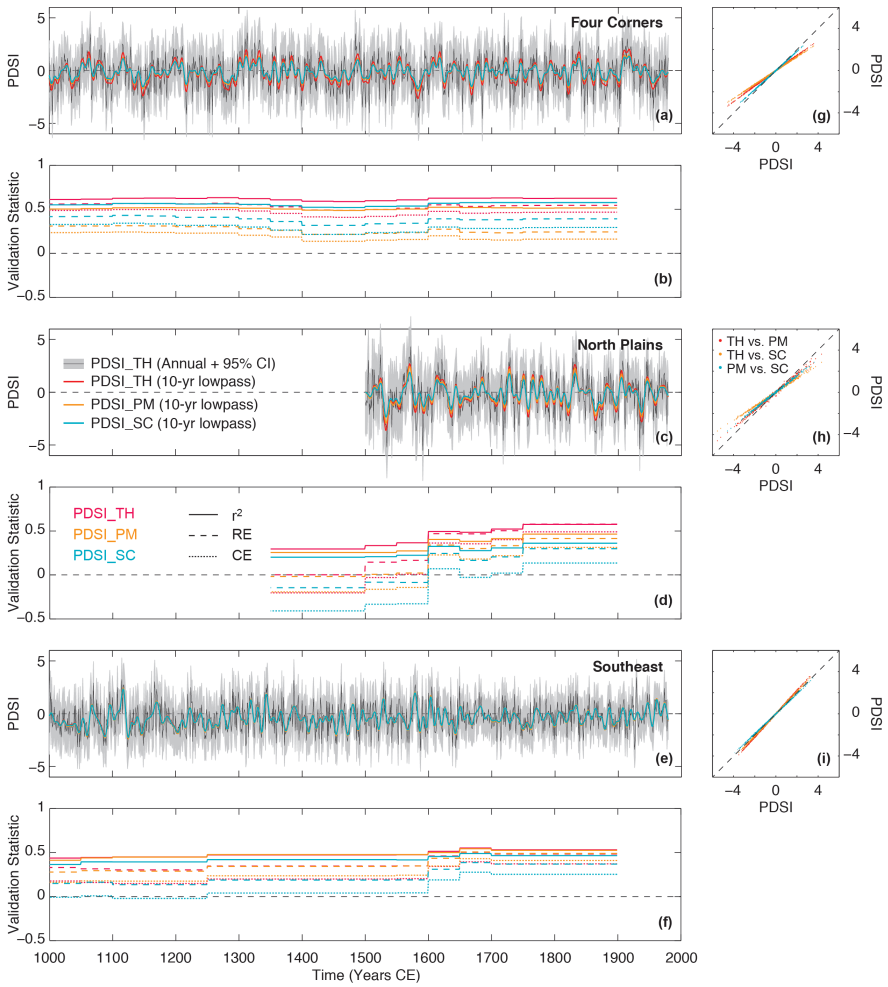

Figure 3. PDSI reconstructions using the CPS method and three regional PDSI target series using the PDSI_TH, PDSI_PM and PDSI_SC formulations (a, c, and e). The 10-yr lowpass time series for each PDSI reconstruction (filtered using a 10-point butterworth filter) are also shown in panels a, c and e, while each panel also plots the annual PDSI_TH and associated 95% confidence intervals for reference. Panels b, d and f plot resolved variance, RE and CE cross-validation statistics for each of the regional PDSI reconstructions as a function of each 50-yr nest. Figure legends for panels a-f are given in panels c and d. Comparisons between the annually-reconstructed PDSI values are shown in the scatter plots in panels g-i, in which each value of reconstructed PDSI is plotted against the other. Scatter plots do not include the calibration/validation interval from 1901-1979 and plot a dashed 1:1 line for reference.

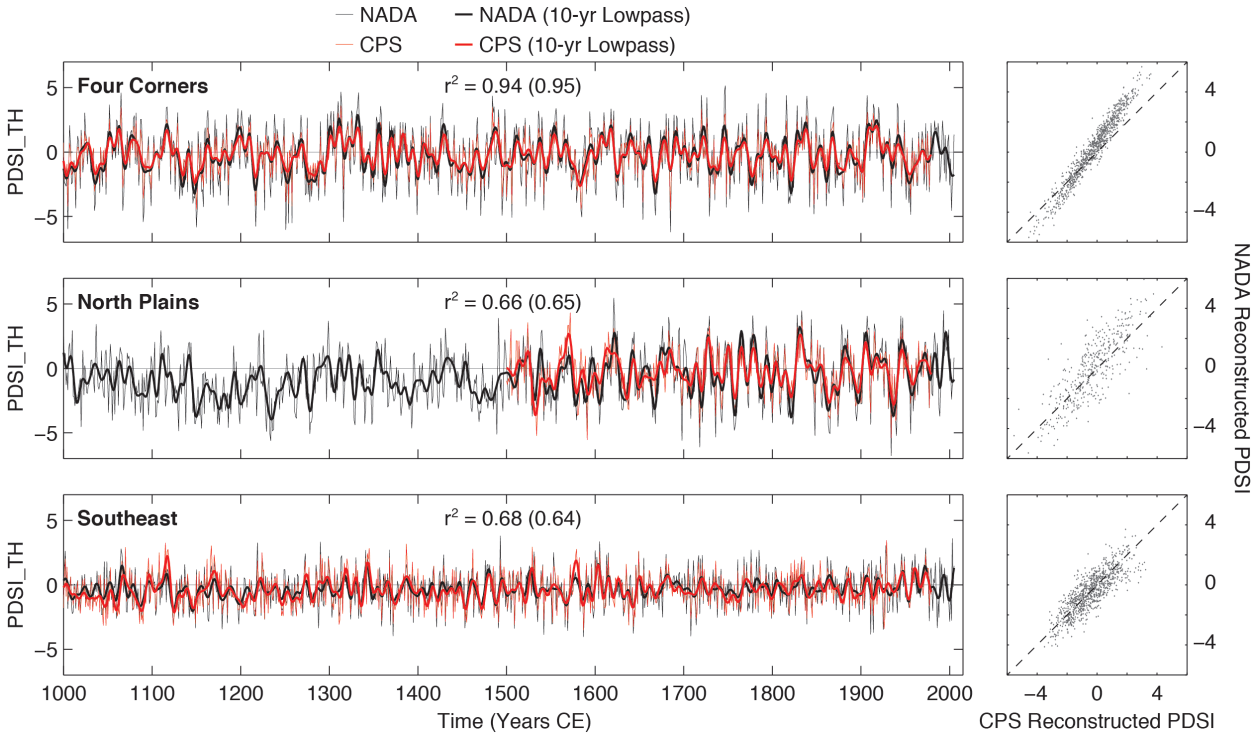

Figure 4. Comparisons of reconstructed PDSI_TH from this study using CPS and the regional averages of the gridded NADA product over the targeted CPS reconstruction domains. Time series plot both annual values and the 10-yr lowpass filtered time series using a 10-point butterworth filter. The shared variance (r^2) between the annual (filtered) CPS and NADA time series over their periods of overlap are shown within each plot in the left column. Scatter plots for the annual CPS and NADA values are shown in the right column of the figure with one-to-one lines shown for reference.

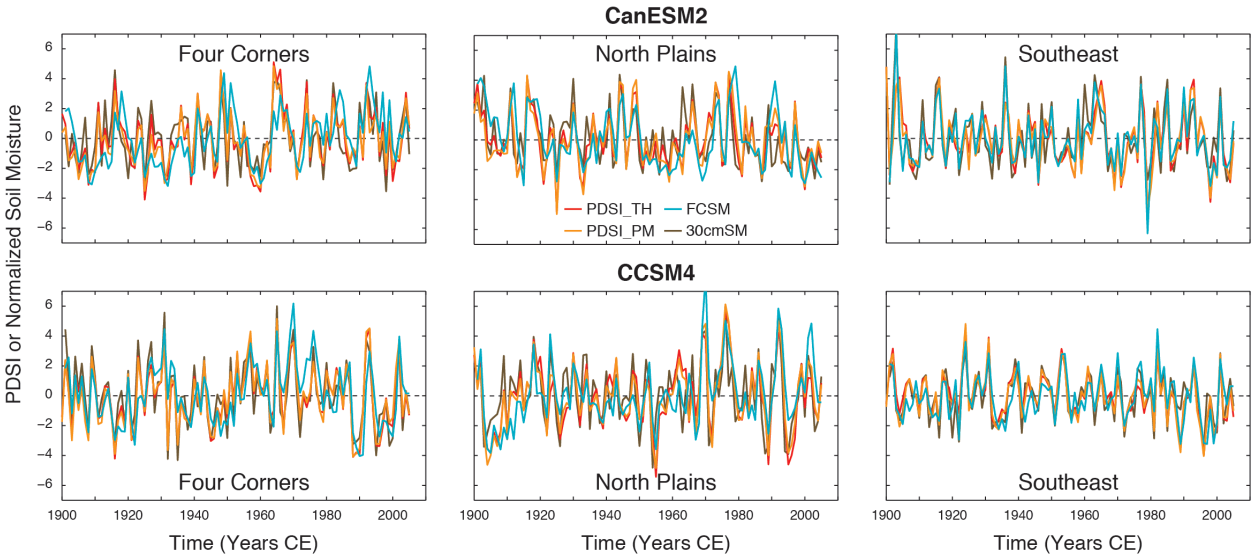

Figure 5: Regional model estimates of PDSI_TH, PDSI_PM, normalized 30cmSM and normalized FCSM during the historical interval (1901-2005 CE) in the CanESM2 and CCSM4 models. The first ensemble member of each model is plotted (interannual r^2 estimates across all five ensemble members are shown in Figures 6 and 7). In all cases, PDSI or soil moisture normalizations used the 1901-2005 CE intervals as a baseline, but time series are recentered from 1901-1979 CE to match the calibration/validation interval of the PDSI reconstructions.

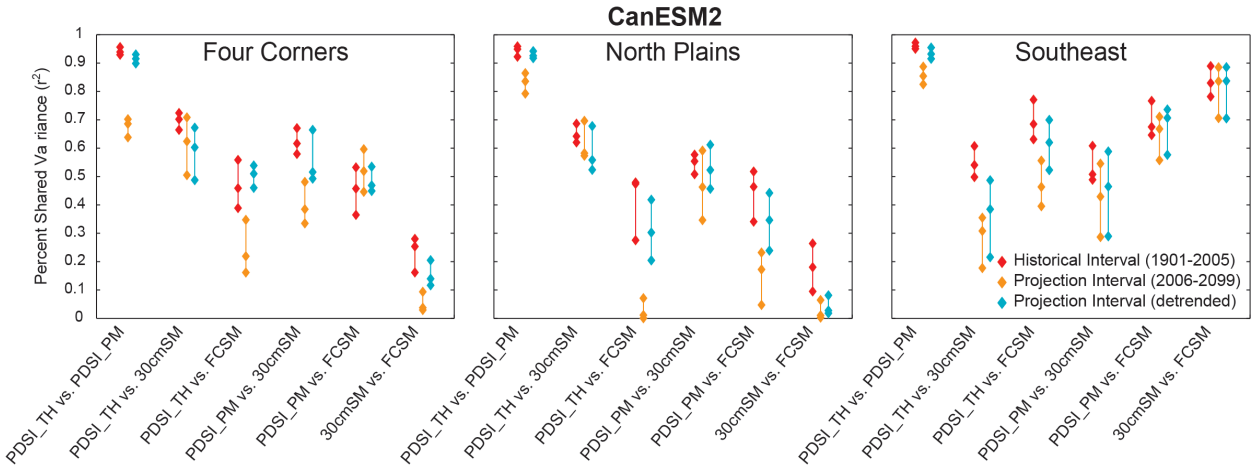

Figure 6: Regional shared variances (r^2) among the collection of soil moisture metrics in the CanESM2 model simulations during the historical (1901-2005 CE) and projection (2006-2099 CE) intervals. Vertical diamond triplets correspond to the maximum, minimum, and median shared variance across the five members of the ensemble. Detrended results have been computed for the projection interval after removing a linear trend over the same period.

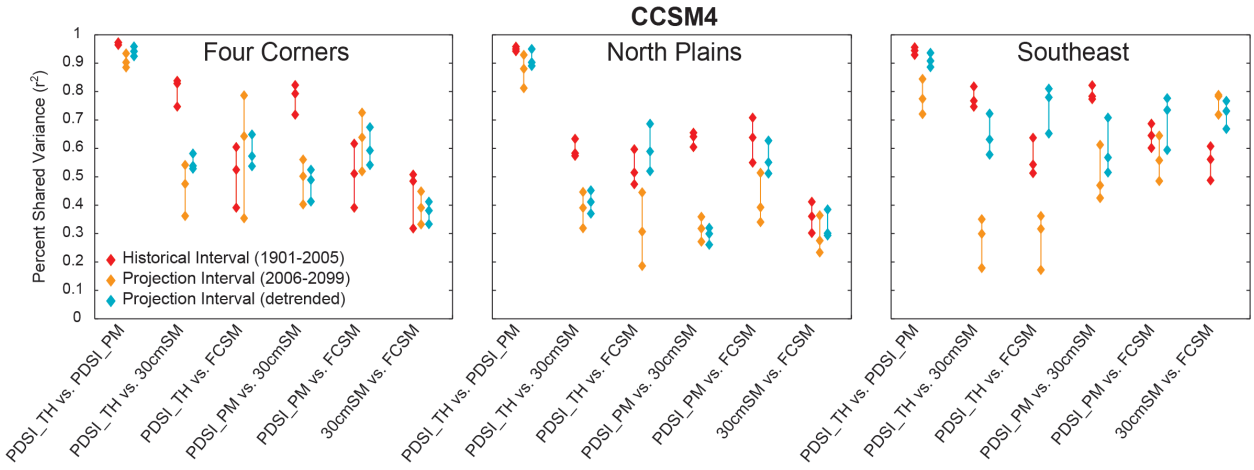

Figure 7: Same as in Figure 6, but for the CCSM4 model simulations. Maximum, minimum and median results were similarly computed from an ensemble of five simulations.

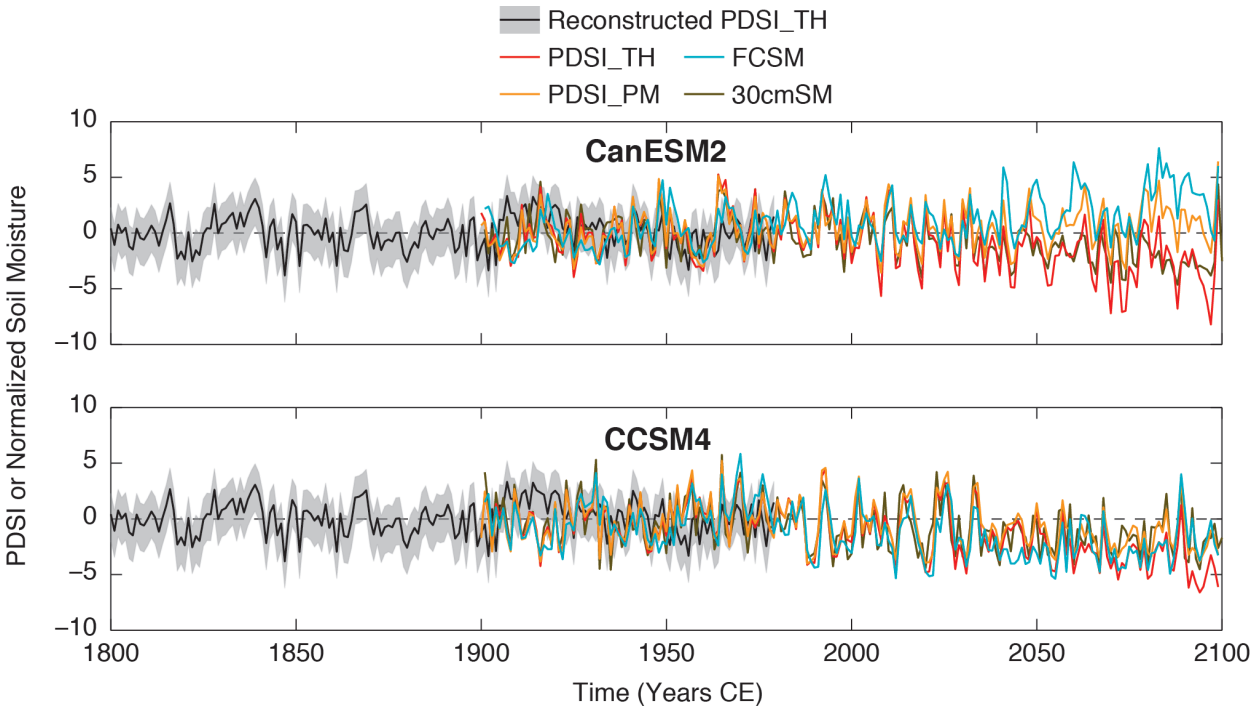

Figure 8: A comparison of reconstructed PDSI_TH in the 4C region (see Figure 1) with collocated model estimates of PDSI_TH, PDSI_PM, normalized 30cmSM and normalized FCSM during the historical and projection intervals. Results are shown for the first ensemble member of the CanESM2 and CCSM4 model simulations to allow direct visual comparisons between the reconstructed and model variance. All time series have been centered over the 1901-1979 interval, the time period of common overlap. Reconstructions extend prior to 1800, but the 1800-2100 CE interval is chosen for ease of visual comparison.

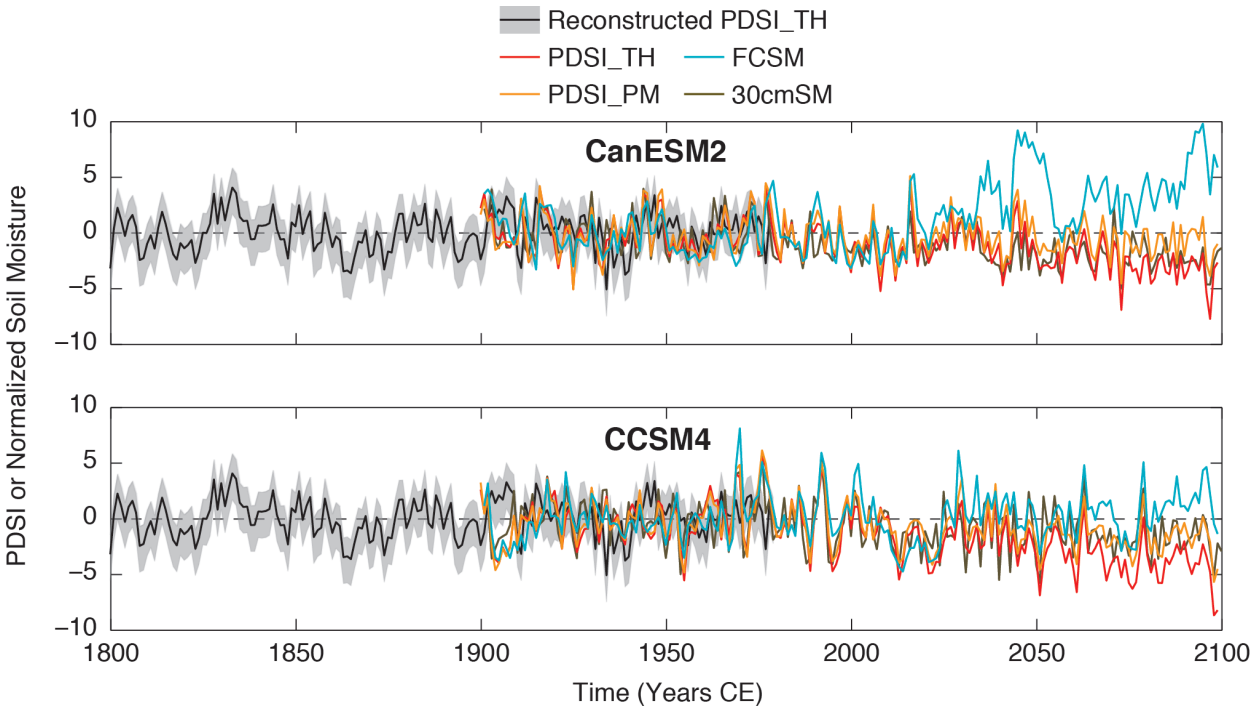

Figure 9: Same as in Figure 8, but for the North Plains region (see Figure 1).

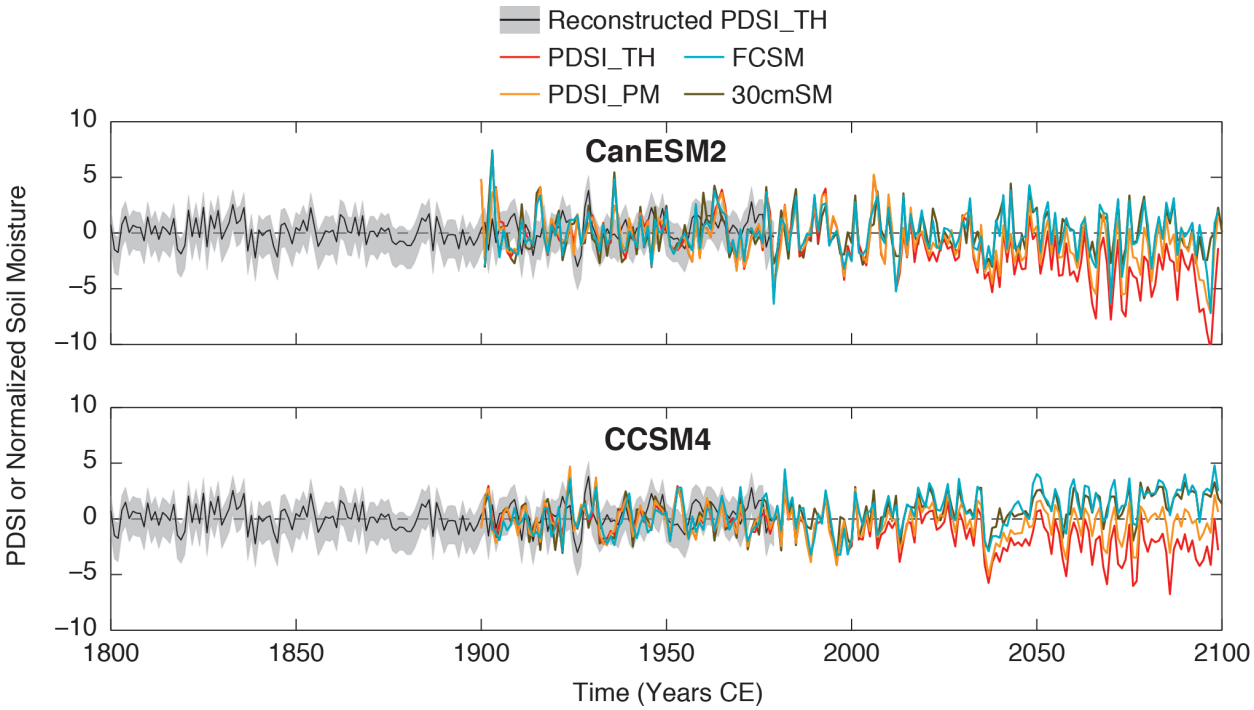

Figure 10: Same as in Figure 8, but for the Southeast region (see Figure 1).

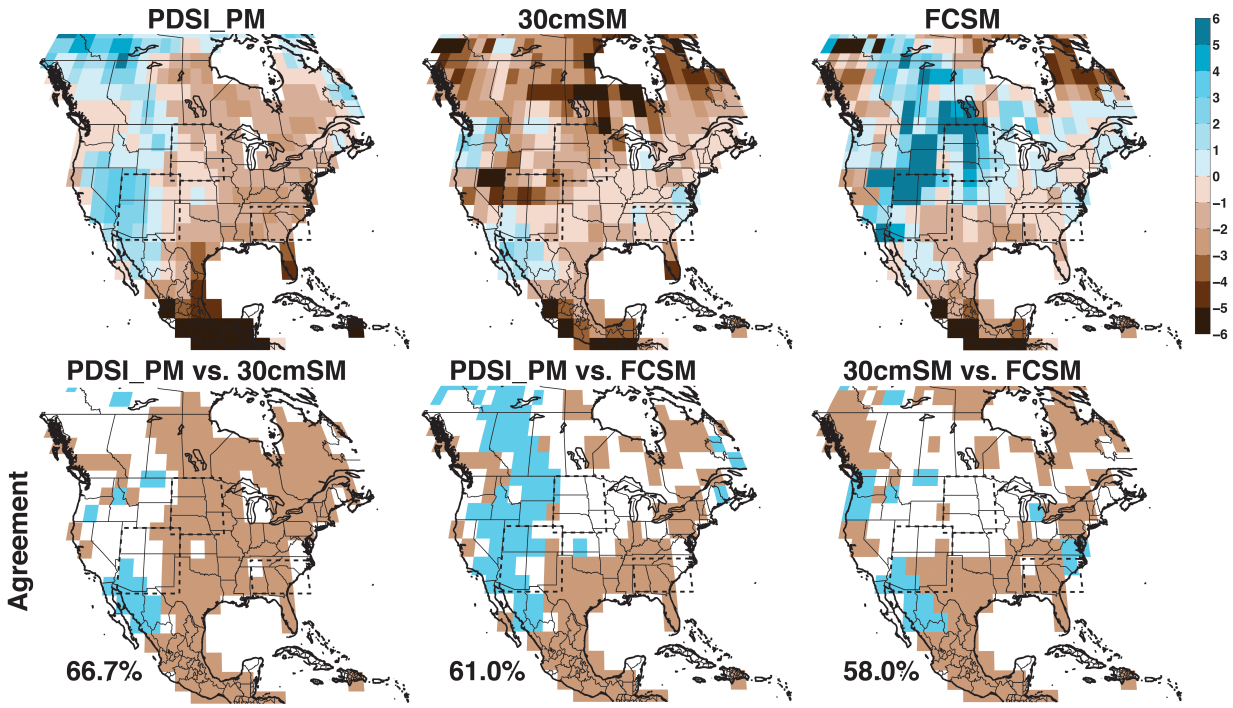

Figure 11: (top row) Mean PDSI_PM, normalized 30cmSM and normalized FCSM for the last two decades (2080-2099 CE) of the 21st-century projection interval (ensemble member 1) from the CanESM2 model. (bottom row) Agreement between the sign of the wetting (blue) or drying (brown) as projected by the three variables in the 2080-2099 CE interval. The total percentage of grid cells that agree on the wetting or drying trends are provided in the lower left hand part of the comparison maps. The boxes defined by dashed lines indicate the regions extracted from the 0.5°x0.5° NADA, observation and model grids as representative of the 4C, NP and SE regions.

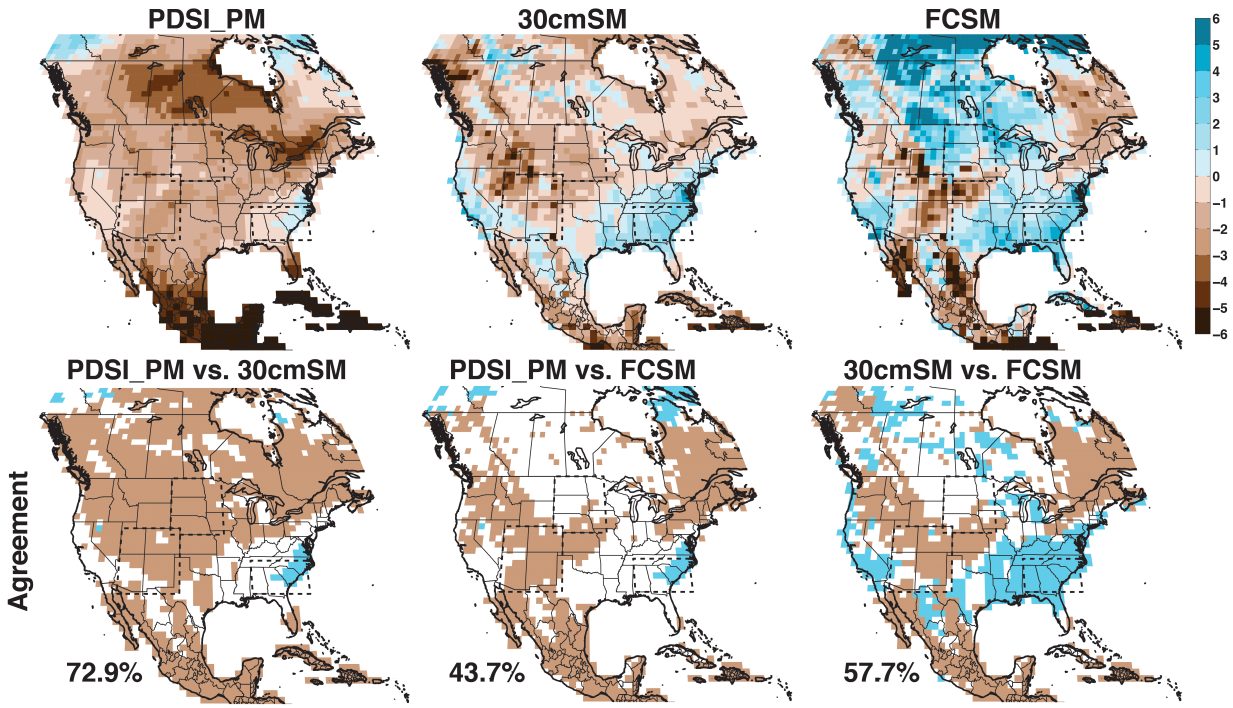

Figure 12: Same as in Figure 11, but for the CCSM4 model.