Journal of Hydrology: Regional Studies 24 (2019) 100606

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Development of a gridded reference evapotranspiration dataset for the Great Lakes region

Michael T. Kiefer[a,*], Jeffrey A. Andresen[a], Dana Doubler[b], Aaron Pollyea[a]

[a] Department of Geography, Environment, and Spatial Sciences, Michigan State University, East Lansing, MI, United States
[b] Toledo Metropolitan Area Council of Governments, Toledo, OH, United States

ARTICLE INFO

Keywords:
Reference evapotranspiration
Potential evaporative demand
Climate
Great Lakes
Gridded dataset

ABSTRACT

Study region: Great Lakes region of North America
Study focus: Reference Evapotranspiration (ETo) is a key variable in water use management and irrigation of agricultural crops. This study describes the development of a gridded historical ETo dataset for the period 1983–2012 derived from North American Data Assimilation System Phase 2 (NLDAS-2) forcing fields for the Great Lakes region of North America. The gridded dataset is intended to fill a gap in the resource toolbox available to growers in this region of rapidly expanding irrigation. As a prerequisite for development of the ETo dataset, a correction procedure is applied to the NLDAS-2 downward solar radiation to account for overall bias and a tendency to underestimate the range of solar radiation on hourly and daily timescales.
New hydrological insights: Analyses of spatial and temporal variability reveal that the lakes play an important role in modulating seasonal and geographical variability in evaporative demand. An example application of the gridded historical ETo dataset to irrigation management is provided. A 30-year climatology of crop irrigation for field maize is developed from the ETo dataset and is applied in a hypothetical irrigation-scheduling scenario. Overall, the study illustrates the utility of the NLDAS-2 ETo dataset in describing spatial and temporal patterns of evaporative demand across the Great Lakes region, and as a source of reference climatological information for irrigation management.

1. Introduction

At a minimum, irrigation managers have three primary decisions to make during the growing season: how much irrigation water to apply, when to begin application, and how long during the season to apply for. These three questions apply to irrigation in humid as well as arid climates. However, irrigation decision-making in humid climates is arguably more complex than in arid climates due to factors both meteorological/climatological (e.g., higher rainfall frequency, higher humidity, greater cloud cover) and structural (e.g., lack of large-scale irrigation water projects) in nature (Yoon et al., 1993; Sadler et al., 2003). First, irrigation managers in humid regions must, to a much greater degree than for production systems located in arid regions, consider the possibility of rain occurring during or shortly after irrigation. Care must be taken to avoid irrigating prior to a substantial rainfall event due to an increased risk of overwatering or flooding the crop and leaching nutrients and/or chemicals from the plants and the rooting zone. Second, a cloudier and more humid climate implies lower evaporative demand by the atmosphere; however, during periods without rain, the weather in

* Corresponding author at: Department of Geography, Environment, and Spatial Sciences, Michigan State University, 673 Auditorium Rd., Room 204, East Lansing, MI, 48824, United States.
E-mail address: mtkiefer@msu.edu (M.T. Kiefer).

https://doi.org/10.1016/j.ejrh.2019.100606
Received 7 August 2018; Received in revised form 23 April 2019; Accepted 22 May 2019
Available online 30 May 2019

humid regions largely mirrors that of arid regions. Thus, understanding how evaporative demand varies both temporally and spatially is of great importance for irrigation managers, particularly in humid climates.

Evapotranspiration (ET) is the combined process through which water evaporates from soil surfaces (evaporation) and through plant stomata (transpiration). ET is sensitive to a variety of factors both meteorological (e.g., downward solar radiation, air temperature, humidity, and wind speed) and non-meteorological (e.g., crop type, development stage, management practices). Reference ET (ETo) is a form of ET in which purely meteorological factors are isolated using a standardized reference grass-covered surface which is not shaded nor for which water is limiting. In other words, ETo is a measure of the potential evaporative power of the atmosphere for the specified reference surface at a given location and time. The standard method of computing ETo is the modified Penman-Monteith equation presented in the Food and Agriculture Organization (FAO) Irrigation and Drainage Paper No. 56 (hereafter, FAO-56: Allen et al., 1998), which considers all parameters that control energy exchange and corresponding evapotranspiration from uniform vegetation.

Traditionally, ETo is computed from data collected at individual weather station sites. However, with the rapid development of gridded weather analyses over the past two decades [e.g., the National Centers for Environmental Prediction (NCEP)/ National Corporation for Atmospheric Research (NCAR) global Reanalysis (Kalnay et al., 1996), the North American Regional Reanalysis (NARR) (Mesinger et al., 2006), and the Real Time Mesoscale Analysis (RTMA) (De Pondeca et al., 2011)], the availability of gridded ETo products has become a reality. In 2001, the U.S. Geological Survey (USGS) Center for Earth Resources Observation and Science (EROS) began generating a global 1-degree resolution daily ETo product based on the Global Data Assimilation System (GDAS) (Senay et al., 2008). Beginning in 2013, the National Oceanic and Atmospheric Administration (NOAA) began producing a continental U.S. daily ETo product based on the 0.125-degree resolution North American Land Data Assimilation System, Phase 2 (NLDAS-2) (Hobbins et al., 2013; Moorhead et al., 2015). In addition to these near real-time products, Abatzoglou (2013) introduced a 4-km resolution gridded hybrid analysis for the period 1979–2010 based on NLDAS-2 and the 4-km resolution Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daly et al., 2008). As an example of the use of the hybrid dataset in ecological modeling, Abatzoglou (2013) produced gridded ETo using the FAO-56 modified Penman-Monteith method.

Validation of each of the earlier gridded ETo datasets has been reported in previous studies [USGS (Senay et al., 2008; Liu et al., 2011), NOAA (Moorhead et al., 2015), NLDAS-PRISM (Abatzoglou, 2013)]. Validation of the USGS ETo product at sites in California (Senay et al., 2008) and Oklahoma (Liu et al., 2011) did not show any systematic biases. However, Senay et al. (2008) did identify considerable positive biases at coastal sites, attributing the overestimation to the inability of the 1-degree GDAS to resolve the sharp temperature gradients between the coast and interior. The inability of GDAS to resolve spatial gradients of ETo input variables (and thus ETo itself) near coastlines is a major limitation of the USGS ETo product. The finer resolution of NLDAS-2 is more suitable for studying ETo in coastal regions (e.g., the Great Lakes region of North America).

However, validation of NLDAS-2–derived ETo products has revealed a systematic positive bias at sites in the central and western U.S., including the Texas High Plains (NOAA: Moorhead et al., 2015) and Pacific Northwest (NLDAS-PRISM: Abatzoglou, 2013). Assessing ETo during the growing season (defined therein as May-September), Abatzoglou (2013) found an average positive bias of 0.5 mm day^{-1}, although a bias of 1–2 mm day^{-1} was reported at some sites. Moorhead et al. (2015) qualitatively described a positive ETo bias in the NOAA product, but only presented a limited suite of statistics. Furthermore, since Moorhead et al. (2015) presented statistics for the whole year, including the dormant season, caution must be exercised when comparing their results to other studies. In both studies, the overestimation of ETo was primarily attributed to variables other than downward solar radiation (i.e., humidity, wind speed, and temperature). The possible impact on ETo estimates of a positive downward solar radiation bias in NLDAS-2 [as reported in Sakamoto et al. (2011), Lewis et al. (2014), and Slater (2016)] was not addressed, nor was any effort made to correct for such a bias.

The primary goal of this study is the development of a gridded historical hourly and daily ETo dataset, for the period 1983–2012, suitable for application in the Great Lakes region of North America. The gridded dataset developed herein is intended to fill a gap in the resource toolbox available to growers in this region. Existing historical datasets are generally station-based or, if they are gridded, lack the spatial resolution needed to resolve climatic gradients resulting from the presence of land-water contrasts and complex coastlines. The presence of complex coastlines in the Great Lakes region motivates the choice of the higher-resolution NLDAS-2 (as opposed to the coarser-resolution GDAS) as the source of gridded meteorological data for the calculation of ETo. The secondary goal of this study is the evaluation of NLDAS-2 downward solar radiation and the development of a correction procedure to account for solar radiation bias and underestimation of temporal variability. This secondary goal is motivated by findings of a positive NLDAS-2 solar radiation bias in earlier studies [e.g., Sakamoto et al. (2011)], and knowledge of the moderating effect of grid cell averaging on the representation of weather extrema in gridded datasets [e.g., Daley (1991); Kitchen and Blackall (1992)]. The premise underlying this secondary goal is that use of corrected NLDAS-2 solar radiation will result in a more accurate ETo dataset, compared to one computed using uncorrected data.

2. Materials and methods

2.1. Study area

This study is focused on the Great Lakes region of North America. This region contains one of the largest fresh water resources in the world and is characterized by strong physiographic gradients of soils, vegetation, and climate (Andresen and Winkler, 2009). Although the region's climate is categorized as humid, irrigation for agriculture has expanded rapidly during the past few decades. For example, in the state of Michigan, irrigated area increased from just over 115,000 ha in 1982 (Bureau of the Census, 1986) to

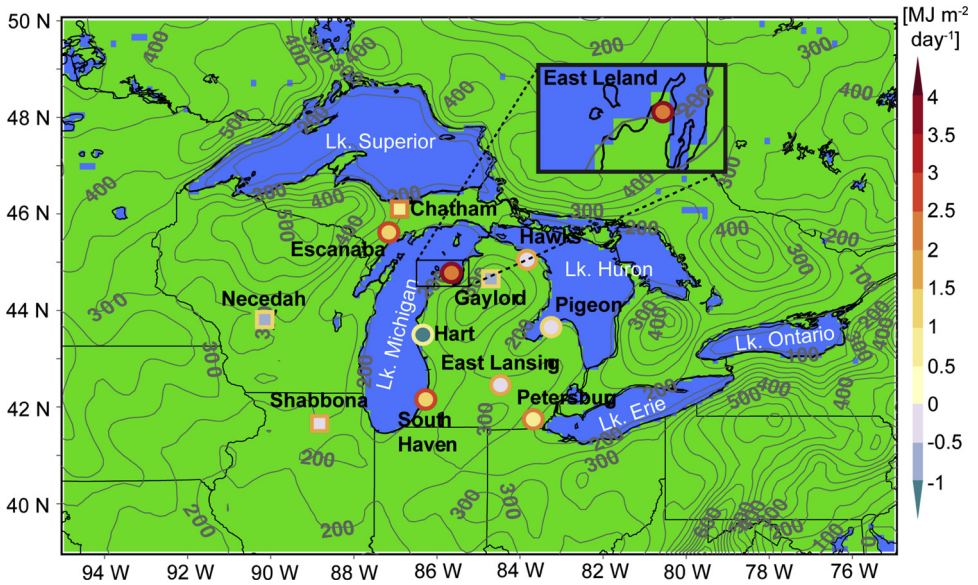

Fig. 1. Study domain, with location of stations used in solar bias and variance correction steps (USCRN – squares) and in solar radiation and ETo validation step (EW –circles). Marker edge (face) color denotes NLDAS-2 daily solar radiation bias before (after) correction procedure (see colorbar). Inset panel depicts the immediate area surrounding East Leland. For details about individual stations, see Table 1. Green (blue) background shading denotes NLDAS-2 land (water) grid cells. Overlaid contours depict surface elevation [meters above mean sea level]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

more than 239,000 ha in 2012 (National Agricultural Statistics Service, 2014). In this study, the Great Lakes region is defined as a rectangle with lower-left coordinates of 38.94 N, 95.06 W and upper-right coordinates of 50.06 N, 74.94 W (Fig. 1). The study area consists of approximately $1.84e^8$ ha, of which about 85% is land and 15% is water. Topography across the study area (Fig. 1) falls mainly into three categories: plains (e.g., Iowa, Illinois), hills and low mountains (e.g., northeast Minnesota, Michigan's western upper peninsula), and Appalachian Mountains topography (e.g., central Pennsylvania). The lowest surface elevations are found across southern Illinois and northern Ontario (north of inset panel in Fig. 1), where surface elevations below 200 m above mean sea level are common; highest elevations in the study area are located along an axis from northern West Virginia to southwestern New York, where surface elevations of 500 m above mean sea level or greater are common.

2.2. Data

NLDAS-2 (Xia et al., 2012a,b) is an atmosphere and land surface hydrology dataset, covering the period 1979–present, consisting of forcing fields derived from NARR (e.g., downward solar radiation, 2-m temperature), and output from four land-surface models driven by the forcing fields. The grid spacing is 0.125 ° (~12 km) and the domain covers the continental US and portions of Canada and Mexico. The NLDAS-2 forcing fields are derived by interpolating 32-km NARR data to the 0.125-degree NLDAS-2 grid, adjusting for elevation differences, and disaggregating the three-hourly NARR to an hourly time scale. The NARR downward solar radiation is bias-corrected using University of Maryland Surface Radiation Budget (SRB) monthly solar data, itself based on Geostationary Operational Environmental Satellite (GOES) data (Pinker et al., 2003). For this study, NLDAS-2 data was downloaded using the National Aeronautics and Space Administration (NASA) Giovanni utility (https://giovanni.gsfc.nasa.gov/giovanni/), for the region defined in Section 2.1, and for the period from 0000 Coordinated Universal Time (UTC) 1 January 1983 to 2300 UTC 31 December 2012 [2000 Eastern Daylight Time (EDT = UTC-4) 31 December 1982 to 1900 EDT 31 December 2012].

2.3. NLDAS-2 solar correction procedure

The remainder of Section 2 focuses on the sequence of steps taken to transform unprocessed NLDAS-2 variables into an hourly and daily ETo dataset; a diagrammatic summary of the procedure is provided in Fig. 2. The procedure consisted mainly of the assessment and subsequent correction of NLDAS-2 downward solar radiation using climate station data. Downward solar radiation is a primary input variable in the series of semi-physical and semi-empirical equations used to compute variables in the FAO-56 Penman-Monteith equation (Allen et al., 1998), as discussed in greater detail in Section 2.4.

Prior to the formal assessment of NLDAS-2 downward solar radiation at climate station sites, a preliminary assessment of spatial patterns of NLDAS-2 downward solar radiation near the Great Lakes revealed a pattern in which artificially high values of solar radiation over the lakes extended 20–30 km inland, in turn creating strong gradients of solar radiation inland of the lakeshore (Fig. S1a-b in Supplementary material). This pattern originates not with NARR but with the SRB data, and appears to be an artifact of the coarse resolution of the underlying GOES dataset (0.5°). A nearest-neighbor approach was utilized to replace erroneous values of

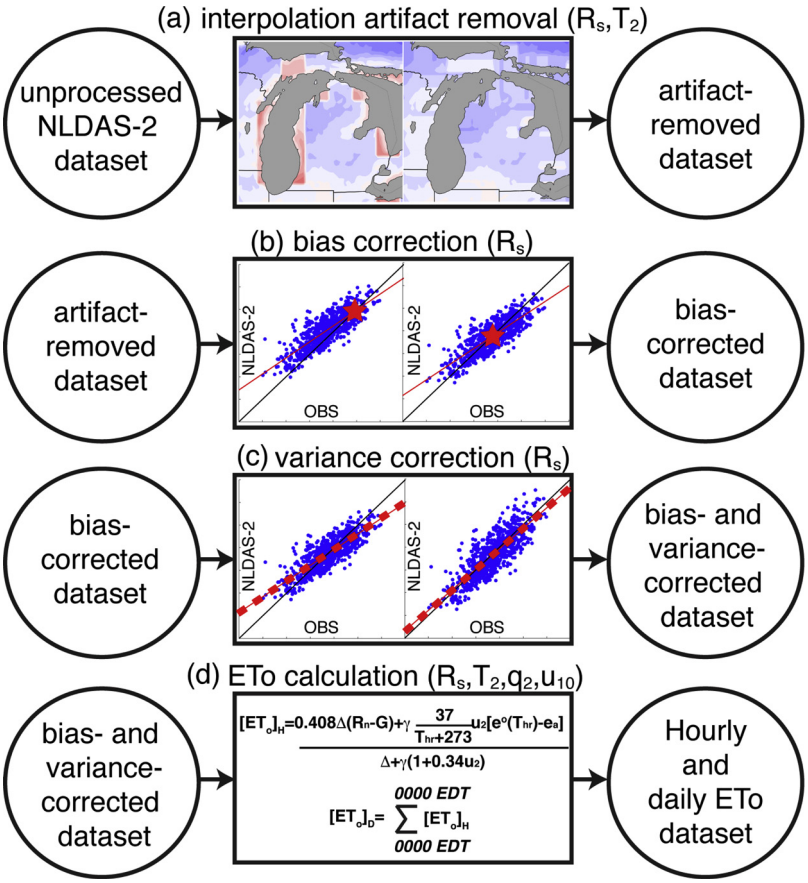

Fig. 2. Diagram of NLDAS-2 ETo dataset construction procedure, where R_s is downward solar radiation, T_2 is 2-m temperature, q_2 is 2-m specific humidity, and u_{10} is 10-m wind speed. The process begins with a nearest-neighbor adjustment of the downward solar radiation and temperature variables to remove artifacts of earlier interpolation from coarser-resolution datasets (GOES and NARR), proceeds to bias correction of downward solar radiation [red star denotes the intercept of the 1:1 (black) and linear regression (red) lines], next proceeds to variance correction of downward shortwave radiation (thick red dashed line highlights the slope of the linear regression line), and concludes with the calculation of hourly and daily ETo using the corrected NLDAS-2 variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

NLDAS-2 downward solar radiation at grid points adjacent to the lakes with values one grid point inland. A similar approach was taken to correct for anomalously high temperatures in far southwestern Ontario (Fig. S1c–d in Supplementary material) that result from interpolation of NARR data to the NLDAS-2 grid at land points between Lakes Erie, Huron, and St. Clair (a small lake between Erie and Huron). The modification of the raw NLDAS-2 downward solar radiation and temperature yielded a set of artifact-removed NLDAS-2 variables (Fig. 2a).

The secondary research goal of this study, as stated in Section 1, is the evaluation and correction of NLDAS-2 solar radiation bias and underestimation of temporal variability. Thus, the artifact-removed NLDAS-2 downward solar radiation was validated using hourly global solar flux measurements from the NOAA U.S. Climate Reference Network (USCRN: Hubbard et al., 2005). USCRN is a system of climate observing stations developed to provide long-term air temperature, precipitation, and soil moisture and temperature observations that are both high quality and collected in stable settings. Validation was performed at four USCRN stations within the Great Lakes region (Necedah, WI; Shabbona, IL; Chatham, MI; and Gaylord, MI), during four growing seasons (2008–2011; Fig. 1, Table 1). However, before proceeding, a word of caution is in order regarding comparisons of point observations and gridded estimates. When comparing primary variables (e.g., downward solar radiation) or derived quantities (e.g., ETo) at a weather/climate station to those extracted at a nearby grid point, as is done in this study, it is important to keep in mind that point values are being compared to spatial averages. Thus, an unknown portion of the differences between the station observations and gridded estimates quoted herein can be attributed to sub-grid variability (i.e., representativeness error) (Daley, 1991; Kitchen and Blackall, 1992). The reader should keep this in mind, particularly when terms such as “error” and “bias” are used.

Upon initial validation of the NLDAS-2 downward solar radiation, a substantial positive bias was discovered at all four stations. Although the focus of the validation was on the hourly timescale, positive bias was also found in the daily solar radiation time series (1.27–2.35 MJ m^{-2} day^{-1}; see square marker edge color in Fig. 1). In the four-station-pooled daily solar radiation scatter plot in Fig. 3a, a mean bias (BIAS) of 1.70 MJ m^{-2} day^{-1} is evident, with the positive bias entirely associated with the left side of the observed

Table 1
Summary of weather station sites used for bias correction and validation of NLDAS-2 solar radiation and ETo. USCRN is the US Climate Reference Network, and EW is the Enviro-weather mesonetwork (see text for details). Variables in column six are limited to those used for correction and validation procedure: downward solar radiation (R_S), 2-m temperature (T_2), 3-m wind speed (u_3), and 2-m relative humidity (h_2).

Station	Network	Lat [° N]	Lon [° W]	Date Range	Variables	Sampling freq. [s]
Necedah	USCRN	44.06	90.17	1 Jan 2008 – 31 Dec 2011	R_S	2
Shabbona	USCRN	41.84	88.85	1 Jan 2008 – 31 Dec 2011	R_S	2
Chatham	USCRN	46.33	86.92	1 Jan 2008 – 31 Dec 2011	R_S	2
Gaylord	USCRN	44.91	84.72	1 Jan 2008 – 31 Dec 2011	R_S	2
East Leland	EW	45.03	85.67	2 May 2003 – 31 Dec 2012	R_S, T_2, u_3, h_2	5
Escanaba	EW	45.86	87.18	17 Jun 2003 – 31 Dec 2012	R_S, T_2, u_3, h_2	5
Hart	EW	43.74	86.36	16 Aug 1996 – 31 Dec 2012	R_S, T_2, u_3, h_2	5
Hawks	EW	45.30	83.85	30 Nov 1999 – 31 Dec 2012	R_S, T_2, u_3, h_2	5
East Lansing	EW	42.67	84.49	1 Jan 1996 – 31 Dec 2012	R_S, T_2, u_3, h_2	5
Petersburg	EW	41.93	83.70	11 Nov 1999 – 31 Dec 2012	R_S, T_2, u_3, h_2	5
Pigeon	EW	43.90	83.27	20 Jul 2000 – 31 Dec 2012	R_S, T_2, u_3, h_2	5
South Haven	EW	42.36	86.29	6 April 2006 – 31 Dec 2012	R_S, T_2, u_3, h_2	5

distribution (BIAS of 3.40 MJ m^{-2} day^{-1} for the bottom 20% of the distribution, but -0.15 MJ m^{-2} day^{-1} for the top 20%). Thus, a decision was made to correct, at the hourly timescale, for bias and the tendency of NLDAS-2 to overestimate (underestimate) solar radiation values on the left (right) side of the observed distribution. The latter correction is referred to hereafter as a variance correction since it accounts for the underprediction of the variance of solar around the median; i.e., the range of solar values (on both hourly and daily timescales) is narrower in NLDAS-2 than the observations suggest. For irrigation scheduling, one must consider not only days within the middle of the distribution, but low and high ETo days as well (Kloss et al., 2012).

Before proceeding, a brief review of existing solar radiation correction methods is in order. In general, existing correction methods reduce bias but do not correct underestimation of temporal variability (e.g., diurnal, seasonal). A common method of bias-correction involves the multiplication of uncorrected solar radiation values by the ratio of a desired mean value to an uncorrected mean value (e.g., Cosgrove et al., 2003; Ngo-Duc et al., 2005). More sophisticated approaches have also been used in which regression models are linked with observed cloud cover to compute bias-correction coefficients (e.g., Sheffield et al., 2006; Weedon et al., 2010). Alternatively, empirical models have been used to rescale downward shortwave radiation using other atmospheric surface variables (e.g., daily maximum temperature) as inputs (e.g., Wei et al., 2014). Lastly, methods have been developed that combine bias correction with corrections for over- or under-estimation of clear-sky radiation and over- or under-estimation of cloudiness (e.g., Slater, 2016). The desire to control both solar radiation bias and variance, and to limit the variables necessary for the correction procedure to solar radiation alone (as opposed to, for example, solar radiation and cloud cover), motivates the development of a correction procedure distinct from those reviewed here.

The correction procedure, restricted to approximate daylight hours [1100–0000 UTC (0700–2000 EDT)] during the 1 April–30 September growing season (applicable to corn and soybeans, primary crops in the Great Lakes region), was developed for each hour as follows. First, four-year-mean values for each growing season day and each station were pooled together, yielding a total of 732 points (183 growing season days x 4 stations). Second, a linear regression model was developed, with the four-year-mean artifact-removed NLDAS-2 downward solar radiation (S_N) as the independent variable, and the four-year-mean observed downward solar radiation (S_O) as the dependent variable,

$$S_O = a + bS_N \tag{1}$$

where a and b are the interecept and slope of the linear regression line, respectively (provided in Table S1 in Supplementary material). Third, the linear regression model was applied to S_N to yield a bias-corrected NLDAS-2 downward solar radiation (S_{Nb}), using slope and intercept values calculated from Eq. (1),

$$S_{Nb} = a + bS_N \tag{2}$$

(Fig. 2b). Fourth, an updated linear regression model was developed, with S_O as the independent variable, S_{Nb} as the dependent variable, and data points limited to 1800 UTC (time of approximate peak solar radiation),

$$S_{Nb} = a' + b'S_O \tag{3}$$

where coefficients a' and b' are 1.2731 and 0.4422, respectively. This step was performed to evaluate the magnitude of over- and under-estimation of solar radiation near the tails of the distribution, around the time of peak solar radiation. Fifth, an adjustment factor was calculated as the ratio of S_{Nb} at the particular hour and S_{Nb} at the pivot point of the 1:1 and linear regression lines (S_{piv}; calculated as 2.28 MJ m^{-2} in the previous step). This adjustment factor was then applied to S_{Nb} to yield a bias- and variance-corrected solar radiation (S_{Nbv}),

$$S_{Nbv} = S_{Nb} * \left(\frac{S_{Nb}}{S_{piv}}\right) \tag{4}$$

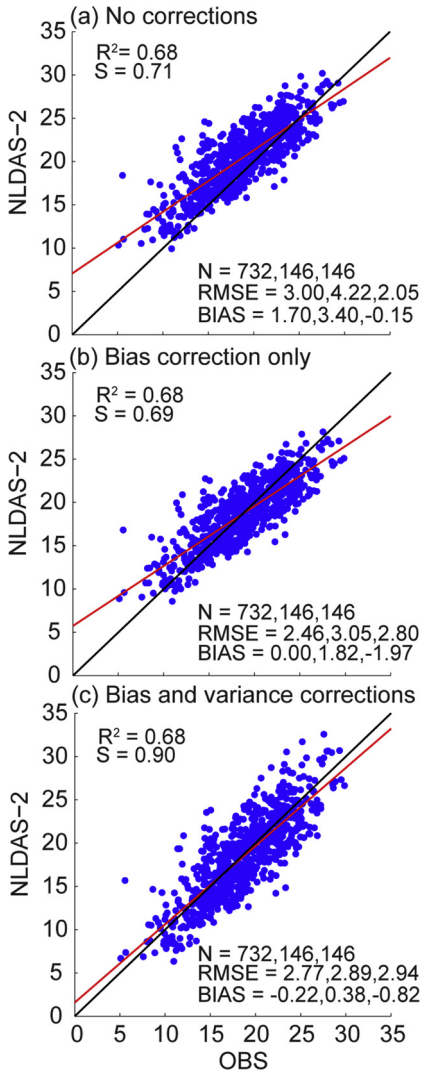

Fig. 3. Scatter plot diagrams of daily downward solar radiation [MJ m^{-2} day^{-1}], (a) uncorrected, (b) with bias correction only, and (c) with bias and variance corrections. Hourly measurements from four USCRN stations [Necedah, WI; Shabbona, IL; Chatham, MI; Gaylord, MI (Fig. 1, Table 1)] are converted to daily values, averaged across four years (2008–2011) and pooled (732 points = 183 growing season days x 4 stations). Black line denotes 1:1 axis and red line denotes linear regression. The following summary statistics are included in each panel: number of points (N), root mean squared error (RMSE), mean bias (BIAS), coefficient of determination (R^2), and slope of regression line (S). For N, RMSE, and BIAS, the statistics provided are for the whole dataset, the bottom 20% of the observed distribution, and the top 20% of the observed distribution, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Fig. 2c). This step was performed to reduce the bias for points closer to the tails of the observed distribution, and was only applied between 1700 and 2000 UTC (1300 and 1600 EDT: peak heating window, when the pivot point is approximately constant). Finally, the hour-specific bias and variance correction procedure was applied to the downward solar radiation variable at all grid points in the NLDAS-2 dataset, and the correction process then advanced to the next hour.

Before proceeding to a description of the FAO-56 ETo calculation, a few words of clarification are in order regarding the solar correction procedure. First, the slope and intercepts used to bias correct the raw NLDAS-2 downward solar radiation are functions of time of day, but not day of year. The underlying assumption of constant bias during the growing season likely limits our ability to eliminate bias, but has some support in the time series of the ratio of NLDAS-2 and USCRN downward solar radiation during the growing season (Fig. S2 in Supplementary material). Second, the correction procedure is based on only four years of data, increasing the likelihood that anomalous conditions in one year influence the corrections applied to all 30 years of data.

2.4. FAO-56 modified Penman-Monteith equation

For hourly reference evapotranspiration from a standardized grass crop (well-watered, 0.12 m crop height, surface resistance of

70 s m^{-1}, and albedo of 0.23), ETo [mm hr^{-1}] is expressed as:

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{37}{T_{hr}+273} u_2 [e^o(T_{hr}) - e_a]}{\Delta + \gamma(1 + 0.34u_2)} \quad (5)$$

(Eq. 53 in Allen et al., 1998). In Eq. (5), R_n is net radiation at the crop surface [MJ m^{-2} hr^{-1}], G is soil heat flux [MJ m^{-2} hr^{-1}], T_{hr} is mean hourly air temperature at 2-m height [°C], u_2 is wind speed at 2-m height [m s^{-1}], e^o is saturation vapor pressure at air temperature T_{hr} [kPa], e_a is mean hourly vapor pressure [kPa], Δ is the slope of the saturated water–vapor-pressure curve [kPa $°C^{-1}$], and γ is the psychrometric constant (kPa $°C^{-1}$).

The variables on the right-hand side of Eq. (5) were derived from NLDAS-2 surface or near-surface atmospheric variables (corrected downward solar radiation, 2-m temperature, 2-m specific humidity, and 10-m wind speed) following the method and procedure described in Chapters 3 and 4 of Allen et al. (1998). In short, downward solar radiation and 2-m temperature were used to estimate net radiation using an albedo appropriate to a short grass surface (0.23), the Stefan-Boltzmann law, and formulas for estimating atmospheric absorption of outgoing longwave radiation; 10-m wind speed was extrapolated to the 2-m level assuming a logarithmic wind profile; and 2-m saturation vapor pressure and actual vapor pressure were derived from 2-m temperature and specific humidity using empirical formulas based on the Clausius–Clapeyron equation. Variables were then calculated and input to Eq. (5) for each hour of the 1983–2012 dataset. Finally, for analyzing evaporative demand during the growing season, daily ETo was computed by integrating hourly values over a 0400 UTC–0400 UTC (midnight–midnight EDT) day. The process of deriving hourly and daily ETo is summarized in Fig. 2d.

Before proceeding, it is important to point out that strictly speaking, weather data used to compute ETo should be collected over or downwind of "dense, fully transpiring grass or similar vegetation exhibiting behavior similar to the definition of the reference surface condition" (Pereira et al., 2015); gridded ETo products cannot satisfy this requirement. Furthermore, effects of local irrigation on temperature and humidity, and thus ETo, are not accounted for in the development of gridded ETo products. This is true for all gridded ETo datasets, including the dataset developed in this study.

3. Results and discussion

3.1. Corrected solar and ETo assessment

The impact of the hourly bias and variance correction procedure on the NLDAS-2 daily downward solar radiation at the four USCRN stations is seen in Figs. 1 and 3. First, note that bias at all four stations is constrained to +/- 1 MJ m^{-2} day^{-1}, although the correction procedure has introduced negative bias at some stations (e.g., Necedah) (compare square marker edge and fill colors in Fig. 1). Second, comparing the full pooled dataset before and after bias corrections it is apparent that the bias correction procedure was successful in eliminating mean bias (BIAS), while also reducing root mean square error (RMSE) from 3.00 MJ m^{-2} day^{-1} to 2.46 MJ m^{-2} day^{-1} (cf. Fig. 3a and b). Nevertheless, the slope of the regression line (S) remains much less than 1.0 (S = 0.69), and consequently, BIAS for the bottom and top 20% of the observed distribution is 1.82 and -1.97 MJ m^{-2} day^{-1}, respectively. The impact of the combined bias and variance corrections is seen in Fig. 3c: BIAS for the bottom and top 20% of the observed distribution has been reduced to 0.38 and -0.82 MJ m^{-2} day^{-1}, respectively, and S has been increased to 0.9. The improvement of the statistics at the shoulders of the distribution does come at a price, however, as RMSE and BIAS for the entire distribution are degraded somewhat compared to the bias correction step alone (cf. Fig. 3b and c).

As this initial assessment was performed at the same sites and with the same dataset used to develop the correction procedure, analysis proceeds to further evaluation of the bias- and variance-corrected downward solar radiation at eight weather stations within the Enviro-weather (EW) mesonetwork in Michigan (Fig. 1, Tables 1 and 2). EW is an interactive information system linking real-time weather data, forecasts, and biological and other process-based models for assistance in operational decision-making and risk management associated with Michigan's agriculture and natural resource industries (Andresen et al., 2012). The need for weather station datasets containing downward solar radiation and the other components of Eq. (5) led to the choice of the EW network as the source of observations for this validation procedure. The earliest records in the EW mesonetwork date back to 1996.

Comparison of the BIAS columns in Table 2 reveals that the eight-station median daily solar radiation BIAS has been reduced from 2.04 MJ m^{-2} day^{-1} (11.23% of mean) to 0.42 MJ m^{-2} day^{-1} (2.34% of mean). Prior to the correction procedure, all stations exhibited a positive BIAS; after the procedure, the sign of BIAS is evenly split between the eight stations, although a small positive overall BIAS remains (Fig. 1; Table 2). To assess changes to daily solar radiation variance due to the correction procedure, the ratio of NLDAS-2 to observed standard deviation (SD) is computed, referred to hereafter as the SD ratio (SDR). Comparison of the SDR columns in Table 2 indicates that the correction procedure has indeed increased the variance of NLDAS-2 daily solar radiation; prior to the correction, the NLDAS-2 standard deviation was 81% of the observed value (as judged by the eight-station median SDR), whereas after the correction this statistic is 100%. Regarding RMSE, the correction procedure has had a smaller impact, reducing the eight-station median RMSE from 4.50 MJ m^{-2} day^{-1} (24.25% of mean) to 4.35 MJ m^{-2} day^{-1} (23.27% of mean). Two-sample t-tests confirm the statistical significance of the changes to BIAS and SDR at all stations; however, changes to RMSE are not universally significant (not shown).

Proceeding to an assessment of NLDAS-2 daily ETo at the EW stations (Table 3), the bias and variance correction of downward solar radiation is found to have a beneficial impact on all ETo statistics. Most notably, the correction procedure has reduced the eight-station median BIAS from 0.21 mm day^{-1} (6.25% of mean) to 0.06 mm day^{-1} (1.81% of mean). Following the correction procedure,

Table 2
Daily downward solar radiation [MJ m^{-2} day^{-1}] summary statistics: mean bias (BIAS), root mean square error (RMSE), standard deviation ratio (SDR: ratio of NLDAS-2 standard deviation to observed standard deviation; for BIAS and RMSE, the statistics normalized by the station mean are included as % BIAS and % RMSE, respectively. Note that N, the number of data points used in the analysis, varies between stations due to different periods of record (Table 1). Columns with dark gray (light gray) shading denote uncorrected (corrected) NLDAS-2 downward solar radiation. Statistics are computed for the 1 April – 30 September period only.

Station	N	BIAS (% BIAS)		RMSE (% RMSE)		SDR	
East Leland	1799	3.90 (23.07)	2.41 (14.26)	5.49 (32.49)	4.96 (29.34)	0.90	1.12
Escanaba	1738	2.86 (16.46)	1.12 (6.47)	4.87 (28.03)	4.43 (25.51)	0.85	1.05
Hart	2961	0.61 (3.12)	-1.11 (-5.70)	4.27 (21.95)	4.43 (22.81)	0.76	0.93
Hawks	2378	1.54 (8.36)	-0.33 (-1.82)	4.32 (23.50)	4.19 (22.77)	0.80	0.98
East Lansing	3099	1.61 (8.75)	-0.08 (-0.45)	4.14 (22.40)	3.94 (21.33)	0.79	0.98
Petersburg	2363	2.47 (13.71)	0.92 (5.13)	4.56 (25.33)	4.27 (23.72)	0.84	1.05
Pigeon	2260	1.45 (7.69)	-0.18 (-0.93)	4.43 (23.51)	4.53 (24.04)	0.83	1.03
South Haven	1270	2.56 (14.01)	1.11 (6.10)	4.56 (24.99)	4.01 (21.97)	0.79	0.98
median		2.04 (11.23)	0.42 (2.34)	4.50 (24.25)	4.35 (23.27)	0.81	1.00

Table 3
As in Table 2, but for daily ETo [mm day^{-1}].

Station	N	BIAS (% BIAS)		RMSE (% RMSE)		SDR	
East Leland	1798	0.16 (5.20)	0.04 (1.21)	0.78 (24.93)	0.73 (23.57)	0.88	0.95
Escanaba	1738	0.44 (13.99)	0.26 (8.16)	0.94 (29.56)	0.85 (26.84)	0.99	1.04
Hart	2892	-0.07 (-2.01)	-0.23 (-6.51)	0.86 (24.35)	0.85 (24.03)	0.81	0.87
Hawks	2371	-0.25 (-7.28)	-0.42 (-12.31)	0.88 (26.11)	0.90 (26.59)	0.77	0.81
East Lansing	3060	0.26 (7.29)	0.08 (2.41)	0.83 (23.52)	0.77 (21.73)	0.89	0.96
Petersburg	2303	0.45 (12.83)	0.30 (8.40)	0.88 (25.03)	0.82 (23.17)	0.94	1.03
Pigeon	2250	0.08 (2.22)	-0.08 (-2.09)	0.80 (22.35)	0.79 (21.99)	0.87	0.94
South Haven	1255	0.58 (17.55)	0.45 (13.45)	0.97 (29.10)	0.88 (26.54)	1.00	1.07
median		0.21 (6.25)	0.06 (1.81)	0.87 (24.98)	0.84 (23.80)	0.89	0.95

all but two sites feature BIAS less than 10% of the mean value, with all sites exhibiting BIAS smaller than what was reported in Abatzoglou (2013) (less than the median value of 0.5 mm day^{-1} reported therein). Comparison of the SDR columns indicates that the correction procedure has increased the variance of NLDAS-2 daily ETo; prior to the correction, the NLDAS-2 standard deviation was 89% of the observed value (as judged by the eight-station median SDR), whereas after the correction this statistic is 95%. Additionally, the eight-station median RMSE has been reduced slightly from 0.87 mm day^{-1} (24.98% of mean) to 0.84 mm day^{-1} (23.80% of mean). RMSE of about 20–30% has been reported for both USGS and NLDAS-2 daily ETo products (Liu et al., 2011; Moorhead et al., 2015). As with solar radiation, two-sample t-tests confirm the universal statistical significance of the changes to BIAS and SDR, but not the change to RMSE (not shown).

In conclusion, the preceding evaluation of NLDAS-2 solar radiation and ETo confirms that the secondary goal of this study has been achieved, that is, the correction of NLDAS-2 solar radiation has yielded a more accurate NLDAS-2 ETo dataset than one computed from uncorrected solar radiation.

3.2. Spatial analysis

Analysis of the NLDAS-2 ETo dataset begins with contoured plots of growing season mean daily ETo before and after application of the solar radiation correction (Fig. 4); recall that the growing season is defined here as 1 April–30 September. Unless otherwise specified, all figures presented in this study display data averaged over the period of record (1983–2012). The bias and variance correction steps described in Section 2.3 reduced mean daily ETo by an average of 5.49% (0.19 mm day^{-1}) across the domain, with the percent reduction varying from a minimum of 2.25% (0.09 mm day^{-1}) in northern Missouri to a maximum of 8.89% (0.26 mm day^{-1}) across central New York (see contoured lines in Fig. 4a).

In general, mean daily ETo tends to decrease in a southwest to northeast gradient across the region, with highest values in northern Missouri and lowest values in southwestern Quebec. The influence of water bodies and surface elevation on this pattern is evident as well, with relatively low values near the lakes (especially the northern lakes, e.g., Superior), and across the higher terrain from West Virginia to upstate New York (cf. Figs. 1 and 4). In addition, a ridge of higher mean ETo is apparent across southern Ontario north of Lakes Erie and Ontario. Overall, these results suggest greatest potential vegetative water needs in southern and western portions of the region, where modification of weather and climate by either large bodies of water or high terrain is minimal. The results also suggest that the Great Lakes play an important role in modulating spatial patterns of evaporative water demand. It is important to note that the spatial resolution of the new NLDAS-2 ETo dataset, combined with the increased accuracy resulting from the solar correction procedure, allows for analysis of spatiotemporal variability of ETo in the Great Lakes region in unprecedented detail.

Analysis now proceeds to an examination of subseasonal variability with monthly averages of daily ETo during the growing season (Fig. 5). Comparing the April and July panels, observe that mean daily ETo increases from a domain-average of 2.41 mm day^{-1} in April to a peak of 4.30 mm day^{-1} in July. The influence of the northern lakes (e.g., Superior) on daily ETo is particularly evident during June and July, when the relatively cool water temperatures (not shown) reduce evaporative demand across the northern Great Lakes region. Percent differences between July and April mean daily ETo (contoured lines in the July panel) vary from 50% to 95% across the domain, with no clear relationship between percent difference and proximity to water. However, the July-April percent differences are reduced at higher elevations; for example, percent differences in central Pennsylvania are 60–65%, compared to 65–75% in the lower-elevation eastern and western portions of the state (cf. Figs. 1 and 5). Lastly, evidence of the impact of the solar correction procedure on monthly mean ETo, and analyses of the spatial patterns of mean solar radiation and 2-m temperature (primary drivers of ETo; Eq. (5)], are shown in Figs. S3–S5, respectively, in the Supplementary material.

With the considerable spatial and subseasonal variability of daily ETo now established, patterns of submonthly variability are examined with contoured maps of SD normalized by the monthly mean (Fig. 6). Overall, the highest (lowest) normalized SD occurs in April (July), with normalized SD generally increasing from southwest to northeast across the domain. Outside of the northeast 20% of the domain (where values are consistently higher), normalized SD decreases from 0.35-0.5 during April to 0.25 or less during July, and increases again to 0.3-0.4 during September. This cycle in submonthly variability is closely related to the corresponding cycle in the magnitude and frequency of shortwave troughs and frontal systems (Sanders and Hoffman, 2002; Payer et al., 2011). As the jet stream weakens and translates northward during the transition from spring to summer, shortwave troughs and frontal systems become weaker and their passage less frequent, and day-to-day variability of the drivers of ETo [e.g., downward solar radiation; Eq. (5)] decreases with time; the opposite trend occurs as the jet stream strengthens and translates southward again during the transition from summer to autumn. On the other hand, the influence of the lakes on normalized SD is generally difficult to discern, except for September when normalized ETo is noticeably lower near the lakes. Lower normalized SD near the lakes in autumn reflects the muting effect of the relatively high heat capacity of the lakes on day-to-day variability of the drivers of ETo.

3.3. Point analysis

Examination of the NLDAS-2 ETo dataset now proceeds to analysis of temporal variability of ETo on timescales ranging from diurnal to interannual, at two locations chosen to represent interior and lakeshore climates: East Lansing, MI and East Leland, MI (Fig. 7; see Fig. 1 and Table 1 for locations). East Lansing is in the south-central part of Michigan's Lower Peninsula, approximately 100 km from the nearest Great Lake, whereas East Leland is in the center of the Leelanau Peninsula in northwest Michigan and is about equidistant (5 km) from Lake Michigan to the west and Grand Traverse Bay to the east (see inset panel in Fig. 1). Before proceeding, it important to note that the NLDAS-2 grid resolution (1/8 °, about 12 km) is approximately equal to the width of the

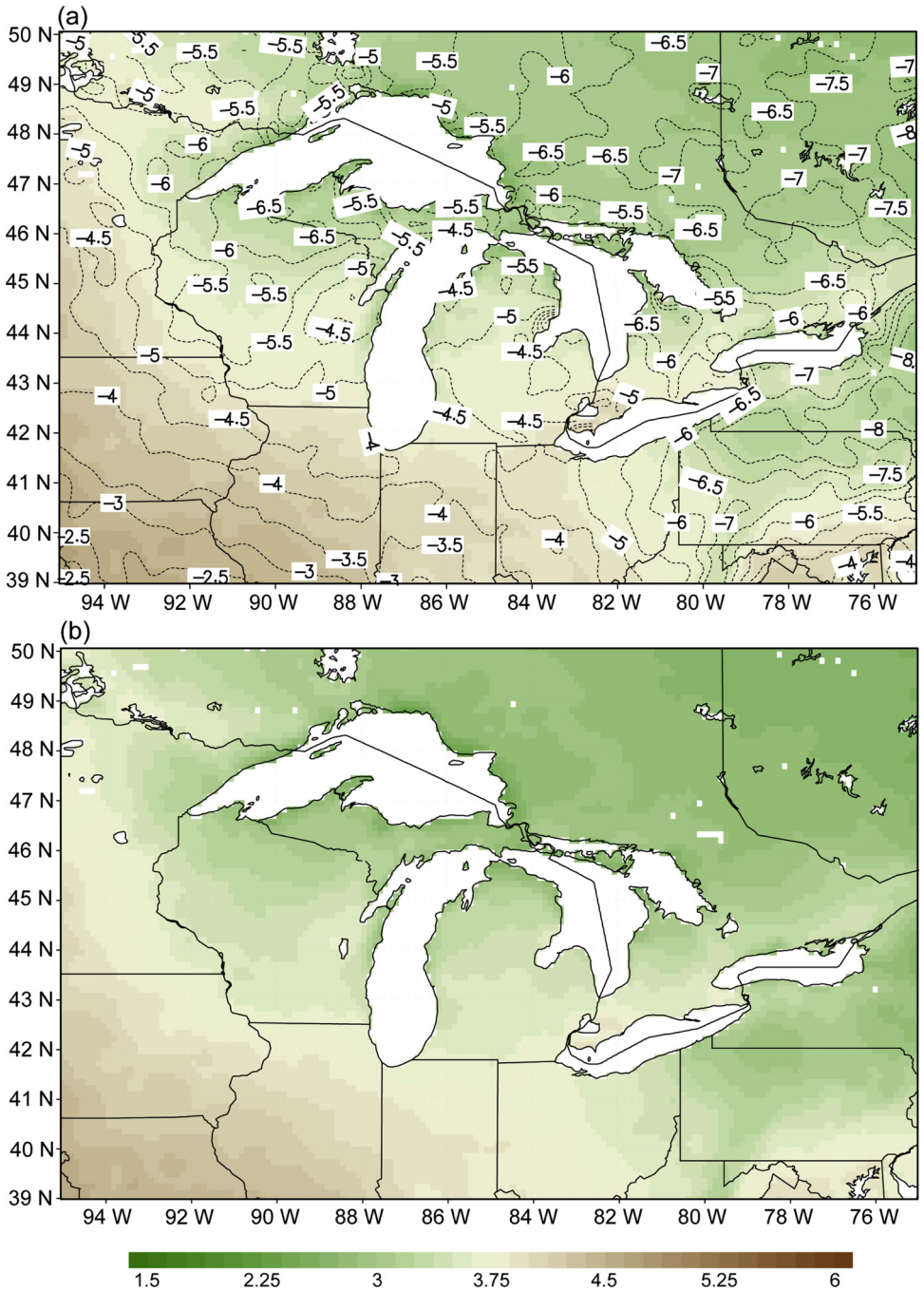

Fig. 4. Growing season mean daily ETo [mm day^{-1}] (a) before and (b) after solar correction (shaded). Percent difference (corrected – uncorrected) overlaid in (a).

Leelanau Peninsula at East Leland. Thus, caution must be exercised as the peninsula is represented by at most 3–4 grid cells, some of which extend partially over water.

Although this study is concerned primarily with daily ETo, it is important to recall that the NLDAS-2 ETo dataset is an hourly product. To help visualize the dataset at the hourly timescale, box and whisker plot time series of hourly ETo are presented in Fig. 7a. Examination of Fig. 7a reveals several prominent characteristics. First, a simple diurnal cycle of ETo is evident at both sites, consisting of a minimum in median ETo and box and whisker plot length between 0900 and 1000 UTC (0500 and 0600 EDT), around sunrise, and a maximum in median ETo and box and whisker plot length at 1800 UTC (1400 EDT), around solar noon. Second, during the day the median hourly ETo as well as the box and whisker plot length is greater at East Lansing than East Leland; the longer plots indicate

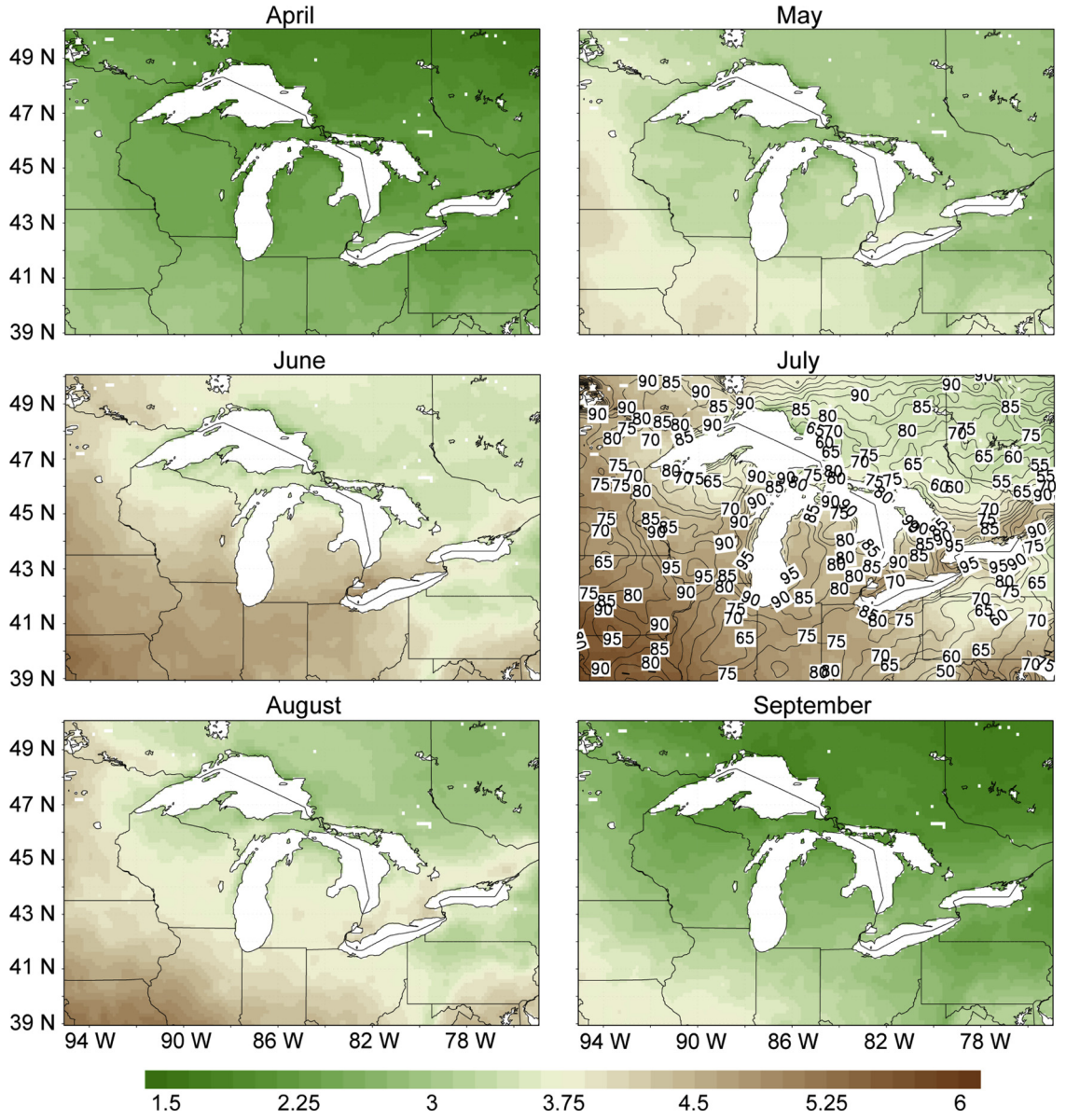

Fig. 5. Monthly mean daily ETo [mm day^{-1}] (with solar corrections), during 1 April–30 September growing season. The percent difference between July and April mean daily ETo is overlaid on the July panel.

greater subseasonal and/or interannual variability at the inland station. The weaker variability at East Leland can be attributed to the proximity of the station to bodies of water to its west and east; the relatively high heat capacity of water moderates the temperature variability at adjacent land points (Legates and Willmott, 1990). Maximum hourly ETo values are just under 1 mm hr^{-1} at East Lansing and just under 0.8 mm hr^{-1} at East Leland; such high values are associated with abnormally warm days during mid-summer. Lastly, the kink in the time series between 2000 and 2200 UTC (1600 and 1800 EDT) is an artifact of the variance correction procedure (Section 2.3): the adjustment was only applied between 1700 and 2000 UTC (1300 and 1600 EDT), i.e., during the period of approximate peak heating.

To compare the growing season cycle of daily ETo at inland and coastal stations, box and whisker plot time series of daily ETo from 1 April to 30 September are presented in Fig. 7b. Consistent with the growing season solar radiation cycle (not shown), the 30-year median daily ETo increases from 1 April until approximately early June, exhibits weak month-to-month variation through approximately early August, and steadily declines thereafter. Submonthly variability of 30-year median daily ETo is highest from about mid-April to early June, and is lower at East Leland than East Lansing, consistent with the closer proximity of East Leland to the stabilizing influence of Lake Michigan. The larger fluctuations in 30-year median daily ETo during spring reflect both the influence of large-scale weather patterns and the impact of one or two anomalous years in the 30-year study period on the trends. The fluctuations

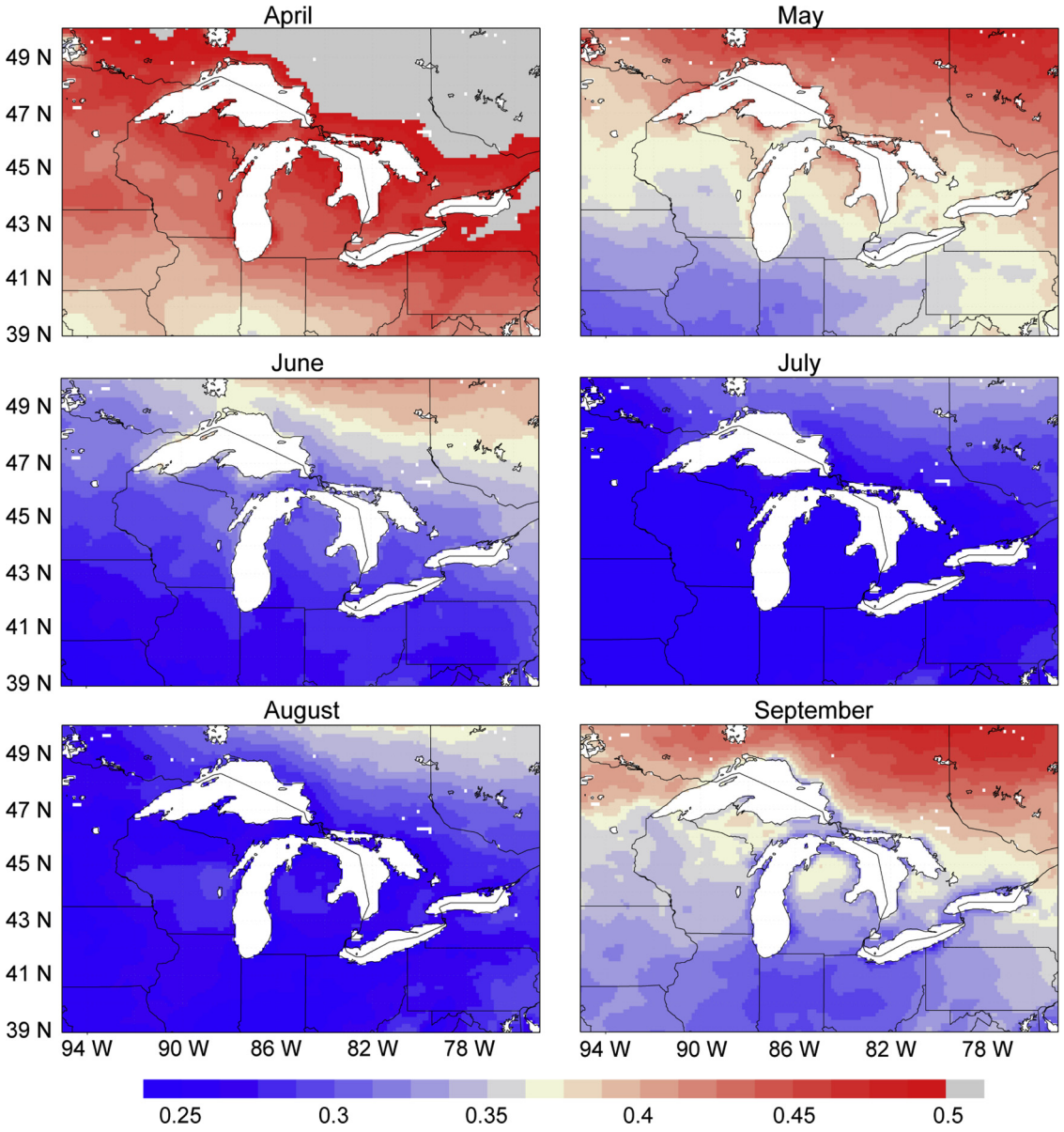

Fig. 6. Standard deviation of daily ETo normalized by monthly mean value [unitless].

of the 30-year median daily ETo, as well as the longer boxes and whiskers, suggest a greater influence of shortwave troughs and frontal systems on submonthly ETo variability during spring; recall the discussion of normalized SD (Fig. 6) and shortwave troughs in Section 3.2. Such variability is reduced during the second half of the growing season, most prominently at East Leland. Lastly, scattered daily ETo values of 7–10 mm day^{-1}, evident from late June to mid-July, occurred in two years only: 1988 and 2012, both of which were substantially impacted by extended drought (Fuchs et al., 2015).

Next, histograms are presented in Fig. 7c to examine the distribution of daily ETo during the growing season; a comparison of the histograms at East Lansing and East Leland reveals considerable differences between inland and coastal station sites. First, both the median and mean daily ETo are about 15% lower at East Leland than East Lansing, a reflection of the spatial gradient of daily ETo between the lakeshore and interior (Fig. 4b). Second, the distribution appears quasi-normal at East Lansing and skewed at East Leland. The skewness statistic confirms this observation, with a value of 0.63 at East Lansing and 0.72 at East Leland. The distribution at East Leland is biased toward lower daily ETo values, with a sharp drop off on the right side of the distribution. The quasi-normal and skewed distributions at East Lansing and East Leland, respectively, were also observed at other inland and coastal station sites examined in this study (not shown), and highlight the importance of lake proximity to weather and climate (Legates and Willmott, 1990). Note also that despite the quasi-normal distribution at East Lansing, a long tail is present on the right side of the distribution; the daily ETo values in the tail are representative of hot, sunny days during drought years, conditions generally absent at East Leland.

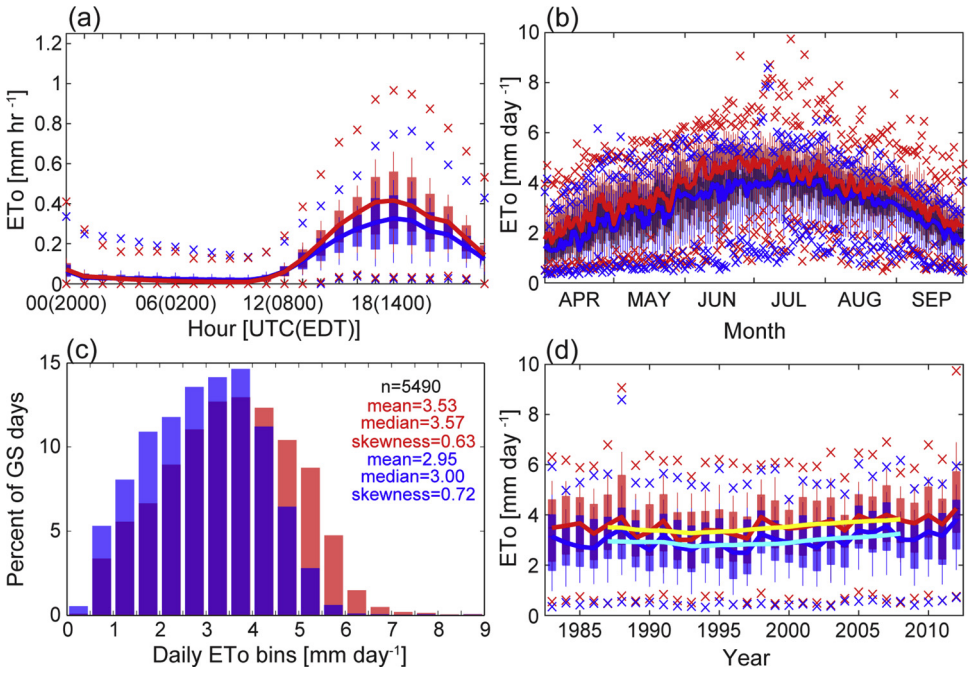

Fig. 7. Point analysis of NLDAS-2 ETo at East Lansing, MI (red) and East Leland, MI (blue): (a) Box and whisker plot time series of hourly ETo, with hour on the x-axis (5490 data points per plot: 183 growing season days x 30 years); (b) box and whisker plot time series of daily ETo with growing season day on the x-axis (30 data points per plot: 30 years); (c) histograms of daily ETo, with summary statistics; (d) box and whisker plot time series of daily ETo, with year on the x-axis (183 data points per plot: 183 growing season days). In panels (a), (b), and (d), the thick line denotes the median, the boxes extend outward to the 25th and 75th percentiles, the whiskers extend outward to the 10th and 90th percentiles, and the "x" symbols indicate the minimum and maximum values. In panel (c), units on y-axis are percent of growing season days. In panel (d), the yellow and cyan lines denote the 9-year moving average of the median at East Lansing and East Leland, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Station sites adjacent to the lakeshore are less likely to experience hot, full-sun days during summer than station sites well inland.

Lastly, questions regarding interannual variability and multi-year trends are addressed with box and whisker plot time series of daily ETo during the period 1983–2012 (Fig. 7d). Examination of the time series at East Lansing reveals two anomalous growing seasons, 1988 and 2012, in which the box and whisker plots extend well above those of the other 28 years (recall earlier discussion of Fig. 7b). Record warmth and notable droughts occurred during both growing seasons; the combination of minimal cloud cover and high temperatures yielded anomalously high daily ETo values. Although evidence of the 1988 anomalous growing season is present in the East Leland time series, the 2012 box and whisker plot is unremarkable compared to other years. A somewhat less obvious feature of Fig. 7d is a period of reduced daily ETo from 1992 to 1997, best seen in the 75th percentiles at East Lansing; note that the minimum in the 9-year moving average (cyan and yellow lines) occurs in 1993 at both East Lansing and East Leland. For clarification, the 9-year period was chosen to filter out higher frequency noise from the annual data and help identify decadal or longer period trends. Beginning with 1993, a positive trend in 9-year average daily ETo is evident through 2008, the last year for which the 9-year moving average is computed; however, the trend appears to continue through the end of the study period in 2012.

3.4. Crop irrigation example

An example application of the NLDAS-2 ETo dataset in irrigation management is now provided: an irrigation climatology for field maize (grown for grain) at East Lansing, MI. Such information may be useful in planning to determine potential irrigation water needs, system design, and operational capacity. Use of ETo in managing irrigation for a specific crop requires the use of a crop coefficient (Kc) to account for differences in the crop canopy and aerodynamic resistance between the crop in question and the grass reference crop (Allen et al., 1998). The product of ETo and Kc is referred to as crop evapotranspiration under standard conditions (ETc), where standard conditions are defined as large fields, optimum soil water, excellent management practices and environmental conditions, and full production.

In this example, the following timeline is assumed: planting date, 1 May; beginning of crop cover, 10 May; effective full cover, 10 July–1 August; maturity, 31 August. Following the procedure outlined in Allen et al. (1998), a crop coefficient curve is constructed using three values of Kc: Kc_{ini} (initial: beginning of crop cover), Kc_{mid} (mid-season: effective full crop cover), and Kc_{end} (end of season: maturity, harvest). Using Table 12 in Allen et al. (1998), and assuming the crop is harvested dry, the three values of Kc are as follows: Kc_{ini} = 0.3, Kc_{mid} = 1.2, and Kc_{end} = 0.35. Linear interpolation of Kc between stages yields the crop coefficient curve presented in Fig. 8a (compare to Fig. 34 in Allen et al., 1998). Multiplication of the daily ETo values at East Lansing in Fig. 7b by the Kc values in

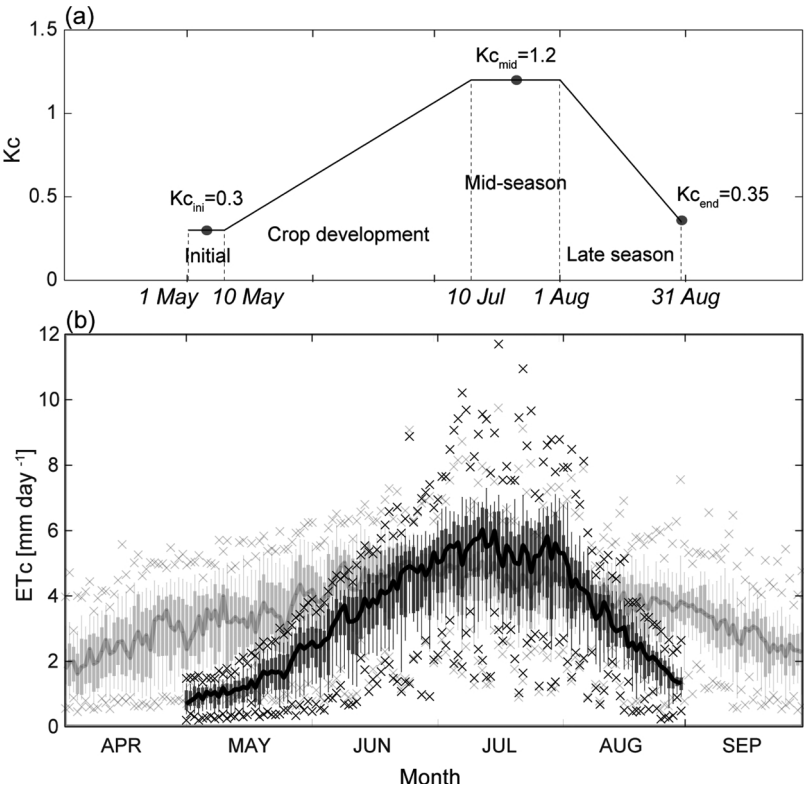

Fig. 8. (a) Crop coefficient curve for field maize and (b) box and whisker plot time series of ETc at East Lansing, MI, assuming the following timeline: planting date, 1 May; beginning of crop cover, 10 May; effective full cover, 10 July–1 August; maturity, 31 August. ETo is reproduced from Fig. 7b and included in the background in (b). Note that each box and whisker plot contains 30 data points (30 years). The thick line denotes the median, the boxes extend outward to the 25th and 75th percentiles, the whiskers extend outward to the 10th and 90th percentiles, and the "x" symbols indicate the minimum and maximum values.

Fig. 8a yields the box and whisker plot time series of daily ETc in Fig. 8b. Application of the Kc scaling factor results in decreased variability during the early and late portions of the growing season, and increased variability during the heart of the growing season in June and July.

Next, accumulated irrigation during the field maize season is examined for a wet and a dry season (2004 and 2012, respectively). First, daily ETo and precipitation measured at the East Lansing EW station during the 2004 and 2012 seasons are input into Microsoft Excel-based irrigation scheduling software (available from http://www.agweather.geo.msu.edu/mawn/irrigation/Irr_cbook_June2014.xls), yielding daily irrigation as output. The software used in this example is based on an irrigation scheduling spreadsheet presented in Allen et al. (1998; Annex 8), and assumes that available water in the root zone is partially depleted before irrigation water is applied (i.e., management allowed depletion of root zone water). The software indicates when supplemental water must be added to match the irrigation demand, for a specific crop type and growing season stage. In this exercise, simulated irrigation applications of 25.4 mm on a given day are initiated whenever daily volumetric soil moisture in the plant rooting zone drops to 50% or less of its maximum value. The soil profile used is Spinks loamy sand, a coarse-textured, commonly irrigated soil type in Michigan with a total of 116 mm of plant extractable water in the top 100 cm of the soil profile. The software internally computes daily ETc from daily ETo using Kc values similar (but not identical) to those in Fig. 8a. Second, NLDAS-2 daily ETo and precipitation are input, for each season between 1983 and 2012, into the irrigation scheduling software, yielding a 30-year climatology of daily irrigation. Accumulation of the daily irrigation totals yields the time series of accumulated irrigation seen in Fig. 9, with the 2004 (green line) and 2012 (brown line) station-data–driven accumulated irrigation overlaid on the NLDAS-2–driven accumulated irrigation climatology (box and whisker plots). In the following discussion, it is assumed that a hypothetical irrigation manager applies water based only on the station-data–driven software output. The user compares the irrigation progress at various points in the 2004 and 2012 seasons to the 30-year climatology provided by NLDAS-2.

Examination of Fig. 9 reveals several notable features. First, a comparison of the station-data–driven irrigation time series shows that considerably more irrigation water was applied during 2012 (279.4 mm) than 2004 (101.6 mm). For reference, the statewide average applied irrigation reported in the 2003, 2008, and 2013 Farm and Ranch Irrigation Surveys varied between 152.4 and 182.9 mm (National Agricultural Statistics Service, 2004, 2009, 2014). The anomalously large amount of water applied in 2012 is consistent with the much higher ETo during the 2012 season (compare 2004 and 2012 in Fig. 7d), as well as much lower rainfall amounts. Observed 1 May–31 August precipitation at East Lansing was 482.87 mm in 2004, but only 214.37 mm in 2012.

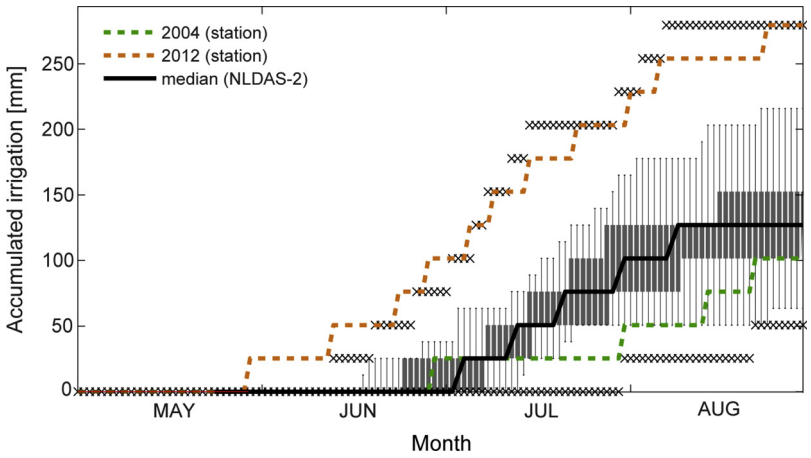

Fig. 9. Accumulated irrigation during the 1 May– 31 August field maize season, based on observed daily ETo and precipitation at East Lansing, MI, for 2004 (green dashed line) and 2012 (brown dashed line). Box and whisker plots depict the accumulated irrigation during the 1 May–31 August field maize season during the period 1983–2012, based on NLDAS-2 daily ETo and precipitation. The thick line denotes the median, the boxes extend outward to the 25th and 75th percentiles, the whiskers extend outward to the 10th and 90th percentiles, and the "x" symbols indicate the minimum and maximum values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Comparison of the 2004 and 2012 observed precipitation at East Lansing to the 1981–2010 average of 327.15 mm at the Lansing Capital City Airport (about 14 km northwest of East Lansing at 42.779 N, 84.587 W) reveals the scale of the precipitation anomalies in 2004 and 2012. Second, it appears that differences between the two years develop early in the season. Specifically, the first application of water in 2012 occurs in early June, about one month earlier than in 2004, well before the median initial application date of 2 July (see thick black line in Fig. 9). This suggests that early application of irrigation water may be, at least in some cases, a precursor to a dry season with heavy irrigation need. Third, variability in the NLDAS-2–driven irrigation climatology increases noticeably during the later portion of the season, on or about 21 July. Such knowledge of increased temporal variability during the later stages of the growing season may be of use in anticipating water needs for an upcoming season and/or planning the capacity of an irrigation system.

4. Conclusions

In this study, a gridded Great Lakes region historical reference evapotranspiration dataset for the period 1983–2012 was developed by inputting NLDAS-2 forcing fields into the FAO-56 ETo equation. Correction of NLDAS-2 solar radiation was shown to yield a more accurate NLDAS-2 ETo dataset than one computed from uncorrected solar radiation. The spatial resolution of the new NLDAS-2 ETo dataset, combined with the increased accuracy resulting from the solar correction procedure, allowed for analysis of spatio-temporal variability of ETo in the Great Lakes region in unprecedented detail. An analysis of the spatial patterns of evaporative demand across the Great Lakes region depicted spatial patterns heretofore unseen, with relatively low ETo values near the lakes, and highest values in southern and western portions of the Great Lakes region, where modification of weather and climate by either large bodies of water or high terrain was minimal. Consistent with the utility of the new ETo dataset to depict spatial variability of evaporative demand, point analysis at two locations chosen to depict lakeshore and interior climates provided insight into temporal variability of ETo not possible with existing station and gridded ETo datasets.

This study has shown the utility of the NLDAS-2 ETo dataset in describing spatial and temporal patterns of evaporative demand across the Great Lakes region; however, the limitations of the study must be kept in mind. First, the correction procedure utilized climate station data from only four stations and was limited to a four-year period, and the coefficients used to bias-correct NLDAS-2 solar radiation were functions of hour of day, but not day of year. Second, the 1/8-degree grid spacing limits the accuracy of the NLDAS-2 ETo dataset, especially near strong physiographic gradients of soils, vegetation, and climate. Third, no gridded dataset can satisfy the requirement that weather data used to compute ETo be collected over or downwind of vegetation with similar behavior to the FAO-56 reference grass definition. However, despite these limitations, the NLDAS-2 ETo dataset is expected to be of use to the agricultural and hydrological communities. Future work includes designing and developing a data extraction utility, linking the ETo dataset with irrigation/water scheduling management applications for on-line user access, developing additional gridded 'integrative' variables (e.g. leaf wetness), and considering projected future trends. Efforts are underway at the time of publication to develop an online access portal on the Great Lakes Integrated Sciences + Assessments (GLISA) website: http://glisa.umich.edu; those interested in accessing the NLDAS-2 ETo dataset are encouraged to check the GLISA News and Events page (http://glisa.umich.edu/news) for updates. Until the data portal is active, all data requests should be forwarded to Enviro-weather at eweather@msu.edu.

Finally, potential applications of the new ETo dataset include, but are not limited to use in agriculture as a tool for estimating potential irrigation water needs and informing irrigation system design, use in hydrology as a tool for managing groundwater

supplies and assessing drought potential, and in both disciplines as a baseline for assessing surface and groundwater changes under future climate scenarios. More broadly, the NLDAS-2 ETo dataset presented here has the potential to serve as a valuable tool for researchers studying various aspects of agricultural climatology and hydroclimatology.

Declarations of interest

None.

Acknowledgements

Funding for this work was provided by Michigan State University AgBioResearch via project no. MICL02274, and by the GLISA Center through NOAA-OAR-CPO grants NA10OAR4310213 and NA15OAR4310148. The color map used for the station symbols in Fig. 1 was developed by Cynthia Brewer at the Pennsylvania State University (http://colorbrewer2.org/). Finally, comments and suggestions from three anonymous reviewers were helpful in revising the manuscript and are greatly appreciated.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrh.2019.100606.

References

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modeling. Int. J. Climatol. 33, 121–131. https://doi.org/10.1002/joc.3413.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage Paper no. 56, Rome, Italy. http://www.fao.org/documents/card/en/c/5de3a877-5547-5b23-9e6f-0c8951c54b5e.

Andresen, J.A., Winkler, J.A., 2009. Weather and climate. In: Schaetzl, R.J., Brandt, D., Darden, J.T. (Eds.), Michigan Geography and Geology. Pearson Custom Publishing, Boston, pp. 288–314.

Andresen, J., Olsen, L., Aichele, T., Bishop, B., Brown, J., Landis, J., Marquie, S., Pollyea, A., 2012. Enviro-weather: a weather-based pest and crop management information system for Michigan. March 27-29. Proceedings of the 7th International Integrated Pest Management Symposium, Center for Innovation in Teaching & Learning. https://ipmsymposium.org/2012/IPM_12_Proceedings_final.pdf.

Bureau of the Census, 1986. 1984 Farm and Ranch Irrigation Survey. 124 pp.. US Department of Commerce. http://usda.mannlib.cornell.edu/usda/AgCensusImages/1982/03/4.pdf.

Cosgrove, B.A., Lohmann, D., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Marshall, C., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Meng, J., 2003. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. J. Geophys. Res. 108, 8842. https://doi.org/10.1029/2002JD003118.

Daley, R., 1991. Atmospheric Data Analysis. Cambridge University Press, Cambridge.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol. 28, 2031–2064. https://doi.org/10.1002/joc.1688.

De Pondeca, M.S.F.V., Manikin, G.S., DiMego, G., Benjamin, S.G., Parrish, D.F., Purser, R.J., Wu, W.-S., Horel, J.D., Myrick, D.T., Lin, Y., Aune, R.M., Keyser, D., Colman, B., Mann, G., Vavra, J., 2011. The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: current status and development. Weather Forecast. 26, 593–612. https://doi.org/10.1175/WAF-D-10-05037.1.

Fuchs, B., Wood, D., Ebbeka, D., 2015. From Too Much to Too Little: How the Central U.S. Drought of 2012 Evolved Out of One of the Most Devastating Floods on Record in 2011. 106 pp.. Drought Mitigation Center Faculty Publications No. 118. http://digitalcommons.unl.edu/droughtfacpub/118.

Hobbins, M.T., Geli, H.M.E., Lewis, C., Senay, G.B., Verdin, J.P., 2013. NOAA introduces its first-generation reference evapotranspiration product. December 9-13. Transactions, American Geophysical Union Fall Meeting 2013.

Hubbard, K.G., Lin, X., Baker, C.B., 2005. On the USCRN temperature system. J. Atmos. Ocean. Technol. 22, 1095–1100. https://doi.org/10.1175/JTECH1715.1.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project. Bull. Am. Meteorol. Soc. 77, 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kitchen, M., Blackall, R.M., 1992. Representativeness errors in comparisons between radar and gauge measurements of rainfall. J. Hydrol. 134, 13–33. https://doi.org/10.1016/0022-1694(92)90026-R.

Kloss, S., Pushpalatha, R., Kamoyo, K.J., Schütze, N., 2012. Evaluation of crop models for simulating and optimizing deficit irrigation systems in arid and semi-arid countries under climate variability. Water Resour. Manage. 26, 997–1014. https://doi.org/10.1007/s11269-011-9906-y.

Legates, D.R., Willmott, C.J., 1990. Mean seasonal and spatial variability in global surface air temperature. Theor. Appl. Climatol. 41, 11–21. https://doi.org/10.1007/BF00866198.

Lewis, C.S., Geli, H.M.E., Neale, C.M.U., 2014. Comparison of the NLDAS weather forcing model to agrometeorological measurements in the western United States. J. Hydrol. 510, 385–392. https://doi.org/10.1016/j.jhydrol.2013.12.040.

Liu, W., Hong, Y., Khan, S., Huang, M., Grout, T., Adhikari, P., 2011. Evaluation of global daily reference ET using Oklahoma's environmental monitoring network – MESONET. Water Resour. Manage. 25, 1601–1613. https://doi.org/10.1007/s11269-010-9763-0.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American Regional Reanalysis. Bull. Am. Meteorol. Soc. 87, 343–360. https://doi.org/10.1175/BAMS-87-3-343.

Moorhead, J., Gowda, P., Hobbins, M., Senay, G., Paul, G., Marek, T., Porter, D., 2015. Accuracy assessment of NOAA gridded daily reference evapotranspiration for the Texas High Plains. J. Am. Water Resour. Assoc. 51, 1262–1271. https://doi.org/10.1111/1752-1688.12303.

National Agricultural Statistics Service, 2004. 2003 Farm and Ranch Irrigation Survey. 216 pp.. US Department of Agriculture. http://usda.mannlib.cornell.edu/usda/AgCensusImages/2002/02/06/Complete%20Report.pdf.

National Agricultural Statistics Service, 2009. 2008 Farm and Ranch Irrigation Survey. 268 pp.. US Department of Agriculture. https://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Farm_and_Ranch_Irrigation_Survey/fris08.pdf.

National Agricultural Statistics Service, 2014. 2013 Farm and Ranch Irrigation Survey. 266 pp.. US Department of Agriculture. https://www.agcensus.usda.gov/Publications/2012/Online_Resources/Farm_and_Ranch_Irrigation_Survey/fris13.pdf.

Ngo-Duc, T., Polcher, J., Laval, K., 2005. A 53-year forcing data set for land surface models. J. Geophys. Res. 110, D06116. https://doi.org/10.1029/2004JD005434.

Payer, M., Laird, N.F., Maliawco, R.J., Hoffman, E.G., 2011. Surface fronts, troughs, and baroclinic zones in the Great Lakes Region. Weather Forecast. 26, 555–563. https://doi.org/10.1175/WAF-D-10-05018.1.

Pereira, L.S., Allen, R.G., Smith, M., Raes, D., 2015. Crop evapotranspiration estimation with FAO56: past and future. Agric. Water Manage. 147, 4–20. https://doi.org/10.1016/j.agwat.2014.07.031.

Pinker, R.T., Tarpley, J.D., Laszlo, I., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Lohmann, D., Cosgrove, B.A., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., 2003. Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) project. J. Geophys. Res. 108 (D22), 8844. https://doi.org/10.1029/2002JD003301.

Sadler, E.J., Camp, C.R., Hook, J.E., 2003. Irrigation management in humid regions. In: Stewart, B.A., Howell, T.A. (Eds.), Encyclopedia of Water Science. Marcel Dekker, New York, pp. 478–482.

Sakamoto, T., Gitelson, A.A., Wardlow, B.D., Verma, S.B., Suyker, A.E., 2011. Estimating daily gross primary production of maize based only on MODIS WDRVI and shortwave radiation data. Remote Sens. Environ. 115, 3091–3101. https://doi.org/10.1016/j.rse.2011.06.015.

Sanders, F., Hoffman, E.G., 2002. A climatology of surface baroclinic zones. Weather Forecast. 17, 774–782. https://doi.org/10.1175/1520-0434(2002)017<0774:ACOSBZ>2.0.CO;2.

Senay, G.B., Verdin, J.P., Lietzow, R., Melesse, A.M., 2008. Global daily reference evapotranspiration modeling and evaluation. J. Am. Water Resour. Assoc. 44, 969–979. https://doi.org/10.1111/j.1752-1688.2008.00195.x.

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. J. Clim. 19, 3088–3111. https://doi.org/10.1175/JCLI3790.1.

Slater, A.G., 2016. Surface solar radiation in North America: a comparison of observations, reanalyses, satellite, and derived products. J. Hydrometeor 17, 401–420. https://doi.org/10.1175/JHM-D-15-0087.1.

Weedon, G.P., Gomes, S., Viterbo, P., Oesterle, H., Adam, J.C., Bellouin, N., Boucher, O., Best, M., 2010. The WATCH forcing data 1958–2001: a meteorological forcing dataset for land surface- and hydrological-models. Tech. Rep. 22 WATCH, 41 pp.

Wei, Y., Liu, S., Huntzinger, D.N., Michalak, A.M., Viovy, N., Post, W.M., Schwalm, C.R., Schaefer, K., Jacobson, A.R., Lu, C., Tian, H., Ricciuto, D.M., Cook, R.B., Mao, J., Shi, X., 2014. The North American carbon program multi-scale synthesis and terrestrial model intercomparison project—part 2: environmental driver data. Geosci. Model. Dev. 7, 2875–2893. https://doi.org/10.5194/gmd-7-2875-2014.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., Mocko, D., 2012a. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. J. Geophys. Res. 117, D03109. https://doi.org/10.1029/2011JD016048.

Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Duan, Q., Lohmann, D., 2012b. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. J. Geophys. Res. 117, D03110. https://doi.org/10.1029/2011JD016051.

Yoon, K.S., Yoo, K.H., Tyson, T.W., Curtis, L.M., 1993. Farmers' irrigation practices in a high rainfall area. Irrig. Drain. Syst. 7, 221–229. https://doi.org/10.1007/BF00881281.