A Physically Based Atmospheric Variables Downscaling Technique

TASNUVA ROUF, YIWEN MEI, AND VIVIANA MAGGIONI
Sid and Reva Dewberry Department of Civil, Environmental, and Infrastructure Engineering, George Mason University, Fairfax, Virginia

PAUL HOUSER
Department of Geography and Geoinformation Science, George Mason University, Fairfax, Virginia

MARGARET NOONAN
Sid and Reva Dewberry Department of Civil, Environmental, and Infrastructure Engineering, George Mason University, Fairfax, Virginia

(Manuscript received 20 May 2019, in final form 13 September 2019)

ABSTRACT

This study proposes a physically based downscaling approach for a set of atmospheric variables that relies on correlations with landscape information, such as topography, surface roughness, and vegetation. A proof-of-concept has been implemented over Oklahoma, where high-resolution, high-quality observations are available for validation purposes. Hourly North America Land Data Assimilation System version 2 (NLDAS-2) meteorological data (i.e., near-surface air temperature, pressure, humidity, wind speed, and incident longwave and shortwave radiation) have been spatially downscaled from their original 1/8\degree resolution to a 500-m grid over the study area during 2015. Results show that correlation coefficients between the downscaled products and ground observations are consistently higher than the ones between the native resolution NLDAS-2 data and ground observations. Furthermore, the downscaled variables present smaller biases than the original ones with respect to ground observations. Results are therefore encouraging toward the use of the 500-m dataset for land surface and hydrological modeling. This would be especially useful in regions where ground-based observations are sparse or not available altogether, and where downscaled global reanalysis products may be the only option for model inputs at scales that are useful for decision-making.

1. Introduction

Hyperresolution (from 100 m to 1 km globally) land surface modeling has recently become available and provides detailed information about the storage, movement, and quality of carbon and water at and near the land surface (Wood et al. 2011). Hyper-resolution land surface data are fundamental for water resources management and for making decisions related to agricultural productivity, crop yield prediction, and hydroclimatic hazards. These hyperresolution land surface data are expected to advance weather forecasting (Senatore et al. 2015), climate prediction (Baker et al. 2017), precise irrigation scheduling (Gibson et al. 2017), quantification of greenhouse gas fluxes (Franz et al. 2017), flood prediction (Maidment 2016), estimation of water scarcity (Zhou et al. 2016), and hydrologic simulations (Ko et al. 2019).

There are numerous challenges in developing a hyperresolution land modeling system, ranging from assessing adequate model physics and computing resources to the representation of human impacts on the land surface. A current barrier is developing a global dataset required to parameterize and dynamically force these models at hyperresolutions. Land surface models typically require a minimum of seven near-surface atmospheric forcing variables provided at every time step (e.g., hourly), including air temperature and humidity, wind speed, incident longwave and shortwave radiation, and precipitation. Additional pressure, precipitation type, and radiation variables may be required for some model classes but are generally easily deduced from the basic set of seven.

Corresponding author: Tasnuva Rouf, trouf@masonlive.gmu.edu

DOI: 10.1175/JHM-D-19-0109.1
© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
Currently, we do not have global subkilometer in situ or satellite observational capabilities from which to derive these forcing variables. Therefore, physical, dynamic, and statistical downscaling approaches have been developed that interpolate the required high-resolution fields from coarser-resolution data incorporating the interactions between the atmosphere and terrestrial surface (Cosgrove et al. 2003; Haylock et al. 2006; Liston and Elder 2006; Giotto et al. 2014; Sunyer et al. 2015; Gaur and Simonovic 2017). For precipitation downscaling, Venugopal et al. (1999) proposed dynamic space–time scaling of rainfall along with a spatial disaggregation scheme at subgrid scales. To account for orographic influences, Badas et al. (2006) considered a modulation function which superimposed to homogeneous and isotropic synthetic fields to take into account for spatial heterogeneity. Using satellite data, Zorzetto and Marani (2019) proposed a downscaling procedure to calculate the point rainfall extreme value distribution and relates it with the ground observation.

Cosgrove et al. (2003) proposed algorithms for developing 0.125°/hourly spatial/temporal resolution products from nine primary forcing fields at a native resolution of 40 km/3 hourly across North America. Their assumption was that the original 0.125° topography differs significantly within a 40-km grid cell and elevation could be used as the prime factor for downscaling temperature, pressure, specific humidity, and longwave radiation. The downscaled variables were successfully validated against ground observations from the Oklahoma Mesonet network and the Atmospheric Radiation Measurement Program Cloud and Radiation Testbed and Surface Radiation observation data.

Another framework was proposed by Liston and Elder (2006) who developed an intermediate-complexity, quasi–physically based, meteorological model (MicroMet) to produce high-resolution (1 km) atmospheric forcings (air temperature, relative humidity, wind speed, incoming shortwave radiation, incoming longwave radiation, surface pressure, and precipitation). They focused on complex terrain regions in Colorado, Wyoming, Idaho, Arctic Alaska, Svalbard, central Norway, Greenland, and Antarctica. Their downscaling approach applied a temperature–elevation relationship and used meteorological stations at hourly resolution.

Fiddes and Gruber (2014) proposed another physically based approach, TopoSCALE, to downscale coarse-grid climate variables to a finescale subgrid forcing data (<100 m), primarily based on a high-resolution digital elevation model (DEM). Elevation and topography correction were estimated by normalizing geopotential heights by gravity at sea level to downscale temperature, humidity, shortwave radiation, and wind speed. This method was tested across the Swiss Alps (characterized by large elevation gradient of 195–4634 m MSL) against a ground-based validation dataset.

More recently, Tao and Barros (2018) developed a framework to derive high-resolution long-term meteorological forcings for hydrologic modeling from mesoscale atmospheric reanalysis products, including topographic and cloud corrections and a new physical parameterization of near-surface wind fields. The downscaling methodology is applied to 3-hourly North American Regional Reanalysis (NARR) fields originally at 32-km spacing to 1-km/hourly resolution for seven years (2007–13) over the southeastern United States. The downscaled datasets were assessed against flux tower observations available in the region and performance statistics, and root-mean-square error (RMSE) had improved for all the variables.

This work builds upon these past studies and goes one step further by including several novelties to the approaches discussed above. We derived dynamic lapse rates based on air and dewpoint temperature data and elevation for the downscaling of near-surface air temperature and dewpoint temperature. These two temperature fields are subsequently used to correct air pressure, humidity and incident longwave radiation. For downscaling wind speed, we assume a log-wind profile and introduce the use of vegetation index for the parameterization of surface roughness and zero-plane displacement height. The downscaling shortwave radiation takes into consideration the optical air depth difference, local illumination, cast-shadowing, portion of the visible sky, and surface reflection to calculate direct, diffusive, and reflected shortwave radiation. These downscaling approaches have been applied to downscale the North American Land Data Assimilation System phase 2 (NLDAS-2; Cosgrove et al. 2003; Mitchell 2004) dataset (original resolution of 1/8°) to a 500-m grid across Oklahoma. The main reason for choosing 500 m as the target resolution is that most physical landscape parameters used to downscale the atmospheric variables are available at 500-m resolution. Achieving finer resolutions is possible, but the supporting landscape parameters would need to be downscaled or use higher-resolution physical parameter sources, adding the opportunity for more uncertainties and errors. Additionally, precipitation is one of the most important inputs in a hydrologic model. All the downscaled variables can be potentially used as predictands in a machine learning algorithm for downscaling precipitation, as shown by Mei et al. (2018). This manuscript represents the first step in this direction and focuses on the methodology to downscale all atmospheric variables except precipitation.
The present manuscript is organized as follows. Section 2 introduces the dataset and study area. A comprehensive step-by-step description of the downscaling algorithms is presented in section 3. The improvement in the downscaled atmospheric forcings with respect to the original resolution NLDAS-2 is assessed using ground observations and is discussed in section 4. Section 5 summarizes our conclusions.

2. Dataset

This work focuses on a domain in Oklahoma, in the midwestern United States during 2015. The region is characterized by a diverse landscape that includes the Great Plains, hills, lakes, and forests. This area is chosen because of the availability of high-quality, high-resolution ground-based data to validate the proposed downscaling approach.

a. NLDAS-2

The NLDAS-2 dataset is available at hourly temporal resolution and 1/8° grid spatial resolution for a period ranging from 1 January 1979 to present over the contiguous United States. The NLDAS-2 datasets are primarily derived from NARR (Mesinger et al. 2006) interpolated from the 32-km horizontal resolution NARR grid to the 1/8° NLDAS-2 grid, adjusted for elevation differences and temporally disaggregated from 3-hourly to hourly time scales (Cosgrove et al. 2003; Mitchell 2004). Surface downward shortwave radiation is derived by bias correcting NARR output using Geostationary Operational Environmental Satellite data (Pinker et al. 2003). The downscaling scheme proposed in this work is applied to the NLDAS-2 2-m air temperature, surface pressure, 2-m specific humidity, downward longwave and shortwave radiation, and wind speed at 1/8° spatial resolution and hourly temporal resolution. Surface albedo is required for the shortwave downscaling, and it is collected from the NLDAS-2 Noah land surface model. All NLDAS-2 variables are projected to 12.5 km under the USA Contiguous Lambert Conformal Conic coordinate system, approximating the original 1/8° resolution.

b. MERRA-2

The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Rienecker et al. 2011; Gelaro et al. 2017) has provided data since 1980. The advancement in the assimilation system replaces the original MERRA dataset that enables the merging of modern hyperspectral radiance and microwave observations, along with GPS radio occultation dataset. The resolution of MERRA-2 is 0.5° and 0.625° in latitude and longitude, respectively, and hourly. Surface roughness and zero-plane displacement height from the MERRA-2 dataset are used to support the downscaling scheme, as described in detail in section 3. These data have been interpolated to match the NLDAS-2 resolution of 12.5 km.

c. SRTM data

The Shuttle Radar Topography Mission (SRTM; CGIAR-CSI 2019) DEM data at 90-m spatial resolution, originally produced by the National Aeronautics and Space Administration (NASA), are a breakthrough in digital mapping and provide high-quality elevation data at the global scale (USGS 2006). The SRTM data are upscaled to 500 m, and NLDAS-2 (12.5 km) resolution using the average pooling operation. The SRTM elevation data play a major role in the proposed downscaling technique, as shown in the methodology section.

d. MODIS

The Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation indices products, MOD13Q1 and MYD13Q1 version 6, at 250-m/16-day resolution are used in the downscaling framework. By considering both MOD13Q1 and MYD13Q1, we obtain normalized difference vegetation index (NDVI) for every 8 days. The 8-daily NDVI is upscaled to 500 m and 12.5 km representing the target and original resolution.

e. Ground observations

The installation of the Oklahoma Mesonet network began in the 1980s, as the result of an ongoing collaboration between Oklahoma State University and the University of Oklahoma to develop a near-real-time, extremely reliable source of surface observational data on local weather conditions across the state (Oklahoma State University 2013a,b). As a result, a statewide network of 121 automated environmental monitoring stations was officially launched in March 1994. These provide regular measurements of air and soil temperature, barometric pressure, rainfall, relative humidity, solar radiation, wind speed and direction, and soil moisture, both direct and calculated, including instrumentation data, all of which are regularly transmitted every 5 min to the Oklahoma Climatological Survey, where the data quality is verified (Van der Veer Martens et al. 2017). Temperature, pressure, specific humidity, longwave radiation and shortwave radiation, and wind speed and direction observations at several stations across Oklahoma are used for validation purposes in this study (Fig. 1; University of Oklahoma 2019).

The Atmospheric Radiation Measurement (ARM) Program, supported by the U.S. Department of Energy,
had an intention to improve the understanding of processes that affect atmospheric radiation and the description of these processes in climate models. To achieve this goal, the ARM Program employed five highly instrumented primary measurement sites at land and ocean locations for up to 10 years, from the tropics to the Arctic. These provide a measurement of longwave and shortwave radiation, along with several meteorological variables including wind velocity, precipitation rate, surface moisture, temperature, and fluxes of sensible and latent heat (Stokes and Schwartz 1994). In 2015, for the launch of the NOAA Geostationary Operational Environmental Satellite R-Series satellite, the sites were improved with new capabilities (Michalsky and Lantz 2016). In this study, we have used longwave radiation data at the Oklahoma site, which are available at hourly resolution. The ground observations used in this study are from the Oklahoma Mesonet network and one ARM site, which are independent of both NLDAS and NARR data (Xia et al. 2012).

3. Methodology

The overarching idea behind the proposed methodology is that topography and vegetation cover have a large impact on surface atmospheric conditions. Therefore, we propose to use high-resolution topography and vegetation information to spatially downscale NLDAS-2 2-m air temperature, surface pressure, 2-m humidity, incident longwave/shortwave radiation, and 10-m wind speed from 12.5 km to a 500-m regular Cartesian grid across Oklahoma. The downscaling approach is composed of two main steps: 1) a statistical interpolation (bilinear interpolation) of all NLDAS-2 variables to match the downscaled terrain resolution of 500 m and 2) deterministic downscaling rules to account for the local scale effects.

a. Air and dewpoint temperature

The downscaling of air and dewpoint temperature are based on a lapse rate correction that accounts for the strong temperature–elevation relationship. These two variables are used to derive the 500-m pressure, humidity, and downward longwave radiation (that will be introduced in section 3b). While 2-m air temperature \( T \) is available from NLDAS-2, dewpoint temperature at the same altitude needs to be derived first. The 2-m dewpoint temperature \( T_d \) is calculated (in K) from vapor pressure using the (Lawrence 2005) method:

\[
T_d = \frac{C_3 \ln \left( \frac{E}{C_1} \right)}{C_2 - \ln \left( \frac{E}{C_1} \right)},
\]

where \( C_1, C_2, \) and \( C_3 \) are constants, and their values are (for water and ice, respectively) 611.21 and 611.15 Pa, 17.368 and 22.452, and 238.88 and 272.55°C adopted from Buck (1981). Variable \( E \) stands for vapor pressure (Pa) and is derived from the NLDAS-2 pressure and specific humidity by reorganizing the following equation that \( E = P/(0.622 + 0.37q) \), where \( P \) and \( q \) represent the NLDAS-2 air pressure (Pa) and specific humidity (kg kg\(^{-1}\)), respectively.

The lapse rate corrections of air and dewpoint temperature are defined as follows:
where a variable with a hat (\( \hat{\) denotes the downscaled resolution and without the original resolution. Variable \( T (K) \) is the NLDAS-2 2-m air temperature; \( Z \) and \( \hat{Z} \) (m MSL) are SRTM terrain elevation at 12.5-km and at 500-m resolution, respectively; and \( \Gamma \) and \( \Gamma_d \) (K m\(^{-1}\)) correspond to the temperature and dewpoint temperature lapse rates, respectively. The elevation and air temperature differences between a target grid cell and its eight nearest neighbors at each time step are calculated, and a line is fitted to describe the \( T-Z \) relationship. The slope of the fitted line is taken as the estimate of \( \Gamma \) for the target grid cell. This process is repeated for all grid cells and time steps and similarly for \( \Gamma_d \). As shown by maps of the lapse rate means and standard deviations, the average \( \Gamma \) is not a constant (although commonly assumed to be \(-6.5 \text{ K km}^{-1}\)) and the standard deviation is not null over time and across the study domain (Fig. 2). The \( \Gamma \) and \( \Gamma_d \) are computed based on NLDAS-2 because the goal is to produce a robust methodology that will generate high-resolution data from any coarse resolution product only with the help of elevation data and MODIS vegetation cover data, which are available all over the world at the target resolution. Figure 3 shows a comparison between the air temperature downscaling technique proposed here that adopts a dynamic lapse rate and a constant lapse rate. Correlation coefficients are also computed between the downscaled air temperature NLDAS-2 and the corresponding MESONET observations across Oklahoma during 2015. First, the correlation coefficient for the dynamic lapse rate case (0.97) is higher than the one obtained with a constant lapse rate (0.93), which is very close to the correlation between the original resolution NLDAS-2 and the in situ observations (0.94, as shown in Table 1), demonstrating minimal improvement due to the downscaling process in the latter case. Second, the downscaled temperature obtained assuming constant lapse rate is consistently higher than the observed temperature, showing an overestimation especially at low temperatures. Third, the deviation around the 1:1 line is largely reduced when the dynamic lapse rate is introduced.

b. Pressure, humidity, and incident longwave radiation

The downscaled 2-m air temperature is used to adjust the surface pressure using the hydrostatic approximation and the ideal gas law [following the methodology developed by Cosgrove et al. (2003)]:

\[
\hat{\rho} = P \exp \left( -\frac{g(\hat{Z} - Z)}{RT_m} \right),
\]

where \( R \) is the ideal gas constant (287 J kg\(^{-1}\) K\(^{-1}\)), \( g \) is the gravitational acceleration (9.81 m s\(^{-2}\)),

\[ \hat{T} = T + \Gamma(\hat{Z} - Z), \]

\[ \hat{T}_d = T_d + \Gamma_d(\hat{Z} - Z), \]
and $T_m$ (K) is the mean air temperature between the $T$ and $T_d$.

The specific humidity is calculated with $\hat{P}$ and $\hat{E}$ as follows:

$$\hat{q} = \frac{0.622 \hat{E}}{\hat{P} - 0.378 \hat{E}},$$  \hspace{1cm} (5)

where $\hat{E}$ is calculated based on the Magnus formula with $\hat{T}_d$ ($\hat{E} = C_1 e^{C_2 T_d/T_d + C_3}$).

The downscaling of downward longwave radiation is based on the Cosgrove et al. (2003) method, which is derived from the Stefan–Boltzmann law:

$$L = \frac{\hat{e}}{\varepsilon} \left( \frac{T}{T_d} \right)^4 L, \hspace{1cm} (6)$$

where $L$ is the NLDAS-2 incident longwave radiation (W m$^{-2}$), and $\varepsilon$ is emissivity. The empirical relationship suggested by Cosgrove et al. (2003) is applied here to calculate $\varepsilon$ and $\hat{e}$.

c. Incident shortwave radiation

The shortwave radiation downscaling technique comprises four steps from different methodologies. First, the global shortwave radiation $S$ is partitioned into direct $S_b$ and diffuse $S_d$ radiation based on the regression model proposed by Ruiz-Arias et al. (2010). This model estimates the solar transmissivity of the atmospheric column by defining the ratio of incident shortwave radiation between the surface level and the top of the atmosphere as the clearness index. Larger values of the clearness index indicate larger weighting for the direct shortwave. As a second step, $S_b$ is adjusted for optical air depth difference, local illumination, and cast-shadowing and $S_d$ is adjusted for sky obstruction (Fiddes and Gruber 2014; Tao and Barros 2018). In the third step, a reflected radiation component is estimated with the

| Table 1. Statistical metrics for each NLDAS-2 atmospheric forcing at their original and downscaled resolution (in italic) with respect to ground observations. |
|-------------------------------------------------|--|------------------|------------------|------------------|------------------|
|                                 | Correlation | Bias     | RMSE     | NSE     | Sample size |
| Temperature (K)                  | 0.94/0.97   | 3.94/1.95 | 5.09/2.69 | 0.74/0.93 | 1 028 409         |
| Pressure (hPa)                   | 0.94/1.00   | 11.6/1.52 | 13.8/2.75 | 0.56/0.99 | 1 030 356         |
| Relative humidity (%)            | 0.69/0.92   | 13.2/7.71 | 17.1/2.6  | 0.25/0.53 | 1 028 328         |
| Longwave radiation (W m$^{-2}$)  | 0.73/0.82   | 37.7/28.1 | 49.4/36.9 | 0.36/0.67 | 8627               |
| Shortwave radiation (W m$^{-2}$) | 0.89/0.93   | 63.7/52.3 | 119/101  | 0.81/0.86 | 1 029 358         |
| Wind speed (m s$^{-1}$)          | 0.67/0.70   | 2.62/1.75 | 3.04/2.19 | −0.27/−0.24 | 1 020 335        |
adjusted $S_b$ and $S_d$, albedo, and a terrain configuration factor (Tao and Barros 2018). Fourth, the downscaled global shortwave radiation $\hat{S}$ results from the summation of the three components:

$$\hat{S} = \delta \cos(\theta) e^{k(\bar{P} - P)} S_b + F_c S_d$$

$$+ AF_t [S_b + (1 - F_e) S_d] \cdot$$

where $\delta$ is a binary shadowing mask indicating whether the location is blocked by the shadow of nearby terrain. The presence of shadow implies that the local horizontal angle on the solar azimuth direction is higher than the solar altitude and vice versa. The $\cos(\theta)$ is the cosine of the solar illumination angle, which indicates if the sun is below or above the local horizon:

$$\cos(\theta) = \cos(\theta_r) \cos(\beta) + \sin(\theta_r) \sin(\beta) \cos(\theta_a - \alpha),$$

where $\theta_r$ and $\theta_a$ are the solar zenith and azimuth, respectively; $\alpha$ and $\beta$ stand for the terrain aspect and slope, respectively; and $k$ is the broadband attenuation coefficient (Pa$^{-1}$) calculated according to Sen Gupta and Tarboton (2016):

$$k = -\frac{\ln(S_T) - \ln(S)}{P},$$

where $S_T$ is the top-of-atmosphere incident shortwave radiation, computed based on $S_T = S^* \cos(\theta_r)$, with $S^*$ being the solar constant ($S^* = 1370 \text{ W m}^{-2}$; Dingman 2015). Variable $F_c$ is the sky-view factor indicating the portion of the visible sky of a location, and $F_t$ is the terrain configuration factor, which is a function of both sky-view factor and slope. Both $F_c$ and $F_t$ are calculated using the SAGA-GIS Sky View Factor Module (Häntzschel et al. 2005). Variable $A$ is the surface albedo from NLDAS-2. This study is conducted over Oklahoma, which is a region characterized by neither frequent snowfalls nor sharp topography. Therefore, the effects of subgrid surface heterogeneity on the albedo are minor. In areas where the frequency of snow is high, albedo at finer resolution should be considered to account for such heterogeneity.

d. Wind speed

The downscaling of wind speed $W$ (m s$^{-1}$) is based on an adjustment for friction velocity, terrain slope, aspect, and curvature (Tao and Barros 2018; Liston and Elder 2006). We start by assuming a logarithmic wind profile:

$$\hat{W} = \frac{\hat{u}_a}{\kappa} \ln \left( \frac{H - \hat{h}_0}{\hat{z}_0} \right).$$

where $\hat{W}$ is wind speed adjusted for the friction velocity $\hat{u}_a$ (m s$^{-1}$), $\kappa$ is the von Kármán constant ($\sim 0.41$), $H$ is the measurement height (m above ground), $\hat{z}_0$ is the surface roughness (m), and $\hat{h}_0$ is the zero-plane displacement height (m). The friction velocity at high resolution is calculated by taking advantage of the dependence of the geostrophic wind component on surface roughness and the assumption that the geostrophic wind component remains constant at different scales (Tao and Barros 2018):

$$\hat{u}_a = u_a \left( \frac{z_0}{\bar{z}_0} \right)^{0.09},$$

where $u_a$ is calculated from the NLDAS-2 wind speed $W$, MERRA-2 surface roughness $z_0$, and zero-plane displacement height $\bar{z}_0$ using Eq. (10). Surface roughness and zero-plane displacement height are related to the type and height of vegetation (Allen et al. 2007; Dong et al. 2001; Bastiaanssen 2000). There are some other factors on which surface roughness and displacement height depend, for example, land cover and snow cover. However, our analysis over Oklahoma found the vegetation index as the most important factor to downscale surface roughness. If other regions of the world are considered, the dependency of surface roughness and displacement height on such factors should be investigated. Given the difficulties in measuring vegetation characteristics, NDVI is often used as a proxy to estimate $z_0$ and $\bar{z}_0$ (Bastiaanssen 2000). In this study, a customized function for $\bar{z}_0$ based on the MODIS NDVI is developed for every time step:
\[ z_0 = e^{a_1 \text{NDVI} + a_2}, \]  \hspace{1cm} (12)

where the coefficients \( a_1 \) and \( a_2 \) are evaluated by log-transferring \( z_0 \) and NDVI at the NLDAS-2 resolution. Coefficients \( a_1 \) and \( a_2 \) have been calibrated for each time step over the entire region \((40 \times 72 \text{ pixels})\); however, future applications of this methodology in more heterogeneous areas should consider a spatial analysis of these two parameters. Figure 4 shows a histogram of the coefficients of determination \( R^2 \) between log-transferred \( z_0 \) and NDVI. Most of the \( R^2 \) values lie within 0.5–0.8, which reveals the strong correlation between the two variables. The term \( h_0 \) is estimated using a transfer function evaluated between \( z_0 \) and \( h_0 \) Nicholas and Lewis (1980):

\[ \log_{10} z_0 = a_3 \log_{10} h_0 + a_4. \]  \hspace{1cm} (13)

The downscaled wind speed is further adjusted by the terrain slope and curvature following Liston and Elder (2006):

\[ \hat{W}_t = \hat{W} \left( 1 + \frac{\Omega_x}{2} + \frac{\Omega_y}{2} \right), \]  \hspace{1cm} (14)

where \( \hat{W}_t \) is the final 500-m wind speed adjusted for the friction velocity and terrain characteristics (\( \text{m} \text{s}^{-1} \)). Parameters \( \Omega_x \) and \( \Omega_y \) are the slope in the wind direction and curvature of terrain, respectively, and are available at 500-m resolution—note that no hat is used for these symbols since they are not available at the NLDAS-2 native resolution of 12.5 km.

4. Results

The annual average for the six atmospheric variables over Oklahoma during 2015 is presented in Figs. 5 and 6 at the original NLDAS-2 resolution (left panels) and at the downscaled 500-m resolution (right panels). Specifically, Fig. 5 presents maps of 2-m air temperature, surface pressure, and 2-m specific humidity, whereas
Fig. 6 shows maps of downward longwave radiation, downward shortwave radiation, and 10-m wind speed. These figures highlight how the spatial patterns of the original resolution NLDAS-2 variables are retained in the downscaled maps. For instance, the temperature gradient (i.e., from colder northwest regions to the warmer southeast plains) is evident at both resolutions (Figs. 5a,b) and is consistent with the topography gradient illustrated in Fig. 1. However, the downscaled resolution variables are able to capture more detail, thanks to the inclusion of physiological feature information (e.g., orography, vegetation) in the downscaling techniques.

Five downscaled variables—air temperature, pressure, specific humidity, downward shortwave radiation, and wind speed—are then validated against ground observations at 118 Mesonet stations, whereas downward longwave radiation is validated against data collected at the ARM station. Specific humidity is converted to relative humidity for comparison with the measurements recorded by the Mesonet network. First, we present scatterplots of NLDAS-2 data (both at their native resolution and the downscaled products) against ground observations (Figs. 7 and 8). Second, we analyze maps of correlation coefficients across all sites in the study region to investigate the spatial variability of the downscaling technique performance (Figs. 9 and 10). Last, we compute overall statistics, including average correlation coefficient, additive bias, RMSE, and Nash–Sutcliffe model efficiency (NSE) to summarize the performance of the proposed approaches for each variable across Oklahoma (Table 1).

The downscaled temperature and pressure products align more closely with the Mesonet observations than the original resolution NLDAS-2 data, as shown by the density scatterplots in Fig. 7.
The dynamic lapse-rate correction largely improves both the air temperature and pressure estimates, moving them closer to the 1:1 line. The correlation coefficient maps presented in Figs. 9a–d corroborate this point: although correlations between the original resolution NLDAS-2 estimates and observations are already high, they are even higher when the downsampled temperature and pressure products are considered. These maps also show that there is no spatial bias in the downscaling algorithm performance since the improvement in the correlation coefficient is evident across the 118 Mesonet stations. For temperature,
FIG. 8. Density scatterplots of NLDAS-2 variables at (left) their original resolution and (right) the downscaled 500-m resolution against the corresponding Mesonet ground observations for (a),(b) downward longwave radiation $L$, (c),(d) downward shortwave radiation $S$, and (e),(f) 10-m wind speed $W$ at hourly resolution.
the correlation coefficient improves from 0.94 for the original resolution to 0.97 for the downscaled product, whereas for pressure, the correlation coefficient goes from 0.94 to 1.00 (Table 1). Although ground observations are quite representative of the coarse grid, the improvement in the correlation coefficients demonstrates that finer resolutions would be even more representative of the in situ observations.

The downscaled relative humidity slightly deviates from the observed data; however, the proposed methodology shows a large improvement over the original resolution NLDAS-2 data (Figs. 7e,f). Overall the correlation improves from 0.69 to 0.92. Similarly, to what observed for temperature and pressure, correlations improve consistently at all stations across the study domain (Figs. 9e,f).

The ARM site is used to validate the downscaled downward longwave radiation product. The sample size is smaller compared to the Mesonet stations, but still significant since data collected at the ARM sites are at hourly resolution and one year of data is considered, with a total of 8627 data points. Although the improvement in this variable is not as great as the previous ones, both the scatterplots (Figs. 8a,b) and the correlation maps (Figs. 10a,b) reveal an improvement in the downscaled product. Specifically, the downscaling scheme shifts the 500-m product closer to the ground observations, and the correlation improves from 0.69 to 0.82. Analogously to longwave radiation, the improvement in the downscaled shortwave radiation is not obvious from the scatterplot (Figs. 8c,d), but Figs. 9c and 9d show higher correlations in the downscaled product over the original NLDAS-2 data.

For wind speed, we do not notice a clear improvement in the downscaled product when analyzing the scatterplots in Figs. 8e and 8f. Similarly, the correlation values presented in Figs. 10e and 10f are very similar to each other at most of the Mesonet stations.
The overall statistics presented in Table 1 confirm the improved performance of the downscaled variables with respect to the original NLDAS-2, in terms of additive bias, RMSE, and NSE, computed against the hourly ground observations recorded at 118 stations (and for the ARM site for longwave radiation). RMSE of pressure is reduced from 13.8 hPa for the original resolution NLDAS-2 to 2.75 hPa for the downscaled product. Moreover, for the downscaled pressure, NSE gets close to 1, showing an almost perfect match with the corresponding ground observations. For longwave radiation, both bias (from 13.18 to 9.63 W m\(^{-2}\)) and RMSE (from 17.1 to 12.6 W m\(^{-2}\)) are lower thanks to the downscaling procedure. For shortwave radiation, bias and RMSE reduce from 63.73 to 52.30 W m\(^{-2}\) and from 119 to 100.68 W m\(^{-2}\), respectively, while the correlation coefficient and the NSE increase from 0.89 to 0.93 and from 0.81 to 0.86, respectively. For wind speed, the overall statistical analysis shows slightly improved results in terms of RMSE for the downscaled product (2.19 m s\(^{-1}\)) with respect to the original resolution NLDAS-2 wind speed (3.04 m s\(^{-1}\)).

5. Conclusions

This study presented a downscaling methodology for six atmospheric variables that are commonly used as land surface model inputs. The proposed downscaling framework comprises a statistical interpolation and a set of deterministic physical rules and is tested on the NLDAS-2 dataset across Oklahoma during 2015. The novelties introduced in this study with respect to previous work can be summarized as follows. First, we adopt a dynamic lapse rate of air and dewpoint temperature that is not constant in space and time. This step is extremely important since (i) lapse rate showed high variability across Oklahoma, and (ii) the dynamic lapse rate is used to downscale air temperature, which is then used to downscale air pressure and air humidity, thus, critically modifying the downscaling of all three...
variables. Second, the method adopted for downscaling shortwave radiation considers multiple factors, including a clearness index, local illumination, cast-shadowing, sky obstruction, and topographic configuration. Third, spatially downscaled surface roughness and zero-plane displacement height are introduced in the downscaling technique adopted for wind speed. In summary, this work combines different methodologies that have been previously proposed and modifies some of them to account for more factors.

To assess the improvement with respect to the original resolution dataset (at 12.5-km resolution), the 500-m downscaled products were compared to high-quality, high-resolution ground observations. The goal of this work is to prove the viability of the proposed methodology where high-quality, high-resolution continuous ground-based data are available, and the Oklahoma Mesonet network offers such a dataset.

For all the atmospheric variables, an improvement was observed in terms of several performance metrics (bias, correlation coefficient, RMSE, and NSE). Except for wind speed, the improvement in the other five downscaled products with respect to the original NLDAS-2 dataset was found substantial in terms of RMSE, correlation, and NSE. Thanks to the adopted dynamic lapse-rate correction, the downscaled temperature and pressure matched almost perfectly with the ground observations. Although the improvements in specific humidity, longwave radiation, and shortwave radiation were not as obvious, the statistical analysis revealed pointedly better performance of the downscaled products with respect to their native resolution version. Overall, the downscaled atmospheric products captured more spatial variability compared to the original NLDAS-2.

During the past few decades, land surface models have become popular for simulating surface energy and water fluxes in response to near-surface atmospheric forcing. The downscaled near-surface atmospheric forcing dataset developed in this study would facilitate the transition of these models to hyperresolution and improve our ability to monitor and predict the Earth’s terrestrial water, energy, and biogeochemical cycles at scales that are relevant for decision-making. Nevertheless, this study only tested the proposed methodology over Oklahoma and during one year. More validation exercises using quality-controlled observations and longer time series should be conducted to understand the potential of such an approach in different regions of the world. Although the downscaling approach here was applied only to the NLDAS-2 dataset, the proposed methodology could be applied to any other dataset. The methodology proposed in this study can be replicated in other regions of the world, by applying the downscaling technique to global datasets, such as the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2; Rienecker et al. 2011; Gelaro et al. 2017) and the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis Fifth Generation [ERA-5; contains modified Copernicus Climate Change Service (2017) information], as presented in Mei et al. (2018). Moreover, the downscaling spatial resolution is limited to 500 m for this work. This method can potentially be applied to finer resolutions if all parameters used in the proposed framework are made available at higher resolutions in the future.

Acknowledgments. T. Rouf, V. Maggioni, and P. Houser were supported by the NASA Science Utilization of the Soil Moisture Active-Passive Mission Program (Award Number NNX16AQ47G). Y. Mei was supported by the NASA High Mountain Asia Program (Award Number NNX16AQ89G). The authors are also grateful to Jing Tao for sharing the downscaling codes. The computations were conducted on ARGO, a research computing cluster provided by the Office of Research Computing at George Mason University, VA. [http://orc.gmu.edu]. The downscaling framework is implemented using MatLab functions available on co-author Mei’s GitHub profile at https://github.com/YiwenMei/AtmDS.

APPENDIX

Nomenclature

This appendix contains Table A1, which is a list of symbols that are used more than three times in the paper.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-m air temperature</td>
<td>$T$</td>
<td>K</td>
</tr>
<tr>
<td>2-m dewpoint temperature</td>
<td>$T_d$</td>
<td>K</td>
</tr>
<tr>
<td>Dewpoint temperature lapse rate</td>
<td>$\Gamma_d$</td>
<td>K km$^{-1}$</td>
</tr>
<tr>
<td>Longwave radiation</td>
<td>$L$</td>
<td>W m$^{-2}$</td>
</tr>
<tr>
<td>Normalized difference vegetation index</td>
<td>NDVI</td>
<td></td>
</tr>
<tr>
<td>Pressure</td>
<td>$P$</td>
<td>Pa</td>
</tr>
<tr>
<td>Shortwave radiation</td>
<td>$S$</td>
<td>W m$^{-2}$</td>
</tr>
<tr>
<td>Specific humidity</td>
<td>$q$</td>
<td>kg kg$^{-1}$</td>
</tr>
<tr>
<td>SRTM terrain elevation</td>
<td>$Z$</td>
<td>m MSL</td>
</tr>
<tr>
<td>Surface roughness</td>
<td>$z_0$</td>
<td>m</td>
</tr>
<tr>
<td>Temperature lapse rate</td>
<td>$\Gamma$</td>
<td>K km$^{-1}$</td>
</tr>
<tr>
<td>Vapor pressure</td>
<td>$E$</td>
<td>Pa</td>
</tr>
<tr>
<td>Wind speed</td>
<td>$W$</td>
<td>m s$^{-1}$</td>
</tr>
<tr>
<td>Zero-plane displacement height</td>
<td>$h_0$</td>
<td>m</td>
</tr>
</tbody>
</table>

TABLE A1. List of symbols.
REFERENCES


