Retrieval of Hourly Records of Surface Hydrometeorological Variables Using Satellite Remote Sensing Data

SANAZ MOGHIM
School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia

ANDREW JAY BOWEN AND SEPIDEH SARACHI
Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California

JINGFENG WANG
School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia

(Manuscript received 31 July 2013, in final form 8 August 2014)

ABSTRACT

A new algorithm is formulated for retrieving hourly time series of surface hydrometeorological variables including net radiation, sensible heat flux, and near-surface air temperature aided by hourly visible images from the Geostationary Operational Environmental Satellite (GOES) and in situ observations of mean daily air temperature. The algorithm is based on two unconventional, recently developed methods: the maximum entropy production model of surface heat fluxes and the half-order derivative–integral model that has been tested previously. The close agreement between the retrieved hourly variables using remotely sensed input and the corresponding field observations indicates that this algorithm is an effective tool in remote sensing of the earth system.

1. Introduction

Impacts of climate change on the natural environment and human society are evident over the past century (Boisvenue and Running 2006; Oppenheimer 2013). Adaptation strategies and impact assessments require high-resolution hydrometeorological data to capture the variability of hydrometeorological processes at local scales (Stamm et al. 1994; Wood et al. 2004; Georgakakos et al. 1998, 2012; van Rheenen et al. 2004; Tanaka et al. 2006; Maurer et al. 2007). The available data products do not always have desired space and time resolutions and hence need to be downscaled to meet the demands of applications. Two approaches to obtain fine-resolution data products from the corresponding coarse-resolution ones are dynamical and statistical downscaling (Fowler et al. 2007; Wilby and Wigley 1997; Xu 1999). Dynamical downscaling often uses regional climate models (RCMs) to simulate high-resolution variables. The performance of RCMs depends on the validity of the boundary condition prepared from the general circulation models (GCMs; Miller et al. 1999; Xue et al. 2007), which usually have biases (Liang et al. 2008). The computational cost of dynamical downscaling makes it “essentially impossible” to produce long-term records (Maurer and Hidalgo 2008). As a result, statistical downscaling methods are often used to generate high-resolution variables. The statistical downscaling is based on the derived relationship or transfer function between modeled outputs and the corresponding observations during the same period of time, assuming that the transfer function derived from historical data remains unchanged for future time (Mearns et al. 1999; Murphy 1999). Yet, this stationarity assumption for the transfer function may not always hold. Mathematically, a transfer function relates predictands to predictors. For example, predictors can be large-scale atmospheric variables and predictands can be the corresponding local surface variables, the parameters of the distributions of local variables, and the frequencies of the extreme local variables (Pfizenmayer and von

Corresponding author address: Sanaz Moghim, Georgia Institute of Technology, 311 Ferst Dr., Atlanta, GA 30332.
E-mail: smoghim@gatech.edu

DOI: 10.1175/JHM-D-13-0127.1

© 2015 American Meteorological Society
The constructed analog (CA) is one of the approaches that use relationships between modeled and observed variables (Hidalgo et al. 2008; Maurer et al. 2010). Maurer et al. (2010) combined the bias correction with the CA (BCCA) to produce bias-corrected and downscaled data. A new bias correction and spatial downsampling (BCSD) method was proposed by Wood et al. (2004). This method uses a quantile mapping approach to remove the biases of the variables (temperature and precipitation) and uses the inverse distance weighting technique for spatial interpolation. Zhang and Georgakakos (2012) proposed the Joint Variable Spatial Downscaling (JVSD) method for bias correction and downsampling of the GCM outputs (temperature and precipitation). The downsampling step uses historical analog for the joint cumulative distribution functions (CDFs) of the increments of the variables obtained by the 12-month differencing process. A relationship between the CDFs of the increments of the modeled and observed variables is defined to correct the biases.

The Intergovernmental Panel on Climate Change (IPCC) Third Assessment Report (AR3) used weather classification, regression models, and weather generators as the main statistical downscaling methods (Giorgi et al. 2001). Weather classification methods group a certain number of weather states or types using cluster analysis or circulation classification schemes (Corté-Real et al. 1999; Huth 2002; Kidson 2000; Bárdossy and Caspary 1990; Jones et al. 1993). The weather states are identified based on their similarity in the nearest neighbors (Hay 1991; Corté-Real et al. 1999). The dominant weather state is assigned as the predictand. Even though the weather classification method can represent some physical processes of climate system, it may not capture intra-annual variability of surface variables (Wilby et al. 2003; Martin et al. 1996). The performance of the weather classification method is similar to the complex regression methods (Zorita and von Storch 1999). Regression models are based on linear or nonlinear relationships between local variables as predictands and large-scale atmospheric forcings as predictors. The widely used regression models are multiple regression (MR; Murphy 1999), canonical correlation analysis (CCA; von Storch et al. 1993), and artificial neural network (ANN). Multiple regression simulates a single dependent variable from multiple independent variables. A more complex version of the MR is the CCA based on the relationships between a set of dependent and independent variables (Lattin et al. 2003). Nonlinear relationships between the predictors and the predictands are used in the ANN model, a computationally efficient black-box model (Crane and Hewitson 1998). Although regression methods are relatively simple, they tend to underestimate the variances of the variables (von Storch 1999; Burger 1996).

Another statistical downscaling technique used in AR3 is weather generator (WG), where daily variables are disaggregated by conditioning the local climate variables on the large-scale atmospheric predictors, such as rainfall, to resolve diurnal cycles (Kilsby et al. 1998; Fowler et al. 2000). For instance, Markov processes are used to describe the precipitation occurrence (rainy or no-rain day) in the WGs, and then the secondary variable, such as temperature, is downscaled by conditioning on the precipitation (Katz 1996; Semenov and Barrow 1997; Wilks and Wilby 1999). Stochastic downscaling and weather-typing methods are widely used in the IPCC Fourth Assessment Report (AR4; Jones et al. 2009).

Statistical downscaling methods are computationally advantageous over dynamical downscaling methods (Hay and Clark 2003; Wilby and Wigley 2000; Wood et al. 2004) although they do have drawbacks. First, statistical relationships between predictors and predictands may be difficult to identify. Second, transfer functions do not always capture the underlying physical mechanisms and variability of climate systems. This study proposes a new algorithm to disaggregate daily surface hydrometeorological variables into hourly variables using physically and statistically based models with input from satellite remote sensing data complemented by ground observations. In section 2, study sites and datasets are described followed by the algorithm formulation in section 3. Section 4 presents the tests of the algorithm, followed by conclusions in section 5.

## 2. Study sites and datasets

Two sites in Brazil [Caxiuanã (CAX) and Reserva Pê-de-Gigante (PDG)] during January and February 2002 and two sites in Arizona [Kendall (Ken) and Lucky Hills (LKH)] during January 2002 were used to test the proposed algorithm. Table 1 lists the characteristics of the sites.

Observations of hourly air temperature for the two sites in Brazil are from the Large-Scale Biosphere–Atmosphere Experiment in Amazonia (LBA) project (de Goncalves et al. 2013). The meteorological data, including air temperature, were measured by sensors installed at 2-m height on a 10-m tower at the sites (more detailed information about the data is available at http://daac.ornl.gov/cgi-bin/dsvviewer.pl?ds_id=1177). Observations of net radiation, sensible heat flux, air temperature, and other meteorological and hydrological data at the two sites in Arizona are from the Walnut Gulch Experimental Watershed (Emmerich and Verdugo 2008). Net radiation and air temperature were measured by
a net radiometer at 3-m height and a temperature probe at 2-m height, respectively. Bowen ratio energy balance technique was used to measure sensible heat flux. Data products are publicly available online (www.tucson.ars.ag.gov/dap/). Daily mean air temperature required as an input to the proposed algorithm is obtained by averaging the observed hourly air temperature at the sites. Satellite-based remote sensing data as the other input used in this study come from the Geostationary Operational Environmental Satellite (GOES) of the National Oceanic and Atmospheric Administration (NOAA; publicly available at www.class.ngdc.noaa.gov). GOES (from GOES-7 to GOES-15) provides visible and infrared images with various spatial and temporal coverages measured by a 19-channel (one visible and 18 infrared) sounder and a five-channel (one visible and four infrared) imager. Visible images of GOES-8 centered at 0.65 μm are used for the retrieval of hourly records of meteorological variables in this study. Spatial and temporal resolutions of visible images are 1 km and 1 h, respectively. Because the proposed algorithm estimates air temperature at hourly resolution and for the sites, the high temporal and spatial resolutions of the satellite data are the requirements for this study.

### 3. Algorithm formulation

The retrieval algorithm of hourly meteorological variables from hourly satellite data and in situ daily air temperature is based on three models: 1) surface net radiation $R_n$ is estimated using hourly albedo $\alpha$ derived from channel one (visible) of GOES (Bisht and Bras 2010), 2) partition of net radiation into fluxes (sensible, latent, and ground heat fluxes) is estimated using the maximum entropy production (MEP) model (Wang and Bras 2011), and 3) hourly surface air temperature is retrieved from sensible heat flux obtained from the second model using the half-order integral model (Wang and Bras 1999).

#### a. Net radiation model

Surface net radiation is expressed as

$$R_n = R^1_S - R^1_S + R^1_L - R^1_L,$$  \hspace{1cm} (1)

where $R^1_S$, $R^1_S$, $R^1_L$, and $R^1_L$ are downwelling shortwave, reflected shortwave, downwelling longwave, and upwelling longwave radiation, respectively. Components of the surface energy budget can be parameterized using near-surface air temperature, humidity, and surface temperature (Brutsaert 1975; Diak and Gautier 1983; Idso 1981; Prata 1996; Zillman 1972; Bisht and Bras 2010). Downwelling shortwave radiation for clear sky $R^1_S^{\text{clear}}$ is expressed as (Zillman 1972)

$$R^1_S^{\text{clear}} = \frac{S_0 \cos^2 \theta}{1.85 \cos \theta + e_o (2.7 + \cos \theta) \times 10^{-3} + \beta},$$  \hspace{1cm} (2)

where $S_0$ is the solar constant (1367 W m$^{-2}$), $\theta$ is the solar zenith angle, $e_o$ (mb) is the near-surface vapor pressure, and $\beta$ is an empirical coefficient set to be 0.1. It has been shown that Eq. (2) tends to overestimate $R^1_S^{\text{clear}}$ (Niemelä et al. 2001a,b; Bisht et al. 2005). In this study, $\beta$ is set to be 0.2 to correct the overestimation (Bisht and Bras 2010), and $e_0$ is approximated as the saturated vapor pressure at air temperature $T_a$. The second term in the denominator of Eq. (2) plays a minor role on $R^1_S^{\text{clear}}$, as it is at least one order of magnitude smaller than the first and third term, allowing a convenient approximation when surface humidity data are not available.

### Table 1. General characteristics of the sites. Caxiuanã and Reserva Pé-de-Gigante are sites of the LBA project; Kendall and Lucky Hills are sites of the Walnut Gulch Experimental Watershed.

<table>
<thead>
<tr>
<th>Site</th>
<th>Location</th>
<th>Lat</th>
<th>Lon</th>
<th>Elev (m)</th>
<th>Biome type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caxiuanã</td>
<td>Para, Brazil</td>
<td>1.72°S</td>
<td>51.46°W</td>
<td>23</td>
<td>Forest</td>
</tr>
<tr>
<td>Reserva Pé-de-Gigante</td>
<td>São Paulo, Brazil</td>
<td>21.62°S</td>
<td>47.65°W</td>
<td>690</td>
<td>Savanna</td>
</tr>
<tr>
<td>Kendall</td>
<td>Arizona, United States</td>
<td>31.73°N</td>
<td>109.94°W</td>
<td>1526</td>
<td>Grass</td>
</tr>
<tr>
<td>Lucky Hills</td>
<td>Arizona, United States</td>
<td>31.74°N</td>
<td>110.052°W</td>
<td>1372</td>
<td>Shrub</td>
</tr>
</tbody>
</table>

Surface downwelling and reflected shortwave radiation are expressed in terms of cloud and surface albedo, $\alpha_c$ and $\alpha_s$, respectively, as

$$R^1_S = (1 - \alpha_c)R^1_S^{\text{clear}}$$  \hspace{1cm} (3)

and

$$R^1_S = \alpha_s R^1_s.$$  \hspace{1cm} (4)

Albedo data used in this study are derived from the GOES visible images. The variable $\alpha_c$ is obtained under cloudy conditions, and $\alpha_s$ is obtained under clear sky conditions. Since the temporal variability of $\alpha_s$ is relatively low (Tsveltsinskaya et al. 2006), the obtained $\alpha_s$ is assigned as the time-invariant parameter for each site.
The other components of net radiation, downwelling and upwelling longwave radiation, are calculated based on the Stefan–Boltzmann law as

\[ R^2_L = \varepsilon_a \sigma T_a^4 \]  

and

\[ R^1_L = \varepsilon_s \sigma T_s^4, \]  

where \( \sigma \) is the Stefan–Boltzmann constant (5.67 \times 10^{-8} \text{W m}^{-2} \text{K}^{-4} ); \( T_a \) is assumed to be equal to surface temperature \( T_s \) as an approximation when hourly \( T_a \) is not available. Surface emissivity \( \varepsilon_a \) is taken as unity because of its small variability over the land (Dickinson et al. 1986). Air emissivity \( \varepsilon_a \) is parameterized using \( e_0 \) (mb) and \( T_a \) (K) as (Prata 1996)

\[ \varepsilon_a = 1 - (1 + \xi) \exp(-\sqrt{1.2 + 3\xi}), \]  

where the dimensionless parameter \( \xi \) is calculated as

\[ \xi = 46.5 \frac{e_0}{T_a}. \]  

\[ \text{b. Sensible heat flux model} \]

The recently developed MEP model of evapotranspiration (ET; Wang and Bras 2011) provides a parameterization of sensible heat flux used in this study. The theory and formulation of the MEP model are described in detail in Wang and Bras (2009, 2011). The MEP model predicts the partition of net radiation into sensible \( H \), latent \( E \), and ground \( G \) heat fluxes according to

\[ R_n = H + E + G, \]  

\[ G = \frac{B(\sigma)}{\sigma} I \int_0^1 H |H|^{-1/6}, \]  

and

\[ E = B(\sigma)H, \]  

where \( I \) is the thermal inertia of the soil, \( I_0 \) is the apparent thermal inertia of the air, and the inverse Bowen ratio \( B \) is given as

\[ B(\sigma) = 6 \left( \sqrt{1 + \frac{11}{36} \sigma - 1} \right) \]  

and

\[ \sigma = \frac{L_s^2}{c_p R \gamma} q_s, \]  

where \( c_p \) is the specific heat of air at constant pressure, \( L_s \) is the latent heat of vaporization of liquid water, \( R \) is the gas constant for water vapor, and \( q_s \) is the surface specific humidity at surface (skin) temperature \( T_s \). For a saturated soil, a dimensionless parameter \( \sigma \) may be expressed as

\[ \sigma = \frac{\Delta}{\gamma}, \]  

where \( \Delta \) is the slope of the saturation vapor pressure at \( T_s \) according to the Clausius–Clapeyron equation and \( \gamma \) is the psychrometric constant (Brunt 1939; Brutsaert 1982).

\[ \text{c. Air temperature model} \]

The diurnal variation of near-surface air temperature may be expressed in terms of a weighted time average (i.e., half-order integral) of sensible heat flux analogous to the half-order integral model of soil temperature and ground heat flux (Wang and Bras 1999; Bennett et al. 2008):

\[ T_a(t) = T_0 + \frac{1}{I} \int_{-\infty}^t \frac{H(s) ds}{\sqrt{\pi(t-s)}}, \]  

where \( T_0 \) is a reference temperature. In this study, \( T_0 \) is determined in such a way that the modeled daily mean air temperature \( T_a \), according to Eq. (15), is equal to the observed daily mean air temperature. The variable \( I \) is the thermal inertia of air. In this study, \( I \) is treated as a fitting parameter to the model described in section 4. The tests below suggest that the integral on the right-hand side of Eq. (15) captures diurnal variation of air temperature reasonably well when \( I \) is treated as a fitting parameter [see Eq. (16)].

\[ \text{4. Tests} \]

\[ \text{a. Procedure} \]

The hydrometeorological variables including hourly net radiation, sensible heat flux, and air temperature are estimated through three steps. In step one, mean daily air temperature and cloud albedo are used to calculate the components of radiation (long- and shortwave radiation), sensible heat flux is computed for a given net radiation using the MEP model, and hourly air temperature is estimated using the half-order integral model. In step two, the same process is repeated by using the estimated hourly temperature as the input to update the calculated radiation and heat flux from step one. In steps one and two, \( I \) is set at 3000 thermal inertia units (tiu; 1 tiu = 1 J m^{-2} K^{-1} s^{-1/2}) as an initial guess. In step
three, the hourly temperature and sensible heat flux obtained from step two are used to estimate $I$ as the regression coefficient between diurnal amplitudes of air temperature $\Delta T_a$ and sensible heat flux $\Delta H$ according to the following equation (Wang et al. 2010):

$$I = \frac{\Delta H}{\Delta T_a \sqrt{\nu_0}},$$

(16)

where $\nu_0 = 2\pi/d$, with $d$ being the length of day (24 h). In step three, the estimated $I$ is used to recalculate the hourly temperature. The schematic diagram of the procedure is illustrated in Fig. 1.

b. Results

The proposed algorithm was tested at two sites in Brazil, Cax and PDG, during January and February 2002 and two sites in Arizona, Ken and LKH, during January 2002. Albedo, as an input to the algorithm, is assigned to each site based on the closest distance of the GOES images to the coordinates (latitude and longitude) of the site. Estimated hourly net radiation and sensible heat flux at two sites in Brazil are illustrated in Fig. 2. The lower negative net longwave radiation at Cax agrees with the fact that more (thicker) clouds over Cax block more surface-emitted longwave radiation from escaping to the space. It suggests that the average cloud albedo over Cax is greater than that over PDG during the test period. Note that narrowband albedo data derived from one visible channel of GOES centered at 0.65 m (0.55–0.75 m) are essentially equivalent to broadband albedo data since the channel covers the range of the high intensity of shortwave radiation energy. Average sensible heat flux over PDG is higher than that over Cax. This is consistent with the fact that lower humidity over the savanna site (PDG) results from higher sensible heat flux compared to those over the forest site (Cax). Figure 3 compares estimated hourly air temperature with the corresponding observations at the two sites. The field tests suggest that the proposed algorithm is able to retrieve hourly records of surface hydrometeorological variables using satellite remote sensing data supplemented by in situ daily mean temperature. The performance of the algorithm is further illustrated through scatterplots in Fig. 4 for the two sites. Correlation coefficients $\rho$ between observed and estimated air temperature of Cax and PDG are 0.90 and 0.81, respectively. The algorithm tends to underestimate lower temperatures and overestimate higher temperatures at Cax, but there are no apparent biases in the
estimated temperature at PDG. This may be due to the fact that the stability of the atmosphere over the forest (Cax) is lower than that over the savanna (PDG). It implies that improvement of parameterization of the thermal inertia is needed for the unstable atmosphere over the forest sites. Since observations of $R_n$ and $H$ at Cax and PDG sites were not available to the authors, the radiation and flux models were tested using field observations from the Walnut Gulch Experimental Watershed. Figure 5 compares estimated hourly net radiation with the corresponding observations at Ken and LKH sites for January 2002. Figure 5 indicates that the estimated hourly $R_n$ is in close agreement with the observed hourly $R_n$. The slightly overestimated $R_n$, in particular, at LKH site may be caused by the use of air temperature as a surrogate of surface temperature, leading to a decrease in $R_L$ and an increase in $R_n$ [see Eqs. (1), (6)]. Figure 6 shows a close agreement between the modeled and observed $H$. Note that unrealistic values that appeared by spurious spikes in Fig. 6 indicate the necessity of the other field experiment to measure
heat fluxes (such as the eddy covariance method). The corresponding scatterplots of $R_n$ and $H$ in Figs. 7 and 8 indicate close agreement between the estimated and observed fluxes. Note that any uncertainty of the results may be caused by the uncertainty of the remote sensing data and the way that albedo is assigned to each site, which is based on the closest distance of the GOES images to the coordinate of the site. Figure 9 indicates that the estimated air temperature agrees closely with the observations. Correlation coefficients between observed and estimated air temperature of Ken and LKH are 0.80 and 0.74, respectively, shown in Fig. 10. The differences between estimated and observed $R_n$ and $H$ obtained from the net radiation and the sensible heat flux models may cause the discrepancies between estimated and observed $T_a$ because the air temperature model requires sensible heat flux obtained from the sensible heat flux model and net radiation from the net radiation model. The disaggregated hourly air temperature is affected by the thermal inertia parameter, daily mean air temperature, and GOES images as inputs to the algorithm. For any given sensible heat flux, greater thermal inertia leads to smaller diurnal amplitude of air temperature and vice versa [see Eq. (15)]. Daily mean

---

**Fig. 7.** Estimated vs observed net radiation corresponding to Fig. 5 for (a) Ken and (b) LKH.

**Fig. 8.** Estimated vs observed sensible heat flux corresponding to Fig. 6 for (a) Ken and (b) LKH.
temperature affects diurnal variations of air temperature through longwave radiation. When daily mean temperature decreases, upward longwave radiation decreases, leading to an increase in net radiation and sensible heat flux and a greater amplitude of the diurnal variation of air temperature, and vice versa. The impact of uncertainty of the GOES data on air temperature was studied at the two sites (Ken and LKH). Relative uncertainty in the GOES albedo data is about 15% (Liang 2008), which leads to the error of 0.06–0.10 in cloud albedo ranging from 0.4 to 0.7. The error is added to the albedo data and the algorithm was used to recalculate net radiation, sensible heat flux, and air temperature. The uncertainty in the albedo data causes the uncertainty in the estimated air temperature through the three models.

1) Net radiation model: Change in albedo impacts down- and upwelling shortwave radiation. For $R_{\text{clear}}^S$ ranging from 900 to 1100 W m$^{-2}$, the error of the GOES data leads to an average error of 30–80 W m$^{-2}$ in net shortwave radiation and net radiation [see Eqs. (3), (4)].

2) Sensible heat flux model: 30–80 W m$^{-2}$ error in net radiation causes an average error of 15–40 W m$^{-2}$ in sensible heat flux through MEP model.

3) Air temperature model: 15–40 W m$^{-2}$ error in sensible heat flux leads to an average error of 1–2 K in air temperature through the half-order integral model.

The close agreement between the observed and estimated air temperature over the sites with different climate (see Table 1) suggests that the MEP and the half-order time integral models used in the proposed algorithm have the potential to improve the estimates of the variability and magnitude of air temperature globally. Although the physically based models (e.g., transfer based models) often use gradients of the variables such as temperature and humidity, the MEP model uses only one level of the inputs, and the vertical profiles of the inputs are avoided. A major advantage of the algorithm is that the algorithm requires only mean daily temperature and satellite-based single-level observations as inputs. The algorithm has lower computational

![Fig. 9. Estimated hourly air temperature (blue dashed line) compared with observations (red solid line) from the Walnut Gulch Experimental Watershed at (a) Ken and (b) LKH in January 2002.](image1)

![Fig. 10. Estimated vs observed air temperature corresponding to Fig. 9 for (a) Ken ($\rho = 0.80$) and (b) LKH ($\rho = 0.74$).](image2)
cost compared to other existing dynamical downscaling methods.

5. Conclusions

This study develops a new physically and statistically based algorithm for retrieving hourly surface hydro-meteorological variables including air temperature, radiation, and heat flux. Tests of the algorithm at multiple sites located in different geographical regions show that hourly estimates of these variables are in close agreement with observations. The proposed algorithm distinguishes itself from the existing approaches by effective use of available information (mean daily temperature, satellite-based images, and the thermal inertia) from only one level. Decision makers and water resources planners can use the retrieved hourly data to drive other land, ecosystem, and coupled climate–hydrology–water resources models.

Acknowledgments. This work was sponsored by the Andes–Amazon Initiative of the Gordon and Betty Moore Foundation. Funding was also provided by NSF Grant EAR-1138611 and ARO Grant W911NF-07-1-0126. We are deeply grateful to Prof. Rafael L. Bras’s support and supervision throughout the project. We thank Dr. Naomi M. Levine for preparing some of the datasets used in this study.

REFERENCES


Hidalgo, H. G., M. D. Dettinger, and D. R. Cayan, 2008: Downscaling with constructed analogues: Daily precipitation


