Environmental Research 159 (2017) 297–312

Contents lists available at ScienceDirect

Environmental Research

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

Estimating daily minimum, maximum, and mean near surface air temperature using hybrid satellite models across Israel

Adar Rosenfeld[a], Michael Dorman[a], Joel Schwartz[b], Victor Novack[c], Allan C. Just[d], Itai Kloog[a,*]

[a] Department of Geography and Environmental Development, Ben-Gurion University of the Negev, P.O.B. 653, Beer Sheva, Israel
[b] Department of Environmental Health, Harvard T. H. Chan School of Public Health, Cambridge, MA, USA
[c] Clinical Research Center, Soroka University Medical Center, Beer Sheva, Israel
[d] Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA

ARTICLE INFO

Keywords:
Air temperature
Surface temperature
MODIS
Epidemiology
Exposure error

ABSTRACT

Meteorological stations measure air temperature (Ta) accurately with high temporal resolution, but usually suffer from limited spatial resolution due to their sparse distribution across rural, undeveloped or less populated areas. Remote sensing satellite-based measurements provide daily surface temperature (Ts) data in high spatial and temporal resolution and can improve the estimation of daily Ta. In this study we developed spatiotemporally resolved models which allow us to predict three daily parameters: Ta Max (day time), 24 h mean, and Ta Min (night time) on a fine 1 km grid across the state of Israel. We used and compared both the Aqua and Terra MODIS satellites. We used linear mixed effect models, IDW (inverse distance weighted) interpolations and thin plate splines (using a smooth nonparametric function of longitude and latitude) to first calibrate between Ts and Ta in those locations where we have available data for both and used that calibration to fill in neighboring cells without surface monitors or missing Ts. Out-of-sample ten-fold cross validation (CV) was used to quantify the accuracy of our predictions. Our model performance was excellent for both days with and without available Ts observations for both Aqua and Terra (CV Aqua R^2 results for min 0.966, mean 0.986, and max 0.967; CV Terra R^2 results for min 0.965, mean 0.987, and max 0.968). Our research shows that daily min, mean and max Ta can be reliably predicted using daily MODIS Ts data even across Israel, with high accuracy even for days without Ta or Ts data. These predictions can be used as three separate Ta exposures in epidemiology studies for better diurnal exposure assessment.

1. Introduction

Ambient temperature or near-surface air temperature (Ta) is considered to be a key factor in a wide range of applications and numerous scientific fields such as environmental epidemiology (Zanobetti and Schwartz, 2008; Basu and Samet, 2002; McGeehin and Mirabelli, 2001; Garske et al., 2013; Focks et al., 1995), climatology (Griggs and Noguer, 2002; Watson and Albritton, 2001), and hydrology (Wang et al., 2009). Many studies in the fields of epidemiology and public health have shown that elevated Ta is a significant factor affecting numerous health outcomes and can cause a variety of medical conditions and diseases. Exposure to elevated Ta can result in morbidity, mortality, respiratory diseases, and cardiovascular diseases. Usually, susceptible groups such as the elderly, infants, sick and populations from low socioeconomic status are more affected compared to healthy adults (Buscail et al., 2012; Kovats et al., 2004; Kestens et al., 2011; Curriero et al., 2002). Exposure to high diurnal temperature variations (DTV, the difference between Max temperature and Min temperature in a specific day) can also affect public health Kan et al. (2007) used daily weather and mortality data from Shanghai during 2001–2004 and found a strong association between DTV and daily mortality. Zeng et al. (2017) found that maternal exposure to high DTV, particularly during the second trimester, increased the risk of pneumonia in offspring.

Ta is usually measured with sensors inside meteorological station shelters at the height of 2 m above the ground surface. These stations collect data with high accuracy and high temporal resolution but rely on regional infrastructure and often their distribution across the rural, undeveloped or less populated areas is sparse. The distribution may also not adequately capture urban heat islands (UHI, a metropolitan area that is significantly warmer than its surrounding rural areas due to human activities). Thus the climatic information meteorological station networks collect is not representative of the true spatial variations of air temperature in a given region (Vancutsem et al., 2010; Prihodko and Goward, 1997). This result in biased Ta exposure estimates and

---

[*] Corresponding author.
E-mail address: ekloog@hsph.harvard.edu (I. Kloog).

http://dx.doi.org/10.1016/j.envres.2017.08.017
Received 9 March 2017; Received in revised form 7 August 2017; Accepted 8 August 2017

exposure error. Epidemiological studies on health effects of temperature are conducted using large geographical areas as units of the analysis (state, nationwide), thus potentially biasing the health effects risks estimates due to exposure measurement error (i.e., assigning inaccurate exposure to each study participant (Zeger et al., 2000; Kloog et al., 2014)). Previous research tried to resolve this lack of high spatio-temporal resolution Ta data, yet most studies used mainly geo-statistical approaches including interpolation of Ta readings between known sites. For example, (Anderson, 2002) evaluated 3 techniques including Spline, Inverse Distance Weighting (IDW) and Kriging and checked their ability to create continuous dataset of Ta across phoenix AZ for a sample day (May 24, 1998). He found that Kriging had the best performance with a root mean square error (RMSE) of 1.62 °C. Yet none of the interpolation methods can adequately estimate actual temperature variability across 110 km^2 from only 11 sampling points and 25 points used to create the interpolated surfaces. Prihodko and Goward (1997) developed the empirical temperature vegetation index (TVX algorithm) which is based on an estimated linear relationship between observed Land Surface Temperature (LST) and a Spectral Vegetation Index (NDVI). A strong correlation ($R^2 = 0.93$) was found between the satellite estimates and measured air temperatures with a mean error of 2.92 °C.

In more recent studies, researchers tried to fill the lack of continuous spatial Ta data with satellite data which offers globally continuous surface temperature (Ts) data at high spatial resolution. Surface temperature as derived from satellites is defined as the "skin temperature" of the ground. It represents the actual reading that returns to the satellite sensor from the ground. One of the first works that tried to combine satellite (special sensor microwave / imager, SSM/I) data and Ta data (Basist et al., 1998) reported that correlation between Ts as measured from the satellite and Ta *in-situ* monitors was greater than 0.95 with standard error of 2 °C. Florio et al. (2004) combined advanced very high resolution radiometer (AVHRR) data with Ta ground monitors measurements and tried several statistical methods (Kriging, linear regressions) finding that the Kriging model performed better than the linear regression with an RMSE of 0.9 °C compared to 1.4 °C respectively. They concluded that models with satellite data performed slightly better under optimal clear-day conditions (cloud free) compared to models without satellite data, Furthermore the many studies (Balling and Brazel, 1988; Dousset, 1991; Lougeay et al., 1996; Su et al., 2008) have shown that the incorporation of land use regression (LUR) and land use variables improved the models due to the strong relation between surface temperature patterns, land use and day-to-day variability of spatial patterns. LUR is an important approach as it predicts long-term average pollution concentrations or temperature exposure at a given site based on surrounding land use and traffic characteristics. Vancutsem et al. (2010) found good accuracy in retrieving Min Ta using both Aqua and Terra night Ts products, with Aqua performing better then Terra as Aqua passes later at night. Aqua is closer to min Ta time with mean absolute error (a quantity used to measure how close forecasts are to the eventual outcomes) MAE = 1.73 °C, standard deviation = 2.4 °C, but failed to accurately retrieve Max Ta from day time Ts data due to major differences caused by local conditions, seasons, microscale advection effects, sky view factor, satellite – sun and surface geometry and thermal properties of the surface. Benali et al. (2012) used a statistical approach to accurately estimate Tmax, Tmin and Tavg for a 10 year period based on remote sensing—Land Surface Temperature (LST) data obtained from MODIS and auxiliary data. The statistical models estimated Tavg with a MEF (Model Efficiency Index) of 0.941 and a RMSE of 1.33 °C. Regarding Tmax and Tmin, the best MEF achieved was 0.919 and 0.871, respectively, with a 1.83 and 1.74 °C RMSE. White-Newsome et al. (2013) assessed the correlations between LST and percent surface imperviousness (SI) images with actual temperature readings from the ground in order to characterize the urban heat island effect, Statistically significant correlations ranging from 0.49 to 0.91 (spearman rank correlation coefficients, r_s) were observed between LST and SI. They presented how Both LST and SI can be used to better understand spatial variation in heat exposure over long periods of time, but less useful for estimating shorter-term actual temperature exposures. Zhang et al. (2011) presented models using MODIS night and day time LST and showed better correlation between models using night time Ts to observed Ta ($R^2 > 0.81$) than models using day time Ts ($R^2 > 0.57$) with significant seasonal variations. Fu et al. (2011a) used MODIS LST and linear regression to estimate air temperature of an alpine meadow on Northern Tibetan Plateau at heights of 1.5–2.1 m. The result show that daily Max and daytime mean air temperatures estimations where not accurate during growing season ($P > 0.01$, $R^2 < 0.10$), during the non-growing season linear relationships between daytime mean and daily max air temperatures to MODIS LST were significant ($P < 0.01$, $R^2 > 0.40$). The data was accurate enough to linearly estimate daily Min and nighttime mean air temperatures ($P < 0.01$, $R^2 > 0.55$). Nieto et al. (2011) improved the TVX (temperature vegetation index) and proposes a novel methodology to estimate NDVI max that uses observed air temperature to calibrate the NDVI max for each vegetation type, Nieto used the new NDVI max and the previous NDVI max with MSG-SEVIRI images in Spain during the year 2005. Results show that the method performed well with a Mean Absolute Error ranging between 2.8 °C and 4 °C. In an attempt to improve day time estimations from MODIS data, (Zhu et al., 2013) also improved the TVX method which lowered the threshold for the negative correlation coefficient between NDVI (a numerical indicator that assess whether the target being observed contains live green vegetation or not) and Ts to 0.1 and managed to predict Max Ta, successfully improving the accuracy of the daily Max air temperature estimations from RMSE = 7.45 °C to RMSE = 3.79 °C and MAE = 6.21 °C to MAE = 3.03 °C with $R^2 = 0.83$. Emamifar et al. (2013) used a M5 model (a model based on a binary decision tree having linear regression functions at the terminal (leaf) nodes, which develops a relationship between independent and dependent variables) to estimate *Ta* in Khuzestan province (in the southwest of Iran). The input variables for the M5 model tree were the daytime and nighttime MODIS *Ts*, extraterrestrial solar radiation and Julian day. The results show good model performance with a RMSE = 2.3 °C and a $R^2 = 0.96$. Xu et al. (2014) used two statistical methods (linear regression and random forest) to estimate Max Ta in British-Columbia, the two model performances were validated with station-by-station cross-validation. Results indicated that the random forest model achieved better accuracy (mean absolute error, MAE = 2.02 °C, $R^2 = 0.74$) than the linear regression model (MAE = 2.41 °C, $R^2 = 0.64$). Ho et al. (2014) introduce a spatial regression approach and calibrated Three regression models, ordinary least squares regression, support vector machine, and random forest using Landsat data and field observations in Vancouver Canada as a case study. Results based on cross validation show that the random forest produced the lowest RMSE = 2.31 °C. A different approach showed that the spatiotemporal regression kriging method incorporating time series of land surface temperatures images from MODIS produced significantly more accurate results compared to spatial interpolation techniques alone, with an accuracy of ± 2.4 °C and 91% of the variability in daily temperatures explained with this method compare to 41% for ordinary kriging (Hengl et al., 2012). The same method of spatiotemporal regression kriging was performed by Kilibarda et al. (2014) using time series of land surface temperatures images from MODIS with topographic layers and geometric temperature trends as covariant to predict temperature for the global land mass. The model presented fits of $R^2 = 0.966$, $R^2 = 0.959$, $R^2 = 0.964$ and standard error of 2.39 °C, 2.70 °C, 2.60 °C for mean min and max temperature estimations respectively. Xu and Liu (2015) estimated near-surface air temperature in Beijing using Landsat/TM satellite images and a statistical model (using parameters: LST, NDVI, altitude & surface albedo) considering that air temperature is related to the present pixel and also affected by the surrounding environment, the estimation accuracy was high ($R^2 = 0.66$ and MAE of 0.87 °C). Chen