Urban Climate 17 (2016) 146–160

Contents lists available at ScienceDirect

Urban Climate

journal homepage: http://www.elsevier.com/locate/uclim

A novel method of urban land type identification utilizing meteorological station network temperature data, as developed for the Baltimore-Washington region

Cassie A. Stearns [a,*], Ricardo K. Sakai [a], Everette Joseph [b]

[a] *Howard University Program in Atmospheric Sciences, 2355 6th Street NW, Washington, DC 20059, United States*
[b] *Atmospheric Sciences Research Center, State University of New York, University at Albany, 251 Fuller Road, Albany, NY, 12203, United States*

ARTICLE INFO

Article history:
Received 23 March 2015
Received in revised form 6 May 2016
Accepted 16 May 2016
Available online xxxx

Keywords:
Urban heat island
Land use type
Meteorological station data
Identification method

ABSTRACT

Changes to urban areas over time and their relationship to precipitation and climate is a topic of growing interest. Cities have existed and expanded for centuries, but datasets which show levels and changes in urbanization prior to recent years are often spatially or temporally limited. It is known urban areas are generally warmer than rural areas, therefore relative values of temperatures can be used to discern land type. In this study, a new method utilizing the statistical k-means clustering technique is proposed to identify urbanization by using temperature data recorded at meteorological weather stations. The new technique was developed and applied to a network of meteorological stations in the greater Washington, DC region for 5-year intervals from 1980 to 2010, where it was able to identify urban, suburban, and rural land types for both large and small cities. The technique was also able to capture general temporal transitions from rural to suburban environments at the edges of urban areas.

1. Introduction

We live in a time of rapid urbanization. In 2009, more than half of the world's population lived in urban areas, and by 2030, that number is expected to be near two thirds (Population Reference Bureau, 2007). Links have been drawn between the effects of urbanization and changes to weather and climate during recent decades (Kalnay, 2008; Seto and Shepherd, 2009; Shepherd et al., 2010). Understanding these changes is

* Corresponding author.
E-mail addresses: c.a.stearns@gmail.com (C.A. Stearns), ricardo.k.sakai@howard.edu (R.K. Sakai), ejoseph@albany.edu (E. Joseph)

http://dx.doi.org/10.1016/j.uclim.2016.05.002
2212-0955

important to infrastructure and safety needs (Shepherd et al., 2010), as well effects on general health (Hondula et al., 2015; Li and Bou-Zeid, 2013).

In order to understand the effects of urbanization on climate and weather, two pieces of information are needed at minimum: data on changes to climate and weather and data on changes in land use. Older urban-weather studies relied mostly on general urban location (Huff and Changnon, 1972a,b), but more recent work has come to show relationships between city size and storm strength (Ashley et al., 2011; Schmid and Niyogi, 2013) as well as interactions between the city edge and the environment (Mote et al., 2007; Shem and Shepherd, 2009). Additionally, green spaces within urban areas have been identified as areas of study (Cheval and Dumitrescu, 2009; Schluenzen et al., 2009). Finally, modern models and land classification schemes include multiple urban/suburban levels depending on the degree of development.

Prior studies examining changes in urban area have used land information datasets such as National Land Cover Database (Hand and Shepherd, 2009), information from studies on land change for the region of interest (Dixon and Mote, 2003) or inferred information via differences between urban stations and known rural stations (Brazel et al., 2000; Schluenzen et al., 2009) or some even inferred from land records (Pielke et al., 2011). Recently, methods have been developed utilizing satellite data to identify urban extent. Cheval and Dumitrescu (2009) used statistical analysis of changes in MODIS-derived surface temperatures and Zhang et al. (2010) used a combination of MODIS data and night-light extent to define impervious surface area (ISA), which was then used to define urban area extent and shape.

While these techniques are promising looking forward in time, understanding can also be gained by looking backward. Unfortunately, urban land use data prior to the era of satellites in the 1970s is more difficult to obtain. Aerial images exist prior to this time period, but high-resolution aerial photos are both spatially and temporally sparse. They also require visual interpretation which takes considerable time and effort. Population has been cited as a way to measure the growth of a city (Marshall, 2007) but population data is often limited to the county level and does not provide information about morphology and urbanization levels (urban core, suburban, etc.). However, temperature data is easily obtained, and surface air temperature data from meteorological stations exist for some locations as far back as the early 1900s.

Temperature differentials have been shown to exist between urban and rural landscapes: a phenomenon known as the Urban Heat Island, or UHI (Shepherd et al., 2010). Dense residential areas, downtown business districts, or industrial areas have been shown to have the highest temperatures, whereas forested areas have the lowest temperatures and fields or low density residential areas have been shown to have temperatures in the moderate range (Saleh, 2010; Weng et al., 2004; Coutts et al., 2007).

In this study, temperature properties at meteorological stations located at known urban, suburban, and rural locations in the greater Washington, DC region were observed and their characteristics utilized to develop a new method for urban land type identification in order to help fill the gap in urban land use knowledge. For this study, the Washington, DC region was chosen for its complex urban morphology, varied rural terrain and known growth throughout recent decades.

2. Data and methods

2.1. Data and preliminary calculations

Temperature data were taken from the National Climatic Data Center TD3200 U.S. Cooperative Summary of Day dataset, published by the CISL Data Support Section at the National Center for Atmospheric Research, Boulder, CO (National Climatic Data Center, 2012a,b). Latitude and longitude and time active were determined by cross-referencing station numbers in the TD3200 dataset with the National Climatic Data Center Cooperative Stations Index (National Climatic Data Center, 2012b)

Surface meteorological station data was chosen because it is ubiquitous with a long history of record. It is accessible throughout the globe and generally easy to obtain at little to no cost. The data is conceptually and computationally simple to work with, compared to satellite or model output data formats, which make it an appealing data source for those with limited resources.

The TD3200 dataset is a quality-controlled daily-resolution dataset for meteorological ground stations in the National Weather Service Cooperative Station Network, published by the U.S. National Climatic Data Center, U.S. National Weather Service, and U.S. Federal Aviation Agency and updated monthly. The NWS Cooperative Station Network is a network of temperature and precipitation observation stations throughout the