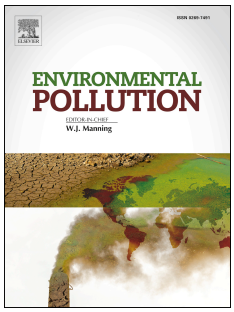
Journal Pre-proof



Estimation of anthropogenic heat emissions in China using cubist with points-of-interest and multisource remote sensing data

Qian Chen, Xuchao Yang, Zutao Ouyang, Naizhuo Zhao, Qutu Jiang, Tingting Ye, Jun Qi, Wenze Yue

PII: S0269-7491(20)31141-6

DOI: https://doi.org/10.1016/j.envpol.2020.115183

Reference: ENPO 115183

To appear in: *Environmental Pollution*

Received Date: 15 February 2020

Revised Date: 3 July 2020

Accepted Date: 3 July 2020

Please cite this article as: Chen, Q., Yang, X., Ouyang, Z., Zhao, N., Jiang, Q., Ye, T., Qi, J., Yue, W., Estimation of anthropogenic heat emissions in China using cubist with points-of-interest and multisource remote sensing data, *Environmental Pollution* (2020), doi: https://doi.org/10.1016/j.envpol.2020.115183.

Author Statement

Qian Chen: Methodology, Formal analysis, Writing - Original Draft

Xuchao Yang: Conceptualization, Writing - Review & Editing, Funding acquisition

Zutao Ouyang: Methodology, Software

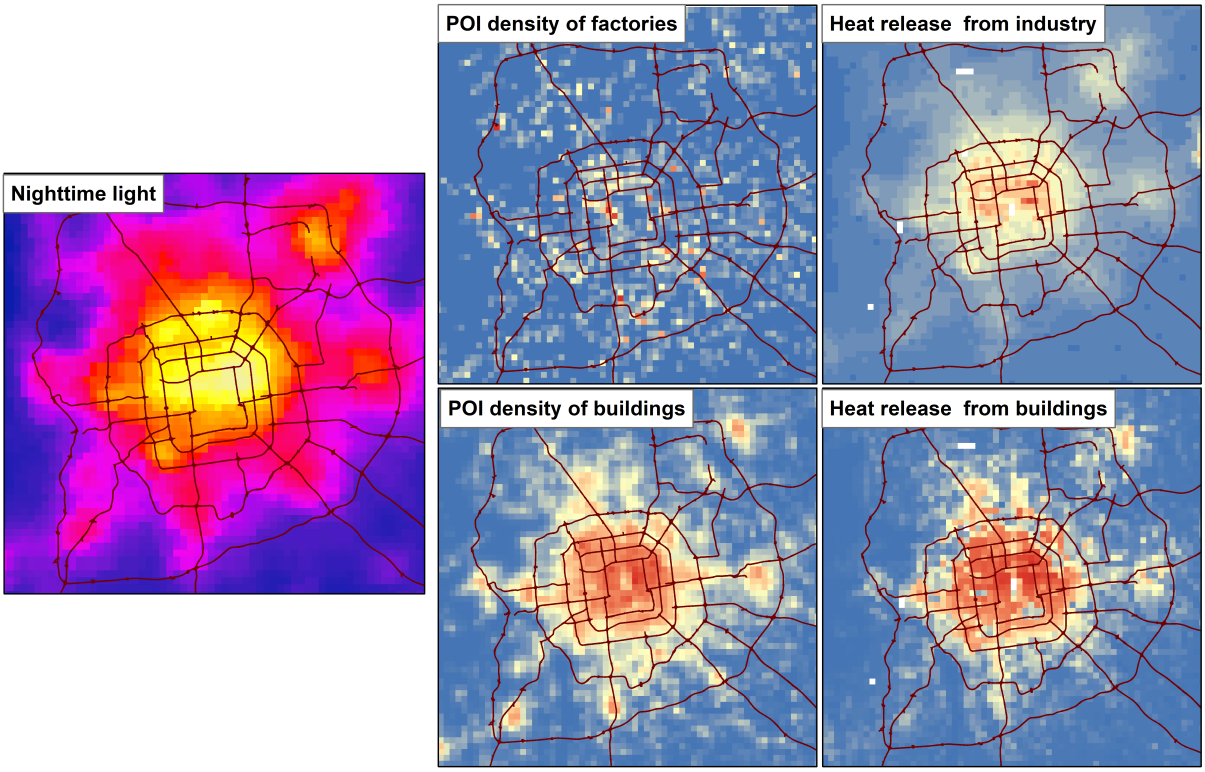
Naizhuo Zhao: Software, Writing - Review & Editing

Qutu Jiang: Software

Tingting Ye: Data Curation

Jun Qi: Resources

Wenze Yue: Writing - Review & Editing, Funding acquisition


Nighttime light
POI density of factories
Heat release from industry
POI density of buildings
Heat release from buildings

Estimation of anthropogenic heat emissions in China using Cubist with points-of-interest and multisource remote sensing data

Qian Chen[a], Xuchao Yang[a*], Zutao Ouyang[b], Naizhuo Zhao[c,d], Qutu Jiang[a] , Tingting Ye[a] , Jun Qi[e,f], Wenze Yue[g]

[a] Ocean College, Zhejiang University, Zhoushan, China

[b] Department of Earth System Science, Stanford University, Stanford, CA, USA

[c] Institute of Land Resource Management, School of Humanities and Law, Northeastern University, Shenyang, China

[d] Division of Clinical Epidemiology, McGill University Health Centre, Montreal, QC, Canada

[e] School of Environment, South China Normal University, Guangzhou, China

[f] Guangdong Provincial Key Laboratory of Chemical Pollution and Environmental Safety & MOE Key Laboratory of Theoretical Chemistry of Environment, South China Normal University, Guangzhou, China

[g] Department of Land Management, Zhejiang University, Hangzhou, China

*Corresponding author:

Dr. Xuchao Yang, Email: yangxuchao@zju.edu.cn, Ocean College, Zhejiang University, Zhoushan Campus, Haike Building 357, 1 Zheda Road, Zhoushan 316021, China. Telephone: +86-13735822563.

Resubmitted to *Environmental Pollution*
July 2020

Abstract

Rapid urbanization and industrialization in China stimulated the great increase of energy consumption, which leads to drastic rise in the emission of anthropogenic waste heat. Anthropogenic heat emission (AHE) is a crucial component of urban energy budget and has direct implications for investigating urban climate and environment. However, reliable and accurate representation of AHE across China is still lacking. This study presented a new machine learning-based top–down approach to generate a gridded anthropogenic heat flux (AHF) benchmark dataset at 1 km spatial resolution for China in 2010. Cubist models were constructed by fusing points-of-interest (POI) data of varying categories and multisource remote sensing data to explore the nonlinear relationships between various geographic predictors and AHE from different heat sources. The strategy of developing specific models for different components and exploiting the complementary features of POIs and remote sensing data generated a more reasonable distribution of AHF. Results showed that the AHF values in urban centers of metropolises over China range from 60 to 190 $W \cdot m^{-2}$. The highest AHF values were observed in some heavy industrial zones with value up to 415 $W \cdot m^{-2}$. Compared with previous studies, the spatial distribution of AHF from different heating components was effectively distinguished, which highlights the potential of POI data in improving the precision of AHF mapping. The gridded AHF dataset can serve as input of urban numerical models and can help decision makers in targeting extreme heat sources and polluters in cities and making differentiated and tailored strategies for emission mitigation.

Keywords: anthropogenic heat; Cubist; Points-of-interest; remote sensing; China

1. Introduction

Energy consumption by human activities in cities leads to the emissions of anthropogenic waste heat into the urban's land–atmosphere system. These emissions arise from industrial processing, space heating and cooling of buildings, vehicle exhausts, and human metabolism (Sailor, 2011). Anthropogenic heat emission (AHE), which plays a key role in the formation of urban heat island (UHI) (Crutzen, 2004; Oke, 1988), is an important but frequently ignored or excessively simplified component of urban climate and environment studies (Sailor, 2011; Sailor and Lu, 2004). Overheating in cities and AHE have a two-way interaction. Rising temperatures caused by AHEs could potentially trigger increasing demands on air conditioning in ill-ventilated cities, which in return lead to considerable AHEs. Moreover, AHEs may substantially alter the evolution of aerosol and air pollutants in urban areas because they are simultaneously released from energy consumption (Crutzen, 2004). Therefore, AHE has potentially significant implications for climate and air quality in urban areas at the regional and global scales (Crutzen, 2004).

Owing to the important role in urban weather, climate, and atmospheric environment, AHE has been regarded as a representative urbanization-related force in mesoscale models. For example, simulations by Fan and Sailor (2005) suggested that the AHE in winter contributes 2 °C–3 °C to the nighttime UHI of Philadelphia, USA. Recent modeling studies have suggested that the AHE in Hangzhou, east China contributes 65.26% and 17.47% to the UHI intensity in winter and summer, respectively (Chen et al., 2016a). The contribution of AHE to the increase in total heat discomfort hours during summer is approximately equal to that because of urban land use change in the urban areas of Yangtze River Delta (Yang et al., 2019). Moreover, incorporating AHE into the modeling system can influence the spatial and vertical

distributions of the simulated air pollutants (Xie et al., 2016a) and improve air quality forecast (Yu et al., 2014; Yang et al., 2018). For example, surface O_3 concentrations increase in the urban areas because of AHE-induced rising temperature (Xie et al., 2016a). Therefore, modeling efforts have aimed to examine the urban climate and atmospheric environment must appropriately characterize the spatial–temporal profiles of AHE. However, such data are lacking in most cities worldwide, thereby increasing the likelihood of modelers simply using the default AHE profiles (Wang et al., 2013; Chen et al., 2014a; Chen et al., 2016b) or even ignoring the AHE effect (Wang et al., 2012; Oleson et al., 2015; Conlon et al., 2016; Morris et al., 2017; Wang et al., 2017).

AHE is closely dependent on the energy consumption in the area of interest and is difficult to determine because its pattern strongly varies and cannot be measured directly. Given the high demand for spatially explicit AHE, increasing efforts have been exerted to estimate anthropogenic heat flux (AHF) leveraging inventory-based method, energy balance residual, building energy models (Sailor, 2011), multimethod (Chow et al., 2014) and statistical regression method (Lee et al., 2014; Lee and Kim, 2015; Koralegedara et al., 2016). Except for the top–down inventory method, most of the abovementioned methods are difficult to implement at the regional to global scale, especially in developing countries because of limited data. The classical top–down inventory-based method, which requires the energy consumption data from major components, namely, industries, buildings, vehicles, and human metabolism, obtained at aggregate scales, has great benefit to provide a wide area coverage for regional and global applications (Sailor and Lu, 2004). The annual total bulk energy consumption data at different administrative units are spatially disaggregated to scales of interest

using gridded ancillary covariate datasets, such as population density (Sailor and Lu, 2004; Flanner, 2009; Allen et al., 2011; Sailor et al., 2015; Lu et al., 2017), nighttime light (NTL) (Chen et al., 2014b; Dong et al., 2017; Yang et al., 2017; Yang et al., 2020), land use (Lee et al., 2009), and multisource remote sensing (RS) data (Chen et al., 2017; Chen et al., 2019; Wang et al., 2019).

The spatial distribution of urban land use is crucial in estimating AHE because the largest mean fluxes generally originate from industrial or commercial areas, which have greater energy consumption compared with residential areas (Chow et al., 2014; Sun et al., 2018). Commercial areas are usually concentrated in urban downtown areas with high population densities, whereas industrial areas are located in suburban areas or the urban–rural boundary. Urban mapping using remote sensing data is limited to land cover monitoring because land use has a strong correlation with human socioeconomic features and is difficult to infer from physical infrastructure, especially on mixed urban environments (Liu et al., 2015; Hu et al., 2016; Liu et al., 2017). Therefore, remote sensing data and population density cannot effectively capture different socioeconomic activities related to AHE. For example, industrial areas and commercial centers have similar NTL brightness and vegetation cover, thereby resulting in the misdistribution of AHF from factories and commercial buildings.

The recent emergence of geospatial big data with spatial interaction and place semantic information provides new opportunities to identify urban land use and improve AHF estimation in complex urban areas. Points-of-interest (POI) data are a type of geospatial big data that is particularly promising for this purpose. The data contain information about location (coordinates), short textual description, and the

category it belongs (Yoshida et al., 2010; McKenzie et al., 2015). POIs highly relate to different human socioeconomic activities and have been increasingly used to identify urban function districts and urban land use types (Jiang et al., 2015; Hu et al., 2016; Kuang et al., 2016; Gao et al., 2017; Liu et al., 2017; Yao et al., 2017; Zhang et al., 2017a; Wang et al., 2018). Therefore, POIs have the potential to better distinguish areas with high AHF from different heat resources, especially in cities. As far as we know, there are no reports on using POI data to refine AHF mapping.

China has experienced rapid economic development and unprecedented urbanization since the start of its economic reforms in the late 1970s. As a result, energy consumption has increased greatly, which implied remarkable increase in AHE. At the same time, the UHI effect has strongly intensified (Zhou et al., 2015) and received much attention in the literature. Recent modeling studies with a simple treatment of AHF have suggested that AHE plays an important role in simulating UHI and urban atmospheric environment in Chinese big cities, such as Beijing, Shanghai, and Hangzhou (Chen et al., 2014a; Yu et al., 2014; Chen et al., 2016a; Xie et al., 2016a; Xie et al., 2016b; Yang et al., 2018; Yang et al., 2019). However, a lack of detailed and accurate AHF at large scales has created a barrier for understanding the relationship between AHE and urban environment (Park et al., 2016). The present work aims to provide an updated and comprehensive view of the AHE across China for satisfying the growing demand for a national database of AHF profiles and fill some gaps of previous studies. A new machine learning-based top–down approach that incorporates multisource remote sensing and POI data in a flexible Cubist estimation technique was presented to generate an accurate gridded AHF benchmark

dataset at 1 km spatial resolution in China. To the best of our knowledge, this study is the first to use machine learning method and incorporate geospatial big data to estimate AHF.

2. Data and preprocessing

This study takes mainland China as the study area. Hong Kong, Macao, and Taiwan are excluded because of their distinct political and economic status from mainland China. The demographic, socioeconomic, and energy consumption data at the provincial and prefecture levels in mainland China for top–down energy inventory approach were collected from the China Statistical Yearbook and China Energy Statistical Yearbook for 2011. Information, such as statistical population, gross domestic product (GDP) from three economic sectors, and the number of civil automobiles were included.

The POI data were derived from the Baidu Map (http://map.baidu.com), which is the largest desktop and mobile map service provider in China (Yao et al., 2017). With the help of the application programming interface provided by Baidu Map Service, we fetched 5,152,850 Baidu POI records within 20 categories for 2010, such as factories, commercial buildings, educational facilities, and residential communities. Location and attribute information in the form of a Chinese phrase were included for each POI record.

The road network data in 2010, which were acquired from the Data Center for Resources and Environmental Sciences of Chinese Academy of Sciences, include China's national highways, expressways, railways, provincial and county-level roads, and city roads. In this study, the Chinese road network data were used to generate the corresponding road density raster layers (Rd-density) and the Euclidean distance to the

173 closest road layers (DtC-Rd) at a spatial resolution of 1 km on ArcGIS 10.4.1.

174 Population density data. The gridded population map in 2010 for mainland China
175 at 100-m spatial resolution from Ye et al. (2019) was used in this study. This dataset
176 shows higher accuracy than the WorldPop dataset in mainland China, especially on
177 highly or lowly populated areas. We aggregated this population map into a new layer
178 with a spatial resolution of 1 km.

179 Multisource remote sensing data, namely, Defense Meteorological Satellite
180 Program Operational Linescan System (DMSP/OLS) NTL, Normalized Difference
181 Vegetation Index (NDVI), digital elevation model (DEM), land surface temperature
182 (LST), and Global Urban Footprint (GUF) product were also used as ancillary covariate
183 datasets in the disaggregation process. The source of these remote sensing data are
184 illustrated in Supplementary Materials S1.

185 **3. Methodology**

186 Although the POIs can provide specific locations of various urban facilities, the
187 availability of energy consumption data at building level is a big challenge for large
188 scale studies, especially in developing countries like China. The top–down approach
189 only requires the accessible energy consumption data from major components obtained
190 at aggregated census units (Sailor et al., 2004; Sailor et al., 2011). The practicality and
191 relative simplicity of this method allow the incorporation of easily constructed and
192 readily available data of urban characteristics with less effort for large spatial scales.
193 Therefore, the top–down inventory-based method was adopted in the present study.

194 **3.1 Estimating prefectural-level AHE from different sources**

195 We first estimate the aggregated AHE from four components at the

196 prefecture-level of China (using AHE_i, AHE_b, AHE_t, and AHE_m for the AHE from
197 industry, buildings, transportation, and human metabolism, respectively). In this study,
198 we considered energy types including coal (raw coal, cleaned coal, other washed coal,
199 briquettes, and other coking products), oil (crude oil, gasoline, kerosene, diesel, fuel
200 oil, and other petroleum products), gas (liquefied petroleum gas, natural gas, coke
201 oven gas, blast furnace gas, and converter gas), electricity and thermal force. To
202 facilitate calculation, all kinds of consumed energy were converted to standard coal
203 equivalents. The total amount of AHE can be expressed as follows:

$$AHE = AHE_i + AHE_b + AHE_t + AHE_m, \quad (1)$$

204 Based on the energy flow chart of China (Li et al., 2006), industrial waste heat
205 not only releases by the final industrial energy consumption but also by the energy
206 losses during the conversion of raw energy to useful energy. The amount of lost
207 energy in this study was calculated based on the efficiency of energy conversion,
208 which is the ratio of the output useful energy to the input primary energy. The
209 provincial AHE_i was calculated firstly and then allocated to prefecture-level cities by
210 industrial GDP, which can be expressed as follows:

$$AHE_i = \eta_i[(1 - E) \times C_1 + C_2] \times C, \quad (2)$$

211 where C_1 is the total primary energy consumption at the province level. E is the
212 total efficiency of energy conversion for the year 2010 and is equal to 72.83% in this
213 study based on the statistics of the China Energy Statistical Yearbook (Department of
214 Energy Statistics, 2011). C_2 is the provincial amount of the final industrial energy
215 consumption. C is the standard coal heat (29,306 $kJ \cdot kg^{-1}$) and η_i is the proportion
216 of industrial GDP of each prefectural-level city to that of the corresponding province.

217 AHE_b, AHE_t, and AHE_m can be calculated as:

$$AHE_b = (\eta_c C_u + \eta_p C_r + \eta_g C_c) \times C, \quad (3)$$

$$AHE_t = N_t \times d \times CE \times NHC \times \rho, \quad (4)$$

$$AHE_m = N_p \times (P_s \times 8 + P_a \times 16) \times 3{,}600 \times 365, \quad (5)$$

where C_u and C_r are the provincial household energy consumption of urban and rural residents, respectively. C_c is the provincial commercial energy consumption. η_c, η_p, and η_g are the proportion of the urban, rural, and total population of each prefecture-level city to that of the corresponding province, respectively. N_t is the number of civil automobiles in each prefecture-level city. d is the annual average driving distance per vehicle (2.5×10^4 km) (Tong et al., 2004). CE is the combustion efficiency (0.127 $L\cdot km^{-1}$). NHC is the net heat combustion (45 $kJ\cdot g^{-1}$). ρ is the combustion density (0.738 $kg\cdot L^{-1}$). N_p represents the population at the prefecture-level, and P_s and P_a denote the metabolic rates of human during sleeping (23:00–7:00, 70 W) and active time (7:00–23:00, 171 W), respectively (Grimmond and Oke, 1999; Quah and Roth, 2012). The estimated prefectural-level AHE from different components were then spatially joined to the corresponding GIS-based administrative boundaries (339 units in total).

3.2 Selection of predictors for AHF mapping

Variable screening under the guidance of prior knowledge (Liang et al., 2015) and evaluation of variable relevance and importance in the implementation of machine learning methods serve as effective mechanism for maximizing model predictability and transferability (Houborg and McCabe, 2018). A suite of independent variables for each Cubist model was accordingly selected. Various POIs that indicate physical infrastructure for various socioeconomic activities can be considered a good proxy of AHE distribution. A region with more factory POIs or close to factory POIs usually have large AHE_i. Similarly, POIs of buildings, such as banks, retail, restaurants, accommodation services, and companies, represent the locations of human activities

related to high AHE_b. Transportation-related POIs (e.g., motor passenger station, bus station, and auto service) indicate the locations of critical transportation hubs and infrastructure with high AHE_t. Correlation analysis (Supplementary Materials S2) shows strong associations between a category of POI and the corresponding AHE. Therefore, POI data were the first predictor for AHF estimation. Each category of the POI was used to create the corresponding raster layer of POIs density (POIs-den) and the distance to the nearest POIs (DtN-POIs) at 1 km × 1 km spatial resolution. Building-related POI categories were combined into raster layers of POIs-den-Build and DtN-POIs-Build for AHE_b estimation and transportation-related POI categories were combined into raster layers of POIs-density-Trans and DtN-POIs-Trans for AHE_t estimation. Factory-related POIs were used to generate raster layers of POIs-den-Fac and DtN-POIs-Fac for AHE_i estimation. The processing of POI data is detailed in Supplementary Materials S3.

Except for POI, AHE also has strong relationships with road network and various RS data that related to human activities, such as NTL (Shin et al., 2015; Liu et al., 2018), LST (Fujimoto et al., 2012; Zhou et al., 2018; Yang et al., 2019; Zhang et al., 2019), NDVI (Chen et al., 2016a; Chen et al., 2019), DEM, and GUF built-up areas. Therefore, factory POIs, distance to railways, NTL, GUF, slope, NDVI, and LST were used in the Cubist model for AHE_i estimation. Building-related POIs, NTL, NDVI, distance to the county and city roads, elevation, and GUF were selected as explanatory variables for AHE_b. Transportation-related POIs, various road density layers, slope, elevation, and NTL data were used as predictors for AHE_t. Gridded population data from Ye et al. (2019) were considered the proxy of the distribution of AHE_m. A summary of the data used for Cubist fitting is illustrated in Supplementary Materials (Table S1).

3.3 Cubist model fitting

Cubist is characterized as a rule-based model tree method proposed by Quinlan (1992, 1993). At each leaf node on the tree of the Cubist model, a multivariate linear regression model is fitted on the basis of the data subset defined by sets of rules. To improve the predictive accuracy, the committee models (a boosting-like scheme for creating iterative model trees in sequence) and instance-based corrections (using nearest neighbors to adjust predictions) were combined to the rule-based Cubist model. The main advantage of the Cubist regression method is its ability to deal with nonlinear and complex relationships between dependent and independent variables using both continuous and categorical variables as inputs. In addition, Cubist can give the relative importance of the predictors, allowing an easy model interpretation. Recently, Cubist has successfully been used in various fields (Ma et al., 2017; John et al., 2018; Xu et al., 2018) and is recognized as an efficient algorithm in terms of computation time (Walton, 2008). In this study, Cubist models were built for different AHE components to effectively capture the complex associations between various geographic variables and the target AHE profile. We aggregated all 1 km geographical predictors to prefectural-level, which were coordinated with bulk AHE estimation. The logarithm of the annual mean AHF value of each prefectural (the dependent variable) and the geographic covariates (the independent variables) were used to fit the Cubist models for different heat sources. The Cubist models representing the relationships between AHF and related geographic covariates at prefectural-level were then applied to the corresponding geographic covariates at 1 km resolution to obtain the primary estimates of AHF. The resulting AHF estimations were used as a weighting layer for a standard dasymetric mapping approach (Stevens et al., 2015) to disaggregate prefectural-level AHE to produce final gridded AHF maps. The workflow of data processing, Cubist model fitting, and dasymetric AHF mapping are shown in Figure 1.

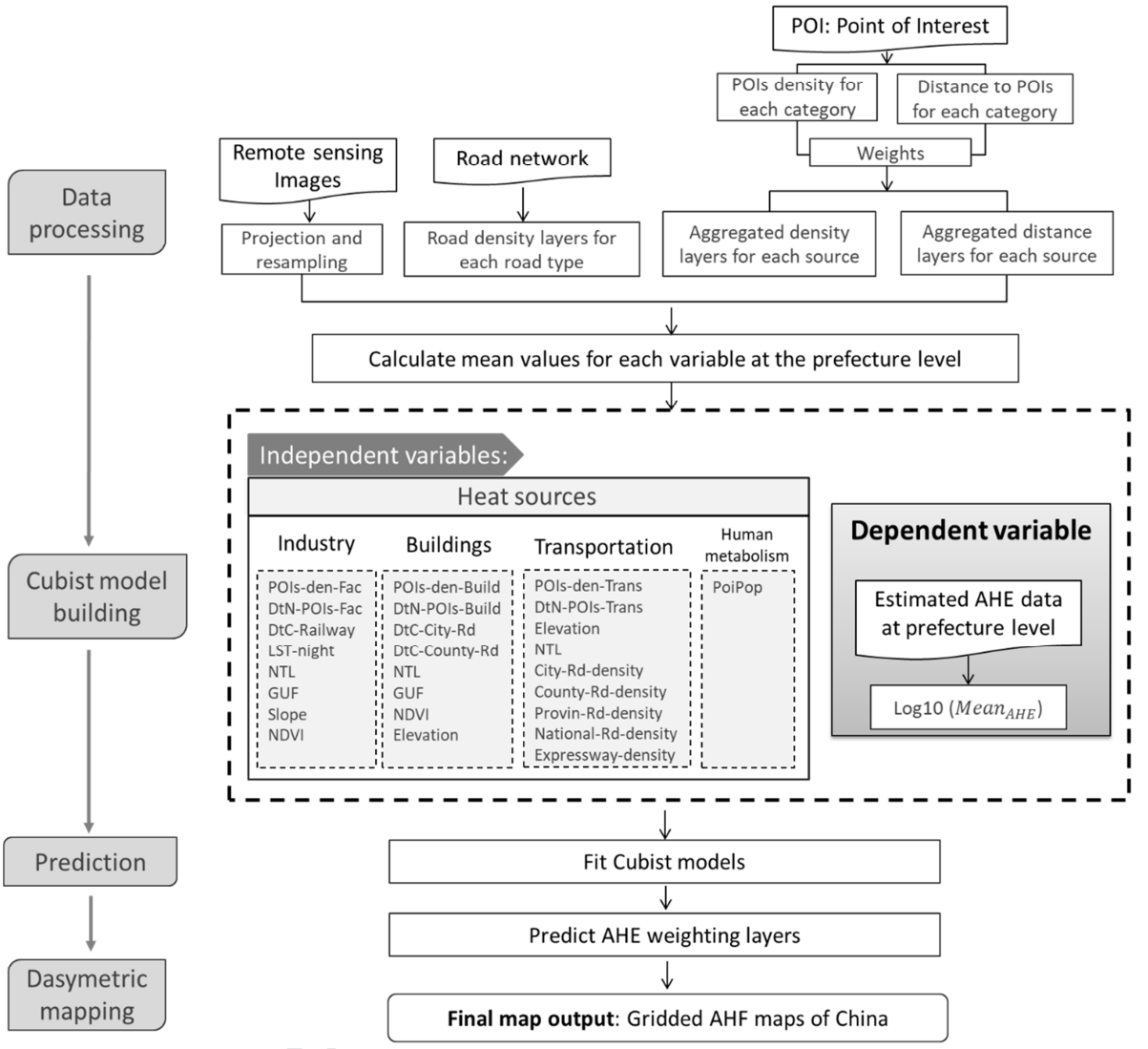


Figure 1. Flow diagram for producing the anthropogenic heat flux maps of China.

4 Results

4.1 AHE in typical prefecture-level cities

Using the top–down inventory-based approach, the total AHE and AHE from four sources were estimated at the prefecture-level for mainland China. For the year 2010, the AHE_i, AHE_b, AHE_t, and AHE_m are 8.4×10^{19}, 9.60×10^{18}, 1.34×10^{19}, and 5.78×10^{18} J, accounting for 74.48%, 8.51%, 11.88%, and 5.13% of the total AHE in China (1.13×10^{20} J), respectively. The AHE_i largely contributes to the total AHE of

China. In prefecture-level cities, AHE from Shanghai was the highest, which reaches 2.41×10^{18} J, followed by Beijing, Chongqing, and Tianjin. As shown in Figure S1, the AHE in typical cities of China mainly came from industries and vehicles, whereas the AHE of Beijing, Shanghai, and Tianjin cities mostly came from industries and buildings. The mean AHF in typical cities ranges from 0.5 $W \cdot m^{-2}$ to 8.5 $W \cdot m^{-2}$. High AHFs generally occurred in large cities (e.g., Shanghai, 8.5 $W \cdot m^{-2}$). However, the AHFs of some cities (e.g., Taiyuan and Zhengzhou) with a number of heavy industries, are higher than those of metropolises (e.g., Beijing and Chongqing) because of the small spatial areas. The results of the estimated annual total AHE and mean AHF of most typical cities are parallel with those reported by Lu et al. (2016).

4.2 Pixel-based distribution of AHF

With the joint use of multisource RS data and POIs in Cubist models, the gridded AHF maps over China for 2010 from industry, buildings, transportation, human metabolism and the total AHF were created with a spatial resolution of 1 km (Figure 2). In general, high AHF pixels from four sources were located in major urban centers and industrial areas in big cities. The AHF in western China was significantly lower than those in eastern China. For the total AHF of China, high AHF is mainly concentrated on economically developed urban areas, especially megacities, such as Beijing, Shanghai, and Guangzhou, with AHF of 100–190 $W \cdot m^{-2}$ across the downtown area (Figure 2d). However, the built-up area in some medium-sized cities (e.g., Handan, Taiyuan, and Shijiazhuang) has the highest total AHF up to 415 $W \cdot m^{-2}$ because of high intensity of industrial activities in small spatial areas. This finding is

consistent with the results of Chen et al. (2019) which found that there was no absolute positive relationship between AHF and economic level.

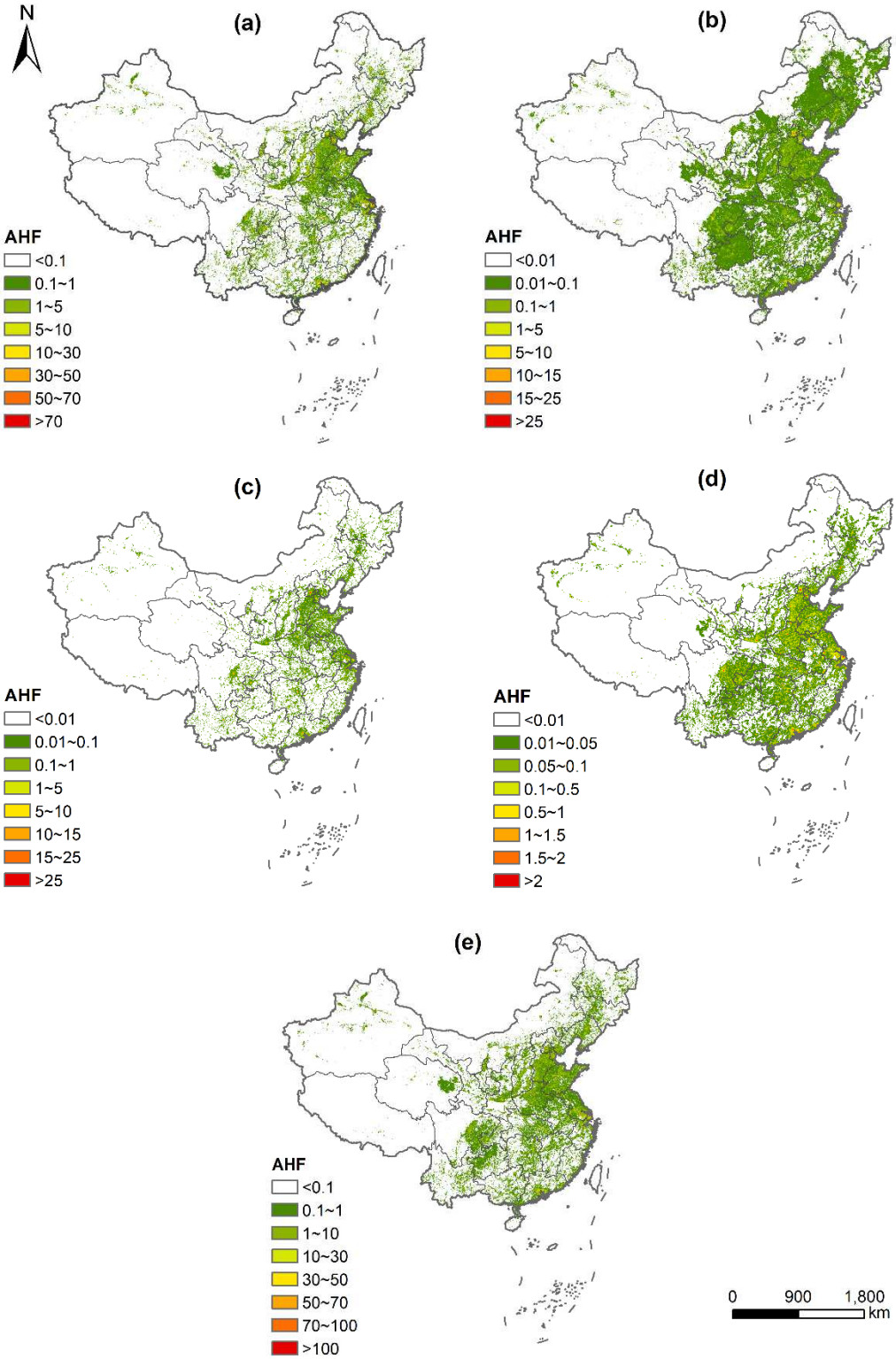
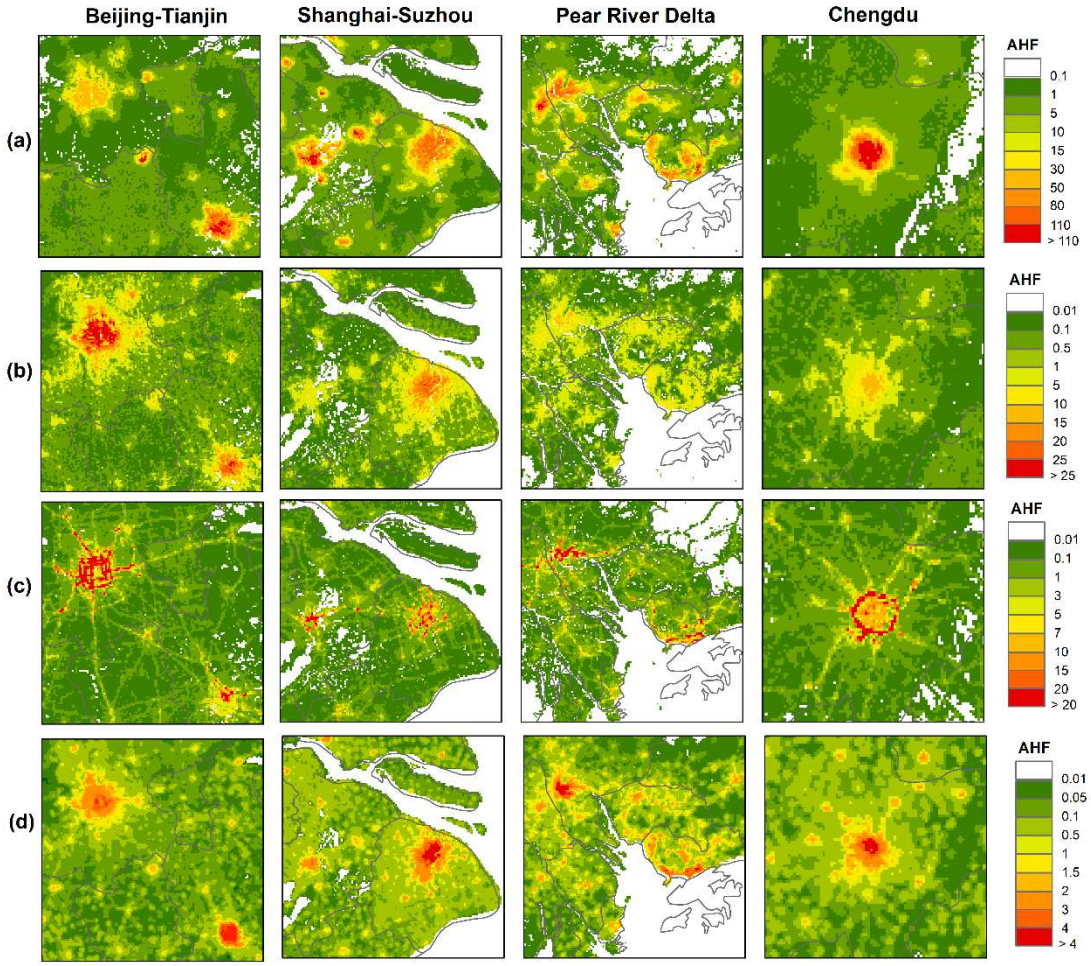


Figure 2. AHF maps for (a) industry (AHF_i), (b) buildings (AHF_b), (c)

transportation (AHF_t), (d) human metabolism (AHF_m), and (e) total AHF in 2010 (unit: $W \cdot m^{-2}$).

As a further demonstration of the advantage of separately modeling China's AHF with joint use of RS data and POIs, Figure 3 shows the predicted AHF from different sources in four metropolitan areas of Beijing–Tianjin, Shanghai–Suzhou, the Pearl River Delta, and Chengdu. Areas with high AHF_i are mainly distributed in the urban fringe of megacities or some medium-sized cities. Although most industries have been relocated away from city centers, there are still some large factories close to the urban center of megacities, such as Chengdu and Chongqing. By contrast, high AHF_b and AHF_m mainly occur in the dense residential and commercial areas in urban centers, exhibiting the characteristic of agglomeration. For the transportation sector, the spatial distribution of high AHF_t is characterized by interconnecting road network in urban areas. For example, evident ringed high AHF_t are found across Beijing and Chengdu. High AHF_t values also can be observed at major roads in the urban centers of Shanghai, Guangzhou, and Shenzhen. The comparison between our AHF maps and other AHF products is shown in Supplementary Materials S4 and Figure S2. The results demonstrate that the proposed methodology effectively captures the main spatial characteristics of AHF from different sources and produces a reasonable spatial distribution of AHF. Comparing the total AHF with mean nighttime LST during summer, we found that the spatial patterns of AHF are in good agreement with those of LST in four cities (Figure S2). Urban centers with larger AHF values generally hold high LST.



349

350 **Figure 3. Comparison of AHF from (a) industry, (b) buildings, (c) transportation,**
351 **and (d) human metabolism for four metropolitan agglomerations in China (unit:**
352 **$W \cdot m^{-2}$).**

353 **4.3 Geographical covariate importance**

354 Variable importance analyses were conducted in order to assess the relative
355 contributions of the geographical covariate to the AHF modeling process. On a large
356 scale across China, two topography-related variables, namely, elevation and slope, are
357 important to predict AHF from three sources, especially for AHF_b. Harsh topography
358 discourages human settlement and socioeconomic activities. High AHF is
359 concentrated in low-elevation, flat areas across China.

A built-up area is a fundamental carrier of industrial activities. At night, ceaseless industrial processes release large amounts of visible and near-infrared radiations and waste heat, thereby lighting industrial areas and changing their land surface thermal radiance (Zhao et al., 2013; Zhang et al., 2017b). Therefore, the proportion of built-up area and the brightness of NTL are the most important variable in mapping AHF_i (Figure 4a). Factories tend to be situated in areas accessible to good road infrastructure (Fan and Chan-Kang, 2008). POI records of factories provide accurate positions of industrial infrastructures. Vegetation cover is closely and negatively correlated with impervious surfaces (Weng et al., 2006). Thus, road networks, NDVI, LST, and POI-related variables can further refine the estimation of AHF_i.

The brightness of NTL, which can be considered a proxy of the intensity of traffic activity (Shin et al., 2015), is the most important predictor in the final Cubist model for AHF_t mapping (Figure 4b). Roadway type and its density are necessary to allocate AHF_t at fine-scale (Sailor, 2007). The density of various types of roads plays certain roles in AHF_t mapping, especially the density of city and county roads. Transportation-related POIs, which provide various types of services to vehicles, also have large contributions to AHF_t mapping (Figure 4b).

Various types of POIs provide accurate positions of buildings, thereby better representing an area with high AHF_b and excluding industrial regions. The distance to the nearest POIs of buildings provides the largest contribution to the final Cubist model of AHF_b, followed by elevation, built-up area, and road networks (Figure 4c). The brightness of NTL is less important in modeling AHF_b because NTL cannot effectively distinguish different functions in complex cities, such as commercial zones and residential areas.

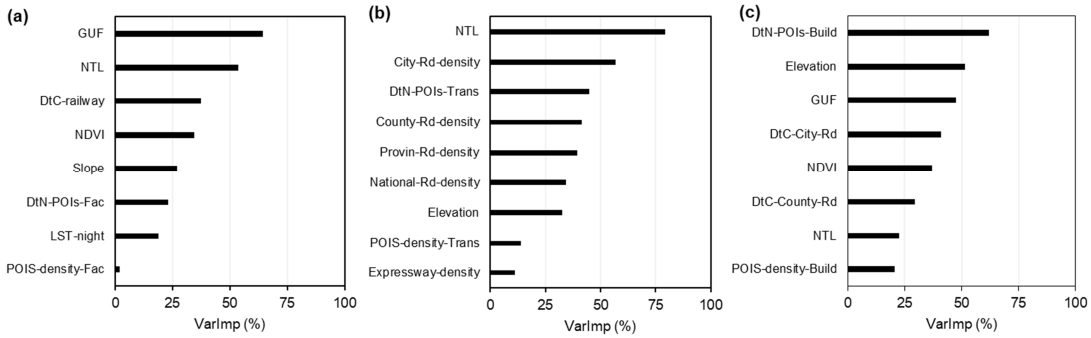


Figure 4. Variable importance of the covariates in the Cubist model for the estimation of AHF from (a) industry, (b) transportation, and (c) buildings

5 Discussion

5.1 Advantages of POIs in refining AHF estimation

Most studies on AHF estimation in China have used the top–down method because of the limited data. Population density (Lu et al., 2016), DMSP/OLS NTL data (Chen et al., 2012; Yang et al., 2014), and a combination of NTL and vegetation index (Chen et al., 2017; Chen et al., 2019; Wang et al., 2019) were recently used as surrogates of AHF. However, the accuracy of resulting AHF datasets suffers from the inherent limitation of RS data and population density. The industrial sector contributed the largest proportion of total AHE in most cities over China (Lu et al., 2016; Chen et al., 2019). Simultaneously, cities in China have experienced industrial adjustments, and many factories, especially heavy industries, have been shifted from urban core areas to suburban or rural areas. Therefore, using NTL or population density as ancillary data for AHF mapping may significantly overestimate the AHF_i in urban centers (Figure S3). In suburbs and some less-urbanized cities, where many factories are located, AHF_i may be underestimated. Moreover, in some regions that have mixed functions, it is very difficult to distinguish AHE from different heat

sources with the sole use of NTL data or population density.

These problems can be alleviated when POI data are incorporated. Compared with population density and RS data, POIs have an inherent advantage in indicating the actual position of different human activities, especially in cities. Thanks to its thematic richness and strong spatial heterogeneity, POI data allow urban functions to be partially discerned, which is beneficial in disaggregating AHE from different sources separately (Figure S3). The density of (or the distance to) a certain type of POI approximately shows the spatial pattern of a certain heat source. Statistical outputs from the Cubist model and resulting AHF maps highlight that POI-related variables can effectively represent the distribution of AHF_b and distinguish it from industrial area (Figures 3 and 4). Information extracted from POIs and multisource remote sensing imagery can complement each other to yield reasonable and precise AHF maps, especially in urban areas. The POI data used in this study were based on commercial navigation data, which are collected by trained persons and undergo rigorous inspections and corrections. As a result, the positional and thematic accuracy of these commercial POI data is reliable. The preeminence of navigation-based POI data as good indicators of various economic activities makes them prominent in spatially disaggregating AHE from different sources to a fine geographic scale.

5.2 Advantages of the Cubist algorithm

The relationships between AHF and the geographical covariates are not constant in space. Previous studies on AHF estimation have used constant relationships, including quadratic functions (Yang et al., 2014; Chen et al., 2017; Chen et al., 2019)

425 or linear functions (Wang et al., 2019), which implies stable relationships between the
426 dependent and explanatory variables in space. In this study Cubist algorithm o□ered a
427 clear advantage in the consideration of the non-stationary relationships between AHF
428 and the geographical covariates, and to select the optimal combinations of explanatory
429 variables for different heat sources. It is also cost effectively due to the utilization a
430 suite of publicly available RS data like DMSP/OLS NTL, DEM, LST, GUF, and
431 NDVI. Another major advantage of Cubist is that it can effectively deal with
432 non-parametric data.

433 **5.3 Limitations**

434 The top–down inventory approach suffers from an inherent limitation that it
435 makes the assumption that energy consumption is equivalent to anthropogenic
436 sensible heat emissions, with no time lag (Sailor, 2011). Despite the advantages of
437 integrating POI and RS data, the accuracy of AHF estimation is potentially limited by
438 the uncertainty of the POIs. Firstly, in urban fringe areas or rural areas, the POI
439 records are significantly less than that in urban areas, and thus have limited capability
440 to estimate AHF in these areas (Chen et al. 2018). Moreover, the POI data can only
441 reflect the geographic and thematic characteristics of various facilities and do not
442 directly indicate the actual extent and intensity of different socioeconomic activities,
443 which also affect the accuracy of AHF estimation. For example, although large steel
444 plants and small factories have distinct AHE, their POIs are treated equally in the
445 modeling process. Further incorporation of building volume and height data would
446 provide supplemental information to identify the extent and intensity of

447 socioeconomic activities in specific fine-scale areas.

448 **6. Conclusions**

449 Based on officially published statistical data, this study adopted the top–down
450 inventory approach for bulk AHE estimation to provide a wider area coverage for
451 regional applications over China. We developed a novel method for gridded AHF
452 mapping through Cubist model with joint use of POIs and multisource remote sensing
453 data. Individual Cubist model was constructed to capture the nonlinear relationships
454 between varied covariates and AHF from each heat source. The unique feature of POI
455 in identifying urban functional regions could help to effectively distinguish the
456 distribution of AHF from different heat components. As a result, we generated a
457 reasonable estimation of pixel-based AHF related to actual human activities over
458 mainland China for the year 2010 at 1 km spatial resolution. The AHF in urban
459 centers of metropolises was in a range of 60 to 190 $W \cdot m^{-2}$. The maximum AHF in
460 some industrial zones reached 415 $W \cdot m^{-2}$. The new gridded AHF dataset with high
461 spatial heterogeneity can serve as a fundamental aspect of the urban energy balance in
462 mesoscale modeling of the urban atmospheric environment. Such modeling will
463 provide valuable insights into how spatiotemporal variability in AHE affects the
464 development of the UHI and its resultant impacts on air quality in major cities over
465 China.

466 **Acknowledgements**

467 The authors are very grateful to the three anonymous reviewers for their helpful
468 comments and constructive suggestions, which led to a significant improvement of the
469 original manuscript. This work was supported by the National Natural Science
470 Foundation of China (No. 41671035, 41671533).

References:

Allen, L., Lindberg, F., & Grimmond, C. (2011). Global to city scale urban anthropogenic heat flux: model and variability. *Int. J. Climatol., 31*, 1990–2005

Chen, B., Dong, L., Shi, G., Li, L.-J., & Chen, L.-F. (2014b). Anthropogenic heat release: estimation of global distribution and possible climate effect. *J. Meteorol. Soc. Jpn. Ser. II, 92A*, 157–165

Chen, B., Shi, G., Wang, B., Zhao, J., & Tan, S. (2012). Estimation of the anthropogenic heat release distribution in China from 1992 to 2009. *Acta Meteorol. Sin, 26*, 507–515

Chen, F., Yang, X., & Wu, J. (2016a). Simulation of the urban climate in a Chinese megacity with spatially heterogeneous anthropogenic heat data. *J. Geophys. Res.: Atmos., 121*, 5193–5212

Chen, F., Yang, X., & Zhu, W. (2014a). WRF simulations of urban heat island under hot-weather synoptic conditions: The case study of Hangzhou City, China. *Atmos. Res., 138*, 364–377

Chen, L., Zhang, M., & Wang, Y. (2016b). Model analysis of urbanization impacts on boundary layer meteorology under hot weather conditions: a case study of Nanjing, China. *Theor. Appl. Clim., 125*, 713–728

Chen, S., & Hu, D. (2017). Parameterizing anthropogenic heat flux with an energy-consumption inventory and multi-source remote sensing data. *Remote Sens., 9*, 1165

Chen, S., Hu, D., Wong, M.S., Ren, H., Cao, S., Yu, C., & Ho, H.C. (2019). Characterizing spatiotemporal dynamics of anthropogenic heat fluxes: a 20-year case study in Beijing–Tianjin–Hebei region in China. *Environ. Pollut, 249*, 923–931

Chen, X., Li, X., Yuan, X., Zeng, G., Liang, J., Li, X., Xu, W., Luo, Y., & Chen, G. (2018). Effects of human activities and climate change on the reduction of visibility in Beijing over the past 36 years. *Environ Int., 116*, 92–100

Chow, W.T.L., Salamanca, F., Georgescu, M., Mahalov, A., Milne, J.M., & Ruddell, B.L. (2014). A multi-method and multi-scale approach for estimating city-wide anthropogenic heat fluxes. *Atmos. Environ., 99*, 64–76

Conlon, K., Monaghan, A., Hayden, M., & Wilhelmi, O. (2016). Potential impacts of future warming and land use changes on intra-urban heat exposure in Houston, Texas. *PLoS One, 11*, e0148890

Crutzen, P.J. (2004). New directions: the growing urban heat and pollution “island” effect-impact on chemistry and climate. *Atmos. Environ., 38*, 3539–3540

Dong, Y., Varquez, A.C.G., & Kanda, M. (2017). Global anthropogenic heat flux database with high spatial resolution. *Atmos. Environ., 150*, 276–294

Department of Energy Statistics. 2011. China Energy Statistical Yearbook. China Statistics Press: Beijing (in Chinese).

Fan, H., & Sailor, D. (2005). Modeling the impacts of anthropogenic heating on the urban climate of Philadelphia: a comparison of implementations in two PBL schemes. *Atmos. Environ., 39*, 73–84

Flanner, M.G. (2009). Integrating anthropogenic heat flux with global climate models. *Geophys. Res. Lett., 36*, L02801

Fujimoto, A., Saida, A., & Fukuhara, T. (2012). A new approach to modeling vehicle-induced heat and its thermal effects on road surface temperature. *J. Appl. Meteorol. Clim., 51*, 1980–1993

Gao, S., Janowicz, K., & Couclelis, H. (2017). Extracting urban functional regions from points of interest and human activities on location-based social networks. *Trans. GIS, 21*, 446–467

Grimmond, C.S.B., & Oke, T.R. (1999). Heat storage in urban areas: Local-scale observations and evaluation of a simple model. *J. Appl. Meteorol., 38*, 922–940

Houborg, R., & McCabe, M.F. (2018). A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning. *ISPRS J Photogramm*, 135, 173–188

Hu, T., Yang, J., Li, X., & Gong, P. (2016). Mapping urban land use by using landsat images and open social data. *Remote Sens., 8*, 151

Jiang, S., Alves, A., Rodrigues, F., Ferreira Jr, J., & Pereira, F.C. (2015). Mining point-of-interest data from social networks for urban land use classification and disaggregation. *Comput., Environ. Urban Syst., 53*, 36–46

John, R., Chen, J., Giannico, V., Park, H., Xiao, J., Shirkey, G., Ouyang, Z., Shao, C., Lafortezza, R., & Qi, J. (2018). Grassland canopy cover and aboveground biomass in Mongolia and Inner Mongolia: Spatiotemporal estimates and controlling factors. *Remote Sens. Environ., 213*, 34–48

Koralegedara, S.B., Lin, C.-Y., Sheng, Y.-F., & Kuo, C.-H. (2016). Estimation of anthropogenic heat emissions in urban Taiwan and their spatial patterns. *Environ. Pollut, 215*, 84–95

Kuang, Y., Zhao, C.S., Ma, N., Liu, H.J., Bian, Y.X., Tao, J.C., & Hu, M. (2016). Deliquescent phenomena of ambient aerosols on the North China Plain. *Geophys. Res. Lett., 43*, 8744–8750.

Lee, S.-H., & Kim, S.-T. (2015). Estimation of anthropogenic heat emission over South Korea using a statistical regression method. *Asia-Pacific J Atmos Sci,*, 1–10

Lee, S.-H., McKeen, S.A., & Sailor, D.J. (2014). A regression approach for estimation of anthropogenic heat flux based on a bottom–up air pollutant emission database. *Atmos. Environ., 95*, 629–633

Lee, S.H., Song, C.K., Baik, J.J., & Park, S.U. (2009). Estimation of anthropogenic heat emission in the Gyeong-In region of Korea. *Theor. Appl. Clim., 96*, 291–303

Li, Z., Fu, F., Ma, L., Ni, W. (2006). Energy flow chart of China based on energy balance sheet. *China Energy*, 28(9), 5e18 (In Chinese). doi:10.3969/j.issn.1003-2355.2006.09.002

Liang, L., Di, L., Zhang, L., Deng, M., Qin, Z., Zhao, S., & Lin, H. (2015). Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. *Remote Sens. of Environ., 165*, 123–134

Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., & Hong, Y. (2017). Classifying urban land use by integrating remote sensing and social media data. *Int J Geogr Inf Sci., 31*, 1675–1696

Liu, Y., Hu, C., Zhan, W., Sun, C., Murch, B., & Ma, L. (2018). Identifying industrial heat sources using time-series of the VIIRS Nightfire product with an object-oriented approach. *Remote Sens. Environ., 204*, 347–365

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., & Shi, L. (2015). Social sensing: a new approach to understanding our socioeconomic environments. *Ann Assoc Am Geogr, 105*, 512–530

Ma, Z., Shi, Z., Zhou, Y., Xu, J., Yu, W., & Yang, Y. (2017). A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed. *Remote Sens. Environ., 200*, 378–395.

McKenzie, G., Janowicz, K., Gao, S., Yang, J.-A., & Hu, Y. (2015). POI pulse: A multi-granular, semantic signature–based information observatory for the interactive visualization of big geosocial data. *Cartographica, 50*, 71–85

Morris, K.I., Chan, A., Morris, K.J.K., Ooi, M.C.G., Oozeer, M.Y., Abakr, Y.A., Nadzir, M.S.M., Mohammed, I.Y., & Al-Qrimli, H.F. (2017). Impact of urbanization level on the interactions of urban area, the urban climate, and human thermal comfort. *Appl. Geogr, 79*, 50–72

Oke, T.R. (1988). The urban energy balance. *Prog Phys Geog, 12*, 471–508

Oleson, K.W., Monaghan, A., Wilhelmi, O., Barlage, M., Brunsell, N., Feddema, J., Hu, L., & Steinhoff, D.F. (2015). Interactions between urbanization, heat stress, and climate change. *Clim. Change,*

129, 525–541

Park, C., Schade, G. W., Werner, N. D., Sailor, D. J., & Kim, C. H. (2016). Comparative estimates of anthropogenic heat emission in relation to surface energy balance of a subtropical urban neighborhood. *Atmos. Environ.*, *126*, 182–191.

Quah, A.K.L., & Roth, M. (2012). Diurnal and weekly variation of anthropogenic heat emissions in a tropical city, Singapore. *Atmos. Environ.*, *46*, 92-103

Quinlan, J.R. (1992). Learning with continuous classes. In, *5th Australian Joint Conference on Artificial Intelligence* (pp. 343–348): World Scientific

Quinlan, J.R. (1993). Combining instance-based and model-based learning. In, *Proceedings of the Tenth International Conference on Machine Learning* (pp. 236-243)

Sailor, D. J., Brooks, A., Hart, M., & Heiple, S. (2007). A bottom–up approach for estimating latent and sensible heat emissions from anthropogenic sources. In *Seventh Symposium on the Urban Environment*, San Diego, California (pp. 10–13).

Sailor, D.J. (2011). A review of methods for estimating anthropogenic heat and moisture emissions in the urban environment. *Int. J. Climatol.*, *31*, 189–199

Sailor, D.J., Georgescu, M., Milne, J.M., & Hart, M.A. (2015). Development of a national anthropogenic heating database with an extrapolation for international cities. *Atmos. Environ.*, *118*, 7–18

Sailor, D.J., & Lu, L. (2004). A top–down methodology for developing diurnal and seasonal anthropogenic heating profiles for urban areas. *Atmos. Environ.*, *38*, 2737–2748

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PloS One*, *10*(2).

Shin, D., Aliaga, D., Tunçer, B., Arisona, S.M., Kim, S., Zünd, D., & Schmitt, G. (2015). Urban sensing: using smartphones for transportation mode classification. *Comput., Environ. Urban Syst.*, *53*, 76–86

Sun, R., Wang, Y., & Chen, L. (2018). A distributed model for quantifying temporal-spatial patterns of anthropogenic heat based on energy consumption. *J. Clean. Prod.*, *170*, 601–609

Tong, H., Liu, H., Sang, J., & Hu, F. (2004). The impact of urban anthropogenic heat on Beijing heat environment. *Clim. Environ. Res.*, *9*, 13

Walton, J.T. (2008). Subpixel urban land cover estimation: comparing cubist, random forests and support vector regression. *Photogramm. Eng. Remote Sens.*, *74*, 1213–1222

Wang, J., Feng, J., Yan, Z., Hu, Y., & Jia, G. (2012). Nested high-resolution modeling of the impact of urbanization on regional climate in three vast urban agglomerations in China. *J. Geophys. Res.: Atmos.*, *117*, D21103

Wang, J., Yan, Z., Quan, X.-W., & Feng, J. (2017). Urban warming in the 2013 summer heat wave in eastern China. *Clim Dynam*, *48*, 3015–3033

Wang, M., Yan, X., Liu, J., & Zhang, X. (2013). The contribution of urbanization to recent extreme heat events and a potential mitigation strategy in the Beijing–Tianjin–Hebei metropolitan area. *Theor. Appl. Clim.*, *114*, 407–416

Wang, S., Hu, D., Chen, S., & Yu, C. (2019). A partition modeling for anthropogenic heat flux mapping in China . *Remote Sens.*, *11*, 1132

Wang, Y., Gu, Y., Dou, M., & Qiao, M. (2018). Using spatial semantics and interactions to identify urban functional regions. *ISPRS Int. J. Geo-Inf.*, *7*, 130

Weng, Q., Lu, D., & Liang, B. (2006). Urban surface biophysical descriptors and land surface temperature variations. *Photogramm. Eng. Remote Sens., 72*, 1275–1286

Xie, M., Liao, J., Wang, T., Zhu, K., Zhuang, B., Han, Y., Li, M., & Li, S. (2016a). Modeling of the anthropogenic heat flux and its effect on regional meteorology and air quality over the Yangtze River Delta region, China. *Atmos. Chem. Phys., 16*, 6071–6089

Xie, M., Zhu, K., Wang, T., Feng, W., Gao, D., Li, M., Li, S., Zhuang, B., Han, Y., Chen, P., & Liao, J. (2016b). Changes in regional meteorology induced by anthropogenic heat and their impacts on air quality in South China. *Atmos. Chem. Phys., 16*, 15011–15031

Xu, Y., Ho, H.C., Wong, M.S., Deng, C., Shi, Y., Chan, T.C., Knudby, A., (2018). Evaluation of machine learning techniques with multiple remote sensing datasets in estimating monthly concentrations of ground-level PM2.5. *Environ. Pollut,* 242, 1417–1426.

Yang, B., Yang, X., Leung, L.R., Zhong, S., Qian, Y., Zhao, C., Chen, F., Zhang, Y., & Qi, J. (2019). Modeling the impacts of urbanization on summer thermal comfort: the role of urban land use and anthropogenic heat. *J. Geophys. Res.: Atmos.*, 124, 6681–6697.

Yang, T., Gbaguidi, A., Zhang, W., Wang, X., Wang, Z., & Yan, P. (2018). Model-integration of anthropogenic heat for improving air quality forecasts over the Beijing megacity. *Air Qual. Res., 18*, 790–802

Yang, W., Chen, B., & Cui, X. (2014). High-resolution mapping of anthropogenic heat in China from 1992 to 2010. Int. J. Environ. Res. Public Health, *11*, 4066–4077

Yang, W., Luan, Y., Liu, X., Yu, X., Miao, L., & Cui, X. (2017). A new global anthropogenic heat estimation based on high-resolution nighttime light data. *Sci. Data., 4*, 170116

Yang, Y., Zhang, M., Li, Q., Chen, B., Gao, Z., Ning, G., Liu, C., Li, Y., & Luo, M. (2020). Modulations of surface thermal environment and agricultural activity on intraseasonal variations of summer diurnal temperature range in the Yangtze River Delta of China. *Sci Total Environ*, 736, 139445.

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., & Mai, K. (2017). Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *Int J Geogr Inf Sci., 31*, 825–848

Ye, T., Zhao, N., Yang, X., Ouyang, Z., Liu, X., Chen, Q., Hu, K., Yue, W., Qi, J., Li, Z., & Jia, P. (2019). Improved population mapping for China using remotely sensed and points-of-interest data within a random forests model. *Sci Total Environ, 658*, 936–946

Yoshida, D., Song, X., & Raghavan, V. (2010). Development of track log and point of interest management system using Free and Open Source Software. Appl. Geomatics, *2*, 123–135

Yu, M., Carmichael, G.R., Zhu, T., & Cheng, Y. (2014). Sensitivity of predicted pollutant levels to anthropogenic heat emissions in Beijing. *Atmos. Environ., 89*, 169–178

Zhang, L., Meng, Q., Sun, Z., & Sun, Y. (2017b). Spatial and Temporal Analysis of the Mitigating Effects of Industrial Relocation on the Surface Urban Heat Island over China. *ISPRS Int. J. Geo-Inf., 6*, 121

Zhang, P., Yuan, C., Sun, Q., Liu, A., You, S., Li, X., Zhang, Y., Jiao, X., Sun, D., Sun, M., Liu, M., & Lun, F. (2019). Satellite-Based Detection and Characterization of Industrial Heat Sources in China. *Environ Sci Technol, 53*, 11031–11042

Zhang, Y., Li, Q., Huang, H., Wu, W., Du, X., & Wang, H. (2017a). The combined use of remote sensing and social sensing data in fine-grained urban land use mapping: A case study in Beijing, China. *Remote Sens., 9*, 865

Zhao, X., Jiang, H., Wang, H., Zhao, J., Qiu, Q., Tapper, N., & Hua, L. (2013). Remotely sensed thermal pollution and its relationship with energy consumption and industry in a rapidly urbanizing Chinese city. *Energy Policy, 57*, 398–406

Zhou, D., Zhao, S., Zhang, L., Sun, G., & Liu, Y. (2015). The footprint of urban heat island effect in China. *Sci. Rep., 5*, 11160

Zhou, Y., Zhao, F., Wang, S., Liu, W., & Wang, L. (2018). A method for monitoring iron and steel factory economic activity based on satellites. *Sustainability, 10*, 1935

654 **Figure captions**

655

656 Figure 1. Flow diagram for producing the anthropogenic heat flux maps of China.

657

658 Figure 2. AHF maps for (a) industry (AHF_i), (b) buildings (AHF_b), (c)
659 transportation(AHF_t), (d) human metabolism (AHF_m), and (e) total AHF in 2010 (unit:
660 $W \cdot m^{-2}$).

661

662 Figure 3. Comparison of AHF from (a) industry, (b) buildings, (c) transportation, and
663 (d) human metabolism for four metropolitan agglomerations in China (unit: $W \cdot m^{-2}$).

664

665 Figure 4. Variable importance of the covariates in the Cubist model for the estimation
666 of AHF from (a) industry, (b) transportation, and (c) buildings

667

668

669

Highlights

1. An annual mean AHF map for China in 2010 at 1km spatial resolution is produced using a top-down approach.
2. Annual anthropogenic heat emission was estimated from four heat sources, based on energy consumption data.
3. POI and multisource remote sensing data were jointly used in Cubist models to estimate AHF.
4. POI data can refine AHF mapping from different heat sources within complex urban areas.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: