

1 Estimating Urban Sensible Heat Flux Using GOES 16 Satellite 2 Data

³ Gabriel Rios^{*1, 2} and Prathap Ramamurthy^{1, 2}

⁴ ¹Department of Mechanical Engineering, CUNY City College, New York, NY 10031, USA

⁵²NOAA CREST Center, CUNY City College, New York, NY 10031, USA

6 Abstract

A model for calculating sensible heat flux (Q_H) – a primary component of the urban surface energy budget – is presented here. Remote sensing data from the NASA/NOAA GOES-16 satellite and a high-resolution land cover dataset are used as inputs to estimate the spatiotemporal variability of Q_H in urban areas. The primary motivation for this model is to present a cost-effective approach to calculate Q_H independent of traditional flux observations and computational methods. The GOES-16 satellite data, which has a moderate spatial and high temporal resolution (2 km square at 5 minute intervals) enables the estimation of Q_H over highly heterogeneous urban areas. The model is constructed using an iterative algorithm that uses a surface layer turbulence parameterization to solve for Q_H as a function of the enterprise GOES-16 Land Surface Temperature product, an urban air temperature model, publicly-accessible ground observations, and the National Land Cover Database (NLCD). Model validation was performed over a year-long period spanning from 2019 June to 2020 May. Three ground flux stations within New York City located in areas with varying degrees of urbanization were used for model validation. Statistics from validation studies found an RMSE of 47.32 W m^{-2} , a mean bias error of 16.58 W m^{-2} , and a R^2 of 0.70. Validation results demonstrate that the algorithm shows good correlation with observed values, suggesting that satellite data can be used as an accessible and cost-effective option to estimate Q_H in urban areas. The satellite-based model outperformed conventional high-resolution urban numerical model.

Keywords: Sensible Heat Flux, Urban, Satellite Remote Sensing, GOES-16, Atmospheric Modeling, GOES-16, National Land Cover Database

**Corresponding author:* Gabriel Rios, Department of Mechanical Engineering, CUNY City College, New York, NY 10031, USA. E-mail: grios001@citymail.cuny.edu

³¹ 1 Background

Sensible heat flux (Q_H) is a key component of the Earth's surface energy balance, as it characterizes the surface-to-atmosphere transport of heat. In urban environments, anthropogenic modification of land cover reduces water retention capacity, increasing the roles of sensible heat and heat storage (Q_S) in the urban surface energy budget. Q_H in cities impacts the urban heat island dynamics, hence, it has significant implications on weather prediction and forecasting, air pollution, and building energy use. (Imran et al. (2018), Schumacher et al. (2019), Vautard et al. (2007)).

38 Q_H is driven by a number of factors - particularly the temperature difference between the land surface
39 temperature (LST) and the air temperature (T_{air}) in the lowest levels of the boundary layer. The LST
40 has been shown to be higher in urban areas than surrounding suburban/rural areas ([Price \(1979\)](#)), which
41 is driven by the heterogeneity in urban land cover and its constituent materials. The increased LST can
42 both increase T_{air} and the temperature difference between the two, resulting in an increased Q_H relative to
43 surrounding areas ([Kato & Yamaguchi \(2005\)](#)).

44 A challenge in understanding the relationship between land cover, LST, T_{air} and Q_H is presented by the
45 techniques used for measurement and estimation of Q_H . This challenge is brought about by a number of
46 factors, including (but not limited to):

- 47 • Computationally-expensive numerical models for estimation purposes ([Best \(2005\)](#), [Zhang et al. \(2015\)](#)),
48 and
- 49 • The lack of well-established measurement networks in rural and urban areas ([Chrysoulakis et al. \(2018\)](#),
50 [Voogt & Oke \(2003\)](#))

51 Numerical models are powerful estimation tools that allow for the understanding of atmospheric processes
52 at much greater spatial extents than possible by measurement and observation alone. However, these models
53 can often feature significant inaccuracies in areas with high spatial heterogeneity, such as urban areas, due to
54 low grid domain resolutions relative to the size and spacing of elements in heterogeneous environments (e.g.
55 buildings, roads, scattered green space and vegetative cover) ([Chen et al. \(2011\)](#), [Hong & Dudhia \(2012\)](#),
56 [Leroyer et al. \(2014\)](#)). Accordingly, it is often that model accuracy can only be improved upon by significantly
57 increasing model resolution to resolve these spatial issues, which risks high time and resource consumption.
58 Meanwhile, measurement networks are vital since observational data is an essential source of validation
59 data for numerical models to ensure their performance. However, accurate measurement of parameters
60 such as Q_H is challenged by the lack of measurement networks with sufficient spatial resolution that can
61 serve as databases for validation efforts. Moreover, this challenge is exacerbated in urban areas due to the
62 aforementioned land cover heterogeneity, which is critical in determining Q_H in localized areas ([Feddeema
et al. \(2005\)](#), [Wang et al. \(2016\)](#)). To address this, remote sensing technologies have been increasingly used
64 to devise estimation methods for Q_H .

65 Several studies in the reviewed literature have estimated heat fluxes using remote sensing methods in
66 rural areas using a variety of methods ([Cammalleri et al. \(2012\)](#), [Kim & Kwon \(2019\)](#), [Miglietta et al. \(2009\)](#),
67 [Mkhwanazi et al. \(2012\)](#), [Ortega-Farías et al. \(2016\)](#)). [Miglietta et al. \(2009\)](#) describes an estimation method
68 using Meteosat land surface temperature and radiation products, as well as aircraft-mounted sensors, to
69 evaluate fluxes over forested areas and cropland between May and June 2005. In [Cammalleri et al. \(2012\)](#),
70 aircraft-mounted multispectral and thermal cameras were used in conjunction with meteorological data to
71 estimate Q_H over 7 days within a 4 month period, with a study area covered by cropland, fallow soil, and
72 bare soil. [Mkhwanazi et al. \(2012\)](#) used Landsat 5 imagery with a bulk parameterization method to evaluate
73 fluxes over an alfalfa field in rural Colorado. [Kim & Kwon \(2019\)](#) and [Ortega-Farías et al. \(2016\)](#) showed
74 promising results using unmanned aerial vehicles (UAVs) to estimate Q_H over a variety of land cover types
75 in rural areas throughout a range of synoptic meteorological conditions, with good agreement between UAV-
76 based estimation results and instrument-based surface observations. These studies all demonstrate great
77 potential for using remote sensing for estimation of surface fluxes, although their temporal frequency and
78 focus on homogeneous land cover types hinders their applicability to urban areas.

79 Fewer studies have been performed to estimate Q_H using remote sensing methods in urban areas, which
80 feature far greater land cover heterogeneity ([Feigenwinter et al. \(2018\)](#), [Liu et al. \(2012\)](#), [Voogt & Grimmond](#)

81 (2000), Xu et al. (2008)). Two studies (Voogt & Grimmond (2000), Xu et al. (2008)) used helicopter-
82 mounted instruments to collect observational data over cities with the goal of estimating Q_H and associated
83 parameters. Voogt & Grimmond (2000) implemented a method for estimating Q_H over a 400 x 300 m
84 sector of Vancouver over 2 days using a helicopter-mounted thermal scanner for surface temperature data
85 collection, using the aerodynamic resistance method for estimation of Q_H . Xu et al. (2008) showed that
86 remote sensing is a viable way to determine the variation of Q_H in urban areas by using an airborne
87 spectrometer to analyze a section of Shanghai to determine land cover information, surface temperature,
88 and other parameters relevant to the calculation of Q_H . Although these methods were able to image urban
89 areas at ultrahigh spatial resolutions, the lack of spatiotemporal variability due to the study areas and image
90 frequency, as well as the expenses associated with the study, prevent them from being a practical method for
91 estimating Q_H for larger areas over extended periods of time. A more recent remote sensing approach that
92 addresses these issues is the use of satellite data over urban areas, as presented in Feigenwinter et al. (2018)
93 and Liu et al. (2012). In Liu et al. (2012), ASTER imagery was used as input to a model to estimate surface
94 fluxes over a 25 km^2 area, encompassing a variety of land cover types that range from highly-developed
95 urban areas to open green space to crop fields. Although study results yielded some correlation with related
96 atmospheric parameters for similar settings in the literature, no surface observation data was used to further
97 validate findings from the study. Additionally, the study was performed for a single point in time, preventing
98 any temporal variability analysis from being performed. In a study by Feigenwinter et al. (2018), Landsat 8
99 and TIRS data was used in conjunction with land cover data to employ the aerodynamic resistance method
100 to estimate sensible and latent heat fluxes in and around Basel, Switzerland over a wide range of land cover
101 types at a very high spatial resolution (100 m). This study presents a comprehensive approach to evaluating
102 spatial variability of fluxes in a heterogeneous study area as well as a relatively robust validation procedure
103 due to the high density of flux towers in an urban setting. Results show generally good agreement at all
104 validation locations, although the temporal frequency of Landsat and TIRS satellite imagery highly limits
105 this method to one estimation every 8 days, at minimum.

106 In this study, a method for estimating Q_H using a combination of open-access remote sensing and ground
107 observational data in a dedicated, cost-effective satellite-based model is introduced. The objective of this
108 method is to use satellite data to provide a large spatial and temporal domain over which Q_H can be
109 accurately estimated. The model uses satellite data from the NOAA/NASA Geostationary Operational En-
110 vironmental Satellite (GOES-16), ground observational data from NWS/FAA/DOD Automated Surface Ob-
111 serving Systems (ASOS) stations, and land cover data from the MRLC 2016 National Land Cover Database
112 (NLCD) to estimate Q_H . The primary advantage to using the GOES-16 satellite for the estimation of Q_H
113 is the spatial extent and high temporal resolution of its collected data. Although GOES-16 satellite data
114 features some limitations such as inability to reliably estimate during periods with significant sky cover and
115 a moderate spatial resolution of 2 km, the benefits provided by remote sensing data for Q_H estimation allow
116 for the limitations of previous studies with similar objectives to be addressed and mitigated.

117 The primary objectives of this paper are:

- 118 • To develop a satellite-based model to estimate the Q_H of urban environments at high temporal and
119 moderate spatial resolutions.
- 120 • To validate and compare the satellite-based estimates of Q_H with ground-based observations, as well as
121 with Q_H derived from high-resolution urban climate models, both temporally and spatially for multiple
122 seasons.

123 This paper will first discuss the theoretical background for the satellite model, including the use of Monin-

124 Obukhov similarity theory ([Monin & Obukhov 1954](#)) and the method for estimation of element roughness
125 heights in urban areas. Next, the paper reviews the use of GOES-16 satellite data and an associated urban
126 air temperature model ([Hrisko et al. 2020](#)) as inputs in the model, as well as how ground stations were
127 used for model inputs and validation. Subsequently, the model results over the year-long study period are
128 presented, along with validation data accompanied by a statistical evaluation of model performance against
129 ground stations. Finally, there is a discussion regarding the performance of the model, potential sources
130 of error within the model and the validation process, as well as application potential and future work to
131 improve the methods presented here.

132 **2 Methodology and data**

133 **2.1 Study area**

134 This study focuses on New York City (see Figure 1), which is the largest city in the United States by popu-
135 lation, with approximately 8.3 million people as of 2019 ([US Census Bureau \(2019\)](#)) and is among the most
136 densely-populated cities in the United States. The city, is composed of 5 boroughs:, the Bronx, Brooklyn,
137 Manhattan, Queens, and Staten Island. The Bronx is made up largely of low- to mid-rise residential and
138 commercial buildings, with decreasing building density and height towards the northern end of the borough.
139 Brooklyn is largely composed of low- to mid-rise buildings, with a concentration of high-rise buildings on
140 the East River, while the southern and eastern portions feature larger proportions of lower-density suburban
141 residential areas. Manhattan is primarily composed of residential and commercial buildings, with mid- to
142 high-rise buildings spanning the entirety of the borough (with the exception of Central Park, which is a
143 mixture of open fields, open water, and deciduous & evergreen forests). Queens is similar in composition to
144 Brooklyn, with the exception of larger spans of lower-density development towards the eastern half of the
145 borough. Staten Island features significantly lower building densities and heights, with expansive wetland
146 and grassy areas on its western edges and a large forested area in the central area of the borough. The com-
147 plex urban landscape, coupled with an array of urban flux towers and weather observation stations within
148 the city, make the city an ideal candidate for implementing and validating the urban-focused Q_H model.

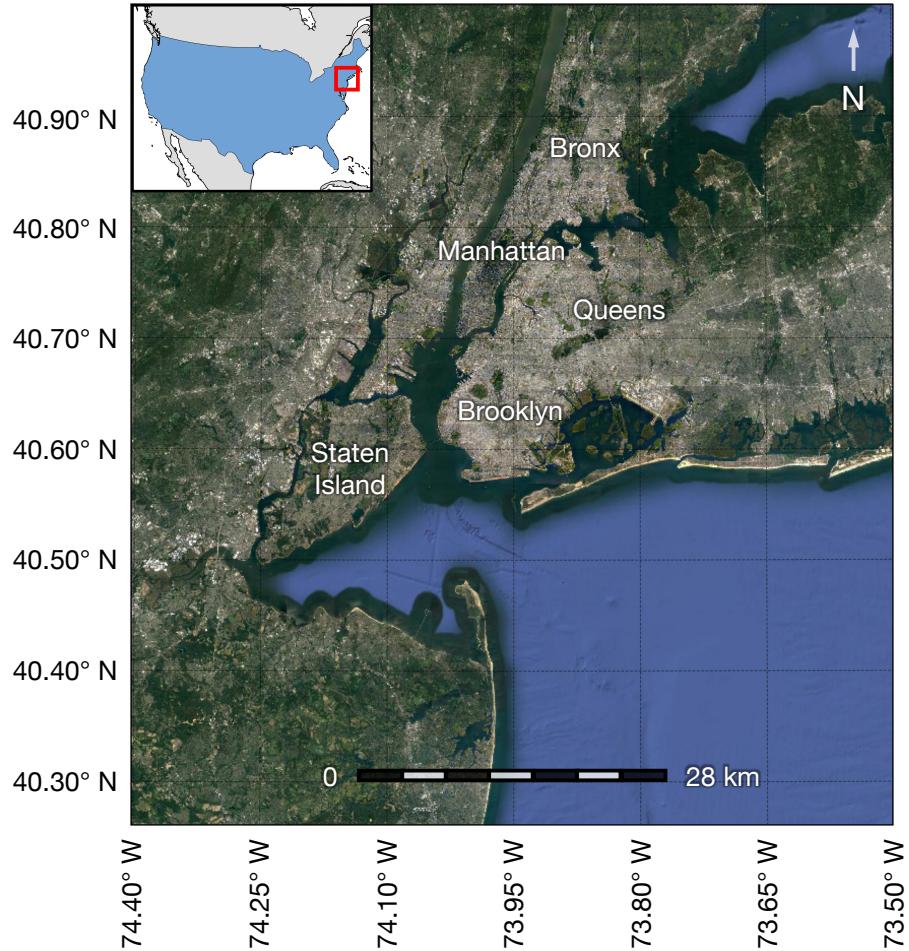


Figure 1: Satellite view of the New York City metropolitan area. New York City, which is composed of 5 boroughs (labeled), is the most-heavily urbanized portion of the metropolitan area, while lower density suburbs and woodlands compose the outer portions of the metropolitan area.

149 2.2 Model overview

150 Q_H and associated parameters are estimated using an iterative algorithm using bulk turbulence parameteriza-
 151 tions based on scaling arguments presented by Monin-Obukhov similarity theory. A flowchart of the model
 152 structure is shown in Figure 2. The model operates with a parallel observational and numerical approach;
 153 ground-based observational data is used for validation purposes, as well as for inputs to the iterative algo-
 154 rithm (specifically, wind speed, u and air pressure, p), while the numerical model receives inputs from the
 155 GOES-16 satellite as well as ancillary datasets (land cover and geographical information). The numerical
 156 model then matches inputs to specified locations, such as the described study area, before using an iterative
 157 algorithm to solve for Q_H and associated parameters.

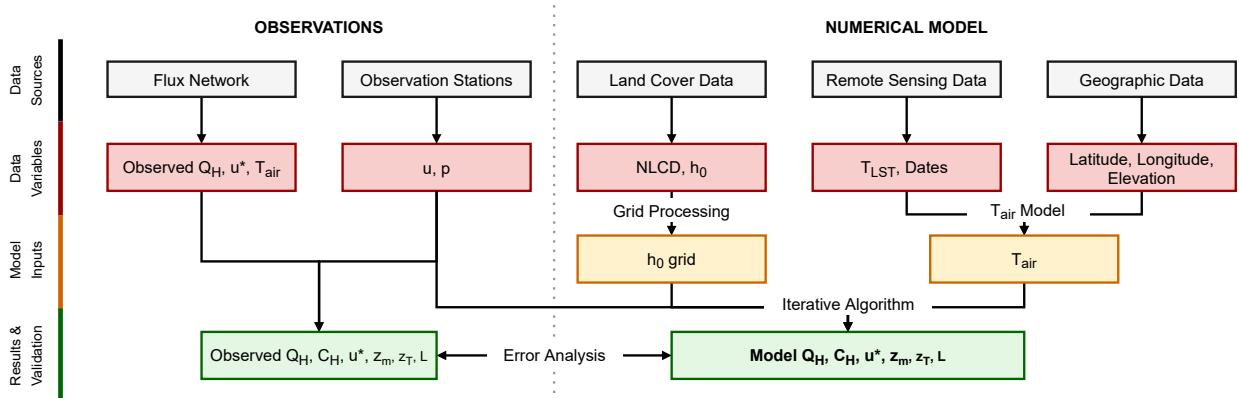


Figure 2: Process flowchart for the sensible heat flux model. Observational data was used for validation of the satellite model as well as inputs to the iterative algorithm. The numerical model used remotely-sensed data from the GOES-16 satellite, as well as ancillary datasets for land cover and geographic data. Error analysis was performed by comparing observational data and model results.

158 **2.2.1 Sensible heat flux iterative algorithm**

159 This section details the variables, equations, and assumptions that constitute the algorithm used to estimate
 160 Q_H . The iterative algorithm in the numerical model is dependent on the convergence of Q_H , which in turn is
 161 dependent on the Obukhov length (L), as is the case in other algorithms found in the literature (Grimmond
 162 & Cleugh (1994), Launiainen & Vihma (1990),). An assumption of a convectively-stable atmosphere ($L \rightarrow$
 163 ∞) defines initial conditions for the model. Momentum and thermal stability parameters, ψ_m and ψ_h ,
 164 are approximately 0 at this initial condition. The following static and dynamic variables - momentum
 165 and thermal roughness heights z_m and z_T , the bulk heat transfer coefficient C_H , the friction velocity u^* ,
 166 the Obukhov length L , and ultimately, Q_H are calculated by iteration, similar to land surface models.
 167 Convergence is defined by a <1 % change in Q_H between iterations.

168 Q_H is directly calculated using Equation 1 (Pond et al. (1974)):

$$Q_H = \rho c_p C_H u (\theta_0 - \theta_r) \quad (1)$$

169 In Equation 1, ρ is air density calculated as a function of air pressure (p) and air temperature at the
 170 reference height of 2 m above ground level (AGL) (T_{air}), c_p is the average specific heat of air (1006 J
 171 $kg^{-1} K^{-1}$) across the range of air temperatures and pressures observed, C_H is a bulk heat transfer coefficient,
 172 u is the observed wind speed at a height of 10 m AGL, and θ_0 and θ_r are potential temperatures at the
 173 surface and at 2 m AGL, respectively. Both θ_0 and θ_r are derived from remotely-sensed data - θ_0 is derived
 174 from remotely-sensed land surface temperature (T_{LST}) and θ_r is derived from a model based on T_{LST} and
 175 several other remotely-sensed parameters (Hrisko et al. (2020)). See Section 2.3 for a detailed discussion
 176 regarding these derived parameters.

177 C_H is calculated using Equation 2 (Monin & Obukhov (1954)):

$$C_H = \frac{\kappa^2}{[\ln \frac{z_r}{z_m} - \psi_m \zeta][\ln \frac{z_r}{z_T} - \psi_h \zeta]} \quad (2)$$

178 In Equation 2, κ is the von Karman constant (assumed to be 0.40), z_r is the reference height of mea-
 179 surement, z_m is the momentum roughness height, z_T is the thermal roughness height, ψ_m and ψ_h are the
 180 momentum and thermal stability parameters, respectively (Businger et al. (1971), Dyer (1974)), and ζ is an
 181 atmospheric stability parameter, defined as $\zeta = \frac{z_r}{L}$.

182 The momentum and thermal roughness heights, z_m and z_T , are calculated using the Raupach [Equation
 183 3] and Zilitinkevich [Equation 5] methods, respectively. The Raupach method (Raupach (1994)) for defining
 184 the momentum roughness height has been found useful in areas with heterogeneous land cover, as it can be
 185 calculated as a function of localized parameters and atmospheric conditions, specifically element roughness
 186 height h_0 and local friction velocity u^* (Voogt & Grimmond 2000). The methodology for the estimation of
 187 h_0 is discussed in detail in 2.2.2. The Zilitinkevich method has been shown to be an effective approximation
 188 method for z_T in areas with tall canopies, such as those present in urban areas, while enabling z_T to be
 189 calculated as a function of local parameters (Chen & Zhang (2009), Zilitinkevich (1995)), as described in Li
 190 & Bou-Zeid (2014).

$$z_m = h_0 \left(1 - \frac{z_d}{h_0}\right) \exp\left[-\kappa \frac{u}{u^*} + 0.193\right] \quad (3)$$

191 where:

$$z_d = \exp[0.9793 * \ln(h_0) - 0.1536] \quad (4)$$

$$z_T = z_m \exp[-\kappa C_{zil} \sqrt{Re_t}] \quad (5)$$

192 where:

$$C_{zil} = 10^{-0.40*h_0} \quad (6)$$

193

$$Re_t = \frac{z_m u^*}{\nu} \quad (7)$$

194 The friction velocity u^* is expressed by Equation 5 ([Monin & Obukhov \(1954\)](#)):

$$u^* = \frac{\kappa u}{\ln \frac{z}{z_m} - \psi_m \zeta} \quad (8)$$

195 The Obukhov length L is expressed by Equation 6 ([Monin & Obukhov \(1954\)](#)):

$$L = \frac{-\rho c_p (u^*)^3 (\theta_0 + \theta_r)}{2\kappa g Q_H} \quad (9)$$

196 The iterative model typically converged within 5 iterations, with convergence having been somewhat
197 dependent on atmospheric stability ζ - the more unstable the atmosphere, the more difficulty the model had
198 in converging.

199 2.2.2 Roughness height estimation

200 Element roughness height is a critical parameter for estimating Q_H , as is evidenced by Equations 3, 4, and
201 6. The element roughness height (h_0) describes the height of objects AGL such as buildings or trees. The
202 element roughness heights are calculated using a weighted average consisting of land cover parameters from
203 the 2016 National Land Cover Database (NLCD) ([Yang et al. 2018](#)) and element roughness height estimates
204 from values specific to urban areas from the Weather Research Forecasting (WRF) model ([Chen et al. \(2011\)](#),
205 [Skamarock et al. \(2008\)](#),).

206 The NLCD data features 20 land cover classes, each with different element roughness heights. The
207 NLCD data is packaged in a 30 x 30 m grid spanning the continental United States (CONUS) and Alaska.
208 To match the 2 x 2 km gridded data presented by the GOES-16 LST product, the NLCD data was upscaled
209 accordingly. Each NLCD grid element, or pixel, is constituted of an array of values ranging from 0 to 1, with
210 each value corresponding to the fraction of pixel that is determined by each land cover class. See Figure 3
211 for the NLCD land cover map of the study area.

212 Element roughness heights used for the WRF model are likewise used for this model for the correspond-
213 ing NLCD classes. Specific h_0 values are used for urban areas, defined as “Developed, Low Intensity”,
214 “Developed, Medium Intensity”, and “Developed, High Intensity” by the NLCD classification system. The
215 corresponding WRF classes are “Low-Density Residential”, “High-Density Residential”, and “Commercial”,
216 respectively. The element roughness heights defined by the WRF for “Low-Density Residential”, “High-
217 Density Residential”, and “Commercial” areas are 5.00, 7.50, and 10.00 m, respectively, as outlined in the
218 description of an urban modeling system for the Weather Research and Forecasting (WRF) model ([Chen](#)
219 [et al. \(2011\)](#)). These values were used in the weighted-averaging scheme to obtain approximate element
220 roughness heights for the model.

221 To estimate the element roughness height corresponding to each 2 x 2 km pixel, an inner product was
222 taken using the land cover class element roughness heights and the land cover class percentages. The results
223 of this estimation method are shown in Figure 4.

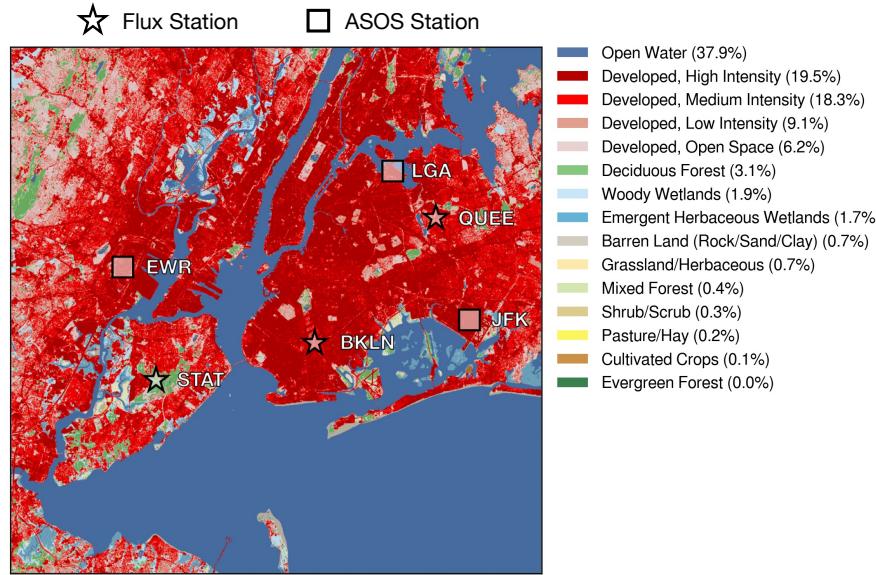


Figure 3: Land cover map of the New York City metropolitan area, per the 2016 National Land Cover Database (Yang et al. (2018)). The legend shows land cover types and the percentage of the study area occupied by each land cover type. Land cover data is shown at a 30 m resolution. Note that flux observation towers and ground weather (ASOS) stations are labeled accordingly.

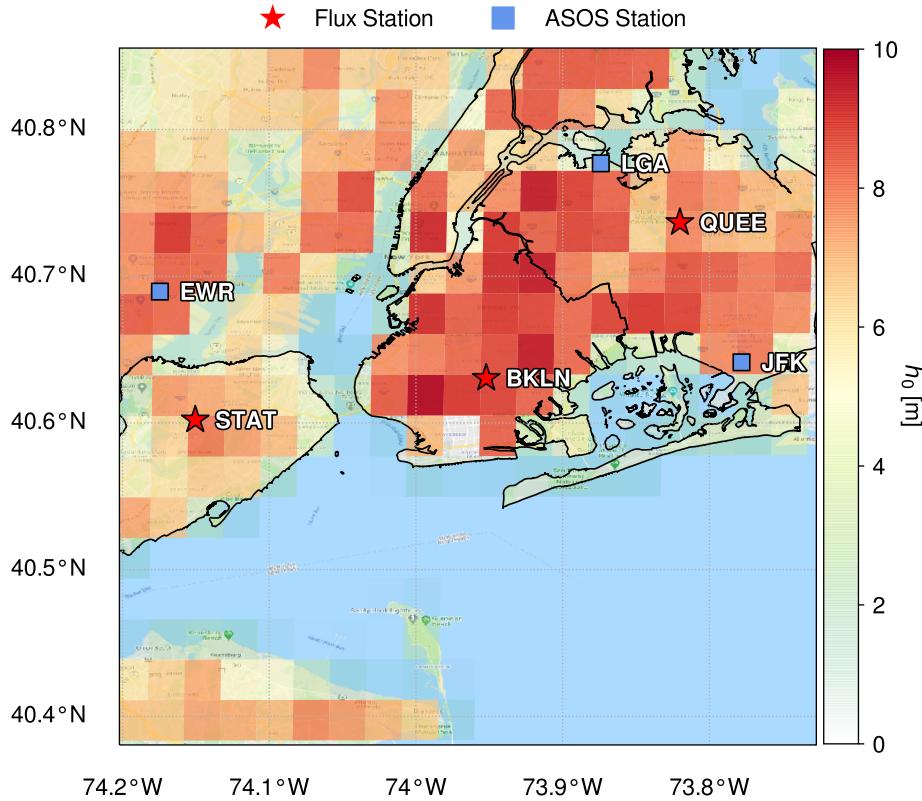


Figure 4: Gridded map of element roughness heights across the New York City metropolitan area. Note that flux observation towers and ground weather (ASOS) stations are labeled accordingly.

2.3 GOES-R land surface temperature (LST) product

The Geostationary Operational Environmental Satellites (GOES-R), GOES-16 and GOES-17, are operated by the National Aeronautic and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA). The GOES-16 satellite, which is used for this study, is located over the western Atlantic Ocean and focuses on observation of North and South America.

A number of products derived from satellite radiance data are offered by the satellite, including a Land Surface Temperature (LST) product, from which T_{LST} (and through derivation, θ_0) is obtained. It is available for public use at a moderate spatial resolution of 2 x 2 km and a high temporal resolution of 5 minutes (NOAA/NESDIS/STAR (2016)). The LST is calculated using GOES-16 infrared bands 14 and

15. This product features a desirable balance of spatiotemporal resolution and high accuracy (<2.50 K) (Valenti (2017)), making it a critical input to the model. The LST product is available in a gridded netCDF (.nc) format, with data corresponding to latitude and longitude mapped over the spatial extent of satellite observations. The data is filtered based on image quality, which is largely dependent on sky conditions (i.e. cloud cover). Therefore, dates within the study timeframe with clear skies or few clouds (<25% sky cover, per METAR (World Meteorological Organization 2008) were selected to ensure high-quality LST data as input to the model. The data used for the model was limited to a 0.50 degree extent encompassing the most heavily-urbanized portion of the New York City metropolitan area, extending from approximately (40.8805 N, 74.2021 W) to (40.3805 N, 73.7021 W), spanning approximately 190 km². On a 2 x 2 grid, this represents approximately 380 pixels over which data was obtained for the metropolitan area.

Another major component of the model is an urban air temperature model that takes GOES-16 LST product data as an input and uses a diurnal regressive algorithm to calculate air temperature at a height of 2 m AGL (Hrisko et al. 2020), from which T_{air} (and through derivation, θ_r) is obtained. The model has been shown to calculate air temperatures in areas featuring a range of land cover classes with high accuracy, specifically in urban areas (RMSE of 2.60 K relative to ground station observations). Inputs to the model are LST, elevation, NLCD land cover class, and coordinates. The model output is a gridded dataset with temperature values. For reference, the data is produced on a 2 x 2 km grid to match the gridded data format of the GOES-16 LST product.

2.4 Ground station observation data

Model inputs for air pressure (p) and wind speed (u_r) were obtained from various Automated Surface Observing System (ASOS) stations in the New York City metropolitan area. The ASOS network, which is operated by NOAA, features over 900 sites in the United States, allowing for weather conditions at many locations within the continental United States to be adequately represented by ASOS data.

Each ASOS station collects a wealth of information regarding weather conditions most relevant for aviation purposes, including air temperature, dew point temperature, air pressure, wind speed and direction, and sky cover. Each station generally records data at a frequency of 5 minutes, providing reasonable spatial and excellent temporal frequencies for model data input. Four stations are located within the spatial domain evaluated in this study (see Figure 4 for reference): John F. Kennedy International Airport (JFK) (40.6413° N, 73.7781° W), LaGuardia Airport (LGA) (40.7769° N, 73.8740° W), Newark Liberty International Airport (EWR) (40.6895° N, 74.1745° W), Central Park (40.7790° N, 73.9693° W). The ASOS stations closest to each observation site are selected for data collection. Specifically, these ASOS stations are JFK (corresponding to Brooklyn), LGA (Queens), and EWR (Staten Island).

The model was validated using the New York State (NYS) Mesonet observation network (Mesonet 2020). The network features 17 flux stations throughout the state of New York, with 3 stations located within New York City - one each in the boroughs of Brooklyn (BKLN) (40.6318° N, 73.9537° W), Queens (QUEE) (40.7343° N, 73.8158° W), and Staten Island (STAT) (40.6040° N, 74.1485° W). The flux network stations record parameters relevant to the surface energy budget, including net radiation R_N , surface latent heat flux Q_L , and surface sensible heat flux Q_H . Each flux station is equipped with a net radiometer (manufactured by Kipp & Zonen CNR4), ground heat flux plates (Hukseflux), and a closed-path eddy covariance system (CPEC200, Campbell Scientific, Inc) consisting of a sonic anemometer and gas analyzer. The net radiometer and eddy covariance system are installed at 33.20 m at the Brooklyn station, 54.60 m at the Queens station, and 33.10 m at the Staten Island station (all heights above ground level). The ground plates are installed

275 at a depth of 6 cm below ground. Station flux measurements are reported every 30 minutes. The eddy
276 covariance system was used to measure Q_H for the duration of the validation period.

277 These stations were used for validation because of their high temporal sampling frequency and their
278 locations in areas of the city with surrounding land cover types representative of their respective boroughs,
279 rendering them useful for validating a model intended to provide output with fine spatial resolution. The
280 Brooklyn station is located in a neighborhood with low- and mid-rise residential and commercial buildings
281 with little open vegetated space (NLCD land cover classification codes "22 - Developed, Low Intensity",
282 "23 - Developed, Medium Intensity", "24 - Developed, High Intensity"). The Queens station is similar to
283 the Brooklyn location, with the exception of a large cemetery directly to the west that serves as an open
284 vegetated space (NLCD land cover classification codes "22 - Developed, Low Intensity", "23 - Developed,
285 Medium Intensity", "24 - Developed, High Intensity"). The Staten Island station is located on a university
286 campus enveloped by deciduous forest on 3 sides and low-density residential on the 4th (NLCD land cover
287 classification codes "22 - Developed, Low Intensity", "23 - Developed, Medium Intensity", "24 - Developed,
288 High Intensity", "41 - Deciduous Forest"). See Figure 3 for a map showing land cover classifications for New
289 York City with flux station locations annotated. Each station is matched by coordinates to a corresponding
290 GOES-16 satellite data pixel such that the pixel envelopes the station and its immediate surrounding area.
291 The limitations of the siting of the validation stations and the station-satellite matching method are discussed
292 later in the paper. NYS Mesonet data used for validation spans a full calendar year, from 1 June 2019 to 31
293 May 2020. All stations were operational and recorded data during the extent of the validation time period.

294 2.5 Urbanized Weather Research and Forecasting (uWRF) Model

295 The Weather Research and Forecasting (WRF version 3.9.9.1) model ([Skamarock et al. \(2008\)](#)) with an
296 urbanization option (uWRF) is used in this study to provide a comparison with the dedicated Q_H model on
297 numerical model estimation performance relative to ground observations. The urbanization option features
298 parameterizations specific to urban areas for better representations of boundary layer processes in cities
299 ([Gutiérrez, Martilli, Santiago & González \(2015\)](#), [Gutiérrez, González, Martilli & Bornstein \(2015\)](#)). This
300 configuration of the WRF model has been used in numerous previous studies to study atmospheric processes
301 in urban areas ([Chen et al. \(2011\)](#), [Gamarro et al. \(2019\)](#), [Gutiérrez, González, Martilli, Bornstein & Arend](#)
302 ([\(2015\)](#), [Hrisko et al. \(2021\)](#), [Ortiz et al. \(2017\)](#)).

303 The uWRF was initialized with the North American Mesoscale (NAM) forecast at 12-km resolution. The
304 uWRF was run on multi-domain mode centered over New York City with the following domain resolutions:
305 9 km (120x120 grid), 3 km (121x121), and 1 km (85x82) with 51 vertical levels; the first level was located
306 at a height of 10 m with 30 additional levels below 1000 m. The uWRF was run for 4 days, chosen to be
307 roughly characteristic of each season: 24 October 2019 (autumn), 23 December 2019 (winter), 20 January
308 2020 (winter), 12 May 2020 (spring). The model was run with the Dudhia scheme ([Dudhia \(1989\)](#)) for
309 shortwave radiation and the Rapid Radiative Transfer Model for longwave radiation ([Mlawer et al. \(1997\)](#)).
310 For the planetary boundary layer (PBL) parameterization, the Mellor-Yamada-Janjić scheme ([Janjić \(1994\)](#))
311 was used while the land surface fluxes for non-urban cover were parameterized using the NOAH scheme ([Niu](#)
312 [et al. \(2011\)](#)). A cumulus parameterization was used for the coarser outer grid domains. For urban fluxes,
313 the coupled Building Environment Parameterization and Building Energy Model (BEP-BEM) was used
314 ([Salamanca & Martilli \(2010\)](#)). Land cover in New York City was represented by the Primary Land Use Tax
315 Lot Output (PLUTO) database.

316 **3 Results**

317 **3.1 Model performance evaluation**

318 Four statistical measures were used to determine model performance relative to ground stations: root-mean-
319 square error (RMSE), mean bias error (MBE), the Nash-Sutcliffe model coefficient (NSC), and the coefficient
320 of determination (R^2). NSC is a commonly-used statistic for model validation as a method to determine the
321 accuracy of model predictions relative to observed data that may be highly variable due to perturbations
322 (such as wind) (Legates & McCabe Jr (1999), Nash & Sutcliffe (1970),). An NSC value greater than 0.50 is
323 considered as denoting satisfactory model performance (Moriasi et al. (2007)).

324 Each measure is defined as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_{H,i,m} - Q_{H,i,o})^2} \quad (10)$$

$$MBE = \frac{1}{N} \sum_{i=1}^N (Q_{H,i,m} - Q_{H,i,o}) \quad (11)$$

$$NSC = 1 - \frac{\sum_{i=1}^N (Q_{H,i,m} - Q_{H,i,o})^2}{\sum_{i=1}^N (Q_{H,i,o} - \bar{Q}_{H,o})^2} \quad (12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (Q_{H,i,m} - \bar{Q}_{H,i,o})^2}{\sum_{i=1}^N N(Q_{H,i,o} - \bar{Q}_{H,o})^2} \quad (13)$$

328 **3.2 Model performance against ground stations**

329 The study period for the model spanned from 1 June 2019 (day of year 152) to 31 May 2020 (day of year 152).
330 Approximately 44 days over the course of the study period were selected for model validation. The selection
331 criteria included sky cover classified as “CLR” (clear sky) or “FEW” (few clouds) at each ASOS observation
332 station continuously over a 24-hour period and operational flux network status. For validation purposes,
333 model runs were initially performed at the latitude and longitude corresponding to each flux station. The
334 corresponding GOES-16 grid location, or pixel, was used for the LST and T_{air} . The closest ASOS station
335 was used to provide inputs of p and u (the distance between the study location and the corresponding ASOS
336 station is a potential source of error that is discussed further). In total, 3 pixels were analyzed for validation
337 purposes at hourly intervals over the selected days, resulting in a total of approximately 3,200 data points.

338 **3.2.1 Overall results**

339 Over the timeframe studied, the model featured a RMSE of 47.32 Wm^{-2} , a bias of 16.58 Wm^{-2} , an NSC
340 value of 0.54, and a R^2 value of 0.70. The statistical results of the study period, as defined in Equations
341 10-13, are shown in Table 1 decomposed by location. The performance statistics suggest the model displayed
342 reasonable agreement with the ground observations and performed satisfactorily, per the definition provided
343 for the Nash-Sutcliffe coefficient by Moriasi et al. (2007). Model performance as a function of spatial and
344 temporal variability will be discussed in this section. Temporal variability will be discussed on two distinct
345 timescales, seasonal and daily, to improve understanding of model behavior and differences driven by changes
346 in time on large and small temporal scales. Nonetheless, the model showed considerable error from the ground
347 observation data, and potential causes will be addressed further in the paper.

Table 1: Model performance statistics against ground station data over the study period (2019 June - 2020 May). RMSE and MBE have units of $W m^{-2}$.

Station	Points	RMSE	MBE	NSC	R^2
BKLN	1149	59.26	29.99	0.26	0.70
QUEE	1165	43.52	15.97	0.63	0.75
STAT	1160	36.21	3.79	0.70	0.73

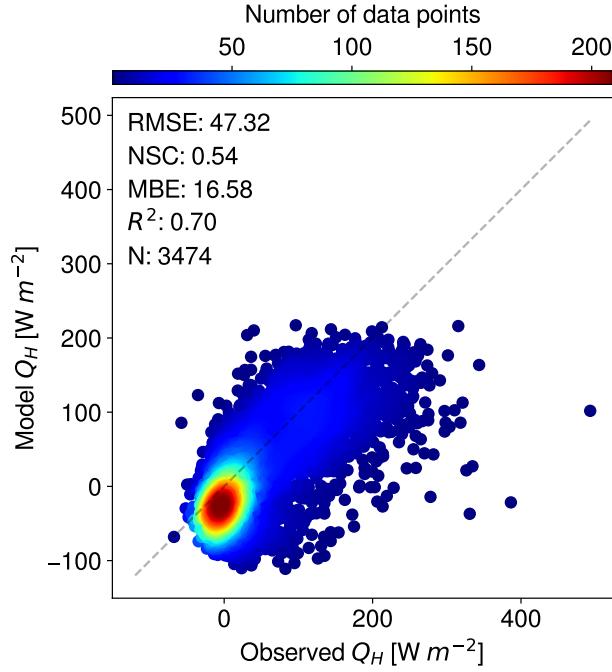


Figure 5: Comparison of model and observed Q_H from 1 June 2019 to 31 May 2020.

3.2.2 Diurnal variability

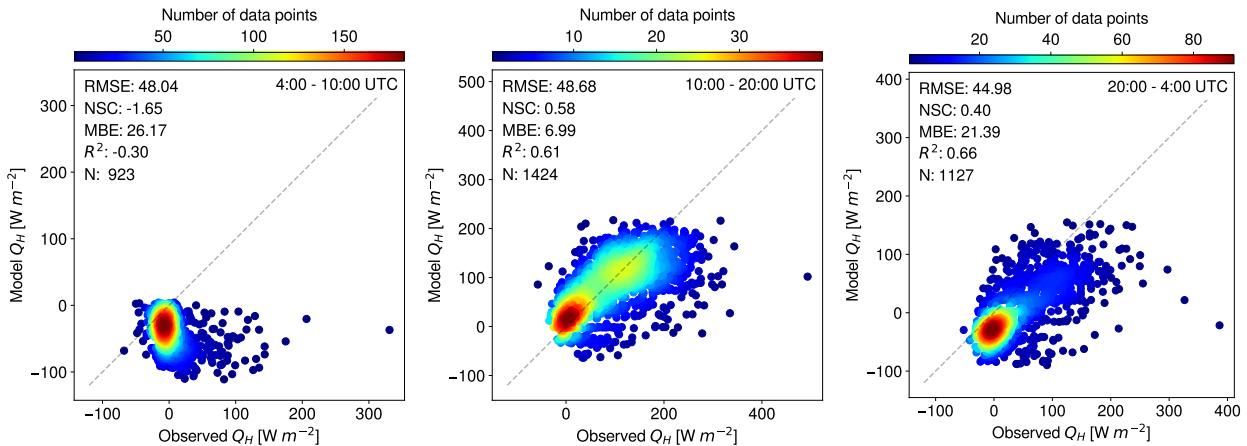


Figure 6: Comparison of model and observed Q_H divided into distinct periods of the day. (from left to right) early morning (pre-sunrise), daytime, evening and late night (post-sunset).

Performance of the model against observed data over distinct portions of the day is displayed in Figure 6. The model performed fairly well during the daytime (10:00 - 20:00 UTC), with good agreement between the model and the observations across all stations ($RMSE = 48.68 W m^{-2}$, $MBE = 6.99 W m^{-2}$, $NSC = 0.58$, $R^2 = 0.61$). However, the model did not perform as well in the early nighttime hours due to an increased bias error (20:00 - 04:00 UTC) ($RMSE = 44.98 W m^{-2}$, $MBE = 21.39 W m^{-2}$, $NSC = 0.40$, $R^2 = 0.66$ across all stations) and poorly during the pre-dawn hours due to a lack of correlation between modeled and observed data (04:00 - 10:00 UTC) ($RMSE = 48.04 W m^{-2}$, $MBE = 26.17 W m^{-2}$, $NSC = -1.65$, $R^2 = -0.30$ across all stations). For reference, New York City is UTC-4 during Eastern Daylight Time (EDT) (applies to months JJA, SON, and MAM) and UTC-5 during Eastern Standard Time (EST) (applies to months DJF). As shown in Figure 7, these diurnal performance trends are further supported by seasonal diurnal averages taken for each season over all stations. The model performs considerably well during daytime hours, especially in

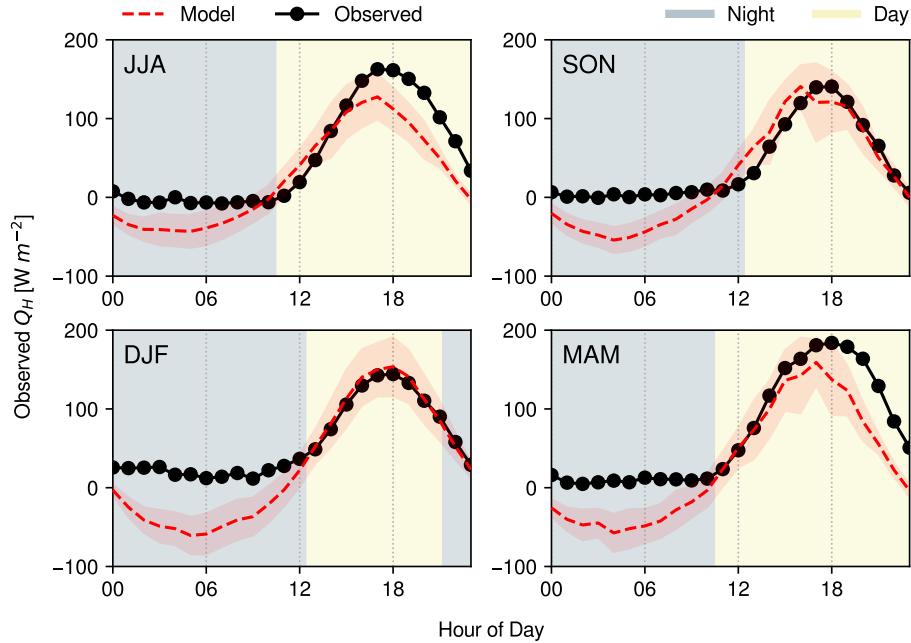


Figure 7: Seasonal averages of daily Q_H , averaged over all stations. For reference, the acronyms reflect their represented months (for example, 'JJA' corresponds to June, July, and August). All times are in each UTC and background shading corresponds approximately to average day and night durations for the respective season.

360 the summer (JJA) and autumn (SON) months. However, the model consistently underestimates Q_H during
 361 nighttime hours - especially so in the winter (DJF) months. The performance error during the pre-dawn
 362 hours is likely due to estimation biases in the remote sensing methods used and will be discussed later in
 363 the paper.

364 3.2.3 Seasonal variability

365 The model exhibited variability in performance relative to the season during which runs were performed.
 366 As shown in Figure 8, the model appears to perform better in warmer months [summer (JJA) and spring
 367 (MAM)] than in cooler months. This is evident in the R^2 values of each season (0.80 and 0.79 for JJA and
 368 MAM, respectively) as compared to cooler months [autumn (SON) and winter (DJF)], which feature lower
 369 R^2 values (0.72 and 0.56, respectively). Model overprediction and model spread were more prevalent in the
 370 winter (DJF) and spring (MAM) months relative to the summer (JJA) and fall (SON) months, as shown in
 371 Figure 9. Analysis regarding seasonal variability in model results and performance is discussed further in
 372 Section 4.2.

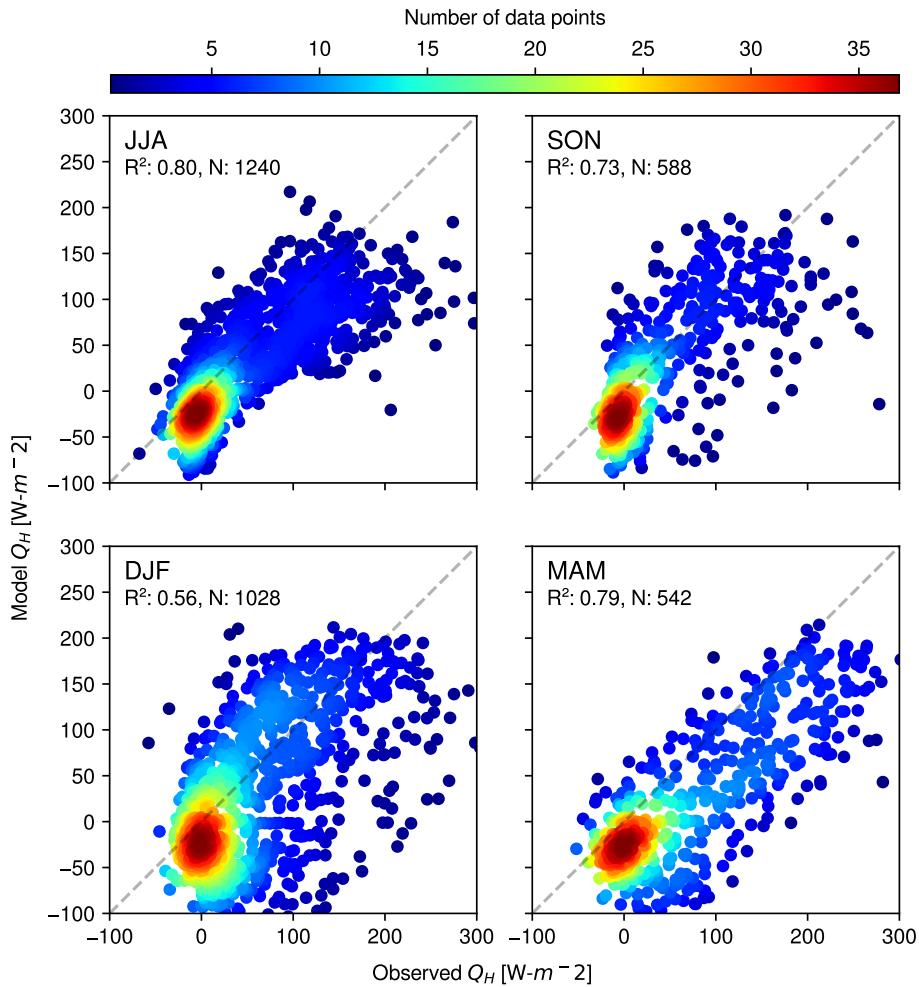


Figure 8: Comparison of model and observed Q_H divided into seasons. Note that the acronyms correspond to months in each season (for example, 'JJA' corresponds to June, July and August).

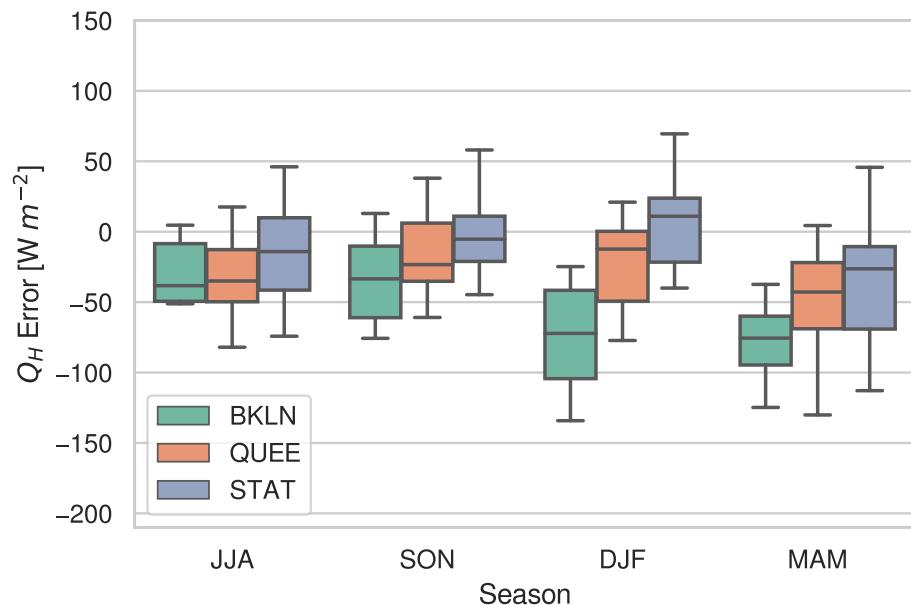


Figure 9: Q_H error between model results and observational data, averaged per season per location. For reference, the acronyms reflect their represented months (for example, 'JJA' corresponds to June, July, and August).

373 3.2.4 Geospatial variability

374 The spatial distribution of Q_H was of interest in this study due to the heterogeneity of the land cover types
 375 present within the metropolitan area. The spatial distribution of Q_H at different times during a day in late
 376 October 2019 is shown in Figure 10. The Q_H at 14:00 UTC (10:00 local time) is shown to be positive at almost
 377 every pixel with relatively low magnitudes ($<100 \text{ W}\cdot\text{m}^{-2}$), which is within the expected range of values for
 378 the mid-morning. At 17:00 UTC (13:00 local time), Q_H is near its peak value in most neighborhoods, with

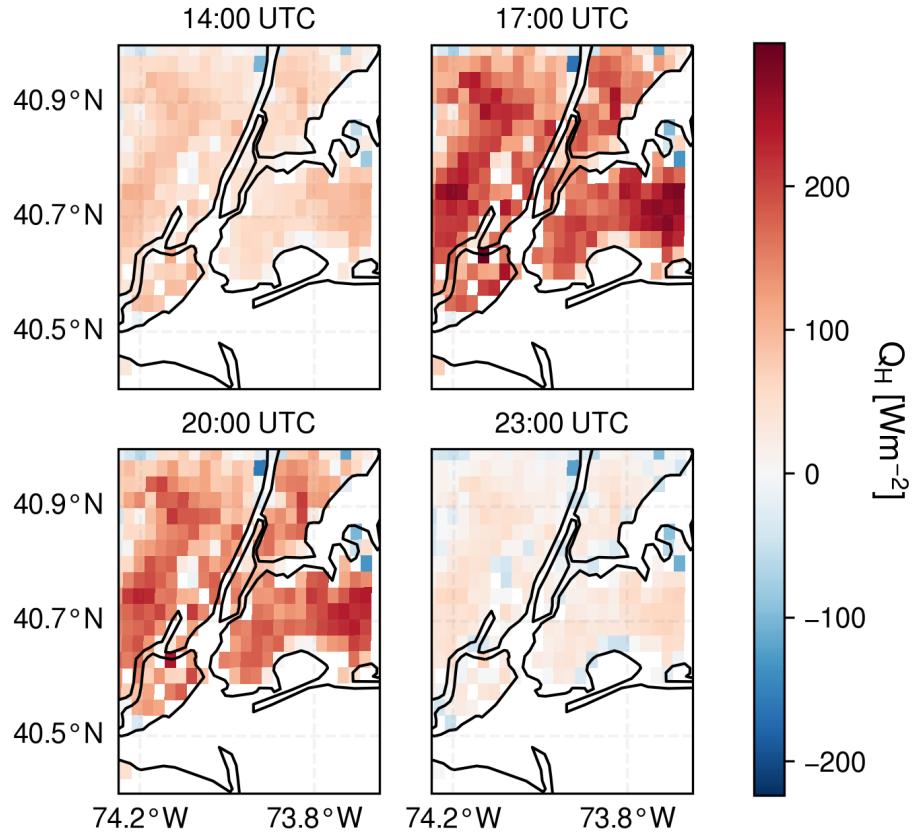


Figure 10: Q_H (W m^{-2}) in New York City on 24 October 2019 shown at 4 different times during the daytime. Note that on this date, New York City is UTC-4.

values nearing 300 W m^{-2} in sections of Queens and Newark. Decreasing values towards the northwestern outer edges (upper-left of the plot) of the metropolitan area correspond to areas with significant vegetative cover in suburban areas. At 20:00 UTC (16:00 local time), Q_H is seen to be decreasing from its peak value, which is a result of the surface layer air temperature increasing and approaching the land surface temperature. At 23:00 UTC (19:00 local time), Q_H nears zero as the sun sets and the land surface temperature decreases. Note that areas on the outer edges of the metropolitan area begin to display negative values of Q_H (equivalent to positive downward surface sensible heat flux) while more highly-urbanized areas continue to demonstrate positive Q_H . This correlates with the elevated heat storage (Q_S) that has been found to last longer into the nighttime hours in urban areas as compared to rural areas (Grimmond & Oke (1999), Hrisko et al. (2021)).

Based on the results presented in Table 1, it is apparent that Staten Island (STAT) features better model correlation than Brooklyn (BKLN) or Queens (QUEE) over all metrics except R^2 , although it is similar in value to the other 2 stations. The Staten Island flux tower is located in a less urbanized vicinity than the others (57.48% developed, per NLCD classifications) compared to Brooklyn (99.73% developed) and Queens (82.16% developed). This disparity in urbanization may indicate that land cover properties may be more homogeneous and extraneous heating sources (i.e. building processes, exhaust from utilities) may play less of a role in near-surface heating than in more heavily-urbanized areas. However, to properly determine any correlation between land cover type and model performance, validation is needed against additional flux towers over a wider range of land cover types within the city to increase confidence in any observed trends.

3.3 Model sensitivity analysis

An analysis of model sensitivity to input parameter was performed to determine the response of the model to inputs to help determine the roles of inputs to the estimation of Q_H . The analysis evaluates 5 input parameters: T_{LST} , T_{air} , h_0 , u_{air} , and z_r . These parameters were chosen due to their presence in every parameter involved in estimating Q_H , whether directly observed or derived. The analysis was performed

402 by independently perturbing the value of each parameter above and below the original value by a defined
 403 quantity, such that the effect of each parameter. Model sensitivity is defined as the percentage error of
 404 Q_H between the model runs with modified input parameter values and unmodified input values. Parameter
 405 perturbation values are shown in Table 2. Results from the sensitivity analysis are shown in Table 2 and
 406 Figure 11 categorized by location and atmospheric stability, ζ . For reference, stability values are categorized
 407 into 'unstable', 'neutral', and 'stable', corresponding to values of $\zeta < -0.25$, $-0.25 \leq \zeta < 0.25$, and $0.25 \leq \zeta$,
 408 respectively.

409 Results show high model sensitivity to T_{LST} , T_{air} , with lesser but significant sensitivity to u_{air} (wind
 410 speed), and minor sensitivity to h_0 (element roughness height) and z_r (reference height). This order of
 411 sensitivity to perturbed parameters is similar to the sensitivity analysis results presented in a study by
 412 Feigenwinter et al. (2018). The high sensitivity of flux estimation to temperature is observed in the literature
 413 (Brenner et al. (2017), Cammalleri et al. (2012), Xia et al. (2016),) and may also be a function of the 2-km
 414 spatial resolution of the surface and air temperature gridded data, as shown by results from both Mott
 415 et al. (2015) and Xu et al. (2008), where an increase in grid resolution significantly altered model results.
 416 With regards to location, model sensitivity to all parameters is similar across all 3 validation locations,
 417 with the exception of higher sensitivity to T_{LST} and T_{air} at Staten Island. A potential cause for this
 418 additional sensitivity is the mixture of land cover types that have very different properties within the same
 419 GOES-16 satellite pixel, namely deciduous forest (NLCD class 41) and medium-density developed land
 420 (NLCD class 23) (Yang et al. 2018). With regards to atmospheric stability, the model was most sensitive to
 421 temperature perturbations during periods of near-neutral conditions, with moderate sensitivity to wind speed
 422 perturbations in non-neutral regimes. It is worth noting that the model is less sensitive to perturbations over
 423 all parameters during periods of instability, likely due to enhanced mixing and homogenization of properties
 424 through the majority of the boundary layer depth (Garratt (1994), Stull (1988)).

Table 2: Model sensitivity analysis inputs, perturbation values, and results.

Parameter	Location	Perturbation Value	1st Quartile Error (%)	3rd Quartile Error (%)
T_{air}	BKLN		-13.15	13.11
	QUEE	± 0.5 K	-13.23	13.21
	STAT		-18.95	18.89
T_{LST}	BKLN		-13.30	13.30
	QUEE	± 0.5 K	-13.15	13.11
	STAT		-19.09	19.10
h_0	BKLN		-4.47	4.56
	QUEE	± 0.5 m	-3.48	3.48
	STAT		-5.04	5.10
u_{air}	BKLN		-11.74	9.30
	QUEE	$1\ ms^{-1}$	-13.13	10.77
	STAT		-12.73	10.53
z_r	BKLN		-2.05	2.20
	QUEE	1 m	-1.00	1.04
	STAT		-1.81	1.93

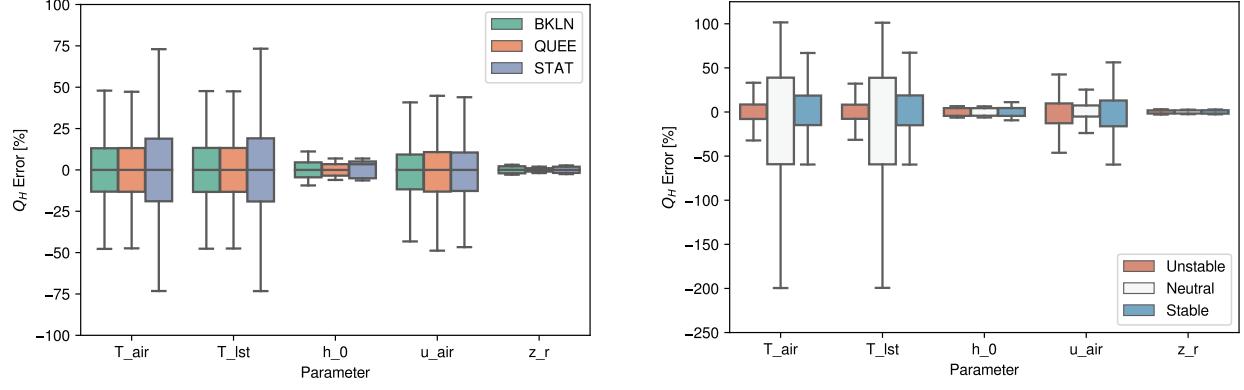


Figure 11: Left: Model sensitivity to each parameter represented by error from baseline values categorized by location (left) and atmospheric stability, ζ (right).

425 3.4 Model performance against uWRF

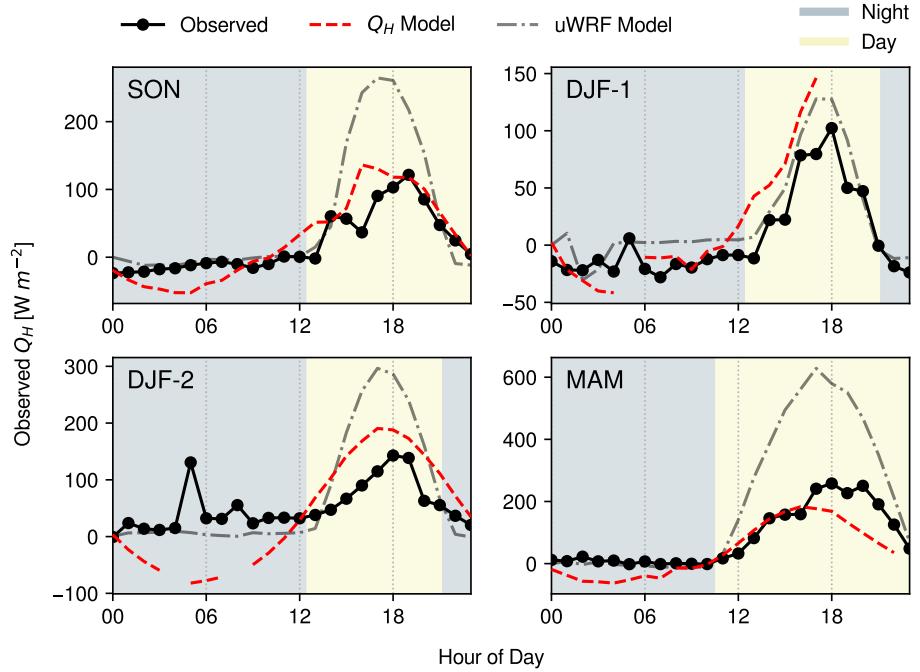


Figure 12: Observed and modeled (uWRF and dedicated) Q_H at days selected for uWRF study at the Queens (QUEE) station. Note that 'DJF-1' represents the date 23 December 2019 and 'DJF-2' represents 20 January 2020.

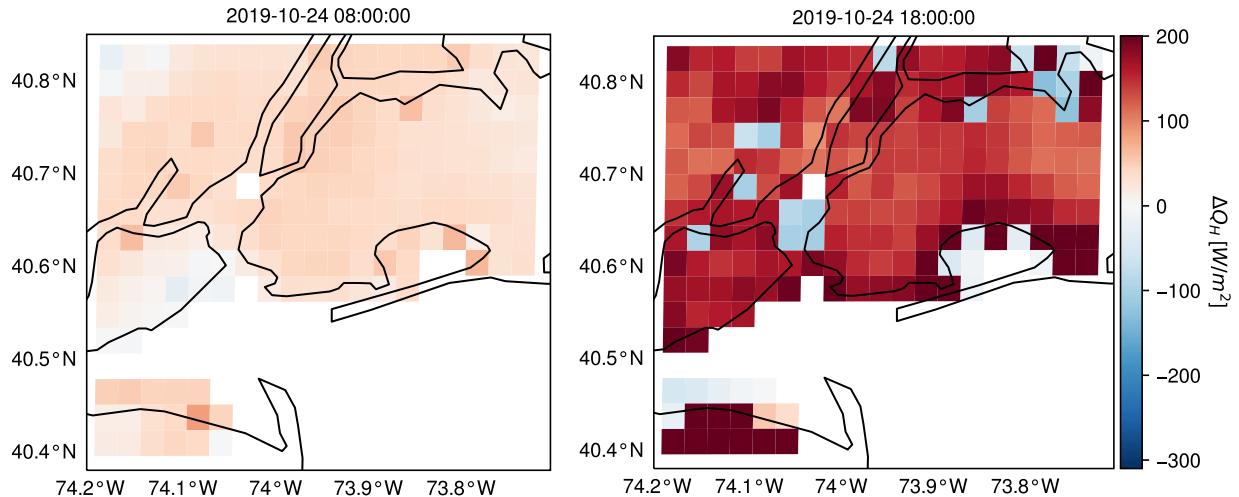


Figure 13: Gridded maps showing error between dedicated Q_H and uWRF models, October 24, 2019 at 08:00 (left) and 18:00 UTC (right). Light blue pixels correspond to areas with mostly covered in water/marsh.

426 The uWRF model was used as an additional metric to evaluate performance of the dedicated Q_H model
 427 relative to observed data. Additionally, comparison with the uWRF model provides the ability to validate the
 428 dedicated Q_H model over a continuous spatial extent not afforded by the single-point observation stations.

429 The model was run on 4 separate days, for 24 hours each: 24 October 2019 (SON), 23 December 2019
430 (DJF-1), 20 January 2020 (DJF-2), and 12 May 2020 (MAM). The spatial pixel nearest to each ground
431 station was used for comparison. Each ground station used in the Q_H to ground station validation study
432 (BKLN, QUEE, STAT) was selected to produce the comparison.

433 Over the days analyzed in the study, the RMSE between the uWRF model and ground observation
434 stations was 108.07 W m^{-2} , with a MBE of -30.11 W m^{-2} , a NSC of -1.47, and a R^2 of 0.63. The performance
435 statistics show considerably poorer performance than the dedicated Q_H model, relative to observed data.
436 The uWRF model consistently overpredicted daytime Q_H (see Figures 12 and 13), although it predicted
437 nocturnal Q_H more accurately than the dedicated Q_H model.

438 4 Discussion

439 4.1 Comparison with other studies

440 There is some difficulty in directly comparing this model with other estimation methods due to the lack
441 of studies evaluating the performance of estimation methods for Q_H in urban areas over a continuous time
442 period using remote sensing methods in the reviewed literature. Although a large body of work exists for
443 proposing and evaluating methods for assessing surface fluxes using remotely-sensed data, these studies
444 primarily focus on the estimation of evapotranspiration and latent heat fluxes in agricultural, forested, or
445 grassland areas. For this reason, these studies are not comparable to the work presented herein. Therefore,
446 this section will attempt to compare the performance of the model described herein to the performance of
447 other studies that estimate Q_H in both urban and rural areas using remote sensing methods.

448 Several studies have used airborne methods to estimate Q_H over rural areas. In [Cammalleri et al. \(2012\)](#),
449 aircraft-mounted multispectral and thermal cameras were used in conjunction with meteorological data to
450 estimate Q_H over 7 days within a 4 month period, with a study area covered by cropland, fallow soil, and
451 bare soil. Using small aperture scintillometers to validate estimated values, the study found errors ranging
452 from -35 to 20 W m^{-2} , which are small but not negligible for the reported flux values. It is worth noting that
453 this study uses 2 distinct numerical methods to estimate Q_H , both featuring reasonable accuracy. [Kim &](#)
454 [Kwon \(2019\)](#) used unmanned aerial vehicles (UAVs) over the course of 11 flights to evaluate fluxes a variety
455 of land cover types in rural areas over a range of synoptic meteorological conditions. A bulk parameterization
456 method was used to estimate Q_H , with eddy covariance and scintillometry used as validation methods. This
457 study found a correlation coefficient (R) of 0.94, with a RMSE of -1.26 W m^{-2} , and a MBE of 19.91 W m^{-2} .
458 [Ortega-Farías et al. \(2016\)](#) described using an unmanned aerial vehicle (UAV) over an orchard over 10 days
459 between February and March 2014 to estimate fluxes. Using eddy covariance methods for validation, the
460 study found a RMSE of 56.00 W m^{-2} and a mean average error (MAE) of 46.00 W m^{-2} . All studies showed
461 promising methods for estimating Q_H using airborne measurements as all features good agreement between
462 estimation and validation methods.

463 Fewer studies have used satellite imagery to estimate Q_H with comprehensive validation measures. [Miglietta et al. \(2009\)](#) describes an estimation method using Meteosat land surface temperature and radiation
464 products, as well as aircraft-mounted sensors, to evaluate fluxes over forested areas and cropland between
465 May and June 2005. Using eddy covariance methods to observe flux values, the study reported reasonable
466 agreement between estimated and observed temperature and net radiation values, although Q_H overestimation
467 ranged up to 30% over the study period. [Mkhwanazi et al. \(2012\)](#) used Landsat 5 imagery with a bulk
468 parameterization method to evaluate fluxes over an alfalfa field in rural Colorado. Despite good correlation
469

($R^2 = 0.80$) and moderate errors ($\text{RMSE} = 59.60 \text{ W}\cdot\text{m}^{-2}$, $\text{MBE} = 31.79 \text{ W}\cdot\text{m}^{-2}$), the Nash-Sutcliffe coefficient was negative, indicating suboptimal model performance. [Feigenwinter et al. \(2018\)](#) used Landsat 8 imagery over an urban area (Basel, Switzerland) over 22 days between 2013 and 2015 with 3 flux towers used as validation. This study provided the most comparable estimation method of fluxes in an urban areas to the knowledge of the authors, with specific information regarding urban land cover types and similar validation methods. The study found good agreement between estimated and observed values, of Q_H , with an overall R^2 value of 0.71 and an RMSE of $54.00 \text{ W}\cdot\text{m}^{-2}$, indicating performance similar to that of the model described herein.

Although these methods present novel and effective ways of estimating Q_H , the main shortfalls include spatial and temporal variability in the results presented. All airborne and most of the satellite-based estimation methods are confined to rural areas, which are dominated by homogeneous land cover types, none of which are urban or built-up land cover types. Additionally, all estimation methods reviewed in the literature rely on temporally infrequent or disjoint data acquisition methods, some of which are especially costly (non-UAV airborne missions). Because of these issues, proper evaluation of model performance as a function of land cover heterogeneity (especially in urban areas) and diurnal and seasonal meteorological conditions are difficult to perform, if at all possible. The work presented in this paper attempts to present a unique approach to address these shortcomings to allow for such an estimation method, and an evaluation with regards to the discussed parameters, to exist.

4.2 Model variability & performance

Model performance will be discussed in this section with regards to diurnal, seasonal, and geospatial characteristics.

The model performs much better during the day, as shown in Figure 3.2.2. The difference in performance could be related to the prevailing atmospheric conditions; daytime transport of heat is facilitated by thermals (buoyancy-generated turbulence) that are highly efficient in mixing heat and other scalars. In urban areas, the largest thermals could be of the size of the boundary layer, on the order of 1-2 km, which is close to the spatial resolution of GOES-16 pixels. In stark contrast, during the nighttime hours, when the urban atmosphere is less convectively active, the transport is mainly dominated by mechanical turbulence through wind shear, which is highly localized. Additionally, the wind field used in the model to derive u^* is obtained from an ASOS station that might be unrepresentative of the nearest GOES-16 LST pixel. This error source is discussed further in the next section.

The model performed best in the summer (JJA) and the worst in the winter (DJF). As referenced in Section 3.2.4, model overprediction was highest in winter (DJF) and spring (MAM) months. Model overprediction in winter and spring months may be explained by a couple of different factors. One potential cause is the increased frequency of midlatitude cyclones affecting the northeastern United States in winter and spring months as compared to summer months ([Gedzelman et al. 2003](#)). The meteorological impacts of these events include increased wind speeds, a shift from westerly to northerly surface winds, and advected air masses from nearby bodies of water (namely, Long Island Sound and the Atlantic Ocean). These events may exacerbate disparities in wind speed differences between the source of wind measurement data, which are all within very close proximity to the coast (JFK is directly on Jamaica Bay and is 3.5 km from the Atlantic Ocean, LGA is directly on Long Island Sound, EWR borders Newark Bay and is 8 km from Upper New York Bay), and the locations where flux estimates are made, which are further removed from the coast (BKLN is 5 km from Jamaica Bay, QUEE is 3 km from Long Island Sound, STAT is 6 km from Lower

512 New York Bay). Additionally, the advected cold air masses during these events tend to cause temperature
513 inversions that may increase errors in estimating T_{LST} and T_{air} by remote sensing methods (Tang et al.
514 2016). This disparity in wind speeds may cause higher u^* values, resulting in overprediction of Q_H since
515 it is proportional to u^* . Another potential cause of model overprediction is the seasonal variability in the
516 mean mixed layer height of the boundary layer. Numerous studies of boundary layer structure climatology in
517 urban areas reveals that mixed layer height is generally lowest in winter months (van der Kamp & McKendry
518 (2010), de Arruda Moreira et al. (2020)). A lower mixed layer height is indicative of less mixing of scalars
519 (such as temperature), leading to a less homogeneous boundary layer where the effects of surface forcings
520 remain more localized as compared to a boundary layer with more mixing and a more uniform composition.
521 This localization of the effects of surface forcings on the mixed layer may result in a disparity in results due
522 to the difference in location between the location of the ASOS observation stations and the locations of the
523 flux towers, the latter of which is where the model is run and validated.

524 The geospatial variability of Q_H in coastal areas is high due to the complex boundary layer formed by
525 the combination of a marine and urban boundary layer (Melecio-Vazquez et al. (2018), Thompson et al.
526 (2007)). Q_H and its transport may be influenced by advective phenomena such as sea breezes, which are
527 further amplified by the sharp transitions between different land cover types, such as the transition from
528 water to a highly-developed urban area, as exists along the portions of New York City bordering the Hudson
529 and East Rivers, Long Island Sound, or the Atlantic Ocean (Bou-Zeid et al. (2020), Lee (2015), Thompson
530 et al. (2007)). The inability to explicitly capture the influence of advection on Q_H is a limitation of this
531 model. Additionally, the 2 km spatial resolution of GOES-16 satellite data is unable to properly capture
532 areas with sharp land cover transitions on land, such as the one at the Staten Island (STAT) flux tower (see
533 Figure 3), which may introduce significant bias into the estimation. This is discussed in greater detail in
534 Section 4.3.

535 4.3 Potential sources of error

536 Numerous assumptions were made in the development of the model that may have contributed to model
537 error.

538 A likely source of model error stemmed from the moderate spatial resolution of the GOES-16 LST product.
539 The LST product features pixels at a spatial resolution of 2 km, which translates to 196 pixels spanning the
540 New York City area (approximate land area of 778 km^2). Although this allows for fragmentation of the city
541 into pixels that can distinguish districts (such as boroughs for New York City) from each other, the lack of
542 higher resolution may fail to capture localized phenomena from smaller neighborhoods in each district, such
543 as the effects of urban street canyons and vegetated spaces, that may influence neighborhood-averaged fluxes
544 (Erell & Williamson (2006), Xiaomin et al. (2006),). This is important due to the highly variable land cover
545 types that exist between neighborhoods in large cities such as New York City (Hamstead et al. (2016)).

546 Another potential source of error arises from biases in estimation methods for T_{LST} and T_{air} used in
547 this model. The GOES-R Land Surface Temperature product has been shown to have an average precision
548 error of 1.58 K when tested over 6 rural locations (Yu et al. 2011). It is worth noting validation for this
549 product has not been performed in urban areas, to the authors' knowledge. The air temperature model used
550 features an RMSE of 2.6 K and a bias of 0.8 K (Hrisko et al. 2020). Additionally, this air temperature model
551 featured a cold bias during nighttime hours and a warm bias during daytime hours, leading to under- and
552 over-prediction respectively. The combined errors, in conjunction with the high model sensitivity to both
553 T_{LST} and T_{air} , can cause significant errors in estimating Q_H , especially during nighttime hours.

554 Another potential source of error comes from the selection of h_0 for land cover types classified as "urban"
555 (Developed, High Intensity; Developed, Medium Intensity; Developed, Low Intensity; per the NLCD). The
556 assumed values were derived from the WRF-ARW model assumptions. However, the values may vary widely
557 from city to city, depending on the average heights of buildings in each. In a city with a large number of
558 tall buildings (>10 m) such as New York City, larger values of h_0 for each class may be more suitable to
559 properly parameterize the boundary layer wind profile.

560 Another additional source of error between the model and the NYS Mesonet observations is the distance
561 between the ASOS stations, where wind speed data is collected, and the Mesonet flux stations, where flux
562 data is collected. The ASOS stations used for data collection are all located at large airports that feature
563 large expanses of flat surfaces surrounding the station instrumentation with minimal upwind obstructions.
564 By comparison, each flux station is located in a moderately- to heavily-urbanized area, reducing the upwind
565 fetch and exposing instrumentation to hyperlocal turbulence that is a direct function of the surrounding
566 geometry (Kastner-Klein et al. (2004)). Additionally, the highly heterogeneous land cover in the New York
567 City metropolitan area has the potential to create highly-localized wind fields due to phenomena such as the
568 urban street canyon effect in areas with tall buildings and sea breezes in coastal areas (Park et al. (2012),
569 Thompson et al. (2007)).

570 4.4 Future work

571 A number of factors from this study motivate future work to improve the accuracy of the model.
572 A possible improvement to the model involves validation at a range of test sites with a variety of land
573 cover types at different locations, both within New York City and in other urban areas. This allows for the
574 model to be evaluated for a wider range of land cover types and permits the evaluation of model sensitivity
575 to land cover type. Moreover, due to the study focusing on one city, atmospheric conditions that are a
576 function of location, such as air pressure or advective fluxes, may not be fully accounted for in this model,
577 potentially requiring a modification of assumptions or parameter values. A related additional improvement
578 to the validation effort would be the use of scintillometry to obtain path-averaged flux observations that are
579 more representative of urban land cover types, as opposed to point-based observations as used in this study
580 (Crawford et al. (2017), Lee (2015), Nadeau et al. (2009)) due to the significant heterogeneity of land cover
581 types in urban areas.

582 Another improvement to the model involves downscaling of the z_m grid. A critical component of the
583 model is the calculation of z_m , which is a factor in nearly every component of the turbulence parameteriza-
584 tion. Due to the highly heterogeneous nature of urban areas, z_m must be calculated at an extremely high
585 spatial resolution to properly represent the corresponding land cover. Although the NLCD has a spatial
586 resolution of 30 meters, the spatial resolution of the z_m calculations is driven by the GOES-16 satellite
587 spatial resolution. Therefore, a higher-resolution satellite or a downscaling algorithm for the GOES-16 LST
588 product would likely improve the calculation of z_m and in turn, the calculation of all dependent parameters.
589 Potential tools for increasing the spatial resolution of Q_H through the fusion of higher-resolution datasets
590 with GOES-16 satellite data, such as incorporation of remotely-sensed surface properties at higher resolu-
591 tions (AVHRR or MODIS infrared band data) (Bala et al. (2019), Bonafoni (2016)) and machine learning
592 techniques (Chrysoulakis et al. (2018), Hrisko et al. (2021)).

593 The estimation of nocturnal sensible heat flux is another critical component to improving model accuracy.
594 The model often underestimates relative to the observation sites despite good approximation during the day.
595 In addition to the bias in estimating T_{LST} and T_{air} as discussed in Section 4.3, the connection between

596 sensible and storage heat is a potential explanation for the near-zero and positive Q_H at night in the
597 urbanized portions of the study area ([Grimmond & Oke \(1999\)](#), [Hrisko et al. \(2021\)](#)). Correction for nocturnal
598 temperature bias and incorporation of the relationship between heat storage and Q_H into the algorithm may
599 improve model accuracy, and this topic requires further exploration.

600 4.5 Application potential

601 The dedicated Q_H model leverages open-access satellite and land cover data that allows for a cost-effective
602 way to analyze sensible heat flux in urban areas. The model enables Q_H to be estimated at any point within
603 the scope of the GOES-16 satellite imagery with reasonable accuracy, removing constraints to single-point
604 observation stations. Consequently, the model can be used to identify a number of factors that contribute to
605 or correlate with the effects of urban heat islands in major cities, which directly relate to the vulnerability
606 of a neighborhood due to the effects of climate. The model is especially valuable in locations that are not
607 in close proximity to flux observation stations. Additionally, the model can be used as a module for high-
608 resolution numerical weather models to improve the spatial resolution of Q_H estimation in areas of interest.
609 Moreover, the geographical extent spanned by the GOES-16 satellite imagery allows the model to be used
610 over wide swaths of the CONUS, allowing for Q_H estimations to be performed efficiently over multiple urban
611 areas using the same imagery data at hourly intervals.

612 5 Conclusions

613 A dedicated satellite-based model using NOAA's GOES-16 data to calculate sensible heat flux in urban areas
614 was introduced. The model couples GOES-16 data and publicly-accessible land cover data in an iterative
615 turbulence parameterization based on MOST to provide a product that is capable of calculating Q_H in areas
616 with highly heterogeneous land cover. The performance of the model was validated using an ample set of
617 ground station observations in New York City. Additionally, the model was compared to an urbanized WRF
618 model and performed significantly better relative to observational data. Accordingly, these validation and
619 comparison procedures suggest that the dedicated model is reasonably accurate in estimating Q_H in urban
620 areas at sub-hourly timescales.

621 Over the duration of the validation period, the RMSE between the model and observational data was
622 47.32 W m^{-2} , with a MBE of 16.58 W m^{-2} , a model NSC of 0.54, and a R^2 coefficient of 0.70. This presents
623 a significant improvement over the numerical WRF model run over fewer days in the same validation period
624 (RMSE of 108.1 W m^{-2} , MBE of -30.1 W m^{-2} , NSC of -1.47 and R^2 of 0.63). The model performed especially
625 well in warmer months (R^2 coefficients of 0.80 and 0.79 for summer and autumn, respectively) and during
626 the daytime and evening hours.

627 The development of a satellite-based Q_H model demonstrates the potential of the use of satellite data for
628 estimating atmospheric processes over large spatial and temporal domains. The ability to leverage this data
629 for use in urban areas is valuable as this method resolves several limitations that are encountered in estimat-
630 ing atmospheric processes in areas with highly heterogeneous land cover and an insufficient observational
631 infrastructure. This ability is especially important considering the impact of heat-related weather events on
632 human populations, allowing for risk assessment and mitigation strategies to become better informed with
633 improved supporting data.

634 Acknowledgments

635 This study is supported and monitored by The National Oceanic and Atmospheric Administration – Cooper-
636 ative Science Center for Earth System Sciences and Remote Sensing Technologies (NOAA-CESSRST) under
637 the Cooperative Agreement Grant #NA16SEC4810008. The authors would like to thank The City College
638 of New York, NOAA-CESSRST program and NOAA Office of Education, Educational Partnership Program
639 for fellowship support for Gabriel Rios. The statements contained within the manuscript/research article
640 are not the opinions of the funding agency or the U.S. government, but reflect the author's opinions. The
641 research was also funded by the Department of Defense Army Research Office Grant No. W911NF2020126.
642 This research is made possible by the New York State (NYS) Mesonet. Original funding for the NYS Mesonet
643 was provided by Federal Emergency Management Agency grant FEMA-4085-DR-NY, with the continued
644 support of the NYS Division of Homeland Security & Emergency Services; the state of New York; the Re-
645 search Foundation for the State University of New York (SUNY); the University at Albany, SUNY; the
646 Atmospheric Sciences Research Center (ASRC) at SUNY Albany; and the Department of Atmospheric and
647 Environmental Sciences (DAES) at SUNY Albany. The authors would also like to thank the reviewers for
648 their feedback, which has helped improve the quality of this paper.

649 References

- 650 Bala, R., Prasad, R. & Yadav, V. P. (2019), 'Disaggregation of modis land surface temperature in urban
651 areas using improved thermal sharpening techniques', *Advances in Space Research* **64**(3), 591–602.
- 652 Best, M. (2005), 'Representing urban areas within operational numerical weather prediction models',
653 *Boundary-Layer Meteorology* **114**(1), 91–109.
- 654 Bonafoni, S. (2016), 'Downscaling of landsat and modis land surface temperature over the heterogeneous
655 urban area of milan', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*
656 **9**(5), 2019–2027.
- 657 Bou-Zeid, E., Anderson, W., Katul, G. G. & Mahrt, L. (2020), 'The persistent challenge of surface hetero-
658 geneity in boundary-layer meteorology: a review', *Boundary-Layer Meteorology* **177**(2), 227–245.
- 659 Brenner, C., Thiem, C. E., Wizemann, H.-D., Bernhardt, M. & Schulz, K. (2017), 'Estimating spatially
660 distributed turbulent heat fluxes from high-resolution thermal imagery acquired with a uav system', *In-
661 ternational Journal of Remote Sensing* **38**(8-10), 3003–3026.
- 662 Businger, J. A., Wyngaard, J. C., Izumi, Y. & Bradley, E. F. (1971), 'Flux-profile relationships in the
663 atmospheric surface layer', *Journal of Atmospheric Sciences* **28**(2), 181–189.
- 664 Cammalleri, C., Anderson, M., Ciraolo, G., D'urso, G., Kustas, W., La Loggia, G. & Minacapilli, M. (2012),
665 'Applications of a remote sensing-based two-source energy balance algorithm for mapping surface fluxes
666 without in situ air temperature observations', *Remote Sensing of Environment* **124**, 502–515.
- 667 Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C., Grossman-Clarke, S., Loridan, T., Man-
668 ning, K. W., Martilli, A., Miao, S. et al. (2011), 'The integrated wrf/urban modelling system: develop-
669 ment, evaluation, and applications to urban environmental problems', *International Journal of Climatology*
670 **31**(2), 273–288.

- 671 Chen, F. & Zhang, Y. (2009), ‘On the coupling strength between the land surface and the atmosphere: From
672 viewpoint of surface exchange coefficients’, *Geophysical Research Letters* **36**(10).
- 673 **URL:** <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL037980>
- 674 Chrysoulakis, N., Grimmond, S., Feigenwinter, C., Lindberg, F., Gastellu-Etchegorry, J.-P., Marconcini,
675 Mitraka, Z., Stagakis, S., Crawford, B., Olofson, F., Landier, L., Morrison, W. & Parlow, E. (2018),
676 ‘Urban energy exchanges monitoring from space’, *Scientific Reports* **8**(1), 11498.
- 677 **URL:** <http://www.nature.com/articles/s41598-018-29873-x>
- 678 Crawford, B., Grimmond, C. S. B., Ward, H. C., Morrison, W. & Kotthaus, S. (2017), ‘Spatial and tempo-
679 ral patterns of surface–atmosphere energy exchange in a dense urban environment using scintillometry’,
680 *Quarterly Journal of the Royal Meteorological Society* **143**(703), 817–833.
- 681 de Arruda Moreira, G., Guerrero-Rascado, J. L., Bravo-Aranda, J. A., Foyo-Moreno, I., Cazorla, A., Alados,
682 I., Lyamani, H., Landulfo, E. & Alados-Arboledas, L. (2020), ‘Study of the planetary boundary layer
683 height in an urban environment using a combination of microwave radiometer and ceilometer’, *Atmospheric
684 Research* **240**, 104932.
- 685 Dudhia, J. (1989), ‘Numerical study of convection observed during the winter monsoon experiment using a
686 mesoscale two-dimensional model’, *Journal of Atmospheric Sciences* **46**(20), 3077–3107.
- 687 Dyer, A. J. (1974), ‘A review of flux-profile relationships’, *Boundary-Layer Meteorology* **7**(3), 363–372.
- 688 **URL:** <http://link.springer.com/10.1007/BF00240838>
- 689 Erell, E. & Williamson, T. (2006), ‘Simulating air temperature in an urban street canyon in all weather
690 conditions using measured data at a reference meteorological station’, *International Journal of Climatology:
691 A Journal of the Royal Meteorological Society* **26**(12), 1671–1694.
- 692 Feddema, J. J., Oleson, K. W., Bonan, G. B., Mearns, L. O., Buja, L. E., Meehl, G. A. & Washington, W. M.
693 (2005), ‘The importance of land-cover change in simulating future climates’, *Science* **310**(5754), 1674–1678.
- 694 Feigenwinter, C., Vogt, R., Parlow, E., Lindberg, F., Marconcini, M., Del Frate, F. & Chrysoulakis, N.
695 (2018), ‘Spatial distribution of sensible and latent heat flux in the city of basel (switzerland)’, *IEEE
696 Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **11**(8), 2717–2723.
- 697 Gamarro, H., Gonzalez, J. E. & Ortiz, L. E. (2019), ‘On the assessment of a numerical weather prediction
698 model for solar photovoltaic power forecasts in cities’, *Journal of Energy Resources Technology* **141**(6).
- 699 Garratt, J. R. (1994), ‘The atmospheric boundary layer’, *Earth-Science Reviews* **37**(1-2), 89–134.
- 700 Gedzelman, S., Austin, S., Cermak, R., Stefano, N., Partridge, S., Quesenberry, S. & Robinson, D. (2003),
701 ‘Mesoscale aspects of the urban heat island around new york city’, *Theoretical and applied climatology
702* **75**(1), 29–42.
- 703 Grimmond, C. & Cleugh, H. (1994), ‘A simple method to determine obukhov lengths for suburban areas’,
704 *Journal of Applied Meteorology* **33**(3), 435–440.
- 705 Grimmond, C. & Oke, T. R. (1999), ‘Heat storage in urban areas: Local-scale observations and evaluation
706 of a simple model’, *Journal of Applied Meteorology and Climatology* **38**(7), 922–940.
- 707 Gutiérrez, E., González, J. E., Martilli, A. & Bornstein, R. (2015), ‘On the anthropogenic heat fluxes using
708 an air conditioning evaporative cooling parameterization for mesoscale urban canopy models’, *Journal of
709 Solar Energy Engineering* **137**(5), 051005.

- 710 Gutiérrez, E., González, J. E., Martilli, A., Bornstein, R. & Arend, M. (2015), ‘Simulations of a heat-wave
711 event in new york city using a multilayer urban parameterization’, *Journal of Applied Meteorology and*
712 *Climatology* **54**(2), 283–301.
- 713 Gutiérrez, E., Martilli, A., Santiago, J. & González, J. (2015), ‘A mechanical drag coefficient formulation
714 and urban canopy parameter assimilation technique for complex urban environments’, *Boundary-Layer*
715 *Meteorology* **157**(2), 333–341.
- 716 Hamstead, Z. A., Kremer, P., Larondelle, N., McPhearson, T. & Haase, D. (2016), ‘Classification of the
717 heterogeneous structure of urban landscapes (sturla) as an indicator of landscape function applied to
718 surface temperature in new york city’, *Ecological Indicators* **70**, 574 – 585. Navigating Urban Complexity:
719 Advancing Understanding of Urban Social – Ecological Systems for Transformation and Resilience.
720 **URL:** <http://www.sciencedirect.com/science/article/pii/S1470160X1500549X>
- 721 Hong, S.-Y. & Dudhia, J. (2012), ‘Next-generation numerical weather prediction: Bridging parameterization,
722 explicit clouds, and large eddies’, *Bulletin of the American Meteorological Society* **93**(1), ES6–ES9.
- 723 Hrisko, J., Ramamurthy, P., Melecio-Vazquez, D. & Gonzalez, J. E. (2021), ‘Spatiotemporal variability of
724 heat storage in major u.s. cities—a satellite-based analysis’, *Remote Sensing* **13**, 59.
725 **URL:** <https://www.mdpi.com/2072-4292/13/1/59>
- 726 Hrisko, J., Ramamurthy, P., Yu, Y., Yu, P. & Melecio-Vázquez, D. (2020), ‘Urban air temperature model
727 using goes-16 lst and a diurnal regressive neural network algorithm’, *Remote Sensing of Environment*
728 **237**, 111495.
729 **URL:** <http://www.sciencedirect.com/science/article/pii/S0034425719305140>
- 730 Imran, H. M., Kala, J., Ng, A. & Muthukumaran, S. (2018), ‘Effectiveness of green and cool roofs in
731 mitigating urban heat island effects during a heatwave event in the city of melbourne in southeast australia’,
732 *Journal of Cleaner Production* **197**, 393–405.
- 733 Janjić, Z. I. (1994), ‘The step-mountain eta coordinate model: Further developments of the convection,
734 viscous sublayer, and turbulence closure schemes’, *Monthly weather review* **122**(5), 927–945.
- 735 Kastner-Klein, P., Berkowicz, R. & Britter, R. (2004), ‘The influence of street architecture on flow and
736 dispersion in street canyons’, *Meteorology and Atmospheric Physics* **87**(1), 121–131.
- 737 Kato, S. & Yamaguchi, Y. (2005), ‘Analysis of urban heat-island effect using aster and etm+ data: Separation
738 of anthropogenic heat discharge and natural heat radiation from sensible heat flux’, *Remote Sensing of*
739 *Environment* **99**(1), 44 – 54. Scientific Results from ASTER.
740 **URL:** <http://www.sciencedirect.com/science/article/pii/S0034425705001707>
- 741 Kim, M.-S. & Kwon, B. H. (2019), ‘Estimation of sensible heat flux and atmospheric boundary layer height
742 using an unmanned aerial vehicle’, *Atmosphere* **10**(7), 363.
- 743 Launiainen, J. & Vihma, T. (1990), ‘Derivation of turbulent surface fluxes — an iterative flux-profile method
744 allowing arbitrary observing heights’, *Environmental Software* **5**(3), 113 – 124.
745 **URL:** <http://www.sciencedirect.com/science/article/pii/026698389090021W>
- 746 Lee, S.-H. (2015), ‘Determination of turbulent sensible heat flux over a coastal maritime area using a large
747 aperture scintillometer’, *Boundary-Layer Meteorology* **157**(2), 309–319.

- 748 Legates, D. R. & McCabe Jr, G. J. (1999), ‘Evaluating the use of “goodness-of-fit” measures in hydrologic
749 and hydroclimatic model validation’, *Water resources research* **35**(1), 233–241.
- 750 Leroyer, S., Bélair, S., Husain, S. Z. & Mailhot, J. (2014), ‘Subkilometer numerical weather prediction in an
751 urban coastal area: A case study over the vancouver metropolitan area’, *Journal of Applied Meteorology
752 and Climatology* **53**(6), 1433–1453.
- 753 Li, D. & Bou-Zeid, E. (2014), ‘Quality and sensitivity of high-resolution numerical simulation of urban heat
754 islands’, *Environmental Research Letters* **9**(5), 055001.
755 **URL:** <https://doi.org/10.1088/1748-9326/9/5/055001>
- 756 Liu, Y., Shintaro, G., Zhuang, D. & Kuang, W. (2012), ‘Urban surface heat fluxes infrared remote sensing
757 inversion and their relationship with land use types’, *Journal of Geographical Sciences* **22**(4), 699–715.
758 **URL:** <http://link.springer.com/10.1007/s11442-012-0957-7>
- 759 Melecio-Vazquez, D., Ramamurthy, P., Arend, M. & Gonzalez-Cruz, J. E. (2018), ‘Thermal structure of a
760 coastal–urban boundary layer’, *Boundary-Layer Meteorology* **169**(1), 151–161.
- Mesonet, N. (2020), New york state mesonet standard network data technical information, Technical report.
URL: <https://www2.nysmesonet.org/documents/NYSMReadmestandard.pdf>
- 761 Miglietta, F., Gioli, B., Brunet, Y., Hutjes, R., Matese, A., Sarrat, C. & Zaldei, A. (2009), ‘Sensible and
762 latent heat flux from radiometric surface temperatures at the regional scale: methodology and evaluation’,
763 *Biogeosciences* **6**(10), 1975–1986.
- 764 Mkhwanazi, M., Chávez, J. L. & Rambikur, E. H. (2012), ‘Comparison of large aperture scintillometer and
765 satellite-based energy balance models in sensible heat flux and crop evapotranspiration determination’,
766 *International Journal of Remote Sensing Applications* **2**(1), 24–30.
- 767 Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J. & Clough, S. A. (1997), ‘Radiative transfer
768 for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave’, *Journal of
769 Geophysical Research: Atmospheres* **102**(D14), 16663–16682.
- 770 Monin, A. & Obukhov, A. (1954), ‘Dimensionless characteristics of turbulence in the surface layer’, *Akad.
771 Nauk SSSR, Geofiz. Inst., Tr* **24**, 163–187.
- 772 Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D. & Veith, T. L. (2007), ‘Model
773 evaluation guidelines for systematic quantification of accuracy in watershed simulations’, *Transactions of
774 the ASABE* **50**(3), 885–900.
- 775 Mott, R., Daniels, M. & Lehning, M. (2015), ‘Atmospheric flow development and associated changes in
776 turbulent sensible heat flux over a patchy mountain snow cover’, *Journal of Hydrometeorology* **16**(3), 1315–
777 1340.
- 778 Nadeau, D. F., Brutsaert, W., Parlange, M., Bou-Zeid, E., Barrenetxea, G., Couach, O., Boldi, M.-O.,
779 Selker, J. S. & Vetterli, M. (2009), ‘Estimation of urban sensible heat flux using a dense wireless network
780 of observations’, *Environmental fluid mechanics* **9**(6), 635–653.
- 781 Nash, J. E. & Sutcliffe, J. V. (1970), ‘River flow forecasting through conceptual models part i—a discussion
782 of principles’, *Journal of hydrology* **10**(3), 282–290.

- 783 Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi,
784 D., Rosero, E. et al. (2011), ‘The community noah land surface model with multiparameterization options
785 (noah-mp): 1. model description and evaluation with local-scale measurements’, *Journal of Geophysical*
786 *Research: Atmospheres* **116**(D12).
- NOAA/NESDIS/STAR (2016), Geostationary operational environmental satellite (goes)-r series advanced
baseline imager (abi) l2+ land surface temperature (lst) beta, provisional, and full validation readiness,
implementation, and management plan (rimp), Technical report.
- URL:** https://goes-r.gov/products/RIMPs/RIMP_ABI-L2FSC_v1.0.pdf
- 787 Ortega-Fariás, S., Ortega-Salazar, S., Poblete, T., Kilic, A., Allen, R., Poblete-Echeverría, C., Ahumada-
788 Orellana, L., Zuñiga, M. & Sepúlveda, D. (2016), ‘Estimation of energy balance components over a drip-
789 irrigated olive orchard using thermal and multispectral cameras placed on a helicopter-based unmanned
790 aerial vehicle (uav)’, *Remote Sensing* **8**(8), 638.
- 791 Ortiz, L. E., Gonzalez, J. E., Gutierrez, E. & Arend, M. (2017), ‘Forecasting building energy demands
792 with a coupled weather-building energy model in a dense urban environment’, *Journal of Solar Energy
793 Engineering* **139**(1), 011002.
- 794 Park, S.-B., Baik, J.-J., Raasch, S. & Letzel, M. O. (2012), ‘A large-eddy simulation study of thermal
795 effects on turbulent flow and dispersion in and above a street canyon’, *Journal of Applied Meteorology and
796 Climatology* **51**(5), 829–841.
- 797 Pond, S., Fissel, D. B. & Paulson, C. A. (1974), ‘A note on bulk aerodynamic coefficients for sensible heat
798 and moisture fluxes’, *Boundary-Layer Meteorology* **6**(1-2), 333–339.
- 799 **URL:** <http://link.springer.com/10.1007/BF00232493>
- 800 Price, J. C. (1979), ‘Assessment of the urban heat island effect through the use of satellite data’, *Monthly
801 Weather Review* **107**(11), 1554–1557.
- 802 Raupach, M. R. (1994), ‘Simplified expressions for vegetation roughness length and zero-plane displacement
803 as functions of canopy height and area index’, *Boundary-Layer Meteorology* **71**, 211–216.
- 804 **URL:** <https://doi.org/10.1007/BF00709229>
- 805 Salamanca, F. & Martilli, A. (2010), ‘A new building energy model coupled with an urban canopy pa-
806 rameterization for urban climate simulations—part ii. validation with one dimension off-line simulations’,
807 *Theoretical and Applied Climatology* **99**(3), 345–356.
- 808 Schumacher, D. L., Keune, J., Van Heerwaarden, C. C., de Arellano, J. V.-G., Teuling, A. J. & Miralles,
809 D. G. (2019), ‘Amplification of mega-heatwaves through heat torrents fuelled by upwind drought’, *Nature
810 Geoscience* **12**(9), 712–717.
- 811 Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., Huang, X.-Y. & Duda, M. (2008),
812 A Description of the Advanced Research WRF Version 3, Technical report, UCAR/NCAR.
- 813 **URL:** <http://opensky.ucar.edu/islandora/object/technotes:500>
- 814 Stull, R. (1988), *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers.
- 815 Tang, B.-H., Zhan, C., Li, Z.-L., Wu, H. & Tang, R. (2016), ‘Estimation of land surface temperature from
816 modis data for the atmosphere with air temperature inversion profile’, *IEEE Journal of Selected Topics
817 in Applied Earth Observations and Remote Sensing* **10**(6), 2976–2983.

- 818 Thompson, W. T., Holt, T. & Pullen, J. (2007), 'Investigation of a sea breeze front in an urban environment',
819 *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied*
820 *meteorology and physical oceanography* **133**(624), 579–594.
- 821 US Census Bureau (2019), 'Population and Housing Unit Estimates Tables'.
822 **URL:** <https://www.census.gov/programs-surveys/popest/data/tables.html>
- 823 Valenti, J. (2017), Goes-r series product definition and users' guide, Technical report.
- 824 van der Kamp, D. & McKendry, I. (2010), 'Diurnal and seasonal trends in convective mixed-layer heights esti-
825 mated from two years of continuous ceilometer observations in vancouver, bc', *Boundary-layer meteorology*
826 **137**(3), 459–475.
- 827 Vautard, R., Yiou, P., D'andrea, F., De Noblet, N., Viovy, N., Cassou, C., Polcher, J., Ciais, P., Kageyama,
828 M. & Fan, Y. (2007), 'Summertime european heat and drought waves induced by wintertime mediterranean
829 rainfall deficit', *Geophysical Research Letters* **34**(7).
- 830 Voogt, J. A. & Grimmond, C. (2000), 'Modeling surface sensible heat flux using surface radiative tempera-
831 tures in a simple urban area', *Journal of Applied Meteorology* **39**(10), 1679–1699.
- 832 Voogt, J. & Oke, T. (2003), 'Thermal remote sensing of urban climates', *Remote Sensing of Environment*
833 **86**(3), 370 – 384. Urban Remote Sensing.
834 **URL:** <http://www.sciencedirect.com/science/article/pii/S0034425703000798>
- 835 Wang, C., Myint, S. W., Wang, Z. & Song, J. (2016), 'Spatio-temporal modeling of the urban heat island in
836 the phoenix metropolitan area: Land use change implications', *Remote Sensing* **8**(3), 185.
- 837 World Meteorological Organization, ed. (2008), *Aerodrome reports and forecasts: a users' handbook to the*
838 *codes*, number no. 782 in 'WMO', 5th ed edn, WMO, Geneva, Switzerland. OCLC: ocn300459644.
- 839 Xia, T., Kustas, W. P., Anderson, M. C., Alfieri, J. G., Gao, F., McKee, L., Prueger, J. H., Geli, H. M., Neale,
840 C. M., Sanchez, L. et al. (2016), 'Mapping evapotranspiration with high-resolution aircraft imagery over
841 vineyards using one-and two-source modeling schemes', *Hydrology and Earth System Sciences* **20**(4), 1523–
842 1545.
- 843 Xiaomin, X., Zhen, H. & Jiasong, W. (2006), 'The impact of urban street layout on local atmospheric
844 environment', *Building and Environment* **41**(10), 1352–1363.
- 845 Xu, W., Wooster, M. & Grimmond, C. (2008), 'Modelling of urban sensible heat flux at multiple spatial
846 scales: A demonstration using airborne hyperspectral imagery of shanghai and a temperature–emissivity
847 separation approach', *Remote Sensing of Environment* **112**(9), 3493–3510.
- 848 Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, C., Dewitz, J., Fry, J.,
849 Funk, M., Granneman, B., Liknes, G. C., Rigge, M. & Xian, G. (2018), 'A new generation of the United
850 States National Land Cover Database: Requirements, research priorities, design, and implementation
851 strategies', *ISPRS Journal of Photogrammetry and Remote Sensing* **146**, 108–123.
852 **URL:** <https://linkinghub.elsevier.com/retrieve/pii/S092427161830251X>
- 853 Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M., Vinnikov, K. Y., Sun, D. & Tian, Y.
854 (2011), 'Validation of goes-r satellite land surface temperature algorithm using surfrad ground measure-
855 ments and statistical estimates of error properties', *IEEE Transactions on Geoscience and Remote Sensing*
856 **50**(3), 704–713.

- 857 Zhang, J., Draxl, C., Hopson, T., Delle Monache, L., Vanyvye, E. & Hodge, B.-M. (2015), 'Comparison of
858 numerical weather prediction based deterministic and probabilistic wind resource assessment methods',
859 *Applied Energy* **156**, 528–541.
- 860 Zilitinkevich, S. S. (1995), 'Air pollution theory and simulation', *Air Pollution* **3**, 53–60.