

1 A novel model to estimate sensible heat fluxes in urban areas using  
2 satellite-derived data

3 Gabriel Rios<sup>\*1, 2</sup> and Prathap Ramamurthy<sup>1, 2</sup>

4 <sup>1</sup>Department of Mechanical Engineering, CUNY City College, New York, NY 10031, USA

5 <sup>2</sup>NOAA CREST Center, CUNY City College, New York, NY 10031, USA

6 **Abstract**

7 A novel model for estimating sensible heat flux in urban areas using satellite data is presented here.  
8 Sensible heat flux ( $Q_H$ ) is a primary component of the urban surface energy budget and is critical to  
9 regulating air temperature in cities. The model employs data from the NASA & NOAA GOES-16 geo-  
10 stationary satellite, ground-based observations from NOAA Automated Surface Observation Stations  
11 (ASOS), and land cover data from the National Land Cover Database (NLCD) as inputs for an iter-  
12 ative algorithm that is based on surface-layer similarity theory for turbulence parameterization. The  
13 application of this model specifically to urban areas is enabled by the spatial extent of the GOES-16  
14 satellite and an element roughness height estimation method based on NLCD land cover data. Model  
15 results were independently validated using three flux towers located in different areas of New York City.  
16 Statistics generated over a year-long validation period from June 2019 to May 2020 show a root-mean-  
17 square error (RMSE) of  $47.32 \text{ W}\cdot\text{m}^{-2}$ , a mean bias error (MBE) of  $16.58 \text{ W}\cdot\text{m}^{-2}$ , and an  $R^2$  correlation  
18 value of 0.70. Model results were also compared to results from the urbanized Weather Research and  
19 Forecasting (uWRF) model relative to flux tower observational data to allow for a comparison between  
20 numerical models. The dedicated  $Q_H$  model outperformed the uWRF model relative to observational  
21 data, with an RMSE reduction of  $63.5 \text{ W}\cdot\text{m}^{-2}$ , an MBE reduction of  $17.5 \text{ W}\cdot\text{m}^{-2}$ , and an  $R^2$  increase  
22 of 0.08. Validation results show good agreement between model and observed values and performance  
23 comparison results show an improvement over a current numerical method for estimation of  $Q_H$ , sug-  
24 gesting the use of satellite data as a cost-effective and accessible option for estimating  $Q_H$  in urban areas.

25  
26 *Keywords:* Sensible Heat Flux, Urban, Satellite Remote Sensing, GOES-16, Atmospheric Modeling,  
27 GOES-16, National Land Cover Database

28  
29 \*Corresponding author: Gabriel Rios, Department of Mechanical Engineering, CUNY City College,  
30 New York, NY 10031, USA. E-mail: grios001@citymail.cuny.edu

31  
32 **1 Background and introduction**

33 Sensible heat flux ( $Q_H$ ) is a key component of the Earth's surface energy balance, as it characterizes the  
34 surface-to-atmosphere transport of heat. In urban environments, anthropogenic modification of land cover  
35 reduces water retention capacity, increasing the roles of sensible heat and heat storage ( $Q_S$ ) in the urban  
36 surface energy budget.  $Q_H$  in cities impacts the urban heat island dynamics, hence, it has significant  
37 implications on weather prediction and forecasting, air pollution, and building energy use (Imran et al.,  
38 2018; Schumacher et al., 2019; Vautard et al., 2007).

<sup>39</sup>  $Q_H$  is driven by a number of factors - particularly the temperature difference between the land surface  
<sup>40</sup> temperature (LST) and the air temperature ( $T_{air}$ ) in the lowest levels of the boundary layer. The LST  
<sup>41</sup> has been shown to be higher in urban areas than surrounding suburban/rural areas (Price, 1979), which is  
<sup>42</sup> driven by the high thermal inertia of urban land cover. The increased LST can both increase  $T_{air}$  and the  
<sup>43</sup> temperature difference between the two, resulting in an increased  $Q_H$  relative to surrounding areas (Kato  
<sup>44</sup> and Yamaguchi, 2005).

<sup>45</sup> A challenge in understanding the relationship between land cover, LST,  $T_{air}$  and  $Q_H$  is presented by the  
<sup>46</sup> techniques used for measurement and estimation of  $Q_H$ . This challenge is brought about by a number of  
<sup>47</sup> factors, including (but not limited to):

- <sup>48</sup> • Computationally-expensive numerical models for estimation purposes (Best, 2005; Zhang et al., 2015),  
<sup>49</sup> • The lack of well-established measurement networks in rural and urban areas (Chrysoulakis et al., 2018;  
<sup>50</sup> Voogt and Oke, 2003)

<sup>51</sup> Numerical models are powerful tools that allow for the understanding of atmospheric processes at much  
<sup>52</sup> greater spatial extents than possible by measurement and observation alone. However, these models can often  
<sup>53</sup> feature significant inaccuracies in areas with high spatial heterogeneity, such as urban areas, due to low grid  
<sup>54</sup> domain resolutions relative to the size and spacing of elements in heterogeneous environments (e.g. buildings,  
<sup>55</sup> roads, scattered green space and vegetative cover) (Chen et al., 2011; Hong and Dudhia, 2012; Leroyer et al.,  
<sup>56</sup> 2014). Accordingly, model accuracy can only be improved upon by significantly increasing model resolution  
<sup>57</sup> to resolve these spatial issues, which risks high time and resource consumption. Meanwhile, measurement  
<sup>58</sup> networks are vital since observational data is an essential source of validation data for numerical models to  
<sup>59</sup> ensure their performance. However, accurate measurement of parameters such as  $Q_H$  is challenged by the lack  
<sup>60</sup> of measurement networks with sufficient spatial resolution that can serve as databases for validation efforts.  
<sup>61</sup> Moreover, this challenge is exacerbated in urban areas due to the aforementioned land cover heterogeneity,  
<sup>62</sup> which is critical in determining  $Q_H$  in localized areas (Feddema et al., 2005; Wang et al., 2016). To address  
<sup>63</sup> this, remote sensing technologies have been increasingly used to devise estimation methods for  $Q_H$ .

<sup>64</sup> Several studies in the reviewed literature have estimated heat fluxes using remote sensing methods in  
<sup>65</sup> rural areas using a variety of methods (Cammalleri et al., 2012; Kim and Kwon, 2019; Miglietta et al., 2009;  
<sup>66</sup> Mkhwanazi et al., 2012; Ortega-Farías et al., 2016). Miglietta et al. (2009) describes an estimation method  
<sup>67</sup> using Meteosat land surface temperature and radiation products, as well as aircraft-mounted sensors, to  
<sup>68</sup> evaluate fluxes over forested areas and cropland between May and June 2005. In Cammalleri et al. (2012),  
<sup>69</sup> aircraft-mounted multispectral and thermal cameras were used in conjunction with meteorological data to  
<sup>70</sup> estimate  $Q_H$  over 7 days within a 4 month period, with a study area covered by cropland, fallow soil, and  
<sup>71</sup> bare soil. Mkhwanazi et al. (2012) used Landsat 5 imagery with a bulk parameterization method to evaluate  
<sup>72</sup> fluxes over an alfalfa field in rural Colorado. Kim and Kwon (2019) and Ortega-Farías et al. (2016) showed  
<sup>73</sup> promising results using unmanned aerial vehicles (UAVs) to estimate  $Q_H$  over a variety of land cover types  
<sup>74</sup> in rural areas throughout a range of synoptic meteorological conditions, with good agreement between UAV-  
<sup>75</sup> based estimation results and instrument-based surface observations. These studies all demonstrate great  
<sup>76</sup> potential for using remote sensing for estimation of surface fluxes, although their temporal frequency and  
<sup>77</sup> focus on homogeneous land cover types hinders their applicability to urban areas.

<sup>78</sup> Fewer studies have been performed to estimate  $Q_H$  using remote sensing methods in urban areas, which  
<sup>79</sup> feature far greater land cover heterogeneity (Feigenwinter et al., 2018; Liu et al., 2012; Voogt and Grimmond,  
<sup>80</sup> 2000; Xu et al., 2008). Two studies (Voogt and Grimmond, 2000; Xu et al., 2008) used helicopter-mounted

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

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

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-

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

## <sup>132</sup> 2 Methodology and data

### <sup>133</sup> 2.1 Study area

<sup>134</sup> The study area used is New York City (see Figure 1), which is the largest city in the United States by  
<sup>135</sup> population, with approximately 8.3 million people as of 2019 ([US Census Bureau, 2019](#)) and is among the  
<sup>136</sup> most densely-populated cities in the United States. The city is composed of 5 boroughs: the Bronx, Brooklyn,  
<sup>137</sup> Manhattan, Queens, and Staten Island. The Bronx is made up largely of low- to mid-rise residential and  
<sup>138</sup> commercial buildings, with decreasing building density and height towards the northern end of the borough.  
<sup>139</sup> Brooklyn is largely composed of low- to mid-rise buildings, with a concentration of high-rise buildings on  
<sup>140</sup> the East River, while the southern and eastern portions feature larger proportions of lower-density suburban  
<sup>141</sup> residential areas. Manhattan is primarily composed of residential and commercial buildings, with mid- to  
<sup>142</sup> high-rise buildings spanning the entirety of the borough (with the exception of Central Park, which is a  
<sup>143</sup> mixture of open fields, open water, and deciduous & evergreen forests). Queens is similar in composition to  
<sup>144</sup> Brooklyn, with the exception of larger spans of lower-density development towards the eastern half of the  
<sup>145</sup> borough. Staten Island features significantly lower building densities and heights, with expansive wetland  
<sup>146</sup> and grassy areas on its western edges and a large forested area in the central area of the borough. The  
<sup>147</sup> complex urban landscape, coupled with an array of urban flux towers and weather observation stations  
<sup>148</sup> within the city, make the city an ideal candidate for implementing and validating the urban-focused  $Q_H$   
<sup>149</sup> model.

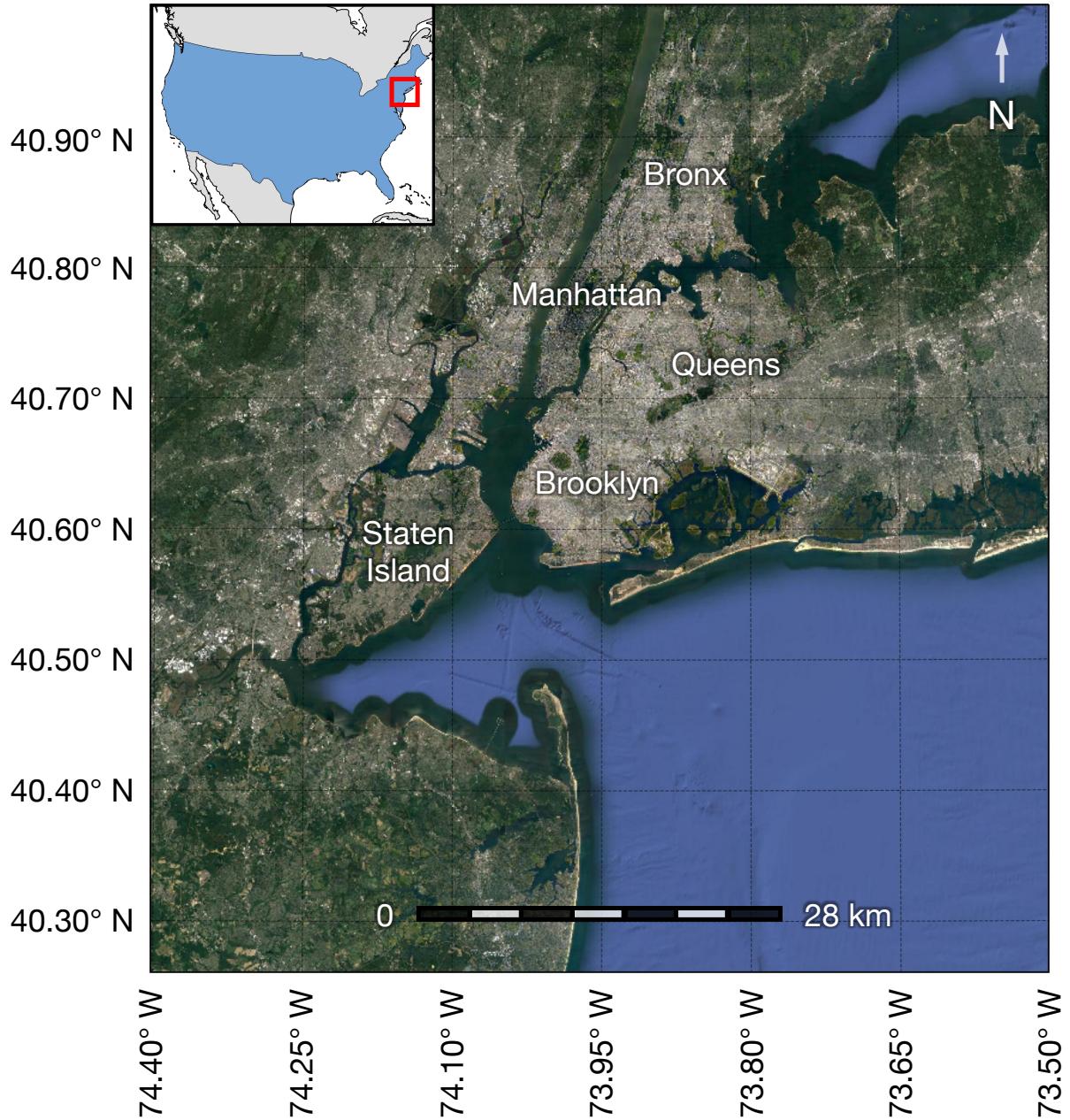


Figure 1: Satellite view of the New York City metropolitan area. New York City, which is composed of 5 boroughs (labeled), is the most-heavily urbanized portion of the metropolitan area, while lower density suburbs and woodlands compose the outer portions of the metropolitan area.

## 150 2.2 Model overview

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

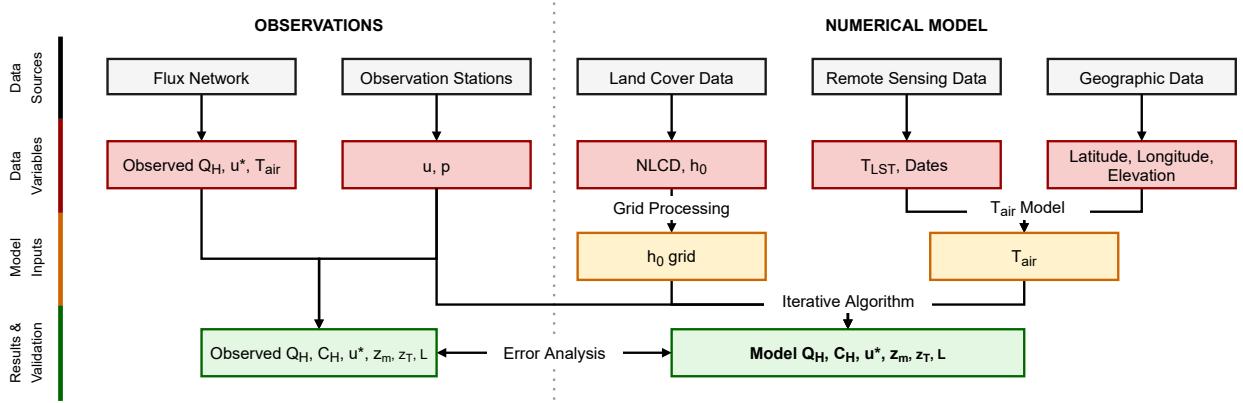


Figure 2: Process flowchart for the sensible heat flux model. Observational data was used for validation of the satellite model as well as inputs to the iterative algorithm. The numerical model used remotely-sensed data from the GOES-16 satellite, as well as ancillary datasets for land cover and geographic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

### 159 2.2.1 Sensible heat flux iterative algorithm

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

169  $Q_H$  is directly calculated using Equation 1 (Pond et al., 1974):

$$170 Q_H = \rho c_p C_H u (\theta_0 - \theta_r) \quad (1)$$

171 In Equation 1,  $\rho$  is air density calculated as a function of air pressure ( $p$ ) and air temperature at the  
 172 reference height of 2 m above ground level (AGL) ( $T_{air}$ ),  $c_p$  is the average specific heat of air (1006 J  
 $kg^{-1} K^{-1}$ ) across the range of air temperatures and pressures observed,  $C_H$  is a bulk heat transfer coefficient,  
 173  $u$  is the observed wind speed at a height of 10 m AGL, and  $\theta_0$  and  $\theta_r$  are potential temperatures at the  
 174 surface and at 2 m AGL, respectively. Both  $\theta_0$  and  $\theta_r$  are derived from remotely-sensed data -  $\theta_0$  is derived  
 175 from remotely-sensed land surface temperature ( $T_{LST}$ ) and  $\theta_r$  is derived from a model based on  $T_{LST}$  and  
 176 several other remotely-sensed parameters (Hrisko et al., 2020). See Section 2.3 for a detailed discussion  
 177 regarding the derivation of these parameters.

178  $C_H$  is calculated using Equation 2 (Monin and Obukhov, 1954):

$$179 C_H = \frac{\kappa^2}{[\ln \frac{z_r}{z_m} - \psi_m \zeta][\ln \frac{z_r}{z_T} - \psi_h \zeta]} \quad (2)$$

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

183 The momentum and thermal roughness heights,  $z_m$  and  $z_T$ , are calculated using the Raupach [Equation  
 184 3] and Zilitinkevich [Equation 5] methods, respectively. The Raupach method (Raupach, 1994) for defining

the momentum roughness height has been found useful in areas with heterogeneous land cover, as it can be calculated as a function of localized parameters and atmospheric conditions, specifically element roughness height  $h_0$  and local friction velocity  $u^*$  (Voogt and Grimmond, 2000). The methodology for the estimation of  $h_0$  is discussed in detail in 2.2.2. The Zilitinkevich method has been shown to be an effective approximation method for  $z_T$  in areas with tall canopies, such as those present in urban areas, while enabling  $z_T$  to be calculated as a function of local parameters (Chen and Zhang, 2009; Zilitinkevich, 1995), as described in Li and Bou-Zeid (2014).

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

where:

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

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

where:

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

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

The friction velocity  $u^*$  is expressed by Equation 5 (Monin and Obukhov, 1954):

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

The Obukhov length  $L$  is expressed by Equation 6 (Monin and Obukhov, 1954):

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

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

## 2.2.2 Roughness height estimation

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

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

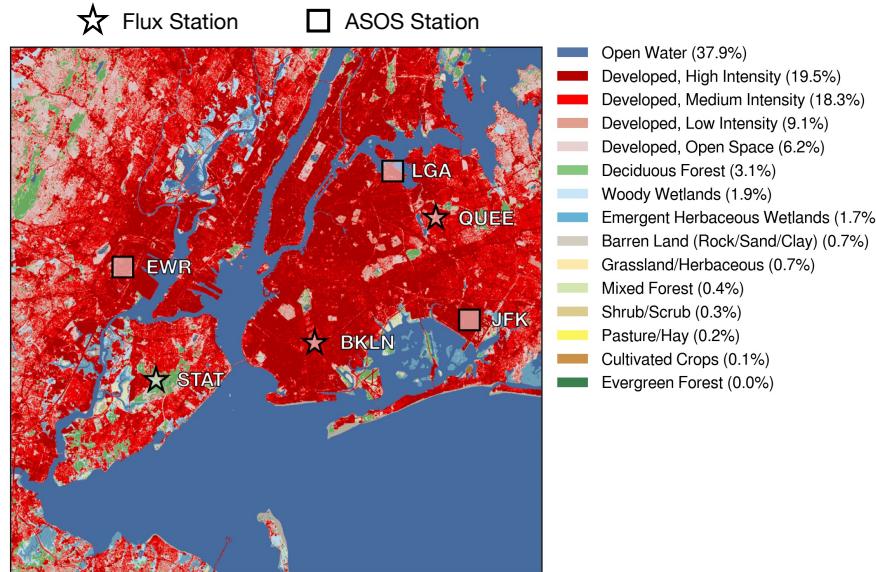


Figure 3: Land cover map of the New York City metropolitan area, per the 2016 National Land Cover Database (Yang et al., 2018). The legend shows land cover types and the percentage of the study area occupied by each land cover type. Land cover data is shown at a 30 m resolution. Note that flux observation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

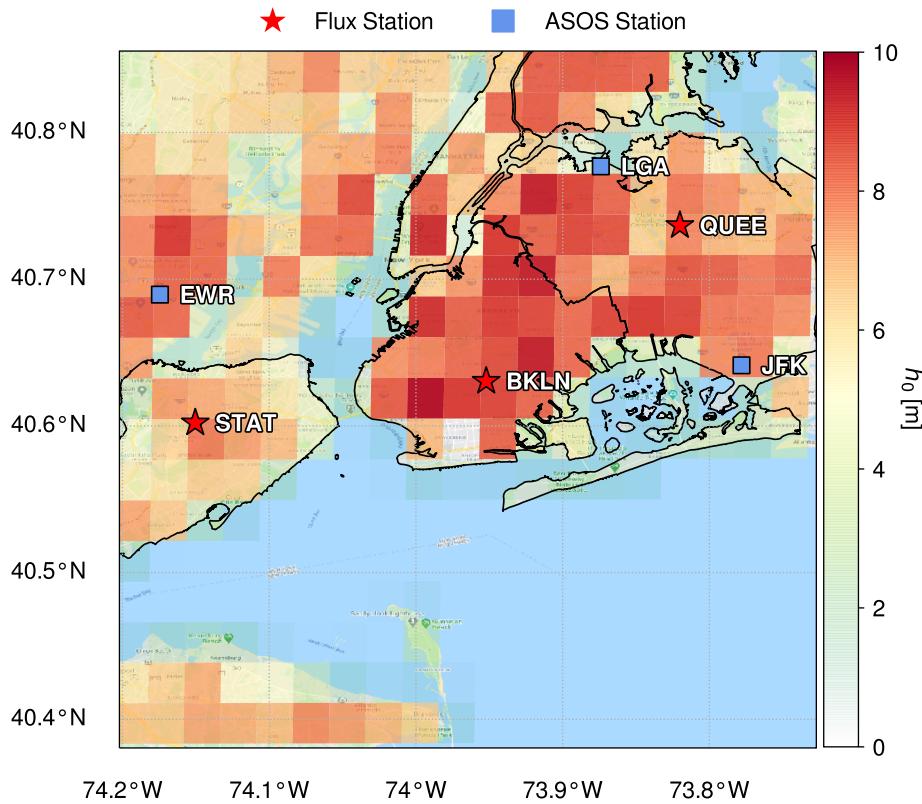


Figure 4: Gridded map of element roughness heights across the New York City metropolitan area. Note that flux observation towers and ground weather (ASOS) stations are labeled accordingly.

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

## 224 **2.3 GOES-R land surface temperature (LST) product**

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

229 A number of products derived from satellite radiance data are offered by the satellite, including a Land  
230 Surface Temperature (LST) product, from which  $T_{LST}$  (and through derivation,  $\theta_0$ ) is obtained. It is  
231 available for public use at a moderate spatial resolution of 2 x 2 km and a high temporal resolution of  
232 5 minutes ([NOAA/NESDIS/STAR, 2016](#)). The LST is calculated using GOES-16 infrared bands 14 and  
233 15. This product features a desirable balance of spatiotemporal resolution and high accuracy (<2.50 K)  
234 ([Valenti, 2017](#)), making it a critical input to the model. The LST product is available in a gridded netCDF  
235 (.nc) format, with data corresponding to latitude and longitude mapped over the spatial extent of satellite  
236 observations. The data is filtered based on image quality, which is largely dependent on sky conditions (i.e.  
237 cloud cover). Therefore, dates within the study timeframe with clear skies or few clouds (<25% sky cover,  
238 per METAR ([World Meteorological Organization, 2008](#)) were selected to ensure high-quality LST data as  
239 input to the model. The data used for the model was limited to a 0.50 degree extent encompassing the most  
240 heavily-urbanized portion of the New York City metropolitan area, extending from approximately (40.8805  
241 N, 74.2021 W) to (40.3805 N, 73.7021 W), spanning a land area of approximately 800  $km^2$ . On a 2 km x 2  
242 km grid, this represents approximately 200 pixels over which data was obtained for the metropolitan area.

243 Another major component of the model is an urban air temperature model that takes GOES-16 LST  
244 product data as an input and uses a diurnal regressive algorithm to calculate air temperature at a height  
245 of 2 m AGL ([Hrisko et al., 2020](#)), from which  $T_{air}$  (and through derivation,  $\theta_r$ ) is obtained. The model  
246 has been shown to estimate air temperatures in areas featuring a range of land cover classes with high  
247 accuracy, specifically in urban areas (RMSE of 2.60 K relative to ground station observations), and is spatially  
248 representative when compared to ASOS observation data (see the next section for more information). Inputs  
249 to the model are LST, elevation, NLCD land cover class, and coordinates. The model output is a gridded  
250 dataset with temperature values. For reference, the data is produced on a 2 x 2 km grid to match the gridded  
251 data format of the GOES-16 LST product.

## 252 **2.4 Ground station observation data**

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

257 Each ASOS station collects a wealth of information regarding weather conditions most relevant for  
258 aviation purposes, including air temperature, dew point temperature, air pressure, wind speed and direction,  
259 and sky cover. Each station generally records data at a frequency of 5 minutes, providing reasonable spatial  
260 and excellent temporal frequencies for model data input. Four stations are located within the spatial domain  
261 evaluated in this study (see Figure 4 for reference): John F. Kennedy International Airport (JFK) (40.6413°  
262 N, 73.7781° W), LaGuardia Airport (LGA) (40.7769° N, 73.8740° W), Newark Liberty International Airport

263 (EWR) ( $40.6895^{\circ}$  N,  $74.1745^{\circ}$  W), Central Park ( $40.7790^{\circ}$  N,  $73.9693^{\circ}$  W). The ASOS stations closest to each  
264 observation site are selected for data collection. Specifically, these ASOS stations are JFK (corresponding  
265 to Brooklyn), LGA (Queens), and EWR (Staten Island).

266 The model was validated using the New York State (NYS) Mesonet observation network ([Mesonet, 2020](#)).  
267 The network features 17 flux stations throughout the state of New York, with 3 stations located within New  
268 York City - one each in the boroughs of Brooklyn (BKLN) ( $40.6318^{\circ}$  N,  $73.9537^{\circ}$  W), Queens (QUEE)  
269 ( $40.7343^{\circ}$  N,  $73.8158^{\circ}$  W), and Staten Island (STAT) ( $40.6040^{\circ}$  N,  $74.1485^{\circ}$  W). The flux network stations  
270 record parameters relevant to the surface energy budget, including net radiation  $R_N$ , surface latent heat flux  
271  $Q_L$ , and surface sensible heat flux  $Q_H$ . Each flux station is equipped with a net radiometer (manufactured  
272 by Kipp & Zonen CNR4), ground heat flux plates (Hukseflux), and a closed-path eddy covariance system  
273 (CPEC200, Campbell Scientific, Inc) consisting of a sonic anemometer and gas analyzer. The net radiometer  
274 and eddy covariance system are installed atop 10 m towers. The towers are mounted on buildings with heights  
275 of 23.20 m at the Brooklyn station, 44.60 m at the Queens station, and 23.10 m at the Staten Island station  
276 (all heights above ground level). For reference, average heights of surrounding buildings are 10.70 m in  
277 Brooklyn, 10.70 m in Queens, and 6.00 m in Staten Island, per New York City zoning areas ([Department of](#)  
278 [City Planning, 2021](#)). Station flux measurements are reported every 30 minutes. The eddy covariance system  
279 was used to measure  $Q_H$  for the duration of the validation period.

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

## 297 2.5 Model performance against ground stations

298 The study period for the model spanned from 1 June 2019 (day of year 152) to 31 May 2020 (day of year 152).  
299 Approximately 44 days over the course of the study period were selected for model validation. The selection  
300 criteria included sky cover classified as "CLR" (clear sky) or "FEW" (few clouds) at each ASOS observation  
301 station continuously over a 24-hour period and operational flux network status. For validation purposes,  
302 model runs were initially performed at the latitude and longitude corresponding to each flux station. The  
303 corresponding GOES-16 grid location, or pixel, was used for the LST and  $T_{air}$ . The closest ASOS station  
304 was used to provide inputs of  $p$  and  $u$  (the distance between the study location and the corresponding ASOS

station is a potential source of error that is discussed further). In total, 3 pixels were analyzed for validation purposes at hourly intervals over the selected days, resulting in a total of approximately 3,200 data points.

## 2.6 Urbanized Weather Research and Forecasting (uWRF) model

The WRF model (Skamarock et al., 2008) with an urbanization option (uWRF) is used in this study as a model-based data set against which the performance of the dedicated  $Q_H$  model can be compared. This supplements the comparison against an observation-based dataset provided by the Mesonet flux towers. The urbanization option features parameterizations specific to urban areas for better representation of boundary layer processes in cities (Gutierrez, Gonzalez, Martilli, Bornstein and Arend, 2015a; Gutierrez, Gonzalez, Martilli and Bornstein, 2015). This configuration of the WRF model has been used in numerous previous studies to study atmospheric processes in urban areas (Chen et al., 2011; Gamarro et al., 2019; Gutierrez, Gonzalez, Martilli, Bornstein and Arend, 2015b; Hrisko et al., 2021; Ortiz et al., 2017).

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

## 2.7 Model performance evaluation

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

Each measure is defined as follows:

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

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

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

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

### 3 Results

#### 3.1 Overall results

In the timeframe studied, the dedicated  $Q_H$  model featured a RMSE of  $47.32 \text{ W m}^{-2}$ , a bias of  $16.58 \text{ W m}^{-2}$ , an NSC value of 0.54, and a  $R^2$  value of 0.70. The overall results are visualized in Figure 5. Statistical results of the study period, as defined in Equations 10-13, are shown in Table 1 decomposed by location. The performance statistics suggest the model displayed reasonable agreement with the ground observations and performed satisfactorily, per the definition provided for the Nash-Sutcliffe coefficient by Moriasi et al. (2007). Model performance as a function of spatial and temporal variability will be discussed in this section. Temporal variability will be discussed on two distinct timescales, seasonal and daily, to improve understanding of model behavior and differences driven by changes in time on large and small temporal scales. Nonetheless, the model showed considerable error from the ground 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  $\text{W m}^{-2}$ .

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

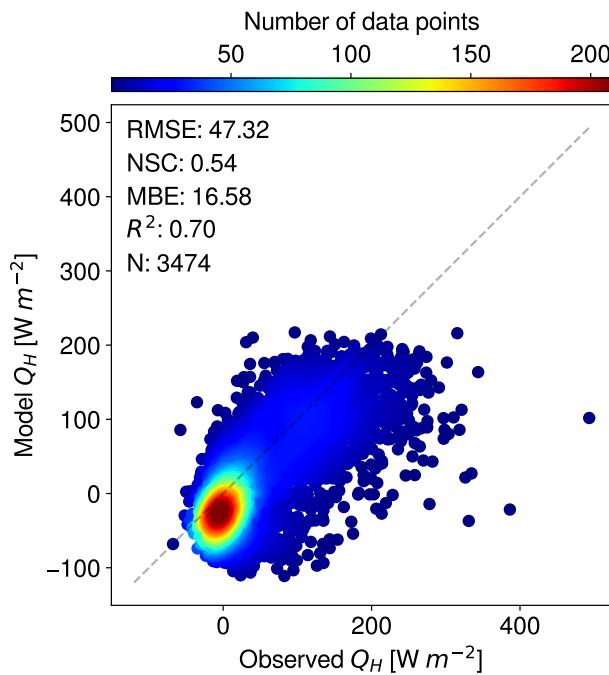


Figure 5: Comparison of model and observed  $Q_H$  for all stations from 1 June 2019 to 31 May 2020.

#### 3.1.1 Diurnal variability

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 (6:00 - 16:00 local standard time), with good agreement between the model and the observations across all stations ( $\text{RMSE} = 48.68 \text{ W m}^{-2}$ ,  $\text{MBE} = 6.99 \text{ W m}^{-2}$ ,  $\text{NSC} = 0.58$ ,  $R^2 = 0.61$ ). However, the model did not perform as well in the early nighttime hours due to an increased bias error (16:00 - 0:00 local standard time) ( $\text{RMSE} = 44.98 \text{ W m}^{-2}$ ,  $\text{MBE} = 21.39 \text{ W m}^{-2}$ ,  $\text{NSC} = 0.40$ ,  $R^2 = 0.66$  across all stations) and poorly during the pre-dawn hours due to a lack of correlation between modeled and observed data (0:00 - 6:00 local standard time) ( $\text{RMSE} = 48.04 \text{ W m}^{-2}$ ,  $\text{MBE} = 26.17 \text{ W m}^{-2}$ ,  $\text{NSC} = -1.65$ ,  $R^2 = -0.30$  across all stations). As shown in Figure 7, these diurnal performance

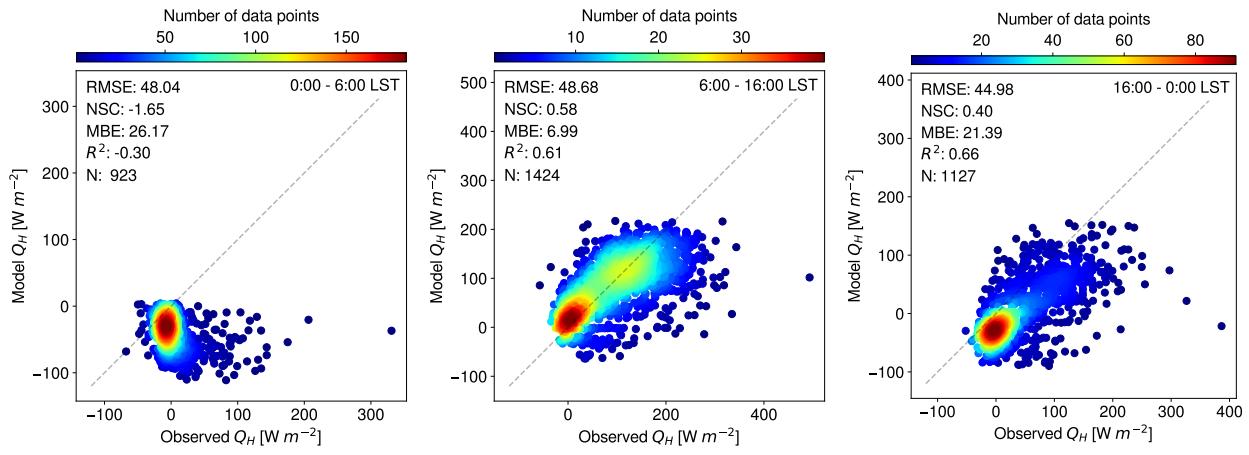


Figure 6: Comparison of model and observed  $Q_H$  for all stations divided into distinct periods of the day. (from left to right) Early morning (pre-sunrise), daytime, evening and late night (post-sunset).

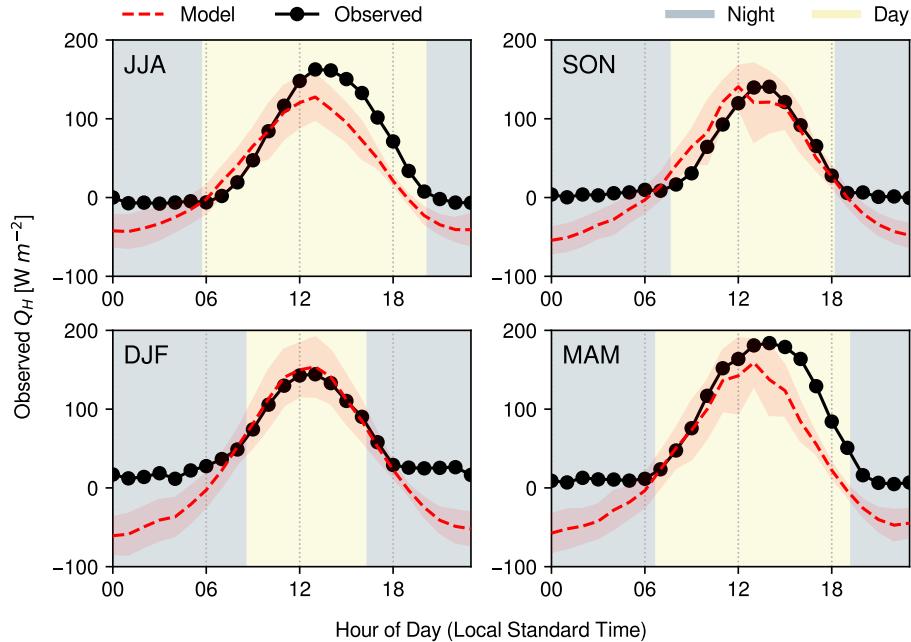


Figure 7: Seasonal averages of daily  $Q_H$ , averaged over all stations. For reference, the acronyms reflect their represented months (for example, 'JJA' corresponds to June, July, and August). All times are in local standard time (LST) and background shading corresponds approximately to average day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

361 trends are further supported by seasonal diurnal averages taken for each season over all stations. The  
 362 model performs considerably well during daytime hours, especially in the summer (JJA) and autumn (SON)  
 363 months. However, the model consistently underestimates  $Q_H$  during nighttime hours - especially so in the  
 364 winter (DJF) months. The performance error during the pre-dawn hours is likely due to estimation biases  
 365 in the remote sensing methods used and will be discussed later in the paper.

### 366 3.1.2 Seasonal variability

367 The model exhibited variability in performance relative to the season during which runs were performed.  
 368 As shown in Figure 8, the model appears to perform better in warmer months [summer (JJA) and spring  
 369 (MAM)] than in cooler months. This is evident in the  $R^2$  values of each season (0.80 and 0.79 for JJA and  
 370 MAM, respectively) as compared to cooler months [autumn (SON) and winter (DJF)], which feature lower  $R^2$   
 371 values (0.72 and 0.56, respectively). Model error was more prevalent in the winter (DJF) and spring (MAM)  
 372 months relative to the summer (JJA) and fall (SON) months largely due to nocturnal model underprediction,  
 373 as shown in Figure 9. Analysis regarding seasonal variability in model results and performance is discussed  
 374 further in Section 4.3.

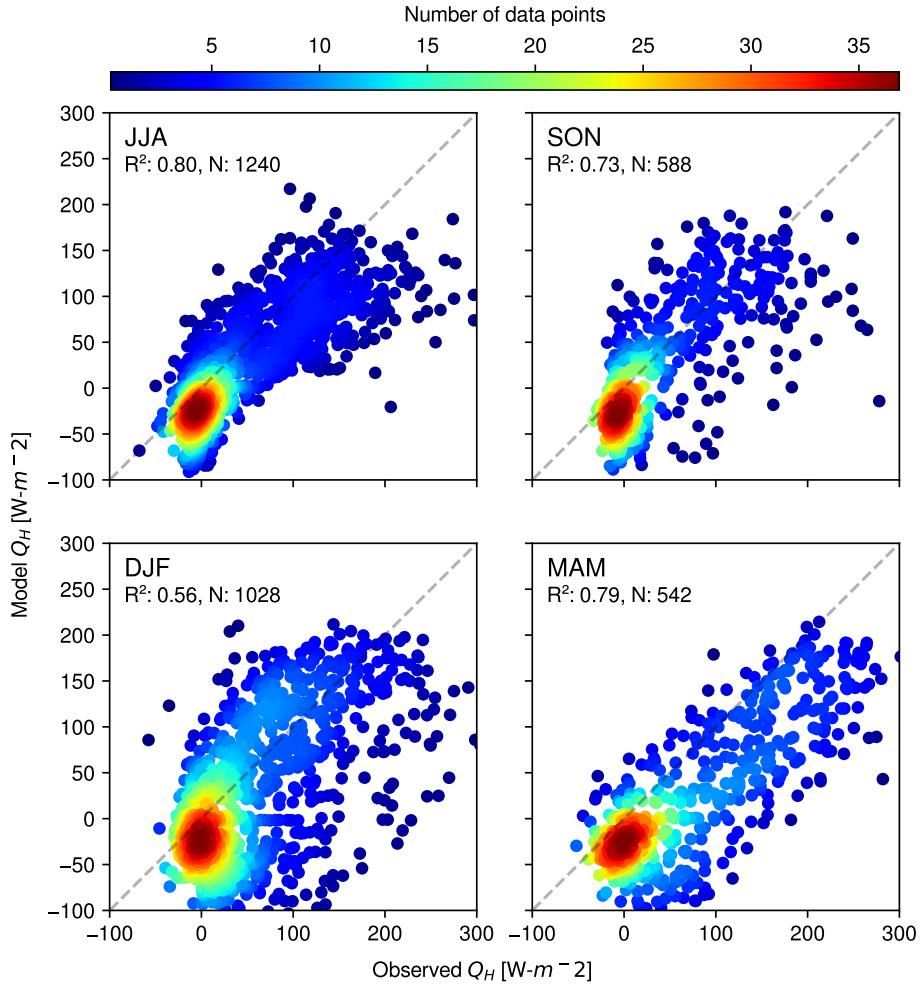


Figure 8: Comparison of model and observed  $Q_H$  divided into seasons. Note that the acronyms correspond to months in each season (for example, 'JJA' corresponds to June, July and August).

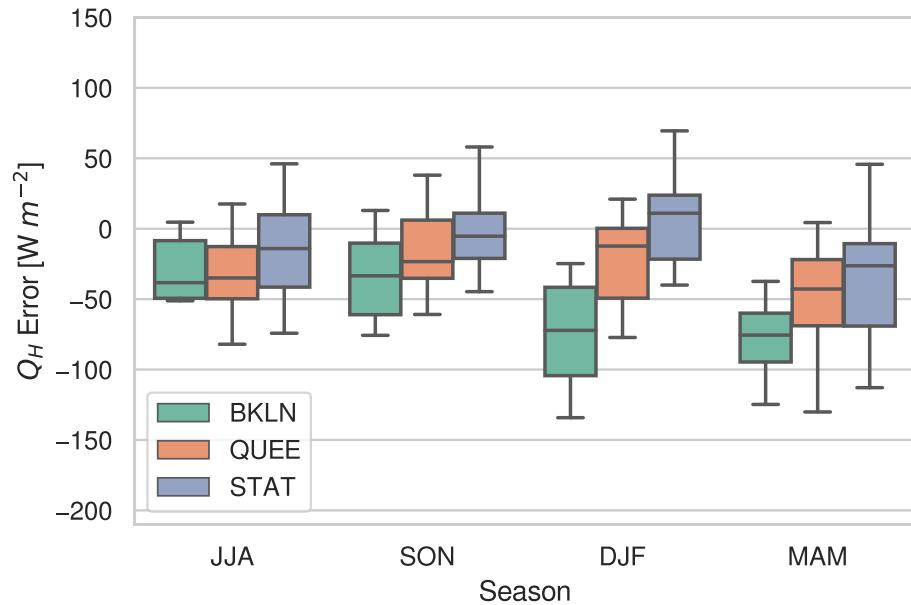


Figure 9:  $Q_H$  error between model results and observational data, averaged per season per location. For reference, the acronyms reflect their represented months (for example, 'JJA' corresponds to June, July, and August).

### 375 3.1.3 Geospatial variability

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

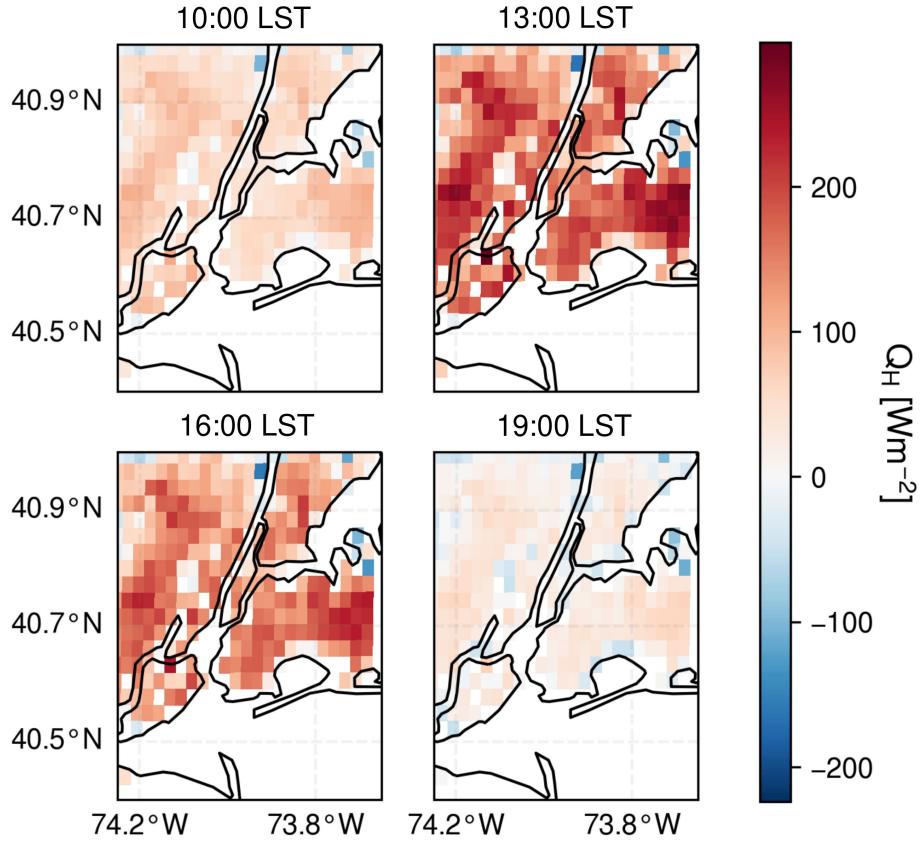


Figure 10:  $Q_H$  ( $\text{W m}^{-2}$ ) in New York City on 24 October 2019 shown at 4 different times during the daytime. Note that all times are in local standard time (LST).

nearing  $300 \text{ W m}^{-2}$  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 16:00 local standard 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 19:00 local standard 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  $Q_H$  (sensible heat flux directed towards the surface) 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 and 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.

### 3.2 Model performance against uWRF

As noted in Section 2.6, the uWRF model was used as an model-based data set against which performance of the dedicated  $Q_H$  model could be compared. Additionally, comparison with the uWRF model provides the ability to validate the dedicated  $Q_H$  model over a continuous spatial extent not afforded by the single-point observation stations. The model was run on 4 separate days, for 24 hours each: 24 October 2019 (SON),

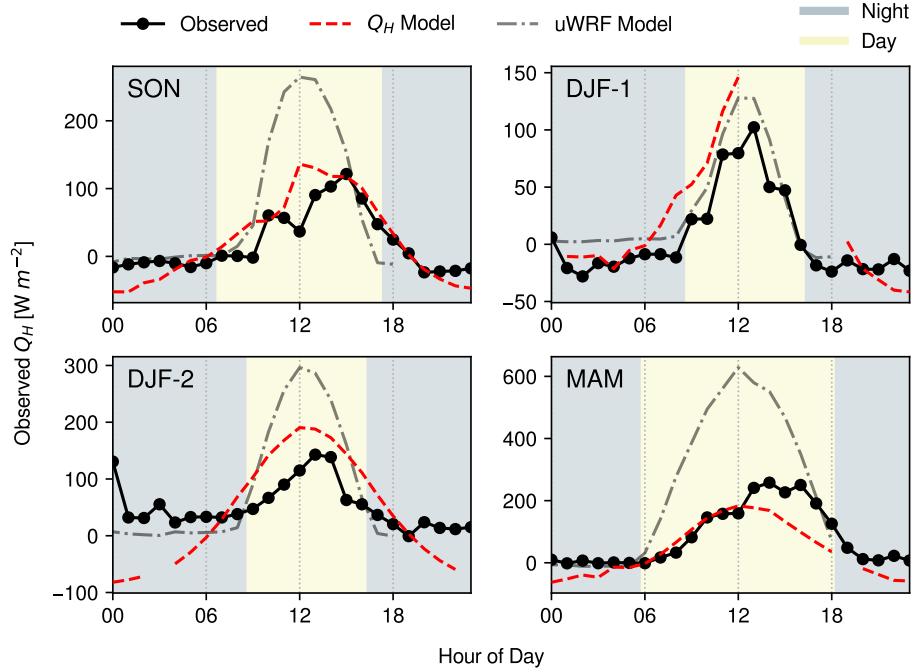


Figure 11: Observed and modeled (uWRF and dedicated)  $Q_H$  at days selected for uWRF study at the Queens (QUEE) station. Note that 'SON' represents the date 24 October 2019, 'DJF-1' represents 23 December 2019, 'DJF-2' represents 20 January 2020, and 'MAM' represents 12 May 2020.

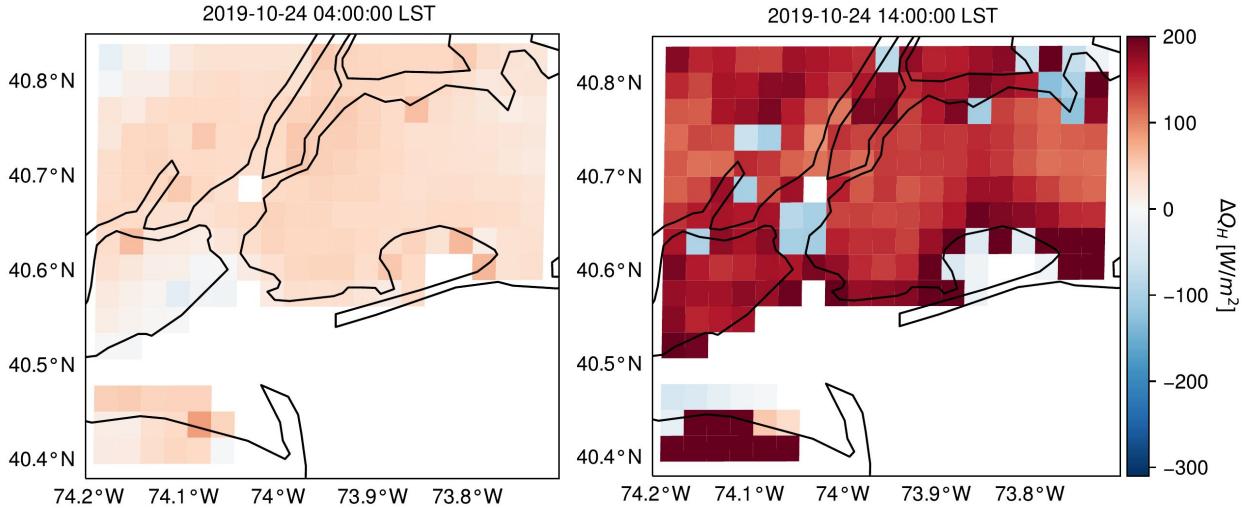


Figure 12: Gridded maps showing error between dedicated  $Q_H$  and uWRF models, October 24, 2019 at 4:00 local standard time (left) and 14:00 local standard time (right). Light blue pixels correspond to areas with mostly covered in water/marsh.

404 23 December 2019 (DJF-1), 20 January 2020 (DJF-2), and 12 May 2020 (MAM). The spatial pixel nearest  
 405 to each ground station was used for comparison. Each ground station used in the  $Q_H$  to ground station  
 406 validation study (BKLN, QUEE, STAT) was selected to produce the comparison.

407 Over the days analyzed in the study, the RMSE between the uWRF model and ground observation  
 408 stations was  $108.07 \text{ W m}^{-2}$ , with a MBE of  $-30.11 \text{ W m}^{-2}$ , a NSC of  $-1.47$ , and a  $R^2$  of  $0.63$ . The performance  
 409 statistics show considerably poorer performance than the dedicated  $Q_H$  model, relative to observed data.  
 410 The uWRF model consistently overpredicted daytime  $Q_H$  (see Figures 11 and 12), although it predicted  
 411 nocturnal  $Q_H$  more accurately than the dedicated  $Q_H$  model. A notable example of overprediction is evident  
 412 in the MAM model run (12 May 2020), where a  $Q_H$  value of  $600 \text{ W m}^{-2}$  was simulated. This significant  
 413 overprediction is primarily a result of overestimated surface wind speeds by the uWRF model (Bauer, 2020).

## 4 Discussion

### 4.1 Model sensitivity analysis

An analysis of model sensitivity to input parameters was performed to determine the response of the model to inputs. 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 by independently perturbing the value of each parameter above and below the original value by a defined quantity, such that the effect of each parameter could be observed individually. Model sensitivity is defined as the percentage error of  $Q_H$  between the model runs with modified input parameter values and unmodified input values. Parameter perturbation values are shown in Table 2. Results from the sensitivity analysis are shown in Table 2 and Figure 13 categorized by location and atmospheric stability,  $\zeta$ . For reference, stability values are categorized into 'unstable', 'neutral', and 'stable', corresponding to values of  $\zeta < -0.25$ ,  $-0.25 \leq \zeta < 0.25$ , and  $0.25 \leq \zeta$ , respectively.

Results show high model sensitivity to  $T_{LST}$ ,  $T_{air}$ , with lesser but significant sensitivity to  $u_{air}$  (wind speed), and minor sensitivity to  $h_0$  (element roughness height) and  $z_r$  (reference height). This order of sensitivity to perturbed parameters is similar to the sensitivity analysis results presented in a study by Feigenwinter et al. (2018). The high sensitivity of flux estimation to temperature is observed in the literature (Brenner et al., 2017; Cammalleri et al., 2012; Xia et al., 2016) and may also be a function of the 2-km spatial resolution of the surface and air temperature gridded data, as shown by results from both Mott et al. (2015) and Xu et al. (2008), where an increase in grid resolution significantly altered model results. With regards to location, model sensitivity to all parameters is similar across all 3 validation locations, with the exception of higher sensitivity to  $T_{LST}$  and  $T_{air}$  at Staten Island. A potential cause for this additional sensitivity is the mixture of land cover types that have very different properties within the same GOES-16 satellite pixel, namely deciduous forest (NLCD class 41) and medium-density developed land (NLCD class 23) (Yang et al., 2018). With regards to atmospheric stability, the model was most sensitive to temperature perturbations during periods of near-neutral conditions, with moderate sensitivity to wind speed perturbations in non-neutral regimes. It is worth noting that the model is less sensitive to perturbations over all parameters during periods of instability, likely due to enhanced mixing and homogenization of properties 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	$\pm 0.5$ K	-13.23	13.21
	STAT		-18.95	18.89
$T_{LST}$	BKLN		-13.30	13.30
	QUEE	$\pm 0.5$ K	-13.15	13.11
	STAT		-19.09	19.10
$h_0$	BKLN		-4.47	4.56
	QUEE	$\pm 0.5$ m	-3.48	3.48
	STAT		-5.04	5.10
$u_{air}$	BKLN		-11.74	9.30
	QUEE	$1\text{ ms}^{-1}$	-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

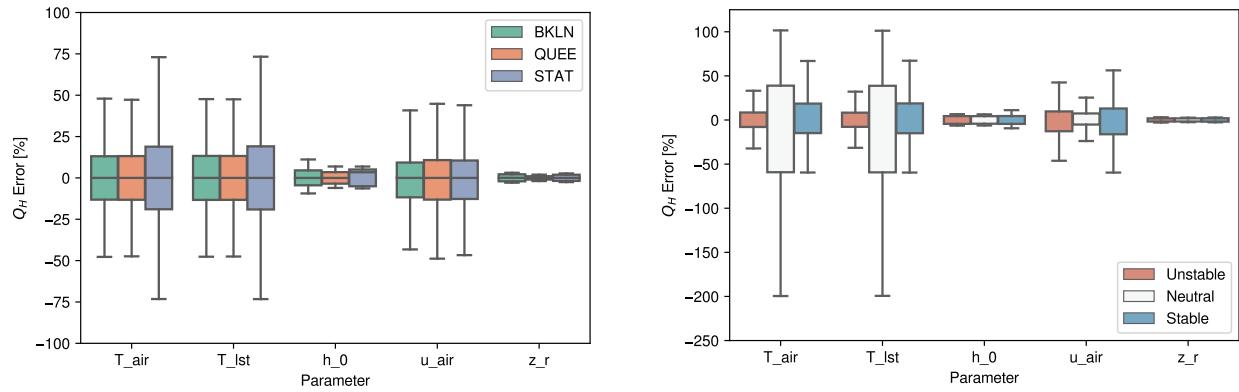


Figure 13: Left: Model sensitivity to each parameter represented by error from baseline values categorized by location (left) and atmospheric stability,  $\zeta$  (right).

## 4.2 Comparison with other studies

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

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

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 products, as well as aircraft-mounted sensors, to evaluate fluxes over forested areas and cropland between May and June 2005. Using eddy covariance methods to observe flux values, the study reported reasonable agreement between estimated and observed temperature and net radiation values, although  $Q_H$  overestimation ranged up to 30% over the study period. [Mkhwanazi et al. \(2012\)](#) used Landsat 5 imagery with a bulk parameterization method to evaluate fluxes over an alfalfa field in rural Colorado. Despite good correlation ( $R^2 = 0.80$ ) and moderate errors (RMSE =  $59.60 \text{ W}\cdot\text{m}^{-2}$ , 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

475 8 imagery over an urban area (Basel, Switzerland) over 22 days between 2013 and 2015 with 3 flux towers  
476 used as validation. This study provided the most comparable estimation method of fluxes in an urban areas  
477 to the knowledge of the authors, with specific information regarding urban land cover types and similar  
478 validation methods. The study found good agreement between estimated and observed values, of  $Q_H$ , with  
479 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  
480 model described herein.

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

### 491 4.3 Model variability & performance

492 Model performance will be discussed in this section with regards to diurnal, seasonal, and geospatial char-  
493 acteristics.

494 The model performs much better during the day, as shown in Figure 3.1.1. The difference in performance  
495 could be related to the prevailing atmospheric conditions; daytime transport of heat is facilitated by thermals  
496 (buoyancy-generated turbulence) that are highly efficient in mixing heat and other scalars. In urban areas,  
497 the largest thermals could be of the size of the boundary layer, on the order of 1-2 km, which is close to  
498 the spatial resolution of GOES-16 pixels. In stark contrast, during the nighttime hours, when the urban  
499 atmosphere is less convectively active, the transport is mainly dominated by mechanical turbulence through  
500 wind shear, which is highly localized. Additionally, the wind field used in the model to derive  $u^*$  is obtained  
501 from an ASOS station that might be unrepresentative of the nearest GOES-16 LST pixel. This error source  
502 is discussed further in the next section.

503 The model performed best in the summer (JJA) and the worst in the winter (DJF). As referenced in  
504 Section 3.1.3, daytime model overprediction and nocturnal model underprediction was highest in winter  
505 (DJF) and spring (MAM) months. Model error in winter and spring months may be explained by a couple  
506 of different factors. One potential cause is the increased frequency of midlatitude cyclones affecting the  
507 northeastern United States in winter and spring months as compared to summer months (Gedzelman et al.,  
508 2003). The meteorological impacts of these events include increased wind speeds, a shift from westerly to  
509 northerly surface winds, and advected air masses from nearby bodies of water (namely, Long Island Sound  
510 and the Atlantic Ocean). These events may exacerbate disparities in wind speed differences between the  
511 source of wind measurement data, which are all within very close proximity to the coast (JFK is directly  
512 on Jamaica Bay and is 3.5 km from the Atlantic Ocean, LGA is directly on Long Island Sound, EWR  
513 borders Newark Bay and is 8 km from Upper New York Bay), and the locations where flux estimates are  
514 made, which are further removed from the coast (BKLN is 5 km from Jamaica Bay, QUEE is 3 km from  
515 Long Island Sound, STAT is 6 km from Lower New York Bay). Additionally, the advected cold air masses  
516 during these events tend to cause temperature inversions that may increase errors in estimating  $T_{LST}$  and

517  $T_{air}$  by remote sensing methods (Tang et al., 2016). This disparity in wind speeds may cause higher  $u^*$   
518 values, resulting in estimation errors of  $Q_H$  since it is proportional to  $u^*$ . Another potential cause of model  
519 error is the seasonal variability in the mean mixed layer height of the boundary layer. Numerous studies  
520 of boundary layer structure climatology in urban areas reveal that mixed layer height is generally lowest in  
521 winter months (van der Kamp and McKendry, 2010; de Arruda Moreira et al., 2020). A lower mixed layer  
522 height is indicative of less mixing of scalars (such as temperature), leading to a less homogeneous boundary  
523 layer where the effects of surface forcings remain more localized as compared to a boundary layer with more  
524 mixing and a more uniform composition. This localization of the effects of surface forcings on the mixed  
525 layer may result in a disparity in results due to the difference in location between the location of the ASOS  
526 observation stations and the locations of the flux towers, the latter of which is where the model is run and  
527 validated.

528 The geospatial variability of  $Q_H$  in coastal areas is high, in part, due to the complex boundary layer  
529 formed by the combination of a marine and urban boundary layer (Melecio-Vazquez et al., 2018; Thompson  
530 et al., 2007).  $Q_H$  and its transport may be influenced by advective phenomena such as sea breezes, which  
531 are further amplified by the sharp transitions between different land cover types, such as the transition from  
532 water to a highly-developed urban area, as exists along the portions of New York City bordering the Hudson  
533 and East Rivers, Long Island Sound, or the Atlantic Ocean (Bou-Zeid et al., 2020; Lee, 2015; Thompson  
534 et al., 2007). The inability to explicitly capture the influence of advection on  $Q_H$  is a limitation of this  
535 model. Additionally, the 2 km spatial resolution of GOES-16 satellite data is unable to properly capture  
536 areas with sharp land cover transitions on land, such as the one at the Staten Island (STAT) flux tower (see  
537 Figure 3), which may introduce significant bias into the estimation. This is discussed in greater detail in  
538 Section 4.4.

#### 539 4.4 Potential sources of error

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

542 A likely source of model error stemmed from the spatial resolution of the GOES-16 LST product. The  
543 LST product features pixels at a spatial resolution of 2 km, which translates to 196 pixels spanning the New  
544 York City area (approximate land area of  $778 \text{ km}^2$ ). Although this allows for fragmentation of the city into  
545 pixels that can distinguish districts (such as boroughs for New York City) from each other, GOES-16 LST  
546 pixel sizes are still considered large relative to the source areas for the flux measurements due to the spatial  
547 heterogeneity the observed areas. Because of the mismatch in pixel and source area size, contributions from  
548 localized phenomena such as urban street canyons and vegetated spaces may not be accounted for (Erell and  
549 Williamson, 2006; Xiaomin et al., 2006). This is important due to the highly variable land cover types that  
550 exist between neighborhoods in large cities such as New York City (Hamstead et al., 2016).

551 Another potential source of error arises from biases in estimation methods for  $T_{LST}$  and  $T_{air}$  used in  
552 this model. The GOES-R Land Surface Temperature product has been shown to have an average precision  
553 error of 1.58 K when tested over 6 rural locations (Yu et al., 2011). It is worth noting validation for this  
554 product has not been performed in urban areas, to the authors' knowledge. The air temperature model used  
555 features an RMSE of 2.6 K and a bias of 0.8 K (Hrisko et al., 2020). Additionally, this air temperature  
556 model featured a cold bias during nighttime hours and a warm bias during daytime hours, leading to under-  
557 and over-prediction respectively. Moreover, the anisotropy of satellite retrievals in mid- to high-latitude  
558 areas (such as New York City) with significant urban canopies may contribute to warm biases in estimations

559 of  $T_{LST}$  (Vinnikov et al., 2012; Wang et al., 2021), especially during the winter months. The combined  
560 errors, in conjunction with the high model sensitivity to both  $T_{LST}$  and  $T_{air}$ , can cause significant errors in  
561 estimating  $Q_H$ , especially during nighttime hours.

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

568 It should be noted that eddy covariance measurements in urban areas are subject to many limitations  
569 (Aubinet et al., 2012). One such limitation arises from the inability to represent the surrounding vicinity  
570 as a single land cover type. Another limitation is caused by the installation of eddy covariance systems in  
571 proximity of or on top of buildings, as the effects of flow distortion can create a local flow environment that is  
572 not representative of the surrounding environment. The presence of obstructions in the flow path may result  
573 in biased data if obstructions of similar height to the station elevation are upstream of the flow, resulting  
574 in some distortion. For the Queens station in particular, easterly flows would result in some bias due to  
575 buildings of similar height due east of the flux tower. However, the winds at each station were primarily  
576 from the west over all days evaluated in the study, with over 75% of all winds recorded possessing a westerly  
577 component. This indicates that a majority of the observational data is not affected by flow distortion due to  
578 nearby obstructions, thus increasing confidence in the observational data as a valid reference. Moreover, the  
579 stations are minimally affected by obstructions upstream due to their prominence over buildings upstream  
580 of the flow. Additionally, it is worth noting that the building upon which the flux tower is installed may  
581 produce some localized flow distortion which could affect observations (Oke et al., 2017; Wang and Benson,  
582 2021). However, due to logistical and safety issues, it is often difficult to find locations in cities that satisfy  
583 all the conditions required for an ideal reference for validation purposes. Therefore, we use the flux data  
584 as reference to compare our model results and fully acknowledge the limitations in using them. Finally,  
585 advances in modeling like the one we have proposed here could pave the way to reduce our dependence on  
586 eddy covariance measurements to quantify heat fluxes in urban areas.

587 A further source of error between the model and the NYS Mesonet observations is the distance between  
588 the ASOS stations, where wind speed data is collected, and the Mesonet flux stations, where flux data is  
589 collected. The ASOS stations used for data collection are all located at large airports that feature large  
590 expanses of flat surfaces surrounding the station instrumentation with minimal upwind obstructions. By  
591 comparison, each flux station is located in a moderately- to heavily-urbanized area, reducing the upwind  
592 fetch and exposing instrumentation to hyperlocal turbulence that is a direct function of the surrounding  
593 geometry (Kastner-Klein et al., 2004) Additionally, the highly heterogeneous land cover in the New York  
594 City metropolitan area has the potential to create highly-localized wind fields due to phenomena such as  
595 the urban street canyon effect in areas with tall buildings and sea breezes in coastal areas (Park et al.,  
596 2012; Thompson et al., 2007). Methods to address limitations in location mismatches and point-based  
597 observational methods are addressed in Section 4.5.

## 598 4.5 Future work

599 A number of factors from this study motivate future work to improve the accuracy of the model.  
600 A possible improvement to the model involves validation at a range of test sites with a variety of land

cover types at different locations, both within New York City and in other urban areas. This allows for the model to be evaluated for a wider range of land cover types and permits the evaluation of model sensitivity to land cover type. Moreover, due to the study focusing on one city, atmospheric conditions that are a function of location, such as air pressure or advective fluxes, may not be fully accounted for in this model, potentially requiring a modification of assumptions or parameter values. A related additional improvement to the validation effort would be the use of scintillometry to obtain path-averaged flux observations that are more representative of urban land cover types, as opposed to point-based observations as used in this study (Crawford et al., 2017; Lee, 2015; Nadeau et al., 2009) due to the significant heterogeneity of land cover types in urban areas.

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

The estimation of nocturnal sensible heat flux is another critical component to improving model accuracy. The model often underestimates nocturnal  $Q_H$  relative to the observation sites despite good approximation during the day. As discussed in Section 4.4, a cold bias in the estimation  $T_{LST}$  during nighttime hours is a potential explanation for the negative  $Q_H$  values predicted by the model. In contrast, observations show near-zero and positive  $Q_H$  values at night as a result of the release of heat stored during the day, especially in the most urbanized portions of the study area (Grimmond and Oke, 1999; Hrisko et al., 2021). Correction for nocturnal temperature bias would better display the relationship between heat storage and  $Q_H$  and improve model accuracy. Therefore, this topic requires further exploration.

## 4.6 Application potential

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

## 640 5 Conclusions

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

649 Over the duration of the validation period, the RMSE between the model and observational data was  
650  $47.32 \text{ Wm}^{-2}$ , with a MBE of  $16.58 \text{ Wm}^{-2}$ , a model NSC of 0.54, and a  $R^2$  of 0.70. This presents a  
651 significant improvement over the uWRF model run over fewer days in the same validation period (RMSE of  
652  $108.1 \text{ Wm}^{-2}$ , MBE of  $-30.1 \text{ Wm}^{-2}$ , NSC of -1.47 and  $R^2$  of 0.63). The model performed especially well in  
653 warmer months ( $R^2$  values of 0.80 and 0.79 for summer and autumn, respectively) and during the daytime  
654 and evening hours.

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

## 662 Acknowledgments

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

## 677 References

678 Aubinet, M., Vesala, T. and Papale, D. (2012), *Eddy Covariance Measurements Over Urban Areas*, Springer.

- 679 Bala, R., Prasad, R. and Yadav, V. P. (2019), ‘Disaggregation of modis land surface temperature in urban  
680 areas using improved thermal sharpening techniques’, *Advances in Space Research* **64**(3), 591–602.
- 681 Bauer, T. J. (2020), ‘Interaction of urban heat island effects and land–sea breezes during a new york city  
682 heat event’, *Journal of Applied Meteorology and Climatology* **59**(3), 477–495.
- 683 Best, M. (2005), ‘Representing urban areas within operational numerical weather prediction models’,  
684 *Boundary-Layer Meteorology* **114**(1), 91–109.
- 685 Bou-Zeid, E., Anderson, W., Katul, G. G. and Mahrt, L. (2020), ‘The persistent challenge of surface het-  
686 erogeneity in boundary-layer meteorology: a review’, *Boundary-Layer Meteorology* **177**(2), 227–245.
- 687 Brenner, C., Thiem, C. E., Wizemann, H.-D., Bernhardt, M. and Schulz, K. (2017), ‘Estimating spatially  
688 distributed turbulent heat fluxes from high-resolution thermal imagery acquired with a uav system’, *In-  
689 ternational Journal of Remote Sensing* **38**(8-10), 3003–3026.
- 690 Businger, J. A., Wyngaard, J. C., Izumi, Y. and Bradley, E. F. (1971), ‘Flux-profile relationships in the  
691 atmospheric surface layer’, *Journal of Atmospheric Sciences* **28**(2), 181–189.
- 692 Cammalleri, C., Anderson, M., Ciraolo, G., D’urso, G., Kustas, W., La Loggia, G. and Minacapilli, M.  
693 (2012), ‘Applications of a remote sensing-based two-source energy balance algorithm for mapping surface  
694 fluxes without in situ air temperature observations’, *Remote Sensing of Environment* **124**, 502–515.
- 695 Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C., Grossman-Clarke, S., Loridan, T., Man-  
696 ning, K. W., Martilli, A., Miao, S. et al. (2011), ‘The integrated wrf/urban modelling system: develop-  
697 ment, evaluation, and applications to urban environmental problems’, *International Journal of Climatology*  
698 **31**(2), 273–288.
- 699 Chen, F. and Zhang, Y. (2009), ‘On the coupling strength between the land surface and the atmosphere:  
700 From viewpoint of surface exchange coefficients’, *Geophysical Research Letters* **36**(10).
- 701 **URL:** <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL037980>
- 702 Chrysoulakis, N., Grimmond, S., Feigenwinter, C., Lindberg, F., Gastellu-Etchegorry, J.-P., Marconcini, M.,  
703 Mitraka, Z., Stagakis, S., Crawford, B., Olofson, F., Landier, L., Morrison, W. and Parlow, E. (2018),  
704 ‘Urban energy exchanges monitoring from space’, *Scientific Reports* **8**(1), 11498.
- 705 **URL:** <http://www.nature.com/articles/s41598-018-29873-x>
- 706 Crawford, B., Grimmond, C. S. B., Ward, H. C., Morrison, W. and Kotthaus, S. (2017), ‘Spatial and tem-  
707 poral patterns of surface–atmosphere energy exchange in a dense urban environment using scintillometry’,  
708 *Quarterly Journal of the Royal Meteorological Society* **143**(703), 817–833.
- 709 de Arruda Moreira, G., Guerrero-Rascado, J. L., Bravo-Aranda, J. A., Foyo-Moreno, I., Cazorla, A., Alados,  
710 I., Lyamani, H., Landulfo, E. and Alados-Arboledas, L. (2020), ‘Study of the planetary boundary layer  
711 height in an urban environment using a combination of microwave radiometer and ceilometer’, *Atmospheric  
712 Research* **240**, 104932.
- 713 Department of City Planning, N. Y. C. (2021), ‘Zoning: District guide - about zoning districts’.
- 714 **URL:** <https://www1.nyc.gov/site/planning/zoning/districts-tools.page>
- 715 Dudhia, J. (1989), ‘Numerical study of convection observed during the winter monsoon experiment using a  
716 mesoscale two-dimensional model’, *Journal of Atmospheric Sciences* **46**(20), 3077–3107.

- 717 Dyer, A. J. (1974), ‘A review of flux-profile relationships’, *Boundary-Layer Meteorology* **7**(3), 363–372.
- 718 **URL:** <http://link.springer.com/10.1007/BF00240838>
- 719 Erell, E. and Williamson, T. (2006), ‘Simulating air temperature in an urban street canyon in all weather  
720 conditions using measured data at a reference meteorological station’, *International Journal of Climatology: A Journal of the Royal Meteorological Society* **26**(12), 1671–1694.
- 722 Feddema, J. J., Oleson, K. W., Bonan, G. B., Mearns, L. O., Buja, L. E., Meehl, G. A. and Washington, W. M. (2005), ‘The importance of land-cover change in simulating future climates’, *Science* **310**(5754), 1674–1678.
- 725 Feigenwinter, C., Vogt, R., Parlow, E., Lindberg, F., Marconcini, M., Del Frate, F. and Chrysoulakis, N. (2018), ‘Spatial distribution of sensible and latent heat flux in the city of basel (switzerland)’, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **11**(8), 2717–2723.
- 728 Gamarro, H., Gonzalez, J. E. and Ortiz, L. E. (2019), ‘On the assessment of a numerical weather prediction  
729 model for solar photovoltaic power forecasts in cities’, *Journal of Energy Resources Technology* **141**(6).
- 730 Garratt, J. R. (1994), ‘The atmospheric boundary layer’, *Earth-Science Reviews* **37**(1-2), 89–134.
- 731 Gedzelman, S., Austin, S., Cermak, R., Stefano, N., Partridge, S., Quesenberry, S. and Robinson, D. (2003),  
732 ‘Mesoscale aspects of the urban heat island around new york city’, *Theoretical and applied climatology*  
733 **75**(1), 29–42.
- 734 Grimmond, C. and Cleugh, H. (1994), ‘A simple method to determine obukhov lengths for suburban areas’,  
735 *Journal of Applied Meteorology* **33**(3), 435–440.
- 736 Grimmond, C. and Oke, T. R. (1999), ‘Heat storage in urban areas: Local-scale observations and evaluation  
737 of a simple model’, *Journal of Applied Meteorology and Climatology* **38**(7), 922–940.
- 738 Gutierrez, E., Gonzalez, J., Martilli, A. and Bornstein, R. (2015), ‘On the anthropogenic heat fluxes using  
739 an air conditioning evaporative cooling parameterization for mesoscale urban canopy models’, *Journal of  
740 Solar Energy Engineering* **137**(5), 051005.
- 741 Gutierrez, E., Gonzalez, J., Martilli, A., Bornstein, R. and Arend, M. (2015a), ‘Simulations of a heat-wave  
742 event in new york city using a multilayer urban parameterization’, *Journal of Applied Meteorology and  
743 Climatology* **54**(2), 283–301.
- 744 Gutierrez, E., Gonzalez, J., Martilli, A., Bornstein, R. and Arend, M. (2015b), ‘Simulations of a heat-wave  
745 event in new york city using a multilayer urban parameterization’, *Journal of Applied Meteorology and  
746 Climatology* **54**(2), 283–301.
- 747 Hamstead, Z. A., Kremer, P., Larondelle, N., McPhearson, T. and Haase, D. (2016), ‘Classification of the  
748 heterogeneous structure of urban landscapes (sturla) as an indicator of landscape function applied to  
749 surface temperature in new york city’, *Ecological Indicators* **70**, 574 – 585. Navigating Urban Complexity:  
750 Advancing Understanding of Urban Social – Ecological Systems for Transformation and Resilience.
- 751 **URL:** <http://www.sciencedirect.com/science/article/pii/S1470160X1500549X>
- 752 Hong, S.-Y. and Dudhia, J. (2012), ‘Next-generation numerical weather prediction: Bridging parameteriza-  
753 tion, explicit clouds, and large eddies’, *Bulletin of the American Meteorological Society* **93**(1), ES6–ES9.

- 754 Hrisko, J., Ramamurthy, P., Melecio-Vazquez, D. and Gonzalez, J. E. (2021), ‘Spatiotemporal variability of  
755 heat storage in major u.s. cities—a satellite-based analysis’, *Remote Sensing* **13**, 59.
- 756 **URL:** <https://www.mdpi.com/2072-4292/13/1/59>
- 757 Hrisko, J., Ramamurthy, P., Yu, Y., Yu, P. and Melecio-Vázquez, D. (2020), ‘Urban air temperature model  
758 using goes-16 lst and a diurnal regressive neural network algorithm’, *Remote Sensing of Environment*  
759 **237**, 111495.
- 760 **URL:** <http://www.sciencedirect.com/science/article/pii/S0034425719305140>
- 761 Imran, H. M., Kala, J., Ng, A. and Muthukumaran, S. (2018), ‘Effectiveness of green and cool roofs in  
762 mitigating urban heat island effects during a heatwave event in the city of melbourne in southeast australia’,  
763 *Journal of Cleaner Production* **197**, 393–405.
- 764 Janjić, Z. I. (1994), ‘The step-mountain eta coordinate model: Further developments of the convection,  
765 viscous sublayer, and turbulence closure schemes’, *Monthly weather review* **122**(5), 927–945.
- 766 Kastner-Klein, P., Berkowicz, R. and Britter, R. (2004), ‘The influence of street architecture on flow and  
767 dispersion in street canyons’, *Meteorology and Atmospheric Physics* **87**(1), 121–131.
- 768 Kato, S. and Yamaguchi, Y. (2005), ‘Analysis of urban heat-island effect using aster and etm+ data: Separation  
769 of anthropogenic heat discharge and natural heat radiation from sensible heat flux’, *Remote Sensing  
770 of Environment* **99**(1), 44 – 54. Scientific Results from ASTER.
- 771 **URL:** <http://www.sciencedirect.com/science/article/pii/S0034425705001707>
- 772 Kim, M.-S. and Kwon, B. H. (2019), ‘Estimation of sensible heat flux and atmospheric boundary layer height  
773 using an unmanned aerial vehicle’, *Atmosphere* **10**(7), 363.
- 774 Launiainen, J. and Vihma, T. (1990), ‘Derivation of turbulent surface fluxes — an iterative flux-profile  
775 method allowing arbitrary observing heights’, *Environmental Software* **5**(3), 113 – 124.
- 776 **URL:** <http://www.sciencedirect.com/science/article/pii/026698389090021W>
- 777 Lee, S.-H. (2015), ‘Determination of turbulent sensible heat flux over a coastal maritime area using a large  
778 aperture scintillometer’, *Boundary-Layer Meteorology* **157**(2), 309–319.
- 779 Legates, D. R. and McCabe Jr, G. J. (1999), ‘Evaluating the use of “goodness-of-fit” measures in hydrologic  
780 and hydroclimatic model validation’, *Water resources research* **35**(1), 233–241.
- 781 Leroyer, S., Bélair, S., Husain, S. Z. and Mailhot, J. (2014), ‘Subkilometer numerical weather prediction in  
782 an urban coastal area: A case study over the vancouver metropolitan area’, *Journal of Applied Meteorology  
783 and Climatology* **53**(6), 1433–1453.
- 784 Li, D. and Bou-Zeid, E. (2014), ‘Quality and sensitivity of high-resolution numerical simulation of urban  
785 heat islands’, *Environmental Research Letters* **9**(5), 055001.
- 786 **URL:** <https://doi.org/10.1088/1748-9326/9/5/055001>
- 787 Liu, Y., Shintaro, G., Zhuang, D. and Kuang, W. (2012), ‘Urban surface heat fluxes infrared remote sensing  
788 inversion and their relationship with land use types’, *Journal of Geographical Sciences* **22**(4), 699–715.
- 789 **URL:** <http://link.springer.com/10.1007/s11442-012-0957-7>
- 790 Melecio-Vazquez, D., Ramamurthy, P., Arend, M. and Gonzalez-Cruz, J. E. (2018), ‘Thermal structure of a  
791 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>

792 Miglietta, F., Gioli, B., Brunet, Y., Hutjes, R., Matese, A., Sarrat, C. and Zaldei, A. (2009), ‘Sensible and  
793 latent heat flux from radiometric surface temperatures at the regional scale: methodology and evaluation’,  
794 *Biogeosciences* **6**(10), 1975–1986.

795 Mkhwanazi, M., Chávez, J. L. and Rambikur, E. H. (2012), ‘Comparison of large aperture scintillometer and  
796 satellite-based energy balance models in sensible heat flux and crop evapotranspiration determination’,  
797 *International Journal of Remote Sensing Applications* **2**(1), 24–30.

798 Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J. and Clough, S. A. (1997), ‘Radiative transfer  
799 for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave’, *Journal of*  
800 *Geophysical Research: Atmospheres* **102**(D14), 16663–16682.

801 Monin, A. and Obukhov, A. (1954), ‘Dimensionless characteristics of turbulence in the surface layer’, *Akad.*  
802 *Nauk SSSR, Geofiz. Inst., Tr* **24**, 163–187.

803 Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D. and Veith, T. L. (2007), ‘Model  
804 evaluation guidelines for systematic quantification of accuracy in watershed simulations’, *Transactions of*  
805 *the ASABE* **50**(3), 885–900.

806 Mott, R., Daniels, M. and Lehning, M. (2015), ‘Atmospheric flow development and associated changes in  
807 turbulent sensible heat flux over a patchy mountain snow cover’, *Journal of Hydrometeorology* **16**(3), 1315–  
808 1340.

809 Nadeau, D. F., Brutsaert, W., Parlange, M., Bou-Zeid, E., Barrenetxea, G., Couach, O., Boldi, M.-O., Selker,  
810 J. S. and Vetterli, M. (2009), ‘Estimation of urban sensible heat flux using a dense wireless network of  
811 observations’, *Environmental fluid mechanics* **9**(6), 635–653.

812 Nash, J. E. and Sutcliffe, J. V. (1970), ‘River flow forecasting through conceptual models part i—a discussion  
813 of principles’, *Journal of hydrology* **10**(3), 282–290.

814 Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi,  
815 D., Rosero, E. et al. (2011), ‘The community noah land surface model with multiparameterization options  
816 (noah-mp): 1. model description and evaluation with local-scale measurements’, *Journal of Geophysical*  
817 *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](https://goes-r.gov/products/RIMPs/RIMP_ABI - L2FSC_v1.0.pdf)

818 Oke, T. R., Mills, G., Christen, A. and Voogt, J. A. (2017), *Urban climates*, Cambridge University Press.

819 Ortega-Farías, S., Ortega-Salazar, S., Poblete, T., Kilic, A., Allen, R., Poblete-Echeverría, C., Ahumada-  
820 Orellana, L., Zuñiga, M. and Sepúlveda, D. (2016), ‘Estimation of energy balance components over a drip-  
821 irrigated olive orchard using thermal and multispectral cameras placed on a helicopter-based unmanned  
822 aerial vehicle (uav)’, *Remote Sensing* **8**(8), 638.

- 823 Ortiz, L. E., Gonzalez, J. E., Gutierrez, E. and Arend, M. (2017), ‘Forecasting building energy demands  
824 with a coupled weather-building energy model in a dense urban environment’, *Journal of Solar Energy  
825 Engineering* **139**(1), 011002.
- 826 Park, S.-B., Baik, J.-J., Raasch, S. and Letzel, M. O. (2012), ‘A large-eddy simulation study of thermal  
827 effects on turbulent flow and dispersion in and above a street canyon’, *Journal of Applied Meteorology and  
828 Climatology* **51**(5), 829–841.
- 829 Pond, S., Fissel, D. B. and Paulson, C. A. (1974), ‘A note on bulk aerodynamic coefficients for sensible heat  
830 and moisture fluxes’, *Boundary-Layer Meteorology* **6**(1-2), 333–339.
- 831 **URL:** <http://link.springer.com/10.1007/BF00232493>
- 832 Price, J. C. (1979), ‘Assessment of the urban heat island effect through the use of satellite data’, *Monthly  
833 Weather Review* **107**(11), 1554–1557.
- 834 Raupach, M. R. (1994), ‘Simplified expressions for vegetation roughness length and zero-plane displacement  
835 as functions of canopy height and area index’, *Boundary-Layer Meteorology* **71**, 211–216.
- 836 **URL:** <https://doi.org/10.1007/BF00709229>
- 837 Salamanca, F. and Martilli, A. (2010), ‘A new building energy model coupled with an urban canopy pa-  
838 rameterization for urban climate simulations—part ii. validation with one dimension off-line simulations’,  
839 *Theoretical and Applied Climatology* **99**(3), 345–356.
- 840 Schumacher, D. L., Keune, J., Van Heerwaarden, C. C., de Arellano, J. V.-G., Teuling, A. J. and Miralles,  
841 D. G. (2019), ‘Amplification of mega-heatwaves through heat torrents fuelled by upwind drought’, *Nature  
842 Geoscience* **12**(9), 712–717.
- 843 Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., Huang, X.-Y. and Duda, M. (2008),  
844 A Description of the Advanced Research WRF Version 3, Technical report, UCAR/NCAR.
- 845 **URL:** <http://opensky.ucar.edu/islandora/object/technotes:500>
- 846 Stull, R. (1988), *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers.
- 847 Tang, B.-H., Zhan, C., Li, Z.-L., Wu, H. and Tang, R. (2016), ‘Estimation of land surface temperature from  
848 modis data for the atmosphere with air temperature inversion profile’, *IEEE Journal of Selected Topics  
849 in Applied Earth Observations and Remote Sensing* **10**(6), 2976–2983.
- 850 Thompson, W. T., Holt, T. and Pullen, J. (2007), ‘Investigation of a sea breeze front in an urban envi-  
851 ronment’, *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences,  
852 applied meteorology and physical oceanography* **133**(624), 579–594.
- 853 US Census Bureau (2019), ‘Population and Housing Unit Estimates Tables’.
- 854 **URL:** <https://www.census.gov/programs-surveys/popest/data/tables.html>
- 855 Valenti, J. (2017), Goes-r series product definition and users’ guide, Technical report.
- 856 van der Kamp, D. and McKendry, I. (2010), ‘Diurnal and seasonal trends in convective mixed-layer heights  
857 estimated from two years of continuous ceilometer observations in vancouver, bc’, *Boundary-layer meteo-  
858 rology* **137**(3), 459–475.
- 859 Vautard, R., Yiou, P., D’andrea, F., De Noblet, N., Viovy, N., Cassou, C., Polcher, J., Ciais, P., Kageyama,  
860 M. and Fan, Y. (2007), ‘Summertime european heat and drought waves induced by wintertime mediter-  
861 ranean rainfall deficit’, *Geophysical Research Letters* **34**(7).

- 862 Vinnikov, K. Y., Yu, Y., Goldberg, M. D., Tarpley, D., Romanov, P., Laszlo, I. and Chen, M. (2012),  
863 ‘Angular anisotropy of satellite observations of land surface temperature’, *Geophysical Research Letters*  
864 **39**(23).
- 865 Voogt, J. A. and Grimmond, C. (2000), ‘Modeling surface sensible heat flux using surface radiative temper-  
866 atures in a simple urban area’, *Journal of Applied Meteorology* **39**(10), 1679–1699.
- 867 Voogt, J. and Oke, T. (2003), ‘Thermal remote sensing of urban climates’, *Remote Sensing of Environment*  
868 **86**(3), 370 – 384. Urban Remote Sensing.  
869 **URL:** <http://www.sciencedirect.com/science/article/pii/S0034425703000798>
- 870 Wang, C., Myint, S. W., Wang, Z. and Song, J. (2016), ‘Spatio-temporal modeling of the urban heat island  
871 in the phoenix metropolitan area: Land use change implications’, *Remote Sensing* **8**(3), 185.
- 872 Wang, D., Chen, Y., Hu, L., Voogt, J. A., Gastellu-Etchegorry, J.-P. and Krayenhoff, E. S. (2021), ‘Modeling  
873 the angular effect of modis lst in urban areas: A case study of toulouse, france’, *Remote Sensing of  
874 Environment* **257**, 112361.
- 875 Wang, Y. and Benson, M. J. (2021), ‘Large-eddy simulation of turbulent flows over an urban building  
876 array with the able-lbm and comparison with 3d mri observed data sets’, *Environmental Fluid Mechanics*  
877 **21**(2), 287–304.
- 878 World Meteorological Organization, ed. (2008), *Aerodrome reports and forecasts: a users’ handbook to the  
879 codes*, number no. 782 in ‘WMO’, 5th ed edn, WMO, Geneva, Switzerland. OCLC: ocn300459644.
- 880 Xia, T., Kustas, W. P., Anderson, M. C., Alfieri, J. G., Gao, F., McKee, L., Prueger, J. H., Geli, H. M., Neale,  
881 C. M., Sanchez, L. et al. (2016), ‘Mapping evapotranspiration with high-resolution aircraft imagery over  
882 vineyards using one-and two-source modeling schemes’, *Hydrology and Earth System Sciences* **20**(4), 1523–  
883 1545.
- 884 Xiaomin, X., Zhen, H. and Jiasong, W. (2006), ‘The impact of urban street layout on local atmospheric  
885 environment’, *Building and Environment* **41**(10), 1352–1363.
- 886 Xu, W., Wooster, M. and Grimmond, C. (2008), ‘Modelling of urban sensible heat flux at multiple spatial  
887 scales: A demonstration using airborne hyperspectral imagery of shanghai and a temperature–emissivity  
888 separation approach’, *Remote Sensing of Environment* **112**(9), 3493–3510.
- 889 Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, C., Dewitz, J., Fry, J.,  
890 Funk, M., Granneman, B., Liknes, G. C., Rigge, M. and Xian, G. (2018), ‘A new generation of the United  
891 States National Land Cover Database: Requirements, research priorities, design, and implementation  
892 strategies’, *ISPRS Journal of Photogrammetry and Remote Sensing* **146**, 108–123.  
893 **URL:** <https://linkinghub.elsevier.com/retrieve/pii/S092427161830251X>
- 894 Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M., Vinnikov, K. Y., Sun, D. and Tian,  
895 Y. (2011), ‘Validation of goes-r satellite land surface temperature algorithm using surfrad ground mea-  
896 surements and statistical estimates of error properties’, *IEEE Transactions on Geoscience and Remote  
897 Sensing* **50**(3), 704–713.
- 898 Zhang, J., Draxl, C., Hopson, T., Delle Monache, L., Vanvyve, E. and Hodge, B.-M. (2015), ‘Comparison  
899 of numerical weather prediction based deterministic and probabilistic wind resource assessment methods’,  
900 *Applied Energy* **156**, 528–541.

902 **List of Figure Captions**

- 903 • Figure 1: Satellite view of the New York City metropolitan area. New York City, which is composed  
904 of 5 boroughs (labeled), is the most-heavily urbanized portion of the metropolitan area, while lower  
905 density suburbs and woodlands compose the outer portions of the metropolitan area.
- 906 • Figure 2: Process flowchart for the sensible heat flux model. Observational data was used for validation  
907 of the satellite model as well as inputs to the iterative algorithm. The numerical model used remotely-  
908 sensed data from the GOES-16 satellite, as well as ancillary datasets for land cover and geographic  
909 data. Error analysis was performed by comparing observational data and model results.
- 910 • Figure 3: Land cover map of the New York City metropolitan area, per the 2016 National Land Cover  
911 Database (Yang et al. (2018)). The legend shows land cover types and the percentage of the study  
912 area occupied by each land cover type. Land cover data is shown at a 30 m resolution. Note that flux  
913 observation towers and ground weather (ASOS) stations are labeled accordingly.
- 914 • Figure 4: Gridded map of element roughness heights across the New York City metropolitan area.  
915 Note that flux observation towers and ground weather (ASOS) stations are labeled accordingly.
- 916 • Figure 5: Comparison of model and observed  $Q_H$  from 1 June 2019 to 31 May 2020.
- 917 • Figure 6: Comparison of model and observed  $Q_H$  divided into distinct periods of the day. (from left  
918 to right) Early morning (pre-sunrise), daytime, evening and late night (post-sunset).
- 919 • Figure 7: Seasonal averages of daily  $Q_H$ , averaged over all stations. For reference, the acronyms reflect  
920 their represented months (for example, 'JJA' corresponds to June, July, and August). All times are  
921 in local standard time (LST) and background shading corresponds approximately to average day and  
922 night durations for the respective season.
- 923 • Figure 8: Comparison of model and observed  $Q_H$ , divided into seasons. Note that the acronyms  
924 correspond to months in each season (for example, 'JJA' corresponds to June, July and August).
- 925 • Figure 9:  $Q_H$  error between model results and observational data, averaged per season per location.  
926 For reference, the acronyms reflect their represented months (for example, 'JJA' corresponds to June,  
927 July, and August).
- 928 • Figure 10:  $Q_H$  ( $\text{W m}^{-2}$ ) in New York City on 24 October 2019 shown at 4 different times during the  
929 daytime. Note that all times are in local standard time (LST).
- 930 • Figure 11: Left: Model sensitivity to each parameter represented by error from baseline values catego-  
931 rized by location (left) and atmospheric stability,  $\zeta$  (right).
- 932 • Figure 12: Observed and modeled (uWRF and dedicated)  $Q_H$  at days selected for uWRF study at  
933 the Queens (QUEE) station. Note that 'DJF-1' represents the date 23 December 2019 and 'DJF-2'  
934 represents 20 January 2020.
- 935 • Figure 13: Gridded maps showing error between dedicated  $Q_H$  and uWRF models, October 24, 2019  
936 at 4:00 local standard time (left) and 14:00 local standard time (right). Light blue pixels correspond  
937 to areas with mostly covered in water/marsh.