JOURNAL OF GEOPHYSICAL RESEARCH: ATMOSPHERES, VOL. 118, 1–18, doi:10.1002/2013JD020657, 2013

Development and evaluation of a mosaic approach in the WRF-Noah framework

Dan Li,[1] Elie Bou-Zeid,[1] Michael Barlage,[2] Fei Chen,[2] and James A. Smith[1]

Received 30 July 2013; revised 3 October 2013; accepted 16 October 2013.

[1] The current Weather Research and Forecasting (WRF)-Noah modeling framework considers only the dominant land cover type within each grid cell, which here is referred to as the "dominant" approach. In order to assess the impact of subgrid-scale variability in land cover composition, a mosaic/tiling approach (hereafter the "mosaic" approach) is implemented into the coupled WRF-Noah modeling system. In the mosaic approach, a certain number (*N*) of tiles, each representing a land cover category, is considered within each grid cell. WRF simulations of a clear sky day and a rainfall period over a heterogeneous urban/suburban setting show that the two approaches generate differences in the surface energy balance, land surface temperature, near-surface states, boundary layer growth, as well as rainfall distribution. Evaluation against a variety of observational data (including surface flux measurements, the MODIS land surface temperature product, and radar rainfall estimates) indicates that, compared to the dominant approach, the mosaic approach has a better performance. In addition, WRF-simulated results with the mosaic approach are less sensitive to the spatial resolution of the grid: Larger differences are observed in simulations of different resolutions with the dominant approach. The effect of increasing the number of tiles (*N*) on the WRF-simulated results is also examined. When *N* increases from 1 (i.e., the dominant approach) to 15, changes in the ground heat flux, sensible heat flux, surface temperature, and 2 m air temperature are more significant during nighttime. Changes in the 2 m specific humidity are more significant during daytime, and changes in the boundary layer height are most prominent during the morning and afternoon transitional periods.

Citation: Li, D., E. Bou-Zeid, M. Barlage, F. Chen, and J. A. Smith (2013), Development and evaluation of a mosaic approach in the WRF-Noah framework, *J. Geophys. Res. Atmos.*, *118*, doi:10.1002/2013JD020657.

1. Introduction

[2] Numerical weather prediction and climate models are typically implemented at horizontal resolutions of 1 km to 100 km. As such, land surface characteristics (e.g., leaf area index, albedo, and roughness length) can vary significantly within a single grid cell. This subgrid-scale variability of land surface characteristics can strongly affect the grid cell averaged surface fluxes and thus the weather and climate dynamics, as well as the hydrological cycle [*Avissar and Schmidt*, 1998; *Chen and Avissar*, 1994; *Chen and Dudhia*, 2001; *Chen et al.*, 2003; *Koster and Suarez*, 1992a; *LeMone et al.*, 2008; *Li and Avissar*, 1994; *Molod and Salmun*, 2002]. Generally, the more heterogeneous the surface is, the more important the subgrid-scale variability of land surface characteristics will be.

[3] The parameterization of subgrid-scale variability of land surface characteristics has been the subject of active research in the past few decades (see *Giorgi and Avissar* [1997] for a review). *Avissar and Pielke* [1989] proposed a "mosaic/tiling" approach to represent the subgrid-scale variability of land surface characteristics, which considers a certain number of patches (or tiles) within a grid cell. The surface fluxes and surface state variables are first calculated for each tile (which is usually assumed to be homogeneous) and then spatially averaged over the whole grid cell; the atmospheric variables continue to be assumed homogeneous within a grid cell. While this approach has been adopted by many mesoscale and global-scale models, the traditional method without considering subgrid-scale variability of land surface characteristics is still used in some modern weather and climate models [see, e.g., *Chen and Dudhia*, 2001]. This latter approach is usually referred to as the "dominant" approach because each grid cell is assumed to be entirely composed of the most abundant tile.

[4] Other methods such as the "composite" approach [*Koster and Suarez*, 1992b] and the "statistical-dynamical" approach [*Avissar*, 1991; *Entekhabi and Eagleson*, 1989; *Famiglietti and Wood*, 1991] are also available for parameterizing subgrid-scale land surface variability. Similar to the dominant approach, the composite approach also considers only one homogeneous tile within each grid cell. However,

[1]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[2]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA.

Corresponding author: D. Li, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08540, USA. (danl.princeton.2009@gmail.com)

2169-897X/13/10.1002/2013JD020657

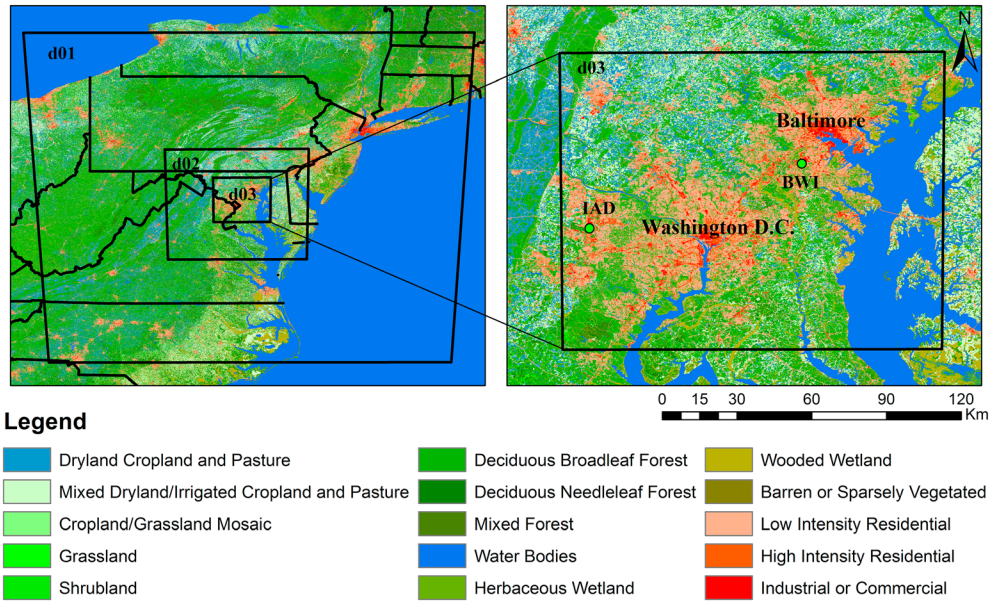

Figure 1. WRF setup and land cover maps over the Baltimore-Washington Corridor.

the surface properties of this tile are aggregated based on the properties of all the tiles within the grid cell, which is different from the dominant approach that imposes the surface properties of a single, most abundant land cover type over the grid cell. The aggregation can be either linear or nonlinear. The statistical-dynamical approach assumes that the land surface parameters that are critical for calculating surface fluxes follow certain probability density functions (PDFs). The PDF of each of these surface parameters is divided into a number of bins and then surface fluxes are calculated using these bin values of surface parameters. The fluxes over the bins are numerically integrated to produce the grid cell averaged fluxes. This approach is in fact similar to the mosaic/tiling approach since different tiles considered in the mosaic approach are likely to have distinct surface parameters, which could alternatively be viewed as different bins in the statistical-dynamical approach. In this study, we will not consider the composite approach and the statistical-dynamical approach and will only focus on the dominant approach and the mosaic/tiling approach.

[5] The Weather Research and Forecasting (WRF) model with the coupled Noah land surface model [*Chen and Dudhia*, 2001; *Ek et al*., 2003; *Skamarock and Klemp*, 2008] is adopted to implement and test the mosaic/tiling approach. In the WRF-Noah modeling framework, the land surface characteristics (e.g., albedo and roughness length) are primarily functions of land cover types. The current WRF-Noah modeling framework considers only the dominant land cover type within each grid cell (i.e., it uses the dominant approach). Given the very wide user base of WRF-Noah and its increasing adoption for numerous climatic and environmental applications [*Rasmussen et al*., 2011; *Talbot et al*., 2012; *Trier et al*., 2004, 2008, 2011; *Wu et al*., 2011] , including application over highly heterogeneous urban settings [*Chen et al*., 2011b; *Li and Bou-Zeid*, 2013; *Li et al*., 2013; *Ntelekos et al*., 2008; *Yeung et al*., 2011; *Zhang et al*., 2009], it has become imperative to develop a mosaic/tiling approach within the WRF-Noah framework and to assess its performance.

[6] Urban environments provide a good test bed for this mosaic/tiling approach due to their considerable surface heterogeneities and the substantial differences between different surface types (e.g., impervious surface and vegetated surface). In this study, a specific focus is placed on the performance of the mosaic/tiling approach as compared to the dominant approach in urban settings, including sprawling suburbs. The hypothesis to be tested is the following: Despite the high resolution used in numerical simulations over urban environments (~1 km), subgrid-scale variability of land surface characteristics continues to be important and should be accounted for. The availability of high-resolution ($\ll$ 1 km) land cover data sets allows assessment of this hypothesis [*Fry et al*., 2011].

[7] The paper is organized as follows: section 2 introduces the WRF-Noah modeling framework and the experimental data sets; section 3 presents the WRF-simulated results with the dominant approach and the mosaic/tiling approach, focusing on the comparison between WRF simulations and observations. Section 4 discusses the different sensitivities of WRF-simulated results with the dominant approach and the mosaic/tiling approach to the spatial resolution of WRF simulations and to the number of tiles considered in a grid cell (N). Section 5 concludes the paper and discusses its implications. The mosaic/tiling approach will hereafter be simply referred to as the mosaic approach.

Table 1. The Land Use Fractions Simulated at the Cub Hill Meteorological Tower Location When the Dominant and Mosaic Approaches are Used

	Urban (Impervious)	Urban (Vegetated)	Cropland[a]	Forest[b]	Shrubland
Dominant	50%	50%	0%	0%	0%
Mosaic	20%	20%	30%	28%	2%

[a]Including dry land/cropland and mixed dryland/irrigated cropland.
[b]Including deciduous broadleaf forest and mixed forests.

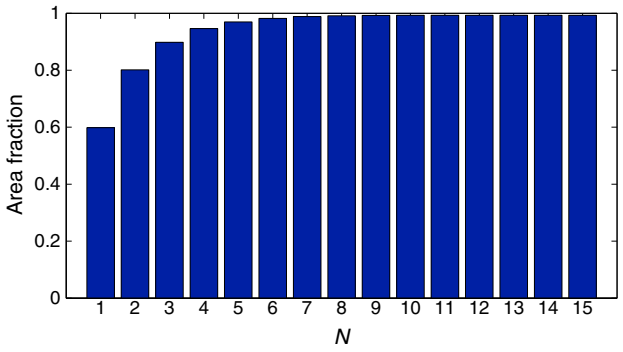

Figure 2. The mean area fractions occupied by N tiles averaged over domain 3.

2. Methodology and Experimental Data Sets

2.1. WRF-Noah Modeling Framework and the Mosaic Approach

[8] WRF is a numerical weather prediction system that has been designed to serve both research and forecasting purposes. It has been widely used for numerous regional climate and environmental research applications (see *Chen et al.* [2011b] for a review of its applications in urban environments). It is also used for operational forecasting at National Centers for Environmental Prediction, the Air Force Weather Agency, and other centers. WRF has several land surface models coupled to it in order to parameterize land surface processes and to provide lower boundary conditions. The Noah land surface model is the most widely used land surface model in WRF [*Chen and Dudhia*, 2001] and it is also used in other numerical weather prediction and climate models (e.g., the Eta model) [see *Ek et al.*, 2003].

[9] As noted in section 1, the current WRF-Noah modeling framework considers only the dominant land cover type within each grid cell. Given that land surface characteristics are principally functions of land cover types, this dominant approach neglects the subgrid-scale variability of land surface characteristics. In order to assess the importance of this variability, a mosaic approach is developed and implemented in this study. The mosaic approach allows users to specify the number of tiles to be considered within each grid cell and calculates surface fluxes and surface state variables for each tile using its surface properties such as albedo and emissivity. The atmospheric properties and soil properties are assumed to be homogenous over the grid cell when surface fluxes and surface state variables are calculated for each tile [*Avissar and Pielke*, 1989] (see *Bertoldi et al.* [2007, 2008] for a discussion of the impacts of this assumption). All prognostic variables are maintained for each tile, some of which are aggregated to yield the grid cell average variables as follows:

$$\overline{\phi} = \sum_{1}^{N} \phi_i A_i^{'}, \tag{1}$$

$$A_i^{'} = A_i \Big/ \sum_{1}^{N} A_i, \tag{2}$$

where i indicates the rank of the tile and N is the number of tiles considered in the grid cell. $A_i^{'}$ is the normalized area fraction while A_i is the area of the ith tile. The tile with the largest normalized area fraction has a rank of 1. ϕ_i represents a surface flux or surface state variable of the ith tile and $\overline{\phi}$ is the grid cell average of that variable. In order to simplify our analysis, N is a constant so each grid cell includes the same number of tiles. Nevertheless, the value of N can be modified by the user. Note that if M land cover types are available in a given land cover data set, N has to be less than or equal to M.

[10] The key grid cell averaged surface variables needed in the mosaic approach for coupling WRF to the Noah land surface model are the following: sensible heat flux, latent heat flux, surface temperature, emissivity, albedo, and momentum roughness length. Other variables, such as soil moisture, are required for each tile, but their grid cell averaged values are not needed. The grid cell averaged surface temperature and emissivity are needed in order to provide the grid cell averaged outgoing longwave radiation, while the grid cell averaged albedo is needed to provide the outgoing shortwave radiation. These grid cell averaged outgoing radiation components are equal to the sum of outgoing components from all tiles, but their computation as grid cell averages simplifies the implementation of the mosaic approach in WRF. The grid cell averaged momentum roughness length is required in order to calculate the grid cell averaged turbulent transfer coefficients, which are needed for the calculation of surface momentum and heat fluxes. The calculations of grid cell averaged surface temperature and roughness length are different from equation (1). For the surface temperature, in order to conserve the outgoing longwave radiation (and hence conserve energy), the grid cell averaged surface temperature is calculated in the following:

$$\overline{T} = \left(\sum_{1}^{N} \varepsilon_i T_i^4 A_i^{'} \Big/ \sum_{1}^{N} \varepsilon_i A_i^{'} \right)^{1/4}, \tag{3}$$

where ε is the emissivity and T is the surface temperature. For the momentum roughness length (z_o), given that its natural logarithm is usually used, its grid cell averaged value is calculated in the following:

$$\overline{z_o} = \exp\left(\sum_{1}^{N} \ln(z_{oi}) A_i^{'} \Big/ \sum_{1}^{N} A_i^{'} \right). \tag{4}$$

[11] More sophisticated models for the average momentum roughness length have been developed [*Bou-Zeid et al.*, 2004, 2007] and could be later implemented, but this simple logarithmic average is sufficient for this assessment. It is also noted that the thermal roughness length (z_{oT}) and water vapor roughness length (z_{oq}) are parameterized in WRF as functions of the momentum roughness length and other variables such as the roughness Reynolds number [see, e.g., *Chen and Zhang*, 2009; *Zilitinkevich*, 1995]; hence, in the mosaic approach, grid cell averaged scalar roughness lengths are calculated based on the grid cell averaged momentum roughness length from equation (4).

2.2. Study Area and WRF Configuration

[12] The study area is the Baltimore-Washington Corridor, which is a heavily urbanized area, as shown in Figure 1. It includes complex land surface features, sprawling and heterogeneous suburbs, mountains to the west, and land-water boundaries to the east. These land surface features are typical of urban settings of the eastern United States, as well as other

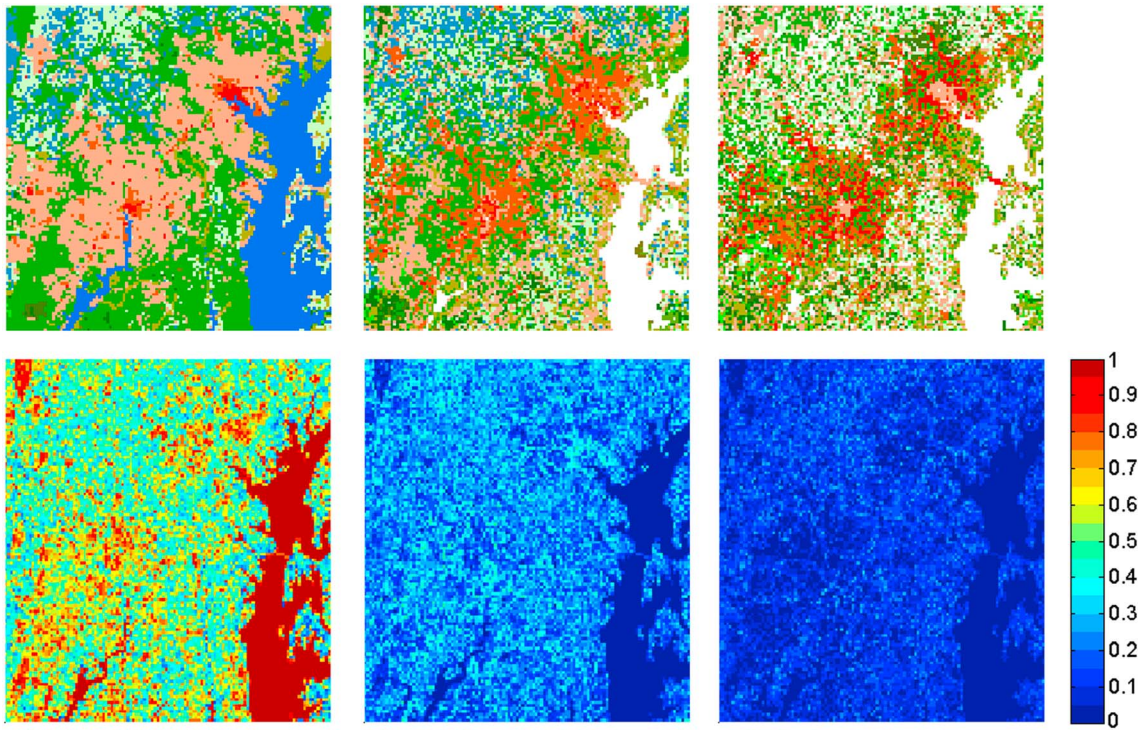

Figure 3. (top) The land cover types of the three dominant tiles within each grid cell in the 1 km domain 3 (the legend for land cover types follows the one used in Figure 1; blank areas indicate that there are no second or third tiles). (bottom) The area fractions of these three dominant tiles within each grid cell in domain 3.

urban environments around the world. This highly urbanized and heterogeneous landscape provides an excellent test bed for the mosaic approach. Figure 1 also shows the land cover map of the region. This land cover information is obtained from the 30 m resolution National Land Cover Dataset of 2006 (NLCD 2006) [*Fry et al*., 2011]. The land cover types are reclassified following the U.S. Geological Survey land cover classification system.

[13] WRF simulations are conducted using three nested domains with horizontal grid spacing of 9 km, 3 km, and 1 km, respectively. All domains use the mosaic approach when it is applied or all use the dominant approach. The three domains have 100, 100, and 121 grid cells, respectively, in both east–west and north–south directions. In the vertical direction, 28 levels are used. The domain configuration is similar to previous studies [*Li and Bou-Zeid*, 2013; *Li et al*., 2013] that have already examined the land surface processes, boundary layer dynamics, and the water cycle in this region. As such, the results of our study can be compared to those of previous applications.

[14] The WRF version 3.5 is used with the following physical parameterization schemes: (1) the Dudhia scheme for shortwave radiation; (2) the Rapid Radiative Transfer Model scheme for longwave radiation; (3) the 2-D Smagorinsky scheme for horizontal mixing; (4) the Mellor-Yamada-Janjic planetary boundary layer scheme; and (5) the Noah land surface model for nonurban surfaces and the original single-layer urban canopy model (UCM) of WRF for urban surfaces. When the dominant approach is used, the UCM is applied to any grid cell that is dominated by an urban category (i.e., the low-density residential, high-density residential, and industrial/commercial; see Figure 1); when the mosaic approach is used, the UCM is only applied to the urban tile instead of the whole grid cell, regardless of its rank within the grid cell. Note that when the UCM is used, an urban grid cell (the dominant approach) or an urban tile (the mosaic approach) is further separated into an impervious fraction and a grass-covered fraction, following the same approach described by *Chen et al*. [2011b] (see section 3.1.1 and Table 1 for an example). This UCM grass fraction is intended to represent urban parks and lawns and thus captures very small scale variability inside the built terrain. It is thus distinct from and does not affect the larger subgrid-scale variability represented by our mosaic implementation. Cumulus parameterization is not used for any of these domains given that most of our analyses are conducted in domain 3 [*Li et al*., 2013; *Yeung et al*., 2011]. In addition, one-way nesting is used since most of our analyses focus on the innermost domain. The initial and boundary conditions for WRF simulations are taken from the North American Regional Reanalysis.

2.3. Surface Characterization

[15] The default number of tiles considered in each grid cell (N) when the mosaic approach is used is 8, unless otherwise specified when the sensitivity to this number is investigated. $N = 8$ is chosen for most analyses because the sum of mean area fractions occupied by eight tiles is over 99%, as can be seen from Figure 2 that shows the mean area fractions covered by different numbers of tiles for the innermost, smallest domain. In order to illustrate the land cover composition and area fractions of different ranks of tiles, Figure 3 depicts the land cover types of the three most abundant tiles and their area fractions within each grid cell in domain 3 (d03). Note that the

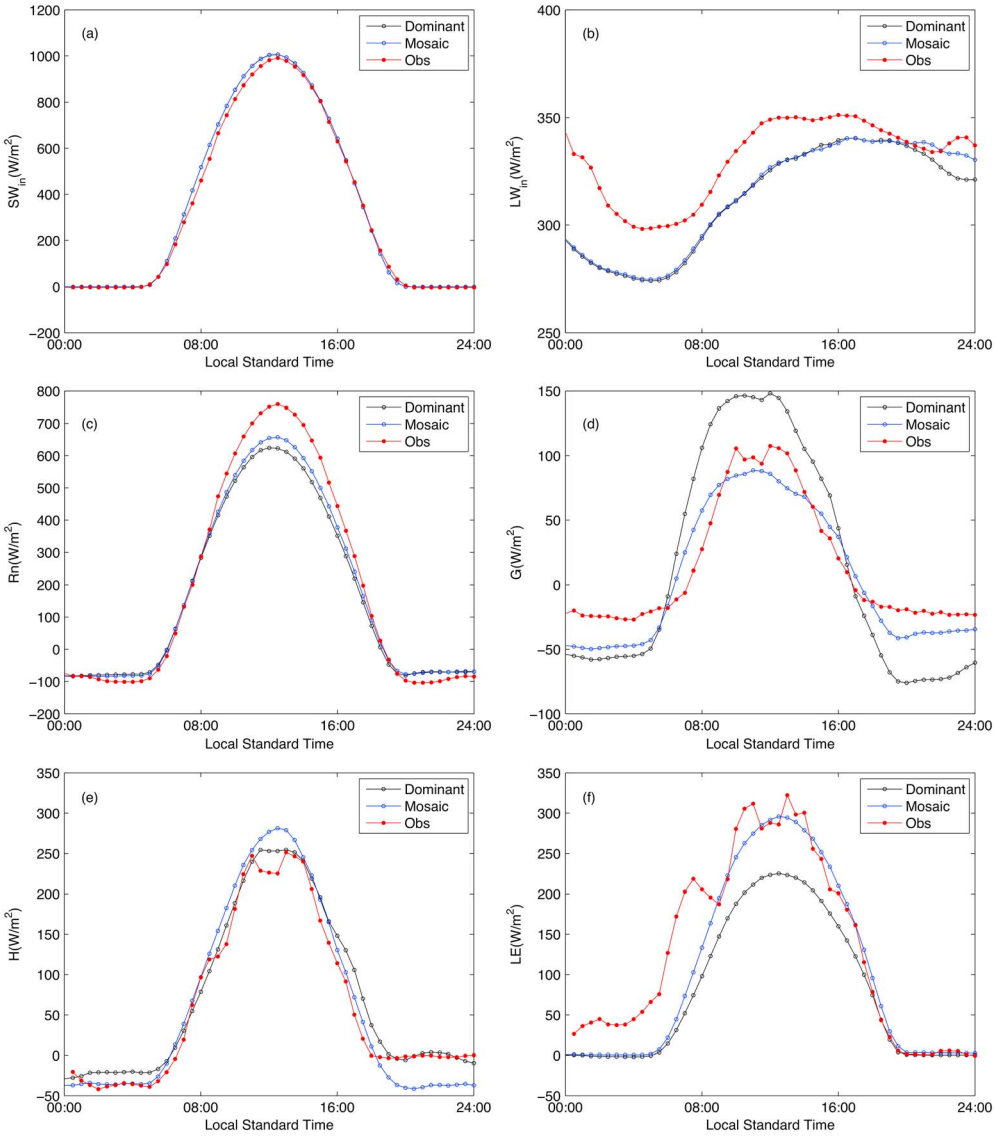

Figure 4. Comparison of the WRF-simulated and measured surface fluxes at the Cub Hill meteorological tower on 14 July 2009: (a) incoming shortwave radiation, (b) incoming longwave radiation, (c) net radiation, (d) ground heat flux, (e) sensible heat flux, and (f) latent heat flux.

area fractions shown in Figure 3 are not the normalized area fractions so that the sum is not unity.

[16] It is clear that the spatial pattern of the first tile within each grid cell is more organized compared to the spatial patterns of the other two tiles. Around the urban cores of Washington D.C. and Baltimore, the first tile is generally an urban land cover (i.e., low-density residential, high-density residential, or industrial/commercial). Nevertheless, the second tile and the third tile show more variability in land cover types. Some grid cells have the first tile as an urban surface, but the second or the third tile as other types of surface. Given that the impervious (urban) surface behaves strikingly differently from the vegetated surfaces (e.g., grassland or forest), it is expected that by considering the second tile and the third tile (and more tiles) within a grid cell whose first tile is an urban tile, the surface energy balances calculated from the mosaic approach will be significantly different from those calculated from the dominant approach.

2.4. Case Descriptions

2.4.1. Case 1: A Clear Day (14 July 2009)

[17] The first case is a clear day on 14 July 2009, where the large-scale weather conditions are characterized by a high-pressure system over the Baltimore-Washington metropolitan area. This case is selected to examine the performances of the dominant approach and the mosaic approach in capturing the surface fluxes and surface temperatures. The simulation starts at 0000 UTC on 14 July and ends at 1200 UTC on 15 July.

2.4.2. Case 2: A Rainfall Period (21–26 July 2008)

[18] The second case concerns a rainfall period from 21–26 July 2008. This case has been studied in *Li et al.* [2013]. The large-scale weather conditions are characterized by a cold front moving from the northwest toward the southeast of the Baltimore-Washington metropolitan area. The simulation starts at 0000 UTC on 21 July and ends at 0000 UTC on 26 July 2008. It is roughly decomposed into three periods: a prestorm

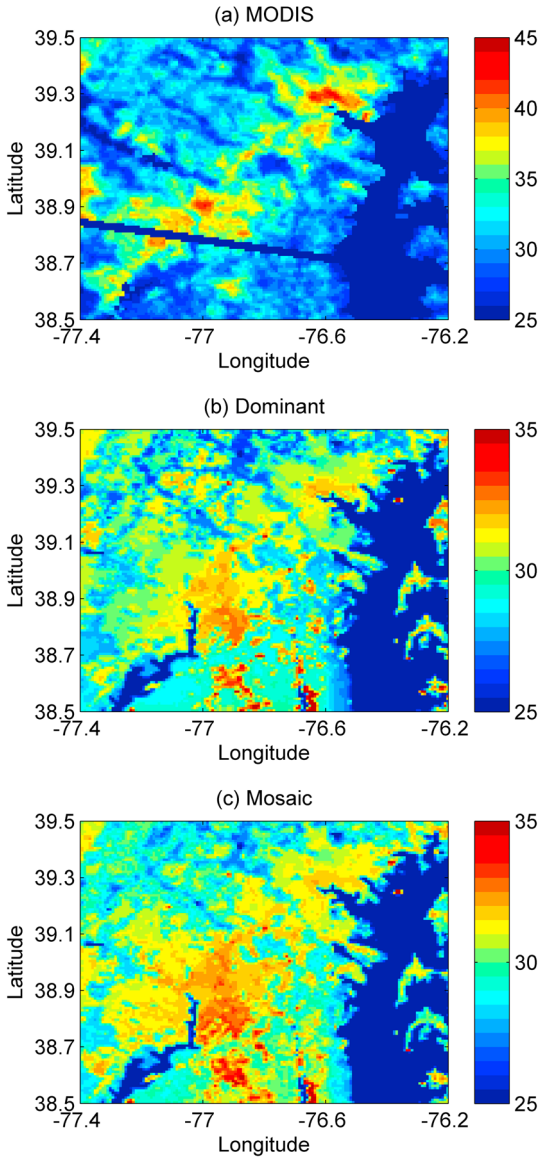

Figure 5. Land surface temperature maps from (a) MODIS observations, (b) WRF simulation with the dominant approach, and (c) WRF simulation with the mosaic approach at about 1230 P.M. local time on 14 July 2009.

period (21 July), a major convective rainfall period (22–24 July), and a dry down period (25–26 July). The maximum rainfall accumulation between 23 July (1200 UTC) and 24 July (1200 UTC) exceeds 100 mm. This case is selected to examine the different impacts of the dominant and mosaic approaches on boundary layer growth and rainfall distribution.

2.5. Experimental Data Sets

[19] The extensive observational data available in this region provide a strong foundation for evaluating the performance of the mosaic approach. Since the variables that are modified at the surface when the mosaic approach is adopted include sensible heat flux, latent heat flux, ground heat flux, surface temperature, albedo, and roughness length, observational data sets of surface fluxes and surface temperatures are primarily used to assess the performance of the mosaic approach compared to the dominant approach. Given that these changes will also induce changes in atmospheric dynamics and the hydrological cycle, vertical profiles of potential temperature in the atmospheric boundary layer and radar rainfall estimates are also used.

[20] The surface flux measurements are taken from the Cub Hill meteorological tower (39.413°N, 76.522°W). Four components of surface radiation (i.e., the incoming shortwave radiation, the outgoing shortwave radiation, the incoming longwave radiation, and the outgoing longwave radiation) are available. Sensible heat flux and latent heat flux are calculated from turbulent measurements of wind speeds, air temperature, and water vapor concentration following the same computation procedure as *Li and Bou-Zeid* [2011]. Ground heat flux is also measured by soil heat flux plates. Details

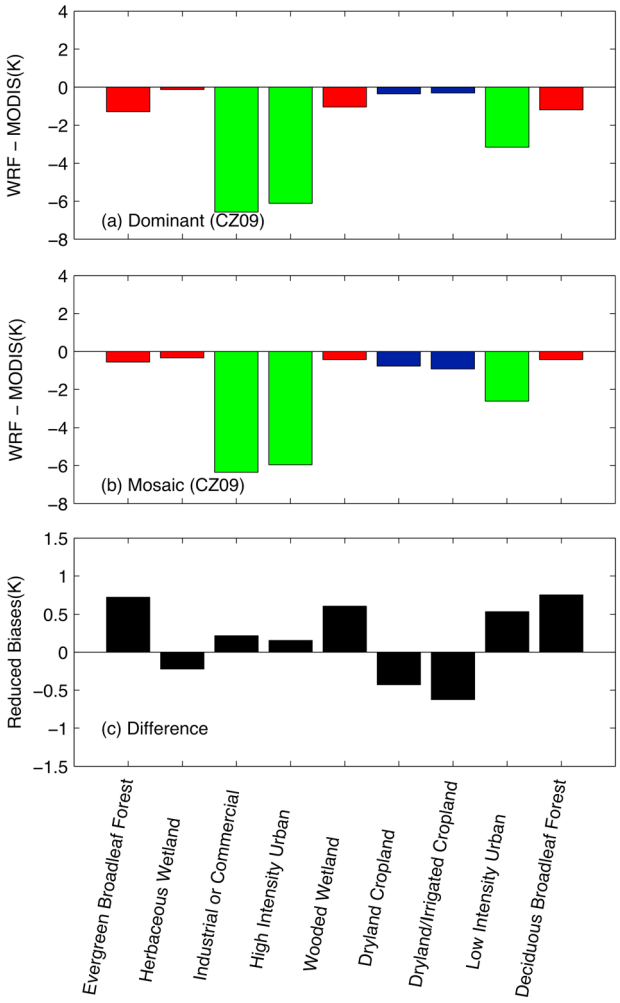

Figure 6. Differences in land surface temperature between WRF simulations and MODIS observations as a function of dominant land cover categories: (a) when the dominant approach is used; (b) when the mosaic approach is used. The red color denotes tall-canopy land use categories, while the blue color denotes short-canopy land cover categories. The urban categories are in green color. (c) The reduced bias shown is defined as $|\text{Bias}|_{\text{dominant}} - |\text{Bias}|_{\text{mosaic}}$.

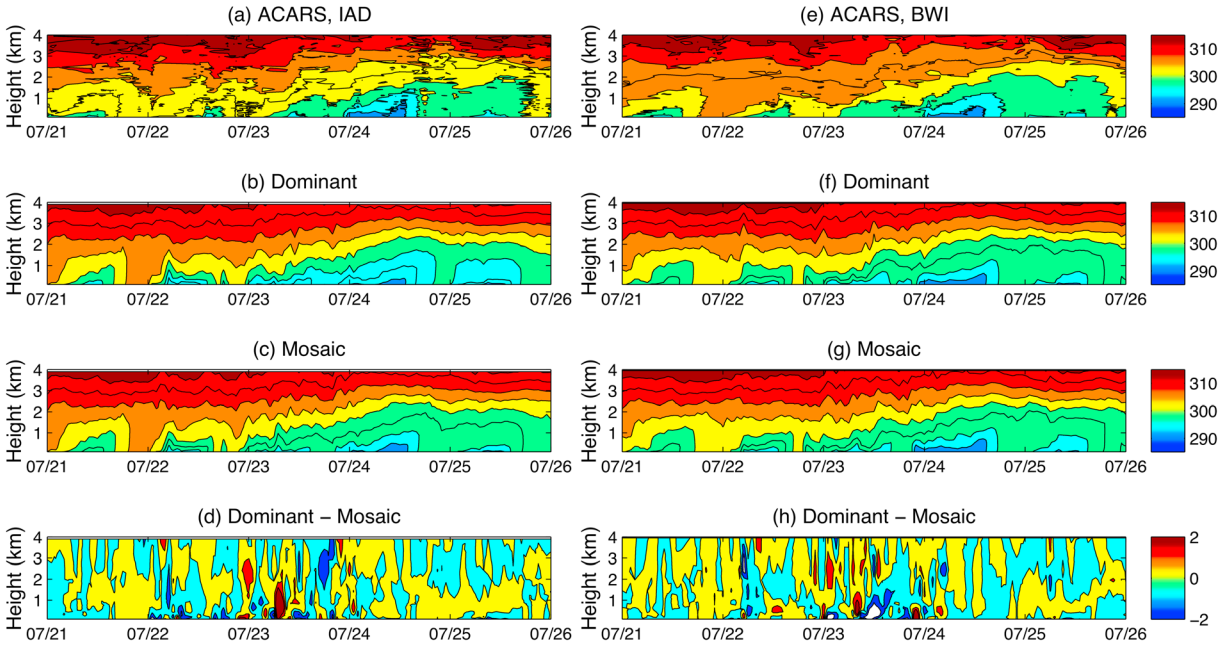

Figure 7. Potential temperature (units: K) in the lower atmosphere (up to 4 km above the ground level) at the (a–e) Dulles International Airport (IAD, left panels) and the (e–h) Baltimore/Washington International Airport (BWI, right panels) from 21 July to 26 July 2008. Figures 7a and 7e are aircraft measurements and Figures 7b, 7c, 7f and 7g are WRF-simulated results with the dominant approach and the mosaic approach, respectively. Figures 7d and 7h are the differences between the dominant approach and the mosaic approach.

of the Cub Hill site and measurements can be found in *Crawford et al*. [2011].

[21] The land surface temperature product from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite observations is used to assess the WRF-simulated land surface temperatures [*Leroyer et al*., 2011; *Zhang et al*., 2011]. The MODIS product used in our study is the MYD11A1 version 5 level 3 product. The cloud-screened MODIS surface skin temperature is usually available twice a day, once during daytime and once during nighttime. In this study, only the daytime MODIS surface temperature is used due to its good quality.

[22] In order to evaluate the impact of the mosaic approach on the boundary layer growth, we use vertical profiles of potential temperature in the lower atmosphere (up to about 4 km above the ground) measured through commercial aircraft-mounted sensors from the Aircraft Communications Addressing and Reporting System (ACARS). These measurements are available at the Dulles International Airport (IAD; see Figure 1) and the Baltimore-Washington International Airport (BWI; see Figure 1); they are interpolated at hourly intervals and at 50 m vertical intervals to facilitate the comparison to WRF-simulated potential temperature profiles.

[23] A long-term (2001–2012), high-resolution radar rainfall data set has been produced for the Baltimore-Washington area [*Smith et al*., 2012]. It is largely based on reflectivity observations from the Weather Surveillance Radar, 1988 Doppler radar in Sterling, Virginia (KLWX). The reflectivity observations are first converted to rainfall rate through the National Weather Service "Z-R relationship" and are then bias corrected using rain gauge observations [*Smith et al*., 2012]. The radar rainfall estimates have been used in previous studies [*Li et al*., 2013] and will be also employed in this study to assess the impact of the mosaic approach on rainfall modeling.

3. Model Validation and the Implications of Representing Subgrid-Scale Variability

3.1. Case 1: A Clear Day (14 July 2009)

3.1.1. Surface Fluxes

[24] The Cub Hill meteorological tower is located within a grid cell whose dominant tile is a low-density residential urban land cover with an area fraction of 40%. Dry land/cropland occupies about 25% of the grid cell, while deciduous broadleaf forest occupies about 22%. The remainder includes about 6% mixed forests, 5% mixed dryland/irrigated cropland, and 2% shrubland. Note that there are only six tiles at Cub Hill in the NLCD 2006 data set despite that $N = 8$ is used when the mosaic approach is adopted. As such, only the dominant tile is an urban tile while the other tiles are all vegetated tiles. If only the dominant tile is considered (i.e., the dominant approach), the whole grid cell is treated as low-density residential urban. When the mosaic approach is used, more vegetated tiles are included. It is again noted that when a UCM is used, a grid cell or a tile that is identified as an urban land will be separated into two parts: an impervious part and a grass-covered part [*Chen et al*., 2011b]. Table 1 presents the land use fractions considered at the Cub Hill meteorological tower location when the dominant and mosaic approaches are used. As can be seen from Table 1, this grid point is in fact mostly covered by vegetation.

[25] Figure 4 compares the WRF-simulated surface fluxes using the dominant and mosaic approaches to the measurements at the Cub Hill meteorological tower location. The

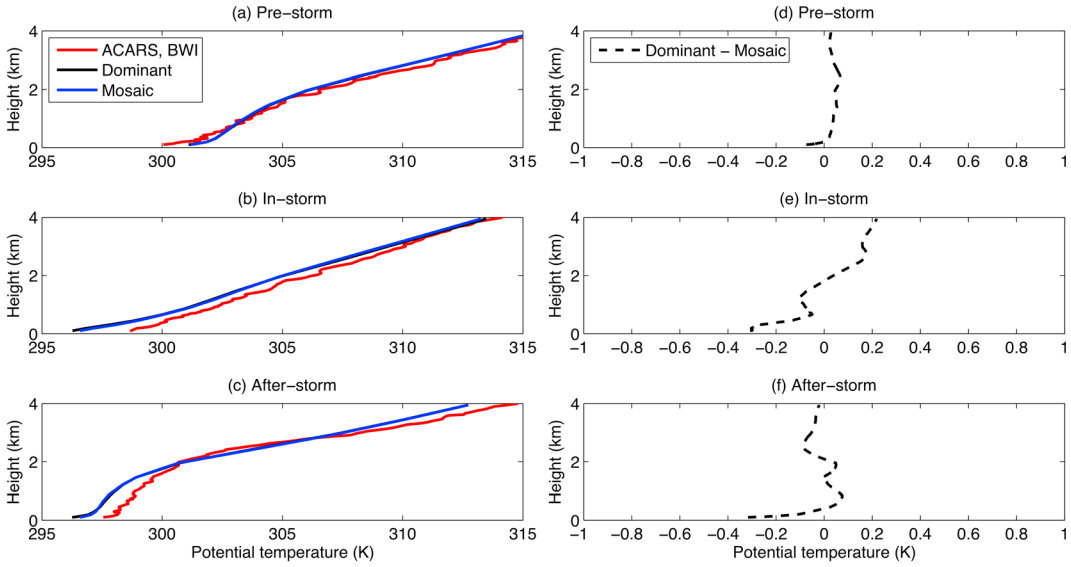

Figure 8. Composite profiles of (a–c) potential temperature at BWI and the (d–f) differences between the dominant approach and the mosaic approach in the lower atmosphere (up to 4 km above the ground level) for prestorm (0000 UTC, 21 July to 0000 UTC, 22 July), in-storm (0000 UTC, 22 July to 1200 UTC, 24 July), and after storm conditions (1200 UTC, 24 July to 0000 UTC, 26 July).

fluxes considered here include incoming shortwave radiation (Figure 4a), incoming longwave radiation (Figure 4b), net radiation (Figure 4c), ground heat flux (Figure 4d), sensible heat flux (Figure 4e), and latent heat flux (Figure 4f). As can be seen from Figures 4a and 4b, the dominant and mosaic approaches yield very similar incoming shortwave and longwave radiation. This is expected since the incoming shortwave and longwave radiation are more dependent on atmospheric conditions than on surface conditions. Subtle differences between the two approaches are observed for the incoming longwave radiation from 16:00 to 24:00, which is caused by the changes in the atmospheric conditions due to changes in the surface fluxes. The WRF-simulated incoming shortwave radiation matches the measurements fairly well, while the WRF-simulated incoming longwave radiation shows biases of approximately 50 W/m^2 compared to the measurements.

[26] The differences between the two approaches are more evident for the net radiation (~ 30 W/m^2), the ground heat flux (~ 70 W/m^2), the sensible heat flux (~ 20 W/m^2), and, most notably, the latent heat flux (~ 100 W/m^2); these values listed between parentheses are the maximum differences between the two approaches over the simulated diurnal cycle. Compared to the dominant approach, the mosaic approach clearly shows better performance in reproducing the net radiation, ground heat flux, and latent heat flux and shows a comparable performance in reproducing the sensible heat flux. In particular, the latent heat flux is significantly increased during daytime with the mosaic approach due to the representation of the vegetated tiles (e.g., dry land/cropland and deciduous broadleaf forest) and thus matches observation better. Note that most of today's urban models have difficulty in reproducing observed latent heat fluxes in cities [*Chen et al.*, 2012; *Grimmond et al.*, 2011], and hence, the mosaic approach has the potential to improve simulations of latent heat fluxes in urban areas. The significant improvements in the simulated ground heat flux (Figure 4d) and latent heat flux (Figure 4f) illustrate that the mosaic approach can capture the surface energy balance more realistically than the dominant approach. The relatively small differences in the simulated sensible heat flux between the two approaches do not imply that the sensible heat fluxes from different tiles are identical. In fact, the eight tiles generate distinct diurnal cycles of sensible heat flux, which are also different from the diurnal cycle of sensible heat flux when the dominant approach is used (not shown here). The comparable sensible heat fluxes are rather due to an accidental match between the flux from the dominant approach and the average flux from the eight tiles of the mosaic approach.

[27] Despite the fact that the mosaic approach shows a better performance than the dominant approach, certain biases still exist in the simulated net radiation. Close inspection reveals that this is due to the fact that the albedo value in the simulations is significantly larger than the measured one, which has also been observed by *Li et al.* [2013]. This implies that an accurate characterization of surface and soil properties remains a key challenge for numerical weather prediction models.

3.1.2. Land Surface Temperatures

[28] Section 3.1.1 relies on measurements at a single location. In this section, the simulated land surface temperature pattern over the whole d03 is compared to satellite observations to yield a more representative assessment of model performance. Figure 5a shows the observed land surface temperature pattern from MODIS at about 12:30 P.M. local time. It is clear that the two cities, namely, Washington D. C. and Baltimore, are significantly hotter than the surrounding rural areas. This reflects the urban heat island effect [*Grimmond*, 2007; *Oke*, 1982; *Shephard*, 2005], which can be also seen in the WRF simulations (Figures 5b and 5c). Note that the 1 km pixel satellite observations are comparable to outputs

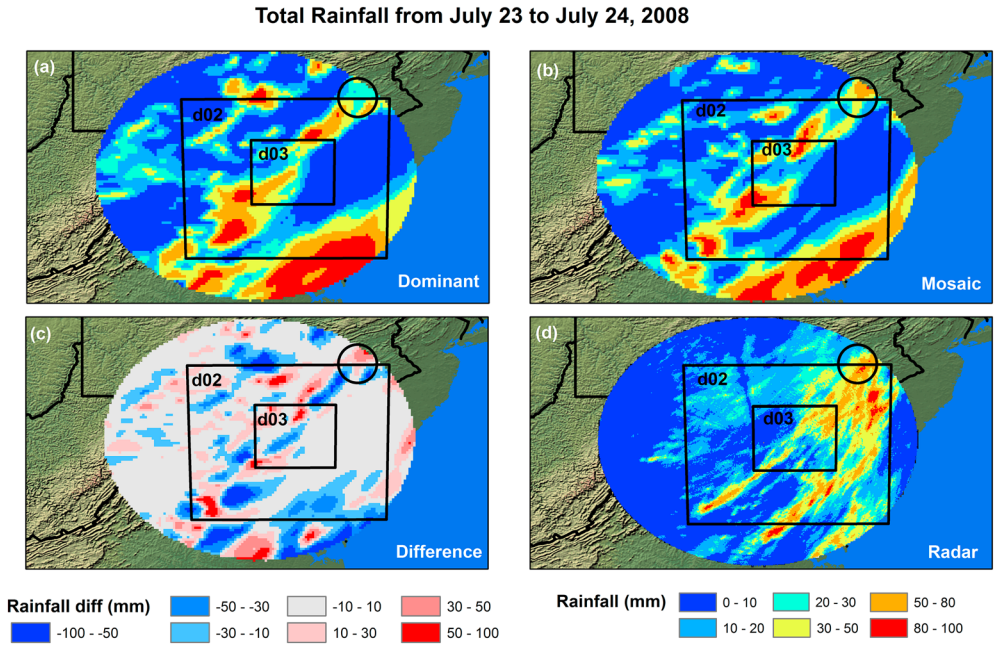

Figure 9. Total rainfall from 23 July to 24 July (12 UTC to 12 UTC) produced by WRF simulations with the (a) dominant approach and (b) the mosaic approach. (c) The differences between the two approaches are shown and (d) measured by the KLWX radar.

from the WRF d03 with a grid spacing of 1 km. As such, the differences between WRF-simulated land surface temperatures with two approaches and satellite-observed land surface temperatures can be quantified.

[29] Previous studies have found that the parameterization of the thermal roughness length (z_{oT}) is an important controlling factor of simulated land surface temperatures during daytime [*Zeng et al*., 2012; *Zheng et al*., 2012]. It also affects surface fluxes and boundary layer growth [*Chen et al*., 1997; *LeMone et al*., 2008, 2010a, 2010b], as well as warm-season precipitation [*Trier et al*., 2011]. *Li and Bou-Zeid* [Quality and sensitivity of high-resolution numerical simulation of urban heat islands, submitted to *Environmental Research Letters*, 2013] shows the modified Zilitinkevich relationship [*Chen and Zhang*, 2009, hereafter CZ09] yields smaller biases compared to the MODIS-observed land surface temperatures, particularly for tall-canopy land use categories (Figure 6, red colors). Therefore, the modified Zilitinkevich relationship (CZ09) is chosen to parameterize the thermal roughness length in this study as well.

[30] Figure 6 shows the domain-averaged differences between WRF-simulated and MODIS-observed land surface temperatures as a function of the dominant land cover category in a grid cell. Figure 6a shows the biases when the dominant approach is used, while Figure 6b shows the biases when the mosaic approach is used. Since in the dominant approach a grid cell is solely composed of the dominant land cover type, the biases shown on Figure 6a are truly the biases from the dominant land cover categories indicated on the *x* axes. This is however not the case when the mosaic approach is used (Figure 6b). Given that each grid cell includes eight land cover types ($N = 8$), the biases shown on Figure 6b also include biases from other land cover categories (ranks 2 to 8) besides the dominant land cover categories indicated on the *x* axes. Note that the land cover categories in the figures are ordered such that, from left to right on the *x* axes, the number of grid cells dominated by this land cover category in d03 increase (the most common are to the far right). More than 60% of the grid cells in d03 are dominated by low-density residential urban or deciduous broadleaf forest.

[31] Since the biases in the surface temperature fields generated by the mosaic approach are different from those generated by the dominant approach, the differences in the absolute biases from the two approaches are quantified and shown in Figure 6c. The reduced biases on the *y* axes are defined as the absolute biases from the dominant approach minus the absolute biases from the mosaic approach: $|\text{Bias}|_{\text{dominant}} - |\text{Bias}|_{\text{mosaic}}$. Thus, when the reduced biases are positive, the mosaic approach generates smaller biases than the dominant approach and vice versa. As one can see, the reduced biases are positive for six of the nine land cover types shown here. In particular, the reduced biases in the two most abundant dominant land cover types (i.e., low-density residential urban or deciduous broadleaf forest) are positive. If the average bias in the whole d03 is defined as the sum of biases from these dominant land cover types weighted by their area fractions, the reduced bias over the whole of d03 is 0.32°C, implying an improvement in the simulation of surface temperature with the mosaic approach. A note is needed here concerning the large biases that persist for urban land cover types. These are related to deficiencies in WRF's original UCM that were detailed and corrected by *Li and Bou-Zeid* (submitted manuscript, 2013) by implementation of the Princeton Urban Canopy Model (PUCM) [*Wang et al*., 2013]. Biases in urban surface temperatures were reduced to ~1.2°C by this improvement. However, so as to not confound the effects of the mosaic approach and of PUCM, and since PUCM is not available in the public release of WRF, we perform this study with the default UCM of WRF. The coupling of the implementations of

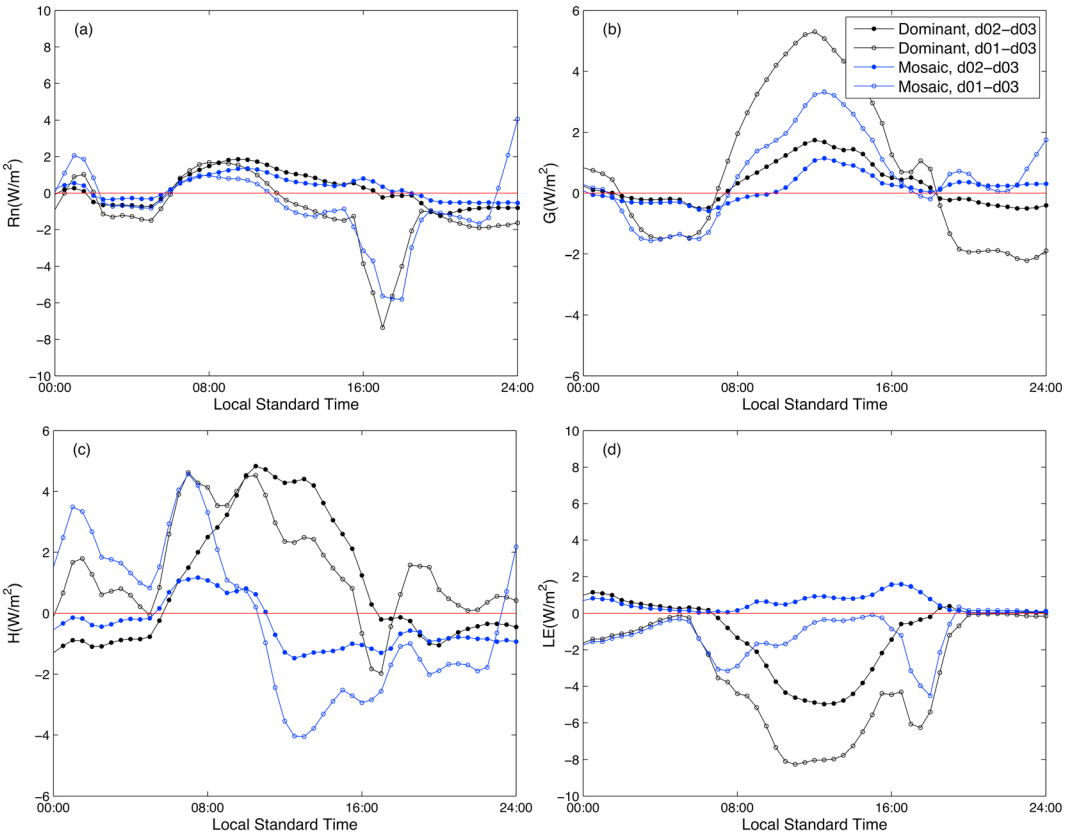

Figure 10. Comparison of WRF-simulated surface energy balances with two different approaches and three different resolutions (i.e., from d01, d02, and d03) on 14 July 2009: (a) net radiation, (b) ground heat flux, (c) sensible heat flux, and (d) latent heat flux. The black lines are with the dominant approach and the blue lines are with the mosaic approach. The filled dots denote the differences between results from d02 and results from d03, while the circles denote the differences between results from d01 and results from d03. Results shown here are averaged over selected grid cells of d03 that display high heterogeneity and exclude water surfaces.

PUCM and the mosaic approach in WRF is under way (the authors are happy to share the relevant subroutine with interested users).

3.2. Case 2: A Rainfall Period (21–26 July 2008)

[32] In this section, the performance of WRF with the dominant and mosaic approaches under rainfall conditions is evaluated. The focus is on the growth of the atmospheric boundary layer and on the rainfall distribution.

3.2.1. Atmospheric Boundary Layer Profiles

[33] Figure 7 compares the potential temperature field (up to 4 km) at the Dulles International Airport (IAD, left panels) and the Baltimore/Washington International Airport (BWI, right panels) during the simulation period from 21 July to 26 July. As can be seen from Figure 7, the potential temperature profiles in the lower atmosphere simulated by WRF with the two different approaches are similar. They also capture the overall evolution of potential temperature well when compared to the aircraft measurements. Some differences are observed between the two approaches, particularly during the rainfall period (23 July to 24 July), as illustrated in lowest row of panels of Figure 7. This confirms that changes in surface fluxes due to the adoption of the mosaic approach have a nonnegligible influence on the whole atmospheric boundary layer.

[34] To further quantify this influence and the biases in the WRF-simulated potential temperature profiles and the differences between the two approaches, the simulation period is separated into three phases: the prestorm phase (0000 UTC, 21 July to 0000 UTC, 22 July), the in-storm phase (0000 UTC, 22 July to 1200 UTC, 24 July), and the after storm phase (1200 UTC, 24 July to 0000 UTC, 26 July). In each phase, the mean potential temperature profiles from WRF simulations with the two different approaches are compared to the mean potential temperature profiles from the aircraft measurements, as shown in Figure 8. Here only the results at BWI are shown since the results at IAD are similar. During the prestorm, the two different approaches produce similar potential temperature profiles and are in fair agreement with the observations, with a slight warm bias close to the surface of about 1 K. However, more important differences between the two approaches, and larger biases for both, are observed in the atmospheric boundary layer (below 2 km) during the in-storm and poststorm phases. This is closely linked to the different rainfall patterns produced by the dominant approach and the mosaic approach, as shall be seen in Figure 9. During the in-storm phase, the biases are larger than those observed before, particularly in the atmospheric boundary layer where the biases are ~1.5 K. During the poststorm phase, the biases remain large in the atmospheric boundary

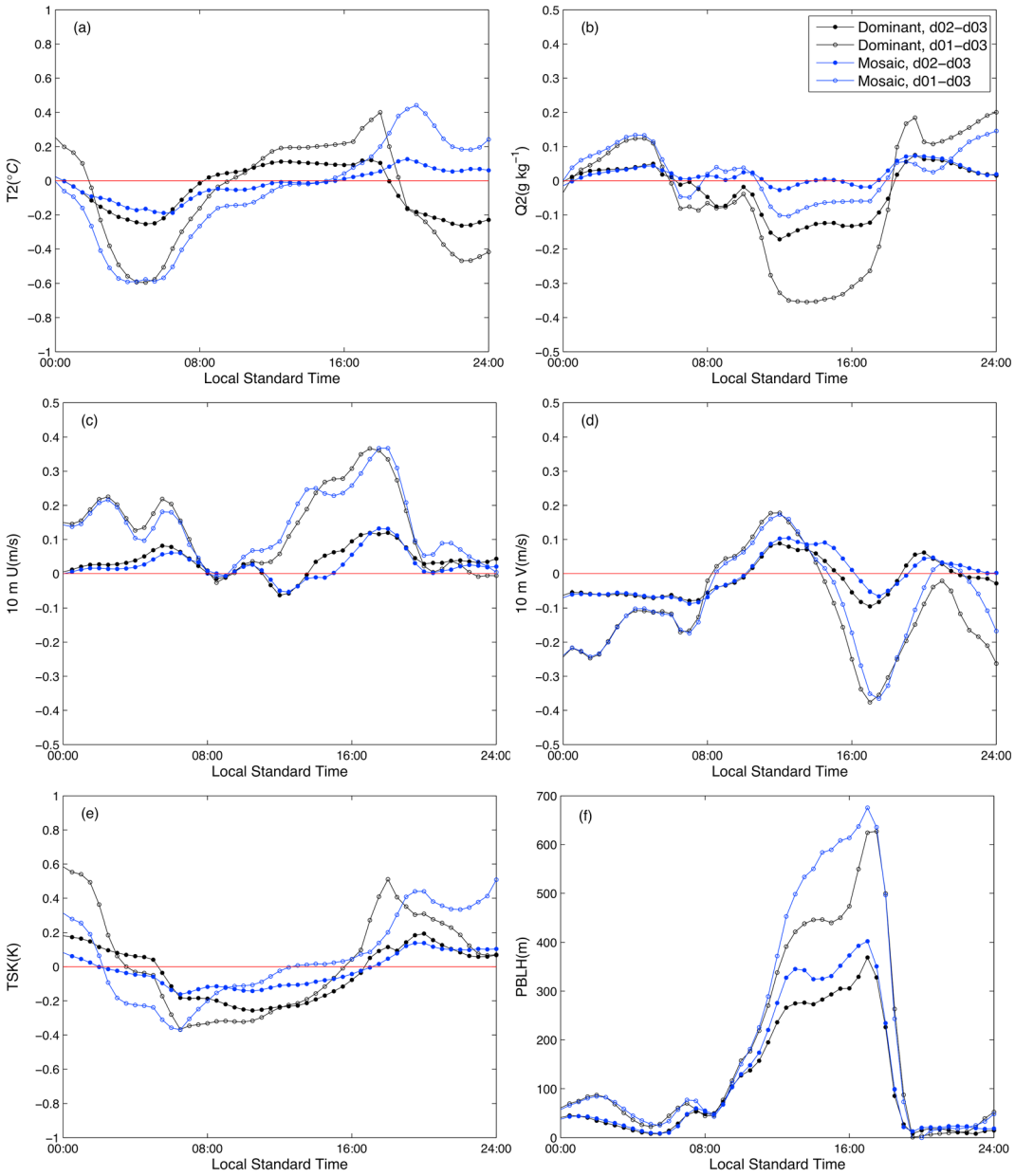

Figure 11. Comparison of WRF-simulated surface and near-surface states and boundary layer heights with two different approaches and three different resolutions (i.e., from d01, d02, and d03) on 14 July 2009: (a) 2 m air temperature, (b) 2 m specific humidity, (c) the east–west component of 10 m wind, (d) the north–south component of 10 m wind, (e) land surface temperature, and (f) boundary layer height. The black lines are with the dominant approach and the blue lines are with the mosaic approach. The filled dots denote the differences between results from d02 and results from d03, while the circles denote the differences between results from d01 and results from d03. Results shown here are averaged over selected grid cells of d03 that display high heterogeneity and exclude water surfaces.

layer. These results overall indicate that the mosaic approach produces different results than the dominant approach even during storm periods when a reduced influence of local surface fluxes on the atmospheric boundary layer dynamics is expected. They also suggest biases in WRF's skill during storm periods that appear to be unrelated to its surface models.

3.2.2. Rainfall Pattern and Distribution

[35] Previous studies have demonstrated that changes in calculation of surface fluxes can result in significant changes in the large-scale rainfall distribution over urban environments [*Chang et al.*, 2009; *Chen et al.*, 2011a; *Li et al.*, 2013; *Miao et al.*, 2011; *Niyogi et al.*, 2011; *Shephard*, 2005; *Yeung et al.*, 2011]. Figure 9 shows the simulated-storm total rainfall fields from WRF using the dominant approach (Figure 9a) and the mosaic approach (Figure 9b). The differences between the two approaches (mosaic-dominant) are shown in Figure 9c. It is clear that the two approaches lead to distinctly different rainfall distributions. The absolute maximum difference at a given location can

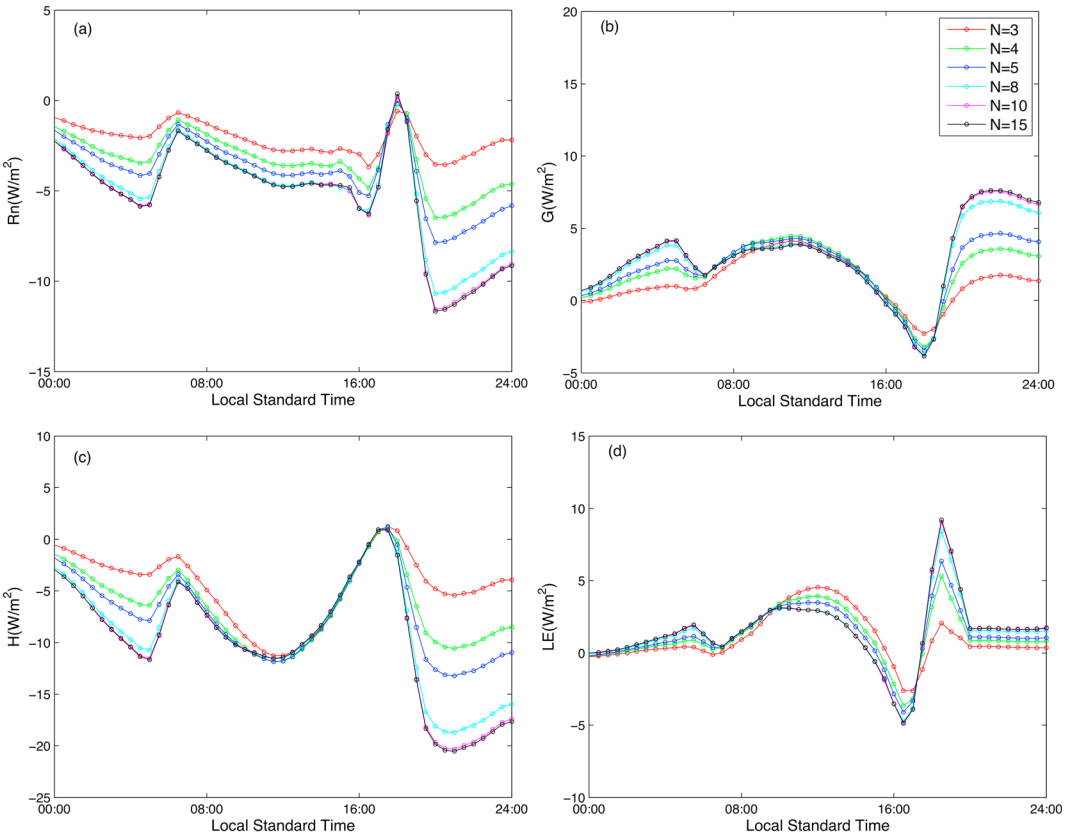

Figure 12. Differences in surface energy components: (a) net radiation, (b) ground heat flux, (c) sensible heat flux, and (d) latent heat flux, between simulations with different N values and the simulation with the dominant approach ($N = 1$) . Results shown here are averaged over selected grid cells of d03 that display high heterogeneity and exclude water surface (similar to Figures 10 and 11).

reach up to 100 mm (note the maximum total rainfall is about 120 mm). Figure 9d shows the rainfall distribution estimated from the radar measurements. Comparing the WRF-simulated and radar-estimated rainfall distributions suggests that the line structures are well reproduced by WRF; nevertheless, WRF generates too much rainfall on each line, which was also observed in previous studies [*Li et al*., 2013]. Another notable feature is that, at the maximum rainfall location shown in the radar estimates (i.e., the northeast corner of d02 as indicated by a black circle on the figure, which is downwind of Baltimore), the mosaic approach generates more rainfall as compared the dominant approach, and the resulting rainfall distribution from the mosaic approach is in closer agreement with the radar estimates. The root-mean-square errors calculated for WRF-simulated total rainfall fields are 34.7 mm and 30.7 mm for the dominant approach and the mosaic approach, respectively. The mean biases are 2.5 mm and 1.7 mm for the dominant approach and the mosaic approach, respectively. Note that in the calculation of root-mean-square errors, only places where rainfall is present in the radar estimates are included, while the whole area covered by radar is included in the calculation of mean biases. Hence, it is evident that large biases still exist in the WRF-simulated rainfall fields when the dominant is replaced by the mosaic approach, suggesting that these biases are likely associated with other parts of the WRF-Noah model or the initial/boundary conditions.

4. Grid Convergence Tests and Sensitivity to the Number of Subgrid-Scale Tiles

4.1. The Sensitivity of WRF-Simulated Results to the Spatial Resolution

[36] One possible weakness of the dominant approach is that the results are likely to change with spatial resolution since the dominant land cover type may change with resolution. For example, if a 3 km × 3 km box is composed of 60% grassland on the west and 40% urban land on the east, a WRF simulation with a grid spacing of 3 km will have only one grid cell over the box and will treat this grid cell as if it is composed of 100% grassland. Nevertheless, a WRF simulation with a grid spacing of 1 km will have nine grid cells over this box and probably six grid cells will be viewed as grassland and the remaining three grid cells on the east will be viewed as urban land. As such, the results from the simulation with 3 km grid spacing are not likely to match the results from the simulation with 1 km grid spacing. The mosaic approach can potentially reduce these discrepancies. In this section, the sensitivity of WRF-simulated results with the dominant approach and the mosaic approach to the spatial resolution is thus examined. This can be viewed as a grid convergence test where a numerical model runs at increasing resolutions until the results cease to change. This would indicate that resolutions finer than this limit are not needed. When complete convergence cannot be attained, such as with

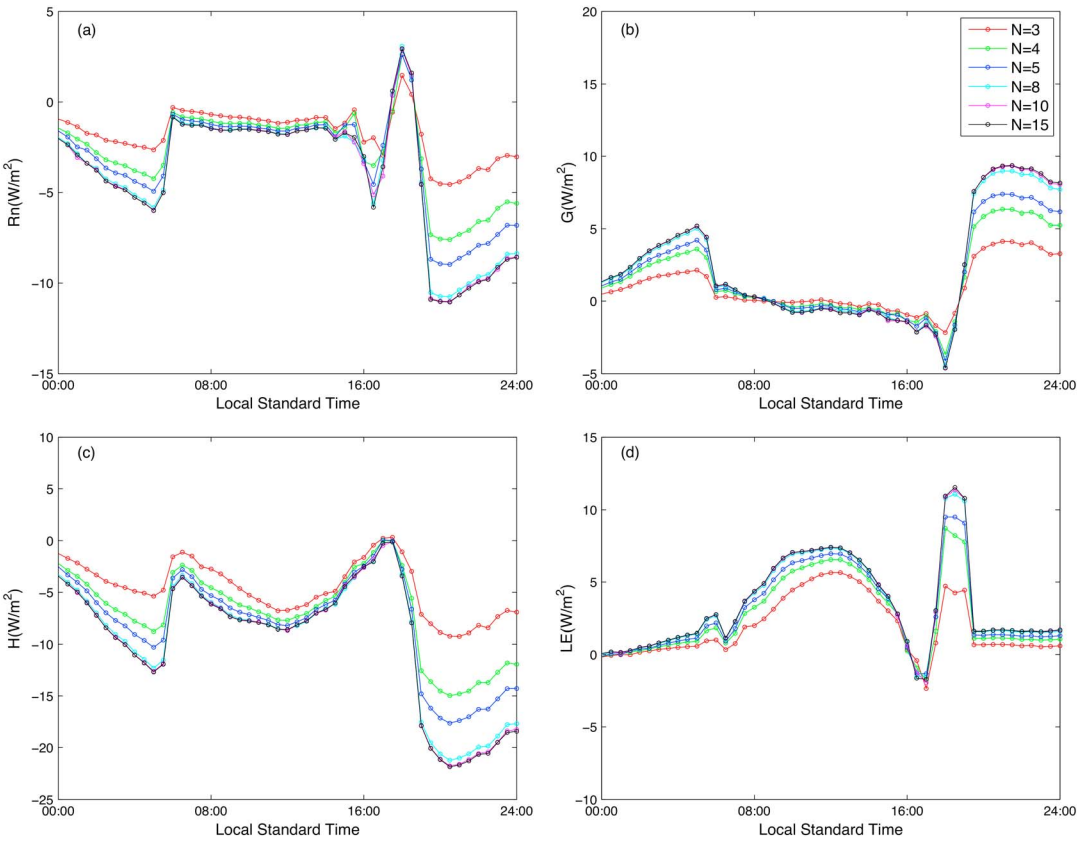

Figure 13. Similar to Figure 12 but the mosaic results are only averaged over the first tiles of those selected grid cells.

the WRF simulations performed here, a smaller change in the outputs with increasing resolution is a positive indicator since, ideally, numerical model should be implemented such that they do not depend on resolution.

[37] Figures 10 and 11 show the differences between the surface energy balances, near-surface states, surface temperatures, and boundary layer heights simulated using different resolutions; these differences are the averages over the area of d03 for the clear day case. The black lines are with the dominant approach and the blue lines are with the mosaic approach. The filled dots denote the differences between results from d02 and results from d03, while the empty circles denote the differences between results from d01 and results from d03. Note that grid cells dominated by water surfaces are excluded in the averaging since the mosaic approach is only applied to grid cells that are dominated by land. In addition, grid cells whose first tile has an area fraction larger than 70% are also excluded since we aim to focus on areas with significant heterogeneity. To make a consistent comparison, when a grid cell in d01 is excluded, the corresponding multiple grid cells in d02 and d03 are excluded. The final resulting area where the averaging takes place is about 9,600 km^2.

[38] The differences between d02 and d03 are overall smaller than those between d01 and d02 suggesting that WRF convergence improves with higher resolution. More importantly for this paper, it is clear that the results with the mosaic approach are generally more consistent and less sensitive to grid resolution compared to those from the dominant approach, particularly for ground heat fluxes (Figure 10b), sensible heat fluxes (Figure 10c), latent heat fluxes (Figure 10d), and 2 m specific humidities (Figure 11b) during daytime. For net radiation (Figure 10a) over the whole diurnal cycle, ground heat fluxes and latent heat fluxes during night, the mosaic approach shows a comparable, if not better, performance when the convergence of results is considered. Note that even small differences shown in Figures 10 and 11 can be substantial given that the averaging area is large. For the 2 m air temperatures (Figure 11a) and surface temperatures (Figure 11e), the mosaic approach yields slightly larger differences between d01 and d03 at night, which is consistent with the larger difference in sensible heat flux between d01 and d03 at night (Figure 10c). Nevertheless, the mosaic gives smaller differences between d02 and d03 in 2 m air temperatures and surface temperatures. For 10 m winds, the two approaches show fairly similar convergence among simulations at different resolutions. It is interesting to note that for the boundary layer height, both approaches generate different results with different resolutions. The differences between the two approaches are minor compared to the differences caused by changing resolutions. Overall, the mosaic approach shows a better or comparable performance when considering the convergence of results with increasing resolution.

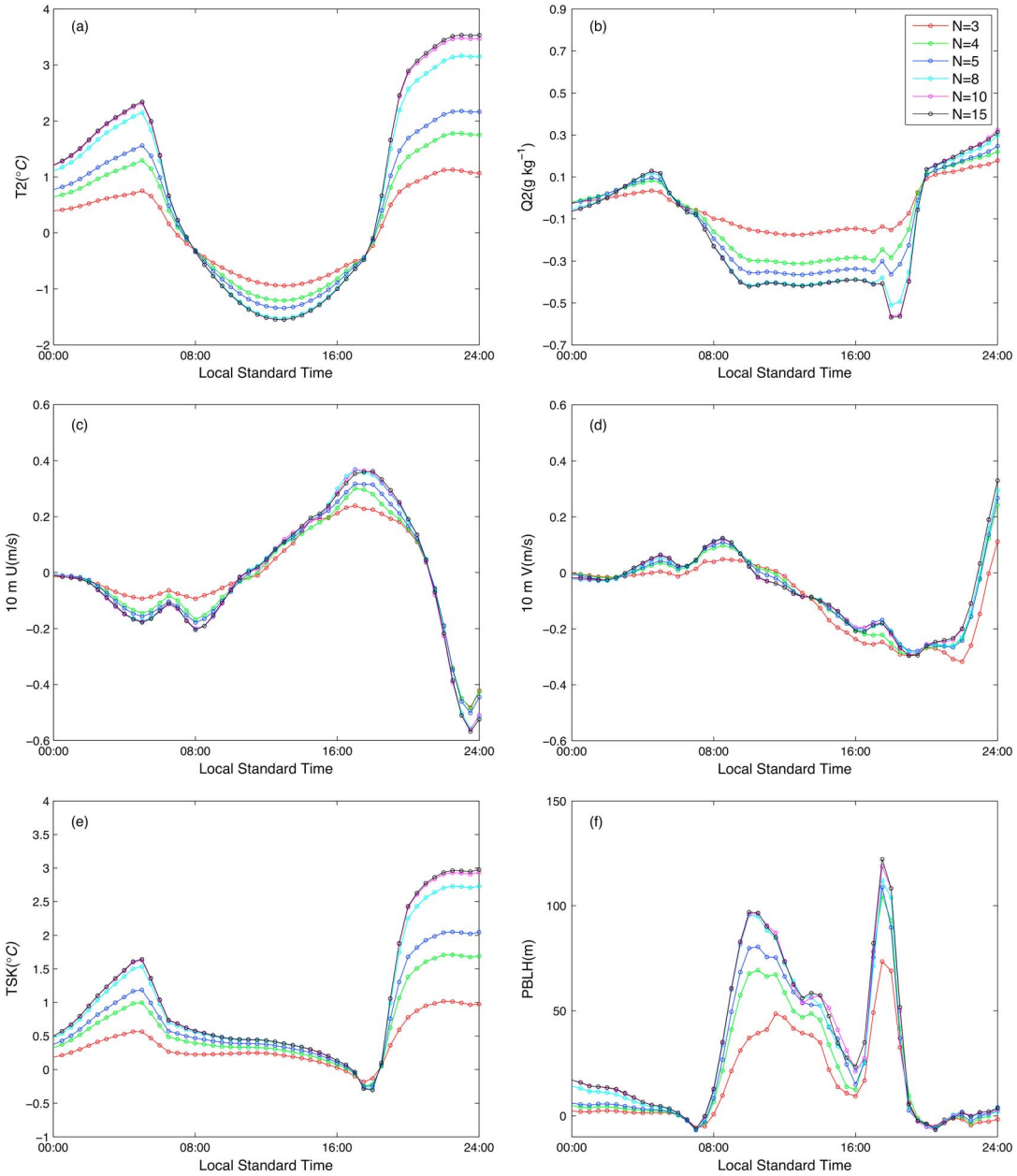

Figure 14. Differences in (a) 2 m air temperature, (b) 2 m specific humidity, (c) 10 m wind speed, (d) 10 m wind direction, (e) surface temperature, and (f) boundary layer height between simulations with different N values and the simulation with the dominant approach ($N = 1$). Results shown here are averaged over selected grid cells of d03 that display high heterogeneity and exclude water surface (similar to Figures 10 and 11).

4.2. The Sensitivity of WRF Results to the Number of Tiles Considered in Each Grid Cell

[39] The results shown in the previous sections (i.e., sections 3 and 4.1) are from WRF simulations with $N = 8$, where N is the number of tiles considered in each grid cell. Theoretically, one expects that as N increases (i.e., more tiles are included), the results converge to the "truth." The aim of this section is to investigate the sensitivity of WRF-simulated results to the number of tiles considered in each grid cell.

[40] Figure 12 shows the differences between results obtained with $N > 1$ (ranging from 3 to 15) and the results obtained with ($N = 1$) (the dominant approach) for surface energy balances averaged over d03 from the simulation with a resolution of 1 km for the clear day case. Note that there are only 15 different land use categories present in our domain (see Figure 1) so the maximum N value used in our simulations is 15. Figure 12a shows that the differences in net radiation range from 0 to −15 W/m^2. The differences represent only 1% changes in net radiation during daytime but 15% changes during nighttime, given that the maximum net radiation exceeds 600 W/m^2 during daytime but is much smaller during nighttime. One can also note that the sensitivity to N is larger during nighttime as the differences in the results with various N values are more prominent. Figure 12b shows changes in ground heat flux when N changes from 1 (i.e., the dominant approach) to 15. Maximum daytime and

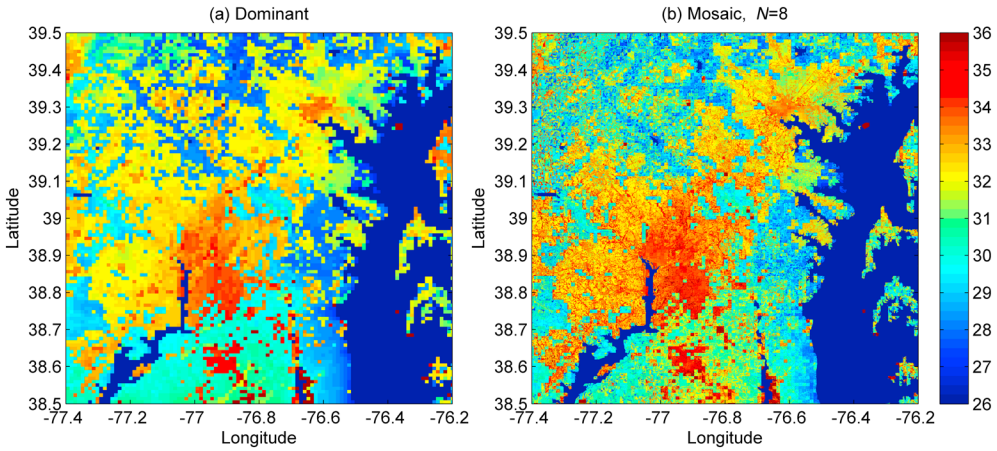

Figure 15. Downscaled land surface temperature maps from (a) WRF simulation with the dominant approach, (b) WRF simulation with the mosaic approach with $N = 8$ at 1400 P.M. local time on 14 July 2009.

nighttime changes exceed 5 W/m^2, which are about 6% of the maximum ground heat flux during daytime and about 10% during nighttime. Figure 12c also shows large differences in the sensible heat flux between the simulation with the dominant approach and those with the mosaic approach ($N > 1$). Increasing N from 1 to 15 results in maximum daytime changes of about 15 W/m^2 and maximum nighttime changes of about 20 W/m^2. Similar to net radiation, relative changes in sensible heat flux during daytime (less than 10%) are smaller than those during nighttime, which can be as large as 100%. In addition, the sensitivity of sensible heat flux to N is also larger during nighttime. For latent heat flux (Figure 12d), using the mosaic approach with different N values yields changes that are slightly larger than 5 W/m^2. Again, for similar reasons, relative changes during nighttime can be very large because of small latent heat fluxes. We also stress that the small differences shown here are important considering the large averaging area (~ 9,600 km^2), and also because they tend to have consistent signs over the diurnal cycle, suggesting they do not cancel out but rather add up over a day.

[41] The sensitivity of surface fluxes to N values is caused by the different surface energy balances over different tiles and by the different atmospheric states. In order to separate these two causes, the surface fluxes averaged only over the first tiles are examined (Figure 13). Comparing Figures 13 and 12 reveals that changes averaged over the first tiles are quite large and comparable in magnitude to the changes averaged over the grid cells. This suggests that as N changes, the atmospheric state changes, altering land-atmosphere interaction even over the dominant tiles and resulting in changes in the simulated surface energy balance over the first tiles, especially during nighttime. The fact that changes in net radiation, ground heat flux, and sensible heat flux are more significant during night suggests that these fluxes are more sensitive to changes in the atmospheric state during nighttime.

[42] Figure 14 presents the sensitivity of domain-averaged near-surface states, surface temperatures, and boundary layer heights to the N value. As can be seen from Figure 14a, WRF-simulated 2 m air temperatures show substantial differences among simulations with different N values. The maximum difference reaches −2°C during daytime and 4°C during nighttime. Furthermore, the sensitivity is larger during nighttime than that during daytime, consistent with the larger sensitivity in sensible heat fluxes at nighttime (Figure 12c). One can further note in Figure 14e that the results of surface temperatures display similar features. Similar to the surface fluxes shown in Figures 12 and 13, these changes are considerably influenced by the changes in fluxes from the first tiles due to changes in the atmospheric fields. The sensitivity of 2 m specific humidity to the N value is on the other hand larger during daytime, as can be seen from Figure 14b. While the 10 m winds (Figures 14c and 14d) converge quickly with increasing N values for the different mosaic simulations (i.e., the lines on Figures 14c and 14d almost collapse), they do indicate significant differences between all the mosaic results, even with $N = 3$, and the dominant approach results. As shown in Figure 14f, the changes in the boundary layer heights display two peaks at the morning and afternoon transitional periods, suggesting large sensitivity of simulated boundary layer heights to the N value when the surface transitions from heating to cooling of the atmosphere or vice versa. During night, the changes in the boundary layer heights are small as N changes, due to the fact that the atmospheric boundary layer is less coupled to the surface and is therefore less affected by the surface conditions.

[43] It can be seen from Figures 12 and 14 that the results with $N = 8$, $N = 10$, and $N = 15$ are fairly close. The differences among these cases are noticeably smaller than the differences among cases with $N = 3$, $N = 4$, and $N = 5$, suggesting that the results with the mosaic approach do converge when N increases. These results in fact motivated our choice of $N = 8$ in the previous analyses.

5. Conclusions and Discussions

[44] This study develops a mosaic approach in the WRF-Noah framework and compares it to the default dominant approach using high-resolution simulations over an urban environment. Various experimental data sets have been used as benchmarks in the comparison. Based on the results, the following conclusions have been reached.

[45] 1. Compared to measurements and observations, the mosaic approach shows better (or, in some cases, comparable) performance than the dominant approach in capturing

the surface energy balance, surface temperatures, boundary layer profiles, and rainfall patterns. In particular, it substantially improves the modeled latent heat flux in urban areas.

[46] 2. The mosaic approach generates more consistent results from simulations with different resolutions than the dominant approach, indicating faster grid convergence.

[47] 3. Compared to the dominant approach, the mosaic approach produces different area-averaged surface fluxes, surface temperatures, near-surface states, and boundary layer heights. When the number of tiles (N) considered in each grid cell increases from 1 to 15, changes in the surface energy balance, surface temperatures, near-surface states, and boundary layer heights are also observed. For ground heat flux, sensible heat flux, surface temperature, and 2 m air temperature, changes are more significant during nighttime, while changes in the 2 m specific humidity are more significant during daytime and changes in the boundary layer heights are most prominent during the morning and afternoon transitional periods. The results with $N = 8$, 10, and 15 are quite similar suggesting convergence of the results with increasing N. The computational costs in terms of wall-clock time increase by 15% when $N = 8$ and by 25% when $N = 15$, as compared to using the dominant approach. However, these findings are specific to our study area and a smaller number of tiles can be considered in other regions depending on the subgrid-scale distribution of land cover type over the domain of interest.

[48] Therefore, the hypothesis presented in the introduction is valid based on our results, that is, despite the fact that numerical simulations over urban environments are usually conducted at high-resolution (with a grid spacing of ~ 1 km), subgrid-scale variability of land surface characteristics continues to be important and should be accounted for, particularly in urban regions with considerable surface heterogeneities and substantial differences between the physical properties of various surface types.

[49] There are a few implications of this study that are important to appreciate. The mosaic approach implemented in this study assumes that the atmospheric properties are uniform over all tiles in a grid cell. This assumption is often made in the literature and some previous studies have examined the impact of this assumption on grid cell averaged fluxes using large-eddy simulations [see, e.g., *Bertoldi et al*., 2007, 2008] or even higher-resolution mesoscale simulations [see, e.g., *Molders et al*., 1996]. *Bertoldi et al*. [2007] discovered the opposing roles of air temperature and wind speed variability in modifying surface flux. For example, their simulations show that the air temperature variability tends to decrease the sensible heat flux over bare soil areas but the variability in wind speed tends to increase the sensible heat flux. Consequently, they conclude that due to a partial canceling of these two effects, the combined effect on flux estimation is minor for operational applications. It is also worth noting that the tiling structure implemented in this study extends to the subsurface, that is, there is a distinct soil type associated with each tile. Nevertheless, due to the unavailability of high-resolution soil properties data sets, the soil properties of all the tiles are assumed identical within a grid cell. Moreover, from a practical perspective, using variable soil types might be challenging since soil variability might not coincide with land cover variability resulting in the need for a large number of tiles with different soil-land-cover combinations.

[50] In this study, the mosaic approach identifies different tiles strictly according to the land use/land cover classification systems that are associated with the land use/land cover data sets. It is also possible to combine some tiles based on their properties. For instance, if one grid cell includes both broadleaf forest and needleleaf forest, one could combine them as one "forest" tile instead of treating them as two separate tiles. This is somewhat similar to the approach adopted by the Community Land Model where each grid cell can be composed of up to five land units: glacier, lake, wetland, urban, and vegetated [*Oleson et al*., 2010]. Nevertheless, in the Community Land Model, each land unit can be further subdivided; for example, the vegetated land unit can include natural and managed land units [*Oleson et al*., 2010].

[51] The mosaic approach clearly shows advantages over the dominant approach in heterogeneous settings such as the urban environment examined in this study. The two cases studied here are designed to represent clear sky conditions and rain conditions, respectively. More case studies and/or longer simulations need to be performed in the future to confirm the encouraging results of these two cases, along with simulations over other regions.

[52] Theoretically, the mosaic approach is expected to perform better than the dominant approach by incorporating more surface information. Nevertheless, other biases and errors are also likely to be introduced when more calculations are included in the WRF-Noah framework. Moreover, in complex numerical models like WRF, reducing biases in one component of the model does not always yield a reduction in total model biases due to the possibility that errors from different component are partially canceling each other. Over heterogeneous areas, this study confirms that the benefits of the mosaic approach outweigh its potential shortcoming.

[53] Despite the fact that the mosaic approach provides improvements in some of the simulated results as compared to experimental data, significant biases still exist in certain aspects of results of WRF when compared to observations, which seem not to be reducible by the mosaic approach. For example, correctly characterizing surface properties such as albedo remains a challenge. The WRF-simulated rainfall distributions do not resemble the radar-estimated rainfall distribution. These biases are potentially linked to errors in the initial/boundary conditions, model dynamics, as well as other model physics (e.g., in the microphysical schemes) that are not related to the earth surface and its modeling.

[54] Although the mosaic method is general and can be applied to any numerical weather/climate model, it requires land use/land cover data sets that are of higher resolutions than the model simulations. In this study, the NLCD land cover data set has a resolution of 30 m and thus is suitable for examining subgrid-scale variability in urban environments. At places where such high-resolution land cover data sets are not available, application of a mosaic approach will not benefit simulations if their resolutions are comparable to those of the land cover data set. However, when such high-resolution land cover data set is available, a potential output of the mosaic approach that is not possible with other approaches including the composite and the statistical-dynamical approach is downscaled surface state maps. To produce such maps, surface parameters such as surface temperature, surface fluxes, and soil moisture of a given land use tile as computed in each grid cell (here 1×1 km) are

taken and assigned to each land use pixel (here 30×30 m from NLCD) in that grid cell, according to its land use. This allows the development of land surface state maps with a resolution equal to that of the finer land cover data rather than the coarser atmospheric model data. Such maps can be used for example in studies focusing on the biogeochemistry of the soils in the area or on surface hydrology where high-resolution information regarding surface conditions is essential. An example of such a map produced from our simulated over d03 is shown in Figure 15b and compared to the corresponding surface map produced at the atmospheric model resolution by the dominant approach (Figure 15a).

[55] **Acknowledgments.** This study was funded by the US National Science Foundation under grant CBET-1058027 and by the Princeton Environmental Institute-Science, Technology, and Environmental Policy fellowship. The authors also like to acknowledge the support from the NOAA CPPA grant NA09OAR4310193, the NOAA Cooperative Institute for Climate Sciences, and the NCAR Water System and BEACHON Programs.

References

Avissar, R. (1991), A statistical-dynamic approach to parameterize subgrid-scale land-surface heterogeneity in climate models, *Surv. Geophys.*, *12*(1–3), 155–178, doi:10.1007/Bf01903417.

Avissar, R., and R. A. Pielke (1989), A parameterization of heterogeneous land surfaces for atmospheric numerical-models and its impact on regional meteorology, *Mon. Weather Rev.*, *117*(10), 2113–2136, doi:10.1175/1520-0493(1989)117<2113:Apohls>2.0.Co;2.

Avissar, R., and T. Schmidt (1998), An evaluation of the scale at which ground-surface heat flux patchiness affects the convective boundary layer using large-eddy simulations, *J. Atmos. Sci.*, *55*(16), 2666–2689, doi:10.1175/1520-0469(1998)055<2666:Aeotsa>2.0.Co;2.

Bertoldi, G., J. D. Albertson, W. P. Kustas, F. Li, and M. C. Anderson (2007), On the opposing roles of air temperature and wind speed variability in flux estimation from remotely sensed land surface states, *Water Resour. Res.*, *43*, W10433, doi:10.1029/2007wr005911.

Bertoldi, G., W. P. Kustas, and J. D. Albertson (2008), Estimating spatial variability in atmospheric properties over remotely sensed land surface conditions, *J. Appl. Meteorol. Climatol.*, *47*(8), 2147–2165, doi:10.1175/2007jamc1828.1.

Bou-Zeid, E., C. Meneveau, and M. B. Parlange (2004), Large-eddy simulation of neutral atmospheric boundary layer flow over heterogeneous surfaces: Blending height and effective surface roughness, *Water Resour. Res.*, *40*, W02505, doi:10.1029/2003WR002475.

Bou-Zeid, E., M. B. Parlange, and C. Meneveau (2007), On the parameterization of surface roughness at regional scales, *J. Atmos. Sci.*, *64*(1), 216–227, doi:10.1175/JAS3826.1.

Chang, H. I., A. Kumar, D. Niyogi, U. C. Mohanty, F. Chen, and J. Dudhia (2009), The role of land surface processes on the mesoscale simulation of the July 26, 2005 heavy rain event over Mumbai, India, *Glob. Planet. Change*, *67*(1–2), 87–103, doi:10.1016/j.gloplacha.2008.12.005.

Chen, F., and R. Avissar (1994), The impact of land-surface wetness heterogeneity on mesoscale heat fluxes, *J. Appl. Meteorol.*, *33*(*11*), 1323–1340, doi:10.1175/1520-0450(1994)033<1323:Tiolsw>2.0.Co;2.

Chen, F., and J. Dudhia (2001), Coupling an advanced land surface—Hydrology model with the Penn State—NCAR MM5 modeling system. Part I: Model implementation and sensitivity, *Mon. Weather Rev.*, *129*(*4*), 569–585, doi:10.1175/1520-0493(2001)129<0569:Caalsh>2.0.Co;2.

Chen, F., and Y. Zhang (2009), On the coupling strength between the land surface and the atmosphere: From viewpoint of surface exchange coefficients, *Geophys. Res. Lett.*, *36*, L10404, doi:10.1029/2009gl037980.

Chen, F., Z. Janjić, and K. Mitchell (1997), Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model, *Boundary Layer Meteorol.*, *85*(3), 391–421, doi:10.1023/a:1000531001463.

Chen, F., D. N. Yates, H. Nagai, M. A. LeMone, K. Ikeda, and R. L. Grossman (2003), Land surface heterogeneity in the cooperative atmosphere surface exchange study (CASES-97). Part I: Comparing modeled surface flux maps with surface-flux tower and aircraft measurements, *J. Hydrometeorol.*, *4*(2), 196–218, doi:10.1175/1525-7541(2003)4<196:Lshitc>2.0.Co;2.

Chen, F., S. G. Miao, M. Tewari, J. W. Bao, and H. Kusaka (2011a), A numerical study of interactions between surface forcing and sea breeze circulations and their effects on stagnation in the greater Houston area, *J. Geophys. Res.*, *116*, D12105, doi:10.1029/2010jd015533.

Chen, F., et al. (2011b), The integrated WRF/urban modelling system: Development, evaluation, and applications to urban environmental problems, *Int. J. Climatol.*, *31*(2), 273–288, doi:10.1002/Joc.2158.

Chen, F., R. Bornstein, S. Grimmond, J. Li, X. D. Liang, A. Martilli, S. G. Miao, J. Voogt, and Y. C. Wang (2012), Research priorities in observing and modeling urban weather and climate, *Bull. Am. Meteorol. Soc.*, *93*(11), 1725–1728, doi:10.1175/Bams-D-11-00217.1.

Crawford, B., C. S. B. Grimmond, and A. Christen (2011), Five years of carbon dioxide fluxes measurements in a highly vegetated suburban area, *Atmos. Environ.*, *45*(4), 896–905, doi:10.1016/j.atmosenv.2010.11.017.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.*, *108*(D22), 8851, doi:10.1029/2002jd003296.

Entekhabi, D., and P. S. Eagleson (1989), Land surface hydrology parameterization for atmospheric general circulation models including subgrid scale spatial variability, *J. Clim.*, *2*(8), 816–831, doi:10.1175/1520-0442(1989)002<0816:Lshpfa>2.0.Co;2.

Famiglietti, J. S., and E. F. Wood (1991), Evapotranspiration and runoff from large land areas: Land surface hydrology for atmospheric general circulation models, *Surv. Geophys.*, *12*(1–3), 179–204, doi:10.1007/Bf01903418.

Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham (2011), Completion of the 2006 national land cover database for the conterminous United States, *Photogramm. Eng. Remote Sens.*, *77*(9), 858–864.

Giorgi, F., and R. Avissar (1997), Representation of heterogeneity effects in earth system modeling: Experience from land surface modeling, *Rev. Geophys.*, *35*(4), 413–437, doi:10.1029/97rg01754.

Grimmond, S. (2007), Urbanization and global environmental change: Local effects of urban warming, *Geogr. J.*, *173*, 83–88, doi:10.1111/j.1475-4959.2007.232_3.x.

Grimmond, C. S. B., et al. (2011), Initial results from phase 2 of the international urban energy balance model comparison, *Int. J. Climatol.*, *31*(2), 244–272, doi:10.1002/Joc.2227.

Koster, R. D., and M. J. Suarez (1992a), Modeling the land surface boundary in climate models as a composite of independent vegetation stands, *J. Geophys. Res.*, *97*(D3), 2697–2715.

Koster, R. D., and M. J. Suarez (1992b), A comparative analysis of two land surface heterogeneity representations, *J. Clim.*, *5*(12), 1379–1390, doi:10.1175/1520-0442(1992)005<1379:Acaotl>2.0.Co;2.

LeMone, M. A., M. Tewari, F. Chen, J. G. Alfieri, and D. Niyogi (2008), Evaluation of the Noah land surface model using data from a Fair-Weather IHOP_2002 Day with heterogeneous surface fluxes, *Mon. Weather Rev.*, *136*(12), 4915–4941, doi:10.1175/2008mwr2354.1.

Lemone, M. A., F. Chen, M. Tewari, J. Dudhia, B. Geerts, Q. Miao, R. L. Coulter, and R. L. Grossman (2010a), Simulating the IHOP_2002 Fair-Weather CBL with the WRF-ARW-Noah modeling system. Part I: Surface fluxes and CBL structure and evolution along the eastern track, *Mon. Weather Rev.*, *138*(3), 722–744, doi:10.1175/2009mwr3003.1.

Lemone, M. A., F. Chen, M. Tewari, J. Dudhia, B. Geerts, Q. Mia, R. L. Coulter, and R. L. Grossman (2010b), Simulating the IHOP_2002 Fair-Weather CBL with the WRF-ARW Noah modeling system. Part II: Structures from a few kilometers to 100 km across, *Mon. Weather Rev.*, *138*(3), 745–764, doi:10.1175/2009mwr3004.1.

Leroyer, S., S. Belair, and J. Mailhot (2011), Microscale numerical prediction over Montreal with the Canadian external urban modeling system, *J. Appl. Meteorol. Climatol.*, *50*(12), 2410–2428, doi:10.1175/Jamc-D-11-013.1.

Li, B., and R. Avissar (1994), The impact of spatial variability of land-surface characteristics on land-surface heat fluxes, *J. Clim.*, *7*(4), 527–537, doi:10.1175/1520-0442(1994)007<0527:Tiosvo>2.0.Co;2.

Li, D., and E. Bou-Zeid (2011), Coherent structures and the dissimilarity of turbulent transport of momentum and scalars in the unstable atmospheric surface layer, *Boundary Layer Meteorol.*, *140*(*2*), 243–262, doi:10.1007/s10546-011-9613-5.

Li, D., and E. Bou-Zeid (2013), Synergistic interactions between urban heat islands and heat waves: The impact in cities is larger than the sum of its parts, *J. Appl. Meteorol. Climatol.*, *52*(9), 2051–2064, doi:10.1175/JAMC-D-13-02.1.

Li, D., E. Bou-Zeid, M. L. Baeck, S. Jessup, and J. A. Smith (2013), Modeling land surface processes and heavy rainfall in urban environments: Sensitivity to urban surface representations, *J. Hydrometeorol.*, *14*(4), 1098–1118, doi:10.1175/jhm-d-12-0154.1.

Miao, S. G., F. Chen, Q. C. Li, and S. Y. Fan (2011), Impacts of urban processes and urbanization on summer precipitation: A case study of heavy rainfall in Beijing on 1 August 2006, *J. Appl. Meteorol. Climatol.*, *50*(4), 806–825, doi:10.1175/2010jamc2513.1.

Molders, N., A. Raabe, and G. Tetzlaff (1996), A comparison of two strategies on land surface heterogeneity used in a mesoscale beta meteorological model, *Tellus Ser. A-Dyn. Meteorol. Oceanol.*, *48*(5), 733–749, doi:10.1034/j.1600-0870.1996.00012.x.

Molod, A., and H. Salmun (2002), A global assessment of the mosaic approach to modeling land surface heterogeneity, *J. Geophys. Res.*, *107*(D14), 4217, doi:10.1029/2001jd000588.

Niyogi, D., P. Pyle, M. Lei, S. P. Arya, C. M. Kishtawal, M. Shepherd, F. Chen, and B. Wolfe (2011), Urban modification of thunderstorms: An observational storm climatology and model case study for the Indianapolis urban region, *J. Appl. Meteorol. Climatol.*, *50*(5), 1129–1144, doi:10.1175/2010jamc1836.1.

Ntelekos, A. A., J. A. Smith, M. L. Baeck, W. F. Krajewski, A. J. Miller, and R. Goska (2008), Extreme hydrometeorological events and the urban environment: Dissecting the 7 July 2004 thunderstorm over the Baltimore MD metropolitan region, *Water Resour. Res.*, *44*, W08446, doi:10.1029/2007wr006346.

Oke, T. R. (1982), The energetic basis of the urban heat island, *Q. J. R. Meteorol. Soc.*, *108*(455), 1–24, doi:10.1002/qj.49710845502.

Oleson, K. W., et al. (2010), Technical description of version 4.0 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-478 +STR Rep.

Rasmussen, R., et al. (2011), High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate, *J. Clim.*, *24*(12), 3015–3048, doi:10.1175/2010jcli3985.1.

Shephard, J. M. (2005), A review of current investigations of urban-induced rainfall and recommendations for the future, *Earth Interact.*, *9*, doi:10.1175/EI156.1.

Skamarock, W. C., and J. B. Klemp (2008), A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, *J. Comput. Phys.*, *227*(7), 3465–3485, doi:10.1016/j.jcp.2007.01.037.

Smith, J. A., M. L. Baeck, G. Villarini, C. Welty, A. J. Miller, and W. F. Krajewski (2012), Analyses of a long-term, high-resolution radar rainfall data set for the Baltimore metropolitan region, *Water Resour. Res.*, *48*, W04504, doi:10.1029/2011wr010641.

Talbot, C., E. Bou-Zeid, and J. Smith (2012), Nested mesoscale large-eddy simulations with WRF: Performance in real test cases, *J. Hydrometeorol.*, *13*(5), 1421–1441, doi:10.1175/jhm-d-11-048.1.

Trier, S. B., F. Chen, and K. W. Manning (2004), A study of convection initiation in a mesoscale model using high-resolution land surface initial conditions, *Mon. Weather Rev.*, *132*(12), 2954–2976.

Trier, S. B., F. Chen, K. W. Manning, M. A. LeMone, and C. A. Davis (2008), Sensitivity of the PBL and precipitation in 12-day simulations of warm-season convection using different land surface models and soil wetness conditions, *Mon. Weather Rev.*, *136*(7), 2321–2343, doi:10.1175/2007mwr2289.1.

Trier, S. B., M. A. LeMone, F. Chen, and K. W. Manning (2011), Effects of surface heat and moisture exchange on ARW-WRF warm-season precipitation forecasts over the Central United States, *Weather Forecast.*, *26*(1), 3–25, doi:10.1175/2010waf2222426.1.

Wang, Z. H., E. Bou-Zeid, and J. A. Smith (2013), A coupled energy transport and hydrological model for urban canopies evaluated using a wireless sensor network, *Q.J.R. Meteorol. Soc.*, *139*, pp. 1643–1657, doi:10.1002/qj.2032.

Wu, L., H. Su, and J. H. Jiang (2011), Regional simulations of deep convection and biomass burning over South America: 1. Model evaluations using multiple satellite data sets, *J. Geophys. Res.*, *116*, D17208, doi:10.1029/2011jd016105.

Yeung, J. K., J. A. Smith, G. Villarini, A. A. Ntelekos, M. L. Baeck, and W. F. Krajewski (2011), Analyses of the warm season rainfall climatology of the northeastern US using regional climate model simulations and radar rainfall fields, *Adv. Water Resour.*, *34*(2), 184–204, doi:10.1016/j.advwatres.2010.10.005.

Zeng, X. B., Z. Wang, and A. H. Wang (2012), Surface skin temperature and the interplay between sensible and ground heat fluxes over arid regions, *J. Hydrometeorol.*, *13*(4), 1359–1370, doi:10.1175/Jhm-D-11-0117.1.

Zhang, C. L., F. Chen, S. G. Miao, Q. C. Li, X. A. Xia, and C. Y. Xuan (2009), Impacts of urban expansion and future green planting on summer precipitation in the Beijing metropolitan area, *J. Geophys. Res.*, *114*, D02116, doi:10.1029/2008jd010328.

Zhang, D. L., Y. X. Shou, R. R. Dickerson, and F. Chen (2011), Impact of upstream urbanization on the urban heat island effects along the Washington-Baltimore Corridor, *J. Appl. Meteorol. Climatol.*, *50*(10), 2012–2029, doi:10.1175/Jamc-D-10-05008.1.

Zheng, W. Z., H. L. Wei, Z. Wang, X. B. Zeng, J. Meng, M. Ek, K. Mitchell, and J. Derber (2012), Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation, *J. Geophys. Res.*, *117*, D06117, doi:10.1029/2011jd015901.

Zilitinkevich, S. S. (1995), Non-local turbulent transport pollution dispersion aspects of coherent structure of convective flows, in *In Air Pollution III Vol. 1, Air Pollution Theory and Simulation*, edited by H. Power, N. Moussiopoulos, and C. A. Brebbia, pp. 53–60, Computational Mechanics Publications, Southampton Boston.