Implementation of a new urban energy budget scheme into MetUM. Part II: Validation against observations and model intercomparison

Article

Accepted Version

the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online
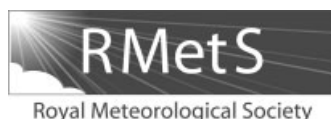
RMetS

Royal Meteorological Society

Implementation of a new urban energy budget scheme into MetUM. Part II: Validation against observations and model intercomparison

A. Porson,[a]* P. A. Clark,[b†] I. N. Harman,[c†] M. J. Best[d†] and S. E. Belcher[a]

[a]*Department of Meteorology, University of Reading, Reading, UK*
[b]*Joint Centre for Mesoscale Meteorology, Met Office, Reading, UK*
[c]*CSIRO Marine and Atmospheric Research, Canberra, Australia*
[d]*Joint Centre for Hydrometeorological Research, Met Office, Wallingford, UK*
*Correspondence to: A. Porson, Met Office, FitzRoy Road, Exeter, EX1 3PB, UK. E-mail: aurore.porson@metoffice.gov.uk
†The contributions of P. A. Clark and M. J. Best were written in the course of their employment at the Met Office, UK, and are published with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

In the first part of this article, we introduced a new urban surface scheme, the Met Office – Reading Urban Surface Exchange Scheme (MORUSES), into the Met Office Unified Model (MetUM) and compared its impact on the surface fluxes with respect to the current urban scheme. In this second part, we aim to analyze further the reasons behind the differences. This analysis is conducted by a comparison of the performance of the two schemes against observations and against a third model, the Single Column Reading Urban model (SCRUM). The key differences between the three models lie in how each model incorporates the heat stored in the urban fabric and how the surface-energy balance is coupled to the underlying substrate.

The comparison of the models with observations from Mexico City reveals that the performance of MORUSES is improved if roof insulation is included by minimizing the roof thickness. A comparison of MORUSES and SCRUM reveals that, once insulation is included within MORUSES, these two models perform equally well against the observations overall, but that there are differences in the details of the simulations at the roof and canyon level. These differences are attributed to the different representations of the heat-storage term, specifically differences in the dominant frequencies captured by the urban canopy and substrate, between the models. These results strongly suggest a need for an urban model intercomparison exercise. Copyright © 2010 Royal Meteorological Society and Crown Copyright.

1. Introduction

In part I of this work, Porson *et al.* (2010) describe a new parametrization of urban areas that they have implemented into the numerical weather prediction and climate versions of the Met Office Unified Model (MetUM). The new scheme is called the Met Office – Reading Urban Surface Exchange Scheme (MORUSES), and represents urban areas as two-dimensional parallel rows of streets and buildings. As shown by Best *et al.* (2006) and Porson *et al.* (2009),

the main features of the energy balance of a system of two-dimensional streets and buildings are well captured by representing the surface using two facets, one that represents the energy balance of the roof and one that represents the energy balance of the street surfaces. Consequently, MORUSES represents the urban surface using two tiles.

The new parametrization is described in detail in Part I (Porson *et al.* 2010), which also shows idealized simulations. Here, in Part II, the performance of MORUSES is tested in two ways: firstly against observations of surface fluxes, and secondly against numerical outputs from another urban energy-balance model, the Single Column Reading Urban Model (SCRUM), described in Harman and Belcher (2006) and Porson *et al.* (2009). The data come from observations of surface fluxes made in Mexico City by Oke *et al.* (1999). These data were selected because this site has a small latent heat flux, and is less complicated by spatial inhomogeneity and topography than many other sites where observations have been taken. The study from Oke *et al.* (1999) shows that the Mexico City site is dominated by a large heat uptake (related to the large thermal mass of the urban facets) and that the energy balance is dominated by urban effects. These characteristics suit the methodology underpinning both the modelling and observational techniques.

MORUSES was constructed based on the physics developed in SCRUM (Harman *et al.*, 2004a, 2004b), but the coupling to the underlying soil is parametrized differently. The soil scheme currently used in the MetUM is not tiled, meaning that diffusion of heat and moisture within the substrate represents a grid-box average. Even with the present high resolutions, as fine as 1 km, most grid boxes that are urban have a substantial fraction that is vegetated. This means that heat stored in the building and street fabric needs to be considered separately from the underlying soil module. As described in Part I (Porson *et al.*, 2010), MORUSES solves this problem by building on the urban parametrization scheme developed by Best (2005), and represents heat storage by the urban fabric through a slab with large heat capacity. SCRUM, in contrast, represents heat storage into the building fabric by diffusion on each of the urban facets (more details on the differences between these models will be given in section 2). By comparing MORUSES with SCRUM, therefore, we are able to assess the ability of the simplified slab representation to capture the behaviour of the more complete diffusion model.

Differences in the parametrizations of the heat-storage and diffusion terms have been shown to be important in simulations of the urban energy balance. For example, Dupont and Mestayer (2006) showed that an urban canopy model including at least two substrate layers would capture better the phase of the surface fluxes. Holt and Pullen (2007) also obtained a significant difference in the period of sensible heat flux when comparing a slab canopy scheme with a multilayer urban scheme.

We centre this evaluation of MORUSES around three questions.

(1) How do MORUSES and the Best scheme compare against real observations of the urban energy balance?
(2) Are the details of the method used to incorporate heat storage in the roof and roof insulation critical to the performance of MORUSES?

(3) Is the surface-energy balance sensitive to the model and/or numerical implementation employed to represent the urban substrate?

The second question is motivated by the increasing prevalence of the use of roof insulation in urban areas, including Mexico City, for the purposes of improving human comfort and reducing energy consumption. The use of roof insulation also implies small conductivity values within the roof substrate and an altered surface-energy balance. We expand on the implications of this point in section 2. From the point of view of numerical weather prediction (NWP), where accuracy and CPU costs are a constraint on new developments (Best, 2006), the third question is important. All three of these questions are best addressed through a comparison study, with all the models being forced by the same atmospheric data and the comparison considering both the aggregate fluxes and the tile (facet) fluxes.

This article is therefore structured as follows: in section 2, we describe briefly the numerical models and observational data set from Mexico City used in this study. The key differences in how SCRUM, MORUSES and the Best scheme address heat storage and the coupling to the underlying substrate are discussed in more detail. In section 3, the performance of MORUSES and the Best scheme against observations of surface fluxes from Mexico City is assessed. As part of this section, the sensitivity of MORUSES to roof insulation is also considered. In section 4, we critically compare the performance of MORUSES and SCRUM against the observations from Mexico City with the aim of assessing the implications of using different methodologies to model heat storage and heat diffusion into the urban fabric. Finally, we summarize and highlight the key implications of this work for the future development of MORUSES and representations of urban areas within NWP models in general.

2. Research tools

2.1. Research models

To allow for a direct comparison with observations, both MORUSES and the Best scheme were used within the Joint UK Land Environment Simulator (JULES). JULES is a community land-surface model that is being developed to be used as a stand-alone scheme as well as the surface scheme within the MetUM. It was initially based upon the Met Office land surface scheme, MOSES (Cox *et al.*, 1999; Essery *et al.*, 2001). JULES is forced with atmospheric driving data of air temperature, wind, pressure, humidity and the downward components of the short-wave and long-wave radiation. Model outputs of surface fluxes are then evaluated against mast measurements that are located in the urban area, above the blending height. The specifications of MORUSES and the Best scheme used here are as described in Part I (Porson *et al.*, 2010).

Figure 1 shows a schematic of the three urban schemes used in this study: the Best scheme, MORUSES and SCRUM. The Best scheme (Best, 2005) consists of a slab urban canopy model, which is radiatively coupled to the soil module of MetUM/JULES. This scheme uses a single, thick facet to represent the entire urban area. The radiative coupling to the underlying soil module is relatively weak which,
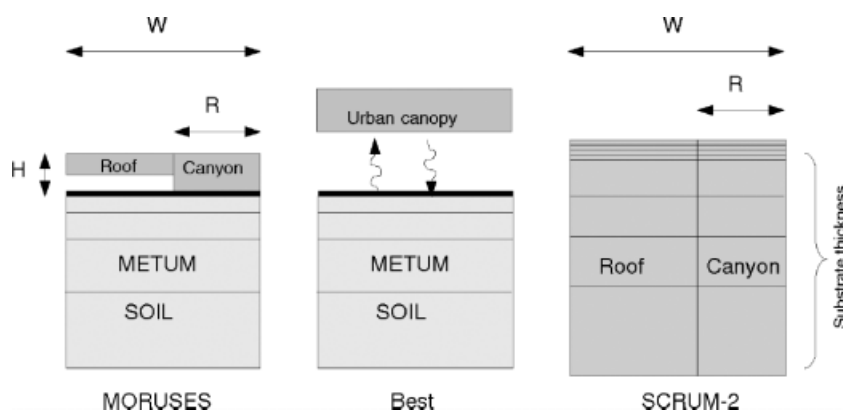
Figure 1. Schematics of MORUSES, Best and SCRUM-2.

together with the high thermal inertia of the facet, allows the urban surface to accumulate considerable amounts of energy during the day and to release it by night. The radiative exchange in the Best scheme uses prescribed values for the surface albedo and emissivity, and the turbulent exchange is given by a parametrization developed from standard surface-layer theory with prescribed values for the roughness lengths for scalar and momentum transfer.

SCRUM exists in a number of different configurations; here we use the two-tile version, denoted SCRUM-2. SCRUM-2 represents the urban surface as the combination of two independent facets (canyon and roof). The energy balance of the canyon facet is obtained by systematically aggregating the energy balances of the wall and street facets of an urban canyon together into a single tile (Porson *et al.*, 2009). MORUSES and SCRUM-2 rely on the same parametrizations for a number of physical processes. The multiple reflection of short- and long-wave radiation within the canyon is captured using the theory of exchange of diffuse radiation (Harman *et al.*, 2004a). A morphometric formulation is used to prescribe the roughness length for momentum and the displacement height (MacDonald, 1998). Finally, the turbulent exchange of scalars (heat) is parametrized using the resistance network developed in Harman *et al.* (2004b), which recognizes the different transport processes occurring in the recirculation and ventilated regions within an urban street canyon.

As mentioned earlier, the principal difference between SCRUM-2 and MORUSES lies in the modelling of the heat-storage term in the energy balance. In SCRUM-2, the temperature diffusion is resolved numerically over a numbers of layers (here eight), reaching a total depth that is greater than the effective damping depth of the facet. This method is used for both the canyon and roof facets. In MORUSES, as explained in Part I (Porson *et al.*, 2010; section 2.5) and illustrated in Figures 1 and 6, the methodology used is substantially different and, furthermore, is different for the canyon and roof facets. For the canyon facet, the heat-storage term is modelled using a single canopy layer and a force-restore model is used to couple this layer to the underlying soil module of MetUM. The force-restore model is used with a depth scale equal to the damping depth of the canyon substrate material. For the roof facet, however, the heat-storage term is modelled using a single canopy layer with a zero-flux bottom boundary condition. This implies no direct coupling to the underlying soil module.

The methodology used to model the roof facet is critical when addressing the issue of roof insulation. Roof insulation is commonly characterized as a material with a low thermal conductivity. Referring to Part I (Porson *et al.*, 2010), the thermal conductivity appears in two ways within MORUSES: firstly in the coupling of the urban canopy to the underlying soil module and secondly through the use of the damping depth to determine the areal heat capacity of the roof and canyon facets. Now since we assume a zero-flux bottom boundary condition for the roof facet, a small thermal conductivity implies a small damping depth (or alternatively a small roof thickness) and a small areal heat capacity. There are therefore two methods available for roof insulation to be incorporated into MORUSES. Firstly, as a control case and denoted 'MORUSES-thick roof', the insulation is captured through the values chosen for thermal conductivity. In this case, these values are determined as the depth-weighted average of the thermal conductivities over the different materials that form the roof layers. The parameter values are given by Masson *et al.* (2002). In this case, the thickness of the roof tile takes the default value. The presence of an insulation layer then acts to reduce the value of the thermal conductivity. Secondly, denoted by 'MORUSES-thin roof', the insulation is modelled by setting the thickness of the roof to be very small. Specifically the thickness of the roof takes a value of 0.02 m instead of the default value of 0.12 m (as given in Table II). Such a value was chosen to be small enough to lead to a significant difference in the surface-energy balance, as shown in Part I (Porson *et al.*, 2010). We will compare both configurations of MORUSES, the Best scheme and SCRUM-2 with the observations from Mexico City in the following sections.

2.2. Observations of Mexico City

The Mexico City data have been used previously to validate urban surface schemes in dry atmospheric conditions (Masson *et al.*, 2002; Best *et al.*, 2006). Mexico City is described as a central city case (Grimmond and Oke, 1999), with a height-to-width aspect ratio of approximately 1.2, in which the buildings occupy approximately 55% of the total planar area (Masson *et al.*, 2002). The three models require values to be specified for parameters associated with the surface characteristics; these are given in Tables I, II and III. For this comparison with observations, we no longer use the Macdonald formulation for roughness length and displacement height. Instead, we prescribe specific values from the literature. The effective roughness length and displacement height are assigned respectively to be

Table I. Input parameters for Mexico City to the Best scheme. The roughness length data are taken from Masson *et al.* (2002). The albedo and emissivity values are calculated using area-weighted averages over the respective albedo and emissivity values of the urban facets (see Table II for details of these values). The heat capacity value as well as the ratio between roughness length (r.l.) of momentum to heat are the default values used in the operational version of the model.

Symbols + units	Value
α	0.20
ε	0.89
C (J m^{-2} K^{-1})	0.28×10^6
Bulk roughness length for momentum (m)	2.2
Ratio r.l. for heat to r.l. for momentum	10^{-7}
Soil temperature (K)	295
Initial temperature (K)	287

Table II. Input parameters for Mexico city to MORUSES. Data are taken from Masson *et al.* (2002). Heat capacity and conductivity values are calculated using depth-weighted averages over four soil layers. Material thickness is computed using a damping depth formulation. Similarly to part I (Porson *et al.*, 2010), the road, wall and roof facets are described respectively with $_{r,w}$ and $_f$.

Symbols + units	Value
C_r (J m^{-3} K^{-1})	1.51×10^6
Δz_r (m)	0.11
λ_r (W m^{-1} K^{-1})	0.69
C_w (J m^{-3} K^{-1})	1.50×10^6
Δz_w (m)	0.13
λ_w (W m^{-1} K^{-1})	0.86
C_f (J m^{-3} K^{-1})	1.09×10^6
Δz_f (m)	0.12
λ_f (W m^{-1} K^{-1})	0.54
λ_{soil} (W m^{-1} K^{-1})	0.22
α_r	0.08
α_w	0.25
α_f	0.20
ε_r	0.95
ε_w	0.85
ε_f	0.90
Building height (m)	18.8
Material roughness length (m)	0.05
H/W	1.18
W/R	0.45
Soil temperature (K)	295
Initial temperature (K)	287

2.2 m (Masson *et al.*, 2002) and 13 m (Oke *et al.*, 1999). Measurements of the atmospheric forcing data and the surface fluxes were taken at 28 m above ground level, giving an effective measurement height of approximately 10 m above the canopy. Further details on the observational

Table III. Input parameters for Mexico city to SCRUM-2. Data are taken from Masson *et al.* (2002). Canyon input parameters are calculated using area-weighted averages from the facet parameters shown in Table II. Similarly to Part I (Porson *et al.*, 2010), the canyon and roof facets are described respectively with $_c$ and $_f$.

| Symbols + units	Value
C_c (J m^{-3} K^{-1})	1.50×10^6
C_f (J m^{-3} K^{-1})	1.09×10^6
λ_c (W m^{-1} K^{-1})	0.81
λ_f (W m^{-1} K^{-1})	0.54
α_c	0.20
α_f	0.20
ε_c	0.88
ε_f	0.90
Building height (m)	18.8
Material roughness length (m)	0.05
H/W	1.18
W/R	0.45
Building internal temperature (K)	293
Initial temperature (K)	287

techniques and experimental design are given by Grimmond and Oke (1999).

Since SCRUM-2 is configured to accept internal building temperatures but MORUSES and Best do not, the initialization of the surface and substrate temperatures are different in the different models. However, from Masson *et al.* (2002), the difference between the internal building temperature and soil bottom temperature is negligible ($\sim$ 2°C).

Each of these three models is used in an identical manner. The same forcing data are used to drive the models and the outputs are aggregated so as to be comparable with hourly averaged observations. The three models are run in dry conditions – moisture processes are neglected for the Best scheme and MORUSES and are not incorporated in SCRUM-2. The small observed values of the latent heat flux are added to the sensible heat flux in order for comparison with the simulated sensible heat flux of the models. The simulations employ a spin-up phase where the first day of observations is used ten times to force the models and the models are run over six full days, from which an average diurnal cycle is made for both model outputs and measurements. Finally, a correction was made to shift the time series of solar forcing by one hour after comparing the time series from observations of downwelling solar fluxes, measured at about 10 km from the site, and net radiation, measured on site (Masson *et al.*, 2002). With this correction we then capture the timing of the increase in radiation at dawn, which is a better metric than the timing of the maximum in net radiation.

3. Comparison of MORUSES and Best schemes against Mexico City data

We first consider the relative performance of MORUSES (thick and thin roof configurations) and the Best scheme against the observations from Mexico City. Figure 2 shows the composite diurnal cycle from six days of simulations. The
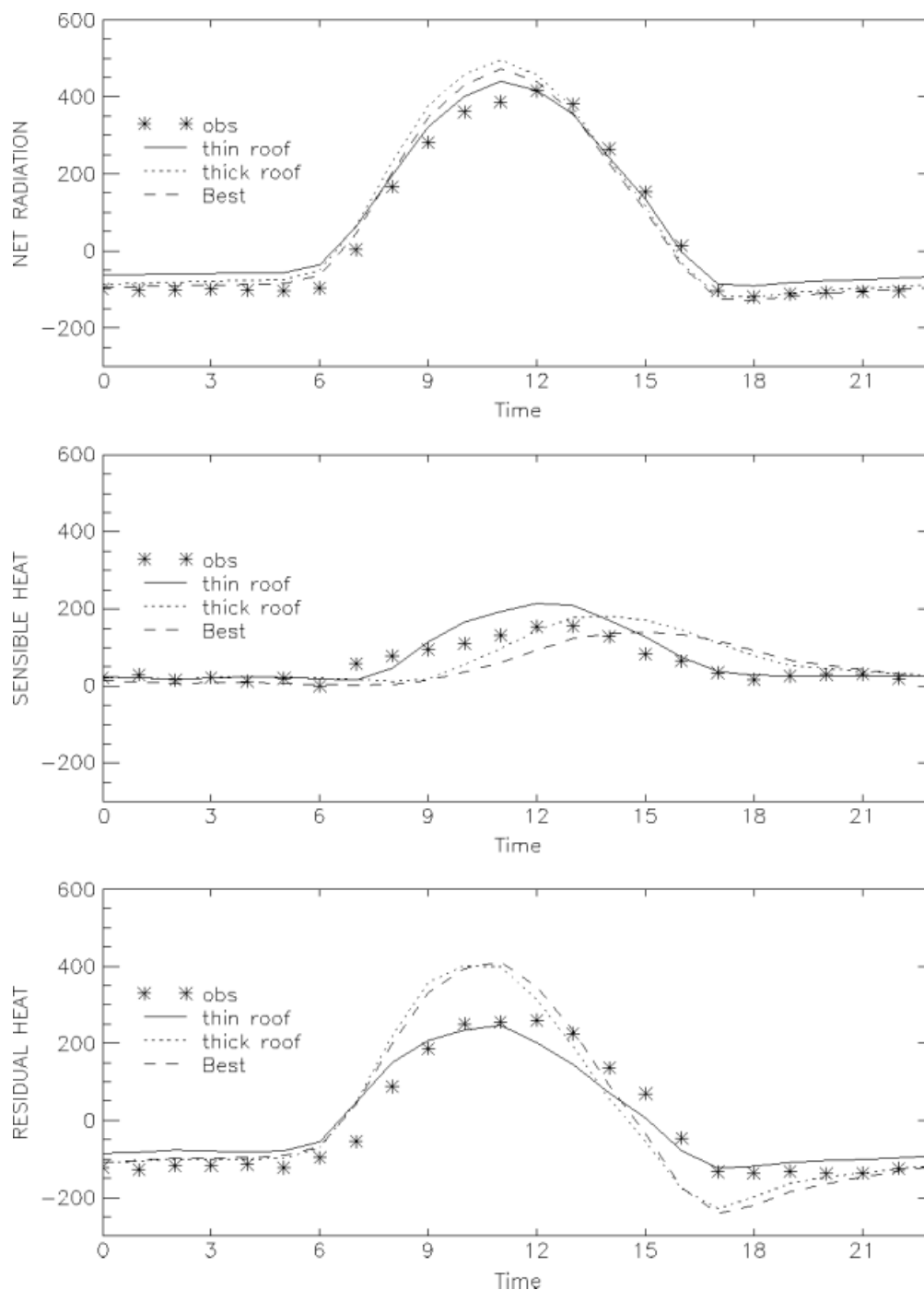
Figure 2. Model outputs of surface fluxes ($W\,m^{-2}$) from MORUSES and Best against observations from Mexico City. Two versions of MORUSES are illustrated (thin roof and thick roof).

Best scheme and MORUSES-thick roof perform similarly (and poorly against the observations of sensible heat and ground heat fluxes) throughout the day. This result is in agreement with the idealized comparison shown in Part I (Porson *et al.*, 2010), for a canyon aspect ratio of $H/W = 1$. However, the two configurations of MORUSES show dramatic differences. The decreased thickness of the roof tile in MORUSES-thin roof allows this model to respond rapidly to the environmental forcing. A smaller roof thickness is related to a smaller damping depth which, in turn, implies that higher frequencies are captured.

This implies that MORUSES-thin roof can simulate the high frequencies and transient forcings (dawn and

dusk) better than MORUSES-thick roof. Again in agreement with the idealized simulations in Part I (Porson *et al.*, 2010), MORUSES-thin roof shows an earlier peak in the sensible heat flux compared with MORUSES-thick roof and, notably, does not lead to significantly negative radiation fluxes during the night. MORUSES-thick roof does simulate the net radiation and ground heat (residual heat) fluxes better than MORUSES-thin roof during most of the night-time, but this is more than offset by the large differences during the evening transition.

The overall performance of the three models is presented in Table IV. The statistics presented therein show the performance of the models for the six-day composite diurnal

Table IV. Rmse values of model outputs of surface fluxes against observations from Mexico City. R_N is net radiation (W m^{-2}), Q_H sensible heat flux (W m^{-2}) and G_N residual heat flux density (W m^{-2}).

rmse	Best	Thick roof	Thin roof	SCRUM-2
All fluxes	57	60	38	37
R_N	35	46	37	31
Q_H	48	43	30	32
G_N	79	82	46	46

cycle, using hourly averages. This reveals that MORUSES-thin roof performs the best against the observations. These results show similar statistical performance to previous studies such as Masson *et al.* (2002) and Best *et al.* (2006). In the latter study, the authors compared MOSES 2.2 against observations of Vancouver and Mexico City and developed a preliminary version of MORUSES-thin roof by splitting the urban tile from Best (2005) into a canyon tile with a large areal heat capacity (0.28 MJ m^{-2} K^{-1}) and a roof tile with a very small areal heat capacity (0.017 MJ m^{-2} K^{-1}).

These simulations show that MORUSES is sensitive to the method chosen to model roof insulation and that its performance can be improved if the higher frequencies in the forcing can be captured. Further insights into why MORUSES behaves in this way are obtained when we compare MORUSES with SCRUM-2 in the next section.

4. Comparison between MORUSES and SCRUM-2

The differences in the way the two schemes model the urban substrate (Figure 1) could potentially lead to notable difference in the energy-balance fluxes and hence provide insight into the optimal method to model the urban heat storage (the third question posed in the Introduction). In this section we compare the performance of MORUSES and SCRUM-2 when forced by the observations from Mexico City. This analysis is firstly centred on the aggregate fluxes and then, secondly, considers the fluxes from the roof and canyon tile separately.

4.1. Aggregate fluxes from Mexico City

We first compare the aggregate fluxes from MORUSES and SCRUM-2. 'Aggregation of fluxes' refers to the total values for the fluxes once the fractional coverage of the surface of roof and canyon is accounted for and, with careful experimental design, aggregate fluxes are equivalent to the observations from mast measurements. The two models are run and their output analyzed as described in section 3.

Figure 3 illustrates the net radiation and sensible heat fluxes for the six-day composite from SCRUM-2, MORUSES-thin roof and MORUSES-thick roof together with the observations. Table IV provides the performance statistics for the three simulations. These demonstrate that SCRUM-2, which implements the heat-diffusion equation with both of the urban tiles, is the best overall model option in these circumstances. After SCRUM-2, the best model option is MORUSES-thin roof. Table IV shows that MORUSES-thin roof outperforms MORUSES-thick roof for all fluxes. However, Figure 3 shows that model

outputs from MORUSES-thin roof capture the phase of the observations correctly, but not the magnitudes of the fluxes. For example, for the sensible heat flux, model outputs from MORUSES-thin roof are in phase with the observations, while MORUSES-thick roof performs better in terms of the magnitude. For the net radiation, MORUSES-thick roof simulated too much radiation during daytime, while MORUSES-thin roof simulates too little radiation at night-time. On balance, the results from MORUSES-thin roof are better than MORUSES-thick roof. The simulations from SCRUM-2 show nicely the benefits of employing the heat diffusion within the substrate.

4.2. Analysis of tile fluxes

An analysis of the fluxes from the two urban tiles of MORUSES and SCRUM-2 is motivated by two considerations. Firstly, analyzing the tile fluxes gives insight into the performance of each model at the aggregate scale. Secondly, as the differences in the methodology to incorporate heat storage within the urban fabric occur at the tile scale, this will also be the scale at which the differences in the simulations are most apparent. This analysis then directly addresses the third of our questions concerning the sensitivity of the model to the choice of method and numerical scheme used to incorporate heat storage.

In order to focus the analysis, we can refine the third research question further. For the canyon facet: Can the force-restore coupling used between the canyon tile and the underlying soil in MORUSES be as efficient at conducting heat as the SCRUM-2 heat-diffusion method? In section 3 we noted that MORUSES-thin roof captures higher frequencies in its response than MORUSES-thick roof and hence has the capacity to adjust to changes in environmental forcing, but at the expense of not capturing lower frequencies. Is MORUSES-thin roof a faithful representation of the roof surface from SCRUM-2 with its multiple layers (which captures both high and low frequencies)?

4.2.1. The canyon facet

Figure 4 shows the composite diurnal variation of the fluxes from the canyon facet of MORUSES and SCRUM-2. The canyon facet of MORUSES-thin roof and MORUSES-thick roof are formulated identically, with the only possible difference in the evolution of the canyon facet in the two configurations arising because of differences in the bulk stability of the surface layer (due to differences in the roof facet surface temperature and sensible heat flux). Hence, for clarity, only fluxes from MORUSES-thick roof are shown. Figure 4 shows that the net radiation from the two models is similar (except for a slight overestimation for MORUSES during daytime), however the sensible heat flux is not: the flux from MORUSES peaks later than that from SCRUM-2. Somewhat surprisingly, however, the night-time sensible heat flux from the two models is similar. This difference in behaviour may be a symptom of a significant deviation between the two schemes, which we explore further.

We next consider the sensitivity of the fluxes from the canyon to the canyon geometry. Most of the physical processes in MORUSES and SCRUM-2 depend on the geometry of the canyon and do so in an identical manner. Figure 5 shows the composite diurnal variation of the
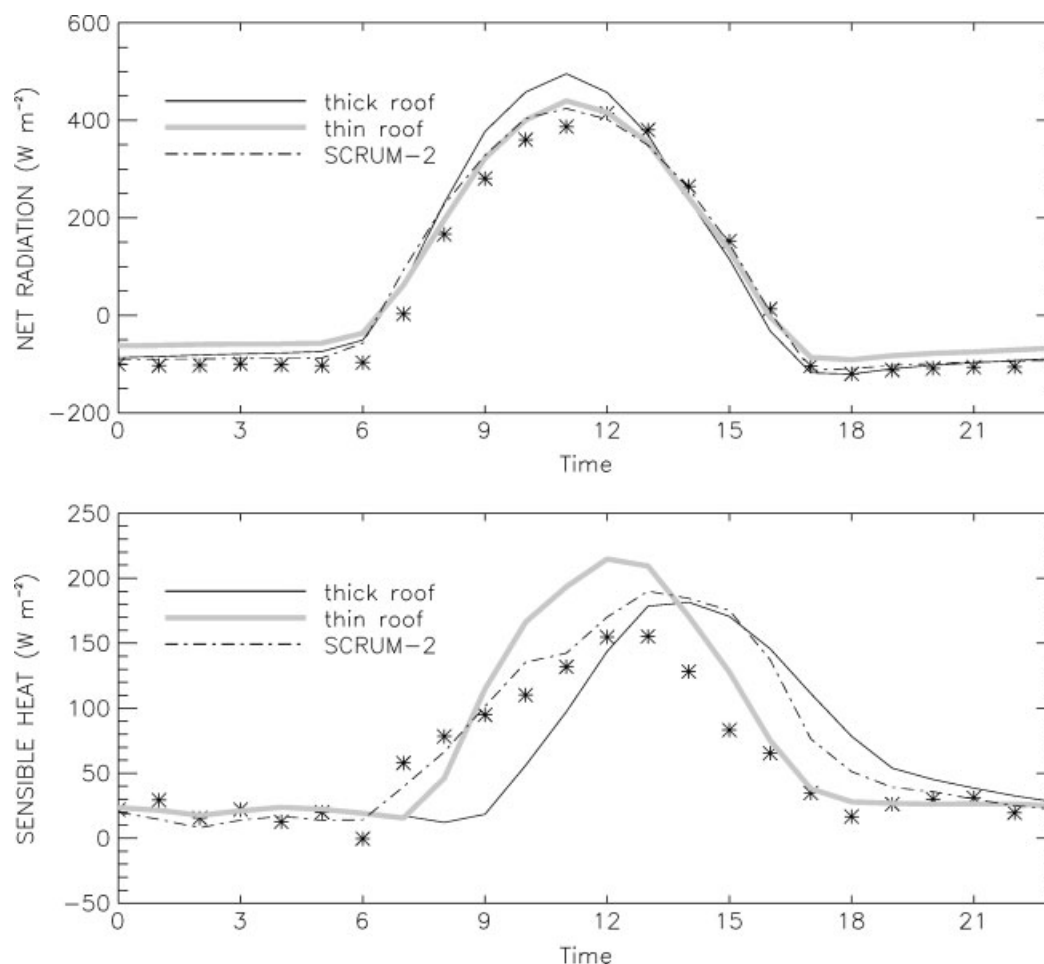
Figure 3. Model outputs of global surface fluxes (W m^{-2}) from MORUSES (thick roof and thin roof) and SCRUM-2. Observations are also illustrated with stars.
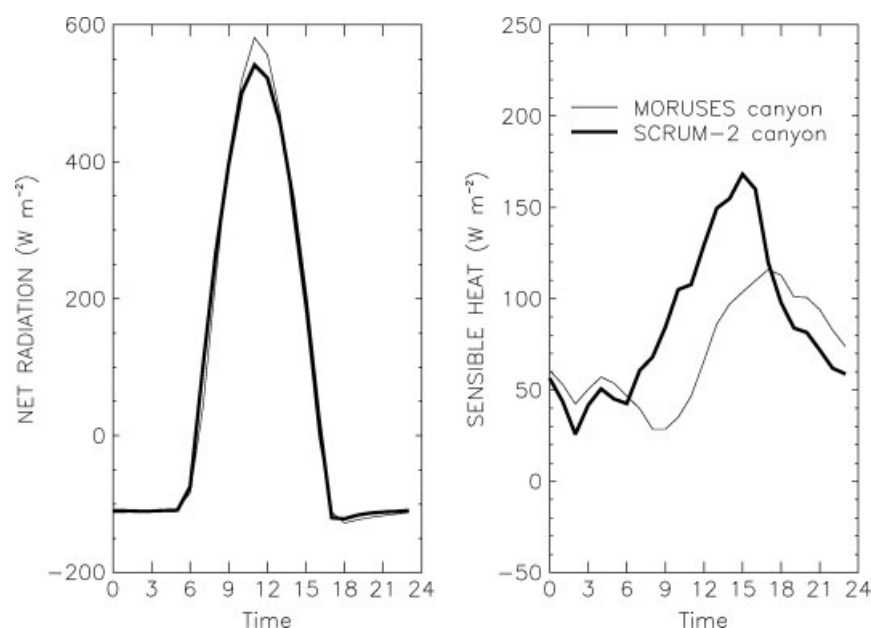
fluxes from the canyon facet as the canyon aspect ratio, H/W, is varied. For these simulations the same atmospheric forcing is used as previously. However, in this figure the models calculate their own values for the effective roughness length for momentum and the displacement height (using Macdonald, 1998) as opposed to using the prescribed values for Mexico City as previously.

Figure 5 shows that the net radiation from MORUSES and SCRUM-2 is similar and indeed converges at high values of H/W. However, the sensible heat flux from the two models differs significantly. On increasing H/W, MORUSES simulates larger magnitudes of the sensible heat flux and, once $H/W > 0.5$, values that peak later. Conversely, SCRUM-2 does not simulate the increased magnitude or phase shift, but the night-time values increase with increasing H/W. For $H/W=3$, the storage capacity of the urban canopy is so large that the sensible heat flux only starts to increase at 1500h local time. It is important to remember that the phase delay, simulated for the canyon tile, is partly corrected by the early phase of the roof tile in order to produce a more realistic behaviour of the urban canopy.

These differences between the canyon facets of SCRUM-2 and MORUSES directly reflect the Fourier frequencies captured by the two models. When the diffusion equation is resolved over multiple layers (as in SCRUM-2), the thickness of each layer is associated with a particular frequency. SCRUM-2 is therefore able to capture multiple frequencies, whereas MORUSES can capture only the

frequency associated with the canyon tile's effective damping depth. By design, the effective damping depth of the canyon tile is relatively thick ($\mathcal{O}(10^{-1}\,\mathrm{m})$) which leads to a much lower frequency being captured by MORUSES than SCRUM-2. For comparison, in SCRUM-2, the urban substrate is resolved over eight layers, with the uppermost layer of the substrate being of $\mathcal{O}(10^{-3}\,\mathrm{m})$ thick; in total the eight layers are $\mathcal{O}(10^{0}\,\mathrm{m})$ thick.

The canyon facet of MORUSES therefore behaves as a thick-slab model (as in Dupont and Mestayer, 2006), with the consequent later peak in the sensible heat flux when compared with SCRUM-2. The similarity in the net radiation is because this term is less sensitive to small differences in the tile temperature than the sensible heat flux. There are other differences between the substrate schemes of MORUSES and SCRUM-2 that could influence these results. For example, MORUSES includes descriptions of substrate moisture dynamics. Additional tests have been undertaken where these soil moisture processes have been neglected; only small differences from the results shown were obtained. Therefore we ascribe the difference in the performance of the canyon tile of MORUSES and SCRUM-2 to the difference in active heat capacity and consequent captured forcing frequency. It is worth noting that, *at this scale*, we cannot assess which implementation method performs the best without having more measurements on the facet properties, or within-canyon fluxes.

Figure 4. Model outputs of canyon surface fluxes (W m^{-2}) from MORUSES and SCRUM-2. Note that, to a first approximation (omitting bulk stability effects), the surface fluxes from the canyon facet are independent of the properties of the roof facet, since the roof is here decoupled from the soil. Note also that the model is driven by atmospheric data.
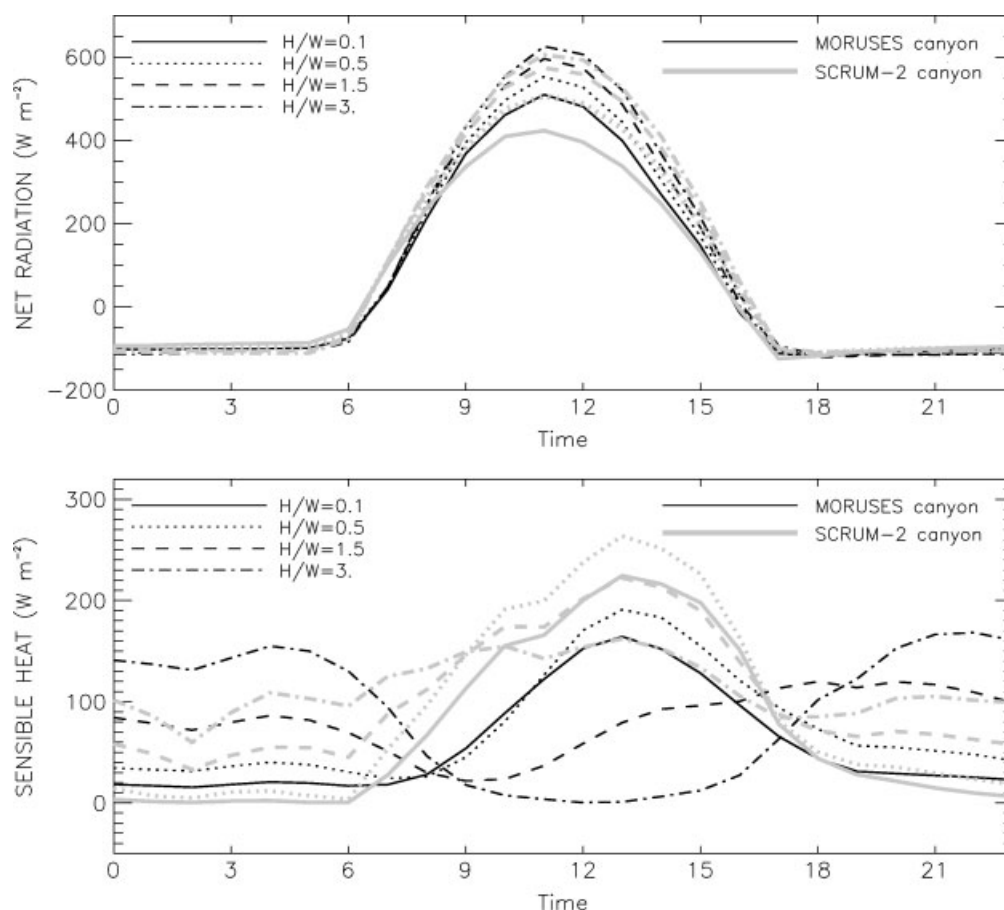
Figure 5. Model outputs of canyon surface fluxes (W m^{-2}) from MORUSES and SCRUM-2, for varying H/W.

4.2.2. The roof facet

Figure 6 shows the fluxes from the roof tile of the two configurations of MORUSES and SCRUM-2. Interestingly, following the analysis at the end of section 4.1, the model outputs from MORUSES-thin roof are in phase with SCRUM-2, while those from MORUSES-thick roof show similar magnitudes to SCRUM-2. This suggests that the two configurations of MORUSES capture extreme behaviours of the roof facet. While the two configurations do imply that different frequencies are captured, this is not the complete picture. Instead, the critical difference lies in the bottom
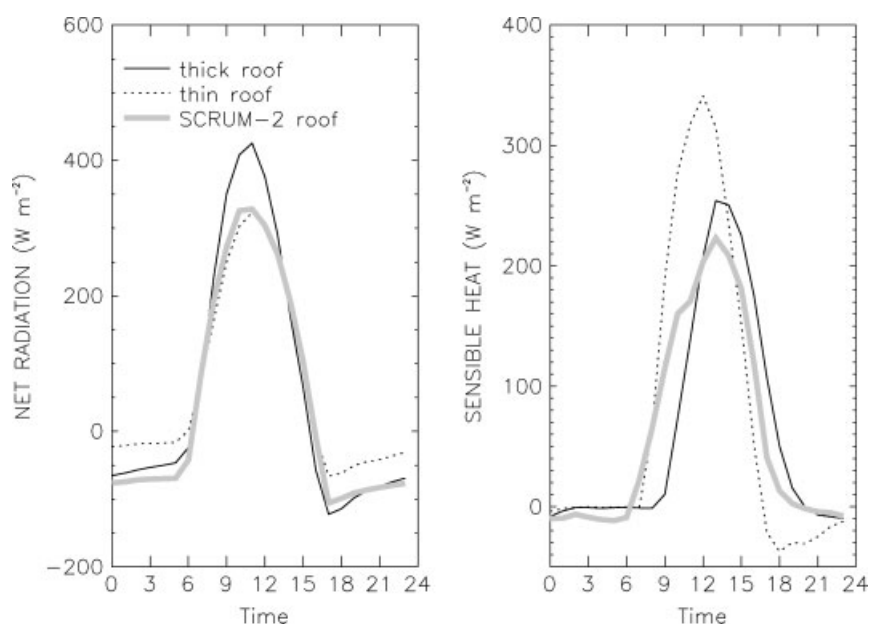
Figure 6. Model outputs of roof surface fluxes (W m^{-2}) from MORUSES (thick roof and thin roof) and SCRUM-2.

boundary condition employed in the two models. Within SCRUM-2 a constant temperature is maintained as the lowest boundary condition; for MORUSES the condition is a zero heat flux. A constant temperature implies a restorative heat flux; it is the absence of this restore term (and the single-layer canopy) that prevents the fluxes from MORUSES from agreeing with those from SCRUM-2, as shown next.

Indeed, there is one characteristic of MORUSES-thin roof that is consistent in the analysis of the aggregate fluxes (Figure 3) and fluxes from the tiles (Figure 6), namely the underestimation of the net radiative fluxes during the night. This characteristic is present with reference not only to the observations but also to the MORUSES-thick roof and SCRUM-2 models. This systematic underestimation strongly suggests that the roof tile is cooling too much and, therefore, that the performance of MORUSES-thin roof could be improved through the addition of a restore term, particularly during the night. In physical terms, the addition of a restore term would represent the presence of a temperature gradient between the base of the roof tile and the interior of the building. This is equivalent to allowing the background temperature T_m, as introduced in section 2.5 of Part I (Porson *et al.*, 2010), to be dependent on z (see Eq. (42) in Part I). The presence of a gradient is justified if energy (heating or air conditioning) is used to maintain an approximately constant building interior air temperature. In Mexico City, it is probably fair to assume that neither air conditioning nor heating would play a major contribution. Instead, ventilation and ground heat exchange could be the dominant factors in maintaining an approximately constant internal building temperature.

Consequently, we introduce two further configurations of MORUSES where such a restore term is included. The restore term is formulated as a radiative exchange term between the roof-tile temperature and a fixed internal temperature, T_b, similarly to the Best scheme (Best 2005). Figure 7 shows a schematic of how this radiative exchange term is implemented in MORUSES. This represents an additional heat flux, and does not conserve energy. For
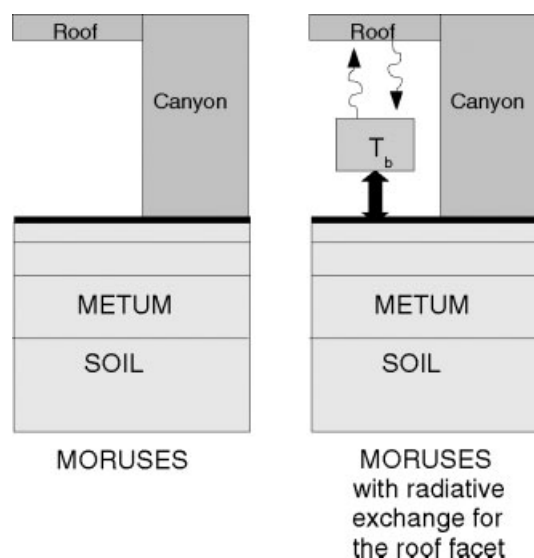

Figure 7. Schematics of MORUSES with and without radiative exchange for the roof facet.

the purposes of this study, a balancing heat flux is taken from the soil to conserve energy in the overall system. This must, in principle, be handled carefully to avoid physically unrealistic behaviour. The internal temperature chosen should be within the diurnal range and not too far from the deep soil temperature so that it acts, effectively, as a very large thermal inertia reservoir similar to the deep soil. Given this constraint, coupling to the internal temperature is not very different from coupling to the soil. For the simulations shown next, the internal temperature T_b is set to 20°C as given by Masson *et al.* (2002). The two configurations are termed 'MORUSES-thin roof rad', where the additional restore term is implemented at all times, and 'MORUSES-thin roof rad night', where the additional restore term is only active during the night. MORUSES-thin roof is characterized by extremely high roof-tile temperature during the day and extremely low roof-tile temperature by night. We therefore expect that the restore term will act
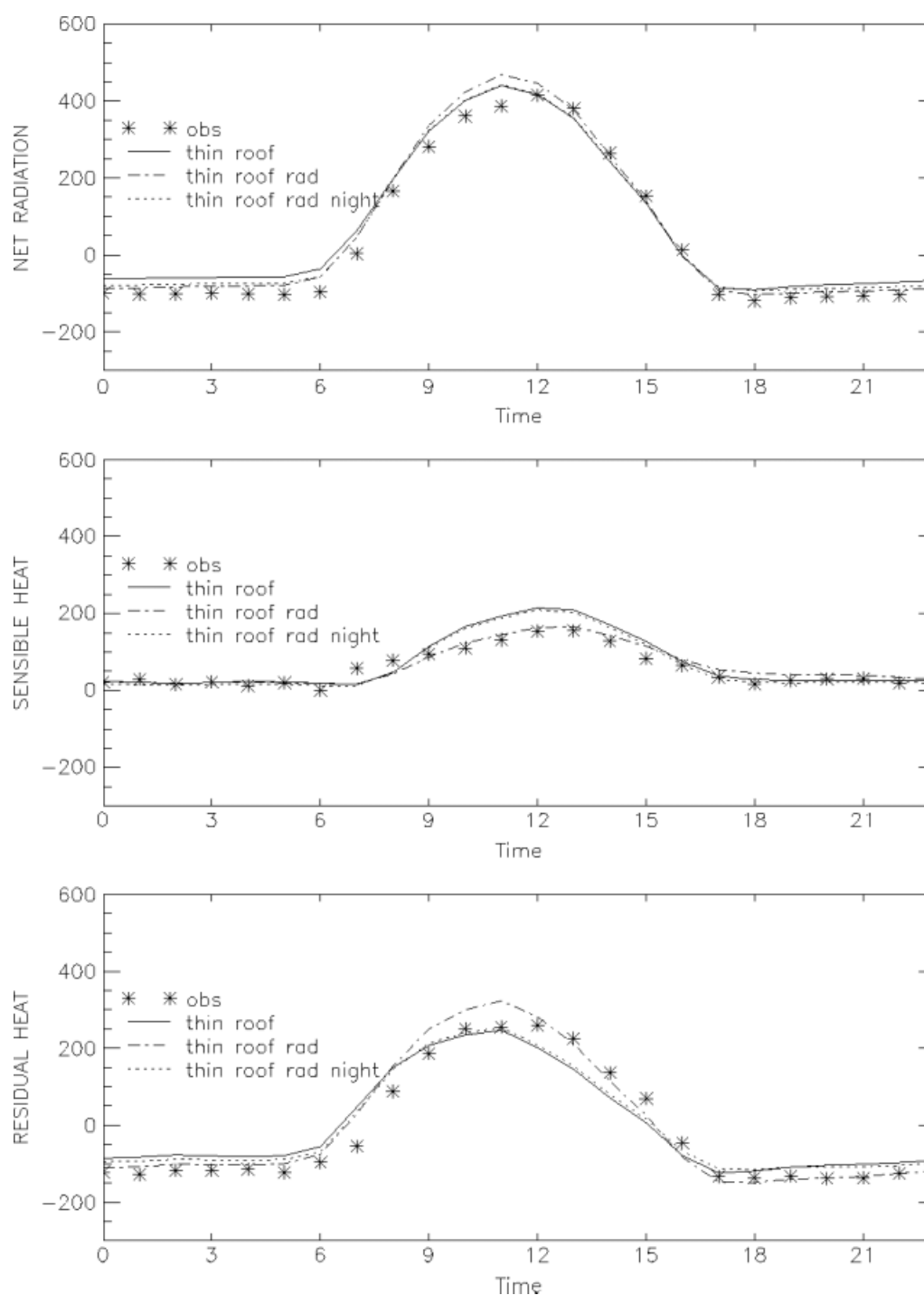
Figure 8. Model outputs of surface fluxes (W m^{-2}) from MORUSES against observations from Mexico City. Three versions of the new scheme are illustrated (thin roof, thin roof rad, thin roof rad night).

as an energy sink during the day and an energy source by night. Note that the restore term arises from the boundary condition of an approximate solution of the heat-diffusion equation and cannot be associated with anthropogenic heat sources as such. It would arise in the case of a passive building, in which case the internal temperature must be chosen to represent the long-time average 'steady-state' temperature. In practice, it would probably be wise to use, for example, a monthly mean rather than a true climatological average.

Figure 8 shows the diurnal composite of the energy balance from these two new configurations, and Table V the associated performance statistics. As expected, both of the new MORUSES configurations improve the overall performance when compared with the MORUSES-thin roof configuration. The new configurations even outperform SCRUM-2 against observations. This improved performance arises through a significantly better simulation of the net radiation. The restore term acts to decrease the roof-tile temperature by day, reducing the outgoing long-wave radiation and therefore increasing the net radiation (as hoped for). By night, the restore term acts to increase the roof-tile temperature, increasing the outgoing long-wave radiation and therefore increasing the magnitude of the net radiation (again as hoped for).

Table V. Rmse values of model outputs of surface fluxes against observations from Mexico City for the thin-roof case, the thin-roof case with radiative exchange during both daytime and night-time and the thin-roof case with radiative exchange only during the night-time. R_N is net radiation $(W\,m^{-2})$, Q_H sensible heat flux $(W\,m^{-2})$ and G_N residual heat flux density $(W\,m^{-2})$.

rmse	Thin roof	Thin rad	Thin rad night
All fluxes	38	29	32
R_N	37	31	29
Q_H	30	17	27
G_N	46	35	39

Table VI. Rmse values of roof model outputs between MORUSES (for different roof surface parametrizations) and SCRUM-2. R_N is net radiation $(W\,m^{-2})$ and Q_H sensible heat flux $(W\,m^{-2})$.

rmse	Thick roof	Thin roof	Thin rad	Thin rad night
R_N	39	37	22	22
Q_H	40	59	32	60

The simulation of the sensible heat flux is also improved if the restore term is active both day and night. Figure 8 shows that the simulated sensible heat flux from 'MORUSES-thin roof rad' performs the best against the observations, but has a slight positive bias during the night. These results are consistent with the changes in roof-tile temperature caused by the restore term.

Figure 9 shows a comparison between SCRUM-2 and the four configurations of MORUSES. As expected, MORUSES-thin roof rad is the closest of the four configurations to SCRUM-2 (see the statistics for the agreement between SCRUM-2 and MORUSES given in Table VI). The results suggest that the roof tile is capable of simulating insulation and should be weakly (radiatively) coupled to an internal building temperature. Future developments of the MetUM and JULES may allow the tiling of the underlying substrate module. In this configuration, these results suggest that the surface under the roof tile should be considered as the building interior, with consequent changes in configuration and interpretation.

A constant internal building temperature is perhaps not what would happen in reality, because air conditioning is not always used to maintain the building interior temperature as a constant. Further research is needed to understand why we obtain a better agreement when the restore term is active during both daytime and night-time.

Finally, Figures 4 and 6 assist us in understanding the performance of all the models against the observations. The good agreement between the sensible heat flux from the MORUSES-thin roof configurations and the observations is explained as the aggregation of a flux from the roof tile, which peaks early during the day, and a flux from the
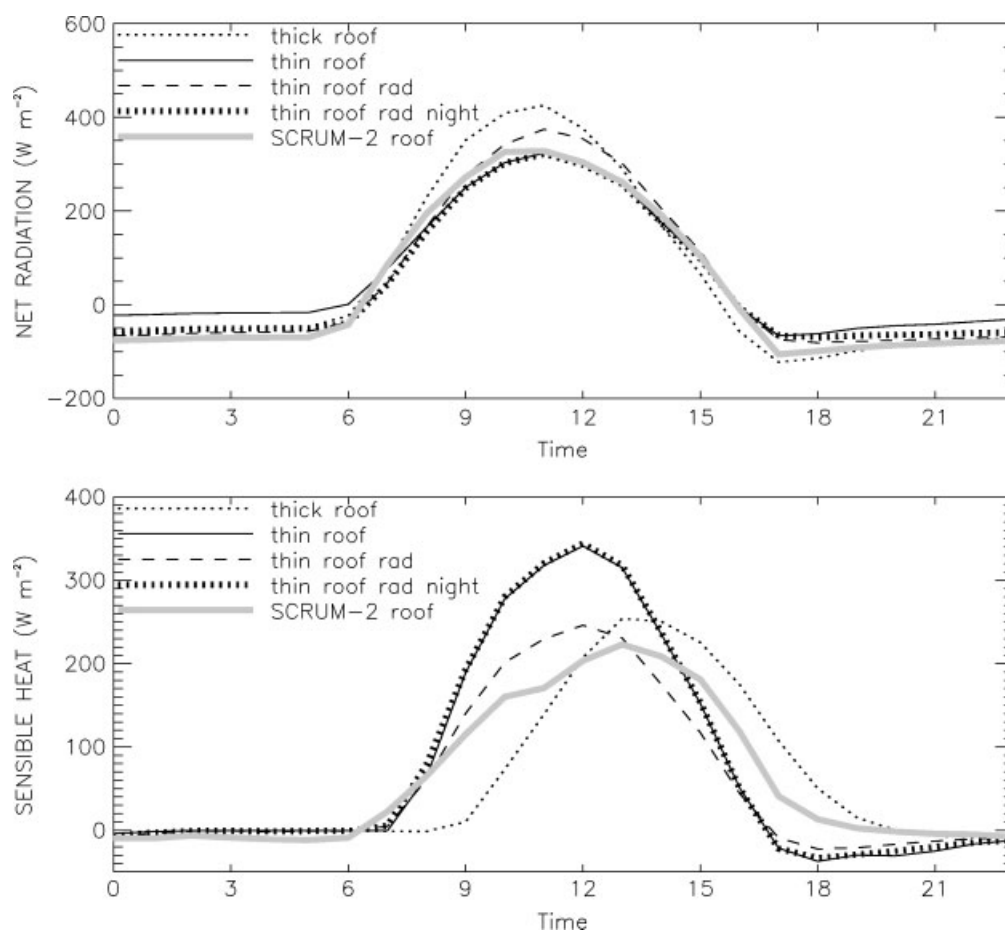

Figure 9. Model outputs of roof surface fluxes $(W\,m^{-2})$ from MORUSES (thick roof, thin roof, thin roof rad, thin roof rad night) and SCRUM-2.

canyon, which peaks later. Due to the inherent frequency response of MORUSES-thick roof, this configuration is not able to simulate the early peak from the roof tile, which leads to poorer agreement with observations. While this explanation is tantalizing, we cannot confirm whether this is what happens in reality without further observations of the fluxes from the individual building facets. Overall, when comparing the model outputs from MORUSES-thin roof and SCRUM-2, we have two models that perform similarly against observations but have very different characteristics at the tile scale. These two models demonstrate that the performance of an urban scheme can depend highly on the methodology and numerical options used to incorporate the heat stored in the building fabric.

5. Conclusions

Part I of this article (Porson *et al.*, 2010) introduced a new urban surface scheme, MORUSES, for use within the MetUM, which disaggregates the urban surface into canyon and roof fractions. Here we analyze the performance of MORUSES against observations from Mexico City, the previous urban surface (Best) scheme and a third model, SCRUM-2, with a higher level of complexity. The default configuration of MORUSES is shown to perform at least as well as the Best scheme against the observations, though not as well as SCRUM-2.

The method used to incorporate roof insulation is critical to the performance of MORUSES. Two optional configurations were considered, firstly by changing the value of the thermal conductivity used and secondly by reducing the thickness of the roof tile. The analytic solution to the heat-diffusion equation shows that this second method applies a control to the frequency at which the roof temperature can vary. Here this method allowed the roof-tile temperature to respond at a higher frequency than either the standard or reduced thermal conductivity configurations of MORUSES. This 'thin roof' option significantly improves the aggregate performance of MORUSES against the observations, particularly with respect to the sensible heat flux and the heat stored in the urban fabric, especially in terms of the phase of the simulations. However, this option also degrades the performance of the net radiation from MORUSES in terms of the magnitude of the simulations. This weakness is traced to the zero-flux condition used as the lower boundary condition on the roof tile; adapting MORUSES further to include a radiative restore term, linked to an internal building temperature, improves the model's performance still further. Interestingly the best performance of MORUSES is obtained if this restore term is utilized during both day and night.

MORUSES and SCRUM-2 are both two-tile models for the urban surface. They are shown to perform similarly at the aggregate scale against observations, but differently at the tile scale. This is despite the radiation and turbulent exchange parametrizations in the two models being the same. This difference in performance is fundamentally related to the difference in methodology and numerical options used to incorporate the heat stored in the building fabric and the coupling with the underlying substrate. These results show that the performance of an urban scheme within NWP models is highly sensitive to the method used to address the heat-diffusion equation and the heat-storage term. Future work includes the comparison of both models through the intercomparison exercise (Grimmond *et al.*, 2010a, 2010b) and the comparison of MORUSES against a set of data over London (UK) within the LUCID project (Bohnenstengel *et al.*, 2010).

Finally, the errors between these urban schemes and the observations are comparable with the current level of agreement between the diurnal surface-energy fluxes from NWP predictions and observations. Edwards *et al.* (2010) found errors of less than $20\,\mathrm{W\,m^{-2}}$ under clear-sky and winter conditions, using a 12 km resolution over flat terrain. Heterogeneity in the land surface may, however, result in a lower agreement between observations and modelled fluxes, as illustrated in Heinemann and Kerschgens (2006) for different types of averaging methods. In particular, further work is needed to assess whether a tiled substrate will improve MORUSES further and hence lead to a more optimal urban surface exchange scheme in terms of accuracy and computational cost.

Acknowledgements

References

Best MJ. 2005. Representing urban areas within operational numerical weather prediction models. *Boundary-Layer Meteorol.* **114**: 91–109.

Best MJ. 2006. Progress towards better weather forecasts for city dwellers: from short range to climate change. *Theor. Appl. Climatol.* **84**: 47–55.

Best MJ, Grimmond CSB, Villani MG. 2006. Evaluation of the urban tile in MOSES using surface energy balance observations. *Boundary-Layer Meteorol.* **118**: 503–525.

Blyth E, Best M, Cox P, Essery R, Boucher O, Harding R, Prentice C, Vidale P, Woodward I. 2006. JULES: A new community land surface model. *Global Change Newsletter* **66**: 1–11.

Bohnenstengel SI, Evans S, Clark PA, Belcher SE. 2010. Simulations of the London urban heat island. *Q. J. R. Meteorol. Soc.* Submitted.

Cox PM, Betts RA, Bunton CB, Essery RLH, Rowntree PR, Smith J. 1999. The impact of the new land surface physics on the GCM simulation of climate and climate sensitivity. *Clim. Dyn.* **15**: 183–203.

Dupont S, Mestayer PG. 2006. Parameterization of the urban energy budget with the submesoscale soil model. *J. Appl. Meteorol. Climatol.* **49**: 1744–1765.

Edwards JM, McGregor JR, Bush MR, Bornemann FJ. 2010. Assessment of numerical weather forecasts against surface observations. *Q. J. R. Meteorol. Soc.* (submitted).

Essery R, Best M, Cox P. 2001. '*MOSES 2.2 Technical Documentation*'. *Hadley Centre Technical Report* No 30. Met Office Hadley Centre: UK.

Grimmond CSB, Oke TR. 1999. Aerodynamic properties of urban areas derived from analysis of surface form. *J. Appl. Meteorol.* **38**: 1262–1292.

Grimmond CSB, Blackett M, Best MJ, Baik J-J, Belcher SE, Beringer J, Bohnenstengel SI, Calmet I, Chen F, Coutts A, Dandou A, Fortuniak K, Gouvea ML, Hamdi R, Hendry M, Kanda M, Kawai T, Kawamoto Y, Kondo H, Krayenhoff ES, Lee S-H, Loridan T, Martilli A, Masson V, Miao S, Oleson K, Ooka R, Pigeon G, Porson A, Ryu Y-H, Salamanca F, Steeneveld GJ, Tombrou M, Voogt JA, Young D, Zhang N. 2010a. Initial results from Phase 2 of the international urban energy balance comparison project. *Int. J. Climatol.* (submitted).

Grimmond CSB, Blackett M, Best M, Barlow J, Baik JJ, S Belcher S, Bohnenstengel SI, Calmet I, Chen F, Dandou A, Fortuniak K, Gouvea ML, Hamdi R, Hendry M, Kawai K, Kawamoto Y, Kondo H, Krayenhoff ES, Lee SH, Loridan T, Martilli A, Masson V, Miao S, Oleson K, Pigeon G, Porson A, Ryu YH, Salamanca F, Steeneveld GJ, Tombrou M, Voogt J, Young D, Zhang N. 2010b. The International Urban Energy Balance Models Comparison Project: First results from Phase 1. *Journal of Applied Meteorology and Climatology* DOI: 10.1175/2010JAMC2354.1.

Harman IN, Best MJ, Belcher SE. 2004a. Radiative exchange in an urban street canyon. *Boundary-Layer Meteorol.* **110**: 301–316.

Harman IN, Barlow JF, Belcher SE. 2004b. Scalar fluxes from urban street canyons. Part II: Model. *Boundary-Layer Meteorol.* **113**: 387–410.

Harman IN, Belcher SE. 2006. The surface energy balance and boundary-layer over urban areas. *Q. J. R. Meteorol. Soc.* **132**: 2749–2768.

Heinemann G, Kerschgens. 2006. Simulation of surface energy fluxes using high-resolution non-hydrostatic simulations and comparisons with measurements for the LITFASS-2003 experiments. *Boundary-Layer Meteorol.* **121**: 195–220.

Holt T, Pullen J. 2007. Urban canopy modeling of the New York City Metropolitan area: A comparison and validation of single- and multilayer parameterizations. *Mon. Weather Rev.* **135**: 1907–1930.

Masson V, Grimmond CSB, Oke TR. 2002. Evaluation of the Town Energy Balance (TEB) scheme with direct measurements from dry districts in two cities. *J. Appl. Meteorol.* **11**: 1011–1026.

Oke TR, Spronken-Smith RA, Jauregui E, Grimmond CSB. 1999. The energy balance of central Mexico City during the dry season. *Atmos. Environ.* **33**: 3919–3930.

Porson A, Clark PA, Harman IN, Best MJ, Belcher SE. 2010. Implementation of a new urban energy budget scheme in the MetUM. Part I: Description and idealized simulations. *Q. J. R. Meteorol. Soc.* **136**: 1514–1529, DOI: 10.1002/qj.668.

Porson A, Harman IN, Belcher SE. 2009. How many facets are needed to represent the surface energy balance of an urban area? *Boundary-Layer Meteorol.* **132**: 107–128.