

THE INTERNATIONAL URBAN ENERGY BALANCE COMPARISON PROJECT: INITIAL RESULTS FROM PHASE 2

CSB Grimmond^{1†}, M Blakett¹, M Best², J-J Baik³, S Belcher⁴, SI Bohnenstengel⁴, I Calmet⁶, F Chen⁷, A Dandou⁸, K Fortuniak⁹, M Gouvea¹, R Hamdi¹⁰, M Hendry², H Kondo¹¹, S Krayenhoff¹², S-H Lee³, T Loridan¹, A Martilli¹³, V Masson¹⁴, S Miao¹⁵, K Oleson⁷, G Pigeon¹⁴, A Porson^{2,4}, F Salamanca¹³, G-J Steeneveld¹⁶, M Tombrou⁸, J Voogt¹⁷, N Zhang¹⁸.

¹King's College London, UK, ²UK Met Office, UK, ³Seoul National University, Korea, ⁴Univ. Reading, UK, ⁶CNRS-Ecole Centrale de Nantes, France, ⁷National Center for Atmospheric Research, USA, ⁸Univ. Athens, Greece, ⁹Univ. Łódź, Poland, ¹⁰Royal Meteorological Institute, Belgium, ¹¹National Institute of Advanced Industrial Science and Technology, Japan, ¹²Univ. British Columbia, Canada, ¹³CIEMAT, Spain, ¹⁴Meteo France, France, ¹⁵IUM, CMA, China, ¹⁶Wageningen Univ, Netherlands, ¹⁷Univ. Western Ontario, Canada, ¹⁸Nanjing Univ., China.

ABSTRACT

Many urban land surface schemes have been developed, incorporating different assumptions about the features of, and processes occurring at, the surface. Here, the first results from Phase 2 of an international comparison are presented. Evaluation is based on analysis of the last 12 months of a 15 month dataset. In general, the schemes have best overall capability to model net all-wave radiation. The models that perform well for one flux do not necessarily perform well for other fluxes. Generally there is better performance for net all wave radiation than sensible heat flux. The degree of complexity included in the models is outlined, and impacts on model performance are discussed in terms of the data made available to modellers at four successive stages.

Key words: Energy balance; fluxes; model complexity

1. INTRODUCTION

The increasingly urbanised global population, predicted to be nearly 70% by 2050, combined with enhanced computer resources, which allow greater spatial resolution to be resolved, have created greater demand for weather and climate information in cities. The surface morphology (existence of buildings, urban canyons), presence of impervious building materials, sparseness of vegetation, anthropogenic heat contributions, and bluff body nature of the buildings, combine to change energy partitioning in urban areas. Land surface schemes, used within numerical models, need to account for the urban influences on the surface-atmosphere exchanges. However, the complexity of these schemes has to be balanced with computational and data demands for particular objectives (e.g. global climate modelling, operational weather forecasting).

Here, first results from Phase 2 of an international comparison of 25 urban land surface schemes (Grimmond et al. 2009a,b) are presented. The fundamental requirement for the models to be included is that they model the urban energy balance fluxes:

$$Q^* + Q_F = Q_H + Q_E + \Delta Q_S$$

The models are forced with the incoming short- ($K\downarrow$) and long-wave fluxes ($L\downarrow$), air-temperature, pressure, specific humidity and wind components. From these the outgoing radiative fluxes ($K\uparrow$, $L\uparrow$) and net all wave radiation (Q^*), turbulent sensible heat (Q_H), turbulent latent heat (Q_E), storage heat flux (ΔQ_S) and the anthropogenic heat (Q_F) may be determined.

Whilst many urban models have been evaluated against observational datasets, these comparisons have not been conducted in a controlled manner that allows robust model inter-comparison. The objective here is to do just that; to undertake a staged and controlled comparison of urban energy balance models and their performance. This will assist in determining those modelling approaches which result in the best model performance.

2. METHODOLOGY

The methodology used here follows PILPS (Project for Intercomparison of Land-Surface Parameterization Schemes) (Henderson-Sellers et al. 1993). This allows the relative importance of key parameters to be determined and provides for an assessment of the level of model complexity required for optimal performance. Individual groups run their model(s) 'offline' using forcing data provided to allow the performance of the land surface schemes to be examined while the atmospheric conditions are fixed and not a function of the performance of a larger scale model.

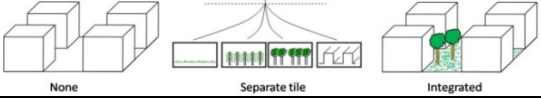
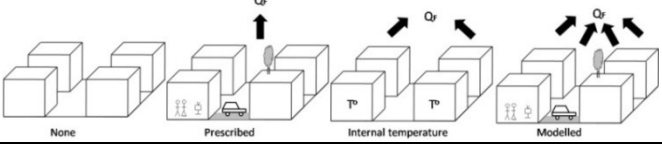
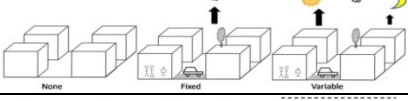
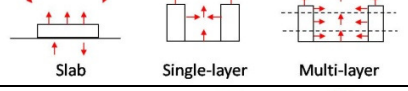
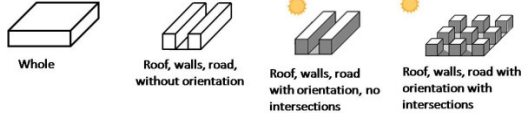
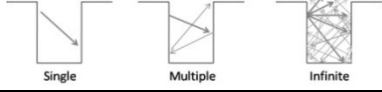
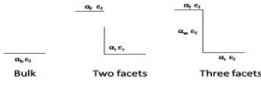
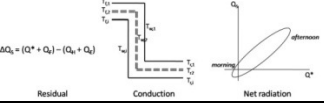
Results presented here are for the first four stages of Phase 2. Phase 1 results are provided in Grimmond et al. (2009). In Phase 2/Stage 1, participants were provided with 15 months of forcing data from an urban site (here termed "alpha") (Table 1). At subsequent stages, additional site information was provided (Table 1): Stage 2 - pervious/impervious fraction, Stage 3 - urban morphology, and Stage 4 -urban materials. From the information provided, further parameters potentially could be derived if needed by individual modelling groups

The models are classified using eight characteristics (Table 2) (see Grimmond et al. 2009b for more details). Each model was assigned a random identifier number so that model performance is anonymous but class performance can be seen. Here the last 12 months of the data set are analysed to allow an initialisation or 'spin-up' period for the models. This includes 8520 30-min periods when all fluxes are observed. Statistics used to compare the models include the root mean square error (RMSE), systematic RMSE (RMSE_s) and unsystematic RMSE (RMSE_u).

Table 1: Data provided at each stage of Phase 2.

Stage	Category	Data provided
1	Observations	K↓, L↓, air temperature, station pressure, specific humidity, wind components, rainfall
	Site	Latitude, Longitude, Measurement height: 6.25 mean roughness height
2	Plan area fraction	Pervious, Impervious.
3	Heights	Instrument height, roughness length for momentum, maximum height of roughness elements, mean building height, height:width, mean wall to plan area
	Plan area fraction	Buildings, concrete, road, vegetation (excl. grass), grass and other (bare, pools)
	Other	Urban climate zone; population density
4	Material characteristics	Thickness, specific heat, volumetric heat capacity, thermal conductivity, type: road, roof and wall layers

Table 2: Description of model classes and number of models (in Stages 1-4, and their capability, Cap) included in the model comparison in Phase 2, for each stage (modified from Grimmond et al. 2009b).

Class	Cap	1	2	3	4	Representation
1 Vegetation (V)	Not included (n)	7	10	7	7	
	Separate tile (s)	16	13	16	16	
	Integrated (i)	2	2	2	2	
2 Q _f (AN)	Negligible or ignored (n)	9	15	15	15	
	Prescribed (p)	5	2	2	2	
	Internal Temp. (i)*	3	3	3	3	
	Modelled (m)	6	3	3	3	
	i,p*	2	2	2	2	
3 Temporal Q _f variation (T)	None (n)	9	15	15	14	
	Fixed (f)	3	3	3	4	
	Variable (v)	13	7	7	7	
4 Urban Morphology (L)	Slab(s)	8	8	8	8	
	Single layer(1)	10	10	10	10	
	Multiple layer (m)	7	7	7	7	
5 Facets & orientation (FO)	Whole (w)	3	3	3	3	
	No orientation (n)	14	14	14	14	
	Orientation (o) no intersections	5	5	5	5	
	Orientation (i) with intersections	3	3	3	3	
6 Reflection (R)	Single (s)	7	8	8	8	
	Multiple (m)	11	10	10	10	
	Infinite (i)	7	7	7	7	
7 Albedo, Emissivity (AE)	Bulk (b)	3	4	3	3	
	Two facets (2)	4	4	4	4	
	Three or more facets (f)	18	17	18	18	
8 ΔQ _s (S)	Residual (r)*	4	5	5	5	
	Conduction (c)	20	19	19	19	
	Net radiation based (n)*	1	1	1	1	

3. RESULTS

When the individual model RMSE are ranked for Stage 1 (Figure 1), the Q* RMSE varies between 10 and 107 W m⁻² - all but four models have RMSE less than 35 W m⁻². A similar pattern is evident for Q_H, although with higher RMSE values (35-130 W m⁻²). Not unexpectedly, the three poorer performing models for Q_H are those that perform poorest for Q*. For Q_E, model performance does vary (40-58 W m⁻²), although without a stepped change in performance. It is evident that models which perform well/poorly for one flux do not necessarily perform well/poorly for other fluxes (e.g. model 16, 43).

Comparison of the model performance between Stages 1-4 (Table 3) shows that for Q* and Q_H the maximum RMSE decreases, i.e. model performance improves, with each stage. The mean RMSE, which gives a measure of overall change for the models, show an improvement in all stages for Q_H and Q_E but only for Stages 1-2 for Q*. For Q* and Q_H the models, on average, have a smaller RMSE_s than RMSE_u which suggests that the mean performance cannot be improved because of a systematic bias. However for Q_E the mean RMSE_s is greater than RMSE_u which clearly is related to whether or not vegetation is included in the models (Table 2).

Taylor plots (Fig. 2) show that for most models there is an improvement in daytime Q*, although the three original outliers are still there by Stage 4. For Q* a number of the models are close to the green square (perfect agreement) and by Stage 4 there is one model that is almost perfect. For Q_H some of the outliers for daytime fluxes are reduced by

Stage 4. For daytime Q_E , model performance suggests least ability to model this flux by Stage 4 but the greatest improvement is seen for this flux. At night the performance is generally poorer for all fluxes. The poorest is clearly Q_H with some models having normalized standard deviations greater than 1.5, so they plot outside of the default graph axes (Fig. 3). Individual models can be tracked between stages and fluxes (Figure 1, 2). For example for night time Q_H model 36 shows clear improvement from Stages 1 to 3 but not for Stage 4. Clearly there are other models that show consistent improvement throughout. For all models at night, for all stages, the correlation coefficient is: less than 0.4 for Q_E ; between 0.4 and 0.8 for Q_H ; and larger than 0.7 for Q^* . So Q^* is again the best modelled flux but the daytime performance is clearly better; for example the best correlation coefficient is still ≤ 0.99 for Stage 4.

It is clear that providing additional parameter information at each stage has resulted in improvements but at the latter stages the (in)ability of individual models to use that data means a continued net improvement is not evident. The latter stages also represent data that it is more difficult to obtain in a spatially explicit way. Here extensive contact with architects, builders, building suppliers and planners in the area was used to obtain realistic values.

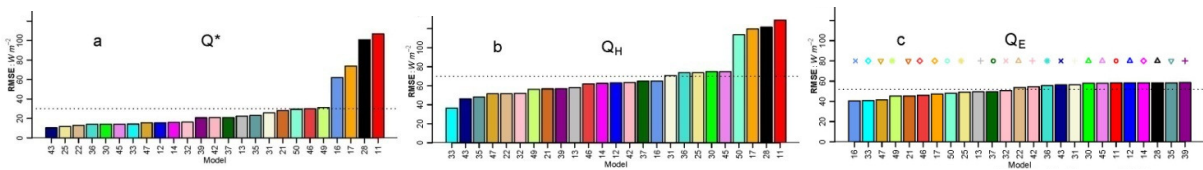


Fig. 1: RMSE for each of the 25 models based on analysis of the last 12 months of the ‘alpha’ data set for: (a) net all wave radiation (b) turbulent sensible heat flux and (c) latent heat flux the fluxes: Q^* , Q_H and Q_E . Also shown in (c) are the coloured symbols used to represent each of the models in Figure 2.

Table 3: Statistics (\bar{x} =mean, σ =standard deviation, R^2 = coefficient of determination, $RMSE_{sys}$ = root mean squared error, systematic, unsystematic) for each flux by Stage (1-4): minimum, maximum and mean performance of the 25 models.

		Q^*				Q_H				Q_E			
		1	2	3	4	1	2	3	4	1	2	3	4
\bar{x} ($W m^{-2}$)	Obs	78.9				37.9				32.5			
	Max	144.6	120.1	120.1	99.4	124.1	125.7	125.7	106.4	43.6	66.8	68.2	68.2
	Min	55.2	62.0	60.7	49.6	35.3	6.0	16.8	5.6	0.0	0.0	0.0	0.0
	Mean	77.6	77.3	75.2	72.6	65.6	50.9	54.6	49.7	11.3	25.1	24.7	25.2
σ ($W m^{-2}$)	Obs	212.2				92.9				48.4			
	Max	252.3	252.3	242.0	241.7	170.6	170.6	170.9	170.0	56.2	86.4	87.1	88.1
	Min	194.7	193.3	191.3	161.1	65.1	30.0	54.7	77.7	0.0	0.0	0.0	0.0
	Mean	214.4	213.0	208.1	201.4	120.3	110.6	107.3	121.1	19.7	32.7	32.6	36.0
R^2	Max	1.00	1.00	1.00	1.00	0.79	0.82	0.83	0.83	0.20	0.29	0.27	0.28
	Min	0.68	0.68	0.68	0.68	0.13	0.13	0.10	0.29	0.00	0.01	0.01	0.02
	Mean	0.96	0.96	0.95	0.95	0.60	0.61	0.60	0.70	0.03	0.13	0.13	0.17
RMSE ($W m^{-2}$)	Max	107.0	107.0	101.8	101.7	129.0	129.0	124.3	122.1	58.6	76.0	77.2	77.9
	Min	10.4	9.5	8.6	10.9	36.5	36.5	36.1	35.2	40.6	36.5	35.0	34.6
	Mean	30.0	28.6	30.1	33.1	69.8	63.1	61.9	58.7	51.8	46.6	45.8	45.8
RMSE _s ($W m^{-2}$)	Max	69.8	44.6	44.6	62.2	94.3	93.8	97.6	82.8	58.3	58.3	58.3	58.3
	Min	1.2	2.3	1.4	7.1	7.8	7.8	8.2	11.1	17.9	12.2	11.0	8.7
	Mean	15.2	13.4	15.7	19.0	40.9	36.2	35.7	31.0	46.4	33.4	33.0	30.7
RMSE _u ($W m^{-2}$)	Max	104.4	104.4	100.9	100.6	99.9	94.8	94.9	94.5	44.5	67.6	68.2	68.9
	Min	6.0	5.8	5.8	6.1	34.9	18.2	27.1	27.7	0.0	0.0	0.0	0.0
	Mean	23.2	23.3	23.4	24.5	54.9	48.9	47.8	48.3	17.7	25.2	25.3	27.4

Acknowledgements: Funds to support this work (Grimmond): UK Met Office (P001550), EU framework 7 (7 FP7-ENV-2007-1) projects MEGAPOLI (212520), BRIDGE (211345). This work contributes to COST728. We would like to thank all who were involved in the collection of the original dataset (alpha, this to remain anonymous until all groups reach stage 5) and organisations and individuals who provided data that supported that work. Funding from BSIK-COM29 (Steenefeld) and the agencies that support the considerable time contributed by participating groups are acknowledged.

REFERENCES

Grimmond et al. 2009a: in *Meteorological and air quality models for urban areas*, eds A Baklanov, S Grimmond, A Mahura, M Athanassiadou Springer Verlag
 Grimmond et al. 2000b: The international urban energy balance models comparison project: first results from phase 1. (in prep).
 Henderson-Sellers A, Yang Z, Dickenson RE 1993: The Project for Intercomparison of Land-surface Parameterization Schemes, *Bull Amer. Met. Soc.*, **74**, 1335-1349.
 Taylor K 2001: Summarizing multiple aspects of model performance in single diagram, *J. Geophys. Res.*, **106**, D7, 7183-7192.

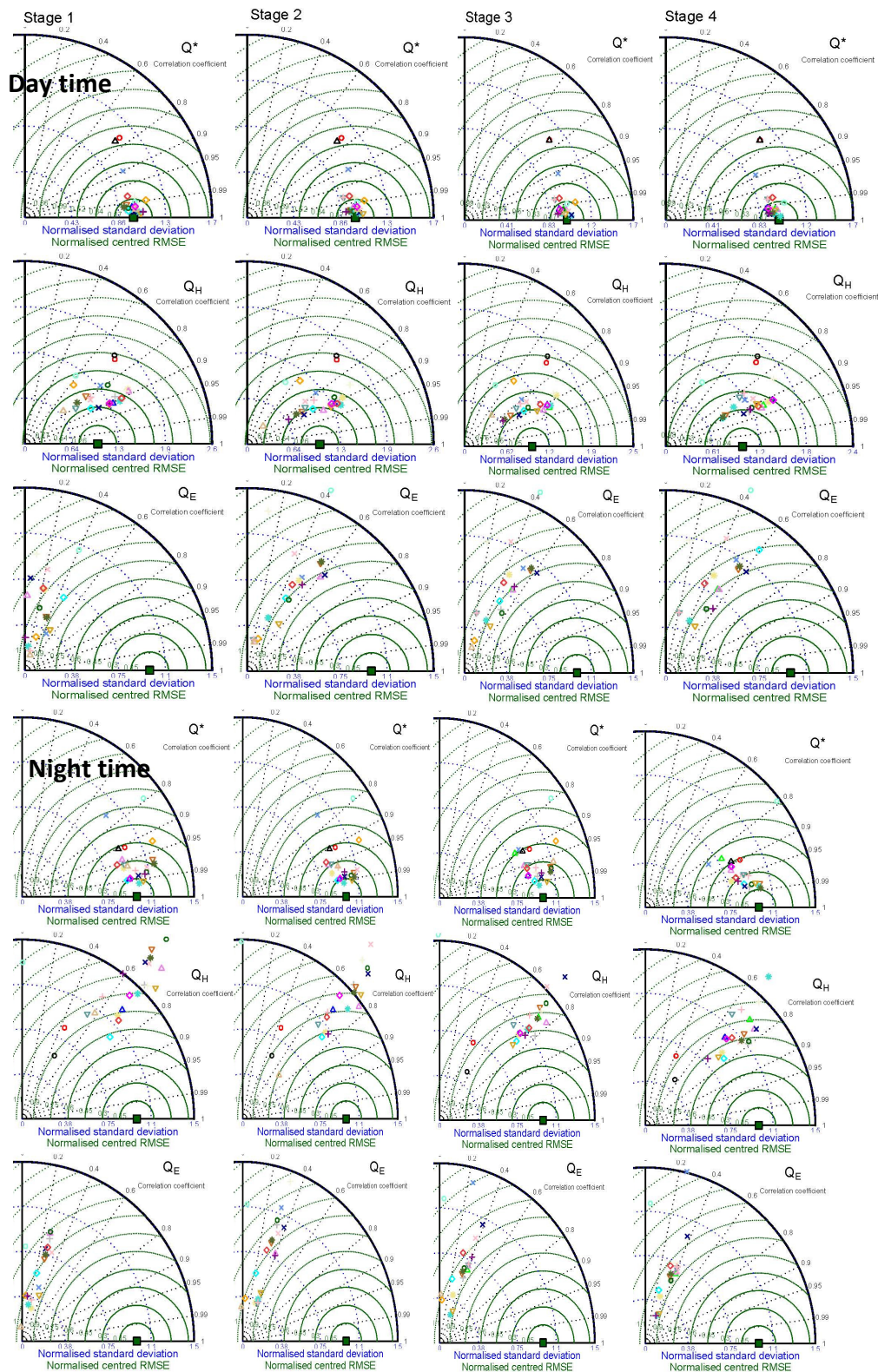


Fig. 2. Normalised Taylor plot with results for each model, for day (upper 3 rows) and night (lower 3 rows) time period, by Q^* (row 1, 4), Q_H (row 2,5) and Q_E (row 3,6) fluxes by Stage (columns 1-4). Plots display the correlation coefficient in relation to the polar axis, the normalised standard deviation in relation to the horizontal axis and the normalised RMSE in relation to the internal circular axes (Taylor, 2001). Day: 1 h after $Q^* \geq 0 \text{ W m}^{-2}$; Night: 1 h after $Q^* \leq 0 \text{ W m}^{-2}$