Numerical Weather Prediction

Assessment of the first version of the Site Specific Forecast Model

Forecasting Research Technical Report No. 213

P.A. Clark, W.P. Hopwood, M.J. Best, C.C. Dunlop and P.E.Maisey

email: nwp_publications@metoffice.com

Forecasting Research
Technical Report No. 213

Assessment of the first version of the Site Specific Forecast Model

by

P. A. Clark, W. P. Hopwood, M. J. Best, C. C. Dunlop, P. E. Maisey

June 1997

Forecasting Research
Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

Abstract

Assessment of the first version of the Site Specific Forecast Model

Summary

The Site Specific Forecast Model (SSFM) development project aims to develop a relatively inexpensive automated model to add site specific forecast skill to NWP products. The model is 1D based upon the UM 'physics' routines, with some small modifications, and driven by output from operational NWP product (currently only the mesoscale model, though this is not a restriction). This report contains the summary report for stage 2 of the project (with a few minor updates from that accepted by the project board).

Stage 2 of the project involved the construction of the first practical version of the model and its testing. This version implemented the following features:

i. Higher resolution in the boundary layer.
ii. Improvements to the surface exchange formulation.
iii. Treatment of alternative surface characteristics more appropriate to the site than the mesoscale model gridbox value.
iv. Implementation of the MOSES soil/surface scheme, with its multi-level soil moisture scheme and with the possibility of higher resolution in the soil temperature scheme.
v. Coupling to mesoscale model output.

The first version of the model has been tested for 14 synoptic sites and one research site (MRU Cardington) over a continuous 3 month period. It is not possible, for this number of sites, to give a subjective assessment of all forecasts, but some clear trends emerged which have relevance to both our model and the UM in general. Statistics were gathered for the synoptic sites showing comparison with observations. The same statistics were gathered for mesoscale forecasts interpolated bi-linearly to the same location for comparison.

Objective analysis of the statistics has shown mixed results but, of primary importance, the SSFM modelwith vegetation canopy reduces RMS errors of nighttime minimum temperatures by up to 0.2-0.3 ˚C compared with the mesoscale model. This, in itself, is significant when compared with improvements that can be achieved by manual intervention. Performance for visibility was similar to the mesoscale model, which was extremely poor during the period covered. Performance for the few radiation fog events captured was often much better than the mesoscale model, and it is thought that the unusually poor performance resulted from the impact of poor UM ice microphysics during the December cold period.

The 15th site, Cardington, was mainly run to provide data for comparison with their routinely run surface site, which is measuring radiative and turbulent fluxes along with time mean temperatures, wind and humidities, and with 'BALTHUM' ascent data. The results from the surface site will be reported in full in a separate document. However, results and insights from both the surface and BALTHUM data have highlighted significant deficiencies in UM ice microphysics and the MOSES surface exchange scheme. The latter has re-confirmed the need to take into account the effect of vegetation on the surface heat balance. Comparison of the MOSES simulation with observed moisture flux has shown deficiencies which may be remedied by modifying the temperature response of C3 grass stomatal resistance and the treatment of canopy water evaporation.

Contents

Assessment of the first version of the Site Specific Forecast Model

Introduction

The Site Specific Forecast model development project aims to develop a relatively inexpensive automated model to add site specific forecast skill to NWP products. The model is 1D based upon the UM 'physics' routines, with some small modifications, and driven by output from operational NWP product (currently only the mesoscale model, though this is not a restriction).

The development project is described in the PID (Clark, 1996). Stage 2 of the project involved the construction of the first practical version of the model and its testing. This version implemented the following features:

i. Higher resolution in the boundary layer.
ii. Improvements to the surface exchange formulation.
iii. Treatment of alternative surface characteristics more appropriate to the site than the mesoscale model gridbox value.
iv. Implementation of the MOSES soil/surface scheme, with its multi-level soil moisture scheme and with the possibility of higher resolution in the soil temperature scheme.
v. Coupling to mesoscale model output.

It must be emphasised that, at this stage, there is no account taken whatsoever of sub-mesoscale grid orography, which will be addressed at the next stage of the project. This includes the treatment of differences in altitude between an actual site and the mesoscale orography at that location. This could be addressed in a relatively simple way at the output stage, but the same correction could be made to mesoscale output, so there seems little point in doing so. It would be more appropriate to adjust the mesoscale forcing data for altitude changes, but doing so raises sufficient questions about other effects of sub-grid orography on the structure of the boundary layer that it has been decided to treat the whole issue together in the next stage of the project. Similarly, no use is being made of local observations. This aspect will be addressed in stage 4 of the project.

These constraints place some limitations on the performance to be expected from this version of the model. In particular, the impact of 'orographic roughness' raises some questions. Since we wish to model explicitly the impact of local surface conditions, using the mesoscale gridbox mean orographic roughness would be inappropriate. The forcing data which we have used (from the mesoscale) already takes account of orographic roughness on the large scale flow, so we have confined ourselves to using the vegetative roughness. However, at this stage we have used a relatively long effective fetch (see below) in order to emphasize the impact of the new surface, with no account taken of local orography in the pressure gradient. As a result, we anticipate somewhat over-estimating the wind speed at most sites. This will be properly rectified in the next stage of the project.

The first version of the model has been tested for 15 sites over a continuous 3 month period.

A number of configurations were tested in order to show the impact of various modifications. The detailed specifications are summarized below, followed by an overview of the performance. The performance was monitored routinely. It is not possible, for this number of sites, to give a subjective assessment of all forecasts, but some clear trends emerged which have relevance to both our model and the UM in general. These will be discussed below. Statistics were gathered for 14 sites showing comparison with synoptic observations. The same statistics were gathered for mesoscale forecasts interpolated bi-linearly to the same location for comparison. The 14 sites used were as follows:

Glasgow Airport	03140	55.87N	4.43W
Manchester Airport	03334	53.35N	2.27W
Waddington	03377	53.17N	0.52W
Cranwell	03379	53.03N	0.50W
Coningsby	03391	53.08N	0.17W
Wittering	03462	52.62N	0.47W
Aberporth	03502	52.13N	4.57W
Birmingham Airport	03534	52.45N	1.73W
Church Lawford	03544	52.37N	1.33W
Northolt	03672	51.55N	0.42W
Odiham	03761	51.23N	0.95W
Beaufort Park	03763	51.38N	0.78W
Heathrow	03772	51.48N	0.45W
Gatwick Airport	03776	51.15N	0.18W

The 15th site, Cardington, was mainly run to provide data for comparison with the routinely run surface site at Cardington, which is measuring radiative and turbulent fluxes along with time mean temperatures, winds and humidities, and with 'BALTHUM' ascent data. The results from the surface site require considerable analysis, which will be reported in a separate document. However, some results and insights from both the surface and BALTHUM data will be briefly summarized below.

Version 1 Model Description

Configurations

A number of different configurations of model physics were run for comparison. The base case was denoted 'as per mesoscale' in that the intention was to produce, as far as possible, the same results as the operational mesoscale configuration. This was run in order to demonstrate that differences from the mesoscale model are due to real changes in physics or surface conditions, not simply from numerical differences, the mode of running or coupling the model.

The 'as per mesoscale' configuration used the same physics and resolution as the mesoscale model. Soil and vegetation and orographic roughness parameters are derived from the mesoscale fields by bi-linear interpolation to the site in question. The first two actually vary very little over the mesoscale domain, and the orographic roughness is relatively small and smooth for the sites under study with the possible exceptions of Aberporth, Glasgow and Manchester.

Initially, a timestep of 5 minutes was used, with radiation called every hour. This corresponds to the physics timesteps of the mesoscale model, and, for the vast majority of occasions, ran without any difficulty. However, runaway cooling of fog took place in a small number of cases (leading to decoupled fog layers at about -50 °C). This was caused by the long

radiation timestep, when coupled to an initially high rate of longwave cooling from the fog top. It is likely that this tendency is controlled in the full UM by dynamical feedback or horizontal diffusion, but the problem should be noted. Instead we reverted to running with 100 s timesteps with radiation called every 5 minutes for safety. In general this gives results indistinguishable from running with longer timesteps, but the difference should be born in mind.

It is, in practice, impossible to run exactly as the mesoscale, as we initialize from the mesoscale analysis. In the full model, data assimilation continues for two hours while in our model no data are assimilated. This additional data assimilation does, of course, have impact via the forcing data, but the absence of a small amount of local forcing should be expected to have a (short lived) impact. It is also impossible to run with exactly the same ozone distribution, as the facility to run with a limited number of ozone layers is not available in the Single Column UM (SC)M without modifying the radiation scheme. These differences are small but further emphasize the need to compare the model run 'as per mesoscale' with the mesoscale model itself.

The version 1 SSFM was built as recommended by the stage 1 report. All configurations had 70 levels, with all the resolution enhancement in the bottom 2 km (the boundary layer levels), which had 53 levels. Several components were varied in order to gauge their contribution to performance changes. These were as follows:

MOSES

The new Met Office Surface Exchange Scheme. This implements multi-layer hydrology, with soil moisture freezing, and Penman-Monteith surface exchange with an interactive vegetation model. Soil types have yet to be properly defined for the sites, and 'medium' was chosen for simplicity. Work is currently under way in MP to define soil types for operational use based upon either ITE or National Soil Survey data, but this information is not yet available. Vegetation data were chosen using the ITE 10 km data available currently in NWP (Jones, 1996), subjectively modified, in some cases, to reduce the urban component. More objective definition of surface type will be addressed in stage 3 of the project, but we anticipate that the data used are more representative than the 17 km mesoscale aggregated data.

Cardington was treated differently, as we are interested in direct comparison with the Cardington surface site. A vegetation type of pasture was chosen, though the chosen effective fetch certainly over-emphasizes the effect of the local surface. This site provides no synoptic data, and so is not included in the objective statistics below.

Vegetation canopy

The simple radiative canopy model was described in the stage 1 report (Clark et al, 1996). Sites were run with and without the scheme.

Monin-Obukhov (MO) surface exchange

While the Monin-Obukhov surface exchange scheme is better founded physically, and appears to give better results at Cardington, a configuration was run with the current UM scheme to gauge the difference.

No RMBL

In all cases, except 'as per mesoscale', the rapidly mixing unstable boundary layer formulation was disabled.

No prognostic convection

As noted in the stage 1 report, the convection scheme has an undesirable effect on the unstable boundary layer. Convective cloud, however, is needed for the radiation scheme, and some convective diagnostic output will be useful. A combined boundary layer/convection scheme is the solution to this dilemma, but, in the meantime, we employed the pragmatic solution of running the scheme in a diagnostic mode only, i.e. without applying the increments derived to the model profiles. Since the impact is then only on radiation, the scheme is only called with the same frequency as the radiation scheme, which, of course, saves substantial CPU time.

9 Soil levels

The MOSES scheme is generally run with four soil levels, but we have shown that, with sufficiently small timestep, it is equally capable of running with greater soil resolution. In the stage 1 report we showed that better simulation of soil temperatures could be achieved with nine levels. This requires a timestep longer than that needed by our high resolution boundary layer, and only adds a few percent onto CPU cost.

Apart from the above, and the minor differences in ozone noted, other aspects of the model were 'as per mesoscale'. The same critical RH was used in the cloud scheme (interpolated in the boundary layer). The correct value to use is unclear. It may be appropriate to move the critical RH towards 100% where we have higher vertical resolution, but this would lead to an inconsistency with the mesoscale model which might not be justifiable. While, in a site specific sense, humidity may be less variable, we still need to simulate the effects of broken cloud within the radiation scheme.

Running with all possible combinations of the above would have been very expensive and rather pointless. Instead, a limited number of combinations were chosen. These were identified by a binary code, with the above factors representing a single bit in the code. The combinations actually implemented, together with the decimal value of the code number, are shown in table 1.

Decimal Code	Configuration
000	Standard mesoscale
025	MOSES + No RMBL + No Convection.
029	MOSES + No RMBL + MO surface exchange + No Convection
031	MOSES + No RMBL + MO surface exchange + No Convection + Vegetation
093	MOSES + No RMBL + MO surface exchange + No Convection + 9 soil levels
095	MOSES + No RMBL + MO surface exchange + No Convection + Vegetation + 9 soil levels

Table 1 Description of configurations tested.

Initialization

The initial data for each run have been derived in two different ways. The first is to derive the data directly from the mesoscale analysis. If the same surface orography is used, this poses little problem in the atmosphere. The prognostic variables are simply interpolated (in log pressure) to the required levels. This, of course, means that the initial data have the same vertical resolution as the mesoscale model, which can cause problems, especially when deep fog layers are already present in the analysis. Problems are greater for the soil variables, as, currently, the mesoscale runs with different soil temperature levels and only one soil moisture level. Soil temperature has been dealt with by interpolation assuming an exponential change with depth from one level to the next, the rate of change depending on mesoscale soil characteristics. Below the bottom mesoscale level the temperature was assumed constant. Subsequent analysis of continuous runs (see below) suggested that a linear interpolation between the last but one and last level would have been more appropriate, but this has negligible impact on short runs.

Soil moisture was generated in the multi-level scheme by assuming a constant soil moisture stress (i.e. the difference between the SMC and wilting SMC divided by the difference between the critical SMC and wilting SMC) equal to that in the mesoscale soil layer. It is impossible to derive an accurate conversion between the two schemes, particularly as we expect a different SMC evolution from MOSES, but this method tends to produce the same evapotranspiration as the single level scheme. It should also be born in mind that in MOSES the soil thermal properties take soil moisture into account, so there is a coupling between the two schemes.

In order to retain soil temperature and moisture structure, especially for the nine level scheme, runs also were performed using the T+6 forecast from the previous run as initial data. Thus, soil moisture and temperature are allowed to free run, and the mesoscale analysis only influences evolution via the forcing data. This may be thought of as the first step towards a ‘data assimilation' scheme, running in continuous mode. It is also a potential mode of

operation for sites without observations. The two modes of operation were denoted 'M' and 'S', i.e. initialized from mesoscale or SSFM.

Configuration Identification

Having specified a three digit model configuration, as in table 1, which we shall denote CCC, a number of levels, LL (generally 31, for mesoscale, or 70), the type of initial conditions, I (either 'M' or 'S') and a run hour HH (denoting the data time, 00, 06, 12 or 18) each overall model configuration has been identified with a code of the form HHICCCLL. Thus, for example, the 12Z run of the continuous 9 soil levels with vegetation canopy, MO surface exchange, no rapidly mixing boundary layer and 'diagnostic' convection, with 70 levels is denoted 12S09570, while the 12Z 'as per mesoscale' is 12M00031. A special case is the dataset extracted directly from the mesoscale, which has been denoted HHM00100 (mainly for compatibility with the mesoscale verification dataset).

Forcing

Since the SCM comprises the physics of the full UM, it could be forced using the total dynamical tendencies from the large scale model it is being coupled to. However, a constraint we have placed upon ourselves is that the forcing should be capable of running with routinely available NWP output. It is difficult to extract the total large scale dynamical tendencies from the UM even when runs are set up specially for the purpose. It would require considerable modification of the UM to output these tendencies at all required points at every timestep, and the amount of output would be huge. An approximate approach is therefore required capable of following the large scale forcing but also able to introduce realistic local variation. To do this we have developed a simple forcing system based upon the UM dynamics.

The forcing terms are applied to the cloud-conserved variables liquid water potential temperature (θ_L) and total water (q_t) defined by:

$$\theta_L = \theta - (L_c q_c + (L_c + L_f) q_f) / (c_p \Pi) \quad \textbf{(1)}$$

$$q_t = q + q_c + q_f \quad \textbf{(2)}$$

where Π is the Exner pressure given by:

$$\Pi = (p/p_0)^{R/c_p} \quad \textbf{(3)}$$

q_c and q_f are the liquid and frozen cloud, and L_c and L_f are the latent heats of condensation and of freezing.

Since the SCM and UM use hybrid vertical coordinates (η) we have decided to work in the same coordinate system. The full set of dynamical equations is detailed in UM documentation paper 10 (Cullen, et al, 1993, hereafter referred to as UMDP 10). For simplicity, we shall start from the same system but omit the secondary metric and Coriolis terms. In so doing, we can also use horizontal Cartesian coordinates, rather than polar coordinates. These notational changes have no significant impact on the model. The basic system of dynamical equations are as follows:

$$\frac{\partial \boldsymbol{u}}{\partial t} = -\boldsymbol{u}.\nabla_\eta \boldsymbol{u} - f\boldsymbol{k}\times\boldsymbol{u} - \frac{1}{\rho}\nabla_z \boldsymbol{p} + \boldsymbol{F}_u \quad \textbf{(4)}$$

$$\frac{\partial \theta_L}{\partial t} = -\boldsymbol{u}.\nabla_\eta \theta_L - \frac{1}{\Pi}\left[(L_c q_c + (L_c + L_f) q_f)/(c_p T)\right](\boldsymbol{R}\, T\, \omega/c_p \boldsymbol{p}) + F_{\theta_L} \quad \textbf{(5)}$$

$$\frac{\partial q_t}{\partial t} = -\boldsymbol{u}.\nabla_\eta q_t + F_{q_t} \quad \textbf{(6)}$$

Here, the F terms represent the diabatic terms calculated by the single column model. The last but one term in eq. (5) is a very small correction to allow for the pressure term in the definition of θ_L (as, strictly, it is T_L rather than θ_L that is conserved when cloud forms). The 'vertical velocity' (in pressure terms) ϖ, is given by:

$$\omega = \frac{\partial p}{\partial t} + \boldsymbol{u}.\nabla_\eta p \quad \textbf{(7)}$$

The advection operator is given by:

$$\boldsymbol{u}.\nabla_\eta = u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + \dot{\eta}\frac{\partial}{\partial \eta} \quad \textbf{(8)}$$

The horizontal pressure gradient in eq. (4) is the strict horizontal gradient (i.e. at constant geopotential height) rather than on η surfaces and is derived from the model surface data as described in UMDP 10. The current UM dynamics are hydrostatic, and so the surface pressure tendency and $\dot{\eta}$ are derived from the integral of the horizontal divergence of velocity.

The large scale model has a grid size with length scale L, and we can define an operator $<>_L$ which performs averaging or filtering on this spatial scale. In other words, the u component of wind, for example, in the large scale model is denoted $<u>_L$, and we can define the subgrid deviation, f_s, of any variable f by:

$$f = <f>_L + f_s \quad \textbf{(9)}$$

Using this decomposition we can rewrite equations (4) to (6) as:

$$\begin{aligned}\frac{\partial \boldsymbol{u}}{\partial t} = {} & -\boldsymbol{u}.\nabla^H_\eta <\boldsymbol{u}>_L - \boldsymbol{u}.\nabla^H_\eta \boldsymbol{u}_s - <\dot{\eta}>_L \frac{\partial \boldsymbol{u}}{\partial \eta} - \dot{\eta}_s \frac{\partial \boldsymbol{u}}{\partial \eta} \\ & - f\boldsymbol{k}\times\boldsymbol{u} - \frac{1}{\rho}\nabla_z <\boldsymbol{p}>_L - \frac{1}{\rho}\nabla_z \boldsymbol{p}_s + \boldsymbol{F}_u\end{aligned} \quad \textbf{(10)}$$

$$\begin{aligned}\frac{\partial \theta_L}{\partial t} = {} & -\boldsymbol{u}.\nabla^H_\eta <\theta_L>_L - \boldsymbol{u}.\nabla^H_\eta \theta_{L_s} - <\dot{\eta}>_L \frac{\partial \theta_L}{\partial \eta} - \dot{\eta}_s \frac{\partial \theta_L}{\partial \eta} \\ & - \frac{1}{\Pi}\left[(L_c q_c + (L_c + L_f) q_f)/(c_p T)\right](\boldsymbol{R}\, T\, \omega/c_p \boldsymbol{p}) + F_{\theta_L}\end{aligned} \quad \textbf{(11)}$$

$$\frac{\partial q_t}{\partial t} = -\boldsymbol{u}.\nabla^H_\eta \langle q_t \rangle_L - \boldsymbol{u}.\nabla^H_\eta q_{t_S} - \langle \dot{\eta} \rangle_L \frac{\partial q_t}{\partial \eta} - \dot{\eta}_S \frac{\partial q_t}{\partial \eta} + F_{q_t} \qquad \textbf{(12)}$$

The gradient terms are the 'horizontal' (on η surfaces) gradients, as the vertical term is shown explicitly and treated separately. It should be noted that the decomposition here is not complete, in that the local values available in the 1D model are used where possible. Thus, the full local horizontal velocity, u, is used in the advection term, but the gradient is decomposed into large and small scale. The diabatic terms are derived using local parametrizations, as far as possible.

Since we are, at this stage, assuming the same orography as the large scale model we ignore the subgrid component of the vertical velocity (i.e. $\dot{\eta}_S = 0$). This is a limitation which will be addressed in future versions of the model. For similar reasons, we also ignore the subgrid component of the horizontal pressure gradient. In so doing, we are assuming that the main impact of perturbations to the surface characteristics is to change the vertical turbulent fluxes in a way which is balanced by subgrid horizontal advection. We are ignoring situations where these perturbations have a significant impact on vertical motion. Such situations include small scale sea or lake breeze circulations and local triggering of convection. Trying to include the latter explicitly in the dynamics would clearly be extremely dangerous., while impacts of thermally induced convergence will be regarded as a local phenomenon to be considered in the next stage of the project. The surface pressure (and, hence, that at model levels) is simply taken from the large scale model.

The remaining terms involve the following:
1) Variables directly available in the 1D model.
2) Values of the large scale field or the gradients thereof.
3) The horizontal gradient of the local perturbation.

The second of these is derived directly from the output from the large scale model or by numerical differentiation of it, and represents the primary coupling term. The last is the only remaining class of terms requiring parametrization.

The last term represents local advection, i.e. the advection of the perturbation from the large scale. If we consider a region with different surface characteristics embedded in the larger scale mean, then a simple parametrization of this term is given by:

$$u\frac{\partial f_s}{\partial x} = \alpha |u| \frac{(f - \langle f \rangle_L)}{l_x} \qquad \textbf{(13)}$$

where l_x is the upwind fetch and, for simplicity, we have aligned the x axis with the wind direction. The factor α is an adjustable parameter of order 1. Here, we argue that at the start of the patch with 'local' characteristics, the variable has a value approximately equal to the large scale mean. Parametrizations based on this idea have been tested in 2D and 1D versions of a detailed boundary layer model (BLASIUS) and have been shown to enable the 1D version to give results quite comparable to the 2D model when simulating flow over a surface inhomogeneity (Grant and Best, unpublished). A value of 1 gives acceptable results, though

the best value apparently depends weakly on fetch. Given that, in practice, defining a precise fetch will be difficult (and of only minor importance) we have absorbed α into the fetch specification.

This term, of course, represents a Newtonian relaxation onto the large scale flow. As well as representing the effect of local advection, it ensures that the solution derived from the large scale forcing terms does not stray too far from the large scale model. In practice, we expect our higher resolution simulation to differ from the grid box mean for reasons other than the different local surface. These differences arise from differences in numerical precision, the impact of minor terms in the dynamics which we have neglected, the impact of horizontal diffusion in the NWP solution which we are not using, and different numerical treatment of the advection. These are all (with the possible exception of diffusion) factors which we would wish to minimise if we run the model with a configuration identical to a corresponding mesoscale grid square. If we run with a different configuration, however, then we expect genuine differences arising from, for example, different radiative flux divergence and different turbulent mixing (especially across inversions) because of different resolution. Below the 'diffusion height' in the boundary layer, we expect to see differences arising from our different surface.

A degree of pragmatic subjective judgement is required when dealing with the relaxation or subgrid advection term. If we treat the large scale forcing data as an exact representation of the large scale mean then we should only expect to see differences below the diffusion height. However, since we also hope to benefit from improved resolution and physics, we need to allow the model to develop differences at least throughout the boundary layer. To account for both model error and genuine physical differences, we have used different effective fetches at different levels. At very high altitude (i.e. stratosphere) the effective fetch is chosen to be short enough to ensure that the single column is essentially identical to the large scale model. (To all intents and purposes, we are simply replacing the SCM profile with the mesoscale). Similarly, in most of the troposphere, a short relaxation timescale of 5 minutes is chosen. In the boundary layer, we have the choice of choosing a relatively long fetch, representing the real fetch over the surface of interest, or turning off the relaxation term altogether. In the latter case, since we are forcing only with the gradients of the large scale fields, provided the errors in the advection terms are small, we can derive the maximum impact of both improved model physics and of initial data.

The final model equations are as follows:

$$\frac{\partial \boldsymbol{u}}{\partial t} = -\boldsymbol{u}.\nabla_\eta \langle \boldsymbol{u} \rangle_L - k_r(|\boldsymbol{u}|,\eta)\,(\boldsymbol{u} - \langle \boldsymbol{u} \rangle_L) - \langle \dot{\eta} \rangle_L \frac{\partial \boldsymbol{u}}{\partial \eta} - f\boldsymbol{k} \times (\boldsymbol{u} - \langle \boldsymbol{u}_G \rangle_L) + \boldsymbol{F}_u \qquad \textbf{(14)}$$

$$\frac{\partial \theta_L}{\partial t} = -\boldsymbol{u}.\nabla_\eta \langle \theta_L \rangle_L - k_r(|\boldsymbol{u}|,\eta)\,(\theta_L - \langle \theta_L \rangle_L) - \langle \dot{\eta} \rangle_L \frac{\partial \theta_L}{\partial \eta} - \frac{1}{\Pi}\left[(L_c q_c + (L_c + L_f) q_f)/(c_p T)\right](\boldsymbol{R} T \omega / c_p p) + F_{\theta_L} \qquad \textbf{(15)}$$

$$\frac{\partial q_t}{\partial t} = -\boldsymbol{u}.\nabla_\eta \langle q_t \rangle_L - k_r(|\boldsymbol{u}|,\eta)\,(q_t - \langle q_t \rangle_L) - \langle \dot{\eta} \rangle_L \frac{\partial q_t}{\partial \eta} + F_{q_t} \qquad \textbf{(16)}$$

The surface pressure (and, hence, that at model levels) is simply taken from the large scale model. Any local pressure perturbation is ignored. Similarly, the horizontal pressure gradient and large scale vertical velocity are derived directly from the large scale model (in fact, a single adjustment step is run in order to ensure exact comparability).

The large scale horizontal gradient has been derived using a local spatially centred difference scheme. However, the horizontal advective tendency of f can be written

$$\frac{\partial \langle f \rangle_L}{\partial t} = -\boldsymbol{u}.\nabla^H_\eta \langle f \rangle_L = -|\boldsymbol{u}|\frac{\partial \langle f \rangle_L}{\partial s} \qquad \textbf{(17)}$$

where s represents the distance travelled in the streamwise direction. Assuming locally uniform velocity, this can be derived numerically from

$$\frac{\partial \langle f(\boldsymbol{x},t) \rangle_L}{\partial t} = -|\boldsymbol{u}|\left(\frac{\langle f(\boldsymbol{x},t) \rangle_L - \langle f(\boldsymbol{x}-\boldsymbol{u}\Delta t) \rangle_L}{|\boldsymbol{u}|\Delta t} \right) \qquad \textbf{(18)}$$

i.e. the value at future time is simply replaced by the appropriate upstream value. We could elect to treat the large scale advection in this way, if necessary, calculating more accurate Lagrangian trajectories.

To illustrate the relationship between the large scale advection and the relaxation terms, we can derive the combined effect using a simple implicit first order timestep (to ensure stability), to obtain the following simple algorithm:

$$\begin{aligned} f(t+\Delta t) = f(t) &- (\langle f(t) \rangle_L - \langle f(t) \rangle_L^{upwind}) - \frac{|\boldsymbol{u}|\Delta t}{l_x}(f(t+\Delta t) - \langle f(t+\Delta t) \rangle_L) \\ &= \frac{\langle f \rangle_L^{upwind} + \alpha \langle f(t+\Delta t) \rangle_L + (f(t) - \langle f(t) \rangle_L)}{(1+\alpha)} \end{aligned} \qquad \textbf{(19)}$$

where $\alpha = |\boldsymbol{u}|\Delta t / l_x$.

This more clearly illustrates that relaxation and advection are essentially the same thing. When the fetch, l_x, is small compared to the advection distance uΔt, so α is large, we expect f(t) to equal $\langle f(t) \rangle_L$ (i.e. we are forcing the mesoscale result in very strongly). When the fetch is long, we advect in the large scale upwind value but maintain any local differences between the SCM and small scale.

While we could implement all the advection terms in this way, thereby needing to extract the 'upstream' values as well as the 'at site' vales from the mesoscale model, we have elected instead to treat the large scale advection directly in the Eulerian form of equation (8), i.e. using the product of velocity and vector spatial gradient. This may, in the absence of the relaxation term, be less accurate and cause a more rapid divergence of the model from the

large scale behaviour, but it is judged an advantage to include explicitly the local velocity in the large scale advection, as then any decoupling near the surface will directly feed back on the large scale advection term. The spatial gradients have been derived using finite differencing of first, second or fourth order, with much the same result provided the relaxation term is included. Tests of the scheme will be reported in more detail separately.

A limitation of the mesoscale operational output is that data are only available every hour. Forcing data for intermediate times are derived by simple linear interpolation in time. This has been tested against runs forced with data extracted every timestep, and the errors involved found to be very small. Clearly, in fog situations, the advective terms are already very small, and small errors are negligible. Results from higher order advection in time will be reported separately, as this may be necessary if we wish to force the model with more temporally sparse data (e.g. every 3 hours).

Diagnostics

The single column version of the UM does not use the STASH diagnostics system of the UM and, as a result, not all UM diagnostics are available. Some work was required to ensure that necessary diagnostics were actually available. Errors in the screen temperature and RH were corrected, a screen dew point added together with 10 m wind speed and direction derived from the model diagnosed 10 m u and v. These, themselves, were modified to ensure that, where the bottom level was below 10 m, the diagnostic was derived by interpolation between levels rather than using an assumed profile to the surface. Work was also required to implement the UM visibility system. Identical code was used, though it is anticipated that improvements will be made in future. The mesoscale ‘aerosol' value was used, derived by bilinear interpolation in space and linear interpolation in time as part of the forcing data. No attempt has, yet, been made to implement the UM cloud base diagnostics, as this was not, at this stage, thought worthwhile. Given the additional information available, in principle, from the SSFM on low cloud, it is though wise to consider developing a new set of cloud diagnostics later in the project.

A new diagnostic output scheme was implemented, which allows a list of diagnostics and output frequencies to be specified very simply. This makes interfacing to any database system very simple. For the purposes of the trial, this output was interfaced to the mesoscale SAS-based verification system, and mesoscale data from this system used directly for comparison.

Trial Description

Objectives

The primary objectives of the trial were as follows:

1) To demonstrate that the SSFM could be run routinely coupled to the mesoscale model.
2) To demonstrate that, when configured in the same way as the mesoscale model, the SSFM produces essentially the same results, i.e. to demonstrate that any differences in performance arise from genuine changes in physical behaviour.

3) To gauge the impact of changes in resolution and model physics on forecast skill, both objectively and subjectively. Tests would concentrate on screen temperatures and visibility/fog.
4) To build an archive of forcing data for a number of sites, to enable subsequent rerun of the model using alternative configurations, additional parametrization of local effects and impact of local data.

Procedure

The trial was run almost exclusively on the HP fr1400 workstation, with routine transfer of the mesoscale analysis to the HP directly from the C90 under control of a hook job, and transfer of model level output data from COSMOS fields files to the HP initiated from the HP. After initial tests were complete, the transfer of data from COSMOS, extraction of forcing data and running of the model were automated under cron control. The system was designed to enable the running and post-processing of all configurations, including the continuous runs, to be automated, but, in practice, it was found that occasional errors in data transfer made the continuous system vulnerable, and these runs were performed retrospectively. All other runs were fully automated. An operational system, of course, is likely to be implemented on the Cray, using data from the model prior to transfer to COSMOS, and these problems will not apply. Other than this, the rest of the system could, if required, be ported to the Cray with very little modification.

The 00Z run was performed very soon after mesoscale output became available on COSMOS, and forecasts for Cardington emailed directly to them. These were used, to some extent, as genuine forecast guidance by Cardington, as well as being compared directly with their 'routine' BALTHUM ascents. These subjective comparisons have provided useful insight into UM errors. During December, a routine graphical output for all sites from the 00Z run was made available on the Met. Office intranet to help with checking. The 12Z and 18Z runs were actually performed late in the day, in order to minimise use of high priority COSMOS units, as timeliness was not a requirement.

Initial tests were started in mid-September 1996, with daily forecasts sent to Cardington. This highlighted some operational difficulties, primarily with data transfer from COSMOS. The trial was formally started for all sites on 1st October 1996. Continuous runs were started 14th October, and, in order to capture a full 3-month's of data, and to include all of the early January 1997 cold spell, the trial continued to 13th January 1997. Data reported below are primarily for this latter period. As the trial commenced, a problem with operational availability of forcing data quickly became apparent. While the 06Z and 12Z mesoscale runs both continue to 12Z the following day (T+30 and T+24 respectively), the model level output fields files are only complete up to T+18. The reason for this is unknown, and the restriction was unanticipated. It was neither practical nor justifiable to change the operational system for the purposes of this trial, so runs were only performed to T+18. The main implication of this is that assessment of likely improvement to the overnight forecast of minimum temperatures will be based on the 12Z run rather than the 06Z run upon which Open Road is based. This is not a very serious restriction, as the improvement in forecast between the 06Z and 12Z runs is on a scale unrelated to improvements we are making with our physics changes, so

improvements to the 12Z forecast at T+18 from our model should give a very good indication of the improvement to be expected in the 06Z forecast at T+24.

All input and output data were archived automatically to COSMOS UABRF cartridge as HP format tar files. These can be recovered quite quickly should the need arise. However, forcing and initial data have, for the short term at least, been kept on-line to facilitate reruns, and the diagnostic output resides in a single SAS dataset.

Subjective Assessment

General Performance.

It is impossible to study in any depth the output for every day of the trial for every site. However, some overall comments can be made. The first is obvious, but worth emphasizing. The purpose of the SSFM is to add site-specific detail to NWP output. It is therefore subject to NWP errors. Subjective assessment of the SSFM performance confirmed that the performance was tightly coupled to the mesoscale model, and the majority of errors arose directly from mesoscale model errors. Some insight into the latter was gained, partly from comparison with Cardington data and partly from simply being able to look at the model output in a way which we are rarely able to do with the full NWP product. These insights will be discussed separately below.

Definite differences in behaviour were evident between the SSFM and mesoscale model which could be directly associated with the differences in resolution and physics. These are discussed in the next section. However, where fog, visibility and screen temperatures are concerned, it was clear that the primary source of error was the mesoscale cloud forecast. On numerous occasions the mesoscale model forecast too little or too much cloud, and the SSFM responded appropriately. The surface physics in the SSFM (in particular the vegetation canopy) have been defined to respond more realistically than the mesoscale model surface physics, which has the side effect that the model is more sensitive to nocturnal cloud cover in light winds. When the cloud cover is small, i.e. on good radiation nights, the model provided a more realistic estimate of screen temperatures as the vegetation surface was allowed to cool further. However, where the absence of cloud was an error, the lower temperatures show up as greater errors.

SSFM configurations compared with mesoscale.

As suggested above, subjectively, the main difference between the SSFM and mesoscale model arose during forecast radiation nights. Those configurations which included the vegetation canopy responded with significantly sharper frosts. Ground frosts were more frequent and severe, and there was a significant impact on screen temperature. Where the absence of cloud was correctly forecast, this produced a significant improvement (measurable in degrees rather than fractions) in minimum temperature. The tendency to produce more frost caused some difficulties for the vegetation scheme, as it does not treat snow-cover properly (the UM treats frost as lying snow) but such problems were of secondary importance. The full scheme under development for the UM should not suffer in the same way.

As expected, winds were generally a little stronger, as no ‘orographic roughness' was used and no local orography taken into account. This was evident particularly at Cardington, where we used a ‘local' grass surface. However, the differences were not severe and probably had only minor impact on surface exchange. A tendency for the mesoscale model itself to over-forecast the strength of very light winds was very evident.

The SSFM proved much more able to simulate the formation and evolution of very shallow radiation fog. This, of course, is extremely difficult to verify, and tends to exist over very small horizontal scales. However, one very successful forecast, which could be subjectively verified by ourselves, for ‘Beaufort Park' (i.e. representative of the Bracknell area) is illustrated in Figure 1. On the night in question, shallow fog (depth around 10 m) was observed over grassland with isolated patches on roads. The fog lifted and cleared between 08Z and 09Z. The benefits of improved vertical resolution are evident. A small difference in the formation, thickness and clearance over grass and non-grass covered surfaces is evident, as is a realistic lifting to stratus. The mesoscale configuration actually makes a surprisingly good attempt, with thick fog in the bottom, 20 m thick, layer, but simply cannot resolve the shallowness of the fog, allowing a small amount of fog to be formed in the second model layer. In spite of this, it clears the fog earlier, essentially as soon as the sun comes up.

Fog performance as a whole through the period was severely contaminated by the behaviour of the UM cloud microphysics below freezing. Some of this was predicted to be a problem in the stage 1 report, but the unusually persistent conditions through December and early January had an extreme impact. These problems will be discussed further below. (The mesoscale model itself had particularly poor subjective verification figures from NMC during December.) There was little difference in behaviour of the SSFM compared with the mesoscale in these conditions. However, it did become evident that where light snow is forecast, with boundary layer temperatures close to zero, the SSFM tends to forecast rather less snow, turning it into rain. This is probably a consequence of the higher boundary layer resolution. It is difficult to verify, but, given the mesoscale model's tendency to under-forecast snow in the first place it is probably not a desirable tendency.

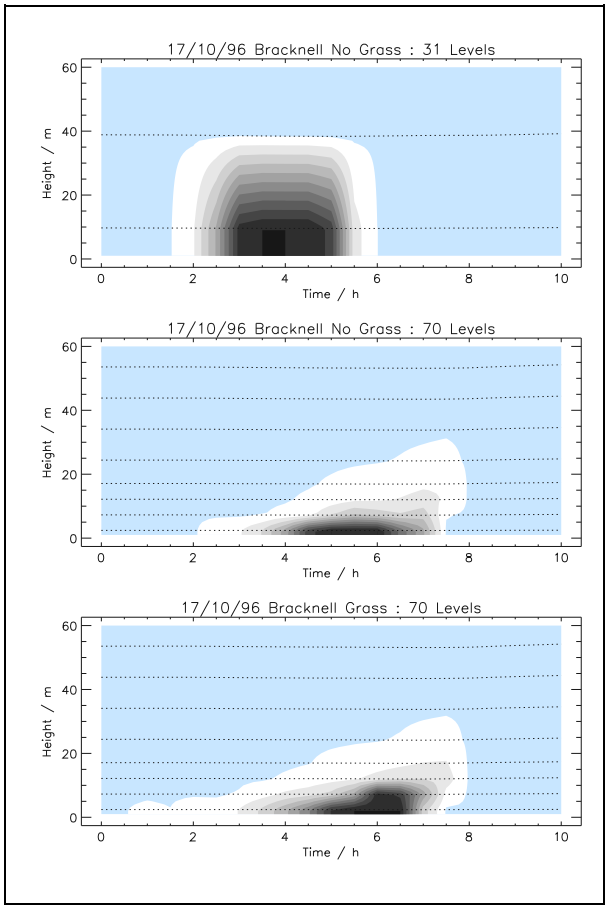

Figure 1 Forecasts of cloud amount (White=>1%, Black=full cover) forecast for Beaufort Park from 00Z 17/10/96. Top figure is ’as per mesoscale’, middle is 70 levels, without a vegetation canopy, bottom is the same with the canopy.

The SSFM showed a distinct tendency to forecast more boundary layer convective cloud than the mesoscale model. As noted above, the scheme has been run diagnostically, and the coupling to the mesoscale relied upon to remove convective instability, so the behaviour of the scheme should be treated with some caution. However, on those occasions studied in more depth, the increase in diagnosed convective cloud at the top of the boundary layer appeared to arise mainly from small increases in the near-surface temperature arising from shutting off the rapidly mixing scheme (which tends to stabilize the surface) coupled with a more sharply defined boundary layer top inversion with higher vertical resolution. Deep fog layers in which the inversion had lifted from the surface were also often diagnosed as unstable over a shallow layer beneath the inversion.

These observations are interesting and, provided the increased convective cloud does not have a negative impact on diagnosed temperatures, suggest a useful by-product of higher resolution. They are clearly no substitute for a better designed boundary layer/convection scheme.

Systematic UM Errors

Microphysics and Ice Fallout

The treatment of ice in the UM was designed primarily to treat cirrus and mid-level frontal cloud in coarse resolution models with relatively long timesteps. It has no separate prognostic variable for ice, instead diagnosing an 'ice fraction' in layer cloud solely on the basis of temperature. This ice is assumed to start falling to the next layer with a velocity representative of quite large crystals (i.e. snow). In the next layer, it may melt, thereby cooling the air and contributing to the rain and cloud, or evaporate. The net effect tends to be a lowering of cloud base. Of course, this is not unrealistic behaviour for glaciated cloud, and the scheme works reasonably well in most circumstances. However, the assumed temperature dependency and relatively fast fallout can produce severe problems in the boundary layer. Though not fixed, the fall velocity is of order 1 ms^{-1}. When compared with a boundary layer depth of a few hundred metres, it is clear that any ice diagnosed in the boundary layer can fall through it very quickly (i.e a few tens of minutes). Amongst other things, persistent freezing fog is an impossibility.

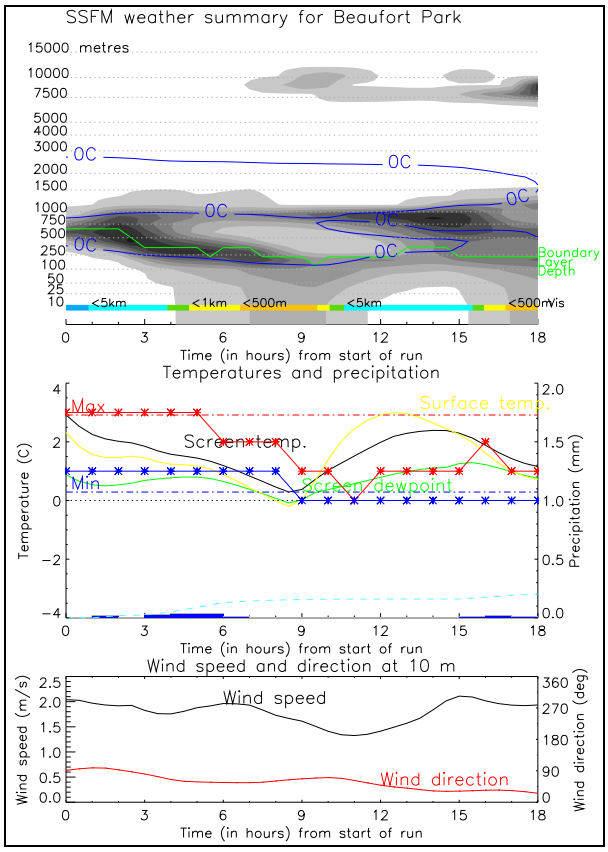

Figure 2 Summary of mesoscale forecast for Beaufort Park, 00Z 12/12/96. Observed temperature (red) and dewpoint (blue) are marked on the middle graph with asterisks.

This behaviour was noted in the stage 1 report, and will not be considered in depth here. Cases of persistent freezing fog are relatively rare, and the model still gives at least some indication of thin fog as some liquid cloud remains, so the impact on verification figures is not very dramatic. However, a reverse problem with the same underlying cause became apparent during December, as a persistent Scandinavian blocking high caused a long period of broadly easterly winds over England, usually with a strong, dry subsidence inversion capping a boundary layer capped with 8/8 stratocumulus. Since air was approaching England from an area with good observational coverage (albeit over a short sea passage), the situation was extremely stationary and, given the time of year, even short wave radiative processes were relatively unimportant, so good forecasts should have been expected. In the event, the mesoscale and SSFM models performed much worse than persistence. The typical behaviour is shown in Figure 2. Starting from midnight, cloud base gradually drops, so that by dawn mist or fog is diagnosed at the surface. This may or may not lift a little during the day. Some drizzle may accompany the cloud base dropping. As the cloud base drops, screen temperature drops and dew point increases. In reality, the cloud cover remained much the same throughout the day, as did screen temperatures and visibility (typically a few km), while any observed drizzle was generally much lighter than forecast.

The cause of this behaviour appears to be related to the fact that the model temperature is below zero beneath the inversion within the stratocumulus deck. This inevitably leads to a downward flux of ice. As the ice falls into sub-saturated air it can evaporate and, eventually (since near surface temperature was above zero most of the time) melt. Both these processes cool the air below and moisten it. Furthermore, as the remaining cloud is re-partitioned between water and ice, some latent heat of freezing is released thereby warming the air. The net effect is to stabilize the boundary layer by both cooling the bottom and warming the top. Observations of the vertical profile confirmed the erosion of the inversion, though they also suggest a cooling above the inversion. This is not yet fully explained, but appears to be related to the erroneous presence of cloud above the inversion. Investigations have shown that this cloud was inserted by the MOPS/NIMROD cloud analysis scheme. In part this was because of a minor error in MOPS which has subsequently been corrected. However, there also exists a coupling between the cloud microphysics and MOPS.

Past versions of MOPS diagnosed the height of cloud top from METEOSAT IR imagery by relating the radiance temperature directly to the model background temperature profile. This could lead to major errors for Sc capped by strong inversions, particularly if this inversion was inadequately captured by the model. To correct this problem, low stratocumulus cloud top is now diagnosed using a conceptual model of stratocumulus topped boundary layers (Hand, 1996). Essentially, the boundary layer is assumed to be well mixed, and the model profile is used to diagnose a representative wet bulb potential temperature (θ_w) for the layer. The associated wet adiabat is then followed upwards until the air temperature matches the observed radiance temperature. This is taken to be cloud top. The method relies, therefore, on a good background value of boundary layer θ_w, but does not require a very accurate temperature profile around the inversion.

These are very reasonable requirements, as the bulk boundary layer temperature profile should not be too sensitive to factors such as model resolution, and has made use of screen

observations and radiosonde profiles. Unfortunately, the effect of melting snow is to lower θ_w slightly, as well as cooling the lower boundary layer. It is possible for a significant fraction of the boundary layer to be warmed as a result. Since assimilation of temperature preserves RH, this will tend to moisten the layer a little, but the combined effect of this, observations of dew point and MOPS cloud on the moisture analysis is difficult to pre-judge. However, if the excessively stable boundary layer is warmed, it will have a higher θ_w and so lead to a higher diagnosed cloud top. This cloud, of course, will tend to cool by radiation, and may add to the downward moisture flux, thereby reinforcing the process at the next assimilation cycle.

Stable Boundary Layer Mixing

The development of the stable boundary layer overnight is clearly of great relevance both to nocturnal cooling and fog formation. The layer is far too shallow be adequately observed by sonde ascents. However, the BALTHUM ascents at Cardington have given some useful insights into the model behaviour. Unfortunately, no data are available showing the development of the layer, but morning ascents have repeatedly shown the model's inability to capture adequately any kind of inversion at the boundary layer top. On the whole, the surface inversion tends to be too sharp, with an excessively deep stable layer above. Sensitivity tests (Derbyshire, personal communication) suggest that this may be the result of the shape of the Richardson number dependence of the mixing coefficients. This has long been acknowledged as a potential problem with the UM boundary layer formulation; the current scheme attempts to take account of some of the larger scale subgrid effects that have no relevance to our model. A sufficient amount of data comparing the SSFM with Cardington profiles is now available to make a refinement of the mixing scheme worthwhile, and this will be attempted during the next stage of the project.

Surface Parametrization

Several month's of data from the Cardington surface site are now available against which we can test the surface exchange scheme. This will be written up fully in a separate document, but the following conclusions have emerged:

1) The main conclusions of stage 1 remain valid with the new data, namely that surface temperatures and fluxes can be estimated more accurately if a simple vegetation canopy is added to the surface energy balance and surface exchange coefficients based upon Monin-Obukhov (MO) similarity theory are used. Of these, the vegetation canopy is by far the most important, and the existing UM surface exchange coefficients are almost as accurate as those based on MO similarity. When the results of the full single column model using MO similarity (without the canopy) are compared with the 'standard' scheme, the differences are rarely significant.

2) Significant errors have been found in the treatment of evaporation of canopy moisture. At present, canopy moisture evaporates essentially at full potential rate. This appears to overestimate substantially the latent heat flux. Parametrization of this is rather arbitrary in the current scheme, designed, as it is, to cope with inhomogeneous canopy cover at

climate scale. It appears possible to improve the scheme substantially within the scope of the current code by small changes to the assumed dependency of evaporation rate on canopy water. This effectively corresponds to changing the assumed proportion of the canopy area which is covered by canopy water. For example, significant improvements can be made by simply assuming that canopy water covers only half the available area. This does not seem, unreasonable for grass, where surface water tends to form discrete drops on leaf surfaces. These changes will be investigated in the next stage of the project. Their main impact is likely to be on screen temperatures and humidity in periods following rain.

3) Significant errors have been found in the temperature dependence of evapotranspiration in C3 grass in the MOSES scheme. Essentially, the current scheme produces too much evaporation at low temperatures (i.e below about 5 °C). Again, this is an area where the MOSES formulation was uncertain and can be re-tuned.

Figure 3 shows a comparison of modelled (using the surface only scheme) and measured latent heat flux over a 4 day period at Cardington. The impact of changing the temperature response of stomatal resistance is clear over much of the period. Figure 4 shows the impact on the modelled latent heat flux of incorporating both a revised temperature dependence and

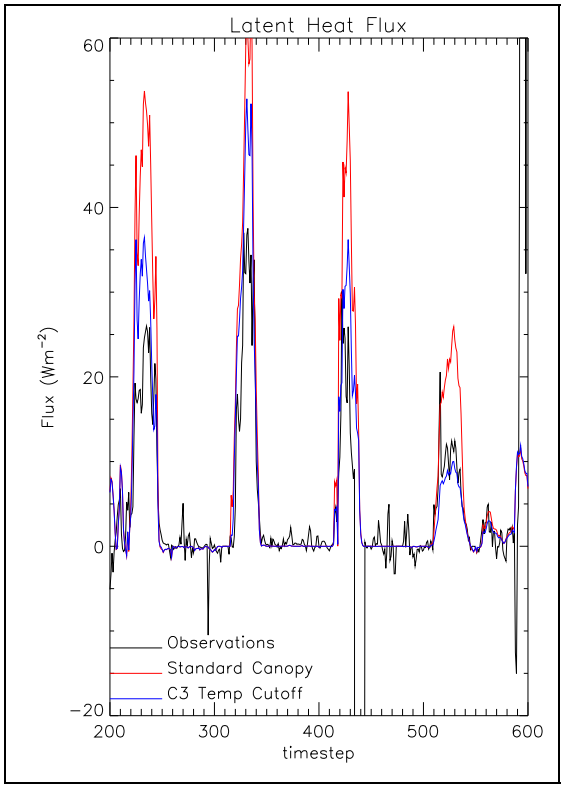

Figure 3 Comparison of 4 consecutive days of latent heat fluxes observed at Cardington with modelled fluxes using the standard and modified temperature response of C3 grass.

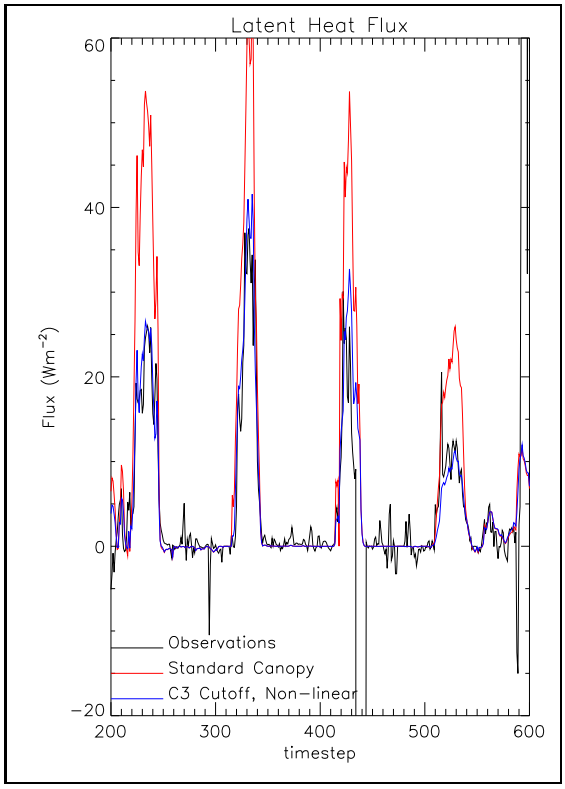

Figure 4 Comparison of 4 consecutive days of latent heat fluxes observed at Cardington with modelled fluxes using the standard and modified temperature response of C3 grass plus non-linear canopy evaporation.

also assuming a non-linear relationship between evaporation from the canopy and available canopy water. The improvements were derived from the period of data illustrated, but they appear to have some generality when applied to other periods (at least for the location).

Soil Temperatures and Soil Moisture

The soil temperature and moisture schemes were allowed to free-run in order to learn about there long-term behaviour. It will ultimately be advantageous if we can run the model with a site-specific soil evolution, either corrected by synoptic observations, or free-running where none are available.

Soil temperatures at Cardington at 0.1 m depth are shown in Figure 5. Here, the initial conditions in the continuous 9 soil level runs (which are T+6 forecasts) have been compared

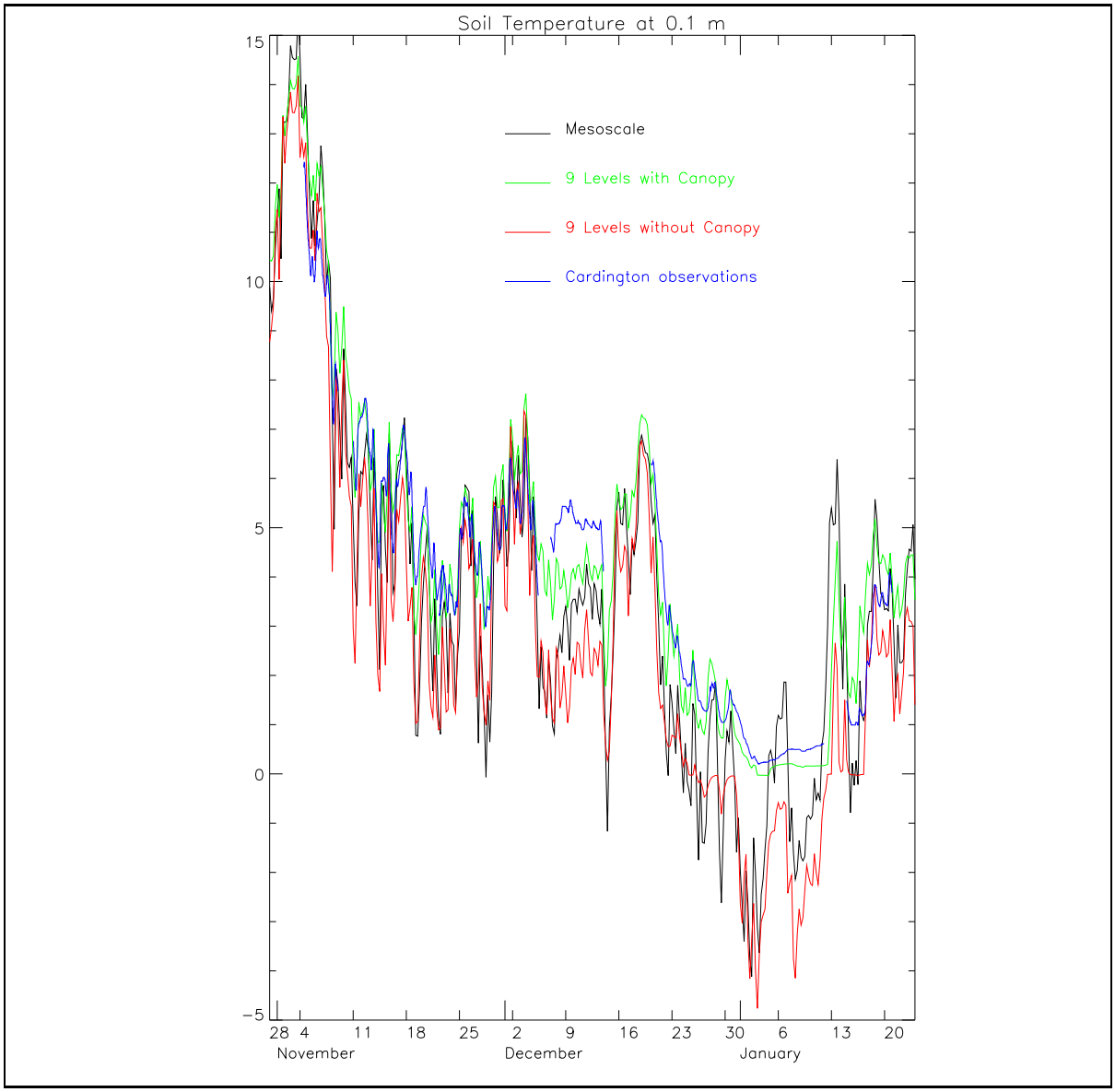

Figure 5 Modelled and measured soil temperatures at Cardington, at a depth of 0.1 m.

with the mesoscale analysis (which includes some impact of screen temperature assimilation on soil temperature) interpolated as described above to 0.1 m, and measurements using a PRT at 0.1 m. Two features are noticeable. First, the nine level runs generally show much less diurnal variation than the mesoscale, which better matches the observations. This is largely a result of a more accurate numerical solution (confirmed by our continuous 4 level MOSES run). Second, the 9 soil level runs are both, generally, warmer than the mesoscale, the run with the canopy more so. With few exceptions, the 9 soil level run with vegetation verifies much better. In particular, radiation nights tend to show a dramatic fall in the 0.1 m soil temperature in both the mesoscale and 9 soil level without vegetation cases, which is not shown by the observations. This implies an energy loss upwards to the air which coincides with the generally warmer air temperatures we find without the canopy. The cold period from mid-December to mid January is particularly noticeable. Soil temperatures (at 0.1 m) in the mesoscale configuration and without the canopy fell well below freezing in spite, in MOSES, of the mitigating impact of soil moisture freezing.

The current single level soil moisture scheme in the mesoscale model is really quite incompatible with MOSES, and so it is hoped that continuous running will produce better results than initialization from the mesoscale. However, bearing in mind that the mesoscale model SMC is MORECS corrected, it is quite possible that our uncorrected continuous runs might drift away from reality more than the mesoscale. In any case, it was anticipated that soil moisture would approach saturation during the trial, and so would be of secondary importance. In the event the autumn was quite dry, so saturation was not achieved at some locations. Available soil moisture evolution for the nine soil level continuous runs at Cardington are shown in Figure 6. The mesoscale soil moisture is shown for comparison, but it must be remembered that the two diagnostics are not comparable. The SMC diagnostic in MOSES represents unfrozen moisture in the root zone available for evapotranspiration. On the whole, the evolution is similar to the mesoscale, with around a factor of three scaling between the two. As no observations are available, it is difficult to comment further. However, the most interesting feature of this figure is the behaviour during the very cold December/January period, where the impact of soil moisture freezing is clear. The vegetation canopy insulates the soil somewhat and so prevents much (though not all) of the freezing.

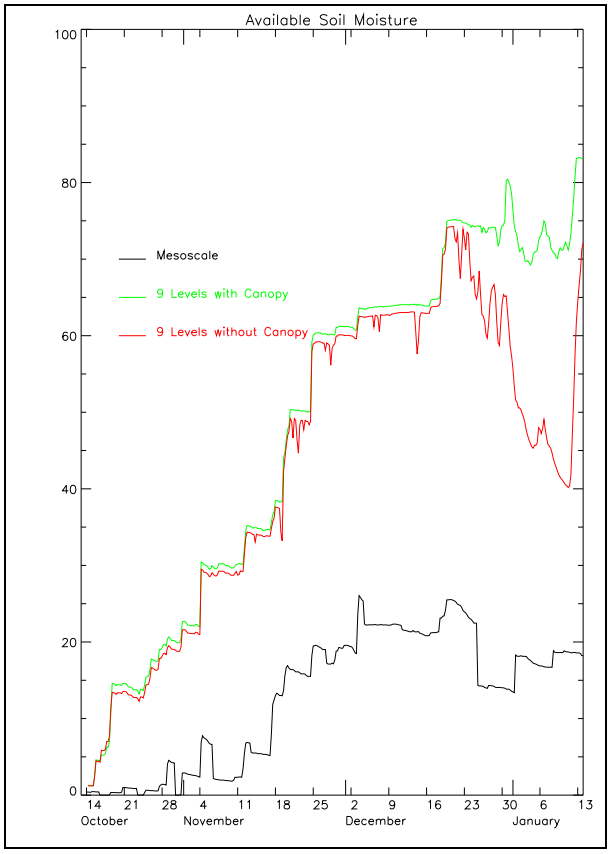

Figure 6 Modelled available soil moisture at Cardington.

Verification against Observations

A statistical analysis of the performance of the Site Specific Forecast Model (SSFM) during a three month trial run from 14th October 1996 to 13th January 1997 is presented. Comparison is made between the operational Mesoscale Model (MES) and the 'as-per-Mesoscale' (APM) configuration of the Single Column Unified Model (SCM) to determine the suitability of this model to replace the MES as the standard in further work, and to assess the accuracy of the coupling between the MES and the SCM. Secondly, the various configurations of the SCM are assessed against the standard and against each other. The choice of configurations allowed the statistical performance of various factors to be considered including:

a) the Monin-Obukhov based surface exchange scheme;
b) the vegetation canopy scheme;
c) interpolation from 4 to 9 soil levels;
d) initialization of the SCM from the MES or 'continuous running' using the previous SCM run for initialization.

The choice of configurations allows the interactions between the different schemes to be investigated, whilst it is still possible to assess the impacts of each scheme individually.

Mean and rms errors, that is to say the mean and rms of the difference between model and observational data in visibility, wind direction and speed, relative humidity and most importantly, screen temperature, were studied, and frequencies of frost occurrence and visibility thresholds are presented.

Temperatures

It should be noted that screen temperature in the MES and APM is assessed as being at 1.5 metres above the underlying surface, whilst in the SCM configurations the screen height has been set at 1.25m, which is more representative of true screen temperature.

Bias and RMS

There is a negative bias (mean error) in all the models, which is much more pronounced in the SCM (excepting the APM configuration). Daily variations in mean and rms errors vary in amplitude strongly between models and run times.

MES against APM configuration

The major results of the comparison between the MES and the APM for all four run times, over the period of the trial when MES data are available, are as follows: a) The largest differences in bias between the two models occur close to T+0, reaching a maximum of ~0.2°C, although the differences are generally much smaller than this. These

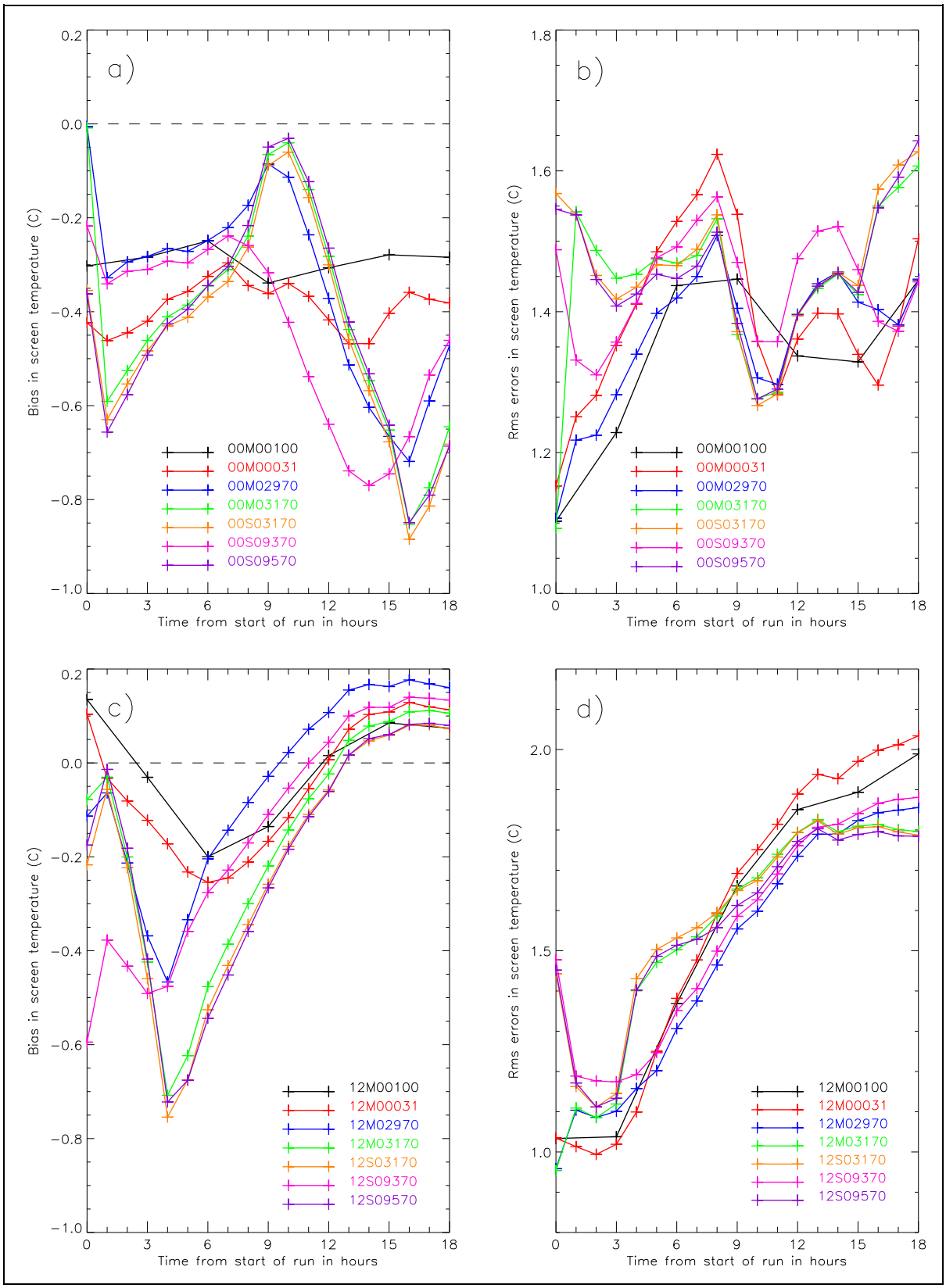

Figure 7 Bias and RMS error in forecast screen temperature over 14 sites from 00Z - a) and b) - and 12Z - c) and d) - forecasts. See page 7 for explanation of key.

relatively large initial differences may be the result of differences in data assimilation.

b) The APM is consistently colder than the MES.

c) The bias is lowest, i.e. closest to zero, at 00Z in all the relevant runs.

d) Rms errors for the two models are very similar, with differences of no more than 0.1°C.

There is a general increase in the rms error with time, with the exception of the period 09-15Z when errors substantially decrease.

SCM configurations together with APM

Results from comparisons between the various SCM configurations, together with the APM configuration of the SCM may be illustrated by Figure 7, which shows the bias and RMS errors over all 14 stations from the 00Z and 12Z runs. The other two runs follow much the same pattern. For an explanation of the key see page 7. The main conclusions may be summarized as follows:

a) Errors, both mean and rms, show a characteristic diurnal pattern irrespective of run time, except in the first 3 hours when initialization still has an impact.

i) The bias is closest to zero at 09Z and most negative at around 16Z.

ii) Rms errors show two distinct maxima at 08Z and 14Z, and a general increase overnight.

b) There is little variation in either bias or rms error between the three configurations containing the vegetation canopy scheme, which show the largest diurnal variations of all the configurations.

i) The vegetative configurations show a negative bias throughout, which reaches a maximum of -0.8 to -0.9°C at 16Z irrespective of data time.

ii) Bias is closest to zero between 09Z and 12Z.

iii) Rms errors are considerably larger than the other configurations between 15Z and 20Z, during the transition from day to night.

c) The other two SCM configurations, the MES-initialised version using the Monin-Obukhov surface exchange scheme (M02970) and the continuous non-vegetative 9 soil levels set-up (S09370), show a less extreme variation in bias and similar rms errors.

i) The maximum negative bias is ~0.3°C closer to zero.

ii) The Monin-Obukhov configuration shows a similar trend in bias to the other configurations, whereas the diurnal cycle in the bias of the non-vegetative 9 soil levels configuration is ~2 hours ahead.

iii) The rms errors of the above two configurations are up to 0.3°C lower during the evening and early part of the night.

iv) Overnight, the rms errors in the SCM configurations are generally lower than those of the APM, and the vegetative configurations are the lowest after ~01Z.

12Z run

Particular interest was paid to the 12Z run, because this provides sufficient time for the SCM to initialise from the MES and to develop the boundary layer through its own physics before the transition from day to night takes place. It also gives a good indication of the performance to be expected from the 06Z run in Open Road applications.

Bias

a) The biases for all the configurations converge after the first hour following initialization (with the exception of the non-vegetative 9 soil levels which converges after ~3 hours) and then increase linearly out to around T+12. After this the rate of increase declines.

b) There is a general increase in the negative bias of all the models and configurations at around T+4, up to a maximum of -0.75°C in the vegetative configurations. This is reduced to -0.5°C in the non-vegetative configurations. This behaviour is present at the same time of day irrespective of data time and corresponds to the evening transition period. This points to errors in the surface exchange and boundary layer parametrizations.

c) Between T+4 and T+16, there is a change in bias towards and then through zero.

i) At T+16, the positive biases in the SCM configurations containing the vegetation scheme are approximately the same as that of the MES, ~0.1°C lower than the other configurations.

ii) The non-continuous Monin-Obukhov configuration (M02970) has the highest positive bias at this time.

Rms errors

a) The rms errors of all the models and configurations show a similar initial convergence, followed by a linear increase with time out to T+12, then a reduced rate of increase.

b) Errors in the configurations containing vegetation are ~0.3°C higher than the APM between T+4 and T+8, whilst the errors in the other two non-vegetative SCM configurations are slightly lower than the APM after T+5. This corresponds to the biases discussed above.

c) After T+8, the errors in the vegetative SCM configurations drop below those of the MES and APM.

d) The rms errors of the two non-vegetative configurations remain the lowest until T+13. By T+18, these configurations are ~0.1-0.15°C lower than the APM, whilst errors in the configurations containing the vegetation scheme are another 0.05-0.1°C lower, an overall improvement of ~0.23°C.

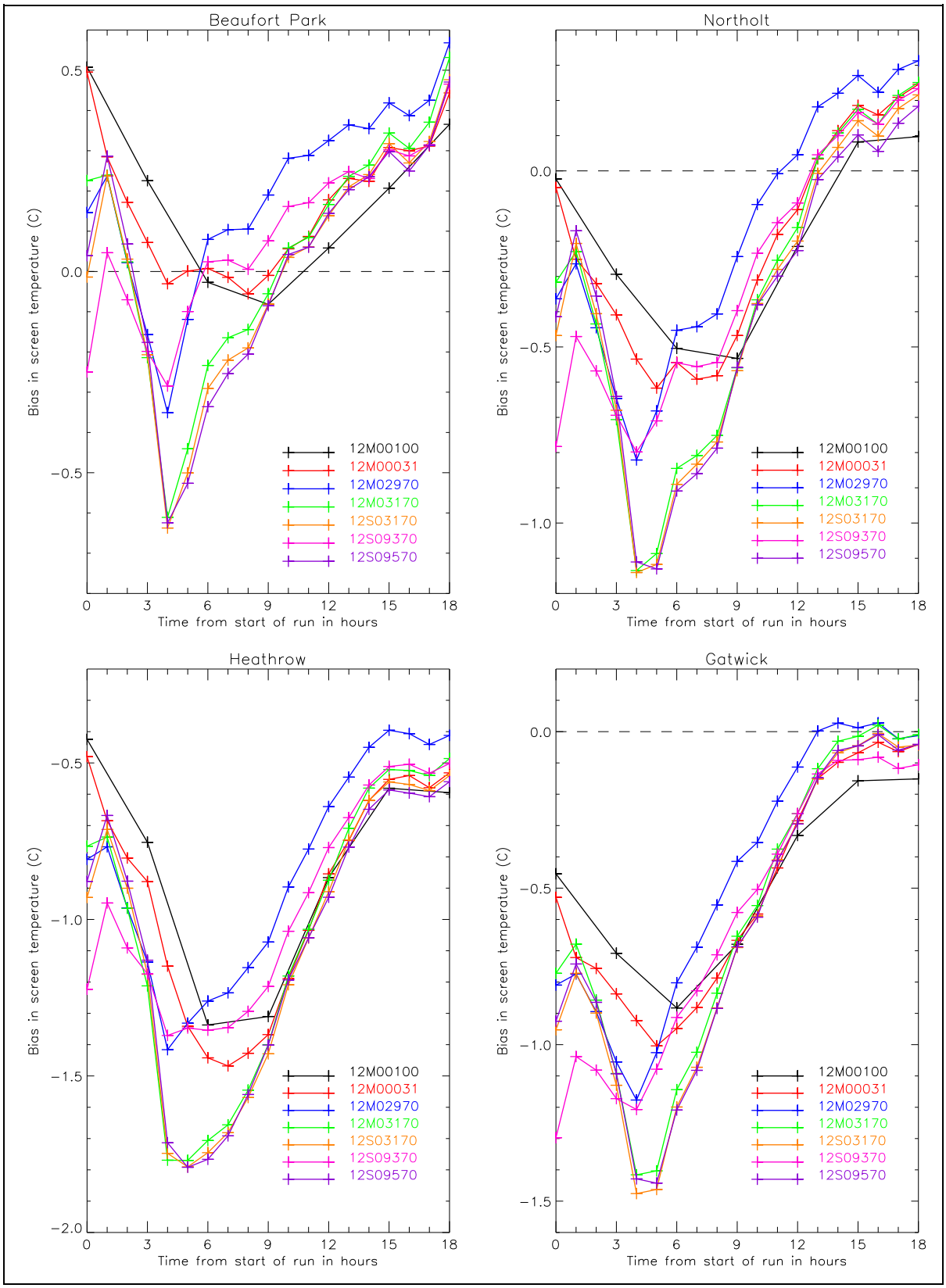

Figure 8 Bias in screen temperature from 12Z forecasts at Beaufort Park, Northolt, Heathrow, and Gatwick. See page 7 for explanation of key.

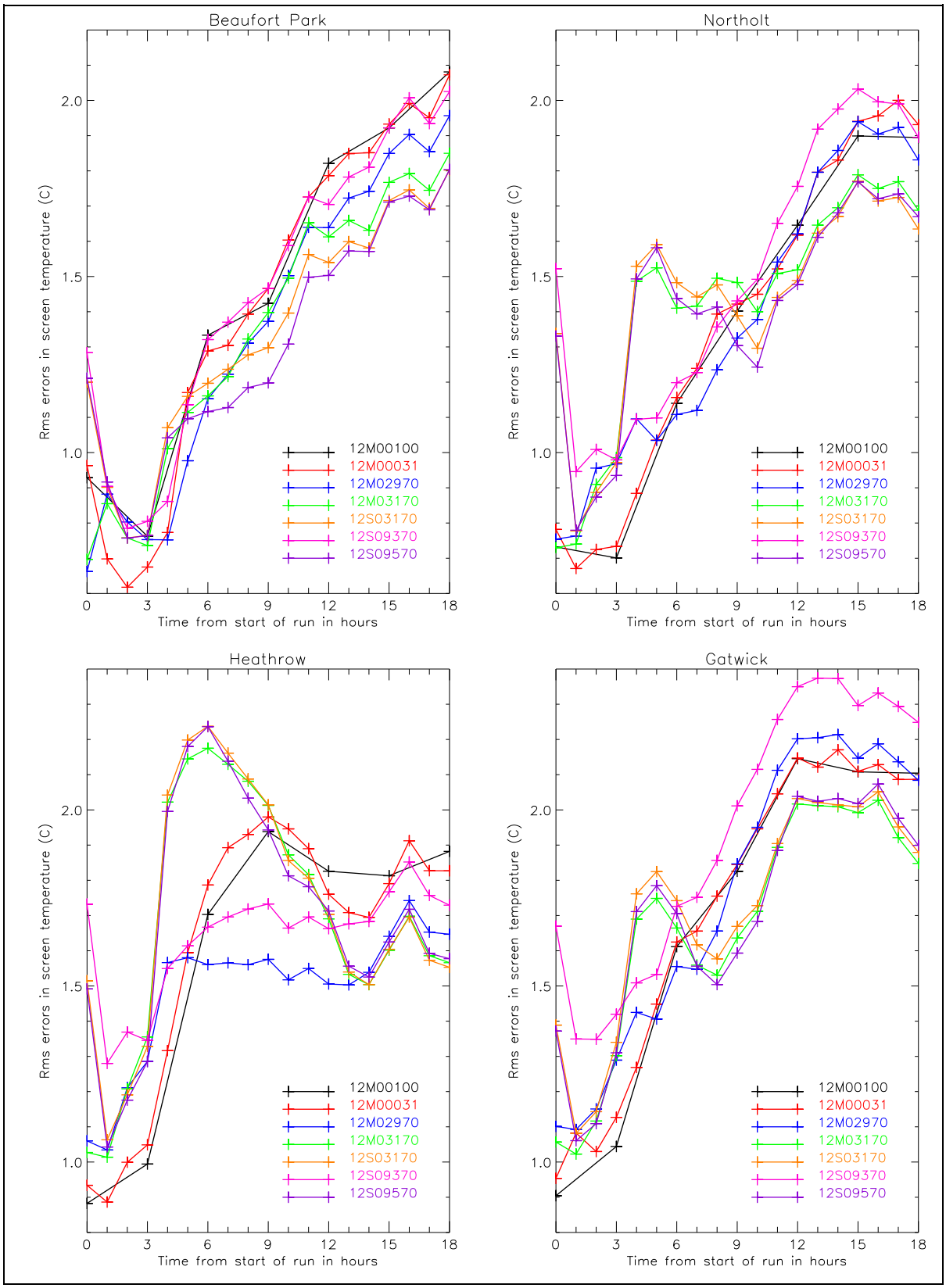

Figure 9 RMS errors in screen temperature from 12Z forecasts at Beaufort Park, Northolt, Heathrow, and Gatwick. See page 7 for explanation of key.

Diurnal Cycle

The diurnal cycle of errors is actually illustrated by Figure 7 as,after the first few hours where the impact of data assimilation is marked, each forecast time was found to show very similar characteristc errors over the diurnal cycle, with a cool bias around 15-17Z (i.e. around dusk) and smallest absolute bias in the late morning. The relative performance of the different configurations is probably of most interest, and this is illustrated by the following table of bias and RMS errors in daily maximum and minimum temperature forecasts. Given the limited forecast period available, these have been derived as follows. The maximum temperature was derived from the 00Z run, taking the daily maximum as the maximum hourly value forecast between 06Z and 18Z, and comparing this with the maximum hourly value observed over the same period. Similarly, the minimum was take as the minimum forecast temperature between 18Z and 06Z from the 12Z forecast, and compared with equivalent observations. The data cover the period 14/10/96 to 13/01/97.

Model configuration	Maximum screen temperature		Minimum screen temperature	
	Bias (°C)	RMS (°C)	Bias (°C)	RMS (°C)
M00031 As per mesoscale	-0.726	1.227	0.158	2.112
M02970 MOSES +MO surface exchange	-0.767	1.312	0.913	2.218
M03170 MOSES + MO surface exchange + canopy	-0.722	1.271	0.270	2.035
S03170 Continuous MOSES + MO surface exchange+canopy	-0.729	1.253	0.159	2.025
S09370 Continuous MOSES + MO surface exchange + 9 soil levels	-0.967	1.373	1.155	2.556
S09570 Continuous MOSES + MO surface exchange + 9 soil levels + canopy	-0.699	1.240	0.145	1.965

These results show that the impact of the different treatments of surface and soil processes (plus atmospheric resolution) is generally small, and that the 'as per mesoscale' configuration performs remarkably well. It is clearly (on average) better for both daily maximum and minimum than the equivalent run with MOSES and Monin-Obukhocv surface exchange. At

the one site at which the standard UM surface exchange was also run (not included in the table), the results were not significantly different, so the presence of this alternative scheme is unlikely to be important. This does not, in itself, prove that MOSES in less accurate than the current UM scheme, not least because we are coupling to a model which uses the current scheme and which assimilates screen observations in a way consistent with the current scheme. It does, however, suggest possible problems, particularly with nocturnal temperatures.

The contrast between this and the same configuration with the canopy scheme added is quite dramatic, again, especially for nighttime minima. The warm bias is substantially improved and both it and the RMS errors are superior to the 'as per mesoscale' (though not by very much). Overall, the continuous, nine soil levels MOSES run with the canopy gives best results at night and very similar results to the mesoscale during the day. Bearing in mind that this has 'free running' soil temperature and so no direct benefit of data assimilation this is extremely encouraging.

Frost Contingency

Contingency table calculation

The contingency tables for (air) frost were calculated by stipulating that any cases when the hourly minimum screen temperature fell to 0°C or below signified that frost had occurred. Only 3 hourly data were available (from the mesoscale verification system) for the MES. The minimum temperature was defined as occurring within a particular time period for each run, to limit the frost occurrence to one nocturnal period:

Run time	Time period for frost occurrence	
00Z	T+0 $\rightarrow$ T+12	00-12Z
12Z	T+6 $\rightarrow$ T+18	18-06Z
18Z	T+0 $\rightarrow$ T+18	18-12Z

The 06Z run was not considered, since it failed to cover a complete nocturnal period. The aim was to use the closest available period to 18-12Z.

Models against observations

a) Overall, the frost hit rate was high as a result of the prevailing synoptic conditions. The beginning of the trial period saw few nights when the temperature even fell close to 0°C and this was quickly followed by a cold spell of weather, during which the temperature remained below freezing for a number of days and the overnight minima were sufficiently low that all models configurations forecast a frost consistently.

b) Comparison of the SCM configurations show that including the vegetation canopy scheme increases the frost hit rate (i.e. the number of times when the model predicted frost correctly) by 2 to 43 (an increase of 3%).

c) The most significant reduction in the hit rate occurred for the non-vegetative 9 soil levels configuration (S09370).

d) Intercomparing the MES and the SCM configurations with observations shows that the three agree in ~88% of cases. Of the remaining cases, the models disagree in only 2-6%.

a)		OBS	
		F	N
MES	F	41	3
	N	6	33

b)		OBS	
		F	N
APM	F	41	3
	N	6	33

c)		OBS	
		F	N
SCM	F	43	3
	N	4	33

d)		OBS	
		F	N
SCM	F	40	3
	N	7	33

e)		OBS	
		F	N
SCM	F	36	2
	N	11	34

Table 2: Air frost contingency tables for the following model configurations against observations: a) Mesoscale model; b) APM; c) the three SCM configurations containing the vegetation canopy scheme; d) the Monin-Obukhov surface exchange; e) and the continuous 9 soil levels.

MES against APM

a) In the 12Z run, there are two cases when the two models do not agree. In both cases, the frost occurrence is marginal, the differences in minimum temperature between the two models being 0.2 and 0.3°C.

b) There are three cases at each of the other two run times (00Z and 18Z) in which the MES and APM disagree. The largest difference between the models is 2.2°C.

SCM configurations together with APM

a) The tendency, in the cases where the APM and the other SCM configurations are not in agreement, is for the APM to predict frost which the SCM configuration does not model.

b) Only 2-4% of cases reside in the above category, except in the case of the non-vegetative 9 soil level configuration(S09370), where 6-7% of cases fit into this category.

c) SCM configurations containing the vegetation scheme most closely follow the APM, again with the exception of the non-continuous Monin-Obukhov configuration (M02970), which disagrees with the APM in up to 4% of cases.

Model configuration	Hit rate (%)	False alarm rate (%)	Skill measure (%)
MES	87.2	6.8	78
APM	87.2	6.8	78
M02970	85.1	7.0	76
M03170	91.5	6.5	82
S03170	91.5	6.5	82
S09370	76.6	5.3	71
S09570	91.5	6.5	82

Table 3: Hit rates, false alarms rates and skill measure (CSI) for the frost occurrences shown in Table 2.

Individual Site Performance

With the exception of Aberporth, performance at individual sites is much the same, qualitatively, as the overall performance above. However, some interesting features arise when each site is studied in the context of its environment. Space does not permit an in-depth description here, but the contrasts are well illustrated by Figure 8 and Figure 9 , which show the bias and RMS errors in screen temperature from 12Z forecasts at Beaufort Park, Northolt, Heathrow, and Gatwick. See page7 for an explanation of key. All show broadly similar behaviour through the night, with the vegetation canopy models performing well, though in the case of Heathrow, the RMS error only recovered in the last 6 hours. Varying degrees of 'evening' bias are visible, however, particularly, though not exclusively, associated with the vegetation canopy. Considering all the sites, they may be classified as showing this weakly, moderately or strongly. The grouping is as follows:

Weak:	Waddington, Cranwell, Coningsby, Wittering, Aberporth, Church Lawford, Beaufort Park
Moderate:	Glasgow Airport, Birmingham Airport (less strong), Northolt, Odiham, Gatwick Airport
Strong:	Manchester Airport, Heathrow Airport

It is not hard to see a pattern of increasing urbanity in this grouping. It seems likely that the effect of the vegetation canopy,

while actually performing well later in the night, is to cause too rapid surface cooling at the more urban sites. At least some of this effect should be properly modelled by using a vegetation model that properly takes the vegetation fraction into account and a 'tiled' approach to the local fetch heat transfer, which will be attempted in the next stage of the project.

Analysis of neighbouring stations.

When discussing 'site-specific' forecast errors, it is useful to divide conceptually an observation (o) into 3 components: the 'mesoscale' (and above) component (m), which we hope to represent accurately with a mesoscale analysis, the 'site-specific' response to this mesoscale forcing (s(m)), which we would normally term a representativity error, and random noise, n, due to measurement or sampling error. Thus:

$$o = m + \mathrm{s}(m) + n \qquad \textbf{(20)}$$

A model attempts to predict this observation, and, in the same way, we can conceptually split the model estimate into 'mesoscale' and the 'site-specific response:

$$\hat{o} = \hat{m} + \hat{\mathrm{s}}(\hat{m}) \qquad \textbf{(21)}$$

where that hats represent model estimates. Clearly, the 'noise' term is absent. In practice, we derive 'mesoscale' estimates for particular sites using some form of interpolation operator on the mesoscale output, but these estimates are still only representative of larger scales, so we are justified in regarding the mesoscale estimate as being defined by the condition $\hat{\mathrm{s}}=0$.

Subtracting the above allows us to write:

$$\text{Forecast Error} = \hat{o}-o = (\hat{m}-m) + (\hat{\mathrm{s}}(\hat{m})-\mathrm{s}(m)) - n \qquad \textbf{(22)}$$

This defines the forecast error at a site labelled i in terms of the mesoscale error, the site-specific error and the measurement noise :

$$(\text{Forecast} - \text{Observed})_i = M_i + S_i - n_i \qquad \textbf{(23)}$$

It also reminds us that even a perfect site-specific model will have site-specific errors if the mesoscale forcing is imperfect, so estimating the 'site-specific' error not straightforward. An attempt has been made at least to establish whether we have made any impact at all on the site-specific error by studying pairs of sites which are close together.

The differences between the SCM errors (for a particular configuration) for two neighbouring sites were compared with the differences in the as-per-mesoscale (APM) errors at these two sites, in an attempt to assess 'local' differences, such as those attributable to the surface forcing. It is assumed that the neighbouring trial sites (labelled 1 and 2) are close enough to each other that the differences in synoptic and mesoscale forcing is minimal both in reality and in the model, so that $M_1=M_2$. Thus:

$$(F-O)_1 - (F-O)_2 = S_1 - S_2 - (n_1-n_2) \qquad \textbf{(24)}$$

The residual term (hopefully) averages out to zero, as the random component of errors should

have the same mean at each site and be uncorrelated. By taking the difference in errors between sites, we are thus deriving the difference between 'site specific' errors in the configurations being studied. If we were completely successful in eliminating the 'site specific' error then $S_1=S_2=0$ so the error difference above will have zero mean and a variance twice the variance of the random noise component. However, it as noted above, a perfect site-specific model will give an erroneous site specific response to errors in forcing, and eventually, as forecast time increases, errors in the mesoscale forecast will still dominate.

The differences between the errors (model minus observation) were calculated for a number of pairs of sites. The SCM configurations run in continuous mode (with MOSES, the Monin-Obukhov surface exchange scheme and 9 soil levels) with and without the vegetation canopy, were compared to the APM configuration of the SCM. The comparisons were carried out only for the 12Z run of the model.

The differences between the APM and the SCM configurations were very small in the comparisons between the three neighbouring Lincolnshire trial sites (Cranwell, Coningsby and Waddington). There were no obvious trends in the comparisons made between Birmingham and Church Lawford. However, the group of sites in the south east of England, Northolt, Heathrow, Odiham and Beaufort Park do show significant differences in the RMS error difference between the APM and SCM. The difference between Northolt and Heathrow is small and probably not significant. However, Northolt-Beaufort Park, Odiham-Beaufort Park and Beaufort Park-Heathrow all show reductions in RMS differences in error over most or all of the forecast.

The comparison between Beaufort Park and Heathrow is illustrated in Figure 10. Between T+6 and T+12, the rms error in the SCM model errors was 0.2-0.3°C lower than that of the APM. Beyond this, the improvement increased to 0.5-0.7°C The vegetative configuration was generally 0.2°C lower than the non-vegetative one after 02Z though higher before 00Z. Bearing in mind that our continuous runs are initialized with previous 6 hour forecasts, it should be noted that the differences are smaller even at short forecast times in the Northolt-Beaufort Park and Beaufort Park-Heathrow comparisons.

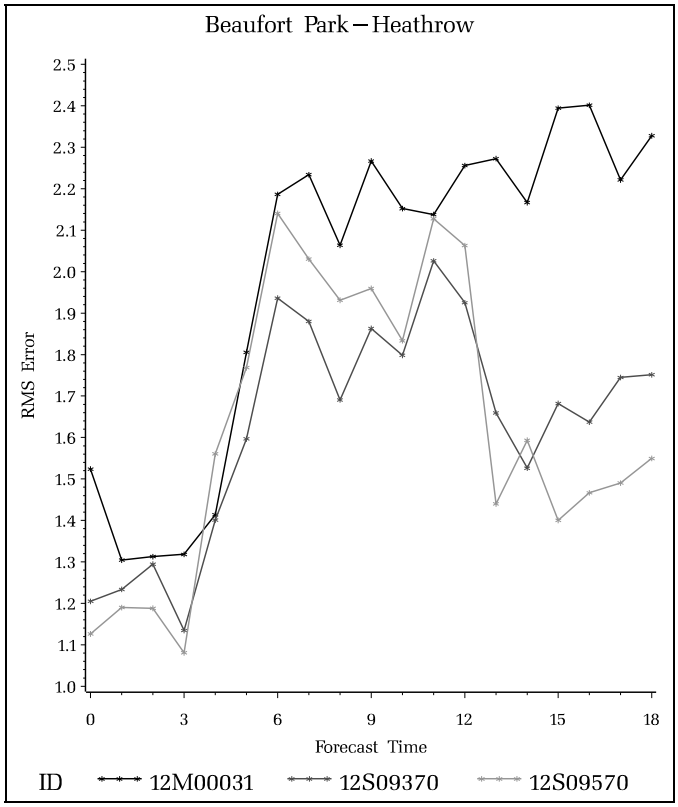

Figure 10: Rms error in the difference between the errors for Beaufort Park and Heathrow for the APM and two configurations of the SCM (for a 12Z run) over the trial period.

This method has clearly shown that we are producing significant 'site-specific' improvements in the area where such improvements might be most dramatic, i.e. where urban and semi rural sites are compared. The method has the limitation that it removes any common bias in local response. As has been seen above, the MOSES simulations do appear to introduce significant biases in afternoon and evening temperatures, but, in spite of this, the site specific model appears to have genuine site-specific skill over the mesoscale model.

Visibility

Bias Factor and RMSF

To account for the skewed distribution of visibility towards higher values, verification is carried out using $\log_{10}$ visibility over the whole range of values, i.e. there is no upper cut-off. One way of expressing the mean and rms errors in $\log_{10}$ visibility is by means of mean factor and rms factor (rmsf):

$$\text{Mean factor} = 10^{\text{Bias in log visibility}}$$

$$\text{Rmsf} = 10^{\text{Rms in log visibility}}$$

These measures give the factor of the error. A bias of 0.3 in log visibility gives a mean factor of 2.0, i.e. the model log visibility is a factor of two larger than the observed value.

The bias is lowest (i.e. most negative) overnight, when it is between -0.35 and -0.5, and peaks at ~-0.2 at 12Z. This corresponds to a mean factor of 0.32-0.45 overnight, and 0.63 at best around 12Z. The rms error increases during the night and reaches a maximum of 0.8-0.85 at 08Z, an rmsf of 6.3-7.1. Since the large rms errors in visibility are apparent in all the models, it appears that the synoptic situation may have been partly responsible. The last month of the trial was characterised by a persistent blocking anticyclone, which maintained an easterly flow and large amounts of low cloud below a strong inversion. This lead to the systematically poor over-prediction of fog discussed above.

MES against APM configuration

The bias in APM visibility is ~0.05 more negative throughout the run (mean factor of 1.12), irrespective of time. The rms error in visibility in the APM configuration is less than 0.1 higher, giving an rmsf of 1.26. This shows that the APM deviates from the MES by no more than 26%, and generally much less.

SCM configurations together with APM

During the day, more precisely between 11Z and 17Z, all the SCM configurations have the most negative biases and the largest rms errors. This ties in with the results from studying RH biases and rms errors in the SCM (see below), since the visibility is calculated using RH and aerosol concentration. Overnight, the SCM configurations lie between the MES and the APM, in terms of both bias and rms errors. There is little variation between the different

configurations, with the exception of the non-vegetative 9 soil level set-up (S09370), which shows a worse performance during the day. The Monin-Obukhov version without the vegetation scheme (M02970) has the smallest bias and rms error of all the configurations.

Fog Contingency

Contingency tables were prepared for each forecast time and data time. In general, as suggested above, the performance of all configurations (including the MES) was similar and poor. Space does not, therefore warrant an in-depth discussion. Results of comparing the APM, SCM configuration and observations for the three visibility ranges (<200m, 200-999m, 1km or greater) may be summed up as follows:

a) Visibility is 1km or greater in 96-97% of observations, so the cases of interest in terms of verifying fog forecasting represent a small percentage of the total number of observations, as would be expected for a set of data taken from this time of year.

b) The number of occurrences of low visibility peaked at T+6, although 3-hourly oscillations between forecast times tended to mask any diurnal variation that may have been present.

i) These 3-hourly oscillations occur because of the 6-hourly overlap between model runs when the data is subdivided purely on the grounds of forecast time.

ii) The lack of a diurnal dependence in visibility may be a result of the synoptic conditions mentioned above in which radiation fog was rare but general low visibility persisted for long periods in anticyclonic conditions.

c) In general, the results show that there are few differences between the APM and SCM configurations. Tables 4 and 5 provide a comparison between two forecast times for the continuous 9 soil levels configuration with the vegetation canopy and Monin-Obukhov surface exchange scheme (S09570), which appears to be the most accurate configuration.

a)		APM			
		<200m	200-999m	≥1km	Total
SCM	<200m	3	0	0	3
	200-999m	3	14	2	19
	≥1km	0	1	13	14
	Total	6	15	15	36

b)		APM			
		<200m	200-999m	≥1km	Total
SCM	<200m	4	4	0	8
	200-999m	1	29	2	32
	≥1km	0	6	17	23
	Total	5	39	19	63

c)		APM			
		<200m	200-999m	≥1km	Total
SCM	<200m	97	77	0	174
	200-999m	5	867	77	949
	≥1km	0	121	3200	3321
	Total	102	1065	3277	4444

Table 4: Visibility threshold contingency tables for the S09570 SCM configuration at FT=15. Comparisons are between the SCM and APM under the constraints: a) visibility observations <200m; b) 200m≥ obs <1km and c) obs ≥1km.

a)		APM			
		<200m	200-999m	≥1km	Total
SCM	<200m	2	2	0	4
	200-999m	1	29	1	31
	≥1km	0	3	18	21
	Total	3	34	19	56

b)		APM			
		<200m	200-999m	≥1km	Total
SCM	<200m	3	2	0	5
	200-999m	0	38	2	40
	≥1km	0	4	18	22
	Total	3	44	20	67

c)		APM			
		<200m	200-999m	≥1km	Total
SCM	<200m	86	92	6	184
	200-999m	10	805	85	900
	≥1km	0	120	3212	3332
	Total	96	1017	3303	4416

Table 5: As table 4, but for FT=18.

Mainly as a result of the synoptic conditions during December which have already been described, the large number of cases for which the model predicted a lower visibility than observed (as shown by the negative bias in visibility) implies a high false alarm rate at both thresholds (Table 6). The large number of underestimates of visibility also produces a high hit rate for visibility less than 1km, but the scarcity of cases of visibility less than 200m (again a feature of the synoptic conditions) accounts for a much lower hit rate.

	Visibility threshold	Hit rate (%)	False alarm rate (%)	Skill measure (%)
FT=15	<200m	8.3	98.4	4.1
	<1km	62.6	94.8	28.8
FT=18	<200m	7.1	97.9	2.7
	<1km	65.0	93.1	32.5

Table 6: Hit rate, false alarm rate and skill measures for the SCM configuration S09570, derived from the visibility thresholds used in Tables 4 and 5.

Other variables

10-metre wind

Speed

Comparison of MES with APM:
a) The biases of the two models agree to within 0.2m/s throughout all the runs.
b) Rms errors show no more than 0.1m/s variations between the MES and APM.
c) Differences between the mean and rms errors of the u and v-components of the wind for the two models are also small.

Comparison of SCM configurations with APM:
a) Averaged over the 18-hour runs, there is little or no overall bias in wind speed in the APM, with a slight positive bias overnight and a slight negative bias during the day. Biases for the SCM configurations are consistently 1m/s greater than the APM.
b) Rms errors for the SCM configurations tend to be around 0.3m/s greater than those for the APM overnight, but comparable during the day. These differences in bias and rms errors between the APM and SCM configurations are almost certainly due to differences in the roughness, since we did not employ orographic roughness in the SSFM configurations.
c) The only noticeable difference between the various configurations of the SCM is that the non-vegetative 9 soil level version (S09370) tends to have slightly lower errors between about 09Z and 15Z.

Direction

Unlike the other variables studied, the main differences in mean and rms errors exist between the MES and all the SCM configurations, including the APM. This suggests that the APM cannot be used for comparisons in place of the MES in this instance.
a) Biases of around 10-15° in the SCM configurations compare with ~1-2° in the MES.
b) Rms errors for the MES are ~60-70°, compared with 80-100° for the SCM configurations.
c) Differences between the SCM configurations and the APM are negligible.

By removing light wind cases, i.e. when the windspeed is less than 5m/s, rms errors in 10m wind direction are reduced by around 20-30° in the MES and ~40° in the SCM configurations. Bias is less positive overall.

Vector wind

Differences between the vector wind rms errors in the MES and APM are no more than 0.05m/s, confirming that the APM can be used in comparisons with the SCM configurations in place of the MES. Differences between the APM and the SCM configurations are very small. Only the non-vegetative 9 soil levels configuration (S09370) diverges from the APM noticeably, but by no more than 0.03,m/s. The rms error increases throughout the night from 15Z onwards, peaking at 08Z. There is a secondary maximum at around 01Z.

Relative humidity

MES against APM

The differences between the two models are small enough that the two can be considered as interchangeable in comparison with differences between the MES and the other SCM configurations. The maximum difference between the bias in the APM and the MES is ~0.5%. Differences in rms errors between the two models are less than 5% of the total rms error.

SCM configurations against APM

a) Continuous and MES-initialised configurations converge within the first hour of the run and are then reasonably similar, with the exception of the non-vegetative 9 soil levels version (S09370), in which the bias is shifted backward by around 1 hour.

b) Bias in 1.5-metre RH peaks at around +6-8% at 15Z in all the SCM configurations, with the magnitude reduced as the time of maximum bias approaches the run time. The APM peaks at around +4-6%.

i) In the 00Z and 06Z runs, the SCM configurations are significantly larger than the APM between 10Z and 18Z.

ii) Beyond 18Z, the SCM configurations have ~0.5-1% lower biases than the APM until 10Z the following day. The non-vegetative configurations are another 0.5% lower.

c) The patterns in rms errors are similar to the mean errors, with a peak of around 11-12% in the APM between 10Z and 18Z, during which time the SCM configurations generally have an rms error 2% larger. In the intervening period, the SCM configurations show ~0.5% improvements over the APM, with the two non-vegetative configurations (M02970 and S09370) showing the greatest improvement.

Conclusions from objective statistics

1) The MES and APM are generally in close enough agreement that the APM can be used in place of the MES in comparisons with the other SCM configurations. Some of the slight differences between the two models result from the differences in data assimilation. In the MES, data are forced in over several hours either side of the run time (when the data are forced in most strongly). However, the APM is relaxed towards MES data which is extracted at run time, so that the APM misses the impact of data assimilation after the start of the run.

2) There was a consistent signal throughout the results that the continuously initialised configuration of the SCM which contained 9 soil levels but excluded the vegetation scheme deviated from the rest of the SCM configurations. In particular, the frost frequencies were significantly lower than for any of the other models or configurations. The explanation for this is given in the section of the report on soil temperatures.

3) Large errors in mean and rms errors occurred initially due to the differences between MES-initialised and continuous configurations. These were, however, short lived (1-3 hours) as the mesoscale forcing took effect.

4) The diurnal pattern of errors showed a clear pattern associated with the degree to which sites might be regarded as urban which can be related to the effective thermal inertia of the 'soil' surface.

5) Wind speed and, more especially, wind direction errors were large. Part of the reason for this seems to be the large number of low wind speed cases, in which measurement of wind direction can be unreliable and velocities may not reach the start-up speed of an anemometer whilst still being modelled. A further likely factor in erroneous wind speeds may be the incorrect assessment of roughness length in the SCM. However, even when occurrences of wind speeds below 5m/s were removed from the data, the mean and rms errors were still large, and a discrepancy seemed to exist between the MES and APM, suggesting that further investigation of this may be necessary.

6) Visibility performance was extremely poor, primarily as a result of poor behaviour of UM ice microphysics. The SSFM showed no objective improvement.

7) A combination of the vegetation canopy scheme, perhaps along with a newly interpolated 9 soil level scheme, in a continuous mode of operation is, on the whole, the best configuration for the SCM, on which to build further adjustment and improvements. This configuration shows significant improvement in nighttime minimum temperature forecasts and frost detection.

Overall Conclusions

The objectives of stage 2 of the project have been fully met, in that:

1.) We are now confident that we can run the SSFM routinely coupled to the mesoscale model.

2.) The coupling system is reliable. With the model configured as per the mesoscale model, it produces essentially the same results, differences being largely attributable to numerical differences and the impracticality of including the full data assimilation cycle. With a different surface and resolution, differences in performance are found consistent with our understanding and expectations.

3.) The performance of the model, at this stage, is mixed, but there is a clear indication that, at long forecast times (i.e. beyond T+12) the model is capable of worthwhile improvements to nocturnal temperature forecasts so long as the radiative vegetation canopy is incorporated into MOSES. A reduction of around 0.1-0.2˚C overnight on an RMS error of about 2˚C over the 14 sites was achieved by the 12Z forecast. This was through a reduction in standard deviation rather than bias, suggesting a genuine improvement in forecast skill. Similarly, a small improvement in skill forecasting overnight frost was achieved.

4.) The system run with the standard MOSES system has a daytime (primarily afteroon) cold bias which is significantly worsened by the incorporation of the canopy, especially at more urban sites. The effect is most profound immediately after sunset. Dusk and evening temperatures with the canopy are therefore significantly worse than 'standard' MOSES and the mesoscale configuration.

5.) These dusk errors will have had an impact on evening fog formation, and, in general, the SSFM performance for visibility is no better, and may be worse, than the mesoscale performance. However, the trial was unfortunate in that the long period of cold easterlies over the December and January period highlighted a serious and systematic error in UM cloud microphysics, which lead to regular over-forecasting of fog. Objective performance of both the mesoscale and SSFM models was much worse than the norm. Subjectively, the SSFM performed much better for genuine radiation fog, though this was very rare during the trial. Objective assessment should probably be postponed until improvements are available for the UM microphysics.

6.) Comparison with Cardington soil temperature measurements has shown that we agree much better with measured soil temperatures than the mesoscale model, in particular showing much less diurnal variation. This confirms that our modified surface scheme is generating a better surface heat flux. Comparisons with surface flux measurements have lead to improvements to detailed parameters in MOSES which will be further tested and should feed back into the overall performance of MOSES in the UM.

Recommendations

1.) The results of stage 2 warrant continuation of the model development to take further account of site-specific factors.

2.) Various minor improvements to the MOSES scheme have been indicated which warrant further investigation. In particular, our current results suggest that MOSES may not simulate the diurnal cycle of screen temperature as well as the current operational scheme unless the effects of a vegetation canopy are included in the surface energy budget.

3.) With the improvements discussed in 1 and 2 in place, results suggest that the model should provide significant improvement to mesoscale overnight temperature forecasts, of a magnitude similar to the improvements currently provided by human intervention. It should be born in mind that relatively few stations were chosen and these may have been more straightforward to forecast than the average.

4.) Performance for radiation fog is subjectively encouraging, but, at this stage, so dominated by errors in the mesoscale development (particularly cloud) and in the UM cloud microphysics that the results are clearly not acceptable. Work should be undertaken, in collaboration with MP group, to rectify the latter.

5.) The SSFM, and the library of forcing data we are building, appears to be a very valuable tool for assessing the long term behaviour of new parametrizations, especially of the surface and boundary layer schemes, and use of it should be promoted widely.

References

Clark, P.A. ,1996, Project Initiation Document - Site Specific Forecast Model (SSFM), Met Office Internal Document.

Clark, P.A.,Hopwood, W.P., Best, M. J., Dunlop, C. C. and Maisey P E., 1996, FR Tech. Report 203. Assessment of the single column UM for use as a local forecasting tool: suitability and recommended configuration.

Cullen, M.J.P., Davies, T. and Mawson, M.H., 1993, Conservative finite difference schemes for a Unified forecast/climate model. UMDP No. 10

Hand, W., 1993, FR Tech. Report 53. A method of improving the analysis of cloud top height for the New Mesoscale Model.

Jones, C.P., 1996, FR Tech. Report 197. Use of high resolution land use data in the UK mesoscale model