UNIFIED MODEL DOCUMENTATION PAPER NO 70

ANCILLARY FILE DATA SOURCES
by

C P JONES

VERSION 10 DATED 02/05/2008

(Corresponds with version 6.6 of the model)

Met R & D (NWP)
The Met Office
Fitzroy Road
EXETER
Devon
EX1 3PB
United Kingdom

Modification Record		
Document version	**Author**	**Description....................**
2	C P JONES	Rewrite and upgrade to v2.3 of UM.
3	C P JONES	Upgrade to v3.2 of UM. New sections on deriving datasets and rearrangement of existing sections.
4	C P JONES	Updates for C90. Inclusion of details about orography. Multi level hydrology datasets and update of file system to UM4.0.
5	C P JONES	Extensive updates
6	C P JONES	Minor updating for UM 5.3
7	C P JONES	Updating for UM 6.0. Add IGBP vegetation and extra details on orography.
8	C P JONES	6.1 release. Further development of creating land sea masks from IGBP vegetation data, addition of soils dust dataset.
9	C P JONES	6.2 release. Addition of IGBP soils dataset and GSWP2 soil moisture dataset.
10	C P JONES	6.4 release. Extra notes added on Cosby soil properties.

INTRODUCTION

Ancillary files are used to supply data fields from an external source to a run of the Unified Model. The data fields being supplied may already exist in the model in which case the ancillary file data will replace that already in the model or may be a totally new field.

In numerical weather prediction models replacing a field would be used either to update a field from an external analysis or to reset a field to climatology. In climate modeling it may be necessary to regularly update a field from an external climatology. This facility is also used when reconfiguring from one resolution to another and it is desired to use externally generated ancillary files instead of using the interpolation within the reconfiguration step. This is particularly important for certain land fields such as vegetation parameters, soils parameters, soil moisture and snow amount as these must be consistent with each other.

If the data field being supplied from the ancillary file doesn't already exist in the model then it is added. This would be used when changing parametrization schemes.

Once reconfigured into the model some ancillary files will remain fixed for the duration of the model run and others will evolve as the model evolves.

It is normal practice to run *reconfiguration* before a model run to read any ancillary files. However, the model itself can regularly update fields during a run if so desired.

This paper describes the data sources used for the master ancillary file datasets and the methods used to generate datasets on model resolutions. In addition to standard references, URL addresses of where datasets are available on the World Wide Web are given. (Note for external users: The Met Office is unable to provide the master datasets, it is up to individual users to obtain the data themselves and reformat to UM ancillary file format.)

Datasets on standard resolutions are held centrally but it is also possible to generate datasets on any desired resolution using the ancillary file generation facility described in UMDP 73.

The datasets covered by this document are;
land sea mask
orography
vegetation parameters
soil parameters
sea surface temperature

sea ice concentration
soil moisture and snow amount
deep soil temperatures
aerosols
ozone

No definition of terms are given nor is any description of how the fields are used in the model. For this see the appropriate documentation paper of the appropriate parametrization scheme.

INTERPOLATION TECHNIQUE

When generating datasets for a given model resolution from the master datasets some kind of interpolation is invariably involved. If the desired resolution is global then area averaging is used otherwise bi-linear interpolation is used, both these methods are described in UMDP S1.

For the sake of interpolation it is generally assumed that the data within the master datasets lie at the centre of grid boxes. On the derived resolution, the data lies at the top left hand corner of the grid box, i.e. a T grid.

Often, ancillary fields have large areas of missing data, e.g. a sea surface temperature climatology will not have data over land areas unless some kind of extrapolation has been performed. This poses a problem when interpolation from one grid to another as there will be a number of points that are unresolved. For these points a gradually increasing spiral search is performed in an attempt to set a data value. If so desired, the radius of the search area can be limited and then a default value is set for any points that remain unresolved.

DESCRIPTION OF DATA SOURCES

Land Sea Mask
PP code 38 STASH code 30 FS Code 74

Optional extra field
River Runoff Outflow Points
PP code 700 STASH code 93

NB : This data source is now considered to be obsolete. It is recommended that the method involving IGBP data be used, see the later section.

Land sea masks are derived from a fractional land cover dataset. Grid boxes with a land fraction greater than some criteria, normally 50%, are classed as land and the remainder as open water, generally sea but may also include lakes.

The dataset used is the US Navy 10’ (1984) dataset (*http://dss.ucar.edu/datasets/ds754.0/*), supplemented by data obtained from the British Antarctic Survey. A dataset at 5’ resolution is also available but this covers north west Europe only.

Before the land sea mask is used in the model it may be desirable perform hand edits to remove features that may cause noise in the lower boundary layer physics. Guidance and utilities for this are described in UMDP 73.

When a land sea mask is generated, a dataset of fractional land cover is also created. This dataset is not used in the model and is provided for interest or diagnostic purposes only. It is also not altered to take into account any manual updates

The river runoff outflow points field is only applicable to the climate model and has been generated by manually inspecting the orography dataset.

[Note: The land sea mask used in hadgem1 runs has been produced using a different method outside the scope of this document]

Orography and Related Fields

	PP	STASH	FS
Orography mean height	1	33	73
standard deviation, $\bullet_h$ (m)	150	34	186
Orographic gradient x component $\bullet H_x$		5	
Oorgraphic gradient y component $\bullet H_y$		6	
xx gradient of standard deviation, $\bullet_{xx}$	152	35	
xy gradient of standard deviation, $\bullet_{xy}$	153	36	
yy gradient of standard deviation, $\bullet_{yy}$	154	37	
silhouette of orography per unit area, A/S	174	17	
$h/2\sqrt{2}$, where h=peak to trough height ($h=2\sqrt{2}\bullet_h$)	175	18	

The main data source used is the dataset known as GLOBE, Global One-km Base Elevation, see URL *http://www.ngdc.noaa.gov/seg/topo/globe.shtml*.

The GLOBE data are at 30" (~1km) resolution. Interim resolutions of 10’ (to be comparable with the old US Navy data) and 1’ are also available but use of these is discouraged.

Very high resolution data is also available covering most of Northern Europe available at 30” (1km) resolution, although a 1’ (2km) version is also available.

The fields are defined as follows.

orographic mean = $\dfrac{\sum H}{n}$

where H is the orographic height

standard deviation

If the mean gradient has been removed (see below) then

$$\sigma = \sqrt{\frac{a\sum H^2}{(a-1)}}$$

otherwise

$$\sigma = \sqrt{\frac{a(\sum H^2 - \overline{H}^2)}{(a-1)}}$$

where a is the area difference between the target and source grid boxes given by

$$a = \frac{\delta\lambda_T}{\delta\lambda_S} \cdot \frac{\delta\phi_T}{\delta\phi_S}$$

•• and •• and the longitudinal and latitudinal spacing of the grids and subscripts T and S denote target and source grids respectively.

The mean gradient is removed by interpolating the interpolated mean field back to the source grid and subtracting from the original source data. The amended source field is then used to calculate the sub-grid scale fields.

The two gradient fields are defined by

$$\partial H_x = \overline{\left(\frac{\delta h}{\delta x}\right)} \quad \partial H_y = \overline{\left(\frac{\delta h}{\delta y}\right)}$$

The three square gradient fields are defined by

$$\sigma_{xx} = \overline{\left(\frac{\delta h}{\delta x}\right)^2} \quad \sigma_{yy} = \overline{\left(\frac{\delta h}{\delta y}\right)^2} \quad \sigma_{xy} = \overline{\left(\frac{\delta h}{\delta xy}\right)^2}$$

where x and y denote the x and y grid spacing, $•_{xy}$ is the grid box diagonal. These fields represent the anisotrophic nature of the orography (the 'shape') within the grid box.

When calculating orographic fields, various filters need to be applied to both the source data and the data calculated on the target grid. The rationale behind the filters that are applied are described in UMDP 74.

The standard deviation and gradient fields are used within the gravity wave parametrization scheme. It should be apparently obvious that in order to calculate the standard deviation and gradient fields, the target grid should be significantly coarser than the grid of the source data. If this is not the case then it is advisable not to activate the gravity wave scheme when running the model.

The remaining two fields are used in the orographic drag parametrization scheme.

A/S is the silhouette of orography per unit area and is calculated through a cross-section using

$$A/S = \frac{\sum H(\delta h)\delta h}{L}$$

H(x) = 1 for x > 0
     = 0 otherwise

Generally, several cross sections are made and an average calculated to find a grid box mean. The diagonals used are shown in figure 1.

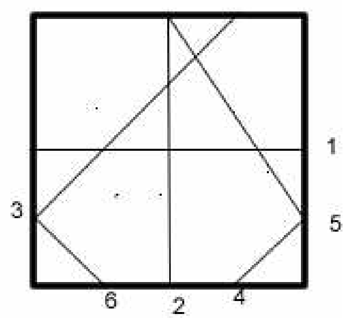

Figure 1: The cross sections used to calculate a value of A/S for each 5′x5′ (see below) grid box. The numbers denote the end of the cross sections.

The actual cross sections chosen are rather arbitrary as long as a good even sample is achieved.

The peak to trough height,h, is parametrized in terms of the standard deviation.

$$h = 2\sqrt{2}\sigma_h$$

For ease of use, these two fields have been pre-calculated using high resolution data on a 5′ x 5′latitude-longitude grid and the fields are then simply interpolated onto the required grid. The A/S field in particular is very sensitive to the resolution of the source data and ideally should be calculated using data no coarser than 3″. However, data at this resolution only exists for a limited area and therefore the field generally used has been calculated using the 30″ GLOBE data and scaled in such a way so that the mean over the area also covered by the 3″ data is conserved. This scaling is applied before the data is interpolated.

It should be noted that the • used to calculate the peak to trough height has been calculated from the high resolution data and is a different • to that used in the gravity wave scheme. Also, for computational convenience the field stored in the ancillary file is actually $\frac{h}{2\sqrt{2}}$.

Sea Surface Temperature (K).
PP code 16 STASH code 24 FS code 91

Sea ice Concentration, f_I.
PP code 37 STASH code 31 FS code 134

Sea ice Fractional Time
PP code 37 STASH code 38

Sea ice 'Equivalent' Thickness, D_I, (m).
PP code 92 STASH code 32

Both the sea surface temperature and sea ice concentration fields have been derived from GISST 2.0 (Global sea-Ice and Sea Surface Temperature) climatology (Parker et al 1995 or *http://www.metoffice.gov.uk/research/hadleycentre/obsdata/GISST.html*). Each has 12 fields valid at the middle of the month. For each sea ice concentration field there is also derived fields of fractional time and sea ice thickness.

The SST climatology was developed using a complete SST background field which was created for each calendar month by averaging the relevant blended satellite and in situ SST fields from 1982 onwards in GISST1.1. In situ SSTs for each month in 1961-90 were then collated with SSTs derived statistically for sea-ice regions using observed sea-ice concentrations, before blending with the background field using a method in which the two-dimensional second derivative of the background field was preserved. The resulting individual monthly SST fields were lightly smoothed before averaging into the final monthly 1 degree dataset.

After interpolation to the required grid, the SST dataset is compared to the corresponding sea-ice concentration dataset. The SST at grid points with a non-zero value for sea-ice concentration is assigned to be 271.35K. At sea points that are not frozen the minimum permissible value for SST is 271.4K.

The sea ice climatology has been created using data obtained from the World Data Center for Glaciology (University of Colorado) supplemented by a Russian sea-ice climatology and a dataset prepared at the University of Illinois by John Walsh. In the WDCG dataset, data for the Arctic covers the period 1972 to 1984 and data for the Antarctic covers the period 1973 to 1984.

The fractional time field is created automatically as part of the interpolation process. It gives the time between one month and the next that the sea ice concentration changes between a non-zero value and zero.

The sea ice thickness field is arbitrarily assigned values of 2m in the Arctic and 1m in the Antarctic. However, a separate dataset exists for the slab model that does have variable sea ice thickness.

Soil Moisture Content in a layer, m, (kgm^{-2}).
PP code 122 STASH code 9 FS Code 191

Snow Amount, S, (kgm^{-2}).
PP code 93 STASH code 23 FS Code 121

Snow fractional time
PP code 93 STASH code 27

There are currently three soil moisture climatologies and two snow climatologies available.

GSWP2: soil moisture only
Willmott et al: soil moisture and snow
AMIP: soil moisture and snow

If GSWP2 soil moisture is chosen then it may be combined with either of the snow climatologies.

It is recommended to use the GSWP2 (Global Soil Wetness Project) (http://www.iges.org/gswp/) soil moisture data. It has been created by running the MOSES-2 scheme as used in the Unified Model off-line with observational forcing data. It is thus a much better representation of the physics of the model than either of the two other datasets.

NB: Ensure that the correct GSWP2 soil moisture climatology is being used according to which soil parameters are being used.

Willmott et al (1985) provide a climatology of the total soil moisture content. This climatology was first scaled to match the vegetation and soil parameters in the model using

$m' = Fm$

where m' is the scaled soil moisture value
m is the original soil moisture value from Willmott et al
and

$F=((\chi_f-\chi_w)D_R\rho)/150$

χ_f is the volumetric soil concentration at field capacity
χ_w is the volumetric soil concentration at wilting point
D_R is the root depth
ρ is the density of water

(NB: these volumetric soil concentrations were for the old single level hydrology scheme).

The soil moisture in a layer, appropriate for the MOSES surface scheme, was then calculated using

$$m_i = \frac{\left(m' \left(\frac{(\chi_c^M - \chi_w^M)}{(\chi_c - \chi_w)} \right) + \rho \chi_w^M D_R \right) \Delta_i}{D_R}$$

m_I is the soil moisture in layer of thickness Δ_i.
χ_c is the volumetric soil moisture concentration at critical point
superscript M denotes values used are for the MOSES surface scheme

The original Willmott et al data was at 1°x1° resolution but the master climatology now used is at N144 resolution.

Willmott et al also provide a climatology of the snow amount and this too has been interpolated to this resolution.

The final climatology has been derived from a 17 year AMIP experiment run at N48 resolution.

For each snow amount field there exists a field of the fractional time that a point changes between a zero and non-zero (or vice-versa) value between one data time and the next.

In calculating soil moisture and snow amount fields, the corresponding soil parameters file is used to ensure that the specification of land ice points is consistent across the datasets. At land ice points, the snow amount is set to 50000kgm^{-2} and the soil moisture is set to 0 kgm^{-2}.

Deep Soil Temperature, T , (K).
PP code 23 STASH code 20 FS Code 190

There are 12 fields for each of the four soil layers at depths 0.1m, 0.25m, 0.65m and 2.0m. The climatology used is that created by a 17 year AMIP run of the climate model.

Soil type dependent fields

There are two sources for soil type classification, Wilson and Henderson-Sellers (1985) (*http://dss.ucar.edu/datasets/ds767.0/*)), hereafter referred to as WHS and IGBP and two soil hydrology schemes, Clapp Hornberger and Van Genuchten. The soil parameters used in the Clapp Hornberger scheme are calculated from fractions of clay/silt/sand using the equation suggested by Cosby et al (1984), they are thus sometimes referred to as Cosby parameters. The Van Genuchten scheme uses different soil parameters and these are found in a lookup table.

It is only possible to calculate Cosby parameters from the WHS data but both Cosby and Van Genuchten parameters may be calculated from the IGBP data.

Cosby Parameters (for Clapp-Hornberger hydrology).

There are a total of ten fields.

Field	**PP**	**STASH**
volumetric soil moisture conc. at wilting point, $\bullet_w$	329	40
volumetric soil moisture conc. at critical point, $\bullet_c$	330	41
volumetric soil moisture conc. at saturation, $\bullet_s$	332	43
Clapp-Hornberger “b” Coefficient, b	1381	207
thermal conductivity of soil, $\bullet_s$ ($Jm^{-1}K^{-1}s^{-1}$)	336	47
saturated hydrological soil conductivity , K_s ($kgm^{-2}s^{-1}$)	333	44
thermal capacity of soil, C_s ($Jm^{-3}K^{-1}$)	335	46
saturated soil water suction (SATHH)	342	48
soil albedo, α	1395	220
soil carbon content, S_c (kgm^{-2})	1397	223

These parameters, except soil carbon, are calculated from fractions of clay/silt/sand in the soil type. If IGBP soils are being used then these fractions are read directly from a lookup table (see section below on Van Genuchten parameters for more information). If WHS soils are being used, then the fractions of clay/silt/sand have to be derived from the soil type given as explained next.

WHS define 22 different soil types according to colour, texture and drainage characteristics, listed in table 1. The drainage characteristics have been ignored. The texture has been used to define the hydrological and thermal properties of the soil and the colour has been used to define the bare soil albedo, used in the calculation of the snow free albedo (see later).

Soil Code	Colour	Texture	Drainage
11	light	coarse	Free
12	light	medium	Free
13	light	fine	Free
14	light	coarse	impeded
15	light	medium	impeded
16	light	fine	impeded
17	medium	coarse	Free
18	medium	medium	Free
19	medium	fine	Free
20	medium	coarse	impeded
21	medium	medium	impeded
22	medium	fine	impeded
23	dark	coarse	Free
24	dark	medium	Free
25	dark	fine	Free
26	dark	coarse	impeded
27	dark	medium	impeded
28	dark	fine	impeded
29	light	-	Poor
30	medium	-	Poor
31	dark	-	Poor
34	ice	**-**	-

Table 1: Soil codes and their properties in the WHS archive.

Using table 1, we first define the percentage of fine, medium and coarse soil for each of the 22 soil classes.

Soil Code	% ice	% fine	% medium	% coarse
11				100
12			100	
13		100		
14				100
15			100	
16		100		
17				100
18			100	
19		100		
20				100
21			100	
22		100		
23				100
24			100	
25		100		
26				100
27			100	
28		100		
29				100
30			100	
31		100		
34	100			

Table 2: Percentage of 4 texture components and their associated soil types.

Each soil texture type has varying fractions of clay, silt and sand, corresponding to varying soil particle size, given by

	Clay	Silt	Sand
Fine	0.52	0.27	0.21
Medium	0.23	0.50	0.27
Coarse	0.05	0.10	0.05

Table 3: Soil particle size fractions, Cosby et al (1984).

WHS provide soil classes on a 1°x1° latitude-longitude grid. For each soil class, the average fraction of each soil particle size is calculated as

$$F_i = \sum_{j=1}^{j=3} \frac{\alpha_{ij} F_j}{100}$$

F_i is the average fraction of soil particle size i in soil type j α_{ij} is the weight as given in table 2.

Note that ice is excluded. Since in each WHS soil class there exists only 1 texture type, there is a direct one to one mapping of WHS soil class to soil particle size fractions.

The soil particle size fractions are then interpolated to the required grid. The interpolated values are then used to calculate the values of the soil parameters using the following equations.

Using the multiple regression relationships of Cosby et al. (1984)**,

$$b = 3.10 + 15.70F_c - 0.3F_s$$
$$SATHH = 0.01e^{(2.17 - 0.63F_c - 1.58F_s)}$$
$$K_s = e^{(-5.55 - 0.64F_c + 1.26F_s)}$$
$$\chi_s = 0.505 - 0.037F_c - 0.142F_s$$

F_c and F_s are the fractions of clay and sand respectively with respect to the total fraction of soil i.e. excluding ice.

** See appendix A.

Calculate χ_w assuming that this corresponds to a suction of –1.5Mpa or an equivalent depth of water of 152.9m.

$$\chi_w = \chi_s \left(\frac{SATHH}{152.9} \right)^{\left(\frac{1}{b}\right)}$$

Similarly, calculate χ_c assuming that this corresponds to a suction of –0.033Mpa or an equivalent depth of water of 3.364m.

$$\chi_c = \chi_s \left(\frac{SATHH}{3.364} \right)^{\left(\frac{1}{b}\right)}$$

The dry soil heat capacity is calculated as

$$C_s = (1 - \chi_s)(F_c c_c + F_s c_s + F_{st} c_{st})$$

where F_{st} is the fraction of silt with respect to the total fraction of soil and c_c, c_s and c_{st} are the heat capacities for air, clay, sand and silt respectively and have the values

$c_s = 2.133 \times 10^6$ $Jm^{-3}K^{-1}$
$c_c = 2.373 \times 10^6$ $Jm^{-3}K^{-1}$
$c_{st} = 2.133 \times 10^6$ $Jm^{-3}K^{-1}$

The values for clay and sand have been chosen to reproduce the dry thermal conductivity and capacity values quoted in table 4.1 of 'The Frozen Earth', Williams and Smith. The value for silt has been set to be the same as for sand.

The thermal conductivity is calculated as

$$\lambda_s = (\lambda_{air})^{\chi^s} \left(\lambda_{clay} \right)^{\left((1-\chi_s) F_c \right)} (\lambda_{sand})^{\left((1-\chi_s) F_s \right)} (\lambda_{silt})^{\left((1-\chi_{st}) F_{st} \right)}$$

where subscripts air, sand, silt and clay denote the thermal conductivity (λ) of air and each of the soil particle sizes respectively and have the values

λ_{air} = 0.025 $Wm^{-1}K^{-1}$
λ_{clay} = 1.16025 $Wm^{-1}K^{-1}$
λ_{sand} = 1.57025 $Wm^{-1}K^{-1}$
λ_{silt} = 1.57025 $Wm^{-1}K^{-1}$

For points with partial or full ice cover the thermal properties are calculated as

$$C_s = F_{soil} C_{s_{soil}} + F_i c_i$$

$$\lambda_s = F_{soil} \lambda_{s_{soil}} + F_i \lambda_i$$

where F_{soil} is the total fraction of soil of heat capacity C_{ssoil} and thermal conductivity λ_{ssoil}. F_I is the fraction of ice and c_I and λ_I are the heat capacity and thermal conductivity of ice respectively and have the values

c_I = 0.63x10^6 $Jm^{-3}K^{-1}$
λ_s = 0.265 $Wm^{-1}K^{-1}$

In the current Unified Model parametrization, the fraction of land ice may only be either 0.0 or 1.0. It thus follows that at land ice points, the thermal quantities are simply set to the values given above. At land ice points, the soil albedo is set to the value for land ice (see below) and all other fields are set to zero.

Calculation of soil albedo

The value of soil albedo is set according to the following table.

	Average soil	Dry soil
Light coloured	0.26	0.35
Medium coloured	0.17	0.25
Dark coloured	0.11	0.15
Ice	0.75	0.75

Table 4: Values of soil albedo according to soil colour and wetness.

WHS also provide a dataset of primary and secondary vegetation type at the same resolution as the soils data and this is used to determine whether the avearge or dry value is used.

If WHS class either the primary or secondary vegetation to be semi arid rough grazing (class 36), or desert (classes 70,71,72,73) then the dry value is used, otherwise the average value is used.

‘Saharan modification’

An option exists to use alternative albedo values for desert regions as the standard values appear to be too low when compared against observational data. If this option is chosen then the following actions are taken.

All points in the area bounded by 5°N, 35°N, 20°W, 60°E are set to a value of 0.4 (or any alternative value specified by the user).

MODIS CLASSIC soil albedo

Houldcroft et al describe a new dataset of bare soil albedo deribed from MODIS data at a resolution of 0.05°x0.05°. Six datasets are available, black-sky and white sky for wavelengths corresponding to visible (0.3•m to 0.7•m), near-infrared (0.7•m to 5.0•m) and total shortwave (0.3•m to 5.0•m) in the spectrum. Black sky radiation is the term given to direct radiation and white sky is direct radiation and reflected radiation.

We use white sky total shortwave.

Each data pixel has an associated error flag, see Houldcroft et al for full details.

1: water
2: no retrieval over land
3: good quality albedo
4: good quality albedo

5: poor quality albedo
6: poor quality albedo

Houldcroft et al recommend that only data flagged as good quality is used. However, we believe that even data marked as poor quality is better than the alternatiev and therefore the default is to use all available CLASSIC data.

The final soil albedo field is a weighted blend with the WHS soil albedo data.

Soil Carbon

Soil carbon data has been derived from Zinke et al (1986), *http://cdiac.esd.ornl.gov/ndps/ndp018.html* or the paper Post et al (1982). Zinke et et al provide data as point values (i.e. observations), and this data has been converted to a regular 0.5°x0.5° grid by Woodward (1995). It is then simply interpolated to the required grid

An alternative soil carbon field is available from the IGBP soil dataset.

Soil type dependent fields (for Van Genuchten hydrology).

(The Van Genuchten hydrology scheme is still being tested and is not yet included as an option in the standard Unified Model).

The ten fields in this case are:

Field	PP	STASH
volumetric soil moisture conc. at wilting point, $\bullet_w$	329	40
volumetric soil moisture conc. at critical point, $\bullet_c$	330	41
volumetric soil moisture conc. at saturation point, $\bullet_s$	332	43
Van Genuchten parameter (1/(n-1))	1381	207*
thermal conductivity of soil, $\bullet_s$ ($Jm^{-1}K^{-1}s^{-1}$)	336	47
saturated hydrological soil conductivity , K_s ($kgm^{-2}s^{-1}$)	333	44
thermal capacity of soil, C_s ($Jm^{-3}K^{-1}$)	335	46
Van Genuchten parameter (1/100•)	342	48*
soil albedo, α	1395	220
soil carbon content, S_c (kgm^{-2})	1397	223

*Currently the same STASH code is being used as fields in the Clapp-Hornberger scheme. These will be given unique STASH codes in a future UM release but it does mean that there will be no checks made in the UM that the soil dataset being supplied matches the hydrology scheme being used. The onus is on the user to be sure. Unusual model behaviour may result if the incorrect dataset is used!

The Van Genuchten parameters n and • are explained below (• here is not albedo).

These parameters are calculated from the IGBP soils dataset (not be confused with the IGBP vegetation dataset described in a later section).

The dataset is obtained on a CD-ROM available for order from the Oak Ridge National Laboratory Distributed Active Archive Centre (*http://www-eosdis.ornl.gov*) .

The documentation supplied with the CD and the references given in this paper describe fully how this dataset was produced and so only a short summary will be given here.

The dataset provides on a regular 5′ by 5′ (approximately 10km by 10km) grid a soil map unit code. The map units are arranged such that numbers 1 to 1972 are reserved for Africa and from 3001 to 6998 for the rest of the world. The soil units are represented by symbol according to the FAO-UNESCO 1974 legend (FAO 1995). For homogeneous soils a map unit will be composed of a single soil unit otherwise there will be a dominant soil type and up to seven other component soils.

In addition to the 106 FAO soil units there are a further 7 miscellaneous units. The list of soil units and their symbol is given below.

Symbol	Soil Type	Comment
Ao	Orthic Acrisols	Acidic soils with a layer of clay accumulation. This class consists only of clays with low cation exchange capacity.
Af	Ferric Acrisols	
Ah	Humic Acrisols	
Ap	Plinthic Acrisols	
Ag	Gleyic Acrisols	
Be	Eutric Cambisols	Soils with slight profile development that is not dark in colour.
Bd	Dystric Cambisols	
Bh	Humic Cambisols	
Bg	Gleyic Cambisols	
Bx	Gelic Cambisols	
Bk	Calcic Cambisols	
Bc	Chromic Cambisols	
Bv	Vertic Cambisols	
Bf	Ferralic Cambisols	
Ch	Haplic Chernozems	Dark soils rich in organic

Ck	Calcic Chernozems	matter.
Cl	Luvic Chernozems	
Cg	Glossic Chernozems	
De	Eutric Podzoluvisols	Soils similar to both Podzols and Luvisols.
Dd	Dystric Podzoluvisols	
Dg	Gleyic Podzoluvisols	
E	Rendzinas	Dark soils rich in organic matter over calcareous material
Fo	Orthic Ferralsols	Highly weathered soils rich in sesquioxide clays and with low catio exchange capacities.
Fx	Xanthic Ferralsols	
Fr	Rhodic Ferralsols	
Fh	Humic Ferralsols	
Fa	Acric Ferralsols	
Fp	Plinthic Ferralsols	
Ge	Eutric Gleysols	Water saturated salts that are not salty.
Gc	Calcaric Gleysols	
Gd	Dystric Gleysols	
Gm	Mollic Gleysols	
Gh	Humic Gleysols	
Gp	Plinthic Gleysols	
Gx	Gelic Gleysols	
Hh	Haplic Phaeozems	Dark soils rich in organic matter.
Hc	Calcaric Phaeozems	
Hl	Luvic Phaeozems	
Hg	Gleyic Phaeozems	
I	Lithosols	Thin soils over rock.
Je	Eutric Fluvisols	Alluvial and floodplain soils with little profile development.
Jc	Calcaric Fluvisols	
Jd	Dystric Fluvisols	
Jt	Thionic Fluvisols	
Kh	Haplic Kastanozems	Dark soils rich in organic matter.
Kk	Calcic Kastanozems	
Kl	Luvic Kastanozems	
Lo	Orthic Luvisols	Soils with strong accumulation of clay in the B-horizon and not dark in colour. These soils have clays with high cation exchange capacity.
Lc	Chromic Luvisols	
Lk	Calcic Luvisols	
Lv	Vertic Luvisols	
Lf	Ferric Luvisols	
La	Albic Luvisols	
Lp	Plinthic Luvisols	
Lg	Gleyic Luvisols	
Mo	Orthic Greyzems	Dark soils rich in organic matter.
Mg	Gleyic Greyzems	
Ne	Eutric Nitosols	Soils with shiny surfaces on structural faces (peds) of the soil.
Nd	Dystric Nitosols	
Nh	Humic Nitosols	
Oe	Eutric Histosols	Soils very rich in organic matter
Od	Dystric Histosols	
Ox	Gelic Histosols	
Po	Orthic Podzols	Soils with a strongly bleached layer and a layer of iron or aluminum cemented organic
Pl	Leptic Podzols	
Pf	Ferric Podzols	

Ph	Humic Podzols	matter.
Pp	Placic Podzols	
Pg	Gleyic Podzols	
Qc	Cambic Arenosols	Sandy soils with little profile development.
Ql	Luvic Arenosols	
Qf	Ferralic Arenosols	
Qa	Albic Arenosols	
Re	Eutric Regosols	Surface layer of rocky material.
Rc	Calcaric Regosols	
Rd	Dystric Regosols	
Rx	Gelic Regosols	
So	Orthic Solonetz	Salty soil with a high concentration of sodium.
Sm	Mollic Solonetz	
Sg	Gleyic Solonetz	
To	Ochric Andosols	Dark soils formed from volcanic materials with little horizon development.
Tm	Mollic Andosols	
Th	Humic Andosols	
Tv	Vitric Andosols	
U	Rankers	Shallow dark soils rich in organic matter and formed from siliceous material.
Vp	Pellic Vertisols	Clayey soils that form deep and wide cracks when dry.
Vc	Chromic Vertisols	
We	Eutric Planosols	Soils with a light coloured layer over a soil layer that restricts water drainage.
Wd	Dystric Planosols	
Wm	Mollic Planosols	
Wh	Humic Planosols	
Ws	Solodic Planosols	
Wx	Gelic Planosols	
Xh	Haplic Xerosols	Aridic soils.
Xk	Calcic Xerosols	
Xy	Gypsic Xerosols	
Xl	Luvic Xerosols	
Yh	Haplic Yermosols	Aridic soils.
Yk	Calcic Yermosols	
Yy	Gypsic Yermosols	
Yl	Luvic Yermosols	
Yt	Takyric Yermosols	
Zo	Orthic Solonchaks	Salty soils with little horizon development.
Zm	Mollic Solonchaks	
Zt	Takyric Solonchaks	
Zg	Gleyic Solonchaks	
RK	Rock debris	
DS	Dune sand	
ST	Salt flat	
WR	Inland waters	
GC	Glacier	
ND	Not determined	

Table 5: List of FAO soil types and general characteristics.

(A useful reference for more information on the soil type classes is the FAO Lecture Notes on the Major Soils of the World available on the FAO website http://www.fao.org/).

It is useful at this stage to refer to the lookup tables that may be found within the ANCIL build at /control/parameters/parameters.

IGBP_SOIL_FRAC gives the percentage of each of the 113 basic types that are present in each of the FAO soil unit codes. For example, code 1 contains 60% Af, 20% Be and 20% I.

Uisng this table, the IGBP soils data is aggregated onto the model grid. If the model grid is rotated then this would be a unrotated temporary grid that covers the domain of the model at the resolution of the model with the proviso that the resolution is not less than twice the resolution of the IGBP dataset.

Soil parameters do not combine linearly for different soils. Therefore, for each model grid box the dominant type is found. The values for each of the Van Genuchten parameters are then read from a lookup table. There are several lookup tables available. (If using IGBP soils to calculate Cosby parameters then average clay/silt/sand fractions may be used)

The default from the disk is IGBP_SOIL_PROPERTIES. This gives various soil properties at the surface and in the sub-soil. We are only interested in the following parameters for the surface.

%clay, %silt and %sand are the relative proportions of clay, silt and sand expressed as a percentage (note they do not necessarily add up to 100%)
•$_r$ is the volumetric water content at the residual point ($m^3.m^{-3}$).
•$_s$ is the volumetric water content at the saturation point ($m^3.m^{-3}$).
• (cm^{-1}) and n (dimensionless) are parameters in the Van Geunchten equation.
K_s is the saturated conductivity (cm/day)

Note. Values for DS and ST are not given on the CD-Rom. Niels Batjes (pers com) provided values for the percentage of clay, silt and sand and the other values have been taken from Qf for DS and Zo for ST as the relative proportions of clay, silt sand are similar. RK, WR, GC and ND are omitted from the data processing. Land ice is set purely using the IGBP vegetation data.

A selection of alternative lookup tables may be found in the subdirectory Rosetta.

Rosetta is a standalone pedotransfer model used to derive soil parameters. Full details may be found at http://www.ars.usda.gov/Services/docs.htm?docid=8953.

The tables called rosetta?.txt (where ? is a number) were obtained by running the Rosetta model for various pedotransfer functions using the clay/silt/sand fractions given in IGBP_SOIL_PROPERTIES. The table called woesten was run for the woesten PFT model. (The Rosetta link above only refers to 5 PFTs, the version we have has several more including the woesten model.)

Alternative clay/silt/sand fractions have been obtained from the WISE dataset (http://www.isric.org/UK/About+Soils/Soil+data/Geographic+data/Global/WISE5by5minutes.htm) . The Rosetta model was rerun using these fractions to produce tables which have _wise in their name.

Using the values extracted from whichever lookup table is being used, the required parameters are calculated according to the following equations.

The Van Genuchten equation is

$$\theta = \theta_r + \frac{(\theta_s - \theta_r)}{\left(1 + (\alpha\psi)^n\right)^m}$$

Where • is the volumetric soil moisture, subscripts s and r are as defined above, • is the matrix water potential (Pa), • and n are Van Genuchten parameters defined in the table above and

$$m = 1 - \frac{1}{n}$$

It is assumed that the residual soil moisture can never be extracted and therefore the UM quantities are calculated thus

$$\chi_s = \theta_s - \theta_r$$

$$\chi_c - \theta_r = \frac{(\theta_s - \theta_r)}{\left(1 + (100\alpha\psi_c)^n\right)^m}$$

$$\chi_w - \theta_r = \frac{(\theta_s - \theta_r)}{\left(1 + (100\alpha\psi_w)^n\right)^m}$$

Where $•_c$ is 33000Pa and $•_w$ is 1500000Pa (and is divided by •g where • is the density of water for insertion into the above equations. The denominator is always positive. (Note that these values are the same as used for calculating the Clapp Hornberger parameters).

The method suggested on the IGBP CD to calculate the two soil thermal fields does not appear to work. Therefore, alternatives have been made available.

The default is to use the fields calculated using WHS data using the Cosby method explained above. However, this means that the thermal fields will be at a much lower resolution than the other parameters and not be consistent with the hydrological parameters and therefore is not recommended.

Peters-Lidard et al (1998) suggest the following equation for soil thermal conductivity

$$\lambda_s = \frac{0.135\gamma + 64.7}{2700 - 0.947\gamma}$$

where • is the soil dry density (kgm^{-3}). This may either be read directly from the lookup table (recommended) or calculated from the porosity.

$$\gamma = (1 - \chi_s) * 2700$$

The soil heat capacity is calculated as

$$C_s = (1 - \chi_s) * 1.942e6$$

However, Lu et al (2007) suggest that Peters-Lidard thermal cnductivity value for soils with high porosity may be too low and suggest a simple linear relationship to porosity.

$$\lambda_s = -0.56\chi s + 0.51$$

The recommended method is the Lu et al method for thermal conductivity and the Peters-Lidard method for heat capacity.

Soil Dust Dataset

This dataset is only produced if IGBP has been chosen as the source of the vegetation data. It is used only by the soil dust parametrisation scheme and may be discarded if not required. Soil heterogeneity is less of an issue and therefore average clay/silt/sand fractions may be used.

The dataset contains 9 fields:

Field	PP	STASH
Dust parent soil clay fraction	1630	418
Dust parent soil silt fraction	1631	419
Dust parent soil sand fraction	1632	420
Dust soil mass fraction division 1	1633	421
Dust soil mass fraction division 2	1633	422
Dust soil mass fraction division 3	1633	423
Dust soil mass fraction division 4	1633	424
Dust soil mass fraction division 5	1633	425
Dust soil mass fraction division 6	1633	426

The clay, silt and sand fractions are as calculated in the course of calculating the physical hydrological and thermal properties of the soil described in the previous section.

The remaining fields are fractions of each of 6 soil size divisions. Further details on the soil dust scheme may be found in Woodward (2001).

Vegetation type dependent fields.

There are three datasets describing the properties of the vegetation cover.

The functional type datasets contains the following fields

Field	**PP**	**STASH**
LAI of functional type broadleaf trees	1392	217
LAI of functional type needleleaf trees	1392	217
LAI of functional type C3 grass	1392	217
LAI of functional type C4 grass	1392	217
LAI of functional type shrub	1392	217
C_h of functional type broadleaf trees	1393	218
C_h of functional type needleleaf trees	1393	218
C_h of functional type C3 grass	1393	218
C_h of functional type C4 grass	1393	218
C_h of functional type shrub	1393	218

pseudo-levels are used to differentiate between the different functional types.

The fractional dataset contains fractions of each of nine surface types.

Field	**PP**	**STASH**
fraction of broadleaf trees	1391	216
fraction of needleleaf trees	1391	216
fraction of C3 grass	1391	216
fraction of C4 grass	1391	216
fraction of shrub	1391	216
fraction of urban	1391	216
fraction of water	1391	216
fraction of soil	1391	216
fraction of ice	1391	216

pseudo-levels are used to differentiate between the different surface types.

The disturbed vegetation dataset contains the fraction of vegetation that is subject to anthropogenic disturbance.

Field	**PP**	**STASH**
fraction of vegetation subject to disturbance	1394	219

Vegetation type is determined using data from the International Geosphere and Biosphere Programme (IGBP) (*http://edcdaac/usgs.gov/glcc/globe_int.html*). The dataset being used is version 2 on the geographical latitude-longitude projection.

The dataset has been derived from AVHRR data covering the period April 1992 to March 1993 and provided at 30 arc-second (~1km) resolution. The data have been classified using various legends and we are using the legend of the IGBP which consists of 17 classes defined in table 9.

Evergreen needleleaf forest
Evergreen broadleaf forest
Deciduous needleleaf forest
Deciduous broadleaf forest
Mixed forest
Closed shrublands
Open shrublands
Woody savannas
Savannas
Grasslands
Permanent wetlands
Croplands
Urban and built-up
Cropland/natural vegetation mosaic
Snow and ice
Barren or sparsely vegetated
Water bodies

Table 11: List of the 17 IGBP land types

It can be seen that the IGBP dataset does not distinguish between inland waters and the open sea. Therefore, we have introduced an additional class of open sea and used the dataset created using the Biosphere Atmosphere Transfer Scheme (BATS) legend which does distinguish inland water from ocean to define these points.

The IGBP data is aggregated onto the model grid (see UMDP 73) for details of methods available) and the fraction of each of the 18 IGBP classes present found.

Not every point on the IGBP grid has been defined a class and thus the final totals are adjusted to remove any areas of missing data. Also, classes that consist of less than 1% of the grid box are eliminated and the area allocated to other classes.

Land ice may only be 0 or 100%. Therefore, grid boxes that have more than a prescribed threshold (normally 50%) of land ice are set to be entirely of land ice. In grid boxes that contained some land ice but below the threshold value, the

land ice is eliminated and the area proportionally added to all the other classes within the grid box.
The fraction totals of the 9 MOSES surface types are then calculated by mapping the IGBP classes to the MOSES surface types using the values given in table 10. Open sea is ignored.

	MOSES surface types								
IGBP class	*Broadleaf*	*Needleleaf*	*C3 Grass*	*C4 Grass*	Shrub	Urban	Water	Bare soil	Ice
Evergreen needleleaf	0.0	70.0	20.0	0.0	0	0.0	0.0	10.0	0.0
Evergreen broadleaf	85.0	0.0	0.0	10.0	0.0	0.0	0.0	5.0	0.0
Deciduous needleleaf	0.0	65.0	25.0	0.0	0.0	0.0	0.0	10.0	0.0
Deciduous broadleaf	60.0	0.0	5.0	10.0	5.0	0.0	0.0	20.0	0.0
Mixed forest	35.0	35.0	20.0	0.0	0.0	0.0	0.0	10.0	0.0
Close shrub	0.0	0.0	25.0	0.0	60.0	0.0	0.0	15.0	0.0
Open shrub	0.0	0.0	5.0	10.0	35.0	0.0	0.0	50.0	0.0
Woody savanna	50.0	0.0	15.0	0.0	25.0	0.0	0.0	10.0	0.0
Savanna	20.0	0.0	0.0	75.0	0.0	0.0	0.0	5.0	0.0
Grassland	0.0	0.0	66.0	15.7	4.9	0.0	0.0	13.5	0.0
Permanent wetland	0.0	0.0	80.0	0.0	0.0	0.0	20.0	0.0	0.0
Cropland	0.0	0.0	75.0	5.0	0.0	0.0	0.0	20.0	0.0
Urban	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0
Cropland/ natural mosaic	5.0	5.0	55.0	15.0	10.0	0.0	0.0	10.0	0.0
Snow and ice	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Barren	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0
Inland water	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0

Table 12: Mapping of IGBP classes to surface types used in MOSES. The surface types in italics are also plant functional types.

A final check is made to ensure that all model grid points have been assigned at least one class, the minimum permitted is 100% bare soil and that the sum of the constituent fractions is 100%.

The leaf area index for each of the plant function types is then calculated as follows.

First, the leaf area index is calculated for each IGBP class.

$$LAI_j = \sum_j f_j \alpha_{ij}$$

where LAI is leaf area index for IGBP class j. • is the fraction of PFT i in IGBP class j as given in table 10 and f_j is the fraction of IGBP class j in the grid box.

Then, the leaf area index for each PFT i is calculated using,

$$LAI_i = \frac{1}{f_i}\sum_j (LAI_j.LAI_{ij})$$

Where f_i is the fraction of PFT i and LAI_{ij} is given by the matrix in table 11. If f_i is zero, then LAI_i is set to the minimum leaf area index value for that plant functional type.

	MOSES Plant Functional Types				
IGBP class	Broadleaf	Needleleaf	C3 grass	C4 grass	Shrub
Evergreen needleleaf	not defined	6.0	2.0	not defined	not defined
Evergreen broadleaf	9.0	not defined	2.0	4.0	not defined
Deciduous needleleaf	not defined	4.0	2.0	not defined	not defined
Deciduous broadleaf	5.0	not defined	2.0	4.0	3.0
Mixed Forest	5.0	6.0	2.0	not defined	3.0
Close shrub	not defined	not defined	2.0	not defined	3.0
Open shrub	5.0	not defined	2.0	4.0	2.0
Woody savanna	9.0	not defined	4.0	not defined	2.0
Savanna	9.0	not defined	not defined	4.0	not defined
Grassland	not defined	not defined	3.0	4.0	3.0
Permanent wetland	9.0	not defined	3.0	not defined	3.0
Cropland	5.0	not defined	5.0	4.03	3.0
Urban	not defined	not defined	not defined	not defined	not defined
Cropland/ natural mosaic	5.0	6.0	4.0	4.0	3.0
Snow and ice	not defined	not defined	not defined	not defined	not defined
Barren	not defined	not defined	not defined	not defined	not defined
Inland water	not defined	not defined	not defined	not defined	not defined

Table 13. Values of leaf area index of each MOSES plant

functional types for each IGBP class. Note that a leaf area index value is not defined for every IGBP class and these are therefore excluded from the summations described above.

The canopy height, C_h, is calculated according to

$$C_H = H_F LAI^{2/3}$$

where H_F is a height factor for each plant functional type given in table 12.

	MOSES Plant Functional types				
	broadleaf	needleleaf	C3 grass	C4 grass	shrub
H_F	6.5	6.5	0.5	0.5	1.0

Table 14. Factors for calculating canopy height from leaf area index values for each PFT.

Seasonal MODIS LAI Values

As an alternative to the fixed LAI values MODIS LAI values for each of the PFTs for each month may be used to create a seasonal varying data.

Land Sea Mask

In addition to calculating vegetation distribution, the IGBP dataset may also be used to calculate a land sea mask. Indeed, it is intended that the IGBP dataset will replace the US Navy fractional land dataset for new configurations as they are introduced.

The IGBP grid boxes are mapped onto the model grid boxes as before but instead of calculating the fractions of individual vegetation types, the total fraction of all non-water types is calculated. It may in some instances be required to count inland water as a land type, as opposed to open sea, and this is possible.

The fractional land field that has been created is then used to define the land sea mask using some threshold, normally 50%. As with using the Navy dataset, it is also possible to perform manual edits to the mask but the fractional land field will not be altered to take into account any changes made.

Ozone
PP code 453 STASH code 60

The climatology used is that created by Shine and Li (1995). This used data from the SBUV (Solar Backscatter UltraViolet instrument) supplemented by data from other satellite instrumentation.

Data is supplied at 2.5°x2.5° resolution and extends from ground to 0.0001mb on 47 levels. The original data was monthly between January 1985 and December 1989 and from this mean values for each calendar month were calculated by averaging across the five years. The original data is in Dobson units but for use in the UM it is converted to mixing ratio. The vertical interpolation is performed, to ξ_h levels, in such a way as to conserve the total ozone in the column.

The vertical distribution is performed on pressure levels. To facilitate this, the model eta_theta values are converted to pressure levels using the ICAO standard atmosphere as detailed in The Meteorological Glossary (1991).

In practice, a zonal mean on all levels is used for global grids. For mesoscale models, data on full fields is used but often only for the levels below which the ozone concentration is fairly uniform.

Atmospheric Aerosols

Total Aerosol Concentration PP ocde 286 STASH code 90
Total Aerosol Emissions   PP code 287 STAHS code 57

These fields are only used for the UK and Balkans mesoscale models and data from a variety of sources have been used. There are a mixture of sources for atmospheric aerosols, low level and high level and sulphur and non-sulphur and all these need to be considered.

For the UK, the Warren Spring Laboratory (WSL) of sulphur dioxide emissions has been used. The data is for the year 1991. The data are in two forms; point sources such as chimneys and therefore assumed to be high level and area averages which are assumed to be low level. The area averages are on a 10km national coordinate grid.

Outside the UK, data from the EMWP inventory is used. This data is on a 150x150km polar stereographic grid and combines both low and high level sources.

The procedure for combining the various datasets to create a dataset for use in the model is as follows.

First of all non-sulphur sources are assigned to the lowest model level. These are arbitrarily set to be 30 tonnes SO_2/year for land points and 10 tonnes SO_2/year for sea points.

Then WSL area sources are interpolated from the source grid to the model grid which has been converted to national coordinates. The interpolated data is scaled so that the mean is conserved, to allow for interpolation errors, and for the ratio in area between the source and model grid boxes.

The WSL point sources are used to set emissions in model levels above the lowest level. The height of the emission source is multiplied by 1.5 to take into plume rise and then the model layer in which it lies is calculated. The emission is then added to the nearest model grid point on that level.

The EMEP data are then interpolated from the source polar stereographic grid to the model grid.

The final emission source field for model level one is found by adding the non-sulphur sources to the WSL area source if present otherwise to the EMEP source.

Once the emission source field has been created then it is possible to calculate an initial total aerosols field. This is done by applying a recursive filter to the emissions field and then distributing the filtered field through the atmosphere.

Appendix A

This appendix provides notes on an error in the calculation of the soil properties from WHS data for the Clapp-Hornberger soil hydrology scheme.

The information given in the text is what has been used since MOSES1 was first implemented (c. 1994). However, recent work by Dharssi (pers com) has discovered that this processing may be incorrect.

The problem originates from the fact that in their paper, Cosby use 'log' without specifying what base is used. Now, the usual convention is that log assumes log to the base 10 but in the processing for MOSES1 log to the base e has been assumed. (Normal convention is to use Ln in this case).

If log to the base 10 is assumed then the expression for SATHH becomes

$$\text{SATHH} = 0.01 \times 10^{(2.17 - 1.58Fs - 0.63Fc)}$$

This change then also changes the values calculated for the wilting point and te critical point.

Assuming a log to the base 10 also changes the expression used for the calculation of the saturated hydraulic conductivity.

The formula given by Cosby et al is

$$\log K_s^c = -0.60 - 0.64 F_c + 1.26 F_s$$

where K_s^c has units of inches per hour.

In the UM, the saturated conductivity (K_s) has units of $kgm^{-2}s^{-1}$, equivalent to mm s^{-1}. To convert between the two units we can use

$$K_s = \frac{25.4}{(60*60)} K_s^c = 10^{-2.15} K_s^c = e^{-4.95} K_s^c$$

(25.4 is the number of millimetres in an inch, 60*60 is the number of seconds in an hour)

Assuming natural logarithm then we obtain the equation currently used

$$K_s = e^{(-5.55-0.64F_c+1.26F_s)}$$

However, using logs to the basse 10 we obtain

$$K_s = 10^{(-2.75-0.64F_c+1.26F_s)}$$

Values calculated using the two sets of equations have been compared against soil properties used by other centres and this suggests that the values calculated using logs to the base 10 are the correct ones.

A word of caution however, do not start using the ‘correct’ values in the UM without first conducting a full assessment of the impact.

Location of datasets.

Datasets for standard UM configurations are held centrally but it is also possible to create datasets for any desired resolution by using the facilities described in UMDP 73.

General Notes:

The datasets are in ancillary file format as described in UMDP F3. The control routines for ancillary fields are described in UMDP C7. All the datasets that are packed as 32 bit method except for WHS types and land sea masks and they are also written to be ‘well-formed’ (even though the SX6/8 has no concept of well formed files, the UM I/O routines will fail if the dataset is not written as such). The UM utilities described in UMDP F5 such as pumf can be used to examine the contents of an ancillary file.

Directory Structure

The centrally held datasets are all stored under the directory structure

$UMDIR/vn$VN/ancil/SUBMODEL_TYPE/RESOLUTION

where
$UMDIR is the unified model path variable

$VN is the unified model version path variable (eg 5.1)

SUBMODEL_TYPE is either **atmos** for the atmospheric model, **ocean** for the ocean model or **slab** for the slab model.

RESOLUTION is the resolution indicator. Met Office users should check the Unified Model directory or Metnet for details of the latest datasets.

Within each directory (6.4 onwards) is Released_Notes file which should be consulted for details of the datasets that the directory contains.

File Naming Conventions

Files have names of the form

qr[TYPE].[CONTENTS]_[SUFFIX}

where

[TYPE] is either:
clim for time varying climatological fields
parm for non-time varying parameters
m[MMM] for single month/seasonal fields. In this case [MMM] denotes the month or season for which data is valid, eg jun, jfm.

[CONTENTS} describes the contents of the dataset

[SUFFIX] is used when alternative versions of datasets are available or when it is not possible to distinguish datasets otherwise. Often it is used to denote a development dataset but if at the time of the next UM release it is to be the main dataset, then the SUFFIX parameter would be dropped. In this instance, the previous datasets may be retained with a SUFFIX of ‘old’.

REFERENCES

Brutsaert,W. (1982). `Evaporation into the atmosphere. Theory, history and applications'. D. Reidel Publishing Company, Dordrecht.

Calder,I.R., Wright,I.R. and Murdiyarso,D. (1986). `A study of evaporation from tropical rainforest - West Java'. *Journal of Hydrology*, **89**, pp13-23.

Cosby, B.J., Hornberger, G.M., Clapp, R.B. and Ginn, T.T. (1984). `A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils.` *Water Resources Research*, **20** pp682-690.

Dickinson, A. and Wilson, C.W. (1991). `Interpolation Techniques and Grid Transformations used in the Unified Model'. (Unified Model Documentation Paper (UMDP) S1. Unpublished paper.)

Dingman,S.L., Barry,R.G., Weller,G., Benson,C., Le Drew,E.F. and Goodwin,C.W. (1980). `Climate, snow cover, microclimate and hydrology'. In An Arctic ecosystem: The coastal tundra at Barrow Alaska, Brown,J., Miller,P.C., Tieszen,L.L. and Bunnel,F.L. (eds). Dowder, Hutchinson and Ross, Stroudsburg, Pennsylvania.

Eagleson (1970) `Dynamic Hydrology'. McGraw Hill, New York.

Evans,R (1995) 'Creation of Data Fields Required for Effective Roughness Parameterisation'. Forecasting Research Technical Report No. 146. (Unpublished Paper).

FAO (1995). ‘The digital soil map of the world’. FAO, Rome.

Folland, C.K. and Parker, D.E. (1990). `Observed Variations of Sea Surface Temperature' in `Climate-Ocean Interaction', M.E. Sclesinger (ed).

Halldin,S., Saugier,B. and Pontailler,F.Y. (1984). `Evapotranspiration of adeciduous forest. Simulation using routine meteorological data'. *Journal of Hydrology*, **75**, pp323-341.

Houldcroft,C.J., Grey,W.M.F., Barnsley,M., Taylor,C.M. and Los,S.O. New vegetation albedo parameters and global fields of backgroudn albedo derived from MODIS for use in climate model. To be published, currently available from Department of Geography, Swansea University, Singleton Park, Swansea, SA2 8PP.

Li,D. and Shine,K.P. (1995) 'A 4-Dimensional Ozone Climatology for UGAMP Models'. UGAMP Internal Report No. 35 (Unpublished paper).

Lindroth,A. (1985). `Canopy conductance of coniferous forests related to climate'. *World Resources Research*, **21**, pp297-304.

Lu, S., Ren, T, Gong, Y and Horton,R (2007) An improved model for predicting soil thermal conductivity from water content at room temperature. *Soil Science Society of America*, **71**, pp8-14.

Mason, P.J. (1985). `On the parameterization of orographic drag', Proc. of the seminar on physical parametrizations for Numerical Models, Reading, 9-13 September 1985, ECMWF.

Meteorological Office (1991) `Meteorological Glossary`, HMSO.

Monteith,J.L. (ed) (1976). `Vegetation and the Atmosphere, 2'. Academic Press, London.

Parker, D.E, Jackson, M., Horton, E.B. (1995) Forthcoming Climate Research Technical Note.
Posey,J.W. and Clapp,P.F. (1964). `Global distributions of normal surface albedo'. *Geofis. Int.*, **4**, pp333-348.

Peters-Lidard,C.D., Blackburn,E., Liang,X and Wood,E.F. (1998) The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures. Journal *of the Atmospheric Sciences*, **55**, pp1209-1224.

Post,W.M. Emanuel,W.R., Zinke,P.J. and Stangenberger,A.G. (1982). 'Soil carbon pools and world life zones'. *Nature* **298** pp156-159.

Robinson,D.A. and Kukla,G. (1984). `Albedo of a dissipating snow cover'. *Journal of Climatology and Applied Meteorology*,

23, pp1626-1634.

Rockwood,A.A. and Cox,S.K. (1978). `Satellite infrared surface albedo over northwestern Africa'. *Journal of Atmospheric Sciences*, **35**, pp513-522.

Scholes,R.J., Skole,D. and Ingram,J.S. (1995). 'A global database of soil properties: Proposal for implementation. IGBP-DIS working paper #10, University of Paris, France.

Shuttleworth,W.J., Gash,J., Lloyd,C.R., Moore,C.J., Roberts,J., Marques Filo,A.O., Silvia Filo,V.P., Ribeiro,M.N.G., Molion,L.C.B., De Sa,D.A., Nobre,J.C.A., Cabral,O.M.R., Patel,S.R. and De Moraes,J.C. (1984). `Eddy correlation measurements of energy partition for the Amazonian forest'. *Quarterly Journal of the Royal Meteorological Society*, **110**, pp1143-1162.

Sud,Y.C. and Smith,W.E. (1984). `Ensemble formulation of surface fluxes and improvement in evapotranspiration and cloud parameterizations in a GCM'. *Boundary-Layer Meteorology*, **29**, pp185-210.

Thompson,N., Barrie,I.A. and Ayles,M. (1981) . `The Meteorological Office rainfall and evaporation calculation system: MORECS. Hydrological Memorandum No. 45, Meteorological Office, Bracknell.

Van Genuchten, M.Th. (1980). 'A closed form of the equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc. Am. J., **44**, pp892-898.

Warrilow, D.A. and Buckley, E. (1989). `The impact of land surface processes on the moisture balance of a climate model'. *Annales Geophysicae*, **7**, (5) pp439-450.

Warrilow,D.A., Sangster,A.B. and Slingo,A. (1986). `Modelling of land surface processes and their influence on European climate'. Meteorological Office, Dynamical Climatology technical note 38, (unpublished paper).

Willmott, C.J., Rowe, C.M. and Mintz, Y. (1985). `Climatology of the Terrestrial Seasonal Water Cycle', *Journal of Climatology*, **5**, pp589-606.

Wilson, M.F. and Henderson-Sellers, A. (1985). `A global archive of land cover and soils data for use in general circulation climate models'. *Journal of Climatology*, **5**, pp119-143.

Woodward,F.I., Smith,T.M. and Emanuel,W.R. (1995). 'A global land primary productivity and phytogeograpy model.' *Global Biogeochemical Cycles*, **9**, pp471-490.

Woodward, S. (2001). ‘Modeling the atmospheric life cycle and radiative impact of mineral dust in the Hadley Centre climate model.’ *Journal of Geophysical Research*, **106**, No D16, pp18155-18166.

Zinke,P.J., Stangenberger,A.G., Post,W.M., Emanuel,W.R. and Olson,J.S. (1986). ‘Worldwide organic soil carbon and nitrogen data’. NDP-018, Carbon Dioxide Information Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.