LEACHE INPUT DATA FILE

John L. Hutson

(29/07/2025)

INTRODUCTION

The format of input data files is similar for all versions of LEACHM with respect to simulation duration, boundary conditions, flow mechanism, output file specifications, soil physical properties and water-related input data items. LEACHM, the generic name for all 1D versions, encompasses a water-only version (LEACHW) and several chemistry-specific versions, e.g.. LEACHP (herbicides and pesticides), LEACHN (nitrogen and phosphorus), LEACHC (major cations and anions), LEACHB (microbial population dynamics) and LEACHE (similar to LEACHP but focussing on PFOS compounds). A 2D or multi-region version of LEACHP is also

available, as well as spatial or GIS versions of LEACHW, LEACHP and LEACHN. This document describes input data for LEACHE, much of which is identical to that required by all versions.

LEACHM has Excel-based input data files. Embedded macros trigger various actions, e.g. the '*Save and Run*' macro saves both the Excel data file and a *txt, prn* or *csv* file which is then read by the LEACHM executable. Additional features are easily added to the Excel file by users, such as plots of soil properties vs depth, or data cells containing equations, for example to relate organic carbon to depth. Data cells in Excel can contain formulae, but be aware that the value read by the Fortran code will be that visible in the cell, not the underlying full-precision value, so ensure that the visible format is sufficiently precise.

Conditional formatting may lead to fading of certain items in the Excel data file, implying that the item is read but not used. This formatting has no effect on a simulation, since all data file items are read and those that are relevant are used. The conditional formatting is intended as an aid when preparing an input data file. For example, if a unit gradient lower boundary condition is selected then all parameters relevant only to a water table lower boundary will be faded. Or if a value for bulk density is defined in the row of uniform profile physical data, depth-variable bulk density values will be faded. Be aware that this conditional formatting is easy to lose if data are copied or inserted from other files, so regard the formatting as a guide and not necessarily definitive. Where possible use the '*Values*' option instead of '*All*' when copying and pasting data using Excel's '*Paste Special*' tool. Conditional formatting can be changed if desired. It is useful to keep a copy of the template data file as a reference.

All data is read free-field from the left-hand side. This means that comments, calculations and graphs can be inserted in the data file to the right of any data to be read. Sections of the data file which can have records (rows) inserted or deleted to match the specified number of soil segments, crops etc. have a beige fill, all others are green. No other rows should be added or deleted, as this would require a code change, and there are no blank rows in the data file. Numerical data items that have to be integers are coloured red; all others, which can be decimal or scientific format, are black. Weather data do not usually display conditional formatting, since those data are often generated by a weather utility which processes daily weather station data, and are appended to a LEACHM data file.

LEACHM input data files are comprised of several sections, some of which are common to all versions of the model, and others which are specific to a particular chemical version. For LEACHE the structure is:

Simulation duration, profile discretization and flow mechanism

Output files

Soil physical properties

Vegetation

Chemical properties

Binder application

Irrigation

Weather

These sections are separated or delineated in the Excel data file by double rows of asterisks. A reason for clear conceptual delineation is that in the spatial, or GIS-linked version of the model (available for LEACHW, LEACHN, LEACHP and LEACHE) the sections are separated; multiple versions of each (library files) can be created to encompass the range of properties encountered in a simulated area.

FILE NAMES AND FORMATS

All Excel input file names consist of a prefix defining the LEACHM version being used, followed by an underscore and an 8-character file name or identifier. For example, *LEACHW_Watetest.xlsm*, is the name of the template file for LEACHW, the water-only version and *LEACHP_Pesttest.xlsm* is the template file for LEACHP, the pesticide version.

File identifier: This name (8 characters, letters or digits, with no extension, e.g. *Watetest*) must be present. The 'Save' macro saves the file as an Excel *.xlsm* file, e.g. *LEACHW_Watetest.xlsm*. The prefix identifies the version of LEACHM used, e.g. LEACHW, LEACHC, LEACHN, LEACHP, LEACHE, LEACHB. The 8-character file identifier originated when there was an 8-character limit on Fortran variable names, but is maintained because it simplifies file storage and searching, is used in many LEACHM utilities, and an 8-digit numerical identifier reflects the library files used in GIS-linked versions of LEACHM.

The flag value adjacent to the file identifier determines the format of the text-based file which will be read by the executable, which can be *.txt*, *.prn* or *.csv* format. Records in *.csv* or *.txt* files can exceed 240 characters, but *.prn* files are limited to 240 characters. LEACHM reads space-, tab-, or comma-delimited files, so *.prn* files must have at least one space between data values. The row width, font and numerical formats in the Excel template input files usually lead to readable *.prn* files. The most common instances of invalid (>240 characters) *.prn* files are encountered in LEACHN and LEACHP files which contain many chemical species. The *.csv* option is most useful in cases where sections of the data file may need to be re-imported into Excel.

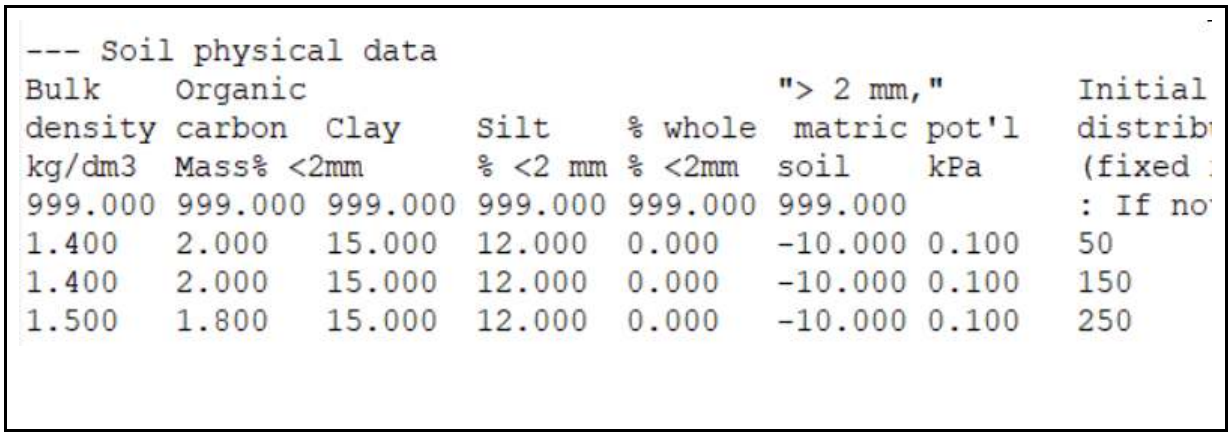
The *Save and run* macro saves both the Excel file (*e.g. LEACHW_Watetest.xlsm*) as well as a version in the designated text format (printer text *Watetest.prn*, comma-delimited *Watetest.csv*, or text *Watetest.txt)* and calls a batch file which executes the model using the designated data file. Be aware that the saved *.txt*, *.csv* or *.prn* files, any of which the LEACHM executable can read, will contain the values that appear in the Excel data cells, not the underlying values which may have greater precision. A cell's numerical format must be appropriate for the desired data precision - 'what you see is what is used'.

The following examples show a snippet of an Excel data file saved as *.txt*, *.csv* and *.prn* files.

A section of an Excel input data file

78	**--- Soil physical data**				
79	Bulk	Organic			> 2 mm,
80	density	carbon	Clay	Silt	% whole
81	kg/dm3	Mass% <2mm	% <2 mm	% <2mm	soil
82	999.000	999.000	999.000	999.000	999.000
83	1.400	2.000	15.000	12.000	0.000
84	1.400	2.000	15.000	12.000	0.000
85	1.500	1.800	15.000	12.000	0.000

The same section saved as a .txt file



```
--- Soil physical data
Bulk    Organic                              "> 2 mm,"       Initial
density carbon  Clay    Silt    % whole  matric pot'l   distrib
kg/dm3  Mass% <2mm      % <2 mm % <2mm   soil    kPa     (fixed
999.000 999.000 999.000 999.000 999.000 999.000         : If no
1.400   2.000   15.000  12.000  0.000   -10.000 0.100   50
1.400   2.000   15.000  12.000  0.000   -10.000 0.100   150
1.500   1.800   15.000  12.000  0.000   -10.000 0.100   250
```

Saved as a .csv file

```
--- Soil physical data,,,,,,,,,,,,,,,,,,,,,,,,,,,,
Bulk,Organic ,,,"> 2 mm,",Initial ,Relative root,,,,,,,,,
density,carbon,Clay,Silt,% whole, matric pot'l,distributic
kg/dm3,Mass% <2mm,% <2 mm,% <2mm,soil,kPa,(fixed roots), ,
999.000,999.000,999.000,999.000,999.000,999.000,,: If not
1.400,2.000,15.000,12.000,0.000,-10.000,0.100,50,,,,,,,,,
1.400,2.000,15.000,12.000,0.000,-10.000,0.100,150,,,,,,,,
1.500,1.800,15.000,12.000,0.000,-10.000,0.100,250,,,,,,,,
```

Saved as a .prn file

```
--- Soil physical data
    Bulk       Organic                                  > 2 mm,     Initial    Relative root
   density      carbon        Clay         Silt        % whole    matric pot'ldistribution for
   kg/dm3     Mass% <2mm     % <2 mm      % <2mm         soil         kPa     (fixed roots)
   999.000      999.000      999.000      999.000      999.000      999.000                 : I
    1.400        2.000       15.000       12.000        0.000      -10.000       0.100
    1.400        2.000       15.000       12.000        0.000      -10.000       0.100
    1.500        1.800       15.000       12.000        0.000      -10.000       0.100
```

While all three saved formats can be read by LEACHM and lead to identical output, the *.prn* file is the easiest to visually comprehend and use in documents or presentations.

DURATION, PROFILE AND FLOW MECHANISM

Date format: Dates must be expressed as a six-digit number with no spaces or delimiters. Do not use any of Excel's date formats. Option 1 (*mmddyy*) requires dates as month, day, year; option 2 (*ddmmyy*) as day, month, year. For option 1, January 10, 1988, is written as 011088; for option 2 it is written 100188. Although the six-digit date format limits simulations to 100 years it is rare that simulations extend for longer periods. However, simulations can be cycled through the same set of boundary conditions many times. LEACHM recognizes leap years and assumes a change of century (from year 99 to year 00) refers to calendar years 1999 to 2000 (the year 1900 was not a leap year).

Starting date: The simulation begins at the start of this day. This must be a six-digit date, and not a day number. This starting date will be day 1 in the simulation.

Ending (date or day number) : The simulation ends at the end of this date or day number. From here onwards, calendar dates and/or day numbers may be used in the data file. Any input date or day number having a numerical value below 010100 (<= 10099) is assumed to be a day number rather than a

calendar date, therefore day numbers can only be used in the input for elapsed times less than 10100 days, or about 27 years. For periods exceeding 10099 days, calendar dates have to be used. Day numbers are convenient for short-term simulations such as laboratory column studies or short field experiments. No date in the data file can be earlier than the starting date (see note below about changes of century); for example, the first date in the crop section, or the first rain event, must not precede the starting date. Dates can extend beyond the ending date. It is useful for example, after setting up a data file which may extent over say, 40 years, to do an initial simulation for just 1 day to ensure that all data can be read.

LEACHM accounts for leap years and a change of century. An end year less than the start year signifies a change of century. For example, to simulate a period from 1945 to 2035, using the day/month/year format, enter the starting date as 010145 and the ending date as 311235. The year '00' is assumed to be 2000, a leap year. (In the unlikely event of running a simulation from, e.g. 1845 to 1935, LEACHM will include a 29th February in the year 1900, which was not a leap year.)

Simulations run from 0h00 on the starting date to 24h00 on the ending date or day number. To simulate only one day, the minimum possible simulation period, use the same date for both starting and last day, or specify the last day number as '1'.

Be aware that day numbers are used in the code instead of dates. Input dates are converted to day numbers immediately after they are read and converted back to dates prior to output. The 10099 day number limit does not apply in the code, only to the input data. Also, time in LEACHM is expressed in units of days, not, for example, seconds. So there is a distinction between 'day number' (an integer and proxy for a calendar date) and 'days', which is the time unit used. For example, midday on day 4 is equivalent to an elapsed time of 3.5 days.

Number of cycles through rainfall, crop and chemical application data: For most purposes set this to '1'. LEACHM can do longer period simulations by repeating the boundary conditions, management and vegetation specifications. At the start of each cycle all variables, with the exception of the current profile water and chemical content distributions, are re-initialized. Thus if three years of rainfall or irrigation data is to be simulated once then set the number of cycles to 1 and the last date or day number three years after the starting date. If one year's data is to be repeated three times then specify the ending date one year after the start and specify 3 cycles. For slowly-changing soil chemical constituents, such as soil carbon or deep soil chloride profiles under conditions of low rainfall, it is sometimes necessary to repeat a 100-year weather data set several times to reach a dynamic equilibrium. This obviously assumes that repeating weather data sets is acceptable.

Ensure that the start and end dates for each cycle are appropriate. Do not for example, specify five cycles, each starting in January and ending in June, since the profile and chemical data at the end of June are unlikely to be appropriate for the next cycle assumed to start in January.

Soil profile

Profile depth (mm): The vertical thickness of the soil profile, or length of a soil column. This

extends from the soil surface to the depth at which the lower boundary condition will be defined.

Segment thickness (mm): The profile depth is divided into segments of equal thickness, therefore the *profile depth* should be a multiple of the segment thickness. For example, if a 2000 mm profile is to be simulated, then 100 mm segments will require data for 20 depth segments to be specified; if the profile depth were 350.7 mm and the segment thickness 16.7 mm, then 21 layers will be required.

Finite-differencing nodes used in LEACHM are located in the centre of the segment (i.e. block centred), which is the reason why depths listed in output files are the mid-points of the segments and not their boundaries. A segment thickness of 25 to 100 mm is usual for field simulations, but can be reduced if necessary. For simulating outflow from small soil cores, for example, segment thickness of 1 mm may be appropriate.

The maximum number of segments (usually about 250), along with other maximum array dimensions are specified in the PARMS.INC file. In the rare event that any array dimension needs to be increased then the value can be changed and the program recompiled.

Water flow and lower boundary condition

Water flow: Water flow can be simulated using either Richards' equation, a mobile/immobile capacity (tipping-bucket) model adapted from Addiscott (1977), or steady-state flow (continuous or interrupted) through a soil column. Capacity flow is an option in LEACHW, LEACHP and LEACHN but not in LEACHC, LEACHB or LEACHE.

Largest time interval within a day: Determines the minimum number of time intervals per day. This is used only for the Richards flow option. A value of 0.05 or 0.1 day is usually satisfactory. The maximum value allowed is 0.1 day, and LEACHM will default to 0.1 day if a value greater than 0.1 is entered. Time steps are automatically reduced during periods of high water flux or to coincide with print specifications, rainfall or chemical application events. A minimum time step (currently 10^{-8} d) is specified in the code. If the Addiscott (capacity model) option is chosen for water flow, then time steps default to daily (with internal 0.1 day steps for ET and temperature calculations). The *Largest time interval* is read, but not used.

Lower boundary: For the Richard's water flow option this flag can be:

1: *Constant potential* or a fixed-depth water table.

2: *A freely draining profile* having unit hydraulic gradient at the lowest node.

3: *Zero flux*, i.e. an impermeable base.

4: *A seepage face*, e.g. a lysimeter from which water drains when the bottom node exceeds a specified matric potential, but is a zero flux condition otherwise. If the seepage face lower boundary condition is selected then the starting *depth to water table* value will be used to determine the matric potential above which water can drain from the profile. Thus specifying a depth to the water table of 1400 mm in a situation where the simulated profile depth was 1100 mm means that the matric potential at the lower boundary will

be a constant -300 mm or -2.94 kPa, mimicking a lysimeter having a drainage system subject to a 'suction' of 300 mm. A water table depth of 1100 mm will result in a lower boundary potential of zero, i.e. a lysimeter having gravity drainage.

5: *Tile drainage* based on the Hooghoudt equations. Currently, as a trial, the height of a 'regional' or equilibrium water table above the lower boundary is defined. When the water table descends to this depth downward flux ceases. A constant in the code (which can be altered if needed) encompasses Hooghoudt parameters such as diameter of the drains, the distance between them and data relating to conductivity and thickness of material below the soil profile. A specified drainage rate (or effective lower boundary saturated conductivity) for a water table height above the regional equilibrium height acts as a 'matching factor' which enables the drainage rate at other water table heights to be calculated during a simulation. The higher the water table the greater will be the drainage flux.

The capacity flow option can have a defined water table, free draining (unit-gradient), zero flux or tile drainage lower boundary. For both Richards and capacity flow options, specifying a low saturated conductivity at the lower boundary can lead to a perched water table which will fluctuate over time, since the drainage rate from the profile cannot exceed the saturated hydraulic conductivity. (The capacity option is currently undergoing revisions, so these options may change.)

OUTPUT FILES

LEACHM has a range of possible output files. They are named using the 8-character LEACHM input file identifier (e.g. *filename*) and an extension unique to the output file type.

- The main output file (*.out*) contains detailed mass balance data for water and chemicals as well as amounts and concentrations at each depth node.

- A 'profile' file (*.prf*) lists soil profile data in a format suitable for generating profile depth property plots, but does not usually contain any information additional to that in the input or *.out* files.

- A summary file (*.sum)* contains a time series, one record per print interval, of mass balance components and fluxes for both water and chemicals. These are listed for the whole profile as well as for each of four macro-segments or observation nodes.

- A breakthrough file (*.btc)* prints time-specific and cumulative water and chemical fluxes across the lower profile boundary.

- A 'grid' file (*.grd)* file records selected depth:time matrices for water content, potential and flux, temperature, and chemical concentrations and flux at time intervals corresponding to those used for the *.sum* file.

- An irrigation *.irr* file records triggered irrigation applications.

The *.prf*, *.out* and *.irr* output files are always generated; the *output files* flag defines which of the remainder will be generated, namely: 1: *.out* only; 2: *.out*, + *.sum*; 3: *.out* + *.sum*, + *.btc*; 4: *.out* + *.sum* + *.btc* + *.grd*.

Be aware that daily prints to any of the output files will be values at the end of a day (midnight).

Temperatures in particular will reflect this so may appear lower than expected.

The *.out* file (Mass balance and profile data)

Units for depth data refer to chemical concentration data in profile segments. In LEACHW, which has no chemical species, this item is included as a placeholder; the value is read but not used. Chemical-related options are described in the input documentation for each model.

Node print frequency is usually 1, i.e. print every node (or segment). A value of 2 will mean that alternate nodes are printed; a value of 3 prints every third node. This is intended to save space when printing data for profiles having many segments.

Tables printed : Three tables in the *.out* file can be printed at each print time: 1) a mass balance table, 2) water content, potential and flux, and 3) soil temperature, root fraction, and water uptake by roots in each segment. This flag is usually set to 3, but if no crop is included it can be set to 2.

Reset .out file cumulative values after each print: Depending on the duration and nature of the simulation it may be desirable to reset cumulative values after each *.out* print. For example if a 100-year simulation were performed with annual prints to the *.out* file, resetting the cumulative values means that each *.out* print will show mass balance data for a single calendar year. Not resetting could lead to extremely high cumulative values, which are difficult to interpret. Cumulative values will nonetheless be available in the *.sum* file if desired. Or prints could be specified for the end of each crop period. For a shorter simulations with daily, weekly or monthly prints, cumulative values are likely to be more useful.

Print times: The *.out* file can be printed at fixed time intervals (1), or at defined times (2) which are not necessarily equally spaced.

For option 1, time intervals between prints can be specified to about 0.001 day. Entering 999 as the time interval leads to default calendar year prints. For printing at irregular time intervals or dates (for example, to match sampling dates and times) specify print option 2 and the number of prints desired. Note: LEACHM will expect to read data for both option 1 and option 2, but the option and values used will depend on the print option flag.

A date (or day number) and time (fraction of a day) must be specified for each of the number of prints specified for *option 2*. Even if option 1 were chosen there need to be dates and times corresponding to the number specified under print option 2, but in this case they will be read and not used, so dates can, e.g., be an arbitrary 999999, a 'date' that will never be encountered.

The *.sum* file (Macro-segment or observation node time series)

A *.sum* file is generated if the *number of output files* is >= 2.

.Sum and .grd file print intervals: Print intervals for both the *.sum* and *.grd* files are specified in days. The smallest print interval is 0.001 day. For example, a value of 1 will trigger daily prints, a value of

7 will trigger weekly prints, a value of 0.042 will trigger hourly prints. Entering 999 will trigger a print at the end of each calendar month.

Print all cycles If the simulation includes more than one cycle through the boundary conditions (see *Number of repetitions of rainfall, crop and chemical application data* above) this flag defines whether output from all cycles is to be printed to the *.sum* file, or that from the last cycle only.

Initialize cumulative values after each .out print: Sometimes it may be desirable to reset all cumulative *.sum* file values after printing to the *.out* file. For example, if prints to the *.out* file were annual and cumulative values in that file were reset after each print. Resetting the *.sum* values at the same time may make data interpretation simpler.

Depths 1, 2, 3 and fourth segment: Data for four 'macro-segments' or 'observation depths' are printed in the *.sum* file. The first three are specified depths (if zeros are entered, the macro-segments will be as close to thirds of the profile as possible). The fourth macro-segment can be 1) the current root zone, 2) the whole profile, 3) the remainder of the profile from depth 3 (if that was not the lower boundary) to the bottom of the profile, or 4) the surface to current water table depth. For example, if the profile depth were specified as 1500 mm, depths 1, 2, and 3 were 100, 300 and 800 mm, and option 3 was selected for the 4th segment, the *.sum* file will contain data for the layers 0 to 100 mm, 100 to 400 mm, 400 to 800 mm and 800 to 1500 mm. If zeros were specified for depths 1, 2 and 3 and the 4th segment option was 2, then the macrosegments would be 0 to 500 mm, 500 to 1000 mm, 1000 to 1500 mm, and the 4th macrosegment would be the whole profile. Depths used by options 1 (the root zone) and 4 (surface to water table) may vary over time. The macro-segments and their definition is described in more detail in the section describing the summary file output.

The *.btc* file (Water and, in other versions, chemical breakthrough data)

A breakthrough file (*.btc*) is generated if the *number of output files* is >=3.

Incremental depth of drainage water prior to print: Prints to the *.btc* file will occur at the end of the time step during which cumulative drainage since the previous *.btc* print exceeds the *incremental depth of water* value specified. The drainage increment recorded may thus be slightly greater than that specified, but both incremental drainage and elapsed time will be correct. Prints will be equally spaced in time for steady-state laboratory column simulations, but for transient flow simulations they will be irregular, since drainage rates vary over time.

The *.grd* file (Depth:time matrices)

A grid file (*.grd*) is generated if the *number of output files* is >=4. The grid file records, at time intervals specified for the .SUM file, and the depth nodes intervals specified for the .out file, water content, water potential, the concentration of a selected chemical species, and temperature The first column in the output matrix lists node depths, and the first two rows in each section lists elapsed time, in both years and days. The depth:time matrices can be used to generate time and/or depth plots, or used for further

statistical analyses. Cells can be shaded using conditional formatting to generate a pictorial view of water and chemical distributions over time.

Start and end dates or day numbers: Since the *.grd* file can be very large there is an option for printing between specified time limits. The difference between these two times, divided by the print interval (that defined for the *.sum* file) determines the number of prints to the *.grd* file. The *.grd* array sizes may be limited by the parameters specified in the PARMS.INC file (usually ~8000, but this can be changed and the code recompiled). The maximum number of columns in Excel, if used for post-processing, can constrain the number of time prints to the *.grd* file, but that number is now 16834. However, if the number of *.grd* time intervals does exceed the code array size (e.g. 8000) then prints to this file will be delayed until the remaining prints are fewer than the code array size, giving preference to later times. If for example, prints to the *.grd* file were daily, a 100-year simulation is performed and the *.grd* array sizes were 11000 prints, only the last 30.1 years (~11000 days) will be written to the *.grd* file. If prints to *.grd* were at 0.1 d intervals then only about 3 years can be recorded. There are no array limits on the number of prints to the *.sum* file.

Records every node to this depth (mm) specifies the depth, from the surface, to which data will be recorded. The *node print frequency* specified for the *.out* file will be used here. The depth limit conserves conserve memory, and we are often more interested in concentrations and fluxes in upper sections of deep profiles, e.g. the root zone.

The *number of chemicals to track* can be specified. To reduce output file size this is usually a selection of the species included in the simulation. Since LEACHW has no chemicals, this section of the data file is merely a placeholder but to preserve the common file format must not be deleted. The Excel template data files usually make provision for up to 12 chemical species, but more cells can be added if required. Unused cells are not read.

Chemical species to track lists the ID numbers of the chemicals to be recorded. They are defined using from the species ID numbers listed in the chemical properties section of the data file. They need not be in numerical order and more than one instance of the same chemical species can be specified in the event that different phases (e.g. soluble, sorbed, total etc.) are desired.

Data types tracked in the .grd file: To save storage space the variables to be tracked are flagged. Physical variables, ie. Water content, matric potential (actual and/or log values), water fluxes and temperature are always recorded. Chemical variables that can be recorded depend on the LEACHM version. For LEACHE they are : 1: Solute, mg/L, 2: Total, mg/dm3, 3: Total, mg/kg, 4: Current retardation coefficient, 5: Sorbed, mg/dm3, 6: Non-interface soln, mg/L, 7: Interface, mg/L.

SOIL PHYSICAL PROPERTIES

Dispersivity is used in convection-dispersion solute transport calculations and the value is merely a placeholder in LEACHW and for any capacity flow option. To avoid numerical dispersion, dispersivity values should lie between about half and twice the node spacing, otherwise adjust the node spacing to suit

the desired dispersivity value.

Particle densities are used, along with bulk density, to calculate porosity. A single value can be entered, assumed to apply to all soil components: clay, silt, sand, organic matter and if present, coarse (>2 mm) material. If the 'single value' is 999, a weighted particle density will be calculated for each soil segment using the density values specified for each component.

Starting matric potential option: Specifying *initial profile matric potentials* (option 1) or *field capacity values* (option 2) will not necessarily represent hydraulic equilibrium, so water fluxes are likely to commence within the profile at the start of the simulation. If hydraulic equilibrium is desired then specify hydraulic equilibrium *based on deepest node listed* (option 3) or on a *depth to a hypothetical water table* (option 4). For options 3 and 4 starting matric potentials are calculated to be in hydraulic equilibrium through the profile, i.e. decreasing towards the surface by the equivalent of the gravitational potential. There may of course be flow across the lower boundary at the start. An initial water table can be above or below the profile lower boundary.

Hypothetical surface crust: This is a new option in LEACHM to mimic the effect of a surface seal or crust on infiltration. Conceptually, the crust is a hypothetical 10 mm saturated soil layer having the saturated hydraulic conductivity (K_{crust}) specified here (mm/d), and a lower boundary corresponding to the current matric potential of the uppermost profile segment. This enables a maximum allowable water flux density across the soil surface to be calculated. Since we rarely know the properties of a crust it should be regarded as an approximation which may mimic the effect of a crust. This applies only during infiltration, not to evaporative fluxes, and does not change specified soil hydraulic conductivity values.

Particle size classification: Particle size classification systems used in different countries or in different disciplines can vary. While most soil particle size systems define clay as < 2 µm diameter, and sand as <2 mm, the upper limit of silt can vary. In the IUSS system, used e.g. in Australia and South Africa, it is 20 µm, in the US it is 50 µm and the UK, 63 µm. Definitions of fine, medium and coarse sand also vary. Pedotransfer functions, which estimate water retention properties from particle size distribution, were each developed assuming a specific size classification system. Some users are not aware of this and select pedotransfer function options which are inappropriate for the particle size data listed. The program now prints a warning message and stops if the specified size classification does not match the pedotransfer function assumption. (Particle size diameters are also used in LEACHE to estimate soil surface areas which in turn are used to estimate air-water interface sorption.)

Soil profile depth data: The soil profile section of the data file consists of two tables, listing soil physical parameters for each segment. Both tables have an initial row containing values which, if not 999, will override the segment values. This is a simple method of making selected soil properties uniform with depth without having to edit the depth-dependent data columns. Conditional formatting in the Excel data file should, when a value other than 999 has been entered in the uniform data row, lead to the underlying column of depth-related data to be faded, which indicates that the value will be read but not used.

Coarse material (>2 mm), i.e. gravel and stones, is not usually included in laboratory particle size

and soil chemical analyses since soils are usually passed through a 2 mm sieve prior to analysis. Predictive tools such as pedotransfer functions are usually based on < 2 mm soil material. Particle size and organic carbon data (excepting coarse material *per se*) in the LEACHM input file should be entered assuming coarse material is absent. Adjustments for the presence of gravel or stones are performed in the model code.

Depth or layer numbers are listed on the right-hand side of the depth tables for the convenience of the user. They are not read so can be any value, numeric or descriptive. Depth values (zero at the surface, negative with depth) are convenient for creating graphs of profile soil properties.

Bulk density is dry bulk density of the <2 mm fraction. If stones are present the dry bulk density is conceptualized as applying to the material between stones. In LEACHM, porosity is calculated from bulk density, the mean or weighted mean particle density values and the volume of gravel and stones, id present.

Humus carbon content (distinct from *organic matter*) is used in some pedotransfer functions, to define initial soil humus carbon content in the nutrient model (LEACHN), and to convert K_{oc} to K_d values in the LEACHP and LEACHE. *Organic matter* is required to calculate weighted particle densities and by other pedotransfer functions. The ratio of organic matter to organic carbon is defined in the model code as 1.70. Humus carbon excludes other undecomposed organic carbon pools such as plant residue and manure, which can be specified in LEACHN. Initial humus carbon or organic matter values are used in pedotransfer functions which are calculated at the start and assumed constant over time. Currently, changes in humus C during LEACHN simulations do not have any effect on bulk density, soil water retentivity and hydraulic conductivity values.

Clay and silt: These values may be used in pedotransfer functions, to calculate porosity using a weighted particle density, and to calculate heat capacity and thermal conductivity. In the event that zero values are specified for both clay and silt, a default value of 5% for each is used in the HEAT.FOR subroutine to enable estimation of thermal properties. In LEACHM it is assumed that the values are percentages of the <2 mm soil fraction, or soil 'fines', excluding organic matter, whereas organic matter is a percentage of the <2 mm fraction, including organic matter. Be aware that how clay, silt and sand are defined varies between laboratories. It can be a percentage of I) the initial soil mass (excluding coarse material), or ii) the initial mass after removal of organic matter. Some laboratories determine clay and silt, then sand by difference, i.e. sand = 100 - (clay + silt). Precipitates such as calcite and gypsum are rarely quantified. Measured particle size distributions also depend also on pretreatment and dispersion procedures as well as the analytical method (sedimentation or laser).

Gravel and stones (>2 mm): This refers to mineral particles larger than 2 mm. While clay, silt and sand fractions are usually expressed as a percentage of the sieved < 2 mm material (laboratory data), coarse material (> 2 mm) is usually expressed as a fraction of the total soil mass.

Starting matric potentials : These values define the starting matric potential distribution through the soil profile. There are four starting matric potential options:

1) Listed matric potentials: The program reads and uses these without further adjustment. The

values listed apply to the nodes, which in LEACHM are in the centre of the segments.

2) Field capacity values: Uniform or depth-related field capacity values are used as starting matric potential values. This is usually selected if an initial 'no-flow' scenario is desired for the capacity flow option.

3) Hydraulic equilibrium relative to the matric potential defined for the deepest node: Hydraulic equilibrium sets up an initial equilibrium scenario for Richards flow, where the decrease in matric or pressure potentials corresponds to the elevation above the lower boundary. The matric potential listed for the deepest segment is the only value used. Matric potentials in segments above the lowest segment are decreased by elevation ($\Delta\psi = -\Delta z/101.2$, where ψ is matric potential, kPa, and Δz is the segment thickness, mm). The values apply to the nodes, which are in the centre of the segments.

4) Hydraulic equilibrium relative to the depth to the water table: Similar to (iii) above, except that hydraulic equilibrium is based on the specified starting depth to the water table, which may be below the profile. At the start of the simulation there will be hydraulic equilibrium both above and below the phreatic surface. Thus for a profile depth of 1100 mm, specifying a depth to water table of 1400 mm means that the starting matric potential at the lower boundary of the profile will be -300 mm water (-2.94 kPa), a depth to water table of 1100 mm will result in a lower boundary potential of 0 kPa, whereas a depth to water table of 700 mm leads to a pressure potential of +400 mm or 3.91 kPa at the lower boundary. Steady-state water flow usually assumes a saturated soil column, so the starting potential would be 0 kPa and matric potentials do not change during the simulation. Saturated columns can have varying porosity as the pore water velocity will vary through the column in response to a constant water flux density. Specifying a water potential other than saturation for steady-state flow is appropriate only for uniform soil columns, as it is physically impossible to have steady-state flow in an unsaturated layered soil column in which the water potential is uniform. Also, while a saturated soil column can have a water flow rate corresponding to a pump rate, a real unsaturated column having a uniform matric potential can only have a flow rate corresponding to the soil's hydraulic conductivity at that matric potential. (The Richards flow option can be used to simulate steady-state water flow through unsaturated layered soil columns).

While not a soil physical property, relative root distribution (fixed roots) is included in these tables since it is a convenient location. If a plant growth option (see below) specifies a fixed root distribution, these values define that distribution. The values are relative, i.e. they can be percentages, relative values etc. After the values are read they are converted to fractions which total to 1 for the whole profile.

Retentivity option: Two water retention functions can be used in LEACHM, the van Genuchten function and the Hutson and Cass (1987) two-part function (which is based on the Campbell function). The parameters for either of these functions can be listed, or parameters for the Hutson and Cass function can be predicted using pedotransfer functions. The retentivity options and corresponding flag values are:

Defined parameters for:

1. Van Genuchten function

2. Hutson and Cass function.

Hutson and Cass parameters using pedotransfer functions:

3: South African soils using clay and silt as independent variables (Hutson, 1986)

4: South African soils using clay + silt as an independent variable (Hutson, 1986)

5: South African soils using clay as an independent variable (Hutson, 1986)

6: UK topsoils (data from Hall et al., 1977)

7: UK subsoils (data from Hall et al., 1977)

8: Australian soils (Minasny, 1999; McKenzie and Cresswell, 2002)

9: USA soils (Rawls and Brakensiek, 1985)

Pedotransfer functions are defined in the RETPRED subroutine, where additional pedotransfer functions can be added if desired. Water retentivity parameters are obtained by fitting the Hutson and Cass function to water contents predicted by the regression equations contained in a specified pedotransfer set.

The water retentivity options need not be the same through the profile, unless specified in the table of uniform soil properties. For example, the Hutson and Cass function can be used for part of the profile and van Genuchten for the remainder, or parameters fitted to measured retention values (external to LEACHM) can be used for part of the profile and pedotransfer function estimates for the remainder. Different pedotransfer functions can be used through the profile, for example, UK topsoil functions in upper layers and UK subsoil functions in lower layers.

These options can be somewhat confusing. All data are read, but the data items used depend on the retentivity option selected. The following is a brief summary:

Listed particle and bulk density values are used to calculate porosity; assumed to correspond to saturated water content. The hydraulic conductivity matching factor, i.e. a K value at a defined matric potential (columns E and F) is always used. If the retentivity option is 1, columns B, C, D are used in the van Genuchten function. If the retentivity option is 2, columns B and C are used in the Hutson and Cass retentivity function, along with column D in the corresponding Campbell K function. If the retentivity option is 3 to 8 then a pedotransfer function, depending on the option value, will use listed particle size data from the table above and columns B and C will be read but not used. Columns D to F will be used to adjust and scale hydraulic conductivity. *Soil retention function parameters:* These values will be used

only if retentivity options 1 or 2 were selected. Their meaning will depend on the water retention and conductivity function option selected, namely a, and *b* (Hutson and Cass), or α, *n* and θ_r (van Genuchten). Ensure that the units of *a* (kPa) or α (kPa^{-1}) are correct, and that the correct conversion factors are used when converting from other units. For example:

-10 mm = -0.0978 kPa (Campbell *a*) -10 mm^{-1} = -10.22 kPa^{-1} (van Genuchten α)

The van Genuchten function has three parameters, whereas the Hutson and Cass function has only two. However, the Campbell hydraulic conductivity function requires a 'pore interaction parameter', *p*, which is listed in the same column as the van Genuchten residual saturation.

If retentivity options 3 to 8 (pedotransfer functions) were selected then van Genuchten or Hutson and Cass retentivity parameters will be read but not used, as pedotransfer functions using listed clay, silt, bulk density and organic carbon values will be used to predict water retention values. The pore interaction parameter will be read and used.

Hydraulic conductivity matching factor: A single known K value at a specified matric potential for positioning or scaling the hydraulic conductivity-water content functions. If the saturated hydraulic conductivity is used then specify the matching matric potential as 0 kPa.

It is possible to use lookup tables, which list clay, silt, bulk density, organic carbon and corresponding water retentivity parameters and hydraulic conductivity values. However, they are usually project-specific and are not described here.

Matric potential at field capacity: The field capacity value is used if an automated irrigation option is replenishment of a specified depth of soil to field capacity. If the capacity flow routine is used then the field capacity matric potential corresponds to the drained upper limit for each soil segment. Be aware that field capacity is defined in several different ways, so it is important to assess the values in terms of the intended model application. (The capacity flow option in LEACHM is currently being revised, so data required may change.)

Addiscott preferential flow threshold (kPa): In the capacity flow option only that fraction of solute contained in water held above this threshold (the mobile fraction) mixes and moves with downward-flowing water, bypassing the immobile fraction. Equilibration between mobile and immobile solute in each soil segment occurs after transport. This value must be equal to, or less (more negative), than the field capacity matric potential.

Runoff

Surface water runoff is calculated using the USDA-NCRS curve number procedure if the slope specified is greater than zero. Runoff is calculated at the start of each rain event and subtracted from total rain for that event. Runoff may also be generated if the profile is saturated and the desired rate of infiltration exceeds the infiltration capacity of the soil. If steady-state water flow was specified there will be

no runoff, analogous to a column having water applied at a constant rate under pressure, and the surface runoff parameters will be read but not used.

The *depth of soil for calculating retention parameter* replaces the *SCS curve number (CN2).* Increasing the soil depth (and hence the amount of water) will reduce the runoff.

A slope (%) of 0 will cause the curve number routines to be bypassed. LEACHM uses slope-dependent runoff equations (Williams, 1991).

Infiltration-excess runoff is generated during a simulation if the amount and rate of water application exceeds the infiltrability of the soil. In some soils a surface seal or crust may accentuate this. A limiting surface hydraulic conductivity can be imposes to mimic a crust. A defined *Max surface K, mm/d* imposes a cap on surface hydraulic conductivity during infiltration. It is not used at any other time. A cap value higher than the soil hydraulic conductivity, e.g. 10 000 mm/d, implies no surface crust or seal.

VEGETATION AND CROPS

LEACHM is not intended to be a plant growth model; functions are used to represent root and cover growth through germination, emergence, maturity, senescence and harvest. Customized plant growth modules can be added to the model if desired. A cover fraction, ranging from zero to one, along with several adjustment factors are used to partition potential or reference ET into potential transpiration and potential evaporation.

In the standard version of the model there is no feedback between soil conditions and plant growth *per se*, but water and nutrient uptake are reduced if the plants are subject to water stress, and both water and nutrient stress indices are calculated. Plant roots and crop cover fraction are defined as either constant or growing. A cereal crop, for example, will have both a growing cover fraction and root distribution, whereas a deciduous plant may have a constant root distribution and a growing cover fraction. Constant roots or cover do not change over a growing season. Growth of roots and crop cover expansion are defined by start times for germination and emergence, root and cover maturity dates, onset of senescence if applicable, and a harvest date.

Plants present: 1 yes, 0 no: A value of 0 leads to all plant-related routines, including growth and transpiration, being ignored. This flag can be used to turn plants 'on' or 'off' without having to change any listed growth parameters..

Number of crops: The number of records of crop data listed. At least one is required. Dates, including germination dates, can extend beyond the simulation end date, but must not precede the simulation start date.

Minimum root water potential (kPa): This is the minimum value of the crown potential. Both this value and the wilting point, which is the soil matric potential below which no water is taken up by

plants, can limit transpiration.

Maximum actual transpiration/potential transpiration: When the soil surface dries, surface evaporation may fall below potential rates. Some of the energy which would have been used in evaporating water from the surface is utilized in increased transpiration. This ratio is the factor by which actual transpiration can be increased to compensate for reduced evaporation. A value of 1 represents no increase; a value of 1.2 represents a possible 20% increase.

Root flow resistance is a factor which determines whether or not, all else being equal, shallow roots absorb more water than deeper roots, reflecting resistance to water flow through the xylem vessels. When calculating the horizontal gradient between bulk soil and root at specific depths, the root potential is assumed to be the plant crown potential value increased by a *depth* x *root flow resistance* term in order to reflect resistance to flow through xylem vessels.

Solute concentration and osmotic potential: Depending on the version of the model, osmotic potential is calculated and used to modify root water uptake. A value of 0 assumes no osmotic effects. In LEACHP, the osmotic potential is calculated by multiplying the concentration of chemical species 1 by this value, so if osmotic effects are to included ensure that species 1 is assigned appropriate properties and concentrations. In LEACHN, the 'tracer' is the osmotic potential-determining solute species. In LEACHC, the osmotic potential is calculated from the concentrations of all dissolved species. Osmotic potentials are not accounted for in LEACHW.

Root water uptake defines the extent to which hydraulic conductivity will be included in the weighting of root water uptake through the profile. Selecting '1' will exclude any hydraulic conductivity weighting. The other options are to include the current K value (2), the cube root of the K value, thus reducing the effects of the highly non-linear K(θ) relationship (3), or using K values but imposing an upper limit of 1 mm/day (4). Option 5 implements the Feddes approach (Bristow et al., 2020), in which matric potentials are classed into categories: 0 to -1 kPa (no uptake), 1 to -2.5 kPa (linear increase), 2.5 to -20 kPa (optimal transpiration), -20 to -800 kPa (linear decrease), <800 kPa (no transpiration) uptake. In most cases the water uptake is somewhat insensitive to this choice, but in cases where a very small fraction of roots have access to continual supply of water (rising from a water table for example) excessive and unrealistic absorption may be calculated if the K values are included in the weighting. Other options are, or can be written into the WUPTAK.FOR subroutine. For example, limiting uptake if the water content or potential becomes too high in order to reflect the so-called 'physiological drought' observed for certain plant species in near-saturated soils.

Growth: There are three vegetation growth options, defined by flags 0, 1 and 2.

Flag 0: Constant or 'no growth' vegetation. A constant root distribution, defined in the *root fraction* column in the table of soil physical properties, and the *Maximum crop cover fraction* is used between the *Germination* and *Harvest* dates.

Flag 1: Both roots and crop cover change over time, using an adaptation of Davidson's (1978) corn

growth functions. The plant growth patterns, determined by the *Planting, emergence, plant and root maturity,* and *harvest dates,* are described in detail in the description of the GROWTH subroutine. Briefly, roots grow from the *germination day or date* to the *root maturity date*. Crop cover fraction increases from zero at the *emergence date* to the *maximum crop cover* (<= 1) at the *crop cover maturity date* and then declines to the *crop cover at harvest* by the *harvest date* or day number. The current crop cover fraction is used to partition reference evapotranspiration (ET_p) into potential transpiration (T_p) and potential evaporation from the soil surface (E_p).

Flag 2: A fixed root distribution and a growing crop cover. It is intended to represent deciduous plants which have an established root distribution at the start of each growing season. This can also be represented by setting the roots to 'grow' to maturity in one or two days.

(In LEACHE, the addition of a binder, together with a soil mixing option, can be specified. These operations are unlikely to occur without vegetation removal. If the date of a binder application is specified to fall within a crop growth period, that crop period will end the day prior to the binder application.)

Crop cover at harvest: To represent senescence, the crop cover is reduced gradually to a defined value at harvest. The reduction starts after 30% into the time interval between the date of cover maturity and harvest. This value is defined in the Growth subroutine.

Mulch effect % : The mulch effect reduces potential surface evaporation (Ep) by the specified percentage during the crop growth period only. Set value to zero for no effect. Hence, $E_p = ET_p$ *x (1 - Crop cover fraction) x (1 - Mulch effect/100).*

Transpiration scaling factor: A multiplier, or scaling factor, which further adjusts potential transpiration T_p only, not evaporation. This factor, which can be less than or greater than 1, is crop-specific, and is applied during the crop growth period only. Set this value to 1 for no further adjustment to T_p. Hence,

T_p = *ETp x Crop cover fraction x Transpiration scaling factor*

Triggered irrigation flag: This flag (if > 0) enables automated irrigation to be triggered according to criteria defined under *Automated irrigation* in the water application section of the data file. Defined or listed irrigation events will always be applied, regardless of the value of this flag or whether or not crops are present, but automated irrigation is triggered only when plants are present, i.e. not during fallow periods between crops.

Fraction of crop period irrigated: Irrigation is sometimes required only for part of the crop growth period, with the later period subject to drying. Triggered irrigation, if required, starts at the planting date and cannot extend beyond the defined fraction of the growing season. This does not override any specified irrigation events that may be listed.

The remaining items in the vegetation section, define nutrient uptake and cycling. They are read and used by LEACHN only. They are not read by other versions of LEACHM, although the columns are usually retained as placeholders. These data are described in the LEACHN guide.

CHEMICAL PROPERTIES

Chemical property data are specific to each version of LEACHM, since chemical species, processes and units vary. These notes describe the data required by LEACHE. Conceptually, LEACHE is similar to LEACHP, the pesticide fate model. LEACHE includes properties unique to per- and polyfluoroalkyl substances (PFASs), e.g. air-water interface sorption, but omits some of the sorption options included in LEACHP, such as Freundlich isotherms and kinetic sorption. Because concentrations of PFASs are usually very low and the compounds are non-volatile, both precipitate and gaseous phases were removed. The option of applying broadcast chemical is replaced by a binder application as a mitigating practise. The aim was to simplify the chemical section of the data file as far as possible, while including phenomena such as interface sorption. Some of the more complex sorption isotherms and kinetic processes can be re-introduced as our knowledge and understanding of PFAS in the soil environment improves.

Number of chemical species: At least one chemical must be specified. The maximum number is usually about 18 (defined in the code) but a practical limit would be about twelve. One reason for this is that a *.prn* file, if selected for saving the data file, can accommodate only 240 characters per row. Another is to limit the size of the output files.

Chemical properties are defined in the next sections of the data file. Chemical species are assigned a numerical ID, which is the sequential column number starting from the left-hand side.

The *Name* of the chemical must be enclosed in single quotes and be not more than 12 characters (including spaces). Using Excel's special format '@' for the chemical name cells simplifies this. This name will be used in output data tables.

Molecular weight (g/mol) is used in the calculation of air-water interface sorption and to calculate the relative mass of parent species loss and transformation species gain.

There are two sorption isotherm options. Option 1 is linear solid:aqueous phase sorption with no air-water interface sorption, and option 4 is linear solid:aqueous sorption as well as air-water interface sorption. All LEACHM models use the same flags to define sorption isotherms. Option 2 (Freundlich) and option 3 (Langmuir) are currently not implemented in LEACHE.

K_{oc} values are defined for solid:aqueous sorption. The listed K_{oc} values are multiplied by the organic carbon mass fraction at each depth in order to obtain a K_d value. Be aware that some online data bases list log K_{oc} values. The organic carbon fractions used are the *Organic carbon %* values specified earlier, divided by 100. The air-water interface sorption functions used in LEACHE (after Guo, et al., 2020, with units adapted for LEACHE) require two additional parameters, a (mmol/L), and b (dimensionless). If the sorption parameter values are related to soil properties other than organic carbon, e.g. clay content, the code can be modified.

Plant uptake 0(no), 1(yes): Chemical can be absorbed by plant roots in each soil segment, in accordance with that segment's transpiration weighting and the prevailing soil solution concentration. The concentration is assumed to be the actual solution concentration, i.e. the concentration in water, excluding chemical on the air-water interface.

TRANSFORMATION AND DEGRADATION

Chemical species may degrade into a product we do not wish to trace, and/or transform into a product we may wish to trace. Degradation and transformation may act simultaneously.

Transformation chains: If the Link flag is set to 1, the chemical species will transform to the species listed on the right-hand side of the parent compound in the data file. This species can in turn be the source of a third species, forming a transformation chain. The chain ends when the *Link* value is set to zero. For example, if chemical 1 is a parent compound of chemical 2, then the value of *Link* for chemical 1 is set to 1, otherwise, 0. A chain of transformation products can be set up in this way, for example chemical 4 could transform into chemical 5, which in turn transforms into chemical 6. This continues until the link flag, in this example for chemical 6, is set to zero.

Base transformation and degradation rate constants (d^{-1}) are the rate constants defined for a base temperature (at which the Q_{10}-based adjustment factor is 1) and an optimum water content. They can refer to the aqueous (non-interface) phase only, or to total chemical (all phases including interface). This provides flexibility; for example a 'transformation' pathway can be made sensitive to both temperature and water content while the simultaneous 'degradation' pathway could be sensitive to water content only. Such a combination may be used for example when temperature-sensitive biodegradation is operating at the same time as less sensitive chemical hydrolysis. Conceptually, both the transformation and degradation pathways follow Michaelis-Menten kinetics, which defaults to first-order kinetics if the *half-saturation constant* is set to 0.

The rate constants during a simulation are the base rate constants adjusted at the start of each simulation time step to reflect current water content and temperature. Currently, the relative response to water content and temperature changes is the same for all chemicals; again, the code could be modified to accommodate different responses. (This is why the adjustment factors are listed in a column rather than a row; it is then possible to define unique properties for each chemical species in adjacent columns.)

The temperature-based rate constant adjustment factor increases linearly from zero at the defined minimum temperature to a point of intersection of the linear and Q10 functions. The rate constants then increase exponentially to the optimum temperature, above which there is a linear decrease to zero at the maximum temperature. The adjustment factor is 1 at the base temperature.

The water-content based rate constant adjustment factor has a value of 1 at water contents within the optimum range of matric potentials. Below the optimum range they decrease linearly to zero at a water content corresponding to the defined minimum matric potential. Above the optimum range they decrease linearly to a defined relative rate at saturation.

There are options for using either the temperature-based rate constant adjustments or the water content-based adjustments. If both are selected the effective adjustment factor is the product of the two.

Diffusion coefficient

Diffusive transport is usually a small component of overall transport but may dominate aqueous phase transport during periods when convective transport is at or near zero. Diffusion coefficients should be set to zero when simulating no-flow batch processes (when each segment may be regarded as a separate batch experiment), or simulating column breakthrough if the simulated breakthrough curve is to be compared with analytical solutions which do not include diffusion. The aqueous diffusion coefficient value is adjusted for tortuosity and water content when calculating diffusivity. Although diffusion coefficients are related to the properties of a molecule or ion, only one value is specified and applied to all chemical species on the assumption that diffusive transport is usually a small component of transport in convective systems, so the effect of uncertainties in the diffusion coefficient are usually small. Large organic molecules have lower diffusion coefficients than, say, chloride, thus it is advisable to ensure the value is representative of the most important class of chemical species in the simulation. The units used for the diffusion coefficient in the data file are $mm^2\ day^{-1}$ for consistency with the rest of the data file.

If necessary, both the rate constant adjustments and diffusion coefficients could be specific for each chemical species.

INITIAL PROFILE CHEMICAL DATA

Starting concentration units: If *mg/kg* units are selected it is assumed that this is a total, which includes the solution, sorbed, and, if present, interface phases. At the start of the simulation the total chemical will be partitioned into each of these phases according to the starting bulk density, water content and organic carbon content, using the listed properties for each chemical species. If solution concentration units (*mg/L*) are specified, then it is assumed that non-interface solution will have the prescribed starting concentration. This reflects a situation in which an unsaturated column is in equilibrium with a constant-flux influent solution which has the defined starting concentration; Both the starting sorbed and interface concentrations will be in equilibrium with the starting solution concentration, but only the actual solution concentration, not the interface, will have the same concentration as the influent. If starting concentrations were specified as µg/L, then the listed values will be converted into mg/L immediately after they are read. LEACHE uses mg for all internal calculations.

Mixing cell: this is intended to mimic the effects of a saturated region below the unsaturated zone. It was included as a simple means of varying the lower boundary solute concentrations when a water table or constant matric potential is specified as a lower boundary condition, that is, when upward water flow or capillary rise is possible. Conceptually, the 'mixing cell' is a depth of water having the *starting concentration* specified. When leaching from the profile occurs, a new mass of chemical in the mixing cell is calculated by adding chemical leached from the profile and subtracting that leached by water flowing from the lower boundary of the cell. The new mean concentration is then assumed to apply to the mixing cell water (which remains at its original volume).

To remove the mixing cell set the depth of water to 0 mm. In this case the concentration at the lower boundary will remain at the specified value, so any water moving up into the profile from below (the 'water table' or constant potential lower boundary is the only option that allows for this) will be assumed to have this concentration.

BINDER APPLICATION

Broadcast applications of chemical are not included in LEACHE, but a single binder application is included. The binder application date, incorporation depth, the increase in effective soil organic carbon content and a nominal degradation period are defined in the input data. For example, if a chemical species having a K_{oc} value of 80 L/kg were added to a soil having an organic carbon content of 2%, the K_d value would be 1.6 L/kg. If a binder equivalent to 30% organic carbon is added, the effective K_d value would be (80 L/kg x 2/100) + (80 L/kg x 30/100), or 25.6 L/kg. Since binders are often physically mixed with a depth of soil, there is also an option for mixing soil (clay, silt, organic carbon, water and chemical species) from the surface to a defined

depth. This option is unique to LEACHE. Also, immediately prior to a binder application there is a print to the *.out* file, and another immediately after the binder application so that the effect of the binder (and mixing) on chemical partitioning and soil properties can be seen.

The *incorporation depth* defines the depth (mm) through which the binder is assumed to be mixed. The depth must lie within the soil profile depth. The amendment is assumed to be allocated uniformly from the surface to the segment closest to the specified incorporation depth. Immediately after the binder addition a new equilibrium between sorbed and solution phases is established.

The '*Soil mixing*' flag triggers an additional step which involves the mixing of all soil components uniformly through a specified soil depth. All chemical species, binder, clay, silt, sand, organic matter and water are summed, weighted by each segment's bulk density. All of these components are then distributed uniformly through the cultivation depth. The new mean bulk density is a value which does not lead to a change in soil depth, i.e. there is no compaction. Immediately after mixing a new equilibrium between sorbed, solution and precipitated species is established. Without mixing there are no changes to water content, soil physical properties or native organic matter content, so the binder is hypothetically distributed through, or 'injected' into a profile.

It is assumed that a binder can degrade over time, but the mechanisms and rates are uncertain. It is unlikely that binder degradation will follow first-order kinetics. Currently, a nominal degradation period (years) is defined, and it is assumed that a constant mass of binder will be lost each year during this period (i.e. zero-order degradation). For example, if the binder K_d is 30, and the degradation period 40 years, then the binder K_d value would decrease by 0.75 per year. Various options for adjusting the binder K_d over time are being tested. The simplest is to decrease the binder K_d at the start of each day. For this example the daily decrease of the binder K_d would be 2.0533×10^{-3} L/kg (0.75/365.25). This decrease would apply to each soil segment containing the uniformly distributed.

As with other transformation and degradation processes it is possible to adjust the daily loss according to current and/or antecedent water content and temperature conditions using current water content and temperature based degradation adjustment factors. The actual degradation period required to decrease the binder K_d to zero will depend on how these adjustments are calculated. The decrease in effective binder K_d will differ between segments because the adjustment values usually vary with depth. This option is included in the input data.

IRRIGATION

Irrigation and weather data have the same format for all versions of LEACHM. Irrigation is applied if automated irrigation was specified for any crop, or if irrigation events are scheduled at any time. This section consists of two parts: automated irrigation and or specified irrigation.

Automated irrigation for a crop is triggered if the crop's irrigation flag is > 0. Automated irrigation can be triggered only if a crop is present and during its specified irrigation period. The time of application is set in the code to be 0.05 day and extends over 0.2 day. If there is a rain event listed for that day, the triggered irrigation is postponed to the next day, if still required. These rules can be changed in AUTOIRRG.FOR if necessary. The *composition of the irrigation water* must be defined for the chemistry versions. The composition is the same for all triggered irrigations, unlike listed irrigations which can have a different composition for each irrigation. Triggered irrigation events are logged in the *.irr* file, which lists the time and amount of water applied.

There are four automated irrigation options:

1) If the soil matric potential (or matric plus osmotic potential) at the specified *sensor depth* drops below the defined threshold, irrigation water is added until the matric potential at the sensor depth rises to a defined value.

2) If the soil matric potential (or matric plus osmotic potential) at the specified *sensor depth* drops below the defined threshold, a depth of water required to replenish the water deficit (the difference between the current water content and the specified field capacity) between the surface and the *replenishment depth* is added in a single event.

3) This option is similar to 2), but instead of replenishment to field capacity, a specified *fixed amount per application* (mm) is used. This applies to irrigation systems such as centre-pivot systems, where a defined amount of irrigation water is applied.

For options 1 and 2, automated irrigation can be prevented, even if the irrigation flag is set, by setting the trigger threshold to a very negative value (e.g. -70000 kPa, which will never be reached. These three options are not realistic for capacity flow.

4) *Cumulative potential or reference ET* can be used to trigger irrigation. For this option irrigation is independent of soil type or soil water potentials. The amount and timing of irrigation will therefore be the same for both the Richards and capacity flow options. An incremental evaporation threshold is specified. At the start of a day when the threshold (including the current day's potential evapotranspiration) exceeds the threshold an amount of water corresponding to the threshold, multiplied by an adjustment factor will be applied.

Specified irrigation events are defined using the same format as used for rainfall. The composition of irrigation water for van be defined. Irrigation can be specified for any time, regardless of whether or not a crop is present. At least one specified irrigation must be entered but

the application date can be set past the end of the simulation (e.g. 999999) to avoid irrigation. If a specified irrigation event falls on the same day as a rain event and their application times overlap, then whichever occurs earliest has precedence and the other event starts immediately after the first event's end time. If scheduled for exactly the same time then irrigation is applied first. Although multiple rain and/or irrigation events can be scheduled for the same day this is unlikely if daily weather data are used to create the input rainfall data. If there are multiple rain or irrigation events scheduled for the same day then the data file should be edited to avoid overlapping.

Specified irrigation can be applied sub-surface. If the *Emitter depth* is set to zero, then irrigation is assumed to be applied on the surface. If the *Emitter depth* is >0, then the irrigation will be applied at the node corresponding most closely to the emitter depth. Currently, triggered irrigation cannot be applied sub-surface, since matric potential criteria that trigger irrigation may not be applicable.

If many regularly-spaced irrigation applications are desired, e.g. daily or weekly, then in place of the application date or day number enter the negative of the time interval between irrigations. For example, if -7 is entered, this irrigation amount and composition will be applied every seven days after the previous irrigation (or the start of the simulation if this is the first) to the date of the next specified irrigation (or the simulation end date if that precedes the next specified irrigation).

In the code, flags are associated with each type of water application to enable summation. These flags are 1) rain, 2) specified surface irrigation, 3) triggered irrigation, and 4) specified sub-surface irrigation.

RAIN AND RAINWATER COMPOSITION

Weather data is appended to the input data file in the form of two tables: I) rain events, including the chemical composition of rain water, and ii) weekly air temperatures, average daily temperature amplitude, and weekly and daily reference evapotranspiration. Data are listed in this format in order to reduce the size of the data file. Although weather data in the input data file can extend past the final date of the simulation, no records can precede the starting date.

The *number of rain events* is defined. These events can extend past the simulation end date. This number acts as a check, in case there are too few events listed, or if there are any typographical errors.

The *starting time*, *amount*, and *rate of application* of rain water is specified, one record per event. The number of rain records (rows) listed must correspond to the number defined. At least one rain event must be specified. Although rain derived from weather data is usually listed as a single event for a rain day, more than one event per day can be listed. Water application dates must be equal to or greater than the start date of the simulation and can extend beyond the end date. Start days and times must be entered in chronological order. The time of day can be specified to 0.001 day. If no rain applications are desired set the application time of the first event beyond the last day of simulation, or set the date to '999999'. If a value of 999999 is specified as

the ‘date’ of the first event, all records will be read but not used since 999999 is a date that is never encountered.

Although the format is the same, specified irrigation events should be specified in the table of irrigation events described above, and not included in the rain data table. The pragmatic reason for this is that it is important to preserve the integrity of what may be a very long rain data set, which may be used in subsequent simulations which do not have irrigation. Also, a flag value is linked to each water application event, so that the various categories of water application (rain, specified or triggered irrigation) and their dissolved chemical species can be listed and summed separately for output files.

The *rate of application* (mm/day) in conjunction with the *amount* (depth) of water applied (mm) determines the duration of the rain event. For example, if 20 mm rain falls at 100 mm/day, starting at 0.3 day, then the duration of the rain event will be 0.2 day and the event will end at 0.5 day. Any desired rate of application can be specified but the actual infiltration flux density may be limited by soil hydraulic gradients and conductivities in the surface segment. If, during infiltration, the specified rate of application exceeds the soil's infiltration capacity then the soil surface is assumed saturated and the remaining water infiltrates at a surface pressure potential of 0 kPa. Positive pressure potentials at the surface owing to ponding are not included. More than one rain event can be specified for the same day, so for example, many short applications can be specified, each having a different application rate, e.g. to simulate the varying rainfall rates during a single storm.

A utility (*Util_Weather.xlsm*) can be used to create weather data files in LEACHM’s format from daily weather station data. This utility assumes one rain event per day, since it reads daily weather data files.

If, at the start of an infiltration event, there is still water remaining to infiltrate from the previous event, the excess water is assigned to ‘infiltration-excess’ runoff and is not added to the new application. Under natural conditions and for most weather data sets, overlapping applications are rare. The model is structured in this way so that when simulating laboratory column experiments or the application of tracers in the field, the times at which influent chemical concentrations change can be defined unambiguously.

Ponded infiltration is defined by specifying 999.9 for the rate of application of water. In this case the soil surface matric potential will be saturated (0 kPa) and the infiltration rate will match the infiltrability of the soil until either all water has infiltrated or the start time of the next specified irrigation or rain event is reached. Currently, no provision is made for a depth of ponded water, i.e. a positive pressure potential at the soil surface.

REFERENCE ET AND TEMPERATURE

Data in the correct format for the weather section of the data file (both rain, ET and temperature) can be generated from daily weather station data by the stand-alone *Util_Weather.xlsm* utility.

There are two control values at the start of the weather data table:

I) A flag which defines whether *weekly reference ET* or *daily reference ET* will be used.

ii) A value for *starting uniform temperature profile* sets a starting uniform soil profile temperature, which will also be used as the lower boundary (at a hypothetical depth of 2 m) for the duration of the simulation. Alternatively, entering a value of 999 means that the average of all temperature data values within the simulation time period will be used as the starting profile temperature, and also be used as the lower boundary temperature for the duration of the simulation. Be aware that the average temperature will depend on the duration of the simulation. For example, a 40-year simulation will calculate the average temperature from 40 years of temperature data, but if the same simulation is run for say, one month, then the starting profile and lower boundary temperature will be that during the first four or five weeks. To avoid unintentional inconsistencies, a good approach is to determine the long term average over the longest simulation time expected and specify this as the starting uniform temperature profile. The long term average is easily determined in the Excel input data file.

In regions where winter air temperatures drop well below freezing, soil temperatures are unlikely to fall to the same level, owing to issues such as latent heat of freezing, snow cover etc. For these reasons the air temperatures used in the weather data utility to calculate average temperatures are, if less than -3 C, raised to that value. These temperatures are used as the surface boundary condition in the heat transport routine and are not used elsewhere as air temperatures *per se*.

For steady-state column simulations the temperature will be set to the defined *starting uniform temperature profile* value and remain constant over time.

The table of ET_{ref} and temperature data contains one record (or row) of data per week. Each record consists of:

Week no., day or date: This can be the week number, day number, or calendar date, whichever is most convenient. This value is read but not used as it is assumed that all weekly data starts at the specified starting date, i.e. the first weekly ET record is the average reference ET for week 1, or days 1 to 7.

Reference (or potential) evaporation (weekly totals): If the flag at the start of the weather section is set to read daily reference ET values, then both weekly and daily values will be read, but the daily values will be used in the simulation. If the flag is set for weekly totals then the weekly totals are used and the daily reference ET used in the model will be one-seventh of the weekly total. The daily reference ET will thus be the same for each day of the week. If daily reference ET values were selected, then the value for each day of the week is listed. If the Util_Weather.xlsm utility is used to process weather data source files, then the listed weekly value will be the sum of the seven daily values. However, the weekly values can be changed, providing a convenient and simple way of comparing other ET_{ref} values without having to edit the daily values. For example, it may be useful to have the weekly value set to a constant 35 mm, or 5 mm per day. Or they can be set to zero, which switches off evapotranspiration. The number of entries in the reference evaporation table must be greater than the number of weeks simulated to allow for a partial week at the end of the simulation, e.g. a year requires 53 weeks of data to encompass 365 days, which is 52.14 weeks.

Mean weekly temperature and daily amplitude: Air temperature is used only to adjust temperature-dependent rate constants, and serves as the upper boundary condition for the heat flow routine. Daily means and amplitude are obtained during a simulation by linearly interpolation between weekly means and amplitudes. Since equating the soil surface temperature boundary condition to air temperature is an approximation, and the only application of temperature data is adjusting somewhat uncertain degradation and transformation rate constants, this is probably sufficiently accurate.