**Data Requirements for the Daisy model**

Daisy requires data about the climate, the crop, the agricultural management and the soil in order to run a simulation. Daisy is to some degree able to adjust to the available data, either by using simpler models internally when less data is available, or by trying to synthesize the missing data from what is available. In the following, we will describe the minimum data requirements for running Daisy at all, the data we recommend for reasonable simulation results, and finally list some of the optional data that you can feed the model for an even more detailed simulation.

Generally, the data requirement level depends on the **purpose of your simulation** **and the required level of details in the model results.** Before starting on the modelling itself, it is useful to try to conceptualize ones understanding of the problem at hand. With your own knowledge of the problem you are studying, what is the level of detail required, what are the main factors of importance, and what are the boundaries of the system you are analysing?

Generally, local measurements improve the model description of the studied area. For some input parameters, like precipitation, local measurements are particularly important, because local variability is large and the influence on the simulation results is significant.

More information about the setup of data in the model can be found in the Daisy tutorial (in the daisy program folder “doc”)

# Climate data requirement,

## The minimal level

Daisy will as a minimum need **daily values** for the following item for the entire simulation period, including the initialization period (see section 3.4).

* Average air temperature.
* Precipitation. (local data important)
* Global radiation.

This will allow the model to use the Makking equation or the Hargreaves-equation for estimation of evapotranspiration and the global radiation will govern the photosynthesis in the model.

Daisy will also need information about the weather station itself, in particular

* the name (for reference)
* elevation of the weather station (used for some of the radiation/energy calculations)
* location: longitude and latitude (used for some of the radiation/energy calculations)
* information about the whether the data is recorded at the field or on a reference surface (may affect reference evaporation)
* the screen height (influences the Penman Monteith evaporation calculation),
* wet and dry deposition of NH4 and NO3 (usually estimated based on national database-information)
* average temperature for the location (can be derived based on long-term records from the area and is used to calculate the lower boundary condition for the calculation of soil temperature)
* the average amplitude of the temperature variation over the year (can be derived as above) , and
* the average day where the maximum temperature is reached (can be derived as above).

All of this information should be supplied in a special climate file, whose format is described in Appendix A of Daisy Program Reference Manual.

## Missing data

All the above data must be supplied to the program in order for the simulation to run. Use the best data available, which means data from the closest available climate station. In some case you will have to supplement local recordings with regional records, **precipitation is the most important to get local numbers for**.

## Predictions

If you are using Daisy for making predictions, you won’t have any climate data. In that case, we recommend that you run the simulation with old data from a number of different years. Be sure to permute the years, as the results are very sensitive to variations not only in the weather, but in the sequence the years appear. Do not use “average” weather data, as that tends to produce more evenly spread rainfall and therefore higher yields and less macropore flow by omitting the extreme events within the specific weather year as well as the yearly variation. For the same reason, be sure that the years you use contain some examples of extreme weather, not just weather from “normal” years.

## Additional climate data

Data for the reference evaporation can be supplied directly in the weather file.

Supplying **hourly values** **or finer** in the climate file will enable a much more detailed simulation. If you furthermore supply information about

* vapor pressure,
* relative humidity and
* wind speed,

Daisy will be able to use the Penman-Monteith equation for evaporation-calculations and select a more advanced bioclimate model.

# Soil data

In Daisy, the column is a one or two dimensional description of an agricultural system, with the weather at the top and the groundwater at the bottom. A soil column consists of soil horizons. A soil horizon is defined as a vertical layer of soil with similar chemical and physical properties. If you make a vertical cut in the soil, there is usually a clear visual distinction between the horizons. This division should determine the number of horizons to be defined, their order and their depth..

## The minimal level

Daisy will need the following information about the soil.

* The soil texture, i.e. the clay, silt and sand fractions of the soil. Daisy needs this information for all horizons in the profile, not just the top horizon.
* The humus content and C/N-ratio in each horizon.
* The maximum root depth of the soil. (soil parameter, not crop-dependent).
* The location of the groundwater and whether soil is drained. There are the following possibilities:

1. The soil is drained. Then we need to know the position of the drains (depth and distance between drains). In this case information about the drainage period or a time series of drain flow, preferably both, is useful. This will be used to calibrate the drainage information.
2. The groundwater is located below 4 meter. If so, we don’t need to know more.
3. The groundwater is located above 4 meter. If so, we would like to know where it is located and better a whole time series of the groundwater table.

If this is all that is available, Daisy will calculate hydraulic functions based on a pedotransfer function (Cosby). If, in addition, the **bulk density** is known, the HYPRES pedotransfer-function is used. For many types of simulations, we have good experience with using HYPRES-generated data for the hydraulic functions.

## Missing data

The texture is very important for water transport, if you can’t measure it, you must use your best estimate. The B horizon (and sometimes the C horizon) may in fact be even more important than the Ap horizon, and more difficult to get information on. Sometimes regional maps are available, or measurements from nearby sites.

Good numbers for the humus content of the top horizon is essential for a reasonable nitrogen balance, the lower horizons are much less important.

The maximum root depth can usually be estimated from knowledge from similar soil profiles.

## Additional Soil data

A key problem in modelling the soil/plant-system is the initialization of the organic pools in the soil. Daisy contains pools describing slow and fast turn-over of added organic matter, soil microbes and soil organic matter – or no turnover of organic matter at all. The size of the pools cannot be measured, so it necessary to estimate their size. The total carbon content and organic nitrogen is considered the total “pool size”.

* Bulk density is a good improvement of predicted hydraulic properties and some of the sorption processes.
* An estimate **for the average yearly carbon input to the soil** in the decades before the simulation period. The yearly carbon input is rarely available; you can use numbers from a predefined table of input from various farm types if someone has created such a table for your area, or better yet run a simulation with a “typical” historical crop rotation to let Daisy calculate the input. This information is important in relation to initializing the organic pools in the soil.
* If possible, you should measure the **nitrogen yield from a plot with a non-fertilized crop**. This provides an indicator of the mineralization of the soil and can be used for model calibration (described in Daisy Tutorial).
* You should measure the C/N ratio for the humus in at least the Ap horizon. The C/N-ratio influences the distribution between organic pools and thus the mineralization. Soils with a high C/N-value are likely to contain a fraction of organic C and N that should be characterized as “SOM3” which is inert.
* You may improve your simulations by measuring the hydraulic properties of the soil. This can be done at different levels of detail:

1. Measurement of (selected) points on a retention curve at different suction (field capacity, wilting point, perhaps other points and saturation based on the bulk density measurement), which may be used to calibrate parameters for the most common hydraulic functions (Brooks and Corey, Campbell, van Genuchten), or – if there are enough points, the values can be used directly.
2. Measurement of saturated hydraulic conductivity,
3. Measurement of unsaturated hydraulic conductivity at different suction, which again can be used for fitting common hydraulic functions (Burdine, Mualem), or can be used directly.

Additional information on carbon turn-over improves simulations considerably. If you have few measurements of hydraulic properties, the experience is that it may not yield better results than using a pedo-transfer function. This is due to high spatial variability of these parameters. The purpose of the simulations is important here – if you are looking very detailed at effects of tillage on transport of solutes, these parameters have to be monitored closely.

In detailed studies, the above information can be supplemented with measurements of:

* moisture content in the soil over time and in different depths,
* drain flow
* nitrate content in soil moisture (soil samples, suction cups)
* nitrate concentration in drain flow.

Such information can be used for calibration.

# Management

## The minimal level

Daisy will need to know the date of the following management operations:

* **Fertilizing:** In addition to the date, we need to know the amount and type of the fertilizer used, and whether it was incorporated or not.
* **Irrigation:** We need also to know the amount of water applied, how it was applied (surface, overhead or subsoil), as well as the amount of mineral fertilizer in the water, if any.
* **Tillage operations:** We need to know the type of operation, e.g. plowing or seed bed preparation.
* **Sowing:** We need to know the type of crop and the selected crop must exist in the Daisy crop library with the required level of parameterisation.
* **Harvesting:** We need to know whether residuals are removed or left on the field.

All this information should be given in the Daisy setup files, or .dai files, in the format described

in Daisy Tutorial, especially section 5.

## Missing data

Daisy needs the above information in order to run. If you don’t have it, you must make estimates. The dates can be essential, for example, if the irrigation date does not match the time the crop need the water the most, the yield may be much lower. You may then actually want to use the techniques for predictions described in the next section instead.

* **Fertilizing**: If you are interested in nitrogen harvest or leaching, it is essential to get the input right. The same is true for yield and to a lesser degree water balance if the crop growth is nitrogen limited. For organic fertilizer types, use the closest match of the already defined fertilizer types from the ‘fertilizer.dai’ file.
* **Irrigation:** Get it right or use “conditions” to let Daisy select the right time for irrigation.
* **Tillage operations:** Use the closest match of pre-defined tillage operations from the ‘tillage.dai’ file.
* **Sowing:** Daisy need a lot of information about the crop, this information have been assembled for a number of crops, which can be found in the ‘crops.dai’ file. If the actual crop is not found there, you may try to substitute with another crop with a similar development cycle and yield, if any such crop exists.
* **Harvesting:** Use your best estimate of harvesting time, stubble height, and fate of residuals.

## Predictions

Daisy allows the manager to “be smart” about the state of the simulation, so instead of using fixed dates it can harvest when the crop is ripe and irrigate when the crop needs water. You can use these techniques to implement your management plan in the simulation. As hinted earlier, this is also useful when you don’t have information about exact dates. This is particular useful when you analyze large areas, where you do not has information from individual field. From statistical information about preferred crop types and knowledge about common praxis in the area, you can create representative crop rotations and run these under various climate conditions, and on the soil types found in the area.

## Additional Management information

* **Initialisation:** As mentioned in section 2.3, the carbon and nitrogen balance of the soil is influenced by what happened earlier on the same plot. We recommend that simulations start *with a “warm up” period of at least 5 years* before the period of interest. You don’t need exact dates for management operations for the warm up period (although having them helps), but can use the techniques described in section 3.3. If no weather data are available for this period, you may have to re-use the data you do have, as described in section 1.3. This may not be correct, but it may still provide better starting conditions for your simulation that running without an initialization period.
* **Fertilizing**: For fertilization, you should specify the amount of carbon, nitrogen and ammonium directly, rather than rely on the values from fertilizer.dai
* **Fertilizing:** You can examine the exact content of organic fertilizer, which tend to vary a lot between farms, and you can estimate the turnover rate with laboratory experiments.
* **Tillage operations:** You can fit the tillage operations to the machinery used at the actual farm, and make sure e.g. the plowing depth is right. Tillage implements are defined by how far down the soil is mixed and the fraction of organic matter on the surface that is incorporated in the soil by this operation.
* **Harvesting**: Measure or estimate for how large a fraction of the various crop parts are left on the field, rather than assume all or nothing.

# Crop data

## The minimal level

At the minimal level, the existing Daisy crop parameter files are used without calibration. The model will function with the information from the Daisy library on which crop is grown. It is evident that such an uncalibrated simulation does not consider the local variety grown.

## Additional crop data

The crop parameterization can be adjusted to local conditions in various ways, depending on the purpose of the simulation.

* **Yield information:** For some simulations, statistical yield data may be useful, but they do not describe the conditions on a given field. Local information of yield may be useful when simulating leaching, because it includes the effects of other liming factors. In general, yield information can be used for either calibration or validation.
* **Biomass development over time**: A simple improvement of the simulation can be obtained by measuring the biomass development over time, and noting the most important growth stages. This can be used for calibration of the crop module (-or validation).
* **LAI**-**values:** can be used directly as input to Daisy, for calibration and for assimilation in Daisy. It can also be used for verification of the model calculations.
* **Root depth of the plant**: Potential root depth for the crop is given as parameter in the crop file. If better local knowledge is available, this can be included in the simulation.

For more detailed calibration of the crop modules or creation of new crop modules, even more detailed information in the form of field experiments with various levels of irrigation and fertilization can be required. The interesting measurements would typically be

* **Dry matter content of the individual crop parts:** (leafs, stem, storage organ, and if you are ambitious, roots) at regular intervals throughout the growth season. Measured RVI can support the calibration exercise.
* **N content of the individual crop parts:** (leafs, stem, storage organ, and if you are ambitious, roots) at regular intervals throughout the growth season.
* **LAI**-**values:** at regular intervals throughout the growth season.
* **Crop height** is used in some process descriptions together with **canopy cover**. This could also be measured at regular intervals.

You can supply information about application of pesticides. Pesticides does not affect the rest of the model (as pests are not part of the model), but Daisy will simulate the fate of the pesticides.

# Other data for calibration or validation purposes

This group of data is rather specific for the type of studies conducted. Particularly for process studies, very specific and detailed values for selected parameters have to be collected. However, it is not uncommon that the following measurements are carried out:

* Moisture content in the soil in a number of depths is measured using e.g. TDR
* Drain flow (Time series)
* Nitrate or pesticide concentrations in drain flow (Time series)
* Nitrate (and pesticide) concentration in water extracted by suction cups. It should be noted that there is a high degree of in field variability on such measurements. This type of measurement is not suited for all pesticides (Several measurements over time).
* Nitrate or pesticide concentration in shallow groundwater (Time series).
* In connection with pesticide studies, particle concentration in drain water may be of interest.