Should the hierarchy be model (crop), data or scenario based? We do need some method to organize data independent of the model. The description table would tie the model to the data. To do this (tie the model run to the data) we may have to make use of filenames more in the table or have a generic automatic file naming system (01, 02 etc). We did this for the glycim data miner. The initialization file would also have to be model specific (I think)

Anyway, I think the hierarchy should be data based but model specific. The data classes should be arranged around the way we are structuring our research – G x M x E

The description table should define the scenario.

Fields:

ID, File names, path to files, geographic location, soil, variety (hybrid), climateID, Location, nitrogen, solute and biology defaults

Some records can be identified by ID, others by name. Items that do not have a lot of variance like hybrids, soils, solute, biology can be identified by name. If we need to create a large number of variations we can give them numbers

Here is my first thought at the hierarchy

2DSOIL (intrinsic?)

gridRatio (need to move planting depth info - needed for root growth in potato)

GridX (depends on row spacing)

Management (M)

irrigation (amount and time)

Planting Time (time)

Fertilization (amount and time)

Soil (properties)

Biology (properties)

Initials (population, dates, autoirrigation (move to irrigation), seed depth, CEC (move to variety) Lat long (move to location - climateID), CO2, outputfile flags

Plant (G)

variety

Root info (move N uptake kinetics to the nitrogen file?)

Environment (E)

climate (not sure what to do with some properties like furrow irrigation (move to irrigation file?

weather

Soil

Organic matter related (Biology and some parameters in the soil table)

Solute

In the root info, under the variety table are some parameters for N uptake kinetics, these seem like they would be variety but may be environmental - Organic matter should be the environment.

There are some data that are strongly scenario dependent and others that are useful over scenarios. The 2DSOIL grid data falls in between - it is dependent on the scenario (row distances) and the soil (depth increments and grid spacing)

Initials is somewhat amorphous. Some of it is management (row spacing, population) some is model dependent (flags for output file generation) some is time dependent (timing) which should probably go into management since it is dependent on management. Lat long might be considered environment - location. Initials is probably just a specific type of management file. It is specific to a scenario.

Have to think about reengineering the initials file. It seems like some of this can be in the description table maybe? The idea of an initials file and management files are related to how we want to use the files. It is all based on a scenario and how we want to vary scenarios.