The LHS/Meta21 Workflow

TBD*

March 31, 2023

Abstract

The brief document describes the workflow for running the combination of the LHS utility with the GAMS-based version of the Meta21 model.

1 The LHS utility

The LHS utility is provided as a Windows-based 'exe' file, `LHS.exe`.[1] The file `Meta21.in` is an input file for the LHS utility, which contains all of the sampling assumptions for the Meta21 model—including the assumed distributions with their parameterization, and for a number of the uncertain parameters it also contains the targeted correlations. In all, there are some 582 uncertain parameters.

The utility produces 1 to 3 output files—depending on the user options.[2] It creates the sample file in one of the five available output formats. If the user has requested the sample histograms (using the option `REPTS HIST`) the utility produces a 'CSV' file, which has the data that can be used to render the frequency histograms for each of the uncertain parameters.[3] The usefulness of the histograms in the listing file is limited with large sample sizes. The third file is produced only if the GDX file is requested. This third file, with the same name as the sample file, but prefixed with `Read`, contains the GAMS code that describes the sample and will read the corresponding GDX file. It is 'included' in the code file `Meta21.gms` to initialize the sample for the Monte Carlo simulations.

2 Data files

Most of the core data for the model is directly embedded in `Meta21.gms`—one of the core model files. Larger data files have been extracted from the relevant Excel files and stored as GDX-formatted GAMS files. Table 1 provides a full list of the external data files. The parameter files are loaded from the GAMS file `Meta21.gms`. The RCP and SSP data are loaded from `Meta21Header.gms`, which is a GAMS file that contains most of the key set definitions for the model.

The SSP database is constructed from a relatively aggregate source database for the SSP data. The data has been extracted for the five broad SSP regions (Asia, Latin America, Middle East

*TBD

[1] The code is available upon request and will eventually be posted to Github. It should be readily ported to Linux and macOS.

[2] It also creates two diagnostic files. The first is the 'listing' file which contains a description of the user-specified sample and optional output such as the sample data, the sample histograms and the sample correlation matrix. The second is the log file, which is mainly used for debugging purposes.

[3] A sample Python program to render the histograms is included in the *LHS Utility User Guide*.

and Africa, OECD, and the reforming economies), and averaged over six reporting models (AIM, GCAM, IMAGE, MESSAGE, REMIND and WITCH).[4] Data for each country is initialized from the World Bank WDI database for the 2010 reference year (in 2010 prices and purchasing power parity exchange rates). Growth from 2010 through 2100 uses the relevant growth rate from the aggregate SSP database, where each country is mapped to one of the five aggregate regions. Additional assumptions are made to generate SSP profiles post-2100.[5]

3 Model files

Table 2 lists the code files used to run the core Meta21 file and the Monte Carlo simulations. The following section will provide a concise description of the workflow.

4 Workflow

4.1 Sequential

The workflow is relatively simple: (1) run the LHS utility to produce the LHS cube; (2) run `Meta21.gms` to setup the model for Monte Carlo simulations; and (3) run `Meta21Stoch.gms` for the Monte Carlo simulations. Steps (2) and (3) can be run for different configurations of the model. The current setup has two configurations: (1) run with the five tipping points; and (2) run without any of the tipping points. This is driven by command line option called `sim` which is set to either `wTP` or `xTP`. In advance, the user should set up folders for the output files.[6] The output folders are defined on lines 50–56 of the `Meta21.gms` file. The file `workflowseq.cmd` summarizes all of these steps in a single Windows command script file. N.B. The full suite of simulations takes from 8-10 hours. The sequential process will have created a single output file for each configuration in the output folder and file name—designated by the user. Both are set in the file `Meta21.gms`, lines 50–56.

4.2 Parallel

The file `Metaqueue.py` uses Python's Queue and multi-processing (MP) capabilities to run a single suite of Monte Carlo simulations in parallel. Instead of the 4 hours of time for the sequential process, the parallel process takes less than 1 hour. The initial step still requires running the LHS utility. The Python code takes care of the other two steps: (1) running the model initialization code, `Meta21.gms`; and (2) running the Monte Carlo simulations. The latter are run using an iterative process in Python with a queue of size 4. When one simulation is done, another enters the queue. The program is not fully bullet proof and will need some additional modifications. First, it does not appear to work with command line arguments. This means that the user will have to modify by hand the simulation code in the Python file. This is done on line 9. To run the first suite with tipping points, set `sim` equal to `wTP`. For the second suite, without tipping points, set `sim` equal to `xTP` Second, the Python script appears to hang after finishing all 10,000 simulations.

[4] The demographic projections are available at the country level and were expected to be harmonized across all models—models could differ nonetheless in country coverage and/or reference year. Similarly, country-based GDP estimates are provided by the OECD and IIASA—see for example [?].

[5] The procedure is described in section 2.4 of the supplemental materials for [?].

[6] With the increasing use of synchronization software such as *OneDrive*, it can be advantageous to store output, at least temporarily, on a non-synchronized drive. This reduces time uploading new versions to the cloud, and storage needs, as versioning is limited.

File name	Description
AMOC_cPARMS.gdx	This file contains the country-specific 'hosing' rates for each of four GCMs labeled 'Hadley', 'BCM', 'IPSL' and 'HADCM'. This parameter, which is not subject to sampling, appears as parameter T^h in equation (??).
cTemp_Parms.gdx	This file contains the country-specific downscaling parameters, which converts the global mean surface temperature to country-specific temperature. This refers to parameters α^g and β^g in equation (??).
cTemp1990.gdx	This file contains the average temperature per country in 1990, in levels in °C. It is the reference year for the temperature-based damage function, see parameter $TL_{c,1990}$ in equation (??). N.B. The original Meta21 package has two versions of the 1990 temperatures. One is in the main Meta21 Excel file. The other, with minor variations is in the 'damage' workbook (BHM betas v3pd.xlsx). The former is used in the simulations.
BetaParm.gdx	This file contains the central parameter values for the damage functions. The parameter Beta_Parm has the temperature-based damage coefficients corresponding to parameters β^1 and β^2 in equation (??), respectively the linear and quadratic coefficients. These are subject to sampling. The parameter SLR_Theta corresponds to the (linear) damage coefficient for sea-level rise— parameter χ^{SLR} in equation (??). It is also subject to sampling.
savrat.gdx	This file contains the country-specific savings rates. These are based on the average savings rates between 2005 and 2015 from the World Bank's WDI database (using indicator NY.GNS.ICTR.ZS). They are constant over time, though time-indexed in the GAMS model code.
RCP_Inputs.gdx	This file contains the exogenous emission and concentration variables for four RCPs. It has CO_2, methane and sulfur emissions, a single exogenous forcing for all gases except for CO_2 and methane, and concentrations for methane and N_2O, the latter which is used to model methane/N_2O interactions in forcing.
SSPDatabase.gdx	This file contains the SSP assumptions for population and GDP. The data is for four SSPs: SSP1, SSP2, SSP4 and SSP5.

Table 1: External data files

File name	Description
`Meta21.gms`	Core simulation file. This file contains the code for parameter and model initialization. It reads `Meta21Header.gms` which has the definitions of most of the key sets. The file is standalone in the sense that it can run the Meta21 model with a given set of parameter values. In most cases, it is used to setup the Monte Carlo simulations. As such, it reads the LHS cube and initializes the seed for the random number generator for each observation. Under normal usage, the file is invoked with the 'save' option as it will be used as the starting point for the Monte Carlo simulations.
`Meta21Header.gms`	File containing the core set definitions. In addition, it will prepare the mapping file, `LHSampleMAP.gms`, which is used to map the country specific parameters in the sample with the model parameters.
`Meta21Mode.gms`	This file contains the model specification. The model is fully recursive and thus can be run as a stand-alone process without the need to invoke a solver.
`Meta21Stoch.gms`	This file is the main driver for the Monte Carlo simulations. It assumes a given starting point, using GAMS' 'restart' feature. The starting point will have been created when invoking the `Meta21.gms` code. It loops over the entire sample: (1) initializes the uncertain parameters and possibly some reference year variables; (2) runs the reference simulation; and (3) runs the simulation with the additional pulse of CO_2 emissions. It will calculate the SCC and save selected results. The code has been designed to run all observations in the sample sequentially, or one at a time. The latter has been used to parallelize the Monte Carlo simulations. Because the random number generator seed is observation specific, the results are intended to be identical whether the Monte Carlo simulations are run sequentially or in parallel.
`initParm.gms`	This file is invoked at the start of each Monte Carlo simulation to initialize the uncertain parameters and re-initialize some variables in the reference year.
`Collect.gms`	This file (and `Collect2.gms`) is used to merge the results when the Monte Carlo simulations are run in parallel.

Table 2: Model files

If the output files for all 10,000 simulations are present, the code can be stopped by closing the terminal window.

The parallel process creates an output file for each observation, designated by the observation index, for example 'S137.gdx'. The file `collect.gms`, together with `collect2.gms`, is used to merge all of the results into a single file.