WP / 23 / 95

GTAP Working Paper No. 95

The META 21 Integrated Assessment Model in GAMS and LHS Sampling

by Dominique van der Mensbrugghe

Global Trade Analysis Project

Center for Global Trade Analysis
Department of Agricultural Economics
Purdue University

First edition: May 2023

Please cite using the following reference:
van der Mensbrugghe, Dominique, 2023.
The META 21 Integrated Assessment Model in GAMS and LHS Sampling
GTAP Working Paper, WP/23/95
Center for Global Trade Analysis, Purdue University
West Lafayette, IN
https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=7036

Center for Global Trade Analysis
Department of Agricultural Economics
Purdue University
West Lafayette, IN 47907
https://www.gtap.agecon.purdue.edu

The META 21 Integrated Assessment Model in GAMS and LHS Sampling

Dominique van der Mensbrugghe[†]

May 8, 2023

Abstract: The META 21 Integrated Assessment Model (Dietz et al., 2021) represents a fairly comprehensive climate change simulation model incorporating a number of features: (a) a recent simple climate model; (b) down-scaling of temperature change to the country level; (c) integration of a number of bio-physical tipping points linked to rising temperatures; and (d) economic damages linked to rising sea levels and temperature using recent country-level estimates from the literature. The original implementation of META 21 was in Excel and linked to the @Risk Excel add-in for performing Monte Carlo-type analysis. This paper reflects a translation of META 21 into GAMS. One of the key purposes is to allow for ready incorporation of many of the features of META 21 into other models—notably CGE-based IAMs. To replace the features provided by the @Risk Excel add-in, this paper also introduces a software package that produces random deviates—and, similar to @Risk, uses the Latin Hypercube Sampling (LHS) approach, which is a stratified sampling technique intended to cover the entire sampling space efficiently and, in addition, orders the sample to reflect a desired correlation matrix for the sampled random variables.

JEL Codes: C15, C53, C63, Q54
Keywords: Integrated assessment models, climate change, tipping points, extreme events, Monte Carlo simulations, Latin hypercube sampling

[†] The Center for Global Trade Analysis, Purdue University, 403 Mitch Daniels Boulevard, IN 47906. Corresponding e-mail: vandermd@purdue.edu.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The META 21 Integrated Assessment Model (Dietz et al., 2021) represents a fairly comprehensive climate change simulation model incorporating a number of features: (a) a recent simple climate model; (b) down-scaling of temperature change to the country level; (c) integration of a number of bio-physical tipping points linked to rising temperatures; and (d) economic damages linked to rising sea levels and temperature using recent country-level estimates from the literature. The original implementation of META 21 was in Excel and linked to the @Risk Excel add-in for performing Monte Carlo-type analysis. This paper reflects a translation of META 21 into GAMS. One of the key purposes is to allow for ready incorporation of many of the features of META 21 into other models—notably CGE-based IAMs. To replace the features provided by the @Risk Excel add-in, this paper also introduces a software package that produces random deviates—and, similar to @Risk, uses the Latin Hypercube Sampling (LHS) approach, which is a stratified sampling technique intended to cover the entire sampling space efficiently and, in addition, orders the sample to reflect a desired correlation matrix for the sampled random variables.

The next section provides a summary description of the META 21 specification. Interested readers are referred to the supplementary materials of Dietz et al. (2021) for a more complete description of the specification and the underlying sampling strategy. The subsequent section describes the sampling strategy for the uncertain parameters. A brief results section describes some results from the Monte Carlo analysis—focusing in particular on the effects of including the tipping points in the simulations. A brief description of the LHS algorithm is provided in the Annex. A summary guide to the LHS utility is provided in van der Mensbrugghe (2023a). The Monte Carlo simulations were run in parallel on a desktop computer using a Python-based script described in van der Mensbrugghe (2023b). The parallel processing decreased the total time for the 10,000 simulations from around 3 hours to 25-30 minutes.

Chapter 2

Model Specification

Integrated assessment models (IAMs) of climate change couple economic models that track GHG emissions with bio-physical models that at a minimum generate changes in the global mean temperature, which subsequently impact the economy via changes in economic potential such as crop yields, infrastructure, labor productivity, etc. Figure 1 provides a schema for the typical IAM model, which has four components: (1) the socio-economic module, which could be a CGE model, such as GTAP (Corong et al., 2017) or ENVISAGE (van der Mensbrugghe, 2019); (2) which generates emissions; (3) that impact atmospheric chemistry and the global energy balance, i.e., temperature; and (4) that impacts economic potential. In the context of ongoing research under the aegis of the Intergovernmental Panel on Climate Change (IPCC), most IAMs are using drivers and storylines called the shared socio-economic pathways (SSPs)—a set of five distinct storylines with variegated pathways for GDP and population growth.[1] The SSPs are coupled with the so-called Representative Concentration Pathways (RCPs)—that represent different pathways for climate, with changes in the global mean temperature as one of the indicators.[2] The system will also be influenced by policies—mitigation polices that will change the emission pathways, and/or adaptation policies that will modify climate-induced impacts.

The subsequent sections will describe one such IAM, the META 21 model (Dietz et al., 2021) and will be used to illustrate the coupling of the IAM with the LHS utility. The META 21 model is in the same spirit as Nordhaus' DICE/RICE suite of models (Nordhaus, 2017), Hope's PAGE model (Hope and Schaefer, 2016) and the FUND model (Anthoff and Tol, 2014).[3] Like these models, META 21 links emissions to the carbon cycle, to forcing and temperature change and then to economic damages. The economic and emissions module of the META 21 model are highly simplified and are largely exogenous in fact. GDP is driven by exogenous assumptions on total factor productivity (TFP) and savings and is only impacted by climate-induced changes in TFP. Similarly, emissions are largely exogenous—not even responding to changes in GDP—though impacted by some of the modeled tipping points such as the release of methane from the oceans.

The focus of META 21 is on potential bio-physical tipping points, which include:

- The thawing of permafrost (PFC), which leads to additional carbon and methane emissions
- The dissolution of methane hydrates (OMH) which leads to additional methane emission
- The Amazon dieback (AMAZ), which leads to a jump in carbon emissions
- Ice sheet melting, which impacts sea-level rise. There are four sources: thermal expansion, small ice-sheets and glaciers, the Greenland ice-sheet (GIS) and the West Antarctic ice-sheet (WAIS)
- Changes in the Atlantic Meridional Overturning Circulation (AMOC), also called the thermohaline circulation. AMOC affects the change in country-level average temperatures
- The weakening of the Indian summer monsoon (ISM), which leads to an additional damage effect for India.

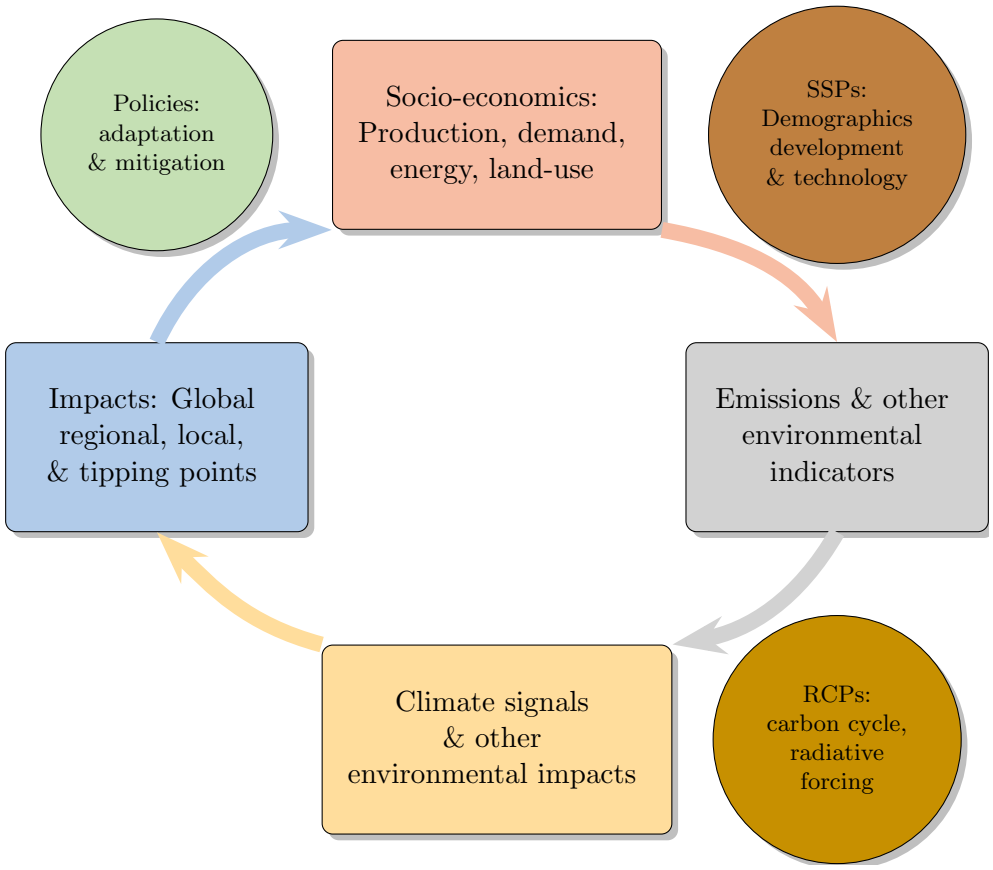
Other notable features of META 21 include: (1) incorporation of a relatively recent simple climate model known as FaIR (Millar et al., 2017); (2) down-scaling of global mean surface temperature change to the country level; (3) country-level damage estimates from changing temperatures and sea-level rise using recent estimates; and (4)

[1] O'Neill et al. (2017) provides a description of the SSP narratives. The quantification in terms of population and GDP pathways for the 5 SSPs are described in KC and Lutz (2017) and Dellink et al. (2017), respectively.

[2] See for example van Vuuren et al. (2011).

[3] Note that these three IAMs were the basis of the U.S. government's initial report on the so-called social cost of carbon (SCC), (Interagency Working Group on Social Cost of Carbon, 2016).

Figure 1: **Integrated Assessment**



calculation of the social cost of carbon using uncertain parameters such as the discount rate and pure rate of time preference.

The following sections are ordered using the IAM circularity concept, starting with emissions, flowing through the carbon cycle and changes in temperature, and impacts in the form of tipping points and changes in economic potential. Note that the original model, implemented in Excel and linked to the @Risk add-in for the Monte Carlo simulations, was structured recursively and thus bypassed the need for a model solver. The GAMS implementation follows the same recursive structure and the model is formulated as a sequence of mathematical expressions with no solver involved.

2.1 The Emission/Climate Module

2.1.1 Carbon emissions

Global carbon emissions in any given year are composed of four items in the META 21 model:

- Exogenous carbon emissions that are linked to one of the specific representative concentration pathways (RCP, Emi^{r})
- Carbon emissions linked to permafrost processes (PCF)
- Carbon emissions linked to Amazon dieback ($AMAZ$)
- A shock to emissions that sets emissions on a different path from the RCP pathway—for example a pulse of emissions in a specific year (Emi^{x}). The latter is used to calculate the social cost of carbon (SCC).

$$Emi_{CO2,t} = Emi^{r}_{CO2,t} + PCF_{CO2,t-1} + AMAZ_{CO2,t-1} + Emi^{x}_{CO2,t} \tag{2.1}$$

Carbon emissions are measured in $GtCO_2$. The impacts of permafrost and the Amazon dieback are lagged (for recursivity purposes) and are described below.

2.1.2 Carbon cycle

The carbon cycle is an adaptation of the FaIR 2.0 model (Leach et al. (2021)). The FaIR model includes four carbon sinks or pools that include geological processes, deep ocean invasion/equilibration, biospheric uptake/ocean thermocline invasion and rapid biospheric uptake/ocean mixed-layer invasion. Though the labels for the pools are suggestive, the pools and their parameterization are mathematical constructs that summarize a complex set of interactions that include chemical reactions in the atmosphere and oceans and the fluxes across a number of carbon sinks that include vegetation and soils. Thus each pool represents a specific impact on part of the atmospheric concentration and the aggregate atmospheric concentration is the sum across all of the pools. The carbon stock in each 'pool', indexed by cb, is equal to the previous period's concentration adjusted by some rate of decay, δ^c, plus some fraction, α^e, of current CO_2 emissions, equation (2.2).[4]

One of the innovations of the FaIR model is to make the decay rate a function of the level of concentration—as the absorption rate in each 'pool' depends on the level of carbon saturation. The coefficient α^c, determined below, is the state-dependent variable that accounts for adjustments in the decay rate. Note that the adjustment is uniform across all pools. Equation (2.3) then determines the total stock of carbon in the atmosphere (in $GtCO_2$), including the pre-industrial stock converted from parts per million to $GtCO_2$, where χ_{co2}^{ppm} is the conversion factor.[5] Equation (2.4) converts the atmospheric stock of carbon into parts per million.

$$\begin{aligned} R_{cb,t} &= \alpha_{cb}^{e} Emi_{CO2,t} + \left(1 - \frac{\delta_{cb}^{c}}{\alpha_{t-1}^{c}}\right) R_{cb,t-1} \\ &= \alpha_{cb}^{e} Emi_{CO2,t} + \left(1 - \frac{1}{\alpha_{t-1}^{c}\tau_{cb}^{c}}\right) R_{cb,t-1} \end{aligned} \tag{2.2}$$

$$StockGt_t = \sum_{cb} R_{cb,t} + \chi_{co2}^{ppm} Conc_{co2,t}^{pre} \tag{2.3}$$

$$Conc_{co2,t} = StockGt_t / \chi_{co2}^{ppm} \tag{2.4}$$

Equation (2.5) determines cumulative anthropogenic CO_2 emissions from the beginning of time. It is initialized at 521.45 gtC in 2009. Equation (2.6) determines the carbon stock levels in all non-atmospheric carbon pools (soils, oceans, etc.) derived from anthropogenic carbon emissions, i.e. it is total cumulative emissions less the contemporaneous level in the atmosphere. This is measured in gtC.

$$CEmi_{co2,t} = CEmi_{co2,t-1} + \chi_{C}^{co2} Emi_{CO2,t} \tag{2.5}$$

$$CStock_{co2,t}^{x} = CEmi_{co2,t} - \chi_{C}^{co2} \sum_{cb} R_{cb,t} \tag{2.6}$$

The original version of the FaIR model calibrated the α^c parameter to two long-term conditions relating to a variable linked to the integrated impulse response function (`iIRF100`), which represents the 100-year average airborne fraction of a pulse of CO_2 (Joos et al. (2013)).

The first of these conditions, derived from a physical relation, is the following:

$$iIRF100_t = \sum_{cb} \alpha_{cb}^{e} \alpha^{c} \tau_{cb} \left[1 - \exp\left(-\frac{100}{\alpha^{c}\tau_{cb}}\right)\right]$$

The second condition, derived from a reduced form statistical relation, is:

[4] In the original FaIR documentation, the motion equation for the pools is given by:

$$\frac{dR}{dt} = \alpha^{e} Emi_t - \frac{1}{\alpha^{c}\tau^{c}} R_t$$

One discrete approximation of this equation is therefore:

$$R_t = \alpha^{e} Emi_t + \left(1 - \frac{1}{\alpha^{c}\tau^{c}}\right) R_{t-1}$$

[5] $7.81 = 2.13 \times 44/12$

$$iIRF100_t = r^0 + r^C\, CStock^x_{co2,t} + r^T\, Temp^{atm}_t \tag{2.7}$$

which links *iIRF100* to the carbon stock level and temperature. The two equations together determine *iIRF100* and α^c, though α^c is only determined implicitly. In the supplementary materials of Dietz et al. (2021) and in Leach et al. (2021), it is posited that α^c is approximately exponential in *iIRF100* and can be expressed as:

$$\alpha^c_t = \alpha^1 \exp\left(\alpha^2\, iIRF100_t\right) \tag{2.8}$$

where α^1 and α^2 are calibrated parameters.

2.1.3 Methane emissions and concentration

Methane emissions, similar to carbon emissions, are on an RCP-based reference trajectory, which in the META 21 model are adjusted by (lagged) methane emissions linked to permafrost processes and to ocean methane hydrates (OMH). Atmospheric concentration of methane is based on a single-box model with a decay rate given by δ^{ch4}. Emissions are converted from metric tons to parts per billion (ppb) using the conversion factor $\chi^{ppbch4}_{mt} = 2.78$.

$$Emi_{ch4,t} = Emi^r_{ch4,t-1} + PCF_{ch4,t-1} + OMH_{t-1} \tag{2.9}$$

$$Conc_{ch4t} = Conc^{pre}_{ch4} + \left(1 - \delta^{ch4}\right)\left(Conc_{ch4,t-1} - Conc^{pre}_{ch4}\right) + Emi_{ch4,t}/\chi^{ppbch4}_{mt} \tag{2.10}$$

2.1.4 Forcing

Radiative forcing is the sum of impacts from a number of drivers. In the case of the META 21 model, there are three drivers—two of which are modeled explicitly: carbon and methane atmospheric concentration— and the third is exogenous, though dependent on the selected RCP.

Carbon

Carbon concentration is assumed to have a logarithmic impact on radiative forcing with no interactions with other forcings. The key parameter is *f2xco2*, which represents the radiative forcing resulting from a doubling of carbon concentration from its pre-industrial level. It is treated as a random variable with a mean value of 3.45. Further discussion of the sampling strategy will be reserved until later as it is linked to the parameterization of the energy balance module (EBM).

$$Forc_{co2,t} = f2xco2 \times \log\left(\frac{Conc_{co2,t}}{Conc^{pre}_{co2}}\right)/\log(2) \tag{2.11}$$

Methane

The radiative impact of methane concentration depends on the square root of the current concentration, relative to the square root of the pre-industrial methane concentration. The key parameter is α^{fch4}. There is nonetheless an interaction effect between methane concentration and the initial concentration of nitrous oxide (N_2O). The interaction expression is given in equation (2.13).

$$\begin{aligned} Forc_{ch4,t} &= \alpha^{fch4}\left[\sqrt{Conc_{ch4,t}} - \sqrt{Conc^{pre}_{ch4}}\right] \\ &- \left[f\left(Conc_{ch4,t}, Conc_{n2o,t0}\right) - f\left(Conc_{ch4,t0}, Conc_{n2o,t0}\right)\right] \end{aligned} \tag{2.12}$$

where the methane/nitrous oxide interaction relationship is given by:

$$\begin{aligned} f\left(Conc_{ch4,t}, Conc_{n2o,t}\right) &= 0.47\log\left(1 + 2.01E^{-5}\left(Conc_{ch4,t}\, Conc_{n2o,t}\right)^{0.75}\right. \\ &+ \left. 5.31E^{-15}\left(Conc_{ch4,t}\, Conc_{n2o,t}\right)^{1.52}\right) \end{aligned} \tag{2.13}$$

2.1.5 Total radiative forcing

Exogenous forcing is provided in the RCP database for four RCPs.[6] In this case we write the total radiative forcing (RF) equation, as the sum of radiative forcing over the three drivers—carbon, methane and the exogenous shifter.

$$RF_t = \sum_{em} Forc_{em,t} \tag{2.14}$$

2.1.6 The Energy Balance Module

Introduction

The energy balance model from DICE, FaIR and others can take the same form:

$$T_t = [1 + A]\, T_{t-1} + M\, RF_t \tag{2.15}$$

where T represents the (change) in temperature in time t and is defined over two or more temperature boxes.[7] In the case of DICE and FaIR 1.5 there are two boxes, essentially the atmosphere (and the surface and shallow ocean) and the (deep) ocean. In FaIR 2.0, the ocean box is split into shallow and deep. The matrix A represents the conveyance of energy across the boxes and the matrix M represents the impact of the contemporaneous radiative forcing on atmospheric temperature. In the FaIR 2.0 model, the matrix A takes the following form:

$$A = \begin{pmatrix} -(\lambda + \kappa_2)/C_1 & \kappa_2/C_1 & 0 \\ \kappa_2/C_2 & -(\kappa_2 + \varepsilon\kappa_3)/C_2 & \varepsilon\kappa_3/C_2 \\ 0 & \kappa_3/C_3 & -\kappa_3/C_3 \end{pmatrix}$$

where the parameter C is the heat capacity of each box, the κ parameters are the heat exchange coefficients, λ is a feedback parameter and ε allows to account for transient warming. M is a 3×1 vector with $1/C_1$ in the first cell.[8]

The META 21 model uses a two-box model, the parameterization of which will be described below. The change in temperature derived from equation (2.15) is further adjusted by a factor influenced by the changing albedo.

Surface albedo feedback

This module is more fully described in the supplementary materials of Dietz et al. (2021) or the primary source (Yumashev et al., 2019). The basic idea is that there are four temperatures: (1) temperature derived from the standard EBM model with no surface albedo feedback (SAF) adjustment, designated $T^{tb,0}$; (2) temperature derived from the standard PAGE-ICE model with no SAF adjustment, designated T^1; (3) temperature from the PAGE-ICE model corrected for the SAF adjustment (including adjustment of the ECS), designated T^2; and (4) the EBM-derived temperature with the SAF adjustment, designated by T^{tb}. The latter adjustment is simply an additive correction to atmospheric temperature, as measured by $T^2 - T^1$.

Equation (2.16) is the core temperature equation in the PAGE-ICE model, where T^1 represents the unadjusted change in atmospheric temperature.

$$T^1_t = T^1_{t-1} + \left(A_{t-1} - FRT\, B_{t-1} - T^1_{t-1}\right)\left(1 - \mathrm{e}^{-1/FRT}\right) + B_{t-1} \tag{2.16}$$

where[9]

$$\begin{aligned} A_{t-1} &= \frac{ECS}{F_{sl}\ln(2)} RF_{t-1} \\ B_{t-1} &= \frac{ECS}{F_{sl}\ln(2)} \left(RF_t - RF_{t-1}\right) \end{aligned}$$

and F_{sl} is the forcing slope (5.5 W/m^2) and FRT[10] is the warming half-life, from a triangular distribution from 10 to 55 with a mode of 20.

The surface albedo feedback is calculated using a quadratic approximation, where SAF decreases more rapidly as temperature increase.

---

[6] In the GAMS code, it is attached to the emission labeled `FGAS` and is fixed.

[7] It is a discrete approximation to the state-space form, which is $\dot{T} = A\,T + M\,RF$.

[8] The DICE parameterization can be converted to this matrix notation by setting κ_2 to `C3` (in the DICE code), λ is equal to `fco22x/t2xo2`, C_1 is equal to `1/C1`, C_2 is equal to `C3/C4` and κ_3 is 0.

[9] The definition of B in the SM appears to be inconsistent with the Excel code.

[10] Also called the e-folding feedback response time.

$$SAF_t = \frac{\frac{\beta_2}{3}\left(T_t^1\right)^3 + \frac{\beta_1}{2}\left(T_t^1\right)^2 + \beta_0 T_t^1 + \gamma\, T_t^1\, \delta - FSAF_0}{T_t^1 - T_{2010}^1} \tag{2.17}$$

where γ is the standard deviation of the SAF quadratic and δ is the non-linearity of SAF, drawn from a symmetric triangular distribution from -1 to 1. $FSAF_0$ is the base year SAF forcing (W/m^2).

$$\Delta FSAF_t = -SAF_t T_{t-1}^2 + \begin{cases} \frac{\beta_2}{3}\left(T_{t-1}^2\right)^3 + \frac{\beta_1}{2}\left(T_{t-1}^2\right)^2 + \beta_0 T_{t-1}^2 + \gamma\, T_{t-1}^2\, \delta & \text{if } T_{t-1}^2 < 10 \\ \psi + \alpha\left(T_{t-1}^2 - 10\right) + \sigma\left(T_{t-1}^2 - 10\right)\delta & \text{if } T_{t-1}^2 \geq 10 \end{cases} \tag{2.18}$$

The adjusted temperature in PAGE-ICE has the same formula as the unadjusted temperature in equation (2.16), with modified ECS and FRT and with the additional forcing $\Delta FSAF_t$.

$$T_t^2 = T_{t-1}^2 + \left(A_{t-1}^{'} - FRT^{'}\, B_{t-1}^{'} - T_{t-1}^2\right)\left(1 - \mathrm{e}^{-1/FRT^{'}}\right) + B_{t-1}^{'} \tag{2.19}$$

where

$$\begin{aligned} A_{t-1}^{'} &= \frac{ECS^{'}}{F_{sl}\ln(2)}\left[RF_{t-1} + \Delta FSAF_t\right] \\ B_{t-1}^{'} &= \frac{ECS^{'}}{F_{sl}\ln(2)}\left(RF_t - RF_{t-1}\right) \end{aligned}$$

and

$$\begin{aligned} ECS_t^{'} &= ECS\left(1 - \frac{ECS\left(SAF_t - \overline{SAF}\right)}{F_{sl}\ln(2)}\right)^{-1} \\ FRT_t^{'} &= FRT\left(1 - \frac{ECS\left(SAF_t - \overline{SAF}\right)}{F_{sl}\ln(2)}\right)^{-1} \end{aligned}$$

The temperature adjustment, $T_t^2 - T_t^1$, is added to the main atmospheric temperature in the model.

The unadjusted temperature is calculated as above using equation (2.15), it will be denoted T^0. N.B. The model tracks both the unadjusted and adjusted temperature. The matrix calculation uses the (lagged) unadjusted temperature to update the unadjusted atmospheric temperature, but uses the adjusted temperature to update the temperatures in the other thermal boxes.

$$T_t^{tb} = \begin{cases} T_t^{tb,0} + \left(T_t^2 - T_t^1\right) & \text{if } tb = \texttt{atm} \\ \sum_{tb'} a_{tb,tb'} T_{t-1}^{tb'} & \text{if } tb \neq \texttt{atm} \end{cases} \tag{2.20}$$

2.1.7 Country temperatures

The change in global mean surface temperature is downscaled to the country level. Equation (2.21) determines the ratio of national mean surface temperature with respect to global mean surface temperature. The term in the logarithm represents the change in the global mean surface temperature from it's pre-industrial level. The first term is the change from pre-industrial level to the level in 2010,[11] and the second term represents the incremental change in temperature in time t relative to 2010. The α^g and β^g parameters are estimated using OLS by Dietz et al. (2021).[12] The temperature level in each country, TL, is then equal to the global temperature adjusted by the downscaling factor μ^g. In addition, each country's temperature is impacted by the the slowdown of the Atlantic Meridional Overturning Circulation, T_c^h, which is described below.

$$\mu_{c,t}^g = \alpha_c^g + \beta_c^g \ln\left[\left(TL_{2010}^{atm} - 18.825\right) + \left(T_t^{atm} - T_{t_0}^{atm}\right)\right] \tag{2.21}$$

$$TL_{c,t} = \mu_{c,t}^g\left(TL_{2010}^{atm} + T_t^{atm} - T_{t_0}^{atm}\right) + T_{c,t}^h \tag{2.22}$$

[11] The notation TL is used to designate the temperature level, whereas the notation T refers to the temperature anomaly.

[12] The downscaling parameters α^g and β^g are fixed.

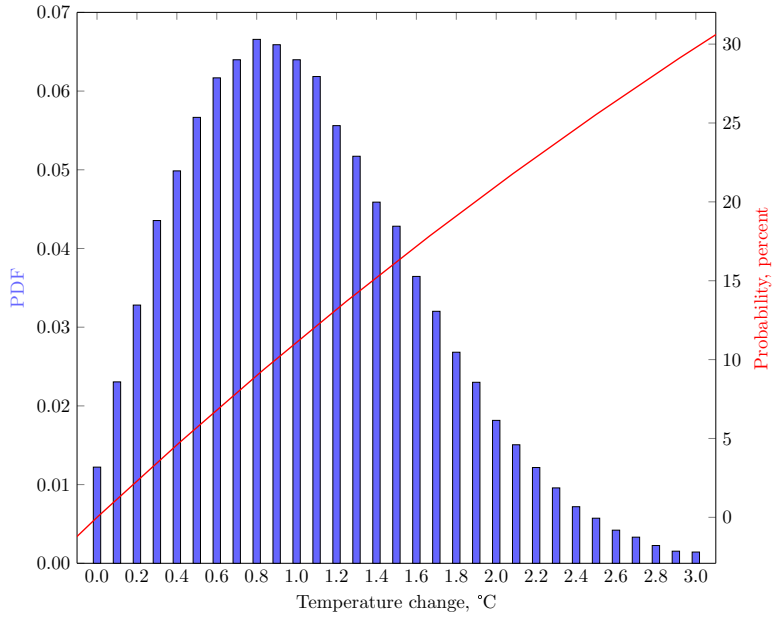


Figure 2: **Sampling for OMH, where** $b^{omh} = 0.118$

2.1.8 Summary

This concludes the emissions/temperature module. The sequence as described above is recursive given the implementation of the lag structure, with the exception of the α^c adjustment used to adjust the decay parameters in the carbon boxes. The formulas for *iIRF100* and α^c should come after updating the temperature equation.

2.2 The Impact/Tipping points Module

2.2.1 Ocean methane hydrates (OMH)

The probability of the beginning of the release of methane from ocean methane hydrates is given in equation (2.23), where b_{OMH} is the so-called hazard rate. Given the probability, in each year t, a single random draw from the binomial distribution is taken with the given probability. The indicator function takes the value of 1 for every year subsequent to a positive draw, i.e. the release of the methane hydrate starts in the first year of the positive draw, and is assumed to be irreversible.

Figure (2) depicts the probability density for a sample of 100,000 random draws for the range of temperature change from 0 to 3°C. The red line shows the level of the probability of the methane release starting for each change in temperature—it starts at 0 and rises to over 30% when the temperature change is 3°C. The mode of the PDF is at around 0.8-0.9°C. Given that the current estimated level of temperature change is already over 1°C, this parameterization implies a relatively early start to the release of the methane hydrates.

$$p_t^{OMH} = 1 - \exp\left(-b^{OMH} T_t^{atm}\right) \tag{2.23}$$

$$I_t^{OMH} = \begin{cases} 1 & \text{if } I_{t-1}^{OMH} = 1 \\ \texttt{randBinomial}(1, p_t^{OMH}) & \text{if } I_{t-1}^{OMH} = 0 \end{cases} \tag{2.24}$$

Once the emissions commence, a constant level of emissions is released each year until the given stock of sequestered methane runs out, equation (2.25), where ${CH4}_{Max}^{OMH}$ is the initial stock of sequestered methane and Δ^{OMH} is the annual release.

$$OMH_t = {CH4}_{Max}^{OMH} / \Delta^{OMH} \quad \text{if } I_t^{OMH} = 1 \text{ and } \sum_{t'=t_0}^{t-1} OMH_t < {CH4}_{Max}^{OMH} \tag{2.25}$$

2.2.2 Permafrost carbon feedback (PCF)

The specification of the permafrost module derives from Kessler (2017). It is based on a two-stage process. In the first stage, near-surface permafrost thaw is a linear function of temperature change relative to the reference year (i.e. 2010). PF^{ext} in equation (2.26) is the area of permafrost remaining at time t, and β^{pf} is the sensitivity of permafrost thaw with respective to temperature change. Note that $PF_t^{ext} = PF_t^{area}/PF_{t0}^{area}$. The amount of carbon in the freshly thawed permafrost is the product of the total stock of carbon locked in the near-surface northern circumpolar permafrost region, C^{pf}, and the freshly thawed permafrost area, equation (2.27).

$$PF_t^{ext} = 1 - \beta^{pf}\left(T_t^{atm} - T_{t0}^{atm}\right) \tag{2.26}$$

$$CThaw_t^{pf} = -C^{pf}\left(PF_t^{ext} - PF_{t-1}^{ext}\right) \tag{2.27}$$

The stock of carbon is divided into two reservoirs: one is passive that releases no carbon, and another that is active that decomposes exponentially. $CumCThaw^{pf}$ in equation (2.28) is the cumulative release of emissions in year t from the active reservoir, where χ^{psv} is the passive proportion of the total reservoir and τ^{pf} is the rate of decomposition of the active reservoir. The decomposed emissions generate carbon and methane emissions in fixed proportion. Equation (2.29) represents the level of both emissions, where φ^{ch4} is the proportion of methane emissions. Permafrost carbon emissions are measured in gtC and are thus converted to $gtCO_2$, the same units as carbon emissions. Methane emissions are converted from gtC to mt NH_4, where the conversion factor is $1000 \times 16/12$.[13]

$$CumCThaw_t^{pf} = (1-\chi^{psv})\sum_{t'=t0}^{t} CThaw_{t'}^{pf}\left(1 - e^{(t'-t)/\tau^{pf}}\right) \tag{2.28}$$

$$PCF_{em,t} = \begin{cases} \left(1-\varphi^{ch4}\right)\left(CumCThaw_t^{pf} - CumCThaw_{t-1}^{pf}\right)/\chi_C^{co2} & \text{if } em = co2 \\ \varphi^{ch4}\left(CumCThaw_t^{pf} - CumCThaw_{t-1}^{pf}\right)\chi_{mtNH4}^{gtC} & \text{if } em = ch4 \end{cases} \tag{2.29}$$

2.2.3 The Amazon rainforest dieback module (AMAZ)

The Amazon rainforest dieback module is implemented similar to methane hydrates. There is a given quantity of carbon sequestered in the forest. When the dieback 'starts', the quantity is released into the atmosphere at a constant level each year until it is exhausted.[14] The probability of the start of the dieback increases with temperature, but the year the dieback starts is a random event based on the probability. The coefficient ψ^x is an interaction effect between the Amazon dieback and other tipping points, which is described further below.

$$p_t^{AMAZ} = \min\left(\psi_{amaz,t}^{x}\left[1 - \exp\left(-b^{AMAZ}\max\left(0, T_t^{atm} - 1\right)\right)\right], 1\right) \tag{2.30}$$

$$I_t^{AMAZ} = \begin{cases} 1 & \text{if } I_{t-1}^{AMAZ} = 1 \\ \texttt{randBinomial}(1, p_t^{AMAZ}) & \text{if } I_{t-1}^{AMAZ} = 0 \end{cases} \tag{2.31}$$

$$AMAZ_t = CO2_{Max}^{AMAZ}/NYear^{AMAZ} \text{ if } I_t^{AMAZ} = 1 \& \sum_{t'=t_0}^{t-1} AMAZ_{t'} < CO2_{Max}^{AMAZ} \tag{2.32}$$

2.2.4 Ice sheets and sea-level rise

The Greenland ice sheet module (GIS)

The specification of the disintegration of the Greenland Ice Sheet is derived Nordhaus (2019). Equation (2.33) posits a relation between the volume of GIS and an equilibrium temperature anomaly, T_t^{eGIS}. The GIS volume is measured as an index between 0 and 1, where the volume is at an initial (and maximum) level at 1, and fully melted at 0. When the equilibrium temperature anomaly is 0 the volume is maximized, i.e. V^{GIS} is equal to 1. On the other hand, when the volume is fully melted, the equilibrium temperature anomaly is equal to T^{max}, which Nordhaus has estimated to be 3.4°C. The default assumption is that this relation is linear and thus η^{eGIS} equals 1. The rate of melting, equation (2.34), depends on the difference between the actual and equilibrium temperature anomalies and the GIS volume. The default value for β^{GIS} is -0.0000106 and $\eta^{GIS} = 0.2$. Equation (2.35) updates the GIS volume,

[13] The atomic weight of carbon is 12 and that of methane is 16.

[14] The factor $NYear^{AMAZ}$ represents the time it will take to release the carbon in years—unlike the methane hydrate factor Δ^{OMH}, which represents the amount released each year.

only if the melt rate is negative, and, in which case, an interaction effect is taken into account. Equation (2.36) posits a linear relation between the GIS volume and sea-level rise, where the default value for SLR_{max}^{GIS} is 7 (m), i.e. when fully melted, sea-level rise would be 7 meters.

$$T_t^{eGIS} = T^{max}\left(1 - V_t^{GIS}\right)^{\eta^{eGIS}} \tag{2.33}$$

$$\Delta V_t^{GIS} = \beta^{GIS}\texttt{sign}\left(T_{t-1}^{atm} - T_{t-1}^{eGIS}\right)\left(T_{t-1}^{atm} - T_{t-1}^{eGIS}\right)^2\left(V_{t-1}^{GIS}\right)^{\eta^{GIS}} \tag{2.34}$$

$$V_t^{GIS} = V_{t-1}^{GIS} + \begin{cases}\psi_{gis,t}^x \Delta V_t^{GIS} & \text{if } \Delta V_t^{GIS} < 0 \\ 0 & \text{if } \Delta V_t^{GIS} \geq 0\end{cases} \tag{2.35}$$

$$SLR_t^{GIS} = SLR_{max}^{GIS}\left(1 - V_t^{GIS}\right) \tag{2.36}$$

The West Arctic ice sheet module (WAIS)

Dietz et al. (2021) cite Diaz and Keller (2016) for the specification of the disintegration of the West Arctic ice sheet (WAIS). Similar to the methane hydrates and the Amazon dieback, the trigger for the start of the disintegration of the West Arctic ice sheet is based on a random start year that is a function of a probability. Once the disintegration starts, the annual sea-level rise generated by WAIS is a constant amount, R^{WAIS}.

$$p_t^{WAIS} = \min\left(b^{WAIS}\psi_{wais,t}^x\left(T_t^{atm} - 0.6\right)^2, 1\right) \tag{2.37}$$

$$I_t^{WAIS} = \begin{cases}1 & \text{if } I_{t-1}^{WAIS} = 1 \\ \texttt{randBinomial}(1, p_t^{WAIS}) & \text{if } I_{t-1}^{WAIS} = 0\end{cases} \tag{2.38}$$

$$\Delta SLR_t^{WAIS} = R^{WAIS} I_t^{WAIS} \tag{2.39}$$

$$SLR_t^{WAIS} = \sum_{t'=t_0}^{t} \Delta SLR_{t'}^{WAIS} \tag{2.40}$$

The sea-level rise module (SLR)

Sea-level rise is generated by four sources: the aforementioned melting of the Greenland and West Arctic ice sheets, the melting of glaciers and small icecaps, and thermal expansion. The latter two have an annual melt that is linear in the temperature anomaly. The variable SLR^{THX} represents the aggregate sea level rise from the latter two sources. Total sea-level rise, SLR^{TOT}, is the sum across all sources.

$$SLR_t^{THX} = SLR_{t-1}^{THX} + \left(\rho^{THX} + \rho^{GSIC}\right) T_t^{atm} \tag{2.41}$$

$$SLR_t^{TOT} = SLR_t^{THX} + SLR_t^{GIS} + SLR_t^{WAIS} \tag{2.42}$$

2.2.5 Changes in the Atlantic Meridional Overturning Circulation (AMOC)

The timing of the tipping point for the slowdown of the Atlantic Meridional Overturning Circulation is based on a probabilistic specification that is a function of the temperature anomaly. The hazard rate is given by b^{AMOC}.[15] The impact from the AMOC slowdown is felt differentially across countries. This has been estimated using four general circulation models (GCMs) for a given impulse of fresh water from higher precipitation and ice melt from Greenland. The results of these GCM runs is a permanent change in the long-run temperature across countries, $T_{c,max}^h$. The long-run temperature change occurs over time, with an annual increment given by Δ^{AMOC}. Equation (2.45) thus provides the country-specific temperature anomaly due to the AMOC slowdown, $T_{c,t}^h$. This anomaly is added to the country temperature anomaly due to the change in mean global temperature, see equation (2.22) above.

$$p_t^{AMOC} = \min\left(\psi_{amoc,t}^x\left(1 - \exp\left(-b^{AMOC}\right)\right), 1\right) \tag{2.43}$$

[15] In the full model the probability may be impacted by interaction effects, ψ_{amoc}^x.

$$I_t^{AMOC} = \begin{cases} 1 & \text{if } I_{t-1}^{AMOC} = 1 \\ \texttt{randBinomial}(1, p_t^{AMOC}) & \text{if } I_{t-1}^{AMOC} = 0 \end{cases} \tag{2.44}$$

$$T_{c,t}^h = T_{c,t-1}^h + \begin{cases} 0 & |T_{c,t-1}^h| \geq |T_{c,max}^h| \\ \dfrac{T_{c,max}^h}{\Delta^{AMOC}} & |T_{c,t-1}^h| < |T_{c,max}^h| \end{cases} \tag{2.45}$$

2.2.6 Weakening of the Indian Summer Monsoon (ISM) module

The specification of this module is due to Schewe and Levermann (2012) and Belaia, Funke, and Glanemann (2017). Equation (2.46) determines the average sulfate burden for India. It is a function of sulfur dioxide emissions—taken from one of the RCP scenarios and adjusted for India's share of Asia's SO_2 emissions, the fractional sulfate yield, H_{so2}, the atmospheric lifetime of sulfate, V_{so2}, divided by the land area, Ω_{IND}. The parameters are fixed with ϕ_{IND}^{so2} equal to 8.5 divided by Asia's RCP-related emissions of SO_2, $H_{so2} = 0.65$, $V_{so2} = 2/365$, and $\Omega_{IND} = 4.2E12m^2$.

$$B_t^{so4} = \phi_{IND}^{so2} Emi_{so2,t}^r \chi H_{so2} V_{so2} \frac{10^{12}}{\Omega_{IND}} \tag{2.46}$$

Equation (2.47) describes the actual planetary albedo. The increment is a function of the fraction of light transmitted by the aerosol layer, φ, the current value of the surface albedo, A^s, and the sulfate burden over India, B^{so4}. The parameter β^{pl} represents the backscatter fraction and α_3^{pl} the mass scattering efficiency. The parameters are fixed with $A_{t_0}^{pl} = 0.47$[16], $\varphi = 0.76$, $A^s = 0.16$, $\beta^{pl} = 0.29$, and $\alpha_3^{pl} = 8.5$.

$$A_t^{pl} = A_{t_0}^{pl} + 2\varphi^2(1 - A^s)^2 \beta^{pl} \alpha_3^{pl} B_t^{so4} \tag{2.47}$$

The critical albedo is an increasing function of CO_2 concentration, equation (2.48). The parameters are fixed with $\alpha_1^{pl} = 0.02$ and $\alpha_2^{pl} = 0.37$.

$$A_t^{pl,crit} = \alpha_1^{pl} \ln(Conc_{co2,t}) + \alpha_2^{pl} \tag{2.48}$$

The probability of having a wet day during the beginning of the season depends on the strength of the so-called Walker circulation, i.e. the Pacific Ocean atmospheric circulation, in May, which is linked to the presence of an El Niño effect. The Walker circulation in May is provided in equation (2.49) and it is a positive function of the incremental increase in the temperature anomaly. It is also subject to an interaction effect, ψ_{ISM}^x. The parameters are fixed with $\beta^{Circ} = -0.29$ and $W_{t_0}^{Circ} = 1009.98$.

$$W_t^{Circ} = \psi_{ISM,t}^x \frac{\beta^{Circ}}{30} \left(T_t^{atm} - T_{t_0}^{atm}\right) + W_{t_0}^{Circ} \tag{2.49}$$

The probability of a wet day during the first days of the season, $P^{init,1}$, is given in equation (2.50), where ϖ', ξ_0 and ϖ_0 are calibrating coefficients—assuming that the actual albedo is less than the critical albedo. The probability of a wet day during the first days of the season, P^{init}, depends therefore on the critical albedo threshold. If the actual albedo exceeds the critical albedo, the probability of a wet day is equal to 18%, 1 minus the maximum probability of a wet day, $P^{max} = 0.82$. The parameters are fixed with $\varpi' = 0.41$, $\xi_0 = 1008.65$ and $\varpi_0 = 0.2$.

$$P_t^{init,1} = \varpi' \left(W_t^{Circ} - \xi_0\right) + \varpi_0 \tag{2.50}$$

$$P_t^{init} = \begin{cases} P_t^{init,1} & \text{if } A_t^{pl} < A_t^{pl,crit} \\ 1 - P^{max} & \text{if } A_t^{pl} \geq A_t^{pl,crit} \end{cases} \tag{2.51}$$

The probability of precipitation on any day during the growing season (of which there are 136 days) depends on the time of the season. For the first days, the probability depends on external events such as albedo and the strength of the Walker circulation. Once the season is under way, the probability is taken from the Uniform distribution.

$$P_{d,t}^{prec} = \begin{cases} \texttt{randBinonmial}(1, P_t^{init}) & \text{if } d \in earlydays \\ \texttt{Uniform}(0,1) & \text{if } d \notin earlydays \end{cases} \tag{2.52}$$

Precipitation on wet days is assumed to increase with warming as there is more vapor in the atmosphere. Precipitation on dry days is constant. The parameters are fixed with $\mu^p = 0.42\,\texttt{mm/d}$, $DailyPrec_{t_0}^{\texttt{wet}} = 9\,\texttt{mm}$ and $DailyPrec_{t_0}^{\texttt{dry}} = 0\,\texttt{mm}$.

[16] Directly embedded in the Excel formulas.

$$DailyPrec_t^c = \begin{cases} \mu^p \left(T_t^{atm} - T_{t_0}^{atm}\right) + DailyPrec_{t_0}^c & \text{if } c = \texttt{wet} \\ DailyPrec_{t_0}^c & \text{if } c = \texttt{dry} \end{cases} \tag{2.53}$$

The probability of a wet day increases with the number of wet days preceding the current day. Finally, the amount of precipitation on any given day depends on the draw and the time of the season. δ^d is set to 16 days. The adjustment factor of 4 is due to the fact that the season has a 4-day time step, not a daily time step.

$$P_{d,t}^{wet} = \max \left[\min \left(\frac{\frac{4}{\delta^d} \sum_{d'=d-\delta^d}^{d-1} Prec_{d',t} - DailyPrec_t^{dry}}{DailyPrec_t^{wet} - DailyPrec_t^{dry}}, P^{max} \right), 1 - P^{max} \right] \tag{2.54}$$

$$Prec_{d,t} = \begin{cases} DailyPrec_t^{wet} & \text{if } \left(P_{d,t}^{prec} = 1 \text{ and } d \in earlydays\right) \\ DailyPrec_t^{wet} & \text{if } \left(P_{d,t}^{prec} < P_{d,t}^{wet} \text{ and } d \notin earlydays\right) \\ DailyPrec_t^{dry} & \text{if } \left(P_{d,t}^{prec} = 0 \text{ and } d \in earlydays\right) \\ DailyPrec_t^{dry} & \text{if } \left(P_{d,t}^{prec} \geq P_{d,t}^{wet} \text{ and } d \notin earlydays\right) \end{cases} \tag{2.55}$$

In summary, for each day of the growing season, a random draw is taken that determines if it is a wet or dry day. At the beginning of the season (the first 16 days) the probability of a wet day is influenced by external events such as the Walker circulation. Once the season is under way, the draw is taken from the Uniform distribution.

The impact of the weakening of the monsoon is measured using the average precipitation over the growing season.

$$Prec_t^{avg} = \frac{1}{136} \sum_{d=1}^{136} Prec_{d,t} \tag{2.56}$$

2.2.7 Interactions

Of the 8 tipping points, 3 have direct interactions that are hard-wired into the model: permafrost (PCF), methane hydrates (OMH) and surface albedo (SAF). The other 5 interact based on a tipping point-specific indicator. The interaction action effect for tipping point *tp* is the product of a parameterization of the cross-tipping point interaction raised to the power of the index, equation (2.58).

$$INDX_{tp,t} = \begin{cases} I_t^{AMAZ} & \text{if } tp = \texttt{AMAZ} \\ 1 - V_t^{GIS} & \text{if } tp = \texttt{GIS} \\ p_t^{WAIS} & \text{if } tp = \texttt{WAIS} \\ I_t^{AMOC} & \text{if } tp = \texttt{AMOC} \\ \max \left(\min \left[10000 \left(1 - \frac{W_t^{Circ}}{W_{t_0}^{Circ}} \right), 1 \right], 0 \right) & \text{if } tp = \texttt{ISM} \end{cases} \tag{2.57}$$

$$\psi_{tp,t}^x = \prod_{tp' \neq tp} \Xi_{tp',tp}^{INDX_{tp',t-1}} \tag{2.58}$$

2.3 The Economic Module

2.3.1 Damages

The META 21 Model has two sources of damage—change in temperature and sea-level rise. In the case of India, there is an additional source of damage derived from the changing of the summer monsoon.

Temperature damage is a quadratic expression in the change in (country) temperature since 1990. The source of the parameter values is Burke, Hsiang, and Miguel (2015). Dietz et al. (2021) describe their transformation of the Burke, Hsiang, and Miguel (2015) estimates. The study further provides an estimate of the mean and standard deviation for the β coefficients. Sampling of the β parameters is undertaken with a supplied correlation matrix across countries and between the two β parameters. There are over 75,000 coefficients in the correlation matrix.

$$Dam_{c,t}^{Temp} = \begin{cases} 0 & \text{if } ifDamage = \texttt{No} \\ \beta_c^1 \left(TL_{c,t} - TL_{c,1990}\right) + \beta_c^2 \left(TL_{c,t} - TL_{c,1990}\right)^2 & \text{if } ifDamage = \texttt{Yes} \end{cases} \tag{2.59}$$

Sea-level rise damage is a linear function of sea-level rise. The sampling strategy is to assume a Triangular distribution of damages based on simulations with the CIAM model (Diaz (2016)).

$$Dam_{c,t}^{SLR} = \begin{cases} 0 & \text{if } ifDamage = \texttt{No} \\ \chi_c^{slr} SLR_t^{TOT} & \text{if } ifDamage = \texttt{Yes} \end{cases} \tag{2.60}$$

Damage from the changing summer monsoon only applies to India (`IND`). There are different levels of damage depending on whether average precipitation falls below a drought threshold or exceeds a flood threshold. The parameters of the ISM module are fixed with the thresholds given by $Prec^{drought} = 2.8667$ and $Prec^{flood} = 7.6667$, and the damages (as a percent of GDP) are $Dam^{drought} = 0.035$ and $Dam^{flood} = 0.0084$.

$$Dam_{c,t}^{ISM} = \begin{cases} 0 & \text{if } (ifDamage = \texttt{No}) \text{ or } (c \neq \texttt{IND}) \\ Dam^{drought} & \text{if } Prec_t^{avg} \leq Prec^{drought} \\ 0 & \text{if } Prec^{drought} < Prec_t^{avg} < Prec^{flood} \\ Dam^{flood} & \text{if } Prec_t^{avg} \geq Prec^{flood} \end{cases} \tag{2.61}$$

2.3.2 Non-market damages

The evaluation of non-market damages is based upon the specification by Manne and Richels (2005). Non-market damage factor is described in equation (2.62). If there is no change in temperature (relative to the change in temperature in 2000) the non-market damage factor is 1. On the other hand, if the change in temperature equals some posited catastrophic change in temperature, T^{cat}, then the non-market damage totally wipes out welfare. The exponent, h, is a weight factor that is a function of per capita income and calibrated to assumptions about the willingness to pay for a certain level of climate damages (given income). The parameters of the non-market damage expressions are all fixed with the willingness to pay, χ^{loss}, set at 0.02 for a reference temperature change, T^{ref} of 2.5°C, the catastrophic temperature change, T^{cat}, is set to 17.67°C, the willingness to pay to avoid the reference temperature change at a level of income of $\$25,000$ is set at 1%, where $\chi^{ref25K} = ln(0.01)/25$. The temperature change in 2000 is 0.413.

$$\varphi_{c,t}^{nmrkt} = \left(1 - \left(\frac{T_t^{atm} - T_{2000}^{atm}}{T^{cat}}\right)^2\right)^{h_{c,t}} \tag{2.62}$$

$$h_{c,t} = min\left[\frac{\log\left(1 - \frac{\chi^{loss}}{1 + 100\mathrm{e}^{-\chi^{ref25K}\, Cons_{c,t}^{pc}}}\right)}{\log\left(1 - \left(\frac{T^{ref}}{T^{cat}}\right)^2\right)}, 1\right] \tag{2.63}$$

2.3.3 GDP and consumption

Pre-damage GDP per capita is described in equation (2.64). It is equal to the previous period's GDP augmented by an exogenous growth rate, γ^x. The exogenous growth rate comes from one of four SSPs.[17] The use of the SSPs and their extension beyond 2100 is described in the supplemental materials of (Dietz et al. (2021)). The equation allows for two cases and a convex combination of the two. In one case, when φ is set to 1, the damages only impact GDP levels as there is no memory in the system relative to damages in the previous period. In other words, pre-damage GDP per capita is equal to the exogenous GDP per capita as given by the SSPs. In the other case, the growth is applied to actual GDP in the previous period, in which case damages influence the growth of GDP, not just the level. The default level for the weight, φ is 0.5, it is sampled using a Triangular distribution bounded by 0 and 1 with a mode of 0.5.

$$y_{c,t}^g = \left(1 + \gamma_{c,t}^x\right)\left[\varphi y_{c,t-1}^x + (1 - \varphi) y_{c,t-1}\right] \tag{2.64}$$

Post-damage GDP per capita is described in equation (2.65), which reflects the compound damage from temperature and sea-level rise (and the changing monsoon in the case of India). The damage from temperature is adjusted for the growth rate, as it is assumed to shift the growth rate additively. Consumption per capita is equal to GDP per

[17] The Shared Socio-Economic Pathways (SSPs) have been developed by the Integrated Assessment Modeling Community (IAMC) for the purposes of assessing the economics of climate change through 2100. Five SSPs were developed by three members of the IAMC—four of the SSPs are available as part of the META 21 distribution. See for example Dellink et al. (2017), for a description of the SSPs developed at the OECD.

capita less savings. The rate of savings is calibrated to 2010 levels derived from the World Bank's WDI databank—and is fixed over time at the reference level.

$$y_{c,t} = y^{g}_{c,t}\left(1 - \frac{Dam^{Temp}_{c,t}}{1+\gamma^{x}_{c,t}}\right)\left(1 - Dam^{SLR}_{c,t}\right)\left(1 - Dam^{ISM}_{c,t}\right) \tag{2.65}$$

$$c_{c,t} = (1 - s_c)\, y_{c,t} \tag{2.66}$$

Utility (on a per capita basis), is given in equation (2.67), which uses a constant-elasticity-of-substitution expression and where η represents the elasticity of the marginal utility of consumption. η is sampled from a Triangular distribution bounded by 0.5 and 2 with a mode of 1.5.

$$u_{c,t} = \frac{1}{1-\eta}\left(c_{c,t}\varphi^{nmrkt}_{c,t}\right)^{1/(1-\eta)} \tag{2.67}$$

2.3.4 Aggregations and the social cost of carbon

The country results can be aggregated to any set of regional aggregations, defined by the set R. Set R must at least contain the world as one aggregate region designated by the label `WLD`. Equations (2.68) and (2.69) define respectively aggregation population and aggregate consumption per capita. Equation (2.70) defines discounted aggregate utility, where the discounting is relative to a reference year, t_r, which is currently set at 2020. The parameter ρ is the pure rate of time preference (PRTP). The Excel file refers to EV as equivalent regional consumption per capita. Global welfare is defined in equation (2.72). The social cost of carbon, defined as $SCC_t = \frac{\partial W/\partial E_t}{\partial W/\partial c_t}$, is approximated by injecting 10 gt of CO_2 emissions in 2020. It is measured as the change in global welfare relative to the reference year global per capita consumption. PRTP is sampled from a Triangular distribution bounded by 0.01% and 2% with a mode of 1%.

$$Pop^{a}_{R,t} = \sum_{c\in R} Pop_{c,t} \tag{2.68}$$

$$c^{a}_{R,t} = \sum_{c\in R} Pop_{c,t}\, c_{c,t} / Pop^{a}_{R,t} \tag{2.69}$$

$$U^{a}_{R,t} = \sum_{c\in R} u_{c,t} Pop_{c,t}\,(1+\rho)^{-(t-t_r)} \tag{2.70}$$

$$EV^{a}_{R,t} = \left(\frac{(1-\eta)U^{a}_{R,t}}{Pop^{a}_{R,t}}\right)^{1/(1-\eta)} \tag{2.71}$$

$$W^{g} = \sum_{t\geq t_r} U^{a}_{\texttt{WLD},t} \tag{2.72}$$

$$SCC = -\frac{\left(W^{g}_{x} - W^{g}_{r}\right)/Emi^{x}_{co2,t_t}}{\left(c^{a}_{\texttt{WLD},t_r}\right)^{-\eta}} \tag{2.73}$$

Chapter 3

Parameterization

This section will mostly focus on the parameters subject to the LHS sampling. Further details on the parameterization can be found in the original META 21 Excel workbook and the supplemental materials of Dietz et al. (2021).

3.1 Carbon cycle

Table 3.1 presents the key parameters for equation (2.2). By default, the model uses the values from the FaIR model—listed in the column labeled *AR5-IR*.[1] Note that α^e_{BIOSF} is always calculated by residual to ensure that the fraction coefficients sum to unity. The table shows the value for the decay rates. The corresponding $\tau = 1/\delta^c$ parameter reflects the lifetime of carbon in the individual pools in years, e.g. 1 million, 394.4, 36.54 and 4.304 years for the *AR5-IR* estimates.

Table 3.1: **Key parameters for the carbon cycle module**

	AR5-IR	AR5-PI	MESMO (lowest decay)	ACC2 (highest decay)
α^e_{GEOPR}	0.2173	0.1545	0.2850	0.1780
α^e_{DPOCN}	0.2240	0.1924	0.2940	0.1650
α^e_{BIOSF}	0.2824	0.2423	0.2380	0.3800
α^e_{MXOCN}	0.2763	0.4108	0.1830	0.2770
δ^c_{GEOPR}	1e-6	1e-6	1e6	1e-6
δ^c_{DPOCN}	0.0025	0.0036	0.0022	0.0026
δ^c_{BIOSF}	0.0274	0.0325	0.0400	0.0271
δ^c_{MXOCN}	0.2323	0.2243	0.4965	0.2686

Table 3.2 provides the main sampling assumptions for the carbon cycle parameters. The column labeled 'X' represents the parameter as used in this document, and the column labeled 'Code' represents the code used for the LHS utility (and corresponds to the naming convention of the META 21 Excel file). The middle panel shows the expected mean and standard deviation for each of the uncertain parameters based on the input parameters, and the right panel shows the actual mean and standard deviation from the sample.[2] In most cases, the differences are small—less than 0.1% in absolute terms. The exceptions are the log-normal distributions where there can be more significant deviations in the standard deviation.

Though the LHS sample is quite good at representing the targeted marginal distribution of each of the individual parameters of the carbon cycle, it does not perform well in capturing the desired correlation across parameters. Tables 3.3 and 3.4 provide the targeted and sample correlations, respectively, for the carbon cycle parameters. The most severe example of the deviation between the target and sample correlation, is the correlation between `alpha3` and `rho2`. The desired correlation is -0.56 and the sample correlation is around 0. This requires further investigation,

[1] These appear to be derived from Table 1 in Millar et al. (2017).

[2] The actual input parameters for the LHS utility can be seen in the `META21.in` file, which is the input file for the LHS utility.

Table 3.2: **Sampling assumptions for carbon cycle parameters**

X	Code	Distribution	Expected		Sample	
			mu	σ	mu	σ
α^e_{cb1}	alpha0	LogNormal(0.235,0.136)	0.235	0.136	0.235	0.136
α^e_{cb2}	alpha1	Logistic(0.224,3.4e-5)	0.224	6.16e-5	0.224	6.17e-5
α^e_{cb4}	alpha3	Levy(0.276,4.0e-7)	∞	∞	0.298	1.95
δ^c_{cb2}	rho1	LogNormal(0.0036,0.0121)	0.0036	0.0121	0.0036	0.0106
δ^c_{cb3}	rho2	Kumaraswamy(1.01,1.7,0.01,0,048)	0.029	0.008	0.029	0.008
δ^c_{cb4}	rho3	LogNormal(0.25,0.30)	0.25	0.30	0.25	0.30
r^0	r_0	Normal(32.4,2.55)	32.4	2.55	32.4	2.55
r^C	r_C	Normal(0.019,0.0015)	0.019	0.0015	0.019	0.0015
r^T	r_T	Normal(4.17,0.33)	4.17	0.33	4.17	0.33

Table 3.3: **Targeted correlations for the carbon cycle parameters**

	alpha0	alpha1	alpha3	rho1	rho2	rho3
alpha0	1					
alpha1	-0.6348	1				
alpha3	0.0876	-0.1113	1			
rho1	0.5673	-0.1757	-0.1047	1		
rho2	-0.0190	0.3920	-0.5641	0.0686	1	
rho3	-0.3432	-0.4030	-0.1319	0.0686	-0.2532	1

but as noted in Swiler and Wyss (2004), there can be inconsistencies in the desired correlations that cannot be overcome when sampling.

3.2 Forcing and energy balance model parameters

The sampling strategy for the forcing and energy balance model parameters is provided in Table 3.5.[3] Similar to the carbon cycle parameters, the sampled distribution modes line up well with the desired modes, with the largest deviation, around 1.1%, for the mean of the C_0 parameter. Tables 3.6 and 3.7 provide, respectively, the targeted and sample correlations for the energy balance model parameters. The deviations between the two are much less pronounced than for the correlations across the carbon cycle parameters—particularly for the two pairs with large correlations, i.e., between xi_3 and f2xco2 and between f2xco2 and t2xco2.

[3] Note that the parameter t2xco2, also denoted as *ECS* is only used in the version of the model that has the Albedo effect on temperature.

Table 3.4: **Sample correlations for the carbon cycle parameters**

	alpha0	alpha1	alpha3	rho1	rho2	rho3
alpha0	1					
alpha1	-0.5922	1				
alpha3	-0.0107	0.0023	1			
rho1	0.3438	-0.0945	-0.0035	1		
rho2	-0.0203	0.3843	-0.0164	0.0398	1	
rho3	-0.2339	-0.1334	-0.0047	-0.0482	-0.1903	1

Table 3.5: **Sampling assumptions for the forcing and EBM parameters**

Parameter	LHS	Distribution	Expected		Sample	
			mu	σ	*mu*	σ
$1/C_1$	`xi_1`	`Pareto(5.9,0.116)`	0.14	0.029	0.14	0.029
C_2	`C_0`	`Pareto(1.7,53.0)`	128.0	∞	126.6	259.6
κ_2	`xi_3`	`Triangular(0.5,0.5,1.24)`	0.75	0.174	0.75	0.174
f2xco2	`f2xco2`	`Normal(3.46,0.437)`	3.46	0.437	3.46	0.437
ECS	`t2xco2`	`Normal(3.25,0.800)`	3.25	0.800	3.25	0.800

Table 3.6: **Targeted correlations for the energy balance model parameters**

	xi_1	C_0	xi_3	f2xco2	t2xco2
xi_1	1				
C_0	-0.0445	1			
xi_3	-0.4372	-0.1198	1		
f2xco2	0.0139	-0.0397	-0.4623	1	
t2xco2	-0.1934	-0.0802	0.0655	0.6512	1

3.3 The SAF module

Most of the parameters for equations (2.17) and (2.18) are static and provided in Table 3.8. The parameter $FSAF_0$ is calibrated in the base period using the numerator of equation (2.17), i.e. the polynomial expression using the atmospheric temperature in the base year. The parameter ψ (also referred to as `SegSwitchFSAF`) is also calibrated using the same expression, with a (change in) temperature value of 10.

The remaining parameters are stochastic with a description of the distributions given in Table 3.9.[4] These are sampled independently.

3.4 Tipping point parameters

3.4.1 Ocean methane hydrates

There are two key sources for the parameterization: Whiteman, Hope, and Wadhams (2013) and Ceronsky et al. (2011). Table 3.10 provides the available range. The Whiteman, Hope, and Wadhams (2013) study parameterizes using 3 distributions and different number of years for exhausting the sequestered stock of methane. The original Excel version of the model uses the Beta+20 parameters. The Ceronsky et al. (2011) is based on the Beta distribution and assumes different sequestered stocks.

3.4.2 Permafrost carbon feedback

The parameterization is sourced from three papers: Kessler (2017)[5], Hope and Schaefer (2016) and Yumashev et al. (2019). Table 3.11 provides the relevant parameterizations. Kessler's estimates are based on a Normal distribution

[4] The ECS parameter is correlated with other temperature parameters and is described above.

[5] The original source.

Table 3.7: **Sampled correlations for the energy balance model parameters**

	xi_1	C_0	xi_3	f2xco2	t2xco2
xi_1	1				
C_0	-0.0184	1			
xi_3	-0.3383	-0.0528	1		
f2xco2	0.0095	-0.0064	-0.4505	1	
t2xco2	-0.1676	-0.0175	0.0623	0.6510	1

Table 3.8: **Static parameters for the SAF model**

Parameter	Name	Value
$\overline{SAF}$	ECS_AVGSAF	0.349590471
F_{sl}	forcingSlope	5.5
β_2	SAF_MEAN_QUAD_SEG_2	-0.001958477
β_1	SAF_MEAN_QUAD_SEG_1	-0.006387192
β_0	SAF_MEAN_QUAD_SEG_0	0.363696355
γ	SAF_STD_QUAD_SEG	0.109096643
α	Linear_SAF_MEAN	0.063605129
σ	Linear_SAF_SDEV	0.029077492

Table 3.9: **Stochastic parameters for the SAF model**

Parameter	Name	Distribution	Mean
δ	nonLinearSAF	Triang(-1,0,1)	0
FRT	warmingHalfLife	Triang(10,20,55)	28.33

with the mean and standard deviation in the rows 'Central' and 'Std. dev.', respectively. The '2.5%' and '97.5%' rows represent the relevant values for the confidence intervals. The values for the other two studies were estimated using least squares by Dietz et al. (2021). The Excel file uses the parameterization of Hope and Schaefer (2016). In principle the Kessler et al. parameterization can be sampled and the @Risk-based sampling formulas are provided in column 'H' of the Excel file.

3.4.3 Amazon rainforest dieback

The hazard rate is taken from Cai, Lenton, and Lontzek (2016).[6] It is equal to 0.00163443. The stock level is assumed to be 183.33 gtCO_2. The Excel sheet assumes that the number of years to deplete the sequestered carbon is a Triangular distribution with a minimum of 10 years, a maximum of 250 years and a mode of 50 years. The current version of the Excel worksheet does not sample this random variable and uses 50 years as the default.

3.4.4 The Greenland icesheet

Nordhaus estimates the melt rate to be -0.005303 with a standard deviation of $2.44E-5$.[7] Both are divided by 5/100—Nordhaus has a 5-year time-step and his volume is in percent. The Excel file allows for the melt rate to be sampled using a Normal distribution with a mean of $-1.06E-5$ and a standard deviation of $4.88E-8$ (or $4.88E-7$ as in the Excel file). The Excel file uses the central estimate of $-1.06E-5$, and does not make use of the sampling. The file also contains a Nordhaus estimate from a study by Robinson, Calov, and Ganopolski (2012) of $-8.8E-6$,[8] the central estimate times 1.1, and the 2.5% and 97.5% confidence range from the Normal distribution.

3.4.5 The West Arctic icesheet

The hazard rate estimate is 0.0043, based on Kriegler et al. (2009). The annual rate of sea-level rise is assumed to be log-Normal with a mean of 3.3 and a standard deviation of 1.65, based on Bamber and Aspinall (2013). Note that the sampled deviate is divided by 1000. In the case of the WAIS module, the sampling is used in the default configuration.

3.4.6 Sea-level rise

Sea-level rise from thermal expansion and glaciers and small ice caps is initialized at 0.04 based on Church and White (2011). The annual melt rates are taken from Diaz and Keller (2016), respectively 0.00079 for thermal expansion

[6] Based on expert elicitation in Kriegler et al. (2009) according to Dietz et al. (2021).

[7] See the supplemental information.

[8] Which is derived from Nordhaus' 5-year time step estimate of -0.0044

Table 3.10: **Parameters for the OMH model**

Source	Years	Stock (mt)	b^{OMH}
Whiteman et al.			
Uniform	10	50,000	1.457
Triangular (mode = 10%)	10	50,000	0.068
Beta	10	50,000	0.093
Uniform	20	50,000	1.290
Triangular (mode = 10%)	20	50,000	0.977
Beta	20	50,000	0.118
Beta	30	50,000	0.178
Ceronsky et al.			
Beta	150	30,000	0.365
Beta	150	267,600	0.244
Beta	150	1,170,000	0.163

Table 3.11: **Parameters for the PCF model**

Source	CH_4 proportion	β^{pf}	C^{pf} (gtC)	Proportion passive	Decomposition Rate (yrs)
Kessler					
2.5%	1.1%	0.121	885	0.2922	11.2
Central	2.3%	0.172	1035	0.4	70.0
Std. dev.	0.006	0.0261	76.53	0.055	30
97.5%	3.5%	0.223	1185	0.5078	128.8
Hope and Schaefer	2.3%	0.066	1160	0.369	30.6
Yumashev et al.	2.3%	0.086	1066	0.407	66.4

and 0.0008164 for glaciers and small ice caps. The Diaz and Keller (2016) estimates are divided by 10 to convert to annual time steps. There is no sampling vis-à-vis these parameters.

3.4.7 Slowdown of the Atlantic Meridional Overturning Circulation

Each set of parameters is indexed by one of the four GCMs. In the Excel file these are labeled 'Hadley', 'BCM', 'IPSL' and 'HADCM'. The user selects one of the four for any given simulation. The hazard rates for each of the four GCMs, estimated by Dietz et al. (2021), are provided in Table 3.12. The long-run impacts on country temperatures are also GCM- and country-specific and are provided in the Excel file (`Parameters!A130:GM133`). The long-run is defined as 35 years once the slowdown starts. There is no sampling for these parameters.

Table 3.12: **Hazard rates for AMOC**

GCM	Rate
Hadley	0.135
BCM	0.611
IPSL	0.54
HADCM	1.6

3.4.8 Tipping point interactions

The parameterization of the interactions module is based on the expert elicitation of Kriegler et al. (2009). The parameters are summarized in Table 3.13. All parameters are assumed to be normally distributed with the means and standard deviations provided in the table. The impacts for each tipping point is read down a column—for

example, the interaction impact for the Amazon dieback has the following weights: 1.05547 (GIS), 1.0 (WAIS), 1.42157 (AMOC) and 1.84143 (ISM).

Table 3.13: **Tipping point interaction parameters**

Source impact	AMAZ	GIS	WAIS	AMOC	ISM
Mean					
AMAZ		1.0	1.0	0.990857	1.13075
GIS	1.05547		1.50381	1.70282	1.0
WAIS	1.0	1.13209		0.984014	1.0
AMOC	1.42157	0.681508	1.02462		1.0003
ISM	1.84143	1.0	1.00003	0.962033	
Std. dev.					
AMAZ		8.66076e-07	7.61999e-07	0.0912496	0.417844
GIS	0.0463336		0.141193	0.16792	6.67783e-07
WAIS	3.44405e-06	0.0845544		0.0816243	5.63738e-07
AMOC	0.340355	0.0836145	0.0370685		0.00145894
ISM	0.265727	3.7256e-06	0.000143963	0.0435226	

3.4.9 Sampling strategy for tipping points

Table 3.14 presents the sampling strategy for the uncertain parameters for the various tipping points. The underlying distribution functions are relatively regular and there are very little deviations between the targeted and samples modes. None of these parameters is subject to a targeted correlation.[9]

3.5 Economic module parameters

Table 3.15 presents the sampling strategy for the uncertain parameters for the economic module. All three of the uncertainty parameters follow a triangular distribution and there is virtually no deviation between the targeted and sampled modes.

In addition to these global variables, there are also two sets of country-specific uncertain parameters. The linear damage parameter for sea-level rise, χ^{slr}, is assumed to be characterized by a Triangular distribution, with country-specific estimates for the total range and the mode. The damage from changing temperatures is specified using a quadratic function. Both the linear and quadratic coefficients, respectively β_1 and β_2 are assumed to be characterized by a Normal distribution, with country-specific estimates of the mean and standard deviation. In addition, Dietz et al. (2021) provides a correlation matrix for all pairs of the coefficients, i.e., $\rho^{i,j}_{c,c'}$, where i and j take the values of 1 and 2, respectively for β^1 and β^2, and c, c' represents all country pairings.

[9] The AMOC tipping point is not subject to a sampling strategy—though there are four sets of estimates for the parameters provided by four GCMs.

Table 3.14: **Sampling assumptions for the tipping point parameters**

X	Code	Distribution	Expected		Sample	
			mu	σ	mu	σ
φ^{ch4}	propCH4	Normal(0.023,0.006)	0.023	0.006	0.023	0.006
β^{pf}	beta_PF	Normal(0.172,0.026)	0.172	0.026	0.172	0.026
C^{pf}	C_PF	Normal(1035,76.53)	1035	76.53	1035	76.55
χ^{psv}	propPassive	Normal(0.4,0.055)	0.4	0.055	0.4	0.055
τ^{psv}	Tau_PF	Normal(70,30)	70	30	70	30
$NYEAR^{AMAZ}$	NYear_AMAZ	Triangular(10,50,250)	103.3	52.5	103.3	52.5
β^{GIS}	Beta_GIS	Normal(-1.06e-5,4.9e-8)	-1.06e-5	4.9e-8	-1.06e-5	4.9e-8
R^{WAIS}	Rate_WAIS	LogNormal(0.0033,0.0017)	0.0033	0.0017	0.0033	0.0017

Table 3.15: **Sampling assumptions for the economic module parameters**

			Expected		Sample	
X	Code	Distribution	mu	σ	mu	σ
φ	`yDamWeight`	`Triangular`	0.5	0.2041	0.5	0.2041
η	`eta`	`Triangular`	1.3333	0.3118	1.3333	0.3118
ρ	`rho`	`Triangular`	0.0103	0.0039	0.0103	0.0039

Chapter 4

Monte Carlo Simulations

This section describes a set of Monte Carlo simulations coupled to the LHS utility. It uses RCP 4.5 for the reference emission trends and SSP2 for the reference population and GDP trends. The global circulation model (GCM) parameterization is taken to be IPSL. The simulation is undertaken 10,000 times. For each sample, the simulation is run twice—the first time with no extra carbon pulse, and the second time with a pulse of 10gtCO_2 in 2020. The uncertain events—for example linked to the start of the tipping points—are only sampled in the first simulation. The with and with-out pulse simulations are used to calculate the social cost of carbon. The results described below are focused on the role of the tipping points. There is one set of simulations with the tipping points active and another set of simulations with the tipping points inactive.

4.1 Sampling

The sampling is done with the LHS utility using the full distributional assumptions from the META 21 model. There are 9 uncertain parameters in the carbon cycle module—some of which have an identified correlation matrix; there are 6 uncertain parameters in the forcing and EBM module—again, some of which are assumed to be correlated; there are 8 independent uncertain tipping point parameters; and there are 3 uncertain parameters in the economic module. The damage module contains a very large number of uncertain parameters due to the fact that the damage parameters—of which there are three—are country specific. In addition the sampling takes into account the correlation matrix for the temperature-based damage coefficients. The input file for the LHS utility has over 75,000 lines in order to take into account the very large correlation matrix across the country-based temperature damage coefficients. Despite the size of the resulting LHS cube, the software operates in under 13 seconds. The resulting cube has 582 random variables sampled 10,000 times.

4.2 Results

This section describes some of the key results from the Monte Carlo simulations of the META 21 model—including assessing the relative role of including the tipping points in the analysis.

4.2.1 The social cost of carbon

With the current configuration for the simulations with tipping points, the simulation fails (for numerical) reasons for 183 observations, i.e. just under 2% of the sample size. Figure 3 presents the histogram for the remaining SCC values greater than 0 and less than $500. This removes some 17 negative values, of which 10 are less than $100 in absolute terms, and 51 values above $500, of which 6 are greater than $1000. There are also some unusual pairings and it might be useful to drill down to see if there are plausible explanations. For example, sample number 1885 has an SCC of $82 without the tipping points, but over $45,000 when the tipping points are included. Another extreme example is for sample number 8947. With tipping points, the SCC is -$89 and around -$6.5 million without the tipping points.

Figure 4 summarizes the SCC for the two configurations—with and without the tipping points. The median value of the SCC with tipping points is $60, and without tipping points is $48. The latter is somewhat lower than the value of $52 in Dietz et al. (2021), part of which could be attributed to differences in prices and exchange rates,

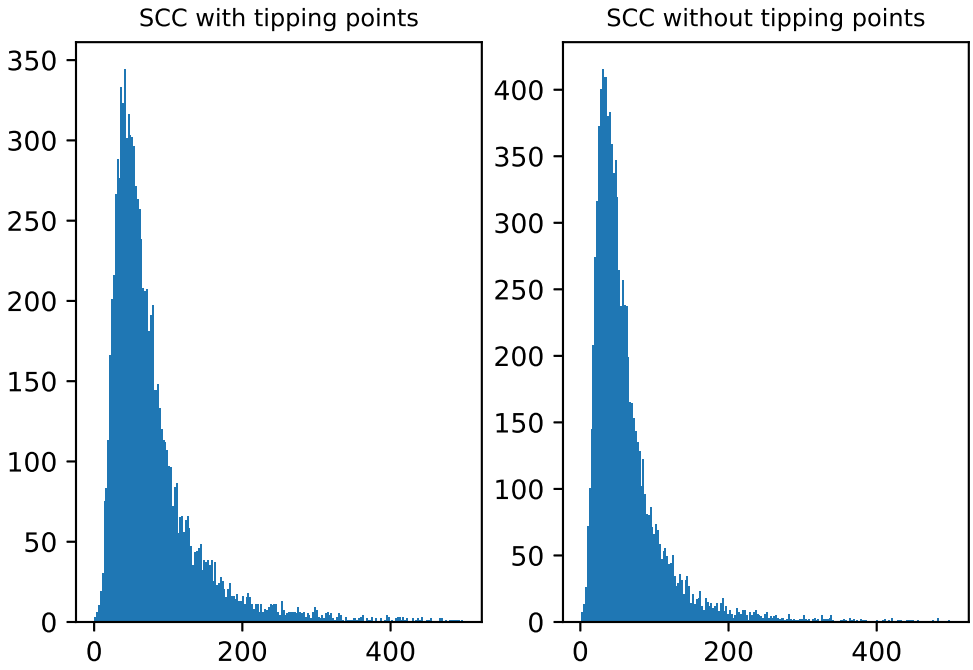


Figure 3: **Histogram of SCC, $2010 per ton** CO_2

Tipping point	Average start year	Frequency
AMOC	2012	9749
OMH	2016	9743
WAIS	2058	9743
AMAZ	2110	4483

Table 4.1: **Tipping points start year**

we are using $2010 prices and purchasing power parity (PPP) exchange rates,[1] with the remainder attributable to changes in sampling strategy and removal of outliers. Thus, the inclusion of these tipping points raises the estimate of the SCC by around 26%, similar in magnitude to the increase in Dietz et al. (2021).

4.2.2 Tipping points start year

Table 4.1 describes the average start year for the uncertain tipping points. The parameterization assumes that the AMOC and OMH starting points are expected to start almost immediately, with the average start year for WAIS around 2058 and a much later start year for the Amazon dieback. The table also provides the frequency of the tipping point within the 2010/2200 time frame over the 10,000 samples. With the exception of the Amazon dieback, the tipping points are present in almost the entire sample, with the Amazon dieback occurring in roughly 50% of the samples. The distribution of the start years can be gleaned from Figure 5. The start years are heavily skewed towards the left, with the exception of the Amazon dieback, which has a relatively uniform distribution.[2]

[1] Dietz et al. (2021) inflate the SCC results by 20% to render the values in 2020 prices and exchange rates. Using the same factor would push our SCC value to $57.

[2] N.B. The x-axis differs across the tipping points.

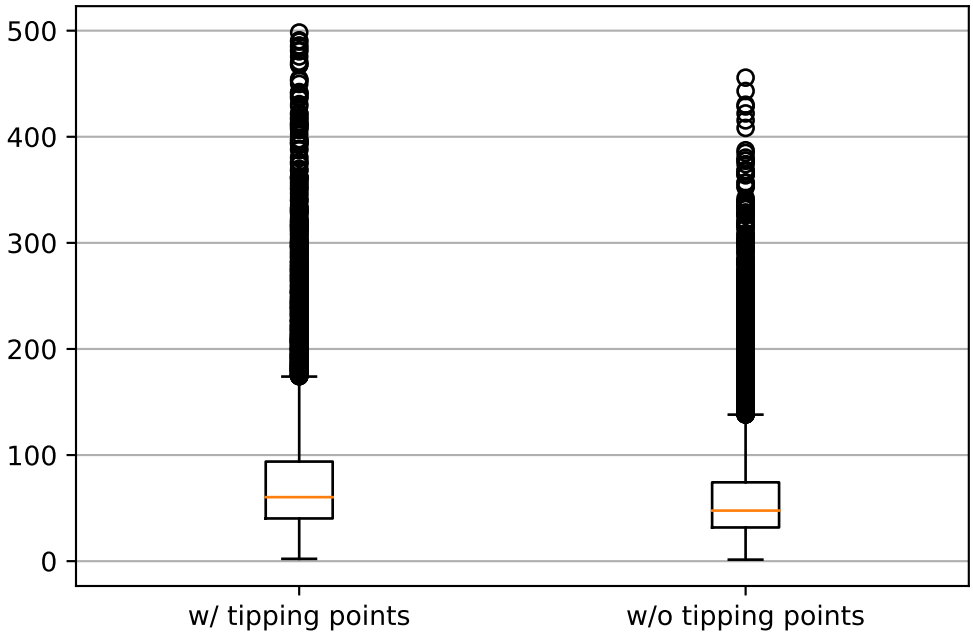


Figure 4: **Box and whisker plot of SCC, $2010 per ton** CO_2

4.2.3 Change in temperature

Figure 6 shows the change in temperature across the two configurations. The outer-most curves show the maximum and minimum temperature profiles across the 10,000 samples. The inner curves show the 95 percentile range. The middle curve shows the average. With the tipping points, the average temperature change could reach 4 °C by 2200, with a probability range between 2-6 °C, and a maximum of 10 °C. Without the tipping points, the ranges are more narrow—notably because 3 of the 5 tipping points have a direct impact on atmospheric concentrations of greenhouse gases.

4.2.4 Impacts on consumption

We conclude by assessing the impacts of damages and the tipping points on consumption. Figure 7 shows the average change in consumption in 2100 across countries relative to a no-damage baseline. There are a handful of countries that could benefit on average from rising temperatures—most visibly in the upper latitudes, e.g., Canada, Finland, Mongolia and Russia, but also China and the United States. The countries that would perceive the highest negative shocks are not visible as these are mostly small island nations. Other countries that would suffer damages in excess of 10% include Guyana and Suriname in South America, and Cameroon, Senegal, Sierra Leone, and Tunisia in Africa.

Figure 8 shows the impacts of excluding the tipping points in the assessment of the average damages. The impact tends to be concentrated in the range of lower damages of between 0 and 2% relative to the model outcomes when the tipping points are included. There are some countries that may see improvement when tipping points are included, e.g., land-locked countries in Central Asia.

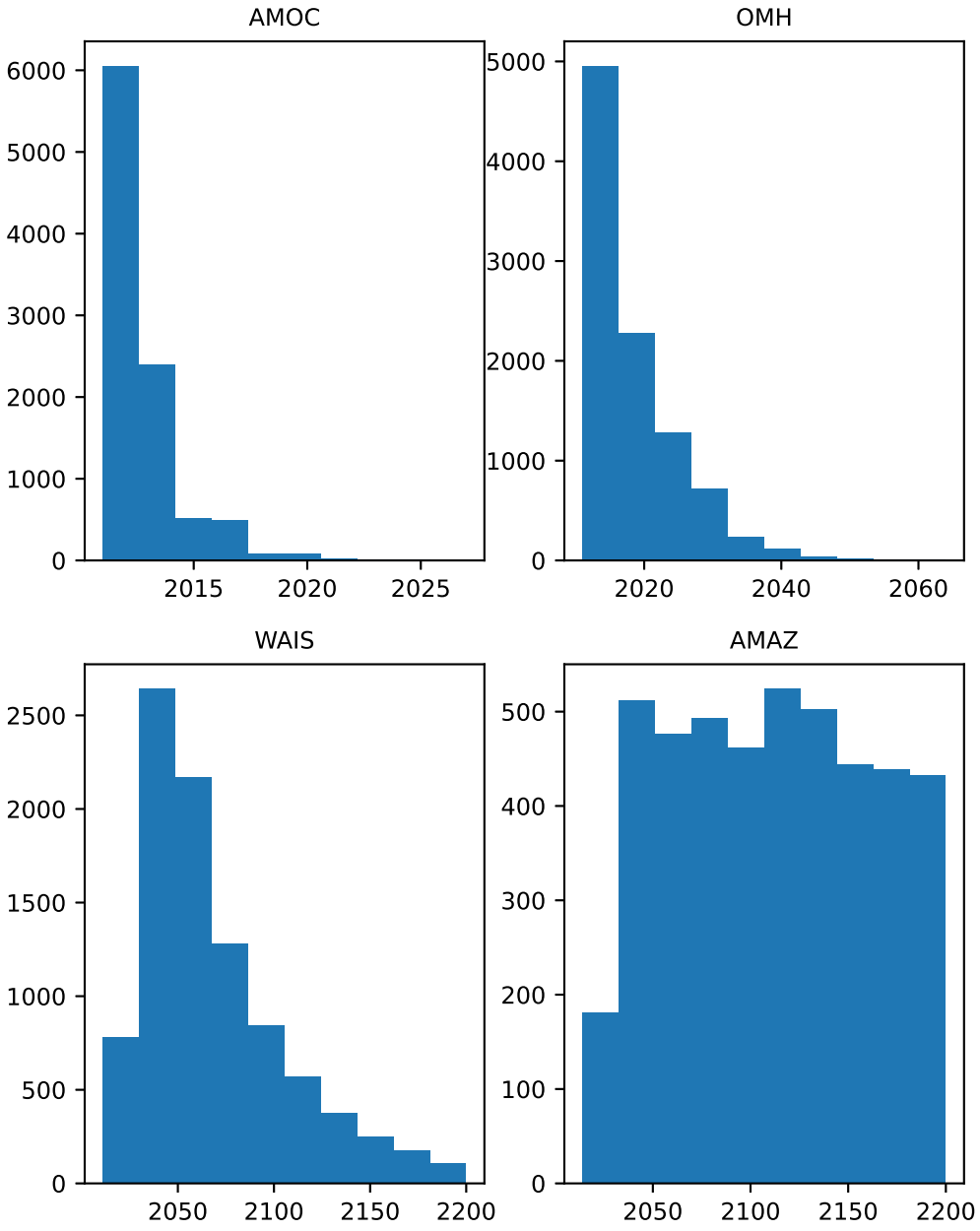


Figure 5: **Tipping points start year**

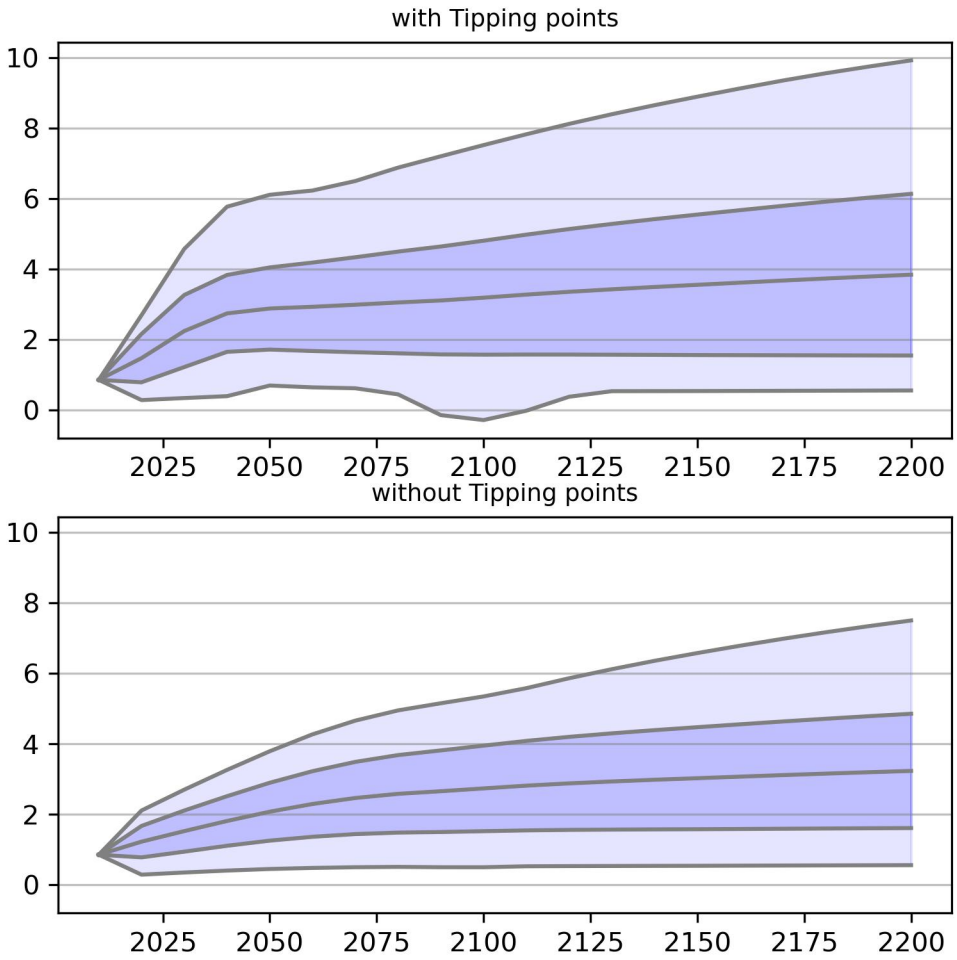


Figure 6: **Change in temperature, °C**

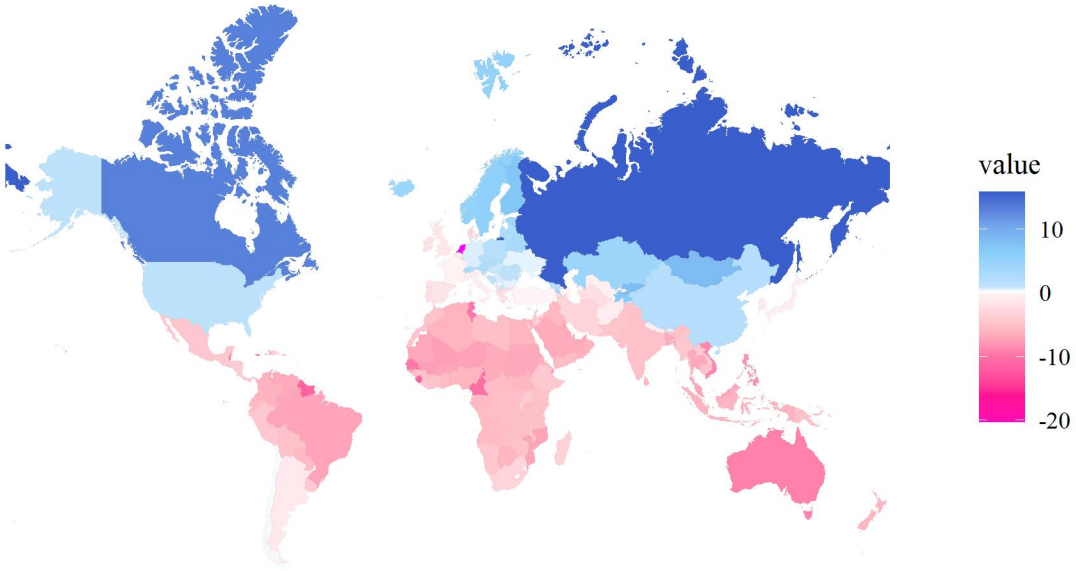


Figure 7: **Average change in consumption in 2100 with tipping points, percent**

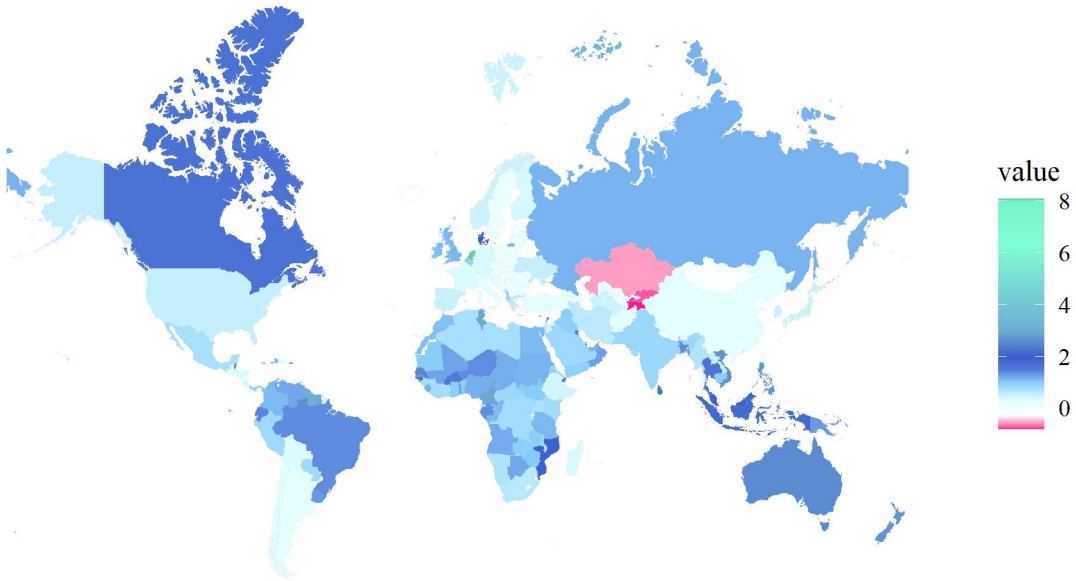


Figure 8: **Average change in consumption in 2100 when excluding tipping points, percentage point difference with the 'full' damages scenario**

Chapter 5

Conclusion

This paper lays out the basic structure of the META 21 Integrated Assessment Model and its implementation in GAMS. It describes how to use the LHS utility to generate a random sample that covers the full range of the underlying distributions and conforms to the expected correlations across the sampled uncertain parameters. The economic and greenhouse gas emission components are largely exogenous in META 21—with the former impacted by climate-related damages and the latter adjusted by tipping-point related events. It would nonetheless be relatively straightforward to graft many, if not most, of the components of META 21 to more complex economic models—such as general equilibrium models. Examples of these include the Envisage model (van der Mensbrugghe, 2019), which has the carbon cycle, temperature and damage functions, though not the tipping point elements. The META 21 model also highlights how one could introduce extreme events in economic models—such as droughts and floods—using the hazard rate approach, where the hazard rate is a function of changing temperatures.

Bibliography

Anthoff, D., and R.S.J. Tol. 2014. "The Climate Framework for Uncertainty, Negotiation and Distribution (FUND), Technical Description, Version 3.9." Mimeo. http://www.fund-model.org/files/documentation/Fund-3-9-Scientific-Documentation.pdf.

Bamber, J.L., and W.P. Aspinall. 2013. "An expert judgement assessment of future sea level rise from the ice sheets." *Nature Climate Change*, 3: 424–427. doi:10.1038/nclimate1778.

Belaia, M., M. Funke, and N. Glanemann. 2017. "Global Warming and a Potential Tipping Point in the Atlantic Thermohaline Circulation: The Role of Risk Aversion." *Environmental and Resource Economics*, 67(1): 93–125. doi:10.1007/s10640-015-9978-x.

Burke, M., S.M. Hsiang, and E. Miguel. 2015. "Global non-linear effect of temperature on economic production." *Nature*, 527: 235–239. doi:10.1038/nature15725.

Cai, Y., T.M. Lenton, and T.S. Lontzek. 2016. "Risk of multiple interacting tipping points should encourage rapid CO_2 emission reduction." *Nature Climate Change*, 6: 520–525. doi:10.1038/nclimate2964.

Ceronsky, M., D. Anthoff, C. Hepburn, and R. Tol. 2011. "Checking the Price Tag on Catastrophe: The Social Cost of Carbon Under Non-linear Climate Response." Economic & Social Research Institute (ESRI), Dublin, ESRI Working Paper No. 392. https://www.esri.ie/publications/checking-the-price-tag-on-catastrophe-the-social-cost-of-carbon-under-non-linear.

Church, J.A., and N.J. White. 2011. "Sea-Level Rise from the Late 19th to the Early 21st Century." *Surveys in Geophysics*, 32: 585–602. doi:10.1007/s10712-011-9119-1.

Corong, E., T. Hertel, R. McDougall, M. Tsigas, and D. van der Mensbrugghe. 2017. "The Standard GTAP Model, Version 7." *Journal of Global Economic Analysis*, 2(1): 1–119. doi:10.21642/JGEA.020101AF.

Dellink, R., J. Chateau, E. Lanzi, and B. Magné. 2017. "Long-term economic growth projections in the Shared Socioeconomic Pathways." *Global Environmental Change*, 42: 200–214. doi:10.1016/j.gloenvcha.2015.06.004.

Diaz, D., and K. Keller. 2016. "A Potential Disintegration of the West Antarctic Ice Sheet: Implications for Economic Analyses of Climate Policy." *American Economic Review*, 106(5): 607–611. doi:10.1257/aer.p20161103.

Diaz, D.B. 2016. "Estimating global damages from sea level rise with the Coastal Impact and Adaptation Model (CIAM)." *Climatic Change*, 137: 143–156. doi:10.1007/s10584-016-1675-4.

Dietz, S., J. Rising, T. Stoerk, and G. Wagner. 2021. "Economic impacts of tipping points in the climate system." *Proceedings of the National Academy of Sciences*, 118(34). doi:10.1073/pnas.2103081118.

Helton, J.C., and F.J. Davis. 2002. "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems." Sandia National Laboratories, Sandia Technical Report No. SAND2001-0417. doi:10.2172/806696.

Hope, C., and K. Schaefer. 2016. "Economic impacts of carbon dioxide and methane released from thawing permafrost." *Nature Climate Change*, 6: 56–59. doi:10.1038/nclimate2807.

Horridge, M., and K. Pearson. 2011. "Systematic Sensitivity Analysis with Respect to Correlated Variations in Parameters and Shocks." Global Trade Analysis Project (GTAP), Department of Agricultural Economics, Purdue University, West Lafayette, IN, GTAP Technical Paper No. 30. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=3496.

Iman, R.L., and W.J. Conover. 1982. "A distribution-free approach to inducing rank correlation among input variables." *Communications in Statistics - Simulation and Computation*, 11(3): 311–334. doi:10.1080/03610918208812265.

Interagency Working Group on Social Cost of Carbon. 2016. "Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866." United States Government, Report. https://www.epa.gov/sites/production/files/2016-12/documents/sc_co2_tsd_august_2016.pdf.

Joos, F., R. Roth, J.S. Fuglestvedt, G.P. Peters, I.G. Enting, W. von Bloh, V. Brovkin, E.J. Burke, M. Eby, N.R. Edwards, T. Friedrich, T.L. Frölicher, P.R. Halloran, P.B. Holden, C. Jones, T. Kleinen, F.T. Mackenzie, K. Matsumoto, M. Meinshausen, G. Plattner, A. Reisinger, J. Segschneider, G. Shaffer, M. Steinacher, K. Strassmann, K. Tanaka, A. Timmermann, and A.J. Weaver. 2013. "Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis." *Atmospheric Chemistry and Physics*, 13(5): 2793–2825. doi:10.5194/acp-13-2793-2013.

KC, S., and W. Lutz. 2017. "The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100." *Global Environmental Change*, 42(42): 181–192. doi:10.1016/j.gloenvcha.2014.06.004.

Kessler, L. 2017. "Estimating the economic impact of the permafrost carbon feedback." *Climate Change Economics*, 8(2): 1–23. https://www.jstor.org/stable/90009412.

Kriegler, E., J.W. Hall, H. Held, R. Dawson, and H.J. Schellnhuber. 2009. "Imprecise probability assessment of tipping points in the climate system." *Proceedings of the National Academy of Sciences*, 106(13): 5041–5046. doi:10.1073/pnas.0809117106.

Leach, N.J., S. Jenkins, Z. Nicholls, C.J. Smith, J. Lynch, M. Cain, T. Walsh, B. Wu, J. Tsutsui, and M.R. Allen. 2021. "FaIRv2.0.0: a generalised impulse response model for climate uncertainty and future scenario exploration." *Geoscientific Model Development*, 14(5): 3007–3036. doi:10.5194/gmd-14-3007-2021.

Manne, A.S., and R.G. Richels. 2005. "Merge: An Integrated Assessment Model for Global Climate Change." In *Energy and Environment*, edited by R. Loulou, J.-P. Waaub, and G. Zaccour. Boston, MA: Springer US, pp. 175–189. doi:10.1007/0-387-25352-1_7.

Matsumoto, M., and T. Nishimura. 1998. "Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator." *ACM Transactions on Modeling and Computer Simulation*, 8(1): 3–30. doi:10.1145/272991.272995.

McKay, M.D., R.J. Beckman, and W.J. Conover. 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics*, 21: 239–245. doi:10.1080/00401706.1979.10489755.

Millar, R.J., Z.R. Nicholls, P. Friedlingstein, and M.R. Allen. 2017. "A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions." *Atmospheric Chemistry and Physics*, 17(11): 7213–7228. doi:10.5194/acp-17-7213-2017.

Nishimura, T. 2000. "Tables of 64-Bit Mersenne Twisters." *ACM Transactions on Modeling and Computer Simulation*, 10(4): 348–357. doi:10.1145/369534.369540.

Nordhaus, W. 2019. "Economics of the disintegration of the Greenland ice sheet." *Proceedings of the National Academy of Sciences*, 116(25): 12261–12269. doi:10.1073/pnas.1814990116.

Nordhaus, W.D. 2017. "Revisiting the social cost of carbon." *Proceedings of the National Academy of Sciences*, 114(7). doi:10.1073/pnas.1609244114.

O'Neill, B.C., E. Kriegler, K.L. Ebi, E. Kemp-Benedict, K. Riahi, D.S. Rothman, B.J. van Ruijven, D.P. van Vuuren, J. Birkmann, K. Kok, M. Levy, and W. Solecki. 2017. "The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century." *Global Environmental Change*, 42: 169–180. doi:10.1016/j.gloenvcha.2015.01.004.

Robinson, A., R. Calov, and A. Ganopolski. 2012. "Multistability and critical thresholds of the Greenland ice sheet." *Nature Climate Change*, 2: 429–432. doi:10.1038/nclimate1449.

Schewe, J., and A. Levermann. 2012. "A statistically predictive model for future monsoon failure in India." *Environmental Research Letters*, 7(4): 044023. doi:10.1088/1748-9326/7/4/044023.

Swiler, L.P., and G.D. Wyss. 2004. "A User's Guide to Sandia's Latin Hypercube Sampling Software: LHS UNIX Library Standalone Version." Sandia National Laboratories, Sandia Technical Report No. SAND2004-2439. doi:10.2172/919175.

van der Mensbrugghe, D. 2019. "The Environmental Impact and Sustainability Applied General Equilibrium (ENVISAGE) Model. Version 10.01." Center for Global Trade Analysis (GTAP), Department of Agricultural Economics, Purdue University, West Lafayette, IN, Unpublished manuscript. https://mygeohub.org/groups/gtap/envisage-docs.

van der Mensbrugghe, D. 2023a. "A Summary Guide to the Latin Hypercube Sampling (LHS) Utility." Global Trade Analysis Project (GTAP), Department of Agricultural Economics, Purdue University, West Lafayette, IN, GTAP Working Paper No. 94. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=7035.

van der Mensbrugghe, D. 2023b. "Using Python for Parallelization." Global Trade Analysis Project (GTAP), Department of Agricultural Economics, Purdue University, West Lafayette, IN, GTAP Working Paper No. 93. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=6826.

van Vuuren, D.P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G.C. Hurtt, T. Kram, V. Krey, J.F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S.J. Smith, and S.K. Rose. 2011. "The representative concentration pathways: an overview." *Climatic Change*, 109: 5–31. doi:10.1007/s10584-011-0148-z.

Whiteman, G., C. Hope, and P. Wadhams. 2013. "Vast costs of Arctic change." *Nature*, 499: 401–403. doi:10.1038/s41467-019-09863-x.

Yumashev, D., C. Hope, K. Schaefer, K. Riemann-Campe, F. Iglesias-Suarez, E. Jafarov, E.J. Burke, P.J. Young, Y. Elshorbany, and G. Whiteman. 2019. "Climate policy implications of nonlinear decline of Arctic land permafrost and other cryosphere elements." *Nature Communications*, 10: 1900. doi:10.1038/s41467-019-09863-x.

Appendix A

The LHS utility

This appendix provides an overview of the LHS algorithm including a small numerical example, and an introduction the LHS utility, which generates the random deviates—ordering the sample to a user-specified correlation matrix (or the identity matrix in the absence of a target). A summary description of the LHS utility is provided as a separate User Guide. A complete guide to the original LHS utility is available in Swiler and Wyss (2004). Note that the former does not include a discussion of user-defined distributions and the latter does not include information on the new features in the LHS utility such as the new distributions and the additional output options.

A.1 LHS algorithm

Monte Carlo methods are well established techniques to explore model or simulation uncertainty when the underlying model is based on uncertain parameters (but with an assumed distribution). For each uncertain parameter, a draw of a random deviate is taken from the known distribution—and if there are k uncertain parameters, each draw will consist of selecting a random deviate for each of the k uncertain parameters. Latin Hypercube Sampling (LHS) (McKay, Beckman, and Conover, 1979) was introduced as an efficient sampling method, where efficiency allows for limiting the size of the sample while at the same time insuring that the entire sample space is covered.

Before explaining LHS, we describe first the technique for generating random deviates. If a distribution has an easily invertible cumulative distribution function (CDF), the first step is to generate a random deviate from the Uniform distribution over the range $[0, 1]$ and then to use the inverse function to generate the random deviate.[1] This is the so-called inverse transform method.

We highlight the technique using the Logistic distribution, which has a readily invertible CDF. Its CDF is given by:

$$F(x) = \frac{1}{1 + e^{-(x-\mu)/s}}$$

where μ is the location parameter and s is the scale parameter. The inverse of the CDF, also known as the quantile function, is easily derived for any value of F, say p:

$$x = Q(p) = \mu + s \ln\left(\frac{p}{1-p}\right)$$

Figure 9 plots an example Logistic CDF with parameters $\mu = 9$ and $s = 3$. We randomly select 5 deviates from the Uniform distribution: $0.12, 0.36, 0.54, 0.79, 0.95$, visible along the y-axis. We then generate the random deviates for the Logistic distribution by locating the Uniform deviates on the y-axis of the CDF and evaluating the corresponding Logistic deviates, respectively 3.0, 7.3, 9.5, 13.0, and 17.8, easily derived from the Logistic quantile function.

In the standard Monte Carlo method, a random deviate is sampled each time from the entire domain of the random variable, i.e. over $[0, 1]$. In the case of selecting a sample of 5, nothing guarantees that the sample will cover the full range. For example, 4 of the numbers could cluster around 0.65 with a fifth at 0.35. LHS sampling breaks the $[0, 1]$ range into 5 equal segments of width 0.2 (for a sample of 5). Sampling is done within each of the sub-ranges, guaranteeing that the sample will cover (almost) the entire range. In the more general case, with a sample size

[1] Not all CDF's have an easy expression for their corresponding quantile function, including the Normal distribution. In the case of the latter, the inverse error function is readily available in most software packages. The Beta distribution is an example where the inverse CDF is not readily available. In these cases, other numerical techniques are deployed to generate random deviates.

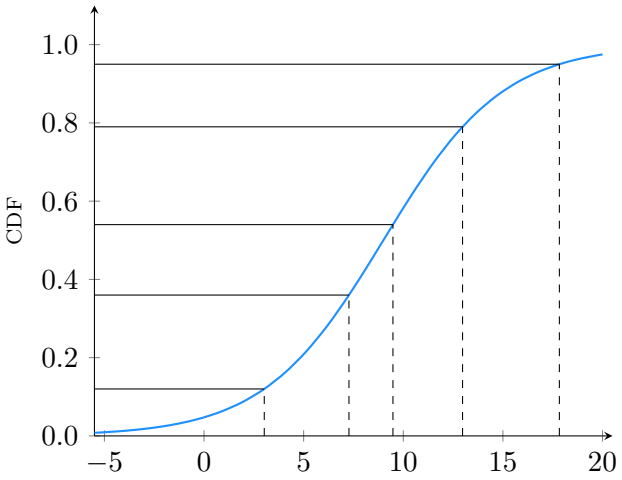


Figure 9: **The Logistic distribution CDF and selecting random deviates**

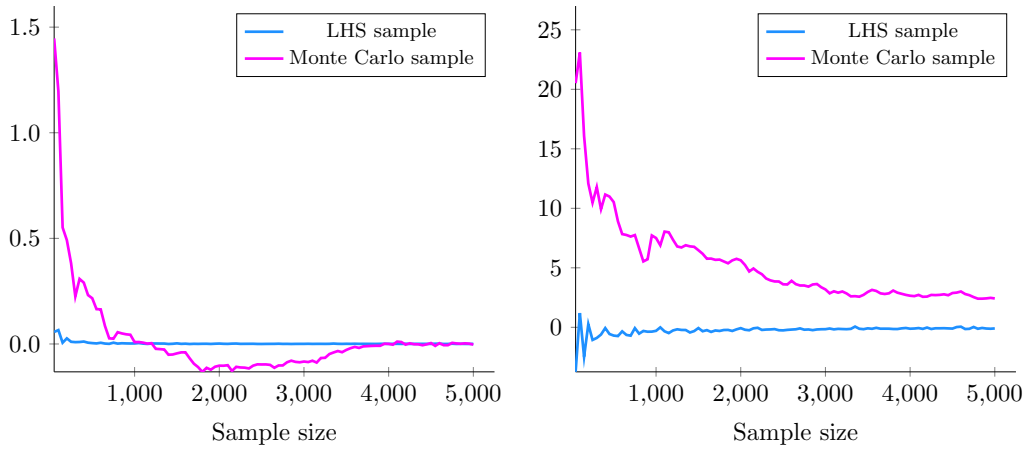
of n, LHS sampling divides the $[0, 1]$ range into equivalent sub-ranges with a width of $1/n$ and samples within the sub-ranges. The efficiency argument is that LHS allows to have a smaller sample size than standard Monte Carlo sampling, while ensuring that the full potential sample range is sampled. The technique is implemented for each random variable.

Figure 10 highlights a simple example of the potential benefits of LHS sampling.[2] The example uses the Logistic distribution with a location parameter of 0 and a scale parameter of 2. The mean of this distribution is 0 and the standard deviation is 3.628.[3] The x-axis of each graph, shows the number of draws for each sample, ranging from 50 to 5000 in steps of 50. The left-panel shows the deviation of the sample mean from the expected mean of 0 and the right-panel shows the percent deviation of the standard deviation from the expected standard deviation. Both figures clearly indicate that the LHS samples very quickly converge to the expected moments of the distribution. The Monte Carlo simulations show persistent deviations in the standard deviation, even with 5,000 draws.[4]

[2] The GAMS code that generates the data for the figure is available upon request.

[3] $\sigma = \sqrt{\frac{(\pi s)^2}{3}}$, where s is the scale parameter.

[4] The advantage of LHS sampling with many uncertain parameters is disputed, though part of the issue relates to memory and computer time, which is disappearing as an issue. A blog post by Lonnie Chrisman (https://lumina.com/latin-hypercube-vs-monte-carlo-sampling/) summarizes part of this discussion and provides additional references.



(a) **Deviation from the expected mean (0)** (b) **Deviation from expected standard deviation, %**

Figure 10: **Comparison of LHS versus Monte Carlo sampling for a Logistic(0,2) distribution**

As coded in the LHS utility, the sample for each variable is in ascending order, as the samples are derived from each sub-range in ascending order. The deviates could be further randomized by sampling without replacement each of the deviates of each variable. The LHS utility does something a bit different. In the absence of user-specified correlations, it re-orders each column of deviates in such a way that the resulting correlation matrix across variables is (nearly) the identity matrix, while preserving the marginal distribution of each column. This is called the Iman-Conover method (Iman and Conover, 1982). The method is explained intuitively and with a numerical example in the Appendix. The same method is used if the user specifies a correlation matrix for all or some of the variables—a correlation of 0 is assumed for all pairs of variables for which a correlation is not user-provided.

In summary, the LHS method for Monte Carlo sampling, uses a method of stratified sampling, which is an efficient method to generate a sample that is nearly sure to cover the entire sample space. In addition, the deviates of each variable are ordered in such a way to (nearly) match a desired correlation scheme, which in the absence of any user-imposed correlations is the identity matrix.

A.2 A numerical example of the LHS algorithm

The origin of LHS is attributed to the work of McKay (McKay, Beckman, and Conover, 1979) and Iman and Conover (Iman and Conover, 1982), which is very nicely described by Helton and Davis (Helton and Davis, 2002). LHS has two components. The first is intended to cover as much of the sampling space as possible by using a simplified version of stratified sampling. The key idea behind this is to reduce the number of samples—particularly for complex models. In random sampling, each random deviate is selected over the entire range of possible outcomes, i.e., over the probability range of $[0, 1]$. The LHS technique divides the $[0, 1]$ range into n sub-intervals of equal length set to $1/n$, where n is the number of deviates being requested in the sample. In this sense, LHS covers the entire probability range. For example, for a sample size of 10, the procedure is certain to select a deviate from the $[0, 0.1]$, whereas their is no guarantee of this outcome from pure random sampling. This simplified stratified sampling is used for each random variable, and hence LHS is sure to cover the entire range of the multi-variate sample.

The second component is the pairing of the deviates so as to match a targeted correlation matrix for the random variables. In the absence of a user-specified correlation matrix, the algorithm pairs the deviates to target a zero-correlation matrix, or as close as possible. The pairing part of the algorithm has several components, which are described below. Let X be a matrix of the initial random deviates. It is of dimension $n \times k$, where n is the sample size and k is the number of random variables. Given the sampling strategy, each column of X is sorted in ascending order. Let C be the targeted correlation matrix, in the absence of a user-specified list of correlations, it is the identity matrix. It has the dimensions of $k \times k$, is symmetric and semi-positive definite. Therefore, it can be split using the Cholesky factorization, which is a lower triangular matrix, say H, with the property $C = HH'$.[5]

[5] Horridge and Pearson (2011) use a similar technique for their method of systematic sensitivity analysis using

The second step involves developing an auxiliary matrix of random deviates, based on the so-called van der Waerden scores. This matrix will be transformed such that it has the same correlation as the target correlation, and once this is accomplished, its rank scores will be used to derive the same rank for the original sample.

The van der Waerden scores represent the inverse of the normal standard deviation and each column will have the same initial distribution, until fully randomized. For example, in the case of $n = 10$, each column has the following values until randomized:

$$\begin{bmatrix} -1.335178 \\ -0.908458 \\ -0.604585 \\ -0.348756 \\ -0.114185 \\ 0.114185 \\ 0.348756 \\ 0.604585 \\ 0.908458 \\ 1.335178 \end{bmatrix}$$

which is the inverse of the cumulative distribution function of the standard normal distribution at each cumulative interval defined by 1/11. After randomization, column by column, the matrix S, may look something like the following:

$$S = \begin{bmatrix} -0.11419 & -1.33518 & -0.34876 & -0.34876 & 0.90846 \\ 0.11419 & 0.34876 & 0.11419 & 0.34876 & -0.90846 \\ -0.90846 & 0.60459 & -0.90846 & -0.60459 & 0.60459 \\ -1.33518 & 0.11419 & 0.60459 & -0.90846 & -0.34876 \\ 0.90846 & -0.60459 & -1.33518 & -0.11419 & -0.11419 \\ 0.60459 & -0.34876 & 0.34876 & 0.90846 & 1.33518 \\ -0.34876 & 1.33518 & -0.60459 & 1.33518 & -1.33518 \\ 1.33518 & -0.11419 & 0.90846 & -1.33518 & -0.60459 \\ 0.34876 & 0.90846 & 1.33518 & 0.11419 & 0.11419 \\ -0.60459 & -0.90846 & -0.11419 & 0.60459 & 0.34876 \end{bmatrix}$$

for a sample of 5 random variables. From the matrix S, which to emphasize, has nothing to do with the sample itself, we can derive its correlation matrix $C^s = Corr(S)$, and the subsequent Cholesky decomposition ($C^s = H^s\,(H^s)'$) and the inverse of the Cholesky decomposition, call the latter $M = (H^s)^{-1}$. The following step transforms S to S^1:

$$S^1 = S\,(HM)'$$

where $(HM)'$ is a projection matrix, based on the targeted Cholesky matrix and the inverse of the Cholesky matrix derived from the van der Waerden scores. S^1 it turns out will have the desired correlation matrix, i.e., its correlation matrix will be C, the targeted correlation matrix. The final step is to assume that the matrix S^1, of transformed van der Waerden scores, has the same rank correlation as the sample of random deviates. So the algorithm proceeds to calculate the rank order of each column of S^1 and uses that rank order to re-order the cells in each column of X.

The following numerical example will show how the algorithm works in practice. The example is for a set of five random variables with the attributes in Table A.1. This is an example for the energy balance model (EBM) of the META 21 model, which drives the change in temperature component. In addition, the components are assumed to be linked to a correlation matrix as below.

$$C = \begin{bmatrix} & xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ xi\_1 & 1.00000 & & & & \\ C\_0 & -0.04451 & 1.00000 & & & \\ xi\_3 & -0.43716 & -0.11978 & 1.00000 & & \\ f2xco2 & 0.01392 & -0.03966 & -0.46228 & 1.00000 & \\ t2xco2 & -0.19343 & -0.08016 & 0.06549 & 0.65122 & 1.00000 \end{bmatrix}$$

The targeted Cholesky matrix is given by:

Gaussian quadrature and allowing for imposing a covariance matrix on the uncertain parameters.

Table A.1: **Example of a sample definition for an EBM model**

Name	Distribution	Parm 1	Parm 2	Parm 3
xi_1	Pareto	5.907	0.11628	
C_0	ParetoAlt	1.7062	53.0	
xi_3	Triangular	0.5	0.5	1.23723
f2xco2	Normal	3.45938	0.43674	
t2xco2	Normal	3.25312	0.80031	

$$H = \begin{bmatrix} & xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ xi\_1 & 1.00000 & & & & \\ C\_0 & -0.04451 & 0.99901 & & & \\ xi\_3 & -0.43716 & -0.13937 & 0.88852 & & \\ f2xco2 & 0.01392 & -0.03907 & -0.51956 & 0.85343 & \\ t2xco2 & -0.19343 & -0.08886 & -0.03540 & 0.74060 & 0.63635 \end{bmatrix}$$

The initial sample of random deviates is given below.[6] The rank correlation is exactly 1 for all pairs of columns as all of the deviates are in ascending order in each column, and the Spearman correlation matrix is also nearly the identity matrix.

$$X = \begin{bmatrix} xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ 0.11646 & 56.14932 & 0.51226 & 2.63749 & 2.11869 \\ 0.11920 & 59.22047 & 0.55074 & 3.05827 & 2.28394 \\ 0.12332 & 65.15213 & 0.58688 & 3.13678 & 2.81773 \\ 0.12652 & 65.61294 & 0.62132 & 3.32291 & 2.85096 \\ 0.12781 & 77.97449 & 0.67191 & 3.45491 & 3.10905 \\ 0.13417 & 82.40018 & 0.76344 & 3.52562 & 3.39424 \\ 0.14078 & 95.69734 & 0.79359 & 3.60265 & 3.55156 \\ 0.14771 & 108.00303 & 0.87995 & 3.78290 & 3.91509 \\ 0.15354 & 153.88674 & 0.94936 & 3.89073 & 3.92973 \\ 0.17777 & 272.65331 & 1.07535 & 4.59178 & 4.34054 \end{bmatrix}$$

The correlation matrix for S is given by:

$$C^s = \begin{bmatrix} & xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ xi\_1 & 1.00000 & & & & \\ C\_0 & -0.16433 & 1.00000 & & & \\ xi\_3 & 0.16526 & 0.17962 & 1.00000 & & \\ f2xco2 & -0.04612 & 0.21876 & -0.21009 & 1.00000 & \\ t2xco2 & -0.01611 & -0.56745 & -0.03087 & -0.03345 & 1.00000 \end{bmatrix}$$

and the corresponding Cholesky matrix is:

$$H^s = \begin{bmatrix} & xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ xi\_1 & 1.00000 & & & & \\ C\_0 & -0.16433 & 0.98641 & & & \\ xi\_3 & 0.16526 & 0.20963 & 0.96371 & & \\ f2xco2 & -0.04612 & 0.21409 & -0.25666 & 0.94136 & \\ t2xco2 & -0.01611 & -0.57795 & 0.09644 & 0.12141 & 0.80104 \end{bmatrix}$$

from which we derive the inverse matrix:

$$M = \begin{bmatrix} & xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ xi\_1 & 1.00000 & & & & \\ C\_0 & 0.16660 & 1.01378 & & & \\ xi\_3 & -0.20772 & -0.22052 & 1.03765 & & \\ f2xco2 & -0.04553 & -0.29068 & 0.28292 & 1.06229 & \\ t2xco2 & 0.17222 & 0.80205 & -0.16781 & -0.16101 & 1.24837 \end{bmatrix}$$

[6] This will depend on the random sample function and the initial seed.

We now have all the elements to transform the S matrix:

$$S^1 = S\,(HM)' = \begin{bmatrix} xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ -0.11419 & -1.36616 & 0.20236 & 0.01005 & 0.19013 \\ 0.11419 & 0.36713 & -0.08598 & 0.23054 & -0.41555 \\ -0.90846 & 0.50155 & -0.45556 & -0.45183 & 0.24884 \\ -1.33518 & -0.04715 & 1.38002 & -1.12544 & -0.66758 \\ 0.90846 & -0.50155 & -1.61302 & 0.46806 & -0.49921 \\ 0.60459 & -0.27950 & 0.04923 & 0.82639 & 1.55553 \\ -0.34876 & 1.30972 & -0.78274 & 1.13261 & 0.14081 \\ 1.33518 & 0.04715 & 0.01497 & -1.35900 & -1.52515 \\ 0.34876 & 0.96259 & 0.69972 & -0.42452 & 0.40673 \\ -0.60459 & -0.99378 & 0.59100 & 0.69313 & 0.56545 \end{bmatrix}$$

The correlation matrix of S^1 matches the targeted correlation matrix for the sample. The rank ordering of each of the columns of S^1 is given by:

$$R^1 = \begin{bmatrix} xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ 5 & 1 & 7 & 5 & 6 \\ 6 & 7 & 4 & 6 & 4 \\ 2 & 8 & 3 & 3 & 7 \\ 1 & 5 & 10 & 2 & 2 \\ 9 & 3 & 1 & 7 & 3 \\ 8 & 4 & 6 & 9 & 10 \\ 4 & 10 & 2 & 10 & 5 \\ 10 & 6 & 5 & 1 & 1 \\ 7 & 9 & 9 & 4 & 8 \\ 3 & 2 & 8 & 8 & 9 \end{bmatrix}$$

Based on R^1, we re-arrange the rows of X to yield:

$$X^1 = \begin{bmatrix} xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ 0.12781 & 56.14932 & 0.79359 & 3.45491 & 3.39424 \\ 0.13417 & 95.69734 & 0.62132 & 3.52562 & 2.85096 \\ 0.11920 & 108.00303 & 0.58688 & 3.13678 & 3.55156 \\ 0.11646 & 77.97449 & 1.07535 & 3.05827 & 2.28394 \\ 0.15354 & 65.15213 & 0.51226 & 3.60265 & 2.81773 \\ 0.14771 & 65.61294 & 0.76344 & 3.89073 & 4.34054 \\ 0.12652 & 272.65331 & 0.55074 & 4.59178 & 3.10905 \\ 0.17777 & 82.40018 & 0.67191 & 2.63749 & 2.11869 \\ 0.14078 & 153.88674 & 0.94936 & 3.32291 & 3.91509 \\ 0.12332 & 59.22047 & 0.87995 & 3.78290 & 3.92973 \end{bmatrix}$$

which is the final sample. Each column has the same sample distribution as the corresponding column in X, but the deviates have been arranged so that the correlation matrix is 'near' the targeted correlation matrix. In this case, the correlation matrix is:

$$C^1 = \begin{bmatrix} & xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ xi\_1 & 1.00000 & & & & \\ C\_0 & -0.17602 & 1.00000 & & & \\ xi\_3 & -0.31037 & -0.26483 & 1.00000 & & \\ f2xco2 & -0.32630 & 0.55056 & -0.29066 & 1.00000 & \\ t2xco2 & -0.26453 & 0.00238 & 0.12102 & 0.45300 & 1.00000 \end{bmatrix}$$

The 'fit' is not so good in this example, in part because the sample size is very small. Using the LHS tool with a sample size of 20,000, we get the following correlation matrix. Though not exactly the target, it is much closer than that with the sample size of 10. As explained in Swiler and Wyss (2004), there are limitations on the targeted correlations.

$$C^2 = \begin{bmatrix} & xi\_1 & C\_0 & xi\_3 & f2xco2 & t2xco2 \\ xi\_1 & 1.00000 & & & & \\ C\_0 & -0.04270 & 1.00000 & & & \\ xi\_3 & -0.42080 & -0.11510 & 1.00000 & & \\ f2xco2 & 0.01130 & -0.04040 & -0.44650 & 1.00000 & \\ t2xco2 & -0.18610 & -0.07590 & 0.05960 & 0.63470 & 1.00000 \end{bmatrix}$$

A.3 An introduction to the LHS utility

The LHS utility is a program that generates a LHS sample based on a user-prepared input file describing the contours of the LHS cube: the list of random variables with their respective distributions and distribution parameters, and optionally correlations across pairs of the random variables. The input file is a simple text file and its contents is described in van der Mensbrugghe (2023a). The output of the LHS utility is a data file containing the sample. The user has several output options including a CSV file, an XML file (readily loaded into Excel), a text-based GAMS data file and a GDX file.[7]

The LHS utility is a full recoding of the original utility, which has been developed in FORTRAN and can be downloaded from https://dakota.sandia.gov/downloads.[8] The re-coding has been done in C/C++ using mostly ANSI-C standards and thus should be readily compilable on other platforms, notably Linux and macOS.[9]

The two versions have been extensively tested and produce virtually the same sample.[10] The key advantage of the new version is the relative ease of coding in C/C++ compared with FORTRAN and the built-in memory management. While the FORTRAN code relies on a set of limits to the size of matrices, the C/C++ code uses dynamic memory management.[11]

The C/C++ version, in addition, has new features including additional distributions and new (and aforementioned) output options that facilitate embedding the utility in a workflow. The specific new features are:

- Ten new distributions: Cauchy, Dagum, Fréchet, Gompertz, Gumbel, Laplace, Lévy, Logistic, Kumaraswamy and Rayleigh[12]
- Four additional output options: CSV, XML, GMS and GDX[13]
- A new optional random number generator based on the Mersenne Twister algorithm described in Matsumoto and Nishimura (1998) and Nishimura (2000)[14]

Most of the features of the LHS utility are described in van der Mensbrugghe (2023a). Users who are interested in the full features of the LHS utility (apart from the new ones) are referred to the original documentation in Swiler and Wyss (2004).

Benchmarking suggests that the C/C++ code is significantly faster than the FORTRAN code. Both are compiled with the Mingw compilers, so we would not expect to see such large differences. The benchmark involves sampling for the META 21 simulation described in the next section. The META 21 sample is large and with a significant number of correlated distributions.[15] The sampling is run 100 times for both the C/C++ and FORTRAN code. It is also run with and without saving the sample. The output file size is over 150MB, so it is possible that C/C++ is more efficient in writing than FORTRAN. The benchmarking results are summarized in Table A.2. The FORTRAN

[7] The new software is fully backward compatible and the default output format is still supported.

[8] The download includes the full suite of software in Sandia National Lab's Dakota package. The LHS code is located in dakota-6.16.0-public-src-cli\packages\external\LHS. Interested users can contact the author for a copy of the FORTRAN code, which includes a few additional features and a *Makefile* to compile with GFortran.

[9] The new Mersenne Twister algorithm for generating random numbers requires a compiler that recognizes non-standard integer types, notably `unsigned long long`.

[10] The largest differences, and they are negligible, relate to deviates for the Beta distribution. The C/C++ code uses a different set of routines for the latter.

[11] The original FORTRAN code in fact crashed when attempting to generate the LHS sample for the IAM model. It had a relatively low limit on the number of potential user-based correlations.

[12] The latest version of the FORTRAN code includes the Fréchet and Gumbel distributions.

[13] We also provide GEMPACK code that automates conversion of the CSV output into a GEMPACK HAR file.

[14] The version of this algorithm in the LHS code is explicitly based on a 64-bit architecture, which may limit its portability.

[15] The full sample is a matrix of dimension 10000×582 and a correlation matrix of dimension 582×582. The benchmark was done under Windows 10 under normal operations (on a non-synced drive) using an Intel i9 12th generation CPU and 64 GB of RAM.

	Average seconds per sampling		Index relative to C/C++ w/ output	
	w/ output	w/o output	w/ output	w/o output
C/C++	8.4	6.0	1.0	0.7
FORTRAN	24.7	21.8	2.9	2.6

Table A.2: **Benchmarking of the LHS software**

version is around 2.9 times slower than the C/C++ code, and the speed difference is not linked to the writing of the output sample.

Appendix B

Workflow for running the Monte Carlo Simulations

B.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 META 21 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.

B.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 B.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 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]

[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*.

[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 Dellink et al. (2017).

[5] The procedure is described in section 2.4 of the supplemental materials for Dietz et al. (2021).

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 (2.45).
`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 (2.21).
`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 (2.59). N.B. The original META 21 package has two versions of the 1990 temperatures. One is in the main META 21 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 (2.59), 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 (2.60). 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 B.1: **External data files**

B.3 Model files

Table B.2 lists the code files used to run the core META 21 file and the Monte Carlo simulations. The following section will provide a concise description of the workflow.

B.4 Workflow

B.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.

B.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 (van der Mensbrugghe, 2023b). 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 size determined by the number of CPU cores. 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. 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.

[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
`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 META 21 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.
`META21Model.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 B.2: **Model files**