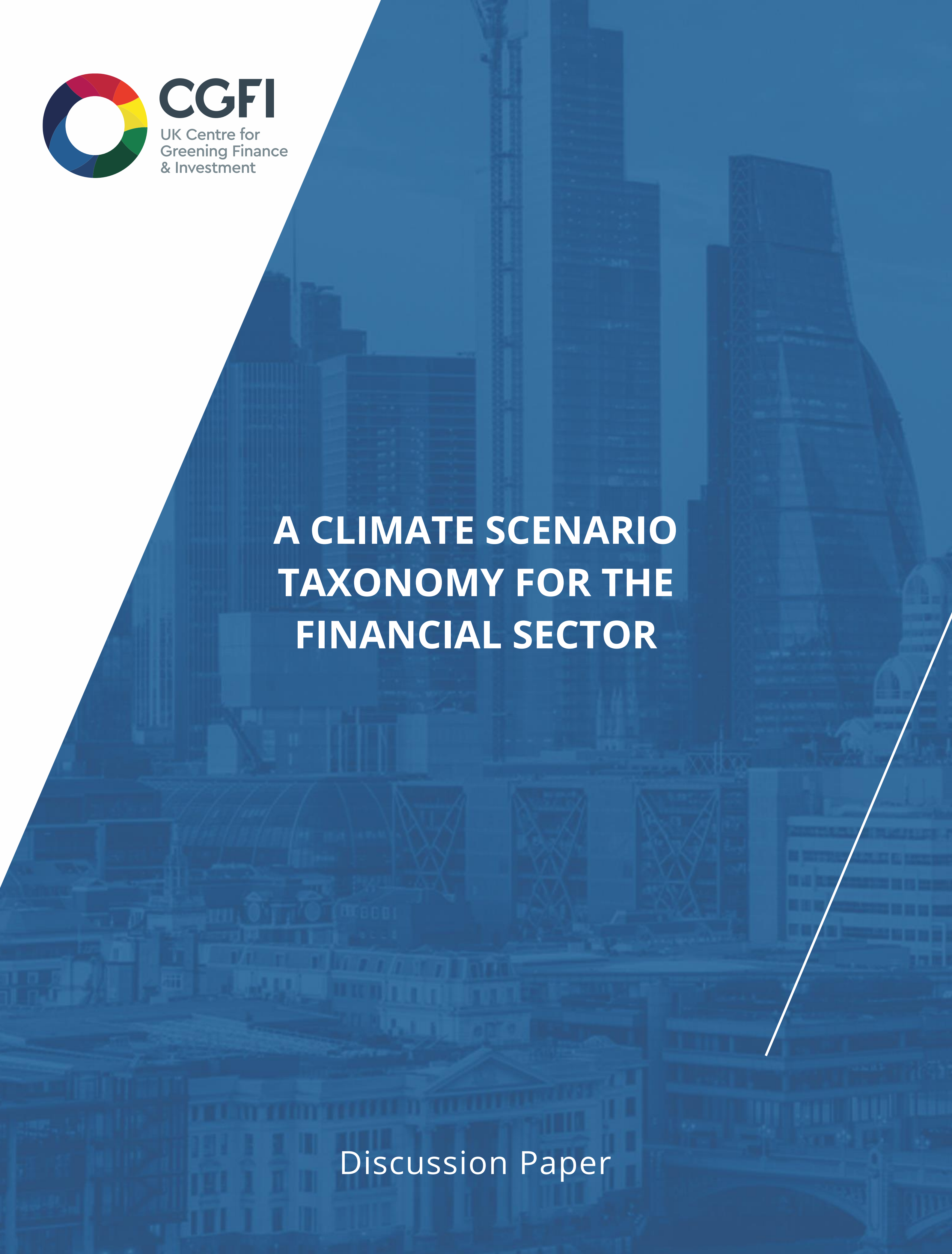
CGFI
UK Centre for Greening Finance & Investment



A CLIMATE SCENARIO TAXONOMY FOR THE FINANCIAL SECTOR

Discussion Paper

University of Bristol | Imperial | University of Leeds | University of Oxford | University of Reading | Catapult Satellite Applications | The Alan Turing Institute | UKRI Science and Technology Facilities Council

Authors

David Wilkinson[1], Moritz Baer[12], Chris Cormack[1], Alexandre C. Köberle[3,4], Jason Lowe[1,5,6], Erica Thompson[1,7]

[1] UK Centre for Greening Finance and Investment
[2] University of Oxford
[3] Centre for Climate Finance and Investment, Imperial College Business School London
[4] NOVA School of Business and Economics, Portugal
[5] University of Leeds
[6] Met Office Hadley Centre
[7] Department of Science, Technology, Engineering and Public Policy, University College London

Acknowledgements

We greatly appreciate the support of Dr Gireesh Shrimali, Head of Transition Finance Research, Oxford Sustainable Finance Group, University of Oxford. As a senior member of the UK Centre for Greening Finance and Investment, Gireesh provided not only sponsorship for this crucial piece of work but also guided our work via detailed reviews.

Abstract

In response to calls to develop a more comprehensive economic climate scenario taxonomy that provides transparent information on scenario characteristics and underlying model construction, we present a flexible and practical taxonomy. Our three dimensional six-level taxonomy, split into three tiers, separates the narrative, macro modelling and micro expansion elements of an economic climate scenario whilst identifying any geographical and sectoral variations.

By developing a common understanding of the assumptions and calibration choices in each step of the scenario construction process, scenario builders and financial end users will be able to position the scenario pathways produced in the range of possible outcomes. This will be possible not just for headline pathways such as global mean temperature, but for lower-level sub-sector and geography level pathways. This will allow financial actors to better match scenario characteristics with individual use cases, ranging from the central expectation scenarios used in the assessment of an individual financial institution's strategic plan, to the tail scenarios required by central banks and supervisors in the assessment of systemic financial stability risk. Having described the taxonomy, we emphasise that evaluation is always needed to understand whether any given modelling approach can meet the needs of a specific decision question and present a pragmatic framework for operational evaluation.

We believe that adoption of the taxonomy will support improved documentation by scenario builders, encourage academic peer review, lead to the development of a broader range of scenarios and ultimately drive the adoption of more decision useful scenario analysis across the financial sector. We encourage central banks and supervisors to develop regulations that require financial institutions to rigorously assess their economic climate scenarios driving investment into improved modelling. This will drive knowledge transfer between the academic and financial sectors, resulting in higher quality and more consistent assessments of climate related risks across the financial sector. International bodies, such as the NGFS, and national bodies, such as the CGFI, have the potential to play a key role in the creation of a broad library of well-documented climate scenarios for use across the financial sector.

Contents

1 Introduction

Climate scenarios are valuable tools for examining the effects of climate change and mitigation policies on micro and macroscale economic aspects of the global system by constructing a range of plausible pathways. However, they are inherently complex to construct, requiring the integration of expertise from different fields of research. This paper responds to recommendations from the Climate Financial Risk Forum (CFRF) that call upon financial institutions (FIs), central banks and supervisors (CB&S) and research institutions to work closely together to develop a more comprehensive scenario taxonomy that provides transparent information on scenario characteristics and underlying model construction, in order to provide an effective mechanism to enhance capabilities around climate scenario analysis in the financial sector.

Climate scenarios in their current form may often not be fit for purpose for specific financial use cases such as stress testing and risk management (Baer et al. 2023). Most scenario deficiencies point towards a systemic underestimation of the risk, potentially giving rise to a false sense of security on how the transition may unfold and the financial consequences of climate change. Examples of such deficiencies are the use of simplistic damage functions that estimate the macroeconomic impact of climate change based on historic relationships between temperature and productivity, and the omission of labour frictions in the modelling of the transition to a low carbon economy.

There will always be a degree of subjectivity in judging the likelihood of different climate pathways. However, for climate scenario analysis to be a decision-relevant tool that drives risk management and capital allocation processes, financial institutions need sufficient understanding to make informed judgements as to where in the probability distribution respective scenarios sit, given the level of conservativeness in key assumptions and the likelihood of assumed policy action. Similarly, CB&S need to understand which scenarios are adequate to support their own risk appetite definition through capital setting and micro prudential frameworks to balance potential short-term disruptions and longer-term financial stability objectives.

The NGFS scenarios, often employed by financial users, are typically described based on the type of policy action assumed (orderly, disorderly, divergent, etc.). A recent USS & University of Exeter paper, *No Time to Lose*, extends the classification of narrative scenarios to two dimensions by considering policy and market dynamism (Cliffe et al. 2023). The NGFS further developed the classification of their scenarios using five risk drivers (NGFS 2021) as illustrated in Figure 1:

Category	Scenario	Physical risk Policy ambition	Policy reaction	Transition risk Technology change	Carbon sequestration	Regional carbon price variation
Orderly	Net Zero 2050	1.5 °C	Immediate	Fast	Medium	Low
	Well below 2°C	1.8°C	Immediate	Moderate	Medium	Low
Disorderly	Divergent net zero policies	1.5°C	Immediate	Fast	Low/Medium	Medium
	Delayed 2°C	1.7°C	Delayed	Slow/Fast	Low	High
Hot house world	Nationally Determined Contributions NDCs	2.5°C	NDCs	Slow	Low	Limited
	Current policies	3°C+	Current policies	Slow	Low	Limited

Macro-financial risk level

Low	Medium	High

Figure 1 Representation of the risk drivers in the NGFS scenarios (NGFS 2021)

We develop a more comprehensive classification system based around the structure of the taxonomy. A detailed understanding of scenario narratives, underpinning assumptions and model components is critical if scenario analysis is to be a decision useful tool for financial users. Typically, financial practitioners are heavily dependent on existing third-party climate scenarios and almost entirely dependent on third party physical climate models and Integrated Assessment Models (IAMs). Expertise in these models, and the scenarios they produce, sits predominantly in the academic community. This knowledge gap needs to be bridged. The development and application of a common scenario taxonomy will provide an effective mechanism to translate academic expertise into practical information to support the development of climate scenario analysis in the financial sector.

We envisage three different applications of the scenario taxonomy:

1) In the creation of end user-friendly documentation by scenario designers.
2) To encourage a standardised output for academic peer review of scenarios.
3) To support a more structured approach to the selection and interpretation of scenarios by financial end users.

By engaging with academic experts, and through multi-disciplinary engagement across the participating institutions, a comprehensive account of the likelihood associated with each individual scenario component, key assumptions, characteristics, and features of most widely used scenarios could be provided. This will allow end users to assess the credibility of a scenario, a vital step if scenarios are to be actively employed in decision making. We consider credibility to be a function of methodological transparency, scenario coherence, and likelihood. Although this assessment will include an element of end-user judgement, a taxonomy will contribute towards mitigating the implications of weak scenario design, communication of scenario information and scenario misuse.

The remainder of this paper is structured as follows:

- The Climate Scenario Taxonomy
  - Taxonomy Overview
  - Taxonomy - Detailed Structure
    - Level 1a – Narrative

    - Lebel 1b – Narrative Parameterisation
    - Level 2a – Modelling Architecture & Feedback Mechanisms
    - Level 2b – Macro Models
    - Level 3 – Microeconomic Expansion Models

- Linking Models with the Real World and the Application of the Taxonomy
- Conclusions

In the level 2 and 3 modelling discussions we present examples of the taxonomy in operation.

We are limited in this analysis by the very obstacles that this this paper seeks to address. Namely, the lack of suitable scenario documentation and peer review analysis to help inform financial end users as to scenario suitability. Therefore, the depth of the analysis in the examples provided is a function of the quality of documentation available and the authors' expertise across the different elements. Although deep review of model code and calibration is possible for open-source models, this level of analysis is beyond the scope of this paper.

Further, a lack of suitable documentation across a range of scenarios makes it challenging to position the output of a particular scenario or scenario component in the range of potential results. It is our hope that the adoption of a scenario taxonomy will lead to the development of a library of scenario documentation that feeds back into improving the documentation and understanding of individual scenarios. We recommend that the role of collating such a library, promoting peer review and identifying gaps in the range of available scenarios is taken up by an international body such as the NGFS with national bodies, such as the CGFI, performing a similar role but with more focus on detailed analysis of local risks.

2 The Climate Scenario Taxonomy

2.1 Taxonomy Overview

We introduce a three-dimensional taxonomy. The primary dimension will capture the steps in the scenario construction modelling chain, descending from the high-level scenario narrative down through the modelling levels (covering climate models, macroeconomic models and sector models) and ending with the construction of financial asset projection curves. The secondary and tertiary dimensions capture variations in the narrative and modelling approach by geographic region and sector. Resolution and calibration in these three dimensions, as well as time, will often play a key role in the applicability of a scenario to certain use cases.

Application of the taxonomy requires the user to drill down to the scenario component level and assess the credibility of the output at this level as illustrated in figure 2. The level of detail included in the assessment of a scenario can be adjusted according to the use case. A scenario builder, documenting their scenario, should include sufficient detail for all potential financial end users to gain the understanding they require. However, a financial end user matching a scenario to a particular use case will only need to consider relevant details. For example, a financial firm that is not involved in the AFOLU sector may need to know how conservative a particular scenario is in its treatment of the emissions mitigation potential of the sector and its adaptation to future climate change, but not the detailed modelling of the various agricultural subsectors (see the later discussion on sectoral modelling).

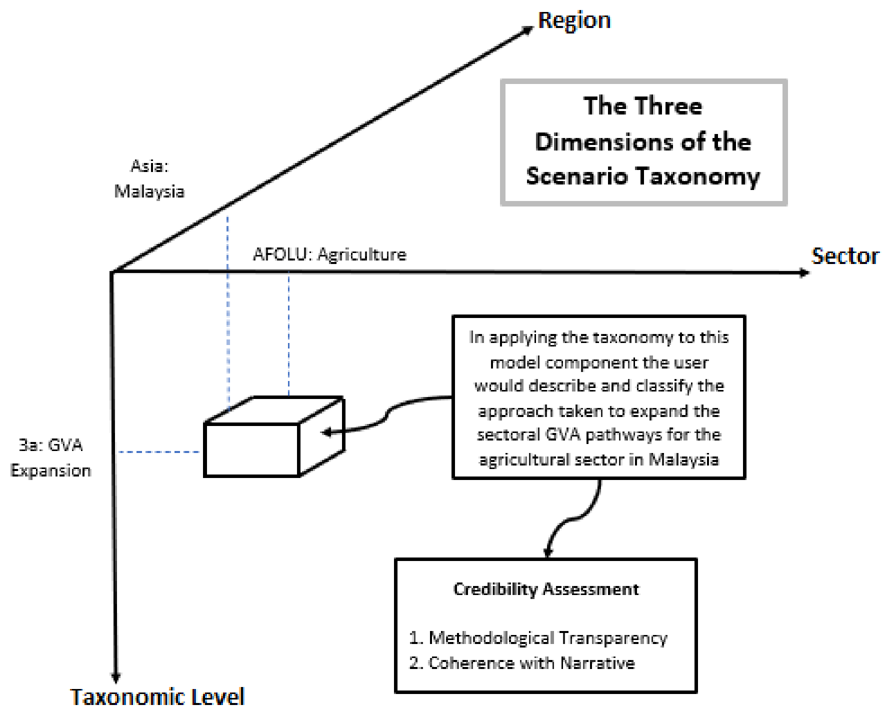


Figure 2: The Three Dimensions of the Scenario Taxonomy

In addition to the overarching structure of the taxonomy, we propose a classification system that provides a high-level description of each component and additional structure for the detailed analysis of each component. Combining the classification tables for each component in a climate scenario provides a convenient one-page summary.

In the full form of taxonomy there are six levels in the primary dimension that drill down into the detailed modelling choices and their calibration. These are listed below:

1 Narrative
  1.1 Description
  1.2 Narrative Parameterisation
2 Macro Modelling
  2.1 Modelling Architecture and Feedback Mechanisms
  2.2 Models
3 Micro (Expansion) Modelling
  3.1 Scenario Designer
  3.2 End User

These six levels correspond to the modelling chain used in regulatory scenario analysis exercises such as the Bank of England climate biannual exploratory scenario (CBES) (Bank of England 2021), illustrated in figure 3 below. The CBES scenarios included some detailed expansion by the scenario designer, the Bank of England (e.g. sectoral GVA and corporate credit spreads), but participants in the exercise were left to perform significant end-user expansion to derive asset level pathways. This structure should provide an effective framework for most climate scenarios employed in the financial services sector.

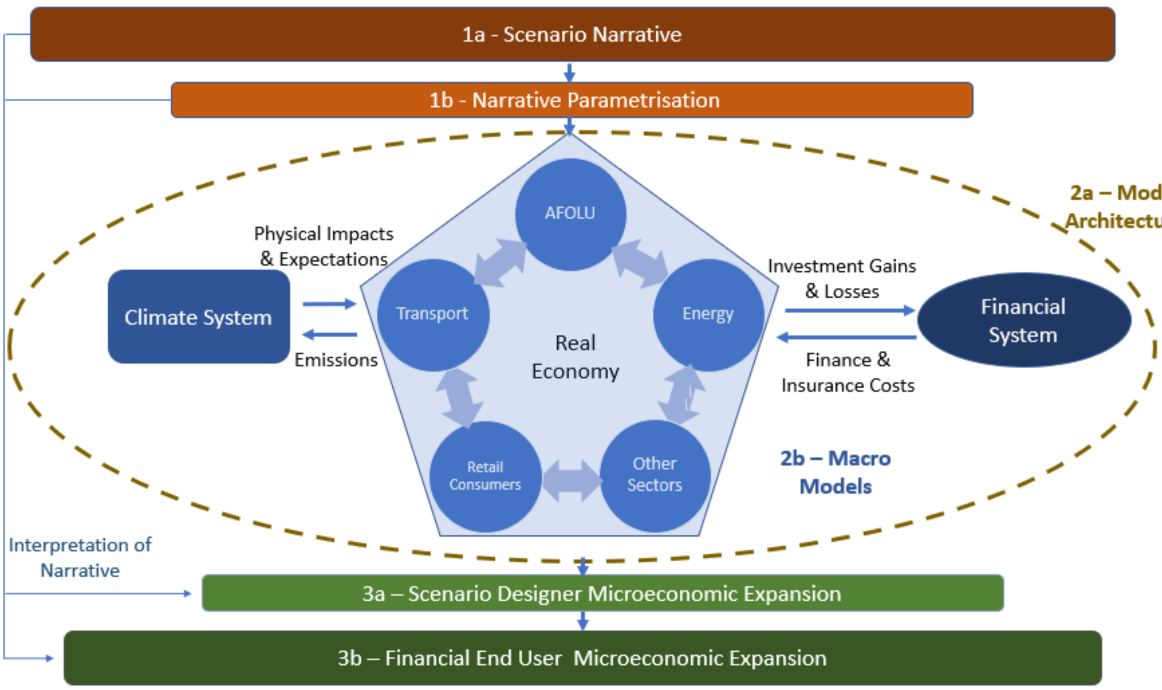


Figure 3: Climate Scenario Modelling Chain

When applying the taxonomy, the user will describe the approach taken in each scenario component and provide a judgement as to transparency of the methodology (not applicable to scenario builders), the level of coherence with the overall narrative and other components, and the relative likelihood of the assumptions given the inputs from higher in the modelling chain. For example, a scenario designed to be a central projection for strategy setting might look to evaluate the expected or most likely pathways for a given high probability narrative across all components. Scenarios where different components are associated with very different probabilities will lack coherence, making them more difficult to associate with specific financial use cases. Scenarios must be well understood and credible to be decision useful.

For financial end users the key output from the taxonomy will be the assessment of the individual asset pathways in terms of positioning in the range of potential outcomes and overall credibility[1]. This analysis will support the user's judgement of scenario likelihood. Ideally, the scenario builder's documentation will provide the majority of the details required. The financial end user may need to fill gaps by developing their understanding of the model employed and will need to cover their own expansion modelling.

[1] Credibility means the belief of a reasonable expert that the output of the assessment follows from the input in a manner that is representative of the real world.

2.2 Taxonomy - Detailed Structure

We propose a standard structure for the primary dimension as follows:

2.2.1 Level 1a – Scenario Narrative

The scenario narrative should capture all the key exogenously defined elements of the scenario that inform the calibration of the underlying modelling of the physical, economic, and financial implications. Many published scenarios will lack the fine detail required to expand down to the asset pathway level and will therefore represent an envelope of scenarios rather than an individual pathway. However, the qualitative narrative might include nuances that go beyond the quantitative parameterisation in level 1b that can inform the expansion modelling choices in level 3.

We identify eight key elements to the narrative that define the scenario as illustrated in figure 4 below (developed from the NGFS classification):

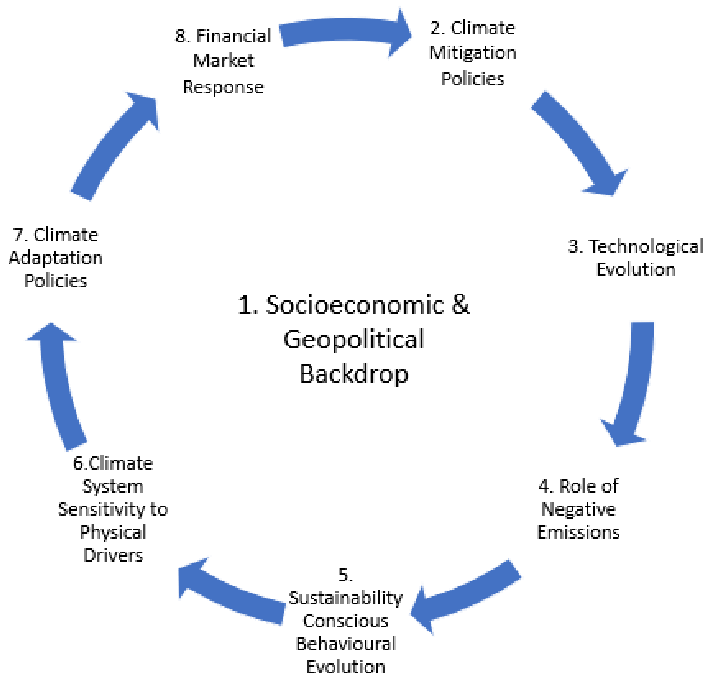


Figure 4: Narrative elements - (the arrows are illustrative: many of the elements interact)

Each key element should be described and assessed (self-assessed in the case of a scenario builder) commenting on the depth, coherence and likelihood of the scenario described. Where a scenario element is not addressed by the scenario builder, this should be recognised in the narrative description so that end users are aware that they may need to make additional assumptions when calibrating their expansion models. To aid the alignment of scenarios with use cases, financial end users should develop the third-party narrative to capture the key assumptions inherent in their own expansion modelling. For example, the role of barriers to new entrants in a sector could be extremely important when modelling the pathway of an individual equity price, even it is not considered when modelling GDP or sectoral GVA pathways. These elements are further described below.

1. **Socioeconomic and Geopolitical Backdrop** – The construction of a climate scenario needs to start with a non-climate backdrop, although there are feedback mechanisms in longer dated scenarios that potentially link all aspects. Most used climate scenarios are based on the current economic landscape combined with general economic projections following the shared socioeconomic pathways (SSPs) developed by the climate scenario modelling community (Riahi et al 2017).

    - SSP1 Sustainability – Taking the Green Road
    - SSP2 Middle of the Road
    - SSP3 Regional Rivalry – A Rocky Road
    - SSP4 Inequality – A Road Divided
    - SSP5 Fossil Fuelled Development – Taking the Highway

    Two socioeconomic pathways that are typically considered exogenous are population growth and the counterfactual GDP pathways. Although population growth might be significantly influenced by climate change, it is generally set based on the "middle of the road" forecasts in SSP2 (Riahi et al 2017, Fricko et al 2017). The counterfactual GDP pathway sets the anticipated growth, in the absence of climate related impacts, and is also often based on the middle of road SSP2. Other socioeconomic pathways captured by the SSPs such as urbanization, income distribution and education may also be relevant for more sophisticated scenarios. However, more stressful scenarios can be constructed by combining climate events with other non-climate stresses, potentially making use of the other SSPs listed above. A detailed framework for assessing non-climate stresses is beyond the scope of this paper but we recognise the need to capture non-climate elements in scenarios.

2. **Climate Mitigation Policies** –The narrative should describe the mitigation policy assumption(s) including any regional and sectoral variations. At the simplest end of the spectrum this might take the form of a globally consistent carbon tax, but more sophisticated approaches will identify combinations of policy instruments with regional variations. Although there may be a target level of warming associated with a suite of policy measures (described as policy ambition by the NGFS), this will be model dependent and will not necessarily coincide with the assumed climate pathway in the scenario. More stressful scenarios should include the failure of a suite of policy options to limit global warming to the intended amount as set in element 6 below.

3. **Technological Evolution** – Advances in technology will play a significant role in the evolution of high emission sectors and their low emission substitutes. The narrative should describe the basis for any assumptions made. This will typically take the form of learning curves with capacity limits and adoption lags. The high-level description may follow an SSP but the detailed interpretation that informs the narrative parameterisation is crucial and should be carefully analysed.

4. **Role of Negative Emissions** – The level of negative emissions plays a key role in determining the relationship between the strength of mitigation policy action and the climate system response. In theory, technological and land use change negative emissions can be calculated endogenously based on mitigation policies, technology factors and behavioural changes. However, given the significance of this component, we recommend that the narrative description draws out which negative emission elements of the scenario are endogenously generated and those that are set exogenously with rationale.

5. **Sustainability Conscious Behavioural Evolution** (Corporate, Domestic and Financial Institution) – Behavioural assumptions can be used to allow consumers/customers to behave in a way that is not purely economically rational by favouring green products or services. The narrative should describe any assumptions made and the rationale on which they are based.

6. **Climate System Sensitivity to Physical Drivers** – The climatic response to a given emissions pathway might be endogenously determined and hence governed solely by model choice. However, it is possible to set the level of conservatism by specifying a confidence interval against a range of model outcomes, typically the IPCC model ensemble. For example, the policy ambition might be set to limit global warming to 1.5^{o}C with 50% confidence but under a given scenario the 90th percentile outcome crystallises. The choice of the 90th percentile is a key exogenous input that sets the realised sensitivity of the climate system and should form part of the scenario narrative. Where a single model has been selected the rationale for the selection of the model should be identified and the positioning of outputs in relation to other models explained. Similarly, the downscaling of chronic climate pathways to acute weather events may involve exogenously setting a level of conservatism. The narrative should describe the extent to which the climate response is endogenously determined by the model architecture and identify any judgmental elements.

7. **Climate Adaptation Policies** – Adaptation policies might include sea defences and government loss cover. This level of detail is often not covered in third party scenario narratives leaving financial end users to make their own assumptions. It can be extremely challenging to associate adaptation responses with top-down macroeconomic modelling approaches.

8. **Financial Market Response –** The financial markets sit at the heart of the economic system and their response can play a significant role in the evolution of a scenario, particularly in the short term. For more sophisticated models this behaviour will need to be calibrated. Market expectation and risk appetite will play a significant role in asset pricing and funding costs. The narrative should describe the approach taken to predicting this behaviour. For simpler models this is likely to be absent with implications for the macroeconomic pathways generated and the calibration of expansion models.

2.2.2 Level 1b – Narrative Parameterisation

Parameterisation of the scenario narrative translates the qualitative narrative into quantitative inputs to the macro models. The level 1b assessment should follow the same structure as level

1a summarising the key parameters used to represent each of the eight narrative elements. The calibration of these parameters should be assessed as to how well they reflect the narrative and whether any bias is introduced by approximations (e.g. the use of a simple carbon price to represent a wider range of mitigation policy tools). It is important to recognise that even two scenarios with a common scenario narrative, and constructed using the same model architecture, might differ in terms of the parameterisation of the narrative.

The full list of exogenous parameters used in the modelling chain, and their calibration, will be captured when identifying the inputs for model component in levels 2 and 3 of the taxonomy. The number and sophistication of the inputs will reflect the sophistication of the modelling, and the range of output pathways covered by the scenario. This might deliver a daunting number of exogenous inputs for sophisticated models with global financial coverage. However, at the other end of the spectrum scenarios based on expert judgement, rather than modelling, may collapse to only two levels with expert judgement translating the narrative into all the required pathways.

Common parameterisation and calibration options for the eight narrative elements are shown in table 1 below, with potential assessment approaches:

Element	Parameterisation	Calibration	Assessment
1. Mitigation Policies	Carbon price	Implied from a chosen IAM based on the policy ambition global temperature rise and timing of policy action in the narrative	Comment on the emission sensitivity of the chosen IAM and the implications of a globally consistent framework
2. Chronic and Acute Climatic Response	Chronic pathway model ensemble confidence interval	Choice of confidence interval	Choice of confidence interval determines positioning in the range of outcomes
3. Climate Adaptation Policies	Sea defences by region with sea-level rise protection level	Current policies	Compare with potential additional response allowing for availability of capital
4. Technological Evolution	Technology cost learning curves, capacity limits, adoption lags	Use of historic data to calibrate learning curves with engineering-based lags and limits	Comparison with academic literature to establish positioning against alternative views
5. Negative Emission Assumptions	Endogenously determined	Driven by calibration of elements 1, 4, 6, 7, and 8	Reflect on calibration of feeder elements and the resultant level of negative emissions
6. Financial Market Response	Increase in equity risk premia and credit spreads by rating	Based on 2008 GFC observations with judgmental sector-specific adjustment for transition	Compare with alternative treatments based on academic literature review
7. Sustainability Driven Behavioural Evolution	Sector-specific carbon price add-on	Judgemental	Compare with alternative treatments based on academic literature review
8. Other SSPs	Population growth, counterfactual GDP, urbanisation %, income distribution	Middle of the road – SSP2	Consider positioning versus other SSPs

Table 1: Common macro model parameterization options

2.2.3 Level 2a – Modelling Architecture and Feedback Mechanisms

Model architectures often consist of a set of model subcomponents that project the physical or socio-economic variables. Consequently, model architectures can be regarded as a combination of a set of modelling approaches, rather than a complete climate model framework. For example, Integrated Assessment Models such as REMIND-MAgPiE and MESSAGE-GLOBIOM combine existing energy system, land use and climate models. These might be further supplemented with more detailed models for specific elements of a scenario. For example, a state-of-the-art physical climate model or macroeconomic model. Even the most sophisticated and granular economic models will need to interact with a physical climate model to evaluate the economic impact of climate change. The architecture presented in figure 5 below is that adopted by the Bank of England for the CBES exercise and is typical of those used across the financial services sector.

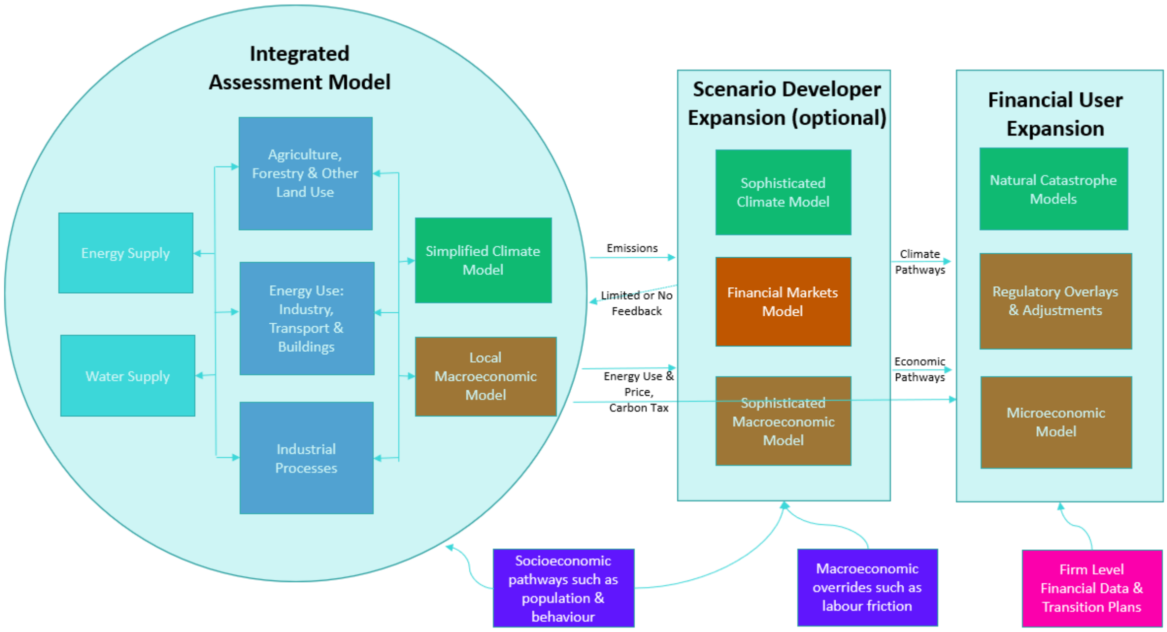


Figure 5: Illustration of a climate economic modelling architecture combining an integrated assessment model with supplementary climate, financial markets and macroeconomic models (from Baer et al 2023)

The choice of modelling architecture plays a significant role in the sophistication with which different elements of the climate-economic system are represented, the degree of pathway expansion left to financial end users and the way in which feedback mechanisms between the elements are captured. In level 2a of the taxonomy we classify the architectural choice, explain the representation and impact of feedback mechanisms, and present an assessment of the limitations and applicability of the chosen approach.

2.2.3.1 Classification System

The classification system presented in table 2 below characterises the modelling architecture based on the time horizons explored, the way that discrete models are linked together, and the approach to capturing inter-component feedback mechanisms. Further, by utilising the component level classifications presented in level 2b to convey the sophistication of the underlying models, the architecture classification provides the experienced reader with a rapid feel for the applicability of the chosen approach.

Scenario		**CBES Scenarios**	**NGFS MESSAGEix-GLOBIOM Scenarios**
Time horizons	Pathways	30 years	80 years
	Market expectations horizon	30 years	N/A
Integrated macro components	Climate system	Reduced Complexity (MAGICC)	Reduced Complexity (MAGICC)
	Macroeconomics	Energy Economy General Equilibrium Model (REMIND)	Computable Generalised Equilibrium Model (MACRO)
	Energy	Energy Economy General Equilibrium Model (REMIND)	Energy Economy General Equilibrium Model (MESSAGEix)
	AFDLU	Partial Equilibrium Dynamic Recursive Model (MAgPIE)	Partial Equilibrium Dynamic Recursive Model (GLOBIOM)
	Others	N/A	N/A
Feedback mechanisms		Unclear	Unclear
Stand-alone macro components	Climate system	Full Complexity Ensemble (90th Percentile Worst Case)	N/A
	Macroeconomics	NiGEM	N/A
	Others	N/A	N/A
Scenario designer expansion components	Sectoral GVA	Emission Proxies and Expert Judgement	N/A
	Sectoral credit spreads	Historic Based Expert Judgement	N/A
	Others	N/A	N/A

Table 2: Model Architecture Classification System (Model architecture is not scenario specific so applies to all scenarios based on this architecture)

Although the description of more sophisticated modelling approaches will be component specific, common classifications can be applied to less numerically intensive methods. For example:

i. Expert judgement
ii. Historic pathways with or without expert overlay
iii. Historic distributions with or without expert overlay

In addition to the high-level classification of the architecture, the 2a assessment should include a more detailed discussion of the representation of feedback mechanisms and provide a summary of the strengths and weaknesses of the chosen approach. However, the more detailed consideration of the underpinning elements is covered in levels 2b and 3 of the taxonomy. At this point, the publicly available documentation may provide insufficient detail to fully understand the representation of feedback mechanisms presenting a credibility issue for financial end users.

2.2.3.2 Example Model Architectures

Inevitably the modelling of the economy and the Earth climate system are usually somewhat separate. Ideally there will be a degree of integration to incorporate feedback mechanisms. However, the computational intensity of advanced climate models means that this will normally be achieved by coupling a reduced complexity climate model with the other components. This is the case for all the commonly used Integrated Assessment models. For example, REMIND-MAgPIE employs the reduced complexity MAGICC climate model. A common architectural choice is the incorporation of a more sophisticated climate model to build out the physical pathways using the emissions pathway determined by the economic modelling. The Bank of England adopted this approach for the CBES.

There is a wider range of architectural options for representing the various elements of the economy. The key choices concern the way that expert judgement, top-down macroeconomic modelling, and bottom-up process-based modelling of economic sectors are combined. As with physical modelling, there is the question of how well-integrated these components are with the potential for additional stand-alone models that expand either the macroeconomic pathways or specific sectors of the economy. Examples of common architectural approaches are as follows:

- **Simplest approaches:** At the simplest end of the spectrum, pathways can be deduced from the narrative using a combination of historic data and expert judgement. Where expert judgement is employed scenario users should endeavour to establish the basis for the judgement as this might effectively be a simple model.

- **Low-granularity approaches:** At the low granularity end of the modelling spectrum, economic pathways are driven by top-down macroeconomic models without any sectoral representation of the economy (e.g. NiGEM), though they may be calibrated with climate variables from feeder models that do include a sectoral representation creating a hybrid approach[2]. For example, the CBES used NiGEM macroeconomic pathways calibrated using climate pathways produced by the REMIND-MAgPIE integrated assessment model. Further sectoral detail was available from REMIND-MAgPIE but not all this information was included in the published pathways.

- **Mid-granularity approaches:** Mid granularity models capture sector specific behaviour, for at least the key climate relevant sectors (such as energy, power and land use). Within this category there is significant variation with respect to the detail with which each sector is modelled. The most sophisticated treatments will include detailed treatment of supply chains, the role of technological innovation and the interaction with the climate system. There may be trade-offs between the level of sophistication offered for different components of the system. The commonly used integrated assessment models are mid-granularity models that capture a degree of detail for the energy and AFOLU sectors but with less sophisticated handling of macroeconomics and the climate system. These weaknesses can, to an extent, be addressed by blending the results with those from more sophisticated macroeconomic and climate models but at the risk of introducing incoherencies.

- **Sophisticated approaches:** At the most sophisticated end of the spectrum agent-based models provide the greatest granularity, with the potential to represent individual firms within a sector. Within this family of models there is still significant variation in the granularity adopted, the sophistication of the agent level decision making rules and the granularity of the input data on which these decisions are based. Agent based models still require the input of chronic and acute climate pathways potentially with an intermediate layer to translate climate/weather data into economic impact.

Within the academic literature the firm agent-based model from Cormack et al. (Cormack et al. 2020) has an interface to a wider scale technology and macro-economic scenario set from integrated assessment models such as those used by the NGFS (the model used in the paper is based on the partial equilibrium model framework GEM-E3 POLES). The IAM in this model architecture provides a temperature pathway because of the energy production and utilization models. These demand and supply pathways are used to drive expected demand predictions for the firms that are modelled. The firms that are modelled are subject to a set of agent rules that define the management of their capital structure and profitability targets based on known data of the firms. Such data involves knowledge of specific costs covering operational, variable and funding as well as information on its credit quality metrics and investor engagement communications covering, for example, dividend payments. Furthermore, the firms engage in micro-competition based on price and capacity to supply. The model framework is integrated to macro-economic factors such as government yields, inflation alongside the detailed specific economic supply and demand drivers. The output of such a model consists of financial data on the firm's performance (e.g. balance sheet, income statement and cash flow statements), information on its non-financial performance such as emissions, physical production and

[2] This hybrid approach, combining a low granularity and mid granularity model, introduces the potential for incoherence between the macroeconomic pathways and the asset pathways developed by the financial end user.

capacity, as well as specific data on firm level physical damage. In terms of financial performance, the model produces a value for the firm’s equity and returns, its probability of default, losses given default, credit rating and funding costs.

Combinations of the four approaches above are possible. For example, the 2023 ACPR insurance stress test (Autorité de Contrôle Prudentiel et de Résolution , 2023) combined an assumed repeat of the 2022 European heat wave in 2023 and 2024, using historic economic data, with a modelled transition policy response based on the NGFS scenarios and the judgmental selection of a dam failure.

2.2.3.3 Feedback Mechanisms

There are many feedback mechanisms to be considered when modelling the combined climate-economic-financial system. For example, the reduction in the planetary albedo, due to Arctic Sea ice retreat, and the release of methane due to the thawing of the Siberian permafrost are positive feedbacks that reinforce global warming. These mechanisms should be captured within the climate model. However, the modelling architecture for climate-economic-financial scenarios will almost inevitably require the integration of discrete models.

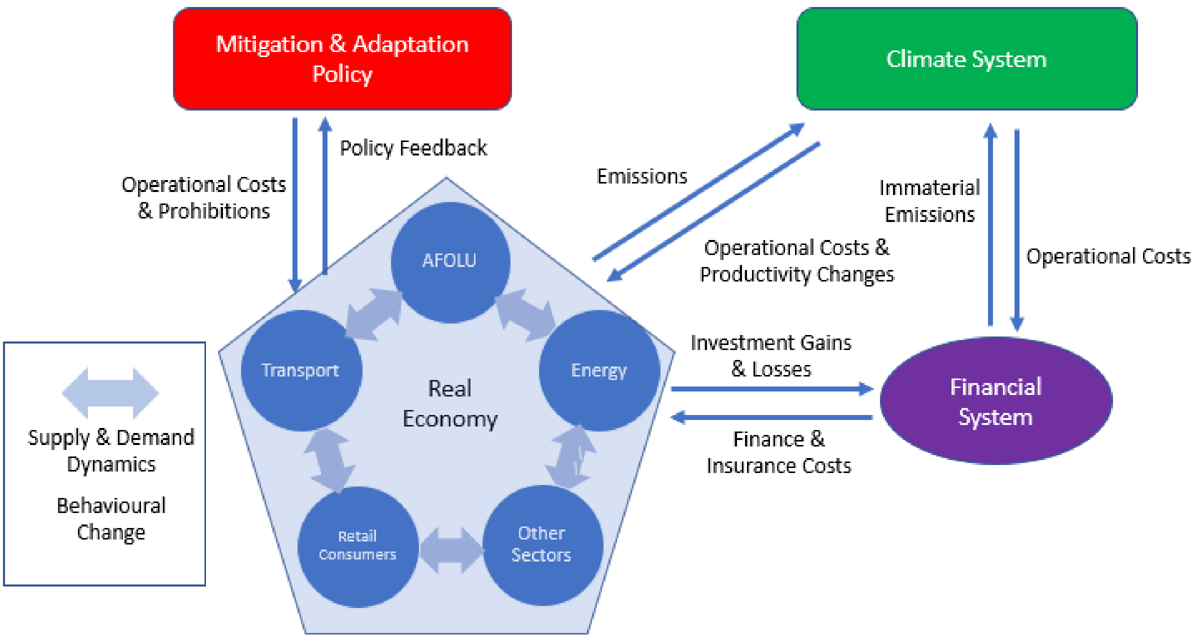


Figure 6 – Schematic Representation of Feedback Mechanisms for a Typical Mosel Architecture

Therefore, special consideration is required for feedback mechanisms that operate between elements of the system represented by separate models. Interactions between the Earth climate system and the AFOLU and energy sectors naturally fit into this category, as these three elements are often modelled separately. The impact of GHG emissions from the energy and AFOLU sectors on the climate system sits naturally at the heart of modelling climate change and should be addressed by any modelling architecture.

However, secondary feedback mechanisms, which may not be captured, could still be material. For example, the impact of climate change on the cost of energy production or agricultural yields will feed through to pricing, demand and subsequently back to future emission levels. The integration of the various models must capture such effects, where material to the intended use case. Secondary energy/AFOLU feedback mechanisms are often captured as the sectors involved are so central, and thus warrant their own models. For sectors that are less clearly recognised in the modelling architecture, it is likely that feedback mechanisms are not well captured and weaknesses are propagated through the modelling chain. For example, the transportation sectors also sit at the nexus of economic activity, energy use and emissions but the associated feedbacks may not be as well represented, and the modelling approaches adopted may be poorly understood by end users.

A subtler feedback mechanism is the result of the failure of policy intervention to deliver the anticipated response. Even disorderly transition scenarios tend to represent transition policy by a smooth increasing carbon price. When considering the full range of potential scenarios, we might wish to include the reaction of public policy makers to the unfolding scenario creating a more disjointed pathway with stress-inducing jumps. This could be addressed in an agent-based architecture with individual nation policy makers responding to the climate, economic and geopolitical landscape.

One feedback mechanism particularly relevant to the financial services sector is the effect of losses in the financial sector, commonly referred to as financial frictions. Losses, particularly in the banking sector, can lead to a reduction in the supply of lending to the real economy and a loss of liquidity in the capital markets. This in turn can depress economic activity reducing GDP and household income and increasing unemployment. The reduction in economic activity will naturally lead to further feedback in the form of a fall in energy use and reduced emissions.

Table 3 below presents a range of commonly recognised inter-model feedback mechanisms:

	Emissions Mitigation Policy	Earth Climate System	Macroeconomy	Energy System	AFOLU	Transportation	Financial System	Other Sectors
Emissions Mitigation Policy		Indirect impact	Top down view of policy impact including use of tax revenue	Carbon price reshapes energy system and reduces overall energy demand	Productivity impacts due to restrictions on fertilizers	Carbon price reshapes transportation and reduces overall demand	Operational costs, asset losses leading to increased financing costs, pension losses	Operational & transition costs
Earth Climate System	Policy response to unexpected climatic outcome		Impact of climate change on labour productivity	Operational & mitigation costs, productivity impacts	Operational & mitigation costs, productivity impacts (e.g. water availability)	Operational & mitigation costs, productivity impacts	Operational costs and asset losses leading to increased financing costs, pension losses	Operational & mitigation costs, productivity impacts
Macroeconomy	Policy response to unexpected macroeconomic changes	Change in global GHG emissions with expansion/contraction of economy		Change in demand - inflationary pressures impact financing of renewables	Change in demand	Change in demand	Change in demand, banking sector losses leading to increased financing costs, pension losses	Change in demand
Energy System	Policy response to unexpected energy system changes	Change in global GHG emissions with transition of the energy sector	Energy price changes impact global economy		Energy price changes impact cost of production influencing profitability and land use change	Energy price changes impact demand for different transportation types impacting global trade	Operational costs, asset losses leading to increased financing costs	Change in operational costs and demand
AFOLU	Policy response to unexpected land use changes	Change in land use impact GHG emissions and hence climate	Change in AFOLU prices impacts price of essential goods impacting residual income with implications for demand in other sectors	Demand for energy by type		Change in AFOLU prices impacts price of essential goods impacting residual income with implications for demand in other sectors	Asset losses leading to increased financing costs, pension losses	Change in AFOLU prices impacts price of essential goods, impacting residual income with implications for demand in other sectors
Transportation	Policy response to unexpected transition of the transportation sector	GHG Emissions	Transportation costs reshape global trade	Demand for energy by type	Transportation cost delta		Asset losses leading to increased financing costs, pension losses	Transportation costs impact production costs and hence demand/profitability
Financial System	Policy response to unexpected technological change	Immaterial direct GHG emissions	Banking sector losses leading to reduced lending to the real economy amplifying macroeconomic impact	Increased perception of risk leads to increased financing and insurance costs	Increased perception of risk leads to increased financing and insurance costs	Increased perception of risk leads to increased financing and insurance costs		Increased perception of risk leads to increased financing and insurance costs
Other Sectors	Policy response to unexpected transition	GHG emissions of varying materiality	Lower materiality impact on GDP	Demand for energy by type		Change in demand for transportation by type	Asset losses leading to increased financing costs, pension losses	

Table 3: Inter-Model Feedback Mechanisms

As with many aspects of climate economic modelling the degree of sophistication and detail that is appropriate will depend upon the intended use. The need to develop and understand more specific feedbacks affecting sub-sectors of the economy will be use case dependent. These should be addressed when applying the taxonomy by users with a particular interest. For example, a bank considering the financing of a new offshore wind installation should be particularly interested in the variation in power output and profitability across the range of potential future climatic conditions. Therefore, the bank would need a clear understanding of the methodology for capturing this feedback mechanism in any scenarios used to inform the decision-making process. However, the poor coupling of physical climate impacts onto the economy is a weakness of many scenarios.

2.2.4 Taxonomy Structure for the Model Levels - 2b and 3

We describe the modelling approach and calibration choices using a high-level classification system (2.2.4.1) supported by more detailed analysis (2.2.4.2). The aim of the high-level classification system is to provide an overview of a scenario component model in a one-page table, with listed options for some characteristics where appropriate. The detailed analysis provides a deeper free-form description of the modelling approach that should focus on the attributes relevant to the documenter or reviewer. Many aspects of this analysis are common across all components, as described in the next section. Component level idiosyncrasies and examples of the application of the taxonomy follow. However, as stated earlier, the examples should be seen as illustrative given the lack of detail in the publicly available documentation. We anticipate that the application of the taxonomy will become more sophisticated over time, particularly with respect to identifying sectoral and regional variations in the positioning of end output versus the range of credible outcomes.

2.2.4.1 A High-Level Classification System

The modelling approach should be classified using a combination of generic descriptors combined with component specific key features. The generic framework is as follows and can be used to present all components of an individual scenario or to compare a particular component across different scenarios. Details of the generic classification framework column entries with potential values follow. Combining the classifications for each component of a scenario with the model architecture classification provides an effective overview of the modelling approach for a full scenario:

Component ID			**Scale and resolution**				**Model type and complexity**				**Interfaces**			**Component specific**
Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	

Table 4: Generic Model Component Classification System

Component ID

The three component id fields position the model component in the taxonomy. The level can be 2b, 3a or 3b. Multiple lines for the same component may be required to capture regional and sectoral variations. The same model may also appear in multiple component lines where it performs roles across multiple layers in the taxonomy. The model type follows on from the level but adds a little more detail with the following options:

- Climate Model – Provides physical climate variables such as temperature, wind stress and precipitation.
- Macroeconomic Model – Provides macroeconomic pathways such as GDP, inflation, and unemployment.
- XXX Sector Model (XXX = sector e.g. Energy) – Provides sector specific information including market volumes, prices and emissions.

- Microeconomic Model – Models the financial performance or value of a firm or an individual asset such as property or another economic agent.
- Financial Market Factor – Provides market factors such as equity prices, credit spreads or default probabilities.

Where one of the above model components forms part of an integrated assessment model the type should have an IAM suffix. For example, Energy Sector Model – IAM would apply to the energy modelling component of the REMIND MAgPIE integrated assessment model.

Scale & Resolution

The four granularity and extent fields describe model coverage and the spatial, temporal and economic resolution at which the component operates though they will not all apply to all components. For more complex models where the resolution varies along a sub-model chain the entry should be based on the output resolution and the assumptions in the chain considered in the detailed analysis.

- Geographic Granularity and Extent- describes the resolution with which physical climate impacts are assessed and area covered. This might typically be expressed as a length or area (km or km^2) but could be post code, country, region, etc. It also describes the extent, which could be global or a specific region. It might also describe if there is averaging over the region.

- Economic Granularity – describes the resolution of economic output pathways such as GDP, unemployment, GVA or P&L. Potential values include global, country, region, firm, consumer. These values should be supplemented with the number of sub-divisions where appropriate, such as G7 countries plus the rest of world.

- Industrial Sector Granularity – provides additional detail for economic pathways with a sectoral analysis. The sectors covered should be listed or a reference to a standard classification system provided.

- Temporal Granularity & Extent – Models that provide outputs at discrete time intervals should be described as multistep with the time period (e.g. - 5 years) included. Models that are called to provide an output for a single set of inputs should be described as a single step. The temporal extent is the time period covered, for example 1860AD to 2100AD.

Model Type & Complexity

The model type and complexity fields provide a deeper understanding of the specific modelling process and the nature of the inputs and outputs. The objective is to provide a more granular description of model components that supports a detailed assessment of larger and smaller modelling components. This facilitates a rapid data driven view of models that quickly highlights model strengths and weaknesses before a full model validation is performed. Therefore, a description of the model and its parameters would be assigned the meta-data classification below for each climate model component defined in section 2.2.5 and 2.2.6.

- Nature of Model Formulation – This aspect of the taxonomy first captures an overall description of the model type. These descriptions will vary with the model component and so are free format.

- Model Process – Defines the evolution process used. Potential values include:
    - Deterministic Process – Deterministic model with no random variables modelled. E.g. linear regression models.
    - Discrete Stochastic Process – Stochastic process for discrete time steps, where one or more of the outputs is a function of a random variable. E.g. A Binomial tree model.
    - Continuous Stochastic Process – Stochastic process for continuous time steps, where one of more of the outputs is a function of a random variable. E.g. The Black-Scholes stochastic differential equation.
    - Probabilistic Timestep – A model that defines the outcome at a future time by sampling from a distribution or set of distributions. The model relies on information from a previous time step to define its output but may not be pathwise continuous.
    - Probabilistic Static – A model that defines its state based on sample from a distribution, where there is no previous time step information required for the output.
    - Climate Model Elements and Dimensions (see 2.2.5.1 for details) – Combinations of atmosphere, ocean, cryosphere etc. and dimensions from 0 to 3D.

- Calibration Type – Describes the nature of the input data used to calibrate the model component. Potential values include:
    - Empirical – Empirical parameters are given from direct objective observation of data; the data may be historical or a point in time.
    - Expert Judgement – Parameters are determined by expert judgment only.
    - Empirical Judgement Combination – Combination of empirical and expert judgment
    - Free format for climate models

- Model Uncertainty – The type of uncertainty or variance introduced to the model Component. A model may have several layers of uncertainty / variance, for example a stochastic process may be driven by a volatility term and the model may permit distribution uncertainty for its calibrated parameters that drive any process whether it is deterministic in the model framework or stochastic. Possible values include:

    - None - No model uncertainty
    - Parametric Discrete Exogenous - A parameter that takes a discrete value that is applied exogenously, for example driven by external inputs, but not varied as part of the model simulation.
    - Parametric Distribution Exogenous - A distribution of values that is fixed as part of the simulation driven by external calibration or other model input, e.g. a fixed Gaussian distribution / stochastic process
    - Parametric Endogenous - A parameter that is part of the simulation that changes the simulation framework, e.g. Values drawn from a Gaussian distribution that may

be driven by the results of the simulation, e.g. the mean of this distribution can be changed because of the simulation.
  - Parametric Expert Distribution - Uncertainty driven by expert judgement and not calibrated from empirical observation. Continuous distribution outcomes
  - Parametric Expert Discrete - Uncertainty driven by expert judgement and not calibrated from empirical observation. Discrete distribution outcomes.
  - Multi-Model Ensemble (MME) – Individual model runs do not incorporate uncertainty, but a range of outcomes is obtained by running the same scenario on multiple models (particularly relevant to climate models).
  - Perturbed Parameter Ensemble (PPE) – As for MME but the ensembles consist of many slightly different versions of the same base climate model in which key model parameters (such as those determining how the land surface or clouds are modelled) are varied to sample response uncertainty.

Interfaces

The interface fields describe how data is passed between the component being analysed and other elements of the modelling architecture covering inputs, outputs and feedback mechanisms.

- Key Inputs – List key inputs to the model such as emissions forecasts, carbon price curves and counterfactual GDP

- Key Outputs – Describes the nature of the model component output. Possible values include:
  - Climate Chronic – Climate pathway
  - Climate Acute – Acute climate event
  - Prices - Price / cost of a good or service
  - Demand - Demand of a good or service
  - Supply - Supply of a good / service
  - Emission Intensity Per Unit Produce - Emissions intensity of the entity per unit of production (where the unit of production must be provided e.g. GW or USD)
  - Market Financial - A traded financial market value
  - Product – e.g. energy in Giga Watts, this can also include waste products from a firm / economic activity. The product is specified in the output list with its unit.
  - GHG Emissions – specific emissions with assigned Scope (1,2,3), this is treated as an entity within this taxonomy outside of the Product classification.
  - Macro - Macro-economic variable
  - Agent Action – data on realisations of AGENT actions/outcomes made by modelled agent components where applicable.

- Integration – Describes the way the model component interacts with other components in the modelling chain and hence its ability to drive and respond to feedback. Potential values include:
  - Exogenous Input - The model relies on a separate exogenous model framework when integrated.

    - Endogenous Feedback - The model component is integrated with its input via feedback mechanism.
    - Calibrated Parameter Static Empirical - The model is driven by a fixed distribution input with no dynamics.

2.2.4.2 Detailed Analysis

Although a combination of the model architecture and component classifications provides a good insight into the design and high-level applicability of a scenario it is unlikely to provide sufficient detail for a financial user to identify all the relevant limitations and conclude on the appropriate uses. The detailed analysis of each component provides this more detailed insight and, when the conclusions for each component in a scenario are combined, allows the final assessment of the overall scenario.

The structure of the detailed analysis is as follows:

- Model Name(s) – The model (or models in the case of an ensemble) should be identified including version number.

- Detailed Model Description – A flexible piece of analysis that varies according to the application of the taxonomy:
    - For model builders this might be the standard model documentation but ideally written with summaries for non-specialists.
    - Peer reviewers might assess the builder's documentation drawing their own conclusions on the assumptions made and comparing them to alternative approaches.
    - Financial end users would analyse the documentation provided by the builder and any peer reviewers to draw out the details relevant to them.

- Limitations, Assumptions & Comparison with Alternative Models: The limitations and assumptions in the modelling approach and its calibration should be discussed, supported by comparison to alternative models and sensitivity analysis to convey the materiality of the weaknesses. The analysis would vary according to the application in a similar way to the detailed analysis. Having assessed the limitations and assumptions in the modelling approach the scenario reviewer should conclude on the credibility of the component and its coherence with the overarching scenario narrative.

- Summary & Recommended Use: The scenario reviewer should summarise the above sections and present a view on the appropriateness of the component for different use cases in the financial services sector. For a financial user this final analysis would focus on the potential use cases relevant to them

The heart of the detailed analysis is the detailed model description. This builds on, or more accurately feeds, the high-level categorisations in the classification table. It should consider:

- Model dynamics – The mathematical equations that define the model with an explanation of the inherent assumptions. Geographical and sectoral variations should be identified. For complex models, with variations in spatial and temporal resolution through the different links in a modelling chain, the impact of upscaling and downscaling should be considered.
- Input parameter calibration – The calibration choices made for the scenario, the range of alternative approaches and the associated input uncertainty. Geographical and sectoral variations should be identified.

- Experimental design – A description of how the selected model(s) is/are applied in the generation of a particular scenario. For example, multiple models (an ensemble) might be run using the same input parameters to provide a range of outcomes, reflecting model uncertainty, from which a single pathway is selected based on a chosen confidence interval. Alternatively, a single model might be run using a range of input parameters to determine a range of potential outcomes that recognise the uncertainties in the input parameters. The experimental design should inform whether pathwise outputs or distributional driven outputs are provided.

2.2.5 Level 2b – Macro Modelling

The level 2b assessment will typically follow the structure of an integrated assessment model considering the following components where applicable:

- Climate System
- Macroeconomics
- Sector Specific Models:
    - Energy System, inc. Carbon Capture and Storage
    - Agriculture, Forestry, Other Land Use (AFOLU)
    - Additional Sector Specific Models
    - Other Sectors – Generic Treatment

In the following sections of the paper, we consider how the taxonomy applies to each of these macro modelling components.

2.2.5.1 Climate System

Climate model information on present day and future hazards forms a potentially key component of physical risk analysis for financial institutions (Fiedler et al., 2021). However, amongst the plethora of climate model data a taxonomy is needed to categorise the type of information that is provided and its suitability.

Here we use the taxonomy shown in table 4. The generic part of the taxonomy includes categorisation of the functional form of the climate modelling approach, and its main characteristics and interfaces. The classification system is not extended beyond the generic elements for climate models, but the detailed analysis focuses on the credibility of the model and is framed in terms of aspects such as the agreement with real world observed climate and weather, comparison with other models, as well as transparency of the model's construction and its theoretical underpinning. These aspects are loosely related to the information categories of Cash et al. (2002).

In the detailed analysis we also recognise that it is not only the characteristics of the model that must be included in the taxonomy but also the details of the climate model set up or experimental design. This can include how the model has been initialised (the start conditions) and whether the model is run as part of an ensemble.

Classification system

Component ID			**Scale and resolution**				**Model type and complexity**				**Interfaces**		
Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration
				N/A	N/A								

Table 5: Climate System Model Component Classification System

Functional categorisation in the generic aspects of the taxonomy as applied to climate system models: The application of the generic taxonomy template to climate models is based on categorising models in terms of spatial and temporal resolution, complexity, and uncertainty representation (e.g. McGuffie and Henderson-Sellers, 2014). Whilst resolution, complexity and ensemble size were often originally discussed in terms of computational limitations this categorisation is also useful here as a way of describing the models in the taxonomy. The application of the generic section of the template differs from the other economic style models and the following definitions adapt those introduced earlier:

Scale and resolution/geographical granularity and extent: The geographical granularity or spatial resolution describes the spacing of the horizontal and vertical grid on which the model data are produced. For instance, global climate models used in CMIP6 have a typical horizontal grid spacing of around 100km and, as the name suggests, cover the entire globe. Regional climate models, such as those in CORDEX, typically have a spatial resolution of around 12km, but cover an area such as Western Europe. A newer class of model, the so-called convective permitting models (e.g. Kendon et al., 2019), have a finer spatial scale that enable the explicit simulation of some aspects of atmospheric convection but often only cover a single nation. At the other end of the spectrum, some reduced complexity climate models are set up to solve equations that represent global mean behaviour. The process of taking results from a coarse model and converting it to information at a finer spatial scale is referred to as downscaling. The simplest, and least satisfactory, approach is simple linear interpolation. More complex statistical and dynamic approaches take account of the physical processes and aspects such as topography that determine the climate outcome on the finer spatial scale.

Temporal granularity and extent: Climate model projection experiments typically begin around year 1850 or 1900 and are driven by historical greenhouse gas concentrations or emissions up to present day, then scenarios into the future. The future scenarios may run for decades, but typically run to 2100 and sometimes 2200 or 2300. Climate model predictions are usually started from present day conditions, e.g. by assimilating the near present current state of the ocean, and then run forward for up to 10 years (Meehl et al. 2009). Physical climate models, which explicitly solve the equations governing the dynamics and thermodynamics of the atmosphere, typically have a characteristic time-step. For the highest spatial resolution models this is usually a few minutes. However, the maximum temporal resolution supplied to the user community is often less than this. For global and regional models this is usually daily information, but for convective permitting models it may be much shorter. This can be important for the consideration of extreme weather events.

Economic granularity and industrial sector granularity: These fields are not typically represented in physical climate models, so these columns remain blank in the typology table.

Model type and complexity

Nature of model formulation: This aspect of the taxonomy first captures an overall description of the model type. Many climate models are physical based models constructed using conservation equations, for instance of energy, momentum, and mass. Some of these models represent the three-dimensional earth system whereas others are referred to as models of reduced complexity or simple climate models. Often these average over one or more spatial dimensions. The simplest models represent global average climate response. There is growing use of model emulators and non-physically based models based around artificial intelligence approaches which might be encountered when completing the taxonomy information. Potential categories for this element of the taxonomy include:

- Physically based,
- AI/ML based or
- Hybrid physical-AI/ML based.

Additional information can be added as freeform text.

Model processes: This category captures aspects of the model complexity. Three-dimensional climate models can come in various forms, notably atmosphere only, ocean only and couple models (which have both atmosphere and ocean). The atmospheric models also typically have a representation of the land surface, whilst ocean models include a treatment of sea-ice. The treatment of atmosphere and ocean will vary in complexity but typically represent circulation and mixing processes, including some treatment of convection. Earth system models include additional complexity to also simulate aspects such as the carbon cycle, which requires a treatment of vegetation, soil and ocean carbon uptake and release processes. Some earth system models are starting to include other cycles such as methane or nitrogen cycles. Those models that do not have adequate earth system processes need to be driven by atmospheric concentrations of greenhouse gases from other gas-cycle models, whereas earth system models can increasingly be driven directly with emissions rather than concentrations. This element of the taxonomy should provide an overview of the model approach. This element of the taxonomy should include:

- The degree of dimensionality of the model chosen from 0D, 1D, 2D or 3D
- The elements of the climate model that are simulated explicitly from atmosphere, land, ocean, cryosphere, and biosphere.

Additional elements or detail can be added as freeform text. This might include clarifying which elements of the cryosphere are modelled, such as sea-ice, glaciers, and the large ice sheets. It might also clarify which aspects of the biosphere are modelled.

Calibration type*:* It is typical for climate models to contain parameters that cannot simply be observed as a best estimate value. Instead, a calibration of tuning approach is used. For a model of reduced complexity such as FAIR (Smith et al. 2018) this might involve selecting values of e.g. the equilibrium climate sensitivity based on literature estimates. For a more complex general circulation model (GCM) it might include activity adjusting several parameters. For example, to achieve an acceptable simulation of some defined aspects of the present-day climate, such as the near-surface temperature and the top of the atmosphere radiative balance, the scheme that represents clouds might require adjustment.

This element of the taxonomy should include a categorisation of the form:

- No tuning/calibration
- Calibration against the observed climate
- Selection of key model parameters from the literature
- Combination of calibration against observations and literature selected values

Additional detail should be freeform description and could include mention of which parameters are tuned or point to a citable reference on the calibration scheme.

Model Uncertainty: All climate model experiments are subject to uncertainty. Uncertainty in the emissions or concentrations of greenhouse gas that are presented to the model are captured by other components in the modelling chain. This element of the taxonomy is for recording the approach to addressing uncertainty within the climate model in the experimental set up that is being used for the particular, financially relevant, application. Climate projections also consider natural variability as an uncertainty, whereas climate predictions attempt to latch on to an observed state of natural variability and follow it forward in time for a season to a few years.

To characterise uncertainty in simulated response climate models are often run as part of an ensemble or collection of models where the type of ensemble will depend on the type of question being addressed. Initial condition ensembles use the same model run many times starting from different initial states and are used to characterise natural variability in the climate system. Larger ensembles enable exploration of longer return period events. Multi-model ensembles typically consist of sets of different models, often designed, and constructed by different research teams, and enable the study of structural uncertainty of model response. Perturbed parameter (or perturbed physics) ensembles consist of many slightly different versions of the same base climate model in which key model parameters (such as those determining how the land surface or clouds are modelled) are varied to sample response uncertainty. In addition to ensemble setup, we also recognise the need to define a scenarios of future greenhouse gas forcing for climate projection experiments that run into the future. These might be selected from the shared socio pathways set (Riahi et al 2017) and might range from the low emissions SSP1-1.9 to the high emission SSP5-8.5. Alternative scenario sets include the older RCPs (Moss et al. 2010) and the NGFS scenarios (NGFS 2021). This element of the taxonomy should at least select which types of climate uncertainty are covered from:

- Emissions uncertainty
- Model response uncertainty (e.g. using a multi-model ensemble)
- Natural variability (e.g. using a perturbed parameter ensemble)

More than one option can be chosen. Further description of the uncertainty approach or appropriate citations should be added as freeform description.

Interfaces

Inputs: For climate models inputs include greenhouse gas emissions or concentrations for a range of species including carbon dioxide, other Kyoto gases, and in many cases additional radiatively active species such as atmospheric aerosol particles. For spatially resolved models sometimes a time varying input is the land surface type, e.g. grass, tree, urban etc. Models that only simulate the land surface and atmosphere also require sea-surface temperatures as inputs. Typically, a climate modeller will also set other parameters as an input, such as details of schemes such as gravity wave drag or convection. However, we consider these an element of the calibration. For spatially resolved limited area models, often called regional climate models, lateral boundary conditions derived from e.g. a global climate model, are also needed as input. The input to the climate model can be freeform text but should at least describe which greenhouse gas forcing is applied.

Outputs: The output from climate models varies with the degree of complexity. For complex earth system models, based on general circulation models, these can include meteorological metrics such as near surface atmospheric temperature, temperature on vertical levels, precipitation, wind speed components, pressure at mean sea-level, vorticity, relative and specific humidity, cloud fraction, convective available potential energy, and radiative flux components. For models with an ocean component, metrics such as salinity are also available. Spatially resolved models output these “fields” covering the domain on some pre-set temporal frequency that is often a multiple of the model time-step. Many studies focus on daily data when considering extremes, although sub-daily data is available from some models. Reduced complexity models tend to have fewer outputs, with the simplest only providing radiative forcing and near surface global mean temperature. A typical complex climate model output list may contain more than 100 quantities. It is suggested that the response to this element is a list of the most relevant outputs for the application in question, such as temperature, windspeed and rainfall.

Integration: The degree of integration describes the links from the climate models to other models in the taxonomy. Climate modellers often use the terms one-way and two-way coupling to describe the degree of integration. One-way coupling implies the climate model receives input from another model but does not provide its output back into other models, alternatively it receives no direct input from another model but does provide output. Two-way coupling implies the climate model can both receive inputs and provide outputs to other models in the modelling chain. The generic categories provided in section 2.2.4.1 can be applied to climate models. Endogenous feedback corresponds to two-way coupling. Exogenous input and calibrated parameter static empirical are forms of one-way feedback.

In practice, most reasonably sophisticated climate scenario generation architectures will incorporate at least a degree of two-way coupling. More freeform detail of the feedback mechanism should be appended to the basic classification, where useful.

Detailed analysis

In addition to the tabular data, it is often useful to provide additional information on the model description, the limitations of the model and the recommended use cases. Each of these can be provided as a freeform response. Considerations applying to climate models are as follows:

The model description should expand on the table entries by providing a description of the model and primary citations for the reader to find more detail. This may include information on how the model has been tested. The model description may include some aspects of describing the theoretical underpinning of the model. Where there is limited description of the modelling methodology and formulation available, or limited evidence of testing then it is the view of the authors that users should proceed with caution.

Limitations of the model should be described, including information on how the model has been tested. In many cases it will be up to the user of the model to assess the limitations in the circumstances for which the model is to be used. For models used to examine the climate change response for a future period it might be prudent to consider both the performance of the model in simulating the present day and its simulations of the future, as put forward by Giorgi (2019), Baumberger et al. (2017) and Baldissera Pachetti et al. (2024) in their assessments of climate model credibility. Reliability describes the performance of a model, when it is used to simulate past or present conditions, with observations of the real world. Models may be found to perform well in some regions and not others, or in simulating some metrics and not others. A difference between the model and observations, known as the bias, may be corrected using a variety of techniques known as bias correction (Gohar et al. 2017, Maraun 2016). However, this must be done with caution as this bias correction can lead to unphysical behaviour. Furthermore, how much a particular bias matters can be dependent on the question being addressed with the climate model. Robustness describes the model performance compared to other models, and is usually applied to the future simulated period. Does the model in question perform similarly to other models of a similar type and construction? Is the model an outlier? And if so, is the cause of this known and how desirable is it for the application? Like reliability, the importance of robustness depends on application.

Recommended use cases can provide a broad guide to the types of application to which the model has been previously applied. Typically, a use case will consist of both the model and the experimental design in which the model is setup and used. When evaluating use cases both aspects must be considered. When considering the recommended use cases, and especially when going beyond existing published use cases it is important to ask questions such as: are the processes that the model is representing understood to be important in the real world included in the model? For instance, a climate model that is intended to examine changes in convective storms would need to be able to simulate appropriate scales of atmospheric convection. A model intended to examine uptake and storage of carbon would need a suitable treatment of the equations describing the carbon cycle.

Taxonomy application example

For illustration we apply the taxonomy to a general circulation model (GCM). As this is for illustration, we select a well-known but older climate model, HadCM3, which is still used for e.g. long simulations requiring a faster GCM. The commonly available IAM based scenarios use less sophisticated climate models. Although scenario developers, such as the central banks, augment IAM output with more detailed GCM derived climate data the way in which the more sophisticated climate model has been integrated is often opaque. We therefore present a generic view of how HadCM3 might be used.

Scenario	**Component ID**			**Scale and resolution**				**Model type and complexity**				**Interfaces**		
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration
Conceptual	2b	HadCM3	Climate	Global coverage - Atmosphere: 19 levels at 2.5 lat by 3.75 long degrees horizontal resolution. Ocean: 20 levels at 1.25 by 1.25 degrees horizontal resolution.	N/A	N/A	Data available down to daily resolution, Policy relevant simulations typically run from 1850 to 2100	Physically based model with parameterisations of sub-gridscale processes	3D model includes atmosphere, ocean and sea ice. Model uses a finite difference approach with sub-gridscale processes parameterized	Tuned to give reasonable agreement with present-day. Biases remain	Uncertainty can be studied using ensemble. Can be run as part of a PPE or MME	Greenhouse gases and precursor concentration	Physical atmosphere, ocean and land metrics such as temperature, humidity and sea-ice concentration	Dependent on scenario specific implementation

Table 6: Climate Model Component Classification for the HadCM3 general circulation model

Scenario: Conceptual (Publicly accessible scenarios are not built directly using HadCM3)
Macroeconomic Model: HadCM3
Detailed Model Description: The model is a coupled atmosphere-ocean general circulation model first used more than 20 years ago. It explicitly simulates the dynamics of the atmosphere and ocean plus also includes treatment of land surface processes and sea-ice. The model is described in detail by Johns et al. (2003), which highlights the key innovation compared to earlier models of not needing flux-corrections to compensate for major errors in heat and moisture transports. The key features are summarized here.

The atmospheric component of the model, HadAM3 (Pope et al. 2000), has 19 levels with a horizontal resolution of 2.5 degrees of latitude by 3.75 longitude, which equates to around 250-

300km. The model includes a radiation scheme with six spectral bands in the shortwave range and eight in the longwave range and treats the effects of minor greenhouse gases as well as CO_2, water vapour and ozone. A penetrative convective scheme is used, modified to include an explicit downdraught and the direct impact of convection on momentum. Parameterisations of orographic and gravity wave drag have been revised to model the effects of anisotropic orography. The large-scale precipitation and cloud scheme is formulated in terms of an explicit cloud water variable following Smith (1990). The effective radius of cloud droplets is a function of cloud water content and droplet number concentration. The atmosphere of HadCM3 also includes the capability to model the transport, chemistry and physical removal processes of anthropogenic sulphate aerosol which is input to the model in the form of surface and high-level emissions of SO_2.

The land surface scheme includes a representation of the freezing and melting of soil moisture, as well as surface runoff and soil drainage; the formulation of evapotranspiration includes the dependence of stomatal resistance on temperature, vapour pressure and CO_2 concentration.

The oceanic component of the model has 20 levels with a horizontal resolution of 1.25 degrees of latitude by 1.25 longitude and uses a rigid lid formulation. At this resolution it is possible to represent important details in oceanic current structures (Wood et al. 1999). Horizontal mixing of tracers uses a version of the adiabatic diffusion scheme of Gent and McWilliams (1990) with a variable thickness diffusion parametrization (Wright 1997; Visbeck et al. 1997). Near-surface vertical mixing is parametrized by a Kraus-Turner mixed layer scheme for tracers, and a K-theory scheme (Pacanowski and Philander 1981) for momentum. Below the upper layers the vertical diffusivity is an increasing function of depth only. The sea-ice model uses a simple thermodynamic scheme including leads and snow-cover. Ice is advected by the surface ocean current, with convergence prevented when the depth exceeds 4m. There is no explicit representation of iceberg calving, so a prescribed water flux is returned to the ocean at a rate calibrated to balance the net snowfall accumulation on the ice sheets, geographically distributed within regions where icebergs are found. To avoid a global average salinity drift, surface water fluxes are converted to surface salinity fluxes using a constant reference salinity of 35 PSU.

Limitations of the HadCM3 model: Like all general circulation model HadCM3 exhibits biases, which are evident when the simulations driven with present day greenhouse gas forcings applied are compared with observations. Some of the major biases are reported by Johns et al. (2003) and Gordon et al. (2000). In terms of atmospheric circulation, a major bias in the Northern Hemisphere is that the Icelandic low is too shallow and the gradient too slack in the Atlantic storm track region in winter. The North Pacific storm track perhaps extends too far to the east compared to the analyses. However, the Asian monsoon trough is well captured in the model, as are the subtropical ocean high pressure systems. In general, the model does a good job of capturing the patterns of mean seasonal precipitation, especially over land areas. However, there is rather a split inter-tropical convergence zone response over the Western Pacific Ocean. Other biases include the model tending to overdo the precipitation in the eastern tropical Atlantic. Collins et al. (2001) show that ENSO-like tropical variability exhibited by the model is reasonably realistic in structure and amplitude, but biases remain. Given that many newer climate models have been developed it is important to keep in mind that compared to more contemporary models the biases in the present day for HadCM3 are typically larger for many aspects than the new models.

When considering the climate change response, attribution studies using optimal fingerprinting approaches show signal strengths for the model response that are compatible with the large-scale warming signals in the real world (Stott et al. 2010). The equilibrium climate sensitivity[3] of the model has been estimated to be around 3.3K, which is in the likely range of more contemporary studies using multiple lines of evidence (e.g. Sherwood et al., 2020). The model captures key large-scale warming features such as the land-sea contrast and larger warming over high northern latitudes.

Whilst the performance of HadCM3 was considered good at the time of its development, it is important to remember that newer global climate models are run with greater spatial detail (often in the range of 50-100km) and include improved parameterisations of a range of physical processes that cannot be represented explicitly at these scales. For instance, many newer models can capture more realistic atmospheric behaviour, including mid-latitude blocking. A particular concern is that some of the apparent quality of the historic climate simulated in HadCM3 may have been due to compensating errors.

As with all coarse global climate models, one should exercise care when using the model to examine weather extremes, such as high (or low) temperatures and precipitation. Furthermore, it is important for many applications to sample uncertainty, and this typically requires running multiple emission scenarios and comparing the model's responses with other climate models, as is done in the IPCC assessment reports. Whilst the base version of the model considered here requires input in the form of greenhouse gas concentrations, a later version HadCM3C includes a carbon cycle module, which allows carbon dioxide emissions to be used instead of concentrations.

Recommended uses cases for the HadCM3 model: Over its lifetime HadCM3 has been used in many hundreds of (possibly more than a thousand) studies. The first class of use cases involve applying the model in detection and attribution studies to explore the emerging signal of climate change in the real world compared to that in the model.

Many studies have used the model to examine the climate response to a range of alternative greenhouse gas concentration futures, to understand the impact of emission mitigation. These ideally compared the results with other structurally different models. Whilst it could be used to examine changes in sea level, looking at the spatial patterns required a considerable amount of post processing to overcome the limitation of the rigid lid ocean formulation and lack of ice sheet modelling. In the UK, the model was used as part of earlier production of UK national climate scenarios, including UKCP09 and a more limited role in UKCP18. For these applications many alternative versions of the model were used with perturbations made to key model parameters to sample uncertainty.

[3] Equilibrium climate sensitivity is defined as the equilibrium temperature response that corresponds to a doubling of atmospheric CO2 concentration.

For applications requiring more local simulations the HadCM3 model was typically dynamically downscaled using regional climate models to provide results at 25-50km (or below) scales. Applying downscaling should be the preferred approach when looking at local weather and climate extremes.

In addition to climate projections, a version of the model was used in the initial DePreSys decadal forecast system. This assimilated the near-current state of the climate system and aimed to simulate the next few years, attempting to capture the effects of both initial conditions and the climate forcing out to 10 years.

Unless computational limits are an issue, we would generally recommend using new models than HadCM3 for most applications today. However, there is still a use for HadCM3, for instance in producing very large ensembles, simulating very long periods in paleo climate studies, or for training purposes.

2.2.5.2 Macroeconomic Modelling

The macroeconomic model projects key economic variables such as GDP, unemployment, interest rates, exchange rates and house prices. The nature of the macroeconomic model is closely associated with the modelling architecture. Many climate scenarios construct the macroeconomic pathway using a top-down approach influenced by a limited range of variables from the other model components. For example, energy prices and temperature changes. At the other extreme, where an agent-based model is employed, the macroeconomic pathways are constructed by aggregating microeconomic performance at the resolution of the agents.

As with climate models many scenarios use multiple macroeconomic models. A low sophistication model to consider the feedbacks with the other elements and a more sophisticated macroeconomic model to produce the final macroeconomic pathways once climate pathways and the quantitative representation of mitigation policies have been set. However, there is often loss of information between the two models with the final macroeconomic model (e.g. NiGEM) using simple physical damage functions and carbon price interpretations that lose the sectoral granularity present at the integrated assessment level. The assessment should consider each macroeconomic model separately considering how the outputs from the different models are used by the wider architecture.

Classification system

Component ID			**Scale and resolution**				**Model type and complexity**				**Interfaces**			**Component specific features**			
Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Climate damage function	Mitigation policy representation	Transition frictions	Financial frictions
				N/A	N/A												

Table 7: Macroeconomic Model Component Classification System

Component specific features including nature of model formulation

Nature of Model Formulation: We identify the following high-level approaches to macroeconomic modelling based on the classifications presented by Monasterolo et al 2022:

- Dynamic Stochastic General Equilibrium (DSGE)
- Computable Generalised Equilibrium (CGE)
- Non-Equilibrium Macroeconometric
- Ramsey Type Optimal Growth Model
- Agent Based Model

Climate Damage Function: The approach to modelling the impact of physical climate change is often a function of the granularity of the economic modelling. The approach should be viewed in combination with the geographic granularity which provides the resolution at which the damage function is evaluated. Possible values include:

- None
- Temperature Based Historic Regression

- Multiple Climate Factor Historic Regression
- Endogenously Determined at Sector Level
- Endogenously Determined at Agent Level

Mitigation Policy Representation: Ideally the parametrisation of mitigation policy should flow down from level 1 of the taxonomy, but the nature of the macroeconomic model may force simplifications. Potential values include (multiple selections are possible):

- Carbon Tax
- Carbon Price – Incorporates taxation and behavioural components.
- Carbon Border Adjustment Mechanism
- Sector Specific Domestic Restrictions
- Sector Specific Import/Export Tariffs and Restrictions

Transition Frictions: An economic transition requires the transfer of labour and capital between sectors and sub-sectors. Simple economic models might assume that this happens in a frictionless way whereby labour moves as required between sectors with no impact on unemployment or wage levels. In practice it is likely that there will be some lag that drives up unemployment and, where competitive advantage is lost, a fall in wages. Additional transition frictional effects may be present in more advanced models. Potential values include (multiple selections are possible):

- Frictionless
- Exogenous Unemployment Factor
- Exogenous Wage Factor
- Endogenous

Financial Frictions: Falls in profitability in the real economy or changes in market expectations for future financial performance can lead to losses in the financial sector. If these losses are significant then this can lead to reduced appetite and associated increases in the cost of finance for bank lending and the capital markets. Potential values include:

- Frictionless
- Judgement based adjustment of financing curves
- Historic regression-based adjustment of financing curves
- Endogenously Determined at Financial Sector Level
- Endogenously Determined at Agent Level

Taxonomy application example

The NGFS scenarios provide the most readily accessible examples of the application of energy sectoral models in the construction of climate scenarios. The NGFS employs three IAMs, GCAM, REMIND MAgPIE and MESSAGE-GLOBIOM. We have chosen the MESSAGE GLOBIOM IAM (Krey et al 2020) to illustrate the application of the taxonomy to the modelling of the macroeconomy, the energy and AFOLU sectors as it has some of the most extensive publicly available documentation. More specifically we have selected the latest generation of the NGFS Net Zero 2050 MESSAGEix- GLOBIOM 1.1 scenario to illustrate the detailed application of the taxonomy.

However, this documentation does not provide full detail and the output variables available through the NGFS portal are limited. The code is publicly available to review and run but this level of analysis is beyond the scope of this paper:

Model component classification

Scenario	Component ID			Scale and resolution				Model type and complexity				Interfaces			Component specific features			
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Climate damage function	Mitigation policy representation	Transition frictions	Financial frictions
NGFS REMIND-MaGPIE 3.2-4.6 Net Zero 2050	2b	MACRO	Macroeconomic	Global with country level macroeconomic projections	6 commercial energy demand categories	Energy only	5-year time steps to 2100	Ramsey type optimal growth	Deterministic	Empirical judgement combination	None	Region temperatures, energy system costs, potential GDP growth rates, projected growth of labour, autonomous energy efficiency improvement coefficients	National GDPs, government revenue, energy system investment, industrial production	Integrated energy sector	Temperature based historic regression	Carbon price	Potentially omitted but unclear	Frictionless

Table 8: Macroeconomic Model Component Classification for the NGFS MESSAGE-GLOBIOM Net Zero by 2050 Scenario

Scenario: NGFS Net Zero 2050 MESSAGEix-GLOBIOM 1.1 M-R12 Net Zero 2050

Macroeconomic Model: MACRO

Detailed Model Description: The MESSAGEix model is integrated with the MACRO general equilibrium macroeconomic model within the MESSAGE-GLOBIOM IAM framework. The model documentation (Krey et al. 2020) describes MACRO as follows:

"MACRO maximises the intertemporal utility function of a single representative producer-consumer in each node (or macro-economic region). The optimization result is a sequence of optimal savings, investment, and consumption decisions. The main variables of the model are the capital stock, available labour, and commodity inputs, which together determine the total output of an economy according to a nested constant elasticity of substitution (CES) production function. End-use service demands in the (commercial) demand categories of MESSAGE are determined within the model, and are consistent with commodity supply curves, which are inputs to the model. Labour supply growth is also referred to as reference or potential GDP growth. In the absence of price changes, energy demands grow at rates that are the approximate result of potential GDP growth rates, reduced by the rates of overall energy intensity reduction. Price changes of the six demand categories, for example induced by energy or climate policies, can alter this path significantly.

MACRO's production function includes six commercial energy demand categories represented in MESSAGE. To optimize, MACRO requires cost information for each demand category. The exact definitions of these costs as a function over all positive quantities of energy cannot be given in closed form because each point of the function would be a result of a full MESSAGE run. However, the optimality conditions implicit in the formulation of MACRO only require the functional values and its derivatives at the optimal point to be consistent between the two models. Since these requirements are therefore only local, most functions with this feature will simulate the combined energy-economic system about the optimal point. The regional costs (of energy use and imports) and revenues (from energy exports) of providing energy in MACRO are approximated by a Taylor expansion to first order of the energy system costs as calculated by MESSAGE. From an initial MESSAGE model run, the total energy system cost (including costs/revenues from energy trade) and additional abatement costs (e.g., abatement costs from non-energy sources) as well as the shadow prices of the six commercial demand categories by region are passed to MACRO. In addition to the economic implications of energy trade, the data exchange from MESSAGE to MACRO may also include the revenues or costs of trade in GHG permits."

The documentation identifies the following two as the key exogenous calibration parameters:

- Projected growth rate of labour (also known as potential GDP) which combines labour growth and labour productivity growth.
- Autonomous energy efficiency improvement coefficients (AEEIs) – the annual rate of reference energy intensity reduction.

However, the detailed calibrations used in the NGFS net zero by 2050 scenario are not provided so we are unable to assess how these have been positioned against the plausible range of values.

Limitations, **Assumptions & Comparison with Alternative Models:** A full analysis of the limitations of the macroeconomic component of the scenario requires detailed analysis of the modelling approach and its calibration. Although this is beyond the scope of this paper, there are commonly recognised limitations of IAM macroeconomics that are well illustrated in figures 7 to 9 below. The model produces a smooth evolution to a new equilibrium that does not recognise the short-term economic stresses that could occur through the transition. Given the speed of decarbonisation assumed, huge stresses are likely. This is particularly true if the scenario is updated to recognise the failure to follow the anticipated emissions pathway between 2020 and 2024 whilst maintaining the same total carbon budget and thus anticipated peak warming. Some of the optimism in the scenario is linked to the speed at which the energy system can transform as carbon taxes are increased. This is considered in the energy modelling section. However, another consideration that hides potential economic stresses is the absence of a financial services sector. As a result, the feedback mechanisms to the macroeconomy, associated with losses in the financial sector, are omitted.

The NGFS explorer provides access to several outputs from the MACRO model for the MESSAGE-GLOBIOM based scenarios. The intermediate outputs do provide some insight into the positioning of the scenario. We have chosen GDP(PPP), Investment (Energy Supply) and Government Revenue (Tax) to illustrate the type of analysis possible. Some very material differences can be seen between the three NGFS IAMs, even though the same scenario is being modelled and the modelling approaches appear superficially similar. Clearly, the definitions of these outputs might vary as even the starting levels differ materially in some cases. When taken in conjunction with the analysis of the energy model intermediate outputs presented later in the paper, even the relative similarity of the GDP projections is thought provoking. MESSAGE-GLOBIOM projects far higher final energy use, despite a relatively similar global GDP. This implies that the global economy is less energy efficient in this model.

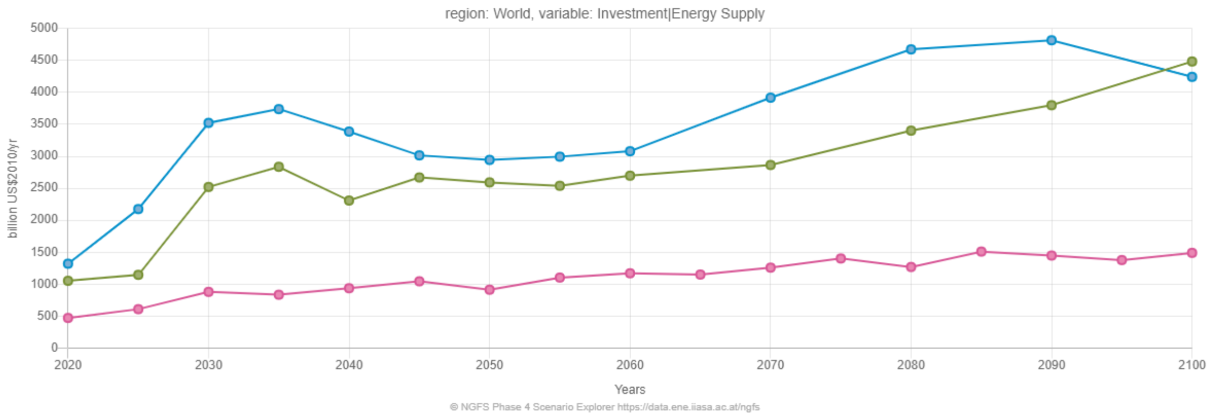


Fig. 7 – NGFS Explorer output showing Investment in Energy Supply for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

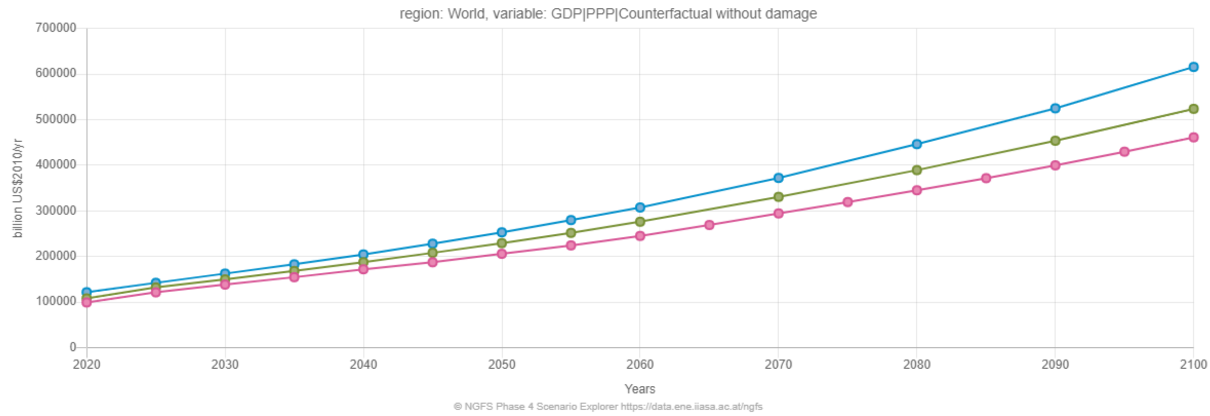


Fig. 8 – NGFS Explorer output showing GDP(PPP) counterfactual without climate damages) for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

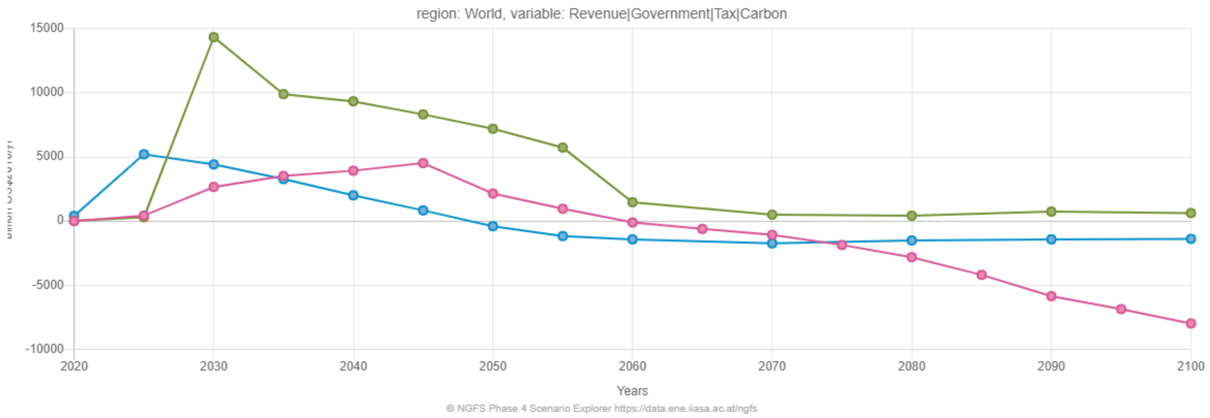


Fig. 9 – NGFS Explorer output showing Government Carbon Tax Revenue for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

Summary & Recommended Use: The macroeconomic approach in the MESSAGE-GLOBIOM IAM has been developed to support policy analysis. In this scenario it provides a smooth view of how a progressively increasing global carbon tax might drive the transition. As such, it provides an optimistic target view of how policy might deliver the desired climate outcome. It is therefore inappropriate for stress testing purposes in the financial services sector as it does not allow for the significant financial stresses that could occur. Although the model's projection of global GDP sits between those of the other two NGFS IAMs, these results are heavily influenced by the counterfactual GDP growth assumptions and the absence of any physical climate change damages. Further, the loss of sectoral information leaves the end user to assess the impact on sectoral level of GVAs

2.2.5.3 Sector Specific Models

The core functionality of an economic climate scenario modelling architecture is the ability to capture the relationships between economic activity, GHG emissions, emissions mitigation policy and climate change. Neither standard macroeconomic models nor physical climate models have this capability, so additional modelling of economic and land use sectors must be performed in sufficient detail to capture these relationships.

The appropriate sophistication of the approach will depend on the materiality of the sector and the intended use of the scenario. Current climate scenarios often use integrated assessment models with sector specific modules, although the required logic could be embedded in a full economy agent-based model or a granular macroeconomic model. Regardless of the architecture adopted we should separate out each sector with specific logic for assessment. However, the architecture does play a significant role in determining the way the various sectors interact within the overall model.

Figure 10 below provides a breakdown of anthropogenic GHG emissions demonstrating the materiality of different elements of the economy from this perspective. This clearly demonstrates why the energy supply and AFOLU sectors typically have dedicated sector specific models. Additionally, there must be generic logic for assessing less material sectors. The sophistication with which the intermediate sectors, such as energy intensive activities like transportation and real estate, are handled will vary. In this paper we present a standard approach to assessing sector specific models and discuss the AFOLU and energy sectors in more detail. The application of the taxonomy to any specific scenario will require the user to consider each sector specific module and any geographic variations.

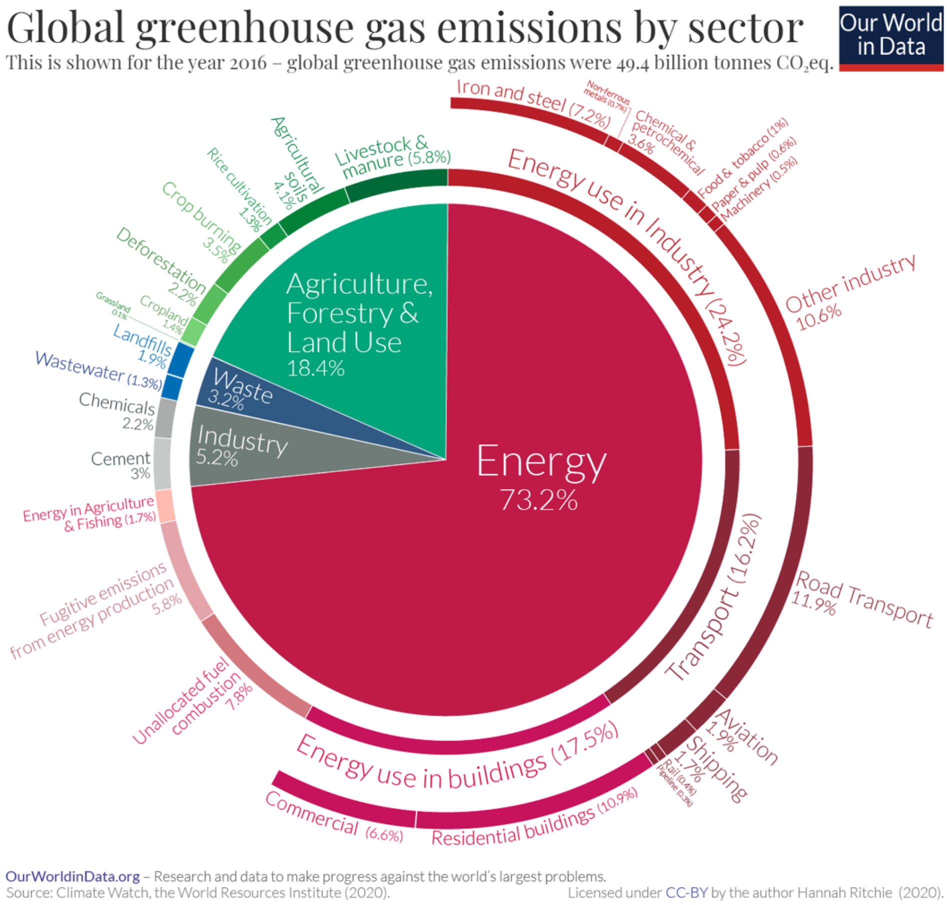


Figure 10: 2016 Global greenhouse gas emissions by sector (Our World in Data, 2016)

Of course, each economic sector has its own idiosyncrasies and modelling the economic and sectoral emissions can be complex, even before climate change is considered. Climate scenarios normally use simple parameterisations of the processes involved but the models for the most significant sectors are still often multi-layered with a bewildering array of input parameters. The challenge for financial users of climate scenarios is to understand the sources of uncertainty and how material they are to their businesses. To understand the sources of uncertainty we must consider:

- The accuracy of any parameterisations used.
- The uncertainty in the inputs that calibrate the parameterisations.
- The sensitivity of the user's business model to the resultant uncertainty.

For example, a biological process that sits at the heart of modelling the AFOLU sector might have a well-established parameterisation, calibrated with tightly defined input parameters, resulting in an intermediate output that is both accurate and precise. However, the response of the process to climate change might be less well defined due to the uncertainty in the projection of physical climate change variables, such as precipitation, which impact the process.

Unfortunately, the detailed assessment of model assumptions and the associated sensitivity analysis of outputs is not commonly available for the sector level models employed in the construction of climate scenarios. We provide illustrative examples for the AFOLU and energy sectors, but more robust application of the taxonomy by financial users will be dependent on scenario builders and academic reviewers developing the quality of documentation and analysis available.

Classification system

A sector specific model[4] must consider the interaction between the physical processes involved in the activity and the market forces that drive demand as presented in figure 11 below.:

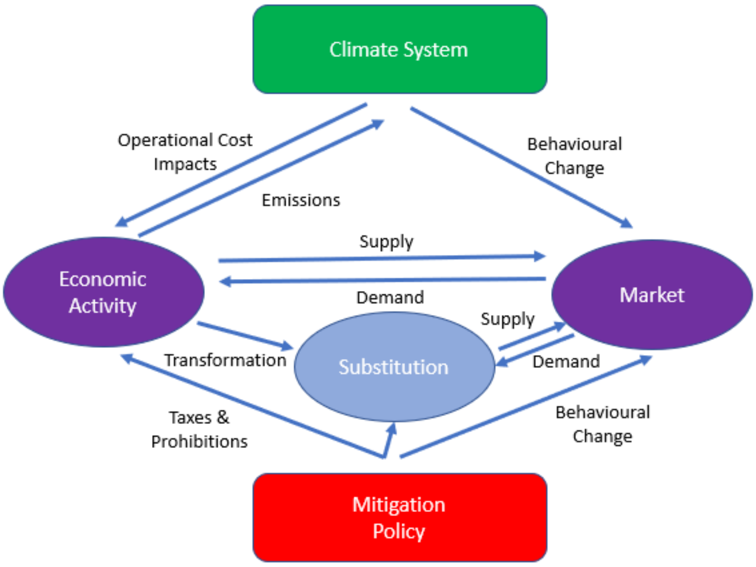


Figure 11: Economic Sector Schematic

This provides a natural set of sector model characteristics to add to the generic classifications presented earlier. These characteristics should be summarised using the table below but explored in more detail in the analysis of the model dynamics and input calibration:

Component ID			**Scale and resolution**				**Model type and complexity**				**Interfaces**			**Component specific credibility**			
Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Climatic feedbacks	Mitigation policy representation	Substitution dynamics	Market dynamics

Table 9: Sector Model Component Classification System

Nature of Model Formulation: When describing a sector model, we must consider the sophistication with which the underlying activity is modelled and the way that market dynamics are represented (captured in the last element of this component taxonomy). Modelling of the underlying activity provides the supply side information for the economic modelling and often

[4] Sector modelling for AFOLU will include areas of land with no economic activity but the treatment of the sector must consider the GHG emissions from these areas and the potential for them to be converted to some form of economic activity.

presents the greatest range in terms of modelling sophistication. At the most sophisticated end of the spectrum, are the process systems engineering models that follow the steps involved in production and distribution of a product (or their biological equivalent). However, such approaches are typically too intensive to include in the generation of climate scenarios so most models will adopt some form of simpler parameterisation. In its simplest form this might be a bulk parameterisation that represents the activity with a single equation. However, more commonly the activity will need be represented by a complex pyramid of parameterisations that bring together many layers of physical and biological processes. The types of models available are sector specific so no we do not propose any common classification.

Climatic Feedbacks: The analysis should list the physical transmission channels impacting productivity and costs. For example, there are the losses due to acute weather events such as flood, drought and storm damage. However, the analysis might also consider chronic impacts on the cost of living such as insurance premia and water supply costs. The evaluation should consider how well the identified transmission channels are modelled as well as the implications of the omissions.

Mitigation Policy Representation: Although mitigation policy and its parametrisation may have been specified in level 1 of the taxonomy the classification should capture any sector specific considerations. This might take the form of sub-sectoral and geographical exclusions to the global policy pathway.

Substitution Dynamics: A simple approach to substitution might allow substitution to occur based on simple economics. However, a more realistic treatment needs to consider the frictional effects created by behavioural inertia, transformation costs and the time lag to build substitute capacity. Possible values include:

- Maximum percentage per annum land use transformation
- Maximum absolute p.a. increase in certain renewable technology energy sources
- Not applicable – simpler representations of a sector might assume that economic activity and GHG emissions scale in line with macroeconomic factors without any substitution logic.

Market Dynamics: The treatment of market dynamics is closely linked to the model architecture. Where a sector is modelled separately the market will often be represented by a partial equilibrium model that only captures the individual sector under consideration in detail. The demand relationship with other sectors must then be handled through exogenously defined parameters. However, the sector-based logic might be incorporated in the logic of a general equilibrium model that covers all sectors of the economy. Alternatively, a non-equilibrium approach might be taken using an agent-based approach that models the behaviour of all agents at each time step. Where such whole market approaches are adopted the market dynamics in the sectoral models will be the same as in the macroeconomic model. Another example of nonequilibrium approaches involves the use of macroeconometric models. Possible values include:

- Partial Equilibrium
- General Equilibrium
- Non-Equilibrium - Agent Based
- Non-Equilibrium – Macroeconometric

In addition to this categorisation, models like integrated assessment models (IAMs) can also be classified as either process-based or cost-benefit models (Weyant, 2017). As the name implies, process-based IAMs (e.g. MESSAGEix-GLOBIOM) represent processes that transform raw materials into consumption goods and services with technological detail that includes conversion efficiencies, load factors, and capacity to use the energy sector as an example. In contrast, cost-benefit IAMs like DICE have aggregated representations of the economy and environment but represent their interlinkages with feedbacks (e.g. climate damages). The variables in process-based models have physical units like MWh, GW, tons per hectare, etc., while cost-benefit models monetize all quantities.

Macroeconomic models (CGEs, macroeconometric) also use monetized quantities in which all variables are converted to a currency value. Such models, especially CGEs, can have high inertia and structural change may only come at very high carbon prices or costs of mitigation. For that reason, they are often hybridised by introducing a process-based module for the energy system embedded in a standard macroeconomic setting (e.g. AIM-CGE, E3ME-FTT) to better represent transition dynamics.

Crucially, these models behave very differently from a model structure point of view, with important consequences for the results. For example, equilibrium models lead to GDP losses when a constraint on emissions is introduced, while non-equilibrium models may show the opposite effect (Köberle et al., 2021).

2.2.5.4 Energy Sector Models

The energy sector is responsible for nearly three quarters of all global greenhouse gas emissions as shown earlier in figure 10. Therefore, the energy model is at the heart of any economic climate scenario, responding to policy input, impacting the climate system through emissions and influencing all sectors of the economy through the different vectors available for delivering energy and the associated prices. Identifying how the energy system is represented in the overall structure of an economic climate modelling chain is a key step in understanding the level of sophistication in the modelling approach, identifying its strengths and weaknesses, and ultimately positioning the outputs in the range of potential outcomes.

At the less sophisticated end of the spectrum the energy model might be somewhat disconnected, simply delivering a relationship between mitigation policy (potentially represented by carbon price curves) and energy emissions, without representing the components making up the energy system and how they evolve. This is the case with all economic pathways derived from the output of a macroeconomic model without sectoral representation such as NiGEM or DICE, with the climate pathway set by a combination of carbon price and a simple representation of physical climate change such as regional temperature. Even with this approach to the macroeconomics, intermediate output (e.g. carbon price, sectoral emissions, bioenergy costs) from the energy model and other sector specific modelling (such as AFOLU as described in the next section) should be used in the expansion modelling that delivers financial asset price pathways, if the scenario is to be coherent.

More sophisticated modelling approaches, that build up the macroeconomic pathways using a sectoral representation of the economy (e.g. E3ME, MESSAGE, REMIND), require the energy model (and the modelling of all other sectors that are individually captured) to be more integrated with the macroeconomic model to the extent that it may be difficult to separate the two. However, the underlying assumptions governing the evolution of the energy sector should still be identifiable and reviewed following the structure of the taxonomy.

These more sophisticated energy models should be considered in two parts, the supply side and the demand side. On the supply side the model must consider the cost to produce and deliver the energy to the final point of consumption, and this can be represented in increasing granularity for the processes governing energy system composition, depending on the model. The basic steps involved are illustrated in Figure 12 below:

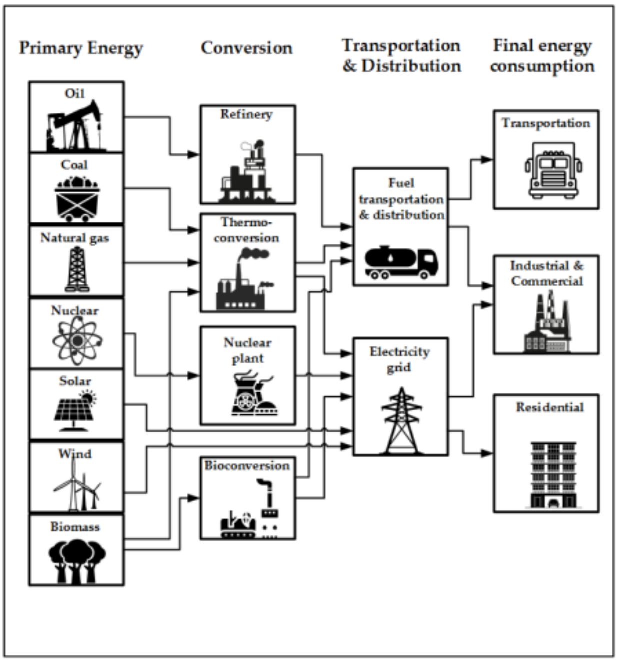


Figure 12: Energy system showing the flow of energy from primary energy to final energy consumption Reproduced from Subramaniam et al 2018

To develop end consumer cost curves the supply side of the modelling chain should consider:

- The current cost of production, conversion, storage, transport and distribution by source.
- The conversion efficiency of technological options available to the model.
- The extent of available reserves and how costs vary with production volume.
- The evolution of these costs with technological progress.
- The impact on costs and renewable capacity of climate change.
- The impact of mitigation policies through taxes and tariffs.
- The introduction of new forms of final energy consumption.
- The impact of emission mitigation policies and cross-border tariffs.
- The role of energy vectors in delivering energy to the various sectors.
- The time and upfront cost required to build new energy capacity.

Inevitably the level of detail with which the supply chain is represented in an economic energy model will be significantly below that found in a process systems engineering model used within the energy sector. In energy systems models, details of the engineering challenges will typically be parameterised through conversion costs or efficiencies, load factors and technology learning curves. As a result, the mathematics in this part of the process are often relatively straight forward and it is the calibration of the inputs that constitute the key scenario design choices. At the most granular end of the energy modelling spectrum are the models at the firm level, representing individual energy and power companies. With this structure the modelling of the energy sector is driven by the rules that govern the behaviour of the individual firms. The sophistication of these rules will, in turn, determine the inputs required. The firms can be represented as agents in agent-based models, or as processes that compete on cost in optimization models.

Energy system models must also include a module that calculates demand for energy services and compares this with the supply cost dynamics. Full economy models can represent market dynamics to project energy prices and production volumes based on assumptions governing the supply-demand balance (equilibrium or non-equilibrium). This might be a separate economic module that is part of a stand-alone energy model, be integrated in the macroeconomic model or result from the behaviour of the agents in an agent -based approach. The structure of the supply and demand side dynamics will determine how the many feedback mechanisms are captured. For example, in the analysis of a scenario we should explain the extent to which physical climate change is considered when assessing energy production capacities and costs or changes in demand. Or describe closure assumptions such as the requirement that investments equal savings that may determine macroeconomic flexibility of the modelling framework.

Taxonomy application example

The NGFS scenarios provide the most readily accessible examples of the application of energy sectoral models in the construction of climate scenarios for financial users. We have selected the implementation of the MESSAGEix model in the latest generation of the NGFS Net Zero 2050 MESSAGEix- GLOBIOM 1.1 scenario to illustrate the detailed application of the taxonomy. However, the documentation does not provide full detail and the output variables available through the NGFS portal are limited. The code is publicly available to review and run and there is extensive documentation available online[5], but this level of analysis is beyond the scope of this paper:

[5] https://docs.messageix.org/en/latest/index.html

Model Component Classification

Scenario	Component ID			Scale and resolution				Model type and complexity				Interfaces			Component specific features			
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Climatic feedbacks	Mitigation policy representation	Substitution dynamics	Market dynamics
NGFS MESSAGEix-GLOBIOM 1.1M-R12 Net Zero 2050	2b	MESSAGEix	Energy sector model – IAM	Global model without geographic climate feedbacks	Energy demand is calculated for 7 energy services across 11 regions	5 energy levels with 4 types of raw resource, y forms of primary energy, 54 conversion technologies, z forms of secondary energy, a forms of final energy and z forms of useful energy.	5 years to 2060 and 10 years beyond	Linear programming energy engineering model	Deterministic	Empirical	None	Demographics: SSP2; Farming yields: FOA Stat; Elasticities: USDA	Emissions, land use, market demands and prices, trade patterns	Exogenous input	Energy related emissions feed the MAGICC climate model but apparently without feedbacks	Policies can be represented as a combination of emission targets, energy shares, capacity or generation targets and macroeconomic targets (taxes and subsidies)	Technology diffusion dynamic constraints	Partial equilibrium linked to MACRO General Equilibrium Model

Table 10: Energy Sector Model Component Classification

Scenario: NGFS Net Zero 2050 MESSAGEix-GLOBIOM 1.1 M-R12 Net Zero 2050

Energy Model: MESSAGEix

Detailed Model Description: The following description of the modelling approach is taken from the MESSAGEix documentation:

> "MESSAGEix provides a framework for representing an energy system with all its interdependencies from resource extraction, imports and exports, conversion, transport, and distribution, to the provision of energy end-use services such as light, space conditioning, industrial production processes, and transportation. The model is designed to formulate and evaluate alternative energy supply strategies consonant with the user-defined constraints such as limits on new investment, fuel availability and trade, environmental regulations and policies as well as diffusion rates of new technologies. Environmental aspects can be analysed by accounting for, and, if necessary, limiting the amounts of pollutants emitted by various technologies at various steps in energy supplies. This helps to evaluate the impact of environmental regulations on energy system development. Because few conversion technologies convert resources directly into useful energy, the energy system in MESSAGEix is divided into 5 energy levels:

- Resources: raw resources (e.g., coal, oil, natural gas in the ground or biomass on the field)
- Primary energy: raw product at a generation site (e.g., crude oil input to the refinery)
- Secondary energy: finalized product at a generation site (e.g., gasoline or diesel fuel output from the refinery)
- Final energy: finalized product at its consumption point (e.g., gasoline in the tank of a car or electricity leaving a socket)
- Useful energy: finalized product satisfying demand for services (e.g., heating, lighting or moving people)

Energy technologies are characterized by numerical model inputs describing their economic (e.g., investment costs, fixed and variable operation and maintenance costs), technical (e.g., conversion efficiencies), ecological (e.g., GHG and air pollutant emissions), and socio-political characteristics. An example for the socio-political situation in a world region would be the decision by a country or world region to ban certain types of technologies (e.g., nuclear power plants). Model input data reflecting this situation would be constraining the use of these technologies or, equivalently, their omission from the data set for this region altogether."

Energy service demands are provided for the following seven sectors:

1. Residential/commercial thermal
2. Residential/commercial specific
3. Industrial thermal
4. Industrial specific
5. Industrial feedstock (non-energy)
6. Transportation
7. Non-commercial biomass.

Baseline demand can be set exogenously or determined endogenously by MACRO. Demand is a log linear function of population and GDP per capita.

As discussed in the introduction the energy model output is therefore dependent on a wide range of parameters calibrated at the base years for the many layers in the model structure. To understand the key calibration choices, we must identify the parameters for which there is uncertainty as to their value which translates into a material variation in the outputs that are material to the end user. A full analysis of all energy inputs to the MESSAGEix model is beyond the scope of this paper. However, calibration of historical years (typically to 2020) is accomplished by constraining the model to values in the IEA and Platts databases, which guide the accuracy with which the initial system is represented. The key calibration parameters are those that govern the evolution of the energy system in terms of supply and demand, and it is important to determine their starting point. In most cases the calibration values are not readily available preventing comparison with other scenarios. For the NGFS implementation of MESSAGEix, we have identified the following key inputs based on a review of the NGFS and IIASA documentation:

- Target Emissions Pathways: Scenario dependent, based on MAGICC[6] outputs.
- Population demographics: SSP2 provides population growth but other demographics can be used by MESSAGEix[7].
- Final Energy Intensity and Energy Service Demands: Calibrated for base years from exogenous sources but method is unclear from the documentation accessed.
- Technological progress: Calibrated based on the SSP2 but the method is not described the in NGFS documentation.
- Resource limits: Exogenously defined but sources are unclear[8].
- Autonomous energy efficiency improvement coefficients: source unclear.

Limitations, **Assumptions & Comparison with Alternative Models:** Example limitations include the following, but the analysis has not been performed with sufficient rigour to identify all major limitations:

- The documentation does not cover any impact on energy production cost and capacity limits due to physical climate change. The resultant scenarios would therefore not be appropriate when considering the impact on weather sensitive energy projects such as wind farms or exposed offshore oil fields. In the case of the net zero by 2050 scenario physical climate change is less pronounced
- Longer-term projections for fossil fuel use are higher than in similar scenarios with a potential impact on fossil fuel asset prices prior to 2050. The scenario is therefore not suitable for identifying long-term strategic vulnerabilities in financial business models with high fossil fuel exposures for chronic risks, but acute risks can be very material in the short term.

The NGFS explorer provides a good range of information allowing comparison of energy use between the three NGFS IAMs for the Net Zero by 2050 scenario as illustrated by figures 13 to 16. MESSGAGE-GLOBIOM projects the highest final energy use, but that this is not replicated in primary energy use highlighting the role of conversion efficiency, particularly up to 2060. Interestingly the additional final energy use is not in the form of electricity. Much of the gap after 2050 is delivered via fossil fuels of which just over half is subject to CCS. By looking at the emissions projections we can see that this additional use of unmitigated fossil fuels is allowed by much more optimistic levels of CO_2 drawdown in the AFOLU sector post 2050. Although the most material variances in this analysis are after 2050, and so not directly relevant to most financial use cases, they illustrate the type of analysis possible by looking at more detailed intermediate level outputs. A more detailed understanding of the underpinning models and their calibration would allow us to identify the sources of these variations and opine on the likelihood of the different pathways. For example:

[6] MAGICC is a simple climate model that runs alongside many integrated assessment models. But different IAMs use other simple climate models, such as FAIR or HECTOR.
[7] Demographics are exogenous drivers of future developments. Default implementation in MESSAGEix is SSP2 but this can be changed to any available demographic projection such as other SSPs or UN for example.
[8] Typically, they can include IEA, IRENA, USGS, BP Statistics

- Is MESSAGE-GLOBIOM at the optimistic end of the plausible range in terms of energy conversion efficiency?
- Are the other two IAMs overly optimistic about the future energy intensity of global GDP?
- Is the GLOBIOM long-term projection for continued CO_2 drawdown by the AFOLU sector realistic?

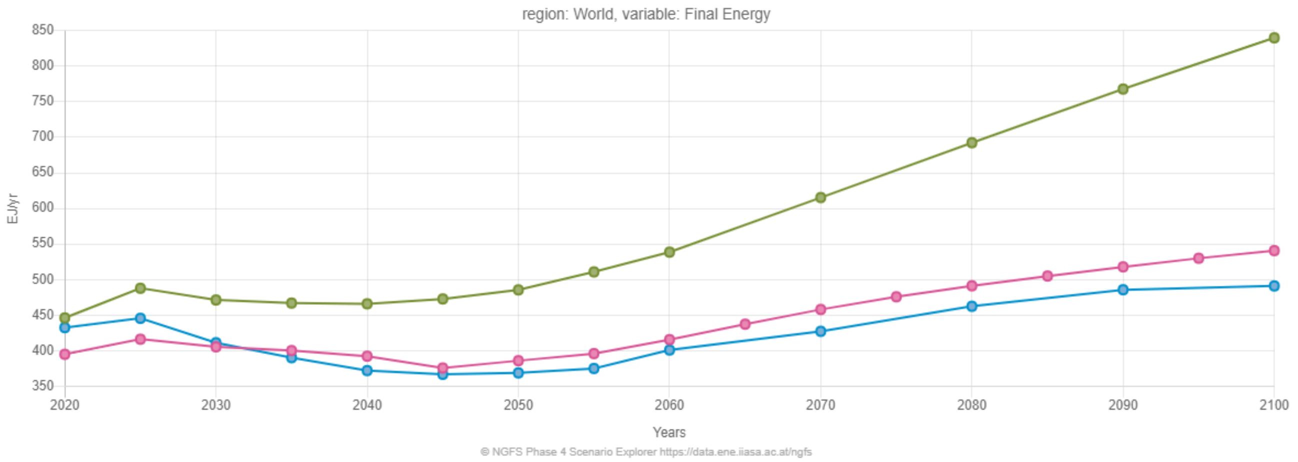


Figure 13 – NGFS Explorer output showing Final Energy Use for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

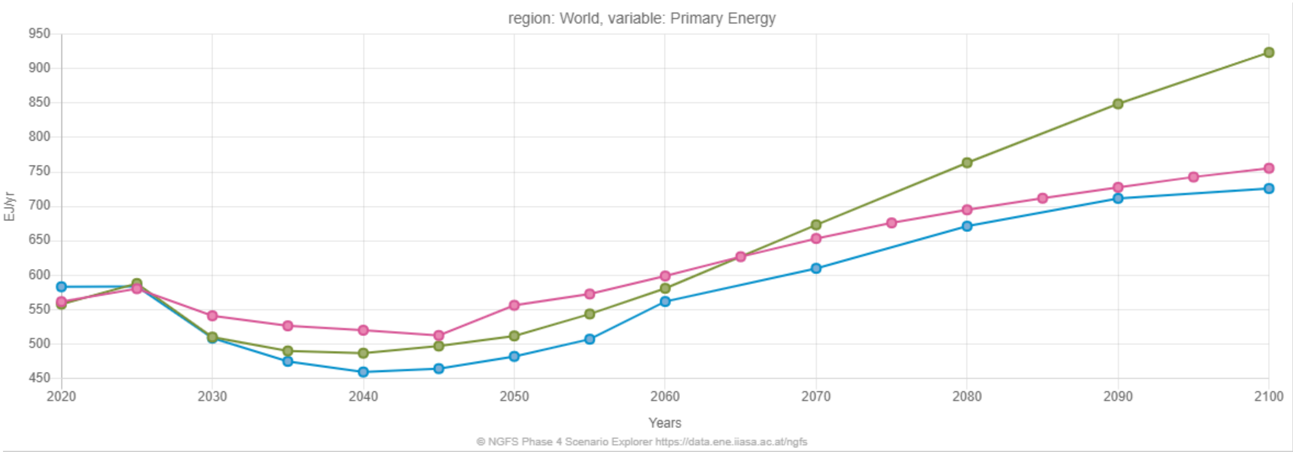


Figure 14 – NGFS Explorer output showing Primary Energy Use for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

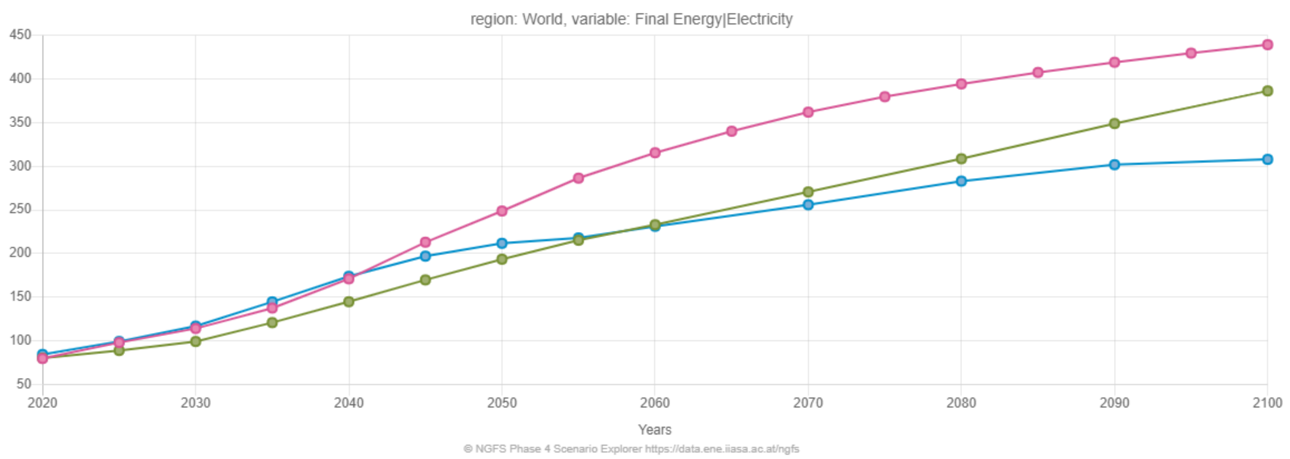


Figure 15 – NGFS Explorer output showing Final Energy Use (Electricity) for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

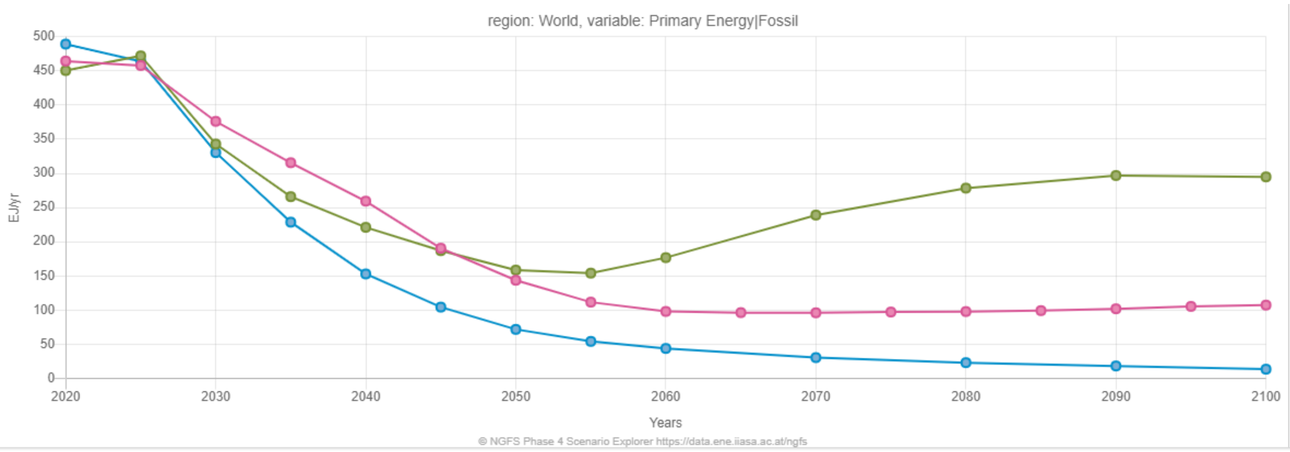


Figure 16 – NGFS Explorer output showing Primary Energy (Fossil Fuels) for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

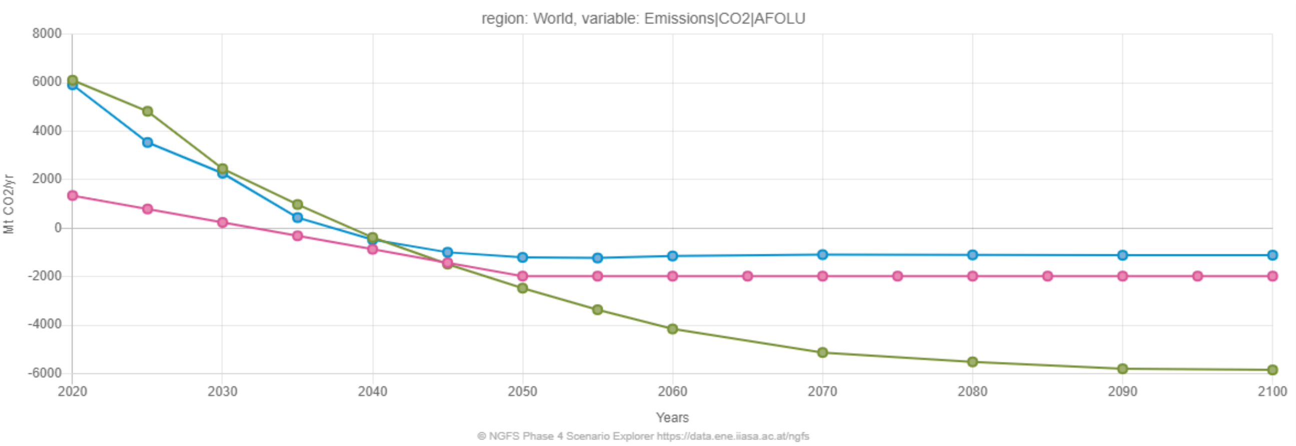


Figure 17 – NGFS Explorer output showing AFOLU sector CO_2 emissions for the Net Zero 2050 scenario run on the three IAMs: REMIND-MAgPIE (blue), MESSAGE-GLOBIOM (green) and GCAM (pink)

Summary & recommended use: MESSAGEix provides a good benchmark for the modelling of the energy sector, being well recognised and employed by many organisations, including the NGFS and the IPCC. The structure and calibration of individual sub-components has been developed over decades and a significant body of literature exists assessing the results.

Despite this, further research is required to better understand the uncertainty in the calibration parameters and the sensitivity of the final output to these calibration choices, resolution and model dynamics. It is tempting to assume that the energy model that underpins MESSAGEix is well established with relatively low levels of uncertainty associated with its output, although ideally this would be quantified in the model documentation. From the viewpoint of a financial end user, it is likely that the key sources of uncertainty will come from the implementation of mitigation policy frameworks, technology learning curves, infrastructure lock-in and the assumptions around demand elasticities, but these assumptions need to be substantiated. However, these are also subject to scenario design, and a single model like MESSAGEix can produce many scenarios that meet net-zero emissions by 2050 but with different underlying energy system dynamics.

Comparison of intermediate outputs between different scenarios produced using MESSAGEix and alternative models such as REMIND will be valuable in understanding the uncertainty in the output parameters relevant to specific financial use cases. The most significant intermediate outputs include the total energy use, energy mix and energy prices and the associated GHG emissions, but others can play important roles as well.

2.2.5.5 Agriculture Forestry and Other Land Use (AFOLU)

The AFOLU sector is the second largest source of GHG emissions after energy, accounting for ca. 20% of global emissions if the intersection with energy is included (see figure 18) and is closely linked with the physical climate system. It is also a sector with enormous mitigation potential (Roe et al 2021). However, this potential is difficult to realise, even at the lower levels depicted in many currently available scenarios. This difficulty is partly due to the high number of small agricultural and forestry activities, and their concentration in areas of the world where dramatic policy intervention is less likely. Further, some of the mitigation potential is associated with changes in consumer behaviour (e.g. a shift to less meat-based diets in the developed economies) which will be difficult to achieve. Therefore, modelling simplifications, such as a globally consistent carbon tax, could deliver mitigation results for the sector at the optimistic end of the spectrum. The assessment should consider how the assumptions impact the overall scenario (see figure 17 in the energy sector analysis).

Modelling the AFOLU sector often involves a further sub-chain of models, as illustrated below based on the approach in GLOBIOM. The structure is like the generic sector specific modelling schematic introduced earlier:

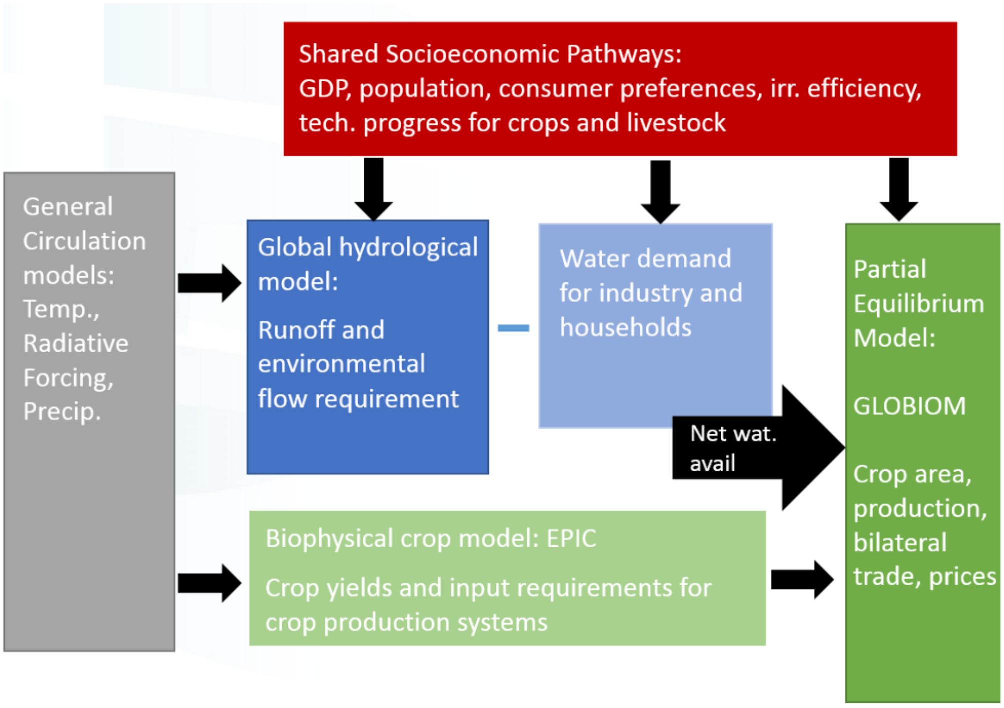


Figure 18: Overview of the GLOBIOM model taken from the model documentation:
(International Institute for Applied Systems Analysis (IBF-IIASA). , 2023)

A combination of the mitigation provided by the AFOLU sector and negative emission technologies, such as direct air capture, significantly impact the relationship between the speed of transition of the energy system and the physical climate pathway. An understanding of the GHG emissions pathway associated with the AFOLU sector is therefore a vital input to the scenario selection and interpretation process for all users. The overly optimistic treatment of emission sinks, both natural and technological, has the potential to underplay the risks faced at both a global and financial institution level. For users with a particular interest in the agricultural and forestry sectors the more detailed modelling of crop yields and profitability will also be of major interest.

Taxonomy Application Example

As discussed earlier the NGFS scenarios provide the most readily accessible examples of the application of AFOLU sectoral models in the construction of climate scenarios. We have selected the implementation of the GLOBIOM model in the latest generation of the NGFS Net Zero 2050 MESSAGEix- GLOBIOM 1.1 scenario to illustrate the detailed application of the taxonomy. The code is publicly available to review and run, with extensive documentation available online[9], but detailed analysis is beyond the scope of this paper.

[9] https://globiom.org/

Model Component Classification

Scenario	Component ID			Scale and resolution				Model type and complexity				Interfaces			Component specific credibility			
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Climatic feedbacks	Mitigation policy representation	Substitution dynamics	Market dynamics
NGFS Messageix-GLOBIOM 1.1M-R12 Net Zero 2050	2b	GLOBIOM 1.1	AFOLU sector model – IAM	5 to 30 arcmin pixels with common characteristics within national borders giving >10,000 units worldwide	Industry sub-sectors aggregated to 37 regions	9 land cover types, 18 crop type globally)27 in the EU) with 4 management systems, 4 livestock species aggregates and 8 production systems. Forestry broken down by biomass, 5 primary wood, 8	5 years to 2060 and 10 years beyond	Bottom up spatially explicit land use model	Deterministic	Empirical	None	SSP2 population demographics, USDA demand elasticities, FAO stat yields and emission curves	Emissions, land use, market demands and prices	Two-way with the climate mode, others unknown	Precipitation, temperature	Global carbon price	Land use transformation with annual limits	Partial equilibrium

Table 11: AFOLU Sector Model Component Classification

Scenario: NGFS Net Zero 2050 MESSAGEix-GLOBIOM 1.1 M-R12

AFOLU Model: GLOBIOM Version 1.1

Detailed Model Description: The GLOBIOM 1.1 documentation provides the following high-level description of the modelling approach:

"The Global Biosphere Management Model (GLOBIOM) is a partial equilibrium model representing main land use sectors, including agriculture, forestry and biofuels. The supply side of the model is built from the bottom (spatially explicit land cover, land use, management systems and economic cost information) to the top (regional commodity markets). Demand is determined at the level of aggregate regions. Food demand projections are based on the interaction of three different drivers: population growth, per-capita income growth, and response to prices (based on consumer preferences), and policies.
The spatial resolution of the supply side relies on the concept of Simulation Units that are aggregates of 5 to 30 arcmin pixels belonging to the same altitude, slope, soil class, and following country borders. For crops, livestock, and forest products, spatially explicit Leontief production functions covering alternative production systems are parameterized using biophysical models like Environmental Policy Integrated Climate Model (EPIC) or Global Forest Model (G4M)."

To understand the model dynamics of GLOBIOM 1.1 we must look at the dynamics of the underlying sub-models:

Biophysical Processes: Biological productivity, the associated GHG emissions, and its response to management techniques and climate change are handled by the EPIC, RUMINANT, Bioenergy and G4M sub-models as presented in the schematic below.

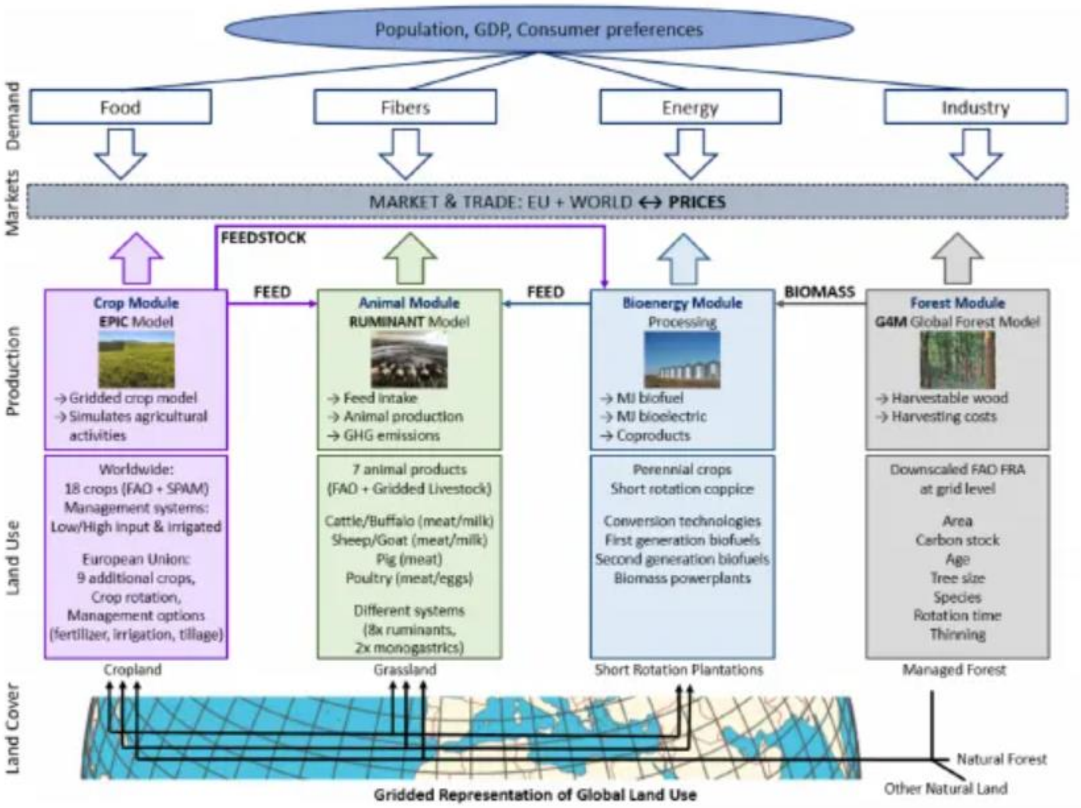


Figure 19: Sub-Model Schematic from GLOBIOM documentation

The sophistication of the dynamics and the resolution with which they are applied determines the overall model's sensitivity to climate forcing and its response to mitigation policies in terms of changes in land management techniques and land use transformation. Satellite data is used to determine current land use at a spatial resolution of between 5 and 30 arcmin creating units with similar physical characteristics in terms of altitude, slope and soil type. The biophysical models then project biological productivity data considering:

- Climate change in terms of temperature and precipitation
- Potential changes in land management
- Potential land use transformations

The productivity results then feed the market dynamics module, within the main GLOBIOM model, to determine the scenario trajectory in terms of land use, land management approach and costs, bioenergy availability and net GHG emissions. Importantly, the net GHG emissions represent the net of sources and sinks for CO_2, and of the emissions of non-CO_2 gases, especially of CH_4 and N_2O which the agricultural sector emits a high share of. If we assume that the biological productivity is well established, the key model dynamics to understand are those that govern future changes in emissions intensity, production costs and profitability that drive land use change. It is important to understand the uncertainty around the calibration choices.

Climate Feedbacks and Mitigation Options: Climate change is captured in terms of precipitation and temperature change. In response the model allows a change in management approach or transformation of land use. Structural management approach options are based on the ability to switch to more GHG efficient practices in terms of crop rotation, irrigation and the use of fertilisers. There are also a set of technological mitigation options to reduce emissions including anaerobic digesters and feed supplements based on the Environmental Protection Agency database. The best economic option is chosen allowing for GHG mitigation policies and land use transformation costs and restrictions. GLOBIOM assumes homogeneous goods with a single market price allowing for transportation, trade tariffs and administration costs. Combining these with the basic carbon tax functionality and land protection and restoration intervention schemes provides the ability to model a range of mitigation policies although the exact calibration in the NGFS scenarios is unclear.

Market Dynamics: GLOBIOM is based around a partial equilibrium market model that maximises the sum of consumer and producer surpluses subject to constraints. Supply side resolution is driven by the combinations of land use and management approach evaluated by the biophysical models. However, interaction with other sectors of the economy is via exogenous inputs passed from other modules of the wider MESSAGE-GLOBIOM integrated assessment model. These exogenous inputs include demand for biofuels and prices of pesticides. The demand is calculated at a regional level (37 regions) driven by population demographics and behavioural change from SSP2. Income and price elasticities are taken from the USDA database with changes in elasticities incorporating behavioural evolution.

Input Calibration: A full analysis of all inputs to the GLOBIOM model is beyond the scope of this paper. However, calibration of historical years (e.g. to 2020) is accomplished by constraining the model to values in the FAOSTAT databases, which guide the accuracy with which the initial system is represented. The key calibration parameters are therefore those that govern the evolution of the AFOLU system in terms of supply and demand, and it is important to determine their starting point. The calibration of elasticities is identified as a source of uncertainty in the GLOBIOM documentation. In most cases actual calibration values are not available, preventing comparison with other scenarios. For the NGFS implementation of GLOBIOM, we have identified the following key inputs based on a review of the NGFS and IIASA documentation:

- Target Emissions Pathways: Scenario dependent, based on MAGICC[10] output
- Population demographics: SSP2 provides the population growth trajectory but other demographics can be used by GLOBIOM.
- GDP projections: from SSP2 projections (Dellink et al., 2017).
- Technological progress: Based on SSP2 assumptions.

---

[10] MAGICC is a simple climate model that runs alongside many integrated assessment models. But different IAMs use other simple climate models, such as FAIR or HECTOR.

- Management System Inputs and Prices: Water, nitrogen and phosphorous requirements and initial prices taken from FAOSTAT.
- Demand Elasticities – Based on USDA data.
- Bioenergy demand: Exogenously determined by energy model (MESSAGEix)

Limitations, **Assumptions & Comparison with Alternative Models:** A full review of the limitations and assumptions of GLOBIOM is beyond the scope of this paper. However, the interaction with the physical climate system provides an illustration of a limitation linked to the structure of the model. The temporal resolution of the GLOBIOM model is low. Although configurable, it is typically several years, for the NGFS scenarios 5 years to 2060 and 10 years beyond that. Therefore, changes in biological productivity and the cost of adaptation are based on chronic climate pathways, omitting the impact of acute weather events. For the net zero 2050 scenario, where climate change is limited, this might be a lesser consideration. However, as plausible scenarios are forced to include more significant levels of climate change this will become increasingly important. The significance to individual financial end users will vary according to their risk profiles. However, in some regions, acute events, such as flooding, storm damage and wildfire, could increase the cost of agricultural production, or even bring production to an end with consequent impacts on the global market.

The NGFS explorer provides the ability to compare a limited selection of GLOBIOM outputs from the NGFS scenarios with those from GCAM and REMIND-MAgPIE. As shown in figure 17 earlier, GLOBIOM assumes a much greater long-term drawdown of CO_2 than the other NGFS IAMs running the same scenario. Figure 20 shows a slightly different problem. Clearly the food demand projected by GLOBIOM is much lower than that by MAgPIE, but the difference arises from the starting point questioning the reliability and comparability of the data.

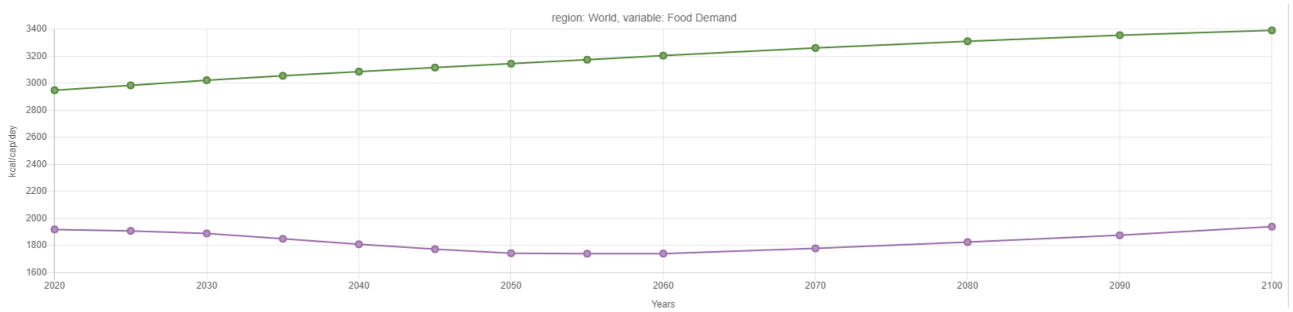


Figure 20 – NGFS Explorer output showing Global Food Demand for the Net Zero 2050 scenario run on two IAMs: REMIND-MAgPiE (green), MESSAGE-GLOBIOM (purple)

Summary & Recommended Use: GLOBIOM provides a good benchmark for the modelling of the AFOLU sector, being well recognised and employed by many organisations, including the NGFS and the IPCC. However, the range of sub-models and the extent of the calibration data involved in the GLOBIOM model are particularly daunting but are necessary to project a complex system. The calibration of individual sub-components has been developed over decades and a significant body of literature exists assessing the results.

Despite this, further research is required to better understand the uncertainty in the calibration parameters and the sensitivity of the final output to these calibration choices, resolution and model dynamics. It is tempting to assume that the biophysical models that underpin GLOBIOM are well established with relatively low levels of uncertainty associated with their output, although ideally this would be quantified in the model documentation. From the viewpoint of a financial end user, it is likely that the key sources of uncertainty will come from the implementation of mitigation policy frameworks and the assumptions around demand elasticities. These assumptions need to be substantiated and are also subject to scenario design. A single model like GLOBIOM can produce many scenarios that meet net-zero emissions by 2050, when combined with various treatments of the other emissions producing sectors.

Comparison of intermediate outputs between different scenarios produced using GLOBIOM and alternative models such as MAgPIE will be valuable in understanding the uncertainty in the output parameters relevant to specific financial use cases. The most significant intermediate outputs are the GHG emissions associated with the AFOLU sub-sectors, but these are not available through the NGFS Scenario Explorer. To better understand the drivers of the variability in emissions projections and understand the economic impacts at a sub-sector level we must delve deeper into the modelling output. We might compare how different scenarios, and alternative models affect land use change, productivity responses to various manifestations of climate change.

2.2.6 Level 3 – Microeconomic Expansion Models

The macro models provide high level economic and climate pathways that must be translated into individual financial asset or liability impacts for use in scenario analysis by the financial sector. The models used in this final step in the construction of a scenario are commonly referred to as microeconomic expansion or risk transmission models. However, where the modelling architecture is built around the most granular agent-based models the process flow may be reversed with the macroeconomic pathways created by aggregating the underlying elements of the economy. In such cases the user of the taxonomy should balance which components of the modelling should be discussed in the level 2b macroeconomic section and which in level 3 to provide the most insightful analysis.

Where expansion is required, this might be performed in whole or in part by the scenario originator, by an intermediary such as a financial regulator or by the end financial user. Level 3a is used to describe the expansion performed by an intermediary and 3b by the financial end user. Although level 3a will not exist where a financial institution builds directly on the output of a model vendor, it is a common step in the development of regulatory scenarios affecting both the regulatory exercise and the on-going use of the resultant scenarios by financial institutions for internal risk management.

2.2.6.1 Challenges in the Design of Microeconomic Expansion Models for Climate Risk

The available microeconomic models designed to assess the impacts of climate physical and transition risks exhibit several features and assumptions that have been observed and assessed to create a wide range of heterogeneous impacts to financial factors such as asset prices (Federal Reserve Board 2024). Assumptions at the macroeconomic level around growth, and related factors such as productivity, applied to ensure exogenous growth may lead to unrealistic (ie. probabilistically infeasible paths) outcomes in terms of the ability to fund decarbonisation for example (McKinsey & Company 2022). Such macroeconomic modelling choices can create further unrealistic outcomes when the resultant macroeconomic pathways are subsequently downscaled to the firm or population level. Such constraints that are path dependent are important to consider in the context of assigning probabilities to such pathways.

For example, in the TRISK model (Baer et al. 2022) modelled firms' capital formation and risk management rules are not explicitly captured. Instead, demand forecasts directly from the input NGFS scenarios drive the capital formation rules. However, models such as Cormack et al (2020) use the macroeconomic scenarios from sources such as the NGFS as expectation forecasts for the modelled firms, but not as prescriptive paths that must be followed. The idiosyncratic capital formation and firm-level risk management rules drive the reasons for such constraints within the Cormack et al (2020) model. Other models that employ agent-based methods, such as (Battiston et al 2023, Dunz et al (2022)), use trajectories from the set of NGFS scenarios (or an equivalent set of scenarios that provides a prediction of the demand and capital costs). This type of framework models the capital formation process from debt issuance with the ability to integrate capital formation rules for firms and with the option to assume that firms maintain constant market shares. Other alternative models outside of the IAM scenario set, such as the INET technology driven concept in Way et al (2022), use a model based purely on the total capital cost of deploying new technology based on empirical motivated estimates of capital costs. These scenarios were used as inputs to the TRISK model described above. The INET model combined with TRISK does not form a complete agent-based model of the energy system as firms are not directly modelled as part of the capital formation process. Whilst it is anticipated that firms would benefit from reduced capital general costs as forecast, further supply side constraints and regional access to technologies in the supply chain would need to be investigated and included. Such factors will inevitably impact costs to firms / consumers as has been recently seen with tariffs imposed on recent imports from China applied by the United States and the EU.

Such discrepancies between the macro-economic scenario and *subsequent* downscaling are currently a direct consequence of the lack of coherence intrinsic in modelling the system's micro-economic agents. Such incoherence manifests itself across several other features that can give rise to the wide disparity in the risk numbers seen in models used to assess risks. As highlighted in Cormack and Shrimali (2023) the challenge of building a valid risk transmission framework for climate risk (and other externalities) requires that the system being modelled reliably reflects the drivers (e.g. demand), economics, price cost, capital formation mechanisms and, importantly the risk management (mitigation) choices of the system being investigated. Without a dependable model of the 'real world' economy (system simulation framework) that provides a clear transmission channel, there is a concern that this leads to an unrealistic assessment of asset values and default risks. When exploring the evolution of models and their approach to capturing the dynamics of climate related financial risks, it is worth looking at the model of a corporate entity, its internal business strategy, risk management, and how it engages in raising finance. In Cormack and Shrimali (2023) the attributes of a firm and its operational characteristics are described. The following characteristics may be regarded as the canonical factors for modelling firms:

- Demand forecasting
- Cost estimation and forecasting, covering materials and energy input.
- Price discovery - reflecting the model for the interplay of competition and cost factors.
- Strategic planning based on demand/price and market strength.
- Internal Strategy and Risk management - across all aspects of a firm's operations, covering:
    - Profit targets - requiring the management of business margins.
    - Investment management - the ability to control the amount of long-term debt to ensure expected earnings are in line with strategy and cover it.
    - Controls on interest rate costs
    - Climate risk management would include using insurance and enhancing a firm's facilities to reduce its vulnerability to climate change (i.e., adaptation expenditure).
    - Hedging strategies across all aspects of the business, from financial market exposure, FX, interest rates, and input commodities to, in some cases, agreeing

- on contracts for sales prices and insurance as an economic hedge against climate risks.
- Control and optimisation of its capital structure.

Within the literature, some models consider several of these aspects, whilst a full review of such features requires a richer analysis and availability of model documentation, especially for commercial frameworks. We can compare how the factors above are captured using macro-economic demand models such as TRISK (where firms demand predictions will following its current relative market share as it exists at the start of the simulation) with models such as Cormack et al (2020), (where firms demand predictions will be linked to the model of the firm's strategy and investment):

Demand Forecasting
Within the current modelling paradigm, IAMs / GVA pathways are used at the macroeconomic level to define outcomes for firms. That is, firms are forced a-priori to follow such trajectories, whilst the principal of such forcing is to create consistency from the macro to micro; such forcing may imply breaking a firm's desired management criteria as highlighted above. It is the case that a firm may not be able to manage the transition effectively, especially if policies are introduced to accelerate transitions. In models such as TRISK, firms engage in capital allocation matching the demand profile of the technology demand prediction without constraints, whereas in Cormack et al (2020) firms will engage in an investment based on an internal assessment of costs using the input demand as a forecast rather than a proscribed path to follow.

Cost and Price Formation
In the paradigm of downscaling prices and costs for firms, the input costs are driven through the IAM for models such as TRISK. Whilst this may be appropriate for market-driven values, this again creates a challenge as price setting for firms is driven through micro competition in the real world, which reflects firms' access to suppliers, regional taxation, and capital costs. Models such as Cormack et al (2020) permit firm price settings to reflect changing product and financial market conditions.

Firm Level Strategy
Within models such as TRISK, a firm has a static strategy whereby its allocation to current and future technologies is constrained to its market share, the same follows for Battison et al (2023) and Dunz at al (2022). Within Cormack et al (2020) firms make their own choices on capital allocation using the external demand forecasts based on internal profit margins, factoring forward-looking views of prices, capital cost, and taxation. Within Cormack et al (2020), it is also possible for firms to define their investment strategies to allocate capital to current and new revenue operations, hence providing a means to assess performance in a communicated transition plan.

Internal Strategy and Risk Management
Within models such as Cormack et al (2020) the modeller can define a set of operating targets that can control internal risk management objectives, such as controls on profit margins in price setting, the amount of debt that can be raised based on expected earnings, and controls

on interest costs. Such control factors are designed to emulate real-world firm behaviours to ensure that the firm does not raise excessive debt, degrade its funding, or control costs. Modelling such factors enables more meaningful estimates of growth, internal costs, and revenue predictions and excludes forced behaviours that are not qualitatively consistent with how the firms currently operate. A model that captures the impact of market forces in FX, interest rates, commodities, and product cost factors will provide an enhanced means to propagate more reliably shocks to these factors than purely exogenous-driven models (such as IAMs) can provide.

Implications for the detailed firm level of Agent-Based Modelling
In contrast to the purely exogenously driven model frameworks, that solely use IAM inputs, the use of heterogeneous agent-based behaviours, such as those applied in Cormack et al (2020), will give rise to deviations between macro-economic inputs (IAM inputs) and the resulting collective impact from the modelled firms. These differences will create notable impacts at the firm level. Indeed, there is evidence of this from reported outputs from several regulatory stress testing activities (e.g. Federal Reserve Board 2024). it also raises questions about the impact of firms on the emerging macro-economy. For example, firms in an industrial sector may face internal idiosyncratic limitations in raising capital, hindering a climate transition in a jurisdiction. This may not be reflected reliably in macroeconomic models if simplified factors are used for rates of capital formation, productivity, prices, etc.

Because of the potential materiality at the firm level, model developers should provide a description of the details as to how firms are modelled. This will allow financial end users to assess how these assumptions align with their own views for strategic planning and risk management. The ability to address and apply clear firm level responses to climate transitions with clear descriptions of capital costs (e.g. abatement expenditure approaches to address emissions through their supply chain and adaption expenditure) will provide a clear expectation of a baseline to investors. Hence where this information is clear (and economically reasonable) will define a valuation baseline, hence defining an expectation on modelling of further scenarios for stress testing.

Summary
The use of agent-based simulations provides a powerful means to capture richer features of the economy from market factors, firm level pricing, demand volatility and supply chain impacts and direct risks from physical events. Capturing, behaviours within these models provides a means of assessing the impacts of direct risks and the potential impact of adaption solutions without resorting to a lower resolution of economic phenomenology that will miss individual firm idiosyncratic risks and highlight gaps in conventional macro-economic models. The process of building micro-economic factors using more complex and detailed frameworks such as Agent Based Models, will clearly provide a wide range of agent features. These features as highlighted in this sub-section currently are not captured by current macro-economic climate models and simple mappings of their demand predictions on firms (or other micro-economic agents such as the population). The use of such models (and their functional equivalents), though requiring mode time to develop, provides the only means to assess firm level (or in general agent level) impacts and are likely to play an important role in future risk assessments.

2.2.6.2 Structure of the Taxonomy for Microeconomic Expansion Models

The taxonomy for describing the microeconomic expansion processes has a common structure across levels 3a and 3b. However, it is important to recognise where the boundary sits between the two and the implications in terms of methodological transparency. Whilst a financial end user has full control over the level 3b expansion, level 3a elements can create challenges where the methodologies employed are opaque. A comparison of recent regulatory scenario analysis exercises demonstrates how the positioning of the line between 3a and 3b can vary:

- The 2021 Bank of England CBES (Bank of England, 2021) exercise was largely based on NGFS scenarios with limited level 3a expansion, such as the provision of UK sectoral GVA pathways. Most of the expansion effort was therefore left to participating firms (i.e. the level 3b component of the scenario).
- The 2021 Bank of Canada/OSFI (Bank of Canada and Office of the Superintendent of Financial Institutions, 2022) exercise specified the methodology firms should use to calculate probability of defaults (PDs) for bonds and loans but left the participating firms to perform the calculations. This creates a hybrid 3a/3b element to the scenario generation.
- The 2021 ECB Economy-Wide Climate Stress Test (Alogoskoufis, et al., 2021) was run as a desktop exercise within the ECB with participating forms only providing exposures. In this approach the regulator becomes the ultimate end user defining all pathways with the whole expansion layer in level 3b.

The Bank of England CBES exercise provides a good example of loss of transparency in the level 3a expansion process. The macroeconomic pathways in the CBES scenarios were generated using NiGEM, calibrated with output from the REMIND-MAgPIE IAM. As the NGFS implementation of NiGEM does not have a sectoral representation of the economy the Bank of England chose to expand the pathways provided with the addition of UK sectoral GVAs and credit spread curves split by financials and non-financials. However, the public guidance for the CBES did not include details of how these pathways were constructed and there are elements that appear inconsistent with the headline pathways.

Unlike the level 1 and level 2 components of scenarios, which are often publicly available with associated documentation (e.g. through the NGFS Scenarios Portal), the expansion methodologies employed by financial institutions and the pathways they produce are not public. This restricts the authors ability to assess these processes, although financial institutions and their supervisors should do so. However, we present a taxonomy for this assessment illustrated with potential approaches.

We further breakdown expansion models into intermediate models and final instrument pricing models. The level 3 modelling assessment should follow the same form as level 2b. Expansion modelling will often make use of non-climate specific stress testing and pricing models. Therefore, consideration should be given to their applicability to climate scenario analysis. Often historic relationships will underpin the methodologies, and the reviewer should consider how likely these relationships are to hold under the overarching climate scenario narrative and to what extent the assumptions might bias the final output pathways.

The specific expansion models required will ultimately be determined by the sophistication of the scenario analysis required as well as the financial end users' business model and balance sheet. However, we present a range of both types to illustrate the application of the taxonomy.

a) Intermediate Models:

Intermediate models add resolution or detail to the macro pathways but without producing final financial asset or liability values. Examples include:

- Downscaling of chronic climate pathways into acute weather events
- National and sectoral GDP/GVA expansion
- Projection of corporate financials

A more complex process might involve a chain of intermediate models. Whether the links in such a chain are evaluated separately will depend on the materiality of the component to the reviewer.

b) Final Financial Asset/Liability Pathway Modelling:

Financial asset or liability pricing models represent the final step in the construction of an economic climate scenario. Financial instrument modelling is a major field, and a full review of possible approaches is beyond the scope of this paper. Examples of the financial instrument valuation models required include:

- Sovereign bonds (expansion for countries outside the macroeconomic model output)
- Corporate equity and debt
    - Equity fair market values
    - Bonds fair market values
    - Traded loan fair market values
    - Loan IFRS 9 impairments
- Real estate
    - Commercial real estate market values and operating costs
    - Residential real estate market values and insurance costs
- Commodities (outside the scope of the macro models)
- Derivative contracts
- Insurance Liabilities

2.2.6.3 Downscaling of Chronic Climate Pathways Iinto Acute Weather Events

Global climate models lack the spatial and temporal resolution to project acute weather events. Expertise in the downscaling of chronic pathways into acute events is primarily found in academia and the Natural Catastrophe (NatCat) modelling industry. These groups will usually be involved in the more sophisticated approaches.

The use of physics-based weather models provides the most rigorous approach to downscaling. Where such downscaling is performed by an independent scenario builder (using a third-party macro model) or by a financial end-user the assessment of the process should be covered in level 3 of the taxonomy. However, simpler approaches, often referred to as statistical methods, combining historic data and expert judgement are common. Such approaches typically lack the capability to recognise non-linearities in the response of the climate system but are much less computationally expensive. Therefore, they are more appropriate for shorter time horizon applications such as general insurance pricing with a 1-year contract length.

Classification System

We suggest the following classification system to summarise the key characteristics of the approach taken:

Scenario	**Component ID**			**Scale and resolution**				**Model type and complexity**				**Interfaces**		
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration
2023 ACPR Insurance Stress Test Short Term Scenario	**3a**	ACPR in-house analysis	Climate	Mainly department level but full historic data available	Department for Health data	N/A	Daily and beyond if required	Historic and expert-narrative based	Deterministic	Empirical and modelled	None	2022 weather/health data and dam burst hydrological model	Acute climate and health data	Unknow

Table 12: Physical Downscaling Classification System

Component Specific Considerations

Nature of Model Formulation: We identify five high level approaches, although a single scenario could blend combinations of these approaches:

- Expert Narrative Based – Experts simply opine on potential events. This is more likely to be employed when a deterministic pathway is required. An example from the 2023 ACPR short term-stress scenario is the assumed French dam collapse with local high-resolution hydrological modelling.
- Historic Data with Expert Adjustment – Historic events are assumed to be repeated but with expert adjustment. For example, NatCat modelers sometimes adjust the historic data set by removing less impactful periods on the basis that they are less likely to repeat in the evolving world. This has the effect of increasing the probability of more extreme events. As this approach produces a multiyear historic dataset it can be used to deliver deterministic or stochastic pathways. The ACPR short term stress that repeats the 2022 European heat wave in 2023 and 2024 is a very simple interpretation of this approach. However, a weakness of this approach is that it does not allow for events that have not previously been seen and climate scientists predict a non-linear increase in both the probability and severity of extreme weather events.

- Historic Data with Chronic Pathway Based Scaling - This approach combines historic data with information from lower resolution chronic climate pathways. For example, historically observed local temperature patterns can be assumed to repeat but scaled for the increase in spatially and temporally averaged temperature taken from a general circulation model. This has the advantage that the intensity of acute weather events can increase with climate change, but this may still fail to account for anticipated non-linear effects.
- Weather Models – This approach uses low resolution pathways from climate models as inputs to high resolution weather models – This is the most sophisticated approach combining climate and weather models of different resolutions.

Model Process: The model process is of particular significance for the downscaling of physical risks. A single deterministic pathway is useful for evaluating the impact of specific tail events but is of less value for projecting the pricing of assets as climate change crystallises and market expectations change. For example, we might examine the economic impact on a single Caribbean Island if it is hit by a category 5 hurricane next year but would not want to assume that there is no asset pricing impact on a nearby island that is unaffected by that event.

Resolution: The model process interacts with the spatial and temporal resolution. Whilst it might be possible to specify an individual acute event with very high-resolution, lower resolution might be appropriate when using probabilistic outputs. The resolution should still be sufficient to distinguish between meaningfully different levels of risk. To extend the Caribbean Islands analogy; there may be no meaningful differentiation in risk between two nearby islands even if they would potentially experience very different impacts from an individual storm. However, as the distance between two islands increases to the extent that there is a meaningful difference in the likelihood, severity and duration of tropical cyclones then the resolution should be sufficient to capture these variations. Insufficient resolution that averages impacts either spatially or temporally could hide risk. For example, peak wind strength or flood depth in a particular location are the likely drivers of economic damage rather than a longer-term average. Where relevant peak data is not available compensation may be required in the final steps of the modelling chain.

Integration: By their nature level 3 expansion processes sit outside the integrated macro models and so it is more challenging to integrate any feedbacks and, when trying to do so, avoid overlaps with the macro modelling. Therefore, it will be common for scenarios to rely on capturing feedbacks at the macro model level. In terms of physical feedbacks this will often mean modelling the feedback response to chronic climate change. However, in some cases scenario users might attempt to include additional feedback due to acute events.

Taxonomy application example

The 2023 ACPR Insurance Stress Test provides a good illustration of downscaling based on a combination of historic data, expert judgement, and detailed hydrological modelling:

Classification: Shown in table 12 above

Scenario: ACPR 2023 Insurance Stress Test – Short-Term Scenario

Model: ACPR Internal Analysis

Detailed Model Description: Historic replication for European weather and health data based on 2022 combined with hydrological modelling for the assumed failure of the Serre-Ponçon dam in 2025.

Limitations, **Assumptions & Comparison with Alternative Models:** The detailed acute weather scenario and associated hazards are specific to France and provides a single deterministic pathway. The repetition of extreme historic events removes the need for modelling. However, comparison with model outputs would allow the probability of such repeat events to be estimated.
An assessment of the hydrological modelling of the dam failure is beyond the scope of this paper.

Summary & Recommended Use: The downscaling is appropriate for testing vulnerabilities in domestic French insurers but lacks broader applicability. Multiple repetition of historic extreme weather events does not assess the vulnerability to the more extreme events that might be anticipated as climate change progresses. The documentation does not express the likelihood associated with the repeat events so users would need to perform their own analysis to understand whether the scenario captures the severity of risks they wish to manage exposure to.

2.2.6.4 National and Sectoral GDP/GVA Expansion

Sectoral GVA pathways are commonly used in the calibration of financial asset pricing models but sit between the macro and microeconomic modelling worlds. Macroeconomic models with sectoral representation (e.g. input-output models) will naturally produce sectoral GVA pathways. More granular agent-based models can also produce GVA pathways by aggregating the relevant agents. In these examples the expansion of GVA pathways might be a matter of adding sectoral or national resolution using the available pathways as proxies. The expansion of GDP curves for countries not resolved by the macro models may also be subject to a similar proxy-based approach.

However, if there is no sectoral resolution in the macro models then the expansion is more challenging and inevitably more judgmental. Assuming a national GDP curve has been produced by the macro models or using a proxy methodology then disaggregation into sectoral impacts might be based on a judgmental assessment of the various risks faced by a particular sector within a national economy. For complex model architectures, which combine IAMs with macroeconomic models (as illustrated in figure 5), there may be sectoral information buried within the IAM that is not utilised in the final macroeconomic models. This is the case with the NGFS combined REMIND-MAGPiE and NiGEM architecture that underpinned the Bank of England's CBES exercise. With such an architecture there is a danger that the assumptions made in the expansion of GVA pathways are inconsistent with those made within the IAMs.

Classification System

Scenario	Component ID			Scale and resolution				Model type and complexity				Interfaces			Component specific	
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Transition risk assessment	Physical risk assessment
BoE CBES Scenarios	3a	Expert judgement	Macroeconomic	National (US, UK, Germany)	National (US, UK, Germany)	9 major sectors plus other	5 years	Expert judgement	Deterministic	Experts judgement	None	Sectoral emission intensity projections	Sectoral GVAs for 9 major sectors plus "other sectors"	None	Emissions intensity based scorecard	Vulnerability and adaptation-based scorecard

Table 13: National GDP/Sectoral GVA Expansion Classification

Component Specific Considerations

Nature of Model Formulation: We identify three basic approaches to the expansion of sectoral GVA and national GDP pathways:

- Expert Judgement – The scenario user judges the relative impact of the various risk transmission channels on the sector or country. For sectoral breakdown of an available GDP pathway this is commonly combined with a simple model to ensure that the sectors aggregate to the GDP.
- Proxy Based – Sectoral or national GVA or GDPs are assumed to follow similar available pathways potentially with adjustment that could be either expert or historic regression based.
- Stand-Alone Macroeconomic Model – the scenario builder or financial end user employs a macroeconomic model with the required sectoral or national representation calibrated with key pathways, such as carbon price and chronic climate indicators from the 3rd party scenario.

Transition and Physical Risk Assessments:

These are both free format fields allowing the user to summarise the extent to which the two main risk transmission channels are captured. Common classifications for simple expansion models are:

- Not captured
- Embedded in historic time series data
- Expert judgement

Taxonomy Application Example

The 2021 CBES exercise is an example of a scenario provider calculating GVA pathways where they have not been provided by the modelling architecture adopted:

Classification: Shown in table 13 above

Scenario: All Bank of England CBES Scenarios

Model: Undisclosed Internal Bank of England Analysis

Detailed Model Description: Details of the model are undisclosed.

Limitations, Assumptions & Comparison with Alternative Models: Unknown

Summary & Recommended Use: Without any disclosure of the methodology employed users cannot judge when and how to use these scenario pathways. Where the CBES scenarios are to be used by financial institutions for internal risk management purposes it is recommended that they perform their own sectoral expansion using a methodology that is appropriate to the intended use.

When regulators set scenarios without full disclosure of the scenario the on-going value of the exercise is reduced, even though the regulator will have the in-house knowledge required when performing its own analysis of the results of the initial regulatory exercise.

2.2.6.5 Projection of Corporate Financials

The detailed bottom-up modelling of the corporate exposures held by financial institutions (equities, bonds, loans and derivative CVA), covered in the next section of the taxonomy, requires the intermediate modelling of corporate financial performance to project the financial statements and metrics (balance sheet, profit and loss, dividend payments and associated ratios) that act as inputs to asset pricing models. This modelling should consider the following (as discussed in section 2.2.6.2):

- The impact of transition policy on operational costs
- The impact of physical climate change on operational costs
- Technological evolution and associated investment requirements
- Customer behaviour
- Competitor behaviour
- Macroeconomic demand effects including non-climate specific scenario elements
- Financing costs including dividend policy

Combining these factors, the model can project the sales volume, operating cost, profit margin and capital investment data that allow forward projection of corporate financials. The approach taken will depend on the range of data provided by the macro models and the sophistication of analysis required. For core sectors such as energy and power the macro models should provide demand and price data leaving the downstream modelling to determine market share. Where this data is not available the downstream models will need to determine the market equilibrium demand and price considering the factors listed above. The modelling assumptions and calibration choices at this stage of the scenario generation process can dramatically impact the results and should be tailored to fit the overall scenario narrative.

Classification System

We propose the following classification system for the modelling of corporate financial performance. Due to the wide range of potential approaches available we leave the individual classifications as free text rather than enumerating a list of options:

Scenario	**Component ID**			**Scale and resolution**				**Model type and complexity**				**Interfaces**			**Component specific credibility**					
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Transition policy	Direct physical damage function	Technological evolution	Macroeconomic considerations	Market dynamics	Finance costs and dividend policy
Baer et al 2022. Power sector stress scenario	3b	TRISK	Micro	N/A	Firm	Power sector	Annual	Fundamental analysis incorporating impact of carbon tax	Deterministic	Expert judgement	None	Carbon prices, corporate transition plans, current market shares by technology	Power sector asset values and leverage ratios	None	Carbon price from NGFS delayed transition scenario	N/A	Corporate plans from asset resolution data followed by IEA SDS pathways	Embedded in the IEA SDS scenario pathways	Constant market share by technology with market size and prices taken from the IEA SDS scenario	N/A

Table 14: Corporate Financial Projection Classification System

Component Specific Considerations:

General: Given the wide array of models available we do not propose any categories for the key component specific features and the fields should be considered free format.

Geographic Resolution: This is a key element for the assessment of physical risks. Basic treatments might assume a single location based on head office location, but more advanced treatments will consider location of individual production assets, supply chain effects and the geography of sales markets.

Economic Resolution: By its nature economic resolution be at least to firm level but more detail will be required to effectively handle conglomerates with multiple business lines with varying climate sensitivities.

Nature of Model Formulation: By its nature the projection of corporate financials is a bottom-up piece of fundamental analysis. Typically, this field will therefore be populated with “Fundamental Analysis" plus any relevant detail. The existence of this modelling step in the architecture provides critical insight into the way asset values are determined.

Transition policy: The representation of policies that restrict or tax GHG emissions. This will typically take the form of a carbon tax or price but could also include production restrictions, bans and cross-border taxes.

Direct physical damage function: The choice of modelling to capture the impact of chronic and acute climate change on the cost of production. The indirect effects of climate change on demand are captured under macroeconomic considerations.

Technological Evolution: The model should consider the impact of technological evolution on emissions, operating costs and capital investment. For example, IEA scenarios provide projections of the different deployment and cost evolution of technologies that determine how the overall energy system is shifting based on shifts in the cost-competitiveness of technologies following the introduction of a carbon tax that alters relative prices. The assumptions made should be consistent with those used in the macro modelling and any incoherencies should be identified.

Macroeconomic considerations: Broader changes in the economy will impact demand for goods and services, resulting in new demand elasticity driven market volume-price equilibria. Where the relevant sector is covered in sufficient detail by the macro modelling these considerations will normally be embedded in the input pathways. However, where there is insufficient granularity in these pathways the expansion models will need to fill any gaps.

Market dynamics: The approach taken to market dynamics should capture the strategy of an individual company in relation to its competitors, allowing for any changes in client behaviour. This will include the models handling of demand and substitution elasticities that drive the evolution of a company's sales volumes and profit margin.

Finance costs and dividend policy: Corporates will need to finance their operations using a combination of bank lending and capital markets. The strategy assumed will impact the leverage ratio of the firm. Recognition of changes in interest rates, credit risk appetite and equity risk premium will all impact financing costs and ultimately a corporate's ability to make the investment required to transition. Scenarios that seek to capture the feedbacks between the real economy and the financial system will need to handle this aspect of a corporate's financial performance in reasonable detail.

Taxonomy Application Example
Most models currently used by financial institutions to project corporate financial performance in climate scenarios are either in-house or provided by external consultants. In neither case are the detailed methodologies publicly available. However, the academic literature does provide an example to which we can apply the taxonomy. Baer et al. (2022) puts forward the TRISK model co-developed with 2DII (now Thea Finance Labs) to project the performance of power sector companies under multiple transition scenarios. The resultant financial statements feed the projection of financial asset values and probability of defaults as covered in the final layer of the taxonomy.

Detailed Analysis

Classification: Shown in table 14 above

Scenario: Baer et al 2022 Power Sector Stress Scenario[11]

Model: TRISK

Detailed Model Description: Market dynamics are embedded in the energy mix, electricity usage and electricity prices taken from the IEA SDS scenario combined with the constant market share by technology assumption. The modelling of corporate performance is centred on profit margin which adjusts with the impact of carbon tax allowing for the individual firm decarbonization strategy. The assets of the corporation are adjusted for stranding of emissions intensive technologies. However, finance costs including dividend payments are not explicitly evaluated and are assumed to scale with the other costs of production when projecting profit margin. The model assumes firm-specific discount rates when modelling net present values of firms, but those remain scenario independent.

Limitations, **Assumptions & Comparison with Alternative Models:** The constant market share assumption does not permit a comparison of a firm's competitive strategic response against its peers. The model projects the NPV of a firm based on its misalignment with a decarbonisation scenario. The results highlight the transition related changes that a firm may face if it continues to follow a business-as-usual strategy, calibrated on the 5 -year capex plans for its business units with market shares held constant to its average technology-market share. There is a simplistic treatment of financing costs embedded in the profit margin projection. There is no impact of physical climate change.

Summary & Recommended Use: TRISK is an intermediate sophistication corporate projection model designed primarily to capture the impact of climate mitigation policy action as represented by policy-induced demand shifts, relative changes in technological unit costs and an additional carbon price. The constant market share assumption is most relevant to mature sectors with high barriers to new entrants but needs to be sensitised for a better understanding of the risks, where sectors are facing a higher penetration of new market entrants and highly competitive dynamics among peers. The absence of any physical impact makes the approach less appropriate for longer term scenarios where physical climate impacts are likely to be material.

[11] The TRISK model extends to the projection of corporate probabilities of default using a Merton framework, but this sits in the final link of the modelling chain discussed in the next section.

2.2.6.6 Final Financial Instrument Valuation

The projection of the value of an individual asset or liability on the balance sheet of a financial institution is the final link in the economic climate modelling chain. The output is the market price of a financial instrument, or a parameter that feeds an accounting model (e.g. probability of default, loss given default for use in the calculation of IFRS9 impairment) or regulatory model (prices, volatilities and correlations).

Classification System

Scenario	**Component ID**			**Scale and resolution**				**Model type and complexity**				**Interfaces**			**Component specific credibility**	
	Taxonomic level	Model name	Model type	Geographic granularity and extent	Economic granularity	Industry sectoral granularity	Temporal granularity and extent	Nature of model formulation	Model processes	Calibration type	Model uncertainty	Key inputs	Key outputs	Integration	Key calibration choices	Market foresight
Conceptual 1	3b	KMV	Financial market factor	N/A	Financial instrument	N/A	Single step	Mixed technical and fundamental analysis	Deterministic	Empirical	None	Corporate leverage ratio, asset volatility, historic default dataset	Corporate probability of default	Non	Asset volatilities derived from historic equity implied volatilities	Embedded in asset volatility
Conceptual 2	3b	Dividend discount model	Financial market factor	N/A	Financial instrument	N/A	Single step	Fundamental analysis	Deterministic	Empirical	None	Corporate financial statements, return on equity	Equity price	Non	Return on equity and dividend payment strategy	Choice of projection of corporate financials beyond forward valuation data

Table 15: Financial Instrument Valuation Approach Classification

Component Specific Considerations

Nature of Model Formulation: We identify three categories of approach to the final stage valuation of a financial instrument:

- Expert Judgement: The translation of the higher-level pathways to the value of a financial asset is based on expert judgement. However, expert judgement is often based on experience and so can share common features with the technical analysis. These elements should be brought out in the analysis of the approach so that the assumptions are understood, and the results can be appropriately interpreted.

- Technical Analysis. In the day-to-day valuation of financial instruments technical analysis refers to the use of the historic charts and trends to identify pricing inconsistencies and hence project price movements. Similarly, we employ the term to capture all methods that use historic price relationships to project the valuation of a financial instrument. This might be some form of regression model or, in a simplified format, the use of a proxy. The nature of regression required will depend on the higher-level pathways available. This approach is typical of financial institutions existing stress testing frameworks.

- Fundamental Analysis – Fundamental analysis is based on projecting the financial statements of an underlying corporation or the projected returns of a specific project or asset for limited recourse lending. Where such detail is available (see intermediary models) the projection of financial instrument values can be used by the wide range of pricing models available in the market. There is a wide body of literature covering the use and suitability of these models but for their use in climate scenario analysis the user must be cognizant both of their general limitations when projecting future values and the specific issues that arise in climate scenarios when historic relationships can be expected to break down.

Key Calibration Choices: The detailed analysis of the modelling approach will identify all inputs and their calibration. However, it is useful to identify the key inputs that impact the values produced and their interpretation. For example, the Merton framework (Merton 1974) is commonly used to determine the probability of defaults on loans and bonds. The inputs to the model are the leverage ratio, the risk-free interest rate and the asset volatility. The first two of these come from the models higher up the chain but asset volatility calibration is key, and arguably a weakness of this model, for projecting financial instrument valuations in scenario analysis. When employing the Merton framework to current or near dated valuations the user might deduce the asset volatility from current the equity price volatility implied from equity option prices, but this does not allow any evolution of market risk appetite. An alternative approach is to use stressed volatilities from appropriate historic time periods, but these may not reflect the nature of a future stress event in which climate risks are either the trigger or an amplifier.

Market Foresight: Market foresight is the extent to which the market is deemed to predict the future scenario pathway beyond the valuation date of the financial instrument. For example, in a late policy action scenario the market might be deemed to have perfect foresight from the point policy action is announced. If the market is deemed to have no foresight it will base its decisions and valuation on the available historic data. The choice of market foresight is typically not a specific parameter but is embedded in the choice of model and the calibration of its inputs. Although market foresight might be considered a temporary valuation issue, that does not impact the long-term performance of an asset, there is potential for feedbacks from the financial system (e.g. the availability and cost of finance to a corporation) that can be captured in more sophisticated scenarios.

Taxonomy Application Examples

There are a vast number of well documented and well understood financial pricing models to which we can apply the taxonomy. We compare two common approaches used in the calculation of corporate default probabilities and equity valuations.

Classification: Shown in table 15 above

	KMV	DDM
SCENARIOS	N/A	N/A
MODEL NAME	Kealhofer, McQuown & Vasicek (KMV) default probability model (Vasicek 1984)	Dividend Discount Equity Pricing Model
DETAILED MODEL DESCRIPTION	Calculates distance to default in terms of number of standard deviations and equates this with probability of default based on historically observed number of defaults for given distance to default. Fundamental analysis could be used to project asset volatilities.	Determines equity price as the net present value of all future dividends discounted at the required return on equity.
LIMITATIONS	No loss given default so potentially use current market standard. Asset volatility is hard to calibrate. The historic relationship between distance to default and probability of default is assumed to hold under the climate scenario pathway.	Dividend discount models are commonly used to value mature companies with steady dividend payouts. The valuation is heavily dependent on the dividend payout assumptions in the financial projections and the expected return on equity required by the market.
SUMMARY & RECOMMENDED USE	The KMV model is a market standard for assessing current probabilities of default. It is more challenging to use in for forward valuations due to the difficulty of projecting the asset volatility. Simply taking historic volatilities is likely to underestimate the risk in sectors that are materially impacted under the chosen scenario narrative, a weakness that magnifies with the time horizon.	Dividend discount models are commonly used to value mature companies with steady dividend payouts. The valuation is heavily dependent on the dividend payout assumptions in the financial projections

Table 16: Comparison of Two Asset Valuation Models Using the Taxonomy Structure

Note - The taxonomy is designed to represent the application of the model component in a specific scenario, allowing the assessment of the calibration in the specific instance. The above examples are more generic as there is a lack of publicly disclosed scenarios that extend to the pricing of financial assets with fully disclosed methodologies.

2.2.6.7 Level 3 - Dynamic Balance Sheet Modelling

For completeness we also consider dynamic balance sheet modelling by financial institutions although a full review is beyond the scope of this paper. Traditional scenario analysis and stress testing assumes that the balance sheet of the financial institution remains fixed. Relaxing this assumption brings in the role of dynamic management of the balance sheet and the generation of future revenues. This modelling is very institution specific. Conceptually it fits well with high resolution agent-based modelling where the individual financial institution is represented as an agent. In practice far more simple approaches are likely. This type of modelling is in its infancy and there is very little in the public domain to review. However, conceptually the taxonomy could be extended to cover the necessary modelling steps. The earlier links in the modelling chain should act as inputs to the modelling of the dynamic balance sheet but the pathways required would include wallet size by market, market share and hedge bleed for derivatives portfolios.

3 Linking Models with the Real World and the Application of the Taxonomy

Despite the focus on models in the foregoing discussion, we know that models are not perfect representations of the real world. After classifying modelling approaches, we must consider how to use imperfect models to provide relevant advice and support decisions in the real world. The examples provided in the previous section include quite simple assessments of the limitations with summary recommendations for the use of the scenario constructed. However, more rigorous analysis is needed if financial institutions are to progress to decision useful scenario analysis. In this this chapter we consider in more detail how the assessment process might work.

This requires evaluation of models to understand where their strengths and weaknesses lie, so that we can make use of valuable information and not be misled by poor information. For complex climate scenario modelling chains this can apply to both the assessments of the individual components covered by the application of the taxonomy and the overall scenario. There are two approaches to model evaluation, both of which are necessary.

3.1 Quantitative Model Evaluation

Quantitative evaluation is performed both during and after model development, by comparing model output with real-world data. During model development, this consists of calibration (tuning) or data assimilation to make the model conform as closely as possible to existing observations of the system. After model development, new data (ideally data not used in the construction of the model) can be used for out-of-sample evaluation.

Example:
Climate models are quantitatively evaluated using a very wide range of data, including:

- Ability to reproduce past climates, from palaeoclimatic history (based on data from ice cores, sediments and tree rings) to the recent past (based on data from physical measurements and satellite observations).
- Ability to represent emergent system behaviours such as atmosphere and ocean circulation patterns, storm tracks, monsoons, and the patterns of polar amplification and land/sea contrast.
- Detailed evaluation of modules or subprocesses with respect to observational data, such as atmospheric chemistry, ice sheet mass balance, physics of cloud formation, etc.

Models are also evaluated by looking at consistency between different modelling approaches, through:

- Sensitivity analyses quantifying the uncertainty within a model due to the range of possible parameters or initial conditions.
- Model intercomparison or use of multi-model ensembles.

3.2 Qualitative (Expert) Model Evaluation

Models cannot be evaluated solely by quantitative methods, essentially because "*past performance is not a guarantee of future success*." Some degree of expert judgement is always required to understand the extent to which we expect the future to be like the past in the ways that matter for the model. When we forecast the weather, we expect that tomorrow's weather will (usually) be within the range of weather that we have experienced in the past, and therefore that our models are forecasting within the domain that they have been well-calibrated for. But when we forecast the climate, we know that many climatic variables are moving rapidly beyond the range that we have experienced or measured in the past and been able to calibrate our models for. We expect the laws of physics to continue to hold, but we do not necessarily expect empirically-derived parameterisations (such as those pertaining to Arctic sea ice, or cloud formation) to continue to be numerically correct in a significantly-different climate. Additionally, the balance of different processes might be expected to change as we look further into the future, which makes it difficult to evaluate some aspects of model performance using behaviour in the present day. For instance, we might expect to see more amplification of warming in future in increasingly moisture limited conditions.

Qualitative evaluation is performed both during and after model development, using the expert domain knowledge of the modeller. During model development, this consists of judgements about the quality of any data used, which systems or processes are necessary to include in the model and how to represent them. After model development, the model outputs can be assessed for overall plausibility in terms of whether they agree with past observations, show expected behaviours, or reproduce notable features of system dynamics.

3.3 Appropriate Processes for Tailored Model Evaluation in Financial Applications

3.3.1 Defining the Questions

Let us say that we have a decision question for which the future climate is a relevant input because it affects the level of risk we are taking on. How should we decide what models to use, and the level of confidence we should have in those model outputs? We are looking for information at the intersection of the following Venn diagram:

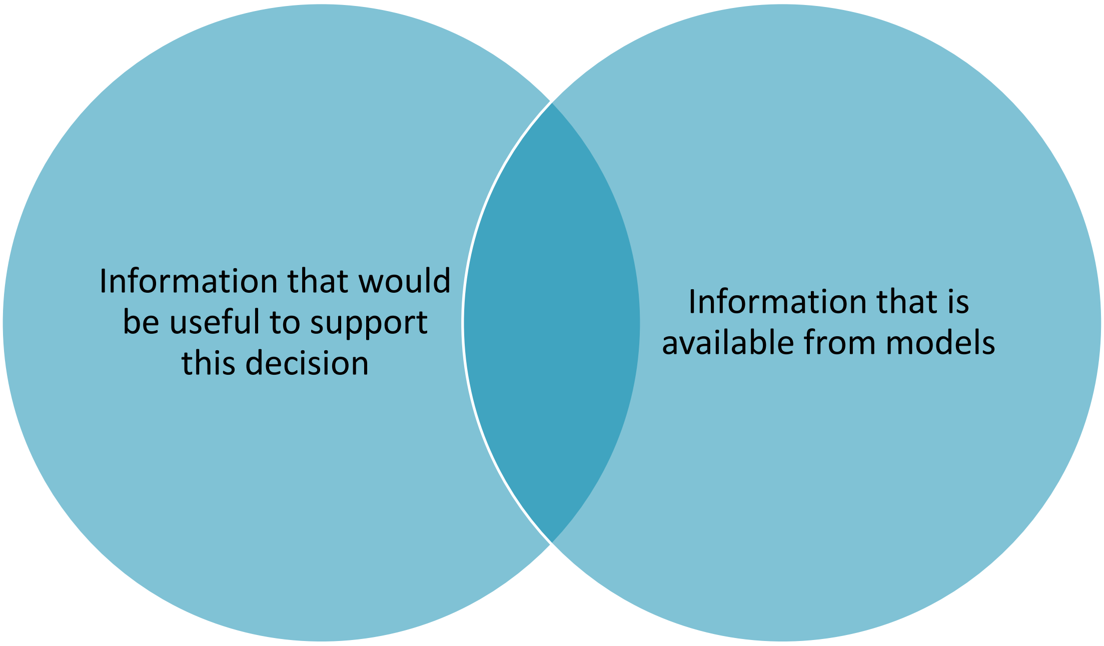


Figure 21: The intersection between model output and decision useful information

In general, there is a lot of information available that is not useful (such as the weather in London in a climate model on 1st July 2080), and a lot of information that would be useful, but is not available (such as the maximum wind gust on a 100m grid). Table 17 outlines the steps required to identify the best data available to support a particular decision:

Step	Example (climate decision, hypothetical)
Define the decision question	How should I adapt my investment portfolio given climate risks?
Define the climate-related aspect of the question	What is the current and future weather risk to a piece of infrastructure in London? Wind, Flood, Heat, Compound event physical hazards at a specific location.
List the climate information that would be maximally informative for this question	100% confident information about • Wind, • Flood, • Heat, and • Compound event physical hazards for this exact location

	• now, and • at a series of future time points.
Given that this is not available, what is the closest available information?	Uncertain information about • Wind, • Flood, • Heat, and other physical hazards, on a 10km grid including this location, • now, and • at a series of future time points, conditional on some specific climate scenario(s); conditional on a certain model of internal variability. Not much information about compound events.
What models are available?	Global climate model (not detailed enough) Downscaled impact models Catastrophe models Statistical models fitting previous data
Which models are useful?	Proceed to evaluation (see Table 3)

Table 17: Climate information

3.3.2 Tailored Evaluation

As discussed above, both quantitative and qualitative evaluation is required. Scientific papers describing the development of a model usually contain quantitative evaluation relative to scientific metrics of interest. Adequacy for the purpose of scientific enquiry, however, is different from adequacy for the purpose of informing a specific real-world decision question. This is why tailored evaluation is important for prospective users of scientific models.

Table 18 shows an evaluative framework for climate information, based on the criteria of Baumberger et al (2017). Quantitative evaluation tends to be the domain of scientific practice and, depending on the kind of relationship a prospective user has with the information provider, it may often not even be possible to request further quantitative evaluation. As such, the qualitative element is vital and should not be underestimated. While some think that quantitative evidence is superior to qualitative judgements, the message here is that quantitative evidence is incomplete without accompanying qualitative evidence.

Qualitative evidence can be built up in a conversational fashion between modeller and decision-maker, following the series of questions in the right-hand column of Table 18. Following that line of questioning, the overall question is “Do we have sufficient evidence that this model provides relevant information to inform our decision?”

B17	Consideration	Quantitative evidence (Provide for inspection)	Qualitative evidence (Discuss)
Empirical accuracy	A. Data	How good are data? • Quality controlled? • Comprehensive? • Relevant? Sensitivity analysis to uncertainties in data	Do we believe the data are reliable? Are the data also model-laden? Are there any other data relevant to our decision question?
	B. Model behaviour	Comparison of hindcasts with observations, using quantitative evaluation metric	Justification for choice of evaluation metric relative to our decision question
	C. Model output (forecasts, projections, or predictions)	Comparison with out of sample dataset, using quantitative evaluation metric	Justification for choice of evaluation metric relative to our decision question
Robustness	D. Uncertainty	Sensitivity analysis to • Initial conditions • Parameters • Other choices?	How much of overall uncertainty do we believe is captured by sensitivity analysis? What uncertainties remain, that are outside the scope of the model?
Coherence with background knowledge	E. Choice of systems/processes	Sensitivity analysis to choice of systems/ processes	Why did we choose these systems/processes to include in the model and why did we exclude others? Are they relevant to our decision?
	F. Representation of systems/processes	Sensitivity analysis to choice of representations (e.g., use of differential equations vs agent-based model)	Why did we choose to represent systems in this way? Are they relevant to our decision?

Table 18: An evaluative framework for climate information, based on the criteria of Baumberger et al (2017).

3.4 Outcomes of Model Evaluation

The outcome of model evaluation should not be a quantitative performance metric. The outcome of model evaluation should be a clear understanding of the use and limitations of a model for the purpose to which you intend to put it (this should be evidenced both by quantitative performance metrics and by expert judgement).

On the one hand, we must be clear about the limitations. It is no use to know that a model is "98% accurate" if the 2% of occasions when it is wrong are those which are business-critical or destroy the planet. Conversely, we must keep in mind that a well-performed model evaluation will always tell us that the model is not perfect, and that there is a possibility that it could be wrong or misleading. This does not mean that it is completely useless or has no information to give.

Table 19 walks through an example of a discussion framed around the qualitative evaluation questions laid out in Table 18, using the same hypothetical example as Table 17. Imagine the decision-maker sitting down with the modeller and asking the questions in the left column, with the modeller providing their expert judgement about the quality of the model and the decision-maker focusing that on topics of relevance to the specific decision.

Evaluation question	**Example (climate decision, hypothetical)**
Do we believe the data are reliable? Are there any other data relevant to our decision question?	Wind; Flood; Heat: previous extreme data available but future events are model-dependent. Wind is physical hazard. Flood depends on future vulnerability, e.g. flood defences (scenarios). Heat is affected by UHI effect (separate model). Compound events not considered by models.
Justification for choice of (hindcast) evaluation metric relative to our decision question	Wind: max gusts are difficult to model so use data to calibrate an empirical fit to extend physical model to extremes (climate + stats). Good fit to past data, but data is of low quality. Flood: depth of water is provided by model (climate + hydrology) with sufficient resolution for decision. Reasonable fit to past data, but not much data available. Heat: max temperature is provided by additional UHI model with urban scenario (climate + impact model). Good fit to past data. Compound event: not available.
Justification for choice of (forecast) evaluation metric relative to our decision question	Some data available from recent academic papers evaluating model performance for a recent flood event, using flood depth as metric. Noted the additional future uncertainty of global sea level rise. No data on other hazards.
How much of overall uncertainty do we believe is captured by	Discussion with modeller suggested that modelled uncertainty ranges reflect 10-90%

sensitivity analysis? What uncertainties remain, that are outside the scope of the model?	confidence intervals. Outcomes beyond this range could be due to (wind) extreme sting jet or local tornado, (flood) failure of flood defences or global SLR beyond modelled ranges, (heat) extended blocking, heat island, soil moisture feedbacks. Discussed extension of models and noted scenarios. Compound events are not considered but important (wind/flood).
Why did we choose these systems/processes to include in the model and why did we exclude others? Are they relevant to our decision?	Discussed the failure modes of model as described above. Some relevant processes are lacking the models, but these can be captured by expert judgement and scenario analysis.
Why did we choose to represent systems in this way? Are they relevant to our decision?	Outside scope to consider alternate representations but fed back to modeller that consideration of compound events would be valuable for decision support.

Table 19: Example notes of a discussion answering the hypothetical question "What is the current and future weather risk to a piece of infrastructure in London?" Details are imagined, to show the kinds of considerations that should be discussed.

Following the above discussion (which will be different in every case), the prospective decision-maker can:

- Assess the relevance of risks quantified by model.
- Understand whether/where additional models are needed to enhance relevance, and the increased level of uncertainty entailed by using further models.
- Understand the degree of confidence attached to various kinds of modelled outcomes, including the case where models simply do not have any relevant information to give.
- Understand the failure modes of a model and the kinds of situation that would cause failure or result in events beyond the modelled range.
- Understand whether additional scenarios are needed to account for model limitations.
- Take a position on how to use the modelled information in a decision process.

Evaluation does not have to be expensive. Depending on the importance of the decision, a different level of resource may be committed to the evaluation process. If we want to do something which is cheap, reversible, and affects only a few people in minor ways (e.g. whether to add shading blinds to windows in a new build), then a basic conversation around existing models is enough. If the decision is expensive, irreversible, or affects many people in profound ways (e.g. a choice of where to locate a new nuclear power station or how to plan urban development and retreat close to sea level), then a comprehensive programme of tailored model development and evaluation for this specific purpose would be merited.

4 Conclusions

The value of the integrated climate economic scenarios currently available to the financial sector continues to be questioned due to the limitations and assumptions inherent in the modelling. Although even highly simplified models can have value, a lack of in-depth understanding of the modelling is a significant obstacle to making climate scenario analysis decision useful in the sector. We believe that the development of a broader range of well understood climate scenarios is vital if climate related financial risks are to be well understood. Developing the understanding of these scenarios across FIs, CB&S and other financial regulators will require the effective and timely transfer of knowledge between these actors.

We have presented a practical taxonomy to describe and classify climate scenarios, the granularity of which can be adjusted according to the needs of the user and the sophistication of the scenario under consideration. We believe that the adoption of a taxonomy to deliver a structured approach to the documentation, assessment and use of climate economic scenarios can play a significant role in promoting the knowledge transfer required to make scenario analysis decision useful.

We urge CB&S to introduce regulation that demands a rigorous and structured approach to the evaluation and use of integrated climate economic scenarios. Such regulation will help to drive investment in developing the broader range of scenarios required and improving the sharing of knowledge between scenario builders and financial end users. We also recommend that an international body, such as the NGFS take on the role of providing a library of peer reviewed scenarios that have been rigorously evaluated using the type of approaches outlined in this paper. National hubs, such as the CGFI, could perform a similar role for more locally focussed scenarios.

5 References

Alogoskoufis, S. et al., 2021. *ECB economy-wide climate stress test - Methodology and results.* [Online]
Available at: https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op281~05a7735b1c.en.pdf

Autorité de Contrôle Prudentiel et de Résolution , 2023. *Scenarios and main assumptions of the 2023 climate stress test exercise.* [Online]
Available at: https://acpr.banque-france.fr/en/scenarios-and-main-assumptions-2023-climate-stress-test-exercise

Baer, M. et al., 2022. *TRISK - A Climate Stress Test for Transition Risk.* [Online]
Available at: https://ssrn.com/abstract=4254114

Baer, M., Gasparini, M., Lancaster, R. & Ranger, N., 2023. Toward a framework for assessing and using current climate risk sceanrios within financial decisions. *UK Centre for Greening Finance and Investment Discussion Paper.*

Baldissera Pacchetti, M., Jebeile, J. & Thompson, E., 2024. For a Pluralism of Climate Modeling Strategies. *Bulletin of the American Meteorological Society,* Volume 105, pp. E1350-E1364.

Bank of Canada and Office of the Superintendent of Financial Institutions, 2022. *BoC-OSFI-Using Scenario Analysis to Assess Climate Transition Risk.* [Online]
Available at: https://www.bankofcanada.ca/wp-content/uploads/2021/11/BoC-OSFI-Using-Scenario-Analysis-to-Assess-Climate-Transition-Risk.pdf

Bank of England, 2021. *The 2021 biennial exploratory scenario on the financial risks from climate change. Discussion Paper 33,* s.l.: s.n.

Battiston, S., Mandel, A., Monasterolo, I. & Roncoroni, A., 2023. *Climate Credit Risk and CorporateValuation.* [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4124002

Baumberger, C., Knutti, R. & Hirsch Hadorn, G., 2017. Building confidence in climate model projections: An analysis of inferences from fit. *WIREs Climate Change,* 8(3), p. e454.

Cash, D. et al., 2002. *Salience, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making.* [Online]
Available at: https://ssrn.com/abstract=372280

Cliffe, M. et al., 2023. *No time to lose - new scenario narratives for action on climate change.* [Online]
Available at: https://greenfuturessolutions.com/news/no-time-to-lose-report-uss-university-of-exeter/

Collins, M., Tett, S. & Cooper, C., 2001. The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments.. *Climate Dynamics 17, 61–81,* Volume 17, pp. 61-81.

Cormack, C., Donovan, C., Köberle, A. & Ostrovnaya, A., 2020. Estimating Financial Risks from the Energy Transition: Potential Impacts from Decarbonization in the European Power Sector. *Journal of Energy Markets,* 13(4).

Cormack, C. & Shrimali, G., 2023. *The Challenge of Climate Risk Modelling in Financial Institutions - Overview, Critique and Guidance', SSRN.* [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4669195

Dellink, R., Chateau, J., Lanzi, E. & Magné, B., 2017. Long-term economic growth projections in the Shared Socioeconomic Pathways,. *Global Environmental Change,* Volume 42, pp. 200-214.

Dunz, N., Naqvi, A. & Monastrolo, I., 2022. *Climate transition risk, climate sentiments, and financial stability in a stock-flow consistent model.* [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4254054

Federal Reserve Board, 2024. *Pilot Climate Scenario Analysis Exercise: Summary of Participants' Risk Management Practices and Estimates.* [Online]
Available at: https://www.federalreserve.gov/publications/2024-may-pilot-climate-scenario-analysis.htm

Fiedler, T. et al., 2021. Business risk and the emergence of climate analytics. *Nature Climate Change,* Volume 11, pp. 87-94.

Fricko, O. et al., 2017. The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century.. *Global Environmental Change,,* Volume 42, p. 251–267.

Gent, P. & McWilliams, J., 1990. Isopycnal mixing in ocean circulation models.. *Journal of Physical Oceanography,* Volume 20, pp. 150-155.

Giorgi, F., 2019. Thirty Years of Regional Climate Modeling: Where We Are and Where Are We Going Next. *Journal of Geophysical Research,* Volume 124, pp. 5696-5723.

Gohar, L. K., Lowe, J. A. & Bernie, D., 2017. The impact of bias correction and model selection on passing temperature thresholds. *Journal of Geophysical Research: Atmospheres,* Volume 122, pp. 12,045–12,061.

Gordon, C. et al., 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments.. *Climate Dynamics,* Volume 16, pp. 147-168.

International Institute for Applied Systems Analysis (IBF-IIASA). , 2023. *Global Biosphere Management Model (GLOBIOM) Documentation 2023 - Version 1.0. Laxenburg, Austria. Integrated Biospheres Futures.* [Online]
Available at: https://pure.iiasa.ac.at/18996

Johns, T. et al., 2003. Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Climate Dynamics,* Volume 20, pp. 583-612.

Kendon, E. et al., 2019. Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale.. *Nature Communications,* Volume 10.

Köberle, A. et al., 2021. The cost of mitigation revisited. *Nature Climate Change,* Volume 11, pp. 1035-1045.

Krey, V. P. H. et al., 2020. *. MESSAGEix-GLOBIOM Documentation – 2020 release. Technical Report,* s.l.: International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria,.

Maraun, D., 2016. Bias correcting climate change simulations - a critical review.. *Current Climate Change Reports,* Volume 2, p. 211–220.

Mcguffie, K. & Henderson-Sellers, A., 2014. In: *The Climate Modelling Primer Fourth Edition ISBN: 978-1-119-94336-5.* s.l.:John Wiley and Co.

McKinsey & Company, 2022. *The net-zer0 transition – What it would cost, what it could bring.* [Online]
Available at:
https://www.mckinsey.com/~/media/mckinsey/business%20functions/sustainability/our%20insights/the%20net%20zero%20transition%20what%20it%20would%20

Meehl, G. et al., 2009. Decadal Prediction. *Bulletin of the American Meteorological Society,* 90(10), pp. 1467-1486.

Merton, R. C., 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance, 29,* Volume 29, pp. 449-470.

Monasterolo, I., Nieto, M. & Schets, E., 2022. *The Good, the Bad and the Hot House World: Conceptual Underpinnings of the NGFS Scenarios and Suggestions for Improvement. Banco Espana Working Paper No. 2302.* [Online]
Available at: https://ssrn.com/abstract=4211384

Moss, R. et al., 2010. The next generation of scenarios for climate change research and assessment. *Nature,* Volume 463, pp. 747-756.

NGFS, 2021. *NGFS Climate Scenarios for Central Banks and Supervisors. Technical Report,* s.l.: s.n.

Pacanowski, R. & Philander, S., 1981. Parameterization of Vertical Mixing in Numerical Models of Tropical Oceans.. *Journal of Physical Oceanography,* Volume 11, pp. 1443-1451.
Pope, V., Gallani, M., Rowntree, P. & Stratton, R., 2000. The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3.. *Climate Dynamics,* Volume 16, pp. 123-146.

Riahi, K. et al., 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change,* 42(ISSN 0959-3780), pp. 153-168.

Roe, S. et al., 2021. Land-based measures to mitigate climate change: Potential and feasibility by country. *Global Change Biology,* 27(23), pp. 6025-6058.

Sherwood, S. et al., 2020. An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence.. *Reviews od Geophysics,* 58(4).

Smith, C. J. et al., 2018. FAIR v1.3: a simple emissions-based impulse response and carbon cycle model,. *Geoscientific Model Development,* Volume 11, pp. 2273-2297.

Smith, R., 1990. A Scheme for Predicting Layer Clouds and Their Water Content in a General Circulation Model.. *Quarterly Journal of the Royal Meteorological Society,* Volume 116, pp. 435-460.

Stott, P. A. et al., 2010. Detection and attribution of climate change: A regional perspective.. *WIREs Climate Change,* 1(2), p. 192.

Subramanian, A., Gundersen, T. & Adams, T. I., 2018. Modeling and Simulation of Energy Systems: A Review.. *Processes,* 6(12), p. 238.

Vasicek, O. A., 1984. *Credit Valuation. KMV Corporation..* [Online]
Available at: http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/c181fb77ee99d464c125757a00505078/$FILE/Credit_Valua t ion.pdf

Visbeck, M., Marshall, J., Haine, T. & Spall, M., 1997. Specification of eddy transfer coefficients in coarse-resolution ocean circulation models. *Journal of Physical Oceanography,* Volume 27, p. 381–402.

Way, R., Ives, M., Mealy, P. & Farmer, J., 2022. *Empirically grounded technology forecasts and the energy transition, INET, University of Oxford.* [Online]
Available at: https://oms-inet.files.svdcdn.com/production/files/energy_transition_paper-INET-working-paper.pdf

Weyant, J., 2017. Some contributions of integrated assessment models of global climate change. *Review of Environmental Economics and Policy, Association of Environmental and Resource Economists,* 11(1), pp. 115-137.

Wood, R., Keen, A., Mitchell, J. & Gregory, J., 1999. Changing spatial structure of the thermohaline circulation in response to atmospheric CO2 forcing in a climate model. *Nature,* Volume 399, pp. 572-575.
Wright, D., 1997. A new eddy mixing parameterization in an ocean general circulation model.. *International WOCE Newsletter,* Volume 26, pp. 27-29.

[redacted] University of Bristol | IMPERIAL | UNIVERSITY OF LEEDS | UNIVERSITY OF OXFORD | University of Reading | CATAPULT Satellite Applications | The Alan Turing Institute | UKRI Science and Technology Facilities Council