ADVANCE Model Diagnostics Exercise – Study Protocol

Final version, 12 December 2014

1. Motivation and Scope

The ADVANCE Model Diagnostics Exercise aims to characterize, compare and classify the behavior of models for climate policy analysis. The experimental setup is dedicated to generate model output that can be used to estimate a set of diagnostic indicators of model response to carbon pricing policies. The ultimate goal is to better understand differences in model behavior, enable fingerprinting of model responses, and classify models along their fingerprints. The feasibility of this approach has been demonstrated in the AMPERE project (http://ampere-project.eu; see Kriegler et al., 2015, *Diagnostic indicators for integrated assessment models of climate policy*, Technological Forecasting and Social Change 90: 45-61) and further explored in the PIAMDDI project (see Wilkerson et al., 2015, *Comparison of integrated assessment models: Carbon price impacts on U.S. energy*, Energy Policy 76: 18–31). Kriegler et al. (2015) established a set of four diagnostic indicators that allows model fingerprinting and classification based on the type of diagnostic experiment proposed here. The ADVANCE project (http://www.fp7-advance.eu) is continuing this work by offering a platform to collect the results from individual energy-economy and integrated assessment modeling teams. The ADVANCE Model Diagnostics Database offers easy access to the diagnostic indicators, their comparison across participating models, and thus allows teams to assess how their model is situated in the space of available models.

The ADVANCE Model Diagnostics Exercise is coordinated by the Potsdam Institute for Climate Impact Research (PIK; Contact: Elmar Kriegler) and the International Institute for Applied Systems Analysis (IIASA; Contact: Volker Krey).

2. Participation

2.1. Who can participate?

Every team with a national, regional, or global energy-economy or integrated assessment model is invited to participate in the diagnostic analysis proposed here. The diagnostic experiments can be run by every climate policy model with a time horizon until 2050 or longer, and which is able to analyze carbon pricing policies. Minimum requirement is the submission of five core diagnostic runs (one baseline and four carbon pricing policies) in a standardized output reporting template (see below).

2.3. How to participate?

The first step is to register your model at the ADVANCE Model Diagnostics Database hosted by IIASA[1]. Model registration requires providing some information about model version, structure, and regional resolution via a registration form (see "About" page of the web-database for details). Once your model is registered, the second step is to run the diagnostic experiment set out in this protocol. As a minimum requirement, five core diagnostic scenarios need to be reported in full. It is recommended to also report five additional scenarios of high relevance for the diagnostic analysis, but this is not a prerequisite for participation. The protocol includes five additional optional scenarios which teams may choose to report if they consider them relevant for an extended diagnostic analysis of their model.

[1] https://tntcat.iiasa.ac.at/ADVANCEWP1DB

Please make sure to implement all specifications in the protocol as requested to ensure comparability of results across models. Baseline and carbon policy scenarios must be derived from the model version that was registered. If your model was updated, and you want to resubmit updated diagnostics, you can do so by registering the new model version separately. The submission of diagnostic scenarios from different versions of the same model will allow investigating the changes in model response due to model updates, and thus help to relate published findings from different stages of model development to each other. Model output needs to be reported in a standardized format, based on the IAMC output reporting template. The template can be downloaded via the web-database page. The completed output report needs to be submitted to the ADVANCE Model Diagnostics Database. Several checks on the fulfillment of minimum requirements, the use of standardized naming and reporting conventions, and consistency with the experimental setup of the diagnostic analysis will be performed at the time of submission. A successful submission establishes the participation of the model in the diagnostic analysis.

The ADVANCE diagnostic database will be open for submission of diagnostic scenarios for a longer period of time. However, a large-scale comparative analysis of diagnostic model output is scheduled to commence in Spring 2015. All teams interested in participating in this analysis need to submit their diagnostic scenarios by **11 March 2015**.

Further questions on the study, the terms of participation and the model implementation of diagnostic scenarios can be directed by email to Elmar Kriegler (kriegler@pik-potsdam.de) and Volker Krey (krey@iiasa.ac.at).

2.4 What are the benefits of participation?

Participants will gain access to the diagnostic database at the time of submission of diagnostic model runs. Data access will allow participating teams to compare their model results and model diagnostics with other models using a Graphical User Interface.

2.5 What are the terms of participation?

With the submission of diagnostic scenarios, modelling teams grant the ADVANCE study on model diagnostics the permission to use and publish their scenario data in the context of the analysis. Modeling teams retain the right to update or withdraw their scenarios at any point prior to submission of a paper summarizing the diagnostic analysis (but not after submission). The diagnostics analysis will be shared with modeling teams at all stages of the analysis. Analysts are required to accommodate comments of modeling teams concerning the presentation and interpretation of their models. Modeling teams are invited to contribute to the analysis beyond the submission of results, which would entitle them to an authorship on the publication of results. Modeling teams retain the right to use and publish their own submitted scenario data individually. The diagnostic database as used for the analysis will be made publicly available at the time of publication of the diagnostic study.

3. The Scenario Setup

3.1. Overview

The diagnostic scenario setup is comprised of a total of sixteen scenarios that are grouped into five mandatory scenarios, six recommended scenarios and seven optional scenarios. In general, the diagnostic scenarios are not expected to reproduce current observations or policy settings. They are constructed with the sole purpose of allowing the academic community to conduct model diagnostics. They are explicitly not intended to provide policy analysis.

Mandatory scenarios: They include a "no policy" baseline run, and four carbon price scenarios. The carbon price scenarios have two different shapes (constant and exponentially growing) and two different levels (USD 30 and 80 per ton CO_2 in the year 2040). This core set is based on the scenarios used in the diagnostic analysis conducted by the AMPERE project. It is sufficient to establish the four diagnostic indicators developed in the AMPERE study and to assess the robustness of resulting model fingerprints across different shapes and levels of assumed carbon price trajectories.

Recommended scenarios: They include three scenarios of price shocks in the year 2040, jumping from zero carbon pricing (in the baseline) to price levels obtained by the core scenarios in 2040. This set of scenarios adds important information for new diagnostic indicators characterizing response times to price shocks, path dependency (different past, same future setting) and system inertia (same past, different future setting). The fourth scenario describes a linear increase of carbon prices enabling an analysis to what extent the widely used Hoteling assumptions of exponentially increasing carbon prices affects the shape of the emissions response and the technology deployment schedules. Finally, the set includes two scenarios with a quantity constraint on cumulative carbon emissions. This allows to compare the impact of carbon prices on emissions (and other variables) with the reverse impact of an emissions constraint on carbon prices (and other variables).

Optional scenarios: This set of scenarios is relevant for particular model types (e.g. models with a time horizon until 2100, models with foresight etc.) and can provide an extended set of model diagnostics. They include a reference (policy) baseline for those models that include climate policies and/or different assumptions about emissions drivers (GDP and population) in their default baseline, and would like to report this baseline in addition to the mandatory no policy baseline. The set also includes two additional carbon price shock scenarios and two additional emissions constraint scenarios with larger CO2 emissions budget to augment the sensitivity analysis of price shock and carbon budget scenarios. A further scenario describes an anticipated jump in carbon prices (rather than an unanticipated shock) which can allow to identify the effect of anticipation in models with foresight. Finally, the set includes a carbon price scenario that switches from constant to exponentially increasing carbon prices, thus allowing the study of inertia in a continuous setting without price shocks.

3.2. General specifications

The following general specifications hold for the diagnostic scenario setup:

- **No-climate-policy baseline (DIAG-Base):** We use a no-climate-policy baseline for the diagnostic exercise with a zero (shadow) price of greenhouse gas emissions. This implies that all existing climate policies implemented in the baseline should be removed if feasible. The 'no-climate-policy baseline' is not expected to match the observed greenhouse gas emissions until 2013, since the observed quantities are already affected by current climate policies or the expectations thereof in some regions. It may be difficult for some models, particularly for national and regional models, to remove all existing climate policies in their baseline. At a minimum, any explicit carbon pricing in the baseline should be removed, and an effort should be made to remove other policies that are explicitly dedicated climate policies penalizing carbon emissions (such as standards on the carbon intensity of fuels) to the extent possible.
- **Model default baseline (DIAG-Base-Def; optional):** In addition, modeling teams may submit their default model baseline if it deviates significantly from the no-policy baseline during some time period. With climate policies being enacted in a number

of world regions to date, models have increasingly included these climate policies in their reference baseline as a means to better capture the existing policy environment and to match the observed greenhouse gas emissions. Due to the recommended harmonization of population and economic growth assumptions in the no policy baseline of global models (see below), the default baseline may also be different from the no-policy baseline in terms of emissions drivers. If the no-climate-policy and default baselines are identical for a given model, only the former should be submitted with a comment that it is also used as a default baseline by the modeling team.

- **Baseline Harmonization:** It is recommended that global modeling teams (but not necessarily national modeling teams) harmonize their GDP and population assumptions in the no-policy baseline to the SSP2 marker scenario[2]. Baseline harmonization is, however, not required in order to keep the barrier for participation as low as possible. Baseline harmonization will enhance the comparability between response patterns of models that treat population and GDP exogenously or semi-exogenously (i.e. GDP is derived from a limited set of model parameters such as labour productivity). GDP and population harmonization can be of more limited value for models that would need to make a number of adjustments to economic activity in individual sectors and energy intensity improvements to match a prescribed GDP trajectory. Modeling teams should evaluate individually whether there is sufficient benefit from harmonizing their population and GDP assumptions, and on this basis make a decision whether to introduce it. This decision should be documented in the submission of scenarios to the IIASA database. Obviously, the no-policy baseline (whether harmonized or not) should serve as the reference point for all primary diagnostic carbon price and carbon budget runs. There is the possibility to submit multiple sets of diagnostic scenarios, e.g. one based on a harmonized no-policy baseline and another one based on alternative baseline assumptions (e.g. the model default baseline with non-harmonized assumptions). See the naming convention for more information on this possibility.

  If modeling teams choose to harmonize population and GDP trajectories in the no-policy baseline they may use their preferred approach for matching the SSP2 population and GDP scenarios on the level of their native model regions. Modeling teams should also use their default method of adapting external PPP GDP projections in their modeling framework.

- **Time horizon:** Carbon price scenarios are specified for the period 2020-2100, but they remain useful for models with shorter time horizon, e.g. to the year 2050. Models with a time horizon shorter than 2100 should adopt the carbon price scenarios until their particular end year. Models with a time horizon extending beyond 2100 should fix the carbon price at the value reached in the year 2100 for

2 The SSP2 population and GDP data can be downloaded at https://secure.iiasa.ac.at/web-apps/ene/SspDb. The SSP database (File SsspDb_country_data_2013-06-12 in the Download tab) provides country level population projections (IIASA-WiC POP, Scenario SSP2-v9-130115: the dataset includes total population, age cohorts, sex and education levels; share of urban population can be taken from NCAR Scenario SSP2-v9-130115) and GDP projections in Power Purchase Parity (PPP) (OECD Env-Growth, Scenario SSP2_v9_130325). Modeling teams should aggregate the country level GDP information to their native model regions in PPP terms before introducing a conversion to Market Exchange Rates (MER) if needed. The database provides historic country level PPP to MER conversion ratios for the year 2005 (File OECD-WB PPP-MER2005_conversion_rates) from which 2005 conversion ratios for the native model regions can be deduced (PPP GDP country sum vs. sum of country GDP converted to MER).

later periods. Carbon budgets are specified for the periods until 2050 and until 2100, and therefore not applicable to models with a shorter time horizon.

- **Regional scale:** The carbon price scenarios can be applied to models on different regional scales, ranging from global to national. The carbon budget scenarios refer to global carbon budgets, and therefore can only be adopted by global models.
- **Where and what flexibility:** The carbon price should be imposed on all regions, and all Kyoto gases represented in the model. 100 yr Global Warming Potentials (GWPs) should be used to convert the carbon price to greenhouse gases prices for non-CO2 greenhouse gases. Likewise, the global carbon budget should be imposed in a way to allow full where and what flexibility, leading to a globally uniform carbon price across regions and sectors. Again, 100 yr GWPs should be used to price non-CO2 GHGs and land use CO2 emissions that are not capped by the budget at the level of the emerging global carbon price. Models should use their default assumption of 100 yr GWPs (SAR, AR4, AR5, …) for the price conversion, and document this assumption in their submission of scenario data.
- **Implementation of carbon price scenarios:** Carbon price scenarios are specified for the period 2020-2100. Since the choice of base year and time steps varies across models, the carbon tax should only be applied from 2020 onwards, and model behavior in earlier years should be fixed to baseline, amounting to the assumption of a zero carbon price (in the case of the no-policy baseline) without anticipation of future climate policy The carbon price for a given model year t is usually assumed to be constant over the length of the time step Δt (either from time t-1 to t or from t-Δt/2 to t+Δt/2, depending on the model). Modelling teams are requested to adjust the carbon price that they apply in the year t to the average carbon price over the constant price period that emerges in the prescribed scenario[3].

  If carbon price levels exceed the domain of applicability of a given model, the modelling team should fix the carbon price at the highest level the model can run in the years after this level was reached, and explain this action in the comments section of the scenario reporting template. However, since fixing the carbon price after some year will severely limit the comparability of model results, such fixing should be avoided where possible.
- **Implementation of carbon budget scenarios:** The carbon budgets in the diagnostic experiment refer to cumulative CO2 emissions from fossil fuel combustion and industry over the period 2011-2100. The calculation of cumulative emissions over a given period should take into account the model discretization, i.e. whether the emissions for a given model year t are assumed to be constant over the length of the time step Δt from time t-1 to t or from t-Δt/2 to t+Δt/2[4].

  As for the carbon price scenarios, model behavior should be fixed to the baseline for model years before 2020. As of 2020, there should be full where and what flexibility

[3] For example, if the carbon price scenario describes a price of \$30 in 2040, increasing at 5% per year since 2020 (i.e. \$11.3 in 2020), the average carbon price over the period 2031-2040 is \$24.3 (calculated as $0.1 * \Sigma_{i=-9,-8,\dots,0}$ \$30*1.05^i) and \$30.3 over the period 2035-2045 (calculated as $0.1 * \Sigma_{i=-4.5,-3.5,\dots,4.5}$ \$30*1.05^i). For the start year 2020, the results of the calculation are as follows: \$1.1 for 2011-2020 (calculated as 0.1 * \$30 * 1.05^(-20)) and \$6.4 for 2015-2025 (calculated as 0.1 * $\Sigma_{i=0.5,1.5,\dots,4.5}$ \$30*1.05^(i-20)).

[4] For example, cumulative emissions over the period 2011-2100 in a model with ten year time steps and 30 $GtCO_2$ in 2010 rising by 10 $GtCO_2$ in each decade are 7200 $GtCO_2$ for constant emissions from t-1 to t (calculated as 10*($\Sigma_{i=1,2,\dots,9}$ (30 $GtCO_2$ + i*10 $GtCO_2$)) or 6750 $GtCO_2$ for constant emissions from t-Δt/2 to t+Δt/2 (calculated as 5*(30 $GtCO_2$ + 120 $GtCO_2$) + 10*($\Sigma_{i=1,2,\dots,8}$ (30 $GtCO_2$+ i*10 $GtCO_2$)).

of emissions reductions leading to a globally harmonized carbon price across all regions and sectors. This represents an idealized policy implementation that is used as a benchmark in most studies. Emerging carbon prices should also be imposed on GHG emissions outside the cap as described above. Models should use their default methodology to implement a carbon budget in their model, including their standard treatment of when flexibility. This can range from fully endogenous carbon pricing in intertemporal optimization models (usually leading to a Hotelling carbon price path) to the iterative adjustment of a pre-scribed carbon tax trajectory or a heuristic approach to convert the budget into a global emissions trajectory (both methods often used by myopic models). Modelling teams should document their method of implementing the carbon budget in their submission of scenario data. Infeasibilities to reach the imposed carbon budget should be reported.

- **Currency:** All carbon tax scenarios are specified in US dollar for the year 2010. Model using a different base year for US dollars should apply an appropriate deflator to convert the carbon prices to the given base year. In particular, models using US 2005 dollars as currency unit should divide the carbon tax values by a uniform deflator of 1.10774 (taken from WG3 AR5). Models using a different currency should convert the carbon tax scenarios using some average market exchange rate of the currency to US dollar in the year 2010. In the case of Euro, the exchange rate varied between 1.22 USD to 1.43 USD per € over the course of 2010. We suggest choosing an exchange rate of 1.30 USD 2010 per € 2010 for the conversion. A similar average exchange rate occurred in the year 2005. For consistency, all price and cost figures from the models should be reported in the same unit, i.e USD 2010, for the submission of the scenario data.

- **Scenario naming convention:** All diagnostic scenario names begin with the identifier DIAG, followed by the individual scenario name. The no-policy baseline is called DIAG-Base, and the primary carbon price/budget scenarios based on the no-policy baseline are called DIAG-[*Identifier carbon price/budget scenario*]. ***This set of the scenarios is the exclusive focus of the ADVANCE open call on model diagnostics to the community.***

  In principle, a secondary set of diagnostic carbon price/budget scenarios based on an alternative baseline assumption can be submitted by extending the proposed naming convention. These scenarios should be named DIAG-[*Identifier carbon price/budget scenario*]-[*Identifier alternative baseline*]. E.g., if the secondary set is based on the model default baseline, they would be named DIAG-[*Identifier carbon price/budget scenario*]-Def, or if based on a no-policy baseline with different GDP and population assumptions (e.g. SSP3 as identifier of the alternative baseline), they would be named DIAG-Base-SSP3 and DIAG-[*Identifier carbon price/budget scenario*]-SSP3. The naming convention for secondary sets of diagnostic runs can be extended to cover variations other than baseline variations. One example is the variation of technology availability, e.g. the limitation of bioenergy use, in the diagnostic scenario set-up. In this case, the diagnostic runs would be named DIAG-Base-LimBio and DIAG-[*Identifier carbon price/budget scenario*]-LimBio. However, we reiterate that such variations of baselines or other scenario characteristics are not the focus of the ADVANCE open call to the community.

4. Definition of individual scenarios

4.1. Mandatory scenarios (see Figure 1; all tax levels in USD 2010 per ton CO_2)

1. Scenario: DIAG-Base

No policy baseline (see above)

2. Scenario: DIAG-C30-const

For t < 2020: Fix to DIAG-Base

For t in [2020, 2100]: Tax(t) = 30 USD

3. Scenario: DIAG-C80-const

For t < 2020: Fix to DIAG-Base

For t in [2020, 2100]: Tax(t) = 80 USD

4. Scenario: DIAG-C30-gr5

For t < 2020: Fix to DIAG-Base

For t in [2020, 2100]: Tax(t) = 30 USD * $1.05^{(t-2040)}$ (USD 30 reached in 2040)

5. Scenario: DIAG-C80-gr5

For t < 2020: Fix to DIAG-Base

For t in [2020, 2100]: Tax(t) = 80 USD * $1.05^{(t-2040)}$ (USD 80 reached in 2040)

4.2. Recommended scenarios (see Figure 1)

6. Scenario: DIAG-C0to30-const

For t < 2040: Fix to DIAG-Base

For t in [2040, 2100]: Tax(t) = 30 USD

7. Scenario: DIAG-C0to30-gr5

For t < 2040: Fix to DIAG-Base

For t in [2040, 2100]: Tax(t) = 30 USD * $1.05^{(t-2040)}$

8. Scenario: DIAG-C0to80-gr5

For t < 2040: Fix to DIAG-Base

For t in [2040, 2100]: Tax(t) = 80 USD * $1.05^{(t-2040)}$

9. Scenario: DIAG-C80-lin

For t < 2020: Fix to DIAG-Base

For t in [2020, 2100]: Tax(t) = 30 USD + 2.5 USD * (t-2020) (USD 80 reached in 2040)

10. Scenario: DIAG-BFC1000 (only for global models with time horizon until 2100)

For t < 2020: Fix to DIAG-Base

Cumulative CO_2 emissions from fossil fuel combustion and industry are limited to 1000 $GtCO_2$ over the period 2011-2100. The budget is chosen to reflect the upper range of budget values for scenarios reaching roughly 450 ppm from the AR5 DB.

11. Scenario: DIAG-BHC1000 (only for global models with a time horizon until at least 2050)

For t < 2020: Fix to DIAG-Base

Cumulative CO_2 emissions from fossil fuel combustion and industry are limited to 1000 $GtCO_2$ over the period 2011-2050. The budget is modeled after the BFC1000 case assuming net carbon neutrality in the second half of the century. If models with time horizon until 2100 choose to run this scenario they should extrapolate the emergent carbon price trajectory that hits the 2011-2050 budget smoothly beyond 2050.

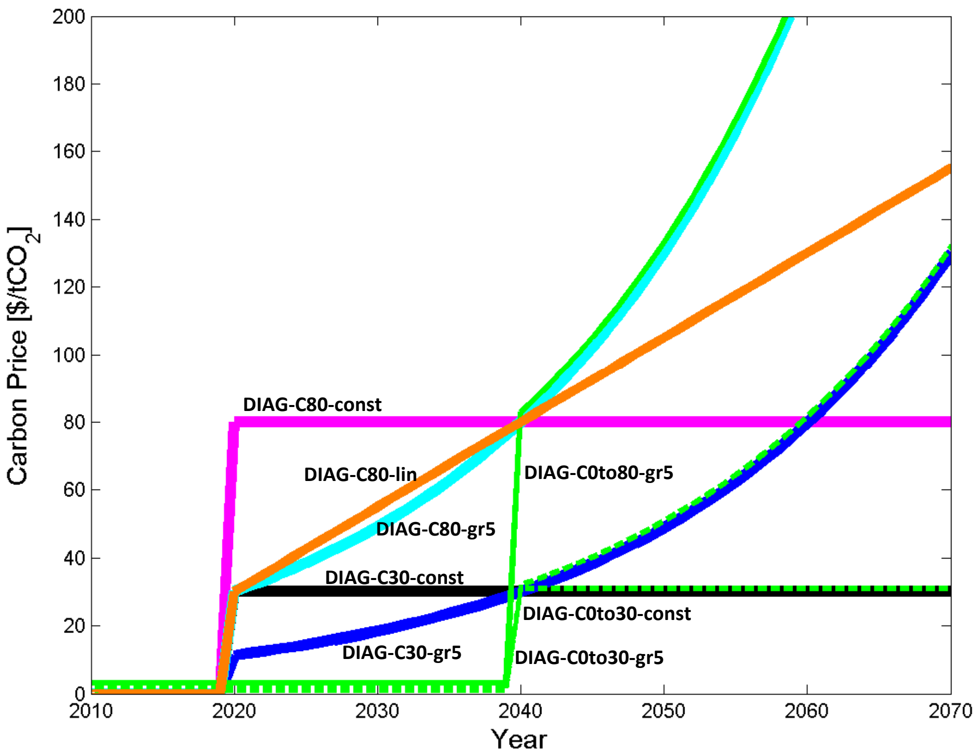

Figure 1: Mandatory and recommended diagnostic scenarios (shown is a close-up of for the period 2010-2070).

4.3. Optional scenarios:

12. Scenario: DIAG-Base-def

Model default baseline if different from no policy baseline (see above)

13. Scenario: DIAG-C0to80-const

For t < 2040: Fix to DIAG-Base

For t in [2040, 2100]: Tax(t) = 80 USD

14. Scenario: DIAG-C0to80-ant (only for models with foresight)

For t < 2020: Fix to DIAG-Base

For t in [2020, 2039]: Tax(t) = 0 USD (but allow anticipation of jump in 2040)

For t in [2040, 2100]: Tax(t) = 80 USD * $1.05^{(t-2040)}$

15. Scenario: DIAG-C0to80-late (only for models with time horizon until 2100)

For t < 2060: Fix to DIAG-Base

For t in [2060, 2100]: Tax(t) = 80 USD * $1.05^{(t-2060)}$

16. Scenario: DIAG-C30-hybrid (only for models with time horizon until 2100)

For t < 2040: Fix to DIAG-C30-const

For t in [2040, 2100]: Tax(t) = 30 USD * $1.05^{(t-2040)}$

17. Scenario: DIAG-BFC1800 (only for global models with time horizon until 2100)

For t < 2020: Fix to DIAG-Base

Cumulative CO2 emissions from fossil fuel combustion and industry are limited to 1800 $GtCO_2$ over the period 2011-2100. The budget is chosen to reflect a mean budget value for scenarios reaching roughly 550 ppm from the AR5 DB.

18. Scenario: DIAG-BHC1200 (only for global models with time horizon until at least 2050)

For t < 2020: Fix to DIAG-Base

Cumulative CO_2 emissions from fossil fuel combustion and industry are limited to 1200 $GtCO_2$ over the period 2011-2050. The budget is modeled after the BFC1800 case assuming a mean value of roughly 600 $GtCO_2$ cumulative emissions in the period 2051-2100 for scenarios reaching 550 ppm in the AR5 DB. If models with time horizon until 2100 choose to run this scenario they should extrapolate the emergent carbon price trajectory that hits the 2011-2050 budget smoothly beyond 2050.