Core Model Proposal #399: Updating the SSP Database (v3.0) (Population, GDP, and Labor Force) and Labor Productivity (KLEM)

Product: Global Change Analysis Model (GCAM)

Institution: Joint Global Change Research Institute (JGCRI)

Authors: Xin Zhao, Pralit Patel, Sheng Di, Jae Edmonds, Stephanie Waldhoff, Allen Fawcett, Christoph Bertram, and Brain O'Neill

Reviewers: Maridee Weber

Date committed: 08/23/2024

IR document number: PNNL-36828

Related sectors: Economy, Agriculture, and Energy

Type of development: Data and Code

Purpose: This Core Model Proposal (CMP) updates the Shared Socioeconomic Pathway (SSP) database to a recent version (v3.0.1; released in 2024) within GCAM. Currently, GCAM relies on socioeconomic drivers, including population, GDP, and labor force projections, from the original SSP database version released in 2013. These projections, provided by independent socioeconomic dynamic models (e.g., multi-dimensional demographic models and macroeconomic models of convergence growth), may need regular updates when (1) near-term observations become available and (2) there are updates and advancements in the socioeconomic modeling. Timely updates of socioeconomic drivers in global economic equilibrium and multisector dynamic modeling will ensure (1) alignment of historical years and near-term projections with observations, enhancing base year calibrations, including calibration parameters and labor productivity, and (2) improvement of long-term projections with updated socioeconomic drivers, which set the scale of the economy. This CMP updates the SSP data (from v2013 to v2024) and also fixes/reconciles historical GDP data sources in GCAM. We investigate the impact of these updates on GCAM projections.

Contents

1. Brief introduction

Data in the Shared Socioeconomic Pathway (SSP) Database[1], such as population, GDP, and labor force projections, are used as socioeconomic drivers in global economic and multisector dynamic models. These projections, provided by independent socioeconomic dynamic models (e.g., multi-dimensional demographic models and macroeconomic models of convergence growth), may be updated[2,3] when (1) near-term observations become available and (2) there are updates and advancements in the socioeconomic modeling.

The socioeconomic drivers set the scales for the projections of sectoral outcomes and the corresponding environmental implications[4,5]. Timely updates of socioeconomic drivers in global economic equilibrium and multisector dynamic modeling will ensure (1) alignment of historical years and near-term projections with observations, enhancing base year calibrations, including calibration parameters and labor productivity, and (2) improvement of long-term projections with updated socioeconomic drivers, which set the scale of the economy.

Currently, GCAM relies on socioeconomic drivers, including population, GDP, and labor force projections, from the initial version of the SSP database released in 2013, which provided projections from 2010 to 2100. Recently, these drivers were updated and released[i] in 2024 with data from 2020 to 2100. Compared with the 2013 version of the data, the updated projections show a higher population and lower GDP by the end of the century, see Figs. S1 & S2.

In this Core Model Proposal (CMP), we update the socioeconomic drivers used in GCAM to ones provided in the most recent version of SSP database (v3.0.1; released in 2024). We also fixes/reconciles historical GDP data sources in GCAM. We investigate the impact of these updates on GCAM projections.

[i] See https://www.iamconsortium.org/event/joint-iamc-iconics-webinar-updated-ssp-socioeconomic-projections/

2. Description of changes (Methods)

2.1. Historical data updates

In the current GCAM Master branch, different sources were used to construct historical GDPs, including USDA, World Bank, and IMF (for near-term updates in gSSPs). The USDA data was initially in 2010 dollars but was converted to 2015 dollars using only the US GDP deflator, whereas regional deflators should have been used. With the updates, we unified the data source by using only FAOSTAT data, which is currently updated to 2023. Note that Taiwan's data was missing in FAOSTAT, so we supplemented it using external sources, including Taiwan Statistics. FAOSTAT data is updated frequently, and we leverage functions in the gcamfaostat[6] R package to download and process the data, including handling regions requiring regional dissolution or aggregation.

During data processing, we made further adjustments to the GDP data for South America and Taiwan for different reasons. For South America_North, there was a significant dip around 2020, so we smoothed the data using a 15-year average to avoid unreasonable TFP calibration. In Taiwan, adjustments were made to prevent a projected GDP drop later in the century, based on IIASA projections, which serve as an alternative source to OECD data (Fig. S3).

For population data, no updates were made except for incorporating the recent historical data from the SSP updates. The labor force data was also updated accordingly.

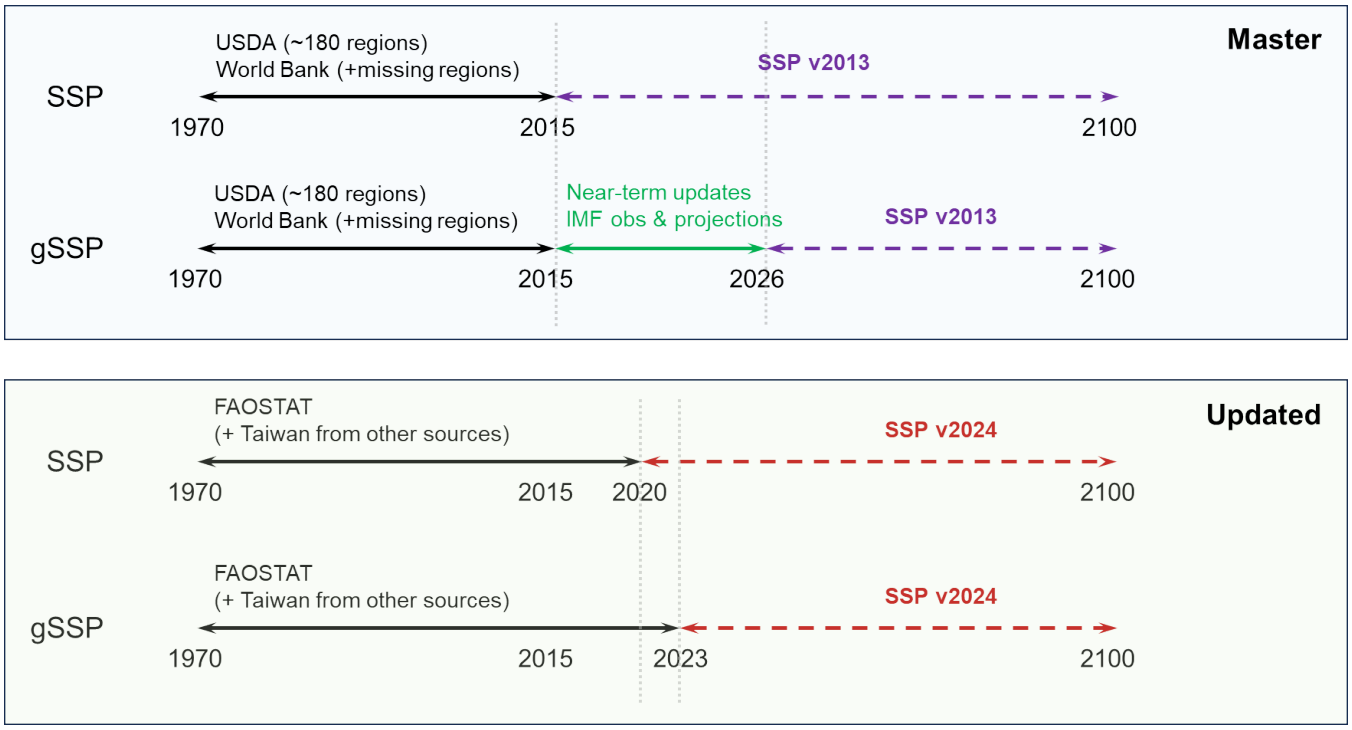

Fig. 1 GDP data source by SSP and gSSP, Master vs. Updated.

2.2. SSP data updates

We are using the 2024 release of the SSP database (v3.0.1), which can be downloaded from the IIASA SSP database (https://data.ece.iiasa.ac.at/ssp). Specifically, we update GDP projections from the OECD ENV-Growth model and population and labor force projections from the IIASA-WiC POP model. For additional information regarding the update, please refer to the IAMC-ICONICS webinar on updated SSP socioeconomic projections (https://www.iamconsortium.org/event/joint-iamc-iconics-webinar-updated-ssp-socioeconomic-projections).

2.3. Overview of key changes in gcamdata

The key data and code changes implemented in GCAM and gcamdata are summarized in **Table 1**. Note that the socioeconomics data folder is restructured and the new PWT data is added (prepared for the KLEAM CMP).

Table 1 key data and code changes made in gcamdata

Data file/R chunk/CPP	Changes made
constants.R	socioeconomics.SSP_DB_BASEYEAR <- 2020 # base year of SSP data base v3.0.1. Other variables specifying the adjustments needed for "ven' and "twn" are also added.
zsocio_data_Maddison_population.R	Deleted; merged into zsocio_L100.Population_downscale_ctry.R
zsocio_L100.Population_downscale_ctry.R	Add new SSP data and process population data from IIASA-WiC POP 2023.
zsocio_L100.GDP_hist.R	Reconcile historical GDP data sources
zsocio_L102.GDP.R	Add new SSP data and process GDP data from OECD ENV-Growth 2023.
zsocio_L180.GDP_macro.R	Update labor force related processing using the new SSP data.
zsocio_L280.GDP_macro.R	More labor force related processing and cleaning.
zsocio_xml_socioeconomics_macro.R zsocio_xml_socioeconomics_SSP.R	XML exporting updates
zsocio_L101.Population.R zaglu_L202.an_input.R zaglu_L162.ag_prodchange_R_C_Y_GLU_irr.R	Minor fixes/improvements
zwater_xml_electricity_water.R zwater_L232.water_demand_manufacturing.R zwater_L203.water_td.R zwater_L173.EFW_manufacturing.R zwater_L132.water_demand_manufacturing.R zenergy_L2391.gas_trade_flows.R zenergy_L1327.paper.R	Fix the 1990 Central Asia calibration issue and gas trade related issues.

2.4. Fix energy calibration issues

While this proposal seemingly does not affect historical energy balances, it is the case that some of the energy balance processing does utilize historical GDP to do sectoral downscaling. And updating the GDP happened to cause a calibration failure in 1990 due to a bug in the energy / water balancing. This is a long standing bug and is being addressed in this CMP to avoid the calibration failure.

The bug arising from subtle definitional issues surrounding industrial water usage and its connection to energy for water (EFW). Ideally, industrial water use and municipal water use would be treated the same. However, the data sets we use have slightly different definitions particularly around accounting for desalinated water usage. For municipal the water volume does include desal, however for industrial it is "self supply" only thus not including desal. Note that in GCAM, given that desal is a subsector within the water_td_ ind / muni sector we must use the convention that the water flow is "self supply" + desal. The energy for water processing correctly differentiates and adjusts for this issue. However, in the level 2 chunks for industrial water use, where we calibrate IO coefficients, we used only the "self supply" water volume. Thus under calculating the intended industrial water volume, and subsequently electricity demands.

To address this issue we simply produce a table from the L173 EFW processing chunk which calculates the total "self supply" + desal and use that in the level 2 industrial water use processing chunks instead of the L132 self supply only. Given municipal water already follows the total including desal convention, no adjustments were needed for that processing.

2.5. Scenarios

There are 32 SPA (shared policy assumption) and macroeconomic (fixed vs. open) scenarios in total (**Table 2**). The current master corresponds to GCAM v7.1 (CMP391). CI321 and CI333 are the pull request folders storing the validation runs.

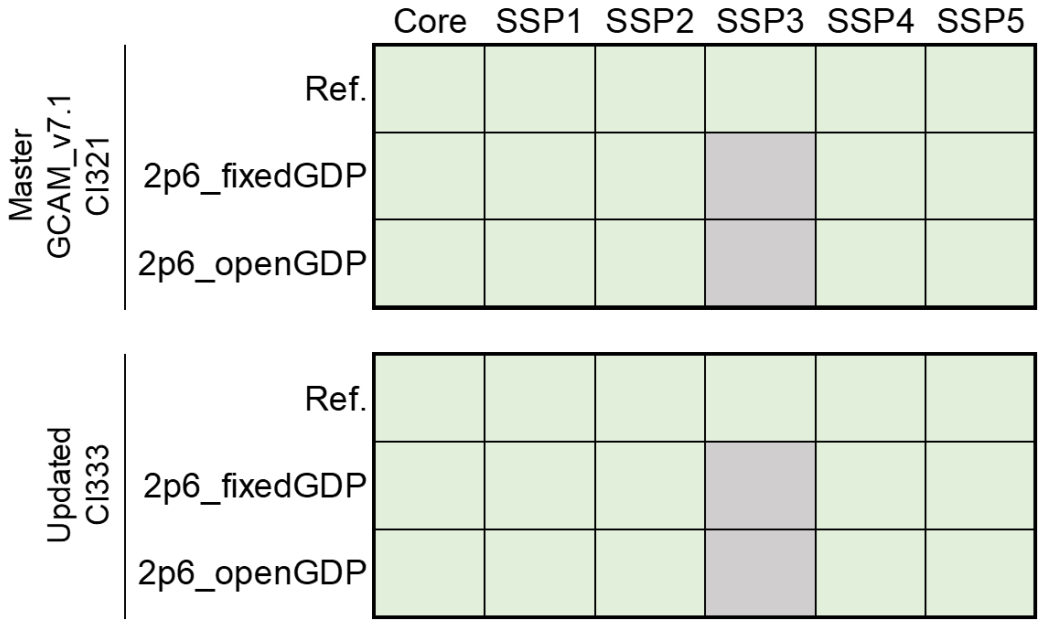

Fig. 2 GCAM validation scenarios

3. GCAM validation runs

In accordance with the GCAM CMP convention, we present GCAM projection results, comparing the Updated (SSP-update) branch with a recent Master branch (CMP-391; GCAM v7.1 release) for reference, RCP 2.6-fixed and RCP 2.6-open scenarios across shared socioeconomic pathways (GCAM core & SSP1-5 assumptions; excluding SSP3-RCP2p6).

3.1. Population and labor force (employment)

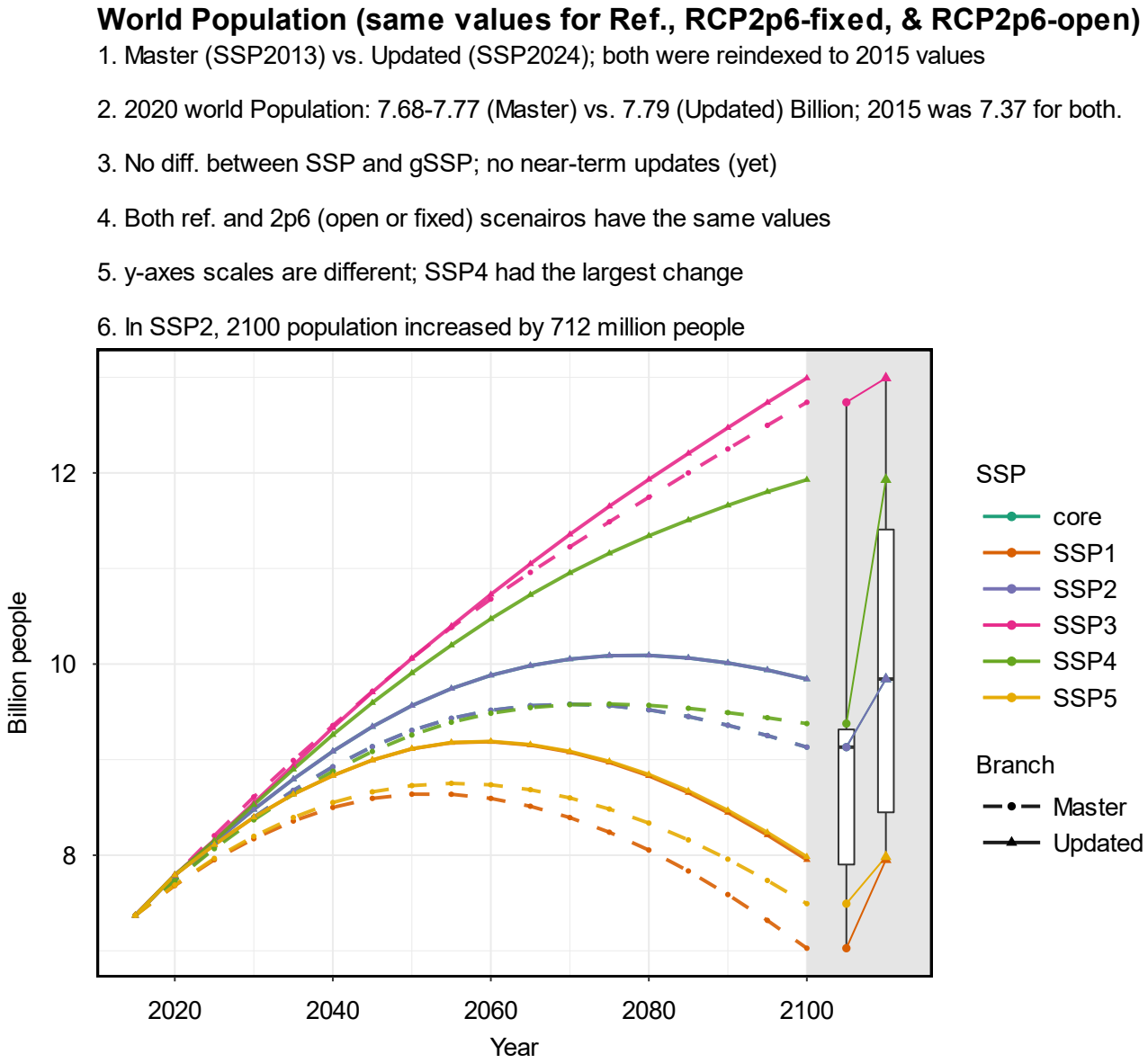

Fig. 3A World population across SPA scenarios. The lines present population projections by 2100 by scenario (SSPs or GCAM core; GCAM-core is identical to SSP2 for population) across branches, Master (dotted lines) vs. Updated (solid lines). The boxplots show distributions of 2100 values across branches Master (left) vs. Updated (right), i.e., including the median values (line), the 1st and 3rd quartiles (boxes),

and the 0-100 percentile ranges (whiskers) of the scenarios. Data of the boxplots are shown (points) and the lines connecting these points across branches show changes due to the updates.

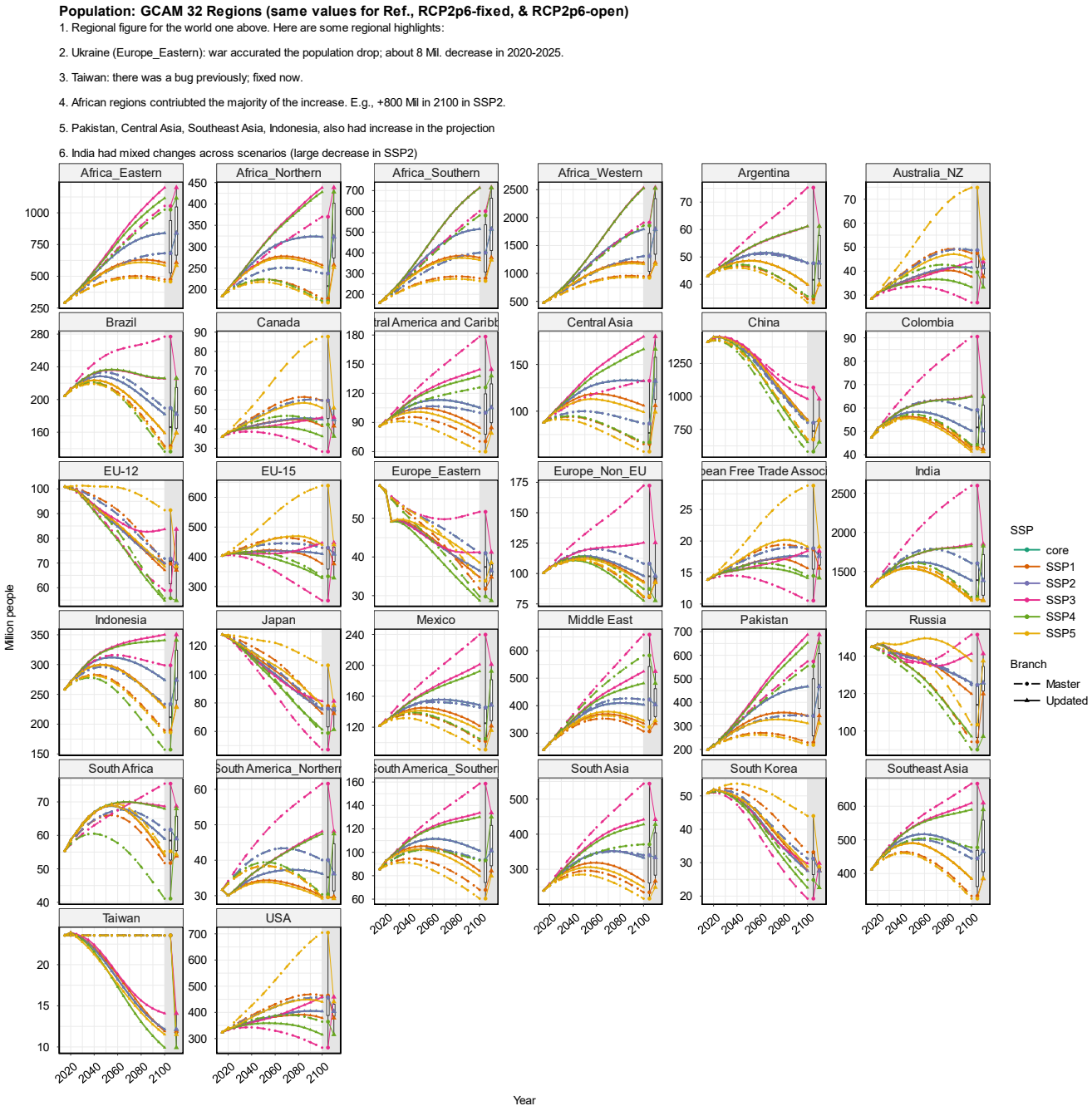

Fig. 3B Regional population projections across SPA scenarios. See Fig. 3A for detailed captions.

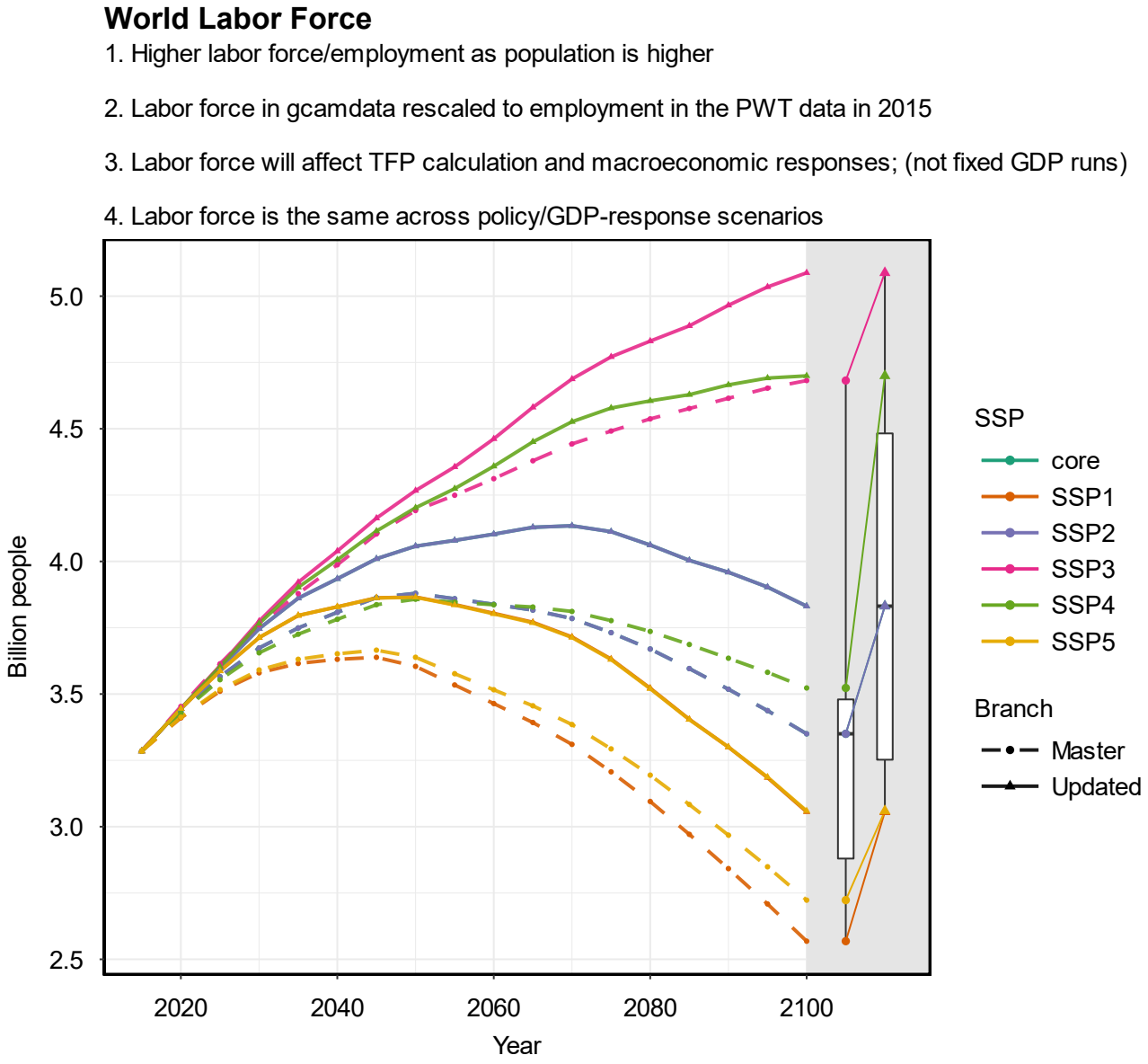

Fig. 3C World labor force (employment) projections across SPA scenarios. See Fig. 3A captions for reading instructions.

3.2. GDP

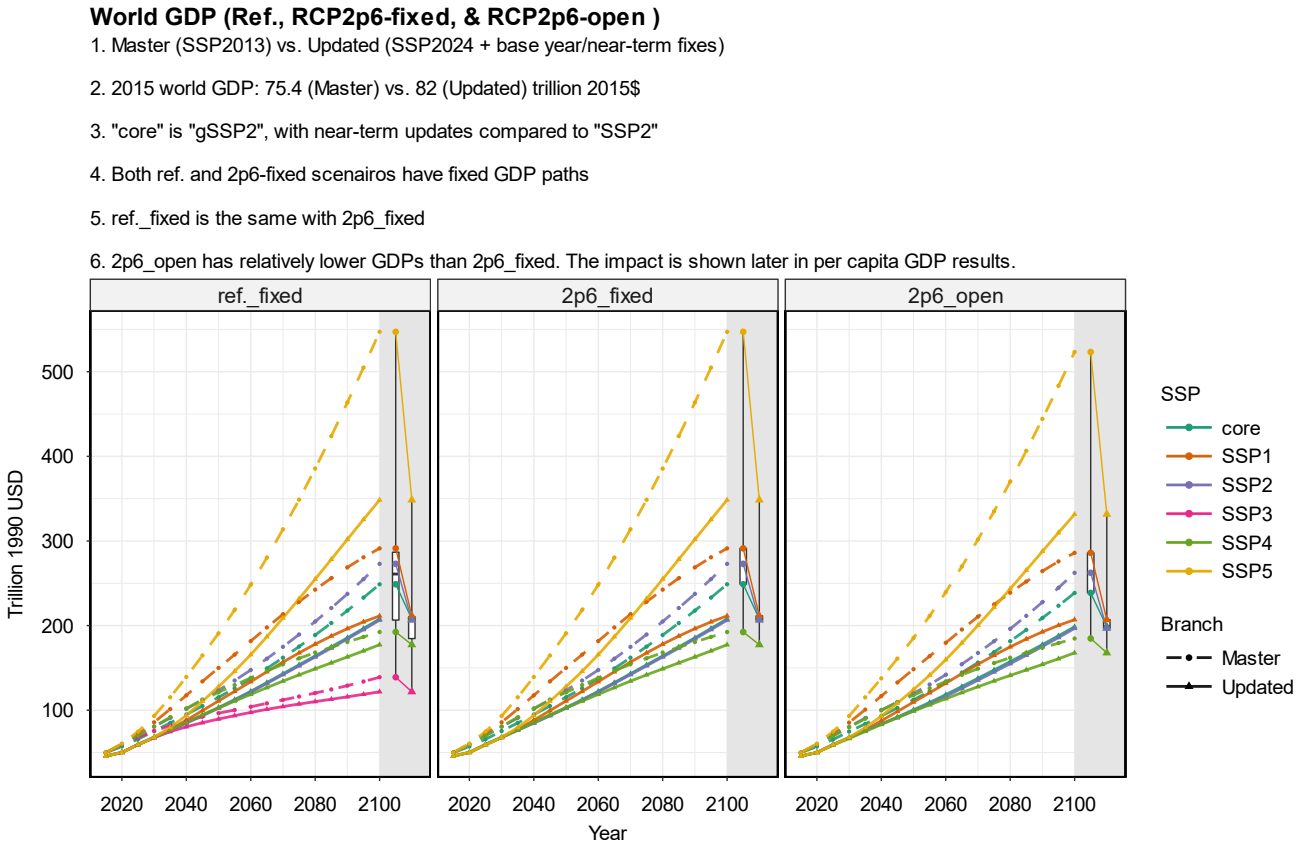

Fig. 4A World GDP across SPA and GDP-response (fixed vs. open) scenarios. The lines present population projections by 2100 by scenario (SSPs or GCAM core; GCAM-core is only slightly different from SSP2 for GDP since 2021-2023 values were updated to observations) across branches, Master (dotted lines) vs. Updated (solid lines). The boxplots show distributions of 2100 values across branches Master (left) vs. Updated (right), i.e., including the median values (line), the 1st and 3rd quartiles (boxes), and the 0-100 percentile ranges (whiskers) of the scenarios. Data of the boxplots are shown (points) and the lines connecting these points across branches show changes due to the updates.

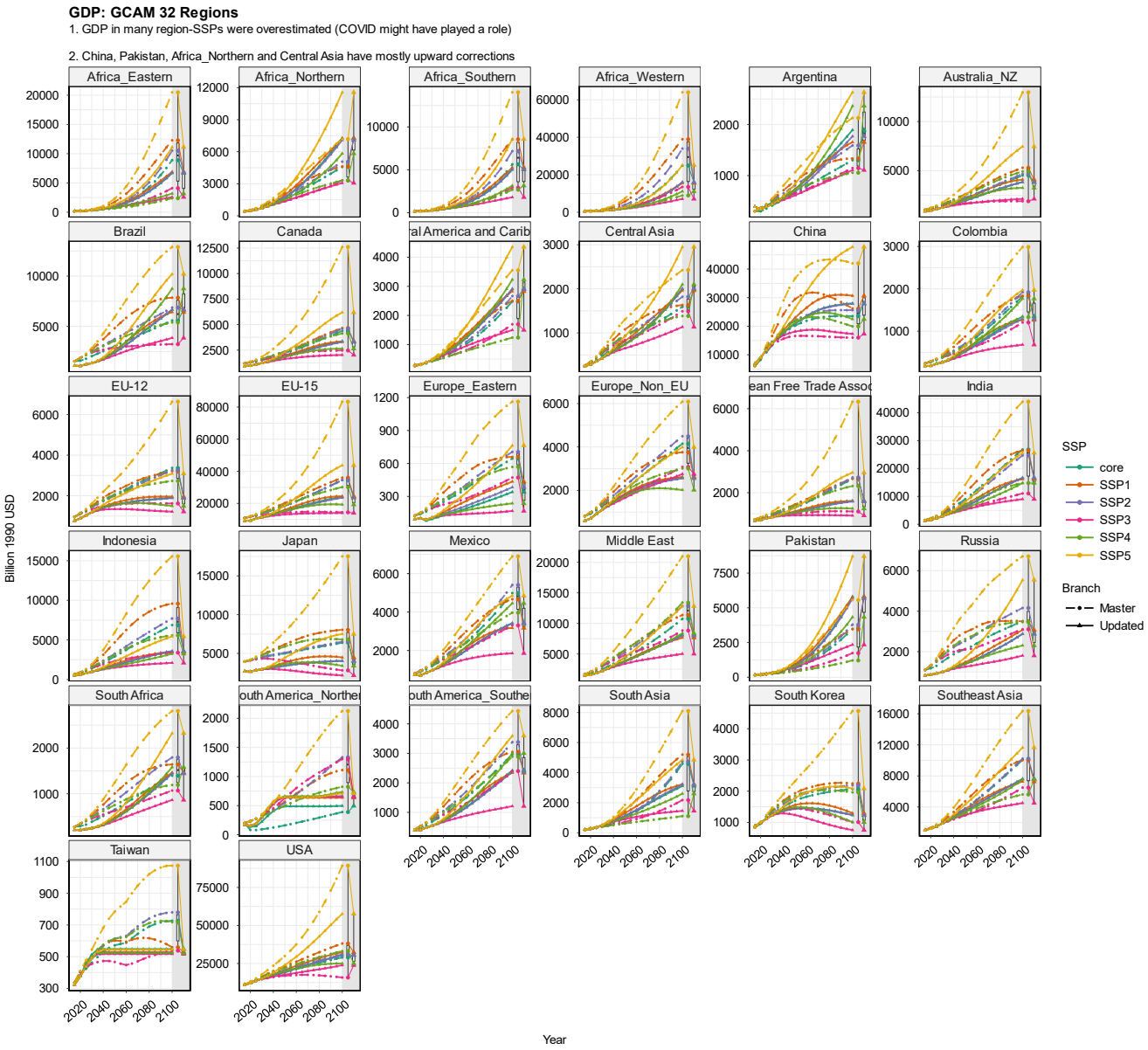

Fig. 4B Regional GDP projections across SPA scenarios. See Fig. 4A for detailed captions.

3.3. Per capita GDP and total factor productivity (TFP)

World per capita GDP (fixed GDP mode)

1. Per capital GDP decreased globally across all SSPs.

2. 2100 per capita GDP decreased by 30% (SSP1) to 7% (SSP3) and ~16% in core.

3. Results here are for ref. or RCP2.6-fixed mode.

4. Both high population and lower GDP projections contributed these changes.

5. Conversely, higher pop and lower per capita GDP led to lower total GDP

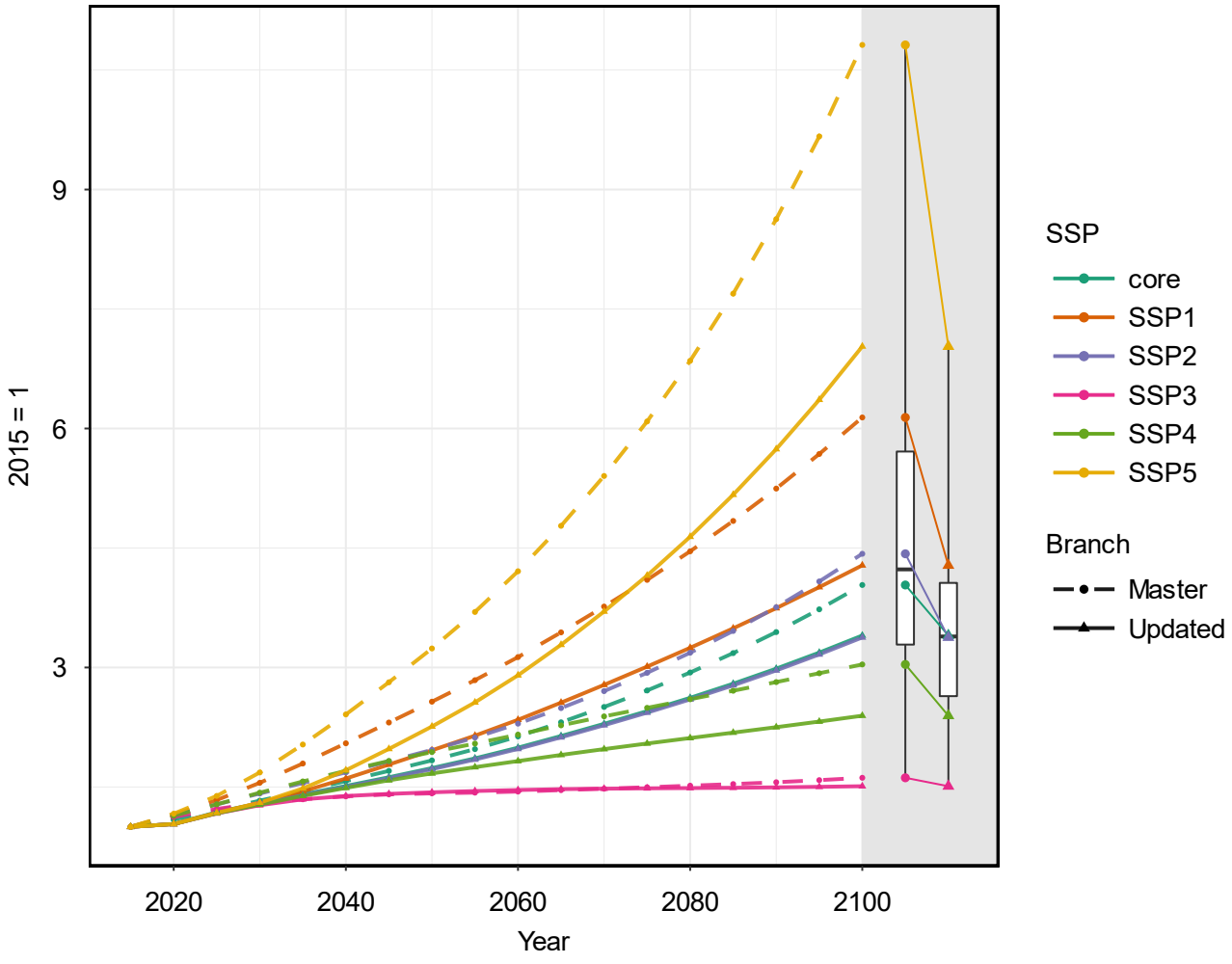

Fig. 5A World per capita GDP across SPA (fixed-GDP mode) scenarios, relative to 2015. The lines present population projections by 2100 by scenario (SSPs or GCAM core; GCAM-core is only slightly different from SSP2 for GDP since 2021-2023 values were updated to observations) across branches, Master (dotted lines) vs. Updated (solid lines). The boxplots show distributions of 2100 values across branches Master (left) vs. Updated (right), i.e., including the median values (line), the 1st and 3rd quartiles (boxes), and the 0-100 percentile ranges (whiskers) of the scenarios. Data of the boxplots are shown (points) and the lines connecting these points across branches show changes due to the updates.

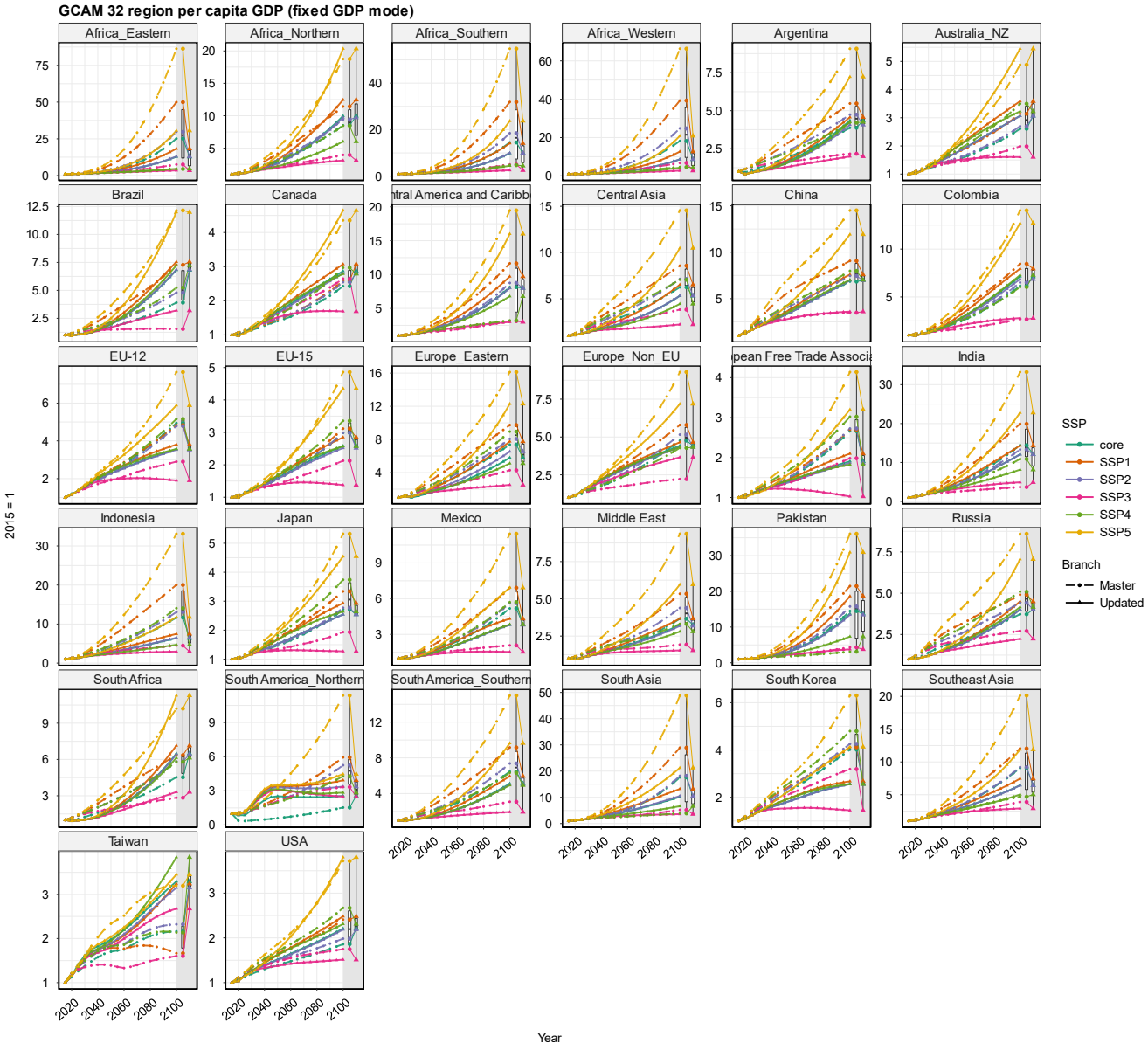

Fig. 5B Regional per capita GDP across SPA (fixed-GDP mode) scenarios, relative to 2015. See Fig. 4A for detailed captions.

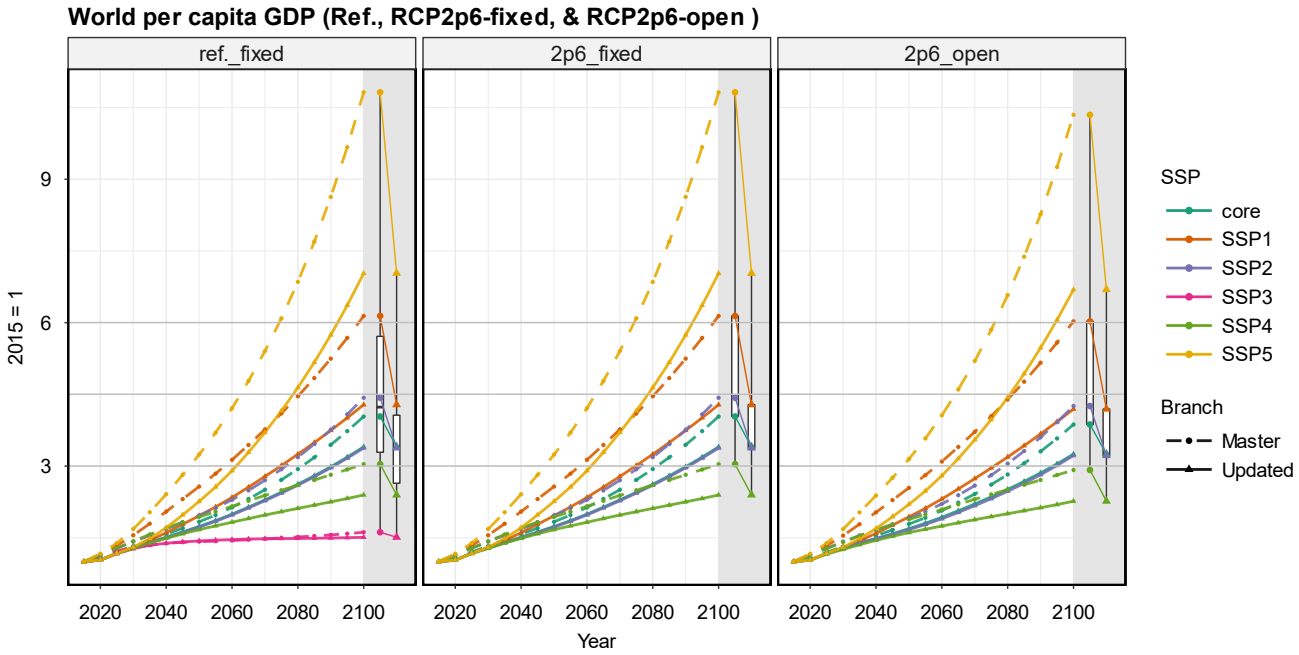

Fig. 5C World per capita GDP across SPA and GDP-response (fixed vs. open) scenarios, relative to 2015. See Fig. 5A for detailed captions.

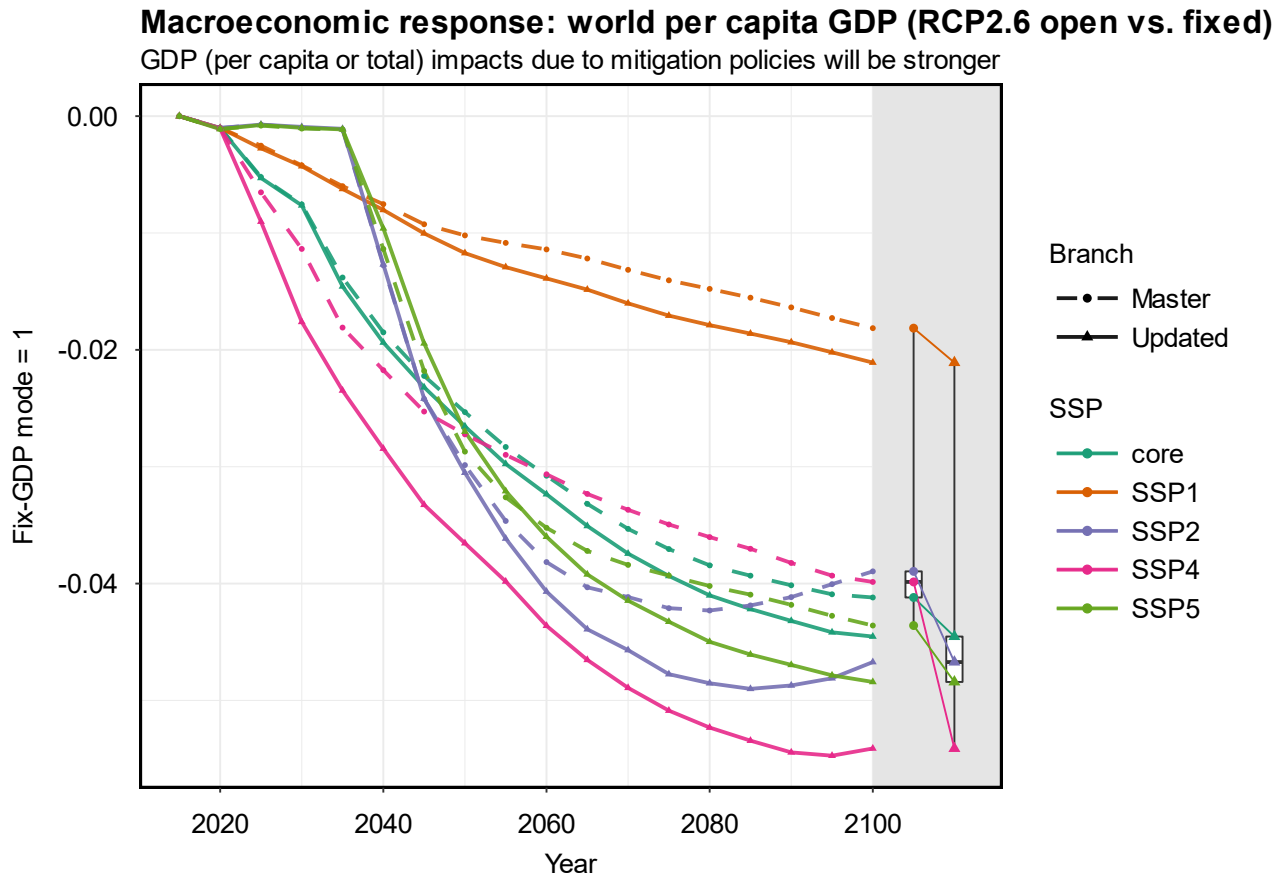

Fig. 5D World per capital GDP impacts due to RCP2p6 mitigation policies (open vs. fixed) across SPA. See Fig. 5A for detailed captions.

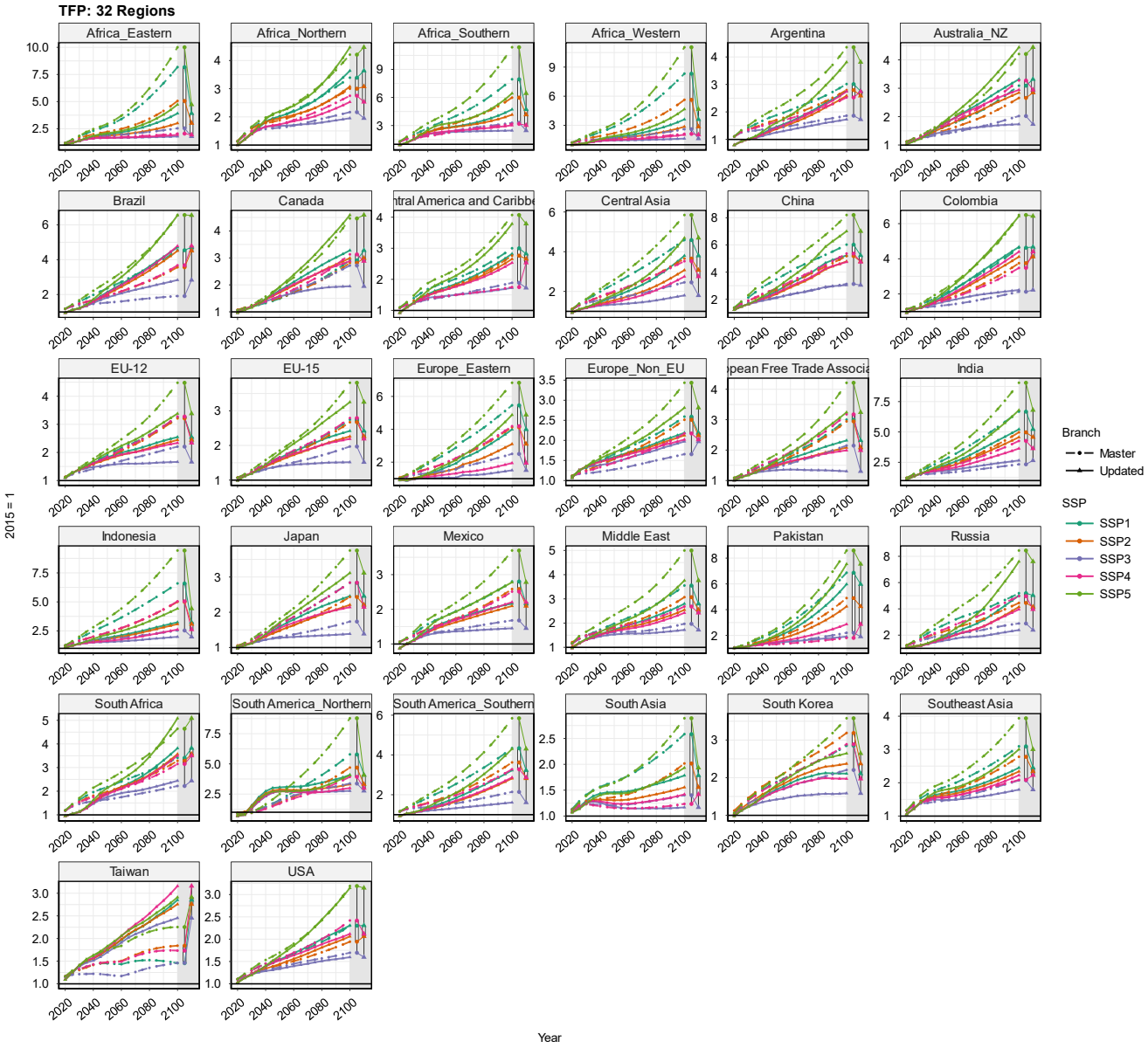

Fig. 5E Calibrated total factor productivity (TFP) by 32 GCAM regions. See Fig. 5A for detailed captions.

The pattern of changes is similar to the changes in per capita GDP shown above (Fig. 5B). Note that the calibrated TFPs are only used in open mode. However, the calibration was based on reference projections, and macroeconomic inputs, such as labor force and savings, also play important roles.

3.4. Radiative Forcing, Temperature, and Carbon Prices

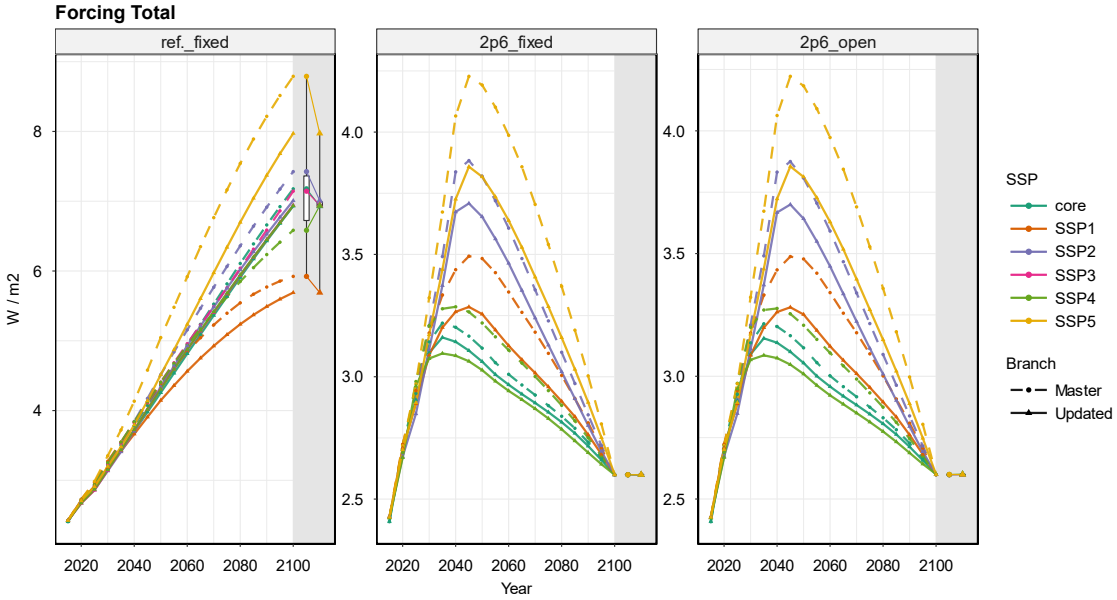

Fig. 6A Total forcing across SPA and GDP-response (fixed vs. open) scenarios.

In reference runs, all scenarios have a lower 2100 forcing except for SSP4. In the RCP 2p6 scenarios, the temporal patterns (e.g., peak forcing time) are also affected. The pattern of these changes is consistent with the changes in emissions (shown later).

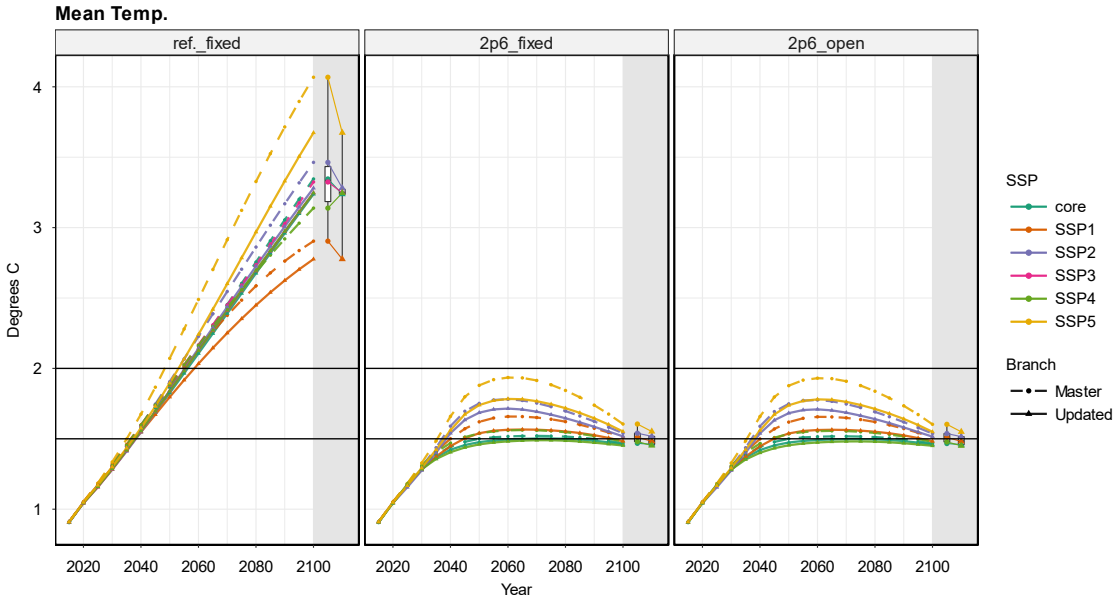

Fig. 6B Global mean temperature change across SPA and GDP-response (fixed vs. open) scenarios.

These results are generally consistent with the forcing results. However, in RCP 2p6 scenarios, 2100 temperature is lower, likely related to other natural climate responses, e.g., terrestrial carbon and ocean heat changes.

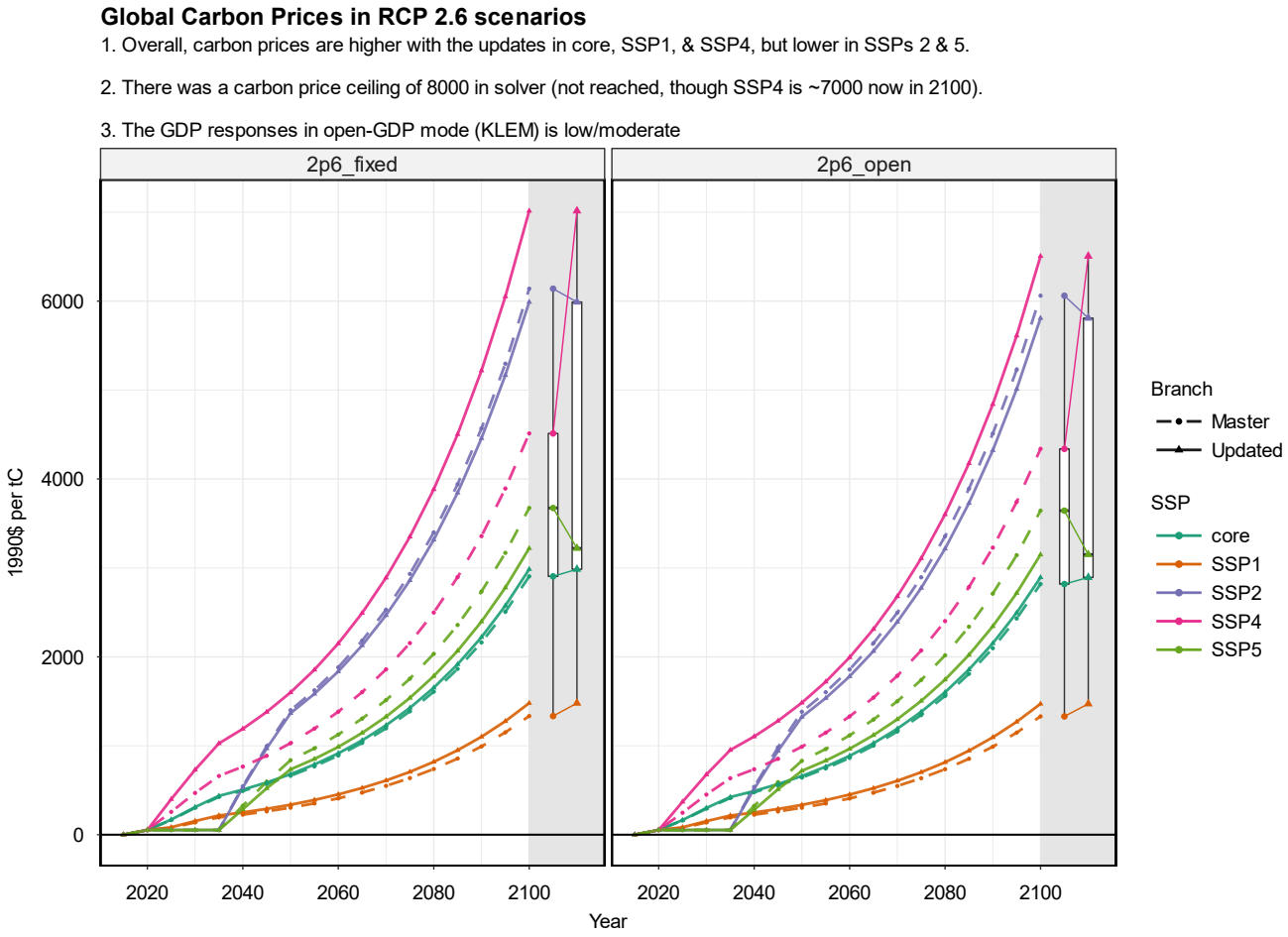

Fig. 6C Global carbon prices across SPA and GDP-response (fixed vs. open) scenarios. There is no SSP3 in this figure, and color codes are different from other figures.

The impacts vary across scenarios due to competing drivers. For instance, when the reference 2100 forcing is lower because of updates (in most scenarios except SSP4), it becomes easier to achieve the 2.6 forcing targets. However, changes in population and lower GDP (due to negative emission technology budget constraints) can make mitigation efforts more challenging and rigid. Generally, the changes in carbon prices are moderate, except for SSP4.

3.5. Emissions

The emissions results can be explained by the changes in energy production, agricultural production, and land use.

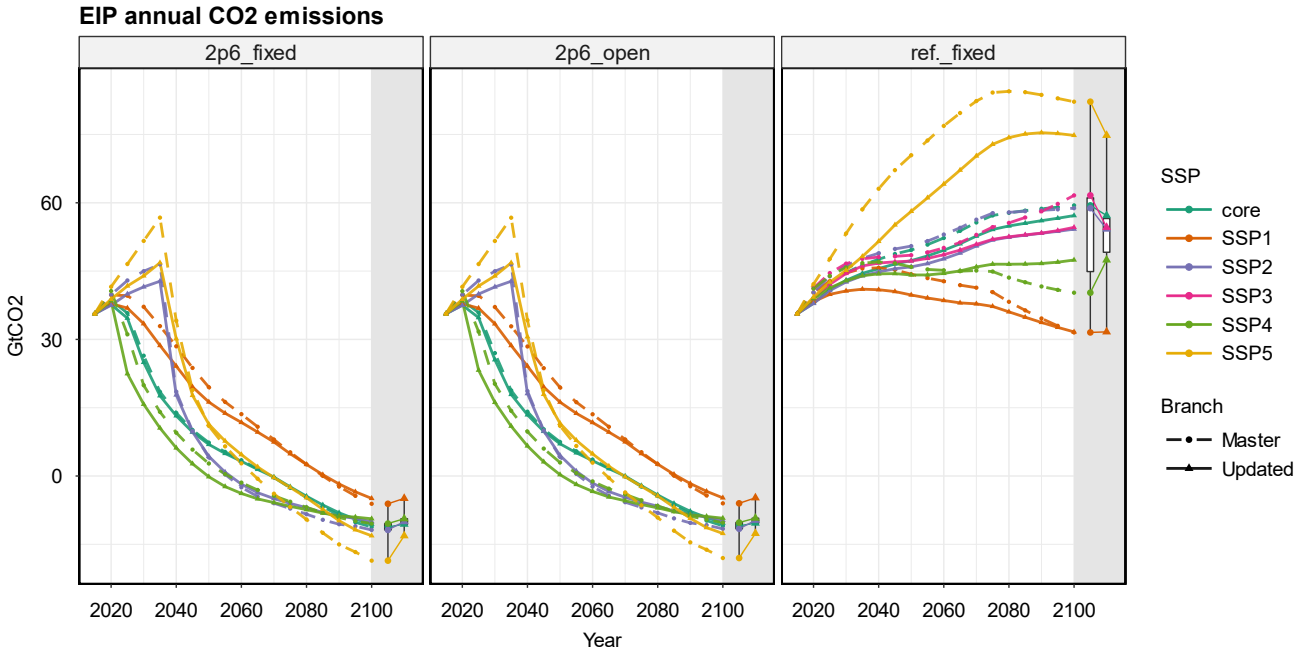

Fig. 7A Global energy and industrial processes (EIP) carbon dioxide emissions across SPA and GDP-response (fixed vs. open) scenarios.

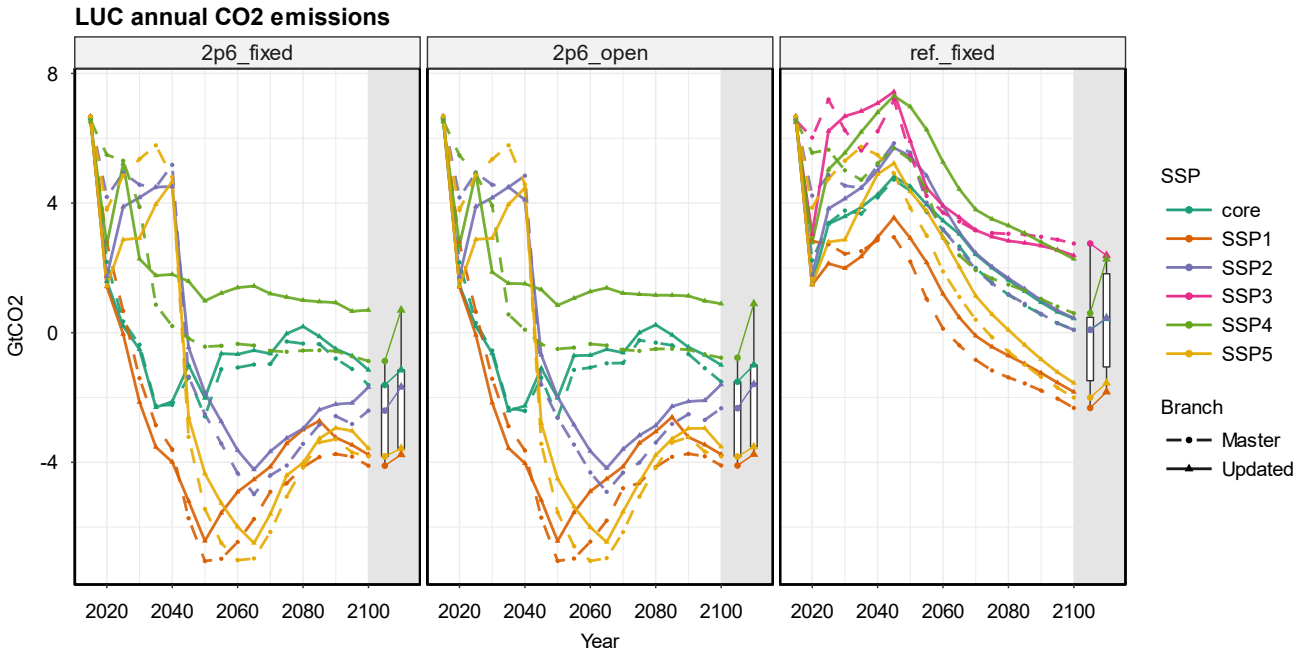

Fig. 7B Global land use change carbon dioxide emissions across SPA and GDP-response (fixed vs. open) scenarios.

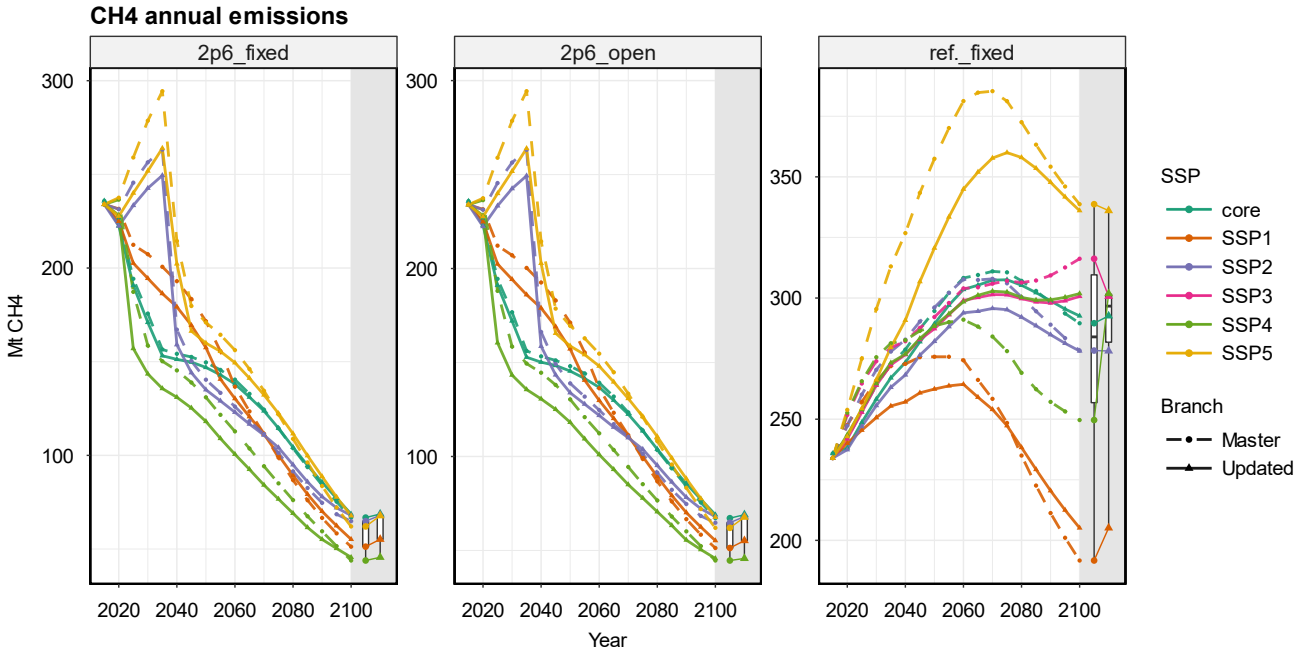

Fig. 7C Global CH4 emissions across SPA and GDP-response (fixed vs. open) scenarios.

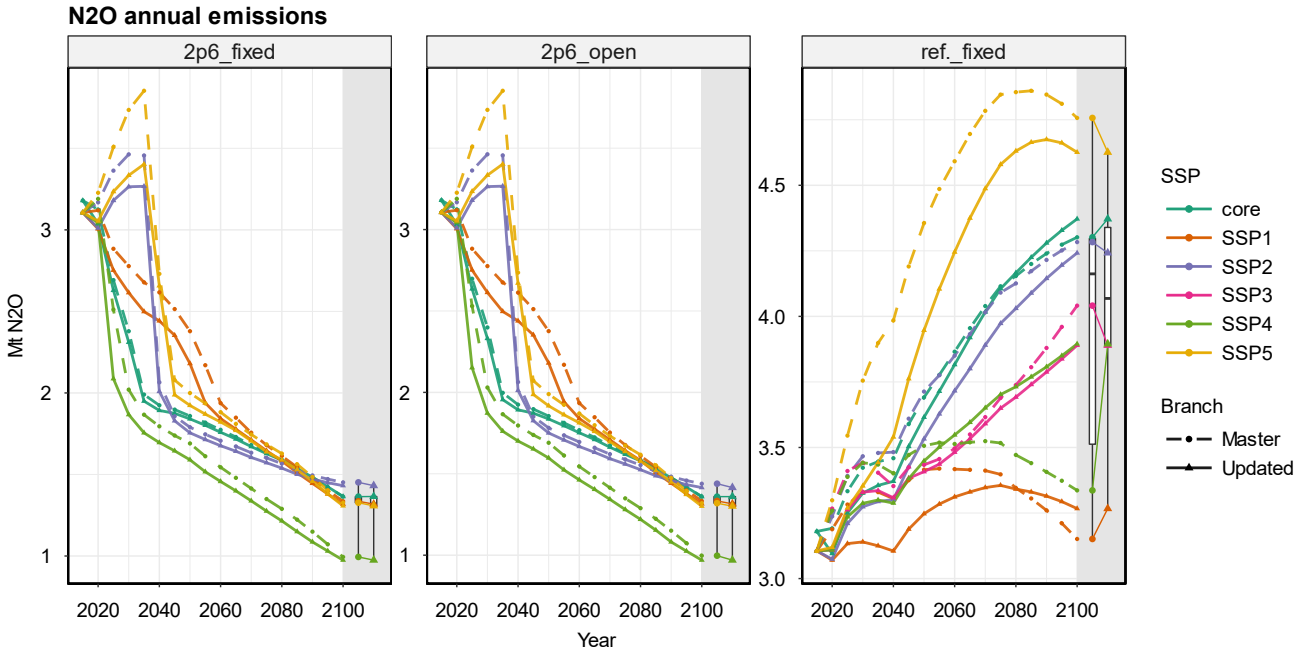

Fig. 7D Global N2O emissions across SPA and GDP-response (fixed vs. open) scenarios.

3.6. Energy

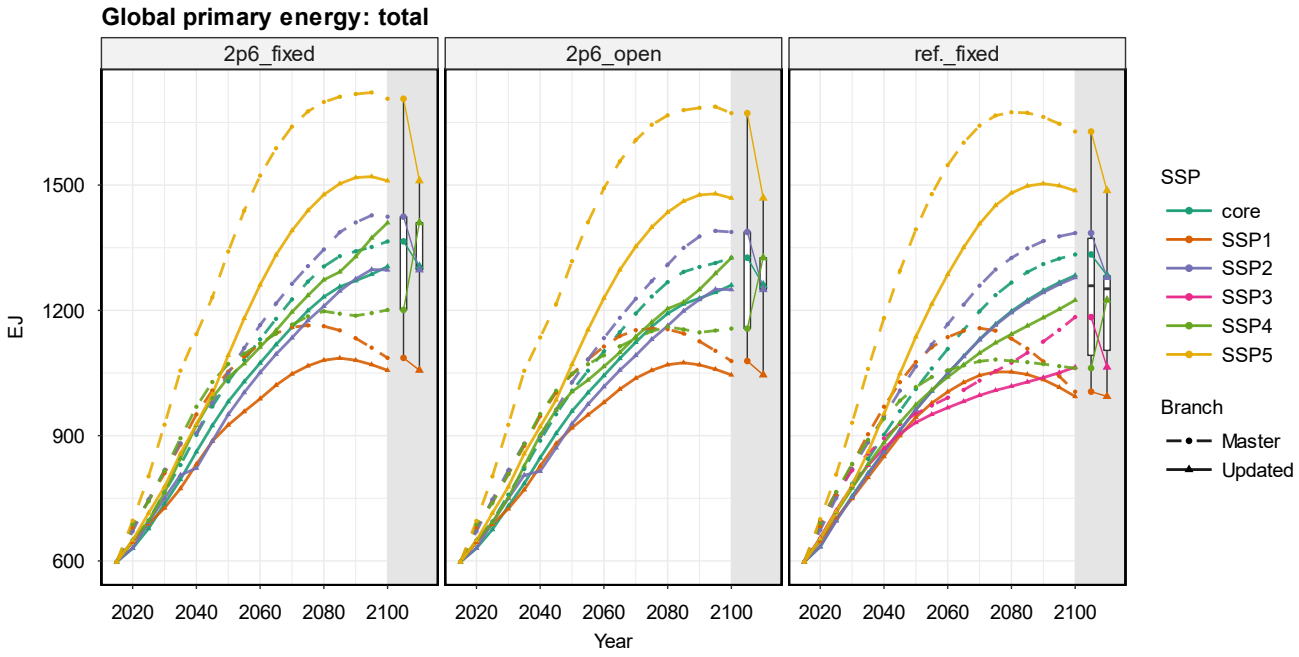

Fig. 8A Global total primary energy across SPA and GDP-response (fixed vs. open) scenarios.

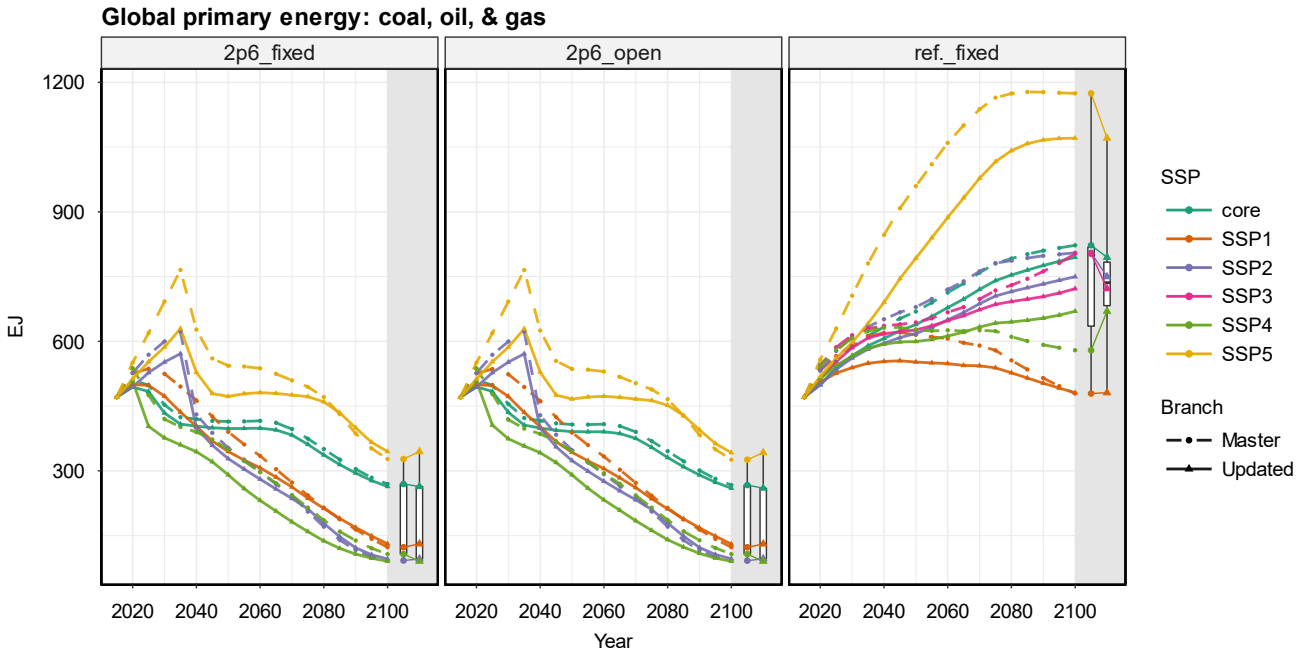

Fig. 8B Global fossil fuels primary energy across SPA and GDP-response (fixed vs. open) scenarios.

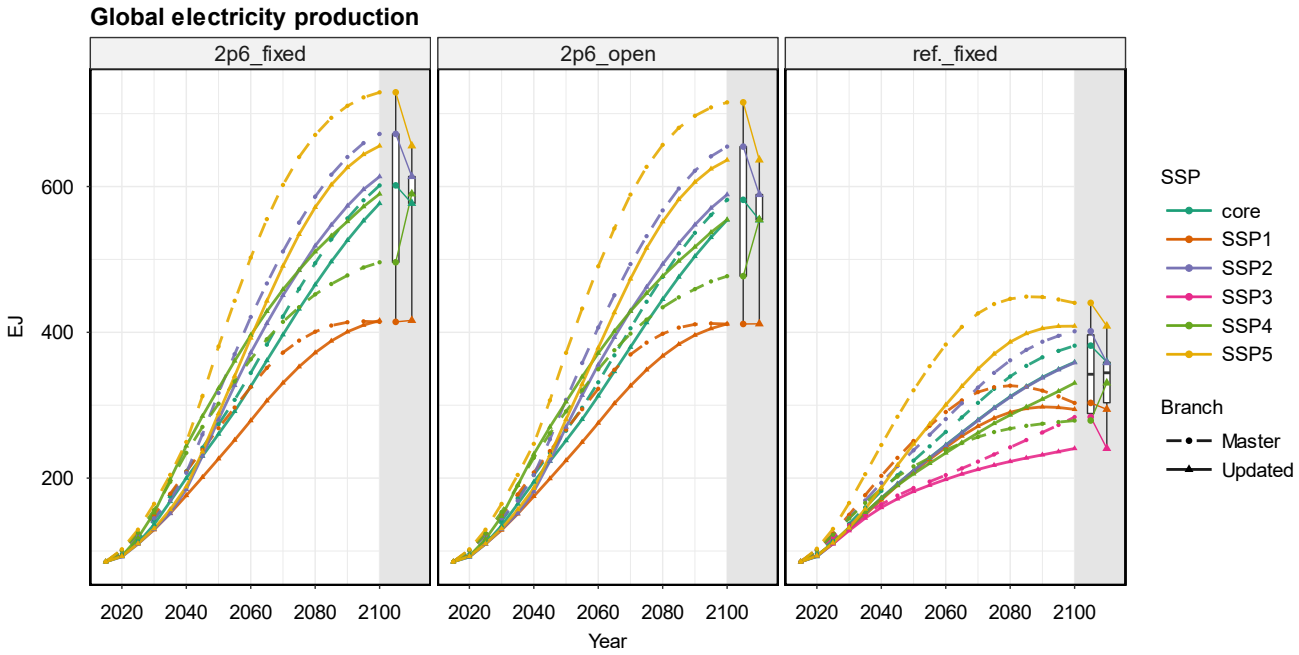

Fig. 8C Global electricity production across SPA and GDP-response (fixed vs. open) scenarios.

3.7. Agriculture and land use

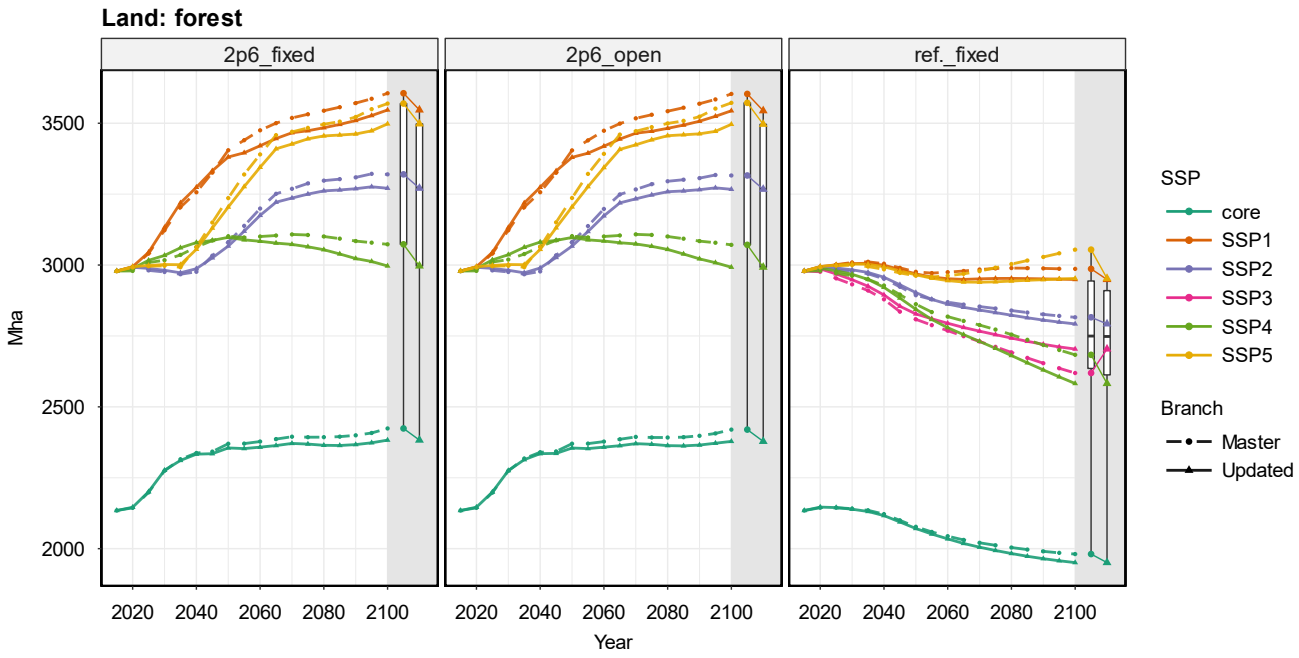

Fig. 9A Global forest area across SPA and GDP-response (fixed vs. open) scenarios. Note that only "unprotected" forest land is shown and core GCAM has higher forest protection.

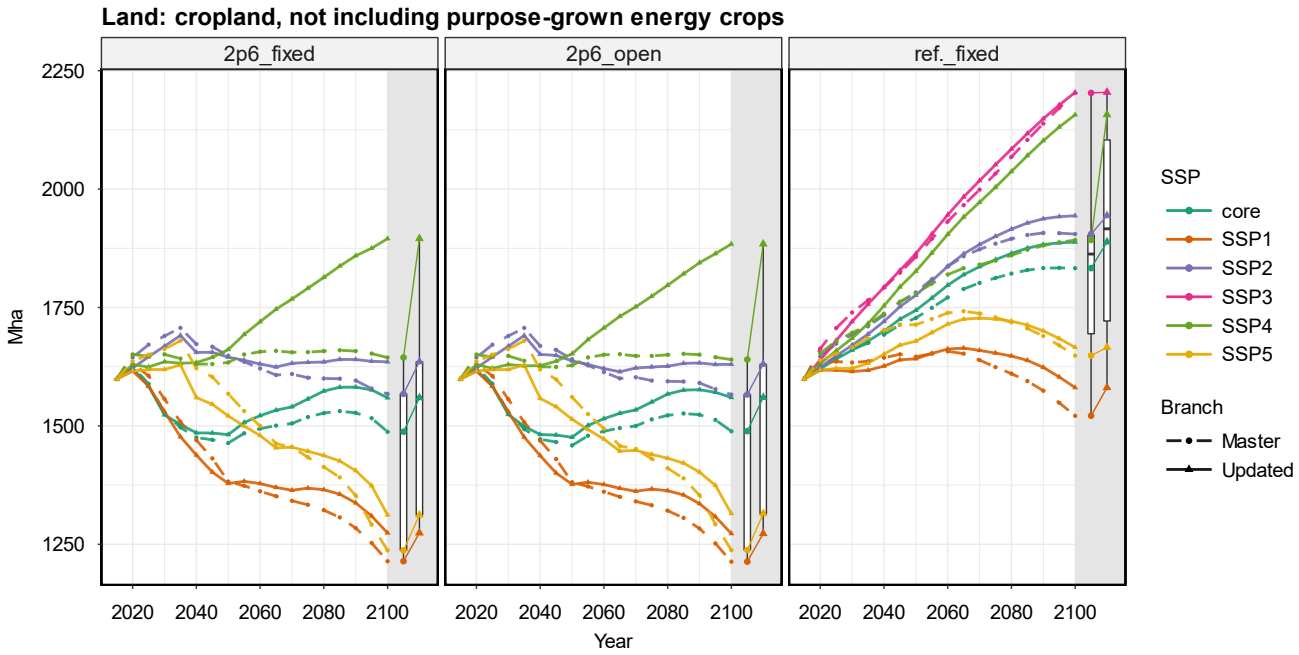

Fig. 9B Global cropland area across SPA and GDP-response (fixed vs. open) scenarios.

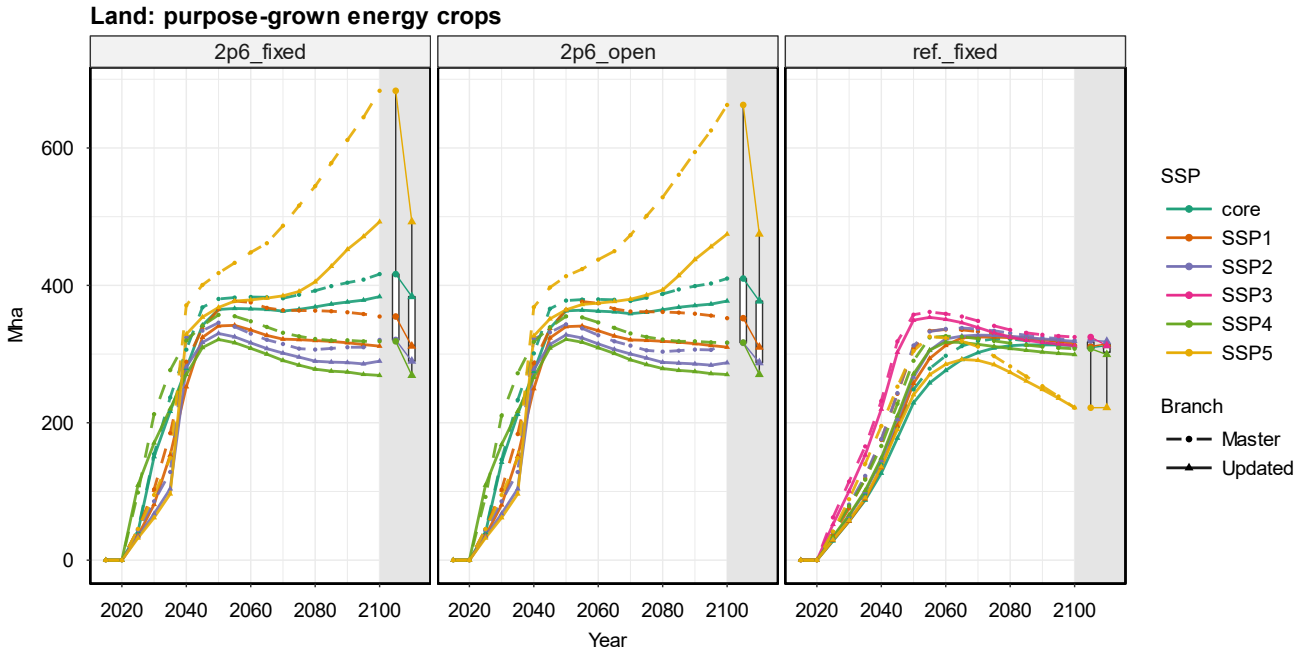

Fig. 9C Global purpose-grown energy cropland area across SPA and GDP-response (fixed vs. open) scenarios.

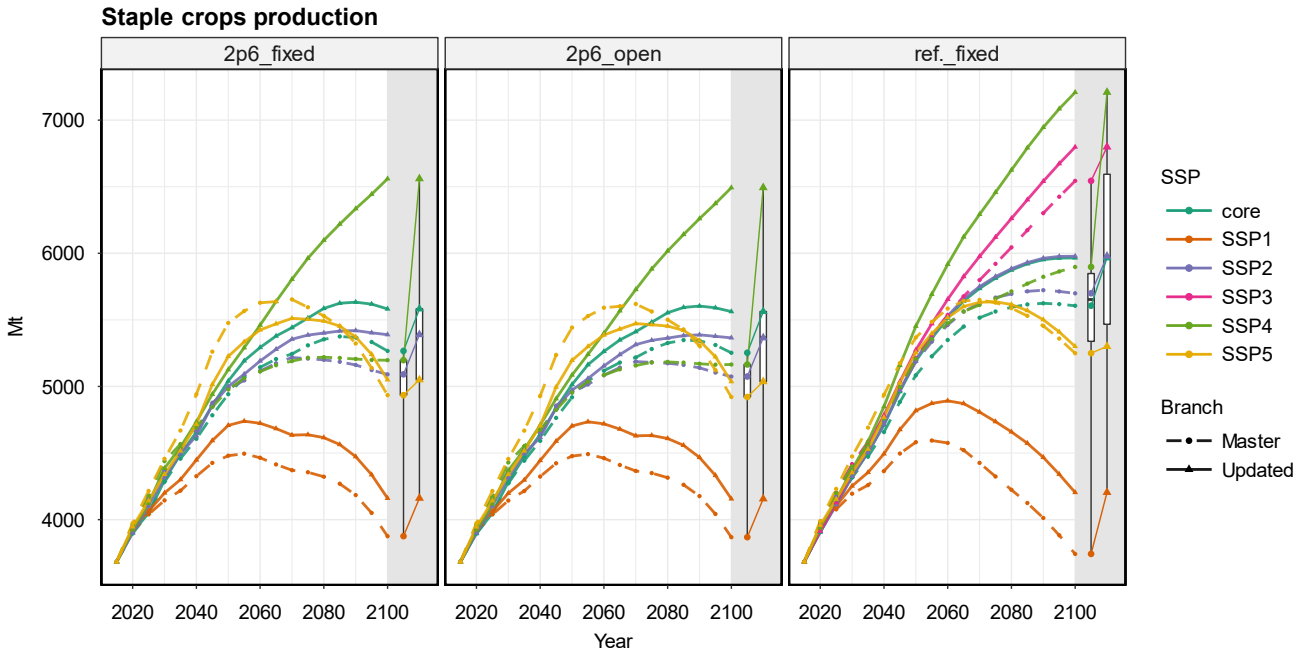

Fig. 9D Global staple crop production across SPA and GDP-response (fixed vs. open) scenarios.

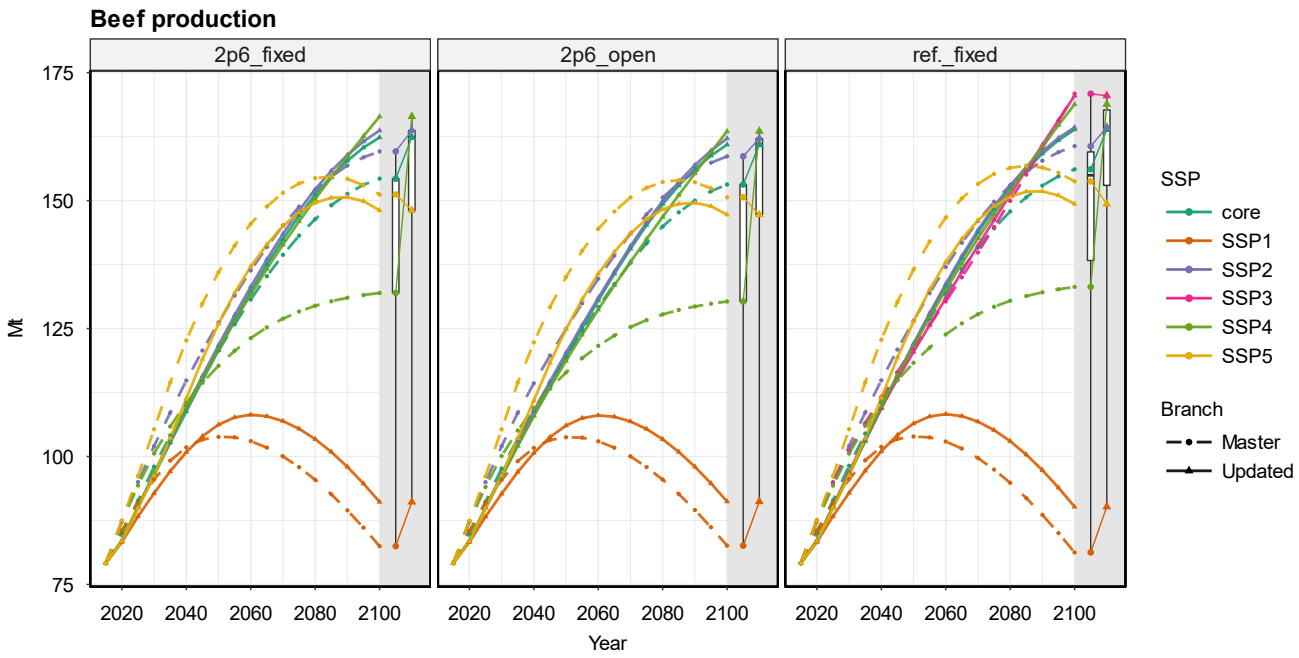

Fig. 9E Global beef production across SPA and GDP-response (fixed vs. open) scenarios.

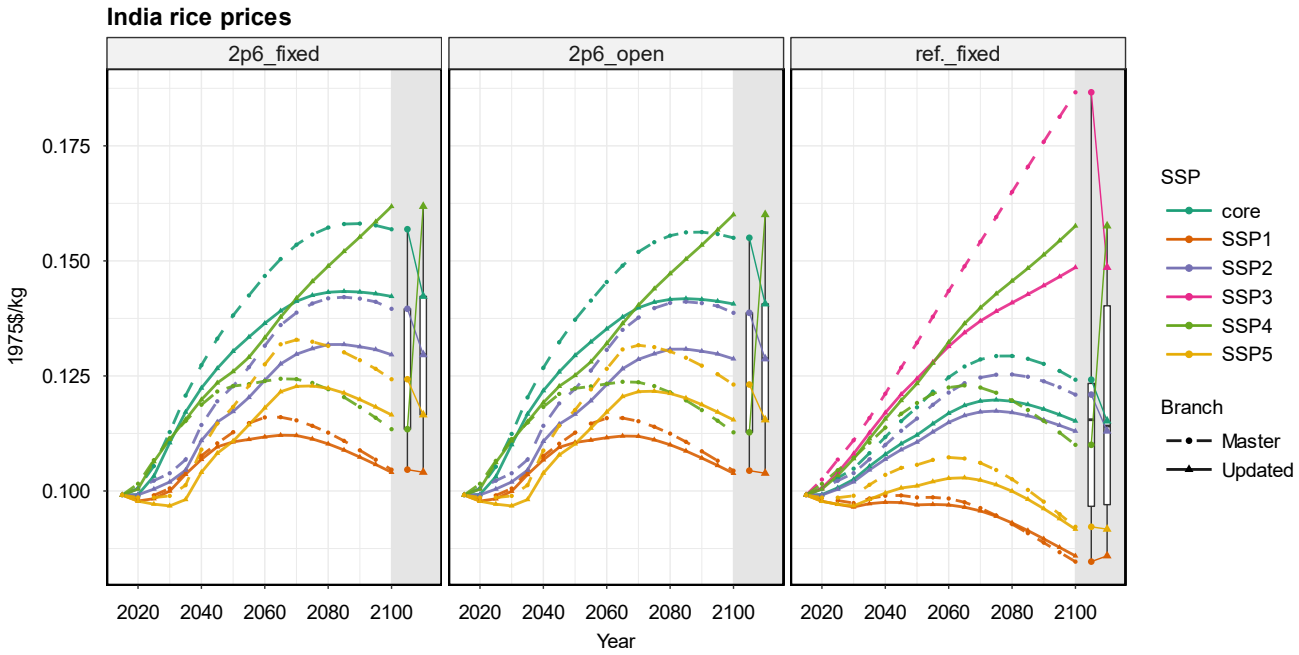

Fig. 9F India rice prices across SPA and GDP-response (fixed vs. open) scenarios.

3.8. Water withdrawals

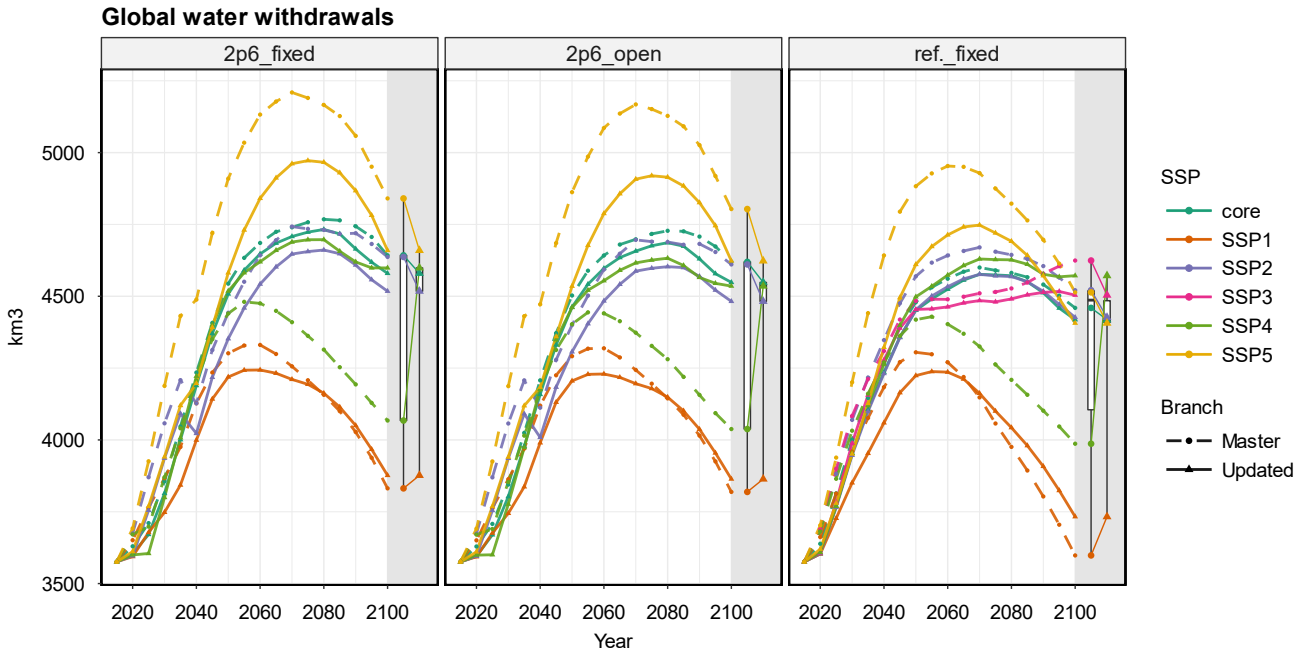

Fig. 10 Global water withdrawals across SPA and GDP-response (fixed vs. open) scenarios.

4. Summary

With updates in the SSP database, in general, population will be higher, and GDP will be lower than previously projected. The updates have had several notable impacts on projections. Overall, they indicate higher food consumption levels but lower energy demand. This was because food demand is less elastic (regarding both price and income) compared to energy demand. While emissions are reduced in the reference scenario, the more rigid systems for adaptation in the RCP 2.6 scenario result in higher GDP impacts. That is, it was easier to reach the climate target due to the lower reference emissions but more difficult and expansive to adapt with more population & lower income.

Both global and regional projections are particularly sensitive to socioeconomic drivers, with distinct differences in economic and environmental responses to population growth and GDP changes. Although GDP feedback is endogenized, the total factor productivity was recalibrated so that the impact on macroeconomic responses were moderate. This underscores the importance of GDP feedback in shaping future trends and raises questions about the potential implications of endogenizing population dynamics.

Finally, it is important for future studies to better document the source and version of socioeconomic driver data used, especially for near-term adjustments (since SSP database starts from 2020), to ensure transparency and accuracy in projections.

5. Supplementary information

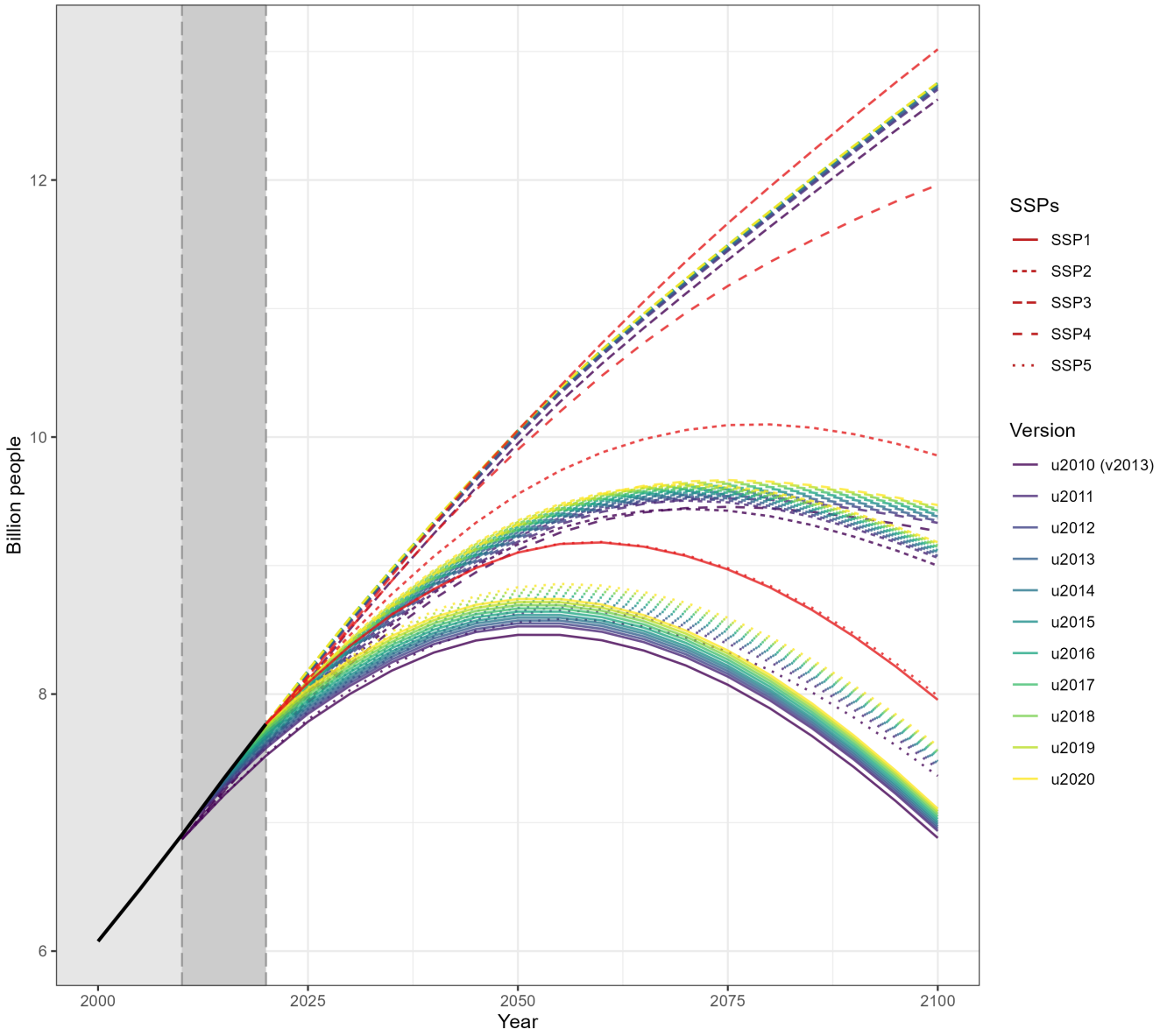

Fig. S1 World population. The black line in 2000 – 2020 shows observations. The red lines (by SSP; line type) show the new SSP v2024 projections. The colored lines (by SSP; line type) show the old SSP v2013 database and the near-term updates.

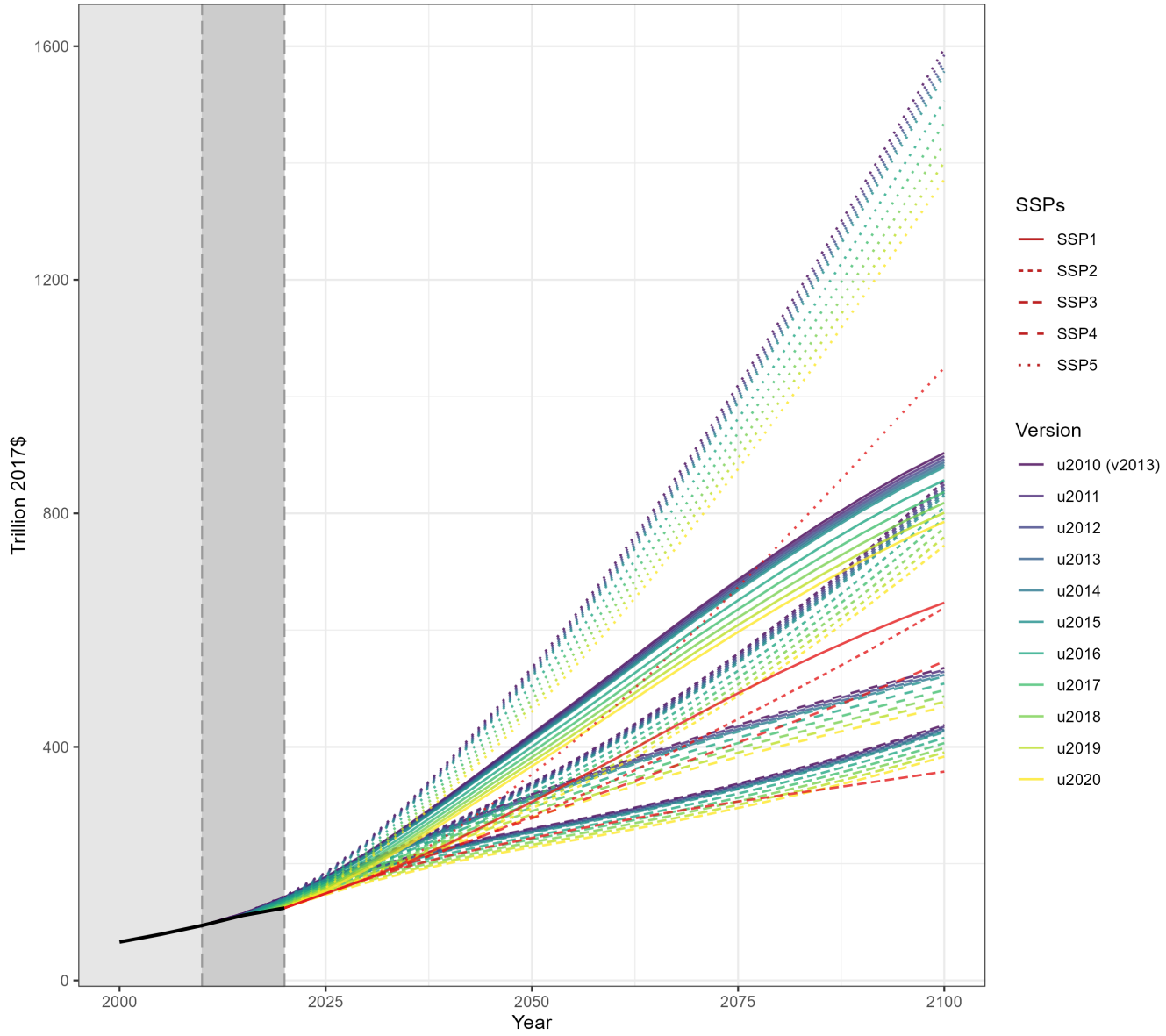

Fig. S2 World GDP. The black line in 2000 – 2020 shows observations. The red lines (by SSP; line type) show the new SSP v2024 projections. The colored lines (by SSP; line type) show the old SSP v2013 database and the near-term updates.

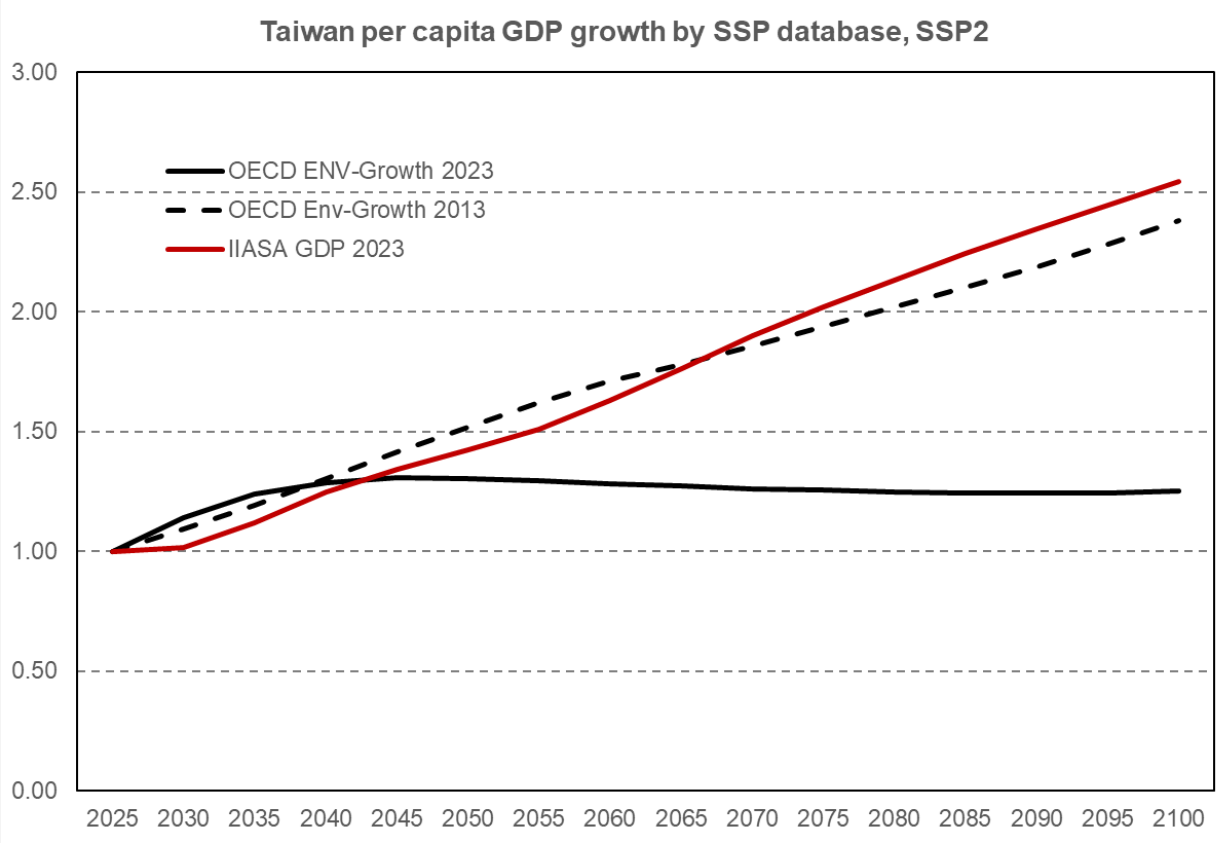

Fig. S3 Taiwan per capita GDP in SSP2 across sources.

References

1. O'Neill, B. C. *et al.* The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change* **42**, 169–180 (2017).
2. O'Neill, B. C. *et al.* Achievements and needs for the climate change scenario framework. *Nat. Clim. Chang.* **10**, 1074–1084 (2020).
3. Koch, J. & Leimbach, M. SSP economic growth projections: Major changes of key drivers in integrated assessment modelling. *Ecological Economics* **206**, 107751 (2023).
4. Patel, P. *et al. Core Model Proposal# 332: GCAM Macro-Economic Module (KLEM Version)*. https://jgcri.github.io/gcam-doc/cmp/332-GCAM_Macro_Economic_Module_KLEM.pdf (2023).
5. Foure, J. *et al.* Macroeconomic Drivers of Baseline Scenarios in Dynamic CGE models: Review and Guidelines Proposal. *Journal of Global Economic Analysis* **5**, 28–62 (2020).
6. Zhao, X. *et al.* gcamfaostat: An R package to prepare, process, and synthesize FAOSTAT data for global agroeconomic and multisector dynamic modeling. *Journal of Open Source Software* **9**, 6388 (2024).