Graduation study
Environmental Engineering Course, Department of Global Engineering

# Extension of Verification and Evaluation Method for Climate Mitigation Scenario: Considering Regional and Historical Information

# 気候緩和シナリオの検証と評価手法の拡張：地域と実績値を考慮した評価

March 1, 2024

Henry Chan
チェン　ヘンリー
2020 enrolled, Environmental Engineering Course
Department of Global Engineering
Faculty of Engineering, Kyoto University

# Extension of Verification and Evaluation Method for Climate Mitigation Scenario: Considering Regional and Historical Information

# 気候緩和シナリオの検証と評価手法の拡張：地域と実績値を考慮した評価

As the urgency for climate change mitigation measures intensifies, there is growing interest among policymakers and researchers in assessing the validity and feasibility of scenarios projected by integrated assessment models (IAMs), which has been used to depict future socio-economic system considering the impact of policies. Validity guarantees scenario quality by ensuring the reproducibility of historical trends and near-term projections, while Feasibility offers practical insights for decision-making by evaluating the degree of projection aligning with societal capacities for change.

Although IPCC Sixth Assessment Report (AR6) demonstrated global vetting for validity and literature-based feasibility assessments, concerns arise due to the lack of regional vetting and doubts about the appropriateness of indicator values. This study aims to extend the current approach by applying (1) regional vetting, which examines if scenarios fall within acceptable range from regional reference values, and (2) feasibility assessment framework which compares scenarios to distribution of historical change, on the AR6 scenario database. The effectiveness of this extension is evaluated by comparing results with the existing approach.

The result of regional vetting shows average passing rate for regions are 25% lower than global counterpart, suggesting necessary review on assumptions and mechanisms regulating regional projection in IAMs with global scope. Result of feasibility assessment shows major concern arises from carbon intensity and electrification rate, proving its versality as it aligns with literature. However, carbon intensity shows decreasing concern in scenarios with limited temperature increase, contrasting to the result of AR6. This suggests necessary adjustment of method to capture characteristics of scenarios.

| Henry Chan | チェン　ヘンリー |
|---|---|

Climate change mitigation, Integrated assessment models, Scenario analysis, Validity, Feasibility

気候変動緩和、統合評価モデル、シナリオ分析、妥当性、実現可能性

## Table of Contents

# 1. Introduction

## 1.1 Background

The Intergovernmental Panel on Climate Change (IPCC)'s Sixth Assessment Report (AR6), highlights that the recorded global average annual carbon dioxide emissions between 2010 and 2019 surpassed those of any previous decade (Dhakal et al., 2022). This observation emphasizes the need to achieve the climate target set in the Paris Agreement, which aims to restrain the rise in global average temperatures to “well below 2°C” above pre-industrial levels and strive “to limit the temperature increase to 1.5°C”. According to calculations based on the "Nationally Determined Contributions" (NDCs), which are greenhouse gas (GHG) reduction targets determined individually by each country, there exists a possibility that temperature increase will be over 1.5°C if the current trend keeps on until 2030. To ensure the temperature increase does not exceed 2°C, further implementation of climate mitigation measures is highlighted as necessary (Lecocq et al., 2022).

Integrated assessment models (IAMs), which model the change in socio-economic state quantitatively, provide analysis of the effectiveness of different climate mitigation policies. This includes renewable energy investment, carbon pricing, use of carbon capture and storage (CCS), etc. In many climate mitigation scenarios (scenarios), which are scenarios projected from IAMs considering the use of climate mitigation policies, generally agrees the utilization of emerging technologies like CCS, to have great importance when strict climate target is to be achieved. However, there are concerns about the feasibility of these scenarios, as most technologies explored are in early development stage and lack historical observation to verify/validate their utilization as suggested.

The concept of validity and feasibility of scenarios has been recently adopted to tackle these concerns. Validity is composed of two parts: (1) reproducibility of historical trends and (2) plausibility for near-future projection. Scenarios considered for in-depth analysis, as demonstrated in Annex III by Working Group III (WGIII) of IPCC AR6, are vetted with historical emission and energy production data and near-term projection. One on hand, this ensures the quality of scenarios used to perform other assessments (for example climate assessment and scenario comparison) later. On the other hand, feasibility, as Brutschin et al. (2021) suggests, is “the degree to which scenarios lie within the boundaries of societal capacities for change in a given period”. Barriers that hinder the achievement of climate targets can be recognized by assessing feasibility of scenarios. This represents a crucial aspect overlooked during quantitative assessments but has practical value in policy decision-making. The difference between validity and feasibility is that the former emphasizes quality control of scenarios for further analysis, while the latter focuses on providing decision-relevant information derived from scenarios. Typically, a validity test comes first in the scenario analysis process, followed by a feasibility evaluation.

However, the justification of the current combined approach, of vetting (verification of validity) and feasibility assessment (evaluation of feasibility) of scenarios, as adopted and shown by IPCC AR6 WGIII, require rigor or needs reasonable justification is not complete or needs to be improved on from two viewpoints. Firstly, only data at world level is being checked during vetting. While growing interest from both academic researchers and policy makers on the regional information contained in scenarios, it is considered more reasonable to include vetting of regional data as well. Secondly, some indicator’s values used in the feasibility assessment framework proposed by Brutschin et al. (2021) are not reasonable. Particularly, indicators for emerging technologies, like scale-up of CCS, use a “conservative” value of 2 percentage points of decadal growth as the medium level of concern. Although such setting is due to the lack of support from current observation, using a value not based on evidence is also not considered appropriate.

## 1.2 Objectives

The objective of this study is to provide an extension to the existing approach due to the following reasons:

1. A pressing need for further acceleration in the implementation of mitigation measures to achieve the 1.5°C climate target as outlined in the Paris Agreement.
2. IAMs provide thorough assessment for the effectiveness of different climate mitigation options, including emerging technologies.
3. Besides the cost-benefit approach, validity and feasibility are also important facets to be considered in the context of policy making and decision making.
4. Room for improvement inside the current scenario assessment approach adopted by IPCC, especially in terms of the lack of regional consideration during vetting and lack of justification of using indicators without strong evidence.

Therefore, the objective of this paper aims to provide an improvement to the existing approach. This paper aims to achieve the following goals:

1. Append regional historical data for vetting; and
2. Apply alternative feasibility assessment framework, specifically, the approach based on historical change as suggested by Nishimoto & Fujimori (2023).

The result obtained from the extended approach will compare with the result obtained from the existing approach, which consists of world-level vetting and feasibility assessment framework proposed by Brutschin et al. (2021). The difference between both results will be discussed in the hope to gain insights about the effectiveness of the extended approach.

# 2. Literature Review

## 2.1 Importance of Scenarios to Decision-Making

Edenhofer & Kowarsch (2015) proposed a model for evaluating environmental policies. Traditional models follow a unidirectional structure where policy objectives are determined first, followed by an evaluation of feasible policy options based on these objectives. In contrast, the newly proposed Pragmatic-Enlightened Model (PEM) allows for the potential revision of policy objectives, rather than solely assessing the direct and synergistic effects of policy options towards achieving a particular policy goal. In other words, there exists a mutual dependence between policy objectives and options, establishing a loop encompassing policy effects. They emphasize the necessity of scenario assessments from perspectives such as model comparisons, exploration of alternative measures, and the timing of policy implementation. They argue that conducting these assessments could offer society a diverse variety of policy options, contribute to more constructive debates, and ultimately aid in policymaking.

Céline Guivarch et al. (2022) noted that utilizing scenario ensembles in analysis yields more robust insights compared to analyzing a single scenario. They highlighted three advantages of employing scenario ensembles: (1) capturing the uncertainty of outcomes; (2) enhancing the reliability and relevance of information; (3) facilitating information sharing among different IAMs. Particularly, they suggested that the former two benefits would be advantageous for decision-making in the financial and political sectors. To maximize these advantages, users of scenario ensembles, including researchers, should consider three aspects: (1) scrutiny and quality management of mitigation scenarios, (2) selection of mitigation scenarios aligned with analysis objectives, and (3) providing efficient information access tools to end-users (such as policymakers and analysts).

## 2.2 Evaluation of IAMs

Wilson et al. (2021) conducted a comprehensive review on common practices when evaluating IAMs and proposed an evaluation framework consists of multiple evaluation practices. They reviewed a total of 6 evaluation methods, including (1) historical simulations, (2) near-term observations, (3) stylized facts, (4) model hierarchies, (5) model inter-comparison and (6) sensitivity analysis. The first three methods (1,2,3) require observational data to compare with the scenario data projected from IAMs. The following two methods (4,5) are based on the comparison between models. The last method (6) is mostly implemented by individual research. Each of these methods has its strengths and weaknesses. For example, historical simulation can provide a quantitative analysis but fails to reflect the predictive reliability of model because future conditions lie beyond the observed range. Therefore, an evaluating approach with a combination of the 6 methods is required to address four types of evaluation criteria, namely (1) appropriateness, (2) interpretability, (3) credibility, and (4) relevance. Appropriateness of a model shows the consistency of a model with the corresponding research question. Interpretability evaluated the difference in the result that can be explained from model's structure and assumptions. Credibility is measured as the performance of the model in fulfilling users' purposes, which is usually tested by comparing historical scenario results with historical data. Relevance informs whether the result of a model deepens scientific understanding and assists decision-making. The aim of the proposed framework is to improve the results simulated from IAMs and to promote "adequacy, legitimacy and usefulness of model applications" among IAMs' users.

## 2.3 Validity of Scenarios

In Chapter 3 of IPCC AR6 WGIII, Riahi et al. (2022) emphasized the importance of quality control in scenario assessment through vetting. The details of the vetting methodology are outlined in Annex III (IPCC, 2022). Scenarios vetting was conducted by evaluating: 1) reproducibility concerning historical trend, and b) consistency with projections for the near future. Reproducibility was assessed by selecting eight indicators across two dimensions — 1) greenhouse gas emissions and 2) energy production. This is to verify whether scenarios aligned within the acceptable ranges of observed reference values for each indicator, taking 2019 as the reference year. The acceptable ranges varied based on the following scenario assessment. For instance, the acceptable ranges for identification of illustrative mitigation pathways (IMPs) were narrower than those for climate category classification. Although the consistency with projections for the near future is also checked, it did not serve to eliminate scenarios even though they fail to fall within the acceptable ranges. The result of vetting showed that out of 2266 mitigation scenarios, only 1686 passed the vetting for climate category classification, and 1202 passed the stricter standards of vetting for identification of IMPs. Reasons for a significant number of scenarios failing were attributed to the fact that some of the scenarios were inherited from the Special Report on 1.5°C (SR1.5), hence lacked data update, while other scenarios are designed with the objective to explore alternative pathway which deviated from history.

Byers et al. (2023) performed vetting at global and regional (specifically European Union (EU)) levels, to distinguish scenarios which align with EU's net-zero goal. Their workflow for scenario assessment in this context drew reference from the previously mentioned Chapter 3 of WGIII. The significant distinction in their vetting was the inclusion of reference values for EU-level indicators, whereas WGIII's Chapter 3 solely addressed global-level indicators. This addition was motivated by the possibility that differing trends may occur between global and EU-level indicators. For instance, the change in CO2 emissions from fossil fuel use between 2015 and 2020 showed an increasing trend globally but a decreasing trend at the EU level. Out of 1062 scenarios subjected to vetting, only 492 met these criteria. Notably, there were 262 scenarios which only met the criteria for either world or EU. This highlights the importance of considering regional variability when vetting scenarios, which possibly identifies scenarios that are not appropriate to be used with regional context.

## 2.4 Feasibility Assessment of Scenarios

Brutschin et al. (2021) endeavored to develop a multidimensional and quantitative framework to assess the feasibility of scenarios. The multidimensional aspect encompassed five dimensions: (1) geographical-physical, (2) technological, (3) economic, (4) social-cultural, and (5) institutional. Within each dimension, they selected numerous indicators, resulting in a total of 24 indicators. Two reference values are defined through literature review and historical data analysis for each indicator, namely medium concern threshold and high concern threshold. By employing these reference values, they could evaluate feasibility concerns of scenarios by classifying them into three categories (low, medium and high level of concern). Indicators are defined based on the decadal change, which enables the assessment of the timing and scales of societal transformations suggested by scenarios. The first application of this assessment to the SR1.5 scenario database revealed that the most concerning aspect of feasibility is institutional. Specifically, there seemed to be a mismatch between governance capabilities and emission reduction targets, potentially leading to an overly optimistic pathway as shown in many scenarios. Additionally, the execution of early mitigation measures was found to cause intensified short-term societal transformations and suppress the long-term scale of societal change.

Nishimoto & Fujimori (2023) attempted to define feasibility of scenarios by utilizing distribution of historical rates of change. They demonstrate the effectiveness of the approach using three indicators—energy intensity, carbon intensity, and electrification rate. They utilized data from the International

Energy Agency (IEA) spanning from 1975 to 2015 at both country and regional levels. These historical trends were contrasted to scenario data obtained from Asia-Pacific Integrated Assessment Model (AIM/Hub; Fujimori et al., 2017) which simulates from 2010 to 2100 for 17 regions. The scenario data is defined as "feasible" only if the change of indicator in scenario data falls within the 90% range of corresponding historical distribution. The 90% range is used to exclude outliers caused due extreme and sudden changes, for example the changes due to famine or war, which are usually included in historical distribution. As a result, under the most stringent climate target (1.5°C) scenario, the passing rates of feasibility assessment for energy intensity, carbon intensity, and electrification rate are 90%, 65% and 59% respectively. This approach contrasts to previously mentioned Brutschin et al. (2021) threshold-based approach in 3 ways: (1) it is independent of model structure, (2) independent of historical threshold, and (3) simple.

## 2.5 Characteristics of This Study

To sum up, scenarios obtained from IAMs are agreed to have their significance when it comes to decision-making. Vetting and feasibility assessment are attracting growing interest as utilization of scenarios are widely recognized, especially under the analysis of net-zero emissions pathways. However, although individual updates of methodology used in vetting and feasibility assessment are seen in some research, such modifications are not explored thoroughly and integrated in the publication from IPCC. Therefore, in this study, regional vetting and the feasibility assessment framework based on historical change is applied to AR6 scenario database, in the hope to validate the versality of such extensions and investigate the difference between the original approach adopted by WGIII, which introduces the possibility to gain insights for the discussion of better scenario processing and assessment methods, potentially leading to improvement of IAMs.

# 3. Methodology

## 3.1 Overview

Historical data for vetting and feasibility assessment, including emissions and energy data, is gathered from a range of data sources, and aggregated to desired regional scope. AR6 scenario database is used as the source of scenarios data which the extended approach is applied to. The result obtained from the extended approach is compared to the original approach used by IPCC AR6 WGIII.

## 3.2 Historical Data

For historical data of emissions, Emissions Database for Global Atmospheric Research (EDGAR; European Commission & Joint Research Centre, 2023) version 8.0 is used. EDGAR Community GHG emissions database is an emissions database provided by the European Commission, with data gathered and supplied by Joint Research Centre (JRC) and IEA. It provides data for historical $CO_2$ emissions originated from fossil fuel combustion, cement production, metal production, urea production, agricultural liming, and solvent use, covering from 1970 to 2022 at sector-specific and country-specific level. $CH_4$ emissions can also be referenced from EDGAR. Alternative data source, IEA's Greenhouse Gas Emissions from Energy 2022 (IEA, 2022), is used instead to obtain $CO_2$ emissions originated from energy use, which is required for the calculation of carbon intensity in feasibility assessment.

For historical data of energy, IEA Energy Balance 2022 (IEA, 2022) is used. It provides detailed energy information, for example primary energy supply and secondary energy supply, covering from 1971 to 2021 at fuel-specific and country-specific level. When accounting for primary energy production of nuclear, geothermal, and non-thermal renewables, IEA adopts "physical energy content" method, which assumes varying primary-to-secondary energy efficiency for each energy source. Specifically, 33% for nuclear, 10% for geothermal and 100% for non-thermal renewables (e.g. solar PV). To align with scenarios' energy data, which assumes one unit of secondary energy is equivalent to one unit of primary energy (i.e. "direct equivalent" method), the following equation is applied when aggregating historical primary energy data from IEA. Note that TES stands for total energy supply (i.e. total primary energy), $PE_i$ stands for primary energy for energy source i.

$$TES_{direct\ equivalent} = \sum PE_i - 0.67\,(PE_{nuclear}) - 0.9\,(PE_{geothermal}) \tag{1}$$

IEA Energy Balance 2022 also contains data of GDP which is required to calculate the energy intensity as used in feasibility assessment. As historical GDP data is presented in constant 2015 USD, it is converted into constant 2010 USD, which aligns with the GDP data presented in scenarios. The conversion involves the use of GDP deflators, which are retrieved from the World Development Indicators (The World Bank, 2023) provided by the World Bank Group. The following equation demonstrates the conversion of constant year of GDP.

$$GDP_{USD2010} = GDP_{USD2015} \times \frac{deflator_{USD2015}}{deflator_{USD2010}} \tag{2}$$

## 3.3 Scenario Data

For scenario data, AR6 scenario database version 1.1 (Byers et al., 2022) is used. It consists of over 3000 scenarios, collected by the International Institute for Applied Systems Analysis (IIASA) from research teams between September 2019 and July 2021. Managed by IIASA, this database is publicly accessible on the internet. Scenarios can be divided into long-term mitigation scenarios covering the global scale, short to medium-term regional-specific scenarios, and sectoral scenarios focusing on the industry, building and transportation sectors. However, as for the focus of this study, long-term mitigation scenarios covering global scale are used.

## 3.4 Regional Vetting

The vetting at global level as showcased in Annex III of IPCC AR6 WGIII consists of two parts: reproducibility of historical trend and consistency to near-term projection, as mentioned in section 2.3. Due to the limitation of data availability, this study focuses on selected indicators used for vetting historical trends, which are shown in Table 1. Reference year is set at 2019, because of the accuracy and consistency for historical data retrieved from data sources. The acceptable ranges used in this study are equivalent to the stricter ranges adopted for the identification of IMPs as shown in Annex III. For the acceptable range for $CO_2$ emissions (energy and industrial process, EIP) percentage change during 2010-2019, a proposed ±15% from observed historical change is used which is deduced Byers et al. (2023) (+5% to -30% for EU; -2.5 to +20% for world). For regional resolution, the R5-region defined by IPCC is used. Aggregation list of R5-region is shown in Table 2. Although finer resolution like R10-region is also considered possible, R5-region allows better scenario coverage.

Table 1 Indicators and acceptable ranges for historical vetting

| **Indicator** | **Acceptable range** |
|---|---|
| $CO_2$ emissions (EIP) | ±10% |
| $CO_2$ emissions (EIP) % change 2010-2019 | ±15% |
| $CH_4$ emissions | ±20% |
| Primary energy | ±10% |
| Electricity: nuclear | ±20% |
| Electricity: solar and wind | ±25% |

Table 2 R5-region aggregation list

| Region | Description | Countries |
| --- | --- | --- |
| OECD & EU (R5OECD+EU) | OECD 90 and EU member states and candidates | Albania, Australia, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Guam, Hungary, Iceland, Ireland, Italy, Japan, Kosovo, Latvia, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, New Zealand, North Macedonia, Norway, Poland, Portugal, Puerto Rico, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States of America |
| Reforming Economies (R5REF) | Countries from the Former Soviet Union | Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan |
| Asia (R5ASIA) | Asian countries excluding Middle East and Japan | Afghanistan, Bangladesh, Bhutan, Brunei Darussalam, Cambodia, China, Democratic People's Republic of Korea, Fiji, French Polynesia, Hong Kong, India, Indonesia, Lao People’s Democratic Republic, Macao, Malaysia, Maldives, Micronesia (Fed. States of), Mongolia, Myanmar, Nepal, New Caledonia, Pakistan, Papua New Guinea, Philippines, Republic of Korea, Samoa, Singapore, Solomon Islands, Sri Lanka, Taiwan, Thailand, Timor-Leste, Vanuatu, Vietnam |
| Middle East & Africa (R5MAF) | Middle East and African countries | Algeria, Angola, Bahrain, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo, Côte d’Ivoire, Democratic Republic of the Congo, Djibouti, Egypt, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Iran, Iraq, Israel, Jordan, Kenya, Kuwait, Lebanon, Lesotho, Liberia, Libyan Arab Jamahiriya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, |

| | | |
|---|---|---|
| | | Mozambique, Namibia, Niger, Nigeria, Occupied Palestinian Territory, Oman, Qatar, Rwanda, Réunion, Saudi Arabia, Senegal, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Swaziland, Syrian Arab Republic, Togo, Tunisia, Uganda, United Arab Emirates, United Republic of Tanzania, Western Sahara, Yemen, Zambia, Zimbabwe |
| Latin America (R5LAM) | Latin and South American countries | Argentina, Aruba, Bahamas, Barbados, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, French Guiana, Grenada, Guadeloupe, Guatemala, Guyana, Haiti, Honduras, Jamaica, Martinique, Mexico, Nicaragua, Panama, Paraguay, Peru, Suriname, Trinidad and Tobago, United States Virgin Islands, Uruguay, Venezuela |

### 3.5 Feasibility Assessment Based on Historical Change

Nishimoto & Fujimori's (2023) feasibility assessment framework based on historical change, as described in section 2.4, is applied. While the regional scope used in the literature is compatible with the 17 regions defined by AIM/Hub (Fujimori et al., 2017), R10-region is considered suitable for this study, as R-10 region has been used as the "least common regional resolution" across IAMs. Table 3 shows the aggregation list for R10-region. Note that the mapping shown in Table 3 is not exhaustive due to the inevitable difference in regional resolution between IAMs. In addition, while scenarios are classified with their average temperature increase by 2100 in the literature, in this study climate indicated by IPCC WGIII is used instead. In IPCC AR6 WGIII, scenarios are classified into climate category by Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC; Meinshausen et al., 2020), a comprehensive climate emulator, regarding their warming level and corresponding likelihood. This also involves a preliminary harmonization and infilling steps for the required emission types.

Table *4* summarizes detailed definition for climate category. The period for feasibility assessment of scenarios is defined as from 2010 to 2100, with a time interval of 5 years. The historical trend is taken from 1975 to 2015, also with time interval of 5 years. Indicators are defined and summarized in Table 5. For energy intensity, rate of change (i.e. in %) is used, while for carbon intensity and electrification rate, change is used. Feasibility is defined as whether change of indicator belongs to a normally observed historical distribution, in other words whether scenario falls within the 90% range of corresponding historical distribution. The reason why 90% range is used is to exclude extreme and sudden changes, as described and reasoned in section 2.4.

Table 3 R10-region aggregation list

| Region | Description | Countries |
|---|---|---|
| Africa (R10) | Africa | Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo, Côte d'Ivoire, Democratic Republic of the Congo, Djibouti, Egypt, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Libyan Arab Jamahiriya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Occupied Palestinian Territory, Rwanda, Réunion, Senegal, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Swaziland, Togo, Tunisia, Uganda, United Republic of Tanzania, Western Sahara, Zambia, Zimbabwe |

| | | |
|---|---|---|
| China+ (R10) | Centrally planned Asia, primarily China | China, Democratic People's Republic of Korea, Hong Kong, Vietnam |
| Europe (R10) | Europe (including Turkey) | Albania, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Kosovo, Latvia, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom |
| India+ (R10) | South Asia, primarily India | India, Bangladesh, Nepal, Pakistan, Sri Lanka |
| Latin America (R10) | Latin America and the Caribbean | Argentina, Bahamas, Barbados, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guadeloupe, Guatemala, Guyana, Haiti, Honduras, Jamaica, Martinique, Mexico, Netherlands Antilles, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Suriname, Trinidad and Tobago, Uruguay, Venezuela |
| Middle East (R10) | Middle East | Bahrain, Iraq, Iran, Israel, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Syrian Arab Republic, United Arab Emirates, Yemen |
| North America (R10) | United States and Canada | Canada, United States of America |
| Pacific OECD (R10) | Pacific OECD | Australia, Japan, New Zealand |
| Reforming Economies (R10) | Former Soviet Union | Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Republic of Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan |
| Rest of Asia (R10) | Other countries of Asia | Afghanistan, Bhutan, Brunei Darussalam, Cambodia, Fiji, French Polynesia, Indonesia, Lao People's Democratic Republic, Malaysia, Maldives, Micronesia (Fed. States of), Mongolia, Myanmar, New Caledonia, Papua New Guinea, Philippines, |

| | |
|---|---|
| | Republic of Korea, Samoa, Singapore, Solomon Islands, Taiwan, Thailand, Timor-Leste, Vanuatu |

Table 4 Climate category defined by IPCC AR6 WGIII

| **Category** | **Description** |
|---|---|
| C1: Limit warming to 1.5°C (>50%) with no or limited overshoot | Reach or exceed 1.5°C during the 21$^{st}$ century with a likelihood of ≤67%, and limit warming to 1.5°C in 2100 with a likelihood >50%. Limited overshoot refers to exceeding 1.5°C by up to about 0.1°C and for up to several decades |
| C2: Return warming to 1.5°C (>50%) after a high overshoot | Exceed warming of 1.5°C during the 21$^{st}$ century with a likelihood of >67%, and limit warming to 1.5°C in 2100 with a likelihood of >50%. High overshoot refers to temporarily exceeding 1.5°C global warming by 0.1°C–0.3°C for up to several decades. |
| C3: Limit warming to 2°C (>67%). | Limit peak warming to 2°C throughout the 21$^{st}$ century with a likelihood of >67%. |
| C4: Limit warming to 2°C (>50%) | Limit peak warming to 2°C throughout the 21$^{st}$ century with a likelihood of >50%. |
| C5: Limit warming to 2.5°C (>50%) | Limit peak warming to 2.5°C throughout the 21$^{st}$ century with a likelihood of >50%. |
| C6: Limit warming to 3°C (>50%) | Limit peak warming to 3°C throughout the 21$^{st}$ century with a likelihood of >50%. |
| C7: Limit warming to 4°C (>50%) | Limit peak warming to 4°C throughout the 21$^{st}$ century with a likelihood of >50%. |
| C8: Exceed warming of 4°C (≥50%) | Exceed warming of 4°C during the 21$^{st}$ century with a likelihood of ≥50%. |

Table 5 Indicator defined in Nishimoto & Fujimori’s (2023) work

| **Indicator** | **Definition** | **Equation** |
|---|---|---|
| Energy Intensity | Energy consumption per unit of GDP. | $\frac{\text{Primary Energy Supply (TES)}}{\text{Activity Level (GDP)}}$ |
| Carbon Intensity | Carbon emissions produced per unit of energy supplied | $\frac{\text{CO2 Emissions (from energy)}}{\text{Primary Energy Supply (TES)}}$ |
| Electrification Rate | Proportion of energy consumed that comes from electricity sources. | $\frac{\text{Electricity (within TFC)}}{\text{Final Energy Consumption (TFC)}}$ |

# 4. Results & Discussion

## 4.1 Results

### 4.1.1 Historical Trend & Scenarios' Projection

Historical trend of indicators aggregated from data sources and projections from scenarios are presented to depict the relationship between observations and scenarios' projection. Figure A1.a shows the historical trend (1970-2022; solid line) and median value of scenarios' projection (2010-2100; dashed line) of $CO_2$ (EIP) emissions in R5 regions and in world. Area in grey represents 90% coverage of scenarios' projection (i.e. between 5$^{th}$ percentile and 95$^{th}$ percentile). In general, the projection follows the historical trend in recent years. The projection range becomes wider in future years compared to that of in recent years. Whilst the projection range is narrow at world level in recent years, it becomes wider for regions like OECD+EU, ASIA, and Reforming Economies (REF; former Soviet Union). For trend and projection of other indicators, please refer to Figure A1.b–A1.g listed in Appendix.

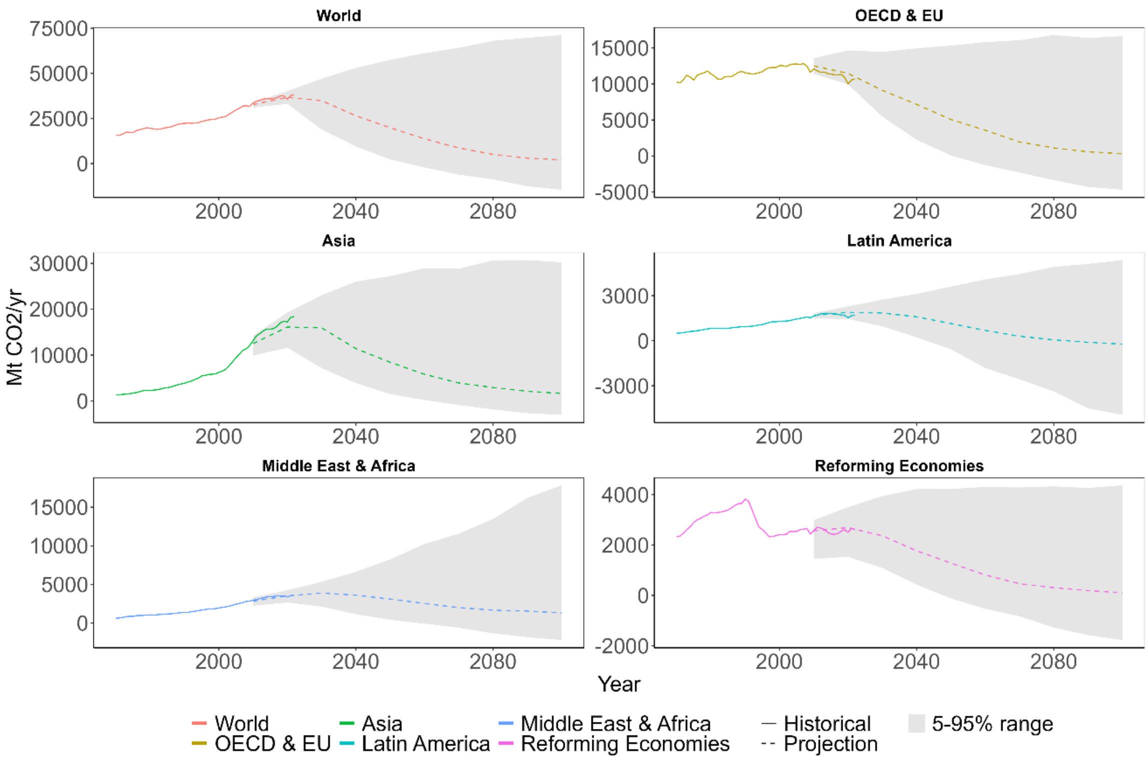


Figure A1.a Historical trend and scenarios' projection of $CO_2$ (EIP) emissions

4.1.2 Regional Vetting

Figure 1a-f shows the distribution of deviation a vetting value is from the defined reference value in each region (World and R5-region). The bottom horizontal line of a box indicates the 1st quartile (i.e. 25th percentile), and correspondingly the top horizontal line indicates the 3rd quartile (i.e. 75th percentile). The horizontal line in the middle indicates the median. Vertical lines extending from both ends of the box show ±1.5 interquartile range (i.e. 3rd quartile - 1st quartile; equivalent to height of box). Dots represent outliers. Note that deviation is calculated as the difference between vetted value and reference value divided by reference value, except for that of $CO_2$ (EIP) emissions percentage change (Figure 1b), which is calculated as the direct difference in percentage point. Deviation = 0 means a perfect match. Red dashed lines indicate the lower and upper boundary of acceptable ranges. Due to the limitation of graph size, extreme values with over ±1 deviation are not shown in the graphs.

Figure 1a shows the distribution of deviation for $CO_2$ (EIP) emissions. For Latin America (LAM), the median duplicate with the upper boundary, meaning that 50% of the total sample lies above the upper boundary. On the other hand, ASIA has a similar situation but in the opposite direction, lying towards the lower boundary.

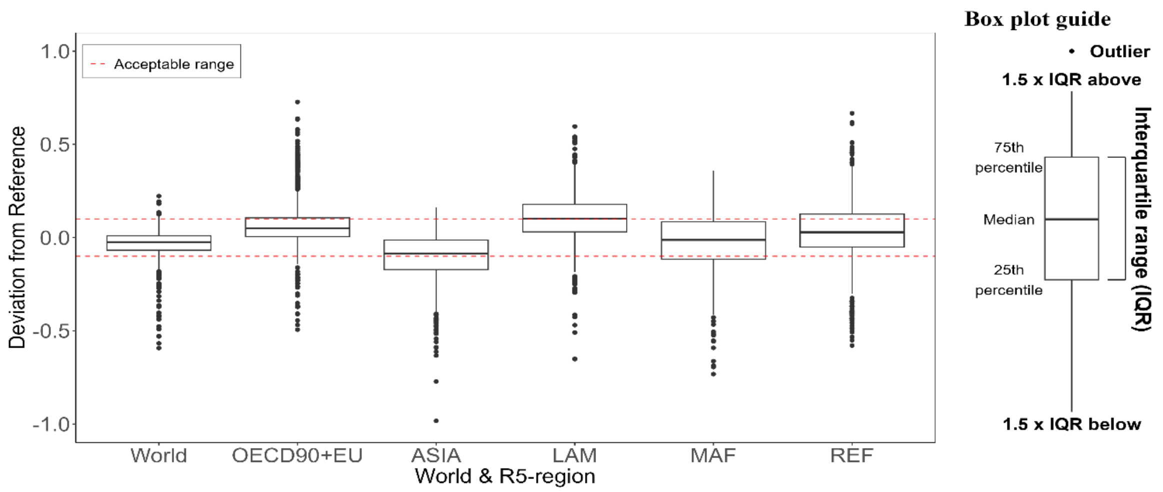


Figure 1a Deviation from reference ($CO_2$ (EIP) emissions)

Figure 1b shows the distribution of deviation for percentage change in $CO_2$ (EIP) emissions. The interquartile range is narrow regardless of regions. Also, compared to Figure 1a, which shows the vetting result of $CO_2$ (EIP) emissions in 2019, the distribution in each region shifts towards zero deviation respectively.

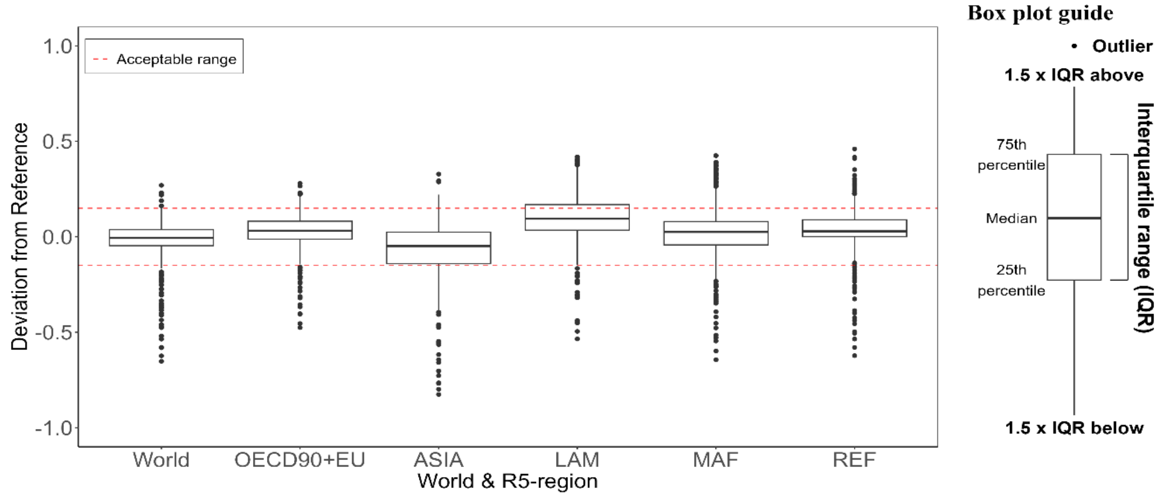


Figure 1b Deviation from reference ($CO_2$ (EIP) emissions % change)

Figure 1c shows the distribution of deviation for $CH_4$ emissions. Most of the regions show similar condensed distribution as to that of the world and fall within the boundaries. However, Middle East and Africa (MAF) and REF show comparatively distributed deviation. Distribution of deviation for MAF lies towards the lower boundary, while that of REF lies towards the upper boundary.

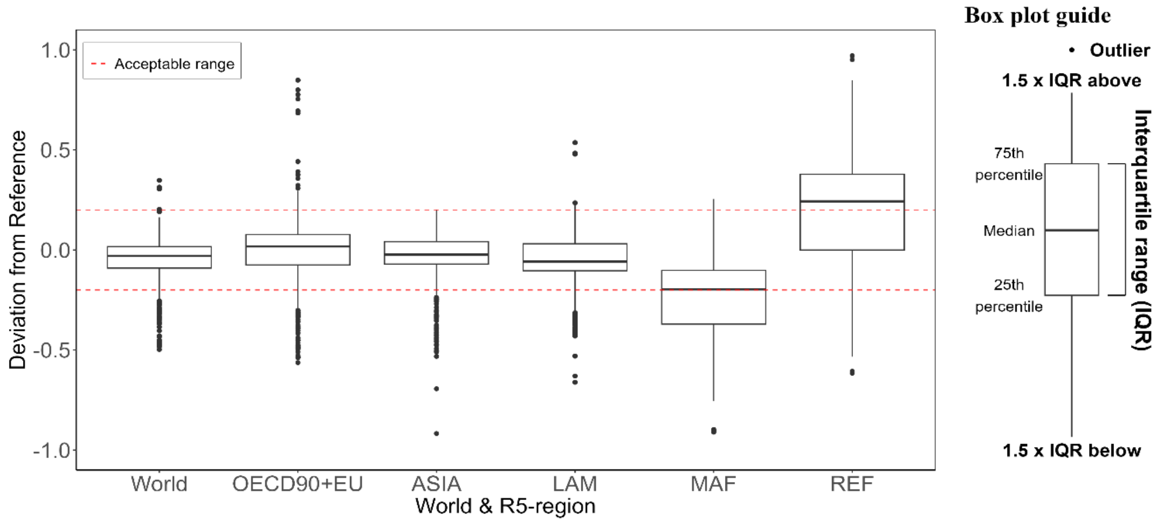


Figure 1c Deviation from reference ($CH_4$ emissions)

Figure 1d shows the distribution of deviation for primary energy. Generally, the distribution of deviation is condensed, regardless of regions. Although the box part of REF is mostly within the acceptable range, it is the most distributed one when compared to its counterpart in other regions.

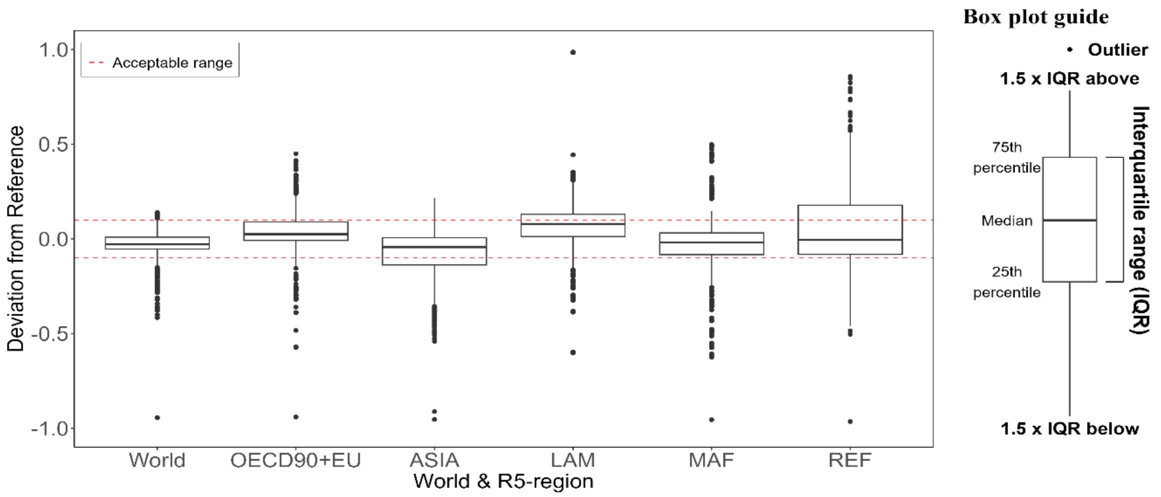


Figure 1d Deviation from reference (primary energy)

Figure 1e shows the distribution of deviation for electricity from nuclear. Compared to other indicators, deviation is relatively distributed for all regions. The most distributed region is MAF, which is widely distributed both beyond and below the acceptable ranges. However, regions other than MAF, have their box part fall within the boundaries.

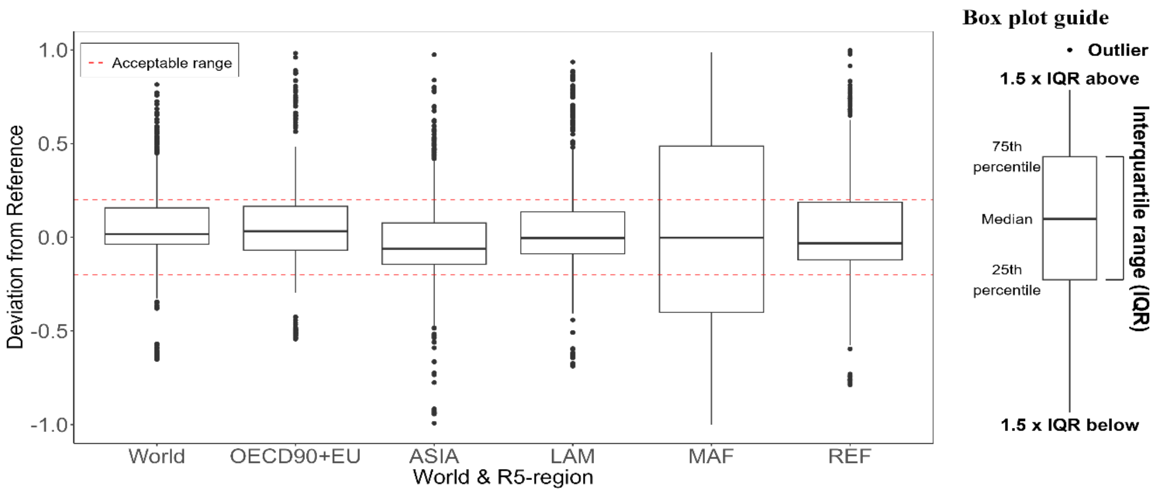


Figure 1e Deviation from reference (electricity from nuclear)

Figure 1f shows the distribution of deviation for electricity from solar and wind. Deviation for all regions is considered distributed, and tends to lie towards the lower boundary, including the world level. The median of LAM and MAF stay near the lower acceptable range. The most distributed region is REF, with outliers, which is not shown in the graph, lying beyond ±1 deviation.

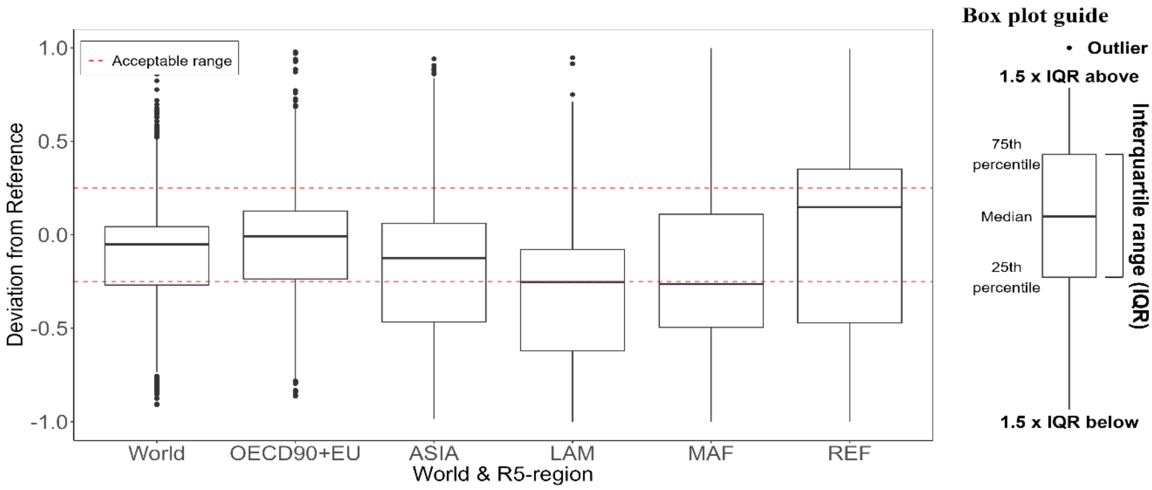


Figure 1f Deviation from reference (electricity from solar and wind)

Figure 2a-f are the corresponding graphs to Figure 1a-f showing the passing rate for each vetting. Passing rate is calculated as number of scenarios within acceptable range divided by the total number of scenarios being vetted. The number n marked under region names indicates the total number of scenarios tested (i.e. either pass or fail).

Figure 2a shows the passing rate of $CO_2$ (EIP) emissions. Passing rate greatly differs between regions. World level vetting shows the highest passing rate which is over 90% whilst LAM and REF show passing rates below 50%.

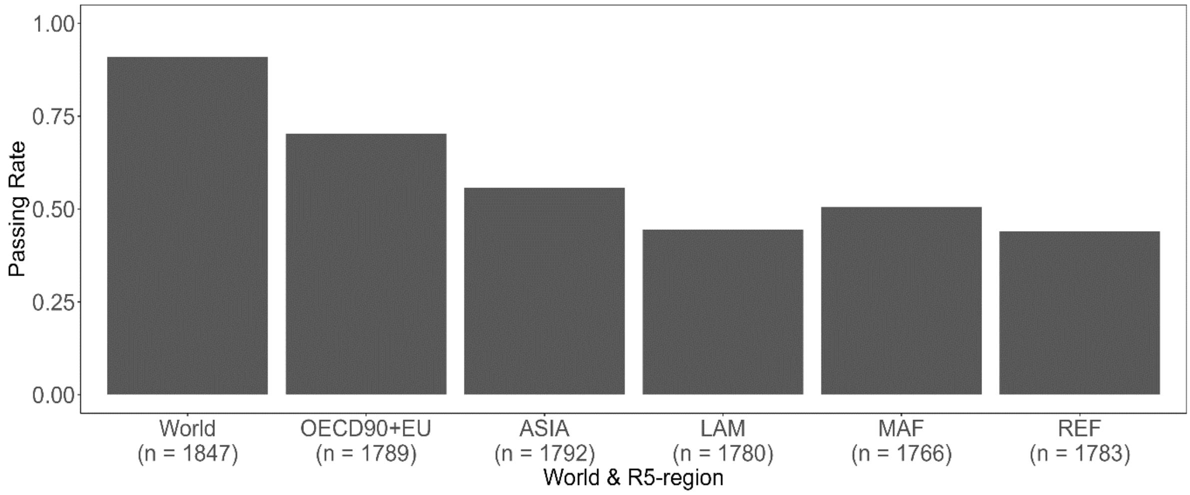


Figure 2a Passing rate of $CO_2$ (EIP) emissions

Figure 2b shows the passing rate of percentage change in $CO_2$ (EIP) emissions. Except LAM, all regions show a passing rate of over 75%. LAM, while being the lowest passing region, is still above 65%.

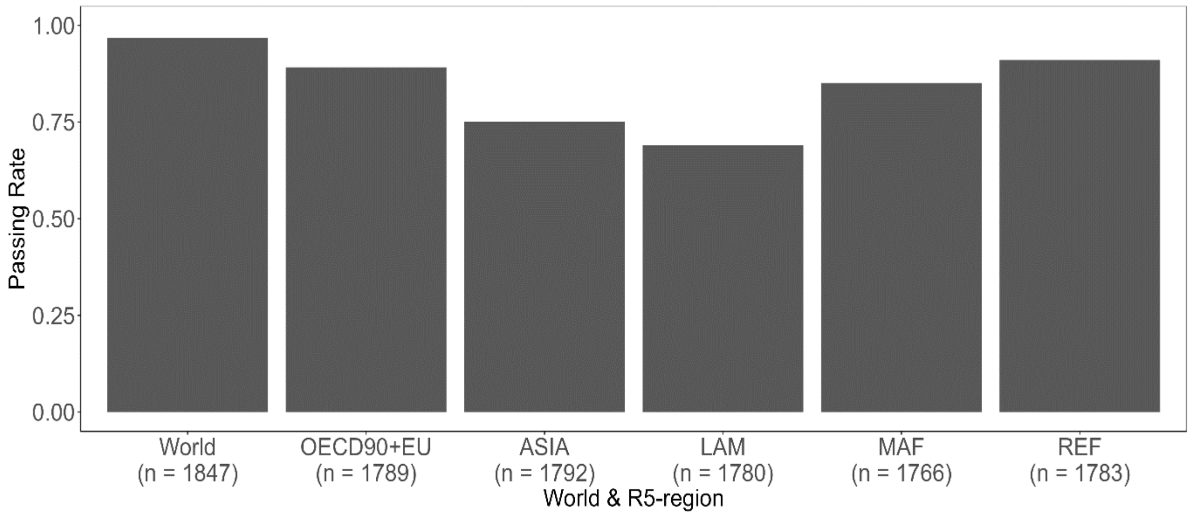


Figure 2b Passing rate of $CO_2$ (EIP) emissions % change

Figure 2c shows the passing rate of $CH_4$ emissions. Most of the regions show consistent results of passing rate, which is above 80%. However, for MAF and REF, passing rates are below average, with REF showing passing rate of only around 30%.

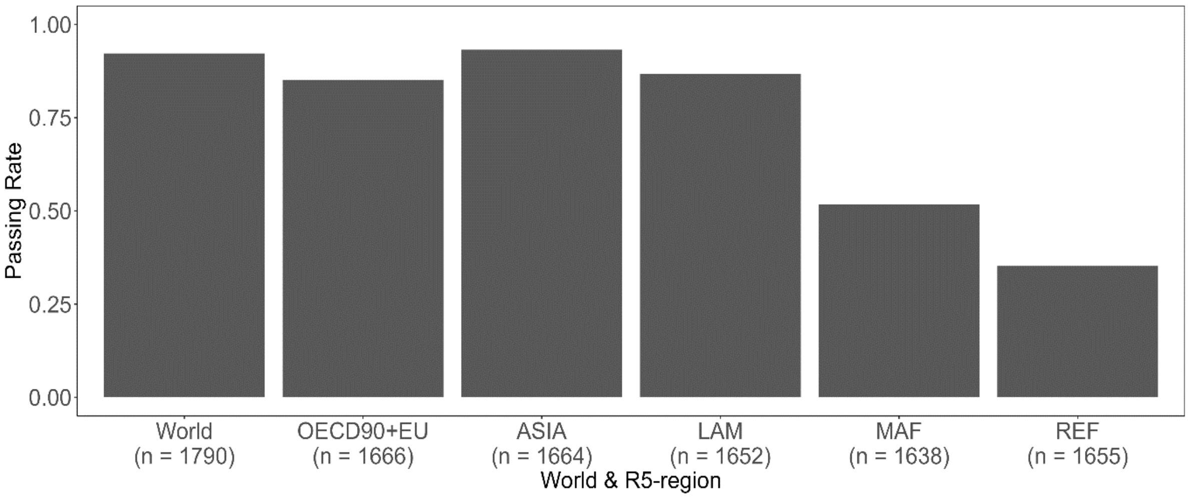


Figure 2c Passing rate of $CH_4$ emissions

Figure 2d shows the passing rate of primary energy. Whilst the passing rate for world vetting exceeds 90%, the passing rate of REF stays below 50%.

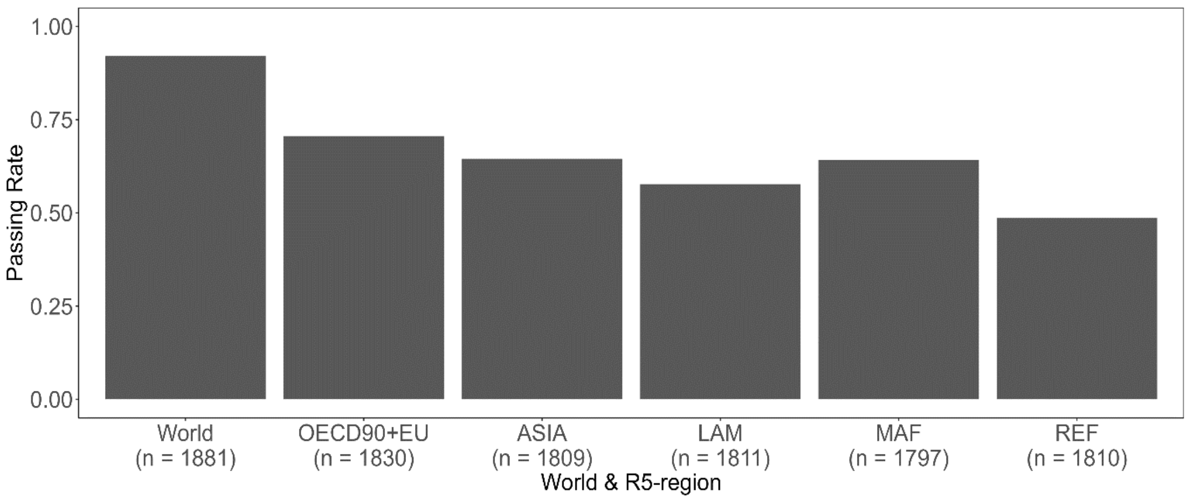


Figure 2d Passing rate of primary energy

Figure 2e shows the passing rate of electricity from nuclear. Passing rates for world, OECD+EU and ASIA are similar, which is around 65%. The passing rate for MAF is the lowest, which is below 25%.

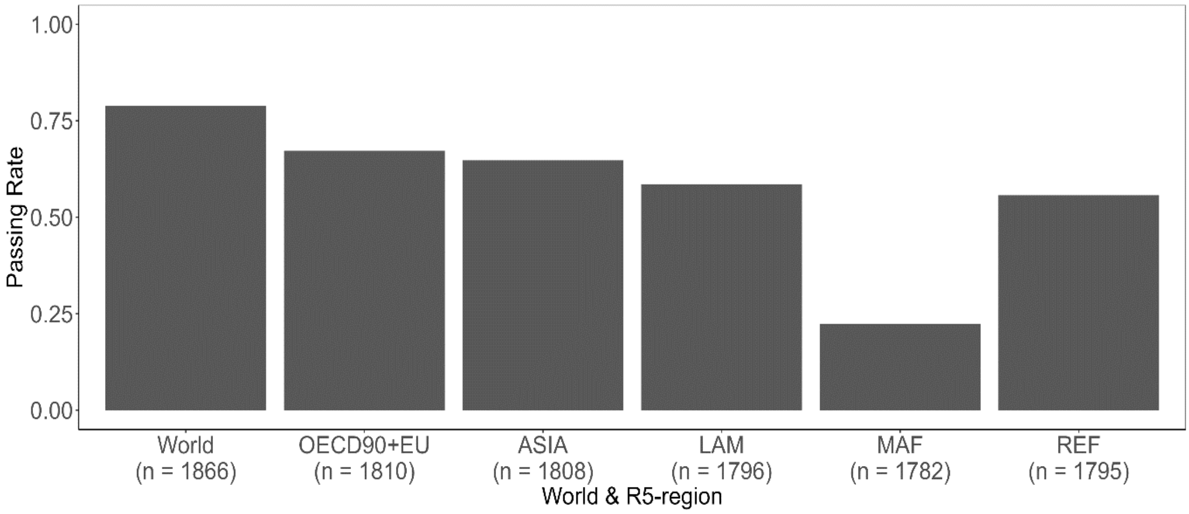


Figure 2e Passing rate of electricity from nuclear

Figure 2f shows the passing rate of electricity from solar and wind. The average passing rate is lower than 50%, which is the lowest compared to that of other indicators. The passing rate of OECD+EU is slightly higher than that of the world.

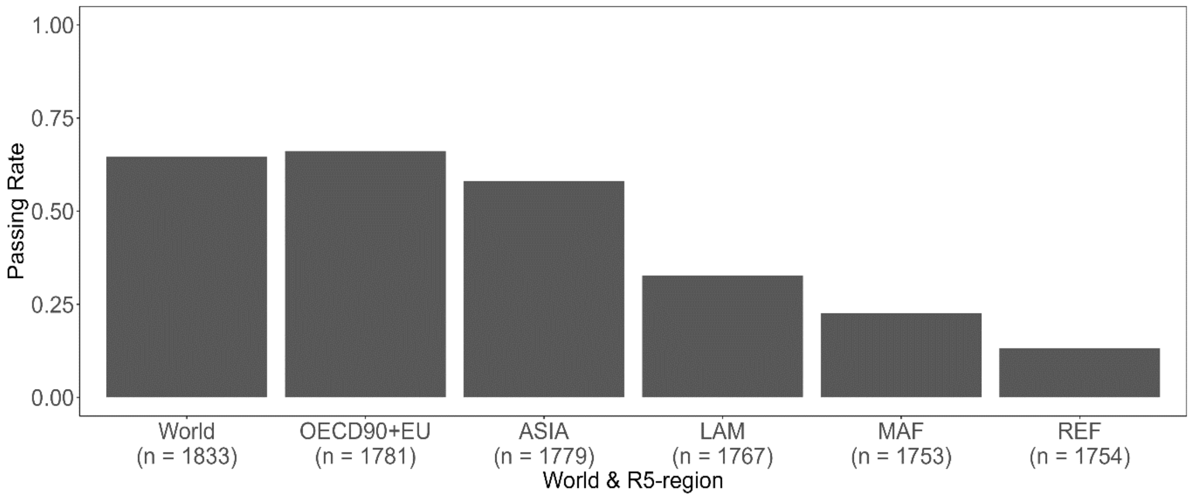


Figure 2f Passing rate of electricity from solar and wind

### 4.1.3 Feasibility Assessment

Figure 3a.I-c.I shows the historical trend and scenarios' projection of the 3 indicators used for feasibility assessment. Solid lines represent historical data and dashed lines represent scenarios' median projection. The shaded area represents 90% coverage of the scenarios' projection. Each subplot represents the climate category (category) scenarios belong to.

Figure 3a.II-c.II shows the historical trend and scenarios' projection of the rate of change or change of indicators. Dots represent historical data, squares represent median of projection across all regions, and vertical lines represent 90% coverage of projection.

Figure 3a.I and Figure 3a.II show the trend and projection of energy intensity and its rate of change respectively. Regardless of categories, all the projection shows negative change which aligns with the historical trend. Scenarios in C7 and C8 show a gentle and consistent change throughout the century, while scenarios with limited temperature increase show more rapid change in near term before 2030.

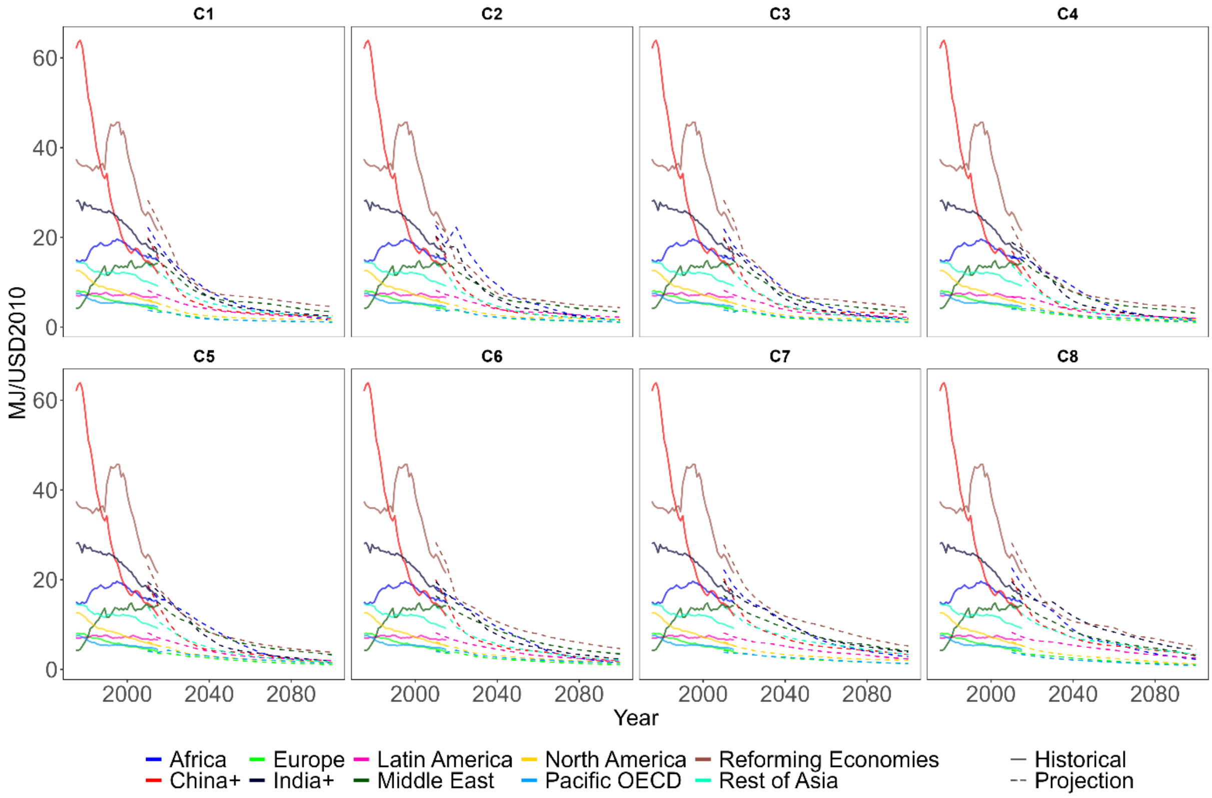


Figure 3a.I Trend and projection of energy intensity

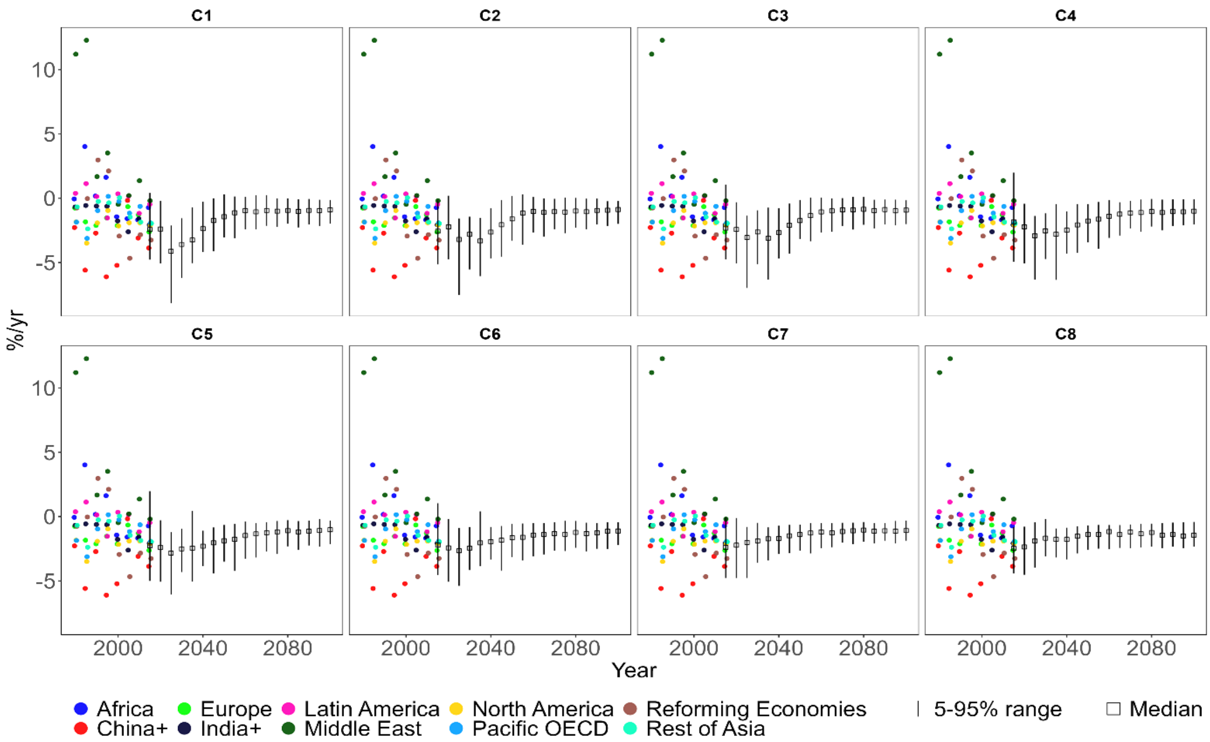


Figure 3a.II Trend and projection of rate of change of energy intensity

Figure 3b.I and Figure 3b.II show the trend and projection of carbon intensity and its change respectively. C8's scenarios continue the current trend which does not reflect an obvious direction of change. However, scenarios in C5 and C6 show carbon intensity drops later in the century, while scenarios in C4 or lower categories project a sudden decrease around 2030.

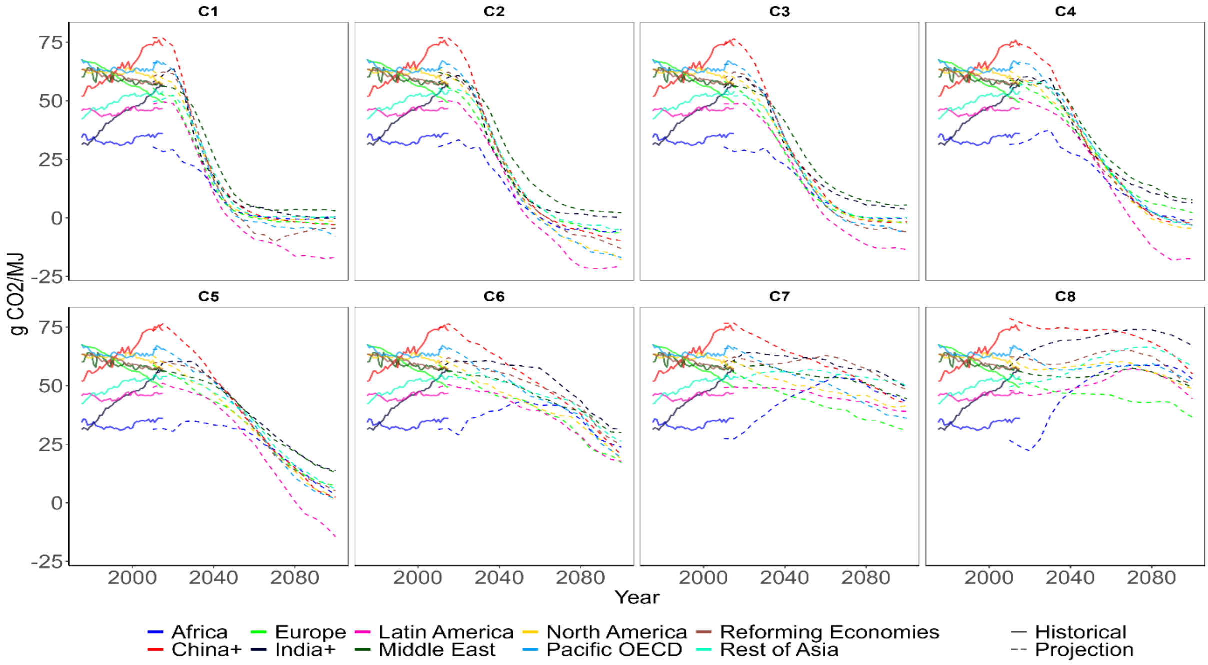


Figure 3b.I Trend and projection of carbon intensity

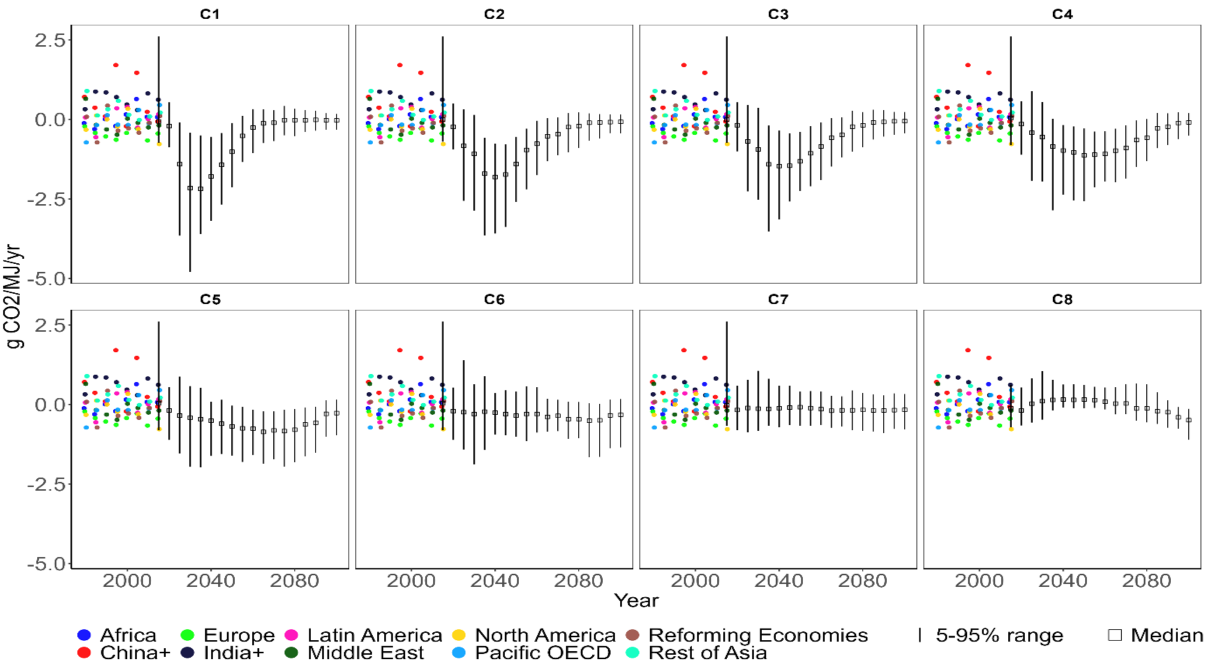


Figure 3b.II Trend and projection of carbon intensity change

Figure 3c.I and Figure 3c.II show the trend and projection of electrification rate and its change respectively. Scenarios in all categories project an increase in electrification rate that agrees with historical trends. With lower the categories are, the more rapid diffusion of electrification is projected which peaks around 2040. However, projected range in strict climate scenarios are more spread than those in C7 and C8.

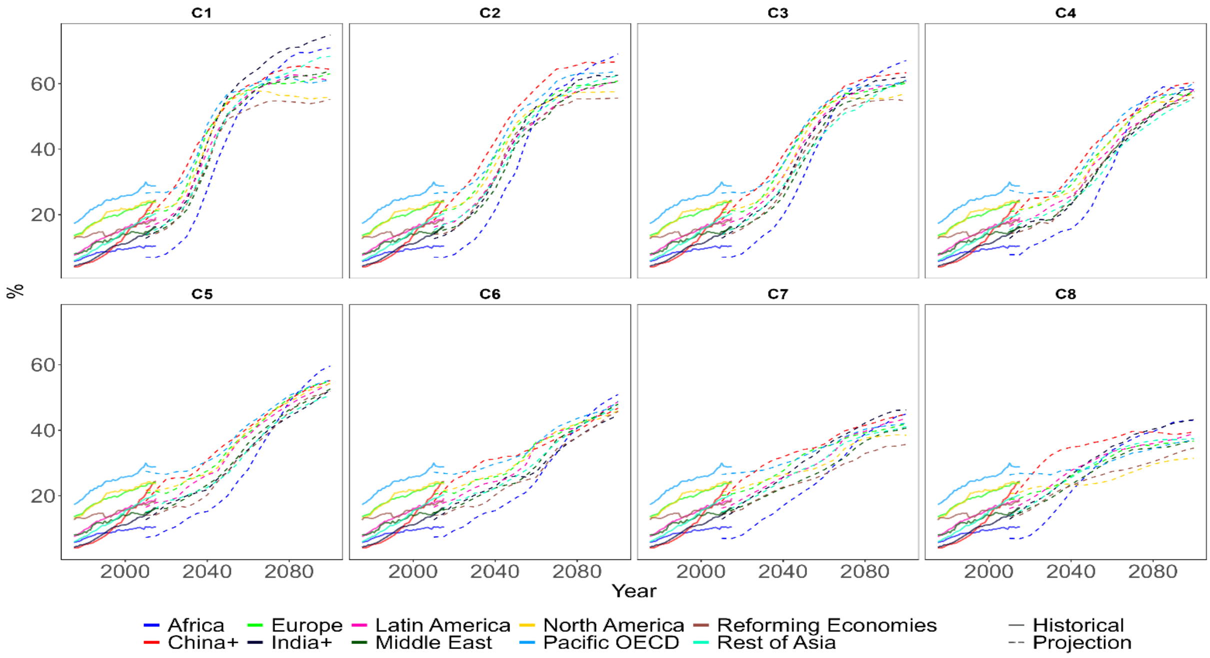


Figure 3c.I Trend and projection of electrification rate

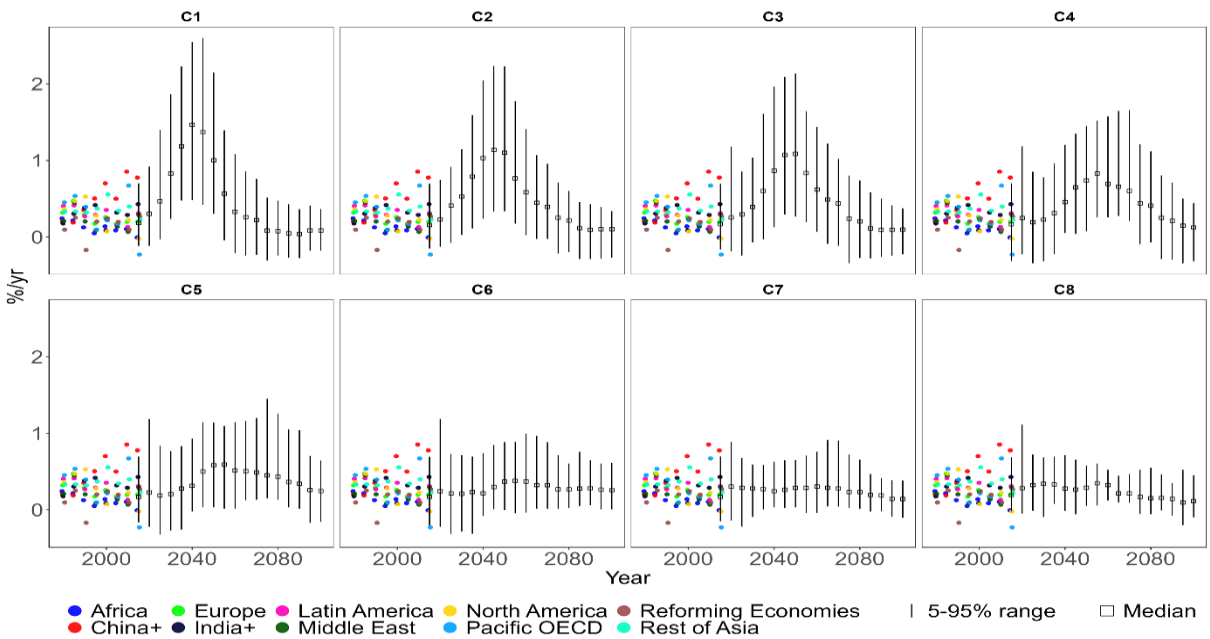


Figure 3c.II Trend and projection of electrification rate change

Figure 4a.I-c.I and Figure 4a.II-c.II show the distribution of (rate of) change of indicators, using density function and box-and-whisker respectively. The 2 vertical dashed lines in black represent the 5$^{th}$ percentile and 95$^{th}$ percentile observed from historical distribution. For Figure 4a.II-c.II, the meaning of box-and-whisker as well as dots is the same as Figure 1a-f, which is described in the beginning of section 4.1.2.

Figure 4a.I and Figure 4a.II show the density function and box-and-whisker diagram for rate of change of energy intensity. While historical change distributes comparatively evenly through both positive and negative sides, scenarios, regardless of categories, tend to concentrate on the negative side.

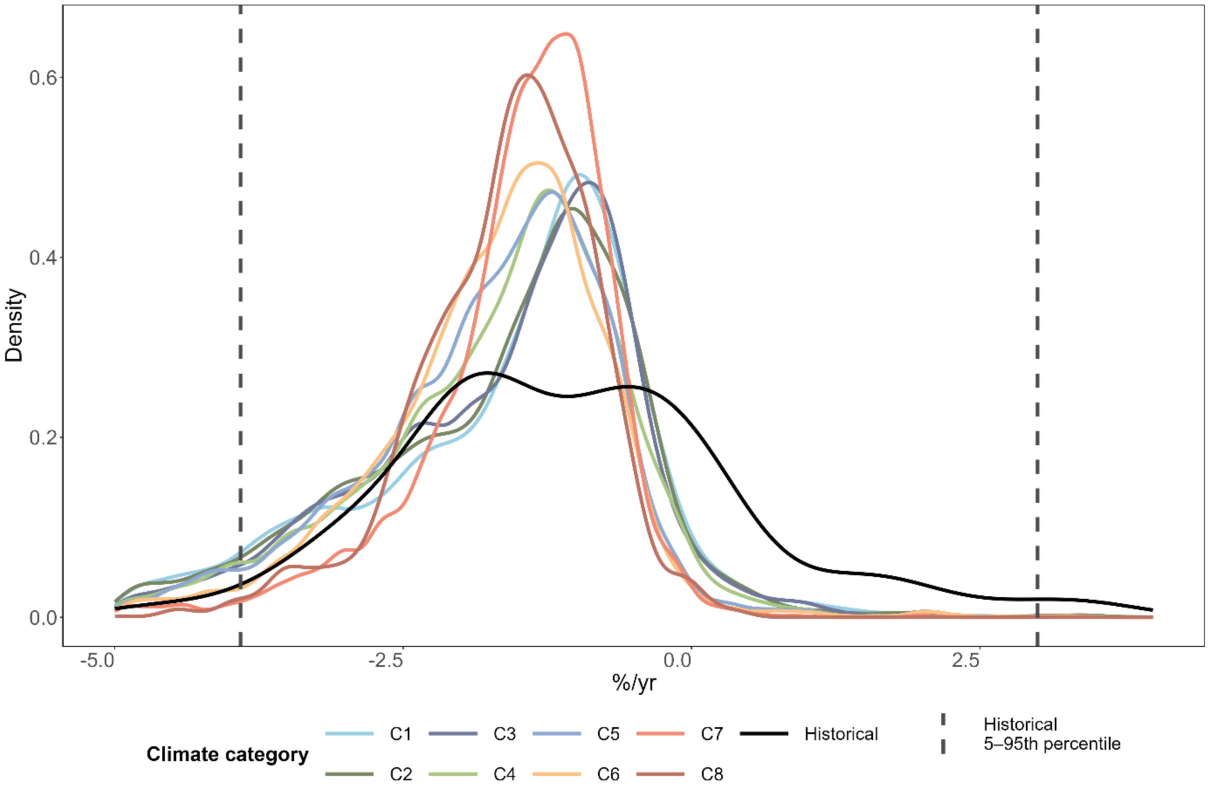


Figure 4a.I Distribution of energy intensity change (density)

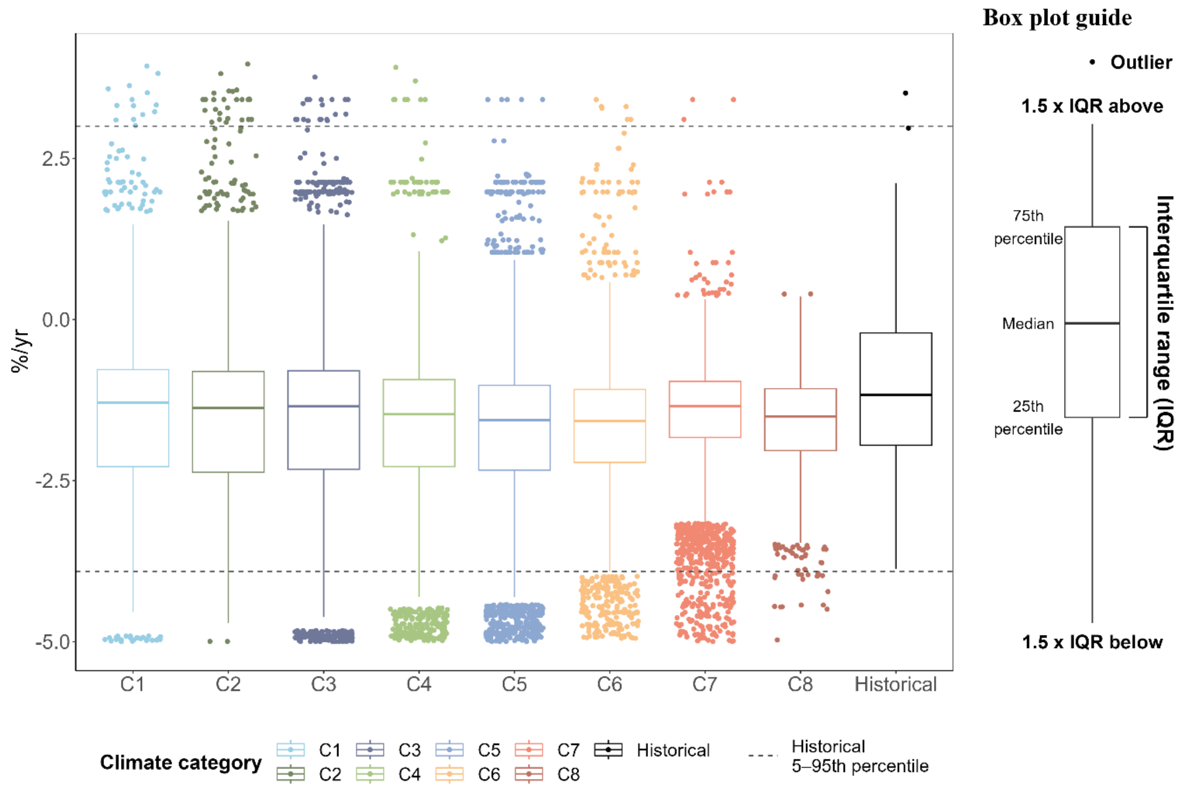


Figure 4a.II Distribution of energy intensity change (boxplot)

Figure 4b.I and Figure 4b.II show the density function and box-and-whisker diagram for carbon intensity change. Except for scenarios in C8, scenarios have their peak values at or to the left-handed side of that of historical change. For scenarios in C5 or lower, 1st quartile values lie below the 5th percentile of historical distribution.

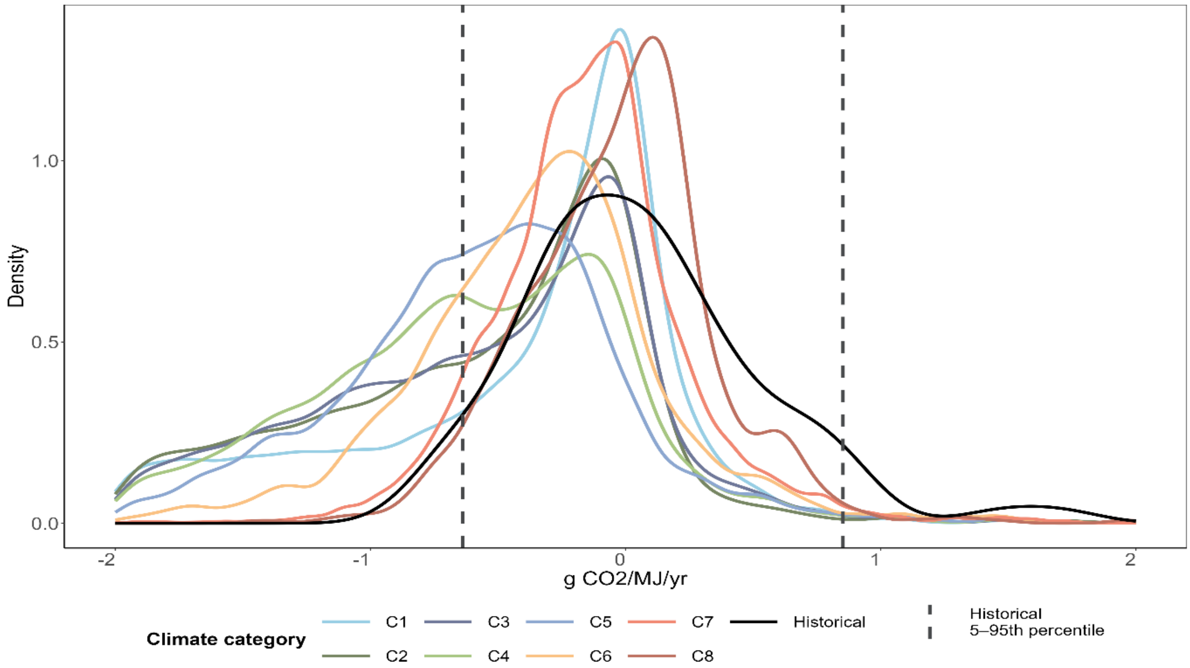


Figure 4b.I Distribution of carbon intensity change (density)

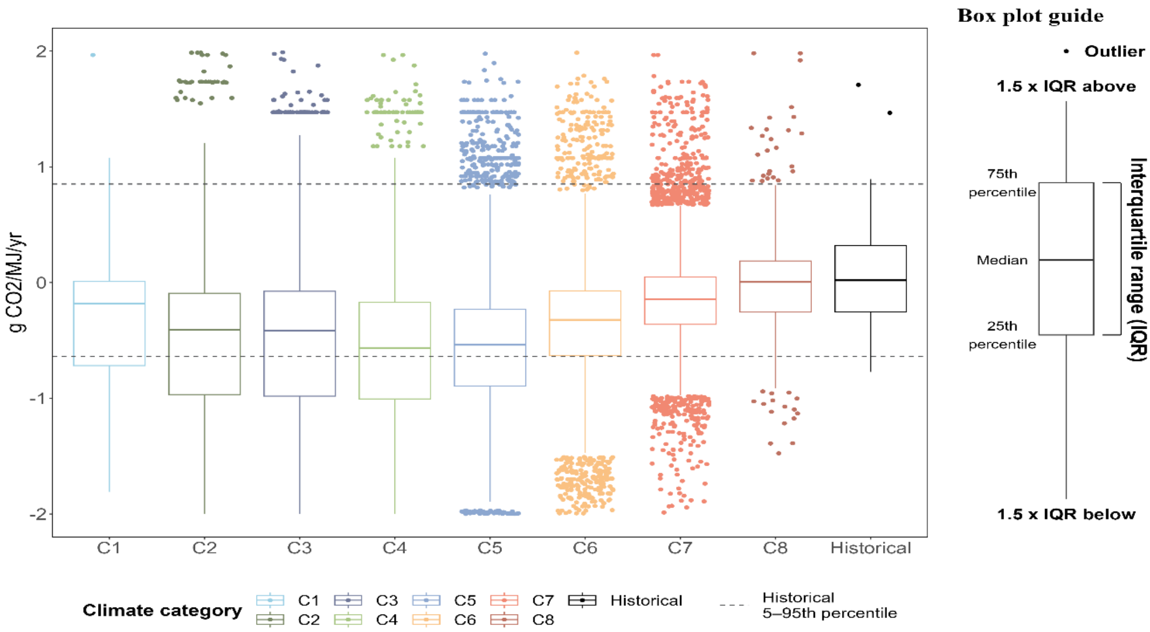


Figure 4b.II Distribution of carbon intensity change (boxplot)

Figure 4c.I and Figure 4c.II show the density function and box-and-whisker diagram for electrification rate change. The lower the categories are, the larger the positive tail of density function is. The same pattern is also seen from boxplot, which since C5 or lower, the 3rd quartile value of scenarios lies above the 95th percentile of historical distribution.

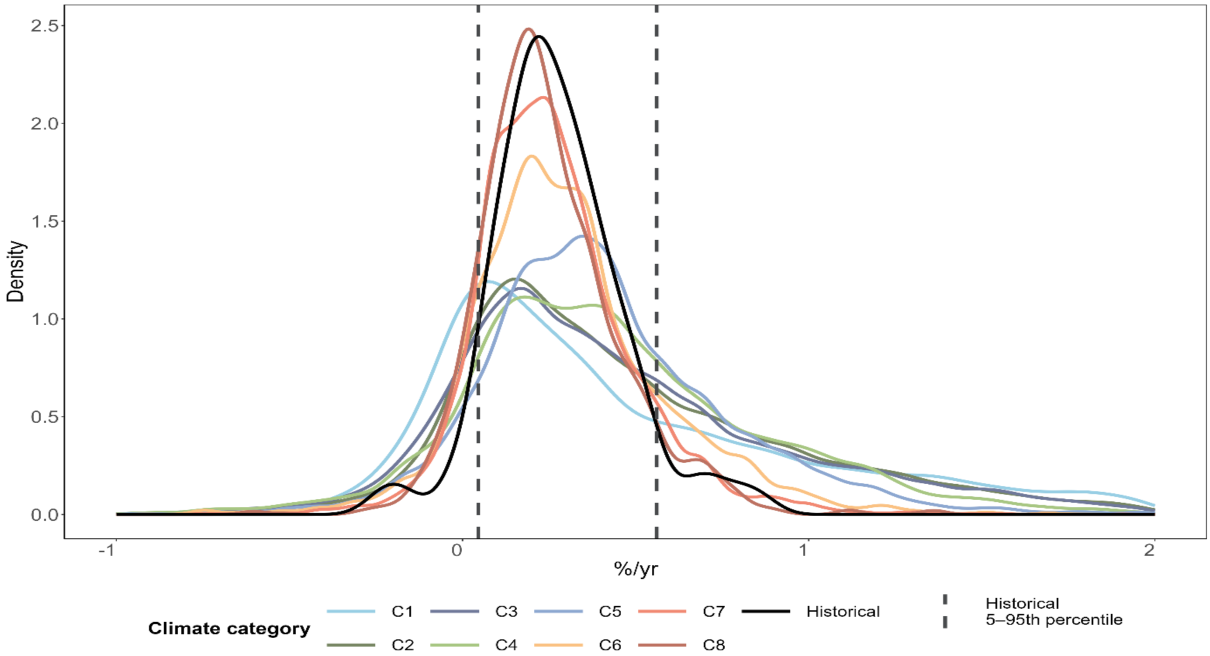


Figure 4c.I Distribution of electrification rate change (density)

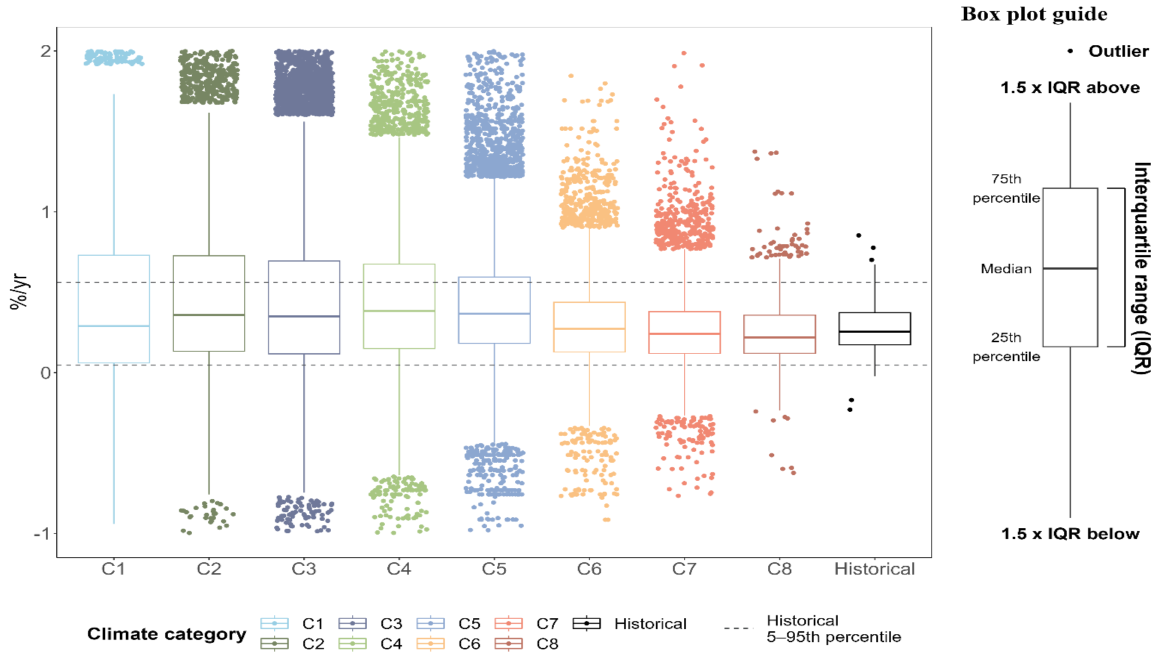


Figure 4c.II Distribution of electrification rate change (boxplot)

Table 6 shows the statistics of change of indicators, grouped by categories. Q1 and Q3 are the $1^{st}$ quartile and $3^{rd}$ quartile values respectively. q5 and q95 are the $5^{th}$ percentile and $95^{th}$ percentile values respectively. N stands for the total number of scenarios. Passing rate implies the proportion of scenarios falling within the historical $5^{th}$ and $95^{th}$ percentile values, in percentage.

Table 6 Statistics of indicators' change

| Category | Min | Q1 | Median | Mean | Q3 | Max |
|---|---|---|---|---|---|---|
| **Energy Intensity (%/yr)** | | | | | | |
| C1 | -10.81 | -2.44 | -1.34 | -1.74 | -0.80 | 5.19 |
| C2 | -10.39 | -2.49 | -1.42 | -1.75 | -0.82 | 4.91 |
| C3 | -11.38 | -2.42 | -1.38 | -1.71 | -0.81 | 6.18 |
| C4 | -9.92 | -2.36 | -1.50 | -1.75 | -0.94 | 7.73 |
| C5 | -10.30 | -2.39 | -1.59 | -1.80 | -1.03 | 10.00 |
| C6 | -9.92 | -2.24 | -1.58 | -1.74 | -1.09 | 11.05 |
| C7 | -10.07 | -1.84 | -1.35 | -1.50 | -0.96 | 3.41 |
| C8 | -5.38 | -2.04 | -1.51 | -1.60 | -1.08 | 0.40 |
| Historical | -6.11 | -2.08 | -1.17 | -0.85 | -0.14 | 12.28 |
| **Carbon Intensity (g CO2/MJ/yr)** | | | | | | |
| C1 | -11.23 | -1.17 | -0.26 | -0.67 | -0.01 | 8.32 |
| C2 | -8.23 | -1.21 | -0.50 | -0.74 | -0.11 | 8.32 |
| C3 | -13.76 | -1.11 | -0.48 | -0.66 | -0.09 | 8.32 |
| C4 | -6.59 | -1.08 | -0.59 | -0.66 | -0.18 | 8.32 |
| C5 | -5.42 | -0.92 | -0.55 | -0.58 | -0.23 | 8.32 |
| C6 | -6.37 | -0.64 | -0.32 | -0.34 | -0.07 | 9.40 |
| C7 | -5.42 | -0.36 | -0.14 | -0.13 | 0.05 | 8.32 |
| C8 | -1.48 | -0.26 | 0.00 | -0.02 | 0.19 | 1.98 |
| Historical | -0.77 | -0.25 | 0.02 | 0.06 | 0.32 | 1.71 |
| **Electrification Rate (%/yr)** | | | | | | |
| C1 | -1.81 | 0.07 | 0.31 | 0.51 | 0.81 | 4.44 |
| C2 | -1.54 | 0.14 | 0.37 | 0.50 | 0.75 | 4.69 |
| C3 | -3.13 | 0.12 | 0.36 | 0.47 | 0.71 | 10.46 |
| C4 | -1.67 | 0.15 | 0.38 | 0.45 | 0.68 | 3.65 |
| C5 | -1.17 | 0.18 | 0.37 | 0.41 | 0.60 | 3.26 |
| C6 | -0.91 | 0.13 | 0.27 | 0.30 | 0.44 | 2.24 |
| C7 | -1.87 | 0.12 | 0.24 | 0.27 | 0.38 | 2.84 |
| C8 | -0.62 | 0.12 | 0.22 | 0.25 | 0.36 | 1.37 |
| Historical | -0.23 | 0.17 | 0.25 | 0.27 | 0.37 | 0.85 |

Table 6(cont.) Statistics of indicators' change

| Category | q5 | q95 | N | Passing (%) |
|---|---|---|---|---|
| | | **Energy Intensity (%/yr)** | | |
| C1 | -4.63 | -0.07 | 11946 | 91 |
| C2 | -4.43 | -0.13 | 16350 | 92 |
| C3 | -4.23 | -0.12 | 35832 | 93 |
| C4 | -4.10 | -0.20 | 14398 | 94 |
| C5 | -4.01 | -0.40 | 22495 | 94 |
| C6 | -3.57 | -0.44 | 9139 | 96 |
| C7 | -3.12 | -0.48 | 11828 | 98 |
| C8 | -3.05 | -0.55 | 1980 | 99 |
| Historical | -3.91 | 3.00 | 80 | - |
| | | **Carbon Intensity (g CO2/MJ/yr)** | | |
| C1 | -2.70 | 0.31 | 12366 | 63 |
| C2 | -2.45 | 0.18 | 17550 | 54 |
| C3 | -2.14 | 0.30 | 42912 | 56 |
| C4 | -1.94 | 0.29 | 18738 | 51 |
| C5 | -1.67 | 0.33 | 28935 | 55 |
| C6 | -1.28 | 0.52 | 11799 | 72 |
| C7 | -0.74 | 0.52 | 13148 | 90 |
| C8 | -0.63 | 0.60 | 1980 | 94 |
| Historical | -0.64 | 0.85 | 80 | - |
| | | **Electrification Rate (%/yr)** | | |
| C1 | -0.17 | 1.84 | 12366 | 43 |
| C2 | -0.11 | 1.53 | 17730 | 50 |
| C3 | -0.17 | 1.51 | 42732 | 48 |
| C4 | -0.16 | 1.25 | 18198 | 53 |
| C5 | -0.09 | 1.07 | 28935 | 61 |
| C6 | -0.07 | 0.82 | 11619 | 70 |
| C7 | -0.04 | 0.67 | 12788 | 79 |
| C8 | -0.02 | 0.63 | 1980 | 81 |
| Historical | 0.05 | 0.56 | 80 | - |

Figure 5 shows the feasibility concern in each category. It is expressed as failing rate per indicator. Feasibility concern generally increases as the category number goes down from C8 to C1. Specifically, the increase is mostly contributed by electrification rate change, followed by carbon intensity change. The increasing trend of feasibility concern between C6 and C7, as well as C5 and C6, is significant. However, for scenarios between C4 to C1, carbon intensity change shows a decreasing trend while energy intensity rate of change and electrification rate change continues increasing.

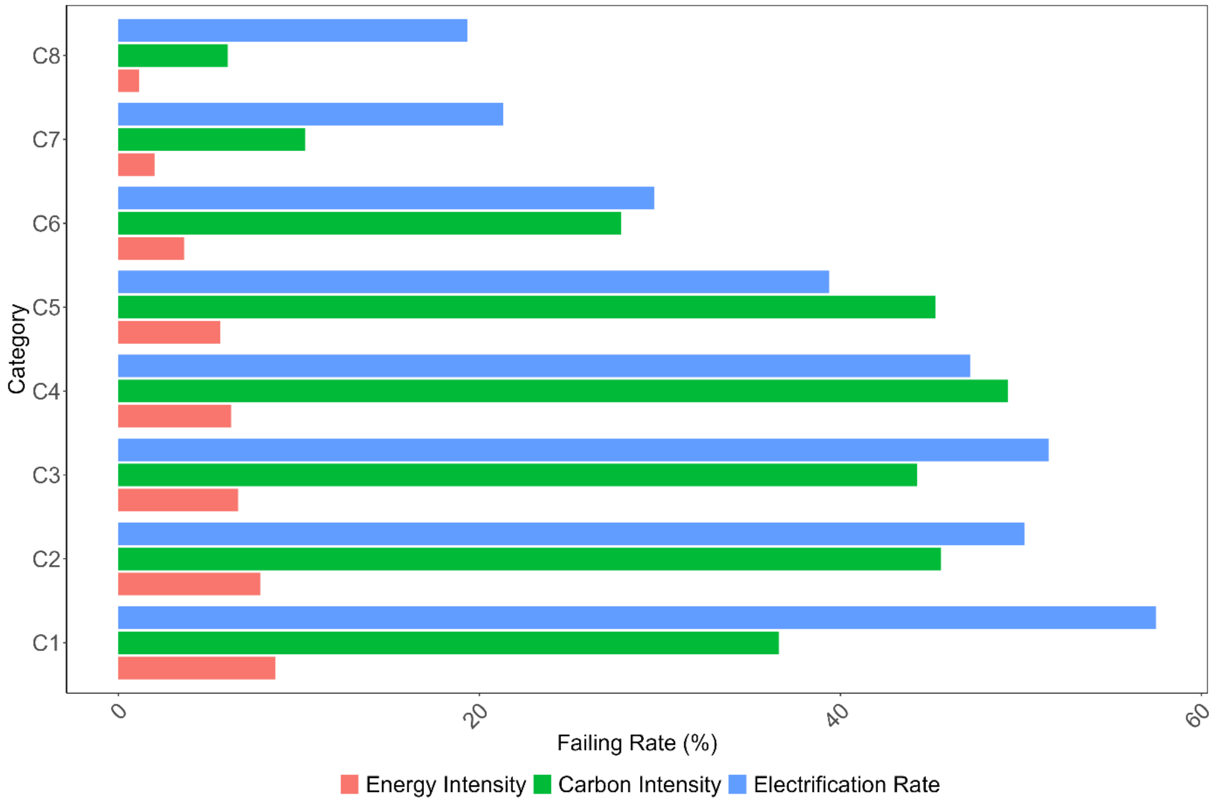


Figure 5 Feasibility concerns of scenarios

Figure 6a shows the feasibility concern of carbon intensity change per regions between C1 to C4, which consists of scenarios with limited temperature increase and thus may reveal a more practical concern. Black square indicates the overall failing rate which is equivalent to the values in Figure 5. The distribution of regional failing rate becomes more concentrated towards overall failing rate as categories moves from C4 to C1. Except C1, Pacific OECD is the region with highest failing rate, while Africa is the lowest counterpart. Most of the regions follow the decreasing trend of overall failing rate from C4 to C1, except for Middle East which shows a comparatively constant failing rate regardless of categories.

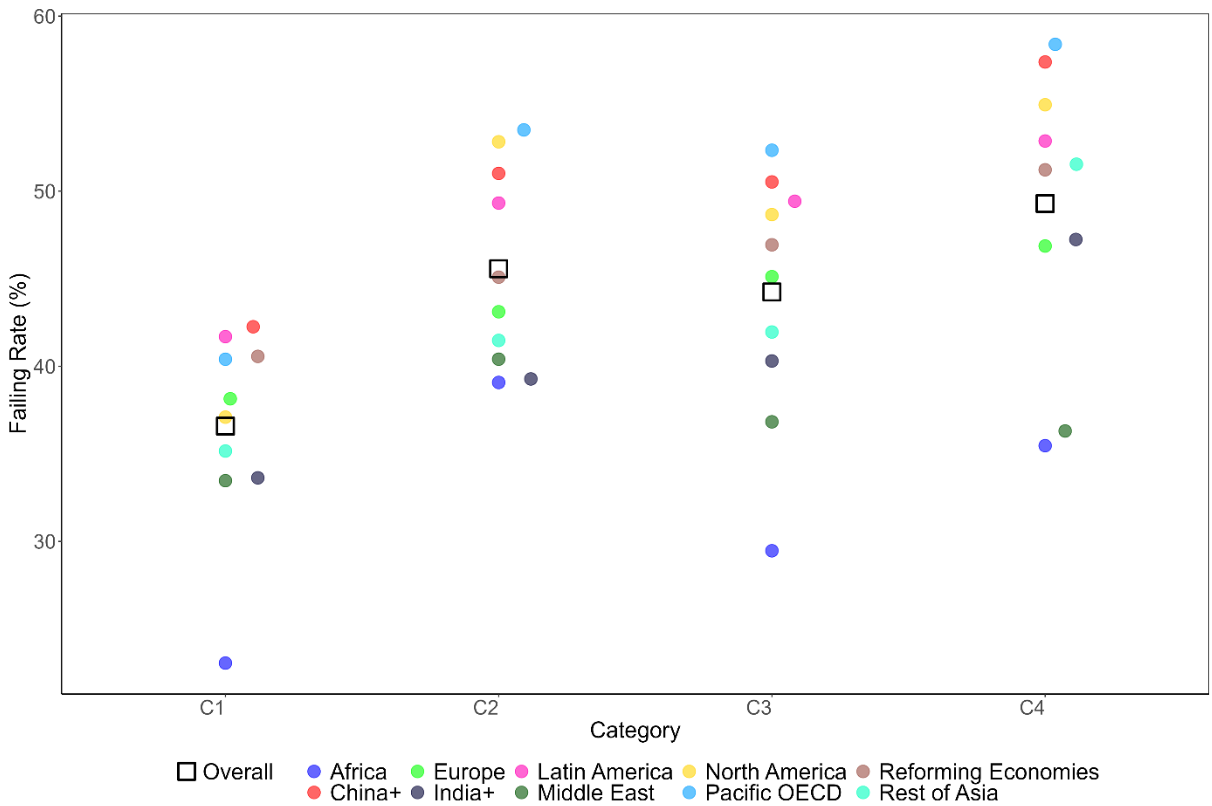


Figure 6a Regional feasibility concern of carbon intensity change

Figure 6b shows the feasibility concern of electrification rate change per regions between C1 to C4. The distribution of regional failing rate becomes more spread away from overall failing rate as categories moves from C4 to C1. Except C4, China is the region with highest failing rate, while Europe is around the lowest counterpart. Most of the regions follow the increasing trend of overall failing rate from C4 to C1, but Africa shows a comparatively constant failing rate regardless of categories.

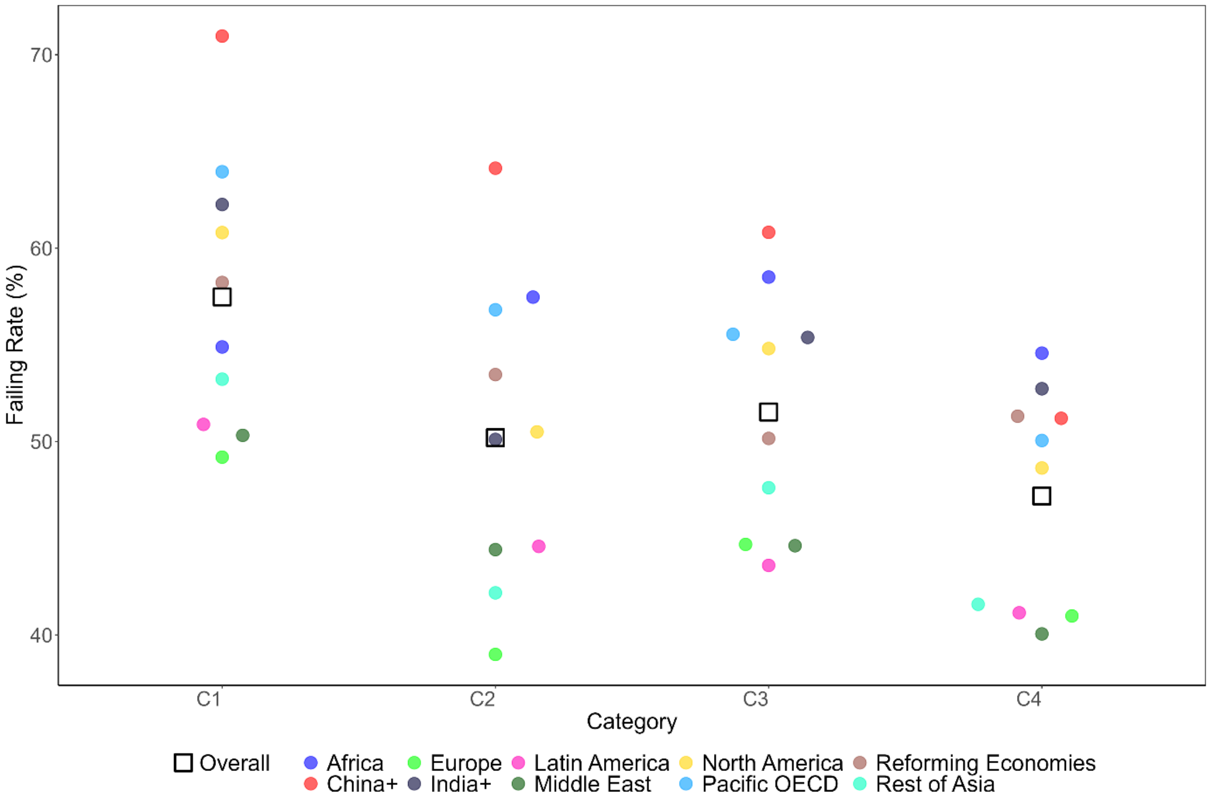


Figure 6b Regional feasibility concern of electrification rate change

## 4.2 Discussion

### 4.2.1 Difference between Global Vetting & Regional Vetting

The difference between global vetting & regional vetting can be discussed in 2 aspects, as shown in Figure 1a-f and Figure 2a-f, namely distribution of deviation from reference and passing rate. Generally, global scenarios show the most concentrated distribution of deviation, followed by OECD+EU and ASIA. The same ranking goes for passing rate as well, although exceptions are also found which will be discussed later. On the other hand, LAM, MAF and REF show the most distributed deviation and tend to rank lower in terms of passing rate.

The main reason for the difference between global vetting and regional vetting could be due to the inappropriate assumption and mechanism deployed in region adopted by IAMs with global scope. As shown in Annex III (IPCC, 2022), among 2266 scenarios which are being vetted, over 65% of them implement globally coordinated climate policies, either immediately or by 2030 onwards. In other words, a significant number of scenarios utilizes global carbon budget as the driver for mitigation, which is to be met by the aggregation of regional effort. When the use of global constraint couples with the uniform assumptions of technological deployment across regions, IAMs might fail to capture the variation of regional characteristic, like the responsiveness to carbon price and local constraints (Köberle et al., 2022).

On the other hand, the reason why OECD+EU and ASIA comparatively perform better than regions could be due to the traditional focus on major regions, for example, OECD+EU comprises of 67% of global electricity from nuclear power while MAF only comprises of only 0.7%. Such focus could also be the consequence of lack of data availability and inconsistency of data quality in less developed countries. This might suggest the limit of scenarios with global scope on the use with regional context, as Köberle et al. (2022) also suggests to refine the treatment of regional data construction when running IAMs with global scope.

In the following, result per indicator will be discussed.

About the $CO_2$ (EIP) emissions (Figure 1a & Figure 2a), vetting result shows the most polarized outcome. While most of the regions show relatively condensed distribution of deviation, the passing rate greatly differs between regions. The fact that deviation lying mostly above the upper boundary and low passing rate seen in LAM reveals a consistent overestimation of $CO_2$ emissions in LAM. REF also suffers from low passing rate, but when considering the percentage change of $CO_2$ (EIP) emissions (Figure 1b & Figure 2b), REF ranks the 3$^{rd}$ in passing rate. This suggests that changes over multiple years might have smoothed out the uncertainties of emissions in a single year. On the other hand, ASIA shows a different story, falling from 3$^{rd}$ place to 5$^{th}$ place in terms of passing rate. This could be attributed to the continuously underestimation of emissions in ASIA 2010 onwards, as revealed by Figure A1.a.

Vetting results of $CH_4$ emissions (Figure 1c & Figure 2c) is considered aligning with the result shown in Annex III of IPCC AR6 WGIII, that most of the scenarios have no concern reproducing $CH_4$ emissions. However, MAF and REF both show below-average passing rates, while showing opposite trends in distribution. Considering Russia being the largest natural gas provider, which could be a potential source for $CH_4$, scenarios may have the tendency to underestimate the momentum insider former Soviet Union's countries to reduce methane emissions.

On the energy side, for the vetting result of primary energy (Figure 1d & Figure 2d), the result for global vetting agrees with the result in Annex III, which also shows high passing rate. However, although the distribution is generally condensed, the passing rate is not as high as for the regional vetting. This highlights the potential of contrasting conclusion between global and regional vetting and requires further investigation for the reason underlying.

For vetting result of electricity produced by nuclear (Figure 1e & Figure 2e), the reason why MAF shows the most distributed deviation, among all vetting tests, could be due to the richness of fossil fuel in the Middle East. In the Middle East, fossil fuel-based power generation is preferable to nuclear power plants. The adoption of nuclear power plants is exceptionally low in the Middle East, only 0.07 EJ electricity generated from nuclear in 2019, while 10 EJ in that of the world. Therefore, given such small scale of electricity from nuclear in the Middle East, it is sensible that MAF deviates the most because scenarios assuming flexible technological scale-up would tend to project significant amount of fossil fuel power plants phasing out due to richness of nuclear potential, while other scenarios assuming sticky elasticity would suggest the level of nuclear adoption remains low and unchanged.

Lastly, vetting results of electricity from solar and wind (Figure 1f & Figure 2f), demonstrate an exceptional case that regional vetting performs better than global vetting. The reason why OECD+EU shows a slightly better passing rate than that of world could be attributed to the implementation of aggressive renewable energy targets, like what is seen in the EU. The fact that distribution of all regions lies towards the lower boundary shows that most scenarios underestimate the scale-up of renewable energy due to the recent cost drop. This also agrees with the result shown in Annex III, that the same vetting showed the worst passing rate among the other 8 historical vetting.

### 4.2.2 Comparison of Results of Feasibility Assessment

To begin with, the result of the feasibility assessment framework found in this study is compared with the results shown in literature (Nishimoto & Fujimori, 2023). In general, the results found in this study are considered aligning with the result in literature. The distribution of historical (rate of) change of indicators is considered similar, as the difference between means in literature and this study are all within 1 standard deviation (0.05 for energy intensity; 0.6 for carbon intensity; 0.1 for electrification rate), even though that the resolution of regions shrinks from 17, which is the native regional resolution used by AIM/Hub, to 10 defined by IPCC. Although the 90% coverage (i.e. $5^{th}$ percentile and $95^{th}$ percentile) for energy intensity ([-3.91, 3.00]) is found to be larger than that of in literature ([-3.36, 2.17]), this does not affect the result that energy intensity ranks the lowest in terms of feasibility concern (Figure 5) in scenarios except baseline (i.e. C7 & C8 equivalent), scenarios, which both this study and literature agree with. For baseline scenarios, while literature points out the feasibility concern is electrification followed by energy intensity, this study finds to be electrification followed by carbon intensity. On the other hand, for scenarios with limited temperature increase, electrification rate and carbon intensity are agreed by both this study and literature to be the major components of feasibility concern, with electrification rate being more concerning than carbon intensity in most of the cases. In addition, the increasing trend in feasibility concern between C8 to C4 mostly aligned with literature, which also shows an increasing concern from baseline scenarios to scenarios which are "well below 2°C", equivalent to C3. However, the increasing trend does not continue from C4 to C1 as carbon intensity drops. As scenarios in C1 and C2 are beyond the scope of literature, the result is hardly comparable to literature. The pattern of feasibility concern between C1 to C3 also contrasts with the result found by IPCC AR6 WGIII which is discussed in the following.

Compared the result of feasibility concern found in this study and the result of feasibility assessment from IPCC AR6 WGIII as shown in Chapter 3 (Riahi et al., 2022), common conclusion could be drawn on some parts, while contrasting parts are also found. As IPCC AR6 WGIII assessed the feasibility of scenarios between C1 to C3, their result shows that in near-term like 2030 the main components of feasibility concerns come from institutional, economic, and technological aspects. Although indicators used in this study might not imply direct interaction with institutional efforts, carbon intensity and electrification rate are considered with clear relationship with economic and technological aspects, which leads to the common conclusion that both this study and IPCC's result agree with. When accounting for feasibility concern in economic aspect in IPCC's analysis, the magnitude of carbon price increase and energy investment ratio compared to baseline scenarios are included as indicators to be assessed with. High carbon price raises the cost for carbon-intensive business, and greater energy investment scale induces ease in employing renewable energy technology which consequently reduces emissions. In other words, these two indicators have indirect impact on lowering carbon intensity. On the other side, when accounting for feasibility concern in technological aspect in IPCC's analysis, scale-up of electricity share in transport is used as one of the indicators, which might be a major contributor of overall electrification rate change. Therefore, it is concluded that both this study and IPCC's result have a common statement on the sources of feasibility concern.

However, in IPCC's result, a general increasing trend of feasibility concern is found between C3 and C1, which contrasts to the result shown in this study as mentioned before. As IPCC AR6 WGIII states, 80% of scenarios in C1 show feasibility concern of medium level or above in technological aspect in 2030, while C2 and C3 only have 30 to 40% in their counterparts. Also, for economic aspect, almost all scenarios in C1 show medium to high level of concern in 2030, but only half of scenario in their counterparts in C2 and C3. On the contrary, for scenarios between C3 to C1 as this study shows, feasibility concern in electrification rate only slightly increases (52% to 57%), while carbon intensity decreases significantly (44% to 37%). Such contrast could be due to the nature of evaluation method. As seen from Figure 3b.II, for carbon intensity change, scenarios with lower temperature increase (e.g. C1) require drastic change in near-term but soon return to neutral while scenarios with higher temperature increase (e.g. C4) require gentle but continuous change. Given that change of indicator is assessed with a 5-year interval, such short-term change might not be captured properly and thus carbon intensity appears to be decreasing. While most feasibility assessments studies focus on scenarios with stringent climate targets (Byers et al., 2023; van de Ven et al., 2023), the feasibility assessment framework adopted in this study might not be an appropriate approach.

In summary, the versality of the feasibility assessment framework based on historical change is proven by to some extent drawing common conclusion with the result seen from Nishimoto & Fujimori (2023) and Chapter 3 from IPCC AR6 WGIII (Riahi et al., 2022). However, limitations of the approach are found specifically on scenarios between C1 and C3, proposing a need to adjust the evaluation method to consider characteristic of scenarios.

### 4.2.3 Regional Variation in Feasibility Concern

As Figure 6a and Figure 6b reveal, regional variation is observed in feasibility concern of carbon intensity change and electrification rate change.

For carbon intensity change, regions like Pacific OECD and China+ are among the most concerning regions, while region like Africa are among the least concerning regions. The reason for that could be due to the gap between historical trend and scenarios' century-end projection. As seen from Figure 3b.I, Pacific OECD and China+ are regions with highest historical carbon intensity in 2010s, while Africa is the lowest. Therefore, Pacific OECD and China+ are projected with pathways of the most drastic change but Africa is comparatively required less effort to achieve the carbon intensity level demanded. Also, the reason why the distribution of regional failing rates changes from a more distributed manner to a more concentrated one from C4 to C1 is possibly due to the characteristics of scenarios. As Figure 4b.II shows, C1 shows a more condensed box area compared to C2-C4, revealing a more uniform magnitude of carbon intensity change is occurring in C1 across regions.

For electrification rate change, China+ is one of the most concerning regions, while Europe is the opposite. The reason for that could be found in Figure 3c.II, where historical electrification rate change of China+ (red dots) has been positioning at the top since 1995. In other words, using the 90% coverage of historical change as the criteria for feasibility assessment potentially treated China+ as outliers. Therefore, scenarios utilizing the historical electrification rate of China+ in recent decades as reference for its future projection might not be considered "feasible" under the context of the evaluation method. On the other hand, as Europe has been positioning in the middle of historical distribution, scenarios might tend to assume a milder pathway for Europe on electrification, and hence ranks the lowest in terms of feasibility concern. In addition, the reason why the distribution of regional failing rates changes from a more concentrated manner to a more spread one from C4 to C1 is possibly due to the characteristics of scenarios. In Figure 3c.II, it is observed that while scenarios in C4 project a steady pace of electrification, scenarios in C1 tend to project a significant shift up in near-term before 2050 but return to neutral or even slightly negative change. Such negative change could be due to the employment of a mix of power sources which limit temperature rise, for example biofuel. Similar pattern is also observed from Figure 4c.II, which C1 shows a more spread distribution towards both ends of cutting points compared to C2- C4.

# 5. Conclusion

## 5.1 Summary

In this study, an extended approach for scenario assessment is proposed, including the use of regional vetting and feasibility assessment based on historical change, and applied to AR6 scenario database. The difference between the extended approach and the existing approach adopted by IPCC AR6 WGIII highlights the effectiveness of the extensions. The main results are:

1. By adding regional vetting, an opposite conclusion is found compared to the current approach. While global level scenarios show high reproducibility for historical data like $CO_2$ emissions and primary energy production, scenarios at regional levels show distributed values and fail to achieve a similar level of passing rate. In addition, regions with comparable value to the world perform comparatively better than regions that are not. These results urge a review of the assumptions and mechanisms regulating regional projection in IAMs and suggest potential limitation of the utilization of global scenario on regional context.
2. By applying the feasibility assessment framework based on historical change, similar conclusion as seen from literature is drawn, regarding the major components of feasibility concern which are carbon intensity and electrification rate. This proves the versality of the framework. However, contrasting results, which is the decreasing trend in feasibility concern of carbon intensity, are observed in scenarios with limited temperature increases, calling for review on the evaluation method to consider characteristic of scenarios.

## 5.2 Limitations

Three major limitations exist in this study. Firstly, while regional vetting is adopted in this study, national vetting is not considered. Although national vetting might provide more relevant information for policy makers, this study adopts R5 regions due to the consideration of representativeness of models and amount of data. Secondly, as mentioned in section 4.2.2, the method used for feasibility assessment does not consider characteristics of scenarios, falling short to properly evaluate scenarios with limited temperature increase. Lastly, differences between IAMs, in terms of model's structure and assumptions, are not thoroughly considered in the discussion of varying pattern seen from feasibility concern. While van de Ven et al. (2023) states that structural differences between models could elaborate the different distribution of feasibility concern, it is beyond the scope of this study and left for future research.

# Bibliography


Brutschin, E., Pianta, S., Tavoni, M., Riahi, K., Bosetti, V., Marangoni, G., & Ruijven, B. J. van(2021): A multidimensional feasibility evaluation of low-carbon scenarios,*Environmental Research Letters*, *16*(6), 064069.

Byers, E., Brutschin, E., Sferra, F., Luderer, G., Huppmann, D., Kikstra, J. S., Pietzcker, R., Rodrigues, R., & Riahi, K.(2023): Scenarios processing, vetting and feasibility assessment for the European Scientific Advisory Board on Climate Change,*International Institute for Applied Systems Analysis,*.

Byers, E., Krey, V., Kriegler, E., Riahi, K., Schaeffer, R., Kikstra, J., Lamboll, R., Nicholls, Z., Sandstad, M., Smith, C., van der Wijst, K., Al - Khourdajie, A., Lecocq, F., Portugal-Pereira, J., Saheb, Y., Stromman, A., Winkler, H., Auer, C., Brutschin, E., ... van Vuuren, D.(2022): *AR6 Scenarios Database* (1.1) [Data set]Intergovernmental Panel on Climate Change.

Dhakal, S., Minx, J. C., Toth, F. L., Abdel-Aziz, A., Meza, M. J. F., Hubacek, K., Jonckheere, I. G. C., Yong-Gun Kim, Gfn., Pachauri, S., Tan, X. C., & Wiedmann, T.(2022): Emissions Trends and Drivers,In P. R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, & J. Malley (Eds.), *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*Cambridge University Press.

Edenhofer, O., & Kowarsch, M.(2015): Cartography of pathways: A new model for environmental policy assessments,*Environmental Science & Policy*, *51*, 56–64.

European Commission & Joint Research Centre(2023): *EDGAR (Emissions Database for Global Atmospheric Research) Community GHG Database* (8.0) [Data set].

Fujimori, S., Hasegawa, T., & Masui, T.(2017): AIM/CGE V2.0: Basic Feature of the Model,In S. Fujimori, M. Kainuma, & T. Masui (Eds.), *Post-2020 Climate Action: Global and Asian Perspectives* (pp. 305–328)Springer.

Guivarch, Céline, Le Gallic, T., Bauer, N., Fragkos, P., Huppmann, D., Jaxa-Rozen, M., Keppo, I., Kriegler, E., Krisztin, T., Marangoni, G., Pye, S., Riahi, K., Schaeffer, R., Tavoni, M., Trutnevyte, E., Van Vuuren, D., & Wagner, F.(2022): Using large ensembles of climate change mitigation scenarios for robust insights,*Nature Climate Change*, *12*(5), 428–435.

IEA(2022a): *Energy Balance 2022* [Data set].

IEA(2022b): *Greenhouse Gas Emissions from Energy 2022* [Data set].

IPCC(2022): Annex III: Scenarios and modelling methods,In C. Guivarch, E. Kriegler, J. Portugal-Pereira, V. Bosetti, J. Edmonds, M. Fischedick, P. Havlík, P. Jaramillo, V. Krey, F. Lecocq, A. Lucena, M. Meinshausen, S. Mirasgedis, B. O'Neill, G. P. Peters, J. Rogelj, S. Rose, Y. Saheb, G. Strbac, ... J. Malley (Eds.), *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*Cambridge University Press.

Köberle, A. C., Daioglou, V., Rochedo, P., Lucena, A. F. P., Szklo, A., Fujimori, S., Brunelle, T., Kato, E., Kitous, A., van Vuuren, D. P., & Schaeffer, R.(2022): Can global models provide insights into regional mitigation strategies? A diagnostic model comparison study of bioenergy in Brazil,*Climatic Change*, *170*(1), 2.

Lecocq, F., Winkler, H., Daka, J. P., Fu, S., Gerber, J. S., Kartha, S., Krey, V., Lofgren, H., Masui, T., Mathur, R., Portugal-Pereira, J., Sovacool, B. K., Vilariño, M. V., & Zhou, N.(2022): Mitigation and development pathways in the near- to mid-term,In P. R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, & J. Malley (Eds.), *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the*

*Sixth Assessment Report of the Intergovernmental Panel on Climate Change*Cambridge University Press.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., ... Wang, R. H. J.(2020): The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500,*Geoscientific Model Development*, *13*(8), 3571–3605.

Nishimoto, H., & Fujimori, S.(2023): An evaluation method of the long-term future climate mitigation scenarios based on the comparison with the historical experience. 歴史的に観察された変化との対比に基づく長期気候緩和シナリオの評価方法,*Japanese Journal of JSCE*, *79*(8), 22–00209.

Riahi, K., Schaeffer, R., Arango, J., Calvin, K., Guivarch, C., Hasegawa, T., Jiang, K., Kriegler, E., Matthews, R., Peters, G. P., Rao, A., Robertson, S., Sebbit, A. M., Steinberger, J., Tavoni, M., & Vuuren, D. P.(2022): Mitigation pathways compatible with long-term goals,In P. R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, & J. Malley (Eds.), *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*Cambridge University Press.

The World Bank(2023): *GDP deflator (base year varies by country) - United States* [Data set]. https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS?locations=US

van de Ven, D.-J., Mittal, S., Gambhir, A., Lamboll, R. D., Doukas, H., Giarola, S., Hawkes, A., Koasidis, K., Köberle, A. C., McJeon, H., Perdana, S., Peters, G. P., Rogelj, J., Sognnaes, I., Vielle, M., & Nikas, A.(2023): A multimodel analysis of post-Glasgow climate targets and feasibility challenges,*Nature Climate Change*, *13*(6), 570–578.

Wilson, C., Guivarch, C., Kriegler, E., van Ruijven, B., van Vuuren, D. P., Krey, V., Schwanitz, V. J., & Thompson, E. L.(2021): Evaluating process-based integrated assessment models of climate change mitigation,*Climatic Change*, *166*(1), 3.

# Appendix

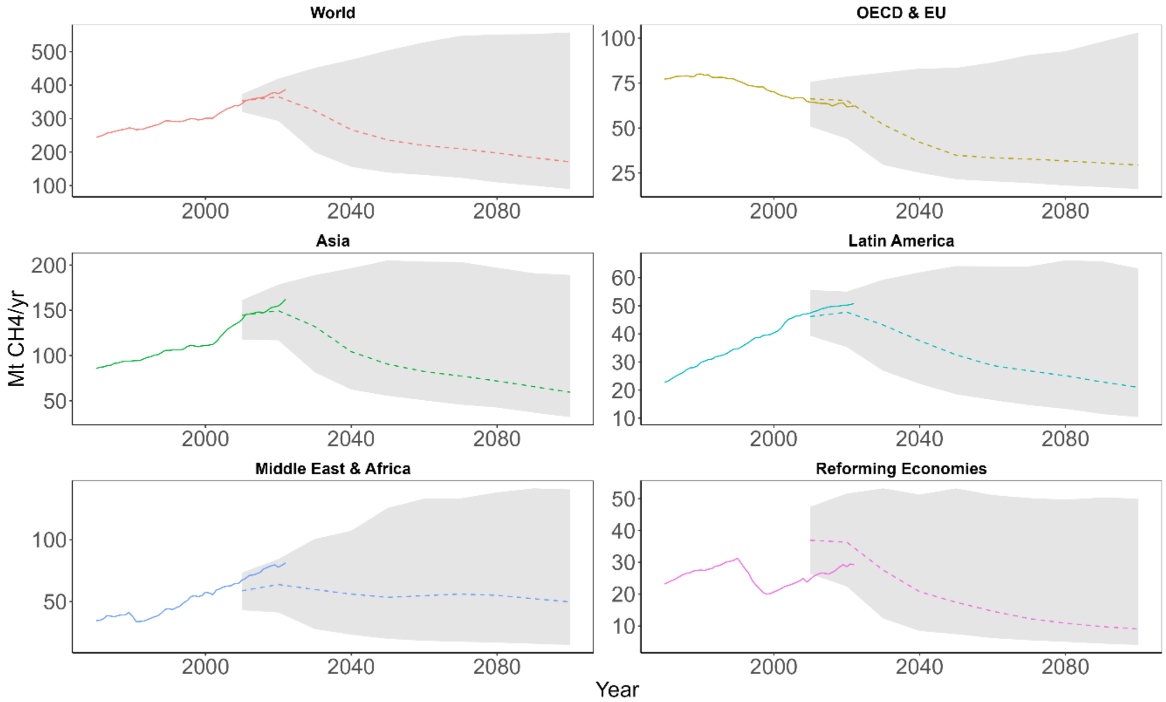


Figure A1.b Trend and projection of $CH_4$ emissions

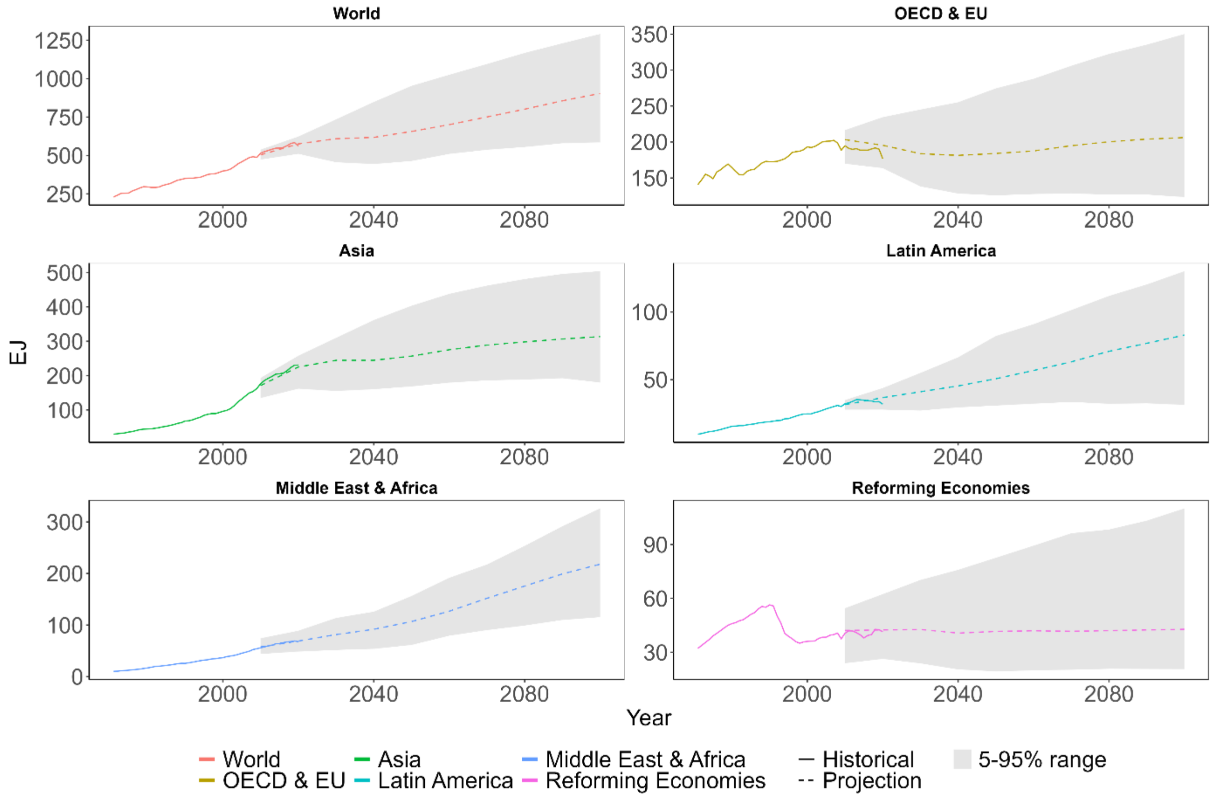


Figure A1.c Trend and projection of primary energy

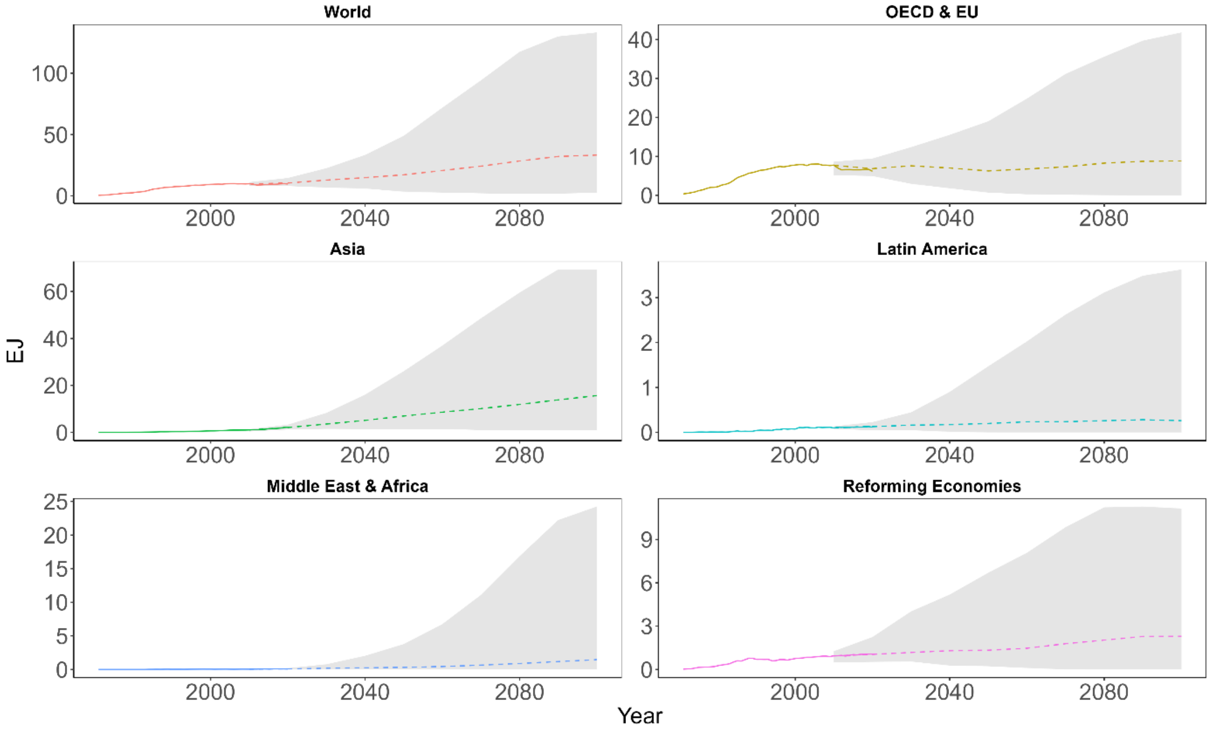


Figure A1.d Trend and projection of electricity from nuclear (1971-2100)

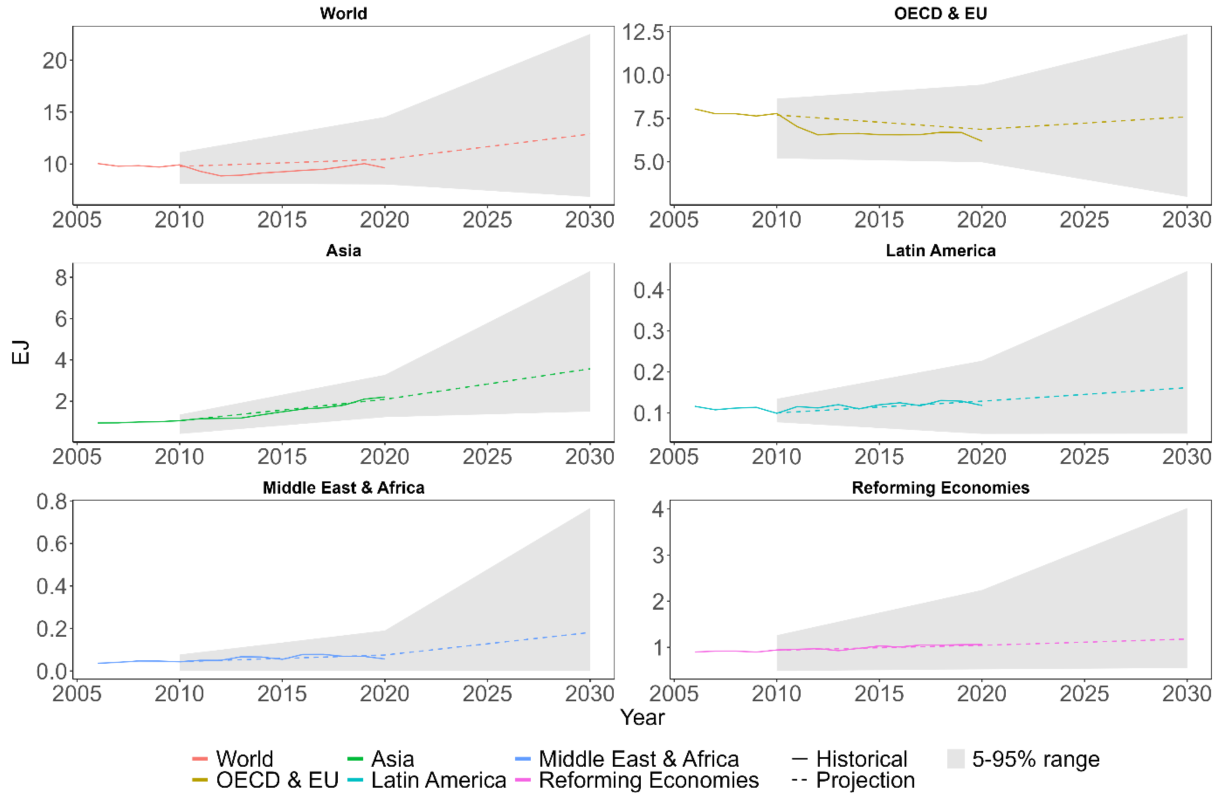


Figure A1.e Trend and projection of electricity from nuclear (2005-2030)

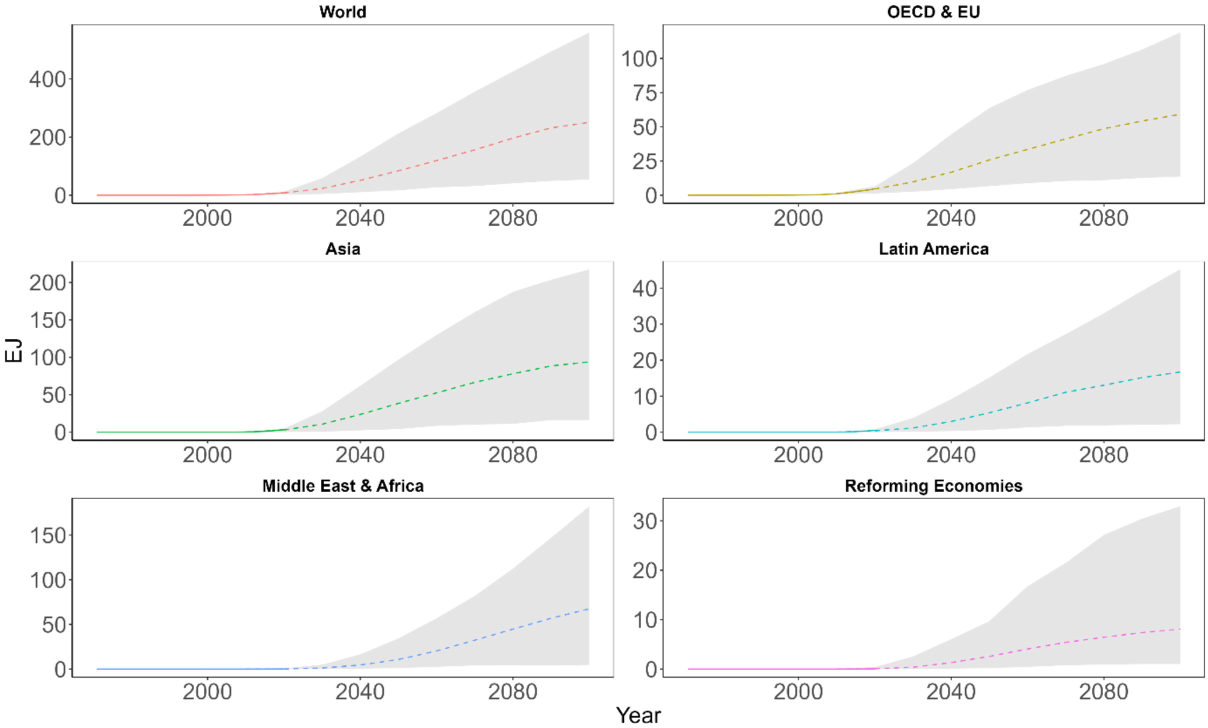


Figure A1.f Trend and projection of electricity from solar and wind (1971-2100)

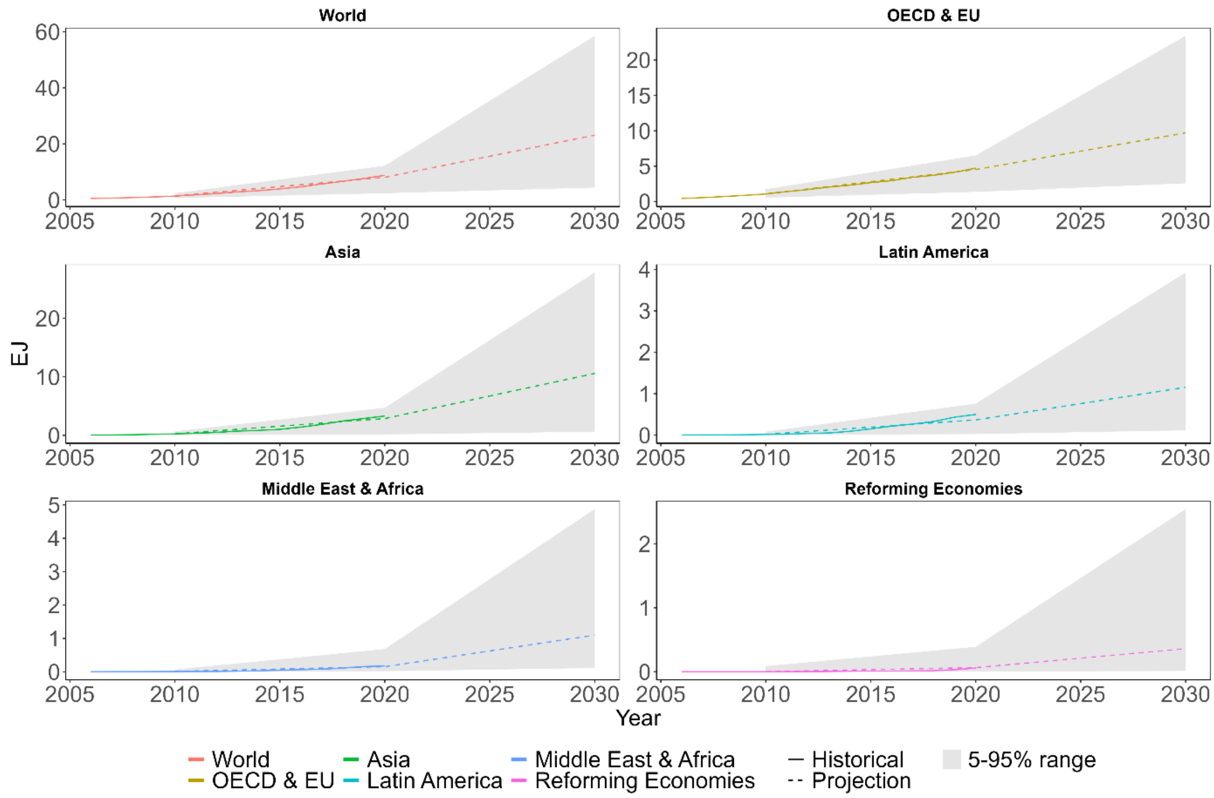


Figure A1.g Trend and projection of electricity from solar and wind (2005-2030)

## Acknowledgements

First and foremost, I would like to express my deepest gratitude to Professor Fujimori for his unwavering support and guidance throughout the entire process of writing this thesis. I am also indebted to Assistant Professor Oshiro for his invaluable advice and insights during the research process, which served as a constant source of inspiration. Special thanks are due to Assistant Professor Vishwanathan for her meticulous reviewing, which undoubtedly enhanced the quality of this thesis.

I am grateful to the members of the Mitigation Team for their continuous support and provision of relevant comments, which greatly contributed to the success of this research. I would also like to express my appreciation to my fellow lab members for their kindness and assistance, especially in creating a comfortable study environment for me as a non-native Japanese speaker.

Lastly, heartfelt thanks to my family and friends for their encouragement and moral support throughout this journey. Their belief in me has been a driving force in overcoming the challenges of this thesis.