ENVIRONMENTAL RESEARCH
LETTERS

PURPOSE-LED PUBLISHING™ Part of the ENVIRONMENTAL RESEARCH SERIES

LETTER • OPEN ACCESS

An exploration and evaluation framework for climate change mitigation scenarios with varying feasibility and desirability

To cite this article: Hamish Beath *et al* 2025 *Environ. Res. Lett.* **20** 074020

IOP Publishing Environ. Res. Lett. 20 (2025) 074020 https://doi.org/10.1088/1748-9326/addd35

ENVIRONMENTAL RESEARCH
LETTERS

RECEIVED
29 August 2024

REVISED
5 May 2025

ACCEPTED FOR PUBLICATION
27 May 2025

PUBLISHED
6 June 2025

LETTER

An exploration and evaluation framework for climate change mitigation scenarios with varying feasibility and desirability

Hamish Beath[1,*], Shivika Mittal[2], Robin Lamboll[1] and Joeri Rogelj[1,3,4]

[1] Centre for Environmental Policy, Faculty of Natural Sciences, Imperial College, London, United Kingdom
[2] CICERO Center for International Climate Research, Oslo, Norway
[3] International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[4] Grantham Institute—Climate Change and Environment, Imperial College, London, United Kingdom
[*] Author to whom any correspondence should be addressed.

E-mail: Hamish.beath16@imperial.ac.uk

Keywords: feasibility of climate change scenarios, energy transition, climate mitigation scenarios, societal transitions, integrated assessment modelling, scenario analysis

Supplementary material for this article is available online

Abstract

Ensembles of climate change mitigation scenarios present users with a collection of strategies for limiting global warming. These strategies may differ in their associated feasibility challenges, mitigation co-impacts, and ultimately their relative societal desirability. Understanding these scenario characteristics is therefore crucial when scenarios are used to inform strategic decisions. One approach to enhance this understanding is to establish scenario archetypes and select contrasting illustrative scenarios from a larger ensemble. We present a new multidimensional framework for the systematic comparison of scenarios at the global or regional level. We illustrate the framework with comparisons in seven dimensions: economic feasibility, mineral resource availability, impacts on societal resilience, near-term scenario robustness, environmental sustainability, interregional fairness, and speed of societal transformation. Using cluster analysis, the framework can be used to select a group of illustrative scenarios with contrasting scores across the dimensions. Beyond the selection of scenarios, our exploration and evaluation framework also allows the identification of gaps in the scenario space that may be of interest but are not covered by the literature. We demonstrate these use cases by applying our framework to a set of mitigation scenarios that limit warming to 1.5 °C. Our results show our framework systematically selects contrasting scenarios, with our illustrative pathways having diverging energy mixes and uses of carbon dioxide removal. Further, we highlight considerable regional differences in the distribution of indicator and dimension scores as a key area for further investigation.

1. Introduction

Understanding the consequences of climate change and strategies to mitigate and adapt to it is crucial for near and long-term policy planning. Quantitative scenario assessments play a key role in this process, providing valuable insights into short- and long-term energy system transformation and the resulting economic impacts [1, 2]. Assessment reports (ARs) of the Intergovernmental Panel on Climate Change (IPCC) extensively discuss insights from these scenarios and are used by policymakers to shape national and international climate policies [3, 4]. These large ensembles of mitigation scenarios have also been utilised in innovative ways by a range of users to suit their specified purposes [5].

There have been recent critiques regarding missing scenarios, policy relevance, and implications of these scenarios [6, 7]. Several studies have highlighted that only selected scenarios are presented and analysed by the IPCC authors, from a potentially infinite array of possible futures aligning with different climate targets [5, 8]. Creating every conceivable scenario is impossible. However, it is important to identify plausible gaps so that modellers can try to address them, extending the usefulness of the literature at large.

In addition, it can be challenging for users to extract relevant information from large collections of scenarios, such as the scenario databases compiled in the context of IPCC assessments [9, 10]. One method consists of selecting a curated set of specific scenarios that aligns with the user's objectives [5, 11]. This approach is used in the Special Report on Global Warming of 1.5° and the most recent Working Group 3 contribution to the IPCC AR6 [3, 4]. The authors selected illustrative pathways to highlight the consequences of diverse mitigation strategies, focusing on low demand, sustainability, and high resource and energy usage dimensions.

Understanding the feasibility and desirability of scenarios adds a further challenge. Brutschin *et al* systematically evaluate scenarios across feasibility dimensions, offering insights to users about implementation challenges and areas of potential concern [12]. Systematic evaluation of scenarios, considering not only feasibility but also relative *desirability* could broaden the usefulness of large scenario collections [13]. Even if assessments of desirability include a certain degree of subjectivity, providing an approach to formalise desirability considerations can improve the transparency of scenario exploration and evaluation. Structured scenario assessment can also enable better scenario selection and allow users to identify gaps in the scenario space where modellers could co-produce new scenarios to fill those gaps and help communicate scenario trade-offs from the onset of an analysis.

Evaluating scenarios on feasibility or desirability criteria is crucial given the significant co-impacts and transition risks mitigation pathways include. For instance, mineral resource needs (e.g. rare earth metals) for a scenario may be far beyond existing production levels, or there may be projected levels of new technologies such as novel carbon dioxide removal (CDR) approaches that are unlikely to be feasible in the near-term [14–16]. Even if feasibility concerns associated with low-carbon technological solutions are overcome, the adoption of some technologies or policies outlined in scenarios may have impacts on human health, economic prosperity, or environmental degradation [17–19]. Some scenario characteristics can more easily be judged by absolute feasibility limits: e.g. whether we have the required resources to achieve the energy system transformation outlined in a scenario. Other characteristics may not necessarily raise feasibility challenges; however, they may be less *desirable* relative to other options [13]. For example, higher mineral use in one scenario over another may be within feasible limits, but less *desirable* in terms of implied environmental damage. Exploration and evaluation of scenarios can guide users of scenario ensembles to make informed choices on appropriate scenarios to suit their needs. Previous studies have focused broadly on feasibility comparisons or constraints, rather than the desirability of mitigation scenarios [12, 20–25]. Tank *et al* emphasise the need for the inclusion of desirability dimensions in systematic scenario assessment [13].

Given the influence of global mitigation scenarios on national and international policy, there is also a need to understand their regional implications. This can help to address the questions of fairness between countries, which are increasingly important issues blocking global climate negotiations [26, 27]. Feasibility and desirability evaluation of mitigation scenarios at the regional level would further enhance the usefulness of the mitigation scenario ensembles [4].

Guivarch *et al* highlight several quantitative and qualitative techniques for selecting scenarios from large ensembles to increase the transparency of the overall scenario selection process [5]. Here, we present a framework for systematically exploring and evaluating scenarios not just based on their feasibility, but also their desirability. We demonstrate how users can apply this framework in different ways and use it to select a diverse set of scenarios from a large ensemble. The framework is flexible, allowing users to add indicators or dimensions as needed. This work demonstrates the possible uses for a broader application of scenario assessment, and we explain possible use cases, by applying the assessment framework to 1.5° scenarios, considering both those with and without temperature overshoot (C1 and C2), within the AR6 scenario database. The following section (Methods) will outline the main components of the framework, and how we apply it to a set of scenarios. We will then present the application of the framework (Results), demonstrating use cases, including illustrative scenario selection and regional comparison. We end with a summary of the findings and a discussion of their significance.

2. Method

2.1. Development of an exploration and evaluation framework

We have developed a multi-dimensional framework to evaluate mitigation pathways across dimensions of feasibility and desirability. Subjectivity exists in the assessment of both feasibility and desirability. Ascertaining thresholds of absolute feasibility is based and can depend on scientific interpretation, risk tolerance or data sources, while assessments of desirability can differ depending on perspective, culture, and context. Here, we present our scenario results in relative terms, avoiding subjectivity in absolute thresholds. For desirability, we present a set of relevant indicators that have been identified in the literature concerning possible mitigation challenges. If necessary, a user of

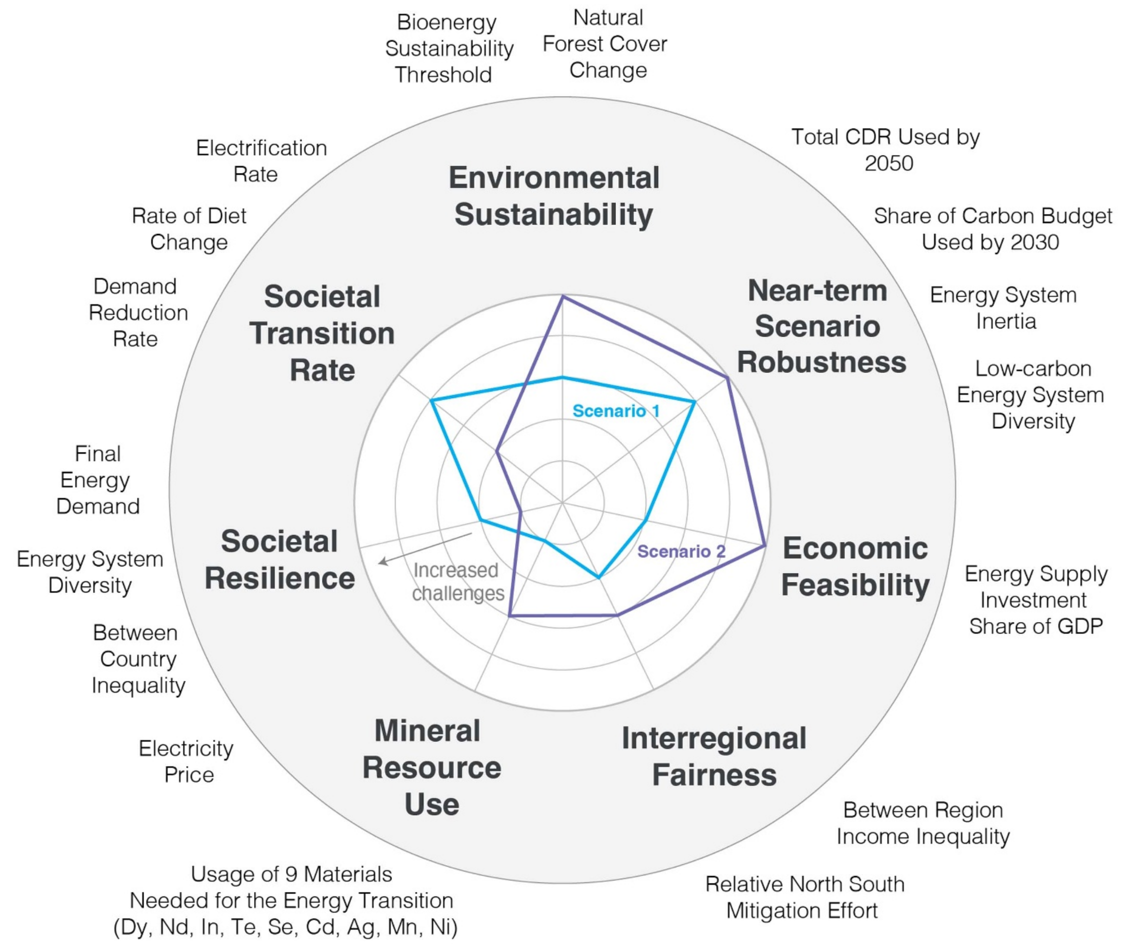

Figure 1. Overview of the exploration and evaluation framework composition, six dimensions and 22 indicators. Details on how each indicator is calculated and combined are given within the methods section and supplementary information. A scenario scoring higher on a given dimension means higher relative challenges in this dimension compared to other lower-scoring scenarios. Plotted scenarios 1 and 2 for illustrative purposes only, not real scenarios.

the framework can decide their relative importance by adjusting weightings depending on specific research questions.

We illustrate feasibility and desirability through the lens of seven dimensions: environmental sustainability, economic feasibility, mineral resource availability, impacts on interregional fairness, societal resilience, near-term scenario robustness and finally societal transition rate (figure 1). These dimensions are selected to show the applicability of the framework to both established and novel, emerging topics in the mitigation scenario literature. We here briefly introduce our seven selected dimensions, while table 1 provides an overview of the technical interpretation and implementation of each dimension.

Projections of *environmental sustainability* may vary widely between mitigation scenarios reaching the same temperature goal, as the use of technologies that require significant land use, such as bioenergy, varies widely [12, 18, 28]. Whilst scenarios that see higher levels of environmental degradation may be feasible, they are likely to be less desirable, all other things being equal.

Insights in *economic feasibility* and the costs of mitigation are key when exploring mitigation strategies [12, 19, 29, 30]. Governments' ability to mobilise resources for investment is limited and the desirability of outcomes needs to be considered together with the accompanying investments required.

Delivering a low-carbon energy transition requires resources. Rare earth minerals are key for several low-carbon technologies and *mineral resource availability* has been highlighted as needing further attention when considering energy system transformation strategies [14–16]. Available reserves and refining capacity of different mineral resources mean there are limits to their plausible use in the energy transition. More resource-intensive scenarios may also be less desirable due to extraction co-impacts, all other things being equal [31].

Interregional fairness is our next dimension. Some regions may have a much higher burden placed

Table 1. Justification for and overview of the method for indicators used in the exploration and evaluation framework. For further detail see supplementary information. The indicators have the letters F and/or D to denote whether they are primarily informing aspects of feasibility, desirability or both.

Indicator	Justification	Method	Weighting
Environmental sustainability			
Natural forest cover change (D)	Increased natural forest cover is considered better for habitats, biodiversity, and the water cycle. Natural forest land cover is a widely reported variable. This is considered a desirability indicator.	We use the land cover total, and land cover natural forest variables and look at the projected change between 2020 (base year) and 2100 for each scenario. We assume that increases in natural forest cover are good for environmental sustainability.	0.5
Bioenergy sustainability threshold (F)	Studies highlight that excess bioenergy use may have negative impacts on environmental sustainability [40]. Primary energy from biomass as it is a widely reported variable. This is considered a feasibility indicator.	Creutzig *et al* [26, 40] indicate that higher than 100EJ of bioenergy would be unsustainable. This indicator is binary, if breached in any year to 2100 = 1. For regional analysis, regional thresholds calculated by regional land area.	0.5
Economic feasibility			
Energy supply investment share of GDP (F, D)	Significant investment in energy system transformation is often required within mitigation scenarios, posing feasibility challenges. We adapt the approach taken by Brutschin *et al* [12]. There may be absolute limits to investment levels (feasibility), however, increased spending may also be less desirable.	Energy supply investment/GDP for each scenario projected to the year 2100, mean value compared with our baseline value. Baseline value used for this study is the mean value of this between 2001–2023, the years for which consistent data is available. Historical GDP from the IMF World Economic Outlook database [41], energy supply investment is taken from the IEA and BNEF [42, 43].	1
Mineral resource availability			
Usage of specific materials for renewables (F, D)	Material bottlenecks are highlighted as a possible constraint for energy transition. We assess 9 materials highlighted by existing work as needed for increased renewables capacity: Neodymium, Dysprosium, Manganese, Nickel, Silver, Cadmium, Tellurium, Selenium, and Indium [14, 16]. Solar and wind capacity growth are widely reported variables, enabling scenario assessment of these materials. This is considered primarily a feasibility indicator, although more intensive resource use may also be less desirable.	Stock and flow model built to estimate total availability of materials to 2050. The approach follows existing studies, with additional data from United States Geological Survey (USGS) [14, 16, 44]. The ratio of historical use of each material's use in solar/wind compared to the maximum five-year mean projected use up to 2050.	0 if usage is less than baseline, otherwise ratio is used (see supplementary information). Ratios summed; values normalised.
Interregional fairness			
Between region inequality (D)	Mitigation scenarios may have a positive or negative impact on between-region inequality. Improved income inequality means improved fairness between regions. GDP is a widely reported variable. This is considered a desirability indicator.	Cumulative GDP (at market exchange rate) and cumulative population taken for each R10 region 2020–2100, Gini values for each scenario calculated based on GDP/POP for each region. This gives a *between-region* Gini value for each scenario.	0.5

(Continued.)

Table 1. (Continued.)

North South relative mitigation effort (D/F)	Higher relative mitigation efforts in the Global South than in the Global North presents fairness challenges given less historical responsibility for climate change in the Global South. This is considered primarily a desirability indicator, although there may be feasibility limits to the Global South mitigation burden.	R10 regions are first split between Global South and Global North (see supplementary information). The remaining Global North and Global South shares of carbon budgets for each scenario are calculated (2030 budget, 50% 1.5 °C). We take the remaining North share minus the remaining South share, giving a scenario value. A lower value equals lower mitigation effort for the Global South, and therefore better fairness.	0.5
Societal resilience			
Primary energy diversity (D)	We assume higher diversity in energy sources will lead to higher societal resilience given a supply disruption impacting one source e.g. oil, or renewables is less likely to impact societies given the availability of others [45]. This is considered a desirability indicator.	We use the Shannon diversity index to provide a value for energy supply diversity for each scenario. The Shannon diversity index is widely used in ecology for species diversity but has also been used in the context of energy systems and energy security [46, 47]. We consider primary energy sources from oil, gas, coal, biomass, nuclear and non-biomass renewables over the scenario timeframe (2020–2100).	0.25
Final energy demand (D)	Resilience (in the form of redundancy) is far easier to achieve in a lower-demand scenario, and therefore we consider higher-demand scenarios to be worse for societal resilience. This is considered a desirability indicator.	We use cumulative final energy demand 2020–2100 for each scenario. For regional analysis, we use final energy demand over GDP, this is so that values are comparable between regions, and so that for historically underdeveloped regions, low energy demand and therefore lower levels of development are not considered beneficial for resilience.	0.25
Between country inequality (D)	A higher level of inequality between countries is likely to result in lower levels of resilience as it suggests that some countries have been left behind and remain poorer. A lower level of inequality suggests a faster catch-up by less economically developed regions and therefore likely to have greater capacity to deal with shocks. This is considered a desirability indicator.	We use country-level GDP and population data which are taken from the SSP the scenario is based on [48–50]. We calculate *between country* Gini values (differing only by SSP) using this data. For regional values, we calculate the coefficient for values for countries within each R10 region providing a *between-country* Gini value for each R10 region.	0.25
Electricity price (D, F)	Electricity prices vary by scenario depending on the degree and timing of energy system transformation. Higher electricity prices will put a strain on household budgets and reduce redundancy and available expenditure for other goods. Lower affordability may contribute to lower resilience. This is considered a desirability indicator, although there may be absolute feasibility limits to the electricity price tolerated by consumers.	Electricity prices are model-dependent and are impacted by initial prices used in the model. For each scenario in our set, we take the matching baseline scenario price and the scenario price at the decadal level to calculate the distance from the baseline. We use the mean value of the 2020–2100 values. When global values are not reported, regional values are used to construct a global price (see supplementary information).	0.25

(Continued.)

Table 1. (Continued.)

Indicator	Justification	Method	Weighting
Near-term scenario robustness			
Share of 2030 carbon budget used (D)	More rapid near-term action leaves in place more contingency for future uncertainties such as in the availability of carbon dioxide removal technologies or unexpected climate impacts. This is considered a desirability indicator.	Fraction of the carbon budget remaining by 2030 for a 50% chance of limiting warming to 1.5 °C, used for each scenario. The remaining carbon budget taken from Lamboll *et al*, scenario emissions are harmonised with emissions data from the Global Carbon Project [51–53]. Regional carbon budgets were calculated using country-level historical emissions and GDP data, using a polluter-pays approach (see supplementary information)	0.25
Low carbon energy system diversity (D)	Diversity in low-carbon primary energy options reduces the risk of disruption to mitigation plans if a single low-carbon energy source is impacted. Over-dependence on a single or fewer low-carbon primary energy sources may decrease the ability to change course under new and unexpected conditions. This is considered a desirability indicator.	We use the Shannon diversity index for primary energy supply from nuclear, biomass and non-biomass renewables. We take the cumulative primary energy from each source from 2020–2100 for each scenario and calculate the Shannon diversity value.	0.25
Energy system inertia (D)	Different types of energy infrastructure have different construction and investment lifetimes. An energy system that has infrastructure with a longer build and lifetime can be considered to have more inertia and will be more difficult to change in the face of the unexpected. This is considered a desirability indicator.	We estimate the average build time and infrastructure lifetime in the primary energy mix for oil, gas, coal, biomass, nuclear and non-biomass renewables.	0.25
Total CDR used by 2050 (F)	Given that CDR has not yet been deployed at scale, a reliance on large ramping-up of near-term CDR in mitigation scenarios is a risky approach. This is considered a feasibility indicator, and there are likely to be absolute limits to the levels of CDR that can be deployed by the mid-century, given the present technology deployment.	We take the total combined scenario assumed CDR for land-based CDR, BECCS, and direct air capture (DAC) (for scenarios that include it) up to 2050. For total CDR, to be able to compare this indicator between regions, we use total CDR over the regional land area.	0.25
Societal transition rate			
Rate of final demand reduction (F, D)	Many pathways include reductions in final energy demand. Depending on the speed of these demand reductions, there may be challenges in implementing them from the perspective of public acceptance, or negative societal impacts such as in health or economic growth. This is considered both a feasibility and a desirability indicator. Demand reductions may have absolute feasibility levels for acceptance and may be less desirable due to knock-on impacts.	We assess the final energy demand per capita changes at the decadal level and use the maximum decadal reduction value for each scenario as the indicator value.	0.33

(Continued.)

Table 1. (Continued.)

Rate of electrification (F, D)	To reduce emissions, decarbonisation pathways typically include some degree of electrification of final energy uses. This necessitates changes in consumer technologies and may see public acceptance challenges or infeasibility in implementation. This is considered both a feasibility and a desirability indicator. Electrification speed may have absolute feasibility limits and a higher rate may be less desirable due to consumer costs etc.	We use the maximum decadal increase in the share of final energy that is electricity as the indicator value.	0.33
Rate of diet change (F, D)	Shifting diet composition to reduce the share of food derived from livestock and increase the share derived from plants leads to reduced emissions [54]. Rapid changes in diet composition may lead to public acceptance challenges. This is considered both a feasibility and a desirability indicator. Public acceptance of dietary changes may have feasibility limits, whilst rapid changes in the food system may be less desirable due to knock-on impacts.	We use the maximum decadal increase in the share of food demand derived from crops as our indicator value.	0.33

on them in terms of mitigation efforts or may see worsening relative income levels due to climate policy [21, 32]. Whilst potentially feasible, less fair scenarios may be less desirable, depending on the perspective.

Strategies to decarbonise the global economy may positively or negatively impact resilience to climate or social disruption. Our framework considers *impacts on societal resilience* as one of the scenario exploration dimensions by looking at scenarios' relative performance in terms of inequality, or the use of energy technologies that may be prone to disruption by climate-related (e.g. extreme weather) or non-climate-related shocks (e.g. conflict) [32–35]. Scenarios driving lower levels of societal resilience are likely to be less desirable.

The *near-term robustness* of scenarios to changes in circumstances is a further dimension. Specifically, we define robustness as how a scenario performs under uncertainties and potential variations in future conditions. For example, how applicable a scenario remains if specific technologies or political actors do not behave as expected. Less robust scenarios may have a strong reliance on singular technologies for mitigation, particularly currently unproven technologies or those with long construction times or may rely on strong future responses due to delayed mitigation [36–38].

Our last dimension is the *societal transition rate.* Here, we seek to assess the degree to which scenarios contain societal changes that may be at a rate where they pose feasibility or desirability challenges. For example, whilst a rapid energy demand reduction may aid in reducing emissions, it could face public acceptance challenges or see negative co-impacts.

The framework is designed to be open-source and flexible and is implemented in Python code. Referred to as the IAMEE (IAM Scenario Exploration and Evaluation) framework, it is available on GitHub for the community to use and further adapt. It utilises the existing Pyam data frame structure used by the IAM modelling community for scenario analysis [39]. Databases in the IAMC data format used in the Pyam data frame structure can be queried for a set of scenarios and then run through the exploration and evaluation framework to produce scores and plots. The framework has seven dimensions, each with indicator(s) assessing feasibility and/or desirability (figure 1). Table 1 outlines a summary of each indicator, with further detail and limitations given in the supplementary information. Note that not all dimensions refer to the same time horizon, as for some considerations scenario characteristics in the near-term are more relevant than for others.

2.2. Weighting and combining of indicators

For each of our dimensions, we typically weight each indicator evenly. These weightings can easily be adjusted if users seek to adjust the emphasis. For dimensions with multiple indicators, indicators are normalised and combined to ensure each dimension has scenarios scoring between 0 and 1, with higher values indicating higher relative challenges in each dimension.

For the resources dimension, we take a different approach. If the usage for a specific material is less than the baseline level, it is not considered (weighted 0). When combining we sum all non-zero usage ratios before normalising to create our resources score, meaning that resources that are higher above baseline usage carry more weight. This ensures materials with significant increases in usage are emphasised, and resource-intensive scenarios are therefore better distinguished.

2.3. Application of framework

We apply the exploration and evaluation framework, at the global and regional level to scenarios from the IPCC AR6 Database that sit within the C1 and C2 temperature categories. These cover scenarios with >50% probability of keeping warming below 1.5 °C by the end of the century. C1 scenarios also limit peak warming below 1.5 °C with ⩾33%. For scenarios to be useful to our analysis, they must report all the required indicator variables (22 in total) at the regional level that we examine (R10 regions). This reduces the number of scenarios included in our analysis to 102 scenarios out of a total of 230 C1 and C2 scenarios; these are from the IMAGE, MESSAGE, REMIND and WITCH models. See supplementary information S1 for a full list of required database variables, their availability, and the scenarios included in the analysis and their respective models.

2.4. Scenario archetypes and selection of illustrative scenarios

The multidimensional framework can be used to identify scenario archetypes and select illustrative scenarios. Within a scenario set, there may be scenarios that fit into clear categories or archetypes based on their scores across the dimensions. We use K-means clustering to group the scenarios into clusters of archetypes and then illustratively select a single scenario from each cluster. To determine the optimal number of clusters, we use the elbow method (see supplementary information section S4). From our data, the optimal number is four clusters. We use the centroid values of each of our clusters as our scenario archetypes. For the selection of illustrative scenarios, we take the closest scenario to the cluster centroid. The result is one scenario from each cluster to serve as an illustrative scenario from each archetype.

3. Results

3.1. Indicator results overview

The results for each indicator allow users to examine differences across the scenario set in absolute terms (figure 2). If all scenarios have scores deemed acceptable for this indicator, then less emphasis can be placed on it when interpreting results. For instance, in the economic feasibility dimension (figure 2(a)), most scenarios project future energy supply investment to be a lower share of GDP than found historically. This suggests that this dimension may be less relevant from the perspective of overall scenario feasibility. However, there is still a considerable spread in the results, with scenarios with mean levels of energy investment from lower than 50%, to higher than 120% of historical levels. These values must also be seen in the context of the likelihood of greater investment needs in other areas alongside the energy system, such as for climate adaptation and healthcare provision.

3.2. Scenario results across the exploration and evaluation framework

Figure 3 illustrates how different scenarios, either across the whole set or the same scenario analysed by different models, perform on various dimensions after combining and normalising the results of each indicator within each dimension. Examining the results across all dimensions allows users to gain insights, consider how different model typologies, scenario assumptions, and climate targets influence the outcomes, and identify potential gaps in the scenario space of interest.

IAMs vary in structure, technological and regional resolution, parameterisation, and solution algorithm [55, 56]. Different scores across models provide insights into patterns that can be attributed systematically to model behaviour. Each model family in the scenario set (see figure 3(a)) has a different number of scenarios (see table S1 in the supplementary information) with the REMIND model family most represented (49 scenarios). The WITCH model family is the least represented (6 scenarios). The uneven distribution of models in the analysed scenario set makes it difficult to robustly quantify the differences in scores and attribute them to model differences, as does the fact that the scenarios are designed under different constraints. However, on the economic feasibility dimension, scenarios from the MESSAGE or WITCH model score much lower than those from the REMIND model family, evident from figures 3(a) and (b). Or, for the near-term scenario robustness dimension, the IMAGE scenarios score much higher than the scenarios by the other three models (figure 3(a)). When we look at scores on a single scenario (EN_NPi2020_400f) within our set that has been run by multiple models (figure 3(b)), model fingerprints can be observed. The scenario run with the REMIND model sees higher scores on most dimensions, with the greatest difference on the economic feasibility dimension, when compared to

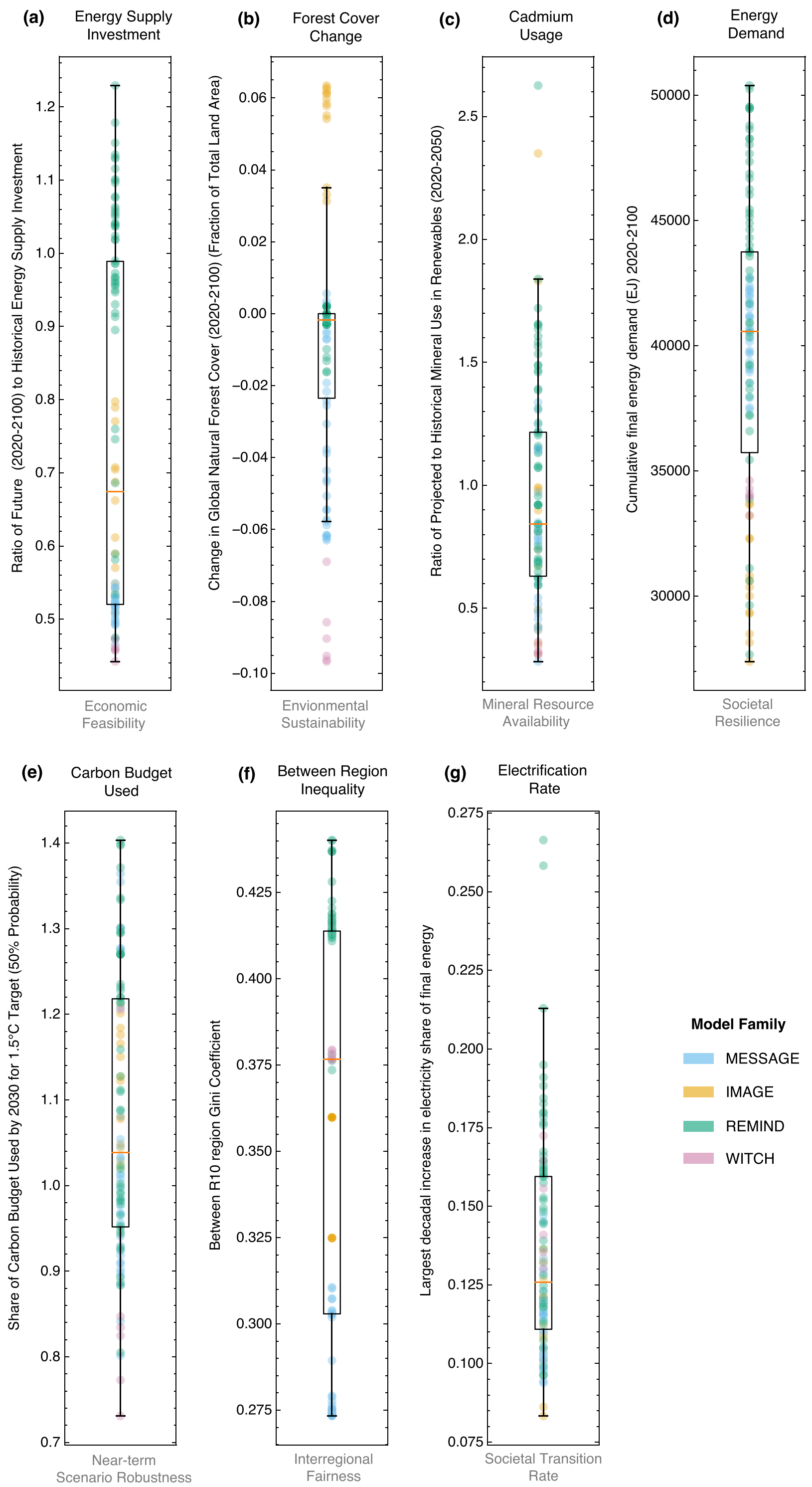

Figure 2. Example indicators from each of the seven dimensions. The distributions shown are the C1 and C2 calculated scenario values, for scenarios reporting all required variables at R10 coverage (see supplementary information). The colours of the dots represent scenarios from each of the four model families represented. (a) shows the distribution of ratios of future to historical energy supply investment as a share of GDP (economic feasibility dimension). (b) shows the distributions of change in forest cover (fraction of total land cover) (environment dimension). A higher score represents a greater change in absolute land area covered by forest. Sub-figure (c) shows the scenario distributions of ratios of projected to historical levels of usage of Cadmium in renewables up to 2050 (resources dimension). Sub-figure (d) shows the cumulative final energy demand of the scenarios in the set (resilience dimension). Sub-figure (e) shows the scenario distributions of the proportion of the remaining carbon budget used by 2030 to have a 50% chance of limiting warming to 1.5° (robustness dimension). Finally, sub-figure (f) shows the distribution of scenario scores for between region (R10) inequality: Gini coefficients. Sub-figure (g) shows the distribution of values for the maximum decadal increase in the electrification rate that occurs between 2020–2100 for each scenario (societal transition rate dimension). For each sub-figure, the median values and interquartile range are given by the central orange line on the boxplot and by the box exterior, respectively. The box whiskers represent the smallest and largest data points that fall within 1.5 times the IQR from Q1 and Q3.

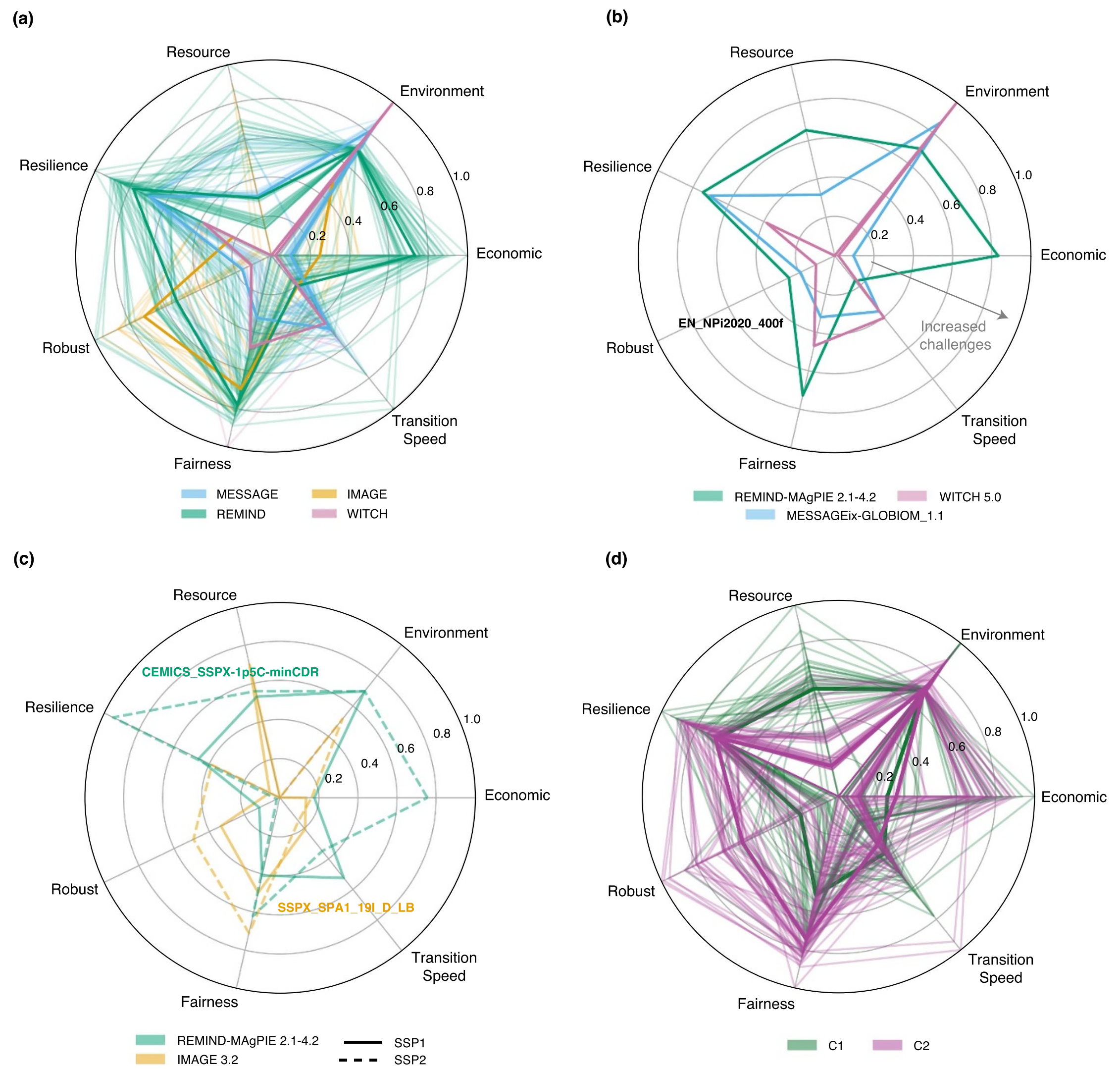

Figure 3. Plots showing the scenario set dimension scores. (a) grouping the scenarios by model family. (b) shows a scenario (EN_NPi2020_400f) run by three models from our scenario set, further highlighting the influence of model fingerprint on the results. (c) plots two examples of pairs of scenarios that are the same but run with SSP1 and SSP2, demonstrating possible SSP differences. (d) shows all scenarios within the set but coloured by temperature categories assessed (C1 and C2). Higher scores represent higher relative challenges.

the IMAGE and WITCH models. On the environmental and transition rate dimensions, IMAGE and WITCH have higher scores, with WITCH being the highest. The differences in figure 3(b) are similar to those observed in figure 3(a), indicating that there is a strong model influence on the results.

Figure 3(c) shows two pairs of SSP1 and SSP2 scenarios, created by the IMAGE and REMIND models, respectively. For the REMIND scenarios, it appears that apart from the near-term robustness and transition rate dimensions, SSP2 has higher relative challenges on all other dimensions. For the IMAGE scenarios, the scores are similar except for the resources and robustness dimension where SSP1 scores higher than SSP2.

Median scores (thick lines, figure 3(d)) for scenarios within the C1 and C2 temperature categories also show some clear patterns. For near-term scenario robustness and fairness challenges, C2 scenarios show higher concerns in the median than C1 scenarios. On the resources and transition rate dimensions, C1 scenarios have a higher median score. This is expected given scenarios without overshoot are likely to require a stronger, earlier push for renewables and more rapid societal transitions.

Low scores across different assessment dimensions might not be achievable at the same time. Scenario scores on our framework can be used to explore associations and possible trade-offs. A good starting point is to examine the pairwise correlations between these dimension scores (see figure 4). In this illustrative application, we see a strong correlation between certain dimension pairs, such as interregional fairness and near-term robustness (0.77), or interregional fairness and economic feasibility (0.66). Conversely, we also see negative correlations,

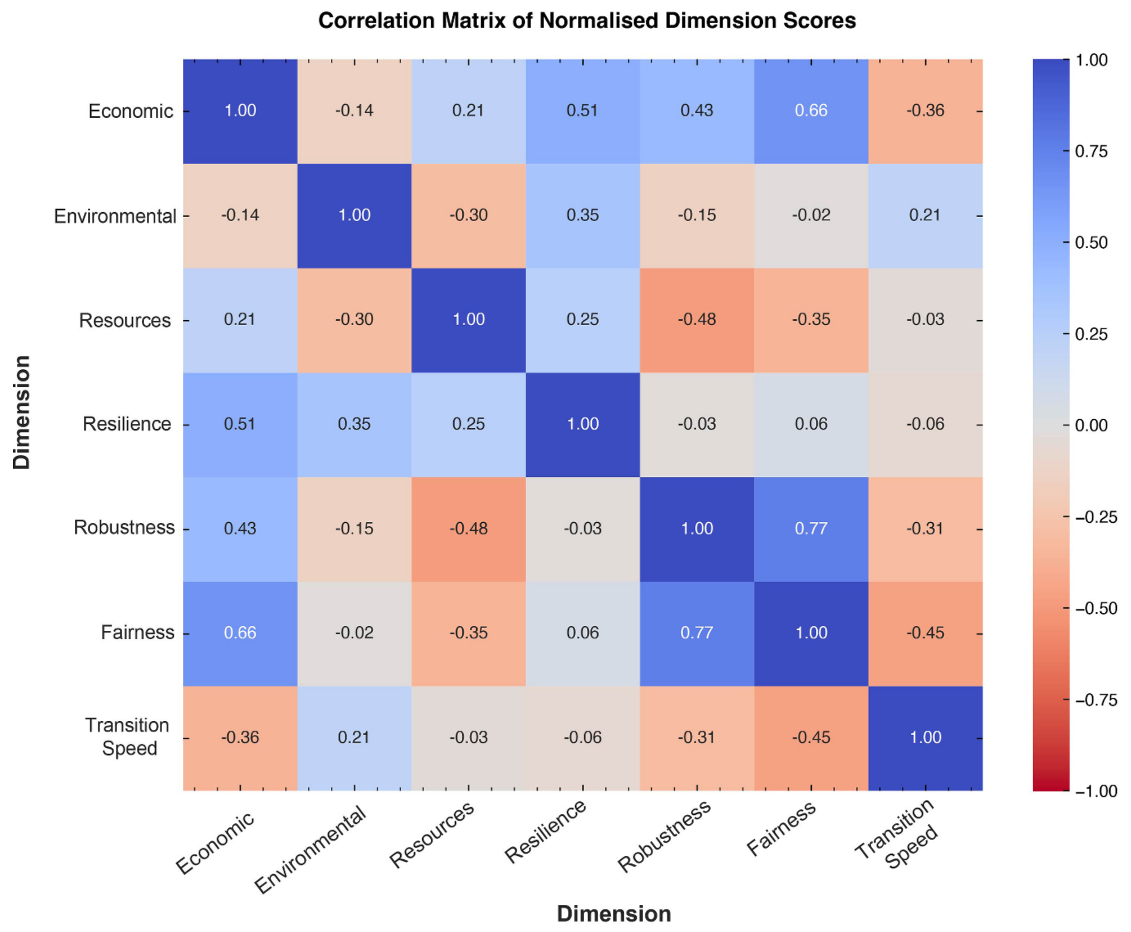

Figure 4. Correlation matrix showing correlations between pairs of dimensions, helping the identification of trade-offs and synergies in scenario characteristics.

indicating possible trade-offs, such as near-term robustness and resource use (−0.48), and interregional fairness with societal transition rate (−0.44). These scores and their correlations thus offer insights into potential trade-offs and co-benefits in mitigation pathways. By highlighting these relationships, the framework can support informed decision-making.

3.3. Exploring gaps in the scenario space

The outputs of the framework can also be used to explore whether there are gaps in the scenario space that it may be possible to address. Gaps can exist because modellers have not explored specific scenarios, or it may be that scenarios are infeasible within models. It is important here to acknowledge that our method provides no way to establish the cause, and further investigation would be needed. However, our framework does provide insight to modellers of where possible gaps exist and can be the impetus for further investigation and experimentation with modelling.

Users may want to explore the dimensions in a way that aligns more directly with their objectives. For example, some users may be interested in scenarios with a specific temperature outcome and simultaneously having lower relative challenges on one or more dimensions. Performing analysis to see the number of scenarios scoring lower on each pair of dimensions gives insights into potential gaps within our chosen scenario set (figure 5). Counting the number of scenarios that sit in the bottom 20% of scores for each combination of dimensions reveals that, for example, whilst there are many scenarios with low relative challenges on both the environmental sustainability and societal resilience dimensions, there are many pairs for which no scenario is found in the analysed ensemble. For example, a user may be looking for scenarios that have lower relative challenges for both fairness and societal resilience; however, this is missing from our scenario set (figure 5). A user may wish to explore whether such a scenario design is achievable.

3.4. Illustrative scenario archetypes

The framework can also be used to select illustrative, contrasting scenarios from a scenario ensemble to demonstrate the choices that exist in mitigation pathways reaching the same climate target. We demonstrate this here using the framework outputs and

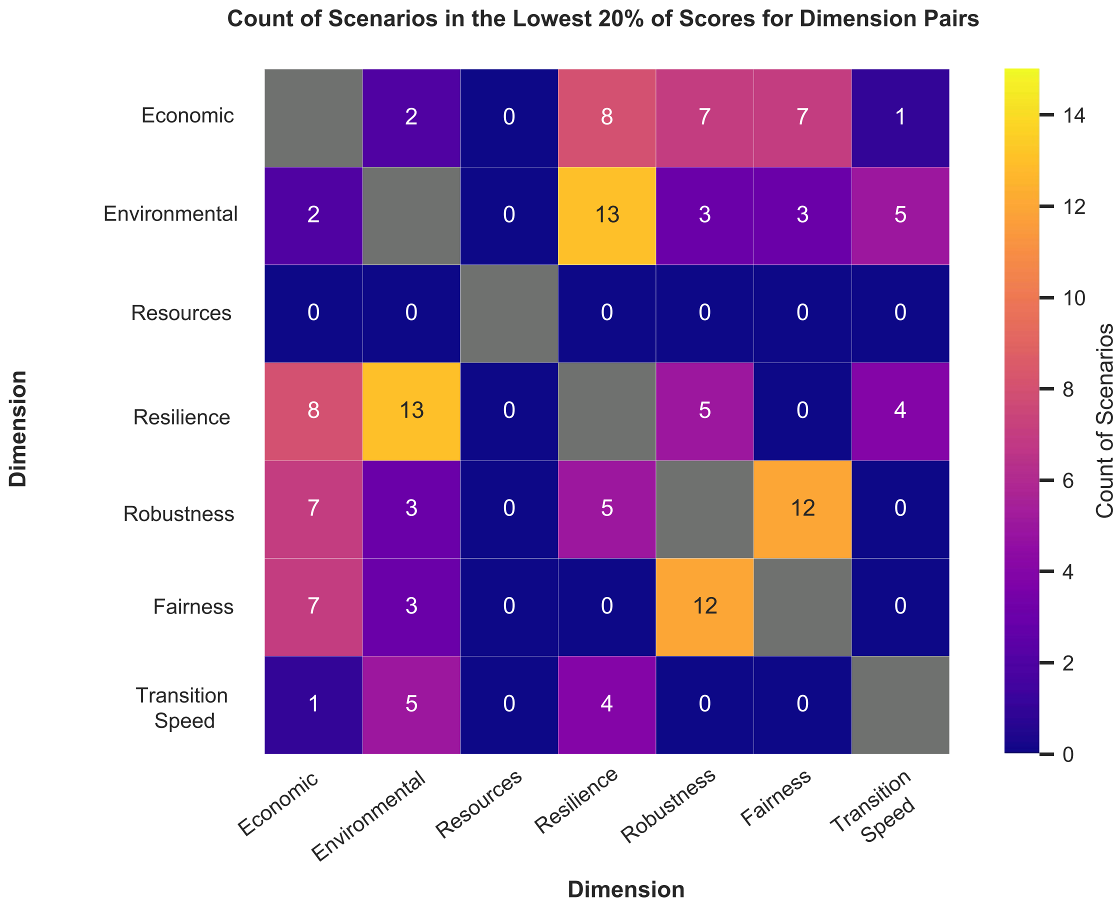

Figure 5. Showing the number of scenarios in the lowest 20% of scores for each pair of dimensions.

cluster analysis (see Methods). Cluster analysis across the seven dimensions provides us with four scenario archetypes (cluster centroids) and associated illustrative scenarios (scenarios closest to centroid) (figure 6). We show how energy system transformation and CDR use are related to the scores on the evaluation dimensions.

3.4.1. Scenario A–high negative mitigation impacts, low near-term risk. High demand, renewables and CDR use from 2050 onwards

This scenario archetype is characterised as having moderate to high scores across resilience, resource, environment, economic and fairness dimensions, with low scores on the robustness and societal transition rate dimensions (figure 6(a)). The illustrative scenario for this archetype is SusDev_SSP2-PkBudg900 from REMIND-MAgPIE 2.1-4.2 [57]. This scenario has a pronounced and consistent increase in energy demand from 2050 onwards (figure 6(b)), which is met primarily by a huge expansion in non-biomass renewables. There is a phasing down of unabated fossil fuel primary energy to around 25% of 2020 levels by 2060. CDR use expands rapidly after 2030, reaching ~6 Gt yr^{-1} by 2050 (figure 6(c)). CDR is predominantly comprised of BECCS. Low energy system inertia (renewables use) near-term demand constraints and fossil phase-down limit near-term robustness challenges.

3.4.2. Scenario B–moderated negative mitigation impacts, low societal burden. Tempered demand, CCS use

This scenario archetype is characterised by low relative challenges across the economic, resource, resilience and transition rate dimensions, but with moderate relative challenges on the robustness and fairness dimensions. This scenario archetype has high relative environmental challenges (figure 6(d)). The illustrative scenario for this archetype is SSP2_SPA1_19I_LIRE_LB from IMAGE 3.2 [58]. This scenario sees primary energy demand falling to 2045, before increasing gradually to slightly higher than present levels by 2100 (figure 6(e)). This is alongside a significant phasing down of unabated fossil fuels to less than a quarter of present use by the mid-century. Fossil with CCS and renewables are used to meet the supply gap. This lower demand trajectory, necessitating a more moderated energy system transformation, explains lower scores across

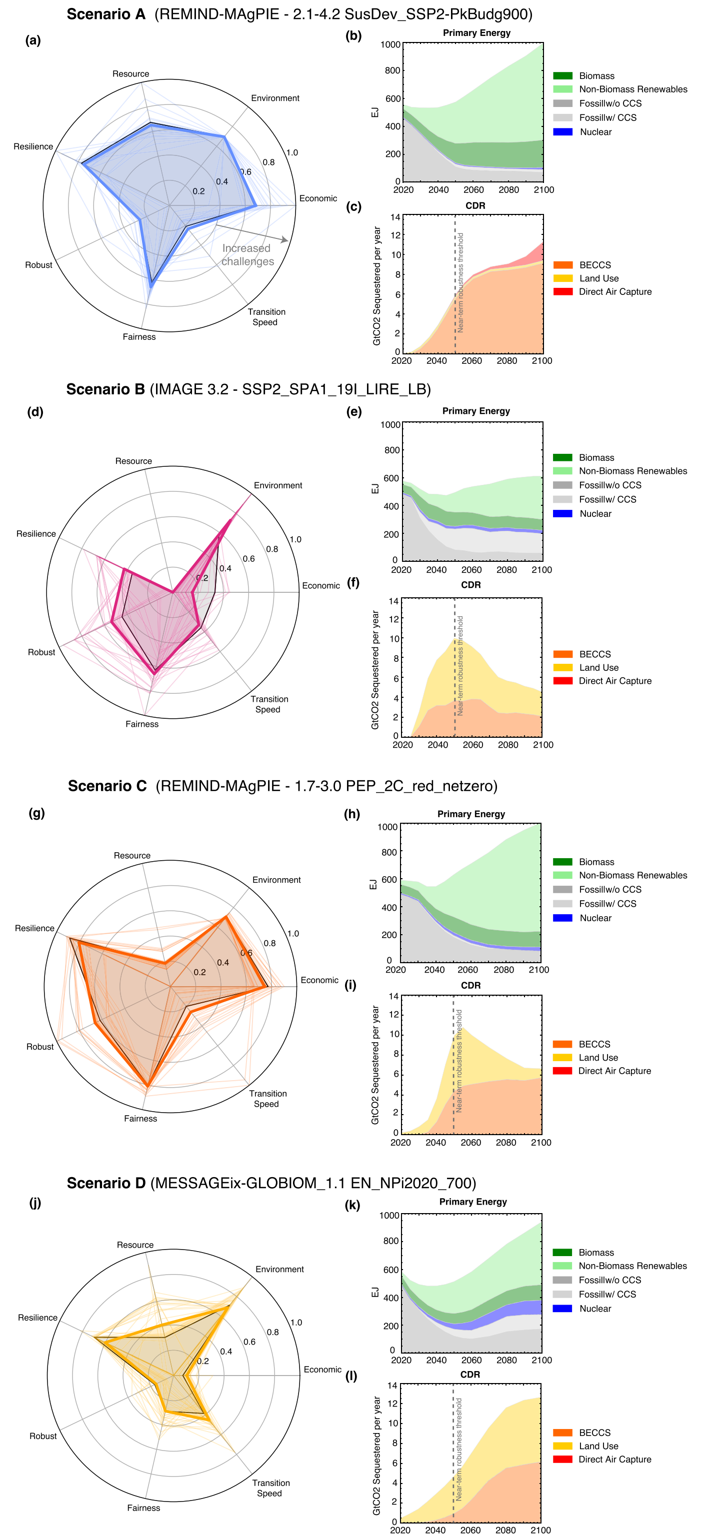

Figure 6. Showing scenario archetypes (thick coloured), corresponding illustrative scenarios (thin black and grey), and remaining scenarios from each cluster (coloured faint lines). The archetypes are the centroid values for each of our clusters. Our illustrative scenarios are scenarios from each cluster that are closest to the centroids. Sub-figures (a), (d), (g) and (j) detail scores for each of our four archetypes and illustrative scenarios. Sub-figures (b), (e), (h) and (k) provide the primary energy shares for each of our illustrative scenarios. Sub-figures (c), (f), (i) and (l) provide the CDR breakdown for each of our illustrative scenarios.

many of our dimensions. There is an immediate, rapid scaling up of land use CDR and BECCS, reaching around 10 Gt per year by 2050 (figure 6(f)). It is the scaling of these CDR types that contributes to this scenario having high relative environmental challenges.

3.4.3. Scenario C–high negative mitigation impacts, regionally unfair. High demand, huge renewables expansion

This scenario archetype is characterised by high relative challenges across the resilience, fairness, economic and environment dimensions, with moderate robustness challenges and low transition rate and resource challenges (figure 6(g)). The illustrative scenario for this archetype is PEP_2C_red_netzero from REMIND-MAgPIE 1.7–3.0 [59]. The scenario has primary energy demand declining slightly up to 2040, then climbing rapidly over the remainder of the century (figure 6(h)). Unabated fossil fuels are phased down, to around 20% of present levels by 2100. Renewable energy expansion meets additional demand and fills the gap from fossil phase down. Low resource challenges are because this occurs later in the century (resources dimension considers up to 2050). However, high demand and significant energy system transformation lead to high challenges across a range of other dimensions, the use of CDR in this scenario is moderate in the near term, climbing rapidly after 2035 to reach $\sim$9 Gt yr^{-1} by the mid-century (figure 6(i)).

3.4.4. Scenario D–lower negative mitigation impacts. Delayed demand growth, diverse energy supply. Low CDR

This scenario archetype has low, or very low relative challenges across all dimensions except resilience, and environmental sustainability, for which moderate challenges are seen (figure 6(j)). The illustrative scenario for this archetype is EN_NPi2020_700 from MESSAGEix-GLOBIOM_1.1 [19]. This scenario sees demand initially falling to 2040 and then climbing consistently, reaching roughly 150% of present levels by 2100 (figure 6(k)). There is a phasing down of unabated fossil fuel use to mid-century, after which it climbs modestly again. In the near term, there is lower overall use of CDR, reaching 4.5 Gt yr^{-1} by 2050 (figure 6(l)). Limited near-term CDR use, and relative diversity in low-carbon primary energy sources, contribute to low relative robustness challenges. However, longer-term land-based and BECCS CDR use (12 Gt yr^{-1} by 2100) contributes towards environmental challenges.

3.5. Considering regional differences in evaluation scores

The exploration and evaluation framework can systematically compare scenarios at the regional level. Where necessary, indicators have been adapted to apply on a regional basis, but where possible they remain the same (see Methods). For the resources dimension, we lack the required data to be able to quantify the indicators at a regional level. The fairness dimension relies on between-region data and is a global measure, it is also not included in the regional analysis.

Regional outputs have a variety of uses, depending in part on how the results are normalised. There may be instances where a user is interested only in the results for a specific region, for example, selecting contrasting scenarios but based on regional scores. In this instance, clustering analysis and scenario selection can be performed on the regional results, in a similar way as demonstrated in the section above.

Within-region normalisation prevents easy comparison of scores between regions.

Alternatively, using across-region normalisation of scores, where the maximum and minimum from all ten regions are used to normalise, may give the user more insights into relative challenges on different dimensions between regions (figures 7(a)–(c)). For example, broadly, regions in the Global South score higher on societal resilience challenges (figure 7(b)) than regions in the Global North. Or, on near-term scenario robustness (figure 7(c)), North America has particularly high challenges when compared to the other regions.

Finally, a user may wish to have an understanding of absolute differences between specific regions for certain indicator scores. We provide an example of this in section S5 in the supplementary information.

4. Discussion and conclusions

Here we have presented a new approach to quantitatively explore, evaluate and select scenarios from larger ensembles. Our framework extends existing work by combining considerations of desirability with considerations of feasibility. This enhances the overall utility for users of scenario ensembles and enables a more comprehensive evaluation of such ensembles for decision-making and planning. Alignment or tensions between individual assessment dimensions can be identified and inform synergies and trade-offs in the pursuit of societal objectives.

Our framework and its application to 1.5 °C scenarios also demonstrate the potential of using quantitative methods to identify scenario archetypes and illustrative scenarios. This complements more qualitative methods that have been used in the past and provides a structured way to choose scenarios that are diverse across a chosen set of dimensions. Additionally, this quantitative framework allows users to identify gaps within the scenario space. Importantly, through its structured approach, the framework allows users to identify gaps within the scenario space that can be subject to further enquiry.

The framework scores, combined with cluster analysis, produced four contrasting scenarios, each with distinct dimension scores and demonstrating different primary energy and CDR trajectories. Our scoring across the seven dimensions also highlighted gaps in the scenario space that modellers could explore, such as scenarios that perform well in both

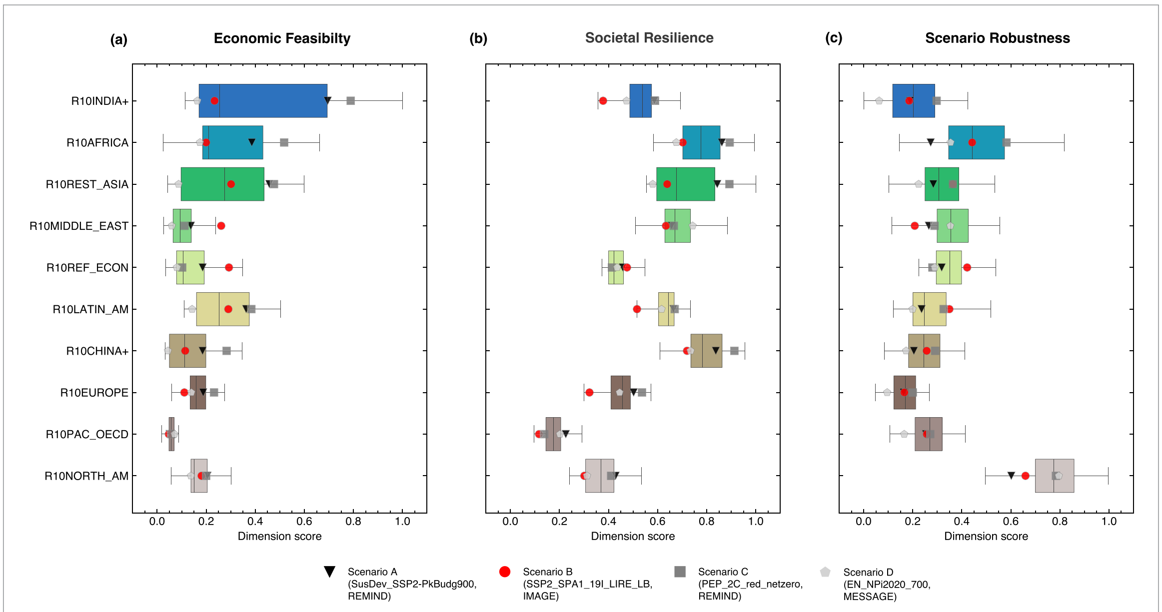

Figure 7. Distributions of the scenario scores for each of the R10 regions, for the economic feasibility (a), societal resilience (b) and near-term scenario robustness (c) dimensions. The data shown is across region normalisation and may indicate relative challenges between different regions on a specific dimension. The regions are arranged starting with the lowest mean GDP/Capita 2020–2100 at the top, to the highest at the bottom. Median values and interquartile range are given by the central line on the boxplot and by the box exterior, respectively. The box whiskers represent the smallest and largest data points that fall within 1.5 times the IQR from Q1 and Q3. Outliers are not shown. The values for the illustrative scenarios are added on top of each sub-figure with icons (see key).

interregional fairness and societal resilience. Future users of this framework can experiment with different scenario sets or focus on specific dimensions of interest. Finally, we demonstrated how the framework could be used to compare regional scenario data, with differences in relative challenges for certain regions, and regional groupings. For example, we see a trend of higher societal resilience challenges for the Global South, than the Global North.

The framework is developed as a tool available to the community, and one of the key strengths is therefore its flexibility. It can be tailored to accommodate different user requirements (such as incorporating additional dimensions or indicators), can be applied both regionally or globally (enhancing its relevance and utility for a wide range of users), and its codebase can be further expanded and re-used.

The present framework comes with identified limitations, for example, in its representation of each assessment dimension. This is in part due to variables reported by scenarios and models. A trivial example here is the limited regional granularity of data available in large ensemble databases. Improved coverage of certain variables would allow them to be utilised in the exploration and evaluation framework, giving a fuller representation of a given dimension. Other dimensions, however, would require the development of additional assessment tools or analysis.

In conclusion, our quantitative approach to scenario exploration, evaluation and selection offers a flexible framework that helps improve the utility of scenario databases. By integrating desirability criteria and emphasising regional evaluations, we provide a tool for scenario analysis that can be adapted to meet the diverse needs of users and has a variety of use cases.

Data availability statement

All data used to generate the results presented in this paper are publicly available. The full source code and associated input/output data used for the paper are archived on Zenodo [60]. The most up-to-date version of the code is available at on Github.

Acknowledgment

This work was partially funded by the Hitachi-Imperial Centre for Decarbonisation and Natural Climate Solutions. S M acknowledges financial support from the Horizon Europe European Commission 'DIAMOND' (Grant No. https://doi.org/101081179).

ORCID iDs

Hamish Beath https://orcid.org/0000-0002-5124-9143
Shivika Mittal https://orcid.org/0000-0003-4718-0064
Robin Lamboll https://orcid.org/0000-0002-8410-037X
Joeri Rogelj https://orcid.org/0000-0003-2056-9061

References

[1] Rogelj J *et al* 2018 Scenarios towards limiting global mean temperature increase below 1.5 °C *Nat. Clim. Change* **8** 325–32

[2] Cointe B 2024 The AR6 scenario explorer and the history of IPCC scenarios databases: evolutions and challenges for transparency, pluralism and policy-relevance *npj Clim. Action* **3** 1–10

[3] IPCC 2022 Summary for policymakers *Global Warming of 15 °C* (Cambridge University Press) pp 1–24

[4] IPCC 2023 *Climate Change 2022—Mitigation of Climate Change* (Cambridge University Press)

[5] Guivarch C *et al* 2022 Using large ensembles of climate change mitigation scenarios for robust insights *Nat. Clim. Change* **12** 428–35

[6] Pielke R and Ritchie J 2021 Distorting the view of our climate future: the misuse and abuse of climate pathways and scenarios *Energy Res. Soc. Sci.* **72** 101890

[7] Werners S E *et al* 2021 Advancing climate resilient development pathways since the IPCC's fifth assessment report *Environ. Sci. Policy* **126** 168–76

[8] Pedersen J T S, van Vuuren D, Gupta J, Santos F D, Edmonds J and Swart R 2022 IPCC emission scenarios: how did critiques affect their quality and relevance 1990–2022? *Glob. Environ. Change* **75** 102538

[9] Byers E *et al* 2022 AR6 scenarios database *Zenodo* (https://doi.org/10.5281/zenodo.5886912)

[10] Huppmann D *et al* 2019 IAMC 1.5°C scenario explorer and data hosted by IIASA *Zenodo* (https://doi.org/10.5281/zenodo.3363345)

[11] Huppmann D, Rogelj J, Kriegler E, Krey V and Riahi K 2018 A new scenario resource for integrated 1.5 °C research *Nat. Clim. Change* **8** 1027–30

[12] Brutschin E, Pianta S, Tavoni M, Riahi K, Bosetti V, Marangoni G and van Ruijven B J 2021 A multidimensional feasibility evaluation of low-carbon scenarios *Environ. Res. Lett.* **16** 064069

[13] Tank L, Voget-Kleschin L, Garschagen M, Boettcher M, Mengis N, Holland-Cunz A, Rehder G and Baatz C 2025 Distinguish between feasibility and desirability when assessing climate response options *npj Clim. Action* **4** 1–8

[14] Wang P, Chen W-Q, Cui X, Li J, Li W, Wang C, Cai W and Geng X 2022 Critical mineral constraints in global renewable scenarios under 1.5 °C target *Environ. Res. Lett.* **17** 125004

[15] Wang P *et al* 2024 Regional rare-earth element supply and demand balanced with circular economy strategies *Nat. Geosci.* **17** 94–102

[16] Schlichenmaier S and Naegler T 2022 May material bottlenecks hamper the global energy transition towards the 1.5 °C target? *Energy Rep.* **8** 14875–87

[17] Rafaj P *et al* 2021 Air quality and health implications of 1.5 °C–2 °C climate pathways under considerations of ageing population: a multi-model scenario analysis *Environ. Res. Lett.* **16** 045005

[18] Hasegawa T *et al* 2021 Land-based implications of early climate actions without global net-negative emissions *Nat. Sustain.* **4** 1052–9

[19] Riahi K *et al* 2021 Cost and attainability of meeting stringent climate targets without overshoot *Nat. Clim. Change* **11** 1063–9

[20] Napp T, Bernie D, Thomas R, Lowe J, Hawkes A and Gambhir A 2017 Exploring the feasibility of low-carbon scenarios using historical energy transitions analysis *Energies* **10** 116

[21] Gambhir A *et al* 2017 Assessing the feasibility of global long-term mitigation scenarios *Energies* **10** 89

[22] Gidden M J, Brutschin E, Ganti G, Unlu G, Zakeri B, Fricko O, Mitterrutzner B, Lovat F and Riahi K 2023 Fairness and feasibility in deep mitigation pathways with novel carbon dioxide removal considering institutional capacity to mitigate *Environ. Res. Lett.* **18** 074006

[23] van de Ven D-J *et al* 2023 A multimodel analysis of post-glasgow climate targets and feasibility challenges *Nat. Clim. Change* **13** 570–8

[24] Bertram C *et al* 2024 Feasibility of peak temperature targets in light of institutional constraints *Nat. Clim. Change* **14** 954–60

[25] Warszawski L *et al* 2021 All options, not silver bullets, needed to limit global warming to 1.5 °C: a scenario appraisal *Environ. Res. Lett.* **16** 064037

[26] Sælen H, Tørstad V, Holz C and Nielsen T D 2019 Fairness conceptions and self-determined mitigation ambition under the Paris Agreement: is there a relationship? *Environ. Sci. Policy* **101** 245–54

[27] Davide M, Parrado R and Campagnolo L Fairness in NDCs: comparing mitigation efforts from an equity perspective

[28] Deprez A, Leadley P, Dooley K, Williamson P, Cramer W, Gattuso J-P, Rankovic A, Carlson E L and Creutzig F 2024 Sustainability limits needed for CO_2 removal *Science* **383** 484–6

[29] Fujimori S, Oshiro K, Hasegawa T, Takakura J and Ueda K 2023 Climate change mitigation costs reduction caused by socioeconomic-technological transitions *npj Clim. Action* **2** 9

[30] Köberle A C, Vandyck T, Guivarch C, Macaluso N, Bosetti V, Gambhir A, Tavoni M and Rogelj J 2021 The cost of mitigation revisited *Nat. Clim. Change* **11** 1035–45

[31] Lèbre É, Stringer M, Svobodova K, Owen J R, Kemp D, Côte C, Arratia-Solar A and Valenta R K 2020 The social and environmental complexities of extracting energy transition metals *Nat. Commun.* **11** 1–8

[32] Rogelj J *et al* 2022 Mitigation pathways compatible with 1.5°C in the context of sustainable development *Global Warming of 15 °C* (Cambridge University Press) pp 93–174

[33] Hegre H, Buhaug H, Calvin K V, Nordkvelle J, Waldhoff S T and Gilmore E 2016 Forecasting civil conflict along the shared socioeconomic pathways *Environ. Res. Lett.* **11** 054002

[34] Hasegawa T *et al* 2018 Risk of increased food insecurity under stringent global climate change mitigation policy *Nat. Clim. Change* **8** 699–703

[35] Andrijevic M, Crespo Cuaresma J, Muttarak R and Schleussner C-F 2019 Governance in socioeconomic pathways and its role for future adaptive capacity *Nat. Sustain.* **3** 35–41

[36] Cardin M A, Mijic A and Whyte J 2021 Flexibility and real options in engineering systems design *Handbook of Engineering Systems Design* ed A Maier, J Oehmen and P E Vermaas (Springer) pp 1–29

[37] Grubler A *et al* 2018 A low energy demand scenario for meeting the 1.5 °C target and sustainable development goals without negative emission technologies *Nat. Energy* **3** 515–27

[38] Rogelj J, Huppmann D, Krey V, Riahi K, Clarke L, Gidden M, Nicholls Z and Meinshausen M 2019 A new scenario logic for the Paris Agreement long-term temperature goal *Nature* **573** 357–63

[39] Gidden M J and Huppmann D 2019 pyam: a Python package for the analysis and visualization of models of the interaction of climate, human, and environmental systems *J. Open Source Softw.* **4** 1095

[40] Creutzig F *et al* 2015 Bioenergy and climate change mitigation: an assessment *GCB Bioenergy* **7** 916–44

[41] IMF. World Economic Outlook database 2023 World Economic and Financial Surveys - World Economic Outlook Database (available at: www.imf.org/en/Publications/WEO/weo-database/2023/October/download-entire-database) (Accessed 4 March 2024)

[42] IEA 2023 World energy investment 2023 *Analysis* (IEA) www.iea.org/reports/world-energy-investment-2023#downloads (Accessed 4 March 2024)

[43] Lubis C, Doherty D and Young W 2022 Investment requirements of a low-carbon world: energy supply

investment ratios (available at: https://assets.bbhub.io/professional/sites/24/BNEF-EIRP-Climate-Scenarios-and-Energy-Investment-Ratios.pdf) (Accessed 4 June 2024)
[44] USGS 2024 Mineral commodity summaries 2024 (available at: https://pubs.usgs.gov/publication/mcs2024) (Accessed 5 March 2024)
[45] Thier P and Pot D'or C 2020 Contribution of diversity to the resilience of energy systems-a literature review *Proc. 30th European Safety and Reliability Conf. and the 15th Probabilistic Safety Assessment and Management Conf.*
[46] Kruyt B, van Vuuren D P, de Vries H J M and Groenenberg H 2009 Indicators for energy security *Energy Policy* **37** 2166–81
[47] Sovacool B K, Von H D F, Suzuki T, Williams J H and Savage T 2010 Measuring security of energy supply with two diversity indexes pp 313–29 (available at: www.taylorfrancis.com/chapters/edit/10.4324/9780203834602-17/measuring-security-energy-supply-two-diversity-indexes-john-kessels)
[48] Benveniste H, Cuaresma J C, Gidden M and Muttarak R 2021 Tracing international migration in projections of income and inequality across the shared socioeconomic pathways *Clim. Change* **166** 1–22
[49] Dellink R, Chateau J, Lanzi E and Magné B 2017 Long-term economic growth projections in the shared socioeconomic pathways *Glob. Environ. Change* **42** 200–14
[50] Lutz W, Stilianakis N, Stonawski M, Goujon A and Samir K C Demographic and human capital scenarios for the 21st century (Publications Office of the European Union) (available at: https://op.europa.eu/en/publication-detail/-/publication/e1853ba8-4444-11e8-a9f4-01aa75ed71a1/language-en) (Accessed 5 March 2024)
[51] Lamboll R D, Nicholls Z R J, Smith C J, Kikstra J S, Byers E and Rogelj J 2023 Assessing the size and uncertainty of remaining carbon budgets *Nat. Clim. Change* **13** 1360–7
[52] Friedlingstein P *et al* 2023 Global carbon budget 2023 *Earth Syst. Sci. Data* **15** 5301–69
[53] Global Carbon Project 2023 Latest data | global carbon budget data (available at: https://globalcarbonbudgetdata.org/latest-data.html) (Accessed 5 March 2024)
[54] Aleksandrowicz L, Green R, Joy E J M, Smith P, Haines A and Wiley A S 2016 The impacts of dietary change on greenhouse gas emissions, land use, water use, and health: a systematic review *PLoS One* **11** e0165797
[55] Harmsen M *et al* 2021 Integrated assessment model diagnostics: key indicators and model evolution *Environ. Res. Lett.* **16** 054046
[56] Kriegler E *et al* 2015 Diagnostic indicators for integrated assessment models of climate policy *Technol. Forecast. Soc. Change* **90** 45–61
[57] Soergel B *et al* 2021 A sustainable development pathway for climate action within the UN 2030 agenda *Nat. Clim. Change* **11** 656–64
[58] Müller-Casseres E, Edelenbosch O Y, Szklo A, Schaeffer R and van Vuuren D P 2021 Global futures of trade impacting the challenge to decarbonize the international shipping sector *Energy* **237** 121547
[59] Kriegler E *et al* 2018 Short term policies to keep the door open for Paris climate goals *Environ. Res. Lett.* **13** 074022
[60] Beath H 2025 hamishbeath/IAMEE: v0.1 - ERL Paper Zenodo (https://doi.org/10.5281/zenodo.15569460) (Accessed 1 June 2025)