IAM COMPACT

Expanding Integrated Assessment Modelling: Comprehensive and Comprehensible Science for Sustainable, Co-Created Climate Action

D3.7 – Open science protocols - Update

WP3 – Exchanging – Open & FAIR science, mutual learning

27/09/2024

Funded by the European Union

www.iam-compact.eu

Disclaimer

The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the European Union. Neither CINEA nor the European Commission is responsible for any use that may be made of the information contained therein.

Copyright Message

This report, if not confidential, is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0); a copy is available here: https://creativecommons.org/licenses/by/4.0/. You are free to share (copy and redistribute the material in any medium or format) and adapt (remix, transform, and build upon the material for any purpose, even commercially) under the following terms: (i) attribution (you must give appropriate credit, provide a link to the license, and indicate if changes were made; you may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use); (ii) no additional restrictions (you may not apply legal terms or technological measures that legally restrict others from doing anything the license permits).

Grant Agreement Number	101056306	**Acronym**	IAM COMPACT
Full Title	Expanding Integrated Assessment Modelling: Comprehensive and Comprehensible Science for Sustainable, Co-Created Climate Action		
Topic	HORIZON-CL5-2021-D1-01-04		
Funding scheme	HORIZON EUROPE, RIA – Research and Innovation Action		
Start Date	September 2022	**Duration**	36 Months
Project URL	https://www.iam-compact.eu		
EU Project Advisor	Andreas Palialexis		
Project Coordinator	National Technical University of Athens – NTUA		
Deliverable	D3.7 – Open science protocols		
Work Package	WP3 – Exchanging – Open & FAIR science, mutual learning		
Date of Delivery	**Contractual** 30/09/2024	**Actual**	27/09/2024
Nature	Report	**Dissemination Level**	Public
Lead Beneficiary	ASOCIACION BC3 BASQUE CENTRE FOR CLIMATE CHANGE - KLIMA ALDAKETA IKERGAI (BC3)		
Responsible Author	Clàudia Rodés-Bachs	**Email**	claudia.rodes@bc3research.org
	BC3	**Phone**	+34 94 401 46 90
Contributors	Russell Horowitz, Dirk-Jan Van de Ven (BC3)		
Reviewer(s)	Natasha Frilingou, Alexandros Nikas (NTUA); Francesco Gardumi, Camilla Lo Giudice (KTH)		
Keywords	open science; open access; FAIR principles; TRUST principles; harmonisation; output management		

EC Summary Requirements

1. Changes with respect to the DoA

No changes with respect to the work described in the DoA.

2. Dissemination and uptake

This report will be used internally by all consortium partners to guide all open science practices of the project, especially in its 2nd Policy Response Mechanism cycle. It can be also used by users outside the consortium as an example for informing their own strategies (especially for other Horizon Europe projects).

3. Short summary of results (<250 words)

This deliverable contains updates to and expansion of deliverable D3.6 – Open Science Protocols. It includes (a) a brief summary of the state-of-art of open science practices; (b) a discussion of the results of a qualitative and quantitative survey that was conducted with all modelling teams from the 1st modelling cycle, and which gathers insights, feedback, and suggestions for the operationalisation of open science principles in the project; (c) the updated Open Science Protocol for IAM COMPACT, which aims to be used in the 2nd Policy Response Mechanism cycle; and (d) suggestions to boost the dissemination and comprehension of the protocol. The updated protocol aims at the integration of the FAIR and TRUST principles in all project activities and by all consortium partners to promote the smooth interconnection of the models and multi-model exercises, and transparently manage the produced outcomes. To that aim, it details clear steps and suggests useful open-source platforms and tools for the different tasks.

4. Evidence of accomplishment

This report.

Preface

IAM COMPACT supports the assessment of global climate goals, progress, and feasibility space, and the design of the next round of Nationally Determined Contributions (NDCs) and policy planning beyond 2030 for major emitters and non-high-income countries. It uses a diverse ensemble of models, tools, and insights from social and political sciences and operations research, integrating bodies of knowledge to co-create the research process and enhance transparency, robustness, and policy relevance. It explores the role of structural changes in major emitting sectors and of political, behaviour, and social aspects in mitigation, quantifies factors promoting or hindering climate neutrality, and accounts for extreme scenarios, to deliver a range of global and national pathways that are environmentally effective, viable, feasible, and desirable. In doing so, it fully accounts for COVID-19 impacts and recovery strategies and aligns climate action with broader sustainability goals, while developing technical capacity and promoting ownership in non-high-income countries.

NTUA – National Technical University of Athens	EL	
Aalto – Aalto Korkeakoulusaatio SR	FI	
AAU – Aalborg Universitet	DK	
BC3 – Asociacion BC3 Basque Centre for Climate Change – Klima Aldaketa Ikergai	ES	
Bruegel – Bruegel AISBL	BE	
CARTIF – Fundacion CARTIF	ES	
CICERO – Cicero Senter for Klimaforskning Stiftelse	NO	
E3M – E3-Modelling AE	EL	
KTH – Kungliga Tekniska Hoegskolan	SE	
POLIMI – Politecnico di Milano	IT	
UPRC – University of Piraeus Research Center	EL	
UVa – Universidad De Valladolid	ES	
WI – Wuppertal Institut fur Klima, Umwelt, Energie GGMBH	DE	
IIMA – Indian Institute of Management	IN	
THU – Tsinghua University	CN	
USMF – University System of Maryland	US	
AAiT – Addis Ababa University	ET	
KEI – International Civic Organisation Kyiv Economics Institute	UA	
RUSL – Raja Rata University of Sri Lanka	LK	
TUM – Technical University of Mombasa	KE	
UNIGE – Université de Geneve	CH	
Imperial – Imperial College of Science, Technology and Medicine	UK	

Executive Summary

Open science is a concept that aims to make scientific research and data available to both the scientific community and non-expert audiences. It has multiple benefits, such as increasing stakeholders' engagement, enhancing scientific collaboration, and ensuring the reproducibility of the work along the way. As open science practices gain traction across various fields, this deliverable outlines the development and application of an open science protocol to be followed by the different modelling teams within the IAM COMPACT consortium.

IAM COMPACT comprises a wide range of models and modelling teams, making transparency and thorough documentation crucial to ensuring the accessibility, clarity, and reproducibility of research outcomes to various audiences. This deliverable includes the key open science principles, including the widely adopted FAIR principles—Findable, Accessible, Interoperable, and Reusable. While these principles have been successful in promoting data exchange and usage in multiple projects, there are evident limitations to their effectiveness. To address these gaps, the TRUST principles—Transparency, Responsibility, User focus, Sustainability, and Technology—were developed to evaluate the trustworthiness of digital repositories. Together, these frameworks ensure that research outputs can be the basis of future scientific work, enabling broad access and reuse.

The deliverable also highlights the infrastructure available to support open science, such as platforms like GitHub, Zenodo, and OpenAIRE, which provide robust tools for code and data documentation, archiving, and ensuring long-term accessibility. A glossary of commonly used open science terms is provided to clarify terminology that may be confusing.

We first review the implementation of the Open Science Protocol proposed in D3.6 during IAM COMPACT's 1st Policy Response Mechanism (PRM), and associated modelling, cycle. It identifies practices embraced by modelling teams as well as challenges encountered when applying certain steps of the protocol. Based on the findings of this review, the deliverable offers an updated Open Science Protocol that better addresses the needs of the teams, accompanied by a checklist to facilitate compliance during the 2nd PRM/modelling cycle. Additionally, it proposes new initiatives, including a workshop and the publication of a manuscript, to further promote and disseminate best open science practices within the wider integrated assessment model (IAM) community.

Funded by
the European Union

Contents

1 Introduction

IAM COMPACT is an international transdisciplinary project that brings together a wide range of models and modelling teams to analyse global climate goals and support the development of post-2030 Nationally Determined Contributions (NDCs) and long-term strategies. With diverse multi-model analyses at its core, the project emphasises the importance of standardising open science practices among the modelling teams to ensure that model outputs are reliable, accessible, transparent, and reusable.

This deliverable builds upon the previous work outlined in deliverable D3.6 – Open Science Protocols (Rodés-Bachs et al., 2023), extending its scope and updating the open science guidelines. As in D3.6, we continue to advocate for open science practices that promote the open availability of scientific research and data. Recognising that every modelling team has unique circumstances, these recommended practices are intended to be adaptable and customisable. While they serve as valuable defaults, they should be refined and tailored to suit the specific needs of each case study and model.

The primary objective of this deliverable is to reinforce the importance of open science and to facilitate its integration into the upcoming 2nd modelling cycle, scheduled for fall of 2024. In addition to expanding on the concepts introduced in D3.6, this report reviews the practices implemented during the 1st modelling cycle. Based on a survey distributed to all modelling teams, we identify the barriers encountered by modellers that hindered the full adoption of the original Open Science Protocol. These insights form the basis for a set of improvements to the above-mentioned protocol, alongside new measures aimed at promoting and easing its adoption. Importantly, all teams reported engaging with open science practices to some degree, signalling a strong foundation and commitment that can be further strengthened.

The deliverable is structured as follows: First, it provides an overview of the current state of open science in research, covering key principles, benefits, and commonly used platforms and tools. This section concludes with a compendium of definitions to clarify recurring concepts throughout the document. Next, the deliverable analyses the open science practices employed by the IAM COMPACT consortium during the 1st modelling cycle, highlighting good practices and identifying barriers based on survey results. Finally, it proposes improvements to overcome these challenges, including an updated version of the IAM COMPACT Open Science Protocol, a checklist to guide teams in the next cycle, and novel initiatives such as hosting a webinar and publishing a manuscript to further disseminate, document, and promote the Open Science Protocol.

For detailed guidance on specific protocol steps, this deliverable references additional reports, such as D4.6 (due in May 2025), D4.5 (Mittal, Giarola, et al., 2023), D3.2 (Mittal, Frilingou, et al., 2023) and D4.4 (Korsbakken et al., 2024). These deliverables offer in-deep insights into crucial modelling points, such as harmonising model inputs and outputs, fostering consistency across the consortium.

By providing this support, we aim to help modelling teams overcome the challenges encountered during the 1st modelling cycle, enabling broader and more effective adoption of open science practices. Ultimately, this will lead to more transparent, accessible, and reproducible research outcomes. Our focus remains on ensuring that modelling teams appreciate the significance of open science and apply it effectively, thereby contributing to the broader goals of the IAM COMPACT project.

Funded by
the European Union

2 Open science practices in scientific research

Open science is a concept that aims to **make scientific research and data available** to scientists, stakeholders, and the general public. The main principle of open science is based on the idea that scientific knowledge should be freely available to everyone and should be easily accessible for reuse, replication, and further development (European Commission, 2019). It rests on a set of core values, including transparency, reproducibility, accessibility, and interdisciplinarity, to ensure that scientific research is conducted in an open and transparent manner, and that the resulting data, methodology, and **findings are made available to the broader scientific community and non-expert audiences** (Munafò et al., 2017). By making scientific research and data available to everyone, open science can help to promote scientific progress and drive innovation (Fecher et al., 2015).

The **benefits of open science are wide-ranging**. It enhances research efficiency, fosters collaboration, and engages the public by making science accessible to non-experts. This was exemplified during the COVID-19 pandemic, where open science accelerated scientific collaboration (Besançon et al., 2021; Fraser et al., 2020). Since then, multiple **open science initiatives** have aroused and best practices and data repositories have been published by specific scientific disciplines (e.g., Nature Protocols[1], BioModels Database[2], and the Environmental Data Initiative[3]). The **Integrated Assessment Modeling Consortium**[4] (IAMC) community, which is closely related to the IAM COMPACT project, also employs open science practices to enhance the accessibility of scenarios from integrated assessment models (IAMC, 2022; Nikas et al., 2021, p. 202), to improve the comprehension and interconnection of models (IAMC, 2022), and to facilitate data analysis and comprehension through interactive visualisation of model outputs, such as the 1.5ºC scenario explorer (Huppmann et al., 2018), the *pyam* package (Gidden & Huppmann, 2019), the Global Hotspots Explorer[5], and the Sixth Assessment Report (AR6) scenario explorer (Byers et al., 2022).

As open data gains prominence, the **FAIR Principles** have emerged as a measurable guideline for assessing the findability, accessibility, interoperability, and reusability of scientific data (Wilkinson et al., 2016). These principles, adopted by institutions like the European Commission, promote data exchange and usage, enabling researchers to build upon existing knowledge. However, there are still some challenges that would need to be addressed. The FAIR Principles face limitations, including a focus on individual practices and short-term data management (Boeckhout et al., 2018; Peng et al., 2022). To address these potential barriers, the **TRUST principles**, based on the pillars of Transparency, Responsibility, User focus, Sustainability, and Technology, were introduced in 2020, ensuring digital repository trustworthiness and long-term data preservation (Lin et al., 2020).

Regarding the **software infrastructure** to develop these practices, researchers are spoilt for choice with a multitude of platforms and tools supporting open science. GitHub[6], for instance, has become a popular venue for version control and collaboration (Finley, 2011). Zenodo[7], hosted by the European Council for Nuclear Research (CERN), and OpenAIRE[8], a European platform for sharing research, provide comprehensive data archiving and curation services (Kaiser & McNeill, 2019; Manghi et al., 2017; Sicilia et al., 2017). While technical barriers exist, numerous accessible tools enhance code reproducibility and transparency, including Docker[9] and JupyterHub[10]. The Commission has also developed platforms like the European Open Science Cloud[11] (EOSC), to promote open access to publicly funded research data and some countries and agencies have introduced internal protocols to standardise open science practices and help others to implement them more effectively (Peroni, 2021).

[1] https://www.nature.com/nprot/
[2] https://www.ebi.ac.uk/biomodels/
[3] https://portal.edirepository.org/nis/home.jsp
[4] https://www.iamconsortium.org/
[5] https://www.hotspots-explorer.org/
[6] https://github.org/
[7] https://zenodo.org/
[8] https://www.openaire.eu/
[9] https://www.docker.com/
[10] http://jupyterhub.org/
[11] https://eosc-portal.eu/

2.1 Glossary

To clarify **key concepts** that will be referenced throughout the deliverable, the following list provides descriptions for each term:

- Model: computer simulation based on a system of postulates, data, and inferences presented as mathematical description of an entity or state of affairs (Merriam-Webster, 2024).
- Integrated assessment model (IAM): model aim at providing policy-relevant insights into global environmental change and sustainable development by providing a quantitative description of key processes in the human and earth systems and their interactions (UNFCCC, 2024).
- Modelling team: a number of persons associated together (Merriam-Webster, 2024) to work with the development and/or usage of a model.
- Data: qualitative or quantitative information that is recorded and analysed. It encompasses facts, numbers, letters, symbols, images, sounds, or measurements. Data is collected from various sources including scientific research, government records, commercial transactions, social media, and more (Bivek Chaudhary, 2023).
- Metadata: description of the corresponding data in a detailed and unambiguous way, including process handling and reusage instructions (Ulrich et al., 2022).
- Open source: software complying with the ability to distribute it freely, the availability of the source code, the right to create derived works through modification, the no discrimination against persons or groups, the no discrimination against fields of endeavour, and distributed under a technology-neutral license that does not restrict other software, is not product specific, and allows distribution (Gacek & Arief, 2004; Open Source Initiative (OSI), 2024).
- Open data: data that is freely available, accessible, machine-readable, and available in open formats (Terzić & Majstorović, 2019).
- Open access (in publishing): free, immediate, online availability of research outputs such as journal articles or books, combined with the rights to use these outputs fully in the digital environment. Open access content is open to all, with no access fees (Springer Nature, 2024).
- Open science: the movement to make scientific research, data, code, educational resources, and publications freely available to everyone without barriers. It promotes transparency, collaboration, and reproducibility in research and encourages the use of open source, open data, and open access publishing. The goal of open science is to increase the accessibility and impact of scientific knowledge and to facilitate faster and more efficient progress (Chiware & Lockhart, 2024).

3 IAM COMPACT open science protocol analysis

After the 1st modelling cycle, we consider of utmost relevance to assess the open science practices employed during this period by the different modelling teams. Through Deliverable 3.6 - Open Science Protocols (Rodés-Bachs et al., 2023), we facilitated the application of several procedures that should be revised for the 2nd modelling cycle. To analyse how the application of these procedures was and which challenges the modelling teams faced, we collected information through a qualitative and quantitative survey regarding the 1st modelling cycle open science practices (available in Appendix A).

3.1 Survey structure and data collection

A survey was sent to all modelling teams in July 2024 containing both quantitative and qualitative questions. The form was divided into three main blocks:

1. **Classifier questions**: This section gathers the models used and the IAM COMPACT modelling studies participated in by respondents. It sets the stage for interpreting the subsequent results.
2. **Usage of D3.6 – Open Science Protocols**: This section aims to unravel the application of the open science protocol during the first modelling cycle. We inquire about the specific practices employed and the software utilised. The questionnaire is divided into three segments, corresponding to the key components of the original Open Science Protocol. The first part focuses on the model and data documentation process, including model documentation, input harmonisation, and models' connectivity. The second part explores the management of project outcomes, covering output standardisation, vetting process, output storage, and results communication. Lastly, we inquire about the tools used or developed to embrace open science principles.
3. **Feedback and suggestions**: This section is an opportunity for respondents to provide valuable insights on D3.6. We encourage them to share which elements were most helpful, which aspects require refinement or expansion in the upcoming D3.7, and any unnecessary sections. This feedback is key for refining our open science practices and enhancing the utility of our protocols.

3.2 Survey results

3.2.1 Section 1: About the respondents

We collected 9 answers (and modelling teams) from 8 institutions. All answers are reported in Appendix B. Here we summarise and discuss key aspects of the answers. The respondents represented 10 different models that cover all research studies from the 1st modelling cycle.

Table 1: Insight of the responding modelling team institutions and the employed models.

Modelling team institution	Models employed
AAU	EnergyPLAN
BC3	GCAM
CARTIF	WILLIAM
E3M	PROMETHEUS
NTUA	GCAM
NTUA	OSeMoSYS, OSEMBE
POLIMI	DyNERIO, MARIO
UNIPI	OSeMOSYS, TEEM
UVa	MEDAS, WILLIAM

3.2.2 Section 2: Usage of D3.6 - Open Science Protocols

All modelling teams were familiarised with open science practices and procedures. Indeed, except for PROMETHEUS and EnergyPLAN, all **models are open source** and available online (see Table 2). They are **all documented** through their respective documentation pages (detailed in Table 2) and the majority are also planning to use the I²AM PARIS platform once the new version is released (D3.9). This platform is currently under maintenance to integrate new features; therefore, not operable for uploading or modifying model documentation, model results, or other data yet. Similarly, all modelling teams that needed to **connect the input/output with other models or tools** for specific studies documented this procedure in the methods or Supplementary Information (SI) section of the related paper, with the exception of one team that did not detail their model connection procedures due to lack of time. Some examples of this step are Moreno et al., 2024 and Nikas et al., 2024. For the **inputs' harmonisation**, the modelling teams referred to the extensive harmonisation protocol for IAM COMPACT studies outlined in deliverable D4.3 – Broad Scenario Logic and recently updated in D4.4 – Broad Scenario Logic Update to complete the harmonisation task. Moreno et al., 2024 is an example of paper with detailed harmonisation procedure.

Table 2. Model repository and documentation links

Model name	Model repository link	Model documentation link
DyNERIO	https://github.com/SESAM-Polimi/DYNERIO-model	(Rinaldi et al., 2023)
EnergyPLAN	-	https://www.energyplan.eu/training/documentation/
GCAM	https://github.com/JGCRI/gcam-core	https://jgcri.github.io/gcam-doc/
MARIO	https://github.com/it-is-me-mario/MARIO	https://mario-suite.readthedocs.io/en/latest/
MEDEAS	https://geeds.es/medeas/	https://medeas.eu/model/medeas-model
OSEMBE	https://github.com/HauHe/OSeMBE	https://osembe.readthedocs.io/en/latest/
OSeMOSYS	https://github.com/OSeMOSYS/OSeMOSYS	http://www.osemosys.org/get-started.html
PROMETHEUS	-	https://e3modelling.com/modelling-tools/prometheus/
TEEM	https://github.com/TEESlab-UPRC/DREEM https://github.com/TEESlab-UPRC/ATOM	http://teeslab.unipi.gr/wp-content/uploads/2022/12/DREEM-Documentation_v1.0.pdf http://teeslab.unipi.gr/wp-content/uploads/2022/12/%CE%91%CE%A4%CE%9F%CE%9C-Documentation_v1.0.pdf
WILLIAM	https://github.com/LOCOMOTION-h2020	https://github.com/LOCOMOTION-h2020/WILIAM_model_VENSIM/wiki/The-WILIAM-model

With the exception of the recently released WILLIAM model and the EnergyPLAN, MARIO, and DyNERIO models, all models report their **outputs according to the IAMC time-series template**, as recommended by D3.6, D4.3 and D4.4 to facilitate model intercomparison. Furthermore, some papers have made their standardised outputs publicly available, such as Fragkos et al., 2024. In the case of the WILLIAM model, the modelling team is in the process of adopting the IAMC time-series template through a new package that will be released soon. On the other hand, the EnergyPLAN, MARIO, and DyNERIO models, which are energy models, usually serve as input sources for other system-wide IAMs and face greater difficulty in standardising their outputs. Their level of detail differs from that of IAMs, making direct comparisons more challenging. As a result, these models are less inclined to produce IAM-compatible output data.

On the validation of the results, three modelling teams currently use **vetting procedures**. For PROMETHEUS, the modelling team manually checks the robustness of the output values (e.g., energy demand and supply, power mix, and CO_2 emissions) for the years 2010 and 2015. In the case of GCAM, the output validation is automated using *gcamreport* (Rodés-Bachs et al., 2024), an R package designed ad hoc for the IAM COMPACT project to

standardise and validate GCAM outputs. This tool asserts that trade flows are robust and that values estimated by the model for the first modelling period (2020) stay within realistic ranges compared to observed data (e.g., emission inventories and energy demands). Other modelling teams reported to try to use the I²AM PARIS validation tool[12] but found some limitations that were reported to the platform developers. Thus, they expect to use it smoothly in the 2nd PRM/modelling cycle.

Regarding the **storage of the produced outputs**, half of the modelling teams store their processed study-specific data and/or analysis code under a DOI on Zenodo. Some examples are Rinaldi et al., 2024, Fragkos et al., 2024, Reisinger et al., 2024, and Sampedro et al., 2024. Other teams are planning to use it and one team indicated that they store data as appendices in scientific publications, but rarely under a DOI, as in Trutnevyte et al., 2024. However, due to multiple challenges (e.g., data coordination), only three teams had stored raw output files under a DOI at the time of the survey. To facilitate the publication of produced outputs during the first modelling cycle, NTUA coordinated data collection, and uploaded all modelling studies to Zenodo[13]. These datasets and/or analysis code are accompanied by metadata, which meticulously details the stored data and the data handling process. One good example is Rinaldi & Ghezzi, 2024. Besides, it is an increasingly common practice to upload standardised datasets, data compendiums, and other types of data not directly tied to a specific paper or study under a DOI. This ensures that other research teams can easily discover and utilise this information, fostering broader collaboration and data reuse. Examples of such data repositories include Mittal, S. & Van de Ven, D.J., 2024 and Van de Ven, D.J. et al., 2024.

The presentation of the results through **interactive user-friendly visualisation tools** is a usual practice by four models of the IAM COMPACT consortium (GCAM, WILLIAM, OSeMOSYS, and TEEM). They either have a model version equipped with a user-friendly interface that produces figures (e.g., WILLIAM web base[14]), an external tool that shows the results in tabular and graphical format (e.g., *gcamreport*[15]), or use EU-funded project platforms such as H2020 SENTINEL[16], H2020 ENCLUDE[17], and I²AM PARIS[18] platform.

All modelling teams used **open-source platforms**, with Zenodo (8 uses), GitHub (7 uses), Jupyter Notebooks (2 uses), Docker (1 use), and GitLab (1 use) being the primary platforms. These tools were mainly used for data storage, model development, release management, model documentation, and report sharing. Additionally, an open-source tool called *gcamreport* (Rodés-Bachs et al., 2024) was created to automatically produce standardise and validated GCAM outputs and to offer interactive and user-friendly visualisations; while another open-source package is being developed to standardise the WILLIAM outputs. Moreover, an **IAM COMPACT community** has been created in Zenodo (https://zenodo.org/communities/iam-compact/) to gather all the uploads related to the project.

3.2.3 Section 3: Feedback and suggestions

In this section some teams acknowledged the protocol's value in enhancing their understanding of open science practices. However, the teams did not emphasise any part of D3.6. Open Science Protocol or suggested improvements for the next round.

3.3 Discussion, protocol modifications and improvements

The collected survey data indicates that while all modelling teams employed various open science practices during

[12] https://validation.i2am-paris.eu/
[13] https://zenodo.org/communities/iam-compact/records?q=&f=resource_type%3Adataset&l=list&p=1&s=10&sort=newest
[14] https://www.locomotion-h2020.eu/locomotion-models/model-explorer/
[15] https://github.com/bc3LC/gcamreport
16 https://sentinel.energy/stories/
17 https://energycitizenship.eu/
18 https://www.i2am-paris.eu/feasibility/scientific_module

the 1st modelling cycle, the full implementation of the protocol across the consortium was hindered by several challenges and barriers, such as limited knowledge and time constraints. Therefore, we suggest some **protocol improvements** and a training workshop to overcome these handicaps and ensure that open science practices can be effectively applied in the 2nd modelling cycle.

The first step is to update the protocol (see Section 4) to include detailed examples of platforms, tools, and software, ensuring it serves as a dependable resource for the modelling teams. This update will also include the output homogenisation as an essential step for the open management of project outcomes. Based on insights from this modelling cycle, we recognise the necessity to homogenise the outputs as a fundamental task for effectively comparing results across different models and for enabling validation through an automated vetting procedure.

To facilitate the protocol application, we added **a checklist** (see Section 4.1) that can serve both as a guidance through the protocol steps and a reminder for best practices. It is structured following the protocol blocks and details all the advisable actions that modelling teams should follow to adhere to open science practices.

For the 2nd modelling cycle the I²AM PARIS platform will be upgraded (D3.9) as well as its validation tool. This will facilitate the publication of harmonisation heatmaps and the results visualisation through user-friendly interfaces in the I²AM PARIS platform, as well as the output vetting through the I²AM PARIS validation tool.

At the beginning of the 2nd modelling cycle, we will organise an **internal workshop** to disseminate the protocol and demonstrate valuable open-source platforms and tools relevant to various protocol steps. The workshop will also address any questions to ensure that teams can effectively apply the acquired knowledge. This approach aims to make the information accessible to all modelling teams in a clear and structured manner. By enhancing their understanding of open science practices, we expect that modelling teams are be able to implement these practices with greater efficiency and less effort.

Finally, there is an ongoing **publication** to document thoroughly this protocol and detail how to overcome barriers that modelling teams can encounter. It will serve as a documentation site not only for consortium members but also to the broad IAM community, encouraging the adoption of open science practices beyond the project's scope.

Funded by
the European Union

4 IAM COMPACT open science protocol update

After the analysis of the first version of the Open Science Protocol, some improvements were introduced to better detail the protocol steps and suggest platforms and tools that help in their implementation. Meanwhile, the updated version of the Open Science Protocol considers the output standardisation as a key task for the open management of model outcomes. Finally, this version incorporates a checklist to facilitate the protocol application in the 2nd modelling cycle.

The new version of the Open Science Protocol for IAM COMPACT, illustrated in Figure 1, guarantees that the intricate models can interact and yield transparent and validated results. By adhering to this protocol, the consortium can produce **reliable outcomes**, serving as a **dependable resource for informed decision-making**. The protocol encompasses two stages, both adhering to the **FAIR and TRUST principles**. The initial stage focuses on harmonising the models and data, ensuring consistency in modelling assumptions and input data. In contrast, the second stage addresses the open management of project outcomes, encompassing validation, quality control, and result visualisation and storage. Both stages are complemented with open-source platforms suggestions to facilitate the adoption of open science practices.

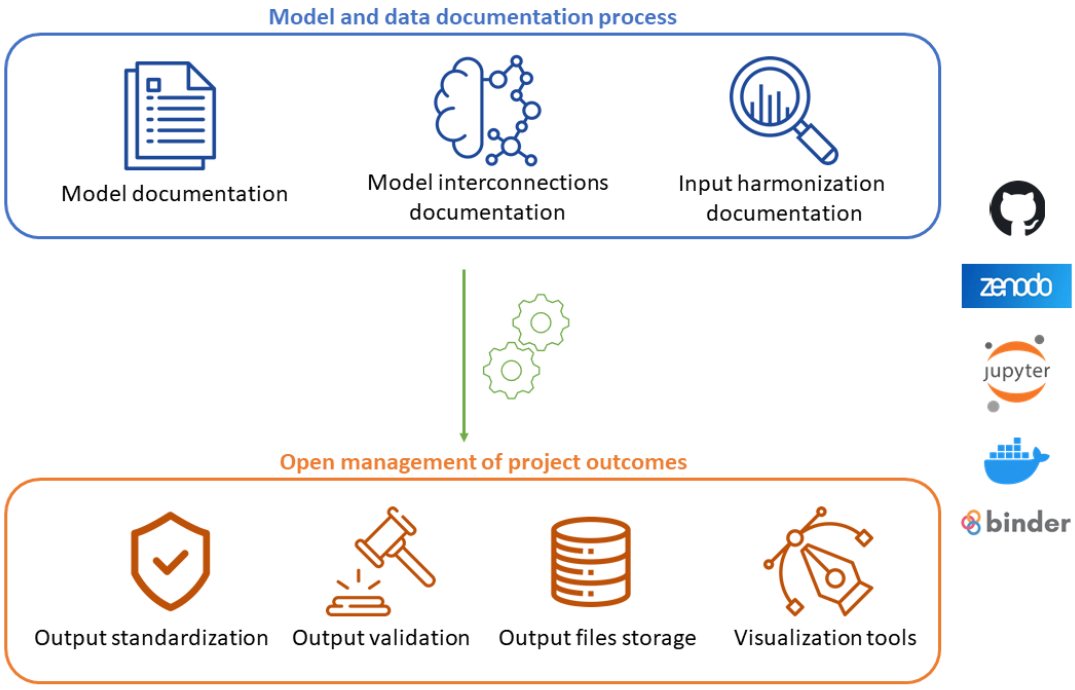


Figure 1: Scheme of the updated Open Science Protocol for IAM COMPACT.

The first stage of the protocol focuses on the **documentation and harmonisation of models and data**. It is essential to ensure that models are understandable, and the modelling assumptions are recorded for future enquiries. To achieve this, we recommend each team to faithfully describe their models in **dedicated platforms** (such as GitHub pages[19] or customised wiki pages, e.g., IAMC wiki[20]), detailing the inputs, outputs, assumptions, and specificities to model versions. Whenever possible, each model version should be linked to a release with an associated Digital Object Identifier (DOI). This ensures the model version is easily discoverable, accessible, and reusable by the research community, while also enabling proper citation in future studies. In the context of IAM COMPACT, all model documentation must also be hosted on **the I²AM PARIS platform**, which was initially

[19] https://pages.github.com/
[20] https://www.iamcdocumentation.eu/IAMC_wiki

developed in the PARIS REINFORCE project (coordinated by NTUA). Based on its initial design, the platform "*seeks to enable modellers to communicate with one another and stakeholders to interact with modelling assumptions, scenarios and results in an informative way and to understand which decarbonisation pathways are the most relevant and realistic, ultimately enhancing the legitimacy of the scientific processes and improving the transparency of the employed methods, models and tools*". This interactive and user-friendly documentation hub will **enable other consortium members and stakeholders to better understand how the models work and how they can be utilised**, ensuring consistency and clarity across all documentation.

For **model interconnections**, **tailor-made documentation** must be uploaded to the appropriate site. On the one hand, the model interconnection, assumptions and input/output interaction can be detailed in specific heatmaps in the I²AM PARIS platform. On the other hand, they should be described in the methodology or Supplementary Information (SI) site of the related publications, which must specify the model versions used. This will promote that models developed by different consortium members can work together in a transparent way.

Finally, to ensure that outputs are comparable **models must homogenise their inputs** in coordination with D4.3 (Korsbakken & Peters, 2023) and its updated version D4.4 (Korsbakken et al., 2024), which define a "broad scenario logic". It describes the common time resolution, input information, or the baseline specifications that modelling teams should implement to the extent possible. It is important that the harmonisation process is documented, for instance, through harmonisation heatmaps provided by the new version of the I²AM PARIS platform that will come out in the following months (D3.9).

The second stage of the protocol focuses on **ensuring the quality and the transparency of produced outcomes**. The starting point is the **homogenisation of the model results** using the time-series data template[21], as described by D4.4 and the *nomenclature* initiative (Daniel Huppmann et al., 2023). This facilitates the cross-comparison of the outputs obtained by different models and modelling teams. Moreover, it enables an automated **vetting procedure** through the automatised I²AM PARIS validation tool[22]. This platform asserts that the model outputs are reliable by comparing early modelled periods to observed data and by ensuring that key balances match, such as demand and supply values or trade flows. Modelling teams can also make use of manual procedures (although not recommended) or model-specific tools (e.g., *gcamreport*[23]). In parallel, model projections require independent ("within-consortium") expert assessment. This process should include a review of the model inputs, assumptions, and outputs. The review should be done in a transparent and collaborative way, with feedback provided to the consortium members to improve the models.

To boost the visibility and impact of key findings and analyses, **model outputs should be made available under a DOI**. This ensures full transparency, accessibility, and reusability of the data, allowing it to serve as a foundation for future analyses. If the raw output data is too large to be uploaded, at least the relevant study-specific results and/or standardised outputs should be deposited in a secure and accessible platform, such as Zenodo. In the same way, **analysis and figure development code must be traceable and open**, in GitHub, Zenodo, or other open-source platforms. These datasets must be accompanied by comprehensive metadata, detailing the contents of the repository, the data-handling processes, and guidelines for reuse. This approach not only preserves the methods and data, but also facilitates its findability, broader use, and code recycling in the research community.

The communication of results could also include **visualisation tools** to effectively disseminate the findings to non-expert audiences. Although developing such tools can be time-consuming, they attract interest from policymakers and stakeholders and **ensure full transparency and acceptance of the results**. Custom websites or R Shiny applications[24] are excellent resources for creating these interactive visualisations. Finally, we note that, in order to ensure the full accessibility and reproducibility of the results, **we are exploring the potential use of additional software platforms** (e.g., Docker, Jupyter Kubernetes, etc.) to solve potential problems related to the reproducibility of the execution environment of the models (when licences allow it) and

[21] https://www.iamconsortium.org/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/
[22] https://validation.i2am-paris.eu/
[23] https://github.com/bc3LC/gcamreport
[24] https://shiny.posit.co/

of the model intercomparison and post-analysis studies.

Overall, **this protocol provides guidance to the consortium members, outlining procedures to ensure that scientific breakthroughs remain accessible and reusable over time.** In parallel, the Open Data Management Plan (D3.2) and Harmonisation protocol (D4.3 & D4.4) provide further details about specific steps, timelines and platforms or software relevant for different phases of the process, facilitating a clear and well-defined protocol application. Given the diversity of models within the project, developing a single, all-encompassing protocol is challenging. Therefore, this protocol establishes a foundational framework for open science practices, which each modelling team can tailor to their specific needs. To ease the implementation of the protocol, we have also created a checklist for use by all modelling teams in the 2nd modelling cycle.

4.1 Protocol checklist

To facilitate the protocol application, the updated Open Science Protocol is accompanied by a **checklist** to be distributed across all modelling teams:

1) **Model and Data Harmonisation Process.**
   a) Model Documentation.
      - A dedicated platform for model documentation, including:
        - Details of inputs.
        - Details of outputs.
        - Details of assumptions.
        - Details of equations.
        - Details of model versions.
      - Assignment of a DOI to model release versions.
   b) Model Interconnections (if applicable)
      - A dedicated section in the paper detailing:
        - Interaction mechanisms between models.
        - Inputs and outputs of the interconnections.
        - Assumptions specific to the interconnections.
        - Model versions used.
   c) Input Harmonisation (if applicable).
      - A dedicated section in the paper detailing:
        - Specific modifications to the inputs and the rationale behind them.
        - Assumptions related to the harmonisation process.
        - Model versions used.
      - Pointer or citation to the specific modified model version, if applicable.
2) **Open Management of Project Outcomes.**
   a) Output Standardisation.
      - Standardisation of model outputs according to the time-series data template.
   b) Vetting Procedure.

- Validation of the first model period against observed data.
- Alignment checks of key outputs, such as demand and supply.

c) Model Output Upload.

- Storage of raw output in an accessible location.
- Storage of study-specific results in an accessible location.
- Availability of analysis code and figure creation code.
- Availability of metadata detailing data origin and handling process.
- Assignment of a DOI to the model outputs and analysis code.

d) Visualisation Tools.

- Provision of comprehensive figures to illustrate the results.
- Development of user-friendly visualisation tools for easy exploration of outcomes.

Funded by
the European Union

Bibliography

Besançon, L., Peiffer-Smadja, N., Segalas, C., Jiang, H., Masuzzo, P., Smout, C., Billy, E., Deforet, M., & Leyrat, C. (2021). Open science saves lives: Lessons from the COVID-19 pandemic. *BMC Medical Research Methodology*, *21*(1), 117. https://doi.org/10.1186/s12874-021-01304-y

Bivek Chaudhary. (2023). *Importance of Data (A Term Paper)*. https://doi.org/10.13140/RG.2.2.23837.69602

Boeckhout, M., Zielhuis, G. A., & Bredenoord, A. L. (2018). The FAIR guiding principles for data stewardship: Fair enough? *European Journal of Human Genetics*, *26*(7), 931–936. https://doi.org/10.1038/s41431-018-0160-0

Byers, Edward, Krey, Volker, Kriegler, E., Schaeffer, R., Riahi, K, Kikstra, J, Lamboll, R, Nicholls, Z, & et al. (2022). AR6 Scenarios Database. *Climate Change 2022: Mitigation of Climate Change, Intergovernmental Panel on Climate Change (978-92-9169-160-9).*, *ISBN: 978-92-9169-160-9*. https://doi.org/10.5281/zenodo.7197970

Chiware, E. R. T., & Lockhart, J. (2024). Open Science. In *Reference Module in Social Sciences*. Elsevier. https://doi.org/10.1016/B978-0-323-95689-5.00061-4

Daniel Huppmann, Laura Wienpahl, Philip Hackstock, & Lucie Castella. (2023). Nomenclature—Working with IAMC-format project definitions (v0.9.1). *Zenodo*. https://doi.org/10.5281/zenodo.7956229

European Commission. (2019). *The European Union's Research and Innovation Performance—Facts and Figures 2019*. https://research-and-innovation.ec.europa.eu/system/files/2019-12/ec_rtd_factsheet_2019_0.pdf

Fecher, B., Friesike, S., & Hebing, M. (2015). What Drives Academic Data Sharing? *PLOS ONE*.

Finley, K. (2011). Github has surpassed sourceforge and google code in popularity. *ReadWrite. Com. Recuperado de Http://Readwrite. Com/2011/06/02/Github-Has-Passedsourceforge*.

Fragkos, P., Van De Ven, D.-J., Horowitz, R., & Zisarou, E. (2024). Analysing the Transformative Changes of Nationally Determined Contributions and Long-Term Targets. *Climate*, *12*(6), 87. https://doi.org/10.3390/cli12060087

Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálfy, M., & Alexis, J. (2020). *Preprinting a pandemic: The role of preprints in the COVID-19 pandemic*.

Gacek, C., & Arief, B. (2004). The many meanings of open source. *IEEE Software*, *21*(1), 34–40. https://doi.org/10.1109/MS.2004.1259206

Gidden, M., & Huppmann, D. (2019). pyam: A Python Package for the Analysis and Visualization of Models of the Interaction of Climate, Human, and Environmental Systems. *Journal of Open Source Software*, *4*(33), 1095. https://doi.org/10.21105/joss.01095

Huppmann, D., Rogelj, J., Kriegler, E., Krey, V., & Riahi, K. (2018). A new scenario resource for integrated 1.5 °C research. *Nature Climate Change*, *8*(12), 1027–1030. https://doi.org/10.1038/s41558-018-0317-4

IAMC. (2022). *IAMC Documentation Wiki*. https://www.iamcdocumentation.eu/index.php/IAMC_wiki

Kaiser, O., & McNeill, G. (2019). OpenAIRE für Repository ManagerInnen—Wie Repository ManagerInnen Open Science unterstützen können. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare*, *72*(2), 373–382. https://doi.org/10.31263/voebm.v72i2.2815

Korsbakken, J. I., Mittal, S., & Peters, G. P. (2024). *IAM_COMPACT_D4.4_Broad_scenario_logic_Update*. Zenodo. https://doi.org/10.5281/zenodo.13132311

Korsbakken, J. I., & Peters, G. (2023). *IAM_COMPACT_D4.3_Broad_Scenario_Logic*. Zenodo. https://doi.org/10.5281/zenodo.10430891

Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., & Jenkyns, R. (2020). The TRUST Principles for digital repositories. *Scientific Data*, *7*(1), 144.

Manghi, P., Artini, M., Atzori, C., Baglioni, M., Bardi, A., La Bruzzo, S., De Bonis, M., Dimitropoulos, H., Foufoulas, I., & Iatropoulou, K. (2017). *OpenAIRE: Advancing open science*. 23–24.

Merriam-Webster. (2024). *Merriam-Webster.com Dictionary*. https://www.merriam-webster.com/dictionary/example

Mittal, S., Frilingou, N., Nikas, A., & Koasidis, K. (2023). *IAM_COMPACT_D3.2_Open_data_management_plan*. Zenodo. https://doi.org/10.5281/zenodo.10849069

Mittal, S., Giarola, S., Nikas, A., Frilingou, N., Horowitz, R., Van de Ven, D.-J., Cui, R. Y., Zhao, A., Garg, A.,

Maheshwari, J. R., Vishwanathan, S. S., Wang, Y., Wang, T., & Fragkos, P. (2023). *IAM_COMPACT_D4.5_National_regional_global_mitigation_pathways*. Zenodo. https://doi.org/10.5281/zenodo.13767330

Mittal, S. & Van de Ven, D.J. (2024). *IAM_COMPACT_NDCs_LTTs_Database* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.12648186

Moreno, J., Campagnolo, L., Boitier, B., Nikas, A., Koasidis, K., Gambhir, A., Gonzalez-Eguino, M., Perdana, S., Van De Ven, D.-J., Chiodi, A., Delpiazzo, E., Doukas, H., Gargiulo, M., Herbst, A., Al-Dabbas, K., Alibaş, Ş., Neuner, F., Le Mouël, P., & Vielle, M. (2024). The impacts of decarbonization pathways on Sustainable Development Goals in the European Union. *Communications Earth & Environment*, *5*(1), 136. https://doi.org/10.1038/s43247-024-01309-7

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Nikas, A., Frilingou, N., Heussaff, C., Fragkos, P., Mittal, S., Sampedro, J., Giarola, S., Sasse, J.-P., Rinaldi, L., Doukas, H., Gambhir, A., Giannousakis, A., Golinucci, N., Koasidis, K., Rocco, M. V., Trutnevyte, E., Xexakis, G., Zachmann, G., Zisarou, E., … Van De Ven, D.-J. (2024). Three different directions in which the European Union could replace Russian natural gas. *Energy*, *290*, 130254. https://doi.org/10.1016/j.energy.2024.130254

Nikas, A., Gambhir, A., Trutnevyte, E., Koasidis, K., Lund, H., Thellufsen, J. Z., Mayer, D., Zachmann, G., Miguel, L. J., Ferreras-Alonso, N., Sognnaes, I., Peters, G. P., Colombo, E., Howells, M., Hawkes, A., Van Den Broek, M., Van De Ven, D. J., Gonzalez-Eguino, M., Flamos, A., & Doukas, H. (2021). Perspective of comprehensive and comprehensible multi-model energy and climate science in Europe. *Energy*, *215*, 119153. https://doi.org/10.1016/j.energy.2020.119153

Open Source Initiative (OSI). (2024). *The Open Source Definition*. https://opensource.org/osd

Peng, G., Gross, W. S., & Edmunds, R. (2022). Crosswalks among stewardship maturity assessment approaches promoting trustworthy FAIR data and repositories. *Scientific Data*, *9*(1), 576. https://doi.org/10.1038/s41597-022-01683-x

Peroni, S. (2021). Review of: 'FAIR RDM (Research Data Management): Italian initiatives towards EOSC implementation'. *Qeios*. https://doi.org/10.32388/LAKK5Q

Reisinger, A., Cowie, A. L., Geden, O., & Al Khourdajie, A. (2024). Science-based targets miss the mark. *Communications Earth & Environment*, *5*(1), 383. https://doi.org/10.1038/s43247-024-01535-z

Rinaldi, L., & Ghezzi, D. (2024). *Exiobase HYBRID extended to model innovative steelmaking routes for EU (v0.1.1) [Data set]*. Zenodo. https://doi.org/10.5281/zenodo.11215808

Rinaldi, L., Ghezzi, D., Colombo, E., & Rocco, M. V. (2024). Assessing environmental and market implications of steel decarbonisation strategies: A hybrid input-output model for the European union. *Environmental Research Letters*, *19*(7), 074059. https://doi.org/10.1088/1748-9326/ad5bf1

Rinaldi, L., Rocco, M. V., & Colombo, E. (2023). Assessing critical materials demand in global energy transition scenarios based on the Dynamic Extraction and Recycling Input-Output framework (DYNERIO). *Resources, Conservation and Recycling*, *191*, 106900. https://doi.org/10.1016/j.resconrec.2023.106900

Rodés-Bachs, C., Sampedro, J., Horowitz, R., & Van de Ven, D.-J. (2023). *IAM_COMPACT_D3.6_Open_Science_Protocols*. Zenodo. https://doi.org/10.5281/zenodo.10464845

Rodés-Bachs, C., Sampedro, J., Horowitz, R., Ven, D.-J. V. de, Cui, R. Y., Zhao, A., Zwerling, M., & Khan, Z. (2024). gcamreport: An R tool to process and standardize GCAM outputs. *Journal of Open Source Software*, *9*(96), 5975. https://doi.org/10.21105/joss.05975

Sampedro, J., Van De Ven, D.-J., Horowitz, R., Rodés-Bachs, C., Frilingou, N., Nikas, A., Binsted, M., Iyer, G., & Yarlagadda, B. (2024). Energy system analysis of cutting off Russian gas supply to the European Union. *Energy Strategy Reviews*, *54*, 101450. https://doi.org/10.1016/j.esr.2024.101450

Sicilia, M.-A., García-Barriocanal, E., & Sánchez-Alonso, S. (2017). Community Curation in Open Dataset Repositories: Insights from Zenodo. *Procedia Computer Science*, *106*, 54–60. https://doi.org/10.1016/j.procs.2017.03.009

Springer Nature. (2024). *The Fundamentals of Open Access and Open Research*.

https://www.springernature.com/gp/open-science/about/the-fundamentals-of-open-access-and-open-research

Terzić, R., & Majstorović, M. (2019). Open data concept, its application and experiences. *Vojnotehnicki Glasnik*, *67*(2), 347–364. https://doi.org/10.5937/vojtehg67-19935

Trutnevyte, E., Sasse, J.-P., Pehle, H., Johannsen, R. M., Thellufsen, J. Z., Rinaldi, L., Rocco, M. V., Nieto, J., Miguel González, L. J., Kleanthis, N., Stavraka, V., Papantonis, D., Manias, N., Flamos, A., & Frilingou, N. (2024). *IAM_COMPACT_D4.9_European_sub_national_deep_dives*. Zenodo. https://doi.org/10.5281/zenodo.13784965

Ulrich, H., Kock-Schoppenhauer, A.-K., Deppenwiese, N., Gött, R., Kern, J., Lablans, M., Majeed, R. W., Stöhr, M. R., Stausberg, J., Varghese, J., & others. (2022). Understanding the nature of metadata: Systematic review. *Journal of Medical Internet Research*, *24*(1), e25440.

United Nations Framework Convention on Climate Change (UNFCCC). (2024). *Integrated Assessment Models (IAMs) and Energy-Environment-Economy (E3) Models*. https://unfccc.int/topics/mitigation/workstreams/response-measures/modelling-tools-to-assess-the-impact-of-the-implementation-of-response-measures/integrated-assessment-models-iams-and-energy-environment-e3-models

Van de Ven, D.J., Sampedro, J, Arenas, J., & Mittal, S. (2024). *IAM_COMPACT_Climate_Policy_Database* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.12648253

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., & Bourne, P. E. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

5 Appendix A

Survey for Deliverable 3.7 – Open Science Protocols

The purpose of this survey is to gather your insights to understand the various methods utilised by the different modelling teams to detect the state of the art, improvement points, and gaps in open science practices. The obtained information will be used for IAM COMPACT Deliverable 3.7 – Open Science Protocols. If you have further comments that what is covered in this survey, you can email Clàudia Rodés (claudia.rodes@bc3research.org) (BC3).

The initial deliverable on open science practices (D3.6) was submitted in June 2023, at the start of the 1st modelling cycle. An updated deliverable (D3.7), due in September 2024, will update and expand on D3.6 while providing updated data on implemented practices and specific recommendations for the second modelling cycle.

This survey is not anonymous, since it is important to be able to follow up and to know which open science practices are being applied by the different modelling teams. If you have any concerns that you do not wish to share directly with BC3, please contact the project administration at NTUA.

In case you select "Other" in any question of the survey, please specify. Use the "Additional Comments" text box below each page if you need to write a longer explanation

* Required

1. About the respondent

This section gathers some data about the modelling team and the contact person.

Question	**Full name**
Question type	Text box

Question	**Email**
Question type	Text box

Question	**Which model do you mainly use in IAM COMPACT?** *If your model is not listed, please select "Other" and specify.*
Question type	Selection boxes
Options	• AIM/Enduse-India • Calliope • China MAPLE • CLEWs • DyNERIO • EnergyPLAN • EXPANSE • GCAM • GCAM-EU • GCAM-USA • MEDEAS • MENA-EDS • MUSE • OSeMOSYS • PROMETHEUS • SLIM-India • TEEM • TIAM • WILLIAM • WISE-ESM • WTMBT • Other

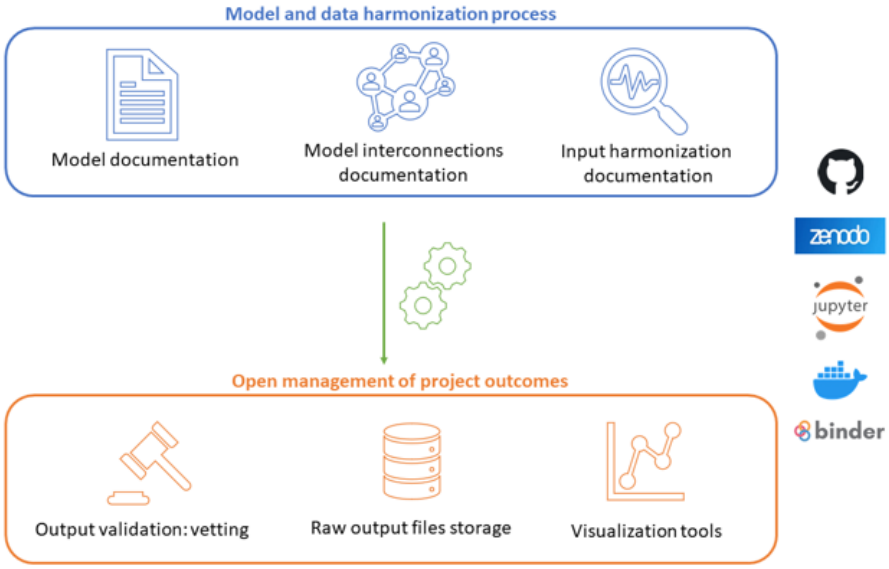
Question	**Which studies, deliverables or other modelling exercises have you participated in during the first modelling cycle?**
Question type	Multi-selection boxes
Options	• Study 1: How does the implementation of the updated draft National Energy and Climate Plans compare to the cost optimal EU approach? • Study 2: What are the different cost, energy security and resilience metrics and how do they compare for different scenarios?

	• Study 3: How could geopolitics affect decarbonisation pathways? • Study 4: What are the implications of EU industry decarbonisation, considering possible off / re-shoring scenarios? • Study 5: What is the contribution of earlier stage technologies if they undergo rapid cost reductions? • Study 6: How do interest rates influence decarbonisation pathways? • Study 7: What are the economic impacts (rather than the cost or effectiveness of policy implementation) of a given behavioral change? • D4.5: National, regional, global mitigation pathways (to include China, India, USA) • D4.7: Sectoral and cross-sectoral analysis • D4.9: European sub-national deep dives • D5.1: Impact assessment of COVID-19 recovery • D5.2: Climate action social & distributional impacts • D5.4: Modelling out-of-ordinary extremes • D5.6: Behavior, social and disruptive innovation • D5.8: Energy crisis modelling • SDG analysis • Other

Question	**Comments and clarifications** *Please add any comments or specifications to what you replied above, or anything else we should know about your modelling activities in IAM COMPACT in order to properly understand your feedback.*
Question type	Text box

2. Usage of D3.6 – Open Science Protocols

This section asks about the Open Science Protocol presented in D3.6 in the first modelling cycle. The protocol structure will be used in the next sections to gather information about open science practices.

Question	**Are you familiarised with the Open Science Protocol suggested in D3.6?** 
Question type	Selection boxes
Options	• Yes • No

1. Block 1: Model and data harmonisation process

This section asks specific questions about the usage of the items detailed in the 1st block of the Open Science Protocol, "Model and data harmonisation process", presented in D3.6 in the first modelling cycle.

i. General questions

Generic questions about the model you are employing in IAM COMPACT.

Question	**Is your model open source?**
Question type	Selection boxes
Options	• Yes • No • The model code yes but not the model data • Other

ii. Model documentation

Models are usually well-documented to help others to reproduce and validate the results; to understand the model purpose, assumptions, and limitations; to provide a roadmap for updating and improving the model; to facilitate new collaborations; and to have version control, among others.

Question	**Is your model documented?** *E.g., GCAM documentation.*
Question type	Selection boxes
Options	• Yes • No

IF YES:

Question	**Which documentation platforms/sites do you use to document your model?**
Question type	Multi-selection boxes
Options	• When available, we are planning to use the I²AM PARIS platform[25] • Models' repository page – please specify • Other

IF NO:

Question	**Why?**
Question type	Multi-selection boxes
Options	• I was unaware of this option • Lack of time

[25] https://www.i2am-paris.eu/#section-works

	• Lack of knowledge • Other

iii. Inputs & Model interconnections documentation

When chaining models, it is common to detail the inputs and outputs of each model and how do they couple together. In the same line, when doing multi-model exercises, scientists usually harmonize the inputs to be able to compare the results.

Question	**If you have ever needed to harmonize your inputs or connect your input/output with other models/tools for specific projects or studies, do you explicitly document these procedures?** *E.g., Gidden et al 2018*
Question type	Selection boxes
Options	• Yes • No • Does not apply

IF YES:

Question	**Which documentation platforms/sites do you or would you use to document model input harmonisation?**
Question type	Multi-selection boxes
Options	• We document the harmonisation/model connection procedure in the methods/SI section of the related paper • When available, we are planning to use the I²AM PARIS platform heatmaps[26] • Models' repository page – please specify • Other

IF NO:

Question	**Why?**
Question type	Multi-selection boxes
Options	• I was unaware of this option • Lack of time • Lack of knowledge • Other

[26] https://www.i2am-paris.eu/harmonisation_map_tool/harmonisation_manual

iv. Additional comments

Question	**Comments and clarifications** *Please add any comments or specifications to what you replied above, or anything else we should know about your open science and data procedures in IAM COMPACT in order to properly understand your feedback.*
Question type	Text box

2. Block 2: Open management of project outcomes

This section asks specific questions about the usage of the items detailed in the 2nd block of the Open Science Protocol, "Open management of project outcomes", presented in D3.6 in the first modelling cycle.

i. General questions

Generic questions about the postprocessing of your model results.

Question	**Do you homogenize your model outputs following the IAMC data series template[27] to facilitate model-intercomparison exercises?** *E.g., homogenised output looks like* Model | Scenario | Region | Variable | Unit | 2005 | 2010 | 2015 | 2020 | 2025 | 2030 GCAM 7.0 | Reference | EU-12 | Agricultural Demand | million t DM/yr | 195.9466191 | 191.68099 | 194.29817 | 204.06022 | 208.64856 | 212.21884 GCAM 7.0 | Reference | EU-15 | Agricultural Demand | million t DM/yr | 778.5847498 | 852.83469 | 914.19359 | 951.80849 | 973.94775 | 995.71437 GCAM 7.0 | Reference | World | Agricultural Demand | million t DM/yr | 974.5313689 | 1044.51568 | 1108.49176 | 1155.86871 | 1182.59631 | 1207.93320
Question type	Selection boxes
Options	• Yes • No • Does not apply

IF NO:

Question	**Why?**
Question type	Multi-selection boxes
Options	• I was unaware of this option • Lack of time

27 https://www.iamconsortium.org/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/

	• Lack of knowledge • Other

ii. Output vetting validation

To be able to validate a model output, it is common to use an exhaustive procedure or an external tool. This gives consistency and legitimacy to the results.

Question	**Do you use an output vetting validation procedure?** *E.g., IAM2Paris validation tool*
Question type	Selection boxes
Options	• Yes • No

IF YES:

Question	**Which kind of vetting do you perform?**
Question type	Multi-selection boxes
Options	• First model period compared to observed data • Values stay within realistic ranges • Other

Question	**Is the vetting procedure automatic?** *tomatic" we refer to embedded validations that do not require from an expert assessment*
Question type	Selection boxes
Options	• Yes • No

Question	**Do you use an external tool to do the vetting?**
Question type	Multi-selection boxes
Options	• Yes, through the IAM2Paris validation tool[28] • Yes, through a package (e.g., gcamreport[29]) • Yes, through another tool – please specify • No

Question	**Which items are vetted?**

[28] https://validation.i2am-paris.eu/
[29] https://github.com/bc3lc/gcamreport/

Question type	Multi-selection boxes
Options	• The whole output • Specific items – please specify • None

Question	**Which year(s) do you vet?**
Question type	Text box

IF NO:

Question	**Why?**
Question type	Multi-selection boxes
Options	• I was unaware of this option • Lack of time • Lack of knowledge • Other

iii. Output files storage

To facilitate the results reproducibility and transparency, the output files are usually stored. We distinguish two different output flows:

- Raw output files: the direct model output files, such as scenario results.

- Study output files: the code and datasets used to generate the figures and analyses of the study.

Study output files are usually associated with a determined publication, while raw output files are usually scenario outputs (not necessarily linked to a specific study).

Question	**Do you usually store your raw output files under a Digital Object Identifier (DOI)?** *v output files" we refer to the datasets directly generated by the model, without any manipulation.* *raham et al 2020.*
Question type	Selection boxes
Options	• Yes • No

IF YES:

Question	**Which platforms/sites do you use to store your raw output files?**
Question type	Multi-selection boxes
Options	• DataHub • Zenodo • Other

IF NO:

Question	**Why?**
Question type	Multi-selection boxes
Options	• I was unaware of this option • Lack of time • Lack of knowledge • Too heavy data • Other

Question	**Do you usually store your <u>study output</u> files (processed data and/or analysis code) under a Digital Object Identifier (DOI)?** *ıdy output files" we refer to the datasets used to generate the figures and analysis code of a certain study, deliverable, or publication.* *mpedro et al 2024.*
Question type	Selection boxes
Options	• Yes • No

IF YES:

Question	**Which platforms/sites do you use to store your <u>study output</u> files?**
Question type	Multi-selection boxes
Options	• GitHub • Zenodo • Other

IF NO:

Question	**Why?**
Question type	Multi-selection boxes
Options	• I was unaware of this option • Lack of time • Lack of knowledge • Too heavy data • Other

iv. Visualisation tools

To engage with the public, facilitate the results analyses, and increase the results transparency, it is common to present the outputs though user-friendly visualisation tools.

Question	**Do you usually present results through interactive user-friendly**

Funded by the European Union

	visualisation tools? *E.g., rfasstUI.*
Question type	Selection boxes
Options	• Yes • No

IF YES:

Question	**Which platforms/sites do you use?**
Question type	Multi-selection boxes
Options	• When available, we plan to use the I²AM PARIS platform (e.g., Post-Glasglow Climate Targets Study[30]) • Community/Team page (e.g., GCAM Foresight[31]) – please specify • RShiny widget – please specify • Other

IF NO:

Question	**Why?**
Question type	Multi-selection boxes
Options	• I was unaware of this option • Lack of time • Lack of knowledge • Other

v. Additional comments

Question	**Comments and clarifications** *Please add any comments or specifications to what you replied above, or anything else we should know about your open science and data procedures in IAM COMPACT in order to properly understand your feedback.*
Question type	Text box

[30] https://www.i2am-paris.eu/feasibility/scientific_module
[31] https://github.com/JGCRI/foresight/

Funded by the European Union

3. Block 3: Tools for open science practices

This section asks specific questions about the tools you used to adhere to the Open Science Protocol presented in D3.6 in the first modelling cycle or you usually employ to develop your studies.

Question	**Which open-source tools do you use to follow the Open Science Protocol?**
Question type	Multi-selection boxes
Options	• GitHub[32] • Zenodo[33] • Docker[34] • Jupyter Notebooks[35] • Binder[36] • Other

Question	**Did you create any open-source tool(s) to follow/facilitate the Open Science Protocol?** *E.g., gcamreport*
Question type	Selection boxes
Options	• Yes • No

Question	**How do you use the above-mentioned tools?** *Please describe in which procedures do you use them and if it is a sporadic or a regular use.*
Question type	Text boxes

Question	**Comments and clarifications** *Please add any comments or specifications to what you replied above, or anything else we should know about your open science and data procedures in IAM COMPACT in order to properly understand your feedback.*
Question type	Text box

[32] https://github.com/
[33] https://zenodo.org/
[34] https://www.docker.com/
[35] https://jupyter.org/
[36] https://mybinder.org/

3. Feedback & Suggestions

This section asks open questions to gather your feedback and suggestions for the Open Science Protocol presented in D3.6.

Question	**What parts of the Open Science Protocol presented in D3.6 did you find particularly useful, what could be improved, and what if anything should be added in D3.7?**
Question type	Text box

Question	**Do you have any concerns about following the protocol?**
Question type	Text box

Question	**Do you have any other feedback or suggestion not covered above?**
Question type	Text box

Funded by
the European Union

6 Appendix B

Survey results

1. About respondent

ID	Full name	Email	Which model do you mainly use in IAM COMPACT? If your model is not listed, please select "Other" and specify.	Which studies, deliverables or other modelling exercises have you participated in during the first modelling cycle?	Comments and clarifications
1	Panagiotis Fragkos	fragkos@e3modelling.com	PROMETHEUS;	Study 1; Study 3; Study 7; D4.5; D5.1; D5.2	
2	Gonzalo Parrado Hernando	gonzalo.parrado@uva.es	MEDEAS; WILLIAM;	Study 4; Study 7; Study 2; D5.1; D5.8	
3	Noelia Ferreras Alonso	noefer@cartif.es	WILLIAM;	Study 7; D5.8	
4	Natasha Frilingou	nfrilingou@epu.ntua.gr	GCAM;	Study 6; Energy crisis modelling;	
5	Thomas Nikolakakis	tnikolakakis@epu.ntua.gr	OSeMOSYS; OSEMBE;	Study 6	
6	Rasmus Magni Johannsen	rmj@plan.aau.dk	EnergyPLAN;	Study 2	
7	Nikos Kleanthis	kleanthis@unipi.gr	OSeMOSYS; TEEM;	Study 1; Study 2; Study 6; Study 7; D4.7; D4.9; D5.1; D5.6	
8	Clàudia Rodés Bachs	claudia.rodes@bc3research.org	GCAM;	Study 1; Study 7: D5.8; Energy crisis modelling; D5.2; Study 2; Study 5; Study 6	
9	Lorenzo Rinaldi	lorenzo.rinaldi@polimi.it	DyNERIO; MARIO;	Study 2; Study 4; Study 3; Energy crisis modelling;	

Funded by the European Union

2A. Usage of the Open Science Protocol

ID	Are you familiarised with the Open Science Protocol suggested in D3.6?	Is your model open source?	Is your model documented?	Which documentation platforms/sites do you use to document your model?	If you chose "Models' repository page", please specify
1	Yes	No	Yes	When available, we are planning to use the I2AM PARIS platform (https://www.i2am-paris.eu/#section-works); Models' repository page	https://e3modelling.com/modelling-tools/prometheus/
2	No	Yes	Yes	Models' repository page; When available, we are planning to use the I2AM PARIS platform (https://www.i2am-paris.eu/#section-works)	https://github.com/LOCOMOTION-h2020/WILIAM_model_VENSIM/wiki/The-WILIAM-model
3	No	Yes	Yes	Models' repository page	https://github.com/LOCOMOTION-h2020/WILIAM_model_VENSIM/wiki
4	Yes	Yes	Yes	When available, we are planning to use the I2AM PARIS platform (https://www.i2am-paris.eu/#section-works); Models' repository page	
5	No	Yes	Yes	Models' repository page	OSEMBE repository page in github (https://github.com/HauHe/OSeMBE) and also there is a publication regarding OSEMBE (https://doi.org/10.1016/j.energy.2021.121973)
6	No	No	Yes	Models' repository page	https://www.energyplan.eu/training/documentation/
7	Yes	Yes	Yes	When available, we are planning to use the I2AM PARIS platform (https://www.i2am-paris.eu/#section-works); Models' repository page	You can find the documentations of the models used in the following pages: 1) TEEM DREEM: https://github.com/TEESlab-UPRC/DREEM | 2) TEEM ATOM: https://github.com/TEESlab-UPRC/ATOM | 3) OSeMOSYS: https://github.com/OSeMOSYS/OSeMOSYS
8	Yes	Yes	Yes	Models' repository page; When available, we are planning to use the I2AM PARIS platform (https://www.i2am-	https://jgcri.github.io/gcam-doc/

				paris.eu/#section-works);	
9	No	Yes	Yes	Models' repository page	https://mario-suite.readthedocs.io/en/latest/

ID	If you have ever needed to harmonize your inputs or connect your input/output with other models or tools for specific projects or studies, do you explicitly document the steps involved in these processes?	Which documentation platforms/sites do you or would you use to document model input harmonisation?	If you chose "Models' repository page", please specify	If not, why?	Comments and clarifications
1	No			Lack of time;	
2	Does not apply				We were not involved in papers about model inter-comparison studies. We are currently developing the necessary code to allow for it
3	Yes	We document the harmonisation/model connection procedure in the methods/SI section of the related paper;			I have experience on some data input harmonisation activities done for some specific sstudies or specifically for including parts of other models in WILIAM in order to add new functionalities (specifically regarding the "climate module" of WILIAM) where data inputs harmonisation has been necessary. However I have not come to a systematic procedure to do it neither I have done this type of activity in IAM COMPACT (data harmonisation has been done by UVa as I have said before).
4	Yes	We document the harmonisation/model connection procedure in the methods/SI section of the related paper;When available, we are planning to use the I2AM PARIS platform heatmaps (https://www.i2am-paris.eu/harmonisation_map_tool/harmonisation_manual);			There is an extensive harmonisation general protocol for IAM COMPACT studies, detailed in deliverable D4.3 - Broad Scenario Logic & updated for the 2nd cycle in deliverable D4.4

5	Yes	We document the harmonisation/model connection procedure in the methods/SI section of the related paper;			OSEMBE is an EU power sector model that we use to develop decarbonisation pathways. However, the electricity demand and carbon cap trajectory are inputs in the model. Those inputs were obtained from (https://doi.org/10.1016/j.energy.2021.121973). Technoeconomics of electricity generation technologies were obtained from WEO 2022.
6	Yes	We document the harmonisation/model connection procedure in the methods/SI section of the related paper;			
7	Yes	We document the harmonisation/model connection procedure in the methods/SI section of the related paper;			
8	Yes	We document the harmonisation/model connection procedure in the methods/SI section of the related paper;When available, we are planning to use the I2AM PARIS platform heatmaps (https://www.i2am-paris.eu/harmonisation_map_tool/harmonisation_manual);			
9	Yes	ZENODO;			

2B. Usage of the Open Science Protocol

ID	Do you homogenize your model outputs following the IAMC data series template [1] to facilitate model-intercomparison exercises?	If not, why?	Do you use an output vetting validation procedure?	Which kind of vetting do you perform?	Is the vetting procedure automatic?
1	Yes		Yes	Output values are robust. E.g., production and demand match;	No
2	No	Lack of time; Lack of knowledge;	No		
3	Does not apply		No		
4	Yes		Yes	First model period compared to observed data;	Yes
5	Yes		No		
6	No	Outputs are not generally relatable to outputs from integrated assessment models, and we have therefore not previously found it relevant.;	No		
7	Yes		No		
8	Yes		Yes	Output values are robust. E.g., production and demand match; First model period compared to observed data;	Yes
9	No	Model is not an IAM;	No		

Funded by
the European Union

ID	Do you use an external tool to do the vetting?	If you chose "through another tool", please specify	Which items are vetted?	If you chose "specific items", please specify (not a full list required, just macro sectors)	Which year(s) do you vet	If not, why?
1	No;		Specific items	energy demand, supply, power mix, CO2 emissions	2010,2015	
2						Lack of time; This step is applied after harmonisation to IAMC format (previous necessary step to be done);
3						I was unaware of this option; Lack of knowledge;
4	Yes, through a package (e.g., gcamreport, https://github.com/bc3lc/gcamreport/);	gcamreport	The whole output		2020	
5						I was unaware of this option; We are not there yet. But when we get results, we will consult the broader team.;
6						Outputs (variables and formatting) differ drastically for EnergyPLAN to IAMs - thus we do not see any value in adopting this form of validation. We typically do vetting and quality checks based on historical statistical data.;
7						I have used the IAM2Paris validation tool before. Even though the file has the correct column format, I get the following error "KeyError: 2020".;
8	Yes, through a package (e.g., gcamreport, https://github.com/bc3lc/gcamreport/);		Specific items	ch4 and co2 emission, primary energy, electricity secondary energy, and trade balances	2020	
9						Lack of knowledge;

ID	Do you usually store your raw output files under a Digital Object Identifier (DOI)?	Which platforms/sites do you use to store your raw output files?	If not, why?	Do you usually store your study output files (processed data and/or analysis code) under a Digital Object Identifier (DOI)?	Which platforms/sites do you use to store your study output files?	If not, why?
1	No		Lack of time; Lack of knowledge;	Yes	Zenodo;	
2	Yes	Zenodo;		Yes	Zenodo;	
3	No		I was unaware of this option; Lack of knowledge;	No		I was unaware of this option; Lack of knowledge;
4	Yes	Zenodo;		Yes	Zenodo;	
5	No		Too heavy data;	No		We are not there yet. But we will do so using Zenodo;
6	No		We typically include these as appendices or additional data in reports and scientific publications, but rarely under their own DOI.;	No		We typically include these as appendices or additional data in reports and scientific publications, but rarely under their own DOI.;
7	No		Lack of time; Lack of knowledge;	No		Lack of time; Lack of knowledge;
8	No		Too heavy data;	Yes	Zenodo; GitHub;	
9	Yes	Zenodo;		Yes	Zenodo;	

ID	Do you usually present results through interactive user-friendly visualisation tools?	Which platforms/sites do you use?	If you chose "Community/Team page", please specify	If you chose "Rshiny widget", please specify	If not, why?	Comments and clarifications
1	No				Lack of time;	
2	No				Lack of time; Lack of knowledge;	
3	Yes	https://www.locomotion-h2020.eu/locomotion-models/model-explorer/;				
4	Yes	When available, we plan to use the I2AM PARIS platform (e.g., Post-Glasglow Climate Targets Study, https://www.i2am-paris.eu/feasibility/scientific_module);				
5	No				I was unaware of this option;	Our study on WACCs is still ongoing; we plan to update our results according to the study protocols
6	No				Lack of time; Lack of knowledge; Most communication of results is done through visualisations included in publications. We have rarely done interactive visualisations, and thus don't have much experience or capacity presently to do so.;	
7	Yes	When available, we plan to use the I2AM PARIS platform (e.g., Post-Glasglow Climate Targets Study, https://www.i2am-paris.eu/feasibility/scientific_module);(Interactive) platforms of EU-funded projects, e.g., H2020 SENTINEL (https://sentinel.energy/stories/) & H2020 ENCLUDE (https://energycitizenship.eu/);				
8	Yes	When available, we plan to use the I2AM PARIS platform (e.g., Post-Glasglow Climate Targets Study, https://www.i2am-		gcamreport		

		paris.eu/feasibility/scientific_module);RShiny widget;				
9	No				Lack of knowledge;	

2C. Usage of the Open Science Protocol

ID	Which open-source tools do you use to follow the Open Science Protocol?	Did you create any open-source tool(s) to follow/facilitate the Open Science Protocol?	Please specify	How do you use the above-mentioned tools?	Comments and clarifications
1	Zenodo (https://zenodo.org/);	No		to store scenario data in Zenodo	
2	GitHub (https://github.com/);Zenodo (https://zenodo.org/);Jupyter Notebooks (https://jupyter.org/);	No		To present the raw data of studies and novelties of the model. Every release of WILIAM is uploaded to Github	
3	GitHub (https://github.com/);gitlab;	Yes	https://github.com/LOCOMOTION-h2020	Github and gitlab we use it to manage WILIAM model and the different versions. Github includes all the related elements of WILIAM as repository (tools, wiliam version, wiki, etc.)	
4	Zenodo (https://zenodo.org/);GitHub (https://github.com/);	No		Documenting datasets supporting scientific publications	
5	GitHub (https://github.com/);Zenodo (https://zenodo.org/);	No		We are not there yet but when our study is conclude we will upload our results in Zenodo and we also aim to come in contact with the OSeMBE team to update the GitHub Repository	
6	Zenodo (https://zenodo.org/);	No		Sharing reports, documentations, data, and model outputs.	
7	GitHub (https://github.com/);Zenodo (https://zenodo.org/);	No		Regular use of GitHub for model documentation and Zenodo for uploading project deliverables	
8	GitHub (https://github.com/);Zenodo (https://zenodo.org/);Docker (https://www.docker.com/);	Yes	gcamreport (https://github.com/bc3LC/gcamreport)	to document procedures, to host code and releases, to store data, to host tools, to create programming environments to run tools	

9	GitHub (https://github.com/);Zenodo (https://zenodo.org/);Jupyter Notebooks (https://jupyter.org/);	No		Support evidences from studies and deliverables with raw input data and model code	

3. Feedback and suggestions

ID	Which parts of the Open Science Protocol presented in D3.6 did you find particularly useful, what could be improved, and what if anything should be added in D3.7?	Do you have any concerns about following the protocol?	Do you have any other feedback or suggestion not covered above?
1			
2	The standardisation and harmonisation of results into IAMC format is very useful and synergetic with other models and projects		
3			
4	The scheme to be used in IAM COMPACT	No, it's quite straight-forward	
5			To be honest I have learned much about the protocol through this questionnaire. I will consult the broader team to inform myself better to continue ahead
6			
7			
8			
9			