Earth4All Modeling Team response to recent questions and criticisms
 (version 5.Feb.2024)

Before going into a detailed discussion about the Earth4All-Global model, we would like to make some general remarks that hold true for any (type of) model. In general computer models are tools that make explicit the modelers' mental models and allow them to run experiments within this. While the famous saying by George Box reads "All models are wrong but some are useful", it has been widely acknowledged that actually all models might be useful for their specific purpose. All models are idealizations of reality, and this often provides a lot of room for criticism. Criticisms are welcome and often a great source of insight in itself as it opens up the debate about different mental models of the systems in question and the underlying assumptions. In fact, some of the most famous and yet most criticized models have not only provided valuable insights into how our world might work but also shown us why some simplifying assumptions about it might lead us astray (e.g. the DICE model by Nordhaus).

Donella Meadows (1982, 2008) wrote several critiques/reflections about models and their role in policy. One of the main conclusions was that models should be designed for a particular purpose/to answer a particular question. This is also what defines the boundaries (i.e. what variables to include, what relationships to include etc) for that specific model.

Some questions about the quality of the E4A-Global-model do not take into account what this global simulation model was made to do. The model was not made to do what many of its critics apparently expect it to do. No model can do what some critics seem to expect E4A-global to do – namely to forecast the future of the real world with "scientific" precision.

What can and does the Earth4all-global model say about humanity's long-term future? The quick answer is "very little" but also "much more than nothing". The model illustrates the story of the Earth4All book by simulating two scenarios resulting from the model's structure and input parameters. Together with the assessments of the 21st-century Transformational Economics Commission (TEC) and the storylines that were outlined in the book, the model paints two consistent but contrasting pictures of potential futures with broad pen strokes.

What is the E4A-global model - and what is it not?

1. This model is a global integrated systems model.
It simulates the development over time of human wellbeing towards 2100 given the model's structure – focused on trends in the wellbeing of the global majority.

2. The E4A-global model is a very rough representation of the real-world system, formulated at a very high level of aggregation. Hence it is incapable of saying anything about local detail. Except in the form of informal deduction from global developments in the model system. This is why an additional regional model has been developed (E4A-regional model to be released in 2024).

3. In principle any model can only say something rigorously about what happens in the model system. The relevance of such statements to the real world depends on whether the model system is similar (in the eyes of its users) to the real system in the aspects that are assessed.

4. The E4A-global model can help keep foresight thinking straight.
That is by ensuring that scenarios are internally consistent, both in space and time

5. The E4A-global model can help clarify future trends – especially in the business-as-usual (BAU) scenario.
That is, it can try to describe what is likely to happen if decision making continues to follow the pattern from the last forty years as represented in the model. For example, what will happen if decision making in the energy field leads to continued decline in energy intensity by the historical 1 % per year.

6. The E4A-global model can provide rough quantitative estimates of central variables in BAU based on the model's structure.
That is, it can give back-of-the envelope sketches of the order of magnitude of central variables, and their rate of change. Such estimates depend (of course) on both the specification of structure and choice of parameter values. Hence it is important to be open about the most important assumption. For example in E4A we connect the estimated values of future population, GDP, energy use, emissions, etc to the underlying assumptions in the form of births, deaths, GDP growth rate, energy use, and emissions as functions of GDP per person (derived from extensive calibration with historical data (1980-2020) achieving statistically significant goodness of fit measures).

7. The E4A-global model can demonstrate the sensitivity of the BAU scenario to variation in parameters and structure.
That is, it can help clarify the effect of change in policy, structure and parameters. For example, the effect of experimentally introducing additional structural cross links can help to explore the magnitude of a (new) causal relationship or driver. E4A-global can also be used to evaluate different 'what-if' scenarios by changing parameters and assessing how trajectories change given the model structure.

Four overarching issues when assessing the Earth4All-Global model

1. Is the purpose of the Earth4All-global model to predict the future?
No, the Earth4All-global model is made to "help increase broad understanding of the system of interest" (structuralist school). It is not the type of models that are made to "forecast the long-term future with high precision" (empirical school, see the table below, excerpted from Clancy et al 2023, https://doi.org/10.1002/sdr.1740).

Table 1. Summary of schools of thought

School	Goal of confidence building	Relationship to and nature of audience	Preferred tools of confidence building	Audience's aesthetic for model presentation	Definition of quality
Empirical	Findings are true and replicable.	Indirect relationship to scientific community	Numerical fit to observed data	Mathematical equations and scientific notation.	Publication in high-quality, rigorous academic journals.
Structural	Findings have explanatory verisimilitude to a real-world system.	Indirect relationship to complex systems they are seeking to explain	Structural fidelity	Vivid explanatory structural diagrams and plain language.	Increase broad understanding of complex phenomena.
Pragmatic	Findings help solve a problem within the constraints provided the sponsor works within.	Direct relationship to stakeholders whose problem they seek to solve.	Defined by stakeholder or client needs.	Day-to-day management tools, including dashboards, flight simulators, and so forth.	Defined by stakeholder or client needs.
Methodological	Findings improve the method of System Dynamics.	Direct relationship to systems scientists, including system dynamicists.	Making a mathematically rigorous connection between behavior and structure and the creation of software tools to facilitate methods.	It varies but includes a mix of mathematical and visual presentations.	How broad is the use of the methods and improve performance of the field?

There are important differences between building "models to help policy making/help change the world" versus building "models that represent specific scientific findings". As a result, some model builders (like us) deliberately choose to downplay quantitative precision and detail, to focus on providing an overarching understanding of the system's main characteristics, namelyits structure and behavior. See Donella Meadow's *Thinking in Systems* (2008, https://www.chelseagreen.com/product/thinking-in-systems/) for an extensive discussion on the choices of both granularity and system boundary of a model.

2. If the Earth4All-global is a "global system model" – does that mean it is an all-round model for all kinds of global issues?
No, quite the opposite. All models are idealized representations that are built to answer very specific questions. The perfect all-round model of the whole wide world does not exist.
Since people see the world differently, model builders (who also are people) differ fiercely in what they view as the best model for a given purpose.
The root cause of such disagreements is the fact that all social large systems are so

complex that it requires (sweeping, heroic) simplifications to avoid ending up with an impenetrable "black box" – streams of code producing voluminous outputs that may be comprehensive or exhaustive but cannot be understood, by people. As the map can never be the territory, a model can never be reality.

Consequently, models should not be used for purposes other than the one they were made to elucidate. And they should certainly not be run further in time than their designed time horizon.

3. Is the Earth4All-global model the oracular source of the main solutions (the "Turnarounds") to create systems change?

The main message of the Earth for All book is that implementing 5 Turnarounds (eliminate poverty, reduce inequality, increase opportunity, switch to regenerative energy and food production) will improve the wellbeing of the global majority – when wellbeing is defined as a combination of a) disposable income, b) public services, c) inequality, d) environmental quality, and e) social tension. You don't need a model to present and defend this conclusion. And clearly the priorities and conclusions will depend on how you weigh the components. One should not trust black-box models, whether system dynamics, equilibrium models, large language models or big-data-models, so much that one starts to believe in conclusions that can not be explained and defended in plain arguments, words and reasoning. Models may help us to spot counter-intuitive connections, see systemic patterns more clearly, and broaden our understanding of the world. But models can never replace values or conversations for guiding policies and common actions to achieve agreed solutions. In sum, the Earth4All-global model illustrates the two scenarios by giving them quantitative backing but does not provide details on how the world would behave.

4. What can a simulation model be used for?

Simulation models can only provide if-then answers: "if we assume this structure and these parameters, then this time development will result". So strictly speaking, a simulation model can only say something about the development of the model system.

Sadly, it is not possible to "prove" that the time development of the model system says something about the time development of the real world. All one can do is to highlight the similarities between model and reality. It helps if the model is capable of reproducing history, if its cause-and-effect links are intuitively plausible and have empirical and theoretical backing, and if the set of assumptions form a logically consistent totality. But none of these prove that "the model is right".

"All models are wrong, but some are useful", George Box, https://doi.org/10.1080/01621459.1976.10480949

More specific responses to quotes from the recent RSHK criticism:

Critical arguments and questions are copy-pasted from their PDF in Bold.
Team **responses** are in normal text

1) Why is the "E4A-model neither (food) mass balance nor energy balance consistent model?"
Our purpose for the E4A project is to answer this research question: *How can humanity achieve more wellbeing within planetary boundaries* (https://www.science.org/doi/10.1126/science.1259855) *this century?* That implies that the model focuses on environmental fates and *sinks,* not on tracking all the sources and material transformation pathways, etc.

In designing the E4A-global (May 2022 version), we chose to focus on a higher level of abstraction for the main economic dynamics. Occam's razor has been ruthlessly applied to make the E4A-global model as parsimonious as possible, in order to have a pedagogical vehicle where the model structure has a strong explanatory verisimilitude to the real world. The E4A-regional model (to be released in 2024) has a climate sub-module (ESMICON) that does have higher granularity, and hence it includes consistent carbon- and heat- stocks.

E4A-global is a recently developed model, and we acknowledge that descriptions of the model released so far may not have been comprehensive enough in communicating the reasons behind the design choices regarding the level of aggregation, the system boundaries, and so on. But in releasing the model as open-source and making it available for anyone who wants to run or modify it, we provide a space for peer scrutiny and experimentation.

[BTW: In comparison, the World7 global model is not open-source, nor is it a full *material*-model. It is a much bigger (at least ~10X bigger) *economic* model]

2) The E4A-global is "dominated by preset values and forcing functions" …
"**It means that a variable that has a set value for 1980 will be forced to get a pre-set value in 2100 by command and smoothing functions to make the commended change appear as reasonable. The implication is that the output value is reached in the "model" because the model builders dictated the output value.**"

It is a bit unclear what is meant by "preset values". If this is a reference to the historical 1980-2020 numbers, then that is correct - the model starts from best-available historic data in 1980, and then recreates main trends to 2020.

And with "forcing functions" we understand Ragnarsdottir et al to mean what we called "guidelines", mainly based on our many f(GDPpp), which are used to estimate future trajectories in main variables. To get toward estimates of how large aggregates of human beings choose to use scarce resources over time at a high level into the future, we want to be evidence-based. We have large amounts of data and documentation on the calculation of the E4A-guidelines. Like many other

scholars, we have observed and documented that GDPpp/yr (as a *proxy for society's wealth*) correlates well with many different human behaviours, particularly for the 1980-2020 period. The justification for using those guidelines is strong; see for instance birth- and death rates in the population report (https://earth4all.life/wp-content/uploads/2023/04/E4A_People-and-Planet_Report.pdf) , and the Collste (2021) for SDG1-7 (https://doi.org/10.1017/sus.2021.26), and Earth3-work (https://eartharxiv.org/repository/view/1200/). We aim to make this documentation more easily accessible and upload our data resources to be more transparent. Several additional technical publications are planned.

In our opinion the word "forcing" is not a suitable descriptor of the use of these evidence-based functions, which is why we recommend speaking of guidelines instead. The huge benefit of using such guidelines is that we can coherently simulate several aspects of human behaviour – average historical national responses and decision making under real-world stress and bounded rationality, with respect to jobs, investments, pregnancy, education, and national development. The model therefore uses these guidelines of how GDPpp/yr correlates with these aggregated behaviours, in order to create future estimates.

So this is not a case of **"This is equivalent to "it has that value of X because I say so!" No reason is given why, commands that are used are all totally unsubstantiated."**

What the critics call "pre-set values to 2100" are actually the policies that we want to explore.

In E4A-Global, the rate of change in the Total Factor Productivity (the TFP factor) is preset to 1%/yr to reproduce the historical trajectory 1980-2020. In the Too Little Too Late scenario, we let it continue as a baseline, all the way to 2100 (building on the conventional Cobb-Douglas-function from mainstream economics). In mov220518-E4A-Global the TFP in TLTL is increased from 2022 over historical levels by a preset value. In review now, it seems like the E4A-Global model also includes a 0.3% increase in the TLTL scenario. The team today sees this is an unfortunate choice, initially made to better align the TLTL with the SSP2 middle-of-the-road scenario used in many climate assessments , but this alignment with other research groups is an issue we may want to revisit in future work.
The E4A-regional model approaches this differently. In E4A-regional, the TFP-rises as a function of investment in public capacity.

To run Giant Leap (GL) in the E4A-global model, obviously there are a lot of normative choices, in order to explore policy outcomes that are central to the GL story in the book. We explore a higher rate of change in TFP in the GL scenario. And for instance, the fraction of cropland worked regeneratively is forced, because we simulate setting a political target and assume there will be sufficient measures and regulations to force this target to be met. In the TLTL scenario however, this is *not* pre-set.

3) "Several of the stocks present in the E4A model have been short-circuited"

In our team we are somewhat confused by what is meant with short-circuited? - And which stocks? Maybe the critics are referring to the fact that delay and smooth functions have been used in which stocks are "hidden". But this is rather common practice, used in many system dynamics models.

4) "Both the GL and the TLTL (BAU) scenarios wreck the world after 2100 because the forcing functions are only set until 2100."

This is a design choice. The stated purpose of the model is to explore well-being within planetary boundaries within this century, with a main focus to 2060 (40 years back + 40 years forward + visualising effects for another 40 years, to 2100).

The time horizon of this model is not intended to go beyond 2100, and it is an error to run it beyond 2100.

5) "The 5 transitions have no representation in the E4A "model"

The modeling approach used in the mov220501-E4AGlobal is to have a main scenario TLTL, and then run what-if-experiments where extraordinary interventions (5 "Turnarounds") are introduced from 2022, through specific levers, involving ~15 variables, that can impact the model behaviour and output over the coming decades. The Turnarounds can be run one by one, or selected in any combination / group one chooses.

6) "E4A is not a proper dynamic model based on causal connections, feedbacks and mass balances"

This comment is unfounded, since the E4A-global model contains many evidence-based causal connections, for instance of GDPpp on fertility, as well as on energy and food demand.
Also, the E4A-global model includes a causal link from government spending on the rate of technological advance, as is clearly seen from the "Public sector" sheet in the published model (mov220501-18 E4A-global).

We agree that the E4A-global is not fully endogenous, but it has nevertheless given useful insights and prompted valuable discussions. As long as people find models useful, they are cordially invited to improve them. We ourselves have taken this approach in the Earth4All regional model which includes many more causal feedbacks than in mov220501-E4A-global. .

7) "The population module in E4A is fatally flawed and does not constitute a valid population model" … "Changing the years of fertility, yields a very strange result. Shorter years of fertility gives a larger population, longer years of fertility a smaller population.

This comment seems to be a case of a muddled misreading of the model.

We looked again at the equations.. The model calculates births as *"Aged 20-40 years Mp" * "Fraction women (1)" * (Observed fertility 1 / Fertile period y)* where *Observed fertility 1* is the number of children a woman has during her lifetime. The units should thus be Children (p) per woman (p) per lifetime (y) = (p/p)/y but the model sets the units to 1 and drops the per year. To then get the births (units p/y) right, the model has introduced a variable that has the units of (per year). The model creates a constant for this, *Fertile period (y)* and divides by this. But, *Fertile period* is also conceptually contained in the product of *"Aged 20-40 years Mp" * "Fraction women (1)"*, so if you change *Fertile period* you <u>must</u> also add to the number of women in the child bearing cohort, which the model does not do automatically. The experimenter thus must do that themself, otherwise they run a nonsensical experiment.

The basic assumption in E4A-global is that all births take place by women between 20 and 40 years of age. They get the desired number of children during this 20-year period. Hence the parameter "Fertile period" cannot be changed from its current value "20 years" without changing the model structure. So, the E4A-global model is correct – both in formulation and units. But it is based on a conceptualization which makes it meaningless to change the value of the parameter "Fertile period" from its current value of "20 years".

" Why was population module from World3 and in principle nearly everything else in World3 ignored?"

There should be no prohibition on learning new things during the past 50 years.

The criticism of how people die in the model is valid, because for design purposes, we decided to simplify to only deaths happening in the last cohort. If we use life expectancy data, thisalready includes non-age-related deaths. Consequently, to include earlier deaths in the model would lead to double counting. Avoiding that would require us to adjust life expectancy data to include only age-related deaths.

Again, in the E4A-Regional model we do this at a higher level of granulation (by using 5-year cohorts per 10 world regions, with age-specific death-rates in the cohorts where deaths historically exceed 2% per year).

The outcomes of these two approaches are largely compatible, as shown in the detailed [E4A Population report](#).

8) "The food and agricultural module in E4A-global lacks basic components and does not constitute a valid agricultural model. The food supply model is absent, and does not constitute a valid model for global food supply."
"The fertility and mortality seem to have no connection to food availability. "

These statements are true, and this is a design choice. This is based on the thinking that in a world with average increasing purchasing power, those with that purchasing power will generate an inelastic demand that food producers will meet. The issue of starvation is not lack of food, but lack of purchasing power.

In the design of E4A-global model we looked at the high-level impact of wealth and emissions, not on feeding the world, as in Gerten et al (2020). Our main interest was in climate change, how wealth improves food security, and looking at nutrient flows and their effect on planetary boundaries.

We agree that the E4A-global food-submodule is very high level. Any criticism can only be relevant compared to the purpose of the model. If one chooses to disagree with the purpose, one should build a new model with one's own, and different, purpose.

We ourselves chose that route when building our Earth4All-regional model. The E4A-regional has a much more granular food sub-module with stocks of different types of land-use (again in order to gain relevance for PBs - emissions, nutrients, land-use). Both in E4A-Global and E4A-regional the resource-stocks are not highest priority because we build on the PB framework, where the critical boundaries are sinks, not sources.

9) "**There is no economic model. The l**a**bor-market module yields more employed people than the working age population** a**nd is not mass balance consistent with the population module.**" Why?

While there is no single "Economy" module, economic mechanisms are simulated through several modules. In particular, economic growth is modeled using a standard Solowian formulation, which is commonly used by most standard economic models of long-term growth. While we agree with RSHK that economic development is a complex phenomenon poorly represented by TFP, the studies referred to by RSHK belong to the economic history tradition, and their granularity is hardly consistent with a highly aggregated model such as E4A-global.

In the model runs presented in the book (and other runs we have been working on), the workforce is always smaller than the working-age population, see figures below. If reviewers have come up with other runs where this is not the case, we ask them to provide us with the details of those runs.
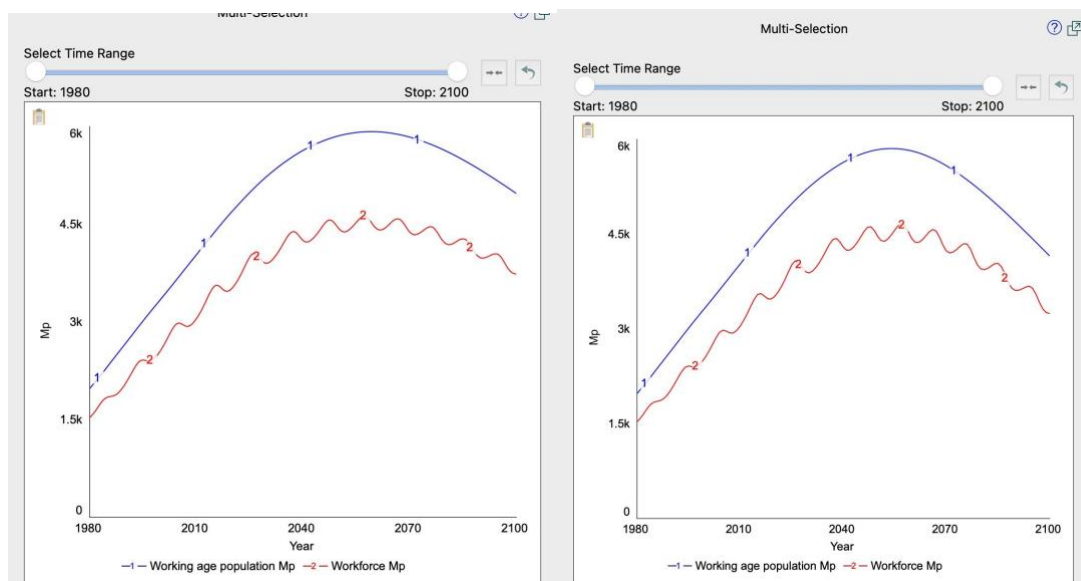
10) "There are *no natural resources* in E4A … Energy is done by creating renewable energy without any metal/material limits. There are no energy balances anywhere in the E4A model." Why?

The main reason is that the Earth4All project focuses on human wellbeing within the planetary boundaries, conceptualised as global sinks, not supplies of natural resources. The main human problem this century is not lack of resources or metals. Usually, warnings about running out of materials or energy are derived from forecasts of population that are ever-increasing, and multiplying this growing population with average per capita consumption of materials and energy of today. However, in Earth4All we foresee different future trajectories in both regards, for these reasons:
1) As people become wealthier, all of humanity will live through the demographic transitions that the rich countries already have behind them. Thus, by 2100 there will be fewer people on Earth than today. As an indicator of this, not only Europe and Japan are getting less populous, even China is.
2) Material and energy consumption follows a sigmoidal limit curve, and with increasing energy and material efficiency this results in a declining per capita consumption for very rich societies. Empirically, this has already been seen for energy in the USA.The model assumes this will also happen for other regions as they get richer.

Thus, until 2100, the primary concern of humanity will not be running out of resources, but the lack of resilience and capacity to deal with the human-caused pressures on the 9 planetary boundaries.

11) "The E4A "model" can neither be used for global policy development, nor for regional development. Without a functioning global model, making regional sub-models from the global one is just nonsense. Making regional models that do not add up to the global one is a sign that something at the fundamental level is wrong."

Valid criticism of a model is most useful when it limits itself to the model and refrains from speculating about the presumed further research agenda of the modelers.

The E4A-regional model, yet to be published, is inspired by the E4A-global model, but has regional sub-models that stand on their own, with a novel structure, and often more granular flows and stocks. It "adds up" to the global historical data.

12) "Poverty eradication is "solved" by cash handouts and printing the money needed. This method is well known from Germany in the 1920'ies, but also from countries like Argentina, Venezuela, Soviet Union or other hyperinflation countries that regularly destroy their currencies and create debts and poverty in the process."

The main driver of poverty mitigation is GDP growth, combined with mitigated demographic pressures. The key policy driver is redistributive policies, reallocating purchasing power from owners to workers. Helicopter money is not relevant for the process. We would also like to point out that the national cases of hyperinflation mentioned are hardly relevant in the context of a global model.

Some concluding reflections:

→ All models are wrong, some are useful. And they are only useful for the specific purpose for which they are made.

→ Disagreeing with the purpose, modeling approach and design choices makes a weak basis for declaring a model "fundamentally flawed" and asking CoR to "distance itself" from work and the scenarios which the model explores.

→ The E4A-regional (2024) model addresses many of the characteristics and concerns that critics find lacking with E4A-global (from 2022). It has higher granulation (with for instance 10 world regions, 20 age cohorts for each), more endogenization, many more stocks and process details, including in the food and natural resouces sectors. .

→ The world and its people are heading in a dangerous direction of pushing Earth systems into increasingly high risk zones. Our aim in the Earth4All project is to facilitate social change in another direction through changes in narrative, investments and policies (the 5 Turnarounds), not to build the perfect, ultra-detailed global system model.