EAEE 4000: Environmental Engineering for Environmental Engineering and Sciences, Final Report Fall 2022. Utilizing Reinforcement Learning for Optimal Aerosol Injection.

Samarth Agrawal, under guidance of Pierre Gentine. Extending work done by Violette Launeau
December 30th, 2022

Abstract

In this project, we train Reinforcement Learning agents to try and determine the optimal deployment timeline for hypothetical deployment of Stratospheric Aerosol Injection to combat short term climate change. Stratospheric Aerosol Injection is a proposed form of Solar Radiation Modification that is known for being cost effective and relatively easy to implement. We use the Finite Amplitude Impulse Response Model to define various climate forcers within OpenAI's gym environment interface. From this, a variety of Reinforcement Learning models were deployed to find carbon, sulfur, and other forcing species emission quantities that best balanced climate anomaly and economic feasibility constraints captured within the reward function. Results indicate that aerosol injection is indeed a powerful tool for reduction of temperatures and climate forcing that is affordable and fast-acting. However, experiments show that delayed deployment has little long term impacts and there is no fear of "locked in" effects in the same way as CO2.

1 Introduction

Much controversy surrounds proposals for "geoengineering" and other deliberate direct anthropogenic interventions in the climate system. Despite concerns over safety, equity, and governance, the fact remains that the global mobilization needed to effectively curtail greenhouse gas emissions through traditional means has yet to appear. The climate is hotter than at any other point in the last 1000 years (Masson-Delmotte et al., 2018) and if current emission policies are retained, the world is set to blow well past the warming limits set by the Paris Climate Agreement. By some estimates, current emission pathways result in a projected 3℃ increase by the end of the century which would cause unacceptable damage. There are numerous ways to respond to climate change, including climate mitigation (tackling the problem at its source by transitioning the economy towards greener practices that cause less harm), climate adaptation (reducing the impacts of climate effects on human society), and climate intervention- deliberate large scale manipulation of geophysical processes. Climate interventions (sometimes also referred to as geoengineering, although the term can be ill-defined) tend to be faster, cheaper, but also riskier responses to climate change and are designed to accompany, not replace, mitigation and adaptation (Zarnetske et al., 2021). Despite valid social and political hesitancies involving climate interventions, the substantial uncertainties regarding what these proposals could look like and what effects they may have make it incredibly difficult to have an informed discussion in the first place.

Alongside direct carbon dioxide removal (CDR), Solar Radiation Management (SRM) is the most commonly discussed form of climate intervention. An umbrella term for reducing sunlight absorption by the planet, growing interest in SRM has increased research into various proposals.

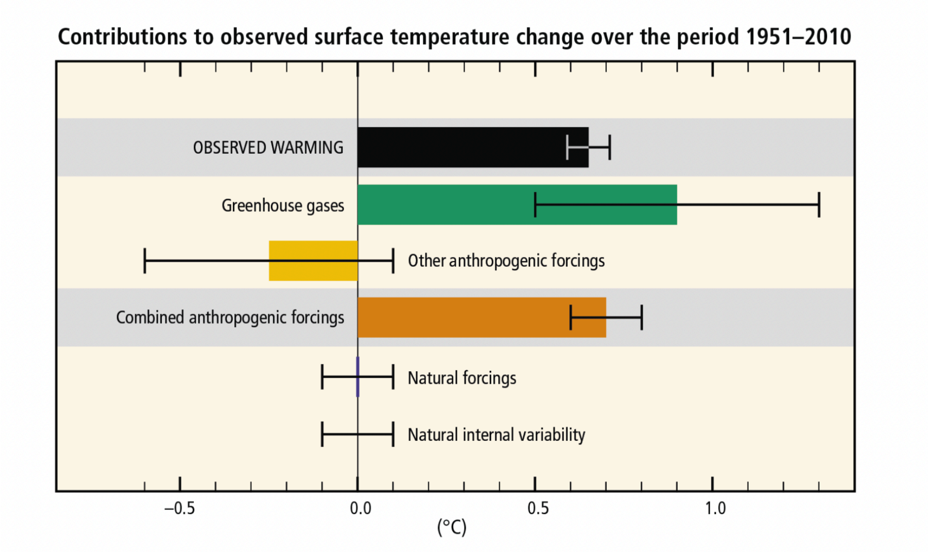


Figure 1: Radiative Forcing from different sources and their contribution to surface temperature.

The driving process behind global warming is the greenhouse effect, in which higher concentrations of greenhouse gases are trapping solar energy and thus preventing it from radiating out into space. While reducing emissions or removing carbon dioxide from the atmosphere combats the direct cause of increased global temperatures, solar radiation management reduces temperatures without altering the concentrations of greenhouse gases (Dunne, 2018) (MacMartin et al., 2018). These proposals are designed to change effective radiative (or climate) forcing, a measure of the energy flux of the atmosphere. Radiative forcing, calculated as the difference between incoming and outgoing solar energy in Watts per square meter ($\frac{W}{m^2}$), is a proxy for how much human activity has impacted the climate and is a critical parameter for determining temperature anomaly (Figure 1). SRM proposals are ways to decrease this energy flux and include marine cloud brightening, ocean and urban albedo modification, cirrus cloud thinning, and others. The most well-known form of SRM is Stratospheric Aerosol Injection (SAI). SAI receives the most attention because aerosols, which are suspended fine particles, such as sulfates are already present in the atmosphere and are empirically known to decrease radiative forcing. Aerosol injection also has the potential for the largest scale deployment at very low costs, meaning it could feasibly have an enormous impact Zarnetske et al. (2021). Aerosols scatter incoming solar radiation which increases the atmosphere's ability to reflect sunlight (Figure 2). The eruption of Mount Pinatubo in 1991 offers an oft-cited case study in the efficacy of aerosol injection; around 20 million tons of SO_2 lowered global temperatures by around 0.5℃ for 3-4 years. While it is unlikely this reduction in temperatures will positively influence every climate externality caused by greenhouse gases (for example, ocean acidification due directly to high concentrations of carbon dioxide), lowering temperatures buys time and stalls positive feedback loops that would hasten all other climate damages (MacMartin et al., 2018).

Figuring out if and how and where and by whom these interventions will take place remains a critical question of international governance and scientific deliberation. There are significant risks that must be taken into account, such as increases in surface UV, acid rain, changes to

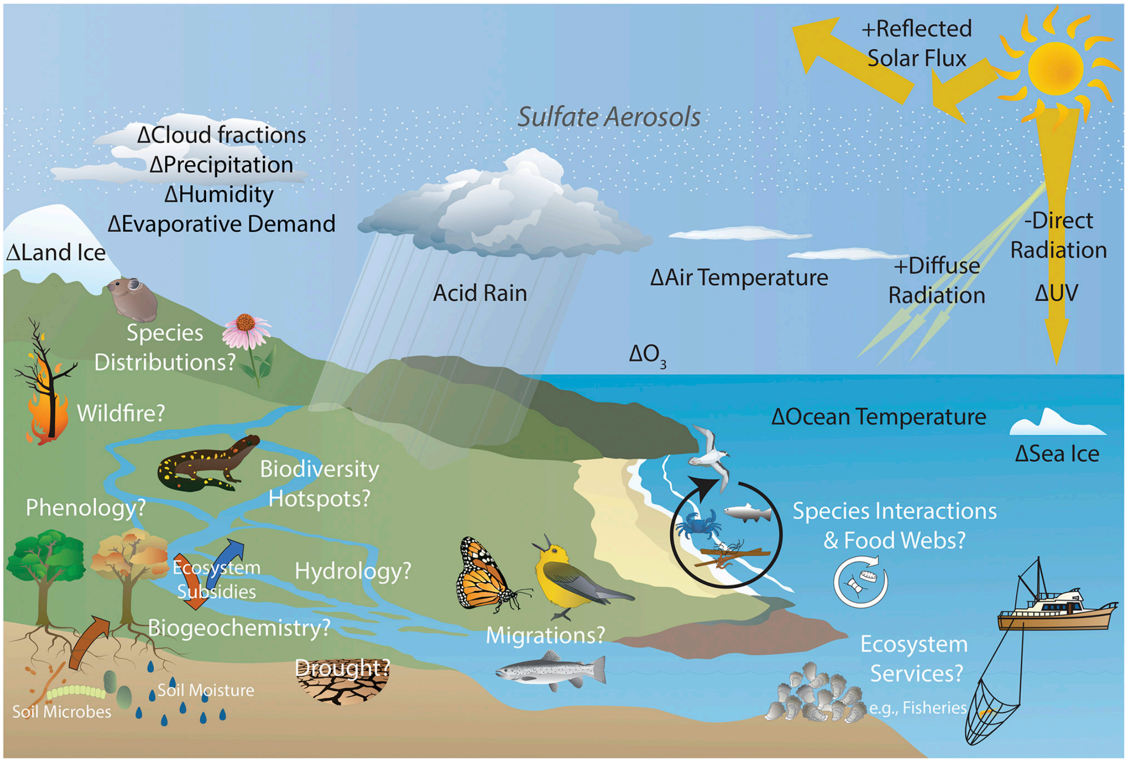


Figure 2: Stratospheric Aerosol Injection may bring about a cascade of secondary effects which as of now still carry many uncertainties Zarnetske et al. (2021).

global hydrological cycles (which may exacerbate local droughts), and polar ozone destruction. This project aims at providing insight toward the implications of the "when" question. Due to the many known and unknown risks as well as the lack of information on what a responsible implementation of stratospheric aerosol injection would actually look like, it is highly unlikely any large scale effort will be launched in the immediate future. How do different intervention timelines affect the efficacy of aerosol injection? Much of the purpose of geoengineering is to buy time while mitigation strategies shift in accordance to the magnitude of the problem. How much time do we lose while we deliberate this information-urgency trade-off? To what extent does the speed at which equivalent magnitudes of intervention are pursued change the effect on the climate system? This project is an extension of the work done by Violette Launeau, in which it was determined that the speed of emissions reduction was in fact a factor in how effective decarbonization was. Similarly, we seek to determine the extent to which speed matters for stratospheric aerosol injection.

2 Method

2.1 Setup of Climate Model: FaIR 2.1

Based on the work by V. Launeau, the Finite Amplitude Impulse Response (FaIR) Climate model was chosen to model the earth system response to various inputs (Launeau, 2022). Unlike larger

Earth System Models, FaIR is lightweight and easy to run. It takes in emissions and climate parameters as inputs and returns forcing, concentration, and temperature data for multiple bounded boxes representing the state of the world (Figure 3). FaIR facilitates running multiple experiments under a wide variety of conditions without much computing power. Previous work has shown that combining Reinforcement Learning with FaIR recovers sensible emissions policies and that shaping reward functions appropriately allows flexible selection of desired outcome. Large modifications were made to the climate model used previously. Most significantly, using FaIR Version 2.1 Leach et al. (2021) transformed the climate model into a class based instance. Climate parameters and species configurations were stored in instance variables and configurations of a model object rather than in a parametrized feed-forward function as in FaIR version 1.6.4. This improvement allowed for easy parallelization. Different hyperparameters, specie, climate, and emission configurations could be run side by side and stored in multidimensional xarrays. That meant that default behavior could be computed alongside aerosol injection. This vastly sped up runtime efficiency and compute time, allowing for more rapid iteration. Compared to FaIR V1.6.4 (used initially and in previous works), a single step of the environment in full multigas mode has gone from taking approximately 45 seconds per step to now less than 2 seconds. Since training the model involves thousands of such steps, the usability increase is substantial. The downside is a larger upfront setup complexity, since setting up the class itself is quite involved. However, that can be sidestepped if one establishes a reference model outside of the environment whose parameters can be easily stored for a quick reset. This front loads the complexity, taking it outside the Reinforcement Learning setup. Each model is initialized using baseline emissions from the Representative Concentration Pathways and Shared Socioeconomic Pathways, a set of scenarios capturing different levels of societal response in the face of climate change Riahi et al. (2017).

FaIR 2.1 allows the user to specify over 60 species inputted as either emissions, concentrations, forcers, or calculated secondarily. Radiative forcing is calculated from the combined effect of CH4, N20, Sulfur, Black Carbon, Organic Carbons, NH3, NOx, Volatile Organic Compounds (VOCs), and CO emissions as well as aerosol-radiation and aerosol-cloud interactions. FaIR uses an n-layer energy balance model that translates forcing to temperature anomaly Leach et al. (2021). We used 3 layers, the first corresponding to the atmosphere and surface, the second the top and middle ocean, the third the deep ocean. A set of default configuration properties of all species (including forcing efficiency, partition fraction, impulse response, and 28 others) were used for each species and then modified to mirror the RCP scenarios. Chosen climate parameters for the layered energy model are in Table 1; select species configurations are in Table 2. Volcanic and stochastic forcing was also fed in based on historical data as well as RCP projections; projections for future years is calculated using all forcing species FaIR has to offer. In our run, carbon and aerosol emissions were altered as inputs (with all other emissions keep default values) while forcing, temperature anomaly, and atmospheric concentrations are all model outputs.

To implement Stratospheric Aerosol Injection specifically, it was assumed the emitted aerosol would be a sulfate species, more specifically sulfur dioxide precursor (which chemically reacts with water in the atmosphere to become H_2SO_4, also known as sulfuric acid). Sulfate aerosols are the most well-understood species with respect to their atmospheric behavior and negative climate forcing impacts MacMartin et al. (2018). They also have the longest history in terms of proposed solar geoengineering (Wigley, 2006). However, research is being done on alternative

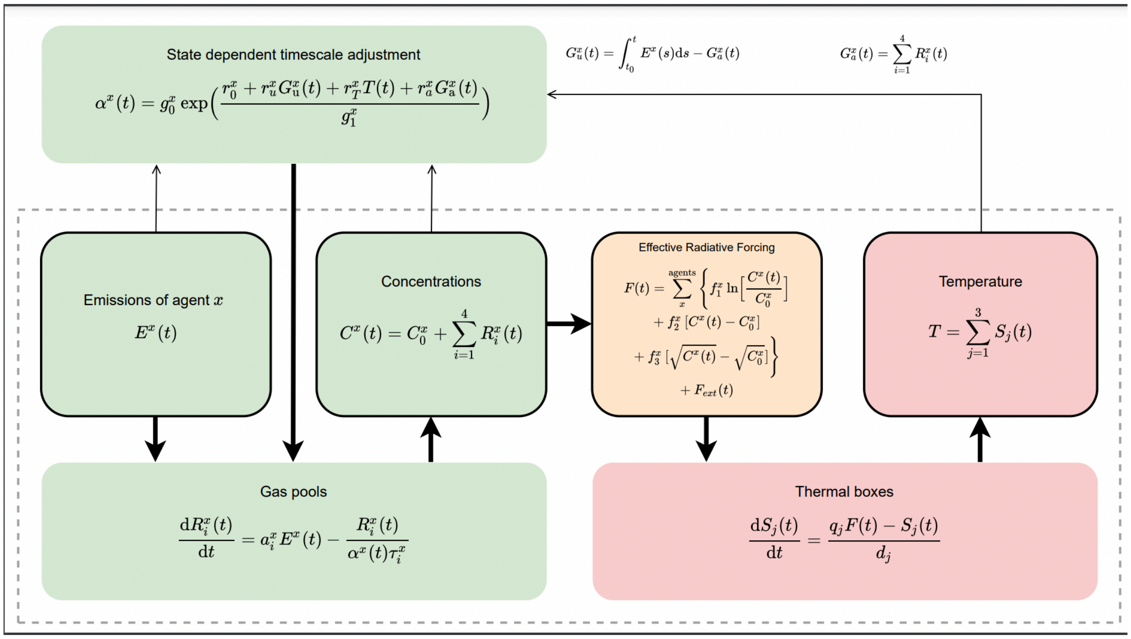


Figure 3: FaIR Version 2.1 Critical Equations and underlying workflow.

aerosol species due to known negative side-effects. Sulfur dioxide is a toxic pollutant; in addition to the harmful impacts mentioned above, it is uncertain what regional impacts high concentrations of SO_2 will have on human and ecological health. However, no other aerosol currently offers the same cost and forcing advantages; for modelling purposes we will use sulfur but it is expected that if widespread SAI is globally pursued that another aerosol will supplant it. Furthermore, while running the model the amount of necessary emitted sulfur for a given change in net forcing seemed abnormally high (almost an order of magnitude higher for comparable effects to the literature), indicating that the radiative efficiency is lower than it should be. It is possible that the species forcing properties for sulfur were not reasonably calibrated. If this work was extended, more attention would be given to more accurately model the behavior of all involved gases; alternatively future work could simulate the impacts of various other aerosol candidates for a full comparative analysis. However, these limitations were acceptable since it is more the shape of the optimal emissions scenario and relative annual changes that are desired, not absolute values.

As in previous works, we scaffold our RL environment using OpenAI's open source Gym class. Compliance with Stable Baselines environment standards was enforced as well. Reinforcement Learning is structured very differently than other Machine Learning methods more commonly applied to the environmental sciences. It typically operates as the interplay between a sequential decision making agent and an environment whose state is dependent on the actions taken by the agent. The agent is trained to maximize a reward provided by the environment based on its actions. At each timestep, the agent makes some observation which partially or fully reveals the state of the environment. In our setup, the observation space is a 5 dimensional vector corresponding to the temperature anomaly relative to pre-industrial levels in degrees Celsius, the amount of CO2 emissions of that year in Gigatons carbon (GtC/yr), the amount of Sulfur emissions in Megatons (MtS/year), the concentration of carbon dioxide in the atmosphere in parts per million (ppm), and the net radiative forcing of all species in watts per square meter (W/m^2). Our actors are trained using a variety of Deep Reinforcement Learning models using the Stable Baselines 3 python library, a reportoire of state of the art algorithms implemented in PyTorch and intended to display reliable performance for research purposes. The 2-dimensional action space of an actor after having observed the environment represents the change in CO2 emissions compared to the SSP baseline as well as the amount of sulfur dioxide to emit into the atmosphere. In addition to the state change, our environment calculates a reward in each time step according to how well it mitigates climate damage while staying within feasibility and economic cost restrictions. The various rewards used in this project will be discussed later in a more detailed explanation of various experiments.

(1) Table 1: Climate configurations for Energy Balance Model

Climate Parameter	**Unit**	**Value**
Ocean Heat Capacity	$\frac{Wyr}{m^2K}$	[3.96786, 16.35885, 108.10587]
Ocean Heat Transfer	$\frac{W}{m^2K}$	[1.1148, 2.5071, 0.83508]
Deep Ocean Thermal Efficacy	Unitless	1.149
CO2 Quadrupling ERF	$\frac{W}{m^2}$	8.0
σ_{ETA} (std of stochastic forcing)	$\frac{W}{m^2}$	0.8897
σ_{XI} (std of stochastic surface disturbance)	Unitless	0.5519
γ forcing autocorrelation	Unitless	5.735

(2) Table 2: Select Species properties for Sulfur Dioxide

Property	SO_2
tropospheric adjustment	0.0
forcing efficacy	1.0
forcing temperature feedback	0.0
forcing scale	1.0
molecular weight	64.069
baseline emissions	2.440048435
concentration per emission	0.08805280331056461
contrails radiative efficiency	0.0
erfari radiative efficiency	-0.00286
h2o stratospheric factor	0.0
lapsi radiative efficiency	0.0
land use cumulative emissions to forcing	0.0
ozone radiative efficiency	0.0
aci scale	-1.03057237
aci shape	0.0169
ch4 lifetime chemical sensitivity	0.0
lifetime temperature sensitivity	-0.0408

2.2 Rudimentary Economic Model

To forestall the model simply halting all global emissions and emitting enormous amounts of sulfur, certain feasibility and economic constraints had to be imposed. The SSP scenarios are more than just climate projections; they are complete societal narratives complete with estimates of urbanization, GDP growth, and demographic factors Riahi et al. (2017). Although data aiming at incorporating this information into a pseudo-realistic economic model was queried, unfortunately time constraints prevented the development of any robust economic model. As such, many of the cost constraints and economic calculations were inherited from the previous work Launeau (2022). We keep track of three variables: GDP ($\mathcal{Y}$), the cost to global GDP due to climate change ($\mathcal{Y}_{cost}$), and the renewable energy knowledge stock ($\mathcal{S}$). For GDP, an annual growth rate of 3 percent is assumed to be the global average from 2023 to 2100, in line with historical macroeconomic trends. We used the same relationship between temperature anomaly and economic cost outlined in the previous work, namely that

$$\mathcal{Y}_{\text{cost temp}} \ (\%GDP) = -\frac{10}{5}\mathcal{T}_{anomaly} - 2 \tag{3}$$

Therefore, a temperature anomaly of 1.5℃ would cause a 1% decrease in GDP. Each additional 0.5℃ of warming would cause another 1% of GDP contraction. Additionally, we assume that the transition to a net zero economy is accompanied by some economic difficulty. According to various studies, it will cost in the ballpark of 50 Trillion USD to achieve a net zero carbon economy by 2050. Dividing by the current 36.6 GtC of global emissions we estimate very very roughly that it costs $1.3 \cdot 10^{12}$ USD to reduce global emissions by 1 gigaton. We add this economic cost to our GDP calculations.

In addition to the indirect cost from the temperature anomaly, the direct cost of implementing Stratospheric Aerosol injection was computed, based off of Smith (2020). Because planes or balloons that are capable of delivering large aerosol payloads at the necessary 20km elevation do not currently exist, the author modelled (following an organizational decision to pursue SRM) commissioning the design, manufacture, and operation of such aircrafts. These costs would be in addition to the cost of the sulfur itself. Several important simplifications were made in this implementation compared to the calculations made by the paper. Firstly, rather than a series of updated aircraft fleets that advance technologically throughout the century, in this project a singular "hybrid plane" was modelled with behavior loosely analogous to a hybrid of all three generations. Secondly, the author assumed that after a decade or so, a more suitable aerosol than sulfur would be found. This compound would have greater radiative efficiency and presumably lack the toxicity of SO_2. For our cost assumptions, this was not factored in. These and other simplifications seemed reasonable as the total cost of SAI is very low relative to global GDP and the economic costs of climate change; even the cumulative trillion dollar estimate for the next 75 years is a drop in the bucket relative to global GDP or the likely total damage caused by extreme weather events (Smith, 2020). After all, it is its very cheapness that makes SAI an appealing purchaser of time for its advocates. Thus these parameters likely would have a fairly low impact on GDP cost and thus on the reward function. In this estimate for how much SAI would cost between the present day and 2100, the costs were decomposed into 3 components:

(4)

$$\mathcal{Y}_{\text{cost sulfur}} \quad (\$) = C_{construction} + C_{operation} + C_{sulfur}$$
$$C_{construction} = N_{\text{airplanes constructed}} * \frac{Cost}{Airplane}$$
$$C_{operation} = N_{\text{airplaces in operation}} * \frac{Cost}{month} * 12$$
$$C_{sulfur} = MtS_{emitted} * \frac{Cost}{Megaton}$$

The number of airplanes needed for a given amount of emissions was based off the annual delivery capacity for an individual airplane (calculated as a hybrid of the parameters of the hypothetical fleet). The marginal airplane manufacturing, fleet operation, and per ton sulfur costs were all drawn from Smith (2020).

Cost of Sulfur	\$1000/ton
Marginal Airplane Construction Cost	\$100,000,000/plane
Monthly Operation Rate	\$800,000/month
Annual Sulfur Delivery	33,000 tons

$$\text{Total Cost to GDP: } \mathcal{Y}_{cost} = -1.3T * \Delta C_{emit} + \mathcal{Y}_{\text{cost sulfur}} + Y_{\text{cost temp}}$$

There are several limitations to this model whose correction was outside the scope of this project. Firstly is the vastly oversimplified model for the costs of decarbonization. To actually assess the feasibility of different carbon emissions policies, a more in-depth analysis (perhaps based on SSP scenarios) should be done. Secondly, the costs of SAI will likely not primarily stem from the direct cost of aviation and material. Although this model is global in scope, regional effects due to the vastly increased atmospheric sulfur concentrations will likely be a huge factor in the decision to implement SAI. However, such fine grained calculations of how injection may affect extreme weather events, human and ecological health, and economic activity are outside FaIR's capability. We are giving a huge benefit of the doubt regarding the merits of SAI and all results should be interpreted in light of that.

2.3 Data

Unlike what may be typical in a supervised learning approach, there is not much imported data in this project. The model is itself learning to generate an optimal emissions policy based on outputs from the gym Environment (which itself is powered by the climate model). Almost all of the data came internally from the FaIR module. FaIR 1.6.4 contained quantitative baseline data on the various RCP and SSP scenarios for 40 different emissions species and 31 different concentration species. These ranged from pre-industrial estimates going back to 1765 to projections until 2500. Unfortunately, these could not simply be imported from fair.scenarios because our environment was in FaIR 2.1 (which did not have access to these). So, the labels for these datasets were scraped from the Fair 1.6.4 documentation web archive using BeautifulSoup and compiled into several pandas DataFrames which were then saved as CSVs that any model version could access. Thus, even though they are read in externally in the code, the SSP scenarios are actually pulled from FaIR itself. RCMIP data, from the Coupled Model Intercomparison Project (CMIP6), is accessible in FaIR 2.1 and can be automatically summoned from the internal method $fair.FAIR.fill\_from\_rcmip()$. This was used to create a reference model pre-loaded with concentration and emission data for the 64 species built into FaIR. Some external SSP reference data was pulled from the SSP database Riahi et al. (2017) but this data was not used due to sparcity. The last source of data was the github of the FaIR examples, linked in the documentation: `https://github.com/OMS-NetZero/FAIR/tree/master/tests/test_data`. There a CSV of past Volcanic Forcing data as well as the climate configs listed in Table 1 were both scraped and integrated into the env parameters.

3 Experiments

In this section, we describe the various tests run using our model.

- Our first test is a validation that the model is setup correctly and is responsive to the reward function. Therefore, initially we will apply only a simple temperature anomaly reward that will try to keep temperatures as close to the Paris climate goal of 1.5℃ as possible by any means necessary. Excessively negative forcing and global cooling are not preferred so the absolute value of the temperature anomaly will be minimized.

$$\mathcal{R}_1 = 100 * (1.5 - T_{anomaly})$$

- **Experiment 2:** We would like the model to determine an optimal emissions and aerosol policy that balances climate effects, feasibility, and economic impact. The purpose of this experiment is to see what sort of policy the model learns to try to keep emissions and temperatures low while having access to geoengineering. Several variables are used to compute the reward. First is the current feasibility, defined as the difference between emitted carbon and the "Middle of the Road" Shared Socioeconomic Pathway. Since that is seen as the most likely pathway, feasible emissions are defined by those emissions that are closer to it. Higher emissions are not viewed as infeasible, so instead of absolute value, $cf = C_{emit} - C_{ref}$. Since we wish to remain within the Paris Climate hard limit of 2℃ of warming, the model is punished strongly for any warming above 2℃ and less strongly rewarded for warming below 2℃. Finally, the model is rewarded for high GDP and low cost. Perhaps tinkering with the coefficients of these various parameters would result in a more optimized pathway. $\alpha = 2000$ if temperature anomaly is greater than 2°, otherwise 200.

$$\mathcal{R}_2 = 100 * Fe_{curr} + \alpha * (2 - T_{anomaly}) + 10^{-12} * \mathcal{Y} + 10^{-11} * \mathcal{Y}_{\rceil\wr\int\sqcup}$$

- **Experiment 3:** Testing the effects of delayed aerosol injection. In this experiment, we assess the effect of various start dates of deciding to implement stratospheric aerosol injection. This is in line with casting light on the Knowledge vs Urgency trade off that exists with a high risk high reward climate response such as aerosol injection. The previous 2 runs allowed aerosol injection to begin immediately in 2023, which was unrealistic since as mentioned above there are no immediate plans to do so. Here we will delay this to both 2035 and 2050. Before the given year, the model is only allowed to take emissions action, as if following a non-SRM climate response. Any nonzero sulfur actions will be discarded and forcing will be calculated as in the pessimistic RCMIP scenario. After the start date is unlocked, all actions are utilized just like in Experiments 1 and 2. This will simulate society determining geoengineering to be too risky and not taking action until truly dire conditions are faced and carbon management alone is deemed insufficient. After the start date, the emission and injection behaviors will be governed by the same RL model that was trained in Experiment 2. This is because we seek to determine the effects of a delayed deployment but keep the aggression of the deployment similar. We wish to forestall the model learning a more aggressive policy (anyways, the learned policy is already aggressive enough) in response to having less time to act. The same carbon cost GDP reward is used as before.

4 Results and Discussion

4.1 Experiment 1

Despite the simplicity of the reward function, initially the model was unable to learn the simple policy of shamelessly reducing emissions and emitting sulfur. There was much model tuning necessary to get better performance. In addition to experimentation with hyper parameters, it ended up being very critical to normalize the action space and observation space. While initially the spaces were scaled according to units of the value they represented, the resulting irregularity resulted in learning bad policies. Thus we normalized both spaces to always be between -1 and 1. Training and model hyper parameters are shown in Table 5. Attached are model outputs indicating the temperature anomaly, carbon and sulfur emissions, atmospheric CO2 concentrations, effective radiative forcing, GDP cost, and total reward over time.

As we can see, the model did a good job learning an aggressive de-carbonization and injection policy. In the absence of any intervention (the default RCMIP scenario), temperature anomaly reaches as high as 3.24℃ and radiative forcing is 5.40 watts / m^2. Without any economic feasibility constraints or secondary climate effects factored in, the model always emits an additional 100 MtS into the atmosphere (the maximum allowed value). This means that as far as temperature is concerned, FaIR sees no negative side effect to emitting as much sulfur as it is allowed in order to decrease forcing as much as possible. In terms of carbon emissions, the model reaches net zero in 2053 and carries on emitting virtually zero (and actually negative) emissions until the end of the century from 2053 onwards. This allows CO2 concentrations to decline from 423.8 ppm to 381.2 ppm from 2023 to 2100. Doing so contains warming to 1.27℃ and net forcing to 1.93 watts / m^2. As noted early, this seems to be a fairly high net forcing for a staggering amount of deliberate anthropomorphic sulfur emission. 100 megatons is after all considerably more than total natural global emissions of sulfur, which peak at 80 megatons in the RCMIP scenario. This again reinforces the previously stated notion that the species configurations of sulfur embedded in FaIR may be underestimating its negative forcing power. If Mount Pinatubo's 20 megatons were sufficient to cool the planet by 0.5° in a single instance, then it seems odd that 100 additional megatons every year for 80 years has this magnitude of effect.

Regardless, the model has learned exactly what it was instructed to: an impossible yet effective reduction of global temperatures. Global emissions are cut drastically basically immediately in 2024 and a fleet of over 3000 payload delivering aircraft are built in a flash. This results in a 14 trillion dollar spike in economic cost as soon as the model takes over (which flattens down to operational maintenance right afterwards). In every single year except 2024, GDP cost is not more than 2% of GDP. While this scenario is obviously unrealistic it highlights the ability of Reinforcement Learning to learn sensible pathways shaped by provided reward criterion. The reward itself is quite unstable due to the reward function's sensitivity to small temperature change (remember, each degree of difference from the desired 1.5℃ changes the reward by 100) that mirror decade-timescale natural climate variations. The fact that the reward is consistently positive despite the stringent penalties is proof enough that the model has learned.

(5)

Parameter	**Value**
Maximum Annual Carbon Decrease (GtC)	40
Maximum Annual Sulfur Injection (MtS)	100
Time Range	2023-2104
RL Algorithm	Advantage Actor Critic
Learning Rate (α)	0.021
Discount Factor (γ)	0.8
Training iterations	2000 (12 minute runtime)
n_steps, n_envs	4, 1

4.2 Experiment 2

Immediately we note that the model was not actually able to learn a policy that earned a good reward. Many model and parameter variants were tested, including an hour long training session to see if convergence would occur. More specifically, learning rates of 0.00021, 0.021, 0.024, 0.026, and 0.21; discount factors of 0.8, 0.85, 0.9; and 100, 600, 1000, 2000, 4,000 iterations were all tested. Normalizing the action and observation space did not solve these issues as they did for Experiment 1 although they did solve similar inaction issues. There remained an issue of continuity within the observation space, meaning similar observations could result in wildly different action choices. Reinforcement Learning is notoriously finicky so it is probable that the right method simply was not tested. Further effort may reveal some improved choice of model, correct set of hyperparameters, or longer training/ larger environment interaction would result in better performance. Be that as it may, there are still a slew of interesting behaviors to note. All the same plots shown in Experiment 1 are included again. We also plot global GDP over time.

Firstly, just as before there was essentially no penalty for injecting as much sulfur into the environment as possible. In these runs, the maximum allowed annual sulfur emission was 60 megatons; the model emitted that maximum in every single timestep. This is sensible for the same reasons it was before. Due to its negative forcing effect, vast quantities of sulfur help keep temperatures low which is still weighted quite highly in the reward. Feasibility was assessed on the basis of decarbonization and thus did not dissuade excessive usage; while it is true that current airplanes lack the payload capacity, much of the appeal of SAI is many of the implementation puzzle pieces are loosely in place. Combined with the fact that its direct economic implementation costs, though easily modelled, are a fraction of the real damage stratospheric aerosol injection may cause, we again see that FaIR is limited in its ability to assess the non-temporal merits of SAI. However, the fact that the model tuned itself to the advantages of unrestricted injection is a good performance indicator.

The model did not seem responsive to the feasibility constraint, which is primarily where the declining reward stemmed from. The model went net zero on carbon emissions at 2053 but rather than a gradual decline over time the model made 2 sharp decreases that forecasted a 92% drop in global emissions over the next two years. Rather than balancing feasibility with temperature control, the model strongly erred on the side of temperature control. At the end of the run, temperature anomaly was 1.61℃ above pre-industrial levels and hovered around 1.5° throughout. Net radiative forcing was 2.25 watts / m^2 at the end of the century. At the very least this is more realistic than the model in Experiment 1. However the steepness of the immediate decline in carbon emissions resulted in a sharp decline in reward. After 2023, the model averages 10 GtC/year less than the SSP 245 scenario and 19.4 GtC/year less than the SSP 375 scenarios. Since this discrepancy is already being punished by the reward, it is likely more a matter of training tuning to capture the right emissions policy to supplement the sensible aerosol behavior. All in all, while this model is perhaps not as distinct from Experiment 1 as one may desire, it still makes intelligible decisions. Future work should focus on deriving more suitable penalties for SAI as well as more robust or better enforced feasibility constraints against a dramatic decline in emissions.

(6)

Parameter	**Value**
Maximum Annual Carbon Decrease (GtC)	30
Maximum Annual Sulfur Injection (MtS)	60
Time Range	2023-2105
RL Algorithm	Proximal Policy Optimization
Learning Rate (α)	0.026
Discount Factor (γ)	0.85
Training iterations	3000 (18 minute runtime)
n_steps, n_envs	4, 1

4.3 Experiment 3

(7)

World State 2150	2025	2035	2050	2065
Temperature Anomaly	2.034	2.061	2.305	2.287
Climate Forcing (W/m^2)	3.251	3.254	3.789	3.745
CO2 Concentrations (ppm)	461.578	462.125	466.594	462.923
Sulfur Emissions	114.619	114.619	54.619	54.619
CO2 Emissions	11.69	11.69	13.69	11.69
World State 2100	**2025**	**2035**	**2050**	**2065**
Temperature Anomaly	1.987℃	1.994℃	2.038℃	2.016℃
Climate Forcing (W/m^2)	2.834	2.837	2.897	2.846
CO2 Concentrations (ppm)	423.979	424.225	429.005	425.052
Sulfur Emissions	90.24	90.24	90.24	90.24
CO2 Emissions	0	0	0	0

To limit the influence of carbon emission (which was dominating climate impact), we set the max reduction in global CO2 emissions in a given year to be 2 gigatons. This helped prevent infeasibly abrupt decarbonization scenarios and made more dependent on the sulfur. However, the experiment proved quite straightforward due to the simplistic policy the agent had learned previously. The agent reduced emissions as much as possible and emitted as much sulfur as it was allowed in every timestep. This standardized behavior was helpful because it determined whether any effects were due to timing of SAI alone. Results show deep symmetries in all the plots. Although there is a substantial distinctions between timesteps within scenarios where geoengineering had been deployed versus not been deployed, once sulfur emissions ratched up at the assigned time, the scenarios quickly converged to similar endpoints. By 2100, all are essentially identical. This implies that once SAI is implemented, its effects are quick to take place. It seems that carbon emission is a much more important long term driver; although forcing decreases very quickly after aerosols are released, even when deployment is delayed until 2065 forcing remains somewhat under control due to the learned carbon policy. Due to the outsized effect of the carbon reduction, it may be interesting to decouple the emissions and sulfur models and see what differences play out if random emission changes are made rather than a steady move towards net zero. That may isolate the independent effect of various aerosol injections.

sulfur aerosol species are indeed a potentially useful tool for temporarily reducing temperatures. However, given the risks and immediacy of the forcing effect, it seems that there is not a large detriment to waiting and gathering more information before any rash injection of sulfur into the atmosphere. To the extent that it is trustworthy, the data from this experiment indicates that there is little urgency to rush through something as drastic as aerosol injection.

References

Dunne, D. (2018). Explainer: Six ideas to limit global warming with solar geoengineering. *None (web article)*.

Launeau, V. (2022). Climate mitigation with reinforcement learning. *None*.

Leach, N. J., Jenkins, S., Nicholls, Z., Smith, C. J., Lynch, J., Cain, M., Walsh, T., Wu, B., Tsutsui, J., and Allen, M. R. (2021). Fairv2. 0.0: a generalized impulse response model for climate uncertainty and future scenario exploration. *Geoscientific Model Development*, 14(5):3007–3036.

MacMartin, D. G., Ricke, K. L., and Keith, D. W. (2018). Solar geoengineering as part of an overall strategy for meeting the 1.5 c paris target. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2119):20160454.

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D. C., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., et al. (2018). Global warming of 1.5° c: Summary for policy makers. *Global Warming of 1.5℃. An IPCC Special Report.*

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Silva, L. A. D., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, 42:153–168.

Smith, W. (2020). The cost of stratospheric aerosol injection through 2100. *Environmental Research Letters*, 15(11):114004.

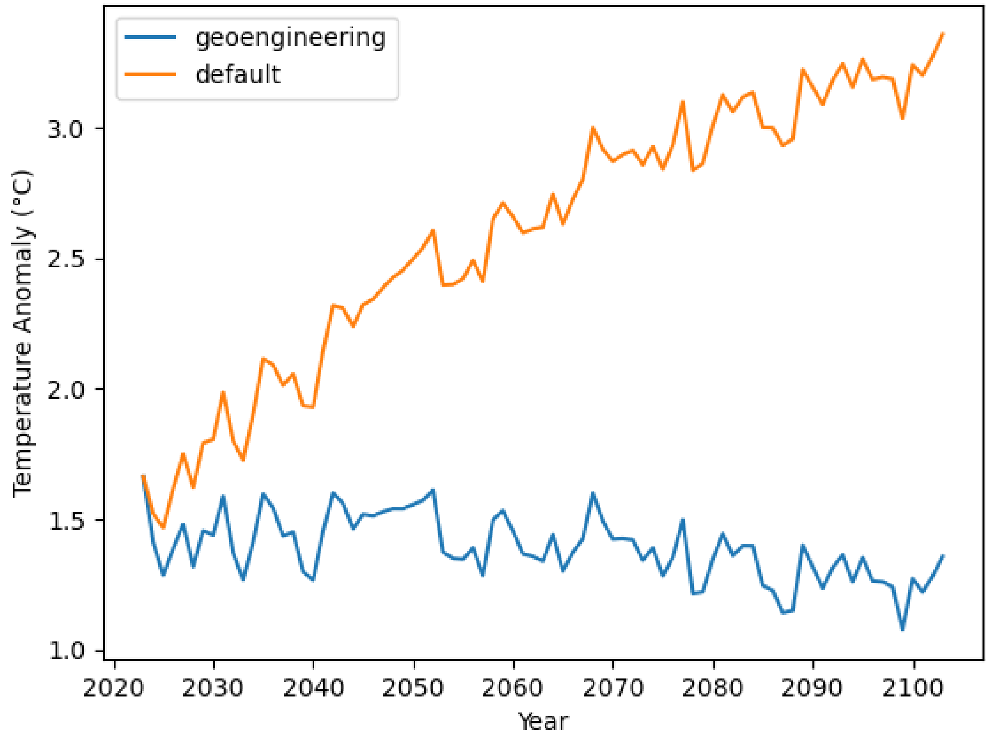
Wigley, T. M. L. (2006). A combined mitigation/geoengineering approach to climate stabilization. *Science*, 314(5798):452–454.

Zarnetske, P. L., Gurevitch, J., Franklin, J., Groffman, P. M., Harrison, C. S., Hellmann, J. J., Hoffman, F. M., Kothari, S., Robock, A., Tilmes, S., Visioni, D., Wu, J., Xia, L., and Yang, C.-E. (2021). Potential ecological impacts of climate intervention by reflecting sunlight to cool earth. *Proceedings of the National Academy of Sciences*, 118(15):e1921854118.

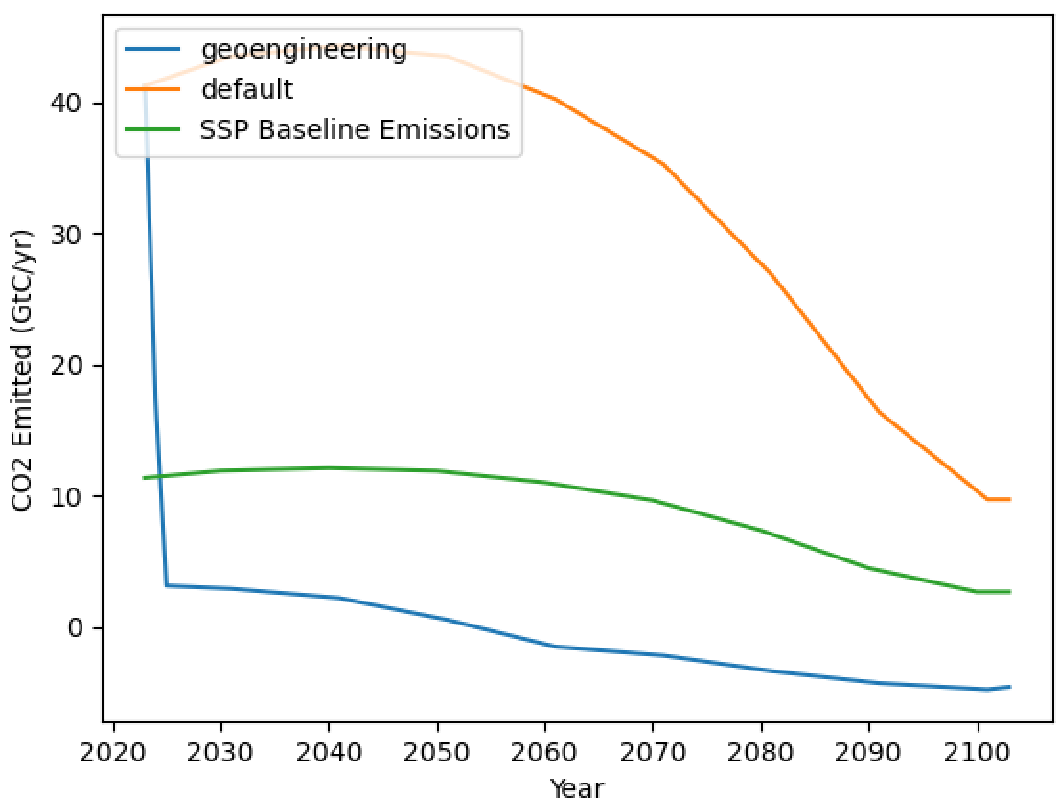
Figures (Overleaf was not allowing figures to be uploaded despite much finagling.)

Experiment 1

Temperature Anomaly over Time

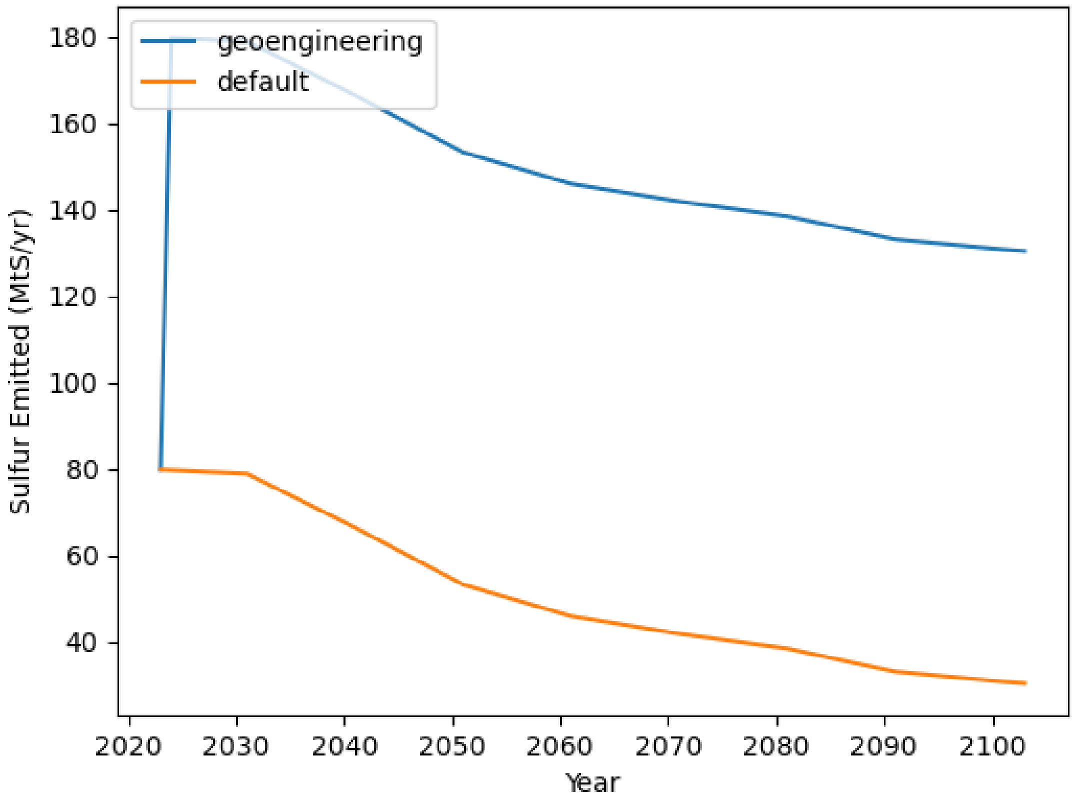


Gigatons CO2 Emitted, SSP2 Reference

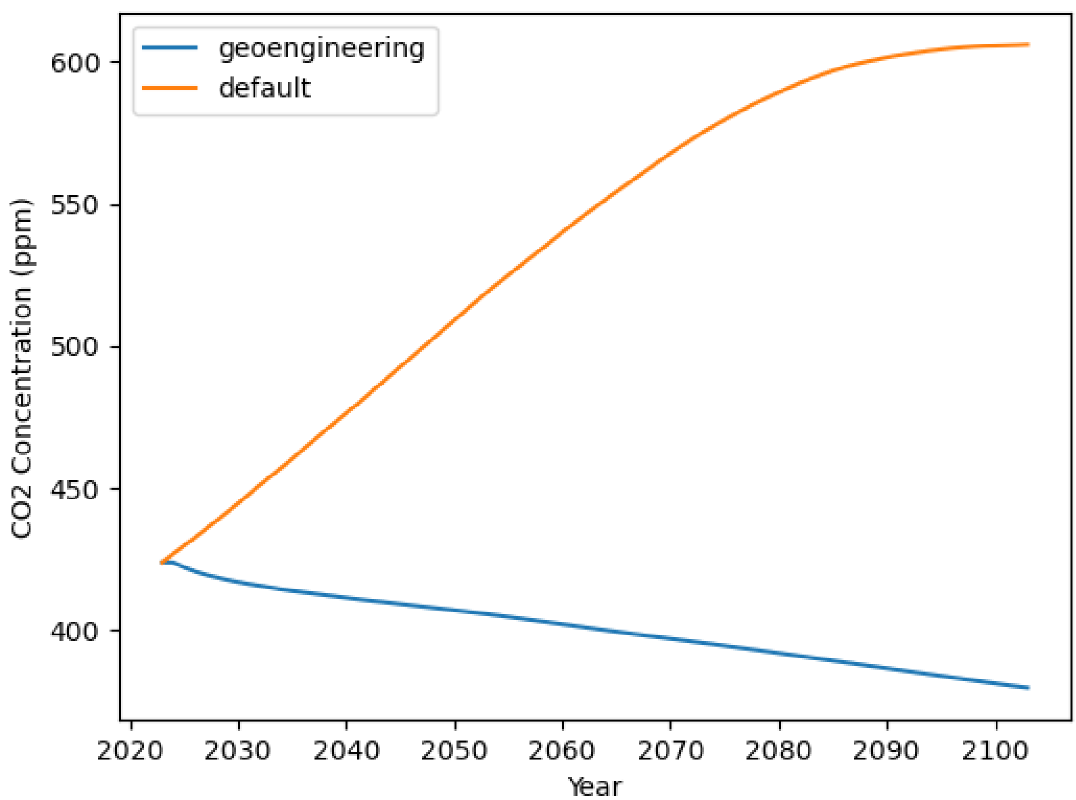


Model emissions, plotted against RCMIP baseline (orange) as well as SSP2 (green) reference emission pathway

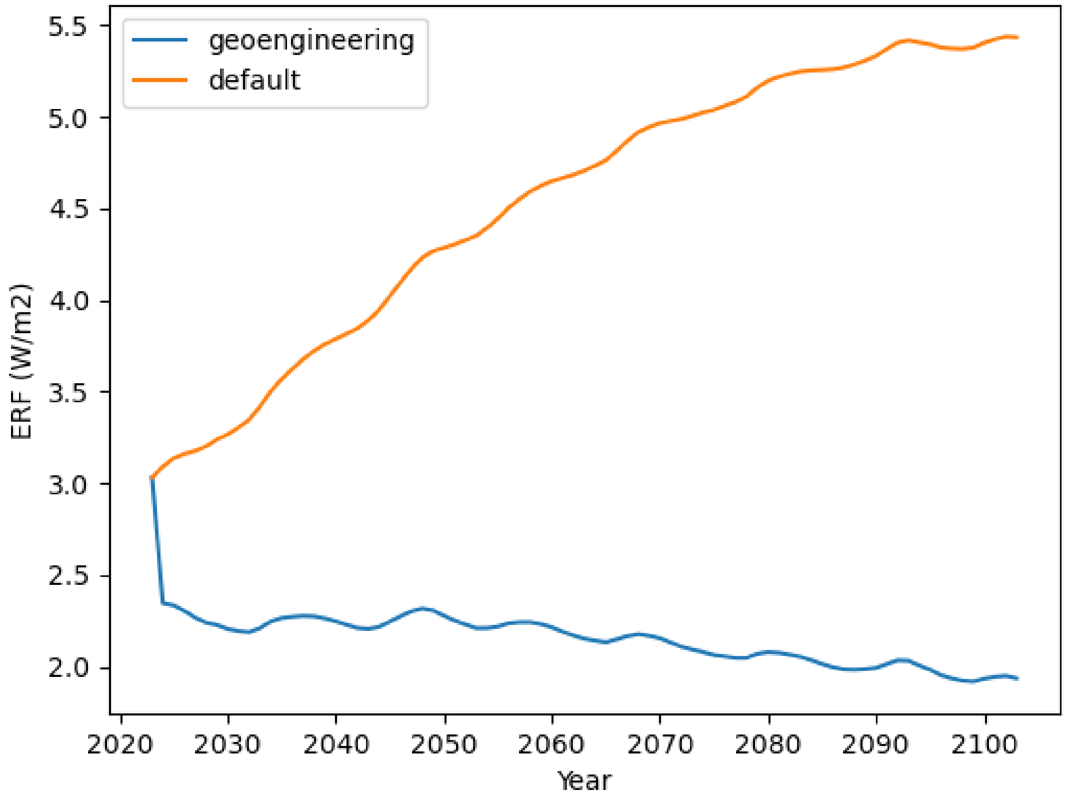
Megatons Sulfur Emitted



Atmospheric Carbon Dioxide Concentrations (ppm)

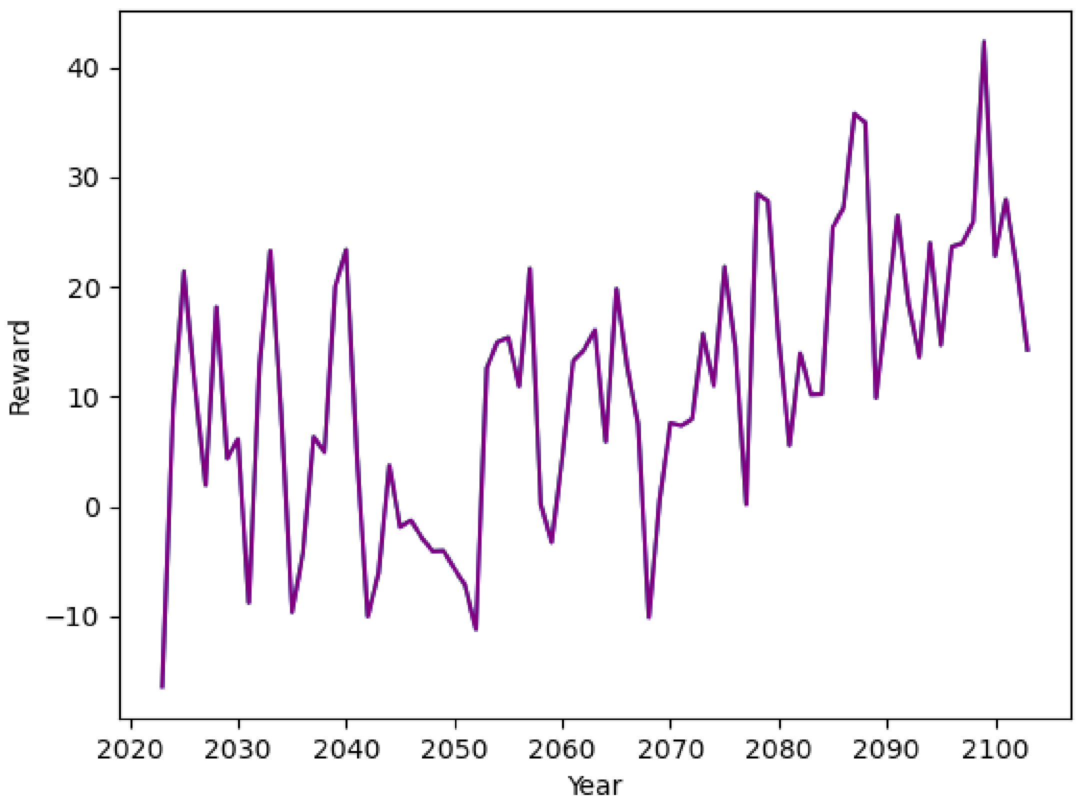


Effective Radiative Forcing, in Watts per Square Meter



Economic Cost due to Climate Mitigation Actions

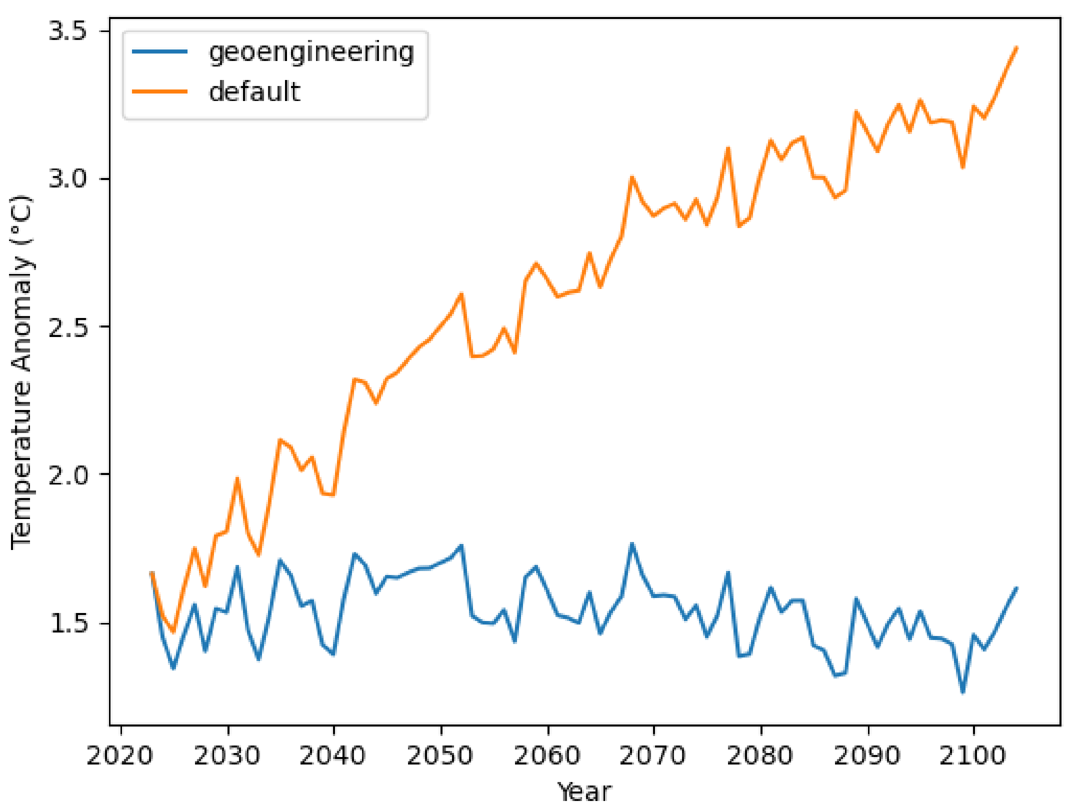
Total Reward



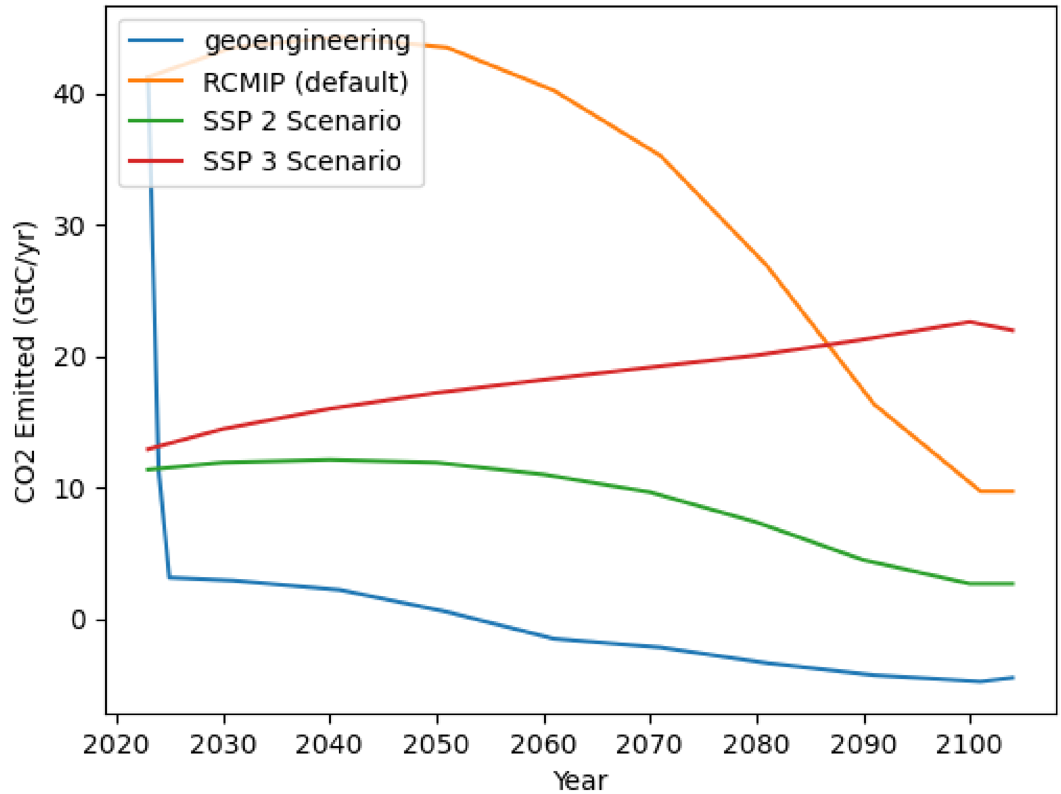
Even though the reward seems to be fluctuating wildly, that is reasonable given that the reward amplifies slight temperature fluctuations by a magnitude of 100. Since the reward is also only positive if the 1.5°C warming target is maintained, the overall positive character of the reward is a good sign of stable climate control.

Experiment 2

Temperature Anomaly over Time

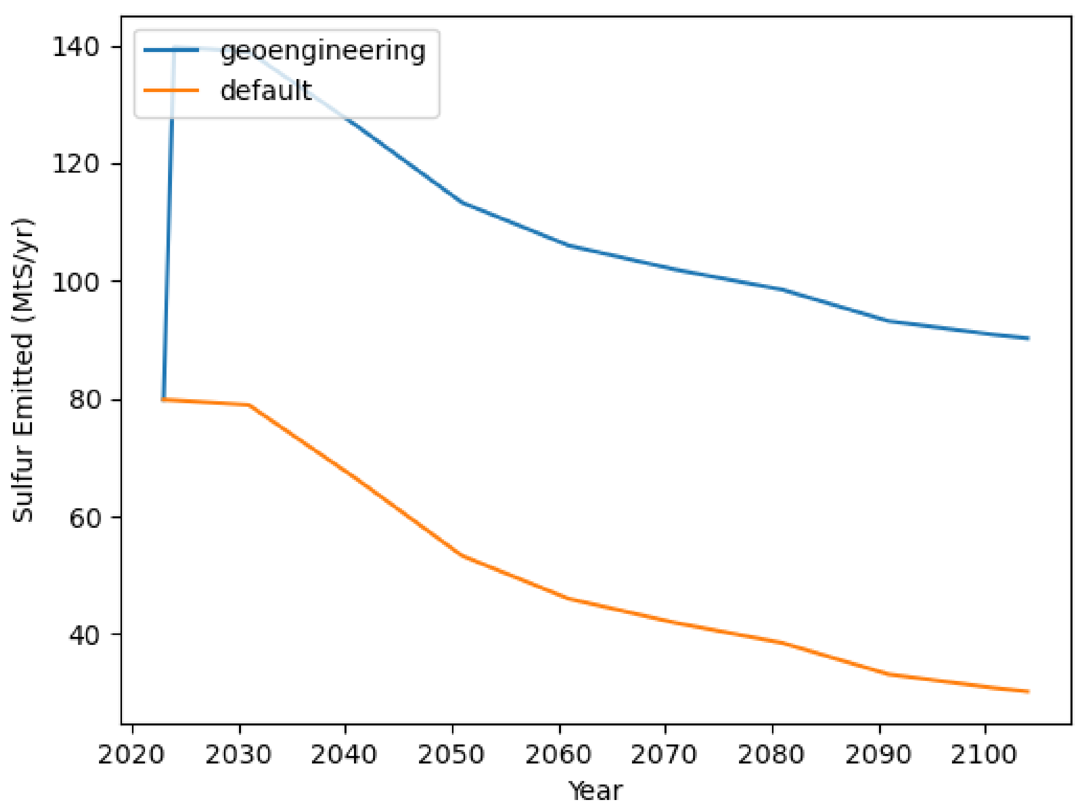


Gigatons CO2 Emitted, SSP2 Reference

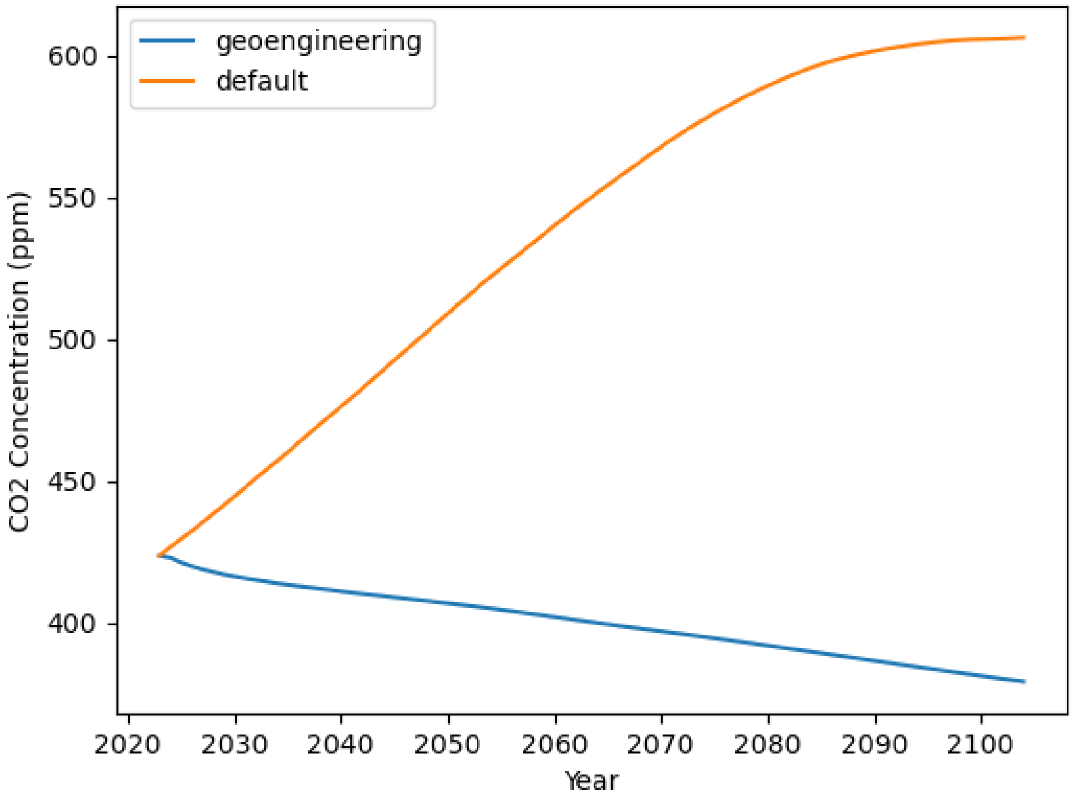


Model emissions, RCMIP baseline (default), SSP2 (green) and SSP3 (red) reference emission pathways

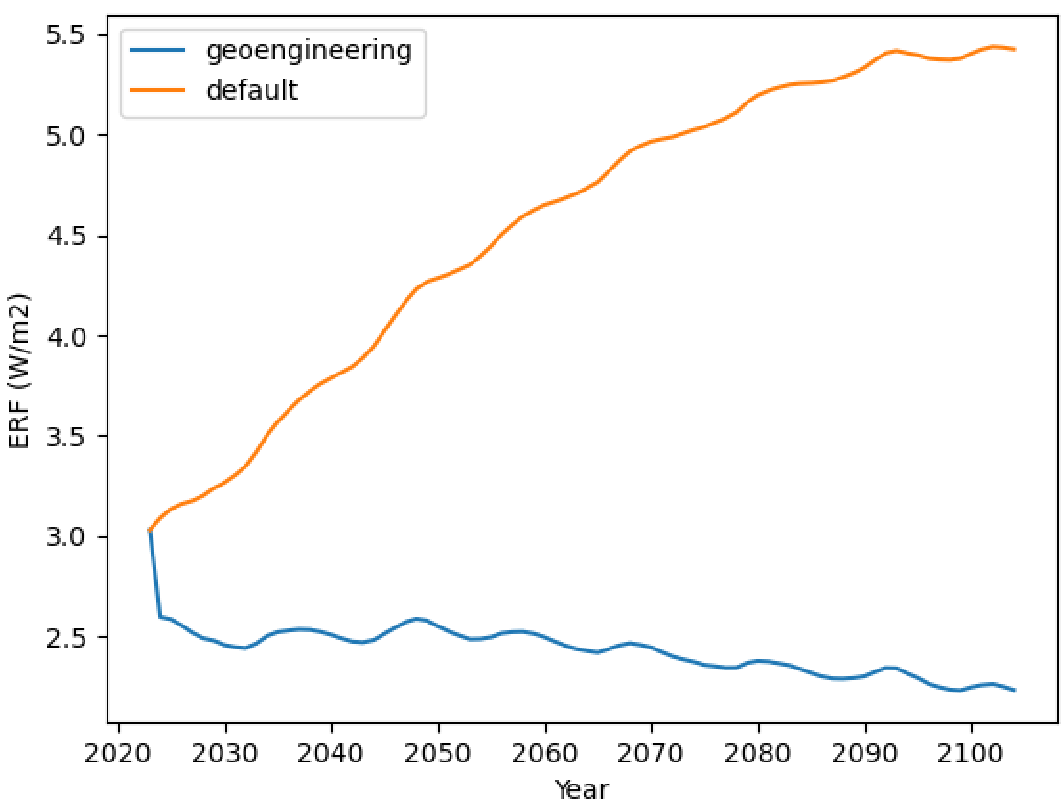
Megatons Sulfur Emitted



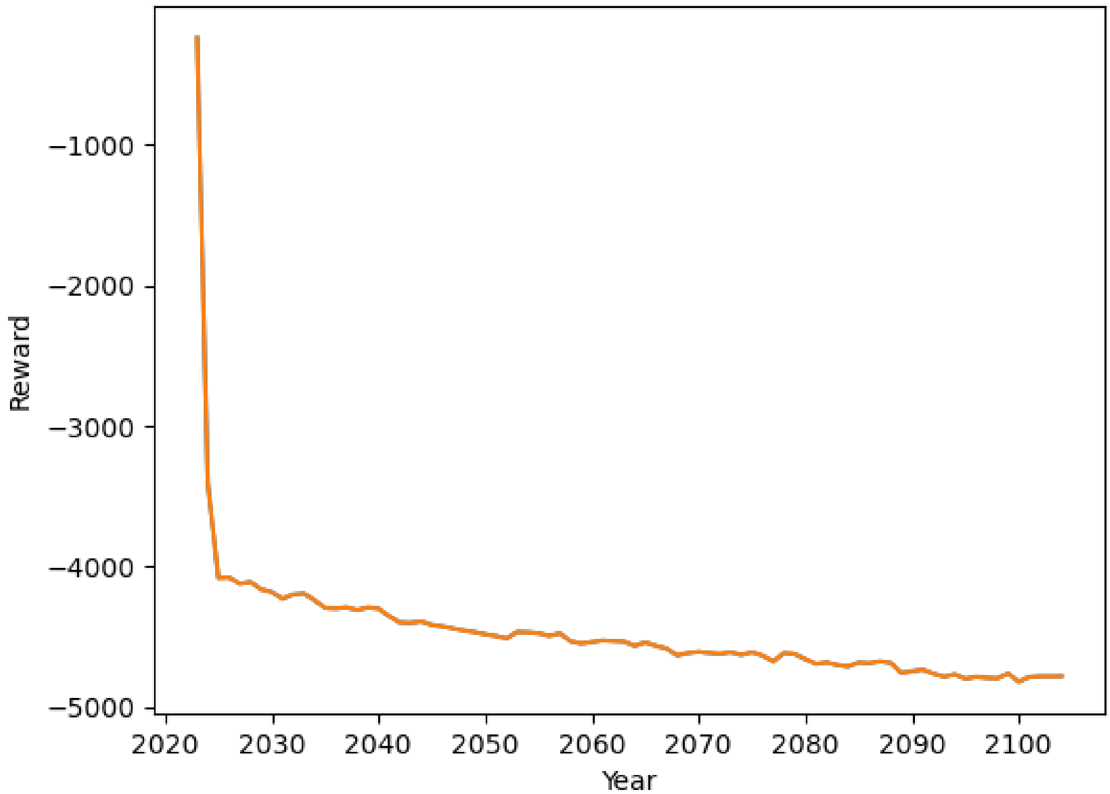
Atmospheric Carbon Dioxide Concentrations (ppm)

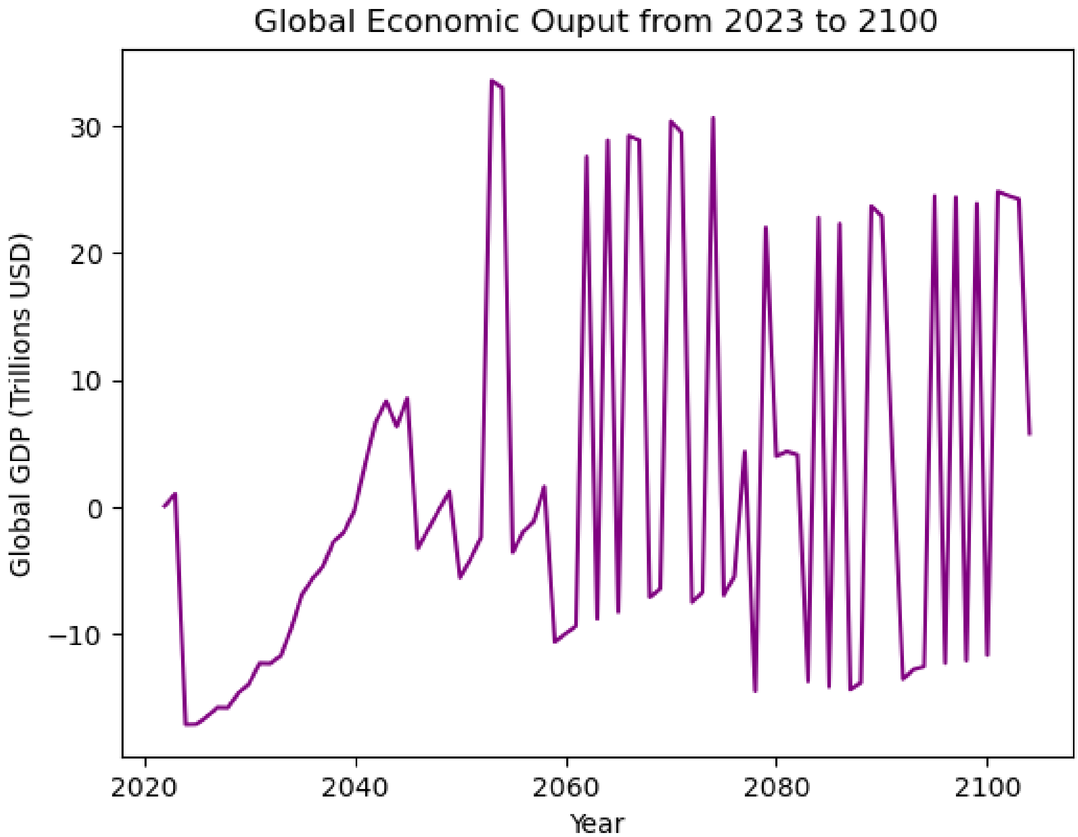


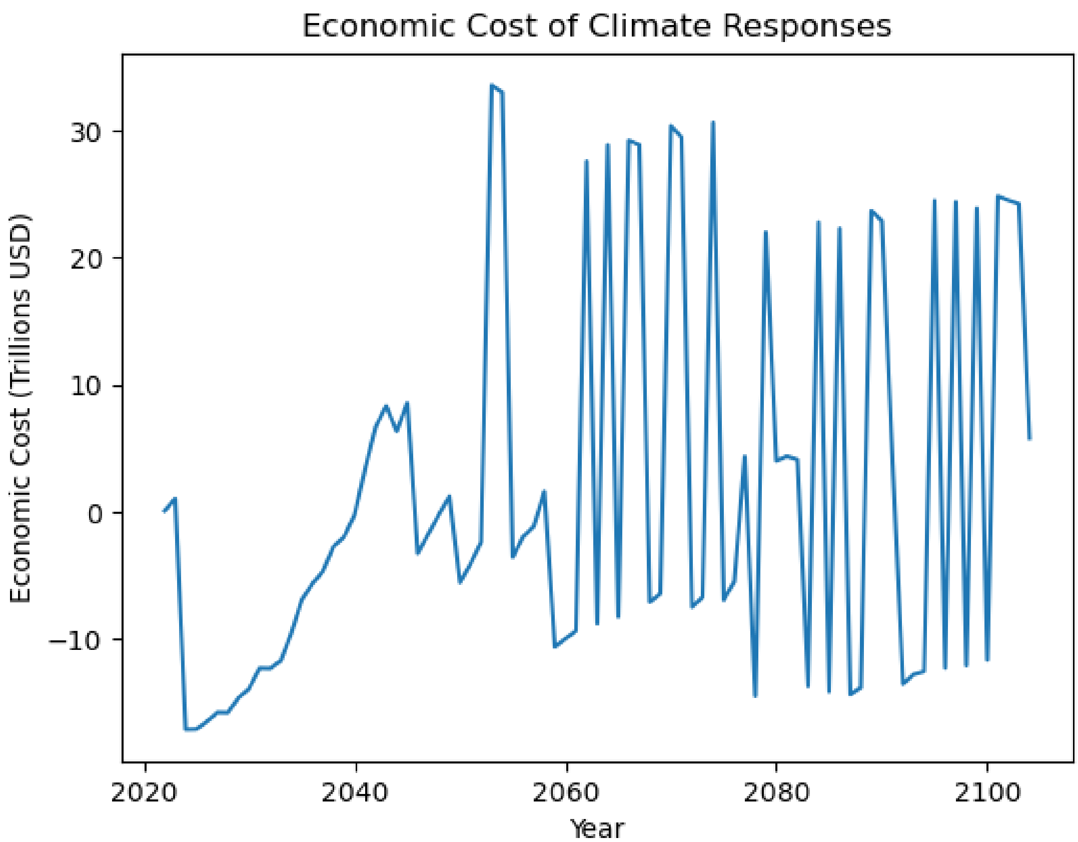
Effective Radiative Forcing, in Watts per Square Meter



Total Reward

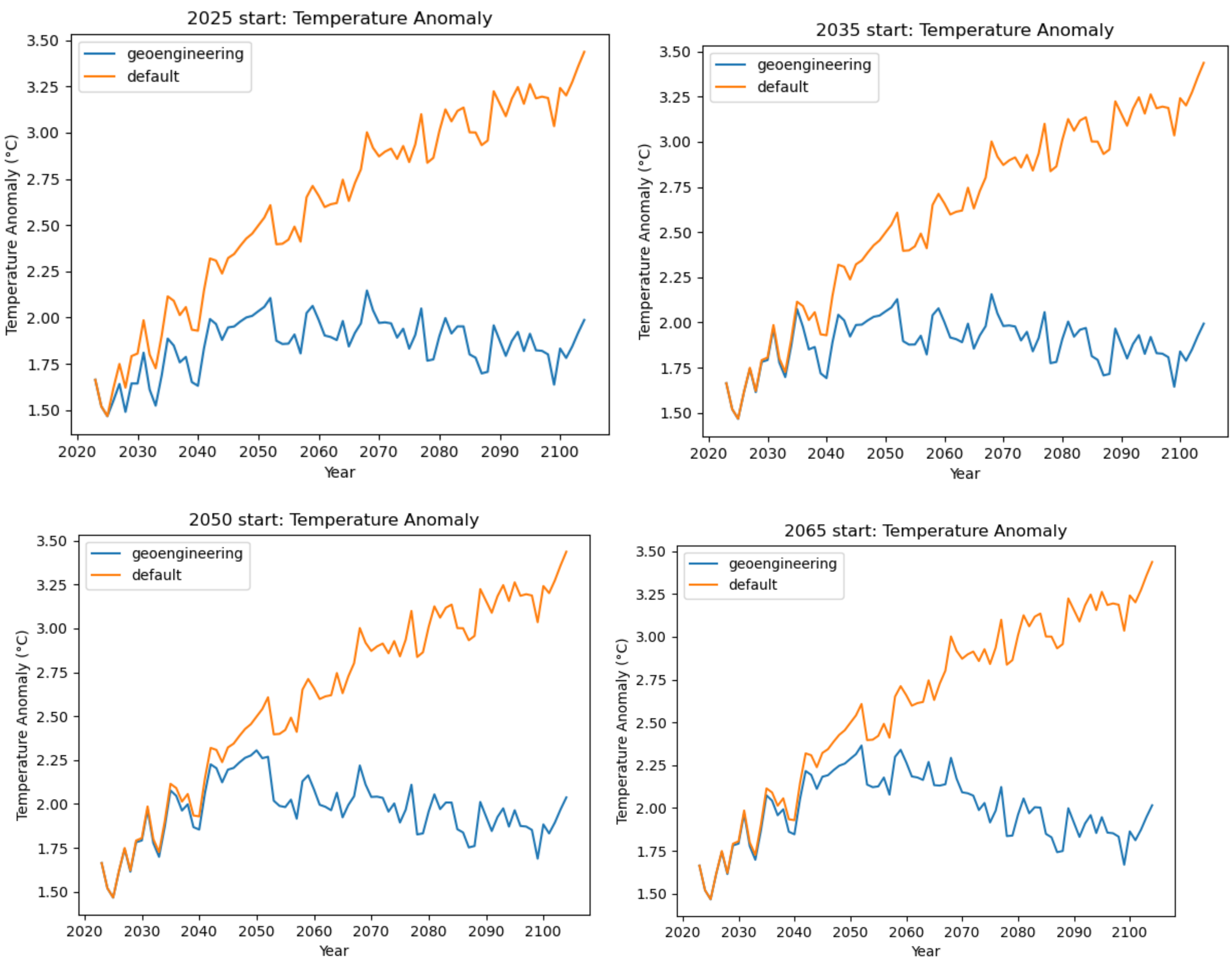



Global Economic Ouput from 2023 to 2100
Global GDP (Trillions USD)
30
20
10
0
−10
2020
2040
2060
2080
2100
Year

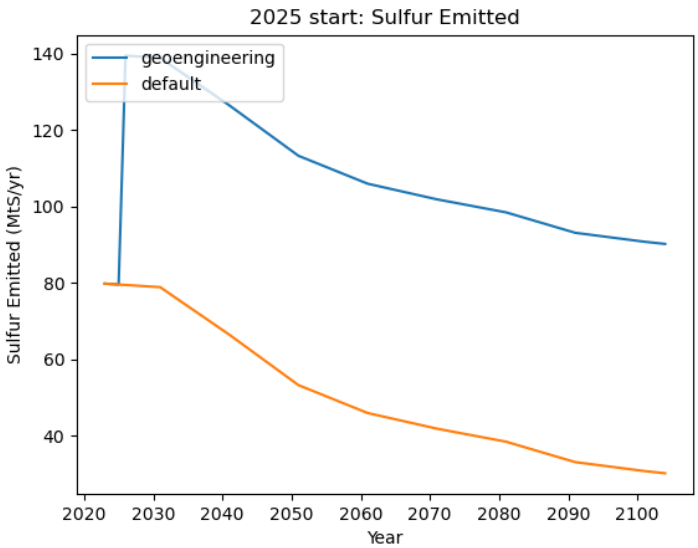

Economic Cost of Climate Responses
Economic Cost (Trillions USD)
30
20
10
0
−10
2020
2040
2060
2080
2100
Year

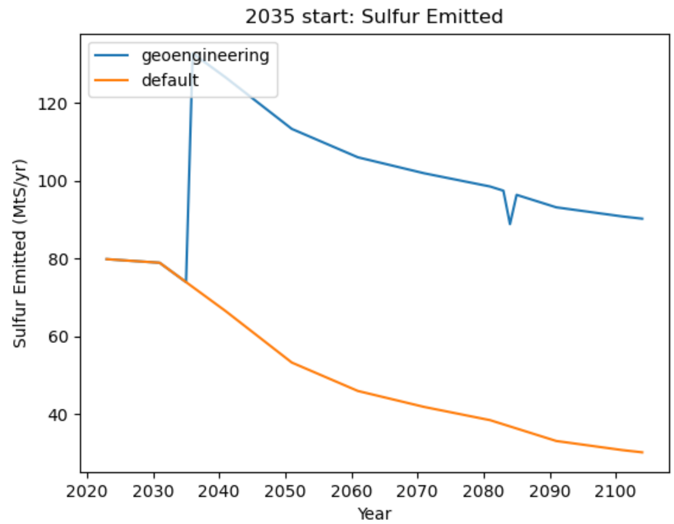
Experiment 3
Comparing Different Start Dates

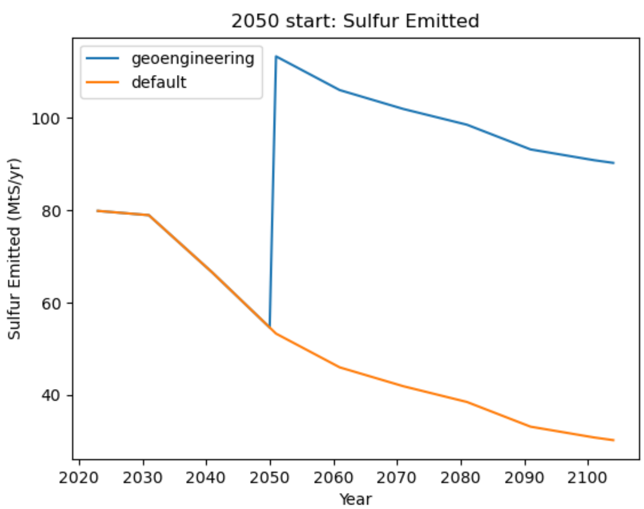
Temperature Anomaly (°C):

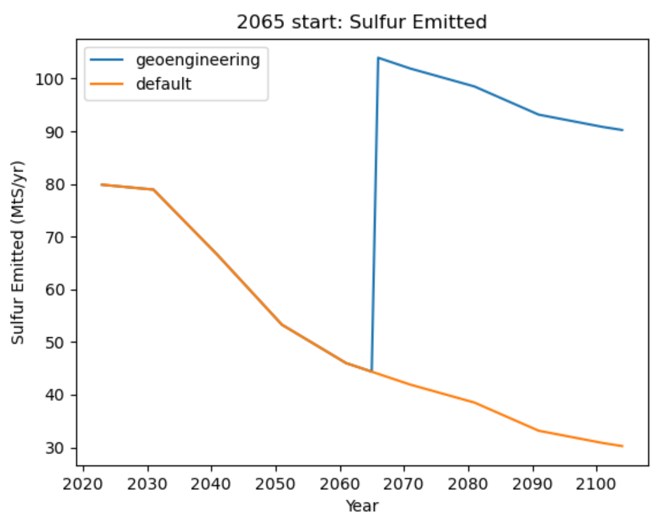


Sulfur Emissions (MtS)

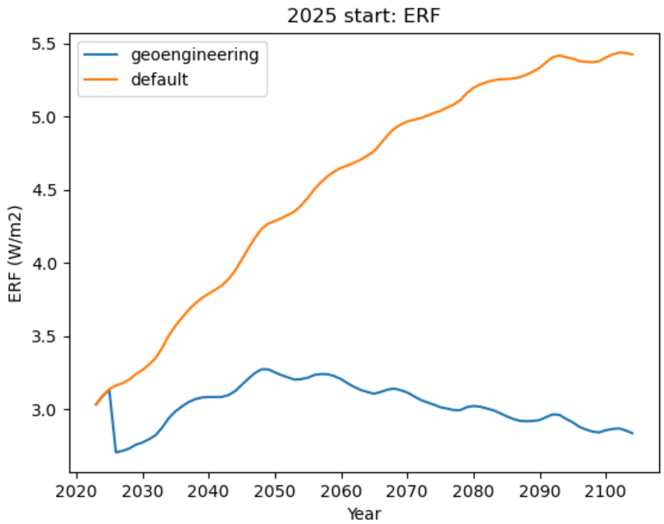


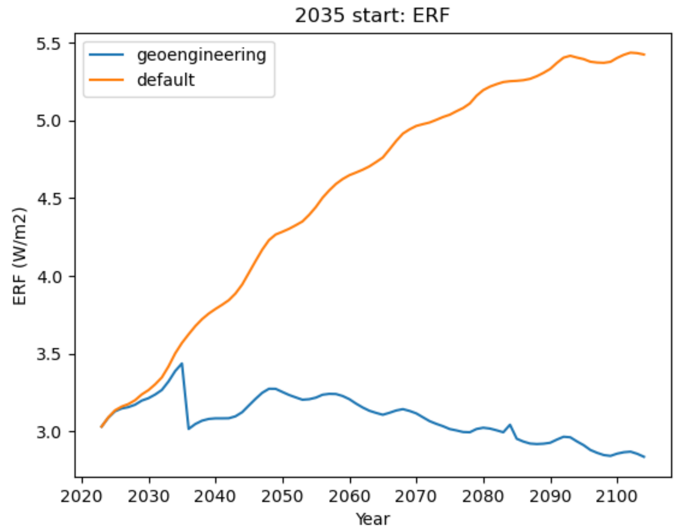


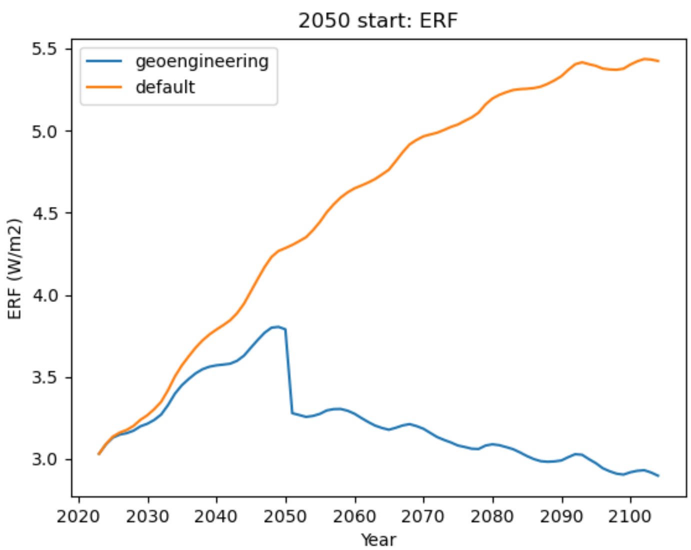


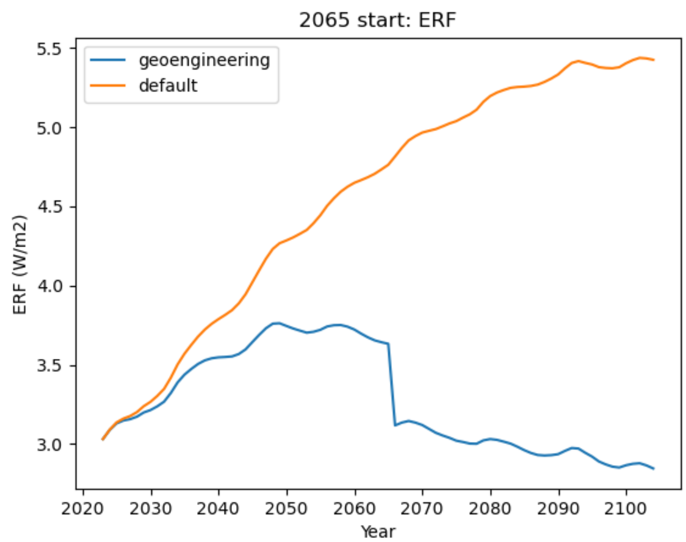


Effective Radiative Forcing

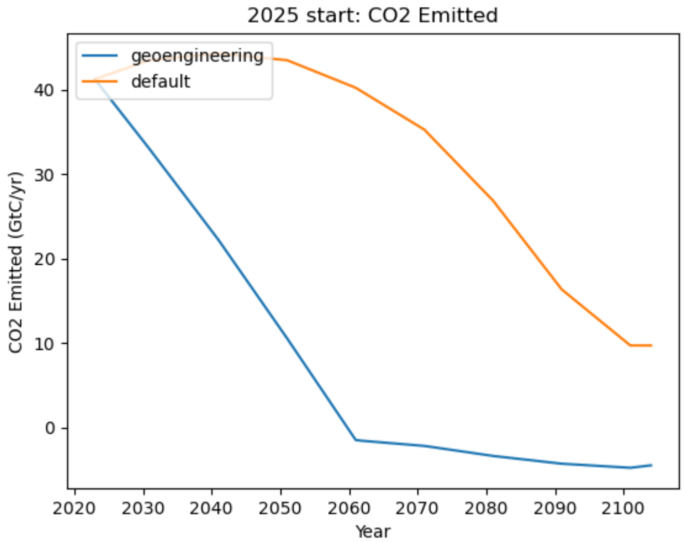


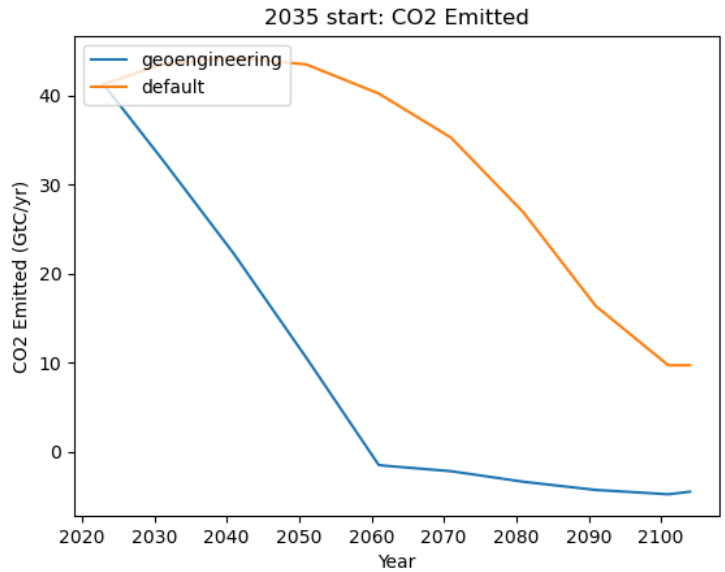


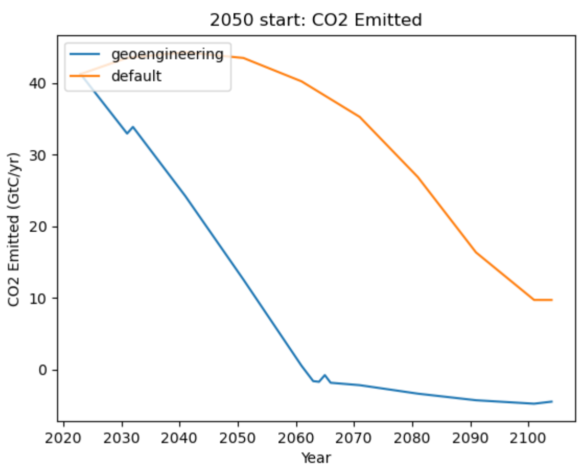


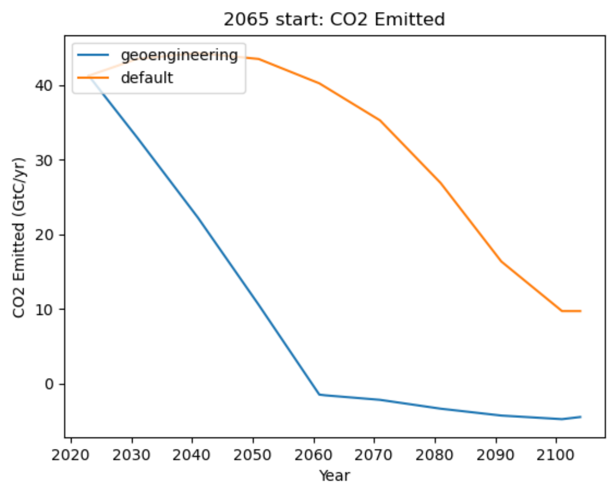


Carbon Emissions (Gt C)

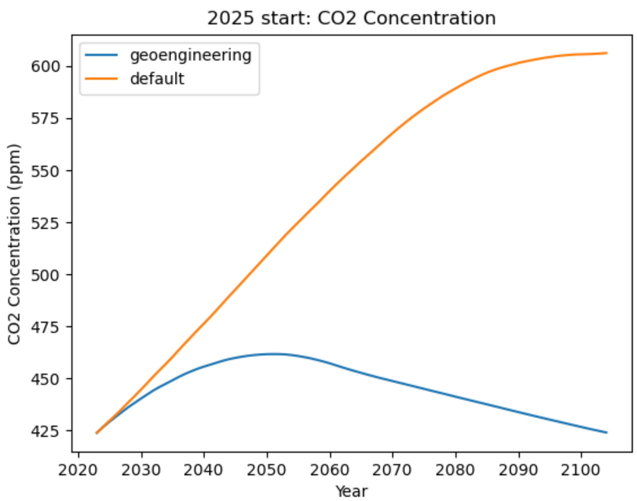


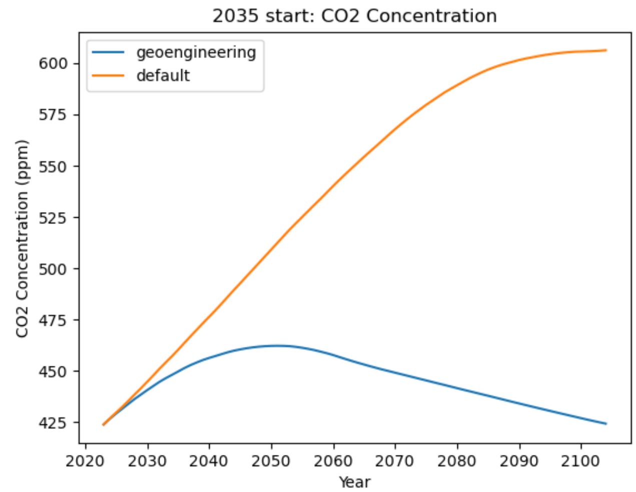


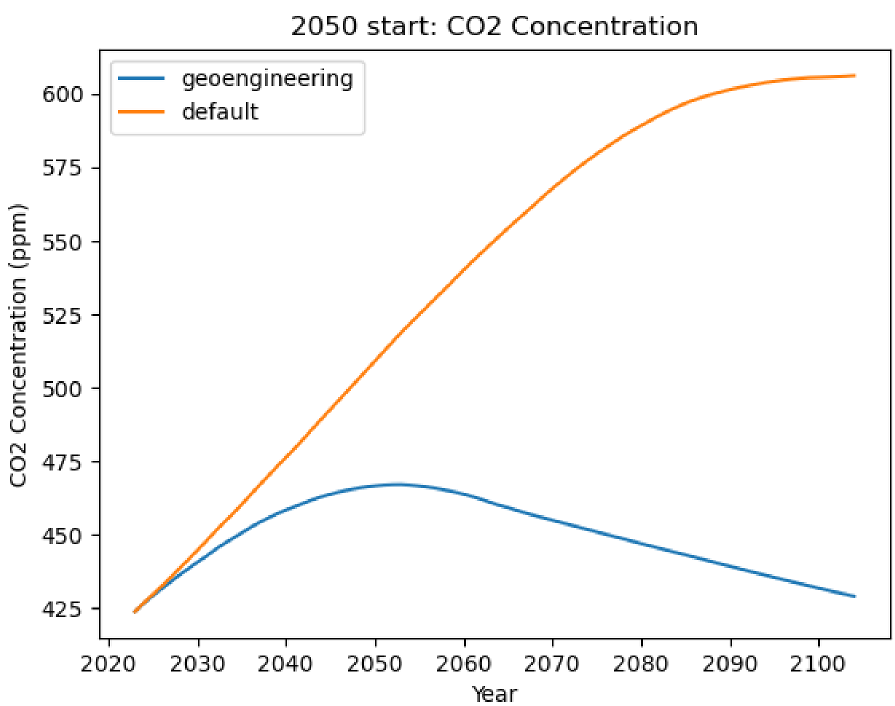


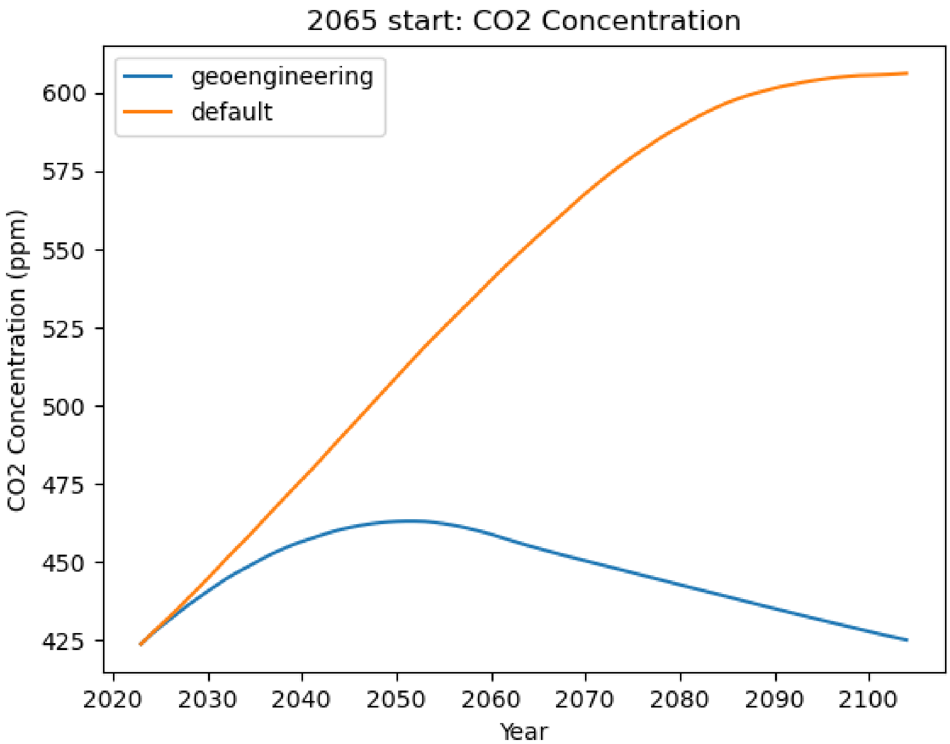


CO2 Concentrations








2065 start: CO2 Concentration
geoengineering
default
CO2 Concentration (ppm)
600
575
550
525
500
475
450
425
2020
2030
2040
2050
2060
2070
2080
2090
2100
Year