Post 7: Why focus so much on global mean temperature?

Is it meaningful to look at global mean temperature?
Yes, it is. Global mean temperature is a useful and convenient indicator in the context of an inhomogeneous and imperfect observational network, and provides an intuitive measure of of the overall state of the climate even in the face of complex interactions among various sub-systems.

What are the advantages to using temperature as an indicator of the state and sensitivity of the climate?
Temperature is relatively easy to measure with long and relatively homogeneous records, and is a basic variable with fundamental relationships to energy, precipitation, and other climatic variables.

Is the global mean temperature an appropriate variable for assessing climate sensitivity to any climate forcing?
Global mean temperature is more appropriate for assessing the response to well-mixed greenhouse gases (which have a relatively uniform spatial response) than the response to, e.g., volcanic eruptions o-r anthropogenic aerosols.

Could including ocean heat content in estimates of global mean temperature help to reduce the decoupling between temperature and the global mean energy budget?
For some purposes, maybe. Scientists have recently plotted time series of OHC, but the observational network for OHC is less extensive in space and less reliable (particularly at earlier times). Estimates of OHC are mainly based on sparse observing networks o-r reanalysis systems that assimilate those sparse observations into a global ocean model.

About discriminant analysis and "improvements" on the global mean temperature
Refer to the blog post, Schneider and Held (2001), and the comments below the blog — or submit your own comment to the blog!

The topography will also influence the noise except the latitude, right?
I'm not sure I understand the question. Topography certainly affects the distribution of surface temperature; however, this is implicitly accounted for in the calculation of the global mean (and in the discriminant analysis introduced in the post). The topography changes only very slowly (we can consider it constant at these time scales), and we typically focus on the global mean temperature anomaly (i.e., the difference relative to the long term mean). Topography therefore has little impact on the global mean temperature time series.

If we use a weighted average method to calculate the global mean, will it be meaningful? How should the coefficient be defined?
We generally use an area weighted mean to calculate the global mean. This is particularly important for regular latitude-longitude grids (e.g., 1 degree by 1 degree), because the area of a grid cell in this type of grid is greater near the equator than near the pole. The weighting coefficient is the cosine of latitude, which goes from 1 at the equator to 0 at the poles.

How does the 4th power in the Stefan-Boltzmann law decouple the global mean temperature and TOA flux?
Prof. Held's argument is that this effect is minor, but I think the rationale is that the 4th power means that the relationship between temperature and the TOA flux is nonlinear even locally. Temperature (and temperature changes) are not spatially uniform, which has the potential to magnify the nonlinearity in the global mean. This is also the physical mechanism behind the Planck feedback, which is a negative feedback associated with the outgoing energy flux becoming more efficient as climate warms.

Why did the author choose 1890–1909 mean as the reference period, rather than a more recent period with better observations?
My guess is that the authors had two reasons for choosing 1890–1909 as the reference, both of which are related to emphasizing the signal they are most interested in (global warming due to increasing CO2). First, this is probably as close as they can get to "pre-industrial" (i.e., before CO2 started to increase). Second, this choice emphasizes the warming in the latter part of the record. If they chose a time period toward the end, or in the middle of the record, then the warming signal (the orange and red colors at the end of the movie) would be less obvious.

What causes polar amplification of the warming signal?
Polar amplification appears to arise from multiple factors, and is not yet fully understood. One likely contributor is the ice albedo feedback, in which warming reduces ice and snow, which reduces the reflection (and increases the absorption) of solar radiation in polar regions. Other possible contributors include changes in the ocean circulation that increase poleward energy transport o-r changes in the distribution of clouds and water vapor that increase the local greenhouse effect, but these are less well understood. The greenhouse effect of CO2 is logarithmic, so that adding the same amount of CO2 has less impact on the greenhouse effect if the background concentration is higher. This is why scientists often consider CO2 increases in terms of doublings (2x, 4x, etc.), rather than absolute increases (e.g., 100 ppmv). It is hard to say whether there is a limit to the temperature increase as CO2 continues to increase.

Can the global mean properly account for different types of noise, or even introduce noise?
The practice of differentiating noise from signal is difficult and differs considerably depending on the goal. In the blog post, noise is all variability at time scales shorter than 10-15 years, and therefore includes everything from individual weather systems to ENSO. In a time series of annual means, noise is everything at time scales shorter than 1 year. Averaging should tend to reduce noise on time scales shorter than the averaging period, though I guess it would emphasize "noise" at the same time scale as the averaging period. This is why it's important to choose the averaging period carefully!

The Stefan-Boltzmann law is formulated for perfect blackbodies, but Earth is not one — why do we still use it?
It is true that the Earth is not a perfect blackbody, but the Stefan-Boltzmann law is still very instructive for understanding the Earth's energy balance and greenhouse effect. This assumption is a common one, and it holds up well in a variety of applications.

Post 9: Summer is warmer than winter

Is the seasonal cycle a response to external forcing or internal variability?
The mean seasonal cycle is a response to the external forcing of variations in the distribution and magnitude of solar radiation (although the seasonal cycle can be and often is influenced by internal variability).

Will the climate forcing associated with increases in CO2 influence the seasonal cycle?
The answer to this question likely varies by location -- the seasonal cycle may change substantially in some places, but not at all in others. Consider, for example, Juneau, Alaska, where the climate (and seasonal cycle) are strongly affected by both nearby glaciers (which have been melting) and the position of the Pacific storm track (which has been moving northward). It is possible that these responses to global warming will significantly change the seasonal cycles of temperature and precipitation. The monsoons are another example -- there is evidence to suggest that the amount of monsoon rainfall could change, but the extent of this change is still unknown.

Will the GHG effect be stronger in summer or winter?
The GHG forcing in high latitudes may be stronger during winter than during summer due to the infrared component of the radiation budget being relatively large during winter (when sunlight is scarce or nonexistent).

Are variations in CO2 an external forcing or internal climate variability?
Anthropogenic increases in CO2 due to fossil fuel burning are considered an external forcing, but CO2 can also change due to natural internal variability.

The seasonal cycle of temperature peaks earlier in the eastern Pacific than in the western Pacific -- why?
This is likely attributable to differences in the surface mixed layer, which is deeper (and therefore takes longer to heat) in the western Pacific.

Why is the departure from the sinusoidal fit to temperature variability in Minneapolis larger during winter than during summer?
This may be due to extensive snow/ice cover and relatively large albedos during winter, or to changes in the atmospheric circulation or thermodynamic structure. Note that the actual annual cycle is lower than the sinusoidal fit during the solstice seasons (summer and winter) but higher during the equinox seasons (spring and fall). This indicates that the semiannual cycle (or second harmonic, with two cycles per year) is also important in Minneapolis.

What is the relationship between the seasonal cycle and intraseasonal oscillations?
Intraseasonal oscillations are variations on time scales shorter than seasons (typically 30-60 days). We expect the intraseasonal oscillations to average out over many years, leaving us with only the seasonal cycle. Note, though, that intraseasonal oscillations can also have seasonal variability (e.g., stronger or more frequent in winter than in summer, or vice versa)!

How might the interannual constant of proportionality be relevant for constraining climate sensitivity?
Prof. Held emphasizes the difficulty of constraining climate sensitivity based on observations of interannual variability. Interannual variability in global mean temperature is dominated by the El Niño-Southern Oscillation (ENSO); however, the spatial pattern of the climate response to ENSO is very different from the spatial pattern of the climate response to increases in CO2. This means that estimates of climate sensitivity derived from observed ENSO variability are probably not appropriate for estimating climate sensitivity to increasing CO2 (this is related to the extent of the decoupling between the global mean temperature and the global mean energy budget). Prof. Held speculates that it may be possible to find a way to compensate for these differences in spatial patterns, but to the best of my knowledge no one has done this -- I expect that it would be a difficult problem to solve!

Post 3: The simplicity of the forced climate response

What is the difference between forced response and internal variability? What do model initial conditions have to do with these?
The forced response is the response of the climate system to a change in radiative forcing, such as changes in the distribution or amount of aerosols, a change in the orbit of the Earth around the sun, or an increase in greenhouse gas concentrations. Internal variability arises naturally within the climate system (although the characteristics of internal variability may change as part of the forced response). From the perspective of a GCM simulation, the evolution of the forced response is not dependent on the initial conditions, but the evolution of the internal variability is strongly dependent on the initial conditions. As a result, the forced response is clearer when we use the average of an ensemble of simulations (where each individual simulation has a different evolution of internal variability). Prof. Held uses the analogy of the macroscopic equation of state: we cannot clearly identify the equation of state by looking at the properties of individual molecules, but we can identify it by looking at the collective properties of many molecules.

What is the difference between radiative flux perturbation and radiative forcing?
In this post, there is no difference between these two terms. Radiative forcing can be defined in a number of ways, and different papers (and subfields of climate science) may use different terminology to specify which definition is being used. Here, radiative forcing is defined as the flux imbalance after the "fast response" of the atmosphere and land surface (both of which have small heat capacities). This definition is appropriate in this case because the atmosphere and land surface respond to changes in the radiative forcing at time scales less than 1 year, while we are interested in the evolution of annual mean global mean surface temperature. This definition of radiative forcing is sometimes called the "radiative flux perturbation".

What does the one box model actually represent?
The one box model simulates the response of the "climate system" to a varying radiative forcing. Conceptually it is easiest if we think of the climate system as a planet covered entirely by an ocean with a specified depth and heat capacity, but in this post c represents the effective heat capacity of the CM2.1 climate system (which also includes the atmosphere, land surface, and sea ice).

Why use τ = 4 years to fit the simple model to the GCM simulation? What would be the effect of using a different value? Which of c and λ should be prescribed, and how?
Prof. Held uses 4 years in this case because it provides the best fit to the GCM simulation. It is also roughly consistent with the expected response timescale of the ocean mixed layer (as indicated by the approximate equivalence in effective heat capacity between the CM2.1 climate system and an ocean 70 m deep). Using larger values (longer relaxation timescales) will result in a smoother variation — if the relaxation timescale is long enough, the responses to changes in the radiative forcing following volcanic eruptions would be effectively eliminated. Compare this with the result when $\tau = 0$: the responses to volcanic eruptions are both very rapid and much stronger. Which variables to prescribe depends on the goal of the simulation. In the post, Prof. Held specifies τ (to optimize the fit with the GCM simulation) and λ (based on estimates of the strength of the relaxation of surface temperature in the GCM). It could be equally valid to prescribe c (for instance, by assuming the ocean has a certain depth) and λ (for instance, λ for a blackbody planet could be the derivative of the Stefan-Boltzmann law with respect to temperature: $\lambda = 4\sigma T^3$).

If we apply the same forcing to two systems with different heat capacities, will they have the same response?
If all other properties are equal (namely the relaxation efficiency λ), the two systems would have the same equilibrium response but would approach it at different rates (slower for higher heat capacities). Two systems with different values of λ (for instance, sand and ocean water, or two planets with vastly different atmospheres) would also have different equilibrium responses.

Why fit the simple model to a GCM result rather than actual climate? What does the close fit between the two mean?
This post is focused on the forced response, which can be difficult to identify in a single realization (such as the actual evolution of climate). In fact, the simple model can only simulate the forced response (i.e., it has no internal variability). The close fit between the two estimates tells us that a 4-member ensemble is sufficient to highlight the forced response (even if it isn't sufficient to isolate it), and it allows us to estimate the effective heat capacity of the CM2.1 climate system for transient climate sensitivity. It also tells us that the ocean mixed layer is probably the primary contributor to this heat capacity (since the effective heat capacity is equivalent to the heat capacity of an ocean 70 m deep).

The fit between the simple model and the GCM is worse in some years than in others — why?
My guess is that the two main factors are internal variability within the GCM (a 4-member ensemble is not enough to eliminate this) and the inability of the simple model to fully represent nonlinearities in the forced response in 'GCM world'.

In discussing the simplicity of the forced response, we have only considered global mean temperature — is the same true for other variables?
To some extent, yes. For example, global mean relative humidity appears to stay roughly constant, which means that the global mean water vapor response is determined by the dependence of water vapor on temperature as expressed in the Clausius-Clapeyron equation (i.e., global mean water vapor increases at approximately 7% per 1K warming). Other variables (such as precipitation) are more difficult to predict using simple theories, although models suggest that global precipitation increases at roughly 1–3% per 1K warming. Note that both relative humidity and precipitation can change substantially locally, even if their global mean changes are relatively small.

What are prescribed variables? Are these variables able to change in time?
Prescribed variables are variables for which the evolution is determined beforehand (i.e., they are not simulated by the model). They are able to change in time, but those changes are independent of the changes in simulated variables. In other words, changes in prescribed variables can act on the model, but do not interact with the model (because the model does not act on them).

Can external forcings change the internal variability?
Yes. For example, polar amplification of the global warming signal reduces the equator-to-pole temperature gradient, which could in turn weaken the trade winds and reduce the ENSO oscillator to a permanent El Niño-like state.

Post 4: Transient vs equilibrium climate response

Why do we care about both equilibrium and transient climate sensitivity? What is the difference between the two?
The transient climate response (TCR) is defined by increasing the CO2 by 1% per year and evaluating the change in global mean surface temperature at the moment of CO2 doubling. The climate system is therefore not in equilibrium when the TCR is evaluated (there is still a radiative forcing that needs to be eliminated before the system can reach equilibrium). The equilibrium climate sensitivity is the change in global mean surface temperature due to a doubling of CO2 after the radiative forcing is completely eliminated. In other words, take a climate system in equilibrium, double the CO2, let it return to equilibrium, and calculate the change in global mean surface temperature between the two equilibria. The equilibrium climate sensitivity (ECS) is independent of the path to CO2 doubling — a lump increase of 100% all at once should give the same ECS as an incremental increase of 1% per year. These two estimates of climate sensitivity are complementary. The TCR may be more "practical" in the sense that it is more relevant to the near future, but the ECS is also important for developing long-term strategies for adapting to or mitigating global warming.

Why is the value of the TCR is less than the value of ECS? What does their ratio indicate?
The TCR is less than the ECS because additional warming is required to equilibrate the radiative flux imbalance that still exists when the TCR is evaluated. To leading order, the ratio of TCR to ECS gives an indication of the efficiency of heat uptake by the deep ocean. If heat uptake by the deep ocean is very efficient, then the ratio of TCR to ECS will be close to 1. If heat uptake by the deep ocean is very inefficient, then the ratio of TCR to ECS will be smaller. The fact that it is difficult to get ratios as small as those found in some GCMs using the two box model suggests that there may be other factors in the coupled climate response that are not represented well in the two box model framework. One possible explanation is changes in the magnitudes of climate feedbacks with time (see also below). If temperature changes associated with these feedbacks accelerate with sustained warming, this could increase ECS while keeping TCR relatively low.

How can the ECS be independent of the parameter γ?
The parameter γ in the two box model represents the efficiency of heat exchange between the surface layer of the ocean and the deep ocean. If the climate system is in equilibrium, there is no net heat exchange between the surface layer and the deep ocean. Although the parameter γ has a strong impact on the time it takes the climate system to reach its new equilibrium, it does not affect the overall temperature change (the ECS).

Under what conditions could we assume that γ equals zero?
1) Perfect mixing. If the surface layer and deep ocean are completely well mixed, then the system reduces to the one box model.
2) No heat exchange between the surface layer and deep ocean. This could happen if the ocean stratification is very strong and the two layers are effectively independent (like oil and water).
3) There is heat exchange, but no net heat exchange. This could happen if heat uptake is perfectly compensated by heat release in other locations.

Do atmospheric composition feedbacks affect TCR and ECS? How much?
Yes, atmospheric composition (and other) feedbacks affect TCR and ECS. The extent to which climate feedbacks influence TCR and ECS depends on the timescale of the feedbacks. Atmospheric composition feedbacks (such as water vapor and clouds) respond quickly to climate changes (for evidence of this, see the water vapor and cloud responses to volcanic eruptions or ENSO variability, which occur on relatively short timescales). My understanding is that these feedbacks should have about the same impact on TCR and ECS. For example, the water vapor feedback roughly doubles the warming (and hence the TCR and ECS) expected from CO2 alone. Feedbacks that occur on long timescales (some land cover feedbacks may fall in this category) would likely have a greater impact on ECS than on TCR. Note that these feedbacks could be positive (decreasing the ratio of TCR to ECS) or negative (increasing the ratio of TCR to ECS)!

What are the 'fast' and 'slow' components of the climate response? What is the 'relaxation timescale'?
In the context of this blog post, the fast component of the climate response is the portion of the climate system that responds to forcing changes within a few years (e.g., the atmosphere, the land surface, and the surface layer of the ocean). The slow component is the portion that takes centuries to equilibrate (mainly the deep ocean, along with some portions of the cryosphere). The relaxation timescale is the exponential decay timescale (i.e., the time required for a ~60% decrease) for the temperature disequilibrium (the difference between the current temperature and the eventual equilibrium temperature). In the two box model framework, this timescale depends on c (the heat capacity), β (the strength of the radiative restoring) and γ (the strength of the 'deep ocean' resistance). In reality, β (which should include the effects of radiative feedbacks) and γ (which depends on the ocean circulation) could both change in time, particularly if the spatial structure of the climate response is time-dependent. The two box framework assumes a tight coupling between the global mean temperature and the global mean energy budget; we know that this assumption is inexact.

What is meant by 'the unperturbed flow transports the perturbed temperatures, and the perturbed flow transports the unperturbed temperatures'?
Heat transport depends on the transport of heat (temperature) by a circulation (flow), which we could write as $H = u \cdot T$. If the response to a climate forcing is approximately linear, we can separate changes in the heat transport into two components. The 'unperturbed flow' and the 'unperturbed temperatures' are the climatological circulation and temperature distribution before the climate forcing is applied. The 'perturbed flow' and the 'perturbed temperatures' are the changes in the circulation and temperature distribution after the climate responds. This sentence suggests that changes in heat transport $\Delta H = \Delta(u \cdot T)$ can be approximated as the sum of new heat transported by the original circulation and original heat transported by the new circulation ($u \cdot \Delta T + \Delta u \cdot T$).

Has the distribution of TCR to ECS ratios changed in the recent IPCC Fifth Assessment Report?
The distribution has changed somewhat (you can find the figure on the course website, and estimates of TCR and ECS from 23 models on page 818 of the AR5 Physical Science Basis document), but not much. The median ratio from these 23 models is 0.57 (as opposed to 0.56 for the 18 CMIP4 models shown in the blog post). The peak of the distribution has shifted right by one bin (from 0.55–0.6 to 0.6–0.65), and the overall distribution is clustered more tightly around its center. Any changes are not due to the new representative concentration pathways (RCPs). TCR is defined according to its own RCP (i.e., CO2 increases by 1% per year until its concentration has doubled), and ECS is defined as the equilibrium response to doubled CO2 and is therefore independent of RCP. Changes are more likely due to changes in model physics and/or numerics.

Post 5: Time-dependent climate sensitivity

What is the difference between models in which CO2 an input and models in which emissions are an input and CO2 is an output?
In the first case, CO2 is a prescribed variable (as discussed in post 3), and its evolution is determined before the model is even run. This type of simulation can be used to see how the model climate system responds to a particular evolution of CO2 concentrations (such as the instantaneous doubling experiment described in this post). In the second case, the model attempts to simulate the evolution of CO2. Emissions are prescribed to represent human use of fossil fuels (or natural CO2 release/uptake, as in some paleoclimate simulations), and the evolution of CO2 is simulated based on the model's representation of the carbon cycle. Climate models that explicitly represent the carbon cycle are often referred to as 'earth system' models. Some modeling groups are even experimenting with simulating emissions (for instance, by simulating the evolution of human societies). Prof. Held sometimes refers to this spectrum as a hierarchy of models, which ranges from the very simple (e.g., the one and two-box energy balance models we have already encountered) to the exceedingly complex (models that attempt to simulate the evolution of human societies), and has argued that progress in climate science will require developing and applying the full range of these models rather than exclusively focusing on the development of models that are increasingly complex and 'complete'.

What does it mean that the dots in the figure are all located below the line?
The straight line represents a linear evolution from the moment of an instantaneous doubling of CO2 (when the radiative forcing is largest and the change in temperature is zero) and the equilibrium climate response (when the radiative forcing is zero and the change in temperature is maximum). The dots represent the evolution of the radiative forcing and the temperature change in individual years. The dots fall below the line because of differences in the radiative restoring (the parameter β) associated with the fast and slow responses. If the restoring is large, then a small change in surface temperature can eliminate a large fraction of the radiative flux imbalance at the top of the atmosphere. This appears to be the case for the fast response in the GFDL CM2.1 model. By contrast, if the restoring is small (as appears to be the case for the slow response in this model), then a much larger change in temperature is required to eliminate the same amount of radiative flux imbalance at the top of the atmosphere. In the opposite case (i.e., if the restoring associated with the fast response were small and the restoring associated with the slow response were large), the dots would be located above the line. In the post, Prof. Held proposes that the difference in radiative restoring is associated with differences between the spatial pattern of the fast response (in which the warming is more spatially uniform) and the spatial pattern of the slow response (in which the warming is largest at high latitudes). The exact reason why the restoring is weaker for a pattern with strong polar amplification is more mysterious, but Prof. Held indicates that it is mainly related to the spatial structure of cloud feedbacks in CM2.1. It is important to recognize that this mechanism may not be an accurate representation of the response in the natural climate system, or even robust among models — cloud feedbacks in particular are often very different among different models. In fact, my understanding is that the polar amplification of warming observed over the past several decades has been stronger than that predicted by most climate model. This might mean that polar amplification is too weak in the fast response simulated by models, or might indicate that model simulations are missing an important forcing (such as the deposition of absorbing black carbon aerosols on ice), or might have a different explanation altogether.

The two (red and blue) simulations match up nicely in year 70 — how are they different before year 70?
Before year 70, the two simulations are different because the evolution of the radiative forcing is different. The atmospheric CO2 concentration is instantaneously doubled at the beginning of the red simulation, which therefore has a very large radiative forcing during the first few years. In the blue simulation, the CO2 concentration (and hence the radiative forcing) are gradually increased over time. As a result, the radiative flux imbalance at the top of the atmosphere in the blue simulation will never be as large as the initial radiative flux imbalance in the red simulation. Note that the equilibrium response to a doubling of CO2 (the ECS) should be the same in both simulations.

What are the time scales of the fast and slow responses and how do we distinguish them?
The time scales of the fast and slow responses emerge from the dynamics of the climate system, and can vary across models. Model estimates of the fast and slow responses, including the time scales involved, may also be different from the actual fast and slow responses in the real climate system. The two box model provides a convenient framework for thinking about the fast and slow responses: the response of the upper box is the fast response, and the response of the lower box is the slow response. We can get a good estimate of the time scales of the fast and slow responses by using this type of model to emulate the evolution of surface temperature and radiative forcing in a full GCM (although it is important to note that the two box model does not allow for differences in the restoring coefficients or spatial patterns of the fast and slow responses). In general, the time scale of the fast response is a few years (maybe 2 to 4), while the time scale of the slow response is much longer (we can see from the figure that it is at least 70 years, and probably longer). The long time scale of the slow response is the reason that the system takes a long time to reach equilibrium. This time scale is affected by the rate of mixing between the surface layer and the deep ocean, the spatial distribution of where this mixing takes place (predominantly in the north Atlantic), and the very slow deep ocean circulation (the transit time of the thermohaline circulation is approximately 1000 years).

Does the efficacy of ocean heat uptake differ between high and low latitudes? Why?
The efficacy of heat uptake by the deep ocean is greater in high latitudes than in low latitudes. This is because heat uptake occurs primarily where the vertical mixing is strong, and vertical mixing is strongest in the subpolar ocean (particularly the North Atlantic). The density of ocean water depends on two factors: the temperature (colder water is denser) and the salinity (saltier water is denser). The gradients of both temperature and salinity are largest in the sub polar oceans. Mixing of cold, relatively fresh water from polar regions and warm, salty water from the subtropical regions creates relatively dense surface waters and enhances vertical mixing. The Atmosphere–Ocean Interactions course next spring will discuss these concepts in more detail.

Why is atmosphere–ocean coupling weaker in high latitudes?
Two reasons are the effects of deep convection in the tropics (which communicates the state of the ocean surface to the entire troposphere) and the presence of sea ice at high latitudes (which insulates the atmosphere from the ocean). The reduction of sea ice extent associated with global warming may increase the strength of atmosphere–ocean coupling at high latitudes. Recent model results even indicate that more open water in high latitude oceans could increase the frequency of convection in high latitudes, which could communicate surface conditions to deeper layers and increase the radiative restoring associated with polar amplification.

Post 8: The recalcitrant component of global warming

What is 'perfect engineering'? How is it different from setting emissions to zero?
'Perfect geoengineering' refers to a magical (and almost certainly unrealistic) scenario in which someone develops a method to instantly remove exactly enough CO_2 from the atmosphere to return CO_2 concentrations to pre-industrial values. In this sense perfect geoengineering is like negative emissions, which is different from setting emissions to zero. If emissions are set to zero, CO_2 concentrations will remain high (because CO_2 is a long-lived gas in Earth's atmosphere) but will gradually decline due to ocean uptake and weathering. These processes take a long time. You can see how long it takes to remove a sudden spike in CO_2 by experimenting with the geological carbon model at http://climatemodels.uchicago.edu/geocarb/.

What is the recalcitrant component of the climate system?
The recalcitrant component of the climate system is the portion of the slow response that has occurred. To leading order, in the experiments described in this post, we can think of the recalcitrant component as the integrated ocean heat uptake from the moment CO_2 began to increase from pre-industrial concentrations to the moment when CO_2 concentrations were returned to pre-industrial values.

What is the mechanism for the rapid cooling in the red curves? Is this the 'transient response'?
The rapid cooling in the red curves is the fast response to a strong negative radiative forcing (namely, the instantaneous removal of a large amount of CO_2 from the atmosphere). This is not the 'transient response' that we have discussed previously, which has a very specific definition: the warming due to a 1% per year increase in CO_2 at the moment of CO_2 doubling. This rapid cooling in global mean temperature also illustrates how risky 'perfect geoengineering' could actually be from a practical standpoint. What would the weather be like during such a rapid release of energy? How would that affect fragile ecosystems and human societies?

What is the difference between the 'fixed concentration commitment' and the 'past emissions commitment'?
The 'fixed concentration commitment' is the warming that would result if CO_2 emissions were reduced to the point that they exactly balance CO_2 uptake (so that CO_2 concentrations are constant in time). The 'past emissions commitment' is the warming that would result if emissions were set to zero, in which case CO_2 concentrations slowly fall due to natural processes (such as ocean uptake and weathering), as described above. You can also experiment with these two scenarios by using the integrated impacts model at http://climatemodels.uchicago.edu/isam/.

What happens in the long time limit after CO_2 is returned to 1860 concentrations?
Assuming all other forcings (orbital parameters, solar output, etc.) stay the same, then the system should approximately return to a state with internal variability around the pre-industrial equilibrium. One of the main points of this post is that this return to pre-industrial conditions will take longer if the recalcitrant component is larger (i.e., the fourth red line will take longer to return to oscillating around 0 than the third red line, the third red line will take longer than the second, and so on). Note that it is possible that other portions of the climate system (such as the cryosphere or biosphere) might change in ways that would slightly (or maybe even significantly) change the pre-industrial equilibrium state. Earth system models (in which ice, the carbon cycle, and the biosphere can evolve dynamically) may be useful for exploring questions like this, but (1) these kinds of models are still in their infancy and (2) it is difficult to think of ways to accurately simulate evolution, especially on global scales.

How could the recalcitrant fraction be about the same even if the change in CO_2 were higher or lower than 700ppm by the end of the century?
My impression is that this comment is speculation, but if I understand it right this speculation is likely justified because the ocean heat uptake is proportional to the radiative forcing. Both the surface temperature change and the ocean heat uptake would be smaller for a smaller radiative forcing ($\Delta CO_2 < 700$ppm), and both quantities would be larger for a larger radiative radiative forcing ($\Delta CO_2 > 700$ppm). Prof. Held's intuition is that the recalcitrant fraction (which we can think of as something like global ocean heat uptake divided by global surface temperature change) would be about the same, regardless of ΔCO_2.

How does the spatial structure of the fast response differ from the spatial structure of the recalcitrant component?
The exact spatial structures of the fast and recalcitrant components will vary among different models, and a complete description is beyond the scope of this class. That said, my understanding is that the spatial structure of the fast response should be much more spatially uniform (because the forcing is spatially uniform), with spatial variability mainly due to the distribution of the oceans and continents. Because the recalcitrant component resides mainly in the deep ocean, I expect that its spatial pattern will be dependent on the ocean circulation (particularly the regions where mixing between the surface layer and the deep ocean preferentially occurs). The timescale of the deep ocean circulation is also very long, which is one reason why the spatial pattern of the recalcitrant component changes in time.

One of the comments below mentioned 'missing heat' — what is the 'missing heat'?
This likely refers to the recent global warming 'hiatus', in which global mean surface temperatures have stayed roughly constant since the late 1990s even as greenhouse gas concentrations have continued to grow. Most scientists believe that this 'missing' heat is being stored in the oceans, but the location of this heat storage and the mechanisms involved are still a matter of debate. We will discuss these concepts further later in the semester (and will cover it in more detail in the Atmosphere–Ocean Interactions course next spring).

Could ocean heat uptake ever be saturated?
In the sense of a continuously increasing radiative forcing, yes — but it is not something we need to spend much time considering right now (except perhaps in thinking about the evolution of climate on other planets). Physically, 'saturating' the ocean heat uptake would mean boiling the oceans until all of the water is released as vapor! A more practical definition of heat uptake 'saturation' for our purposes might refer to the asymptotic reduction in heat uptake as the climate system approaches equilibrium (at equilibrium, net heat uptake should be zero).

Post 6: The transient climate response

Why does a 5-year running mean remove much of the effects of ENSO variability?
The period of ENSO variability is 4 to 7 years, which means that most 5 year periods should contain both a warm phase and a cold phase of ENSO. A five-year running mean will therefore smooth a lot of ENSO variability, because it will average warm and cold phases together. It is important to note, though, that the amplitude of contiguous warm and cold phases could be very different — if the warm phase during a given 5-year period is much stronger than the cold phase in the same period, then the mean for that 5-year period will still reflect the ENSO warming (just not as strongly as an annual mean time series). Moreover, some 5 year periods may contain only a warm phase, or only a cold phase, or even no ENSO variability at all.

How can we determine the range of natural internal variability?
In climate models, we can run and analyze long simulations under equilibrium conditions (such as pre-industrial control simulations run for CMIP3 and CMIP5), or run multiple forced simulations and remove the mean forced response (i.e., the red lines minus the blue line in the figure for this post). In the real world it is more difficult. For instance, the forced responses to greenhouse gases and anthropogenic aerosols must be removed from recent data to explore the effects of natural internal variability. This is often done by removing long-term linear trends, which is probably reasonable to leading order for WMGGs, but not necessarily for aerosols. Various methods of time series analysis can also be used to isolate variability on different timescales. Paleoclimate proxy time series offer another alternative, but temperature reconstructions based on these data have substantial uncertainties (particularly at annual timescales) and one still needs to controls for changes in solar forcing (due to changes in orbit, precession, etc.).

Internal variability in climate models is dependent on initial conditions — how do we get the "correct" internal variability?
In free-running models, we cannot control the internal variability, which largely emerges from the coupling between the atmosphere and ocean. We can use these free-running coupled models in a few different ways. The way we have focused on so far is based on running three (or more) simulations and averaging them so that the forced response is emphasized. In this case, the internal variability is "noise" and the forced response is the "signal". We could also remove any long term trends and focus on the characteristics of internal variability (phases, amplitudes, etc.) — in this case, the internal variability is the "signal", and the forced response is "noise". If we are interested in reproducing the observed time series of internal variability as closely as possible, then we should probably not use a fully coupled model. We can simulate the evolution of atmospheric conditions by prescribing the SST, or simulate the evolution of the ocean by prescribing winds and fluxes at the atmosphere–ocean interface. In both cases, we can

Why is the simulated response to increases in WMGGs larger than the observed (GISS) warming?
The blog post offers two explanations: the transient climate response may be too large in the climate model, or the additional warming may have been counteracted by processes not included in the model simulations (most likely increased aerosol emissions). It is difficult to quantify which of these is more important or more likely, especially because we do not know the transient climate response of the real climate system. The median TCR based on the CMIP5 models (~1.8K) is about the same as the median TCR based on the CMIP3 models (~1.9K).

How long is the lifetime of CO2? Was the CO2 that was in the atmosphere in 2005 still in the atmosphere in 2012?
The lifetime of CO2 in the atmosphere is on the order of 30-100 years. This is considerably longer than the 7 years between 2005 and 2012, so most of the CO2 that was in the atmosphere in 2005 was still in the atmosphere in 2012, supplemented also by new CO2 from fossil fuel (and natural) emissions between 2005 and 2012 (about 40% of the emissions each year stay in the atmosphere). Meanwhile, some of the 2005 CO2 would have been removed from the atmosphere by ocean uptake and other processes.

What contributes to the differences in CO2 concentrations between the Northern and Southern Hemispheres?
Differences in emissions (which are largest in the more heavily populated NH), differences in ocean uptake of CO2 (ocean surface area is larger in the SH, but the strongest mixing between the surface and deep ocean takes place in the North Atlantic), differences in uptake by vegetation (which plays a very important role in the seasonal variability of CO2) or ice sheets (via snowfall) — all of these processes likely play a role.

Why isn't water vapor considered a forcing agent?
Water vapor is the most abundant and influential of the greenhouse gases, its amount is tightly controlled by temperature (it condenses and falls out as rain when its concentration is larger than the temperature permits). Most water vapor in the atmosphere comes from ocean evaporation, and its lifetime in the atmosphere is short (1-2 weeks in the troposphere — water vapor is not well-mixed!). As a result, human emissions of water vapor have very little impact on the overall atmospheric concentration, and we do not consider it as part of the anthropogenic forcing. However, the temperature dependence of the maximum (or saturation) water vapor concentration means that air can 'hold' more water vapor as the temperature increases. Changes in water vapor therefore represent a strong positive climate feedback.

Would including the effects of methane oxidation on stratospheric water vapor change the results?
Methane oxidizes in the stratosphere, with each molecule of methane producing approximately two water vapor molecules. Since water vapor is a strong greenhouse gas, increases in stratospheric water vapor result in warmer surface temperatures. I don't know exactly how the greenhouse impact of two water vapor molecules compares to that of one methane molecule, though I would guess that the greenhouse effect of the water vapor should be slightly stronger (because water vapor absorbs across more of the spectrum). Including these processes would probably not change the results too much.

Can we account for different anthropogenic or volcanic aerosol emissions scenarios in climate models?
Yes, it is easy to change the prescribed temporal and spatial distributions of atmospheric aerosols in models. Some models can even simulate the evolution of aerosols in the atmosphere.

Why are we only discussing climate dynamics from a model perspective?
Prof. Held is a climate modeler and his main expertise is in climate modeling, so he focuses mainly on climate model experiments and results. Some of the later posts will spend more time discussing observations and comparisons of model results and observations. Issues with the accuracy of observations (particularly as we go back in time) probably do play a role in this choice, but more often it is because we are interested in exploring the results of repeatable and perturbable experiments. We cannot run these experiments in the real world, so we use 'model world' instead.

Posts 49 and 50: Volcanoes and the transient climate response

What is the relationship between climate feedbacks and the parameter β in the two-box model?
The parameter β is the radiative restoring, and represents how efficient the climate system is at eliminating a radiative flux imbalance at the top of the atmosphere (see also the response to the second question in post 5). In a positive feedback (such as the water vapor feedback), an increase in temperature results in a secondary response that reinforces the radiative imbalance (e.g., a warmer atmosphere holds more water vapor; water vapor is a strong greenhouse gas; an increase in greenhouse gases increases the radiative forcing). A negative feedback is the opposite — an increase in temperature results in a secondary response that reduces the radiative imbalance. Negative feedbacks reduce the temperature change required to eliminate the radiative forcing; positive feedbacks increase it. Therefore, positive feedbacks make β smaller, and negative feedbacks make β larger.

How are TCR and ECS related to the fast and slow responses?
Both TCR and ECS contain contributions from the fast and slow (and intermediate) responses of the climate system. The difference is in the relative magnitude of these contributions. The ECS is based on the full response of the system, and therefore contains the full measures of the fast and slow responses. The TCR is based on the partial response of the system at the time of CO2 doubling, so that the proportion of the TCR that is attributable to the fast response is higher than the proportion of the ECS that is attributable to the fast response (even though the total temperature change due to the fast response is similar or even higher at equilibrium), while the proportion of the TCR that is attributable to the slow response is lower.

Why does the estimate of TCR based on the response to a volcanic eruption underestimate the TCR?
Several factors may contribute to this. Perhaps the most important is that the response to volcanic forcing is biased toward the fast response, while the intermediate and even slow responses may contribute substantially to TCR. This effect is particularly obvious in the differences between CM2.1 and CM3 (for which the intermediate response is much stronger). Second, the spatial distribution of the forcing due to a volcanic eruption is also different from that of the forcing due to CO2. Third, some aspects of the response to a negative forcing (volcanic aerosols) might be different from the response to a positive forcing (CO2). Note that these latter two do not necessarily imply that the response to volcanic forcing should be weaker than the response to CO2, but they are reasons that it might be different.

Why account for ENSO in the response to volcanic eruptions but not the response to a doubling of CO2?
The aerosol forcing following volcanic eruptions is relatively short-lived, and therefore the ENSO state at the time of the eruption might affect the climatic response. By contrast, the forcing due to a doubling of CO2 is long-lived and will persist through a large number of ENSO cycles.

Why are the responses to 2 × CO2 and (–)0.5 × CO2 so similar? What causes the differences between them?
These two responses are similar because the greenhouse impact of CO2 (and other greenhouse gases) is approximately logarithmic, so that the forcing due to a doubling of CO2 has about the same impact no matter what the initial value of CO2 is (i.e., doubling and halving CO2 approximately result in equal and opposite forcings). I don't know the exact cause of the differences between them, but I guess that differences in the spatial structure of the slow response may play a role: heat uptake by the deep ocean and heat release by the deep ocean may occur in different locations or at different rates.

How can we determine when the forcing due to a volcanic eruption has decayed to the point that it can be ignored?
For climate model simulations in which the only forcing is due to the volcanic eruption this is easy — we wait for the radiative imbalance at the top of the atmosphere to be eliminated and the system to return to varying around its equilibrium state. Things are slightly more complicated in the real world (because the volcanic radiative forcing is superimposed on other forcings and observations of net radiation at the top of the atmosphere are incomplete), so we might instead choose to monitor the evolution of volcanic aerosols in the stratosphere (since these provide the bulk of the forcing).

Why are the results using CM2.1 and CM3 so different?
Prof. Held indicates that the primary source of this difference is the different parameterizations of sub-grid scale convection in CM2.1 and CM3. I think his point is that the similarity between estimates of CM2.1 TCR based on the response to volcanic eruptions and the actual CM2.1 TCR may be specific to CM2.1; in particular, an estimate of TCR based on the response to a volcanic eruption in 'CM3 world' would not be a very good estimate of the actual CM3 TCR. This raises the question: can we use the real-world response to volcanic eruptions to estimate the real-world TCR?

Could climate change affect volcanic eruptions or other geological movements?
This question is still under debate. Volcanic eruptions certainly affect climate (the eruption of a 'supervolcano' like that underneath Yellowstone National Park in the USA could alter climate so drastically that we probably would not survive!), but recent research indicates that climate change can also affect volcanic activity (volcanic eruptions may be more common during periods of rapid ice melt). So should we include these effects in climate models? One could argue that yes, we should aspire to build the most complete (and realistic) climate models that we can. On the other hand, the importance of these effects is not yet firmly established, and their inclusion would be difficult and computationally expensive. To my knowledge, no one has yet tried.

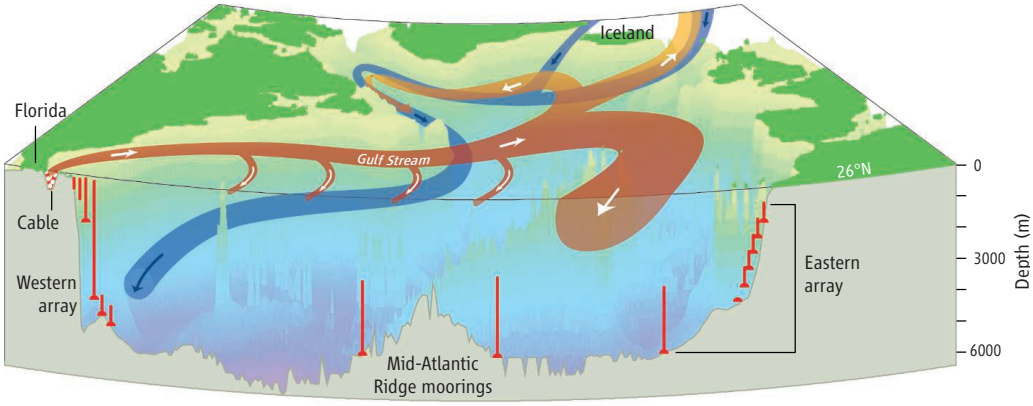
How can differences in the representation of sub-grid scale moist convection have such a strong impact on climate sensitivity?
The representation of sub-grid scale moist convection has important effects on the distribution of clouds (both vertically and horizontally), and therefore is a major factor in determining the strength of the cloud feedback. Cloud feedbacks are the largest source of uncertainty in simulated climate sensitivity (see, e.g., Fig. 9.43 in the IPCC Fifth Assessment Report).

Volcanic eruptions release a lot of heat, so why do they cool the climate?
The release of heat by volcanic eruptions into the atmosphere is local and efficiently dissipated by radiation (the time scale for the atmosphere to relax back to equilibrium following a sudden heating is about one month) and the atmospheric circulation (heat can be converted to kinetic energy and change the local weather). The cooling effect is caused by the injection of volcanic aerosols (and aerosol precursors) into the stratosphere, where they spread out and reduce the amount of energy (sunlight) entering the climate system. These aerosols can persist for several years, and the associated cooling overwhelms any warming due to heat (or greenhouse gases) released from inside the earth (particularly at time scales longer than a month).

What are the major factors in determining ocean heat uptake efficiency?
The ocean heat uptake efficiency can be affected by deep convection, upwelling/downwelling, diffusion, internal wave breaking (turbulence), and other processes in the ocean. The occurrence frequency and strength of deep convection in the North Atlantic and upwelling in the Southern Ocean are likely to be particularly important factors. Overall, the ocean circulation, which is driven primarily by surface winds, determines ocean basin-mode internal variability, and hence the heat uptake.

How does a weaker Atlantic Meridional Overturning Circulation (AMOC) result in a colder global mean temperature?
We should start by clarifying the definition of the AMOC. Warm water emerges from the Florida Strait and flows northward along the east coast of North America in the Gulf Stream. After leaving the coast of North America, the warm water flows northeastward and then diverges, with some water continuing northward beyond Iceland and the remainder returning in the broad southward branch of the subtropical gyre in the upper kilometer or so of the ocean. As it flows northward, the warm water loses heat to the atmosphere and becomes denser. This colder, denser water sinks at high latitudes and returns southward at depths of 2 to 5 km. This northward flow of near-surface, warm water and southward flow of deeper, colder water is known as the North Atlantic meridional overturning circulation. According to this definition of the AMOC, it is clear that a weaker AMOC will reduce the northward heat transport in the ocean. In other words, heat becomes trapped in the tropics. The tropical atmosphere is warm and moist, while the middle and high latitude atmosphere is relatively cold and dry. This difference means that heat release from the ocean to the atmosphere is more efficient at high latitudes. A weaker AMOC reduces this heat release, and therefore results in a colder global mean temperature.

What kind of data sets or indices can be used to monitor AMOC variability?
The Atlantic Multidecadal Oscillation and related AMOC variability can be monitored using sea surface temperature anomalies (see, e.g., HadISST or COBE). The warm phase of the AMO corresponds to a strengthening of the AMOC, and vice versa — note, however, that the AMO is only partially caused by changes in the AMOC. Other aspects of the overturning, including poleward flow in the Gulf Stream, can be measured using buoy arrays (see, e.g., Cunningham et al. 2007). The UK National Environmental Research Council (with support from U.S. agencies) deployed the RAPID/MOCHA (Rapid Climate Change/Meridional Overturning Circulation and Heat Flux Array) buoy array in March 2004 to continuously monitor the AMOC at 26°N (see figure below). The RAPID/MOCHA array measures bottom pressure, temperature (T), and salinity (S) at the eastern and western boundaries of the North Atlantic. These observations are sufficient to monitor variability in the AMOC because horizontal pressure differences across the basin are proportional to variations in the integrated northward flow. The ocean interior measurements are complemented by the Gulf Stream transport estimated from the undersea cable data through the Florida Strait, and the surface wind-driven transport can be calculated from QuikScat satellite measurements of wind stress.



Is the AMOC an example of processes that contribute to the intermediate response time scale?
Yes, but the AMOC does not only represent the intermediate response time scale. The AMOC includes both surface and deep circulations and therefore changes on both fast and slow time scales, with variability that ranges from daily, to seasonal, to multi-decadal, and even longer time scales (up to at least 1000 years).

Post 27: Estimating TCR from recent warming

Why estimate TCR using only the most recent 30 years?
First, the rate of warming in the Northern Hemisphere has been large relative to the rate of warming in the Southern Hemisphere over the past 30 years, which suggests that the warming signal can be separated into two components: one that affects the two hemispheres roughly equally (the WMGG forcing) and one that affects the Northern Hemisphere but not the Southern Hemisphere (everything else, AMOC variability or a reduction in aerosols being the prime candidates). Second, if we go back further than 30 years, the rate of warming was smaller in the NH than in the SH (in fact, the NH was cooling on average, rather than warming). This suggests either that the forcing was fundamentally different (e.g., increasing rather than decreasing aerosol forcing) or that natural variability was in a different phase. Both of these complicate the simple separation into two components (although this can be addressed using the approach described in Post 38). Third, observations of temperature and the global energy budget have been much more complete over the past 30 years. This allows the use of a broader set of variables in investigating TCR (which, in this case, is of interest mainly because it is the reason that several published estimates of TCR focus on the most recent 30 years).

The major source of internal variability on 30-year time scales is the pole-to-pole overturning circulation. What is this circulation? Why does it affect the Northern and Southern Hemispheres differently?
The pole-to-pole overturning circulation that is discussed in this post consists primarily of the Atlantic Meridional Overturning Circulation (see Q&A for Posts 49 and 50). Variability in this circulation affects the Northern and Southern Hemispheres differently because it transports heat from the SH into the NH, and has a stronger impact on mean temperatures in the NH than on mean temperatures in the SH because this heat mostly originates in the tropics.

Why is it simplest to assume that neither internal variability nor aerosols affect the trend since 1980?
This is the approach used by Gregory and Forster (2008). It is the simplest case because it means that the time series of global mean temperature is only dependent on the evolution of the WMGG forcing — in the context of Post 38, this assumption is equivalent to assuming that $A(t) = 0$ and the entire signal is due to $G(t)$. However, examining the time series of NH and SH mean temperature separately suggests that things are not this simple.

What is the cause of the relative flattening of the temperature time series between 1940 and 1980?
The prime candidates are (1) a decreasing in the growth of the forcing due to WMGGs, (2) an increase in cooling due to aerosols, and (3) natural variability associated with the AMOC. The last two are likely to affect the Northern Hemisphere more strongly than they affect the Southern Hemisphere, which is the physical basis for focusing on the difference in warming between the two hemispheres and separating variability in the temperature time series into two contributions with different spatial patterns. In particular, the contribution of changes in WMGGs should be relatively similar between the two hemispheres (hence the small value of g in Post 38), while the effects of other processes (aerosol forcing or natural variability associated with AMOC) are expected to affect the Northern Hemisphere more strongly than they affect the Southern Hemisphere (hence the constraint that $a >> g$).

Could we also define TCR and ECS in terms of other forcing agents?
Yes, although the definitions we should use depend on why we are interested in evaluating the response to other forcing agents in the first place. If we are interested in evaluating the response to the evolution of a different forcing, and do not intend to compare this response to the TCR or ECS due to a doubling of CO2, then we can create new definitions that are appropriate for our study (probably it is better to avoid confusion by not calling these quantities TCR and ECS). If we want to compare the response to a different forcing agent and the response to CO2, then the most appropriate approach would be to set the total radiative forcing due to the other forcing agent to be the same as the radiative forcing due to a doubling (or halving) of CO2. We can then evaluate and compare the equilibrium responses to each forcing agent. We could do the same thing with TCR (i.e., allow the other forcing agent to evolve in a way that keeps the evolution of the global forcing equivalent to a 1% per year increase in CO2), but the usefulness of this experiment is less clear (this framework doesn't make much sense for a volcanic or solar forcing — maybe geoengineering?).

How can we determine the value of heat uptake efficiency?
It is relatively easy to determine the value of heat uptake efficiency in a GCM because we know the exact evolutions of the radiative forcing, the surface temperature, and the ocean heat content. It is much more difficult to evaluate the heat uptake efficiency in the real world because observations (particularly observations of ocean heat content) are more uncertain and more scattered in space and time.

Is the value of ξ affected by internal variability?
I suspect that it is: if we think about the physical meanings of β (the radiative restoring) and γ (the heat uptake efficiency), both parameters are likely to be affected by internal variability. We don't consider the time dependence of ξ in the simple model framework — should we? I would argue no. First, the TCR and ECS are functions of the forced response, and are therefore independent of the exact evolution of internal variability (i.e., the TCR and ECS are independent of the path taken to reach them). Second, the integrated effects of internal variability on ξ are included when we estimate β and γ from climate model output, provided we use a sufficient number of simulations or simulations that are sufficiently long.

Could differences between the NH and SH warming be due to a lag in the SH response rather than aerosols or internal variability?
A lag in the SH response to changes in radiative forcing could explain the differences in the rate of warming between the NH and SH over the past 30 years — perhaps the NH just responds more rapidly to increases in WMGGs. Furthermore, as one of you noted, the amplitude of the changes is consistently larger in the NH, so the exaggerated cooling in the NH prior to 1980 could fit with a global reduction in radiative forcing. However, if a lag in the SH response is the dominant cause of the SH–NH difference, then the assumption is that the forcing is roughly globally uniform. Aerosols don't fit this assumption (because aerosols are short-lived, and were predominantly emitted in the NH), and WMGGs increased pretty much monotonically throughout the twentieth century (and are therefore inconsistent with a reduction in radiative forcing between 1940 and 1980). In my opinion, the evidence is more consistent with either a secondary forcing (e.g., aerosols) or natural variability (e.g., the AMO) that disproportionately affects the NH.

Note: Prof. Held presents the TCR in units of both °C and K. These two units are interchangeable in this case, because the TCR is defined as the difference between two temperatures.

Post 38: NH-SH differential warming and the TCR

Why was the NH temperature anomaly lower than the SH temperature anomaly near 1970 but higher near 2000?
In formulating the simple model, we assume that these differences are because of variations in the non-WMGG component *A(t)*, which affect the NH more strongly than they affect the SH (because $a > 0$). The two most likely possibilities for the non-WMGG component are aerosol forcing (aerosols disproportionately affect the NH because they are short-lived and mainly emitted in the NH; if the negative radiative forcing due to aerosols was larger around 1970 and then began to decrease, this would explain the difference) and natural variability (the Atlantic Multidecadal Oscillation disproportionately affects the NH because it affects northward heat transport out of the tropics into the North Atlantic; if the AMO were in a cold phase around 1970 but a warm phase around 2000, this would explain the difference).

Why do the curves differ more toward the end than toward the beginning?
The NH temperature anomaly *N(t)*, the SH temperature anomaly *S(t)*, and the difference between the two temperature anomalies *N(t) – S(t)* are all larger near the end of the record than near the beginning. This is because the anomalies are defined relative to the beginning of the record (the base period is not specified, but must be the average of some set of years near the beginning of the time series because that's when the anomalies are close to zero). Both *G(t)* and *A(t)* should be larger (and more sensitive to the values of *g* and *a*) when *N(t), S(t)* and the difference between them are larger.

Why does a small value of *g* suggest that the non-WMGG component is dominated by internal variability?
A small value of *g* indicates that the non-WMGG component is dominated by internal variability because it requires that the non-WMGG component account for a large fraction of recent warming in the global mean temperature (in this case, the NH response to WMGGs must be almost the same as the SH response to WMGGs, so the non-WMGG component must have caused strong warming in the NH). It is difficult to attribute strong warming in the NH to aerosol forcing, but easy to explain it as a consequence of natural variability (e.g., a stronger AMOC results in a warmer NH).

Why does a large value of *g* indicate that non-WMGG component is dominated by aerosols?
A large value of *g* means that the response to WMGGs is stronger in the NH than in the SH. If the value of *g* is large enough, then the non-WMGG component (which is concentrated in the NH) could have had a negative impact on NH (and global mean) temperature despite the larger rate of warming in the NH. This picture is consistent with the non-WMGG component being dominated by aerosols (which have represented a net negative forcing throughout the latter half of the twentieth century).

How does the parameter *a* impact the estimate of TCR?
The parameter *a* determines how much the non-WMGG component affected the SH. If $a = 1$, this means that the non-WMGG component had no impact on the SH (i.e., the SH temperature change was solely due to WMGGs). By contrast, if $a = 0.5$, this means that the warming in the SH due to WMGGs was larger than the observed warming (assuming that *A(t)* was negative), because the actual response due to WMGGs was reduced by $0.5 \times A(t)$. The TCR depends only on *G(t)*, so smaller values of *a* indicate larger values of TCR (note that the opposite is true if *A(t)* is positive, as is apparent in the figures).

Why do changes in *a* affect the estimate of TCR more strongly when *g* is large?
When *g* is small, the TCR is strongly constrained by the temperature series in the SH (because the response to WMGGs must be nearly the same in the NH and SH). As a result, changes in *a* do not have much impact on the estimated TCR. As *g* increases, the constraint it places on TCR decreases because more of the increased warming in the NH can be explained by the response to WMGGs. The importance of *a* for determining TCR therefore increases.

The parameters *g* and *a* are both assumed to be positive in this post — could one or both be negative?
Yes, but our physical understanding of *G(t)* and *A(t)* suggests that both *g* and *a* should have been positive during the recent past. In particular, *G(t)* is the response to WMGGs. Emissions and concentrations of WMGGs are larger in the NH than in the SH. The ultra-fast response to changes in WMGGs may also be larger in the NH, because of the distribution of land and ocean. The main contributors to *A(t)* are expected to be aerosol forcing and/or the AMO, both of which also affect the NH more strongly than the SH. This is also the physical basis for assuming that the WMGG and non-WMGG components have different impacts in the Northern and Southern Hemispheres.

Should the parameters *g* and *a* be time-dependent?
One could make an argument that these two parameters depend on time (as one of you noted, the difference between the NH and SH temperature time series is influenced by anthropogenic activity, land use changes, geomorphology, natural variability, etc.). However, representing this time dependence would be exceedingly difficult, and even if we could, doing so would severely complicate the model and make it both less understandable and less useful. If we want to simulate all of these dependencies, it would be better to abandon the simple model and use an Earth System Model instead.

Based on the figures, can we conclude that the non-WMGG component mainly has a cooling effect?
We can conclude that the non-WMGG component mainly had a cooling effect over the past 50 years, but we cannot conclude that it will always be this way. For instance, in the absence of aerosol forcing, the non-WMGG component should be dominated by natural variability that alternates between warming and cooling. Note that the WMGG and non-WMGG components do not have to have opposite signs!

What is the basis for assuming that the temperature anomaly is a linear combination of the WMGG and non-WMGG components?
This assumption is based on the false but useful idea that the WMGG and non-WMGG components of temperature change are effectively independent of each other (i.e., the WMGG response does not depend on the non-WMGG component and vice versa) and spatially consistent in a hemispheric sense (i.e., variability in the two components is described entirely by the global mean amplitudes *G(t)* and *A(t)*, with *a* and *g* independent of time).

Is the initial concentration of CO2 important for the value of the TCR?
No (at least not at leading order). This is because of the logarithmic dependence of the radiative impact of CO2 on its concentration. For evidence, look at the similarity between the simulated response to doubling CO2 and the simulated response to halving CO2 shown in Post 49.

Posts 16, 17, and 44: Heat uptake and internal variability

How can we define "internal variability"?
Internal variability, which may also be called "natural variability", can be defined as changes that arise from the non-linear coupling between components of the climate system (primarily the atmosphere and ocean), and can be contrasted with external forcings (which arise from "external" factors such as changes in solar radiation or human activity; see also notes for previous posts). The El Niño–Southern Oscillation (ENSO) is the most famous example of internal variability, which generally occurs in the form of oscillations with positive and negative phases. These oscillations are generally associated with natural processes that affect the rate of heat uptake by (or heat release from) the deep ocean.

Why do these posts focus on low-frequency internal variability? What are some examples of low-frequency internal variability?
These posts focuses on low-frequency internal variability (i.e., variability with time scales of decades or longer) because we want to evaluate whether internal variability can explain the observed warming trend over recent decades. This warming trend extends over a long enough period that it could not be explained by internal variability with time scales shorter than a few decades. In other words, we would not expect global warming to be the result of a single ENSO cycle, because ENSO has a characteristic period of 2~7 years, much shorter than the trend. These posts focus primarily on low-frequency variability in the sub-polar oceans, such as the Atlantic Multi-decadal Oscillation (AMO) and the Pacific Decadal Oscillation (PDO).

Why is the maximum amplitude of internal variability located in the sub-polar oceans?
As mentioned above, internal variability typically results from processes that change the heat uptake by (or heat release from) the deep ocean. Mixing between the surface and deep layers of the ocean is strongest in the sub-polar oceans, so that the amplitude of internal variability is largest in these regions.

What is the relationship between heat uptake and global warming? Why do we know more about heat uptake over the past 50 years?
Heat uptake slows the rate of global warming. From the two-box model framework, we can think of heat uptake as reducing the transient climate response, but not affecting the equilibrium climate sensitivity (the ECS is independent of the parameter γ; see notes for post 4). Our improved knowledge of heat uptake over the past 50 years comes mainly from an expanding set of buoys and buoy arrays (such as ARGO).

Does "heat uptake" include any other components besides ocean heat uptake?
Yes, but total heat uptake is dominated by the ocean component because the heat capacity of the ocean is much larger than the heat capacity of any other component of the climate system. There are also components of heat uptake associated with the cryosphere (think of the melting of the Antarctic ice sheet, which takes up a lot of energy and has likely played an important role in limiting recent surface warming in Antarctica) and with the land surface and atmosphere (but recall that these have very small heat capacities, which is why their forced responses are considered "ultra-fast").

Can we also explore relationships between heat uptake, internal variability and global warming using observations?
Of course — but it is very difficult to separate the forced response from internal variability in observations, and our knowledge of heat uptake based on observations is less complete and more uncertain (in models, we know the heat uptake everywhere in the ocean exactly).

What is the rationale for assuming the same radiative restoring ratio β for both external forcing and internal variability?
The radiative restoring ratio β is dependent on the spatial structure of the surface warming. In post 16, we begin by assuming that the surface warming is caused by either external forcing or internal variability (rather than a combination of the two). These two limiting cases have the same radiative restoring ratio, because the spatial pattern of the surface warming is in both cases defined to be the observed spatial pattern of surface warming. This assumption is unrealistic, but the results are still useful. A more realistic approach is to assume that the observed spatial pattern of temperature change is the superposition of the two spatial patterns of temperature change due to external forcing and internal variability. In this case the radiative restoring of the forced response β_F need not equal the radiative restoring of the internal variability β_I, as described in post 44.

What is the physical meaning of the efficacy of heat uptake (β_F/β_H)?
The efficacy of heat uptake is defined here as the ratio of the radiative restoring of the forced response to the radiative restoring of the heat uptake. Prof. Held mentions that this ratio is always greater than 1 in climate models, indicating that $\beta_F > \beta_H$. This makes sense because the radiative restoring is strongest in the tropics (due to ventilation by deep convection). Heat uptake primarily takes place in the sub-polar oceans (as described above), while the forced response is more spatially homogeneous, with a stronger tropical component. Heat uptake would be less effective if the ratio were less than 1, because surface cooling due to heat uptake would be restored more strongly than the forced warming (i.e., the climate system would ventilate the cooling more effectively than the heating). To the extent that β_H is approximately equal to β_I (the radiative restoring of the low-frequency internal variability, which is also concentrated in the sub-polar oceans), I would guess that a larger value of β_H would also reduce the characteristic time scales of internal variability like the AMO and PDO, making them more like ENSO.

How do models consider low-frequency variability?
Low-frequency variability in a model emerges from coupled model dynamics, especially the nonlinear coupling between the model atmosphere and ocean. Low-frequency variability in a model thus represents the physics of "model world", and may be different from the low-frequency variability that emerges in other models or the real world (among other things, β_I may be different). Confidence in model representations of low-frequency internal variability reflects confidence in the representation of the physics of coupled atmosphere–ocean interactions and surface–deep ocean mixing in the model. These physics are complex and poorly understood, so we are not very confident about model simulations of low-frequency internal variability.

The author writes that the dominant time scale of variability in the North Atlantic is shorter than 20 years — isn't this too short for the AMO?
Prof. Held mentions that the North Atlantic dominates variability on time scales of less than ~20 years *in this model* — as mentioned above, this result emerges from the model dynamics and does not necessarily match variability in the real world. The characteristic time scale of AMO variability is estimated to be 50~90 years; however, proxy reconstructions using tree rings indicate that individual positive or negative phases of the AMO have can last anywhere from a few years (less than 10 y) to several decades (more than 60 y), and most of these individual phases (~70%) have durations less than 20 years. In this context, it is probably reasonable that the North Atlantic dominates variability at time scales shorter than ~20 years in this model.

The author mentioned observational constraints on climate sensitivity — what does "constraint" mean?
Constraints are like uncertainty bounds or confidence intervals. An observational constraint on a variable narrows the range of possible values for that variable (i.e., values that don't meet criterion 'X' are inconsistent with the observations). For example, an observational constraint on climate sensitivity would allow us to conclude with high confidence that climate sensitivity is less than some maximum value (or greater than some minimum).

Can you give an example of the interaction between internal variability and external forcing?
The potential interactions between global warming and ENSO covered in posts 41 and 45 (an initial increase in La Niña events as surface layer warming increases the contrast between temperatures in the mixed layer and temperatures in upwelling water, following by an increase in El Niño events as polar amplification reduces the contrast between temperatures in the mixed layer and temperatures in upwelling water) is an example of this kind of interaction. Potential changes in the AMO caused by freshening of the North Atlantic due to the melting of the Greenland ice sheet are another example.

Post 13: The strength of the global hydrologic cycle

What variables should we use to measure the strength of the hydrologic cycle?
As Prof. Held notes, this depends on what we mean by "the strength of the hydrologic cycle". Here, the focus is on global precipitation and evaporation. These variables are constrained by the energy cycle, and useful variables include precipitation (P), evaporation (E), and free tropospheric radiative cooling (changes in radiative cooling must balance changes in latent heating associated with changes in P). Studies of regional precipitation changes might also look at changes in the convergence/divergence of moist static energy. O'Gorman et al. (2012) provide an excellent review of these concepts.

What causes the differences between the simulated and predicted changes in P minus E?
The predicted changes rely on two assumptions: (1) relative humidity doesn't change and (2) the circulation doesn't change. The difference between the simulation and the prediction therefore indicates errors associated with the two assumptions. One or both of the assumptions breaks down, particularly in the tropics (where the simulated increase is greater than the predicted increase) and the subtropics (where the simulated decrease is greater than the predicted decrease). My guess is that this difference is related to a breakdown in assumption (2). In particular, rather than remaining constant, the tropical Hadley cell intensifies and expands poleward. This expansion means that the evaporation that feeds precipitation in the ITCZ is drawn from further away and from a larger area, so that decreases in P minus E are larger than predicted in the subtropics (particularly on the poleward edge) and increases in P minus E are larger than predicted in the tropics. Assumption (1) may break down for similar reasons: enhanced subsidence in the subtropics could reduce near-surface relative humidity and increase evaporation relative to the constant RH prediction.

The model suggests that the subtropics become drier under global warming — does this agree with observations?
The subtropical drying is generally attributed to intensification and expansion of the tropical Hadley circulation (and associated poleward shifts in the dry downward branches of the circulation). Recent observations are consistent with tropical expansion and subtropical drying (one example is the trends in sea surface salinity shown in post 14; see also reviews by Lucas et al., 2013 and Birner et al., 2014), but attribution of tropical expansion to global warming is still uncertain (other possible explanations and/or contributing factors include changes in aerosols and stratospheric ozone).

Why is subtropical drying asymmetric between the southern and northern hemispheres?
Although there are a number of reasons to expect asymmetric drying in the southern and northern subtropical regions (e.g., differences in the spatial distribution of P minus E, asymmetries in the ITCZ, asymmetries in the storm tracks, and differences in the distributions of aerosols and ozone), the model used to generate the figure does not include any of these factors. My guess is that the asymmetric response shown in the figure reflects chaotic effects (i.e., weather specific to the model run) — if the model was run many times or for a longer time, I think that these asymmetries would disappear.

What factors control the strength of the hydrologic cycle? How do these differ from the factors that control the abundance of water vapor?
The strength of the hydrologic cycle is controlled by the atmospheric energy budget: latent heating from precipitation formation must be balanced by radiative cooling. This is one reason it is useful to report changes in evaporation and precipitation in units of W m^{-2} (i.e., as energy fluxes). By contrast, the abundance of water vapor is controlled by temperature. The Clausius–Clapeyron equation that relates saturation vapor pressure and temperature is nonlinear (i.e., changes in saturation vapor pressure are not exactly proportional to changes in temperature), but at current temperatures the rate of increase in saturation vapor pressure is about 7% per Kelvin. Increases in greenhouse gas concentrations (including water vapor) also increase radiative cooling rates in the free troposphere (where most precipitation forms), but this increase is smaller than the simultaneous increase in saturation vapor pressure. Precipitation therefore increases at a slower rate (1~4% per Kelvin) than water vapor under global warming.

Why are regional precipitation changes more closely related to global mean temperature than to the global mean hydrologic cycle?
This is expected if dry places get drier and rainy places get rainier: precipitation in dry places has a negative correlation with global mean temperature, while precipitation in rainy places has a positive correlation with global mean temperature. Meanwhile, global mean precipitation is largely decoupled from global mean temperature (and hence regional precipitation) because of compensation between changes in dry and rainy places. This result also illustrates the importance of the decoupling between global mean temperature and the global mean energy budget (see post 7).

The model in this post uses a slab ocean — is this a reasonable assumption?
In this post, the ocean is basically an infinite source of water vapor. This source is determined primarily by relative humidity in the surface layer, with some dependence on SST. SST is of course affected by ocean dynamics; however, given the other simplifications in the model (no continents, zonal mean annual mean solar radiation), we wouldn't expect realistic ocean dynamics even if we used a (very expensive) full ocean model. The slab ocean approach eliminates this computational expense while still allowing us to look at how E and P respond in this model framework.

Why does radiative cooling destabilize the troposphere?
Radiative cooling can be thought of as a relaxation of the temperature toward the radiative equilibrium temperature (i.e., absorption equals emission). The radiative equilibrium temperature profile is very unstable (see, e.g., the figure in post 19). This destabilizing effect might also be considered in terms of mass or energy balance, at least in the tropics. Radiative cooling balances adiabatic warming due to compression during subsidence. This downward flux balances the upward flux in convection. Stronger radiative cooling/subsidence permits additional convection, which requires instability.

How does the ITCZ form? Are there other convergence zones?
The ITCZ forms over the maximum surface temperatures. Air is lighter and surface pressure is lower over warmer surface temperatures. Low surface pressure drives convergence, which drives upward motion and condensation, which in turn reinforces low-level convergence. This relationship also governs the formation of other convergence zones (e.g., the South Pacific Convergence Zone or the summer monsoons), as well as changes in the locations of convergence zones (e.g., during ENSO)

Why are changes in evaporation under increased CO2 small relative to changes in precipitation?
First, it is important to note that changes in evaporation are small relative to changes in precipitation only in a local sense — in other words, changes in precipitation are more spatially variable than changes in evaporation. In fact, changes in evaporation should be slightly larger than changes in precipitation at the global scale until the system reaches equilibrium (to account for the simultaneous increase in atmospheric water vapor).

Post 14: Surface salinity trends

What are the relationships among precipitation, evaporation and salinity?
Precipitation and evaporation represent fluxes of water between the ocean and the atmosphere, while salinity is defined as the concentration of salt ions in sea water. Evaporation increases salinity (by increasing the ratio of ions to water molecules), while precipitation decreases salinity.

Why is the Atlantic saltier than the Pacific?
The difference in salinity between the Atlantic and the Pacific is primarily a function of differences in fresh water fluxes (mainly precipitation but also inflow from rivers and runoff). Evaporation exceeds precipitation over the North Atlantic, but precipitation exceeds evaporation over the North Pacific. The evaporation maxima in the subtropics and the easterly orientation of the subtropical trade winds play an important role in this difference: evaporation from the Atlantic is easily transported to the Pacific, but evaporation from the Pacific must cross several continents before it reaches the Atlantic.

How does surface salinity respond to global climate change?
As far as I know, this is an open question. Salinity responds both to changes in the distribution of P minus E and to changes in the ocean circulation. Both of these factors are likely to change under global climate change, but the spatial patterns and magnitudes of these changes are uncertain.

What is the justification for ignoring ocean circulation changes in relating salinity trends to changes in P minus E?
This assumption is based on the idea that salinity responses to changes in P minus E and salinity responses to changes in the ocean circulation occur at different time scales (i.e., stronger sensitivity to P minus E at shorter time scales, such as the 50-year period used in the study). This is likely true in a climate change sense, but may be problematic with respect to internal variability (e.g., ENSO or the AMO). The authors of this study controlled for ENSO (assuming a linear relationship), but do not seem to have controlled for variability at longer time scales. This is a key uncertainty in their analysis.

Why has salinity increased in regions where salinity was already high?
The hypothesis is that salinity has increased in regions where it was already high because the process that caused salinity to be high (E greater than P) is amplified under global warming (increased E, decreased P, larger increases in E than P, or some combination of these).

What does it mean that the trends in salinity are larger than expected based on Clausius–Clapeyron?
Prof. Held makes this statement based on calculations that he did himself, which are not included in the post. In the simplest sense, it means that he thinks the trends in salinity are larger than they should be if they are only due to changes in P minus E that scale with global mean temperature (i.e., Clausius–Clapeyron). In a deeper sense, if his calculations are correct, it means that this difference must be explained before we can truly be confident that (1) the salinity trends are dominated by changes in P minus E and (2) they support our understanding of how P and E change under global warming. The difference could be related to several factors that affect salinity trends but are not related to P and E (e.g., ocean circulation changes and changes in the instrumentation used to make the observations, among others). It could also indicate that P and E change at a faster rate than currently expected. One previous study of precipitation trends indicated that global mean precipitation increased much faster than expected during recent decades (about 7% per Kelvin); however, most other studies (including studies based on newer and reprocessed versions of the data used in that study) indicate that precipitation has increased at a rate that is consistent with expectations (1~4% per Kelvin).

How can total salt stay the same if salinity changes by multiplying by a constant?
Since the salinity (S) is always positive, multiplying it by a constant would result in increases everywhere — panel b shows that this is clearly not the case. Evaporation and precipitation do not change the amount of salt, only its concentration (evaporation increases salinity and precipitation decreases salinity, but both only involve fluxes of fresh water — salt is not transferred). The quantity that is (approximately) multiplied by a constant is therefore not the salinity but the anomaly in salinity relative to the global mean (i.e., $S - \langle S \rangle$). In other words, if the local salinity is originally less than the global mean, then it decreases proportionally to the amount by which it is less than the global mean (local freshening). The opposite is true if the local salinity is greater than the global mean: it increases approximately proportionally to the amount by which it exceeds the global mean (local salinization).

How do changes in salinity impact the ocean circulation?
Changes in salinity affect vertical and horizontal density gradients, and can therefore drive changes in the ocean circulation. Liu Qun described some of the ways changes in salinity can affect the ocean circulation in his presentation.

Can changes in surface salinity affect precipitation or evaporation? What about climate sensitivity?
I doubt that changes in surface salinity would have a strong direct impact on P or E. Increases in salinity do decrease the latent heat of vaporization for water, so that an increase in salinity requires an increase in the amount of water vapor transferred for the same surface flux of latent heat, but variations in surface salinity are sufficiently small (and the relative humidity constraint on E is sufficiently strong) that these differences should be negligible. Indirect changes (due to salinity-driven changes in the circulation) may impact P or E in more profound ways. Changes in the ocean circulation due to changes in surface salinity could also affect climate sensitivity by changing the heat uptake efficiency (see notes for previous posts).

Why does accounting for the presence of land include the constraint that P must be greater than E on the time scales of interest?
Most of the precipitation over land comes from ocean precipitation: total evaporation exceeds precipitation over the oceans by about 11 cm/year, while precipitation exceeds evaporation over land by about 27 cm/year. The balance is closed by runoff from the land to the ocean. My understanding is that P must be greater than E at the time scales of interest because the runoff component introduces a delay in closure.

Why is salinity the primary factor controlling density in the subpolar oceans? What about the tropical oceans?
The temperature profile in the subpolar oceans is often almost constant (i.e., the deep water and the surface water are approximately the same temperature). As a result, the vertical profile of salinity can have a strong impact on the vertical profile of density. This is particularly important with respect to the in-mixing of relatively warm, salty water transported north from the subtropics in the Gulf Stream and Kuroshio currents — this water can lose its heat without mixing, but cannot change its salinity. In the tropics, the sea surface is much warmer than the deep ocean, so that the vertical profile of density is primarily dependent on the vertical profile of temperature (surface waters are less dense because they are warmer).