
AMERICAN PHYSICAL SOCIETY
CLIMATE CHANGE STATEMENT REVIEW WORKSHOP

NEW YORK UNIVERSITY
CENTER FOR URBAN SCIENCE AND PROGRESS
One MetroTech Center
19th Floor
Brooklyn, New York 11201
January 8, 2014
8:20 A.M.

TRANSCRIPT OF PROCEEDINGS

IN ATTENDANCE:

AMERICAN PHYSICAL SOCIETY STAFF:

DR. FRANCIS SLAKEY, Associate Director of Public Affairs
JEANETTE RUSSO, Office Manager and POPA Studies Administration Specialist

APS CLIMATE CHANGE STATEMENT STEERING COMMITTEE:

DR. ROBERT JAFFE, Chair
DR. MALCOLM BEASLEY, APS President
DR. KATE KIRBY, Executive Officer

APS CLIMATE CHANGE STATEMENT REVIEW SUBCOMMITTEE:

DR. STEVEN KOONIN, Chair
MR. PHILIP COYLE
DR. R. SCOTT KEMP
DR. ROBERT ROSNER
DR. SUSAN SEESTROM

NYU-CUSP STAFF:

DR. MICHAEL HOLLAND, Chief of Staff
DR. ARI PATRINOS, Deputy Director for Research

EXPERTS:

DR. JOHN CHRISTY
DR. WILLIAM COLLINS
DR. JUDITH CURRY
DR. ISAAC HELD
DR. RICHARD LINDZEN
DR. BENJAMIN SANTER

Transcribed by JOSHUA B. EDWARDS, RPR, CRR

INDEX

DR. KOONIN: Welcome to Brooklyn, everybody, and to this space, which is part of NYU’s Center for Urban Science and Progress. It's an organization that I have been building for the last two years. And during the break, I am happy to tell you more about it.

Thanks, of course, for taking the time to help out the American Physical Society, which is convening this meeting in thinking through its statement on climate change.

The history and context of what we would like to accomplish today were covered in the pre-read material that we sent around, and so I am not going to take time to go through much of that.

But let me just note that this meeting is one intermediate step in an orderly, open, and substantive process to create an APS stance on climate change.

The meeting is convened by the APS subcommittee that is charged with reviewing the statement.

And the meeting's purpose is to explore through expert presentations and discussion the state of climate science, both the consensus view as expressed by several thousand pages of the IPCC AR5 Working Group 1 report that came out three months ago, but also the views of experts who credibly take significant issue with several aspects of the consensus picture.

In doing this, the subcommittee hopes to illuminate the certainties and the gaps in our understanding of the physical basis of climate change for the subcommittee itself, for the APS leadership who are present here as observers, and, through a transcript, for the APS membership and the broader public.

Let me start with introductions

around the room. I would ask each of you to just state your name and the institution and the capacity in which you are here today.

And as you introduce yourselves, you have the option of using your quota of one weather-related remark, after which we will ban all further discussions of weather!

So, I am Steve Koonin and I am Chair of the subcommittee that is responsible for reviewing the statement and making recommendations up the chain.

And I am a professor, of civil and urban engineering in the engineering school here at NYU and a professor of information, operations and management in the NYU business school. And I have never taken a course in either of those subjects!

DR. KEMP: I am Scott Kemp. I am assistant professor of nuclear science and engineering at MIT, where I direct a Lab for Nuclear Security and Policy. I also have a position in the policy schools at Princeton and Harvard. And I am here as a member of the subcommittee.

DR. LINDZEN: I am Dick Lindzen, emeritus professor at MIT in atmospheric sciences.

DR. CHRISTY: John Christy, professor of atmospheric science at the University of Alabama in Huntsville. My one weather comment was made 25 years ago on the Weather Channel when I said, "If it happened before, it will happen again, but probably worse."

DR. CURRY: I am Judy Curry from Georgia Tech, earth and atmospheric sciences.

DR. COLLINS: I am Bill Collins. I head the weather science department at Berkeley. I also teach earth and air science at Berkeley.

And I guess my role here today is as one of the lead authors of chapter 9 in the fifth IPCC report.

DR. SANTER: I am Ben Santer. I am from Lawrence Livermore National Laboratory. And I will be talking today about detection and attribution work and the stasis.

DR. HELD: I am Isaac Held. I am with NOAA's Geophysical Fluid Dynamics Laboratory. And I also teach at Princeton in the program in atmospheric oceanic sciences.

MS. RUSSO: I am Jeanette Russo with the American Physical Society. I am the office manager with the Office of Public Affairs in Washington, D.C. and administrator for meetings like this.

DR. ROSNER: I am Bob Rosner. I am professor of physics and astrophysics at the University of Chicago and chair of the Panel on Public Affairs at the American

Physical Society, which is the parent committee of the committee that Steve chairs.

DR. HOLLAND: Mike Holland, I am the chief of staff here at CUSP.

DR. SLAKEY: Francis Slakey, associate director of public affairs for APS.

DR. JAFFE: I am Bob Jaffe. I am a professor of physics at MIT.

DR. SEESTROM: I am Susan Seestrom. I am a senior fellow at Los Alamos National Laboratory and a member of the subcommittee.

MR. COYLE: Philip Coyle, member of the Panel on Public Affairs, of course, and most recently associate director for National Security and International Affairs at OSTP. I am currently with the Center for Arms Control and Non-Proliferation.

DR. BEASLEY: I am Mac Beasley, currently president of the American

Physical Society and I am a professor of applied physics at Stanford.

DR. KIRBY: Kate Kirby, executive officer of the American Physical Society and formerly senior research physicist at Harvard Smithsonian Center for Astrophysics.

DR. KOONIN: And our Court Reporter?

THE REPORTER: Joshua Edwards, good morning.

DR. KOONIN: I am sure he urges us again to speak up.

DR. BEASLEY: I have sympathy for this gentleman.

DR. KOONIN: We are going to organize our discussion around the agenda [next page] you have seen in outline form. And now I have tried to put in a batting order for our outside experts. My suggestion is we start with Bill, go to Judy, take a break, then have Ben and Dick, we will have a brief break, pick up lunch and then

CCSR workshop agenda

0800-0830	Opening Remarks (over breakfast)
0830-0915	Collins (30 min presentation + 15 min discussion)
0915-1000	Curry
1000-1015	Break
1015-1100	Santer
1100-1145	Lindzen
1145-1215	Break (including pick up lunch)
1215-1300	Christy
1300-1345	Held
1345-1400	Break
1400-1600	Panel discussion (speakers + subcommittee)
1600-1700	Contingency time

John and then finish up with Ike.

If you all have thought about a different batting order, I am fine with that, but this seems about as good as any. And I am sure we will thread themes throughout the discussion. And then again, another break and we will run a panel discussion for as long as it seems useful or until people have to leave. And as you can see, there is an hour of contingency built in, (which we are not up to using yet!).

We will have a transcript and each of the participants (the experts and the subcommittee) within a week, I hope, will have an opportunity to clarify the transcript, which will eventually be made public according to the procedures that we sent around.

To help in the transcription, I am going to try to make sure that only one person at a time is speaking. And if you want to be recognized, I think we are trying to get some pieces of paper that you can wave around.

And we will try to get you (indicating stenographer) the names and a seating chart so that you will be able to know who is talking.

This workshop and its transcript will likely not be the final technical input to what we are about, as the issues raised will no doubt be discussed further by the subcommittee and the broader APS membership.

The scope today: I would like to really keep rigorously to Working Group 1, namely the physical basis for climate change and focus on the science. As important as they might be, we are not going to cover other broader issues like programmatics, communications, climate impacts or societal responses, except perhaps we will touch on them a bit during the panel discussion.

While not all or even most of the APS membership are experienced in climate, it's important to realize that physicists do bring a body of knowledge and set of skills that are directly relevant to assessing the physical basis for climate science. Radiation transfer, including the underlying atomic and molecular processes, fluid dynamics, phase transitions, all the underpinnings of climate science are smack in the middle of physics.

Physicists also have a deep expertise in the handling of large observational data sets and in modeling complex physical systems. And indeed, there has been enough APS interest among the membership that a topical group on the physics of climate was established two years ago.

Those of you who know me know I am not inexperienced in wielding a gavel. And so I won't hesitate to cut off remarks that are out of scope, that go on too long, or that are unproductive toward the goals that we are trying to establish.

As you go about the day, you might just bear in mind that unsupported appeals to authority just aren't going to fly with the APS membership. And our discussions today are going to be read and commented upon by an extraordinarily technically literate and experienced group of more than 50,000 physicists from all over the world. So, in that sense, this is on the record.

Finally, the real practicalities; there is ongoing coffee available over there, and there is even stronger coffee in the pantry which you probably all walked by. Don't hesitate to just step out and grab whatever you need. Signs in the hallway are pointing to the restrooms.

We are also not expecting any fire drills today, but if the alarm does sound and we need to evacuate, just follow one of the locals down the stairs in the center of the floor and then out of the building. People with the yellow hats are particularly important if that exercise should happen.

With that, I think we are ready to start unless somebody else has any questions or comments?

Okay, Isaac?

DR. HELD: Can we expect the presentations to be more or less uninterrupted?

DR. KOONIN: Oh, I missed that, yes, the flow, I'm sorry. I had notes here and just didn't read them. What I would like is that during the 30 minutes of the talk, we will take clarifying questions only from the subcommittee.

After that, we will do subcommittee questions and then we will open it up to the experts. And what I hope will be a productive freeform dialogue.

Okay, Bill…?

DR. COLLINS: Thank you. Good morning. So first, thanks for inviting us to talk with you about the recent findings of IPCC and the scientific context for them.

I think this is a particularly timely time to have this conversation because, as you know, the first volume of the fifth IPCC assessment was issued electronically to the world sort of in two stages, in late September, first the summary for policymakers on September 27th, and then the electronic version of the Working Group 1 report which deals with the science and physics of climate change on September 30th.

Those reports are still going through a set of final edits to get them ready for publication. But this is a good time to be talking about the findings. And I oriented my presentation --

DR. KOONIN: Bill, I think that podium mic is live.

DR. COLLINS: How about that? Much better, yes. And I think I elected to sort of hew to the questions that you raised in your notes that you sent to us. So, my presentation actually deals somewhat specifically with several of the topics that came up that you raised in connection with this report.

So again, to reiterate my role in the IPCC, I have served now twice as lead author, once for the chapter dealing with projections in the fourth assessment and now as a lead author on the chapter dealing with evaluation of models, chapter 9 in the fifth IPCC assessment.

And I have also been heavily involved in constructing one of the climate models that has been used in these reports for a number of different iterations and I am still directing effort in that direction. So, I do climate modeling basically in my professional life.

And I would be happy to both ask for your input on that and also answer any questions you might have about modeling. And there are a number of us here who do that for one of our day jobs.

I thought I would start with the issue of radiative forcing because, after all, this is a forced problem that we are looking at. And just to remind you what the current state of that forcing information looks like [next page], one of the issues that you raised in your notes repeatedly

Natural forcing (TSI) << anthropogenic (topic I.2)

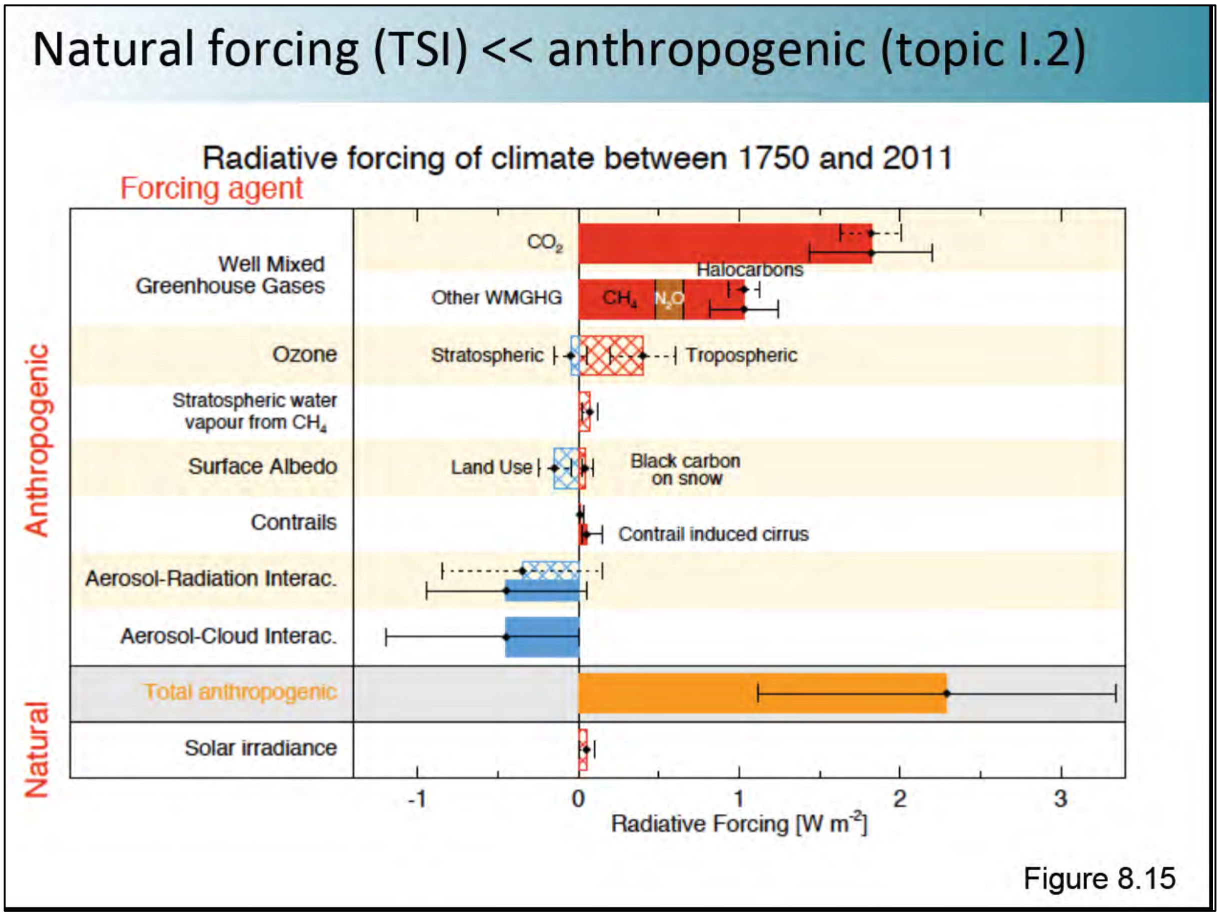


Figure 8.15

was the difference between change that is forced by evolution of the boundary conditions?

And we considered the radiative forcing for this problem, radiative forcing meaning the net radiative balance of the earth's climate system at the tropopause, the boundary between the troposphere and the stratosphere.

We regard that essentially as a boundary condition problem to which the climate system responds. And a number of your questions dealt with the issue of whether or not change in the climate system is forced by evolution in these boundary conditions or by essentially uncertainty in the initial conditions.

And I think a number of us will touch on that topic today in our presentations, because this was a thread in the comments that you brought back to us from the Working Group 1 report.

This is the way that the climate science community looks at radiative forcing. And this is in watts meters squared. Just to sort of set the scale here, and this was also noted in your questions back to us, the global annual incident solar radiation at the top of the earth's atmosphere is about 340 watts per meter squared.

So, all of these numbers that you see here are less than one percent of the incident solar at the top of the earth's atmosphere. And approximately 70 percent of that is absorbed by the climate system.

So again, that sort of just sets the scale. These numbers, and these are, perturbations to the energy budget are about one percent of incident solar.

And one of the questions that

you raised, of course is, is this enough to actually force climate change? I will come back to the issue of where these numbers come from in a moment because there is another important issue. And I think perhaps there is a little bit of, I think, a hint of a misunderstanding in some of the questions coming back to us.

This is broken out. All the numbers to the right of zero, of course, are terms where greenhouse gases have added and have reduced the amount of emission to space and enhanced the greenhouse effect of the earth's atmosphere.

And several of these deal with well-mixed greenhouse gases like CO_2 and other gases which have lifetimes in the troposphere of 100 years plus.

They are effectively very well mixed compared to the mixing time for the troposphere of about a month. And WMGHG stands for "well-mixed greenhouse gases."That includes methane, nitrous oxide, halocarbons and molecular carbons a bunch of other carbons. Those are all in red and those represent heating for the climate system.

The aerosol interactions and aerosol climate interactions, which were highly uncertain, you will note that, because of the large error bars in blue and represent, we believe, slight coolings in the climate system.

One of the reasons why the aerosol radiative interactions (and this is just the direct effect of scattering of absorption of sunlight) actually has a slight uncertainty is because of the large uncertainty of the amount of black carbon in the earth's atmosphere.

That's very hard to sense remotely from space and that has proven to be a major source of uncertainty of these estimates.

The final feature of this is the total anthropogenic, which is a summary of everything above it, has a very large error bar. This is going to come back to haunt us when we talk about the estimate of transient climate response which appeared in your notes because, I'm sorry to say, that error bar was not propagated into that calculation, and it's a large error bar. So, we will come back that to that point in a bit.

The main reason I wanted to show you this graph is to emphasize how large the anthropogenic part is of the estimate. And again, this is a model estimate. And it is an unknown to you relative to the solar radiance, which is the number immediately below. So, these differ by, well, easily over an order of magnitude.

And this is one of the reasons why the climate community, and we have records of this thanks to Galileo that are quite good because we can count sunspot number and correlate that to the sunspot of SOHO with the solar radiance back for 350 years plus.

And so, this number for the solar radiance variations over the last 500 years is -- I will show you in a moment -- there is still some uncertainty, but it is not huge.

One other thing I want to call out to your attention about this graph so that you are all aware of it is that these are model calculations. These are not measurements. In many cases, they are based on observations.

So, for example, we have very good records down to parts per million of the well-mixed greenhouse gases, et cetera. There are the radiative transfer codes which are backstopped by Maxwell's equations.

So I would argue, and we have very good evidence, that the radiative forcing by CO_2 and well-mixed greenhouse gases on this figure are quite good. But I am happy to take that point of discussion if you wish.

This is radiative forcing. And the main thing I want to call out to you is, this is the boundary condition on the climate system. Yes, the changes are small, but the one component in this that is -- well, there are two components.

The other one that is not shown is volcanic and that turns out to be even smaller than solar.

Both of those are dwarfed by our estimates of the anthropogenic. So, that is one of the reasons we think if this is a boundary condition force problem, the IPCC is quite confident that the anthropogenic component of this is the main driver.

These are various reconstructions of the total solar radiance.

I just wanted to show this [next page] to you to kind of get this off the table. These are time series where you can clearly see the solar cycle built into the oscillations. This time series runs back to the introduction of the steam engine.

But, of course, if we take it back another 400 years, thanks to Galileo, the reconstructions differ because of sort of the means by which you interpret the modern sunspot record and its relation to solar radiance in time.

There are other ways of constructing this from isotope proxies. But in any case, these numbers of uncertainty in total solar radiance are tiny.

TSI uncertainty small relative to GHG forcing (I.2)

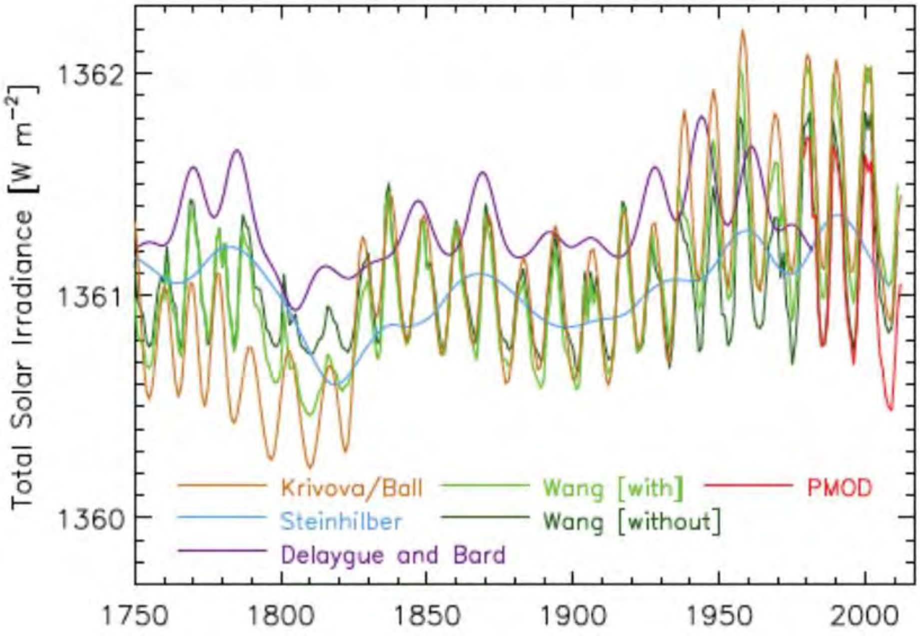


Figure 8.11

So, what is done for the projections [next page] that also form some of the topics for discuss today is that we use that information for the historical climate record and take climate models that have been brought into quasi-equilibrium so they are in equilibrium state so that they are not varying very much in time at the start of industrialization.

So, we build climate models. We assume when we construct those models that the net energy balance of the planet was identically zero or effectively zero at the start of industrialization.

We ensure that the climate models produce a steady-state climate for the millennia under those conditions, and then we begin subjecting them to the historical time series of forcing, bring them up to the present day, and then we spawn a series of model runs off the end of

Scenarios for projections branch from historical (IV)

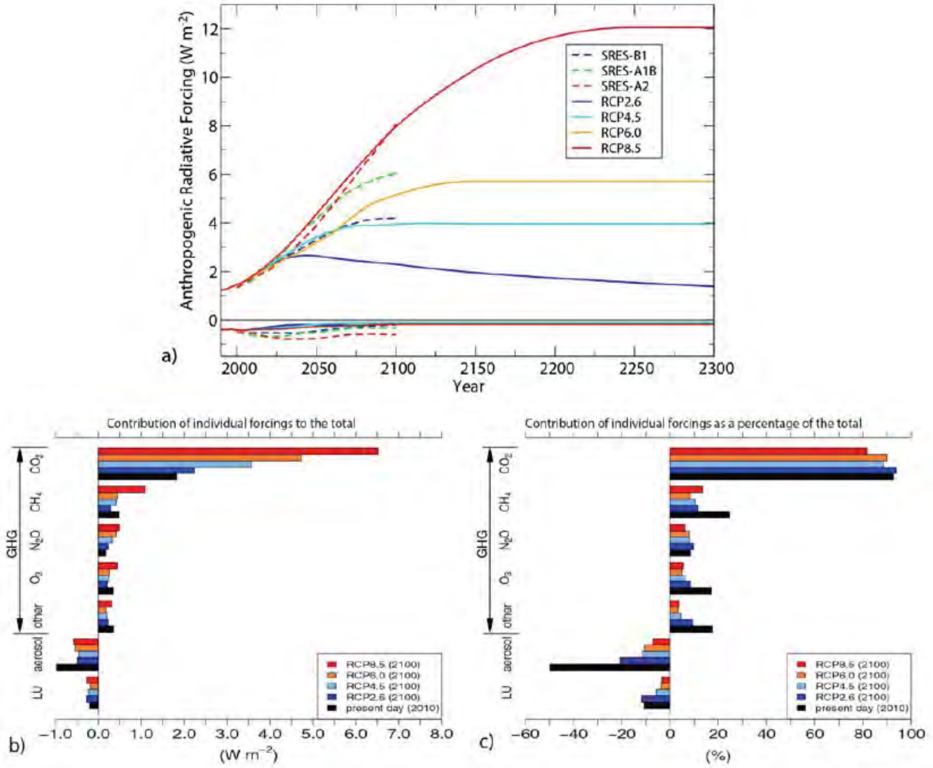


Figure 12.3

the present day.

This also came up in connection with, what was the experimental design here? I should also say that these runs are not commissioned by the IPCC.

These are actually done as a service to the IPCC, but it's done through the Working Group for Climate Modeling, which is part of the World Climate Research Program.

So actually, IPCC does not commission these runs. I will come back to this point in a minute. But they are sort of done for the IPCC. And we are trying to deal with several sources of uncertainty.

One of them is the huge uncertainties even in historical forcing. And I want to highlight to you the graph in the lower right which shows the contribution of the forcing from various components and how large the negative component of the aerosol is. It's about minus 40 percent.

This number is one of the most uncertain in this figure. And this will also come back to haunt us, I think, a little bit in terms of the interpretation of the historical record.

We have very poor information -- we have essentially no measurements of aerosol radiative forcing that go back of any utility back much further than about 40 to 50 years.

And our information regarding the concentration of aerosols in the atmosphere becomes quite problematic once you go back more than a few decades. At that point, we are literally relying on high school records. So, the aerosol number in this graph is particularly uncertain.

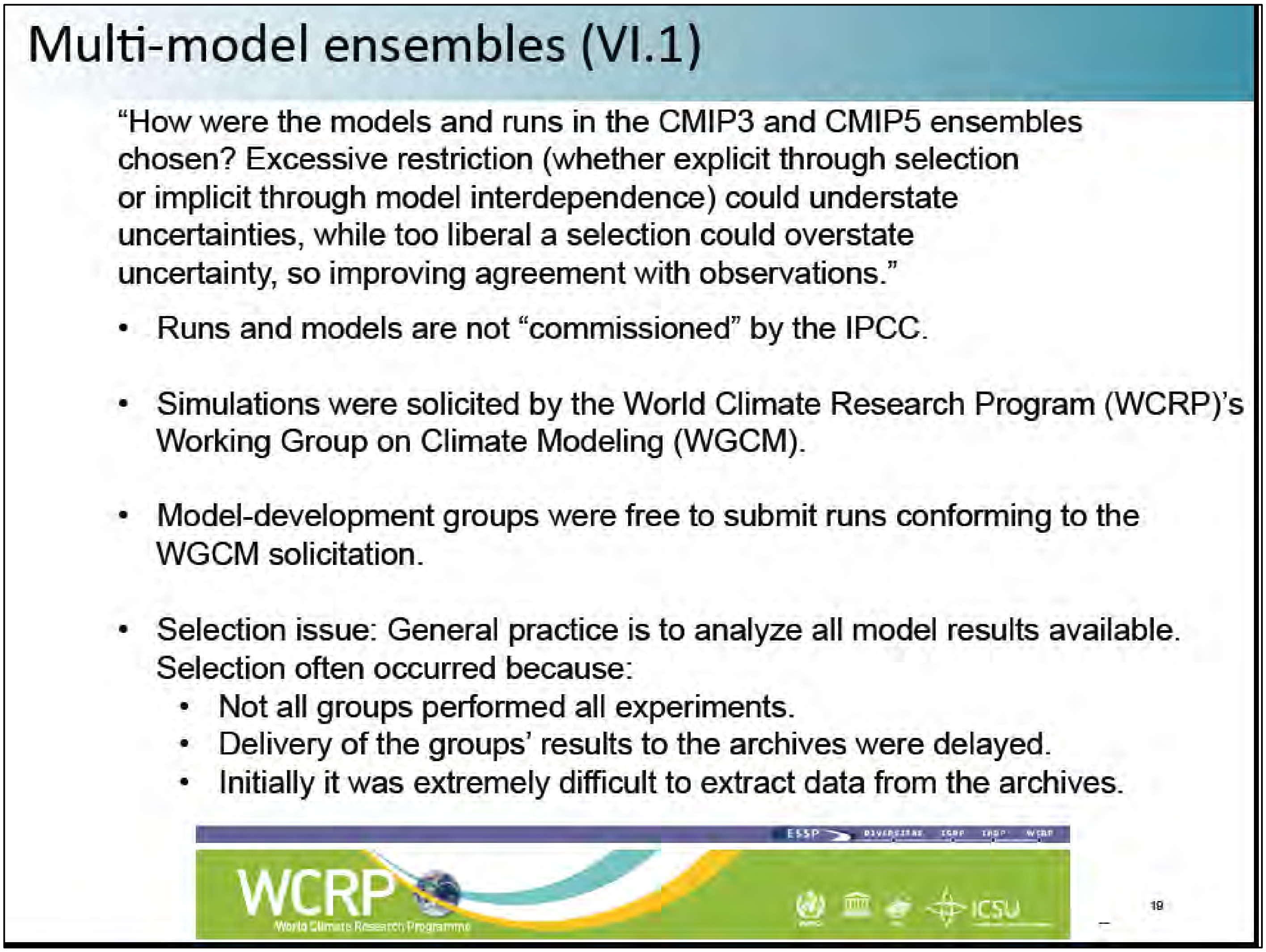
What we do is take the models up to the present day and then spawn several different runs. In this case we did four. And they are called representative concentration pathways, RCPs. And they have a number on the end. So, throughout the report, you will see RCP 8.5, 2.6, et cetera.

That 8.5 refers to the anthropogenic radiative forcing in watts per meter squared. So, that is what "RCP" means.

DR. KOONIN: At the end of some time period?

DR. COLLINS: In 2100. What we are trying to do, so, what we have, what is done is that we accumulate models from around the world [next page]. There were 45 plus, I think, that participated in the round of model comparisons that form the basis for what I am going to show you.

We do that in order to account for structural uncertainty among the climate models, because there are a



Multi-model ensembles (VI.1)

"How were the models and runs in the CMIP3 and CMIP5 ensembles chosen? Excessive restriction (whether explicit through selection or implicit through model interdependence) could understate uncertainties, while too liberal a selection could overstate uncertainty, so improving agreement with observations."

- Runs and models are not "commissioned" by the IPCC.
- Simulations were solicited by the World Climate Research Program (WCRP)'s Working Group on Climate Modeling (WGCM).
- Model-development groups were free to submit runs conforming to the WGCM solicitation.
- Selection issue: General practice is to analyze all model results available. Selection often occurred because:
  - Not all groups performed all experiments.
  - Delivery of the groups' results to the archives were delayed.
  - Initially it was extremely difficult to extract data from the archives.

number of processes in the climate system we just do not understand from basic physical principles.

For example, let me be careful how I state that exactly. We understand a lot of the physics in its basic form. We don't understand the emergent behavior that results from it. And so, a good example for that would be cumulus convection.

Well, we know, okay, it's anisotropic turbulence occurring, anisotropic because it's dealing with a buoyancy gradient. It's got an internal heat engine fluid in the form of condensation of water vapor.

So, it's nasty, it's turbulence, it's anisotropic and it has a heat engine at intervals physics across twelve orders of magnitude. So, it’s a multiphysics problem.

We account for the structural uncertainties by using 45 different climate models. Those are not selected -- essentially, it's a very democratic system. And I will come back to that point, too. That's one source of uncertainty.

The second source of uncertainty is what mankind is going to do. And we are not going to talk too much about that today. And the solutions on this graph don't really separate out until 2040 or so.

Most of the climate change between now and 2040 is committed from historical emissions, about two-thirds of the common signal.

Robert?

DR. ROSNER: So, are you saying that you accounted for model uncertainty by basically assuming that all these models were created independently, that they explore the parameters of basic possible models?

DR. COLLINS: They do not. And this has been now examined carefully by Reto Knutti and colleagues. There has been sort of a careful analysis of the clustering models as they appeared in the literature, that appeared, actually before the fifth IPCC report went to press.

So yes, there are a number of these models are certainty cousins, first or second cousins. But they are cousins for sure.

I can see from the way I am consuming time on the introduction that some of the material I have in my talk will be covered by the speakers.

In particular, I have some slides I have borrowed from Ben without his knowing, and he will show those. And parts of the talk where I am going to cover stuff by other people, I will go quickly. And Ben mentioned that he is going to talk about the hiatus.

I did want to say here [next page] that

Perspectives on the hiatus (II)

*"IPCC suggests that the stasis can be attributed in part to "internal variability."
Yet climate models imply that a 15-year stasis is very rare (von Storch et al., 2013) and models cannot reproduce the observed GMST even with the observed radiative forcing [See figure immediately below from the AR5 WG1 report]."*

Response:

- With a couple of important exceptions, there is no "observed forcing".
- Forcing in IPCC ***is calculated***, e.g. using benchmark radiative transfer models.
- Each climate model computes its own estimate of the forcing.
- Major uncertainties (aside from solar):
  - Emissions of black carbon aerosol – major warming agent
  - Higher cloud albedo from anthropogenic aerosols (Twomey effect).

just, so, the nit I wanted to pick with one statement, and I am going request to pick nits with you; I'm sorry. This is going to be a give-and-take here a little bit. So, I put your statement in quotes with regards to the hiatus. And I specifically want to point out the issue of radiative forcing.

You say, "Models cannot reproduce the observed global mean surface temperature even with the observed radiative forcing." The reason I went through this whole exercise on forcing is that it is not observed. It is calculated. And the aerosol competent of that is highly uncertain.

The models we use for the greenhouse gases, those are really good, but the aerosol component is uncertain. Dealing with uncertainty in chemistry, the microphysics of the aerosol is a mess. It's basically the physics the dirt, quite literally. So, it's messy.

And each model is computing its own radiative forcing. We do not prescribe that information. We hand them concentrations. They are asked to compute forcing there that.

And even under controlled circumstances, we can show that something like maybe a quarter, in fact, about a quarter of the response variation we see in the ensemble is just due to uncertainties in the forcing.

Even though we try to control for that, even though we claim we are handing them exactly the same climate conditions, we are handing them chemical boundary conditions and not radiative forcing boundary conditions to compute from that the radiative forcing.

And that's about a quarter of the variation we see in response across the model ensemble. Major uncertainties associated with black carbon and particularly with cloud-radiative interactions for reasons I will be happy to come back to. It's called the Twomey effect.

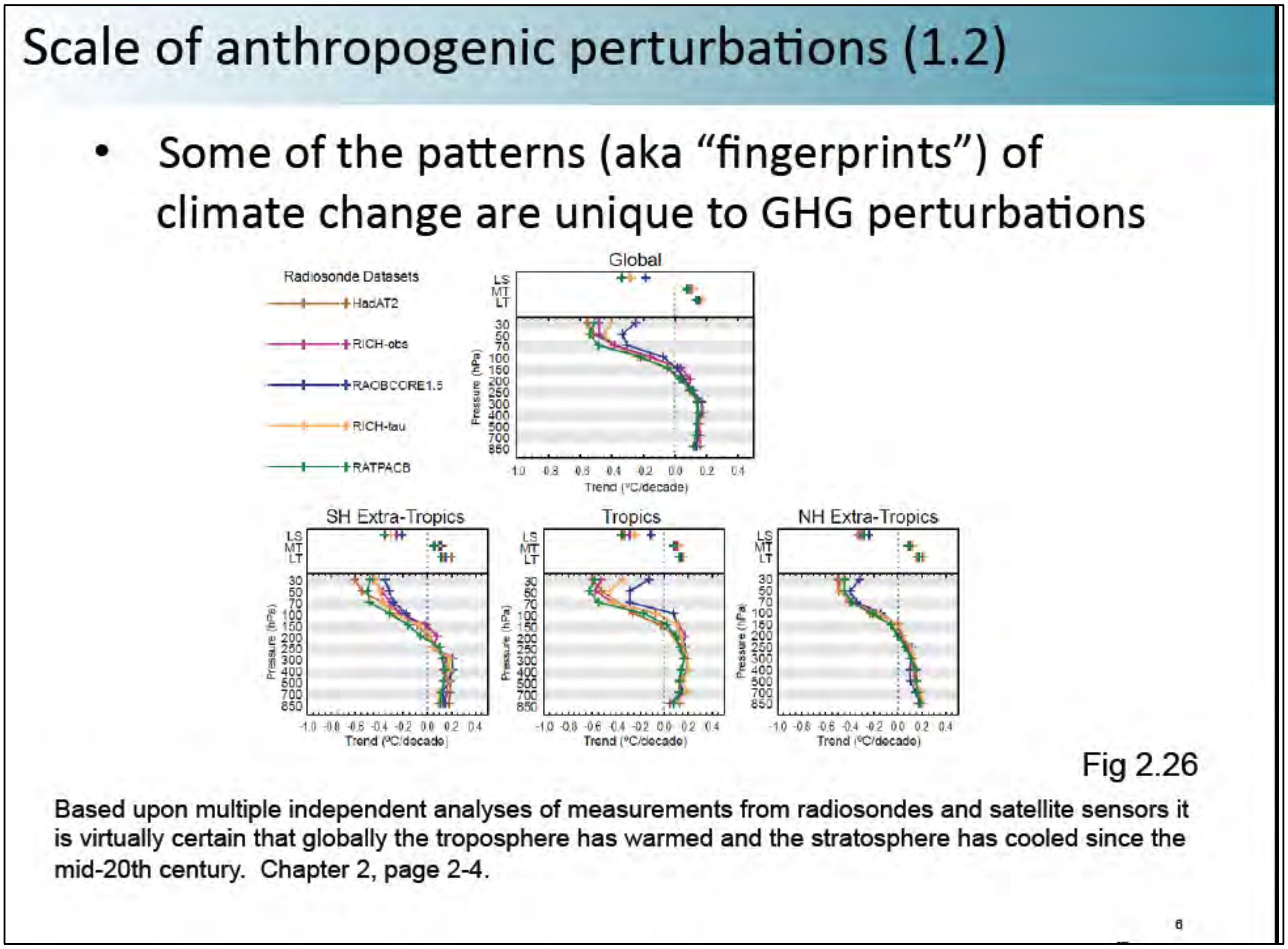
So, one of the issues that you raised because of the tininess of this perturbation of the boundary conditions is [see slide], how can you be sure, given the fact that uncertainties and fluxes in the climate system are quite large and these perturbations and boundary conditions are small, how can you be sure that, when you look at a field like temperature which has a lot of stuff, a lot of different processes that contribute to its variations, how can you be sure that you are correctly interpreting the influences?

This is also a drawing on work that Ben and his colleagues pioneered. But you can use gradients in the temperatures as a clue and a means to get through that thicket of different influences on the temperature.

And one of the most powerful tools -- and a number of people in this room have contributed to this literature John, Ben, others -- has to do with the vertical gradients in temperature in the earth's atmosphere.[next page]

And one of the particularly strong fingerprints for global warming is a dipole, warming of the troposphere, cooling of the stratosphere due to the physics of the radiative transfer and the interactions between the two and the effect of carbon dioxide on the stratosphere.

The reason I am showing this to you is that this is also a tiny signal. I don't know if you can read these numbers, but they are tenths of



Scale of anthropogenic perturbations (1.2)

- Some of the patterns (aka "fingerprints") of climate change are unique to GHG perturbations

Fig 2.26

Based upon multiple independent analyses of measurements from radiosondes and satellite sensors it is virtually certain that globally the troposphere has warmed and the stratosphere has cooled since the mid-20th century. Chapter 2, page 2-4.

a degree Celsius per decade. So, these are small numbers, but this dipole is a very robust pattern.

And more interestingly, and this is the insight Ben had almost 20 years ago, 25 years ago, this pattern is very hard to get from sources of climate change other than well-mixed greenhouse gases.

So, I won't go through this graph. We can come back to it at the end in discussion. But one can compute, for example, what would happen if the sun increased its luminosity.

The earth's atmosphere includes a lot of gases that are quite effective absorbers of infrared in which he finds that, when you increase luminosity, you heat the whole column.

You don't heat just the troposphere and the close the stratosphere. You heat everything because, of course, the ozone in the stratosphere is also an excellent absorber.

Volcanos have sort of the same effect, although in the opposite direction. So essentially, we don't know the physical mechanism for getting this dipole out of solar variations, volcanic variations. The only way that we can explain it is with well-mixed greenhouse gases.

DR. KOONIN: And again, the dipole is focused on the warming troposphere?

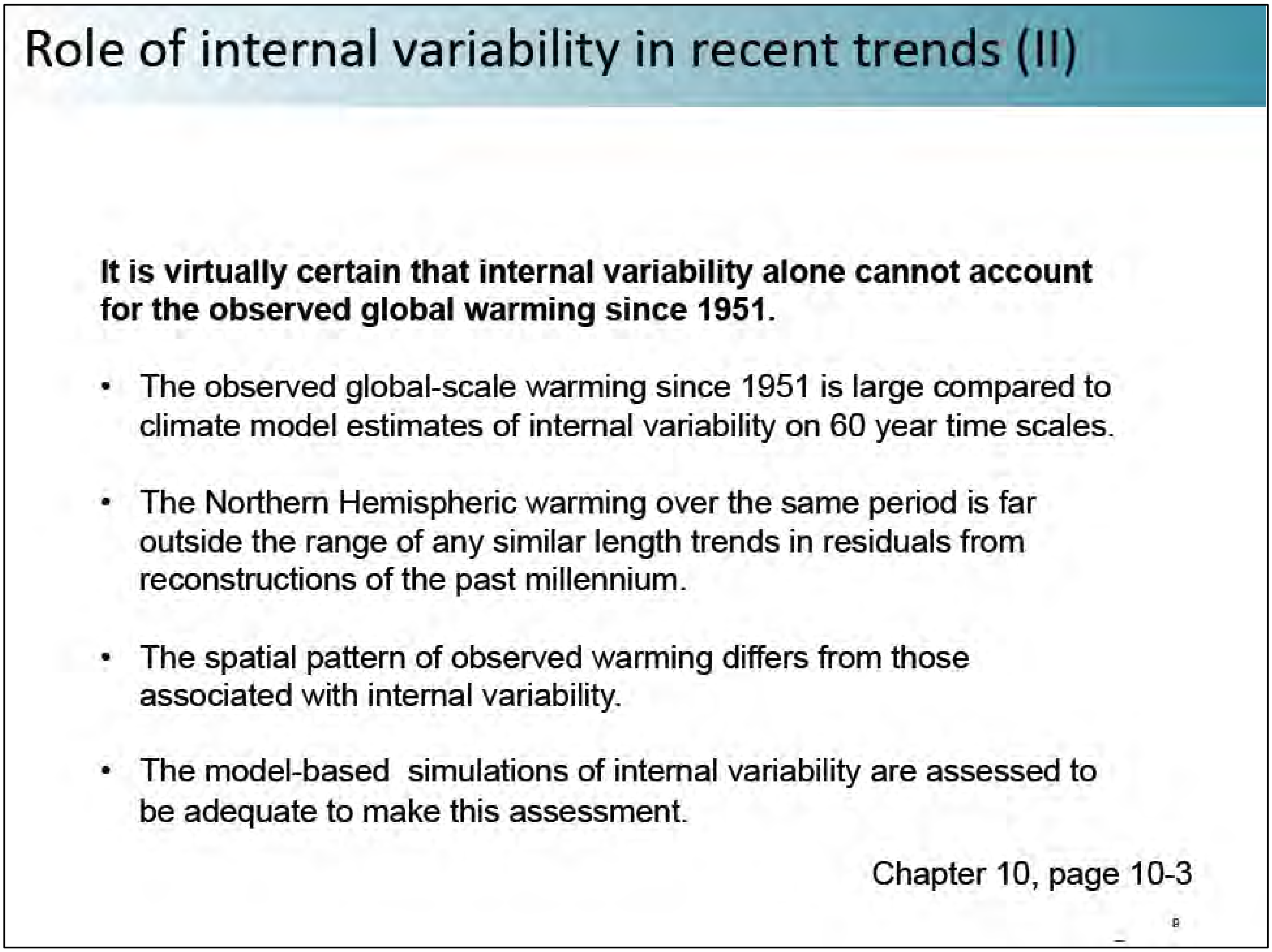
DR. COLLINS: That's correct. That's right. This is one of the reasons why this tiny gradient is actually a very big fingerprint for climate change and one of the reasons we think they can sort this problem out by looking at a signal of well-mixed greenhouse gases that is essentially, we don't know of a mechanism for getting it from natural sources, including natural variability, by the way.

So, natural variability, for example, if you were to heat the ocean's surface, because of some internal mode, it would not produce the signal that we are seeing with this dipole. So, that's another reason we are having these boundary conditions.

So, I will skip this (indicating slide).

So, this [next page] is the reason why there are such strong statements in the report that, "It is virtually certain that internal variability alone," because just heating the ocean alone will not produce this dipole, "cannot account for the observed warming since 1951."

There are some other reasons why this warming is large compared to climate model estimates, internal variability. And I will come back to



Role of internal variability in recent trends (II)

It is virtually certain that internal variability alone cannot account for the observed global warming since 1951.

- The observed global-scale warming since 1951 is large compared to climate model estimates of internal variability on 60 year time scales.
- The Northern Hemispheric warming over the same period is far outside the range of any similar length trends in residuals from reconstructions of the past millennium.
- The spatial pattern of observed warming differs from those associated with internal variability.
- The model-based simulations of internal variability are assessed to be adequate to make this assessment.

Chapter 10, page 10-3

this point in a minute. We are dealing in the climate system with a difficult system. We are looking at an integrated -- we have one instantiation of it.

So, we don't have a parallel where we can go run experiments, although I have had some interesting discussions with people about using Mars for this purpose. But at the moment, we are limited to just Earth and we have to sort of take the omelet we have and unscramble it.

We do use models for that. And we should talk about whether or not the models are a suitable tool for unscrambling. That is an issue.

Statements like this are predicated on the idea that we can look at, we can assess the internal variability to the climate system, essentially setting variations in the boundary conditions aside.

So, we can sort of explore how

big the natural variability in the climate system will be in all its different modes while holding the boundary conditions fixed, and use that essentially as our means of driving signal-to-noise statements that we make throughout this report.

And so, one of the key questions, I think, is, are the models doing a decent job in reproducing internal variability?

By "internal variability," I just mean the behavior of a dynamical system to explore limit cycles if you let it loose. That's what we are talking about.

Now, the reason why this is a tough problem for us, this internal variability, is that some of the modes of it are quite long. They have long periods of 60 to 100 years.

We have an inadequate record with which to constrain the climate models sufficiently to make sure we have exactly the right initial conditions.

And there are portions of the climate system that have long memory. So, land surface moisture has memory scales of 300 years. The ocean turnover time is about 3,000 years. We have grossly inadequate observations of the salinity and dynamical structure of the ocean that makes it very difficult for us to nail down the initial conditions.

So, there is some discussion in your notes about, well, why is this such a difficult issue? Internal variability is an internal mode.

It's a coupled oscillation of the climate system. That's not mysterious. What is hard to us to nail down is the initial conditions, amplitude and phase of these things when we put up our climate model runs.

DR. KOONIN: So, some people

explore “if I initialize the ocean in a different way –“

DR. COLLINS: Yes.

DR. KOONIN: They do?

DR. COLLINS: We have certainly done so. They do that. And, in fact, the ensembles, one of the other dimensions in this data set that we produced are perturbed initial condition ensembles.

So, on top of all the multiplicity, multiple scenarios, multiple models, each model is typically initialized with five to ten different initial conditions and then run forward in time so that we can average out the effects of uncertainty in the initial conditions. And so, that is explored.

And the ocean, that is typically done in a separate mode. But yes, that has also been explored and it's not a major driver for the results I am going to show you.

Anyway, the IPCC concluded, "Internal variability is unlikely the explanation." And I will show you a little bit more of that in a moment. This is actually from a figure by Jones that was then quoted in the IPCC report. [next page]

And you will get a copy of this in my presentation along with the source citation and the notes that went along with it.

But the top line of this figure shows temperature change over various periods of time from a temperature reconstruction. And what you will notice is that, if you take a run called historical in the middle which is the next row down.

And then these are model runs, the second row from the top, and apply to it our best knowledge of radiative forcing. You will qualitatively reproduce those

Patterns of SST change

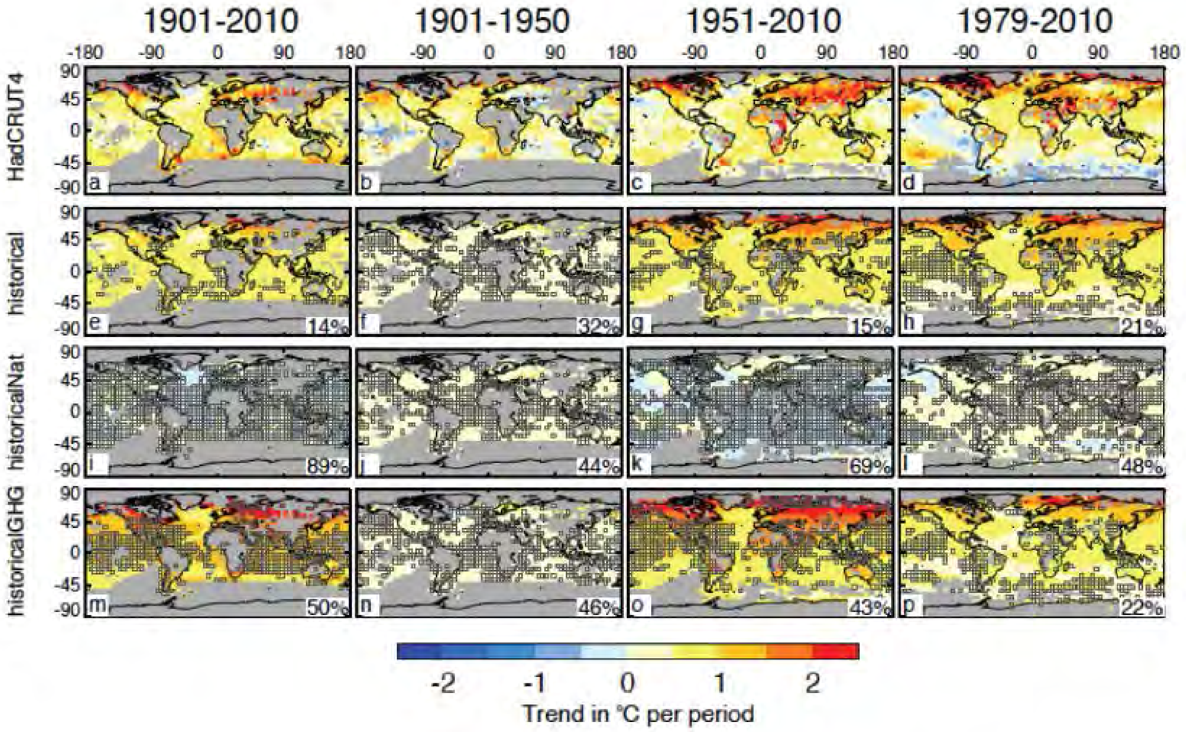


Figure 10-2

patterns.

Places without hashing are places where the data and the model are in agreement. If you go down to the third row from the top and just apply reconstruction of natural forcings from volcanos and from the sun, you notice that most of the figures are hashed and we cannot reproduce, according to our models, we cannot reproduce the historical record.

Again, let's be very clear. These statements hinge on the fidelity of models. That's the reason why I included this statement. We did look at this issue. How badly would the models have to be wrong for these statements of attribution to be blown?

And chapter 10, which deals with detection and attribution concluded that we have to be underestimating the variability by a factor of three. And we don't see any evidence for that. Let me show you one of the sources for this statement.

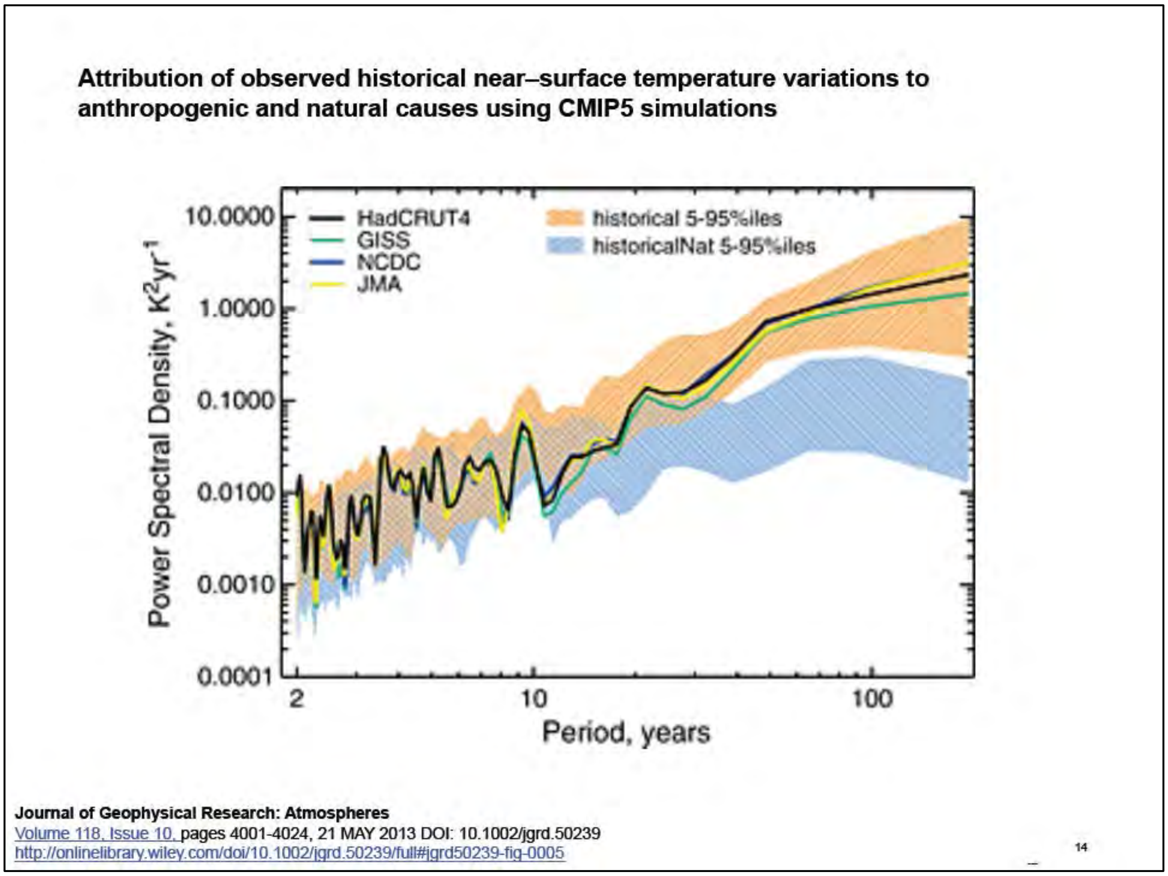
So, the lower graph is the one I wanted to focus on. [next page] It's CMIP5. It's a measure of standard deviation in temperature, standard deviation in time. And the observations are plotted on top of model results.

And you can see that there is no -- and we will go through exact error analysis here -- but the evidence shows on the basis of this graph, at least, the observations and the models lie on top of each other in terms of their estimates of temporal variation in temperature.

DR. KOONIN: Bill, we want to finish in five minutes.

DR. COLLINS: Yes, I know. I am going to accelerate.

This [next page] is also another way of looking at the same problems. This



Attribution of observed historical near–surface temperature variations to anthropogenic and natural causes using CMIP5 simulations

Journal of Geophysical Research: Atmospheres
Volume 118, Issue 10, pages 4001-4024, 21 MAY 2013 DOI: 10.1002/jgrd.50239
http://onlinelibrary.wiley.com/doi/10.1002/jgrd.50239/full#jgrd50239-fig-0005

is spectral power difference, now on a much longer time scale. This is over a century. And the data that is shown in black, green, blue and yellow are observational estimates of that power spectrum.

The orange on this figure are the historical reconstructions of models which overlay the model estimates. And it you take out the variations in the boundary conditions, you get the light blue period.

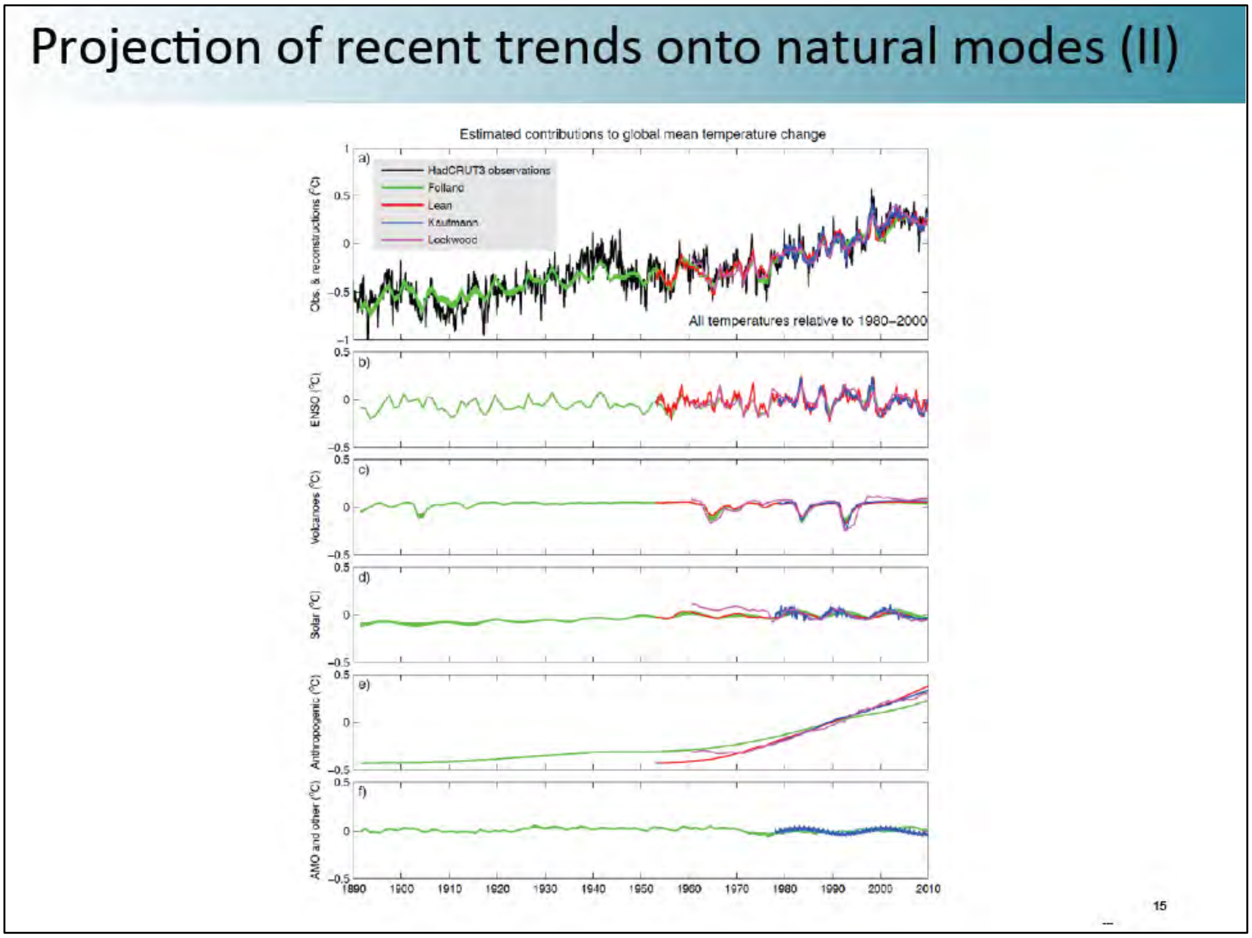
And what you find is you start really misrepresenting or underestimating the power in the climate system once you get out beyond about 20 years, 20 to 30 years.

There is a real departure between a run with and without anthropogenic influences, especially in longer time periods. Again, this is the evidence that we think we are getting variability about right.

And this [next page] is the projection of the natural modes. I won't go through this. But this is, again, evidence, what this graph shows, and I will be happy to talk about this in questions, we don't think there is a lot of power associated with these longer modes. We did leave open that question to IPCC.

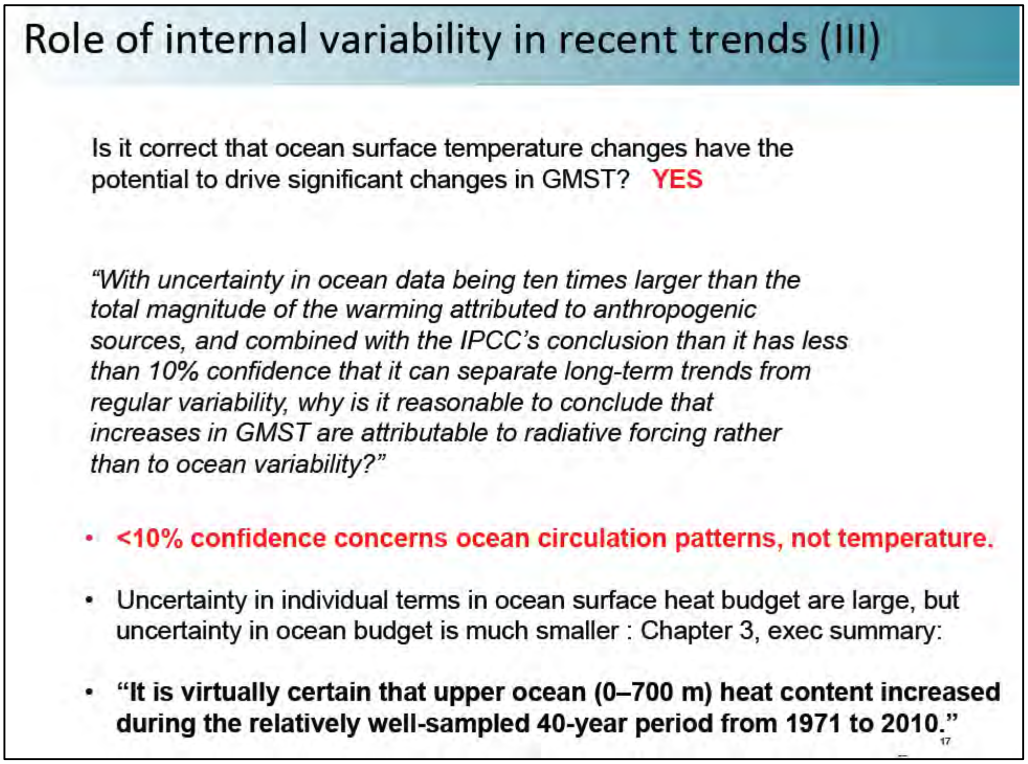
Okay, there are a few nits I want to pick here. So, one [next page] of them, you were looking at a chapter dealing with the ocean and said well, look, we only have 10-percent confidence we separate long-term trends from regular variability.

You are looking at a section of the report that dealt with ocean dynamics and not with temperature. So that, I think, was a point of perhaps slight misreading of the IPCC report on the part of people who put together those notes.



Projection of recent trends onto natural modes (II)

Estimated contributions to global mean temperature change



Role of internal variability in recent trends (III)

Is it correct that ocean surface temperature changes have the potential to drive significant changes in GMST? **YES**

"With uncertainty in ocean data being ten times larger than the total magnitude of the warming attributed to anthropogenic sources, and combined with the IPCC's conclusion than it has less than 10% confidence that it can separate long-term trends from regular variability, why is it reasonable to conclude that increases in GMST are attributable to radiative forcing rather than to ocean variability?"

- **<10% confidence concerns ocean circulation patterns, not temperature.**
- Uncertainty in individual terms in ocean surface heat budget are large, but uncertainty in ocean budget is much smaller : Chapter 3, exec summary:
- **"It is virtually certain that upper ocean (0–700 m) heat content increased during the relatively well-sampled 40-year period from 1971 to 2010."**

We are virtually certain that the ocean heat content has increased. This is a graph [next page] showing you in tens of zettajoules the ocean heat content. The heat content for the upper two kilometers of the ocean is shown in red with error bars and the coverage of the globe is shown in that light blue in the bottom of the graph.

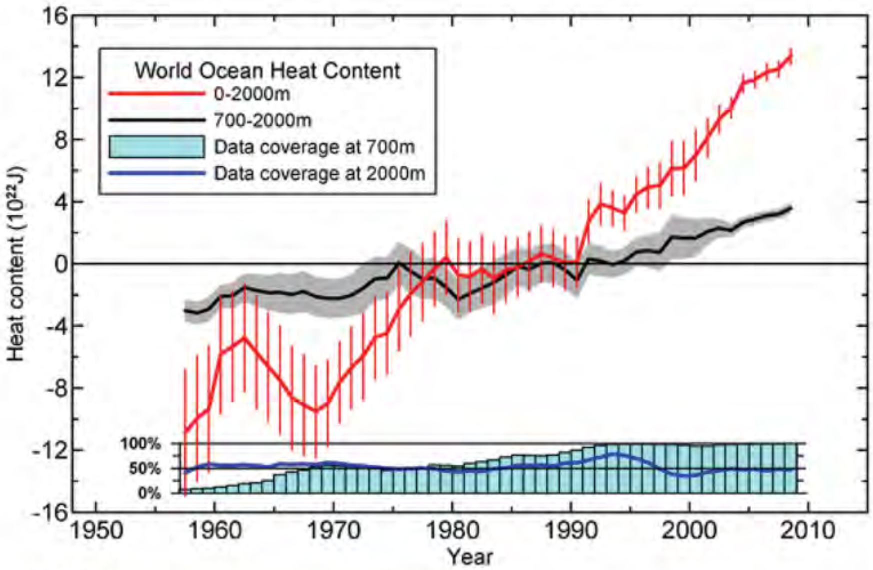
We have very good coverage. There is a reason why those error bars come down so sharply by the year 2010. And we are quite confident that the ocean heat content has been increasing since the start of this record 50 years ago.

I am not going to have time to talk about model ensembles. I will be happy to come back to this in question. So, I have some discussion here about how we constructed these, how we dealt with model means.

I do want [next page] to point out that

Recent evolution of ocean heat content (III)

World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010



Geophysical Research Letters
Volume 39, Issue 10, L10603, 17 MAY 2012 DOI: 10.1029/2012GL051106
http://onlinelibrary.wiley.com/doi/10.1029/2012GL051106/full#grl29030-fig-0001

IPCC guidance on weighting models (VI.1)

- "No general all-purpose metric (either single or multiparameter) has been found that unambiguously identifies a 'best' model; multiple studies have shown that different metrics produce different rankings of models."
- "There are however few instances of diagnostics and performance metrics in the literature where the large intermodel variations in the past are well correlated with comparably large intermodel variations in the model projections."
- "To date a set of diagnostics and performance metrics that can strongly reduce uncertainties in global climate sensitivity has yet to be identified."

Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections

you asked whether or not we weight these things. We don't. So, we looked at this issue in detail two years, three years before the report came -- two years before the report came out and decided essentially not to weight the models.

Now, there are graphs are you pointed out where some of the models are not included. But we typically did not weight them.

And what we found, in fact, is that, for reasons that are still under investigation, averaging across the ensemble, including all the structural uncertainties, seems to have compensating errors that cancel.

So, the multimodel average actually does better than any single member of the realization. That is what is shown in these figures [next page] from Peter Gleckler at Lawrence Livermore National Laboratory.

I have a couple more minutes,

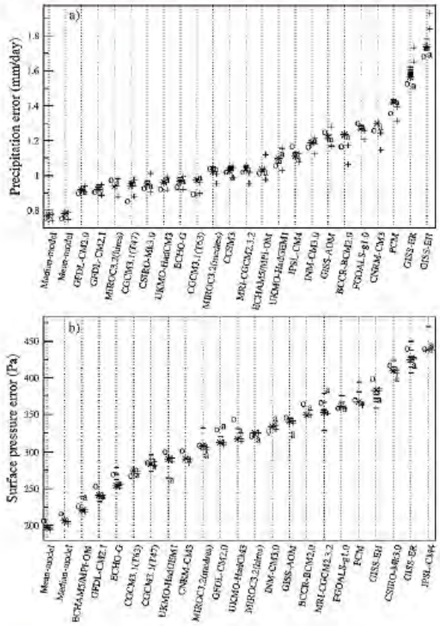
IPCC guidance on weighting models [cont.] (VI.1)

Using a variety of performance metrics, a number of studies have shown that **a multi-model average often out-performs any individual model compared to observations.**

This has been demonstrated for
- mean climate …
- detection and attribution… and
- statistics of variability

Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections

Performance metrics for climate models



Journal of Geophysical Research: Atmospheres
Volume 113, Issue D6, D06104, 20 MAR 2008 DOI: 10.1029/2007JD008972
http://onlinelibrary.wiley.com/doi/10.1029/2007JD008972/full#jgrd14137-fig-0006

and so I wanted to also point out to you since you raised the issue what metrics you used to assess improvement. And this is a figure from my chapter.

So, what is shown here [next page] are metrics for mean state trends, variability, extremes. I will of happy to come back to this in questions. Each of these acronyms is backstopped by a section in this chapter.

What this is showing you essentially is that the improvements between the last ensemble and the current one is a little bit of a mixed bag. There are many instances in the orange color where there is essentially no improvement.

And green is where we think there was some improvement and in some cases, that improvement was quite modest. But these statements are backstopped qualitatively in

Metrics used to assess improvement (VI.1)

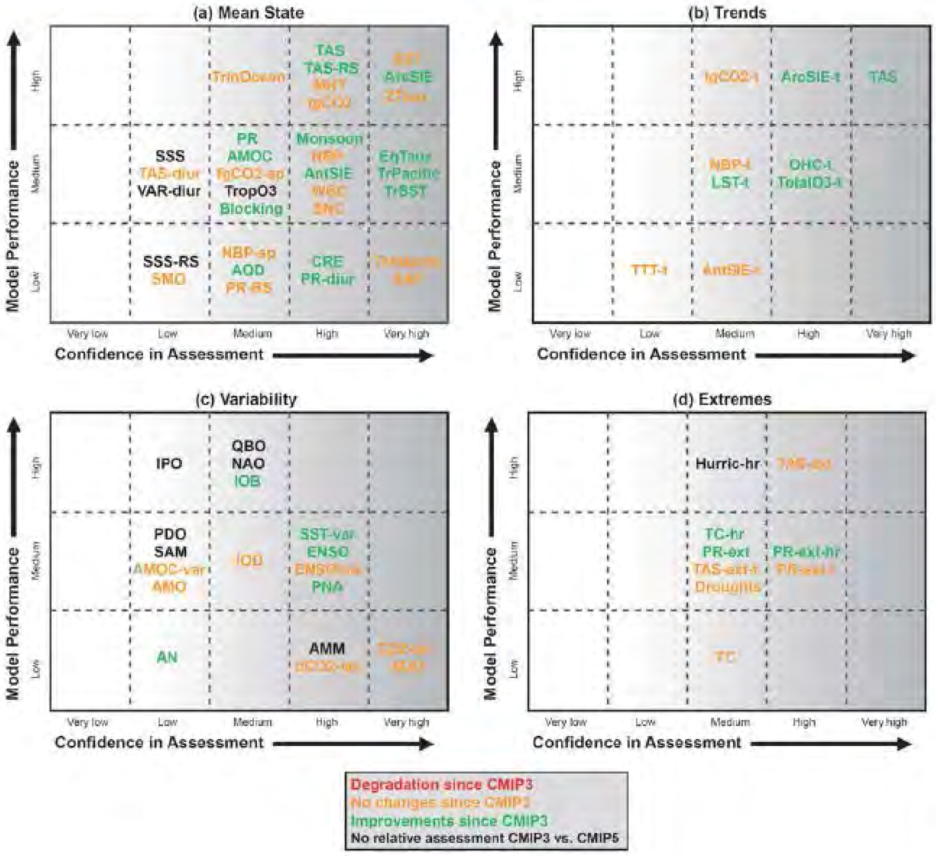


Figure 9.44

chapter 9. I will be happy to come back to this point in question.

And further evidence that the ensemble will improve in time, this [next page] is actually from a paper by Reto Knutti showing errors in precipitation between this ensemble two generations ago and the current one, showing how the mean and the range has been collapsing with time.

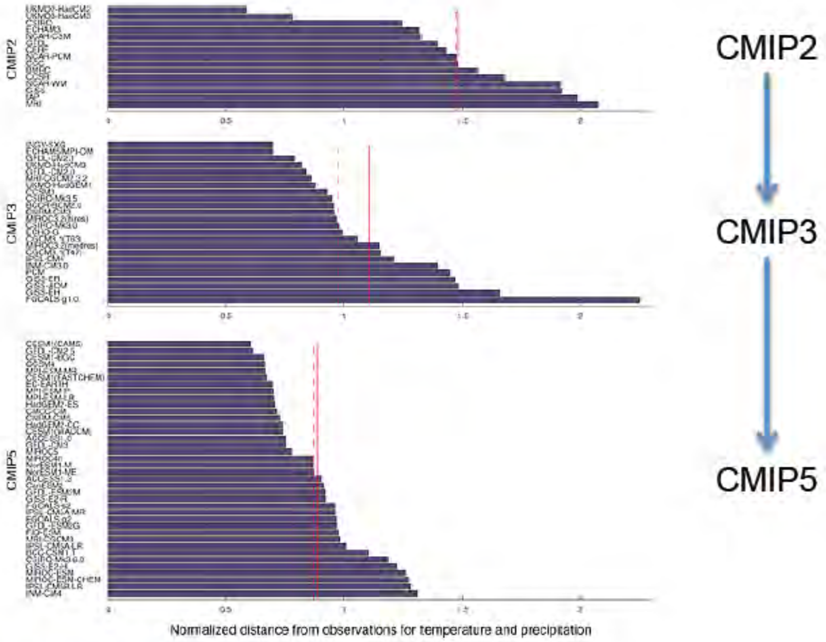
These are errors in precipitation and temperature, so, direct evidence that the model ensemble has been improving.

I will skip this because Ben is going to cover it (indicating slide).

I am going to use my last 40 seconds wisely. So [next page], one of the statements that was in your notes was, "Please comment on the cause and significance of these model overestimates of equilibrium sensitivity, particularly for projections of future anthropogenic

Metrics used to assess improvement [cont.] (VI.1)

Climate model genealogy: Generation CMIP5 and how we got there



Geophysical Research Letters
Volume 40, Issue 6, pages 1194-1199, 26 MAR 2013 DOI: 10.1002/grl.50256
http://onlinelibrary.wiley.com/doi/10.1002/grl.50256/full#grl50256-fig-0003

Equilibrium climate sensitivity (VI.2)

"Please comment on the cause and significance of these model overestimates of ECS, particularly for projections of future anthropogenic impacts"

To be frank, the characterization of the clustering of ECS in the upper portion of the stated 1.5 to 4.5C range as an "overestimate" is a misreading of the quoted text, which the chapter 9 coordinating lead author and I wrote.

As Figure 12.2 demonstrates, the range of ECS is consistent with the climatological constraints and with ~50% of the instrumental ranges.

Basic theory shows that given reasonable estimates of uncertainty in feedbacks that the distribution of ECS should be asymmetric and fat tailed towards high values.

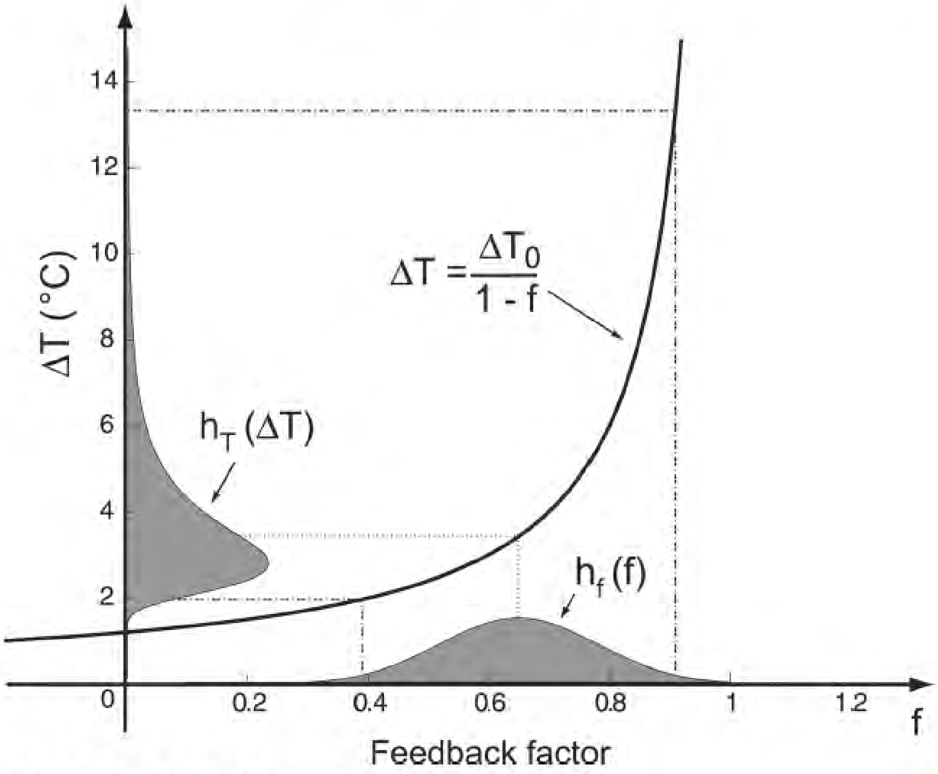
impacts."

I have to say because the statement you were quoting from the IPCC report is a sentence that I wrote that I think that, I'm sorry to say, but I think you may have misquoted that text, at least misread it, because as we demonstrated in a figure that appeared in chapter 12, the range of equilibrium climate sensitivity is consistent with climate constraints and about 50 percent of the instrumental ranges.

So, we didn't see there as being a problem. And there is plenty of evidence in the literature.

This [next page] is from a paper by Roe and Baker that, "The distribution of climate sensitivity has to be fat-tailed toward the high end. This is an intrinsic feature of the math of the feedbacks that are part of the equilibrium climate system

Fig. 1. Demonstration of the relationships linking hT(ΔT) to hf(f). ΔT0 is the sensitivity in the absence of feedbacks.



G H Roe, and M B Baker Science 2007;318:629-632

Published by AAAS

calculation."

So [next page], you are going to tend to see models sort of turning toward this fat-tail to a high climate sensitivity seems to be just part of the math. You can't avoid it.

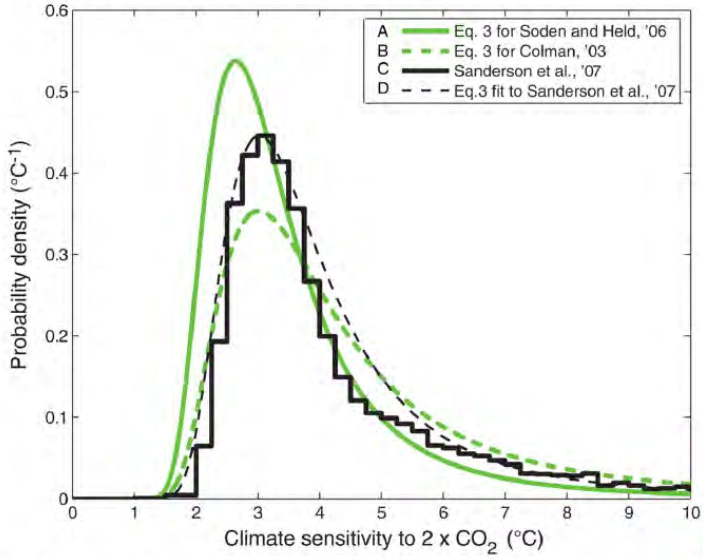
The last thing I want to end on, and this [next page] is my last slide, is, and I think a number of other people will go to bat on this particular issue. I am just going to deal with the issue, a very simple one.

There are more sophisticated analyses you will hear later in the day about error propagation.

So, one of the interesting footnotes in your notes was dealing with your estimate of transient climate response, which is how much climate response, say, of doubling of carbon dioxide, say, 70 years just in relation.

This is not the asymptote. This is the intermediate response to

Fig. 3. Climate sensitivity distributions: (A) from (18), which calculated (f, σf) of (0.62, 0.13) from a suite of GCM simulations; (B) from (17), which found (f, σf) of (0.7, 0.14) from a different suite of models; and (C) from the ~5700-member multiensemble climateprediction.net (9, 10) for different choices of cloud processes. [Data were provided courtesy of B. M.



G H Roe, and M B Baker Science 2007;318:629-632

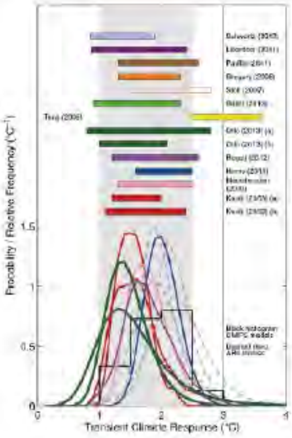
Published by AAAS

Science AAAS

Transient climate sensitivity (VI.2)

As the observational value of TCR is simply estimated to be approximately 1.3 C, it appears that the models overestimate this crucial climate parameter by almost 50%.

"From 1950 to 2011, GMST rose by 0.6C, while Figure SPM.4 shows that total anthropogenic forcing rose by 1.7 W/m2 over this same period. Since the forcing corresponding to doubling CO2 is 3.7 W/m2, the TCR is easily estimated to be 0.6C X (3.7/1.7) = 1.3 C. This value is in accord with the more sophisticated observational values shown in Box 12.2, Figure 2."



Response: Forcing is ***much*** less certain than claimed.

Forcing in 1950 is 0.57 ± 0.28 W/m^2.
Forcing in 2011 is 2.29 (-1.16 + 1.04) W/m^2.
Resulting estimate would be 1.7 W/m^2 ± 1.07 W/m^2.

the climate system to being forced with greenhouse gases. And we typically use a number for what would happen if you doubled carbon dioxide.

And you posed the question or at least raised the issue it appears the models are overestimating this by about 50 percent relative to the back-of-the envelope calculation that you have in your notes, which I quoted here.

You used the central estimate of the forcing. This is one of the issues with this estimate. There will be others. But you used the central estimate of the forcing for that calculation. And the forcing, as I pointed out, is much less certain than claimed.

So, I am quoting you here, now, the numbers from the report. You looked at the difference between 2010 and 1950 and said, oh look, it's 1.7. What if, in fact, it's 1.7 plus or minus one watt per meter squared?

If you do the most naïve propagation of that, through the calculation you will find that you get, the range of this explodes toward the high end.

And, in fact, there is no, as far as we can tell, no issue with the time and transient climate response, at least based on this using this kind of back-of-envelope calculation.

So, this is one of those places where I think we greatly appreciate all the attention that you paid to the report. You clearly read it very carefully, disturbingly carefully.

This is one of the places where I think this simple addition to your calculation would, I think, would help improve the interpretation of the results.

DR. KOONIN: Will these uncertainties in the forcings get propagated into the projections for the next several decades or into the projection after 2100?

DR. COLLINS: They are sort of inadvertently in the following sense. We have looked carefully at the range of aerosol radiative forcing, which is the major driver for the uncertainty in the present day in the climate models.

And it's actually larger than one. I think it is one watt per meter squared that is quoted in here.

So, the models are started from actually a quite diverse set of estimates for the aerosol radiative forcing under present-day conditions relative to preindustrial.

So yes, in some sense, it was propagated, although I have to say, sort of unintentionally, but it has been propagated into the ensemble.

With that, let me conclude and see what questions you have for me. Thank you very much.

DR. KOONIN: Thanks for getting us off to a good start.

DR. COLLINS: Good, Thank you.

DR. KOONIN: The floor is open. Subcommittee first. Go ahead, Bob.

DR. ROSNER: So, I guess I don't understand the issue of the weighting.

DR. COLLINS: Yes.

DR. ROSNER: Or not weighting.

DR. COLLINS: Yes.

DR. ROSNER: Clearly, the models, I have seen graphs that show the various capabilities, claimed capabilities of the models, and they are remarkably diverse.

DR. COLLINS: Yes.

DR. ROSNER: And having uniform weighting seems, to me, surprising, to say the least. So, I just don't get it.

DR. COLLINS: One of the key questions, I think, that the subtext, I think, for your question is how contingent are findings, well, first off, would weighting be a sensible idea?

I am going to answer a second question which you implicitly asked first. How robust are our results to the presence or absence of weighting? In other words, that's one way of putting it.

DR. ROSNER: Let me add something to it. I asked earlier about the models because obviously, there are two kinds of errors, right, the errors with the data that you spoke about and then the errors having to do with model uncertainties.

DR. COLLINS: That's correct.

DR. ROSNER: And to me it's completely unclear which dominate, especially if you don't have really good estimates for what the model errors would be.

DR. COLLINS: That's right.

Ben, may I just check, Isaac, and I will come back to you in a second. Could I just check with you, are you going to discuss your PNAS paper showing the robustness of the water vapor attribution to scrambling model error and the ranking of the models? Is that something you are going to show?

DR. SANTER: I suspect I am going to run into the same difficulty that you did. So, I do have it in my talk, but it's right at the end. So, if I don't cover the detection and attribution and the hiatus, I won't get to it.

DR. KOONIN: We will make sure to ask about it.

DR. COLLINS: Robert, before I take Isaac's point, one of the figures I had to rush over because Ben is the author and I defer to him, we have an example of attributing change in atmospheric moisture.

And the question is what if you rank the models or weighted them or used a subset of them, depending on which fidelity to which metric, how robust are the results to that?

And you can show actually through a careful common study that the results are remarkably robust regardless of how you rank the models according to whatever weighting scheme you want. And Ben explored several.

So, this is an example where the detection and attribution of anthropogenic signal is remarkably insensitive to how one precisely weights the models, which I would regard as a confidence-building measure because that weighting is highly subjective.

Let me come back to Ike's question.

DR. KOONIN: Ike doesn't get to talk yet!

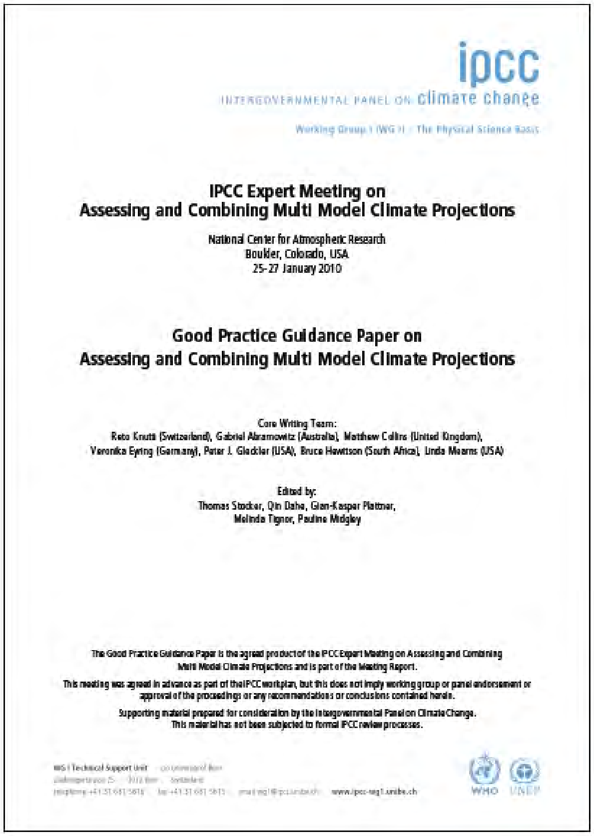
So, I have a question. What is the gateway for getting included in CMIP5 ensemble? If I write a model, it would probably be pretty lousy and pretty simple, if I could do it at all. Can I get included? Who decides what gets included?

DR. COLLINS: So, you have to meet some experimental protocols. But there is a statement. One of the statements in this good guidance document [next page] is that there is -- so, I will be honest with you. It sort of shocked me.

One of the statements in this good guidance document, and you can find it yourselves, so I am just going to quote it to you, "There is no minimum fidelity requirement for inclusion in the ensemble."

DR. KOONIN: So, how was the ensemble, in fact, constructed? Is it just everybody who came forward and said "I have got a model," or was

IPCC Dealt Extensively with Ensemble Means (VI.1)



ipcc
INTERGOVERNMENTAL PANEL ON climate change

Working Group I (WG I) / The Physical Science Basis

IPCC Expert Meeting on
Assessing and Combining Multi Model Climate Projections

National Center for Atmospheric Research
Boulder, Colorado, USA
25-27 January 2010

Good Practice Guidance Paper on
Assessing and Combining Multi Model Climate Projections

Core Writing Team:
Reto Knutti (Switzerland), Gabriel Abramowitz (Australia), Matthew Collins (United Kingdom),
Veronika Eyring (Germany), Peter J. Gleckler (USA), Bruce Hewitson (South Africa), Linda Mearns (USA)

Edited by:
Thomas Stocker, Qin Dahe, Gian-Kasper Plattner,
Melinda Tignor, Pauline Midgley

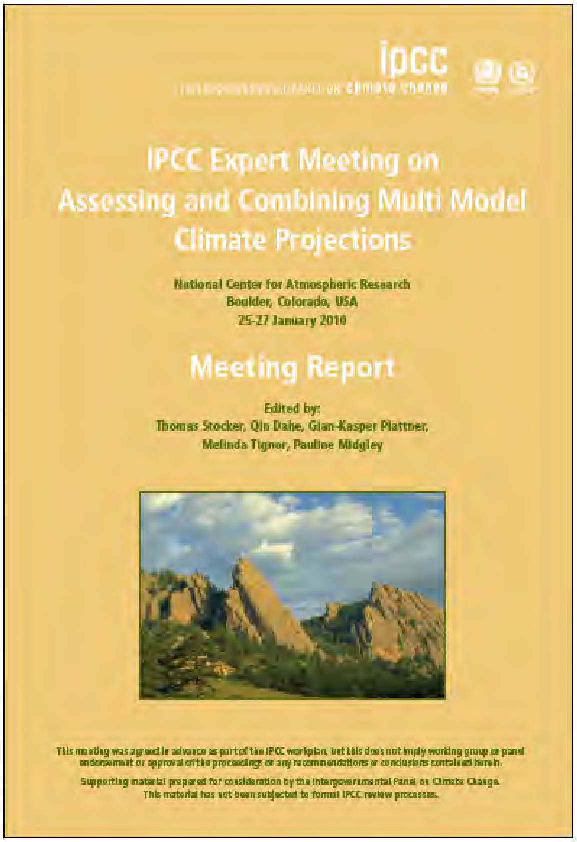
The Good Practice Guidance Paper is the agreed product of the IPCC Expert Meeting on Assessing and Combining Multi Model Climate Projections and is part of the Meeting Report.

This meeting was agreed in advance as part of the IPCC workplan, but this does not imply working group or panel endorsement or approval of the proceedings or any recommendations or conclusions contained herein.

Supporting material prepared for consideration by the Intergovernmental Panel on Climate Change.
This material has not been subjected to formal IPCC review processes.

WG I Technical Support Unit
www.ipcc-wg1.unibe.ch

WMO UNEP



ipcc
INTERGOVERNMENTAL PANEL ON climate change

IPCC Expert Meeting on
Assessing and Combining Multi Model
Climate Projections

National Center for Atmospheric Research
Boulder, Colorado, USA
25-27 January 2010

Meeting Report

Edited by:
Thomas Stocker, Qin Dahe, Gian-Kasper Plattner,
Melinda Tignor, Pauline Midgley

This meeting was agreed in advance as part of the IPCC workplan, but this does not imply working group or panel endorsement or approval of the proceedings or any recommendations or conclusions contained herein.

Supporting material prepared for consideration by the Intergovernmental Panel on Climate Change.
This material has not been subjected to formal IPCC review processes.

there some hurdle that you had to get over?

DR. COLLINS: Well, the models, of course, we are not inviting models that have been scrawled down on somebody's shower wall.

DR. KOONIN: I understand that.

DR. COLLINS: Right? So, assuming these models, they are backstopped by peer-reviewed literature.

So, the Working Group 1 climate models issues letter of invitation to the major modeling centers in the world, and these entities are well-known, to submit findings to the IPCC. So, there are about 25 of these letters that go out.

And there are new groups that submitted runs or runs that weren't directly commissioned as part of the CMIP5 for analysis. So, it is actually a quite democratic process.

DR. KOONIN: Yes.

DR. COLLINS: So, the results are not biased by some sort of preconceived notion of the fact that the model has to be exhibiting nice, robust climate change, just to sort of take that off the table.

DR. KOONIN: The ensemble consists of how many models altogether?

DR. COLLINS: 45, roughly.

DR. KOONIN: I cannot believe that you or Ben or other people who look at them closely don't have some favorites.

DR. COLLINS: Oh, sure we do.

DR. KOONIN: You must have favorites because you think they do better?

DR. COLLINS: Well, we don't think. We know.

DR. KOONIN: So, what happens if you take only the models that do better and look at all the kinds of results you have been showing us?

DR. COLLINS: So, may I, I am going to elect to deflect -- may I, Steve, to deflect that question to Ben?

DR. KOONIN: Okay. You can answer now, Ben, or when you have the floor in an hour or so.

DR. SANTER: In the study that Bill mentioned or paper published in PNAS 2009, we looked at that question, whether it made a difference in terms of our ability to identify a human fingerprint on changes in atmospheric moisture over oceans if one used just the top ten models in some Letterman-type sense, or the bottom ten.

And we selected those top ten and bottom ten models in 70 different ways looking at a whole bunch of different metrics, how well these models captured today's mean state, seasonal cycle and amplitude and pattern of variability for water vapor and ocean surface temperature.

It turned out that in that particular study, it didn't make much of a difference because the underlying physics was very simple. Essentially, you heat the surface. You heat the lower atmosphere. Water vapor increases.

Because of the non-linearities, you get the biggest bang for your buck over the warmest areas of the ocean in equatorial regions. And that sort of equatorial amplification for water vapor is very different from the dominant pattern of natural variability which has this El Niño-like, horseshoe-type pattern.

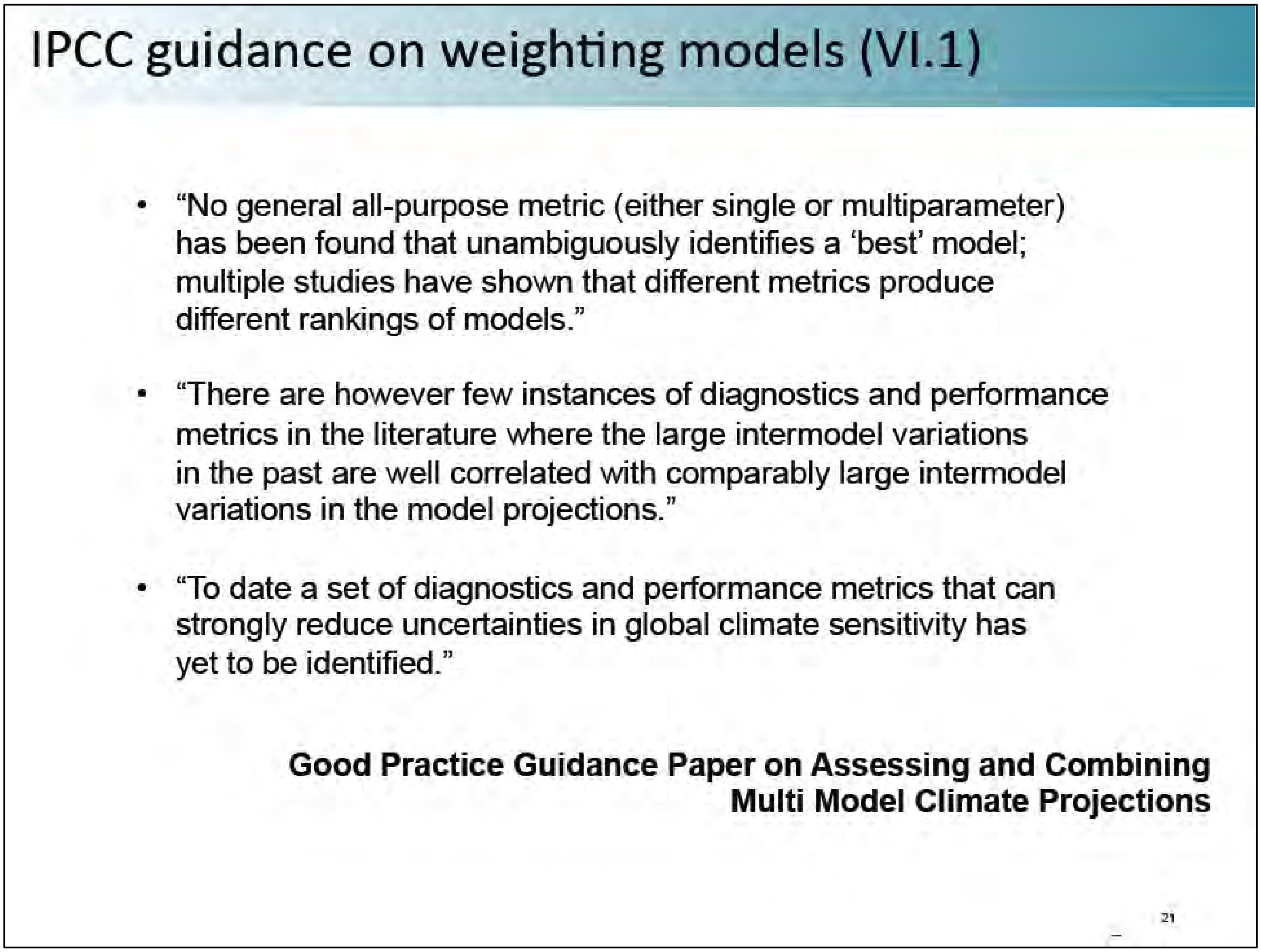
DR. KOONIN: How about if you go to projections over two decades, five decades? Presumably the width gets narrower in the dispersion of the ensemble, among the best, or not?

DR. SANTER: Well, you are saying if one looks for clever transfer functions between present-day observables and the projection uncertainty. So, lots of people have been trying to do that.

DR. COLLINS: That turns out to be, I quoted statements from this expert document. But let me just show those to you. This again [next page], this is the guidance on weighting models, which is this good practice paper. And you can get it off the IPCC's website. So, let me just read these so they are on the record.

(Reading): "No general, all-purpose metric has been found that unambiguously identifies a best model. Multiple studies have shown that different metrics produce different rankings of models."

And so, for example, some models do a great job of reproducing internal variability. Other models do a great job reproducing a time-mean climatology. In many cases



IPCC guidance on weighting models (VI.1)

- “No general all-purpose metric (either single or multiparameter) has been found that unambiguously identifies a ‘best’ model; multiple studies have shown that different metrics produce different rankings of models.”
- “There are however few instances of diagnostics and performance metrics in the literature where the large intermodel variations in the past are well correlated with comparably large intermodel variations in the model projections.”
- “To date a set of diagnostics and performance metrics that can strongly reduce uncertainties in global climate sensitivity has yet to be identified.”

Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections

those are not the same model.

(Reading): "There are few instances of diagnostics where larger intermodel variations in the past are well-correlated with comparably large intermodel variations in the model projections."

It actually turns out to be very hard to use past as prologue. That's the bottom line here. And believe me, a lot of people are looking.

And there are some spectacular examples. For example, snowfall, that is possible, or snow coverage. But there are very few examples in literature. And this has been done exhaustively using ensembles of hundred-thousand member ensembles; very little luck there so far.

Finally, and this is perhaps the core thing for a group like this, we don't have a first-principles theory that tells us what we have to get right in order to have an accurate projection.

So, let's just make sure that that's clear. We do not have a first-principles theory for that. This is sort of an emergent knowledge base.

So, that's the translation of this last statement, "To date, a set of diagnostics and performance metrics that can strongly reduce uncertainties in global climate sensitivity," a la projections, "has yet to be identified."

DR. KOONIN: I am happy to take one more question, but I want to move on so we can try to stay on time.

Ike, did you have --

DR. COLLINS: Isaac had a point, I think.

DR. HELD: No.

DR. KOONIN: Phil?

MR. COYLE: I understand that Ben is going to talk about the hiatus later.

DR. SANTER: Yes.

MR. COYLE: But while you have the floor, do you have any comments you want to make? You must get questions about that all the time.

DR. COLLINS: Well, yes. That actually was dealt with by chapter 9, which is the chapter I was on. I think you accurately captured the state of the field currently.

We are unsure about what -- we know that there are several possible causes. And they are stated in the report. And also, you capture them correctly as well.

They could be errors in the forcing. It could being a mode of natural variability that the models are not correctly reproducing. And it could be cases or it could be that the models are overly sensitive. And so, all three are noted in that the IPCC report and will be actively investigated.

I do not have an opinion. We thought while we were writing this report that it was aerosols. And there were a number of -- people became very alarmed. There were four meetings that went into this report, four face-to-face meetings.

As of the second, we were having these frantic meetings between people like myself on radiative forcing and the later chapters that were looking at these projections saying oh, my God. The models are running hot. Why are they running hot? By "running hot," I mean running hot for 2011, 2012 as we were writing the report.

So, there was a lot of speculation that the projections had sort of overcooked the level of air pollution controls that were going to cause aerosol loading to decrease in the near future. That is a plausible explanation.

Other people have looked at subtle amounts of volcanic activity that have since gone undetected. This is work by Susan Solomon, other changes in the stratosphere. This is one of those topics that I think is going to have to be sorted out.

Now, I am hedging a bet because, to be honest with you, if the hiatus is still going on as of the sixth IPCC report, that report is going to have a large burden on its shoulders walking in the door, because recent literature has shown that the chances of having a hiatus of 20 years are vanishingly small.

DR. KOONIN: Okay, thank you.

DR. COLLINS: Thank you.

DR. KOONIN: All right. I have got to say, I come away, Bill, and thanks for being so clear, that this business is even more uncertain than I thought, uncertainties in the forcing, uncertainties in the modeling, uncertainties in historical data. Boy, this is a tough business to navigate.

DR. COLLINS: Can I respond to that?

DR. KOONIN: Yes, please.

DR. COLLINS: I mean, yes and no. The first calculations of greenhouse gas warming done by Arrhenius were done using the tools of the trade circa 1880.

And he got most of the facts right because he knew, obviously, how to alter the greenhouse effect of the climate system and could write down essentially a zero-dimensional model of the climate system which reproduces a lot of the qualitative behavior we see here.

So yes, we are asking the climate models to do things that -- we are no longer looking at this as a point problem, which is the way

Arrhenius looked at it. We are looking at the model in detail as we have in the past.

But I think to come away with the fact this whole thing is highly uncertain, we fail to recognize both the insight that Arrhenius had, which I think still holds true today, and the fact that the climate models, despite the fact that they have those uncertainties, have on a number of cases predicted behavior that was subsequently verified, which is certainly a nice thing to see in cosmology. And it's very nice to see in the climate.

There is actually a beautiful book written by Ray Pierrehumbert called "The Warming Papers." I strongly urge you to look at that book because it deal with -- there are a number of cases where the climate models anticipated behavior the observing systems at the time could not see and they subsequently saw. And these include changes of large scale beyond the earth's atmosphere due, we think, to the influence of energy.

So, I want to make sure, Steve, we don't come away with too much.

DR. KOONIN: That's good.

Okay, Judy, you might as well start taking the podium.

DR. LINDZEN: I think there are certain things here that are a little bit peculiar, the business of the fingerprint. The only thing you are saying is when you are nearly transparent to space, you are going to have cooling to space.

And when you get further in, you are deep and then you will get warming, but that depends on the feedback. And there is no signature that will distinguish different sensitivities in that. So, it's a little bit awkward.

DR. KOONIN: We are going to get onto that, I suspect.

DR. LINDZEN: Also, black carbon isn't the only reason you can get the sign wrong.

DR. KOONIN: Sure.

DR. LINDZEN: Aerosols can, for instance, cause condensation of ice and change the character.

DR. KOONIN: That was the indirect aerosol?

DR. COLLINS: Yes, Dick is exactly right. There are a number of reasons why the science can change.

DR. KOONIN: Judy …?

DR. CURRY: I would like to start off, as a member of the Executive Committee of the Topical Group on Climate Change, I would like to applaud this committee for the process that you are undertaking.

It is much better than anything that I anticipated and I think it's very good. And this workshop is a real good step and I would like to thank you for inviting me to be a part of it.

For a little bit of context for where I am coming from on this issue, I am not involved in the IPCC. I had some minor involvement in a third assessment report as a contributing author and as a reviewer. The more recent ones, I have not been involved at all.

I am not a climate modeler, although I use climate models, and some parameterizations for my research group on cloud microphysics and sea ice have made it into a few climate models. But I am not what you would call a climate modeler.

My areas of expertise are in clouds including cloud aerosol interactions, sea ice, air/sea interactions and the climate dynamics of extreme events. So, my perspective comes from a little bit more the observational theoretical side than climate modeling. So, it's sort of a counterpoint to some of the other people in the group.

And in trying to decide what to cover in 30 minutes, I decided to keep it very focused on your questions and also tried to pick topics that I anticipated other people wouldn't cover that I could speak to with some sort of expertise.

So to start [next page], I will address your two first questions which sort of gives you a little bit of a perspective of where I am coming from and the rest of my presentation.

So, what do I consider to be the greatest advance? And it's really the narrowing of uncertainty in the aerosol indirect effect.

I think this is the biggest deal. It's an important scientific advance, but it has a number of

What do you consider to be the greatest advances in understanding of the physical basis of climate change since AR4?

- Narrowing of uncertainty in the aerosol indirect effect

What do you consider to be the most important gaps in current understanding?

- Solar impacts on climate (including indirect effects)
- Multi-decadal natural internal variability
- Mechanisms of vertical heat transfer in the ocean
- Fast thermodynamic feedbacks (water vapor, clouds, lapse rate)

implications because we can't so easily blame all the uncertainties on aerosols anymore. It's getting tougher.

DR. KOONIN: The uncertainty narrowed and the mean shift also?

DR. CURRY: Yes, it's not as big overall as we thought. Yes, there is some canceling of black carbon and other stuff. So, there is less wiggle room and we can't blame everything on aerosols anymore.

DR. KOONIN: But it also means that the aerosols are contributing a bit less to cooling. If you start to tune the models, it means the sensitivity is too high.

DR. CURRY: Right, it has those implications. I am sure we will be hearing more about that.

What do I consider to be the most important gaps in current understanding? The solar impacts, and this is really indirect effects, a whole host of things some of which I will mention later.

The issue of multidecadal natural internal variability, mechanisms of vertical heat transfer in the ocean and the fast thermodynamic feedbacks, water vapor clouds and lapse rate, these are the issues that I regard as the biggest outstanding uncertainties.

You asked the question with all the things that have to go on in a climate model to get any kind of a plausible agreement between observations and the climate model output, it's fairly amazing when you think about it.

Not only do the climate models have to be working, but you have to have confidence in your forcing and in the observations against which you are comparing it with.

So, [next page] what do we derive our

CLIMATE MODELS: How can one understand the IPCC's expressed confidence in identifying and projecting the effects of such small anthropogenic perturbations in view of such difficult circumstances?

Confidence in climate models derives from:

- Model relation to theory and physical understanding of processes
- Convergence of different climate models and agreement of successive generations of climate models
- Verification history of numerical weather prediction models
- Success in simulating observed global temperature anomaly trend during 1975-2000.

Curry and Webster 2011: Climate science and the uncertainty monster. *Bull Amer Meteorol. Soc.,* 92, 1667-1682.

confidence in climate models from? We have heard some from Bill already. But we have the model relation to theory and the physical understanding of the processes. Again, these aren't statistical models. They are based on thermodynamics and fluid dynamical equations.

The convergence of different climate models and agreement of successive generations of climate models, and then the verification history of numerical weather prediction models also play into this because that is the heritage of the atmospheric piece of this.

You know, the fact that we can predict weather using these models trickles down to the confidence of climate models.

DR. KOONIN: Just to clarify, a weather model takes the SSTs as a boundary condition?

DR. CURRY: It's a boundary

condition. It's one piece. It's the atmospheric piece of it, just the atmospheric piece of it.

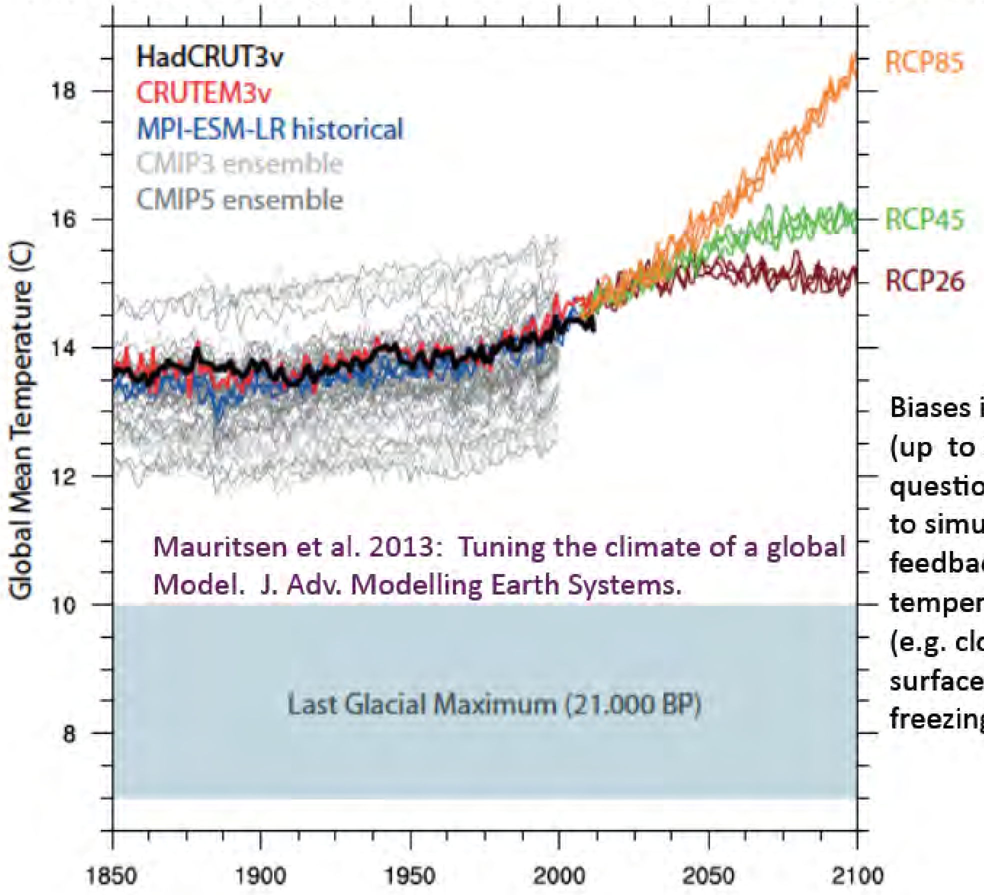
DR. KOONIN: The biosphere, ocean dynamics are not in there.

DR. CURRY: And the other thing is, and I was particularly impressed in the AR4 report is a success in simulating the observed global temperature anomaly trend during the period 1975 to 2000.

Now, the anomaly trend, climate model results are presented usually in terms of anomaly trends. If you actually look at the absolute temperatures from the climate models, it doesn't look so pretty.

This [next page] is from a recent article. I guess my references are given at the back, from Fyfe, et al. And you can see there is a spread of several degrees centigrade amongst the CMIP5 ensemble, the actual model climate. Some of them do a pretty good job of

Absolute temperatures from climate model simulations



Mauritsen et al. 2013: Tuning the climate of a global Model. J. Adv. Modelling Earth Systems.

Biases in modeled climate (up to 2C) raise serious questions about the ability to simulate processes and feedbacks that are temperature dependent (e.g. cloud formation, surface evaporation, sea ice freezing/melting) reproducing, but some of them are off by several degrees.

And you say, well, it's just the anomaly in the trend that matters. But again, to the extent that thermodynamics is important like the melting temperature of snow and sea ice and the formation of clouds and the Clausius-Clapeyron Equation is temperature-dependent, you know, these temperature errors do matter.

And I would just love to know because some of these are very far off, at what temperature do they actually melt sea ice? I would like to know.

DR. KOONIN: The implication being that it has been tuned?

DR. CURRY: Yes, there is a lot of tuning that goes on. And it's actually, I think, a relatively difficult thing to even get the climatology. I think the better models that Bill discussed and his favorites, I suspect, are ones that do a reasonable job of model climatology. I don't know.

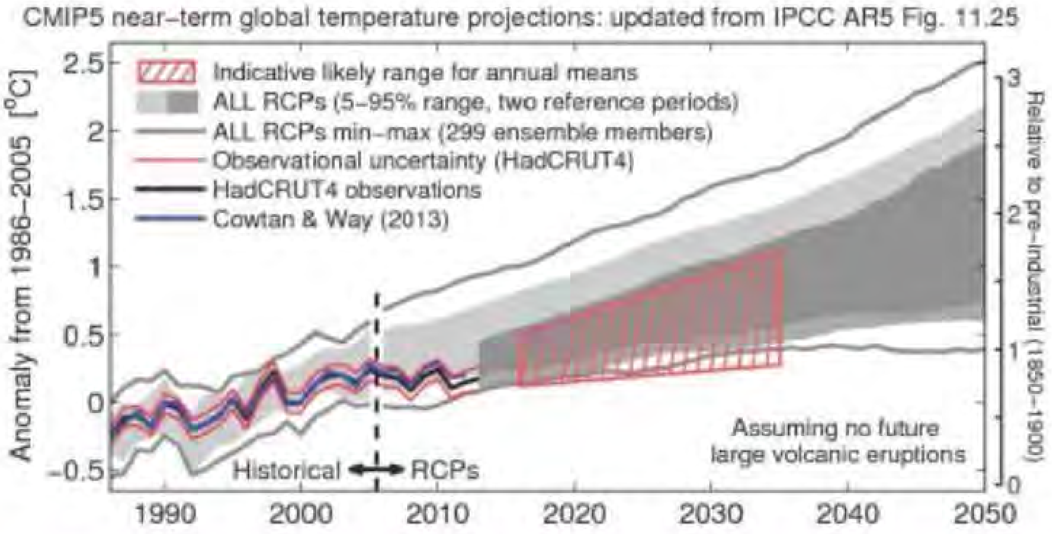
But this is the spread of more than 2C is larger than the trend that we have seen. So, this is something to keep in mind that it's a very hard thing to do to get all of this right.

The stasis,[next page] okay, to what do I attribute the stasis? Well, I think it's predominantly an issue of natural internal variability. And I will talk more about this in a minute. There is potential for solar effects, but again, this is in known/unknown territory.

I am not convinced by arguments related to Chinese power plants, reductions in CFCs and volcanic activity. I don't think these are very convincing. When I was at the AGU meeting, American Geophysical Union, people were talking about oh, the hiatus has gone away. It's not a

To what would you attribute the stasis?

- Natural variability (internal)
- Solar effects
- I am not convinced by arguments related to Chinese power plants, reductions in CFCs, volcanic activity



http://www.climate-lab-book.ac.uk/2013/updates-to-comparison-of-cmip5-models-observations/

problem.

There was a paper published by Cowtan and Way who extended the temperature analysis of the UK group to fill in for the Arctic and they said well, this has gone away.

Well, I don't know if you can see this far, but this is a diagram from Ed Hawkins. This is from figure 11.25 from the AR5. And Ed Hawkins, who did the original figure, redid it with Cowtan and Way.

And you can hardly tell the difference between the blue and the black line. And the difference in trend in Cowtan and Way is a little lower in '98 and a little higher since 2005. So, it's really in the noise of the observation. This doesn't make the pause go away. The hiatus is still there.

Okay, on to internal variability.[next page] As Bill mentioned, pure internal variability is associated

What is the definition of “internal variability”?

- Pure internal variability is associated with nonlinearities and chaotic nature of the coupled ocean/atmosphere system. External forcing projects onto the modes of internal variability and so influences the amplitude, tempo and phasing of the internal modes
- There is some predictability on decadal timescales of the multidecadal modes of climate variability, particularly the AMO

Are there any other possible multidecadal modes of variability besides ENSO? If so, how is that variability accounted for?

- AMO & PDO; stadium wave
- This variability is not accounted for in attribution studies

If non-anthropogenic influences are strong enough to counteract the expected effects of increased CO2, why wouldn’t they be strong enough to sometimes enhance warming trends, and in so doing lead to an over-estimate of CO2 influence?

- I have argued that non anthropogenic influences (e.g. solar; warm phases of PDO and AMO) have enhanced the warming in the latter quarter of the 20th century with non-linearities and chaotic nature of the coupled atmosphere ocean system.

Now, it's very difficult to separate internal variability from forcing because my understanding is that the external forcing projects onto the modes of variability. So, this is not as easily separable as you would like or would hope.

In terms of predictability, yes, the models can simulate oscillations that look something like the modes. When you combine it with external forcing, Bill showed a figure where you get something reasonable.

But in terms of the ones that we care about in actually getting the timing right, it's very hard to predict these.

DR. KOONIN: Clarification: the models don't get the timing of ENSO?

DR. CURRY: Yes, even with

initialization and the decadal simulations, it looks like there is some predictability of the Atlantic multidecadal oscillation, maybe out to ten years, but Pacific multidecadal oscillation just --

DR. KOONIN: It's not in the model?

DR. CURRY: Yes, just fell apart. So, apart from ENSO, I mean, the other modes, the longer modes, the Atlantic multidecadal oscillation, Pacific decadal variability are important ones on time scales that we were concerned about, and also the stadium wave, which I will mention in a minute.

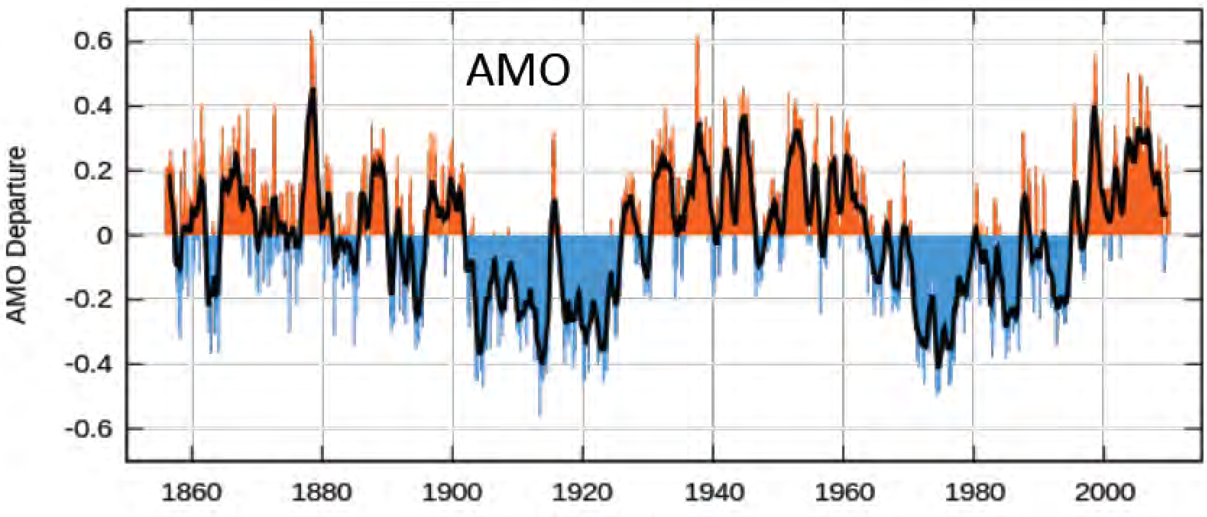
And while the models do produce oscillations that sort of look like them, the timing of the variability isn't right. And this kind of variability doesn't get explicitly included in attribution studies.

So, you asked a question, "If

nonanthropogenic influences are strong enough to counteract the expected effects of increased CO_2, why wouldn't be they strong enough to sometimes enhance the warming trends and in so doing lead to an overestimate of CO_2 influence?"

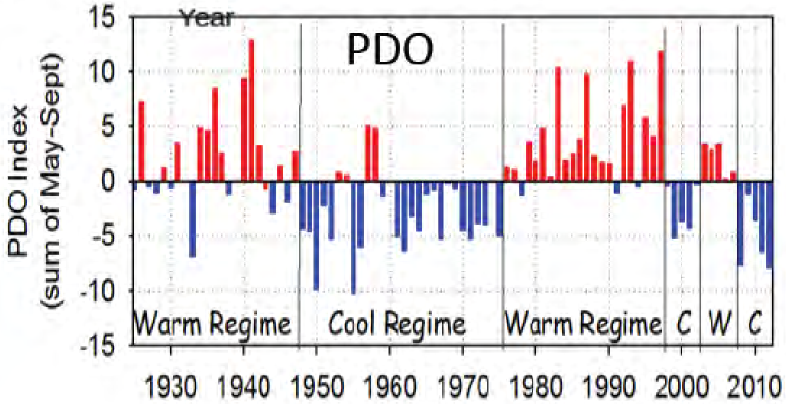
Well, if you are attributing the hiatus to natural internal variability, this immediately raises a question well, what about the warming from '75 to 2000? And so, I think that was probably juiced to some extent by natural variability.

Just to show you what I am talking about, the Atlantic multidecadal oscillation is shown on the top one.[next page] Pacific decadal oscillation is shown on the bottom one. The Pacific decadal oscillation temperatures in the Pacific are, well, unreliable before 1980, but really unreliable before about 1920. So, other than proxies, we don't have



Currently:
- Warm AMO
- Cool PDO

Previous analogue:
- 1946-1964



really good estimates going back.

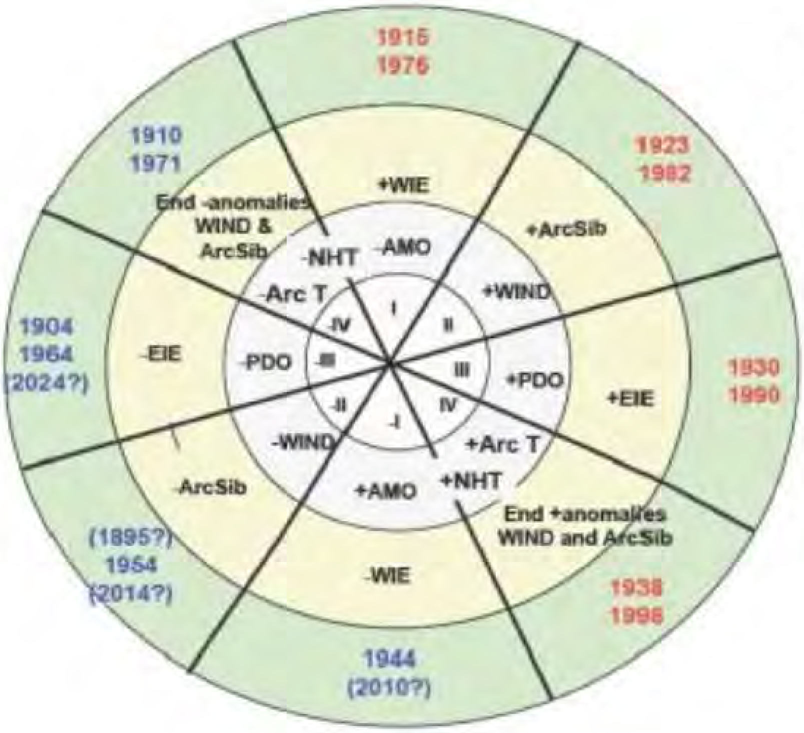
DR. KOONIN: These are constructed from differences in pressures or temperatures or something?

DR. CURRY: Yes, this is really constructed from mostly temperatures and patterns and stuff. So, since 1995, we have been in the warm phase of the Atlantic multidecadal oscillation.

And we started flickering in the PDO going to the cool phase in the latter years of the 20th century. And we have been decisively in the cool phase for the last couple of years.

Now, [next page] a recent paper that I coauthored called "The Stadium Wave," what it does is it takes a bunch of these teleconnection indices of natural variability and linked them into a network. And we saw a progression of all of these things

Stadium Wave



The 'stadium wave' climate signal propagates across the NH through a network of ocean, ice, and atmospheric circulation regimes that self-organize into a collective tempo.

Wyatt & Curry, 2013: Climate Dynamics

integrating. I don't have time to explain it here, but this is a simplified version of the diagram.

Here we see peak AMO, negative AMO, peak PDO and negative PDO. And in the second row are some indices related to sea ice and around this outer ring. So, for the past several hundred years, we have sort of seen a little bit of a repeat in the progression of how this goes.

And the implications: I think we have seen a transition. We are in the midst of a little transition from here to here (indicating slide). And this makes a couple of sort of simplified, if you continue the network, it makes predictions.

The bottom line is that we could see the hiatus if it is natural variability continuing into the 2030s and starting to see a mini-sea ice recovery like the Western Eurasian Arctic. Kara Sea around there I think has bottomed out already and is starting to recover a little bit.

So, there is a natural variability component here. To what extent this is important relative to the forest, I mean, to me, that's the big request question.

But there is a component here. And the natural variability piece of this will tend, I think, if this is right, to want to keep cooling for another two decades, potentially.

DR. KOONIN: The world has been around this wheel at least once?

DR. CURRY: Okay, with proxies, we have gone back 300 years. For the last maybe 150 years it has been nominally a 60, 65, but it's shorter in previous times. It's more in the 50 years.

So, I think the external forcing does change the tempo and things like that. Again, this is just an idea, but it's a potential explanation for the hiatus. And if this is correct, we could be seeing it for another two decades.

Now, with regards to the climate models, like as Bill said, vanishingly small after 20 years. And I think somebody has smoked out a CMIP3 simulation that showed a hiatus looking similar to the current one of 21 years.

But that is, yes, so that's sort of it. So, you know, time will tell as to whether this general idea is right. But I think this helps organize the modes of natural variability.

It doesn't answer the question the relative magnitude of the natural variability versus the forced variability. But what we are seeing happening right now could continue for another two decades.

So,[next page] "How would the model's underestimate of internal variability

How would the models’ underestimate of internal variability impact detection and attribution?

- Incorrect simulations of natural internal variability results in biasing detection and attribution in favor of external forcing as the cause of any variability; in the latter half of the 20th century, the dominant external forcing is anthropogenic.

What are the implications of this stasis for confidence in the models and their projections?

- Models are not useful on timescales of of up to 2 decades; serious implications for decadal projections and attribution analysis on time scale of decades (including 1975-2000 period of warming)

How long must the stasis persist before there would be a firm declaration of a problem with the models?

- Stasis persistence beyond 20 years would support a firm declaration of a problem with the models.

impact detection and attribution?"

Well, incorrect simulations of the natural internal variability, even if you have the amplitude right getting the phasing and the timing right, it results in biasing detection and attribution in favor of external forcing as the cause of any variability.

And in the latter half of the 20th century, the dominant external force is anthropogenic. So, this is potentially how it could lead to an overestimate.

So, "What are the implications of the hiatus or stasis for confidence in the models and their performance?"

Well, to me, this tells me the models aren't useful on time scales of two decades or less, because if they are regarding natural internal variability as unpredictable, we are sort of seeing evidence that they are not predicting it. And it may mean potentially even longer. They only get the very longer-term trend.

If we have high-amplitude stuff, 60 years, then the climate models aren't going to give us terribly useful predictions on decadal time scales.

And so, "How long must the stasis persist before there would be a firm declaration of a problem with the model?"

I would say 20 years. When you actually start it at '98 or at 2001, when I think was the more fundamental shift in the circulation patterns. We can debate, but I don't think we will be splitting hairs. Either it is going to turn around quickly or it's going to stay for a while.

So, it will be interesting. By the time of the sixth assessment report, I think we will have gotten to an interesting time. It's either turned around or it's still flat. So, I am not signing up for that assessment report.

DR. KOONIN: What is the expected timing on that?

DR. COLLINS: 2020, roughly. We will all have new day jobs by then.

DR. CURRY: Now,[next page] if that occurs, what would the fix entail? To me, this is a really fascinating question. I don't think it's an issue of tuning. Again, the model fundamentals are sound, so there is something in between. Well, what?

I think the problem is the ocean circulation and the coupling to the atmosphere. Higher resolution, I think, can solve some of this. But I suspect getting it down for the ocean, you need really high resolution. So, even if we get to the desired resolution, is that going

If that occurs, would the fix entail: A retuning of model parameters? A modification of ocean conditions? A re-examination of fundamental assumptions?

- The problems are with the ocean circulation and coupling with the atmosphere. No easy fixes, although higher resolution would help. Some fundamental unknowns in terms of how the ocean rapidly transports heat in the vertical.

What do you see as the likelihood of solar influences beyond TSI? Is it coincidence that the stasis has occurred during the weakest solar cycle in about a century?

- Solar effects beyond TSI are the major known unknown (e.g. cosmic rays, global electric circuit, magnetic field). We simply don't know, but I wouldn't be surprised if they are important. It is not known to what extent solar effects have caused the stasis, this may be coincidence or not to solve the problem? Not really.

I mean, the key issue is how the oceans transport heat in the vertical. And the models are keeping the heat in the upper ocean. And we are seeing lots of stuff going on in the deeper ocean.

And there are some ideas on how that occurs, but actually getting that into the climate models in a sensible way is a challenge.

You asked the questions of solar influences beyond TSI. This is a subject that intrigues me greatly. It's sort of in the known/unknown category. All you can do is sort of speculate on ideas.

That might be cosmic rays, global electric circuit, magnetic field. We simply don't know. But I wouldn't be surprised if they are important.

And when I talked to people doing planetary atmospheres, a question that -- why don't you people pay attention to the magnetic field in planetary atmosphere? This is a big deal.

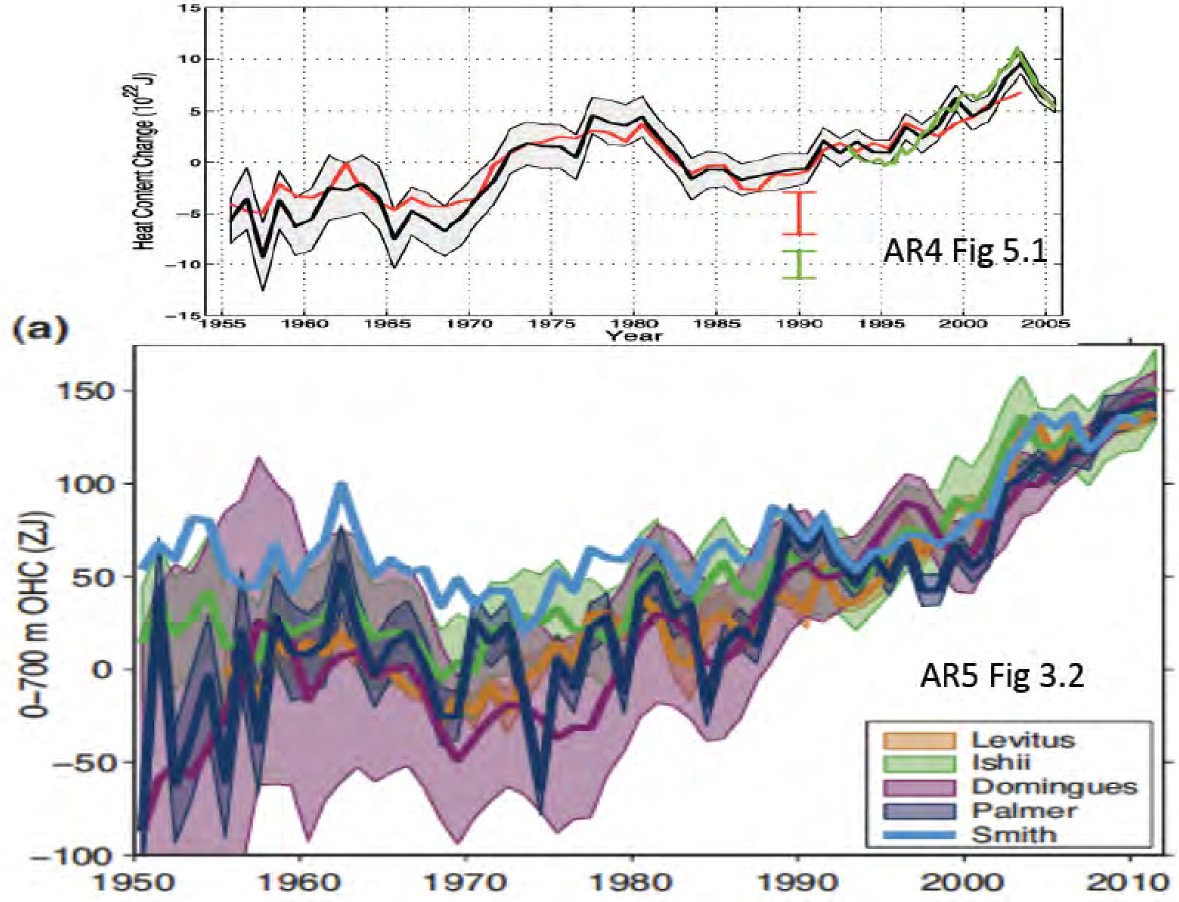
And we don't really pay much attention to the magnetic field in context. So, there are things like this, questions that we haven't really asked. So, I don't know, but some very intriguing possibilities.

Issues [next page] related to ocean heat content and the measurements, I want to give you a sense of the uncertainties. The top figure is the ocean heat content zero to 700 meters from the AR4. You see a very narrow uncertainty range.

You also see there is a bump (indicating 1975-1985). What is this bump? Now, we look at the same figure from the AR5, much broader range of uncertainty and the bump disappeared, okay.

Well, the issue is that, apart

Ocean Heat Content 0-700 m

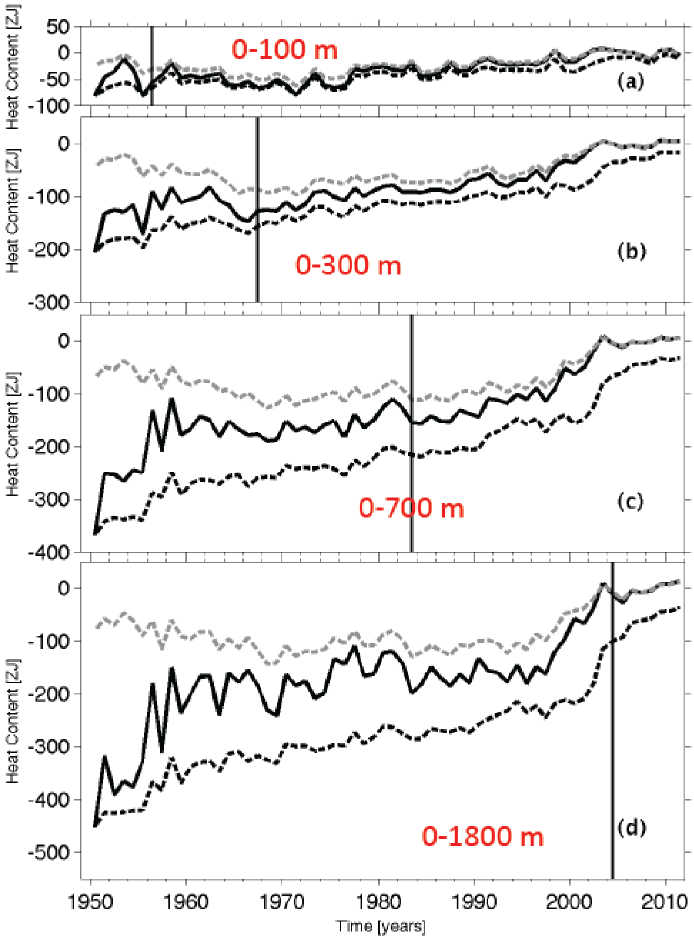


from the issue of spatial coverage, which I will get to on the next slide, it is not simple to process these measurements, especially the expendable bathythermographs with little things that are just dropped. You have figure out how to calculate it from the voltages kind of thing.

So, in the '70s and '80s, it was a lot of reliance on these XBTs as people are trying to figure out how to process. And before that, it was NBTs. And there are questions about how to process that, also.

So, a lot of this, apart from spatial coverage, there is a lot of uncertainties in how you do the calibration and the processing.

Now, the next figure [next page] gives you a sense of the impact of the sampling. This is a recent paper that I like. We have different layers in the ocean. And the vertical line is the date when you



Ocean Heat Content

(observations constrained by SSH)

Lyman & Johnson, 2014:
J. Climate, in press

had 50-percent coverage of the ocean.

So, we see since the '50s, we had it in the upper ocean. But it's really only been very recently with the Argo that we have any kind of coverage really below 700 meters.

And so, the different curves represent different assumptions that you make about the stuff that you can't measure. So, this gives you a crude estimate of the uncertainty in the coverage in trying to make a global estimate.

So, what do you see? A feature, most climatologies agree that there was a little peak in 2003 and since 2003, it has been relatively flat, although there is uncertainty there.

A big increase, really, since 1995 to 2003, it's a big part of the increase, and then relatively flat in the stuff before 1960 is probably pretty implausible. So, this is sort of what the observations seem.

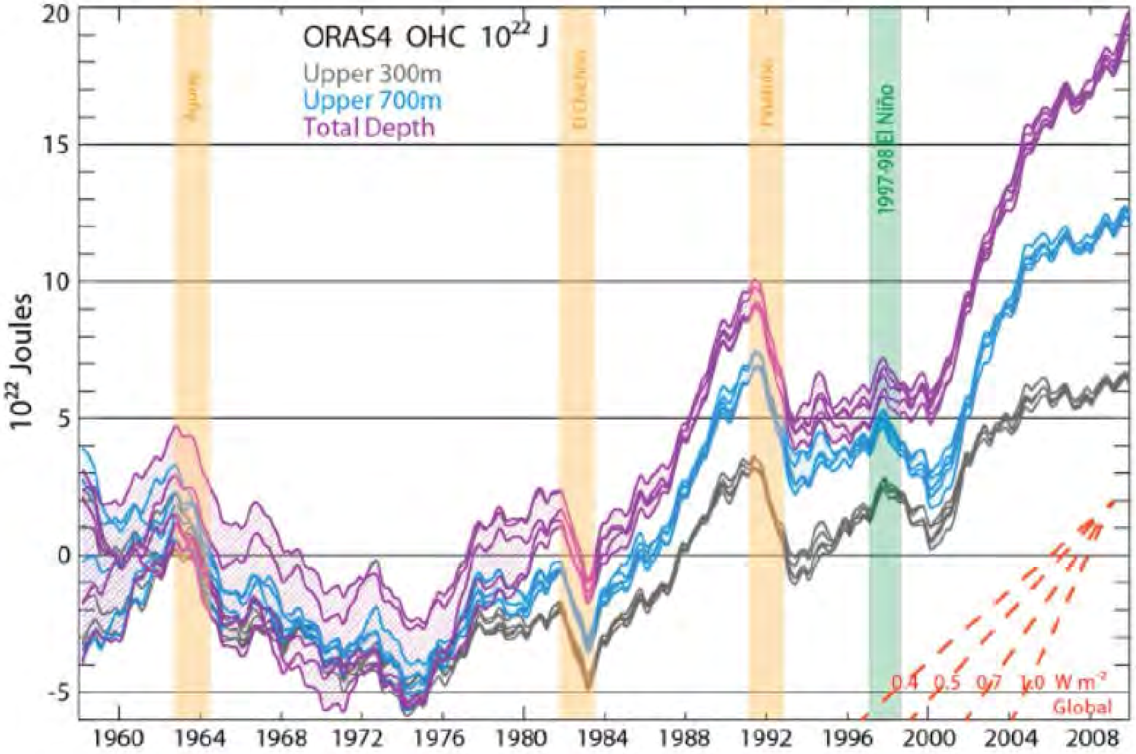
Now, the other way of sort of filling in the gaps is through ocean reanalysis.[next page] And this is done by the way you initialize inverse -- okay, I don't need to explain how this goes.

But this is probably the most reliable of the ocean reanalyses. This is from ECMWF. And so, for the first time, we get something below 1800 down to the deep ocean down to several thousand meters.

And so, they are effectively filling in the gaps through the ocean's circulation model, which seems like a sensible thing to do. But we see some features that don't look all that much like the observed.

We see this big spike which we see around '92, which we didn't really see anything there. We see in the observations we were seeing starting around '95 there was a big increase and a relatively flat right

Ocean Reanalysis – ECMWF



Balmaseda et al. 2013: J. Geophys Res.

here.

DR. KOONIN: Just so I understand: Again, this is some combination of models driven by observations?

DR. CURRY: Yes, models that simulate the observation in the same way that a numerical weather prediction initializes the weather model. They are sort of initializing the ocean weather model, if you will.

DR. KOONIN: Regarding the model data as --

DR. CURRY: Right. So, there is a background circulation and then they assimilate observations where it's available. So, I think this is eventually a very promising approach, but it doesn't quite have the fidelity to the observations yet.

But you see a lot of heat going into this layer that goes down to the total depth. And so, how is that heat getting there? The models keep it all in the upper layers. They are not sending it down below.

So to me, this is one of the big issues. The ocean seems to transfer heat vertically more rapidly than we know how to do it in the model.

So, [next page] as far as your question goes, some have suggested that the missing heat is going into the deep ocean. Okay, if you average this heat over the depth of the ocean, it's .05 kelvins since 1960. So, it's the big heat reservoir, so there is not a big temperature change.

So, why would the heat sequestration have turned on at the turn of this century? Well, if this is a robust thing, presumably it has something to do with natural internal variability.

And so, what could make it turn off? Natural internal variability. And if this is related to the stadium

Some have suggested that the "missing heat" is going into the deep ocean, causing mK temperature rises. Are deep ocean observations sufficient in coverage and precision to bear on this hypothesis quantitatively?

- No. There are substantial uncertainties in data coverage and calibration, and reanalysis estimates disagree quantitatively with each other and with data only analyses. (uncertainty estimates)

Why would the heat sequestration have "turned on" at the turn of this century?

- Presumably associated with natural internal variability

What could make it "turn off" and when might that occur?

- Same; the next shift in the stadium wave is expected in the 2030's

Is there any mechanism that would allow the added heat in the deep ocean to reappear in the atmosphere?

- The deep ocean has warmed approximately 0.05K; if the heating is well mixed in the ocean, there is no way for warming in the atmosphere to occur beyond 0.05K wave idea, whatever, we could see this dynamic changing sometime in the 2030s.

Now, in the media, I think that the phrase has been used the heat will come back to haunt us. Well, if this heat is genuinely well-mixed in the ocean, you have got the second law of thermodynamics on your side. This heat is not going to, other than .05 kelvin, this heat isn't coming back.

So, the question is, to what extent is this well-mixed or is it indiscreet plumes or whatever? So, this whole issue of ocean mixing, to me is, like, one of the biggest issues out there.

So, I don't think too much of this. I mean, this is actually quite a way that people hadn't thought of sequestering heat in the deep ocean. If you can well-mix it, that's an interesting way to sequester it.

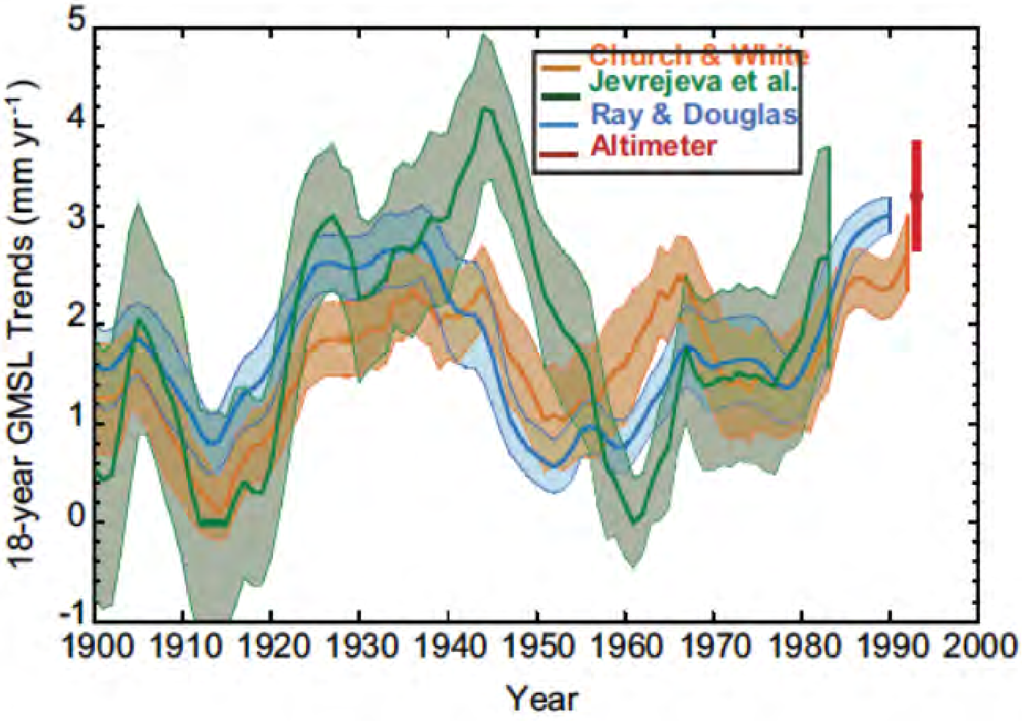
On to the issue of sea level rise, now, I am not a particular expert on the methods of determining this, but I want to remark on something. Again, what this is, it's in your document, this figure. [next page]

The issue is this bump here (indicating 1945). And it's the same rate here as here, basically (indicating red bar). And again, sea-level rise is one of the things we have done in the stadium wave, and it really does fit with that kind of an explanation in the context of natural variability.

So, you are seeing this big signal of natural internal variability in the sea-level rise data as well.

DR. KOONIN: So, you would say from the stadium wave it is going to come back down again?

DR. CURRY: Yes. At some point, by 2040, the natural



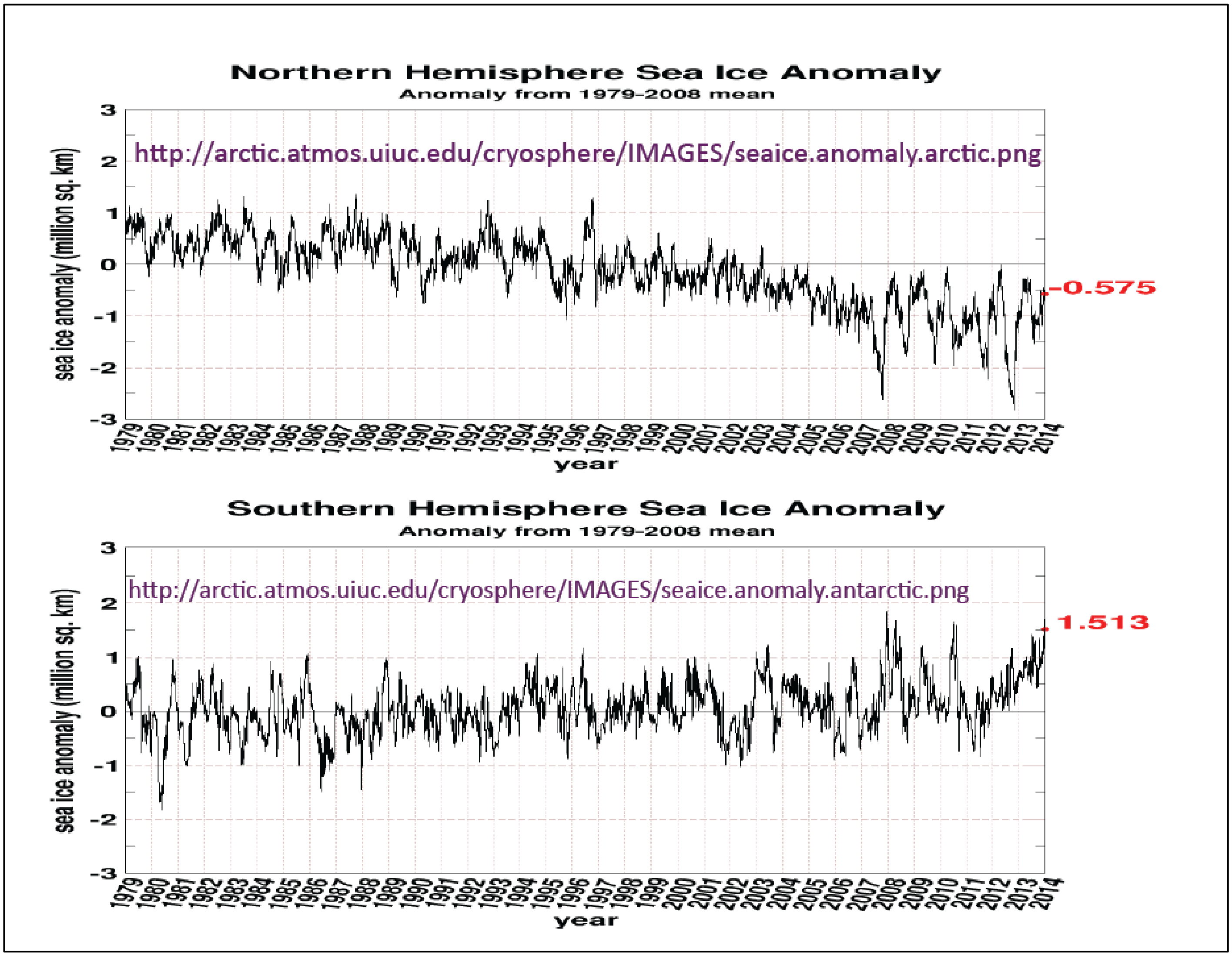
20th century sea level trends co-varies with the AMO and the stadium wave variability would support warming, which would enhance, yes. So, at some point it will turn around.

But again, so, if you look at this from the perspective of 2100, all this may look like noise. But from where we sit right now, it doesn't feel like noise and is challenging, you know, the climate models.

With regards to sea ice, this [next page] is the anomaly of Arctic sea ice. And you see the decline particularly over the last two decades. You see the two record-breaking years.

Now, Antarctic shows a slight positive trend with some of the biggest values in the last decade. There is almost sort of an "anti" with the two hemispheres.

Now, to what extent is this natural variability versus forced variability, particularly the Arctic? You can't tell just looking at data


Northern Hemisphere Sea Ice Anomaly
Anomaly from 1979-2008 mean
http://arctic.atmos.uiuc.edu/cryosphere/IMAGES/seaice.anomaly.arctic.png
sea ice anomaly (million sq. km)
3
2
1
0
-1
-2
-3
-0.575
1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
year
Southern Hemisphere Sea Ice Anomaly
Anomaly from 1979-2008 mean
http://arctic.atmos.uiuc.edu/cryosphere/IMAGES/seaice.anomaly.antarctic.png
sea ice anomaly (million sq. km)
3
2
1
0
-1
-2
-3
1.513
1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
year

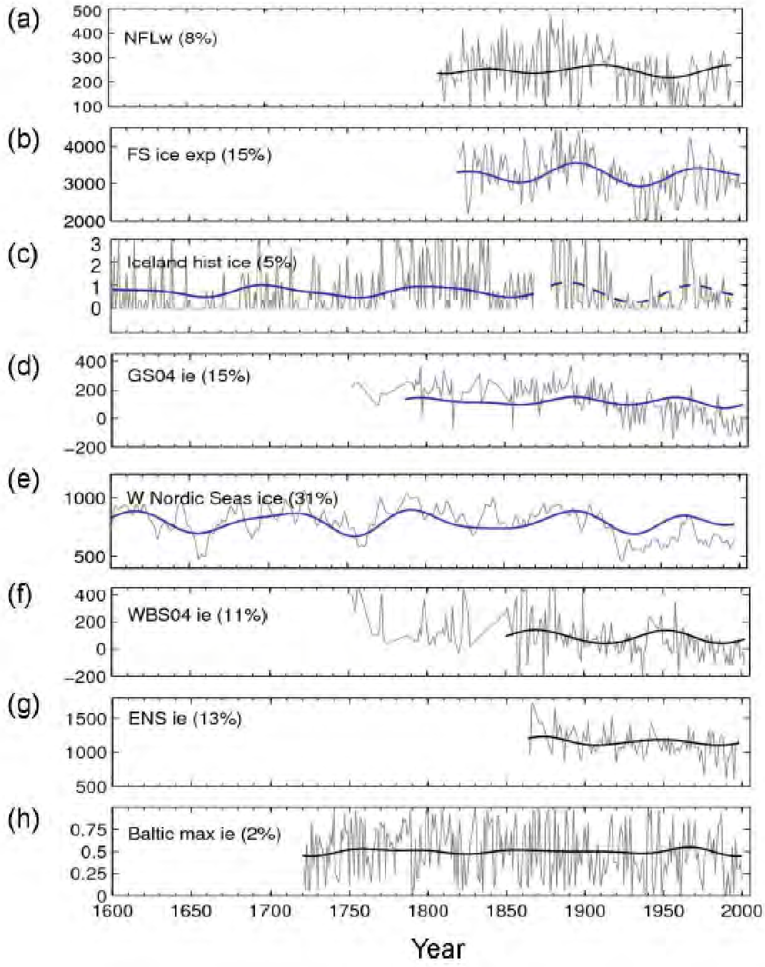
from 1979. So, you want data to go farther back and there isn't a heck of a lot of it.

This [next page] was a paper that is in press that, I think, synthesizes the data sets that are available going back in time with some sort of sufficient resolution.

Some of the paleo stuff doesn't have good enough resolution for you to resolve something out of the decadal time scale. And what they were particularly looking for was some signal from the Atlantic multidecadal oscillation.

And you certainly see it in the Fram Strait. You see it in the Atlantic Arctic. You do see a pretty big signal of the Atlantic multidecadal oscillation.

So again, this is early days of trying to sort of out what the internal variability piece might be. So, at this point, we don't know to



Arctic sea ice variability

Miles et al. 2014:
J. Geophys. Res., in press

what extent --

DR. KOONIN: Five minutes.

DR. CURRY: Yes, I am getting very close to being done.

So,[next page] what extent do we believe the recent Arctic decline is unusual? It's probably unusual. The extent to what is natural variability versus forced, we still don't know. The thing that raises questions is the models predict the Antarctic to be declining, not increasing.

So, the fact that we don't understand that one, there are some ideas related to hydrological cycle, wind patterns and stuff that might explain that. But we don't have a good understanding and the models don't get it right.

So, if you don't get it right in both hemispheres, do you understand what is going on, either? And I would argue that I am concerned as to whether we really understand

To what extent do you believe the recent Arctic decline to be unusual: "There is medium confidence that the current ice loss and increasing SSTs in the Arctic are anomalous at least in the context of the last two millennia."?

- Determining sea ice extent prior to the satellite era is very challenging, using proxies and historicla data. Much more work is needed on this topic, and I find the 'medium confidence' to be wholly unconvincing.

Please comment on the ability of the models to reproduce the Arctic trend, but not the Antarctic trend.

- ~ 47-60% of the Arctic sea ice decline is natural (Stroeve et al. 2012); climate models that reproduce the observed trend without correct natural variability have CO2 sensitivities that are too high (e.g. two wrongs make a right).
- Antarctic sea ice increase is complex interplay between the hydrological cycle, winds, and ocean mixed layer, which models do not correctly simulate (Liu, Curry et al. PNAS 2010)

what is with going on.

And the interplay between natural variability and forced variability in the sea ice is fascinating, but we need more data. And trying to piece this together is key, not simple.

I already mentioned that predicting from natural variability a gradual recovery of the Arctic sea ice progressing from the Eurasian Arctic around the Russian Arctic that we might see over the next 20 years.

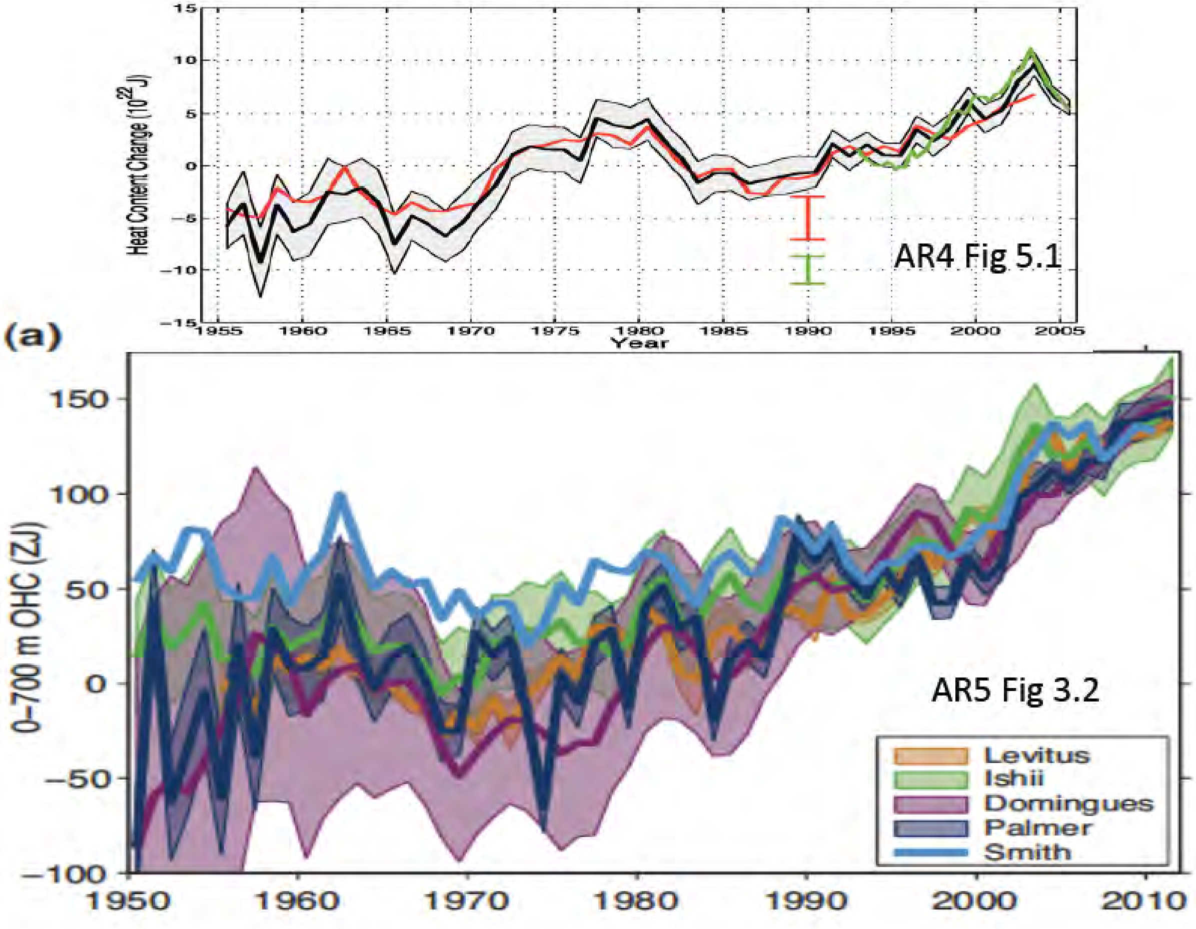
Okay, I think that's all I want to cover.

DR. KOONIN: Thank you. It's open for questions.

DR. ROSNER: Earlier on, you showed a plot of the data for the ocean warming. Could you go back to that slide. [next page] And what you showed us a plot that dates back to the fourth assessment, yes, that one.

So, I have to say that that

Ocean Heat Content 0-700 m



graph, that picture surprised me hugely because I am used to progression advanced in science.

And I would think that, as we get better in measuring, that we do better in our error estimates and in our assessment of what the data really is. And this seems to state the opposite.

Could you comment on that? I guess what I am asking is, do we understand the errors or not?

DR. CURRY: We are starting to. We are starting to. Even in surface temperature, I would say it's only literally a paper in the last two months by John Kennedy at the UK Met office did a really good error analysis of sea surface temperatures, much better than anything we have seen.

Every time somebody does a really good job, the error bars get bigger because they are incorporating more sources of error and better understanding of what the errors actually are.

So again, these aren't laboratory measurements. In physics, what you say holds. But when you have a natural system that you are just sort of dealing with the wild cards that you have been dealt, it's difficult to decipher.

DR. LINDZEN: Also indirect measurements.

DR. CURRY: Indirect measurements, yes, these are indirect measurements. These are inferences. And even if they are direct measurements, they are not direct measurements of what you really want.

DR. KEMP: At the beginning of your talk, you reminded us that the models are rooted
in more than just statistics, that they are based in physics, and therefore they are somehow

inherently, I don't want to say good, but in the right direction; and that the discrepancies that we might observe right now are about timing of the internal variability; but that in the long run, because the physics is basically right, it seems like what you might be saying is in a 100-year time scale, model-predicted things like ECS would be basically right; is that correct?

DR. CURRY: There are two other big uncertainties on my list on the first page. One is the solar indirect effect, a wild card we don't know. And the other one --

DR. LINDZEN: Sensitivity.

DR. CURRY: Sensitivity, yes, the sensitivity the fast feedbacks, water vapor, cloud, lapse rate. Again, this is the big wild card, big wild card. I mean, this is the name of the game. And all of these things are related to subgrid-scale parameterizations.

So, the things related to the fast feedbacks aren't things we have good, fundamental equations for at that scale, because it's related to very small-scale processes that are hidden in the parameterizations that are subject to a lot of tuning.

DR. KEMP: Those are additional issues. But my question is this: To what extent are the models predictive if we are predicting outside of the range in which we can calibrate them?

Say if it were a purely statistical model, the answer would be there or not. But with these models there is some amount of physics — but then there is also calibration parameters which are based on historical observations — can you give me a sense to what extent 100-year, 200-year, 300-year predictions with exogenous forcing can be predicted by the models?

DR. CURRY: This is a big question. It depends on what time scales. I mean, you have already seen that we have got noise on 60-year time scales that the model can't really predict.

So, once you go beyond a century or two centuries, then it relies on forcing and correct feedbacks in the model. And so, that's the big question, I mean, what you asked.

So, a lot of these parameterizations are regime-dependent. If we go into a regime that is very different than our current climate, then it depends on how robust those degrees of freedom are.

And that's the big unknown. I mean, if you are taking the climate to something very different, ten degrees, I would think all bets are off.

Whether two or three degrees is sufficiently close to the regime for which the model can handle, I don't know, maybe more likely than ten degrees. But that's the big question.

DR. KEMP: It seems like the historical observations are so poor that even two or one degree goes outside really the calibration to me. You only have calibration data of meaning in the last decade.

DR. CURRY: It could be, yes.

DR. KOONIN: Phil?

MR. COYLE: The stadium wave analysis that you showed, I think it's very interesting. From where I am sitting, I couldn't read all of the notations and all the rings. To what extent does that analysis include human activity?

For example, does it include at all much more CO_2, much more methane?

DR. CURRY: We remove a secular

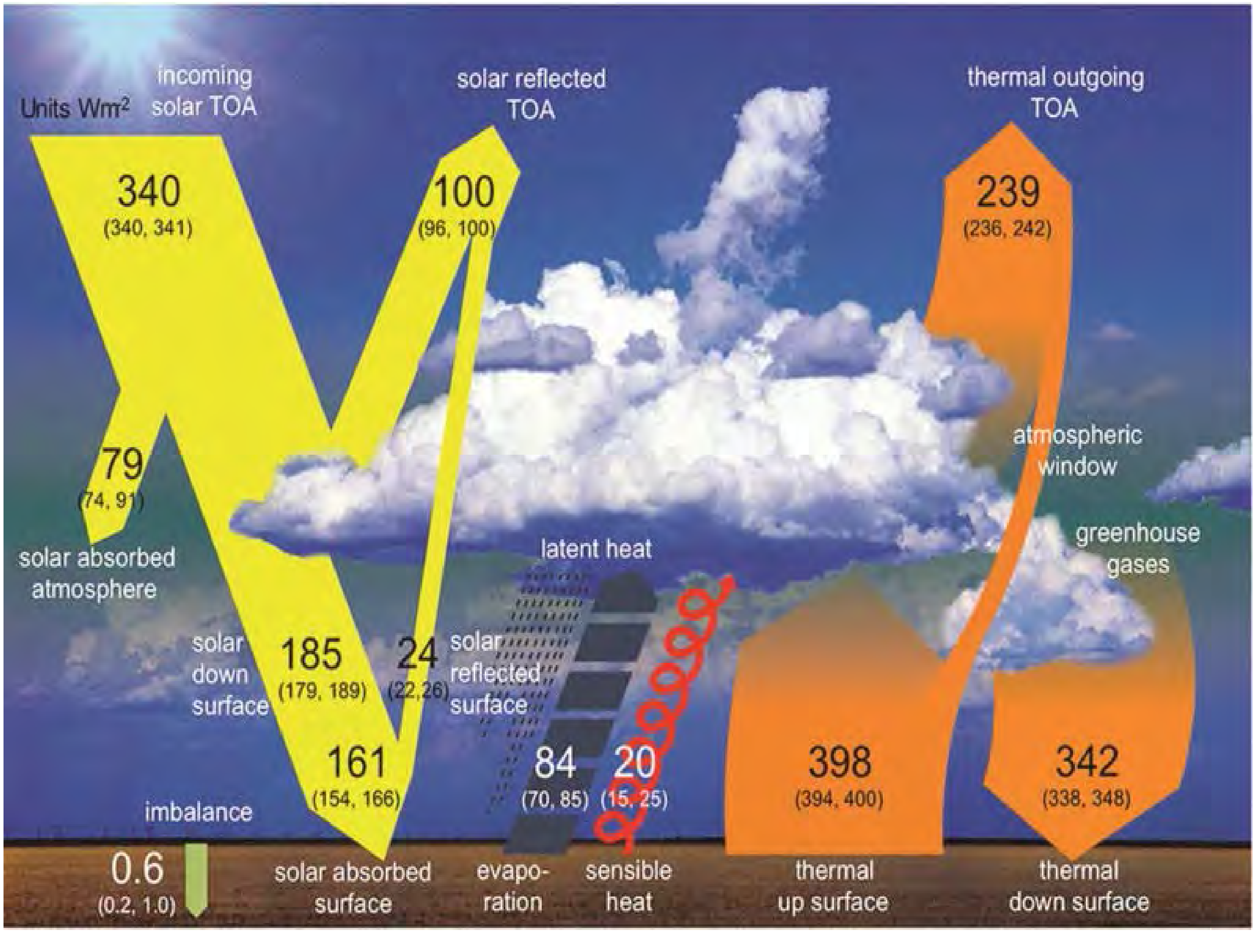
trend, whatever caused it. We removed the secular trend and look at the variability. So, it says nothing about external forcing other than there is a secular increase.

MR. COYLE: Have you thought about a way to include it?

DR. CURRY: Not yet. Some other people have used the stadium wave in observationally determined attribution-type sensitivity-type studies. So, a couple of people have tried it and have shown it to me. Nothing has been published yet. So, other people are trying it.

MR. COYLE: Thank you.

DR. KOONIN: Judy, since you raised the ocean heat content, I want to ask a question. I want to put a picture up there to frame it.[next page] It probably takes you back to the first or second week of the courses you all teach about the climate system. But it's something I don't understand and I


Units Wm2
incoming solar TOA
340
(340, 341)
solar reflected TOA
100
(96, 100)
thermal outgoing TOA
239
(236, 242)
79
(74, 91)
solar absorbed atmosphere
atmospheric window
greenhouse gases
latent heat
solar down surface
185
(179, 189)
24
(22,26)
solar reflected surface
161
(154, 166)
84
(70, 85)
20
(15, 25)
398
(394, 400)
342
(338, 348)
imbalance
0.6
(0.2, 1.0)
solar absorbed surface
evapo-ration
sensible heat
thermal up surface
thermal down surface

expect other people don't as well.

I want to talk about energy balance in the system as a whole. Ocean heat content: I sort of got that little green arrow at the bottom left, 0.6 watts per square meter, I calculated proudly is ten zettajoules per year.

So, it works. That is good. So, I understand where that number came from. That's the slope of the ocean heat content over the last decade.

There are other numbers floating around in watts per square meter that I don't quite understand. So, there is radiative forcing, right? We heard about that, two and a half with big error bars. And as Bill said, that is the net, the change in the net flux downward at 60 years or whatever.

Suppose I rolled the clock back to 1750. Then the radiative balance is zero at the top, by definition, right?

DR. COLLINS: I will say maybe not identically zero. There has been some evidence now converging that heat from burning rice paddy work from the Chinese produces enough methane to interpret it down slightly.

DR. KOONIN: I will give you half a watt for that, all right. Let's go back to 1000 or maybe older. That number goes to zero, right? There isn't much energy stored in the atmosphere in the surfaces, I understand it. Most of the storage is in the ocean.

So, does that mean that the .6 number goes down to minus one and a half or something like that? What do we expect for that number down at the bottom in preindustrial times? Judy had the floor first, but if she wants to --

DR. CURRY: No, no, go ahead.

DR. HELD: I think you are getting the concept of radiative forcing wrong.

DR. KOONIN: Thank you. Please tell me.

DR. HELD: It's a hypothetical quantity how much the balance would change if you fixed temperature. It's not showing up on this picture.

DR. KOONIN: So, the temperature would be very different in 1700?

DR. HELD: Colder.

DR. CURRY: Colder, yes.

DR. KOONIN: So then, maybe the second question related to that, the ocean is cold. Isn't the ocean always warming as a result of, I mean, the long-term average heat flow from the surface of the ocean. Is it always in that direction?

I am trying to understand to what extent we believe the 0.6 (or 0.8

number in more recent analysis). Is it unusual or not?

DR. HELD: If it was sustained, the ocean would have a big temperature gradient.

DR. KOONIN: And it doesn't?

DR. HELD: It doesn't. We can go back to measurements of the deep ocean from the Challenger expedition and changes are in the hundredths of a degree.

DR. KOONIN: But I don't understand the mixing in the deep ocean.

DR. HELD: I think we have CFCs. We have radiocarbon. We have a lot of things we look at, not just heat. So, it's not that simple.

DR. COLLINS: We have a nice many choices, in other words.

DR. LINDZEN: The issue of deep water formation is still a little dicey.

DR. KOONIN: All right, other

questions for Judy, or comments from everybody?

DR. BEASLEY: Judy, you talked about solar influences. And I thought that Bill nailed one of the influences, which is changes in the solar. It's hard to understand how that would happen. I can really see that.

But then you mentioned a bunch of others that kind of surprised me, quite frankly. And so, for example, the magnetic field, I can't resist picking that one. Do you have a physics notion of what --

DR. CURRY: Okay, this is known/unknown. Some people with publishing papers speculating.

DR. ROSNER: This is based on they are certain that there is an effect or they have a physical process in mind that actually would do something?

DR. CHRISTY: There is the

cosmic --

DR. CURRY: Well, that is one example. I can't recite all the arguments off the top of my head. But people are publishing papers that present some intriguing possibilities. These are obviously not in the mainstream. But we have only really started looking at these kind of topics.

If you are interested, I can send you a list of papers I have been recently. But this is known/unknown category.

DR. LINDZEN: They all relate to particle processes influencing cloud condensation.

DR. ROSNER: Right.

DR. LINDZEN: And that has long been, that has been about 40 years that people have identified cloud condensation as the big magnifier, potentially.

DR. KOONIN: Scott?

DR. KEMP: Just a clarification question [to Dr. Collins]: Earlier you pointed out that, when we sent in our questions from the ocean-circulation chapter, which perhaps not everyone agrees with it, that the confidence in separating the variability from the trends was very good?

DR. COLLINS: The reason I was calling you on that was that statement was specifically ocean dynamics, not with thermal structure.

DR. KEMP: This is my question. Is that not related to understanding AMO and PDO trends? Or is it related to understanding AMO and PDO trends?

DR. CURRY: It's related to understanding how the whole processes on those time scales work. You can calculate the AMO and PDO out understanding the deep ocean.

But in terms of understanding the processes of how all this would influence sea ice, for example, you need to understand those circulations.

DR. COLLINS: Please understand when the projection is done on these modes of variability, what we are looking at is a mode of variability that we know the phase is left indeterminant.

Specifically, we asked whether or not the observations can be projected onto that mode with the phase as a degree of freedom in that projection.

So, I think there is a little bit of disagreement about whether or not the phase matters. But I just wanted to make -- I didn't want to pick nits over the issue. I want to be clear that the thermal structure is better understood.

And I do have a point to your question about modeling that I would like to come back to.

DR. KOONIN: I think Ben?

DR. SANTER: Judy, I guess my question relates to your claim that models can't capture AMO and PDO variability. You showed that the observations were uncertain by SST. You showed that there is some projection of model external forcing onto the modes of variability.

We also know that the human influence isn't just a simple linear trend. If you look at the ice core record, there are very large changes in anthropogenic sulfates over the 20th century. So, that deconvolution --

DR. CURRY: Is not external.

DR. SANTER: -- of external forcing on internal variability is not straightforward, very, very difficult.

DR. CURRY: Oh, I agree.

DR. SANTER: Given the short observational records, it is kind of difficult to uniquely determine what the characteristics of a 60-year mode of variability are.

DR. CURRY: I agree. I agree.

DR. SANTER: I think that's very difficult to make the statements models cannot do this when we have such a poor observational record.

DR. CURRY: The decadal, my comment there was based on the decadal simulations from CMIP5.

And I published a paper on it, Kim Webster and Curry, that basically showed that we didn't look at all the models, only the ones that were available early, but found that they were able to hang onto the AMO for about eight years even after being initialized. But even after one year, they weren't able to hang onto an initialized PDO.

So, that was the context that that statement was made. That said, just running a model for multicentury runs, you will get oscillations that resemble those.

DR. SANTER: But how well you capture the observed PDO or AMO is critical how you initialize how much subsurface information you get --

DR. CURRY: Absolutely, absolutely.

DR. SANTER: -- and how representative that is of what happened in the real world. And as you showed, even now that's problematic.

DR. CURRY: I agree. So, is the default position the models do it right or the models do wrong? I think the right interpretation is there is a whole lot of uncertainty in all of this.

DR. LINDZEN: There is a quip among oceanographers that the PDO is not an oscillation, it's not decadal, but it is in the Pacific. But one of the things that I think has to be remembered is the coupling of the atmosphere in the ocean, at least as far as heat goes, is a function of climate sensitivity. So, the higher the sensitivity, the weaker the coupling.

Gerard Roe put forward an interesting suggestion which we followed up, namely things like the PDO are pretty much an AR1 process and they have a time constant, response time associated with them much shorter than a decade. It's about 15 months, something like that in the data.

We went through the preindustrial historic runs in the CMIP. And that time scale for Pacific temperature, North Pacific temperature is about double what it is in the data. So, there is a suggestion the coupling isn't right.

DR. KOONIN: Bill?

DR. COLLINS: Scott, I wanted to return to the issue that you raised about model calibration and also what the models do. Just so everybody understands what a climate model is, the two major components traditionally were atmosphere and ocean. And I'm sorry. I am not dissing anybody. There are several other components of it.

They are solving the Euler equations for the fluid. They are solving fluid equations. There is a scale separation issue. You might imagine that we are dealing with a multiphysics situation that extends over 14 orders of magnitude.

So, we do have to parameterize just as one would have to in a multiphysics model of the operation of the universe. So, it's the same class of problem, almost the exact same major scales. And one must parameterize in that instance.

The issue that you raised about model preparation is a tricky one.

It is one that we are acutely aware of. And although we don't have multiple instantiations of the earth in the present day, we can use the paleo record as a means of probing how the models will do out of sample.

And the out-of-sample that we use there is the Milankovitch cycle, {*garbled transcription*: the variations of all the powers of the earth, the dry variations and solar insolation and its position on the surface as a function of seasonal cycle, and its particularly larger use of __ to the earth to disappear from the sun} and the orientation of the lower hemisphere landmasses to the sun during the summer.

And there is an extensive amount of literature on that work. The models are exercised extensively using the paleo record. There is a very large amount of work that is done to analyze models of the

samples.

I really want to make clear to you these are regimes which have a very different thermal characteristic than the present day. So, the models are exercised routinely and tested routinely by a very large cottage industry out of sample.

So, just please be aware of that before concluding that the record is so lousy over the last 30 years that we are in danger of extrapolating wrong.

And the final thing I want to point out to you, there is a lot more known about the physics on small scales that we haven’t been incorporating in the models because of computational limitations.

These problems are inherently too long in time because the model time scales are long. They are millennia. So, a lot of what we know and get emulated in process models and test against measurements has to be used in statistical fashion regarding the models.

But nonetheless, the understanding at process level is there, for example, the formation of stratus clouds.

So again, I don't want the community to come away with the fact that there is this -- model is resting on a large amount on mystery meat. They are not. There is mystery meat for sure.

But there is a very large amount of process modeling and process observations and backstop data that we can't incorporate simply because of computational limitations.

DR. KOONIN: Bill, as long as you raised the Milankovitch cycle, is there a way to phrase the Milankovitch forcing in watts per square meter so that one can compare it with the current anthropogenic influences? How would that comparison go?

DR. COLLINS: Well, in some cases, six watts.

DR. KOONIN: Six, roughly?

DR. COLLINS: Yes, it's quite large.

DR. KOONIN: But not an order of magnitude?

DR. COLLINS: No.

DR. CHRISTY: But high latitudes can be much larger than six.

DR. LINDZEN: Averaged over the globe and over the years, it's small.

DR. COLLINS: That's right, but locally --

DR. LINDZEN: Locally, it's 100 watts per meter squared in the Summer Arctic.

DR. COLLINS: It's big.

DR. KOONIN: I think we have reached a time when we should take a break. Why don't we break until

10:30 and pick up then. Thank you.

(Whereupon, a recess was taken.)

DR. KOONIN: Okay Ben, you're on.

DR. SANTER: Thank you very much for giving me this opportunity. I would like to talk about a couple of things. Since a number of your questions related to detection and attribution, I thought I would give an example of a recent study that Bill mentioned.

Then I am going to spend some time talking about the stasis. Since, again, it figured prominently in your questions, I wanted to present some work that is currently under review at Nature Geoscience about that; finally, some conclusions.

If I get time, I would likely to revisit this issue that turned up after Bill's presentation of multimodel ensembles and how we actually exploit them for what I do, detection and attribution work, and do differences in model quality really matter for the work do I? Do they affect our ability to identify an anthropogenic fingerprint on climate?

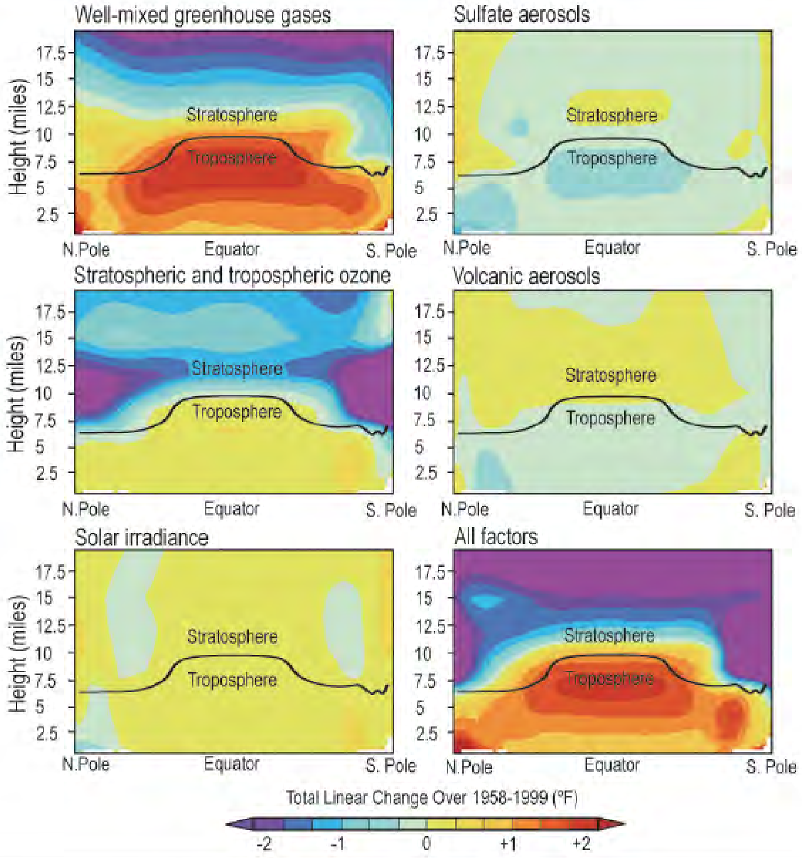
These[next page] are slices through the earth's atmosphere. These are all model calculations. This is from the so-called parallel climate model that was developed jointly at the National Center for Atmospheric Research in Los Alamos.

And in each case, this particular model was run with changes in just one factor alone, except in the bottom-right panel.

And that one factor was changes in well-mixed greenhouse gases, changes in volcanic aerosols, changes in the sun's energy output, changes in anthropogenic sulfate aerosols,

Different factors that influence climate have different “fingerprints”

Source: Global Climate Change Impacts in the United States (Karl *et al.*, 2009; modified from Santer *et al.*, *Nature* (1996)



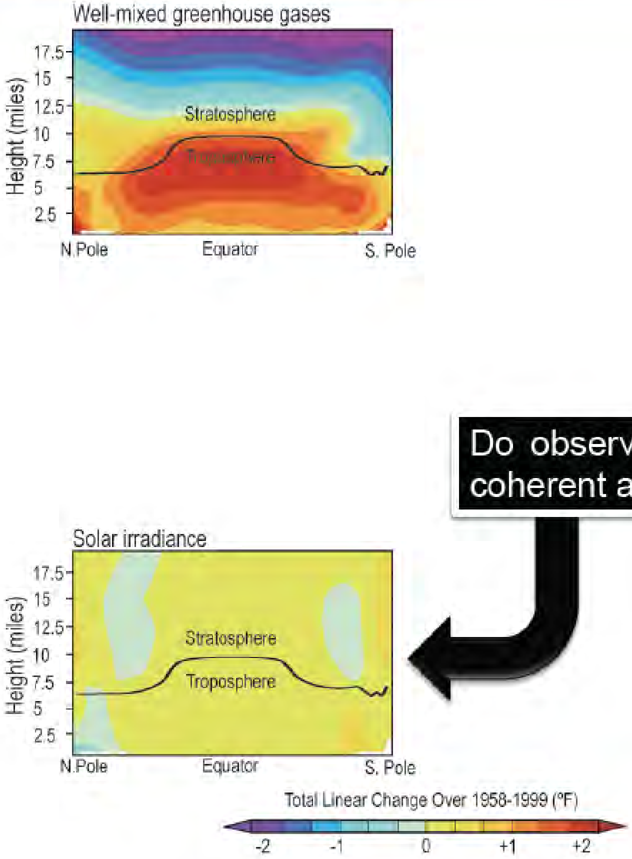
changes in tropospheric and stratospheric ozone according to our best understanding of how these things actual did change over the 20th century.

And I won't get into the details of the differences between these pictures here. But what they show you is that in fingerprinting, we don't just look at global mean changes.

We probe beyond one global mean number. And understanding a discriminatory power comes in looking at complex geographical, or in this case, altitudinal patterns of climate change.

Now, much of the attention has focused on these two patterns [next page], the vertical pattern of the response to human-caused changes in CO_2 and other greenhouse gases, and the vertical pattern of change associated with the dialing up of the sun.

Different factors that influence climate have different fingerprints



Do observations show vertically coherent atmospheric warming?

As Bill pointed out, although our estimates of solar radiance changes over the 20th century are uncertain, people think that there may have been some small low-frequency increase in total solar radiance over the 20th century.

If that happened, then we would expect to see heating throughout the full vertical extent of the atmosphere.

Now, we have known since the 1960s, since Suki Manabe, Warren Washington and others performed the first simulations where they doubled preindustrial CO_2 that the vertical fingerprint of response to human-caused changes in greenhouse gases is very different, and it involves this dipole as we discussed, the warming of the troposphere, the cooling of the stratosphere.

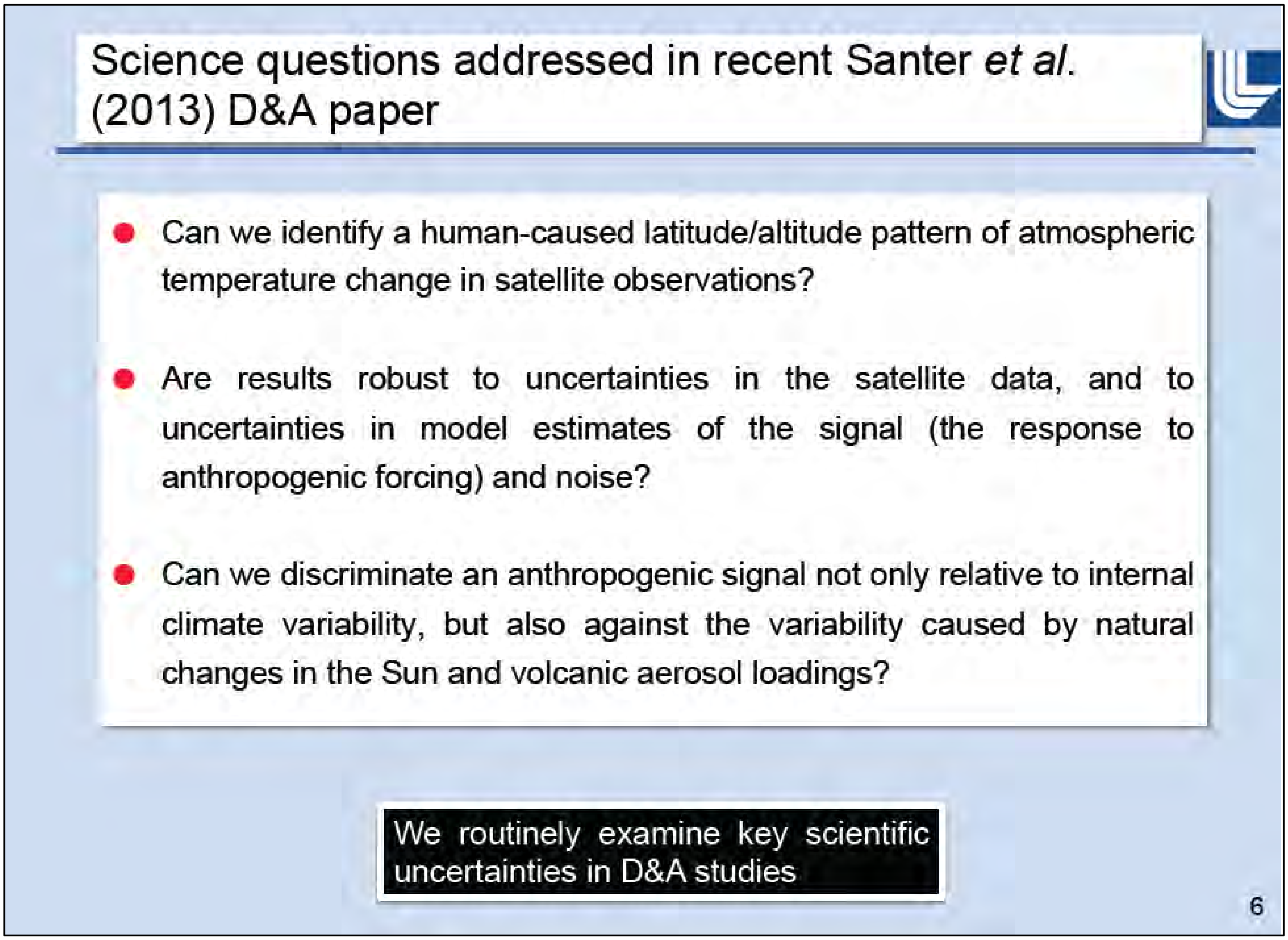
I just wanted to point out here that folks often say models are not falsifiable. They cannot make predictions that we can actually test. That's not true.

Back in the '60s when Suki Manabe and Warren and others did these simulations, we really didn't have the satellite data and we had sparse weather balloon data.

It was not possible to determine back then whether there were sustained multidecadal changes in the temperature of the troposphere and the stratosphere. These early pioneers could have been wrong. I will try and convince you that they were not.

So, one of the questions that we will get onto is, do observations actually show vertically-coherent atmospheric warming, do they look like sun fingerprint or do they look like the CO_2-increase fingerprint?

I am going to give you an example of a recent study. [next page]



Science questions addressed in recent Santer *et al.* (2013) D&A paper

- Can we identify a human-caused latitude/altitude pattern of atmospheric temperature change in satellite observations?
- Are results robust to uncertainties in the satellite data, and to uncertainties in model estimates of the signal (the response to anthropogenic forcing) and noise?
- Can we discriminate an anthropogenic signal not only relative to internal climate variability, but also against the variability caused by natural changes in the Sun and volcanic aerosol loadings?

We routinely examine key scientific uncertainties in D&A studies

Bill mentioned this. It came out of a few months ago in PNAS. And we wanted to look at the following science questions.

The first was revisiting some of the early work we did about 15 years ago. Can we identify some human-caused pattern of climate change in the vertical structure of atmospheric temperature?

Another question was uncertainties. Judy has raised the question of uncertainties. And there is some, I think, misperception that detection and attribution studies sweep these uncertainties under the carpet. I will try and convince you that that is not the case.

In fact, we wouldn't be able to get this kind of work published if we did not routinely and comprehensively look at uncertainties in model estimates of the response to forcing, uncertainties in model estimates of internal variability and at uncertainties in the observations themselves.

And the real opportunities to do that now, as I will show you in a few minutes, one of the groups we work with, Remote Sensing Systems in Santa Rosa has developed an ensemble of observations for atmospheric temperature.

So, they played through all of uncertainties in data set construction, how you account for satellite orbital drift, the impact of that drift on the sampling of Earth's diurnal cycle, how you account for inter-instrument calibration biases using a nice Monte Carlo approach.

And they generate a 400-member ensemble model of observations that you can use in this kind of fingerprinting work and see whether your ability to detect is sensitive to those uncertainties in the observations.

Another thing that we do in this study that is a little unusual, typically fingerprint work tests against internal variability alone that they estimate from models.

We are also going to ask the question, given these new world-without-us simulations in CMIP5, so, the simulations that Bill mentioned that have natural external forcing, the sun and volcanos from 1850 through to the present and some of them over the last millennium.

You can ask this sort of worse-case scenario statistical significance testing question; could larger solar radiance changes or the recovery from larger volcanic eruptions that have occurred over the past 1,000 years screw up anthropogenic signal detection?

Could we misidentify that

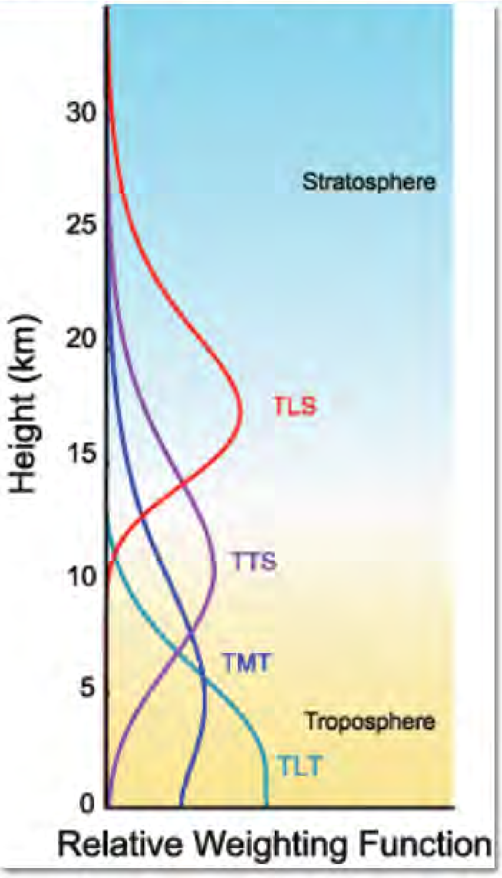
dipole pattern of tropospheric warming and stratospheric cooling? Could it really be due to something else?

Okay,[next page] we are going to do all of this in satellite space. So, as John Christy, I am sure, will talk about later, the microwave sounding unit estimates of atmospheric temperature change which he and Roy Spencer pioneered look at the temperature changes over broad layers of the atmosphere based on the microwave emissions from oxygen molecules.

We are going to be working in this vertically-smooth space looking at the temperature of the lower troposphere. That's the cyan curve, the temperature of the mid to upper troposphere and the temperature of the lower stratosphere.

And what we have done is we have actually calculated synthetic satellite data from the model

Our fingerprint study uses zonal-mean changes in the temperature of broad atmospheric layers



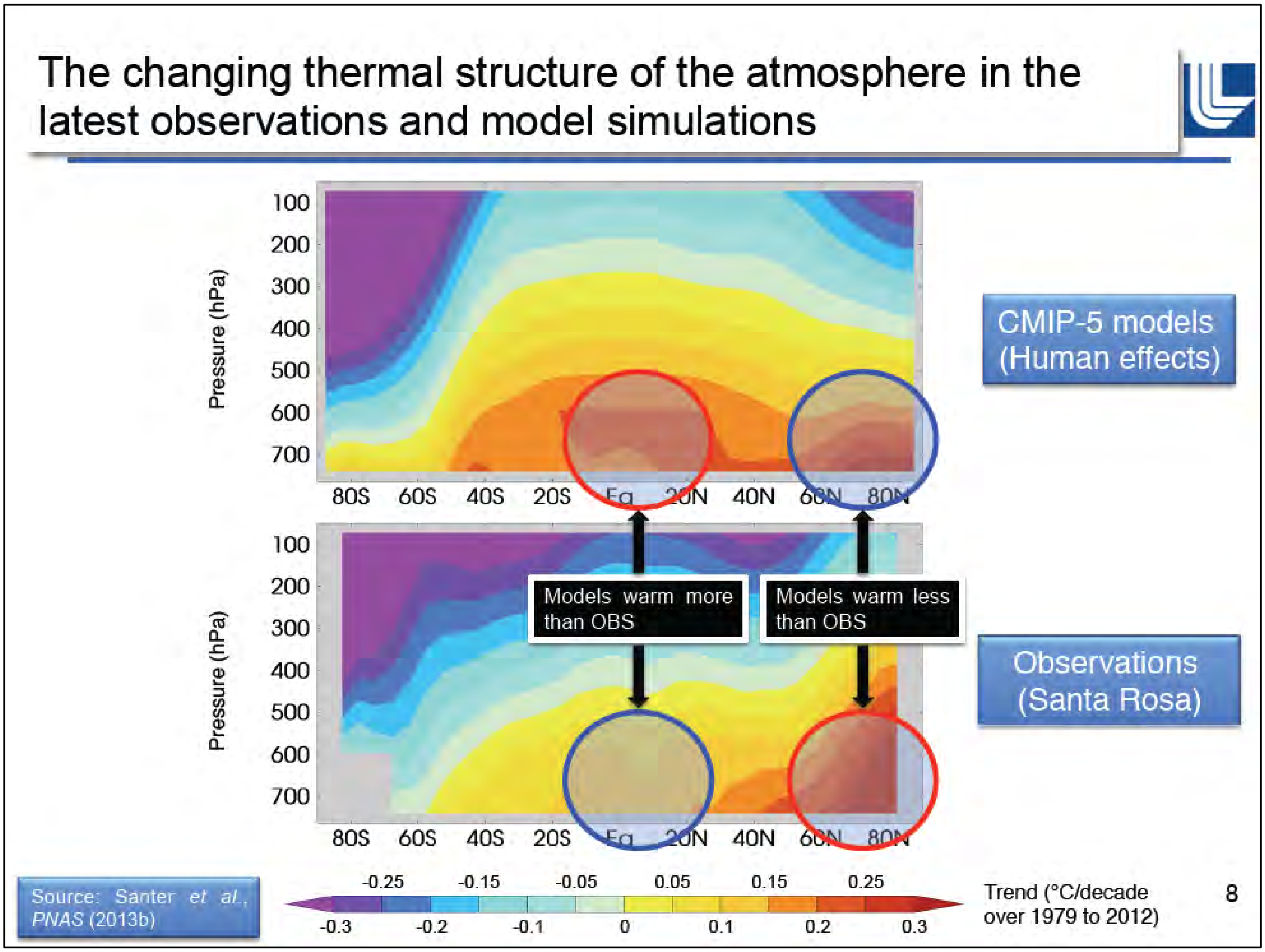
simulations in order to facilitate a comparison between the two.

This [next page] was the title slide. This shows in that vertically-smooth space, then, the temperature changes in the average of 28 CMIP5 models. These are from simulations with human effects.

These are over the full satellite era, so 1979 through to 2012. And the bottom panel is the publically-available version of the Santa Rosa Remote Sensing Systems observations.

You can see that both show this dipole, first of all, this cooling of the stratosphere, warming of the troposphere over this 34-year record. But there are some noticeable differences.

Over the Arctic, and this is true both of the Santa Rosa data and the University of Alabama data, the observations warmed more than the



The changing thermal structure of the atmosphere in the latest observations and model simulations

CMIP-5 models (Human effects)

Observations (Santa Rosa)

Models warm more than OBS

Models warm less than OBS

Trend (°C/decade over 1979 to 2012)

Source: Santer *et al.*, *PNAS* (2013b)

8

models. Over the tropics, it is the other way around.

In the lower troposphere, the models actually warmed more than the observations. And we can get into possible causes for these smaller-scale differences later.

DR. KOONIN: Ben, a clarification?

DR. SANTER: Sure.

DR. KOONIN: When I read IPCC AR5, I learned about scaling factors in detection and attribution. Do these graphs have scaling factors in them?

DR. SANTER: No.

DR. KOONIN: So they are just raw out of the box?

DR. SANTER: These are just, in the top panel, the multimodel average. So, these are the least-squared linear trends over this 384-month period of time, January 1979 through to

December 2012, models and observations.

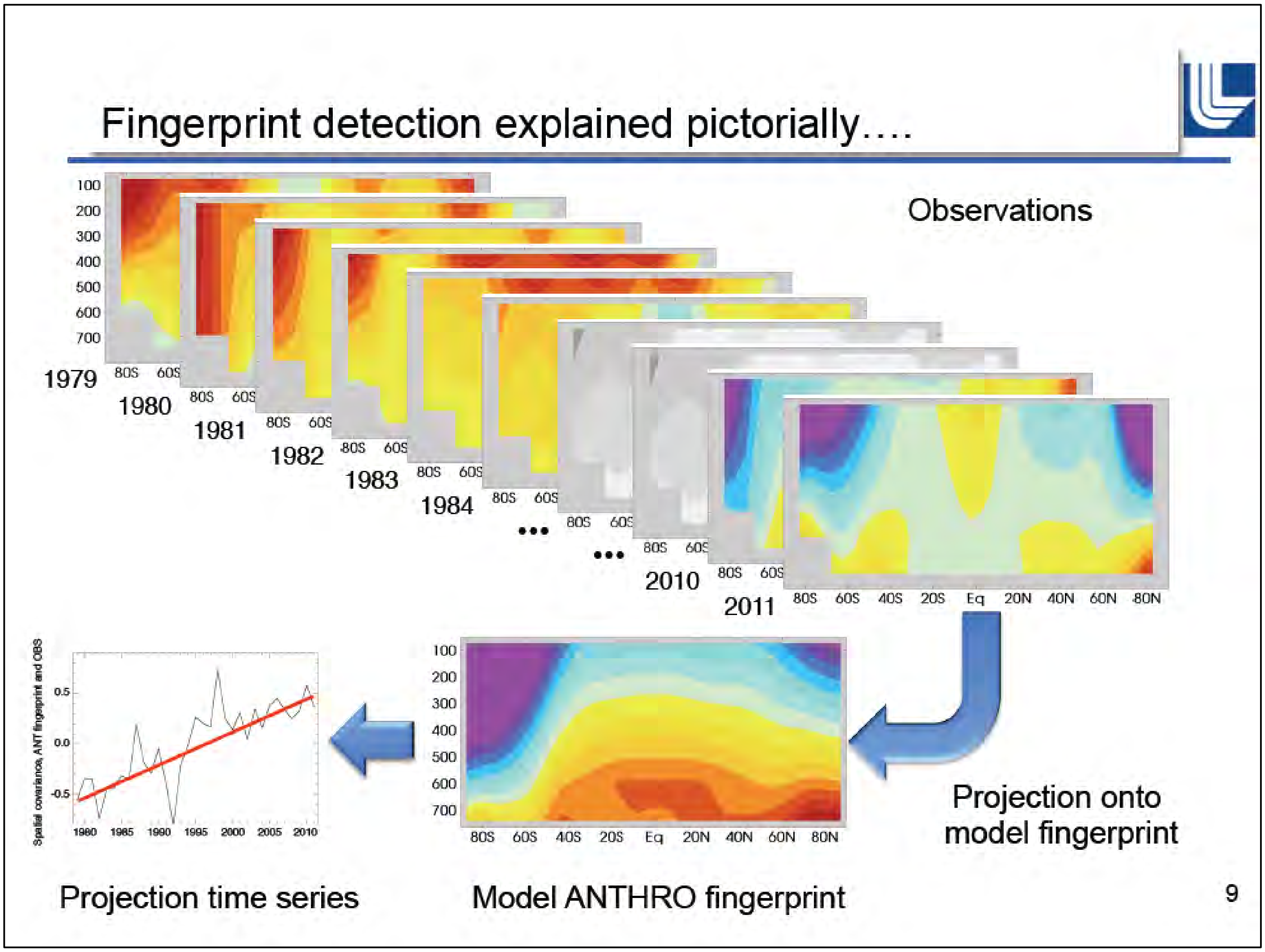
The models as an average over 41 realizations can, I believe, performed with about 28 different parts.

DR. KOONIN: We will have more discussion in the question period, but thanks. That's good.

DR. SANTER: Okay. So, just briefly then, how do we actually compare models and observations? This [next page] is fingerprint detection explained pictorially.

Imagine we have from these 28 models some estimate of the response to total anthropogenic forcing. And we are going to search for that in the time-varying observational record.

So, here we have observational microwave sounding unit data from '79 through to, in this case, 2011. We calculate some major spatial



Fingerprint detection explained pictorially….

Observations

Projection onto model fingerprint

Projection time series

Model ANTHRO fingerprint

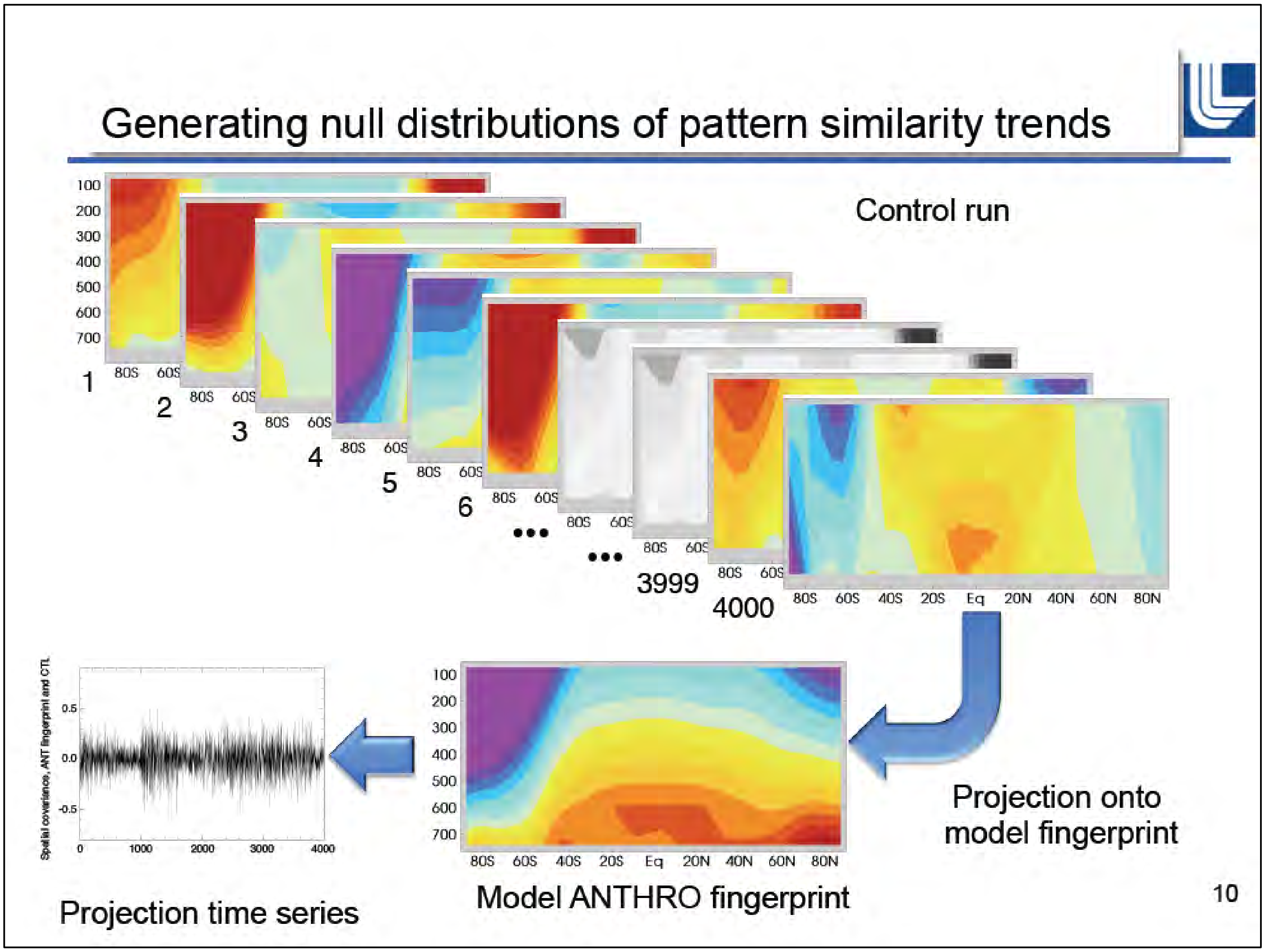
similarity between the models and the observations. And that gives us this -- you can't see it very well -- that gives us this trend.

And there is a lot of wiggles. We understand some of those wiggles. You can see around '91, '92, there is a big dip down. That is because of Pinatubo. Pinatubo warmed the stratosphere, cooled troposphere.

That's the converse of the expected fingerprint. There's a bump in '98. That is the big El Niño in '97, '98. So, we understand a lot of the variability superimposed on that trend.

But the issue is, is that trend, say, over this 34-year period of records, statistically significant? And in order to address that question, we generate null distributions of trends. [next page]

With these models, we have control simulations with no changes



Generating null distributions of pattern similarity trends

Control run

Projection onto model fingerprint

Model ANTHRO fingerprint

Projection time series

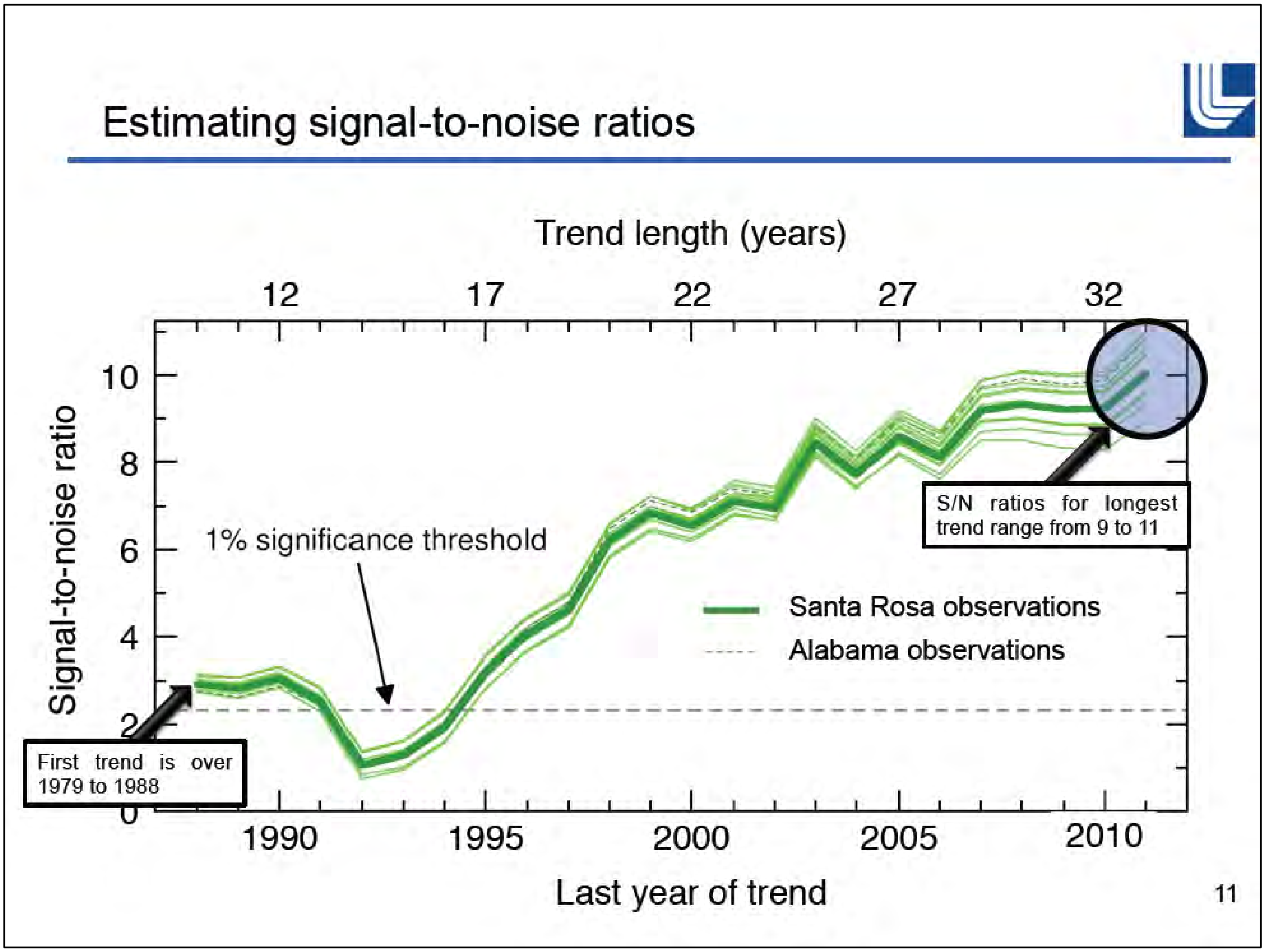
in external forcing. We are going to use results from over 4,000 thousand years of simulation. And you can do the same thing.

You can look at the pattern agreement that you might expect by chance between the model unforced variability and the searched-for anthropogenic signal.

And [next page] you get some projection time series. You can look at trends on any time scale in that projection series, and then can you look at signal to noise.

You can look at the observed trend that I showed you in the previous picture relative to these unforced trends and pattern similarity. And that enables you to look at signal to noise as a function of time scale.

Now, the first trend is for the first ten years, '79 to 1988. Since the satellite record starts in 1979,



Estimating signal-to-noise ratios

First trend is over 1979 to 1988

S/N ratios for longest trend range from 9 to 11 the longest trend is over the full period of satellite record, '79 in this case, through to 2011.

The light green lines there show you these realizations from the Santa Rosa results. I have used the five, ten, 15, percentiles of that 400-member ensemble of observations to be able to look at the uncertainty in the observations and how that projects onto our ability to detect.

The key thing here is that if you look over the full period of the satellite record, remember that plot that I showed you before with the tropospheric warm and stratospheric cooling, natural internal variability can't give you that. The signal-to-noise ratio is nine to eleven.

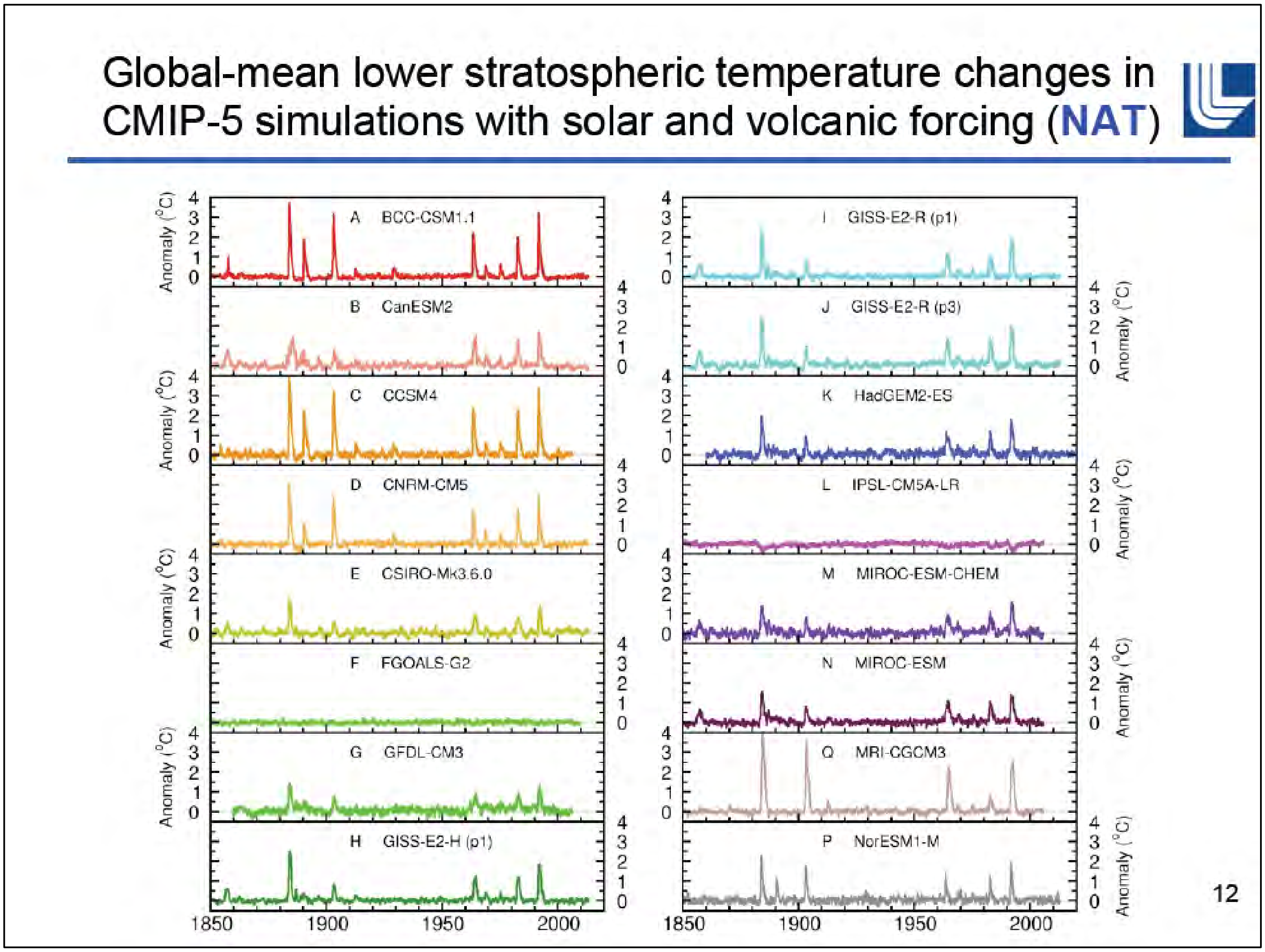
It's kind of interesting to compare that, say, with the big discussion that we have had in the last year or two about the significance of a five-sigma result with the detection of the Higgs boson.

So, basically this says that model-based estimates of internal variability, if credible, cannot give you this kind of result.

You can also, then, [next page] look at the second question, whether model responses to solar forcing and very large volcanic eruptions could mimic the kind of things we see in the observations there.

We know we did have two big volcanic eruptions, El Chichón, Pinatubo. We know we have had changes in solar radiance. How about if we looked in the deep past, if we looked at things like Krakatoa here?

This is the stratospheric temperature changes here, MSU Channel 4 from 16 different models. And you can see that most of these have some representation of volcanic aerosol



Global-mean lower stratospheric temperature changes in CMIP-5 simulations with solar and volcanic forcing (**NAT**)

changes and solar changes from 1850 through to the present.

So, you can use this as your noise basis for trying to do signal detection. And [next page] you can look back deeper in time at these last millennium runs that typically start in 850 AD and have very large eruptions like this in 1258 here.

And again, when you have a big eruption, you warm the stratosphere. You cool the troposphere. You have some recovery time scale, which we will get into in discussion of the stasis.

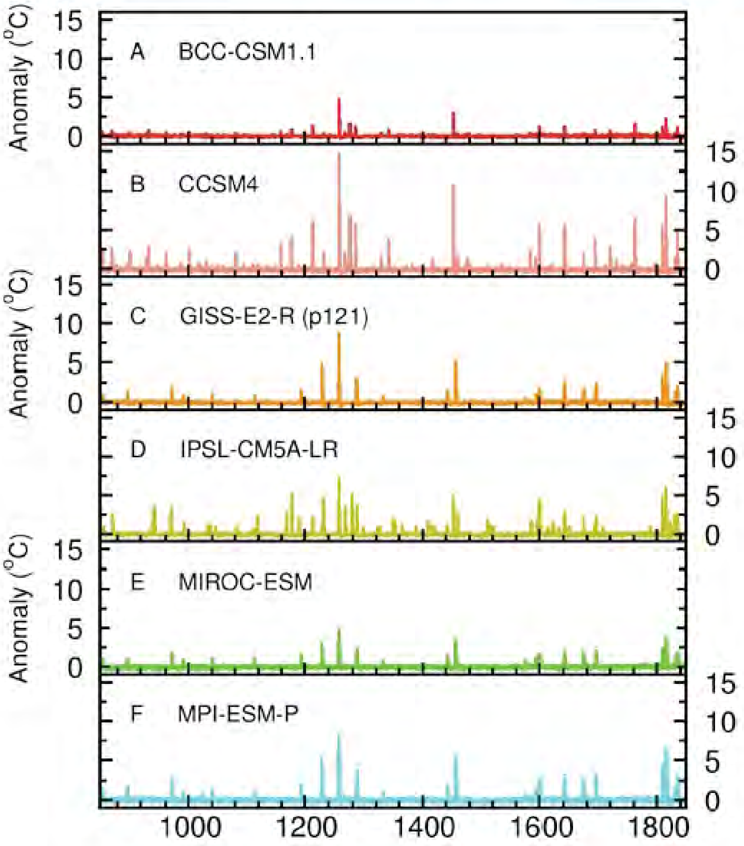
So, the question is whether that warming that you see in the recovery phase could cause you to misidentify anthropogenic fingerprints.

DR. KOONIN: So, just hold on for a minute. Go back.

DR. SANTER: Sure.

DR. KOONIN: The models have

Global-mean lower stratospheric temperature changes in CMIP-5 "Last Millennium" simulations (**P1000**)



different responses?

DR. SANTER: Very different responses, yes, that's true.

DR. KOONIN: Right. So, do you just average them all together?

DR. SANTER: No, we use all of them. We use all of them. So, we concatenate these control runs and we look at all of these model-based noise estimates.

I should say that for the control runs what we do is we make sure that, since model control runs are a different length, we have the same length control run from each model so that we are not preferentially giving weight to one model relative to another.

DR. KOONIN: I look at CCSM4, for example. So, it's pretty responsive to some of the --

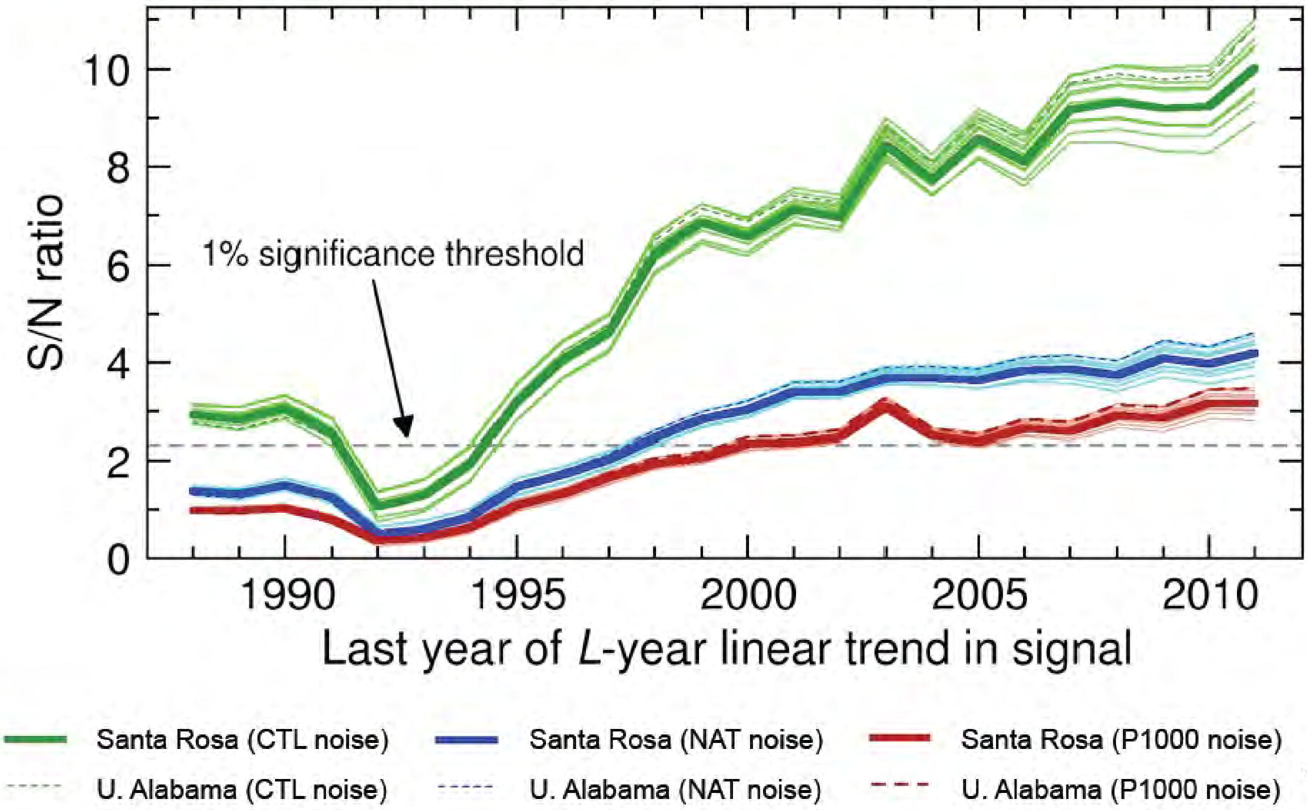
DR. SANTER: Well, that's right. It has global mean 15 degrees C warming of the lower stratosphere after this big eruption in 1258, and about four degrees Celsius cooling of the global mean troposphere.

So, that's a very big eruption. And the question is, could that interfere with anthropogenic signal detection and recovery from that very large eruption? And the answer is no.

So, [next page] the blue and the red now are testing against this world without us, but with solar and volcanic forcing. The blue lines are from 1850 through to the present, these naturally-forced simulations, and the red is the last millennium simulations.

Since the red has very much larger volcanic eruptions and larger solar radiance changes around the time of the modern millennium, signal to noise goes down, but it's still in every case above the one-percent

Estimating signal-to-noise ratios



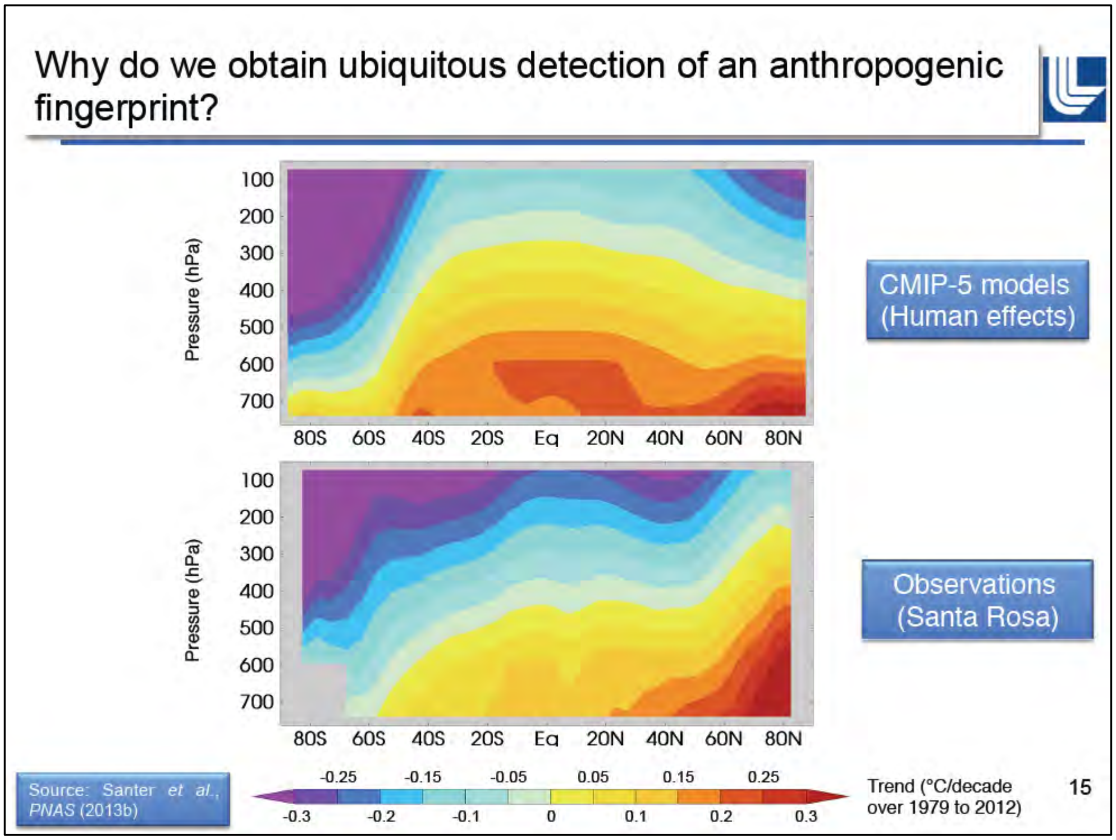
significance threshold.

Incidentally, this little dip here [indicating 1992] is the effect of Pinatubo in the observations which, again, warmed the lower stratosphere, cooled the troposphere. That's the converse of the expected anthropogenic signal.

So,[next page] the question is why? Why do we get those results? Again, let me take you back to these patterns here. And they do show latitudinally pretty coherent cooling of the stratosphere and warming of the troposphere.

And [next page] it turns out that when you look at these 28 control runs and you do an EOF analysis and look at the dominant modes of variability, they don't do that.

They don't generate sustained warming of the troposphere and cooling of the lower stratosphere on these long time scales, nor do the naturally-forced runs. They can't



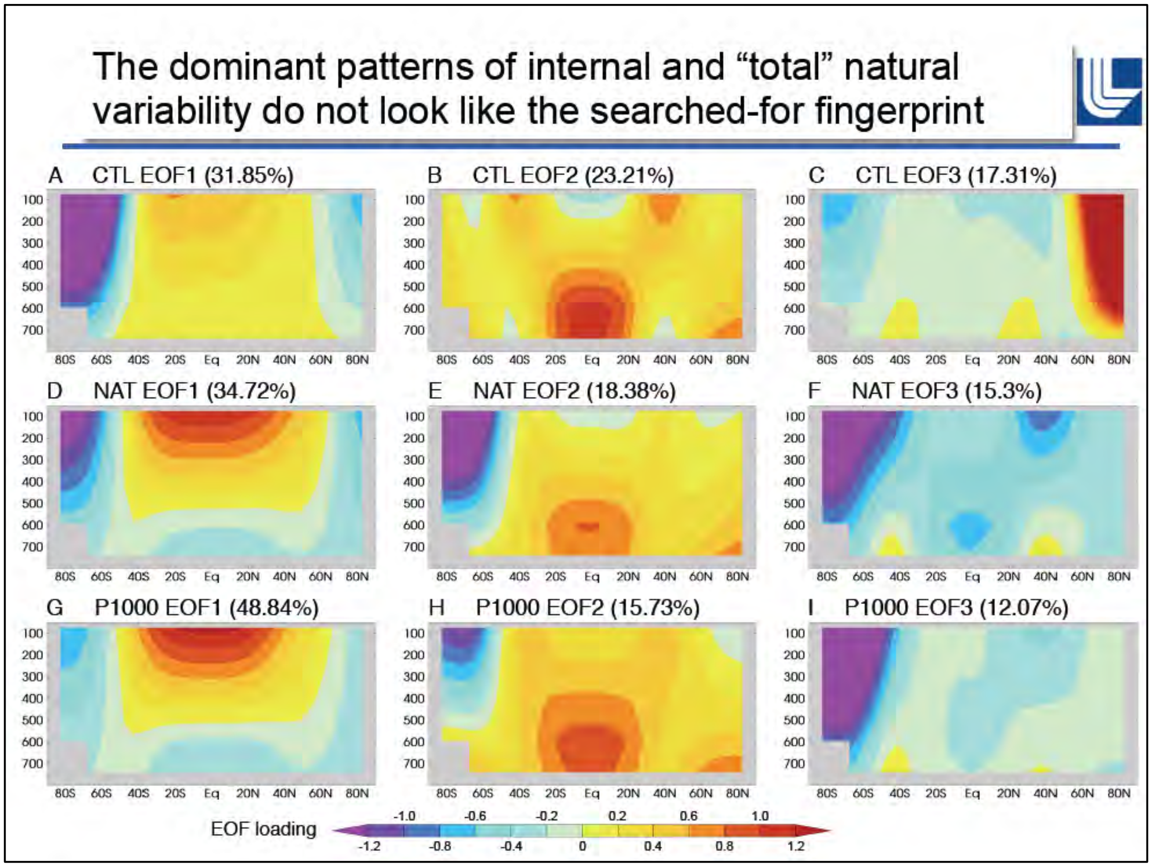
Why do we obtain ubiquitous detection of an anthropogenic fingerprint?

CMIP-5 models (Human effects)

Observations (Santa Rosa)

Source: Santer *et al.*, *PNAS* (2013b)

Trend (°C/decade over 1979 to 2012)



The dominant patterns of internal and “total” natural variability do not look like the searched-for fingerprint

A CTL EOF1 (31.85%)

B CTL EOF2 (23.21%)

C CTL EOF3 (17.31%)

D NAT EOF1 (34.72%)

E NAT EOF2 (18.38%)

F NAT EOF3 (15.3%)

G P1000 EOF1 (48.84%)

H P1000 EOF2 (15.73%)

I P1000 EOF3 (12.07%)

generate those kind of patterns, and nor do the last millennium runs.

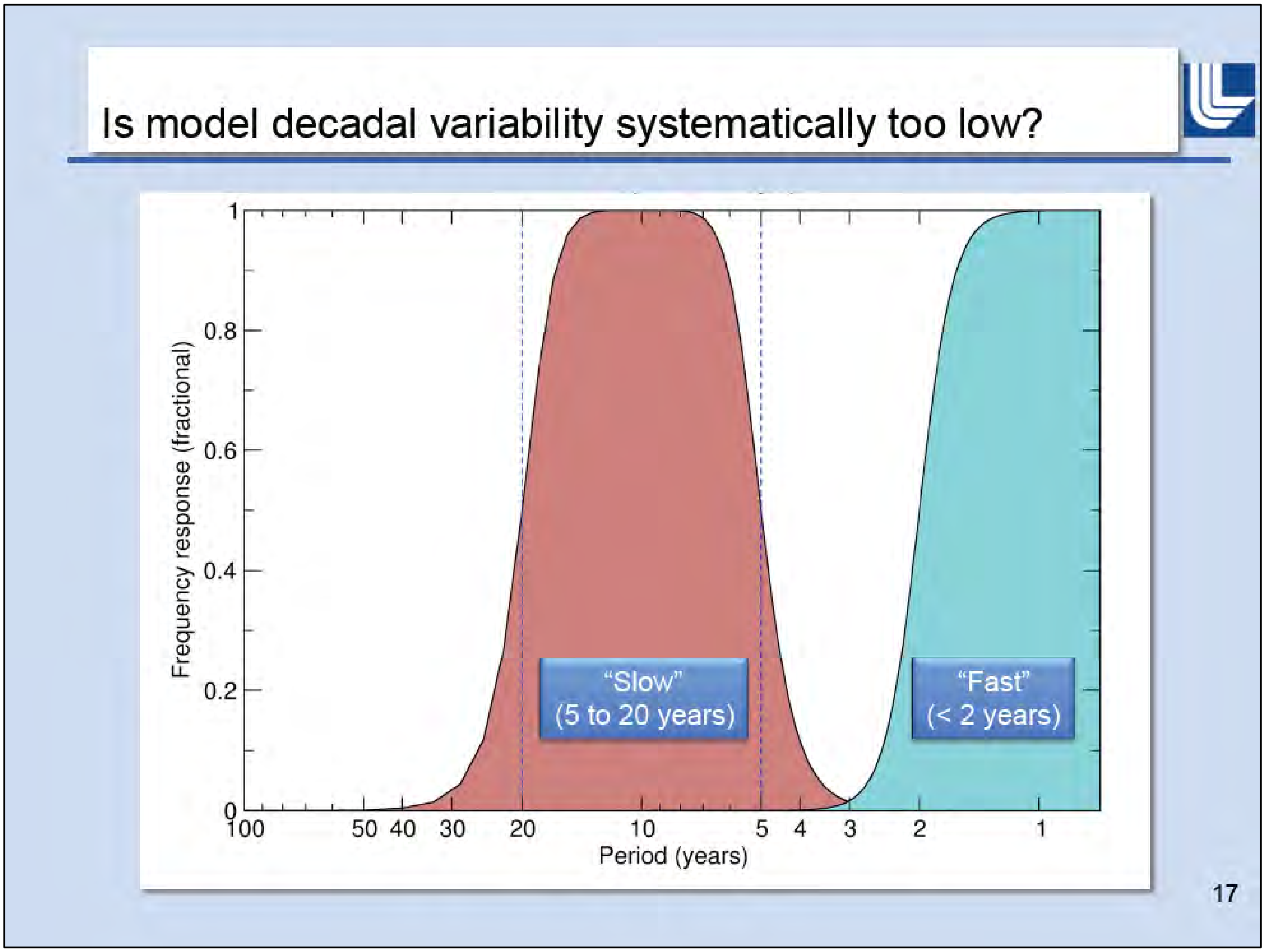
Now, Judy raised the question here of model-based estimates of internal variability. They are a crucial underpinning this have work.

So, if we systematically underestimated the true unforced variability, particularly on these multidecadal time scales that are crucial to the identification of an anthropogenic system, then the signal-to-noise ratio would be biased high. It would be systemically too high.

So, let's look at that. [next page] We have done some band-pass filtering for ocean surface temperature, tropospheric temperature, stratospheric temperature.

Basically what we have done is we have windowed in on variability of time scales of ten years. Recall, again, that the microwave sounding

Is model decadal variability systematically too low?



units record is about 35 years long now. So really, we can't go much longer than that.

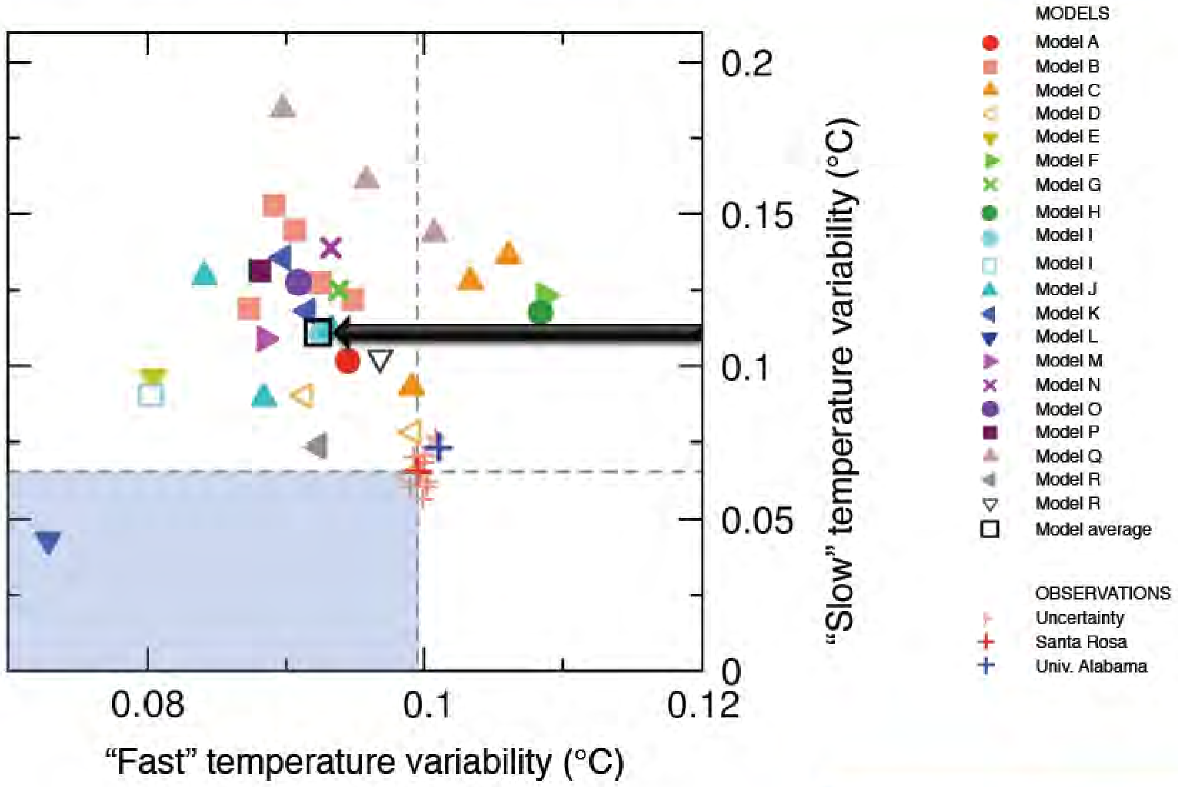
And we focused on the variability of the time scales below two years. What we did is we band and high-pass-filtered all of the model and observational data.

And I am going to plot now,[next page] this is, again, global mean tropospheric temperature, the sub-two-year time scale variability against the five-to-20-year time scale variability.

The cross hairs are on the observations. And if we were in that blue quadrant of doom, we would be in trouble because that would mean that the model systemically underestimated the amplitude particularly of the crucial low-frequency variability.

Let's start adding things in now. So, here you see the 400-member ensemble from Santa Rosa. Most of

Do CMIP-5 models underestimate the observed “slow” variability of tropospheric temperature?



Source: Santer *et al.*, *PNAS* (2013a)

the uncertainty is in the low-frequency direction there pretty much as expected, because that's where different decisions in how to adjust for satellite orbital drift effects really become manifest.

Now we are going to start adding in model results. You can see, as Bill mentioned, that some of these have multiple realizations. That's why I say in the case of Model B, there are five squares there.

Adding in a bunch of models still, you can see that only one is in the quadrant of doom and that on average, the CMIP5 multimodel average actually overestimates the low-frequency variability by about 40 to 50 percent.

We found this for SST as well as Bill showed from the spectrum, there is no real evidence, at least on these kind of time scales, of some fundamental error in the amplitude of variability.

That, of course, doesn't get at the pattern-of-variability issue which is equally important in detection and attribution work.

Okay, the stasis.
[charts following have been removed at Santer’s request to avoid prepublication release]

So this, again, is global mean change in lower tropospheric temperature from both Remote Sensing Systems and from John's group.

And you can see that it's there in tropospheric temperature, too. This is not something that is confined to surface temperature.

And remember Bill mentioned and Judy mentioned, I think, the Cowtan and Way paper that looks at these coverage issues for surface temperature.

If that alone were the explanation for the stasis, then MSU, which pretty much has global coverage, would not show this kind of behavior, but it does.

So, science questions here. In the observations, what factor or factors have contributed to the stasis in tropospheric and surface warming and why are the tropospheric temperature trends in CMIP5 models, on average, larger than those observed over the stasis period?

You can see if we go back to this figure that -- well, you actually can't see this pink envelope very well. You can see that a couple of models actually simulate behavior that looks reminiscent of the stasis, but very few. Most of them are systemically above.

DR. KOONIN: Well, actually, do we know that? Because it may be one model is at the bottom of the distribution for some years and then goes to the high end of the distribution for other years and so on.

DR. SANTER: It is a couple of different models. It isn't multiple realizations of the same model. And I can tell you which models are at the bottom of the distribution later.

So, as Bill mentioned, a number of different explanations have been posited. One is model sensitivity errors, and John had mentioned this in his Congressional testimony.

Another is forcing errors. Even the perfect model, the hypothetical perfect model with perfect representation of all the physical processes that drive the real-world climate system, if you give it the incorrect external forcings, it will get the wrong spatiotemporal resolution.

There are concerns about the stratospheric ozone depletion. We think that, on average, the models that specified stratospheric ozone changes over the observational period probably underestimated the changes, even in the tropics down to about 150 to 200 hectopascals.

There are concerns about volcanic aerosols that we will get onto in a minute, and their representation after Pinatubo in 1991. There are concerns that Bill mentioned about anthropogenic sulfate aerosols and possible underestimate of Chinese sulfate aerosol pollution.

And there are concerns about solar forcing in that most of these models do not have the unusually broad solar minimum over the last solar cycle.

Also, there are concerns about residual errors in observational temperature data both in the tropospheric temperature data and in the surface temperature data.

And then there is this issue of an unusual manifestation of natural variability in the observations associated with ENSO, the PDO or some combination thereof.

So, let's look at this first. Do ENSO effects explain the stasis? So, what we did is we used some iterative regression-based method to remove ENSO effects.

Turns out you have got to be a little careful because there's co-linearity between ENSO and Volcanos. And that matters over this period of record.

So, if you just plug everything into some multiple regression framework, you get the wrong answer. So, using this method, we remove ENSO effects and the hiatus is still there.

So, at least when you statistically remove ENSO effects, you cannot fully explain this discrepancy between models and observations or the failure of the observations to warm much over the last 15 years.

All right, next question.

DR. KOONIN: When you remove ENSO effects, do you do that model by model to the extent that --

DR. SANTER: Yes, sir.

DR. KOONIN: -- that the models show ENSO?

DR. SANTER: We do it model by model for every model, and we do it in a whole bunch of different ways. It turns out that one of the unknowns is this what we call tau, the recovery time scale, which is related to the transient climate response and the equilibrium sensitivity.

We do that removal both with each model's individual estimated value of tau based on their estimated equilibrium sensitivity from the four-time CO_2 runs.

And we also do it with stipulated values of tau that span an ECS range of one degree to about five and a half degrees. It doesn't make much difference. That uncertainty in tau does not make much difference to these results that I am going to show you here.

So, the next question is, do CMIP5 models capture the observed changes in warming rate after El Chichón and Pinatubo, El Chichón, again, in 1982, Pinatubo in 1991?

What we are looking at here is maximally-overlapping ten-year trends. So, it's another noise filter. We have reduced some of the noise by removing ENSO effects. Now we are going to look at overlapping ten-year trends.

And a gentleman at a meeting at the Royal Society in London presented something like this for surface temperature and was saying well, we really need to get away from looking just at one specific period. We have got to look at many, many different overlapping trends.

And he obtained very similar twin peaks there without any understanding why they were there. Well, we do know why they were there.

So, the models do capture, on average, at least, all of the slowdown in warming after both El Chichón and Pinatubo and the gradual recovery thereafter.

You can actually see that the conditional probability of getting a ten-year warming trend around this time and this time is obviously critically dependent on this, on where you start the ten-year trend relative to the peak volcanic cooling.

So, if you are starting at the peak volcanic cooling associated with Chichón or Pinatubo, you have background anthropogenic forcing acting in concert with this recovery phase.

DR. KOONIN: And again, in these comparisons or these analyses of models, no scaling factors for aerosols, nobody reduces the aerosols by 20 to 30 percent model by model to match something?

DR. SANTER: No.

DR. COLLINS: Not that I know.

DR. SANTER: No. And I think that for the sulfate aerosols -- correct me if I am wrong here, Bill and Isaac -- what was done is that most groups used the same history of sulfate aerosol emissions.

That is put through some atmospheric chemistry transport model in order to calculate spatiotemporal changes in atmospheric burdens of sulfur dioxide.

DR. KOONIN: Then I am really confused about something in the IPCC. I will show you later.

DR. SANTER: Maybe we can get

onto that a little bit later and I can finish this.

DR. CURRY: The AR4 was, I think, a lot squishier about what they used for aerosols, although this was tightened up in AR5.

DR. KOONIN: Okay, good.

DR. SANTER: So, in the final third of the satellite record, this good agreement that we saw over the first two-thirds breaks down. Why? What is going on there?

How can you, on the one hand, successfully capture the amplitude and phase of the temperature response to El Chichón and Pinatubo with the first two-thirds of the record, but get this divergence over the final third?

And if this divergence is really due to some fundamental errors in model physics and ocean heat uptake and, therefore, in sensitivity as, say, John Christy, has posited, then why don't you see that in the response to El Chichón and Pinatubo? It seems like a real conundrum.

So, let me try and convince you, Judy. You said you weren't really very convinced by post-Pinatubo or recent volcanic aerosol forcing.

This is a record of stratospheric aerosol optical depth; beautiful measurements. So, these things look at the occultation of sunlight and moonlight at different wavelengths.

This is roughly from about 15 to 35, in the case of Sato, kilometers, 15 to 40 kilometers in the case of Vernier, et al., a whole bunch of different satellite instruments that are spliced together in different ways.

And you can see that each of these vertical lines is an eruption. The solid lines are tropical eruptions within 20 north to 20 south. The dashed lines are extratropical eruptions.

And this was the assumption in CMIP5 that, after Pinatubo, stratospheric aerosol optical depth decayed to zero or to background values by the end of the 20th century.

Now, that's not what happened in the real world. In the real world, there were, as Tim Barnett likes to call it, a swarm of over 17 eruptions with a volcanic explosivity index of three to four after Pinatubo. So, this is an instance of a systematic error in volcanic aerosol forcing.

All right, this is now looking in the tropics specifically at stratospheric aerosol depth. Again, vertical lines are eruptions. You can see the signatures of these early 21st-century eruptions across the electromagnetic spectrum.

So again, we see them in the visible in stratospheric aerosol optical depth. Look at the two largest here, at Tavurvur in Indonesia in 2006, and Nabro in Africa in 2011.

You can see that the increase in stratospheric aerosol optical depth leads to this increase in net reflected shortwave at the top of the atmosphere. That backscattering is the primary signature we are picking up here.

You can see it, too, in the MSU data in the tropics after you statistically remove ENSO effects. Again, after each of these eruptions there is cooling of the lower troposphere in the tropics within, say, three to six months. So, that's in the microwave.

Now, you can also ask the question well, okay, how about if I look at the correlation between stratospheric aerosol optical depth and lower tropospheric temperature, we lag things because there is some lag between forcing and response?

How about if we look at the instantaneous correlation between stratospheric aerosols and reflected shortwave at the top of the atmosphere? That is pretty much instantaneous.

And because volcanic activity non-stationary, we look at these things in 60-month sliding windows. And what you see is that, during the Pinatubo period, you have this very strong, highly significant negative relationship where aerosol optical depth leads to cooling.

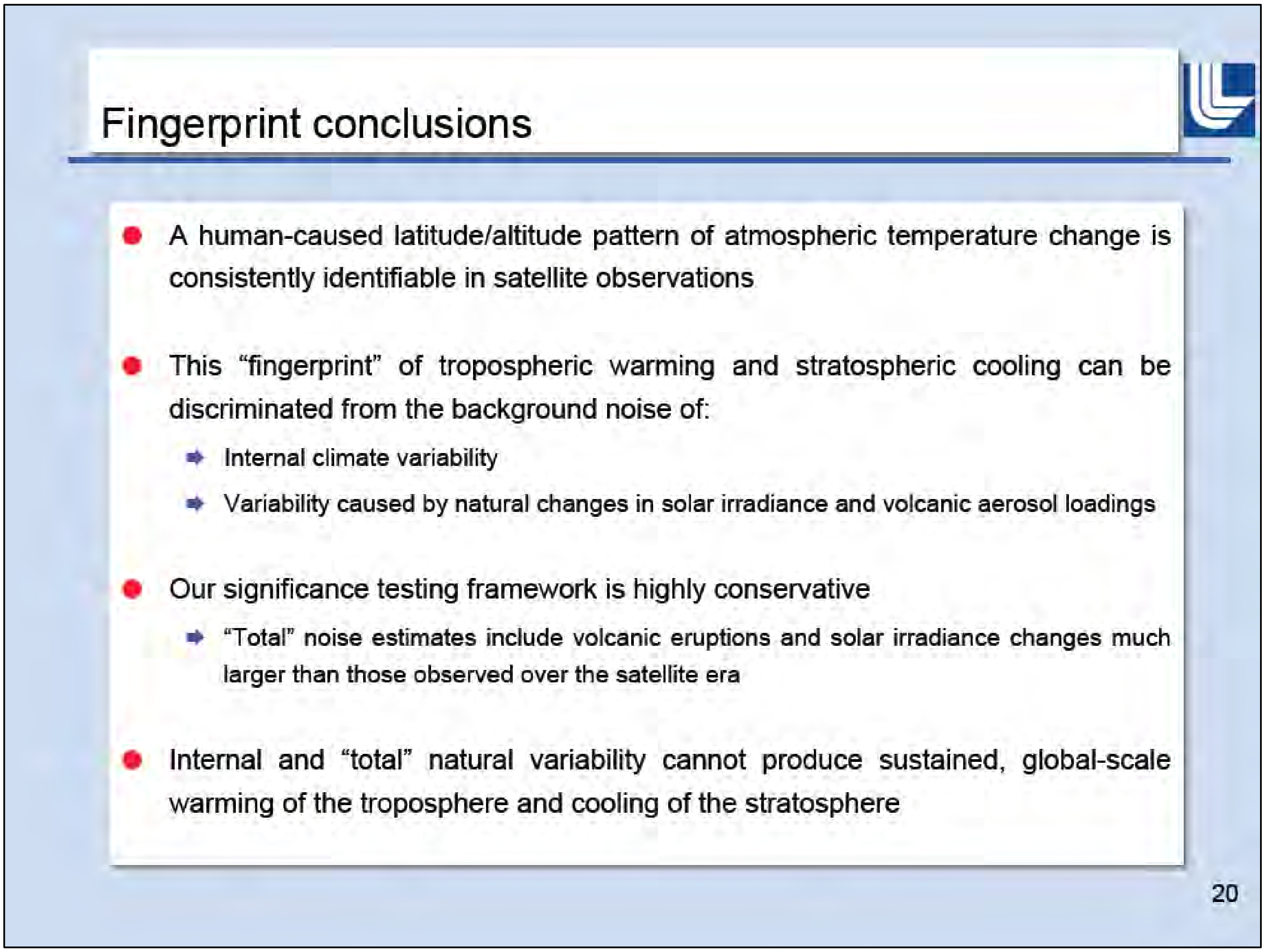
But even in the most recent period here where you have these big three, Manam, Tavurvur and Nabro, we have highly significant cooling of the lower troposphere associated with increases in aerosol optical depth.

And we also get significant results. This is the series shortwave record there which only goes back to March 2000. So, you can't push it back as far in time.

But there, too, we see statistically significant relationships between the recent aerosol optical depth changes and the shortwave changes. So, there clearly is some signal there.

Okay, conclusions. [next page] From the fingerprinting, we find that some human-caused latitude/altitude pattern of atmospheric change is consistently identifiable in the satellite observations.

And we can discriminate this not only from the background noise of internal variability in the models, but also from the larger total variability caused by changes in volcanic forcing and solar radiance.



Fingerprint conclusions

- A human-caused latitude/altitude pattern of atmospheric temperature change is consistently identifiable in satellite observations
- This “fingerprint” of tropospheric warming and stratospheric cooling can be discriminated from the background noise of:
  - Internal climate variability
  - Variability caused by natural changes in solar irradiance and volcanic aerosol loadings
- Our significance testing framework is highly conservative
  - “Total” noise estimates include volcanic eruptions and solar irradiance changes much larger than those observed over the satellite era
- Internal and “total” natural variability cannot produce sustained, global-scale warming of the troposphere and cooling of the stratosphere

And this significance testing strategy is highly conservative because we are looking at much larger changes in solar radiance, say around the time of the Maunder Minimum, much larger volcanic eruptions than we have actually observed.

So, I think the bottom line here is that internal and total natural variability in the CMIP5 suite of models just can't produce patterns of change like we have actually seen in the observations.

And I think that is what we are seeing, the direct radiative signature in the stratosphere of ozone depletion, to a lesser extent over the last 35 years of CO_2 increases and of the troposphere of greenhouse gas increases.

Stasis. Anthropogenic changes in greenhouse gases have this slowly evolving tropospheric warming signal which is superimposed on background volcanic cooling.

And it's this juxtaposition of the anthropogenic and volcanic signals that leads to decadal changes in warming rates after El Chichón and Pinatubo.

I note that Richard Muller in an op-ed in The New York Times a few months ago claimed that volcanos have no impact on decadal warming rates. I think this analysis clearly shows that he is wrong.

After removing ENSO signals, many aspect of the observed temperature response to El Chichón and Pinatubo were well captured by of CMIP5 multimodel average.

And again, for me at least, this is difficult to reconcile with a claim that we fundamentally screw up, on average, in estimates of transient climate response to external forcing.

However, there are still important questions. As I showed, there are still differences over the last 15 years during what you call the stasis period. Models don't show this. And the hiatus is still present in observations even after removal of ENSO effects.

Clearly, the missing volcanic forcing contributes to that discrepancy between modeled and observed behavior. The question is how much? Susan Solomon and colleagues have estimated about 25 percent. We get something similar to that.

There is a lot of uncertainty. And much of that uncertainty relates to the representation of volcanic aerosol effects in models. Eruptions are different. If I have learned anything over the last six months, it's that.

You can't take Pinatubo as a model for every other eruption in terms of the particle size distributions, the optical and chemical properties. We know from direct measurements these eruptions are different.

And it's very encouraging to me that, based on this work, modeling groups at NCAR and actually around the world are now trying to look more closely at the decisions that have to be made in translating observational estimates of aerosol optical depth into a volcanic radiative forcing.

I think that the bottom line is this. The stasis is not due to one factor alone. It's not internal variability alone. It's not external forcing alone. It's some combination of multiple factors.

And the real scientific challenge as I see it is to reliably quantify the contributions of different factors to the stasis and to the differences that we see between models and observations.

DR. KOONIN: Good, thank you. Phil?

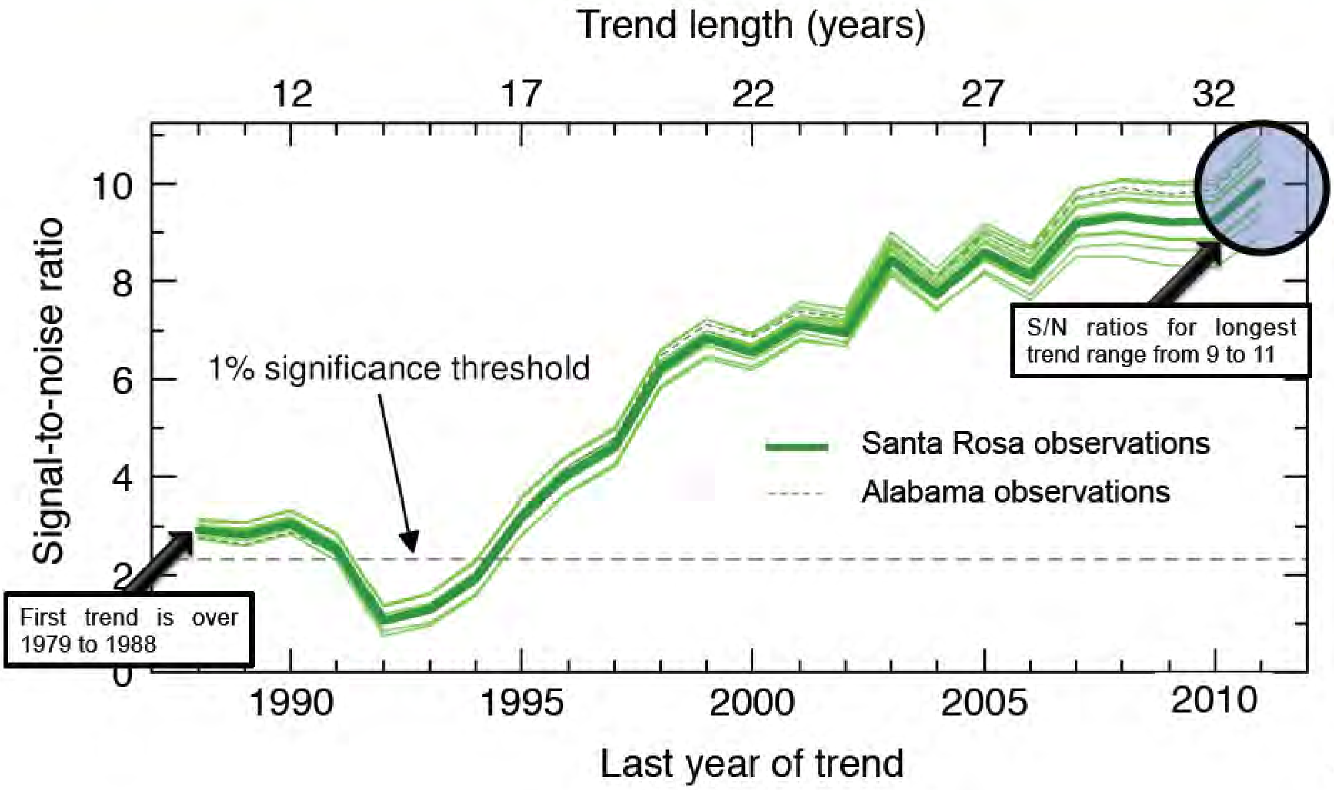
MR. COYLE: You showed us the publically-available data from Santa Rosa. Can you say what the differences would be if you were able to show us the proprietary date, what would cause the proprietary data to be different than what the public data is? Are you able to comment about that?

DR. SANTER: Yes, let's go back to one of these here. Yes, so this figure here [next page] shows in the signal-to-noise display both the publically available version of the data. That's the bold line, and the individual realizations.

Again, I looked at not the full 400-member ensemble, but I looked at the five to 95th percentile range. Now, those are publically available, too, I should point out.

Remote Sensing Systems has made

Estimating signal-to-noise ratios



First trend is over 1979 to 1988

S/N ratios for longest trend range from 9 to 11 them available. This is the publically-available version of the University of Alabama data set there in the dashed line.

So, to me, this is kind of cool because for these questions we have been discussing, model evaluation, detection and attribution, we can now compare distributions of observational results with distributions of model results.

That's new. For many, many years, we had one or two observational data sets that were regarded as sort of targets.

And I think what we have realized both more atmospheric temperature, ocean surface temperature, water vapor, is that in making these kind of assessments of model performance, it's very valuable to be able to fold observational uncertainty into the mix and to do model ranking as well.

You really want to know whether the results of some ranking are dependent on which observational realization you select.

DR. KOONIN: Scott?

DR. KEMP: Can you go to your slide 32. You went past it very quickly the first time, and I just thought there was something interesting there.

So, it seems that… Is it that the orange lines are the ten-year running averages?

DR. SANTER: Of the individual realizations.

DR. KEMP: And are they spreading as you go into the 2000 era?

DR. SANTER: I think they are, yes.

DR. KEMP: But they all have the same input data?

DR. SANTER: No, they don't.

Dr. KEMP: Okay, that was the

question.

DR. SANTER: No, they don't, unfortunately. Say for the volcanic aerosol forcing, not every group did exactly the same thing. Some groups used the so-called Ammann, et al., volcanic aerosol forcing. Some use Sato, et al., which I showed you a little later. Some used modified versions of Sato.

I think with one of the issues here is that observational estimates of changes in aerosol optical depth are themselves uncertain.

Just like in the MSU arena, different groups emerge these occultation instruments from different satellites in different ways. You have cirrus contamination effects, too, that you have to deal with.

Turns out that which wavelength you measure at is important in terms of the estimated aerosol optical death and which attitude range you look at. So, different groups do this in different ways and they get different results. And I think that is some of this spread that we see here.

Also, there are other things that are not identical. For stratospheric ozone, a number of these models actually have integrated stratospheric ozone chemistry models.

So, they compute historical changes in stratospheric ozone rather than actually specifying them. That's another reason for some of the differences that we see here. And I am sure that Bill and Isaac can expound on those issues.

DR. KEMP: Is this kind of saying that, as models get better and data gets more detailed, it’s actually pushing the models apart, but that with time, maybe all the models will eventually adopt all of the enhancements and maybe they will come back together?

DR. SANTER: Well, I think the issue is this is an ensemble of opportunity. And unfortunately, the forcings, so the estimates of changes in natural and anthropogenic constituents in the atmosphere, are not identical.

They are for CO_2 to first order. They are not for ozone. They are certainly not for volcanic aerosols.

They are probably pretty close for solar, although some models just look at TSI. Others actually spectrally resolve the changes in solar radiance over the solar cycle.

So, this is a sort of fundamental dilemma in what we do, that we are convolving intermodel differences in forcing with intermodel differences in response.

And that makes it a little bit

more difficult to figure out, well, are these particularly a manifestation of forcing error or response error?

Again, as I showed you for the volcanic aerosols, it is pretty clear that, in all models, there is this systematic error in forcing here.

Essentially, we flat-lined and that's not what the real world did. In the real world, background stratospheric aerosol increased by about four to seven percent per year from 2000 through to 2009.

DR. ROSNER: From small eruptions?

DR. SANTER: From this concatenation, this series of small eruptions that you see in blown-up form here.

DR. ROSNER: Could you go back to the slide that Scott just asked about, the previous one. Yes, that one, okay. So, presumably there is a subset of the models that really treated the aerosols exactly the same; is that true?

DR. SANTER: Sorry, which aerosols are we talking about here?

DR. ROSNER: The contributions from the volcanic eruptions.

DR. SANTER: There is a subset of models that used the same estimate of historical changes in aerosol optical depth. They then made probably very, very different --

DR. ROSNER: No, no, that's good enough. So, if you had plotted the results from just that set --

DR. SANTER: Sure.

DR. ROSNER: -- you would have then revealed what the differences are in the models? What is the answer to that question?

DR. SANTER: That's a great point. And actually what we have done is we have retrospectively calculated the radiative forcing in each of these models associated with volcanos.

And it turns out that for the two volcanic forcings I mentioned, Ammann developed at NCAR and Sato developed at GISS, there are differences, quite substantial differences in the peak radiative forcing around the time of El Chichón and Pinatubo.

So, if one does that stratifying, you should be able to pick up those kind of things. And we think that some of these differences here are associated with those fundamental differences in the forcing.

We are not showing them in this analysis here, but they are very relevant to this issue of trying to estimate transient climate response from the response to volcanos.

And in order to do that reliably, you really need to know about the differences in the volcanic aerosol forcing.

DR. KOONIN: Good. So, I want to open it up to our experts and then after that, I have a question that is related but it will take us off in a slightly different pursuit.

Sue?

DR. SEESTROM: I have a question about this same picture. So, it's clear that the spread in the models get bigger, but the red and the blue curves are two observations?

DR. SANTER: Yes.

DR. SEESTROM: Why is the difference so much greater in the period 1997 on than anywhere else in the historical record between the two sets of observations?

DR. SANTER: I don't know the answer to that question. These two groups, John Christy presumably will --

DR. CHRISTY: It's a difference

in the way we make a correction for the diurnal drift of the satellite.

DR. SANTER: Basically our knowledge of the diurnal cycle is incomplete. As a function, the diurnal cycle and temperature is a function of latitude, altitude, season.

And in order to adjust for the effects of satellite orbital drift on the sampling of the diurnal cycle, you have to have some model of what you think the diurnal cycle actually is.

And differences in how groups treat that diurnal cycle are responsible for some of these differences.

DR. SEESTROM: Did they correct for that change?

DR. CHRISTY: We did not use a model, by the way. We did not use a model. We used empirical evidence for calculation of the diurnal effect. Because it turns out some satellites don't drift, and so you can use those as references to show the difference.

DR. SANTER: Remote Sensing Systems doesn't use models exclusively. They also use GPS. They do other things, too.

DR. CHRISTY: Right, but the climate model is the basis for the --

DR. SANTER: No, it doesn't, actually. They use multiple. In their Monte Carlo-based technique --

DR. CHRISTY: No, I am talking about the original diurnal, the fundamental diurnal correction that is applied.

DR. SANTER: In their ensemble of observations, one of the reasons they get that spread is because they have different estimates of what the diurnal cycle is.

DR. CHRISTY: Right. What you described as the publically-available data set is the one that has a different diurnal correction. We use exclusively empirical. That's a nit.

DR. KOONIN: Right, right.

DR. SANTER: Anyway, the key thing here is that as in the surface temperature, as Judy showed, there are residual uncertainties in the observations. Different groups get different results.

To me, it seems important to incorporate that kind of information when comparing with model results.

DR. KOONIN: Fair enough.

Bill?

DR. COLLINS: So, just a couple things to note with regards to some of the discussion about the volcanics.

Also, the volcanic aerosols have effects both in the shortwave part of the spectrum and they also have effects in the infrared. So, they affect the thermal emission to space.

There is actually quite appreciable differences in whether or not the models even include those infrared effects and if so, how they do it. So, that's another difference here.

Even if they specify exactly the same optical depth, how they treat it, that is more likely to make sure that they are all the same in the shortwave. But there could still be diversity in the longwave.

And the other thing to keep in mind just to place this uncertainty in the volcanic aerosols in context is that we would be delighted if we understood tropospheric aerosol optical depth to a level of accuracy of .01.

This actually is a testament to their ability to remotely sense the stratosphere certainly down to that level.

One final point just to translate that into watts per meter squared, that is on the order of about two to three tenths of a watt per meter squared. So, it's about one tenth of the anthropogenic forcing. It's a small number, but nonetheless those small numbers do make a difference.

DR. KOONIN: John and then Scott.

DR. CHRISTY: I just noticed that these volcanic incidents in the past ten years or so on the order of .01 optical depth or less, that is much, much less, much, much less than Pinatubo and El Chichón.

And I think those little excursions on the MSU data there, I don't think you can identify those as volcanic. They are on the bottom, because, look, There are others.

DR. SANTER: I disagree completely. I mean, the message from this is you can. What matters here, since we are looking at 60-month sliding windows, is whether you could by chance simultaneously get five coolings after five of these early 21st-century eruptions. You can't.

We have looked at this very, very carefully. And the same with the shortwave up here, it's clear that, after these volcanic eruptions, you can see this, particularly if you look at this geographically. You can see the pancakes of these things in the reflected shortwave.

DR. CHRISTY: What caused the 2002 and 2004 cooling?

DR. SANTER: Excuse me?

DR. CHRISTY: What caused the 2002 to 2004 cooling?

DR. SANTER: Well, remember, we are looking at a given altitude range. Not all of these things make it up into the stratosphere. We are looking --

DR. CHRISTY: No, no, I am talking about the temperature.

DR. SANTER: Excuse me. Let me finish, please.

DR. CHRISTY: Bottom line.

DR. SANTER: Yes, I know. Not all of these volcanic eruptions like this one here, Anatahan, actually make it up into the stratosphere.

Some of them have a significant component to the troposphere. There are forest fires that have signatures in stratospheric aerosol optical depth.

The largest ones, again, these measurements look at different wave lengths, different altitude ranges. It is not surprising to me that there is some evidence of residual noise here that is uncorrelated with the stratospheric aerosol.

But what we actually do, again, is look at the probability of getting cooling after this guy, after this guy, after Sarychev, after all of the major eruptions in the 21st century.

And when you do that, then residual noise is a very poor explanation for the simultaneous cooling that we see after multiple events.

DR. KOONIN: So, it's the correlation of all three of these measurements, which are independent?

DR. SANTER: Yes, they are all independent measurements. And again, they clearly show some multivariate signal of early 21st-century volcanic activity.

DR. KOONIN: John, you want to respond?

DR. CHRISTY: Look how rapid that temperature bounces back. That would not really be a volcanic signature if it did that. Talking about months.

DR. SANTER: These are small eruptions. These are not sustained for years, John. Look at the stratospheric aerosol optical depth.

DR. CHRISTY: They are little blips.

DR. SANTER: This isn't Pinatubo. This isn't El Chichón lasting for years. No wonder it's responding quickly.

DR. KOONIN: I have got a question that sort of leverages off of the detection and attribution, a little bit of the models and then onto projection.

This is probably the right time to raise it, since you are the detection and attribution guy. I have got about two or three slides I would like to just show to set it up, and maybe we will do that discussion for five minutes.

DR. SANTER: Sure.

DR. KOONIN: Let me put up those charts.

DR. KEMP: While you are doing

that, can I ask a quick question?

DR. KOONIN: Sure.

DR. KEMP: Ben, it seems that a general approach of the attribution studies is to look for spatial fingerprints and say you cannot recreate these spatial patterns through natural forcing. What is the statistical probability that this is just a natural effect?

Is there any interest in actually taking that a step further and saying let's compute the coefficients on the natural and anthropogenic forcings with the spatial patterns?

Are you not hearing me or not following what I am saying?

DR. SANTER: I didn't hear the last part of what you said.

DR. KEMP: Is there any interest in moving beyond just asking the statistical question about what is the chance that this pattern is anthropogenic or not anthropogenic and moving to actually trying to estimate the coefficients?

DR. SANTER: Yes. So, what I do is quite difference from what people like Myles Allen and others do.

They cast all of this in some regression framework and they actually estimate what they call the betas, the scaling factors on some model-based spatiotemporal signal, say, associated with greenhouse gases only or sulfate aerosols only.

And then they try and estimate that beta in the observations and see whether the model, on average, gets the right strength of that particular response to forcing in the observations, or whether that has to be scaled down or scaled up.

So, they cast all of this in terms of explicitly estimating the strength of individual model signals in observations.

I don't do that. One of the issues associated with trying to do that individual signal search and quantification is degeneracy. If you have things that look like each other, then that's a bit of a problem.

And to me, when you put things together into one combined spatiotemporal factor, you lose this kind of pattern information.

It is not easy to decompose it again after the fact into altitudinal patterns or geographical patterns and figure out why you do or do not get correspondence between models and observations.

That, I think, is one of the advantages of what we do. But the advantage of what they do is that explicit quantification of individual factors and observations.

DR. KOONIN: Good. So, let me

ask some questions about what is in the IPCC report. I understand it may be different than what you have just talked about.

DR. SANTER: Can I sit down?

DR. KOONIN: Yes, might as well. This is going to be a discussion.

So, look,[next page] I find in IPCC chapter 10 -- I understand, Bill, this is not your chapter, but you should know something about this. I would assume everybody -- IPCC is a consensus, so presumably everybody agrees.

And it's about the scaling factor discussion. And I have highlighted in red the relevant piece here.

(Reading): "Responses to individual forcings can be scaled up or down in order to be consistent with observations."

And then I look at one of the

Scaling factors in attribution

IPCC AR5 WG1 10.2.1:

So-called 'fingerprint' detection and attribution studies characterise their results in terms of a best estimate and uncertainty range for 'scaling factors' by which the model-simulated responses to individual forcings can be scaled up or scaled down while still remaining consistent with the observations, accounting for similarities between the patterns of response to different forcings and uncertainty due to internal climate variability.

figures, 10.4, [next page] and I see for a range of models shown there, this is for GMST, the surface temperature, that there are a set of scaling factors for the greenhouse gases, which are the green bars, the anthropogenic, other anthropogenic, and natural.

And there are three numbers for each the models. And the surprising things are, (A), they are not one. I understood, Ben, that the corresponding things in the studies you showed, they were all one. You didn't scale things.

DR. SANTER: I don't do any scaling at all.

DR. KOONIN: I understand. But here, evidently IPCC needs to scale in order to match the observations; second, that many of the scaling factors are not consistent with one; they are smaller than one. The tightest ones are smaller than one. And there is a fair bit of

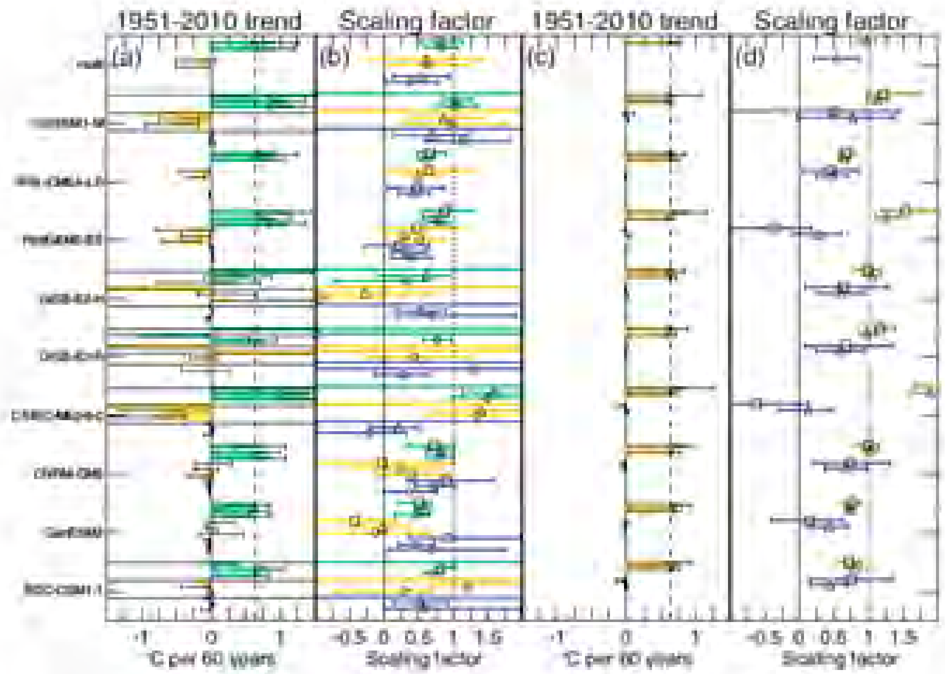


Figure 10.4: (a) Estimated contributions of greenhouse gas (green), other anthropogenic (yellow), natural (blue) forcing components to observed GMST changes over the 1951–2010 period. (b) Corresponding scaling factors by which simulated responses to greenhouse gas (green), other anthropogenic (yellow) and natural forcings (blue) must be multiplied to obtain the best-fit to HadCRUT4 (Morice et al., 2012) observations based on multiple regressions using response patterns from nine climate models individually and multi-model averages (multi). Results are shown based on an analysis over the 1901–2010 period (squares, Ribes and Terray, 2013), an analysis over the 1861–2010 period (triangles, Gillett et al., 2013) and an analysis over the 1951–2010 period (diamonds, Jones et al., 2013). c) and d) As for a) and b) but based on multiple regressions estimating the contributions of total anthropogenic forcings (brown) and natural forcings (blue) based on an analysis over 1901–2010 period (squares, Ribes and Terray, 2013) and an analysis over the 1861–2010 period (triangles, Gillett et al., 2013). Coloured bars show best estimates of the attributable trends (a and c) and 5–95% confidence ranges of scaling factors (b and d). Vertical dashed lines in a) and c) show the best estimate HadCRUT4 observed trend over the period concerned. Vertical dotted lines in b) and d) denote a scaling factor of unity.

variability in them.

Go ahead, Ben.

DR. SANTER: Excuse me. I think for the greenhouse gas component, many of them are consistent with one.

DR. KOONIN: Well, there are some, for example (indicating slide). There are a couple over there. There is one up there, the greens that are not consistent with one.

And the model mean is up there and it's about 75 percent. Yes, it's consistent with one, but the mean is less. In some cases, there are negative scaling factors.

And then I go into chapter 11. [next page] And I asked, “did you account for that scaling?” In other words, did you calibrate the model? And then did you use that when you went to the decadal projections?

I find, in general, no. But they have this method, right,

Scaling in decadal projections?

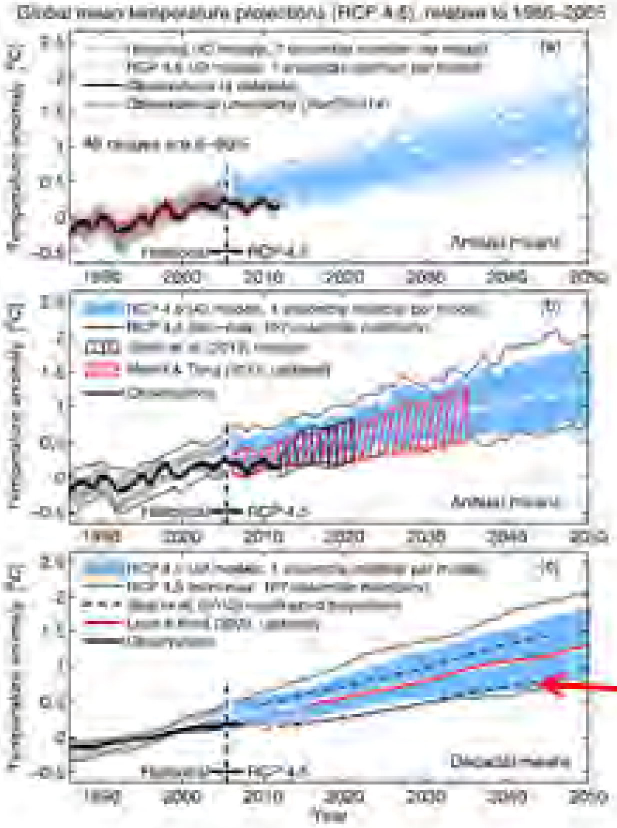


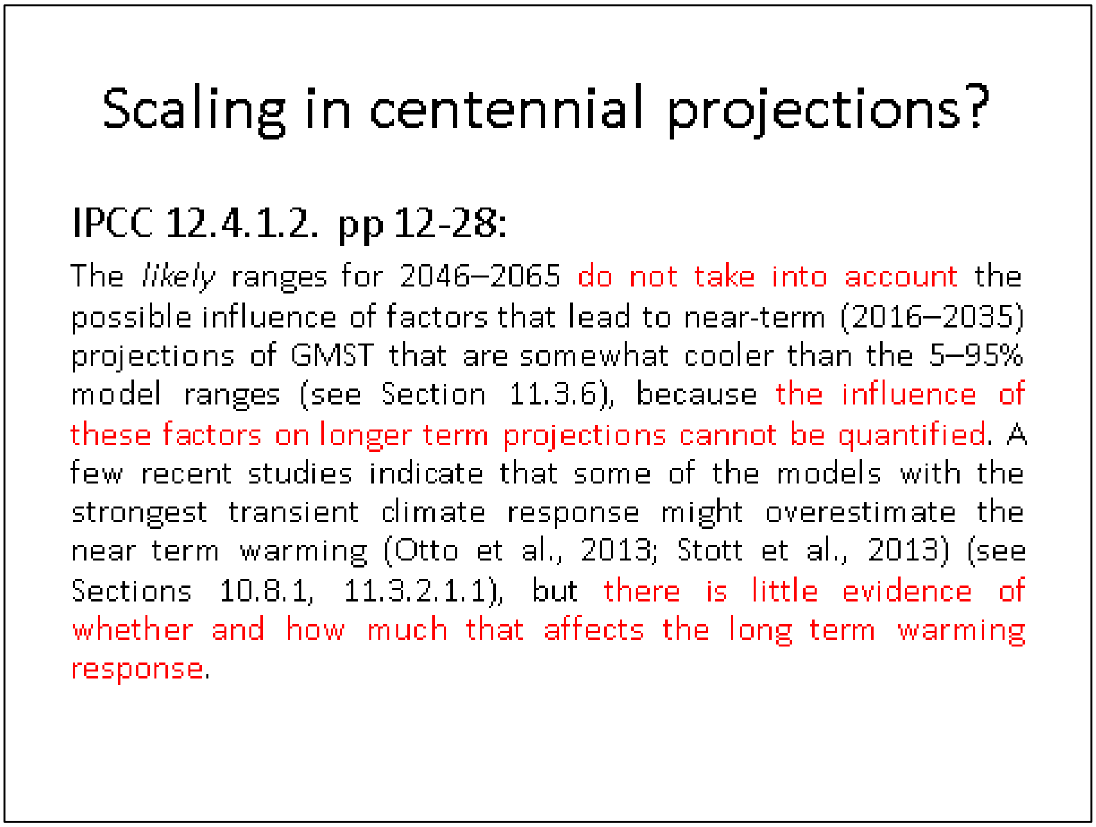
Figure 11.9: a) Projections of global mean, annual mean surface air temperature 1986–2050 (anomalies relative to 1986–2005) under RCP4.5 from CMIP5 models (blue lines, one ensemble member per model), with four observational estimates (HadCRUT3: Brohan et al. (2006); ERA-Interim: Simmons et al. (2010); GISTEMP: Hansen et al., 2010; NOAA: Smith et al. (2008)) for the period 1986–2011 (black lines); b) as in a) but showing the 5–95% range (grey and blue shades, with the multi-model median in white) of annual mean CMIP5 projections using one ensemble member per model from RCP4.5 scenario, and annual mean observational estimates (solid black line). The maximum and minimum values from CMIP5 are shown by the grey lines. Red hatching shows 5–95% range for predictions initialized in 2006 for 14 CMIP5 models applying the Meehl and Teng (2012) methodology. Black hatching shows the 5–95% range for predictions initialized in 2011 for 8 models from Smith et al. (2013b). c) as a) but showing the 5–95% range (grey and blue shades, with the multi-model median in white) of decadal mean CMIP5 projections using one ensemble member per model from RCP4.5 scenario, and decadal mean observational estimates (solid black line). The maximum and minimum values from CMIP5 are shown by the grey lines. The dashed black lines show an estimate of the projected 5–95% range for decadal mean global mean surface air temperature for the period 2016–2040 derived using the ASK methodology applied to 6 CMIP5 GCMs (from Stott et al. (2013)). The red line shows a statistical prediction based on the method of Lean and Rind (2009), updated for RCP4.5.

where they can do it for some subset of the models. And as Bill remarked, most of the warming is already committed over the next decades, so it doesn't matter much. And that's the dashed lines up there.

But then I go to the centennial projections in chapter 12 and it says that,[next page] "The likely ranges do not take into account these factors because the influence of these factors on the long-term projections cannot be quantified."

So, to me, it looks like they set a calibration against the historical data and then they wiped out that calibration in doing the centennial projections resulting in probably a 25, 30 percent overprediction of the 2100 warmings.

So, is that right? Am I reading IPCC right or have they done what I would have thought is the scientifically correct thing to do?



Scaling in centennial projections?

IPCC 12.4.1.2. pp 12-28:

The *likely* ranges for 2046–2065 do not take into account the possible influence of factors that lead to near-term (2016–2035) projections of GMST that are somewhat cooler than the 5–95% model ranges (see Section 11.3.6), because the influence of these factors on longer term projections cannot be quantified. A few recent studies indicate that some of the models with the strongest transient climate response might overestimate the near term warming (Otto et al., 2013; Stott et al., 2013) (see Sections 10.8.1, 11.3.2.1.1), but there is little evidence of whether and how much that affects the long term warming response.

DR. COLLINS: So, the place where you are getting the overestimate is that you are concluding that it has been overestimated because, during what looks like, I supposed could be inferred as a calibration exercise in chapter 10?

DR. KOONIN: Correct.

DR. COLLINS: The models had to be scaled down in their greenhouse gas component, which is the dominant thing by 2100, leading you to conclude that one should apply similar scaling for the projections into 2100.

DR. KOONIN: Probably 25, 30 percent, maybe more. I don't know. So, have I understood what they did right or not? Have I correctly understood what they did and if so, is that the right thing to have done?

If you wanted to do it right,

you would need to do the ASK method for the centennial simulations. And they haven't done that because they have only done it for a few of the decadal simulations.

Judy?

DR. CURRY: It was a relatively last-minute thing to do that. If you look at the second order draft, they hadn't done any of the --

DR. KOONIN: To do which?

DR. CURRY: The second order draft of the Working Group 1 report, you didn't see any sign of that downscaling. So, it was something that was done relatively last-minute by, I guess, the Chapter 11 authors.

DR. KOONIN: This downscaling? (Indicating slide.)

DR. CURRY: I'm not sure about that one, but I am talking about -- go back -- this one, that red-hashed box. That was a new addition.

DR. KOONIN: It was kind of

halfway step or even a third of the way.

DR. CURRY: Yes, so it was done by the chapter 11 authors. And I think it was generally a sensible thing to do. But chapter 12, it did not trickle into chapter 12, that kind of thinking.

DR. KOONIN: At least I conclude now from what I understand that the centennial scale projections of temperatures are probably high?

DR. CURRY: I think so.

DR. KOONIN: By 30 percent, at least for RCP8.5 which is dominated by greenhouse gases?

DR. COLLINS: Well, I would be unwilling to do sort of an error assessment of this in public without having looked at it a lot more closely.

As Ben pointed out, one of the issues with what chapter 10 did is that you have signals that have similar spatial patterns. They are using, so, in some cases the uncertainty in aerosols, for example, is quite large. It is the dominant source of uncertainty during this time period.

The extent to which that has been properly accounted for in the error propagation, frankly, I don't know how chapter 10 did this exercise. I haven't looked at it closely enough to be able to answer your question. I think that is an interesting question.

But some of the issue about whether or not the relationship of that scaling factor to one hinges on their ability to deconvolve aerosol forcing from greenhouse gas forcing, as a for-instance.

DR. KOONIN: We know they are coupled, right?

DR. COLLINS: They were coupled, yes, because when you burn fossil fuel, you are emitting sulfur dioxide and you are emitting carbon dioxide.

DR. KOONIN: I had it more in the sense that they were coupled when you tried to reproduce the historical data.

DR. COLLINS: Yes. Well, they are coupled in two ways, yes.

DR. KOONIN: The models overpredict the Pinatubo response, but they overpredict the CO_2 response as well.

DR. SANTER: We don't find overprediction, not significant overprediction of Pinatubo, as I was trying to show here.

DR. KOONIN: So, you are one of the models that are in the wings, not in the bulk? You are a high outlier?

DR. SANTER: I would say two things, Steve. One thing is that, as I tried to show and as discussed in the IPCC hiatus box, it's clear that there are some systematic errors in the forcing over the last 15 years.

We underestimated the cooling associated with post-Pinatubo volcanic aerosols. We underestimated the cooling associated with the unusually broad solar minimum in the last solar cycle.

We probably underestimated systematically some of the cooling associated with stratospheric ozone.

So, if you are estimating some beta that from the observations, and the models simulations do not incorporate those negative influences that the real world experienced --

DR. KOONIN: You are not going to get it?

DR. SANTER: Yes, you are not going to get the right beta. So, to me, that's the scientific challenge, to deconvolve the errors in beta that arise from incorrect simulation of internal variability or an unusual manifestation of internal variability that we didn't capture from bona fide errors in model response and errors in model forcing. All of the above are at play.

DR. COLLINS: I would have made a similar conclusion to the statement you just made for a zeroth-order physics error that was made in the very first assessments where they did not include aerosol forcing.

So, what happened in the early days of the assessments, the aerosols were not such a big player.

They were essentially looking at a system where you had the solar boundary condition, changes in the well-mixed greenhouse gases and no aerosols.

And they found, sure enough, that they were overestimating the warming without looking at the historical record.

That's sort of a zeroth-order forcing error because they had left out something we know we can see by eye in the earth's atmosphere, and would have led to a similar conclusion which is sure, you have the models. If you leave out aerosols, they are overestimating in the historical record.

I think the important thing to recognize is that the historical record is different from what we think will be happening at the end of the 21st century.

So, up until now, we have been dealing with a signal where there are a strong influences of both positive and negative from greenhouse gases in the positive and aerosols in the negative, as I showed you, about 40 percent of the signal currently.

By the year 2100, we believe that people will be wisely improving air quality, but that's removing the shielding effect of aerosols. And so we think that, by the year 2100, the forcing -- and this is just a projection -- will be dominated by well-mixed greenhouse gases.

DR. KOONIN: But if the model tells you that you got the response to the forcing wrong by 30 percent, you should use that same 30-percent factor when you project out a century.

DR. COLLINS: Yes. And one of the reasons we are not doing that is that we are not using the models as statistical projection tool.

DR. KOONIN: What are you using them as?

DR. COLLINS: Well, we took exactly the same models that got the forcing wrong and which got sort of the projections wrong up to 2100.

DR. KOONIN: So, why do we even show centennial-scale projections?

DR. COLLINS: Well, I mean, it

is part of the assessment process. And the uncertainty, I think there is a point not to get confused about what the driving uncertainties there are. By the year 2100, it's not --

DR. KOONIN: If you calibrated the model against historical data, discovered you needed .7 to be applied to the greenhouse gas, you should keep that same .7 when you run it forward, no?

DR. COLLINS: No.

DR. KOONIN: You keep all the other parameters. You don't change any of the other parameters.

DR. COLLINS: No, that calibration factor is due to an error in the boundary condition.

DR. KOONIN: Which boundary?

DR. COLLINS: In the aerosol boundary condition. Beta is accounting for an error in a boundary condition.

DR. KOONIN: You can't untangle

the aerosol in greenhouse gases well enough? Is that what you are telling me?

DR. COLLINS: I think that's a large source of uncertainty.

DR. LINDZEN: I think he is saying there is a specific assumption that the aerosol will disappear.

DR. KOONIN: Well, that's in RCP and whatever the RCP is, that's what it is. That's a boundary condition.

But the greenhouse part of RCP, which is dominant in 8.5, you should take the greenhouse sensitivity that you determined from the historical data, shouldn't you?

DR. SANTER: Can I respond to that. So, the kind of thing that you mentioned has been done by Peter Scott, Myles Allen and colleagues where they calculate some beta for their model results over some calibration period and then apply that beta to the projections.

DR. KOONIN: That's the ASK method?

DR. SANTER: Right. And what they show is they shrink the uncertainty range in the projections.

DR. KOONIN: And bring it down a bit?

DR. SANTER: Yes. And I think what Bill is saying, and what I agree with, is that it is clear the reason we need to scale down is not only associated with some fundamental model error insensitivity.

That's possible, but we know beyond a shadow of the doubt that we got some of the forcing wrong systematically.

So, some of the that downscaling is associated with incorrect representation of cooling influences that the real world experienced but that the CMIP5 multimodel archive did not.

Now, to me, when I look at that

figure, I showed you, the first two-thirds agreement, last two-thirds disagreement, if modelers were really so skilled and so focused on tuning to get a desired result, we would have done a lot better job than that.

There is no way, there is no way we would have gotten that fundamental disconnect.

DR. KOONIN: You said you are not representative of what IPCC does. Certainly some modelers are well-focused on tuning and they discovered they need .6, .5, .7 in the greenhouse gas response in order to tune properly.

And what bothers me is that they throw away that tuning when they project out a century. That's what I am worried about.

DR. SANTER: Again, to me the real problem as a scientist here is in partitioning forcing error from the response error. It's not easy to do that with this ensemble. I think what you need are experiments where you systemically explore some of the these forcing uncertainties. And we have not done a good job of that.

We have done a good job exploring parameter uncertainty that I would argue that we have not done a comparably good job exploring forcing uncertainty --

DR. CURRY: Thank you. That's very important.

DR. SANTER: -- which is large and as I indicated, affects critically the correspondence between models and observations.

So, if you care about the parameter uncertainty, you ought to care equally about the forcing uncertainty. And that forcing uncertainty affects the betas that you are concerned with.

DR. KOONIN: Absolutely, yes. Thank you. Dick?

DR. LINDZEN: I am the one who stands between you and lunch. Thank you, Steve.

DR. KOONIN: Before you launch in, Isaac, did you want to say something?

DR. HELD: No.

DR. KOONIN: You will get to it?

DR. HELD: I still have my chance later.

DR. BEASLEY: The last word.

DR. KOONIN: All right, Dick?

DR. LINDZEN: At any rate, thank you for having this. I think it is a good idea to discuss this instead of assert.

I find a little bit of strangeness in the incompatibility between major uncertainties in understanding sensitivity and so on and the kind of bookkeeping approach that I include two percent here and one percent there.

I am more or less in the first category. And the question I am addressing -- it's already been addressed -- is what gives rise to the large uncertainties in sensitivity?

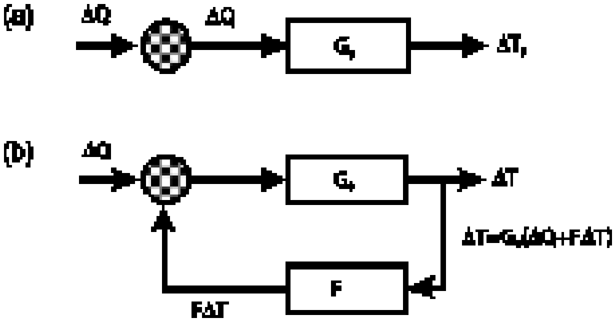
And secondarily, how is the IPCC expression of increasing confidence in the detection attribution consistent with the persistent uncertainty? Wouldn't detection of anthropogenic signal necessarily improve estimates of the response?

At any rate, let's start with the first question. And it has been pointed out somewhat obscurely, it's intrinsic to feedback systems.

So, you have this diagram. [next page] So, you know, you have a forcing and the node here and the zero-feedback gain and so you get the zero-feedback response. If you have a feedback then you have this circuit here, APS questions concerning equilibrium climate sensitivity.

- What gives rise to the large uncertainties in this fundamental parameter of the climate system?

- How is the IPCC's expression of increasing confidence in the detection/ attribution/ projection of anthropogenic influences consistent with this persistent uncertainty? Wouldn't detection of an anthropogenic signal necessarily improve estimates of the response to anthropogenic perturbations?



$$\Delta T_0 = G_0 \Delta Q$$

$$\Delta T = G_0(\Delta Q + F\Delta T)$$

$$\Delta T = \frac{\Delta T_0}{1-f}$$

$$f = F\,G_0$$

The uncertainty comes from the fact that each model essentially starts with f=0.5 due to water vapor, so that any small addition leads to huge increases in sensitivity.

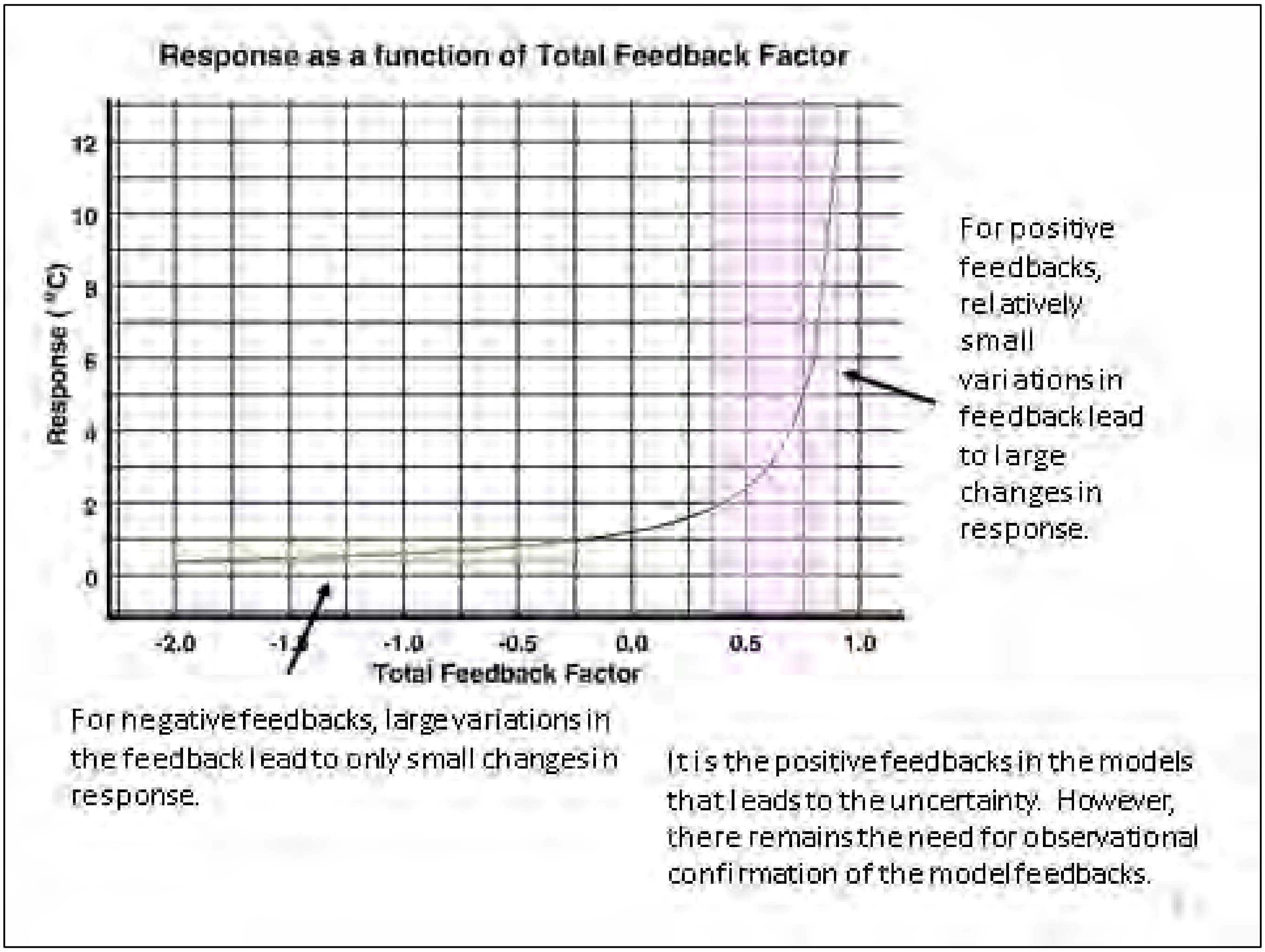
adding to the forcing.

And so, you solve and you have delta T naught over one minus f. At any rate, the uncertainty comes from something that was mentioned, the Manabian water vapor feedback.

Early on in the '70s, there was the discovery that if you assumed relative humidity stayed constant, you could double the response to CO_2 with water vapor simply because as temperature increases relative humidity is fixed, so specific humidity must increase.

Once you start out with 0.5 for f, of course anything you add to it including 0.5, which will bring you to infinity, gives you the range. What you have is a curve like this. [next page]

For positive feedbacks, relatively small variation in the feedback lead to large changes in the response. But it's equally true that if you didn't have the strong



For positive feedbacks, relatively small variations in feedback lead to large changes in response.

For negative feedbacks, large variations in the feedback lead to only small changes in response.

It is the positive feedbacks in the models that leads to the uncertainty. However, there remains the need for observational confirmation of the model feedbacks.

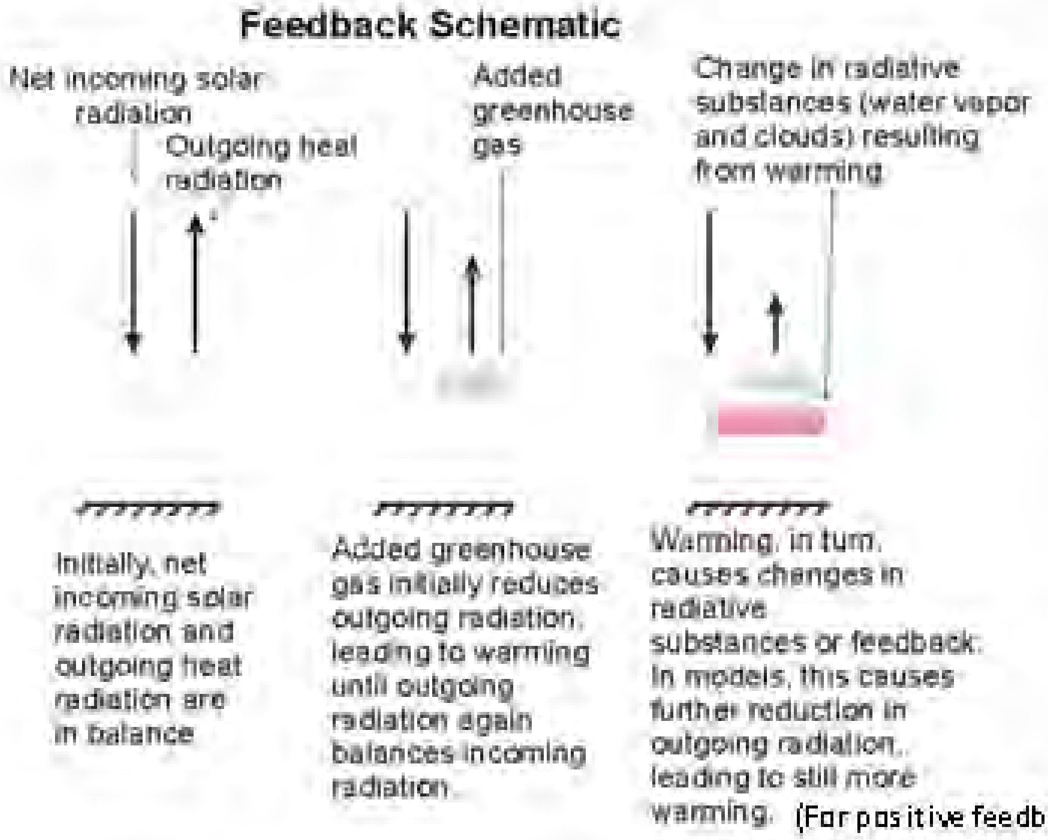
positive feedback to begin with, you would be in a very much more constrained region.

The point is, it is the existence in the models of a basic positive feedback that leads to the uncertainty. And this would lead to the suggestion that you would like an observational basis for the feedbacks.

And [next page] a number of people, including myself and Choi, Spencer and Braswell, Trenberth and Fasullo, Gregory and others, have tried to find this in looking at the outgoing radiation from ERBE and CERES and so on, various satellites in recent years.

And the idea is simple enough. I mean, these pictures are not that helpful, but they are describing the feedback. You start out in equilibrium. You add some greenhouse gas.

An obvious approach to measuring feedbacks would be to see how outgoing radiation responds to surface temperature fluctuations, but it, too, has difficulties. It is, nevertheless, the approach tried by Lindzen and Choi, Spencer and Braswell, and Trenberth and Fasullo, Gregory and Foster, and others.



Initially, net incoming solar radiation and outgoing heat radiation are in balance

Added greenhouse gas initially reduces outgoing radiation, leading to warming until outgoing radiation again balances incoming radiation.

Warming, in turn, causes changes in radiative substances or feedback. In models, this causes further reduction in outgoing radiation, leading to still more warming. (For positive feedbacks).

Of course, there are also short wave feedbacks due to the albedo of clouds and of variable snow cover. The reflectivity of ice is more difficult since wet ice is like water. Also, snow at high latitudes is mostly in winter where, at least in the arctic, there is no sunlight.

You now reduce the emission temperature because you have raised the level. And then if you have a feedback that is positive, you do this more. This is the longwave feedback. There are shortwave feedbacks due to albedo and variability and so on.

And these are actually it turns out much tougher to deal with. But they are the ones that are giving you a lot of this uncertainty. We’ll come back to it.

If you want to measure it, [next page] basically what you are saying is, a feedback doesn't care where the temperature change came from. So, you look at fluctuations in temperature.

If you get more response in terms of outgoing radiation than you would get from zero feedback, which may be Planck black body, then you have negative feedback. If you get

The crucial point about the feedbacks is that they respond to surface temperature fluctuations regardless of the origin of the fluctuations.

The basis of the approach is to see if the satellite measured outgoing radiation associated with short term fluctuations in Sea Surface Temperature (SST) is larger or smaller than what one gets for zero feedback. Remember that a positive feedback will lead to less outgoing radiation, while a negative feedback will lead to more.

It turns out that the model intercomparison program has the models used by the IPCC, forced by actual SST, calculate outgoing radiation. So one can use the same approach with models, while being sure that the models are subject to the same surface temperature fluctuations that applied to the observations.

less, you have a positive feedback.

So, you have something to look at. And you have model comparisons to look at. All the models have not only CMIP but AMIP so you even have models that are forced by exactly the same temperatures you are looking at.

Okay, so you do that. In principle, [next page] it sounds straightforward. In practice, it's not. First of all, there are obvious considerations of time scale.

So, for instance, if you have a perturbation in temperature and you wait forever and the system equilibrates, you now have a change in temperature without a change in flux. That's a bias. So, you have to make it shorter than that.

You need to consider the process. Most of the feedbacks we are looking at involve very short-term changes in water vapor, cloudiness and so on. They are

In principle, this should be a straightforward task. However, in practice, it is rather difficult. The first two difficulties involve basic physical considerations.

First, not all time scales are appropriate for such studies. Greenhouse warming continues until equilibrium is reestablished. At equilibrium, there is no longer any radiative imbalance. If one considers time intervals that are long compared to equilibration times, then one will observe changes in temperature without changes in radiative forcing. The inclusion of such long time scales thus biases results inappropriately toward high sensitivity. Equilibration times depend on climate sensitivity. For sensitivity on the order of 0.5C for a doubling of CO_2, it is on the order of years, and for higher sensitivities it is on the order of decades. In order to avoid biasing sensitivity estimates, one should restrict oneself to time intervals less than a year.

There is also the need to consider time intervals long enough for the relevant feedback processes to operate. For water vapor and cloud feedbacks, these time scales are typically on the order of days. For practical time resolution, this is generally not a problem.

Time scales on the order of 1-3 months are, thus, certainly appropriate for sensitivity studies. Longer time scales also involve 'pollution' from seasonal effects, etc. Restricting consideration to such short time scales is the approach taken in Lindzen and Choi (2009, 2011).

associated with things on the order of days. And so, certainly you want your period to be longer than that.

As a practical matter, this is not a problem. Time scales on the order of one month, three months are fine. The problem with the equilibration, by the way, is it depends on sensitivity itself.

So, for instance, if you have a sensitivity of five degrees for doubling of CO_2, time scale is many decades. But if your sensitivity was only a half degree, the time scale would be on the order of a year. So, you have a range that you don't want.

Okay, you have other problems, seasonal effects so on. And each of the papers I mentioned deals with this.

The problem [next page] that is hardest to deal with, though, and that has to deal with the shortwave mostly, is

The **second** problem is more difficult. Outgoing radiation varies (especially in the visible) for reasons other than changing surface temperature (volcanoes, non-feedback cloud fluctuations). Such changes are not responses to surface temperature fluctuations but they do cause surface temperature fluctuations.

Apart from basic physical issues, there are other practical problems such as the presence of significant gaps in the outgoing radiation data. Also, the radiation data involves two satellite systems (ERBE and CERES) with different properties.

Lindzen and Choi, 2011, describes our attempt to deal with these issues. **Here, I will simply describe the signature of the second problem: namely, when one has an unambiguous feedback, a plot of r^2 and/or $\Delta F/\Delta T$ v. Lag has a single maximum at a small lag. If, however, the non-feedback variations are large, then these relations have an S-shape, and the regression at zero lag can be completely misleading because it consists primarily in artifacts from the fact that there is a finite decorrelation time for the non-feedback variations.**

outgoing radiation is not simply a matter of the temperature perturbation.

There are changes in outgoing radiation going on all the time because of large changes in clouds that have nothing to do with feedbacks.

And this normal variability, short-term variability, not the long-term variability we are talking about, in turn induces changes in temperature in the surface.

And so, there is a distinct need to consider lags and so on to make sure you are looking at responses. And even then, there are decorrelation times which screw things up.

At any rate, you can go through the list of problems with incompatibility between CERES and ERBE. And we deal with it and you get a result.

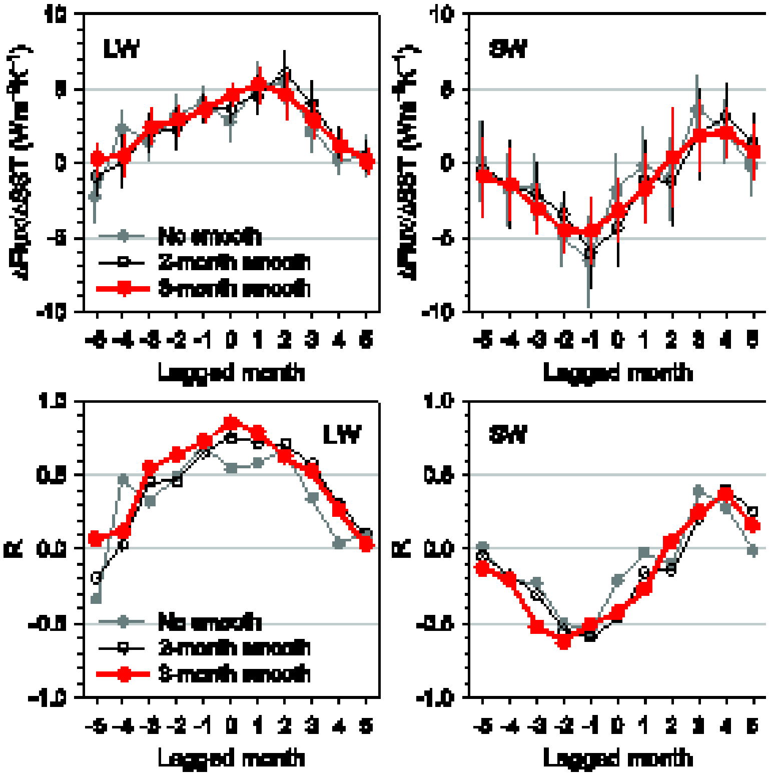
These are results. [next page] The top diagram is delta flux over sea surface temperature as a function of lag. And the bottom is correlation as a function of lag. And the left is longwave and the right is shortwave.

What you will notice is, for the longwave, you have a single peak, reasonably well-defined. Now, this is largely tropical. You get poorer results globally, but I think there are good reasons to focus on the tropics and both I think are fairly unambiguous.

And they unambiguously show a negative feedback with an uncertainty. So, it's like replacing f=.5 from Manabe’s water vapor feedback with minus 0.3 plus or minus 0.2.

On the other hand, when you get to the shortwave, you have this kind of S-pattern. And it's clear that at zero lag, you are still getting what looks like a positive feedback, but

Here are our results based primarily on SST and tropical radiation (Lindzen and Choi, 2011). In evaluating feedbacks, we require that radiative imbalances in the tropics be shared with the globe.



Note the relatively unambiguous longwave ***negative*** feedback as well as the S-shape for the short wave feedback.

that could have been produced by a nonfeedback change in temperature with the decorrelation time.

Recently, Choi and Hee-Je Cho -- I am a distant author on this paper -- did a couple of thousand Monte Carlo runs with noise and so on to see what happens.

And what we found was,[next page] if you had sufficiently low noise, you got the curves like you got for the longwave. As you increased the noise, a product of the noise was the S-shape.

And so, that leaves me with a fairly pessimistic view of our ability at this point to detect the shortwave feedback. But the important point is the longwave feedback was essential to the huge uncertainty.

And so, for example,[next page] if you start out with a longwave feedback of .5, and you have a shortwave feedback

Choi et al (2013) show that S-curve is mostly an indicator that noise is dominant. By noise, we simply mean that cloud variations arise mostly from processes other than feedbacks.

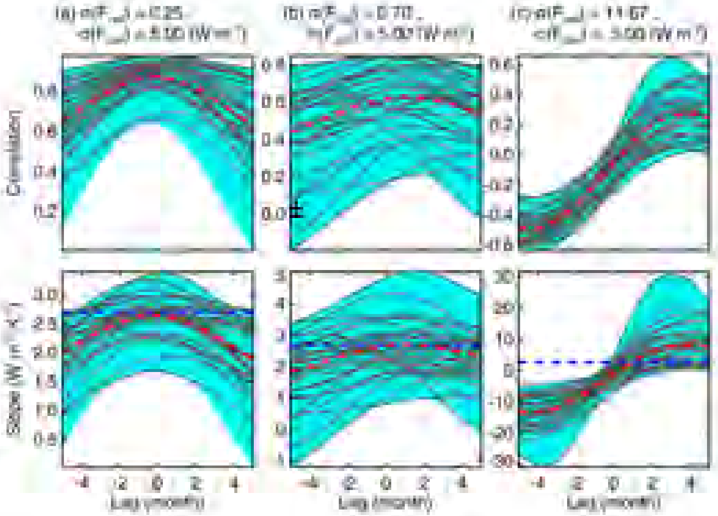


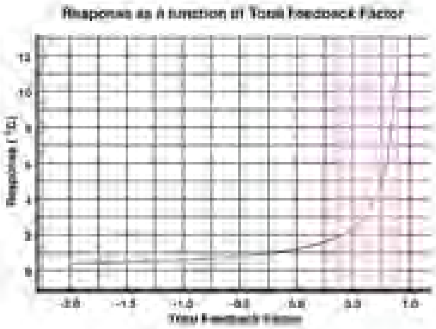
Fig. 2 The lagged linear correlation coefficient and regression slope of ΔR versus ΔTs from the zero-dimensional energy balance model simulations. The blue shaded area indicates the range of the simulated values from 1,000 simulated realizations, and the red dashed line indicates their average. The assumed climate feedback 2.7 W m−2 K−1 is superimposed by the blue dashed line in the lagged slope graph. The thin solid lines are 30 randomly selected examples. The forcings are purely radiative (a) and non-radiative (b), respectively

However, longwave feedback (including the crucial water vapor feedback) is essentially negative or zero. Previous result gave $f_{lw}=-0.3 \pm 0.2$. Absence of longwave feedback has been confirmed independently by Spencer and Braswell, and even by Trenberth and Fasullo.

Let us take the zero gain sensitivity as ~ 1C

f_{lw}	f_{sw}	Sensitivity
0.5	0-0.3	2-5
-0.5 - 0	0.3	0.83 – 1.4
-0.5 - 0	0	0.67 - 1

Remember:



Note: One really can't decompose lw feedback for water vapor and upper level cirrus because where there are upper level cirrus, the water vapor feedback doesn't work, and the area covered by upper level cirrus changes as part of the cloud feedback.

that is between zero and plus 0.3, you, of course, get a range of equilibrium sensitivity of two to five degrees.

But if instead you had the longwave being uncertain between minus 0.5 and zero, let's say, and you had a shortwave feedback that was 0.3, that would give you 0.83 to 1.4.

If you had no shortwave feedback, you would be at 0.67 to 1. But you would be in this constrained range of the feedback behavior.

I should mention that longwave feedback here is what we deal with, not water vapor feedback. And the reason is, you cannot disentangle the two.

So, for instance, the feedback depends on changing the emission level for infrared. Where you have upper-level clouds, thin cirrus, the water vapor doesn't matter. Clouds determine the emission level.

So, you cannot get the longwave feedback independently for the clouds and the water vapor. The area is varying all the time. These are very large changes. I will come back to that.

If you look at normal variance of clouds, you know, for instance, to equal three and a half watts per meter squared, let's say, it would be like ten percent in upper-level cirrus, a fraction of a percent you have in your document for the lower level

It would be a 500 meter change in altitude for upper level cirrus. If you look at the normal variations, they are much larger than that.

So, these things are happening all the time. So, that is the reason for the uncertainties in sensitivity and where we may be way off.

Now, attribution, that has been discussed here.[next page] And the problem

The problem with connecting sensitivity with attribution stems from the fact that models are tuned. The use of a simple energy balance model with a mixed-layer and diffusive thermocline ocean illustrates the issue. Incidentally, such simple models, with sensitivity assumed, replicate the response to explicit radiative forcing of GCMs with the same sensitivity quite well, and are, in fact, used for IPCC scenario construction.

of identifying sensitivity with attribution is terrifically distorted by the ability to adjust aerosols.

And it becomes really difficult. We will see some aspects of it. Now, one can address this much more simply than with complex models. And that's always been funny to watch.

Yes, if you want to know what the feedback factors are, if you want to know about ENSO and so on, you are not going to get it from a simple model. But you are not getting it terribly well from the big models, either.

If what you want to know is the response to the specified globally-averaged forcing, it's long recognized that simple energy balance models, if tuned to the same sensitivity as the larger models, can for a simple ocean model do a fairly good job of replicating the forced mean response. And in fact, if you go to chapter 13 of Working Group 1, they are still doing that for the scenario-building.

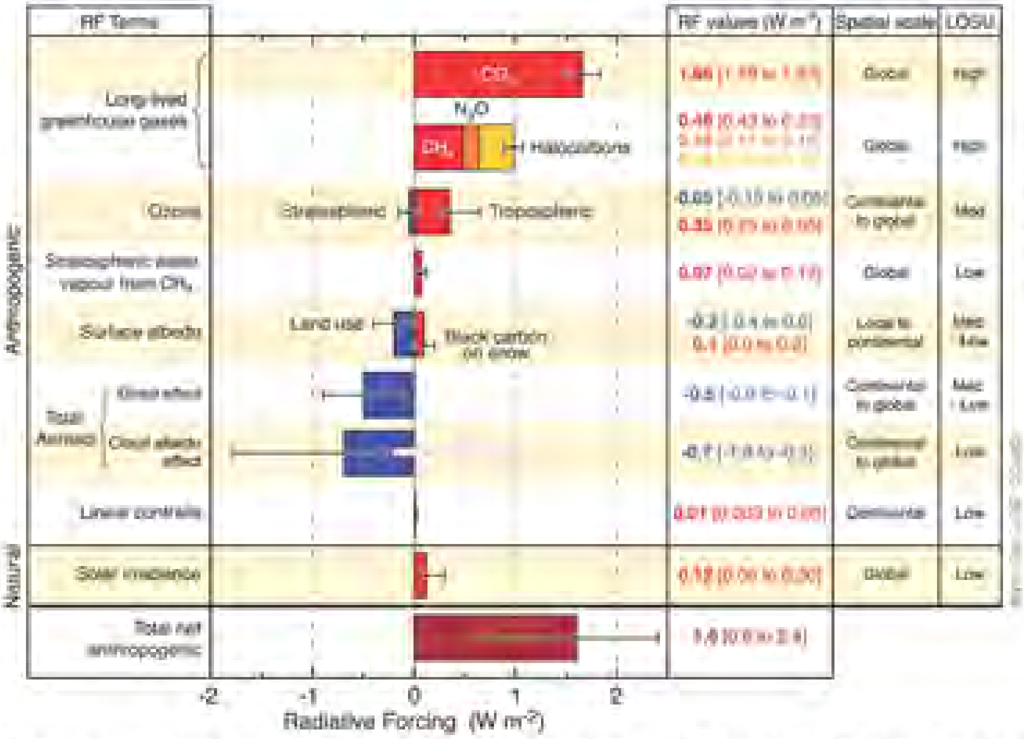
So, that's all I am going to do here is go with that. This [next page] is probably from AR4, but it doesn't much matter for present purposes. You have the uncertain aerosols in blue. You have the greenhouse --

DR. KOONIN: Just to ask, I mean, Judy started today or it was the second talk of the day by saying the aerosol uncertainty has been reduced significantly.

DR. LINDZEN: Yes, oh, yes. That's going to be important in this. But part of it is increasing it, actually. And this is the point I mentioned.

If you go to the indirect effect, you will notice I have, if you can see it here; I don't know. I'm sorry sort of blind. But you

Here are the IPCC's estimates of various 'climate forcings.'



Note that total GHG forcing is about 80% greater than that due to CO_2 alone.

Note that most observed aerosols are natural.

Note also that the uncertainty on the negative end for aerosols is at the limit of possibility. Also, as the IPCC text notes, the cloud effect can actually be positive for clouds at levels where temperatures are below 0C (which is the case for a large part of the cloud cover).

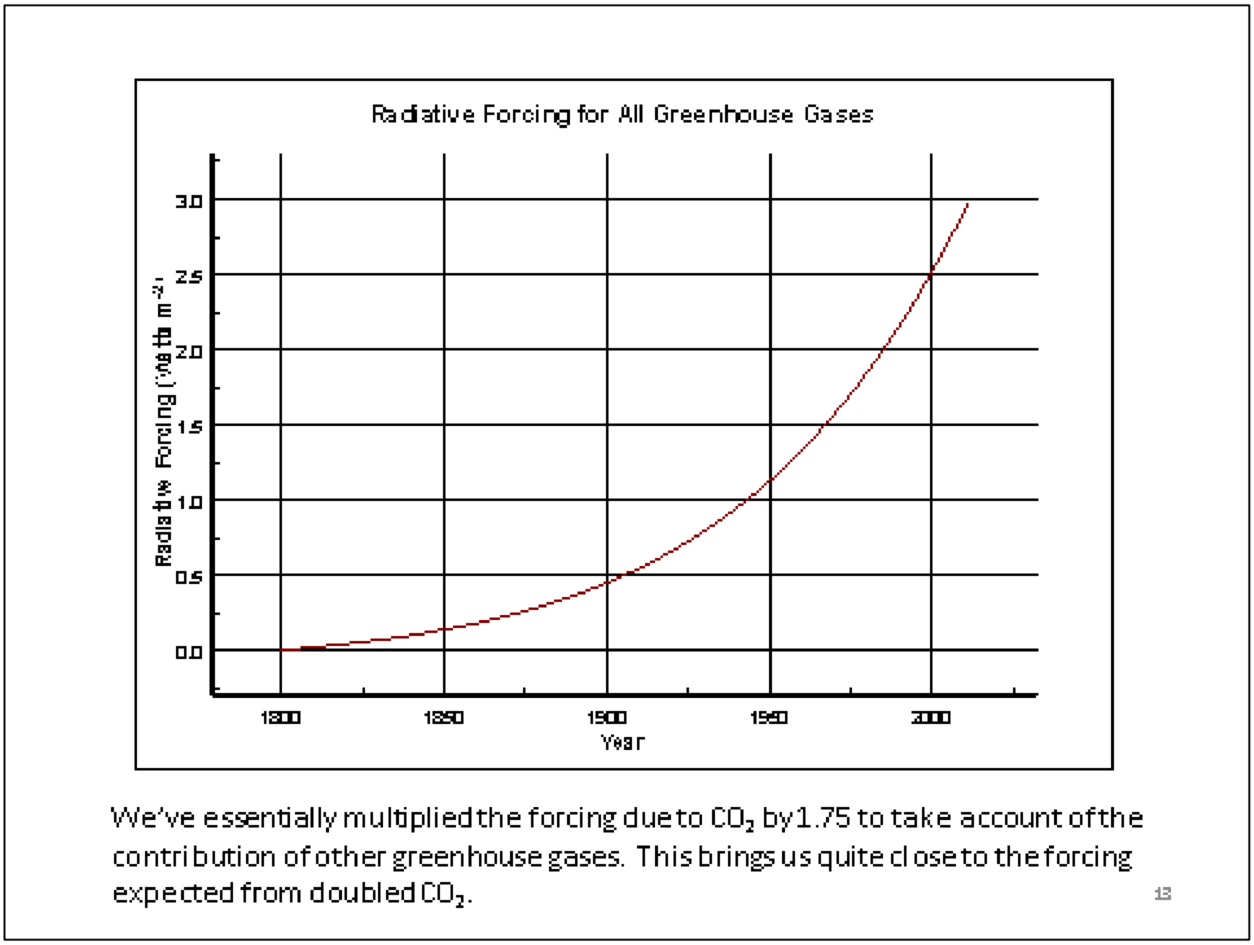
notice this thing here (indicating).

You have that floating around and then you have the soot in the direct effect. So, it could be going every which way.

In any event, the greenhouse part of it is interesting because that is already about much greater than CO_2 alone and pretty close to what you would expect for a doubling of CO_2. So, it is not in some remote future we are looking at a doubling.

Also, with aerosols, you have Calipso and other satellites looking at them. It's not necessarily relevant, but most of the aerosols you see are natural. So, this is sort of interesting.

Now you have a simple picture of the radiative forcing. [next page] It has been increasing over time. You could have every detail in it, but this is roughly what you are doing. Multiply the CO_2 by 1.75.



We've essentially multiplied the forcing due to CO_2 by 1.75 to take account of the contribution of other greenhouse gases. This brings us quite close to the forcing expected from doubled CO_2.

Sato's picture of the aerosol forcing,[next page] the volcanic forcing is used by a lot of the models. I put it in here. It doesn't have the latest ones, but that isn't relevant to what I am talking about.

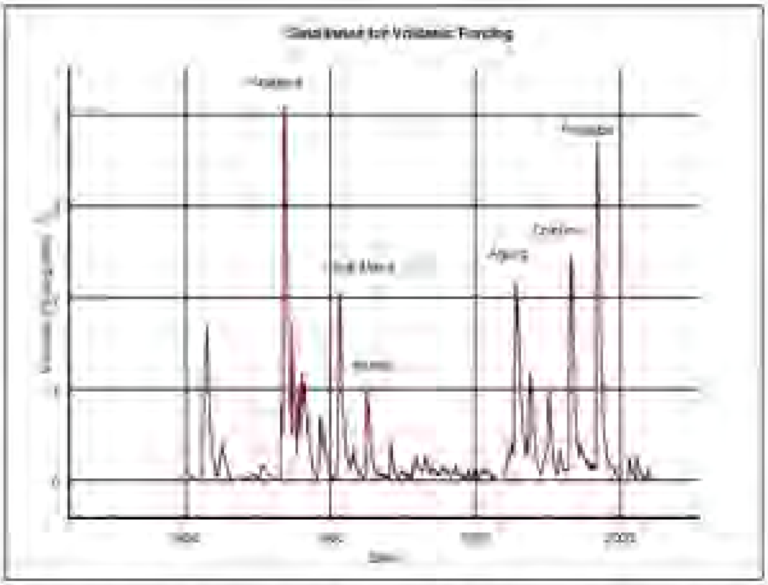
It's often been pointed out they cluster. This is, however, probably a property of random processes. They cluster. It's one of these oddities.

At any rate, the response to the volcanos depends on the sensitivity of the model. So, here [next page] you have different models with different sensitivities ranging from 0.75 to five degrees.

For 0.75, you also have very short response times. So, you only see the blips in the red. As you go down to the higher sensitivities, you begin seeing a secular effect.

Now, if you look at this literatures, you know from the UK Met

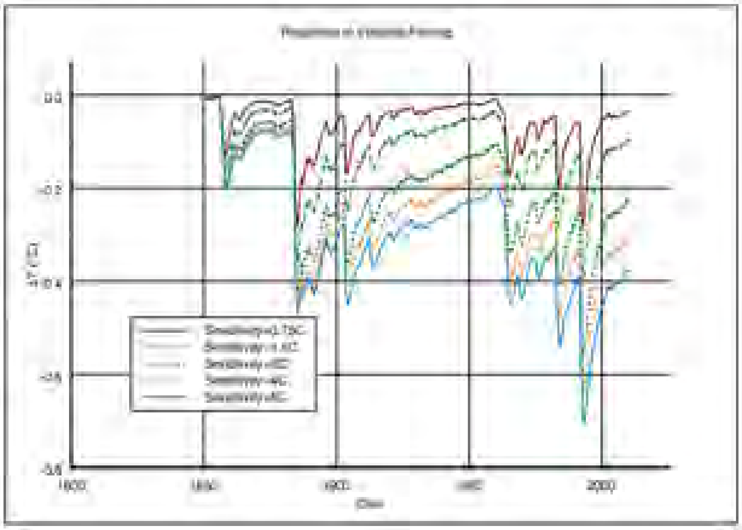
Radiative forcing due to volcanoes.



It turns out that most models use the estimate for volcanic forcing developed by Sato at the Goddard Institute for Space Studies by Sato. This is probably as good as any estimate, though there is substantial uncertainty. Note that there are two clusters of volcanic activity separated by a period of relative quiet. Clustering is, in fact characteristic of random processes.

Here is the response to the volcanic forcing for models with different sensitivity. Climate sensitivity is defined as the equilibrated response to a doubling of CO_2.

Note the persistence of volcanic cooling in sensitive climates. Such persistence is not evident in the observations, but is found in some models.



Note that knocking sensitivity down to 0.75C gains about 0.3C relative to models with sensitivity of about 3C.

Office, Gregory and others have been complaining that their model shows an influence of Krakatoa to the present. Now, the question is, is this real or not?

DR. HELD: In sea level.

DR. LINDZEN: In sea level, but they also are seeing it in other things.

DR. SANTER: Not on surface temperature.

DR. HELD: Not on surface temperature.

DR. LINDZEN: May not.

DR. SANTER: No, definitely not.

DR. LINDZEN: Well, this would not be a big thing on that issue, either. It's 0.3 degrees. It's saying you would get something on the order of a third of a degree cooling that you might not have in a high-sensitivity model.

In any event, the persistence

is a thing that would be itself a reasonable test. And I have it at the end, the slide, but I wasn't planning it on showing it, on response time where you can look at the processes, assume they are AR1, look for the response time, compare data and models.

You know, it's not perfect. None of these things are perfect. You don't quite know what the oceans are doing in each of the models.

You have a simple model. You have you have a certain time scale for that. But nevertheless, there is a fair systematic appearance of a longer time scale in the models.

In any event, this is simply saying if the response time is short compared to the intervals, the average interval between volcanos, you will see blips. If it's the opposite, you will see secular trend.

DR. KOONIN: And the response

time is correlated with the sensitivity?

DR. LINDZEN: Yes.

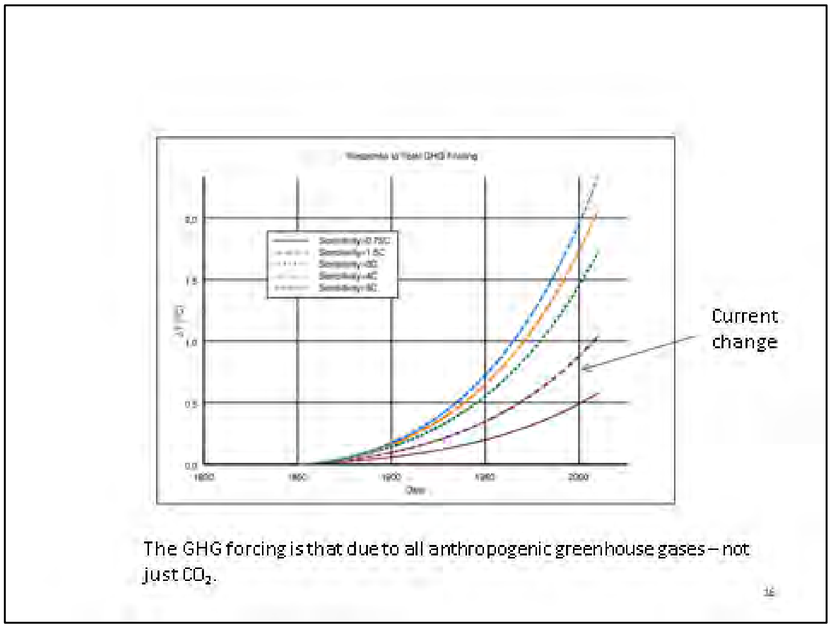
DR. KOONIN: So, you are getting an indirect measurement?

DR. LINDZEN: Essentially, sensitivity is the ratio of a flux to a delta T at the surface. And so, that is the coupling.

Okay, now here [next page] is just the response to the greenhouse forcing for such a simple model. And the current change is where this arrow is.

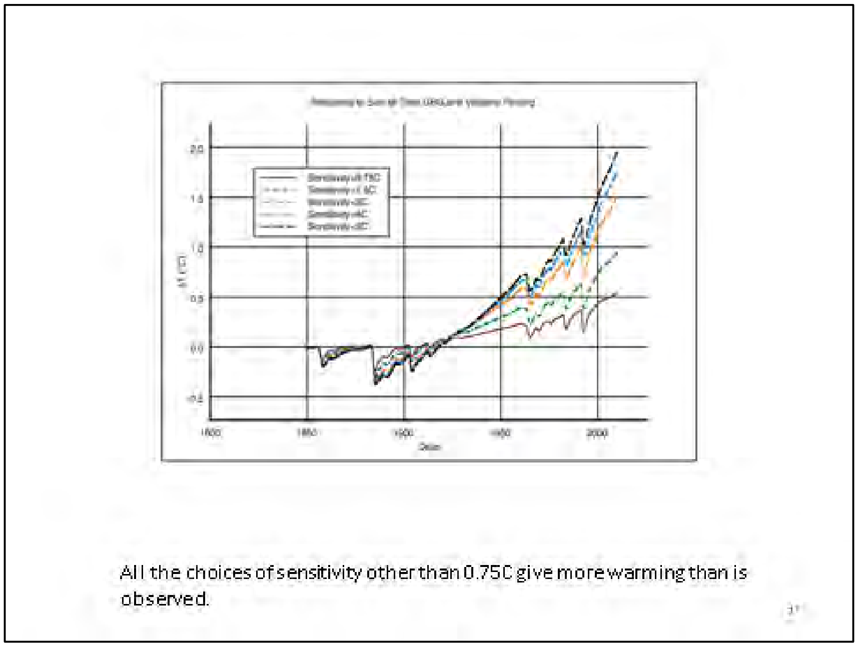
It's obviously looking closer to the lower. When you add in the volcanos [next page] , it reduces the difference but you still have a significant overestimate.

On the other hand, until AR4, most models ended up describing what you saw, and that was the aerosols. So, you had something like this. [next page] They look fairly similar.

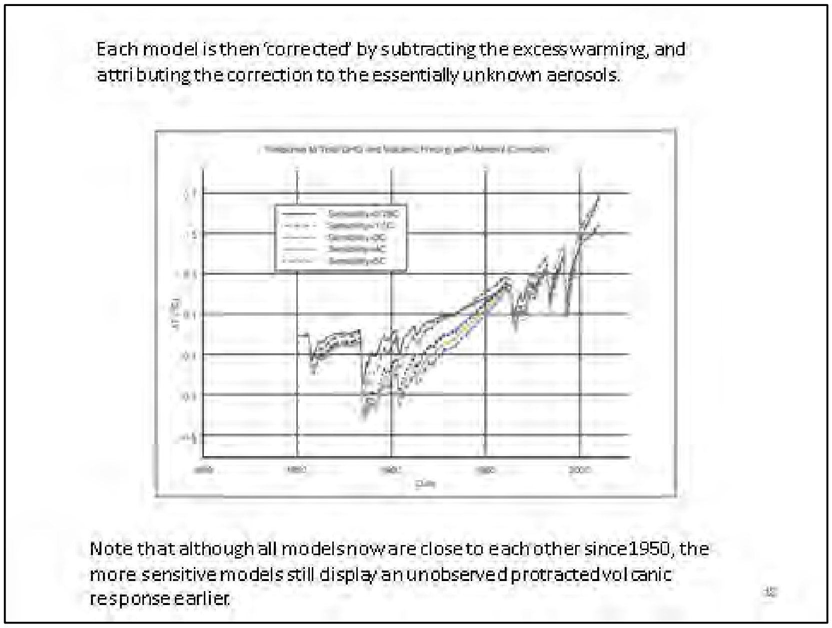


Current change

The GHG forcing is that due to all anthropogenic greenhouse gases – not just CO_2.



All the choices of sensitivity other than 0.75C give more warming than is observed.


Each model is then 'corrected' by subtracting the excess warming, and attributing the correction to the essentially unknown aerosols.

Note that although all models now are close to each other since 1950, the more sensitive models still display an unobserved protracted volcanic response earlier.

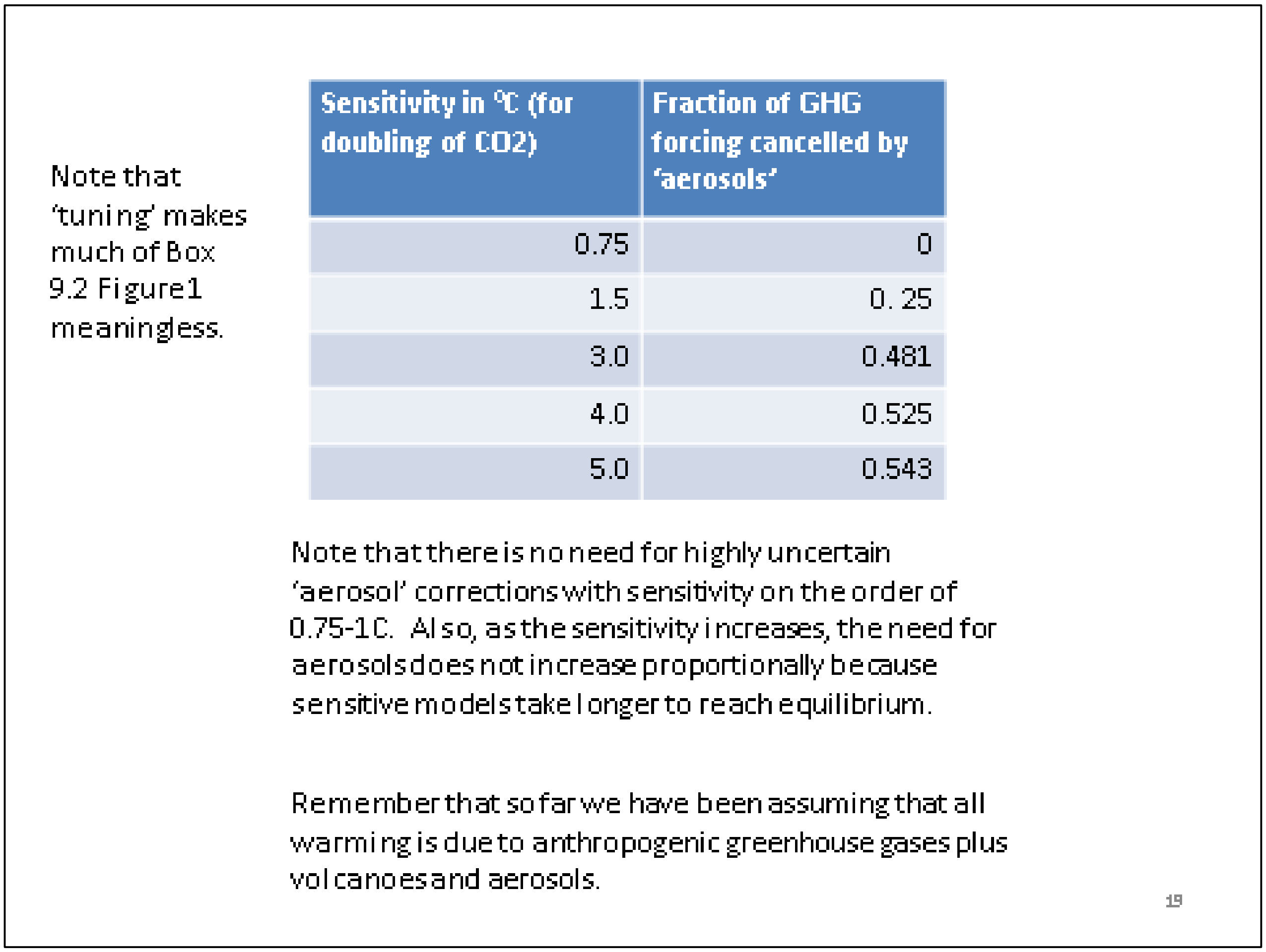
And the only difference is that you might have had a slight difference in the response to volcanos in the in-between period, but people rarely focus on that.

And the question is how much did you have to subtract? [next page] And so, obviously, if you had 0.75, you didn't subtract anything.

By the time you had one and a half degrees for doubling, you had to take 25 percent out. And then for the rest, it really didn't matter much.

You had to take about half out. And that's because you are in that part of the sensitivity curving. It changes a lot for a little.

In any event, that's where you are at. And you are so far assuming everything is due to the specified forcing. But there have been a number of papers in recent years -- this stuff is from Tung and Zhao from



Note that ‘tuning’ makes much of Box 9.2 Figure1 meaningless.

Sensitivity in °C (for doubling of CO2)	Fraction of GHG forcing cancelled by ‘aerosols’
0.75	0
1.5	0. 25
3.0	0.481
4.0	0.525
5.0	0.543

Note that there is no need for highly uncertain ‘aerosol’ corrections with sensitivity on the order of 0.75-1C. Also, as the sensitivity increases, the need for aerosols does not increase proportionally because sensitive models take longer to reach equilibrium.

Remember that so far we have been assuming that all warming is due to anthropogenic greenhouse gases plus volcanoes and aerosols.

PNAS.

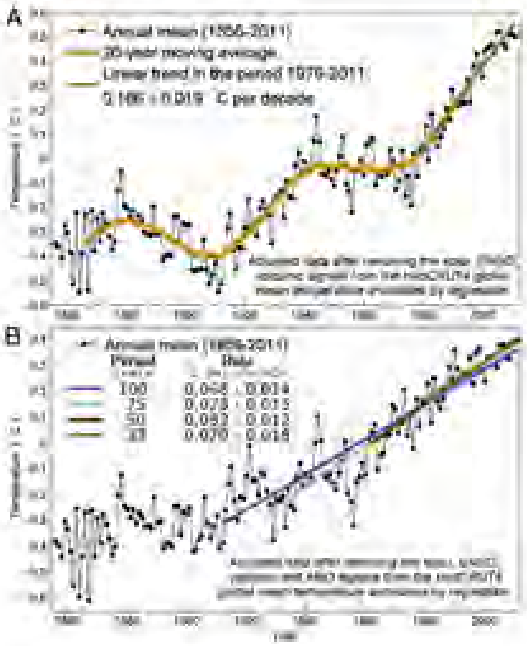
There has also been a set of papers by Tsonis and Swanson and others who are trying to estimate how much comes from internal variability. The general conclusion is it's on the order of half, [next page] although you could account for more, depending on what model you wanted to use.

And, of course, each of these things puts more and more constraints on the attribution and the related sensitivity.

Now, I would suggest that most independent attempts to find sensitivity end up with less sensitivity than the models are displaying. But paleo is an interesting exception. [next page]

There, the fact that Milankovitch parameters, orbital parameters are giving you no change in mean insolation, essentially. And you are getting a big climate

However, recent work shows that on the order of half of the warming occurring over the past 50 years or so is actually due to natural internal variability. Thus, the need for aerosol correction becomes excessive -- especially given the new more restricted estimates – when sensitivities are high.



There have, of course, been attempts to attach probabilities to the various adjustments so that extended deviations from observations do imply reduced sensitivity, but the IPCC seems reluctant to pursue this.

Tung and Zhou, PNAS, 2013. Similar results have been been found by Swanson & Tsonis.

There are, in fact quite a few approaches to estimating sensitivity. Most point to lower sensitivity, but some estimates of high sensitivity come from paleoclimatic data. For example, in the cycles of glaciation over the past 800 kyrs, there was almost no change in globally and annually averaged insolation suggesting potentially high sensitivity.

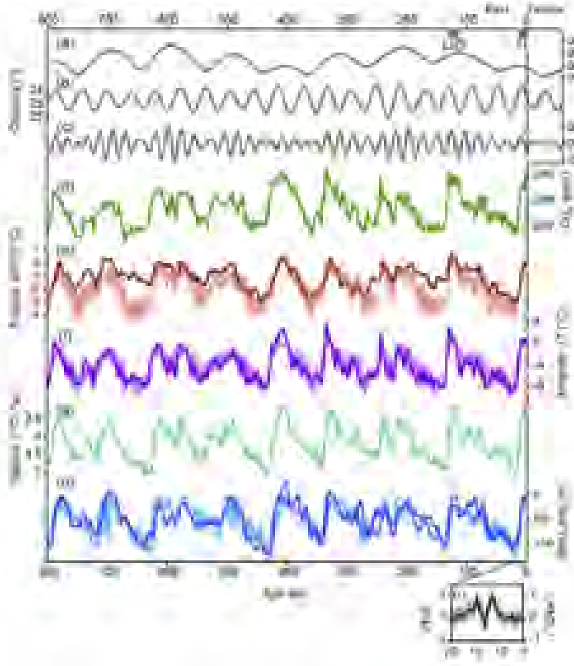


Figure 5.3: Orbital parameters and proxy records over the past 800 kyr. (a) Eccentricity, (b) obliquity, (c) precessional parameter, (d) atmospheric concentration of CO2 from Antarctic ice cores, (e) tropical SST stack, (f) Antarctic temperature stack based on up to seven different ice cores, (g) stack of benthic δ18O, a proxy for global ice volume and deep-ocean temperature, (h) reconstructed sea level. Lines represent orbital forcing and proxy records, shaded areas represent the range of simulations with climate, climate-ice sheet models of intermediate complexity and an ice sheet model forced by variations of the orbital parameters and the atmospheric concentrations of the major greenhouse gases. (i) Rate of changes of global mean temperature during Termination I. [See WG1 report for full caption with references.]

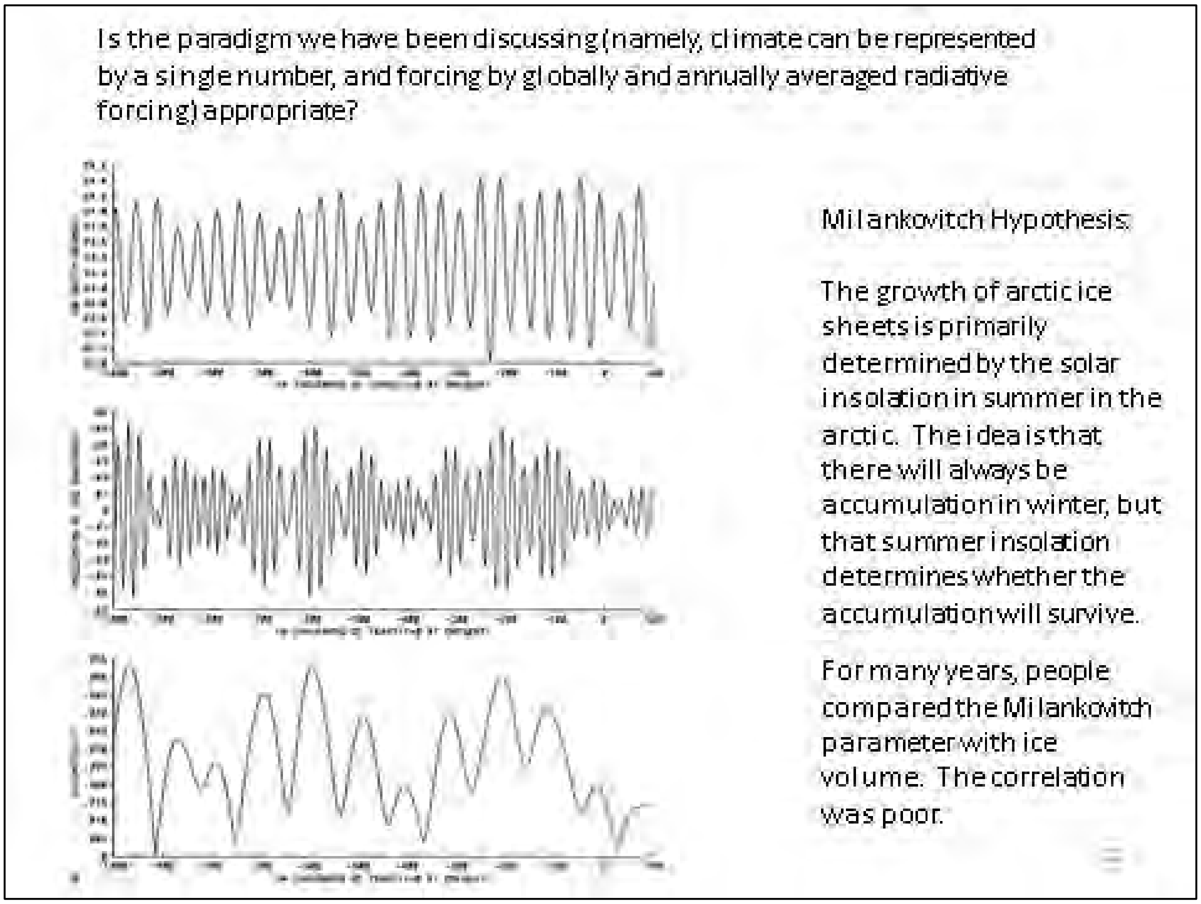
change, which suggests sensitivity.

And here, it's interesting Isaac is here because he was a post-doc with me and he was the person who got me interested in the Milankovitch thing. I had not thought about it much.

And it seemed to me very interesting that you had almost no forcing and you were getting a big response. And I worked on this for a few years and suddenly realized I am thinking wrong.

This [next page] is not a problem with globally averaged, annually averaged forcing. Milankovitch was probably right. What Milankovitch did was simply say you have these orbital variations, the obliquity, the eccentricity, precession of the equinoxes.

But what was important for glaciers was the insolation in the Arctic in summer. Almost every



Is the paradigm we have been discussing (namely, climate can be represented by a single number, and forcing by globally and annually averaged radiative forcing) appropriate?

Milankovitch Hypothesis:

The growth of arctic ice sheets is primarily determined by the solar insolation in summer in the arctic. The idea is that there will always be accumulation in winter, but that summer insolation determines whether the accumulation will survive.

For many years, people compared the Milankovitch parameter with ice volume. The correlation was poor.

glaciologist will say that. Essentially, you will always get snow in the winter.

It will always accumulate. What determines whether you build up an ice sheet over a long period of time how much survives the summer.

Now, at first, people looked at that. And this is a funny field and all of us make errors that are pretty gross in retrospect.

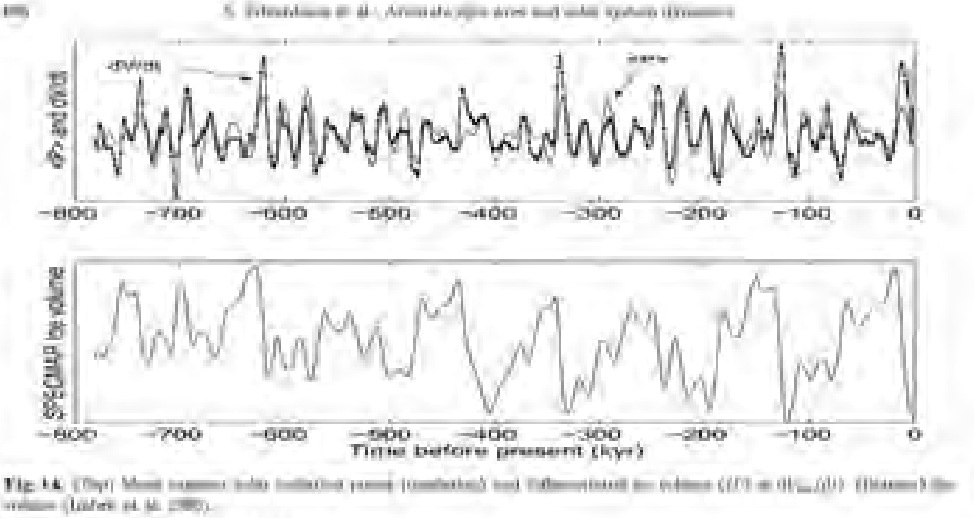
But what happened with the CLIMAP program is they compared Milankovitch parameter with ice volume and they didn't get an awfully good correlation.

Eventually, I feel embarrassed because I realized at some point I was looking at tropical influence. Three Swedish astronomers, Edvardsson and some other names studied this. [next page]

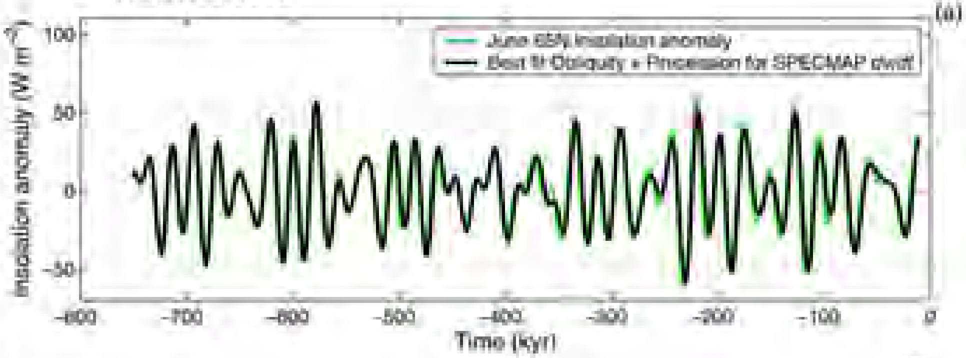
They did the obvious thing, which was to look at the time derivative of the ice volume versus

However, in 2003, several Swedish physicists (Edvardsson et al) noted the obvious fact that we should compare time rates of change of ice volume (rather than the ice volume, itself) with the Milankovitch parameter. This was also noted by Roe in 2006. The results were striking.

Note that the Milankovitch parameter varies by about 100 Wm⁻²!



From Roe, 2006



the Milankovitch parameter. And what you get is at the top there.

I mean, I don't know of a better correlation in geophysics. And at the bottom, you see the ice volume itself. Of course, that doesn't look nearly as good.

Other people have independently discovered this because Edvardsson, et al. was the astronomical literature and nobody saw it.

But they also went so far as to ask whether the range of variability of insolation due to the Milankovitch parameter was compatible with the heat of fusion for the ice volume.

And even that was very, very close. Just to give you an idea of the range, that's in the bottom. Gerard Roe's paper had that. It's 100 watts per meter squared.

DR. KOONIN: Over what region is that?

DR. LINDZEN: Pardon me?

DR. CHRISTY: 65 north.

DR. LINDZEN: Yes.

DR. KOONIN: Wow!

DR. LINDZEN: That's the Arctic. So, this is big time.

Now, the question is, is the current paradigm reasonable? [next page] Is it true that there is a profound problem with the Milankovitch hypothesis because the orbital parameters leads to almost no change in globally or annually averaged insulation?

Is it really that one and a half watts per meter squared that is at issue? And I think that makes no sense.

What you have, and this is what we saw in the sensitivity measurements from space, you have huge amount of variability in clouds and other things. And they are not feedbacks.

Why aren't they degrees of

Not only does the Milankovitch parameter vary by about 100 Wm^{-2}, but Edvardsson et al showed that this was approximately the energy required to account for the melting and freezing of the ice over ice age time scales.

However, according to the currently fashionable paradigm, there is a profound problem with the Milankovitch hypothesis. The orbital parameters lead to almost no change in globally and annually averaged insolation. It is claimed that CO_2 changes are needed to actually produce the cycles of ice ages. The CO_2 changes seen in the ice cores produce about 1.5 Wm^{-2} of radiative forcing.

This seems truly absurd, given how easily the climate system can adjust to this trivial imbalance.

freedom that the system has to adjust to the small imbalances here? And I think that is probably the way one ought to look at the climate system.

This is from a paper on different models. [next page]I mean, the range of variability they are getting in precipitation in cloud radiative effects is huge compared to what you need.

But it is also on the order of our uncertainty on this and probably on the order of the normal variability.

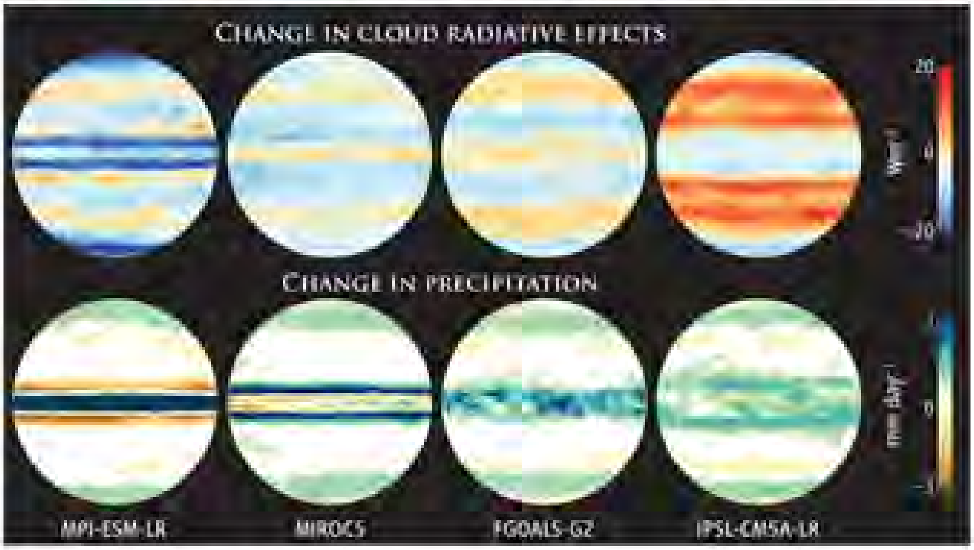
Okay, I will end it at that. I don't want to keep people from lunch.

DR. KOONIN: If I could try to summarize my own words. What I just heard in the last two minutes is that the CO_2 feedbacks are too small to plausibly play a significant role in driving the Ice Ages?

DR. LINDZEN: My feeling is that the CO_2 effects are not as

The variations among models strongly suggests that such processes are still outside of model capabilities.

Wide variation. The response patterns of clouds and precipitation to warming vary dramatically depending on the climate model, even in the simplest model configuration.



B Stevens, and S Bony Science 2013;340:1053-1054

Such changes (as well as naturally occurring changes in cloud properties) are much larger (in terms of radiative impact) than what is needed to accommodate the small radiative imbalances associated with Milankovitch forcing. Cloud fluctuations are more like degrees of freedom than feedbacks.

focused as Milankovitch. Milankovitch is telling you whether the ice survives or not.

Then you are saying, if I have ice over this and it is changing the thermodynamic balance, does the system have the capacity adjust to that? And the answer I think is almost certainly yes.

DR. KOONIN: Good, thank you. Other questions from the subcommittee, comments? From our experts? Ben has a question.

DR. SANTER: Two quick points, Dick. One, you said that in the observations, there is not much evidence of some longer-term, multiyear response to volcanic eruptions.

We certainly see that and so have many other studies, even the original Christy and McNider paper back in 1994 that statistically removed ENSO effects from lower tropospheric temperature better reveals that that long tail, the long goodbye.

And many, many studies not only with satellite data but also with weather balloon data show that longer-term response there.

So, I would disagree with the premise that there isn't some long response in the observations.

DR. LINDZEN: You think it's as big as these tails?

DR. SANTER: Well, again, we tried to look at this issue.

DR. LINDZEN: I mean, what I found was dealing with one volcano, for instance, the tail was too small to really be significant in the data.

I found that, for instance, if I looked at a single volcano, given the uncertainties, it was hard to distinguish one sensitivity from another. And people using two years to distinguish were probably, I thought, stretching things.

The only place we saw the significant tail, regardless of sensitivity, was the sequence of volcanos. If the sequence was such that one volcano came sufficiently soon after another one so that the response time included it, then you started building a secular trend.

DR. SANTER: Just to follow up on that, we have looked at this in the same way that you have with simple energy balance models and looked at the expectation of volcanic parameters, the maximum cooling, the timing of the cooling after El Chichón and what happens for different plausible ranges of sensitivity from one to five and a half.

And there are many, many things that you can see and may be able to discern occasionally that, as I tried to show there, lead me to suspect there is not some real big fundamental error in ocean heat changes after El Chichón and Pinatubo, and therefore, not some fundamental error in TCR.

DR. LINDZEN: Perhaps; I don't know.

DR. SANTER: The other thing was, there is this paper by Piers Forster, et al., that has looked at this tuning issue you mentioned.

So, they looked at total anthropogenic aerosol forcing and the relationship between that and global mean surface temperature changes over the 20th century.

As you may remember, Jeff Kiehl looked at this at CMIP3 and showed that there was some evidence of a relationship there. But Forster, et al., don't find that at CMIP5. So, I don't think there is strong evidence --

DR. LINDZEN: You are saying

for CMIP5, the aerosol adjustments are not related to the sensitivity?

DR. SANTER: Are not related to what Forster, et al., looked at, which was the size of the global mean surface temperature trend over the 20th century. I think they looked at a couple of other things as well.

But there was no evidence of that strong functional relationship that Jeff had found looking at CMIP3 results. This appeared a year or two ago in JGR, I think?

DR. COLLINS: Yes.

DR. LINDZEN: So, you are saying in CMIP5, that relation that Jeff found disappeared?

DR. SANTER: What I am saying is, there is not evidence for some strong relationship between what each modeling group did with anthropogenic aerosol forcing, total forcing and that model's global mean temperature change over the 20th century.

DR. LINDZEN: It is something that has to be looked at more closely because obvious the time scales differ according to the ocean modeling.

DR. SANTER: I agree, but it's a very different result from the CMIP3 result.

DR. LINDZEN: Interesting.

DR. KOONIN: Bill?

DR. COLLINS: This is just more a point of information. But I think Dick shared an intriguing analysis of the outgoing longwave. There is quite a lot of literature on both the how cloud changes in the tropics occur, and in water vapor feedback.

So, we saw a particular aspect this morning. There is a large body of literature on this topic. Let me just sort of end that discussion there. And I think, with all due respect, I think there is some diversity of opinion on this topic.

DR. LINDZEN: It is interesting. There is a large body of literature there and there are particularities. But I mentioned the Trenberth and Fasullo paper. When you break it into longwave and shortwave --

DR. COLLINS: Which I have done, yes.

DR. LINDZEN: -- I find it's the shortwave where you have most of the uncertainty.

DR. COLLINS: Well, in any case, I just wanted to point this out to the committee.

DR. KOONIN: Scott?

DR. KEMP: This is a general question. You wrote down and mentioned several times, and again here, and that is if you assume that the feedback parameter is normally distributed, then you get this tail in ECS?

DR. LINDZEN: If it's normally

distributed about a not high-value to begin with.

DR. KEMP: And this is used to explain why the range of ECS is large. But why has the range of ECS not changed since --who am I trying to think of -- 1984, basically? Charney. Thank you.

DR. COLLINS: To give you a very quick answer, we don't know what we can't know, and we can't go back and fix a lousy observational record. And it's just, we can't do it.

In the absence of knowing having that information on how aerosol radiative forcing, for example, has changed over the 20th century, we are stuck.

And that is a place where you can't, with temperature, you can't go back and take instruments out of the Naval Observatory in Greenwich and calibrate them against modern instruments and figure out how the temperature records, how to compare bucket records from the 18th century to the present day.

We have no such data for aerosols except for high school records.

DR. KEMP: Aerosols?

DR. COLLINS: Well, it's one of them. I don't want to be cavalier about it. But remember we are looking at a situation where, in essence, you are solving delta T which we think we know reasonably well equals lambda times delta F.

So, lambda equals delta T over delta F. And uncertainties in delta F are a problem because they appear in the denominator.

DR. KOONIN: Ben, did you want to comment?

DR. SANTER: I think I will defer to Isaac here.

DR. HELD: I was going to respond to the same question, because

I think the answer is different depending on whether you are talking about top-down or bottom-up constraints.

And Bill is talking about the top-down constraint. You have observed warming. You are trying to understand it. And there the problem is aerosols. The problem is forcing, basically, by definition.

But as far as bottom-up, I think the answer is also one word. It's clouds. It's clouds that prevent us from fundamentally in some reductive fashion understanding the climate system. They are two different things.

DR. COLLINS: Yes.

DR. KOONIN: Okay, good. A good morning. Let us take 20 minutes to grab lunch and begin eating and we will do whatever else we need to do and pick up about 12:30 or 12:35.

(Whereupon, a luncheon recess

was taken.)

DR. KOONIN: All right, we are going to continue. For those of you who did not get a chance to sample the cookies or the brownies, in a few minutes we are going to have cheesecake coming in from Junior's.

All right, John?

DR. CHRISTY: It is a real pleasure to be here. And I particularly was pleased to see the way you had framed the discussion here and the questions that you had.

There are many that those of us in the climate field do have and wonder about when something like the IPCC presents a report as it did.

My main aspect in this endeavor is that I am one of those people that builds climate data sets. So, whether it is the digital count from a microwave sensor in space or a dusty archive in the UK Met office, I get those data to create climate data sets to basically tell us what is happening with the climate as best we can.

So, I boiled some of your framing questions by this blue expression here. [next page]

One of the things you asked is, "Why did confidence regarding the assertion that human influences dominate the climate system increase in AR5 when (A), so many of the climate processes are poorly known and modeled, and (B), the global temperature failed to warm as expected?"

And it kind of filtered in through several of those places in the framing document. And really, the truth is the answer must come from the convening lead authors of the IPCC AR5 because I am baffled. And that is exactly what I told the Congressional committee just a month ago.

Why did “confidence” regarding the assertion that human influences dominate the climate system, *increase* in AR5 when (a) so many of the climate processes are poorly known and modeled, and (b) the global temperature failed to warm as expected?

[Sections I.1, I.2, II, IV of Framing Document]

Answer must come from the Convening Lead Authors of the IPCC AR5 because I am baffled. (Much of IPCC background text is reasonably done.)

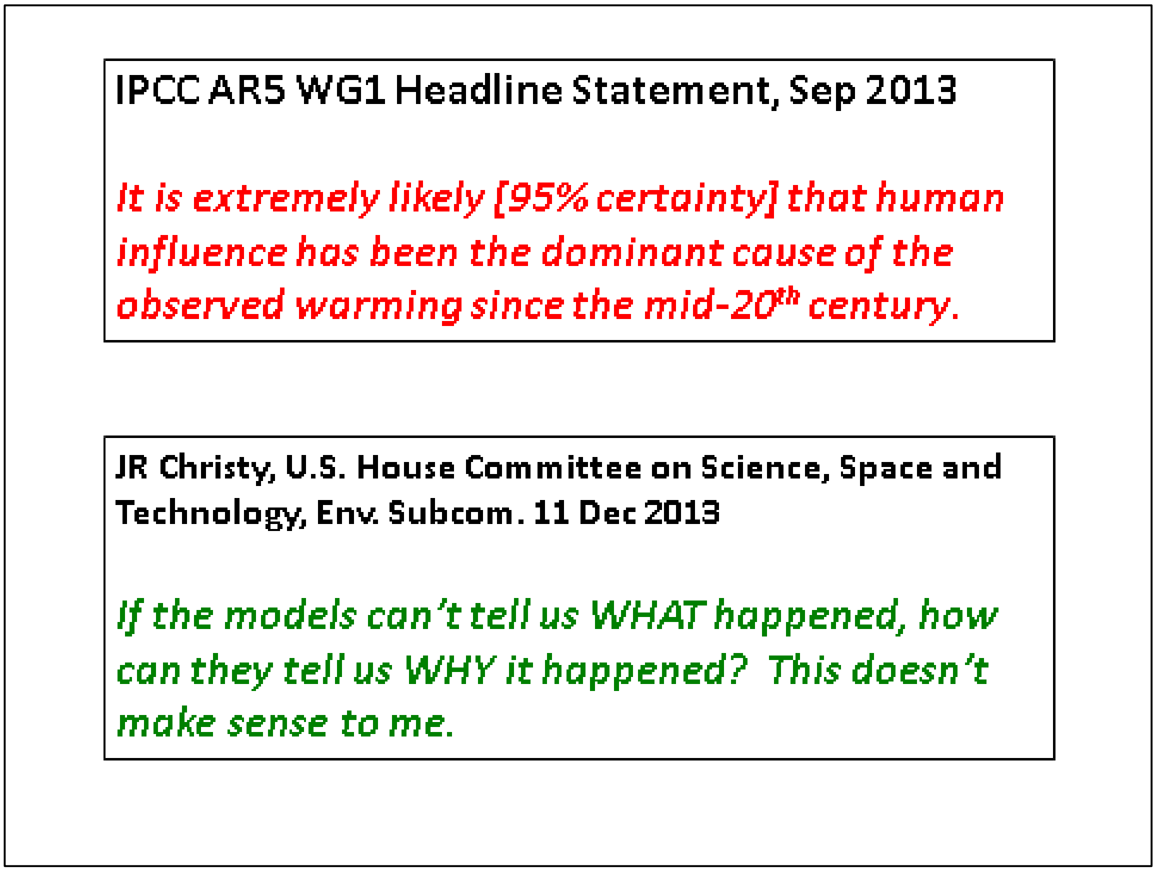
The only way to tell how much is human vs. natural is through model simulations.

Now, as you probably saw much of the background material and the text of the IPCC was reasonable. It has lots of caveats and concerns and so on. But when it came down to the final statements, it really wandered.

Well, the only way to tell how much global warming is due to human or natural is basically through model simulations because we found out that we can't put a thermometer out there that will say this much was due to Mother Nature and this much was due to Mankind.

We just don't have instruments like that. So, using models is the way to do this.

And the statement [next page] that is explicit in this from the IPCC is, "It is extremely likely," and that meant 95 percent certainty, "that human influence has been the dominant cause of observed warming since the mid-20th century."



IPCC AR5 WG1 Headline Statement, Sep 2013

It is extremely likely [95% certainty] that human influence has been the dominant cause of the observed warming since the mid-20th century.

JR Christy, U.S. House Committee on Science, Space and Technology, Env. Subcom. 11 Dec 2013

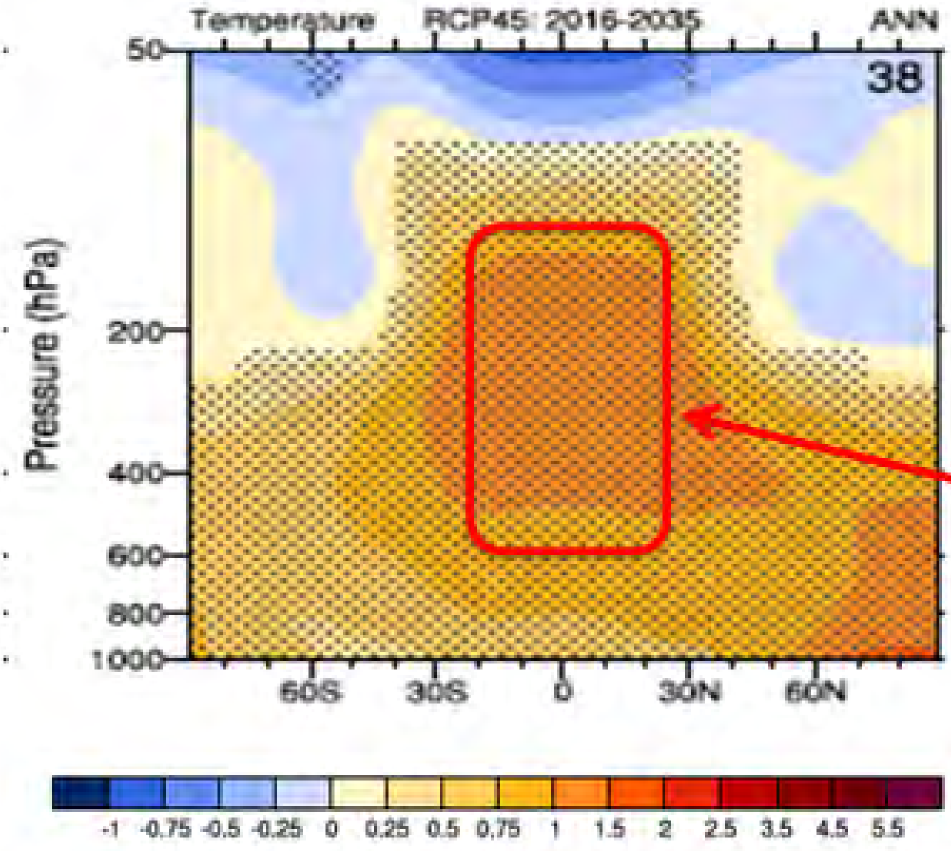
If the models can't tell us WHAT happened, how can they tell us WHY it happened? This doesn't make sense to me.

So, as I said a month ago, if the models can't tell us what happened, how can they tell us why it happened? This doesn't make sense to me. So, I will explain to you why this doesn't make sense to me.

Ben showed this or a similar figure to this. [next page] This is a cross-section of the atmosphere, so the North Pole, South Pole surface, stratosphere. This is the tropics. Huge amount of mass right here.

If you want to look at something that has a greenhouse signature from model simulations, that would be the place to do it because it has the biggest signal, the most mass.

So, now we are talking about the joules, the most joules of energy that are going to affect the system. And so right there it's commonly called the tropical hot spot response in climate models.



Tropical tropospheric "Hot Spot" is a major signature of GHG response in models

CMIP-5 RCP4.5 Model Runs

DR. KOONIN: This is one pole of the dipole that Ben was talking about?

DR. LINDZEN: No, it really isn't, and that has been bothering me a little. Point of information, the hot spot is the temperature maximum near the upper troposphere in the tropics. That's due to the moist adiabatic.

The dipole is the difference between warming in the troposphere and cooling.

DR. KOONIN: That's right. As I said, it's one pole. The upper part of it is one pole.

DR. LINDZEN: But the structure of the lower part is the hot spot.

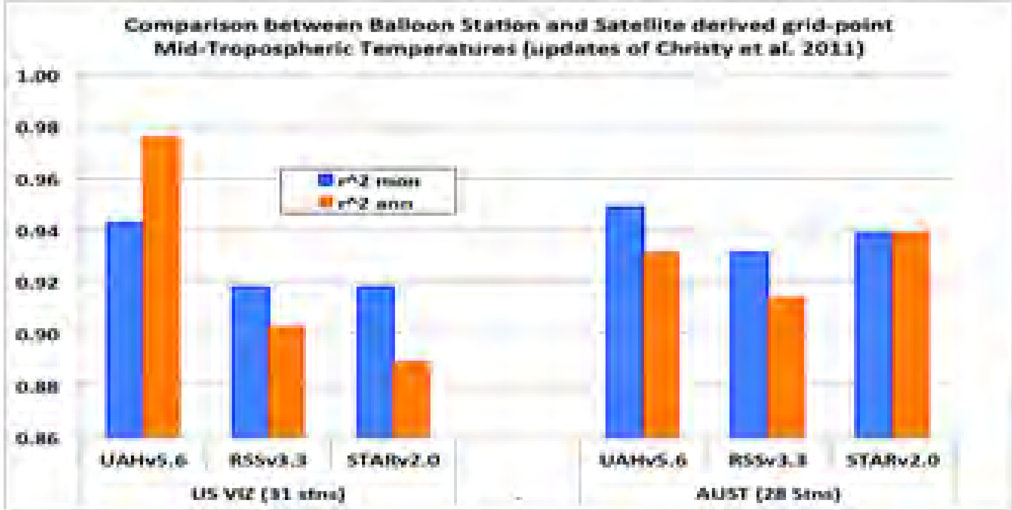
DR. CHRISTY: Yes, just right now we are looking at that part because it's a big signal. Just look at the picture and you will say that is a target that we ought to be able to hit because it's so big and strong.

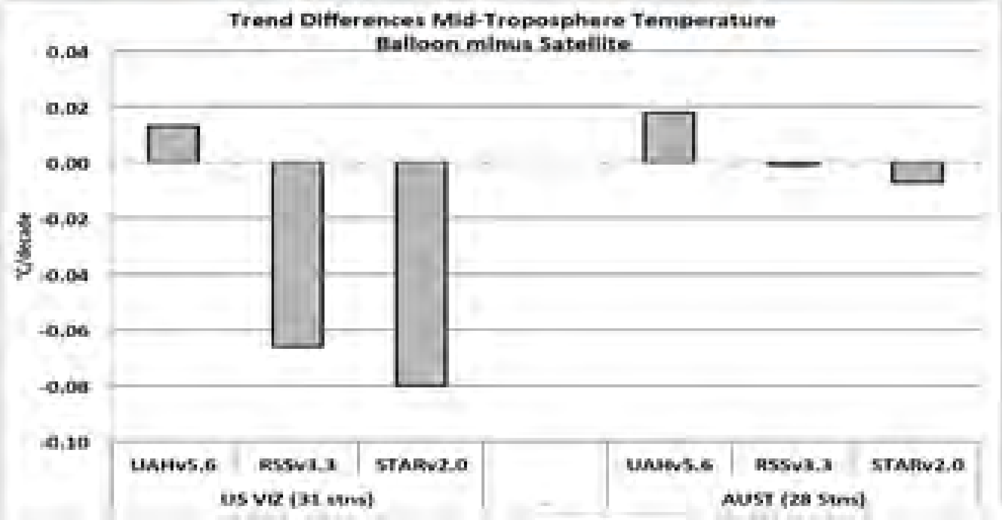
Now, to do that, we can use radiosonde balloons which are balloons that go up and take the temperature at every elevation. You can get the bulk temperature that way, or microwave emissions from oxygen molecules tell us the intensity or their intensity is proportional to temperature.

And up here,[next page] I just have a small comparison studies that were published earlier. And you can see, the main thing I want you to see is what the R squares are here.

Balloons at a particular station, what the satellite sees at that same place, so just, how well do they compare? And these R squares are in the mid-.9s and above for the three different satellites data sets we are showing here.

So, my view is that we have tremendous skill at understanding





AR5 claimed to have "only low confidence" in observations (2-4, 9-30, 10-28)

what the tropical tropospheric temperature is doing because of these kinds of independent measurements.

The IPCC said, however, we have only low confidence in the observations. And that bled into later chapters where they said well, the models and observations don't agree, but that could largely be due to poor observations.

But I don't think that's the case. I think we do have good information on observations and we have pretty good confidence.

DR. KEMP: This is not because of a question of old radiosonde data? Is this comparing the old weather balloon data to current data only?

DR. CHRISTY: No, the one on the left is comparing the United States VIZ stations. So, that is a single type of radiosonde that was launched from the tropics to Port Barre, Alaska, 31 stations.

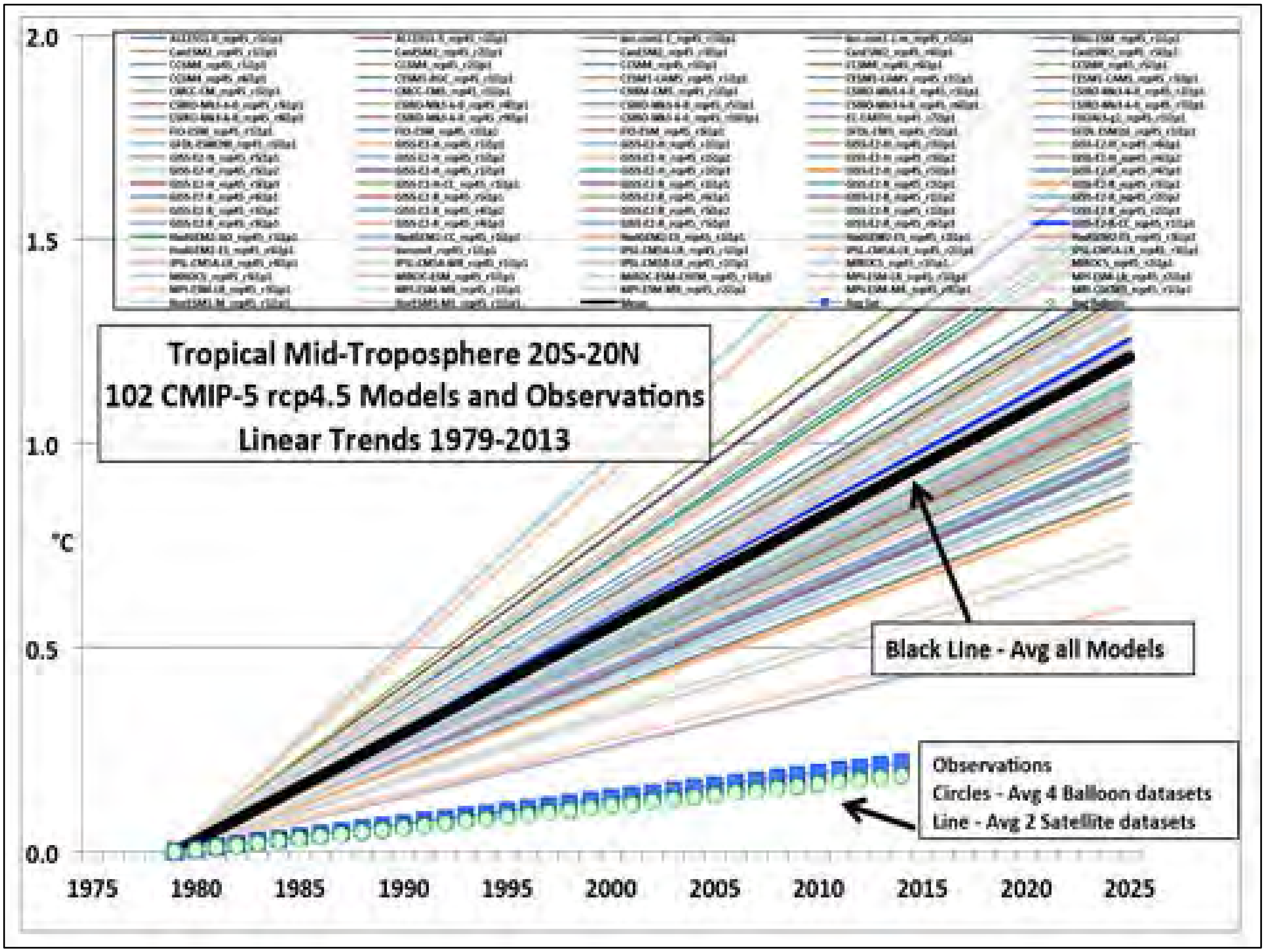
At those points is where we take our satellites measurements as well. And they line up. I mean, getting an R squared of .98, for atmospheric science, this is another planet.

DR. KEMP: I thought the AR5 statement was related to historical radiosonde data. That's what I thought.

DR. CHRISTY: Well, these data do go back to 1979 when the satellite launched, so you could call them historical in that sense.

The key thing is, IPCC is correct. A lot of radiosonde data is not very good at all. But where you take the best radiosonde data, do the comparisons with the satellite data, then you get this kind of result.

So,[next page] let's look at climate model simulations just in the simple metric of linear trend from 1979 when satellites started onward. And there


Tropical Mid-Troposphere 20S-20N
102 CMIP-5 rcp4.5 Models and Observations
Linear Trends 1979-2013
Black Line - Avg all Models
Observations
Circles - Avg 4 Balloon datasets
Line - Avg 2 Satellite datasets
°C
2.0
1.5
1.0
0.5
0.0
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
2025

I have the spread of 102 RCP 4.5 model runs with both the balloons and the satellites below.

In every case, all 102, they are much warmer than the observations showed. So, this is a 35-year trend depiction. It is not a 15-year trend projection. So, this is over a third of a century we are looking at here.

And I think what you see is the observations from the two independent data sets are almost on top of each other, whereas the models have a huge spread and every single one is warmer than the observations, and the average is quite a bit warmer.

Ben?

DR. SANTER: Sorry, John. Are the model trends plotted out to 2025? Are they estimated over '79 to 2025?

DR. CHRISTY: Yes, they were just extended from 2013.

DR. SANTER: So, you calculated the model trends over '79 to 2025?

DR. CHRISTY: Yes, all trends from 1979 to 2013. And those are just extrapolated. But the next picture will demonstrate that --

DR. KOONIN: Could I just go back. This is RCP4.5.

If you used the lower one, what is it, 2.8 or something like that?

DR. CHRISTY: They were all the same.

DR. KOONIN: They were locked in, basically?

DR. CHRISTY: Right. They don't start diverging until about 2030, 2040.

DR. KOONIN: And again, just to really emphasize what Ben asked: the models have actually been run from 1980 to 2013, and you just extended the lines up?

DR. CHRISTY: The models have 1860 to 2100.

DR. KOONIN: But you calculated the trend in this picture?

DR. CHRISTY: Apples to apples in this picture, apples to apples.

DR. SANTER: So, these are synthetic MSU temperatures?

DR. CHRISTY: Yes.

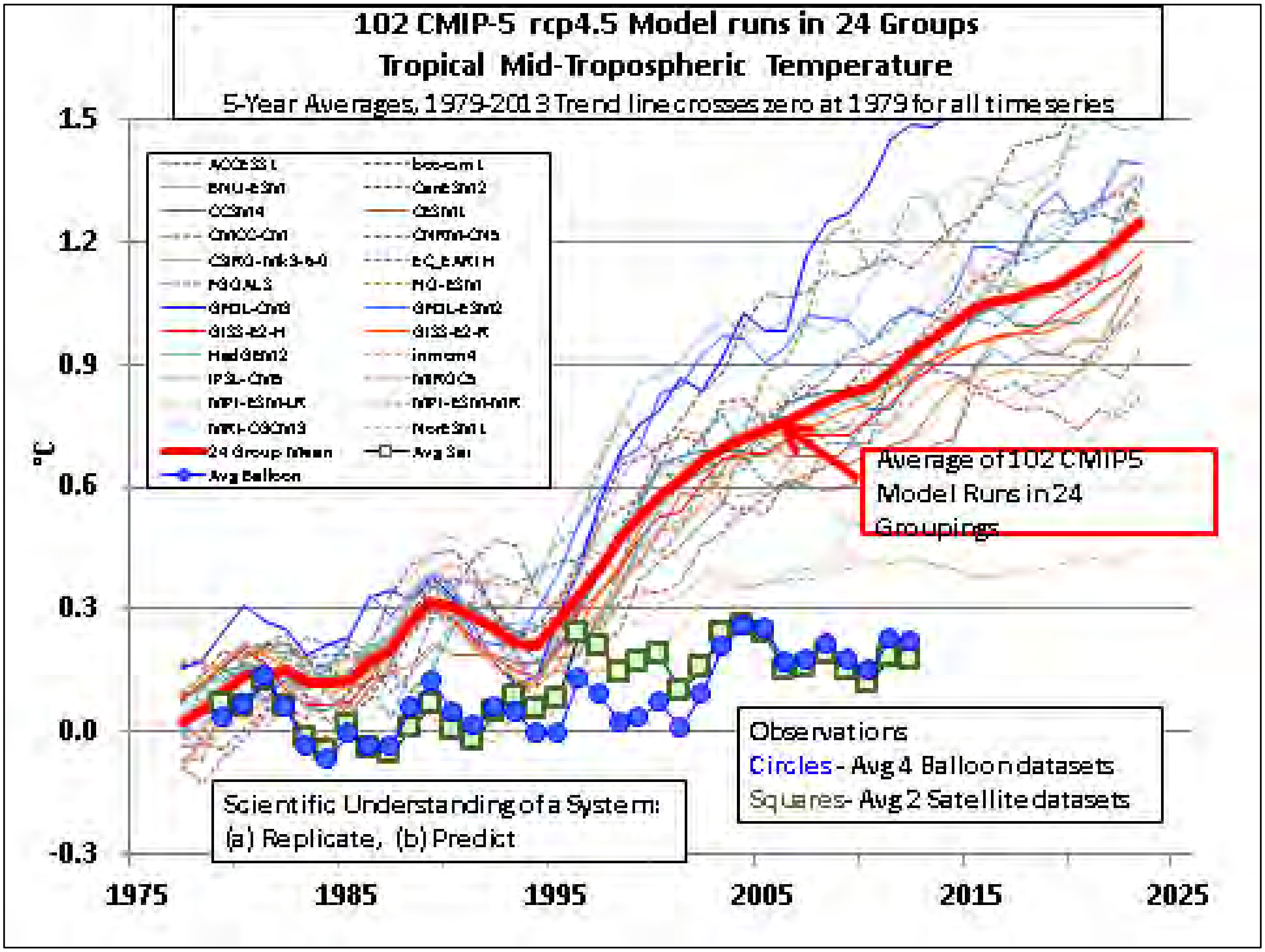
DR. SANTER: That you have calculated from the models?

DR. CHRISTY: Yes.

DR. SANTER: Using a global mean weighting function-type approach?

DR. CHRISTY: A tropical mean weighting function which, by the way, was also done with the balloons. So, they were identical in that way.

This [next page] now gets you to the five-year running average of all those things, except instead of 102 realizations, and those were all that were available to me at the time, I added them together into 24 families of models so that you can see how the spread occurs from the different families in five-year averages.


102 CMIP-5 rcp4.5 Model runs in 24 Groups
Tropical Mid-Tropospheric Temperature
5-Year Averages, 1979-2013 Trend line crosses zero at 1979 for all time series
1.5
1.2
0.9
0.6
0.3
0.0
-0.3
°C
1975
1985
1995
2005
2015
2025
ACCESS1
bcc-csm1
BNU-ESM
CanESM2
CCSM4
CESM1
CMCC-CM
CNRM-CM5
CSIRO-Mk3-6-0
EC_EARTH
FGOALS
FIO-ESM
GFDL-CM3
GFDL-ESM2
GISS-E2-H
GISS-E2-R
HadGEM2
inmcm4
IPSL-CM5
MIROC5
MPI-ESM-LR
MPI-ESM-MR
MRI-CGCM3
NorESM1
24 Group Mean
Avg Sat
Avg Balloon
Average of 102 CMIP5 Model Runs in 24 Groupings
Observations
Circles - Avg 4 Balloon datasets
Squares - Avg 2 Satellite datasets
Scientific Understanding of a System:
(a) Replicate, (b) Predict

And I think you can see that, for this particular system, there is a lot of concern because none of the models were able to come within the range of observations there.

And the general rule is, if you have a good, confident understanding of a system, you ought to be able to at least replicate what it does and then predict what it does.

And I think can you see here that really none of the models were able to do that. And most, a great majority of them did not do it closely at all.

DR. KOONIN: These models also do not reproduce the surface temperature; is that correct?

DR. CHRISTY: Oh, they do.

DR. KOONIN: They do reproduce surface temperature?

DR. CHRISTY: Yes. Oh, I mean, they have the surface temperature in them. I don't have a surface temperature plotted here.

DR. KOONIN: I am asking whether the same models reproduce GMST and the error is in the vertical structure, or they also do a bad job on GMST?

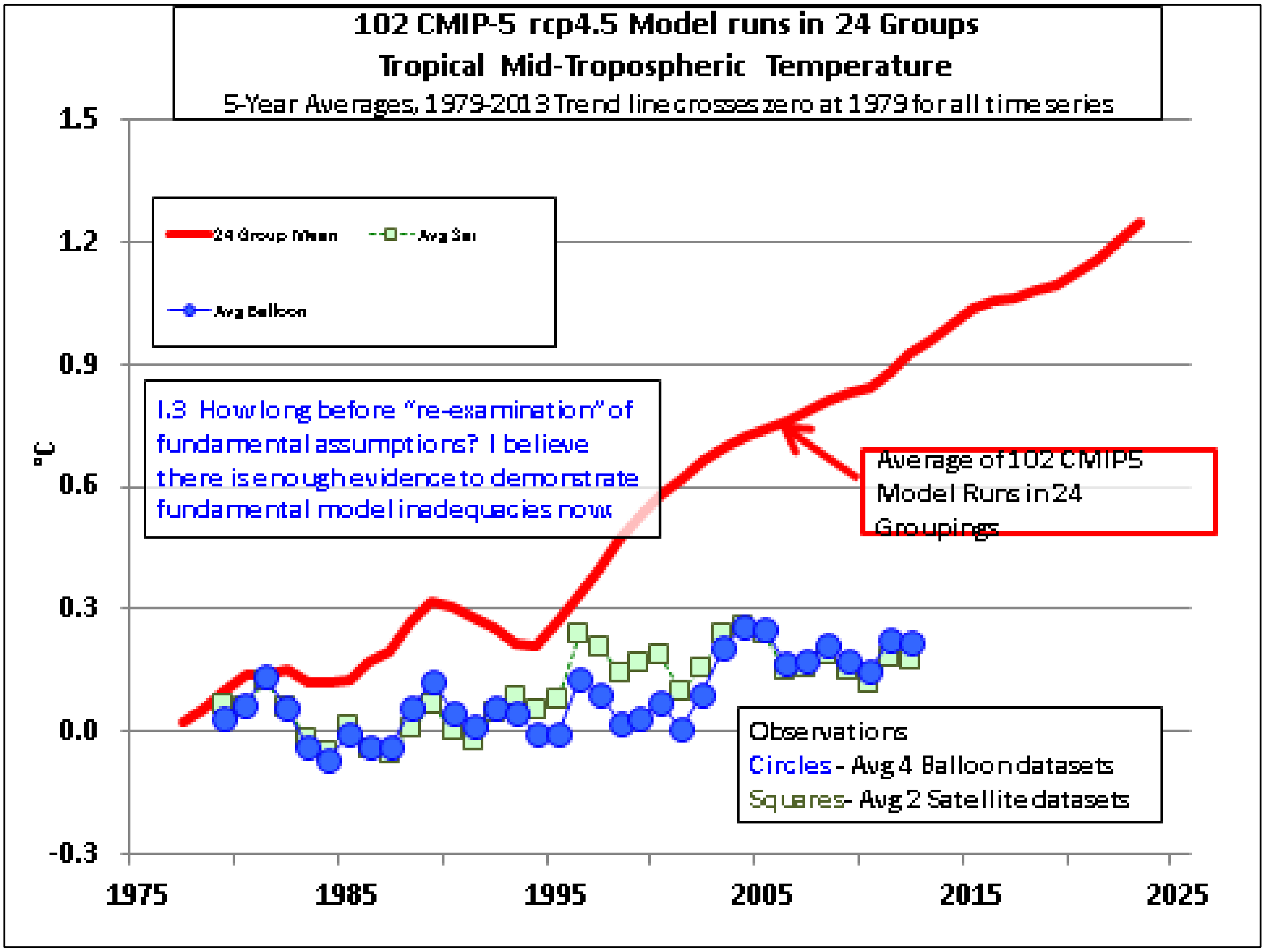
DR. CHRISTY: It is not as bad as this on GMST. It looks more like this (indicating slide).

DR. KOONIN: Yes. Judy showed some of that. We have some of that.

DR. CHRISTY: But as physicists, I hope you would understand what I am looking at is the big, mass bulk of the atmosphere where there is lots of kilograms of air and lots of joules are going to make a difference.

So, that's what you want to measure, whereas I will show a little later why I don't like to use surface temperature for these kinds of studies.

So, [next page] that's just taking away all



102 CMIP-5 rcp4.5 Model runs in 24 Groups
Tropical Mid-Tropospheric Temperature
5-Year Averages, 1979-2013 Trend line crosses zero at 1979 for all time series

1.3 How long before “re-examination” of fundamental assumptions? I believe there is enough evidence to demonstrate fundamental model inadequacies now

Average of 102 CMIP5 Model Runs in 24 Groupings

Observations
Circles - Avg 4 Balloon datasets
Squares - Avg 2 Satellite datasets

the different realizations and showing you just the average. And I think from your question 9.3, "How long before reexamination of the fundamental assumptions," I believe we are already there, that the fundamental assumptions need to be examined because, before the most recent 15-year hiatus occurred, models were already over what the atmosphere was doing.

DR. SEESTROM: Question, what was the basis for the groupings, difference science in the models?

DR. CHRISTY: Oh, no, the organization. So, like, GFDL I think had two groupings of their model runs.

DR. KOONIN: Can you go back one?

DR. CHRISTY: So, when you get this, you can see GFDL. I have two. They are the blue ones. They were pretty hot, by the way.

DR. KOONIN: If I were to phrase this in terms of TCR?

DR. CHRISTY: I really don't want to get into that, but I have one slide about that.

DR. HELD: The key thing is tropical versus global.

DR. KOONIN: Right, okay, fair enough, fair enough.

DR. HELD: It's not just the vertical dimension.

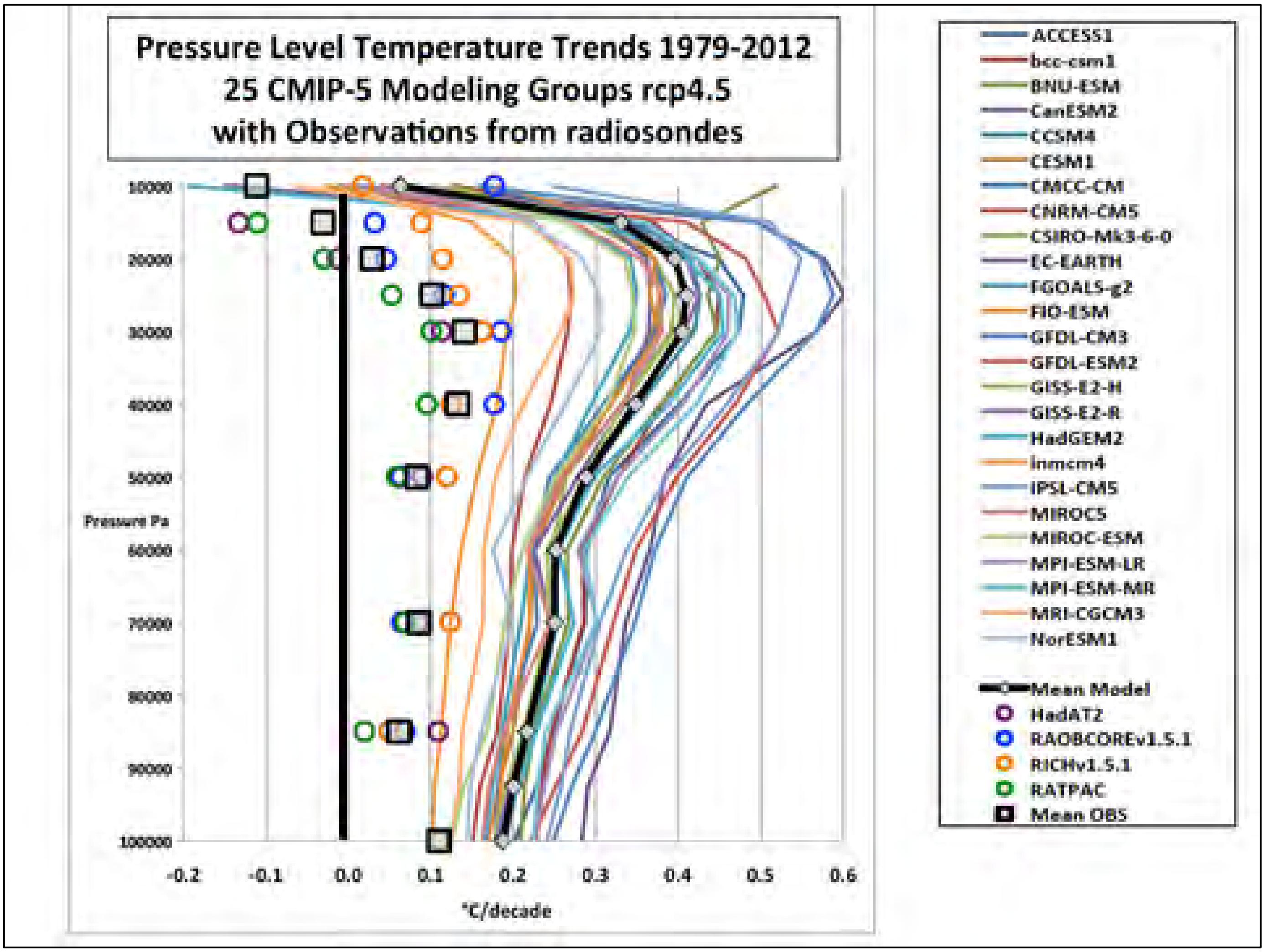
DR. CHRISTY: The reason I do the tropical is that's where the signal is.

DR. HELD: I am just trying to clarify.

DR. KOONIN: Good, I learned something!

DR. CHRISTY: So, we are through with that.

Now let's go in the vertical dimension. [next page] It's still in the tropics, but from the surface all the way up to the stratosphere. And you


Pressure Level Temperature Trends 1979-2012
25 CMIP-5 Modeling Groups rcp4.5
with Observations from radiosondes
10000
20000
30000
40000
50000
Pressure Pa
60000
70000
80000
90000
100000
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
°C/decade
ACCESS1
bcc-csm1
BNU-ESM
CanESM2
CCSM4
CESM1
CMCC-CM
CNRM-CM5
CSIRO-Mk3-6-0
EC-EARTH
FGOALS-g2
FIO-ESM
GFDL-CM3
GFDL-ESM2
GISS-E2-H
GISS-E2-R
HadGEM2
inmcm4
IPSL-CM5
MIROC5
MIROC-ESM
MPI-ESM-LR
MPI-ESM-MR
MRI-CGCM3
NorESM1
Mean Model
HadAT2
RAOBCOREv1.5.1
RICHv1.5.1
RATPAC
Mean OBS

can see that, yes, you can say you have low confidence in a balloon measurement at 500 millibars because at the .01-degree C per decade, you have low confidence.

But how can you say you have low confidence when all the range of results are here and these are all the model projections?

And so, I think I would fault the IPCC for saying since we have low confidence, we are not going to talk a lot about the disagreement that we see in this diagram. In my reviews, I hit and hit. I wanted this picture in there.

DR. KOONIN: Is there a diagram like this in IPCC?

DR. CHRISTY: No. And you can read and when the reviews will be published sometime way down the road when all of this has blown over, you will see people like me and others are saying please show pictures like this to demonstrate where the models are right now.

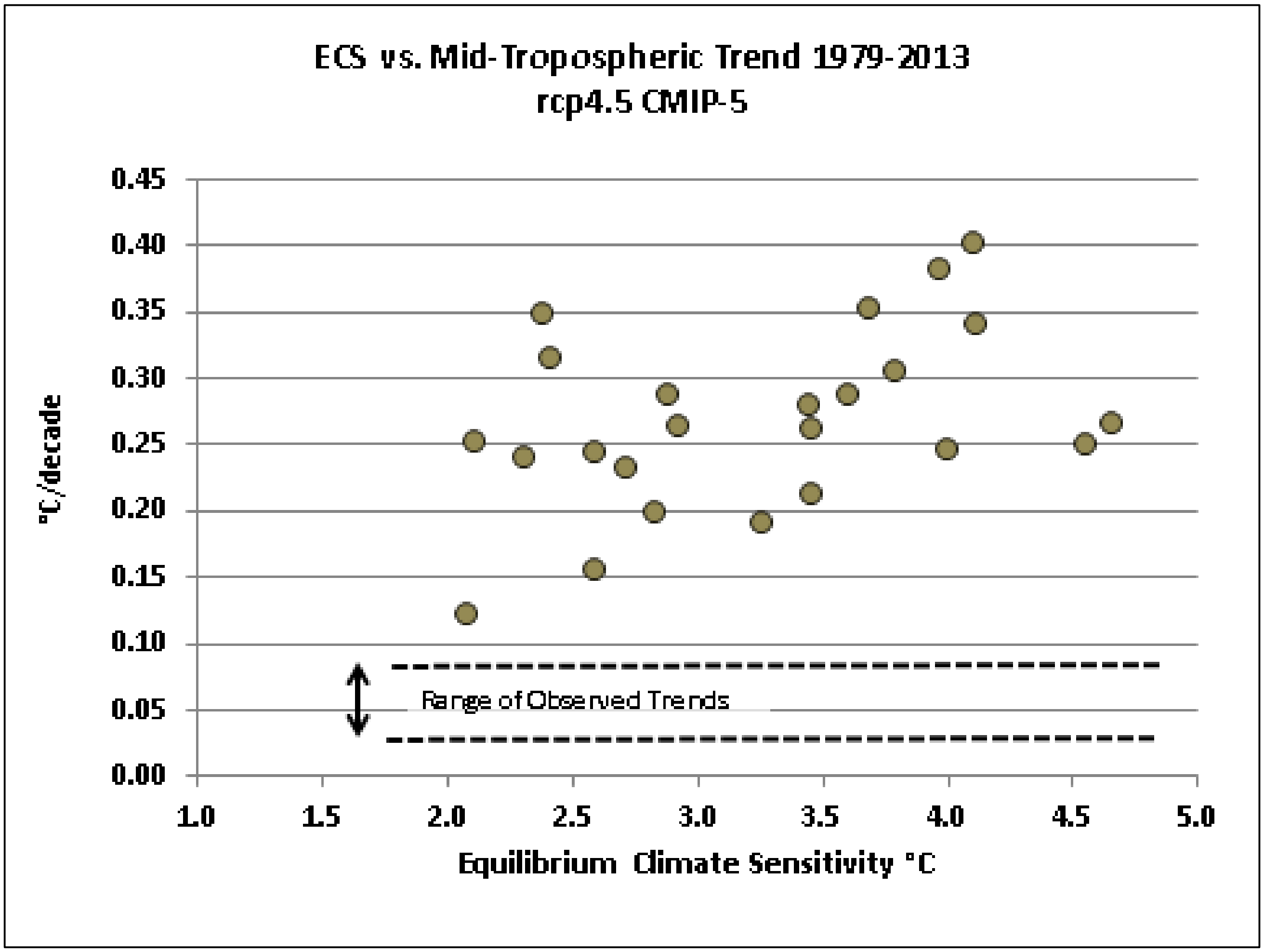
And they were pretty much ignored. There was one that started in 1961, but it's an odd one. No satellite comparisons were done.

Then, what about sensitivity? This [next page] is those same trends, but now organized by equilibrium climate sensitivity. And it's pretty simple. I mean, the more sensitive the model was the worse it did in terms of reproducing the action.

And this is the entire range of the observations. And I think this one is too hot. That one is probably too cold. But that is the entire range of the observations, both balloon and satellite.

DR. KOONIN: And the horizontal axis is the global equilibrium of climate sensitivity in terms of the model run?

DR. CHRISTY: Right, right.


ECS vs. Mid-Tropospheric Trend 1979-2013
rcp4.5 CMIP-5
°C/decade
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Range of Observed Trends
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Equilibrium Climate Sensitivity °C

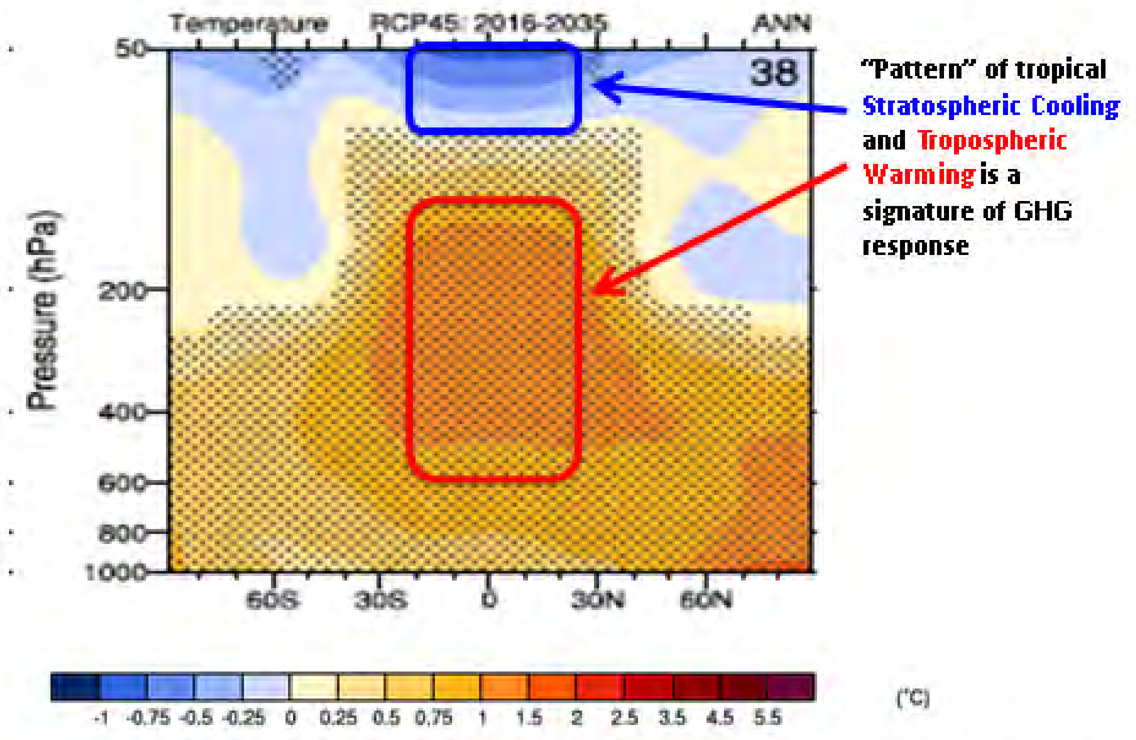
DR. KOONIN: It's not as great a correlation, at least, as I would have thought.

DR. CHRISTY: Well, you could throw a line through it, but I didn't. I did throw a line in there to see where it would intersects, and it didn't come out with a good picture.

Now, Ben brought up the diagnostic tool of the pattern stratosphere cooling and the tropospheric warming which is very strong here. [next page]

But if you actually look at the real data like this, [next page] you will find that, in the stratosphere, we have the warming from El Chichón and Pinatubo here. And since that has happened, nothing has happened. In fact, the global is also no trend in the last 20 years.

DR. KOONIN: Even the stasis, while we can say 1999 or something is

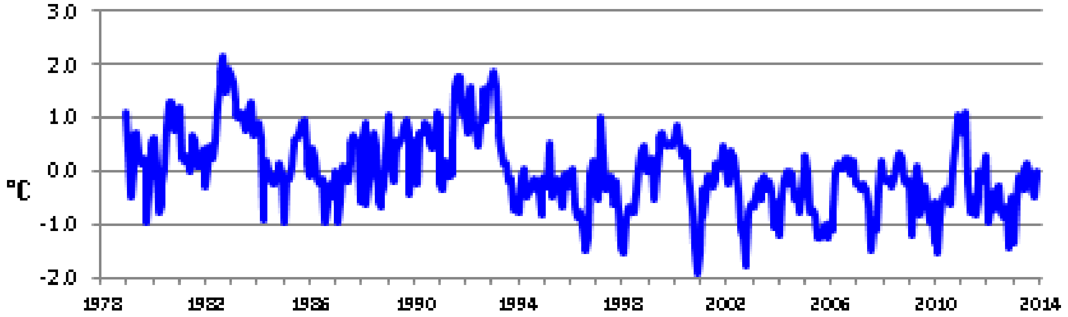


"Pattern" of tropical Stratospheric Cooling and Tropospheric Warming is a signature of GHG response

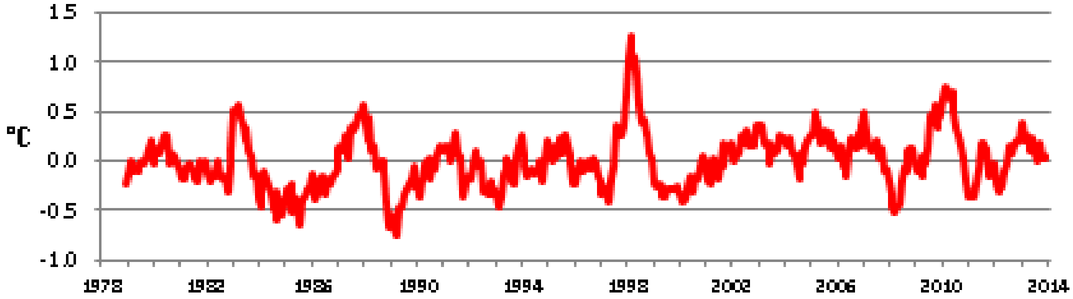
CMIP-5 RCP4.5 Model Runs

The "Greenhouse" Pattern of Stratospheric Cooling and Tropospheric Warming has been absent in the past 20 years

Tropical Lower Stratosphere (UAHv5.6)



Tropical Lower Troposphere (UAHv5.6)



when the stasis started, it was already dropping in --

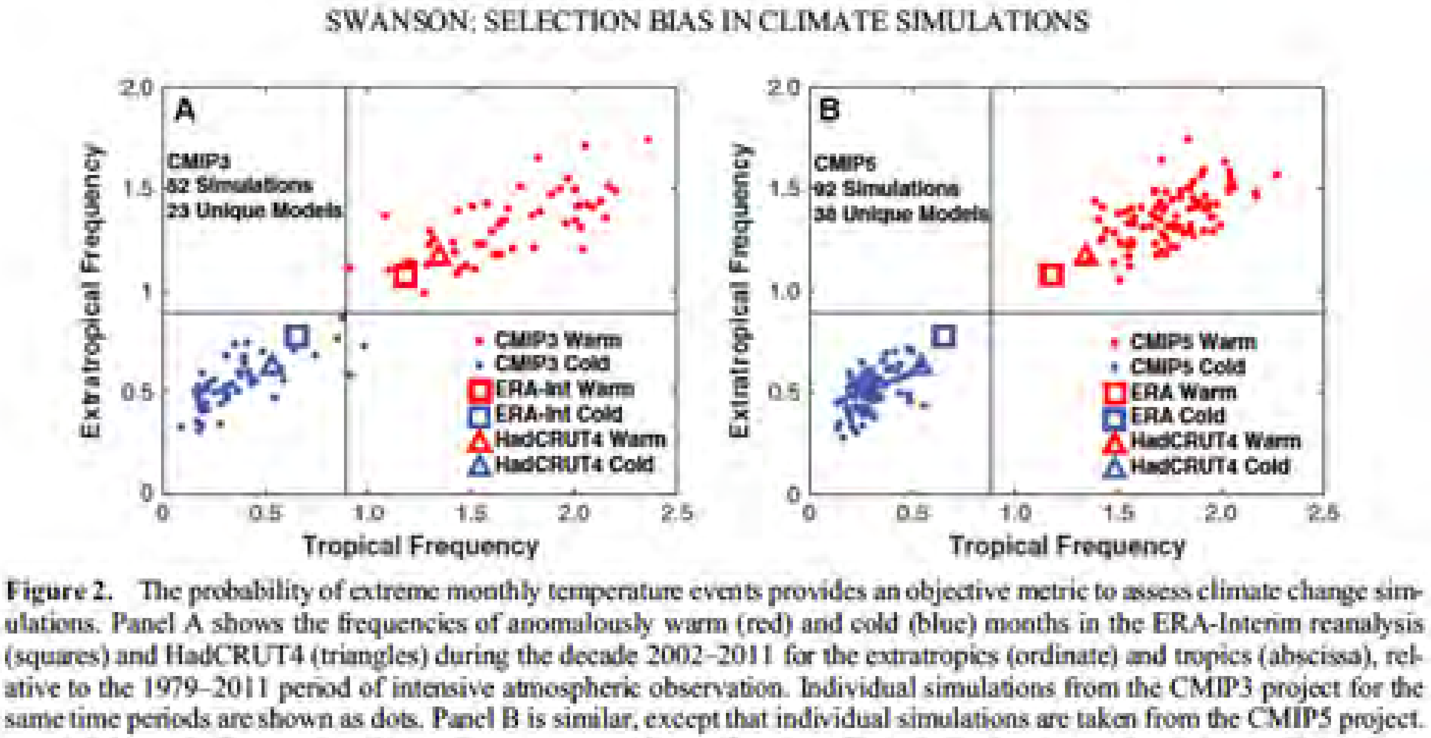
DR. CHRISTY: This is stratosphere. This is the cold part.

DR. KOONIN: Oh, sorry, okay.

DR. CHRISTY: So, what you do see here in the tropics is, boy, lots of ups and downs, a huge amount of variance just explained by the El Niño southern oscillation. And not much has happened at all. In fact, there is hardly any trend there overall.

This [next page] is a paper by Swanson I think that someone referred to before. Just one of the interesting conclusions -- it came out about a month ago -- is that, in his analysis, the CMIP5 models are worse than the CMIP3 models because they cluster further away from observational metric.

This is just an odd metric. But there is the metric there and the



SWANSON: SELECTION BIAS IN CLIMATE SIMULATIONS

Figure 2. The probability of extreme monthly temperature events provides an objective metric to assess climate change simulations. Panel A shows the frequencies of anomalously warm (red) and cold (blue) months in the ERA-Interim reanalysis (squares) and HadCRUT4 (triangles) during the decade 2002–2011 for the extratropics (ordinate) and tropics (abscissa), relative to the 1979–2011 period of intensive atmospheric observation. Individual simulations from the CMIP3 project for the same time periods are shown as dots. Panel B is similar, except that individual simulations are taken from the CMIP5 project.

Swanson 2013, GRL. "The situation here, with convergence apparently rooted in the desire to capture one particular regional signature [Arctic warming] is difficult to justify."

VI. What metrics led to selection or rejection of models?

The climate modeling industry needs to be subject to independent "red teams" evaluation processes.

observations. The model results in CMIP5 cluster and go away from what had happened in CMIP3.

And so, I don't know how you would select the best models under something like that.

DR. KOONIN: One of the questions comes to me as I listen to the discussion about the troposphere is, how important is it that the models get that right?

To phrase it maybe in a crude way, are you picking some minor feature of the climate system that, okay, it doesn't really matter whether you get it right or not, or is this kind of the nut of the problem we are talking about?

DR. KEMP: Can I add onto that? If the stratosphere cooling is the signature, but you are going back further, I was just trying to figure out if you would expect it to disappear with the stasis if the stasis were to perform internal variability?

DR. CHRISTY: Well, what surprises me is this doesn't have a more downward trend because of the ozone issue. That is really what is driving it.

DR. LINDZEN: What height is that?

DR. CHRISTY: That's about 70, 60 millibars is the average. It has a piece of the troposphere in it.

DR. KOONIN: Could I get a clean response to the question I asked without Scott's addendum? Do you think that the models don't reproduce the observations?

DR. CHRISTY: That was what I tried to -- well, no, let's just go right to it.

DR. KOONIN: The answer is "yes"?

DR. CHRISTY: That's the biggest target you have to shoot at.

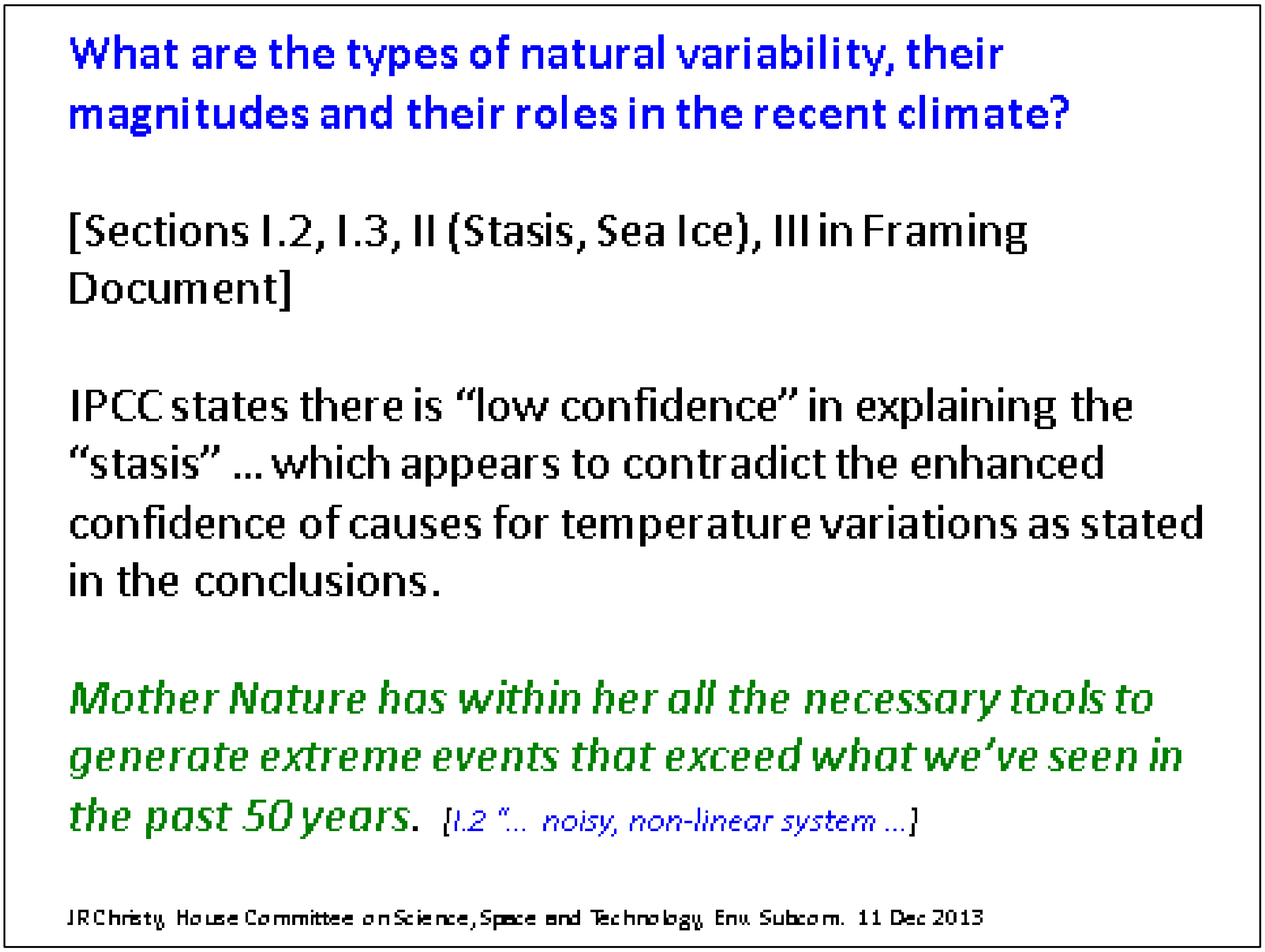
And what I would say is that there have been 112 shots, 110 shots taken at that target and they all shot high. And yet you come up and say, we have more confidence than we are getting it right. It doesn't make sense to me.

DR. KOONIN: Okay.

DR. CHRISTY: What about natural variability, their magnitudes and roles in the recent climate? [next page] You mentioned this in these sections here.

The IPCC states, "There is low confidence in explaining the stasis." And I am right there with them. I can't explain it for you either.

My comment to the committee when something like that was asked to me a month ago was, "Mother Nature has within her all the necessary tools to generate extreme events that exceed what we have seen in the past 50 years."



What are the types of natural variability, their magnitudes and their roles in the recent climate?

[Sections I.2, I.3, II (Stasis, Sea Ice), III in Framing Document]

IPCC states there is “low confidence” in explaining the “stasis” … which appears to contradict the enhanced confidence of causes for temperature variations as stated in the conclusions.

Mother Nature has within her all the necessary tools to generate extreme events that exceed what we’ve seen in the past 50 years. *[I.2 “… noisy, non-linear system …]*

J R Christy, House Committee on Science, Space and Technology, Env. Subcom. 11 Dec 2013

So, whatever we have seen out there, Mother Nature already has the ability to do it.

And to back that, let's start with millennial and centennial variability. You see here [next page] Greenland temperatures for the last 10,000 years.

Here is the Medieval warm period, even warmer than it is today. From borehole temperatures, much warmer in the mid-Holocene period.

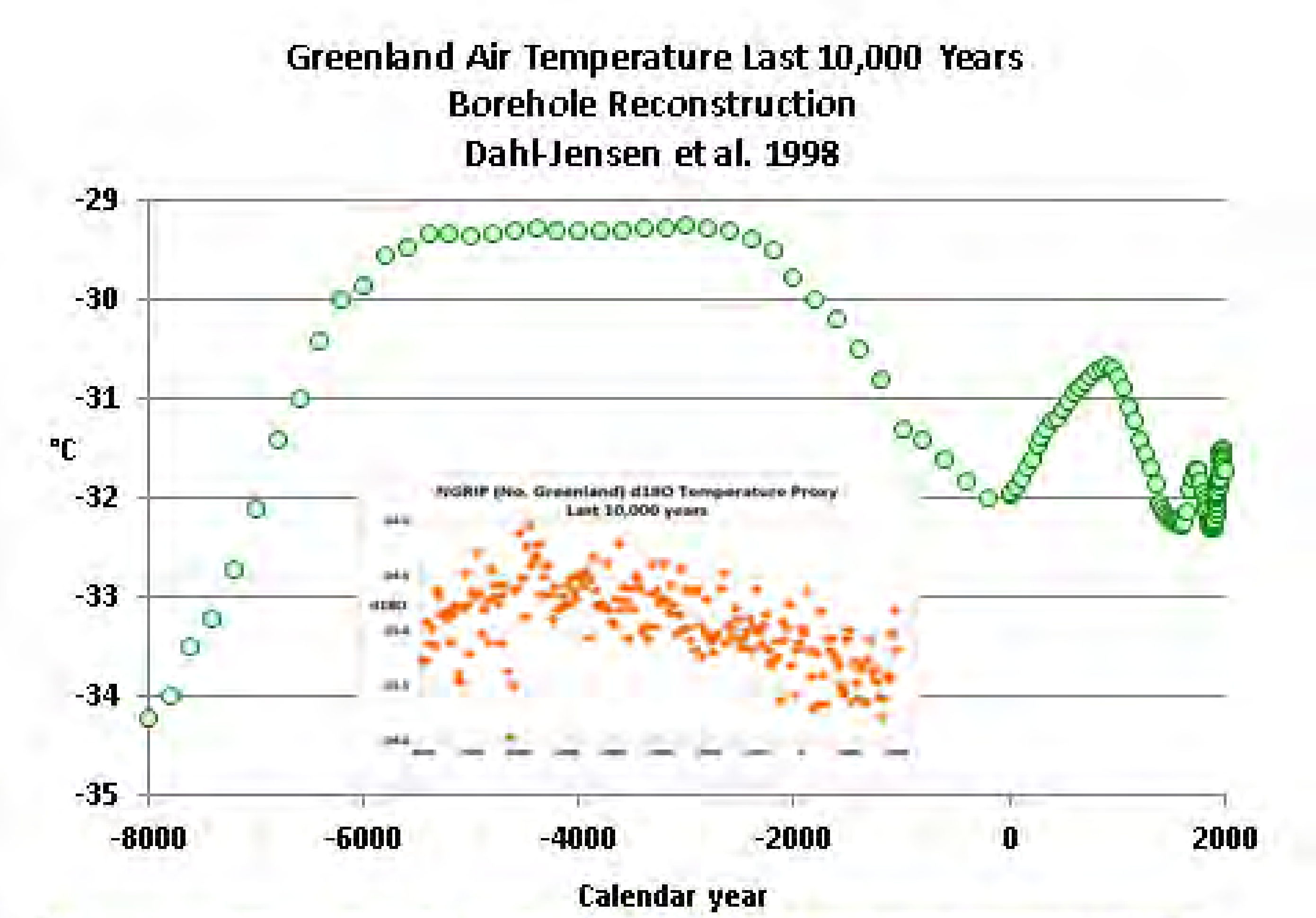
And this is the same scale that scrunched up the last 10,000 years from a completely independent measurement, the oxygen isotope temperature measurement.

DR. KOONIN: Do the guys who made the green points believe in error bars?

DR. CHRISTY: What's that?

DR. KOONIN: Do the green guys believe in error bars?

DR. CHRISTY? I took this from their chart. They have this kind of



If snowcover does not melt in summer in No. Quebec, Ice Age may be coming [I.3] thing going with it, but I did it really fast.

DR. KOONIN: Okay, but they do.

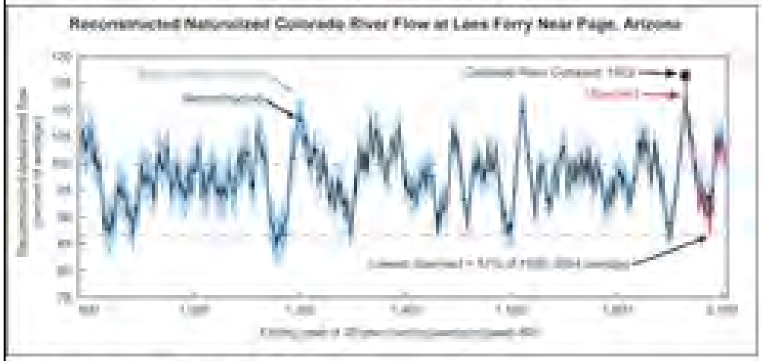
DR. CHRISTY: But this is outside the error from this. So, as we mentioned earlier, your question 9.3, "If snow cover does not melt in the summer in northern Québec, Ice Age might be coming."

That was millennial/centennial scale. Here [next page] is sort of a centennial scale in terms of climate. This is drought and water resource problems.

This is from the Colorado River flow. And you see in the past there have been centuries that the so-called megadroughts, 11th, 12th 13th century shown here in a tree-ring reconstruction as well, huge droughts that occurred.

But I like this picture the best. I like to show two because these are taken in Alpine Lakes in California in the Sierra Nevada that

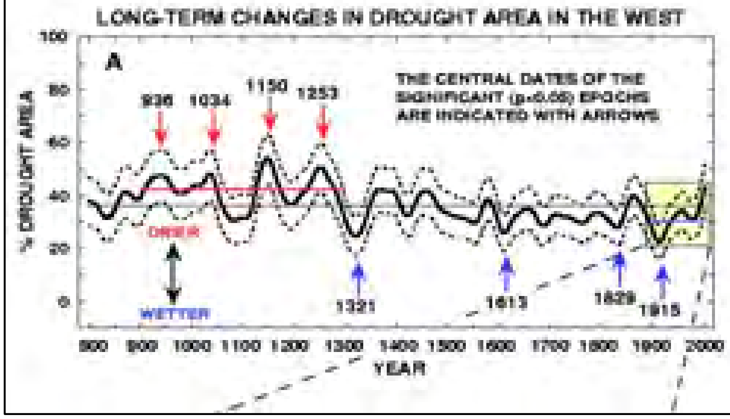
Variability in Western U.S. Precipitation

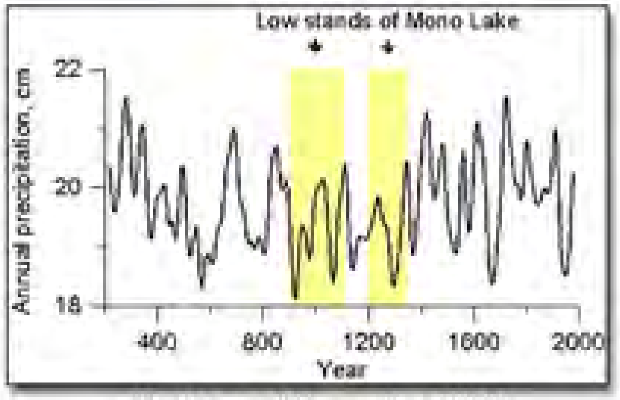


Meko et al. 2007

Tree stand in Alpine Lake, Sierra Nevada

Cook et al. 2004





Hughes and Graumlich 1996

shows trees grew on what are now lakes.

It was so dry for so long back in the megadrought 900 years ago that huge conifers grew year after year, hundreds of years old, or a hundred old, that shows climate or Mother Nature has huge centennial-scale things going on with it as well.

We were talking about 60-year scales and 30-year. Millennial, 100-year scales are going on in the background as well.

Let's go down to Antarctica. This [next page] paper is about a month old as well. 300 years of West Antarctic temperatures here and you see variability down there is huge.

I mean, year to year, it just goes up and down. But there is also this decadal variability you see quite strongly here. And most recently, it's actually come down a bit.

THOMAS ET AL.: WEST ANTAI

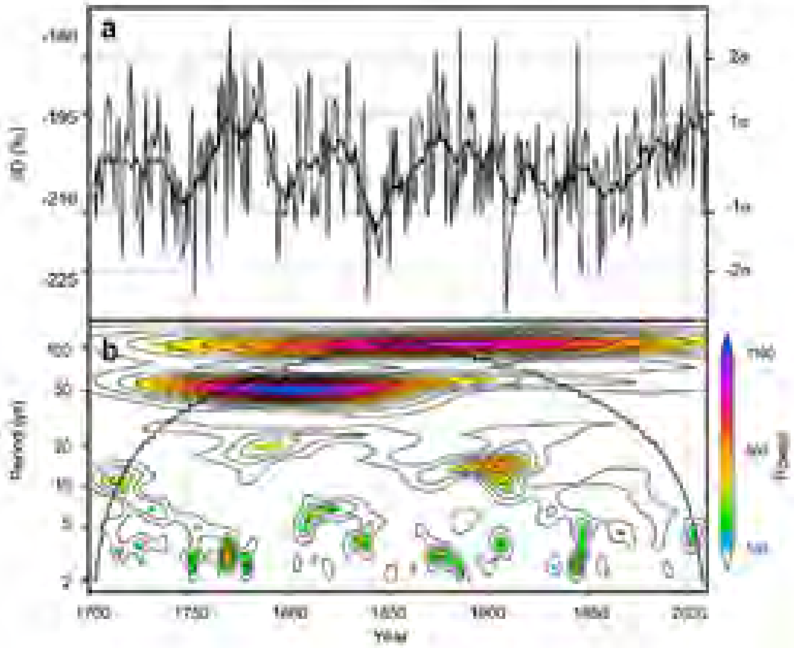


Figure 1. (a) Ferrigno (F10) annual average δD (January–December) and running decadal mean (thick line). Horizontal dashed lines represent one and two standard deviations (σ) above and below the mean. (b) Order 6 Morlet wavelet analysis of the detrended annual average δD. Black line indicates the cone of influence; color shading indicates >95% confidence levels.

… this warming trend is not unique. More dramatic isotopic warming (and cooling) trends occurred in the mid-nineteenth and eighteenth centuries suggesting that at present, ***the effect of anthropogenic climate drivers at this location has not exceeded the natural range of climate variability*** *in the context of the past ~300 years.*

Thomas et al. 2013

VI. What metrics led to selection or rejection of models for Arctic/Antarctic analysis?

The authors just said you can't see any kind of greenhouse gas signal at all in the temperature scale. It's gotten hotter. It's warmed faster in the past than it has now.

And so, as I like to say when I am talking in some venues, in terms of a legal system, it's very hard to convict carbon dioxide of a crime here when you can go back and see the same crime committed when there was no way carbon dioxide could have been the one forcing the crime.

It's a very expectative defendant argument, by the way.

DR. KOONIN: Do we understand who the perpetrators were in the previous incidents?

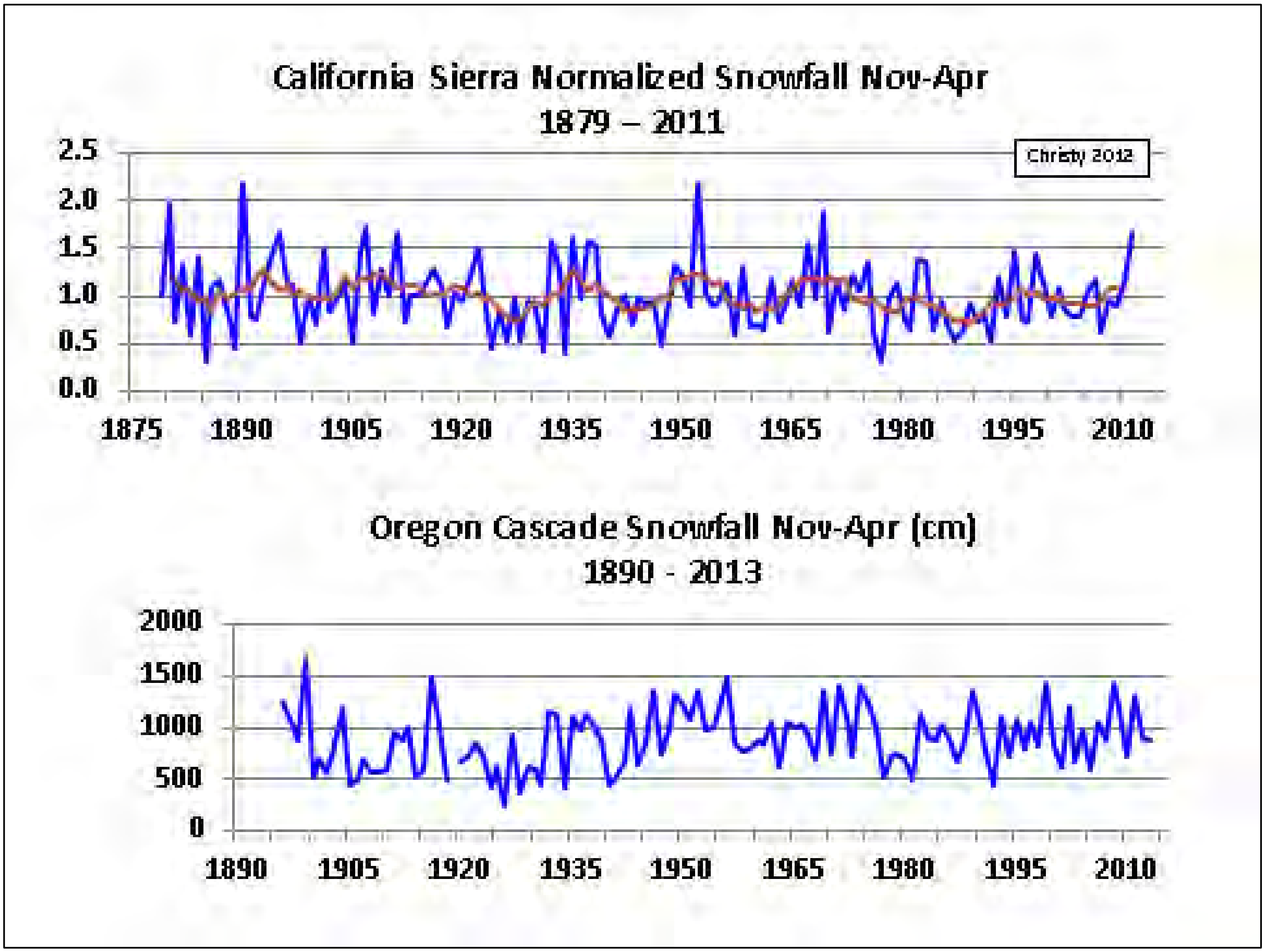
DR. CHRISTY: Natural variability. All I would say is Mother Nature. That's right. I don't know.

But I think -- I am back to my original thing -- what has happened with the climate system? We can say these kinds of things happened in the past where carbon dioxide was not the driver.

Then going down to, say, a smaller scale, a couple of years, I built this [next page] data set from snowfall records when I heard that some predictions were that by now, the snowfall in California should have pretty much gone away at the lower elevations and so on.

And so I built a data set. I actually got the Southern Pacific Railroad records because way back in 1878, they had to know how much snow was there before the sent the trains over the pass. So, they were just meticulous in the records they kept.

So, I was able to build long, 130-year time series which don't show any loss in terms of snow. But you can see that you have four- or five-year periods of huge droughts



California Sierra Normalized Snowfall Nov-Apr
1879 – 2011

Oregon Cascade Snowfall Nov-Apr (cm)
1890 - 2013

that happen, single years followed by other years that are hugely wet.

When this year comes in, this looks like another 1977 right now out in California. It's just very, very dry right now.

Cascades, also same sort of thing that the interannual variability is huge. It is the biggest signal for metrics like this with which the population has to contend.

Now, [next page] the last section I have is what affects the surface temperature? You kind of addressed that in a couple of your questions.   And I don't think -- well, I do think it's a poor proxy for assessing what the thermal content of the climate system is.

Surface temperature is just something measured about a meter and a half off the ground. When I showed you the satellite balloon thing,

What affects the surface temperature?

[Sections II, III in Framing Document]

Surface temperature (GMST), as now measured and reported, is a poor proxy for assessing the thermal content of the climate system

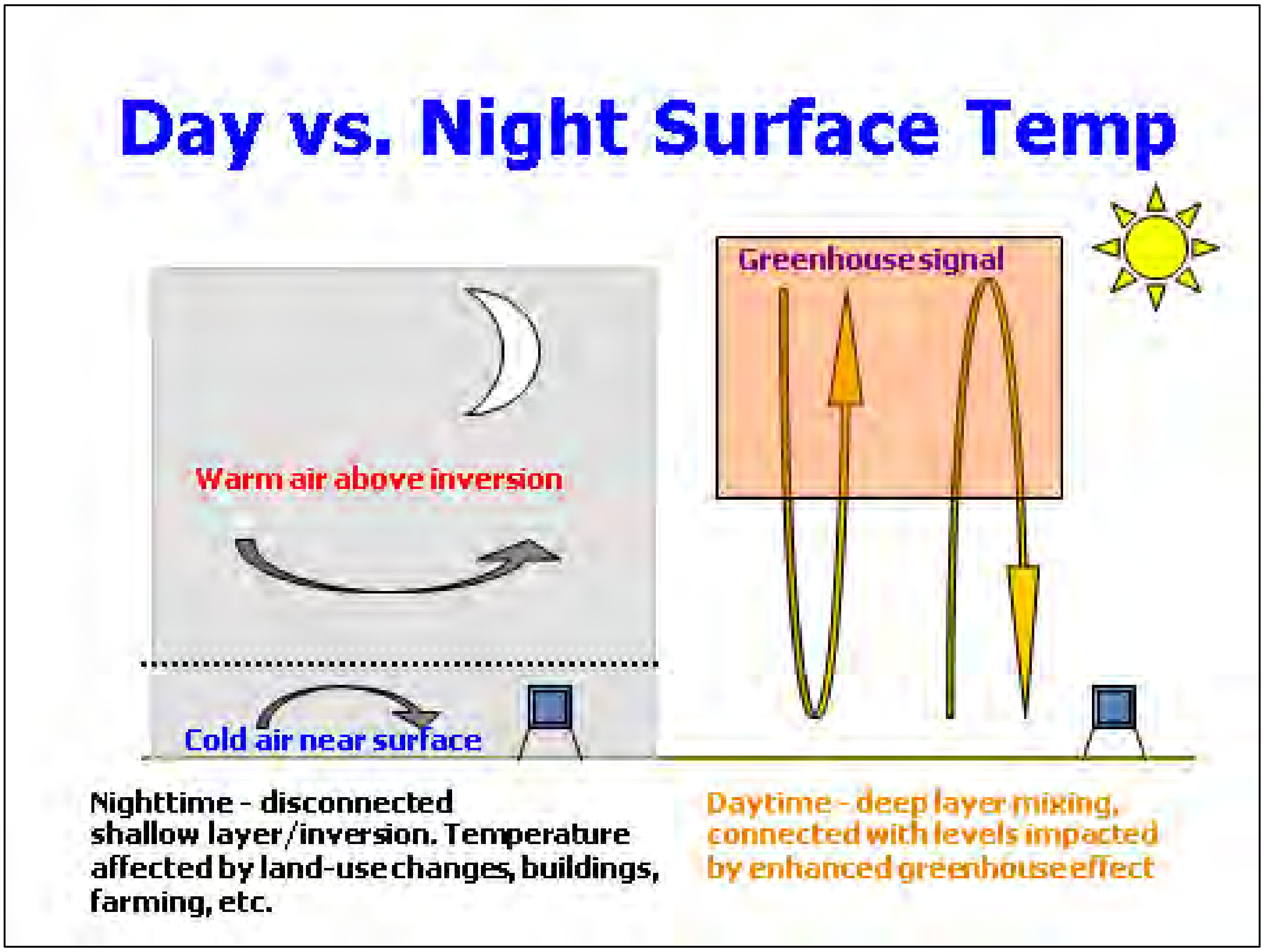
GMST over land is the average of daytime hi (TMax) and nighttime lo (TMin) there were volumes, massive atmosphere was being used in that.

And I will show you why for land, GMST over land is the average of the daytime high which you are all familiar with, and the nighttime low, which around here last night was something like eight degrees or something like that.

And here is the problem. [next page] In a pristine situation, the general rule is that, at night, the boundary layer decouples from the air above. It cools by radiation rapidly. It settles.

And so, you have two types of atmospheres. You have the cold boundary layer. And that's where the thermometer shelter is. It's in the cold boundary layer. The air doesn't change temperature up here much at all.

The daytime, however, when you have heating of the surface and



Day vs. Night Surface Temp

Greenhouse signal

Warm air above inversion

Cold air near surface

Nighttime - disconnected shallow layer/inversion. Temperature affected by land-use changes, buildings, farming, etc.

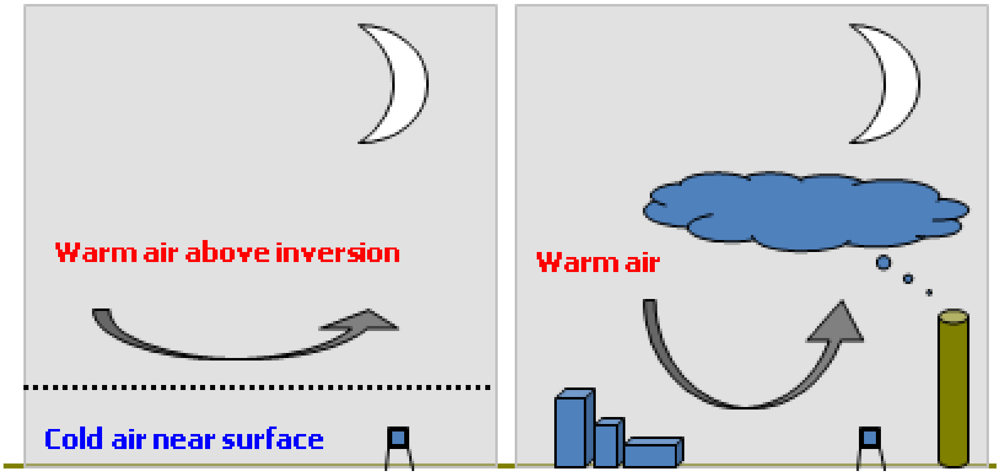
Daytime - deep layer mixing, connected with levels impacted by enhanced greenhouse effect convective turbulence and all that stuff, your daytime temperature really does represent a larger mass of the atmosphere.

And therefore, if you want to measure a surface measurement that kind of gives you a clue with what might be happening upstairs, you would want to measure the daytime maximum.

Now, what happened is that this [next page] situation [left panel] has gone to this situation [right panel] around most of our weather stations. There has been surface development. And it can be anything. If you build buildings, you now have created a different sort of roughness parameter that creates a turbulence that keeps that warm air mixed.

When you launch aerosols into the atmosphere, now the radiative cooling cannot occur because those infrared photons hit this stuff and come right back.

Night Surface Temp



Nighttime - disconnected shallow layer/inversion. But this situation can be sensitive to small changes such as roughness or heat sources.

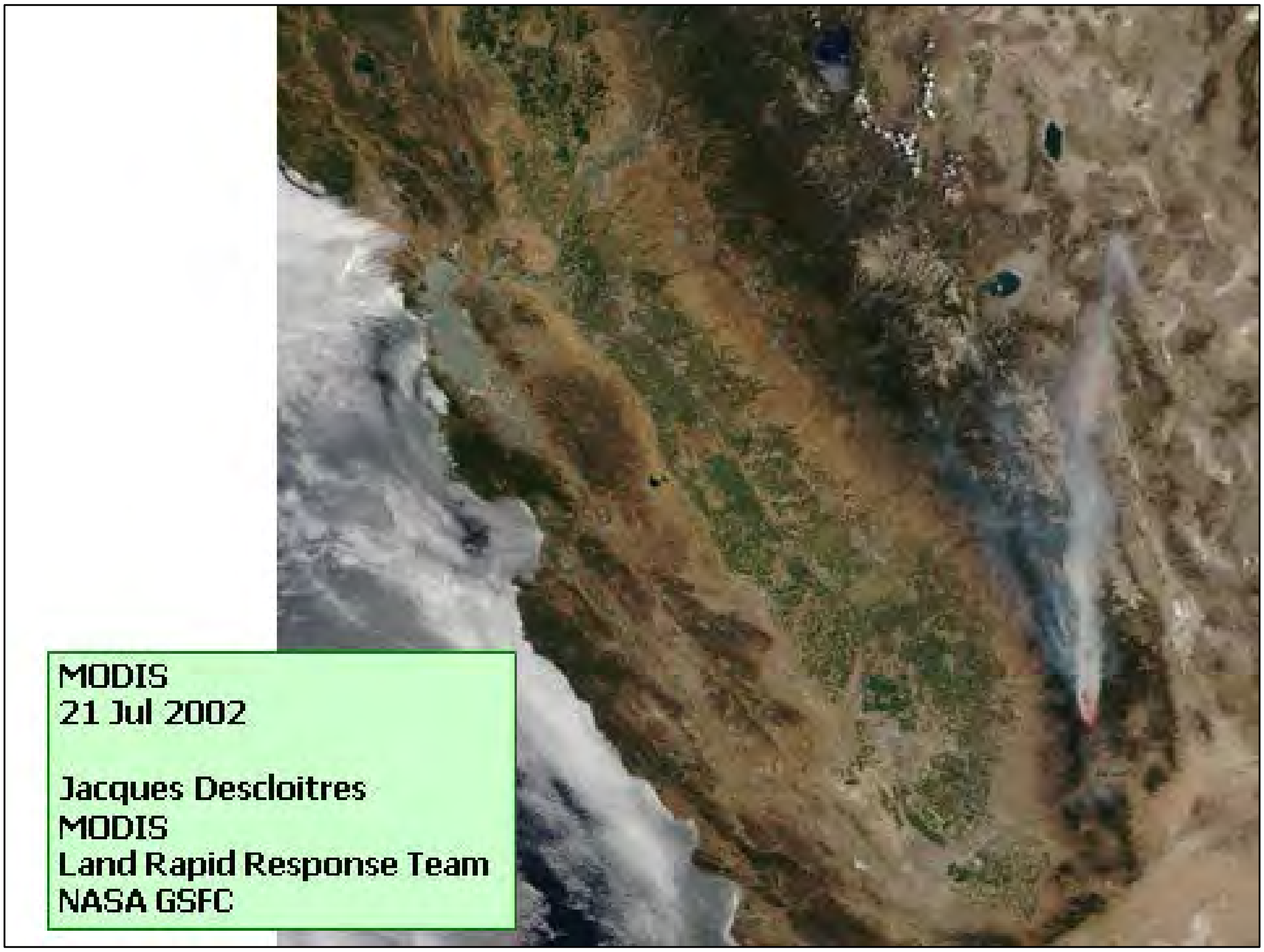
Buildings, heat releasing surfaces, aerosols, greenhouse gases, etc. can disrupt the delicate inversion, mixing warm air downward - affecting TMin.

Tremendous issue in the developing world. And I lived in Kenya for a while. And every night, they would light up the cook fires, the dung fires and so on and you would see that pall of aerosol. Well, that's where all the measurements nearly are being taken. So, that system affects irrigation.

DR. KOONIN: John, I thought that the BEST folks did a pretty thorough study of urban heat island and convinced themselves and me, at least.

DR. CHRISTY: I cannot reproduce their results. I tried and the type of warming they have, and I will show you in Africa, I just can't reproduce it. I don't know what they are doing there.

So, let's just go to California, my home state. This [next page] is San Joaquin Valley. And you can see this green. I don't know. The



MODIS
21 Jul 2002

Jacques Descloitres
MODIS
Land Rapid Response Team
NASA GSFC

colors might not be too great there. This is the Pacific Ocean, Sierra Nevada here, very developed.

This should be that color right there. Before human habitation, it was a desert. It is a desert. I used to chase tumbleweeds when I was a kid growing up here.

So, this surface has been changed significantly. And so, I thought this is a good experiment. I will build a data set here of what's happening in the valley versus what's happening in the foothills right next. That's an experiment-control kind of thing.

This took me a long time, years, because I had to go through something like 1,500 pages, physical pages of information about the data, about the instrument that took temperature in various places.

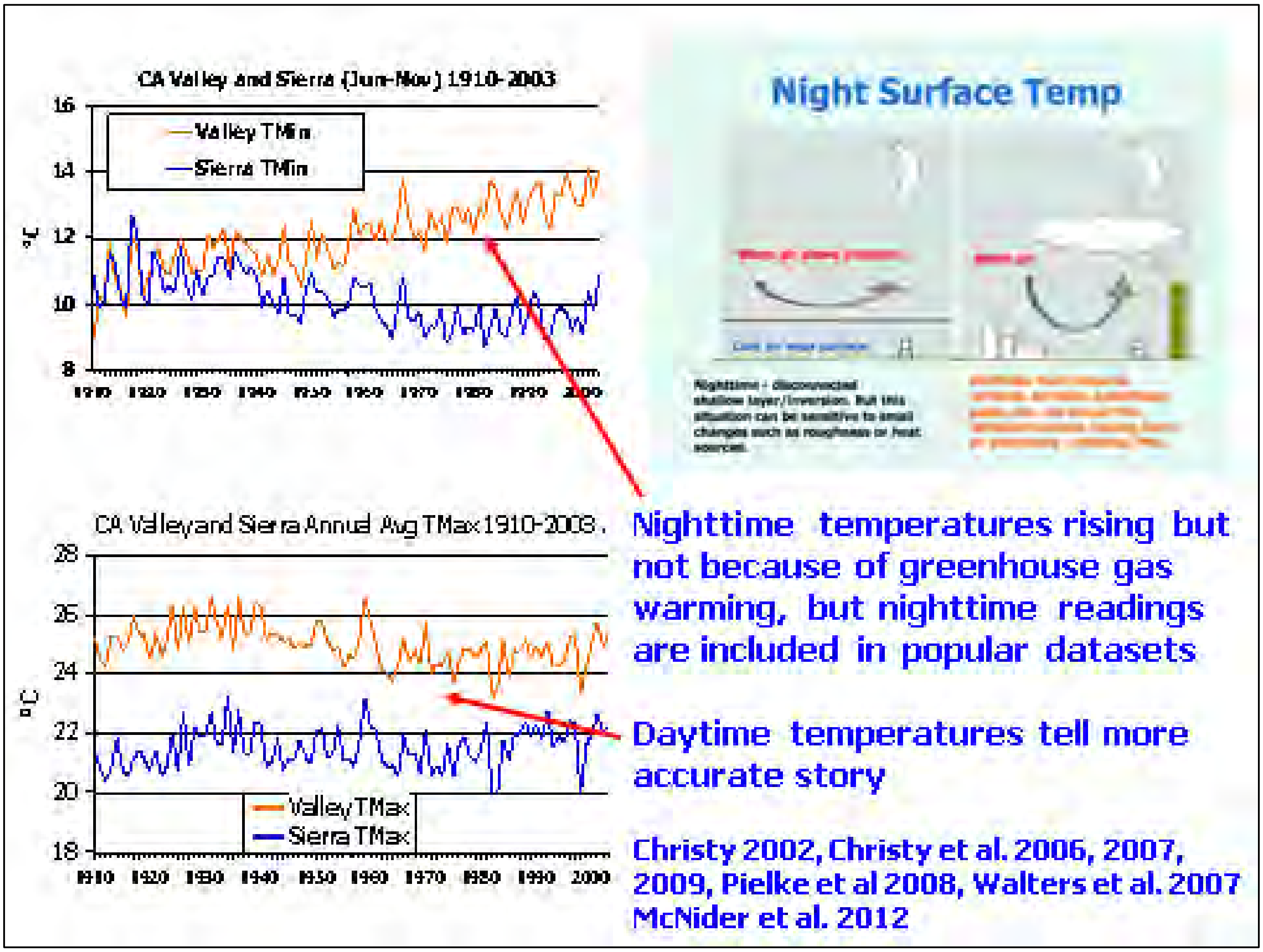
So, I would read about we moved the thermometer shelter 20 feet because a sprinkler was hitting in the afternoon watering time.

Or a great story about a guy in Sequoia National Park up in the mountains. And he said the forest ranger wouldn't let us put up the white screen that reflects sunlight in order to measure the temperature.

We had to paint it dark green to match the forest which, you know, that's a problem for temperature measurement. So, all that stuff I read through and took care of.

So, I built these data sets. And what we find, experiment, this [next page] is the trend in the valley [upper panel], with the actual annual temperatures.   This was in the foothills right next to it. So you see, when you disturb the surface by building buildings, having farmland, I mean, those crops are green and they are wet.

And so, that sun is just absorbed all day long because there



CA Valley and Sierra (Jun-Nov) 1910-2003

CA Valley and Sierra Annual Avg TMax 1910-2003

Night Surface Temp

Nighttime - disconnected shallow layer/inversion. But this situation can be sensitive to small changes such as roughness or heat sources.

Nighttime temperatures rising but not because of greenhouse gas warming, but nighttime readings are included in popular datasets

Daytime temperatures tell more accurate story

Christy 2002, Christy et al. 2006, 2007, 2009, Pielke et al 2008, Walters et al. 2007 McNider et al. 2012

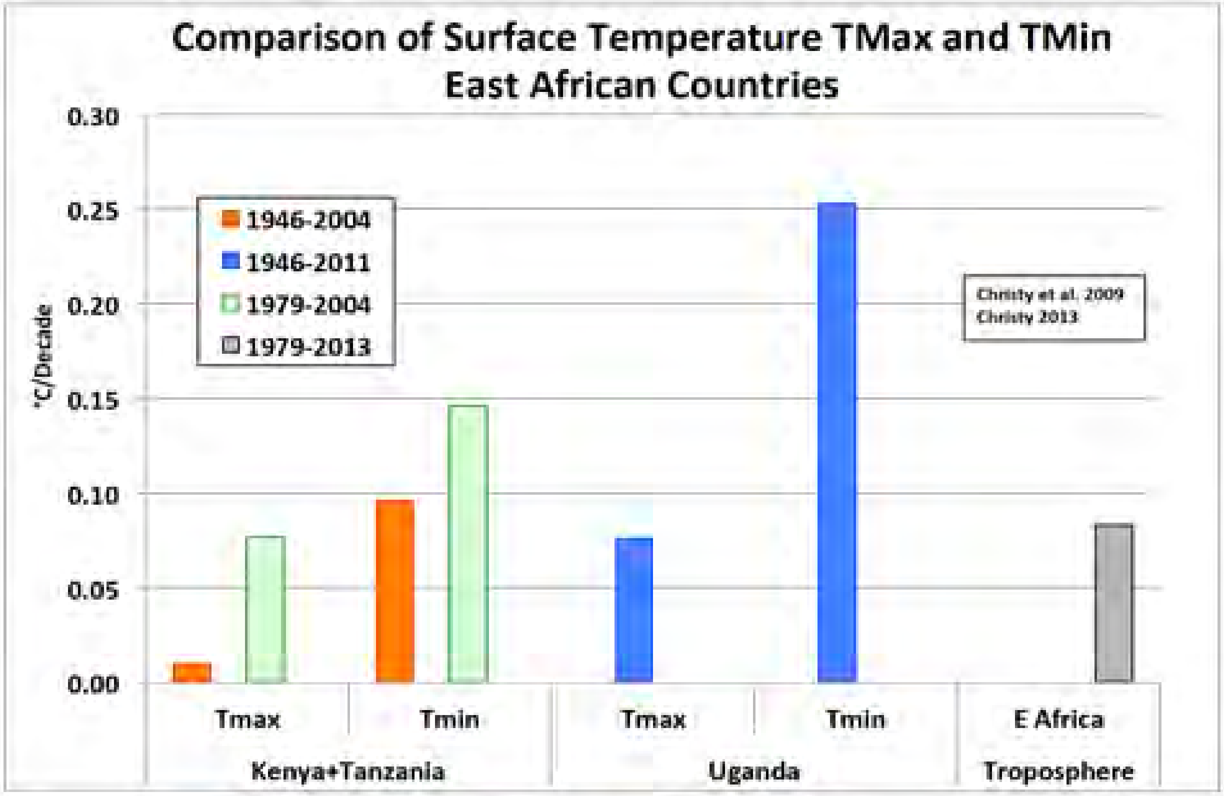
is no clouds during the summer there, especially. But it warms up the night and prevents this boundary layer from forming like it should.

Well, I have another experimented control. I can look at the daytime maximum. And this [lower panel]is what I found, that you see the same temperature variations there, in fact.

And correlation is very high because the scale, the large-scale effects like a hot summer and so on affected both equally. So, this is one of the reasons I don't like to use GMST because of the nighttime warming there.

I built data sets for Uganda, Kenya and Tanzania. This one [next page] was just published three or four months ago. Again, I went through thousands of pages of old documents to get the numbers necessary to produce this.

They have different time frames



Compare like colors for TMax and TMin trend differences
GMST is not the best metric for GHG response detection

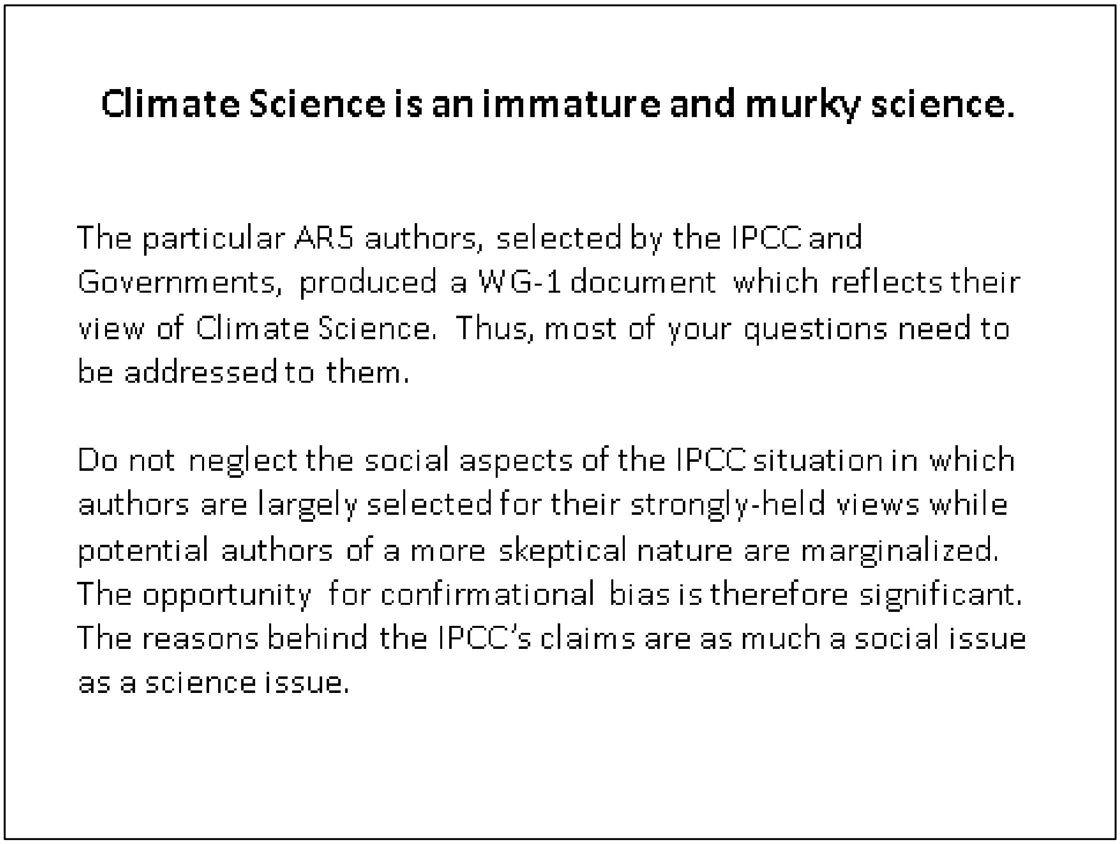
on the trends because I published them at different times, but so just compare the like colors.

And so, on the like colors, you see the nighttime warmer than the daytime, nighttime warmer than the daytime, nighttime warmer than the daytime.

And so, there again, I think it's the aerosol effect of all the fires that are burning around stations and so on that causes that.

Well, the point there is something other than greenhouse effect is causing a temperature rise in the common data sets now being used. I like the bulk atmospheric measurement. That's why that picture is there.

So finally, to wrap it up,[next page] to me, and I built my first climate data set 50 years ago, (reading): "Climate science is an immature and murky science."



Climate Science is an immature and murky science.

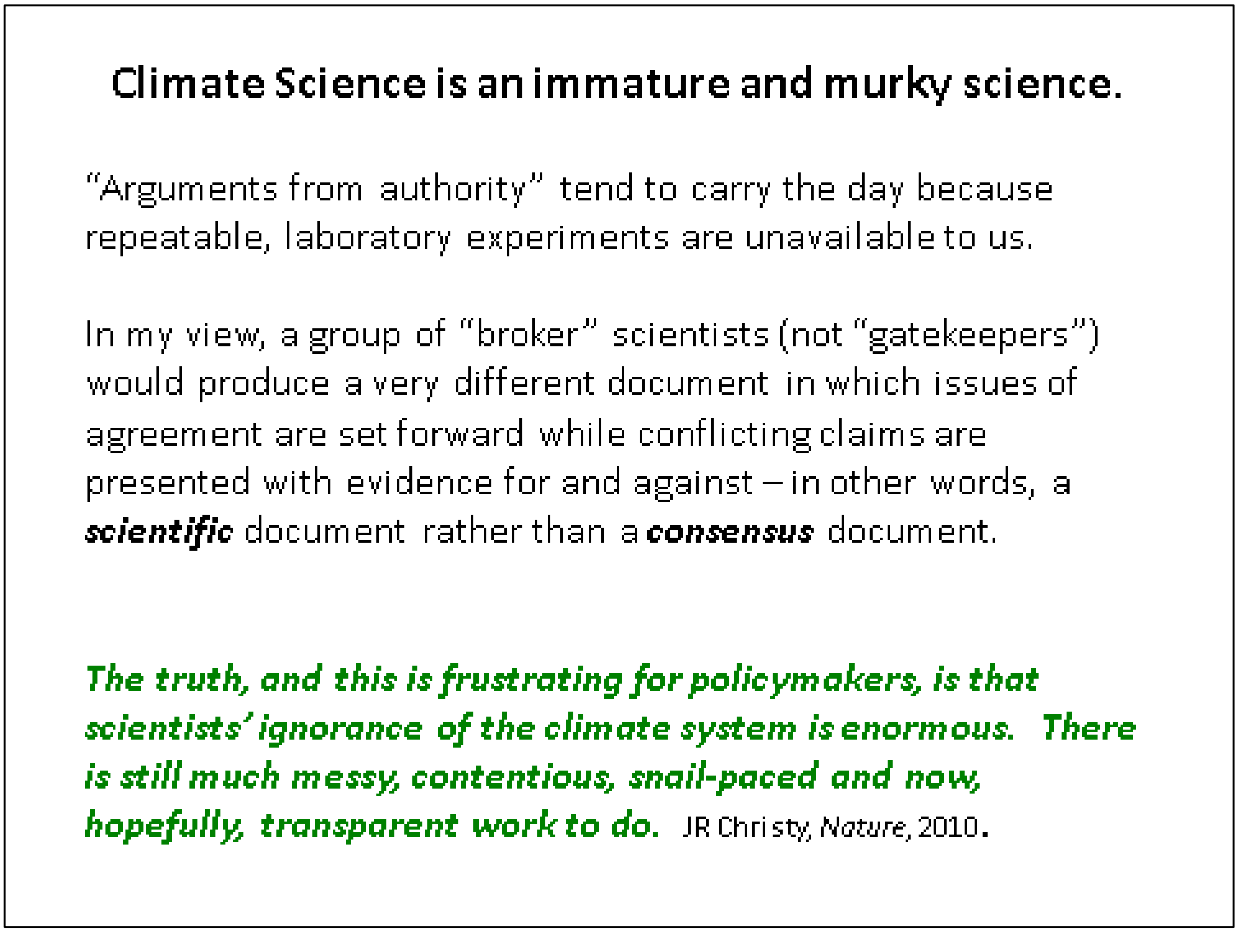
The particular AR5 authors, selected by the IPCC and Governments, produced a WG-1 document which reflects their view of Climate Science. Thus, most of your questions need to be addressed to them.

Do not neglect the social aspects of the IPCC situation in which authors are largely selected for their strongly-held views while potential authors of a more skeptical nature are marginalized. The opportunity for confirmational bias is therefore significant. The reasons behind the IPCC's claims are as much a social issue as a science issue.

The particular AR5 authors selected by the IPCC and governments produced a Working Group 1 document which reflects their view of climate science. Thus, most of your questions really need to be addressed to them that you had about the IPCC.

(Reading): "But do not neglect the social aspects of the IPCC situation in which authors are largely selected for their strongly-held views while potential authors of a more skeptical nature are marginalized. Therefore, the opportunity for conformational bias is therefore significant."

The reasons behind the IPCC's claims are as much a social issue as a science issue. And [next page] like I said early today, arguments from authority unfortunately in our science tend to carry the day in a lot of places because laboratory experiments are just not available to us.



Climate Science is an immature and murky science.

"Arguments from authority" tend to carry the day because repeatable, laboratory experiments are unavailable to us.

In my view, a group of "broker" scientists (not "gatekeepers") would produce a very different document in which issues of agreement are set forward while conflicting claims are presented with evidence for and against – in other words, a ***scientific*** document rather than a ***consensus*** document.

The truth, and this is frustrating for policymakers, is that scientists' ignorance of the climate system is enormous. There is still much messy, contentious, snail-paced and now, hopefully, transparent work to do. JR Christy, *Nature*, 2010.

So in my view, "A group of broker scientists, not gatekeepers, would produce a very different documents in which issues of agreement are set forward while conflicting claims are presented with evidence for and against," in other words, a scientific document rather than a consensus document.

And in a sense, you are going to be, if I understand the APS charge to you, you are going to be the brokers of what this is. You don't have a dog in the fight or anything like that.

You want to, as best you can, understand what the evidence says and what you can be sure about and what you cannot be sure about. Unfortunately, I don't see that in the IPCC because of the author selection process and the point of which consensus has to be driven home for.

Well, in "Nature," they asked me to write a little op-ed one time and I closed it with this statement.

(Reading): "The truth, and this is frustrating for policymakers, is that the sciences' ignorance of the climate system is enormous. There is much messy, contention and snail-paced and now, hopefully" -- that was a prayer at the time -- "transparent work to do."

But it didn't quite pan out that way. So with that, I am through.

DR. KOONIN: Thank you. Subcommittee?

Phil?

MR. COYLE: Do you have a view about what we, the United States, should do differently or additionally than what we are doing now given your observations? Do you have a view?

DR. CHRISTY: That's a science question, right?

MR. COYLE: Yes, yes.

DR. KOONIN: We are still on science. We can get to the social and political aspects of that in the panel, but yes, science now, please.

DR. CHRISTY: My view, because I have seen the problem, is the degradation of the observational network. I mean, if we can't understand what is happening with the climate system, how are we ever going to figure out why it's happening?

I might have to close down a climate reference network here this year because of funding cuts. So, that would be the main thing I would say is the observational network, especially at the surface and the balloon network, would be one thing I would like to see.

DR. LINDZEN: The balloon network has degraded over the years steadily, hasn't improved as a result of interest.

DR. KOONIN: Although you could argue that the satellite observations --

DR. LINDZEN: No. For instance, for weather forecasting, resolution is very important for both vertical and horizontal. The satellites have no vertical resolution.

DR. KOONIN: Right, good, good.

DR. CHRISTY: They are good at bulk numbers. But like Dick said, they are not going to tell the difference between 10,000 feet and 15,000 feet.

DR. KOONIN: Scott?

DR. KEMP: Do your mid-troposphere data match ocean data?

DR. CHRISTY: They are pretty close in terms of the 35-year trend.

DR. KEMP: The trends, do they follow?

DR. CHRISTY: Yes, they do,

yes. They are pretty close, yes.

DR. KOONIN: Isaac?

DR. HELD: Just on that point, I think it is important, and there is some literature on this, but there is a lot of work in progress that, if you take these models which look so bad in John's pictures and you impose the observed ocean surface temperatures as a boundary condition on those models, the fit is much better.

We argue about whether it's completely consistent. There still may be some discrepancy, but it's not like -- the problem is more in whether the ocean hasn't warmed as much as the models, their surface temperature, not so much the vertical structure.

DR. CHRISTY: I would say yes and no on that. Yes, if you took all these models and shoved them back to where they had this as the bottom boundary condition (indicating) -- I have actually written a paper on this -- they still go warmer, but not as much.

DR. HELD: Right. So, we have some work that is coming out that gives the opposite answer to that question. We do very well with our model at simulating the RSS temperature trends.

DR. KOONIN: Bob?

DR. ROSNER: I may be getting inserted right into the middle of the argument here. But I was struck by your comment that if you focus on what you would expect most of the physics to be, which are where most of the mass is, that the models do so much worse.

So, I am just curious whether or not there is an understanding of why that is. What is it that the models are missing? Is there even a hint of an understanding what the issue is?

DR. CHRISTY: I imagine there is, but it probably varies from model to model.

DR. KOONIN: Ben, Bill?

DR. SANTER: Let me tackle that one. I think one of the issues is this forcing uncertainty.

Again, as I mentioned this morning, all of the models in CMIP5 that did not have interactive stratospheric ozone chemistry specified historical changes in ozone, and they used something called the Chioni, et al., database to do that.

That has subsequently been compared, that database of ozone loss with more recent estimates. Susan Solomon's group has done this.

And what it shows is that there are profound differences in the tropics, just where John is looking, extending down as far as 200 hectopascals.

Our best estimate from these kind of comparisons is that Chioni, et al., underestimated the observed ozone loss over the satellite era, which certainly factors into some of these differences that John is seeing here.

So, my problem with this kind of comparison is that it presents only one explanation for the model versus observed differences-sensitivity error.

I don't think one can make a single interpretation of those discrepancies when there is a priori evidence that we got some of the forcings wrong, particularly in these key regions.

Now, what people are doing now is using improved ozone data sets to rerun simulations. And I would point out that the whole problem with developing ozone data sets is you are fitting to sparse observational data. You have some rocket sondes. You have some SAGE measurements.

And people fit statistical models with some volcano terms, some anthropogenic terms, some QBO terms. And they do it differently and they get different results.

And that is part of the reason for these discrepancies that John is showing here, not just model sensitivity error.

DR. CHRISTY: I would just like to say I did not say everything goes on sensitivity here. In fact, I just answered your question that I did not.

DR. SANTER: John, I think in your Congressional testimony, all you discuss is model sensitivity error. In your testimony from a couple of years ago, you said this shows that models are two to three times too sensitive to anthropogenic greenhouse gas increases.

I don't think one can make that kind of inference from this comparison.

DR. KOONIN: The related question is, if you sought to adjust the forcing, are the adjustments plausibly within the uncertainties that you would need in order to bring things up?

DR. SANTER: Well, again, this is why I said earlier I think in tandem with the exploration of parameter uncertainty, what we need to do is not just have this ensemble of opportunity, but have individual modeling groups look a little bit more carefully at the sensitivity of stuff we really care about like the vertical structure of atmospheric temperature change to plausible uncertainties in some of these key external forces.

DR. KOONIN: Bill?

DR. COLLINS: I did want to revisit the issue of global mean surface temperature just briefly. So, John was presenting some analysis with respect to heat island effects, as it's known.

DR. CHRISTY: It's much more than heat island.

DR. COLLINS: Right.

DR. CHRISTY: It's irrigation, everything.

DR. COLLINS: Something that is at issue, of course, one can also look at this issue over the open ocean, including the remote open ocean.

And I would just like to think this sort of partitioning has been done. And I think we can reasonably assert that this is the issue of -- that is probably not a major player in the Southern Pacific Ocean. And the temperature trends do look quite similar.

DR. CHRISTY: Between what and what?

DR. COLLINS: Between the land and the ocean. So, this issue of whether or not the global mean surface temperature record that we are looking is biased by heat island effects, has been partitioned the data and looked at remote ocean regions and the signal that we are seeing is very coherent and it shows up in very remote ocean regions.

DR. CHRISTY: You do now show land warming more than the ocean, though, right?

DR. COLLINS: No. The temperature transfers are essentially both three quarters of a kelvin.

DR. CHRISTY: What data sets are you talking about?

DR. COLLINS: Pick one of four. So, it doesn't really matter. The IPCC analyzed four different temperature reconstructions. They all came to the exactly the same conclusion.

DR. CHRISTY: I don't agree with that as a data set builder.

DR. KOONIN: We can look.

DR. CHRISTY: You can look.

DR. KOONIN: I have to just say the statement about what fraction of the earth's surface is occupied by people, it seems to me urban heat island is probably not an issue.

DR. CHRISTY: Well, actually, it's where the thermometer is. That's what counts.

DR. KOONIN: Well, you don't use ten thermometers in the same city. You take one and you know the correlation methods of 600 or 1,000 kilometers or something like that. So, you sample them at some sensible intervals.

DR. CHRISTY: It's very clear that the land has shown more warming in the data set. The land has shown more warming than the ocean.

DR. KOONIN: And I thought that that is physics. I could be wrong.

DR. CHRISTY: It could be, but if you take the data sets that we built and others that just have the maximum temperature, it comes out better in comparison with the upper air.

DR. KOONIN: As long as you are on T-max, BEST has this funny compression of the diurnal range and then it turns around about a decade ago. Maybe it's a little more; I can't remember. Bill is smiling. He probably remembers that.

DR. COLLINS: No, I am smiling for other reasons involving BEST, but let's not go there.

DR. CHRISTY: You know, I looked at the African data from there, and I probably had three to four times as much data stations as BEST had. And I cannot reproduce.

They have a very large warming in the last 30 years in the daytime of the maximum temperature there. And it's just not there in the data. I don't know where it's coming from.

DR. KOONIN: I see, okay, good. Scott?

DR. KEMP: I guess, there is a question of, if these issues, persistent issues with model forcings and so on remain an issue today, what is the alternative approach of doing a regression and not having the underlying physics in the model?

Is that an appropriate way to proceed in lieu of agreement on the models?

DR. SANTER: I think one of the issues there, again, is this co-linearity that in the real world, at least, you have co-linearity between ENSO, volcanos, some solar terms.

So, if you just plug everything

into a multiple regression framework in the observations or in the model and you are looking at short periods of record, it's very difficult to do an unambiguous separation of these individual terms.

One thing I would say about two other points that John raised here is one on the claim of cessation of lower stratospheric cooling. I think one of the issues there is, indeed, this signal from early 21st-century volcanic eruptions.

So, that increase of four to seven percent per year in background stratospheric aerosol optical depth is warming the lower troposphere in the observations. How much a contribution that is, I don't know. But clearly that is part of it.

The other thing relates to the R squares that you showed initially, John. They are very impressive, but I would contend that they are largely dictated by that high-amplitude, high-frequency monthly time scale variability, and not necessarily indicative of whether there is really good agreement in terms of the lower frequency changes.

DR. CHRISTY: Did you see that the annual was higher than the monthly?

DR. SANTER: Well, the annual is going to be affected by ENSO time scale variability, so I am not really surprised.

DR. CHRISTY: Not so much in these.

DR. SANTER: Those were tropical?

DR. CHRISTY: No, no, those were from the tropics to Port Barre. Those were all the VIZ stations in the U.S. network.

DR. SANTER: But still, it's primarily that large-amplitude, high-frequency variability that is dictating those large R squares that you are getting there.

It doesn't necessarily tell you all that much about agreement or lack thereof the low-frequency changes.

DR. KOONIN: So, those R squares, just to understand what those are, are correlation coefficients between satellites and balloons measurements? Balloons go up four times a day or something like that or how often?

DR. CHRISTY: Twice a day, usually.

DR. KOONIN: Twice a day? So, it's correlation at the twice-a-day level?

DR. CHRISTY: For monthly average.

DR. KOONIN: Monthly mean?

DR. CHRISTY: And then also annual average. I did both of them to show the shorter and the longer time scale.

DR. KOONIN: So, there is a filtering that is going on there. Other comments? That's good enough.

All right. Isaac, you get the last word. For those of you who haven't noticed, the cheesecake has arrived.

DR. HELD: Thanks for the invitation. Just to introduce myself, I work a lot on climate models, but I think of myself as a physicist. My background is in physics and I am interested in the fundamentals. I want to understand the climate system. That's my motivation.

I think I have a little cold. I may be losing my voice a bit. I will try to stay close to the mic.

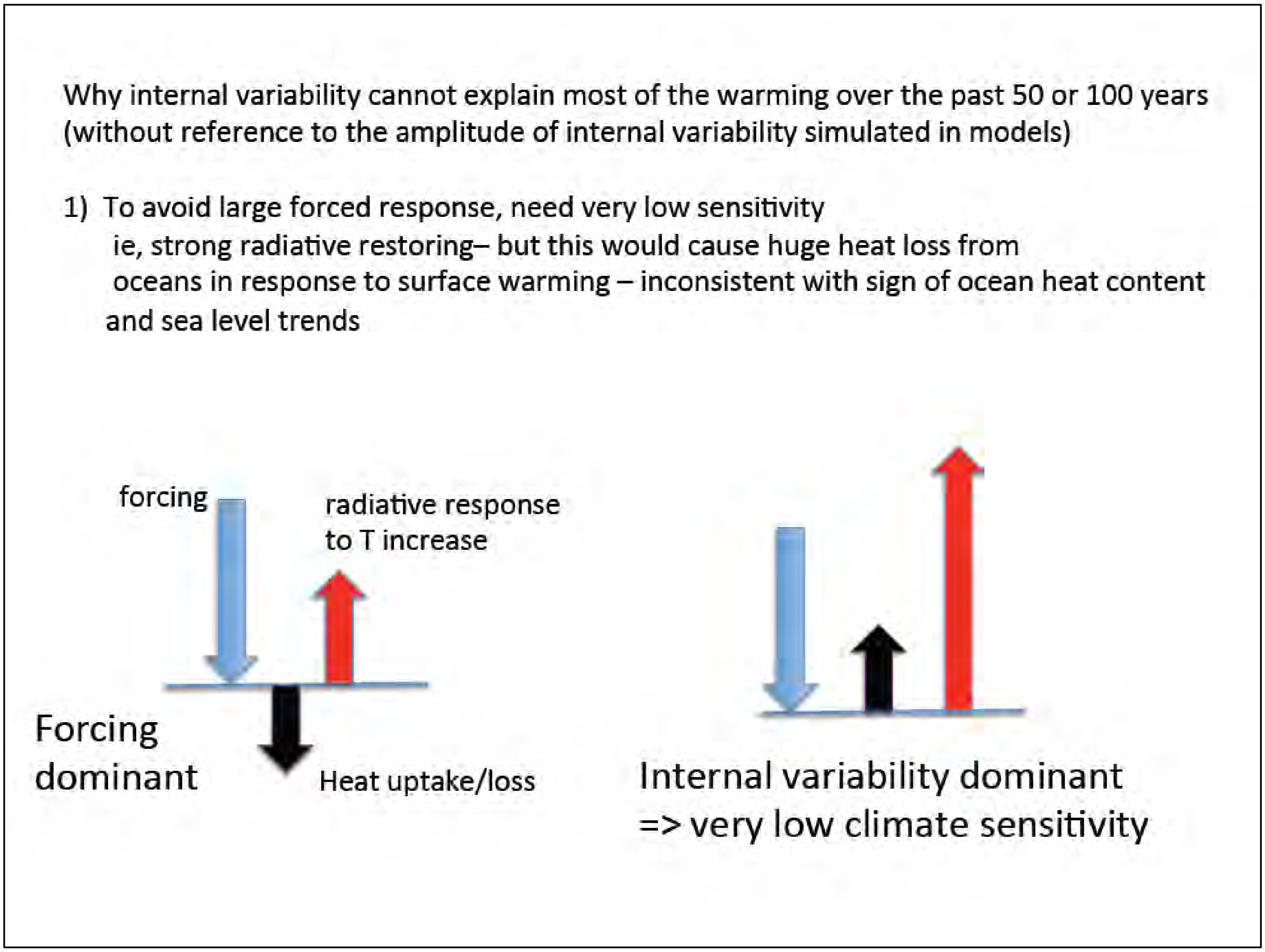
Here is an argument. The way I am presenting this is sort of a crosscut across the questions that you asked rather than focusing question by question.

I think it tries to be responsive. And this [next page] is the way I think about the problem of forced versus internal variability.

And this is sort of, this is really my starting point which, independent of any estimate of internal variability for models, convinces me at a very high level of certainty that the warming we have seen over, say, 50 years or 100 years was mostly forced rather than internal. So, let's just start with that basic fact.

It doesn't say anything about climate sensitivity, per se, directly, because the forcing could be due to other things than greenhouse gases. But let's suppose it's mostly forced.

Then what kind of picture do we have? And this is meant to be the ocean surface here. I will try to move over here. We have some



Why internal variability cannot explain most of the warming over the past 50 or 100 years (without reference to the amplitude of internal variability simulated in models)

1) To avoid large forced response, need very low sensitivity
   ie, strong radiative restoring– but this would cause huge heat loss from oceans in response to surface warming – inconsistent with sign of ocean heat content and sea level trends

Forcing dominant

Internal variability dominant
=> very low climate sensitivity

forcing. And this was, I think, a little bit of the confusion that Steve was referring to.

This is four watts per meter squared or three or whatever, or over the historical period, say two, and this, going into the ocean is 0.6.

The rest of it has been radiated away as a response to the warming. And the amplitude of this is the climate sensitivity, effectively. So, this is a consensus picture that mostly is forced.

So, what would things look like conceivably if I was completely wrong? And let's go to the extreme limit, that it's pretty much all internal variability.

Well, first of all, you are talking immediately about a low-climate sensitivity, much lower than the consensus picture, because otherwise the forced response would be there.

If you are saying it's mostly internal variability, you are talking about a very low-sensitivity system compared to the consensus picture, which means that we are talking about, say observed warming.

So, you would have a huge outgassing of heat from the ocean because that's what you mean by "low-sensitivity model." For the same warming, you get a huge output of energy trying to restore that.

And for the same forcing, I am assuming the forcing estimate is not uncontroversial. You have heat coming out of the ocean. That's the bottom line. We don't see that.

We can argue about whether the heat going into the ocean is accelerating or who knows what. But all the estimates are that the ocean is gaining heat over this time period. We have sea level going back for longer periods.

There is no way I can construct a simple model that would give me heat going into the ocean if the response is basically internal. I am willing to discuss that with panel members.

DR. KOONIN: Okay, keep going.

DR. HELD: But that's one point. The other one is spatial structure. [next page] And these things, again,are implicit in various fingerprinting studies.

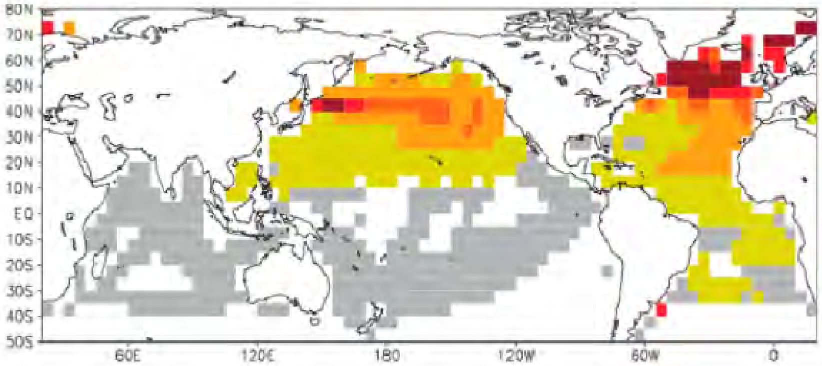
And the first point is sort of implicit in studies where people take simple models and just vary the parameters all over the place and see what they can do.

Where do you expect low-frequency variability to emerge in the coupled climate system? It's going to emerge at high latitudes because, where you have memory on these multidecadal time scales is in the deep ocean.

Why internal variability cannot explain most of the warming over the past 50 or 100 years (without reference to the amplitude of internal variability simulated in models)

1) To avoid large forced response, need very low sensitivity
   ie, strong radiative restoring– but this would cause huge heat loss from oceans in response to surface warming – inconsistent with sign of ocean heat content and sea level trends

2) Simple arguments point to subpolar oceans as the regions where low frequency variability originates (confirmed by models) – so very hard to get the more uniform observed trends. In contrast, spatial pattern of greenhouse gas response pretty good fit to observed trends



DelSole,2012

Maximizing integral
time scale of variability
In CMIP3 models

And where is the deep ocean coupled efficiently to the surface? It's in subpolar regions. That's where the ocean is least stratified.

The tropics are just too strongly stratified for those time scales. You look at where models predict their lowest-frequency variability.

There was a nice paper by Del Sole looking at the models. And he finds a pattern in all the models that have the largest integral time scale or decorrelation time.

They are at high latitudes, especially the northern North Atlantic. This plot doesn't go to the Southern Ocean, but you would see high variability in the Southern Ocean as well.

And that's just the opposite of what you see in reality. In fact, in a forced response, you expect to see an orthogonal pattern, more or less, because those are the regions that are coupled strongly to the deep ocean.

You basically have big heat capacity. So, you have the smallest response to the forcing-- so, it shouldn't be that hard to separate internal variability from forced patterns. They tend to have the opposite structures.

And this doesn't look -- and basically, we don't see subpolar-dominated warming over the ocean. In the subpolar North Atlantic, we have seen very little warming over the century time scale. Over the last 20 years, we have seen quite a bit.

And that's why a couple of people talked about there might have been some contribution to the recent ramp-up, say, of the last 30 years from the North Atlantic variability, which is entirely plausible.

But you don't see a magnified

subpolar Atlantic warming over the 50- to 100-year time scale. So, to combine those two things, without any reference to the magnitude of internal variability in the models, it's pretty inconceivable to me. And we haven't seen it.

I don't think it's a -- it's not a mystery to me that no one has produced a model that gives you something that looks like the warming over the last 50 or 100 years from internal variability.

I just don't think you can do it. I haven’t tried to put a number on it. I don't know if I come up with 95 percent or 90 percent or what. I am not holding my breath.

So, I think there is a point of confusion here. [next page] Does the IPCC actually say that the level, that our confidence has increased from this 90 to 95 percent level? Actually, it doesn't. There is no statement to

Ch. 10/WG1/AR5:
More than half of the observed increase in global mean temperature from 1951 to 2010 is very likely due to the observed anthropogenic increase in greenhouse gases

It is extremely likely that human activities caused more that half of the observed increase in global mean surface temperature from 1951-2010

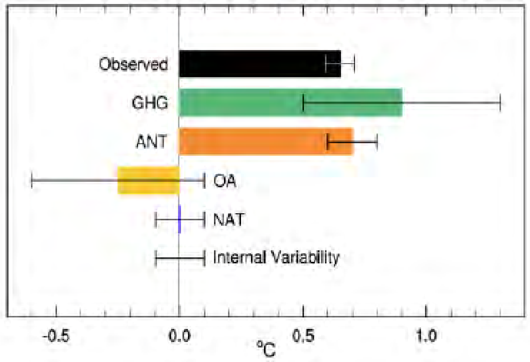


Figure 10.5: Assessed *likely* ranges (whiskers) and their mid-points (bars) for attributable warming trends over the 1951–2010 period due to well-mixed greenhouse gases (GHG), other anthropogenic forings (OA), natural forcings (NAT), combined anthropogenic forcings (ANT), and internal variability. The HadCRUT4 observations are shown in black with the 5–95% uncertainty range due to observational uncertainty in this record (Morice et al., 2012).

that effect, unfortunately.

If you go into chapter 10, you could say this is a communication problem, but I'm not sure I want to talk about communication. We are talking about science. These two statements are different. I didn't label this exactly correctly. This is the AR4 statement.

Very likely that more than half of the 20th-century warming was due to an increase in greenhouse gases. That means well-mixed greenhouse gases.

The statement didn't say "well-mixed," but if you read the text. Now this is the statement in AR5, "extremely likely." That seems like a stronger statement.

DR. KOONIN: Just what is the origin of the two different statements?

DR. HELD: This one is about human activities. This one is about greenhouse gases.

DR. KOONIN: They are both in AR5?

DR HELD: No -- well, yes, they are. I'm sorry. They are both in chapter 10 of AR5. In fact, they are both right next to each other in the summary of chapter 10.

And so, for people who read chapter 10, these are two different statements. And it's discussed in some detail in chapter 10.

DR. CURRY: The issue is what showed up in the summary for policymakers.

DR. LINDZEN: And the press release.

DR. CURRY: And the press release, yes.

DR. KOONIN: That's not science, but it's important.

DR. HELD: I want to stick to the science. I am not saying it's not important.

And so, what is going on here is a statement about, this statement is dividing up the response into two pieces, the human activities and everything else. "Everything else" is basically internal variability and natural forces, volcanos, solar.

This is dividing up into, like, three pieces. There is the greenhouse gas and there is the other forcing, as well as the natural forcing internal variability.

And this is one of the pictures from chapter 10. I think this is kind of a detail. I'm not sure. So, why is it that this statement, which is different than this one, is more popular than this one?

Well, it has to do with this. This is their error bar on the total anthropogenic forcing, which is equivalent to an error bar on internal natural variability, because it's just a two-part decomposition.

DR. KOONIN: Where are the anthropogenic aerosols in that picture?

DR. HELD: Well, it's greenhouse gas plus aerosols will give you this. So, this error bar is a lot smaller than this one.

You have this issue that we have been talking about, compensation between greenhouse and aerosol. And this error bar is basically similar to these.

DR. KOONIN: To the extent that they compensate well, the top statement becomes even more accurate.

DR. HELD: This is a statement about natural variability. This is a statement about greenhouse gases. I don't know if we have enough time to go into exactly why one is stronger than the other.

But I would encourage you to read chapter 10 if you are interested in focusing on this. I don't focus that much on 90 versus 95 percent. To me, I don't get into a tizzy about that sort of thing.

DR. KOONIN: Nobody should, but the media do.

DR. HELD: Yes, they do. Just, I don't.

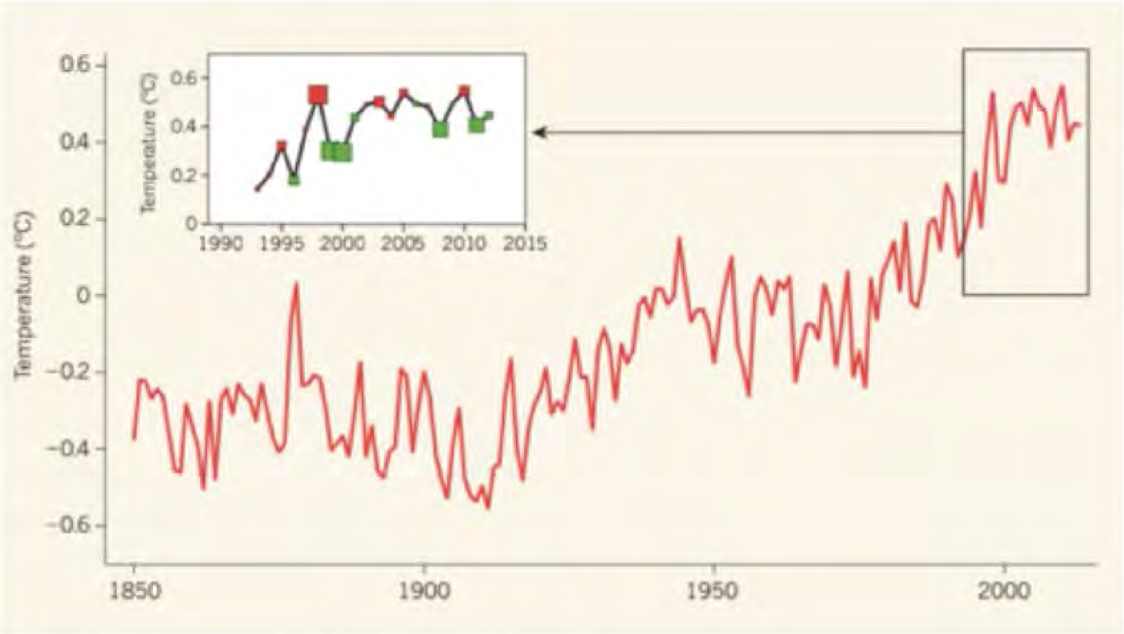
How about the hiatus? I like the word "hiatus." This [next page] is the way I plot it. This is a standard Hadley Center data set.

And this is just a blowup showing El Niño years and La Niña years. The red is El Niño and the green is La Niña by some standard definition. And the magnitude of these boxes is the magnitude of the El Niño or La Niña.

And you can plot this in different ways. Superficially, a lot of the hiatus is, it looks like an extended La Niña-like period.

And this also gives you a little flavor of what we are talking about here. We are not talking about

Hiatus resembles extended La Nina-like period



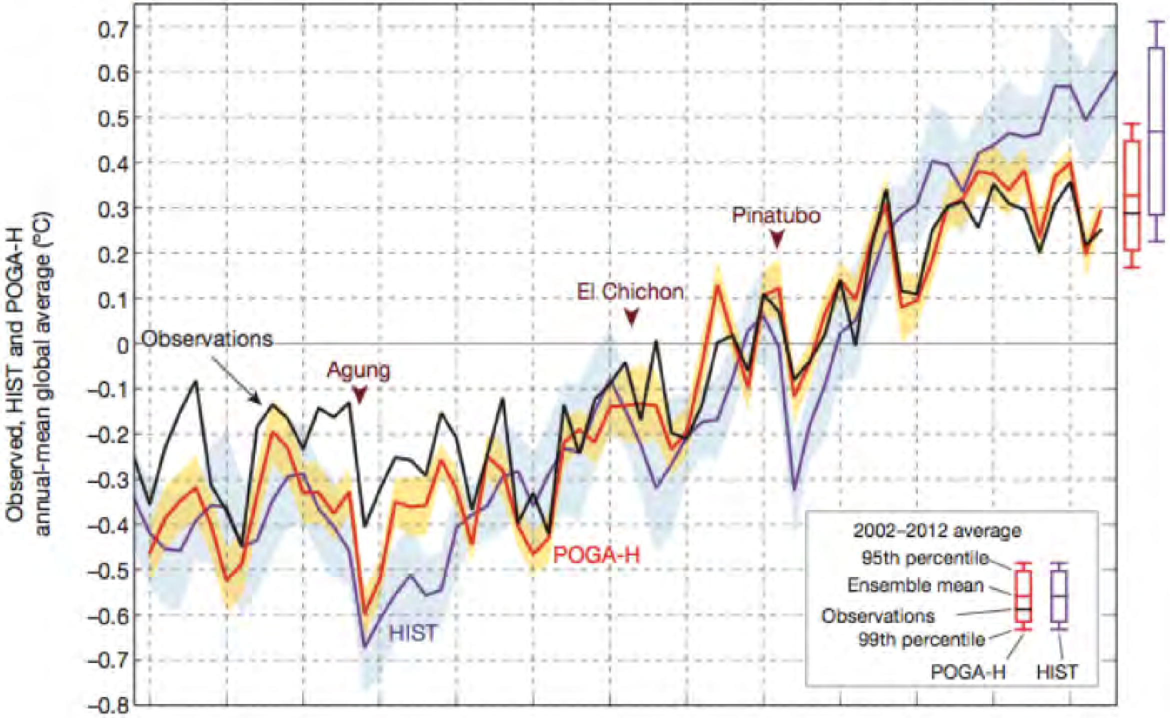
something that is the order of magnitude of the 20th-century warming. I am not saying it's not important.

I like this paper [next page] by Kosaka and Xie that came out in "Nature" in the past year. I would encourage people to read it. They took a model. It happened to be our model. And this model warms too much.

This is global mean surface temperature. And this is the blue line. If you just let the model run freely, it warms too much over this hiatus period. And you do ten runs. None of them look very much like the hiatus, unfortunately, for better or worse, I guess for worse.

You can think of it in a couple of different ways. I think it was motivated. Let's let the model filter out El Niño for us. And we can come back to what the simulation actually shows.

What happens if you constrain eastern equatorial Pacific temperatures to vary as observed in a coupled model?



Kosaka + Xie, Nature 2013 => explanation for hiatus must pass through eastern equatorial Pacific

Let's go into the Eastern Equatorial Pacific in a region which is in the order of, I don't know, 12 percent or something of the global area of the earth, and just specify the ocean temperatures in that region to be those that were observed in that period.

And then you get the observed here in black and the red line is what you get from this constrained model, just constrained in the Equatorial Eastern Pacific.

So, you are forcing the model, if you like, to absorb some heat to keep those temperatures from warming, at least in that little region. And it looks pretty good.

It even improves Pinatubo. It's a way of removing ENSO effects. People say oh, the model is overestimating the response to Pinatubo. If you constrain the Eastern Equatorial Pacific, it looks pretty good.

DR. KOONIN: If you had taken some other region of the ocean, the Atlantic somewhere, would it have done the same thing?

DR. HELD: No, it wouldn't work very well. The tropical Pacific is powerful.

DR. KOONIN: That's where the action is.

DR. HELD: And also it's motivated. This isn't arbitrary. It's motivated by the fact that it looks like the recent past has been La Niña-like. And so, just go in and let the model do it for you. And it looks pretty good.

And you can focus on that we still have some discrepancy. The model is still warming a little bit too much if you go back. That could be due to forcing errors as we have been discussing or sensitivity errors.

But the hiatus period is

captured pretty well. So, not -- what is it? Is it some kind of clumping of La Niña events that happened randomly? Is there something going on in the tropical Eastern Pacific, a signature of a slower mode of the ocean?

If it's not just a random clumping, is there actually something significant going on there that might even be predictable? Or could it even be some forcing that, for whatever reason, is influencing the Eastern Equatorial Pacific temperatures? I don't know.

All I can say from this paper, which, as I said, I like a lot, is the explanation has to go through Eastern Equatorial Pacific. If it is going to be forcing, then the forcing is going to have to change the Equatorial Eastern Pacific, especially.

So anyway, that's just a point of, I don't think this is particularly model-specific.

DR. CHRISTY: What section of the Eastern Tropical Pacific was that?

DR. HELD: I don't have a picture here, but it's a non-trivial part. I don't remember how far it goes west, but I think something like 10 north to 10 south. I just don't remember the east or west.

If you look at, if you take out the forcing from the model, if you take out the time evolution of the forcing -- I didn't bring that picture with me -- and just specify the temperature in that region, you get essentially nothing.
So, most of this is due to forcing. So, I think this is a clue as to what is going on. You have to explain what is happening in the Eastern Pacific to explain the

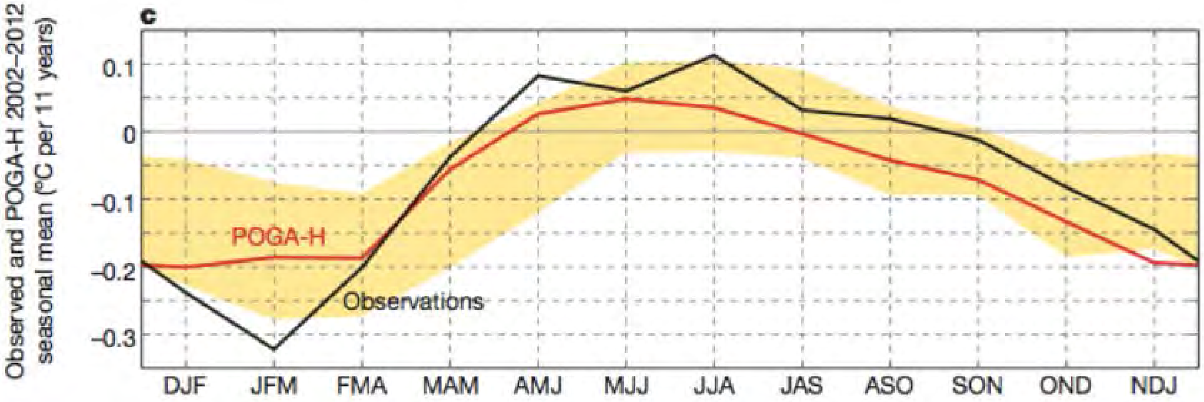
hiatus. That's what it's telling me.

And here [next page] is something else that I found interesting in that paper. Other people have written about this, but it doesn't seem to get focused on. The hiatus is a wintertime phenomenon. If you look in the summer, there is no hiatus. There is still warming.

Global mean temperatures are increasing in northern summer. And this is the seasonal cycle of the observed trend over this hiatus period. And their model is pretty good. It captures that roughly, not perfectly.

And that is a pretty high bar for explanation in terms of forcing. I am not saying it's impossible, but you have to explain that seasonal cycle from volcanos or something. I think it's hard. Volcanos could be a part of it. But I don't think they could do it all, clearly.

Seasonal cycle of hiatus



Kosaka and Xie, Nature 2013

Okay, so that's my hiatus talk. Some of the questions that came through in your background document I thought were a little off, if I can be frank --

DR. KOONIN: That's fine. We are not experts.

DR. HELD: -- in the sense that they don't conform to my picture of how the climate system works. So, I have my null hypotheses. And I have been doing this for over 30 years, so I have developed a lot of hypotheses. Some of them turn out to be wrong.

I don't like this argument from complexity saying oh, it's a chaotic system. There is all sorts -- you can get a nonlinear system to do anything you want. That just doesn't tell me anything.

But whenever I look at the forced response of the climate system, it looks linear to me. And what is the best example we have of forced responses? The seasonal cycle. Seasonal cycles are remarkably linear-looking.

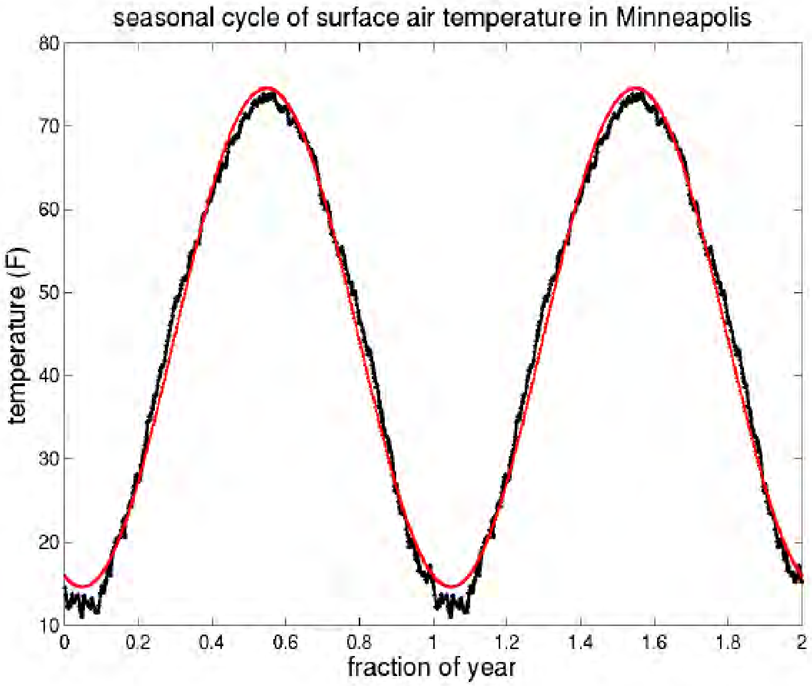
I grew up in Minneapolis which is why I plotted Minneapolis here. [next page] I just repeated it twice for clarity. This is just the seasonal cycle. It's almost perfectly inside the squiggle.

There is an awful lot of nonlinear fluid dynamics and cloud formation stuff going on underneath this. My analogy here is the thermodynamic limit of statistical mechanics.

The smaller response, you seem to worry about the fact that the external forcing is so small, but that just makes it more likely to be linear.

DR. KOONIN: Although, in real thermodynamics, since you have a good separation of scale, there is a

"The argument from complexity"



Analogy is thermodynamic limit that responds linearly to external forces, emerging from complex molecular dynamics small parameter or a big parameter, right? The size of the atoms or the number of atoms or something?

DR. HELD: I am not saying it is as good as thermodynamics, but that's my underlying picture.

One other example of forced response that Dick referred to, we have Milankovitch. We don't have anything really in between -- I mean, we have the sunspots, but that's hard to see, it's so small.

So, we have the seasonal cycle and Milankovitch. Those are both changes in our orbit. And that looks pretty linear, too, at least in the sense that you see the periods of the orbital changes coming out.

DR. KOONIN: If you take a given model, one of the ones in the middle of the pack, and start doing the linear study on one or several of the forces, start cranking up the solar constant or the aerosol loading or CO_2, does it behave in a linear way?

DR. HELD: Yes.

DR. KOONIN: Over the range of what we are talking about?

DR. HELD: A lot of people looked at that. It's very linear.

DR. COLLINS: Yes, it is very linear.

DR. HELD: The whole language, the whole forcing-feedback language we look at is assuming that this linear picture is useful. Otherwise, what is forcing and what is feedback? I don't even know where to start.

DR. COLLINS: At the risk of breaking protocol, may I?

DR. KOONIN: Yes.

DR. COLLINS: You can force the model separately with different forcing agents, look at the separate response, add the response and then add the forcings and compare the total response to the total forcings. That has been done ad nauseam, not a problem.

DR. HELD: The models look pretty linear. The observed seasonal cycle, that looks linear. Even if in the Ice Age times, things look pretty linear. We don't know that much about it.

So, why should I assume that things are, gee, the anthropogenic CO_2 pulse is going to interact in some exotic way with internal modes of variability? Well, it's conceivable. But I am not convinced. I don't think that is particularly relevant.

DR. KOONIN: But to come back to my earlier hobbyhorse, that means that the sensitivity you determined to, let's say, CO_2 from the last 30 years, you should use in extrapolating out of next century?

DR. HELD: Yes, I don't think

there is much evidence that there is much secular variation in sensitivity.

DR. LINDZEN: But I think this is important. For instance, when I presented the simple analysis, I was assuming it was all due to anthropogenic. Sensitivity is a separate question. And I think in conflating the two issues, we are confusing things.

DR. HELD: I was trying to separate them here. I don't think there is so much a collection of sensitivity as you are saying.

I just think if you want internal variability to be important, you have to be in a low-sensitivity model by definition. And then you are going to have the heat going in the wrong direction. It's just so basic to me, I don't see why we talk about it.

DR. KOONIN: Some of us haven't

spent 30 years.

DR. HELD: You asked about our models getting better. I actually don’t think this is a big issue for this group.

It's a hard problem and if you go into chapter 9, there is a frequently asked question. All the chapters have these things, and one of them is are models getting better?

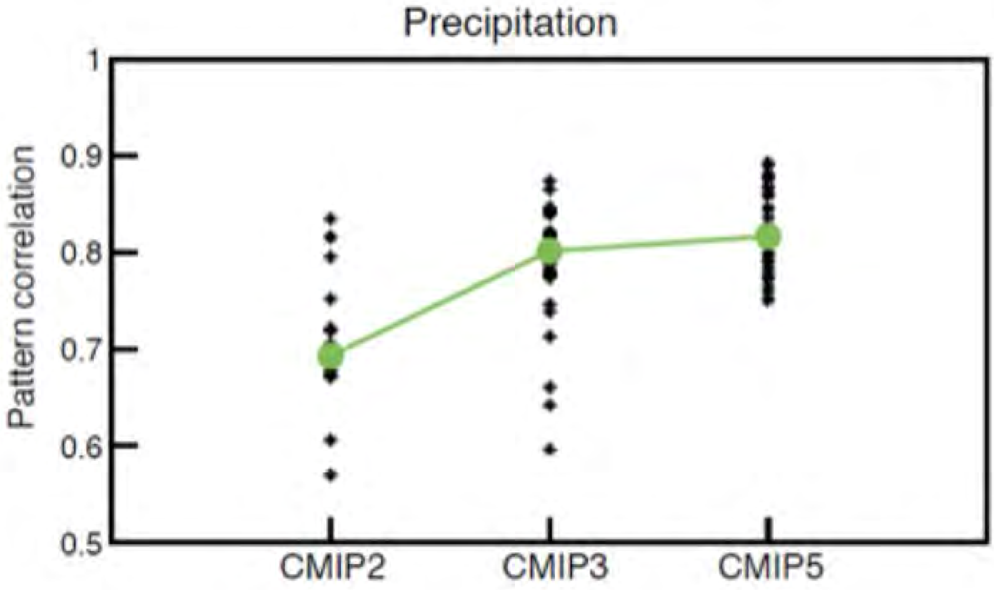
And this is a figure [next page] I took from the answer to that frequently asked question showing precipitation correlation with observations, CMIP2, CMIP3, CMIP5.

DR. KOONIN: As we have discussed, the correlation coefficients depend on what frequency band you are looking in?

DR. HELD: This is a spatial correlation, nothing to do with time, just space.

One thing that has happened is that poor models disappear. You say how do models get

Ch. 9, WG1, AR5: FAQ “Are models getting better?”



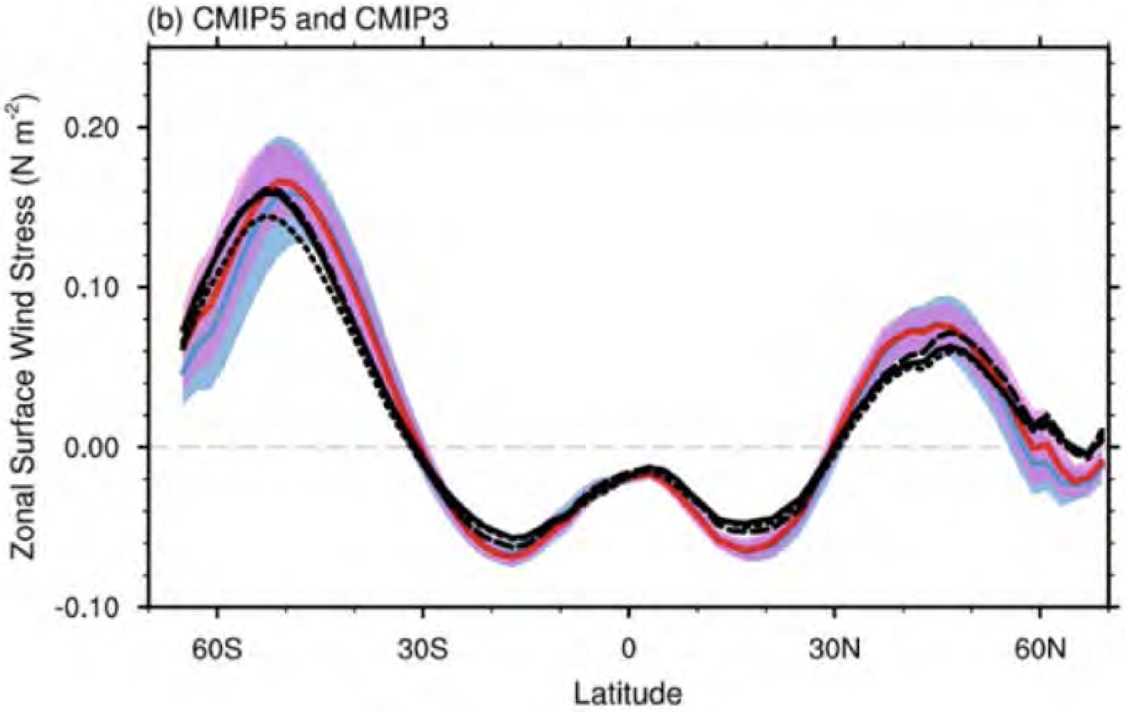
selected? People get embarrassed at zeroth order. People put models in. They look really bad when they put them into these databases. They just drop out and you end up with better models just by public relations. It's harder to say that the best models are getting better.

DR. LINDZEN: I don't know what Isaac's experience is. I know in Paris at LMD if they send something in to CMIP that's too far out, they get a telephone call, "How come it's so far out? Can't you do something about it?"

DR. HELD: Well, that's a quality-control issue. The kind of thing I am interested in is in this next figure, [next page] this is the kind of detail. This is directly out of chapter 9.

And this is pretty hard to see, but the blue here is the range of the CMIP3 models, and the purple is the range of the CMIP5 models. This is

Zonally averaged wind stress on oceans – Ch 9 AR5 WG1



the zonal-average wind stress on the ocean. That turns out to be a pretty interesting quantity.

DR. KOONIN: Is that anomaly relative to some absolute?

DR. HELD: This is just the climatology. This is the absolute.

DR. KOONIN: Which is the data?

DR. HELD: The two estimates are these black lines, two different estimates.

DR. KOONIN: Good.

DR. HELD: And one thing that happens, a lot of older models have the Surface Westerlies, the roaring 40s too far equatorward.
It turns out to be important
for stratosphere and troposphere
coupling in response to the ozone hole.

You get a rather different wind response to the ozone hole just because the ozone hole is happening in the Antarctic; not too surprising. That turns out to be important.

What we have left kind of, and this turns out to be important as well, is that the models are giving us too-strong stresses on average, which you can sort of vaguely see.

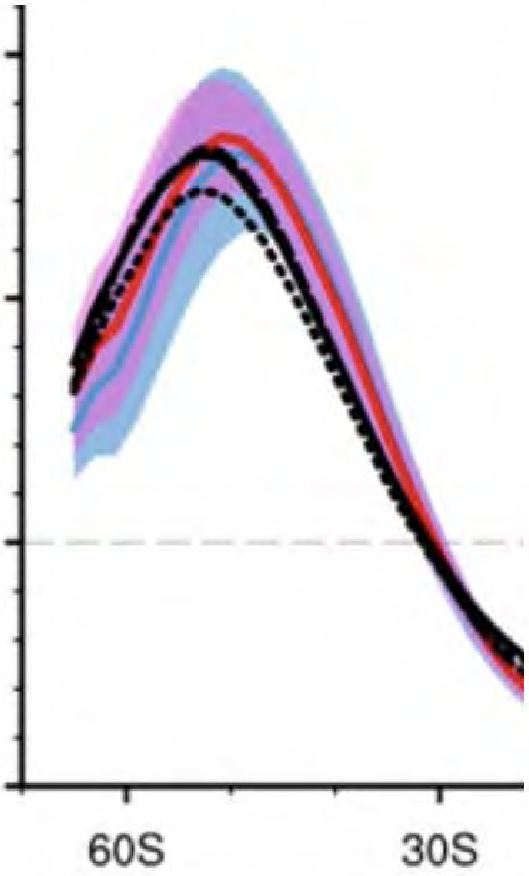
Why is that? I think that is responsible for some other biases we have. It's a signature of it. I think it's a mystery.

As the models have gotten better, they sort of converge on the value of the wind stress. That means the model is transporting too much angular momentum horizontally on average.

So, that's the kind of thing we focus on. It's at higher order. You can see the models are getting a little better when you look at these kinds of things-- here is another. [next page] I just blew up the location of the ozone hole.

I think this is one of the great success stories of climate

Zonally averaged wind stress on oceans – Ch 9 AR5 WG1



Important for response
of troposphere to
Ozone Hole

modeling in the last ten years. It doesn't have too much to do with greenhouse warming.

But there are observational studies that suggest, and pretty strongly, that when the ozone hole developed, the Surface Westerlies moved poleward. But, how are you going to prove something like that?

Current-generation models do that very robustly. Every single model, when you put in the ozone hole, differ almost by a factor of two in how much the Westerlies move.

But that is arguably the biggest circulation change we have seen, because the ozone hole developed so fast, it just had a big effect, a little easier to discern. We think the effect on the circulation of the ozone hole and greenhouse gases are comparable.

We talked about how do you decide what metrics to use when you are trying to rank models or weight a model? That is a huge question. There is a lot of question about that. The problem is there are literally thousands of things you can compare. How do you choose between all those?

And I think a good argument is, to start with, to use those things that are relevant for what you are trying to predict. How do you know what metric is relevant for what you are trying to predict?

Well, you can start just by looking at your ensemble of models and say, within that ensemble, what distinguishes between, say, a model that dries the Sahel in the future and those that don’t. You can look at something that is observable in the present-day simulation or 20th-century historical simulation that distinguishes between those models that dry the Sahel and those that make the Sahel wetter. And that is essentially defining a metric, I think, and then you use the observations.

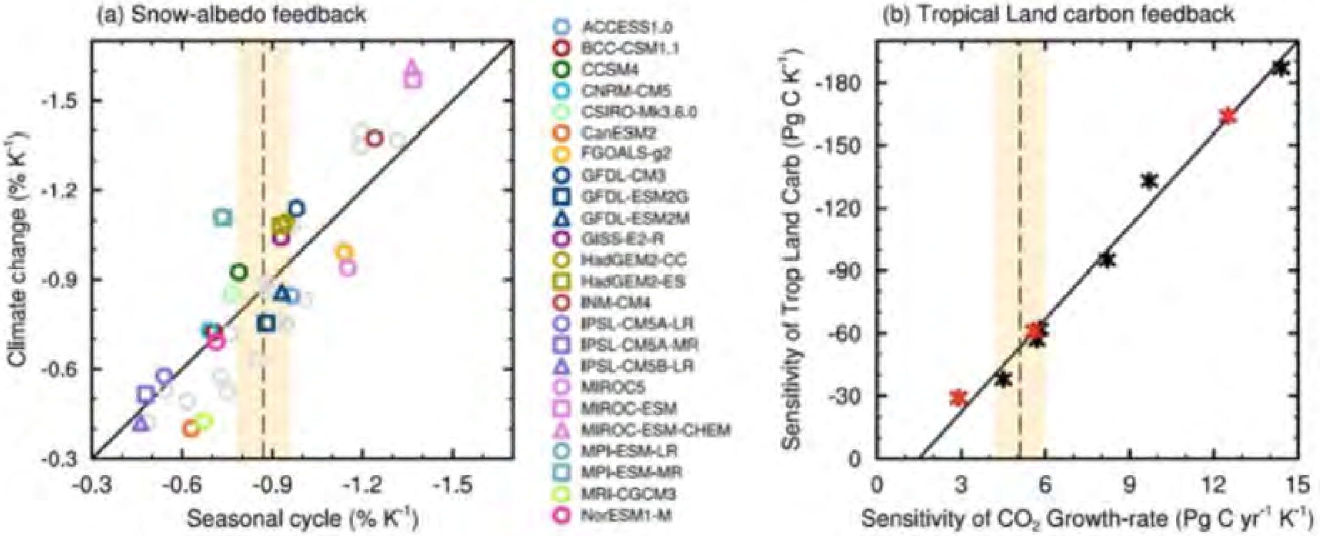
And I think this approach [next page] is becoming very popular. Bill referred to it. Sometimes it's called emerging constraints.

And Bill referred to this one [left panel]which is usually the poster boy, referring to snow albedo feedback which, they are looking at the seasonal cycle of snow cover in models, saying that correlates very well with the changes in snow cover as you go forward in the projections.

And then you can look at the observational constraint on the seasonal cycle of snow cover and you can improve your projection by potentially an order of magnitude.

We have to believe that the ensemble of models is certainly

“Emerging constraints”
or how to find relevant metrics for model quality



capturing this functional relationship. And then there is some cancellation of errors or something.

And this [right panel] is another one that came up recently that is attracting a lot of attention and a lot of skepticism.

This relates to the question of how much carbon the land surface is going to take up in the future, which is actually a pretty big uncertainty in the big picture here, almost as important as climate sensitivity uncertainty, maybe as important.

And this claim was, actually, I don't have the references here. I'm sorry. The claim here is that you can look at El Niño and look at how much carbon the land is taking up due to El Niño.

And that turns out to be in this small ensemble. And here there is a lot more skepticism with the models at capturing all the physics of snow. Snow albedo is simpler.

This works the same way. Use the observed changes in CO_2 on ENSO time scales, which are mostly a land absorption of carbon. Even though El Niño is in the ocean, it is mostly the land fluxes of the carbon that are dominating the response of CO_2 to El Niño and use that as a constraint on the models. And that seems to constrain. Also, the model ensemble mean is biased pretty high there.

So, that's one way of trying to find constraints or metrics with which to weight models. I don't use the word "weight." In fact, some of these outliers are valuable in determining the functional relationship. You want variations in models.

So, you can try the same thing with climate sensitivity, I have done this, a lot of people have. I haven't done this with CMIP5, though. Can you use the historic, something about historical temperature change?

In fact, if you use the historical temperature change, the models predict for greenhouse gas only or something to try to -- I mean, use observed temperature change to predict future temperature change. Doesn't seem to work. Certainly doesn't work as well as these things.

So, that's kind of, you have to be careful about what metrics you use. If you are interested in precipitation over the Sahel, we tried all sorts of things.

I happen to be interested in that. We haven't found a way to distinguish between models that dry the Sahel and models that don’t. We don't have a metric for that.

DR. KOONIN: So, the ability to then reproduce historical data is neither necessary nor sufficient to predict the future? Is that what I understand?

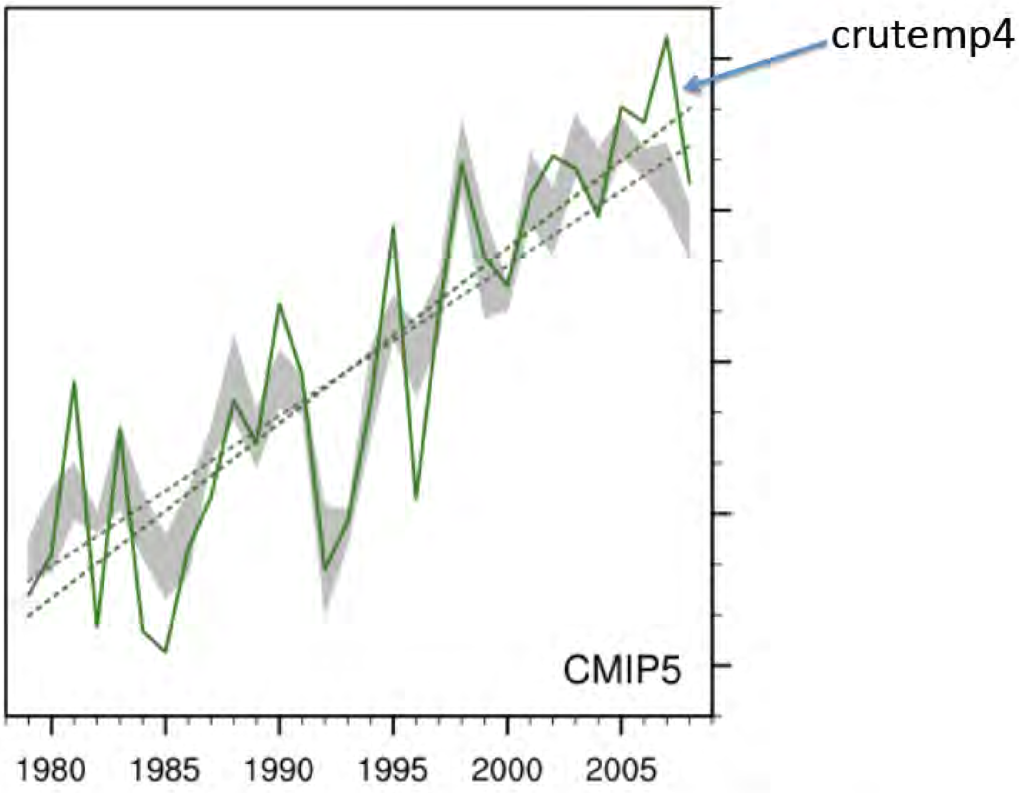
DR. HELD: Well, there might be something else you can use. And people are looking at things. There is a paper I just brought. I had a copy of "Nature" on the train to read the latest article on the subject.

People are trying to use information about the simulation of the observed cloud field to distinguish between models with high- and low-climate sensitivity.

So, let me talk about Arctic/Antarctic. I didn't check the time when I started, so just cut me off, Steve.

DR. KOONIN: Let's go for another two or three minutes and then we will take some discussion and then take a break.

DR. HELD: This [next page] will be my last. Let me mention this one, because this is one place I disagree with John. So, we do a lot of these



Mean land temperature in ensemble mean of CMIP5 prescribed SST simulations

simulations where I call it CMIP5 here.

These are what we call AMIP simulations. We just constrain the ocean temperature to be the observed ocean temperature. And we are looking at the atmosphere and land. So, the atmosphere and land are being predicted, but we are constraining the ocean.

But here, this is the land temperature, one of these CRU data sets. And this is the ensemble mean land temperature evolution you get, but from all of these AMIP simulations where you impose the ocean surface temperature. Looks pretty good to me.

I mean, you can worry about some discrepancies. I don't know if John would argue this was a coincidence. There is no information about the land surface in here.

How much of this is driven by

ocean temperature and how much is driven by forcing? This relates to your question of, can the ocean drive things -- three-quarters of this is driven by the ocean temperature. It has nothing to do with forcing. It's the ocean influencing the land.

So, our ocean and land temperatures are redundant to zeroth order, which is great, because that's what you want, is redundancy. We don't seem to be in disagreement.

And this is, I think John went over this period. Certainly from 1980, the land has warmed a lot more than the ocean. But that's captured perfectly well by the CMIP5 models, just specifying the ocean temperature.

I can stop or talk about Arctic/Antarctic.

DR. KOONIN: Why don't we stop and just take some more comments, questions, discussions.

Phil?

MR. COYLE: I wanted to hear what you were going to say about the Arctic and Antarctic, so I will ask a question about it.

DR. HELD: I was going to paint a big picture that has been very robust over the history of climate models. The two hemispheres are totally differently.

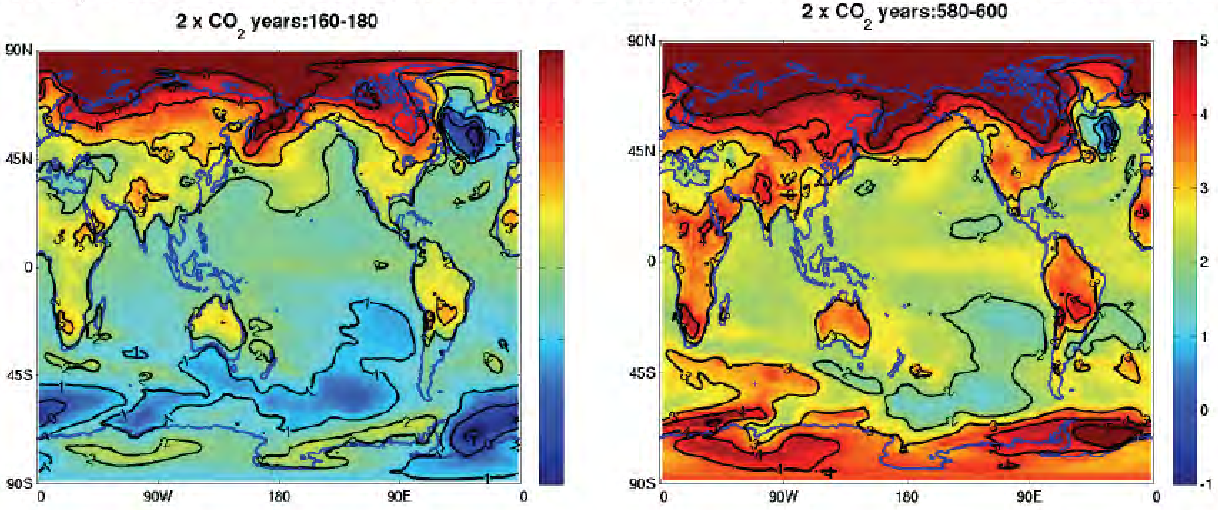
You couldn't design two hemispheres that are more different, especially in high latitudes where you have a polar continent versus the Polar Ocean.

And the transient time scales, this [next page] is showing shorter time scales.I happened to pick this figure. This is from a particular model, but it's fairly robust.

The transient response to increase in CO_2 is very northern-hemisphere dominated and it doesn't have to do with the amount of

Arctic vs Antarctic – hemispheric asymmetry in transient warming

Response to instantaneous doubling of CO2 in a particular model



160-180 years

600 years

land versus ocean. What that has to do with, is that on this time scale, the surface ocean, down to a few hundred meters has plenty of time to adjust.

It's really where your deep ocean is coupled strongly to surface ocean. So, northern North Atlantic and the Southern Ocean are where we have strong coupling to really deep water.

DR. CHRISTY: What color is the zero?

DR. HELD: Pardon?

DR. CHRISTY: What color is the zero?

DR. HELD: The blue is not cooling here. I'm sorry. That is kind of a poor choice. But the very dark blue is a cooling. Is that your question?

DR. CHRISTY: Yes, I just wondered about zero.

DR. HELD: Sorry about that. Yes, the numbers get cut off.

DR. SEESTROM: On the right you can see it.

DR. HELD: The zero is sort of the blue. Sorry. This isn't actually cooling except in this particular model. And then you go out to equilibrium. Of course you get a picture with more symmetry.

DR. KOONIN: So again, the essential physics here is the coupling of the surface of the ocean down to the deep --

DR. HELD: To the deep ocean. And "deep," you really mean deep here on these time scales. So, you are starting with a prediction that, for CO_2, that you expect the Arctic to be warming. The Antarctic isn't going to do a whole lot on this time scale. That's the starting point.

Now, this is actually a bigger gradient between the hemispheres that you see in observations, and that suggests there are some aerosols.

It's another piece of evidence for aerosol cooling in the north.

But so, zeroth order says, well, you don't expect too much would happen with Antarctic sea ice. And I think with zeroth order, that's kind of what you see. But you do see this increase which the models I think have a hard time getting.

So, what's going on with Antarctic sea ice? The variability is mostly wind-driven. That's thin ice. It's not multiyear ice like we used to have in the Arctic.

And you get stronger winds and we are getting stronger winds from the ozone hole, for example. You can blow that ice through Ekman Drift, as we call it. Combination of the stress and the rotation drives the ice further equator and then it sort of fills in by cooling.

But the zeroth order is just the wind drift driving gives you a lot of variability in this thin Antarctic ice. And it may be that the models aren't responding strongly enough to the ozone hole. It might be the coupling with the ocean is off or something.

We are not getting that increase in Antarctic ice in most models, anyway. But I think the zeroth order, the picture from greenhouse gases, you don't expect the Antarctic to do that much.

DR. KOONIN: Phil, you got your answer?

MR. COYLE: Thank you.

DR. KOONIN: Ben?

DR. SANTER: Just adding to that, Isaac, that's one region over Antarctica where there are also big differences between these different ozone forcing data sets, Chioni, et al., and this new one that Susan and colleagues have been working on.

DR. HELD: I get the impression

the models are underestimating the ozone hole.

DR. SANTER: So, that could be part of it.

I had one more question about the Kosaka and Xie paper. So, in the abstract of their paper, they claim that the hiatus is basically internal variability alone.

And I would just point out that, of course, in the observed SST changes over this region of the Eastern Equatorial Pacific that they are prescribing, it is possible that there are volcanic signatures in that. As I mentioned, most of these eruptions are tropical.

We do see signatures of those tropical eruptions in MSU lower tropospheric temperature after removing ENSO effects.

And given the tight coupling between tropical SST and tropical lower-tropospheric temperature, I would expect something to be there in the SSTs that they are specifying.

DR. HELD: I think it's mostly internal variability myself.

DR. SANTER: Well, you may be right. But I think where are they are wrong is claiming that it was all internal variability.

DR. HELD: I don't think you can claim that based on their experiment.

DR. SANTER: Yes.

DR. HELD: You can claim, as I tried to say, that the explanation has to flow through the Eastern Equatorial Pacific one way or the other.

DR. SANTER: It has to, but their experimental design does not cleanly separate bona fide internal variability.

DR. HELD: I agree.

DR. KOONIN: Okay, anybody else?

DR. LINDZEN: In the models, what causes the 1919 to 1940 warming?

DR. HELD: The models tend to underestimate it. There is no model shown here.

DR. LINDZEN: No, no, I am saying what do they do?

DR. HELD: Well, they miss the peak of the warming. The greenhouse gases leveled off in the World War II years.

DR. LINDZEN: It's not going to be greenhouse.

DR. HELD: I think there is some internal variability there in the models. And there are some models that can produce this with internal variability.

So, it's not implausible that some of this hiatus period or most of it is internal variability as well, which I think is what Kosaka and Xie point to as well.

DR. KOONIN: Let us take a

15-minute break and we will convene for some group discussion.

(Whereupon, a brief recess was taken.)

DR. KOONIN: So, we are in what we are calling a panel discussion among the speakers and subcommittee. And we would like to keep it largely, but not exclusively focused on Working Group 1 science.

But as I said at the beginning, we will allow excursions into programmatics and policy and so on to the extent that they don't get out of hand.

I would like to start off by seeing to what extent among our experts we can get agreement, beyond-consensus agreement, to a number of statements of increasing import and complexity.

And so, I will just ask as I start to read through these, if anybody has any objections, please speak up and that will precipitate a discussion.

I think the first is that the global temperature has risen certainly from, let's say, 1980 to 1998 or so in a fairly steep way, and that post-1998 or '99, we have seen a moderation of that trend, if not flat-lining of the temperature.

In other words, there isn't much disagreement about what the global mean surface temperature record is now compared to, let's say, ten years ago or so.

DR. LINDZEN: I don't disagree with the statement. But I think it is still terribly important to keep in mind how dicey the data is, and too, how small the temperature change is we are talking about. When you talk about sharp increase, it's a sharp increase of a few tenths of a degree.

DR. KOONIN: Yes, good. Even

to say at a higher level, there isn't much disagreement in the community about what the global temperature record is.

DR. LINDZEN: Including even if you go to the UK Met Office and so on, that there is an error bar that isn't far off from 50 percent.

DR. KOONIN: Far off from 50 percent of …?

DR. LINDZEN: The change over that period.

DR. KOONIN: Really?

DR. LINDZEN: Plus or minus .2 I don't think would be considered off.

DR. HELD: Over that time period, that sounds too big to me.

DR. SANTER: Certainly the error bars get much bigger as you go back in time. But in this period of time, I think that sounds rather large to me as well.

DR. KOONIN: Let me try to

phrase it yet another way. The IPCC has in either the SPM or I think chapter 2, a graph of GMST annual values. Would anybody dispute that there is a problem with those numbers with the uncertainties that are indicated?

DR. LINDZEN: Yes. Could I ask, are you talking about statistical uncertainty or systematic uncertainty?

DR. JAFFE: If you took the data and you ran a line through it and calculated the statistical uncertainty --

DR. LINDZEN: I defer to John on this. But if you looked at the temperature change latitude band by latitude band for the period from '79 to '98, what was the contribution from the tropics vis-à-vis high latitudes?

DR. CHRISTY: Well, the high latitudes are pretty much driven at both.

DR. LINDZEN: So, the tropics were small?

DR. CHRISTY: Fairly small, yes. But you were talking about statistical uncertainty.

And what we call statistical, which is exactly how you defined, how well does the trend line depict some kind of modes versus measurement uncertainty, and other kinds of structural uncertainties and data, spatial coverage and all that kind of stuff.

DR. JAFFE: Right, that was the question.

DR. CHRISTY: Yes, those are two different ones. And for that period, they are probably about the same, probably less than a tenth of a degree for an annual average.

DR. KOONIN: All right. Let me try another statement. This one is going to be interesting, that certainly the atmospheric CO_2 has gone up over the last century, and it is largely, almost exclusively an anthropogenic increase due mostly to burning fossil fuels.

DR. CHRISTY: I think people like me would say that's what we have been told, and I don't see any reason to disagree with it very strongly.

DR. KOONIN: You know, even I, no expert, I can cite what I think are several reasons why there is a good reason to believe it is anthropogenic. Northern hemisphere is bigger than southern hemisphere. The isotopes ratio is consistent with fossils.

DR. LINDZEN: But that is one measurement, I believe.

DR. KOONIN: Well, the isotope ratios get measured all the time, almost daily.

DR. LINDZEN: Yes.

DR. KOONIN: That's one method.

DR. CURRY: It's beyond the scope of what I can critically evaluate.

DR. KOONIN: You don't disagree?

DR. CURRY: No, I don't have any reason to disagree with that.

DR. KOONIN: Let me try a third. Sea-level rise has continued over the last, I don't know, let's say 60 years. It has been going up.

The current rate over the last decade is higher than it has been historically, but not at all unprecedented in the record, again, consistent with the uncertainties that we have seen in the figures.

DR. CHRISTY: How far back are you going with this?

DR. KOONIN: Oh, I don't know. Let's go in the last 100 years. That was the extent of the graph that we saw for the rate and all that's in the framing document.

DR. CHRISTY: So, the rate right now is higher than the average rate over the past 100 years?

DR. KOONIN: But not unprecedented if you go back to 1940 or so.

DR. CHRISTY: As Judy said, that's not my area of expertise, but I believe that.

DR. KOONIN: None of you are out on the street saying this is wrong?

DR. LINDZEN: I think Carl Wunsch's statement was probably the most accurate. It's impossible to say if it is significantly different from the long-term trend, because there are different instruments.

You have all sorts of comparisons. It's just hard to say anything. The statements that it's unambiguously accelerated and so on, I don't think there is a basis for it; may be right.

DR. KOONIN: That's yet a step further than I was going. I mean, it's certainly been going up over the last century.

DR. LINDZEN: You know, you have all sorts of new data and they are measuring different things. The problem with sea level is, until you had satellites, you were measuring differentials between land and sea governed by tectonics.

So, that was a huge mess. Now you have something that is more absolute. How do you compare the two?

DR. KOONIN: Right. Let's try another one. We have signatures of anthropogenic influence on the climate, but there is disagreement as to how strong that influence is and what it will be in future decades.

Let me not say "disagreement," but there is uncertainty. You have signatures, but what is going to happen in the future is uncertain.

DR. HELD: It's an empty statement unless you quantify it.

DR. KOONIN: Well, propose a quantification and let's see. The IPCC said 50 percent, right, half?

DR. CURRY: No, more than 50 percent.

DR. LINDZEN: Well, 51 you are not going to argue about.

DR. CURRY: I really don't think it's 51. When they say "most," they are really thinking it is more.

DR. KOONIN: So, help me here. What we can have an eruption, right?

DR. LINDZEN: What I was suggesting is, if it's 100 percent, that leaves you open to any sensitivity down to about .75. If it's less than 100 percent, then you go down proportionately as a possible lower bound.

The thing that I find peculiar about the IPCC statement is, it's sort of a red herring. It's made to the public. It's immediately interpreted as meaning a disaster is around the corner.

But the statement itself is compatible with a wide range of possibilities, some of which are totally benign.

So, it's a case where the scientific community is permitted to say something sort of reasonable with the assurance that the advocacy community will interpret it as it wishes.

DR. KOONIN: Let me try; we are going to go back again. Would anybody disagree with the statement that we have seen anthropogenic influence on the climate?

DR. CHRISTY: By "climate," you mean temperature?

DR. KOONIN: That's probably the simplest interpretation, and I mean more than regional.

DR. CHRISTY: I agree with that, but I don't know how much.

DR. KOONIN: That's the second step -- how much?

DR. CHRISTY: We don't have a thermometer that says Mother Nature did this much and humans did that much.

DR. KOONIN: I get a sense we are starting to approach our limit of agreement here.

DR. LINDZEN: In a sense, we wouldn't have needed any data or proof or anything for that agreement.

DR. KOONIN: Because?

DR. LINDZEN: The physics says you should have something. You have to have a huge negative feedback to not have anything. You know, normally there would be something.

DR. KOONIN: So, a lot of the discussion today has been just how much?

DR. LINDZEN: Yes, that's a

common physics concern.

DR. KOONIN: Well, there are places you can do experiments it's a lot easier to answer that question.

DR. LINDZEN: Yes.

DR. KOONIN: So, evidence in the historical record, how much of the historical record is greenhouse gases versus aerosols versus natural variability versus we just don't understand about what the forcing is, et cetera?

DR. LINDZEN: Consistent with what sensitivity.

DR. KOONIN: All of that.

Bill, you look like you are about to speak.

DR. COLLINS: No, no, I am just nodding.

DR. KOONIN: Does anybody want to propose a statement that goes further to see if we can get consensus? Well, let's try the IPCC half. More than half --

DR. CURRY: To me, half is a very awkward divider, because I probably think it's 50 percent plus or minus a bit. Once you say "more than half" --

DR. KOONIN: How about if we put the "half" in quotes?

DR. CURRY: Or maybe divide it up into three.

DR. KOONIN: Terciles!

DR. LINDZEN: The interpretative statement of the IPCC would probably go a long way to clarifying the issue, namely, a statement that a significant part, say, half in quotes, whatever you want, of the observed temperature change over the last 50 years or whatever is likely to be anthropogenic. But that leaves open the sensitivity over a wide range.

DR. KOONIN: You said "likely to be anthropogenic." IPCC says --

DR. LINDZEN: I think Isaac is

right. I mean, that didn't come from any statistical analysis.

DR. BEASLEY: But I would say when addressing this question, what fraction is anthropogenic, I think I would like to hear the comments about what is in the IPCC report or a statement that we feel it's better to do it this way.

In other words, you can't just leave that hanging there, because there is a big gorilla out there, right? So, I think it needs to be addressed as a straw man if nothing else. I am not saying you have to accept that as the best characterization.

But I think from the point of view of thinking through from an APS point of view, and if you all think that's not the best characterization and this would be better, I want to know.

DR. CURRY: Even if you say

it's somewhere between 51 and 95, that's a huge --

DR. KOONIN: Range.

DR. CURRY: -- range. And yes, so it's not, to me, I never thought it was a useful statement, because it's a huge range. And whether it's 51 versus 95 makes a huge difference.

So, I don't think it's useful. And this is the key question that we don't know. We don't know how much is natural and how much is anthropogenic.

DR. LINDZEN: But also, after that, we still don't know the sensitivity.

DR. KOONIN: That's a different issue.

Ben, do you have anything?

DR. SANTER: I think Isaac was going to go first.

DR. HELD: I think the AR4 statement was a statement of sensitivity. The AR5 statement in the SPM is not, for better or worse.

DR. CHRISTY: My answer is I don't know.

DR. KOONIN: Ben?

DR. SANTER: What gives me confidence in the reality of detection of the human effect on climate is the internal and physical consistency of the evidence.

Back around the time of the second assessment report, one of the criticisms of the balance of evidence finding, just viable criticisms was, you folks have essentially only looked at surface temperature.

If there really is some human-cause-to climate-change signal lurking in the system, go after it in water vapor. Look at ocean heat content. Look at circulation changes. Look at a bunch of different things.

And that's what has happened. And to me, the power of that sort of work is that it's not a house of cards resting on one surface temperature data set.

People have interrogated very, very different observational estimates of ocean heat content change, moisture over oceans, circulation changes.

And the bottom line in all of that is, there is internal and physical consistency. To me, that is very powerful.

DR. COLLINS: I would echo that I think something that reflects the multiple lines of evidence and analytical techniques that point to what appear to be at least a plausible common cause would be an accurate reflection of the information that you have seen presented by the IPCC and heard from several of us today.

But my understanding from our discussion over lunch or actually when we were milling around in the hallway is that this is not going to be a statement, APS's assessment of the IPCC assessment, right?

DR. KOONIN: No, we are not doing that.

DR. COLLINS: Nor are you going to come up with necessarily with sort of a well -- let's -- oh, greater-than-half-less-than-half number while we are sitting here sipping coffee. A great deal more work went into all the different estimates you heard from us today than that.

So, I am a little nervous about, I know what you are trying to do, but it makes me -- even putting the number "half" in air quotes, I think, it is probably a disservice to the amount of work that has gone into --

DR. CURRY: But a lot of this, at the end of the day, the "half" comes from expert judgment, right? It was a different group of people sitting around a different table.

DR. KOONIN: With a different set of coffee cups.

DR. COLLINS: Well, but remember, the IPCC is an assessment of literature and of model runs. Our job is not to do research. We are not sitting in that room sort of guesstimating what those numbers were, Judy.

DR. CURRY: More than half, more than half, more than half, that is an expert judgment. That's not anything that popped out of any statistical analysis. It even says in there "expert judgment."

DR. KOONIN: I might point out one social or political comment. The Interacademy Council review of IPCC when it happened, said that, among other things, the IPCC should give a clear line of reasoning for whatever statements it made. And "expert judgment" is, for me, a little bit too coarse a reason.

DR. CHRISTY: A reason I don't have confidence in these model things and the consistencies that some people find is I don't find consistencies in the metrics I check.

And I think I am checking some pretty basic metrics. And the models just can't tell me why what has happened.

DR. KOONIN: Isaac and then Ben.

DR. HELD: It gets back to physical consistency. I worry, are there smoking guns out there that will change the consensus? I think we focused on two of them, the hiatus and the tropical, upper tropospheric warming. Those are real issues. I don't have the answer.

I think they are related, although I think the problem with the satellite data on the upper tropospheric warming seems to go back a little bit earlier as John showed. But I think to zeroth order, they may be the same problem, that the tropics isn't warming up very much.

Over the satellite era, the models have overestimated tropical warming. And at least in the ensemble mean forced response, they are overestimating Arctic warming. There are some interesting things.

DR. SANTER: They are underestimating the observed Arctic warming in the lower troposphere.

DR. HELD: I was thinking of sort of a normalized -- anyway, this is focused on the tropics. They overestimated. It could be something else.

Let me just describe another side of the coin, which is a little, I don't think it's esoteric. But when we try to force a model to have something different than, not warming as much as the moist adiabatic in the tropics, we get a huge increase in activity in the tropics, increase in hurricanes.

This is the model which produces pretty good distributions, spatial, seasonal distribution of tropical cyclogenesis. This has consequences. You are destabilizing the atmosphere pretty dramatically if you take this at face value.

There are arguments in the literature that that is not happening, that we don't see anything like that. The moist adiabatic assumption is the most conservative one you could possibly make as far as minimizing the impact of warming on the tropics.

DR. CHRISTY: It works on a monthly scale with the data as well. You see it on a monthly scale just like that. It does work.

DR. HELD: The models, as far as upper tropospheric warming during ENSO events, the models are pretty much spot-on. What is the difference in the physics? The atmosphere has a time scale of a month or something.

DR. LINDZEN: Without a great deal of experience on it, I have often wondered at the upper troposphere where you have the Rossby radius.

I know from personal data and analysis that you have far more horizontal variance in the surface boundary there. So, if you have a sampling problem, the sampling problem is worse at the surface.

DR. HELD: That has motivated a lot of ongoing work on whether the surface data is consistent with the upper-air data. A lot of that is still in press, but it's an open question.

DR. LINDZEN: Maybe we assume

the surface is good and the upper is bad.

DR. HELD: I am interested in whether they are consistent, as you are.

DR. KOONIN: Ben?

DR. SANTER: Getting back to this issue of the interpretation of "most" and what that means, Judy mentioned that that largely is coming from expert judgment, not wholly, I would say.

There are studies like a study Tom Wigley and I published in "Climate Dynamics" where we used the very same sample models that Dick was talking about and comprehensively explored forcing uncertainty space, uncertainty space in ocean diffusion, uncertainty space in climate sensitivity and looked at fitting two observed surface temperature data over various periods of time.

And the results of that study

was that the IPCC finding was likely conservative. And indeed, it's extremely likely that most of the observed warming observed over the second half of the 20th century is due to anthropogenic greenhouse gas concentrations.

It was very difficult to find combinations of climate sensitivity and aerosol forcing and diffusion, ocean diffusion that would give you something substantially less than half of the observed warming.

DR. CURRY: To me, that is circular reasoning where you define multidecadal natural internal variability out of existence. It is just defined out of the problem.

DR. SANTER: You are not defining it out of the existence.

DR. CURRY: No, out of your analysis, though, in terms of attribution.

DR. SANTER: Well, you are

looking at surface temperature data that are affected by both external forcing and by internal variability. So, it's not entirely out of the analysis.

DR. CURRY: But there have been dozens of those kind of analyses and they come up with different kinds of results depending on how you frame it and what you look at.

So, I don't find that, to me, all that convincing. I am still in the camp we don't know. And I don't rule out a 50-percent kind of answer, actually lower.

DR. SANTER: Well, I am just pointing out that that kind of conclusion doesn't arise from expert judgment alone.

There is actually a substantial amount of work with the simple models that tries to directly address the question how much is due to anthropogenic greenhouse gas increases.

DR. KOONIN: My sense is that we are just about hitting the limit of what we can get everybody to agree on here.

DR. CHRISTY: A lot of "I don't know”'s coming out.

DR. CURRY: Some people know; some people don't.

DR. KOONIN: We could take this in another way. We could get on to “How do we make the science better going forward?”

DR. CURRY: To me, this is what I think the APS statement should be out, not trying to judge stuff which is really outside, in many ways, the expertise of the Society.

But what do you see from all this where the Society and the membership can contribute going forward? To me, this is the big contribution that APS can make.

DR. CHRISTY: We would love to

hear from you.

DR. CURRY: Yes.

DR. KOONIN: What we think?

DR. CHRISTY: Yes, you seem to have a menagerie here.

DR. KOONIN: Well, I think we haven't quite milked all the information out of you all, yet.

One thing I heard is that a longer, more consistent, more precise, better coverage in the data of all the relevant variables is extraordinarily important.

DR. CURRY: And uncertainty in error assessments in the data, better assessments.

DR. KOONIN: Yes, a lot of the "I don't know"’s were couched in “well, we don't have good enough data back far enough.” And so, it seems to me that if anybody is recommending anything to the decision-makers, it's make sure the data get covered in a continuous, precise way.

DR. CHRISTY: Along with that, for 15 years, I have been going to Congress trying to get this done, a red-team assessment of the climate modeling enterprise.

These are truly independent, I mean, it would be great to have people from of APS or the engineering societies who know about modeling and simulation and how to test and so on. If that independent group could come and then look at the insides of these things.

DR. SANTER: Excuse me, John, but that was the rationale behind the setting of the PCMDI that the Department of Energy over 20 years ago wanted to set up a group that had no horse in the race, that was not actively involved in model development to independently analyze fidelity with which models capture important features.

DR. CHRISTY: From my view,

that was mainly an advocacy group that came out of there.

MR. KOONIN: I will let Ari speak since he has a personal involvement in that.

DR. PATRINOS: No, I don't think it was an advocacy group.

DR. SANTER: Our concern is getting the science right. I would dispute that characterization that PCMDI and other model evaluation centers are advocacy centers. The greatest good for us is getting the science right.

DR. ROSNER: If anything, the Department of Energy has always been accused of being partial in proving that this is not a problem. That's the battle we used to fight with the department all the time.

DR. CURRY: I think there is a disconnect between what Ben and John suggest. I think John is suggesting people looking at how the models are constructed and the process of how they are evaluated rather than just the actual verification and statistics, looking at a more meta-kind of look.

DR. KOONIN: At some of these experiments that we have been talking about, numerical experiments?

DR. CURRY: Right.

DR. COLLINS: I would like to point out that the committee also is not -- we are not handing results back to the international community out of black boxes.

I think it's very important for the committee to understand that, at least in a number of cases, what you have seen today certainly coming out of the work that Ben did with PCM, the work that I have done with the CCSM and Isaac's work with the GFDL model, these models are all -- we live in a glass house.

So, the source code is

available. The input is available. The output is available. The models are completely described in peer-reviewed literature. If you want access to it, I can show you how to download the model and can run it on your Mac.

But these models, they are not black boxes. They are subject to a great deal of public scrutiny.

DR. CHRISTY: But there is virtually no funding to do that.

DR. COLLINS: Fair enough. But I want to make it very clear --

DR. CHRISTY: So, it's not done. It's not done.

DR. COLLINS: Yes, but the community is at least living up to its side of the transaction in terms of being extremely open.

DR. KOONIN: For the U.S. models?

DR. COLLINS: Also increasingly, too, for the European.

Hadley Centre is distributing their model. The ECHO model been in the public domain since the 1990s.

DR. LINDZEN: But it's not always a practical issue.

DR. COLLINS: I understand that. But I am just saying that the community, we understand there may not be a partner there, John, with whom to handshake.

But that is essentially, we have negotiated our side of the transaction. This information is entirely in the public domain. So, if you have an issue with it, have at it. It's all there.

DR. KOONIN: John said the word "red team." One of the things that characterizes a red team is the attitude that it goes in with, which is, "I am going to show what's wrong with this." And we all know that's a very important attitude to have in science.

DR. ROSNER: But if I may, also, we did have that with JASON.

DR. COLLINS: That's right.

DR. ROSNER: Within 15, 20 years, maybe it doesn't represent the entire non-climate community, but it came into the picture with a jaundiced eye and wanted to critically look at it, and it did.

Now, it may be time to do it again or may be time to do it again on a different scale. But it's not that it wasn't done. And I take issue with the fact that it was an advocacy group.

DR. CHRISTY: When I say that, I mean there was very little that came out that was critical of the models, of the type of analyses and stuff that I had done.

DR. SANTER: That's not true.

DR. KOONIN: Ben or Bill?

DR. SANTER: Sorry, I would really dispute that.

DR. CHRISTY: I am sure you would.

DR. SANTER: I think we have not swept differences between models and observations under the table.

DR. KOONIN: Whoa, whoa, whoa.

DR. SANTER: Excuse me. Let me finish, please.

I would say, John, that unlike you, who just presented these discrepancies and threw up his hands and said, oh, we don't understand these things, we have actually tried to understand why the differences exist and whether they are bona fide model response errors, whether they are forcing errors, whether they are internal variability errors.

So, I don't think it's sufficient to just do the kind of analysis you have done, show discrepancies and say this proves that all models are wrong or are too sensitive to anthropogenic greenhouse gas concentrations. That is not helpful in advancing the state of the science.

DR. KOONIN: So, I was surprised. I thought one of the heating profiles John put up in the model comparisons, that is a pretty powerful figure.

There are perhaps reasons, inadequate forcings, why that kind of discrepancy exists. But it does exist and I cannot find a hint of it, or maybe there is, in the IPCC documents. That is a failure.

DR. COLLINS: It's hard for me to respond to that since I don't know. John, what is the status of that publication, those results and their peer review?

DR. CHRISTY: We have already have one of those out.

DR. COLLINS: But the heating profile paper, the profile results?

DR. CHRISTY: That was in 2011,

I believe.

DR KOONIN: Should have been there.

DR. COLLINS: Okay.

DR. SANTER: Steve, I would point out that we published similar results showing vertical profiles in science in 2005 showing the CMIP3 results compared with satellite and weather balloon profiles, and in 2008 as well. So, it's not some startling, new thing.

DR. KOONIN: So again, I ask, as we saw it today. Gosh, that's pretty interesting. Somebody needs to explain that discrepancy. First time I have ever heard about it. It's not in the IPCC report.

DR. CHRISTY: It's in the IPCC. But as I said in the observation chapter, they called it, well, the observations said, "We have only low confidence in the observations." That allowed the chapters 9 and 10 not to really address it.

DR. KOONIN: That's interesting because I heard from Ben and Bill it's the forcing we don't have confidence in, not that we don't have confidence in the observations.

DR. CHRISTY: In chapter 2, which is about observations, probably not observations on forcing, but that's what they said. And I just don't agree with that. I think for this problem, those data are certainly good enough.

DR. KOONIN: Isaac?

DR. HELD: We can focus on the AR5, but this problem has been recognized for a long time. And there was an academy committee specifically devoted to addressing this issue.

DR. COLLINS: In 2000.

DR. HELD: Way back, yes.

DR. SANTER: And a USCCSP report in 2006.

DR. HELD: You can't get a model to depart from the moist adiabatic very much. As far as in the atmosphere, the models are very stiff. You can get the ocean temperature to do different things.

DR. KOONIN: I am going to go back to a higher level. We sort of entered this little discussion about model red team or closer scrutiny of the models.

Okay, that's one thing one can imagine the Society will be opining on, the data and others. We have talked about that. Are there other things?

Phil?

MR. COYLE: Well, I would like to go back to the question I asked Dr. Christy, namely, what do you all think we, the United States, should do that we are not doing now, do differently, do additionally if the APS says anything about this, which has not yet been determined? We are early in the process.

But if the APS says anything about this, it may have policy implications. So, if our experts agree on certain things that you think the country ought to be doing, that would be important to know.

DR. HELD: This is in relationship to science or in relationship to politics?

DR. KOONIN: Science, science, science first, science.

DR. CHRISTY: Back to the observations --

MR. COYLE: Is it funding for balloons? What is it?

DR. CHRISTY: This country could establish the right kind of balloon stations, for example, in places that can't afford it, that don't have the infrastructure to do this.

DR. KOONIN: John, sorry, but

the right kind of balloons?

DR. CHRISTY: Balloon stations and other kind of remote sensing that is ground-based. Then, and Judy might know about this.

You're still on the NASA NAC?

DR. CURRY: No, thank God.

DR. CHRISTY: Oh.

The satellite systems are, they are threatened. And the satellite systems that we have are really the only way to get some global pictures of this stuff.

DR. BEASLEY: Can I ask you, just a rough answer, the balloons or the ground-based stuff you are talking about, that doesn't strike me as something that is hugely expensive, hundreds of millions?

DR. CHRISTY: No, it would be less than that. But, you know, are you going to get a guy out of Kerguelen Islands in the South Indian Ocean to do it?

DR. BEASLEY: No, no, I understand. But it is not like the diseconomy in energy physics where it's up there in the billions now?

DR. CHRISTY: No, no, no.

DR. ROSNER: But, I think, to be fair, you have to scale the needs of a program like that to the size of the budgets of the agency that supports that work.

DR. BEASLEY: That's fine.

DR. ROSNER: NSF budgets are different than NASA budgets.

DR. KOONIN: Again, we shouldn't worry about if the APS says anything at all about this. It should not worry about that.

Bill and then Ben.

DR. COLLINS: I think several people have mentioned that we are entering an era where we need to keep the observational networks running.

The other tension here that we haven't talked about so much is the fact that these networks are primarily built or have been traditionally built for operational weather forecasting. And those observations don't need to capture long-term trends.

So, they are inherently not designed to be accurate over longer time scales. And what inevitably happens, what is happening now, I shouldn't even mention agency names, but the agency that -- you frequently see climate sacrificed to weather. That's basically the sacrifice that happens.

And what would be nice is a statement that says both observations are intrinsically valuable and it would be nice to have observations that are useful for both weather and climate, since they are both end-proposition, regardless of what the observing system is.

And what that means, to

translate it into English, is that the sensors need to be carefully characterized so that we can build long-term series for them and ideally be a little bit more accurate than they currently are. But a good network, maintain the network.

DR. CHRISTY: We have written in our reports about this very thing, exactly what you are talking about.

DR. COLLINS: That's right.

DR. KOONIN: Didn't some agency that shall remain unnamed establish a national climate service?

DR. COLLINS: No, they did not.

DR. HELD: They wanted to.

DR. KOONIN: They did not?

DR. COLLINS: No, they did not. That went down in flames.

DR. SANTER: I just wanted to point out that if the APS committee is going to make some statement along the lines John mentioned regarding red teams and the need to subject models, their development to more scrutiny, I hope the APS will do the same with observations, particularly with satellite-based estimates and weather balloon-based estimates of atmospheric temperature change.

One of the things that I have learned over the last 15 years in my involvement with the MSU issue is the extraordinary uncertainties.

It's a very, very difficult job to construct climate-quality data sets from well over a dozen drifting satellites with all of these very, very complex orbital drift effects that affect the sampling of the diurnal cycle, uncertainty in the diurnal cycle.

And we saw that today. Somebody asked the question well, in those overlapping trends, why do you get the difference between the red and the blue lines?

To me, we have only a couple of groups that are looking at these satellites-based estimates of atmospheric temperature changes. They yield different results.

I think that having a better understanding of why they differ and what the real residual uncertainties are in those measurements and in the balloon measurements with their equal difficulties with changes in instrumentation, the thermal shielding of the sensors, those are real things.

And oddly, there are far fewer groups looking at those issues than there are climate modeling centers. So, if you are going to say something about the need to red-team climate model development and analysis, I would hope you would say the same about the development of observational data sets.

DR. BEASLEY: What about in the

ocean?

DR. SANTER: You know, the ocean, it's the same thing. As Judy mentioned, there are XBTs. There are these Argo floats. There are buoys. There are a whole bunch of different measurement systems and they change in a spatially and temporally nonrandom way.

So, when you try and identify biases in each of these and adjust for these spatially and temporally nonrandom changes over time, it's tough.

You want to do ocean reanalysis and depending on the ocean model you use, you get different results if you use an ocean model to fill in the gaps and assimilate those observations. It's an equally tough problem.

DR. BEASLEY: Probably more expensive.

DR. JAFFE: I would like to

deflect the question back to this question of policy. APS statements and reports are fundamentally scientifically based, but they do make policy recommendations.

And to quote from a famous APS statement, "The APS also urges governments, universities, national laboratories and its membership to support policies and actions that will reduce the emission of greenhouse gases."

And it seems to me that if we write a report that makes motherhood statements about supporting -- not "we," but you, and then the APS adopts a report that makes motherhood recommendations about supporting more research, makes acceptable statements about holding up models to greater scrutiny and data collection to greater scrutiny, the elephant in the room will be well, what do you think about policy recommendations on greenhouse gases?

And I wonder whether that should be a subject for discussion here while we have some wonderful experts.

DR. KOONIN: We, of course, talked about that in subcommittee. And when you do that you start to get into issues that extend beyond the expertise of physicists.

You are into certainly ecosystem. You are into economics. You are into mitigation technologies. You are into value judgments. How much do you value today versus tomorrow?

And it's now obvious to me that we will have a discussion, I am sure all of us, in the course of putting a statement together, about to what extent do physicists have a special claim on that kind of knowledge or that kind of expertise.

DR. JAFFE: You make very

strong statements about nuclear weapons.

DR. KOONIN: Well, we can argue about that. Judy, I want to hear what she has to say.

DR. CURRY: Personally, I don't think the scientific societies should make statements about those kind of, what I would call public policy that is not related to the policy of science like we need more observing systems and things like that.

Apart from the expertise, I mean, you should only speak to where your expertise is and as you describe. And I am not even sure that APS has sufficient expertise on the climate issue to be making a statement at all. That is my personal opinion.

And the AAAS held a workshop and I did mail that to the committee I think indirectly through -- I think Steve must have gotten it -- about the AAAS had a workshop on what is the appropriate thing for advocacy, for individual scientists and what is irresponsible advocacy and for institutions?

And it gave some criteria. And the first one is speak to your area of expertise. And so, not only is mitigation policy outside your expertise, I would even argue that the whole issue of global warming climate change is broadly outside the expertise of the Society.

Now, APS has been very active in talking about vaccines and all sorts of things that I imagine they don't have any expertise at all. So, that has been sort of the history, a lot of advocacy.

But I think the Society should step back and think about, you know, what defines responsible advocacy for your society. And if you are going to make a statement, I think it would have a far greater impact if you do really stick to the things like better observing systems, disagreement, this is how we can sort it out, and some specific rules for the Society and topical areas that you want to get into.

Then I think you have a more powerful and useful statement than advocating for greenhouse gas emission policy. That's my personal take on it.

And it's speaking to your expertise. And once you go outside of your expertise as a society, apart from what the members, they can raise their hand and vote oh, we all want to say something about mitigation, but at the end of the day, it is outside their expertise.I would be very careful.

DR. KOONIN: Dick and then Bill.

DR. LINDZEN: I would like to

see, getting back to the science issue of improved observational networks and so on, could we do, or could somebody do the Gedanken experiment of saying, let's say we had that system and it was running for five years or ten years. How would it change our assessment of anything?

DR. COLLINS: I think the issue, Dick, is I am after longer records.

DR. LINDZEN: Yes, yes, and so on. So, if you are asking for something, it probably would pay to show explicitly what it would resolve over what time.

DR. KOONIN: You can do pseudodata experiments, right?

DR. LINDZEN: Sure.

DR. KOONIN: Take your model and generate 1,000 years of data.

DR. LINDZEN: 1,000 years?

DR. KOONIN: All right, 100.

But then you assume you know everything or you know only 50 percent, et cetera, or even ten years? Have those things been done, that kind of thing?

DR. LINDZEN: No.

DR. SANTER: Yes.

DR. COLLINS: Well, yes, they have. We have run climate observing system simulation experiments for other applications, Dick, for things like some of the NASA satellites.

We have looked at that issue, tundra detection and climate change using observing system networks. So, it has been done.

DR. KOONIN: Bob and then Bill.

DR. ROSNER: I want to follow up on what Judy said. So, it struck me that one area of physics uncertainty has to do with deep ocean sampling.

The question I have is, so, what would be the experimental campaign that you would need to mount in order to actually set up the physics to be able to improve the models? What is involved?

DR. CURRY: Okay, well, there is tracers, argon and various other things that are used to look at that, gravity, wave breaking associated with bottom topography.

It's something that people are working on it. But to me, this is looming as if the ocean ate the global warming, we have to understand some mechanisms.

DR. ROSNER: But I am asking, what would you need to do? That's what I am asking.

DR. LINDZEN: I think even in the oceanographic community, the people I know would not have something ready at hand saying "if only we had this."

DR. CURRY: Right. It's very subtle. A lot of these things, yes, it's a tough problem. But I think thinking about it in a meta way would be beneficial.

DR. ROSNER: So, for example, if you wanted to understand thermohaline mixing, say, deeper down, are there experiments that people have --

DR. HELD: Ongoing experiments. They are quite expensive.

DR. ROSNER: Never mind that. I am just curious what has been talked about?

DR. HELD: There's tracer release experiments. People go out, release sulfur hexafluoride and come back five years later to measure it.

DR. ROSNER: Where is it?

DR. HELD: Yes, where is it? And there are natural tracers that are arguably even more useful. CFCs are the best.

DR. LINDZEN: And our field benefitted greatly from the nuclear tests.

DR. KOONIN: You sold out!

DR. HELD: I would just be careful. I don't know if I still have the floor here?

DR. KOONIN: Yes, you do.

DR. HELD: We have heard discussions of mixing being important for getting the heat down. That's not necessarily the case.

You can get heat down below a certain level just by adiabatic rearrangement of water, just tilting the isoclines of temperature.

You have to be careful when you talk about mixing, quote/unquote. It's not clear that's what is going on on these time scales at all.

DR. CURRY: That's a question whether to what extent it is mixed versus not in terms of --

DR. HELD: That has a big effect whether it is going to come out quickly or not.

DR. KOONIN: Yes, yes, yes, some better characterization of the deep oceans.

DR. ROSNER: That's what I am getting at.

DR. KOONIN: We don't have to get into that.

DR. HELD: I think one idea is getting the Argo float program to go down to the bottom of the ocean. Right now it doesn't. That would help.

DR. KOONIN: I want to come back to the question Bob raised and Judy addressed a little bit is, how appropriate is it for the Physical Society to go beyond the obvious scientific expertise?

And we have heard one instance cited already, Bob, in the nuclear weapons example. We can have a group discussion about how effective that particular set of statements has been in modulating U.S. nuclear policy.

You know, if we had experts in deterrence, if we had experts in geopolitical doctrine, et cetera, et cetera, then it probably would have been more effective. The Society does not have that, at least collectively. But its individual members do, some of whom we know well.

DR. JAFFE: When we studied critical materials and made policy recommendations, we had a committee that had geologists, economic geologists, physical chemists and so on.

So, we did do what I thought Judy was suggesting we do in getting a group of expertise. That was the full-fledged focus, a small group getting together and making recommendations.

DR. KOONIN: We were just five random POPA members none of whom are climate experts here.

DR. BEASLEY: Bob, this is a

conversation for APS for another day, but I think whether we have to or need to or want to get involved in the bigger crosscutting thing is something POPA needs to deal with.

DR. ROSNER: Yes, that's a POPA discussion.

DR. KOONIN: Yes. But as somebody said, it's good to get the opinions of our experts, which is why we're doing this.

DR. BEASLEY: No, no, I understand. We sort of know that, but we haven't done anything.

DR. KOONIN: Bill?

DR. COLLINS: In response to your question, I mean, I think there is an issue. The things which are certainly within the APS's range of expertise like radiative transfer and spectroscopy where you would be on extremely safe ground enumerating some of the things that are obvious and undisputed.

We have changed the chemistry of the earth's atmosphere. We have CO_2 levels that are as big as they have been in three million years. If we double that again, they are going to be higher than they have been in 34 million years. We know the CO_2 is anthropogenic because of isotopic analysis.

We know exactly what it does to the radiative transfer budget. It's been verified by satellite. It's been verified at the ground. We know that CO_2 is a greenhouse gas. We know that we are draining the energy budget out of the earth's system appreciably.

Those are all things that are completely within the APS's sphere of expertise and extremely safe statements to make. And even those would be regarded as sort of shockingly novel, I think, in certain circles of the scientific community.

But they are extremely solid statements and completely within your area of expertise.

And by the way, many of us, as Isaac pointed out, are trained as physicists. It's not as if we are two disjointed communities. I was trained as an astrophysicist and a cosmologist in Bob's department.

DR. KOONIN: My God!

DR. BEASLEY: That explains a lot.

DR. COLLINS: In credit to me, I quickly abandoned particle cosmology and moved on to something that was a little bit more reputable.

DR. LINDZEN: But, you know, any such statement would be misleading if it were not accompanied by the fact that that alone does not tell you A, B and C, which we need to know for the policy.

And one problem with this issue, and the IPCC statement is difficult. As far as I was concerned, that statement, even if it were true, was not ominous. It said it might be ominous. It might not be. And that was left unclear.

So, if you have a statement, yes, we know there is a lot of CO_2 and it is more than it has been, less than most of the earth's history, et cetera, et cetera, so what have you said?

DR. COLLINS: Well, this is a point where we disagree. I would actually say that you said quite a lot. But this is a point of disagreement.

DR. LINDZEN: It's a lot about science, but the reason anyone is interested in the statement is the policy projection, and it hasn't been that relevant to policy.

DR. COLLINS: Well, I was starting to get partly down the road to addressing your question.

DR. KOONIN: It was a useful bit. But to follow on maybe where we didn't go is therefore, we expect significant perturbations of the climate system in the future and you had better start thinking about adaptation or mitigation, certainly, if not adaptation. That's the Full Monte, so to speak.

DR. LINDZEN: Adaptation is the safer bet.

DR. KOONIN: I would agree. It's also much more likely to happen than mitigation.

DR. LINDZEN: It will happen.

DR. CHRISTY: Of course it will happen, by definition.

DR. KOONIN: I was going to go even further into national policy.

DR. HELD: Can I?

DR. KOONIN: Isaac?

DR. HELD: This is a little bit of a tangent, but what I would like to see in the statement, I don't really care about a motherhood statement on observation systems. I think there really is motherhood, but you can say that if you like.

DR. KOONIN: Your mother is happy when you do!

DR. HELD: From my perspective, as Bill said, I think of myself as a physicist. I haven't changed fields. But for some reason, physicists haven't adopted this problem as a core problem in physics. This is basically a problem in physics.

Everything we have been talking about today, except maybe when we talked about carbon uptake by land. I think I may have mentioned that. That's a little more biology than physics, but a lot of it was physics.

But why hasn't this been adopted as one of the key core problems in physics? And why not have the statement related to education or promoting this in physics departments among graduate students as a problem to focus on?

DR. KOONIN: This is an important problem you have something to contribute with the schools.

DR. HELD: Educationally.

DR. LINDZEN: But Isaac just mentioned the funding situation. You can't hire post-docs. So, whoever we train won't have work.

DR. CURRY: The rationale or the charter for the topical group, they listed the number of areas where they felt that physicists could make a big contribution.

And I think reiterating that in the policy statement would be, you know, these are key issues of uncertainties where the expertise of physics can be brought to bear and the Physics Society is going to adopt this and have sessions, whatever.

DR. HELD: It's fluid dynamics. The whole thing is physics.

DR. KOONIN: Yes, go ahead.

DR. SANTER: I just wanted to point out on that same vein that there was a somewhat similar meeting between the American Statistical Society and a bunch of climate scientists at NCAR a while ago.

And the bottom line was statisticians wanted to know how they could contribute and where the opportunities were. And I think that was extremely useful.

And some good things had come out of that in training more statisticians in the analysis of observational and of model data, model evaluation, detection and attribution.

So, I see this as an opportunity. That's one of the reasons I am here.

DR. KOONIN: Good.

Phil, I know that there were things beyond the framing document you wanted to see us discuss. Have we hit them all?

MR. COYLE: Well, I think we are getting there, yes. One question. There is a view, I think, that it doesn't matter what experts think, that the threat from global warming has become accepted by the general public, by the media and all and internalized by the general public and the media.

And so, what we need to address is what difference can we make given that situation, given the situation that the so-called threat from global warming has been so widely accepted?

What contributions could we make that would help to educate people better, even the APS membership itself, for example, or the general public or the media?

DR. LINDZEN: You could indicate the degree to which there are questions. That would be a phenomenal service. I mean, this Society is not just supposed to be "me, too."

DR. KOONIN: That's why we are having this kind of meeting.

DR. LINDZEN: Yes. I am saying there is a positive use for this.

DR. KOONIN: I want to go off the record.

(Whereupon, an off-the-record discussion was held.)

DR. KOONIN: Back on the record. I will turn to my fellow subcommittee members. What particular lines of discussion do you want to take on?

DR. ROSNER: I asked the question about deep-ocean mixing. I am wondering whether, from the point of view of improving the models and dealing with the data, where else in the modeling do you see worthwhile investments?

For example, necessarily, you

showed the vertical stratification issue. Is there an issue with vertical mixing, for example, that is not well-adjusted?

DR. CHRISTY: In the ocean or the atmosphere?

DR. ROSNER: No, I am talking about the atmosphere.

DR. LINDZEN: The data that you showed, it is convection.

DR. ROSNER: I mean, is it governed by episodic mixing? It is governed by tuning?

DR. LINDZEN: It's clusters.

DR. ROSNER: Bill? Feel free, Bill.

DR. COLLINS: I will give a two-word answer then turn the floor over to Isaac, but vertical velocity, right? We have great measurements in the horizontal.

DR. ROSNER: Yes.

DR. COLLINS: A lot of this is bouncy-driven. We have squat in the vertical. That was more than two words.

DR. ROSNER: How do you calibrate vertical mixing lines? I don't get that.

DR. HELD: Well, for example, we have what are sometimes called process models, very high-resolution models of moist convective turbulence. And those are compared against field programs. DOE supports a lot of this effort.

And so, it's a multistep process. You use those very high-resolution models of field experiments and try to fall back into the global models. It's difficult to do. It's not just global models and trends.

DR. ROSNER: Yes, but there has to be a huge difference. There has got to be a huge difference, mixing, say, about thunderstorms, for example, huge cells that have scales of tens of thousands of feet as opposed to, say, boundary mixing, gravity wave lengths, internal waves and all that.

DR. HELD: I am sure there are field programs focused on each of the topics that you have mentioned.

DR. LINDZEN: Each is separately parametrized.

DR. ROSNER: And separately calibrated?

DR. LINDZEN: Oh, yes.

DR. HELD: Well, they are studied with models of -- all sorts of variety of models. And some of those are directly comparing against field programs designed to test those particular parts of the models.

DR. ROSNER: Do you guys feel confident that you know what you are doing?

DR. LINDZEN: No. Can I give you an example that is innocuous? There is a phenomenon that I was involved in many years ago called the quasibiennial oscillation. You have the wind going from one direction one year, another direction the other year back and forth.

And in the late '60s, early '70s, it was recognized that this was essentially waves interacting with the flow causing the wind to change and descend. Fine.

Almost no model comes close to showing this phenomenon. And it's understood the models don't represent the equatorial gravity waves and smaller gravity waves.

So, increasingly models now make models a flux of gravity wave that they suppose sometimes is related to other things and tune it so that a QBO emerges. This is not a terribly satisfying thing, but they are not going to resolve the waves.

What is bothersome to me about it is, if you do this, there are still things you can get. So, one of the things with the model I have been looking at is, when we look at these tropical waves today, we look at the infrared space data and we see them as cold patterns.

When one of the models that gets a QBO tries to do this with its outgoing longwave, they don't see them. That immediately allows you to see something wrong with that model, maybe not other models.

But this degree of interaction and understanding with the implications is not widespread.

DR. KOONIN: Go ahead. You want to follow up?

DR. ROSNER: So, given that, doesn't it bother you that this level of misunderstanding or not understanding, if I were doing this in what I do, I would be --

DR. HELD: It's turbulence. It's a hard problem.

DR. ROSNER: It's not just turbulence. It's beyond turbulence, right? There is no turbulence.

DR. LINDZEN: You have a variety of things going on.

DR. ROSNER: It's complicated.

DR. SANTER: Again, what Isaac's work has shown is that if you give at least the GFEL model, the observed changes in ocean surface temperature, it does not produce that error structure in the way that John showed. It actually is much closer to the estimated, observed changes. That tells us something useful there.

Another things is this time, scale and variance issue that, if you look at amplification of surface temperature changes in the deep tropics on monthly, on annual, on El Niño time scales, models and observations are not in fundamental disagreement.

So, one aspect of the physics

is time scale invariance in a big way. It's on those long decadal time scales where the observation results were most sensitive to the adjustments that you make with things like orbital drift.

DR. ROSNER: So, here is the thing that struck me. So John, during your discussion, the way I read your talk was that you were struck by the fact that the band of models was way off from what the data was.

What struck me was something else, which was the huge dispersion among the models. And where does that come from?

DR. KOONIN: Ben?

DR. SANTER: Ocean.

DR. KOONIN: The ocean?

DR. SANTER: Ocean, but another thing. What John showed was the mid to upper troposphere in the tropics. That has a nontrivial contribution from the cooling stratosphere. And actually the model --

DR. CHRISTY: No, he is talking about the radiosonde one.

DR. ROSNER: Right, the radiosonde, yes.

DR. CHRISTY: Level by level.

DR. SANTER: I thought you meant the band of changes that John was showing for 102 models and the observation being completely outside.

DR. ROSNER: The height and then temperature brought in horizontally. And there is this band of solutions that went sort of like this (indicating).

DR. CHRISTY: Ben was right when he said the ocean, because if you looked at the surface --

DR. ROSNER: All the different points.

DR. CHRISTY: -- they had the spread there.

DR. HELD: They are all --

DR. ROSNER: They are all normalized in totally different ways, really?

DR. HELD: I didn't say "normalized." It's because clouds are giving you different -- are changing in different ways in the different models. They are causing the tropical ocean to warm in different ways. It's influencing the tropical atmospheric profile. That's the picture, the zeroth order picture.

DR. KOONIN: Let me go in a slightly different direction. One thing that physicists care about and some of the people sitting at this end of the table care a lot about is advancing high-performance computing.

To what extent would Exa-scale capability improve what one can do in science?

DR. HELD: This is something a lot of us have thought about. I would say as long as it's not monolithic in the sense that it's a small number of people controlling that facility, but it's available to be used in novel ways by the climate community as a whole, I think it would be fabulous.

DR. KOONIN: You said “No, it's not as though, my gosh, I've got this model. If I can just get a factor of 100 more computing power then a breakthrough?” That's not the case?

DR. LINDZEN: Where does the Japanese program stand on this?

DR. HELD: There is an example of the earth simulator. If I look at the science there, it looks very promising but, if I were to ask has it revolutionized anything, the answer is no.

The Japanese community is wonderful but it's small. They have a certain number of things they are interested in.

But if you increase the computer capacity of the field as a whole and it's open to new ideas and younger people, then I think you will get something.

DR. KOONIN: Capacity?

DR. HELD: Yes, okay, that's the word.

DR. KOONIN: So, Ben was first and then you, Judy.

DR. SANTER: I will let Judy go.

DR. CURRY: A couple of things to advise you. First, you have the potential for a much larger ensemble size. You have the potential for a much larger ensemble size rather than this ad hoc ensemble of opportunity.

The other one, cranking down on the horizontal resolution is important to get the natural internal variability right and to get the blocking patterns. And you are never going to get extreme weather events from, of course, resolution models.

So, you can explicitly get some of those extremes. And cranking down the resolution is particularly critical for the ocean because the resolution that we are doing at the ocean right now is extremely crude.

DR. KOONIN: What is it now, sixth of a degree?

DR. CURRY: Yes, but given relative to the Rossby radius --

DR. ROSNER: But you have the data to do the calibration?

DR. CURRY: So, it's not the answer to everything, but it's an answer to some things in terms of really seeing what kind of information we can extract from this type of model that we really need a bigger ensemble and higher resolution before we can feel like we really explored this path that we have been on.

Sorry, Ben.

DR. SANTER: No, I agree. I want a larger ensemble. I want more systematic exploration of forcing uncertainty. To me, it seems like one of the issues here is that every modeling group wants to put their best foot forward in IPCC. They want to have the best possible physical model of the climate system.

I think much less attention is devoted to the construction of forcing data sets, both natural and anthropogenic. They come in kind of at the end in the process of performing simulations for IPCC. To me, that's where the scientific understanding comes.

It's not sufficient, again, just to show some discrepancy between models and observations and say models are wrong. We need to understand why those differences exist.

And in order to get that understanding, if I were king for a day, I would use that computational power to more systematically explore forcing.

DR. CURRY: Thank you. That's actually very, very important.

DR. SANTER: And increase model sizes as well. Some of these simulations that I mentioned that people are performing now with more realistic representation of 21st-century volcanic aerosols, they have got ensemble sizes of five.

This is a relative weak forcing. In order to better estimate of signal and beat down the noise, we need larger ensembles.

Finally, what I would do is I would go after the seasonal stuff, as Isaac showed for the hiatus. All the detection and attribution work essentially looks either at decadal mean changes or it looks at an annual changes.

What you lose, then, is the effects of different forcings on the seasonal cycle. For ozone, that's profound. You look at stratospheric ozone depletion and its impacts on the lower stratospheric other Antarctica, it's huge. You get this huge signal in October or November that is clearly beyond anything that you can generate with noise alone.

Now, many of these radiative forcings that we have been talking about like, say, biomass burning, fires up in the Congo and in the Amazon at certain times of year, very, very specific regional and seasonal signatures.

We need to look at that kind of thing, in my opinion, in detection and attribution work in order to better discriminate between different anthropogenic forcings. You lose that seasonal specificity when you do detection and attribution work with decadal mean changes.

DR. KOONIN: John?

DR. CHRISTY: The only thing, Dick showed a picture of four models run in this experimental mode. And this is a really neat experiment where the authors had runs from a water-earth, very simple earth, current temperature of water, warm it up four degrees. How does the model respond?

So, that kind of fundamental test could be done so that you could see the dispersion of how the models create clouds and radiation, how different they are, and perhaps come up with a better way to understand why they are different, what could be done to better characterize that process, that kind of experiment, a fundamental experiment that would enlighten us about how these models.

DR. KOONIN: Judy?

DR. CURRY: Yes, the point I want to make is basically the same, but I will reemphasize it. Because of the cost of running these big models for the CMIP5 and arguably the IPCC production runs, there is no room left over for the creative, imaginative experiments to really test understanding.

And again, you need large ensembles, very long runs, whatever, sensitivity to a variety of things. Forcing, I agree, is very important.

There is not enough horsepower left over to do these things. And we are selling ourselves short by not being able to do that.

DR. KOONIN: When you say "horsepower," both cycles, but also people?

DR. CURRY: Cycles and people.

DR. CHRISTY: Expensive.

DR. CURRY: Cycles and people.

DR. LINDZEN: I wonder

sometimes if it's an excuse. If you suggest something to a modeling group, one of the convenient answers is, "We would love to do it, but we are doing CMIP projects."

DR. KOONIN: Ultimately, it boils down to, if you will excuse me, program direction and what the funders try to nudge the system to do.

DR. COLLINS: I think one of the things that could be done with such a capability -- sorry, I have been wrestling with flight itineraries here which is why I keep running out of the room -- but exploration of uncertainties that we heard about.

For example, Ben discussed systematic exploration of uncertainties, but perhaps not with his simple models of, but models of the ilk we have been using for the

IPCC.

That is another thing one can do with this capability that would be extremely fruitful, a systematic exploration of parametric uncertainty and forcing uncertainties so really, we can construct an error budget for a climate model.

And I think that would be another constructive use. Besides, the natural tendency, I think, would be to add a lot more complex physical processes and to take them out into very high resolution.

And I think one could argue that a complementary activity that is sort of saying let's assess what we have got. We want to understand the foundations for further development. And also do hypothesis testing, again, in sort of an exploratory mode with this capability would be extremely useful.

And increasingly, in at least,

if I could speak for the United States, the luxury of having both of people and the computing power to do that has become harder. But it would be a really constructive use of the cycles.

DR. ROSNER: Would you agree with the following statement, that increasing the fidelity of models without a corresponding increase in the data collection capabilities is a waste of time?

DR. COLLINS: No, for the following reason. No, in the short term. This is a time scale question.

One of the difficulties we have had with climate modeling is that, at the moment, the models are being run at length scales where we have to construct effective and often less-than-ideal empirical theories about how things work.

We actually understand how things work at smaller scales. And this gets back with to some of the topics that Isaac was raising earlier. We raise observations at these small scales.

The moment there is a big enough gap that we have to fill in the middle with sort of, it's physical theory, statistical physical mechanics, which is often a very fraught exercise, driving the models down to the native skill and observing networks and the process models, which we will be able to do soon, would be extremely useful.

Because at that point, we will be able to essentially test the models deterministically against observation, against observational networks. The climate community should stop throwing up its hands and saying, "We do climate, weather." We should perhaps do both.

So, it's an initial value-driven problem and a boundary-value driven potential.

So, I actually would argue that the exponentially increasing computing power for the time being is buying us something in the sense that it's going to hold the climate modeling community's feet to the fire, I hope, further.

DR. ROSNER: So, you are saying there is still a lot of space between the grid resolutions for models and the resolution at which you do data sampling; is that right?

DR. COLLINS: Well, you have seen satellite observations that are conducted globally, but some of the pertinent aircraft observations are made at small scales, et cetera. It's not clear to me that going necessarily to the end result, extremely small scales, is the relevant issue.

But what is relevant is the ability to test the model in a way where we can, for example, say, let me run the model with observed initial meteorological conditions and ask which process fails, or maybe you can get a convolution of both large-scale atmospheric state and the process. And that's currently the problem that we face.

So, I would say in the long-term they need to advance commensurately, because the answers will not come out of Silicon.

DR. ROSNER: Right, exactly, yes. That's why I'm asking.

DR. SANTER: Just to follow up from what Bill said, some of that work is going on. So, at PCMDI and elsewhere, modeling groups are running the GFDL model, the NCAR model in weather forecast mode assimilating observations, making forecasts comparing, say, what the high temporal resolution ARM measurements and learning a lot of useful things about errors in certain model parameterizations that manifest very quickly and then propagate into climate time scales.

And that has been extremely useful, I would argue, in trying to really put your finger on causes of differences between models and observations for some aspects of these simulations.

DR. LINDZEN: What about mesoscale modeling efforts? They also have very limited success. They have extremely high resolution, but they are a small phenomenon.

DR. COLLINS: It's not a panacea, I completely agree.

DR. LINDZEN: Pardon me?

DR. COLLINS: It's not a panacea. Resolution is not a panacea.

DR. ROSNER: You guys will be in business for a long time.

DR. LINDZEN: The best way to avoid it is not to depend on the weather.

DR. COLLINS: I would like to get back to a point that Isaac raised about the engagement, sort of, perhaps a part of the statement that could address the engagement of the physics community on this problem.

And I know from when I was a graduate student at the University of Chicago, the physics community benefitted tremendously from the influx of -- well, applied mathematicians are very interested in the chaos problem.

One could point to similar examples involving general relativity and the work of Roger Penrose that kind of transformed general relativity in the 1970s. So, this kind of crosstalk can be tremendously beneficial.

And my sense, to be honest with you, is that, and I think this all makes us a little bit nervous, climate is not a problem that is amenable necessarily to reductionist treatment.

That's a problem. And there are aspects in which it's messy and it's hard to do simple -- some of the simple, low-hanging fruit is also gone.

And so, there are ways in which this does not look appealing. But it's a really important problem.

And I think we would benefit tremendously from engagement of people who want to think critically about how to do the error right, the measurement right and the modeling right.

DR. KOONIN: I think you just said “it's a mature, messy problem.”

DR. COLLINS: But there are also examples of physicists getting deeply involved in the life sciences.

DR. KOONIN: Bob?

DR. JAFFE: This is, I guess, a follow-up your question. There was a discussion about horsepower in terms of flops and also people. I wonder, what is your workforce problem like? Where did your graduate students come from?

Are there graduate students flocking to your door or do they come from physics? Do they come from earth sciences? Do they come from oceanography?

What is that structure like and do you need the recommendations as this is a field which needs workforce development?

DR. COLLINS: Well, so, our graduate students do come from the physics community. Currently in my department, we have three former string theorists as graduate students.

DR. JAFFE: They don't count.

DR. COLLINS: Is that on the record?

DR. JAFFE: I am afraid it is.

DR. BEASLEY: Following your inspiration.

DR. KOONIN: Dr. Jaffe, you will have the opportunity to clarify your response.

DR. COLLINS: Sorry, string theorists, physicists, applied mathematicians, civil engineers, those are several different departments from which I have drawn personally and my department has drawn recently.

I actually think, I'm not sure if the attraction of this field because it's a hot topic is necessarily the issue. But somehow I think the problem is furthering along people's careers, right?

So, the issue is how does one get -- this problem looks messy. It's non-reductionist. How do you do the right thing to get tenure in a physics department doing a problem in climate?

That's the reason why I am

thinking this is a problem a little bit downstream. You see what I am saying?

DR. LINDZEN: If I could make a suggestion. Now, I think the business of reductionism is extremely important and appealing. One problem with the current, quote, practical climate problem, greenhouse gases and so on, is it has drained the energy from phenomenology.

It would be terrific to have students understand the Eocene, to work on the glaciation cycles. There are plenty of well-defined problems in climate. Why did the cycle of glaciation begin about 700,000 years ago?

These are, in a way, traditional problems, almost 19th century, and they are exciting. And the oxygen has been drained from them by the environmental issue.

DR. KOONIN: Plus you have

modern modeling tools that you can --

DR. LINDZEN: Well, you have that, but thought is --

DR. KOONIN: Think before you compute!

DR. BEASLEY: I don't want to get too school mom-ish here. But as a condensed matter physicist, we thrive on phenomenology. So, it's not the existence of phenomenology that is not attractive to students.

To throw it back to you all rhetorically, what is needed is a clear statement of what are the fundamental problems or what are the interesting outcomes that all of this could lead to?

And I know you are busy and you have got all this. But I think that's part of the problem. Because if you don't, if you don't get that into the students' minds, then they will stay close to home.

But if they see is that

excitement, they will go out and go to a mechanical engineering department. At least they will at Stanford.

DR. KOONIN: One thing we haven't talked about, we are the American Physical Society, although there is a big international component. How does the U.S. stack up in the science relative to EU, China, Japanese?

Are we doing enough and do we understand enough to be able to hold our own in international discussions of climate issues?

DR. LINDZEN: Alas, yes.

DR. KOONIN: Okay.

DR. CHRISTY: In many of these observational data sets, we are driving the bus. We are kind of the ones that started the whole satellite movement and many of the other networks.

DR. BEASLEY: Will that be true

in ten years?

DR. CHRISTY: I don't know. I really don't. It doesn't look promising right now.

DR. COLLINS: There was an issue I think back in the early '90s that dealt with this concerning climate modeling. But it is true that the U.S. now has multiple, very strong climate modeling efforts.

And the U.S. has actually maintained a strength in diversity in quite deliberately in this area and has, as John said, been really a leader along with the EU in building satellites to look at, to examine the earth system.

But certainly, NASA right now is -- I was just at NASA. I was talking to them about their upcoming decadal survey and observations. And there is a real risk to next generation of satellites. That is a very, very concrete risk.

Setting aside our posture within the international community, just setting a fairly high bar for ourselves, I think we are at risk of not grasping the bar in the next decade because of the risks to, in particular, the satellite systems.

So, the EU is drawing strength in doing federated intercomparisons in a way that we do not do in the United States. I am thinking of Prudence and Ensembles, for example, these huge intercomparisons they do.

I think there are differences, but I'm not sure if they are leading to qualitative or dramatically different outcomes in terms of scientific quality.

DR. KOONIN: Anybody else? We reached --

DR. COLLINS: The asymptote.

DR. SANTER: Hiatus.

DR. COLLINS: The hiatus.

DR. KOONIN: Maybe we can just

kind of close out by final remarks, last shots?

DR. BEASLEY: Well, on behalf of APS, I really want to thank you all. Well done. I don't know whether to give Steve credit or you all, but there was more discipline in addressing the questions posed than I have been able to manage in my own field. So, thank you very much.

DR. KOONIN: So yes, of course, for me, too. But I still want to give people an opportunity. I can summarize. Maybe I will start with that.

You know, at the same time in some dimensions there is more confidence, greater certainty in some of these issues, but in other dimensions, more uncertainty.

The uncertainty in the forcings, which almost from the beginning of the day became a theme is something that I am now educated about and more concerned about. It makes all of this just a bit shakier than it was for me to start. That's something I wanted to say.

DR. LINDZEN: I think there is one field that was omitted here. And I was reminded of it by your statement what we are confident on, which is geochemistry. There are plenty of gaps in our understanding of carbon dioxide budget.

And that, of course, enters into the forcing issue, but also into all sorts of attribution.

How should I put it, the one thing I feel and I think that you don't want to use the word "incontrovertible" unless you know what you are talking about.

DR. ROSNER: That was an early recognition.

DR. KOONIN: Yes.

DR. BEASLEY: A well-analyzed problem.

DR. KOONIN: Sue?

DR. SEESTROM: So, something I didn't hear you pose in your set of statements that people might agree on, but I think could be useful, comes out of the interaction between the climate models and the natural multidecadal oscillations, is the fact that there is complexity there that makes it hard for the models to be predictive on one- or two-decade time scales, because I think for people who haven't studied this as a newcomer, the fact that you hear a lot about this hiatus.

And it seems to me the hiatus has a high probability of being able to be described by interactions with these natural oscillations, just pulling that out and telling it to the membership I think would be useful.

DR. KOONIN: All right. So, I will offer my thanks to all of you for, I thought, a really good day and good discussion of the science, very productive, collegial, and thanks.

And I hope the world will review what we did and it will be beneficial.

(Whereupon, at 3:49 P.M., the workshop concluded.)