What is a climate model?

Hello and welcome to a new episode of Larscience. Thanks for all of you who gave me feedback on the first two: I've tried to make sure I've addressed all of your points – I've written a script for starters, so hopefully that will help.

This one is going to be about a fundamental concept, but also a question you don't hear answered very often in my field: what's a climate model? You hear them talked about all the time, but rarely defined. So I'm going to try and do that, without turning you off with a whole bunch of maths.

***

I don't normally like to do this, but I think I'm going to have to start with a quote, because it's just so great. The statistician George Box is credited with first making this observation that underpins all modelling work.

He said in 1976: "all models are wrong, but some are useful".

You can interpret that how you will, but what I take from it is this. You can never represent reality perfectly, unless you create an experiment that is exactly as complex as the world we live in (and, we are nowhere near understanding our world completely, so that puts a spanner in the works somewhat). Clearly, the limitations of our knowledge and abilities preclude us doing that, and I'm not sure we'd even want to create such a model.

So, whatever we do to simplify reality into a model, we're always going to be smoothing over some kind of subtlety. That's what makes them "wrong". However, as long as we are aware of these limitations, they can still be immensely useful.

In fact, we can use these simplifications to our advantage, to test the relative importance of certain processes and factors away from the confusion and noise of the real world, which we don't understand perfectly.

Ok, so so far I've jumped straight in to talking about models without defining them. Terrible scientific practice: I apologise. The concept of a model is a bit confusing – it can be many things. You can have conceptual models, statistical models, numerical models, climate models… all sorts. What they all have in common though is that they represent the world, or some element of the world, in a simplified form.

In the context of my work, I will primarily be referring to numerical weather prediction and climate models when I talk about 'models', because those are what I use. In fact, the model I use for my PhD is both of those – a climate and numerical weather prediction model – in one. More on that later.

Models are scary if you think of them as black boxes that do mysterious stuff with data and spit out a garbled mess of unintelligible numbers. That's what I thought models did when I first encountered them. But, nowadays I try not to be intimidated by them. I like to think of models as giant calculators. That's what they are - they represent the world with maths, but amazingly, they only use a few equations to simulate the whole Earth system.

You might be surprised that you'd probably recognise those equations from school science lessons, at least if you were paying attention. They describe the fundamental laws of physics – like the conservation of energy, which states that no energy can be created nor destroyed, or the Navier Stokes equations, which describe how fluids (like air) move.

The reason that models are so complex, and have to be solved on extremely beefy supercomputers, is because Earth is pretty huge, and the model has to calculate each variable you're interested in at each point in space, and *then* figure out how to evolve it over time. Equations that calculate the rate of change with time are called differential equations, and this is primarily what is going on behind the scenes in a model simulation.

To do this, the model setup divides the planet up into small, bitesize chunks of space in a grid. It then performs calculations in each square, or gridbox. Of course, the world is 3D, so there are gridboxes in the vertical direction too – this all combines to make many, many 3D gridboxes.

You can see how it all gets quite complicated very quickly. It gets even more complicated when you add all the components of the system in. There are many different elements of the Earth or climate system, such as the atmosphere, ocean, ice, vegetation etc. Sometimes only one of these will be simulated, and other times they will be connected together, in what is called a 'coupled' model. These are more realistic, but often take a lot longer to run because the calculations are so complicated!

In my work, I use an atmosphere-only configuration of the Met Office's Unified Model, with prescribed values for things like sea ice concentrations that are important for driving the development of the atmospheric variables I'm interested in, like temperature and humidity. What I mean by prescribed is that I have one set of input data, that doesn't change, that the model uses to calculate other things, rather than the model generating those numbers itself, which would take a lot longer. It's not perfect, but it's functional.

I said I'd come back to that, didn't I? The Unified Model, or UM to you and me, is the UK Met Office's operational forecast model that they use to generate their weather forecasts. However, they use an approach called 'seamless modelling', which means that they use the same model to predicting the weather, and to project and explore the climate. That's what makes the UM a numerical weather prediction model as well as a climate model.

I don't want to get too in-depth at this stage because it can get quite involved quite quickly, but I just wanted to touch on one other thing. It's all very well and good having a model that accurately represents the system you're interested in, but a model also relies on having good quality data to drive it.

A model requires input data to provide the starting conditions. Basically, you have to tell it what all of those variables were doing at the beginning of the time period you're looking at (which is called the 'run') so that it can perform those calculations on the data.

So, if you're putting rubbish data in, you're going to get rubbish data out.

Scientists can sometimes get quite protectionist about their disciplines. You know – modellers only working with other modellers and experimentalists only working with other experimentalists. However, I think this demonstrates perfectly the need to work together across disciplines. Modellers like me need to understand the need for good quality observational data, and how that kind of data is obtained, and observationalists need to understand how other people use their data, and therefore what's most useful.

There's a whole infrastructure around climate and weather prediction, and models require vast data assimilation networks, which pull in many types of data from across Earth and space and 'assimilate' it, or bring it together, into a complete package that can be used to drive models, amongst other things. That includes satellite observations, ground-based station data, aircraft measurements, radar, lidar, you name it – whatever is available will be tested, filtered for quality and compiled.

On that note, I think I'll end – models are an important tool, but they are only as good as what you use to set them off, and we need to be aware of what they can't do as much as what they can do. So next time you hear someone saying that model data definitively shows something – remember to take it with a pinch of salt, because there's still a lot we don't know.

See you next time.