Virtual CP4C Workshop Notes
April 22, 2022 (2nd Day)

Panel discussion 1- Status and development plans for CanESM

Facilitator: Paul Kushner
Time: 13:10 pm
Panelists:

- Vivek Arora
- Jason Cole
- Yanping Li
- Paul Myers
- Ivy Tan
- Neil Swart

Main Takeaways:

Jason:

- Take a piece of the parametrization; IMSI gives some opportunity to revive climate models

Ivy Tan:

- The single Column version model will allow for good testing of ice microphysics, aerosol-cloud interactions, etc.
- Suggestion: develop radiation kernel for CanESM, which leads to many applications, e.g. diagnostics on radiation
- Some items are low-hanging food, some items are not readily available (e.g. cloud model), may take 3 years to make available
- With P3, something just started, there is work that can be done concurrently, and green's function is something more readily available.

Jason:

- IMSI gives the opportunity to use new configurations. There is an interest in implementing AQUAPLANET
- Climate intervention work, so would like easier use/development of middle atmosphere model, chemistry, aerosol
- That is something the community can easily run; it is just a single setting, and people can do a lot of analysis

Ivy Tan:

- Clouds and climate sensitivity. No aerosol effects on cloud microphysics. Single column version of the model would be a good tool for testing the implementation of new physics parameterizations (ice clouds, microphysics, etc.). New aerosol sources can be improved
- Green's functions work to explore pattern effects.
- Offline radiative transfer model can be used to develop radiative kernels
- Understand contributions of different radiation bands to the climate feedbacks nudged experiments useful for sensitivity tests (constrain meteorological variables for easier comparison to observations, intermodal comparisons

What are the constraints to doing some of that work? What are the items that will be done no matter what, and what requires additional resources?

Ivy:

- Radiative kernel is low hanging fruit re: potential applications since it is close to implementation
- SC version is not going to be easy to do right away (not readily available).

Jason:

- Something that is being explored but not close to being complete (needs GEM to be ready(?))

Paul:

- To identify ideas that can be done based on current work vs those that would require more investment to complete but would be more novel.

Ivy:

- Ice nucleation and cloud things can be done simultaneously. Green's functions developed for a MIP

Jason:

- Green's function can be easily taken up by the external community since it is not a special configuration, it just needs a lot of simulations

Ivy:

- Many low hanging fruits, but the P3 scheme will be the most strenuous task and cloud microphysics.

Paul:

- 3-year time frame is informal, is the typical time frame for CC allocation. No set funding, modest funding, thus focus on the technical side. It's a long-term goal

Neil:

- Need to train people to use the model before expanding the development of more specific parts (microphysics).

Philip Austin/Paul:

- Opening the model is important for graduate student training, allowing them to get experience working with the model

Land Model:

Yanping:

- Current working with the US land surface model (NOAA MP)
- Land surface models are used a lot to help future projections.
- NOAA MP does not include important processes for Canada, rather than the US (e.g. wheat crop type)
- Looking toward improving the Canadian land model with ideas from other land models.
- Hoping to be able to contribute to improvements to CLASSIC with experience with NOAA MP
- Other research directions: wetland representation, Groundwater scheme
- Current NOAA MP lacks representation of permafrost, boreal wetland; something that exists in CLASSIC
- Need to learn from different models.
- We develop surface models for different uses. We don't need to re-invent the wheels. In the future, we can adapt more to the Canadian land surface model. Models have limitations. We should see how the 3 models can be combined to do certain things. model developments, learn from different models and hope our work can contribute.

Vivek

- CLASSIC has simple crop representation because there is no carbon sequestration thus not a focus relative to the forest, other vegetation
- Wetlands are dynamic in CLASSIC - wetlands are offline (?) or the emissions?
- Not much progress on atmosphere-land coupling work in CLASSIC, mainly offline land modelling
- Welcoming all types of collaboration/contribution
- Info from university community re: CLASSIC performance will be taken as feedback by ECCC development, welcoming changes to the model
- Wetland emissions (ammonia etc.) are offline but hope to one day make it online in the coupled model
- Knut asked to move dust out of the interactive model to the offline model.
- To do: bring dust into CanESM5, the interaction of methane, nitrogen emissions from land to be used for aerosol chem

- Crops get harvested every year. Crops do not store a lot of carbon, so it is not reflected in the global mode well. We do not have any representation of groundwater and radiation. They are offline.
- They do have effects on global temperature. We have a good connection with the offline community; we welcome all collaborations.
- You are welcome to use CLASSIC on the crop side. We will receive feedback and make improvements. We will do control runs to see if it breaks anything and incorporate those improvements.
- Incorporating aerosol coupling effects into CanESM, we haven't implemented yet. We can do the offline analysis first and then perform the analysis with CanESM when the network is ready. To do: bringing dust into CanESM.

John Scinocca:

- Online, offline division is important; if you are thinking about helping with a phys param, this feedback onto the prognostic variables. Need to consider if changes to params are more advanced in terms of representation vs the current model and thus might not be a priority.
- Need communication between CCCMA and the university community to ensure priorities are right/coordinated
- We need to be careful; would it be meaningful to implement something higher-order when some important physical processes (such as ice physics) have not been implemented yet? We need to prioritize the tasks.

Paul:

- Need to establish the priorities and the current projects going on.

Ocean Model

Paul M:

- Coupling of Nemo and biogeochem.
- Large NEMO user community can do evaluations that can be fed back into CanESM.
- Operational prediction work is being done.
- Downscaling of NEMO around the coast of Canada. A lot of work on the NEMO system is on top of the CanESM work.
- Want to collaborate between CANESM and NEMO communities
- Biogeochemical work linking observations to modelling work
- Sea ice locking, iceberg work.
- NEOPAR (?) is working on upgrading to NEMO v4
- Work on process representation in NEMO, hopefully, can be useful in the coupled model system.
- Hope to see collaboration with the ocean community and see how those changes impact the climate in the coupled system.

Neil:

- Why use CanESM instead of CESM? A lot of value in bringing model diversity to CMIP exercises. Most modelling groups don't care about the Canadian climate in particular. Downscaling activities over Canada in CanESM bring huge value.
- Need to emphasize the value of using CanESM vs other models
- The more work is done, the more EC will take an interest in collaborative work with univ community
- The Ocean group has been relatively small and never developed its ocean model for Canadian models. Currently, a gap in the ocean modelling group
- The university community can help support the weak points at CCCma. Opening CanESM allows more ppl to be trained on the model, making it easier to recruit into CCCma
- Coupled downscaling is important to capture full effect and would bring additional value.
- A lot of other groups do not care about the regions of Canada. We have a mandate to look at regions of Canada specifically. Currently, there is a gap in the ocean group. There is no talk on the ocean group in this workshop. What should we do to successfully guide the users to use the model, not just for specific physics but generally as a coupled process?
- The coupled system will give a lot of values

Kirsten:

- CanESM is a great tool with the new land carbon representation. Improving CanESM is important for evaluating carbon-climate feedback and its relevance for Canada. Funding is ELEPHANT in the room. The limited ability of ECCC to fund
- Usage of CanESM is limited due to a lack of funding.
- An important use for Canada is understanding the role of carbon removal. How can the community come together and make a case for funding?

Paul:

- Have to show the need for a Canadian model and the use of a Canadian model.

James Christian:

- In Canada, we don't have a large group of ocean modelling (large scale/global)
- Different priorities in processes to represent CanESM vs regional model.
- Different academic groups are doing regional modelling. The parameters being prioritised for the regional and international models are very other. Hopefully CanESM motion model will be more accessible; as the HPC pipelines go online, people will start using it.
- Not that many people are being trained in global ocean modelling

PAUL K.:

- Might help to have more communication about global model processes
- At the current resolution of CanESM, there is still priority focused on bias correction, and hopefully will help

PANEL 2: CP4C - Science Applications and Collaborative Environment

Facilitator: Neil Swart

Time: 13:10 pm

Panelists:

- Joe Melton
- Andrew Shao
- Hansi Singh
- Clint Seinen
- Jack Virgin
- John Scinocca

Main Takeaways:

What types of policies and guidelines should be established to guide CP4C?

Andrew Shao:

- Crucial to identify users who are interested in full pipelines vs users who are just interested in one part.

Clint Seinen:

- There is a need for a testing pipeline to make sure people can see if something is going as expected or for evaluation. In the end, it comes to reproducibility; we are working on it.
- Reproducibility needs to be at the core of what generates some policies.

Joe Melton:

- 2 very different user groups (government - long time frames, low turnover vs university - short time frames set by funding) need to have clear guidelines for how to share work, ensure work is taken up by the rest of the modelling community
- There are many users, e.g. universities which have some funding in the short term for some projects. Need to establish some ground rules for both sides to benefit from this collaboration.

Andrew Shao:

- NEMO consortium had a similar problem. In practice, we need to find an inside advocate to make changes to the model. Another method is to take in all developments, which can introduce problems in the model. There is a need to make clear the process/tests for inclusion developments into CanESM. Unspoken and soft rules are the hardest to enforce.
- Welcoming all developments also brings in management challenges. Need to set rules like what you need to demonstrate to be incorporated.

Hansi:

- Experience with CESM - This model goes forward through working groups (each has a university person and a NCAR person). Brings development and scientific ideas together at the same time. Allows for the great collaboration done there.
- Version control is important because smaller branches of the model need a more direct method for involving the university community.
- CESM - once you get it running, you can do anything you want; very decentralised, but CanESM will be different, so you need to make sure there are good connections to EC.
- Developments will go through working groups. Some developments can be a branch and use version control to manage. This may get university users to get involved. When it comes to a model that that is difficult to know how to run it.
- It is important to get researchers in universities to involve.

John Virgin:

- One of the most exciting things is to have the possibility to be able to update CanESM locally by yourself. Having enough documentation to do day-to-day tasks without requiring back and forth with ECCC people.
- Possibility to make CanESM more autonomously. Infrastructure to support users, a lot of stress on documentation, would be great..

Neil:

- That is in line with what we do. Currently not perfect, but with the resources, we can improve that. A forum would be great to support users.

John Scinocca:

- Make a distinction between the application and development of CanESM. Development is complicated with testing and legacy issues. But application provides a lot of opportunities that should be relatively free.
- Starting with applications is easier regarding policies and can build into development feedback as CP4C develops.
- Important to distinct applications and developments. Developments involve a lot of work. Applications, there are a lot of interesting projects and experiments. Should get people comfortable running the models and doing interesting science as a starting point.
- We can then document how things come back upstream. We do not need to solve everything right now.

Michael Lazare:

- Went through this about 30 years ago and split the work in 6 groups. There was one support persona (Michael) who was outside of CCCMA.

We do not have a lot of people to support users. How do we prioritize which groups to support first?
Neil:

- NCAR is much larger and has an explicit mandate to support external work.
- Scientist to scientist collaboration is probably a better model within the capacity of CCCMA.
- CCCma is a small group compared to the size of users. Setting up an IMSI working group may be a solution.

Joe

- New tools allow more capacity for supporting external developers. Documentation is crucial for avoiding the developer's time being occupied with support work.
- People may leave the group, and documentation is the thing left.

Neil:

- We can automate to make people just click a button to run the model. But if they want to do slightly more complicated things, documentation or a recipe would be important.

Paul K.

- RPP - need a management strategy for the project. Going to establish a management team
- The management strategy was part of the proposed work, collaboration with workshop organisers, etc.
- need more PIs to help with management, maybe do rotating team, administrative work ahead, but should be doable with current work
- CanESM working groups, just like the CESM working groups, would be a great thing.
- We do need people. Would like to see the rotating team, given that resources are available. People train researchers over 10 years period. The research cycle does not change much. We have a lot of things in common in terms of research cycles and structures.

Neil:

- CanESM application/running is made easier by full diagnostics pipeline, including CMORizing data. Is it perhaps almost too easy? How do we ensure data from simulations is properly vetted - needs attribution etc. CMOR data looks official - how do we make sure ppl know what to trust?
- Should there be a policy on how CMOSized models are used for publishing?

Andrew:

- If people CAN do whatever they want, should they be allowed to? Should set up ways to authenticate data using, e.g. encrypted keys. Some way to "mint" data is as being officially from CCCma.
- Put authentication key to meta file to confirm the runs come from the group.

Joe:

- Mostly, people just run the model; they would eventually link back to the experts if they need to drill into the model.

Paul K:

- Our plans include identifying users, having them agree to a simple code of conduct that includes something about intellectual property, setting up a database of simulations as they are produced for provenance and documentation. This could include authentication information, which is a great idea.

Joe Melton:

- Few people are doing really whatever they want with models without connecting back into the CanESM people. You get connected back during the review process.

John S:

- Wary of open use of CanESM. Currently, CanESM data in CMIP is from CCCma. But what if ppl can contribute to CMIP w/ CanESM that isn't verified by CCCma. Don't want to restrict what science can be done too much.
- It is difficult to draw a line to restrict what people can do and not do about science.

Neil:

- Generated large parameter ensembles for tuning and testing. But there was debate, should the parameter ensemble be released? Some ppl put the data online, but someone might not know that the data is not the "correct" version. Need to ensure there is clear documentation of the simulations, avoid mistakes/misinterpretations of available data
- If we have guidance on how to process the data, how to validate the data, it would be great

Hansi:

- CESM deals with this issue with specific scientifically validated configurations. Scientific validation would be an easy way to deal with this.
- There is a small pipeline of people who have knowledge of HPC, slurmp.

Clint:

- HPC is a major barrier to entry which means that users will have some measure of expertise. Ensure information for reproducibility is readily available.

Neil:

- CP4C DECK simulations are a good example, will be validated etc
- We can provide guidance (e.g. diagnostics produced on Niagara); if they get very different results, they know something is wrong.

Michael:

- If porting the models to different systems, eventually, some bits would be lost. It would be difficult to define what discrepancy is expected.

James C:

- Provide users with a set of benchmarks. Platform to upload results of benchmarks.

Vivek:

- Code of conduct and a licensing agreement address many concerns raised here.

Joe M:

- External experts to help with the funding could help generate the right type of interest in the proposals.
- A lot of things can be learnt from external experts & government organisation.

Clint:

- We can use the tree structures to manage relationships with universities. The PI structure can fall off easily. A working group structure would be particularly useful.

Neil:

- The structure already exists. We can open the issue trackers a bit to see if people are interested in contributing.

Paul:

- How working groups would work would be worth discussing. This would take more than two meetings, but it is a good starting point.

John:

- If we can focus on one scientific project, having a successful case would be great to highlight how portable the project is. Collaboration between university and CCCma.

Hansi:

- The working groups should be more diverse in terms of users vs developers.
- Need to give out the computing hours we have to people who will use it! It should not sit unused.
- CCCma is focused on development, and the university is less interested. CCCma and university overlaps manage allocations to have some allocations for science. Would like to share some allocations for science.

Michael:

- Have to be very strict on reproducibility. If not, judging an “acceptable” climate in non-coupled mode may not be so in coupled scenarios where long simulations need to be done. Moreover, if the lack of bit-for-bit is due to a memory issue, this will have major issues down the line in different model configurations.