**Motivation**

A main challenge for the upcoming CMIP6 is to design metrics that reflect the ability of climate models to simulate processes rather than numbers. We take a first shot at it by proposing three Arctic sea ice diagnostics that quantify how a variety of models (with different levels of complexity) simulate essential processes such as the ice-albedo feedback, the ice growth-thickness feedback and the sea ice thickness-drift relationship. Note that a companion poster describes similar metrics for Antarctic ocean and sea ice simulations (Louvem and colleagues).

**Example 1 – Ice Melt: The Open Water Formation Efficiency**

Mechanism investigated: the ice-albedo feedback

The Open Water Formation Efficiency (OWFE) was first introduced by Holland et al. (GRL, 2006) to understand the origins of simulated abrupt sea ice reductions over the 21st century. It quantifies the ability of a model to produce open water during the melt season and is therefore a proxy for the ice albedo feedback. Moreover, it depends inversely on the baseline sea ice thickness: thin ice melts away more easily.

The OWFE is estimated for each year at every grid point, as the regression between sea ice concentration and volume over the melt season. Note that these changes account for both thermodynamic (melt) and dynamic (including advection, ridging and rafting) processes.

**Example 2 – Ice Growth: The Heat Conduction Index**

Mechanism investigated: the ice-growth-thickness feedback

Like the OWFE, the HCI diagnostic is very stable w.r.t. interannual variations of the forcing. This is because it is preserved under proportional increases of snow and ice thickness.

**Example 3 – Ice Dynamics**

Mechanism investigated: the sea ice thickness-drift relationship

Sea ice is a complex body, for which better rheologies are currently being tested. One basic relationship that a model should simulate is the “thickness-drift” relationship: Packed (typically >80% concentration) and thick ice drifts slower, all other things being equal.

**Conclusions, lessons learned and recommendations:**

- It remains to be seen how these process-based metrics relate to larger-scale, climate metrics (e.g., extent, volume, their respective trends) and how well they explain spread between projections.
- Process-based diagnostics do not imply to derive single numbers to quantify performance. An analysis of appropriate figures is equally justifiable. Expert judgement also has a strong role in that respect.
- Given the stability of the proposed diagnostics, targeted measurements over one year (e.g., the Year of Polar Prediction) could be enough to apply emergent constraints to the simulations.
- One critical aspect over the coming months will be to insert these diagnostics directly in the model codes (“on-line” metrics) and/or to rethink the list of variables to be saved.