Notes on NEMO Verification & Validation (VV) Webex: Fri 11 October 2019

Present: Mike B, Massimiliano D, Claire L , Nicolas M, Seb M, Pierre M, Simon M, Julien le S,

Apologies: Dorotea I

The focus of the meeting was on the structure of the roadmap for V&V activities that the team will write. **It was agreed to aim to write a first draft by Xmas**.

Structure of the V&V roadmap

A structure loosely based on the talk given to the NDC was discussed:

1) Introduction (objectives etc.)
2) What do we mean by V&V?
3) What do we have now?
4) What could we have long-term?
5) What could we improve in the short-term?
6) Road-map

Section 2 on the definition of V&V should be extended to give a brief overview of "textbook" V&V methods, their strengths and weaknesses and how they fit together (complement each other).

Section 3 will cover the SETTE & Trusting tools, the test cases, and the validation of new releases, with a focus on what aspects of the NEMO code and its V&V are and are not covered.

Section 4 should consider methods deployed in other model systems (e.g. the new GFDL code, CROCO, ROMS, WaveWatchIII) and how unit testing and test cases might be applied to NEMO. It should also aim to provide a conceptual framework for NEMO V&V (it has to supply the link between sections 2 and 6).

Section 5 should suggest short-term practical steps that can be taken to improve what we have now. So its starting point is section 3. Some of these practical steps might be included in the 2020 NST WorkPlan.

Section 6 needs to provide a bridge between sections 4 and 5. It needs to consider the resources that will be required and the order of priorities. Evidence about and an asssessment of the importance, urgency and value of the work also needs to be given.

Points raised

We need to think about our ideal V&V system as well as having a practical focus (the above structure tries to capture that).

We need to consider separately 1) what tools and processes we need and 2) how we can get there.

There are (or could be) strong links between unit tests & test cases. For both, the expected results have to be supplied.

We need to specify what functionality we are verifying. For example it could include HPC performance.

There are 600 diagnostics available within NEMO. Will we verify all of them?

The resources (human & computational) and time-scales (1 year, 5 year ?) need to be clarified; both what is required and what could realistically be made available.

It will be useful to capture experience from the past (both successes and failures)

We should aim to identify quick wins (section 5) but need to consider how they fit into the longer term plan.

A common view on definitions is helpful. People mean different things by reproduceability.

Our approach to a continuous integration suite is within the scope of V&V.

[Not discussed at the meeting but pretty obvious! This roadmap will be longer than a few pages. This does not particularly matter. We can write a summary once we've got the main points clearly established.]

Proposed contributions

It will be most efficient to take an iterative approach to development of this document. We don't have to write carefully crafted sentences in the first instance in some of the sections; bullet points on what is most important will be enough to get started.

We will write in a shared googledoc. Julien set this up immediately after the meeting.

Pierre M will contribute material on SETTEE for sections 2 (what do we have) and section 4 (short-term improvements). **Nicolas M** will contribute similar material on the Trusting tool.

Seb M and **Mike B** will explore the GFDL approaches to V&V. Seb will also look at CROCO.

Simon M will propose ideas for a simple infrastructure to support unit testing (section 4).

Julien le S will help with sections 2 and 4 (the organisation of the framework).

Claire will read through draft text.

Mike B will start with section 2 and then move onto the components of a long-term framework (section 4).

Points raised

There is a tension between getting on with short-term improvements as soon as possible (e.g. including them in the NST 2020 Workplan) and taking some time to produce a coherent roadmap. We don't have a clear view on how to do both at this point [we don't know which of our systems we will want to replace in the medium term].

When writing, where ever possible, specific examples should be given. This will make things clearer.

We need to work out how to better integrate testing into the development process.

Timeline

We'll meet again at 15:00 (CEST) on Tues 12th November. We'll discuss initial contributions and what to do next.