Where our science ambitions meet computing and data handling limitations?

Peter Bauer
IFS: government needs extra £5bn to fund spending promises

Thinktank says Wednesday’s spending review will amount to boost of £9bn, with £4.5bn already accounted for

The Treasury will need to find an extra £5bn from next year to fund the government’s spending promises, according to the Institute for Fiscal Studies (IFS), which said the increase could undermine the Treasury’s commitment to reduce borrowing.

Boris Johnson said on Monday that Wednesday’s Whitehall spending review, which will set out department funding for 2020/21, would be the “most ambitious spending round for more than a decade”.

A few days ago...
[Courtesy O Marsden]
So what are our ambitions?

**Sub-seasonal**

**Progress** in past 5 years:
- Initialisation with ERA-5, better medium-range model & reforecasts, resolution, ocean and sea-ice
- Improvement in tropics (gained 3 days lead time MJO over Indian Ocean)
- Improvement in reforecasts (but not in calibrated forecasts)
- Neutral (at best) in extra-tropics (NAO, stratosphere)
- Much better diagnostics, skill metrics
- S2S database crucial for inter-model comparisons, multi-model output

**Key issues:**
- MJO propagation across maritime continent
- Weak tele-connection between MJO and mid latitudes (depends on MJO in initial conditions)
- Lack of link between Rossby-wave breaking and blocking
- Large SST biases in western boundary currents
So what are our ambitions?

**Seasonal**

**Progress** in past 5 years:
- Initialisation with ERA-I/ORA-S5, better medium-range model & reforecasts, resolution, ocean and sea-ice
- Improvement for ENSO (4.5 months lead time for SST anomalies for Nino3.4), NAO in winter, tropical rainfall
- Improvement in reforecasts; both bias and variance = f(resolution, lead time, seasonal cycle)
- Neutralish in extra-tropics
- Much better diagnostics, skill metrics
- Copernicus CDS crucial for inter-model comparisons, multi-model output

**Key issues:**
- Spring prediction barrier for ENSO
- Lack of tele-connection between Indian Ocean and North Atlantic variability, inter-basin links
- Lack of stratospheric variability (e.g. QBO) and link to troposphere
- Large regional biases in ocean reanalyses (e.g. gulf stream), snow and soil moisture

[from various talks earlier thus week]
So what do we need to achieve our ambitions?

Options:

- **Horizontal resolution** in the atmosphere and interaction between convection, eddies & circulation
- **Horizontal resolution** in the ocean
- **Vertical resolution**, especially in tropics and UTLS
- **Lagged ensembles** instead of burst ensembles, ensemble size
- **Adding chemistry** (ozone) and aerosols
- **Initialising** with better ocean and sea ice (thickness), soil moisture and snow

- ... and the same for **reanalyses and reforecasts**

- ... and better **observations**, of course, allowing us to measure processes across scales
So what do we need to achieve our ambitions?

Is a seamless prediction system the answer?

• Definition of ‘seamless’:
  • Palmer et al. 2008 (BAMS): “predictions across the range of weather and climate time scales”
  • Brunet et al. 2010 (BAMS): “the spatial-temporal continuum of the interactions among weather, climate and the Earth system”
  • Ruti et al. 2019 (BAMS): “extend the concept of seamless prediction … to improve the development of knowledge and services”

• Here: same initial condition / forecast model / reforecast infrastructure as in medium range, something like a leg-C (leg-A = day 0 – 15, leg-B = day 16 – 46)

Challenges are opportunities:
• ease of maintainability & pace of upgrades
• single model & performance at all scales
• …
• technical & scientific
Cost-benefit from ‘configuration’ upgrades

- Initialisation with ERA-5 (instead of ERA-I)
- 137 vertical levels (instead of 91)
- Daily update with 51 lagged members (instead of twice weekly)
- Daily update with 25 lagged members (instead of twice weekly)
- 18 km resolution (instead of 36 km)
- 9 km resolution (instead of 36 km)

Similar experience sub-seasonal and seasonal

[Courtesy F Vitart]
Cost-benefit from adding atmospheric composition

Added cost:
- at CAMS resolution (= TL511) x2.5-3 for forecasts (x2 for analyses)
- At TCo1279 CO$_2$ and CH$_4$ x1.1 for forecasts
- At TCo1279 full chemistry and aerosols (150-200 tracers) x8-10

(50% chemistry, 50% advection, global norms, I/O)

Benefit:
- Not well known as no feedback to meteorology in operational systems
- Research exps. suggest some benefit, but in non-NWP configurations

[Courtesy J Fleming]
Cost-benefit from high-resolution ocean & sea ice

TCo639 ENS cost distribution:

- 19.7%
- 16.8%
- 4.6%
- 8.0%
- 3.7%
- 0.8%
- 29.8%
- 16.6%

Ocean + Waves = 36.5%

**Added cost:**
- x1.2 at TCo639, x1.1 at TCo1279 (same ocean model)

**Benefit:**
- Obvious at all scales
High-resolution GCMs or global storm resolving models?

**Resolution:** Add spatial resolution to eliminate parametrisations → Computing?

**Traditional:** Add model complexity, rely on parametrisations → Realism?

**Technology:** Use ML to replace parametrisations → Training?
High-resolution GCMs or Global storm resolving models?

They are not the same:

[Courtesy Bjorn Stevens]
Global storm resolving models

- Representation of the global mesoscale
- Multi-scale scale interactions of convection
- Circulation-driven microphysical processes
- Turbulence and gravity waves
- Synergy with satellite observations
- Downscaling for impact studies
- Etc.

[Courtesy Bjorn Stevens]
Present capability @ 1.45km: IFS (atmosphere)

$O(3-10)$ too slow (atmosphere only, no I/O)

$O(100-250)$ too slow (still no I/O)

$O(1000)$ incl. everything (ensembles, Earth system, etc.)

[Schulthess et al. 2019]
Present capability @ 1km: NEMO (ocean)
But we don’t have to move to 1km to be worried

Computing:

https://www.ecmwf.int/sites/default/files/ECMWF_Strategy_2016-2025.pdf: “[...] An ambitious target that depends on scientific, computing and scalability advances is for this ensemble to have a horizontal resolution of about 5 km by 2025. [...]”

HPC cost growth = \frac{Cost \ (50 \ members, \ 5 \ km, \ 200 \ levels)}{Cost \ (50 \ members, \ 18 \ km, \ 137 \ levels)} = O(100)

(more if we count significant ocean model upgrades and atmospheric composition)

Data:

Public access per year:
• 40 billions fields
• 20 PB retrieved
• 25,000 users

Total activity (Member States and commercial customers) per day:
• 450 TBytes retrieved
• 200 TBytes archived
• 1.5 million requests

Total volume in MARS: 220 PiB

[Courtesy T Quintino]
HPC history at ECMWF

x100 ≈ 10 years (in the past!)

[Courtesy M Hawkins]
Why?

The end of Moore’s law and Dennard scaling for CPU

The diversity of algorithms and their use of memory hierarchy, bandwidth, interconnects, storage etc.
Example from another sector
Reality check!

Real task graph of NCAR CESM POP ocean model

[Courtesy van Werkhoven]
We have 2 problems right now

1. We can’t pack more transistors on a chip anymore, clock-speed does not increase anymore
   (because of the end of Moore’s law and Dennard scaling)

2. Our codes have become increasingly inefficient on HPC systems
   (because of increasing forecasting system and code complexity)

→ from where will we get efficiency we need for fulfilling our strategy?
1st order solutions

1. We buy bigger computers (…just like in the past)
   a. No, because they will be unaffordable to buy
   b. No, because they will be unaffordable to power

2. We hand our codes over to software people (…just like in the past)
   a. No, because the existing codes do not have algorithmic flexibility
   b. No, because the existing codes do not have hardware flexibility
   c. No, because significant scientific judgment is needed to do less, to do it cheaper, and to eliminate bottlenecks
**ECMWF Scalability Programme – Holistic approach (1)**

- **Data acquisition**
  - Lean workflow in critical path
  - Object based data store
  - Load balancing obs-mod
  - Quality control and bias correction with ML
  - OOPS control layer
  - Algorithms: 4DV, En4DV, 4DEnVar
  - Models: IFS, NEMO, QG
  - Coupling
  - Surrogate models with ML

- **Data assimilation**
  - IFS-ST & IFS-FVM on same grid and with same physics
  - Coupling
  - Separation of concerns
  - Surrogate models with ML

- **Forecast**
  - Lean workflow in critical path
  - Object based data store
  - Use deep memory hierarchy
  - Broker-worker separation
  - Integration in Cloud
  - Data analytics with ML

- **Product generation**
  - Integration in Cloud
  - Workload prediction with ML

- **Dissemination**
  - RMDCN
  - Internet

- **Web services**
  - Internet

- **Archive**
  - Data Handling System
ECMWF Scalability Programme – Holistic approach (2)

**Software**
- Suite parallelism & concurrency
- Scheduling
- I/O
- File system access

**Hardware**
- Task parallelism & concurrency
- Memory hierarchy (size & latency)
- Scheduling
- Load balancing
- I/O
- File system access

- Task parallelism on distributed memory
- Thread parallelism on shared memory
- Load balancing
- On and off-node data transfers
- Memory hierarchy (size & latency)
- Precision
Wouldn’t it be nice, if...

- applied scientists could work with **Mathlab, Python**;
- applied scientists would **not have to worry about computing** aspects;
- applied scientists could try **very different algorithms**;
- applied scientists and end-users could **analyse data on the fly**;
- computational scientists could avoid **uncomputable science choices**;
- computational scientists could have **full access to present and future technology**;
- organisations could avoid **vendor lock-in**

ECMWF Scalability Programme – Aim
ECMWF Scalability Programme – Do less and do it cheaper

Single precision (Vana et al. 2017, MWR; Dueben et al. 2018, MWR):
• running IFS with single precision arithmetics saves 40% of runtime, IFS-ST offers options like precision by wavenumber;
• storing ensemble model output at even more reduced precision can save 67% of data volume;
→ to be implemented in operations asap (capability + capacity)

Concurrency:
• allocating threads/task (/across tasks) to model components like radiation or waves can save 20% (gain increases with resolution);
→ to be implemented in operations asap (capability + capacity)

Overlapping communication & computation:
• through programming models (Fortran co-array vs GPI2 vs MPI), gave substantial gains on Titan w/Gemini,
• on XC-30/40 w/ Aries there is no overall performance benefit over default MPI implementation;
→ to be explored further
ECMWF Scalability Programme – Use new processor technology

Spectral transforms @ 2.5 km, 240 fields on GPUs of the biggest HPC system in the world:

- x20 in time to solution
  (but a real pain to programme)

Cloud scheme on FPGAs on Maxeler test system:

- x3 in time to solution
  (but an excruciating pain to programme)
ECMWF Scalability Programme – Use new memory technology

Running ensembles and reading/writing to NVRAM produces no bottlenecks and scales well!

[Courtesy O Iffrig, T Quintino, S Smart]
ECMWF Scalability Programme – Ultimately, touch everything

Wouldn’t it be nice, if:

- ...

This comes at a cost:

- New technology ain’t cheap
- To migrate existing forecasting systems to new state requires work – and rethinking

There is no silver bullet … but the most important ingredient is flexibility:

- **Algorithms**: Way too science centric, need to be more computing focused
- **Programming**: Redesign millions of lines of code – separation of concerns (but in steps)
- **Computing**: Prepare for full hardware flexibility minimising energy to solution
- **Workflows**: Minimise data movement, post-process on the fly, recompute, explore Cloud
Programming: Traditional way of working

Mathematical description

Wind $\rho \vec{v} = -\nabla p + \rho g - 2\Omega \times (\rho \vec{v}) + \mathbf{F}$

Pressure $\dot{p} = -\left(\frac{c_{pd}}{c_{vd}}\right) p \nabla \cdot \mathbf{v} + \left(\frac{c_{pd}}{c_{vd}} - 1\right) Q_h$

Temperature $\rho c_{pd} \dot{T} = \dot{p} + Q_h$

Water $\rho_0 q^u = -\nabla \cdot \mathbf{F}^u - (I^l + I^f)$

$\rho_0 q^{\ell f} = \nabla \cdot (p^{\ell f} + \mathbf{F}^{\ell f}) + I^{\ell f}$

Density $\rho = p \left[ R_d \left( 1 + \left( \frac{R_v}{R_d} - 1 \right) q^u - q^l - q^l \right) \right]^{-1}$

Domain science and applied mathematics

$$\text{lap}(i, j, k) = -4.0 \times \text{data}(i, j, k) + \text{data}(i+1, j, k) + \text{data}(i-1, j, k) + \text{data}(i, j+1, k) + \text{data}(i, j-1, k)$$

Physical model

Algorithmic description

Imperative code

Compilation

Computer engineering

[Schulthess 2015]
Programming: Separation of concerns

Mathematical description:
- Wind $\nu = - \nabla \rho + \rho g = 2 \nabla (\rho v) + F$
- Pressure $p = -(\gamma_{w}/\gamma_{d}) p \nabla v + (\gamma_{w}/\gamma_{d} - 1) Q_{0}$
- Temperature $\rho(T - \beta + Q_{0})$
- Water $\rho w = - \nabla \cdot \rho w = - (I + I')$
- $\rho d = - \nabla \cdot \rho d = - (I + I')$
- Density $\rho = \rho(R_{g}(1 - \rho / R_{g}) q^{n} - q^{n} - q^{n}) T^{-1}$

Science applications using a descriptive and dynamic developer environment

Algorithmic description

Imperative code

Compiler front-end

Optimization / low-level libraries / runtime

Architecture specific back-ends

Architecture 1

Architecture 2

Architecture N

Multidisciplinary co-design of tools, libraries, programming environment

↑ science specific code

↓ generic code

[Schultess 2015]
So, where are we with all this?

Open questions:

- What about code that is not in our control, e.g. NEMO?
- Do we have sufficient expertise – collaboration?
- Do we have sufficient funding?
Conclusions

At present:
• Long-standing skill issues with no apparent solution:
  → are incremental business-as-usual upgrades sufficient?
• No seamlessness, because of model & initial condition, test & operational system differences (both scientific & technical):
  → are impact sensitivity experiments useful that work with configurations that are untenable, and that perform different than integrated algorithms/solutions?

Urgent next steps*:
• Invest in common & future-proof software infrastructures:
  → generic and scientifically sound framework for performing science-cost trade off
• Invest in seamless diagnostics and cycle upgrades:
  → effective transfer of research to operations
  → identification of key sources of model error = sources for gaining predictive skill

(*at lot of this does not lead to Nature publications, but is essential for the future)
So is a seamless prediction system the answer?

• Yes, regarding a **unified software infrastructure** that allows to:
  • Produce and test **objective science** across scales
  • optimally exploit **HPC & BD** technology
  • enable cost-effective **research-to-operations-to-research** transfer
  • enable cost-effective **science-to-impact** transfer

• No, in terms of running forecasts in a **single configuration** across all scales (mostly because it’s not cost effective)

• But seamless thinking is crucial: science – computing – impacts