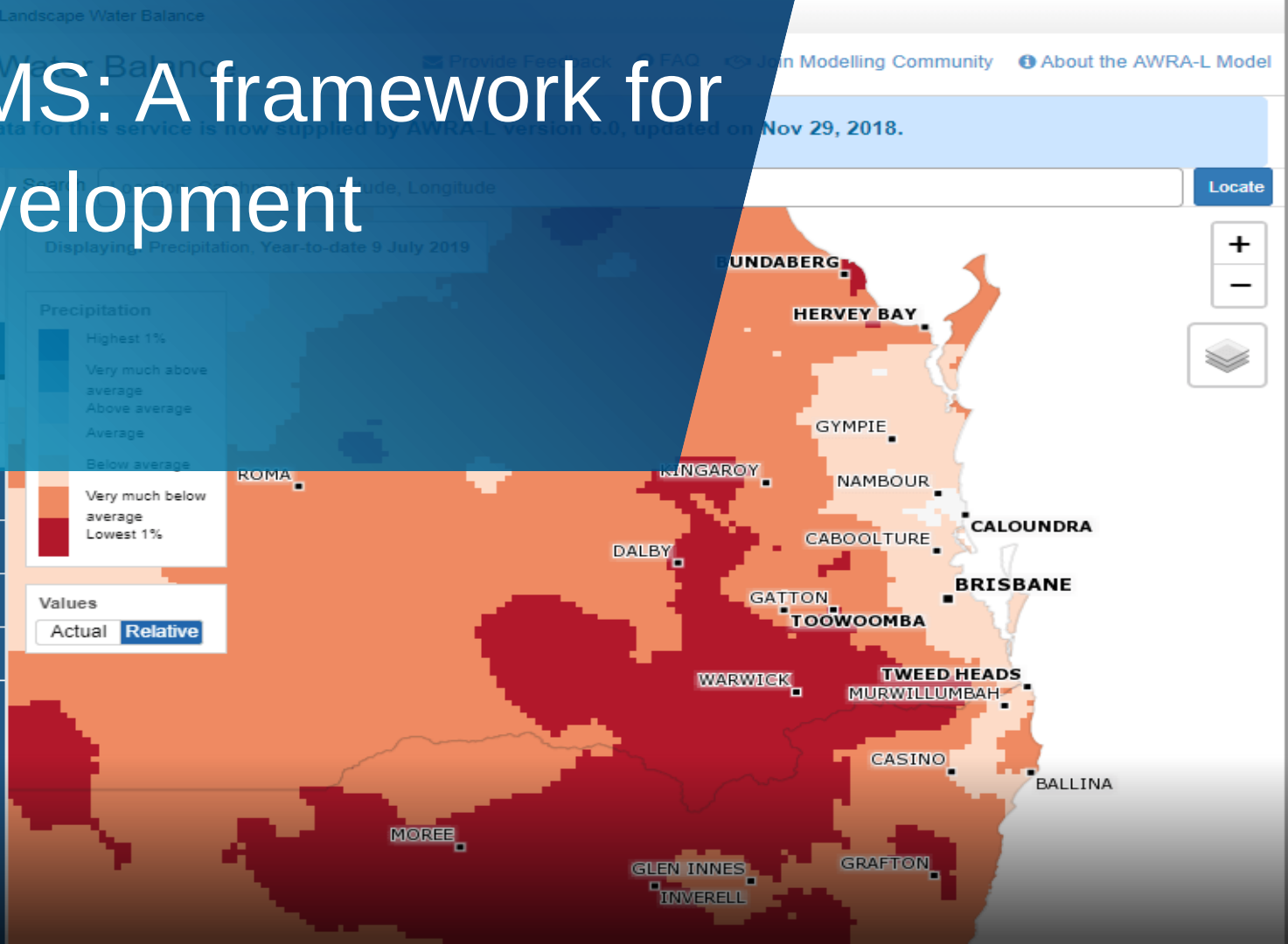
AWRA-CMS: A framework for model development

December 2019



Australian Government
Bureau of Meteorology

Overview

Background

What is the AWRA-CMS

Who are our users?

Enabled by HPC + good software design

- FAST
- FLEXIBLE
- ROBUST

Tools to aid model development

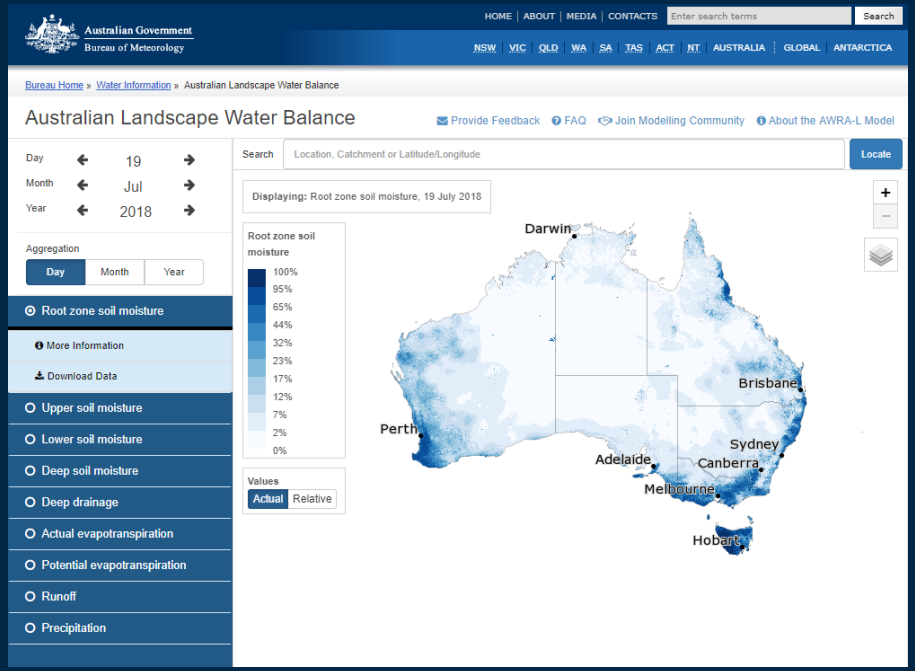
- Workflow
- Implementation
- Calibration
- Validation
- Analysis

What is the AWRA-CMS?

- Operational framework for the Bureau's national hydrological model

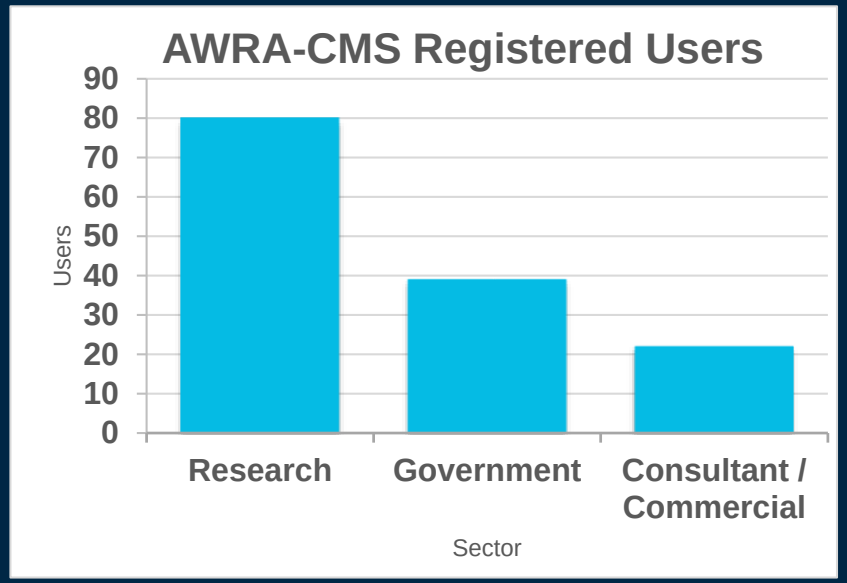
- Products:
  - Soil moisture
  - Runoff
  - ET
  - Deep drainage

- Wider community:
  - Reduced feature public package on GitHub



Who are our users?

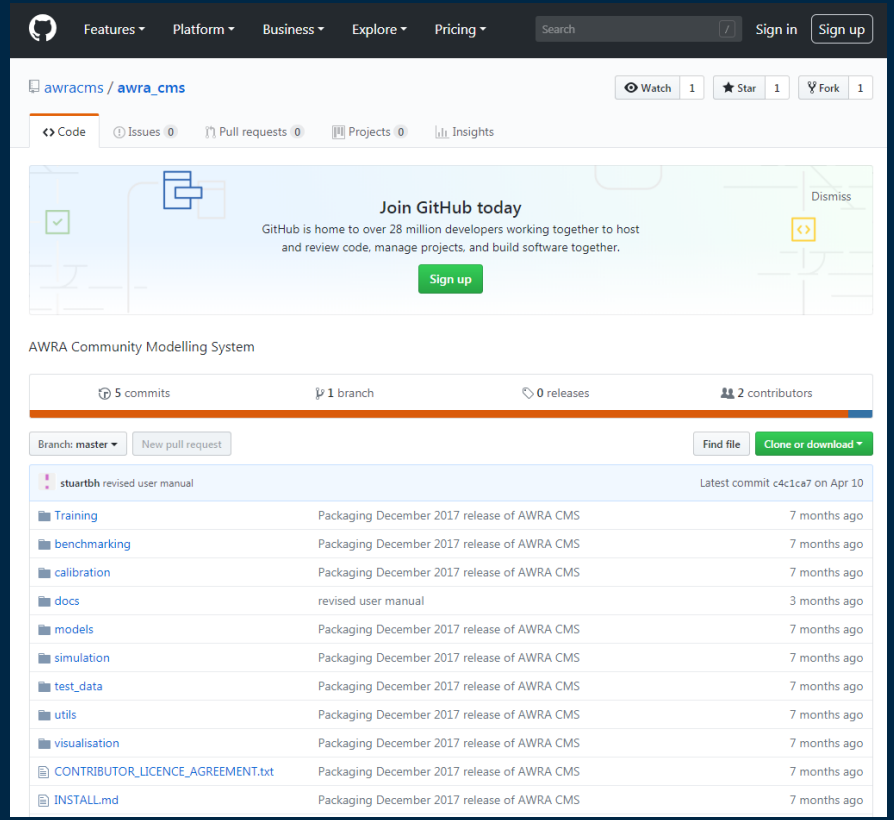
- 144 Registered Users of the AWRA CMS (since 2017):
  - 80 Research
  - 39 Government – State and Federal
  - 22 Consultants / Commercial – which includes entities such as Snowy Hydro, eWater, COFCO Agri, HARC Consulting and Jacobs



How do we collaborate?

- Training:
  - Half day model overview for data users
  - One and half days technical instruction including tutorials for model users
  - Training notebooks and dataset which get shipped with the codebase

- Collaboration tools:
  - GitHub: Functionality to review and merge model changes (https://github.com/awracms/awra_cms)
  - Notebooks: Easy for novice user to combine code, documentation and visualizations

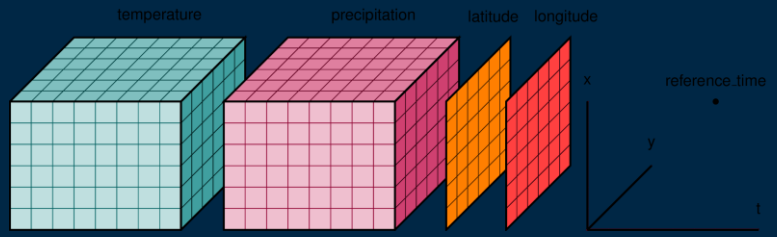


Enabled by HPC and good software design: **FAST**

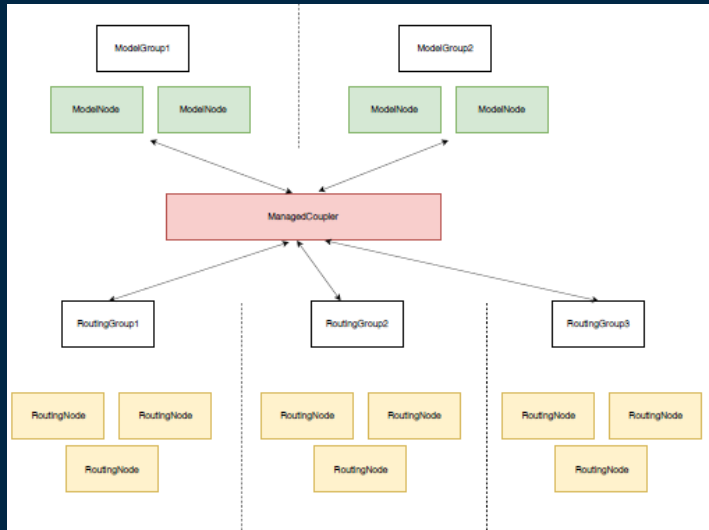
How?

- Use of
  - Vectorisation
  - Parallel compute across and within nodes
  - Parallel IO
  - Big data: use of next generation Python data science libs

- On the fly
  - Data transformation
  - Templating and compilation
  - Model coupling (e.g. model to routing)



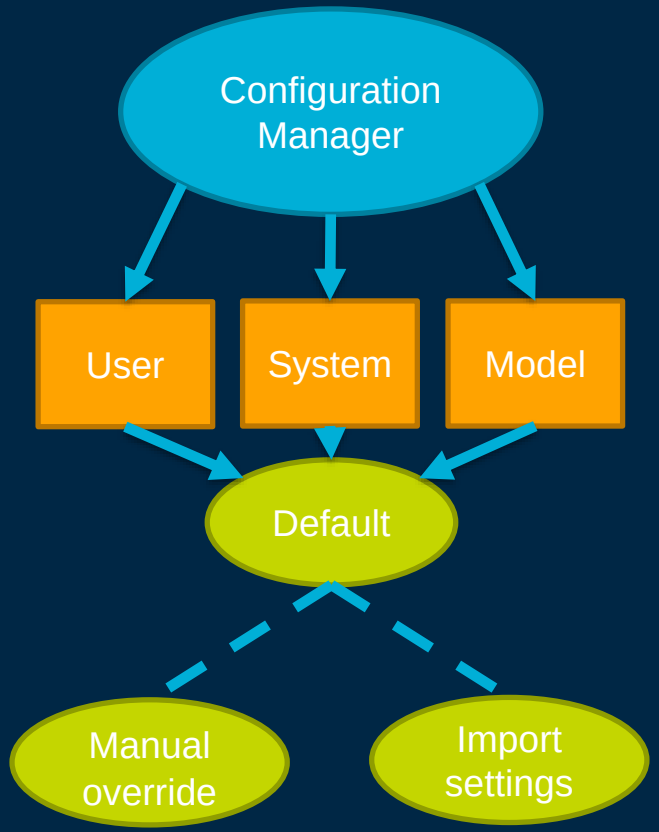
Source: http://martin-jung.github.io/post/2018-xarrayregression/



Enabled by HPC and good software design: **FLEXIBLE**

How?

- Configuration manager
  - One installation script; installs training data, sets paths and settings based on installation location
  - Training notebooks run out of the box with training data
  - Complexity abstracted for the user
  - Easier for external users to collaborate



Enabled by HPC and good software design: **FLEXIBLE**

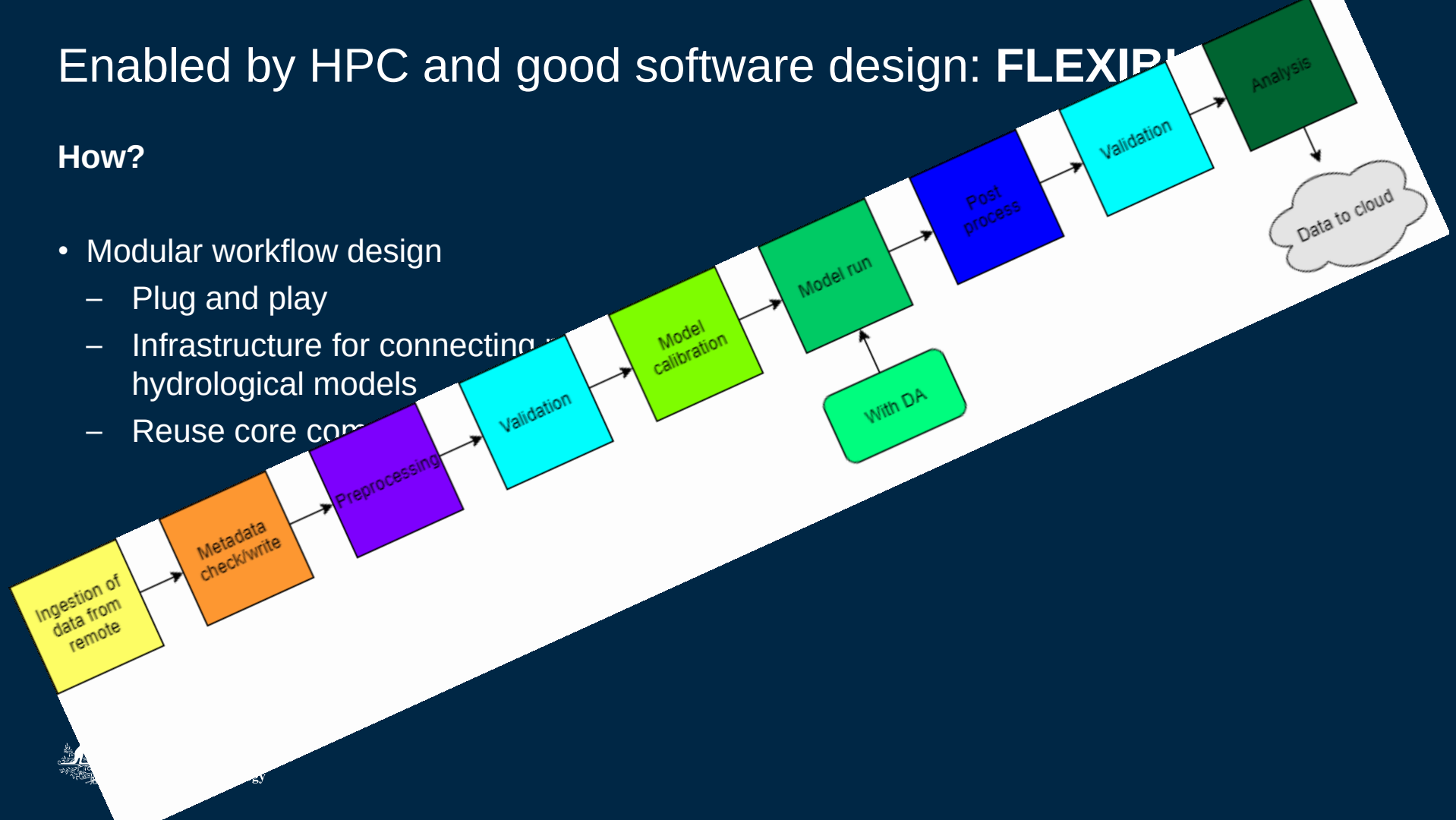
How?

- Modular workflow design
  - Plug and play
  - Infrastructure for connecting new hydrological models
  - Reuse core components / inheritance

Enabled by HPC and good software design: **FLEXIB**

How?

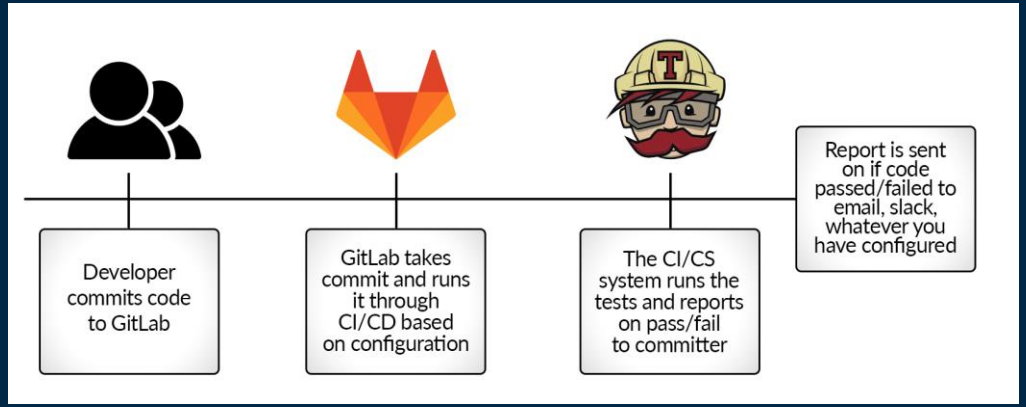
- Modular workflow design
  - Plug and play
  - Infrastructure for connecting hydrological models
  - Reuse core com



Enabled by HPC and good software design: **ROBUST**

How?

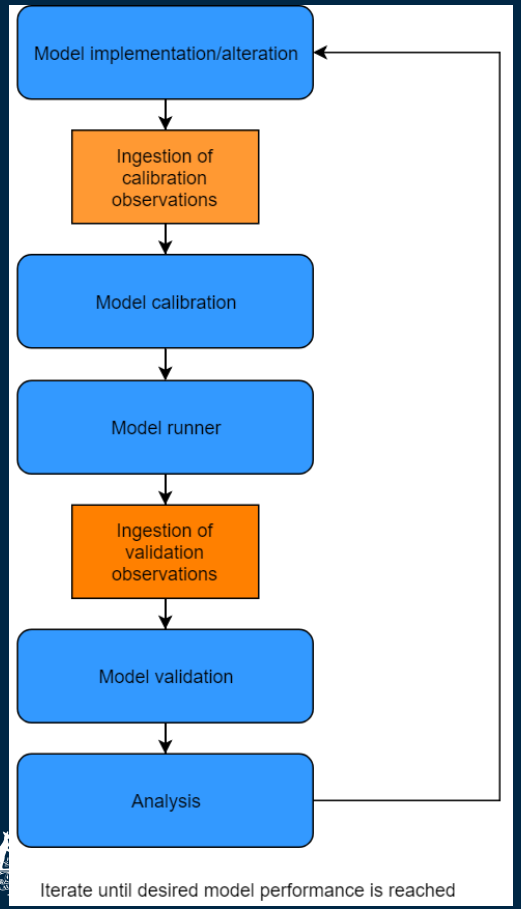
- Continuous integration
  - Training notebooks, plus tests are part of a continuous integration pipeline
  - Metrics collected for performance, stability etc
  - Linting
  - Check for security vulnerabilities
  - Run on multiple platforms including HPC



Source: https://www.hurricanelabs.com/splunk-tutorials/splunk-app-development-continuous-integration

- Testing
  - Ensure adequate test coverage
  - Stress testing

Model development workflow:



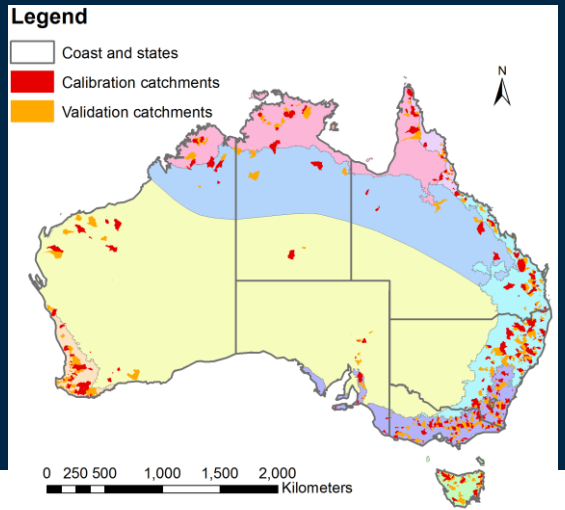
Field variable	Observation data set
Streamflow	Streamflow gauge timeseries (Zhang et al. 2013)
Evapotranspiration	OzFlux, CMRSET
Soil moisture	ASCAT, SASMAS, OzNet, AMSR-E
Ground water deep drainage	Recharge timeseries (Crosbie et al. 2010, 2015)

Model development workflow: Model implementation

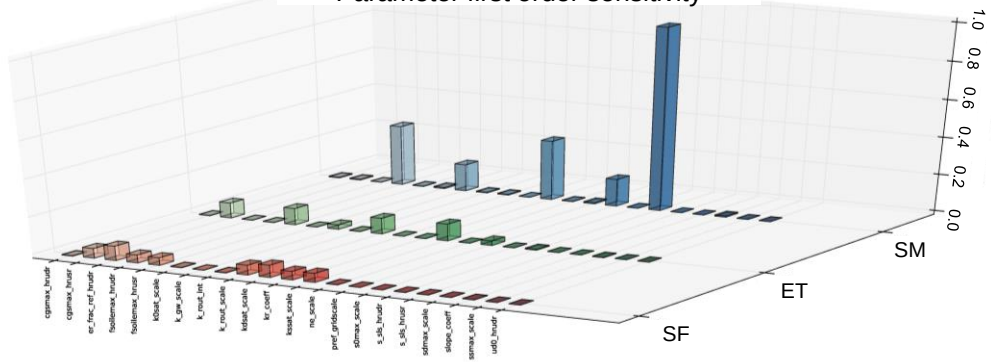
- Wrappers and templating:
  - Implement your own hydrological model in the CMS using python wrapper files which setup the inputs and outputs to your model code
  - Model code templating functionality for the user to change/set inputs and outputs on the fly
  - Can also use templating functionality to make changes to the model code itself

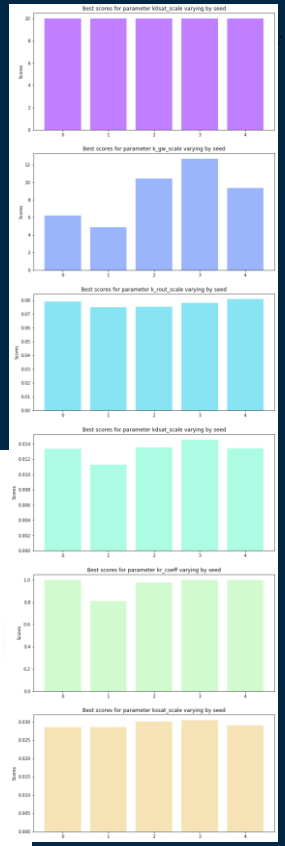
```
def gen_templates(tdict):


    return {
        'STRUCT_DEFS': gen_struct_defs(tdict),
        'WRITE_HRU': gen_hru_writes(tdict['OUTPUTS_HRU']),
        'INIT_COMBINED': gen_init_combined(tdict['OUTPUTS_AVG']),
        'WRITE_COMBINED': gen_combined_writes(tdict['OUTPUTS_AVG']),
        'WRITE_CELL': gen_cell_writes(tdict['OUTPUTS_CELL']),
        'LOAD_FORCING': gen_forcing_loads(tdict['INPUTS_FORCING']),
        'LOAD_CONST': gen_const_loads(tdict['INPUTS_SCALAR']),
        'LOAD_SPATIAL': gen_spatial_loads(tdict['INPUTS_SPATIAL']),
        'LOAD_HRUSPATIAL': gen_hruspatial_loads(tdict['INPUTS_SPATIAL_HRU']),
        'LOAD_HRUCONST': gen_hruconst_loads(tdict['INPUTS_SCALAR_HRU'])
    }
```

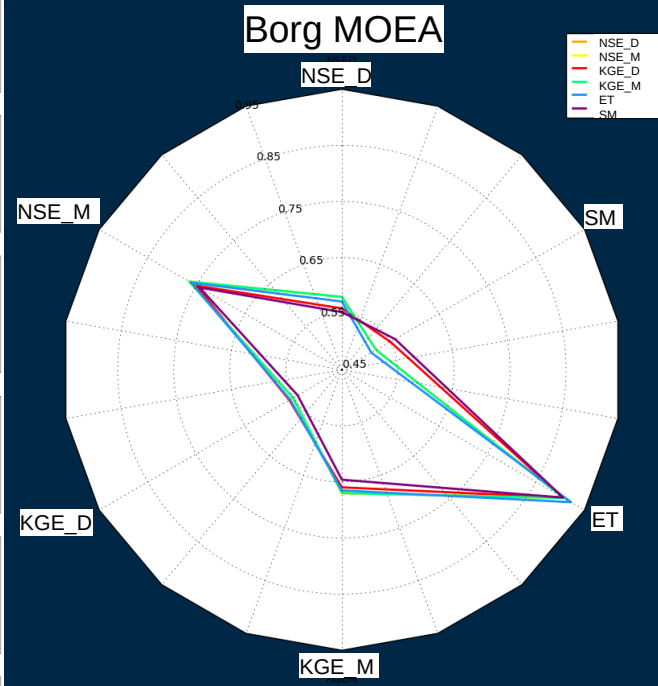
Model development workflow: Model calibration



Parameter first order sensitivity

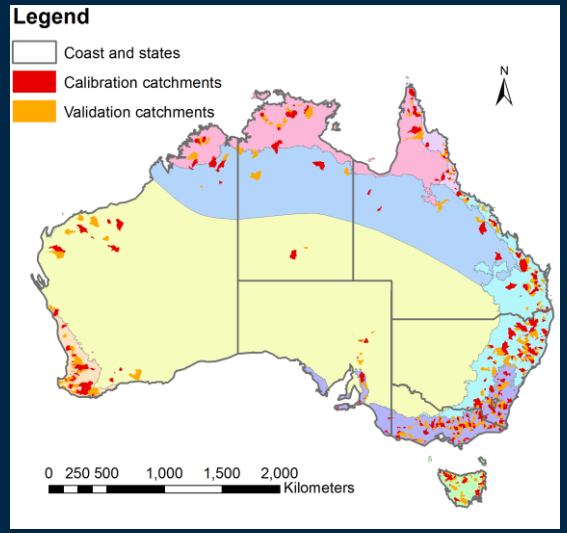


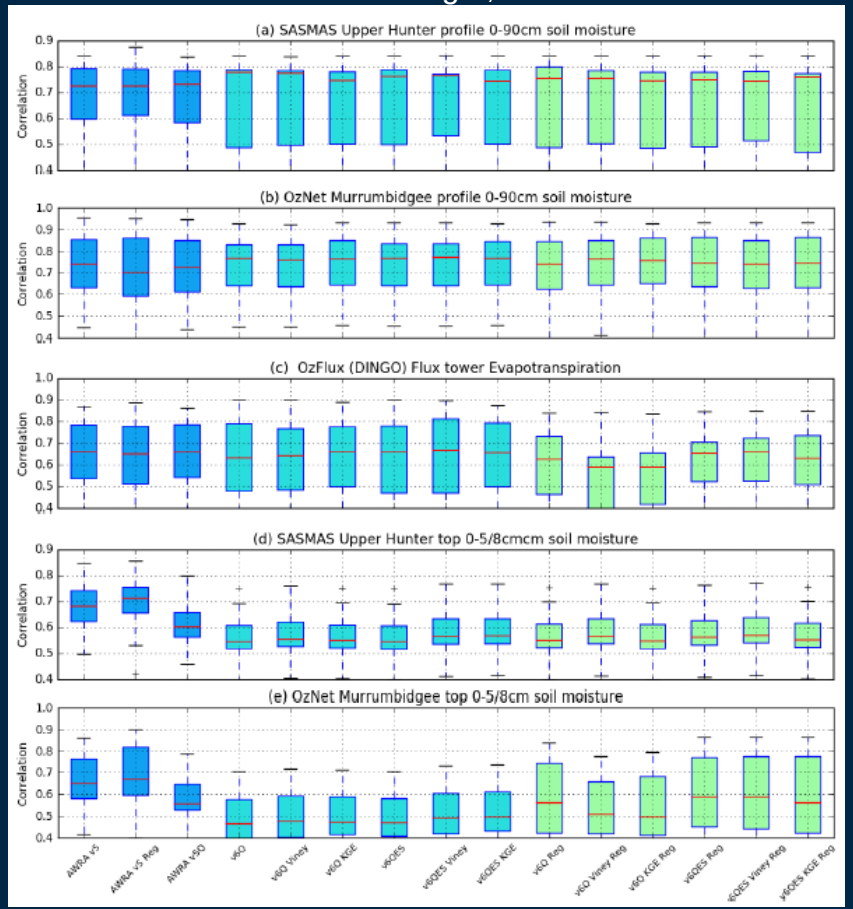




Model development workflow: Model validation

Frost & Wright, 2018





Australian Government
Bureau of Meteorology

Model development workflow: Model analysis

Data and statistic	Perc. (%)	Target	Best	CABLE	WaterDyn	v5	v5 Reg	v5 Q	v6Q	v6Q Viney	v6Q KGE	v6QES	v6QES Viney	v6QES KGE	v6Q Reg	v6Q VineyReg	v6Q KGE Reg	v6QES Reg	v6QES VineyReg	v6QES KGE Reg
Streamflow																				
Daily NSE	5%	0	0.01			-0.30	0.00	-0.56	-0.40	-0.28	-0.89	0.01	0.00	-1.88	-0.21	-0.07	-1.55	-0.23	-0.30	-1.19
	50%	0.5	0.53			0.45	0.53	0.49	0.52	0.52	0.48	0.48	0.48	0.44	0.51	0.52	0.47	0.51	0.51	0.49
Monthly NSE	5%	0	-0.02	-0.38	-1.76	-0.23	-0.02	-0.30	-0.38	-0.21	-0.68	-0.21	-0.14	-0.61	-0.66	-0.20	-0.47	-0.38	-0.61	-0.43
	50%	0.5	0.74	0.31	0.60	0.70	0.74	0.70	0.68	0.68	0.68	0.69	0.67	0.67	0.72	0.70	0.69	0.70	0.71	0.69
Monthly Relative Bias	5%	>-0.5	-0.41	-0.56	-0.46	-0.46	-0.49	-0.53	-0.41	-0.46	-0.43	-0.44	-0.45	-0.44	-0.44	-0.47	-0.46	-0.43	-0.44	-0.41
	25%	>-0.3	-0.16	-0.27	-0.16	-0.18	-0.20	-0.20	-0.19	-0.21	-0.17	-0.23	-0.20	-0.17	-0.18	-0.21	-0.20	-0.17	-0.16	-0.18
	75%	<0.3	0.25	0.32	0.60	0.30	0.26	0.28	0.31	0.32	0.36	0.25	0.27	0.32	0.31	0.29	0.30	0.30	0.32	0.34
	95%	<1.5	1.10	1.44	2.38	1.28	1.21	1.10	1.53	1.53	1.29	1.37	1.39	1.25	1.44	1.16	1.35	1.34	1.26	1.36
Soil moisture 0-90cm																				
SASMAS Daily correlation	50%	0.75	0.78			0.73	0.72	0.73	0.78	0.77	0.75	0.76	0.76	0.74	0.75	0.75	0.75	0.75	0.75	0.76
SASMAS Monthly correlation	50%	0.75	0.75	0.69	0.49	0.72	0.71	0.73	0.75	0.74	0.75	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.73	0.75
OzNet Daily correlation	0.50%	0.75	0.77			0.74	0.70	0.72	0.77	0.76	0.76	0.77	0.77	0.77	0.74	0.76	0.76	0.74	0.74	0.74
OzNet Monthly correlation	50%	0.75	0.79	0.79	0.77	0.76	0.71	0.75	0.78	0.78	0.79	0.78	0.79	0.79	0.74	0.77	0.77	0.75	0.75	0.75
Actual Evapotranspiration																				
Monthly correlation	5%	0.5	0.61	0.32	0.61	0.49	0.44	0.47	0.42	0.38	0.41	0.39	0.41	0.40	0.37	0.34	0.37	0.42	0.40	0.42
	50%	0.8	0.86	0.85	0.86	0.82	0.83	0.83	0.83	0.83	0.86	0.85	0.85	0.83	0.73	0.71	0.75	0.81	0.80	0.79
Daily correlation	5%	0.5	0.37			0.37	0.34	0.35	0.25	0.24	0.26	0.18	0.19	0.26	0.28	0.29	0.33	0.31	0.30	0.34
	50%	0.75	0.67			0.66	0.65	0.66	0.63	0.64	0.66	0.66	0.67	0.66	0.63	0.59	0.59	0.65	0.66	0.63
Recharge																				
Long Term Average Relative bias*	25%	<0	0.60	0.48	0.60	-0.57	-0.75	-0.62	-0.30	-0.32	-0.35	-0.30	-0.20	-0.27	-0.59	-0.20	-0.28	-0.25	-0.39	-0.34
	75%	>0	2.35	11.50	13.98	3.96	2.35	3.27	6.17	5.99	5.25	5.97	6.77	6.19	4.77	6.71	6.09	7.06	5.28	5.99
Annual correlation	5%	0.5	0.51	0.44	0.47	0.22	0.17	0.24	0.31	0.29	0.41	0.30	0.30	0.46	0.41	0.39	0.47	0.32	0.34	0.51

Frost & Wright, 2018

Summary

To aid model development:

=>We've tried to build a fast, flexible and robust modelling system which is easy to use

1. **Fast**

- Good software design (optimization, parallelism, load balancing etc.)

2. **Flexible**

- Build your own workflows
- Can be installed on different platforms
- Different models can be implemented in the same framework

3. **Robust**

- CI pipeline and extensive test coverage

4. **Ease of use**

- Use of interactive python to set up and run simulations, calibrations and validations
- Training notebooks for self directed learning

Thanks for listening

CREDITS:

Andrew Frost, David Shipman, Stuart Baron-Hay, David Smith, Ashkan Shokri, David Wright, Chantal Donnelly

Bureau of Meteorology

Australian Government
Bureau of Meteorology