Harnessing Professional Communities for Model Development

Chantal Donnelly, Bureau of Meteorology, Feb 2020

Australian Government
Bureau of Meteorology

Contact: Chantal Donnelly
Chantal.donnelly@bom.gov.au

Overview

- What is a Community Modelling System?
- Why model as a community?
- Experiences with the HYPE and AWRA
- Keys to Success
- Long-term vision for AWRA-CMS
- How can the community contribute?

Community Modelling Systems

- Leverage the efforts of university researchers, institutions, consultants and others to provide continuous model development
- Efforts range from completely open source to interagency agreements
- Common in climate, national weather prediction and earth system modelling communities
- Emerging in Hydrological Modelling Communities, particularly large scale hydrology

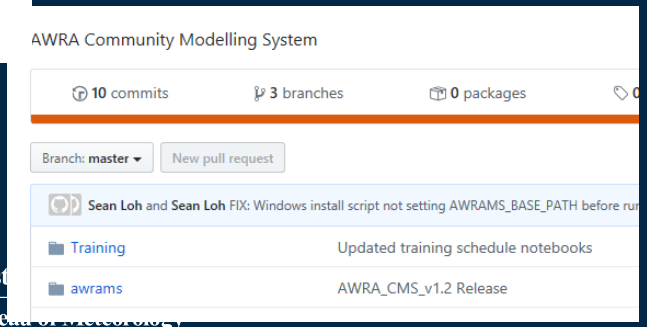
Australian Community Climate and Earth System Simulator (ACCESS)

SWAT Soil & Water Assessment Tool

CABLE: The Community Atmosphere Biosphere Land Exchange Model

High Resolution Limited Area Model Hirlam

Why GitHub?



WATER MODEL COMMUNITY

R

python™

Aust
Bureau of Meteorology

EC EARTH

Why Develop a model as a Community?

- More efficient: With every PhD student (almost!) making their own model, there is lots of repeated effort, and going over and over the same scientific ground.
- We leverage expertise, and thus drive science forward better and faster
- Stimulate discussion on key process disagreements in hydrology
- International application means all important processes will need to be explicitly modelled – more robust models in time and space, situations in which model's fail are highlighted
- Would further push data standardisation and quality in the field

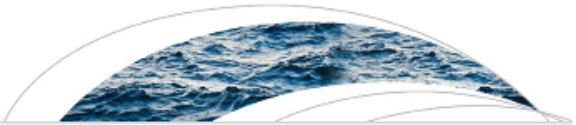
Counterarguments

- Applications of hydrological modelling vary widely and complexity is often guided by the application
- Might only be useful for furthering science, not for engineering application?
- Too computationally heavy for some applications
- Can we actually agree on the modelling concepts as a community?

A call for Community Hydrological Modelling

AGU PUBLICATIONS

Water Resources Research



COMMENTARY
10.1002/2014WR016731

Correspondence to:
M. Weiler,
markus.weiler@hydrology.uni-freiburg.de

Citation:
Weiler, M., and K. Beven (2015), Do we need a Community Hydrological Model?, *Water Resour. Res.*, *51*, 7777–7784, doi:10.1002/2014WR016731.

Received 26 NOV 2014
Accepted 6 AUG 2015
Accepted article online 18 AUG 2015
Published online 12 SEP 2015

Do we need a Community Hydrological Model?

Markus Weiler[1] and Keith Beven[2,3,4]

[1]Faculty of Environment and Natural Resources, University Freiburg, Freiburg im Breisgau, Germany, [2]Lancaster Environment Centre, Lancaster University, Lancaster, UK, [3]Department of Earth Sciences, Uppsala University, Uppsala, Sweden, [4]IDYST, University of Lausanne, Lausanne, Switzerland

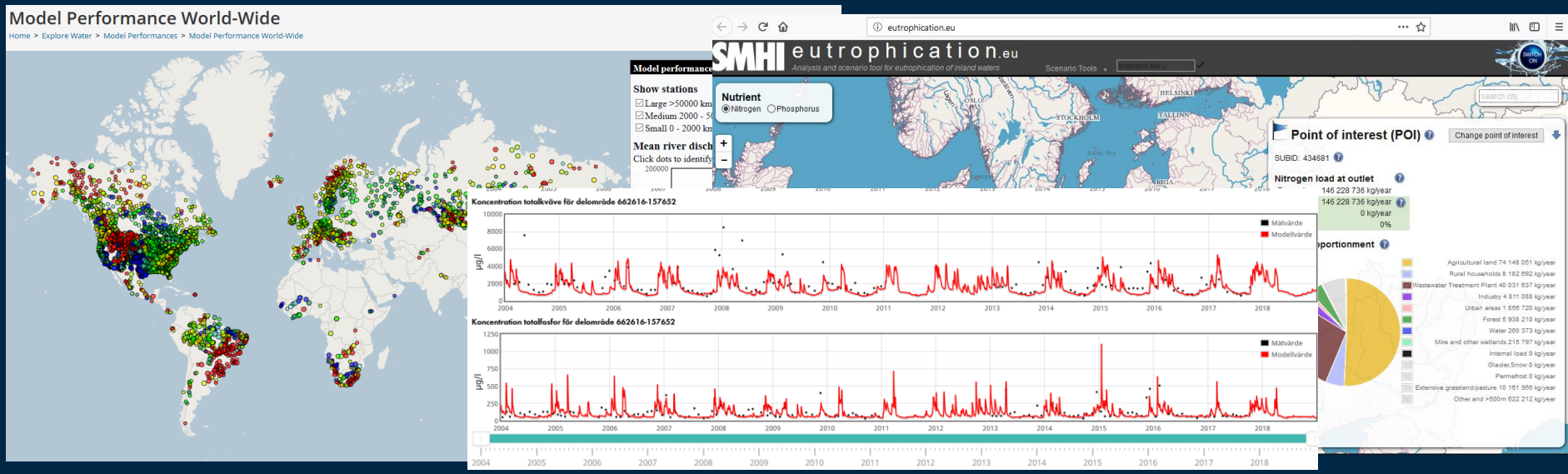
Abstract We believe that there are too many models in hydrology and we should ask ourselves the question, if we are currently wasting time and effort in developing another model again instead of focusing on the development of a Community Hydrological Model. In other fields, this kind of models has been quite successful, but due to several reasons, no single community model has been developed in the field of hydrology yet. The concept, strength, and weakness of a community model were discussed at the Chapman Conference on Catchment Spatial Behaviour and Complex Organisation held in Luxembourg in September 2014. This discussion as well as our own opinions about the potential of a community models or at least the necessary discussion to establish one are debated in this commentary.

1. There are Too Many Models in Hydrology

This is partly because every generation of PhD students thinks they could do better, partly because there is a wide range of different types of applications with different needs in hydrology, and partly because there is no agreement on a common set of concepts for the process representations. However, many of those models are similar in structure (for both the "bucket" model and the "physically based" model cases) so that the same, or very similar, concepts have been programmed and tested over and over again. There are

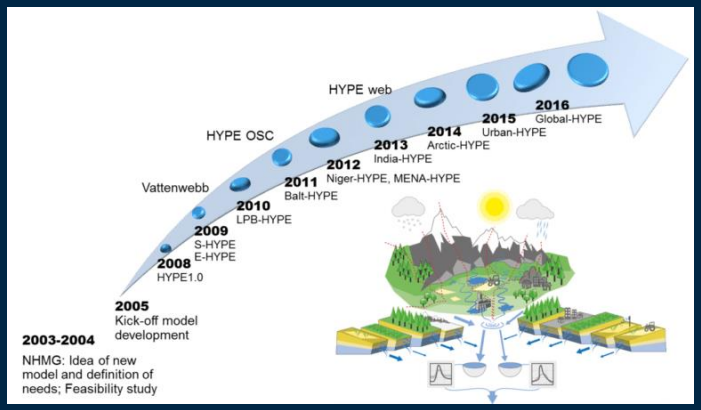
HYPE: Hydrological Predictions for the Environment

- Initially developed by the Swedish Meteorological and Hydrological Institute over the period 2004-2008, to support national reporting on water and nutrients for the European Water Framework Directive.
- Simulates the flow sand transformation of water, nutrients, organic carbon and sediment in the landscape and in the surface water system. It includes the impacts of irrigation, damming, points sources, floodplains and aquifers.
- Between 2008 and 2018, the model has been applied at subcatchment, catchment, river basin, national, continental and global scale
- Used for scenario modelling, nutrient source apportionment, historical means, forecasting and climate change projections



HYPE

- First applications emerged 2008
- Released as open source Oct 2011 on SourceForge (HYPE v 3.5.3) and with a Kick-off and Training workshop
- Currently at HYPE v5.9.0
- Code that can flexibly be applied at any spatial scale, anywhere in the world
- SMHI have developed and run the model for Sweden, La Plata Basin, Aland Islands, Europe, India, Niger, Arctic Basin, Chile, South America
- Can be used as a scenario tool, in nowcasting, forecasting or projections mode



https://hypeweb.smhi.se/about-us/about-the-model/

Ongoing Internal Improvements/additions

- DOC
- Aquifers,
- Generalised lake outlet and reservoir routines,
- Irrigation
- Wetlands and overbank flow and nutrient processes
- Sediment transport
- Glacier modelling
- Data Assimilation
- Source Apportionment
- Subdaily time-step
- + many process improvements for water and nutrients

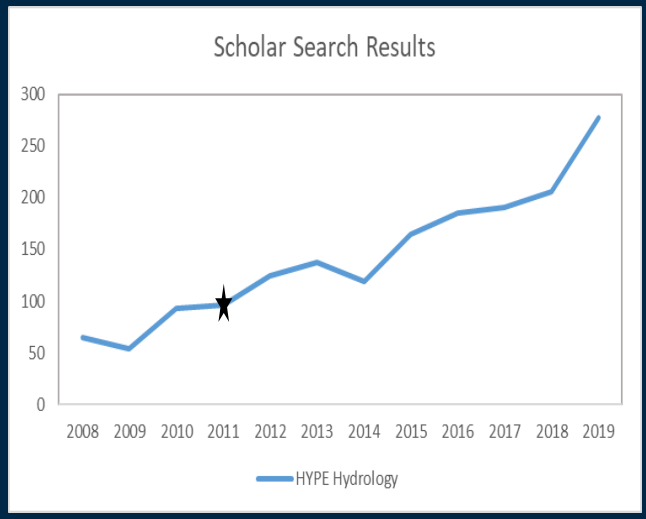
The HYPE Community

External Community Contributions

- Urban runoff improvements – Tanouchi et al., 2016, Japan
- Crop rotations for nutrient inputs – Yin et al., 2016, China
- Sediment – de Lavenne et al. 2019,
- Microbial contamination – Lewerin et al. 2019
- Spatial validation – Ghaffar et al. 2018, Germany
- Runoff generation improvements – Sinha et al. 2016, Germany
- Permafrost – Russia
- Kettle Lakes - Canada

- Routed HRUs – Hankin et al., United Kingdom
- Microplastic transport – Unice et al. 2019, France

Independent model applications in Russia (multiple), Latvia, India, Poland, South Africa, UK, Norway, Vietnam, Germany, UK, Niger, Chad, Chile, Turkey,



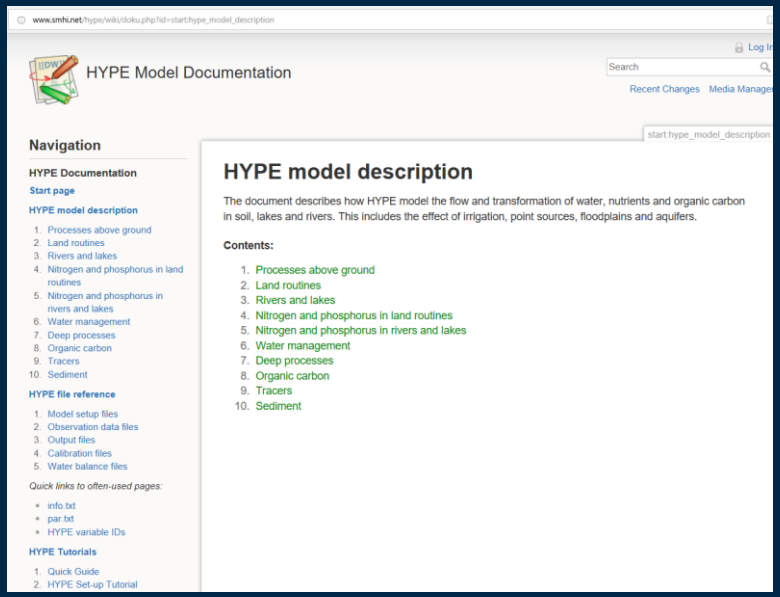
- Independent external publications emerge 2015
- External contributions to model code from about 2016
- Took about 5 years to build up community, still building up strongly

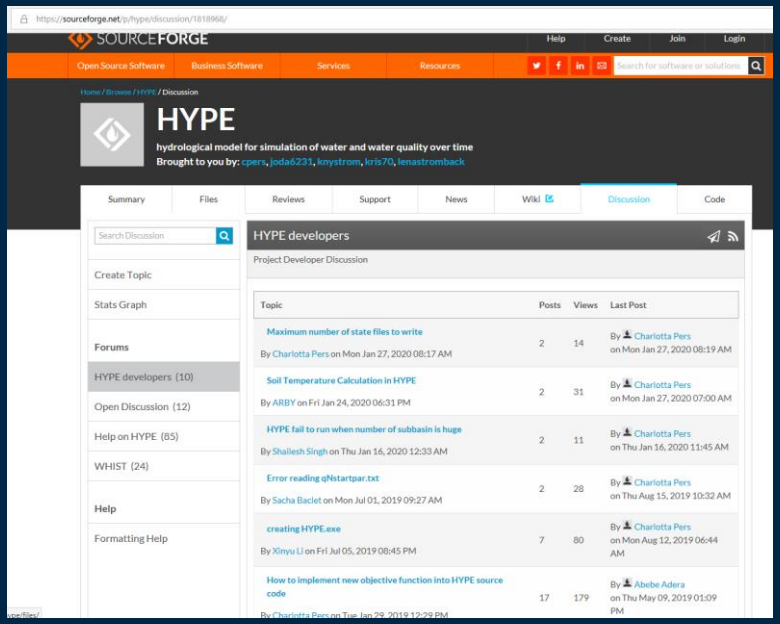
Caveat: Scholar search includes host institute's own contributions, other non-relevant papers

Australian Government
Bureau of Meteorology

HYPE community – Why is it working?

- A committed, sponsoring institution with a strong internal community
- Annual, free training
- Comprehensive wiki covering the code and how to use it
- A supported forum
- First build up model users, then model developers
- International visibility and networking in projects, at conferences and in publications



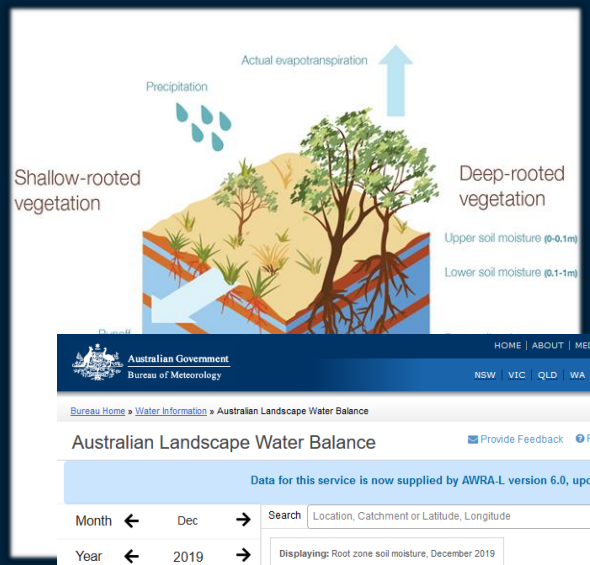


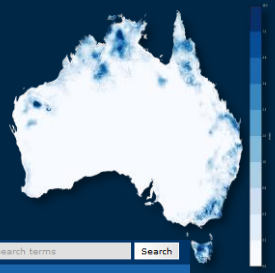
Australian Government
Bureau of Meteorology

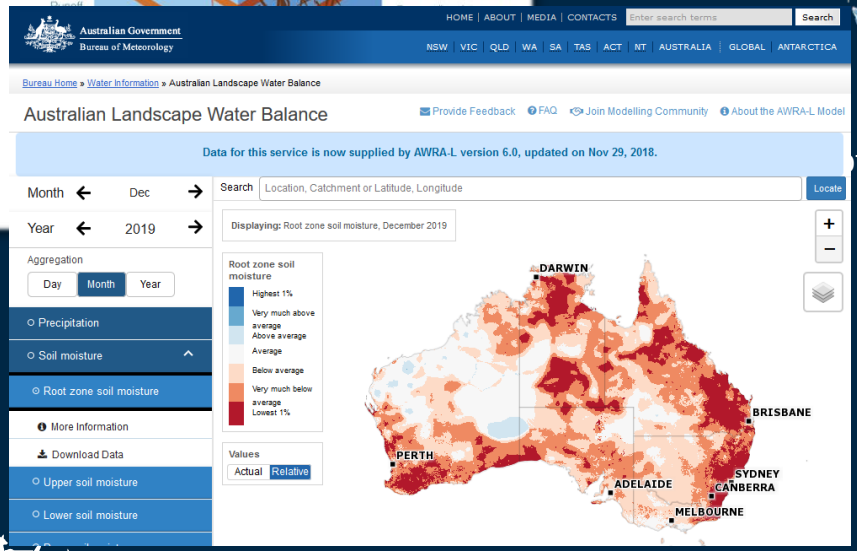
AWRA Community Modelling System

Modelling System with Python Wrappers and a C core

National Application, AWRA-L







- Co-developed by Bureau and CSIRO (WIRADA) since 2010
- Launched as a community model (AWRA-CMS), in Dec 2017
- Used operationally by the Bureau of Meteorology since 2016
- Forecasting and future projection applications coming 2020.

Australian Government
Bureau of Meteorology

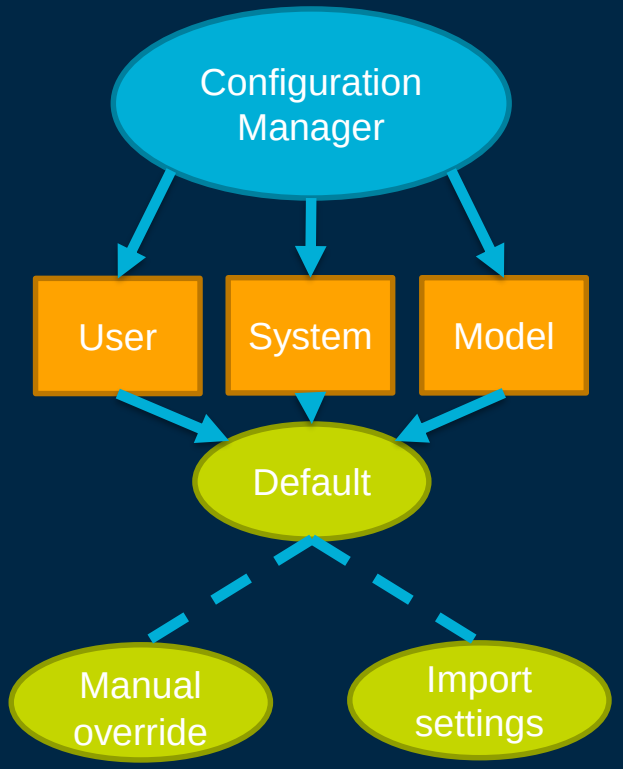
www.bom.gov.au/water/landscape

AWRA – CMS Enabled by HPC and good software design:

How?

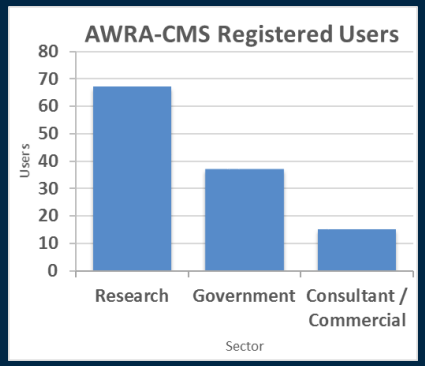
- Configuration manager
  - One installation script; installs training data, sets paths and settings based on installation location
  - Training notebooks run out of the box with training data
  - Complexity abstracted for the user
  - Easier for external users to collaborate
  - FLEXIBLE!

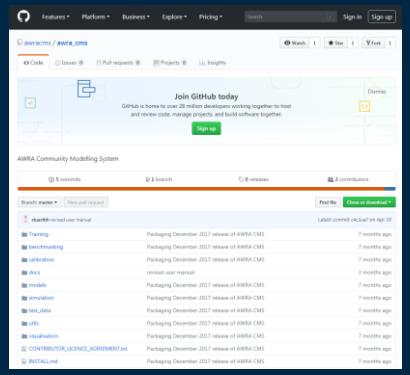
- Modular workflow design
  - Plug and play
  - Infrastructure for connecting new hydrological models
  - Reuse core components / inheritance



- 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

AWRA User Community





Training based on python notebooks:
2016 – 1 course
2017 – 2 course
2018 – 3 courses
2019 – 2 courses
2020 – 1 course
Total 150 trained in CMS so far

140 Registered Users of the AWRA CMS (since 2017):

- 67 Research
- 37 Government – State and Federal
- 15 Consultants / Commercial – which includes entities such as Snowy Hydro, eWater, COFCO Agri, HARC Consulting and Jacobs

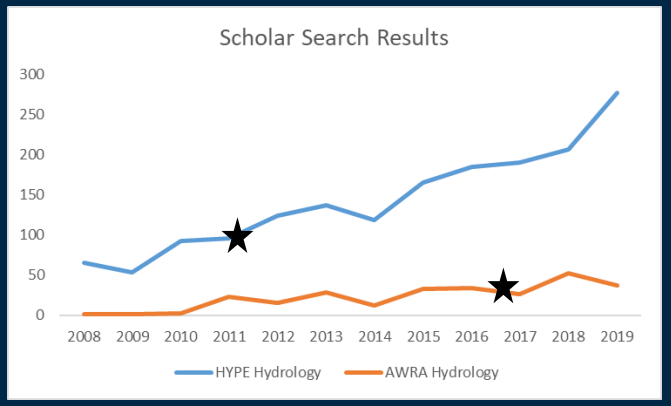
Australian Government
Bureau of Meteorology

How we're building the community

- Scientific Steering Committee (AWRA CMS Reference Group) with good representation nationally
- Regular training
- Python notebooks based training material
- Making installation easier (containerised version)
- Providing some support
- Fund R&D with the model
- Encouraging PhD students to work with the model
- *Training material for undergraduate courses?*

AWRA User Community

- Bureau of Meteorology – Founding partner via WIRADA
- CSIRO – Founding partner via WIRADA
- ANU – Develop a branch model and contribute to the AWRA trunk (assimilation, post-processing)
- University of Melbourne/Monash – Improving groundwater calibration, validation and process descriptions
- University of Queensland – Improving deep drainage and baseflow
- Students at many other universities including UNSW, University of Newcastle, Adelaide University, University of Sydney,



CSIRO

THE UNIVERSITY OF MELBOURNE

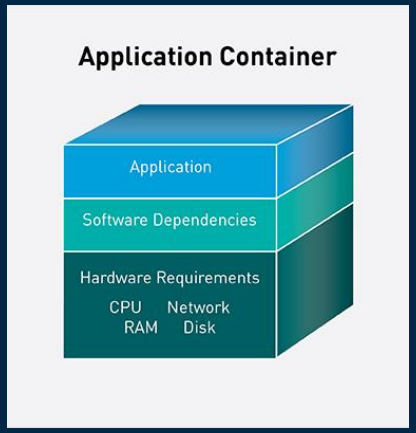
Australian National University

UNSW SYDNEY

AWRA-CMS News!

Containerised Installation

A container encapsulates an application with its own operating environment



What's the appeal?

Some findings using AWRA-CMS as a case study

- Less than 100 lines of code in a docker file vs 700 in bash script
- Can be run on Linux, Mac and Windows
- With good internet speed (BoM) and hardware:
  - Download prebuild image from docker hub and run it in <15 mins (Previous build and setup time ~45 mins for a new user assuming everything goes smoothly!)
- Seamless software builds, tests, packaging, deployment and distribution
- Easy to maintain

https://mapr.com/blog/overview-of-application-containers-part-1-of-4

Keys to Success

- Sponsor organisation(s)
- Scientific Steering Committee
- Open source code
- Training and Support (Wikis, forums, training events)
- Good Documentation
- Easily available training material/tutorials for self help
- Visibility at Conferences, in publications and in international projects
- Flexible models – can be applied in different domains, time-steps, applications
- Time – at least 5 years to build up momentum!

Thanks for your attention

Chantal Donnelly
Head of Water Investigations team
Chantal.donnelly@bom.gov.au
07 3239 8767

Australian Government
Bureau of Meteorology