Development of cloud-based virtual desktop environment for synthesis and analysis for ecosystem science community
Siddeswara Guru, Hoang Nguyen, Shilo Banihit, Matt Mulholland, Kim Olsson & Tim Clancy

Presentation by Siddeswara Guru
Outline

• Motivation
• System Overview
• Application
• Conclusion
Typical Analysis and Synthesis

1. Formulate a problem
2. Define hypothesis
3. Gather data to conduct experiment
4. Test the outcome
5. Analyse data
6. Make inference
7. Publish results
Computational Options

Option 1:
• Tools: Matlab, R, python, statistical packages etc.
• Individual programs run separately to conclude the meta-analysis.

Advantages
• Develop in any desktop environment
• Lower barrier for running analysis

Disadvantages
• Reproducibility and transferability is not that easy
Computational options

Option 2
Use of scientific workflow

• Series of structured interconnected computational activities
• Visual front-end to build experiments using components.
• Components can be implemented in high-level and/or scripting languages.
Computational options

Popular Workflow systems

• Kepler, Taverna, Galaxy, VisTrails, Triana

Advantages:

• GUI interface to compose workflow
• Ability to create components in different programming languages
• Easy to run workflows as distributed tasks
• Ability to track and collect provenance information
• Platform independent re-usable experiments
• Possibility of repeatability and reproducibility
Computational Options

Disadvantages:

• New tool to learn
• Strong informatics focus
Challenge - Research infrastructure

NeCTAR and RDSI

• Cloud-based research infrastructure
• Based on merit-allocation
  • Apply for compute allocation from NeCTAR research cloud
  • Apply for storage allocation from RDSI
• Based on Infrastructure as a Service
  • Compute and storage resource over the Internet
• Users build and manage platform.
No need to do any of these things but able to access the platform via a browser
Virtual Desktop on Browser
Public and Personal workspace

Please visit [http://www.coesra.org.au/#/accesspolicy](http://www.coesra.org.au/#/accesspolicy) for more information.
Publish and share your workflow
Publish your workflow

Publish your workflow
Application
International Union for Conservation of Nature Red lists of Ecosystems
Application of IUCN Red List of Ecosystems Categories and Criteria

**Initial Assessment**
- NE: Describe ecosystem
- DD: Compile information

**Data Type**
- Distribution
- Degradation and disruption

**Data Analysis**
- Gather and process spatial data (satellite images, aerial photos, paper maps, field accounts, historical documents)
- Classify, validate data and create time series of ecosystem maps
- Identify appropriate abiotic (C) and biotic (D) variables, and describe for each one the collapsed state. Estimate the extent and relative severity of degradation (C) or disruption (D) for each variable
- Implement a probabilistic ecosystem model, based on ecosystem compartments, fluxes and processes, to estimate probability of collapse under plausible scenarios of the future

**Apply Criteria**
- Criterion A: Declining distribution
  - Calculate extent over time
    - Past (50 y)
    - Future (next 50 y)
    - Future (any 50 y period)
    - Historic (1750)
- Criterion B: Restricted distribution
  - Calculate current
    - AOO
    - EOO
    - Number of discrete occurrences
- Criterion C: Abiotic degradation
- Criterion D: Biotic disruption
- Criterion E: Probability of collapse
  - Model probability of collapse
    - In 50 y (for CR, EN)
    - In 100 y (for VU)

**Assessment Outcome**
- CR
- EN
- VU
Mountain Ash Forest Ecosystem Risk Assessment

- Unique biodiversity,
- World tallest flowering plants (over 100 m),
- Contribute to water and timber production
- subject to wildfire.
- Apply IUCN Red List of Ecosystem criteria for risk assessment²

IUCN Ecosystem Risk Assessment of Mountain Ash Forests, Southeastern Australia

The projectdir is intended to be the primary workspace for all the output data and results generated during this Risk Assessment.

The Victorian Mountain Ash stags and fire data are restricted datasets. Please contact Professor David Lindenmayer to request access to the datasets.
Contact email: David.Lindenmayer@anu.edu.au
Once you have these files save them into a directory and edit this variable to show the location (directory path) of these files.

Each group of subcriteria is evaluated in the separate workflows and all will be evaluated by this main workflow.

The options are available to:
- ‘evaluate’ (allowing the option to run only one or a few of all five criteria), or
- ‘recreate from source’ (i.e. take raw data inputs and do pre-processing such as polygon union etc). Setting this to NO will result in a faster run, but will only reproduce the final steps of the workflow.

All the available results will be summarised at the end of the workflow and written into the file at this location:

```
outfile
{results/subcriteria_summary_table_showing_all.csv}
```

Sub- Workflow for Criteria A
Workflow

The project directory is intended to be the primary workspace for all the output data and results generated during this Risk Assessment. The workflow is designed to evaluate subcriteria in a hierarchical manner, starting with the primary biotic variable. Each group of subcriteria is evaluated in separate workflows, and all will be evaluated by this main workflow.

The options available to:

- Use a specific library
- Read data from a file
- Apply a function to the data

Ecosystem Risk Estimates


TERN - Terrestrial Ecosystem Research Network
Provenance Information
Governance

• Accessible from Google chrome browser
• For research purpose only
• Keep only research related data in personal and public folder
• We have rights to refuse access to any users.
CoESRA

- Access to cloud-based Linux desktop via a browser
- Virtual desktop comes with Kepler scientific workflow and other tools
- Both personal and public storage space
- Ability to distribute the execution
- Free to use, build and/or execute workflows.

Register/login
Access to virtual desktop
Access your virtual desktop by visiting

www.coesra.org.au

Contact us if your institution is not part of AAF
Project Sponsors

TERN is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy and the Super Science Initiative

collaborators
More information

TERN website:  www.tern.org.au,
               www.coesra.org.au

Thank you  
   s.guru@uq.edu.au