Evaluation of the Vegetation Optimality Model along the North-Australian Tropical Transect using a fully Open Science approach

R.C. Nijzink¹, J. Beringer², L. Hutley², C. Ramakrishnan⁴, R. Roskar⁴, S. Schymanski¹

¹ Luxembourg Institute of Science and Technology, Belvaux, Luxembourg
² University of Western Australia, Crawley, Australia
³ Charles Darwin University, Darwin, NT, Australia
⁴ Swiss Data Science Center, Zurich, Switzerland

Supported by the Luxembourg National Research Fund (FNR) ATTRACT programme (A16/SR/11254288)
Land Surface Models today:

...produce very different results for unclear reasons

**Whitley et al. (2015): Biogeosciences 13**
STATE-OF-THE-ART

Land Surface Models today:
...produce very different results for unclear reasons

(a) Howard Springs

Seasonality underestimated

Fluxes underestimated

Differences due to different structure, parametrization…?
STATE-OF-THE-ART

Land Surface Models today:

...produce very different results for unclear reasons

...commonly rely on past observations

Whitley et al. (2015): Biogeosciences 13
STATE-OF-THE-ART

Whitley et al. (2015): Biogeosciences 13

Land Surface Models today:
...produce very different results for unclear reasons
...commonly rely on past observations
...cannot simulate full response to change due to prescribed properties
Land Surface Models today:

...produce very different results for unclear reasons

...commonly rely on past observations

...cannot simulate full response to change due to prescribed properties

**Whitley et al. (2015): Biogeosciences 13**

And how about reproducibility and repeatability?
VEGETATION OPTIMALITY

Net Carbon Profit:
Difference carbon uptake and carbon costs

Motivation
Hypotheses
Methods
Results
Conclusions
Knowledge graph

Evaporation
Assimilation
Carbon costs
Root uptake
VEGETATION OPTIMALITY

**Net Carbon Profit:**
Difference carbon uptake and carbon costs

**Vegetation Optimality Model**
Optimizes vegetation properties to maximize NCP

Net Carbon Profit: Difference carbon uptake and carbon costs

Motivation
Hypotheses
Methods
Results
Conclusions
Knowledge graph

**Renku workflow**

- Assimilation
- Root uptake
- Evaporation
- Carbon costs
VEGETATION OPTIMALITY

Net Carbon Profit:
Difference carbon uptake and carbon costs

Vegetation Optimality Model
Optimizes vegetation properties to maximize NCP

Net Carbon Profit:
Difference carbon uptake and carbon costs

Vegetation Optimality Model
Optimizes vegetation properties to maximize NCP

No vegetation data needed
VEGETATION OPTIMALITY

Net Carbon Profit:
Difference carbon uptake and carbon costs

Vegetation Optimality Model
Optimizes vegetation properties to maximize NCP

No vegetation data needed

Dynamically adapts to change
VEGETATION OPTIMALITY

Net Carbon Profit:
Difference carbon uptake and carbon costs

Vegetation Optimality Model
Optimizes vegetation properties to maximize NCP

And how about reproducibility and repeatability?

Dynamically adapts to change

No vegetation data needed
VEGETATION OPTIMALITY

Net Carbon Profit:
Difference carbon uptake and carbon costs

Assimilation
Evaporation
Root uptake
Carbon costs

Vegetation Optimality Model
Optimizes vegetation properties to maximize NCP

And how about reproducibility and repeatability?

Renku 連句
Software platform for reproducible and collaborative science

No vegetation data needed
Dynamically adapts to change
Evaporation is quite right!
Evaporation is quite right!

With less data!
Most data comes in at the end for evaluation!
Evaluation of the Vegetation Optimality Model along the North-Australian Tropical Transect using a fully Open Science approach

R.C. Nijzink¹, J. Beringer², L. Hutley³, R. Roskar⁴, S. Schymanski¹

Motivation
State-of-the art
Vegetation Optimality

Hypotheses
VOM
NATT

Methods
Initial Results
Model Comparison
Hydrology and carbon costs
Vegetation dynamics

Results

Conclusions

Knowledge graph

¹ Luxembourg Institute of Science and Technology, Belvaux, Luxembourg,
² University of Western Australia, Crawley, Australia
³ Charles Darwin University, Darwin, NT, Australia
⁴ Swiss Data Science Center, Zurich, Switzerland
Land Surface Models today:

...produce very different results for unclear reasons

...commonly rely on past observations

...cannot simulate full response to change due to prescribed properties

Whitley et al. (2015): Biogeosciences 13

And how about reproducibility and repeatability?
VEGETATION OPTIMALITY

Net Carbon Profit:
Difference carbon uptake and carbon costs

Vegetation Optimality Model
Optimizes vegetation properties to maximize NCP

And how about reproducibility and repeatability?

Renku 連句
Software platform for reproducible and collaborative science

Previous  Home  Next
• Observed vegetation dynamics in tropical savanna sites can be explained by the maximization of **Net Carbon Profit**.

• Optimization of vegetation properties for the **Net Carbon Profit** leads to reduced data requirements for Land Surface Models

• **Hydrological formulation** of Land Surface Models matters for flux exchanges
VEGETATION OPTIMALITY MODEL

Optimized constants:
- Tree cover fraction
- Tree rooting depth
- Grass rooting depth
- Water use strategies

Dynamically optimized variables:
- Grass cover fraction
- Photosynthetic capacity
- Stomatal conductances
- Fine root surface area

Highlights

Root distributions

Tree cover

Grass cover

Tree rooting depth

Grass rooting depth

Knowledge graph

Motivation

Hypotheses

Methods

Results

Conclusions

Renku workflow

Previous  Home  Next
VEGETATION OPTIMALITY MODEL

Getting the hydrology right:
- Groundwater influences root water uptake
- Water balance should be correct
- Modelling for the right reason

Free draining conditions
- Conventional approach
- Large unsaturated zone (Zr = 30m)
- No influence of groundwater table

Dynamic water tables
- Unsaturated zone from 5 – 30m
- Based on max. elevation and stream elevation
- Drainage depends on slopes
- Influence of groundwater tables

Getting the hydrology right:
- Groundwater influences root water uptake
- Water balance should be correct
- Modelling for the right reason

Free draining conditions
- Conventional approach
- Large unsaturated zone (Zr = 30m)
- No influence of groundwater table

Dynamic water tables
- Unsaturated zone from 5 – 30m
- Based on max. elevation and stream elevation
- Drainage depends on slopes
- Influence of groundwater tables
NORTH AUSTRALIAN TROPICAL TRANSECT

- Mean annual rainfall: 500-1800 mm
- Pronounced wet season: Nov-Feb
- Evergreen trees + seasonal grass
- Evaporation and CO$_2$ fluxes >10 years
Renku is an environment for collaborative, reproducible data science

**Concepts**
- Tracking of scientific steps to create data lineage, i.e. a knowledge graph
- Updating of out-dated results
- Tool to re-use or re-run analyses
- Sharing of analyses

**Features**
- Renku is based on:
  - Gitlab
  - JupyterHub
  - Kubernetes
  - Keycloak
  - Common Workflow Language
**Methods**

**Results**

**Conclusions**

**Knowledge graph**

**Motivation**

**Hypotheses**

RENKU  連句

---

*Renku run* makes sure the workflow is tracked

*Renku status* shows if all outputs are generated from the most recent input data

*Renku update* re-runs everything to have all outputs based on the most recent inputs

*Renku log* shows how a file is generated. In other words, it shows the knowledge graph.

---

Go to workflow of this experiment →
INITIAL RESULTS

Reasonable match observed and modelled fluxes

But there is room for improvement...
• Similar results with less data

Graph modified from:
Whitley et al. (2015): Biogeosciences 13
MODEL COMPARISON

- Similar results with less data
- Correct seasonal amplitude in most cases

Hypotheses

Motivation

Results

Conclusions

Knowledge graph

Graph modified from: Whitley et al. (2015): Biogeosciences 13
• Similar results with less data
• Correct seasonal amplitude in most cases
• Improvements still needed:
  → Assimilation too high
  → Especially for drier sites

Graph modified from: Whitley et al. (2015): Biogeosciences 13
Higher cost factor:  
• Improves seasonal signal  
• Especially for wetter areas

Costfactor for water transport is unknown, and may need refinement

Graph modified from: Whitley et al. (2015): Biogeosciences 13
MODEL COMPARISON

Hydrology is parameterized for free draining and non-free draining conditions

- Some improvements
- Some deteriorations
- Mostly similar results

Uniqueness of place?
Robust model?

Graph modified from: Whitley et al. (2015): Biogeosciences 13

Highlights

- Some improvements
- Some deteriorations
- Mostly similar results

Motivation
Hypotheses
Methods
Results
Conclusions
Knowledge graph
IMPROVING HYDROLOGY AND CARBON COSTS

Hydrology differs strongly, but has hardly any influence on fluxes

**Motivation**

**Hypotheses**

**Methods**

**Results**

**Conclusions**

**Knowledge graph**

**Dynamic water table**

**Free drainage**

**Howard Springs**

**Adelaide River**

**Daly Uncleared**

**Dry River**

**Sturt Plains**

- cpcf=1.0 μmol m$^{-3}$ s$^{-1}$, free drainage
- cpcf=1.2 μmol m$^{-3}$ s$^{-1}$, free drainage
- cpcf=1.4 μmol m$^{-3}$ s$^{-1}$, free drainage
- cpcf=2.0 μmol m$^{-3}$ s$^{-1}$, free drainage
- cpcf=1.0 μmol m$^{-3}$ s$^{-1}$, dynamic water table
- cpcf=1.2 μmol m$^{-3}$ s$^{-1}$, dynamic water table
- cpcf=1.4 μmol m$^{-3}$ s$^{-1}$, dynamic water table
- cpcf=2.0 μmol m$^{-3}$ s$^{-1}$, dynamic water table

See more...
Improvements by the hydrology:

- Only small improvements at end of dry season

**Motivation**

**Hypotheses**

**Methods**

**Results**

**Conclusions**

**Knowledge graph**
• Higher values of the water transport cost parameter improve assimilation
• Only small differences for dynamic water tables
PERFORMANCES

- Assimilation increasingly over-estimated for drier areas
- Evaporation still okay

![Graph showing PERFORMANCES with red and blue markers indicating evaporation and assimilation respectively, with increasing dryness indicated by an arrow pointing to the right.]
VEGETATION DYNAMICS

- Temporal signal largely reproduced
- Timing improves for higher cost factor
- Higher minimum cover for non-freely draining conditions
CONCLUSIONS

- Optimizing for the **Net Carbon Profit** leads to similar vegetation dynamics as observed

- Similar performances as conventional models, with **less data**

- Not a clear influence of the **hydrological formulation**

- **Cost factor** for water transport needs to be refined

- Reproducible science with **Renku!**
Knowledge graph

Data

Tasks

Tools

Motivation

Hypotheses

Methods

Results

Conclusions

Most data comes in at the end for evaluation!
CARBON COSTS

- **Root respiration** is a function of respiration rate ($c_{R_n} \text{ mol s}^{-1} \text{ m}^{-3}$), fine root radius ($r, \text{ m}$), root surface area per unit ground area ($S_{Ar}, \text{ m}^2 \text{ m}^{-2}$):
  \[ R_r = c_{Rr} \left( \frac{r}{2} S_{Ar} \right) \]

- **Leaf area costs** are a function of vegetated fraction ($M_A, -$), clumped leaf area index (2.5, -), average carbon investment ($0.22 \mu\text{mol s}^{-1} \text{ m}^{-2}$):
  \[ R_l = 2.5 \times 0.22 \mu\text{mol s}^{-1} \text{ m}^{-2} M_A \]

- **Water transport costs** are a function of rooting depth ($y_r$), vegetated fraction ($M_A, -$) and a cost factor ($c_{pcff}, \text{ mol s}^{-1} \text{ m}^{-3}$):
  \[ R_v = c_{pcff} * M_A y_r \]

The cost factor $c_{pcff}$ is rather unknown, and may need refinement.
Improvements by the hydrology:
- Dynamic groundwater tables improve evaporation at end of dry season
- Still just small improvements
Improvements by the hydrology:

- Dynamic groundwater tables do not show a strong improvement
• Higher values of the water transport cost parameter improve assimilation
• Dynamic groundwater tables do not help for assimilation
Improvements by the hydrology:

- Dynamic groundwater tables do not show strong improvements
HYDROLOGY AND CARBON COSTS

- Higher values of the water transport cost parameter improve assimilation
- Dynamic groundwater tables do not help for assimilation
• Higher values of the water transport cost parameter improve assimilation
• Dynamic groundwater tables do not help for assimilation
IMPROVING HYDROLOGY AND CARBON COSTS

- Higher values of the water transport cost parameter improve assimilation
- Dynamic groundwater tables do not help for assimilation
Higher values of the water transport cost parameter improve assimilation.
Dynamic groundwater tables do not help for assimilation.
PERFORMANCES

- Assimilation increasingly over-estimated for drier areas
- Evaporation still okay

Increasing dryness

HowardSprings   AdelaideRiver   DalyUncleared   DryRiver   SturtPlains

- Assimilation increasingly over-estimated for drier areas
- Evaporation still okay

See more ...

Rel. Err. Annual Means
Rel. Err. Mean Dry Season
Rel. Err. Mean Wet Season
Kling-Gupta Efficiency
PERFORMANCES

- Assimilation increasingly over-estimated for drier areas
- Evaporation still okay
PERFORMANCES

- Assimilation increasingly over-estimated for drier areas
- Evaporation still okay

**Increasing dryness**

![Graph showing PERFORMANCES](image)

- **Kling-Gupta Efficiency**
- **Rel. Err. Annual Means**
- **Rel. Err. Mean Dry Season**
- **Rel. Err. Mean Wet Season**
- **Kling-Gupta Efficiency**
VEGETATION DYNAMICS

- Temporal signal largely reproduced
- Timing improves for higher cost factor
- Similar minimum cover for dynamic water tables

Frac. Cover. fPar

- $\text{cpcff}=1.0 \ \mu\text{mol m}^3\ \text{s}^{-1}$
- $\text{cpcff}=1.2 \ \mu\text{mol m}^3\ \text{s}^{-1}$
- $\text{cpcff}=1.4 \ \mu\text{mol m}^3\ \text{s}^{-1}$
- $\text{cpcff}=2.0 \ \mu\text{mol m}^3\ \text{s}^{-1}$

Better timing

Similar minimum

Free draining conditions

Dynamic water table
**VEGETATION DYNAMICS**

- Temporal signal largely reproduced
- Timing improves for higher cost factor
- Similar minimum cover for dynamic water tables

**Motivation**

**Hypotheses**

**Methods**

**Results**

**Conclusions**

**Knowledge graph**
VEGETATION DYNAMICS

- Temporal signal largely reproduced
- Timing improves for higher cost factor
- Higher minimum cover for non-freely draining conditions

Frac. Cover, fPar
- \( \text{cpcff} = 1.0 \, \mu\text{mol m}^{-3} \, \text{s}^{-1} \)
- \( \text{cpcff} = 1.2 \, \mu\text{mol m}^{-3} \, \text{s}^{-1} \)
- \( \text{cpcff} = 1.4 \, \mu\text{mol m}^{-3} \, \text{s}^{-1} \)
- \( \text{cpcff} = 2.0 \, \mu\text{mol m}^{-3} \, \text{s}^{-1} \)

Better timing
Higher minimum

Howard Springs
Adelaide River
Daly Uncleared
Dry River
Sturt Plains

Free draining conditions
Dynamic water table
VEGETATION DYNAMICS

- Temporal signal largely reproduced
- Timing improves for higher cost factor
- Higher minimum cover for non-freely draining conditions

Previous

Home

Next

Motivation

Hypotheses

Methods

Results

Conclusions

Knowledge graph

VEGETATION DYNAMICS

- Temporal signal largely reproduced
- Timing improves for higher cost factor
- Higher minimum cover for non-freely draining conditions

Howard Springs

Adelaide River

Daly Uncleared

Dry River

Sturt Plains
**Renku log** shows how a file is generated. In other words, it shows the knowledge graph.
Renku status shows if all outputs are generated from the latest inputs. Renku log shows how a file is generated. In other words, it shows the knowledge graph. Renku update re-runs everything to have all outputs based on the most recent inputs. Renku run makes sure the workflow is tracked.
Renku log shows how a file is generated. In other words, it shows the knowledge graph.

Renku status shows if all outputs are generated from the most recent input data.

Renku update re-runs everything to have all outputs based on the most recent inputs.

Renku run makes sure the workflow is tracked.

Go to workflow of this experiment →
Renku run makes sure the workflow is tracked.

-renko@ERIN-RNI-30243:~/renku_egu
remko@ERIN-RNI-30243:~/renku_egu
renku log --format dot evap_adelaide.png | dot -Tpng > ../.../knowledge_graphs/evap_adelaide.png

Show knowledge graph
Atmospheric CO$_2$ levels are needed as input for the VOM model. Therefore, weekly data is taken from the Mauna Loa observatory.
Atmospheric CO$_2$ levels are needed as input for the VOM model. Therefore, weekly data is taken from the Mauna Loa observatory.
Atmospheric CO$_2$ levels are needed as input for the VOM model. Therefore, weekly data is taken from the Mauna Loa observatory.
Meteorological data are needed as input for the VOM model. Data is taken from the Australian Silo weather office.
Meteorological data are needed as input for the VOM model. Data is taken from the Australian Silo weatheroffice.
Meteorological data are needed as input for the VOM model. Data is taken from the Australian Silo weatheroffice.
CREATE MODEL INPUT

The meteorological data and the atmospheric CO₂ data need to be processed and formatted in order to serve as input for the model.
The meteorological data and the atmospheric CO\textsubscript{2} data need to be processed and formatted in order to serve as input for the model.
The Shuffled Complex Evolution algorithm is used in order to derive the vegetation properties that maximize NCP.

- Sample s points
- Rank points
- Partition into complexes
- Evolve complex
- Shuffle complexes
- Check convergence

The Shuffled Complex Evolution algorithm is used in order to derive the vegetation properties that maximize NCP.

The model runs with the 5% highest NCP are selected in order to construct uncertainty bounds.

To assess the model performance, several independent datasets are used:

- DINGO fluxdata
- Fraction vegetation cover from fPar

The performances are assessed by:

- Timeseries with uncertainties
- Relative errors annual and seasonal means
- Kling-Gupta efficiencies
- Residuals
The model runs with the 5% highest NCP are selected in order to construct uncertainty bounds.

To assess the model performance, several independent datasets are used:

- DINGO fluxdata
- Fraction vegetation cover from fPar

The performances are assessed by:

- Timeseries with uncertainties
- Relative errors annual and seasonal means
- Kling-Gupta efficiencies
- Residuals
Fluxes are derived with the DINGO algorithm from the flux towers at the study sites.

HowardSprings

AdelaideRiver

DalyUncleared

DryRiver

SturtPlains

Create renku dataset
Add data
Back to knowledge graph
Fluxes are derived with the DINGO algorithm from the flux towers at the study sites.

```bash
remko@ERIN-RNI-30243:~/renku_egu$ renku dataset create DINGO
Creating a dataset ... OK
remko@ERIN-RNI-30243:~/renku_egu$
```
Fluxes are derived with the DINGO algorithm from the flux towers at the study sites.

DINGO

HowardSprings
AdelaideRiver
DalyUncleared
SturtPlains

Fluxes are derived with the DINGO algorithm from the flux towers at the study sites.
The Enhanced Vegetation Index from MODIS is used to compare the modelled and observed vegetation dynamics.

Fractional cover derived from satellite observed fPAR are used to compare with modelled vegetation cover.
FRACTIONAL COVER

The Enhanced Vegetation Index from MODIS is used to compare the modelled and observed vegetation dynamics.

Create renku dataset
Add data
Back to knowledge graph
The Enhanced Vegetation Index from MODIS is used to compare the modelled and observed vegetation dynamics.

Fractional cover derived from satellite observed fPAR are used to compare with modelled vegetation cover.
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
Switching the vegetation properties for the two hydrological situations leads to similar results:

- Hydrological formulation not of influence
HYDROLOGY AND CARBON COSTS

Motivation
Hypotheses
Methods
Results
Conclusions

Knowledge graph

Howard Springs
Adelaide River
Daly uncleared
Dry River
Sturt Plains
HYDROLOGY AND CARBON COSTS

[Graph showing data on HYDROLOGY AND CARBON COSTS]
HYDROLOGY AND CARBON COSTS

Motivation
Hypotheses
Methods
Results
Conclusions
Knowledge graph