Seasonal precipitation forecasting for the Melbourne region using a Self-Organizing Maps approach

Ross Pidoto, Markus Wallner, Uwe Haberlandt

Leibniz Universität Hannover

1. Motivation & Objectives

- Australia experiences often long periods of drought.
- Reduced runoff leads to water restrictions within cities. Need for better prediction methods for water resources planning.
- ***Problem***: High inter- and intra-annual variability of P. Melbourne lies in a zone less effected by the SOI.
- ***Solution***: seasonal rainfall forecasting model based on a SOM (fig. 3) using gridded SST data, antecedent precipitation and climate indices as training data.

2. Study Area & Data

- **Target variable**: seasonal precipitation anomaly
  - Three month moving average
  - Arithmetic mean over study area
- **Predictor variables:**
  - Climate indices: SOI, N3.4, DMI, ONI, SOI phase
  - Antecedent P, drought index, lowest 10% index
  - Gridded aggregated SST anomalies:
    - 210 cells (2° x 2°) around south eastern Australia
    - 100 cells (12° x 6°) across the tropical Pacific/Indian

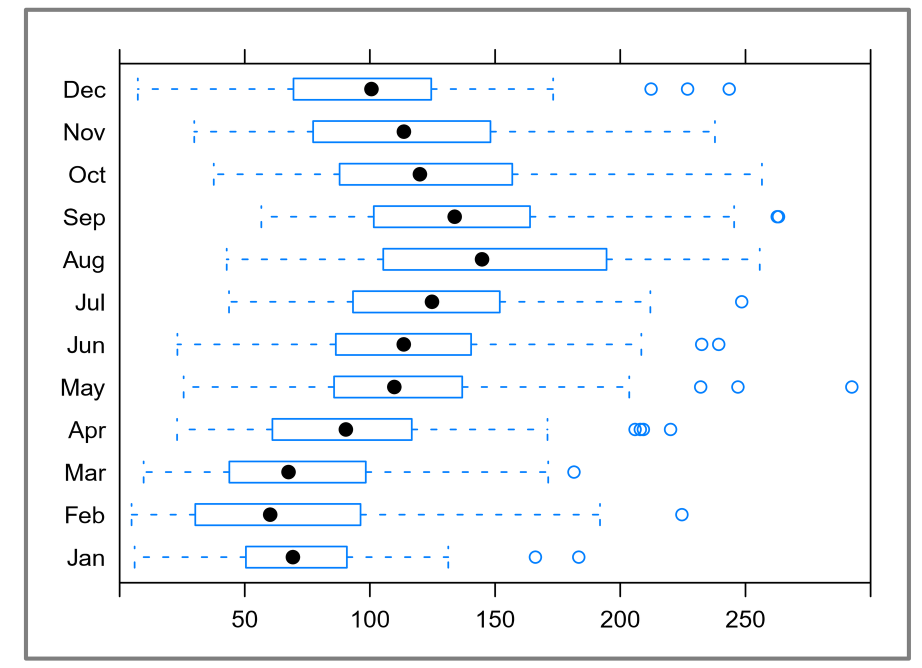

Figure 1: Mean pcp. (mm) over study area

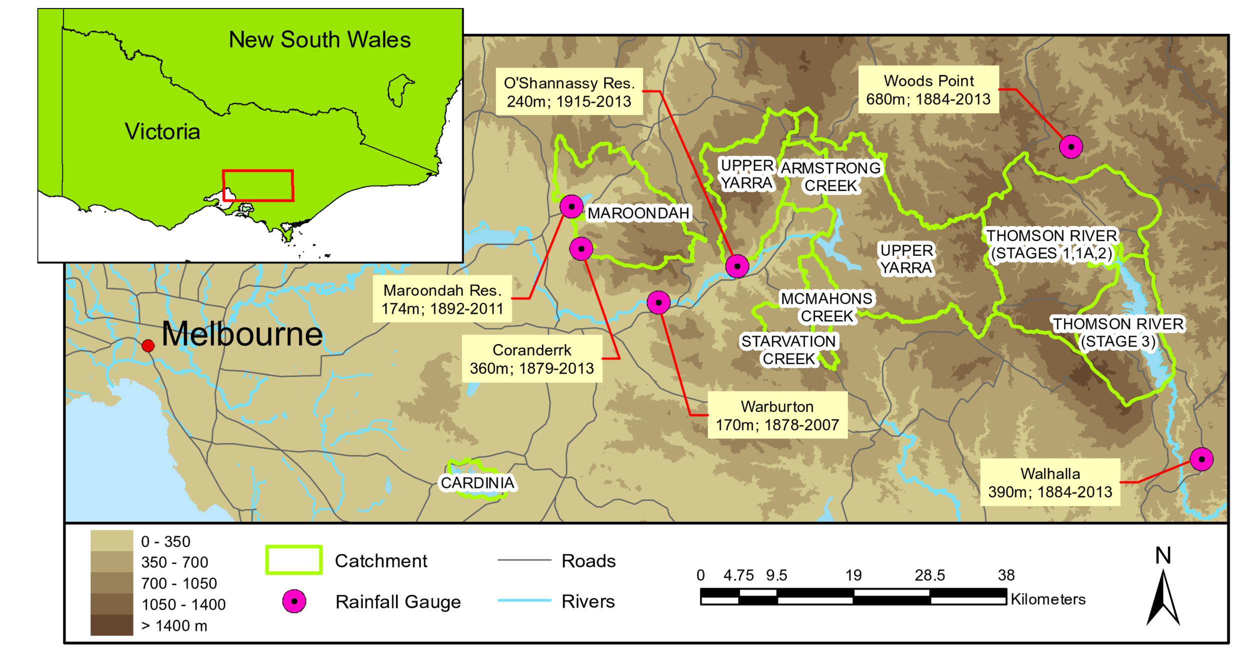

Figure 2: Study area outlining rain gauge locations

3. Methodology

Predictor selection via Random Forests
Breiman (2001)

- SST datasets at a lead of 3 months (Fig. 4)
- Climate indices at leads of 3-12 months
- Each season was assessed separately, plus all observations.

Self-Organizing Map (SOM)
Kohonen (1990)

- 5 models: the four seasons and annual
- Hexagonal grid, large neuron count
- Stepwise predictor testing step to determine optimal predictor ensemble

Assessment of model performance

- Leave-one-out cross validation
- Performance criteria:
  - Pearson's correlation, R
  - Nash–Sutcliffe model efficiency (NSE)
  - Root mean square error (RMSE)

References: Breiman, L. (2001) Random Forests. *Machine Learning*, 45(1):5-32.
Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, 78(9):1464-1480.
Mekanik, F., et al. (2013). Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. Journal of Hydrology, 503:11–21.

4. Results

- Final predictor variable selction:
  - Annual: **13** vars – P_3, SOI_3, $N3.4_{1\&3}$, 2 SST cells, drought, $P_{10\%}$
  - Spring: **9** vars – P_3, SOI_3, DMI, N3.4, 4 SST cells, $P_{10\%}$
  - Summer: **8** vars – P_3, DMI, $N3.4_3$, 3 SST cells, drought , $P_{10\%}$
  - Autumn: **6** vars – P_3, $N3.4_{1\&3}$, 1 SST cell, drought , $P_{10\%}$
  - Winter: **12** vars – SOI_3, DMI_3, $N3.4_{1\&3}$, 6 SST cells, drought
  - Most lead times (73%) were at the minimum of 3 months
- Final model performance:

Criteria	Annual	Spring	Summer	Autumn	Winter
Correlation, r	0.54	**0.61**	0.52	0.41	0.57
NSE	0.14	**0.31**	0.10	0.00	0.22
RMSE [mm]	29.0	28.7	**24.9**	29.8	29.7
RMSE / SD_{obs}	0.92	**0.83**	0.95	1.00	0.88

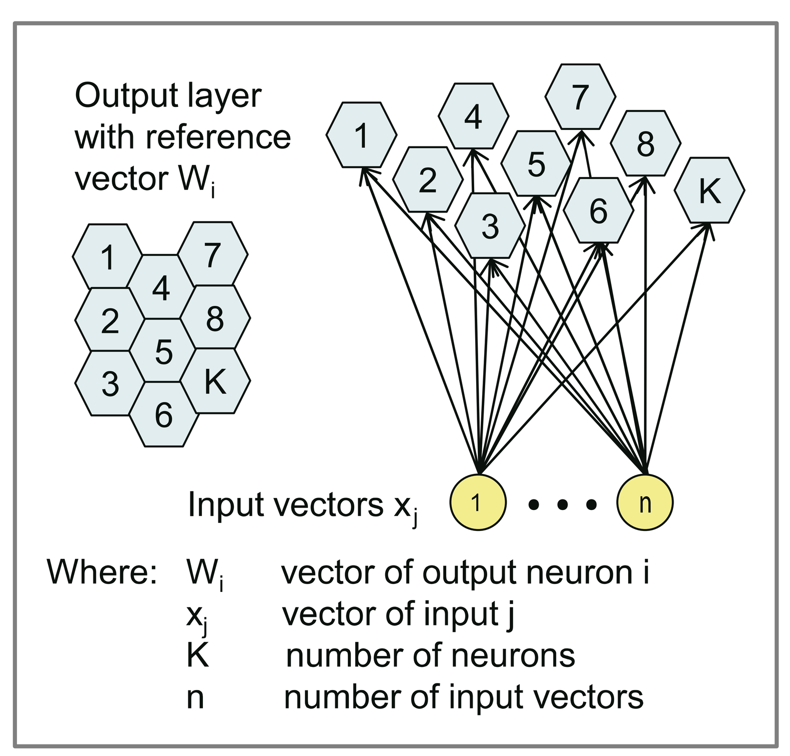

Figure 3: Schematic of a SOM

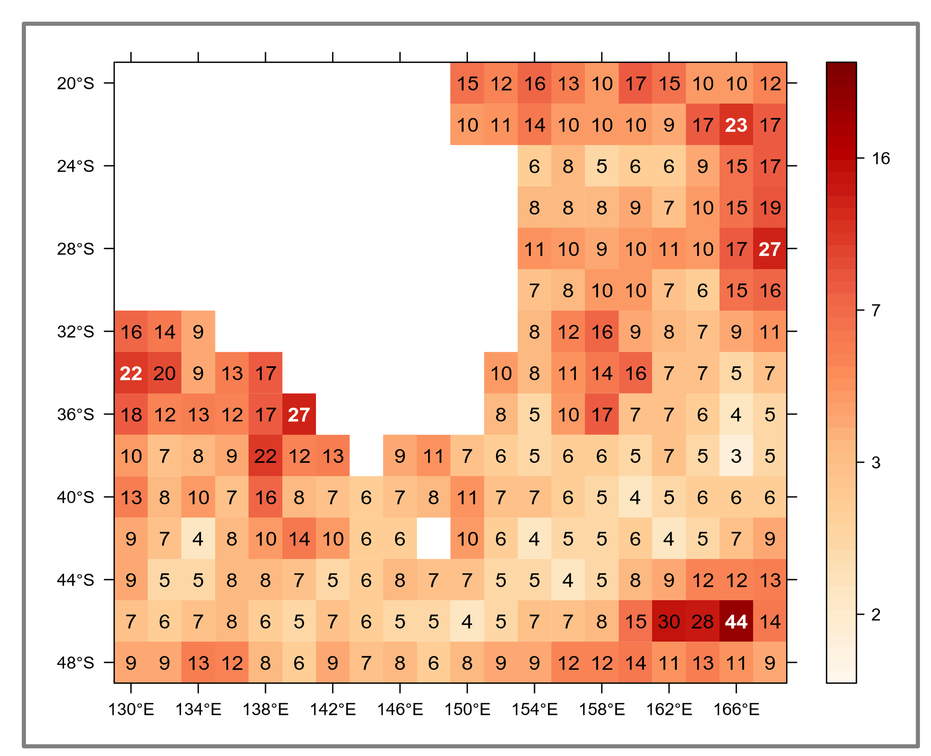

Figure 4: Example output of the Random Forest predictor selection – spring, local SST.

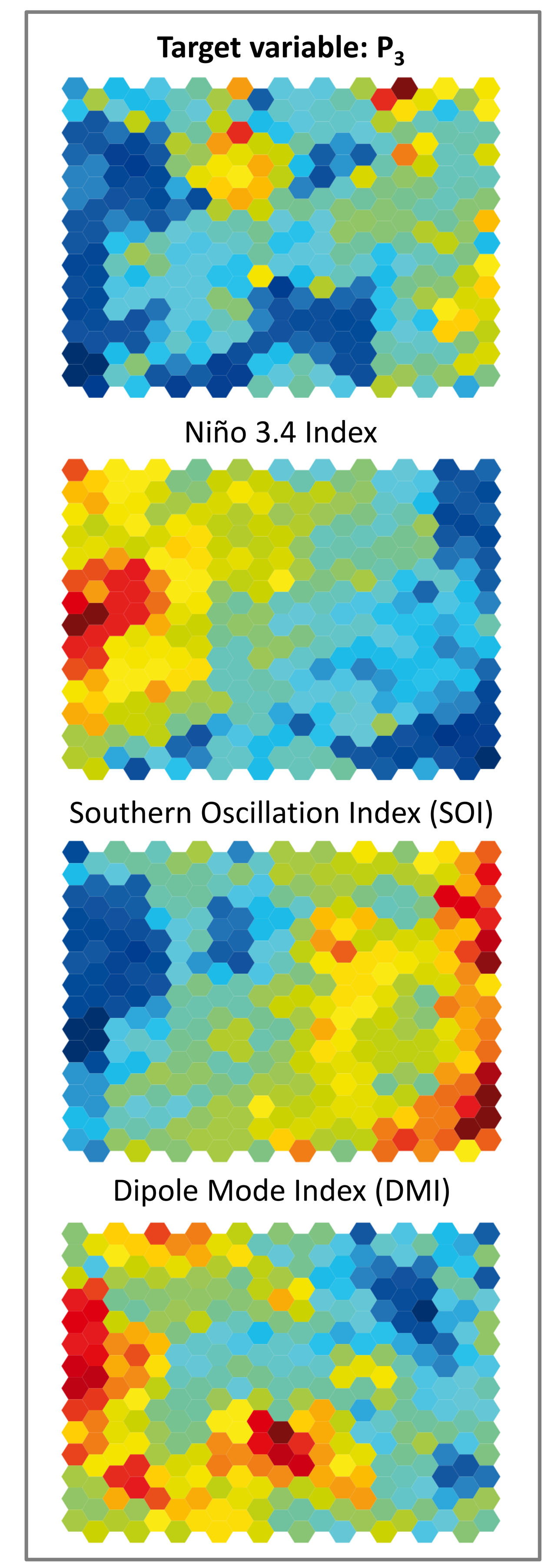

Figure 5: Component planes of the spring SOM. Red indicates high values, blue low.

5. Summary and Conclusions

- Results obtained were favourable compared to other studies (Mekanik 2013)
- The SOM had difficulties handling the large number of predictor variables. This also led to high computational times.
- The SST dataset was relatively ineffective compared to the climate indices. The fixed SST variable lead time of 3 months is not realistic on a global scale. The climate indices were often with leads > 3 months.
- Future inclusion of a step to reduce collinearities.

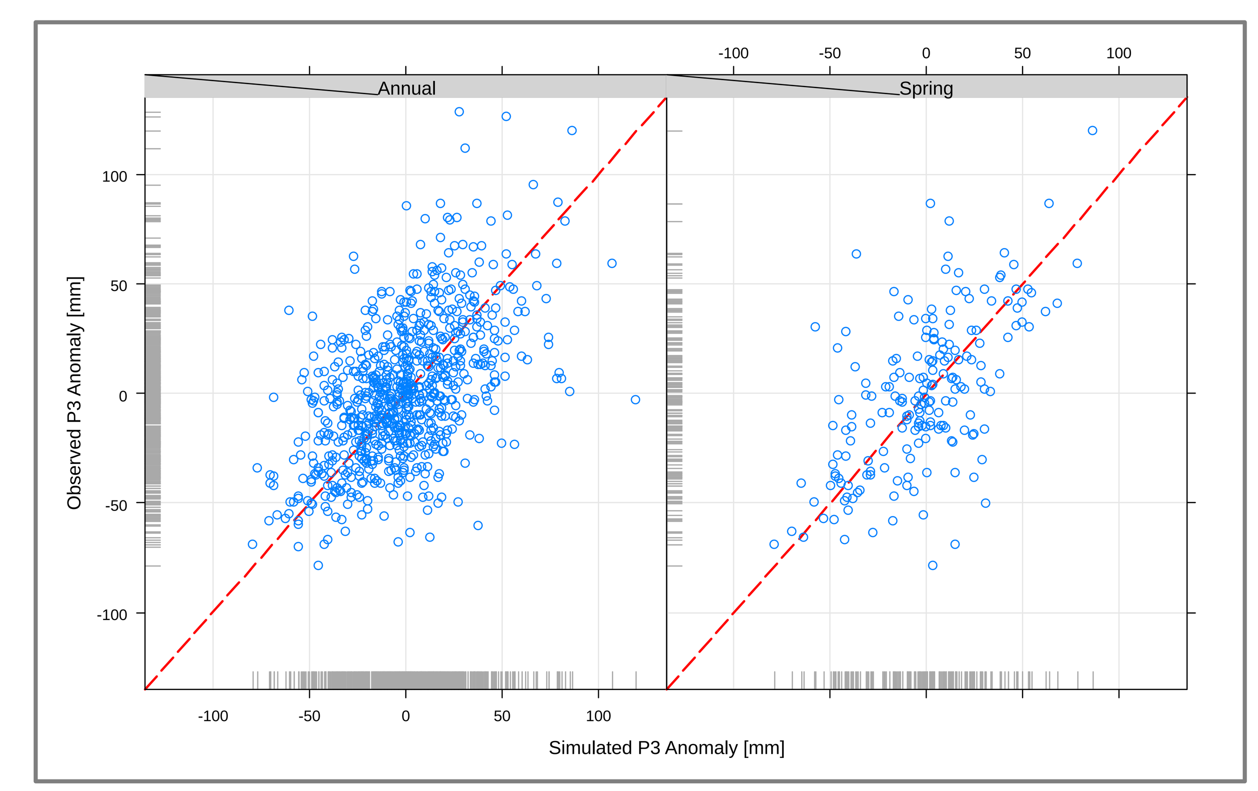

Figure 6: Scatter plots of the observed seasonal precipitation versus the simulated seasonal precipitation for both the spring and annual SOM models.

EGU General Assembly, April 23 – 28, 2017, Vienna

Institute of Hydrology and Water Resources Management
pidoto@iww.uni-hannover.de