STAHY 2019

# Drought prediction for improved water resource management: A wavelet-based system prediction approach

Ze Jiang[1], Ashish Sharma[1] and Fiona Johnson[1]

[1] Water Research Centre, School of Civil and Environmental Engineering, UNSW Sydney

# Why cannot use wavelets for prediction?

Wavelets need the information from “future” – does it make sense to use the future characterise the past?

But we “know” the future using GCM (Global Climate Model) simulations – the aim is to characterize it well!

# Key Research Questions

- How to characterise drought using GCM simulated data?
- How to extract useful information?
- How to use GCM future projections?

# How to extract useful information?

The hypothesis:

If the spectral variance structure of the predictor is similar to that of the response, the predictive model using that predictor will exhibit better accuracy than otherwise.

# How to extract useful information?

Background: Wavelet Transform

**Additive Decomposition:**

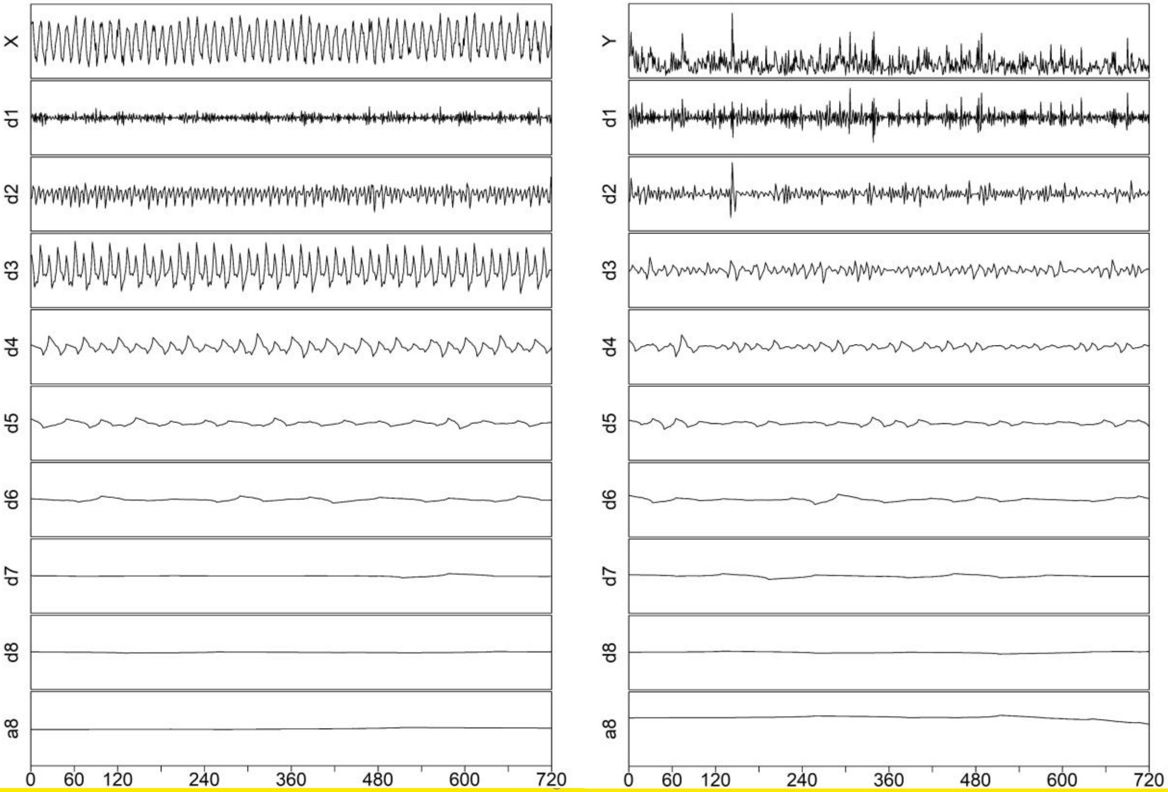
**(Multiresolution Analysis, MRA)**

$$\boldsymbol{X} = \sum_{j=1}^{J} \boldsymbol{d}_j + \boldsymbol{a}_J$$

$$\sigma_X^2 = \sum_{j=1}^{J} \sigma_{d_j}^2 + \sigma_{a_J}^2$$

Predictor Variable X (EPT)

Target Response Y (Rainfall)




water@
UNSW
water research centre

# How to extract useful information?

**MRA:**

$$\boldsymbol{X} = \sum_{j=1}^{J} \boldsymbol{d_j} + \boldsymbol{a_J}$$

$$\sigma_X^2 = \sum_{j=1}^{J} \sigma_{d_j}^2 + \sigma_{a_J}^2$$

**Matrix form:**

$$\boldsymbol{X} = \tilde{\boldsymbol{R}}\boldsymbol{I}$$

where $\tilde{\boldsymbol{R}}$ is normalized reconstructions matrix.

$$\boldsymbol{R} = [\boldsymbol{d_1}, \ldots, \boldsymbol{d_J}, \boldsymbol{a_J}] \qquad \boldsymbol{I} = [\sigma_{d_1}, \ldots, \sigma_{d_J}, \sigma_{a_J}]^{\mathrm{T}}$$

**What we are looking for:**

$$\boldsymbol{X}' = \tilde{\boldsymbol{R}}\boldsymbol{\alpha}$$

$$\boldsymbol{\alpha} = \sigma_X \tilde{\boldsymbol{C}}$$

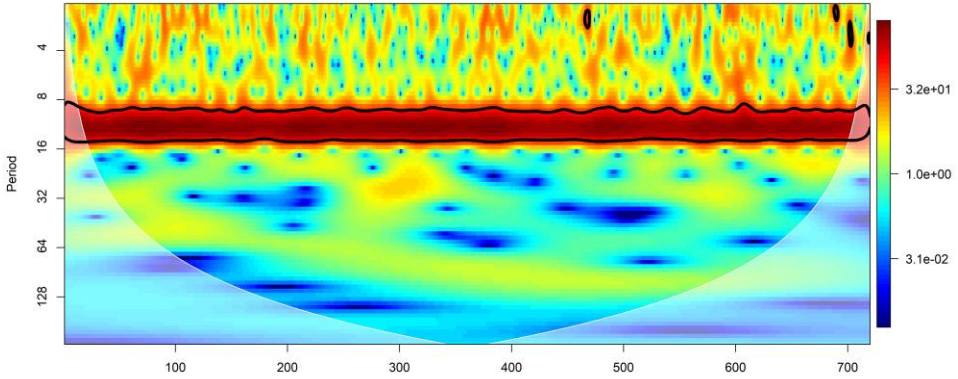
where $\tilde{\boldsymbol{C}}$ is the normalized covariance matrix for the variable set ($\boldsymbol{Y}$, $\tilde{\boldsymbol{R}}$)

$$\boldsymbol{C} = \frac{1}{n-1} \boldsymbol{Y}^T \tilde{\boldsymbol{R}} = \left[ S_{Y\tilde{d}_1}, \ldots, S_{Y\tilde{d}_J}, S_{Y\tilde{a}_J} \right]$$

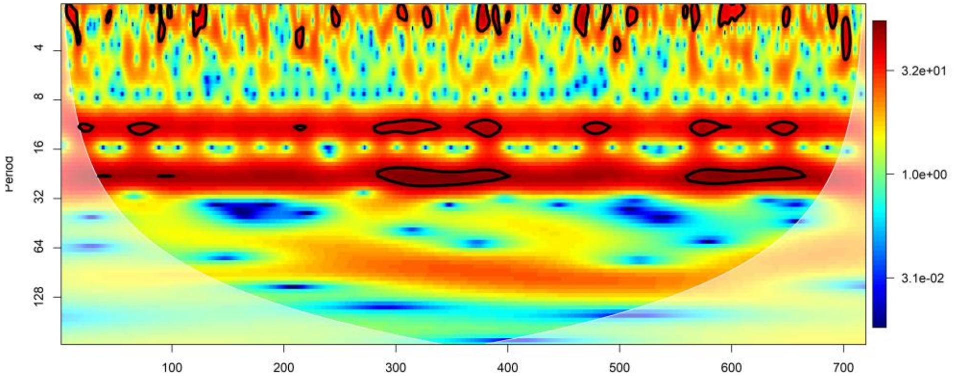
$$RMSE_{\min} = \sqrt{\frac{n-1}{n}(\sigma_Y^2 - \|\boldsymbol{C}\|^2)}$$

# How to extract useful information?

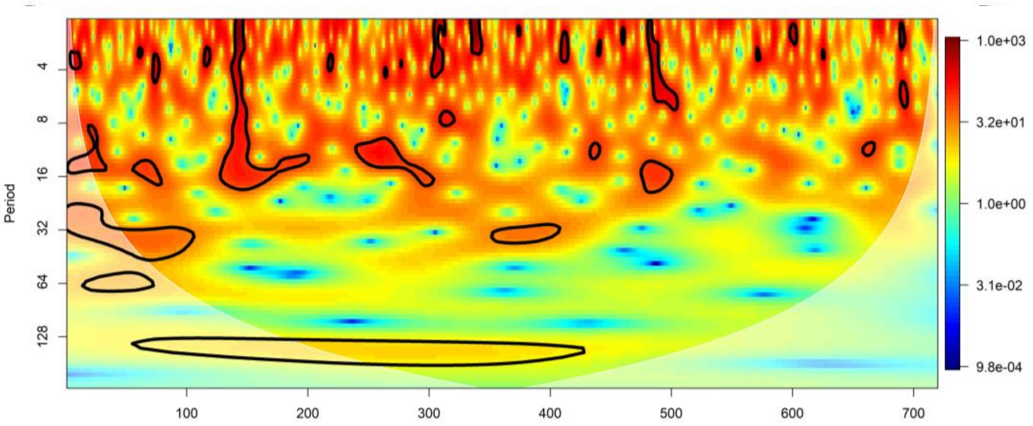
Predictor variable (Before variance transformation, VT)



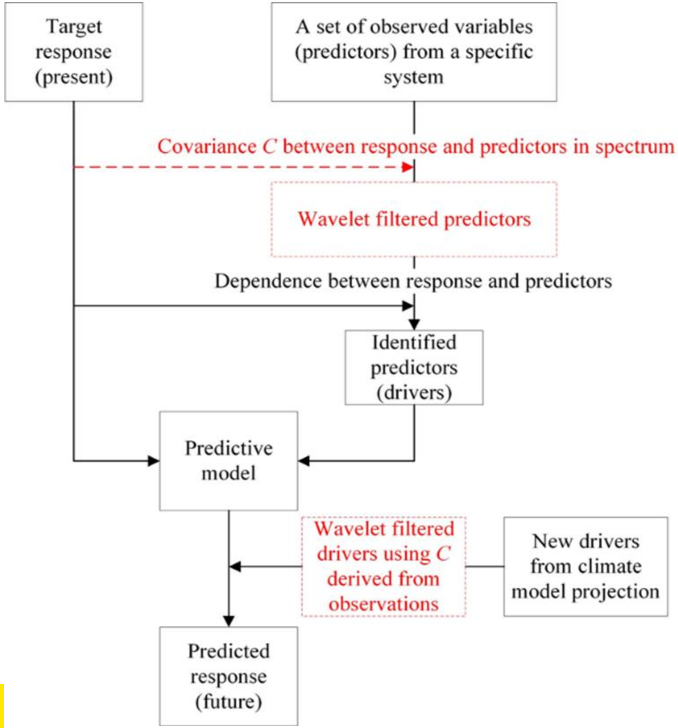

Predictor variable (After variance transformation, VT)




Target Response



# How to use GCM future projections?




The proposed wavelet-based system prediction framework

Step 1 - identify best possible drivers from large numbers of climatic variables (inputs)

Step 2 - form a predictive model based on the identified drivers, estimate the model parameters that best fit to the data

Step 3 - predict the system response for new inputs.

Wavelets: obtain filtered new climate variables

# Results – Synthetic example

A dynamic example (Rössler system):

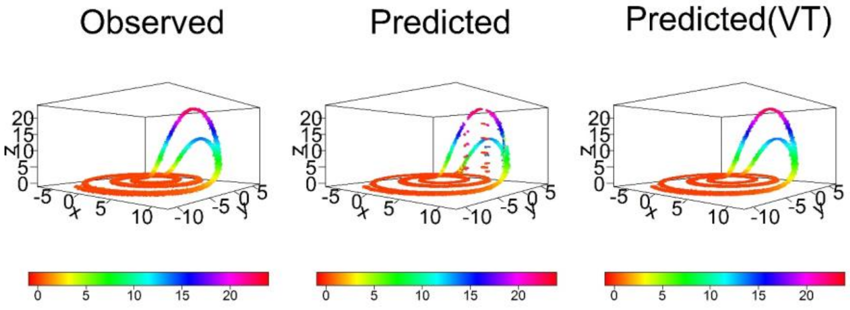
$$\dot{x} = -y - z,$$

$$\dot{y} = x + ay,$$

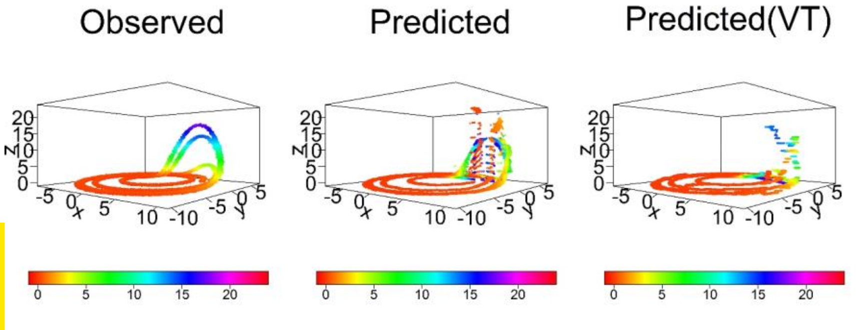
$$\dot{z} = b + z(x - c).$$

Use x and y to predict z

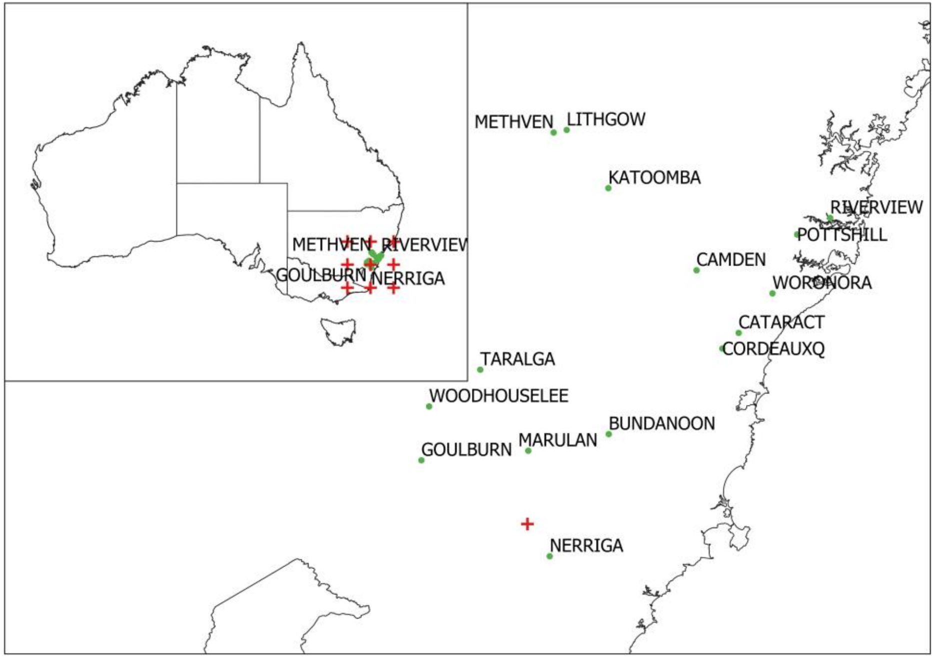
**Calibration: RMSE** = 0.113 against 1.189




**Validation: RMSE** = 2.550 against 4.493



# Results – Real example




- Sydney Region Rainfall Stations
- NCEP-NCAR Reanalysis grid

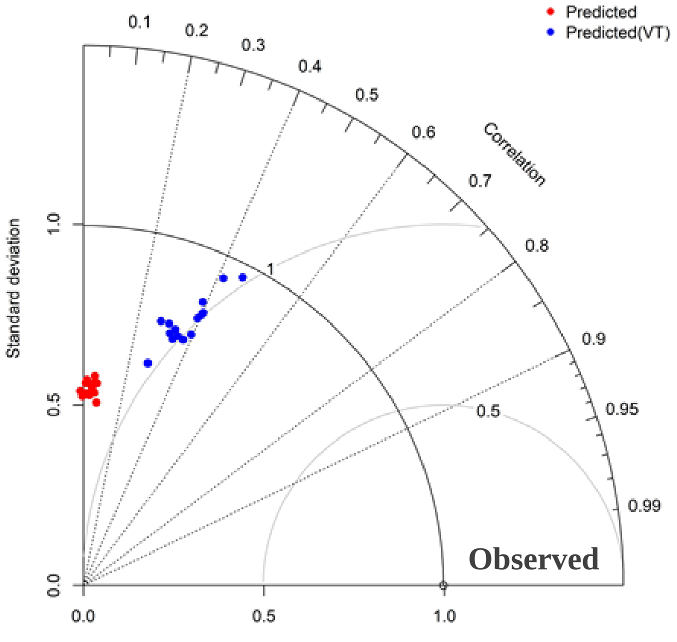
**Target response: Drought Index (SPI12)**

| Predictor No. | Predictor Name |
|---|---|
| 1 | Geopotential heights (m) at 925 hPa (GPH@925) |
| 2 | Temperature depression (degree C) at 700 hPa (TDP@700) |
| 3 | Temperature depression (degree C) at 500 hPa (TDP@500) |
| 4 | Equivalent potential temperature (Kelvin K) at 500 hPa (EPT@500) |
| 5 | Zonal Wind (m/s) at 500 hPa (UWND@500) |
| 6 | Meridional Wind (m/s) at 500 hPa (VWND@500) |
| 7 | N–S gradient of mean sea level pressure (NS-MSLP) |

# Results – Real example

Taylor diagram of evaluating model performance by
the standard deviation, centered RMSE and correlation coefficient.

## Cross-Validation



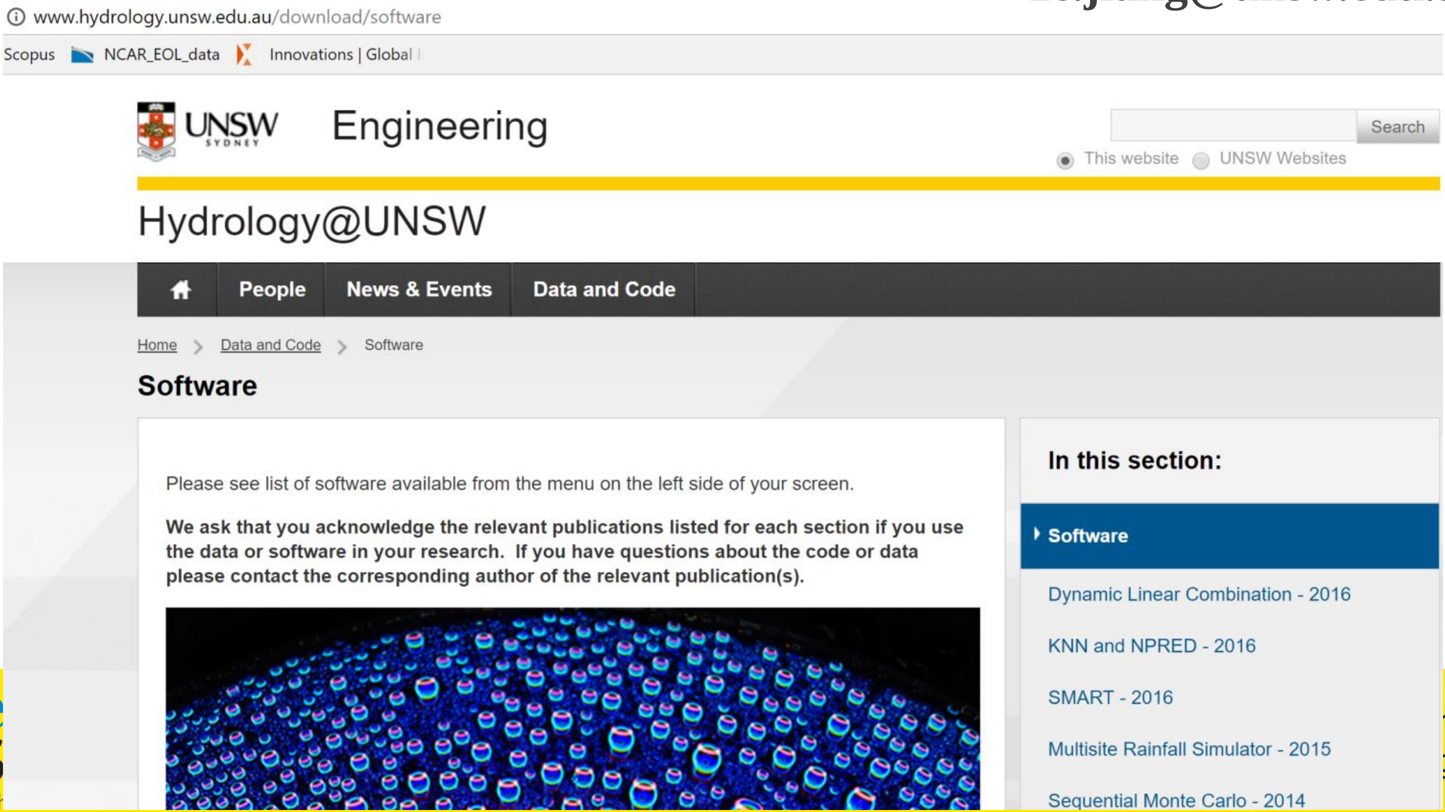

| Station No. | Predicted | Predicted (VT) | Reduced RMSE |
|---|---|---|---|
| **1** | 1.12 | 1.02 | 0.10 |
| **2** | 1.09 | 1.01 | 0.08 |
| **3** | 1.12 | 1.00 | 0.12 |
| **4** | 1.14 | 1.01 | 0.13 |
| **5** | 1.12 | 1.01 | 0.11 |
| **6** | 1.11 | 1.06 | 0.05 |
| **7** | 1.13 | 1.01 | 0.12 |
| **8** | 1.14 | 1.04 | 0.10 |
| **9** | 1.12 | 1.05 | 0.06 |
| **10** | 1.12 | 1.03 | 0.09 |
| **11** | 1.13 | 1.04 | 0.09 |
| **12** | 1.12 | 1.02 | 0.10 |
| **13** | 1.14 | 1.08 | 0.06 |
| **14** | 1.13 | 1.02 | 0.11 |
| **15** | 1.11 | 1.04 | 0.07 |

# Conclusions

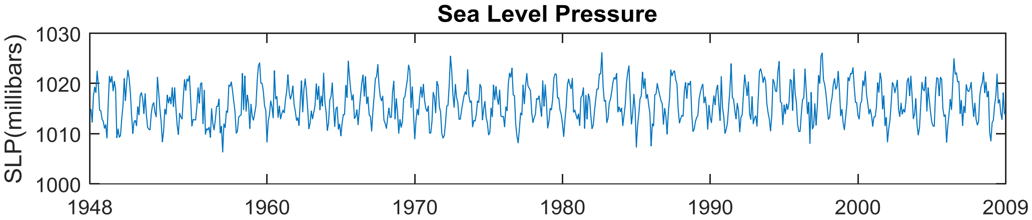
- A unique variance transformation is identified for each predictor variable that explains maximal information in the corresponding response.
- Results of a dynamic example and a real application show clear improvements in predictability compared to the use of unfiltered predictors.
- It is a generic method and not limited to the hydro-climatology system.

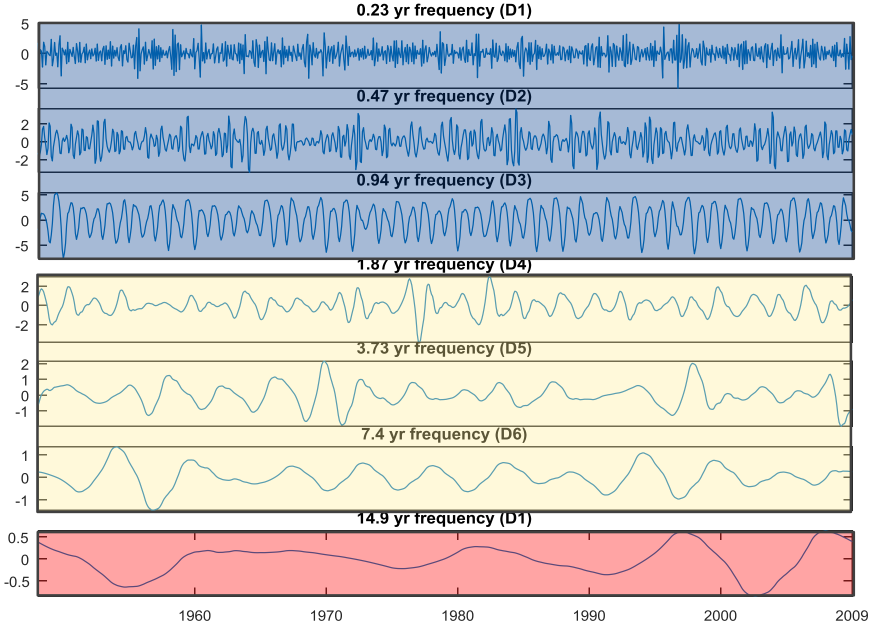
The open-source R-package WASP is available for download from the following website “**http://hydrology.unsw.edu.au/download/software/** WASP”. Source codes are available, along with help-files and example real datasets used to generate the outcomes reported.

**ze.jiang@unsw.edu.au**



# Thank you!


Sea Level Pressure
SLP(millibars)
1030
1020
1010
1000
1948
1960
1970
1980
1990
2000
2009


0.23 yr frequency (D1)
0.47 yr frequency (D2)
0.94 yr frequency (D3)
1.87 yr frequency (D4)
3.73 yr frequency (D5)
7.4 yr frequency (D6)
14.9 yr frequency (D1)
1960
1970
1980
1990
2000
2009