

Forecast Linear Augmented Projection (FLAP)

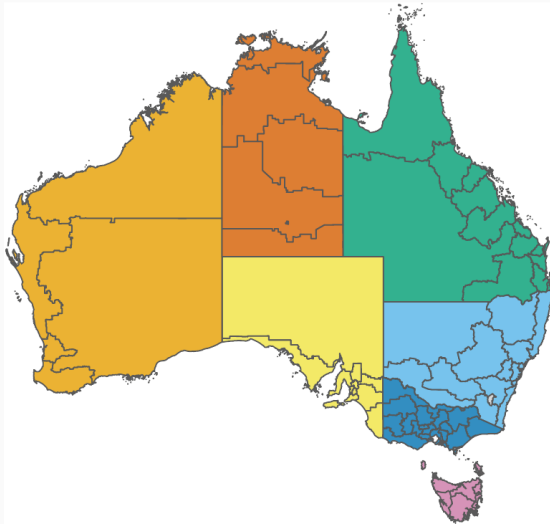
Fin Yang, **George Athanasopoulos**, Rob J Hyndman, Anastasios Panagiotelis



MONASH University

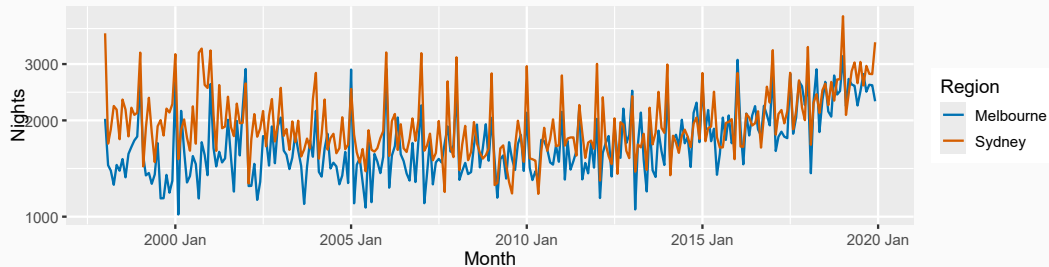
Photo by Stanley Cheung on Unsplash

Australian tourism regions

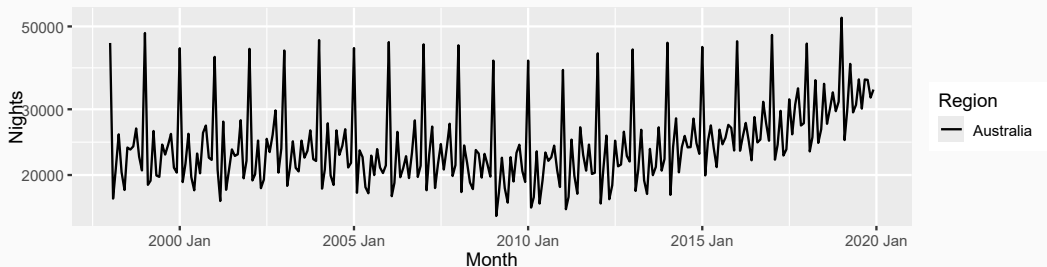
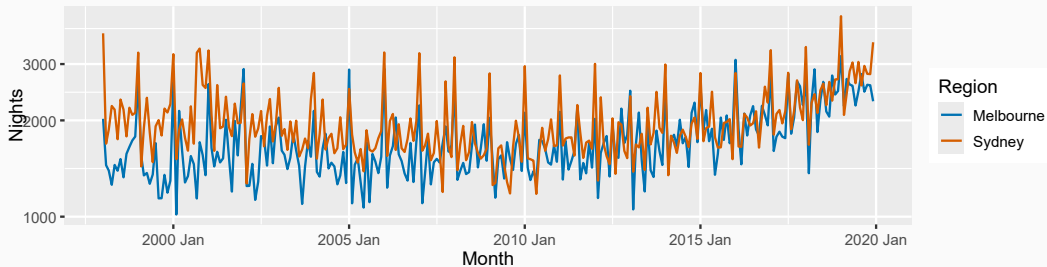


- Visitor Nights
- Monthly time series
- 1998 – 2019
- **77 regions**

Regions V Aggregate



Regions V Aggregate



FLAP Intuition

We have multivariate times series:

- which share similar patterns;
- with a better signal-noise ratio in the linear combination.

FLAP Intuition

We have multivariate times series:

- which share similar patterns;
- with a better signal-noise ratio in the linear combination.

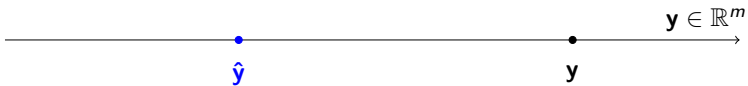
Can we find components that:

- 1 are easy to forecast (or easier than the original series);
- 2 can capture possible common signals;
- 3 can improve forecast of original series.

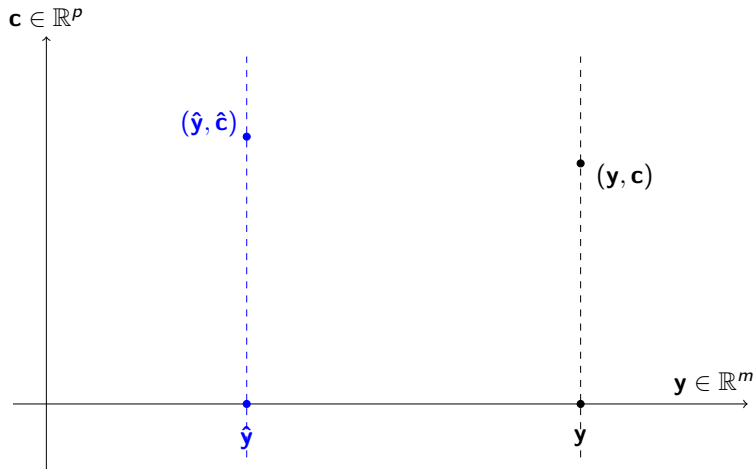
Outline of FLAP Implementation

- We want to forecast a multivariate series $\mathbf{y}_t \in \mathbb{R}^m$.
- Construct many linear combinations $\mathbf{c}_t = \Phi \mathbf{y}_t \in \mathbb{R}^p$ of the multivariate series.
- Produce univariate forecasts of all series $\hat{\mathbf{y}}_{t+h}$ and all linear combinations $\hat{\mathbf{c}}_{t+h}$.
- Project forecasts onto the \mathbb{R}^m coherent subspace, resulting in $\tilde{\mathbf{y}}_{t+h}$.

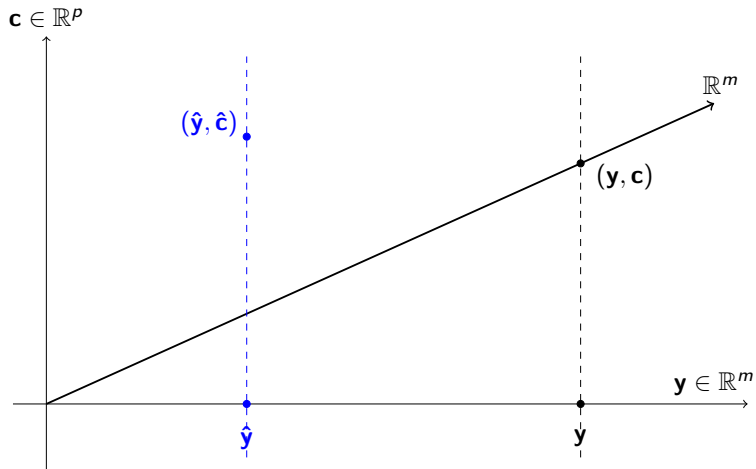
Geometry of FLAP



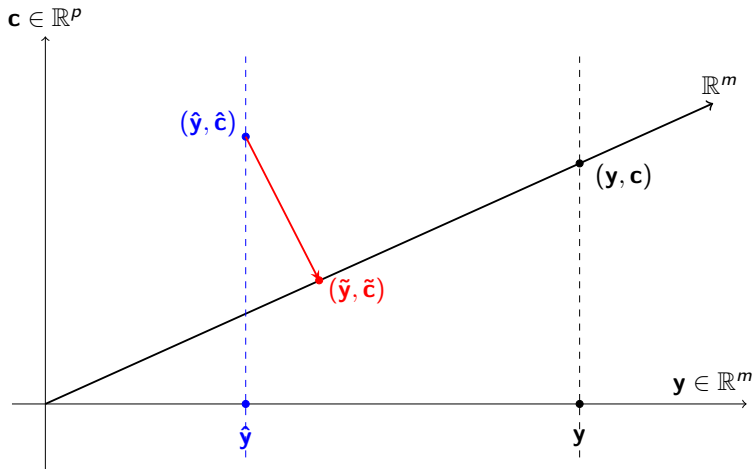
Geometry of FLAP



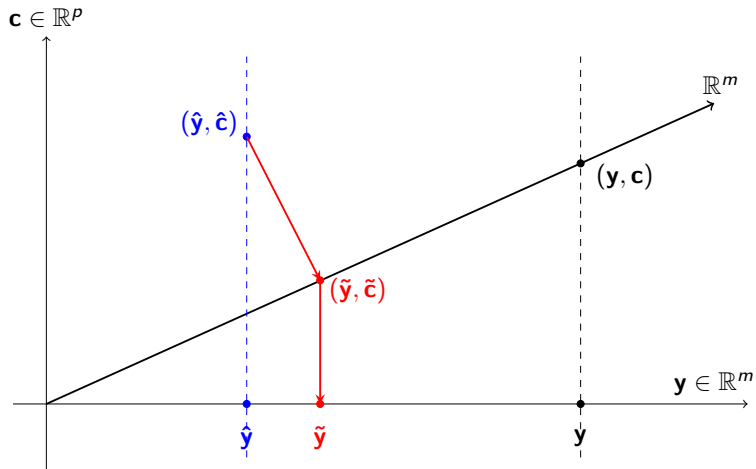
Geometry of FLAP



Geometry of FLAP



Geometry of FLAP



FLAP Projection

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{y}_t \\ \mathbf{c}_t \end{bmatrix}, \quad \hat{\mathbf{z}}_{t+h} = \begin{bmatrix} \hat{\mathbf{y}}_{t+h} \\ \hat{\mathbf{c}}_{t+h} \end{bmatrix}, \quad \tilde{\mathbf{z}}_{t+h} = \mathbf{M} \hat{\mathbf{z}}_{t+h}$$

where \mathbf{M} is a projection matrix onto the \mathbb{R}^m coherent subspace

$$\mathbf{M} = \mathbf{I}_{m+p} - \mathbf{W}_h \mathbf{C}' (\mathbf{C} \mathbf{W}_h \mathbf{C}')^{-1} \mathbf{C}$$

$$\mathbf{C} = \begin{bmatrix} -\Phi & \mathbf{I}_p \end{bmatrix}$$

$$\mathbf{W}_h = \text{Var}(\mathbf{z}_{t+h} - \hat{\mathbf{z}}_{t+h})$$

FLAP Projection

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{y}_t \\ \mathbf{c}_t \end{bmatrix}, \quad \hat{\mathbf{z}}_{t+h} = \begin{bmatrix} \hat{\mathbf{y}}_{t+h} \\ \hat{\mathbf{c}}_{t+h} \end{bmatrix}, \quad \tilde{\mathbf{z}}_{t+h} = \mathbf{M} \hat{\mathbf{z}}_{t+h}$$

where \mathbf{M} is a projection matrix onto the \mathbb{R}^m coherent subspace

$$\mathbf{M} = \mathbf{I}_{m+p} - \mathbf{W}_h \mathbf{C}' (\mathbf{C} \mathbf{W}_h \mathbf{C}')^{-1} \mathbf{C}$$

$$\mathbf{C} = \begin{bmatrix} -\Phi & \mathbf{I}_p \end{bmatrix}$$

$$\mathbf{W}_h = \text{Var}(\mathbf{z}_{t+h} - \hat{\mathbf{z}}_{t+h})$$

$$\tilde{\mathbf{y}}_{t+h} = \mathbf{G} \hat{\mathbf{z}}_{t+h} = \mathbf{J} \mathbf{M} \hat{\mathbf{z}}_{t+h}$$

$$\mathbf{J} = \begin{bmatrix} \mathbf{I}_m & \mathbf{O}_{m \times p} \end{bmatrix}$$

Forecast error variance reduction

- The variance reduction $\text{Var}(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{t+h}) - \text{Var}(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{t+h})$ is **positive semi-definite**.
- The diagonal elements of $\text{Var}(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{t+h}) - \text{Var}(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{t+h})$ are non-decreasing as the number of components increases.

Minimum variance of individual series

The projection is equivalent to the mapping

$$\tilde{\mathbf{y}}_{t+h} = \mathbf{G}\hat{\mathbf{z}}_{t+h},$$

where $\mathbf{G} = [\mathbf{g}_1 \ \mathbf{g}_2 \ \dots \ \mathbf{g}_m]' \in \mathbb{R}^{m \times (m+p)}$ is the solution to

$$\arg \min_{\mathbf{G}} \text{tr}(\mathbf{G}\mathbf{W}_h\mathbf{G}') \quad \text{s.t. } \mathbf{G}\mathbf{S} = \mathbf{I}$$

or

$$\arg \min_{\mathbf{g}_i} \mathbf{g}_i' \mathbf{W}_h \mathbf{g}_i \quad \text{s.t. } \mathbf{g}_i' \mathbf{s}_j = \mathbf{1}(i = j),$$

$$\text{where } \mathbf{S} = \begin{bmatrix} \mathbf{I}_m \\ \Phi \end{bmatrix} = [\mathbf{s}_1 \ \dots \ \mathbf{s}_m].$$

Key results

- 1 The forecast error variance is **reduced** with FLAP.
- 2 The forecast error variance **monotonically** decreases with increasing number of components.
- 3 The forecast projection is **optimal** to achieve minimum forecast error variance of each series.

Key results

- 1 The forecast error variance is **reduced** with FLAP.
- 2 The forecast error variance **monotonically** decreases with increasing number of components.
- 3 The forecast projection is **optimal** to achieve minimum forecast error variance of each series.

In practice, we need to:

- Estimate $\mathbf{W}_h = \text{Var}(\mathbf{z}_{t+h} - \hat{\mathbf{z}}_{t+h})$. (Use in-sample residuals, shrink variances to their median, covariances to zero.)
- Construct the components, Φ .

Construction of Φ

Principal component analysis (PCA)

Finding the weights matrix Φ so that the resulting components **maximise variance**

Simulation

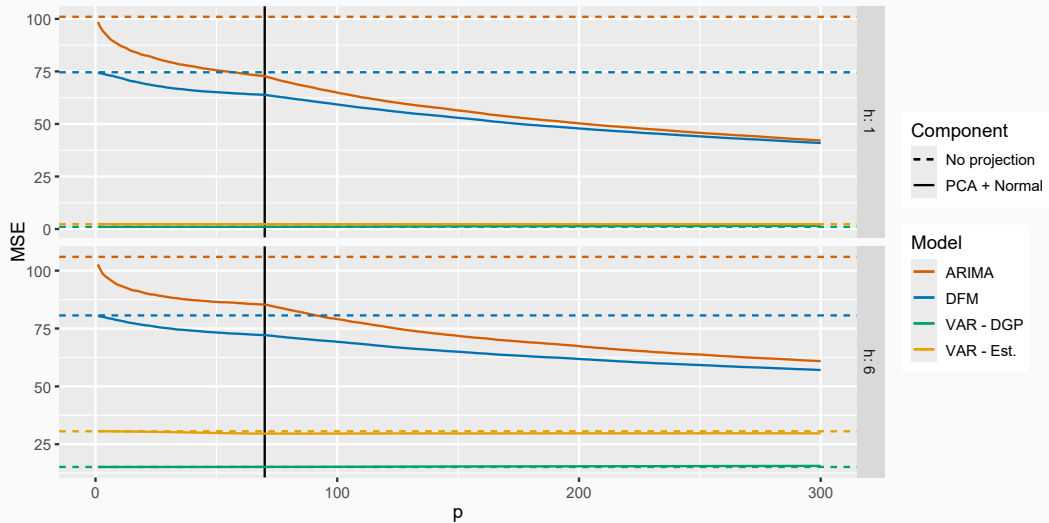
Generating values of Φ from a random distribution and normalising them to unit vectors

- Normal distribution
- Uniform distribution
- Orthonormal matrix

Simulation

- Data generating process: VAR(3) with $m = 70$ variables
- Innovations $\sim N(0, I_m)$
- Sample size: $T = 400$
- Number of repeated samples: 220
- Base forecasts:
 - ▶ ARIMA models using AICc (`auto.arima()` in forecast package).
 - ▶ DFM structure using BIC (different model for each horizon).

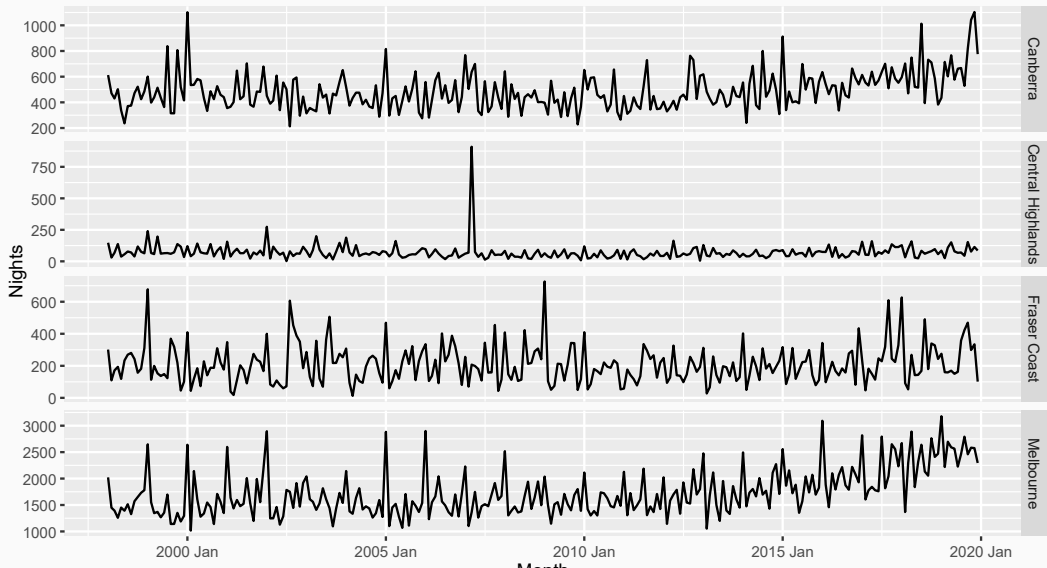
Simulation



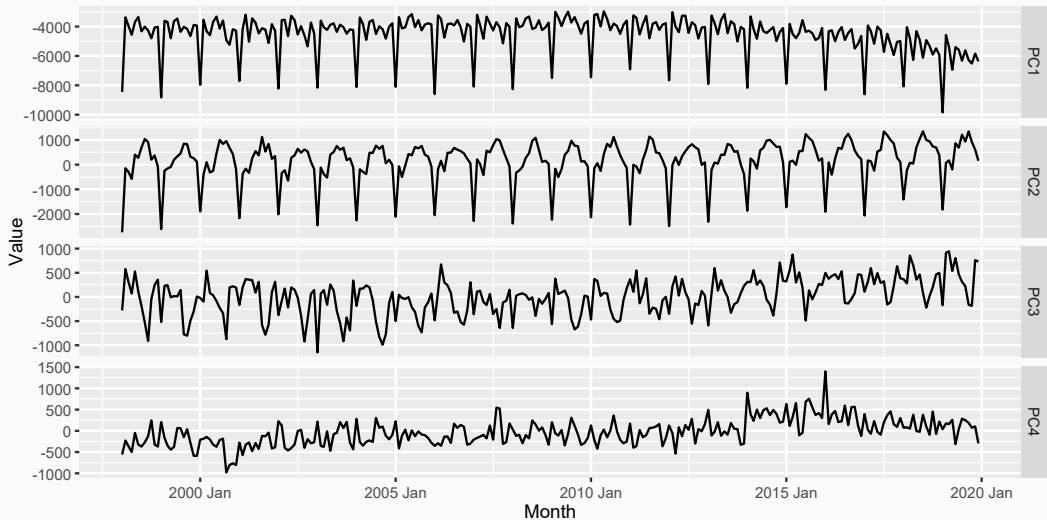
Monthly Australian regional tourism

- Monthly Australian tourism data by region giving 77 series, from Jan 1998 to Dec 2019
- Use expanding window time series cross-validation with $T = 84$ observations in first training set, and forecast horizons $h = 1, 2, \dots, 12$.
- Estimate `ets()` models using the `forecast` package.

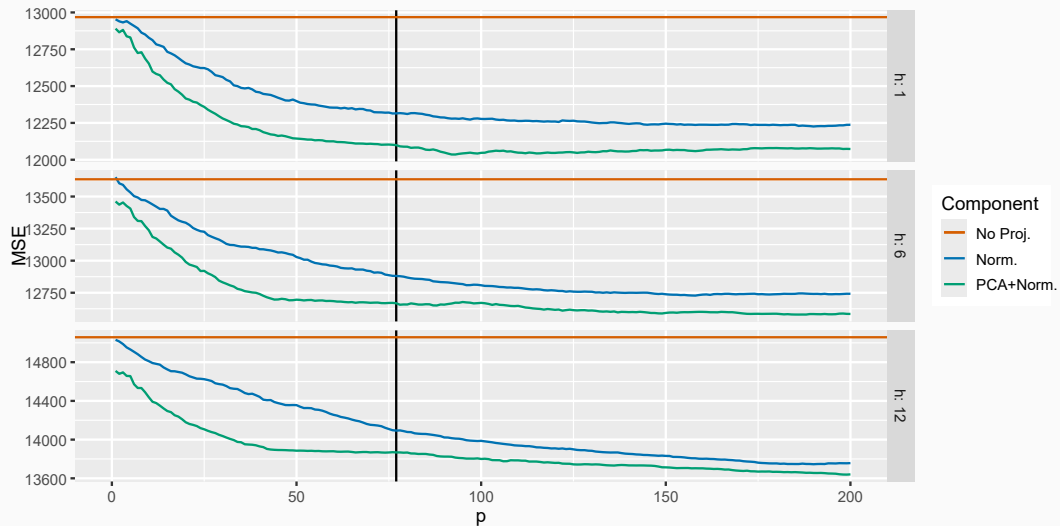
Monthly Australian regional tourism



Monthly Australian regional tourism

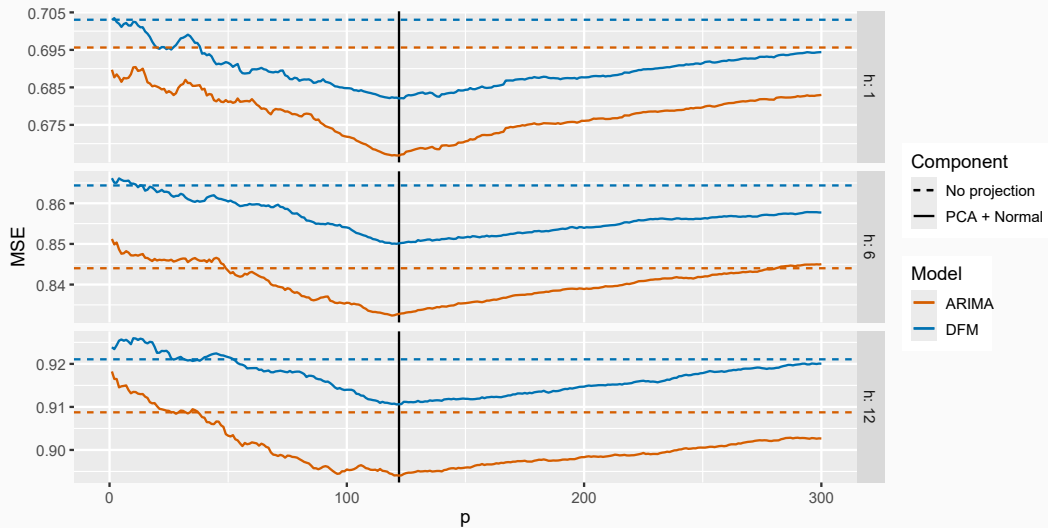


Monthly Australian regional tourism - ets()



- Monthly data of macroeconomic variables (McCracken and Ng, 2016).
- Data from Jan 1959 – Sep 2023. 777 observations on 122 series.
- Same cleaning process as per McCracken and Ng (2016).
- All series scaled to have mean 0 and variance 1.
- Expanding time series cross-validation with initial size of 25 years and forecast horizon 12 months.

FRED-MD



Future research directions

- Investigate why PCA performs better than random weights
- Find other components that are better than PCA
- Find optimal components by minimising forecast error variance with respect to Φ
- Use forecast projection and forecast reconciliation together

Working Paper and R Package

YF Yang, G Athanasopoulos, RJ Hyndman, and A Panagiotelis (2024). **“Forecast Linear Augmented Projection (FLAP): A free lunch to reduce forecast error variance”**. Department of Econometrics and Business Statistics, Monash University, Working Paper Series 13/24. URL: <https://www.monash.edu/business/ebs/research/publications/ebs/2024/wp13-2024.pdf>

You can install the stable version from CRAN

```
## CRAN.R-project.org/package=flap  
install.packages("flap")
```

or the development version from Github

```
## github.com/FinYang/flap  
# install.packages("remotes")  
remotes::install_github("FinYang/flap")
```



Yang, YF, G Athanasopoulos, RJ Hyndman, and A Panagiotelis (2024). **“Forecast Linear Augmented Projection (FLAP): A free lunch to reduce forecast error variance”**. Department of Econometrics and Business Statistics, Monash University, Working Paper Series 13/24. URL: <https://www.monash.edu/business/ebs/research/publications/ebs/2024/wp13-2024.pdf>.