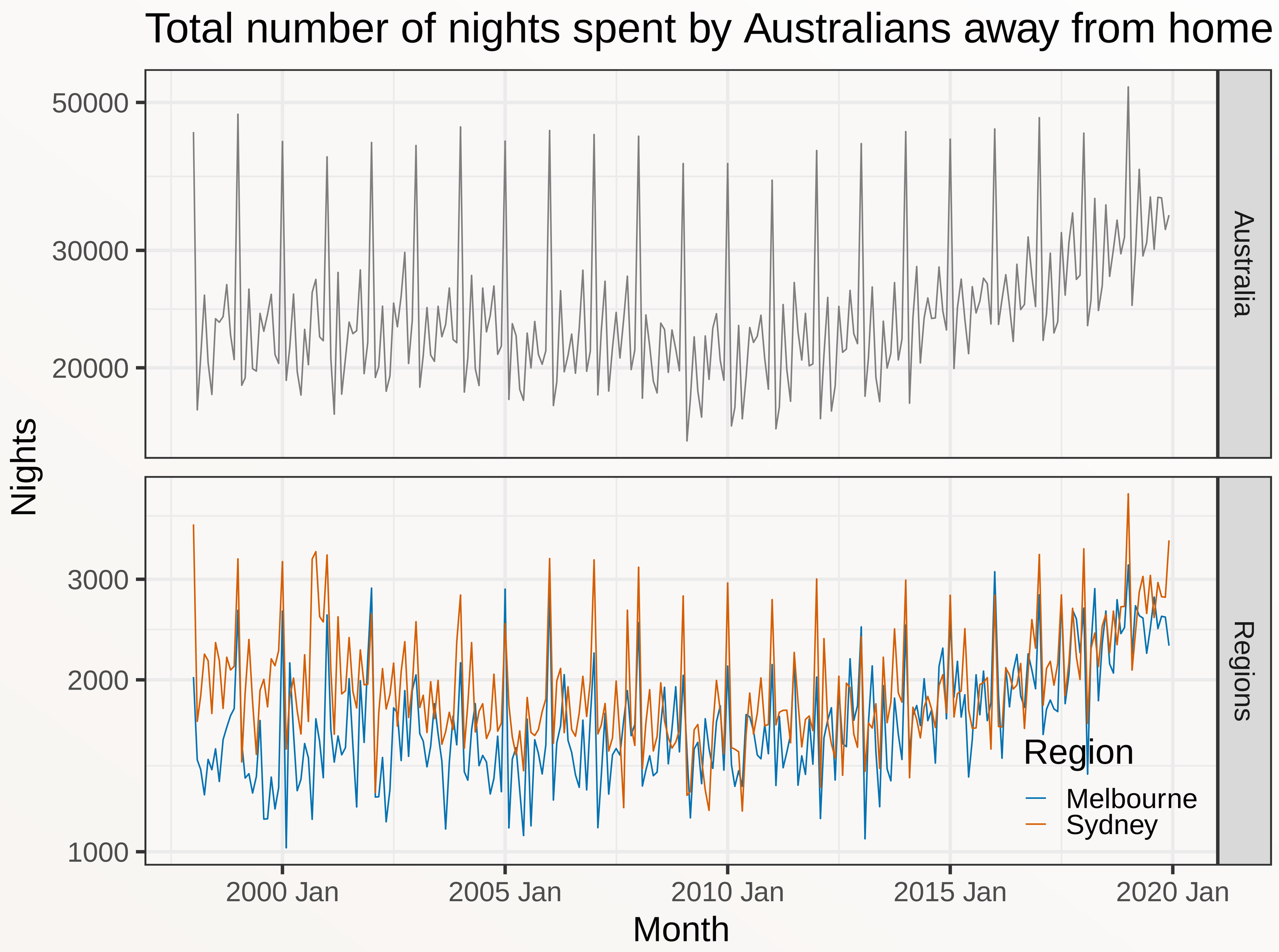
Forecast Linear Augmented Projection



We have multivariate time series:

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

Can we find components that:

1. are easier to forecast;
2. can capture possible common signals;
3. **can improve forecast of original series.**

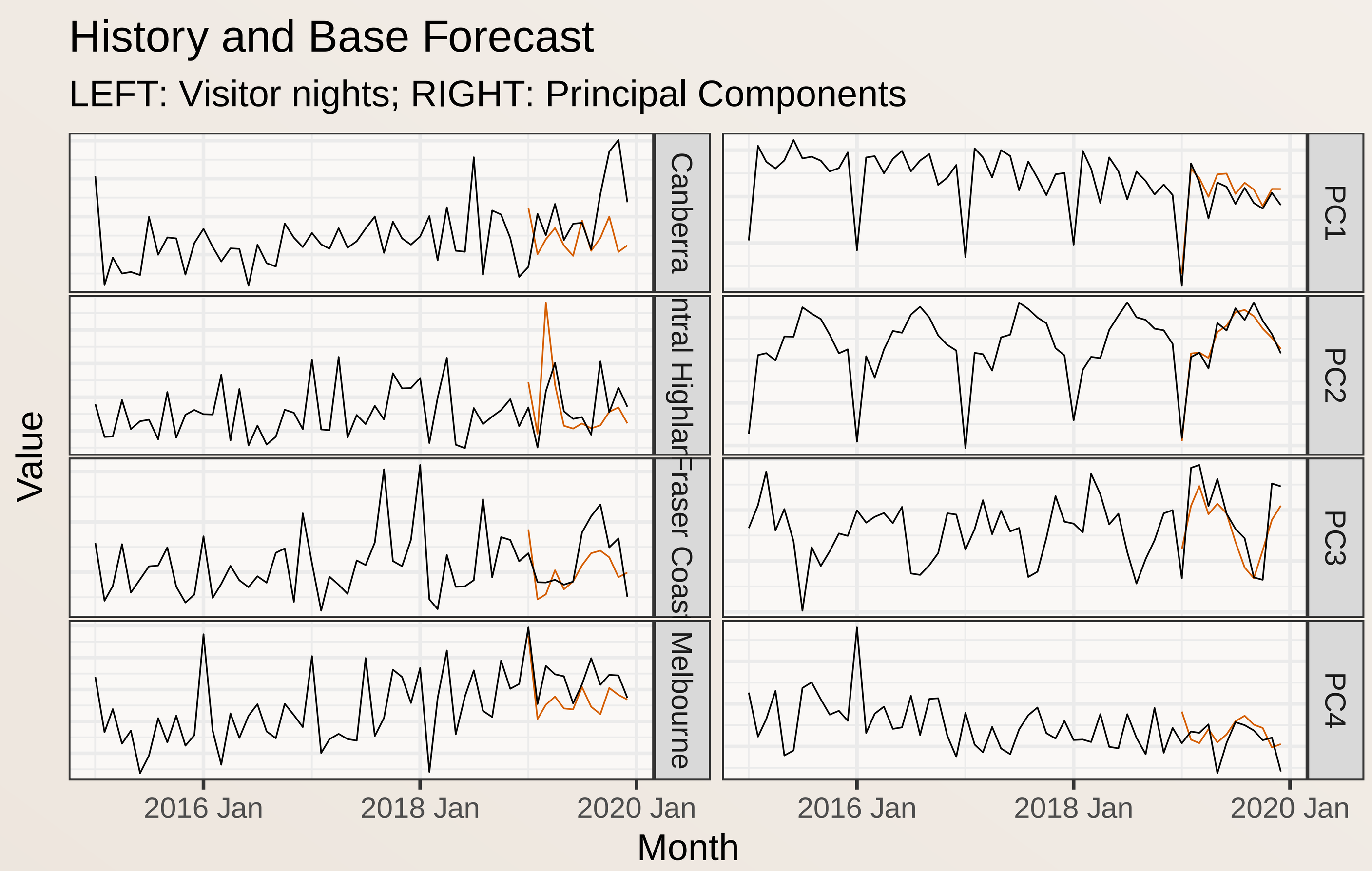
A free lunch to reduce forecast error variance

Yangzhuoran Fin Yang • George Athanasopoulos • Rob J Hyndman • Anastasios Panagiotelis

Implementation

Let $\boldsymbol{y}_t \in \boldsymbol{R}^m$ be a vector of m observed time series we are interested in forecasting. The FLAP method involves three steps:

1. **Form components.** Form $\boldsymbol{c}_t = \boldsymbol{\Phi}\boldsymbol{y}_t \in \boldsymbol{R}^p$, a vector of p linear combinations of $\boldsymbol{y}_t$ at time t, where $\boldsymbol{\Phi} \in \boldsymbol{R}^{p\times m}$. We call $\boldsymbol{c}_t$ the components of $\boldsymbol{y}_t$ and the component weights $\boldsymbol{\Phi}$ are known in the sense that they are chosen by the user of FLAP. Let $\boldsymbol{z}_t = \begin{bmatrix}\boldsymbol{y}_t', \boldsymbol{c}_t'\end{bmatrix}'$ be the concatenation of series $\boldsymbol{y}_t$ and components $\boldsymbol{c}_t$. $\boldsymbol{z}_t$ will be constrained in the sense that $\boldsymbol{C}\boldsymbol{z}_t = \boldsymbol{c}_t - \boldsymbol{\Phi}\boldsymbol{y}_t = \boldsymbol{0}$ for any t where $\boldsymbol{C} = \begin{bmatrix}-\boldsymbol{\Phi} & \boldsymbol{I}_p\end{bmatrix}$ is referred to as the constraint matrix.
2. **Generate forecasts.** Denote as $\hat{\boldsymbol{z}}_{t+h}$ the h-step-ahead base forecast of $\boldsymbol{z}_t$. The method used to generate forecasts is again selected by the user, and **any** prediction method can be used.



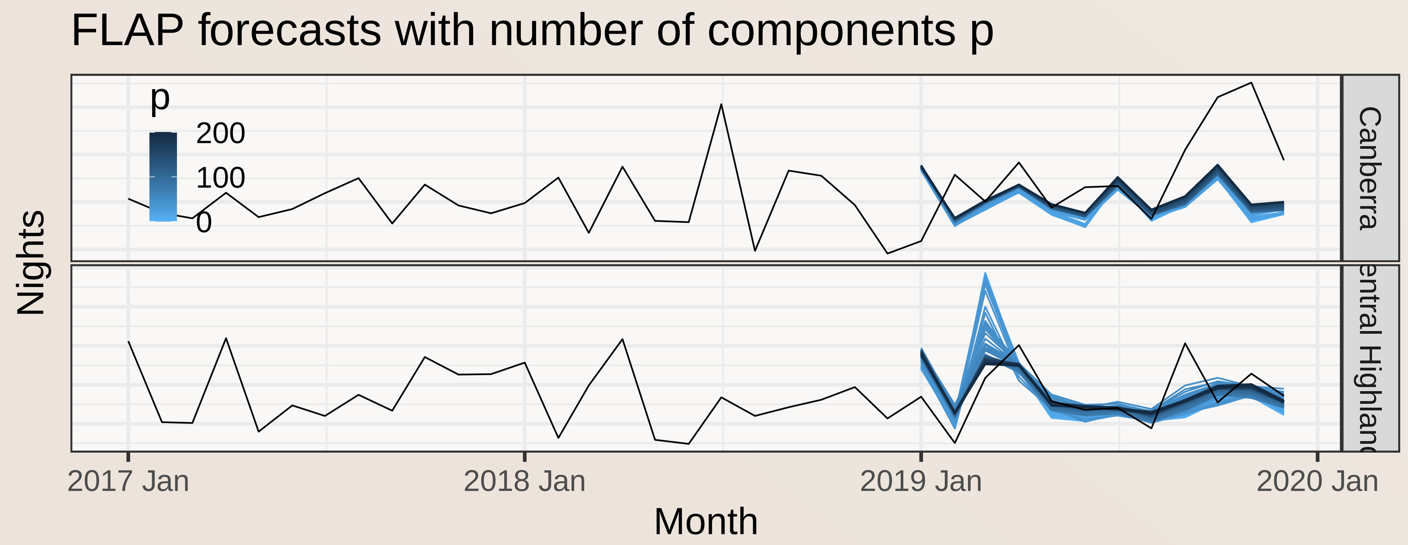
3. **Project the base forecasts.** Let $\tilde{\boldsymbol{z}}_{t+h}$ be a set of projected forecasts such that,

$$\tilde{\boldsymbol{z}}_{t+h} = \boldsymbol{M}\hat{\boldsymbol{z}}_{t+h} \qquad (1)$$

with projection matrix

$$\boldsymbol{M} = \boldsymbol{I}_{m+p} - \boldsymbol{W}_h\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{W}_h\boldsymbol{C}')^{-1}\boldsymbol{C}, \qquad (2)$$

where $\text{Var}(\boldsymbol{z}_{t+h} - \hat{\boldsymbol{z}}_{t+h}) = \boldsymbol{W}_h$ is the forecast error covariance matrix. In practice a plug-in estimate can be used.



Theoretical Properties

Key results

1. The forecast error variance is **reduced** with FLAP. The variance reduction matrix is positive semi-definite:

$$\begin{aligned}\text{Var}(\boldsymbol{y}_{t+h} - \hat{\boldsymbol{y}}_{t+h}) - \text{Var}(\boldsymbol{y}_{t+h} - \tilde{\boldsymbol{y}}_{t+h}) \\ = \boldsymbol{J}\boldsymbol{W}_h\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{W}_h\boldsymbol{C}')^{-1}\boldsymbol{C}\boldsymbol{W}_h\boldsymbol{J}'\end{aligned} \qquad (3)$$

2. The forecast error variance **monotonically** decreases with increasing number of components. The value of each of the diagonal elements of

$$\text{Var}(\boldsymbol{y}_{t+h} - \hat{\boldsymbol{y}}_{t+h}) - \text{Var}(\boldsymbol{y}_{t+h} - \tilde{\boldsymbol{y}}_{t+h})$$

is non-decreasing as p increases.

3. The forecast projection is **optimal** to achieve minimum forecast error variance of each series. The projection is equivalent to the mapping

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

where $\boldsymbol{G} = \begin{bmatrix}\boldsymbol{g}_1 & \boldsymbol{g}_2 & \ldots & \boldsymbol{g}_m\end{bmatrix}' \in \mathbb{R}^{m\times(m+p)}$ is the solution to

$$\arg\min_{\boldsymbol{G}} \boldsymbol{G}\boldsymbol{W}_h\boldsymbol{G}' \qquad \text{s.t. } \boldsymbol{G}\boldsymbol{S} = \boldsymbol{I}$$

or

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

where $\boldsymbol{S} = \begin{bmatrix}\boldsymbol{I}_m \\ \boldsymbol{\Phi}\end{bmatrix} = \begin{bmatrix}\boldsymbol{s}_1 \cdots \boldsymbol{s}_m\end{bmatrix}$.

Positive condition

For the first component to have a **guaranteed reduction of forecast error variance**, the following condition must be satisfied:

$$\boldsymbol{\phi}_1\boldsymbol{W}_{y,h} \neq \boldsymbol{w}_{c_1y,h},$$

where $\boldsymbol{\phi}_1$ is the weight vector of the first component, $\boldsymbol{W}_{y,h} = \text{Var}(\boldsymbol{y}_{t+h} - \hat{\boldsymbol{y}}_{t+h})$, and $\boldsymbol{w}_{c_1y,h}$ is the forecast error covariance between the first component and the original series.

A new component reduces the error variance as long as the its forecast covariance with the original series cannot be expressed as a linear combination of the forecast covariance between the already existing time series, in which case it adds no information.

Example $\boldsymbol{W}_h = \boldsymbol{I}_{m+p}$

$$\begin{aligned}\text{Var}(\boldsymbol{y}_{t+h} - \hat{\boldsymbol{y}}_{t+h}) - \text{Var}(\boldsymbol{y}_{t+h} - \tilde{\boldsymbol{y}}_{t+h}) \\ = \boldsymbol{J}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-1}\boldsymbol{C}\boldsymbol{J}' \\ = \boldsymbol{\Phi}'(\boldsymbol{\Phi}\boldsymbol{\Phi}' + \boldsymbol{I})^{-1}\boldsymbol{\Phi}\end{aligned}$$

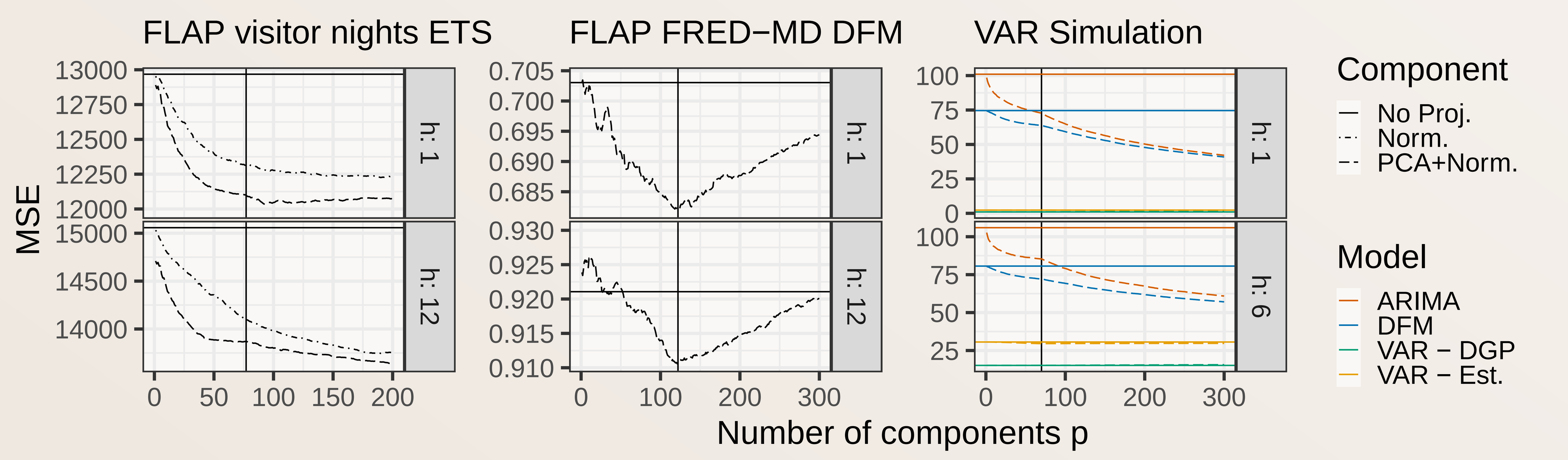
Let $\boldsymbol{\Phi}$ consist of orthogonal unit vectors:

$$\boldsymbol{\Phi}\boldsymbol{\Phi}' = \boldsymbol{I}_p \text{ when } p \leq m$$
$$\boldsymbol{\Phi}'\boldsymbol{\Phi} = \boldsymbol{I}_m \text{ when } p = m.$$

$$\begin{aligned}\text{tr}(\text{Var}(\boldsymbol{y}_{t+h} - \hat{\boldsymbol{y}}_{t+h}) - \text{Var}(\boldsymbol{y}_{t+h} - \tilde{\boldsymbol{y}}_{t+h})) \\ = \frac{1}{2}tr(\boldsymbol{\Phi}'\boldsymbol{\Phi}) = \frac{1}{2}p\end{aligned}$$

Applications

We estimate $\boldsymbol{W}_h$ using a shrinkage estimator to ensure positive definitness and numerical stability.
We construct $\boldsymbol{\Phi}$ using principal component analysis (PCA) and simulations from random distributions.




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

I am looking for a job! ➜

yangzhuoranyang.com
@FinYang
@yangzhuoranyang.bsky.social
yangzhuoran.yang@monash.edu

Projection matrix The matrix $\boldsymbol{M}$ is a projection onto the space where the constraint $\boldsymbol{C}\boldsymbol{z}_t = \boldsymbol{0}$ is satisfied.