

Improving forecasts via subspace projections

George Athanasopoulos
CRETE2025, Tinos, Greece



MONASH University

Outline

- 1 Improving hierarchical forecasts
- 2 Improving univariate forecasts
- 3 Improving cross-temporal forecasts
- 4 Improving multivariate forecasts
- 5 Final comments

Outline

1 Improving hierarchical forecasts

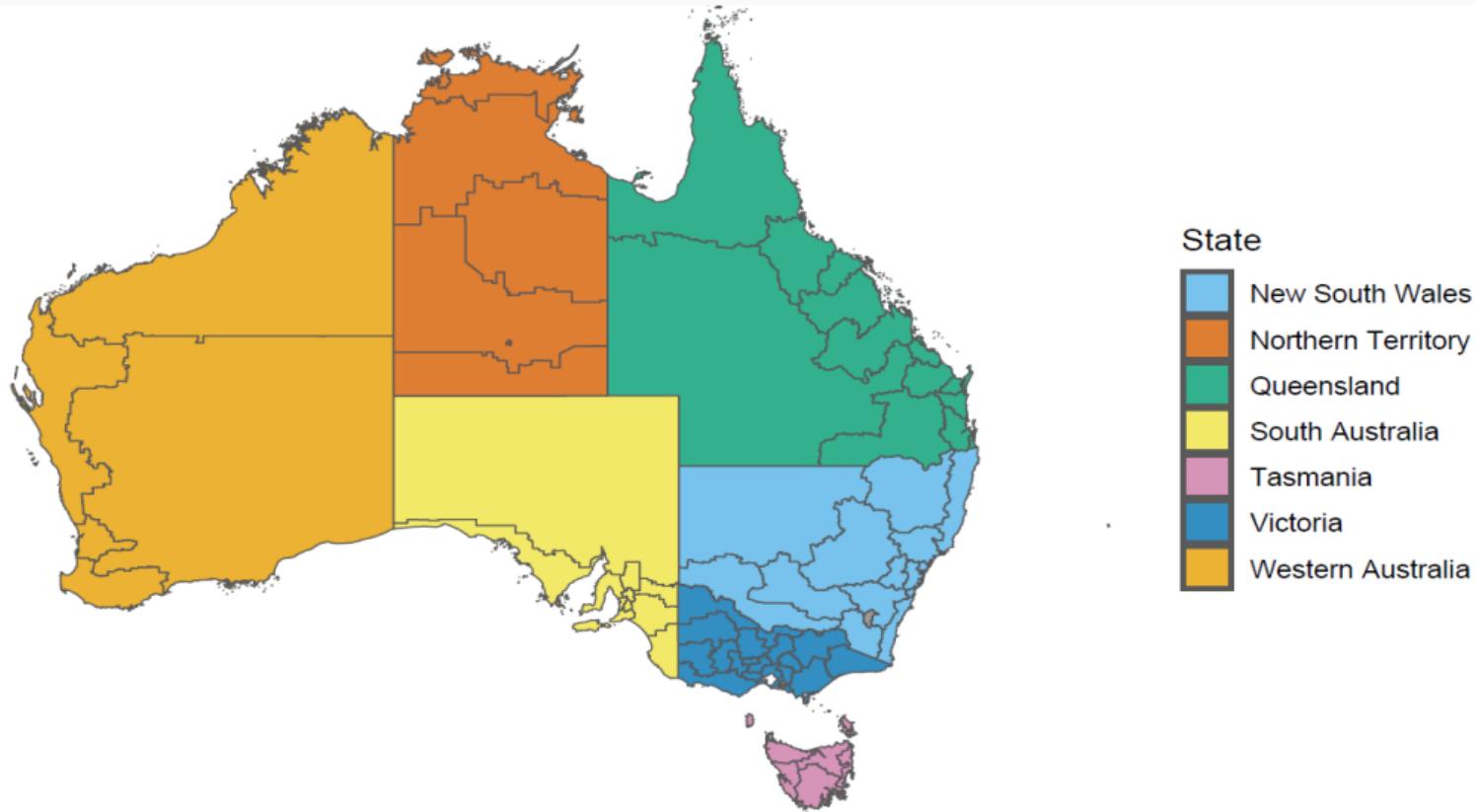
2 Improving univariate forecasts

3 Improving cross-temporal forecasts

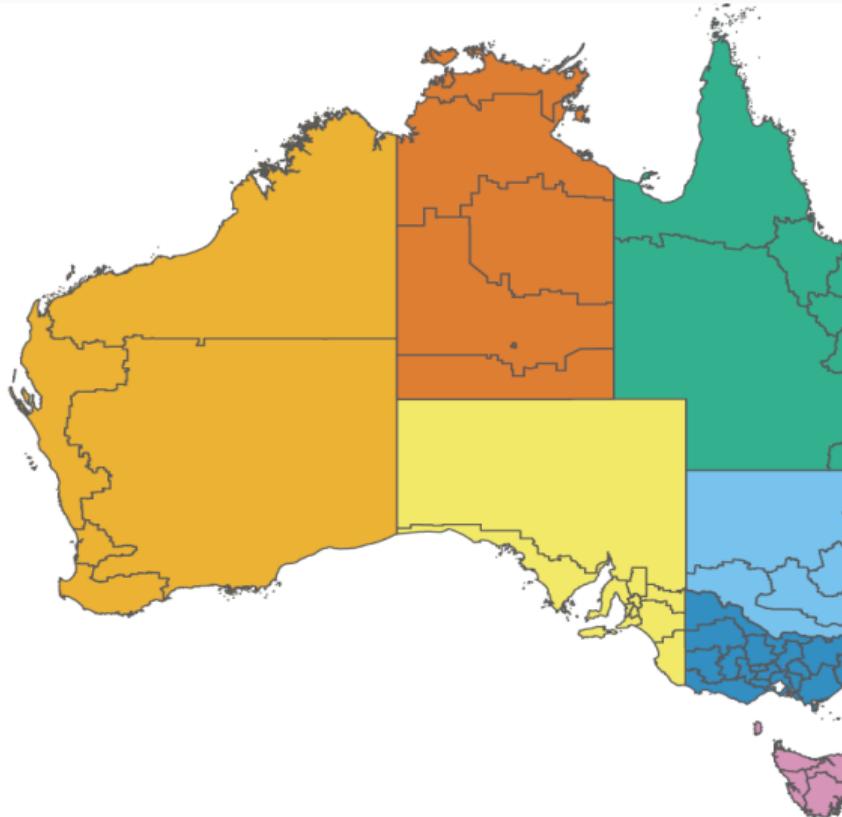
4 Improving multivariate forecasts

5 Final comments

Australian tourism regions



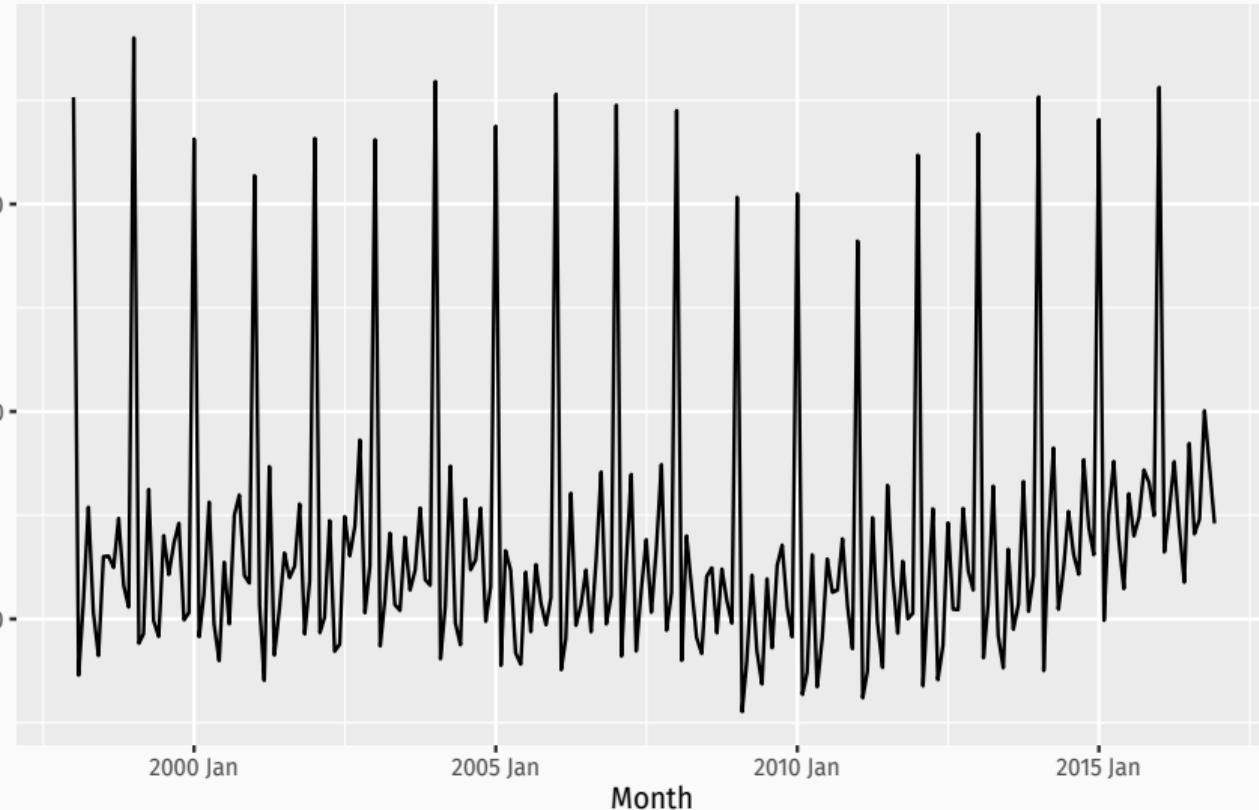
Australian tourism regions



- Monthly data on visitor night: 1998 – 2019
- From *National Visitor Survey*, annual interviews of 120,000 Australians aged 15+.
- Geographical hierarchy split by
 - ▶ 7 states
 - ▶ 27 zones
 - ▶ 75 regions

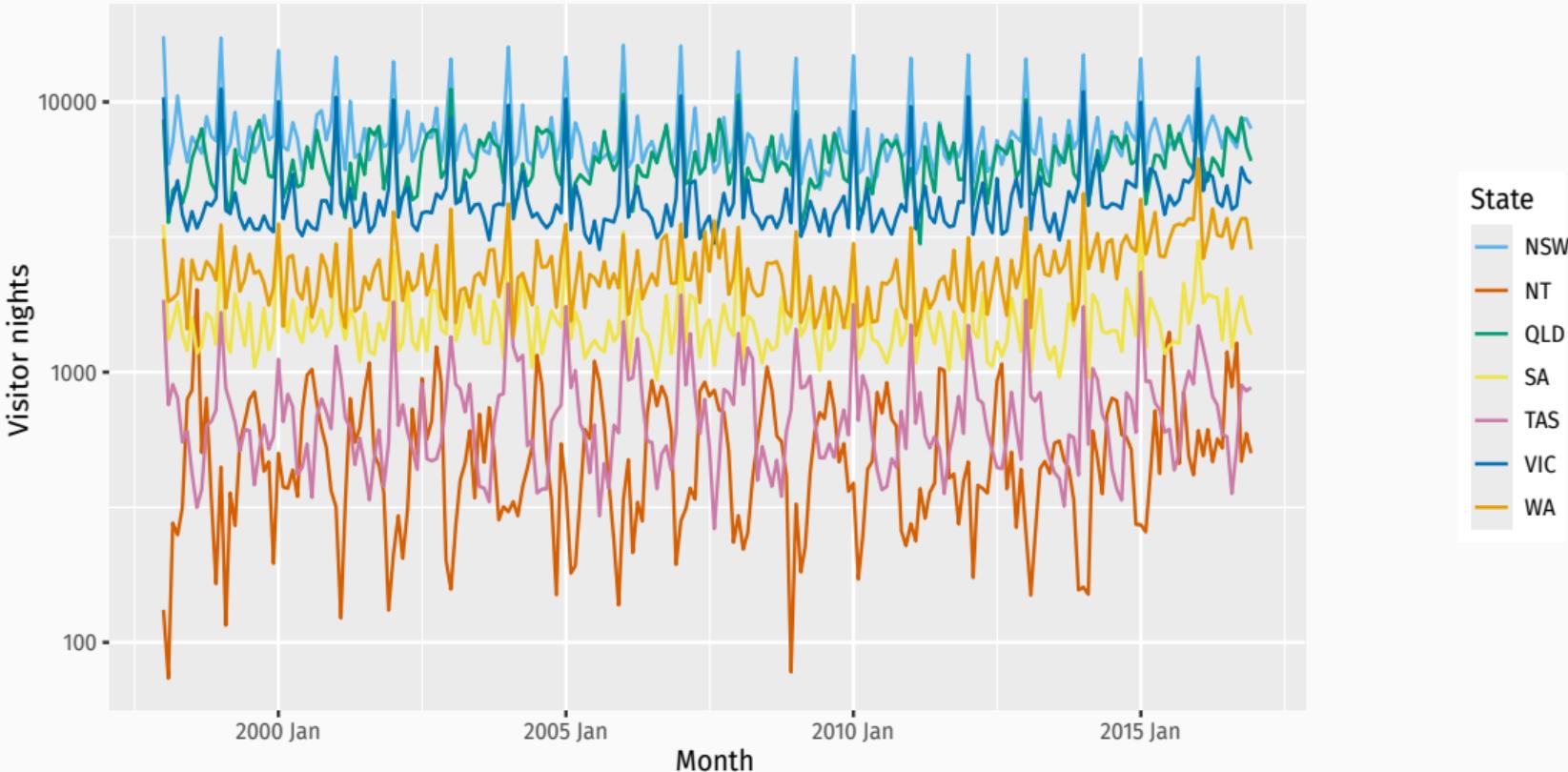
Australian tourism data

Total domestic travel: Australia



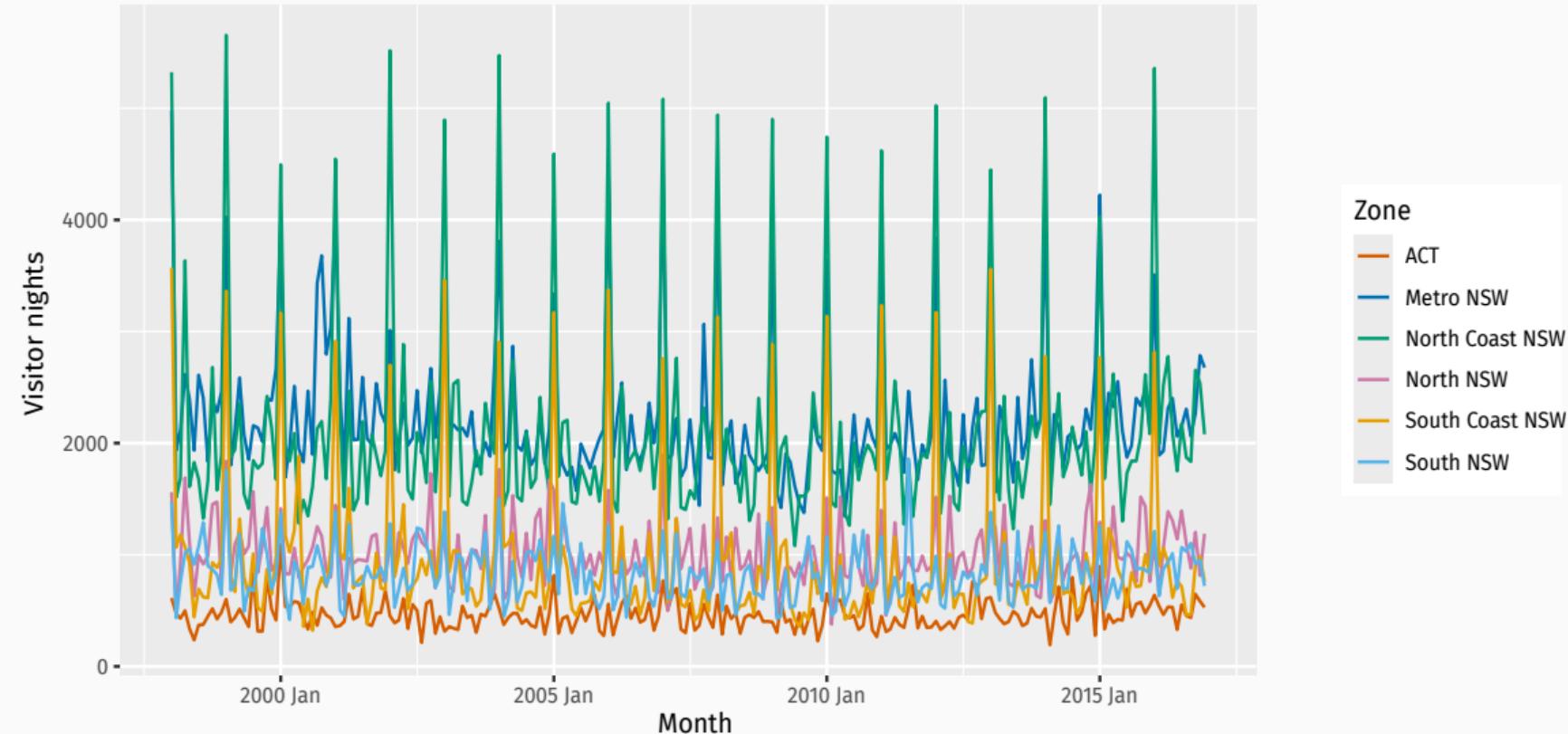
Australian tourism data

Total domestic travel: by state



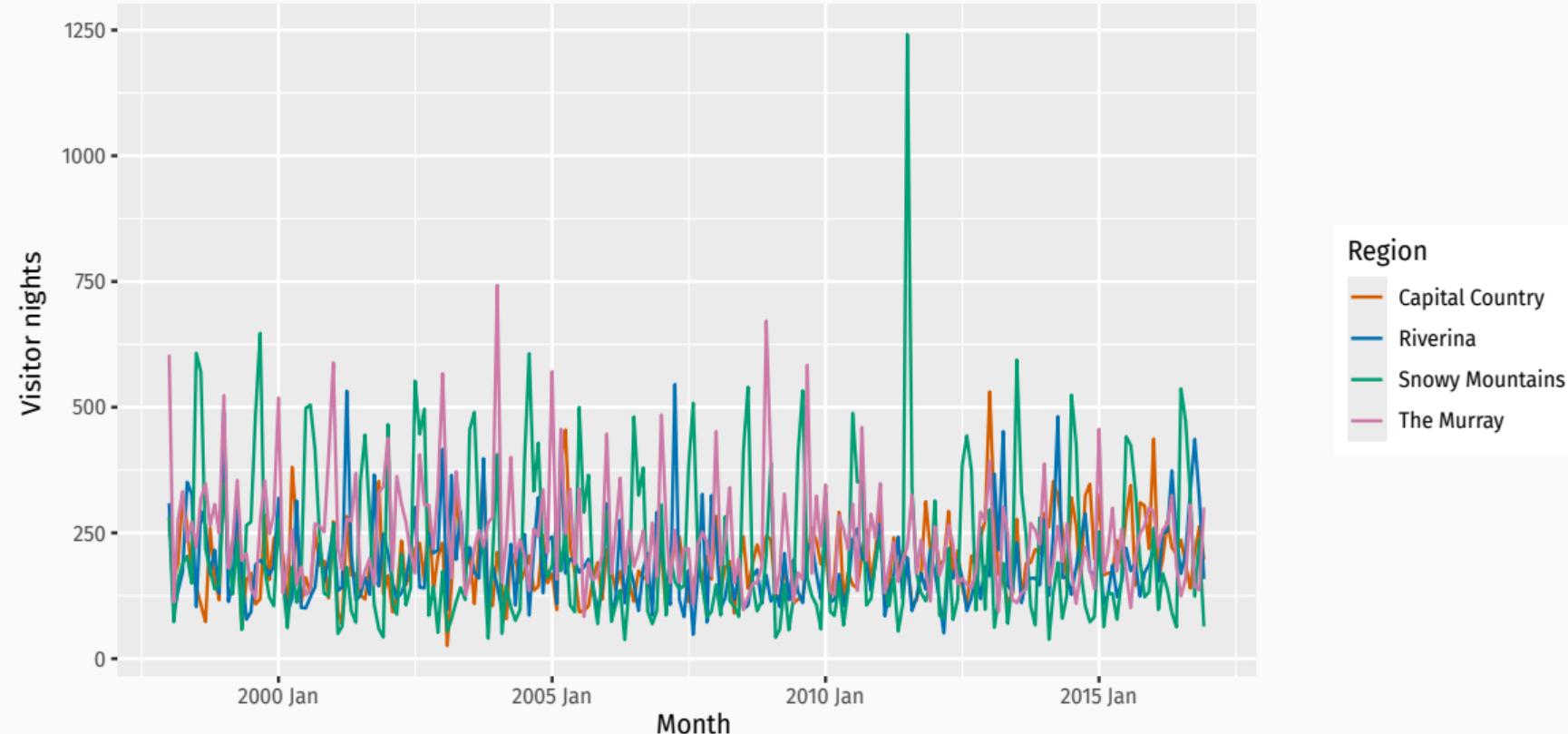
Australian tourism data

Total domestic travel: NSW by zone



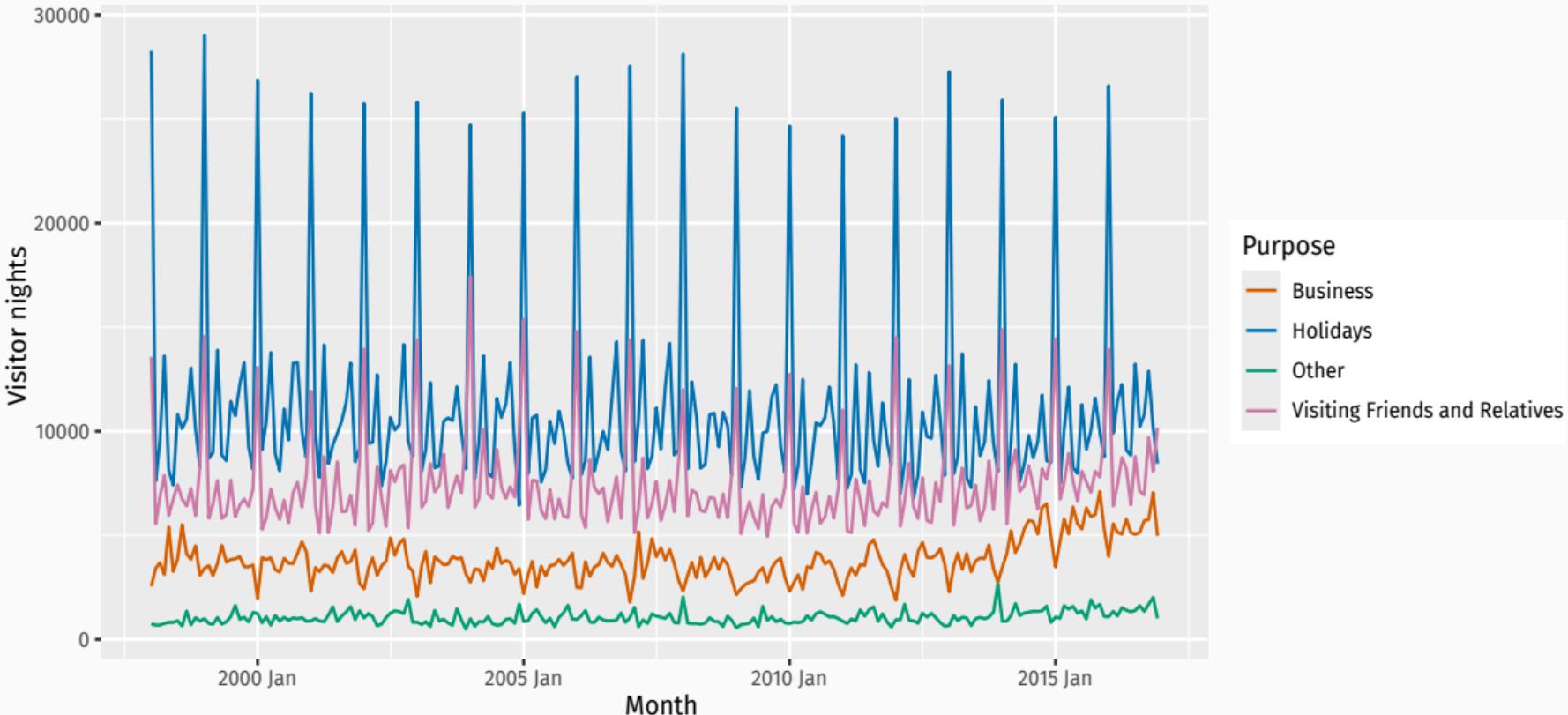
Australian tourism data

Total domestic travel: South NSW by region



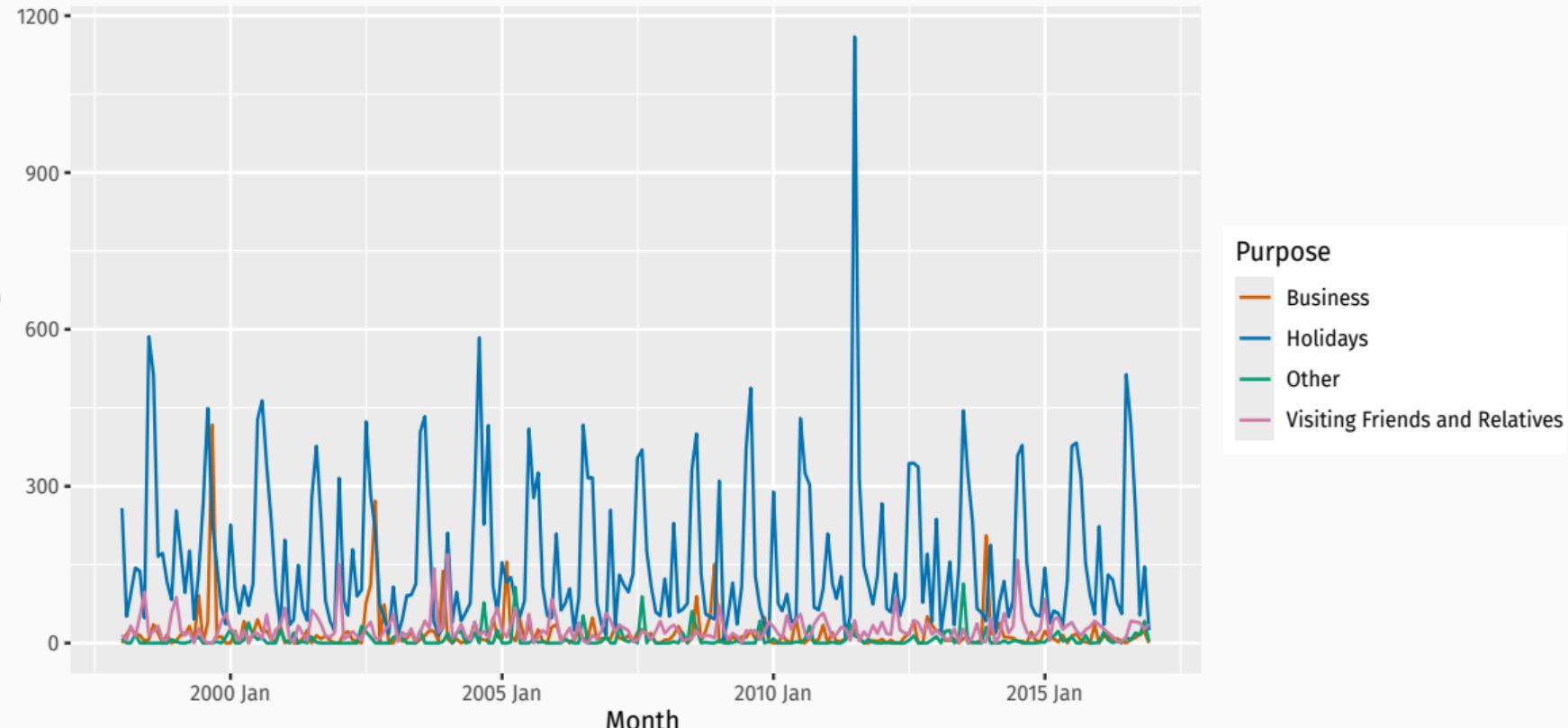
Australian tourism data

Total domestic travel: by purpose of travel



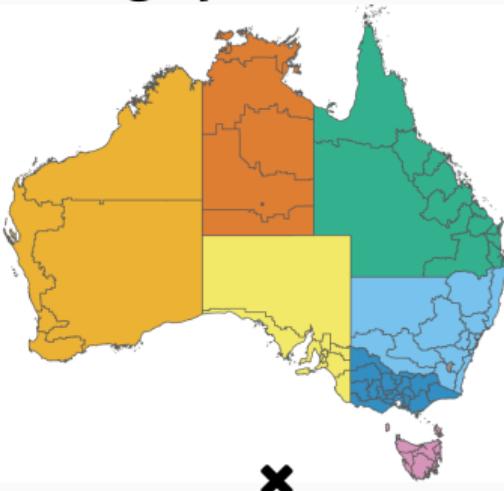
Australian tourism data

Total domestic travel: Snowy Mountains by purpose of travel



Australian tourism data

Geographical division



Purpose of travel

Holiday, Visiting friends & relatives, Business, Other

Grouped time series

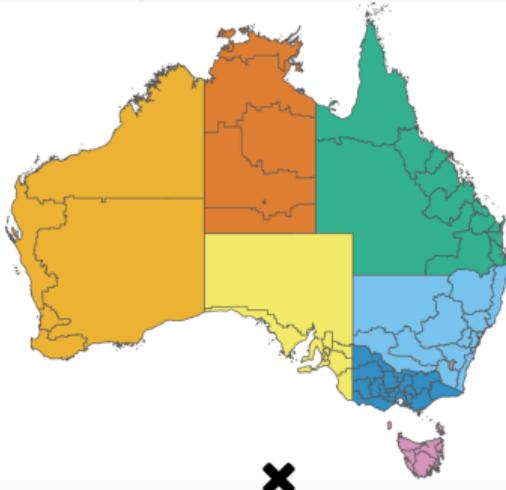
(geographical divisions × purpose of travel)

	AUS	States	Zones	Regions	Tot
geographical	1	7	21	76	105
purpose	4	28	84	304	420
total	5	35	105	380	525

m = 304 and **n = 525**

Australian tourism data

Geographical division



Purpose of travel

Holiday, Visiting friends & relatives, Business, Other

Grouped time series

(geographical divisions \times purpose of travel)

	AUS	States	Zones	Regions	Tot
geographical	1	7	21	76	105
purpose	4	28	84	304	420
total	5	35	105	380	525

m = 304 and **n = 525**

- Need forecasts at all levels of aggregation.
- Forecasts generated by different agents/models will not adhere to aggregation constraints.

Key idea

- Traditional single level approaches: bottom-up, top-down or middle-out.

Key idea

- Traditional single level approaches: bottom-up, top-down or middle-out.
- Forecast all series. ([Base forecasts](#)).

Key idea

- Traditional single level approaches: bottom-up, top-down or middle-out.
- Forecast all series. ([Base forecasts](#)).
- Project onto coherent subspace. ([Reconciled forecasts](#)).

Key idea

- Traditional single level approaches: bottom-up, top-down or middle-out.
- Forecast all series. (**Base forecasts**).
- Project onto coherent subspace. (**Reconciled forecasts**).



G Athanasopoulos, R Ahmed, and R Hyndman (2009). “**Hierarchical forecasts for Australian domestic tourism**”. In: *International Journal of Forecasting* 25, pp. 146–166. URL: <https://doi.org/10.1016/j.ijforecast.2008.07.004>



RJ Hyndman, R Ahmed, G Athanasopoulos, and H Shang (2011). “**Optimal combination forecasts for hierarchical time series**”. In: *Computational Statistics & Data Analysis* 55.9, pp. 2579–2589. URL: <https://doi.org/10.1016/j.csda.2011.03.006>

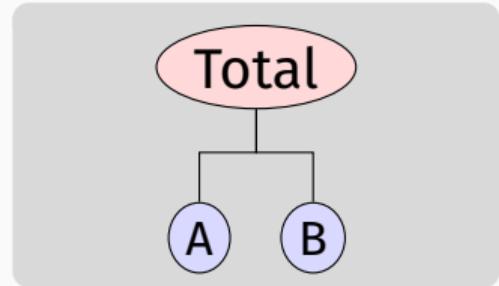


RJ Hyndman and G Athanasopoulos (2021). **Forecasting: principles and practice**. 3rd Edn. Melbourne, Australia. URL: <http://otexts.com/fpp3/>

Notation

$$\mathbf{y}_t = \mathbf{S}\mathbf{b}_t$$

- \mathbf{y}_t = vector of all series at time t
- \mathbf{b}_t = vector of most disaggregated series at time t
- \mathbf{S} = “structural matrix” containing the linear constraints.

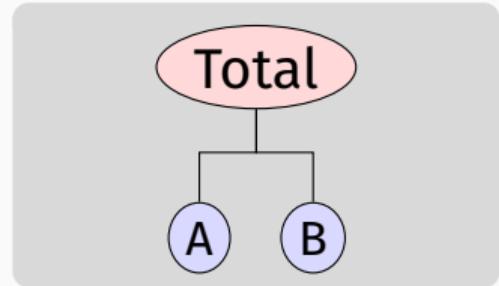


$$\begin{aligned}\mathbf{y}_t &= \begin{pmatrix} y_{\text{Total},t} \\ y_{A,t} \\ y_{B,t} \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}}_S \underbrace{\begin{pmatrix} y_{A,t} \\ y_{B,t} \end{pmatrix}}_{\mathbf{b}_t}\end{aligned}$$

Notation

$$\mathbf{y}_t = \mathbf{S}\mathbf{b}_t$$

- \mathbf{y}_t = vector of all series at time t
- \mathbf{b}_t = vector of most disaggregated series at time t
- \mathbf{S} = “structural matrix” containing the linear constraints.
- m – number of bottom-level series
- n – number of all series



$$\begin{aligned}\mathbf{y}_t &= \begin{pmatrix} y_{\text{Total},t} \\ y_{A,t} \\ y_{B,t} \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}}_S \underbrace{\begin{pmatrix} y_{A,t} \\ y_{B,t} \end{pmatrix}}_{\mathbf{b}_t}\end{aligned}$$

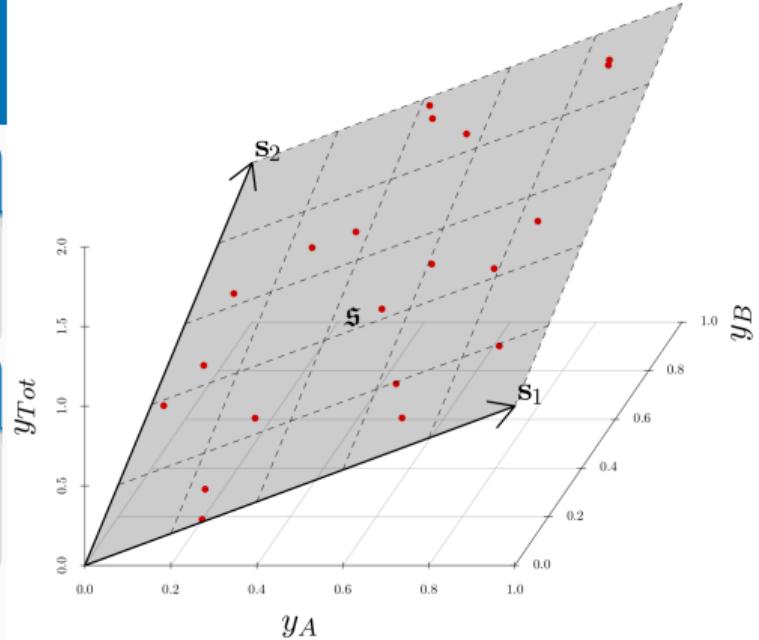
The coherent subspace

Hierarchical time series

Multivariate time series $\mathbf{y}_t \in \mathbb{R}^n$, bound by linear constraints.

Coherent subspace

m -dimensional linear subspace $\mathfrak{s} \subset \mathbb{R}^n$ for which linear constraints hold for all $\mathbf{y}_t \in \mathfrak{s}$.



$$y_{Tot} = y_A + y_B$$

The coherent subspace

Hierarchical time series

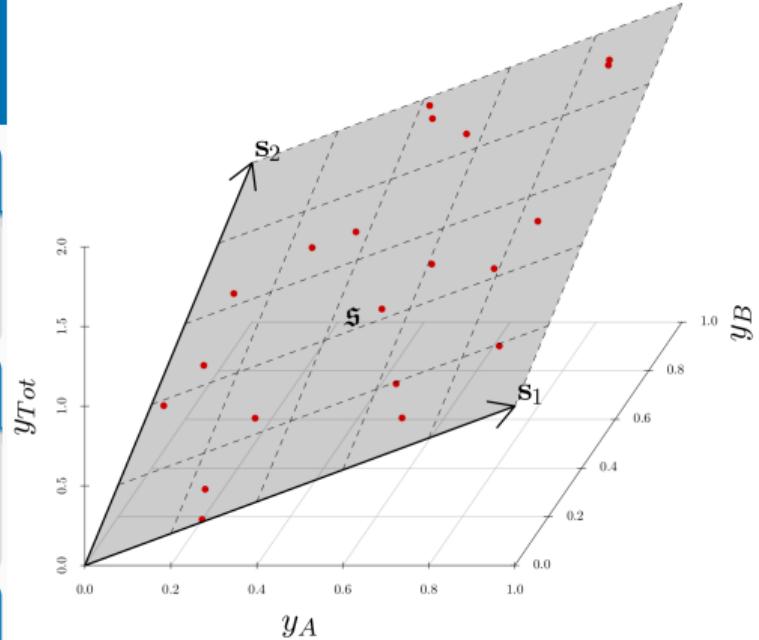
Multivariate time series $\mathbf{y}_t \in \mathbb{R}^n$, bound by linear constraints.

Coherent subspace

m -dimensional linear subspace $\mathfrak{s} \subset \mathbb{R}^n$ for which linear constraints hold for all $\mathbf{y}_t \in \mathfrak{s}$.

Coherent point forecasts

$\tilde{\mathbf{y}}_{T+h|T}$ is *coherent* if $\tilde{\mathbf{y}}_{T+h|T} \in \mathfrak{s}$.



$$\mathbf{y}_{Tot} = \mathbf{y}_A + \mathbf{y}_B$$

The coherent subspace

Hierarchical time series

Multivariate time series $\mathbf{y}_t \in \mathbb{R}^n$, bound by linear constraints.

Coherent subspace

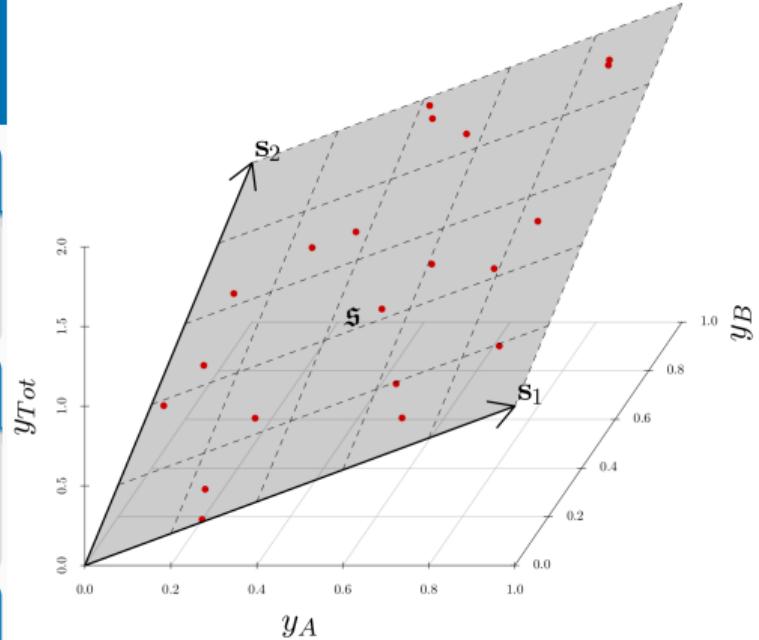
m -dimensional linear subspace $\mathfrak{s} \subset \mathbb{R}^n$ for which linear constraints hold for all $\mathbf{y}_t \in \mathfrak{s}$.

Coherent point forecasts

$\tilde{\mathbf{y}}_{T+h|T}$ is *coherent* if $\tilde{\mathbf{y}}_{T+h|T} \in \mathfrak{s}$.

Base forecasts

Let $\hat{\mathbf{y}}_{T+h|T}$ be vector of *incoherent* initial h -step forecasts.



$$y_{Tot} = y_A + y_B$$

The coherent subspace

Hierarchical time series

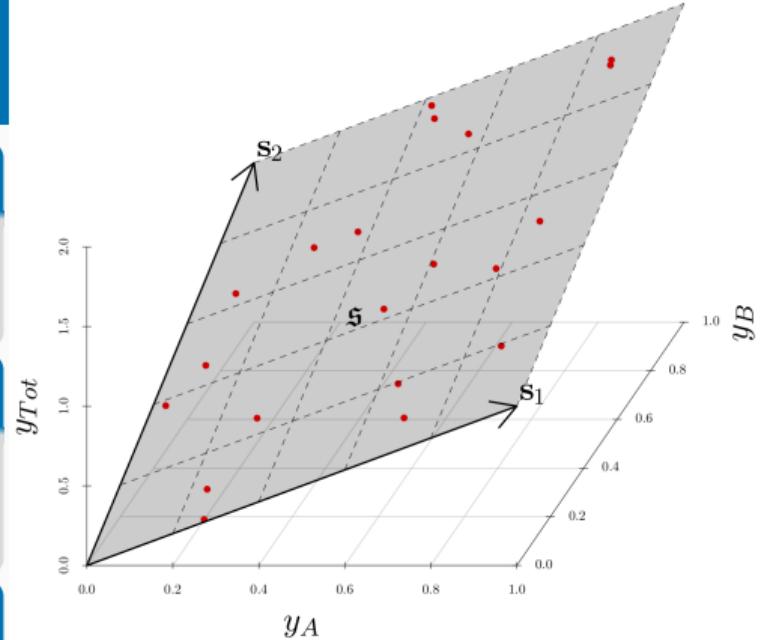
Multivariate time series $\mathbf{y}_t \in \mathbb{R}^n$, bound by linear constraints.

Coherent subspace

m -dimensional linear subspace $\mathfrak{s} \subset \mathbb{R}^n$ for which linear constraints hold for all $\mathbf{y}_t \in \mathfrak{s}$.

Coherent point forecasts

$\tilde{\mathbf{y}}_{T+h|T}$ is *coherent* if $\tilde{\mathbf{y}}_{T+h|T} \in \mathfrak{s}$.



$$y_{Tot} = y_A + y_B$$

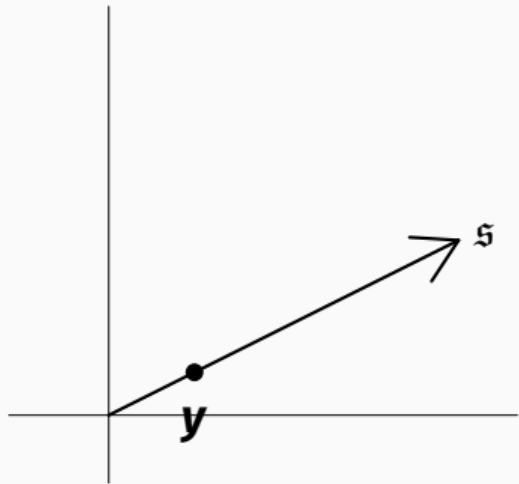
Base forecasts

Let $\hat{\mathbf{y}}_{T+h|T}$ be vector of *incoherent* initial h -step forecasts.

Reconciled forecasts

Let \mathbf{SG} be a projection matrix.
 $\tilde{\mathbf{y}}_{T+h|T} = \mathbf{SG}\hat{\mathbf{y}}_{T+h|T}$ “reconciles” $\hat{\mathbf{y}}_{T+h|T}$.

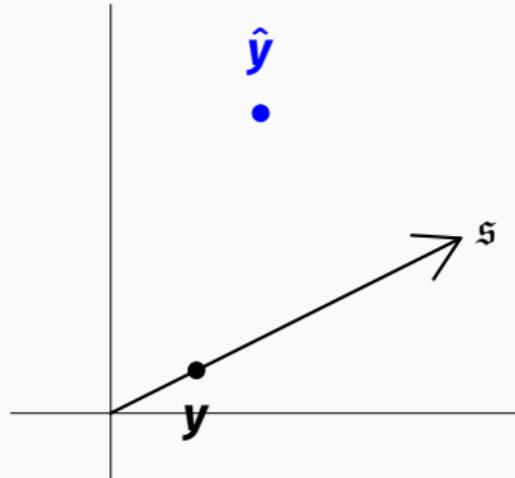
Geometry of forecast reconciliation



Geometry of forecast reconciliation

Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(.)$



Geometry of forecast reconciliation

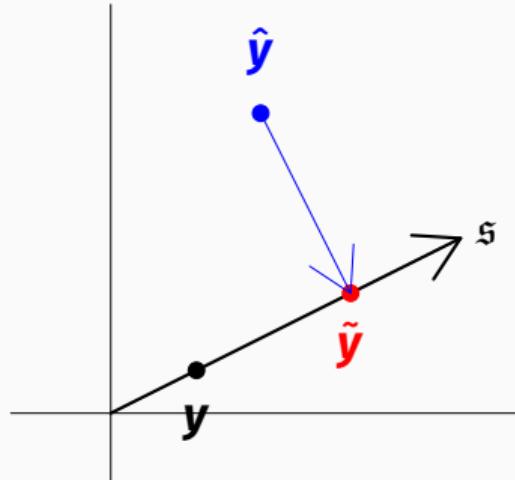
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Geometry of forecast reconciliation

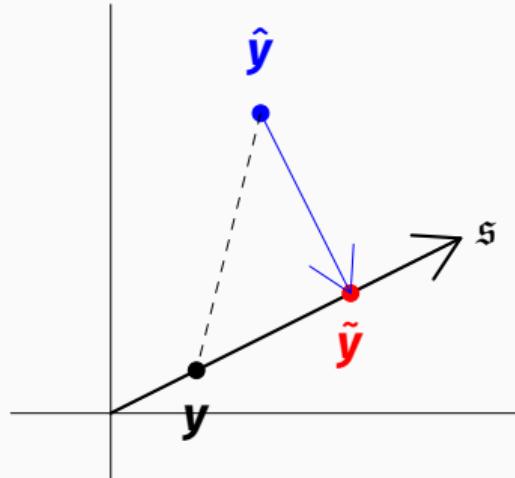
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Geometry of forecast reconciliation

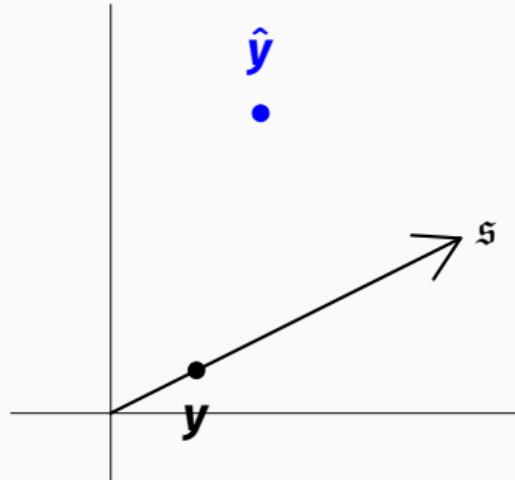
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Geometry of forecast reconciliation

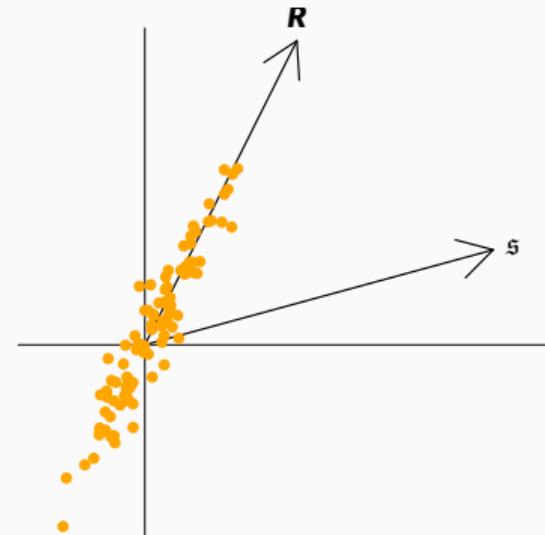
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Oblique Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$, $\mathbf{G} = (\mathbf{S}'\Psi\mathbf{S})^{-1}\mathbf{S}'\Psi$ achieves

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\|_\Psi \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|_\Psi$$

Geometry of forecast reconciliation

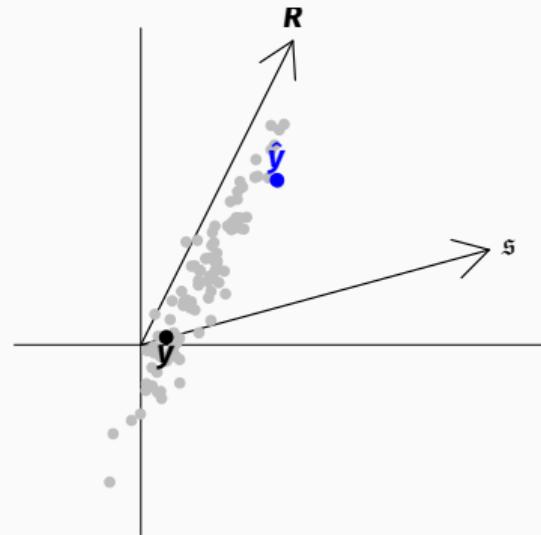
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Oblique Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$, $\mathbf{G} = (\mathbf{S}'\Psi\mathbf{S})^{-1}\mathbf{S}'\Psi$ achieves

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\|_\Psi \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|_\Psi$$

Geometry of forecast reconciliation

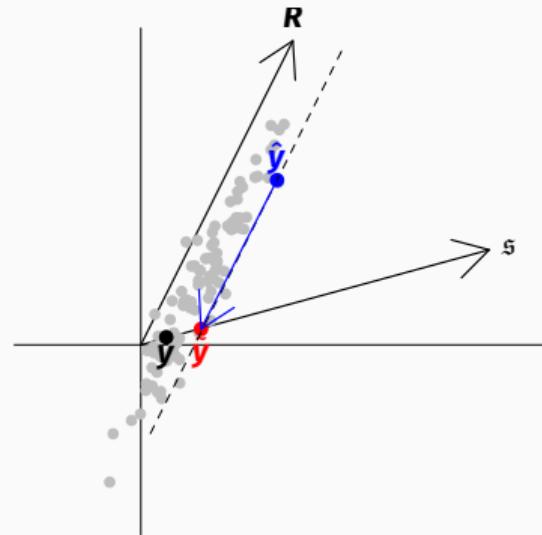
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Oblique Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$, $\mathbf{G} = (\mathbf{S}'\Psi\mathbf{S})^{-1}\mathbf{S}'\Psi$ achieves

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\|_\Psi \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|_\Psi$$

Geometry of forecast reconciliation

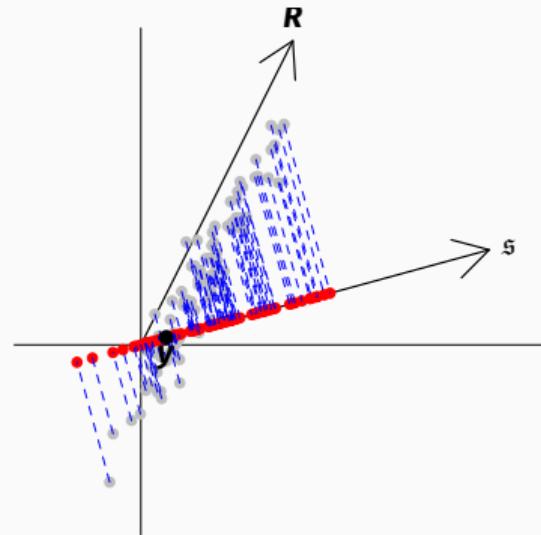
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Oblique Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$, $\mathbf{G} = (\mathbf{S}'\Psi\mathbf{S})^{-1}\mathbf{S}'\Psi$ achieves

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\|_\Psi \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|_\Psi$$

Geometry of forecast reconciliation

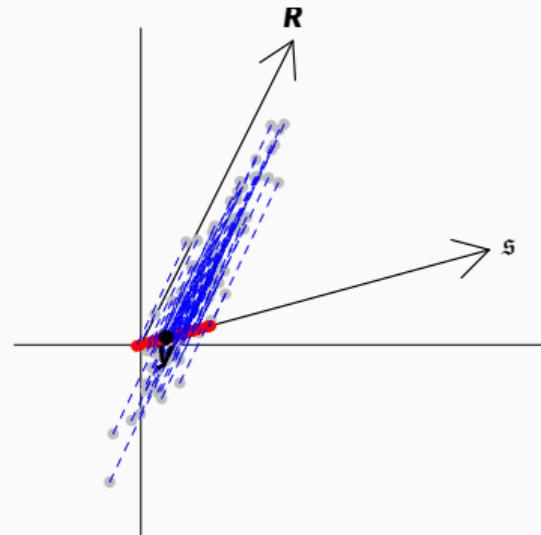
Reconciled point forecasts

Let ψ be a mapping, $\psi : \mathbb{R}^n \rightarrow \mathfrak{s}$. The point forecast $\tilde{\mathbf{y}}_{T+h|T} = \psi(\hat{\mathbf{y}}_{T+h|T})$ “reconciles” a base forecast $\hat{\mathbf{y}}_{T+h|T}$ with respect to the mapping $\psi(\cdot)$

Orthogonal Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$ where $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ guarantees

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\| \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|$$



Oblique Projection

$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$, $\mathbf{G} = (\mathbf{S}'\Psi\mathbf{S})^{-1}\mathbf{S}'\Psi$ achieves

$$\|(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{T+h|T})\|_\Psi \leq \|(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{T+h|T})\|_\Psi$$

Minimum trace (MinT) reconciliation

- How to choose the best Ψ ?

Minimum trace (MinT) reconciliation

- How to choose the best Ψ ?
- Let $\mathbf{W}_h = \text{Var}[\mathbf{y}_{T+h} - \hat{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T]$ be the covariance matrix of the base forecast errors.
- Then $\mathbf{V}_h = \text{Var}[\mathbf{y}_{T+h} - \tilde{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T] = \mathbf{S}\mathbf{G}\mathbf{W}_h\mathbf{G}'\mathbf{S}'$ is the covariance matrix of the reconciled forecast errors.

Minimum trace (MinT) reconciliation

- How to choose the best Ψ ?
- Let $\mathbf{W}_h = \text{Var}[\mathbf{y}_{T+h} - \hat{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T]$ be the covariance matrix of the base forecast errors.
- Then $\mathbf{V}_h = \text{Var}[\mathbf{y}_{T+h} - \tilde{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T] = \mathbf{S}\mathbf{G}\mathbf{W}_h\mathbf{G}'\mathbf{S}'$ is the covariance matrix of the reconciled forecast errors.

Minimum trace (MinT) reconciliation

If \mathbf{SG} is a projection, then trace of \mathbf{V}_h is minimized when $\Psi = \mathbf{W}_h$, so that

$$\mathbf{SG} = \mathbf{S}(\mathbf{S}'\mathbf{W}_h^{-1}\mathbf{S})^{-1}\mathbf{S}'\mathbf{W}_h^{-1}$$

Minimum trace (MinT) reconciliation

- How to choose the best Ψ ?
- Let $\mathbf{W}_h = \text{Var}[\mathbf{y}_{T+h} - \hat{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T]$ be the covariance matrix of the base forecast errors.
- Then $\mathbf{V}_h = \text{Var}[\mathbf{y}_{T+h} - \tilde{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T] = \mathbf{S}\mathbf{G}\mathbf{W}_h\mathbf{G}'\mathbf{S}'$ is the covariance matrix of the reconciled forecast errors.

Minimum trace (MinT) reconciliation

If \mathbf{SG} is a projection, then trace of \mathbf{V}_h is minimized when $\Psi = \mathbf{W}_h$, so that

$$\mathbf{SG} = \mathbf{S}(\mathbf{S}'\mathbf{W}_h^{-1}\mathbf{S})^{-1}\mathbf{S}'\mathbf{W}_h^{-1}$$

- MinT is L_2 optimal amongst linear unbiased forecasts.

MinT linear projections

$$\tilde{\mathbf{y}}_{T+h|T} = S\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$$

- How to estimate $\mathbf{W}_h = \text{Var}[\mathbf{y}_{T+h} - \hat{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T]$?

MinT linear projections

$$\tilde{\mathbf{y}}_{T+h|T} = \mathbf{S}\mathbf{G}\hat{\mathbf{y}}_{T+h|T}$$

- How to estimate $\mathbf{W}_h = \text{Var}[\mathbf{y}_{T+h} - \hat{\mathbf{y}}_{T+h|T} \mid \mathbf{y}_1, \dots, \mathbf{y}_T]$?

Reconc. method $\mathbf{S}\mathbf{G}$

OLS $\mathbf{S}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$

WLS(var) $\mathbf{S}(\mathbf{S}'\Lambda_v\mathbf{S})^{-1}\mathbf{S}'\Lambda_v$ where $\Lambda_v = \text{diag}(\hat{\mathbf{W}}_1)^{-1}$

WLS(struct) $\mathbf{S}(\mathbf{S}'\Lambda_s\mathbf{S})^{-1}\mathbf{S}'\Lambda_s$ where $\Lambda_s = \text{diag}(\mathbf{S}\mathbf{1})^{-1}$

MinT(sample) $\mathbf{S}(\mathbf{S}'\hat{\mathbf{W}}_1^{-1}\mathbf{S})^{-1}\mathbf{S}'\hat{\mathbf{W}}_1^{-1}$

MinT(shrink) $\mathbf{S}(\mathbf{S}'\hat{\mathbf{W}}_{\text{shr}}^{-1}\mathbf{S})^{-1}\mathbf{S}'\hat{\mathbf{W}}_{\text{shr}}^{-1}$

- $\hat{\mathbf{W}}_{\text{shr}}$ is shrinkage estimator $\tau \text{diag}(\hat{\mathbf{W}}_1) + (1 - \tau)\hat{\mathbf{W}}_1$
where τ selected optimally.

MinT and Geometry papers



SL Wickramasuriya, G Athanasopoulos, and RJ Hyndman (2019). “Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization”. In: *Journal of the American Statistical Association* 114.526, pp. 804–819. URL: <https://doi.org/10.1080/01621459.2018.1448825>.



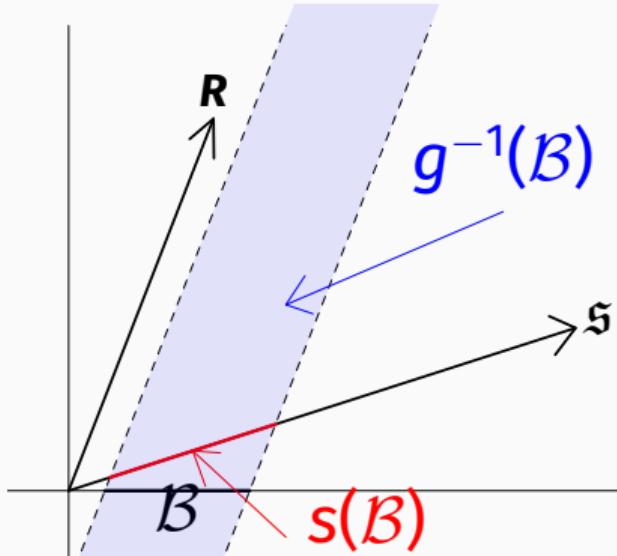
A Panagiotelis, G Athanasopoulos, P Gamakumara, and RJ Hyndman (2021). “Forecast reconciliation: A geometric view with new insights on bias correction”. In: *International Journal of Forecasting* 37.1, pp. 343–359. URL: <https://doi.org/10.1016/j.ijforecast.2020.06.004>

Probabilistic forecast reconciliation

A probability triple $(\mathfrak{s}, \mathcal{F}_{\mathfrak{s}}, \check{\nu})$ is coherent with the bottom probability triple $(\mathbb{R}^m, \mathcal{F}_{\mathbb{R}^m}, \nu)$, if

$$\check{\nu}(s(\mathcal{B})) = \nu(\mathcal{B}) \quad \forall \mathcal{B} \in \mathcal{F}_{\mathbb{R}^m}$$

- Random draws from coherent distribution must lie on \mathfrak{s} .
- The probability of points not on \mathfrak{s} is zero.
- The reconciled distribution is a transformation of the base forecast distribution that is coherent on \mathfrak{s} .

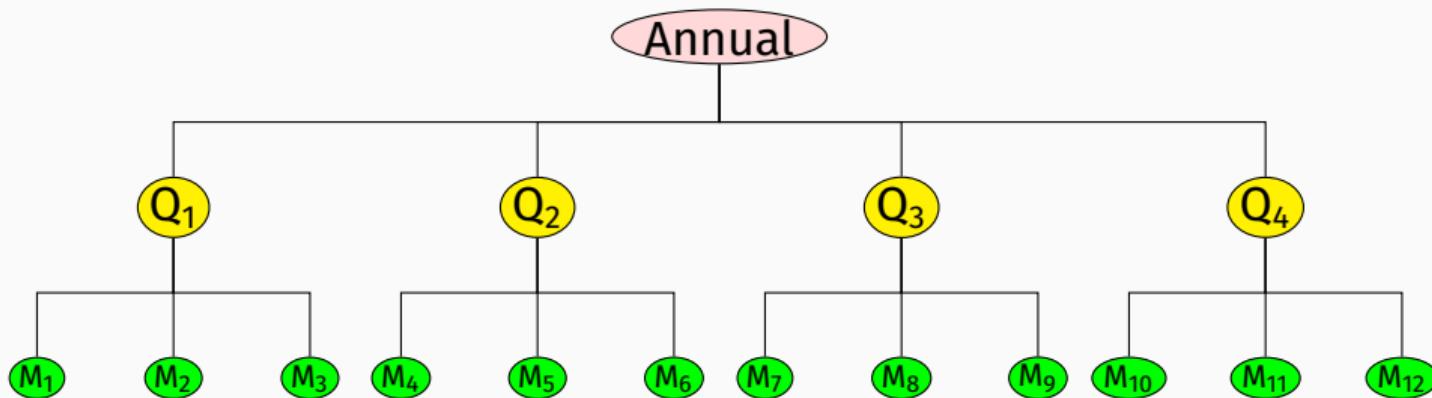


A Panagiotelis, P Gamakumara, G Athanasopoulos, and RJ Hyndman (2023). “**Probabilistic forecast reconciliation: Properties, evaluation and score optimisation**”. In: *European Journal of Operational Research* 306.2, pp. 693–706. URL: <https://doi.org/10.1016/j.ejor.2022.07.040>

Outline

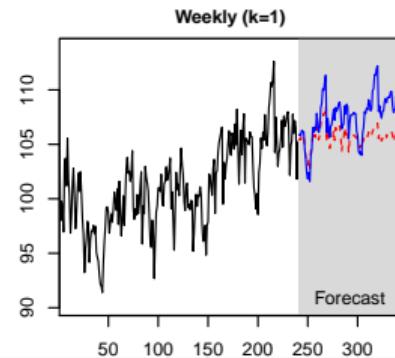
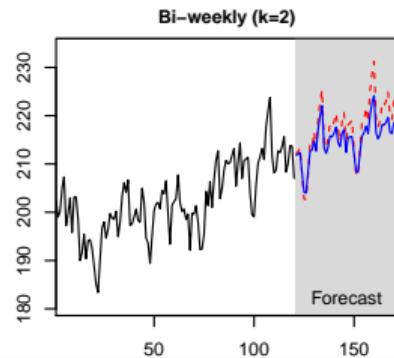
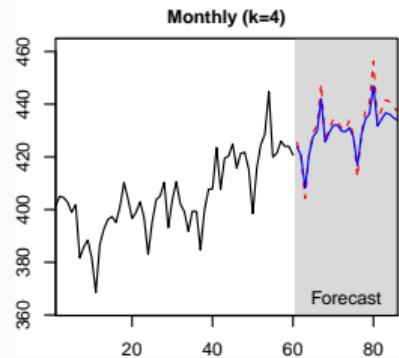
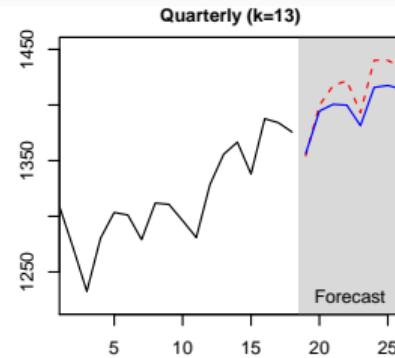
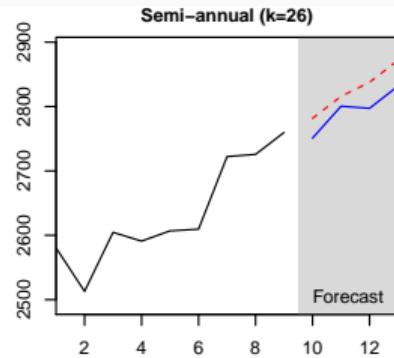
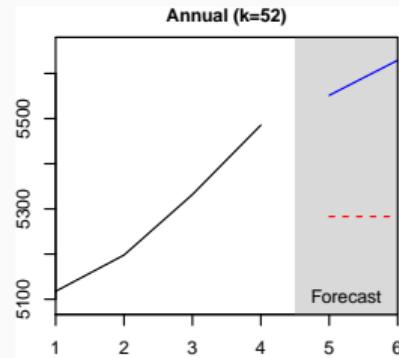
- 1 Improving hierarchical forecasts
- 2 Improving univariate forecasts
- 3 Improving cross-temporal forecasts
- 4 Improving multivariate forecasts
- 5 Final comments

Temporal reconciliation: monthly data



- Forecast series at each available frequency.
- Optimally combine forecasts within the same year.

Example: Accident & emergency services demand



Temporal Hierarchical Forecasting - THieF



G Athanasopoulos, RJ Hyndman, N Kourentzes, and F Petropoulos (2017). “**Forecasting with Temporal Hierarchies**”. In: *European Journal of Operational Research* 262, pp. 60–74. URL: <https://doi.org/10.1016/j.ejor.2017.02.046>



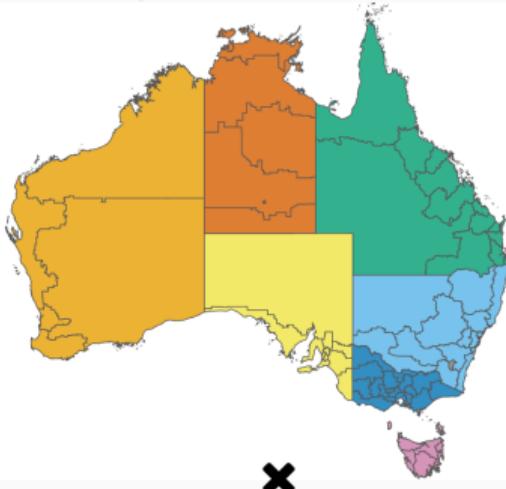
N Kourentzes and G Athanasopoulos (2021). “**Elucidate structure in intermittent demand series**”. In: *European Journal of Operational Research* 288.1, pp. 141–152. URL: <https://doi.org/10.1016/j.ejor.2020.05.046>.

Outline

- 1 Improving hierarchical forecasts
- 2 Improving univariate forecasts
- 3 Improving cross-temporal forecasts
- 4 Improving multivariate forecasts
- 5 Final comments

Cross-temporal reconciliation

Geographical division



Purpose of travel

Holiday, Visiting friends & relatives, Business, Other

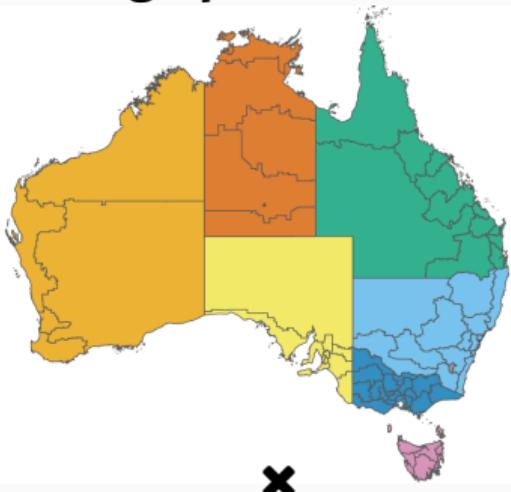
■ Cross-sectional aggregations

(geographical divisions × purpose of travel)

	AUS	States	Zones*	Regions	Tot
geographical	1	7	21	76	105
purpose	4	28	84	304	420
total	5	35	105	380	525

Cross-temporal reconciliation

Geographical division



Purpose of travel

Holiday, Visiting friends & relatives, Business, Other

■ Cross-sectional aggregations

(geographical divisions × purpose of travel)

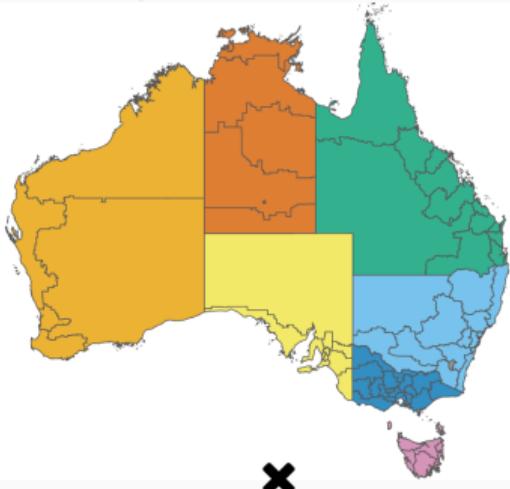
	AUS	States	Zones*	Regions	Tot
geographical	1	7	21	76	105
purpose	4	28	84	304	420
total	5	35	105	380	525

■ Temporal aggregations, frequencies:

- ▶ Monthly
- ▶ Bi-Monthly
- ▶ Quarterly
- ▶ Four-Monthly
- ▶ Semi-Annual
- ▶ Annual

Cross-temporal reconciliation

Geographical division



Purpose of travel

Holiday, Visiting friends & relatives, Business, Other

Total: 3150 Series

■ Cross-sectional aggregations

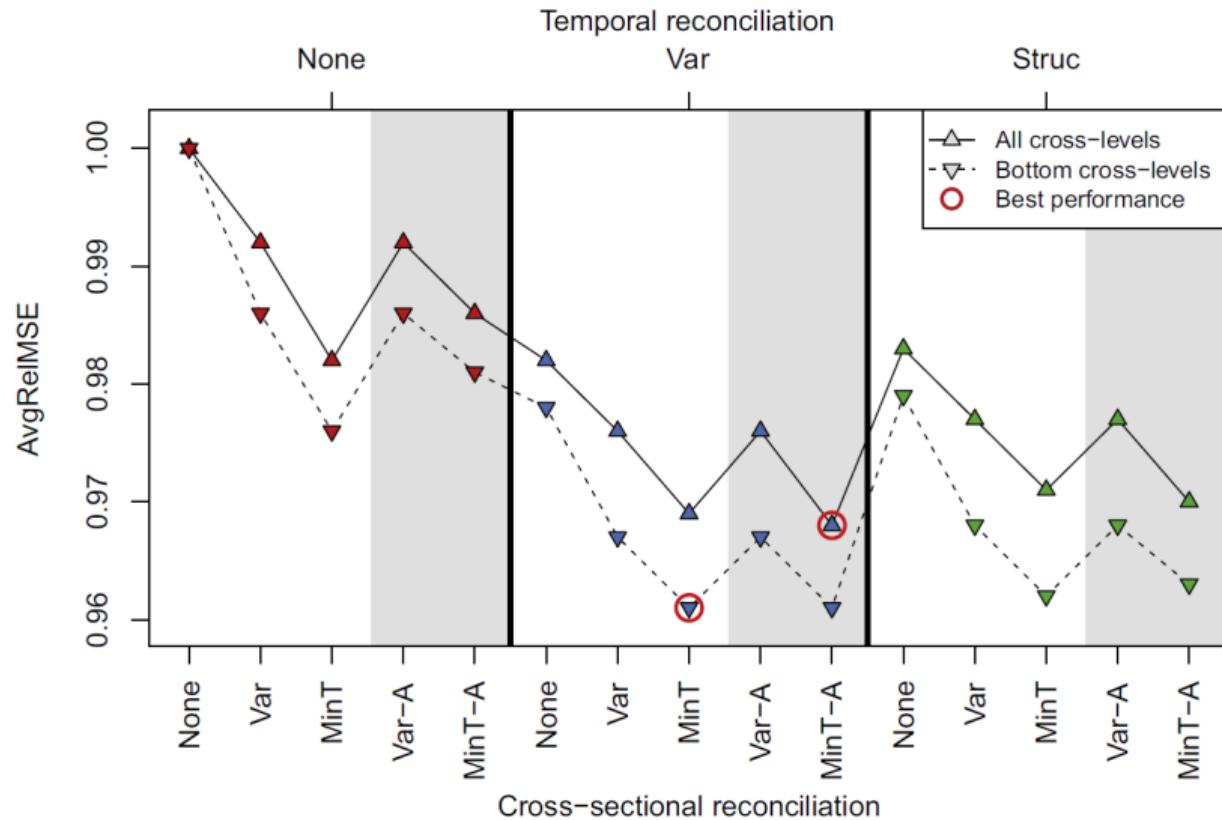
(geographical divisions × purpose of travel)

	AUS	States	Zones*	Regions	Tot
geographical	1	7	21	76	105
purpose	4	28	84	304	420
total	5	35	105	380	525

■ Temporal aggregations, frequencies:

- ▶ Monthly
- ▶ Bi-Monthly
- ▶ Quarterly
- ▶ Four-Monthly
- ▶ Semi-Annual
- ▶ Annual

Cross-temporal reconciliation



Cross-temporal reconciliation



N Kourentzes and G Athanasopoulos (2019). “**Cross-temporal coherent forecasts for Australian tourism**”. In: *Annals of Tourism Research* 75, pp. 393–409. URL: <https://doi.org/10.1016/j.annals.2019.02.001>



D Girolimetto, G Athanasopoulos, T Di Fonzo, and RJ Hyndman. (2024). “**Cross-temporal probabilistic forecast reconciliation: methodological and practical issues**”. In: *International J Forecasting* 40.3, pp. 1134–1151. URL: <https://doi.org/10.1016/j.ijforecast.2023.10.003>



G Athanasopoulos, RJ Hyndman, N Kourentzes, and A Panagiotelis. (2024). “**Forecast reconciliation: a review**”. In: *International J Forecasting* 40.2, pp. 430–456. URL: <https://doi.org/10.1016/j.ijforecast.2023.10.010>

Outline

- 1 Improving hierarchical forecasts
- 2 Improving univariate forecasts
- 3 Improving cross-temporal forecasts
- 4 Improving multivariate forecasts
- 5 Final comments

FLAP (Forecast Linear Augmented Projection)

Intuition:

- Suppose we are interested in multivariate forecasting but do not have linear (or non-linear) constraints.
- Can reconciliation help?

FLAP (Forecast Linear Augmented Projection)

Intuition:

- Suppose we are interested in multivariate forecasting but do not have linear (or non-linear) constraints.
- Can reconciliation help?

Can we find linear components that:

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

Outline of FLAP Implementation

- We want to forecast a multivariate series $\mathbf{y}_t \in \mathbb{R}^m$.

Outline of FLAP Implementation

- We want to forecast a multivariate series $\mathbf{y}_t \in \mathbb{R}^m$.
- Construct synthetic linear components $\mathbf{c}_t \in \mathbb{R}^p$ where
$$\mathbf{c}_t = \Phi \mathbf{y}_t.$$

Outline of FLAP Implementation

- We want to forecast a multivariate series $\mathbf{y}_t \in \mathbb{R}^m$.
- Construct synthetic linear components $\mathbf{c}_t \in \mathbb{R}^p$ where $\mathbf{c}_t = \Phi \mathbf{y}_t$. The choice of Φ is arbitrary.

Outline of FLAP Implementation

- We want to forecast a multivariate series $\mathbf{y}_t \in \mathbb{R}^m$.
- Construct synthetic linear components $\mathbf{c}_t \in \mathbb{R}^p$ where $\mathbf{c}_t = \Phi \mathbf{y}_t$. The choice of Φ is arbitrary.
- The augmented vector $(\mathbf{y}'_t, \mathbf{c}'_t)'$ $\in \mathbb{R}^m$ coheres to known linear constraints.

Outline of FLAP Implementation

- We want to forecast a multivariate series $\mathbf{y}_t \in \mathbb{R}^m$.
- Construct synthetic linear components $\mathbf{c}_t \in \mathbb{R}^p$ where $\mathbf{c}_t = \Phi \mathbf{y}_t$. The choice of Φ is arbitrary.
- The augmented vector $(\mathbf{y}'_t, \mathbf{c}'_t)' \in \mathbb{R}^m$ coheres to known linear constraints.
- Produce forecasts for original series and components $\hat{\mathbf{y}}_{t+h}$ and $\hat{\mathbf{c}}_{t+h}$.

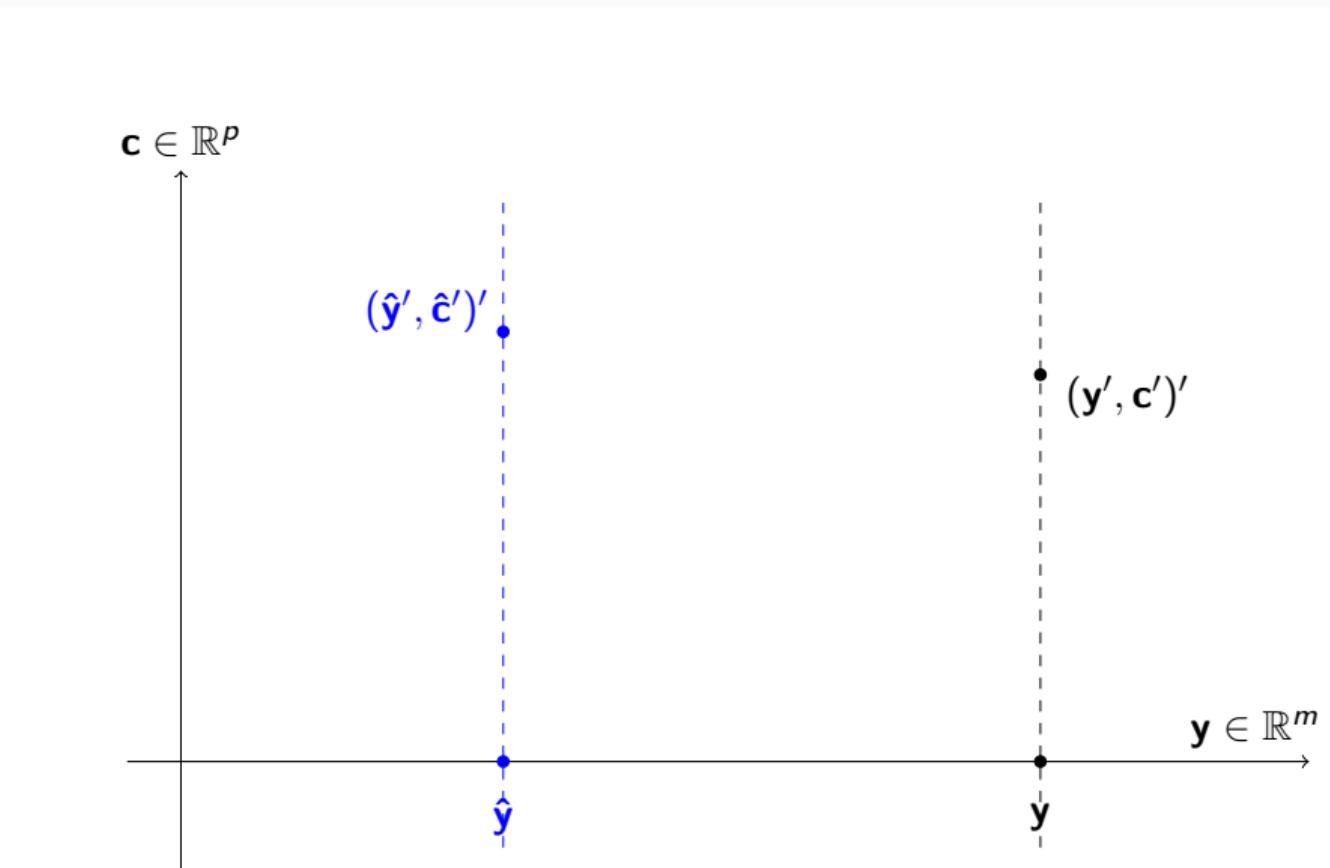
Outline of FLAP Implementation

- We want to forecast a multivariate series $\mathbf{y}_t \in \mathbb{R}^m$.
- Construct synthetic linear components $\mathbf{c}_t \in \mathbb{R}^p$ where $\mathbf{c}_t = \Phi \mathbf{y}_t$. The choice of Φ is arbitrary.
- The augmented vector $(\mathbf{y}'_t, \mathbf{c}'_t)' \in \mathbb{R}^m$ coheres to known linear constraints.
- Produce forecasts for original series and components $\hat{\mathbf{y}}_{t+h}$ and $\hat{\mathbf{c}}_{t+h}$.
- Project forecasts onto the \mathbb{R}^m coherent subspace using MinT, resulting in $\tilde{\mathbf{y}}_{t+h}$.

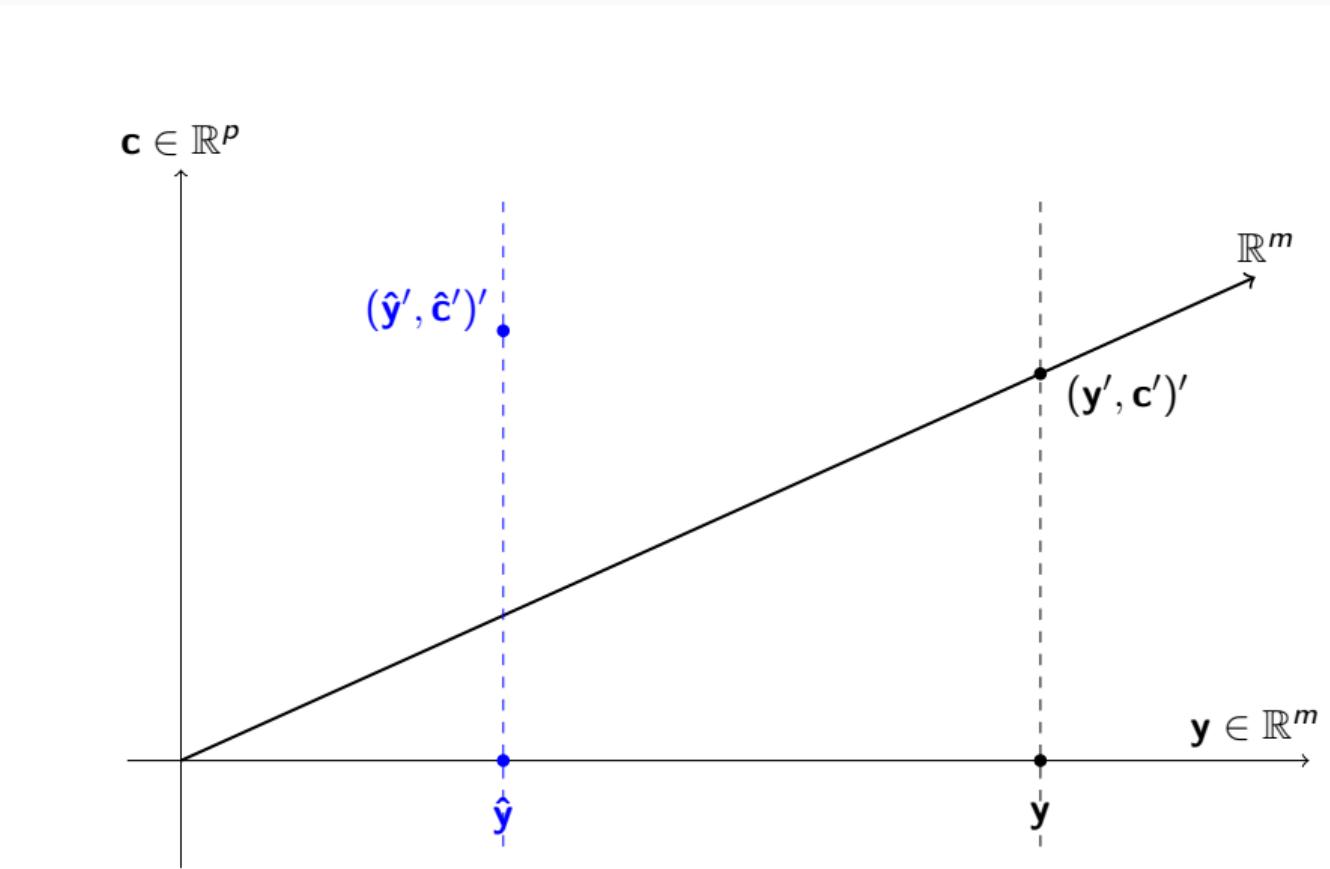
Geometry of FLAP



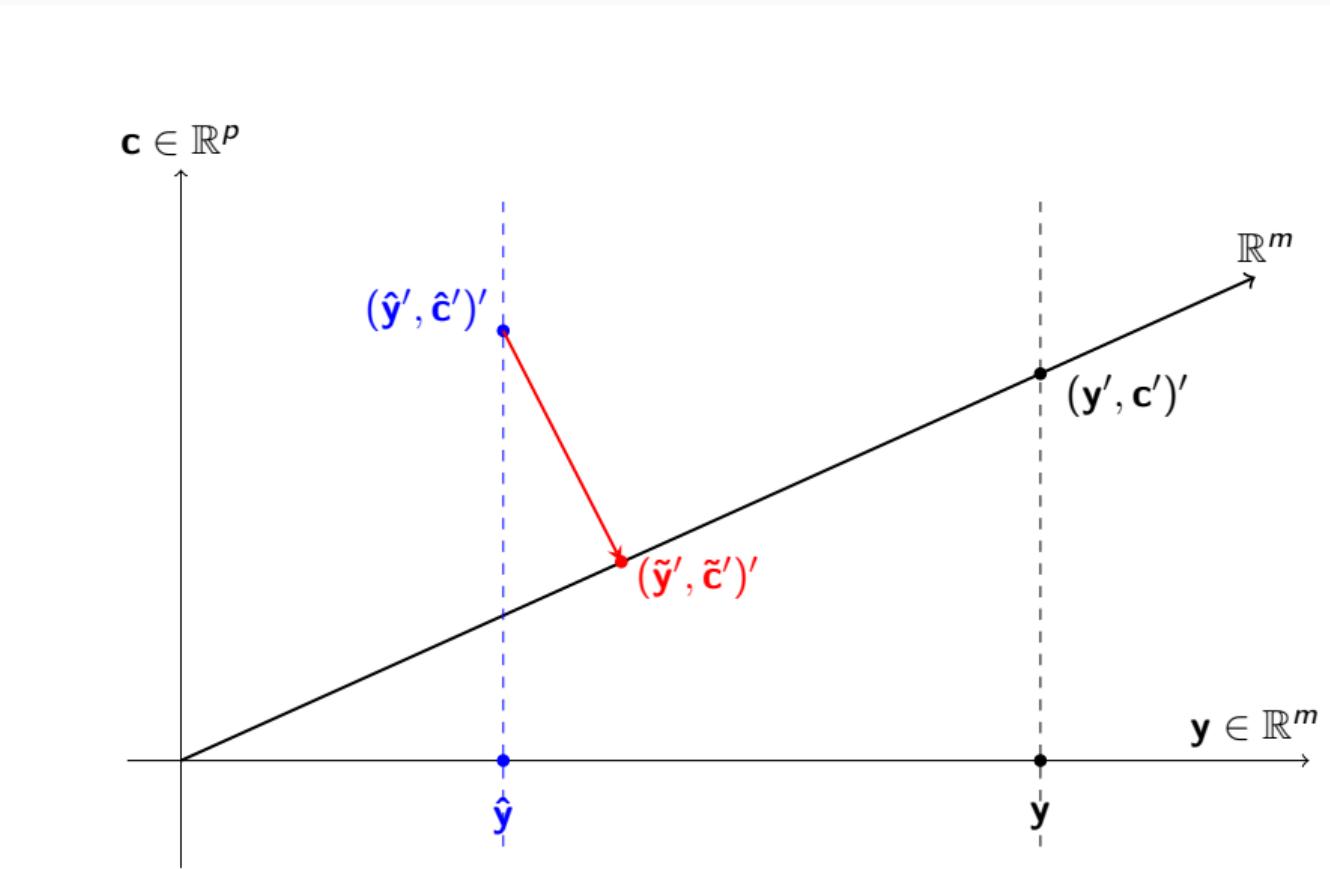
Geometry of FLAP



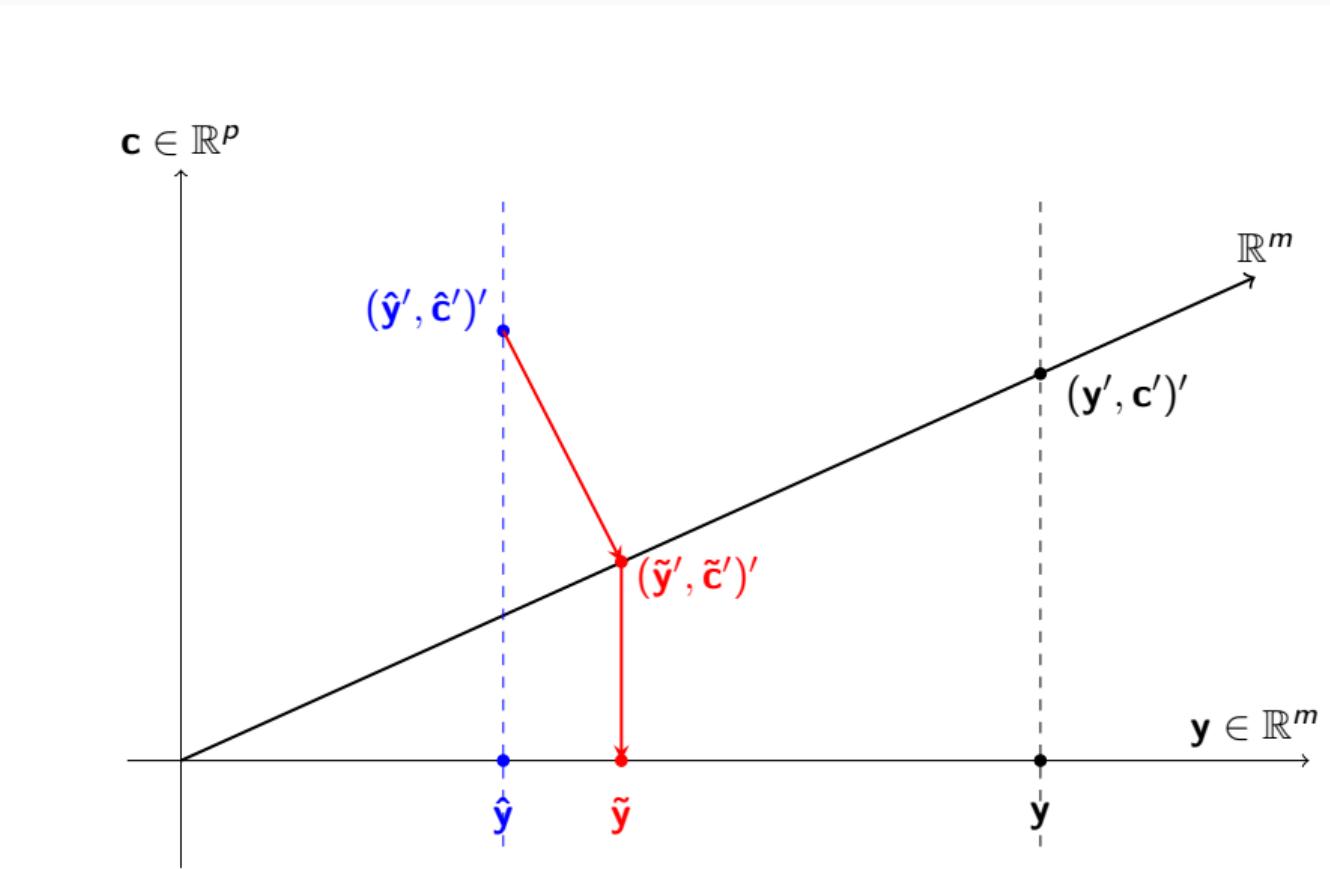
Geometry of FLAP



Geometry of FLAP



Geometry of FLAP



Key results based on MinT

- 1 The forecast error variance is **reduced** with FLAP
 - ▶ $\text{Var}(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{t+h}) - \text{Var}(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{t+h})$ is also **positive semi-definite**.

Key results based on MinT

- 1 The forecast error variance is **reduced** with FLAP
 - ▶ $\text{Var}(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{t+h}) - \text{Var}(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{t+h})$ is also **positive semi-definite**.
- 2 The forecast error variance **monotonically** decreases with increasing number of components
 - ▶ 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.

Key results based on MinT

- 1 The forecast error variance is **reduced** with FLAP
 - ▶ $\text{Var}(\mathbf{y}_{t+h} - \hat{\mathbf{y}}_{t+h}) - \text{Var}(\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_{t+h})$ is also **positive semi-definite**.
- 2 The forecast error variance **monotonically** decreases with increasing number of components
 - ▶ 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.
- 3 The forecast projection is **optimal** to achieve minimum forecast error variance for each series.

No free lunch

- In practice we need to estimate $\mathbf{W}_h = \text{Var}(\mathbf{z}_{t+h} - \hat{\mathbf{z}}_{t+h})$.
- Theory works with known \mathbf{W}_h .

No free lunch

- In practice we need to estimate $\mathbf{W}_h = \text{Var}(\mathbf{z}_{t+h} - \hat{\mathbf{z}}_{t+h})$.
- Theory works with known \mathbf{W}_h .
- The quality of covariance matrix estimates deteriorate with higher dimension.

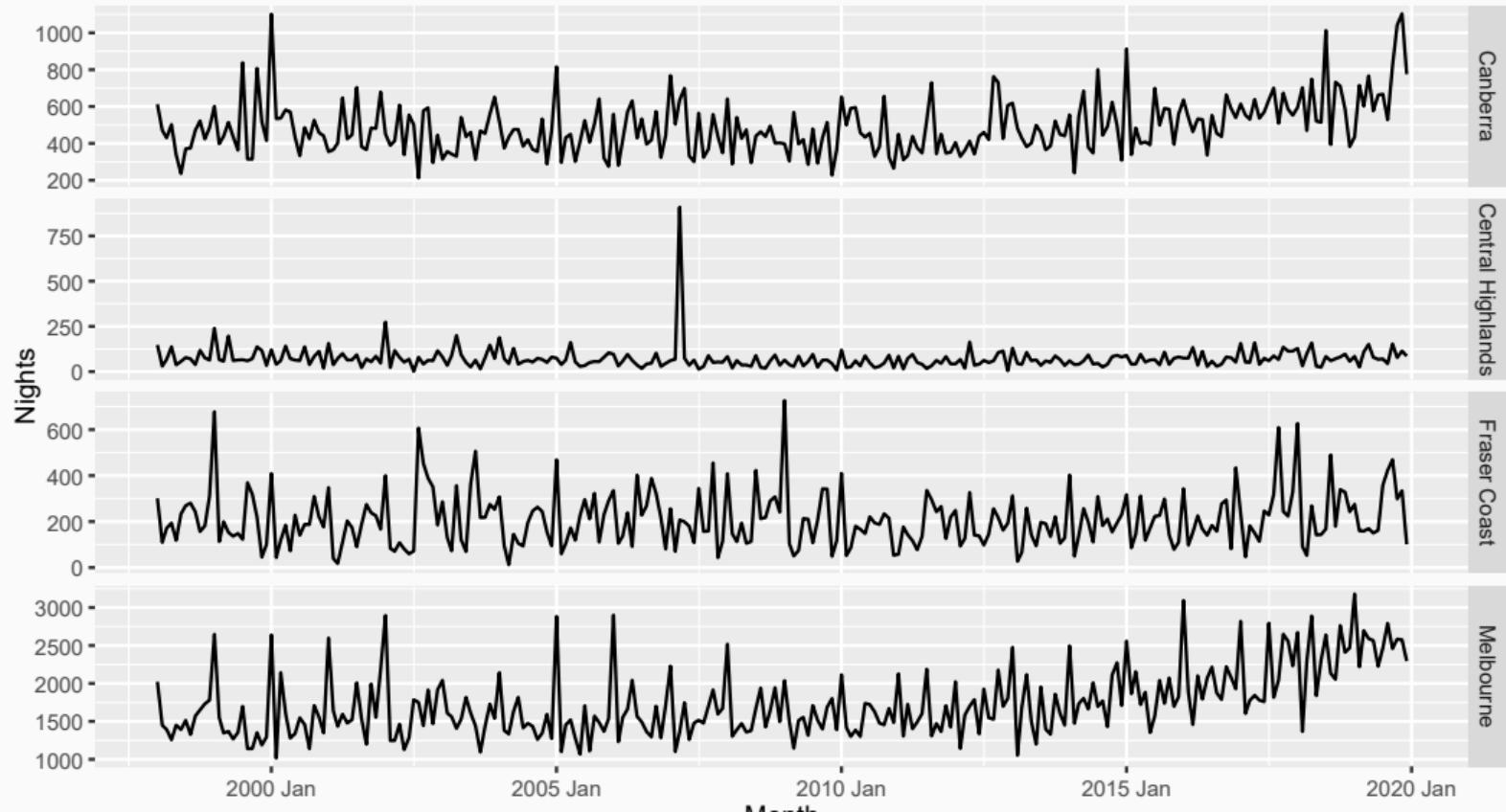
No free lunch

- In practice we need to estimate $\mathbf{W}_h = \text{Var}(\mathbf{z}_{t+h} - \hat{\mathbf{z}}_{t+h})$.
- Theory works with known \mathbf{W}_h .
- The quality of covariance matrix estimates deteriorate with higher dimension.
- However for finite dimension, the benefit of FLAP outweighs errors in estimating covariance matrix.

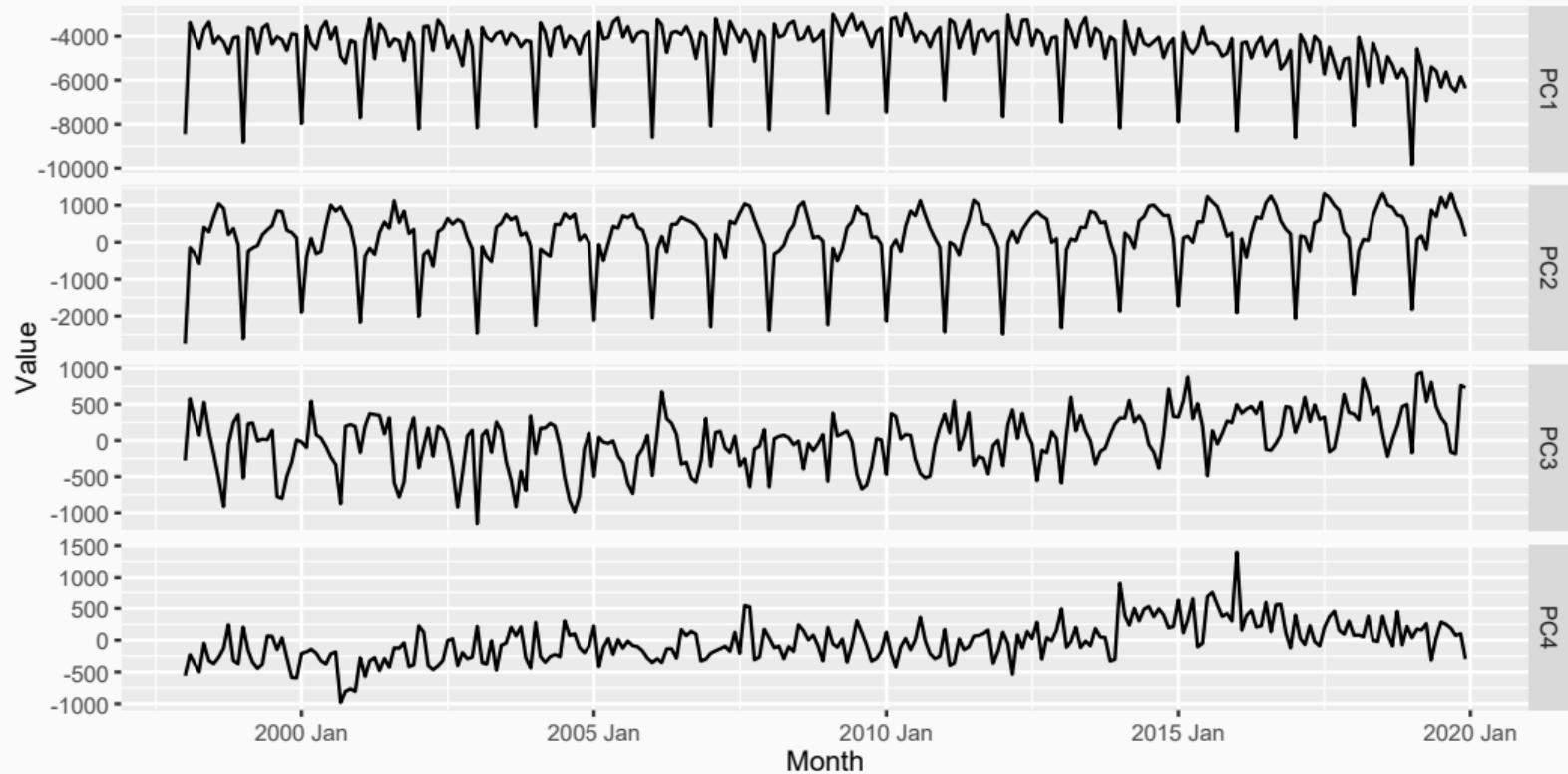
Monthly Australian regional tourism

- 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.

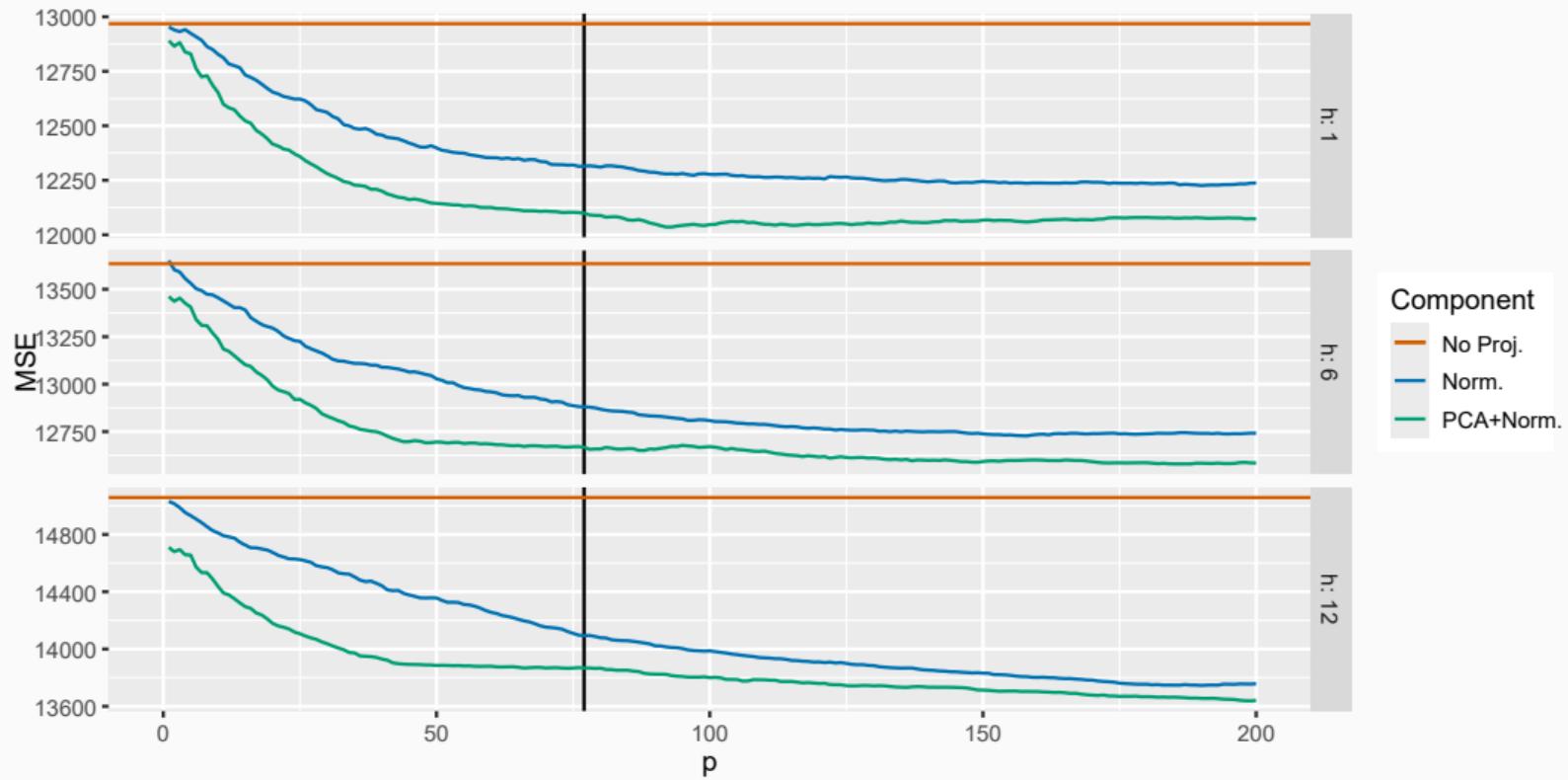
Monthly Australian regional tourism



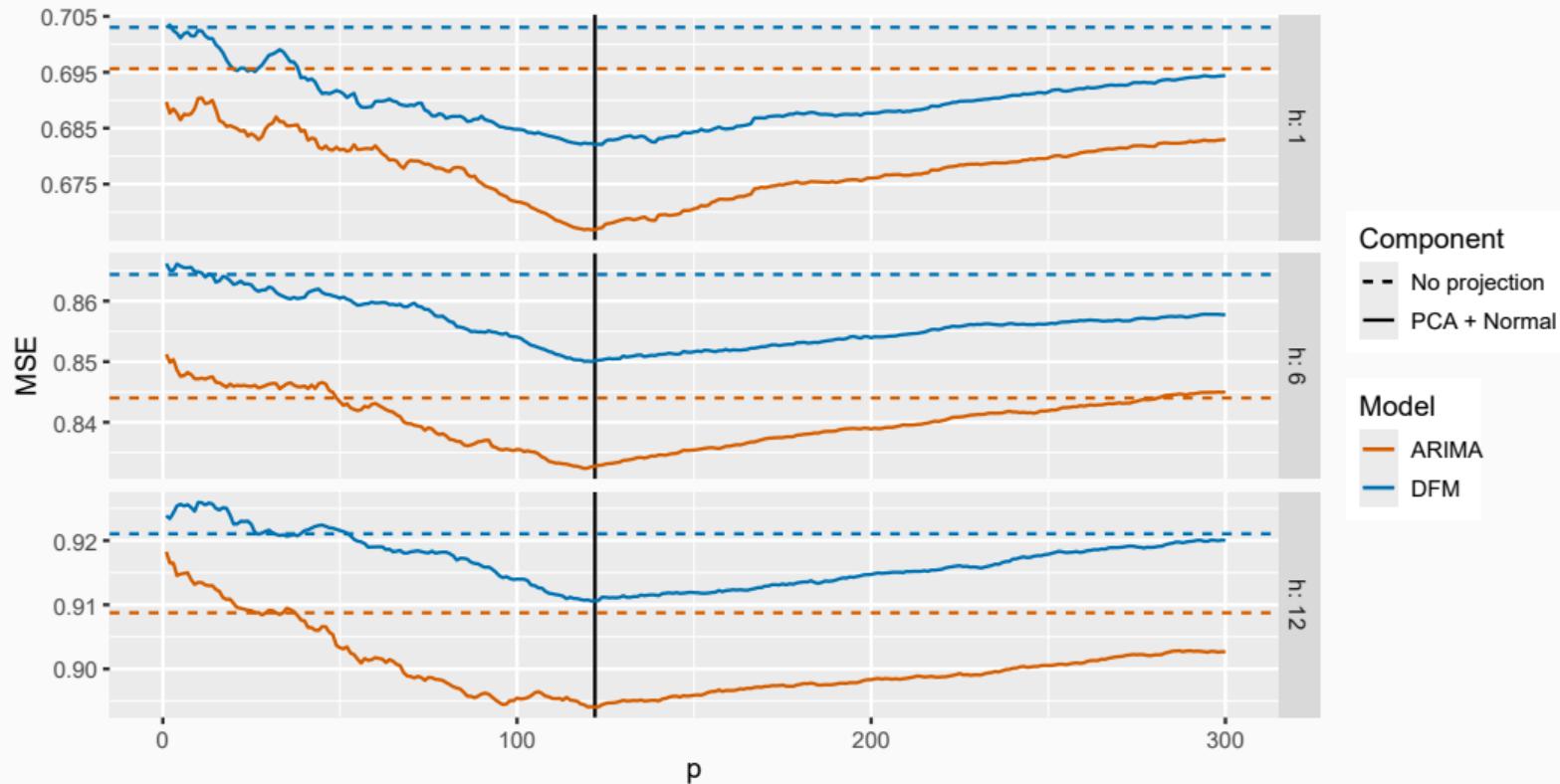
Monthly Australian regional tourism



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.



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")
```

Outline

- 1 Improving hierarchical forecasts
- 2 Improving univariate forecasts
- 3 Improving cross-temporal forecasts
- 4 Improving multivariate forecasts
- 5 Final comments

Forecasting: Principles and Practice

0Texts.com

Rob J Hyndman
George Athanasopoulos

FORECASTING PRINCIPLES AND PRACTICE

A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.



Rob J Hyndman
George Athanasopoulos

FORECASTING PRINCIPLES AND PRACTICE

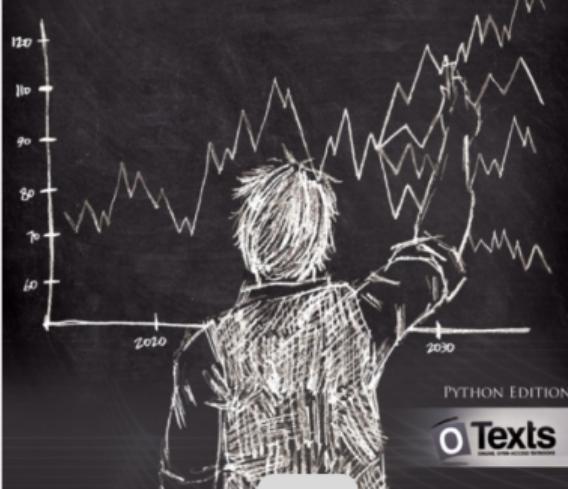
A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.



Rob J Hyndman, George Athanasopoulos,
Azul Garza, Cristian Challu,
Max Mergenthaler, Kin G Olivares

FORECASTING PRINCIPLES AND PRACTICE, THE PYTHONIC WAY

A comprehensive introduction to the latest forecasting methods using Python. Learn to improve your forecast accuracy using dozens of real data examples.



Forecasting: Principles and Practice

0Texts.com

Rob J Hyndman
George Athanasopoulos

FORECASTING PRINCIPLES AND PRACTICE

A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.



Rob J Hyndman
George Athanasopoulos

ΠΡΟΒΛΕΨΗ ΑΡΧΕΣ ΚΑΙ ΠΡΑΚΤΙΚΗ

Επιστημονική Επιωλεία - Απόδοση
Ιωάννης Α. Νίκας
Αθανάσιος Κούτρας



Rob J Hyndman, George Athanasopoulos,
Azul Garza, Cristian Challu,
Max Mergenthaler, Kin G Olivares

FORECASTING PRINCIPLES AND PRACTICE, THE PYTHONIC WAY

A comprehensive introduction to the latest forecasting methods using Python. Learn to improve your forecast accuracy using dozens of real data examples.



Links

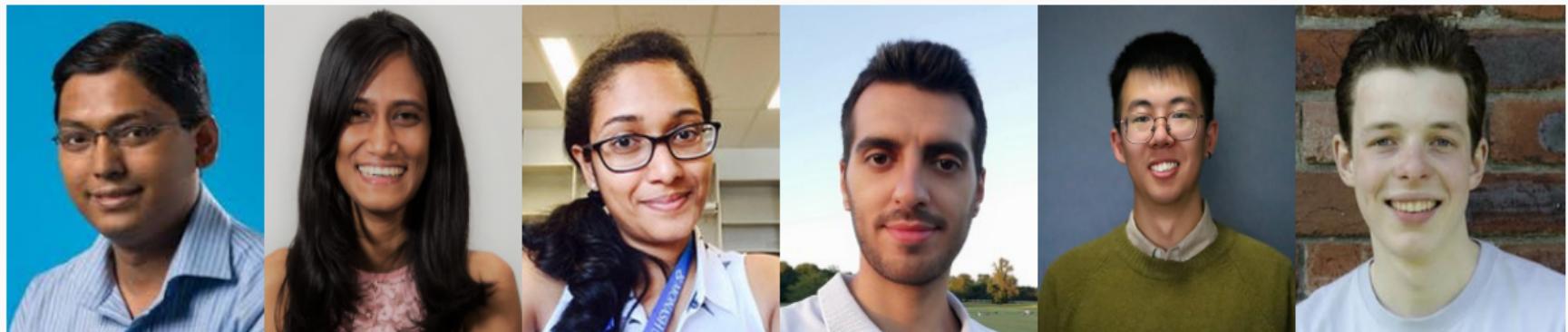


Postdoc opportunity



Link to slides

Thank you to



Other information

Forecasting: Principles and Practice <https://otexts.com/fppgr/>
(Thank you to [Ioannis Nikas](#) and [Athanasios Koutras](#))

Monash webpage

<https://research.monash.edu/en/persons/george-athanasopoulos>

Thank you!

-  Athanasopoulos, G, R Ahmed, and R Hyndman (2009). “**Hierarchical forecasts for Australian domestic tourism**”. In: *International Journal of Forecasting* 25, pp. 146–166. URL: <https://doi.org/10.1016/j.ijforecast.2008.07.004>.
-  Athanasopoulos, G, RJ Hyndman, N Kourentzes, and A Panagiotelis (2024). “**Forecast reconciliation: a review**”. In: *International J Forecasting* 40.2, pp. 430–456. URL: <https://doi.org/10.1016/j.ijforecast.2023.10.010>.
-  Athanasopoulos, G, RJ Hyndman, N Kourentzes, and F Petropoulos (2017). “**Forecasting with Temporal Hierarchies**”. In: *European Journal of Operational Research* 262, pp. 60–74. URL: <https://doi.org/10.1016/j.ejor.2017.02.046>.

- Girolimetto, D, G Athanasopoulos, T Di Fonzo, and RJ Hyndman (2024). “**Cross-temporal probabilistic forecast reconciliation: methodological and practical issues**”. In: *International J Forecasting* 40.3, pp. 1134–1151. URL: <https://doi.org/10.1016/j.ijforecast.2023.10.003>.
- Hyndman, RJ, R Ahmed, G Athanasopoulos, and H Shang (2011). “**Optimal combination forecasts for hierarchical time series**”. In: *Computational Statistics & Data Analysis* 55.9, pp. 2579–2589. URL: <https://doi.org/10.1016/j.csda.2011.03.006>.
- Hyndman, RJ and G Athanasopoulos (2021). **Forecasting: principles and practice**. 3nd Edn. Melbourne, Australia. URL: <http://otexts.com/fpp3/>.

-  Kourentzes, N and G Athanasopoulos (2019). “**Cross-temporal coherent forecasts for Australian tourism**”. In: *Annals of Tourism Research* 75, pp. 393–409. URL: <https://doi.org/10.1016/j.annals.2019.02.001>.
- (2021). “**Elucidate structure in intermittent demand series**”. In: *European Journal of Operational Research* 288.1, pp. 141–152. URL: <https://doi.org/10.1016/j.ejor.2020.05.046>.
-  Panagiotelis, A, G Athanasopoulos, P Gamakumara, and RJ Hyndman (2021). “**Forecast reconciliation: A geometric view with new insights on bias correction**”. In: *International Journal of Forecasting* 37.1, pp. 343–359. URL: <https://doi.org/10.1016/j.ijforecast.2020.06.004>.

-  Panagiotelis, A, P Gamakumara, G Athanasopoulos, and RJ Hyndman (2021). “Forecast reconciliation: A geometric view with new insights on bias correction”. In: *International J Forecasting* 37.1, pp. 343–359.
- (2023). “Probabilistic forecast reconciliation: Properties, evaluation and score optimisation”. In: *European Journal of Operational Research* 306.2, pp. 693–706. URL:
<https://doi.org/10.1016/j.ejor.2022.07.040>.
-  Wickramasuriya, SL, G Athanasopoulos, and RJ Hyndman (2019). “Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization”. In: *Journal of the American Statistical Association* 114.526, pp. 804–819. URL: <https://doi.org/10.1080/01621459.2018.1448825>.



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>.