Forecast Congruence

A Quantity to Align Forecasts and Inventory Decisions

Kandrika Pritularga and Nikos Kourentzes

Centre for Marketing Analytics, Lancaster University

Skövde AI Lab, University of Skövde

Seminar at Drexel University

6th of December 2024

Marketing Analytics and Forecasting | Lancaster University Management School

Outline

- A tale between forecasting and decisions
- From users' perspectives…
- Forecast congruence
- Congruence and inventory decisions
- Conclusions

CMAF

CMAF

A tale between forecasting and decisions

- Typical focus on forecast accuracy
  - A "forecasting microcosm"!
- Forecasts → Decisions
  - Actions taken sequentially
  - Accuracy is a convenient (but incomplete) proxy
  - Different loss/objective functions become prominent
- Forecasts used for decisions depend on the choice of error metrics
  - Inconsistent findings on the relationship between the quality of forecasts and decisions
  - Choice of accuracy metric 'should' match the decision context (Athanasopoulos & Kourentzes 2024)

A tale between forecasting and decisions

- **(Quantile) Forecasts** are often used as an input for **inventory control**
  - Accurate point (mean) forecasts → Useful quantile forecasts → Inventory optimisation → Low inventory costs
- Accuracy (MSE) is used as a proxy of the standard deviation of the forecasts
  - The error follows a normal distribution, and its variance is estimated well
- What happens if the model is wrong?
  - The error won't follow a normal distribution → variance?
- Some argue that forecasting performance needs to be assessed with inventory decision metrics

A tale between forecasting and decisions

However, the literature in the interface between forecasting and inventory management has a different story…

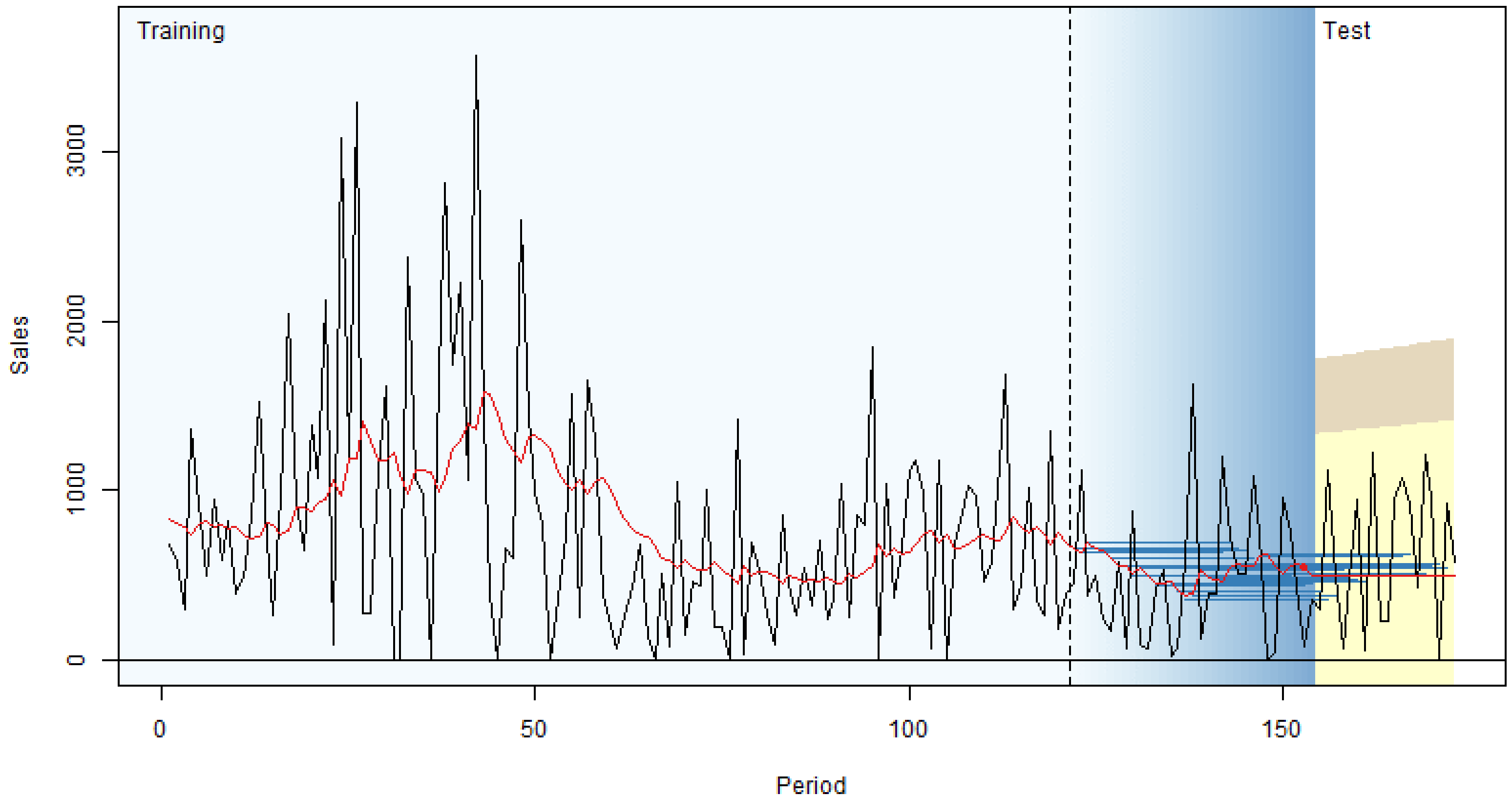
- **A low inventory cost** is characterized by **biased forecasts** (Kourentzes et al 2020)
- The relationship between accuracy and inventory costs depends on **the uncertainty of the time series and the simulation design** (Fildes & Kingsman 2013)

Problems in reality:

- We do not know the demand
- We need an alternative measure to assess forecasts that relates well with inventory decisions (not only accuracy)

From users' perspectives…

Rolling origin for forecast evaluation



Svetunkov (2024) & Kourentzes (2024)

CMAF

From users' perspectives…

"[T]he rolling forecast signals are way too noisy and have no value to us. We would rather ignore those troublesome data in forecasting processes."

Chuang et al (2021)

Chuang, H.H., Chou, Y. and Oliva, R. (2021) 'Cross-item learning for volatile demand forecasting: An intervention with predictive analytics', Journal of Operations Management, 67(7), pp. 828–852. Available at: https://doi.org/10.1002/joom.1152.

From users' perspectives…

- We conducted a small online survey amongst practitioners to understand the prevalence of the practice in firms

Table 1: Summary survey results (sample size = 21)

Forecasting model review interval	Responses	Are forecasts adjusted to be more 'stable' over time?	Responses
Every time	71.43%	No	19.05%
Longer review intervals	28.57%	Yes, in an ad-hoc manner	33.33%
		Yes, rule-based changes	47.62%

- Few of them update their models infrequently
- Adjustments to achieve 'stable' forecasts

So what?

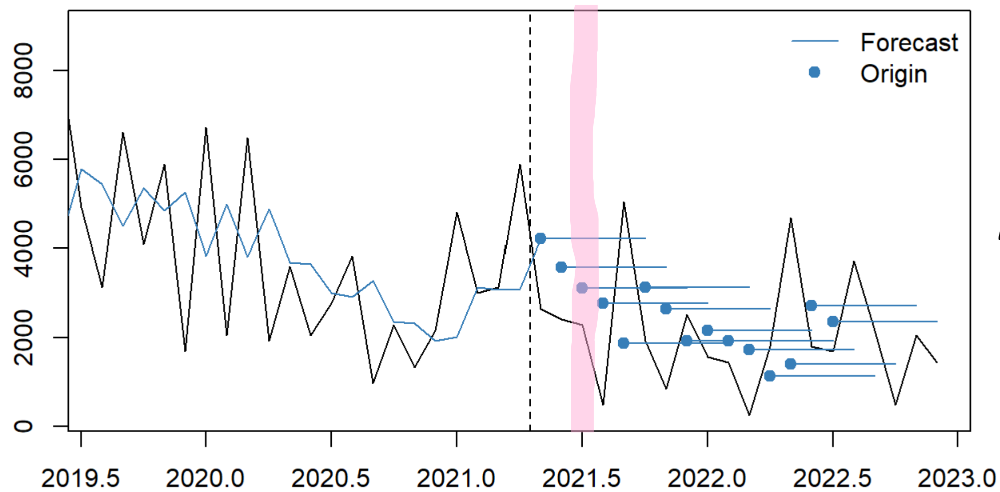
- In practice, accuracy might not be sufficient enough to achieve desirable inventory decision
- Decision makers prefer forecasts that are not so ‘jittery’ across origins

We need a way to define and measure this ‘overlooked’ property of forecasts

“Congruence”

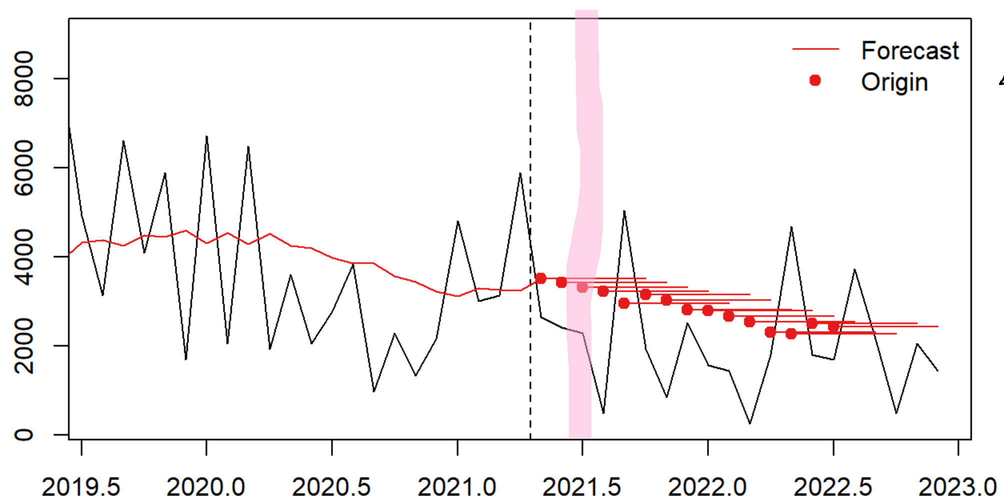
CMAF

Forecast congruence



Less congruent

More congruent



Congruence as error measure

- Let's think about how we produce forecasts

Observation

Horizon

$$
\begin{array}{c}
 \\ t+1 \\ t+2 \\ t+3 \\ t+4
\end{array}
\begin{array}{c}
\begin{array}{cccc} h=1 & h=2 & h=3 & h=4 \end{array} \\
\begin{bmatrix}
\hat{y}_{t+1|t} & \hat{y}_{t+1|t-1} & \hat{y}_{t+1|t-2} & \hat{y}_{t+1|t-3} \\
\hat{y}_{t+2|t+1} & \hat{y}_{t+2|t} & \hat{y}_{t+2|t-1} & \hat{y}_{t+2|t-2} \\
\hat{y}_{t+3|t+2} & \hat{y}_{t+3|t+1} & \hat{y}_{t+3|t} & \hat{y}_{t+3|t-1} \\
\hat{y}_{t+4|t+3} & \hat{y}_{t+4|t+2} & \hat{y}_{t+4|t+1} & \hat{y}_{t+4|t}
\end{bmatrix}
\end{array}
$$

Take a variance of forecasts across horizons

$$\hat{y}_{t+1|t} = y_{t+1} + e_{t+1|t}$$

$$
\begin{array}{c}
 \\ t+1 \\ t+2 \\ t+3 \\ t+4
\end{array}
\begin{array}{c}
\begin{array}{cccc} h=1 & h=2 & h=3 & h=4 \end{array} \\
\begin{bmatrix}
y_{t+1} + e_{t+1|t} & y_{t+1} + e_{t+1|t-1} & y_{t+1} + e_{t+1|t-2} & y_{t+1} + e_{t+1|t-3} \\
y_{t+2} + e_{t+2|t+1} & y_{t+2} + e_{t+2|t} & y_{t+2} + e_{t+2|t-1} & y_{t+2} + e_{t+1|t-2} \\
y_{t+3} + e_{t+3|t+2} & y_{t+3} + e_{t+3|t+1} & y_{t+3} + e_{t+3|t-2} & y_{t+3} + e_{t+1|t-1} \\
y_{t+4} + e_{t+4|t+3} & y_{t+4} + e_{t+4|t+2} & y_{t+4} + e_{t+4|t-3} & y_{t+4} + e_{t+1|t}
\end{bmatrix}
\end{array}
$$

$$\mathrm{Var}_h(\hat{y}_{t+1|t-h}) \approx \mathrm{Var}(y_{t+1}) + \mathrm{Var}_h(e_{t+1|t-h})$$

Zero!

Forecast congruence

- We can then construct a rolling-origin forecast error matrix
  - Column-wise (horizon) → accuracy (MSE)
  - Row-wise → congruence

Horizon

$$
\boldsymbol{e} = \begin{array}{c c}
 & \begin{array}{cccc} h=1 & h=2 & h=3 & h=4 \end{array} \\
\begin{array}{c} t+1 \\ t+2 \\ t+3 \\ t+4 \end{array} &
\begin{bmatrix}
e_{t+1|t} & e_{t+1|t-1} & e_{t+1|t-2} & e_{t+1|t-3} \\
e_{t+2|t+1} & e_{t+2|t} & e_{t+2|t-1} & e_{t+1|t-2} \\
e_{t+3|t+2} & e_{t+3|t+1} & e_{t+3|t-2} & e_{t+1|t-1} \\
e_{t+4|t+3} & e_{t+4|t+2} & e_{t+4|t-3} & e_{t+1|t}
\end{bmatrix}
\end{array}
$$

Observation

Congruence

$$
\begin{array}{c c}
\begin{pmatrix}
e_{1+l|l} & e_{1+l|l-1} & \cdots & e_{1+l|1+l-(l-1)} & e_{1+l|1} \\
e_{2+l|1+l} & e_{2+l|l} & \cdots & e_{2+l|2+l-(l-1)} & e_{1+l|2} \\
\vdots & \ddots & & & \vdots \\
 & & e_{t|t-h} & & \\
\vdots & & & \ddots & \vdots \\
e_{n-1|n-2} & e_{n-1|n-3} & \cdots & e_{n-1|n-1-(l-1)} & e_{n-1|n-1-l} \\
e_{n|n-1} & e_{n|n-2} & \cdots & e_{n|n-(l-1)} & e_{n|n-l}
\end{pmatrix}
&
\begin{array}{l}
\rightarrow \tau^2_{l+1} \\
\rightarrow \tau^2_{l+2} \\
\\
\quad\vdots \\
\\
\rightarrow \tau^2_{n-1} \\
\rightarrow \tau^2_{n}
\end{array} \\
\begin{array}{ccccc}
\downarrow & \downarrow & & \downarrow & \downarrow \\
\sigma^2_1 & \sigma^2_2 & \cdots & \sigma^2_{l-1} & \sigma^2_l
\end{array} &
\end{array}
$$

One-step ahead MSE

l-step ahead MSE

Forecast congruence

- We can then construct a rolling-origin forecast error matrix
  - Column-wise (horizon) → accuracy (MSE)
  - Row-wise → congruence

Congruence

Observation

$$
\begin{array}{c}
\begin{pmatrix}
e_{1+l|l} & e_{1+l|l-1} & \cdots & e_{1+l|1+l-(l-1)} & e_{1+l|1} \\
e_{2+l|1+l} & e_{2+l|l} & \cdots & e_{2+l|2+l-(l-1)} & e_{1+l|2} \\
\vdots & \ddots & & & \vdots \\
 & & e_{t|t-h} & & \\
\vdots & & & \ddots & \vdots \\
e_{n-1|n-2} & e_{n-1|n-3} & \cdots & e_{n-1|n-1-(l-1)} & e_{n-1|n-1-l} \\
e_{n|n-1} & e_{n|n-2} & \cdots & e_{n|n-(l-1)} & e_{n|n-l}
\end{pmatrix}
\begin{matrix}
\to & \tau^2_{l+1} \\
\to & \tau^2_{l+2} \\
 & \\
 & \vdots \\
 & \\
\to & \tau^2_{n-1} \\
\to & \tau^2_{n}
\end{matrix}
\\
\begin{matrix}
\downarrow & \downarrow & & \downarrow & \downarrow \\
\sigma^2_1 & \sigma^2_2 & \cdots & \sigma^2_{l-1} & \sigma^2_l
\end{matrix}
\end{array}
$$

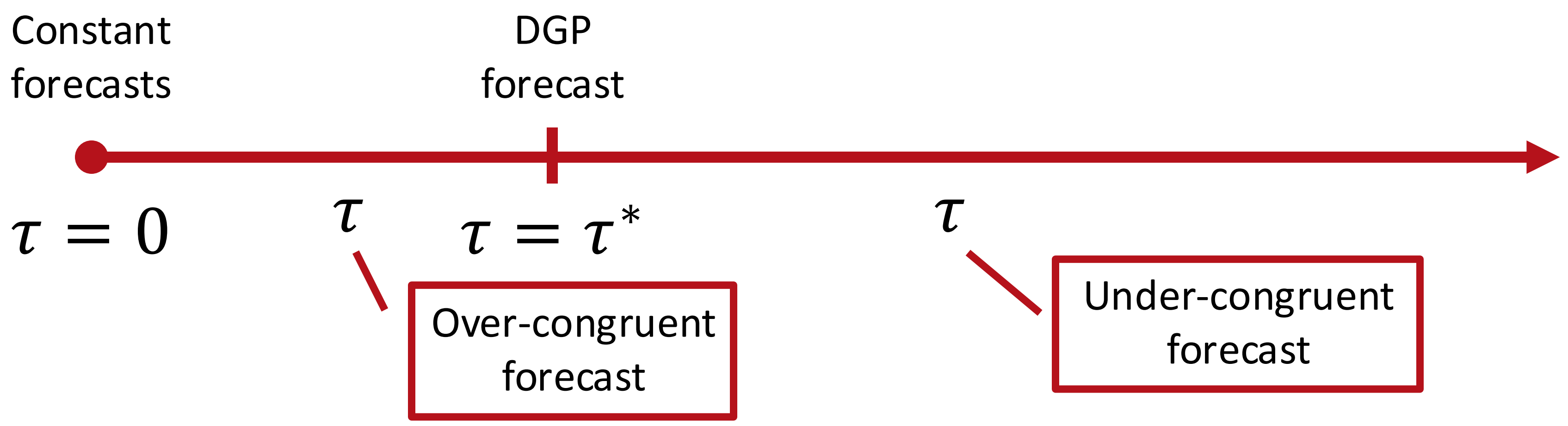
MSE

Horizon

- MSE and congruence are constructed from a matrix
  - Differ in value due to different lengths in rows and columns

Under- and over-congruent forecasts

How can we assess congruence?



What do under- and over-congruent forecasts mean?

- $\tau < \tau^*$, then the forecasts are over-congruent
- $\tau > \tau^*$, then the forecasts are under-congruent
- It's inspired from the idea of over-fitting and under-fitting

Under- and over-congruent forecasts

- Can we calculate the true congruence or τ^*?
- If we know the data generating process, we can!
- If the process is an AR(1)

$$\tau_t^* \approx \sqrt{\sum_{i=1}^{l}(l-i+1)\alpha^{2i}\,\varepsilon_{t-i+1}^2 + \left(\sum_{i=1}^{l}(l-i+1)\alpha^{i}\varepsilon_{t-i+1}\right)}$$

- Congruence is affected by
  - Model parameters (α)
  - Forecast horizon (l)
  - Variance of the innovations (ε_{t-i+1}^2)

Experiments

- We aim to observe a relationship between:
  - Accuracy metrics v congruence
  - Accuracy metrics v inventory decision metrics
  - Congruence v inventory decision metrics
- We need to design two tasks:
  - Forecasting task → Inventory task
- Data for analysis
  - Simulated data
    - Different processes, from simple AR to non-stationary seasonal processes
  - UK-based FMCG dataset

Experiments

- Standard order-up-to inventory policy with lost sales
- Target service levels: 90%, 95%, and 99%
- Lead time: 3 and 5 periods
- Review: 1 period **(?)**

$$S = \sum_{i=t}^{t+L} \hat{y}_{t+i|t} + Q_L(q),$$

- $Q_L(q)$ is the buffer stock for anticipated demand
  - Empirical prediction intervals
- Implemented for both datasets

Experiments

Forecasting models/ methods

- Naive
- ETS (re-estimate on each origin)
- ETS-Static (no re-estimation)
- ETS-Combination (Kolassa)
- ETS-Shrinkage (Pritularga et al)
- MAPA (Kourentzes et al)
- Demand generating process (DGP)

Demand generating processes

- ARIMA(1,0,0)
- ARIMA(1,0,1)
- ARIMA(1,1,1)
- SARIMA(1,0,0)(1,0,1)
- SARIMA(1,1,0)(1,0,1)

400 series for each process, with 132 monthly observations

- 36 obs: burn-in period for inventory
- 48 obs: training set
- 48 obs: test set

Evaluation: rolling-origin approach

- Increased training set
- Until test set is all used

Metrics we use

Forecasting metrics

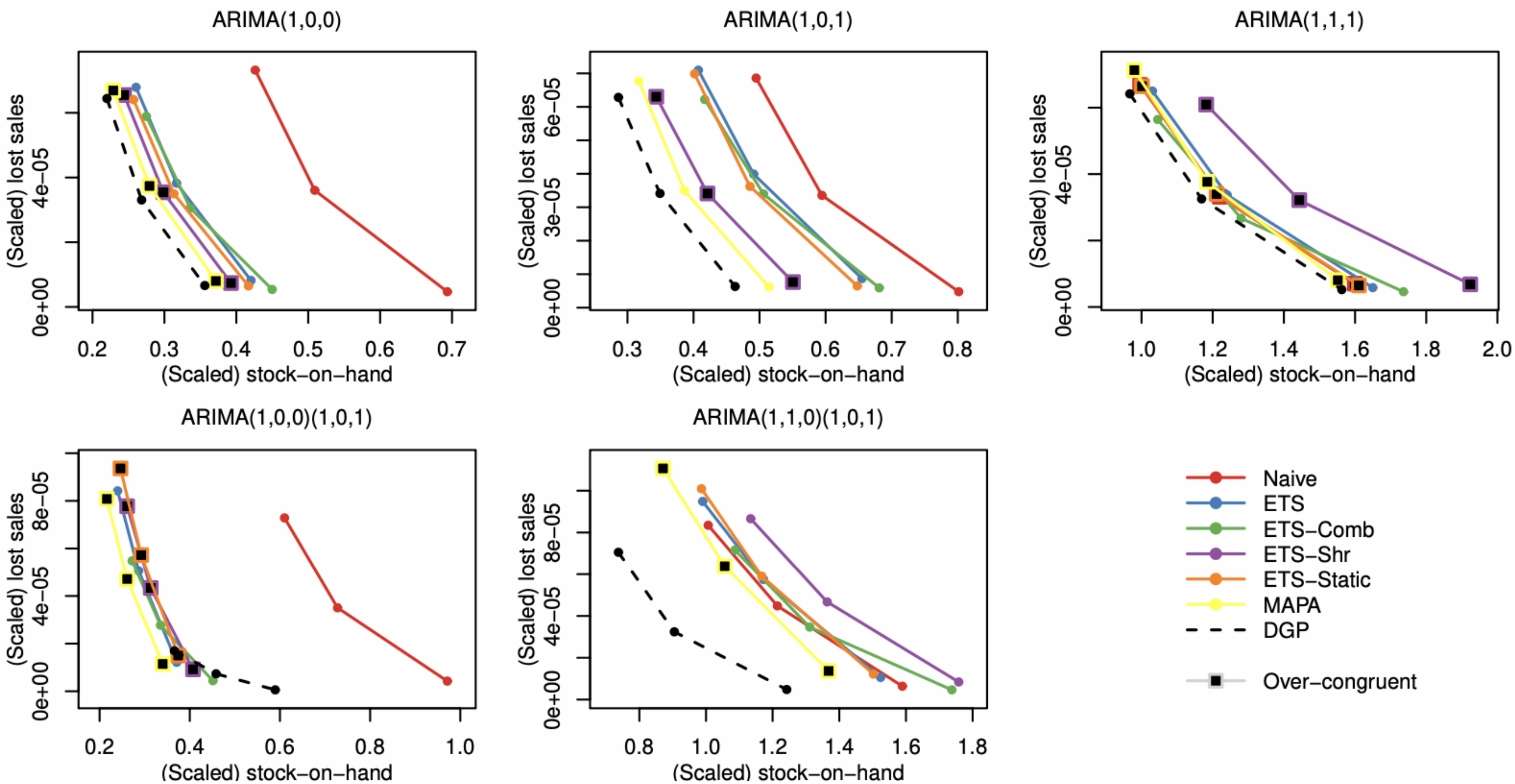
- RMSE (scaled): point forecast accuracy
- Pinball loss (scaled): ‘accuracy’ for prediction intervals
- Congruence

Inventory decision metrics

- CSL difference
- Lost sales
- Mean stock-on-hand
- <u>SD(stock-on-hand)</u>
- <u>SD(orders)</u>
- %periods that order placed

Simulation findings

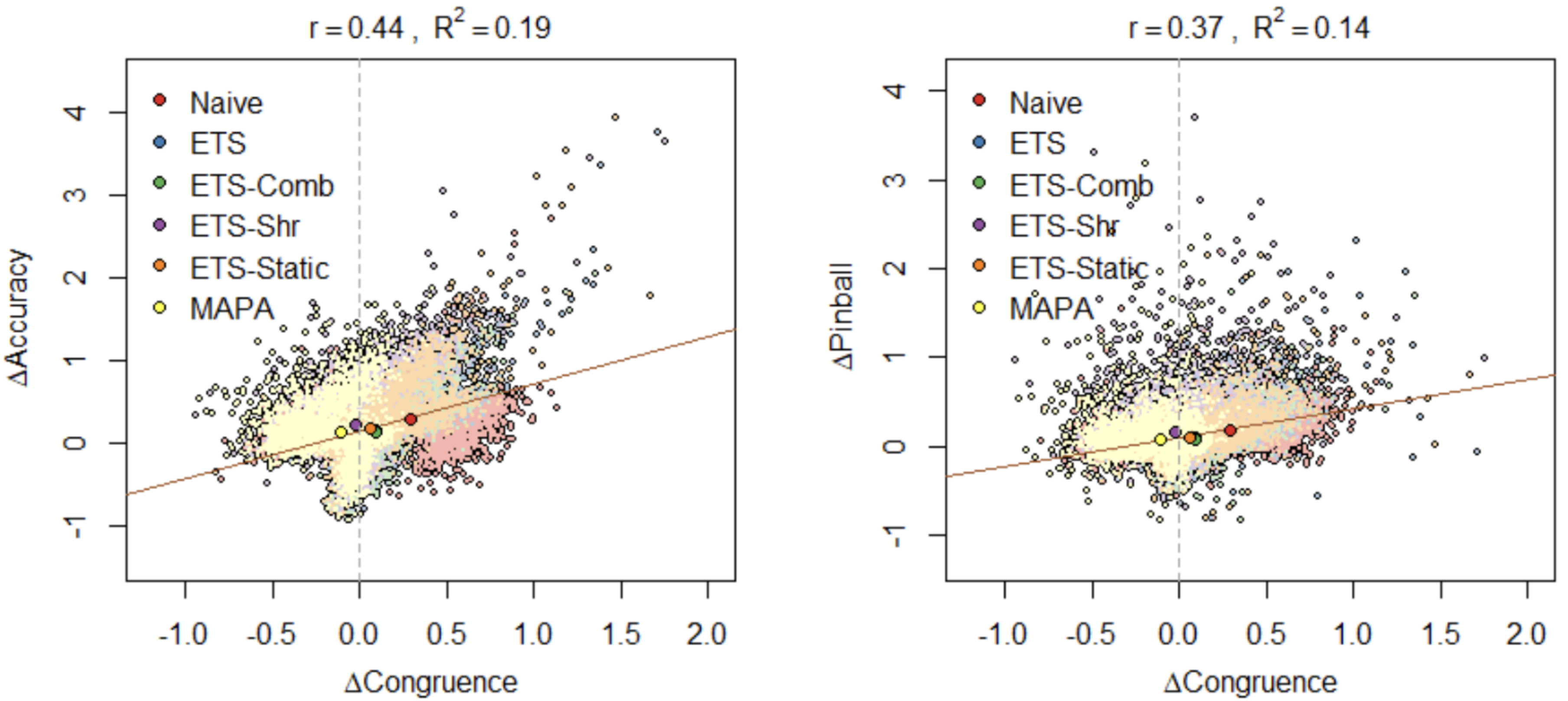
- Assessed forecasting models with an inventory decision trade-off
- The closer to zero, the better the inventory decisions are



- Over-congruent forecasts perform well in this trade-off

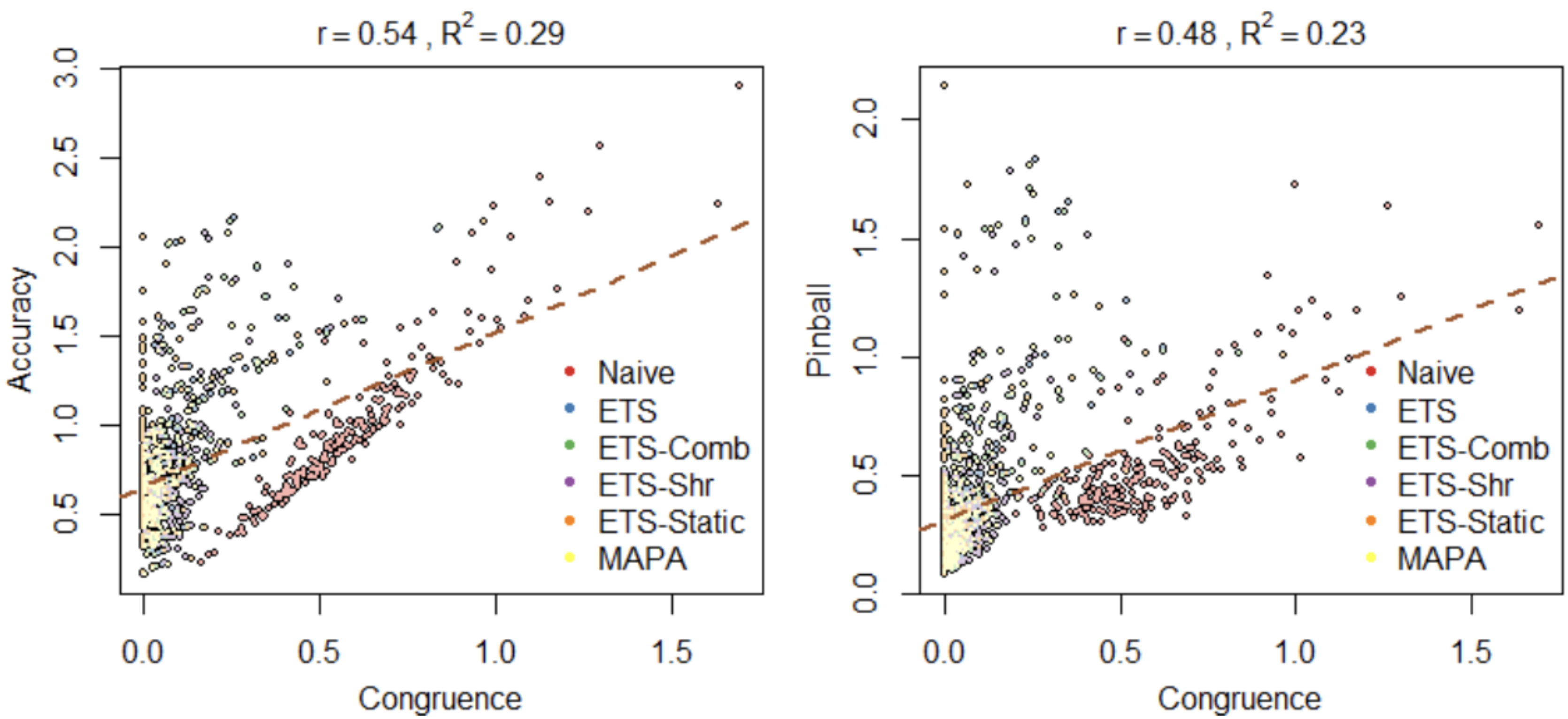
Simulation findings

- We control the DGP by taking the difference of each method with DGP, shown with Δ
  - We know the true congruence as we know the DGP
- Low correlations between Congruence v Accuracy, Congruence v Pinball



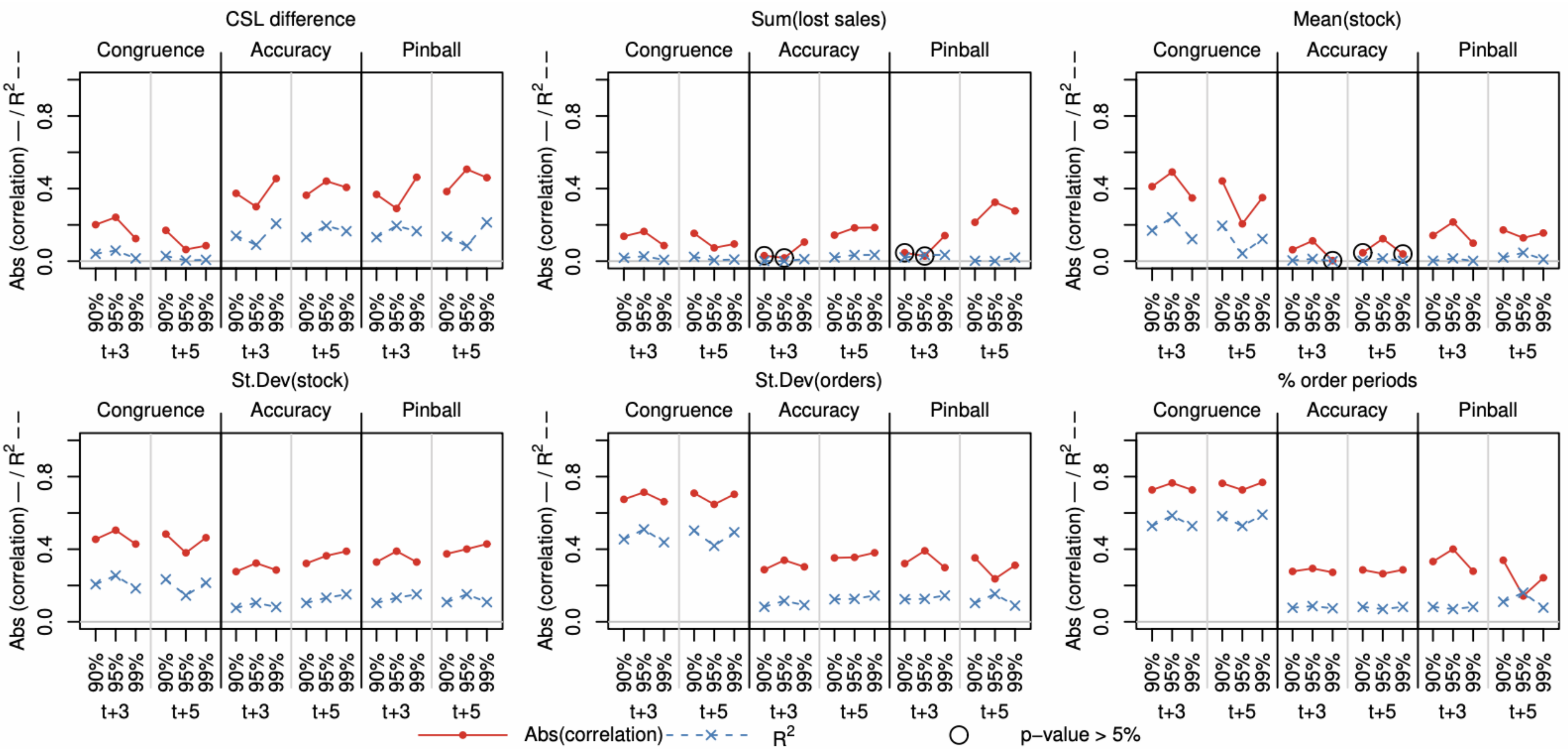
FMCG findings

- FMCG dataset confirms our simulation findings
- Low correlations between Accuracy, Pinball, and Congruence
- Congruence contains a different set of information, given that accuracy and congruence come from the same error matrix



FMCG findings

- Accuracy and Pinball help us achieve desirable CSL
- Congruence affects SD(order) and SD(stock) significantly
- Congruence affects how often orders are placed



Discussion (1)

- Congruence does not replace accuracy, but it **complements**!
- Select forecasts that are congruent to the point where congruence does not harm accuracy
  - We have a set of accurate forecasts, pick the most congruent ones!
  - **Over-congruence** is **not too bad**; **under-congruence** is **bad**!
- There are models and methods that achieve some levels of congruence
  - MAPA → smoothing effects due to temporal aggregation make the forecasts congruent
  - ETS-Shr → shrinking parameters in dynamic models but use it carefully as it tends to produce over-congruent forecasts

Discussion (2)

- Congruence has an implication for managing bull-whip effect; useful for supply chain managers!
- Bull-whip effect is measured

$$\mathrm{BR} = \frac{\mathrm{var}(o_t)}{\mathrm{var}(d_{t-1})},$$

- If we think the number of orders as an estimated conditional value:

$$\mathrm{BR} = \frac{\mathrm{var}(\hat{o}_t|\hat{y}_{t|t-h}, \mathcal{I}_{t-h})}{\mathrm{var}(\hat{y}_t|\mathcal{I}_{t-h})},$$

- Then, congruence comes to play!

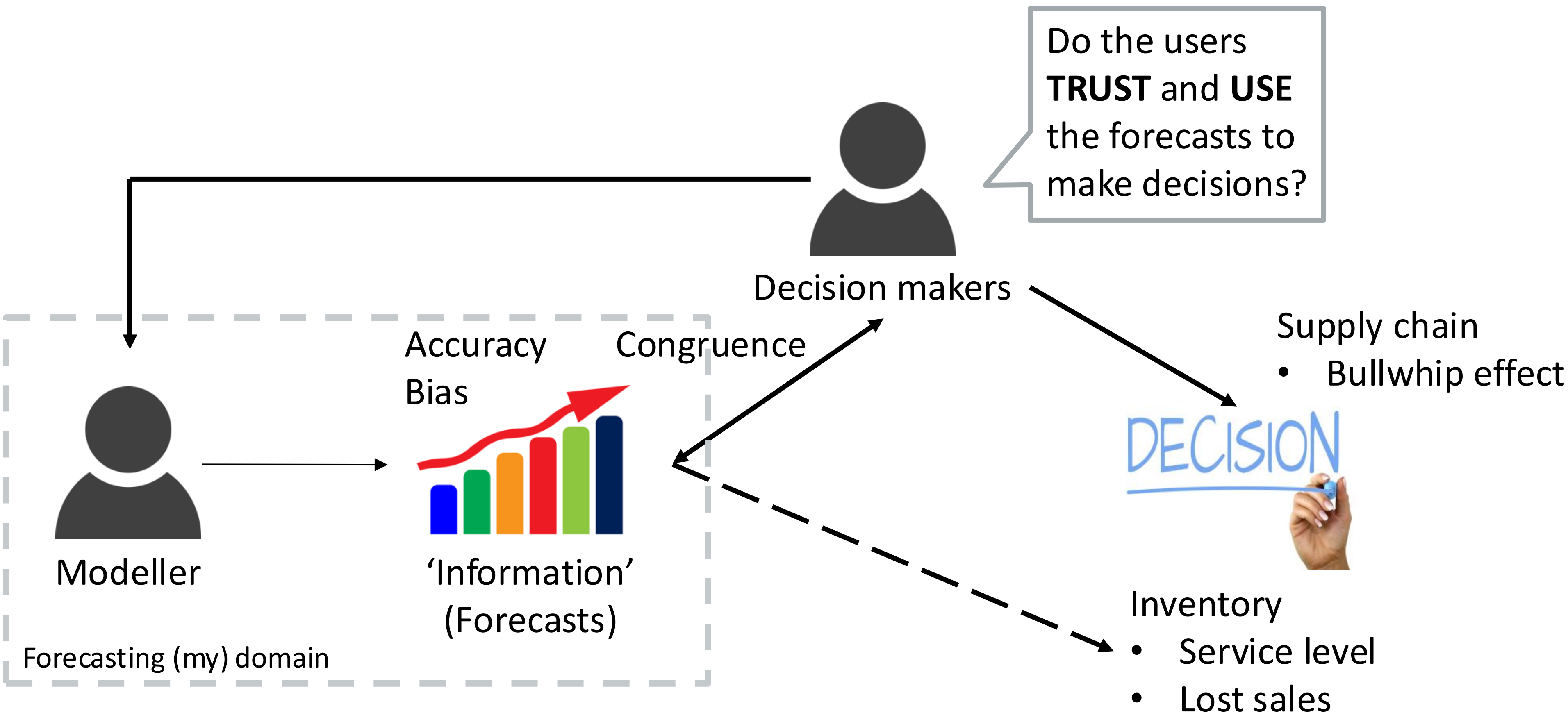
$$\mathrm{var}(\hat{o}_t|\hat{y}_{t|t-h}, \mathcal{I}_{t-h}) \approx \alpha + \beta\tau_{\mathcal{I}_{t-h}}$$

Decision-oriented forecast evaluation

- Accuracy is not the end goal of a forecasting task
  - Follow-up questions: Do users use the forecasts? Do the forecasts result in ‘better’ decision outcomes?
- Looking at forecasting from a broader perspective from
  - Modeler → the current establishment
  - User and decisions → limited exploration
- Challenge the notion of **accurate forecasts**

  accurate forecasts → useful forecasts
- Think of “**usefulness**” as a concept that has dimensions:
  - The first one is **congruence**!
  - The second one is… future research ☺

Decision making processes



Thank you for your attention!

QR code for our paper:

Q&A?!

k.pritularga@lancaster.ac.uk
nikolaos.kourentzes@his.se

Marketing Analytics and Forecasting | Lancaster University Management School

Stability or congruence?

- The literature offers a similar concept called 'forecast instability' but we believe that the terminology is misleading
- Stability is a well-defined concept, at least in statistical forecasting models especially in single source of error state-space framework (Hyndman et al, 2008, p. 36).
- A model is forecastable if

$$\hat{y}_{t|t-1} = a_t + \sum_{j=1}^{t-1} c_j y_{t-j}$$

$$\lim_{t\to\infty} a_t = a \text{ and } \sum_{j=1}^{\infty} |c_j| < \infty$$

$$\text{where } a_t = \boldsymbol{w}'\boldsymbol{D}^{t-1}\boldsymbol{x}_0 \text{ and } c_j = \boldsymbol{w}'\boldsymbol{D}^{j-1}\boldsymbol{g}$$

- A model is stable if (1) $\boldsymbol{D}$ converges to a null matrix as j increases and (2) a_t and c_j decay exponentially
- Stability characterises a model, not a forecast. Thus, we are careful in using the term 'stability' in forecasts.

CMAF