A taxonomy of Multiple Source of Error models for demand forecasting

PTS — Power · Trend · Seasonal

Diego J. Pedregal Juan R. Trapero Ivan Svetunkov

Universidad de Castilla–La Mancha

Lancaster University Management School

ISF 2026 · Montreal, Canada

Marketing Analytics and Forecasting | Lancaster University Management School

Two error formulations

Exponential smoothing was embedded in a state space framework (Snyder, 1985) – but two formulations exist for where the randomness enters.

SSOE – Single Source of Error

- One shared disturbance drives every component. Exponential smoothing (ETS) lives here.
- Free R toolboxes since inception → wide adoption.

MSOE – Multiple Sources of Error

- Each component carries its own disturbance. Structural time series and DLMs live here.
- Historically no open-source tools → confined to econometrics.

The two frameworks are close in theory; MSOE has not been standardised as SSOE.

What we are suggesting

1. **A taxonomy inspired by ETS.** A compact Power–Trend–Seasonal taxonomy that names any model in the family.
2. **A free R/Python toolbox for MSOE models.** The `muse` package and its `pts()` function — available in R and Python – bringing MSOE estimation and forecasting to the same ground as ETS.
3. **A fair MSOE-vs-SSOE comparison – and more.** Both frameworks now run in R/Python, enabling like-for-like benchmarks, plus outlier detection, interpolation, missing-data handling and signal extraction.

From the Kalman filter to structural time series

1960 **Kalman** — recursive filtering for dynamic systems under noise.

1989 **Harvey** – MSOE structural models: level, trend, seasonality, cycle as interpretable components.

1997 **West and Harrison** — Bayesian Dynamic Linear Models for non-stationary series.

1991 **de Jong** – diffuse initialisation.

2012 **Durbin and Koopman** – likelihood by prediction-error decomposition.

Mature theory – smoothing, outlier diagnostics, censored and missing data – but scattered across closed software (e.g. STAMP; Koopman et al., 2009).

Two traditions, and why they diverged

MSOE – born in engineering

- Kalman (1960) – optimal recursive estimation of hidden states; reaches statistics via Harvey (1984, 1989).
- Powerful and interpretable – but *hard to estimate*, so its use stayed confined to econometrics.

SSOE – built for simplicity

- Snyder (1985) recast it with a single source of error – far easier to estimate.
- Ord et al. (1997) added multiplicative forms; Hyndman et al. (2002, 2008) unified it as ETS with IC selection and free R software.

The bridge between them. Muth (1960) and Theil and Wage (1964) showed that additive exponential smoothing has an underlying MSOE model – the same forecasts, reached two ways.

Free software + simplicity made SSOE the standard.

MSOE's strengths stayed underused: the gap PTS closes.

Power-Trend-Seasonal model in the MUSE package

Not this MUSE!

More like this...

A three-letter code, in the spirit of ETS

Power

Box–Cox parameter λ

- $\lambda = 1$ additive
- $\lambda = 0$ multiplicative (logs)
- estimated via modified Guerrero

Trend

- **N** None
- **L** Local (stochastic)
- **G** Global (deterministic)
- **D** Damped

Seasonal

- **N** None
- **D** Discrete (indices)
- **T** Trigonometric (Fourier)

Example. `PTS(0,L,D)`: logs, local trend, discrete seasonality, behaves like ETS(M,M,M).

4 trend $\times$ 3 seasonal $\rightarrow$ **12** base models $\rightarrow$ **24** with $\lambda \in \{0, 1\}$

Connection with ETS

PTS(1,N,N)

Local level · like ETS(A,N,N)

$$y_t = \alpha_t + \varepsilon_t$$
$$\alpha_{t+1} = \alpha_t + \eta_{\alpha,t}$$

A random-walk level plus observation noise — Simple Exponential Smoothing.

PTS(1,L,N)

Local trend · like ETS(A,A,N)

$$y_t = \alpha_t + \varepsilon_t$$
$$\alpha_{t+1} = \alpha_t + \beta_t + \eta_{\alpha,t}$$
$$\beta_{t+1} = \beta_t + \eta_{\beta,t}$$

A second component for the slope – Holt's linear method.

+ Seasonal + Power

$S \in \{$D, T$\}$ · Box–Cox λ

Seasonal adds γ_t – seasonal indices (D) or Fourier harmonics (T).
Power transforms the measurement equation:

$$\frac{y_t^{\lambda}-1}{\lambda} = \alpha_t + \gamma_t + \varepsilon_t$$

$\lambda = 1$ *keeps the model additive;* $\lambda = 0$ *turns it multiplicative.*

One state space form, three recursions

$$y_t = \mathbf{w}'\boldsymbol{\nu}_t + \varepsilon_t$$
$$\boldsymbol{\nu}_{t+1} = \mathbf{F}\boldsymbol{\nu}_t + \mathbf{G}\boldsymbol{\eta}_t$$

Every PTS model is one choice of the system matrices F, G, w.

$+$ support for ARMA innovations in ε_t.

Kalman filter

Optimal one-step estimates, innovations, and the likelihood for estimation.

Fixed-interval smoother

Best component estimates using the entire sample – signal extraction, interpolation.

Disturbance smoother

Standardised residuals per component – diagnostics and outlier detection.

Estimation is maximum likelihood by prediction-error decomposition, with a diffuse start for unknown initial states.

Built on the MSOE literature

Each model in the taxonomy draws on decades of MSOE structural / unobserved-components research. Some resources (not an exhaustive list):

Trends

- Harvey (1984)
- Sbrana and Silvestrini (2020)
- Sbrana and Silvestrini (2025)

Seasonality

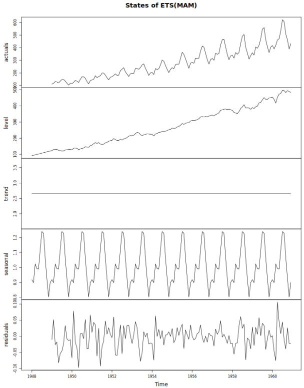
- Harvey (1989)
- West and Harrison (1997)
- Proietti (2000)
- Pedregal and Young (2006)

Box–Cox transform

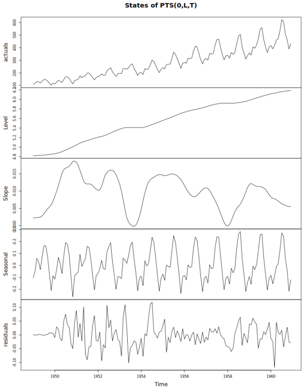
- Box and Cox (1964)
- Proietti and Riani (2009)

Two ways to decompose the same series

ETS(M,A,M)



PTS(0,L,T)



Beyond point forecasts

1. **Diffuse initialisation.** Exact (Durbin and Koopman, 2012) or augmented (de Jong, 1991) starts for unknown states.
2. **Missing data, natively.** Skip the update step; the backward pass interpolates from the full sample.
3. **Outlier detection.** Additive outliers, level shifts and slope changes from auxiliary residuals (Harvey and Koopman, 1992).
4. **Exogenous regressors.** Added to the measurement equation, concentrated out of the likelihood.
5. **Automatic model selection.** AIC / AICc / BIC / BICc across the taxonomy.
6. **Signal extraction.** Smoothed level, trend and seasonal components for interpretation.

Real data evaluation

Real data evaluation

Design

- **325** series, **72** months each (6 yrs).
- **12** rolling origins, $h = 12$ months ahead.
- Expanding window, in-sample 49–60 obs, seasonality 12.
- Every model fitted at each origin; results pooled across all series and origins.

Models & packages

- **pts** – MSOE model
  muse · github.com/config-i1/muse
- **adam** – SSOE benchmark
  smooth · CRAN & PyPI · Svetunkov (2023)
- **ets** – conventional ETS
  forecast · CRAN · Hyndman and Khandakar (2008)

Error measures (lower is better); Δ_t = in-sample first differences, e_{t+j} = forecast errors:

$$\mathrm{RMSSE} = \frac{\mathrm{RMSE}}{\sqrt{\frac{1}{T-1}\sum_{t=2}^{T}\Delta_t^2}} \qquad \mathrm{SAME} = \frac{\mathrm{AME}}{\frac{1}{T-1}\sum_{t=2}^{T}|\Delta_t|}$$

pts leads on point accuracy

Full distribution

	Q1	Med	Q3	Mean
RMSSE				
pts	**0.170**	**0.289**	**0.583**	**0.616**
adam	0.323	0.588	0.906	0.831
ets	0.348	0.612	1.048	0.972
SAME				
pts	**0.070**	**0.182**	**0.451**	**0.685**
adam	0.142	0.353	0.738	0.728
ets	0.169	0.409	0.867	0.914

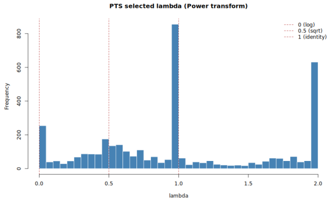
Computational time per series (s)

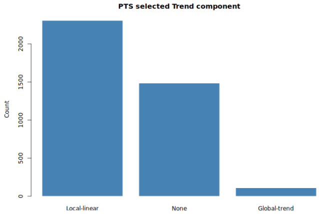
	Q1	Med	Q3	Mean
pts	0.086	0.116	0.131	0.111
adam	0.071	0.087	0.102	0.094
ets	0.401	0.470	0.512	0.426

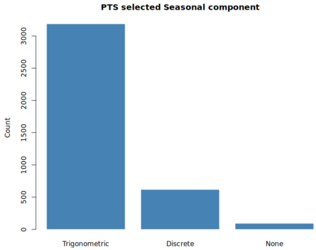
pts $\approx$ adam ($\sim$0.1 s/series); ets is 4–5$\times$ slower.

Which components pts selects – and why it wins

Pooled across all 325 series and 12 origins, pts adapts per series – choosing the power transform, trend and seasonal form that fit each one.

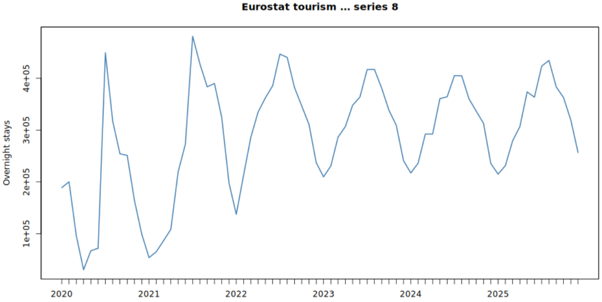






Power: λ clusters at 0 (log), 1 (identity), 2. **Trend**: mostly local-linear or none. **Season**: trigonometric dominates.

How the data looks



A typical Eurostat monthly series: strong seasonality, a sharp COVID collapse in 2020, then a steady recovery.

And here is the real reason!

Full distribution

	Q1	Med	Q3	Mean
RMSSE				
snaive	**0.146**	**0.258**	**0.488**	**0.487**
pts	0.170	0.289	0.583	0.616
adam	0.323	0.588	0.906	0.831
ets	0.348	0.612	1.048	0.972
SAME				
snaive	**0.053**	**0.147**	**0.339**	**0.413**
pts	0.070	0.182	0.451	0.685
adam	0.142	0.353	0.738	0.728
ets	0.169	0.409	0.867	0.914

Computational time per series (s)

	Q1	Med	Q3	Mean
snaive	**0.008**	**0.009**	**0.009**	**0.009**
pts	0.086	0.116	0.131	0.111
adam	0.071	0.087	0.102	0.094
ets	0.401	0.470	0.512	0.426

Conclusions

Conclusions

- MSOE models are powerful and potentially more flexible than SSOE.
- Building on the existing work, we created a unified taxonomy, **PTS**.
- The taxonomy uses an ETS-style vocabulary and relies on automatic selection.
- PTS delivers outlier detection, missing-values handling and signal extraction seamlessly.
- On Eurostat tourism, pts is more accurate than the SSOE benchmarks at comparable speed.

Thank you!

Any questions?

R/Python package: `muse·pts(y, model = "ZZZ")`

Authors: Pedregal · Trapero · Svetunkov

Marketing Analytics and Forecasting | Lancaster University Management School

References I

George E. P. Box and David R. Cox. An analysis of transformations. *Journal of the Royal Statistical Society. Series B*, 26(2):211–252, 1964.

Piet de Jong. The diffuse Kalman filter. *The Annals of Statistics*, 19(2):1073–1083, 1991.

Alysha M. De Livera, Rob J. Hyndman, and Ralph D. Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496):1513–1527, 2011.

James Durbin and Siem Jan Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, 2nd edition, 2012.

Andrew C. Harvey. A unified view of statistical forecasting procedures. *Journal of Forecasting*, 3(3):245–275, 1984.

Andrew C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, 1989.

Andrew C. Harvey and Siem Jan Koopman. Diagnostic checking of unobserved-components time series models. *Journal of Business & Economic Statistics*, 10(4):377–389, 1992.

References II

Charles C. Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1):5–10, 2004. Reprint of the 1957 ONR memorandum.

Rob J. Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3):1–22, 2008.

Rob J. Hyndman, Anne B. Koehler, Ralph D. Snyder, and Simone Grose. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3):439–454, 2002.

Rob J. Hyndman, Anne B. Koehler, J. Keith Ord, and Ralph D. Snyder. *Forecasting with Exponential Smoothing: The State Space Approach*. Springer, 2008.

Rudolph E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

Siem Jan Koopman, Andrew C. Harvey, Jurgen A. Doornik, and Neil Shephard. *STAMP 8.2: Structural Time Series Analyser, Modeller and Predictor*. Timberlake Consultants Press, 2009.

John F. Muth. Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, 55(290):299–306, 1960.

J. Keith Ord, Anne B. Koehler, and Ralph D. Snyder. Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association*, 92(440):1621–1629, 1997.

References III

Diego J. Pedregal and Peter C. Young. Modulated cycles, an approach to modelling periodic components from rapidly sampled data. *International Journal of Forecasting*, 22(1):181–194, 2006.

Tommaso Proietti. Comparing seasonal components for structural time series models. *International Journal of Forecasting*, 16(2):247–260, 2000.

Tommaso Proietti and Marco Riani. Transformations and seasonal adjustment. *Journal of Time Series Analysis*, 30(1):47–69, 2009.

Steven A. Roberts. A general class of Holt-Winters type forecasting models. *Management Science*, 28(7): 808–820, 1982.

Giacomo Sbrana and Andrea Silvestrini. Forecasting with the damped trend model using the structural approach. *International Journal of Production Economics*, 226:107654, 2020.

Giacomo Sbrana and Andrea Silvestrini. The structural theta method and its predictive performance in the M4 competition. *International Journal of Forecasting*, 41(3):940–952, 2025.

Ralph D. Snyder. Recursive estimation of dynamic linear models. *Journal of the Royal Statistical Society. Series B*, 47(2):272–276, 1985.

Ivan Svetunkov. *Forecasting and Analytics with the Augmented Dynamic Adaptive Model (ADAM)*. Chapman & Hall/CRC, 2023. smooth package.

References IV

Henri Theil and Sebald Wage. Some observations on adaptive forecasting. *Management Science*, 10(2): 198–206, 1964.

Mike West and Jeff Harrison. *Bayesian Forecasting and Dynamic Models*. Springer, 2nd edition, 1997.