Demand forecasting model taxonomy for short seasonal time series using cross-sectional information

Huijing Chen

John Boylan, Ivan Svetunkov

ISF 2019

Agenda

- Research motivation
- Theoretical framework
- Empirical analysis
- Simulations
- Conclusions
- Further research

Research motivation ---- Practical needs

- Difficulty in estimating seasonality with short data history
- Multiple products with similar seasonal patterns
- Same/similar products in different locations
- Exploring cross-sectional information
- Empirical evidence: Dalhart (1974), Withycombe (1989), Bunn and Vassilopoulos (1993), Ouwehand et al. (2007), Chen and Boylan (2008)

Research motivation ---- Theoretical development

- Theoretical underpinnings:
  - Chen and Boylan (2007): stationary seasonality
  - Ouwehand *et al*. (2007): multivariate ETS(MAM) model, risk of negative demand (Akram *et al.* 2009)
- Need of a multiplicative demand modelling framework
- Multivariate framework to model inter-relationships
  - de Silva *et al*. (2010): vector innovations structural time series (VISTS) framework
  - VISTS models beneficial but experiments using models containing level and trend components only
  - Current research framework focusing on the seasonal component

Research motivation ---- vector exponential smoothing (VES)

- VES model assuming single source of error (based on VISTS)
- Taxonomy on individual or common:
  - Seasonal component: Ouwehand *et al*. (2007)
  - Smoothing parameters:
    - Level and trend: Fildes *et al*. (1998)
    - Seasonality: Ouwehand *et al*. (2007)
  - Seasonal seed values: no previous studies
- Original contribution: ALL three elements examined together comprehensively

Model framework

Model from Ouwehand *et al.* (2007)

$$Y_{it} = (l_{i,t-1} + b_{i,t-1})S_{t-m}(1 + \varepsilon_{it})$$

$$l_{it} = (l_{i,t-1} + b_{i,t-1})(1 + \alpha_i \varepsilon_{it})$$

$$b_{it} = b_{i,t-1} + \alpha_i \beta_i (l_{i,t-1} + b_{i,t-1})\varepsilon_{it}$$

$$S_t = S_{t-m}\left(1 + \gamma \sum_{i=1}^{N} w_i \, \varepsilon_{it}\right)$$

Our proposed model

$$Y_{it} = l_{i,t-1} b_{i,t-1} S_{i,t-m} \delta_{it}$$

$$l_{it} = l_{i,t-1} b_{i,t-1} \delta_{it}^{\alpha_i}$$

$$b_{it} = b_{i,t-1} \delta_{it}^{\beta_i}$$

$$S_t = S_{t-m} \prod_{i=1}^{N} w_i \delta_{it}^{\gamma}$$

*can be log-transformed into additive format

Cross sectional information on seasonality: seasonal components, seasonal smoothing parameters and seed values

Taxonomy of vector exponential smoothing (VES) models

Most restrictive model

Seasonal components	Smoothing parameters	Seasonal seed values	Abbreviation
Common	Common (Seasonal)	Common	CC_SC
Individual	Common (All)	Common	IC_AC
Individual	Common (All)	Individual	IC_AI
Individual	Common (Seasonal)	Common	IC_SC
Individual	Common (Seasonal)	Individual	IC_SI
Individual	Individual	Common	IIC
Individual	Individual	Individual	III

Benchmark model

$$S_t^* = S_{t-m}^* + \gamma \sum_{i=1}^{N} w_i \delta_{it}^*$$

$$S_{it}^* = S_{i,t-m}^* + \gamma \delta_{it}^*$$

$$S_{it}^* = S_{i,t-m}^* + \gamma_i \delta_{it}^*$$

Common (Seasonal): same seasonal smoothing parameter across series

Common (All): all three smoothing parameter values are the same across series

Weights in the CC_SC model

$$S_t^* = S_{t-m}^* + \gamma \sum_{i=1}^{N} w_i \delta_{it}^*$$

- Equal weights: adopted in current study, variance equal across series

$$S_t^* = S_{t-m}^* + \frac{1}{N} \gamma \sum_{i=1}^{N} \delta_{it}^*$$

- Unequal weights:
  - Reasonable to assign low weights to noisy series
  - Requires further research

$w_i = \frac{\sigma_i^{-1}}{\sum_{j=1}^{N} \sigma_j^{-1}}$ or $w_i = \frac{\sigma_i^{-2}}{\sum_{j=1}^{N} \sigma_j^{-2}}$

Empirical analysis

- 218 series from 7 product families
- Monthly demand data for different types of light bulbs
- 5 years history: 1 year hold out; 2, 3 or 4 years estimation

	No. of SKUs
Group 1	66
Group 2	27
Group 3	61
Group 4	45
Group 5	2
Group 6	12
Group 7	5
Total	**218**

Estimation

- All models implemented in smooth package for R (Svetunkov, 2019)
  - Version 2.4.8
  - ves() function implements all main VES models
  - gsi() function implements CC_SC
- Models estimated via minimisation of trace of covariance matrix $\boldsymbol{\Sigma}$
  - Full likelihood can only be applied to small groups
  - de Silva et al. (2010) reported no deterioration in forecasting performance

Error measures

Mean Absolute Error (MAE): $\frac{\sum_{j=1}^{h} |e_{t+j}|}{h}$

h: forecasting horizon
t: forecast origin
j=1 to 12

Measure overall accuracy

Mean Squared Error (MSE): $\frac{\sum_{j=1}^{h} e_{t+j}^2}{h}$

All measures are relative to the benchmark model (III)

Absolute Mean Error (AME): $\left|\frac{\sum_{j=1}^{h} e_{t+j}}{h}\right|$

Measures bias

Empirical findings

	4 years			3 years			2 years		
	RelMAE	RelMSE	RelAME	RelMAE	RelMSE	RelAME	RelMAE	RelMSE	RelAME
CC_SC	0.72	0.49	0.67	0.88	0.76	0.74	0.78	0.62	0.78
IC_AC	**0.56**	**0.31**	**0.43**	**0.63**	**0.41**	**0.57**	**0.65**	**0.43**	**0.56**
IC_SC	0.69	0.46	0.67	0.74	0.54	0.79	0.70	0.48	0.69
IIC	0.80	0.60	0.87	0.85	0.69	0.87	0.82	0.66	0.90
IC_AI	0.80	0.66	0.74	0.89	0.80	0.74	0.94	0.87	0.76
IC_SI	0.97	0.95	1.04	1.02	1.03	1.00	0.91	0.84	0.97
III	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

All VES models perform better than III (benchmark)

IC_AC is the best performing VES model

Improvement on III greater as length of history increases

Seed values: common seeds beneficial

Smoothing parameters: $*C_S*/*C_A*$ beneficial, more so IF seed values are common

Seasonal components: common beneficial but may be too restrictive

Simulations: Data generation

- Level, trend and seasonality specified in additive form (sim.ves() function), then taking exponent
- **Initial level**: random ∈ $[4, 10]$. This reflects a wide range of mean demand values
- **Initial trend**: 0
- **Initial seasonal seeds**: randomly generated between -1 and 1, then normalised
- **Noise**: δ_{it} ~ LogN (0,σ_i^2), σ is set to be 0.09 and equal
- **Forecasting horizon**: 1 to 12 months ahead
- **Number of series**: 2 or 100
- **Length of history**: 36 or 60 observations
- **Replication**: 1000

Simulations: Smoothing parameters

Group 1: DGP--ICAI & ICAC

- Common low: [0.1, 0.05, 0.1]
- Common high: [0.4, 0.3, 0.4]

Group 2: DGP-- ICSI, ICSC

- Common γ: 0.1 or 0.4
- α_i, β_i: random ∈ $[0.1, 0.4]$
- $\beta_i < \alpha_i$
- $\gamma < 1 - \alpha_i$
- CCSC: not included in the data generation yet

Group 3: DGP-- III & IIC

- All smoothing parameters random ∈ $[0.1, 0.4]$

Using most commonly applied values rather than the whole parameter space

Number of parameters to estimate

2 series

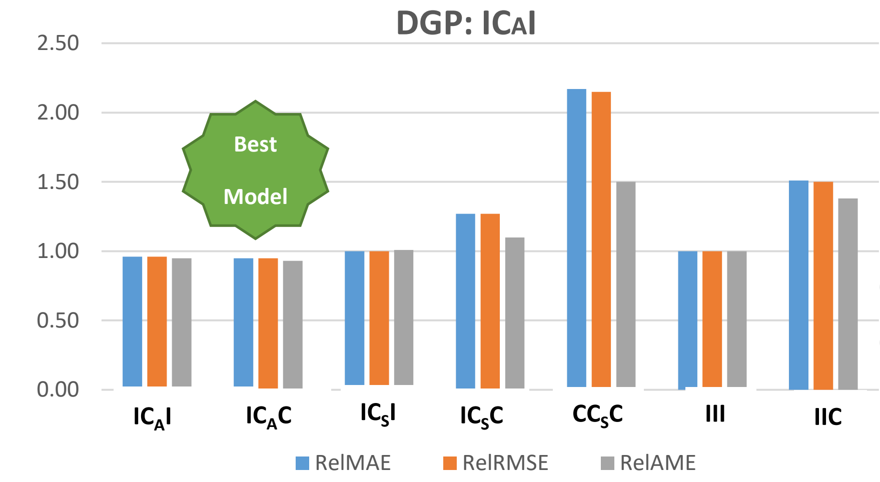
	Parameters per series		Parameters overall	
Model	likelihood	trace	likelihood	trace
III	19.0	18.0	37.0	36.0
IIC	13.0	12.0	25.0	24.0
ICSI	18.5	17.5	36.0	35.0
ICSC	12.5	11.5	24.0	23.0
CCSC	12.5	11.5	24.0	23.0
ICAI	17.5	16.5	34.0	33.0
ICAC	11.5	10.5	22.0	21.0

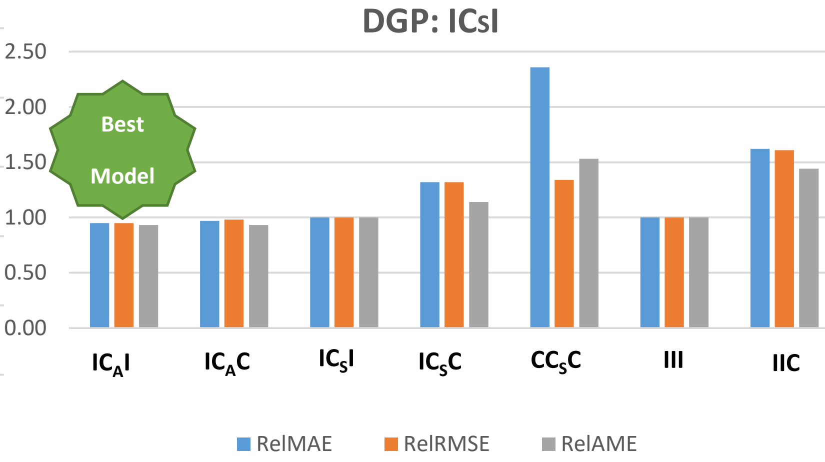
100 series

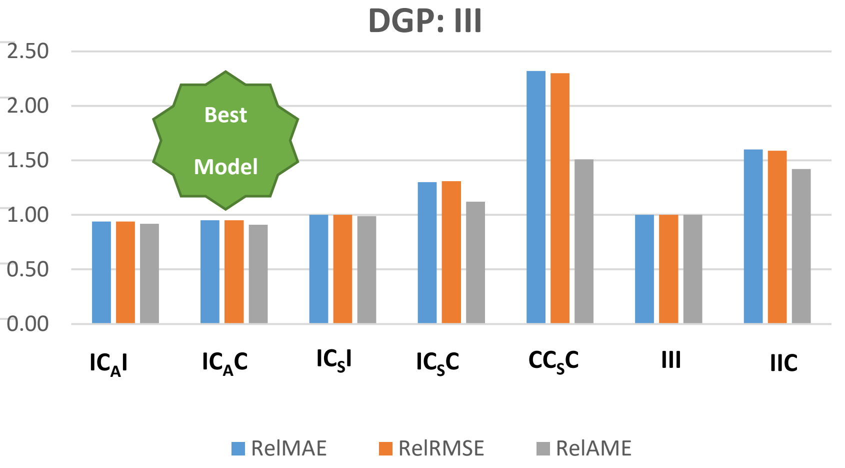
	Parameters per series		Parameters overall	
Model	likelihood	trace	likelihood	trace
III	117.0	18.0	6750.0	1800.0
IIC	105.1	6.1	5562.0	612.0
ICSI	116.0	17.0	6651.0	1701.0
ICSC	104.1	5.1	5463.0	513.0
CCSC	104.1	5.1	5463.0	513.0
ICAI	114.0	15.0	6453.0	1503.0
ICAC	102.2	3.2	5265.0	315.0

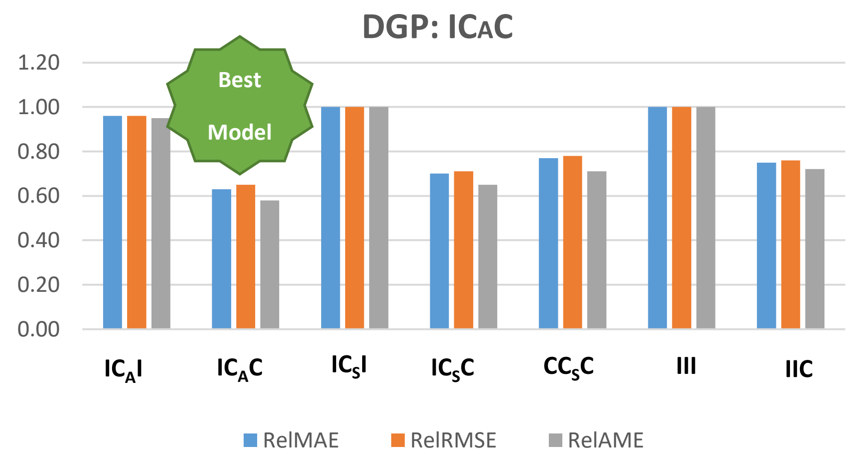
Number of parameters to estimate per series is larger than the sample size (max 48 obs)

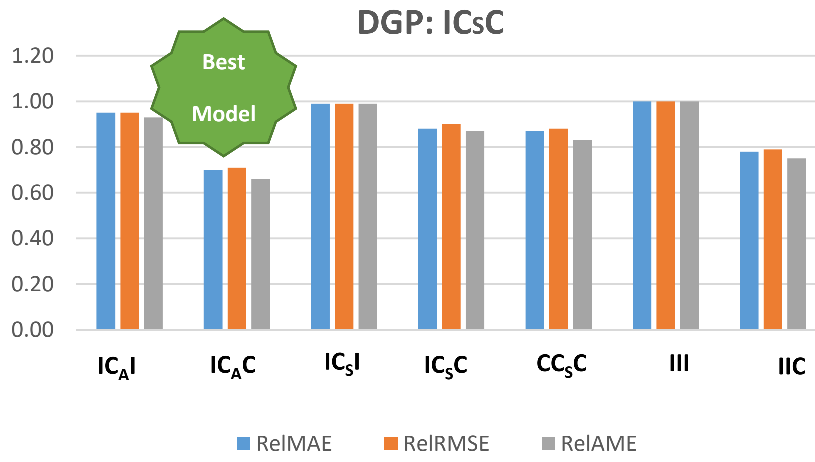
Simulations findings: point forecasts (1)

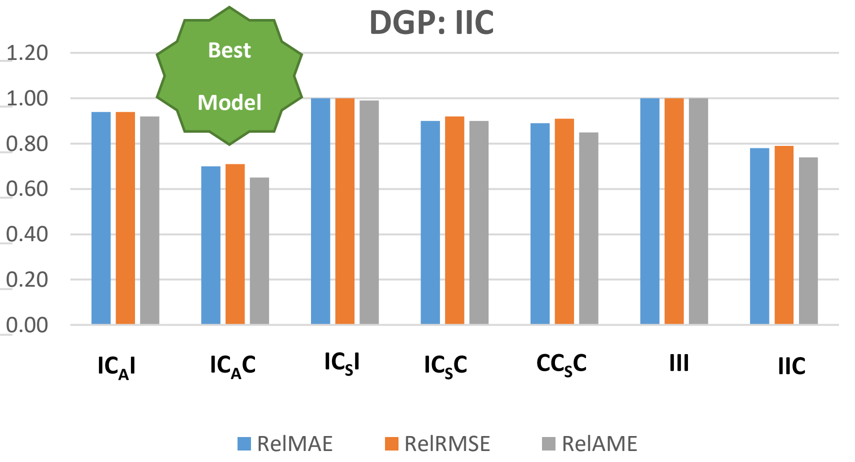












Simulations findings: point forecasts (2)

- 2 series

	RelMAE	RelRMSE	RelAME
IC_AI	0.95	0.95	0.94
IC_AC	1.02	1.02	1.03

	RelMAE	RelRMSE	RelAME
IC_AI	0.95	0.95	0.95
IC_AC	0.71	0.72	0.68

Red indicates data generation process

Parameters to estimate:
IC_AI : 33
IC_AC: 21
III: 36

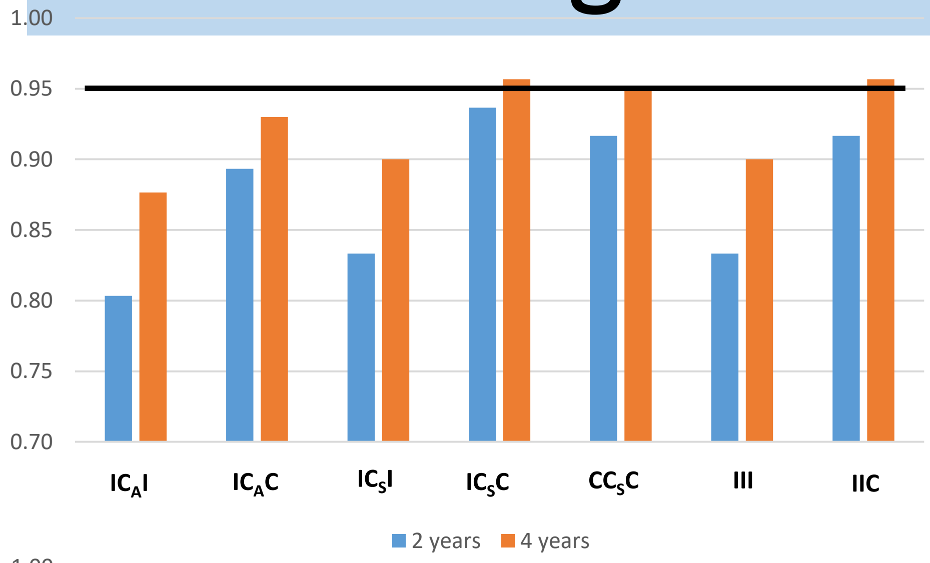
- 100 series

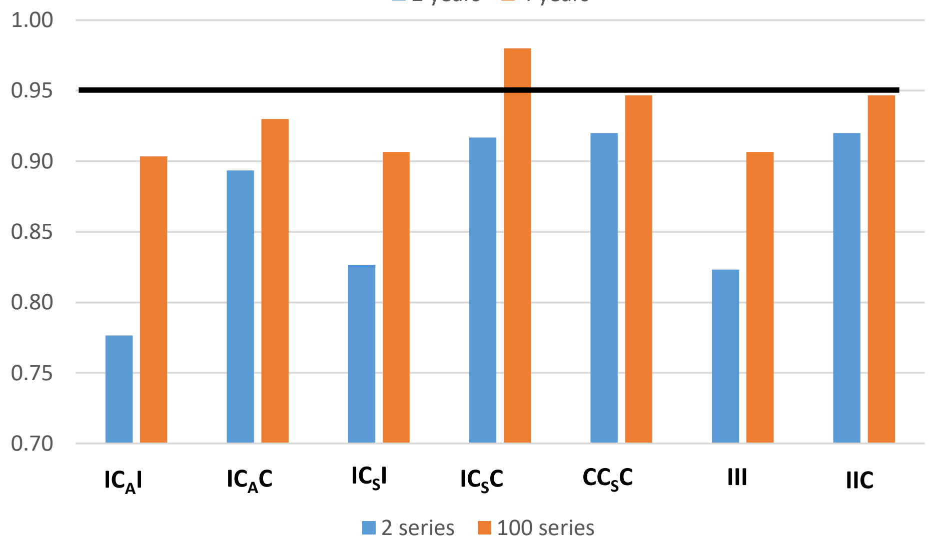
	RelMAE	RelRMSE	RelAME
IC_AI	0.97	0.97	0.96
IC_AC	0.88	0.89	0.83

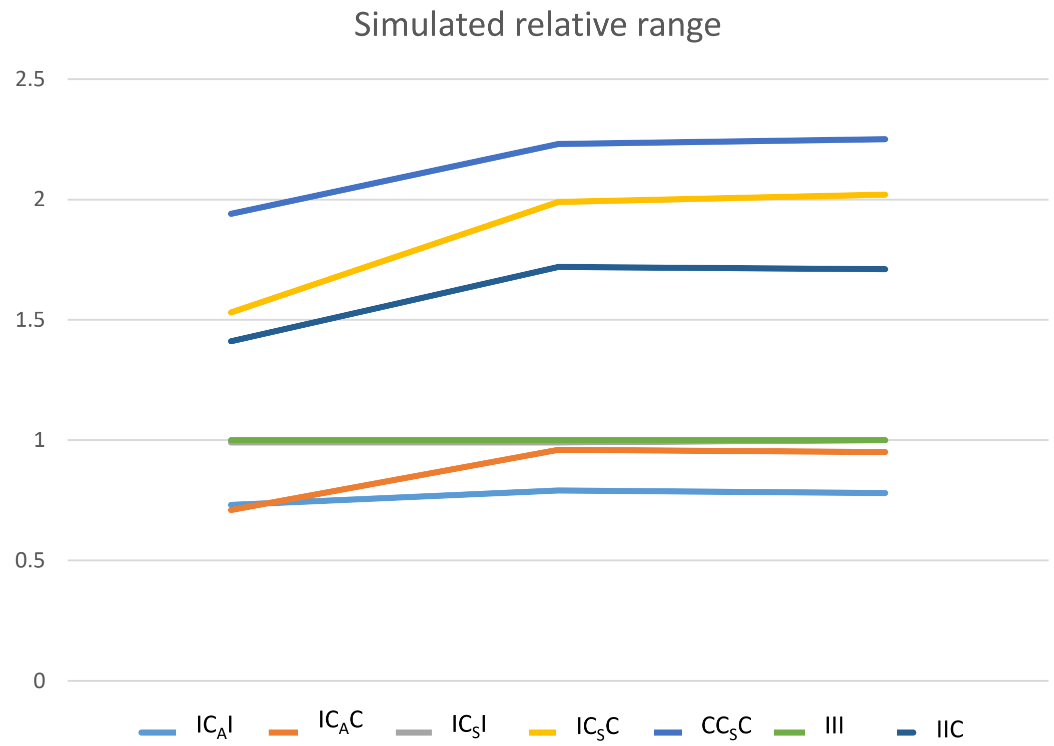
	RelMAE	RelRMSE	RelAME
IC_AI	0.97	0.97	0.95
IC_AC	0.57	0.58	0.50

Parameters to estimate:
IC_AI : 1503
IC_AC: 315
III: 1800

Simulations findings: 95% interval coverage & relative range







Conclusions

- VES models perform well in simulations and empirical analysis
- Parsimonious and flexible models perform strongly
- $*C_A*$ models provide the lowest forecasting errors
  - perform consistently in different data generation processes
  - IC_AC Performance against benchmark improves when number of series available increases
- Coverage:
  - Improves when more observations available
  - Improves when more series available
  - **C models closer to 95% coverage than **I models
- Relative range: $*C_A*$ models have narrower range than the III benchmark

Future research

- Simulate data generation process of CC_SC
- Allow optimised weights in CC_SC
- Explore benefits (if any) of full likelihood
- Consider shrinkage estimation for VES models
- Assess VES models with new empirical data
- Model selection within the taxonomy of VES models
- VES models extendable to multiple seasonal patterns

UNIVERSITY OF PORTSMOUTH

Marketing Analytics & Forecasting

Lancaster University Management School

Thank you for your attention!

References

- Akram, M., Hyndman, R. J., Ord, J. K., Stat, A. N. Z. J., Kram, M. U. A., Yndman, R. O. B. J. H., & Rd, J. K. E. O. (2009) Exponential smoothing and non-negative data. *Australian and New Zealand Journal of Statistics*, *51*(4), 415–432. https://doi.org/10.1111/j.1467-842X.2009.00555.x
- Bunn, D.W. and Vassilopoulos A.I. (1993) Using group seasonal indices in multi-item short-term forecasting. *International Journal of Forecasting, 9*(4), 517-526. https://doi.org/10.1016/0169-2070(93)90078-2
- Chen, H. and Boylan, J.E. (2007) Use of individual and group seasonal indices in subaggregate demand forecasting. *Journal of the Operational Research Society, 58*(12), 1660-1671. https://www.jstor.org/stable/4622863
- Chen, H. and Boylan, J.E. (2008) Empirical evidence on individual, group and shrinkage seasonal indices. *International Journal of Forecasting, 24*(3), 525-534. https://doi.org/10.1016/j.ijforecast.2008.02.005
- Dalhart, G. (1974) Class seasonality – a new approach. *American Production and Inventory Control Society 1974 Proceedings*. Reprinted in *Forecasting*, 2nd Ed., American Production and Inventory Control Society, Washington DC, 11-16.
- de Silva, A., Hyndman, R. J., & Snyder, R. (2010). The vector innovations structural time series framework. *Statistical Modelling: An International Journal*, *10*(4), 353–374. https://doi.org/10.1177/1471082X0901000401
- Fildes, R., Hibon, M., Makridakis, S. & Meade, N. (1998) Generalising about univariate forecasting methods: further empirical evidence. *International Journal of Forecasting, 14*(3), 339-358. https://doi.org/10.1016/S0169-2070(98)00009-0
- Ouwehand, P., Hyndman, R.J., de Kok, T.G., & van Donselaar, K.H. (2007) A state space model for exponential smoothing with group seasonality. Department of Econometrics & Business Statistics Working Paper Series 07/07, Monash University. https://www.monash.edu/business/ebs/research/publications/ebs/wp07-07.pdf
- Svetunkov I. (2019). smooth: Forecasting Using State Space Models. R package version 2.4.8 https://github.com/config-i1/smooth
- Withycombe, R. (1989) Forecasting with combined seasonal indices. *International Journal of Forecasting, 5*(4), 547-552. https://doi.org/10.1016/0169-2070(89)90010-1