**Institution:** Lancaster University

**Unit of Assessment:** 17, Business and Management

**Title of case study:** Improving forecasting in organisations with statistical and machine learning methods

**Period when the underpinning research was undertaken:** 2000 - 2020

**Details of staff conducting the underpinning research from the submitting unit:**

| Name(s): | Role(s) (e.g. job title): | Period(s) employed by submitting HEI: |
|---|---|---|
| Robert Fildes | Professor | 1995 - present |
| John Boylan | Professor | 2015 - present |
| Sven Crone | Lecturer | 2004 - present |
| Nikolaos Kourentzes | Professor | 2009 - present |
| Ivan Svetunkov | Lecturer | 2016 - present |
| Fotios Petropoulos | Research Associate | 2012 - 2014 |
| Devon Barrow | Research Associate | 2013 - 2014 |

**Period when the claimed impact occurred:** 2013 - 2020

**Is this case study continued from a case study submitted in 2014?** N

## 1. Summary of the impact

Lancaster University Management School's (LUMS) Centre for Marketing Analytics and Forecasting (CMAF) works with major international companies and public sector organisations to develop innovative predictive analytics techniques to improve forecast accuracy and reduce operational costs and waste. The CMAF's holistic approach includes developing new statistical methods/algorithms and software and training experts. The impact of this work includes the following:

1) Achieved forecast accuracy and efficiency improvements averaging 10% with companies such as Amazon, Beiersdorf; Bekaert, [text removed for publication], Demand Works, Johnson & Johnson (J&J), Lego, Volvo, leading to economic impacts in excess of GBP100 million since 2013.

2) These innovations have informed the design of nine open source software packages that facilitate organisational efficiency and sustainability on a global scale. These packages have been downloaded >1.2 million times and integrated into commercial software programs (including Solventure LIFe and Demand Works).

3) Improving forecasting processes requires changes in management practices beyond statistical and technological innovations. In response, the CMAF has designed and delivered executive training courses and corporate mentoring programmes involving >40 major organisations. Since 2013, the management practices of approximately 600 demand planners and forecasters have been impacted by the CMAF's bespoke user engagement activities.

## 2. Underpinning research

The last six years have seen rapid growth in data availability and computing power, offering organisations the potential to make major analytical improvements in operational efficiency in such aspects as service quality and inventory. Research from the CMAF combines statistical innovations with an organisational perspective, resulting in an interdisciplinary approach to these modern challenges. This unique approach has been mainly applied in the following areas.

### Improved statistical forecasting methods

The CMAF has a long tradition of researching new statistical forecasting methods. In recent years, it has led in the development of multiple new methods that have already been adopted into practice. Established methods have focused on analysing data collected for a particular variable at a given frequency. With an award-winning paper [R1], Kourentzes and colleagues have been at the forefront of researching temporal hierarchies, connecting aggregate and disaggregate data models, and enabling a holistic consideration of the

forecasting problem. Overcoming existing limitations in the field in data prolixity, the resulting models draw on both micro- and macro-views, incorporating substantially more information into the end forecast than conventional approaches, which has led to the following gains: (i) aligned decision making across planning horizons and organisational departments or divisions and (ii) consistently more reliable and accurate forecasts, facilitating the automation of massive forecasting processes that are now commonplace in industry (e.g. large retailers, such as [text removed for publication], require millions of daily forecasts).

From 2014, Kourentzes and colleagues worked with industrial partners and developed a methodology to automatically identify useful macro-economic leading indicators from a large pool of possibilities to augment strategic/tactical forecasts, with gains of 15% [R2]. These are then disaggregated into detailed operational forecasts through the use of hierarchical forecasting. These two streams of research led to the first statistically sound generation of cross-temporally coherent forecasts, which result in the sought-after 'one number forecast', which provides a common view of the future across functions within organisations.

Fildes, Kourentzes and colleagues have contributed substantially to model selection, combination, and pooling [R2, R3] using meta-learning approaches. These methods aim to improve predictions and increase the reliability of forecasting processes in organisations by mitigating the risk of using inappropriate forecasts. Boylan, Kourentzes and Svetunkov have also contributed substantially with algorithmic innovations in intermittent demand forecasting, which is relevant to slow-moving items (such as retailing and spare parts) but also of growing importance due to the increased frequency of decision making that results in intermittent observations [R4]. Crone, Kourentzes and colleagues have developed methodologies for the automatic building of neural networks and other machine learning methods for business forecasting in the area of artificial intelligence (AI) and for tackling high-frequency forecasting problems [R5] in areas such as call centre forecasting and demonstrated the superiority of AI over conventional statistical methods.

**Improved utilisation of expert judgement and better forecasting support systems (FSSs)**

Following a three-year (2004-2007) EPSRC-funded project into how forecasters use their expert judgement to respond to special effects beyond the data available to a statistical model, Fildes and colleagues (including University of Bath) established a strong influential research agenda through an ambitious data collection, analysis and modelling of forecasts from retailers and suppliers to show that substantial deficiencies in judgement can be rectified in part by improved FSS design to improve accuracy [R6]. These results show how changes in design involving better integration of statistical methods with experts can impact how judgements are made, statistical forecasts and the final accuracy [R6].

**Forecasting in the retailing and supply chain contexts**

Forecasting the impact of promotional events has been identified in survey evidence as particularly problematic for organisations. Fildes and colleagues have conducted research into data-intensive approaches, utilising estimation methods capable of handling the large number of variables and the vast quantities of retail data at the product-store level.

**3. References to the research**

[R1] **Kourentzes, N., Petropoulos, F.,** & Trapero, J. R. (2014). Improving forecasting by estimating time series structural components across multiple frequencies. *International Journal of Forecasting*, *30*(2), 291-302. https://doi.org/10.1016/j.ijforecast.2013.09.006. (125 citations on Google Scholar, Altmetric score: 6)

[R2] Sagaert, Y. R., Aghezzaf, E. H., **Kourentzes, N.**, & Desmet, B. (2018). Temporal big data for tactical sales forecasting in the tire industry. *Interfaces*, *48*(2), 121-129. https://doi.org/10.1287/inte.2017.0901. (15 citations on Google Scholar)

[R3] **Fildes, R., & Petropoulos, F.** (2015). Simple versus complex selection rules for forecasting many time series. *Journal of Business Research*, *68*(8), 1692-1701. https://doi.org/10.1016/j.jbusres.2015.03.028 (58 citations on Google Scholar)

[R4] **Kourentzes, N.** (2014). On intermittent demand model optimisation and selection. *International Journal of Production Economics*, *156*, 180-190. https://doi.org/10.1016/j.ijpe.2014.06.007. (63 citations on Google Scholar)

[R5] **Kourentzes, N., Barrow, D. K., & Crone, S. F.** (2014). Neural network ensemble operators for time series forecasting. *Expert Systems with Applications*, *41*(9), 4235-4244. https://doi.org/10.1016/j.eswa.2013.12.011. (196 citations on Google Scholar)

[R6] **Fildes, R.**, Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, *25*(1), 3-23. https://doi.org/10.1016/j.ijforecast.2008.11.010. (403 citations on Google Scholar, Altmetric score: 38)

[R7] **Svetunkov, I., & Boylan, J. E.** (2020). State-space ARIMA for supply-chain forecasting. *International Journal of Production Research*, *58*(3), 818-827. https://doi.org/10.1080/00207543.2019.1600764

## 4. Details of the impact

The Centre for Marketing Analytics and Forecasting's (CMAF) improved forecasting algorithms, model selection method, and integration of judgement and systems have made significant contributions to a wide range of organisations through improved forecast accuracy, substantial cost/service-level improvements over the supply chain, and positive indirect impacts on consumer service and society. Impact has been achieved in the following ways.

### Forecasting improvement projects with organisations

The CMAF has implemented research increasing forecast accuracy in various industrial projects, which in turn have increased service levels and reduced operational costs and waste. Since January 2014, the CMAF has worked with >40 major companies in the UK and internationally, including Amazon, Beiersdorf, Bekaert, H&M, Heineken, Jaguar Land Rover, J&J, Lego, Morrisons and [text removed for publication], on dedicated projects implementing CMAF research [S1]. Examples of the contributions provided by these projects include:

- Beiersdorf: In a presentation at a CMAF Workshop (2017) supported by Crone, process improvements of 5% arising equally from novel statistical model selection and judgemental adjustment were outlined. With a stock level of EUR616 million, this improvement freed up an estimated EUR30 million in 2017 alone [S2].
- Bekaert: Implementing CMAF innovations has led to an accuracy improvement of 15% with a respective reduction in working capital, estimated at around EUR80 million [S3]. This impact is attested to by their Director of Global Marketing and Business Development in a 2018 verification letter, "*The model provided so far on average over 15% improvement of accuracy, as measured in mean absolute percent of forecast error. This benefits our supply chain management in their tactical decision making. We strongly believe that the benefits obtained by this tool have the potential in helping us in reducing the overall inventory in our global supply chain"* [S3].

Other notable users of CMAF research include Google's Technical Infrastructure Division, with the Resource Efficiency Data Science team lead stating in December 2020 that the work of the centre, in particular that of Kourentzes and Fildes, has been something he has learned a lot from over the several years of collaboration [S4].

Furthermore, CMAF research has led to approximately 60 improvement projects set up as part of the master's programme in operational research (OR, now Business Analytics) at LUMS.

### Software and package design that implements CMAF research innovations

The CMAF has published nine open source R packages [S5] that implement the research innovations in temporal hierarchies and aggregation, automatic model building, neural

networks, intermittent demand, diffusion and life cycle modelling, and modelling online search behaviours. These packages have been downloaded >1.2 million times between 2014 and 2020 [S5], making CMAF a worldwide leader in R package development. These packages can save years of dedicated development time and allow practitioners to quickly run state-of-the-art analyses, limiting development costs and overcoming implementation barriers. Numerous global organisations have made use of the R packages developed by CMAF. The Data and Analytics Leader and Director in [text removed for publication] said, "*The CMAF's R packages have been instrumental for our predictive analytics capabilities and consulting activities, providing the only implementations to address a number of business analytics problems, such as intermittent demand forecasting*" [S6].

Many established software houses now facilitate the use of CMAF open source packages, further extending the research reach. Proprietary code developed by the CMAF is included in forecasting software commercialised by Demand Works, leading to annual inventory savings estimated at USD20 million, totalling USD60 million since 2017 [S7]. Complementary research on the forecasting process and the role of judgement has influenced large software firms, such as SAS, where the work of CMAF has "*proved important in analysis of the fundamentals of this value-adding process*" [S8]. In turn, ideas generated by the CMAF have led to SAS's Forecast Value Added module, which aims to disentangle the separate effects of statistical modelling and expert managerial judgement [S8].

Commercial software implementing methods developed by research primarily carried out in the CMAF includes:

- Iqast, developed by Crone, offers AI forecasting solutions to industry and is used by major companies such as Beiersdorf, Johnson & Johnson, Hapag-Lloyd, Sanofi Aventis, Bayer, Virgin Atlantic Cargo, and Anheuser-Busch InBev.
- Solventure LIFe models the influence of leading indicators on sales, improving forecasts. This solution implements research methods developed by the CMAF [R2] and is used in companies such as Bekaert, Volvo and Solvay [S9].
- Demand Works implemented an automatic state-space method for seasonal autoregressive integrated moving average (ARIMA) models [R7] for supply chain applications [S7].

**Design and delivery of workshops, executive training courses, and corporate mentoring, diffusing current CMAF research into practice**

Improving the forecasting process requires changes in management practices beyond any statistical innovations. Therefore, CMAF run biannual executive training courses and practitioner workshops. The courses have been both open (involving 40 organisations) and tailored to individual organisations [S1]. At Beiersdorf, the CMAF has trained approx. 165 Demand Planners between 2013 and 2020 and is recognised to have contributed substantially to increasing forecasting accuracy [S3]. Other participating companies include Arco, British Gypsum, Colruyt, [text removed for publication], Howdens, McBride, [text removed for publication], BSkyB, and Wilko [S1, S6, S11].

The biannual workshops cover CMAF's core research and regularly attract between 50 and 100 participants with uptakes into projects from for example, Lego and Wilko [S1]. A 2018 webinar on forecasting software, sponsored by the OR Society, attracted 60 attendees, the highest number achieved at these events: the 2019 webinar had >100 registered participants.

Members of the CMAF have presented their research to sector leaders, such as [text removed for publication], Amazon and Google [S1, S2], and are often speakers in practitioner workshops (for example, SAS analytics academies [S8]), influencing the future development of practice through their widely used software.

The Head of Data Science & AI at Global Logistics, [text removed for publication], said of the impact of these user engagement activities, "*The work by Kourentzes and colleagues on*

*hierarchical forecasting and the forecasting training by CMAF have been influential in the rethinking of our global forecasting practices and is now a blueprint for our ongoing implementation*" [S10].

**5. Sources to corroborate the impact**

[S1] List of executive training workshops in the impact period.

[S2] Presentation and public material connected with Beiersdorf, 2017.

[S3] Published paper, Bekaert (see [R2]) and annual reports, 2014, 2015, 2016.

[S4] Email from lead of Google's Resource Efficiency Data Science team, 2020.

[S5] R packages, CRAN evidence. See R script.

[S6] Letter from the Data & Analytics Head and Director in [text removed for publication] and public marketing material, 2020.

[S7] Letter from the President of Demand Works, 2020.

[S8] Letter from the Product Marketing Manager of SAS, 2020.

[S9] Public material, Solventure, available (with registration)

[S10] Letter from the Head of Data Science & AI, Global Logistics, [text removed for publication], 2020.