DYNAMIC COVARIANCE MODELS FOR WIND POWER AND NET-DEMAND FORECASTING

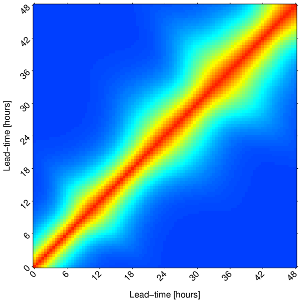
Wind Energy Science Conference

Jethro Browell, Vincenzo Gioia, Matteo Fasiolo, Ciaran Gilbert

24 May 2023

School of Mathematics and Statistics
University of Glasgow
jethro.browell@glasgow.ac.uk

CONTENTS



This work was supported by UKRI via EPSRC Innovation Fellowship *System-wide probabilistic energy forecasting* (EP/R023484/2).

MOTIVATION

Decision-makers (TSOs, DSOs, traders) require forecasts of multiple quantities to operate efficiently and manage risk:

- How much demand will be met by wind and solar power tomorrow?
- What is the chance of power flows exceeding network capacity?
- Will I get a better price if I sell my power day-ahead or intraday?

Probabilistic forecasts quantify uncertainty by expressing predictions as probability density functions.

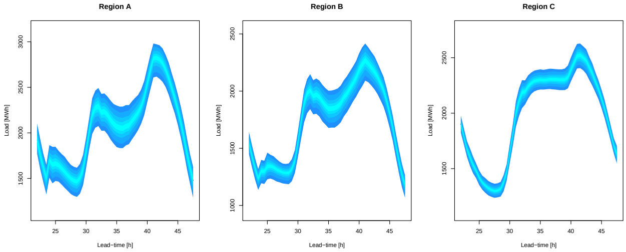


Figure 1: Fan plots of density forecasts for three locations and 48 time periods

⚠ These forecasts do not describe spatial or temporal dependency.

Probabilistic forecasts quantify uncertainty by expressing predictions as probability density functions.

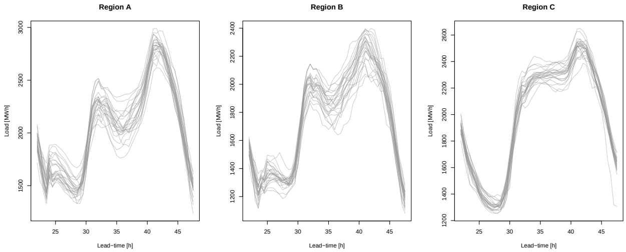


Figure 2: Space-time trajectories (or scenarios/samples) drawn from multivariate probabilistic forecast. These forecasts contain dependency information but are difficult to visualise.

This quickly becomes a high-dimensional problem!

Gaussian copulas provide a suitable framework for describing such high-dimensional predictive distributions:

- Margins of the copula are the familiar density forecasts
- Dependency structure specified by a covariance matrix, Σ
- Scales well (compared to other copulas) but limited by estimation of the covariance matrix

The remainder of this talk is concerned with this covariance matrix and the possibility that it:

1. has a complex structure, and/or
2. varies over time, perhaps as a function of some covariate.

COVARIANCE FUNCTIONS AND MATRICES

Consider a random process $Z_t(\boldsymbol{s}, l)$ at location $\boldsymbol{s}$, forecast lead-time l, and forecast issue time t.

A covariance function, C_t, is *stationary* if the covariance

$$\mathrm{cov}(Z_t(\boldsymbol{s}, l), Z_t(\boldsymbol{s} + \boldsymbol{h}, l + u)) = C_t(\boldsymbol{h}, u) \tag{1}$$

depends only on separation $(\boldsymbol{h}, u)$. Furthermore, C_t is *isotropic* if it is invariant to the direction of $\boldsymbol{h}$ and u

$$\mathrm{cov}(Z_t(\boldsymbol{s}, l), Z_t(\boldsymbol{s} + \boldsymbol{h}, l + u)) = C_t(\|\boldsymbol{h}\|, |u|) \quad . \tag{2}$$

Table 1: Some parametric classes of isotropic covariance functions where $C(\boldsymbol{h})$ takes the form $C(\|\boldsymbol{h}\|;\boldsymbol{\xi})$. The Whittle–Matérn covariance is defined in terms of the modified Bessel function of the second kind K_ν.

Class	Function $C(r;\boldsymbol{\xi})$	Parameters $\boldsymbol{\xi}$
Powered Exponential	$\sigma^2 e^{-(\theta r)^\gamma}$	$0 < \gamma \leq 2;\ \theta > 0;\ \sigma \geq 0$
Whittle–Matérn	$\sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)}(\theta r)K_\nu(\theta r)$	$\nu > 0;\ \theta > 0;\ \sigma \geq 0$
Cauchy	$\sigma^2(1+(\theta r)^\gamma)^{-\nu}$	$0 < \gamma \leq 2;\ \nu > 0;\ \theta > 0; \sigma \geq 0$
Spherical	$\sigma^2 \left(1 - \frac{2}{\pi}(\frac{r}{\theta}\sqrt{1-(\frac{r}{\theta})^2} + \sin^{-1}\frac{r}{\theta})\right)$	$c(r) = 0$ if $r > \theta;\ \sigma^2 \geq 0;\ \theta > 0$

Given the separation, $(||\boldsymbol{h}||, |u|)$ between all pairs of variables i and j, given by $R_{i,j}$, the dynamic (time-dependent) covariance matrix $\boldsymbol{\Sigma}_t$ may be formed as

$$\boldsymbol{\Sigma}_t = \begin{pmatrix} C_t(R_{1,1}) & C_t(R_{1,2}) & \dots & C_t(R_{1,p}) \\ C_t(R_{2,1}) & \ddots & & \vdots \\ \vdots & & & \\ C_t(R_{p,1}) & \dots & & C_t(R_{p,p}) \end{pmatrix} . \tag{3}$$

Therefore, we can specify a covariance matrix of arbitrary size by a covariance function, with a small number of parameters, and the known separation matrix, R.

Alternatively we can consider a matrix decomposition of the covariance matrix or its inverse, in this case the Modified Cholesky Decomposition of $\boldsymbol{\Sigma}^{-1}$ [Pourahmadi, 1999],

$$\boldsymbol{\Sigma}^{-1} = \mathbf{T}^{\top}\mathbf{D}^{-2}\mathbf{T}\,, \tag{4}$$

where $\mathbf{D}^2$ is a diagonal matrix with $\mathrm{D}^2_{jj} = \exp(\eta_{j+d})$, for $j = 1, \ldots, d$, and

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \eta_{2d+1} & 1 & 0 & \cdots & 0 \\ \eta_{2d+2} & \eta_{2d+3} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \eta_{q-d+2} & \eta_{q-d+3} & \cdots & \eta_q & 1 \end{pmatrix}. \tag{5}$$

However, what if the dependency structure we would like to model is

- Non-stationary, i.e. $C_t(\cdot)$ depends on specific location $\boldsymbol{s}$ or lead time l, or
- Dynamic, evolves over time or via a random process or via dependence on a time-varying covariate?

We can model these behaviours in a parsimonious fashion by allowing the parameters to covariance functions $C_t(\cdot)$, or elements of the modified Cholesky decomposition η_i to be additive models of covariates (which may include $\boldsymbol{s}$ and/or l).

FLEXIBLE COVARIANCE MODELLING

Let $C(r;\boldsymbol{\xi})$ be a covariance function parametrised by the m-dimensional parameter vector $\boldsymbol{\xi}$. The elements of $\boldsymbol{\xi}$ are modelled via

$$g_j(\xi_j) = \mathbf{A}_{j,t}\boldsymbol{\beta}_j + \sum_i f_{j,i}(\mathbf{x}_t^{S_{j,i}}), \quad \text{for} \quad j = 1, \ldots, m, \tag{6}$$

a Generalised Additive Model, the parameters of which (including regularisation) are to be estimated. Details in [Browell et al., 2022].

In the MCD case, we model a subset of the elements of $\boldsymbol{\eta}$ in exactly the same way, selection precisely of which η_i to model is made via a boosting algorithm. Details in [Gioia et al., 2022].

EXAMPLES

The only difference between models is the covariance structure, all margins/density forecasts are the same.

We use standard scoring rules for multivariate probabilistic forecasting:

- Multivariate Energy Score (generalisation of CRPS)
- Log (or Ignorance) Score
- Variogram Score (with $p = 0.5$ and $p = 1$)

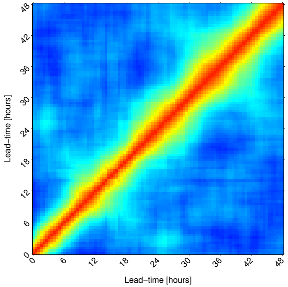


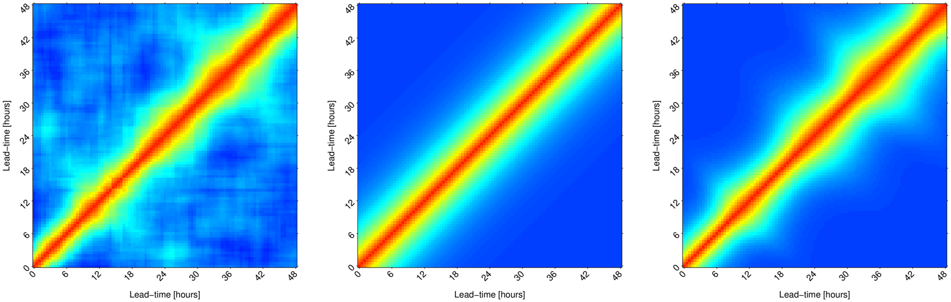
Figure 3: Empirical temporal dependency structure of wind power forecasts

The temporal dependency structure of wind power forecast is non-stationary and complex.

Modelled with exponential correlation function and cubic splines: θ becomes a smooth function of lead-time

$$\theta = \hat{\theta}_{\mathrm{cr}}(d) = \beta_0 + f_{\mathrm{cr}}(d) \quad .$$

where d is distance along the diagonal.



(a) Empirical (b) Constant $\hat{\theta}$ (Stationary) (c) GAC $\hat{\theta}_{cr}(d)$

Figure 4: Temporal dependency structure of wind power forecasts from 0 to 48 hours-ahead. Forecasts have a visible non-stationary structure. The width of the diagonal ridge indicates how long forecast errors are likely to persist for in time.

Table 2: Results different temporal dependency models for wind power forecasting. <u>Underline</u> indicates that the corresponding skill score relative to the GAC model are not significantly different from zero.

Name	Energy	Log	VS-0.5	VS-1
Empirical	<u>7.139</u>	Inf	<u>1409</u>	<u>5444</u>
Constant	<u>7.142</u>	19.86	<u>1409</u>	<u>5439</u>
GAC	7.137	**15.46**	1406	5433

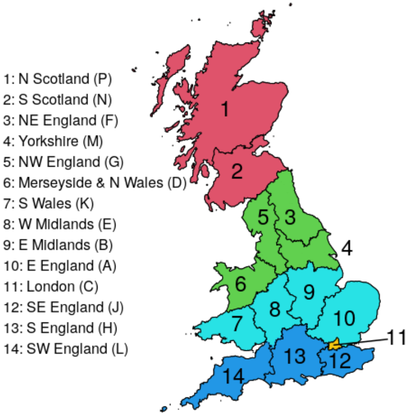


Figure 5: A map of the regions (Grid Supply Point groups) forming GB's electricity grid.

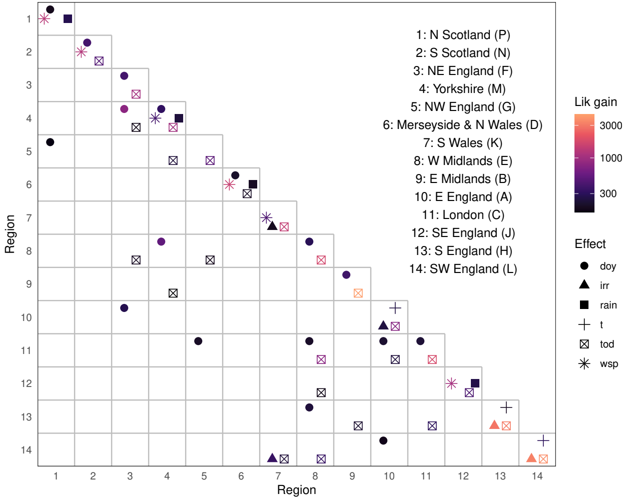


Figure 6: Model selection results. The diagonal corresponds to the elements of **D**, the rest to those of **T**.

	Scot - Rest		South - Rest		Lon - Neigh.	
	Log	CRPS	Log	CRPS	Log	CRPS
Indep	5879	6169	4495	4310	2842	1491
Static	4645	5790	4206	4221	2850	1489
Cal	4543	5701	<u>4117</u>	<u>4150</u>	2715	1454
Cal+Ren	<u>4541</u>	<u>5698</u>	4121	<u>4150</u>	<u>2695</u>	<u>1450</u>
Full	4545	5703	4122	4153	2703	1452

Table 3: Day-ahead performance when forecasting the marginal distribution of differences in net-demand across key boundaries. The best score in each column has been <u>underlined</u>.

SUMMARY

SUMMARY

- We present two approaches to model dynamic and non-stationary covariance structures flexibly
- Doing so may substantially improve the quality of multi-variate probabilistic forecasts (and other covariance-based models!)
- There is much still to be done to understand and improve model selection and estimation...

Full details including more examples, code and data can be found in [Browell et al., 2022, Gioia et al., 2022].

Papers, slides, code and more linked from www.jethrobrowell.com

References:

- Browell, J., Gilbert, C., and Fasiolo, M. (2022).
  Covariance structures for high-dimensional energy forecasting.
  Electric Power Systems Research, 211:108446.
- Gioia, V., Fasiolo, M., Browell, J., and Bellio, R. (2022).
  Additive Covariance Matrix Models: Modelling Regional Electricity Net-Demand in Great Britain.
- Pourahmadi, M. (1999).
  Joint Mean-Covariance Models with Applications to Longitudinal Data: Unconstrained Parameterisation.
  Biometrika, 86(3):677–690.