Identifying the effect of public holidays on daily demand for gas

Sarah Heaps

Joint work with Malcolm Farrow and Kevin Wilson

School of Mathematics, Statistics and Physics, Newcastle University, UK

StanCon 2018 Helsinki,
August 31, 2018

Outline

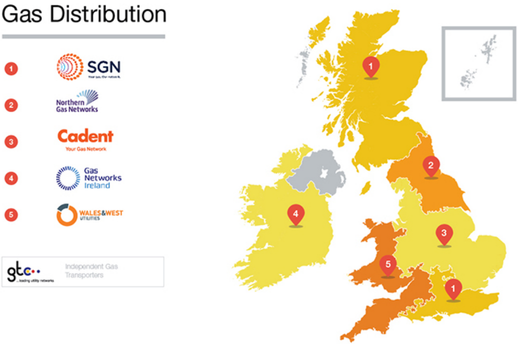
1. Introduction
2. Data
3. Hidden Markov model
4. Prior distribution (time permitting)
5. Posterior computation
6. Results
7. Conclusion

The UK energy sector

- The energy sector in the UK is changing.
- The UK Climate Change Act 2008 requires greenhouse gas emissions to be reduced to 80% of their 1990 levels by 2050.
- At present, natural gas is a comparatively low cost source of energy. Unlike some renewables, e.g. wind, it provides reliability in supply and storability.
- In future, decarbonisation of the gas network is likely to involve switching to carbon-neutral bio-methane or hydrogen.
- Both now and in the future, gas has a vital role to play in the energy mix provided costs remain low.
- It is more important than ever that the gas distribution network operates as efficiently as possible.

Gas distribution in the UK

National Grid is the sole owner and operator of the gas transmission infrastructure in the UK.



National Grid works with the regional distribution networks to balance local supply and demand for gas.

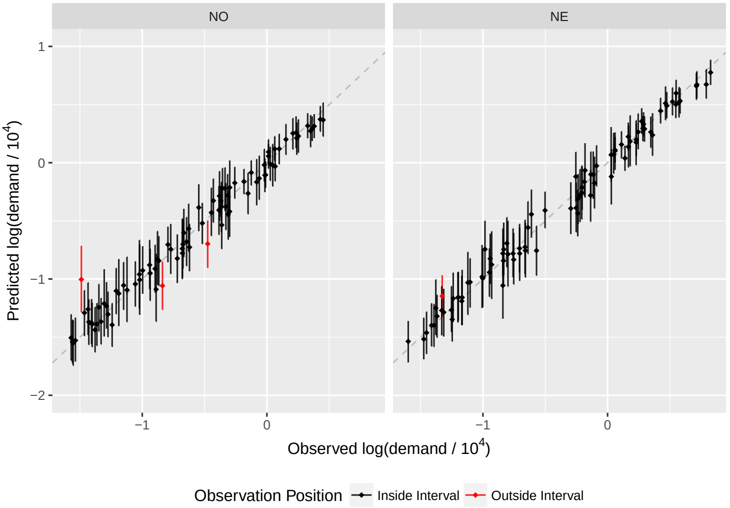
Forecasting the demand for gas

- Gas distribution companies, like Northern Gas Networks (NGN) need to be able to forecast the demand for gas at various levels of aggregation and temporal resolution, over a range of horizons. For example:
  - Short-term, hourly forecasting for individual “offtakes”: needed to allow steady removal of gas from national transmission system throughout the day.
  - Medium-term, daily forecasting for large geographical regions: needed to support investment and planning processes.
- Accurate forecasts help to reduce operational costs, e.g. through lower investment in storage facilities, avoidance of regulatory penalties.

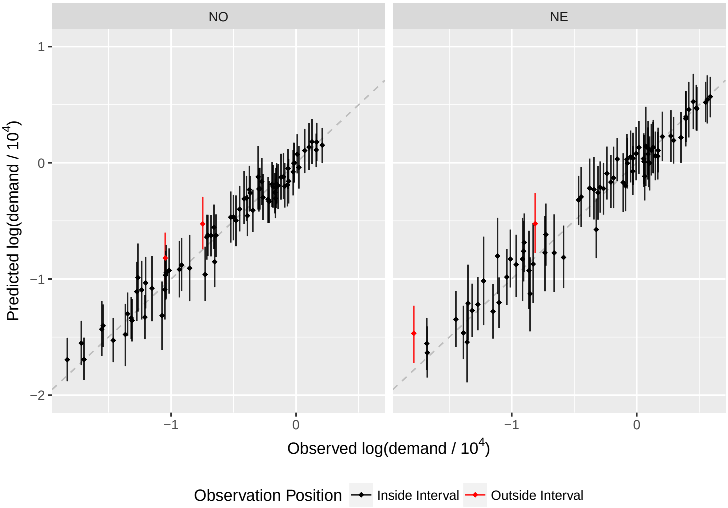
Modelling and forecasting daily gas demand

- Three types of consumers: residential, commercial and industrial.
- Amongst residential and commercial consumers, gas is predominantly used for heating and cooking.
- Most models for daily gas demand have components:
  1. Weather-related predictors: primitive variables (e.g. temperature), derived variables;
  2. Seasonal and calendar effects: day-of-the-year, day-of-the-week, public holidays and proximity days;
  3. Possibly an interaction between (1) and (2);
  4. Allowance for autocorrelation amongst residuals.
- Proximity days: days in neighbourhood of public holiday, allowing for a protracted effect, e.g. due to changes in cooking and travel arrangements, commercial slow-down, etc.

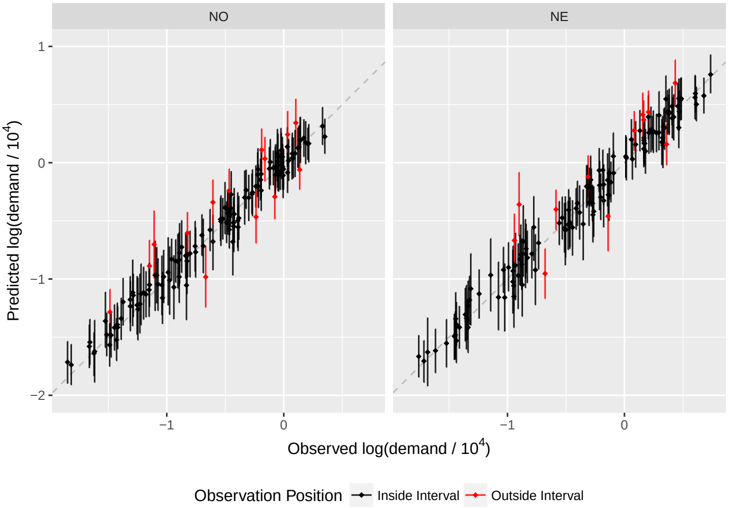
Model fit: "ordinary" day (10 days from public holiday)



Model fit: public holiday



Model fit: "proximity" day (1 day from public holiday)



The proximity effect

- Public holidays often occur during periods where the demand for gas is at its highest. At these times it is essential to have accurate forecasts and a correct quantification of uncertainty, e.g., to allow appropriate investment to support a severe winter.
- In the literature, proximity days are defined as those within a fixed window around each public holiday. How wide? Same for all holidays? Symmetric?
- We investigate a flexible non-homogeneous hidden Markov model, with cyclic dynamics, which allows the existence and duration of proximity effects to be unknown.
- The results of a preliminary version of this model are already being used by NGN in its annual medium-term forecasting exercise.
- General modelling idea can be applied in other applications to smooth the transition between two observed states, e.g. disease state models.

Northern Gas Networks (NGN)

- NGN is a regional distribution network, responsible for gas distribution to 2.7 million homes and businesses across a $25000\mathrm{km}^2$ region in the North East of England, Northern Cumbria and Yorkshire.
- National Grid requires medium and long term demand forecasts for thirteen local distribution zones (LDZs).
- NGN is responsible for the Northern (NO) and North East (NE) LDZs.

Gas demand data

- Every consumer meter point or offtake is categorised according to its load band, based on how much gas it uses.
- We focus on the first three non-daily metered (NDM) load bands:

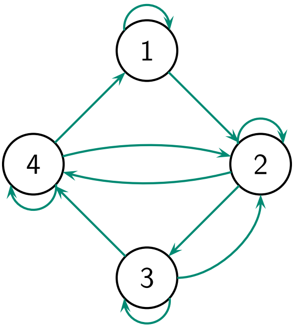
Index	Load Band MWh / year	Example
1	0-73 MWh	Domestic, e.g. a single house
2	73-732 MWh	Commercial premises
3	732-5860 MWh	Small industrial premises

- For both LDZs and each of NDM load band, daily gas consumption data are available from 1st January 2008 to 18th February 2017.

Notation

- $\tilde{Y}_{t,j}$: gas demand in tenths of a gigawatt-hour (GWh) from a given NDM load band on day t in the region j ($j = 1$ refers to NO, $j = 2$ refers to NE).
- Define $Y_{t,j} = \ln(\tilde{Y}_{t,j})$ and $\boldsymbol{Y}_t = (Y_{t,1}, Y_{t,2})'$.
- Explanatory variables:
  - $w_{t,j}$: composite weather variable (CWV) on day t in region j and $\boldsymbol{w}_t = (w_{t,1}, w_{t,2})'$;
  - $n_t \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$: number of days to the next public holiday;
  - $p_t \in \mathbb{N}_0$: number of days since the previous public holiday;
  - $r_t \in \{1, 2, 3\}$: "type" of the nearest public holiday. Go
- Rather than defining proximity days as deterministic functions of (p_t, n_t), we take a more flexible approach and allow the number of proximity days on either side of each public holiday to be unknown, taking values in $\mathbb{N}_0$.

Hidden states

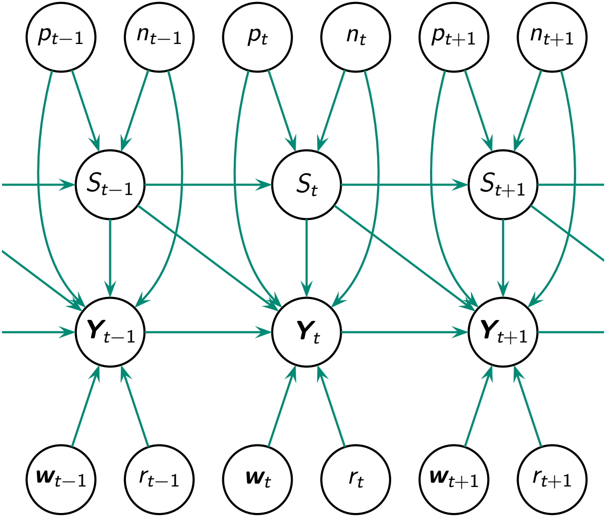


- We introduce a discrete-valued stochastic process $\{S_t : t = 0, 1, \ldots, T\}$ where $S_t \in \mathcal{S}_s = \{1, 2, 3, 4\}$.
- State 2 can only apply to public holidays and is observable. States 1, 3, and 4 are "hidden".
- State 1 allows for "pre-holiday" proximity days.
- State 3 allows for "post-holiday" proximity days.
- State 4 is a baseline for "normal" days.

Hidden state process

- How can we allow State 2 to be observable and use our knowledge of the dates of future holidays?
  - Homogeneous pre-process: start with homogeneous Markov chain, then condition on $S_t = 2$ if day t is a public holiday and $S_t \neq 2$ otherwise.
  - Non-homogeneous process: allow transition probabilities to vary so $S_t = 2$ is certain if day t is a public holiday (impossible otherwise) and transitions into proximity state depend on distance to holiday.
- Expert judgement from engineers at NGN: any proximity effect will last for one or two days on either side of a public holiday; longer periods are very unlikely.
- Capturing this effect demands non-geometric decay of the sojourn time in the proximity states.

Directed acyclic graph (DAG)



Modelling the states

- Suppose that the states S_t follows a first order, non-homogeneous Markov chain with

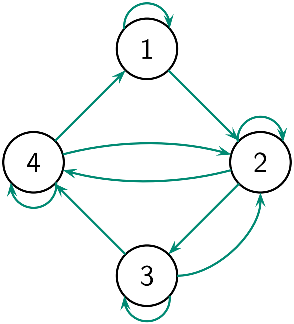
$$\Pr(S_t = k | S_{t-1} = j, n_t, p_t) = \lambda_{j,k}(n_t, p_t), \qquad (j, k) \in \mathcal{S}_s^2$$

for $t = 1, 2, \ldots, T$, and initial distribution

$$\Pr(S_0 = k | n_0, p_0) = \ell_k(n_0, p_0), \qquad k \in \mathcal{S}_s.$$

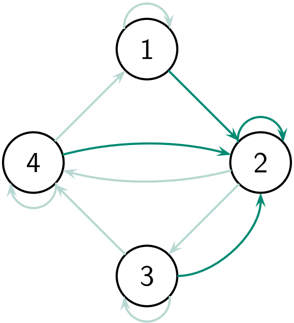
- State 2 is observable, so if day 0 is a public holiday, then $n_0 = p_0 = 0$ and $\ell_2(0, 0) = 1$.
- Otherwise $(n_0, p_0) \in \mathbb{N}_0^2 \setminus \{(0, 0)\}$ and we assign $\ell_k(n_0, p_0) = 1/3$ for $k \in \mathcal{S}_s \setminus \{2\}$. Note: $\ell_2(n_0, p_0) = 0$.

Modelling the states cont'd



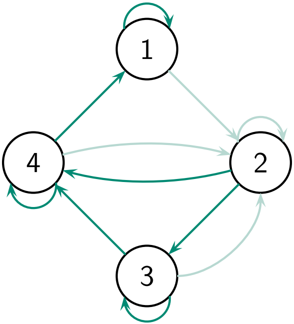
- If day t is a public holiday, then $n_t = p_t = 0$ and $\lambda_{j,2}(0,0) = 1$ for any $j \in \mathcal{S}_s$.
- Otherwise $\lambda_{j,2}(n_t, p_t) = 0$ and we impose cyclic dynamics on the state evolution.
- We need to define $\lambda_{2,3}(n_t, p_t)$, $\lambda_{3,4}(n_t, p_t)$ and $\lambda_{4,1}(n_t, p_t)$.
- Note: an uninterrupted spell of proximity days between two public holidays are post-, rather than pre-, holiday days.

Modelling the states cont'd



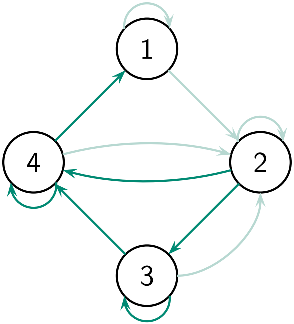
- If day t is a public holiday, then $n_t = p_t = 0$ and $\lambda_{j,2}(0,0) = 1$ for any $j \in \mathcal{S}_s$.
- Otherwise $\lambda_{j,2}(n_t, p_t) = 0$ and we impose cyclic dynamics on the state evolution.
- We need to define $\lambda_{2,3}(n_t, p_t)$, $\lambda_{3,4}(n_t, p_t)$ and $\lambda_{4,1}(n_t, p_t)$.
- Note: an uninterrupted spell of proximity days between two public holidays are post-, rather than pre-, holiday days.

Modelling the states cont'd



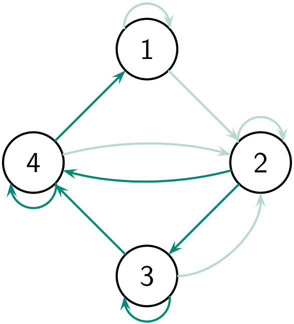
- If day t is a public holiday, then $n_t = p_t = 0$ and $\lambda_{j,2}(0,0) = 1$ for any $j \in \mathcal{S}_s$.
- Otherwise $\lambda_{j,2}(n_t, p_t) = 0$ and we impose cyclic dynamics on the state evolution.
- We need to define $\lambda_{2,3}(n_t, p_t)$, $\lambda_{3,4}(n_t, p_t)$ and $\lambda_{4,1}(n_t, p_t)$.
- Note: an uninterrupted spell of proximity days between two public holidays are post-, rather than pre-, holiday days.

Modelling the states cont'd



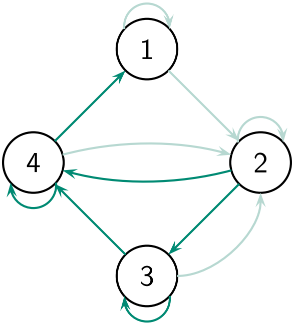
- If day t is a public holiday, then $n_t = p_t = 0$ and $\lambda_{j,2}(0,0) = 1$ for any $j \in \mathcal{S}_s$.
- Otherwise $\lambda_{j,2}(n_t, p_t) = 0$ and we impose cyclic dynamics on the state evolution.
- We need to define $\lambda_{2,3}(n_t, p_t)$, $\lambda_{3,4}(n_t, p_t)$ and $\lambda_{4,1}(n_t, p_t)$.
- Note: an uninterrupted spell of proximity days between two public holidays are post-, rather than pre-, holiday days.

Modelling the states cont'd



- If day t is a public holiday, then $n_t = p_t = 0$ and $\lambda_{j,2}(0,0) = 1$ for any $j \in \mathcal{S}_s$.
- Otherwise $\lambda_{j,2}(n_t, p_t) = 0$ and we impose cyclic dynamics on the state evolution.
- We need to define $\lambda_{2,3}(n_t, p_t)$, $\lambda_{3,4}(n_t, p_t)$ and $\lambda_{4,1}(n_t, p_t)$.
- Note: an uninterrupted spell of proximity days between two public holidays are post-, rather than pre-, holiday days.

Modelling the states cont'd



- If day t is a public holiday, then $n_t = p_t = 0$ and $\lambda_{j,2}(0,0) = 1$ for any $j \in \mathcal{S}_s$.
- Otherwise $\lambda_{j,2}(n_t, p_t) = 0$ and we impose cyclic dynamics on the state evolution.
- We need to define $\lambda_{2,3}(n_t, p_t)$, $\lambda_{3,4}(n_t, p_t)$ and $\lambda_{4,1}(n_t, p_t)$.
- Note: an uninterrupted spell of proximity days between two public holidays are post-, rather than pre-, holiday days.

Modelling the transition probabilities

- Expert judgement from engineers at NGN: any proximity effect will last for one or two days on either side of a public holiday; longer periods are very unlikely.
- How should we model $\lambda_{4,1}(n_t, p_t)$, i.e. Pr("normal"→"pre-hol.")?
- How about $\lambda_{4,1}(n_t, p_t) = \lambda_{4,1}$? No good: want higher probability when public holiday is near.
- We take

$$\text{logit}\{\lambda_{4,1}(n_t, p_t)\} = \nu_{4,1,1} + \frac{\nu_{4,1,2}}{10}(n_t - 1)^{1/2}$$

- The prior we induce for $\lambda_{4,1}(n_t, p_t)$ is unimodal for all feasible values of n_t.

Modelling the transition probabilities

- Expert judgement from engineers at NGN: any proximity effect will last for one or two days on either side of a public holiday; longer periods are very unlikely.
- How should we model $\lambda_{4,1}(n_t, p_t)$, i.e. Pr("normal"→"pre-hol.")?
- How about $\lambda_{4,1}(n_t, p_t) = \lambda_{4,1}$? No good: want higher probability when public holiday is near.
- We take

$$\text{logit}\{\lambda_{4,1}(n_t, p_t)\} = \nu_{4,1,1} + \frac{\nu_{4,1,2}}{10}(n_t - 1)^{1/2}$$

- The prior we induce for $\lambda_{4,1}(n_t, p_t)$ is unimodal for all feasible values of n_t.

Modelling the transition probabilities

- Expert judgement from engineers at NGN: any proximity effect will last for one or two days on either side of a public holiday; longer periods are very unlikely.
- How should we model $\lambda_{4,1}(n_t, p_t)$, i.e. Pr("normal"→"pre-hol.")?
- How about $\lambda_{4,1}(n_t, p_t) = \lambda_{4,1}$? No good: want higher probability when public holiday is near.
- We take

$$\operatorname{logit}\{\lambda_{4,1}(n_t, p_t)\} = \nu_{4,1,1} + \frac{\nu_{4,1,2}}{10}(n_t - 1)^{1/2}$$

- The prior we induce for $\lambda_{4,1}(n_t, p_t)$ is unimodal for all feasible values of n_t.

Modelling the transition probabilities

- Expert judgement from engineers at NGN: any proximity effect will last for one or two days on either side of a public holiday; longer periods are very unlikely.
- How should we model $\lambda_{4,1}(n_t, p_t)$, i.e. Pr("normal"→"pre-hol.")?
- How about $\lambda_{4,1}(n_t, p_t) = \lambda_{4,1}$? No good: want higher probability when public holiday is near.
- We take

$$\text{logit}\{\lambda_{4,1}(n_t, p_t)\} = \nu_{4,1,1} + \frac{\nu_{4,1,2}}{10}(n_t - 1)^{1/2}$$

- The prior we induce for $\lambda_{4,1}(n_t, p_t)$ is unimodal for all feasible values of n_t.

Modelling the transition probabilities cont'd

- How should we model $\lambda_{3,4}(n_t, p_t)$, i.e. Pr("post-hol."→"normal")?
- How about $\lambda_{3,4}(n_t, p_t) = \lambda_{3,4}$? No good: geometric distribution of sojourn time inflexible.
- We take

$$\text{logit}\{\lambda_{3,4}(n_t, p_t)\} = \nu_{3,4,1} + \frac{\nu_{3,4,2}}{10}(p_t - 2)^{1/2} + \nu_{3,4,3}\mathbb{I}(n_t = 1).$$

- The final term in n_t allows the transition probability to differ over a weekend between two public holidays, e.g. period between Good Friday and Easter Monday.
- Similarly we model $\lambda_{2,3}(n_t, p_t)$, i.e. Pr("hol."→"post-hol."), using

$$\text{logit}\{\lambda_{2,3}(n_t, p_t)\} = \nu_{2,3,1} + \nu_{2,3,2}\mathbb{I}(n_t = 2).$$

Modelling the conditional demand for gas

- We model log demand $\boldsymbol{Y}_t$, conditional on S_t and S_{t-1}, as a first-order vector autoregression:

$$\boldsymbol{Y}_t - \boldsymbol{\mu}_t = \Psi(\boldsymbol{y}_{t-1} - \boldsymbol{\mu}_{t-1}) + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}_2(\boldsymbol{0}, \Omega_t^{-1})$$

  for $t = 2, \ldots, T$ with initial distribution:

$$\boldsymbol{Y}_1 = \boldsymbol{\mu}_1 + \boldsymbol{\epsilon}_1, \quad \boldsymbol{\epsilon}_1 \sim \mathrm{N}_2\left(\boldsymbol{0}, V(\Psi, \Omega_1)\right).$$

  where $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_T$ are independent, zero-mean bivariate normal random vectors.
- Dependence on the states comes through the time-dependent mean $\boldsymbol{\mu}_t$ and precision matrix Ω_t.
- $\Psi \in \mathcal{S}$ is a (2×2) real-valued matrix, constrained to satisfy the stationarity condition of a bivariate AR(1) process.
- $V(\Psi, \Omega_1)$ is the stationary variance of $\{\boldsymbol{Y}_t - \boldsymbol{\mu}_t : t = 1, 2, \ldots\}$ if $\Omega_t = \Omega_1$ for all t, i.e. V solving $V = \Psi V \Psi' + \Omega_1^{-1}$.

Time-dependent mean, μ_t

Conditional on the state, $S_t = s_t$, the mean for LDZ j, is given by

$$\mu_{t,j} = \alpha_j + B_{t,j}\beta_{j,r_t} + \Gamma_{t,j} + \Delta_{t,j} + (\eta_{j,1} + \eta_{j,2} w_{t,j})\tilde{w}_{t,j}.$$

- $\tilde{w}_{t,j} = w_{t,j} - m_{d(t),j}$, where $m_{d(t),j}$ is a smoothed average for the CWV in region j on day t of the year;
- $\Delta_{t,j}$ is constructed to give a day-of-the-week effect:

$$\Delta_{t,j} = \sum_{k=1}^{3} \left\{ \delta_{j,1,k} \cos\left(\frac{2\pi kt}{7}\right) + \delta_{j,2,k} \sin\left(\frac{2\pi kt}{7}\right) \right\}$$

- $\Gamma_{t,j}$ provides a seasonal effect:

$$\Gamma_{t,j} = \sum_{k=1}^{K_\gamma} \left\{ \gamma_{j,1,k} \cos\left(\frac{2\pi kt}{365.25}\right) + \gamma_{j,2,k} \sin\left(\frac{2\pi kt}{365.25}\right) \right\}.$$

- $B_{t,j}\beta_{j,r_t}$ provides a public holiday or proximity effect.

Time-dependent mean, μ_t, cont'd

Conditional on the state, $S_t = s_t$, the mean for LDZ j, is given by

$$\mu_{t,j} = \alpha_j + B_{t,j}\beta_{j,r_t} + \Gamma_{t,j} + \Delta_{t,j} + (\eta_{j,1} + \eta_{j,2} w_{t,j})\tilde{w}_{t,j}.$$

- $\beta_{j,1}$, $\beta_{j,2}$ and $\beta_{j,3}$ indicate the effect of each type of public holiday.
- $B_{t,j}$ depends on s_t and controls if and how the public holiday effect is included:

$$B_{t,j} = \begin{cases} \rho_{\beta,j}^{n_t}, & \text{if } s_t = 1 \text{ (``pre-hol.'')}, \\ 1, & \text{if } s_t = 2 \text{ (``holiday'')}, \\ \rho_{\beta,j}^{\min(n_t,p_t)}, & \text{if } s_t = 3 \text{ (``post-hol.'')}, \\ 0, & \text{if } s_t = 4 \text{ (``normal'')}, \end{cases}$$

  where $\rho_{\beta,j} \in (0,1)$ so that the holiday effect β_{j,r_t} is scaled down on proximity days.
- The scaling factor decays to zero with increasing separation.

Time-dependent precision matrix, Ω_t

- Ω_t depends on S_t, with different values in the holiday ($S_t = 2$) and normal ($S_t = 4$) states.
- Like the mean $\boldsymbol{\mu}_t$, Ω_t is modelled on proximity days (states 1 and 3) by interpolating between its "holiday" and "normal" values, with weights dependent on n_t and p_t.
- This is more natural in $\mathbb{R}^3$ than the constrained space of 2×2 covariance matrices, so Ω_t is reparameterised in terms of its square-root free Cholesky decomposition, $\Omega_t = T'_\Omega D_\Omega^{-1} T_\Omega$, as in Pourahmadi (1999).
- The new parameters, $(\omega_{t,1}, \omega_{t,2}, \omega_{t,3})'$ can be interpreted as autoregressive coefficients and log error precisions in the marginal / conditional factorisation of the density of ϵ_t.

Time-dependent precision matrix, Ω_t

Conditional on the state, $S_t = s_t$, component i of the reparameterised precision matrix is given by

$$\omega_{t,i} = \eta_i + \Theta_{t,i}\theta_i + K_{t,i}$$

- $K_{t,i}$ provides a seasonal effect:

$$K_{t,i} = \sum_{k=1}^{\mathrm{K}_\kappa} \left\{ \kappa_{i,1,k} \cos\left(\frac{2\pi kt}{365.25}\right) + \kappa_{i,2,k} \sin\left(\frac{2\pi kt}{365.25}\right) \right\}$$

- θ_i is a public holiday effect; $\Theta_{t,i}$ depends on s_t and controls if and how the effect is included:

$$\Theta_{t,i} = \begin{cases} \rho_\theta^{n_t}, & \text{if } s_t = 1 \text{ (``pre-hol.'')}, \\ 1, & \text{if } s_t = 2 \text{ (``holiday'')}, \\ \rho_\theta^{\min(n_t, p_t)}, & \text{if } s_t = 3 \text{ (``post-hol.'')}, \\ 0, & \text{if } s_t = 4 \text{ (``normal'')}, \end{cases}$$

where $\rho_\theta \in (0, 1)$.

Prior distribution

- We highlight one aspect of our prior distribution.
- Consider the conditional model for log demand:

$$\boldsymbol{Y}_t - \boldsymbol{\mu}_t = \Psi(\boldsymbol{y}_{t-1} - \boldsymbol{\mu}_{t-1}) + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}_2(\mathbf{0}, \Omega_t^{-1}).$$

- We focus on the prior for the autoregressive coefficient matrix Ψ.

Prior for autoregressive coefficient matrix

- Consider a univariate AR(1) model:

$$Y_t - \mu_t = \Psi(y_{t-1} - \mu_{t-1}) + \epsilon_t, \quad \epsilon_t \sim \mathrm{N}(0, \Omega^{-1}).$$

- The stationarity condition demands that the root λ of the characteristic equation,

$$|\Psi - \lambda| = 0 \quad \iff \quad \lambda = \Psi$$

  lies inside the unit circle. That is, $|\Psi| < 1$.

- Now consider a bivariate AR(1) model:

$$\boldsymbol{Y}_t - \boldsymbol{\mu}_t = \Psi(\boldsymbol{y}_{t-1} - \boldsymbol{\mu}_{t-1}) + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}_2(\boldsymbol{0}, \Omega^{-1}).$$

- Again, the stationarity condition demands that the roots λ of the characteristic equation,

$$|\Psi - \lambda I_2| = 0,$$

  lie inside the unit circle.

Prior for autoregressive coefficient matrix

- Consider a univariate AR(1) model:

$$Y_t - \mu_t = \Psi(y_{t-1} - \mu_{t-1}) + \epsilon_t, \quad \epsilon_t \sim \mathrm{N}(0, \Omega^{-1}).$$

- The stationarity condition demands that the root λ of the characteristic equation,

$$|\Psi - \lambda| = 0 \quad \iff \quad \lambda = \Psi$$

lies inside the unit circle. That is, $|\Psi| < 1$.

- Now consider a bivariate AR(1) model:

$$\boldsymbol{Y}_t - \boldsymbol{\mu}_t = \Psi(\boldsymbol{y}_{t-1} - \boldsymbol{\mu}_{t-1}) + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathrm{N}_2(\mathbf{0}, \Omega^{-1}).$$

- Again, the stationarity condition demands that the roots λ of the characteristic equation,

$$|\Psi - \lambda I_2| = 0,$$

lie inside the unit circle.

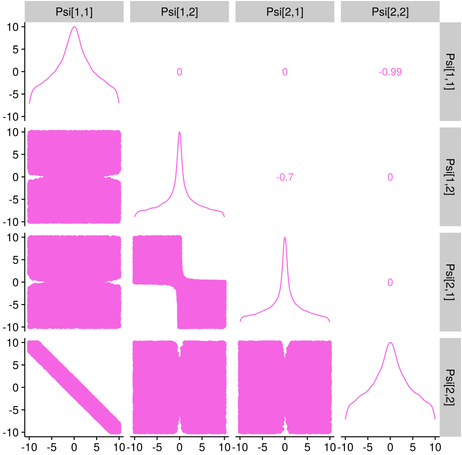
The stationary region

- This corresponds to a subset $\mathcal{S}$ of $\mathbb{R}^4$ bound by the following constraints:

$$\begin{aligned} \operatorname{tr}(\Psi) + \det(\Psi) &> -1, \\ \operatorname{tr}(\Psi) - \det(\Psi) &< 1, \\ \det(\Psi) &< 1. \end{aligned}$$

- We want to construct a prior for Ψ which only has non-zero density in $\mathcal{S}_0 \subseteq \mathcal{S}$.
- The region has $\mathcal{S}$ has an unusual geometry.

The stationary region: bivariate marginal spaces



Reparameterising Ψ

- If we constrain $\Psi_{11} = \Psi_{2,2} = \Psi_{on}$ and $\Psi_{1,2} = \Psi_{2,1} = \Psi_{off}$, the constraints are just

$$|\Psi_{on} + \Psi_{off}| < 1 \quad \text{and} \quad |\Psi_{on} - \Psi_{off}| < 1.$$

- We can easily specify a prior over this region by reparameterising as

$$\xi_1 = (\Psi_{on} + \Psi_{off} + 1)/2 \quad \text{and} \quad \xi_2 = (\Psi_{on} - \Psi_{off} + 1)/2$$

  and then assigning (ξ_1, ξ_2) a distribution over the unit square.
- For example,

$$\xi_i \sim \mathrm{Beta}(a_{\xi,i}, b_{\xi,i}), \quad \text{independently for } i = 1, 2.$$

Posterior distribution

- The posterior distribution of the model parameters follows from Bayes Theorem as

$$\pi(\Pi, \Lambda | \boldsymbol{y}, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}) \propto p(\boldsymbol{y} | \Pi, \Lambda, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}) \pi(\Pi) \pi(\Lambda).$$

- The observed data likelihood, $p(\boldsymbol{y} | \Pi, \Lambda, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w})$, is calculated using a filtering algorithm which computes $\Pr(S_t = k, \boldsymbol{y}_{1:t} | \Pi, \Lambda, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w})$, $k \in \mathcal{S}_s$, recursively for $t = 0, \ldots, T$ and finally

$$p(\boldsymbol{y} | \Pi, \Lambda, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}) = \sum_{k=1}^{4} \Pr(S_T = k, \boldsymbol{y}_{1:T} | \Pi, \Lambda, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}).$$

Posterior computation

- The posterior cannot be evaluated in closed form. A numerical approximation is obtained using Hamiltonian Monte Carlo (HMC).
- And, of course, `rstan` (Stan Development Team, 2016; Carpenter et al., 2017) is used to implement the HMC algorithm. ☺
- The hidden states S_t are not sampled. Their marginal posteriors can be approximated by Rao-Blackwellisation through:

$$\hat{\Pr}(S_t = k | \boldsymbol{y}, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}) = \frac{1}{M} \sum_{i=1}^{M} \Pr(S_t = k | \boldsymbol{y}, \Pi^{[i]}, \Lambda^{[i]}, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}), \ k \in \mathcal{S}_s,$$

for $t = 0, 1, \ldots, T$ using the full sample smoothed probabilities.

Posterior computation

- The posterior cannot be evaluated in closed form. A numerical approximation is obtained using Hamiltonian Monte Carlo (HMC).
- And, of course, `rstan` (Stan Development Team, 2016; Carpenter et al., 2017) is used to implement the HMC algorithm. ☺
- The hidden states S_t are not sampled. Their marginal posteriors can be approximated by Rao-Blackwellisation through:

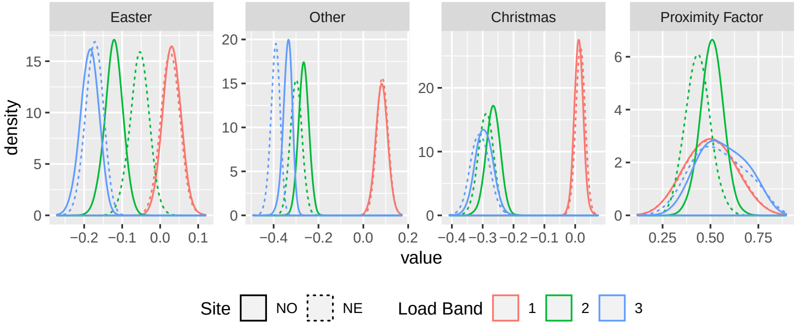
$$\hat{\Pr}(S_t = k | \boldsymbol{y}, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}) = \frac{1}{M} \sum_{i=1}^{M} \Pr(S_t = k | \boldsymbol{y}, \Pi^{[i]}, \Lambda^{[i]}, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{r}, \boldsymbol{w}), \; k \in \mathcal{S}_s,$$

for $t = 0, 1, \ldots, T$ using the full sample smoothed probabilities.

Application to NGN data

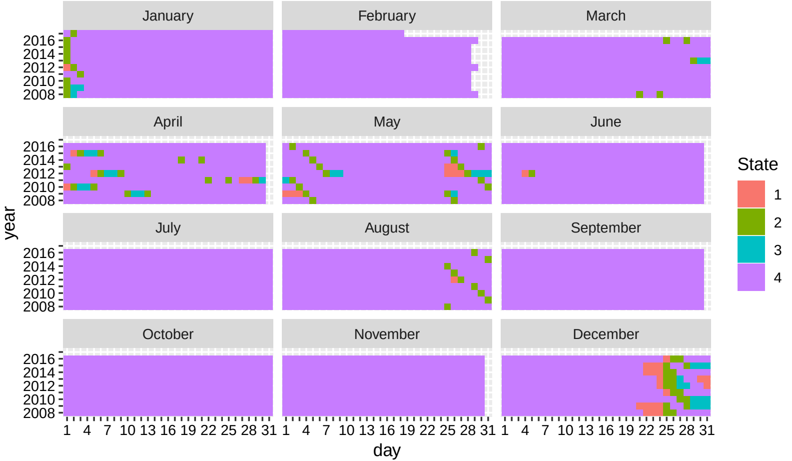
- The model was fitted to data from all NDM load bands.
- The number of harmonics in the seasonal components of the time-varying mean and precision matrix were truncated at $\mathrm{K}_\gamma = 6$ and $\mathrm{K}_\kappa = 12$. Higher harmonics had negligible amplitude.
- Using `rstan`, four HMC chains, initialised at different starting points, were run for 10000 iterations, discarding the first half as warm-up, and thinning to every 10 to reduce computational storage overheads.
- The usual graphical and numerical diagnostics gave no evidence of any lack of convergence.

Posterior for public holiday effects in the mean

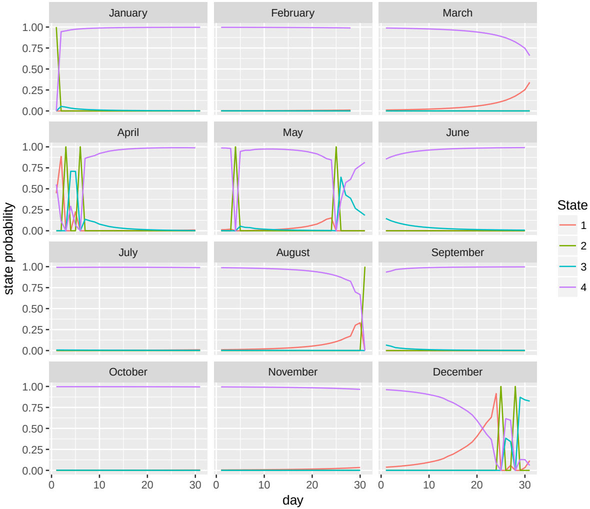


- We now focus on the results for load band 3 (small industrial customers).

Pointwise posterior mode of the states



Posterior uncertainty about the state allocation (2015)



Posterior predictive checks: 10 days from public holiday

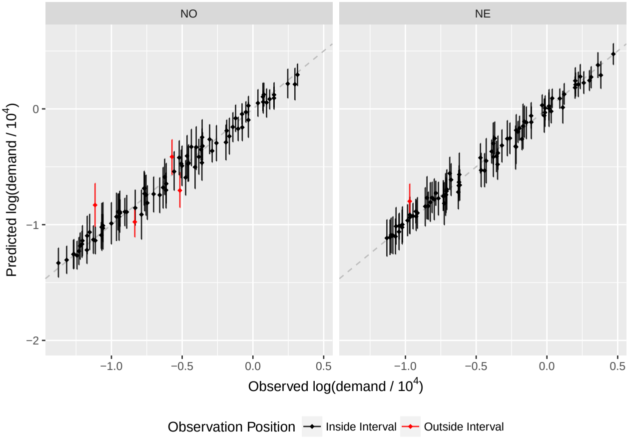


Figure: Two observed states, no proximity effect.

Posterior predictive checks: 10 days from public holiday

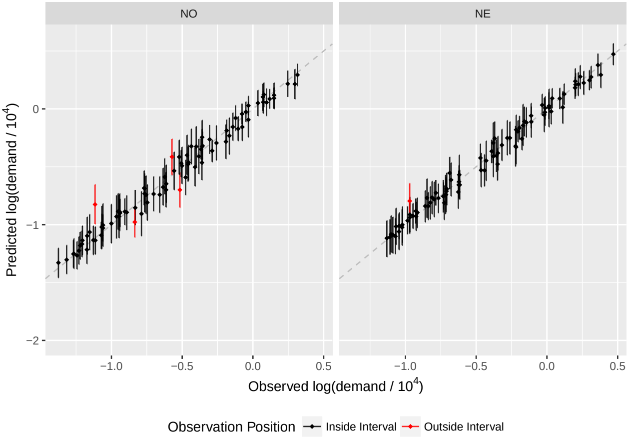


Figure: Non-homogeneous hidden Markov model.

Posterior predictive checks: 1 day from public holiday

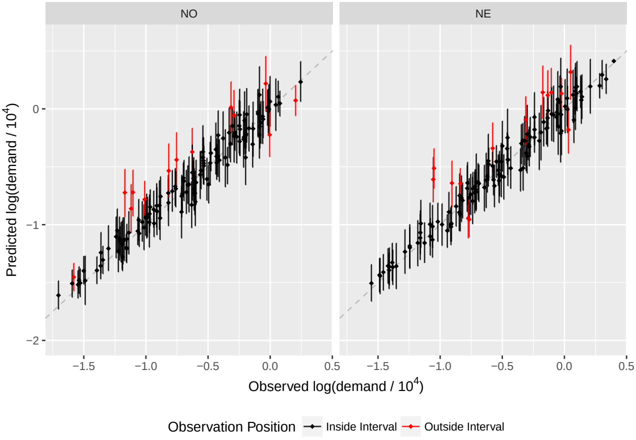


Figure: Two observed states, no proximity effect.

Posterior predictive checks: 1 day from public holiday

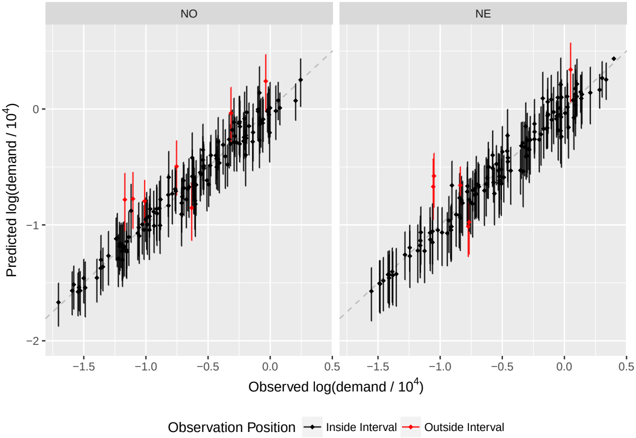


Figure: Non-homogeneous hidden Markov model.

Conclusion

- Changes in the energy industry mean that gas distribution networks need to operate more efficiently. Improved forecasting of the demand for gas is one way of doing this.
- Public holidays have a pronounced effect on the demand for gas that spreads into neighbouring days.
- We have developed a flexible, non-homogeneous hidden Markov model which allows for a protracted effect, of unknown size and duration, on either side of public holidays.
- We are exploring the problem of constructing a prior over the stationary region for the autoregressive coefficient of a bivariate AR(1) process in the more general, non-symmetric case.
- In the pipeline, ☺
    - Short-term forecasting of hourly offtakes;
    - Monitoring of alarms.

References

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell (2017). Stan: A probabilistic programming language. *Journal of Statistical Software 76*(1), 1–32.

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation. *Biometrika 86*, 677–690.

Stan Development Team (2016). *RStan: the R interface to Stan*. R package version 2.12.1.

Types of public holidays in England

Type	Name	Days
1	"Easter"	Good Friday, Easter Monday
2	"Christmas"	Christmas Day, Boxing Day, New Year's Day
3	"Other"	May Day, Spring bank holiday, Summer bank holiday, one-off holidays

Back