

Forecasting the mobility frequency with a dynamic linear model

Trond

1 Introduction

Long-term projections of internal migration are key inputs to regional-level demographic population forecasts. The long-term projections rely on the identification of a long-term trend in data, thus filtering out short-run fluctuations due to the business cycle. However, such detrending analyses are complicated if there is no clear trend in the data, if business cycles themselves stretch into the medium- to long run or if the data point represent the top or the bottom of the cycle (see e.g Canova 1998, and @hamilton2018you). Therefore it is informative to analyse the short- to medium dynamics of demographic components, forecasting along the business cycle. This paper develops a state space model for short-to medium term univariate forecast of monthly frequency of mobility in the Netherlands.

As is the case in most developed countries, the main driver of regional-level population change in the Netherlands is migration (external and domestic). Due to the close relation between migration decisions and labour- and housing market conditions, time series of internal migration flows often move along the business cycle. However, the volume of internal migration may also be affected by a myriad other factors. For example, a recent change to the Dutch student finance system seems to have caused a decrease in mobility for an otherwise very mobile group.¹ Finding the precise determinants of changes to migration flows can be very complicated. This justifies the use of univariate forecasting techniques, which make use of basic time series patterns to form a forecast (see e.g. Zietz and Traian 2014).

Forecasting in general entails decomposing a time series into elements such as trend, cycle and seasonality. One popular type of model used for this purpose are Dynamic Linear Models (DLM), which is a type of State Space models (Petrus, Petrone, and Campagnoli 2009; Durbin and Koopman 2012). As the name suggests, DLMS are linear models with a Gaussian error structure, where the relevant inferences are carried out using the Kalman filter. Unknown parameters of the model are estimated by either maximum likelihood or Bayesian techniques.

In this paper we focus on the national-level frequency of mobility, defined as the sum of intra- and inter-municipal moves per 1000 inhabitants, representing a rough measure of the intensity of internal migration in the Netherlands. Data from the Statistics Netherlands allow us to calculate monthly time series of the frequency of mobility, from January 1995 to December 2018, meaning we have in total 288 observations at our disposal.

In this paper I firstly show the time series data and discuss its patterns. Next I present the model, discuss its elements and the estimated parameters. Finally I use the model for a series of out-of-sample forecasts around a peak in the time series. In order to assess the performance, the prediction accuracy of the DLM model is compared to that from forecasts obtained with three other popular models for univariate forecasting.

2 Data

The frequency of mobility y_t in month t , is in this paper defined as the sum of inter- en intramunicipality moves between the first and the last day of month t , divided by the population on the last day of t . Figure 1 shows the time series from January 1995 to December 2018 as well as the seasonal differences $y'_t = y_t - y_{t-12}$.

¹<https://www.cbs.nl/nl-nl/nieuws/2018/04/studenten-gaan-minder-op-kamers>

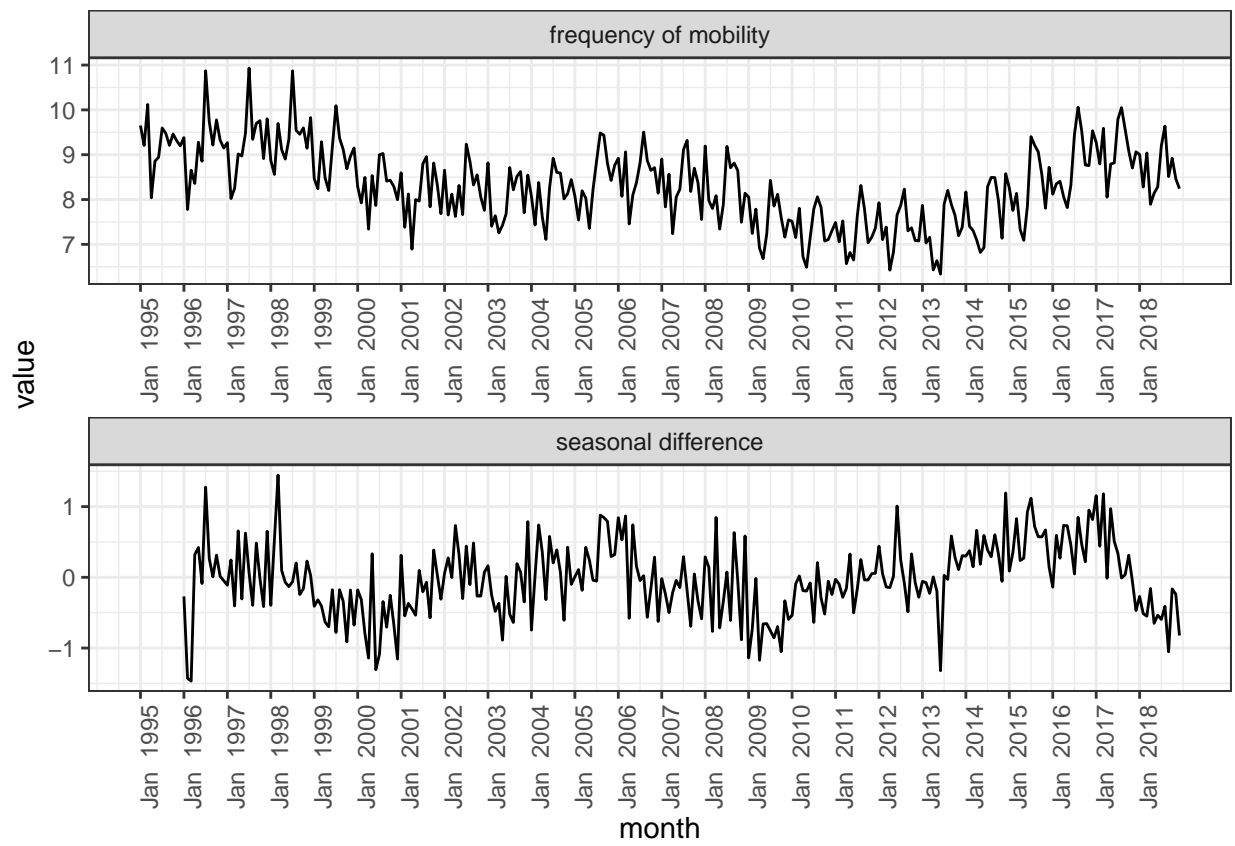


Figure 1: Mobility frequency (upper panel) and seasonal differences

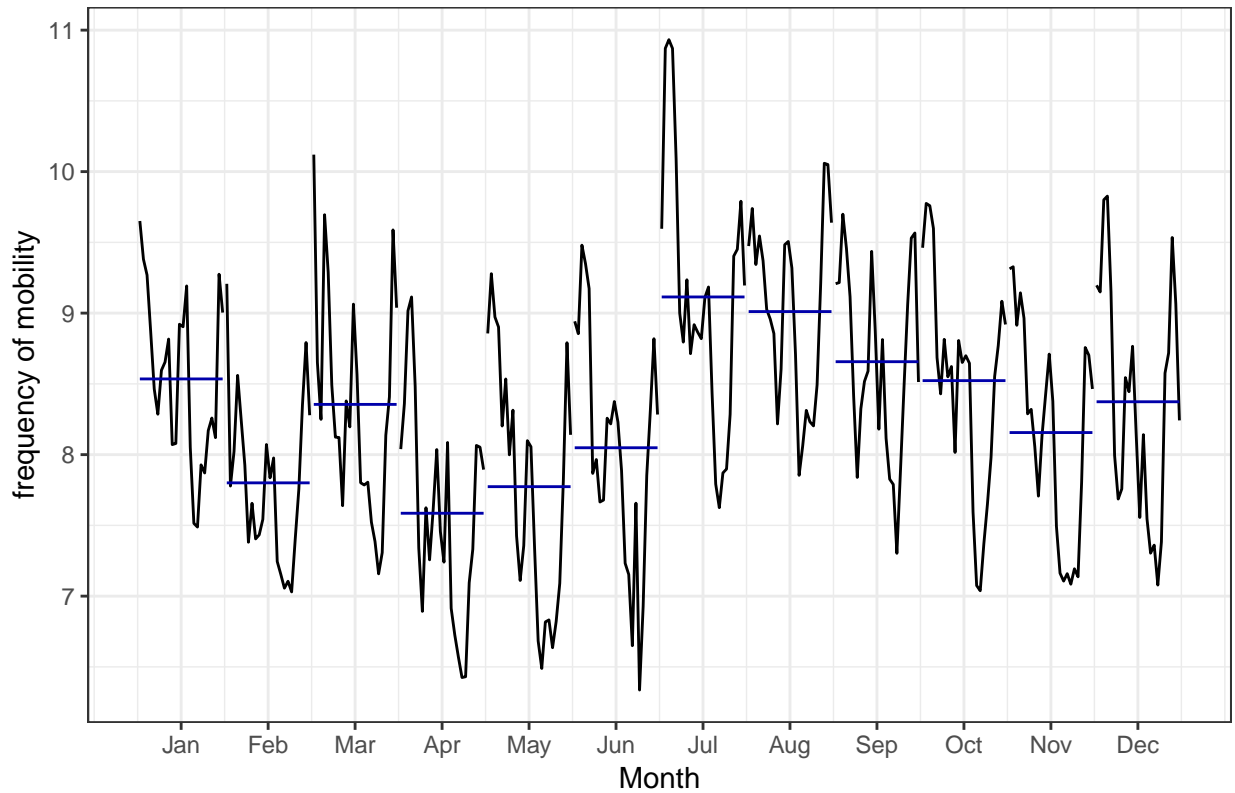


Figure 2: Frequency per month and year

The upper panel shows that there are between 11 and 7 moves per 1000 inhabitants per month. The plot suggests that the frequency of mobility is cyclical, without any clear trend. The time series is clearly non-stationary. The financial crisis, which hit the Dutch housing market hard, can be seen in the dip from 2008 onwards. We also see that the recovery from the crisis set in during the year of 2014. The lower panel, showing the seasonal differences, shows that the year on year changes are negative from January 2009 with a careful recovery in 2012 and a new dip in 2013 before a real recovery from 2014.

Interestingly the lower-panel figure reveals that the year on year changes are negative from medio 2017. This is also reported by Statistics Netherlands:² in 2018 there were 5 % fewer moves than in 2017, with the decline concentrated primarily among people younger than 50 years old. This reflects thus an end of an increasing trend from 2014 onwards. It is precisely this apparent trend-change that we use to motivate this study.

In addition to the cycle, there are strong seasonal patterns that seem to vary over time. Figure 2 shows the distribution of the frequency of mobility per month, with the year of each observation indicated with colours (upper panel), and a the movement per year and average per month (lower panel). The figure shows that the highest frequency is, on average, in August and July while the lowest is in April. We also see there is quite some variation between the years: the differences between the months were more pronounced in the early part of the time series than in later years. The lower panel suggests a sinusoidal pattern with two peaks within one year - one peak in the summer and one at the end of the year.

Finally, we check whether there is a linear relationship between lagged variables of the time series (autocorrelation). Figure 3 reveals a large and positive autocorrelation for small lags, since observations nearby in time

²<https://www.cbs.nl/nl-nl/nieuws/2018/26/minder-mensen-verhuisd-in-eerste-kwartaal-2018>

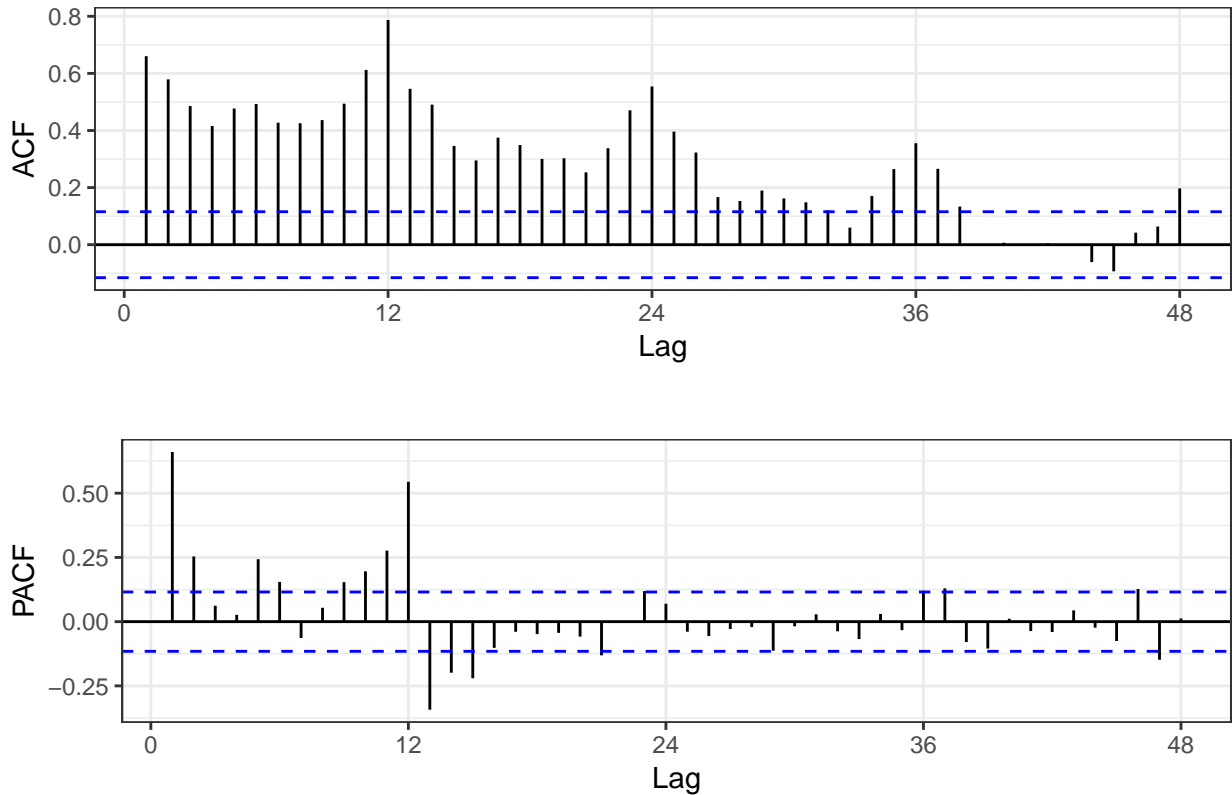


Figure 3: Autocorrelation function the frequency of mobility

tend to be similar in size. We also see that strong autocorrelation in lags that are multiples of the seasonal frequency (12, 24, and 36), which is due to the seasonality discussed above. The slow decline is related to the trend-cycle while the ‘scalped’ pattern is related to the seasonality. The significant spikes at the first and second lag in the partial autocorrelation plot in the lower figure suggests that we should include at least two autoregressive terms.

Summing up, the decomposition of the time series suggests that there is no linear trend, rather a cycle with peaks roughly every 10 years. There are furthermore strong seasonal effects with variations between the years, meaning the seasonality contains noise. Furthermore, the seasonality and the trend-cycle lead to autocorrelation. The next section shows how these elements are addressed in the state space model.

3 Model description

As mentioned in the Introduction, state space models are modular in the sense that the components discussed in the previous section can be included as separate components in the model. The essence of a state space model is that an unobserved dynamic process θ_t governs the observed y_t which is a noisy measurement of the process. Forecasting and inference is done with the recursive Kalman filter algorithm. Residuals are calculated using the one-step ahead predictions from the Kalman filter. There are many ways of describing state space models - the model is implemented in *R* using the *dlm* package, therefore the description in this section follows the notation used in Petris, Petrone, and Campagnoli (2009).

Let y_t denote the logarithm of the frequency of mobility in month t , where y_t are the observed values

of underlying (unobserved) vector of states θ_t . The DLM assumptions are that errors are Gaussian and independent. We can write this DLM as:

$$\begin{aligned} y_t &= F\theta_t + v_t & v_t &\sim N(0, V_t) \\ \theta_t &= G\theta_{t-1} + w_t & w_t &\sim N(0, W_t) \\ \theta_0 &\sim N(m_0, C_0) \end{aligned}$$

In this DLM the observed frequency of migration, y_t , are conditionally independent given the state θ_t : $P(y_t|\bar{\theta}_t)$. The state thus represents a latent Markov process, mean that the probability of moving to the next state depends only on the previous state: $P(\theta_t|\bar{\theta}_{t-1})$. The first equation is called the *observation equation* and the second the *state equation*. v_t and w_t are uncorrelated Gaussian errors, where the observation variances are gathered in the $m \times m$ matrix V_t and system variances in the $p \times p$ matrix W_t . F and G are known system matrices of dimension $p \times m$ and $p \times p$ respectively. The initial θ_0 are normally distributed with means m_0 and variances C_0 .

Since the frequency of mobility is clearly non-stationary, we use as a starting point the linear growth or local linear trend model, defined as follows:

$$\begin{aligned} y_t &= \mu_t + v_t & v_t &\sim N(0, \sigma_v^2) \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + w_{t,1} & w_{t,1} &\sim N(0, \sigma_\mu^2) \\ \beta_t &= \beta_{t-1} + w_{t,2} & w_{t,2} &\sim N(0, \sigma_\beta^2) \end{aligned}$$

The local linear trend model is a DLM with $\theta_t = \mu_t$, where y_t is modelled as a noisy observation of the level μ_t . μ_t has a time-varying slope, governed by β_t . As is seen in this equation, we assume that the frequency of mobility is a non-stationary process growing with β_{t-1} . In line with the notation in Petris, Petrone, and Campagnoli (2009) as:

$$\theta_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} \quad G_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad F_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad W = \begin{bmatrix} \sigma_\mu^2 & 0 \\ 0 & 0 \end{bmatrix}$$

Seasonality is dealt with by including a trigonometric seasonal components to the model. One big advantage with the trigonometric specification, relative to simpler seasonal dummies, is that we allow autocorrelation to last through more lags, resulting in a smoother seasonal pattern. This way we filter out some of the noise seen in the figures in the previous section. By including non-zero variances, the seasonal pattern is allowed to change trough time. Seasonality is incorporated in the DLM model by extending the state equation with an underlying periodic process $g(t)$

$$\mu_t = \mu_t + g(t) + v_t v_t \sim N(0, \sigma_v^2)$$

where g_t defined as the sum of j harmonics. For the j th harmonic, we can write the evolution of the seasonal effects as:

$$\begin{aligned} S_j(t+1) &= S_j(t) \cos \omega_j + S_j^*(t) \sin \omega_j \\ S_j^*(t+1) &= -S_j(t) \sin \omega_j + S_j^*(t) \cos \omega_j \end{aligned}$$

with $\omega_j = \frac{2\pi t j}{s}$. We include two harmonics, meaning that the trigonometric specification can be written as:

$$\theta_t = \begin{bmatrix} S_1(t) \\ S_1^*(t) \\ S_2(t) \\ S_2^*(t) \end{bmatrix} \quad G = \begin{bmatrix} \cos \omega_1 & \sin \omega_1 & 0 & 0 \\ -\sin \omega_1 & \cos \omega_1 & 0 & 0 \\ 0 & 0 & \cos \omega_2 & \sin \omega_2 \\ 0 & 0 & -\sin \omega_2 & \cos \omega_2 \end{bmatrix} \quad F = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix} \quad W = \begin{bmatrix} \sigma_{S_1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{S_1^*}^2 & 0 & 0 \\ 0 & 0 & \sigma_{S_2}^2 & 0 \\ 0 & 0 & 0 & \sigma_{S_2^*}^2 \end{bmatrix}$$

Finally, residual autocorrelation not accounted for by the trend and seasonal components is dealt with by the introduction of an autoregressive elements into the model. The PACF from Figure 3 suggested that an AR(2) model, of the following form, might be sufficient:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + u_t \quad u \sim N(0, \sigma_u^2)$$

In order to make sure that observations Y_t are independent of past values of itself, an AR(2) model is written as a DLM as follows:

$$\theta_t = \begin{bmatrix} \mu_{1,t} \\ \mu_{2,t} \end{bmatrix} \quad G = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 1 \end{bmatrix} \quad F = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad W = \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & 0 \end{bmatrix}$$

The resulting composite model (trend, seasonal and autoregressive) has a state vector μ_t of length 8. The model contains a number of unknown parameters that need to be estimated. Besides the observation variance σ_v^2 , the matrix W contains 6 potentially non-zero variances. We choose the integrated random walk formulation for the trend-cycle component, meaning that we let $\sigma_\beta^2 = 0$. In addition, a residual analysis revealed that a model where only the variance of the first harmonic is zero performs best. Finally, the G matrix contains two parameters for the AR(2) part that need to be estimated. These parameters are estimated subject to stationarity restrictions:

$$-1 < \phi_2 < 1, \phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1$$

Consequently there are in total 6 unknown parameters in the model: σ_v^2 , σ_μ^2 , $\sigma_{S_2}^2$, σ_u^2 , ϕ_1 and ϕ_2 . These parameters are estimated using maximum likelihood (see Petris, Petrone, and Campagnoli 2009, ch. 4).

A residual analysis of the model described so far revealed that there was still some autocorrelation not captured by the seasonal effects; namely in lags 7 and 12. These effects can be accounted for by including addition autoregressive terms to the state equation. This extension of the model will obviously improve the within-sample fit of the model, however it is not a priori not clear that the forecasting accuracy is also improved. In addition to the model described above, hereafter referred to as *DLM1*, we therefore also estimate a model (*DLM2*) with 4 autoregressive terms. This model has two additional parameters that need to be estimated: ϕ_7 and ϕ_{12} .

After determining the values of unknown parameters we use the Kalman filter to obtain the filtered distribution of θ_t conditional on the observed series up to t : $\theta_t \bar{y}_1, y_2, \dots, y_t$. In addition we can estimate a smoothing distribution of the components, representing the past values of the states given all observed values $\theta_t \bar{y}_1, y_2, \dots, y_s$ where $s \geq t$.

4 Estimated model

The estimated parameters are shown in Table 1. We see that the observation variance σ_v^2 is very close to zero, indicating a high level of precision on the observation. Note however, that the variances refer to the log of the frequency of mobility. The small size of the observation variance relative to the system variance results in a fairly large signal-to-noise ratio W/V , meaning again that a relatively large weight is put on the forecast in the Kalman algorithm. We furthermore see that both of the AR parameters are negative, with the absolute value of the parameter for the first lag larger than that of the second lag.

The one-step ahead predictions from the Kalman filter, or filtered values, from both of the models are displayed in Figure 4, together with the (within-sample) Root Mean Square Error (RMSE) of the predictions. Note that we have taken the anti-log of the filtered values, since this is our primary interest here. As is seen in the figure, in the first time periods the volatility of the predicted time-series exceeds that of the observed values, and the filtered values drift quite far off the observed values. This reflects the *adaptive* nature of the

Table 1: Estimated variances and AR parameters

par	d1m1	d1m2
σ_v^2	0.00000002	0.00000002
σ_μ^2	0.00018418	0.00015379
$\sigma_{S_2}^2$	0.00000552	0.00000753
σ_u^2	0.00132805	0.00099461
ϕ_1	-0.51527247	-0.48772863
ϕ_2	-0.32662361	-0.20892364
ϕ_7		-0.18020803
ϕ_{12}		0.40300435

Kalman algorithm, and when calculating the RMSE the two first years are discarded. It is also clear from the figure that DLM2, with the two higher-order autoregressive terms, performs better than the DLM1.

Figure 5 shows the autocorrelation function of the residuals from the filtered values shown in Figure 4. As is readily seen in the left panel, the residuals from DLM1 exhibits autocorrelation in the 7th and 12th lag not captured by the seasonal component. We see that the inclusion of the two additional autoregressive terms (right panel) effectively accounts for the residual autocorrelation in DLM1.

The smoothing estimates for the trend-cycle component and for the seasonal effects are shown in Figure 6. From the upper panel we can pin-point the recent peak in the trend-cycle to March 2017. The lower panel shows how the seasonal effects vary over time; at the start of the time series the within-year cycle exhibits one pronounced maximum (July) and one minimum (March). However, from around the middle of the series, there is gradually another local maximum within each year (January). This means that the seasonal patterns in the model captures well the development of the seasonal factors from the raw data.

5 Cross validation: evaluation over a rolling forecasting origin

In this section we present the results of the evaluation of forecasting accuracy of the two models described in the previous sections. Evaluating forecasting accuracy is thus akin to common out-of-sample testing, where a model is estimated on a training data set and evaluated on an independent test data set. In order to control for effects arising from the composition of the training data, we carry out the evaluation using a cross validation approach (see e.g. Hastie, Tibshirani, and Friedman 2009). Due to the serial autocorrelation in time-series, we can not partition the data randomly. More precisely, in time-series cross validation the test data can not contain observations occurring prior to the observations in the training data. The approach chosen here is the evaluation over a *rolling forecasting origin* (Tashman 2000), where we successively extend the training data and produce forecasts. The last point of the training data in each iteration is referred to as the *origin* T , since the origin on which the forecast is based rolls forward in time. From the origin forecasts are generated for time periods $T + 1, T + 2, \dots, T + N$.

This procedure is shown in Figure 7 (Hyndman and Athanasopoulos 2018). The movement along the x-axis illustrates how the origins (blue dots with red border) are “rolling” forward in time, where the model is estimated on the training data (blue and red dots) and evaluated on the test data (grey dots). The model performance is then evaluated by averaging the prediction errors over the forecast horizon (grey dots with black border).

We evaluate forecasting performance on three time horizons - $N = 6$, $N = 12$ and $N = 18$ - meaning that, for each origing, we calculate three indicators of accuracy; one for a half year forecast, one for a year and one for one and a half year. In order to assess the performance around the changepoint of the trend-cycle in March 2017 from Figure 6, we apply the following strategy: firstly, since we forecast 18 months ahead and the last observation is on December 2018, the last possible origin is June 2017. Secondly, we make sure the first origin is far enough away from March 2017 that the models do not pick up the changepoint, but close

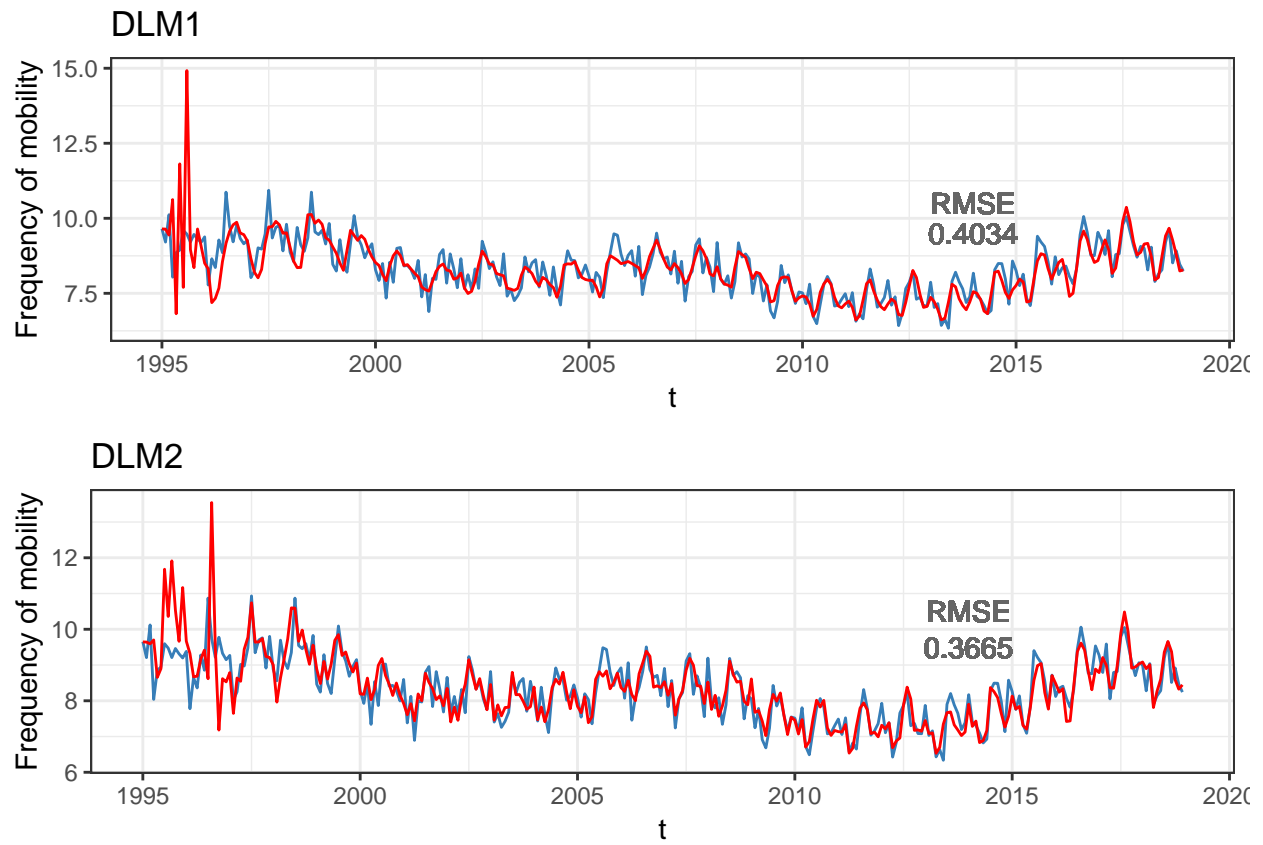


Figure 4: One-step-ahead predictions (blue line), observed data (red) and RMSE (calculated from January 1998)

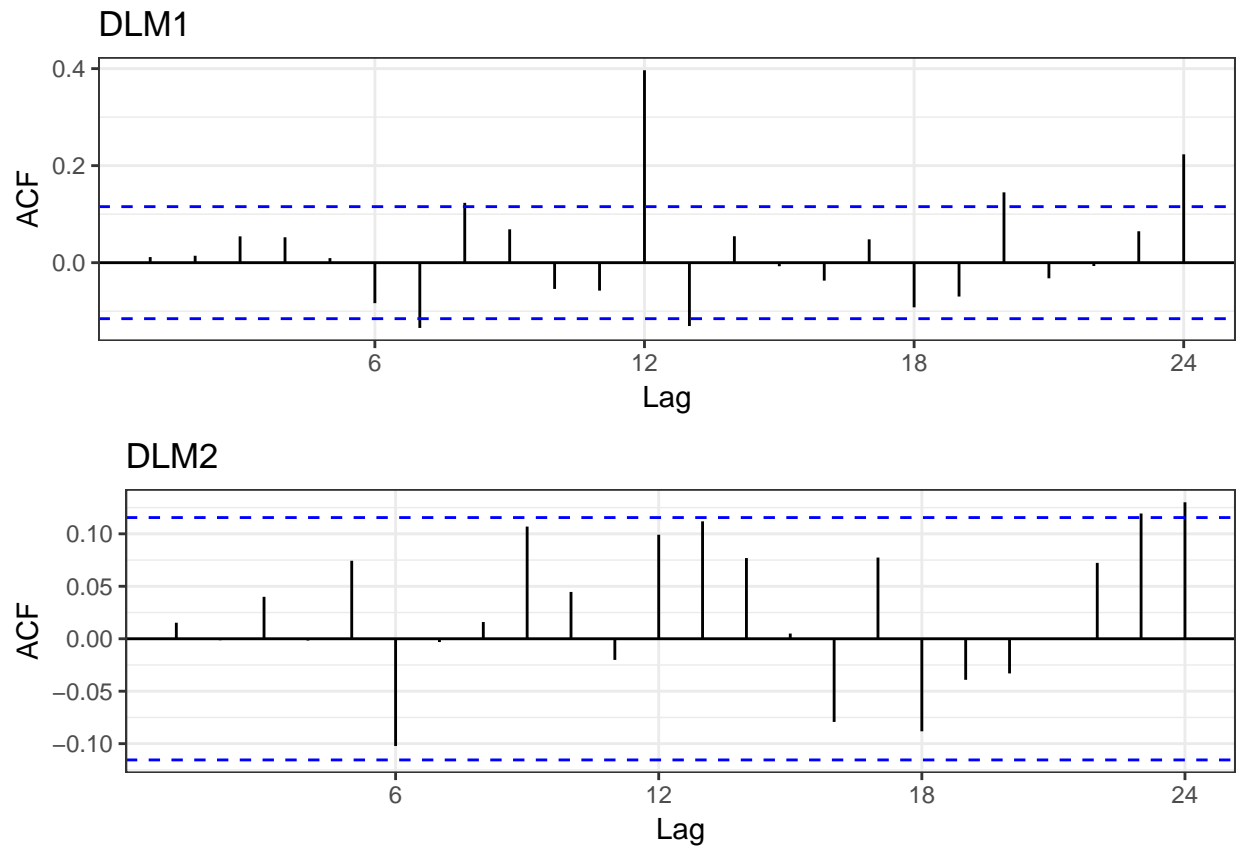


Figure 5: Autocorrelation function for the residuals

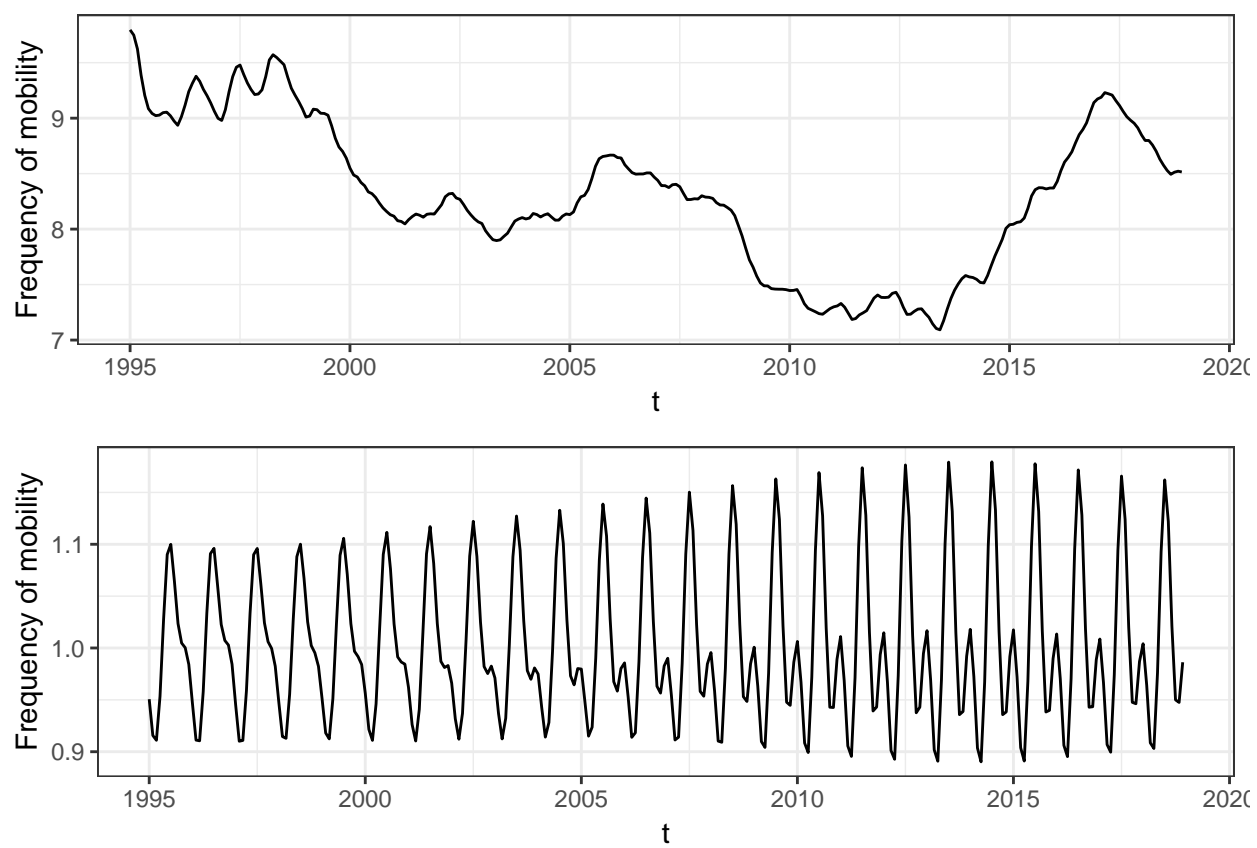


Figure 6: Smoothing estimate of the trend-cycle and seasonal effects

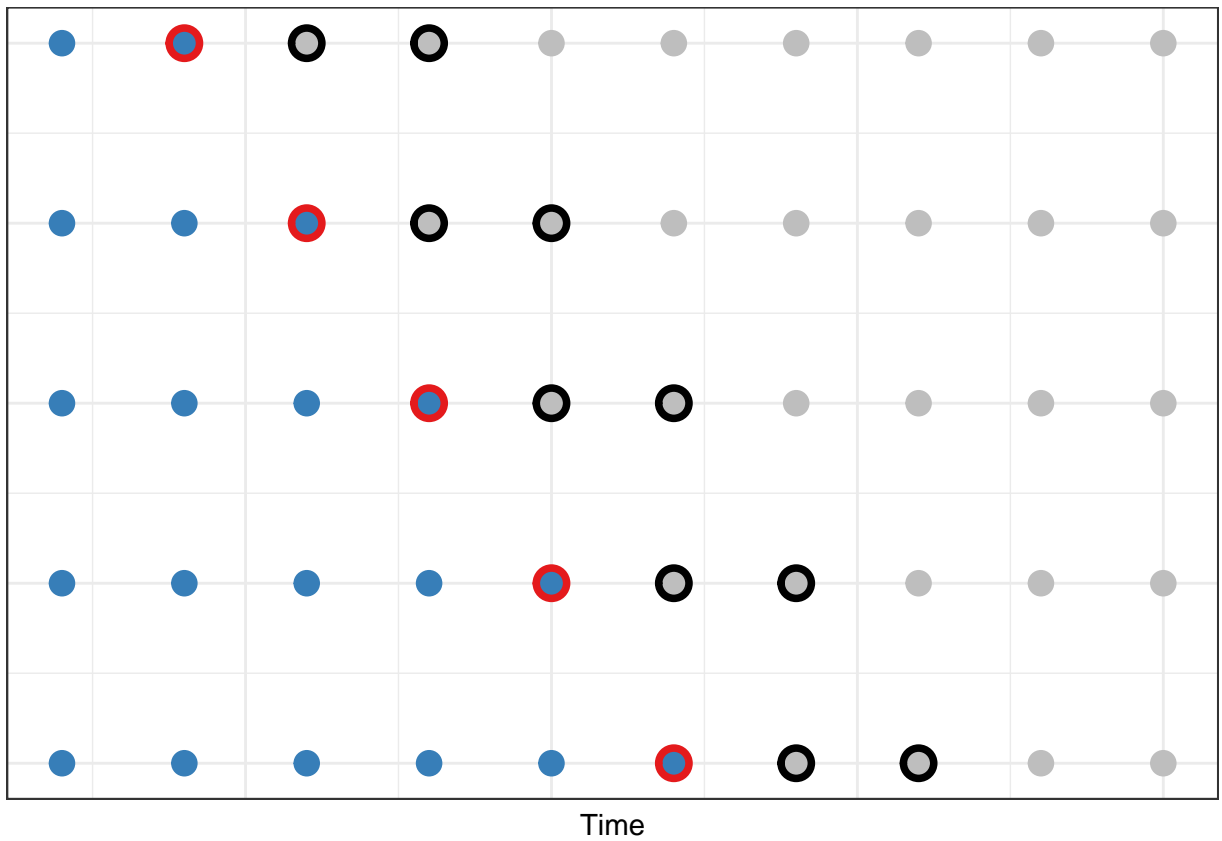


Figure 7: Illustration of rolling forecasting origin: training data (blue), origin (red), forecast horizon (black) and test data (grey)

Table 2: Mean and standard deviation of the estimated parameters

Parameter	DLM1 Mean	DLM1 SD	DLM2 Mean	DLM2 SD
σ_v^2	0.00000002	0	0.00000002	0
σ_μ^2	0.00018418	0	0.0001538	0
$\sigma_{S_2}^2$	0.00000552	0	0.00000754	0
σ_u^2	0.00132805	0	0.00099469	0.00000002
ϕ_1	-0.51527277	0.0000001	-0.48765618	0.000005
ϕ_2	-0.32662374	0.00000008	-0.20883421	0.00000144
ϕ_7			-0.18023308	0.00000311
ϕ_{12}			0.40306025	0.000001

enough so that the forecast of the changepoint is evaluated. Although several options are available, we opt for January 2016, yielding in total 18 forecasting origins: $T = \text{January 2016, February 2016, } \dots, \text{ July 2017}$.

In order to ensure independence between the test and training sample, we recalibrate the DLM models before each origin forecast. The resulting mean and standard deviations, shown in Table 2, reveal that the choice of training data has very limited effect on the estimated parameters. In fact, the estimated parameters are more sensitive to the initial values for the maximum likelihood estimation than to the amount of data. This gives us confidence that there is enough data to estimate the DLM, even on the first origin.

Performance of the DLMs is assessed by comparing forecasting accuracy with that of three other popular models. The first of these is a naïve seasonal model, where the forecast of a specific month is simply the value of the same month of the previous year. Despite its simplicity this model is useful with data with no clear trends. Since the model assumes no trend, we expect this model to perform particularly well around the changepoint.

The next model is the Holt-Winters method with multiplicative seasonal components. This model belongs to the family of exponential smoothing model, where forecasts are weighted averages of past observations, with the weights decaying exponentially for observations further back in time. The smoothing parameters of the model are selected automatically, for each origin, with the default initial values using the function `HoltWinters()` in base *R*. This model type performs particularly well on data with a clear trend, and we therefore expect it to forecast accurately before the changepoint and increasingly worse around and after it.

The final model used for comparison is as seasonal ARIMA model. Identification of the model is carried out using the `auto.arima()` function on the whole sample (Hyndman and Khandakar 2008), and results in an ARIMA(2,1,5)(2,1,1). This is thus a model with double differencing, with yearly and monthly MA(2), and yearly AR(5) and monthly AR(2). Since any ARIMA model can be written as a DLM, differences in performance is mainly due to order selection and treatment of seasonality. This means that we primarily investigate how the manual selection of the DLMs compares to the automatical selection of the ARIMA models.

Table 3 shows the overall forecast accuracy across all origins for all models. Forecasting accuracy is here calculated as the mean RMSE of each of forecasting horizon (6 months, 12 months, 18 months). In order to asses the stability of the accuracy, the table also shows the standard deviation of the RMSE.

From the table we see that the mean RMSE of the DLM1 and DLM2 is quite similar across all forecast horizons, however with a slight advantage of DLM2. Not suprisingly, accuracy deteriorates as N increases, although not dramatically so: in the case of DLM2, the mean RMSE for $N=6$ is 0.43, compared to 0.56 for $N=18$. Compared to both the ARIMA and the Holt-Winter model, the limited deterioration in performance is remarkable: both of the DLMs perform substantially worse than the ARIMA and the Holt-Winters for the six-months forecast; the yearly forecast accuracy is similar across all models, and in the case of the 18 months forecast both of the DLMs outperform the others. For the 18 months forecast horizon the standard deviations of the DLMs are also substantially lower than those of the ARIMA and Holt-Winters models, providing further evidence of this result. The mean RMSEs of the naïve model is remarkably similar across

Table 3: Mean and standard deviation of the RMSE for three forecasting horizons

Origin	ARIMA	DLM1	DLM2	Holt-Winters	Naïve
Mean N=6	0.3841	0.4390	0.4279	0.3641	0.6317
SD N=6	0.0777	0.1338	0.1029	0.1696	0.1848
Mean N=12	0.4924	0.5344	0.5061	0.5077	0.6191
SD N=12	0.1275	0.1399	0.1242	0.2395	0.1323
Mean N=18	0.5988	0.5795	0.5618	0.7227	0.6380
SD N=18	0.2428	0.1425	0.1263	0.3829	0.1474

all forecast horizons, and it is, in general, the least accurate model.

Figure 8 delves further into the forecasting performance, showing the RMSE per origin across all models and forecast horizons. Starting with the shortest forecast horizon, we see that both the ARIMA and Holt-Winters outperform the DLMs until January 2017. In particular we see that the latter performs very well for the first origins, where the 6 first months of the test data follows an upward trend. However, as the origin moves forward and the test data gradually consists of more data after the changepoint, we see the performance of the Holt-Winters model deteriorating. In contrast, the accuracy of the DLMs increases from the origin November 2016 onwards. A similar development is seen for the naïve model, where we observe substantial prediction errors until the same origin, and a decline therefore.

We observe a similar pattern for the yearly forecast horizon, where the Holt-Winters and ARIMA both outperform the DLMs for the first origins. However, here this is reversed two months earlier and we see performance of the Holt-Winters model deteriorating at an earlier point than for a shorter forecast horizon. For the 18 months forecast horizon the DLMs have the lowest RMSE already from November 2016, being overtaken by the Naïve forecast in March 2017. Summing up, while the ARIMA and Holt-Winters models struggle with replicating the declining trend present in the data from March 2017 onwards, the DLMs to some follow the performance of the naïve model which excels when the data does not exhibit any clear trend.

Next we zoom in on the forecasting performance around and after the changepoint. Figure 9 compares the forecasts of all models in the period January 2017 to December 2018. The grey dots represent forecast from different origins, the red lines are the mean forecast across all origins and the blue line is the data. The figure shows clearly the divergence in forecasting performance between the DLMs and the ARIMA and Holt-Winters models. The Holt-Winters model systematically overpredicts the migration frequency already from June 2017, meaning that the forecasted frequency exceeds the data for all origins. For the ARIMA model this occurs from October 2017. In the case of the DLMs the last forecasted point with a value lower than that of the data occurs in May 2018.

6 Conclusion

In this study we have developed a DLM...

- The DLMs outperform, on average, other models on the longer forecast horizon.
- Differences particularly visible in forecasts from March 2017 onwards.
- DLMs appear to offer a middle ground between naïve and ARIMA/Holt-Winters
- Difference DLMs: although DLM2 is clearly better within-sample, difference in terms of forecasting performance is not particularly big.

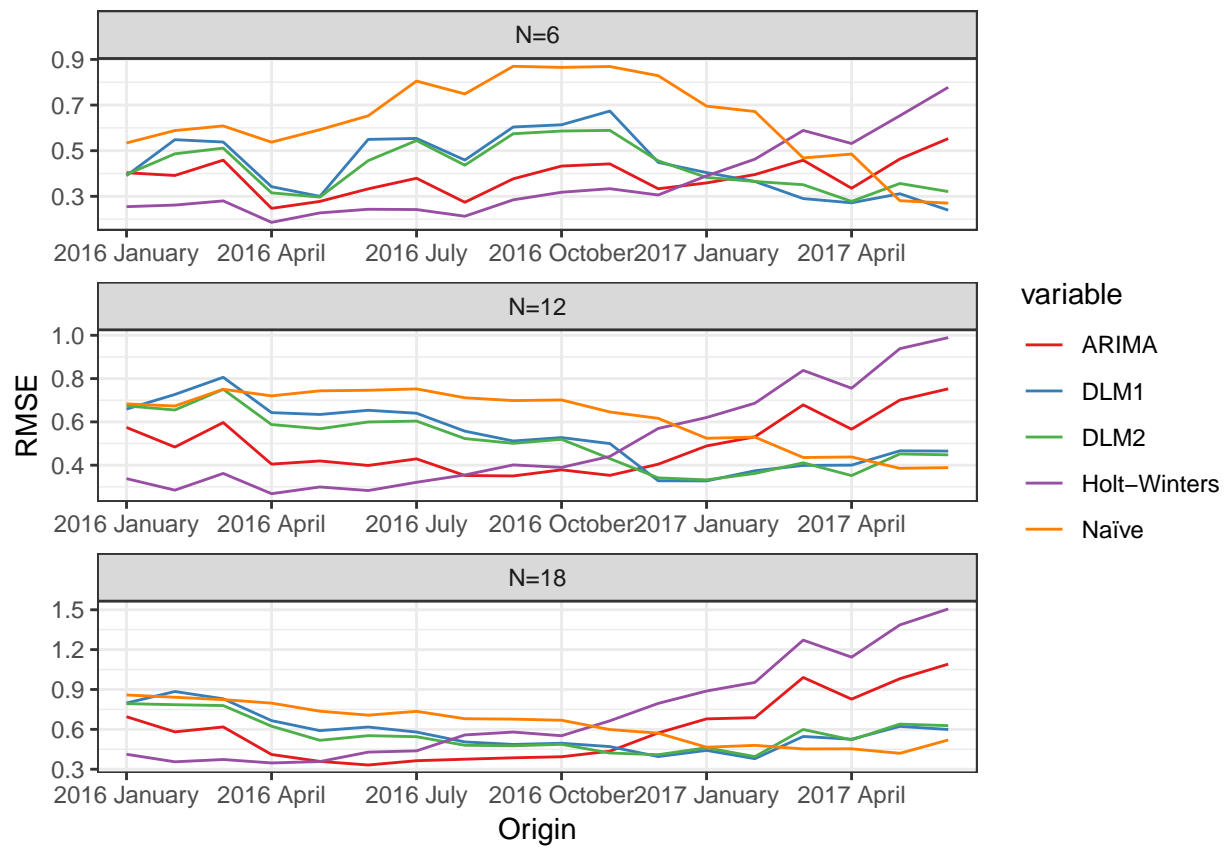


Figure 8: RMSE of each origin for three forecast horizons

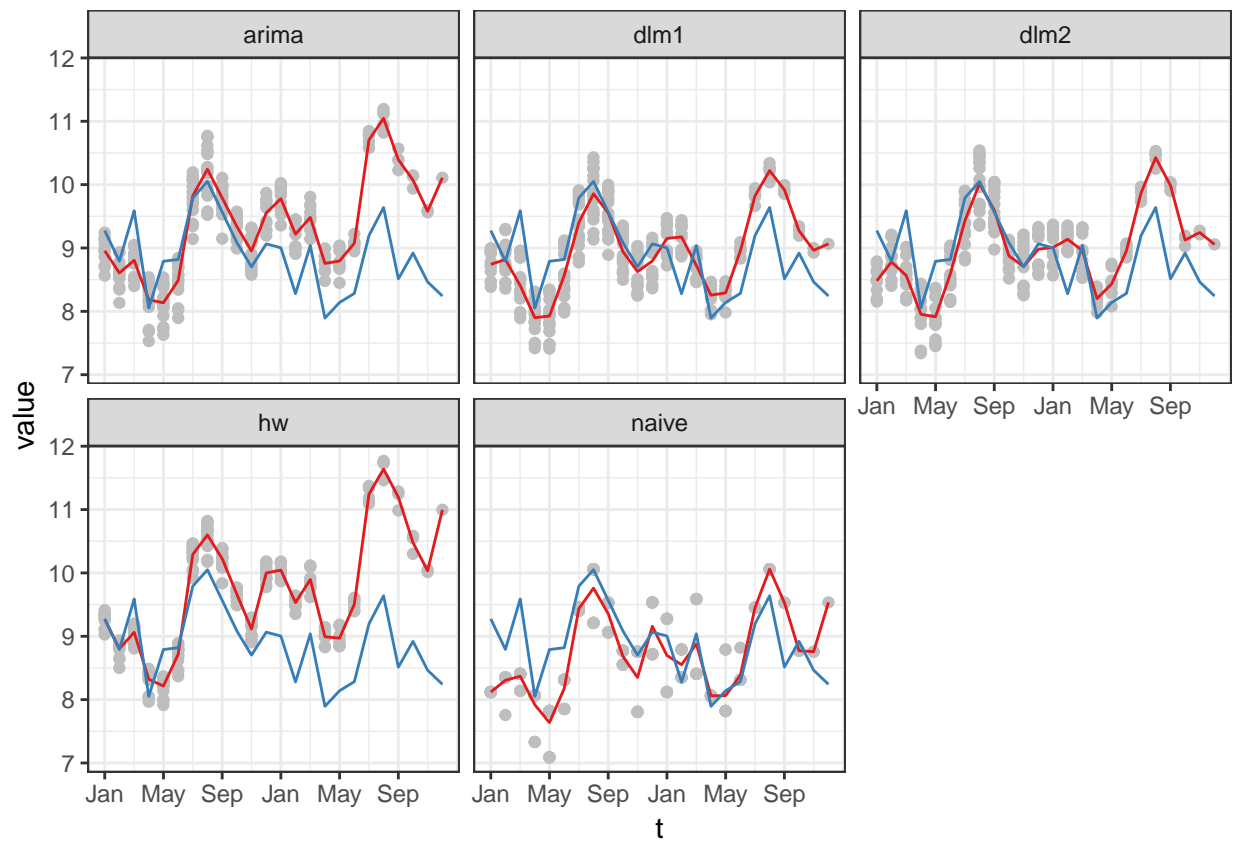


Figure 9: Forecasted and actual frequency of mobility (2017 and 2018)

Bibliography

- Canova, Fabio. 1998. “Detrending and Business Cycle Facts.” *Journal of Monetary Economics* 41 (3). Elsevier: 475–512.
- Durbin, James, and Siem Jan Koopman. 2012. *Time Series Analysis by State Space Methods*. Oxford university press.
- Hamilton, James D. 2018. “Why You Should Never Use the Hodrick-Prescott Filter.” *Review of Economics and Statistics* 100 (5). MIT Press: 831–43.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. “Model Assessment and Selection.” In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 219–59. New York, NY: Springer New York. doi:10.1007/978-0-387-84858-7_7.
- Hyndman, Rob J, and George Athanasopoulos. 2018. *Forecasting: Principles and Practice*. OTexts.
- Hyndman, Rob J, and Yeasmin Khandakar. 2008. “Automatic Time Series Forecasting: The Forecast Package for R.” *Journal of Statistical Software* 26 (3): 1–22. <http://www.jstatsoft.org/article/view/v027i03>.
- Petris, Giovanni, Sonia Petrone, and Patrizia Campagnoli. 2009. *Dynamic Linear Models with R*. Springer.
- Tashman, Leonard J. 2000. “Out-of-Sample Tests of Forecasting Accuracy: An Analysis and Review.” *International Journal of Forecasting* 16 (4): 437–50. doi:[https://doi.org/10.1016/S0169-2070\(00\)00065-0](https://doi.org/10.1016/S0169-2070(00)00065-0).
- Zietz, Joachim, and Anca Traian. 2014. “When Was the Us Housing Downturn Predictable? A Comparison of Univariate Forecasting Methods.” *The Quarterly Review of Economics and Finance* 54 (2). Elsevier: 271–81.