Chapter 4

An assessment of the comparative accuracy of time series forecasts of patent filings: the benefits of disaggregation in space or time

Nigel Meade

Tanaka Business School, Imperial College London, London, UK

Module: E
Software: STAMP, TSP

1 Introduction

This work is concerned with methods for forecasting the filing of patents and was carried out in conjunction with the European Patent Office. The filings data were subdivided by:

- Blocs – European Patent Convention countries, Japan, US and the Rest of the World.
- Industries – main Fields of Technology according to headings A–H of the International Patent Classification (WIPO 2000).

The issues addressed are: the benefits of multivariate models versus univariate ones in exploiting any correlations between the filings in different blocs or industries; the effect of aggregation over time (from monthly to annual data) and the effect of aggregation by bloc or by industry on the accuracy of the forecast of total EPO filings. Two approaches are used: the ARIMA framework and the dynamic linear model (DLM) in both univariate and multivariate modes.

The main results are: monthly data does tend to provide greater accuracy in annual forecasts; there are no significant benefits to be gained by multivariate modelling and no significant benefits are found from aggregating over blocs or industries. There are benefits from using monthly data, rather than annual data. The best modelling approaches are, for monthly data, the univariate dynamic linear model; for annual data either the univariate ARIMA or DLM could be used. The recommended forecasting approach provides a benchmark against which other forecasts drawing on different data sources can be compared.

The filing of an application for a patent is the first step in achieving protection for intellectual property. The three major patent offices are the European (EPO), the Japanese (JPO) and the US (USPTO) offices. The examination of each application is a labour intensive process, involving technological and legal expertise. The motivation for this study of the forecasting of patent filings with the EPO is that the forecasts are a prerequisite for manpower planning at an aggregate scale and at the level of availability of expertise in different technologies. The data used in this study was made available in 2002 as part of an EPO research programme. The programme looks at five different approaches to forecasting patent filings, see Chap. 2. These approaches include: survey methods; micro-level studies of firms' patenting practice; effects of inter-firm competition on patenting behaviour; consideration of the flows of filings between patent offices; time series modelling of aggregate filing data. It is this last topic that is dealt with here.

The levels of aggregation at which patent filings are analysed are:

- Bloc level – filings are from the EPC contracting states, US, Japan and the rest of the world.
- Industry level – nine industry groupings are used.
- Total filings within the EPO.

The objective is to identify the most accurate means of forecasting filings within these categories. This analysis of patent filings, at several levels of aggregation, uses two time series modelling frameworks.

The basic methodologies used are:

- ARIMA model.
- Dynamic linear model.

Three central issues are:

- Is there information in the filings within blocs or industries that would lead to increased forecasting accuracy via the use of multivariate forecasting models?
- Does disaggregation over blocs or industries lead to improved forecasting?
- Does disaggregation over time lead to improved forecasting? Does the use of monthly data, rather than annual data, lead to greater forecasting accuracy?

In other words, is it better to forecast the filings from the blocs (or industries) individually and then consolidate them; or to simply forecast the consolidated figures?

Out of sample forecasting accuracy over a one year ahead horizon is the criterion upon which model performance is judged.

This chapter is divided into the following sections: a description and exploratory analysis of the data; a discussion about the forecasting methods to be used; the application of the forecasting methods to the data; an analysis of forecasting accuracy; forecasting accuracy over a longer horizon followed by a summary and conclusion.

2 Data description

The data provided represent filings via two mechanisms: direct applications to the EPO (euro-direct) and via the Patents Cooperation Treaty (euro-PCT-IP). For further details, see WIPO (2006a).

Data are available on both a monthly and a yearly basis. For the analysis, the data were consolidated into bloc or industry series shown in Table 4.1. The relative proportions of filings in the different world blocs and industry groups at the EPO are shown in Fig. 4.1. Industry is used as shorthand for the main field of technology according to the International Patent Classification, see WIPO (2000).

Table 4.1. The countries / blocs and industries for which data are available

Bloc	Code
European Patent Convention Signatories	EPC
Japan	JP
United States	US
Rest of world	ROW
Total	EPO
Industry	**Code**
Human Necessities	A
Performing Operations/Transport	B
Chemistry; Metallurgy	C
Textiles; Paper	D
Fixed Constructions	E
Mechanical Engineering	F
Physics	G
Electricity	H
Other	other

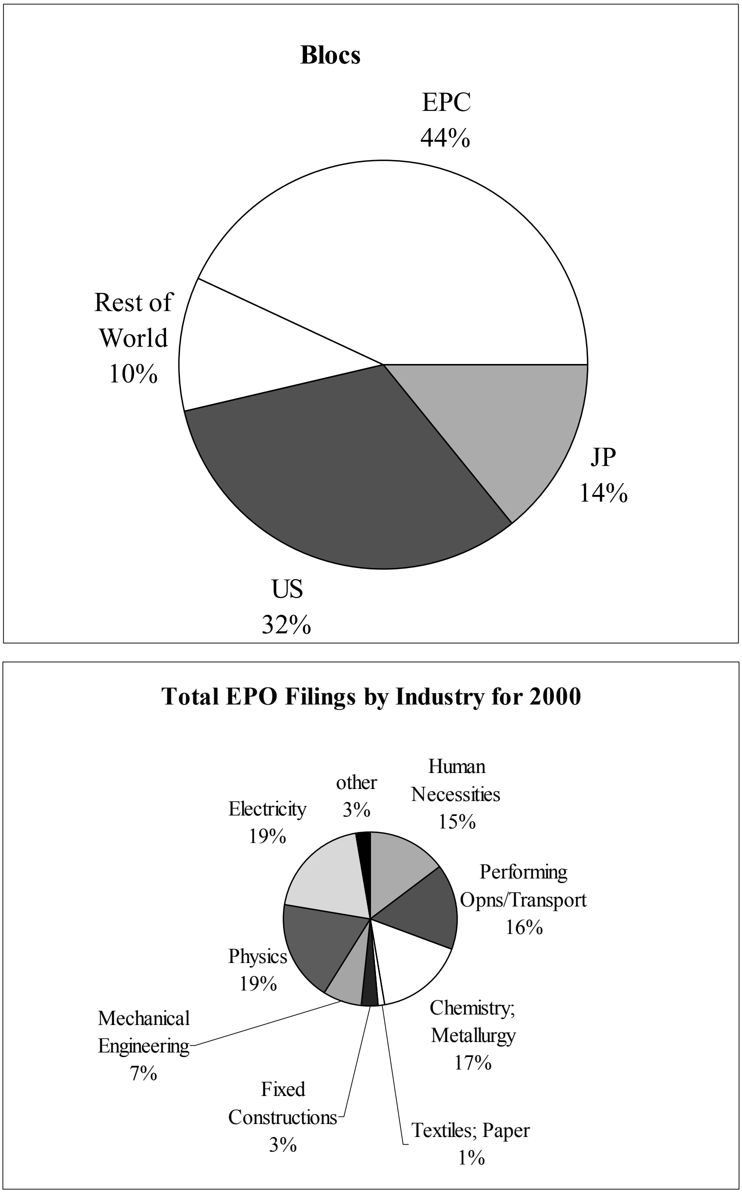

Fig. 4.1. Proportions of filings by bloc and by industry (IPC main fields)

An important issue in this project is the extent to which filings from one bloc or industry affect filings in another bloc or industry as this will impact on the performance of multivariate models. The filings data over the period available has trended upward throughout with the occasional hesitation. The data are graphed in Fig. 4.2, giving annual filings by blocs and total EPO filings; and in Fig. 4.3, giving the filings by industry groupings.

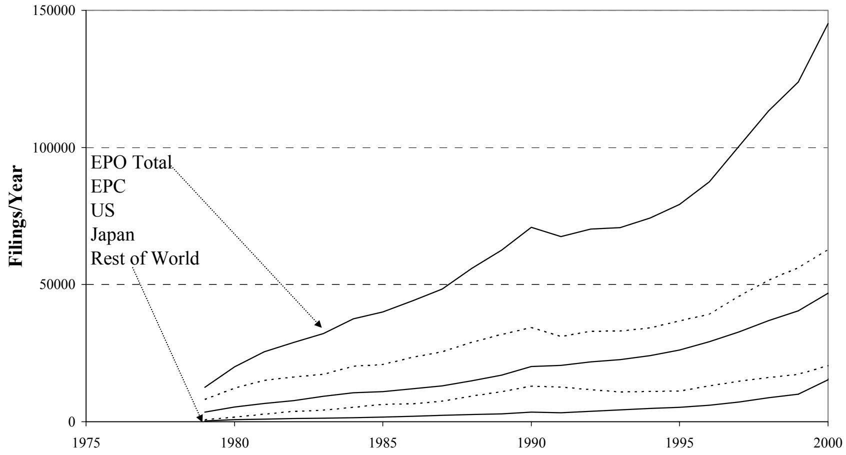

Fig. 4.2. Annual filings for EPO blocs

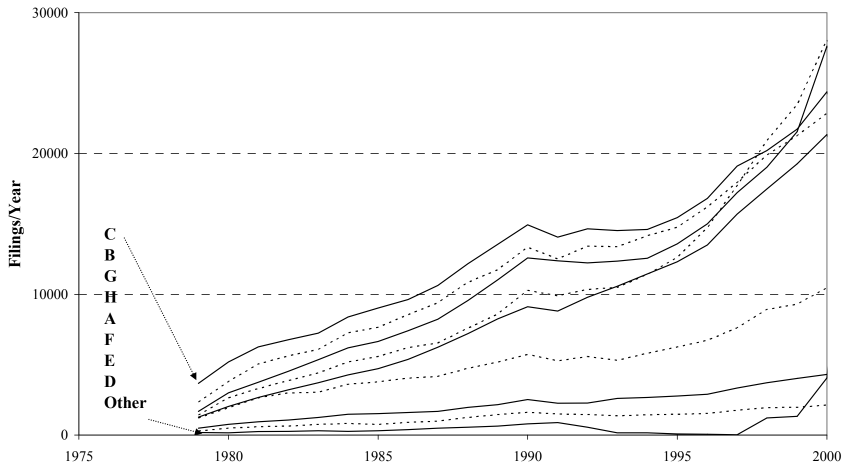

Fig. 4.3. Annual filings by industries

Here we are interested in any evidence of one industry or bloc *leading* another. This is different to the patent family approach, see Hingley and Park (2003), where the progress of filings at different offices is modelled. The strong trend apparent in the data (X_t) means that filings by bloc or industry will be strongly correlated simply because of the common trend. For correlations to be meaningful, the data need to be stationary, that is their mean and variance should not change over time. Thus, it is more informative to look at the correlation between annual changes in filings $(X_t - X_{t-1})$ as these changes are stationary. The correlations are tabulated in Table 4.2.

Table 4.2. Correlations between annual changes in filings using annual data

By Bloc

Correlations between coincident annual changes

	EP	JP	US	ROW
EP	1.00			
JP	0.66	1.00		
US	0.92	0.62	1.00	
ROW	-0.67	-0.22	-0.78	1.00

Inter-bloc correlations are significant, correlations with rest of world and other blocs either not significant or negative

No significant correlations were found at other lags

By Industry

Correlations between coincident annual changes

	A	B	C	D	E	F	G	H	other
A	1.00								
B	0.95	1.00							
C	0.95	0.96	1.00						
D	0.86	0.92	0.91	1.00					
E	0.92	0.93	0.91	0.87	1.00				
F	0.88	0.94	0.91	0.89	0.85	1.00			
G	0.79	0.76	0.85	0.72	0.73	0.80	1.00		
H	0.90	0.88	0.90	0.79	0.83	0.91	0.93	1.00	
other	-0.89	-0.84	-0.82	-0.77	-0.81	-0.70	-0.51	-0.64	1.00

Virtually all correlations significant. No significant correlations were found at other lags

Unfortunately, there was no evidence of a lagged relationship between blocs or industries; that is, there was no evidence that a change in one section preceded change in another. The different filings groups seem to respond to the same stimuli at the same time.

In order to examine the possible presence of a lagged relationship further, the exercise is repeated using monthly data, this extra 'definition' in the data might reveal something hidden in the annual data. In order to achieve stationarity, the seasonal effect is removed by 'seasonal differencing'. The data used are (X_t - $X_{t\text{-}1}$ - $X_{t\text{-}12}$ + $X_{t\text{-}13}$) and the correlations are shown in Table 4.3.

Table 4.3. Correlations between monthly changes in filings using monthly data

By Bloc

Correlations between coincident monthly changes

	EP	JP	US	ROW
EP	1.00			
JP	0.73	1.00		
US	0.66	0.64	1.00	
ROW	0.02	0.21	0.00	1.00

Correlations with rest of world and other blocs either not significant or negative. No significant correlations were found at other lags

By Industry

Correlations between coincident monthly changes

	A	B	C	D	E	F	G	H	other
A	1.00								
B	0.79	1.00							
C	0.83	0.84	1.00						
D	0.62	0.68	0.64	1.00					
E	0.61	0.74	0.63	0.49	1.00				
F	0.81	0.83	0.82	0.64	0.68	1.00			
G	0.82	0.88	0.87	0.65	0.68	0.81	1.00		
H	0.79	0.80	0.86	0.67	0.60	0.79	0.90	1.00	
Other	-0.21	-0.08	-0.09	-0.02	-0.13	-0.15	-0.12	-0.10	1.00

Virtually all correlations significant. Several slightly significant correlations were found at other lags

The correlations for monthly changes are typically lower than for the annual data. This is due to the greater stringency that is asked for here, because changes have to happen not just in the same year but in the same month. However, there was still little evidence of lagged effects.

3 Review and description of forecasting methods

The forecasting methods discussed here are extrapolative methods, this means that the information set used for forecasting is the history of the relevant variable or variables. Extrapolation in its widest context is discussed by Armstrong (2001). The three central issues mentioned in the introduction will be discussed one by one here. Firstly, the value of using a multivariate model rather than a set of univariate models will be examined. If there is correlation between the set of series being forecast, then the multivariate model may be expected to capture this extra structure and hence produce more accurate forecasts. Preez and Witt (2003) compared univariate and multivariate ARIMA and state space model based approaches to forecasting international tourism (the dynamic linear model is a particular implementation of a state space model). The data were the numbers of visitors to the Seychelles from four European countries. They found that their multivariate (state space) model was uniformly least accurate and that univariate ARIMA was, on average, most accurate. Their findings demonstrate that the use of multivariate models can involve loss of accuracy, perhaps due to constraints in their structure, as well as the opportunity of increased accuracy.

Secondly, the issue of cross-sectional aggregation will be examined. Aggregation across the blocs of the EPO is a form of spatial aggregation, a topic addressed by Miller (1998). He points out that the econometric literature is mainly concerned with the effects of aggregation on the density functions of parameter estimates rather than forecasting accuracy. In his study of forecasting economic variables, such as unemployment, in regions within a US state, he found no real difference in accuracy between forecasts using disaggregated and aggregate data. Individual industry forecasts are analogous to *bottom up* forecasts in the context of business forecasting, while the alternative is the *top down* forecast where an aggregate forecast is sub-divided. In this context, Schwartzkopf et al (1998) and Dangerfield and Morris (1992) found that disaggregated (bottom up) forecasts tended to be more accurate than aggregated forecasts.

The third issue raised was the effect of temporal aggregation, from the patent filings perspective using annual data when monthly data is available. The issue of time deformation is discussed by Stock (1988). He considers the situation where there is an "operational time scale" rather than a calendar scale. Essentially, events influence the rate of change of the process being studied. Although a time transformation is not considered feasible here, it is possible that the evolution of the patent filing process is not fully summarised by annual data and that monthly data provides more useful information. This argument influenced Funke (1990), who used monthly Vector AutoRegression (VAR) models to "capture the current economic outlook as quickly as possible" when forecasting industrial production in OECD countries. Rossana and Seater (1995) find evidence of a substantial loss of information when temporal aggregation of monthly or quarterly data to annual data is performed. The information lost described low frequency cyclical variation. Perhaps as a consequence of this, they found that the aggregated annual data showed more long run persistence than the underlying higher frequency data.

In order to try and illuminate these issues, two different time series frameworks were used for modelling the data. These are ARIMA modelling and the dynamic linear model. Software for the analysis is Time Series Processor (TSP[1]) and Structural Time series Analyser, Modeller and Predictor (STAMP[2]) respectively. These frameworks will both be used in univariate and multivariate modes to address the first issue. To address the third issue, models will be derived using both monthly and annual data. The second issue of cross-sectional aggregation will be examined by aggregating over blocs and industries.

ARIMA modelling is well described in Box and Jenkins (1976) or Pankratz (1983). For the stationary process, Z_t, the general ARMA process has p autoregressive terms and q moving average terms.

$$Z_t - \phi_1 Z_{t-1} \ldots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} \ldots - \theta_q a_{t-q} \qquad (1)$$

The variable, a_t, is a white noise term where:

$$a_t \sim N\left(0, \sigma^2\right) \quad \text{and} \quad \rho\left(a_t, a_{t-k}\right) = 0 \quad \text{for} \quad k \neq 0$$

We use the conventional notation where B is the backward difference operator, B, where

$$BX_t = X_{t-1} \quad \text{so} \quad B^d X_t = X_{t-d} \quad \text{and} \quad \left(1-B\right)X_t = X_t - X_{t-1}$$

In this context, X_t represents the number of patents filed at time t, and Z_t represents a stationary transformation of X_t, after finding a suitable value

1 See http://www.tspintl.com.

2 See http://stamp-software.com.

for d. The estimation of the coefficients is achieved by maximum likelihood, the necessary first step is the identification of p and q. This identification is achieved by using an information criterion. The criterion balances the marginal value of increasing the number of parameters; say by increasing p by one, against the reduction in the variance of the noise term, a_t. The underlying intuition is that the criterion allows the extraction of the underlying structure of the data without over-fitting the model. In the following analysis, the values of p and q in ARMA(p,q) were chosen by minimising Schwarz's Bayesian Information Criterion (see, for example, Mills, 1999). The criterion is:

$$BIC = \ln\left(\hat{\sigma}^2\right) + m\frac{\ln(T)}{T} \tag{2}$$

where there are T observations used to calculate the standard error $\hat{\sigma}$, an estimate of σ where $V(a_t) = \sigma^2$ and $m = p + q$ coefficients estimated.

Monthly data are likely to exhibit seasonality, a pattern that tends to repeat itself every year. The most common ARIMA model for handling monthly data is called the multiplicative model. It is a combination of two models, one representing within year behaviour, with L observations per year, and one representing annual behaviour. This can be denoted by ARIMA(p,d,q)x(P,D,Q), and its equation is:

$$\begin{aligned}&\left(1-B^d\right)\left(1-B^{DL}\right)\left(1-\phi_1 B\ldots-\phi_p B^p\right)\left(1-\Phi_1 B^L\ldots-\Phi_P B^{PL}\right)X_t = \\ &\left(1-\theta_1 B\ldots-\theta_q B^q\right)\left(1-\Theta_1 B^L\ldots-\Theta_Q B^{QL}\right)a_t\end{aligned} \tag{3}$$

The identification and estimation procedure is analogous to that for non-seasonal series.

The multivariate version of ARIMA modelling used here is VAR. The model is:

$$Z_t = \mu + \Phi_1 Z_{t-1} + \ldots \Phi_p Z_{t-p} + \varepsilon_t \tag{4}$$

where Z_t is now a vector of length k of observations of k stationary time series and Φ_i (i = 1,… p) are k x k matrices of coefficients, μ is a vector of constants and ε_t is a vector of white noise terms with zero mean and a non-singular covariance matrix. The vector time series, Z_t, is formed by taking first differences of the raw data. The order of the process, p, is normally determined by likelihood ratio tests or by information criteria. However, in the case of yearly data, shortage of observations dictates the maximum value of p.

The dynamic linear model is founded on the simple hypothesis that a time series can be decomposed into a level, a trend (called a slope in the software used), seasonality (if relevant) and a random component. The model is a 'state space' model in that it is described by an observation equation and several more equations describing the evolution of the states.

The observation equation is:

$$X_t = \mu_t + \varepsilon_t \qquad \varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right) \tag{5}$$

where the level, μ_t, is defined by

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \qquad \eta_t \sim N\left(0, \sigma_\eta^2\right) \tag{6}$$

and the trend, β_t, is defined by

$$\beta_t = \beta_{t-1} + \zeta_t \qquad \zeta_t \sim N\left(0, \sigma_\zeta^2\right) \tag{7}$$

The random changes to level, η_t, and trend, ς_t, and the random disturbances, ε_t, about the level are each independent of each other. As a new observation becomes available, estimates of the current level and trend are revised. The forecast for *k* periods ahead is the current estimate of the level plus *k* times the current estimate of the trend, i.e.

$$E\left(X_{t+k} \mid X_t\right) = \hat{\mu}_t + k\hat{\beta}_t \tag{8}$$

This formulation is also called an unobserved components model. If the noise terms identified earlier, η_t for level and ς_t for trend, are each related linearly to the random disturbance, ε_t, then the dynamic linear model is said to be a single source of error model. In this case, the model described is identical to the ARIMA(0, 2, 2) model. This type of parallel is discussed by Gardner and McKenzie (1988). A dynamic linear model with multiple sources of error as defined above cannot be well represented by the ARIMA structure simply because it only admits one source of error. This is discussed in Meade (2000). A framework of dynamic linear models with a single source of error is given by Hyndman et al (2002). This covers additive trend (the case used here), multiplicative trend and an absence of trend. This framework also includes seasonal components that are modelled as either additive or multiplicative terms.

STAMP, the dynamic linear modelling software, offers several options for modelling seasonal variation. Seasonal factors can be fixed constants (estimated from the data) or they can be modelled as stochastic trigonometric factors. The former alternative is equivalent to the additive seasonal factor in Hyndman et al's framework of state space models. However, the latter alternative was used for this analysis, because a model of the whole seasonal profile is used rather than a set of disjoint factors. The seasonal variation within a year is represented as a weighted sum of sine and cosine waves (the fitting procedure is a form of Fourier analysis). The additive seasonal factor for month *m* (=1, 12) is represented below.

$$\text{Seasonal factor}_m = \sum_k \beta_{S,k} \sin\left(\frac{2\pi mk}{12}\right) + \sum_k \beta_{C,k} \cos\left(\frac{2\pi mk}{12}\right) \quad (9)$$

The Fourier series contains as many terms as necessary to capture the seasonal profile, subject to there being no more terms than months. In other words, this representation is usually more parsimonious (requires fewer estimates) than individual monthly factors.

The multivariate version of the dynamic linear model is a form of the seemingly unrelated time series equations (SUTSE) model. SUTSE implies a set of linear models linked only by their disturbances, which may be correlated. If the linear disturbances are uncorrelated, then the procedure is equivalent to separate linear models. The greater the correlation, the greater the efficiency gain in estimation using this approach. Note that multivariate dynamic linear model implies SUTSE, not vice versa.

SUTSE is a generalisation of (5,6 and 7) where vectors replace the dependent variable and the level and trend parameters. The variances of univariate disturbances, $\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\varsigma^2$ become *k* x *k* covariance matrices, $\mathbf{\Sigma}_\varepsilon, \mathbf{\Sigma}_\eta, \mathbf{\Sigma}_\varsigma$. A homogeneous model is used, where the covariance matrices of the disturbances are assumed proportional to one another. That is $\mathbf{\Sigma}_\eta = q_\eta \mathbf{\Sigma}_\varepsilon$ and $\mathbf{\Sigma}_\varsigma = q_\varsigma \mathbf{\Sigma}_\varepsilon$ where q_η and q_ς are scalar parameters.

4 Application of forecasting methods

For the analysis, all the data series are split into two regions. In the estimation region, the data are used for model identification and parameter estimation only. The remaining data are used for out of sample forecasting. Since forecasting is done recursively on an annual basis, after a year has been forecast, that year's data is then included for parameter estimation for the forecast of the following year. The estimation and forecast regions are defined in Table 4.4.

Table 4.4. Definition of estimation and forecast regions

		Yearly	Monthly
Total availability	Start date	1979	1978 June
	Finish date	2001	2001 Dec
Number of observations		23	283
Data used only for estimation	Start date	1979	1978 June
	Finish date	1994	1994 Dec
Number of observations		16	199
Data used for out of sample forecasting and subsequent estimation	Start date	1995	1995 June
	Finish date	2000	2000 Dec
Number of observations		6	72

Although data were available for 2001, their quality was suspect and these data were not used in the evaluations of forecast accuracy.

Once the model has been identified for each series, the forecasting process is carried out recursively. Forecasts from the model estimated using data in the estimation region are prepared for up to a year ahead. This additional data is then used for parameter re-estimation; forecasts are prepared for a year ahead and so on. In summary, forecasts with a maximum horizon of one year are prepared sequentially for six origins at annual intervals.

These sets of forecasts provide accuracy information on a multi-origin and multi-horizon basis, as recommended by Fildes (1992). The measures of accuracy used are root mean square error (rmse) and mean absolute percentage error (mape). Both of these measures are calculated over different origins $i = 1,\cdots, I$, over horizons $h = 1,\cdots, H$, (for monthly data only in this section, multiple annual horizons are examined in Sect. 4) and overall. If the forecast origin is $T + (i - 1)M$ (where: for annual data $T = 16$, $M = 1$, $I = 5$ and for monthly data $T = 199$, $M = 12$, $I = 5$); then the h step ahead forecast $\hat{Y}_{T+(i-1)M,h}$ is an estimate of the observation $Y_{T+(i-1)M+h}$ and the error is $e_{i,h} = Y_{T+(i-1)M+h} - \hat{Y}_{T+(i-1)M,h}$. The percentage error is

$$p_{i,h} = \frac{100 e_{i,h}}{Y_{T+(i-1)M+h}} .$$

The measures of accuracy are given in Table 4.5.

Table 4.5. Definitions of accuracy measures used

root mean square error (rmse)	
For origin i	$RMSE_{Origin}(i) = \sqrt{\frac{\sum_{h=1}^{M} e_{i,h}^2}{M}}$
For horizon h	$RMSE_H(h) = \sqrt{\frac{\sum_{i=1}^{I} e_{i,H}^2}{I}}$
Overall O	$RMSE_O = \sqrt{\frac{\sum_{h=1}^{M}\sum_{i=1}^{I} e_{i,h}^2}{I\,M}}$
mean absolute percentage error (mape)	
Overall O	$MAPE_O = \frac{\sum_{h=1}^{M}\sum_{i=1}^{I} \lvert p_{i,h} \rvert}{I\ \ M}$

As a general rule, the rmse is reported as the prime measure of accuracy and mape is used as a supporting figure, often for comparison between series.

Accuracy over a horizon of one year is used as the main criterion for comparison between models for two reasons. Firstly, a year is the shortest common horizon for the annual and monthly data sets. Secondly, the coefficients of both modelling procedures are estimated to best describe the relationship between an observation at time *t* and those observations available at time *t-1*. Each model has the objective of minimising the one period ahead error variance.

The focus in this study is on a one-year forecasting horizon, because in practice the forecasting exercise will be repeated at intervals of no more than a year. Some analysis of forecasts for horizons of up to four years is given in Sect. 6.

In the analysis, root mean square errors are used to estimate the one-step ahead error standard deviation.

The application of the forecasting methods is described in the following sub-sections.

4.1 ARIMA univariate (annual data):

The values of p and q, for ARIMA(p, d, q), were chosen using the Bayesian Information Criterion. In two (monthly) cases, non-differenced models were used. Firstly, the results of forecasting annual filings are discussed, then those for monthly data.

Aggregate forecasts of total EPO filings are produced; these forecasts are the sum of forecasts of the components of the total EPO filings. These components are either the four blocs (EPC, Japan, US, Other countries) or industries (IPC main Section headings A – H). A comparative analysis of this type produces a large volume of results, the policy adopted here is to first give example tables for the ARIMA models using annual data; subsequently summary results are given. The example tables give an idea of the relative difficulty of forecasting the different series, the summary tables show the relative overall accuracy between methods.

For the four blocs (see Table 4.6), the US has the lowest mape; the rest of the world has the highest. Forecasts labelled origin 1 are for 1995, based on data up to 1994. Forecasts labelled origin 2 are for 1996, based on data up to 1995, etc. The forecast based on aggregated blocs gives a similar rmse to the direct forecast of total EPO filings. The rmses associated with Origin 6 are typically high. These relate to the forecasts made at the end of 1999 for 2000. Total filings increased by 15% from 1999 to 2000 in contrast with the 7% growth experienced over the last ten years, 1990 to 2000. All four univariate forecasts of total EPO filings are shown in Fig. 4.4.

Table 4.6. Summary accuracy of ARIMA forecasts of total EPO annual filings forecasts by blocs

		EPC	JP	US	Rest of World	Blocs aggregated	Total EPO
	Std Error	1900	849	722	224		3211
Accuracy by origin (rmse)	1	895	97	659	52	1405	1129
	2	575	1780	1759	310	4424	3754
	3	4706	355	2221	841	8124	9750
	4	3927	103	983	749	5762	2352
	5	539	144	278	172	845	1839
	6	1863	2260	2796	3833	10753	10593
Overall rmse		2660	1186	1699	1638	6286	6212
Overall mape		4.29	4.91	4.11	8.91	4.68	4.38

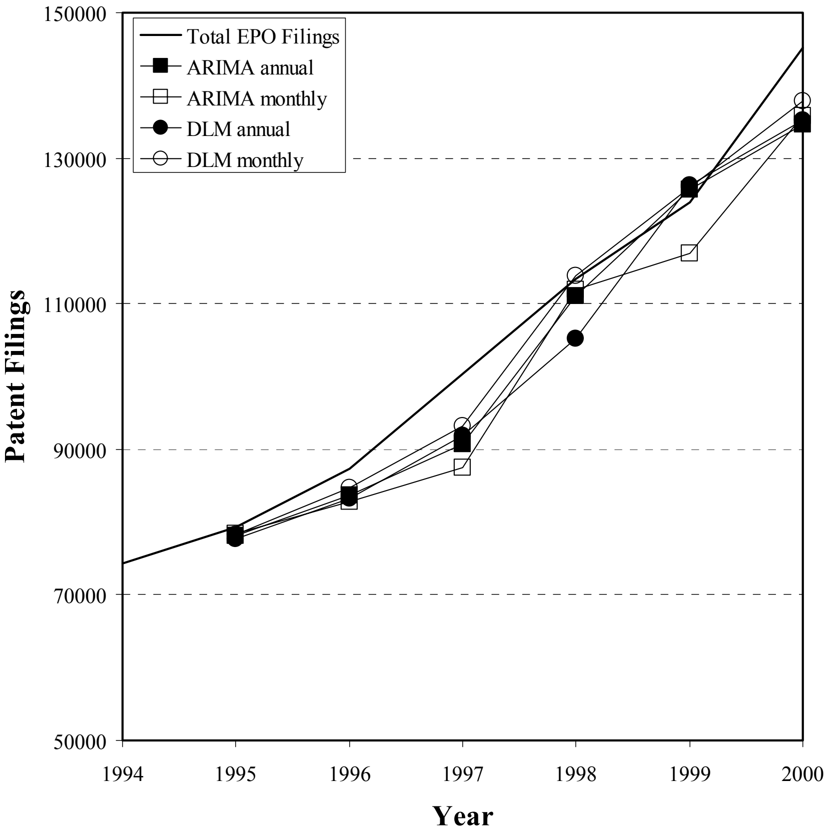

Fig. 4.4. One year ahead forecasts of total EPO filings for origins 1994 to 1999 from the four univariate models

The results of examining filings on an industry basis for the EPO are summarised in Table 4.7. industries G (physics) and H (electricity) are the most difficult to forecast with both high root mean square errors and high mean absolute percentage errors. The residual category 'other' has a high mape and high rmse showing that this balancing item is difficult to forecast (see Fig. 4.3). Aggregation of industry forecasts to give an overall EPO forecast is on average slightly better than the forecast of the total. There were five out of six origins where the aggregate forecast was more accurate than the direct forecast. It is worth remembering that industry specific forecasts have value, in terms of matching skills to future workload, even if aggregation does not lead to greater overall accuracy.

Table 4.7. Summary accuracy of ARIMA forecasts of EPO annual filings forecasts by Industry

		Industry										
		A	B	C	D	E	F	G	H	other	Aggre-gated	Total EPO
	Std Error	347	660	720	118	173	353	528	428	91		3211
Accuracy	1	256	197	582	2	47	119	846	404	324	1640	1129
by	2	554	711	553	51	8	167	529	1834	56	4540	3754
origin	3	1224	905	1697	183	320	546	979	1288	91	7050	9750
(rmse)	4	505	1190	651	44	241	988	350	1056	1166	4189	2352
	5	61	71	272	121	177	606	930	189	439	14	1839
	6	339	179	1330	152	41	609	3646	2262	2938	11498	10593
Overall rmse		612	685	982	113	181	585	1645	1382	1310	6093	6212
Overall mape		3.13	2.98	4.31	4.89	3.90	5.83	5.80	6.14	168.7	4.32	4.38

A: human necessities, B: performing operations/transport, C: chemistry; metallurgy, D: textiles; paper, E: fixed constructions, F: mechanical engineering, G: physics, H: electricity.

For this section, results for blocs are summarised in Table 4.8, results for industries are summarised in Table 4.9 and results for total EPO filings are summarised in Table 4.10. In order to judge comparative accuracy, the rank of each method for overall rmse is shown.

The lower the rank, the more accurate is the forecasting method. In addition, in order to gain an idea of how important the differences in rank are, a geometric mean is given. The geometric mean for approach i is defined below, where n filings series have been forecast:

$$GM_i = \sqrt[n]{\prod_{j=1}^{n} \frac{RMSE \quad \text{for} \quad \text{series} \quad j \quad \text{for} \quad \text{approach} \quad i}{\min(RMSE) \quad \text{for} \quad \text{series} \quad j}}$$

The reason for this measure is that it shows an average proportion indicating how far an approach deviates, on average, from the most accurate result.

4.2 ARIMA univariate (monthly data)

Aggregating monthly forecasts over time to provide annual forecasts does not provide greater accuracy, compared to forecasts based on annual data, for any bloc (see Table 4.8). Although aggregation across blocs gives a more accurate forecast of total EPO filings than the direct forecast of total

EPO filings; these are both less accurate than their forecasts based on annual data (see Table 4.10). In Fig. 4.4, the greater accuracy of the annual ARIMA forecast of total EPO filings is visible. Aggregation over the monthly forecasts leads to greater accuracy than the annual industry models for only three out of the nine industries (see Table 4.9). Aggregation of these industry forecasts gives a less accurate forecast of total EPO filings than the aggregated direct annual forecast of total EPO filings (see Table 4.10).

Table 4.8. Overall rmse levels for blocs (Rank of method shown in *italics*)

rmse	EPC		JP		US		Rest of World		Average rank	GM
Univariate ARIMA using annual data	2660	*5*	1186	*5*	1699	*7*	1638	*2*	4.8	1.29
Univariate ARIMA using monthly data	2721	*7*	1353	*6*	1675	*6*	1709	*6*	6.3	1.35
VAR using annual data	5226	*8*	2081	*8*	4935	*8*	1873	*7*	7.8	2.37
VAR using monthly data	1684	*1*	1866	*7*	1243	*2*	2095	*8*	4.5	1.27
Univariate DLM using annual data	2674	*6*	1077	*2*	1581	*5*	1647	*3*	4.0	1.24
Univariate DLM using monthly data	1898	*3*	1084	*3*	1380	*3*	1659	*4*	3.3	1.10
Multivariate DLM using annual data	2346	*4*	1152	*4*	1575	*4*	1681	*5*	4.3	1.23
Multivariate DLM using monthly data	1811	*2*	986	*1*	1169	*1*	1636	*1*	1.3	1.02

rmse defined in Table 4.5.

4.3 ARIMA multivariate (annual data)

For VAR modelling, the order of the process p is normally determined by likelihood ratio tests or by information criteria. However, in the case of yearly data, shortage of observations dictates the maximum value of p. For the four (k = 4) EPO bloc series, there are fifteen years (T = 15) of differenced data. The constraint on the number of degrees of freedom is:

$$(T-p)-kp>0\text{, thus } p<T/_{k+1}.$$

This means that the maximum value for p is two for blocs. For modelling individual industries (where k = 9) the situation is worse with a maximum value for p of one.

For blocs, the forecast accuracy of the VAR model over the last available six years is poor, and it is uniformly worse for each country than the univariate result (see Table 4.8). The shortage of degrees of freedom may have contributed to this poor performance. Repeating the exercise for industries (see Table 4.9) results in a poor forecasting performance by the VAR models. The shortage of degrees of freedom is more severe here as there are more industries than blocs.

4.4 ARIMA multivariate (monthly data)

In order to use stationary data, the data were transformed to $(1\text{-}B)(1\text{-}B^{12})X_t$. The number of lagged terms to include in the VAR was decided by Bayesian Information Criterion (BIC). Although the maximum value for the number of lags was of the order of twenty or so, the criterion led to the choice of low values. For the blocs, the number of lagged terms was 2, for industries, the number of lagged terms was 1. Note that the information criterion indicates that recent annual changes are useful in prediction, i.e. lags of 1 or 2 means that the data used for the forecasts are between one and two years old.

In addition, it should be noted that the forecasting of monthly data (in VAR) is more problematic than for annual data. Since no data later than the current origin are available for forecasting, the lagged data required in the forecast have to be replaced by forecasts. This means that monthly forecasts for all countries (or industries) are made recursively one month at a time.

The accuracy per bloc is similar to univariate ARIMA (monthly), where there are improvements for the EPC and the US (see Table 4.8). For the overall EPO filings, the forecast based on the aggregated bloc forecasts yields a rmse of 5 307 (see Table 4.10). Similar results for the industries within the EPO are shown in Table 4.9. Industry by industry, the monthly VAR accuracy is similar to monthly univariate accuracy. Aggregating to forecast annual filings, the monthly VAR is more accurate than the corresponding univariate figure for four out of nine industries (industries C, F, G and H). For the total EPO filings, there is an increase in accuracy with a rmse of 4 462 compared to 6 743 for aggregated monthly ARIMA (see Table 4.10).

4.5 DLM univariate (annual data)

Forecasts were computed recursively, using the same initial estimation region as defined in Table 4.4. The accuracy of forecasts using bloc annual data is summarised in Table 4.8. The accuracy of the dynamic linear model for the industry groupings is summarised in Table 4.9. Broadly, the DLM forecasts are similar to the comparable ARIMA ones.

4.6 DLM univariate (monthly data)

As explained earlier, the dynamic linear model was used on monthly data with stochastic trigonometric seasonal factors. Here, for blocs, the monthly rmse values are consistently lower than the comparable ARIMA rmse values. Aggregating across blocs to get an overall EPO filings forecast was more accurate than the direct univariate forecast for both aggregated monthly data and annual (see Table 4.10). The accuracy of forecasts by industry is summarised in Table 4.9. For six out of nine industries, the DLM using monthly data was more accurate than the annual DLM. We see there is no additional accuracy to be gained here by aggregating across industries (see Table 4.10). In Fig. 4.4, we can see that this forecast of total EPO filings tends to be more accurate than the annual DLM and the ARIMA models.

4.7 DLM multivariate (annual data)

The model chosen for forecasting was the homogeneous model that constrains the elements of the covariance matrix driving the disturbances to be proportional to one another. There is a halfway house of *trend homogeneity*, where the constraint of proportionality only applies to level and slope disturbances, not the error term. Experimentation showed that full homogeneity led to more accurate forecasts, so only these are reported. The accuracy of the multivariate dynamic linear model using annual data for EPO blocs is summarised in Table 4.8 and for industries in Table 4.9. The multivariate DLM is of similar accuracy to the univariate DLM.

4.8 DLM multivariate (monthly data)

For monthly data, the accuracy is summarised in Table 4.8 for EPO blocs and Table 4.9 for EPO industries. For both blocs and industries this model has a similar level of accuracy to the univariate DLM.

Table 4.9. Overall rmse levels for industries (IPC main fields of technology). Ranks in *italics*

	A		B		C		D		E		F		G		H		other		Average rank	GM
Univariate ARIMA using annual data	612	*1*	685	*4*	982	*4*	113	*4*	181	***1***	585	*5*	1645	*6*	1382	***6***	1310	*7*	4.2	1.28
Univariate ARIMA using monthly data	907	*5*	626	*3*	1252	*6*	119	*5*	390	***6***	537	*4*	1458	*3*	785	***2***	1384	***8***	4.7	1.38
VAR using annual data	1406	*8*	1712	*8*	2303	*8*	231	*7*	527	***8***	967	*8*	1861	*7*	1600	*7*	1284	***4***	7.2	2.27
VAR using monthly data	1243	*7*	781	*6*	1212	*5*	254	*8*	394	*7*	469	*3*	1016	*1*	749	***1***	1289	***5***	4.8	1.48
Univariate DLM using annual data	639	*2*	740	*5*	957	*3*	92	*1*	220	***4***	587	*6*	1630	*5*	1380	***5***	1198	***2***	3.7	1.28
Univariate DLM using monthly data	676	*3*	539	*1*	859	*1*	108	*3*	216	***3***	408	*1*	1457	*2*	809	***3***	1300	***6***	2.6	1.13
Multivariate DLM using annual data	1050	*6*	1061	*7*	1360	*7*	132	*6*	226	***5***	707	*7*	2629	*8*	2397	***8***	1239	***3***	6.3	1.76
Multivariate DLM using monthly data	717	*4*	546	*2*	863	*2*	103	*2*	195	***2***	437	*2*	1623	*4*	1190	***4***	1074	***1***	2.6	1.16

A: human necessities, B: performing operations/transport, C: chemistry; metallurgy, D: textiles; paper, E: fixed constructions, F: mechanical engineering, G: physics, H: electricity. rmse defined in Table 4.5.

Table 4.10. Overall rmse levels for different aggregated forecasts of total EPO filings

Type of Aggregation	Approach	rmse	Rank	GM
	Univariate ARIMA using annual data	6212	13	1.44
As a univariate time series	Univariate ARIMA using monthly data	7374	17	1.70
	Univariate DLM using annual data	6623	16	1.53
	Univariate DLM using monthly data	4478	3	1.03
	Univariate ARIMA using annual data	6286	14	1.45
	Univariate ARIMA using monthly data	6511	15	1.50
	VAR using annual data	7517	18	1.74
As an aggregate across blocs	VAR using monthly data	5307	6	1.23
	Univariate DLM using annual data	5796	9	1.34
	Univariate DLM using monthly data	4328	1	1.00
	Multivariate DLM using annual data	5989	10	1.38
	Multivariate DLM using monthly data	4578	5	1.06
	Univariate ARIMA using annual data	6093	11	1.41
	Univariate ARIMA using monthly data	5691	8	1.31
	VAR using annual data	9501	19	2.20
As an aggregate across industries	VAR using monthly data	4462	2	1.03
	Univariate DLM using annual data	6126	12	1.42
	Univariate DLM using monthly data	4567	4	1.06
	Multivariate DLM using annual data	10098	20	2.33
	Multivariate DLM using monthly data	5417	7	1.25

rmse defined in Table 4.5.

5 Analysis of comparative accuracy

In order to judge whether the observed differences in accuracy matter it is helpful to test the significance of the differences. Here we use the Friedman test for this task. The ranking of each treatment (forecasting method) is calculated across the available blocks (forecasting origin and time series)[3]. The test first evaluates if the differences in the ranks can be explained as random variation between methods of similar accuracy, or if the differences in ranks are due to greater accuracy of one or more methods. If the hypothesis of no difference in accuracy can be rejected, the signifi-

[3] Note that *bloc* is a national grouping; *block* is a grouping of data in this hypothesis testing exercise.

cantly more accurate methods can be identified. For details of the Friedman test, see Conover (1999).

We carry out three comparisons, over blocs, over industries and over total EPO filings. For EPO planning purposes, forecasts of filings by industry and of total filings are of particular interest.

In each case, the block data are the errors for a one-year ahead forecast over six origins and over the available series. For example for blocs, there are four time series (EPC, JP, US and rest of the world) and six origins giving twenty-four blocks.

For blocs, the multivariate version of the dynamic linear model using monthly data is most accurate, but it is only significantly better than the multivariate ARIMA using annual data, using the VAR model. In this case, the main result is that the annual VAR is significantly worse than the other methods.

For industries, the univariate monthly DLM is significantly more accurate than all methods except the multivariate monthly DLM. The multivariate annual versions of ARIMA and the DLM are significantly worse than the remaining models.

For the total EPO filings series, the methods include univariate forecasts of the series and forecasts aggregated over blocs or industries. Three of the four most accurate methods use the monthly univariate DLM, either on the total filings directly or aggregating across blocs or across industries. The use of monthly data dominates the rankings by accuracy.

One of the main issues of this study concerns the advantages of a multivariate model over a univariate model. Is there evidence of information contained in the time series of filings for one bloc or industry that will improve forecasting of filings in other blocs or industries? The answer is no; for blocs a multivariate model was most accurate but not significantly more accurate than its univariate counterpart; for industries there was no evidence of greater accuracy from a multivariate model. This conclusion is conditional on the industrial classification used here. It is not definitive for all possible classifications.

For the issue of aggregation over time, the use of monthly data, rather than annual data, does lead to greater accuracy in terms of ranking. In terms of significantly greater accuracy, this is achieved for industries where the most accurate two methods are significantly more accurate than their versions using annual data. (DMU vs DAU and DMM vs DAM). A similar statement applies to forecasts of total EPO filings, (Bloc DMU vs Bloc DAU and Total DMU vs Total DAU).

Table 4.11. Methods in descending order of accuracy by bloc, industry and total EPO filings

For blocs

Framework	Frequency			Average Rank	Is significantly more accurate than:
DLM	Monthly	Multivariate	(DMM)	3.6	AAM
ARIMA	Annual	Univariate	(AAU)	4.0	AAM
DLM	Monthly	Univariate	(DMU)	4.2	AAM
DLM	Annual	Univariate	(DAU)	4.2	AAM
ARIMA	Monthly	Multivariate	(AMM)	4.4	AAM
ARIMA	Monthly	Univariate	(AMU)	4.5	AAM
DLM	Annual	Multivariate	(DAM)	4.5	AAM
ARIMA	Annual	Multivariate	(AAM)	6.5	

For industries

Framework	Frequency			Average Rank	Is significantly more accurate than:
DLM	Monthly	Univariate	(DMU)	3.3	DAU, AMU, AMM, AAU, AAM, DAM
DLM	Monthly	Multivariate	(DMM)	3.7	AAM, DAM
DLM	Annual	Univariate	(DAU)	4.2	AAM, DAM
ARIMA	Monthly	Univariate	(AMU)	4.4	AAM, DAM
ARIMA	Monthly	Multivariate	(AMM)	4.5	AAM, DAM
ARIMA	Annual	Univariate	(AAU)	4.5	AAM, DAM
ARIMA	Annual	Multivariate	(AAM)	5.6	
DLM	Annual	Multivariate	(DAM)	5.8	

For EPO total

Framework	Frequency			Average Rank	Is significantly more accurate than
DLM	Monthly	Univariate	(Bloc DMU)	5.0	Bloc AMU, Bloc AAU, Ind DAU, Total AMU, Bloc DAM, Bloc DAU, Bloc AAM, Total DAU, Ind AAM, Ind DAM
DLM	Monthly	Univariate	(Total DMU)	5.2	Bloc AMU, Bloc AAU, Ind DAU, Total AMU, Bloc DAM, Bloc DAU, Bloc AAM, Total DAU, Ind AAM, Ind DAM

Table 4.11. (cont.)

For EPO total					
Framework	Frequency			Average Rank	Is significantly more accurate than
DLM	Monthly	Multivariate	(Bloc DMM)	6.2	Bloc AAM, Total DAU, Ind AAM, Ind DAM
DLM	Monthly	Univariate	(Ind DMU)	6.2	Bloc AAM, Total DAU, Ind AAM, Ind DAM
ARIMA	Monthly	Multivariate	(Ind AMM)	8.2	Ind AAM, Ind DAM
ARIMA	Monthly	Univariate	(Ind AMU)	8.7	Ind AAM, Ind DAM
ARIMA	Monthly	Multivariate	(Bloc AMM)	9.5	Ind DAM
DLM	Monthly	Multivariate	(Ind DMM)	9.7	Ind DAM
ARIMA	Annual	Univariate	(Total AAU)	10.5	Ind DAM
ARIMA	Annual	Univariate	(Ind AAU)	10.5	Ind DAM
ARIMA	Monthly	Univariate	(Bloc AMU)	11.7	
ARIMA	Annual	Univariate	(Bloc AAU)	11.8	
DLM	Annual	Univariate	(Ind DAU)	11.8	
ARIMA	Monthly	Univariate	(Total AMU)	12.2	
DLM	Annual	Multivariate	(Bloc DAM)	12.2	
DLM	Annual	Univariate	(Bloc DAU)	12.2	
ARIMA	Annual	Multivariate	(Bloc AAM)	12.7	
DLM	Annual	Univariate	(Total DAU)	13.2	
ARIMA	Annual	Multivariate	(Ind AAM)	15.2	
DLM	Annual	Multivariate	(Ind DAM)	17.7	

For the issue of aggregation over blocs or industries leading to more accurate forecasts of total filings, there was no significant evidence of this.

To achieve the greatest accuracy with one forecasting method, the univariate DLM with monthly data is the most appropriate choice.

If by some diktat, the use of monthly data was forbidden, a similar analysis can be carried out just using annual data. For blocs, the VAR model with annual data is significantly less accurate than other methods. For industries, the univariate ARIMA and DLM are significantly more accurate than their multivariate counterparts. For total EPO filings there was no significant difference detected between the methods. To achieve the greatest accuracy using annual data, either univariate ARIMA or univariate DLM could be used.

6 Forecast accuracy over a longer horizon

The analysis has concentrated on accuracy for the one year ahead forecast. One explanation is that accuracy over short horizons is a guide to accuracy over longer horizons. Another is that a wide-ranging comparison of accuracy over so many approaches and so many different series would have become (even more) unwieldy with more than one horizon. At this stage, two possible univariate forecasting strategies have been identified:

- Univariate DLM using monthly data.
- Univariate ARIMA using annual data.

A third method can be added to represent multivariate modelling:

- Multivariate DLM using monthly data

A comparison across horizons up to four years ahead is carried out using these methods.

The accuracy of the two approaches when applied to blocs is summarised by horizon in Table 4.12. Note that the rmses by horizon are calculated using fewer errors as the horizon increases (6 errors for one year ahead, 3 errors for four years ahead). The accuracy of the approaches for industries is summarised in Table 4.13. The accuracy deteriorates with length of horizon for all methods, from a mape of around 5% at one year to around 30% for four years ahead. The one to four year horizon forecasts of total EPO filings by the DLM (monthly) are shown in Fig. 4.5. It can be seen that the deterioration of accuracy with horizon is due to the forecasts with origins of 1994 to 1996. By 1997, the forecasts have adapted to the steeper trend.

Table 4.12. Comparison of the accuracy of the two broad approaches over a 4 year horizon (blocs)

	Approach	Horizon	EP	JP	US	Rest of World	blocs aggregated to give EPO forecast	Total EPO
rmse	ARIMA	1	2424	1174	1467	1629	5799	5751
		2	5018	2068	2448	1898	10086	9976
		3	9520	3252	5057	3028	19806	19137
		4	15325	4881	8512	4518	32588	32494
	DLM (monthly Univ.)	1	1898	1084	1380	1659	4328	4478
		2	4321	1739	2824	2141	8002	8818
		3	7618	1626	5875	3326	14478	14649
		4	12128	2746	7388	4603	24991	26957
	DLM (monthly Multiv.)	1	1811	986	1169	1636	4578	
		2	4007	1543	2431	2468	8678	
		3	7066	1837	4204	4080	14680	
		4	11398	2901	7728	6354	26915	
mape	ARIMA	1	4.61	4.69	3.29	8.48	4.50	4.60
		2	8.66	10.89	6.66	13.34	8.69	8.31
		3	17.44	18.01	13.05	23.03	16.73	16.13
		4	27.02	27.29	20.86	32.36	25.66	25.65
	DLM (monthly Univ.)	1	3.16	6.45	3.62	8.87	3.10	3.15
		2	8.07	10.02	7.64	15.18	6.20	7.24
		3	14.17	8.56	14.09	24.38	12.06	11.62
		4	21.42	15.11	18.07	32.73	19.30	20.54
	DLM (monthly Multiv.)	1	2.91	5.21	3.00	11.37	3.23	
		2	7.28	7.88	6.67	20.18	7.19	
		3	13.06	8.42	10.32	28.77	11.73	
		4	20.01	15.21	18.80	36.38	20.67	

rmse, mape, defined in Table 4.5.

Table 4.13. Comparison of the accuracy of the two broad approaches over a 4 year horizon (industries)

	Approach	Horizon	A	B	C	D	E	F	G	H	other	Industries aggregated to give EPO forecast
rmse	ARIMA	1	525	643	1014	112	188	534	1636	1165	1323	5673
		2	1196	1418	2117	198	460	852	2859	2206	1318	10385
		3	2198	2632	3121	312	811	1472	3817	4302	2217	19186
		4	361	407	5085	468	1187	2294	6694	7213	2449	32419
rmse	DLM (Monthly Univ.)	1	676	539	859	108	216	408	1457	809	1302	4567
		2	1420	1169	1750	193	364	799	2740	2435	814	8747
		3	2387	1921	2534	261	490	1203	4707	3883	2221	15626
		4	3810	2948	3485	376	689	2016	6097	6455	2564	27731
rmse	DLM (Monthly Multiv.)	1	717	546	863	103	195	437	1623	1190	1082	5417
		2	1543	1279	1679	173	328	821	2743	2850	1111	10885
		3	2640	2172	2627	243	496	1291	4004	4947	2221	19005
		4	4197	3413	4160	376	716	2132	6807	8419	2564	32676

Table 4.13. (cont.)

	Approach	Horizon	A	B	C	D	E	F	G	H	other	Industries aggregated to give EPO forecast
mape	**ARIMA**	1	2.23	3.23	4.93	4.92	4.28	5.94	5.73	5.75	145.27	4.26
		2	5.42	6.74	9.93	7.78	11.32	9.27	11.10	10.33	391.92	8.72
		3	10.76	12.13	13.77	14.46	20.57	16.49	18.14	19.54	683.84	16.24
		4	18.60	19.10	23.18	23.06	29.29	23.74	28.71	30.48	85.27	25.52
mape	**DLM (Monthly Univ.)**	1	2.50	2.71	3.61	4.78	5.29	3.92	5.41	3.95	93.55	3.30
		2	5.96	5.00	8.38	9.71	8.75	7.52	11.80	11.74	81.45	7.07
		3	10.96	8.12	11.61	13.04	10.77	11.79	19.01	16.91	100.00	12.96
		4	18.42	13.69	15.46	17.98	15.84	20.83	24.65	27.02	100.00	21.58
mape	**DLM (Monthly Multiv.)**	1	2.79	2.30	3.63	4.69	4.23	4.74	5.40	5.27	77.83	4.03
		2	6.57	5.70	8.32	8.64	6.63	7.74	11.33	12.51	88.45	8.59
		3	12.55	9.74	11.86	11.87	10.70	13.47	18.44	21.42	100.00	15.74
		4	20.87	16.02	18.60	16.94	16.67	22.11	27.56	34.31	100.00	24.98

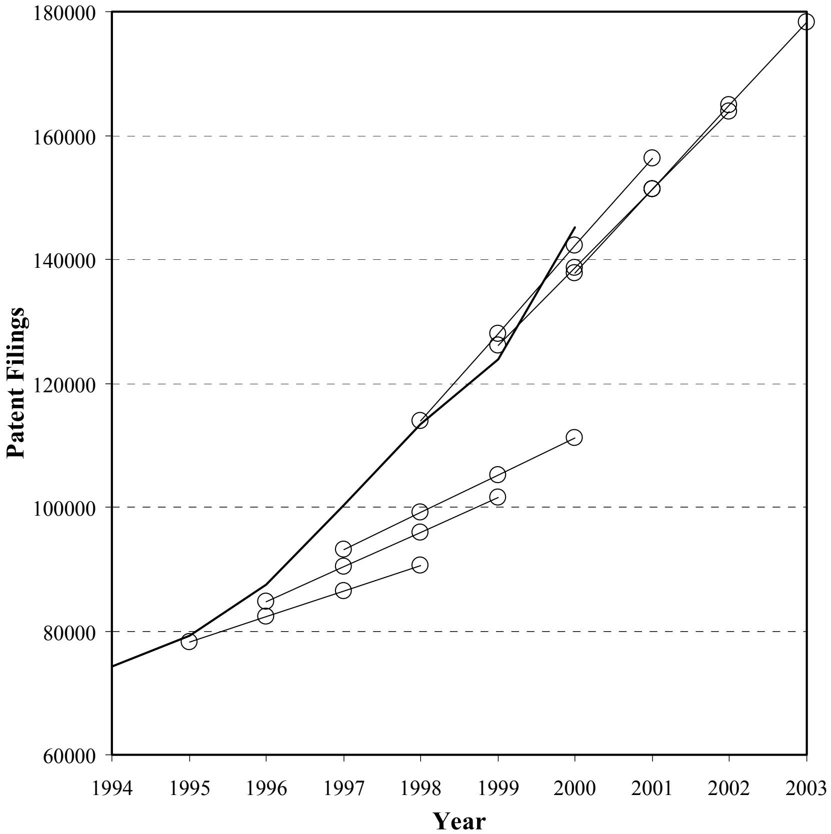

Fig. 4.5. DLM (Monthly) forecasts of total EPO filings with one to four year horizons for origins 1994 to 1999

In Table 4.14, the geometric mean of the rmses of the methods compared to the most accurate is shown for each method. It can be seen that the univariate monthly DLM is more accurate on average than the monthly multivariate DLM, which is in turn more accurate than the annual ARIMA for each of the horizons. In addition the most accurate forecast of total EPO filings is the aggregated blocs forecast by univariate DLM (for all horizons).

Thus the results over horizons up to four years ahead support the findings for a one-year horizon. The DLM approach maintains its superior accuracy over longer horizons.

Table 4.14. Comparison of geometric mean of rmse/minimum(rmse)

	Horizon (years)			
Method	1	2	3	4
ARIMA annual	1.19	1.17	1.25	1.24
DLM monthly univariate	1.06	1.06	1.06	1.01
DLM monthly multivariate	1.08	1.11	1.09	1.12

7 Conclusions

This study has looked at the comparative accuracy of methods of forecasting patent filings at the EPO on four different bases. The first basis for comparison is the issue of whether there is a correlation structure between filings by industry, or bloc, that can be exploited by using a multivariate rather than a univariate forecasting model. For the more accurate models, the monthly DLMs, no significant difference in accuracy was found between comparable models. The implication here is an absence of a useful correlation structure, the series respond similarly to common stimuli, but do not affect each other.

The second basis for comparison is the effect of cross-sectional aggregation across blocs or industries. Again, concentrating on the monthly DLMs, no significant difference in the accuracy of forecasts for total EPO filings was found between the direct forecasts and aggregate forecasts over blocs or industries. This corresponds with Miller's (1998) findings.

The third comparison concerns temporal aggregation, here our findings support those of Rossana and Seater (1995). Forecasts based on monthly data ranked highest in all comparisons. For industries, monthly DLMs were significantly more accurate than their annual counterparts.

The fourth basis of comparison is between the ARIMA and DLM frameworks. For monthly data, the DLM is more accurate, for industries and total EPO filings significantly more so, than monthly ARIMA. For annual data, there was no significant difference between the univariate models. The seasonal modelling of the DLM seems to have better captured the within year variation than the ARIMA model used. This has allowed the DLM to then detect information about changing trends more effectively than the ARIMA model.

It was also demonstrated that the greater accuracy of the univariate monthly DLM persisted over horizons of up to four years.

This univariate forecasting exercise provides a benchmark against which other forecasts drawing on different data sources can be compared. In addition the DLM lends itself to further development by the introduction of possible explanatory variables, such as research and development expenditure, into the state equations.