Europäisches Patentamt
European Patent Office
Office européen des brevets

FINDING A WELL PERFORMING BOX-JENKINS FORECASTING MODEL FOR ANNUALISED PATENT FILINGS COUNTS

Peter Hingley[1] & Gerhard Dikta[2]

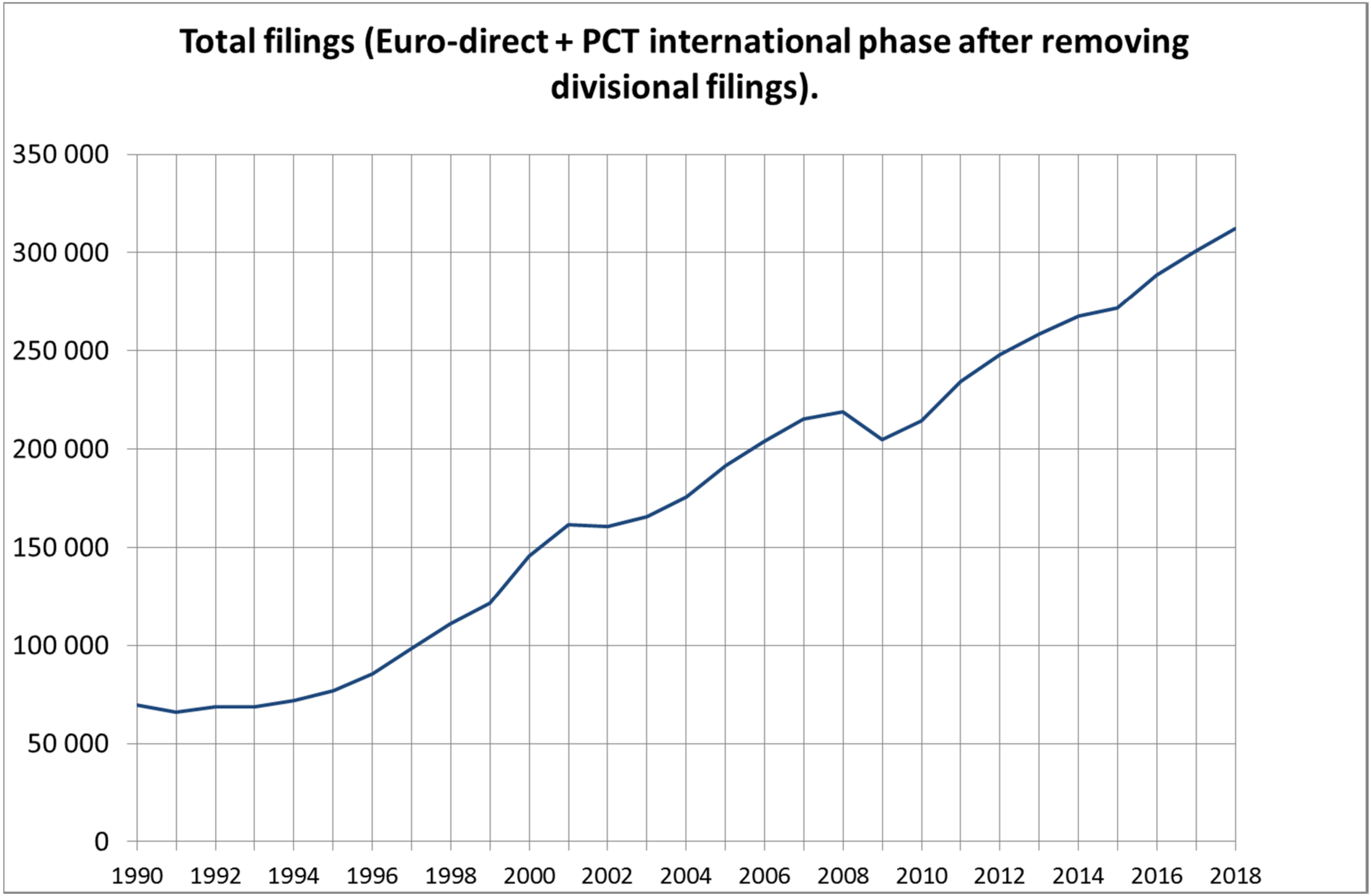
International Symposium on Forecasting, Thessaloniki, Greece, June 2019

1. European Patent Office, phingley@epo.org 2. Aachen University of Applied Sciences, dikta@fh-aachen.de

Contents

Disclaimer: The forecasts contained in this presentation are not official forecasts of the EPO

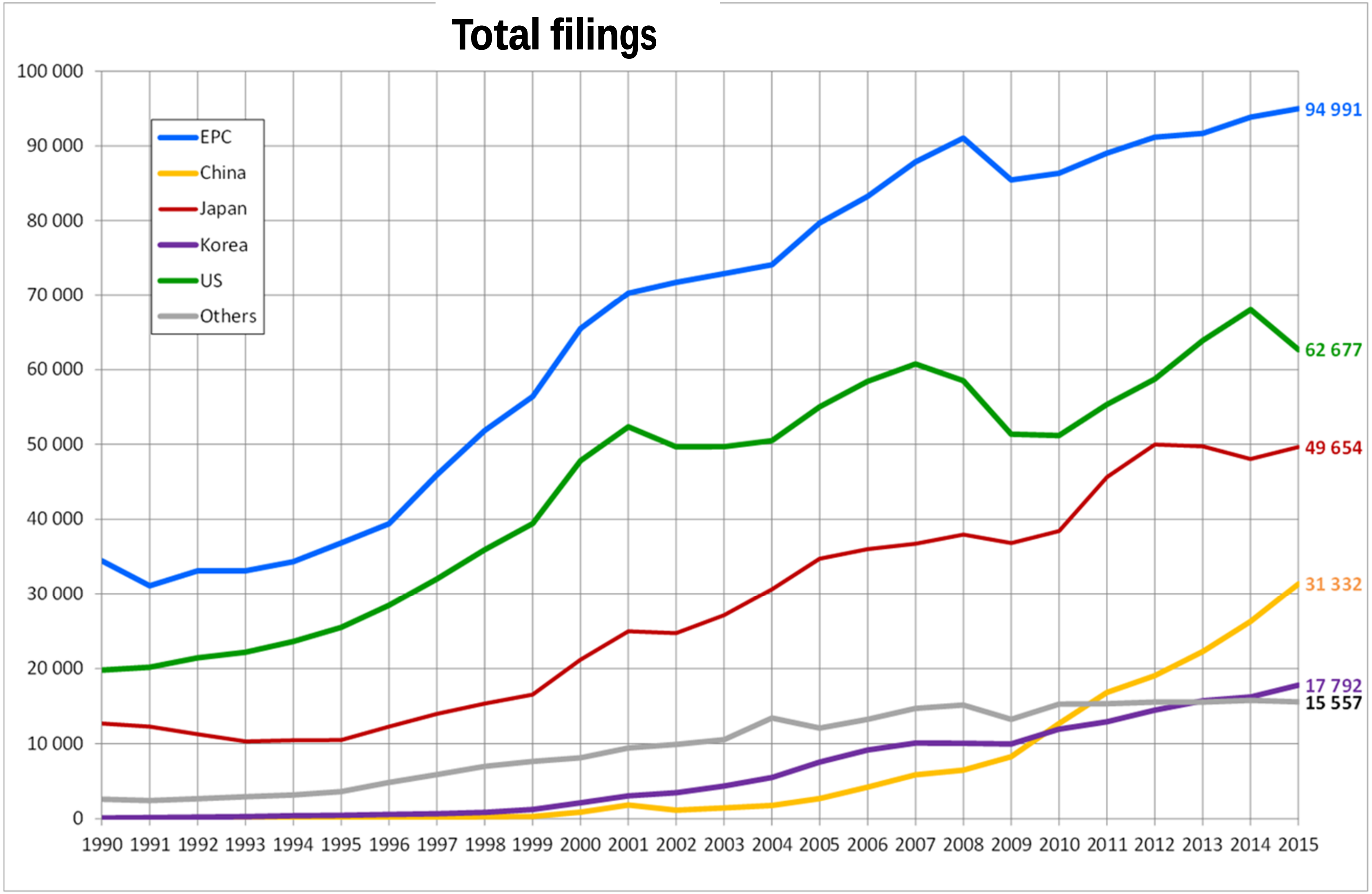
1. The evolution of EPO Total Filings (TFs).



Forecasting is done as an annual exercise on EPO Total filings (TFs henceforward).

The evolution of EPO Total Filings (TFs).

Breakdowns by Geographical blocs/countries



2. Forecasting TFs, annual for next few years.

Forecasting methods

See Hingley (2016) for a review.

Trend / scenario analyses:-

Direct (TFs vs time)

Transfer (First Filings, transfer coefficients to TFs)

Patent Filings Survey among applicants EPO (2019)

Econometric models:-

Dynamic log linear model Hingley & Park (2017)

Box Jenkins based times series models (ARIMA, ADL):- Dikta (2006)

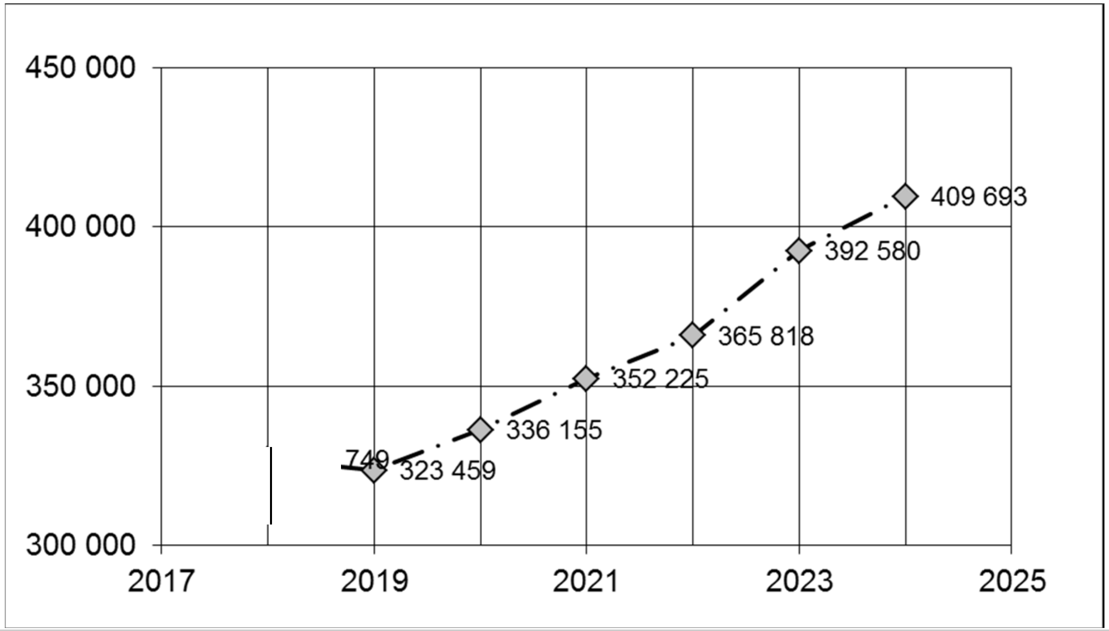
EPO TFs are considered as one time series

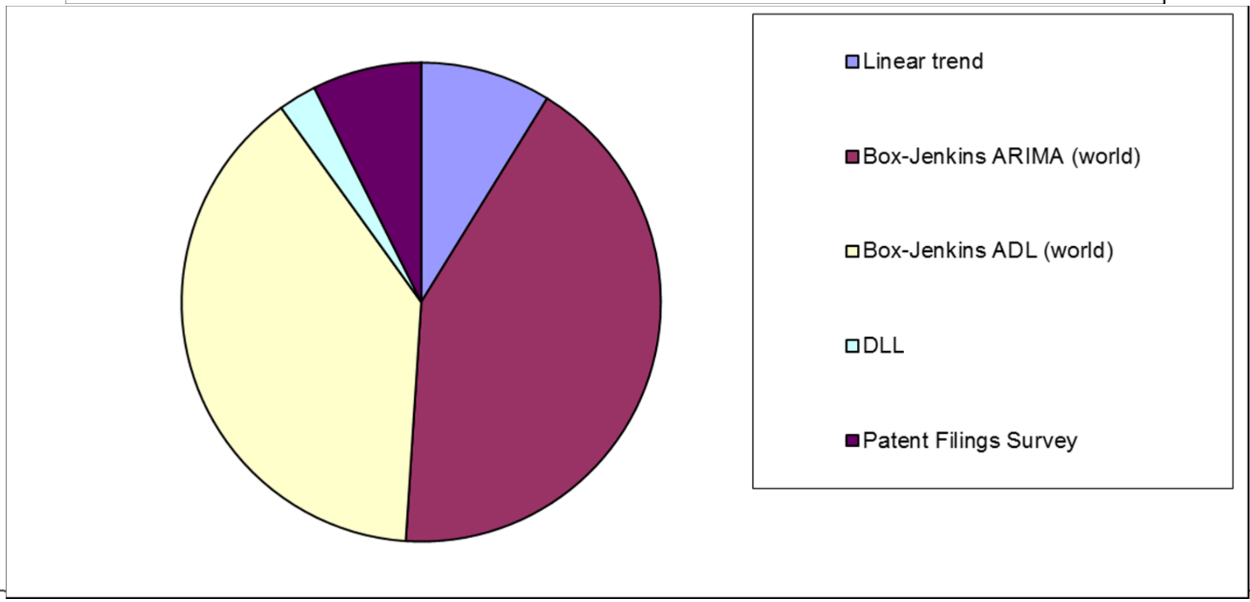
1. **ARIMA (Auto-regressive Integrated Moving Average)**
2. **ADL (Auto-distributed lags, using the predictors R&D expenditures of source countries and GDP of source countries)**

Automatic model selection program.

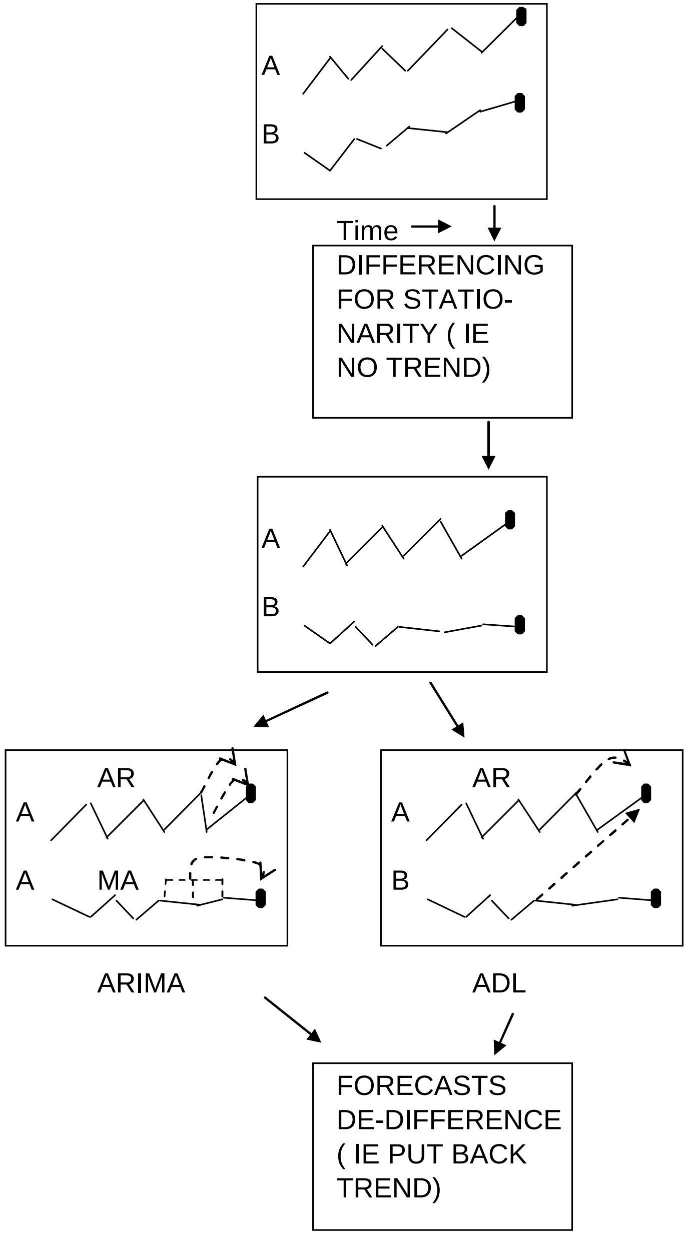
Forecasting TFs, annual for next few years.

Performance based average scenario (PBAS).





3. Box-Jenkins approach.



The program first determines the degrees of differencing required by all series under study. (KPSS tests at a default level of P > 0.1)

Kwiatkowski et al. (1992)

Then, for each method, the program runs alternative models and selects the best model via Akaike's information criterion (AIC).

Burnham & Anderson (1998)

Output of the selected fit is supported by normality and autocorrelation tests on the residuals.

Models including Lags up to 4 years are allowed.

Box-Jenkins approach.

All allowed models at the degree of differencing of TFs that passes the KPSS test are reported. (The “Trace”).

The selected model is the one in the Trace that minimises AIC.

Different models can minimise AIC for successive years.

So we also impose the same model over a set of successive years.

An imposed model is chosen that appears in the Trace each year (as far as possible) and that has a low AIC.

A RETROSPECTIVE FORECASTING EXERCISE

Forecasts were made for data starting in 1994, and ending in 2013, 2014, 2015, 2016 and 2017.

KPSS always selects first order differences for TFs.

In ADL , R&D growth rates and GDP Growth rates were added as predictors for TFs.

KPSS always selects zero order differences for R&D Growth and for GDP Growth with ADL.

Box-Jenkins approach.

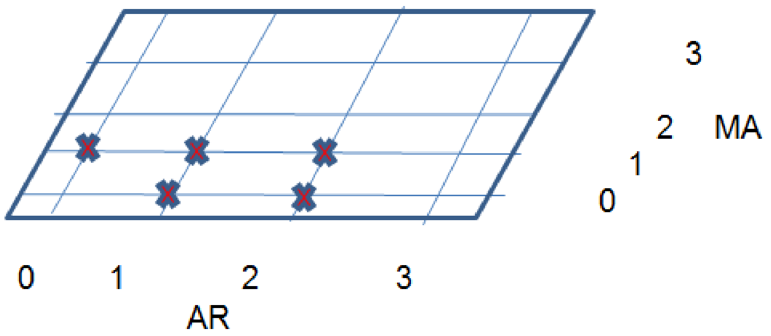
For example, for 1994 to 2017 for ARIMA models, the following models appear in the Trace.

Model	ar	diff	ma	aic
1	1	1	0	18.21
2	2	1	0	18.39
3	0	1	1	17.91
4	1	1	1	18.29
5	2	1	1	18.50
6	4	1	1	18.91

← Minimum AIC

← Not analysed

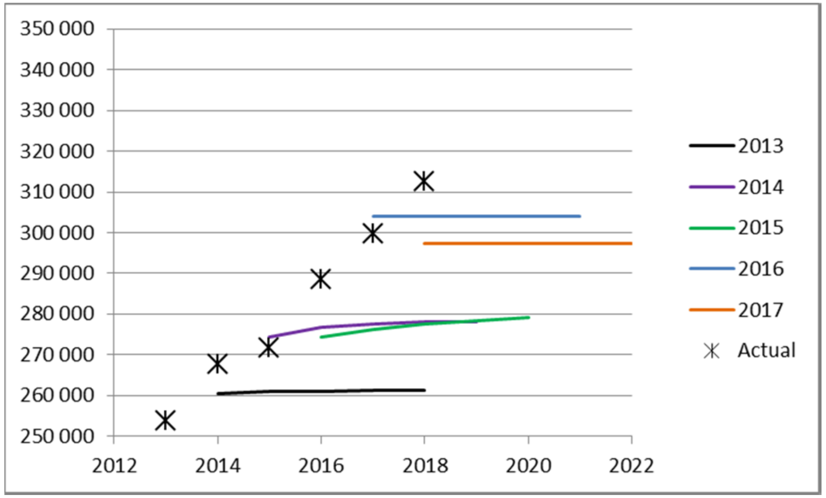
TFs differenced once



Box-Jenkins approach.

ARIMA models, Selected model for each year. First differences

ARIMA	Selected model for each year. First differences.				
Year	AR	Diff	MA	AIC	SE 1year
2013	1	1	1	18.3	8 909
2014	1	1	1	18.3	8 789
2015	1	1	0	18.2	8 844
2016	0	1	1	17.9	9 289
2017	0	1	1	17.9	9 117



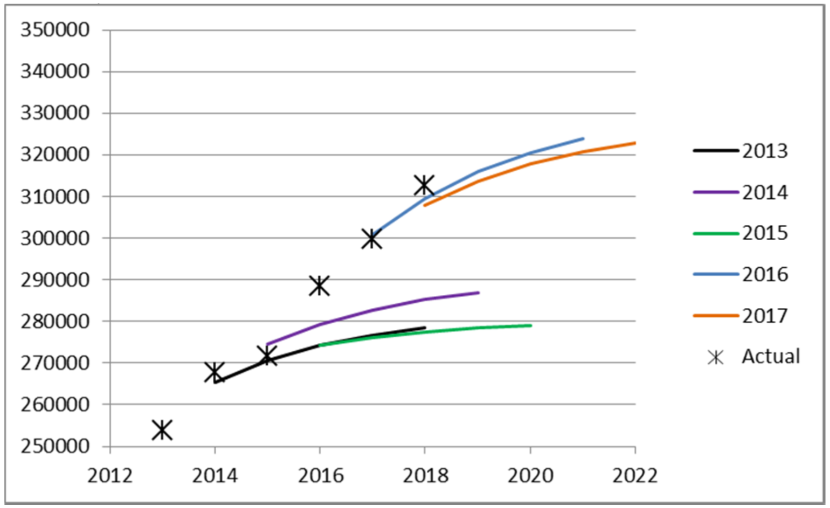
Years ahead	MAPE	MPE
1	3.0%	-2.0%
2	4.7%	-4.7%
3	9.4%	-9.4%
4	12.0%	-12.0%

Always under-forecasts

Box-Jenkins approach.

ARIMA models, Imposed (1,1,0) model for each year. First differences

ARIMA	Imposed (1,1,0) model for each year				
Year	AR	Diff	MA	AIC	SE 1year
2013	1	1	0	18.3	9 280
2014	1	1	0	18.3	9 046
2015	1	1	0	18.2	8 844
2016	1	1	0	17.9	9 167
2017	1	1	0	17.9	8 956



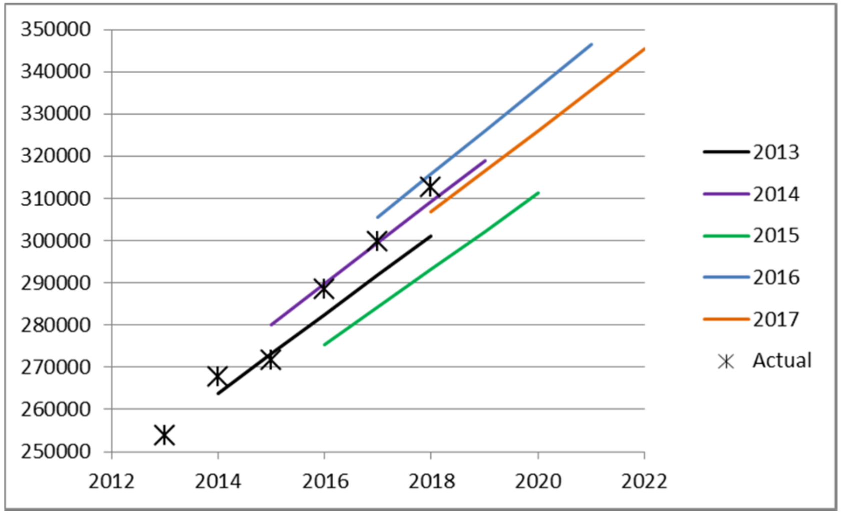
Years ahead	MAPE	MPE
1	1.8%	-1.2%
2	3.1%	-3.1%
3	7.3%	-7.3%
4	8.2%	-8.2%

Improves.
Still usually under-forecasts

Box-Jenkins approach.

ARIMA models, Selected model for each year. Second differences

ARIMA	Selected model for each year. Second differences.				
Year	AR	Diff	MA	AIC	SE 1year
2013	0	2	2	18.1	7 634
2014	0	2	2	18.1	7 485
2015	0	2	2	18.1	7 462
2016	0	2	2	18.1	7 794
2017	0	2	2	18.1	7 677



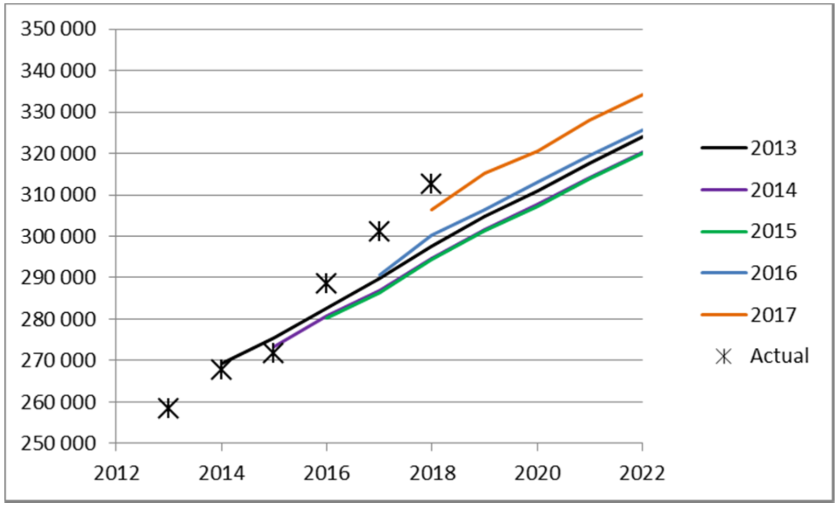
Years ahead	MAPE	MPE
1	2.6%	-0.6%
2	1.8%	-0.8%
3	2.8%	-2.8%
4	1.9%	-1.9%

Big improvement.
Same model always selected.

Box-Jenkins approach.

ADL models, Selected model for each year. First differences

ADL	Selected model for each year. First differences.				
Year	TF	R&D growth rate	GDP growth rate	AIC	SE 1year
2013	2,3	0,1	0,2	307.7	2 420
2014	3	0,1,3	0,2,3	325.8	2 260
2015	3	0,1,3	0,2	344.2	2 396
2016	3	0,1,3	0,2	372.7	2 967
2017	3	0,1,3	0,2	392.2	3 167



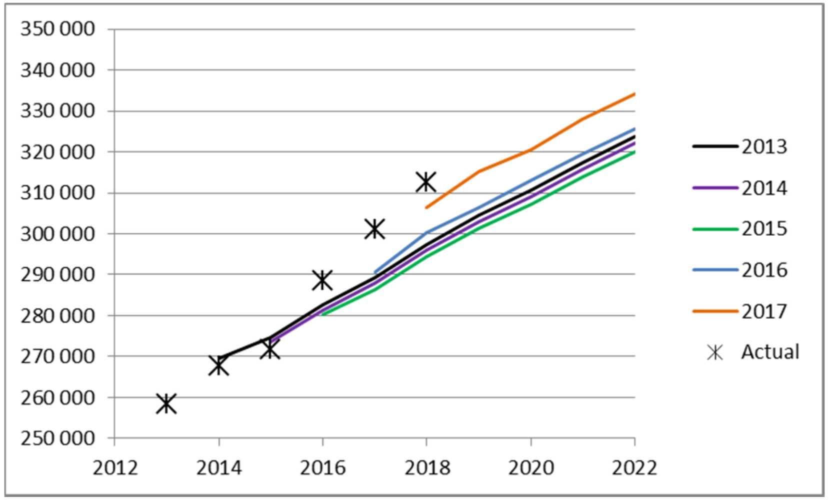
Years ahead	MAPE	MPE
1	1.8%	-1.3%
2	3.1%	-2.4%
3	4.0%	-4.0%
4	4.5%	-4.5%

Usually under-forecasts.
Better than free fit with ARIMA

Box-Jenkins approach.

ADL models, Imposed model for each year. First differences

ADL	Imposed (3;0,1,3;0,2) model for each year. First differences.				
Year	TF	R&D growth rate	GDP growth rate	AIC	SE 1year
2013	3	0,1,3	0,2	NA	2 467
2014	3	0,1,3	0,2	326.0	2 409
2015	3	0,1,3	0,2	344.2	2 396
2016	3	0,1,3	0,2	372.7	2 967
2017	3	0,1,3	0,2	392.2	3 167



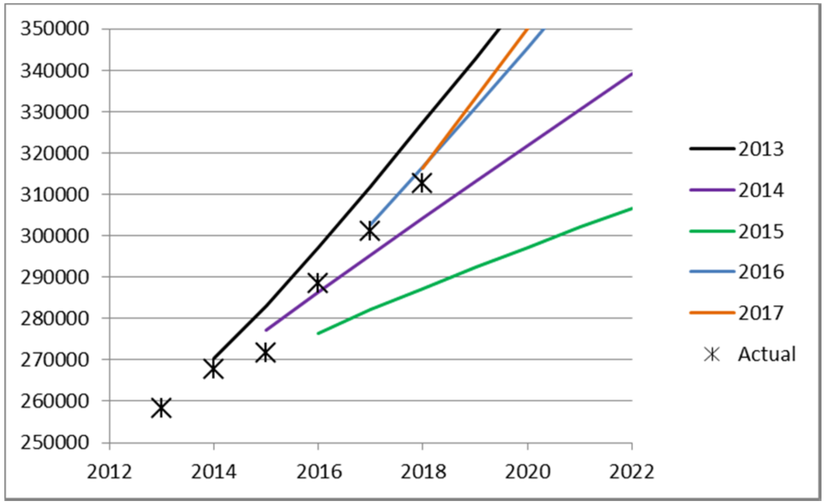
Years ahead	MAPE	MPE
1	1.8%	-1.3%
2	3.0%	-2.5%
3	3.9%	-3.9%
4	4.4%	-4.4%

Slight improvement.
Changes only to 2013 and 2014.

Box-Jenkins approach.

ADL models, Selected model for each year. Second differences

ADL	Selected model for each year. Second differences.				
Year	TF	R&D growth rate	GDP growth rate	AIC	SE 1year
2013	1,2,3	1,2	0,1,2,3	337.9	2 860
2014	1,3	0,1,2	0,1,2	359.2	3 249
2015	1,3	0,1,2	0,1,2	359.1	3 207
2016	3	0,2,3	0,1	405.2	4 390
2017	1,3	2	0,2	423.8	4 491



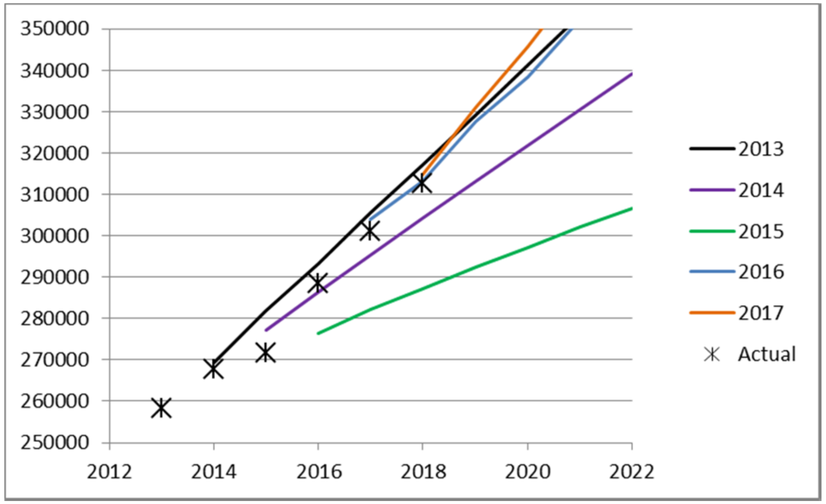
Years ahead	MAPE	MPE
1	1.8%	0.1%
2	3.1%	-0.4%
3	4.3%	-2.4%
4	3.1%	0.4%

Similar MAPEs.
No under-forecasting bias.

Box-Jenkins approach.

ADL models, Imposed model for each year. Second differences

ADL	Imposed (1,3;0,1,2;0,1,2) model for each year. Second differences.				
Year	TF	R&D growth rate	GDP growth rate	AIC	SE 1year
2013	1,3	0,1,2	0,1,2	NA	3 060
2014	1,3	0,1,2	0,1,2	359.2	3 249
2015	1,3	0,1,2	0,1,2	359.1	3 207
2016	1,3	0,1,2	0,1,2	NA	4 246
2017	1,3	0,1,2	0,1,2	NA	4 139



Years ahead	MAPE	MPE
1	1.7%	0.0%
2	2.7%	-0.8%
3	3.9%	-2.8%
4	2.1%	-0.6%

Further improvement.

4. Exact Information Criterion EIC.

OTHER INFORMATION CRITERIA.

AIC is well known and original, but approximate. Several other information criteria are available that might select better models for forecasting (EG Bayesian Information Criterion – BIC).

ARIMA only was studied. For now, AIC was accepted for candidate model selections (IE the Trace of candidate models is built on AIC). Pairwise comparisons of the trace models was made using an Exact Information Criterion (EIC). Hingley (2008)

EXACT INFORMATION CRITERION EIC.

Consider ARIMA models as General Linear models with given Design matrix X and Variance matrix V.

Say that the true model for the data w is multivariate Normal with design matrix X and variance matrix V_0, with a p x 1 parameter vector θ_0. $MN_w(X_0\theta_0, V_0)$.

Estimation model is $MN_w(X_1\phi_1, V_1)$ with a q x 1 parameter vector ϕ.

Exact Information Criterion EIC.

The EIC $H(\hat{\theta}, \hat{\phi})$ is based on the idea of Kullbach-Leibler Information and is given by

$$H(\hat{\theta}, \hat{\phi}) = log[\frac{m_A(\hat{\theta})}{m_B(\hat{\phi})}]$$

$m_A(\hat{\theta})$ is the analytic density of the maximum likelihood estimate $\hat{\theta}$ under the true model.

$m_B(\hat{\phi})$ is the analytic density of the maximum likelihood estimate $\hat{\phi}$ under the estimation model when the data are generated by the true model.

$$m_B(\hat{\phi}) = \frac{1}{\sqrt{(2\pi)^q abs(|X_1^T V_1{}^{-1} V_0 V_1{}^{-1} X_1|)}}.$$
$$\exp\Big[\frac{-1}{2}.((X_1^T V_1{}^{-1}(X_1\hat{\phi} - X_0\hat{\theta}))^T.$$
$$(X_1^T V_1{}^{-1} V_0 V_1{}^{-1} X_1)^{-1}.$$
$$X_1^T V_1{}^{-1}(X_1^T V_1{}^{-1}(X_1\hat{\phi} - X_0\hat{\theta})))\Big]$$

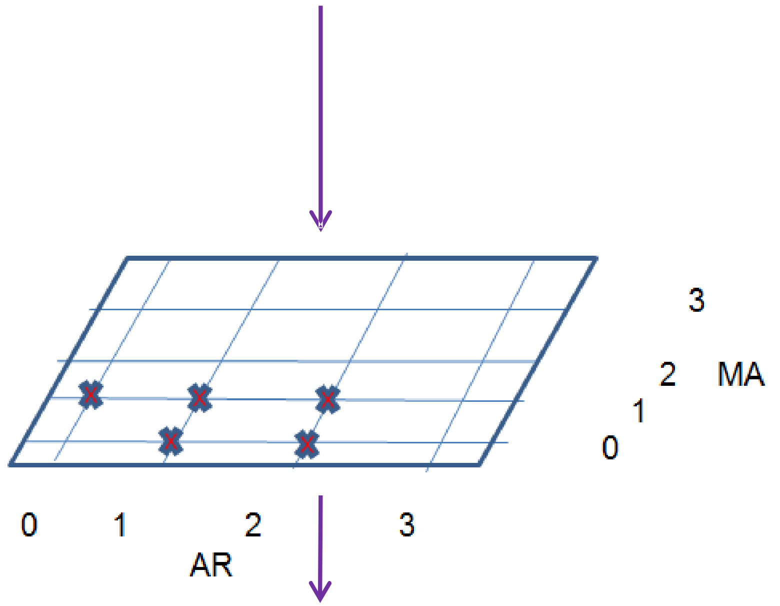
Exact Information Criterion EIC.

KPSS only allows a certain degree of differencing of TFs (here first differences). The candidate models are limited to the Trace at this degree of differencing. Pairwise comparisons by EIC were made for this limited set of models. Each model can be the DGM (true model) or the Estimation Model.

KPSS

TFs not differenced

TFs differenced once



TFs differenced twice

Not checked,
overfitted?

Exact Information Criterion EIC.

RESULTS SUGGEST THAT THE ARIMA (1,1,1) MODEL MAY WORK BEST AS ESTIMATION MODEL WITHIN THE CANDIDATES.

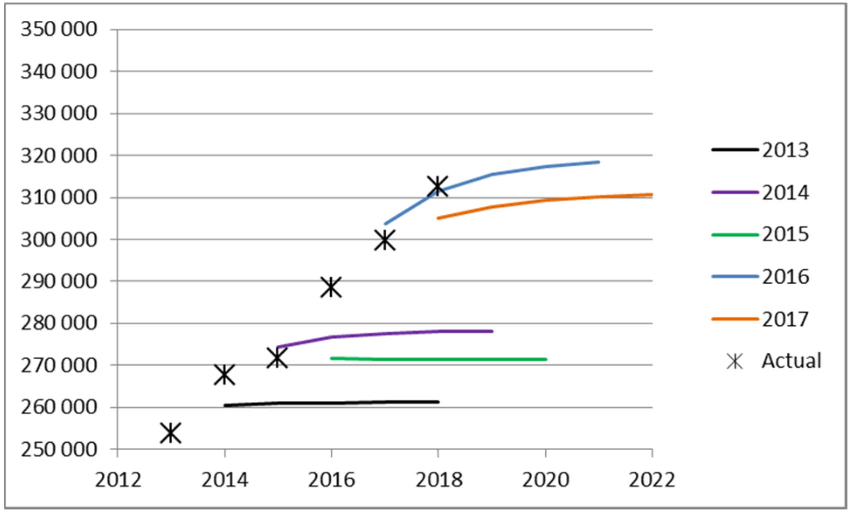
True model	Estimation model					
		A. 1,1,0	B. 2,1,0	C. 0,1,1	D. 1,1,1	E. 2,1,1
A. 1,1,0	log(mB)	1.71	-1.79	27.41	-0.45	24.13
log(mA) = 1.71	**EIC**	**0.0**	**3.5**	**-25.7**	**2.2**	**-22.4**
B. 2,1,0	log(mB)	-1.54	3.46	-61.76	-1.09	-1.97
log(mA) = 3.46	**EIC**	**5.0**	**0.0**	**65.2**	**4.6**	**5.4**
C. 0,1,1	log(mB)	-2.01	-3.56	2.55	-2.48	-4.53
log(mA) = 2.55	**EIC**	**4.6**	**6.1**	**0.0**	**5.0**	**7.1**
D. 1,1,1	log(mB)	-1.09	-2.21	-2.98	5.28	-8.85
log(mA) = 5.28	**EIC**	**6.4**	**7.5**	**8.3**	**0.0**	**14.1**
E. 2,1,1	log(mB)	-1.76	-3.29	-0.16	-2.58	5.49
log(mA) = 5.49	**EIC**	**7.2**	**8.8**	**5.6**	**8.1**	**0.0**
	Standard Deviation	13084	13121	**12530**	13489	13302
	AIC	18.2	18.4	**17.9**	18.3	18.5
	Lowest absolute EIC (except 0)	4.6	3.5	5.6	2.2	5.4
	Highest absolute EIC	7.2	8.8	65.2	8.1	22.4
	Average absolute EIC (except 0)	5.8	6.5	26.2	**5.0**	12.3

↑ Best by AIC ↑ Best by EIC

Exact Information Criterion EIC.

ARIMA models, Imposed (1,1,1) model for each year. First differences

ARIMA	Imposed (1,1,1) model for each year				
Year	AR	Diff	MA	AIC	SE 1year
2013	1	1	1	18.3	8 909
2014	1	1	1	18.3	8 789
2015	1	1	1	18.2	8 582
2016	1	1	1	18.3	9 104
2017	1	1	1	18.3	8 908



Years ahead	MAPE	MPE
1	2.7%	-1.8%
2	4.4%	-4.4%
3	10.0%	-10.0%
4	12.0%	-12.0%

Not as good as ARIMA(1,1,0).
Always under-forecasts.

5. Conclusions

ARIMA:- The automatic models under-forecast future TFs in this experiment;

In 1st differences, (1,1,0) works best, but this is a rather uninformative model;

In 2nd differences, forecasting power improves, but are these models allowed?

ADL:- This model forecasts better than ARIMA does;

In 1st differences, there is still a tendency to under-forecast with ADL;

In 2nd differences, forecasts improve and the imposed model is even better;

But, as in ARIMA, are 2nd differences allowed when KPSS selects 1st differences?

Conclusions

EIC vs AIC:- So far looked at for ARIMA only;
EIC suggests (1,1,1) model while AIC suggests (1,1,0) model;
But forecasting ability of (1,1,1) is less than for (1,1,0);
EIC needs more exploration.

Further work to do:- Test more time periods, do preliminary results above always apply?;
Try alternative fitting methods (maximum likelihood as well as conditional least squares);
Extend the scope of EIC comparisons, because models in the Trace do not remain constant from year to year, and also consider other criteria (EG BIC).

6. References.

1. P. Hingley (2016), " Forecasting total numbers of filings at the European Patent Office:- How useful are breakdowns by countries and technologies", 18th IIF Workshop, Milan, http://www.innovazionesistematica.it/index.php?option=com_content&view=article&id=107&Itemid=1

2. EPO (2019), "Patent Filings Survey 2018", https://www.epo.org/service-support/contact-us/surveys/patent-filings.html

3. P. Hingley & W. Park (2017), "Do business cycles affect patenting; evidence from European Patent Office filings", Technological Forecasting and Social Change, 116, p76-86.

4. G. Dikta (2006), "Time series methods to forecast patent filings", in "Forecasting Innovations", eds. P. Hingley & M. Nicolas, Springer, p95 - 124.

5. D. Kwiatkowski, P. Phillips, P. Schmidt & Y. Shin (1992), "Testing the null hypothesis of stationarity against the alternative of a unit root", J. of Econometrics, 54, p159-178.

6. K. Burnham & D. Anderson (1998), "Model selection and multimodel inference", 2nd edition, Springer.

7. P. Hingley (2008), "Distributions of maximum likelihood estimators and model comparisons", in "Current Themes in Engineering Science 2007", ed. A. Korsunsky, American Institute of Physics Proceedings 1045, p111-122.