World Patent Information 35 (2013) 130–140

Contents lists available at SciVerse ScienceDirect

World Patent Information

journal homepage: www.elsevier.com/locate/worpatin

Optimization of prediction methods for patents and trademarks in Spain through the use of exogenous variables

Antonio Hidalgo [a,*], Samuel Gabaly [b]

[a] Dept. Business Administration, Universidad Politécnica de Madrid, c/José Gutiérrez Abascal, 2, 28006 Madrid, Spain
[b] Gabaly Diseño S.L., Madrid, Spain

Keywords:
Patent time series
Trademark time series
Intelligent transfer function model
Exogenous variables
Polynomial distributed lag

A B S T R A C T

An accurate forecast of patent and trademark application filings is strategic for resource planning at the Spanish Patents and Trademarks Office and other patent offices, national and supranational. The need for reliable forecasts of patents and trademarks application filings has been accentuated by the current situation of budgeting rationalization imposed by the economic crisis. In this study we have evaluated the suitability and effectiveness of different methodologies for advanced data analysis to predict the number of national patent and trademark applications in the short and medium terms (2011–2014), including the use of exogenous variables or predictors which help to understand the changes in these variables. The inclusion of exogenous variables which explain the behavior of patent and trademark application filings, in particular the investment in R&D and GDP, and the use of advanced predictive analysis techniques, amongst which the most notable are Polynomial Distributed Lags and Intelligent Transfer Function models, have all achieved an improvement upon the prediction and modeling power possessed by the models formerly used to predict trademark and patent series based only on the analysis of time series.

1. Introduction

The Spanish Patents and Trademarks Office (SPTO) requires accurate analyses and predictions regarding the changes in demand for its services, especially in terms of national patent and trademark applications, in order to ensure high-quality, properly dimensioned service for its clients. Given the current situation of budgeting rationalization imposed by the economic crisis, this factor is highly relevant because prior knowledge of changes in patent and trademark applications may contribute to an optimization of planning, an improvement in cost rationalization and greater efficiency in providing services to clients in the coming years.

With this objective in mind, in 2010 the Spanish Office of Patents and Trademarks promoted a research project aimed at developing a methodology for predicting changes in the number of national patent and trademark application filings. The methodology used consists of three stages: the objective of the first stage (developed in 2010) was to predict the changes in the number of national patent and trademark applications for a time horizon of one to four years (short and medium term), using regression models of trends and advanced time series models; the second stage (developed in 2011) sought the same time objective, but using advanced econometric methods to identify indicators correlated with the changes in the national patent and trademark applications; last of all, the third stage (pending development) is oriented towards predicting the long-term changes (with a horizon of more than five years) using data at the company and sector levels, as well as identifying potential transfer functions through the use of multiple techniques, such as surveys among patent applicants, long-term econometric models and signals analysis.

The results found in the first stage of research made it clear that it is feasible to model the series of national patent and trademark applications with different models of time series and that the advanced time series models, in particular ARIMA (Auto-Regressive Integrated Moving Average) [1], are better adjusted to the real values of the series than the regression models of trends with satisfactory results in terms of the fit of models and relatively low error levels [2]. In fact, a comparison of the prediction of patent and trademark applications performed using real data from the year of 2010 displays a high level of reliability, as demonstrated by the data in Table 1.

The goal of this study (second stage) is to evaluate the suitability and effectiveness of different methodologies for advanced data

* Corresponding author. Tel.: +34 913363210.
E-mail addresses: ahidalgo@etsii.upm.es (A. Hidalgo), sgabaly@gmail.com (S. Gabaly).

0172-2190/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wpi.2012.12.009

Table 1
Comparison of Spanish patent and trademark application predictions with real data in 2010.

Year 2010		Prediction	Real
Patents	Prediction	3.740	3.702
	SCI 95%	4.290	
	ICI 95%	3.246	
Trademarks	Prediction	55.071	64.948
	SCI 95%	65.595	
	ICI 95%	45.900	

analysis to predict the number of national patent and trademark applications in the short and medium terms (2011–2014), including the use of exogenous variables or predictors which help to understand the changes in these variables. The starting point of the research was made up of the aggregate annual data on national patent applications (period 1986–2010) and trademark applications (period 1979–2010).

As for the exogenous variables or predictors which have been taken into consideration, they were selected due to their recurring use in the research literature [3–5], and because they are being used at present by entities such as the United States Patents and Trademarks Office (USPTO), the European Patent Office (EPO) and the World Intellectual Property Office (WIPO). These variables are as follows: Gross Domestic Product (in millions of euros), annual GDP growth rate (%), per capita GDP (in millions of euros), total investment in R&D (in millions of euros and as percentage of GDP), firm investment in R&D (in millions of euros and as percentage of GDP), number of researchers (full time equivalent), the Industrial Production Index (IPI), the IBEX-35 stock market index (annual average), gross capital formation (in millions of euros), venture capital investment (percentage of GDP), prices of national patent and trademark applications (in euros), and the number of companies in Spain. All economic variables were considered at constant prices in order to avoid including the effects of inflation in the prediction models.

The article is organized into four sections. First of all, a review is performed of the prediction methods used at other patent offices, national and supranational. The second section provides an analysis of trends in national patent and trademark applications, as well as identifying the exogenous variables which will be used for the prediction models, on the basis of their relationship with the target variables and the number of cases in which values are provided. In the third section, the prediction analysis models are explained and the results thereof are shown, identifying those which possess the highest level of overall fit and, therefore, the greatest predictive power. Last of all, the main conclusions of the study are summarized and the predictions of future values are provided for the series of national patents and trademarks using the model which has displayed the highest degree of fit.

2. Revising patent and trademark prediction methods

The use of accurate and effective prediction models is of great importance for any patent and trademark office, at the levels of both the individual country and at the supranational intellectual property entities such as WIPO and EPO. The prediction methods examined in this study are used by some or all of the following patent offices: the European Patent Office (EPO), the World Intellectual Property Office (WIPO), the United States Patents and Trademarks Office (USPTO), the Japanese Patent Office (JPO), the Korean Intellectual Property Office (KIPO) and the Swiss Federal Institute of Intellectual Property (SFIIP) (Table 2).

The set of prediction techniques identified can be classified as either quantitative or qualitative. The quantitative prediction techniques are as follows:

- Simple extrapolation of trends based on the projection into the future of average growth rates in patent and trademark applications detected in recent time periods.
- Linear and quadratic trend fit methods with different time horizons. These methods are often combined to form prediction clouds so as to select the final prediction on the basis of the average predictions estimated using different types of models.
- Exponential smoothing methods such as Holt-Winters, whether simple or for damped trend. These methods, with semi-automatic estimation of parameters, are usually a first step in the analysis of time series for later application of more complex models.
- Advanced time series models such as AR, ARMA and ARIMA.
- Methods of structural state-space models with unobserved components and Kalman filters similar to those calculated using STAMP (Structural Time Series Analyzer, Modeler and Predictor) software.
- Econometric methods which include time series with regressors and which use exogenous variables, also including the

Table 2
Summary of patent and trademark prediction models.

Agency	Prediction models	Remarks
World Intellectual Property Office (WIPO)	Combination of time series, econometric and survey-based models.	There is a wide range of prediction models proposed by the WIPO for calculating PCT patents, and for coordination between EPO, USPTO and JPO.
European Patent Office (EPO)	Simple trend models, time series and econometric models, first application transfer models. ARIMA transfer function methods. Surveys amongst clients and consensus of experts.	Use of linear and quadratic models with different horizons. They perform the model cloud fit and use a domestic patent transfer model (first filings). Surveys amongst applicants and consensus of experts are carried out periodically.
United States Patents and Trademarks Office (USPTO)	Simple trend models, exponential smoothing models, ARIMA and econometric models with regressors. Client surveys and Delphi method.	Use of Forecast Pro software for time series models (annual and quarterly predictions). Regressors: price of patent application and R&D mobile media. Use of surveys and Delphi of experts on a recurring basis over time.
Japanese Office of Patents and Trademarks (JPO)	Client surveys and time series prediction models.	They use surveys amongst applicants and companies according to size and perform propensity calculations to patent on the basis of these surveys.
Swiss Federal Institute of Intellectual Property (SFIIP)	Structural models in state-space from with and without regressors.	Use of STAMP software for the analysis of time series. Prediction of trademarks using the variables Dow Jones, SMI, SPI and Swiss consumer index as regressors.
Korean Intellectual Property Office (KIPO)	Model of trend extrapolation based on the application of average inter-annual growth rates to the recent data.	Recently a consensus was reached with the JPO, the Office for Harmonization in the Internal Market (OHIM) and the USPTO to share statistical indicators and prediction models.