



## TECHNOLOGY FORECASTING VIA PUBLISHED PATENT APPLICATIONS AND PATENT GRANTS

Dar-Zen Chen

*Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.*

Chang-Pin Lin

*Department of Mechanical and Mechatronic Engineering, National Taiwan Ocean University, Keelung, Taiwan, R.O.C., b0195@mail.ntou.edu.tw*

Mu-Hsuan Huang

*Department of Library and Information Science, National Taiwan University, Taipei, Taiwan, R.O.C.*

Yi-Tung Chan

*Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.*

Follow this and additional works at: <https://jmstt.ntou.edu.tw/journal>



Part of the [Engineering Commons](#)

### Recommended Citation

Chen, Dar-Zen; Lin, Chang-Pin; Huang, Mu-Hsuan; and Chan, Yi-Tung (2012) "TECHNOLOGY FORECASTING VIA PUBLISHED PATENT APPLICATIONS AND PATENT GRANTS," *Journal of Marine Science and Technology*. Vol. 20: Iss. 4, Article 1.

DOI: 10.6119/JMST-011-0111-1

Available at: <https://jmstt.ntou.edu.tw/journal/vol20/iss4/1>

This Research Article is brought to you for free and open access by Journal of Marine Science and Technology. It has been accepted for inclusion in Journal of Marine Science and Technology by an authorized editor of Journal of Marine Science and Technology.

---

## TECHNOLOGY FORECASTING VIA PUBLISHED PATENT APPLICATIONS AND PATENT GRANTS

### Acknowledgements

The authors acknowledge the supports from the National Science Council, Taiwan, with grant no. NSC-98-3114-P-002-005.

# TECHNOLOGY FORECASTING VIA PUBLISHED PATENT APPLICATIONS AND PATENT GRANTS

Dar-Zen Chen<sup>1</sup>, Chang-Pin Lin<sup>2</sup>, Mu-Hsuan Huang<sup>3</sup>, and Yi-Tung Chan<sup>1</sup>

Key words: technology forecasting, published patent application, patent grants.

## ABSTRACT

The objective of this research is to study and establish the relationships between published patent applications and patent grants for exploring the technology development trend on a specific technology since more and more patents have their applications published before they are granted. Two modeling algorithms based on the patent grant/publish ratio as well as one long-term modeling algorithm based on the average publish-to-grant lag, were developed accordingly. The relationships between patent grants and published patent applications were constructed through two case studies on Magnetic Random Access Memory and Organic Light-Emitting Diode technologies and corresponding forecasts were then conducted. Comparing to the traditional time-series Autoregressive Integrated Moving Average method, the predicting power of the modeling algorithms based on the patent grant/publish ratio was satisfactory. On the other hand, the modeling algorithm based on the characteristic of average publish-to-grant lag has shown superior predicting power. Results from these two applications help us to validate the proposed methods and appropriate tools for forecasting the patent grants.

## I. INTRODUCTION

Technology forecasting is defined as the prediction of future developments on a particular technology. It is conducted to support decision makings by identifying technology bottlenecks, establishing feasible rates of progress, providing references and warning signals, and indicating achievable

alternatives. Many different technology forecasting methods have been developed and addressed accordingly. For example, Martino [15] showed that computer and mathematical modeling was the most commonly used approach, with Delphi second and Trend extrapolation third. Porter and Rossini [23] suggested that all forecasting techniques fit into five families: Monitoring, Expert Opinion, Trend Analysis, Modeling, and Scenarios. Assuming that technology changes can be explained by factors, Porter *et al.* [22] categorized technology forecasting methods as to whether they are direct, correlative, or structural. Levary and Han [13] presented most popular methods: Delphi, Nominal group process, the case study method, Growth curve, Trend analysis, Correlation analysis, Analytic hierarchy process, Cross impact analysis, Relevance trees, Scenario writing, and System dynamics. Martino [16] concluded that the passage of 30 years has seen little change in the most popular methods of technology forecasting, and Porter [21] considered recent methodological contributions in terms of six prominent methodological approaches to technology forecasting: Creativity methods, Monitoring, Trend analysis, Modeling, Expert opinion, and Scenarios. More recently Martino [17] again presented developments in Environmental scanning, Models, Scenarios, Delphi, Extrapolation, Probabilistic forecasts, Technology measurement, and Chaos theory.

As mentioned above, based on many technology forecasting methods some researcher studied appropriateness of technology forecasting for different kinds of technologies. However, it is difficult to forecast emerging technologies as there is little historical data available. Therefore, some experts used patent data to derive information about a particular industry or technology in forecasting. A patent is the property right to a knowledge asset. Bibliographic data of patent provides enormous information and is accessible with low cost which represents a comprehensive, in-depth technological activity information resource. Just as other similar bibliographic sources of data, patent data offer a lot of methodological and technical advantages. Patent systems were developed to store enormous and continuously expanding patent data, and are easy to be accessed and analyzed. Hence, patent data can be effectively used to provide indispensable technological information. Furthermore, recognition is given to the theory that

---

Paper submitted 06/12/09; revised 12/23/10; accepted 01/11/11. Author for correspondence: Chang-Pin Lin (e-mail: b0195@mail.ntou.edu.tw).

<sup>1</sup>Department of Mechanical Engineering and Institute of Industrial Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

<sup>2</sup>Department of Mechanical and Mechatronic Engineering, National Taiwan Ocean University, Keelung, Taiwan, R.O.C.

<sup>3</sup>Department of Library and Information Science, National Taiwan University, Taipei, Taiwan, R.O.C.

using patent statistics as a technology indicator successfully measures technology changes and developments. Various researchers have conducted studies by using patent statistics to illustrate the process of innovation and technology changes [2, 6, 14, 20, 25, 26]. Campbell [3] showed that patent indicators provide a very useful forecasting tool for decision makers in the public and private sectors. Moguee [18] concluded that statistical analysis of international patent records is a valuable tool for corporate technological analysis and planning. Ernst [5] used an S-shape curve to forecast the development of CNC-technology and assessed the suitability of patent data for forecasting technological developments. Palmer *et al.* [19] showed that the Fish-Pry model is fitted well in the electronics industry and rates of technological progress are correlated to numbers of patents filed. Numerous other researchers have also presented valuable work on correlating patent numbers to technology developments [1, 4, 7, 9-11].

There are many possible factors relevant to the development of patent grants on a specific technology, such as: number of inventors, number of patent applications, granting procedure alterations, workload and budget approval at patent office, and etc. However, due to changes of the publication practice of the USPTO, the published patent applications have become a distinct indicator for possible patent grants since more and more patents have their application published before they are granted. The USPTO calculated an average total pending period, which measures the average time in months from filing until the application is issued, to be 21.1 months in 2005 [28]. Understanding the trend of patent grants for a specific technology is vital and common practice for resource planning at companies making products with that technology. Heyman [8] found a basic competitive monitoring program that can be established by searching public databases of granted patents and published patent applications at a regular interval. Therefore, the number of published patent applications on a specific technology reflects earlier the significance of that technology than the number of granted patents. Therefore, it is our goal to establish the relationships between published patent applications and patent grants for exploring the technology developments, especially on the newly developed technologies.

A patent application will be published after the expiration of an 18-month period following the earliest effective filing date or a prior date claimed by the applicant. Following the publishing, the application of the patent is no longer held in confidence by the Office and any member of the public may request access to the entire file history of the application. Published patent applications provide a preview of soon-to-come patents and reveal the technology in advance. Ragusa [24] concluded that the study of published patent applications permits industry to develop improvements upon the published technologies and further stimulates innovations within that industry. Moreover, the published application system will help inventors and the corporations that often support them by reducing the needless supplication of research efforts which

waste time and money. Kotabe [12] compared patent systems established at the United States and Japan. Accordingly he suggested that US firms could access Japanese patent applications published within eighteen months of filing to keep abreast of innovations originating from Japan. Silverman [27] showed that the benefits of giving notice to potential infringers after publishing may be sought.

In this paper, we focus on studying and establishing the relationships between published patent applications and patents grants. Assuming that published application database is considered the pilot model of granted patent database, the number of published patent applications and the number of granted patents should be correlated. Two distinctive technologies were studied in this paper, Magnetic Random Access Memory (MRAM), a new memory technology that promises to provide non-volatile, low-power, high speed and low-cost memory. Although MRAM has many advantages over virtually every existing memory type, it is still in its infancy and will potentially generate considerable amount of patents for a long period of time. Organic Light-Emitting Diode (OLED) technology was invented by Eastman Kodak in the early 1980s. It is beginning to replace LCD technology in handheld devices such as PDAs and cellular phones because the technology is brighter, thinner, faster and lighter than LCDs. In addition, OLED uses less power, offer higher contrast and are cheaper to manufacture.

The results can be used to reflect a specific technology trend in order to capture the future in that industry, and decision makers can use this information to understand their competitive advantages. Two modeling algorithms were developed to predict the patent grants by modeling the trends of patents' grant/publish (G/P) ratios. Furthermore, an effective modeling algorithm for establishing long-term relationships between published patent applications and patents grants was developed based on the average publish-to-grant (P2G) lag. Finally two case studies on MRAM and OLED technologies by applying proposed modeling algorithms were detailed and discussed.

## II. RESEARCH METHODS

Patent applications are automatically published 18 months after their effective filing date (or earlier, if requested by the applicants). By analyzing the granted patents with previously published applications, we assume that the number of granted patents and the number of published patent applications are correlated.

The USPTO granted patent and published application databases were chosen to establish models for granted patent number. Two important aspects have to be considered during the modeling process.

- a. Published patent application database is available only after March 2001, therefore, patent search of granted patent database should begin then.

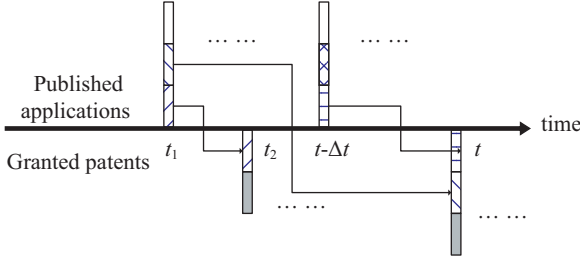


Fig. 1. Different time lags between patent granted and published dates along a timeline.

b. Patent applications published after they were granted should be disregarded.

In this study, two modeling algorithms, ‘(1) model based on G/P ratios at  $t'$  and ‘(2) model based on G/P ratios across  $\Delta t'$ , were developed to predict the patent grants by modeling the trends of patents’ G/P ratios. Furthermore, an effective modeling algorithm for establishing long-term relationships between published patent applications and patents grants was developed based on the average P2G lag  $\Delta t$ . The time frames defined for granted patents with and without previously published applications are illustrated as in Fig. 1. The P2G lag  $\Delta t$  is defined as the period between granted date and published date of a patent.

1. Modeling by G/P Ratios

The number of patents granted at time  $t$  with previously published applications can be obtained by summing up the number of patents granted at time  $t$  with previous applications published at time  $t-\Delta t$  for  $\Delta t$  from 1 to  $t-t_1$  and is written as:

$$G_p^t = \sum_{\Delta t=0}^{t-t_1} G_{p(\Delta t)}^t \tag{1}$$

The total number of patents granted at time  $t$  can be expressed by the sum of the numbers of patents granted at time  $t$  with and without previously published applications:

$$G^t = G_p^t + G_n^t \tag{2}$$

The number of patent applications published at time  $t$  ( $t_1 \leq t$ ) is expressed as:

$$P^t = \sum_{i=0}^{\Delta t} P_{g(i)}^t + P_{n(i)}^t \tag{3}$$

To establish relationships between the number of published patent applications and the number of granted patents with previously published applications, a G/P ratio of the number of patents granted at time  $t$  with previous applications published at time  $t-\Delta t$  to the total number of patent applications published at time  $t-\Delta t$  is defined as:

Table 1. Patent G/P ratio Matrix.

$\Delta t/i$	$t_2$	$t_2 + 1$	$\dots$	$t - 1$	$T$
0	$z_{p(0)}^{t_2}$	$z_{p(0)}^{t_2+1}$	$\dots$	$z_{p(0)}^{t-1}$	$z_{p(0)}^t$
1	$z_{p(1)}^{t_2}$	$z_{p(1)}^{t_2+1}$	$\dots$	$z_{p(1)}^{t-1}$	$z_{p(1)}^t$
2	$z_{p(2)}^{t_2}$	$z_{p(2)}^{t_2+1}$	$\dots$	$z_{p(2)}^{t-1}$	$z_{p(2)}^t$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t_2 - t_1 - 1$	$z_{p(t_2-t_1-1)}^{t_2}$	$z_{p(t_2-t_1-1)}^{t_2+1}$	$\dots$	$z_{p(t_2-t_1-1)}^{t-1}$	$z_{p(t_2-t_1-1)}^t$
$t_2 - t_1$	$z_{p(t_2-t_1)}^{t_2}$	$z_{p(t_2-t_1)}^{t_2+1}$	$\dots$	$z_{p(t_2-t_1)}^{t-1}$	$z_{p(t_2-t_1)}^t$
$t_2 - t_1 + 1$		$z_{p(t_2-t_1+1)}^{t_2+1}$		$z_{p(t_2-t_1+1)}^{t-1}$	$z_{p(t_2-t_1+1)}^t$
$\vdots$			$\ddots$	$\vdots$	$\vdots$
$t - t_1 - 1$				$z_{p(t-t_1-1)}^{t-1}$	$z_{p(t-t_1-1)}^t$
$t - t_1$					$z_{p(t-t_1)}^t$

$$z_{p(\Delta t)}^t = G_{p(\Delta t)}^t / P^{t-\Delta t} \tag{4}$$

In general, forecasting uses an established model to predict the objective values for the upcoming time step. Two algorithms have been developed for modeling and forecasting the number of patents granted at the next time step: model based on G/P ratios at  $t$ ; and model based on G/P ratios across  $\Delta t$ . A G/P ratio matrix with row of  $\Delta t$  from 0 to  $t-t_1$  and column of time from  $t_2$  to  $t$  is shown as in Table 1.

Prior to executing these two algorithms, three actions listed below have to be completed to prepare proper data files from patent databases chosen.

- Action 1: Choose a target technology.
- Action 2: Select a suitable patent database. Patent databases in USPTO, EPO, and JPO are the most commonly used. In this research patent database in USPTO is chosen.
- Action 3: Execute a complete patent search, and then compile patent statistics depending on the time step used; i.e., week, month, season, or year.

The two algorithms are then executed by the following steps.

A. Model based on G/P ratios at  $t$ :

By using Eqs. (1)-(4), a G/P ratio matrix  $Z$  as shown in Table 2 could be obtained from the existing patent data. The predicted array of G/P ratios at time  $t+1$  could then be calculated as a function of  $k$  given as:

$$\tilde{D}_{t+1} = f(D_{t+1-1}, \dots, D_{t+1-k}) \tag{5}$$

Since the array of G/P ratios at time  $t+1$  is predicted, the predicted number of patents granted at time  $t+1$  with previously published applications could be calculated by multiplying the predicted array of G/P ratios at time  $t+1$  to the corresponding numbers of published patent applications:

**Table 2. Statistical analysis results for MRAM.**

Variable	Triple-Point Model F-ratio 9.14		Single-point Model F-ratio 29.41	
	Parameter Estimate	T for H <sub>0</sub>	Parameter Estimate	T for H <sub>0</sub>
Intercept	0.360660	0.16	-0.552383	-0.32
P <sup>t-10</sup>	0.290815	1.04	-	-
P <sup>t-9</sup>	0.484059	1.64	1.186026	5.83
P <sup>t-8</sup>	0.094235	0.33	-	-
P <sup>t-102</sup>	-0.012492	-1.13	-	-
P <sup>t-9</sup> P <sup>t-10</sup>	-0.001562	-0.10	-	-
P <sup>t-92</sup>	-0.012273	-1.68	-0.019466	-3.90
P <sup>t-8</sup> P <sup>t-10</sup>	0.020630	1.23	-	-
P <sup>t-8</sup> P <sup>t-9</sup>	0.012187	0.79	-	-
P <sup>t-82</sup>	-0.012016	-1.22	-	-

$$\tilde{G}_p^{t+1} = \sum_{\Delta t=0}^{t+1-t_1} \tilde{D}_{t+1}(\Delta t) \times P^{t+1-\Delta t} \quad (6)$$

Using the regression analysis, the relationship between the predicted number of granted patents with previously published applications and the actual number of granted patents with previously published applications from time  $t_2+1$  to  $t$  is established as:

$$G_p^{t+1} = \tilde{G}_p^{t+1} + \Delta e_{t+1} \quad (7)$$

Ratios of the number of patents granted at time  $t$  with previously published applications to the number of all patents granted at time  $t$  could be expressed by:

$$X^t = G_p^t / G^t \quad (8)$$

By analyzing the ratios from time  $t_2$  to  $t$ , the ratio at time  $t+1$  could be established. The ratio at time  $t+1$  could be represented as a function of previous ratios:

$$X^{t+1} = f(X^{t_2}, \dots, X^t) \quad (9)$$

Finally, the number of patents granted at time  $t+1$  could be calculated by:

$$G^{t+1} = G_p^{t+1} / X^{t+1} \quad (10)$$

**B. Model based on G/P ratios across  $\Delta t$ :**

From Eq. (4) and Table 2, ratios of the number of patents granted at time  $i$  ( $t_2 \leq i \leq t$ ) with previous applications published at time  $i-\Delta t$  to the number of patent applications published at time  $i-\Delta t$  is calculated. For each P2G lag  $\Delta t$  ( $0 \leq \Delta t \leq$

$t-t_1$ ), the trend of individual row in matrix  $Z$  could be found. The predicted array of G/P ratios at time  $t+1$  could be decomposed by  $\Delta t$ . Each element of the predicted array of G/P ratios at time  $t+1$  could be estimated by the trend of the corresponding row. The predicted array of G/P ratios at time  $t+1$  could be represented by function of each element as:

$$\tilde{z}_{p(\Delta t)}^{t+1} = f(z_{p(\Delta t)}^{t_2}, \dots, z_{p(\Delta t)}^t) \quad (11)$$

Eq. (11) and the moving average algorithm were applied and the period of moving average is chosen to be 5. The average value can not be calculated when the number of the data is less than 5. Eq. (11) can be rewritten as:

$$\tilde{z}_{p(\Delta t)}^{t+1} = \frac{1}{5} \sum_{n=t-4}^t z_{p(\Delta t)}^n \quad \text{for } \Delta t = 0, \dots, t+1-t_1-5 \quad (12a)$$

$$\tilde{z}_{p(\Delta t)}^{t+1} = 0 \quad \text{for } \Delta t = t+1-t_1-4, \dots, t+1-t_1 \quad (12b)$$

Since the predicted array of G/P ratios at time  $t+1$  is calculated, the predicted number of patents granted at time  $t+1$  with previously published applications could be calculated by:

$$\tilde{G}_p^{t+1} = \sum_{\Delta t=0}^{t+1-t_1} \tilde{z}_{p(\Delta t)}^{t+1} \times P^{t+1-\Delta t} \quad (13)$$

The predicted array of G/P ratios at time  $i$  could be calculated by:

$$\tilde{z}_{p(\Delta t)}^i = f(z_{p(\Delta t)}^{t_2}, \dots, z_{p(\Delta t)}^{i-1}) \quad (14)$$

Using the regression analysis, the relationship between the predicted number of granted patents with previously published applications and the actual number of granted patents with previously published applications from time  $t_2+1$  to  $t$  is established. Same as in model based on G/P ratios at  $t$ , the number of patents granted at time  $t+1$  with previously published applications could be calculated by Eq. (7). The ratio at time  $t+1$  could be represented as the function of previous ratios in Eq. (9). Finally, the number of patents granted at time  $t+1$  could be calculated by Eq. (10).

**2. Modeling by Average P2G Lag**

In general, long-term forecasting predicts period for 5, 10, or more time steps. Traditional long-term forecasting methods are mostly non-quantitative, i.e. Delphi and Scenarios, or using extrapolation for extending periods of prediction, i.e. Trend Extrapolation. The number of granted patents at time  $t+1$  with corresponding published applications could be forecasted by the short-term forecasting method. The number of granted patents at time  $t+2$  with corresponding published applications could be analogized by assuming that it at time  $t+1$  is given. It takes the risk of the cumulative error to

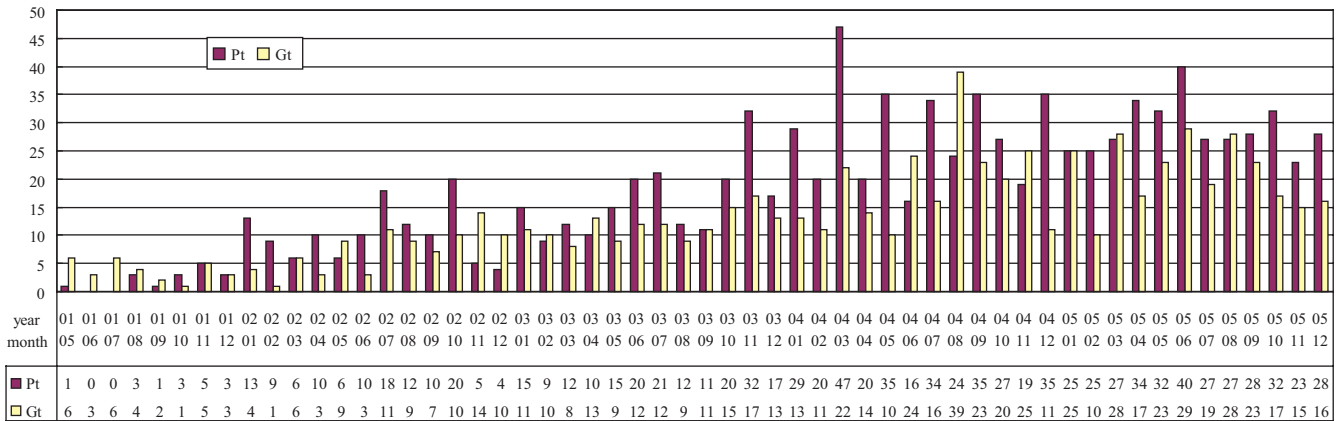


Fig. 2. Histogram of granted patents and published patent applications on MRAM.

practice long-term forecasting. Since the published application at time  $t-\Delta t$  issues at time  $t$ , we assume that the number of published applications at time  $t-\Delta t$  and the number of granted patents at time  $t$  with corresponding published applications are correlated. By analyzing the characteristic of  $\Delta t$ , the representative time or time interval to proceed long-term forecasting can be found.

By defining  $t_{lag}$  as the average P2G lag, the number of patent applications published at time  $t-t_{lag}$  and the number of patents granted at time  $t$  with previously published applications are verified for their correlations in some technologies.

### III. CASE STUDIES – MRAM & OLED

Two technologies were chosen, Magnetic Random Access Memory (MRAM) and (OLED), as the target technologies to be studied. In order to verify the patent data to be stationary time-series data, Augmented Dickey-Fuller (ADF) tests were conducted for both MRAM and OLED data sets. The results show that DF statistic value to be -6.3 for MRAM and -5.4 for OLED, and p-value to be 0.021 for MRAM and 0.037 for OLED. Therefore, the time-series data for both technologies can be considered to be stationary which requires more negative on DF statistic value and lesser on p-value.

#### 1. Case 1: MRAM

We first chose Magnetic Random Access Memory (MRAM) as the target technology to be studied. MRAM is a modern technology on memory device that uses electron spin to store information. MRAM was developed by potentially combining the density of DRAM with the speed of SRAM and non-volatility of FLASH memory or hard disk, along with consuming a very low amount of power. MRAM is a solid state device and, as such, has much greater speed and durability. It can resist high radiation, and can operate in extreme temperature conditions. Like conventional RAM, MRAM is composed of transistors but, instead of electrical charges, it uses magnetic charges to store information. Automotive applications using sensors can benefit from MRAM. Since sensors

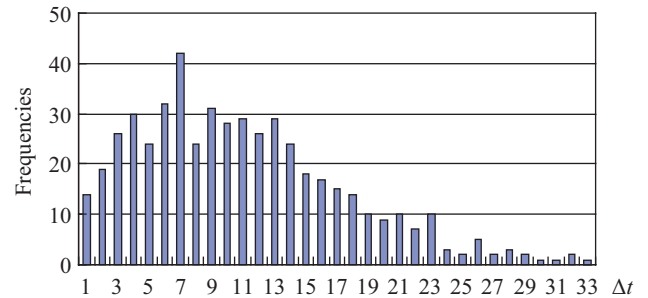


Fig. 3. Probability density function of P2G Lag for MRAM.

write data continuously, flash memories have difficulty keeping up with such data flow. New airbag systems also have sensors to detect and record passenger weight, interactions with other safety devices on the vehicle and the impact of collision. Further MRAM technology improvements can radically change embedded systems architecture. MRAM has the potential to replace RAM and flash memory used in embedded MCUs for data storage and program memory, respectively.

Patent data was drawn from the patent database for both U.S. granted and U.S. published application. The search data covers from March 2001 to December 2005. The number of granted patents found is 742 and the number of published patent applications is 1022. The histogram of granted patents and published patent applications on MRAM technology is shown in Fig. 2.

Fig. 3 shows the distribution of frequencies for  $\Delta t$ . The distribution is skew. So  $t_{lag}$  is chosen to be the median:  $t_{lag} = 9$ .

Running regression analysis, the relationship of the number of patent applications published at time  $t-9$  and the number of patents granted at time  $t$  with previously published applications can be established and the coefficient of determination was calculated to be 0.5721. The relationship of the number of patent applications published at time  $t-10$ ,  $t-9$ , and  $t-8$  and the number of patents granted at time  $t$  with previously published applications can be established and the coefficient of determination was calculated to be 0.6956.

Since the relationship of the number of patent applications





The number of patents granted at time  $t+1$  with previously published applications was then predicted by regression analysis. G/P ratios of patents for MRAM from time  $t_2$  to  $t$  could then be calculated and the G/P ratio at time  $t+1$  was obtained to be 88.34% by regression analysis. The number of patents granted at time  $t+1$  was calculated to be 16.76.

Next, we used model based on the array of G/P ratios across  $\Delta t$  to implement the forecast. The predicted array of G/P ratios at time  $t+1$  was calculated from Eq. (12):

$$\begin{aligned} \tilde{z}_{p(0)}^{t+1} &= \frac{1}{5} \sum_{n=52}^{56} z_{p(0)}^n = 0.024534 \\ \tilde{z}_{p(1)}^{t+1} &= \frac{1}{5} \sum_{n=52}^{56} z_{p(1)}^n = 0.014815 \\ &\vdots \\ \tilde{z}_{p(57)}^{t+1} &= 0 \\ \tilde{z}_{p(58)}^{t+1} &= 0 \end{aligned}$$

The predicted number of patents granted at time  $t+1$  with previously published applications was calculated from Eq. (13):

$$\tilde{G}_p^{t+1} = \tilde{z}_{p(0)}^{t+1} \times P^{t+1} + \tilde{z}_{p(1)}^{t+1} \times P^t + \dots + \tilde{z}_{p(58)}^{t+1} \times P^1 = 18.79$$

Similarly the predicted array of G/P ratios and the predicted number of granted patents with previously published applications from time  $t_2+5$  to  $t$  were calculated.

The number of patents granted at time  $t+1$  with previously published applications was then predicted by regression analysis. G/P ratios of patents for MRAM from time  $t_2$  to  $t$  could be calculated and the G/P ratio at time  $t+1$  was then obtained to be 88.34% by regression analysis. The number of patents granted at time  $t+1$  was calculated to be 19.52.

Furthermore, we have applied ARIMA to predict the number of MRAM patent granted at a later time. The most popular time series method of is Box-Jenkins method. The class of models used is the autoregressive integrated moving averages (ARIMA) processes. The Box-Jenkins modeling approach suggests that first differencing of the data was appropriate. ARIMA (1,1,1) was applied to forecast the number of patents granted at time  $t+1$  with previously published applications based on the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots on MRAM technology which show significant spike only at lag 1 as in Fig. 5.

2) Modeling by Average P2G Lag

Running regression analysis, the relationship of the number of patent applications published at time  $t-9$  and the number of patents granted at time  $t$  with previously published applications can be established from Table 2 as below and the coefficient of determination was calculated to be 0.5721.

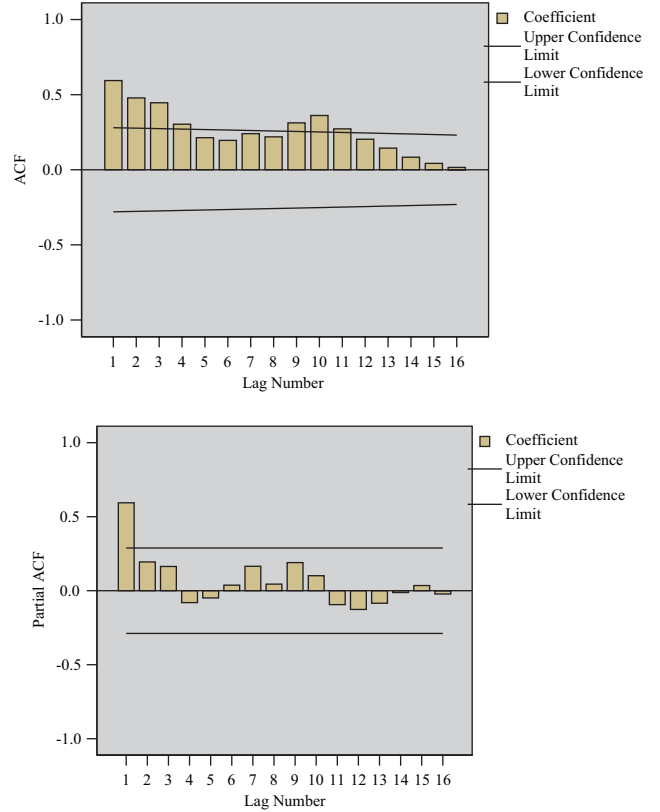


Fig. 5. ACF and PACF plots for the time series of granted patents on MRAM.

$$G_p^t = -0.552 + 1.186P^{t-9} - 0.019P^{t-9^2}$$

From the statistic analysis results as shown in Table 2 with sample size 56, the standard values in  $F$  distribution and  $t$  distribution can be found from their corresponding tables to be:  $F_{0.05}(2,52) = 3.196$  and  $t_{0.05}(55) = 1.675$ . The computed  $F$  value for this model is 29.41 which means this model is statistically applicable for this relationship of the number of patent applications published at time  $t-9$  and the number of patents granted at time  $t$  with previously published applications. The most dominant patent variables are the number of patent applications published at time  $t-9$  and its squared value.

The relationship of the number of patent applications published at time  $t-10$ ,  $t-9$ , and  $t-8$  and the number of patents granted at time  $t$  with previously published applications can be established from Table 2 as:

$$\begin{aligned} G_p^t &= 0.361 + 0.291P^{t-10} + 0.484P^{t-9} + 0.094P^{t-8} \\ &\quad - 0.002P^{t-10}P^{t-9} + 0.021P^{t-10}P^{t-8} + 0.012P^{t-9}P^{t-8} \\ &\quad - 0.012P^{t-10^2} - 0.012P^{t-9^2} \\ &\quad - 0.012P^{t-8^2} \end{aligned}$$

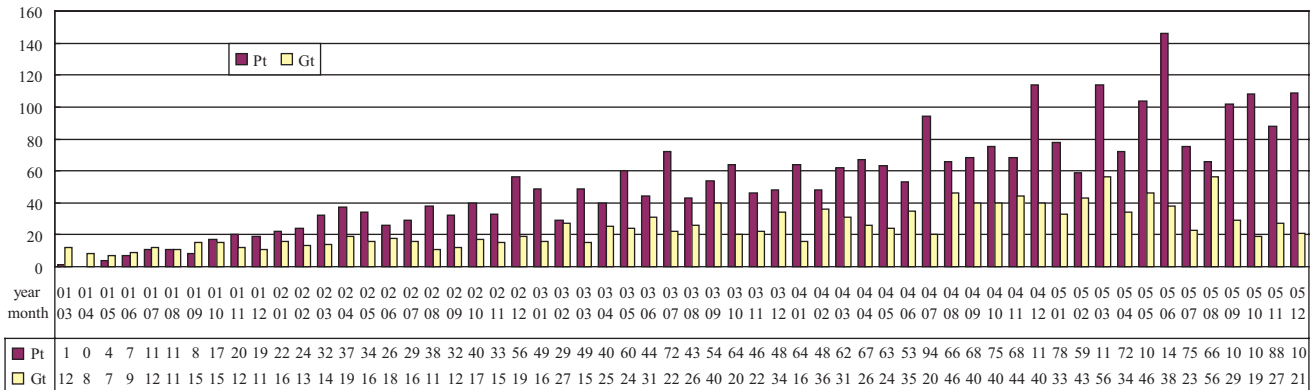


Fig. 7. Histogram of granted patents and published patent applications on OLED.

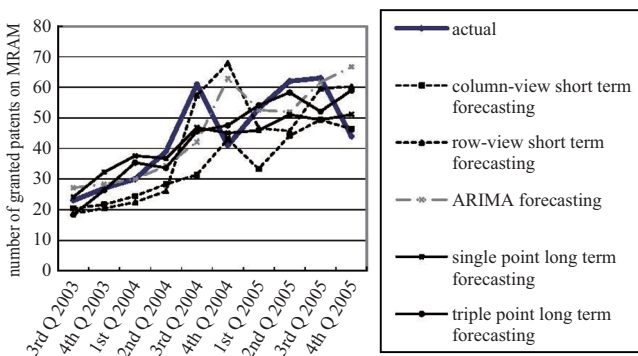


Fig. 6. Actual and predicted numbers of granted patents on MRAM.

The coefficient of determination was calculated to be 0.6956. Since they are correlated, the predicted number of granted patents from time  $t+1$  to  $t+9$  with previously published applications can be calculated. From the statistic analysis results as shown in Table 2 with sample size 54, the standard values in  $F$  distribution and  $t$  distribution can be found from their corresponding tables to be:  $F_{0.05}(9,44) = 2.14$  and  $t_{0.05}(53) = 1.676$ . The computed  $F$  value for this model is 9.14 which means this model is statistically applicable for this relationship of the number of patent applications published at time  $t-10$ ,  $t-9$ , and  $t-8$  and the number of patents granted at time  $t$  with previously published applications. The most dominant patent variable is the squared value of the number of patent applications published at time  $t-9$ .

The actual number of granted patents on MRAM from 2003 to 2005 were shown in Fig. 6, associated with numbers predicted by ARIMA time series method and other four forecasting methods developed in this paper.

2. Case 2: OLED

Next, we chose OLED as the target technology to be studied. OLED is a flat display technology, made by placing a series of organic thin films between two charged electrodes, one a metallic cathode and one a transparent anode, usually being glass. They operate on the attraction between positively

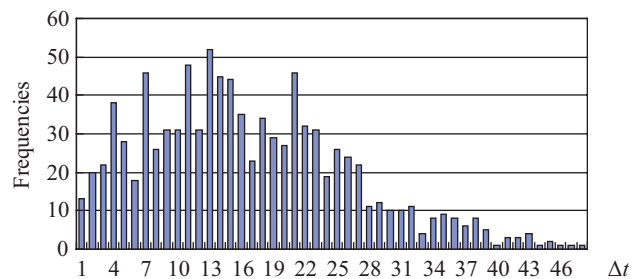


Fig. 8. Probability density function of P2G Lag for OLED.

and negatively charged particles. When electrical current is applied, one layer becomes negatively charged relative to another transparent layer. As energy passes from the negatively charged (cathode) layer to the other (anode) layer, it stimulates organic material between the two, which emits light visible through an outermost layer of glass. OLED technology enables ultra-thin, flexible or transparent displays with brighter screens and a fuller viewing angle, and makes very durable displays that can operate in a broader temperature range.

Patent data was drawn from the patent database for both U.S. granted and U.S. published application. The search data covers from March 2001 to December 2005. The number of granted patents found is 1423 and the number of published patent applications is 3062. The histogram of granted patents and published patent applications on OLED technology is shown in Fig. 7. Fig. 8 shows the distribution of frequencies for  $\Delta t$ . The distribution is skew. So  $t_{lag}$  is chosen to be the median:  $t_{lag} = 14$ .

Running regression analysis, the relationship of the number of patent applications published at time  $t-14$  and the number of patents granted at time  $t$  with previously published applications can be established and the coefficient of determination was calculated to be 0.5884. The relationship of the number of patent applications published at time  $t-15$ ,  $t-4$ , and  $t-13$  and the number of patents granted at time  $t$  with previously published applications can be established and the coefficient of determination was calculated to be 0.6906.

The histogram of granted patents and published patent

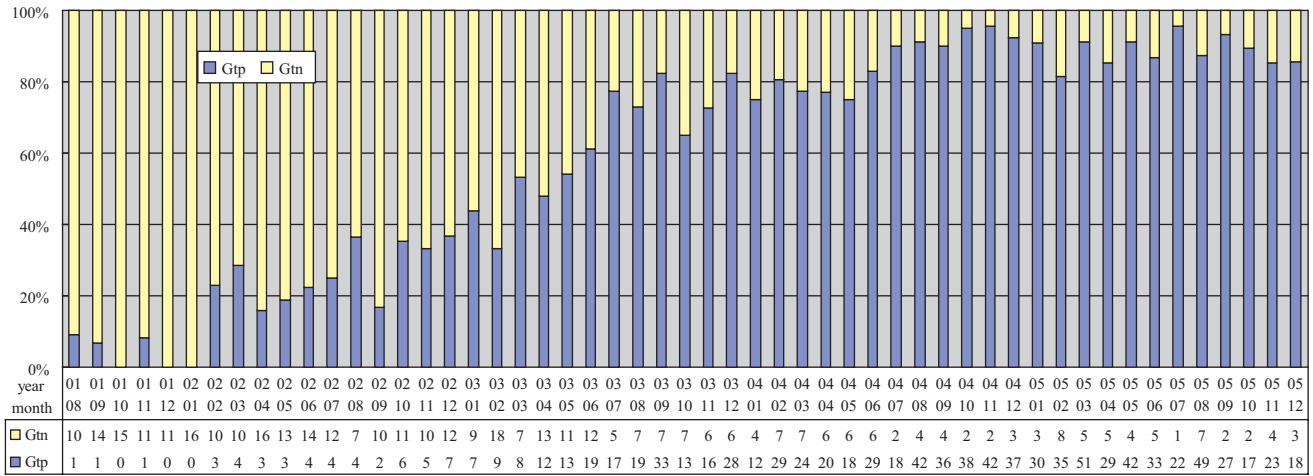


Fig. 9. Granted patents with and without previously published applications on OLED.

applications on OLED technology is shown in Fig. 6. The number of granted patents with previously published applications and the number of granted patents without previously published applications are shown in Fig. 8. The total number of granted patents with previously published applications is 963. There are 3 patents with publishing date after granting date. When we calculate the average P2G lag, they are ignored. From Fig. 9, we found that the G/P ratio of patents for OLED is growing and approaching unity. This ratio ranges from 0% to 95.7% and averages 60.1%.

1) Modeling by G/P Ratios

For OLED technology, the date of first published application being granted,  $t_1$ , is March 2001 and the date of first granted patent with a previously published application,  $t_2$ , is September 2001. From Eqs. (1) to (4), the G/P ratio matrix Z was built.

Model based on the array of G/P ratios at  $t$  or model based on the array of G/P ratios across  $\Delta t$  forecasting methods were used to obtain the number of patents granted at time  $t+1$  and the results were compared to traditional short-term forecasting method: ARIMA. First, we use model based on the array of G/P ratios at  $t$  to implement the forecast.

The array of individual column was established by a 2nd-order polynomial and the predicted array of G/P ratios at time  $t+1$  was built from Eq. (5):

$$\tilde{D}_{t+1}(\Delta t) = -1.3013 \times 10^{-5} + 0.0007142\Delta t - 0.00045303\Delta t^2$$

The predicted number of patents granted at time  $t+1$  with previously published applications was calculated from Eq. (6):

$$\tilde{G}_p^{t+1} = D_t(0) \times P^{t+1} + \dots + D_t(t+1-t_1) \times P^{t_1} = 19.27$$

Similarly the predicted array of G/P ratios  $s$  from time  $t_2+1$  to  $t$  were expressed as:

$$\begin{aligned} \tilde{D}_t(\Delta t) &= -1.4839 \times 10^{-5} + 0.00069467\Delta t - 0.0029436\Delta t^2 \\ &\vdots \\ \tilde{D}_{t_2+1}(\Delta t) &= -0.0089286 + 0.080357\Delta t - 0.10714\Delta t^2 \end{aligned}$$

And the predicted number of granted patents with previously published applications from time  $t_2+1$  to  $t$  were calculated as:

$$\begin{aligned} \tilde{G}_p^t &= D_t(0) \times P^t + \dots + D_t(t-t_1) \times P^{t_1} = 24.52 \\ &\vdots \\ \tilde{G}_p^{t_2+1} &= D_t(0) \times P^{t_2+1} + \dots + D_t(t_2+1-t_1) \times P^{t_1} = 1.59 \end{aligned}$$

The number of patents granted at time  $t+1$  with previously published applications was then predicted by regression analysis. G/P ratios of patents for OLED from time  $t_2$  to  $t$  could then be calculated and the G/P ratio at time  $t+1$  was obtained to be 91.44% by regression analysis. The number of patents granted at time  $t+1$  was calculated to be 25.39.

Next, we used model based on the array of G/P ratios across  $\Delta t$  to implement the forecast. The predicted array of G/P ratios at time  $t+1$  was calculated from Eq. (12):

$$\begin{aligned} \tilde{z}_{p(0)}^{t+1} &= \frac{1}{5} \sum_{n=54}^{58} z_{p(0)}^n = 0 \\ \tilde{z}_{p(1)}^{t+1} &= \frac{1}{5} \sum_{n=54}^{58} z_{p(1)}^n = 0.0022727 \\ &\vdots \\ \tilde{z}_{p(57)}^{t+1} &= 0 \\ \tilde{z}_{p(58)}^{t+1} &= 0 \end{aligned}$$

The predicted number of patents granted at time  $t+1$  with previously published applications was calculated from Eq. (13):

**Table 3. Statistical analysis results for OLED.**

Variable	Triple-Point Model <i>F</i> -ratio 8.18		Single-point Model <i>F</i> -ratio 29.30	
	Parameter Estimate	T for H <sub>0</sub>	Parameter Estimate	T for H <sub>0</sub>
<i>Intercept</i>	-2.974892	-0.59	-3.668402	-0.95
$P^{t-10}$	0.559064	1.11	-	-
$P^{t-9}$	0.749403	1.40	1.011020	5.09
$P^{t-8}$	-0.367140	-0.61	-	-
$P^{t-10^2}$	0.000992	0.11	-	-
$P^{t-9}P^{t-10}$	-0.009485	-0.68	-	-
$P^{t-9^2}$	0.001114	0.16	-0.007269	-3.12
$P^{t-8}P^{t-10}$	-0.004744	-0.26	-	-
$P^{t-8}P^{t-9}$	-0.005602	-0.38	-	-
$P^{t-8^2}$	0.010634	1.33	-	-

$$\tilde{G}_p^{t+1} = \tilde{z}_{p(0)}^{t+1} \times P^{t+1} + \tilde{z}_{p(1)}^{t+1} \times P^t + \dots + \tilde{z}_{p(58)}^{t+1} \times P^1 = 30.283$$

Similarly the predicted array of G/P ratios and the predicted number of granted patents with previously published applications from time  $t_2+5$  to  $t$  were calculated.

The number of patents granted at time  $t+1$  with previously published applications was then predicted by regression analysis. G/P ratios of patents for OLED from time  $t_2$  to  $t$  could be calculated and the G/P ratio at time  $t+1$  was then obtained to be 91.44% by regression analysis. The number of patents granted at time  $t+1$  was calculated to be 31.59.

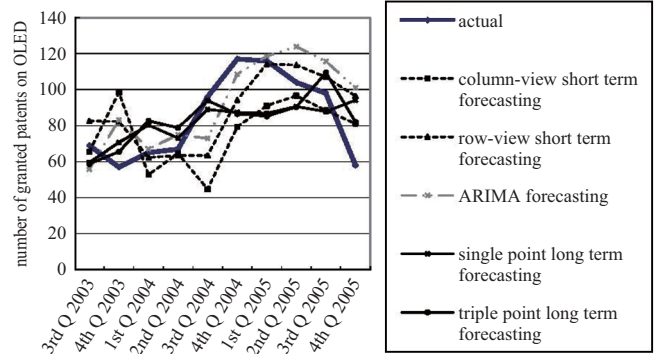
Same as in previous case study, we have applied ARIMA to predict the number of MRAM patent granted at a later time. ARIMA (1,1,1) was also used to forecast the number of patents granted at time  $t+1$  with previously published applications based on the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots on OLED technology which show significant spike only at lag 1 as in the MRAM case.

2) Modeling by Average AG2P Lag

Running regression analysis, the relationship of the number of patent applications published at time  $t-14$  and the number of patents granted at time  $t$  with previously published applications can be established from Table 3 as below and the coefficient of determination was calculated to be 0.5884.

$$G_p^t = -3.668 + 1.011P^{t-14} - 0.007P^{t-14^2}$$

From the statistic analysis results as shown in Table 3 with sample size 56, the standard values in  $F$  distribution and  $t$  distribution can be found from their corresponding tables to be:



**Fig. 10. Actual and predicted number of granted patents on OLED.**

$F_{0.05}(2,52) = 3.196$  and  $t_{0.05}(55) = 1.675$ . The computed  $F$  value for this model is 29.30 which means this model is statistically applicable for this relationship of the number of patent applications published at time  $t-9$  and the number of patents granted at time  $t$  with previously published applications. The most dominant patent variables are the number of patent applications published at time  $t-9$  and its squared value.

The relationship of the number of patent applications published at time  $t-15$ ,  $t-14$ , and  $t-13$  and the number of patents granted at time  $t$  with previously published applications can be established from Table 3 as:

$$G_p^t = -2.975 - 0.367P^{t-13} + 0.749P^{t-14} + 0.559P^{t-15} - 0.006P^{t-13}P^{t-14} - 0.005P^{t-13}P^{t-15} - 0.009P^{t-14}P^{t-15} + 0.011P^{t-13^2} + 0.001P^{t-14^2} + 0.001P^{t-15^2}$$

The coefficient of determination was calculated to be 0.6906. Since they are correlated, the predicted number of granted patents from time  $t+1$  to  $t+14$  with previously published applications can be calculated. From the statistic analysis results as shown in Table 3 with sample size 54, the standard values in  $F$  distribution and  $t$  distribution can be found from their corresponding tables to be:  $F_{0.05}(9,44) = 2.14$  and  $t_{0.05}(53) = 1.676$ . The computed  $F$  value for this model is 8.18 which means this model is statistically applicable for this relationship of the number of patent applications published at time  $t-10$ ,  $t-9$ , and  $t-8$  and the number of patents granted at time  $t$  with previously published applications. The most dominant patent variable is the squared value of the number of patent applications published at time  $t-9$ .

The actual number of granted patents on OLED from 2003 to 2005 were shown in Fig. 10, associated with numbers predicted by ARIMA time series method and other four forecasting methods developed in this paper.

**IV. DISCUSSION AND CONCLUSION**

In the first case study, the G/P ratio for MRAM patents is growing and approaching unity. Therefore, we conclude that

**Table 4. Comparison of the short-term forecasting models for MRAM.**

Model based on the array of G/P ratios at $t$	Model based on the array of G/P ratios across $\Delta t$		ARIMA (1,1,1)
7.42	6.46	RMSE	5.46
5.31	4.85	MAE	3.92
43.81%	33.74%	MAPE	34.37%

**Table 5. Comparison of the long-term forecasting models for MRAM.**

	Single point time interval ( $t_{lag} = 9$ )	3 points time interval ( $t_{lag} = 8,9,10$ )
RMSE	4.89	4.02
MAE	3.72	2.76
MAPE	40.62%	25.87%

more and more patents on MRAM technology have their applications published before they are granted. The relationship between the number of published patent applications and the number of granted patents is established and verified by means of our forecasting models and regression analysis. The number of patents granted at time  $t+1$  in short-term forecasting and the number of granted patents from time  $t+1$  to  $t+9$  with previously published applications in long-term forecasting were successfully predicted.

To compare the predictive power of each method, some indicators are available: root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Table 4 shows the comparison of the short-term forecasting models. In this case, the model based on the array of G/P ratios across  $\Delta t$  is better than the model based on the array of G/P ratios at  $t$  and ARIMA, particularly the relationship between published patent applications and granted patents is constructed. Table 5 shows the comparison of the long-term forecasting models. In this case, using 3 points time interval to proceed long-term forecasting is superior to using a single time point.

In second case study, the G/P ratio for OLED patents is also growing and approaching unity. Therefore, we obtained the same conclusion that more and more patents on OLED technology have their applications published before they are granted. The relationship between the number of published patent applications and the number of granted patents is also established and verified by means of our forecasting models and regression analysis. The number of patents granted at time  $t+1$  in short-term forecasting and the number of granted patents from time  $t+1$  to  $t+14$  with previously published applications in long-term forecasting were successfully predicted.

Table 6 shows the comparison of the short-term forecasting models. In this case, model based on the array of G/P ratios across  $\Delta t$  is better than model based on the array of G/P ratios

**Table 6. Comparison of the short-term forecasting models for OLED.**

	Model based on the array of G/P ratios at $t$	Model based on the array of G/P ratios across $\Delta t$	ARIMA (1,1,1)
RMSE	13.86	10.36	9.95
MAE	9.50	8.13	8.06
MAPE	42.47%	37.24%	36.62%

**Table 7. Comparison of the long-term forecasting models for OLED.**

	Single point time interval ( $t_{lag} = 14$ )	3 points time interval ( $t_{lag} = 13,14,15$ )
RMSE	8.6	7.32
MAE	6.91	5.84
MAPE	45.34%	36.93%

at  $t$  and ARIMA, particularly the relationship between published patent applications and granted patents is constructed. Table 7 shows the comparison of the long-term forecasting models. In this case, using 3 points time interval to proceed long-term forecasting is better than using a single time point.

From this research work, two modeling methods based on G/P ratios and a long-term modeling method based on average P2G lags were developed to predict future patent granted numbers from prior patent published application numbers. Two case studies on MRAM and OLED were conducted to establish its technology forecasting models in order to verify their forecasting capabilities. Comparing to traditional time series ARIMA method, the forecasting methods based on G/P ratios have similar forecasting capability and the long-term forecasting method based on average P2G lags was proven to be superior in predicting capability. Patent growth generally follows a similar trend that can resemble s-shaped growth. In early stages of a technology the number of patents issued is very limited. A fast-growing period then follows when the number of patents filed and issued increases and then a plateau is reached [28]. Because the patent approval process is costly and can take several years, filing a patent generally means there is optimism in economic or technical contribution granted. Therefore, by analyzing the trend of patent granted numbers predicted by the published patent applications, we can foresee the changes of technology trend, especially on the newly developed technologies. Furthermore, we can model and predict the number of citations of certain patent by using the relationship that we have built and then infer the technology mainstream by identifying the essential patents.

## NOMENCLATURE

- $t_1$  : the starting published date of a patent application  
 $t_2$  : the starting granted date of a published patent ap-

plication

$G^t$ : number of patents granted at time  $t$

$G_n^t$ : number of patents granted at time  $t$  without previously published applications

$G_p^t$ : number of patents granted at time  $t$  with previously published applications

$\tilde{G}_p^{t+1}$ : predicted number of patents granted at time  $t+1$  with previously published applications

$G_{p(\Delta t)}^t$ : number of patents granted at time  $t$  with previous applications published at time  $t-\Delta t$

$P^t$ : number of patent applications published at time  $t$

$P_{g(\Delta t)}^t$ : number of patent applications published at time  $t$  which are later granted at time  $t+\Delta t$

$P_{n(\Delta t)}^t$ : number of patent applications published at time  $t$  which have not yet been granted at time  $t+\Delta t$

$X^t$ : ( $= G_p^t / G^t$ ) ratio of the number of patents granted at time  $t$  with previously published applications to the total number of patents granted at time  $t$

$z_{p(\Delta t)}^t$ : ( $= G_{p(\Delta t)}^t / P^{t-\Delta t}$ ) ratio of the number of patents granted at time  $t$  with previous applications published at time  $t-\Delta t$  to the total number of patent applications published at time  $t-\Delta t$

$D_t$ : ( $= \{ z_{p(\Delta t)}^t \mid 0 \leq \Delta t \leq t-t_1 \}$ ) grant/publish ratio array for patents granted at time  $t$

$\tilde{D}_{t+1}$ : predicted grant/publish ratio array for patents granted at time  $t+1$

$\Delta e_{t+1}$ : ( $\tilde{G}_p^{t+1} - G_p^{t+1}$ ) predicting error

## ACKNOWLEDGMENTS

The authors acknowledge the supports from the National Science Council, Taiwan, with grant no. NSC-98-3114-P-002-005.

## REFERENCES

- Adams, K., Kim, D., Joutz, F. L., Trost, R. P., and Mastrogianis, G., "Modeling and forecasting U.S. patent application filings," *Journal of Policy Modeling*, Vol. 19, pp. 491-535 (1997).
- Basberg, B. L., "Patents and the measurement of technological change: a survey of the literature," *Research Policy*, Vol. 16, pp. 131-141 (1987).
- Campbell, R. S., "Patent trends as a technological forecasting tool," *World Patent Information*, Vol. 5, pp. 137-143 (1983).
- Dehon, C. and Van Pottelsberghe, B., "Implementing a forecasting strategy for PCT applications at WIPO," *Proceedings of the WIPO—OECD Workshop on Statistics in the Patent Field*, Geneva, Switzerland (2003).
- Ernst, H., "The use of patent for technological forecasting: the diffusion of CNC-technology in the machine tool industry," *Small Business Economics*, Vol. 9, pp. 361-381 (1997).
- Griliches, Z., "Patent statistics as economic indicators: a survey," *Journal of Economic Literature*, Vol. 18, pp. 1661-1707 (1990).
- Harhoff, D., *Improvements of Methods for Forecasting Patent Filings*, Research Program announced by the European Patent Office, Background Information Document (2001).
- Heyman, P. D., "Using a patent portfolio to defend against a patent infringement suit," *Intellectual Property & Technology Law Journal*, Vol. 17, pp. 9-15 (2005).
- Hingley, P. and Nicolas, M., "Methods for forecasting numbers of patent applications at the European Patent Office," *World Patent Information*, Vol. 26, pp. 191-204 (2004).
- Joutz, F. L., "Forecasting USPTO patent application filings," *Proceedings of the WIPO—OECD Workshop on Statistics in the Patent Field*, Geneva, Switzerland (2003).
- Joutz, F. L., "Forecasting patent applications at the European Patent Office: a bottom-up versus top-down approach," *Proceedings of the WIPO—OECD Workshop on Statistics in the Patent Field*, Geneva, Switzerland (2004).
- Kotabe, M., "A comparative study of U.S. and Japanese patent systems," *Journal of International Business Studies*, Vol. 23, pp. 147-168 (1992).
- Levary, R. R. and Han, D., "Choosing a technological forecasting method," *Industrial Management*, Vol. 37, pp. 14-18 (1995).
- Liu, S. and Shyu, J., "Strategic planning for technology development with patent analysis," *International Journal of Technology Management*, Vol. 13, pp. 661-680 (1997).
- Martino, J. P., "Technology forecasting: an overview," *Management Science*, Vol. 26, No. 1, pp. 28-33 (1980).
- Martino, J. P., "Thirty years of change and stability," *Technological Forecasting and Social Change*, Vol. 62, pp. 13-18 (1999).
- Martino, J. P., "A review of selected recent advances in technological forecasting," *Technological Forecasting and Social Change*, Vol. 70, pp. 719-733 (2003).
- Mogee, M. E., "Using patent data for technology analysis and planning," *Research Technology Management*, Vol. 34, pp. 43-49 (1991).
- Palmer, P. J., Williams, D. J., and Hughes, C., "Observations and models of technology trends within the electronics industry," *Engineering Science and Education Journal*, Vol. 8, pp. 233-240 (1999).
- Pavitt, K., "Uses and abuses of patent statistics," *Handbook of Quantitative Studies of Science and Technology*, Elsevier Science Publishers, Amsterdam, Holland (1988).
- Porter, A. L., "Tech forecasting an empirical perspective," *Technological Forecasting and Social Change*, Vol. 62, pp. 19-28 (1999).
- Porter, A. L., Roper, A. T., Mason, T. W., Rossini, F. A., and Banks, J., *Forecasting and Management of Technology*, Wiley, Inc., New York, USA (1991).
- Porter, A. L. and Rossini, F. A., *Encyclopedia of Systems and Control, in Technology Forecasting*, Singh, M. G. (Ed.), Pergamon, Oxford, UK (1987).
- Ragusa, P. A., "Eighteen months to publication: should the United States join Europe and Japan by promptly publishing patent applications?," *The George Washington Journal of International Law and Economics*, Vol. 26, pp. 143-180 (1992).
- Scherer, F., "Firm size, market structure, opportunity and the output of patented inventions," *American Economic Review*, Vol. 55, pp. 1097-1125 (1965).
- Schmookler, J., *Invention and Economic Growth*, Harvard U. P., Cambridge, UK (1966).
- Silverman, A. B., "Can publishing your U.S. patent application reveal confidential information," *JOM*, Vol. 53, pp. 64 (2001).
- USPTO, *Performance and Accountability Report for Fiscal Year 2005*, United States Patent and Trademark Office, US (2005).