(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

International Journal of Forecasting 28 (2012) 587–606

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

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

A population dependent diffusion model with a stochastic extension

C. Michalakelis *, T. Sphicopoulos

Department of Informatics and Telecommunications, University of Athens, Greece

ARTICLE INFO

Keywords:
Innovation diffusion
High technology markets
Technology estimation and forecasting
"Population" diffusion model (PDM)
Stochastic diffusion models (SDM)

ABSTRACT

Diffusion modeling is rather broad in nature, and is important in the areas of estimation and forecasting. Conventional models do not incorporate parameters that explicitly take into account the size of the population, or, equivalently, the size of the potential market. As a consequence, the models often fail to provide accurate forecasts, especially when the diffusion process is in the take-off stage. Furthermore, since diffusion is not a strictly deterministic process, forecasts should provide a measure of the underlying uncertainty of the process by incorporating a stochastic process into the formulation of the models.

The aim of the present work is to fill this gap by proposing an aggregate diffusion model, the "population" diffusion model (PDM), which incorporates the potentially varying market size as a function of the corresponding population. This model realization provides more accurate estimations and future forecasts of the diffusion process, especially when compared to the conventional aggregate diffusion models.

1. Introduction

In the context of the contemporary competitive global market, there are tremendous pressures on companies and organizations, due to exponential growth and the introduction of new technologies. Sustainable and disruptive technologies play a critical role in shaping the success of companies.

The introduction of innovative products into a market is frequently connected to heavy investments, requiring critical business planning in order to meet market demand and competition. Industrial plans rolled out in an attempt to attract and retain customers must be forecasted precisely, in order to obtain the expected level of adoptions and market shares. A failure to produce accurate forecasts often has dramatic results for the supply, whether oversupply and unnecessary over-investment, or under-utilization of a firm's capacities.

The most representative case is probably the telecommunications sector. It is one of the most significant contemporary investments, which refers to new technologies and services that are subject to competition. Privatization and deregulation, along with the effects of increasing competition and the introduction of new services, will tend to introduce new research problems. Diffusion forecasting is among the most important of such problems, and aims to face the high level of uncertainty and the consequent need for risk management.

Moreover, estimating and forecasting the diffusion patterns of the innovations is considered important for all kinds of high technology markets. Characteristic examples in the literature include the works of Teng, Grover, and Guttler (2002), who suggest general diffusion patterns for information technology innovations; and Linton (2002), who presents a thorough overview of the literature, before proposing a method for forecasting disruptive and discontinuous innovations. In addition, Fildes and Kumar (2002) review the telecommunications demand forecasting literature.

Although forecasting models for established products and services are well developed, new opportunities have emerged due to the nature of the high technology product market. Therefore, further methodological work should be

* Correspondence to: Department of Informatics and Telecommunications, University of Athens, Panepistimioupolis, Illisia, Greece. Tel.: +30 210 7275 319.

E-mail address: michalak@di.uoa.gr (C. Michalakelis).

0169-2070/$ – see front matter © 2012 Published by Elsevier B.V. on behalf of International Institute of Forecasters.
doi:10.1016/j.ijforecast.2012.03.002

carried out to identify the gaps that have opened up as a result of changes in the market's scope and structure.

The remainder of the paper is organised as follows. Section 2 provides a short presentation of the background to diffusion theory and models, together with the research objectives and contribution of the present work. Section 3 describes the deterministic process for the "population" diffusion model (PDM). The model's evaluation results are presented in Section 4. The stochastic analogue of the "population" diffusion model (PDM) is presented in Section 5, and its evaluation results are illustrated in Section 6. The conclusions are presented in Section 7, together with their limitations and directions for future research.

2. Background and objectives

2.1. Models for innovation diffusion

One of the central themes of the innovation field is the mathematical modeling of innovation diffusion, which pertains to different types of innovations under different assumptions. The main findings can be summarized as a bell-shaped curve depicting the frequency of adoption against time and an S-shaped curve representing the cumulative number of adopters. During the initial stages of the innovation life cycle (the introduction stage), the rate of adoption is relatively low. This is followed by the take-off stage, characterised by a high rate of adoption. Finally, the peak of the bell curve is reached, corresponding to the inflection point of the cumulative adoption. After this point, the adoption rate decreases until the market saturation level is met asymptotically and the maximum number of adopters has been reached. This situation corresponds to the end of the life cycle of the innovation, after which, in the case of high technology products, it is usually replaced by its descendant generation.

Apart from the pioneering work of Gompertz (1825), the works of Bass (1969) and Rai (1999) represent the early contemporary efforts to capture the diffusion dynamics. These two works, together with the logistic family of models (Bewley & Fiebig, 1988), such as the linear logistic or Fisher–Pry model (Fisher & Pry, 1971), are the most widely used diffusion models employed for estimating and forecasting market demand.

All of the above models provide "S-shaped curve" estimations to describe the cumulative diffusion, which is used to estimate and forecast the diffusion of innovations at the aggregate level. This approach describes the total market response, in contrast to econometric models, which focus on the factors affecting the adoption of the studied innovation by individuals.

Aggregate models are generally able to provide reliable estimations of diffusion processes with respect to the adoption of innovations into a market of reference. One of their main fields of application is the sector of high technology, and especially telecommunications. An informative review of the forecasting of demand in telecommunications is provided by Fildes and Kumar (2002). Moreover, important research results on the development and evaluation of diffusion models are presented by Bewley and Fiebig (1988), Geroski (2000), Jain and Rai (1988), Mahajan and Muller (1979), Mahajan, Muller, and Bass (1990), Michalakelis, Varoutas, and Sphicopoulos (2008) and Skiadas (1987). Extensions of these models include cross-national influence (Kumar, Ganesh, & Echambadi, 1998; Kumar & Krishnan, 2002; Michalakelis, Dede, Varoutas, & Sphicopoulos, 2008) and marketing variables (Mahajan & Peterson, 1979; Ruiz-Conde, Leeflang, & Wieringa, 2006), in an attempt to describe the process of adoption in more detail. Most of the resulting models are derived by incorporating functional adjustments into the original formulation of a diffusion model's equation.

2.2. Research objectives and contribution

None of the approaches described above, despite employing a number of parameters and marketing variables, explicitly take into consideration the influence of the population size on the diffusion process, and consequently incorporate it into the appropriate mathematical formulations. The market potential, or saturation level, used in the formulation of the models does not coincide with the actual size of the market population, nor does it depend on it. Moreover, no relationship between these two quantities is derived. In fact, they may lead to substantial divergences in forecasting, especially at certain stages of the diffusion process, such as the take-off stage. This may turn out to be a major deficiency in the context of diffusion analysis, especially in the case of rapid take-offs. The reason for this is that the population size imposes a constraint on the further acceleration of the innovation diffusion. This is common in the case of a technological product, where the probability of adopting another unit of the product decreases after the first adoption, although it is not totally eliminated. Mobile phones, broadband connections and personal computers are some representative examples. Conventional diffusion models do not usually provide accurate forecasts when the diffusion process accelerates fast, since their forecasts tend to diverge from the actual values recorded later, sometimes substantially. For some important contributions to the task of appraising the influence of the population on the diffusion process, see Gruber (2001) and Gruber and Verboven (2001).

Without limiting the value of the above contributions, the first objective of the present work is to study the influence of the target market's size on the diffusion process, by developing an aggregate diffusion model that includes the population size of the market as an explicit parameter which influences both the diffusion rate and the market potential. This is an important contribution, especially at the take-off stage of the diffusion life cycle, where the inflection point has not been reached. At this particular stage, the process is characterised by a high level of skewness and flexibility. In such cases, forecasting can contain a high level of uncertainty and have a major effect on strategic planning. Since the life cycle of a typical product is expected to be rather short, due to substitution by its descendant generation, forecasting is usually based on a limited amount of historical data.

A second objective is the development of a diffusion model that is capable of providing a flexible inflection point which is able to capture the flexibility and skewness of the process and which is affected by the market size and expressed as a function of it. Since the inflection point of the "traditional" diffusion models occurs at a certain level of capacity K (e.g., 50% of K for the logistic growth model), there is a need for a new model which can describe the diffusion process of high technology products, which are usually characterized by asymmetric shapes and a high level of skewness. This is described in more detail in the next section, where the development of the proposed model is presented.

The third objective of this research is to provide a measure which can quantify the level of forecast uncertainty. It is worth mentioning that most of the research in the literature focuses on describing diffusion as a deterministic process in time. However, diffusion processes should often be modeled as dynamic phenomena, which can be described by stochastic differential equations. One work that raised this need is that of Eliashberg and Chatterjee (1986), who present some arguments on the necessity of using stochastic models, together with ways of introducing randomness into a diffusion process study. In the end, an estimation of the level of uncertainty in forecasts can provide valuable input with respect to a firm's available actions, predictions for the future and even expectations for competition.

Another important objective here is the application of the proposed diffusion model to a number of historical data series, in order to evaluate its performance and its ability to provide accurate forecasts. This is especially important when the peak point is not included in the observations available, which usually limits the forecasting ability of the conventional models, due to the underlying uncertainty of the process.

Given the above considerations, the present work is devoted to the construction of a new diffusion model, the "*population-diffusion model*", or the PDM. The model turns out to be able to capture the influence of the population size on the procedure of estimating and forecasting penetration in the case of high technology products. Two formulations of the model are constructed, presented and evaluated: the deterministic model and its stochastic analogue. Evaluation was performed using historical data on mobile phone subscribers from 22 countries over the wider European area. The results demonstrate that not only can the PDM provide accurate diffusion estimations over a dataset that includes inflection points, but it can also forecast future values quite accurately. In order to compare the results of the PDM with the performances of conventional diffusion models, the most representative cases of the latter (the logistic, Gompertz and Bass models) are also evaluated. Moreover, the evaluation of the stochastic realization provided a range of forecasted values that are also quite accurate, since they indicate lower and upper bounds within which future recorded values are expected to fall. The proposed diffusion model takes into account the number of connections (services) provided to the customer, not the physical number of devices. The probability of purchasing the first connection is higher than the probability of purchasing a second connection, and so on.

The data used to evaluate the PDM were extracted from the ITU (International Telecommunication Union) database, describing the diffusion patterns of the countries considered.

The accomplishment of the objectives of this work will contribute significantly to both research and practice. Research will benefit from the provision of new directions for the development of a diffusion forecasting framework by incorporating a number of influential parameters. This will help improve our understanding of the diffusion shapes in the high technology market, especially when combined with our study of the underlying uncertainty. For practitioners, our research findings will be very useful for strategic planning and decision making, as they can be applied in an increasingly competitive environment, since more accurate a priori estimates of the diffusion pattern can be derived, enhanced by the estimation of the level of uncertainty provided by the stochastic analysis.

However, it should be clarified that the proposed model and the corresponding analysis do not attempt to provide a method for the estimation of time-variant saturation levels, based on the population size, but only show the effect of the population size on the diffusion process.

3. Deterministic "population" diffusion model (PDM)

3.1. Development of the model

Aggregate diffusion models, which describe cumulative penetration, are derived by a differential equation of the following general form (Seber & Wild, 2003):

$$\frac{dN(t)}{dt} = d\left(N(t), \bar{p}\right) \cdot \left[f\left(K, N(t)\right)\right]. \quad (1)$$

In the above equation, $N(t)$ represents the total penetration at time t, while K is the saturation level, or the maximum expected cumulative penetration of the innovation. $d(N(t), \bar{p})$ is a function which represents a factor of proportionality. The quantity $f(K, N(t))$ represents the function of the remaining market potential at time t, depending on the saturation level of K and the number of adopters $N(t)$ at this time t. Finally, $\bar{p}$ is the vector of the model parameters, which are considered as constants over the period of the study.

The most widely used aggregate diffusion models are the linear logistic and Bass models, which create symmetric S-curves. In Fig. 1, the diffusion shapes of some of the most common models are illustrated for the same saturation level of K, together with their inflection points, where the diffusion rate becomes equal to zero before starting to decrease. As was observed, the linear logistic and Bass models are described by a symmetric diffusion shape, with their inflection points at $K/2$. However, in the case of high technology products, the corresponding diffusion shapes are usually right skewed rather than symmetric, since they follow a high initial rate of diffusion, slowing down after that. Therefore, symmetric models often fail to describe such cases adequately.

The non-symmetric Gompertz model, which exposes an inflection point at the 37% of K, is sometimes considered as more appropriate for describing a diffusion process. Bemmaor and Lee (2002) made an important contribution regarding the changes in parameter estimates of the Bass

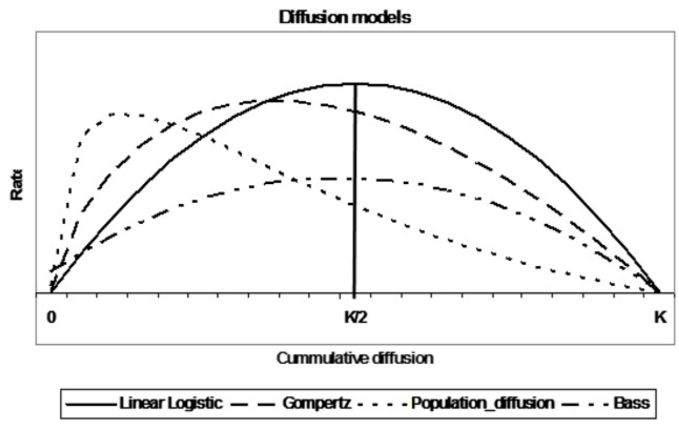


Fig. 1. Diffusion rates of the most popular aggregate models.

model, the coefficients of innovation, imitation and the market penetration rate as a result of the underlying heterogeneity of the population. They found significantly opposite patterns in these parameters, depending on the skew of the diffusion curve detected by a generalized model.

According to the general formulation of diffusion models described by Eq. (1), the linear logistic model is described by:

$$\frac{dN(t)}{dt} = rN(t)\,(K - N(t))\,, \tag{2}$$

and the Gompertz model by:

$$\frac{dN(t)}{dt} = rN(t)\ln\left(\frac{K}{N(t)}\right). \tag{3}$$

In the above equation, r is the coefficient of proportionality, and K is the market potential (the maximum expected level of diffusion). However, in the case of rapid diffusion rates, the asymmetric Gompertz model often fails to provide accurate forecasts, as it does not take the effect of the population size into account; this is particularly a problem when the available historical data do not include the inflection point. Thus, introducing a Gompertz-like model would provide the ability to incorporate the effect of the population size into an appropriate formulation. These considerations constitute the underlying explanation for the construction of the proposed model.

Thus, in accordance with Eq. (1), the main assumptions of the proposed model are:

$$d(N(t), \bar{p}) = rN(t)\ln\left(a + b\frac{P}{N(t)}\right), \tag{4}$$

and the $f\,(K, N(t))$ function is given by the following equation:

$$f\,(K, N(t)) = \ln\left(\frac{K}{N(t)}\right). \tag{5}$$

Accordingly, the PDM is described by the following differential equation:

$$\frac{dN(t)}{dt} = rN(t)\ln\left(a + b\frac{P}{N(t)}\right)\ln\left(\frac{K}{N(t)}\right), \tag{6}$$

where $N(t)$ is the number of adopters (i.e., the number of mobile connections in a mobile company, not the number of devices) at time t, K is the saturation level, P is the population size, and r is a coefficient of proportionality. It should be noted that the population does not necessarily refer to the number of individuals, but can describe the number of households, or any otherwise defined unit of adopters. It is obvious that if the nonnegative quantities a and b take values of e and 0, respectively, Eq. (6) reduces to the well known Gompertz model, thus eliminating the effect of the population on the diffusion process. In Eq. (6), the saturation level K may be equal to, less than, or greater than the population size P.

The proposed model described by Eq. (6) is the most appropriate, since, as Fig. 1 shows, it is capable of providing a flexible inflection point which is affected by the market size. More specifically, by observing Eq. (6), it becomes clear that the $\ln\left(a + b\frac{P}{N(t)}\right)$ part of the equation is capable of describing a fast acceleration of diffusion, as long as $N(t)$ is smaller than the population size, P. However, when $N(t)$ becomes equal to or greater than P, the diffusion rate does not stop, but only slows down, thus providing a better estimate of the consequent value of the adoption rate. This reflects the possibility of one unit of the population adopting a second or third unit of the product. Moreover, the choice of the logarithmic function retains the asymmetric property of the diffusion shape, thus allowing the model to easily adapt and describe the process, as illustrated in Fig. 2. For certain values of the parameters a and b (e and 0, respectively), the model coincides with the Gompertz model, as indicated by the heavy line.

Consequently, based on its construction, the proposed model is characterized by a high level of flexibility, which is able to accommodate the peculiarities of any described diffusion process. This is extremely important for the stage before the inflection point, where the conventional models are not flexible enough to describe the non-symmetric nature of the process. In addition, and due to the nature of its construction, the PDM incorporates skewness into the diffusion process, which also contributes to the flexibility of the model.

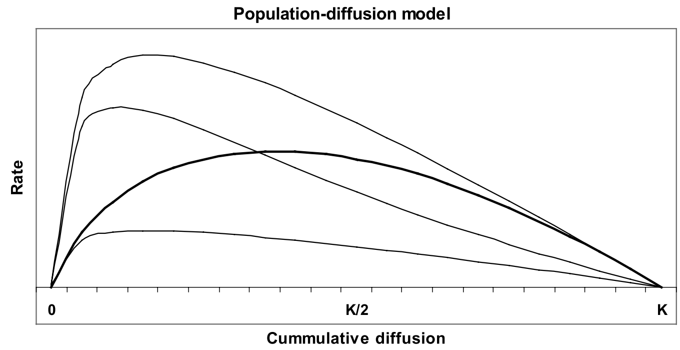


Fig. 2. Diffusion shape of the PDM for different values of the parameters and the same K.

To obtain the solution of the autonomous differential equation (6), the following transformation of variables is used:

$$u = \ln\left(\frac{N}{K}\right), \tag{7}$$

and consequently,

$$N = Ke^{u} \Rightarrow \frac{dN}{dt} = Ke^{u}\frac{du}{dt}; \tag{8}$$

then, by using the equalities of Eq. (8), the initial Eq. (6) is transformed to:

$$\frac{dN}{dt} = Ke^{u}\frac{du}{dt} = -ruKe^{u}\ln\left(a + \frac{bP}{Ke^{u}}\right). \tag{9}$$

The simplification of the terms in Eq. (9) leads to:

$$\frac{du}{dt} = -ru\ln\left(a + \frac{bP}{Ke^{u}}\right). \tag{10}$$

In addition, the expansion of the $\ln\left(a + \frac{bP}{Ke^{u}}\right)$ term into a MacLaurin power series (Wrede & Spiegel, 2002) yields the formulation of the differential equation, to be solved as:

$$\frac{du}{dt} = -ru\left[\ln\left(a + \frac{bP}{K}\right) - \frac{bP}{aK + bP}u\right]. \tag{11}$$

Since $\ln\left(a + \frac{bP}{K}\right)$ and $\frac{bP}{aK+bP}$ are constant quantities, for computational convenience they can be replaced by the constants x and y, respectively.

Thus, by setting

$$x = \ln\left(a + \frac{bP}{K}\right), \qquad y = \frac{bP}{aK + bP}, \tag{12}$$

the initial equation, Eq. (11), after incorporating the substitution in Eq. (12), becomes:

$$\frac{du}{dt} = -ru(x - yu). \tag{13}$$

Eq. (13) is an autonomous differential equation with the following solution (Boyce & DiPrima, 2005):

$$u = \frac{xe^{-rxt}C}{1 + ye^{-rxt}C}. \tag{14}$$

Eq. (14) provides the general solution of the differential equation, and C is an arbitrary constant, known as the constant of integration. However, in the case of an initial value problem, C can be determined by substituting the initial values into the general solution and solving the equation with respect to C.

The substitution of the solution into the initial variable transformation, described by Eq. (8), yields:

$$N = Ke^{\frac{xe^{-rxt}C}{1+ye^{-rxt}C}}, \tag{15}$$

since it is an initial value problem and the integration constant C can be calculated, as mentioned above, by incorporating the initial condition: $N(0) = N_0$, into Eq. (15). Thus,

$$N_0 = Ke^{\frac{xC}{1+yC}} \quad \text{or} \quad C = \frac{\ln\left(\frac{N_0}{K}\right)}{x - y\ln\left(\frac{N_0}{K}\right)}. \tag{16}$$

The substitution of C into the initial equation, Eq. (15), after performing the necessary calculations, yields:

$$N = Ke^{\frac{x\ln\left(\frac{N_0}{K}\right)e^{-rxt}}{x + y\ln\left(\frac{N_0}{K}\right)\left(e^{-rxt}-1\right)}}. \tag{17}$$

Finally, reversing the substitution described by Eq. (12), or, equivalently, substituting x and y for $\ln(a + \frac{bP}{K})$ and $\frac{bP}{aK+bP}$, respectively, leads to the general formulation of the proposed model, described by:

$$N = Ke^{\frac{\ln\left(a+\frac{bP}{K}\right)\ln\left(\frac{N_0}{K}\right)e^{-r\ln\left(a+\frac{bP}{K}\right)t}}{\ln\left(a+\frac{bP}{K}\right)+\frac{bP}{aK+bP}\ln\left(\frac{N_0}{K}\right)\left(e^{-r\ln\left(a+\frac{bP}{K}\right)t}-1\right)}}. \tag{18}$$

The solution presented in Eq. (18) can be represented graphically by an S-shaped curve, and reduces again to the formulation of the Gompertz model if a and b become equal to e and 0, respectively.

4. Evaluation of the model

The evaluation of the PDM was performed using historical data describing the diffusion of both 2G and 3G mobile telephony over the population, for 22 countries in the wider European area. The evaluation data were collected from the International Telecommunication Union

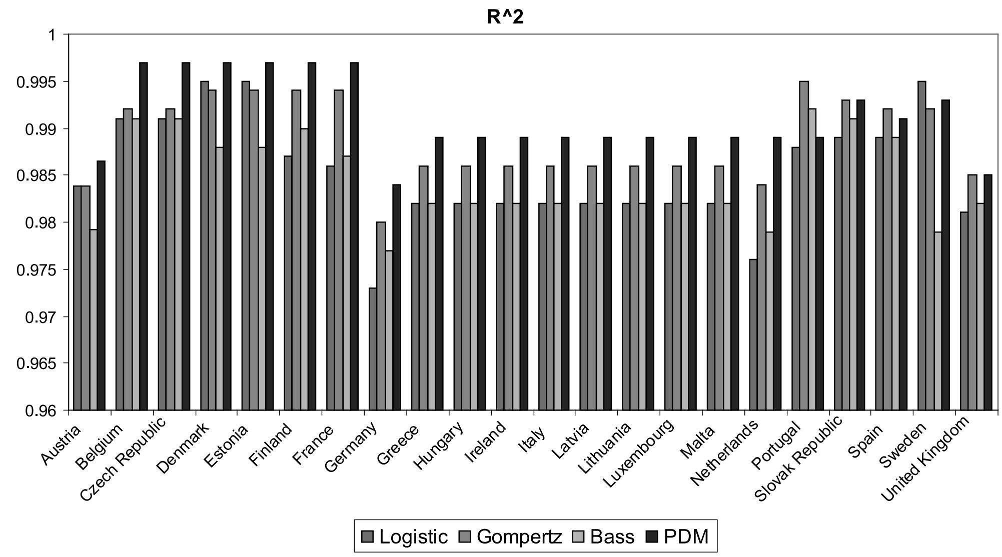


Fig. 3. R^2 values for the evaluated cases.

(ITU, http://www.itu.int), and correspond to the period from 1995 to 2007. In addition to this, population data were extracted from the Eurostat and ITU databases.

The procedure includes evaluating the model's performance in terms of its ability to estimate the corresponding diffusion processes. Moreover, three of the most widely used models are also included in the evaluation process, in order to provide comparative results regarding the PDM performance. The estimation of the models' parameters in each case was performed by the nonlinear least squares (NLS) estimation method.

The other models involved in the evaluation process are the Gompertz model (Rai, 1999), the linear logistic model (Bewley & Fiebig, 1988; Fisher & Pry, 1971) and the Bass model (Bass, 1969), and their specifications are given by the following formulations, which describe the number of adopters $N(t)$ at time t:

$$N(t) = \frac{K}{1 + e^{-a-bt}} \tag{19}$$

$$N(t) = Ke^{-e^{-a-bt}} \tag{20}$$

$$N(t) = K\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}. \tag{21}$$

Eqs. (19) and (20) are the solutions of the models described by Eqs. (2) and (3), respectively. In all of the models above, the parameter K represents the saturation level of the market. In Eqs. (19) and (20), the parameters a and b correspond to the rate and shape of the diffusion, whereas the parameters p and q in the Bass model (Eq. (21)) correspond to the coefficients of innovation and imitation, respectively. p describes the probability of potential adopters proceeding to an early adoption, and q describes the probability of adopting the product or service as a result of influential interactions with others who have already adopted.

A number of characteristic results, corresponding to cases with different diffusion rates, are illustrated in

Table 1
Evaluation of the participating diffusion models: statistical measures of accuracy.

		Min	Max	Mean
Linear logistic	R^2	0.973	0.995	0.981
	MSE	1E−4	0.004	0.002
	MAPE	0.035	9.28	0.205
Gompertz	R^2	0.981	0.994	0.99
	MSE	0.001	0.004	0.002
	MAPE	0.03	8.93	0.22
Bass	R^2	0.977	0.991	0.987
	MSE	1E−4	0.003	0.002
	MAPE	0.04	6.46	0.32
Population-diffusion	R^2	0.984	0.997	0.99
	MSE	1E−4	0.003	0.002
	MAPE	0.019	5.423	0.186

Figs. 5–8. The detailed results of the evaluation, together with the historical data used, are included in Appendix A, where the performance of the PDM is illustrated.

The statistical measures of accuracy calculated for all of the evaluated datasets are presented in Table 1, where the minimums (min), maximums (max) and means (mean) of the calculated values are given. As was observed earlier, the coefficient of determination (R^2) for the proposed model is quite high, showing that it is managing to describe the observed data adequately. In addition, the calculated error measures, namely the Mean Squared Error (MSE) and the Mean Absolute Percentage Error (MAPE), are at quite acceptably low levels. The calculated statistical measures are illustrated in the plots of Figs. 3–5.

In addition to the evaluation of the PDM's ability to estimate diffusion accurately, its forecasting ability was also evaluated. This was achieved by splitting the available datasets into two parts, the "training" and "holdback" (or holdout) datasets. The former were used for training the model and estimating its parameters, whereas the latter were used for comparing the actual recorded values with the forecasts from the models. The holdback samples include historical data from 1995 until one year before the inflection point (the maximum recorded penetration

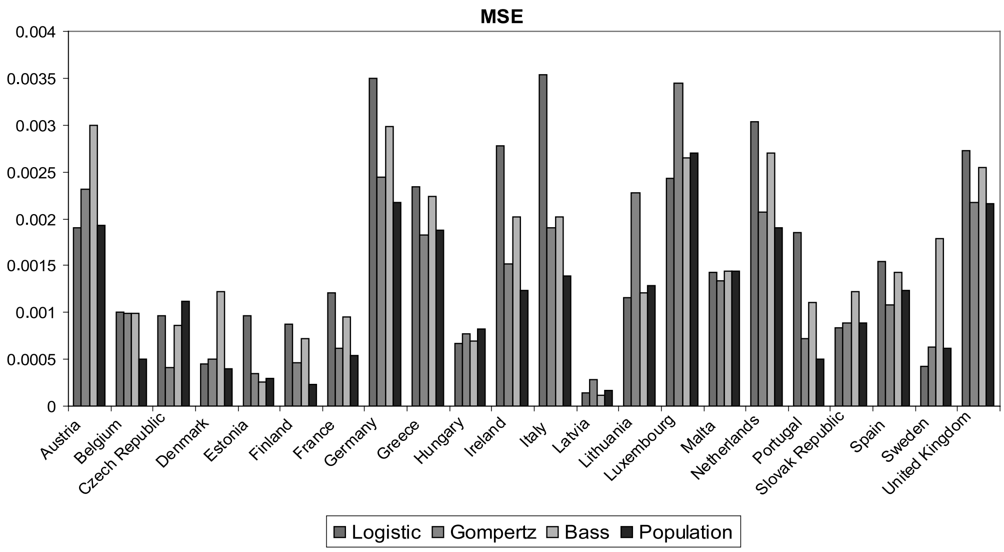


Fig. 4. MSE values for the evaluated cases.

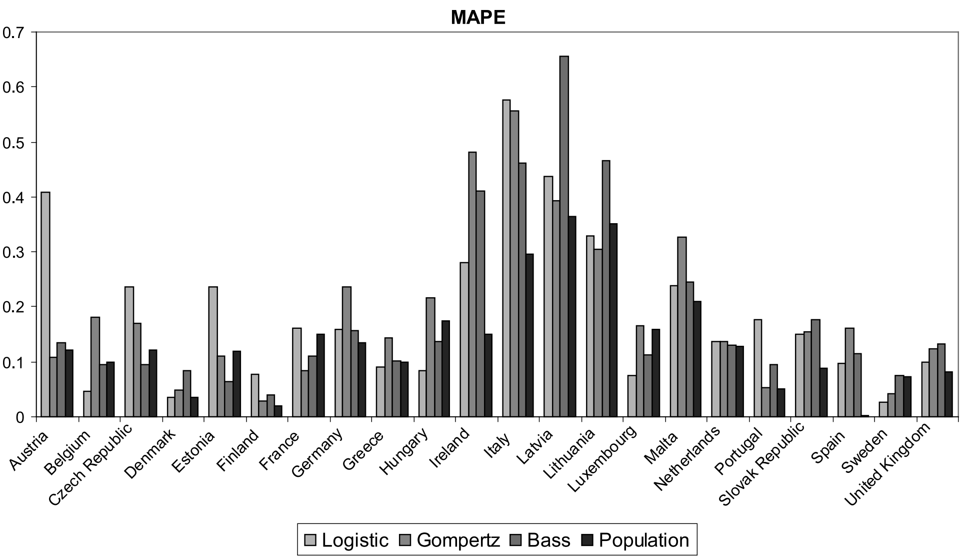


Fig. 5. MAPE values for the evaluated cases.

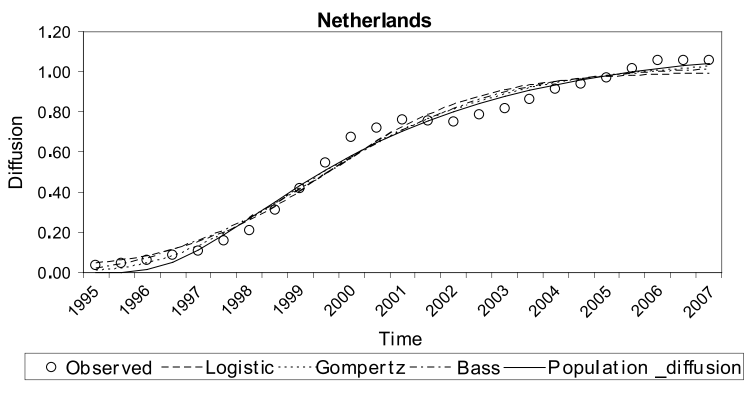


Fig. 6. Diffusion estimation of the population-diffusion model, using actual data for the Netherlands, compared with the performances of other diffusion models.

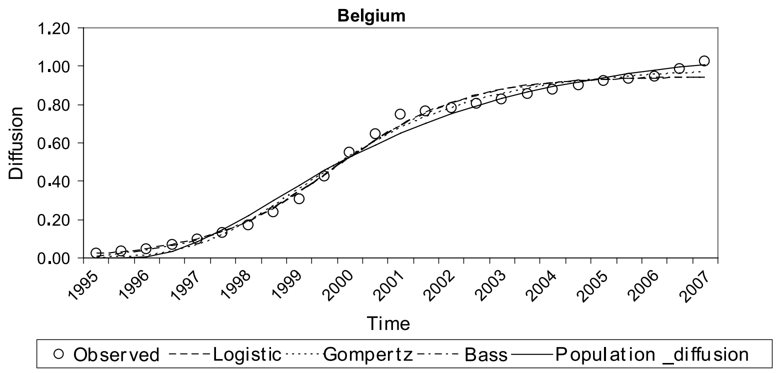


Fig. 7. Diffusion estimation of the population-diffusion model, using actual data for Belgium, compared with the performances of other diffusion models.

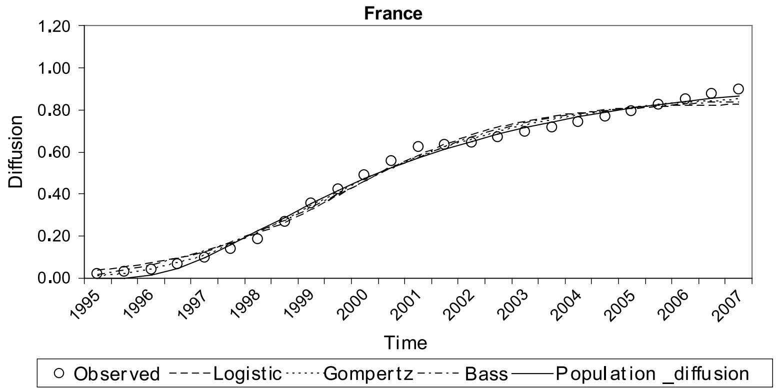


Fig. 8. Diffusion estimation of the population-diffusion model, using actual data for France, compared with the performances of other diffusion models.

value, before the diffusion process slowed down), in each case considered. The model parameters were then estimated again by the NLS estimation method. Some of the evaluation results are illustrated in Figs. 9–12, showing the diffusion (y-axis) over time (x-axis). The y-axis corresponds to the penetration level (percentage of adoptions over population). As observed, the PDM manages to forecast the diffusion process accurately, even before the inflection point has been reached, while the rest of the participating models fail. From Figs. 9 to 12, it is clear that the PDM forecasts are close to the observed data, particularly if the simulation is continued after 2007.

Detailed results of the evaluation of the models' forecasting abilities for all of the datasets are presented in Appendix B.

The main finding of the last evaluated case, with respect to the forecasting ability of the proposed model, is that attempting to forecast using conventional models before the inflection point provides results which diverge observably from the actual future values. This is because the population size of the market that slows down the diffusion rate does not participate as a parameter in any of the other evaluated benchmark models.

In Table 2, the estimated market potentials for all of the evaluated models and the participating countries are presented. By observing the corresponding values, it can be shown that the PDM can forecast the market potential quite accurately, especially when the conventional models provide observably diverging estimates.

As a final step, and in order to demonstrate the importance of the population size P, a sensitivity analysis is performed. More specifically, the changes in the forecasts are estimated for different values of P. The corresponding results are presented in Fig. 13 for the case of Belgium.

In the above diagrams, the model is evaluated for the actual population, PDM (P), for 20% and 50% increases in the population, PDM $(1, 2^*P)$ and PDM $(1, 5^*P)$ respectively, and for a 20% decrease, PDM $(0, 8^*P)$. As observed, there is a substantial change in the estimated forecasts with respect to both the adoption rate and the saturation level. Sensitivity results reveal the necessity for the precise specification of the population parameter, P, in the proposed model.

5. Stochastic "population" diffusion model (PDM)

5.1. The need for stochastic analysis

Despite the fact that diffusion models often succeed in describing a diffusion process, stochastic considerations are usually ignored. Nevertheless, they are of major importance, due to the existence of rapidly changing environmental socioeconomic factors, which in turn affect the diffusion's characteristics by adding randomness to the adoption pattern (Eliashberg & Chatterjee, 1986). Moreover, the deterministic realizations of diffusion models are able to provide discrete estimations of the process, whereas

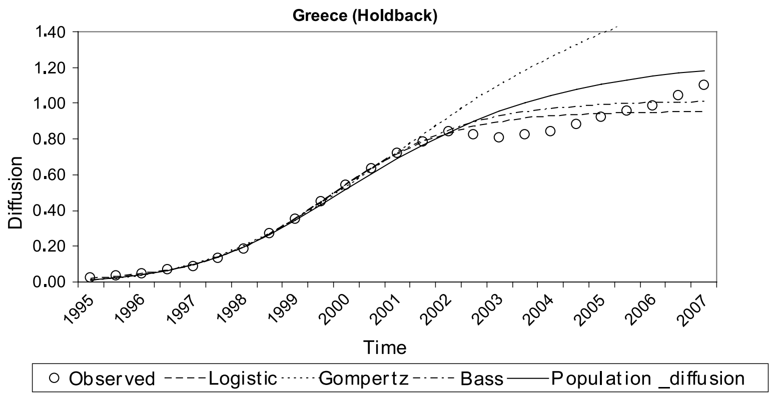


Fig. 9. Forecasting evaluation of the "population-diffusion" model, using actual data for Greece with a holdback sample, compared with the performances of the other participating diffusion models.

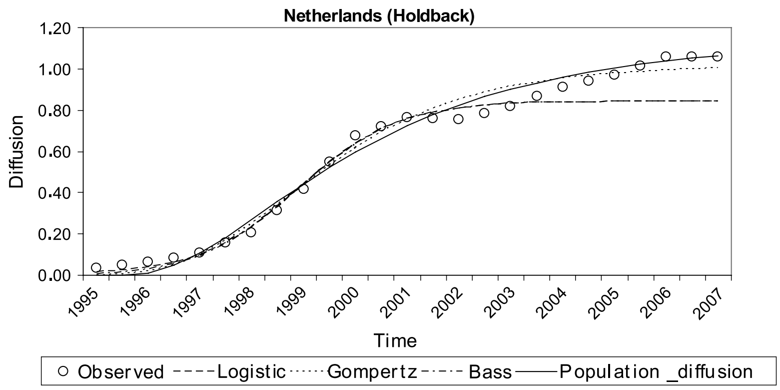


Fig. 10. Forecasting evaluation of the "population-diffusion" model, using actual data for the Netherlands with a holdback sample, compared with the performances of the other participating diffusion models.

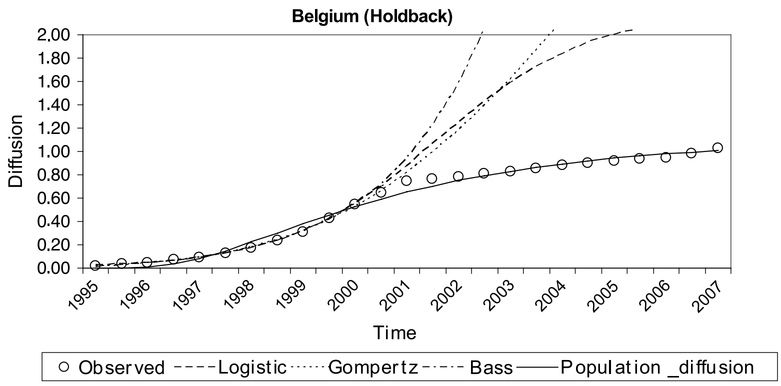


Fig. 11. Forecasting evaluation of the "population-diffusion" model, using actual data for Belgium with a holdback sample, compared with the performances of the other participating diffusion models.

stochastic models are capable of quantifying the uncertainty caused by factors which can be either internal or external to the system, thus providing a set of possible scenarios of the process at each point of time. In addition, no matter how sophisticated a deterministic model is, it cannot include every single factor that could possibly affect the process. Since many of the external parameters are random in nature, they cannot be estimated accurately and used for forecasting purposes. In addition, a stochastic perspective is especially vital, given the long-term forecasts that the diffusion models can potentially provide and the existence of several rapidly changing factors in the environment, as well as in the interior of the system (Giovanis & Skiadas, 1999; Skiadas & Giovanis, 1997).

There are two main ways of introducing randomness into a deterministic diffusion model in order to describe the uncertainties that accompany a diffusion process. The first is based on the assumption that the parameters of an aggregate diffusion model follow a stationary stochastic process (Karmeshu & Pathria, 1980). An interesting study

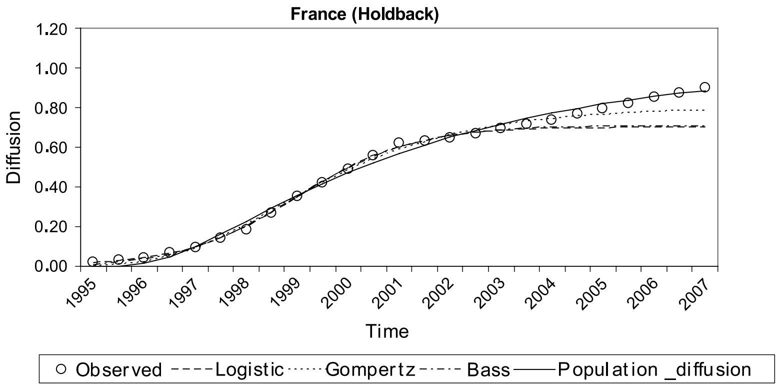


Fig. 12. Forecasting evaluation of the "population-diffusion" model, using actual data for France with a holdback sample, compared with the performances of the other participating diffusion models.

Table 2
Comparison of the estimated market potential for all models.

	Logistic	Gompertz	Bass	Population
Austria	1.4709	28.385	3.011	1.15
Belgium	2.2012	11.266	28124.645	1.00
Czech Republic	2.2220	2.222	1.033	1.06
Denmark	1.4563	4.559	13.464	1.30
Estonia	0.6869	1.098	0.827	1.47
Finland	0.9526	1.082	1.695	1.12
France	0.7003	0.798	0.707	1.03
Germany	0.8332	0.912	0.834	1.13
Greece	0.9513	1.876	1.010	1.24
Hungary	1.6750	132.934	2.220	1.13
Ireland	0.8847	1.169	0.901	1.21
Italy	1.0991	1.576	1.205	1.52
Latvia	0.8469	2.895	1.003	1.81
Lithuania	1481.0500	1.E+21	107755.236	1.83
Luxembourg	1.3478	2.695	1.510	1.33
Malta	6.7024	9.E+24	3.903	1.00
Netherlands	0.8421	1.015	0.842	1.14
Portugal	0.9128	1.184	0.949	1.20
Slovak Republic	0.8454	2.039	0.885	1.79
Spain	0.8829	1.311	0.884	1.18
Sweden	1.3296	2.528	2.635	1.21
United Kingdom	0.9152	1.086	0.910	1.18

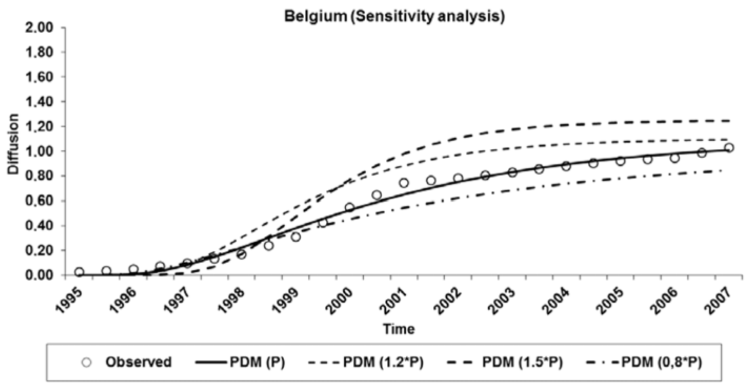


Fig. 13. Sensitivity analysis of the "population-diffusion" model over actual data for Belgium, for different values of the population, P.

was undertaken by Debecker and Modis (1994), to quantify the uncertainties of the parameters determined by logistic S-curve fits. The second approach involves assuming that the future remaining growth of the underlying process is not known with certainty, but can be modeled using an appropriate stochastic process through an Ito's stochastic differential equation (SDE), taking into account the internal and/or external fluctuations. Such a framework for a stochastic substitution model was developed by Meade (1989), whereas stochastic innovation diffusion models involving the introduction of stochasticity into the Bass and logistic diffusion models, respectively, were proposed by Skiadas and Giovanis (1997) and Giovanis and Skiadas (1999). In the context of the present work, a stochastic realization of the PDM is constructed and evaluated based on the second approach described above, providing a valuable way of measuring the uncertainty regarding the forecasted future values.

5.2. Model development

Stochastic differential equations (SDEs) are used quite frequently, and in a wide range of fields of application. As was mentioned in the introductory section, the work presented by Eliashberg and Chatterjee (1986) was one of the earliest studies to propose the need to employ SDEs in the estimation and forecasting of diffusion processes.

This section is devoted to the development of a method of stochastic analysis for aggregate models, and more specifically for the proposed population diffusion model, PDM. The analysis is based on the addition of a noise term to the initially developed model, which is represented by a Wiener process. Wiener processes play a vital role in stochastic calculus and diffusion processes. This analysis is important as a measure of the volatility of the deterministic process, since diffusion is frequently characterized by uncertainty.

Stochastic differential equations are ordinary differential equations which are parameterized by means of Wiener processes (Gardiner, 2004; Oksendal, 2003), in addition to time. A Wiener process, W_t, is a non-differentiable random function of time t, which is obtained by sampling the normal probability density:

$$\frac{1}{\sqrt{2\pi t}}e^{-W_t^2/2t}. \tag{22}$$

A general formulation of a stochastic differential equation is given by the following relationship:

$$dN = f(N,t)dt + g(N,t)dW(t). \tag{23}$$

In Eq. (23), $dW(t)$ is 1-dimensional "white noise", the time derivative of the Wiener process, and $f(N,t)$, $g(N,t)$ are known functions (Evans, 2005). $g(N,t)$ represents the volatility, or the width of the noise of the process.

Recalling the general formulation of an aggregate diffusion model, as described by Eq. (1), the addition of a stochastic term yields:

$$dN(t) = d(N(t),\bar{p}) \cdot \{[f(K) - f(N(t))]dt + g \cdot dW(t)\}, \tag{24}$$

or, equivalently:

$$dN(t) = d(N(t),\bar{p}) \cdot [f(K) - f(N(t))]dt + d(N(t),\bar{p}) \cdot g \cdot dW(t). \tag{25}$$

The application of the formulations described in Eqs. (24) and (25) over the population model, as is described by Eq. (6), provides its equivalent stochastic realization:

$$dN = rN\ln\left(a + b\frac{P}{N}\right)\left[\ln\left(\frac{K}{N}\right)dt + gdW(t)\right], \tag{26}$$

which, after performing the expansion of terms, becomes:

$$dN = rN\ln\left(a + b\frac{P}{N}\right)\ln\left(\frac{K}{N}\right)dt + rN\ln\left(a + b\frac{P}{N}\right)gdW(t). \tag{27}$$

The stochastic differential equation described by Eq. (27) is a nonlinear stochastic differential equation. Therefore, in order to estimate its parameters, a suitable transformation should be applied so that the sde can be approximated by an equivalent linear sde, based on the available discrete time observations.

5.3. Local linearization

5.3.1. Function transformation

A number of approaches are available in the literature for the linearization of a nonlinear diffusion process (Bergstrom, 1990; Dembo & Zeitouni, 1987; Milstein, 1995; Nikolau, 2005; Ozaki, 1985; Singer, 1993). For the needs of the present work, the local linearization method proposed by Shoji and Ozaki (1998) is adopted, as it provides a better performance and more accurate results. According to this method, the original stochastic differential equation is locally approximated by a linear stochastic differential equation, which is analytically tractable, since it can be solved easily and the solution is expressed as a sample path of the stochastic process. Thus, the discretized process can be obtained by the discretization of the sample path.

The basic idea of the local linearization method is based on approximating a stochastic differential equation of the form

$$dN = f(N,t)dt + \sigma dW(t), \tag{28}$$

using a linear differential equation

$$dN = L_s N dt + \sigma dW(t), \tag{29}$$

with L_s being a real valued function and σ the volatility of the process.

The local linearization methodology is applied over the stochastic formulation of the population model expressed by Eq. (27). Again, as in the deterministic formulation of the model and for the sake of computational simplification, the constants x and y are substituted for the constant quantities $\ln(a + \frac{bP}{K})$ and $\frac{bP}{aK+bP}$, respectively. Therefore, the corresponding stochastic differential equation that describes the stochastic diffusion process becomes:

$$\frac{dN}{dt} = rN(x - yN)\ln\left(\frac{K}{N}\right)dt + rN(x - yN)gdW(t). \tag{30}$$

In order to express Eq. (30) in a more tractable form, similar to that of Eq. (28), which has a constant coefficient for the diffusion term (volatility), a transformation $n_t = \varphi(N_t)$ can be applied to Eq. (30). $\varphi(N_t)$ must satisfy the ordinary differential equation:

$$grN(x - yN)\frac{d\varphi}{dN} = \sigma, \quad \sigma \text{ constant}. \tag{31}$$

The solution of Eq. (31) is:

$$\varphi = \frac{\sigma}{grx} \ln \left(\frac{N}{x - yN} \right), \quad (32)$$

and N can be obtained by

$$N = \frac{x}{y + e^{-\frac{ngrx}{\sigma}}}. \quad (33)$$

Thus, Eq. (32) is the transformation that should be applied over Eq. (30) in order to derive a constant volatility stochastic differential equation. The application of Ito's formula (Evans, 2005; Oksendal, 2003; Shoji & Ozaki, 1998) gives:

$$dn = \left[rN(x - y) \ln \left(\frac{K}{N} \right) \frac{d\varphi}{dt} + \frac{g^2 r^2 N^2}{2} (x - yN)^2 \frac{d^2\varphi}{dN^2} \right] dt + \sigma dW(t). \quad (34)$$

$\varphi(N_t)$ is twice continuously differentiable, with the first and second derivatives being described by Eqs. (35) and (36), respectively:

$$\frac{d\varphi}{dN} = \frac{\sigma}{gr(x - yN)N} \quad (35)$$

$$\frac{d^2\varphi}{dN^2} = -\frac{\sigma}{gr} \frac{x + 2yN}{(x - yN)^2 N^2}. \quad (36)$$

After substituting Eqs. (33), (35) and (36) into Eq. (34), the latter becomes:

$$dn = \left[rN(x - y) \ln \left(\frac{K}{N} \right) \frac{\sigma}{grN(x - yN)} + \frac{g^2 r^2 N^2}{2} (x - yN)^2 \frac{\sigma^2}{gr} \left(\frac{x - 2yN}{N^2 (x - yN)^2} \right) \right] dt + \sigma dW(t) \quad (37)$$

or

$$dn = \left\{ \ln \left(\frac{K}{N} \right) - \frac{\sigma^2 gr}{2} \left[\ln \left(a + b\frac{P}{K} \right) - 2 \frac{bP}{aK + bP} N \right] \right\} dt + \sigma dW(t). \quad (38)$$

Therefore, the initial stochastic differential equation described by Eq. (27) can be expressed in the form

$$dn = \tilde{f} dt + \sigma dW_t, \quad (39)$$

where

$$\tilde{f} = \frac{\sigma}{g} \ln \left[\frac{K}{x} \left(y + e^{-\frac{ngrx}{\sigma}} \right) \right] - \frac{gr\sigma}{2} \left(x + \frac{2yx}{y + e^{-\frac{ngrx}{\sigma}}} \right). \quad (40)$$

5.3.2. Linearization

The linearization of Eq. (38) is derived by applying the methodology proposed by Shoji and Ozaki (1998), which yields

$$n_t = n_s + \frac{\tilde{f}}{L_s} (e^{L_s(t-s)} - 1) + \frac{M_s}{L_s^2} \left\{ \left[e^{L_s(t-s)} - 1 \right] - L_s(t - s) \right\} + \sigma \int_s^t e^{L_s(t-u)} dW_u. \quad (41)$$

In Eq. (41), $\tilde{f}$ is given by Eq. (40), and L_s and M_s by

$$L_s = \frac{\partial \tilde{f}}{\partial n} \quad \text{and} \quad (42)$$

$$M_s = \frac{\sigma^2}{2} \frac{\partial^2 \tilde{f}}{\partial n^2} + \frac{\partial \tilde{f}}{\partial t}. \quad (43)$$

The second part of the right-hand-side of Eq. (43) is equal to 0, since it is the case of an autonomous equation. Moreover, the integral term in Eq. (41) has a Normal distribution, with mean 0 and variance given by

$$\mathrm{Var}_s(n_t) = \sigma^2 \left(\frac{e^{2L_s(t-s)} - 1}{2L_s} \right). \quad (44)$$

After solving Eq. (41), corresponding values of N_t can be calculated by applying the relationship given by Eq. (33).

5.4. Parameter estimation of the discretized process

Since the discretized process locally follows a Gaussian distribution, the log-likelihood can be obtained easily. Therefore, the maximum likelihood method can be used to estimate the model's parameters. The log-likelihood function is given by

$$\log(p(n_1, n_2, \ldots, n_k)) = \sum_{j=1}^{k-1} \log(p(n_{j+1}|n_j)) + \log(p(n_1)). \quad (45)$$

However, since n was derived by applying the transformation $\varphi(t)$ to the initial stochastic differential equation of Eq. (27), the same transformation should be applied to Eq. (45). This is necessary in order to remove the large variance, caused by changes in the ns, and is achieved by applying the transformation rule of a density function

$$p(N_1, N_2, \ldots, N_k) = p(n_1, n_2, \ldots, n_k) \times \left| \frac{\partial(n_1, n_2, \ldots, n_k)}{\partial(N_1, N_2, \ldots, N_k)} \right|, \quad (46)$$

where $\left| \frac{\partial(n_1, n_2, \ldots, n_k)}{\partial(N_1, N_2, \ldots, N_k)} \right|$ is the Jacobian, and is equivalent to $\prod_{j=1}^{k-1} \left| \frac{d\varphi}{dN} \right|_{N=N_j}$.

Finally, the log-likelihood for the stochastic population-diffusion model, based on the linearized approach of Eq. (41), is given by

$$\log(p(N_1, N_2, \dots, N_k)) = -\frac{1}{2}\sum_{j=1}^{k-1}\left\{\frac{(n_{j+1}-E_j)^2}{V_j} + \log(2\pi V_j)\right\} + \log(p(n_1)) + \sum_{j=1}^{k-1}\log\left(\left|\frac{d\varphi}{dN}\right|_{N=N_j}\right), \quad (47)$$

where

$$E_j = n_j + \frac{f}{L_j}(e^{L_j\Delta t}-1) + \frac{M_j}{L_j^2}\left\{(e^{L_j\Delta t}-1) - L_j\Delta t\right\} \quad (48)$$

$$V_j = \frac{(e^{L_j\Delta t}-1)}{2L_j}\sigma^2 \quad (49)$$

$$L_j = \frac{\partial f}{\partial n} \quad (50)$$

$$M_j = \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial n^2} + \frac{\partial f}{\partial t} \quad \text{and} \quad (51)$$

$$f = \tilde{f}\frac{\partial\varphi}{\partial N} + \frac{g^2}{2}\frac{\partial^2\varphi}{\partial n^2}. \quad (52)$$

Explicit formulas for the calculation of the parameter that maximizes the log-likelihood function do not always exist, as is the case for the function in Eq. (47). Therefore, heuristic methods, such as Genetic Algorithms, are usually employed for the estimation of a model's parameters. Genetic algorithms, as posited by Goldberg (1989), are search algorithms based on the mechanisms of natural selection and natural genetics. The key points of the process are reproduction, crossover and mutation, which are performed according to a given probability, just as happens in the real world. Reproduction involves copying (reproducing) solution vectors, crossover involves swapping partial solution vectors, and mutation is the process of randomly changing a cell in the string of the solution vector, thus preventing the possibility of the algorithm being trapped. The process continues until it reaches the optimal solution of the fitness function. Genetic algorithms have been applied for high technology demand estimation (Venkatesan & Kumar, 2002).

In the context of the present work, the fitness function to be optimized is the log-likelihood function, described by Eq. (47), seeking to find its maximum value, based on the historical data. In order to evaluate the model's accuracy, estimation was based on data from the period from 1994 to 1999. After 1999, diffusion slowed down, as the market was about to reach saturation, and the rest of the models evaluated estimated extremely high saturation levels which did not correspond to the actual values recorded in later years.

The application of the genetic algorithms was based on initial values equal to the ones estimated in the deterministic case of the model. In addition, other, randomly chosen, initial values were also used, in order to avoid the algorithm getting trapped at local maxima. After performing about 100,000 iterations, the process ended up with a saturation level value (K) of about 1.013. This can be considered as quite a consistent result, since it does not diverge from the value calculated for the deterministic analogue (1.031).

6. Results of the stochastic model

Forecasts of the stochastic model are provided by simulating the solution given in the preceding section, based on the parameters estimated by the maximum likelihood method. Realizations of ΔW_t were generated using the well-known Box–Muller method (Box & Muller, 1958).

The model was evaluated for the case of Greece, and the corresponding results are shown in Fig. 14. As observed, the stochastic formulation of the model indicates the upper and lower bounds of the expected demand values. However, the deterministic analysis only provides the lower values. This simulation is performed using parameter values estimated using data for the period from 1995 to 1999 (holdback sample). The important outcome that accompanies the results is the estimation of the uncertainty of the forecasted values, together with the possible values diffusion can take, according to the diffusion dynamics. As observed, the deviation from the mean value can be quite high. In the case of Greece, the deterministic PDM forecasted a saturation level of 1.24. However, the stochastic analysis revealed that possible values vary from about 1.06 to 1.35. This stochastic analysis, together with the corresponding results, can be a valuable guide for the construction of business plans and investments, since it provides a measure of the deviation of the future expectations.

7. Conclusions and future work

The aim of this research was to contribute to the existing knowledge regarding diffusion estimation and forecasting, for both research and practice. This was attempted by developing a diffusion model which explicitly incorporates the size of the market, as expressed by the corresponding population. The two formulations of the proposed PDM, both deterministic and stochastic, provide quite accurate results in terms of diffusion estimation and forecasting, especially when the rest of the aggregate models failed to do so, such as when the diffusion is described by a high adoption rate.

In such cases, the widely used diffusion models cannot predict the saturation level accurately, mainly because the population is not taken into account. The proposed model was evaluated using historical data from 22 European countries, and provided accurate forecasts, even from the early stages of the corresponding diffusion processes.

This study helped to move the consideration of innovation diffusion towards the development of a diffusion forecasting framework, incorporating a number of driving factors and decision variables, and also helped increase the understanding of high technology market trends. Moreover, the estimation of the underlying level of uncertainty will stimulate further research, in order to produce more accurate forecasts.

The model provides critical inputs for strategic planning and decision making, in an increasingly competitive

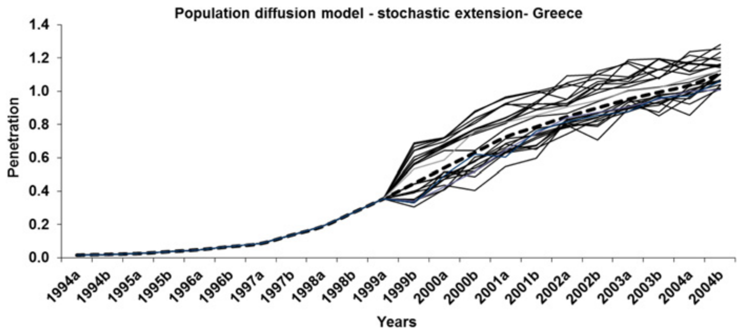


Fig. 14. Simulation of the stochastic formulation of the population model. The dashed line corresponds to the mean value of the process, which coincides with the results of the deterministic formulation of the model.

Table 3
Percentage diffusion over the population of mobile telephony for countries in the wider European area.
Source: ITU.

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Austria	0.05	0.07	0.14	0.28	0.52	0.76	0.81	0.83	0.89	0.98	1.06	1.13	1.19
Belgium	0.02	0.05	0.09	0.17	0.31	0.55	0.75	0.78	0.83	0.88	0.92	0.94	1.03
Czech Republic	0.00	0.02	0.05	0.09	0.19	0.42	0.68	0.84	0.95	1.05	1.15	1.22	1.25
Denmark	0.15	0.24	0.27	0.36	0.49	0.63	0.74	0.83	0.88	0.95	1.00	1.07	1.14
Estonia	0.02	0.05	0.11	0.18	0.29	0.41	0.48	0.65	0.78	0.94	1.09	1.25	1.48
Finland	0.20	0.29	0.41	0.54	0.63	0.72	0.80	0.87	0.91	0.95	1.00	1.08	1.15
France	0.02	0.04	0.10	0.19	0.35	0.49	0.62	0.65	0.69	0.74	0.79	0.85	0.90
Germany	0.05	0.07	0.10	0.17	0.28	0.59	0.68	0.72	0.78	0.86	0.96	1.04	1.18
Greece	0.02	0.05	0.08	0.19	0.35	0.54	0.72	0.84	0.81	0.84	0.92	0.99	1.10
Hungary	0.03	0.05	0.07	0.11	0.16	0.30	0.49	0.68	0.78	0.86	0.92	0.99	1.10
Ireland	0.00	0.07	0.13	0.23	0.41	0.65	0.77	0.76	0.87	0.95	1.03	1.13	1.16
Italy	0.00	0.11	0.20	0.35	0.52	0.73	0.89	0.94	0.98	1.08	1.23	1.35	1.35
Latvia	0.00	0.01	0.03	0.07	0.12	0.17	0.28	0.39	0.52	0.66	0.81	0.95	0.97
Lithuania	0.00	0.01	0.05	0.08	0.10	0.15	0.29	0.47	0.61	0.89	1.27	1.38	1.45
Luxembourg	0.06	0.10	0.15	0.29	0.46	0.70	0.93	1.06	1.19	1.02	1.10	1.17	1.30
Malta	0.03	0.03	0.04	0.06	0.09	0.29	0.61	0.70	0.73	0.77	0.81	0.86	0.91
Netherlands	0.03	0.06	0.11	0.21	0.42	0.68	0.76	0.75	0.82	0.91	0.97	1.06	1.06
Portugal	0.03	0.06	0.14	0.29	0.45	0.65	0.78	0.84	0.96	1.01	1.09	1.16	1.27
Slovak Republic	0.01	0.01	0.04	0.09	0.12	0.23	0.40	0.54	0.68	0.79	0.84	0.91	1.13
Spain	0.02	0.07	0.10	0.15	0.35	0.60	0.72	0.81	0.88	0.91	0.99	1.06	1.09
Sweden	0.22	0.28	0.35	0.46	0.57	0.72	0.81	0.89	0.98	0.98	1.01	1.06	1.14
United Kingdom	0.10	0.12	0.15	0.25	0.46	0.74	0.79	0.83	0.92	1.00	1.10	1.17	1.18

environment, by making accurate a priori estimates of the diffusion pattern available. The results are enhanced by the estimation of the level of uncertainty, provided by the stochastic analysis. As observed, the stochastic formulation of the model provides an indication of the upper and lower bounds of the expected diffusion values, in contrast to the deterministic analysis, which provides only the lower limits. Thus, the model will enhance our ability to develop effective strategies for introducing and adopting new technologies.

As with any model, this model has certain limitations that need to be investigated in future work, the first of which is the population size, which was considered to be constant over the evaluation period. Therefore, incorporating the population rate of change would be expected to improve the forecasting and estimation of the market saturation level. Thus, a framework accommodating the flexibility of the market ceiling, due to a varying population, should be developed.

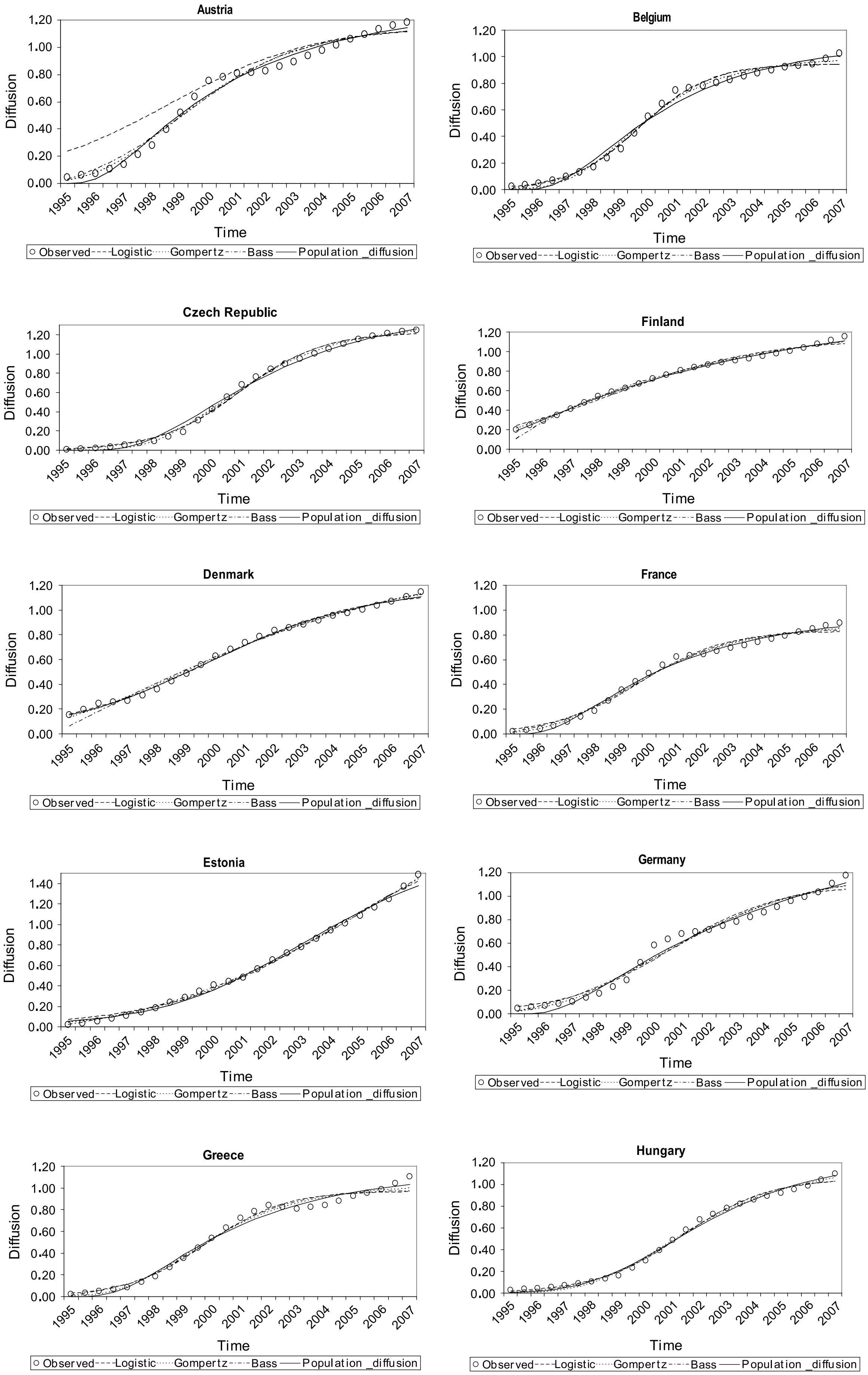
A second limitation of the current model is the assumption that decision variables, such as price, are exogenous to the system. Allowing these variables to be endogenous to the proposed model would improve the forecasting further. In this direction, the development of appropriate price elasticity functions and their incorporation into the model's structure would provide significant information regarding the impact of the price on the diffusion process.

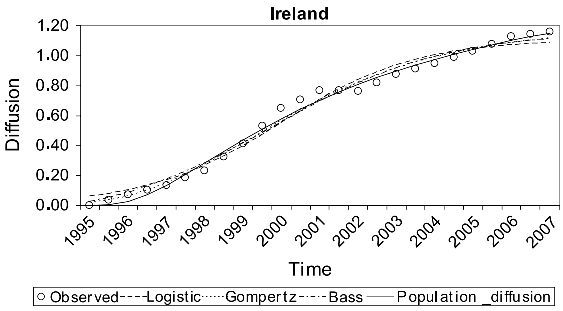
The inclusion of the above considerations in the diffusion framework would provide important managerial insights and new directions for research.

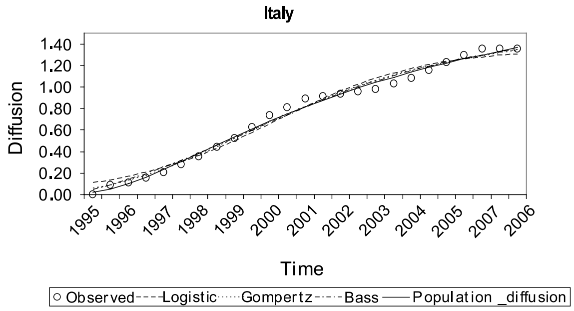
Appendix A. PDM evaluation results

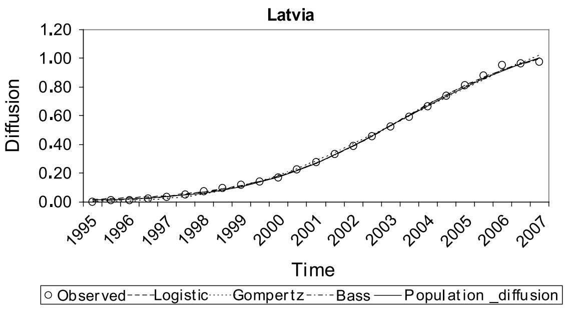
This appendix presents the detailed evaluation results from the PDM.

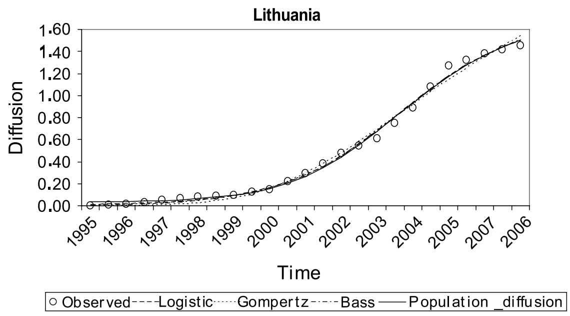
Moreover, the evaluation results from the Gompertz, linear logistic and Bass models for the same dataset are also included, for the sake of comparison. In addition, the historical data for the evaluation process are given in Table 3.

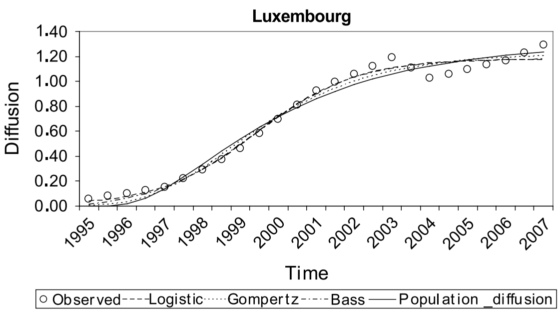

Austria
Belgium
Czech Republic
Finland
Denmark
France
Estonia
Germany
Greece
Hungary
Diffusion
Time
Observed
Logistic
Gompertz
Bass
Population _diffusion

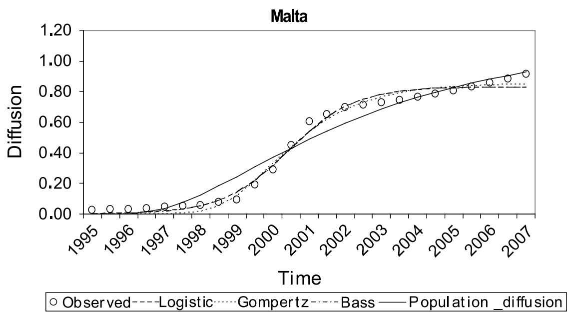

Ireland
Diffusion
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion

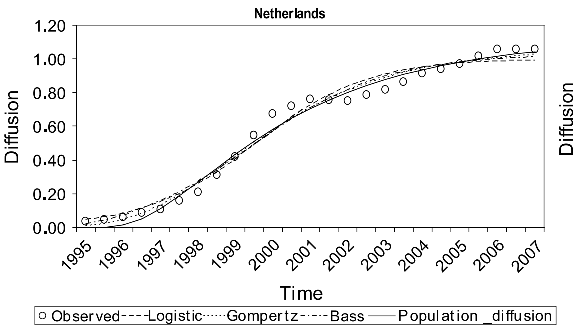

Italy
Diffusion
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2006
Time
Observed Logistic Gompertz Bass Population _diffusion

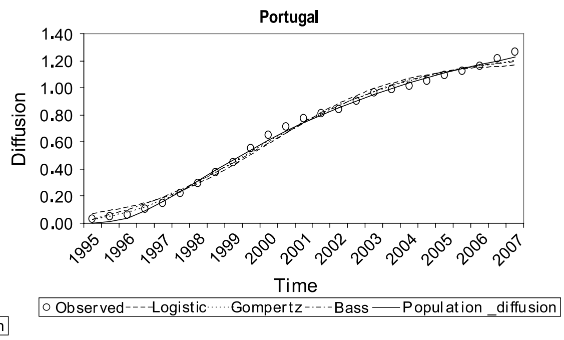

Latvia
Diffusion
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion

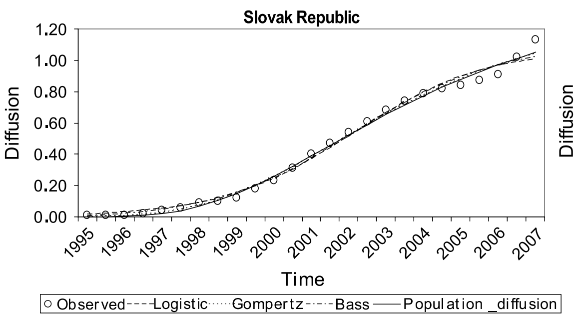

Lithuania
Diffusion
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2006
Time
Observed Logistic Gompertz Bass Population _diffusion

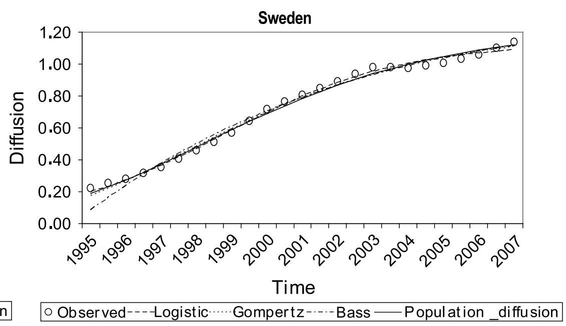

Luxembourg
Diffusion
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion

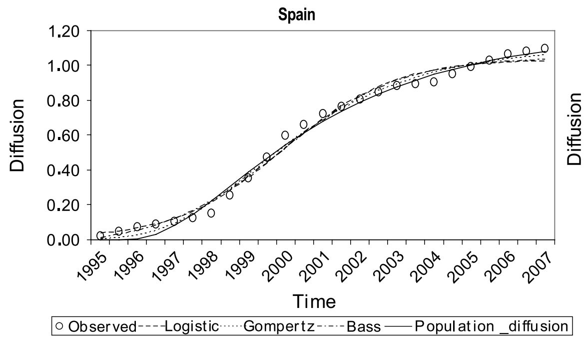

Malta
Diffusion
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion

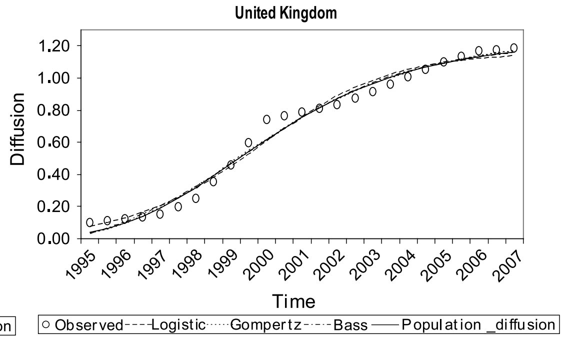

Netherlands
Diffusion
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion


Portugal
Diffusion
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion


Slovak Republic
Diffusion
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion

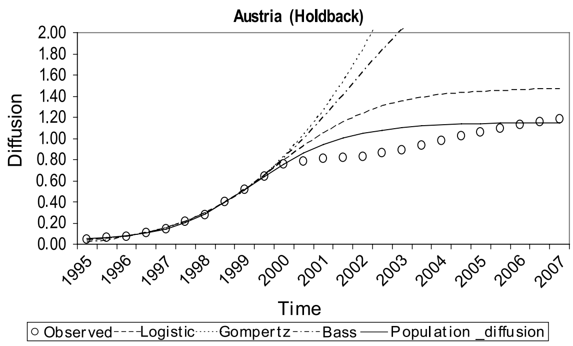

Sweden
Diffusion
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Time
Observed Logistic Gompertz Bass Population _diffusion

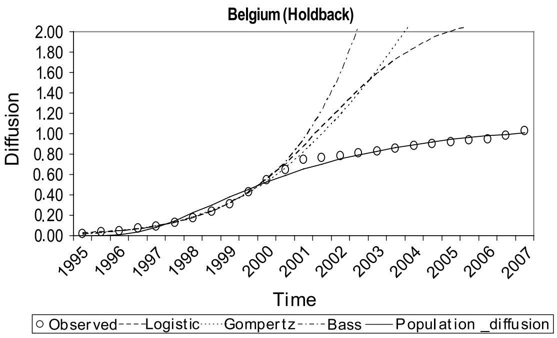


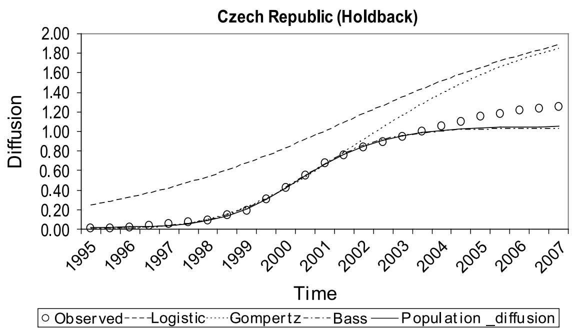


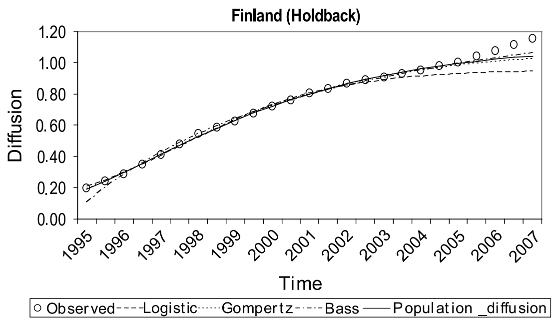
Appendix B. PDM forecasting results

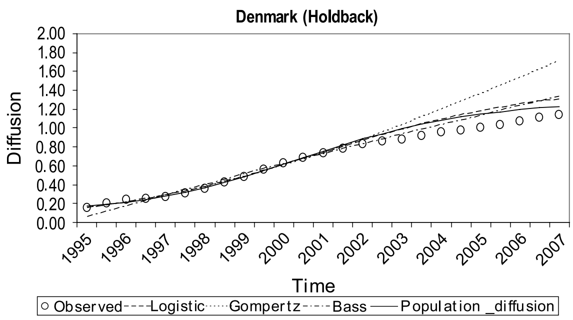
In this appendix, the forecasting ability of the proposed model is evaluated, together with the results of the other participating models, namely the Gompertz, linear logistic and Bass models.

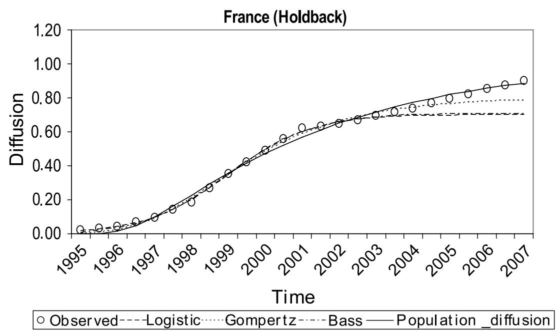


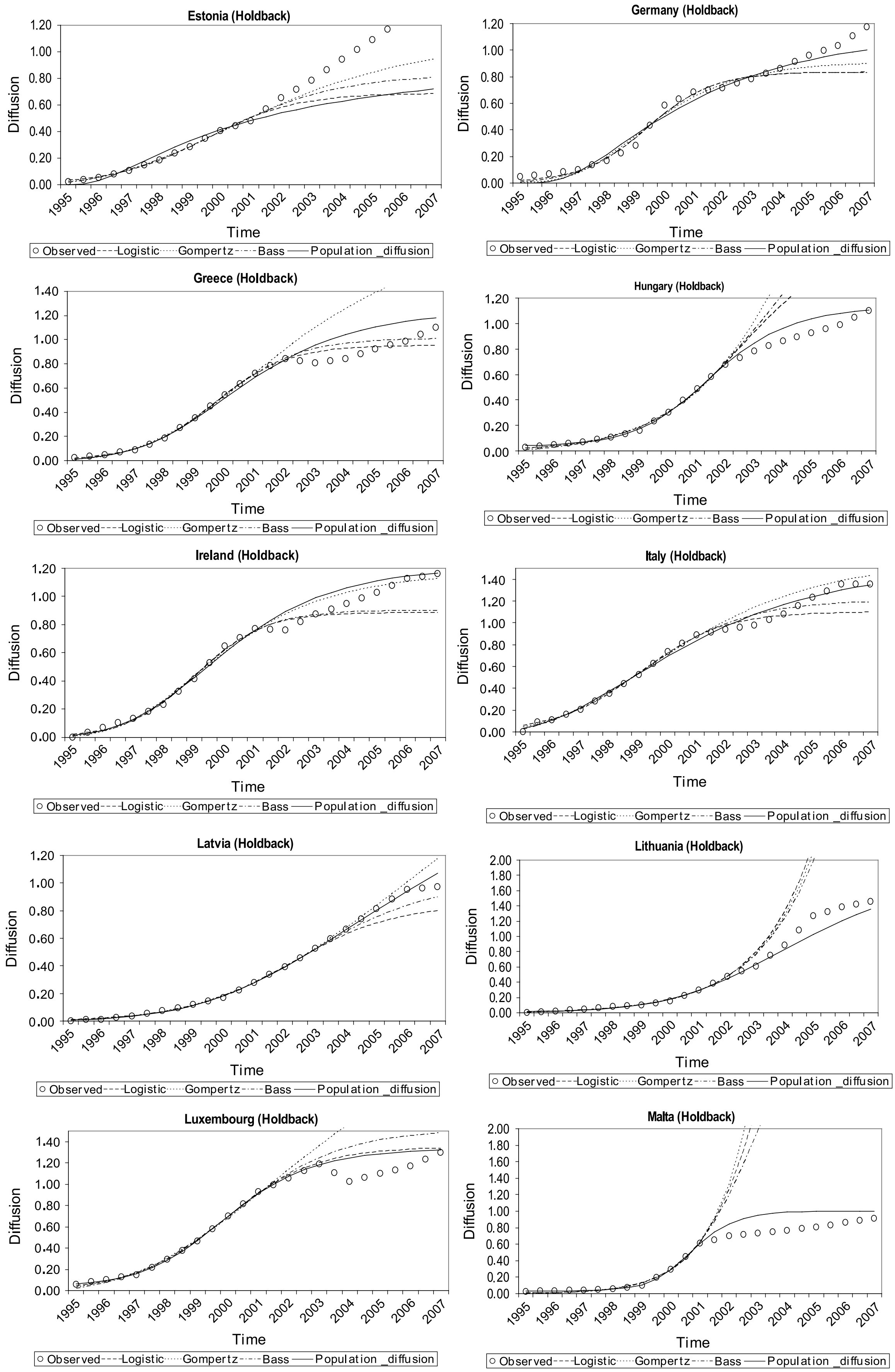


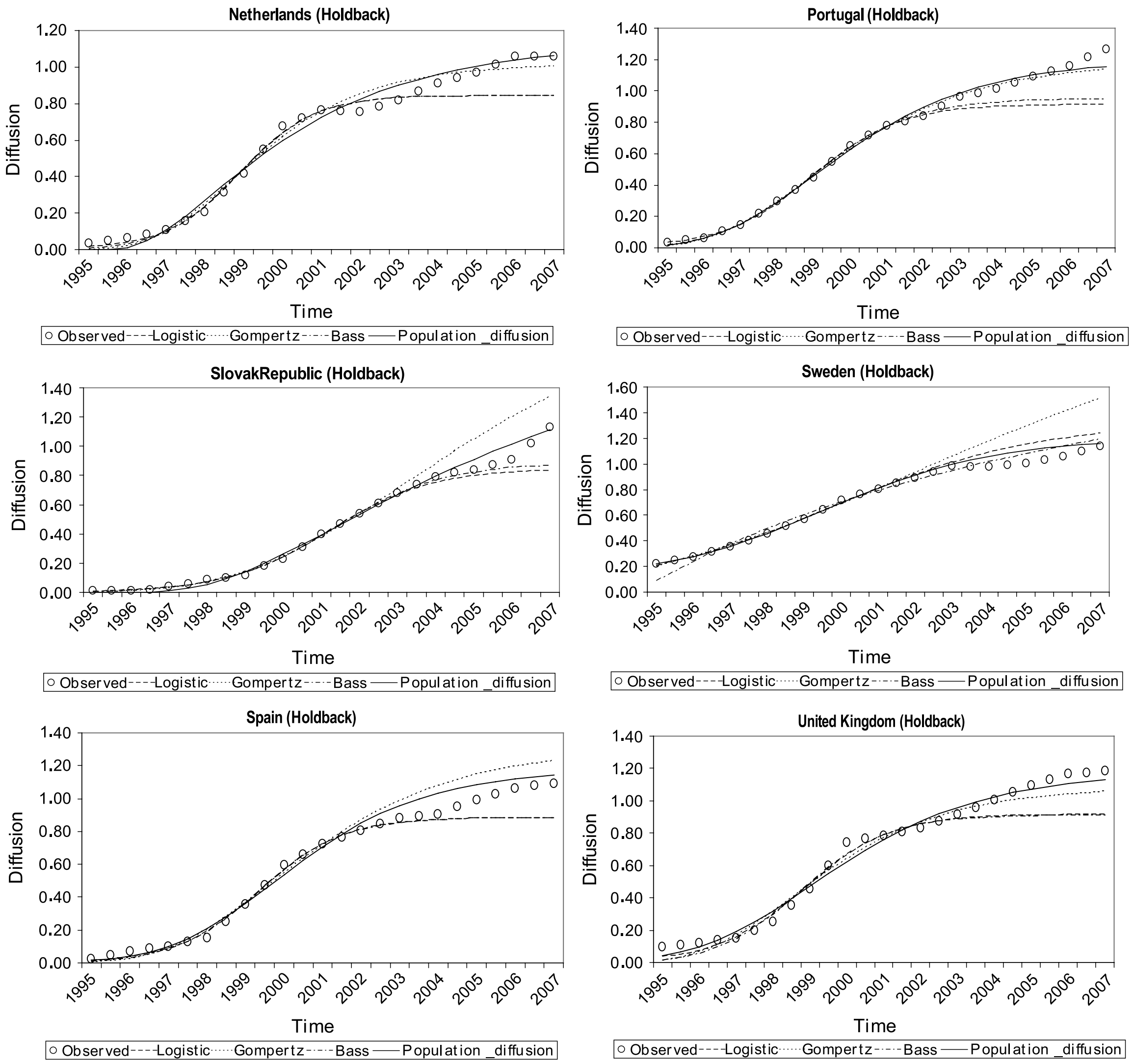









Estonia (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Germany (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Greece (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Hungary (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Ireland (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Italy (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Latvia (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Lithuania (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Luxembourg (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion
Malta (Holdback)
Diffusion
Time
Observed Logistic Gompertz Bass Population _diffusion



References

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science*, *15*, 215–227.

Bemmaor, A. C., & Lee, J. (2002). The impact of heterogeneity and ill-conditioning on diffusion model parameter estimates. *Marketing Science*, *21*, 209–220.

Bergstrom, A. R. (1990). *Continuous time econometric modelling*. Oxford: Oxford University Press.

Bewley, R., & Fiebig, D. G. (1988). A flexible logistic growth-model with applications in telecommunications. *International Journal of Forecasting*, *4*, 177–192.

Box, G., & Muller, M. (1958). A note on the generation of random normal variables. *The Annals of Mathematical Statistics*, *29*, 610–611.

Boyce, W. E., & DiPrima, R. C. (2005). *Elementary differential equations and boundary value problems* (8th ed.). Hoboken, NJ: Wiley.

Debecker, A., & Modis, T. (1994). Determination of the uncertainties in S-curve logistic fits. *Technological Forecasting and Social Change*, *46*, 153–173.

Dembo, A., & Zeitouni, O. (1987). On the parameters estimation of continuous-time ARMA processes from noisy observations. *IEEE Transactions on Automatic Control*, *32*, 361–364.

Eliashberg, J., & Chatterjee, R. (1986). Stochastic issues in innovation diffusion models. In V. Mahajan, & Y. Wind (Eds.), *Innovation diffusion models of new product acceptance* (pp. 151–199). Cambridge, MA: Bullinger Publishing Company.

Evans, L. C. (2005). *An introduction to stochastic differential equations. Lecture notes*, UC Berkeley, Department of Mathematics. Available at: http://math.berkeley.edu/~evans/SDE.course.pdf.

Fildes, R., & Kumar, P. (2002). Telecommunications demand forecasting—a review. *International Journal of Forecasting*, *18*, 489–522.

Fisher, J. C., & Pry, R. H. (1971). A simple substitution model of technological change. *Technological Forecasting and Social Change*, *3*, 75–88.

Gardiner, C. W. (2004). *Handbook of stochastic methods for physics, chemistry, and the natural sciences* (3rd ed.). Berlin: Springer.

Geroski, P. A. (2000). Models of technology diffusion. *Research Policy*, *29*, 603–625.

Giovanis, A. N., & Skiadas, C. H. (1999). A stochastic logistic innovation diffusion model studying the electricity consumption in Greece and the United States. *Technological Forecasting and Social Change*, *61*, 235–246.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Reading, Mass: Addison-Wesley.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, *115*, 513–585.

Gruber, H. (2001). Competition and innovation: the diffusion of mobile telecommunications in Central and Eastern Europe. *Information Economics and Policy*, *13*, 19–34.

Gruber, H., & Verboven, F. (2001). The diffusion of mobile telecommunications services in the European Union. *European Economic Review*, *45*, 577–588.

Jain, A., & Rai, L. P. (1988). Diffusion-models for technology-forecasting. *Journal of Scientific and Industrial Research*, *47*, 419–429.

Karmeshu, & Pathria R. K., (1980). Stochastic-evolution of a non-linear model of diffusion of information. *Journal of Mathematical Sociology*, *7*, 59–71.

Kumar, V., Ganesh, J., & Echambadi, R. (1998). Cross-national diffusion research: what do we know and how certain are we? *Journal of Product Innovation Management*, *15*, 255–268.

Kumar, V., & Krishnan, T. V. (2002). Multinational diffusion models: an alternative framework. *Marketing Science*, *21*, 318–330.
Linton, J. D. (2002). Forecasting the market diffusion of disruptive and discontinuous innovation. *IEEE Transactions on Engineering Management*, *49*, 365–374.
Mahajan, V., & Muller, E. (1979). Innovation diffusion and new product growth-models in marketing. *Journal of Marketing*, *43*, 55–68.
Mahajan, V., Muller, E., & Bass, F. M. (1990). New product diffusion-models in marketing—a review and directions for research. *Journal of Marketing*, *54*, 1–26.
Mahajan, V., & Peterson, R. A. (1979). First-purchase diffusion-models of new-product acceptance. *Technological Forecasting and Social Change*, *15*, 127–146.
Meade, N. (1989). Technological substitution—framework of stochastic-models. *Technological Forecasting and Social Change*, *36*, 389–400.
Michalakelis, C., Dede, G., Varoutas, D., & Sphicopoulos, T. (2008). Impact of cross-national diffusion process in telecommunications demand forecasting. *Telecommunication Systems*, *39*, 51–60.
Michalakelis, C., Varoutas, D., & Sphicopoulos, T. (2008). Diffusion models of mobile telephony in Greece. *Telecommunications Policy*, *32*, 234–245.
Milstein, G. N. (1995). *Numerical integration of stochastic differential equations*. Dordrecht: Kluwer.
Nikolau, J. (2005). Methods for simulating non-linear stochastic differential equations in R1. *Journal of Statistical Computation and Simulation*, *75*(8), 595–609.
Oksendal, B. (2003). *Stochastic differential equations: an introduction with applications* (6th ed.). Berlin: Springer.
Ozaki, T. (1985). Non-linear time series models and dynamical systems. In E. J. Hannan (Ed.), *Handbook of statistics, vol. 5* (pp. 25–83). Amsterdam: North-Holland.
Rai, L. P. (1999). Appropriate models for technology substitution. *Journal of Scientific and Industrial Research*, *58*, 14–18.
Ruiz-Conde, E., Leeflang, P. S. H., & Wieringa, J. E. (2006). Marketing variables in macro-level diffusion models. *Journal für Betriebswirtschaft*, *56*, 155–183.
Seber, G. A. F., & Wild, C. J. (2003). *Nonlinear regression*. Hoboken, New Jersey: Wiley.
Shoji, I., & Ozaki, T. (1998). Estimation for nonlinear stochastic differential equations by a local linearization method. *Stochastic Analysis and Applications*, *16*, 733–752.
Singer, H. (1993). Continuous-time dynamical systems with sampled data, errors of measurement and unobserved components. *Journal of Time Series Analysis*, *14*, 527–545.
Skiadas, C. H. (1987). Two simple-models for the early and middle stage prediction of innovation diffusion. *IEEE Transactions on Engineering Management*, *34*, 79–84.
Skiadas, C. H., & Giovanis, A. N. (1997). A stochastic Bass innovation diffusion model for studying the growth of electricity consumption in Greece. *Applied Stochastic Models and Data Analysis*, *13*, 85–101.
Teng, J. T. C., Grover, V., & Guttler, W. (2002). Information technology innovations: general diffusion patterns and its relationships to innovation characteristics. *IEEE Transactions on Engineering Management*, *49*, 13–27.
Venkatesan, R., & Kumar, V. (2002). A genetic algorithms approach to growth phase forecasting of wireless subscribers. *International Journal of Forecasting*, *18*, 625–646.
Wrede, R., & Spiegel, M. (2002). *Theory and problems of advanced calculus* (2nd ed.). New York: McGraw-Hill.

Christos Michalakelis holds a degree in Mathematics (University of Athens, Department of Mathematics), an M.Sc. degree in Software Engineering (The University of Liverpool, UK), and an M.Sc. degree in "Administration and Economics of Telecommunication Networks" from the National and Kapodistrian University of Athens (Department of Informatics and Telecommunications, Interfaculty course of the Departments of Informatics and Telecommunications and Economic Sciences). He also holds a Ph.D. degree in the area of technoeconomics, and with a focus on the demand estimation and forecasting of high technology products.

He is currently a Lecturer on telecommunications technoeconomics in the Department of Informatics and Telematics at the Harokopio University of Athens. Previously, he was with the Greek Ministry of Education for 7 years, in the Managing Authority of Operational Program for Education and Initial Vocational Training, as an IT manager. Mr. Michalakelis has participated in a number of projects concerning the design and implementation of database systems, and now participates in several technoeconomic activities for telecommunications, networks and services, such the CELTIC/ECOSYS project pricing and regulation. He has also developed or made a major contribution to the development of a number of information systems and applications.

Mr. Michalakelis' research interests include the study of the diffusion of high technology products and services, the application of economic theories to telecommunications, and the wider area of informatics and price modeling and indexing. He has authored a number of papers presented at conferences, and has published many papers in scientific journals and books. He serves as a reviewer for a number of scientific journals.

Thomas Sphicopoulos received a degree in physics from Athens University in 1976, the D.E.A. degree and a Doctorate in Electronics, both from the University of Paris VI, in 1977 and 1980 respectively, and the Doctorat Es Science from the Ecole Polytechnique Federale de Lausanne in 1986. From 1976 to 1977, he worked in Thomson CSF Central Research Laboratories on microwave oscillators. From 1977 to 1980, he was an Associate Researcher in Thomson CSF Aeronautics Infrastructure Division. In 1980, he joined the Electromagnetism Laboratory of the Ecole Polytechnique Federal de Lausanne, where he researched applied electromagnetism. Since 1987, he has been at Athens University, engaged in research on broadband communications systems. He was elected an Assistant Professor of Communications in the Department of Informatics & Telecommunications in 1990, an Associate Professor in 1993, and a Professor in the same department in 1998.

His main scientific interests are optical communication systems and networks, and technoeconomics. He has led about 50 national and European R & D projects.

Professor Sphicopoulos has published more than 150 articles in scientific journals and conference proceedings, and is an advisor for several organizations.