

Petre, Konstantin; Varoutas, Dimitris

## Conference Paper

# On the application of Machine Learning in telecommunications forecasting: A comparison

31st European Conference of the International Telecommunications Society (ITS): "Reining in Digital Platforms? Challenging monopolies, promoting competition and developing regulatory regimes", Gothenburg, Sweden, 20th - 21st June 2022

### Provided in Cooperation with:

International Telecommunications Society (ITS)

*Suggested Citation:* Petre, Konstantin; Varoutas, Dimitris (2022) : On the application of Machine Learning in telecommunications forecasting: A comparison, 31st European Conference of the International Telecommunications Society (ITS): "Reining in Digital Platforms? Challenging monopolies, promoting competition and developing regulatory regimes", Gothenburg, Sweden, 20th - 21st June 2022, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/265665>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# On the application of Machine Learning in telecommunications forecasting: A comparison

Konstantin Petre and Dimitris Varoutas  
Department of Informatics and Telecommunications  
National and Kapodistrian University of Athens  
Athens, Greece  
[kpetre,d.varoutas@di.uoa.gr](mailto:kpetre,d.varoutas@di.uoa.gr)<sup>1</sup>

## Abstract

Over the past few decades, a large number of research papers has published focused on forecasting ICT products using various diffusion models like logistic, Gompertz, Bass, etc. Much less research work has been done towards the application of time series forecasting in ICT such as ARIMA model which seems to be an attractive alternative. More recently with the advancement in computational power, machine learning and artificial intelligence have become popular due to superior performance than classical models in many areas of concern. In this paper, broadband penetration is analysed separately for all OECD countries, trying to figure out which model is superior in most cases and phases in time. Although diffusion models are dedicated for this purpose, the ARIMA model has nevertheless shown an enormous influence as a good alternative in many previous works. In this study, a new approach using LSTM networks stands out to be a promising method for projecting high technology innovations diffusion.

## Keywords

Diffusion models, ARIMA, LSTM, broadband penetration forecasting

## 1. Introduction

As can be seen over the past decades, there have been published numerous research papers exploring different innovation forecasting techniques. In [1], Meade and Islam review literature associated with modeling and forecasting innovation diffusion, highlighted many previous works of comparative forecasting accuracy. Most of them used general data sets to examine, however there are cases with specific purpose in telecommunication products. In a previous work [2], Meade and Islam investigated 17 diffusion models represented by 25 time series describing telephone penetration in 15 different countries. An early study of Gottardi and Scarso in [3], compared the forecasting accuracy of the Box-Jenkins and diffusion models, on the basis of many different time series. However, many of the data sets were found to be out of scope, as they described consumption or production, rather than diffusion. The use of time series ARIMA models has not been widely investigated in the case of forecasting the diffusion of innovations. In [4], Christodoulos *et al* presented a methodology of providing short-term forecasts for the world broadband and mobile telecommunications penetration. The paper focused on the improvement of the short-term prediction by combining ARIMA and diffusion models. As opposed to [3], ARIMA model found to be superior than classic diffusion model whereas the combined model yielded better results. Recently, in [5] Panigrahi and Behera constructed hybrid models by suitably combining linear models like ARIMA with nonlinear models like artificial neural network (ANN). Sixteen datasets and five different models from the

---

<sup>1</sup> *Disclaimer: The positions illustrated in this article do reflect the personal opinions of the authors and not the official positions of the organizations that the authors are employed to. These official positions of the organizations are presented via their public announcements and official documents such as decisions, etc.*

literature were considered and in most of the cases hybrid model outperformed individual models. Another study with machine learning explores the application of machine learning methodologies to forecast video subscribers [6]. Their results showed that boosted trees, similar to XGBoost, outperformed SVRs, ANNs and traditional random forests.

## 2. Dataset and Methods

### 2.1 The Dataset

The dataset used in this study consists of annual observations from 38 OECD countries for the period between 1998 and 2020 published by The World Bank. The countries included in the analysis are listed in Table 1 with the actual number of subscriptions. Figure 1 shows the corresponding penetration diagram per population for the comparison to be relevant. As can be seen Switzerland is among the countries with highest penetration relative to population (47%) but on the contrary there are also countries with very low penetration that may affect the results of this research and could well be omitted (Turkey, Poland, Chile, Costa Rica and Colombia).

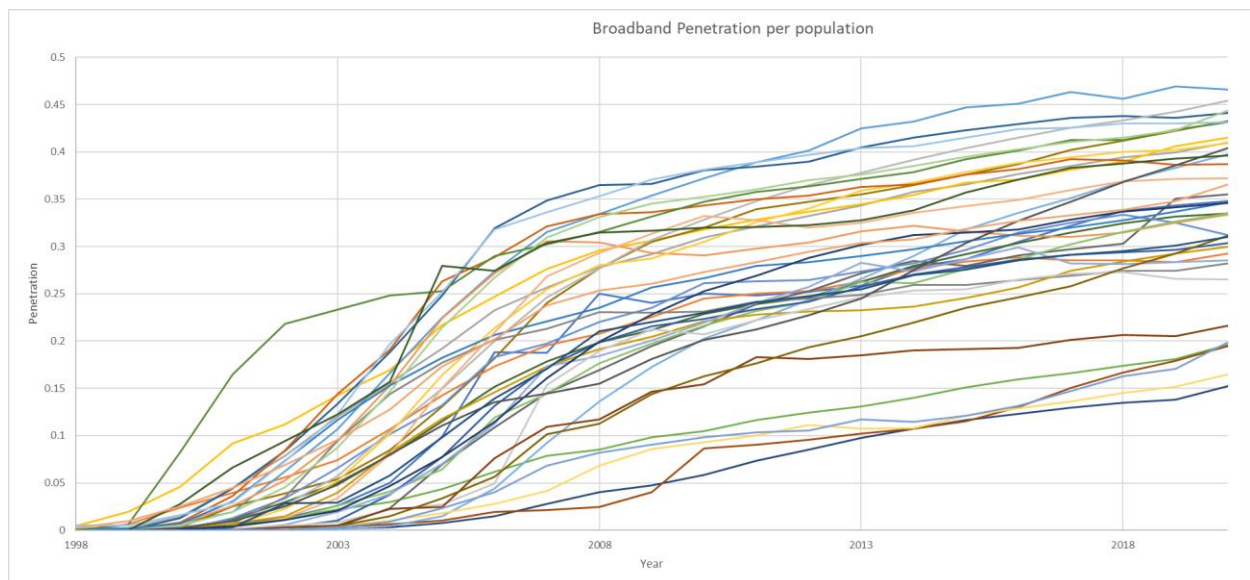


Figure 1: Broadband penetration per population

Country Name	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Australia				122800	258100	516800	1012000	2016000	3900000	3900000	5315000	5221000	5510000	5552000	5735000	5981000	6536000	6828000	7374000	7922000	8427316	8705523	8937550
Austria		50900	190500	320600	451000	601000	870000	1174000	1432000	1622000	1729000	1878524	2050400	2097700	2130200	2232500	2359000	2455500	2523300	2511200	2521100	2519000	2606000
Belgium	10924	23000	144215	460011	815418	1242928	1619944	2010584	2451585	2729624	2982500	3153262	3373143	3543759	3692009	3828918	4011201	4121049	4270309	4378973	4502950	4590707	4734210
Canada	140000	582000	1410932	2836000	3515000	4513000	5416000	7004000	8044000	9075322	9842303	10290000	10817103	11282579	11689914	12094461	12568000	13115376	13386496	13923805	14445606	15273496	15776602
Switzerland			56416	140000	455220	783874	1227397	1669217	2050184	2380588	2556297	2741199	2913517	3078338	3212371	3438094	3535986	3700563	3775916	3916720	3883878	4022520	4023000
Chile			7680	66723	188454	352234	478883	708564	1011646	1302310	1427178	1654676	1789355	2011244	2166330	2295169	2489717	2719439	2904580	3062392	3250888	3429668	3751227
Colombia			8872	13830	34888	64436	127113	318683	628077	1207090	1772088	2114782	2643033	3348142	3938657	4537621	5028595	5525763	5912968	6331032	6678543	6949852	7764772
Costa Rica					8989	14878	27931	44914	83327	95000	109939	182654	396262	419782	448594	484883	516337	558656	639087	744041	834784	904734	992725
Czech Republic			2500	6200	15300	34690	235996	709063	1112500	1496720	1759586	2036110	2261179	2509100	2656980	2856168	2994818	2946626	3069970	3146091	3222835	3739614	3802644
Germany			265000	2100000	3205000	4470000	7000000	10786800	14977200	19751400	22710136	24964600	26161950	27257096	27956700	28641961	29572818	30707429	31861900	33243300	34152033	35071539	36040739
Denmark			67000	238000	451297	718299	1017594	1343855	1735317	1903541	2005681	2022709	2112406	2143247	2179545	2272401	2341759	2404962	2461403	2511871	2535904	2536508	2571736
Spain			76358	466600	1247496	2121930	3401411	5035203	6739110	8055780	9135959	9800597	10652372	11167809	11524543	12252061	13004969	13542906	14112657	14668212	15176954	15616585	15850358
Estonia				17474	46500	90300	138677	179200	246800	264949	294660	313625	347883	348946	349507	360948	371009	390275	414478	428453	441173	431251	415610
Finland			35000	134000	273500	491100	800000	1174200	1429000	1617000	1618000	1565600	1559400	1606000	1647600	1720200	1758500	1729897	1712000	1710000	1737000	1797000	1846000
France	13464	55000	196601	601500	1655000	3569381	6561035	9471000	12711000	15750000	17830000	19852000	21337000	22749000	23980000	24940000	25969000	26867000	27680000	28480000	29100000	29760000	30627000
United Kingdom			52890	330960	1356481	3113702	6123907	9898653	13013000	15606000	17310023	17877223	19151508	20588901	21685668	23039803	23729800	24663109	25473619	26043431	26587110	26871963	27490258
Greece					10476	51455	160113	488180	1017475	1507000	1916630	2252653	2464282	2689428	2913191	3156071	3439034	3615029	3778268	3961864	4111278	4257026	
Hungary			3400	31384	111458	264311	411111	651689	1199190	1454511	1770574	1976723	2159421	2313092	2396035	2597159	2580537	2718794	2814523	2956585	3079549	3189689	3265308
Ireland					10600	41800	152100	322500	561700	758722	891243	976381	1019964	1070776	1112082	1189212	1258758	1309467	1360204	1398798	1430160	1462549	1516252
Iceland		81	2358	10424	24270	41571	55764	78017	87738	100026	106017	107072	109212	111584	113420	117467	119647	124436	128023	134624	137989	139241	141816
Israel				43865	231663	633100	980000	1229626	1421000	1528500	1684000	1723000	1762000	1879029	1937000	2003000	2131000	2173000	2258000	2342000	2435000	2481000	2602079
Italy			115000	390000	850000	2250000	4724500	6822210	8497422	10122126	11276262	12084537	13098028	13518710	13763362	14012976	14382487	14900171	15563279	16586377	17157901	17470489	17855620
Japan	32000	216000	854655	3835000	9397426	14917165	19557146	23301105	26438867	28303003	30117679	32863203	34101778	35696214	36132405	36918650	37788929	38872579	39805586	40532466	41496293	42502489	43633244
Korea, Rep.	14000	278000	3870000	7818020	10405486	11178499	11921439	12190711	14042698	14709998	15474931	16348571	17194272	17859522	18252661	18737514	19198934	20024419	20555683	21195918	21285858	21906172	22330182
Lithuania			2427	20000	66790	129051	234081	368715	559036	590103	633779	679321	727699	768861	836200	801863	833298	857761	798769	788743	791805	796814	
Luxembourg				1215	5827	15351	36500	70100	98930	128680	143200	156120	168368	169753	169700	176500	186800	195300	203100	214600	224300	230100	235155
Latvia			284	3235	10000	19533	49147	60770	109674	338559	412108	454714	434876	457422	476141	496000	505000	503503	519154	526222	525995	508894	503804
Mexico			15000	50000	231486	428371	1057282	1922352	3020000	4504422	7532633	9641168	10582865	11566069	13077276	12747873	13032519	14757686	15923971	17000482	18359028	19355208	21206911
Netherlands	75500	75500	260000	466200	1170966	1988000	3206000	4100000	5192200	5507000	5805000	6129000	6329000	6498007	6654000	6792000	6851000	7029097	7222770	7289580	7406700	7459000	7525016
Norway			23297	88541	205307	398758	671666	991349	1244536	1457265	1577430	1668231	1723355	1785676	1857724	1910720	1977129	2050460	2106975	2165221	2205861	2260605	2387661
New Zealand			4658	17267	43500	83000	191695	321000	480000	680000	850000	980000	1100000	1180000	1270000	1340000	1410000	1450000	1500000	1580000	1650000	1700000	1760000
Poland				12000	121684	195752	875865	945159	2911152	4174003	4462686	5596000	5858498	6972130	6887676	7031055	7233854	7265546	7327918	7630957	7851422	7783887	8212601
Portugal		297	25154	99316	260591	502030	838371	1165440	1423687	1527050	1635427	1911827	2126689	2243036	2390778	2563356	2857776	3142188	3375597	3574638	3784684	3967699	4160895
Slovak Republic					4083	22461	78764	181538	304615	546664	604688	775056	876953	953098	1045490	1109060	1191216	1273812	1336541	1404751	1507998	1596581	1701561
Slovenia				5500	56735	57992	115069	196650	279814	344677	425294	494030	470527	494372	509336	524621	555815	569026	589324	601820	612745	627939	651604
Sweden			249000	587000	840000	1095000	1410000	2522000	2489000	2780298	2902833	2945093	2999537	3027459	3073174	3147615	3281408	3496240	3679768	3854837	3942088	4038725	4101078
Turkey				10915	21205	199324	577931	1589768	2773685	4753757	5756965	6450287	7098163	7591367	7868968	8893391	8866361	9504594	10499692	11924905	13407226	14231978	16734853
United States	705900	2754286	7069874	12792812	19881549	27744342	37352520	51156350	60237701	71704000	77130000	79992000	84522000	88317000	92514000	96032000	97810000	102212000	105714000	108200000	110756000	114269000	120531000

Table 1. Broadband penetration per country (OECD)

For the purposes of our research, the following categories of models will be used comparatively: diffusion models, ARIMA model as well as Machine Learning models.

## 2.2 Diffusion Models

Diffusion models are mathematical functions, mainly of time, used for estimating the adoption of technological innovations or other products or services. The cumulative diffusion shapes of innovations are often described by sigmoid growth patterns. The aggregated S-shaped diffusion models can be derived from a differential equation such as

$$\frac{dN(t)}{dt} = f(N(t)) * (S - N(t)) \quad (1)$$

where  $N(t)$  represents the total penetration at time  $t$ ,  $S$  the saturation level of the specific technology and  $f(N(t))$  is the specific mathematical function that shapes the form of diffusion model. The most common family of  $f(N(t))$  function consists of two parts, a simple constant  $a$  and the product  $b*N(t)$ .

$$f(N(t)) = a + b * N(t) \quad (2)$$

If  $f(N)$  has only the constant term, we obtain the so-called external influence model which concludes with the modified exponential model. In this scenario, the driving power of diffusion consists only of innovators where information about an innovation reaches them only from external sources like mainstream media. On the other hand, if the specific function consists only of the second term we derive the internal influence model where the adoption is based on interactions between prior and potential adopters. Its mathematical form derives from the traditional biological studies on the spread of a disease through a population. This case leads to the well-known logistic model. The Bass model contains both parts in the specific function which takes into account both external and internal influence forces. A more complicated model allows  $f(N(t))$  to be an explicit function of time. The log-logistic formulation for example used by Tanner [7] and Bewley and Fiebig [8], converts  $f$  function as

$$f(N(t)) = b * \frac{N(t)}{S * t} \quad (3)$$

that allows the point of inflection to be data-determined as a desirable outcome the model needs to have that could accommodate different patterns of diffusion. However this model possesses a certain degree of arbitrariness because of invariance of time scale option.

Another form of the differential equation which leads to Gompertz model is

$$\frac{dN(t)}{dt} = b * N(t) * (\ln S - \ln(N(t))) \quad (4)$$

where both Gompertz and Logistic can be transformed and represented as linear functions of time

$$\ln\left(\frac{N(t)}{S-N(t)}\right) = a + b * t \quad (5)$$

$$-\ln\left(-\ln\left(\frac{N(t)}{S}\right)\right) = a + b * t \quad (6)$$

a) Logistic model: 
$$N(t) = \frac{S}{1 + \exp(a - b * t)} \quad (7)$$

b) Bass model: 
$$N(t) = S * \frac{1 - \exp(-(q+p)*t)}{1 + \frac{q}{p} * \exp(-(q+p)*t)} \quad (8)$$

c) Log-Logistic model: 
$$N(t) = \frac{S}{1 + \exp(a - b * \ln(t))} \quad (9)$$

d) Gompertz model: 
$$N(t) = S * \exp(-b * \exp(-c * t)) \quad (10)$$

In all the above cases we can test the relative 4 parameter model that induces the initial level of adoption or otherwise the low level asymptote of the s-shaped curve.

e) Logistic model 4 parameters: 
$$N(t) = D + \frac{S - D}{1 + \exp(a - b * t)} \quad (11)$$

Another technique that can be found in [2] transforms the above models in a manner that the growth curve equations are rewritten in order to represent market development in terms of past observations rather than time. This approach overcomes inconsistency between recent observations and forecasts. For example, if the most recent observation is greater than the forecast from the model for this time slot, then it is conceivable that the next estimate should be less than the next coming observation. So, the Logistic equation in a) it is converted to

f) Local Logistic: 
$$N(t) = S * \frac{N(t-1)}{b1 + b2 * N(t-1)} \quad (12)$$

where  $b1 = S * \exp(-a)$  and  $b2 = 1 - \exp(-a)$

The Gompertz function in the same manner is converted to

g) Local Gompertz: 
$$N(t) = b1 * N(t - 1)^{b2} \quad (13)$$

where  $b1 = S * (1 - \exp(-c))$  and  $b2 = \exp(-c)$

## 2.3 ARIMA model

Introduced by Box and Jenkins, the ARIMA model has been one of the most popular approaches to time-series forecasting. In an ARIMA, the future value of a variable is assumed to be a linear function of several past observations plus random errors. The linear function is based upon three parametric components: autoregression (AR), integration (I), and moving average (MA) and can be denoted by ARIMA(p,d,q), where p is the number of autoregressive terms, d is the number of non-seasonal differences, and q is the number of lagged forecast errors in the prediction equation. Given a time-series of data  $X_t$  where t is an integer index and  $X_t$  are real numbers, corresponding to values at time t, then an ARIMA (p,d,q) model is described by

$$(1 - \sum_{k=1}^p a_k B^k) * (1 - B)^d X_t = (1 + \sum_{k=1}^q b_k B^k) * e_t \quad (14)$$

where B is the backward shift operator, expressing the length of previous data the model uses to provide forecasts,  $a_i$  are the parameters of the autoregressive part of the model, the  $b_i$  are the parameters of the

moving average part and  $\epsilon_t$  are error terms. The error terms  $\epsilon_t$  are generally assumed to be independently, identically distributed variables (iid) sampled from a normal distribution with zero mean. The  $d$  integer is positive and controls the level of differencing. If  $d = 0$ , then the ARIMA is equivalent to an ARMA model. In simple words, AR stands for “autoregressive” and describes a stochastic process that can be described by a weighted sum of its previous values and a white noise error, while MA stands for “moving average” and describes a stochastic process that can be described by a weighted sum of a white noise error and the white noise error from previous periods.

## 2.4 Machine learning models

### 2.4.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory is a special case of Recurrent Neural Network (RNN) method that was initially introduced by Hochreiter and Schmidhuber in [9]. A common recurrent neural network (RNN) consists of an input layer, a hidden layer and the output layer. The hidden layer is considered as the storage area, where information captured in earlier stages persists there for future use. A typical problem arises with RNN is that network is able to remember only few earlier steps and thus fail to do so for longer patterns of data. As Hochreiter and Schmidhuber mentioned error signals flowing backwards in time tend to either blow up or vanish, where in first case may lead to oscillating weights while in second case learning to bridge long time lags takes a prohibitive amount of time or does not work at all.

An LSTM layer consists of a set of recurrently connected blocks, which contain one or more memory cells. These blocks can be thought of as a differentiable version of the memory of personal computer. Each cell contains three multiplicative units, the input, output and forget gate which have similar roles to read, write and reset operations. The gates are the only method the network can interact with cell. Hence, the gates, which are based on sigmoidal neural network layer, enable the cells to optionally let data pass through or disposed. All the gates involved in LSTM are focusing at taking control of the state of the cell. Forget gate decides to keep data or ignore them whether the value of the sigmoid is closer to 1 or 0 respectively. Input gate chooses which data will be stored in the cell or not and finally output gate decides whether the information will be yield out or not.

### 2.4.2 SVR

Support vector machines (SVM) developed by Vapnik [10] is a learning technique that can be seen as a method for training polynomial or Gaussian functions classifier. Initially SVM were created for classification purposes but later with the introduction of  $\epsilon$ -sensitive loss function have been applied to solve non-linear regression problems. It can lead to great performance in time series forecasting problems due to the structure risk minimization principle which seeks to minimize an upper bound of the generalization error consisting of the sum of the training error and a confidence interval. This induction principle is different from the empirical risk minimization principle which only minimizes the training error. The objective of SVR model is to determine a function  $f(x)$ , so as to predict accurately the desired target. Given a dataset of points  $(x_i, y_i)$ , SVR approximates the function using the following equation

$$f(x) = w * \varphi(x) + b \tag{15}$$

where  $\varphi(x)$  denotes a mapping function in the feature space. The coefficients  $w$  and  $b$  are estimated by minimizing the objective risk function

$$\frac{1}{2}|w|^2 + C * \frac{1}{n} * \sum_{i=1}^n L(y_i, f(x_i)) \quad (16)$$

$$L = \{|y_i - f(x_i)| - \varepsilon, |y_i - f(x_i)| \geq \varepsilon\} \quad (17)$$

0, otherwise

Minimizing the first term  $\frac{1}{2} * |w|^2$  will make the function as flat as possible which controls the trade-off between the complexity and the approximation accuracy of the regression model. The second term is the  $\varepsilon$ -sensitive loss function which indicates the fact that it does not penalize errors below  $\varepsilon$ .  $C$  is a regularized constant determining the trade-off between the training error and model flatness. To get  $w$  and  $b$  equation () is transformed using the slack variables to

$$\frac{1}{2}|w|^2 + C * \frac{1}{n} * \sum_{i=1}^n \xi_i + \xi'_i \quad (18)$$

with constraints

$$\begin{aligned} y_i - w * \varphi(x_i) - b &\leq \varepsilon + \xi_i \\ w * \varphi(x_i) + b - y_i &\leq \varepsilon + \xi'_i \\ \xi_i &\geq 0, \xi'_i \geq 0 \end{aligned}$$

Finally using Langrange multipliers, the function() has the following form

$$f(x) = \sum_{i=1}^n (a_i - a'_i) * K(x_i, x) + b \quad (19)$$

where  $a_i$  and  $a'_i$  are the Langrange multipliers and  $K(x_i, x)$  is called the kernel function.

### 3 Methodology

For every year in the dataset, each model will try to predict every country's penetration using only its history, and averaged results will be obtained in every stage of diffusion for 5 steps ahead of prediction. For a country with  $n$  years of data, while using  $n-m$  data for regression,  $m$  steps ahead of predictions can be obtained. Cases with less than 6 data of observations will be omitted for practical reasons.

As for the convergence of diffusion models and ARIMA is nowadays an easy and well-known procedure, the opposite holds for ML techniques. In diffusion models, a fitting curve function is applied to produce the coefficients of the model. Bound restrictions have been set, in an effort to restrict outliers. In a similar manner, each data array will be explored in the case of ARIMA using 3 nested for loops, one for every parameter( $p,d,q$ ). For every combination, data are split into training and testing sets with a ratio of 2:1, and a root mean squared index will choose the best case in the testing set.



A simple LSTM configuration was selected, named Vanilla architecture, as is most suitable for predictions in short univariate time series. A typical Vanilla network consists of an input layer, a fully connected hidden layer, and a simple output layer. Each time the network gets an input sequence produces a single output estimation. In this case, after a long run of simulations, it is found that the data sequence should be split into 3 parts. So, a sequence of  $n$  data  $x_i \{x_1, x_2, \dots, x_n\}$  should be divided into 3 parts  $\{x_1, x_2, \dots, x_{n-3}\} \rightarrow \{x_{n-2}\}$ ,  $\{x_2, x_3, \dots, x_{n-2}\} \rightarrow \{x_{n-1}\}$ ,  $\{x_3, x_4, \dots, x_{n-1}\} \rightarrow \{x_n\}$  and these must be fed into the network for learning. After the completion of the learning procedure, a forecasting estimation can be obtained by applying the following sequence  $\{x_4, x_5, \dots, x_n\} \rightarrow \{x_{n+1}\}$ . Our network is constructed each time from  $n-3$  hidden unit cells and a single dense output cell. Drop out parameter found to not affect the efficiency of prediction, so its value was left simply at zero. The rest parameters epochs and batch play also an essential role in the fitting procedure. Epochs have been set to 200 but an early stopping mechanism is forced every time the network detects no further progress. As for the batch, which value has a close relationship with the number of a data sequence, also found from the simulation that should be  $n-3$ . This design is a simple construction, found to achieve impressive results and by no means is it the best and only case of architecture. Numerous parameterizations exist already in the literature and further investigations should shed light on the effectiveness of machine learning algorithms in time series forecasting.

As the SVR algorithm is a regression algorithm, the only independent variable that will influence our desired target is the sequence of years. There are many kernels of SVR models and the rationale here is to fit a Gaussian one. Except for  $C$  and epsilon parameters that have been already discussed earlier, the RBF kernel defines an extra control parameter gamma which is related to the standard deviation of Gaussian distribution. As there is no clear evidence of what these parameters should be, a random search will investigate the appropriate combination. A triple nested loop will be invoked once again. The sequence of data here is the desired output and the arithmetic sequence will feed the input to the model. Inconsistency with LSTM, the desired output sequence  $x_i \{x_1, x_2, \dots, x_n\}$  will fed with an input sequence  $\{1, 2, \dots, n\}$ . A testing split ratio of 20% will be also used.

In the literature different performance measures of forecasting accuracy are used: the best known and most employed is the Mean Absolute Percentage Error (MAPE). The performance analysis presented has been carried out on the basis of this criterion.

## 4 Results and discussions

Table 2 contains the outcomes of all diffusion models and ARIMA for one step ahead prediction. In the last line is highlighted the number of times a model failed to converged while in the last column the total number of cases the models run for each size of data sequence. Despite the fact that Logistic model ranked worst among all the diffusion models, it shows a strong convergence ability. It succeeded in all the cases to produce reliable coefficients without a violation in the predefined bounds. However, this unexpected worst case comes in contradiction with findings of Islam and Meade [11] who compared mobile diffusion in 25 countries, concluding logistic as most accurate method. One possible explanation for this finding perhaps is related to the fact that the average performance of logistic model affected by cases where the other models did not converge. As can be seen Local Gompertz and log-logistic are among models with best performance from diffusion category, while ARIMA model starts to show superiority when enough data are used. In figure 2, 5 models selected for comparison to highlight all the above findings, where it is obvious that ARIMA behaves better when 14 data points or more are used for prediction.

Data points	Logistic	Logistic 4	Logistic 5	Gompertz	Gompertz 4	Bass	Log-Logistic	Log-Logistic 4	ARIMA	Local Logistic	Local Gompertz	Local Bass	Cases
6	14.64	10.72	16.73	8.63	7.12	9.65	7.8	8.36	12.06	14.86	8.41	7.68	38
7	10.93	9.64	11.59	6.88	8.68	9.27	8.62	8.58	15.14	10.79	6.78	8.65	38
8	8.48	6.32	6.62	5.88	5.54	7.04	6.04	5.39	8.15	8.25	5.44	6.07	38
9	6.07	3.95	6.6	4.6	3.96	4.81	4.1	3.82	7.67	5.82	3.88	3.95	38
10	6.25	4.9	4.49	4.18	4.89	5.46	4.89	4.22	4.84	5.49	3.62	4.62	38
11	5.6	4.8	4.5	4.25	3.83	4.85	4.16	3.81	4.05	4.49	3.34	3.93	38
12	6.14	4.9	4.55	3.64	3.5	4.9	3.57	3.49	2.89	4.87	2.71	3.65	38
13	6.86	5.34	4.91	4.11	4.19	5.5	3.82	3.88	3.16	5.21	3.27	4.25	38
14	7.44	5.72	5.18	4.48	4.5	5.9	3.82	4.02	2.37	5.24	3.14	4.31	38
15	8.66	6.74	6.19	5.65	5.55	6.77	4.81	5	2.23	6.05	3.96	5.08	38
16	9.3	7.11	6.63	6.52	6.2	7.22	5.43	5.45	2.07	6.18	4.11	5.01	38
17	9.79	7.06	6.81	7.1	6.47	7.61	5.79	5.64	1.1	6.06	3.98	4.82	38
18	9.28	6.14	6.18	6.69	5.72	6.74	5.17	4.89	1.39	5.08	3.22	3.86	37
19	9.28	6.26	6.31	6.85	5.73	6.47	5.49	5.2	1.72	5.32	3.64	3.94	34
20	9.2	6.31	6.47	6.6	5.86	6.86	5.09	4.91	1.98	4.97	3.2	3.93	26
21	8.48	5.92	5.77	5.96	5.39	5.51	4.08	4.02	1.23	4.58	2.95	3.14	10
22	9.54	6.55	6.6	7.07	6.1	6.74	5.23	5.1	1.01	4.96	3.35	3.6	7
Converged	0	26	18	30	35	27	38	39	16				

Table 2. 1 step ahead prediction MAPE %

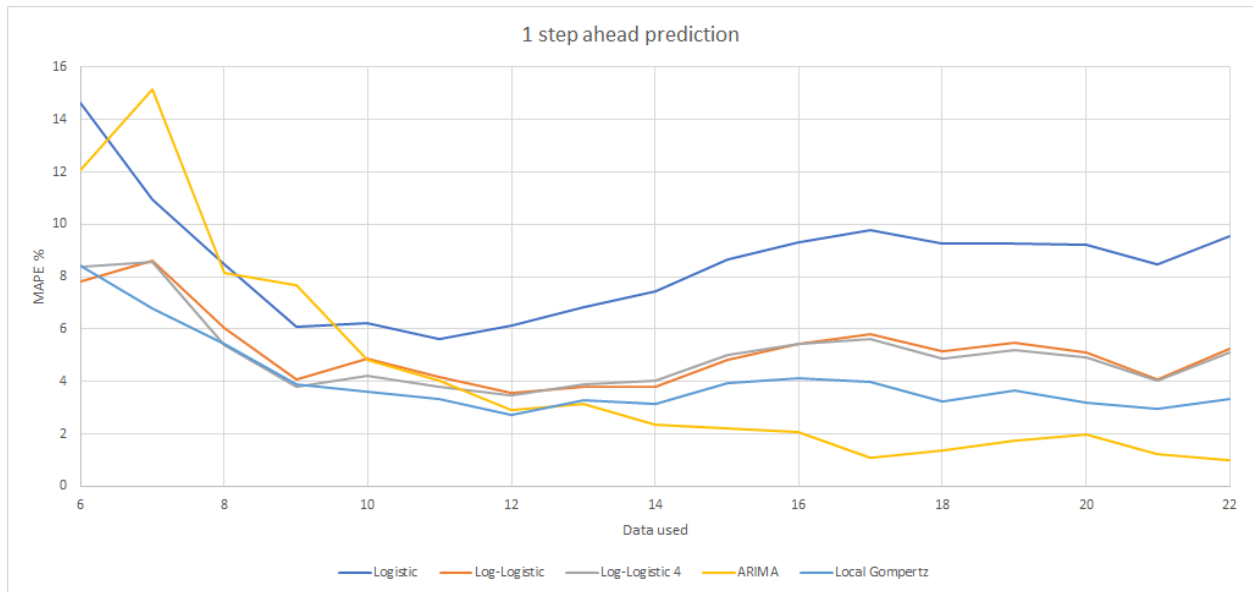


Figure 2

The same outcomes are presented in next table 3 and figure 3 for 5 steps ahead prediction. Full results outcomes are presented in Appendix. All previous outcomes are still valid except the fact that Local Gompertz now is less effective than log-logistic case. This is a more general outcome and for every diffusion model a local style transformation produces better estimations for short step ahead predictions while in turn as uncertainty increases with longer forecasts diminishes the effect (figure 4). One possible reason is that in cases where saturation level is underestimated, classic diffusion models keep repeating the same value. On the other hand, these local transformed diffusion models rely on last observation as reference so the effect of underestimation is reduced. Another conclusion that is safe to yield, is that diffusion models require at least 8 data points in order to foreseen reliable forecasts.

Data points	Logistic	Logistic 4	Logistic 5	Gompertz	Gompertz 4	Bass	Log-Logistic	Log-Logistic 4	ARIMA	Local Logistic	Local Gompertz	Local Bass
6	50.24	38.17	48.79	24.73	27.54	35.48	31.21	30.06	48.72	50.36	24.77	35.28
7	28.83	28.26	32.61	25.43	28.41	25.95	27.98	26.12	50.93	28.83	25.33	26.16
8	21.59	20.84	17.79	15.14	13.39	19.35	17.26	13.96	34.65	21.62	15.04	19.31
9	25.02	16.62	22.14	12.71	12.27	16.67	14.39	13.22	46.21	25.15	12.59	16.38
10	18.37	17.53	16.22	11.51	16.07	16.44	14.74	14.39	26.25	18.33	11.46	16.53
11	18.81	17.38	16.74	11.76	13.16	16.50	11.34	12.70	18.37	18.77	11.67	16.29
12	19.32	17.34	16.08	12.45	13.36	16.89	11.73	12.62	13.76	19.25	12.39	16.61
13	19.70	17.55	16.48	14.12	14.43	17.44	12.62	13.31	13.38	19.59	14.02	17.18
14	19.46	17.28	16.18	14.65	14.67	17.38	12.90	13.34	10.83	19.30	14.42	17.10
15	18.77	16.55	15.43	14.57	14.34	16.47	12.61	12.84	8.35	18.56	14.25	16.18
16	18.11	15.35	14.11	13.69	13.24	15.50	11.48	11.56	9.62	17.80	13.20	15.10
17	16.75	12.78	11.94	12.54	11.21	12.85	9.58	9.16	4.88	16.30	11.85	12.25
18	16.50	13.57	12.57	12.73	11.84	13.29	10.03	9.50	3.81	16.02	12.04	12.76

Table 3. 5 steps ahead prediction MAPE %

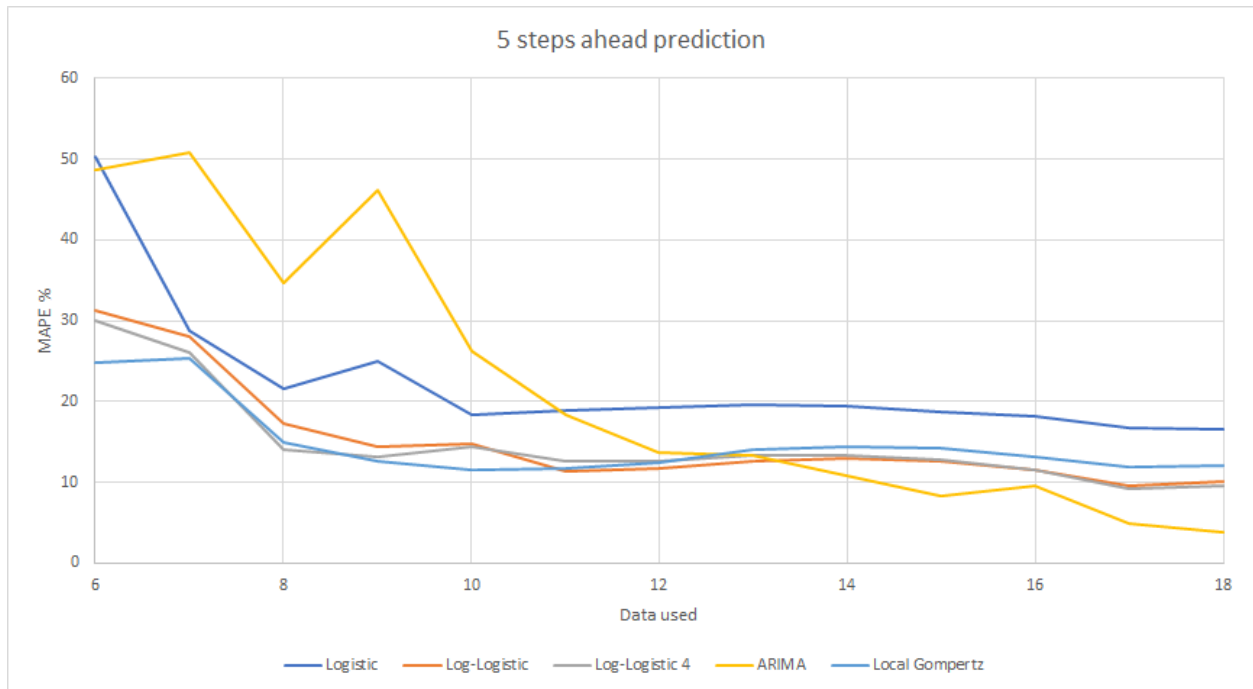


Figure 3

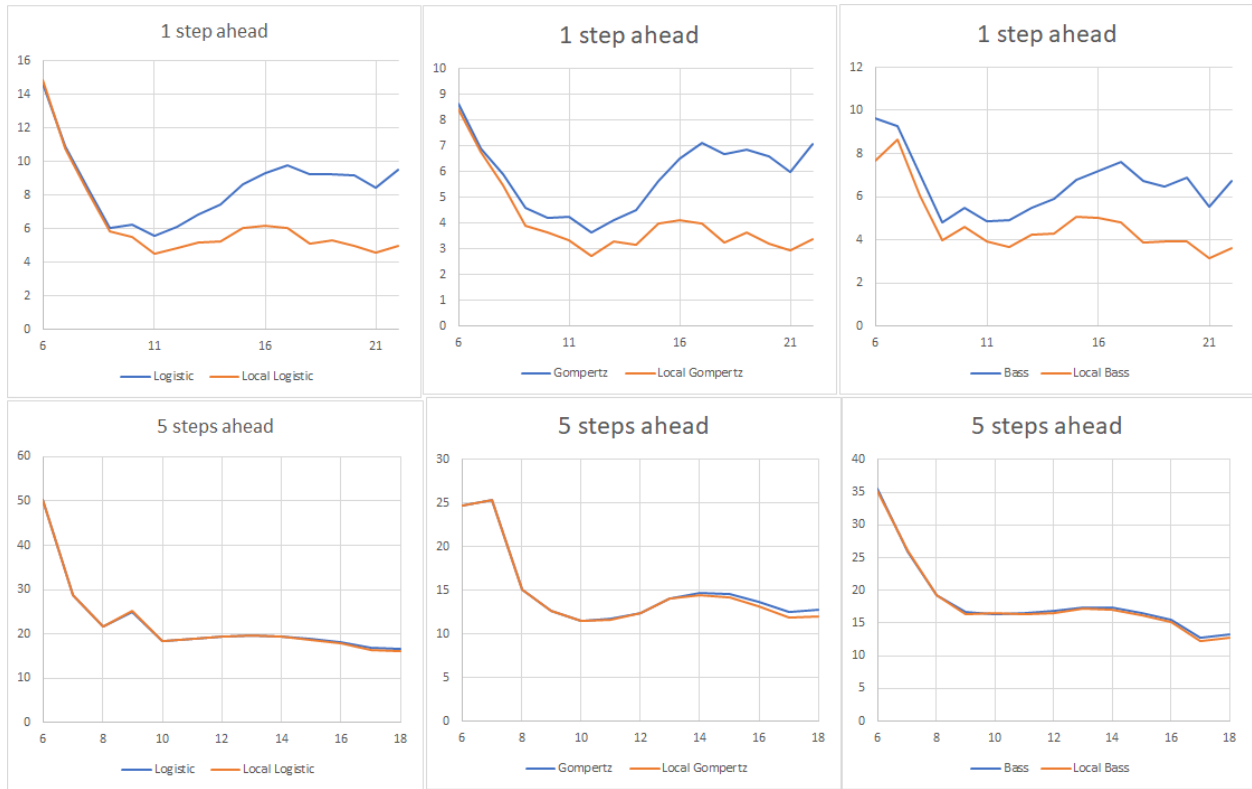


Figure 4

LSTM model and Vanilla implementation succeeded to achieve spectacular results despite the fact of its simple structure. As can be seen from the next table 4 and figure 5 is slightly better than ARIMA model in most cases. Each sequence of all cases trained independently and only the structure was common among them.

Data points	LSTM 1	LSTM 2	LSTM 3	LSTM 4	LSTM 5	ARIMA 1	ARIMA 2	ARIMA 3	ARIMA 4	ARIMA 5
6	12.69	18.68	25.48	24.24	26.84	12.06	18.36	29.27	38.54	48.72
7	11.27	19.90	27.63	26.95	25.81	15.14	23.32	32.51	41.22	50.93
8	8.63	14.11	20.29	20.84	22.59	8.15	13.82	20.19	27.27	34.65
9	4.72	9.66	15.17	16.55	15.53	7.67	15.98	26.69	36.40	46.21
10	4.34	7.06	9.13	10.60	11.54	4.84	9.74	14.64	20.54	26.25
11	2.59	3.88	5.65	7.53	9.46	4.05	7.70	10.70	14.50	18.37
12	2.05	3.33	5.66	6.90	8.54	2.89	5.06	7.87	10.85	13.76
13	2.44	3.40	4.87	7.82	9.60	3.16	5.35	7.71	10.51	13.38
14	1.99	3.53	5.82	7.33	8.07	2.37	5.12	7.10	9.24	10.83
15	2.25	4.10	5.62	6.65	7.25	2.23	4.22	6.01	7.49	8.35
16	1.71	3.11	4.22	5.29	5.45	2.07	4.11	6.00	8.17	9.62
17	1.38	2.43	3.94	4.33	4.56	1.10	2.03	3.56	4.09	4.88
18	1.38	2.90	3.40	2.08	2.55	1.39	3.13	4.04	3.61	3.81
19	1.92	2.70	1.54	1.48		1.72	2.48	2.96	2.54	
20	1.48	1.97	2.11			1.98	1.22	1.80		
21	1.28	1.85				1.23	1.73			
22	0.85					1.01				

Table 4 LSTM – ARIMA comparison MAPE %

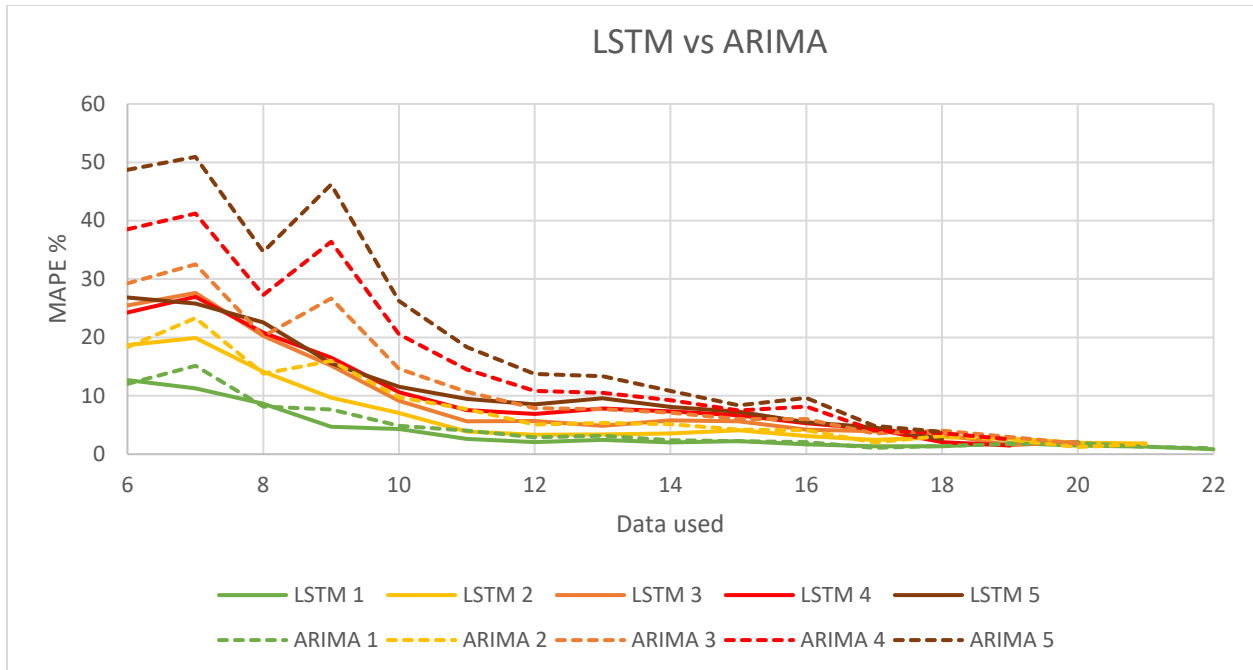


Figure 5. LSTM vs ARIMA 5 steps ahead

On the contrary, the SVR model failed to achieve impressive results as in LSTM case. However, SVR can offer comparable results if it is investigated which combination of coefficients can achieve the best performance. This approach is tricky and does not come in line with the methodology we have presented. Probable reasons have to do with the fact that the coefficients (C, gamma, epsilon) were not investigated in-depth, and presumably, there should be a combination to fit better. In next table 5 and figure 6 are presented the results of SVR with the tricky option, where the same combination of coefficients is applied to all countries.

Data points	SVR 1	SVR 2	SVR 3	SVR 4	SVR 5
6	33.72	45.29	51.08	54.05	56.35
7	24.94	33.07	37.32	40.69	43.21
8	16.66	22.37	26.76	29.89	32.68
9	11.15	15.69	19.40	22.62	25.33
10	7.41	11.05	14.34	17.19	19.80
11	5.26	7.84	9.94	12.69	16.20
12	4.29	5.73	7.94	11.31	14.38
13	3.66	4.95	7.90	11.28	14.26
14	3.21	5.29	8.40	11.57	13.77
15	3.90	6.04	8.87	11.10	12.67
16	4.02	6.39	8.69	10.29	11.77
17	3.72	5.82	7.88	9.55	9.95
18	3.50	5.50	6.91	7.43	8.99
19	3.38	4.73	5.55	7.12	
20	2.85	3.33	4.95		
21	1.98	3.21			
22	1.96				

Table 5. SVR prediction MAPE %

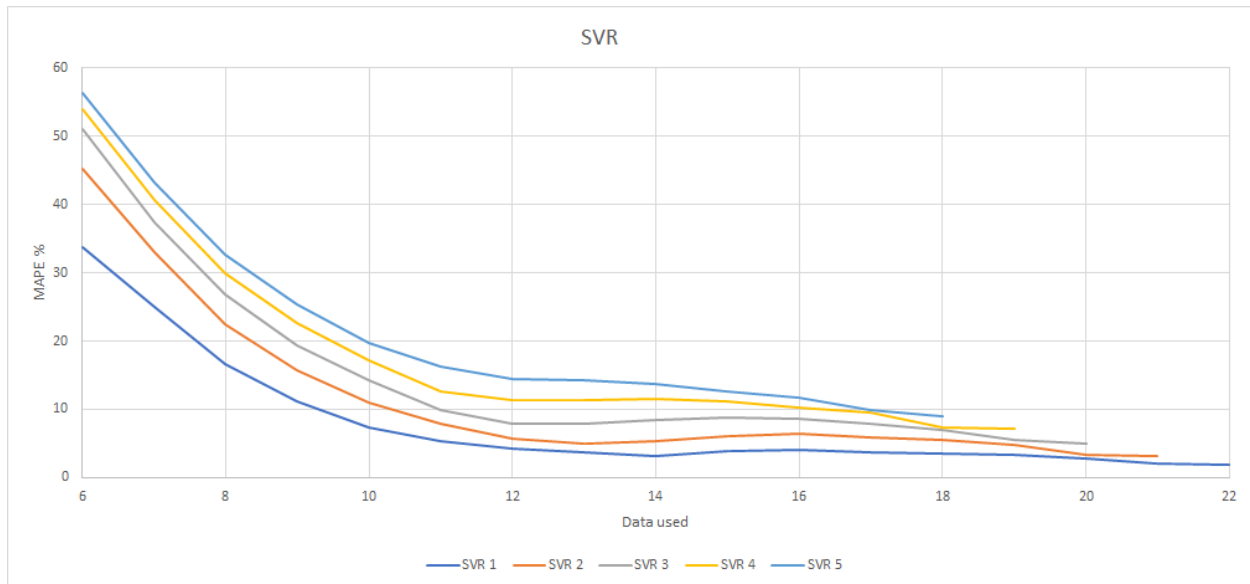


Figure 6

## 5 Conclusions

This paper presented a comparison between different methods for delivering short-term forecasts of broadband penetration in OECD countries. From the results, it is obvious that diffusion models perform better for short time series (<12 data), while ARIMA and LSTM do better for longer an outcome that comes in line with the previous study of Christodoulos, Michalakelis, and Varoutas [4]. Log-logistic and Gompertz were found to perform better among diffusion models, while LSTM stands out as a promising method for projecting the diffusion of high technology innovations in the community, as it slightly outperforms ARIMA. All the step ahead forecasts of each method were compared based on Mean Absolute Percentage Error (MAPE) as it is considered the main measure from similar studies.

Future research on this topic includes more complex structures in the architecture of LSTM networks and a further investigation of SVR failure. Other cases of high technology innovations in ICT sector should be also examined, like mobile broadband penetration. A recent study by Sima Siami-Namini, Neda Tavakoli, and Akbar Siami Namin [12], who made a comparison between ARIMA and LSTM in financial time series, confirms our findings, where the average reduction in error rates obtained by LSTM was between 84 - 87 percent when compared to ARIMA indicating the superiority of LSTM to ARIMA. However financial time series are out of scope from the diffusion of innovation, so we can not use them for reference. Another aspect of future investigation concerns whether reliable hybrid models can offer some significant improvement in forecasts.

### Acknowledgments.

The work presented in this article has been partially funded by the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens (N.K.U.A.), Athens, Greece.

## Bibliography

- [1] N. Meade and T. Islam, "Modelling and forecasting the diffusion of innovation - A 25-year review," *International Journal of Forecasting*, 2006, doi: 10.1016/j.ijforecast.2006.01.005.
- [2] N. Meade and T. Islam, "Forecasting with growth curves: An empirical comparison," *International Journal of Forecasting*, vol. 11, no. 2, pp. 199–215, Jun. 1995, doi: 10.1016/0169-2070(94)00556-R.
- [3] G. Gottardi and E. Scarso, "Diffusion models in forecasting: A comparison with the Box-Jenkins approach," *European Journal of Operational Research*, vol. 75, no. 3, pp. 600–616, Jun. 1994, doi: 10.1016/0377-2217(94)90300-X.
- [4] C. Christodoulos, C. Michalakelis, and D. Varoutas, "Forecasting with limited data: Combining ARIMA and diffusion models," *Technological Forecasting and Social Change*, vol. 77, no. 4, pp. 558–565, May 2010, doi: 10.1016/j.techfore.2010.01.009.
- [5] S. Panigrahi and H. S. Behera, "A hybrid ETS–ANN model for time series forecasting," *Engineering Applications of Artificial Intelligence*, vol. 66, pp. 49–59, Nov. 2017, doi: 10.1016/j.engappai.2017.07.007.
- [6] J. Baker, A. Pomykalski, K. Hanrahan, and G. Guadagni, "Application of machine learning methodologies to multiyear forecasts of video subscribers," in *2017 Systems and Information Engineering Design Symposium, SIEDS 2017*, May 2017, pp. 100–105. doi: 10.1109/SIEDS.2017.7937697.
- [7] J. C. Tanner, "Long-Term Forecasting of Vehicle Ownership and Road Traffic," *Journal of the Royal Statistical Society: Series A (General)*, vol. 141, no. 1, pp. 14–41, Jan. 1978, doi: 10.2307/2344775.
- [8] R. Bewley and D. G. Fiebig, "A flexible logistic growth model with applications in telecommunications," *International Journal of Forecasting*, vol. 4, no. 2, pp. 177–192, Jan. 1988, doi: 10.1016/0169-2070(88)90076-3.
- [9] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [10] "The Nature of Statistical Learning Theory - Vladimir Vapnik - Google Libri." [https://books.google.it/books?hl=it&lr=&id=sna9BaxVbj8C&oi=fnd&pg=PR7&ots=oqLbJYkpg9&sig=nYfR-\\_Z0VY-fHOLhiwFT2rFHVuc&redir\\_esc=y#v=onepage&q&f=false](https://books.google.it/books?hl=it&lr=&id=sna9BaxVbj8C&oi=fnd&pg=PR7&ots=oqLbJYkpg9&sig=nYfR-_Z0VY-fHOLhiwFT2rFHVuc&redir_esc=y#v=onepage&q&f=false) (accessed Jun. 03, 2022).
- [11] N. Meade and T. Islam, "Forecasting in telecommunications and ICT—A review," *International Journal of Forecasting*, vol. 31, no. 4, pp. 1105–1126, Oct. 2015, doi: 10.1016/J.IJFORECAST.2014.09.003.
- [12] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, pp. 1394–1401, Jan. 2019, doi: 10.1109/ICMLA.2018.00227.

## Appendix A

Data points	Logistic	Logistic 4	Logistic 5	Gompertz	Gompertz 4	Bass	Log-Logistic	Log-Logistic 4	ARIMA	Local Logistic	local Gompertz	local bass	LSTM	Cases
6	14.64	10.72	16.73	8.63	7.12	9.65	7.80	8.36	12.06	14.86	8.41	7.68	12.69	38
7	10.93	9.64	11.59	6.88	8.68	9.27	8.62	8.58	15.14	10.79	6.78	8.65	11.27	38
8	8.48	6.32	6.62	5.88	5.54	7.04	6.04	5.39	8.15	8.25	5.44	6.07	8.63	38
9	6.07	3.95	6.60	4.60	3.96	4.81	4.10	3.82	7.67	5.82	3.88	3.95	4.72	38
10	6.25	4.90	4.49	4.18	4.89	5.46	4.89	4.22	4.84	5.49	3.62	4.62	4.34	38
11	5.60	4.80	4.50	4.25	3.83	4.85	4.16	3.81	4.05	4.49	3.34	3.93	2.59	38
12	6.14	4.90	4.55	3.64	3.50	4.90	3.57	3.49	2.89	4.87	2.71	3.65	2.05	38
13	6.86	5.34	4.91	4.11	4.19	5.50	3.82	3.88	3.16	5.21	3.27	4.25	2.44	38
14	7.44	5.72	5.18	4.48	4.50	5.90	3.82	4.02	2.37	5.24	3.14	4.31	1.99	38
15	8.66	6.74	6.19	5.65	5.55	6.77	4.81	5.00	2.23	6.05	3.96	5.08	2.25	38
16	9.30	7.11	6.63	6.52	6.20	7.22	5.43	5.45	2.07	6.18	4.11	5.01	1.71	38
17	9.79	7.06	6.81	7.10	6.47	7.61	5.79	5.64	1.10	6.06	3.98	4.82	1.38	38
18	9.28	6.14	6.18	6.69	5.72	6.74	5.17	4.89	1.39	5.08	3.22	3.86	1.38	37
19	9.28	6.26	6.31	6.85	5.73	6.47	5.49	5.20	1.72	5.32	3.64	3.94	1.92	34
20	9.20	6.31	6.47	6.60	5.86	6.86	5.09	4.91	1.98	4.97	3.20	3.93	1.48	26
21	8.48	5.92	5.77	5.96	5.39	5.51	4.08	4.02	1.23	4.58	2.95	3.14	1.28	10
22	9.54	6.55	6.60	7.07	6.10	6.74	5.23	5.10	1.01	4.96	3.35	3.60	0.85	7
Converged	0	26	18	30	35	27	38	39	16					

Table A1. 1 step ahead prediction MAPE %

Data points	Logistic	Logistic 4	Logistic 5	Gompertz	Gompertz 4	Bass	Log-Logistic	Log-Logistic 4	ARIMA	Local Logistic	local Gompertz	local bass	LSTM	Cases
6	26.20	20.45	31.45	14.55	14.09	19.33	15.72	16.71	18.36	26.45	14.51	17.52	18.68	38
7	17.04	16.36	18.86	13.98	15.29	14.75	15.72	14.12	23.32	16.97	13.85	15.27	19.90	38
8	13.10	10.76	10.21	9.10	8.52	11.01	9.30	8.12	13.82	12.94	8.77	10.32	14.11	38
9	11.55	7.14	14.05	7.62	6.36	7.87	7.18	6.02	15.98	11.62	7.11	6.63	9.66	38
10	9.66	8.19	7.88	6.25	7.90	8.67	7.64	6.73	9.74	9.28	5.92	8.23	7.06	38
11	8.82	7.94	7.40	6.11	5.59	7.52	5.97	5.65	7.70	8.37	5.47	6.97	3.88	38
12	9.45	7.90	7.19	5.71	5.92	7.79	5.39	5.53	5.06	8.84	5.10	6.93	3.33	38
13	9.96	8.12	7.44	6.13	6.31	8.16	5.49	5.75	5.35	9.14	5.52	7.24	3.40	38
14	11.05	9.04	8.29	7.12	7.16	9.23	6.17	6.51	5.12	9.92	6.28	8.15	3.53	38
15	11.84	9.75	9.03	8.40	8.24	9.76	7.20	7.39	4.22	10.50	7.25	8.68	4.10	38
16	12.50	10.02	9.38	9.25	8.85	10.15	7.84	7.84	4.11	10.77	7.57	8.72	3.11	38
17	12.05	8.98	8.61	8.93	8.20	9.59	7.26	7.09	2.03	9.88	6.71	7.70	2.43	37
18	11.63	8.24	8.29	8.67	7.85	9.15	7.13	6.80	3.13	9.18	6.38	7.14	2.90	34
19	11.31	8.28	8.32	8.26	7.48	8.29	6.56	6.39	2.48	8.91	6.05	6.59	2.70	26
20	10.18	7.42	7.05	7.26	6.59	6.82	5.06	4.93	1.22	7.67	4.94	5.19	1.97	10
21	11.25	8.26	8.00	8.47	7.54	8.37	6.39	5.84	1.73	8.61	6.06	6.53	1.85	7
Converged	0	24	15	30	33	27	38	36	16					

Table A2. 2 steps ahead prediction MAPE %

Data points	Logistic	Logistic 4	Logistic 5	Gompertz	Gompertz 4	Bass	Log-Logistic	Log-Logistic 4	ARIMA	Local Logistic	local Gompertz	local bass	LSTM	Cases
6	36.89	29.02	42.67	20.20	20.87	27.46	22.73	24.10	29.27	37.11	20.18	26.20	25.48	38
7	21.91	21.19	23.49	19.05	21.09	19.05	20.70	19.21	32.51	21.88	18.96	18.93	27.63	38
8	16.91	14.88	13.25	11.71	10.30	14.77	11.93	9.82	20.19	16.79	11.46	14.07	20.29	38
9	16.35	10.40	17.35	9.34	8.21	10.72	9.65	8.56	26.69	16.51	9.09	10.06	15.17	38
10	12.63	11.12	11.00	8.16	10.61	11.66	10.00	9.01	14.64	12.46	7.96	11.27	9.13	38
11	12.10	11.05	10.40	8.39	8.35	10.66	8.04	8.14	10.70	11.92	8.07	10.38	5.65	38
12	12.65	10.86	9.92	7.63	8.09	10.56	7.11	7.50	7.87	12.36	7.35	9.97	5.66	38
13	13.54	11.49	10.56	8.46	8.81	11.48	7.58	8.16	7.71	13.10	8.14	10.88	4.87	38
14	14.22	12.05	11.20	9.90	9.91	12.21	8.64	8.99	7.10	13.63	9.34	11.57	5.82	38
15	15.02	12.77	11.90	11.20	11.00	12.79	9.72	9.91	6.01	14.30	10.43	12.12	5.62	38
16	14.78	12.17	11.37	11.18	10.75	12.28	9.48	9.49	6.00	13.80	10.01	11.36	4.22	37
17	14.21	11.18	10.59	10.85	10.11	11.80	9.05	8.86	3.56	12.99	9.35	10.78	3.94	34
18	13.50	10.42	9.73	9.98	9.25	11.04	8.14	7.85	4.04	12.13	8.50	9.77	3.40	26
19	12.15	9.11	8.50	8.80	7.95	8.42	6.39	6.07	2.96	10.65	7.15	7.31	1.54	10
20	12.91	9.91	9.42	9.81	8.89	9.92	7.48	6.90	1.80	11.30	8.10	8.62	2.11	7
Converged	0	22	13	30	32	27	38	35	16					

Table A3. 3 steps ahead prediction MAPE %



Data points	Logistic	Logistic 4	Logistic 5	Gompertz	Gompertz 4	Bass	Log-Logistic	Log-Logistic 4	ARIMA	Local Logistic	local Gompertz	local bass	LSTM	Cases
6	45.25	34.35	49.26	23.07	25.04	31.77	27.87	27.86	38.54	45.42	23.13	31.24	24.24	38
7	25.84	25.34	28.84	22.42	24.51	23.02	24.03	23.08	41.22	25.84	22.34	23.18	26.95	38
8	18.82	17.62	15.89	13.65	11.78	17.20	14.64	11.85	27.27	18.84	13.48	16.89	20.84	38
9	20.81	13.56	19.76	11.43	10.46	13.97	12.30	11.09	36.40	20.97	11.28	13.54	16.55	38
10	15.90	14.58	13.81	9.90	13.81	14.30	12.72	12.01	20.54	15.82	9.81	14.44	10.60	38
11	15.28	13.98	13.39	10.04	10.74	13.59	9.44	10.18	14.50	15.20	9.88	13.47	7.53	38
12	16.24	14.31	13.18	9.87	10.74	13.91	9.51	10.19	10.85	16.10	9.72	13.51	6.90	38
13	16.67	14.52	13.52	11.27	11.59	14.47	10.09	10.75	10.51	16.45	11.12	14.05	7.82	38
14	17.30	15.10	14.12	12.72	12.72	15.25	11.18	11.57	9.24	16.99	12.36	14.81	7.33	38
15	17.24	14.99	13.96	13.15	12.97	14.96	11.44	11.67	7.49	16.85	12.64	14.51	6.65	37
16	16.60	14.07	13.15	12.86	12.42	14.10	10.92	10.95	8.17	16.07	12.11	13.52	5.29	34
17	15.84	12.73	11.77	11.90	11.18	13.49	9.87	9.67	4.09	15.18	11.00	12.79	4.33	26
18	14.33	10.92	10.03	10.55	9.39	10.34	7.91	7.17	3.61	13.43	9.38	9.46	2.08	10
19	14.67	11.70	10.99	11.25	10.35	11.58	8.72	8.15	2.54	13.78	10.13	10.78	1.48	7
Converged	0	21	12	30	31	27	38	34	17					

Table A4. 4 steps ahead prediction MAPE %

Data points	Logistic	Logistic 4	Logistic 5	Gompertz	Gompertz 4	Bass	Log-Logistic	Log-Logistic 4	ARIMA	Local Logistic	local Gompertz	local bass	LSTM	Cases
6	50.24	38.17	48.79	24.73	27.54	35.48	31.21	30.06	48.72	50.36	24.77	35.28	26.84	38
7	28.83	28.26	32.61	25.43	28.41	25.95	27.98	26.12	50.93	28.83	25.33	26.16	25.81	38
8	21.59	20.84	17.79	15.14	13.39	19.35	17.26	13.96	34.65	21.62	15.04	19.31	22.59	38
9	25.02	16.62	22.14	12.71	12.27	16.67	14.39	13.22	46.21	25.15	12.59	16.38	15.53	38
10	18.37	17.53	16.22	11.51	16.07	16.44	14.74	14.39	26.25	18.33	11.46	16.53	11.54	38
11	18.81	17.38	16.74	11.76	13.16	16.50	11.34	12.70	18.37	18.77	11.67	16.29	9.46	38
12	19.32	17.34	16.08	12.45	13.36	16.89	11.73	12.62	13.76	19.25	12.39	16.61	8.54	38
13	19.70	17.55	16.48	14.12	14.43	17.44	12.62	13.31	13.38	19.59	14.02	17.18	9.60	38
14	19.46	17.28	16.18	14.65	14.67	17.38	12.90	13.34	10.83	19.30	14.42	17.10	8.07	37
15	18.77	16.55	15.43	14.57	14.34	16.47	12.61	12.84	8.35	18.56	14.25	16.18	7.25	34
16	18.11	15.35	14.11	13.69	13.24	15.50	11.48	11.56	9.62	17.80	13.20	15.10	5.45	26
17	16.75	12.78	11.94	12.54	11.21	12.85	9.58	9.16	4.88	16.30	11.85	12.25	4.56	10
18	16.50	13.57	12.57	12.73	11.84	13.29	10.03	9.50	3.81	16.02	12.04	12.76	2.55	7
Converged	0	20	11	30	30	27	38	34	17					

Table A5. 5 steps ahead prediction MAPE %