Time Series Forecasting for Healthcare Diagnosis and Prognostics with the Focus on Cardiovascular Diseases

Conference Paper *in* IFMBE proceedings · June 2018

DOI: 10.1007/978-981-10-4361-1_138

Anh Vo
New Jersey Institute of Technology

Trung Le
University of South Florida

A review on time series forecasting for healthcare diagnosis and prognostics with the focus on cardiovascular diseases

C. Bui[1], N. Pham[1], A. Vo[1], A. Tran[2], A. Nguyen[1] and T. Le[1]

[1] Biomedical Engineering, International University-Vietnam National University, Ho Chi Minh City, Vietnam
[2]Abbott Diagnostics Division-Abbott Laboratories S.A, Ho Chi Minh City, Vietnam

Abstract: **Time series forecasting has always been a prosperous field of science because it not only is widely applied in every aspect of life, but also poses a great challenge in method development due to the complexity of the underlying processes and the accuracy of assessing techniques. In medical applications, time series forecasting models have been successfully applied to predict the progress of the disease, estimate the mortality rate and assess the time dependent risk. However, the vast availability of many different techniques, in which each type excels in particular situations, makes the process of choosing an appropriate model more challenging. Therefore, the aim of this paper is to summarize and review different types of forecasting model that has been tremendously cultured for medical purposes using time series based forecasting methods. For each type of model, we will list the current related research papers, briefly describe the underlying theories, and discuss its advantages and disadvantages within different clinical situations. At the end, this paper also provides a robust and purpose-oriented classification of about 60 different forecasting models, therefore providing a comprehensive references for scientists and researchers to determine the suitable forecasting models that excel in their case of study.**

Keywords: **short-term, long-term, forecasting,**

I. Introduction

Time series forecasting can be defined as the estimation of future values of a time series whose methods of calculation are built on top of different mathematical and statistical models which may or may not rely on the past values and usually are accompanied with some assumptions about the system [1, 2]. In short, it is a method of transforming past values or measurements into the estimates of the future. There are hundreds of different types of forecasting models, each type of which relies on different methods, excels in different situations and has very different assumptions about the variation and evolution of the series over the time [3]. Assessing the outcome, time series forecasting has made its success within various fields of science, some of which include business operation [4], electrical engineering [5], power management [6], information systems [7] and medical diagnosis applications [8-11], which are the main focus of this review paper. Despite the fact that time series forecasting is a very broad field of science, it can be subdivided into two categories: short-term and long-term forecasting. Short-term forecasting performs/is used for intensive analysis and calculations of the underlying characteristics and variations of the time series to provide a robust and precise prediction of the future up to hours ahead of time [12]. In contrast, long-term prediction generally analyses the trend of the available data and the effect of the associated parameters to provide estimates for years in the future [13]. As the technique requires tremendous analysis and calculations, short-term forecasting techniques are not suitable for long-term prediction. Because of their differentiated abilities, their potentials can be applied in different clinical situations. Short-term forecasting, for example, is extremely useful in assessing patients' mortality in emergency care unit where immediate action is crucial, allowing doctors to make immediate response before a vital situation can take place [14]. Meanwhile, long-term forecasting thrives at assessing health condition for many years after hospital discharge taking into account the effects of different types of treatment and the associated risks, thus allowing doctors to provide suitable healthcare services for these patients [13]. In this paper, we will review all of the current methods using time series for short term and long term forecasting. The organization of the paper is summarized in Figure 1.

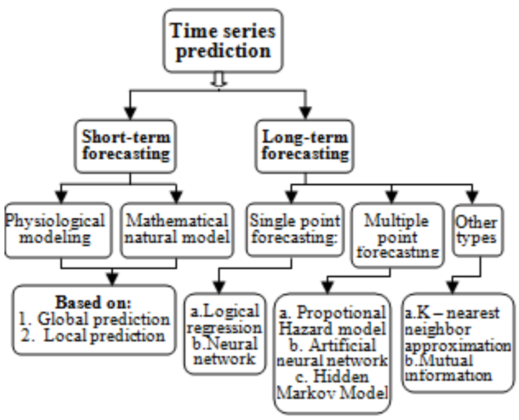

Fig. 1: Summary of the forecasting methods presented in the paper

II. Methodology

A. *Short-term forecasting techniques*

The short-term forecasting techniques can be further divided into 2 types of mathematical-based modeling and physiological-based modeling.

a)Mathematical natural model:

Mathematical-based forecasting model describes the variation of the signal collected from the system over time by using the mathematical representation and its dynamic variations to estimate the future states. In this paper, we categorize this forecasting model into three types of approach. The first group of approach involves forecasting the quantitative indices which have been constructed to characterize physiological or pathophysiological condition. The second group of approaches is related to the use of symbolic mismatch to forecast the system progress with the assumption of medical knowledge**.** The third group of forecasting approaches studies the association among state variables using system identification techniques.

The first approach forecasting the quantitative indices are constructed to classify broad physiological or pathophysiological conditions. For instance, these indices which are ratios of low to high frequency spectral power have been used to evaluate 'sympatho-vagal balance'[15] whereas Mayer wave frequency power has been used to classify congestive heart failure patients [16] or predict death after heart attacks [17] [18]. In terms of sympatho-vagal balance, it is generally accepted that the sinoatrial node or the natural pace-maker of the heart is innervated by the tenth cranical nerve or the vagus nerve of the parasympathetic nervous system and by the T1-4 spinal nerve of the sympathetic nervous system. Power spectral analysis of the heart period (R-R) reveals that while the power of the high frequency (HF) component of R-R activity corresponds to vagal activity or parasympathetic activity of the sinoatrial node, the power of low frequency (LF) of R-R activity is a marker of sympathetic activity of the sinoatrial node. It has been proved that the ratio of LF and HF components of R-R variability is an effective index to evaluate the state of the sympatho-vagal balance of the sinoatrial node in numerous physiopathological conditions. In the context of Mayer wave frequency power, each of four bands in the heart period power spectrum has one or two corresponding variables in the time domain. Time domain measurement of heart period variability, in particular ultra-low or low-frequency power, are strongly and independently associated with death during two weeks after acute myocardial infarction. Another example is the application of the slope of very low frequency blood pressure oscillations in predicting the cardiovascular mortality [19]. In details, traditional time and frequency domain heart rate variability indexes along with short-term fractal-like correlation properties of R-R intervals (exponent α) and power-law scaling (exponent β) were studied in 159 patients with depressed left ventricle function after an AMI. After follow-up years, the number of patients who died or were alive was counted. A general conclusion was that short-term scaling exponent α and power-law slope β differed between survivors and dead patients.

The second approach is symbolic mismatch that has been used to solve the problems related to patient-specific differences. Symbolic mismatch [20] is based on the principle of comparing long-term time-series with the predefined symbolic sequences. The differences between the morphology and the frequency of prototypical functional units with the predefined sequences have been used for the state prediction. There are three different methods using symbolic mismatch to figure out high risk patients in a population. They are one-class support vector machine, nearest neighbor analysis and hierarchical clustering. An application of symbolic mismatch approach on clinical cardiology is to identify patients who are at high risk of adverse cardiovascular events after an acute coronary syndrome (ACS). Quantifying variability in electrocardiogram (ECG) morphology has been proved to improve risk stratification post-ACS. A new metric has been developed to quantify beat-to-beat morphologic variability in beat space (MVB) [21]. Another application is to estimate the risk of death in patients after acute coronary syndrome. The biomarker was derived from time-series analyses of continuous electrocardiographic data collected from patients in the TIMI-DISPERSE2 clinical trials and then evaluated in a blinded, prespecified, and fully automated study on more than 4500 patients in the MERLIN-TIMI36. The results showed a strong association between the cardiac biomarker generated by symbolic mismatch and cardiovascular death in the MERLIN-TIMI36 trial over a 2-year period after acute coronary syndrome [22].

The third approach is statistical identification. It consists of Autoregressive-AR (p), Moving average-MA (q), Mixed-ARMA (p,q), Autoregressive integrated moving average-ARIMA (p,d,q), Autoregressive conditional heteroskedastic-ARCH (q), and Generalized autoregressive conditional heteroskedastic-GARCH (p,q). Parameter p is the number of autoregressive parameters, parameter q is the number of moving average parameters and d is the number of differentiations for the series to be stationary. In the context of Autoregressive (AR) models, they have the potential of quantifying essential delays and gains from blood pressure, cardiac interval, sympathetic activity and lung volume. Empirical observation suggests that cardiovascular oscillations are complex and mutable, resulting in two main characteristics of cardiovascular system- non-linear and non-stationary. The expansive versions of this model like PAR, ARMA, ARIMA or BL model become necessary to meet these requirements of the system. In terms of ARIMA model [23], it is a general statistical model which is widely used in the field of time series analysis. Alternatively, BL model [24] seems to be suitable for heart rate time series while requiring more extensive investigation. Non-parametric autoregressive time series [25] or threshold autoregressive [26] models can be used as exploratory tools for parametric models and wavelet non-stationary time series identification [27] and can sharply delineate temporal variations. However, careful statistical analysis models should be routine, whereas alternative models should be fitted to the same data set for comparison.

b) Physiological-based model:

It is noticeable that mathematical models for cardiovascular diseases are likely to obtain input data from Electrocardiography (ECG) and other biosignals such as blood pressure, pulse rate, and respiration. However, the usage of only this signal cannot reflect completely the cardiovascular system. It is due to the fact that the body contains organ systems which have effect on each other. Hence, it is necessary to develop appropriate models to solve this problem. This part shows a kind of that model— physiological-based model. There are three types of approach that would be discussed in this paper. We name the first one patient-specific time series modeling. The second one is state-space model, a hybrid of math natural model and physiological model. The last one is a novel terminology ─ mechanobiology.

The first approach is to apply time series models to characterize the physiological systems. For example, heart displacement caused by diaphragm motion is problematic in real-time (RT) registration during respiration. Chung et al. [28] developed a method to detect spatial alignment of RT 2D slices and a prior 4D volume using contour tracking during respiration. Cardiac respiratory motion exhibits different patterns in different conditions. Patient drifts and variations in both respiration period and amplitude cause non-stationary in respiratory motion. The contribution of this work is RT cardiac image registration during respiration via time series predictors [14]. Briefly, a periodic pattern of heart respiratory motion is measured first as a 4D time series, and then a set of predictors are calibrated to predict this pattern. The predictors are then employed to speed up RT registration with such accuracy and speed superior to the current techniques. Another example is the application of time series prediction in real time cardiac image registration during respiration [29]. In this case, time series predictors were used to speed up real time registration to predict the next position of the heart. Two kinds of this technique— Extended Kalman filter and linear adaptive filter— were investigated for respiratory motion prediction. Also, conditional distribution was applied to predict short cardiovascular variability signals [30]. In details, the predictor is defined as the median of the distribution conditioned by a sequence of L-1 previous sample. A function referred to the corrected mean squared predictor error is set to avoid the division of the whole set of data in learning and test sets. This function displays a minimum which is taken as a measure of the series predictability. This kind of modeling is applied to analyze the cardiovascular variability series of the heart period (RR interval) and the systolic arterial pressure (SAP).

The second method involves the forecasting of N-dimensional time series y_t using state space method. These time series are the observed data to an m-dimensional state vector α_t, and a Markovian transition equation that describes the evolution of the state space over time. The formulas of this kind of model contain correlation matrices which reflect the association between in-system variables. State-space model includes Markov chain (Random walker), Linear-Gaussian state-space models, Kalman filter and their extended versions (e.g. Extended Kalman Filter and Kalman smoother and so on. These models are hybrid forms and used for time series forecasting. Three components of state-space model ─ trend, seasonality, and noise - have been investigated to estimate in time series of coronary heart disease events [31]. The method is using a non-linear trend, allow multiple seasonal components, and carefully examining the residuals from the fitted model. Another method is to vary the seasonal effects over time and figure out how this helps to understand the association between coronary heart disease and varying temperature patterns. Furthermore, a hierarchical extension for the linear Gaussian state space model [32] was applied to study the individual differences in the dynamics of emotional systems. In particular, applying it to the Oregon adolescent interaction data resulted in interesting discussions about hypotheses on the links between cardiovascular processes, emotion dynamics and depression.

The last approach of physiological model is the mechanobiology modeling [33]. This kind of model has been applied to guide cell culture and test hypotheses related to the role of biomechanical factors in vascular diseases. Furthermore, it has enabled the application of cardiovascular mechanics which is outcome prediction of alternate therapeutic interventions for individual patients. A novel mathematical model of arterial chemo-mechano-biology has been developed to remodel and grow arteries. In details, the arterial wall is modeled as a bilayer cylindrical non-linear elastic membrane while fibroblast-mediated collagen growth is represented using a biochemical pathway model.

B. Long term forecasting methods

In some clinical situations, it is more favorable to assess patient health for a particular long period of time. While short term forecasting models excel at predicting the onset of catastrophic event, these models are not appropriate when dealing with long term forecasting. The reason is that short term forecasting models make heavy calculation to capture the dynamic detection of change points as they continuously alter over the time. Applying this practice, therefore, for long term purposes, results in a tremendous yet unnecessary computational cost that reduces the effectiveness of the model. For the above reasons, long term forecasting models are developed to give overall assessment about the disease condition, without relying heavily on capturing the hidden dynamics to give an appropriate prediction for long term purposes [13]. In this session, we are going to investigate the implementation of long term forecasting models for clinical assessment. These models can be classified into 2 categories:

a). Single point forecasting:

Single point forecasting models involve calculating the risk of an event or disease condition within a fixed time frame in the future. The result is usually given in the form a statistical value representing the degree of confidence whether or not the event will occur. Some remarkable single point forecasting models are Logistic Regression and Neural Network.

Logistic Regression

Logistic regression is a parametric model that calculates the probability for an event to happen based on the data that have been captured in the past known as experience [34]. By applying logistic calculation, this model can analyze the strength of the relationship between categorized dependent parameters and the desired variable. After performing this calculation, Logistic Regression model can determine the core parameters that greatly affect the output of the system, and in addition, provide a Logistic Function to calculate the output using these core parameters. When applied in clinical studies, Logistic Regression proves to be very powerful in predicting mortality after hospital discharge based on some crucial information regarding the patients' health [10].

One remarkable application of Logistic Regression in the field of medical diagnosis is the development of a simple risk score for assessing clinical severity of Acute Myocardial Infarction after Hospitalization [35], by Jacob, PhD and Henry, MD. This article strives to evaluate long term mortality risk for patients with acute myocardial infarction after hospital discharge within 6 years. The study found out strong correlation between mortality and clinical parameters including shock, heart failure, ECG finding, kidney function, and age. In general, patients who have their risk score greater than 16 points are 22 times more likely to die within the next 6 years than whose score ranges from 0 to 1. The result was compared with actual death certificate and the model proved to be very accurate.

Artificial Neural Network

Artificial Neural Network (ANN) is the general term for a group of Biological Neural Networks models that mimics human cognitive ability to make future predictions based on the past experience [36]. Similar to Logistic Regression, ANN models also learn how to calculate the provided inputs in order to give the final estimate. However, the main difference is that while Logistic Regression uses Logistic Calculation to analyze the strength of parameters' relationship, ANN treats each input parameter as an interconnected neuron that exchanges information with one another. Each neuron also contains an adaptive weight representing for its degree of importance on the final result [37]. Then by applying a predefined Training Function, ANN can predict the desired parameter based on all of the input parameters. This ability, therefore, allows ANN to base the prediction on very large amount of available input parameters which render other forecasting models poorly performed because of the tremendous calibration associated [38]. However, its strength also implies its weakness. In order to perform ANN with larger amount of input parameters, a computer system with strong processing power and large data storage must be used [38], thus making these models not available for small scale analysis.

In real - life application, ANN excels as a method for many scientific purposes including classification [34], pattern recognition [39], data processing and robotic control [40]. In the field of time series analysis for clinical diagnosis and prognosis, ANN thrives as a long term forecasting model that predicts accurately the outcome of diseases given with large amount of input parameters representing patient current condition. One noticeable study is the "Prediction of protein stability changes upon single-point mutations" [41], as described by Emidio, Pierro and Rita. This study involves creating an ANN system on top of a dataset of 1615 mutations documented with numerous input parameters and outcomes. At the end, this model was capable of analyzing the whole system and giving a prediction up to 90% in accuracy about the changes of protein stability. In another study, the authors Stephan and Lucia make "A comparison about the methodology and clinical application of Logistic Regression and Artificial Neural Network" [34]. The result is quite interesting, which states that ANN is the generalized version of Logistic Regression and both perform well in the field of Biomedical Diagnosis. However, one worth mentioning weakness of ANN over Logistic Regression is that, as described above, the former takes up much more computer resources to perform calibration than the latter.

Other types of forecasting models

Beside the two most commonly used long term forecasting models described above, the following methodologies are also very famous for their application in medical prognostics: k-nearest neighbor approximation (k-NN) and mutual information. Similarly, these predictive models also require good selection of inputs and give accurate estimates about the probability of an event to happen.

For a robust description, k-NN approximation method is a nonparametric model used for classification and regression. The method is powerful yet simple to apply. The principle of the technique is that similar inputs will create similar outputs and be utilized to provide estimate for time series. When applied in clinical situation for long term forecasting, k-NN analyzes the trend of the past data and creates a collection of dataset categorized with some similar properties [42]. This model then determines the most similar data point to the current data point and provides the next stage of this data point as the prediction. The advantage of this model is that k-NN approximation model is easy to apply and the associated hardware system does not need strong computational power. However, the drawback is that it cannot analyze multiple datasets at the same time [42].

In probability theory and information theory, mutual information measures the mutual dependence between two variables. As the name implies, this model allows prediction of a parameter given another parameter as input. Mutual Information model is suitable for small scale analysis and does not require strong computational skills to perform the forecasting. One drawback, however, is its inability to execute complex prediction usually encountered in real-life situations [43].

b) Multiple point forecasting models

In some clinical studies, in addition to determining the probability of an event to happen, researchers also wish to learn about how the disease progresses over the time or compare different types of clinical treatments to find out the best suited one for a particular patient. In these circumstances, Single point Forecasting model cannot be applied as it lacks the ability to give multiple outputs and therefore cannot capture the disease evolution or performing comparison. For the above requirements, a long term Multiple point Forecasting model will be performed.

Multiple point forecasting models are statistical methods that allow researchers to construct a mathematical function that predicts the system output with the varying time as one of the associated variables. In addition to the variable time, there are other parameters that represent both patient's current health condition and external effects [44]. This method is particularly useful when the interest of the study is to learn about disease progression within different groups of patients [10]. Some famous and mostly implemented models are Proportional Hazard Model and Artificial Neural Network (Multiple-point).

Proportional Hazard Model

Proportional Hazard Model, one part of the Survival Analysis, is the analysis of data from a time origin to an end time when an event happens. To be more precise, it involves calculating the probability of having an event as a function of time [45]. In this model, the first term is the probability distribution of an event without having any treatment applied. This value is predetermined by using simple statistical calculation about the past data that had been collected regarding this event. After that, the term is multiplied with an exponential value representing the effectiveness of a particular type of treatment. A treatment will comprise of two parts: a covariance vector and an effectiveness matrix. Covariance vector is an array of quantitative values ranging from -1 to 1 that manifests patient's personal and clinical characteristics. In addition, the effectiveness matrix contains a set of values, each of which corresponds to a value in the covariance vector and represents for its importance in affecting the final result. Because the model can make multiple and continuous outputs in the form of time series as well as represent both patient clinical characteristics and the effectiveness of the treatment applied, Proportional Hazard Model excels at comparing different types of treatment for a group of patients and analyzing disease evolution over the time [45]. The greatest advantage of using proportional hazard model is that it can handle censored data and capture disease evolution [46]. However, the major drawback of this technique is the complicated process to create one and the theory of proportional relationship must hold for the dataset [45].

For medical application, Proportional Hazard Model is developed with the aim to provide a method for researchers to learn about disease progression and compare the effectiveness of different types of treatment applied. In the article of "Predicting Survival In pulmonary arterial hypertension (PAH)" [47], a group of scientists and analysts have successfully developed a survival model that measures patient mortality each month after hospital discharge given an associated clinical treatment. This study involves analyzing data from 2716 patients with PAH enrolled consecutively and the final model developed is mortality assessment one year after hospital discharge. The result demonstrates that if all clinical activities listed in the paper are properly followed, one year survival could go up to 90% with 95% confidence interval. The authors also want to turn this model into a guideline that helps clinical centers give appropriate treatment for PAH patients.

Multiple point Neural Network

Multiple point Neural Network (MPNN) is a machine learning method belonging to the Neural Network family that is applied for long term prediction purposes. Similar to other Neural Network models, Multiple point Neural Network (NN) is also a parametric model that analyzes the past dataset to give prediction of the future event probability [48]. The main difference is that, however, the outputs are given as continuous values in the form of a time series. Elaborately, Multipoint Neural Network contains the input values known as neurons that exchange information to one another through an interconnected network mimicking human cognitive system. The biggest characteristics of Multiple Points Neural Network that differentiates it from other models belonging to the same family is that it contains the recurrence method. This methodology involves making the previously created output become the input for the next calibration, thus allowing MPNN to base prediction of the future event on the past prediction unlike other models of Neural Network family [37].

In real-life application, Multiple point Neural Network thrives at creating time series prediction for power management [49], pattern classification [39] and signal processing [50]. Although this model is not very well known for medical application, the same technique can be applied for creating a predictive medical prognostic model. For accuracy demonstration, in the paper "Multi-point tidal prediction using artificial neural network (ANN) with tide-generating forces" [51], the authors (Hsien-Kuo and Li-Ching) use ANN to create a model that simulates tides at multiple-point considering tide-generating forces function. The proposed model is then examined to estimate tides at some predefined single point and shows very good prediction accuracy. The authors also state that the extended application of this model could predict tides at multiple points in neighbor to the original point and the result is as accurate as the NAO.99b numerical model for tide prediction.

III. Discussion and conclusion

Short-term method can provide the forecast for a short period of time from some days to several weeks. It only requires few of input data. The outcome of long-term method, in contrast, is for months or years. This method requires many sources of information.

All in all, each type of forecast has its advantages and disadvantages. In terms of short-term prediction, it is good at forecasting quantity-related values and uses the correlation parameter to verify the result. Thanks to this method, responses to disease abnormalities can be made immediately. However, its drawbacks are the computationally intensive. On the other hand, long-term forecast is suitable for processing probability-related values. It uses ROC curve and R square to verify the result. This method helps to make decision and learn the dynamics of the diseases. In comparison with short-term forecast, long-term forecast is less flexible and requires less calculation.

In addition, for long-term prediction, risk factors are the most important elements. In details, these factors may come from life span, gender, habits like smoking, and so on. In long-term prediction, serious errors may occur corresponding to the big changes of risk factors in the system like the numbers of population or lifestyle. In contrast, short-term prediction shows few research articles on predicting heart disease. It may be due to the complication of the models and the complex methods for identifying the parameters. Moreover, for modeling physiological processes, researchers must have the deep understanding about the cardiovascular operations. In fact, it is not easy to handle both algorithm and physiology. Hence, in comparison with long-term forecast, current short-term methods do not have many applications on cardiovascular systems. However, according to the widespread of acute illness, short-term predictions are currently developed and promise to be useful tools for accurately predicting diseases. This information may help patients to access to the treatment in early stages and even allows intervention in preventing the disease' symptoms from occurring. A summary table of the methods, its limitations and applications is provided in the Appendix 1

The most important point of time series analysis is to select the appropriate model for a particular system. The selection of a method depends on many factors—the context of the forecast, the relevance and availability of historical data, the degree of accuracy desirable, the time period of forecasting, the cost or benefit of the forecast, and the time available for making the analysis. The main reason why time series methods applied to short-term forecast have high appropriation and accuracy is the past behavior of a particular variable is a good indicator of its future behavior, at least in the short-term. It is recognized that the most significant discrepancy between mathematical natural model and physiological model is that while the former can make the forecast based on parameters of the system, the latter can consider factors from other systems because human body comprises organ systems which act in synchronization.

References

1. Chatfield, C., *Time-series forecasting*. 2000: CRC Press.
2. Shumway, R.H. and D.S. Stoffer, *Time series analysis and its applications*. 2013: Springer Science & Business Media.
3. Hamilton, J.D., *Time series analysis*. Vol. 2. 1994: Princeton university press Princeton.
4. Hamilton, J.D., *A new approach to the economic analysis of nonstationary time series and the business cycle.* Econometrica: Journal of the Econometric Society, 1989: p. 357-384.
5. Park, D.C., et al., *Electric load forecasting using an artificial neural network.* Power Systems, IEEE Transactions on, 1991. **6**(2): p. 442-449.
6. Taylor, J.W., P.E. McSharry, and R. Buizza, *Wind power density forecasting using ensemble predictions and time series models.* Energy Conversion, IEEE Transactions on, 2009. **24**(3): p. 775-782.
7. Reis, B.Y. and K.D. Mandl, *Time series modeling for syndromic surveillance.* BMC Medical Informatics and Decision Making, 2003. **3**(1): p. 2.
8. Soni, J., et al., *Predictive data mining for medical diagnosis: An overview of heart disease prediction.* International Journal of Computer Applications, 2011. **17**(8): p. 43-48.
9. Getzen, T., *Forecasting health expenditures: short, medium and long (long) term.* Journal of Health Care Finance, 2000. **26**(3): p. 56-72.
10. Kirkwood, B.R., *Essentials of medical statistics*. 1988: Blackwell Scientific Publications.

11. Knaus, W.A., et al., *The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults.* Chest Journal, 1991. **100**(6): p. 1619-1636.
12. Rünstler, G., et al., *Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise.* Journal of forecasting, 2009. **28**(7): p. 595-611.
13. Armstrong, J.S., *Long-range forecasting*. 1985: Wiley New York ETC.
14. Cohen, M.A. and J.A. Taylor, *Short-term cardiovascular oscillations in man: measuring and modelling the physiologies.* The Journal of physiology, 2002. **542**(3): p. 669-683.
15. Malliani, A., M. Pagani, and F. Lombardi, *Physiology and clinical implications of variability of cardiovascular parameters with focus on heart rate and blood pressure.* The American journal of cardiology, 1994. **73**(10): p. C3-C9.
16. Teich, M.C., et al., *Heart rate variability: measures and models.* Nonlinear Biomedical Signal Processing: Dynamic Analysis and Modeling, Volume 2, 2000: p. 159-213.
17. Bigger, J.T., et al., *Frequency domain measures of heart period variability and mortality after myocardial infarction.* Circulation, 1992. **85**(1): p. 164-171.
18. Bigger, J.T., et al., *Correlations among time and frequency domain measures of heart period variability two weeks after acute myocardial infarction.* The American journal of cardiology, 1992. **69**(9): p. 891-898.
19. Mäkikallio, T.H., et al., *Fractal analysis of heart rate dynamics as a predictor of mortality in patients with depressed left ventricular function after acute myocardial infarction.* The American journal of cardiology, 1999. **83**(6): p. 836-839.
20. Syed, Z. and J.V. Guttag. *Identifying patients at risk of major adverse cardiovascular events using symbolic mismatch*. in *Advances in Neural Information Processing Systems*. 2010.
21. Liu, Y., et al., *ECG morphological variability in beat space for risk stratification after acute coronary syndrome.* Journal of the American Heart Association, 2014. **3**(3): p. e000981.
22. Syed, Z., et al., *Computationally generated cardiac biomarkers for risk stratification after acute coronary syndrome.* Science translational medicine, 2011. **3**(102): p. 102ra95-102ra95.
23. Box, G., G. Jenkins, and G. Reinsel, *Time series analysis: Forecasting and control. 3rd Prentice Hall.* Englewood Cliffs, NJ, 1994.
24. Christini, D.J., et al., *Application of linear and nonlinear time series modeling to heart rate dynamics analysis.* Biomedical Engineering, IEEE Transactions on, 1995. **42**(4): p. 411-415.
25. Fan, J. and I. Gijbels, *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*. Vol. 66. 1996: CRC Press.
26. Tong, H., *Non-linear time series: a dynamical system approach*. 1990.
27. Mallat, S., G. Papanicolaou, and Z. Zhang, *Adaptive covariance estimation of locally stationary processes.* Annals of Statistics, 1998: p. 1-47.
28. Chung, D., et al. *Real-time registration by tracking for MR-guided cardiac interventions*. in *Medical Imaging*. 2006. International Society for Optics and Photonics.
29. Esteghamatian, M., et al., *Real time cardiac image registration during respiration: a time series prediction approach.* Journal of real-time image processing, 2013. **8**(2): p. 179-191.
30. Porta, A., et al., *Prediction of short cardiovascular variability signals based on conditional distribution.* Biomedical Engineering, IEEE Transactions on, 2000. **47**(12): p. 1555-1564.
31. Barnett, A. and A. Dobson, *Estimating trends and seasonality in coronary heart disease.* Statistics in medicine, 2004. **23**(22): p. 3505-3523.
32. Lodewyckx, T., et al., *A hierarchical state space approach to affective dynamics.* Journal of mathematical psychology, 2011. **55**(1): p. 68-83.
33. Taylor, C.A. and C. Figueroa, *Patient-specific modeling of cardiovascular mechanics.* Annual review of biomedical engineering, 2009. **11**: p. 109.
34. Dreiseitl, S. and L. Ohno-Machado, *Logistic regression and artificial neural network classification models: a methodology review.* Journal of biomedical informatics, 2002. **35**(5): p. 352-359.
35. Jacobs, D.R., et al., *PREDICT: A Simple Risk Score for Clinical Severity and Long-Term Prognosis After Hospitalization for Acute Myocardial Infarction or Unstable Angina The Minnesota Heart Survey.* Circulation, 1999. **100**(6): p. 599-607.
36. Gurney, K., *An introduction to neural networks*. 1997: CRC press.
37. Hagan, M.T., et al., *Neural network design*. Vol. 20. 1996: PWS publishing company Boston.
38. Hassoun, M.H., *Fundamentals of artificial neural networks*. 1995: MIT press.
39. Pao, Y., *Adaptive pattern recognition and neural networks.* 1989.
40. Brooks, R.A., *A robot that walks; emergent behaviors from a carefully evolved network.* Neural computation, 1989. **1**(2): p. 253-262.
41. Capriotti, E., P. Fariselli, and R. Casadio, *A neural-network-based method for predicting protein stability changes upon single point mutations.* Bioinformatics, 2004. **20**(suppl 1): p. i63-i68.
42. Sorjamaa, A., et al., *Methodology for long-term prediction of time series.* Neurocomputing, 2007. **70**(16): p. 2861-2869.
43. Baldi, P., et al., *Assessing the accuracy of prediction algorithms for classification: an overview.* Bioinformatics, 2000. **16**(5): p. 412-424.
44. Harrell, F.E., K.L. Lee, and D.B. Mark, *Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.* Statistics in medicine, 1996. **15**: p. 361-387.
45. Sy, J.P. and J.M. Taylor, *Estimation in a Cox proportional hazards cure model.* Biometrics, 2000. **56**(1): p. 227-236.
46. Schoenfeld, D., *Partial residuals for the proportional hazards regression model.* Biometrika, 1982. **69**(1): p. 239-241.
47. Benza, R.L., et al., *Predicting survival in pulmonary arterial hypertension insights from the registry to evaluate early and long-term pulmonary arterial hypertension disease management (REVEAL).* Circulation, 2010. **122**(2): p. 164-172.
48. Rojas, R., *Neural networks: a systematic introduction*. 2013: Springer Science & Business Media.
49. Ruddin, S., E. Karatepe, and T. Hiyama, *Artificial neural network-polar coordinated fuzzy controller based maximum power point tracking control under partially shaded conditions.* Renewable Power Generation, IET, 2009. **3**(2): p. 239-253.
50. Miller, A. and B. Blott, *Review of neural network applications in medical imaging and signal processing.* Medical and Biological Engineering and Computing, 1992. **30**(5): p. 449-464.
51. Chang, H.-K. and L.-C. Lin, *Multi-point tidal prediction using artificial neural network with tide-generating forces.* Coastal Engineering, 2006. **53**(10): p. 857-864.
52. Semmlow, J.L., M. Akay, and W. Welkowitz, *Noninvasive detection of coronary artery disease using parametric spectral analysis methods.* Engineering in Medicine and Biology magazine, IEEE, 1990. **9**(1): p. 33-36.
53. Chia, T.L., P.-C. Chow, and H.J. Chizeck, *Recursive parameter identification of constrained systems: An application to electrically stimulated muscle.* Biomedical Engineering, IEEE Transactions on, 1991. **38**(5): p. 429-442.
54. Liu, Q., et al., *Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model.* BMC infectious diseases, 2011. **11**(1): p. 218.
55. Abdel-Aal, R. and A. Mangoud, *Modeling and forecasting monthly patient volume at a primary health care clinic using univariate time-series analysis.* Computer Methods and Programs in Biomedicine, 1998. **56**(3): p. 235-247.
56. Reaz, M., M. Hussain, and F. Mohd-Yasin, *Techniques of EMG signal analysis: detection, processing, classification and applications.* Biological procedures online, 2006. **8**(1): p. 11-35.
57. Semmlow, J. and K. Rahalkar, *Acoustic detection of coronary artery disease.* Annu. Rev. Biomed. Eng., 2007. **9**: p. 449-469.
58. Arnsparger, J.M., et al., *Adaptive control of blood pressure.* Biomedical Engineering, IEEE Transactions on, 1983(3): p. 168-176.
59. Van Vliet, R.C., *Predictability of individual health care expenditures.* Journal of Risk and Insurance, 1992: p. 443-461.
60. Ge, D., N. Srinivasan, and S.M. Krishnan, *Cardiac arrhythmia classification using autoregressive modeling.* Biomedical engineering online, 2002. **1**(1): p. 5.

61. Paiss, O. and G.F. Inbar, *Autoregressive modeling of surface EMG and its spectrum with application to fatigue.* Biomedical Engineering, IEEE Transactions on, 1987(10): p. 761-770.
62. Anderson, C.W., E.A. Stolz, and S. Shamsunder, *Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks.* Biomedical Engineering, IEEE Transactions on, 1998. **45**(3): p. 277-286.
63. Kelwade, J. and S. Salankar, *Prediction of Cardiac Arrhythmia using Artificial Neural Network.* International Journal of Computer Applications, 2015. **115**(20).
64. Baxt, W.G., *Use of an artificial neural network for data analysis in clinical decision-making: the diagnosis of acute coronary occlusion.* Neural computation, 1990. **2**(4): p. 480-489.
65. Segovia, F., et al., *Early diagnosis of Alzheimer's disease based on partial least squares and support vector machine.* Expert Systems with Applications, 2013. **40**(2): p. 677-683.
66. Kerhet, A., et al., *A SVM-based approach to microwave breast cancer detection.* Engineering Applications of Artificial Intelligence, 2006. **19**(7): p. 807-818.
67. Tapak, L., et al., *Real-data comparison of data mining methods in prediction of diabetes in Iran.* Healthcare informatics research, 2013. **19**(3): p. 177-185.
68. Yau, C., et al., *Bayesian non‐parametric hidden Markov models with applications in genomics.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2011. **73**(1): p. 37-57.
69. Karplus, K., C. Barrett, and R. Hughey, *Hidden Markov models for detecting remote protein homologies.* Bioinformatics, 1998. **14**(10): p. 846-856.
70. Uğuz, H., A. Arslan, and İ. Türkoğlu, *A biomedical system based on hidden Markov model for diagnosis of the heart valve diseases.* Pattern Recognition Letters, 2007. **28**(4): p. 395-404.
71. Coast, D.A., et al., *An approach to cardiac arrhythmia analysis using hidden Markov models.* Biomedical Engineering, IEEE Transactions on, 1990. **37**(9): p. 826-836.
72. Andreão, R.V., B. Dorizzi, and J. Boudy, *ECG signal analysis through hidden Markov models.* Biomedical Engineering, IEEE Transactions on, 2006. **53**(8): p. 1541-1549.
73. Tarvainen, M.P., et al., *Time-varying analysis of heart rate variability signals with a Kalman smoother algorithm.* Physiological measurement, 2006. **27**(3): p. 225.
74. Oikonomou, V.P., et al., *The Use of Kalman Filter in Biomedical Signal Processing*. 2009: INTECH Open Access Publisher.
75. Wu, W., et al., *Modeling and decoding motor cortical activity using a switching Kalman filter.* Biomedical Engineering, IEEE Transactions on, 2004. **51**(6): p. 933-942.
76. Ting, C.-M., et al., *Spectral estimation of nonstationary EEG using particle filtering with application to event-related desynchronization (ERD).* Biomedical Engineering, IEEE Transactions on, 2011. **58**(2): p. 321-331.
77. Lee, J. and K.H. Chon, *Time-varying autoregressive model-based multiple modes particle filtering algorithm for respiratory rate extraction from pulse oximeter.* Biomedical Engineering, IEEE Transactions on, 2011. **58**(3): p. 790-794.
78. Dunson, D.B., *Nonparametric Bayes applications to biostatistics.* Bayesian nonparametrics, 2010. **28**: p. 223.
79. Wakefield, J., *The Bayesian analysis of population pharmacokinetic models.* Journal of the American Statistical Association, 1996. **91**(433): p. 62-75.
80. Durichen, R., et al. *Multi-task Gaussian process models for biomedical applications*. in *Biomedical and Health Informatics (BHI), 2014 IEEE-EMBS International Conference on*. 2014. IEEE.
81. Blanco-Velasco, M., B. Weng, and K.E. Barner, *ECG signal denoising and baseline wander correction based on the empirical mode decomposition.* Computers in biology and medicine, 2008. **38**(1): p. 1-13.
82. Echeverria, J., et al., *Application of empirical mode decomposition to heart rate variability analysis.* Medical and Biological Engineering and Computing, 2001. **39**(4): p. 471-479.

Appendix 1: Summary of current time series forecasting methods, application fileds, and limitations.

Type	Describe	Application	Limitation
Parametric model	A function describes the relationship between inputs and outputs.	Predict and classifies disease [52]	False positives due to miss-specified model
ARIMA models.	Assumes a linear relationship between the lagged variables	Parameter identification [53], Forecasting [54], Time series modeling [7, 55]	Limited to capturing the first-order non-stationarity
Reviewed AR and ARMA	A coarse approximation to real world complex systems ARMA	Signal processing, classification [56, 57], adaptive control [58]	Fail to accurately predict the evolution of nonlinear and non-stationary processes.
AR, ARMA, PAR, BL Models	AR is a model whose input variable switches smoothly between two regimes. ARMA is the linear extension and PAR is the nonlinear extension of the AR model. PL is the nonlinear extension of the ARMA	Signal analysis and prediction [59], arrhythmia classification [60], signal modeling [61], and signal classification [62]	Parameters have to be optimized to build a consistent time-varying STAR model, and no efficient analytical optimization is available. PL model is a bilinear structure, but it has the ability to adequately model nonlinear systems
Neural Networks	Approximate any continuous function to an arbitrary precision	Health informatics [63], clinical decision-making [64]	One major disadvantage is that there is no formal systematic model building approach
SVM-based forecasting methods	Use a class of generalized regression models, such as Support Vector Regression.	Early diagnosis of Alzheimer disease [65], Breast cancer detection [66]. Prediction of Diabetes [67]	A common disadvantage of non-parametric techniques such as SVMs is the lack of transparency of results
Hidden Markov Models	A class of models, where the observed time series is treated as a function of the underlying, unobserved states vector.	Applications in genomics [68, 69], Disease diagnosis [70, 71] and signal analysis [72]	Sensitive to the order of the Markovian employed to represent the state. The parametric form can be unwieldy
State space model	Sate space models such as Kalman Filter (KF) and Particle Filter (PF; Arulampalam	Signal analysis and processing [73, 74], Modeling and decoding [75]	Provides accurate results only for Gaussian and linear models. For non-Gaussian and non-linear models, particle filtering (PF) is the most appropriate approach
PF (Particle Filter) models	Allow structured approximation via Bayesian estimation, most effective for nonlinear time series forecasting	Spectral estimation of non-stationary EEG using PF [76] Respiratory rate extraction from pulse oximeter [77]	Computational requirements much higher than of the Kalman filters. Problems with nearly noise-free models, especially with Accurate dynamic models. Very hard to find programming errors
Bayesian model	A process of incorporating prior information to render posterior inference; estimating the conditional distribution p(θ|y) of the hidden model	Applications in biostatistics [78], genomics [68], Analysis of population [79]	Bayesian nonparametric inference has witnessed considerable advances. However, these advances have not received a full critical and comparative analysis of their scope, impact and limitations in statistical modeling
GP model (MPG model)	A GP model seeks to establish a mapping f of the form y = f (x) + ε, where x can be constructed from historical realizations	Multi-task Gaussian process models for biomedical applications [80]	Require computational expense to perform. Compared with classic GP models, LGP and DPMG models can be applied to non-stationary time series
Empirical Mode Decompositi on (EMD)	Decompose non-stationary time series into a finite number of components called Intrinsic Mode Functions (IMFs	ECG signal de-noise and baseline wander correction [81], Heart rate variability analysis [82]	Although attractive for non-linear and non-stationary forecasting, EMD poses some mathematical challenges due to the end effects.

Author: Trung Le
Previous
Institute: International University- Vietnam National University
Address: Quarter 6, Linh Trung, Thu Duc District
City: Ho Chi Minh
Country: Vietnam
Email: trung.le@hcmiu.edu.vn/ trung.q.le@ndsu.edu