Current Rheumatology Reports (2020) 22:27
https://doi.org/10.1007/s11926-020-00907-6

RHEUMATOID ARTHRITIS (L MORELAND, SECTION EDITOR)

Viewpoint on Time Series and Interrupted Time Series Optimum Modeling for Predicting Arthritic Disease Outcomes

Hossein Bonakdari[1,2] · Jean-Pierre Pelletier[1] · Johanne Martel-Pelletier[1]

Abstract

Purpose of Review The propose of this viewpoint is to improve or facilitate the clinical decision-making in the management/treatment strategies of arthritis patients through knowing, understanding, and having access to an interactive process allowing assessment of the patient disease outcome in the future.

Recent Findings In recent years, the time series (TS) concept has become the center of attention as a predictive model for making forecast of unseen data values. TS and one of its technologies, the interrupted TS (ITS) analysis (TS with one or more interventions), predict the next period(s) value(s) of a given patient based on their past and current information. Traditional TS/ITS methods involve segmented regression-based technologies (linear and nonlinear), while stochastic (linear modeling) and artificial intelligence approaches, including machine learning (complex nonlinear relationships between variables), are also used; however, each have limitations.

Summary We will briefly describe TS/ITS, provide examples of their application in arthritic diseases; describe their methods, challenges, and limitations; and propose a combined (stochastic and artificial intelligence) procedure in post-intervention that will optimize ITS modeling. This combined method will increase the accuracy of ITS modeling by profiting from the advantages of both stochastic and nonlinear models to capture all ITS deterministic and stochastic components. In addition, this combined method will allow ITS outcomes to be predicted as continuous variables without having to consider the time lag produced between the pre- and post-intervention periods, thus minimizing the prediction error not only for the given data but also for all possible future patterns in ITS. The use of reliable prediction methodologies for arthritis patients will permit treatment of not only the disease, but also the patient with the disease, ensuring the best outcome prediction for the patient.

Keywords Data-driven · Time series · Interrupted time series · Arthritis · Clinical decision-making · Management/treatment strategies

Introduction

Precision medicine for arthritis patients is an emerging approach that tailors decisions/treatments to individual patient variability. This entails better diagnoses, earlier interventions, more efficient therapies and customized treatment plans. A long-standing dilemma of arthritis has been the challenge of identifying patients who will respond to a therapeutic approach, whereby treatments are often given on the basis of trial and error. In short, prediction is one of the most important and difficult tasks for various pathologies, including the treatment and management of arthritic diseases. The heterogeneous nature of many arthritic diseases and pharmacokinetic drivers of drug response are only two variables affecting a patient's "poor" outcome in response to a given treatment. Another disadvantage in clinical practice is the significant

This article is part of the Topical Collection on *Rheumatoid Arthritis*

✉ Johanne Martel-Pelletier
jm@martelpelletier.ca

Hossein Bonakdari
hossein.bonakdari@fsaa.ulaval.ca

Jean-Pierre Pelletier
dr@jppelletier.ca

1 Osteoarthritis Research Unit, University of Montreal Hospital Research Centre (CRCHUM), 900 Saint-Denis, R11.412, Montreal, QC H2X 0A9, Canada

2 Department of Soil and Agri-Food Engineering, Laval University, 2425 rue de l'Agriculture, Québec, QC G1V 0A6, Canada

Published online: 20 May 2020

delay or "lag time" between the onset of the intervention and the impact of the intervention. Determining this unknown delay in response to the implemented treatment or better predicting the result for a given patient is of great importance, as it is well known that early and accurate intervention can lead to improved long-term patient outcomes.

The use of modeling tools is indispensable to help accurately distinguish patients who will benefit from an intervention. Existing epidemiological research strategies such as cohort and case-control research studies illustrate essential disease etiology knowledge, but because of limitations such as confounding factors based on group differences, they are not as useful as intervention studies. Randomized controlled trials are still considered the benchmark design for determining whether an intervention is effective, yet such studies are not always a viable option. Moreover, there is often a need to evaluate interventions that have already been put in place, either without randomization or without any control. Due to their immense potential as a predictive model, time series (TS) and one of its methodologies, interrupted time series (ITS) (TS with one or more interventions), provide exceptional vehicles to achieve such a goal. As examples for health care quality improvements, Table 1 summarizes some studies that analyze ITS modeling to investigate the impacts of the implemented intervention. In this viewpoint, we will briefly describe TS and ITS; provide examples of the potential of TS applications in arthritic diseases; describe TS and ITS methods, their challenges and limitations; and propose a novel ITS procedure in post-intervention that optimizes ITS modeling.

Description of Time Series and Interrupted Time Series

TS and ITS analysis uses a patient's historical data and current information to anticipate that patient's outcome, predicting the patient's future by understanding the past. These methodologies allow each sampling unit to serve as its own control without stripping contextual and temporal factors from the analysis. Moreover, such analysis circumvents the need to specify a causal relationship and to predict the values of exploratory variables. In short, TS is a series of data points in time order [22]; it is a sequence taken at successive equally spaced points in time and their associations with other trends or events, taking account of the temporal structure of such data. A TS could be affected by four main components (trend, seasonality, cyclic and residual), which can be separated from the observed data.

ITS is a quantitative, statistical method in which multiple, repeated observations are made at regular intervals before and after an intervention. In an ITS study, a time series of an outcome is employed to analyze an underlying trend, which is interrupted by the implementation of an intervention at a known point of time. Importantly, in ITS, the intervention(s) at a change point(s) splits the outcome into two or more separate pre- and post-intervention periods. Several studies investigated the impacts of implemented intervention on arthritic diseases and healthcare quality improvements, and examples are shown in Table 1 [1–15, 16•, 17–21]. This table illustrates that each intervention has possible impacts by changing the level and/or trend without and/or after a time lag on the post-intervention period. It is used to assess the impact of an implemented intervention on an outcome immediately and over time, instantly or with delay, transiently or long-term, and whether elements other than the intervention could account for the change.

Applications of Time Series/Interrupted Time Series in Arthritic Diseases

The following are examples of the potential of these methodologies' applications in arthritic diseases and, although the possibilities are limitless, also included are some potential future studies using TS/ITS modeling.

TS analysis was employed to propose a system for monitoring rheumatoid arthritis (RA) based on the TS concept to optimize the number of face-to-face patient consultations [23]; predict patient use of outpatient hospital resources after a mandatory switch from originators to biosimilars in inflammatory arthritis healthcare use and costs or disease activity [24, 25]; evaluate in veterans with knee or hip osteoarthritis, national trends in non-opioid analgesic prescribing following guidance limiting or decreasing use of opioid therapy [21]; and to assess the health authority recommendation to the use of knee arthroscopy in patients with knee osteoarthritis [17].

For its part, ITS was utilized to assess whether treatment guidelines change the management of early RA patients [6]; investigate the impact of the electronic reminders with linked order sets, physician auditing and feedback, patient outreach, and optional printed prescriptions on clinic-level changes in influenza vaccination, pneumococcal vaccination and herpes zoster vaccine rates in patients with RA [7]; study the impacts of the introduction of infliximab and etanercept biosimilars on the utilization of biological disease-modifying antirheumatic drugs (bDMARDs) and the National Health Service budget impact in the UK [9]; investigate impacts of the UK National Institute for Health and Care Excellence (NICE) approval of tumor necrosis factor inhibitor therapies on the incidence of total hip and knee replacement in RA patients [14]; evaluate the effect of health information on the performance of RA disease activity measures and outcomes [16•]; and investigate the potential for individualized and personalized rheumatic medicine by introducing a dose-reduction intervention of bDMARDs and to forecast subjects at risk of arthritis-related pain in subsequent years [20]. Following the introduction of new disease-modifying osteoarthritis drugs (DMOADs) or new (DMARD/bDMARD), ITS analysis may enable the

Table 1 Interrupted time series modeling that investigates the impact of the implemented intervention on health care quality improvements including arthritis

Study	Intervention	Impacts on post-intervention period			Design strategy to analyze the interrupted time series outcome
		Trend	Level	Time lag	
Gasparrini et al. [1]	The smoking bans in indoor public places, mainly in public venues and workplaces in the Tuscany region in Central Italy	Yes	No	No	By assuming Poisson distribution of the outcome, a nonlinear model was adjusted
Desai et al. [2]	A simple point-of-care paper reminder form as a quality improvement (QI) strategy on the rate of being up-to-date with pneumococcal vaccination in a rheumatology practice	Yes	Yes	No	Linear segmented regression was employed
Robinson et al. [3]	The government of New Zealand implemented an intervention for the hospital discharge program with nurse telephone follow-up and referral among hospitals in Auckland to improve care and reduce hospital readmissions	No	No	No	Three linear, quadratic, and cubic models segmented regression models were verified, none of which demonstrated a significant discontinuity
Linden [4]	A large-scale tobacco control program launched in California in 1988 by raising state excise tax on cigarettes	Yes	Yes	No	Linear segmented regression was employed
Wang and Bhattacharyya [5]	The prevalence of prescriptions for non-steroidal anti-inflammatory drugs (NSAIDs) in the USA between 1996 and 2009	Yes	Yes	No	Using autoregressive error a model with up to fifth-order autocorrelations was adjusted
Judge et al. [6]	The British Society for Rheumatology and British Health Professionals in Rheumatology Guideline for the Management of Rheumatoid Arthritis in 2006	Yes	Yes	No	Linear segmented regression was employed
Baker et al. [7]	Electronic reminders with linked order sets, physician auditing and feedback, patient outreach, and optional printed prescriptions for zoster vaccination at an outside pharmacy to improve influenza, pneumococcal, and herpes zoster vaccination among patients with rheumatoid arthritis	Yes	No	No	Using autoregressive errors a linear regression model was fitted
Parkinson et al. [8]	The self-reported incident osteoarthritis on outlines of health services use in a representative group of community living 'baby boomer' Australian women using person-level health survey data linked with administrative data.	Yes	Yes	No	Linear segmented regression was employed
Aladul et al. [9]	The infliximab and etanercept biosimilars modifying Biological disease-modifying antirheumatic drugs (bDMARDs) utilization and National Health Service (NHS) budget in the UK	Yes	Yes	No	Linear segmented regression was employed
Bernal et al. [10]	The Italian smoking ban in public areas on hospital admissions for acute coronary events	Yes	Yes	No	A non-linear segmented regression model based on time-varying confounders was designed
Lin et al. [11]	The Opioid Safety Initiative (OSI) was implemented by The Veterans Health Administration leadership to promote safer opioid-related prescribing practices associated with adverse outcomes in US	Yes	No	No	Linear segmented regression was employed
Zombré et al. [12]	The user fee exemption on healthcare utilization for children under-five intervention was delivered within the context of population health by the government of Burkina Faso	Yes	Yes	No	By comparison between different segmented regression models, the quadratic functional form for the post-intervention trend was designed
Cordtz et al. [13]	The biological disease-modifying anti-rheumatic drugs (bDMARDs) and associated rheumatoid arthritis (RA) management guidelines on the incidence of total hip (THR) and knee replacements (TKR) in Denmark	No (TKR) Yes (THR)	Yes	Yes (12 months)	Linear segmented regression was employed

Table 1 (continued)

Study	Intervention	Impacts on post-intervention period			Design strategy to analyze the interrupted time series outcome
		Trend	Level	Time lag	
Hawley et al. [14]	The UK National Institute for Health and Care Excellence (NICE) approval of tumor necrosis factor inhibitor therapies on the incidence of total hip (THR) and knee (TKR) replacement in rheumatoid arthritis patients	Yes	No (TKR) Yes (THR)	Yes (12 months)	Linear segmented regression was employed
Williams et al. [15]	The prescription fees on hospital admissions and prescribed medicines launched in 2008 in Scotland	Yes	Yes	0, 1-, 2-, and 3-month lags	Using moving average process a linear regression model was fitted
Gandrup et al. [16•]	The impact of three health-IT initiatives including electronic health record (EHR) flowsheet to input scores, peer performance reports, and an HER SmartForm including a Clinical Disease Activity Index (CDAI) calculator on performance of RA disease activity measures and outcomes	Yes	Yes	No	Using autoregressive errors a linear regression model was fitted
Kiadaliri et al. [17]	The Swedish health authority guideline for musculoskeletal diseases against the use of knee arthroscopy in patients aged ≥ 40 years with knee osteoarthritis	Yes	Yes	Yes (6 months)	Linear segmented regression was employed
Langaas et al. [18]	An academic detailing program in primary care on the prescribing rate of diclofenac, naproxen and non-steroidal anti-inflammatory drugs (NSAIDs) in Norway	Yes	Yes	No	Linear segmented regression was employed
Majka et al. [19]	The impacts of four multifaceted interventions including clinician education, point-of-care decision support, performance feedback to clinicians and care management on cardiovascular disease risk factor changes in rheumatoid arthritis	Yes	No	No	Using autoregressive errors a linear regression model was fitted
Meyer et al. [20]	A dose-reduction intervention of biological disease-modifying anti-rheumatic drugs (bDMARDs) in patients in remission was implemented in 2009 in Denmark	Yes	Yes	No	Linear segmented regression was employed
Trentalange et al. [21]	A system-wide Opioid Safety Initiative (OSI) implemented in October 2013 by Veterans Health Administration (VHA) to educate prescribers about safer opioid prescribing practices, followed by a sustained organizational effort to attenuate opioid prescribing	Yes	No	No	Linear segmented regression was employed

prediction of disease outcomes such as the incidence of joint replacement, as recently published for bDMARD in patients with RA [13], or on the incidence rate of upper limb joint replacements among newly diagnosed RA patients [26•]. With respect to orthopedic surgery, it might also facilitate investigation of whether repaired meniscal lesions in knee osteoarthritis patients can predict a delay of early disease and/or its progression, data that are not well documented presently. It could also be employed for the evaluation of patterns over time through modeling and analyzing the dynamics of change in reaction to an intervention, bridging the gap between practitioners and researchers in pre- and post-treatment.

Models of Interrupted Time Series

The vast amounts of data collected from patient historical records require certain assessments when selecting an approach for future data prediction. ITS modeling could eliminate problems caused by unknown and/or unmeasured parameters, which could affect the target parameter estimation. Clearly, care should be taken to use an appropriate model fitting. There are three main well-known approaches in ITS analysis. The traditional one involves segmented regression-based methods (linear and nonlinear), while the second uses artificial intelligence (AI) (nonlinear) approaches, and the third uses stochastic methods (linear).

Segmented Regression-Based Methods

Segmented regression is a method in regression analysis in which the independent variable (outcome) is partitioned into intervals and a separate line segment is fitted to each interval; it minimizes least square errors between each segment, thus between prediction and the true outcome. Indeed, by affecting an intervention at a change point, there are different possible outcome patterns in the post-intervention period for both level and trend parameters. As illustrated in Table 1, by applying an intervention, a change may occur in the trend and/or level immediately or with a lag period between intervention/program implementation and the assumed impact (see examples below). In addition, the intervention may lead to a temporary change in the outcome [10].

The most basic and common form of regression analysis used in predictive analysis is the linear regression approach, which is an attempt to model the relationship between two variables by fitting a linear equation to observed data. Also used is the nonlinear analysis, for which one of the most commonly employed forms is logistic regression. Nonlinear segmented regression is used when at least one of its parameters appears nonlinearly. In brief, in nonlinear segmented regression, a function that is a nonlinear combination of the model parameters, is used to model observational data and requires one or more independent variables. A method of successive approximations is used for fitting the data. The nonlinear terms of regression equations in each of the pre- and post-intervention (segment) could be applied for each segment, but such a model could be difficult to interpret [27]. Unlike the linear regression method, with this approach, for the best fitting parameters there is no closed-form expression [28]. Hence, such a regression approach may be used to measure the rate of outcome changes but it is unable to reliably predict the outcome.

Wagner et al. [29] introduced a strong, quasi-experimental approach to evaluate longitudinal effects of interventions by looking at the impacts of the policy and administrative interventions on the improvement of the quality of medication use and/or the containment of costs. Its approach was based on a variable indicating pre- or post-intervention periods and four different coefficients: two coefficients represent the baseline level trend and the other two, the slope and the level changes following the intervention implementation. The choice of each segment was based on the change point with the possible additional time lag in some cases to ensure that the intervention shows its impact on the outcome. In this approach, a linear (or nonlinear) equation should be adjusted to pre- and post-intervention outcomes. Therefore, accurate values of the pre-specified time lag parameters (to allow the implemented intervention to be effective) are required to define an appropriate equation for each segment [30]. However, as there is no specific rule to determine time lag period between pre- and post-intervention, once the intervention has been implemented, segmented regression restricts the analysis to one outcome from one unit. Thus, assessing the impact of an intervention with segmented regression requires a separate analysis for each individual unit. Of note, assessing the intervention impact on more than one unit via segmented regression methods ignores shared characteristics across units, in particular the similarity between characteristics influencing the change point. Since most interventions often have a dynamic impact on the outcome and they depend on several unknown parameters, accurate values for the time lag are unspecified. In addition to these limitations in this approach, it has been suggested [31] that various other factors like total amount of time points, effect size, and location of intervention should be considered in ITS modeling by segmented regression-based models to ensure accuracy of the obtained results.

Importantly, a change in time lag may lead to a change in accuracy of the prediction of the trend in post-intervention results. To illustrate such an impact of the choice of the time lag on the results of the segmented regression approach, the following is a hypothetical example of ITS in which we included different lag time between pre- and post-intervention, showing the effect on the post-intervention segmented linear regression (Fig. 1). In Fig. 1a, there is no lag between pre- and post-intervention which yield a correlation coefficient (R) for

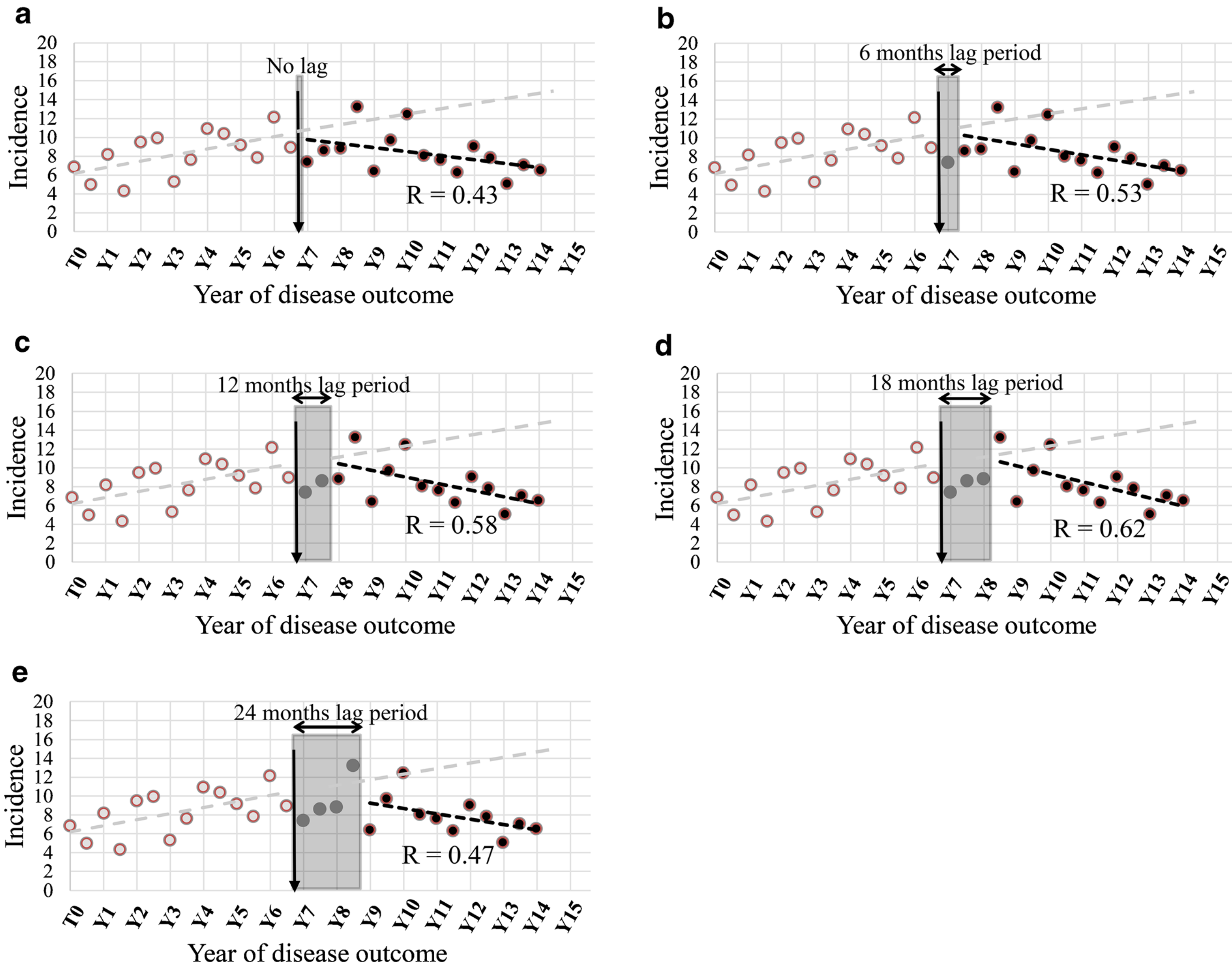

Fig. 1 An example of the possible impacts of different time lags on the segmented linear regression approach of a hypothetical incidence rate (dots). The linear regression (gray dashed lines) and linear regression **a** without or **b**–**e** with lag time (bold black dashed lines) for **b** 6 months, **c** 12 months, **d** 18 months, and **e** 24 months. *R* correlation coefficient

the segmented linear regression at post-intervention of 0.43. However, by considering a time lag of 6 months, the accuracy of prediction will be increased to 0.53 (Fig. 1b), and more so by applying 12 ($R = 0.58$) or 18 ($R = 0.62$) months lag time (Fig. 1c, d). However, it should be noted that by considering 24 months' time lag (Fig. 1e), compared to 6–18 months' lag time (Fig. 1b–d), there will be a considerable decrease ($R =$ 0.47) in the accuracy of the prediction.

Artificial Intelligence (AI)-Based Methods

The AI approach has been recommended as an alternative to the TS prediction technique. The objective is to build a reliable model that can identify patterns and recurrences in the input data and produce generalized results based on learned experience and prior knowledge. A recent report [32] highlights and discusses the significant potential of machine learning and AI approaches in increasing the accuracy and interpretability of arthritis prediction models.

An excellent feature of AI, when applied to TS forecasting problems, is its inherent capability in nonlinear modeling without any presumption about the statistical distribution followed by the observations [33]. These properties make such a model more precise and practical in modeling complicated patterns of data. The design of the reliable model is based on the given data: AI is data-driven and self-adaptive.

Another important advantage of AI approaches for ITS modeling is that they do not require any knowledge about the time lag [34••]. However, choosing an appropriate approach among several AI methods for a given ITS dataset is a challenging task, as there are no documented rules for such a selection; it is based only on the experiences of users, in addition to trial and error.

The basic function of an AI approach in ITS is the transformation of historical data into an outcome. To reach the point where an AI approach can predict, it is important to create a structure of how the best input combinations relate to the outcome. The values of the autocorrelation of TS with its passed values (lagged values) can be obtained by the autocorrelation function (ACF). The correlation of the residuals with time series lags can be captured by partial autocorrelation function (PACF). To select the best input combinations and the appropriate time lags, the ACF and PACF plots are used, as they illustrate how well the present value of the series relates to its earlier values. ACF considers all the TS components while finding correlations of present with lags time. PACF for its part, finds correlation of the residuals with the next lag value, helping to retain only the relevant features. This contrasts with the ACF, which does not control for other lags. These ACF and PACF plots are useful in defining the relevant range of inputs for creating different input combinations for ITS modeling.

By identifying effective lags in studied ITS via ACF and PACF plots, the different input combinations can be identified. Then, the modeling parameters of the nonlinear approach (different AI/machine learning methods) should be determined. Training dataset is utilized to build up the nonlinear approach and testing dataset is employed to evaluate the model performance. Another major issue is that too many or too few network parameters may lead to poor precision in predicted results due to over- and under-fitting issues. Thus, a resampling technique (cross-validation) is needed. Finally, a reproducibility experiment should be performed with a hold back or a new dataset to ensure the validity of the obtained results.

Stochastic Models

The stochastic model refers to any model involving a probability based on a linear regression and random sampling from historical data. One of the most common linear stochastic methods is the autoregressive integrated moving average (ARIMA) model used to forecast the outcome, which has found applications in medicine in recent years [35–38]. This model is implemented under the assumption that the considered TS is linear and follows a specific and known statistical distribution, for example, normal distribution.

ARIMA model, as a well-known statistical approach, employs TS data to better analyze the data values or to predict future trends. An ARIMA model is a set of regression analysis that measures the strength of one dependent variable compared to other changing variables. The objective of the model is to anticipate future moves by checking the differences between values in the series rather than through actual values. To simplify the development of stochastic processes, the concept of stationarity (when statistical properties such as the mean, variance, and autocorrelation do not change over time) was constructed, in which the TS is expected to be stationary to ensure adequate prediction. Unfortunately, this is not always the case; therefore, ARIMA models are applied where data show evidence of non-stationarity, and/or where an initial differencing step can be used one or more times to remove the non-stationary term.

Although stochastic models are relatively easy to apply, some general concepts should be considered in order to produce a model with the highest accuracy. Each dataset in a TS comprises deterministic (trend, periodical, random or jump) and autoregressive components. After assuring that throughout the calibration period, the TS length includes the extreme TS values, the existence of deterministic components should be investigated and eliminated [39]. The best model based on various possible scenarios (stationarization techniques and normalization transforms) are selected from sub-scenarios [40]. According to the pre-processing technique applied, the ACF and PACF of the obtained series are plotted to determine the modeling parameters for each sub-scenario. ARIMA's greatest limitation is that the associated TS' pre-assumed linear form is impractical in many scenarios [39, 40].

When Does Interrupted Time Series Stochastic/Artificial Intelligence-Based Modeling Fail?

Although the segmented regression-based approach is useful for predicting the level and trend in the outcome of interest after intervention implementation, it requires exact lag time values between pre- and post-intervention. In addition, this approach cannot reliably predict the outcome. Table 2 recapitulates some of the main challenges with the above described ITS modeling methods. In general, in conventional ITS modeling, different approaches are tried and the best is selected. However, due to many potential factors influencing patient's disease activity/progression, for example, medication interaction, lag time between fulfillment of the intervention/program, and the hypothesized impact, it is difficult to conclude whether any single model is the best and if it will be able to capture all possible future patterns in ITS. A reliable model should be designed with small error values on sample (training) data, but also on out-of-sample (testing) data. Furthermore, it is still challenging to determine whether stochastic methods are more effective than AI methods with out-of-sample prediction data in ITS modeling, due to many potentially influencing factors such as sampling variation, model uncertainty, and structural changes.

To overcome these limitations and to improve the accuracy of both of these methods, a new combined methodology could be used (Fig. 2). This methodology will not only model complex autocorrelation structures of ITS, but the challenges of the stochastic and AI-based models, as in Table 1, will also be addressed. Figure 2 presents the flowchart of such a combined methodology (stochastic and AI model integration). The goal

Table 2 Interrupted time series modeling models: challenges and key factors for success

Method	Challenge	How to address
Segmented regression-based	Selecting linear or nonlinear approach	Minimize least squared error between prediction and the true interested outcome
	Selecting the most appropriate time lag	There is *no specific rule* to define the time lag produced between the pre- and post-intervention periods
	Temporary change of outcome in post-intervention period	There is *no specific rule* to determine temporary change in level and trend of interested outcome
AI-Based	Selecting the best input combination	Use ACF and PACF graphs to define appropriate time lags
	Selecting an appropriate nonlinear approach	Trial and error to find appropriate methods
	Optimizing non-linear model parameters	Trial and error to adjust the best AI architecture to a given ITS with appropriate method parameters
	Poor performance due to over- and under-fitting	Resampling technique (cross-validation) and reproducibility with a hold back or new dataset
Stochastic	Minimum length of time series	Collect the data over more time
	Linear vs. nonlinear data	Applying an appropriate pre-processing technique
	Removal of time series components (jump, trend, period)	Use ACF and PACF graphs to define appropriate time lags
	Time series stationarization	Verify differencing, seasonal standardization, spectral analysis methods
	Time series normalization	Apply combined transformation
	Selecting parameters of the linear stochastic models	Increase stationarization steps

ACF autocorrelation function, *AI* artificial intelligence, *ITS* interrupted time series, *PACF* partial autocorrelation function

is to obtain data which graphically depicts the closeness between the original and predicted observations.

Below is a brief description for undertaking the new proposed TS prediction modeling (Fig. 2) using linear, nonlinear,

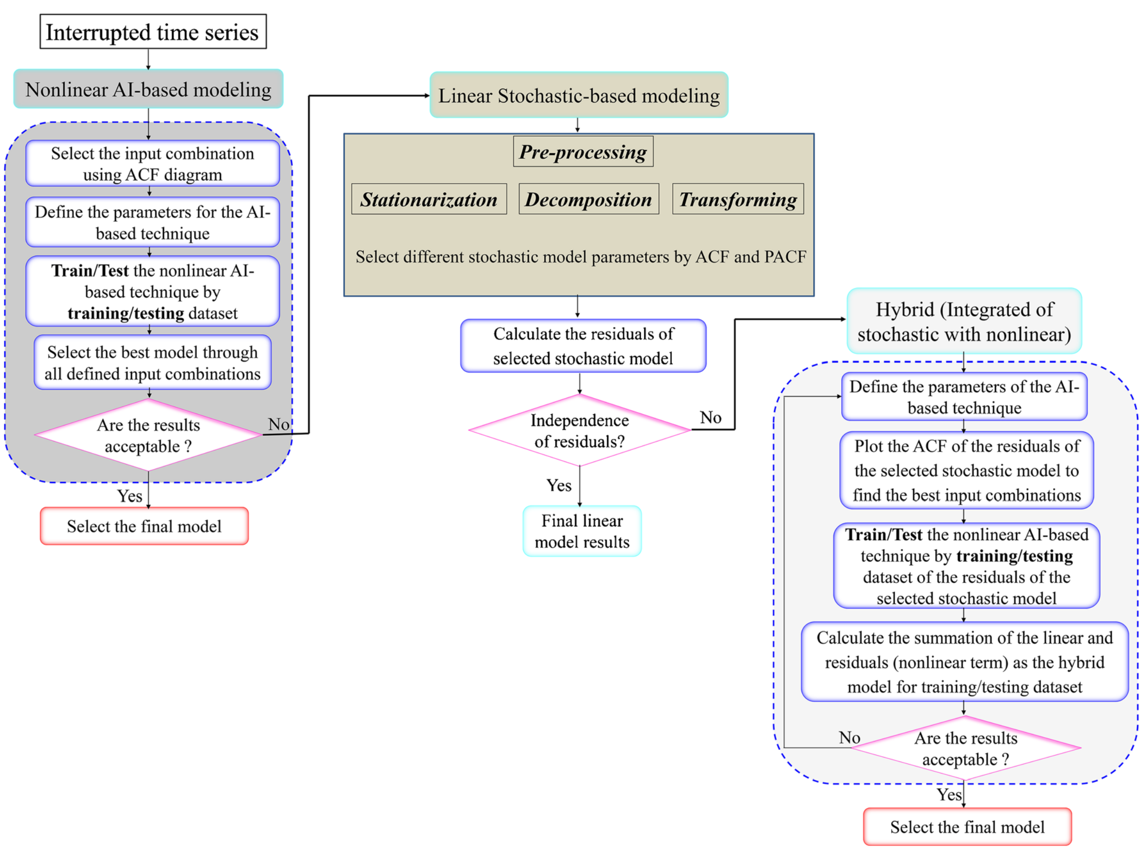

Fig. 2 Flowchart of the proposed combined methodology. *ACF* autocorrelation function, *AI* artificial intelligence, *PACF* partial autocorrelation function

and hybrid models. This methodology suggests starting ITS with AI-based modeling, whereby in plotting an ACF diagram, the different input combinations are defined. ACF reflects how the outcomes of interest in ITS are related to each other for determining the order of necessary time lags to create different input combinations. Based on the selected AI method, a suitable architecture should be obtained for the training and testing stages. This step necessitates trial and error to adjust the optimum values of the AI parameters, such as the number of hidden layers, the number of iterations, the training algorithm, etc. By calculating the respective statistical indices, the best model is selected through all defined input combinations. If the results obtained with this model are acceptable, the final model for ITS modeling is introduced. Otherwise, stochastic-based modeling is applied, whereby the time series components are identified by pre-processing. In a stationary ITS, statistical properties such as mean, variance and autocorrelation are constant over time. The normalization rescales ITS data from the original range into the range of 0 and 1. The ITS decomposition is a statistical task that deconstructs a TS into its components including stochastic and deterministic terms.

Therefore, stationarization, decomposition, and transformation are applied to create different stochastic sub-scenario models for ITS modeling. Next, the ACF and PACF plots are used to select the parameters of these models. By comparing the results obtained from these stochastic models and the observed ITS outcomes of interest, the values of the modeled residuals and the actual ITS values are calculated. If the results are not acceptable, a hybrid model that integrates the stochastic approach with a nonlinear model should be considered. Then, the stochastic model residuals are used as the new TS for AI modeling and nonlinear modeling is carried out with the new TS (residual series). After residual series modeling is completed, the predicted values are combined with the linear results and reported as the final hybrid method results. The advantages of this methodology are as follows: (i) increasing the accuracy of ITS modeling by profiting from the advantages of both stochastic and nonlinear models to capture all ITS deterministic and stochastic components; (ii) predicting the outcomes of interest that no segmented approach can predict; and (iii) modeling ITS as continuous variables without having to consider the time lags between, in order to answer a crucial challenge in the segmented regression approach regarding the time lag produced between the pre- and post-intervention periods. The latter answers the main challenge of segmented regression in ITS modeling.

As an example of a real case of arthritis disease and to illustrate the performance of the combined methodology in ITS simulation, Bonakdari et al. [34••] applied a continuous nonlinear model to predict the incidence of total hip (THR) and knee (TKR) replacement in RA patients. In brief, Hawley et al. [14] evaluate the impact of the NICE in UK approval of tumor necrosis factor inhibitor therapy on the temporal trends of THR and TKR among RA patients in the UK. As the regulation came into effect, two distinct linear regression models were fitted with a time lag of 12 months to calculate the change in the subsequent outcome levels and trends. The root mean square error (RMSRE) of the ITS post-intervention was 0.21 for THR and 0.73 for TKR and the R was 0.67 and 0.58, respectively. By applying the proposed continuous nonlinear model (no need to identify the change point and intervention lag time as the methodology simulated ITS continually throughout modeling), ITS predictive value was improved, in which there was a decrease in the RMSRE values (0.11 and 0.12 were found, respectively) and a considerable increase in R values (0.97 and 0.78). These results confirm not only that the model is able to identify the trend of the outcome of interest in ITS, but that it can also make high-accuracy predictions. As another illustration of the benefit of the proposed methodology, a dataset from Bernal et al. [10] evaluated the rate of acute coronary events (ACEs) associated with the implementation of a smoking ban in all indoor public places to calculate the change in the subsequent outcome levels and trends using a TS regression-based method. In this case study, by applying the proposed combined methodology, we found that the RMSRE and R values improved the prediction capability by 25% and 14%, respectively (authors' personal information).

Conclusion

Precision medicine for arthritis patients is an emerging approach of tailoring decisions/treatments to individual patient variability. This entails better diagnoses, earlier interventions, more efficient therapies and customized treatment plans. However, to achieve such an evolution in clinical management/treatment strategies, it is crucial that physicians and health professionals know about, understand, and have access to interactive processes so they may adapt therapeutic approaches in the most appropriate and efficient manner to provide superior precision medicine.

TS methodologies can be used to achieve such an aim. They can predict next period(s) value(s) using past and current information, by using, for example, a longitudinal database to predict a long-term outcome for an early arthritis patient. Moreover, a proper selection of the model, number of inputs and modeling architecture, and parameters are crucial for successful prediction.

However, each of the most commonly used TS methodologies still present major challenges. When the data is such that the common technologies cannot provide a strong and accurate prediction, a combined stochastic and AI approach, which considers the linear and nonlinear elements and minimizes prediction error by combining multiple different and distinct methods, could be used for greater prediction accuracy. In addition, as the combined method uses a continuous ITS modeling procedure in which the time lag produced between the pre- and post-intervention periods is not required, applying it will solve this major problem. By using such a tool, for a given patient, the intervention will be more appropriate and, among other benefits, overtreatment could be avoided.

In conclusion, the use of prediction methodologies such as TS and ITS for arthritis patients will allow us to treat the patient rather than just the disease, thus providing the best outcome for the patient.

Acknowledgments We would like to thank Jacqueline Brunet for her assistance in the preparation of this manuscript.

Author Contributions All the authors were equally involved in the conception, design, writing, and final approval of this viewpoint article.

Funding Information This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with Ethical Standards

Conflict of Interest J Martel-Pelletier declares no conflicts of interest with the content of this paper. She is a shareholder of ArthroLab Inc., and reports grants and lecture fees from TRB Chemedica, outside the submitted work.

J-P. Pelletier declares no conflicts of interest with the content of this paper. He is a shareholder of ArthroLab Inc., reports speaker fees from Mylan, speaker fees and grants from TRB Chemedica, and advisor fees from UCB Advisory Board, outside the submitted work.

H. Bonakdari has nothing to disclose.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

References

Papers of particular interest, published recently, have been highlighted as:

- • Of importance
- •• Of major importance

1. Gasparrini A, Gorini G, Barchielli A. On the relationship between smoking bans and incidence of acute myocardial infarction. Eur J Epidemiol. 2009;24(10):597–602. https://doi.org/10.1007/s10654-009-9377-0.
2. Desai SP, Lu B, Szent-Gyorgyi LE, Bogdanova AA, Turchin A, Weinblatt M, et al. Increasing pneumococcal vaccination for immunosuppressed patients: a cluster quality improvement trial. Arthritis Rheum. 2013;65(1):39–47. https://doi.org/10.1002/art.37716.
3. Robinson TE, Zhou L, Kerse N, Scott JD, Christiansen JP, Holland K, et al. Evaluation of a New Zealand program to improve transition of care for older high risk adults. Australas J Ageing. 2015;34(4):269–74. https://doi.org/10.1111/ajag.12232.
4. Linden A. Conducting interrupted time-series analysis for single- and multiple-group comparisons. Stata J. 2015;15(2):480–500. https://doi.org/10.1177/1536867X1501500208.
5. Wang Z, Bhattacharyya T. Trends of non-union and prescriptions for non-steroidal anti-inflammatory drugs in the United States, 1993-2012. Acta Orthop. 2015;86(5):632–7. https://doi.org/10.3109/17453674.2015.1028860.
6. Judge A, Wallace G, Prieto-Alhambra D, Arden NK, Edwards CJ. Can the publication of guidelines change the management of early rheumatoid arthritis? An interrupted time series analysis from the United Kingdom. Rheumatology (Oxford). 2015;54(12):2244–8. https://doi.org/10.1093/rheumatology/kev268.
7. Baker DW, Brown T, Lee JY, Ozanich A, Liss DT, Sandler DS, et al. A multifaceted intervention to improve influenza, pneumococcal, and herpes zoster vaccination among patients with rheumatoid arthritis. J Rheumatol. 2016;43(6):1030–7. https://doi.org/10.3899/jrheum.150984.
8. Parkinson L, Moorin R, Peeters G, Byles J, Blyth F, Caughey G, et al. Incident osteoarthritis associated with increased allied health services use in 'baby boomer' Australian women. Aust N Z J Public Health. 2016;40(4):356–61. https://doi.org/10.1111/1753-6405.12533.
9. Aladul MI, Fitzpatrick RW, Chapman SR. Impact of infliximab and etanercept biosimilars on biological disease-modifying zntirheumatic drugs utilisation and NHS budget in the UK. BioDrugs. 2017;31(6):533–44. https://doi.org/10.1007/s40259-017-0252-3.
10. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol. 2017;46(1):348–55. https://doi.org/10.1093/ije/dyw098.
11. Lin LA, Bohnert ASB, Kerns RD, Clay MA, Ganoczy D, Ilgen MA. Impact of the opioid safety initiative on opioid-related prescribing in veterans. Pain. 2017;158(5):833–9. https://doi.org/10.1097/j.pain.0000000000000837.
12. Zombré D, De Allegri M, Ridde V. Immediate and sustained effects of user fee exemption on healthcare utilization among children under five in Burkina Faso: a controlled interrupted time-series analysis. Soc Sci Med. 2017;179:27–35. https://doi.org/10.1016/j.socscimed.2017.02.027.
13. Cordtz RL, Hawley S, Prieto-Alhambra D, Hojgaard P, Zobbe K, Overgaard S, et al. Incidence of hip and knee replacement in patients with rheumatoid arthritis following the introduction of biological DMARDs: an interrupted time-series analysis using nationwide Danish healthcare registers. Ann Rheum Dis. 2018;77(5):684–9. https://doi.org/10.1136/annrheumdis-2017-212424.
14. Hawley S, Cordtz R, Dreyer L, Edwards CJ, Arden NK, Delmestri A, et al. Association between NICE guidance on biologic therapies with rates of hip and knee replacement among rheumatoid arthritis patients in England and Wales: an interrupted time-series analysis. Semin Arthritis Rheum. 2018;47(5):605–10. https://doi.org/10.1016/j.semarthrit.2017.09.006.
15. Williams AJ, Henley W, Frank J. Impact of abolishing prescription fees in Scotland on hospital admissions and prescribed medicines: an interrupted time series evaluation. BMJ Open. 2018;8(12):e021318. https://doi.org/10.1136/bmjopen-2017-021318.
16.• Gandrup J, Li J, Izadi Z, Gianfrancesco M, Ellingsen T, Yazdany J, et al. Three quality improvement initiatives improved performance of rheumatoid arthritis disease activity measures in electronic health records: results from an interrupted time series study. Arthritis Care

Res (Hoboken). 2019;72(2):283–91. https://doi.org/10.1002/acr.23848. **An important contribution for evaluating the effect of health information on the performance of rheumatoid arthritis disease activity outcomes.**
17. Kiadaliri A, Bergkvist D, Dahlberg LE, Englund M. Impact of a national guideline on use of knee arthroscopy: an interrupted time-series analysis. Int J Qual Health Care. 2019;31:G113–8. https://doi.org/10.1093/intqhc/mzz089.
18. Langaas HC, Hurley E, Dyrkorn R, Spigset O. Effectiveness of an academic detailing intervention in primary care on the prescribing of non-steroidal anti-inflammatory drugs. Eur J Clin Pharmacol. 2019;75(4):577–86. https://doi.org/10.1007/s00228-018-02611-y.
19. Majka DS, Lee JY, Peprah YA, Lipiszko D, Friesema E, Ruderman EM, et al. Changes in care after implementing a multifaceted intervention to improve preventive cardiology practice in rheumatoid arthritis. Am J Med Qual. 2019;34(3):276–83. https://doi.org/10.1177/1062860618798719.
20. Meyer MK, Andersen M, Ring T, Andersen GN, Ehlers LH, Rasmussen C, et al. Personalized rheumatic medicine through dose reduction reduces the cost of biological treatment - a retrospective intervention analysis. Scand J Rheumatol. 2019;48(5):398–407. https://doi.org/10.1080/03009742.2019.1585940.
21. Trentalange M, Runels T, Bean A, Kerns RD, Bair MJ, Brody AA, et al. Analgesic prescribing trends in a national sample of older veterans with osteoarthritis: 2012-2017. Pain. 2019;160(6):1319–26. https://doi.org/10.1097/j.pain.0000000000001509.
22. McDonald S, Quinn F, Vieira R, O'Brien N, White M, Johnston DW, et al. The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview. Health Psychol Rev. 2017;11(4):307–23. https://doi.org/10.1080/17437199.2017.1316672.
23. Hendrikx J, Fransen J, van Riel PL. Monitoring rheumatoid arthritis using an algorithm based on patient-reported outcome measures: a first step towards personalised healthcare. RMD Open. 2015;1(1):e000114. https://doi.org/10.1136/rmdopen-2015-000114.
24. Glintborg B, Sorensen IJ, Loft AG, Lindegaard H, Linauskas A, Hendricks O, et al. A nationwide non-medical switch from originator infliximab to biosimilar CT-P13 in 802 patients with inflammatory arthritis: 1-year clinical outcomes from the DANBIO registry. Ann Rheum Dis. 2017;76(8):1426–31. https://doi.org/10.1136/annrheumdis-2016-210742.
25. Glintborg B, Ibsen R, Bilbo REQ, Lund Hetland M, Kjellberg J. Does a mandatory non-medical switch from originator to biosimilar etanercept lead to increase in healthcare use and costs? A Danish register-based study of patients with inflammatory arthritis. RMD Open. 2019;5(2):e001016. https://doi.org/10.1136/rmdopen-2019-001016.
26.• Cordtz R, Hawley S, Prieto-Alhambra D, Hojgaard P, Zobbe K, Kristensen LE, et al. Reduction in upper limb joint surgery among rheumatoid arthritis patients: an interrupted time-series analysis using Danish health care registers. Arthritis Care Res (Hoboken). 2020;72(2):274–82. https://doi.org/10.1002/acr.23835. **An important study illustrating the relevance of interrupted time-series results in rheumatoid arthritis patients.**
27. Gillings D, Makuc D, Siegel E. Analysis of interrupted time series mortality trends: an example to evaluate regionalized perinatal care. Am J Public Health. 1981;71(1):38–46. https://doi.org/10.2105/ajph.71.1.38.
28. Terrell C. Predictions in time series using regression models. 1st ed. Waltham Abbey Essex: ED-Tech Press; 2019.
29. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther. 2002;27(4):299–309. https://doi.org/10.1046/j.1365-2710.2002.00430.x.
30. Penfold R, Zhang F. Use of interrupted time series analysis in evaluating health care quality improvements. Acad Pediatr. 2013;13(6 Suppl):S38–44. https://doi.org/10.1016/j.acap.2013.08.002.
31. Hawley S, Ali MS, Berencsi K, Judge A, Prieto-Alhambra D. Sample size and power considerations for ordinary least squares interrupted time series analysis: a simulation study. Clin Epidemiol. 2019;11:197–205. https://doi.org/10.2147/CLEP.S176723.
32. Jamshidi A, Pelletier J-P, Martel-Pelletier J. Machine-learning-based patient-specific prediction models for knee osteoarthritis. Nat Rev Rheumatol. 2019;15(1):49–60. https://doi.org/10.1038/s41584-018-0130-5.
33. Gharabaghi S, Stahl E, Bonakdari H. Integrated nonlinear daily water demand forecast model (case study: city of Guelph, Canada). J Hydrol. 2019;579:124182. https://doi.org/10.1016/j.jhydrol.2019.124182.
34.•• Bonakdari H, Pelletier JP, Martel-Pelletier J. A reliable time-series method for predicting arthritic disease outcomes: New step from regression toward a nonlinear artificial intelligence method. Comput Methods Prog Biomed. 2020;189:105315. https://doi.org/10.1016/j.cmpb.2020.105315. **This study demonstrated, for the first time, that by using a new interrupted times-series (ITS) nonlinear method, there is no need to identify the change point and intervention lag time for ITS modeling.**
35. Pratyaksa H, Permanasari AE, Fauziati S, Fitriana I, editors. ARIMA implementation to predict the amount of antiseptic medicine usage in veterinary hospital. 2016 1st International Conference on Biomedical Engineering (IBIOMED); 2016; Yogyakarta, Indonesia.
36. Revels S, Kumar SA, Ben-Assuli O. Predicting obesity rate and obesity-related healthcare costs using data analytics. Health Policy Technol. 2017;6(2):198–207. https://doi.org/10.1016/j.hlpt.2017.02.002.
37. Petukhova T, Ojkic D, McEwen B, Deardon R, Poljak Z. Assessment of autoregressive integrated moving average (ARIMA), generalized linear autoregressive moving average (GLARMA), and random forest (RD) time series regression models for predicting influenza a virus frequency in swine in Ontario, Canada. PLoS One. 2018;13(6):e0198313. https://doi.org/10.1371/journal.pone.0198313.
38. Su K, Xiong Y, Qi L, Xia Y, Li B, Yang L et al., editors. City-wide influenza forecasting based on multi-source data. IEEE International Conference on Big Data (2018); 2018; Seattle, USA.
39. Bonakdari H, Moeeni H, Ebtehaj I, Zeynoddin M, Mahoammadian A, Gharabaghi B. New insights into soil temperature time series modeling: linear or nonlinear? Theor Appl Climatol. 2019;135(3–4):1157–77. https://doi.org/10.1007/s00704-018-2436-2.
40. Zeynoddin M, Bonakdari H, Azari A, Ebtehaj I, Gharabaghi B, Riahi MH. Novel hybrid linear stochastic with non-linear extreme learning machine methods for forecasting monthly rainfall a tropical climate. J Environ Manage. 2018;222:190–206. https://doi.org/10.1016/j.jenvman.2018.05.072.