

A new numerical method for processing longitudinal data: clinical applications

Stura Ilaria⁽¹⁾, Perracchione Emma⁽²⁾, Migliaretti Giuseppe⁽¹⁾, Cavallo Franco⁽¹⁾

(1) Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino, Italia

(2) Dipartimento di Matematica Tullio Levi-Civita, Università di Padova, Italia

CORRESPONDING AUTHOR: Ilaria Stura, Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino, Italia, email: Ilaria.stura@unito.it

DOI: 10.2427/12881

Accepted on May 14, 2018

ABSTRACT

Background: Processing longitudinal data is a computational issue that arises in many applications, such as in aircraft design, medicine, optimal control and weather forecasting. Given some longitudinal data, i.e. scattered measurements, the aim consists in approximating the parameters involved in the dynamics of the considered process. For this problem, a large variety of well-known methods have already been developed.

Results: Here, we propose an alternative approach to be used as effective and accurate tool for the parameters fitting and prediction of individual trajectories from sparse longitudinal data. In particular, our mixed model, that uses Radial Basis Functions (RBFs) combined with Stochastic Optimization Algorithms (SOMs), is here presented and tested on clinical data. Further, we also carry out comparisons with other methods that are widely used in this framework.

Conclusion: The main advantages of the proposed method are the flexibility with respect to the datasets, meaning that it is effective also for truly irregularly distributed data, and its ability to extract reliable information on the evolution of the dynamics.

Key words: Statistical Method; Radial Basis Function; Stochastic Optimization Algorithm; Longitudinal Data.

INTRODUCTION

Longitudinal data are often object of study in many fields, e.g. sociology, meteorology and medicine. In medicine, repeated measurements are used to monitor the patients' behaviours and also to adjust the therapies accordingly. However, many problems occur when these data are analysed. Indeed, each time series could have a different number of observations and not equally spaced. In addition, the sampling period could vary from patient to patient, measurement errors and also missing data often occur. Thus, since in these cases common methods,

such as linear regression, usually fail, the recent research is directed towards more robust statistical methods. For instance, longitudinal data are commonly analysed using parametric models such as Bayesian ones [1] and Functional Data Analysis (FDA) [2,3]. In both cases, many data are required in order to model the behaviour of the studied variable(s). These methods, in fact, try to find an 'average curve' using all the data, including truncated series and observations with missing information.

However, in clinical applications the estimate on the future dynamics of a single series, given few previous values, could be needed; think for instance to tumour volumes during

a treatment, height/weight of children during growth, concentration of some substance in the body. Each patient is different and could have different growth behaviour and different growth parameters, so an 'average curve' could not be sufficient. An important information could be, for example, the possible future development of the subject, given his/her previous growth and the clinical background (e.g. treatments). These data could be compared with the real dynamics, in order to see if the response of the patients to the treatment is stable (the parameters do not vary in the future) or not (change in the parameters).

The aim of our work is to propose our numerical tool that can provide information on the future dynamics given few follow-up data. Thus, we first model longitudinal data via widely used mathematical models in population dynamics. Therefore, on one hand we aim at validating such model by approximating the parameters involved in the dynamics. On the other one, we are also interested in giving reliable information on the future dynamics of the curves.

In order to achieve our goal, we propose our numerical tool based on optimization methods coupled with interpolation techniques. Specifically, we approximate the parameters involved in the dynamics by means of Stochastic Optimization Algorithms (SOMs) [4-7]. Moreover, for each data series, we improve the performances of the optimization tools by means of Radial Basis Function (RBF) interpolation; see [8] for a general overview and [9,10] for particular instances on the topic and applications. In the interpolation process, we also take into account the critical computational issue of carrying out stable computations. For this reason, and since data are subject to noise, we adopt a kind of Tikhonov regularization (see [11]).

The method, namely RBF-SOM, is here tested on two different datasets:

- Height measurements of children with a diagnosis of Growth Hormone Deficiency (GHD) during treatment,
- Prostate Specific Antigen (PSA) values of prostatectomized patients with a recurrence of prostate cancer.

The paper is organized as follows: in Section 2 the RBF-SOM technique is described. In Section 3 the two datasets used for the validation are presented. Section 4 is devoted to the numerical results and it is divided into two subsections: in the first one, all the data of each series are considered in order to reconstruct the curves and approximate the parameters, while, in the second one, only few initial data of each series are used to predict the curve behaviour. In Sections 5 and 6 discussion and conclusions are presented.

METHODS

This section is devoted to describe the method used to fit a given data series and to approximate the parameters

involved in the dynamics.

Given several scattered measurements $\{y_i\}_{i=1}^N$, sampled at different times $\{t_i\}_{i=1}^N$, the basic idea of the RBF-SOM here proposed consists in considering the theoretical function f , depending on the time t and on several parameters $\lambda = (\lambda_1, \dots, \lambda_p)$, and to approximate such parameters in order to obtain reliable information on the biological or physical phenomenon.

In the proposed examples, we use, as theoretical growth curve f , the so-called *Gompertzian function*:

$$f(t, \lambda_1, \lambda_2) = \lambda_2 \exp \left\{ -\log \left(\frac{\lambda_2}{f_0} \right) \exp[\lambda_1(t - t_0)] \right\},$$

where f_0 is the measurement at time t_0 (i.e. the first measurement), λ_1 is the growth rate and λ_2 is the carrying capacity, i.e. the maximum value that can be asymptotically achieved by f .

The Gompertzian function is characterized by a fast-growing initial period and by a progressive slowdown, reaching a carrying capacity after a certain time. This curve, depending on the values of the parameters, is able to model a variety of types of growth, from human to cancer cells ones, see [12-16] for details. For this reason, we will use in Section 5 the same function for both datasets. Moreover, its form is particularly suitable in this study because the parameter estimation is not possible via simple methods like Least Square Approximation.

Trivially, the parameters are approximated by finding

$$\tilde{\lambda} = \min_{\lambda} \left(\sum_{k=1, \dots, N} (y_i - f(t_i, \lambda_1, \lambda_2))^2 \right).$$

Note that we need optimization methods that can be used in case of non-linearity of f , as in the considered cases. In particular, we direct our research throughout stochastic methods. They have been designed by considering analogies with natural phenomena. The most popular are evolution strategy and genetic algorithms, both based on competition among individuals. On the opposite, other methods proposed in the last decades mainly focus on cooperation. Among them, Particle Swarm Optimization (PSO), Cuckoo Search (CS) and ant colony are widely used techniques, based on the mutual interaction and exchange of information between individuals. In particular, here we will consider PSO and CS briefly described in what follows.

PSO has been firstly introduced in [4] by Kennedy (social psychologist) and Eberhart (electrical engineer) and further developed by other researches, see for instance [6,7,17]. In order to describe it, let us consider a group of particles or birds which are represented as points in the space. At first, we need to model their way of flying. Then, taking into account that the target of birds consists in looking for the maximum availability of food, i.e. the minimum of the objective function f , we can easily find its minimum.

The main objective consists in simulating the trajectories of the single birds by considering their selfish behaviour (which is the ability of a bird of randomly flying away from the flock) and their social behaviour (which is the ability of a bird of staying in the group). With these simple considerations, it is possible to simulate the way of moving of a group of birds, taking also into account that particles avoid collisions.

To explain how we can find the minimum of the objective function interpreting the latter as food, let us first suppose that a bird discovers some food. Then, the other birds have two alternatives:

- get out of the flock and reach the food (selfish behaviour);
- stay in the flock (social behaviour).

If a good trade-off between the two behaviours is allowed, then the flock can reach the minimum. Indeed, if a bird can move towards some food then other birds can change their directions towards the same place. Acting in this way, the flock gradually changes its direction until the best place, i.e. the minimum, is reached.

As concerns CS, it was developed by Yang [18] and it simulates the behaviour of the cuckoo, a bird that does not incubate its eggs but tries to put them in nests of other species. The problem of this conduct is that, in some cases, the egg is removed by the nest's owner. The cuckoo, then, searches a nest in which its egg can be 'confused' with the others. Therefore, in this algorithm the minimum of the function is the nest in which more cuckoos can put their eggs without being discovered.

As for the PSO, the user needs to give a set of possible initial solutions. They are usually randomly initialized. Indeed, if the initial solutions are chosen so that they are feasible, the stochastic methods do not fail into local minima and thus the methods are not truly sensitive with respect to the initial conditions. The main difference with respect to the PSO approach is that, at each iteration, a fraction of nests, which are far from the minimum, are abandoned and new ones close to the minimum are built.

Note that both PSO and CS approaches can be performed in order to minimize the target function, but unfortunately the cardinality of the samples in concrete applications is really small. Thus, in order to improve the performance of the optimization methods, we first reconstruct the growth curves by means of a RBF-based interpolation scheme; see (Fasshauer and McCourt 2015; Wendland 2005) "plainCitation": "(Cavoretto, de Rossi, and Perracchione 2017; Fasshauer and McCourt 2015; Wendland 2005. In doing so, we also take into account the instability problems arising in applications. An example of RBF reconstruction can be seen in Fig.1a-b (big coloured dots).

Moreover, the so-reconstructed function can be used to estimate $\lambda_t, t > N$, i.e. the evolution of the considered quantity.

As accuracy indicator, we use the following Root Mean Square Error:

$$\epsilon = \sqrt{\frac{\sum_{i=1}^m (f(t_i, \lambda_1, \lambda_2) - y_i)^2}{m}},$$

where m denotes the number of patients. This indicator represents the standard deviation of the differences between predicted and observed values.

DATA

In order to assess strength and weaknesses of the discussed methods, we use two different datasets presented below. Both datasets are real patients' data. Each patient has a different number of irregularly spaced measurements in a different time interval.

Children data

Different problems can occur during growth. Here we consider paediatric patients with a diagnosis of Idiopathic Growth Hormone Deficit (IGHD), i.e. a low or absent production of GH for unknown causes. These children are treated with rhGH (a synthetic GH) and monitored during growth.

Our dataset is composed by 121 male IGHD patients, treated in "Ospedale Infantile Regina Margherita (OIRM)" in Turin between January 2000 and January 2016 [16,20] few studies analyze the effect of GH therapy on height, preferring a more indirect approach, where factors influencing the total pubertal and pre-pubertal growth in GH-deficient patients are evaluated and subsequently used to estimate the overall effect at the end of the therapy; unfortunately, this approach does not quantify the real growth gain in treated patients. Using a non-parametric Empirical Bayes approach, our study analyzes the growth response to GH treatment in a homogeneous cohort of 317 patients with pituitary GH deficiency who were enrolled during their pre-pubertal stage in the GH Piedmont Registry (Italy). The measurements are collected each 4-6 months from the beginning of the therapy (age between 3 and 14 years) to the adulthood (18-20 years old). As explained in [12], each period of human growth can be modelled with a Gompertzian law. We therefore use, as theoretical function, the Gompertz curve explained in the section before.

Note that each series is monotonic, strictly growing, not (or slightly) affected by measurement errors and it has very few missing data. Therefore, we expect robust performances of the RBF-SOM method for this clinical test.

Prostate cancer data

Data released from clinicians about the PSA value (which is a mirror of the mass of the prostate cancer) are

needed in order to have a reliable estimate the cancer's evolution. After a radical prostatectomy, the PSA turns out to be a good biomarker and it could be used to monitor a possible relapse. In fact, only prostate cells produce the PSA and, after surgery, there should be no prostate cells in the body. Hence, the PSA value should be very small, close to zero. If its value is bigger than 0.2 ng/mL, then PSA-producer cells are present, i.e. a relapse (a local or distal metastasis) occurs.

Here we use a subset of the Eureka1 study collection [21]. Eureka1 is a retrospective study on Italian patients who had a prostatectomy in the last 15 years. Our subset contains follow up data (PSA values series) of relapsed patients who did not undergo an adjuvant therapy.

In general, cancer growth is very fast and it is modelled with an exponential function. However, prostate cancer is a very slow-growing tumour, so Gompertzian [22-24] and West [25-27] laws are often used. In the following section, we will use the Gompertzian one.

Note that these series are not monotonic: PSA values can grow, be stable for months or decrease. PSA values should be sampled each 4-6 months, while in some series only 1-2 values are reported in 1-2 years. The PSA value strictly depends on the accuracy of instruments used in laboratory: some machines have a precision of 0.1 ng/mL, while others of 0.01 ng/mL. Moreover, the precision could change from value to value in the same series. Therefore, for the RBF-SOM method this dataset is more challenging than the previous one.

RESULTS

Curve reconstruction

The tests carried out in this subsection are devoted to assess the robustness of the RBF-SOM method as descriptive tool. Indeed, we consider each growth curve and we approximate the two parameters involved in the dynamics, i.e. the growth rate and carrying capacity. We remark that, at first, we reconstruct the curve by means of RBF interpolants and we then apply PSO or CS methods. The aim of these experiments is to obtain a feedback about the accuracy of the proposed approach in approximating the parameters.

Figure 1 shows two examples of the methods in the case of one unknown parameter. In particular, Fig.1a shows the reconstruction of a growth curve of one child (GH Database), while Fig.1b shows the reconstruction of a PSA growth curve. The circles represent real data, the big coloured dots are the RBF interpolation and the straight and dotted lines are the PSO and CS reconstructions respectively.

For both datasets, we consider as theoretical growth curve the Gompertzian function. Note that, differently from the growth rate, the carrying capacity (the maximum value

that can be achieved asymptotically) can be supposed to be fixed and known for all patients in both datasets, i.e. only one parameter should be estimated. In what follows, we test the methods with both one and two unknown parameters. As expected, when only the growth rate needs to be estimated, we obtain better results.

More precisely, as concerns only one unknown parameter λ_1 (while λ_2 is fixed as 200 cm in GH and 200 ng/mL in PSA database) both RBF-PSO and RBF-CS are truly performing and show the same accuracy ($\epsilon = 1.2$ cm in GH and $\epsilon = 0.9$ ng/mL in PSA database). Results are shown in Table 1. This means that data follow the theoretical distribution predicted by the Gompertzian and that the parameter is accurately estimated.

As concerns two unknown parameters, RBF-CS is more effective than RBF-PSO. In fact, for the former the $\epsilon = 0.80$ cm and $\epsilon = 0.89$ ng/mL in GH and PSA database, respectively, while for the latter the $\epsilon = 2.66$ cm and $\epsilon = 0.91$ ng/mL. Results are shown in Table 1. Note that this happens also enlarging the basin of possible solutions in the CS algorithm.

Dynamics evolution

This subsection is addressed to test the method for modelling the evolution of the dynamics. Thus, we apply the RBF-SOM using only the first four values of each series. We select the first four values for several reasons. First of all, the data availability: indeed, only few patients (in both databases) have more than six values in their series. Then, selecting four data in the databases means that they represent about two years of follow-up, that is a reasonable period to understand the future dynamics.

Again, by fixing the carrying capacity, both methods give the same results ($\epsilon = 3.01$ cm in GH and $\epsilon = 3.20$ ng/mL in PSA database). Results are shown in Table 1. Of course, the error is larger than the one shown in the previous subsection. However, since here we only take four values, such error indicates a meaningful accuracy on the estimation, especially as concerns the GH case. In fact, it means that, after four measurements, we are able to estimate the final height with a precision of about 3 cm. Looking at the other dataset, the error seems truly large, is $\epsilon = 3.20$ ng/mL. But, the majority (73.46%) of estimates differ less than 1 ng/mL to the real value. Unfortunately, the irregularity of the measurements and the errors due to the machine precision might cause problems for several series. Indeed, in few cases (6.12%), results differ more than 3 ng/mL from the real value.

Figure 2 shows some examples of the output of the method, in case of one unknown parameter: the circles are the real data (heights or PSA during visits), dots are the RBF interpolation on the first four values and PSO and CS provisions are the straight and dotted lines. Fig.2a-b concern on GH database, while Fig.2c-d on PSA

FIGURE 1. One unknown parameter: curve reconstruction of a) the height of a GH patient; b) the PSA values of a prostatectomized patient.

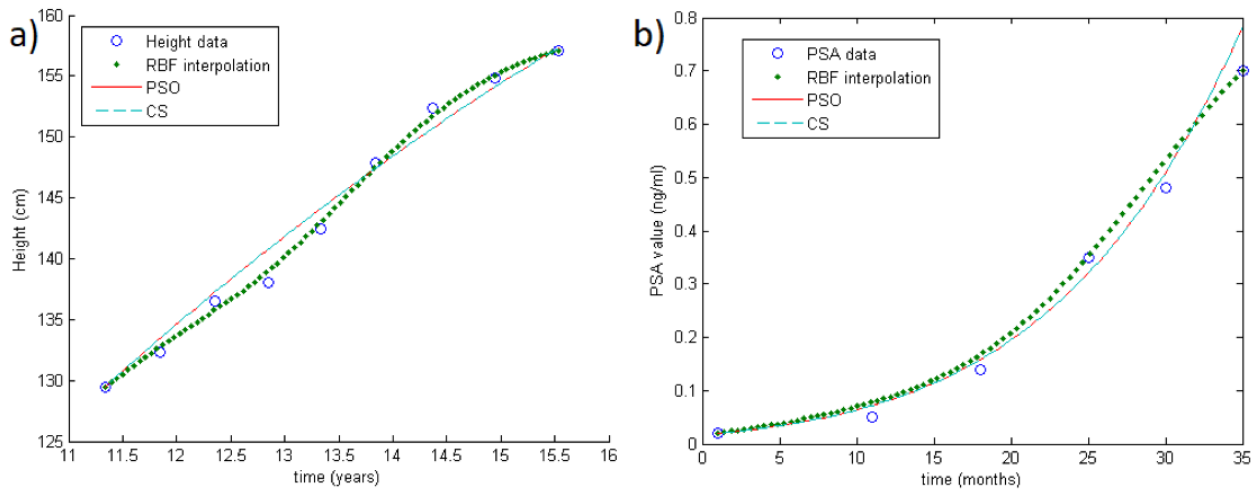


TABLE 1. Summary of the accuracy of the methods on test data

MODEL	N° UNKNOWN PARAMETERS	DATABASE	RBF-PSO (E)	RBF-CS (E)
Curve reconstruction	1	GH	1.2 cm	1.2 cm
		PSA	0.9 ng/ml	0.9 ng/ml
	2	GH	2.66 cm	0.80 cm
		PSA	0.91 ng/ml	0.89 ng/ml
Dynamics evolution	1	GH	3.01 cm	3.01 cm
		PSA	3.20 ng/ml	3.20 ng/ml

database. In particular, Fig.2a shows only one growth period, while in Fig.2b there is the combination of two growth periods. Note in Fig.2d that the estimation error grows during time, because of the irregularity of the data.

As for the two parameters (carrying capacity and growth rate), the estimate on the final height cannot be performed with only four values. Indeed, it is well known that there are no optimization tools robust enough to accurately approximate two parameters given only four samples. However, we remark once more that more measurements can be added to the model and in this sense, after also approximating the carrying capacity, the estimate on the final value can be sensibly improved.

DISCUSSION

From the numerical experiments, we note that both RBF-CS and RBF-PSO provide reliable approximations, even if CS seems to be less affected by the quality of input data.

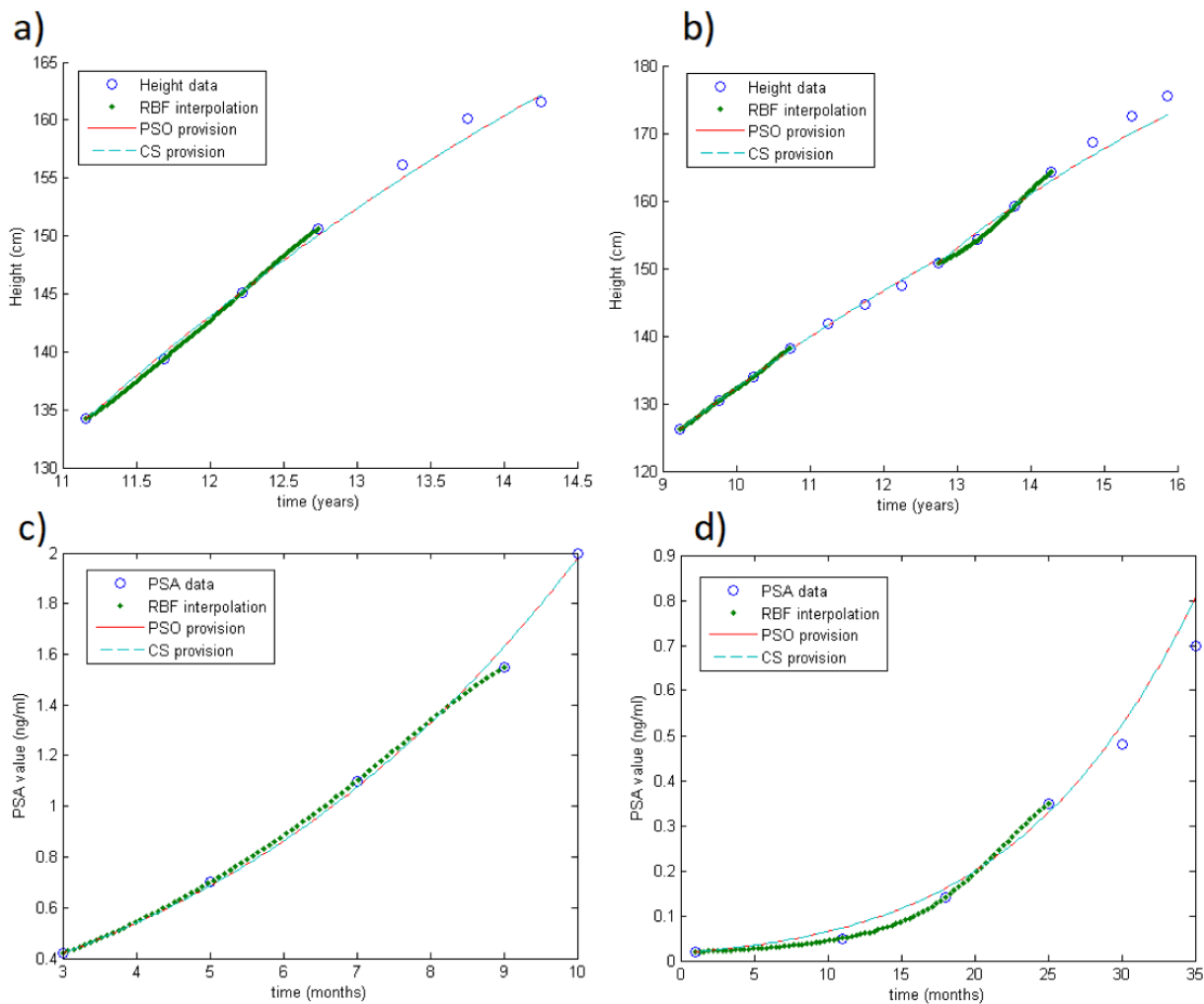
The novelty of the presented RBF-SOM method is the type of output. Indeed, it provides a (continuous) growth curve allowing the analysis of each growth function

independently of the others. This can be an advantage for doctors: for example, they can understand if the treatment is effective and, if not, change the therapy accordingly. In this sense, this approach is opposite to the most used techniques such as Bayesian methods or FDA. They analyse a (hopefully) large amount of similar data, from different patients, and try to find a common behaviour (or groups of them), in order to make predictions about new 'standard' patients. RBF-SOM, instead, starts from very few data of a single patient and estimates the curve, with a 'personalized medicine' approach.

RBF-SOM and FDA have the same starting point: each series is considered as a curve [3,22,28], while Bayesian methods take into account data points. RBF-SOM and FDA find coefficients that describe the growth, while Bayesian methods are non-parametric. FDA gives a polynomial reconstruction, i.e. the coefficients represent velocity, acceleration and other derivatives, while RBF-SOM finds the biological parameters of the chosen theoretical function. In some cases, a simple description of the growth velocity could be sufficient. RBF-SOM is more problem-specific, but this means that a theoretical function, such as the Gompertzian one, needs to be known.

As concerns the reconstruction, FDA is similar to

FIGURE 2. One unknown parameter: dynamics evolution of a-b) height of GH patients; c-d) PSA values of prostatectomized patients.



RBF. Our method introduce in addition errors due to the parameter approximation and the adhesion of data to the theoretical function. Bayesian methods are flexible in clustering and curve reconstruction, because are non-parametric techniques and do not need theoretical functions like the Gompertzian one. Hence, in order to reconstruct the whole curve and/or make clusters, FDA and Bayesian methods are preferable to RBF-SOM.

Concerning the estimate of the future dynamics of the curve, RBF-SOM is proved to be reliable and robust even with few initial data. Moreover, the result is personalized on the single patient, which is not possible using Bayesian methods. FDA cannot be used in this way, while it can compare different versions of the same phenomenon and distinguish phase and amplitude variations (see for example [28]).

CONCLUSIONS

Longitudinal data can be analysed in different ways. In the majority of cases, the goal of the analysis is to find

trend, or cluster, that a new series would follow, given many similar series and some additional parameters.

In this work, the goal is slightly different: to find the exact future dynamics of the single curve, given the previous data (of the curve) and the theoretical function that it should follow.

This approach could be useful in order to both consider biological meaning on the parameters of the curve and predict the future of the series without considering a large amount of similar data.

Moreover, RBF-SOM is a good tool for the reconstruction of curves, given its expected shape. It could be used in a large variety of cases, for both analysing pre-existent data as well as producing estimates of future scenarios, in a personalized medicine framework.

Ethics approval and consent to participate

Informed consent and ethical approvals were obtained from the participants of both studies.

Data are anonymous and were used in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Authors' contributions

IS and EP have created the mathematical method, tested it on data and been involved in drafting the manuscript.

GM and FC have been involved in drafting the manuscript and revising it critically for important intellectual content.

All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

- Rao CR. Prediction of Future Observations in Growth Curve Models. *Stat Sci*. 1987;2(4):434–47.
- Ji H, Müller H-G. Optimal designs for longitudinal and functional data. *J R Stat Soc Ser B Stat Methodol*. 2017;79(3):859–76.
- Functional Data Analysis | James Ramsay | Springer [Internet]. [cited 2017 May 29]. Available from: <http://www.springer.com/br/book/9780387400808>
- Kennedy J, Eberhart R. Particle swarm optimization. *Proc 1995 IEEE Int Conf Neural Netw*. 1995;1:1942–8.
- Parsopoulos K, Vrahatis M. Particle swarm optimization method for constrained optimization problems. *Front Artif Intell Appl*. 2002;76:214–20.
- Pedersen M, Chipperfield A. Simplifying particle swarm optimization. *Appl Soft Comput*. 2010;10:618–28.
- Shi Y, Eberhart R. A modified particle swarm optimizer. *Proc 1998 IEEE Int Conf Evol Comput*. 1998;1:69–73.
- Fasshauer GE, McCourt MJ. *Kernel-based Approximation Methods using Matlab*. Singapore: World Scientific; 2015.
- Cavoretto R, De Rossi A, Perracchione E. Optimal Selection of Local Approximants in RBF-PU Interpolation. 2017.
- Cavoretto R, De Rossi A, Qiao H. Topology analysis of global and local RBF transformations for image registration. *Math Comput Simul*. 2018 May 1;147:52–72.
- Cancelliere R, Gai M, Gallinari P, Rubini L. OCREP: An Optimally Conditioned Regularization for pseudoinversion based neural training. *Neural Netw*. 2015;71:76–87.
- Giozzi A, Guiot C, Delsanto P, Iordache D. A novel approach to the analysis of human growth. *Theor Biol Med Model*. 2012;9(1):1–16.
- Gompertz B. On the nature of the function expressive of the law of human mortality and a new mode of determining life contingencies. *Phil Trans R Soc*. 1825;115:513–585.
- Stura I, Venturino E, Guiot C. A two-clones tumor model: Spontaneous growth and response to treatment. *Math Biosci*. 2015;10(1):19–28.
- Stura I, Gabriele D, Guiot C. A two population model of cancer growth with fixed capacity. In: *Proceedings of the 2014 6th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation*. 2014.
- Migliaretti G, Ditaranto S, Guiot C, Vannelli S, Matarazzo P, Cappello N, et al. Long-term response to recombinant human growth hormone treatment: a new predictive mathematical method. 2018.
- Qasem S, Shamsuddin S. Radial basis function network based on time variant multi-objective particle swarm optimization for medical diseases diagnosis. *Appl Soft Comput*. 2011;11:1427–38.
- Yang X, Deb S. Cuckoo search via Lévy flights. In: *2009 World Congress on Nature and Biologically Inspired Computing, NABIC 2009 - Proceedings*. 2009.
- Wendland H. *Scattered Data Approximation*. Cambridge Univ. Press, Cambridge: Cambridge Monogr. Appl. Comput. Math.; 2005.
- Migliaretti G, Berchiolla P, Borraccino A, Gregori D, Cavallo F, Commissione GH Piemonte. A mathematical model in the analysis of the response to growth hormone treatment in pediatric patients with diagnosis of growth hormone deficiency. *J Endocrinol Invest*. 2012 Feb;35(2):209–14.
- Gabriele D, Porpiglia F, Muto G, Gontero P, Terrone C, Annoscia S, et al. Eureka-1 Database: an Epidemiological Analysis. *Minerva Urol Nefrol*. 2015;1(1):9–15.
- Perracchione E, Stura I. An RBF-PSO based approach for modeling prostate cancer. In 2016.
- Guiot C, Degiorgis PG, Delsanto P, Gabriele P, Deisboeck T. Does Tumor growth follow a “universal law”? *J Theor Biol*. 2003;225:147–51.
- Retsky M, Swartzendruber D, Wardwell R, Bame P. Is Gompertzian or exponential kinetics a valid description of individual human cancer growth? *Med Hypotheses*. 1990;33:95–106.

25. von Bertalanffy L. Quantitative laws in metabolism and growth. *Q Rev Biol.* 1957;32:217–231.
26. West GB, Brown JH, Enquist BJ. A general model for ontogenetic growth. *Nature.* 2001;413:628–31.
27. Deisboeck TS, Mansury Y, Guiot C, Degiorgis PG, Delsanto PP. Insights from a novel tumor model: Indications for a quantitative link between tumor growth and invasion. *Med Hypotheses.* 2005;65(4):785–90.
28. Ramsay JO, Silverman BW. *Applied Functional Data Analysis: Methods and Case Studies.* Springer New York; 2007. 194 p.

