Prediction of time series using wavelet Gaussian process for wireless sensor networks

Jose Mejia, Alberto Ochoa-Zezzatti, Oliverio Cruz-Mejía & Boris Mederos

Wireless Networks
The Journal of Mobile Communication, Computation and Information

ISSN 1022-0038

Wireless Netw
DOI 10.1007/s11276-020-02250-1

Springer

Wireless Networks
https://doi.org/10.1007/s11276-020-02250-1

Prediction of time series using wavelet Gaussian process for wireless sensor networks

Jose Mejia[1] · Alberto Ochoa-Zezzatti[1] · Oliverio Cruz-Mejía[2] · Boris Mederos[1]

Abstract
The detection and transmission of a physical variable over time, by a node of a sensor network to its sink node, represents a significant communication overload and consequently one of the main energy consumption processes. In this article we present an algorithm for the prediction of time series, with which it is expected to reduce the energy consumption of a sensor network, by reducing the number of transmissions when reporting to the sink node only when the prediction of the sensed value differs in certain magnitude, to the actual sensed value. For this end, the proposed algorithm combines a wavelet multiresolution transform with robust prediction using Gaussian process. The data is processed in wavelet domain, taking advantage of the transform ability to capture geometric information and decomposition in more simple signals or subbands. Subsequently, the decomposed signal is approximated by Gaussian process one for each subband of the wavelet, in this manner the Gaussian process is given to learn a much simple signal. Once the process is trained, it is ready to make predictions. We compare our method with pure Gaussian process prediction showing that the proposed method reduces the prediction error and is improves large horizons predictions, thus reducing the energy consumption of the sensor network.

Keywords Sensor networks · Time series · Gaussian process

1 Introduction

Sensor networks generated time series are increasingly significant for emerging applications that analyze this data, however, the acquisition for long periods of time depends on the network making a proper management of its energy resources. One of the main sources of energy expenditure occurs in the transmission of data within the network [1, 2]. Efforts to reduce radio transmissions communication include data aggregation [3, 4] and data reduction via prediction of the sensed magnitude [5], the latter consists in the use of algorithms to analyze and predict the sensed magnitude, in this way if the prediction is within a certain range of error, the data is not transmitted, allowing energy savings. This scheme depends on the model used to predict the time series. The analysis of time series is an important area of research in general, in the last decades there has been a growing activity in trying to develop and improve time series forecast models [6].

Time series prediction on sensor networks has been analyzed with different statistical methods, from classical prediction methods such as autoregressive moving averages (ARMA) and integrated autoregressive moving averages [7], Kalman filters [8] to deep learning LSTM [9]. The latter has taken a great boom with the introduction of deep learning, especially the use of extensive short-term memory networks [10], however despite the success using deep learning methods, for most sensor networks, the computational cost is still prohibitive and prediction with classical methods is still applicable and desirable in comparison to deep learning methods that generally require a large amount of training data, and increase complexity [11]. Here we are interested in probabilistic modeling such as Gaussian processes [12], these methods were applied to sensor networks, for deciding when and which sensor to acquire reading [13] and modeling for sensor localization [14]. Also, Gaussian processes have been applied to a wide

✉ Oliverio Cruz-Mejía
ocruzm@uaemex.mx

[1] Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Mexico

[2] Universidad Autónoma del Estado de México, Nezahualcoyotl, Mexico

Published online: 08 January 2020

variety of problems outside the sensor networks context, such as noise reduction, image super resolution and classification, in addition, they have also been used in the problem of time series prediction [15], for example, in [12] the authors consider the problem of predicting multiple steps in time series using a non-parametric Gaussian process model. The method is focused on minimizing the variance of error in the prediction in long horizons and noisy inputs, this is achieved by adapting the variance each time a prediction is made on a sample; in [16] the adaptive control of a domestic heating system is addressed in order to minimize costs and carbon emissions within an intelligent network, Gaussian processes are used to predict environmental parameters 24 h in advance, by On the other hand, [17] models based on Gaussian processes for forecasting electric consumption is examined. The work in [18] used a multiresolution model developed in the context of spatial modeling based on regression ideas with Kriging and Markov random fields, the authors propose a sum of independent processes at different scales to model a family of larger processes, however it is not suited for application in sensor networks.

In this work, we contemplate a new approach for the application of Gaussian processes to time series forecasting in a sensor network that implements the dual prediction scheme explained in [1]. In this context, the sensor network samples each observation with a specified accuracy, consequently a data is accepted by a sensor when it lies within an error bound (ε). Specifically, under this scheme a sensor node can use the time series prediction model to transmit a sample, x, when the prediction is inside the interval [x− ε, x + ε], otherwise nothing is transmitted. In the case that the sample is not transmitted the sink node uses the predicted value, here is assumed that the model is known by the entire network.

The approach that we propose to carry out predictions in a sensor node is the Gaussian processes technique as in [12, 18], but under a new domain without using Markov modeling, this type of analysis has been proposed also in [19, 20] using neural networks and ARMA, under a simple multiresolution decomposition of two levels with basic *Haar* filters of two coefficients, in this work we extend the application to an arbitrary number of levels and using any filter with multiresolution analysis capability, in addition, we consider the application of Gaussian processes at each level using a kernel that is a combination of radial and exponential basis functions, for easy implementation in sensor networks. One of the advantages of using Gaussian process is that they can give a reliable estimate of the uncertainty of the predicted value, this can be used to decide whether or not the sample will be transmitted.

2 Multiresolution and Gaussian processes

This section offers a brief introduction to the wavelet transform and the theory of Gaussian processes. For a more detailed treatment of these topics, the reader can consult [19, 21] and [22] for the topic of wavelets and GP, respectively.

2.1 Wavelet transform

The wavelet transform, $W\{.\}$, consists of a decomposition of a continuous signal, $f(x)$, using basis functions, ψ_{mn}. This family of bases, is obtained by translations, n, and dilations, m, of a specified basis function, $\psi(x)$, known as mother wavelet. Thus, any basis function is specified as [23]

$$\psi_{mn}(x) = 2^{\frac{m}{n}}\psi(2^m x - n) \tag{1}$$

where m and n, are integers that specify translations and dilations of the mother wavelet function.

An important feature of the mother wavelet is that it can be constructed from a scaling function, $\phi(x)$, that meets the property

$$\phi(x) = \sqrt{2}\sum_{l=-\infty}^{\infty} h(l)\phi(2x - l) \tag{2}$$

where $h(l)$ are scalar factors [19, 23, 24], which are the coefficients of some filter h. Using the scaling function, it is possible to express $\psi(x)$ as

$$\psi(x) = \sqrt{2}\sum_{l=-\infty}^{\infty} (-1)^l h(1 - l)\phi(2x - l). \tag{3}$$

It is possible to generate the decomposition of a signal without explicitly using the wavelet basis functions $\psi(x)$, this is done through the coefficients h, which define digital filters that can be used to represent the wavelet transform of discrete signals [25], in this case, the wavelet transform in discrete time, for a discrete signal $f(m)$, is defined as a series of subband signals y_k of a filter bank given by

$$y_k(n) = \sum_{m=-\infty}^{\infty} f(m)h_k\left(2^{k+1}n - m\right) \quad k = 1, \ldots, K \tag{4}$$

Where, K is the number of levels of the transform and the filter h is related to h_k in the Fourier domain as

$$H_k\left(e^{j\omega}\right) = H\left(e^{j2^k\omega}\right) \tag{5}$$

where $H(\cdot)$ is the Fourier transform of $h(\cdot)$.

2.2 Gaussian process

A random process $f(x)$ is a Gaussian process if any vector $(f(x_1), f(x_2), \ldots, f(x_m))$ at finite number of points $x_1, x_2, \ldots x_m$ has a multivariate normal distribution. The Gaussian process is fully characterized by its mean $m(x)$ and a covariance function $c(x, x')$ [22].

So, given a sequence of points $x_1, \ldots, x_N$, the sequence of the values of f, evaluated at those points, $f = \{f_1, \ldots, f_N\}$, is distributed as a multivariate Gaussian, that is

$$\boldsymbol{f} \sim N(m, C) \tag{6}$$

where $C = \left(c\left(x_i, x_j\right)\right)$ is the covariance matrix and m is the mean vector.

It is possible to use Gaussian processes to predict a value, $f^*(x)$, at a new point at location x^*, through the knowledge of noisy observations $y = (y_1, y_2, \ldots, y_m)$, which are relate to GP by the expression $y = f + w$, where $w \sim N(0, \sigma I)$ is Gaussian noise.

Under these circumstances, the calculation of new predictions is made through a predictive distribution p (f* | y) which turns out to be normally distributed, making it possible to obtain a solution in analytical form of the mean and variance at the prediction point x^* as follows

$$m(x^*) = c(X, x^*)^T \left(C + \sigma^2 I\right)^{-1} y \tag{7}$$

$$var(x^*) = c(x^*, x^*) - c(X, x^*)^T \left(C - \sigma^2 I\right)^{-1} c(X, x^*) \tag{8}$$

where $c(X, x^*)^T = (c(x_1, x^*), c(x_2, x^*), \ldots, c(x_m, x^*))^T$.

The Gaussian process covariance function determines many of the characteristics that the process will have and is generally selected during the observation of the data.

3 Methodology

It is common that in certain applications of time series or signals, the samples represent a complex sequence difficult to model, so it is necessary to resort to statistical models. Gaussian processes, as seen in Sect. 2.2, can provide a model that is easily adaptable to a large class of practical signals. However, due to the complex forms that the signal can acquire, it is in general difficult to adapt a covariance function in these types of signals. In the present work it is proposed to transform the signal to be analyzed to the wavelet domain and then adapt a Gaussian process to each of the subbands of the decomposition. Using this approach, it is expected that the representation of the signal in each subband has of less complexity and therefore it will be much easier to adapt the covariance function of the Gaussian process. Thus, our proposed method has the following advantages:

- It is adaptable to specific signal by estimating a covariance function for that signal.
- It can handle variable complexity of the signals by incrementing the level of decomposition of the wavelet transform.
- Our method can give a reliable estimate of the uncertainty of the predicted value, this can be used to decide whether or not the sample will be transmitted given the error bound (ε) of a dual prediction scheme [1].
- As can be seen from the experiments the proposed method has less error than the other methods with which the comparison was made, thus using the dual prediction scheme [1], the amount of data transmitted over the network would decrease as compared to the other methods.

3.1 Prediction model

The first step of the proposed model is to transform the signal of interest to the wavelet domain, a series or discrete signal *f(m)* will be assumed to have a wavelet representation given by a collection of subbands y_k. The number of subbands, K, will depend on the complexity of the signal. The determination of the optimal number of levels can be performed empirically or by measuring the error between the Gaussian process model and the original signal.

A Gaussian process is then adjusted to each subband using a suitable kernel or covariance. In this work we assume that the subbands resulting from the transformation consist of collections of random variables, of which any finite set has a multivariate Gaussian joint distribution,

$$y_k(1), \ldots, y_k(N_k) \sim N(0, C_k) \quad k = 1, \ldots, K \quad n = 1, \ldots, N_k \tag{9}$$

where N_k is the number of elements in subband k and C_k is the specific covariance matrix at level k, here each level or subband defines a different Gaussian process.

The prediction sought is forward in time to a given horizon, that is, to predict values using the last known value of the signal. So the next step is to adjust the number of decomposition coefficients to the number of desired samples to predict for this, the following formula adapted from [24] is used:

$$N_{k+1} = floor\left(\frac{N_k - 1}{2}\right) + \frac{N_f}{2} \tag{10}$$

where N_k is the length of the subband adjusted to the number of samples to predict with N_0 the length of the signal to be predicted, N_f is the size of the subband filter k.

Next, the necessary points in the future are predicted for each subband using their respective Gaussian processes and

finally the subbands are transformed with the inverse wavelet transform, to obtain a time domain signal with an expanded horizon.

4 Results

This section presents the results of a series of experiments focused on the validation and comparison of the proposed algorithm. All experiments were performed on a computer with microprocessor i7-6500U CPU @ 2.50 GHz, with 8 G of RAM.

For a first experiment, data come from the database of [26]. This database contains 9358 samples of average hourly responses from five chemical metal oxide sensors embedded in an air quality sensor device. The device was located in a significantly contaminated area of an Italian city. Registration data began in March 2004 until February 2005. For more details regarding the database, the reader can consult [26].

The experiment consists of a comparison with a pure Gaussian processes (GP), the implementation of the algorithms was made in python language with the library of [27]. In order to make a better comparison, the same kernel, radial basis function (RBF) with exponential square sinusoidal was used, in both implementations, the proposed method and the algorithm using only GP:

$$kernel(x_i, x_j) = \frac{1}{2} RBF(x_i, x_j, pScala) + \exp\left(-2\left(\frac{\sin\left(\frac{\pi * abs(x_i, x_j)}{pPeriod}\right)}{pScala}\right)^2\right) \quad (11)$$

Where the *pScala* and *pPeriod* parameters establish the scale and period of the process, this was automatically established by optimization algorithms provided with [27]. The metric used to quantitatively evaluate the results was the root mean square error (RMSE), defined as:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (\hat{y}(t) - y(t))^2}{n}} \quad (12)$$

Where n is the number of samples, y(t) is the original signal and $\hat{y}(t)$ is the estimated sample.

To select the number of decomposition levels of the wavelet we observed the effect that decomposition levels have on the prediction error. From the database of [26], 329 samples were taken, of these samples 290 samples were taken for training. A prediction horizon of 39 samples was used. Once data was selected, the wavelet transform is used for its analysis and prediction. This process was repeated using $J = 2,3,4$ and 5 levels of decomposition with the wavelet transform. For each level selection the prediction error was calculated using the RMSE, the results can be

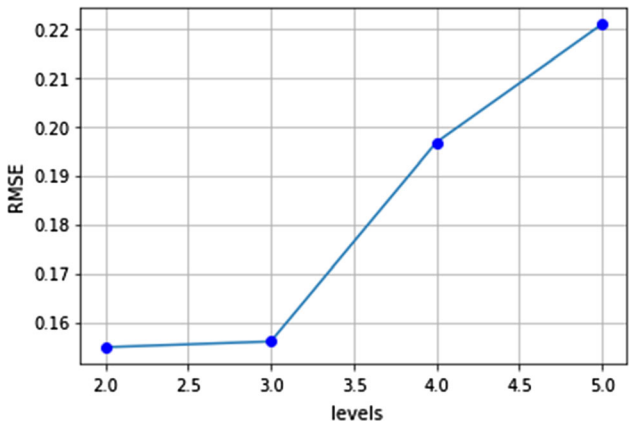

Fig. 1 Prediction error of the proposed method using 1, 2, 3, 4, and 5 levels of decomposition with wavelet transform

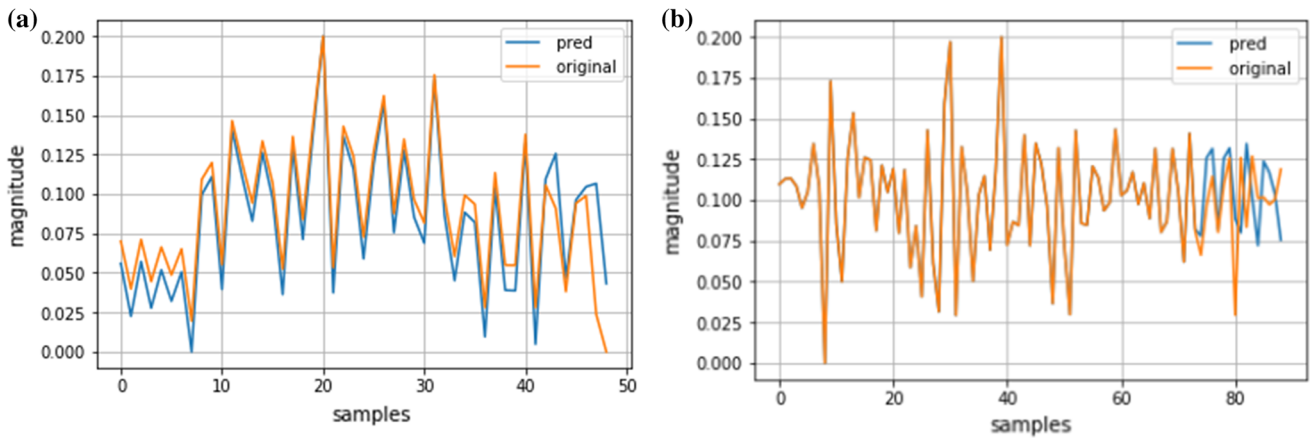

Fig. 2 Decomposition with three-level wavelet transform. **a** Prediction in the aproximation subband, **b** prediction in the last level, detail subband

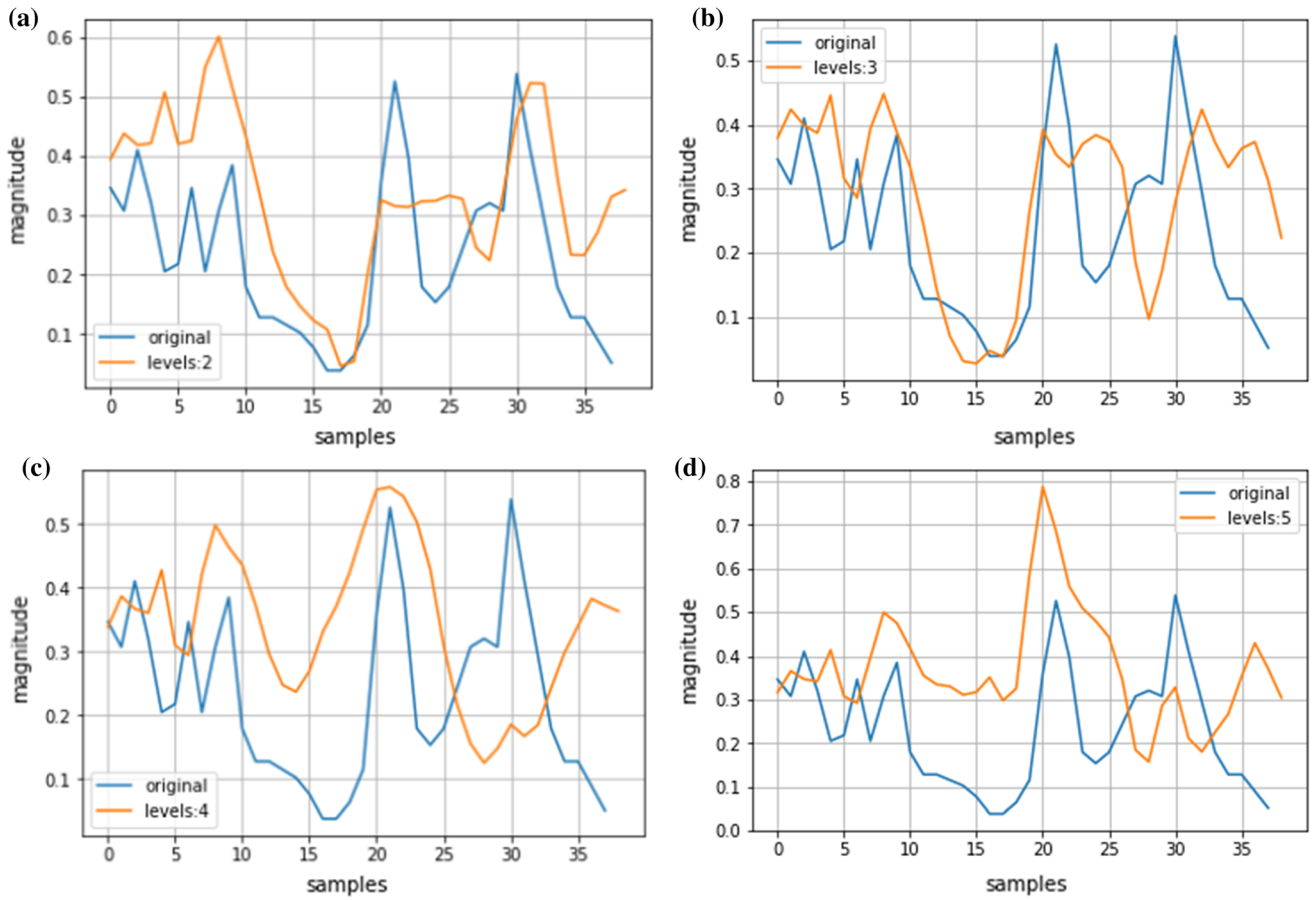

Fig. 3 Data predicted with the proposed method using: **a** two levels, **b** three levels, **c** four levels and **d** five levels

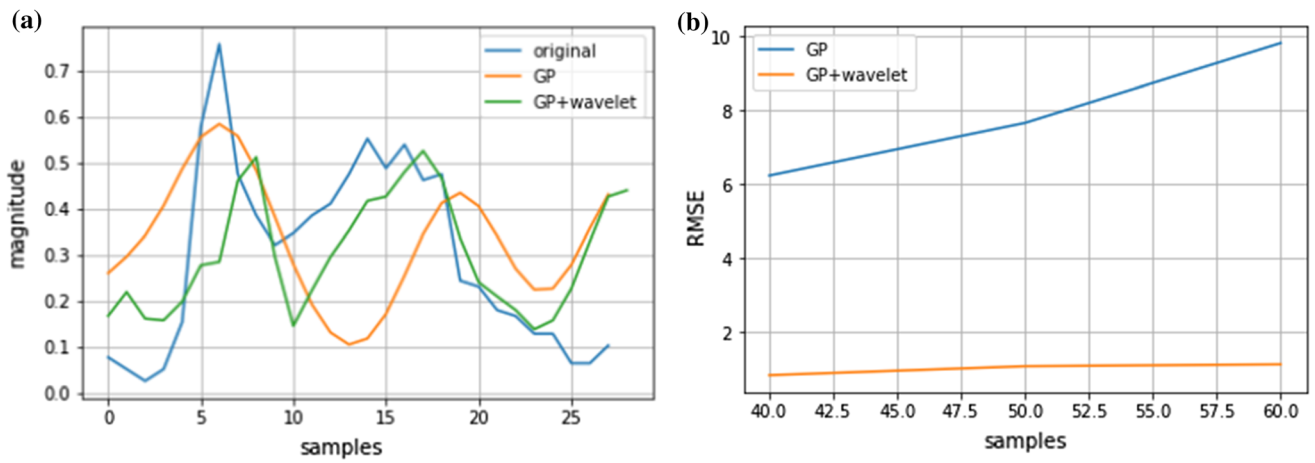

Fig. 4 **a** Prediction with a horizon of 27 samples. **b** Prediction error with different prediction horizons

seen in the graph of Fig. 1, it can be seen that the RMSE increases as the number of levels used increases, this may be due to different causes, One of them could be that by increasing the number of levels, the information in each subband decreases what could adversely affect the training of the Gaussian process causing a deterioration in the prediction.

In Fig. 2, the prediction in the last band is shown, using a three-level wavelet transformation (J = 3). Figure 2b shows a prediction error in the last samples due to a poor

optimization of the Gaussian process parameters in that subband.

In Fig. 3, the predicted data are shown against the original data for the reconstructed data from the wavelet transform resulting from the predictions with the proposed method.

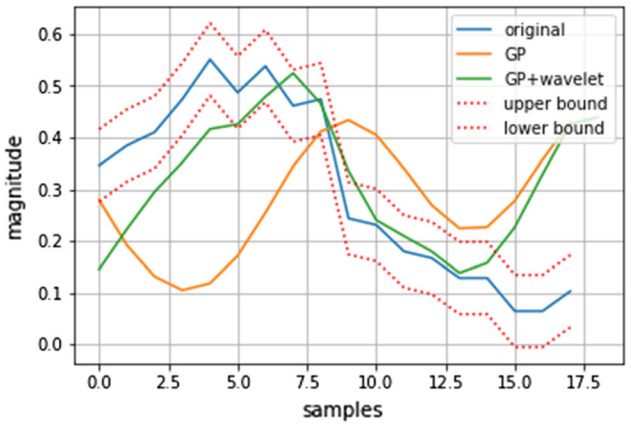

Fig. 5 Signal with bounds $\pm\ \varepsilon = 0.05$, in zones where the predicted signal is within the bound, the samples are not transmitted

Next, we compare the proposed algorithm with prediction using only Gaussian processes (GP), a prediction horizon of 27 samples were used, for the case of the proposed method a two-level wavelet decomposition was selected. Prediction was done using a window of five samples, results are shown in Fig. 4a, where it can be seen that the proposed method (GP + wavelet) exceeds the method using GP directly, this is shown quantitatively in Fig. 4b where it is shown a graph of the prediction error as the samples to be predicted increase. Figure 5, shows the original signal an a bound of $\pm\ \varepsilon = 0.05$, it can be seen that the proposed model has more intervals where the prediction is within bound, than using GP alone, meaning that under the dual prediction scheme [1], using the proposed method there could be more energy savings than using GP alone.

For the next experiment, a more complex dataset was used, it consists of a time series of greenhouse gas concentrations measured by grid cells. Data was created using simulations of the Weather Research and Forecast model with Chemistry, for details see [28]. Comparisons were made with the following methods: a neural network (NN) of five layers with 6, 3, 2 and 1 neurons with rectified linear

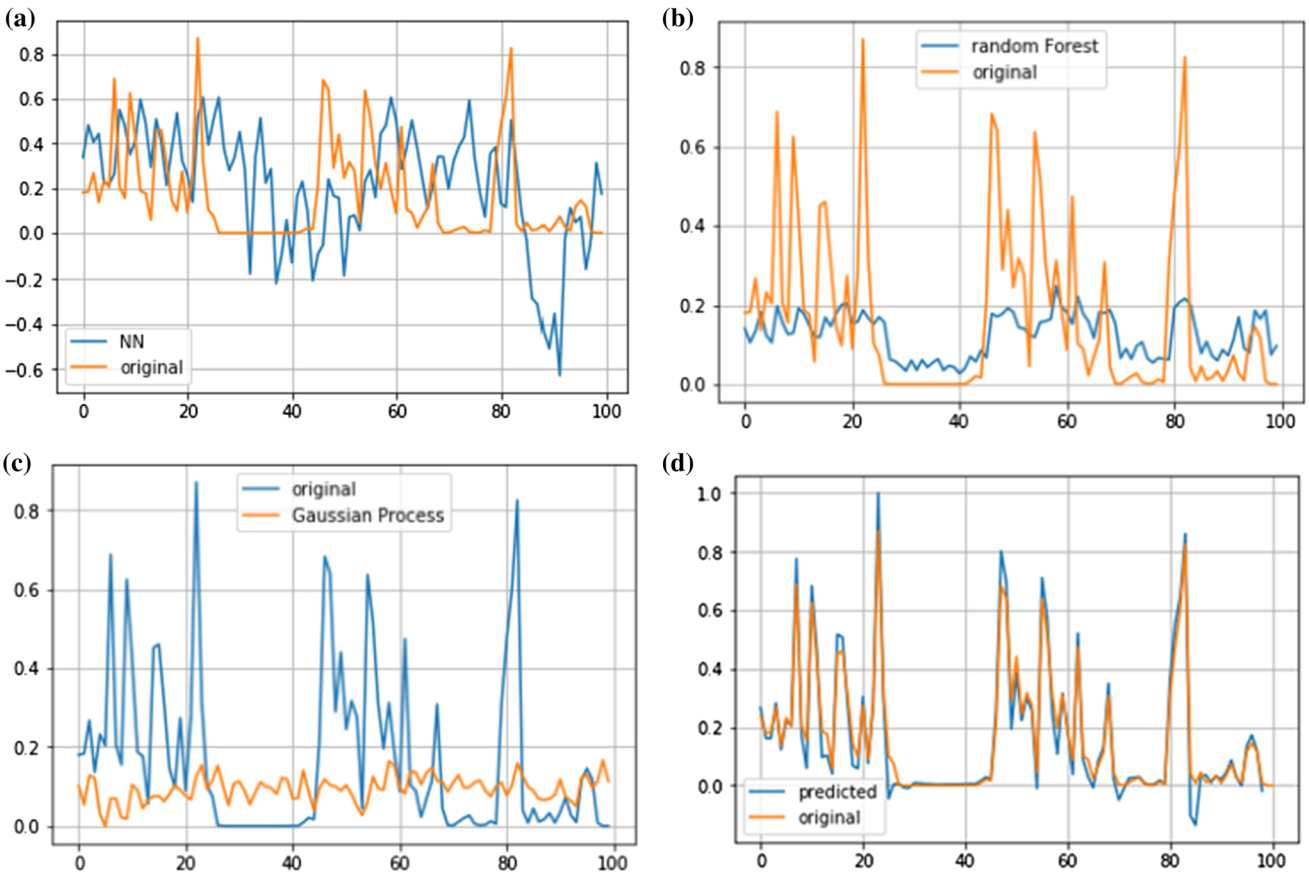

Fig. 6 The Greenhouse Gas Observing Network Data Set [28]. Prediction with **a** NN, RMSE = 0.302, **b** random forest, RMSE = 0.188, **c** GP, RMSE = 0.223, and **d** proposed method, RMSE = 0.047

activation function, except the last layer which has a linear activation; and a random forest with 220 trees. Training for all methods was done with 2000 samples, for the NN 600 epoch were used and for GP 19 iterations. Predictions were made using windows of 100. Figure 6a, b, show the prediction using NN and random forest respectively, the random forest was unable to follow the abrupt changes that the series presented. Figure 6c shows the GP prediction, here the GP has the same parameters as the proposed method (Fig. 6d), however it was unable to adapt to changes in the signal, there may be other kernels that are better suited to the signal, however, one of the purposes of the experiment is to show the advantage of using GP in the wavelet domain, without a more accurate analysis of the types of kernels.

In Fig. 6c, it is shown the results of the proposed method, despite the greater complexity of the series, the larger number of samples and the use of a larger prediction window than the one used in the previous experiment, helped to have a prediction quite accurately compared to the other methods presented. Although in general, using a larger number of samples for the prediction, significantly increases the accuracy, the training of the estimators, including the proposed method, becomes computationally prohibitive for low-end devices, however since the training is done by the central node, a device with more computing power for that node can be used, such as a small single board computer.

5 Conclusions

An algorithm for signal prediction to operate under the dual prediction scheme was presented. The proposed method makes use of the wavelet transform and Gaussian processes to obtain estimates of the signal. The algorithm first transforms the signal to the wavelet domain and the prediction by Gaussian processes is made in each subband. The signal in each subband is expected to be easier to analyze and predict by Gaussian processes than if the entire signal was taken in its entirety, this was more evident when using the method in complex series, as is shown in the last experiment where the proposed method was able to adapt to the series much more precisely than using GP without a multi-resolution analysis. The method is compared with prediction using Gaussian processes, NN, and random forest on the original signal resulting that the proposed algorithm obtains better prediction results, based on the RMS measured with the test signal. As a future work in the future, we plan to analyze the algorithm behavior using other multiresolution transforms apart from wavelets and try different decomposition filters. Also, an implementation of the algorithm in a physical sensor network.

Acknowledgements Funding was provided by Sistema Nacional de Investigadores, CONACYT, México.

References

1. Le Borgne, Y.-A., Santini, S., & Bontempi, G. (2007). Adaptive model selection for time series prediction in wireless sensor networks. *Signal Processing, 87*(12), 3010–3020.
2. Lazaridis, I., & Mehrotra, S. (2003). Capturing sensor-generated time series with quality guarantees. In *Proceedings 19th international conference on data engineering* (Cat. No. 03CH37405). IEEE.
3. Madden, S., Franklin, M. J., Hellerstein, J. M., & Hong, W. (2002). Tag: a tiny aggregation service for ad-hoc sensor networks. *SIGOPS Operating Systems Review, 36*(SI), 131–146.
4. Nath, S., Gibbons, P. B., Seshan, S., & Anderson, Z. R. (2004). Synopsis diffusion for robust aggregation in sensor networks. In *SenSys*.
5. Goel, S., & Imielinski, T. (2001). Prediction-based monitoring in sensor networks: Taking lessons from MPEG. *SIGCOMM Computer Communication Review, 31*(5), 82–98.
6. Längkvist, M., Karlsson, L., & Loutfi, A. (2014). A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognition Letters, 42,* 11–24.
7. Tulone, D., & Madden, S. (2006). PAQ: Time series forecasting for approximate query answering in sensor networks. In *European workshop on wireless sensor networks*. Berlin: Springer.
8. Ji, H., et al. (2017). Distributed information-weighted Kalman consensus filter for sensor networks. *Automatica, 77,* 18–30.
9. Ma, X., et al. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies, 54,* 187–197.
10. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780.
11. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems* (pp. 1097–1105).
12. Girard, A., Rasmussen, C. E., Candela, J. Q., & Murray-Smith, R. (2003). Gaussian process priors with uncertain inputs application to multiple-step ahead time series forecasting. In *Advances in neural information processing systems* (pp. 545–552).
13. Osborne, M. A., et al. (2008). Towards real-time information processing of sensor network data using computationally efficient multi-output Gaussian processes. In *2008 international conference on information processing in sensor networks (IPSN 2008)*. IEEE.
14. Richter, P., & Toledano-Ayala, M. (2015). Revisiting Gaussian process regression modeling for localization in wireless sensor networks. *Sensors, 15*(9), 22587–22615.
15. Roberts, S., Osborne, M., Ebden, M., Reece, S., Gibson, N., & Aigrain, S. (2013). Gaussian processes for time-series modelling. *Philosophical Transactions of the Royal Society A, 371,* 20110550.
16. Rogers, A., Maleki, S., Ghosh, S., and Jennings, N. R. (2011). Adaptive home heating control through Gaussian process prediction and mathematical programming. eprints.soton.ac.uk. Nychka, D.
17. Leith, D. J., Heidl, M., & Ringwood, J. V. (2004). Gaussian process prior models for electrical load forecasting. In *International conference on probabilistic methods applied to power systems* (pp. 112–117).

18. Bandyopadhyay, S., Hammerling, D., Lindgren, F., & Sain, S. (2015). A multiresolution Gaussian process model for the analysis of large spatial datasets. *Journal of Computational and Graphical Statistics, 24*(2), 579–599.
19. Percival, D. B., & Walden, A. T. (2006). *Wavelet methods for time series analysis* (Vol. 4). Cambridge: Cambridge University Press.
20. Soltani, S. (2002). On the use of the wavelet decomposition for time series prediction. *Neurocomputing, 48*(1), 267–277.
21. Strang, G., & Nguyen, T. (1996). *Wavelets and filter banks*. Cambridge: SIAM.
22. Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian processes for machine learning*. Cambridge: MIT Press.
23. Chang, T., & Kuo, C. C. J. (1993). texture analysis and classification with tree-structured wavelet tranform. *IEEE Transactions on Image Processing, 2*(4), 429–441.
24. Mallat, S. (1999). *A wavelet tour of signal processing*. Cambridge: Academic Press.
25. Vaidyanathan, P. P. (1993). *Multirate systems and filter banks*. Upper Saddle River: Prentice Hall.
26. De Vito, S., Massera, E., Piga, M., Martinotto, L., & Di Francia, G. (2008). On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario. *Sensors and Actuators B: Chemical, 129*(2), 750–757.
27. Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., et al. (2013). Design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD workshop: Languages for data mining and machine learning* (pp. 108–122).
28. Lucas, D. D., Yver Kwok, C., Cameron-Smith, P., Graven, H., Bergmann, D., Guilderson, T. P., et al. (2015). Designing optimal greenhouse gas observing networks that consider performance and cost. *Geoscientific Instrumentation, Methods and Data Systems, 4*(1), 121–137.

Jose Mejia received the M.Sc. degree in electrical engineering and a Ph.D. in Engineering from the Universidad Autonoma de Ciudad Juarez (UACJ), Mexico. He has also worked for the consumer electronics and automotive industry as an electronic designer. His research interests include sensor networks, medical image reconstruction, and game theory. He is currently a Research Professor in the Department of Electrical and Computation Engineering at the UACJ.

Alberto Ochoa-Zezzatti participates in the organization of several International Conferences. He has review of four important journals from highlighting Applied Soft Computing and Computers in Human Behavior. His research interests include ubiquitous compute, evolutionary computation, natural processing language, social modeling, anthropometrics characterization and Social Data Mining. In May 2016 begin a Sabbatical internship year in Barcelona Supercomputing Center with a project related with simulation of human stampedes and refugee boat sinking improved with Artificial Intelligence. Actually has Level 2 in SNI membership at Conacyt. He has a project with European Union specially with Genoveva Vargas (France) related with Smart Cities.

Oliverio Cruz-Mejía is a Professor of Industrial Engineering at the Universidad Autónoma del Estado de México, México. He holds a Ph D in Management Science form the Lancaster University, UK, a master's degree in Manufacturing Engineering from the University of Syracuse, USA and a B. S. in Mechanical Engineering from National Polytechnic Institute of Mexico. He is board review member of the Journal of Supply Chain Management (JSCMP). He has published in the Journal of Physical Distribution and Logistics Management (JPDLM), Mathematical Problems in Engineering (MPE), Journal of Ambient Intelligence and Humanized Computing (JAIHC), International Journal of Applied Metaheuristic Computing and Nova Scientia. Oliverio has contributed to book chapters on Simulation for Industry 4.0 (Springer) and Order Picking Performance in Logistics (IGI Global).

Boris Mederos has a Ph.D. in Applied Mathematics from the Instituto de Matemática Pura e Aplicada (IMPA), Brazil. He is currently a Full-Time Professor at the Autonomous University of Ciudad Juárez, México. His research interests include medical image processing, machine learning, calculus of variation, optimization and numerical methods.