



MULTI-OUTPUT GAUSSIAN PROCESSES APPLIED TO MODELING AND REDUCING DIMENSIONALITY OF DATA TRANSMITTED IN A WIRELESS SENSOR NETWORK

Procesos gaussianos multi-salida aplicados al modelamiento y reducción en la dimensionalidad de los datos transmitidos en una red inalámbrica de sensores

José-Aicardo Ortega-Vela 
Universidad del Quindío (Armenia, Colombia). 
jaortega@uniquindio.edu.co

Pablo-Andrés Muñoz-Gutiérrez 
Universidad del Quindío (Armenia, Colombia). 
pamunoz@uniquindio.edu.co

Hernán-Darío Vargas-Cardona 
Pontificia Universidad Javeriana de Cali (Cali, Colombia). 
hernan.vargas@javerianacali.edu.co

Received: 26-09-2024

Accepted: 12-01-2025



ABSTRACT

This work presents data modelling from Wireless Sensor Networks (WSN) applying Multiple Output Gaussian Processes (MOGP). The objective, in addition to describing the dynamics of sensed magnitudes (temperature, relative humidity, atmospheric pressure, and soil humidity), is to exploit the probabilistic learning nature of MOGPs and use the variance (uncertainty) provided by this method to reduce dimensionality in the network data, meaning the elimination of redundant data. It is compared with other machine learning models such as simple Gaussian Processes (GP), Support Vector Regressor (SVR), Neural Networks (NN), and Random Forest (RF), in terms of Root Mean Squared Error (RMSE) for modeling real data sensed by the WSN located at Universidad del Quindío. Results demonstrate that MOGPs are highly accurate supervised learning algorithms, flexible in modeling any physical magnitude, and capable of detecting redundant data, in some cases achieving a reduction greater than 50 %.

Keywords: data processing; dimensionality reduction; Multiple Output Gaussian Processes (MOGP); resource optimization; supervised learning; Wireless Sensor Networks (WSN).

RESUMEN

En este trabajo se presenta el modelamiento de datos provenientes de redes de sensores inalámbricos (WSN), aplicando procesos Gaussianos de múltiples salidas (MOGP). El objetivo, además de describir la dinámica de las magnitudes medidas (temperatura, humedad relativa, presión atmosférica y humedad de suelo), es aprovechar que los MOGP son máquinas de aprendizaje probabilísticas y usar la varianza (incertidumbre) que entrega este método para reducir la dimensionalidad en los datos de la red, es decir, eliminar datos redundantes. Se hace una comparación con otras máquinas de aprendizaje como procesos Gaussianos simples (GP), regresor basado en máquina de soporte vectorial (SVR), redes neuronales (NN), y *Random Forest* (RF), donde se evalúa el rendimiento en términos del RMSE, en el modelamiento de datos reales medidos por la WSN

How to cite: J. A. Ortega-Vela, P. A. Muñoz-Gutiérrez, and H. D. Vargas-Cardona, "Multi-Output Gaussian Processes Applied to Modeling and Reducing Dimensionality of Data Transmitted in a Wireless Sensor Network". *Revista Facultad de Ingeniería*, vol. 34, no. 71, e18219, 2025. <https://doi.org/10.19053/01211129.v34.n71.2025.18219>

ubicada en la Universidad del Quindío. Los resultados prueban que los MOGP son algoritmos de aprendizaje supervisado de alta exactitud, flexibles para modelar cualquier magnitud física y, además, detectan datos redundantes, en algunos casos con reducción mayor al 50%.

Palabras clave: aprendizaje supervisado; optimización de recursos; procesamiento de datos; procesos gaussianos de múltiples salidas; redes de sensores inalámbricos; reducción de dimensionalidad.

PROCESSOS GAUSSIANOS MULTI-SAÍDA APLICADOS AO MODELAMENTO E REDUÇÃO DE DIMENSIONALIDADE DOS DADOS TRANSMITIDOS EM UMA REDE SEM FIO DE SENSORES

RESUMO

Este trabalho apresenta o modelamento de dados provenientes de redes de sensores sem fio (WSN), aplicando processos Gaussianos de múltiplas saídas (MOGP). O objetivo, além de descrever a dinâmica das grandezas medidas (temperatura, umidade relativa, pressão atmosférica e umidade do solo), é aproveitar que os MOGP são máquinas de aprendizado probabilísticas e usar a variância (incerteza) fornecida por esse método para reduzir a dimensionalidade dos dados da rede, ou seja, eliminar dados redundantes. Realiza-se uma comparação com outras máquinas de aprendizado, como processos Gaussianos simples (GP), regressão baseada em máquina de vetor de suporte (SVR), redes neurais (NN) e Random Forest (RF), onde o desempenho é avaliado em termos do RMSE, no modelamento de dados reais medidos pela WSN localizada na Universidade do Quindío. Os resultados comprovam que os MOGP são algoritmos de aprendizado supervisionado de alta precisão, flexíveis para modelar qualquer grandeza física e, além disso, detectam dados redundantes, em alguns casos com redução superior a 50%.

Palavras-chave: aprendizado supervisionado; otimização de recursos, processamento de dados; processos gaussianos de múltiplas saídas; redes de sensores sem fio; redução de dimensionalidade.

1. INTRODUCTION

Wireless Sensor Networks (WSNs) are key for intelligent, real-time data collection monitoring. Their range of applications is vast; they contribute significantly to industries such as manufacturing, science, transportation, security, and infrastructure [1], [2]. While it is well-known that WSN systems are highly useful in various sectors and fields, there are unresolved design and operational issues related to the number of transmissions required by the nodes within the sensor network [3] or the amount of data that must be transmitted to ensure that there is not a significant loss of information [4]. Dimensionality reduction or data reduction in transmission involves key aspects such as redundancy, information rotation mode, speed, and sampling rate [3], [5]. Additionally, the most critical factors related to the high dimensionality of acquired data are excessive energy consumption, the large bandwidth required for data transmission, and the substantial hardware volume needed for storage [6], all of which limit communication range and processing power capabilities [7]. Therefore, it is evident that there are unresolved issues in WSNs regarding the high amount of redundant data acquired per time instance [1], [8]-[9].

Some latent issues in WSNs are, for example, environmental measurements made by each sensor can be misleading, meaning that the occurrence of events is not properly examined, which leads to careless use of the CPU (Central Processing Unit) [5]; the repetition of irrelevant information generated by different nodes significantly increases the amount of data transmitted across the sensor network [3]; methods for data compression are not applied to reduce the number of transmissions between nodes [6]; finally, transmitting large volumes of data results in excessive energy consumption, increased bandwidth usage, and greater storage hardware requirements, while significantly decreasing the speed at which the sensor network delivers the final data [10].

In recent years, methodologies for processing data from WSNs using data fusion theory have been proposed, e.g., the Internet of Things (IoT), fuzzy logic, information fusion, clustering, Bayesian networks, and the Kalman filter, among others [7]- [11]. These approaches allow the processed information to be communicated via messages to fusion-type agents and later analyzed for information quality, always supported by the mathematical theory of communication, thus resulting in a cooperative process provided by each sensor [3]- [5]. However, none of these methods emphasize mathematical modeling or the reduction of dimensionality. This represents a significant research gap when processing multichannel (multi-source) data, such as in WSNs.

Conversely, there has been limited research into the possibility of simultaneously modeling data from different sources (sensors). In other words, no studies or techniques have been developed that fit mathematical model to the time series data produced by a wireless sensor network [6]. Undoubtedly, this could open new avenues for research related to the description of the behavior of non-stationary variables, such as temperature, pressure, humidity, etc. The literature includes advanced methods based on machine learning theory like linear and nonlinear regression models [12], support vector machines [13], and unsupervised learning [14]. These methods could model data from individual sensors and perform dimensionality reduction through training with a reasonable amount of data. However, it would require training a number of learning models equivalent to the number of sensors in the network, which would be critical for computational complexity. In this case, a multitask learning method could be a viable alternative for this purpose, as a single model can mathematically describe different signals from the sensors.

This work presents a multi-output method as a viable alternative for this purpose, as a single model can mathematically describe different signals from sensors. Specifically, a probabilistic method based on multi-output Gaussian Processes (MOGP) is proposed; it offers significant advantages such as accuracy, robustness, and dimensionality reduction by achieving drift estimation with a mean squared error below 10%. Results obtained from a real WSN demonstrate that MOGPs are supervised learning algorithms with high accuracy and flexibility to model any physical magnitude, they outperform other machine learning techniques such as single-output Gaussian Processes, Support Vector Machines for Regression (SVR), and Random Forest (RF). Furthermore, MOGPs can detect redundant data, achieving reductions of over 50% in some cases.

The remainder of this article is organized as follows: Section 2 describes the materials and methods, Section 3 presents the results and discussion, and, Section 4 provides the conclusions of the study.

2. MATERIALS AND METHODS

A. Database

The real database consists of a set of 10 sequences of 75 or 100 data points (for each of the 4 sensors) acquired using the WSN located at Universidad del Quindío. The timing of the measurements is random (minutes, hours, or days). The original file, named *Datos.csv*, was processed to obtain the 10 sequences in *.mat* format, with the following names: *datosprueba1.mat*, *datosprueba2.mat*, ..., *datosprueba10.mat*. Each file contains variables **D**, **D2**, **D3**, and **D4**, which correspond to temperature (°C), relative humidity (%), atmospheric pressure (Bar), and soil humidity (%), respectively. Figure 1 shows the time series for sequence 1, where data align with standard values for temperature, relative humidity, atmospheric pressure, and soil humidity, considering that the altitude of the city of Armenia is approximately 1550 meters above sea level.

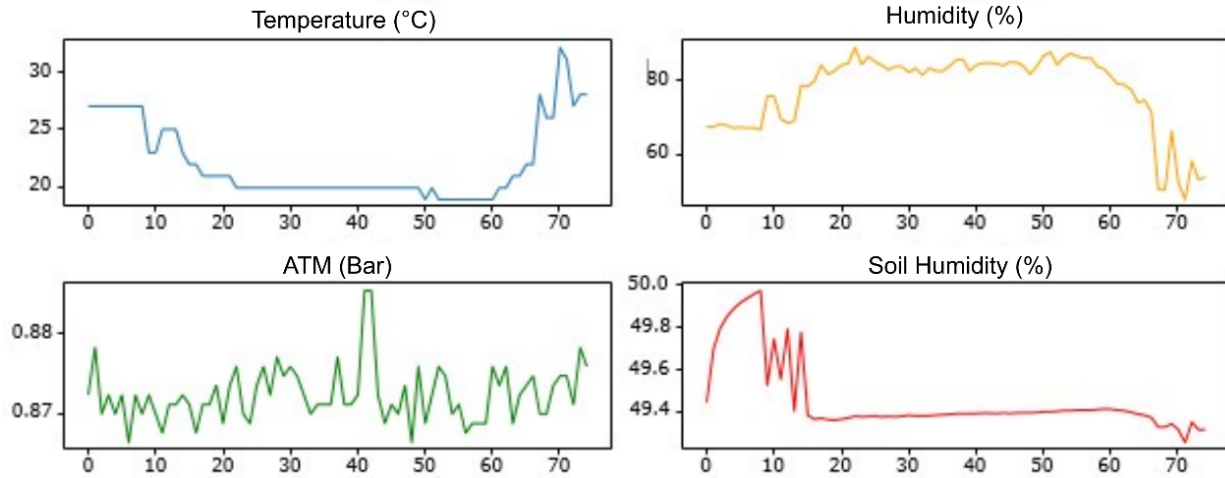


Figure 1. Data Frame 1

B. Wireless Sensor Network (WSN)

A wireless sensor network (WSN) is defined as a self-configurable network composed of a small number of sensor nodes, also known as motes, which are spatially distributed and interconnected. Recently, a new method has been developed to enhance data performance, achieving a reliability range sufficient for continuous monitoring in systems [15]. Additionally, WSNs serve as an important communication bridge between the virtual and the real physical world. Their wide range of applications makes them a significant contributor to fields such as industry, science, transportation, security, infrastructure, the military, and even commercial applications [16], [17].

C. Approach based on Multi-Output Gaussian Processes

1) Gaussian Process (GP). A Gaussian Process is a collection of random variables where a finite number of them follow a joint Gaussian distribution [18]. A GP is completely defined by its mean function $m(x)$, and its covariance function $k(x, x')$, such that $f(x) \sim GP(m(x), k(x, x'))$. In supervised learning, the radial basis kernel (RBF) is commonly employed as the covariance function, and is given by:

$$k(x, x') = \sigma^2 \exp\left(-\frac{\|x - x'\|^2}{2\theta^2}\right) \quad (1)$$

Where θ and σ^2 are the hyper-parameters of scale and variance, respectively. Consider a noisy data, indexed by an independent variable (may be time) and its related values $\{(x_i, y_i)\}_{i=1, \dots, n}$, where the y_i values follow the standard regression model $y_i = f_i + \varepsilon$, with $f_i = f(x_i)$ and a Gaussian noise $\varepsilon \sim N(0, \sigma^2 n)$. The joint distribution of the training values y , and the validation values, y^* , is given by:

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0^* \end{bmatrix}, \begin{bmatrix} K_{yy} + \sigma^2 n I & K_{yy^*} \\ K_{y^*y} & K_{y^*y^*} + \sigma^2 n I^* \end{bmatrix}\right) \quad (2)$$

Being K_{y,y^*} , $K_{y,y}$ and K_{y^*,y^*} the covariance matrix between the training-validation, training-training, and validation-validation points, respectively. Here, the conditional distribution for the validation points is given by $y^*|X,y,X^* \sim N(y^*, \text{cov}(y^*))$, where:

$$\bar{y}^* = K_{y^*,y}^T [K_{yy} + \sigma^2 n I]^{-1} y \quad (3)$$

$$\text{cov}(y^*) = K_{y^*,y^*} + \sigma^2 n I^* - K_{y^*,y}^T [K_{yy} + \sigma^2 n I]^{-1} K_{yy}^* \quad (4)$$

2) Multi-Output Gaussian Processes (MOGP). The general method for Multiple Output Gaussian Processes (MOGP) describes D output tasks $\{f_d(x)\}_{d=1}^D$, $x \in \mathbb{R}^p$, through convolution integrals of latent functions $\{u_q^i(x)\}_{q=1, i=1}^Q, R_q$, with soft kernels $\{G_{d,q}^i(x-z)\}_{D=1, q=1, i=1}^{DQ, R_q}$

$$f_d(x) = \sum_{q=1}^Q \sum_{i=1}^{R_q} \int G_{d,q}^i(x-z) u_q^i(z) dz \quad (5)$$

Assuming that the latent functions $u_q^i(x)$ are independent Gaussian processes with covariance function $k_q(x,x')$, the outputs $f_d(x)$ conform to a joint Gaussian process with covariance function $k_{d,d'}(x,x')$ with $d,d' = 1, \dots, D$, given by:

$$\sum_{q=1}^Q \sum_{i=1}^{R_q} \int \int G_{d,q}^i(x-z) G_{d',q}^i(x'-z') k_q(z,z') dz dz' \quad (6)$$

Which is known as convoluted multiple output covariance or CMOC. If it is assumed that $G_{j,q}^i(x-z) = a_{j,q}^i \delta(x-z)$, being $\delta(x)$ the Dirac delta function, there is a particular case for the covariance function known as the linear model of coregionalization (LMC). The Covariance $K_{jj'}(x,x')$ simplifies to:

$$K_{jj'}(x, x') = \sum_{q=1}^Q \sum_{i=1}^{R_q} a_{jq}^i a_{j'q}^i k_q(x, x') = \sum_{q=1}^Q b_{jj'}^q k_q(x, x') \quad (7)$$

Where $b_{jj'}^q = \sum_{i=1}^{R_q} a_{jq}^i a_{j'q}^i$. In this case, RBF kernels (see Equation (1)) are used to build the LMC. Once the LMC is developed, the classical inference method based on maximum likelihood is used to find the posterior distribution over the outputs (x). For a detailed explanation of Multiple Output Gaussian Processes [19]. In the context of a WSN, each output $f_j(x)$ corresponds to the data delivered by each sensor.

D. Classical Machine Learning Techniques

1) Support Vector Machine. The Support Vector Machine for Regression (SVR) is a variant of the Support Vector Machine (SVM) used for regression problems rather than classification [12].

2) Random Forest (RF). It is a supervised learning algorithm used for both classification and regression tasks in machine learning. Instead of relying on a single tree to make decisions, RF combines

the predictions from multiple trees to improve accuracy. Its main feature is randomization during tree construction and decision making. In training, a subset of the features is randomly selected, and variations are also introduced into the training data.

E. Experimental Setup

The details of each algorithm, as well as the evaluation and partitioning of the database for training and validation, are described as follows:

Training partition: The dataset was split into 80% for training and 20% for testing across four atmospheric variables: temperature, relative humidity, atmospheric pressure, and soil humidity.

MOGP: The linear model of co-regionalization was used to compute the posterior distribution, along with a radial basis function (RBF) kernel for the covariance matrix calculation. Training was set to 150 iterations with 4 latent functions.

GP: An RBF kernel and maximum likelihood estimation were applied to compute the posterior distribution. Training was set to 50 iterations.

SVR: An RBF kernel was used with parameters $C=20$ and $\text{Gamma}=0.5$, determined via grid search.

RF: An ensemble of 100 decision trees was employed.

The Root Mean Square Error (RMSE) was calculated on the validation sets for all methods.

3. RESULTS AND DISCUSSION

A. MOGP Results

Figure 2 shows data modeling with MOGP for Frame 1. The red line corresponds to the real data of the WSN, the blue line to the regression with MOGP, and the green points are the data necessary for reconstruction, that is, it indicates the dimensionality reduction having a given RMSE.

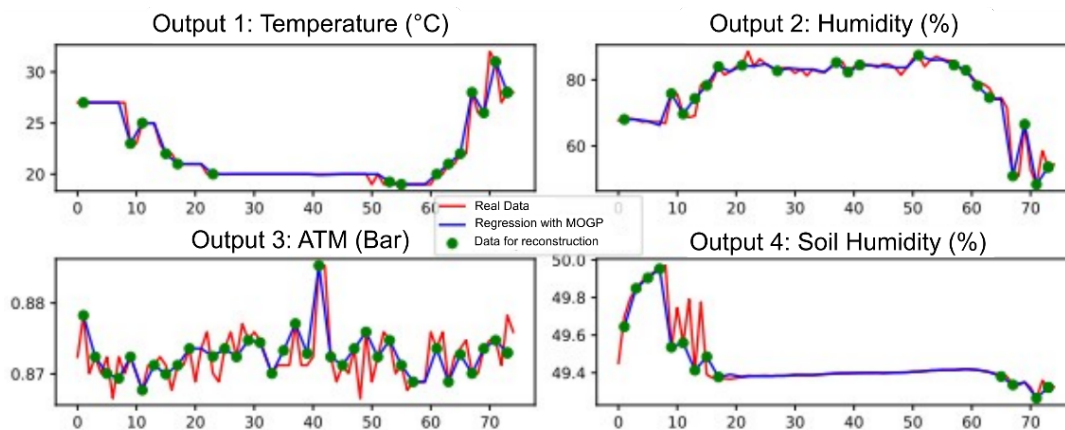


Figure 2. MOGP regression for Frame 1

Tables 1 and 2 show RMSE values for all frames and dimensionality reduction percentages, the average value and the corresponding standard deviation.

Table 1. RMSE results for MOGP

	S1:Temperature	S2:Humidity	S3:ATM	S4:Soil humidity
Frame 1	0.043	1.040	1.25	0.018
Frame 2	1.486	2.606	0.35	0.132
Frame 3	0.781	1.831	0.75	0.129
Frame 4	0.788	2.275		0.183
Frame 5	0.255	3.807		0.155
Frame 6	0.691	3.616		0.197
Frame 7	0.453	3.324	0.8	0.109
Frame 8	0.690	1.841	0.4	0.093
Frame 9	0.673	3.415		0.139
Frame 10	0.374	3.337		0.077
Average	0.623±0.390	2.709±0.931	(1.09±0.52)	0.123±0.052

Table 2. Data reduction (dimensionality) results with MOGP given in percentage (%).

	S1:Temperature	S2:Humidity	S3:ATM	S4:Soil humidity
Frame 1	78.7	73.3	54.7	84.0
Frame 2	57.3	64.0	53.3	52.0
Frame 3	53.3	66.7	58.7	53.3
Frame 4	56.0	65.0	56.0	50.0
Frame 5	53.0	54.0	55.0	51.0
Frame 6	52.0	56.0	56.0	52.0
Frame 7	56.0	66.0	58.0	51.0
Frame 8	53.0	64.0	57.0	51.0
Frame 9	64.0	63.4	56.0	49.5
Frame 10	56.0	60.0	58.0	52.0
Average	57.9±8.1	63.2±5.5	56.3±1.7	54.6±10.4

B. Simple GP Results

Figure 3 shows the data modeling with GPs for Frame 1. The color coding is the same as described for MOGP.

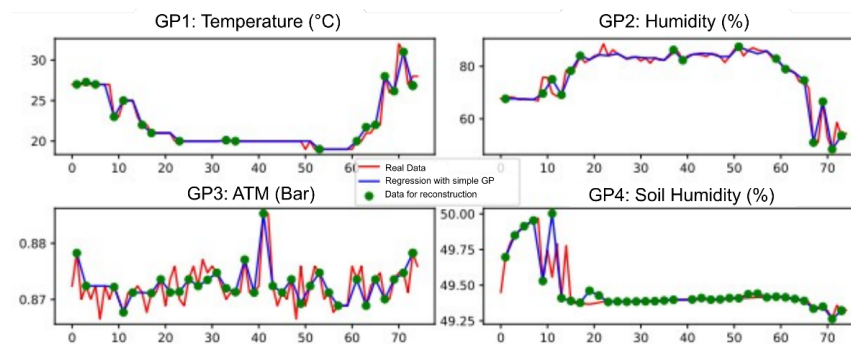


Figure 3. Simple GPs regression for Frame 1

C. SVR Results

Figure 4 shows data modeling with independent SVRs for Frame 1. The color coding is the same as in the previous cases; however, there are no green points, since SVR does not allow for redundant data reduction.

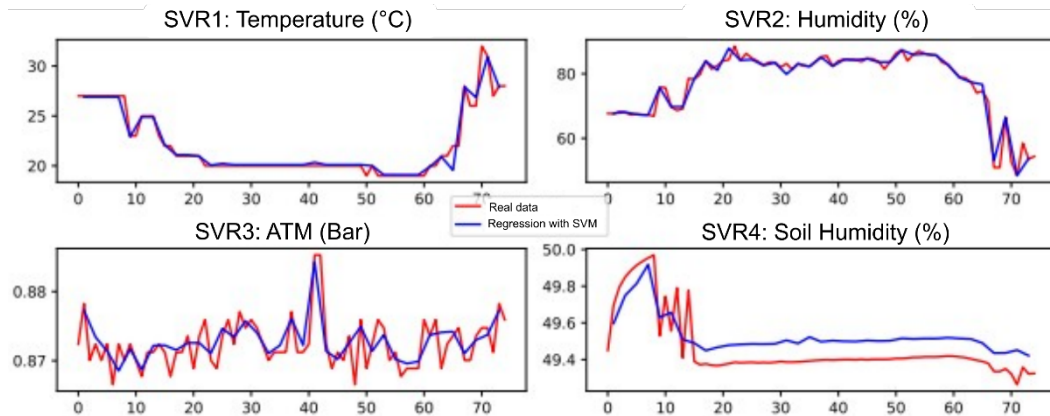


Figure 4. SVR regression for Frame 1

D. RF Results

Figure 5 shows the data modeling with Random Forest for Frame 1. The color coding is the same for SVR.

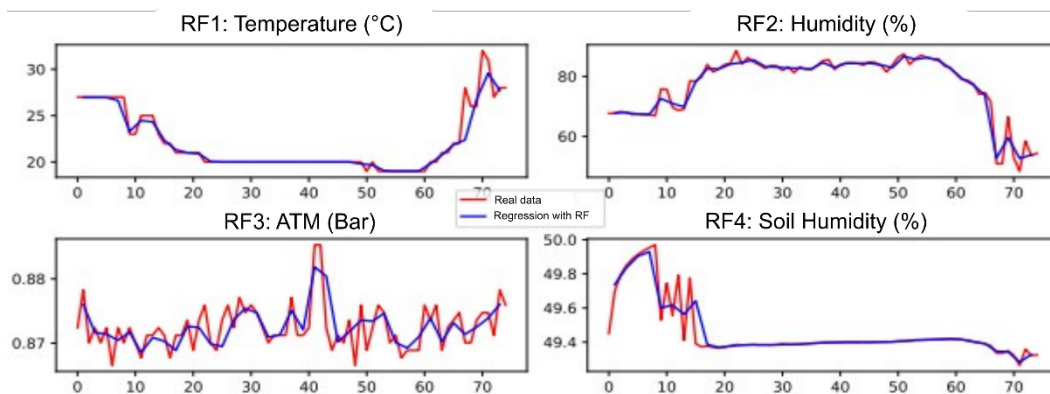


Figure 5. RF regression for Frame 1

E. Comparison Among Models

1) **Analysis in terms of RMSE.** Table 3 provides comparative results of the Root Mean Square Error (RMSE) for different machine learning methods (MOGP, GP, SVR, RF) applied to data from four sensors (S1, S2, S3, S4), each related to different environmental magnitudes (temperature, humidity, atmospheric pressure and soil humidity).

Table 3. Comparative results (average) of RMSE for each machine learning method

	S1:Temperature	S2:Humidity	S3:ATM	S4:Soil Humidity
MOGP	0.623±0.390	2.709±0.931	(1.09±0.52)	0.123±0.052
GP	0.808±0.324	2.879±1.772	(1.16±0.53)	0.132±0.027
SVR	1.311±0.996	5.115±3.494	(2.01±1.08)	0.232±0.085
RF	0.786±0.281	3.574±1.301	(1.55±0.41)	0.179±0.054

According to the results described in tables and figures, the following can be established:

MOGP (Multi-Output Gaussian Process): This method exhibits the lowest RMSE across all sensors: S1 (temperature), S2 (humidity), and S4 (soil humidity). This indicates that MOGP is the most accurate approach for predicting these magnitudes compared to other methods. For S3 (atmospheric pressure), MOGP also achieves the lowest RMSE, although the difference with some other methods is less pronounced. The accuracy of MOGP is particularly high, as it shows the lowest RMSE in all four scenarios. This suggests that MOGP is robust and well-suited for adapting to different environmental magnitudes. Additionally, MOGP demonstrates low variability, which indicates greater robustness and reduced sensitivity to changes in input data or model configuration.

GP (Gaussian Process): This method performs well for S1 (temperature) and S4 (soil humidity), though not as effective as MOGP. Overall, it ranks second in terms of accuracy. Simple Gaussian Processes remain a strong alternative for regression problems; however, in this context, as many independent GPs are required as there are sensors in the WSN. In contrast, MOGP requires only a single model. Whether single-output or multi-output, Gaussian Processes are particularly efficient for modeling relationships in data with potentially nonlinear patterns.

SVR (Support Vector Regression): This method tends to show intermediate performance across all scenarios. It is neither the most accurate nor the least, which may suggest that it is a more balanced model and less prone to overfitting. SVR follows the temporal trends of the WSN measurements; however, it struggles to capture abrupt changes. It is observed that SVR exhibits higher standard deviations compared to other methods, thus reflecting greater variability in its results. This could be attributed to its sensitivity to hyperparameter configuration.

RF (Random Forest): Similar to GP, it has a fairly consistent performance in all scenarios but fails to outperform MOGP in terms of accuracy. RF is an interesting model, simple and easy to implement. It is certainly a good learning machine for modeling WSN data, with acceptable errors.

2) Analysis of Redundant Data Reduction (Dimensionality). Results in [Table 4](#) provide relevant information about the redundant data reduction capability using GP and MOGP in a WSN.

Table 4. Average results of redundant data reduction percentage (dimensionality) in WSN for MOGP and GP

	S1:Temperature	S2:Humidity	S3:ATM	S4:Soil Humidity
MOGP	57.9±8.1	63.2±5.5	56.3±1.7	54.6±10.4
GP	55.6±7.1	67.7±5.0	56.6±2.5	49.7±1.0

In summary, results support the effectiveness of Gaussian process-based models (MOGP and GP) for dimensionality reduction in wireless sensor network environments by using uncertainty information effectively.

4. CONCLUSIONS

MOGP is the most outstanding method for modeling WSNs in this experiment, demonstrating consistency and accuracy in predicting various sensed environmental magnitudes. This method was compared with single-output Gaussian Processes (GP), Support Vector Regression (SVR), Multilayer Neural Networks (NN), and Random Forest (RF). These results support its applicability in real-world scenarios where resource optimization is essential, for instance, in WSNs with energy and bandwidth constraints.

A database collected using the WSN located at Universidad del Quindío was acquired and organized. This database is available for other researchers in the field to conduct similar experiments or replicate the results presented in this study. The database can be accessed at the following link: https://github.com/Nancho027/WSN_UQ_Dataset

AUTHORS' CONTRIBUTION

José-Aicardo Ortega-Vela: Data curation, research, methodology, software development, writing-original draft.

Pablo-Andrés Muñoz-Gutiérrez: Supervision and leadership in planning and execution of activities, writing-review and editing.

Hernán-Darío Vargas-Cardona: Software development, results analysis, writing – review and editing.

REFERENCES

- [1] M. S. Mahmoud, Y. Xia, *Networked Filtering and Fusion in Wireless Sensor Networks*, 2015.
- [2] A. Ali, Y. Ming, S. Chakraborty, S. Iram, "A Comprehensive Survey on Real-Time Applications of WSN," *Future Internet*, vol. 9, pp. 77-87, 2017. <https://doi.org/10.3390/fi9040077>
- [3] J. A. López, *Integración y fusión multisensorial en robots móviles autonomos*, Doctoral Dissertation, Universidad Complutense de Madrid, Spain, 1998.
- [4] H. Mitchell, *Data Fusion: Concepts and Ideas*, Springer, 2012.
- [5] F. Castanedo, *Fusión de datos distribuida en redes de sensores visuales utilizando sistemas multi agente*, Doctoral Dissertation, Universidad Carlos III de Madrid, Spain, 2010.
- [6] A. Mostafa, X. Jiang, *Wireless Sensor Multimedia Networks: Architectures, Protocols, and Applications*, 2015. <https://doi.org/10.1201/b19230>
- [7] A. Almasri, A. Khalifeh, S. Al-Agtash, "Scsap: Spiral clustering based on selective activation protocol for industrial tailored wsns," *Journal of Industrial Information Integration*, vol. 27, e100332, 2022. <https://doi.org/10.1016/j.jii.2022.100332>
- [8] M. Liggins II, D. Hall, J. Llinas, *Handbook of Multisensor Data Fusion: Theory and Practice*, 2008. <https://doi.org/10.1201/9781420053098>
- [9] J. R. Raol, *Multi-Sensor Data Fusion with MATLAB®*, 2009.

- [10] H. Sardar, "American Society of Mechanical Engineers," in *Proceedings of the ASME Dynamic Systems and Control Division*, 2003.
- [11] L. D. Avendano-Valencia, L. E. Avendano, J. M. Ferrero, G. Castellanos-Dominguez, "Improvement of an extended Kalman filter power line interference suppressor for ECG signals," in *Computers in Cardiology*, 2007, pp. 553-556. <https://doi.org/10.1109/CIC.2007.4745545>
- [12] Bishop, *Pattern Recognition and Machine Learning*, 2006.
<http://research.microsoft.com/en-us/um/people/cmbishop/prml/>
- [13] A. Smola, *Regression estimation with support vector learning machines*, 1996.
- [14] C. Zhai, J. Lafferty, "A study of smoothing methods for language models applied to information retrieval," *ACM Transactions on Information Systems*, vol. 22, pp. 179-214, 2004.
<https://doi.org/10.1145/984321.984322>
- [15] M. C. Vuran, I. F. Akyildiz, *Wireless Sensor Networks*, John Wiley & Sons, 2010.
- [16] S. El khediri, A. Benfradj, A. Thaljaoui, T. Moulahi, R. Ullah Khan, A. Alabdulatif, P. Lorenz, "Integration of artificial intelligence (ai) with sensor networks: Trends, challenges, and future directions," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 1, e101892, 2024.
<https://doi.org/10.1016/j.jksuci.2023.101892>
- [17] S. K. Jaiswal, A. K. Dwivedi, "A Security and Application of Wireless Sensor Network: A Comprehensive Study," in *International Conference on IoT, Communication and Automation Technology (ICICAT)*, 2023, pp. 1-5. <https://doi.org/10.1109/ICICAT57735.2023.10263644>
- [18] C. E. Rasmussen, K. I. Williams, *Gaussian Processes for Machine Learning*, 2006.
- [19] M. Alvarez, N. D. Lawrence, "Sparse convolved gaussian processes for multi output regression," in *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods ICPRAM*, 2016, pp. 437-445. <https://doi.org/10.5220/0005700504370445>