ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

Soft Sensor Development for the Measurement of CO_2

Abhishek Deshmukh[1], Jayesh Petkar[2], Aniruddha Shukla[3]

Alumnus, Dept. of Electrical Engineering, Zeal College of Engineering and Research, Pune, Maharashtra, India[1]

Alumnus, Dept. of Electrical Engineering, Zeal College of Engineering and Research, Pune, Maharashtra, India[2]

Alumnus, Dept. of Computer Engineering, Marathwada Mitramandal's College of Engineering, Pune, Maharashtra, India[3]

ABSTRACT:A soft sensor is developed using Gaussian Process Regression. A data set of excess air and carbon dioxide in a boiler is plotted. Using the changes in the amount of excess air, the amount of carbon dioxide is predicted using the available data set. Accurate predictions require optimized value of the squared exponential kernel, which is used in the Gaussian Process model. Optimal value for the squared exponential kernel is found out to make accurate predictions. The optimized values are then put into the Gaussian Process Regression model and the regression/prediction is found out of the amount of carbon dioxide for the required value of excess air.

KEYWORDS: Carbon Dioxide, Gaussian Process Regression, Squared Exponential Kernel, Meanand Covariance.

I.INTRODUCTION

Soft sensor or virtual sensor is a common name for softwarethat is used to process and analyze several measurements together.Soft sensors can deal with measurements numbering in the dozens or the hundreds. Interactions of multiple signals can be used to calculate quantities relevant to the process which need not be measured separately[8][10]. This combination of different measurements is known as data fusion, and can be implemented in soft sensors. They can be useful for fault diagnosis and can even output a control signal[1][9].

In this paper,Gaussian Process Regression is used to develop asoft sensor for measuring the carbon dioxide in the exhaust of a boiler. It is a mathematical model applied in a ladder logic or a computer program to obtain the required result. The amount of carbon dioxide depends upon the excess air provided for the combustion process. As the excess air increases, the carbon dioxide in the exhaust decreases[12]. The relationship between the excess air and CO_2 is given in the following topics.

II.MODELLING OF THE GAUSSIAN PROCESS FOR REGRESSION

Gaussian Process Regression is used for the development of the soft sensor for the measurement of carbon monoxide.Gaussian processes (GPs) extend multivariate Gaussian distributions to infinite dimensionality. Formally, a GP generates data located throughout some domain such that any finite subset of the range follows a multivariate Gaussian distribution[11]. Now the n observations in an arbitrary data set, y={y1,y2….yn} can always be imagined as single point sampled from some multivariate (n-variate) Gaussian distribution[6]. Hence, working backwards, this data set can be partnered with a GP. Frequently, the mean of the GP is assumedto be zero everywhere. What relates one observation to another in such cases is the covariance function, k(x, x'). A popular choice of the covariance function/kernel is the 'squared exponential'[2][7]

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

$$K(x_i, x_j) = \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{2h^2}\right) \quad \text{where } h > 0 \qquad ....1$$

The 'h' in equation 1 is the smoothing parameter σ in equation 2[3].

$$\sigma = \left(\frac{\sum_{n=1}^{n} (x_i - \hat{\mu})^2}{n}\right)^{1/2} \quad \text{where } \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad ....2$$

The purpose of calculating 'h' using equation 3 is to find out the optimal value for the squared exponential kernel. Optimized value of 'h' gives rise to accurate prediction/ regression.

Sometimes the data are noisy from measurement errors and so on. Each observation y can be thought of as related to the underlying function f(x) through a Gaussian noise model[2].

$$y = f(x) + \mathcal{N}(0, \sigma_n^2) \qquad ....3$$

For simplicity, it is assumed that the measurements of excess air content are accurate. Hence the noise term can be eliminated from the regression calculations.

So now given n observations y, our objective is to predict y*, not the actual f*. To prepare for Gaussian Process Regression, we calculate the covariance function (equation 1) for all possible combinations of the following points summarizing our findings in three matrices[2].

$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{bmatrix}$$

$$K_* = \begin{bmatrix} k(x_*, x_1) & k(x_*, x_2) & \cdots & k(x_*, x_n) \end{bmatrix} \qquad K_{**} = k(x_*, x_*)$$

In order to perform regression, the assumption is that our data can be represented as a sample from a multivariate Gaussian distribution, we have that

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & K_*^{\mathrm{T}} \\ K_* & K_{**} \end{bmatrix}\right)$$

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

Where T indicates the transpose of a matrix. We want the conditional probability P(y*|y), meaning given the data, how likely is a certain prediction for y_*[2]. The probability follows a Gaussian distribution as given in equation 4

$$y_*|\mathbf{y} \sim \mathcal{N}(K_*K^{-1}\mathbf{y},\ K_{**} - K_*K^{-1}K_*^{\mathrm{T}}) \quad \ldots 4$$

Our best estimate for y* is the mean of this distribution given by equation 5

$$\overline{y}_* = K_*K^{-1}\mathbf{y} \quad \ldots .5$$

And the uncertainty of our estimate is captured in its variance given by equation 6

$$\mathrm{var}(y_*) = K_{**} - K_*K^{-1}K_*^{\mathrm{T}} \quad \ldots .6$$

Here, K is an n*n matrix;
y is a 1*n matrix;
x is an n*1 matrix
and x* and y* are scalar quantities

III.OBSERVATIONS AND RESULTS

Fig. 1 shows the graph of Excess Air vs Carbon Dioxide. This relationship is for when fuel oil is used as a fuel in the combustion process of a boiler. Seven observations of excess air were taken, the corresponding change in the amount of CO_2 in the exhaust of the boiler was measured and the graph of excess air vs CO_2 was plotted as shown below. The assumption made here is that no other parameter affects the amount of CO_2 present in the exhaust of the boiler except the quantity of excess air.

Here, x= [10 20 30 40 50 70 90];

x*= 45;

h=26.10; ….(from equation 2)

y= [14

13

12.1

11

10.1

9.2

8.4];

The value of Y* at X* is found out using the Gaussian Process Regression Model.

 DOI:10.15662/IJAREEIE.2016.0507079

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

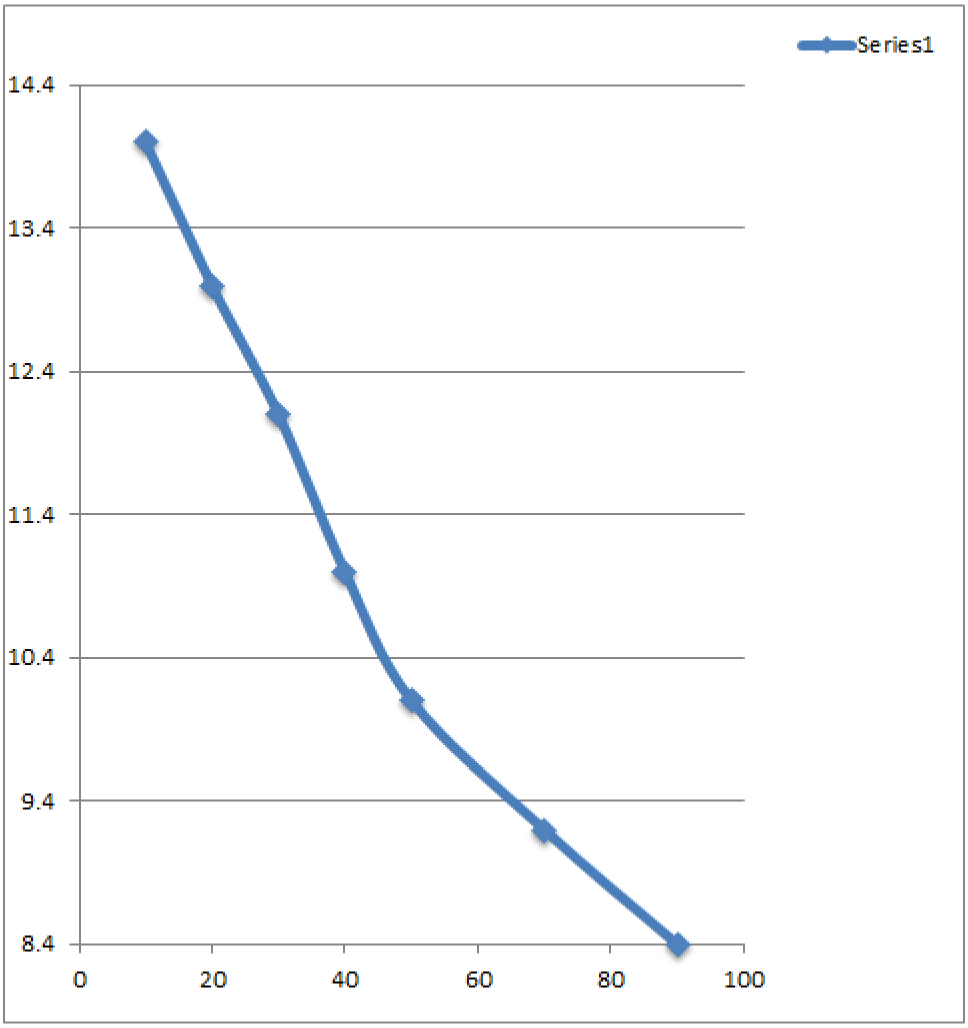

Fig. 1: X-axis: Excess Air and Y-axis: CO2%

Using the squared exponential function at equation 1, the values of k(x, x') were calculated as follows

$$
\begin{bmatrix}
1.0000 & 0.920 & 0.740 & 0.51 & 0.31 & 0.071 & 0.0091 \\
0.9200 & 1.000 & 0.920 & 0.74 & 0.51 & 0.160 & 0.0270 \\
0.7400 & 0.920 & 1.000 & 0.92 & 0.74 & 0.310 & 0.0071 \\
0.5100 & 0.740 & 0.920 & 1.00 & 0.92 & 0.510 & 0.1600 \\
0.3100 & 0.510 & 0.740 & 0.92 & 1.00 & 0.740 & 0.3100 \\
0.0710 & 0.160 & 0.310 & 0.51 & 0.74 & 1.000 & 0.7400 \\
0.0091 & 0.027 & 0.071 & 0.16 & 0.31 & 0.740 & 1.0000
\end{bmatrix}.
$$

In order to calculate the value of y*, the inverse of the above matrix was calculated. The inverse is, i.e. K^{-1}, is as shown below[4]

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

$$\begin{bmatrix} 16.611 & -26.431 & 5.941 & 14.724 & -10.573 & 1.007 & 0.697 \\ -26.037 & 54.091 & -27.893 & -16.284 & 16.745 & 0.943 & -4.309 \\ 4.763 & -25.049 & 40.507 & -29.680 & 12.975 & -9.164 & 7.853 \\ 16.888 & -25.208 & -18.618 & 70.548 & -55.532 & 19.553 & -7.883 \\ -12.888 & 27.844 & -3.500 & -47.003 & 48.755 & -18.770 & 5.687 \\ 2.916 & -8.337 & 4.803 & 11.822 & -18.023 & 11.250 & -4.465 \\ -0.651 & 2.130 & -1.667 & -3.052 & 5.831 & -5.019 & 3.355 \end{bmatrix}$$

From equation 5, the best estimate or the mean for y* was calculated. The value of the best estimate or the mean is 10.518 [4]. From the graph, it can be seen that the result is pretty accurate.

The covariance is -0.003. When the greater values of one variable mainly correspond to the lesser values of the other, i.e., the variables tend to show opposite behaviour, the covariance is negative. And if both the variables show similar behaviour, then the covariance is positive.

IV.FUTURE PROSPECTS

Gaussian Process Regression is a universal model which means that it can be used for any applications containing a normally distributed data set. In boilers alone, a lot of sensors are used unnecessarily to measure various parameters like CO_2, carbon monoxide, NO_x-SO_X, steam temperature, stack temperature etc. Many of these parameters are interdependent i.e. if one parameter varies, then so does the other. This interdependency of the parameters can be used to develop various soft sensors which have the potential to overcome the drawbacks of a physical sensor. The model developed in this paper is very simple and easy to apply in various applications and the cost of its development is minimal.

VI.CONCLUSION

As seen from the results, a highly efficient and accurate soft sensor has been developed to measure the amount of carbon dioxide in the exhaust of a boiler. A Gaussian Process model has been used for regression.Matrix calculations were done using an online matrix calculator and the prediction results were found out.

REFERENCES

[1] A. Verikas. (2009, September) *Soft Sensors for Monitoring* [Online] Available: https://www.hh.se/download/18.70cf2e49129168da0158000149268/soft_Sensors.pdf
[2] M. Ebden, (2008, August) *Gaussian Processes for Regression: A Quick Introduction* [Online] Available: http://www.robots.ox.ac.uk/~mebden/reports/GPtutorial.pdf
[3] G. Felici and C. Vercellis, "Kernel Width Selection for SVM Classification: A Meta-Learning Approach," in Mathematical Methods for Knowledge Discovery and Data Mining,Hershey PA, Information Science Reference2008
[4] "Online Matrix Calculator". Bluebit.gr. N.p., 2016. Web. 16 July 2016.
[5] J. Hensman*et al.*, (2014, April 14)*Fast nonparametric clustering of structured time-series*[Online] Available: https://arxiv.org/pdf/1401.1605.pdf, doi: arXiv:1401.1605v2 [cs.LG]
[6] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006
[7] Joaquin Quinonero-Candela, Carl Edward Rasmussen, and Christopher K. I. Williams. *Approximation Methods for Gaussian Process Regression. In Large Scale Kernel Machines*. MIT Press, 2007
[8] S. Chen, Z. Xi, W. Yan and D. Zhang, *Soft Sensor Modeling for Oxygen-Content in Flue Gasses in 1000MW Ultra-superficial Units*, Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), 2015 International Conference on, Wuhan, 2015, pp. 164-167.doi: 10.1109/ICIICII.2015.124
[9] L. Fortuna, *Soft sensors for monitoring and control of industrial processes*, 2007, Springer Verlag
[10] J. Maa, X. Cuia , X. Fub, *On line measurement of carbon nanotube growth rate based on soft sensor technology*,New Paradigm of Particle Science and Technology Proceedings of The 7th World Congress on Particle Technology, Volume 102, 2015, Pages 1742-1751
[11] J. Zheng, B. Song, Y. Wang, H. Wang, *Adaptive Filter Updating for Energy-Efficient Top-k Queries in Wireless Sensor Networks Using Gaussian Process Regression*, International Journal of Distributed Sensor Networks June 2015 vol. 11 no. 6 304198, doi: 10.1155/2015/304198
[12] J. Piper, Operations and Maintenance Manual for Energy Management, Routledge (May 14, 1999)