Predictive Maintenance in Power Transformers

COMS7009A - Research Report

J.M. Atherfold

Student number: 708298
Supervisor: Dr T. van Zyl

A research report submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Science in Computer Science.

Johannesburg, February 2019

University of the Witwatersrand, Johannesburg

School of Computer Science and Applied Maths

INDIVIDUAL DECLARATION WITH TASK SUBMITTED FOR ASSESSMENT

I declare that this research report is my own, unaided work, except where otherwise acknowledged. It is being submitted for the degree of Master of Science in Computer Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other university.

Signed this 7th day of February 2019

J.M. Atherfold

Acknowledgements

Thank you to LH Marthinusen: Electro-Mechanical Engineers, who sponsored this research.

Abstract

Power transformers are the link between power generation systems, and the clients that are supplied with power. They are one of the most important components in a power distribution network, and hence their reliability is of the utmost importance. Maintenance data from the power transformers is in the form of Dissolved Gas Analysis (DGA) time series data. This research attempts to consolidate industry knowledge on the maintenance of power transformers and time series forecasting techniques into a coherent system for the purpose of predictive maintenance of power transformers. In addition to this, the generalisability of forecasting models is investigated by measuring performance of single models across multiple transformers, and hence, multiple data sets. The data preprocessing method utilised an exponential smoothing technique specifically designed for the type of raw data received (aperiodically sampled, noisy data). Additional features were engineered based on research in the field, and added to the data set. The preprocessing technique is novel in the context of DGA forecasting. The forecasting method developed utilises a Least-Squares Support Vector Machine (LS-SVM), and the hyper-parameters of the model were chosen and optimised using a Particle Swarm Optimiser (PSO). Performance of the developed models was measured against various metrics, including previously developed univariate forecasting techniques; built-in Support Vector Regressors; Naive and Mean forecasts; and ARIMA and Exponential Smoothing forecasters. Two sets of experiments were run (LS-SVM(1) and LS-SVM(2)) which differed in how the Training, Validation, and Testing sets were chosen. LS-SVM(1) held out an entire transformer for Testing; and LS-SVM(2) held out the final h points of each transformer, $h \in [1:4]$. These experiments were run for different input vectors; the Original input vector, and an input vector Augmented with the additionally engineered features. The statistical significance between the difference of error distributions produced by the different input vectors was discussed, and found to be highly feature-dependent. The LS-SVM(2) model outperformed all other models, when the distributions of the Testing errors were considered. In terms of the generalisability of the models, it was found that the models trained across all transformers outperformed the per-transformer models that treated the data as univariate time series. Recommendations were made in order to potentially improve the performance of future developed models, such as changing how the Training, Validation, and Testing sets are chosen; changing the fitness function used in the PSO; and training individual models per transformer.

Published Work

Aspects of this research report have been submitted for publication in the following references:

1. J.M. Atherfold and T. Van Zyl, 'A Novel Method for DGA Forecasting in Power Transformers Using Least-Squares Support Vector Regression', *IEEE Transactions on Power Systems*, Submitted for Publication.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms

ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
DGA	Dissolved Gas Analysis
EMA	Exponential Moving Average
ES	Exponential Smoothing
LOO	Leave-One-Out
LS-SVM	Least Squares Support Vector Machine
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
PSO	Particle Swarm Optimisation
RBF	Radial Basis Function
SVR	Support Vector Regression
UV	Univariate

Greek Symbols

$\boldsymbol{\alpha}$	Lagrange Multipliers/LS-SVM Model Parameters
ϵ	Minimum Error in ϵ-SVR
γ	Kernel Coefficient for RBF
λ	Inertia Parameter for PSO
$\lambda_{max}, \lambda_{min}$	Initial and Final Inertial Weights Respectively
$\boldsymbol{\Omega}$	Kernelised LS-SVM Model Input
μ	Mean
μ_{MSE}	Mean MSE of a Feature
ω	Parameter in SVM Formulation
ϕ	Set of Non-linear Transformations Operating on an Input

σ Standard Deviation

τ Fixed Length Rolling Window

ξ, ξ^* Slack Variables for SVM Formulation

Symbols

$\hat{f}$ Estimator of True Value

$\mathbb{B}$ Backshift Operator

$\overline{\mathbf{a}}$ Smoothed and Resampled Time Series

$\overline{\mathbf{a}}_i^j$ Values in Time Series $\overline{\mathbf{a}}^j$ from indices i to $i+m-1$

$\mathbf{a}$ Unevenly Spaced Time Series

$\mathbf{b}_i^{k_p}$ i^{th} Feature Vector Augmentation of Input Vector

$\mathbf{c_1}, \mathbf{c_2}$ Relative Velocity Regulators

$\mathbf{g}_{best}(t)$ Global Best Position at t^{th} Time Step

$\mathbf{P}_i(t)$ Position of i^{th} Particle at t^{th} Time Step

$\mathbf{P}_{best}(t)$ Personal Best Position of i^{th} Particle at t^{th} Time Step

$\mathbf{v}_i(t)$ Velocity of i^{th} Particle at t^{th} Time Step

$\mathbf{x}$ LS-SVM Model Input Vector (Query Point)

a_{t_i} Value in Unevenly Spaced Time Series

b Model Bias

C Regularisation Parameter

d Dimensionality of Model Input Vector, $\mathbf{x}$

e Model Error

G Total Number of Generations

h Number of Points used in Testing Set

I Iterative Variable - Transformer Number

i Iterative Variable - Case Specific

j Forecast Feature Number

k	Iterative Variable - Case Specific
$K(\mathbf{x}, \mathbf{x}_i)$	Mercer Kernel Function Operating on $\mathbf{x}$ and $\mathbf{x}_i$
k_p	Feature from List of Features s.t. $k_p \neq j, p \in [1 : M-1]$
l	Number of Training Pairs in a Training Set
M	Number of Features
m	Hyper-parameter Representing How Many Lags of Feature j are in Input Vector
N	Number of Data Points in a Distribution
n	Number of Points in Smooth Resampled Time Series
r_1, r_2	Random Numbers Drawn from Uniform Distribution
S	PSO Population Size
T	Total Number of Transformers
t_i	Time at which Value a_{t_i} in Aperiodic Time Series Occurs
v	Number of Points used in Validation Set
y	Target Value of Model

1 Introduction

1.1 Background and Motivation

Power transformers are vital components of power distribution networks. The reliability of such a component is of paramount importance; not only for industrial clients, but also for the economic operation of a utility. Unscheduled downtime of a transformer can cost a company significant amounts of money in terms of production interruption, as well as externally imposed penalties [1].

Other concerns related to power transformer failures include: operational complications to the electrical system; leaking of flammable insulating oil that is hazardous to the environment; interruption to service; and the deterioration of the lifetime of the transformer. Service interruption is a major concern, since transformers are large and expensive components, and spares are not kept by the utility in the event of a breakdown. Transformer lifetime is dependent on the insulation and cooling system, since thermal and electrical stresses are the source of transformer ageing. These stresses subject the transformer to various faults [2].

Transformer maintenance is therefore a high priority. Scheduled maintenance ensures operational reliability of components by leaving the system unavailable for the shortest possible time period. In general, there are four main types of component maintenance [1]:

- **Corrective maintenance** - Planned or unplanned maintenance that is undertaken after a fault has already occurred
- **Preventive maintenance** - Seeks correction of faults before they occur, and may result in the premature replacement of parts in a component
- **Predictive maintenance** - Corrects faults before they occur through the use of data analysis techniques, such as analysing trends and finding patterns through historical analysis
- **Proactive maintenance** - Determines the root causes of failure in components, and seeks to deal with these issues before problems occur

The power transformer industry in South Africa currently performs corrective and preventive maintenance, depending on the context of the failure. Unplanned corrective maintenance is the worst case scenario, where failure of a power transformer has occurred, resulting in high costs. There is also the potential threat to the environment, depending on the nature of the failure. Failure also puts stress on the transformer maintenance company, since other projects are delayed because of unscheduled maintenance.

Preventive maintenance is less disruptive to the general operational activities of both the client and maintenance company. Planned downtime is accounted for, and maintenance turnaround time is expected to be shorter. The disadvantage of preventive maintenance is that maintenance may be performed prematurely. Premature servicing of the transformer results in unnecessary costs.

In both preventive and planned corrective maintenance, various data are obtained from components currently in operation, in order to determine when the component needs to be replaced or maintained, as well as the nature of the maintenance. In the case of power transformers, the data is in the form of Dissolved Gas Analysis (DGA). These data are subject to analyses within the maintenance company, according to various international standards, such as ASTM D3612-02 [3]. The outcome of the analysis dictates which failures, if any, have occurred within the transformer, and what action should be taken to ensure the transformer's continued operation.

In addition to the standards, there is knowledge developed and utilised within the maintenance company that's indicative of incipient faults, and the overall health of the transformer. These additional features that indirectly relate to the health of the transformer aid the company in their diagnosis of the system, should the system not be performing as expected. These features have the potential to be utilised in a machine learning model, to potentially increase the accuracy of DGA forecasts.

1.2 Research Question

- Is it possible to incorporate industry and expert knowledge into machine learning models, with the intention of increasing the performance of these models, for the purposes of DGA forecasting?
- To what extent can forecasting models be generalised across transformers? How will the models perform on data that is completely unseen?

1.3 Research Aims and Objectives

1.3.1 Aims

The aims of the research are as follows:

- Create a novel method of DGA forecasting in power transformers.
- Compare this novel method with various other methods of forecasting in this context.

1.3.2 Objectives

The aims of the research are achieved by completing the following objectives:

- Obtain raw Dissolved Gas Analysis data and pre-process it appropriately.
- Investigate existing industry and expert knowledge through literature and consultation.
- Develop forecasting models incorporating this additional industry knowledge.
- Investigate models already developed in literature and compare the results of these to the newly developed models.
- As an additional metric, compare the results to standard ARIMA and exponential smoothing forecasting techniques.
- Conclude which method(s) are best.

1.3.3 Delineations, Limitations, and Assumptions

The following assumptions are made about the data and the research, in order to better define the scope of the research in question:

- The research is restricted to forecasting in the time domain, and does not include the classification of the various transformer faults that occur.
- As is discussed in Section 3.1.2, the raw data needed significant cleaning, both from a re-encoding perspective, and a smoothing and resampling perspective. One of the concerns with resampling time series is the fact that future data points might be used in the resampling process. This effectively makes the data dishonest, as discrete events

such as failures are pre-empted by a trend that is created, but does not exist in reality. This has been avoided where possible by using forward filling, and a smoothing technique that utilises past data points only [4]. A resample interval has been chosen in a way that is realistic and practical when considering the reality of the problem.

- This research does not claim to replace the current methods of maintenance scheduling, it is merely an investigation into the viability of incorporating machine learning techniques in industry for the purpose of data forecasting.

1.4 Outline

Section 1 introduces the problem and gives reasons of the significance of the research in industry, and specifies the direction in which the research developed. Section 2 introduces the type of data used in the research, and specifies the methods used for this type of data forecasting in previous works, as well as the optimisation technique used in this research. Section 3 specifies the methods used for preprocessing of the data, and specifies the forecasting procedures developed for the purpose of this research. Also introduced are baseline models for comparison against the developed models. Section 4 presents and discusses the results of the research, and their implications. Section 5 presents concluding remarks from the research, and Section 6 summarises recommendations for further development. Finally, Appendix A summarises results that were too cumbersome to include in the main body of the work presented.

2 Literature Survey

2.1 Dissolved Gas Analysis

2.1.1 Key Gas Analysis

The current method of fault diagnosis is an analysis of the dissolved gases in the transformer oil. As the transformer operates, the oil and/or insulation paper degrades, thereby producing some combustible gases that dissolve in the oil [2]. The gases produced are mostly hydrocarbons, along with other by-products of combustion. Table 1 shows the types of gases that could be dissolved in the oil, along with the current limit and corresponding fault type. Currently, if the gas limit is exceeded, the transformer is assessed and repaired.

Table 1: Key Gasses and their Associated Fault Types

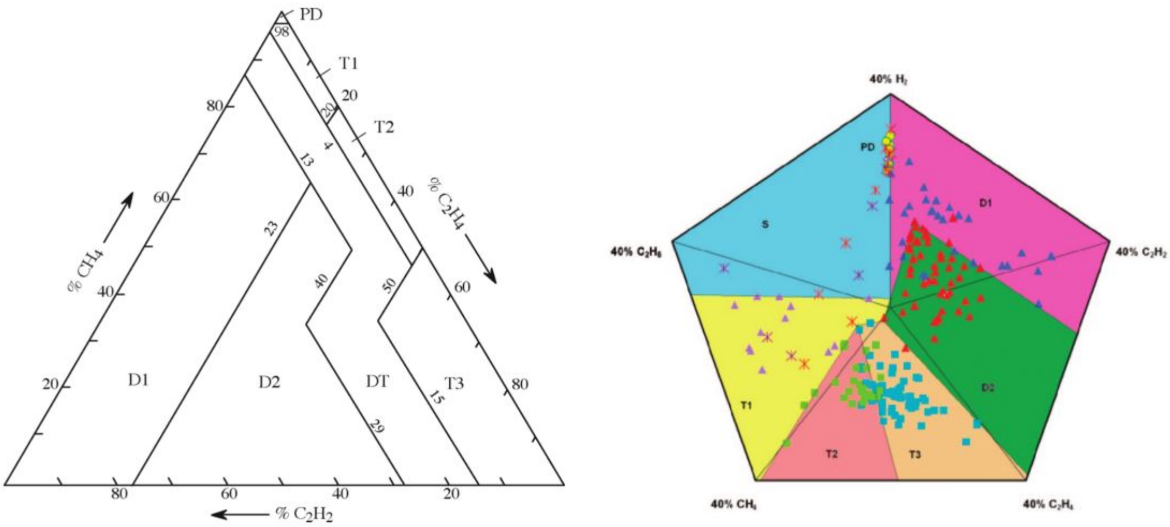
Key Gas	Limit	Fault Type	Expected Degradation Type
Oxygen (O_2)	-	Non-Fault Condition	-
Nitrogen (N_2)	-		
Methane (CH_4)	>100	Low Temperature Overheating of Oil	Continuous
Ethane (C_2H_6)	>100		
Ethylene (C_2H_4)	>80	High Temperature Overheating of Oil	Continuous
Carbon Monoxide (CO)	>1000	Overheating of Cellulose Insulation	Continuous
Carbon Dioxide (CO_2)	>10000		
Hydrogen (H_2)	>150	Corona (Low Intensity Partial Discharge)	Discrete
Acetylene (C_2H_2)	>15	Arcing	Discrete

Temperature faults such as overheating are expected to occur after a significant period of time, and are therefore classified accordingly. Electrical faults such as corona discharge and arcing are instantaneous faults that happen at a discrete time.

The difference in these fault types is significant when considering the data; the gas concentrations corresponding to continuous faults are expected to increase predictably over time, whereas the concentrations of gases associated with discrete faults are expected to increase spontaneously.

2.1.2 Duval Triangle Method

An additional popular method of fault diagnosis in power transformers is the Duval Triangle Method [5], and the Duval Pentagon Method [6]. These methods investigate the relationship between different dissolved gases, and use these relationships and empirical charts to diagnose the fault. Examples of the Duval triangle and pentagon are shown in Figure 1.



(a) (b)

Figure 1: a) Duval Triangle 1 [5], and b) Duval Pentagon 1 [6]

The concentration of each gas of interest is plotted on its corresponding axis. Once this is done, the centroid of the resulting shape is found. The region of the chart corresponding to the centroid's position indicates the type of fault the transformer has undergone. There are five Duval triangles, and two pentagons; each comparing different dissolved gases in order to diagnose multiple faults.

The relationship between dissolved gases can be utilised in the analysis of the data; the ratios between the gases in the relevant triangles and pentagons were used as additional features. This non-linear transformation could potentially reveal structure in the data that would have been otherwise obscured.

2.2 Forecasting in Power Transformers

Several techniques have been used previously in the forecasting of DGA data, including: Support Vector Machines (SVM) [7–10]; various types of Artificial Neural Networks (ANN) [11–13]; Grey Models and Grey Extension Models [14, 15]; and Fuzzy Linear Regression Models [16]. The approach chosen in this report follows two different SVM models, and some baseline comparisons. SVM models are chosen because of their superiority over other machine learning techniques in this context of forecasting [7, 10].

Results from other SVM related papers claim Testing errors of between 1% and 10% [7–10]. A direct comparison between these methods and the current research is difficult, since the data sets, as well as interpolation and resampling techniques are different. Instead, the experimental procedure specified by Liao et al. [7] will be implemented on the current data sets, and the results compared with those of the current research. This experimental procedure is chosen because it closely mimics the procedure specified in this research; the key differences being how data was preprocessed, and the investigation into the generalisability of the model across multiple transformers.

2.2.1 ϵ-SVR

Support vector machines are a classification tool that can be used to classify linearly separable data. If the data is not linearly separable, it is mapped to a higher dimensional feature space where it is separable. Support vector machines can be adapted to perform regression, allowing the analyst to obtain quantitative responses [17, 18].

The purpose of the regression is to deduce an estimate, $\hat{f}(\mathbf{x})$, of the true value, y, as well as the relationship between a vector of observations, $\mathbf{x}$, and the output value, y, from a given set of training samples. This is done by projecting the data from the original feature space to a higher dimensional space in order to increase the flatness of the function, and approximate it in a linear way. The approximation of the function is given as

$$\hat{f}(\mathbf{x}) = \omega^T \phi(\mathbf{x}) + b, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^d$, $\hat{f}(\mathbf{x}) \in \mathbb{R}$, $\omega \in \mathbb{R}^d$, $\phi(\mathbf{x})$ denotes a set of non-linear transformations, and $b \in \mathbb{R}$ is the model bias.

Consider a set of training samples $\{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\} \subseteq \left(\mathbb{R}^d \times \mathbb{R}\right)$, where $\mathbf{x}_i$ is the input, and y_i is the corresponding target value, $i \in \{1, ..., l\}$. The primal optimisation problem for ϵ-SVR is defined as [18]:

$$\min \quad \psi(\omega, \xi, \xi^*) = \frac{1}{2}\omega\omega^T + C\sum_{i=1}^{l}(\xi_i + \xi_i^*), \tag{2}$$

subject to

$$\begin{cases} y_i - \omega^T\phi(\mathbf{x}_i) - b \leq \epsilon + \xi_i \\ \omega^T\phi(\mathbf{x}_i) + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, ..., l \end{cases}, \tag{3}$$

where ξ and ξ^* are the slack variables, C is the regularisation parameter, and ϵ is the minimum penalisable error in the ϵ-SVR. The hyper-parameters for this model are typically associated with the type of kernel function used, $\phi(\mathbf{x})$, and the regularisation parameter, C. The exact formulation of the SVR in the context of forecasting is specified in Section 3.2.

2.2.2 Least Squares - Support Vector Machine (LS-SVM)

This particular machine learning model reformulates the classic SVM development. It applies linear least squares criteria to the loss function, and replaces the inequality constraints with equality ones [7, 19]. In the primal space, the LS-SVM formulation can be described as

$$\min \;\; \psi\left(\omega, e\right) = \frac{1}{2}\omega\omega^T + C\sum_{i=1}^{l}\left(e_i^2\right), \tag{4}$$

subject to the equality constraints

$$y_i = \omega^T\phi(\mathbf{x}_i) + b + e_i, \quad i = 1, 2, ..., l. \tag{5}$$

From this formulation, a Lagrangian function can be constructed for the purpose of finding conditions for optimality [7]. The Lagrangian function is given by

$$L(\omega, b, e, \alpha) = \psi(\omega, e) - \sum_{i=1}^{l}\alpha_i\left(\omega^T\phi(\mathbf{x}_i) + b + e_i - y_i\right), \tag{6}$$

where $\alpha_i, \quad i \in \{1, ..., l\}$ are Lagrange multipliers. The Karush-Kuhn-Tucker conditions for optimality are given by

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \implies \omega = \sum_{i=1}^{l}\alpha_i\phi(\mathbf{x}_i) \\ \frac{\partial L}{\partial b} = 0 \implies \sum_{i=1}^{l}\alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \implies \alpha_i = Ce_i, \quad i = 1, 2, ..., l \\ \frac{\partial L}{\partial \alpha_i} = 0 \implies \omega^T\phi(\mathbf{x}_i) + b + e_i - y_i = 0, \quad i = 1, 2, ..., l \end{cases}. \tag{7}$$

After elimination of the variables ω and e_i, the optimisation problem presents the following linear solution

$$\begin{bmatrix} 0 & \mathbf{1}_{\boldsymbol{l}}^T \\ 1_l & \mathbf{\Omega} + C^{-1}\mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}, \tag{8}$$

where $\mathbf{1}_{\boldsymbol{l}}$ denotes a column vector of length l, $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_l]$, and $\mathbf{y} = [y_1, ..., y_l]$. The regularisation parameter, C, is a hyper-parameter that must be chosen a priori. The Mercer's condition is applied to the matrix $\mathbf{\Omega} = \{\Omega_{ij}\}_{l \times l}$ as follows:

$$\Omega_{ij} = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j). \tag{9}$$

The LS-SVM model for regression therefore becomes

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{l} \alpha_i K(\mathbf{x}, \mathbf{x}_i) + b, \tag{10}$$

where α_i's and b are solutions to the Equation 8. The Mercer kernel function chosen as the radial basis function (RBF)

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(\frac{||\mathbf{x} - \mathbf{x}_i||^2}{2\gamma^2}\right). \tag{11}$$

The RBF is chosen since it is a commonly used kernel in this forecasting context, and is reported to yield the best results [7–11]. The kernel has one hyper-parameter, γ. The exact formulation of the LS-SVM in the context of forecasting is specified in Section 3.2.

2.3 Particle Swarm Optimisation

Particle swarm optimisation (PSO) is a heuristic optimisation algorithm that mimics the flocking of birds [20]. Consider a population of particles, where S denotes the set of all particles in the population. The i^{th} particle in the population is characterised by its current velocity, $\mathbf{v}_i(t)$, its personal best historical position, $\mathbf{P}_{best}(t)$, and its current position, $\mathbf{P}_i(t)$, at the t^{th} time step, for all $i \in S$. The final parameter specified in the problem is the global best position of all particles at the t^{th} time step, $\mathbf{g}_{best}(t)$.

To search for the optimal solution, the velocity and positions of each particle are updated by the following equations:

$$\begin{cases} \mathbf{v}_i(t+1) & = \lambda \mathbf{v}_i(t) + r_1 \mathbf{c_1} \left(\mathbf{P}_{best}(t) - \mathbf{P}_i(t) \right) + r_2 \mathbf{c_2} \left(\mathbf{g}_{best}(t) - \mathbf{P}_i(t) \right) \\ \mathbf{P}_i(t+1) & = \mathbf{P}_i(t) + \mathbf{v}_i(t+1) \end{cases}, \tag{12}$$

where $\mathbf{c_1}$ and $\mathbf{c_2}$ are vectors containing acceleration constants which regulate the relative velocities with respect to the personal best and global best positions respectively; r_1 and r_2 are scalers of random numbers drawn from a uniform distribution, i.e. $r_1 \sim U(0,1)$, $r_2 \sim U(0,1)$; and λ is an inertia parameter given by

$$\lambda = \lambda_{max} - \frac{\lambda_{max} - \lambda_{min}}{G} t, \tag{13}$$

where λ_{max} and λ_{min} are the initial and final inertial weights respectively, and G is the total number of generations. The implementation of the algorithm is shown in Figure 2.

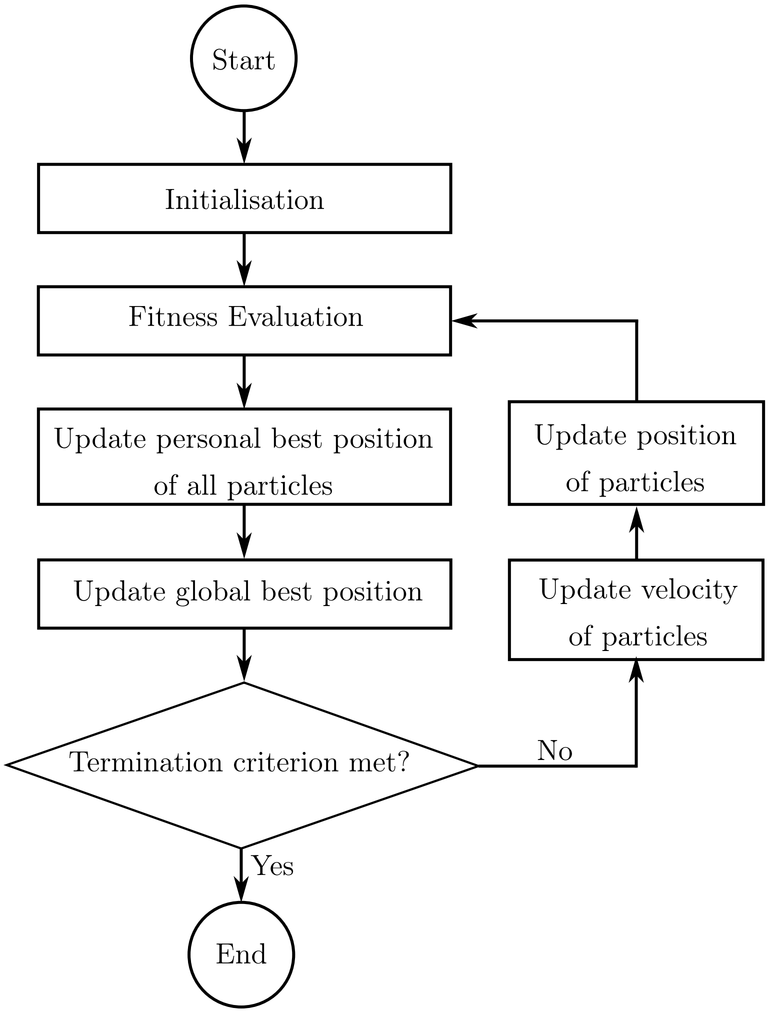


Figure 2: PSO Algorithm

The initialisation of the algorithm consists of initialising the population, either randomly or by other means. The fitness of each member in the population is then evaluated through the use of some fitness function. At each iteration, the personal best position of each particle, as well as the overall best position is updated. If the termination criterion has not been met yet, the velocity and position of each particle is updated by Equation 12. The termination criterion is a set number of iterations for which the algorithm runs, G.

3 Method

The data provided for this research consists of multivariate time series data for multiple power transformers. The raw data was sampled aperiodically, and was very noisy. This data was smoothed and resampled appropriately before being processed further. The data set is relatively large: it consists of time series of 5 - 25 points; has multiple features (discussed and described in Section 3.2); and contains data from 18 power transformers. The most transformers used in various literature sources for time series forecasting so far has been three [7]. Since this data set is relatively large, two main experimental methods were able to be conducted, and the results from each were subsequently compared. The different experimental methods are specified in Sections 3.3 and 3.4.

3.1 Nature of the Data

3.1.1 Original Features

The data was obtained from the power transformer company that sponsored this research. The data consists of multivariate time series, which consist of the key gasses typical in DGA, and some other discrete indicators which provide an indication of the state of the transformer. After consultation with an expert, the features used in the analysis were chosen, and are summarised in Table 2 [21]. There are seven key gasses, which are the features of interest that are forecast, and eleven additional features that indicate the occurrence of important events in the transformers' life span, such as oil filtering or oil replacement.

Continuous and discrete features are inherently treated slightly differently. For instance, the interpolation method used for missing values of continuous features is linear interpolation. For the discrete set of features, it physically makes no sense to interpolate linearly, and therefore forward filling is used for missing values. In some cases, discrete features need to be encoded in order to be interpreted by the machine learning techniques used. This encoding is summarised in Table 3.

Table 2: Features Chosen for Forecasting Model

Feature	Reason for Choice	Feature Type
Hydrogen	Key Gas. Indicates Failure.	Continuous
Ethane		
Methane		
Ethylene		
Acetylene		
Carbon Dioxide		
Carbon Monoxide		
Acid	May indicate Oil Change	
Nitrogen	May Indicate Oil Filtering/Change	
Oxygen	May Indicate Oil Filtering/Change	
Total Furfuaral and Furans	May Indicate Failure	
Moisture	May Indicate Oil Filtering/Change, or Reason for Drop in Gas Concentration	Discrete
Dielectric Strength	May Indicate Oil Filtering/Change	
Top Oil Temperature	Proportional to Moisture Content	
Colour	May Indicate Oil Change	
Dirt Sediment	May Affect Dielectric Strength	
Oil Level	May Indicate Excess Gassing	
Oil Leaks	May Indicate Excess Gassing	

Table 3: Encoding of Discrete Features

Feature	Raw Data Type	Encoding Method
Moisture	Float	No encoding required
Dielectric Strength	Float	No encoding required
Top Oil Temperature	Float	No encoding required
Colour	Scale from 1.0 to 4.0	No encoding required
Dirt Sediment	‘None’ or ‘Yes’ or ‘Carbon Residue’	0.0 or 1.0 or 2.0
Oil Level	‘Normal’ or ‘Low’	1.0 or 0.0
Oil Leaks	‘Yes’ or ‘No’	1.0 or 0.0

3.1.2 Exponential Smoothing and Resampling

As part of the preprocessing of the data, an exponential smoothing method was utilised. The function of smoothing is to eliminate unwanted noise that could negatively impact the accuracy of the forecasts. The source of this noise is inconsistencies in the physical testing of the transformer oil; for example, in the case of multiple sampling, all values are present in the data instead of potentially erroneous ones being omitted. Not enough information was gathered at the time of data collection in order to quantify the extent or scale of the noise.

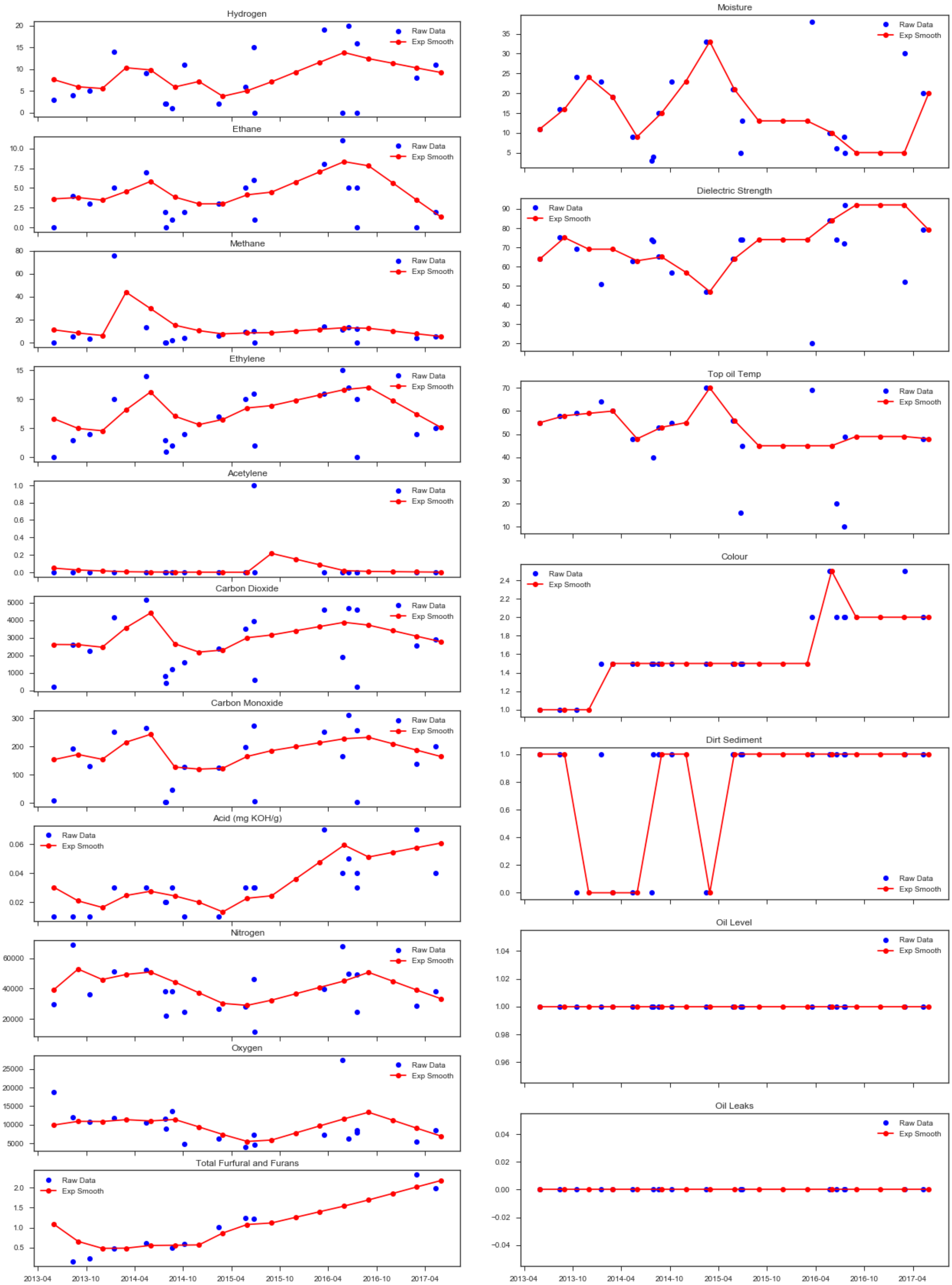
With the data in question, the challenge was finding an appropriate smoothing method for aperiodically sampled data, such that noise is smoothed out but the overall variation of the data is maintained. This specific formulation is recommended for this application, since it is used to analyse discrete events, such as failures [4]. Other works incorporating this smoothing technique on irregularly sampled data include papers in the fields of biology [22], data analytics [23], and mechanical engineering [24].

The exponential moving average method uses a rolling time series operator with a rolling window of fixed length $\tau > 0$. Let $\mathbf{a} = \left\{a_{t_1}, a_{t_2}, ..., a_{t_{N(\mathbf{a})}}\right\}$ denote an unevenly spaced time series, with observation times $T(\mathbf{a}) = \{t_1, ..., t_{N(\mathbf{a})}\}$, and length $N(\mathbf{a})$. For a time point, $t_i \in \mathbb{R}$, a_{t_i} denotes the most recent observation at or before time t_i. Samples before t_1 are taken to be a_{t_1}. The exponential moving average can be expressed as

$$\mathrm{EMA}(a,\tau)_t = \begin{cases} a_{t_1} & \text{, if } t = t_1, \\ \left(1 - \exp\left(-\frac{\Delta t_i}{\tau}\right)\right) a_{t_i} + \exp\left(-\frac{\Delta t_i}{\tau}\right) \mathrm{EMA}(a,\tau)_{t_{i-1}} & \text{, if } t > t_1, \end{cases} \tag{14}$$

where Δt_i is the time difference between successive points. The free parameter τ is effectively a scaling parameter that dictates the degree of variability that is conserved between the original data and the smoothed data. Because the target data sampling period is quarterly, the parameter τ is chosen to be 4, even though it is not period dependent. This allows the exponent in the smoothing operation to be raised to the power of the number of quarters between successive data points. Once the exponential smoothing was complete, linear interpolation was used to transform the continuous features to a fixed interval, and forward filling was used to transform the discrete features. The data was then normalised appropriately before further processing. This exponentially smoothed data is used in the SVM, and the corresponding time series of each individual feature is denoted by $\bar{\mathbf{a}}$.

Figure 3 illustrates a time series plot of all the features for a particular transformer, and shows the effect of smoothing and resampling on the data.



(a) (b)

Figure 3: Time Series Example of a) Continuous, and b) Discrete Features

The blue points represent the raw data obtained from the transformer company, and the red points and line represent the smoothed and resampled points. In an effort to mitigate the effects of the first points in the time series being outliers, the first point of each time series was chosen to be the mean value of that time series, i.e. $\overline{a}_{t_1} = \frac{a_{t_1}+a_{t_2}+...+a_{t_{N(\mathbf{a})}}}{N(\mathbf{a})}$. From the figure, it is clear that some of the raw data is noisy, and therefore exponential smoothing was necessary. The aperiodic nature of the raw data is also apparent.

3.1.3 Correlation Between Features

Figure 4 shows the scatter matrix of the seven failure gases for a single transformer, but instead of considering the correlation between features at the same time step, the correlation between lagged values is considered. On the y-axis, the feature value is plotted, and on the x-axis, the corresponding first lag is plotted. The lag, or backshift, operator operates on an element of a time series such that it produces the previous element in that time series:

$$\mathbb{B}\overline{\mathbf{a}}_{t_n} = \overline{\mathbf{a}}_{t_{n-1}}, \tag{15}$$

or similarly, keeping consistent with the labels in Figure 4:

$$\mathbb{B}\text{Feature} = \text{Feature}_{\text{lag}_1}. \tag{16}$$

There is definite structure in the lagged feature scatter matrix, and there is therefore definite correlation between the first lags of some of the features, and the current values. For this reason, the first lags of every feature forms part of the input vector of the model, in order to take advantage of the evident correlation in the lags of the features. The second and third lags of the additional features were also considered, but no apparent correlation existed between the features. These lags were therefore not included in the model.

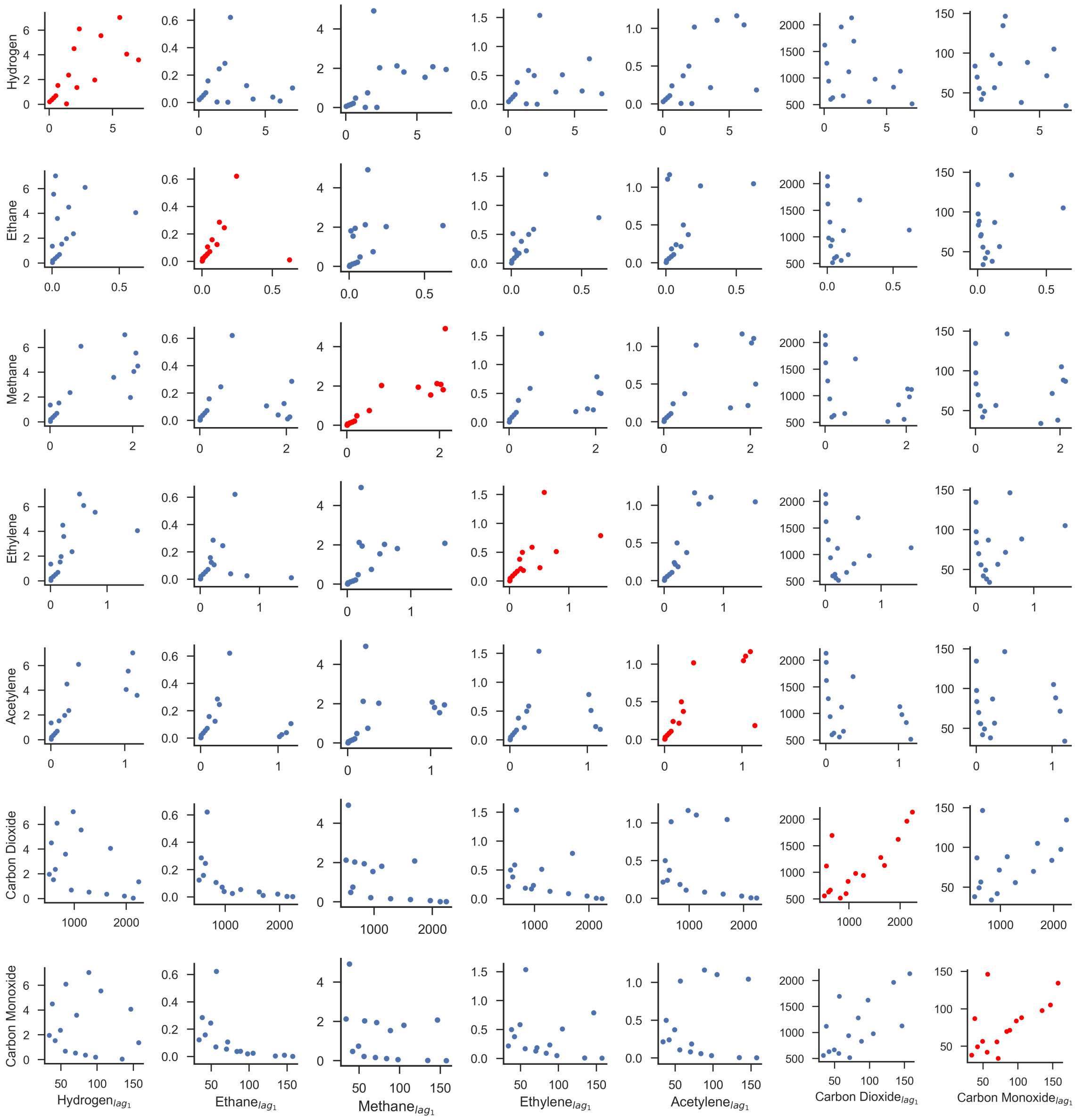


Figure 4: Scatter Matrix of Lagged Feature Correlations in a Single Transformer

3.1.4 Additionally Engineered Features

With the intention of assisting the forecasting techniques with the accuracy of their predictions, additional features were added to the data in order to exploit any potential non-linearities. Using the empirical knowledge provided by the Duval Triangles and Pentagons [5, 6], it is known that there is some relationship between the ratios of some of the key gasses and mechanisms of failure. These eleven ratios were augmented on to the original input vectors of the model, in the hope that they may improve the forecasting accuracy. These are all the ratios used in the empirical method of classifying failures. The ratios provided were as follows:

- $\frac{CH_4}{H_2}$
- $\frac{C_2H_4}{C_2H_6}$
- $\frac{C_2H_2}{CH_4}$
- $\frac{C_2H_4}{H_2}$
- $\frac{C_2H_2}{C_2H_6}$
- $\frac{CO}{CO_2}$
- $\frac{C_2H_6}{H_2}$
- $\frac{C_2H_6}{CH_4}$
- $\frac{C_2H_4}{CH_4}$
- $\frac{C_2H_2}{C_2H_4}$
- $\frac{C_2H_2}{H_2}$

For a fair comparison, the experiments were run with both the original input vectors, and the input vectors augmented with the gas ratios. These two input vectors will be refered to as the original and augmented input vectors respectively. The results from the experiments were compared directly.

3.2 Procedure for Forecasting

The LS-SVM model has been fully defined for a general case. The procedure for utilising the model for forecasting purposes, as well as the procedure for optimising the model's hyper-parameters, is therefore specified. The procedure is fairly common when forecasting DGA results using machine learning techniques [7, 8]. The difference in this work is the inclusion of additional features in the input vector, in order to take advantage of the evident relationship between features.

Consider the time series of the j^{th} feature for a certain transformer, $\bar{\mathbf{a}}^j = \{\bar{a}_1, ..., \bar{a}_m, ..., \bar{a}_n\}$, where $\bar{a}_i,\ i \in \{1, ..., n\}$, is the smoothed, resampled, normalised data point obtained from the pre-processing procedure, and j indicates the feature to be forecast, where $j \in \{1, ..., r\}$. There are seven features of interest - the key gases in Table 2 - therefore $r = 7$. It is reiterated that the model training samples are expressed as $\{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\} \subseteq \left(\mathbb{R}^d \times \mathbb{R}\right)$. The time series $\bar{\mathbf{a}}$ needs to be split into input vectors, $\mathbf{x}_i$, and target values, $y_i,\ i \in \{1, ..., l\}$.

Multivariate time series forecasting is performed. The input vectors consisted of the concatenation of m-lagged values of the target feature, and 1-lagged values of all other features in Table 2. The hyper-parameter m represents how many lagged points of feature j are chosen to be in the input vector when training the model, and is tuned using the PSO algorithm.

The formulation of the input vector can be expressed mathematically; let the first m entries of the j^{th} feature be expressed as $\bar{\mathbf{a}}_1^j = \{\bar{a}_1, ..., \bar{a}_m\}$. The vector $\bar{\mathbf{a}}_1^j$ is the time series vector, and forms the first m entries of the input vector $\mathbf{x}_1$. The next $(d - m)$ entries of the input vector are the 1-lagged values of the non-target features. This feature vector can be expressed mathematically as $\mathbf{b}_1^j = \left\{b_m^{k_1}, b_m^{k_2}, ..., b_m^{k_p}\right\}$, where the k_p^{th} feature is one such that $k_p \neq j$, and

$p \in \{1, ..., M-1\}$, where M is the total number of available features. For the experiments conducted with original features only, $M = 18$ (seven key gasses and eleven additional features). For the experiments conducted with the augmented input vectors, $M = 29$ (eighteen features from the original set, and eleven engineered features). The corresponding target value of the model is $y_1 = \{\overline{a}_{m+1}\}$.

The input vector is the concatenation of the time series vector and the feature vector, expressed as $\mathbf{x}_1 = \left[\overline{\mathbf{a}}_1^j, \mathbf{b}_1^j\right]$. This can be expressed in general for the i^{th} input vector ($i \in \{1, ..., l\}$) as $\mathbf{x}_i = \left[\overline{\mathbf{a}}_i^j, \mathbf{b}_i^j\right]$, where $\overline{\mathbf{a}}_i^j = \{\overline{a}_i, ..., \overline{a}_{i+m-1}\}$ and $\mathbf{b}_i^j = \left\{b_{i+m-1}^{k_1}, ..., b_{i+m-1}^{k_p}\right\}$. In general, the target value of the model is given as $y_i = \{\overline{a}_{i+m}\}$. Validation and Testing sets differed based on which experiment was being conducted, and are specified in Sections 3.3 and 3.4 respectively.

The total number of hyper-parameters in the formulation is therefore three; the regularisation parameter in the LS-SVM model, C; the free parameter in the RBF kernel function, γ; and the number of lagged points in the input vector of the model, m. Each feature has its own set of hyper-parameters. These three hyper-parameters are optimised across all transformers using the PSO algorithm. The fitness function used in the PSO is the mean squared error between the forecast value and the actual value in the Validation set. For v validation points, the fitness of the i^{th} particle at the t^{th} time step can be expressed as:

$$\text{Fitness}_i(t) = \frac{1}{v}\sum_{k=1}^{v}\left(y_k - \hat{f}\left(\boldsymbol{x}_k\right)\right)^2, \tag{17}$$

where y_k is the target value that is to be approximated, and $\hat{f}\left(\boldsymbol{x}_k\right)$ is the model's estimate of the target value.

3.3 Experiment 1 - Leave-One-Out Across Transformers

The data provided from the transformer company consists of data from one of their clients, with 18 transformers. In all previously found literature, the maximum number of transformers tested has been four, and all transformers were treated as completely independent, without any hyper-parameters shared across transformers [7–11, 25]. As stated previously, the difference between experiment types in this research is in how the Validation and Testing sets are chosen. The methods used for choosing how Validation and Testing sets are chosen is similar to that specified by Hyndman [26]. This is further discussed in Section 6.

In Experiment 1, the models were trained, and the hyper-parameters optimised per feature across all transformers. The data of a whole single transformer was held-out as the Testing set. The Training set of the j^{th} feature is given as:

$$\text{Training: } \left\{ (\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l) \right\}_j . \tag{18}$$

For Experiment 1, the number of training pairs is given as $l = n - m - v$. The Validation set of the j^{th} feature is defined to be:

$$\text{Validation: } \left\{ \left(\mathbf{x}_{(n-m-v)+1}, y_{(n-m-v)+1}\right), ..., (\mathbf{x}_{n-m}, y_{n-m}) \right\}_j . \tag{19}$$

In the method used for Experiment 1, the number of points held-out for Validation varied between one and four, i.e. $v \in \{1, ..., 4\}$. This effectively increased the amount of data used to optimise the hyper-parameters of the model. As stated previously, the Training and Validation sets consisted of data from all transformers except one, which was held out for Testing. The Testing set can therefore be described in a similar manner to the Validation set:

$$\text{Testing: } \left\{ \left(\mathbf{x}_{(n-m-h)+1}, y_{(n-m-h)+1}\right) \right\}_j , \tag{20}$$

where h is the number of points held out for Testing on the Testing transformer, $h \in \{1, ..., 4\}$. The approach is still a Leave-One-Out (LOO) approach, since only one point is forecast per instance of the experiment. Increasing the number of points being held out effectively increases the size of the Testing set. Each of the 18 transformers were held out as the Testing set in turn. Due to the stochastic nature of the optimisation algorithm, each experiment was repeated 20 times, in order to obtain an estimate of the distribution of the forecasting errors.

Figure 5 is a visual representation of how the Training, Validation, and Testing sets were defined for the j^{th} feature in Experiment 1. For a total number of T transformers, the I_T^{th} transformer is withheld as the Testing set, and all the other transformers are used for Training and Validation.

$$\text{Transformer}_{I_1} = \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\}_j$$

$$\text{Transformer}_{I_2} = \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\}_j$$

$$\vdots$$

$$\text{Transformer}_{I_{T-1}} = \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\}_j$$

(a) Training Set

$$\text{Transformer}_{I_1} = \left\{\left(\mathbf{x}_{(n-m-v)+1}, y_{(n-m-v)+1}\right), ..., (\mathbf{x}_{n-m}, y_{n-m})\right\}_j$$

$$\text{Transformer}_{I_2} = \left\{\left(\mathbf{x}_{(n-m-v)+1}, y_{(n-m-v)+1}\right), ..., (\mathbf{x}_{n-m}, y_{n-m})\right\}_j$$

$$\vdots$$

$$\text{Transformer}_{I_{T-1}} = \left\{\left(\mathbf{x}_{(n-m-v)+1}, y_{(n-m-v)+1}\right), ..., (\mathbf{x}_{n-m}, y_{n-m})\right\}_j$$

(b) Validation Set

$$\text{Transformer}_{I_T} = \left\{\left(\mathbf{x}_{(n-m-h)+1}, y_{(n-m-h)+1}\right)\right\}_j$$

(c) Testing Set

Figure 5: Experiment 1 Set Division

3.4 Experiment 2 - Leave-One-Out Within Transformers

Experiment 2 is similar to Experiment 1 in the way that the hyper-parameters are optimised for a single feature across all transformers. The difference was in how the Testing set was defined; the final h points of each transformer were held out, $h \in \{1, ..., 4\}$. This was expected to outperform Experiment 1, since the Testing set wasn't from an entirely unseen transformer. As before, the Training set of the j^{th} feature is given as:

$$\text{Training: } \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\}_j . \tag{21}$$

For Experiment 2, the number of training pairs is given as $l = n-m-v-h$. The data remaining after the Testing set was withheld was split 75%-25% between Training and Validation sets respectively. The Validation set of the j^{th} feature is defined to be:

$$\text{Validation:}\ \left\{\left(\mathbf{x}_{(n-m-v-h)+1}, y_{(n-m-v-h)+1}\right), ..., \left(\mathbf{x}_{n-m-h}, y_{n-m-h}\right)\right\}_j. \tag{22}$$

In the method used for Experiment 2, the number of points held-out for Validation was $v \in \{1, ..., \lfloor 0.25 * (n - m - h) \rfloor\}$. As stated previously, the Training and Validation sets consisted of data from all transformers. The Testing set can therefore be described in a similar manner to the Validation set:

$$\text{Testing:}\ \left\{\left(\mathbf{x}_{(n-m-h)+1}, y_{(n-m-h)+1}\right)\right\}_j, \tag{23}$$

The approach is still a LOO approach, since only one point is forecast per instance of the experiment. Increasing the number of points being held out effectively increases the size of the Testing set. Due to the stochastic nature of the optimisation algorithm, each experiment was repeated 30 times, in order to obtain an estimate of the distribution of the forecasting errors.

Figure 6 is a visual representation of how the Training, Validation, and Testing sets were defined for the j^{th} feature in Experiment 1. For a total number of T transformers, the final h points of each transformer are held out as a Testing set.

$$\begin{aligned}
\text{Transformer}_{I_1} &= \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\}_j \\
\text{Transformer}_{I_2} &= \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\}_j \\
&\vdots \\
\text{Transformer}_{I_T} &= \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_l, y_l)\}_j
\end{aligned}$$

(a) Training Set

$$\begin{aligned}
\text{Transformer}_{I_1} &= \left\{\left(\mathbf{x}_{(n-m-v-h)+1}, y_{(n-m-v-h)+1}\right), ..., \left(\mathbf{x}_{n-m-h}, y_{n-m-h}\right)\right\}_j \\
\text{Transformer}_{I_2} &= \left\{\left(\mathbf{x}_{(n-m-v-h)+1}, y_{(n-m-v-h)+1}\right), ..., \left(\mathbf{x}_{n-m-h}, y_{n-m-h}\right)\right\}_j \\
&\vdots \\
\text{Transformer}_{I_T} &= \left\{\left(\mathbf{x}_{(n-m-v-h)+1}, y_{(n-m-v-h)+1}\right), ..., \left(\mathbf{x}_{n-m-h}, y_{n-m-h}\right)\right\}_j
\end{aligned}$$

(b) Validation Set

$$\begin{aligned}
\text{Transformer}_{I_1} &= \left\{\left(\mathbf{x}_{(n-m-h)+1}, y_{(n-m-h)+1}\right)\right\}_j \\
\text{Transformer}_{I_2} &= \left\{\left(\mathbf{x}_{(n-m-h)+1}, y_{(n-m-h)+1}\right)\right\}_j \\
&\vdots \\
\text{Transformer}_{I_T} &= \left\{\left(\mathbf{x}_{(n-m-h)+1}, y_{(n-m-h)+1}\right)\right\}_j
\end{aligned}$$

(c) Testing Set

Figure 6: Experiment 2 Set Division

The main experiments conducted for the research are therefore listed as follows:

- **LS-SVM(1)** - Least Squares Support Vector Machine trained using the method specified by Experiment 1 in Section 3.3.
- **LS-SVM(2)** - Least Squares Support Vector Machine trained using the method specified by Experiment 2 in Section 3.4.
- **ϵ-SVR(2)** - SK-Learn's in-built support vector regression [27] trained using the method specified by Experiment 2 in Section 3.4.

All experiments are conducted for both the input vectors containing original features only, as well as input vectors containing original features augmented with the ratios of different gases, specified in Section 3.1.4.

3.5 Baseline Models for Comparison

Several standard methods are utilised as baseline comparisons, in order to gauge the effectiveness of the newly developed models. The testing results of each method are compared with those of the relevant experiment, both qualitatively and quantitatively. The following forecasting methods are used as baseline comparators:

- **Mean** - Forecast point is the cumulative mean of the feature values up until that point.
- **Naive** - Forecast point is equal to the previous value in the time series.
- **Univariate LS-SVM (LS-SVM-UV(2))** - Least Squares Support Vector Machine trained using the method specified by Experiment 2 in Section 3.4, except hyper-parameters are trained on a per-transformer basis - not across transformers; and only the time series portion of the input vector is used in the LS-SVM ($\mathbf{x}_i = \bar{\mathbf{a}}_i^j$). The additional feature vector augmentation ($\mathbf{b}_i^j$) is omitted. These results are directly comparable with those produced in other published works [7–9] since the method used mimics methods in previous research.
- **ARIMA** - An in-built Autoregressive Integrated Moving Average forecaster from Python's Statsmodels library was used as a baseline for comparison [28]. Various models were trained and validated, and the best model per feature per transformer was taken as the forecast value.
- **Exponential Smoothing (E.S.)** - An in-built Exponential Smoothing Forecaster from Python's Statsmodels library was used as a baseline for comparison [28]. This forecasting method uses simple exponential smoothing, and optimises two hyper-parameters by splitting the data into Training, Validation, and Testing.

The method specified by Experiment 2 is easily implementable across all types of models and baseline tests specified, and it mimics a more traditional method of splitting time series data [26]. The testing procedure consisted of obtaining and comparing results from all models using Experiment 2. The best model was then chosen and trained using the method specified by Experiment 1. This would give a sense of the extent of the generalising ability of the model.

4 Results and Discussion

The research thus far specifies methods used to forecast single data points per feature, using differing experiment types. The data from 18 power transformers of various sizes was taken, and split accordingly into Training, Validation and Testing sets. The data was split differently according to the method used and the type of experiment that was run. For a detailed explanation on how the data was split for each experiment, refer to Sections 3.3, and 3.4.

4.1 Optimisation Results

In LS-SVM(1), LS-SVM(2), and LS-SVM-UV(2), the method of choosing hyper-parameters (C, γ, and m) was by Particle Swarm Optimisation (PSO). This was done by splitting the data into Training, Validation and Testing sets, and optimising the hyper-parameters according to the Validation set. The methods used in choosing how the sets were split in each experiment were specified in Sections 3.2 and 3.5.

Since there are three hyper-parameters that require optimisation, the dimensionality of the PSO is three. The initial population was generated randomly using upper and lower bounds for each of the three hyper-parameters. Equation 24 illustrates how the i^{th} member of the population was generated at the $t = 0$ time step. The parameter, *rand*, represents a number drawn from a uniform distribution, i.e. rand $\sim U(0, 1)$, and the function, *round()*, rounds off the input to its nearest integer.

$$\mathbf{P}_i(0) = \begin{bmatrix} C_{min} + (C_{max} - C_{min}) * \text{rand} \\ \gamma_{min} + (\gamma_{max} - \gamma_{min}) * \text{rand} \\ \text{round}\left(m_{min} + (m_{max} - m_{min}) * \text{rand}\right) \end{bmatrix}^T, \in (\mathbb{R} \times \mathbb{R} \times \mathbb{N}) \tag{24}$$

Equation 12, utilised in the PSO, is reiterated for convenience:

$$\begin{cases} \mathbf{v}_i(t+1) & = \lambda \mathbf{v}_i(t) + r_1 \mathbf{c_1} \left(\mathbf{P}_{best}(t) - \mathbf{P}_i(t)\right) + r_2 \mathbf{c_2} \left(\mathbf{g}_{best}(t) - \mathbf{P}_i(t)\right) \\ \mathbf{P}_i(t+1) & = \mathbf{P}_i(t) + \mathbf{v}_i(t+1) \end{cases}.$$

The values of the parameters used in the PSO are summarised in Table 4.

Table 4: Summary of Parameters used in PSO Algorithm

Parameter	**Value**
S	20
G	100
$\mathbf{c_1}$	[0.5, 0.5, 2.0]
$\mathbf{c_2}$	[0.5, 0.5, 2.0]
λ_{min}	0.4
λ_{max}	0.9

Parameter	**Value**
C_{min}	0.0
C_{max}	5000.0
γ_{min}	10.0
γ_{max}	200.0
m_{min}	1
m_{max}	5

The fitness function chosen to evaluate the fitness of each particle in the swarm is the mean squared error of the validation points. For v validation points, the fitness of the i^{th} particle at the t^{th} time step can be expressed as:

$$\text{Fitness}_i(t) = \frac{1}{v}\sum_{k=1}^{v}\left(y_k - \hat{f}\left(\boldsymbol{x}_k\right)\right)^2, \tag{25}$$

where $\boldsymbol{x}_v$ is the input vector to the LS-SVM, y_v is the target value, and $\hat{f}\left(\boldsymbol{x}_v\right)$ is the model's estimate of the target value.

LS-SVM(1) was run 20 times per holdout transformer, and both LS-SVM(2) and LS-SVM-UV(2) were run 30 times per holdout point. From the repetitions, a distribution of errors could be estimated. Figure 7 illustrates the convergence to an optimal solution for a single experiment. The figure shows the overall convergence of the population on a local minimum, as well as how the fitness of the global best member of the population changed throughout the experiment.

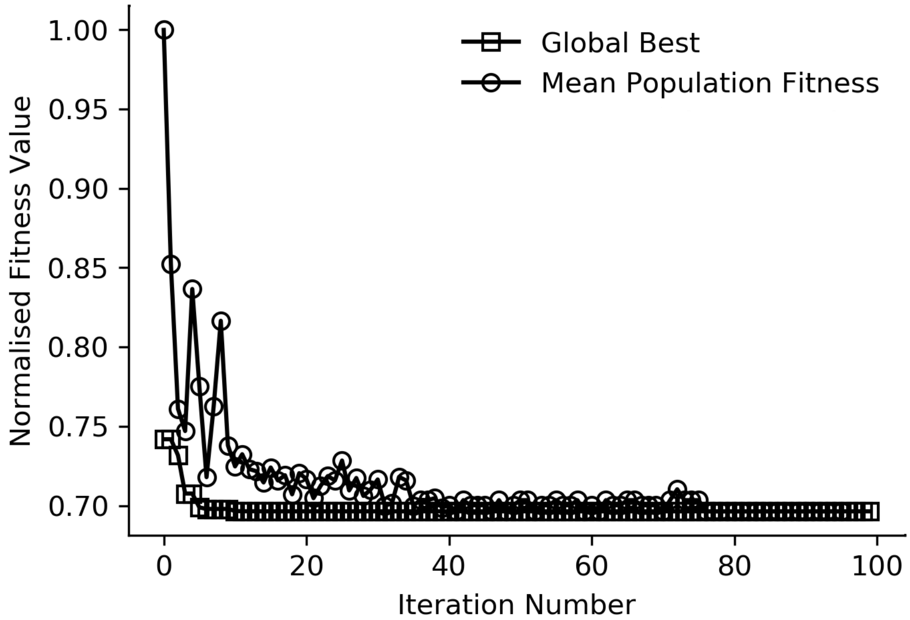


Figure 7: Convergence Curves for Hydrogen for a Particular Transformer

In this particular case, it is clear that a minimum solution was found after a small number of iterations (less than 20). This is indicated by the constant global best solution, where the fitness doesn't change in the later iterations. However, due to the stochastic nature of the problem, it is important that the experiment run to completion, in case a better solution is found in later iterations.

4.2 Validation and Testing Errors

Table 5 shows the mean Validation and Testing MSE's across all transformers per feature, for both original and augmented input vectors. The highlighted cells in the table indicate which model performed best for each feature in the Testing experiments. The purpose of comparing Validation and Testing errors within experiments is to determine whether models have been over-fit to the particular dataset, and also to observe whether the errors in each set are as expected relative to one another.

It is noted that in general, the Validation and Testing errors for LS-SVM(1) are such that the Validation errors are smaller than the Testing errors, but still within one order of magnitude. This is a positive result because it rules out model over-fitting in the experiment.

If the method behind LS-SVM(1) is considered, the difference between Validation and Testing errors is expected; the Testing set was a completely unseen transformer, potentially with unique characteristics that the model hadn't been trained to handle. In cases like this, the performance of this model would be comparable to a Mean or Naive forecast, but this is case-specific, and difficult to generalise. The phenomenon would be more easily observable in a time series plot.

The relatively low Validation errors of LS-SVM(1) are also expected when the method is considered; LS-SVM(1) had both the largest Training set per transformer, as well as potentially the largest Validation set when compared to the method of Experiment 2. The model therefore has enough information to properly fit the data, without over-fitting. This well-fit Validation set is a useful result, as it gives a lower bound on the error of an unseen transformer, assuming nothing about the transformer history is known. A potential application of this model would be for a brand new transformer, where the current number of data points is too few to train a new model, but enough such that a forecast is possible.

In LS-SVM(2), in general, the Testing errors are smaller than the Validation errors, but are still within one order of magnitude. This is an unusual result, since the Validation set is always expected to outperform the Testing set. Possible causes of this phenomenon could be due to the design of the Experiment 2 method; if a large number of points were held out for testing relative to the number of points used for validation, the Validation set would be smaller than the Testing set. This could possibly imply under-training of the model, and a better result could potentially be obtained with a larger Validation set. Despite this inconsistency, LS-SVM(2) generally outperforms all other models in Testing.

Both the ϵ-SVR(2) and LS-SVM-UV(2) models display signs of over-fitting in some cases. These are the cases where the Validation and Testing errors differ by over one order of magnitude. Both of these models were trained with the method specified by Experiment 2. The fact that neither of them performed well relative to the LS-SVM models indicates that neither of them are a good model choice for the forecasting of this particular data.

Table 5: Mean Validation and Testing Error Summary across All Transformers

Input Type	**Feat**	**LS-SVM (Exp. 1)**				**LS-SVM (Exp. 2)**				**ϵ-SVR (Exp. 2)**				**LS-SVM-UV (Exp. 2)**			
		Valid		**Test**		**Valid**		**Test**		**Valid**		**Test**		**Valid**		**Test**	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
Original	H_2	0.181	0.396	0.425	0.652	1.112	1.050	0.278	0.521	2.266	1.505	2.829	1.675	6.412	2.532	160.8	12.67
	C_2H_6	0.304	0.501	0.534	0.707	0.383	0.575	0.217	0.420	0.370	0.562	0.978	0.968	64.50	8.030	4.182	2.026
	CH_4	0.072	0.228	0.186	0.426	0.198	0.376	0.053	0.207	0.427	0.590	0.176	0.417	11.74	3.420	5.437	2.263
	C_2H_4	0.584	0.735	0.632	0.776	0.365	0.588	0.672	0.773	0.398	0.603	1.183	1.053	2.733	1.627	7.135	2.668
	C_2H_2	0.164	0.389	1.451	1.204	0.885	0.941	0.510	0.702	10.93	3.304	456.2	21.35	99.54	9.935	226.1	15.04
	CO_2	0.326	0.548	0.495	0.680	0.387	0.617	0.331	0.535	1.798	1.315	2.790	1.667	796.3	28.22	278.0	16.67
	CO	0.283	0.531	0.348	0.586	1.046	1.008	0.224	0.472	2.417	1.551	2.695	1.640	15.17	3.889	17.99	4.239
Augmented	H_2	0.189	0.435	0.638	0.785	1.056	1.026	0.336	0.574	1.621	1.273	1.199	1.080	6.412	2.532	160.8	12.67
	C_2H_6	0.350	0.567	0.579	0.750	0.420	0.553	0.290	0.520	1.331	1.140	43.12	6.556	64.50	8.030	4.182	2.026
	CH_4	0.044	0.184	0.094	0.302	0.156	0.333	0.058	0.232	0.101	0.274	78.04	8.808	11.74	3.420	5.437	2.263
	C_2H_4	0.353	0.573	0.614	0.775	0.203	0.424	0.748	0.837	0.213	0.435	0.890	0.917	2.733	1.627	7.135	2.668
	C_2H_2	0.320	0.564	2.630	1.622	0.719	0.847	0.533	0.706	2.706	1.622	335.2	18.31	99.54	9.935	226.1	15.04
	CO_2	0.290	0.510	0.377	0.604	0.377	0.597	0.349	0.584	0.797	0.884	1.111	0.981	796.3	28.22	278.0	16.67
	CO	0.276	0.514	0.313	0.559	0.849	0.901	0.263	0.513	0.986	0.985	0.414	0.626	15.17	3.889	17.99	4.239

4.3 Visualisation of Testing Errors

The visualisation of prediction error distributions in each experiment is useful because it can immediately be seen which features are performing better. Desirable properties of the distributions include centring around zero, a relatively small standard deviation, and an approximately normal distribution. Figure 8 shows the prediction error histograms plotted for Hydrogen in Experiment 1. The prediction errors are plotted for this feature using the two specified input types. The resulting prediction error distributions are compared with each other, as well as error distributions obtained from the Mean and Naive forecasts. In Figure 8, the left hand figure was obtained using only the original input vector, and the right hand figure was obtained using the augmented input vector. For the unabridged set of results, refer to Appendix A.1.1.

The resulting prediction error distributions from both original and augmented input vectors show high error densities close to zero. The original input vector results display a relatively large bias. The standard deviation of the original input vector distribution is significantly lower than that of the augmented input vector distribution. The error density around the mean is higher for the original input vector distribution. It is inconclusive to say which set of input vectors performs better without having considered quantitative methods. These are discussed in Section 4.4.4.

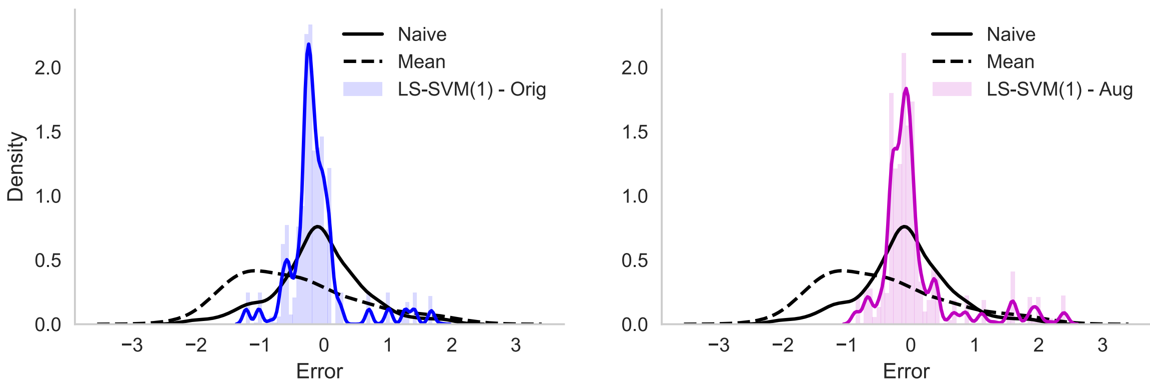


Figure 8: LS-SVM(1) Testing Error Distribution - Hydrogen

A notable property in general is the spike in density of the outliers for various discrete values. The increased density of these outliers may indicate that for a particular testing transformer, the set of hyper-parameters found to be optimal were identical for several instances of the experiment, and hence the testing errors for this particular parameter set were identical for each instance. This is not an ideal occurrence.

The purpose of running repeated experiments with a global optimiser with different initialisations is to find multiple solutions to a problem. This procedure is done in order to obtain a good minimum as opposed to many local minima in the space. Using multiple runs, it is useful to obtain an error distribution; ideally one that is normally distributed. An error density function that exhibits a complex mix of distributions is undesirable, since useful statistical tools in forecasting, such as confidence intervals, are based on normal error distributions. The occurrence of arbitrary error density functions could potentially be mitigated by having a cleaner set of data to work with, or by changing the fitness function of the optimiser to one that punishes outliers to a greater extent.

4.4 Comparison of Methods

Table 6 shows the Mean Squared Error results obtained for the Testing errors of all experiments conducted and all baseline tests. The table shows the mean MSE's and standard deviations of errors obtained across all transformers, for both original and augmented input vectors. The highlighted cells indicate which model performed best for each feature, for the same input vector type (either the original input vector, or the augmented input vector). The entire set of results can be found in Appendix A.2.3.

From the table, it can be observed that LS-SVM(2) generally outperformed all other models in the majority of cases, for both input vector types, except for one feature in the original input case and two features in the augmented input case where it was beaten by the Naive forecast. This is an interesting result. In general, if the Naive forecast beats all other forecasts, the data is a random walk, with no discernible pattern or trend that could be identified by the models.

Table 6: Results Summary - Mean Testing Results for Original and Augmented Features

Input Type	**Feat**	**LS-SVM(1)**		**LS-SVM(2)**		**ϵ-SVR(2)**		**LS-SVM-UV(2)**		**ARIMA**		**E.S.**		**Naive**		**Mean**	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
	H_2	0.425	0.652	0.278	0.521	2.829	1.675	160.8	12.67	491.7	21.08	48.17	6.939	0.626	0.790	1.213	1.016
	C_2H_6	0.534	0.707	0.217	0.420	0.978	0.968	4.182	2.026	$22e^{+3}$	141.2	37.78	6.142	0.598	0.771	1.256	1.013
	CH_4	0.186	0.426	0.053	0.20	0.176	0.417	5.437	2.263	1122	32.07	250.4	15.72	0.673	0.816	1.257	0.989
Original	C_2H_4	0.632	0.776	0.672	0.773	1.183	1.053	7.135	2.668	$21e^{+3}$	142.0	1040	31.88	0.564	0.750	1.252	1.018
	C_2H_2	1.451	1.204	0.510	0.702	456.2	21.35	226.1	15.04	3650	57.03	10.65	3.262	0.512	0.713	0.895	0.900
	CO_2	0.495	0.680	0.331	0.535	2.790	1.667	278.0	16.67	$87e^{+3}$	295.2	$16e^{+4}$	389.7	0.878	0.936	1.246	1.113
	CO	0.348	0.586	0.224	0.472	2.695	1.640	17.99	4.239	502.7	22.05	630.9	25.09	0.842	0.917	1.235	1.104
	H_2	0.638	0.785	0.336	0.574	1.199	1.080	160.8	12.67	491.7	21.08	48.17	6.939	0.626	0.790	1.213	1.016
	C_2H_6	0.579	0.750	0.290	0.520	43.12	6.556	4.182	2.026	$22e^{+3}$	141.2	37.78	6.142	0.598	0.771	1.256	1.013
	CH_4	0.094	0.302	0.058	0.232	78.04	8.808	5.437	2.263	1122	32.07	250.4	15.72	0.673	0.816	1.257	0.989
Augmented	C_2H_4	0.614	0.775	0.748	0.837	0.890	0.917	7.135	2.668	$21e^{+3}$	142.0	1040	31.88	0.564	0.750	1.252	1.018
	C_2H_2	2.630	1.622	0.533	0.706	335.2	18.31	226.1	15.04	3650	57.03	10.65	3.262	0.512	0.713	0.895	0.900
	CO_2	0.377	0.604	0.349	0.584	1.111	0.981	278.0	16.67	$87e^{+3}$	295.2	$16e^{+4}$	389.7	0.878	0.936	1.246	1.113
	CO	0.313	0.559	0.263	0.513	0.414	0.626	17.99	4.239	502.7	22.05	630.9	25.09	0.842	0.917	1.235	1.104

4.4.1 LS-SVM vs. ϵ-SVR

Other popular support vector regression models, such as ϵ-SVR, have been used as comparative models in previous work [7]. This is a good comparative model to the LS-SVM because they both use the same primal machine learning technique; the only difference between them being how the parameters of the model are solved.

The LS-SVM(2) and ϵ-SVR(2) models may be compared directly because they are both given identical input data - data that was split into Training, Validation, and Testing sets using the method for Experiment 2. The only difference in the models is therefore how the parameters within each model were solved. Comparing these two models therefore indicates which method of solving an SVR is better suited to the data in question.

Figures 9 and 10 show the error distributions of LS-SVM and ϵ-SVR models for the feature Hydrogen. These error distributions were obtained using the original input vector. Figure 9 compares both models with the Mean and Naive forecasting errors, and Figure 10 compares the models directly.

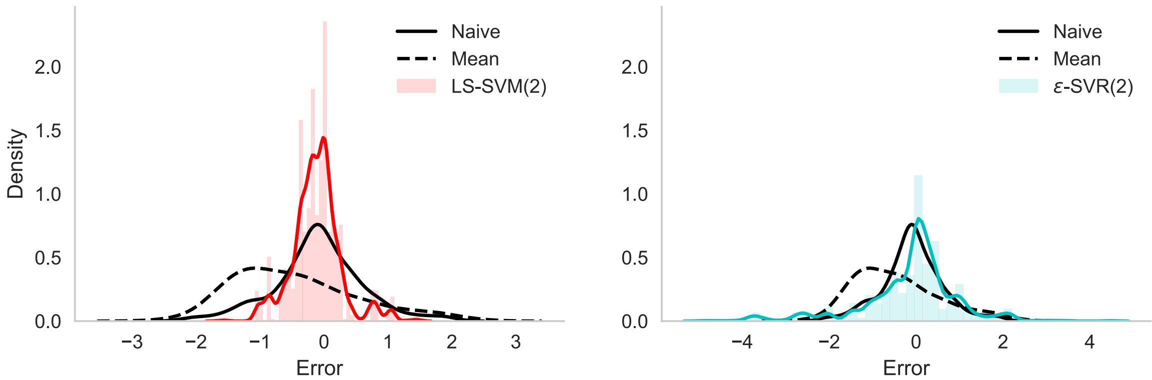


Figure 9: LS-SVM(2) vs ϵ-SVR(2) - Hydrogen

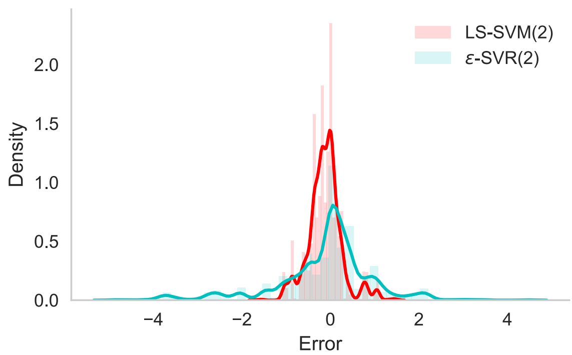


Figure 10: LS-SVM(2) vs ϵ-SVR(2) - Direct Comparison

From Table 6, the mean MSE values for the Testing set of the feature Hydrogen for the LS-SVM(2) and ϵ-SVR models are 0.278 and 2.829 respectively. The LS-SVM(2) model outperformed the ϵ-SVR model by 90.2% according to the mean MSE value of each distribution.

The source of the large MSE value for the ϵ-SVR model is the magnitude of the outliers in its Testing errors. This is clearly observed in Figure 9; the density of large outliers is significantly larger in the ϵ-SVR when compared to the LS-SVM(2) plot. Another noteworthy property of the ϵ-SVR model is its tendency to closely follow the Naive forecast errors.

The difference in standard distributions is clearly noted in Figure 9. The testing errors of the ϵ-SVR model are significantly more spread out when compared to those of LS-SVM(2). Furthermore, the density centred around the mean is significantly higher for LS-SVM(2) than it is for ϵ-SVR. This indicates that in general, the errors are expected to be smaller for LS-SVM(2).

4.4.2 LS-SVM vs. LS-SVM-UV

In all previously encountered works, features within transformers are treated to be very separate, and only univariate time series forecasting has been attempted. Transformers are also treated very separately, and there has never been an attempt to consolidate transformers by having a single set of hyper-parameters between multiple transformers [7–11, 25]. If having a single set of hyper-parameters across transformers yields good forecasting results, this could significantly reduce training time of the models, since the models could be generalised across several transformers.

The focus of this research has been to investigate to what extent the machine learning models in question are generalisable; to investigate if a single set of hyper-parameters can be chosen such that they can be applied to multiple transformers without compromising forecasting accuracy. To investigate the ability of the model to generalise, a model was designed similar to the ones specified in literature [7]. These models treat each transformer separately, i.e. hyper-parameters are trained on a per-transformer basis, and features are treated as completely independent univariate time series. Mathematically, the input vector is reduced to $\mathbf{x}_i = \overline{\mathbf{a}}_i^j$. The additional feature vector augmentation ($\mathbf{b}_i^j$) is omitted. This model is called LS-SVM-UV (Univariate). In this way, the performance of reportedly good models can be measured against the models presented in this research.

Figures 11 and 12 show the error distributions of LS-SVM(2) and LS-SVM-UV(2) for the feature Hydrogen. These error distributions were obtained using the original input vector.

Figure 11 compares both models to the Mean and Naive forecasting errors, and Figure 12 compares the models directly.

These models are directly comparable because they use identical techniques for finding both the parameters and hyper-parameters (LS-SVM). The only difference is how the data is split, and how the hyper-parameters are optimised; LS-SVM(2) allows for multiple features in its input whereas LS-SVM-UV(2) only allows for a single input vector of the feature being forecast. LS-SVM(2) optimises a single set of hyper-parameters across all transformers whereas LS-SVM-UV(2) optimises a different set of hyper-parameters for each transformer.

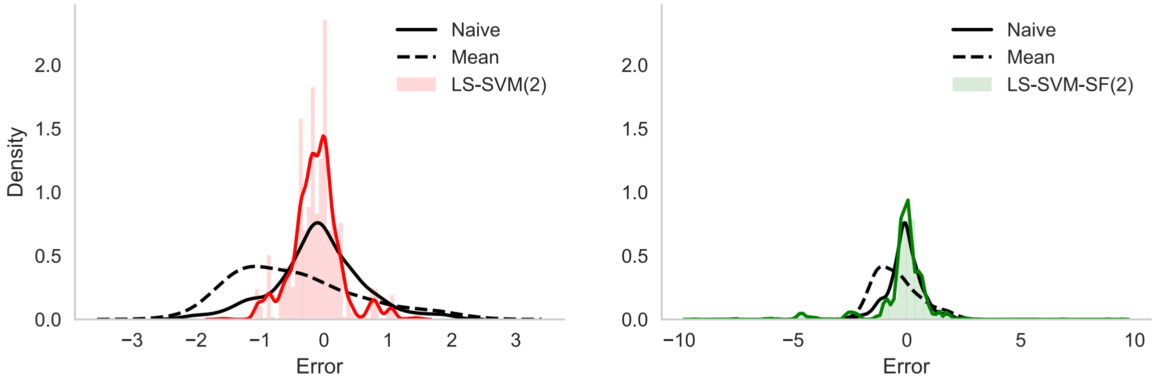


Figure 11: LS-SVM(2) vs LS-SVM-UV(2) - Hydrogen

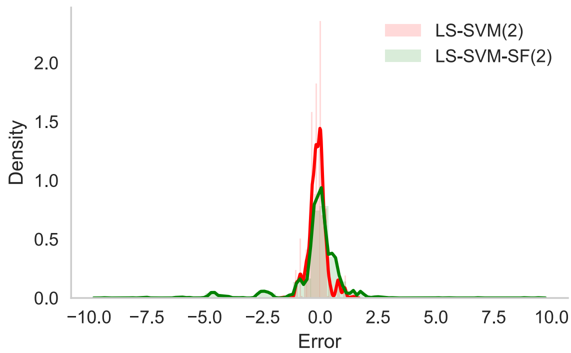


Figure 12: LS-SVM(2) vs LS-SVM-UV(2) - Direct Comparison

From Table 6, the mean MSE values for the Testing set of the feature Hydrogen for the LS-SVM(2) and LS-SVM-UV(2) models are 0.278 and 160.8 respectively. The LS-SVM(2) model outperformed the LS-SVM-UV(2) model by 99.8% according to the MSE value of each distribution.

Despite the fact that the LS-SVM-UV(2) model is reported to perform very well in previous work, it generally does not perform well on the current set of data. The reported range of Mean Absolute Percentage Errors (MAPE) in previous works were between 1% and 10% [7–10]. The current model yields MAPE's of between 170% and 260%. This could be due to several reasons: the resampling techniques used in this research are different from other previously used resampling techniques; the experimental method is slightly different, as previous works did not include both Validation and Testing sets; and the obvious fact that the data is different. It is therefore not surprising that the model did not yield identical results.

4.4.3 LS-SVM - Experiment 1 vs. Experiment 2

The LS-SVM model has been used in two distinct experiments throughout the research. In both Experiments 1 and 2, the LS-SVM was used to forecast single data points per feature. The difference in models is due to how the data was split into Training, Validation, and Testing sets. In Experiment 1, a LOO approach was taken, and the data of an entire transformer was held out as the Testing set. The purpose of running this experiment was to observe how well the model performed on a completely unseen transformer. If this is successfully implemented, training times could be reduced since a new model wouldn't be required for unseen transformers.

In Experiment 2, the final h points of each transformer were held out as the Testing set, $h \in [1 : 4]$. This method would be closer to reality, as it was expected to perform better than Experiment 1 since some of the transformer data had been seen by the model. A major disadvantage of this experiment is that the Validation set is potentially significantly smaller than that of Experiment 1. This could result in model under-training, and the only possible mitigation is obtaining more data, or redesigning the experiment.

Figures 13 and 14 show the testing error distributions for the feature Ethane, for both LS-SVM models. The left hand distribution is for Experiment 1, and the right hand distribution is for Experiment 2. The distributions are from the results obtained using the original input vector - not the input vector augmented with gas ratios.

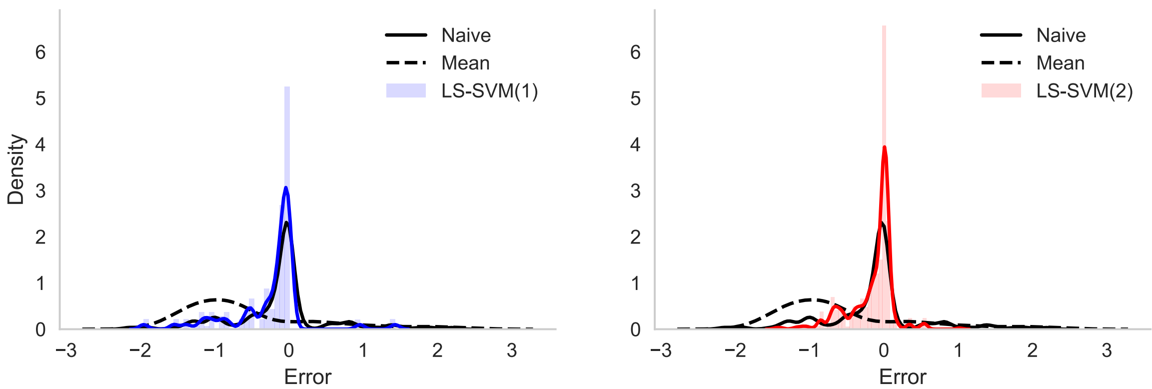


Figure 13: LS-SVM(1) vs LS-SVM(2) - Ethane

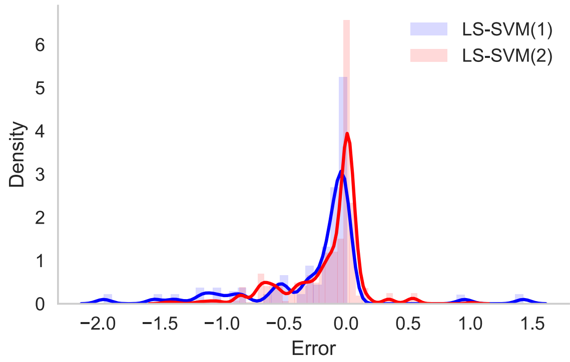


Figure 14: LS-SVM(1) vs LS-SVM(2) - Direct Comparison

From Table 6, the Original Input experiment yielded mean MSE values of 0.534; 0.217; and 0.598 for Ethane for the LS-SVM(1), LS-SVM(2) and Naive forecasts respectively; and standard deviation values of 0.707; 0.420; and 0.771 respectively. Figure 13 illustrates the error distributions corresponding to those experiments, and Figure 14 compares the experiments directly for the same feature.

The LS-SVM(1) experiment for this feature only narrowly beat the Naive forecast by 10.7%, according to their differences in mean MSE. This is clearly reflected in the distribution. The shape of the Testing error distribution closely follows that of the Naive error distribution; the only major difference being the density of values centred around the distribution peaks. Even the standard deviations of the LS-SVM(1) and Naive forecasts are similar, having an 8% difference. This may illustrate a tendency of the model to mimic the Naive forecast if the characteristics of the unseen transformer are foreign to the model. As a result, it can be inferred that the model effectively defaults to the Naive forecast when it is under-trained.

The LS-SVM(2) experiment beat the Naive forecast by 63.7%, according to their mean MSE difference, and 45.5% according to their standard deviation difference. The difference in these standard deviations can be directly observed in Figure 13, as the outliers of the Testing error are closer to the peak of the distribution.

The outliers of the Naive method are significantly more spread out. Since the metric being used is the MSE and not MAPE, the standard deviation of an error distribution does have an effect on the measure of how good a model is, since the MSE punishes outliers significantly. The consequence of this can be observed in Figure 13; the peaks of the LS-SVM(2) and Naive distributions do not have a significant bias relative to each other, but LS-SVM(2) is classified as being significantly better according to the MSE.

Figure 14 compares LS-SVM(1) and LS-SVM(2) directly. According the the MSE's of the methods, LS-SVM(2) is 59.4% better than LS-SVM(1). The resulting distributions are clear as to why LS-SVM(2) performs better; the peak of the LS-SVM(2) experiment is significantly closer to zero when compared to LS-SVM(1), and the density of extreme outliers is significantly lower for LS-SVM(2).

The fact that LS-SVM(2) beat LS-SVM(1) significantly is a clear indication that there is information in the data within transformers that cannot be perfectly translated across transformers. For the purpose of implementing this system, Experiment 2 should therefore be considered a more reliable method of splitting data for forecasting than Experiment 1, since it includes all transformers in Training and Validation, and hence allows the models access to more information.

4.4.4 Original vs. Augmented Input Vectors

Expert knowledge can be an important part of designing successful machine learning systems, since it may give a good indication as to which features contain important information about the system, and hence should be incorporated into the models in an effort to improve them. In the context of forecasting, an improvement is quantified as how well a particular model performs relative to how it would have performed if the expert knowledge was not available.

In the case of DGA, the expert knowledge comes from literature that suggests the ratios between specific dissolved gases is indicative of different potential failure mechanisms that occur [5, 6]. The input vector was augmented with these ratios and utilised in various machine learning models. The results were subsequently compared to identical models where the original input vector was used. The purpose of this was to observe the effect of including expert knowledge into forecasting methods. Student's t-test is used as a metric to adequately compare the methods by investigating the statistical significance of the difference between their respective error distributions [29]. The null hypothesis in this case is such that there is not statistically significant difference between the samples. The two tailed t-value between original and augmented input vector result is found as follows:

$$t_{value} = \frac{|\mu_{orig} - \mu_{aug}|}{\sqrt{\frac{\sigma_{orig}^2}{N_{orig}} + \frac{\sigma_{aug}^2}{N_{aug}}}}, \tag{26}$$

where μ is the mean of the distribution in question, σ is the standard deviation of the same distribution, and N is the number of data points in the distribution. If t_{value} is less than the critical value 1.96, then the null hypothesis is not rejected, and there is no statistically significant difference between the two distributions in question. If t_{value} is not less than the critical value, then the null hypothesis is rejected, which implies that there is a statistically significant difference between the distributions being compared.

Table 7 summarises the mean results of the original and augmented input vectors for the three methods that allowed for input vector augmentation; LS-SVM(1), LS-SVM(2), and ϵ-SVR(2).

Table 7: Input Vector Comparison Summary for All Experiments across All Transformers

Exp. Type	**Feat**	**Original**				**Augmented**				**t-value**
		Valid		**Test**		**Valid**		**Test**		
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	
LS-SVM (Exp. 1)	H_2	0.181	0.396	0.425	0.652	0.189	0.435	0.638	0.785	6.459
	C_2H_6	0.304	0.501	0.534	0.707	0.350	0.567	0.579	0.750	1.936
	CH_4	0.072	0.228	0.186	0.426	0.044	0.184	0.094	0.302	1.000
	C_2H_4	0.584	0.735	0.632	0.776	0.353	0.573	0.614	0.775	1.959
	C_2H_2	0.164	0.389	1.451	1.204	0.320	0.564	2.630	1.622	0.037
	CO_2	0.326	0.548	0.495	0.680	0.290	0.510	0.377	0.604	2.859
	CO	0.283	0.531	0.348	0.586	0.276	0.514	0.313	0.559	4.513
LS-SVM (Exp. 2)	H_2	1.112	1.050	0.278	0.521	1.056	1.026	0.336	0.574	0.059
	C_2H_6	0.383	0.575	0.217	0.420	0.420	0.553	0.290	0.520	4.108
	CH_4	0.198	0.376	0.053	0.207	0.156	0.333	0.058	0.232	6.017
	C_2H_4	0.365	0.588	0.672	0.773	0.203	0.424	0.748	0.837	2.228
	C_2H_2	0.885	0.941	0.510	0.702	0.719	0.847	0.533	0.706	2.458
	CO_2	0.387	0.617	0.331	0.535	0.377	0.597	0.349	0.584	7.143
	CO	1.046	1.008	0.224	0.472	0.849	0.901	0.263	0.513	2.455

The highlighted cells in the table indicate which Testing sets performed better per feature for the same model, for the two different input types. The t-values between Testing error distributions are also indicated in the table. The highlighted cells in the t-value column indicate for which features the input type makes no statistically relevant difference. In LS-SVM(1), when the mean MSE values are considered, it seems that the model performed better overall when given the augmented input vectors. However, upon closer inspection, the statistically significant difference between four of the features is negligible.

For the features CO_2 and CO, the augmented input vector outperformed the original input vector when the mean MSE results are considered, and for the feature H_2, the original input vector outperformed the augmented one. This is an interesting result, as the increase of forecasting accuracy seems feature-dependent for LS-SVM(1), when input types are considered.

When the Validation error results are considered, the LS-SVM(2) performed better using the augmented input vector. This is indicated by the lower values of mean MSE, and standard deviation of the error distributions. However, when the Testing errors are considered, the LS-SVM(2) model performed better when the original input vector is used. This is evident from the lower mean MSE values for all features, as well as the fact that for six of the seven features, the difference in distributions is statistically significant.

The differences in results show that although the model trained better according to the Validation errors, it failed to generalise, and yielded worse results in the Testing set. This is an unexpected result - it was initially expected that the addition of expert knowledge to the model would increase the forecasting accuracy, but the results indicate that the model performs better without the additional features. A possible explanation to this result could be that these additional features were originally designed for fault classification, and not for forecasting. It is apparent that the knowledge may not be extended to forecasting data.

4.4.5 Time Series

Figure 15 illustrates an example of a time series plot with the forecast points using several methods that were implemented and discussed. The data is from the feature Ethylene (C_2H_4) of Transformer No. 4. The figure shows the actual time series being forecast, as well as the Testing results from four of the best performing techniques - LS-SVM(1), LS-SVM(2), ϵ-SVR, and Naive forecast.

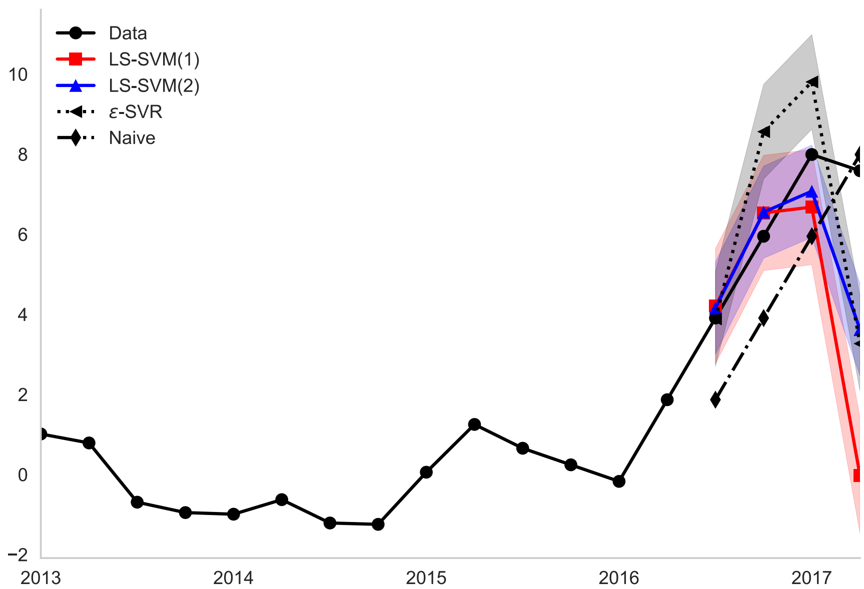


Figure 15: Time Series Forecast - Ethylene

The shaded regions associated with each forecast is that forecast's 95% forecast interval. This is obtained by using the estimated standard deviation of the data up until that point. The 95% forecast interval can be expressed as [26]:

$$\hat{f}(\boldsymbol{x}) \pm 1.96\sigma, \tag{27}$$

where σ is the estimated standard deviation. In this case, σ is taken as the standard deviation of the Validation errors. The final h points of the time series were held out for testing, $h \in [1:4]$, and these are the points that are forecast. It is once again emphasised that the forecasts conducted are one step forecasts.

In the particular case being plotted, it is clear that LS-SVM(1) and LS-SVM(2) perform very similarly for the first three points that are forecast. In this case, both LS-SVM models outperformed the Naive forecast. The errors associated with the ϵ-SVR model are generally larger when compared to the LS-SVM models. Once again, this emphasises the performance of the LS-SVM models over the ϵ-SVR model.

5 Conclusions

The following conclusions can be drawn from the discussion of the results:

- The hyper-parameter sets generally took a small number of iterations to optimise (less than 20, Figure 7). The optimisation portion of the experiment was therefore drawn out, and training times could have effectively been reduced if the total number of training iterations was reduced.
- Increasing the size of the Validation set by decreasing the size of the Testing set may improve forecasting results.
- Training models per-transformer instead of across multiple transformers may improve forecasting results at the expense of model generality. This is likely to reduce outliers in particular transformer forecasts.
- Performance of the models on unseen transformers was significantly worse than that on transformers that were previously encountered by the model, even though both sets of data were unseen. This is a clear indication that there is structure within the data of a particular transformer that assists in the forecasting of data of that transformer.
- When considering original and augmented input vectors, no generalisable conclusion could be reached with respect to the enhancement of forecast accuracy. The value of additional features is subjective according to both the model type and feature. As a result, new models should be tested with these different input types, or when applying models in reality, it should be noted which input types perform better for which feature.
- The LS-SVM(2) model generally outperformed the ϵ-SVR(2) model. This confirms the findings of other papers [7] that report on the superiority of LS-SVM models over ϵ-SVR models when forecasting DGA results.
- The LS-SVM(2) model also outperformed the LS-SVM-UV(2) model. This shows the advantage of including multiple features in the input vector to the model, and not treating time series as completely independent.

6 Recommendations

Recommendations for further research and development of the project are as follows:

- Incorporate mean absolute percentage error into the cost function of the optimiser. This would potentially increase the accuracy of the forecasts by further modification of the hyper-parameters.
- Work closer with power transformer companies in order to obtain good, clean, periodically sampled data. All research conducted thus far holds an underlying assumption that the data being used is correct and accurate. This is not necessarily the case; when working with industrial data, the quality of the data is vital in accurate forecasting.
- In the event that cleaner data is unavailable, the free parameter used in the smoothing stage of the data pre-processing, τ, should be optimised. This parameter was chosen on the basis that is related the time between successive data points to the number of target periods, despite the fact that it acts as a scaling parameter, and is not period dependent. It should be chosen in such a way that eliminates the noise in the data without compromising the trends within the time series.
- The purpose of the research was to compare novel methods across all transformers to methods that are transformer specific. As a result, the novel method of how data was split in both experiments was not extensively tested on a per-transformer basis. This method of experimentation may result in less general hyper-parameters, but may improve forecast accuracy.
- Multi-step forecasting models should be developed. It would be useful to have models that forecast more than one time step into the future, especially from the perspective of an industry where the components being analysed are vital to the functioning of the system.
- Use a more traditional cross-validation method specifically designed for time series [26]:
  - Assuming m is the minimum number of observations for a Training set, Select observation $m+i$ for the Validation set, using observations at times $t = 1, 2, ..., m+i-1$ to estimate the model.
  - Compute the error for time $m + i$. Repeat for $i = 0, 1, 2, ..., n - m - h$, where n is the total number of points in the time series and h is the number of points held out for Testing.
  - This method would definitely increase the size of the Validation set, prevent under-training, and possibly result in a higher forecast accuracy.

References

[1] H. De Faria, J. G. S. Costa and J. L. M. Olivas, 'A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis', *Renewable and Sustainable Energy Reviews*, vol. 46, pp. 201–209, 2015, ISSN: 18790690. DOI: 10.1016/j.rser.2015.02.052.

[2] H. C. Sun, Y. C. Huang and C. M. Huang, 'A Review of Dissolved Gas Analysis in Power Transformers', *Energy Procedia*, vol. 14, no. 2011, pp. 1220–1225, 2012, ISSN: 18766102. DOI: 10.1016/j.egypro.2011.12.1079.

[3] ASTM International, 'ASTM D3612-02, Standard Test Method for Analysis of Gases Dissolved in Electrical Insulating Oil by Gas Chromatography', ASTM International, West Conshohocken, PA, Tech. Rep., 2017. DOI: 10.1520/D3612-02R17.

[4] A. Eckner, 'Algorithms for Unevenly Spaced Time Series : Moving Averages and Other Rolling Operators', January 2010, 2015, pp. 1–11.

[5] M. Duval, 'The Duval Triangle for Load Tap Changers , Non-Mineral Oils and Low Temperature Faults in Transformers', *IEEE Electrical Insulation Magazine*, vol. 24, no. 6, pp. 22–29, 2008.

[6] M. Duval and L. Lamarre, 'The Duval Pentagon - A New Complementary Tool for the Interpretation of Dissolved Gas Analysis in Transformers', *IEEE Electrical Insulation Magazine*, vol. 30, no. 6, pp. 1–4, 2014.

[7] R. Liao, H. Zheng, S. Grzybowski and L. Yang, 'Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers', *Electric Power Systems Research*, vol. 81, no. 12, pp. 2074–2080, 2011, ISSN: 03787796. DOI: 10.1016/j.epsr.2011.07.020.

[8] S. W. Fei and Y. Sun, 'Forecasting dissolved gases content in power transformer oil based on support vector machine with genetic algorithm', *Electric Power Systems Research*, vol. 78, pp. 507–514, 2008. DOI: 10.1016/j.epsr.2007.04.006.

[9] S. W. Fei, M. J. Wang, Y. B. Miao, J. Tu and C. L. Liu, 'Particle swarm optimization-based support vector machine for forecasting dissolved gases content in power transformer oil', *Energy Conversion and Management*, vol. 50, no. 6, pp. 1604–1609, 2009, ISSN: 0196-8904. DOI: 10.1016/j.enconman.2009.02.004.

[10] Y. Y. Zhang, H. Wei, Y. D. Yang, H. B. Zheng, T. Zhou and J. Jiao, 'Forecasting of Dissolved Gases in Oil-immersed Transformers Based upon Wavelet LS-SVM Regression and PSO with Mutation', *Energy Procedia*, vol. 104, pp. 38–43, 2016, ISSN: 1876-6102. DOI: 10.1016/j.egypro.2016.12.008.

[11] L. Wenjun, Z. Yu, G. Weiqiang and L. Liegen, 'Forecasting Dissolved Gases Content in Power Transformer Oil Based on Particle Swarm Optimization-based RBF Neural Network', in *ISECS International Colloquium on Computing, Communication, Control, and Management Forecasting*, 2009, pp. 153–155, ISBN: 9781424442461.

[12] K. Shaban, 'A Cascade of Artificial Neural Networks to Predict Transformers Oil Parameters', *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 16, no. 2, pp. 516–523, 2009.

[13] F. J. Chang and Y. C. Chen, 'Estuary water-stage forecasting by using radial basis function neural network', *Journal of Hydrology*, vol. 270, pp. 158–166, 2003.

[14] M. H. Wang and C. P. Hung, 'Novel grey model for the prediction of trend of dissolved gases in oil-filled power apparatus', *Electric Power Systems Research*, vol. 67, no. 1, pp. 53–58, 2003, ISSN: 03787796. DOI: 10.1016/S0378-7796(03)00047-6.

[15] M. H. Wang, 'Grey-Extension Method for Incipient Fault Forecasting of Oil-Immersed Power Transformer', *Electric Power Components and Systems*, vol. 32, no. 10, pp. 959–975, 2004, ISSN: 1532-5008. DOI: 10.1080/15325000490257999.

[16] C. Kao and C.-l. Chyu, 'A fuzzy linear regression model with better explanatory power', *Fuzzy Sets and Systems*, vol. 126, pp. 401–409, 2002.

[17] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*, Second Edition. Stanford: Springer, 2008, ISBN: 978-0387848570.

[18] C. M. Bishop, *Pattern Recognition and Machine Learning*, First Edition, M. Jordan, J. Kleinberg and B. Scholkopf, Eds. Cambridge: Springer, 2006, ISBN: 9780387310732.

[19] J. Suykens and K. U. L. Esat-scd-sista, *Least Squares Support Vector Machines*, July. 2002, ISBN: 9812381511.

[20] J. Kennedy and R. Eberhart, *Swarm Intelligence*, First Edition. Morgan Kaufmann, 2001, ISBN: 1-55860-595-9.

[21] R. de Bruyn and J. Atherfold, *Interview: 1 February 2018*, Johannesburg, 2018.

[22] D. Sitaram, A. Dalwani, A. Narang, M. Das and P. Auradkar, 'A Measure of Similarity of Time Series Containing Missing Data Using the Mahalanobis Distance', *Proceedings - 2015 2nd IEEE International Conference on Advances in Computing and Communication Engineering, ICACCE 2015*, pp. 622–627, 2015. DOI: 10.1109/ICACCE.2015.14.

[23] L. Li, F. Noorian, D. J. M. Moss and P. H. W. Leong, 'Rolling Window Time Series Prediction using MapReduce', *Ieee Iri*, no. September, 2014. DOI: 10.1109/IRI.2014.7051965.

[24] J. Schroeder, C. Berger, A. Knauss, H. Preenja, M. Ali, M. Staron and T. Herpel, 'Predicting and Evaluating Software Model Growth in the Automotive Industry', *IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pp. 584–593, 2017. arXiv: arXiv:1708.02884v1.

[25] K. I. Annapoorani, 'FORECASTING THE NATURE OF POWER TRANSFORMER INSULATION OIL USING CHEMICAL PROPERTIES OF DISSOLVED GASES', *Int. J. Chem. Sci.*, vol. 13, no. 4, pp. 1777–1786, 2015.

[26] Rob J Hyndman, *Forecasting: Forecasting: Principles & Practice*, September. 2014, p. 138, ISBN: 9780987507105. DOI: 10.1017/CBO9781107415324.004.

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, 'Scikit-learn: Machine Learning in Python', *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[28] S. Seabold and J. Perktold, *Statsmodels: Econometric and statistical modeling with python.* 2010.

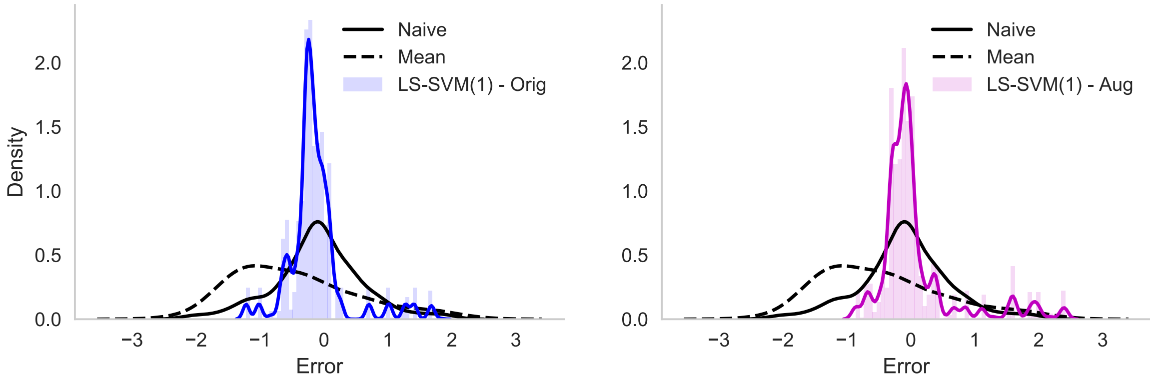
[29] J. K. Fotheringham, 'The probable error of a water-clock', *The Classical Review*, vol. 37, no. 7-8, pp. 166–167, 1923, ISSN: 14643561. DOI: 10.1017/S0009840X00085267. arXiv: 9603010 [alg-geom].

A Unabridged Results

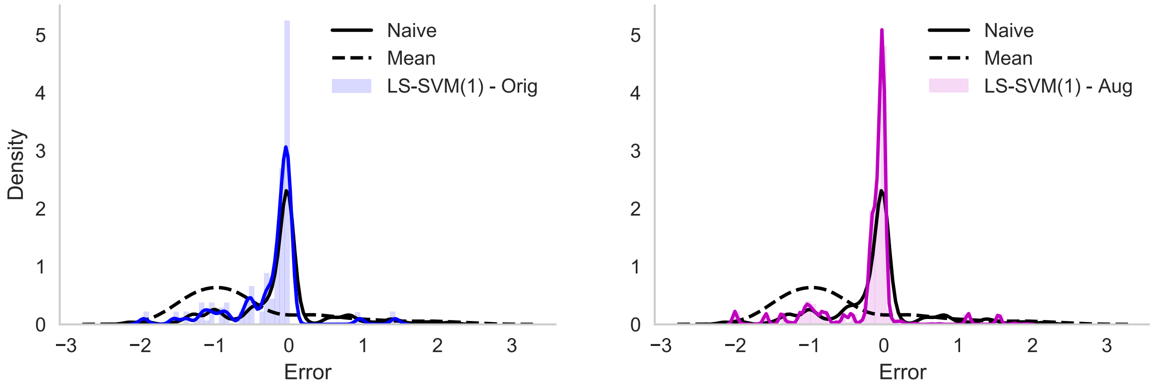
A.1 Distribution of Errors

A.1.1 Experiment 1

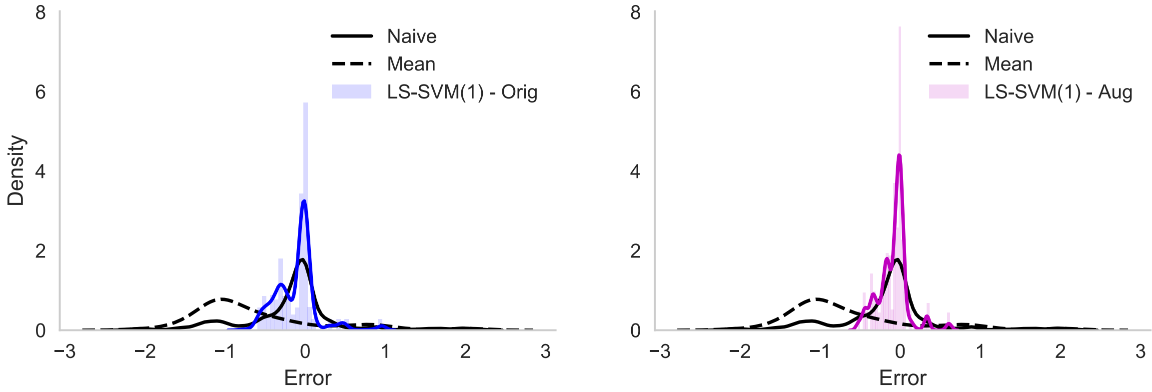
Figure 16 shows the unabridged error histograms plotted for Experiment 1. In each case, the testing error histograms and density estimators are plotted for each feature. These are directly compared with the errors obtained using the Mean and Naive forecasts. For each feature, the left hand figure was obtained using the original input vector, and the right hand figure was obtained using the augmented input vector.



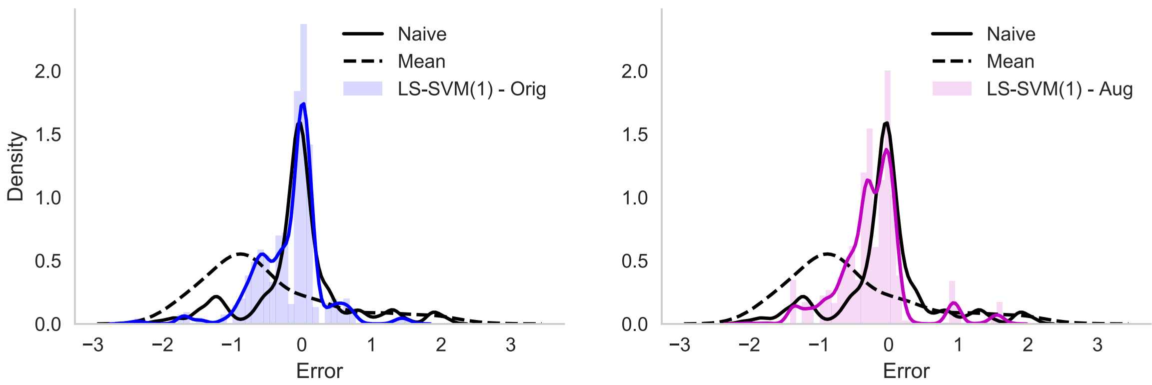
(a) Hydrogen



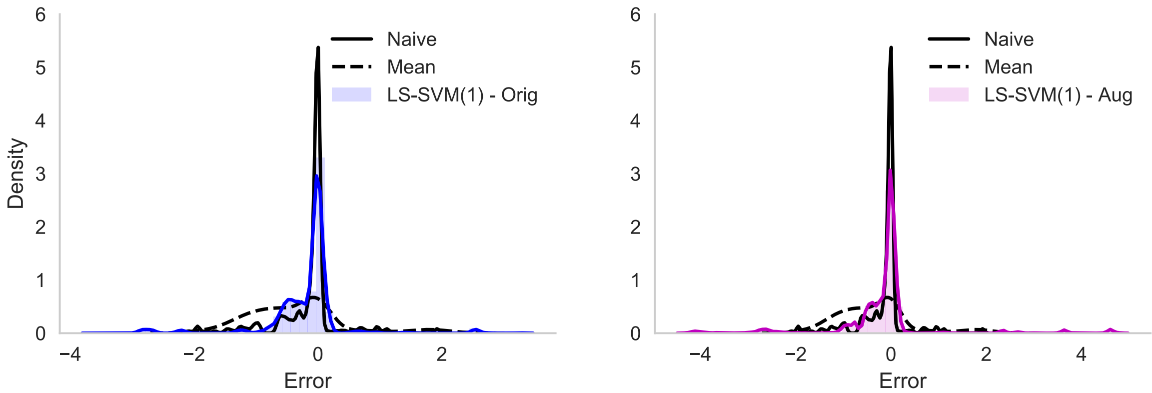
(b) Ethane



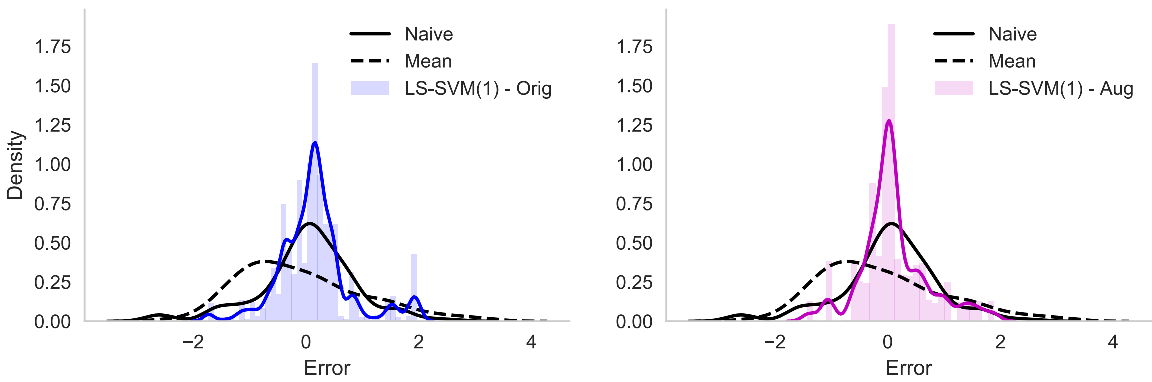
(c) Methane



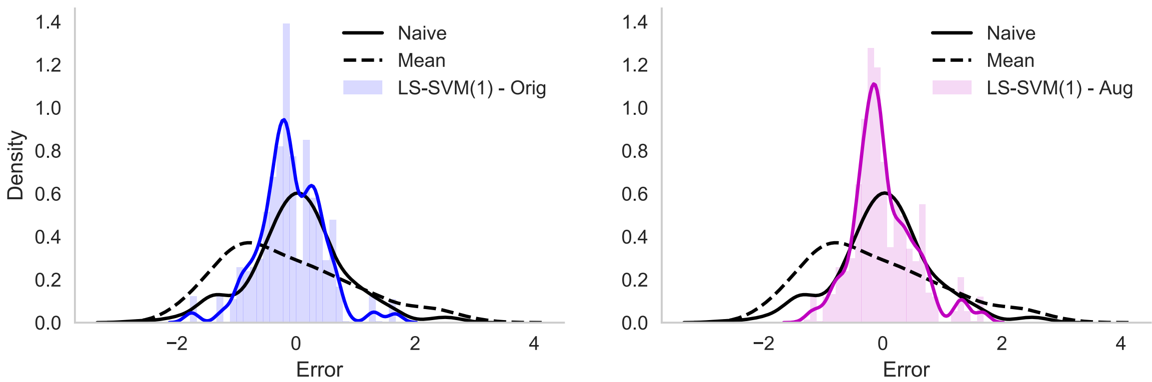
(d) Ethylene



(e) Acetylene



(f) Carbon Dioxide

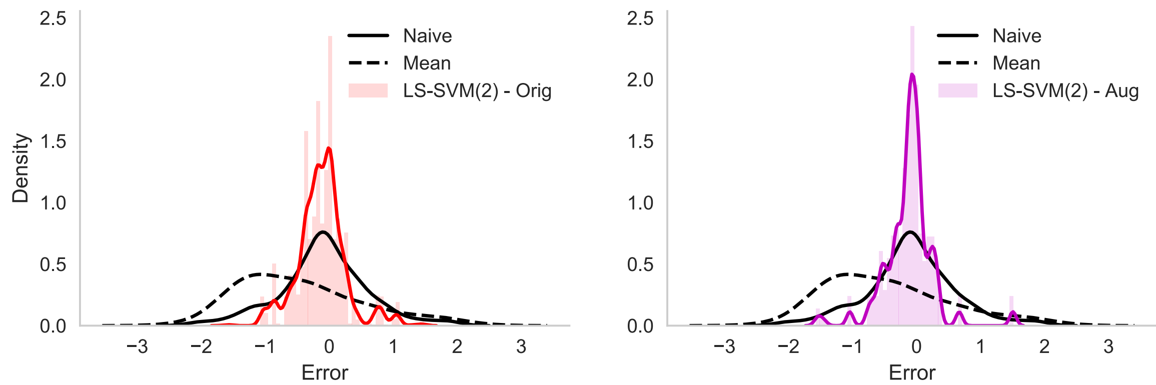


(g) Carbon Monoxide

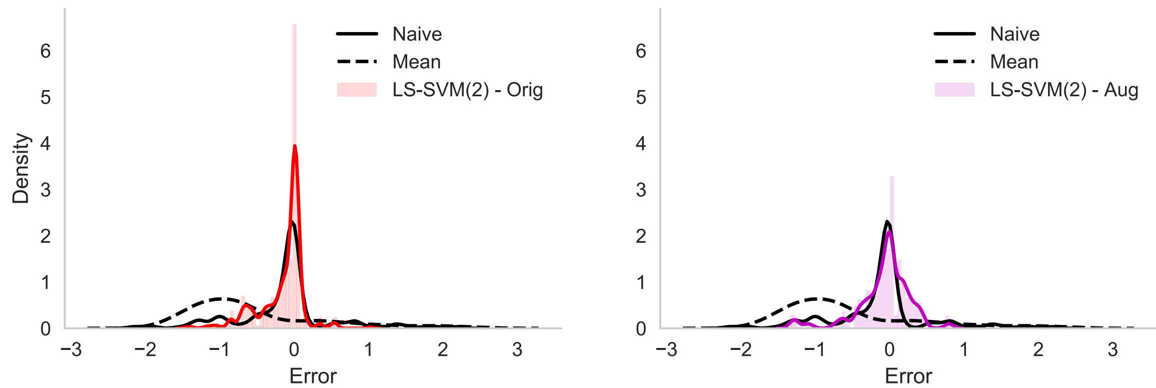
Figure 16: Experiment 1 Results - All Features

A.1.2 Experiment 2

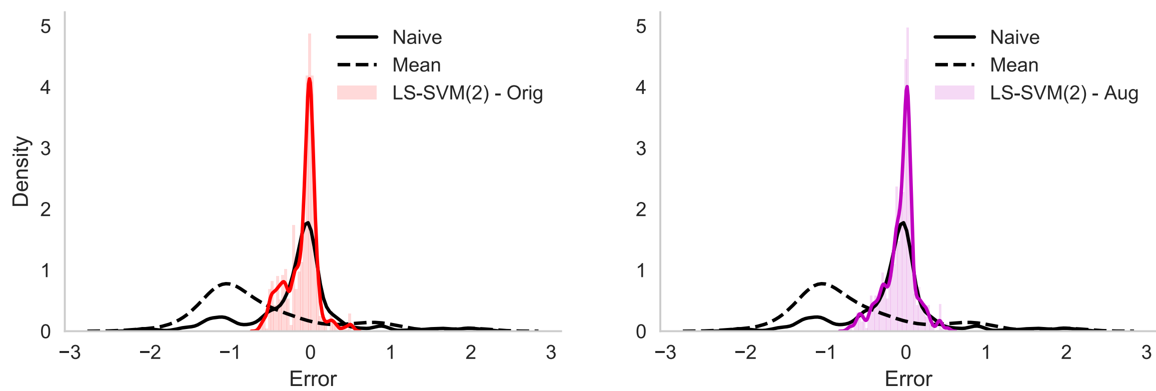
Figure 17 shows the unabridged error histograms using LS-SVR as a forecasting tool. In each case, the testing error histograms and density estimators are plotted for each feature. These are directly compared with the errors obtained using the Mean and Naive forecasts. For each feature, the left hand figure was obtained using the original input vector, and the right hand figure was obtained using the augmented input vector.



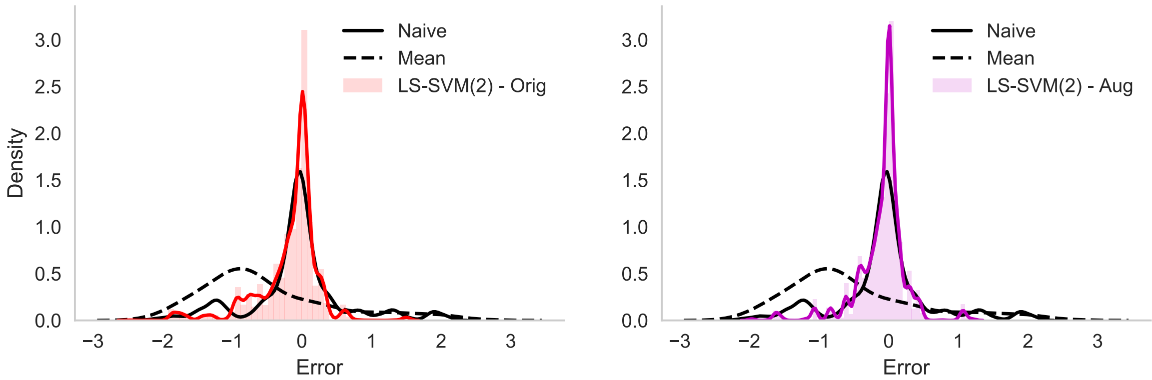
(a) Hydrogen



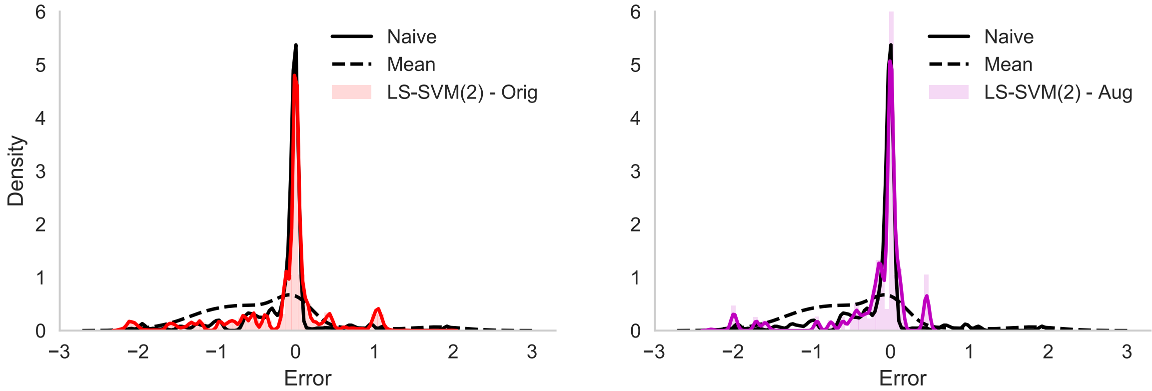
(b) Ethane



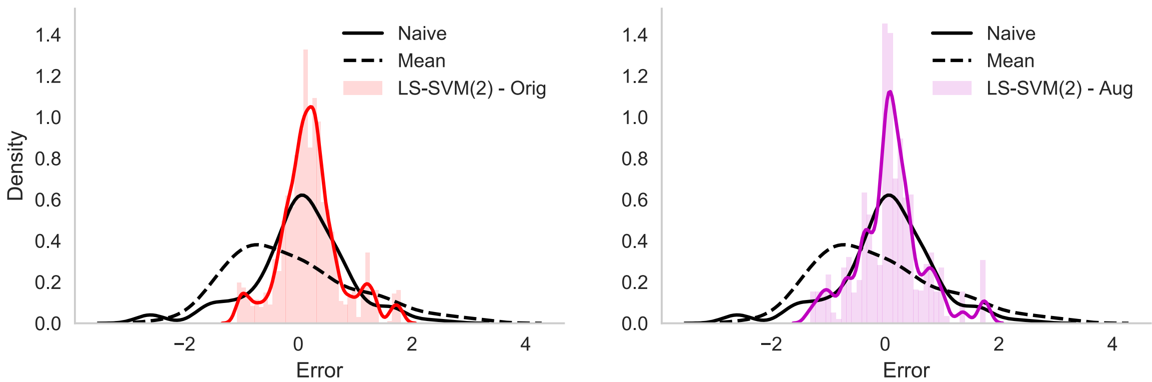
(c) Methane



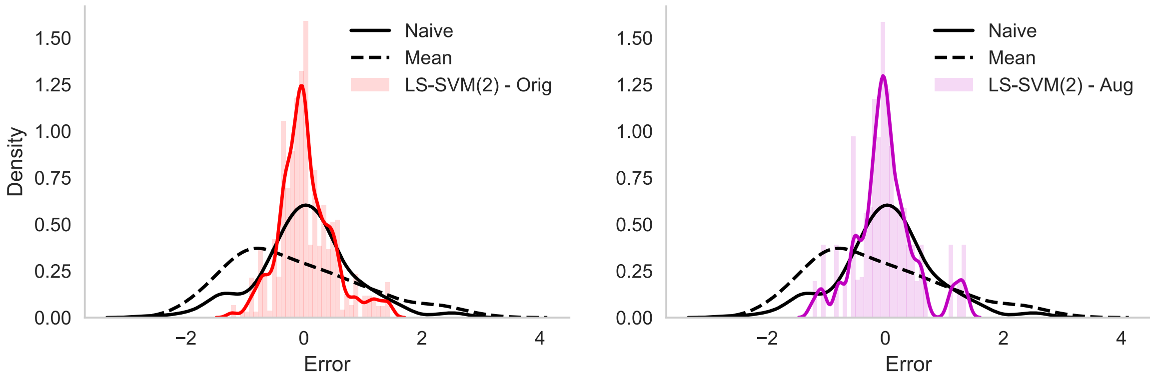
(d) Ethylene



(e) Acetylene



(f) Carbon Dioxide

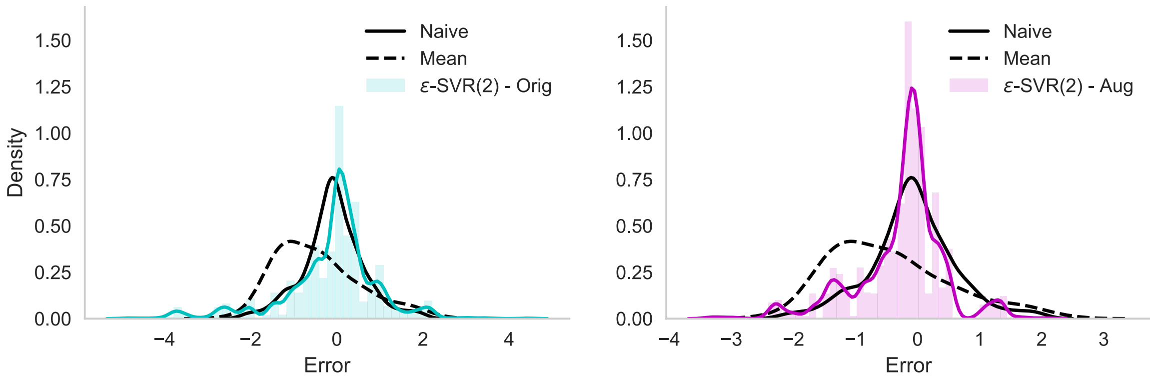


(g) Carbon Monoxide

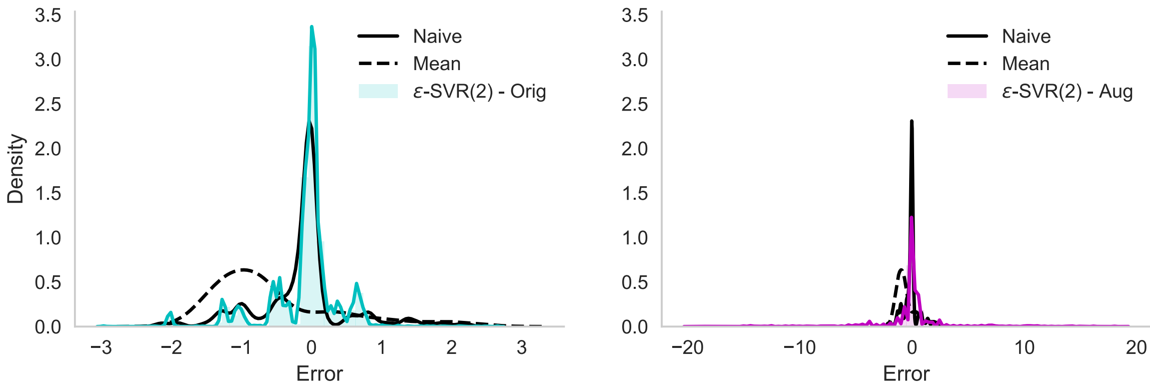
Figure 17: Experiment 2 Results - All Features

A.1.3 ϵ-SVR

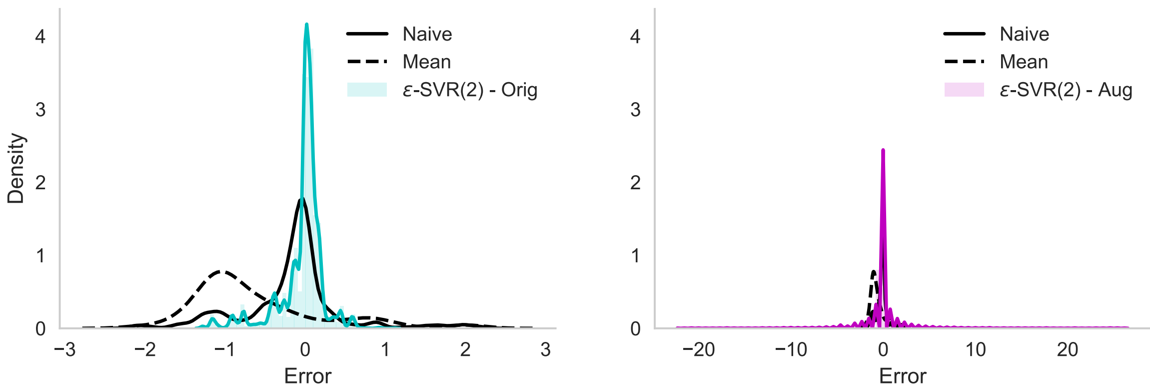
Figure 18 shows the unabridged error histograms using SK-learn's ϵ-SVR as a forecasting tool [27]. In each case, the testing error histograms and density estimators are plotted for each feature. These are directly compared with the errors obtained using the Mean and Naive forecasts. For each feature, the left hand figure was obtained using the original input vector, and the right hand figure was obtained using the augmented input vector.



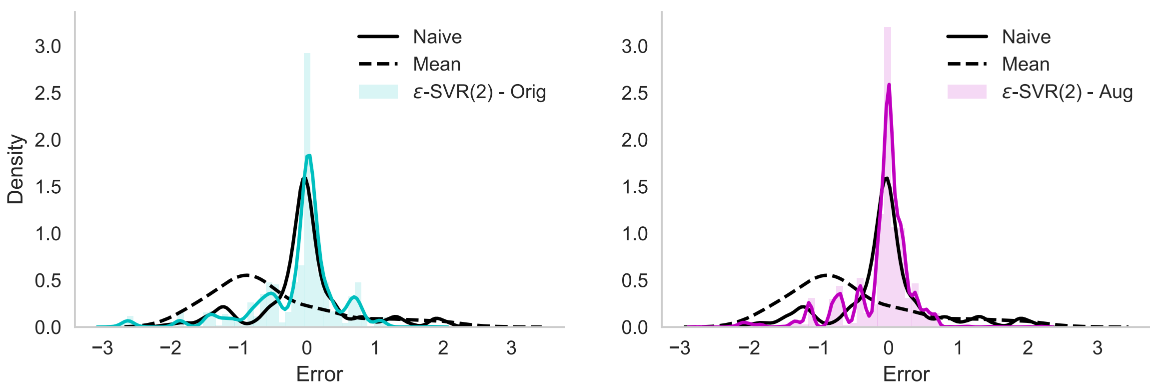
(a) Hydrogen



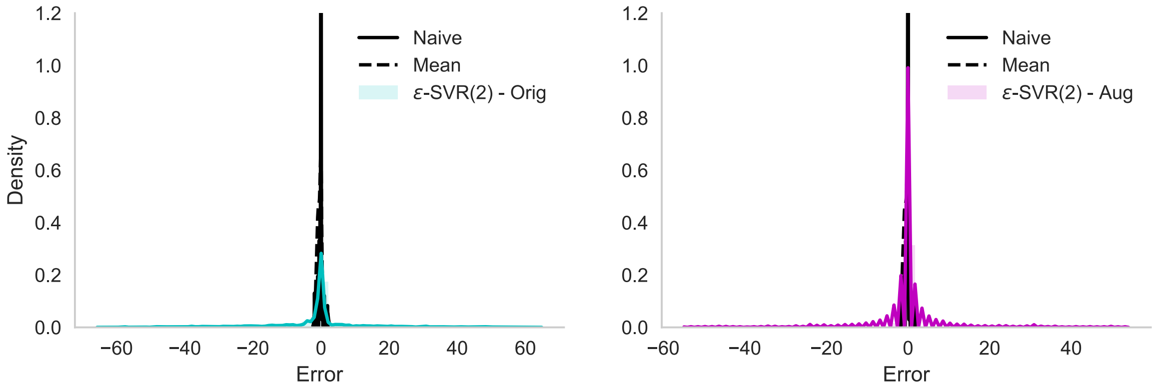
(b) Ethane



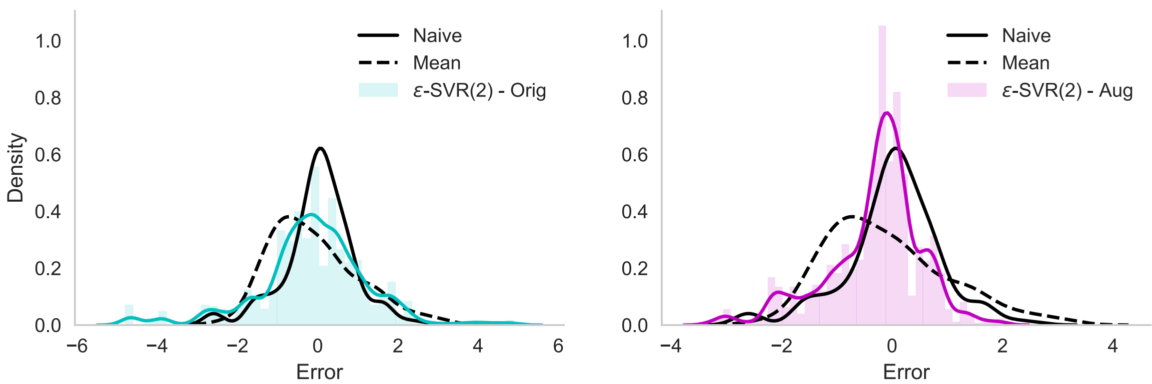
(c) Methane



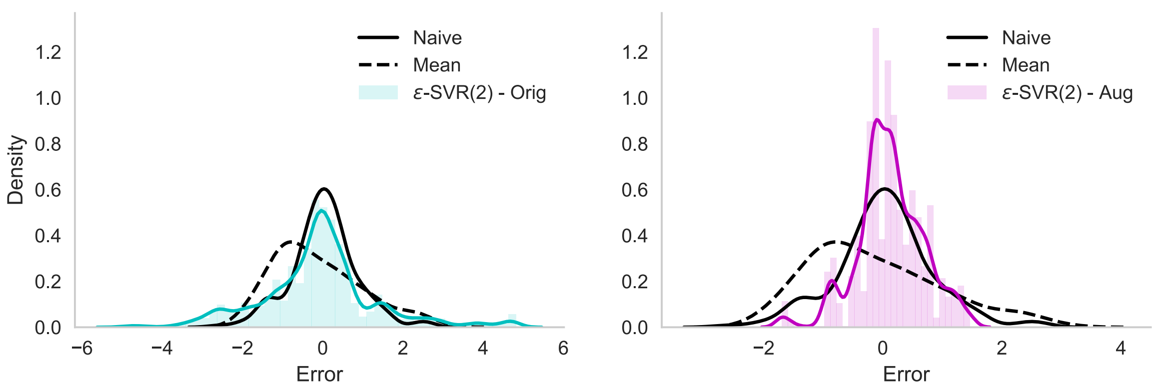
(d) Ethylene



(e) Acetylene



(f) Carbon Dioxide



(g) Carbon Monoxide

Figure 18: ϵ-SVR Results - All Features

A.2 Forecasting Results

A.2.1 Validation Errors

Tables 8 and 9 show all of the Mean Squared Error results obtained for all experiments conducted, using the original and augmented input vectors respectively. The highlighted cells in each table indicate the experiment that performed best in Validation and Testing sets for each feature. At the end of each table, it indicated which experiments performed best per feature across all transformers, according to the mean MSE.

Table 8: Validation and Testing Results Summary - Original Features

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
1	H_2	0.651	0.563	0.655	0.565	0.492	0.356	0.323	0.479	0.441	0.605	1.866	0.605	0.641	0.495	0.477	0.600
	C_2H_6	0.710	0.263	0.856	0.228	0.141	0.252	0.807	0.767	0.036	0.189	1.379	0.189	0.381	0.149	0.559	0.179
	CH_4	0.206	0.122	0.207	0.122	0.118	0.175	0.148	0.144	0.071	0.217	0.020	0.217	0.059	0.204	0.061	0.116
	C_2H_4	1.052	0.358	0.314	0.415	0.512	0.228	0.893	0.661	0.261	0.105	0.888	0.105	0.345	0.296	0.091	0.266
	C_2H_2	0.085	0.257	0.327	0.196	0.151	0.291	0.341	0.270	0.152	0.386	1309	0.386	2097	43.75	1310	36.190
	CO_2	1.646	0.392	3.700	0.242	0.521	0.701	1.399	0.498	0.436	0.638	1.376	0.638	0.435	0.645	0.921	0.516
	CO	1.074	0.624	1.748	0.707	0.266	0.470	0.776	0.583	0.259	0.509	3.351	0.509	2.519	1.240	27.27	5.187

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
2	H_2	4.9e^{-3}	0.064	0.011	0.097	0.131	0.296	0.029	0.067	0.157	0.344	24.09	0.344	0.387	0.593	0.109	0.301
	C_2H_6	9.555	3.027	9.647	3.025	0.832	0.328	10.83	3.107	1.783	0.615	13.32	0.615	3.132	1.769	17.02	3.837
	CH_4	0.097	0.300	0.096	0.300	0.613	0.424	0.082	0.287	0.734	0.569	0.209	0.569	6.201	2.468	6.859	2.615
	C_2H_4	45.56	6.190	47.62	6.143	0.132	0.304	47.02	6.107	0.268	0.459	42.64	0.459	0.181	0.425	45.36	6.162
	C_2H_2	232.7	13.25	233.7	13.25	4.2e^{-3}	0.024	232.9	13.27	0.079	0.201	1230	0.201	6.1e^{-4}	0.023	234.1	13.25
	CO_2	0.388	0.528	0.519	0.598	0.406	0.614	0.404	0.595	0.445	0.618	3.874	0.618	6.230	2.453	3.247	1.203
	CO	0.106	0.321	0.112	0.331	0.350	0.453	0.116	0.303	0.173	0.415	0.601	0.415	1.675	1.075	1.268	1.098
3	H_2	6.0e^{-3}	0.015	6.0e^{-3}	0.015	1.9e^{-5}	4.3e^{-3}	7.9e^{-4}	6.1e^{-3}	3.4e^{-4}	0.017	9.3e^{-3}	0.017	9.5e^{-4}	0.019	3.3e^{-4}	0.017
	C_2H_6	3.5e^{-4}	1.5e^{-3}	3.5e^{-4}	1.5e^{-3}	9.6e^{-4}	0.030	1.1e^{-3}	0.034	4.3e^{-4}	0.011	0.313	0.011	0.051	0.109	0.056	0.105
	CH_4	1.8e^{-4}	2.7e^{-3}	1.8e^{-4}	2.7e^{-3}	2.8e^{-3}	0.022	6.1e^{-4}	0.013	2.5e^{-3}	0.036	0.077	0.036	4.8e^{-3}	0.066	4.6e^{-3}	0.067
	C_2H_4	1.5e^{-3}	0.037	1.4e^{-3}	0.034	6.4e^{-4}	0.024	3.6e^{-4}	0.018	5.8e^{-4}	0.016	8.6e^{-5}	0.016	3.3e^{-3}	0.048	3.1e^{-3}	0.043
	C_2H_2	0.109	0.112	0.135	0.150	4.9e^{-3}	0.017	2.3e^{-5}	3.5e^{-3}	57.00	4.983	2252	4.983	1.4e^{-3}	0.037	6.6e^{-4}	0.025
	CO_2	0.722	0.514	0.874	0.723	0.362	0.362	0.540	0.687	2.329	0.361	6.127	0.361	3.351	1.798	13.37	3.294
	CO	0.577	0.656	0.483	0.669	0.628	0.246	0.410	0.640	1.570	0.547	1.435	0.547	0.849	0.382	0.596	0.768

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
4	H_2	0.450	0.668	12.74	1.379	0.681	0.389	1.620	1.235	0.190	0.205	17.70	0.205	0.312	0.252	11.71	2.347
	C_2H_6	0.063	0.228	18.81	0.646	0.416	0.631	6.310	1.129	0.419	0.646	11.87	0.646	0.671	0.768	10.64	2.849
	CH_4	0.541	0.314	5.061	2.246	0.548	0.606	5.032	1.717	0.493	0.682	10.18	0.682	1.508	0.833	10.68	2.389
	C_2H_4	6.122	2.405	8.087	2.737	0.861	0.912	8.912	1.949	0.893	0.943	14.61	0.943	25.31	4.996	17.98	3.487
	C_2H_2	1.034	0.824	56.35	7.030	1.451	1.203	2.161	1.466	1.470	1.182	171.0	1.182	1.628	1.274	4.549	2.129
	CO_2	1.602	1.060	4.194	1.544	0.397	0.593	0.608	0.740	0.204	0.445	6.952	0.445	0.335	0.554	1.047	0.940
	CO	0.184	0.314	1.379	0.449	2.796	1.154	0.403	0.630	1.591	0.995	17.80	0.995	3.099	1.659	163.5	11.76
5	H_2	0.042	0.023	0.078	0.130	0.074	0.262	0.029	0.029	0.052	0.226	0.227	0.226	0.263	0.416	0.341	0.314
	C_2H_6	3.8e^{-5}	6.0e^{-3}	3.8e^{-5}	6.0e^{-3}	0.022	0.112	3.2e^{-3}	0.011	2.5e^{-3}	0.039	3.8e^{-4}	0.039	0.034	0.100	3.3e^{-3}	0.051
	CH_4	9.6e^{-4}	0.014	9.8e^{-4}	0.014	0.029	0.058	9.9e^{-4}	0.030	0.019	0.096	1.1e^{-3}	0.096	0.096	0.124	4.2e^{-3}	0.064
	C_2H_4	0.242	0.196	0.296	0.051	0.027	0.075	1.6e^{-3}	0.029	0.011	0.054	9.9e^{-3}	0.054	0.152	0.234	0.056	0.236
	C_2H_2	9.8e^{-3}	0.057	0.015	0.056	0.522	0.715	1.076	0.058	4.290	1.597	1913	1.597	0.871	0.737	0.777	0.731
	CO_2	0.105	0.307	0.105	0.301	0.149	0.384	0.117	0.338	4.523	0.998	0.484	0.998	0.437	0.646	2513	49.90
	CO	0.185	0.176	0.213	0.174	0.344	0.496	0.179	0.198	1.670	0.723	0.441	0.723	0.429	0.626	0.266	0.371

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
6	H_2	0.085	0.035	0.085	0.035	0.053	0.116	0.035	0.026	0.052	0.227	0.792	0.227	0.038	0.191	0.719	0.846
	C_2H_6	0.030	0.021	0.030	0.021	0.025	0.082	0.019	0.035	0.011	0.102	0.013	0.102	7.7e^{-3}	0.059	9.0e^{-3}	0.016
	CH_4	2.2e^{-3}	0.015	2.2e^{-3}	0.015	0.059	0.183	1.7e^{-3}	0.041	0.054	0.230	0.032	0.230	0.186	0.427	7.3e^{-3}	0.085
	C_2H_4	0.041	0.153	0.041	0.152	0.054	0.134	0.015	0.115	0.039	0.161	5.4e^{-3}	0.161	0.046	0.103	0.013	0.109
	C_2H_2	0.015	0.048	0.012	0.049	8.0e^{-4}	0.028	7.8e^{-3}	0.027	5.104	1.565	799.8	1.565	0.011	0.088	2.3e^{-3}	0.046
	CO_2	0.025	0.152	0.080	0.157	0.214	0.373	0.042	0.202	0.166	0.400	0.113	0.400	0.757	0.791	1318	36.20
	CO	0.019	0.135	0.053	0.224	0.180	0.297	0.029	0.163	0.102	0.309	0.038	0.309	0.105	0.304	4.0e^{-3}	0.051
7	H_2	0.021	0.022	0.026	0.052	0.061	0.034	3.9e^{-3}	8.9e^{-3}	0.033	0.172	1.017	0.172	0.226	0.156	0.033	0.129
	C_2H_6	0.011	0.027	0.011	0.027	0.013	0.110	2.9e^{-3}	0.053	0.012	0.092	2.126	0.092	0.021	0.089	0.017	0.109
	CH_4	2.1e^{-3}	0.032	2.1e^{-3}	0.032	2.5e^{-3}	0.048	2.5e^{-3}	0.048	3.1e^{-3}	0.043	0.044	0.043	0.051	0.112	0.061	0.156
	C_2H_4	0.038	0.138	0.043	0.125	6.7e^{-3}	0.059	5.7e^{-3}	0.072	8.2e^{-3}	0.089	3.169	0.089	7.4e^{-3}	0.078	0.013	0.104
	C_2H_2	3.7e^{-3}	0.042	0.024	0.152	1.622	0.314	0.017	0.041	7.242	2.683	1788	2.683	0.896	0.721	0.214	0.048
	CO_2	0.032	0.032	0.018	0.075	7.0e^{-3}	0.075	0.017	0.054	0.124	0.239	0.530	0.239	0.039	0.066	0.059	0.068
	CO	0.076	0.106	0.028	0.086	0.019	0.119	0.010	0.099	0.108	0.113	0.354	0.113	0.067	0.199	0.059	0.198

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
8	H_2	1.097	0.885	1.865	0.873	0.670	0.392	1.011	0.880	1.720	0.538	0.857	0.538	0.253	0.408	1.385	0.964
	C_2H_6	4.1e^{-3}	5.8e^{-3}	4.1e^{-3}	5.8e^{-3}	0.012	0.074	9.7e^{-3}	0.096	7.3e^{-3}	0.071	3.3e^{-3}	0.071	0.352	0.072	0.150	0.123
	CH_4	3.8e^{-3}	0.011	3.8e^{-3}	0.011	0.013	0.029	2.1e^{-3}	0.018	6.3e^{-3}	0.072	3.2e^{-3}	0.072	0.379	0.254	0.018	0.073
	C_2H_4	0.016	0.120	0.012	0.085	0.478	0.174	0.056	0.119	0.474	0.177	0.144	0.177	1.900	0.220	0.093	0.068
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.126	0.340	0.120	0.322	0.026	0.159	0.133	0.273	0.093	0.058	0.239	0.058	0.150	0.121	0.134	0.284
	CO	0.046	0.194	0.048	0.191	0.102	0.315	0.051	0.185	0.165	0.406	0.058	0.406	0.081	0.279	0.046	0.215
9	H_2	0.073	0.030	0.073	0.031	0.116	0.280	0.139	0.100	0.342	0.315	0.103	0.315	0.537	0.575	0.330	0.527
	C_2H_6	4.6e^{-3}	0.022	4.6e^{-3}	0.022	5.1e^{-3}	0.035	9.8e^{-3}	0.051	0.079	0.274	3.350	0.274	0.024	0.081	17.71	4.176
	CH_4	0.101	0.049	0.103	0.048	0.041	0.194	0.097	0.135	0.159	0.227	0.063	0.227	161.3	12.34	0.049	0.217
	C_2H_4	3.2e^{-3}	0.029	2.4e^{-3}	0.023	5.8e^{-3}	0.075	0.023	0.068	6.8e^{-3}	0.063	6.8e^{-3}	0.063	0.130	0.353	0.089	0.284
	C_2H_2	8.4e^{-5}	9.2e^{-3}	9.9e^{-5}	9.7e^{-3}	0.010	0.045	4.4e^{-3}	0.015	104.3	10.16	1422	10.16	0.018	0.070	0.053	0.229
	CO_2	0.025	0.154	0.232	0.357	0.634	0.673	0.198	0.377	5.444	1.521	2.131	1.521	1.3e^{+4}	113.5	170.1	13.03
	CO	0.027	0.162	0.314	0.268	1.784	1.143	0.193	0.286	5.564	2.063	3.977	2.063	1.886	1.339	0.994	0.961

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
10	H_2	7.4e^{-4}	0.016	7.3e^{-4}	0.016	2.6e^{-3}	0.045	7.0e^{-4}	0.020	0.230	0.478	7.387	0.478	4.1e^{-3}	0.052	1.5e^{-3}	0.035
	C_2H_6	5.3e^{-6}	1.7e^{-3}	5.0e^{-6}	1.6e^{-3}	7.2e^{-4}	0.026	1.5e^{-4}	0.011	1.8e^{-4}	0.011	0.140	0.011	2.6e^{-3}	0.050	1.1e^{-3}	0.031
	CH_4	1.8e^{-4}	4.4e^{-3}	1.9e^{-4}	4.4e^{-3}	2.5e^{-4}	0.013	5.8e^{-5}	5.7e^{-3}	0.014	0.115	6.3e^{-3}	0.115	0.017	0.109	5.0e^{-3}	0.071
	C_2H_4	1.1e^{-3}	0.027	1.3e^{-3}	0.027	1.6e^{-4}	0.011	1.6e^{-4}	7.3e^{-3}	6.8e^{-5}	8.1e^{-3}	1.2e^{-3}	8.1e^{-3}	0.014	0.104	6.1e^{-3}	0.075
	C_2H_2	0.119	0.220	0.130	0.271	1.2e^{-3}	0.015	5.2e^{-4}	0.014	960.7	28.08	7291	28.08	5.0e^{-3}	0.058	2.4e^{-3}	0.045
	CO_2	0.029	0.071	0.032	0.161	0.102	0.247	0.050	0.209	3.107	0.956	18.78	0.956	460.9	21.39	7.526	2.741
	CO	0.496	0.118	0.231	0.146	3.393	1.278	0.041	0.200	10.78	2.800	17.88	2.800	210.8	14.16	52.28	7.230
11	H_2	0.072	0.108	0.072	0.108	39.59	4.427	0.108	0.217	72.74	7.279	177.2	7.279	35.78	4.676	0.125	0.310
	C_2H_6	9.8e^{-5}	5.0e^{-3}	1.0e^{-4}	5.0e^{-3}	3.8e^{-4}	0.019	3.7e^{-4}	6.8e^{-3}	8.4e^{-4}	0.023	1.677	0.023	6.5e^{-3}	0.075	0.017	0.131
	CH_4	4.4e^{-4}	0.021	4.4e^{-4}	0.021	0.014	0.107	0.030	0.173	0.221	0.358	0.165	0.358	0.038	0.188	0.036	0.122
	C_2H_4	0.014	0.116	0.014	0.119	0.061	0.075	7.7e^{-3}	0.088	0.296	0.439	0.590	0.439	0.017	0.130	0.048	0.167
	C_2H_2	0.065	0.204	1.711	1.292	22.37	2.819	0.657	0.805	265.6	12.77	967.9	12.77	20.70	3.124	0.346	0.575
	CO_2	0.032	0.131	0.107	0.151	0.261	0.455	0.191	0.207	3.696	1.688	2.518	1.688	1.766	1.322	0.809	0.869
	CO	0.527	0.427	0.544	0.412	0.391	0.488	0.655	0.442	5.948	1.592	3.239	1.592	3.616	1.208	0.532	0.510

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
12	H_2	0.089	0.137	0.184	0.221	0.175	0.382	0.022	0.148	0.630	0.410	0.444	0.410	0.654	0.299	0.401	0.109
	C_2H_6	$5.7e^{-4}$	$4.5e^{-3}$	$5.8e^{-4}$	$4.5e^{-3}$	0.093	0.253	$1.1e^{-3}$	0.032	$8.9e^{-3}$	0.067	0.464	0.067	0.079	0.192	0.021	0.134
	CH_4	$4.2e^{-5}$	$5.0e^{-3}$	$4.0e^{-5}$	$5.0e^{-3}$	0.042	0.100	$1.9e^{-4}$	0.014	0.065	0.186	0.134	0.186	0.887	0.655	0.068	0.186
	C_2H_4	0.080	0.175	0.082	0.184	0.058	0.236	0.031	0.162	0.072	0.254	0.279	0.254	1.961	0.981	2.841	1.679
	C_2H_2	0.213	0.156	0.203	0.154	25.89	4.013	0.148	0.339	9.621	3.075	383.2	3.075	38.33	4.634	0.430	0.490
	CO_2	0.063	0.081	0.134	0.317	0.052	0.206	0.041	0.140	0.221	0.255	0.311	0.255	2.345	1.194	0.064	0.207
	CO	0.063	0.123	0.041	0.107	0.330	0.376	0.011	0.079	2.582	1.256	0.079	1.256	0.270	0.187	0.011	0.067
13	H_2	1.487	0.623	1.726	0.696	5.924	2.051	2.638	0.652	15.33	3.887	10.44	3.887	7.264	2.211	2867	53.17
	C_2H_6	$8.2e^{-3}$	0.040	$8.3e^{-3}$	0.040	$7.9e^{-3}$	0.063	$3.2e^{-3}$	0.055	0.025	0.054	23.71	0.054	0.741	0.346	0.225	0.275
	CH_4	0.020	0.106	1.744	0.529	0.145	0.173	0.028	0.141	4.579	2.107	4.512	2.107	1.310	0.340	4.768	1.466
	C_2H_4	0.479	0.669	0.394	0.605	0.365	0.307	0.618	0.756	0.482	0.694	2.169	0.694	2.261	0.663	0.314	0.516
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.773	0.558	0.619	0.584	1.591	0.581	0.329	0.527	0.509	0.648	0.465	0.648	3.969	1.650	0.975	0.865
	CO	0.546	0.608	0.949	0.771	2.561	1.592	0.528	0.666	6.334	2.016	13.06	2.016	2.902	1.704	2.117	1.271

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
14	H_2	0.029	0.155	0.027	0.137	2.028	0.275	0.040	0.106	2.102	0.352	1.811	0.352	78.33	8.851	8.816	2.927
	C_2H_6	1.790	1.141	1.724	1.098	5.095	0.730	0.585	0.492	4.642	1.202	0.088	1.202	1449	37.91	27.57	4.356
	CH_4	0.043	0.167	0.044	0.167	0.014	0.095	0.022	0.128	$3.8e^{-3}$	0.050	0.012	0.050	1.079	0.850	1.181	0.917
	C_2H_4	0.495	0.599	0.572	0.635	1.042	0.466	0.630	0.448	0.895	0.581	0.233	0.581	1.829	0.636	0.415	0.644
	C_2H_2	1.035	0.108	3.983	0.896	1.212	0.131	3.678	0.523	0.612	0.509	194.2	0.509	72.43	8.464	3.380	0.949
	CO_2	0.253	0.503	0.158	0.382	1.012	0.334	0.116	0.302	1.040	0.676	1.194	0.676	1.837	0.735	0.223	0.230
	CO	0.412	0.632	0.309	0.554	1.117	0.194	0.253	0.491	2.057	0.869	1.782	0.869	1.492	0.535	0.329	0.553
15	H_2	0.188	0.148	0.192	0.153	0.118	0.051	0.407	0.291	0.170	0.248	1.450	0.248	0.059	0.241	2.748	1.182
	C_2H_6	0.105	0.157	0.106	0.157	0.189	0.098	0.016	0.110	0.159	0.210	$3.1e^{-3}$	0.210	0.099	0.109	0.059	0.120
	CH_4	0.064	0.162	0.064	0.162	0.204	0.148	0.071	0.171	0.174	0.228	$7.6e^{-3}$	0.228	0.138	0.310	0.086	0.274
	C_2H_4	0.029	0.158	0.029	0.159	0.081	0.123	0.021	0.132	0.094	0.197	0.014	0.197	0.034	0.087	0.040	0.188
	C_2H_2	0.218	0.148	0.221	0.150	0.190	0.201	0.948	0.391	0.434	0.377	202.2	0.377	3.688	1.920	2517	50.05
	CO_2	0.189	0.035	0.088	0.244	0.343	0.347	0.113	0.153	0.508	0.503	13.23	0.503	11.76	3.429	709.1	26.62
	CO	0.056	0.074	0.060	0.094	0.022	0.144	0.061	0.145	0.040	0.115	2.343	0.115	2.971	1.701	56.43	7.509

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
16	H_2	0.235	0.247	0.235	0.247	0.591	0.147	0.095	0.193	0.999	0.235	0.826	0.235	0.268	0.363	0.501	0.704
	C_2H_6	0.082	0.030	0.083	0.031	1.407	0.271	0.057	0.138	1.590	0.363	0.281	0.363	0.833	0.247	0.042	0.171
	CH_4	0.099	0.024	0.099	0.024	0.808	0.230	0.050	0.106	1.790	0.452	0.271	0.452	0.270	0.141	0.234	0.465
	C_2H_4	0.063	0.150	0.062	0.151	1.050	0.238	0.100	0.252	1.280	0.264	0.277	0.264	0.878	0.186	57.80	7.420
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.360	0.403	0.376	0.451	0.198	0.355	1.089	0.531	1.593	0.523	1.361	0.523	32.53	5.147	2.903	1.633
	CO	0.039	0.183	0.082	0.260	0.071	0.248	0.155	0.335	1.792	0.675	5.824	0.675	26.16	5.113	17.74	4.000
17	H_2	0.043	0.030	0.071	0.126	0.230	0.108	0.196	0.096	0.650	0.173	0.685	0.173	1.688	0.725	0.057	0.199
	C_2H_6	0.248	0.100	0.248	0.100	2.507	0.376	0.181	0.144	3.965	0.862	1.904	0.862	2.519	0.365	0.229	0.303
	CH_4	0.268	0.058	0.268	0.058	1.221	0.221	0.193	0.085	3.446	0.669	1.179	0.669	1.569	0.329	0.624	0.715
	C_2H_4	0.384	0.508	0.481	0.576	2.564	0.432	0.276	0.336	3.927	0.729	0.982	0.729	3.582	0.469	1.021	0.873
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.372	0.399	0.542	0.702	0.210	0.458	0.963	0.913	5.941	0.816	3.058	0.816	2.195	0.847	1.304	1.139
	CO	0.124	0.109	0.190	0.102	0.337	0.270	0.092	0.283	4.572	0.203	2.310	0.203	0.737	0.225	0.162	0.132

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
18	H_2	0.156	0.246	0.155	0.230	4.964	1.123	0.047	0.103	2.467	1.183	0.381	1.183	2.734	1.137	0.105	0.321
	C_2H_6	1.392	1.149	1.321	1.114	0.477	0.260	0.391	0.573	0.380	0.475	0.189	0.475	1.202	1.096	0.948	0.774
	CH_4	0.082	0.273	0.082	0.273	0.016	0.072	0.040	0.200	3.9e−3	0.055	0.021	0.055	33.201	3.695	73.11	5.760
	C_2H_4	0.619	0.731	0.594	0.721	2.768	1.235	0.126	0.287	2.860	1.376	0.085	1.376	11.15	1.497	2.240	1.233
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.396	0.510	0.094	0.306	2.594	0.620	0.133	0.304	1.899	1.326	0.780	1.326	80.346	6.771	261.7	16.174
	CO	0.536	0.717	0.310	0.503	5.134	0.765	0.350	0.589	0.784	0.884	0.287	0.884	0.034	0.167	0.162	0.256
All	H_2	0.181	0.396	0.425	0.652	1.112	1.050	0.278	0.521	2.266	1.505	2.829	1.675	6.412	2.532	160.8	12.67
	C_2H_6	0.304	0.501	0.534	0.707	0.383	0.575	0.217	0.420	0.370	0.562	0.978	0.968	64.50	8.030	4.182	2.026
	CH_4	0.072	0.228	0.186	0.426	0.198	0.376	0.053	0.207	0.427	0.590	0.176	0.417	11.74	3.420	5.437	2.263
	C_2H_4	0.584	0.735	0.632	0.776	0.365	0.588	0.672	0.773	0.398	0.603	1.183	1.053	2.733	1.627	7.135	2.668
	C_2H_2	0.164	0.389	1.451	1.204	0.885	0.941	0.510	0.702	10.93	3.304	456.1	21.35	99.54	9.935	226.1	15.04
	CO_2	0.326	0.548	0.495	0.680	0.387	0.617	0.331	0.535	1.798	1.315	2.790	1.667	796.3	28.22	278.0	16.67
	CO	0.283	0.531	0.348	0.586	1.046	1.008	0.224	0.472	2.417	1.551	2.695	1.640	15.18	3.889	17.99	4.239

Table 9: Validation and Testing Results Summary - Augmented Features

Trans No.	**Feat**	**LS-SVM (Exp. 1)**				**LS-SVM (Exp. 2)**				**ϵ-SVR (Exp. 2)**				**LS-SVM-UV (Exp. 2)**			
		Valid		**Test**		**Valid**		**Test**		**Valid**		**Test**		**Valid**		**Test**	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
1	H_2	0.330	0.513	0.330	0.513	0.378	0.353	0.227	0.443	0.425	0.595	6.564	0.595	0.641	0.495	0.477	0.600
	C_2H_6	0.697	0.257	0.721	0.209	0.534	0.254	0.194	0.439	0.618	0.227	187.6	0.227	0.381	0.149	0.559	0.179
	CH_4	0.160	0.108	0.161	0.108	0.097	0.170	0.118	0.137	0.053	0.192	3.9e^{-3}	0.192	0.059	0.204	0.061	0.116
	C_2H_4	0.762	0.334	0.916	0.232	0.530	0.176	0.089	0.265	0.726	0.075	0.168	0.075	0.345	0.296	0.091	0.266
	C_2H_2	0.015	0.075	0.673	0.177	0.118	0.263	0.204	0.215	0.233	0.401	4801	0.401	2097	43.75	1310	36.19
	CO_2	1.416	0.322	2.417	0.289	0.592	0.711	1.876	0.485	0.463	0.591	0.710	0.591	0.435	0.645	0.921	0.516
	CO	1.188	0.600	1.713	0.682	0.238	0.419	0.847	0.565	0.108	0.328	0.394	0.328	2.519	1.240	27.27	5.187
2	H_2	4.6e^{-3}	0.063	4.7e^{-3}	0.063	0.216	0.294	4.2e^{-3}	0.063	0.181	0.391	21.19	0.391	0.387	0.593	0.109	0.301
	C_2H_6	12.16	3.465	9.661	3.015	0.155	0.319	10.61	2.984	2.181	0.950	28.67	0.950	3.132	1.769	17.02	3.837
	CH_4	0.128	0.304	0.127	0.304	1.086	0.534	0.134	0.337	0.871	0.600	4905	0.600	6.201	2.468	6.859	2.615
	C_2H_4	46.78	6.147	47.76	6.131	0.109	0.328	46.03	5.886	0.150	0.376	43.56	0.376	0.181	0.425	45.36	6.162
	C_2H_2	231.7	13.26	234.6	13.25	3.0e^{-3}	0.021	232.6	13.27	0.088	0.215	1072	0.215	6.1e^{-4}	0.023	234.1	13.25
	CO_2	0.471	0.555	0.616	0.593	0.430	0.582	0.586	0.592	0.496	0.679	0.678	0.679	6.230	2.453	3.247	1.203
	CO	0.115	0.315	0.111	0.327	0.174	0.322	0.121	0.345	0.162	0.388	0.094	0.388	1.675	1.075	1.268	1.098

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
3	H_2	4.2e^{-5}	5.1e^{-3}	3.2e^{-4}	0.017	2.0e^{-3}	4.5e^{-3}	1.3e^{-3}	0.013	5.1e^{-3}	7.7e^{-3}	0.048	7.7e^{-3}	9.5e^{-4}	0.019	3.3e^{-4}	0.017
	C_2H_6	5.6e^{-4}	0.016	5.5e^{-4}	0.016	5.8e^{-4}	0.016	3.1e^{-5}	4.8e^{-3}	3.113	1.694	10.11	1.694	0.051	0.109	0.056	0.105
	CH_4	7.6e^{-5}	8.4e^{-3}	7.7e^{-5}	8.4e^{-3}	3.4e^{-4}	0.018	7.9e^{-4}	0.027	1.4e^{-3}	0.032	5.1e^{-4}	0.032	4.8e^{-3}	0.066	4.6e^{-3}	0.067
	C_2H_4	1.4e^{-3}	0.037	1.3e^{-3}	0.032	8.8e^{-4}	0.023	3.9e^{-4}	0.012	5.6e^{-3}	0.013	0.011	0.013	3.3e^{-3}	0.048	3.1e^{-3}	0.043
	C_2H_2	5.7e^{-3}	0.053	7.0e^{-3}	0.056	2.6e^{-3}	0.011	5.6e^{-5}	4.3e^{-3}	95.44	9.416	1260	9.416	1.4e^{-3}	0.037	6.6e^{-4}	0.025
	CO_2	0.331	0.478	0.307	0.446	0.126	0.348	0.368	0.451	2.385	0.598	0.713	0.598	3.351	1.798	13.37	3.294
	CO	0.396	0.623	0.277	0.511	0.307	0.164	0.290	0.485	0.814	0.835	0.589	0.835	0.849	0.382	0.596	0.768
4	H_2	0.651	0.804	21.842	1.681	0.769	0.460	1.669	0.776	0.334	0.297	6.053	0.297	0.312	0.252	11.71	2.347
	C_2H_6	0.444	0.615	21.29	0.801	0.522	0.690	4.642	1.284	0.549	0.727	10.03	0.727	0.671	0.768	10.64	2.849
	CH_4	0.513	0.396	3.293	1.429	0.565	0.705	4.218	1.575	0.500	0.707	8.958	0.707	1.508	0.833	10.68	2.389
	C_2H_4	1.402	0.765	6.553	1.370	1.020	0.927	11.68	1.998	0.994	0.954	16.11	0.954	25.31	4.996	17.98	3.487
	C_2H_2	4.237	1.104	40.33	3.801	1.356	1.148	4.074	1.984	1.754	1.238	2.304	1.238	1.628	1.274	4.549	2.129
	CO_2	0.577	0.741	0.567	0.752	0.669	0.611	0.528	0.538	1.106	0.600	0.760	0.600	0.335	0.554	1.047	0.940
	CO	0.116	0.176	0.120	0.146	2.941	1.153	0.459	0.279	2.462	1.219	0.922	1.219	3.099	1.659	163.5	11.76

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
5	H_2	0.016	0.026	0.036	0.117	0.080	0.277	0.011	0.038	0.088	0.297	0.072	0.297	0.263	0.416	0.341	0.314
	C_2H_6	0.170	0.403	0.158	0.392	0.947	0.209	0.120	0.043	1.009	0.468	0.043	0.468	0.034	0.100	3.3e^{-3}	0.051
	CH_4	1.5e^{-3}	0.013	1.5e^{-3}	0.013	0.227	0.372	0.029	0.166	7.5e^{-3}	0.070	8.2e^{-4}	0.070	0.096	0.124	4.2e^{-3}	0.064
	C_2H_4	0.113	0.046	0.119	0.048	0.082	0.172	0.016	0.037	0.030	0.068	2.8e^{-3}	0.068	0.152	0.234	0.056	0.236
	C_2H_2	8.5e^{-3}	0.066	0.010	0.093	0.114	0.332	0.208	0.015	0.151	0.383	0.046	0.383	0.871	0.737	0.777	0.731
	CO_2	0.087	0.295	0.074	0.272	0.276	0.472	0.081	0.283	0.424	0.353	0.337	0.353	0.437	0.646	2513	49.90
	CO	0.118	0.171	0.116	0.156	0.307	0.533	0.080	0.156	0.303	0.488	0.083	0.488	0.429	0.626	0.266	0.371
6	H_2	0.070	0.030	0.070	0.030	0.048	0.123	0.108	0.030	0.053	0.229	0.030	0.229	0.038	0.191	0.719	0.846
	C_2H_6	0.048	0.146	0.035	0.096	0.028	0.047	0.050	0.052	0.014	0.033	699.1	0.033	7.7e^{-3}	0.059	9.0e^{-3}	0.016
	CH_4	0.030	0.016	0.030	0.016	0.079	0.223	0.095	0.198	0.019	0.135	1465	0.135	0.186	0.427	7.3e^{-3}	0.085
	C_2H_4	0.348	0.163	0.335	0.176	0.032	0.124	0.014	0.113	0.027	0.143	0.058	0.143	0.046	0.103	0.013	0.109
	C_2H_2	0.173	0.028	0.439	0.600	1.2e^{-3}	0.034	0.026	0.034	3.116	1.016	2940	1.016	0.011	0.088	2.3e^{-3}	0.046
	CO_2	0.022	0.137	0.018	0.122	0.258	0.349	0.080	0.086	0.146	0.374	0.242	0.374	0.757	0.791	1318	36.20
	CO	0.071	0.142	0.079	0.118	0.143	0.242	0.014	0.082	0.083	0.288	0.268	0.288	0.105	0.304	4.0e^{-3}	0.051

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
7	H_2	8.2e^{-3}	0.030	0.015	0.069	0.118	0.069	5.8e^{-3}	0.018	0.044	0.122	2.910	0.122	0.226	0.156	0.033	0.129
	C_2H_6	0.169	0.405	0.063	0.230	0.154	0.021	0.024	0.018	0.188	0.125	36.24	0.125	0.021	0.089	0.017	0.109
	CH_4	2.7e^{-3}	0.023	2.6e^{-3}	0.023	0.033	0.163	2.8e^{-3}	0.034	1.7e^{-3}	0.038	171.8	0.038	0.051	0.112	0.061	0.156
	C_2H_4	0.076	0.087	0.069	0.135	0.062	0.133	0.026	0.116	7.9e^{-3}	0.051	0.314	0.051	7.4e^{-3}	0.078	0.013	0.104
	C_2H_2	0.159	0.110	0.343	0.361	1.484	0.082	5.7e^{-3}	0.047	86.98	8.818	1978	8.818	0.896	0.721	0.214	0.048
	CO_2	0.014	0.052	2.3e^{-3}	0.047	9.1e^{-3}	0.068	7.2e^{-3}	0.026	0.102	0.116	0.084	0.116	0.039	0.066	0.059	0.068
	CO	0.149	0.075	0.082	0.036	0.019	0.112	0.015	0.074	0.040	0.136	2.523	0.136	0.067	0.199	0.059	0.198
8	H_2	1.425	0.939	1.451	0.943	0.141	0.359	0.965	0.920	0.332	0.401	0.537	0.401	0.253	0.408	1.385	0.964
	C_2H_6	0.050	0.198	0.117	0.325	0.032	0.100	0.020	0.038	0.036	0.120	0.012	0.120	0.352	0.072	0.150	0.123
	CH_4	6.7e^{-3}	0.013	6.7e^{-3}	0.013	0.020	0.057	0.015	0.076	9.8e^{-3}	0.041	3.7e^{-3}	0.041	0.379	0.254	0.018	0.073
	C_2H_4	0.261	0.075	0.241	0.128	0.332	0.232	3.3e^{-3}	0.050	0.192	0.238	0.025	0.238	1.900	0.220	0.093	0.068
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.140	0.308	0.122	0.298	0.036	0.128	0.114	0.291	0.011	0.103	0.064	0.103	0.150	0.121	0.134	0.284
	CO	0.066	0.168	0.039	0.171	0.062	0.245	0.029	0.169	0.092	0.296	0.029	0.296	0.081	0.279	0.046	0.215

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
9	H_2	0.099	0.015	0.098	0.020	0.125	0.271	0.262	0.067	0.292	0.217	0.511	0.217	0.537	0.575	0.330	0.527
	C_2H_6	6.0e^{-3}	0.069	5.6e^{-3}	0.059	5.6e^{-3}	0.043	4.1e^{-4}	0.020	293.0	16.76	71.36	16.76	0.024	0.081	17.71	4.176
	CH_4	0.084	0.071	0.085	0.070	0.055	0.235	0.121	0.106	0.073	0.194	0.165	0.194	161.3	12.34	0.049	0.217
	C_2H_4	3.0e^{-3}	0.055	3.0e^{-3}	0.055	7.0e^{-3}	0.081	4.4e^{-3}	0.026	0.020	0.137	2.6e^{-3}	0.137	0.130	0.353	0.089	0.284
	C_2H_2	0.011	9.3e^{-3}	0.011	0.020	0.014	8.0e^{-3}	0.011	0.027	3.047	1.732	93.73	1.732	0.018	0.070	0.053	0.229
	CO_2	0.176	0.233	1.807	0.494	0.916	0.675	0.791	0.358	4.330	1.497	8.574	1.497	1.3e^{+4}	113.5	170.1	13.03
	CO	0.036	0.163	0.738	0.331	1.092	1.006	0.853	0.334	2.999	1.651	1.969	1.651	1.886	1.339	0.994	0.961
10	H_2	4.0e^{-4}	0.020	4.0e^{-4}	0.020	2.0e^{-3}	0.042	5.5e^{-4}	0.021	3.4e^{-3}	0.042	8.1e^{-3}	0.042	4.1e^{-3}	0.052	1.5e^{-3}	0.035
	C_2H_6	0.017	0.132	6.0e^{-3}	0.077	6.2e^{-4}	0.017	1.9e^{-4}	0.010	1.7e^{-3}	9.2e^{-3}	25.25	9.2e^{-3}	2.6e^{-3}	0.050	1.1e^{-3}	0.031
	CH_4	4.0e^{-5}	3.5e^{-3}	4.1e^{-5}	3.5e^{-3}	7.0e^{-4}	0.020	2.0e^{-4}	8.9e^{-3}	3.5e^{-4}	0.015	186.7	0.015	0.017	0.109	5.0e^{-3}	0.071
	C_2H_4	1.4e^{-5}	2.7e^{-3}	2.0e^{-5}	4.3e^{-3}	1.6e^{-4}	6.8e^{-3}	6.9e^{-5}	7.8e^{-3}	1.2e^{-4}	2.5e^{-3}	0.024	2.5e^{-3}	0.014	0.104	6.1e^{-3}	0.075
	C_2H_2	3.2e^{-5}	2.6e^{-3}	2.3e^{-3}	0.046	3.4e^{-4}	0.016	2.4e^{-4}	0.011	8.4e^{-3}	8.3e^{-3}	3479	8.3e^{-3}	5.0e^{-3}	0.058	2.4e^{-3}	0.045
	CO_2	7.8e^{-3}	0.085	8.9e^{-3}	0.082	0.064	0.149	4.7e^{-3}	0.068	1.545	0.329	0.521	0.329	460.9	21.39	7.526	2.741
	CO	0.462	0.112	0.752	0.364	5.311	1.424	0.043	0.206	4.317	2.064	0.461	2.064	210.8	14.16	52.28	7.230

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
11	H_2	0.069	0.061	0.066	0.068	39.93	4.462	0.113	0.245	54.07	5.042	0.780	5.042	35.78	4.676	0.125	0.310
	C_2H_6	1.9e^{-3}	0.043	2.3e^{-3}	0.047	1.0e^{-3}	0.015	2.4e^{-4}	7.6e^{-3}	2.9e^{-3}	0.050	0.524	0.050	6.5e^{-3}	0.075	0.017	0.131
	CH_4	3.9e^{-4}	0.018	3.8e^{-4}	0.018	0.025	0.067	1.2e^{-3}	0.033	0.067	0.102	8.6e^{-3}	0.102	0.038	0.188	0.036	0.122
	C_2H_4	5.0e^{-3}	8.3e^{-3}	4.8e^{-3}	0.014	0.055	0.095	3.9e^{-3}	0.043	0.079	0.115	0.141	0.115	0.017	0.130	0.048	0.167
	C_2H_2	0.067	0.064	0.633	0.685	22.92	3.095	0.829	0.521	29.39	3.202	59.41	3.202	20.70	3.124	0.346	0.575
	CO_2	0.011	0.093	0.022	0.058	0.196	0.306	0.059	0.044	0.882	0.923	0.078	0.923	1.766	1.322	0.809	0.869
	CO	0.589	0.417	0.577	0.419	0.905	0.403	0.634	0.396	0.425	0.592	0.732	0.592	3.616	1.208	0.532	0.510
12	H_2	0.066	0.195	0.104	0.320	0.131	0.345	0.032	0.174	0.203	0.275	0.053	0.275	0.654	0.299	0.401	0.109
	C_2H_6	2.9e^{-3}	0.037	1.0e^{-3}	6.3e^{-3}	0.601	0.542	5.5e^{-3}	0.025	1.822	1.310	43.25	1.310	0.079	0.192	0.021	0.134
	CH_4	6.9e^{-4}	3.3e^{-3}	6.7e^{-4}	3.3e^{-3}	0.181	0.315	1.6e^{-3}	0.039	0.024	0.075	344.780	0.075	0.887	0.655	0.068	0.186
	C_2H_4	0.082	0.069	0.078	0.072	0.195	0.286	0.010	0.097	0.041	0.194	0.053	0.194	1.961	0.981	2.841	1.679
	C_2H_2	0.275	0.231	0.410	0.517	31.281	4.299	0.077	0.202	55.67	7.441	1725	7.441	38.33	4.634	0.430	0.490
	CO_2	0.015	0.094	8.1e^{-3}	0.090	0.046	0.172	0.042	0.100	0.068	0.181	0.040	0.181	2.345	1.194	0.064	0.207
	CO	8.3e^{-3}	0.042	0.028	7.0e^{-3}	0.120	0.213	1.2e^{-3}	9.6e^{-3}	7.6e^{-3}	0.080	6.9e^{-3}	0.080	0.270	0.187	0.011	0.067

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
13	H_2	1.645	0.219	3.686	0.332	7.601	2.108	6.440	0.457	16.04	3.227	27.33	3.227	7.264	2.211	2867	53.17
	C_2H_6	0.014	0.096	0.013	0.074	1.127	0.068	0.055	0.023	0.560	0.118	3053	0.118	0.741	0.346	0.225	0.275
	CH_4	0.013	0.105	1.446	0.472	0.631	0.425	0.558	0.650	0.183	0.119	5097	0.119	1.310	0.340	4.768	1.466
	C_2H_4	0.728	0.819	0.821	0.836	0.123	0.322	0.424	0.651	0.483	0.559	0.442	0.559	2.261	0.663	0.314	0.516
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.627	0.434	0.698	0.419	1.637	0.563	0.802	0.337	0.448	0.660	0.170	0.660	3.969	1.650	0.975	0.865
	CO	0.592	0.548	1.079	0.612	2.468	1.532	1.667	0.530	4.989	2.222	4.454	2.222	2.902	1.704	2.117	1.271
14	H_2	0.231	0.337	0.231	0.337	1.959	0.195	0.063	0.112	2.216	0.372	0.786	0.372	78.334	8.851	8.816	2.927
	C_2H_6	3.371	1.825	2.341	1.440	5.806	0.777	0.930	0.595	6.235	1.502	0.551	1.502	1449	37.91	27.57	4.356
	CH_4	0.042	0.170	0.042	0.170	0.015	0.124	0.025	0.151	$7.9e^{-3}$	0.060	0.013	0.060	1.079	0.850	1.181	0.917
	C_2H_4	1.118	1.054	0.970	0.923	1.329	0.590	0.593	0.347	1.994	0.857	0.800	0.857	1.829	0.636	0.415	0.644
	C_2H_2	0.680	0.094	6.617	1.303	1.011	0.083	3.993	0.625	0.096	0.249	507.6	0.249	72.43	8.464	3.380	0.949
	CO_2	0.205	0.451	0.141	0.374	1.022	0.120	0.169	0.316	1.250	0.720	1.016	0.720	1.837	0.735	0.223	0.230
	CO	0.396	0.627	0.301	0.536	1.625	0.168	0.248	0.470	1.697	0.750	0.132	0.750	1.492	0.535	0.329	0.553

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
15	H_2	0.095	0.108	0.102	0.116	0.161	0.018	0.174	0.244	0.101	0.174	0.758	0.174	0.059	0.241	2.748	1.182
	C_2H_6	0.362	0.520	0.340	0.502	0.089	0.049	0.097	0.152	0.038	0.123	70.68	0.123	0.099	0.109	0.059	0.120
	CH_4	$5.5e^{-3}$	0.051	$5.4e^{-3}$	0.050	0.034	0.076	$8.1e^{-3}$	0.084	0.022	0.110	$5.7e^{-4}$	0.110	0.138	0.310	0.086	0.274
	C_2H_4	0.010	0.060	0.012	0.063	0.012	0.106	0.048	0.155	0.046	0.188	0.075	0.188	0.034	0.087	0.040	0.188
	C_2H_2	0.023	0.048	20.81	4.286	$1.8e^{-3}$	0.041	$9.9e^{-3}$	0.077	0.051	0.135	709.0	0.135	3.688	1.920	2517	50.05
	CO_2	0.178	0.137	0.017	0.131	0.658	0.201	0.011	0.086	0.623	0.550	7.778	0.550	11.76	3.429	709.1	26.62
	CO	0.015	0.076	0.019	0.134	0.066	0.133	$4.2e^{-3}$	0.065	0.062	0.146	0.049	0.146	2.971	1.701	56.433	7.509
16	H_2	0.032	0.173	0.034	0.180	0.119	0.193	0.022	0.145	0.651	0.436	0.972	0.436	0.268	0.363	0.501	0.704
	C_2H_6	0.048	0.192	$1.2e^{-3}$	0.023	0.420	0.117	0.028	0.147	15.37	3.561	403.9	3.561	0.833	0.247	0.042	0.171
	CH_4	$2.0e^{-3}$	0.045	$2.1e^{-3}$	0.046	0.324	0.311	0.028	0.144	0.637	0.293	62.14	0.293	0.270	0.141	0.234	0.465
	C_2H_4	0.015	0.046	0.016	0.047	0.273	0.054	0.141	0.133	0.385	0.093	$6.0e^{-3}$	0.093	0.878	0.186	57.80	7.420
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.554	0.323	0.322	0.352	0.252	0.382	0.322	0.344	0.969	0.838	4.531	0.838	32.53	5.147	2.903	1.633
	CO	0.093	0.143	0.031	0.168	0.110	0.149	0.029	0.154	0.222	0.224	0.291	0.224	26.16	5.113	17.74	4.000

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
17	H_2	0.034	0.062	0.042	0.084	0.061	0.146	0.035	0.060	0.050	0.116	0.058	0.116	1.688	0.725	0.057	0.199
	C_2H_6	0.029	0.071	0.024	0.057	1.440	0.216	0.044	0.197	4.430	2.046	199.9	2.046	2.519	0.365	0.229	0.303
	CH_4	0.025	0.016	0.025	0.016	0.416	0.224	0.020	0.079	0.300	0.239	1.754	0.239	1.569	0.329	0.624	0.715
	C_2H_4	0.083	$7.8e^{-3}$	0.084	0.008	0.638	0.195	0.064	0.253	0.630	0.230	0.143	0.230	3.582	0.469	1.021	0.873
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.177	0.421	0.372	0.338	0.254	0.331	0.359	0.414	0.893	0.828	1.391	0.828	2.195	0.847	1.304	1.139
	CO	0.013	0.104	0.056	0.107	0.177	0.135	0.034	0.167	0.103	0.215	0.107	0.215	0.737	0.225	0.162	0.132
18	H_2	0.081	0.241	0.083	0.247	4.197	1.110	0.035	0.072	2.839	1.241	0.549	1.241	2.734	1.137	0.105	0.321
	C_2H_6	1.787	1.332	1.647	1.273	0.899	0.675	0.558	0.690	1.683	0.843	0.303	0.843	1.202	1.096	0.948	0.774
	CH_4	0.089	0.286	0.088	0.285	0.012	0.104	0.117	0.317	$1.2e^{-3}$	0.031	0.025	0.031	33.20	3.695	73.11	5.760
	C_2H_4	0.965	0.966	0.524	0.713	2.047	1.189	0.082	0.236	1.757	1.112	0.134	1.112	11.15	1.497	2.240	1.233
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.216	0.423	0.135	0.345	0.709	0.276	0.080	0.283	0.535	0.651	0.159	0.651	80.35	6.771	261.7	16.17
	CO	0.538	0.715	0.331	0.570	2.187	0.522	0.270	0.440	2.700	1.123	0.277	1.123	0.034	0.167	0.162	0.256

Trans No.	Feat	LS-SVM (Exp. 1)				LS-SVM (Exp. 2)				ϵ-SVR (Exp. 2)				LS-SVM-UV (Exp. 2)			
		Valid		Test		Valid		Test		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
All	H_2	0.189	0.435	0.638	0.785	1.056	1.026	0.336	0.574	1.621	1.273	1.199	1.080	6.412	2.532	160.8	12.67
	C_2H_6	0.350	0.567	0.579	0.750	0.420	0.553	0.290	0.520	1.331	1.140	43.12	6.556	64.50	8.030	4.182	2.026
	CH_4	0.044	0.184	0.094	0.302	0.156	0.333	0.058	0.232	0.101	0.274	78.04	8.808	11.74	3.420	5.437	2.263
	C_2H_4	0.353	0.573	0.614	0.775	0.203	0.424	0.748	0.837	0.213	0.435	0.890	0.917	2.733	1.627	7.135	2.668
	C_2H_2	0.320	0.564	2.630	1.622	0.719	0.847	0.533	0.706	2.706	1.622	335.2	18.31	99.54	9.935	226.1	15.04
	CO_2	0.290	0.510	0.377	0.604	0.377	0.597	0.349	0.584	0.797	0.884	1.111	0.981	796.3	28.22	278.0	16.67
	CO	0.276	0.514	0.313	0.559	0.849	0.901	0.263	0.513	0.986	0.985	0.414	0.626	15.18	3.889	17.99	4.239

A.2.2 Input Type Comparison

Tables 10, 11 and 12 compare the results of Experiments 1 and 2, and the in-built ϵ-SVR, for the original and augmented input vectors respectively.

Table 10: Input Vector Comparison - LS-SVM (Experiment 1)

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
1	H_2	0.651	0.563	0.655	0.565	0.330	0.513	0.330	0.513
	C_2H_6	0.710	0.263	0.856	0.228	0.697	0.257	0.721	0.209
	CH_4	0.206	0.122	0.207	0.122	0.160	0.108	0.161	0.108
	C_2H_4	1.052	0.358	0.314	0.415	0.762	0.334	0.916	0.232
	C_2H_2	0.085	0.257	0.327	0.196	0.015	0.075	0.673	0.177
	CO_2	1.646	0.392	3.700	0.242	1.416	0.322	2.417	0.289
	CO	1.074	0.624	1.748	0.707	1.188	0.600	1.713	0.682
2	H_2	$4.9e^{-3}$	0.064	0.011	0.097	$4.6e^{-3}$	0.063	$4.7e^{-3}$	0.063
	C_2H_6	9.555	3.027	9.647	3.025	12.16	3.465	9.661	3.015
	CH_4	0.097	0.300	0.096	0.300	0.128	0.304	0.127	0.304
	C_2H_4	45.55	6.190	47.62	6.143	46.78	6.147	47.76	6.131
	C_2H_2	232.7	13.25	233.7	13.25	231.7	13.26	234.6	13.25
	CO_2	0.388	0.528	0.519	0.598	0.471	0.555	0.616	0.593
	CO	0.106	0.321	0.112	0.331	0.115	0.315	0.111	0.327
3	H_2	$6.0e^{-3}$	0.015	$6.0e^{-3}$	0.015	$4.2e^{-5}$	$5.1e^{-3}$	$3.2e^{-4}$	0.017
	C_2H_6	$3.5e^{-4}$	$1.5e^{-3}$	$3.5e^{-4}$	$1.5e^{-3}$	$5.6e^{-4}$	0.016	$5.5e^{-4}$	0.016
	CH_4	$1.8e^{-4}$	$2.7e^{-3}$	$1.8e^{-4}$	$2.7e^{-3}$	$7.6e^{-5}$	$8.4e^{-3}$	$7.7e^{-5}$	$8.4e^{-3}$
	C_2H_4	$1.5e^{-3}$	0.037	$1.4e^{-3}$	0.034	$1.4e^{-3}$	0.037	$1.3e^{-3}$	0.032
	C_2H_2	0.109	0.112	0.135	0.150	$5.7e^{-3}$	0.053	$7.0e^{-3}$	0.056
	CO_2	0.722	0.514	0.874	0.723	0.331	0.478	0.307	0.446
	CO	0.577	0.656	0.483	0.669	0.396	0.623	0.277	0.511

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
4	H_2	0.450	0.668	12.74	1.379	0.651	0.804	21.84	1.681
	C_2H_6	0.063	0.228	18.81	0.646	0.444	0.615	21.29	0.801
	CH_4	0.541	0.314	5.061	2.246	0.513	0.396	3.293	1.429
	C_2H_4	6.122	2.405	8.087	2.737	1.402	0.765	6.553	1.370
	C_2H_2	1.034	0.824	56.35	7.030	4.237	1.104	40.33	3.801
	CO_2	1.602	1.060	4.194	1.544	0.577	0.741	0.567	0.752
	CO	0.184	0.314	1.379	0.449	0.116	0.176	0.120	0.146
5	H_2	0.042	0.023	0.078	0.130	0.016	0.026	0.036	0.117
	C_2H_6	$3.8e^{-5}$	$5.9e^{-3}$	$3.8e^{-5}$	$6.0e^{-3}$	0.170	0.403	0.158	0.392
	CH_4	$9.6e^{-4}$	0.014	$9.8e^{-4}$	0.014	$1.5e^{-3}$	0.013	$1.5e^{-3}$	0.013
	C_2H_4	0.242	0.196	0.296	0.051	0.113	0.046	0.119	0.048
	C_2H_2	$9.8e^{-3}$	0.057	0.015	0.056	$8.5e^{-3}$	0.066	0.010	0.093
	CO_2	0.105	0.307	0.105	0.301	0.087	0.295	0.074	0.272
	CO	0.185	0.176	0.213	0.174	0.118	0.171	0.116	0.156
6	H_2	0.085	0.035	0.085	0.035	0.070	0.030	0.070	0.030
	C_2H_6	0.030	0.021	0.030	0.021	0.048	0.146	0.035	0.096
	CH_4	$2.2e^{-3}$	0.015	$2.2e^{-3}$	0.015	0.030	0.016	0.030	0.016
	C_2H_4	0.041	0.153	0.041	0.152	0.348	0.163	0.335	0.176
	C_2H_2	0.015	0.048	0.012	0.049	0.173	0.028	0.439	0.600
	CO_2	0.025	0.152	0.080	0.157	0.022	0.137	0.018	0.122
	CO	0.019	0.135	0.053	0.224	0.071	0.142	0.079	0.118
7	H_2	0.021	0.022	0.026	0.052	$8.2e^{-3}$	0.030	0.015	0.069
	C_2H_6	0.011	0.027	0.011	0.027	0.169	0.405	0.063	0.230
	CH_4	$2.1e^{-3}$	0.032	$2.1e^{-3}$	0.032	$2.7e^{-3}$	0.023	$2.6e^{-3}$	0.023
	C_2H_4	0.038	0.138	0.043	0.125	0.076	0.087	0.069	0.135
	C_2H_2	$3.7e^{-3}$	0.042	0.024	0.152	0.159	0.110	0.343	0.361
	CO_2	0.032	0.032	0.018	0.075	0.014	0.052	$2.3e^{-3}$	0.047
	CO	0.076	0.106	0.028	0.086	0.149	0.075	0.082	0.036

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
8	H_2	1.097	0.885	1.865	0.873	1.425	0.939	1.451	0.943
	C_2H_6	4.1e^{-3}	5.8e^{-3}	4.1e^{-3}	5.8e^{-3}	0.050	0.198	0.117	0.325
	CH_4	3.8e^{-3}	0.011	3.8e^{-3}	0.011	6.7e^{-3}	0.013	6.7e^{-3}	0.013
	C_2H_4	0.016	0.120	0.012	0.085	0.261	0.075	0.241	0.128
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.126	0.340	0.120	0.322	0.140	0.308	0.122	0.298
	CO	0.046	0.194	0.048	0.191	0.066	0.168	0.039	0.171
9	H_2	0.073	0.030	0.073	0.031	0.099	0.015	0.098	0.020
	C_2H_6	4.6e^{-3}	0.022	4.6e^{-3}	0.022	6.0e^{-3}	0.069	5.6e^{-3}	0.059
	CH_4	0.101	0.049	0.103	0.048	0.084	0.071	0.085	0.070
	C_2H_4	3.2e^{-3}	0.029	2.4e^{-3}	0.023	3.0e^{-3}	0.055	3.0e^{-3}	0.055
	C_2H_2	8.4e^{-5}	9.2e^{-3}	9.9e^{-5}	9.7e^{-3}	0.011	9.3e^{-3}	0.011	0.020
	CO_2	0.025	0.154	0.232	0.357	0.176	0.233	1.807	0.494
	CO	0.027	0.162	0.314	0.268	0.036	0.163	0.738	0.331
10	H_2	7.4e^{-4}	0.016	7.3e^{-4}	0.016	4.0e^{-4}	0.020	4.0e^{-4}	0.020
	C_2H_6	5.3e^{-6}	1.7e^{-3}	5.0e^{-6}	1.6e^{-3}	0.017	0.132	6.0e^{-3}	0.077
	CH_4	1.8e^{-4}	4.4e^{-3}	1.9e^{-4}	4.4e^{-3}	4.0e^{-5}	3.5e^{-3}	4.1e^{-5}	3.5e^{-3}
	C_2H_4	1.1e^{-3}	0.027	1.3e^{-3}	0.027	1.4e^{-5}	2.7e^{-3}	2.0e^{-5}	4.3e^{-3}
	C_2H_2	0.119	0.220	0.130	0.271	3.2e^{-5}	2.6e^{-3}	2.3e^{-3}	0.046
	CO_2	0.029	0.071	0.032	0.161	7.8e^{-3}	0.085	8.9e^{-3}	0.082
	CO	0.496	0.118	0.231	0.146	0.462	0.112	0.752	0.364
11	H_2	0.072	0.108	0.072	0.108	0.069	0.061	0.066	0.068
	C_2H_6	9.8e^{-5}	5.0e^{-3}	1.0e^{-4}	5.0e^{-3}	1.9e^{-3}	0.043	2.3e^{-3}	0.047
	CH_4	4.4e^{-4}	0.021	4.4e^{-4}	0.021	3.9e^{-4}	0.018	3.8e^{-4}	0.018
	C_2H_4	0.014	0.116	0.014	0.119	5.0e^{-3}	8.3e^{-3}	4.8e^{-3}	0.014
	C_2H_2	0.065	0.204	1.711	1.292	0.067	0.064	0.633	0.685
	CO_2	0.032	0.131	0.107	0.151	0.011	0.093	0.022	0.058
	CO	0.527	0.427	0.544	0.412	0.589	0.417	0.577	0.419

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
12	H_2	0.089	0.137	0.184	0.221	0.066	0.195	0.104	0.320
	C_2H_6	$5.7e^{-4}$	$4.5e^{-3}$	$5.8e^{-4}$	$4.5e^{-3}$	$2.9e^{-3}$	0.037	$1.0e^{-3}$	$6.3e^{-3}$
	CH_4	$4.2e^{-5}$	$5.0e^{-3}$	$4.0e^{-5}$	$5.0e^{-3}$	$6.9e^{-4}$	$3.3e^{-3}$	$6.7e^{-4}$	$3.3e^{-3}$
	C_2H_4	0.080	0.175	0.082	0.184	0.082	0.069	0.078	0.072
	C_2H_2	0.213	0.156	0.203	0.154	0.275	0.231	0.410	0.517
	CO_2	0.063	0.081	0.134	0.317	0.015	0.094	$8.1e^{-3}$	0.090
	CO	0.063	0.123	0.041	0.107	$8.3e^{-3}$	0.042	0.028	0.007
13	H_2	1.487	0.623	1.726	0.696	1.645	0.219	3.686	0.332
	C_2H_6	$8.2e^{-3}$	0.040	$8.3e^{-3}$	0.040	0.014	0.096	0.013	0.074
	CH_4	0.020	0.106	1.744	0.529	0.013	0.105	1.446	0.472
	C_2H_4	0.479	0.669	0.394	0.605	0.728	0.819	0.821	0.836
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.773	0.558	0.619	0.584	0.627	0.434	0.698	0.419
	CO	0.546	0.608	0.949	0.771	0.592	0.548	1.079	0.612
14	H_2	0.029	0.155	0.027	0.137	0.231	0.337	0.231	0.337
	C_2H_6	1.790	1.141	1.724	1.098	3.371	1.825	2.341	1.440
	CH_4	0.043	0.167	0.044	0.167	0.042	0.170	0.042	0.170
	C_2H_4	0.495	0.599	0.572	0.635	1.118	1.054	0.970	0.923
	C_2H_2	1.035	0.108	3.983	0.896	0.680	0.094	6.617	1.303
	CO_2	0.253	0.503	0.158	0.382	0.205	0.451	0.141	0.374
	CO	0.412	0.632	0.309	0.554	0.396	0.627	0.301	0.536
15	H_2	0.188	0.148	0.192	0.153	0.095	0.108	0.102	0.116
	C_2H_6	0.105	0.157	0.106	0.157	0.362	0.520	0.340	0.502
	CH_4	0.064	0.162	0.064	0.162	$5.5e^{-3}$	0.051	$5.4e^{-3}$	0.050
	C_2H_4	0.029	0.158	0.029	0.159	0.010	0.060	0.012	0.063
	C_2H_2	0.218	0.148	0.221	0.150	0.023	0.048	20.81	4.286
	CO_2	0.189	0.035	0.088	0.244	0.178	0.137	0.017	0.131
	CO	0.056	0.074	0.060	0.094	0.015	0.076	0.019	0.134

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
16	H_2	0.235	0.247	0.235	0.247	0.032	0.173	0.034	0.180
	C_2H_6	0.082	0.030	0.083	0.031	0.048	0.192	$1.2e^{-3}$	0.023
	CH_4	0.099	0.024	0.099	0.024	$2.0e^{-3}$	0.045	$2.1e^{-3}$	0.046
	C_2H_4	0.063	0.150	0.062	0.151	0.015	0.046	0.016	0.047
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.360	0.403	0.376	0.451	0.554	0.323	0.322	0.352
	CO	0.039	0.183	0.082	0.260	0.093	0.143	0.031	0.168
17	H_2	0.043	0.030	0.071	0.126	0.034	0.062	0.042	0.084
	C_2H_6	0.248	0.100	0.248	0.100	0.029	0.071	0.024	0.057
	CH_4	0.268	0.058	0.268	0.058	0.025	0.016	0.025	0.016
	C_2H_4	0.384	0.508	0.481	0.576	0.083	$7.8e^{-3}$	0.084	$7.9e^{-3}$
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.372	0.399	0.542	0.702	0.177	0.421	0.372	0.338
	CO	0.124	0.109	0.190	0.102	0.013	0.104	0.056	0.107
18	H_2	0.156	0.246	0.155	0.230	0.081	0.241	0.083	0.247
	C_2H_6	1.392	1.149	1.321	1.114	1.787	1.332	1.647	1.273
	CH_4	0.082	0.273	0.082	0.273	0.089	0.286	0.088	0.285
	C_2H_4	0.619	0.731	0.594	0.721	0.965	0.966	0.524	0.713
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.396	0.510	0.094	0.306	0.216	0.423	0.135	0.345
	CO	0.536	0.717	0.310	0.503	0.538	0.715	0.331	0.570
All	H_2	0.181	0.396	0.425	0.652	0.189	0.435	0.638	0.785
	C_2H_6	0.304	0.501	0.534	0.707	0.350	0.567	0.579	0.750
	CH_4	0.072	0.228	0.186	0.426	0.044	0.184	0.094	0.302
	C_2H_4	0.584	0.735	0.632	0.776	0.353	0.573	0.614	0.775
	C_2H_2	0.164	0.389	1.451	1.204	0.320	0.564	2.630	1.622
	CO_2	0.326	0.548	0.495	0.680	0.290	0.510	0.377	0.604
	CO	0.283	0.531	0.348	0.586	0.276	0.514	0.313	0.559

Table 11: Input Vector Comparison - LS-SVM (Experiment 2)

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
1	H_2	0.492	0.356	0.323	0.479	0.378	0.353	0.227	0.443
	C_2H_6	0.141	0.252	0.807	0.767	0.534	0.254	0.194	0.439
	CH_4	0.118	0.175	0.148	0.144	0.097	0.170	0.118	0.137
	C_2H_4	0.512	0.228	0.893	0.661	0.530	0.176	0.089	0.265
	C_2H_2	0.151	0.291	0.341	0.270	0.118	0.263	0.204	0.215
	CO_2	0.521	0.701	1.399	0.498	0.592	0.711	1.876	0.485
	CO	0.266	0.470	0.776	0.583	0.238	0.419	0.847	0.565
2	H_2	0.131	0.296	0.029	0.067	0.216	0.294	$4.2e^{-3}$	0.063
	C_2H_6	0.832	0.328	10.83	3.107	0.155	0.319	10.61	2.984
	CH_4	0.613	0.424	0.082	0.287	1.086	0.534	0.134	0.337
	C_2H_4	0.132	0.304	47.02	6.107	0.109	0.328	46.03	5.886
	C_2H_2	$4.2e^{-3}$	0.024	232.9	13.27	$3.0e^{-3}$	0.021	232.6	13.27
	CO_2	0.406	0.614	0.404	0.595	0.430	0.582	0.586	0.592
	CO	0.350	0.453	0.116	0.303	0.174	0.322	0.121	0.345
3	H_2	$1.9e^{-5}$	$4.3e^{-3}$	$7.9e^{-4}$	$6.1e^{-3}$	$2.0e^{-3}$	$4.5e^{-3}$	$1.3e^{-3}$	0.013
	C_2H_6	$9.6e^{-4}$	0.030	$1.1e^{-3}$	0.034	$5.8e^{-4}$	0.016	$3.1e^{-5}$	$4.8e^{-3}$
	CH_4	$2.8e^{-3}$	0.022	$6.1e^{-4}$	0.013	$3.4e^{-4}$	0.018	$7.9e^{-4}$	0.027
	C_2H_4	$6.4e^{-4}$	0.024	$3.6e^{-4}$	0.018	$8.8e^{-4}$	0.023	$3.9e^{-4}$	0.012
	C_2H_2	$4.9e^{-3}$	0.017	$2.3e^{-5}$	$3.5e^{-3}$	$2.6e^{-3}$	0.011	$5.6e^{-5}$	$4.3e^{-3}$
	CO_2	0.362	0.362	0.540	0.687	0.126	0.348	0.368	0.451
	CO	0.628	0.246	0.410	0.640	0.307	0.164	0.290	0.485
4	H_2	0.681	0.389	1.620	1.235	0.769	0.460	1.669	0.776
	C_2H_6	0.416	0.631	6.310	1.129	0.522	0.690	4.642	1.284
	CH_4	0.548	0.606	5.032	1.717	0.565	0.705	4.218	1.575
	C_2H_4	0.861	0.912	8.912	1.949	1.020	0.927	11.68	1.998
	C_2H_2	1.451	1.203	2.161	1.466	1.356	1.148	4.074	1.984
	CO_2	0.397	0.593	0.608	0.740	0.669	0.611	0.528	0.538
	CO	2.796	1.154	0.403	0.630	2.941	1.153	0.459	0.279

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
5	H_2	0.074	0.262	0.029	0.029	0.080	0.277	0.011	0.038
	C_2H_6	0.022	0.112	3.2e^{-3}	0.011	0.947	0.209	0.120	0.043
	CH_4	0.029	0.058	9.9e^{-4}	0.030	0.227	0.372	0.029	0.166
	C_2H_4	0.027	0.075	1.6e^{-3}	0.029	0.082	0.172	0.016	0.037
	C_2H_2	0.522	0.715	1.076	0.058	0.114	0.332	0.208	0.015
	CO_2	0.149	0.384	0.117	0.338	0.276	0.472	0.081	0.283
	CO	0.344	0.496	0.179	0.198	0.307	0.533	0.080	0.156
6	H_2	0.053	0.116	0.035	0.026	0.048	0.123	0.108	0.030
	C_2H_6	0.025	0.082	0.019	0.035	0.028	0.047	0.050	0.052
	CH_4	0.059	0.183	1.7e^{-3}	0.041	0.079	0.223	0.095	0.198
	C_2H_4	0.054	0.134	0.015	0.115	0.032	0.124	0.014	0.113
	C_2H_2	8.0e^{-4}	0.028	7.8e^{-3}	0.027	1.2e^{-3}	0.034	0.026	0.034
	CO_2	0.214	0.373	0.042	0.202	0.258	0.349	0.080	0.086
	CO	0.180	0.297	0.029	0.163	0.143	0.242	0.014	0.082
7	H_2	0.061	0.034	3.9e^{-3}	8.9e^{-3}	0.118	0.069	5.8e^{-3}	0.018
	C_2H_6	0.013	0.110	2.9e^{-3}	0.053	0.154	0.021	0.024	0.018
	CH_4	2.5e^{-3}	0.048	2.5e^{-3}	0.048	0.033	0.163	2.8e^{-3}	0.034
	C_2H_4	6.7e^{-3}	0.059	5.7e^{-3}	0.072	0.062	0.133	0.026	0.116
	C_2H_2	1.622	0.314	0.017	0.041	1.484	0.082	5.7e^{-3}	0.047
	CO_2	7.0e^{-3}	0.075	0.017	0.054	9.1e^{-3}	0.068	7.2e^{-3}	0.026
	CO	0.019	0.119	0.010	0.099	0.019	0.112	0.015	0.074
8	H_2	0.670	0.392	1.011	0.880	0.141	0.359	0.965	0.920
	C_2H_6	0.012	0.074	9.7e^{-3}	0.096	0.032	0.100	0.020	0.038
	CH_4	0.013	0.029	2.1e^{-3}	0.018	0.020	0.057	0.015	0.076
	C_2H_4	0.478	0.174	0.056	0.119	0.332	0.232	3.3e^{-3}	0.050
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.026	0.159	0.133	0.273	0.036	0.128	0.114	0.291
	CO	0.102	0.315	0.051	0.185	0.062	0.245	0.029	0.169

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
9	H_2	0.116	0.280	0.139	0.100	0.125	0.271	0.262	0.067
	C_2H_6	$5.1e^{-3}$	0.035	$9.8e^{-3}$	0.051	$5.6e^{-3}$	0.043	$4.1e^{-4}$	0.020
	CH_4	0.041	0.194	0.097	0.135	0.055	0.235	0.121	0.106
	C_2H_4	$5.8e^{-3}$	0.075	0.023	0.068	$7.0e^{-3}$	0.081	$4.4e^{-3}$	0.026
	C_2H_2	0.010	0.045	$4.4e^{-3}$	0.015	0.014	$8.0e^{-3}$	0.011	0.027
	CO_2	0.634	0.673	0.198	0.377	0.916	0.675	0.791	0.358
	CO	1.784	1.143	0.193	0.286	1.092	1.006	0.853	0.334
10	H_2	$2.6e^{-3}$	0.045	$7.0e^{-4}$	0.020	$2.0e^{-3}$	0.042	$5.5e^{-4}$	0.021
	C_2H_6	$7.2e^{-4}$	0.026	$1.5e^{-4}$	0.011	$6.2e^{-4}$	0.017	$1.9e^{-4}$	0.010
	CH_4	$2.5e^{-4}$	0.013	$5.8e^{-5}$	$5.7e^{-3}$	$7.0e^{-4}$	0.020	$2.0e^{-4}$	$8.9e^{-3}$
	C_2H_4	$1.6e^{-4}$	0.011	$1.6e^{-4}$	$7.3e^{-3}$	$1.6e^{-4}$	$6.8e^{-3}$	$6.9e^{-5}$	$7.8e^{-3}$
	C_2H_2	$1.2e^{-3}$	0.015	$5.2e^{-4}$	0.014	$3.4e^{-4}$	0.016	$2.4e^{-4}$	0.011
	CO_2	0.102	0.247	0.050	0.209	0.064	0.149	$4.7e^{-3}$	0.068
	CO	3.393	1.278	0.041	0.200	5.311	1.424	0.043	0.206
11	H_2	39.59	4.427	0.108	0.217	39.93	4.462	0.113	0.245
	C_2H_6	$3.8e^{-4}$	0.019	$3.7e^{-4}$	$6.8e^{-3}$	$1.0e^{-3}$	0.015	$2.4e^{-4}$	$7.6e^{-3}$
	CH_4	0.014	0.107	0.030	0.173	0.025	0.067	$1.2e^{-3}$	0.033
	C_2H_4	0.061	0.075	$7.7e^{-3}$	0.088	0.055	0.095	$3.9e^{-3}$	0.043
	C_2H_2	22.37	2.819	0.657	0.805	22.92	3.095	0.829	0.521
	CO_2	0.261	0.455	0.191	0.207	0.196	0.306	0.059	0.044
	CO	0.391	0.488	0.655	0.442	0.905	0.403	0.634	0.396
12	H_2	0.175	0.382	0.022	0.148	0.131	0.345	0.032	0.174
	C_2H_6	0.093	0.253	$1.1e^{-3}$	0.032	0.601	0.542	$5.5e^{-3}$	0.025
	CH_4	0.042	0.100	$1.9e^{-4}$	0.014	0.181	0.315	$1.6e^{-3}$	0.039
	C_2H_4	0.058	0.236	0.031	0.162	0.195	0.286	0.010	0.097
	C_2H_2	25.89	4.013	0.148	0.339	31.28	4.299	0.077	0.202
	CO_2	0.052	0.206	0.041	0.140	0.046	0.172	0.042	0.100
	CO	0.330	0.376	0.011	0.079	0.120	0.213	$1.2e^{-3}$	$9.6e^{-3}$

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
13	H_2	5.924	2.051	2.638	0.652	7.601	2.108	6.440	0.457
	C_2H_6	$7.9e^{-3}$	0.063	$3.2e^{-3}$	0.055	1.127	0.068	0.055	0.023
	CH_4	0.145	0.173	0.028	0.141	0.631	0.425	0.558	0.650
	C_2H_4	0.365	0.307	0.618	0.756	0.123	0.322	0.424	0.651
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	1.591	0.581	0.329	0.527	1.637	0.563	0.802	0.337
	CO	2.561	1.592	0.528	0.666	2.468	1.532	1.667	0.530
14	H_2	2.028	0.275	0.040	0.106	1.959	0.195	0.063	0.112
	C_2H_6	5.095	0.730	0.585	0.492	5.806	0.777	0.930	0.595
	CH_4	0.014	0.095	0.022	0.128	0.015	0.124	0.025	0.151
	C_2H_4	1.042	0.466	0.630	0.448	1.329	0.590	0.593	0.347
	C_2H_2	1.212	0.131	3.678	0.523	1.011	0.083	3.993	0.625
	CO_2	1.012	0.334	0.116	0.302	1.022	0.120	0.169	0.316
	CO	1.117	0.194	0.253	0.491	1.625	0.168	0.248	0.470
15	H_2	0.118	0.051	0.407	0.291	0.161	0.018	0.174	0.244
	C_2H_6	0.189	0.098	0.016	0.110	0.089	0.049	0.097	0.152
	CH_4	0.204	0.148	0.071	0.171	0.034	0.076	$8.1e^{-3}$	0.084
	C_2H_4	0.081	0.123	0.021	0.132	0.012	0.106	0.048	0.155
	C_2H_2	0.190	0.201	0.948	0.391	$1.8e^{-3}$	0.041	0.010	0.077
	CO_2	0.343	0.347	0.113	0.153	0.658	0.201	0.011	0.086
	CO	0.022	0.144	0.061	0.145	0.066	0.133	$4.2e^{-3}$	0.065
16	H_2	0.591	0.147	0.095	0.193	0.119	0.193	0.022	0.145
	C_2H_6	1.407	0.271	0.057	0.138	0.420	0.117	0.028	0.147
	CH_4	0.808	0.230	0.050	0.106	0.324	0.311	0.028	0.144
	C_2H_4	1.050	0.238	0.100	0.252	0.273	0.054	0.141	0.133
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.198	0.355	1.089	0.531	0.252	0.382	0.322	0.344
	CO	0.071	0.248	0.155	0.335	0.110	0.149	0.029	0.154

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
17	H_2	0.230	0.108	0.196	0.096	0.061	0.146	0.035	0.060
	C_2H_6	2.507	0.376	0.181	0.144	1.440	0.216	0.044	0.197
	CH_4	1.221	0.221	0.193	0.085	0.416	0.224	0.020	0.079
	C_2H_4	2.564	0.432	0.276	0.336	0.638	0.195	0.064	0.253
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.210	0.458	0.963	0.913	0.254	0.331	0.359	0.414
	CO	0.337	0.270	0.092	0.283	0.177	0.135	0.034	0.167
18	H_2	4.964	1.123	0.047	0.103	4.197	1.110	0.035	0.072
	C_2H_6	0.477	0.260	0.391	0.573	0.899	0.675	0.558	0.690
	CH_4	0.016	0.072	0.040	0.200	0.012	0.104	0.117	0.317
	C_2H_4	2.768	1.235	0.126	0.287	2.047	1.189	0.082	0.236
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	2.594	0.620	0.133	0.304	0.709	0.276	0.080	0.283
	CO	5.134	0.765	0.350	0.589	2.187	0.522	0.270	0.440
All	H_2	1.112	1.050	0.278	0.521	1.056	1.026	0.336	0.574
	C_2H_6	0.383	0.575	0.217	0.420	0.420	0.553	0.290	0.520
	CH_4	0.198	0.376	0.053	0.207	0.156	0.333	0.058	0.232
	C_2H_4	0.365	0.588	0.672	0.773	0.203	0.424	0.748	0.837
	C_2H_2	0.885	0.941	0.510	0.702	0.719	0.847	0.533	0.706
	CO_2	0.387	0.617	0.331	0.535	0.377	0.597	0.349	0.584
	CO	1.046	1.008	0.224	0.472	0.849	0.901	0.263	0.513

Table 12: Input Vector Comparison - ϵ-SVR (Experiment 2)

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
1	H_2	0.441	0.605	1.866	1.294	0.425	0.595	6.564	2.451
	C_2H_6	0.036	0.189	1.379	1.174	0.618	0.227	187.6	13.70
	CH_4	0.071	0.217	0.020	0.058	0.053	0.192	$3.9e^{-3}$	0.028
	C_2H_4	0.261	0.105	0.888	0.815	0.726	0.075	0.168	0.410
	C_2H_2	0.152	0.386	1309	32.65	0.233	0.401	4802	64.39
	CO_2	0.436	0.638	1.376	1.070	0.463	0.591	0.710	0.753
	CO	0.259	0.509	3.351	1.638	0.108	0.328	0.394	0.409
2	H_2	0.157	0.344	24.09	4.484	0.181	0.391	21.19	4.555
	C_2H_6	1.783	0.615	13.32	3.371	2.181	0.950	28.67	4.503
	CH_4	0.734	0.569	0.209	0.412	0.871	0.600	4905	69.84
	C_2H_4	0.268	0.459	42.64	5.730	0.150	0.376	43.55	5.614
	C_2H_2	0.079	0.201	1230	34.90	0.088	0.215	1072	32.29
	CO_2	0.445	0.618	3.874	1.474	0.496	0.679	0.678	0.526
	CO	0.173	0.415	0.601	0.674	0.162	0.388	0.094	0.307
3	H_2	$3.4e^{-4}$	0.017	$9.3e^{-3}$	0.096	$5.1e^{-3}$	$7.7e^{-3}$	0.048	0.219
	C_2H_6	$4.3e^{-4}$	0.011	0.313	0.554	3.113	1.694	10.11	3.071
	CH_4	$2.5e^{-3}$	0.036	0.077	0.276	$1.4e^{-3}$	0.032	$5.1e^{-4}$	0.015
	C_2H_4	$5.8e^{-4}$	0.016	$8.6e^{-5}$	$7.6e^{-3}$	$5.6e^{-3}$	0.013	0.011	0.106
	C_2H_2	57.00	4.983	2252	47.42	95.44	9.416	1260	34.24
	CO_2	2.329	0.361	6.127	1.126	2.385	0.598	0.713	0.843
	CO	1.570	0.547	1.435	1.147	0.814	0.835	0.589	0.762
4	H_2	0.190	0.205	17.70	2.730	0.334	0.297	6.053	1.289
	C_2H_6	0.419	0.646	11.87	1.789	0.549	0.727	10.03	1.886
	CH_4	0.493	0.682	10.18	1.756	0.500	0.707	8.958	1.260
	C_2H_4	0.893	0.943	14.61	2.570	0.994	0.954	16.11	1.907
	C_2H_2	1.470	1.182	171.0	13.06	1.754	1.238	2.304	1.214
	CO_2	0.204	0.445	6.952	2.471	1.106	0.600	0.760	0.633
	CO	1.591	0.995	17.80	4.219	2.462	1.219	0.922	0.463

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
5	H_2	0.052	0.226	0.227	0.351	0.088	0.297	0.072	0.268
	C_2H_6	$2.5e^{-3}$	0.039	$3.8e^{-4}$	$6.2e^{-3}$	1.009	0.468	0.043	0.129
	CH_4	0.019	0.096	$1.1e^{-3}$	0.023	$7.5e^{-3}$	0.070	$8.2e^{-4}$	0.027
	C_2H_4	0.011	0.054	$9.9e^{-3}$	0.066	0.030	0.068	$2.8e^{-3}$	0.051
	C_2H_2	4.290	1.597	1914	43.74	0.151	0.383	0.046	0.076
	CO_2	4.523	0.998	0.484	0.688	0.424	0.353	0.337	0.580
	CO	1.670	0.723	0.441	0.486	0.303	0.488	0.083	0.223
6	H_2	0.052	0.227	0.792	0.887	0.053	0.229	0.030	0.061
	C_2H_6	0.011	0.102	0.013	0.033	0.014	0.033	699.1	25.51
	CH_4	0.054	0.230	0.032	0.174	0.019	0.135	1465	38.25
	C_2H_4	0.039	0.161	$5.4e^{-3}$	0.065	0.027	0.143	0.058	0.151
	C_2H_2	5.104	1.565	799.8	27.60	3.116	1.016	2940	54.21
	CO_2	0.166	0.400	0.113	0.291	0.146	0.374	0.242	0.491
	CO	0.102	0.309	0.038	0.190	0.083	0.288	0.268	0.516
7	H_2	0.033	0.172	1.017	0.989	0.044	0.122	2.910	1.689
	C_2H_6	0.012	0.092	2.126	1.379	0.188	0.125	36.24	5.682
	CH_4	$3.1e^{-3}$	0.043	0.044	0.154	$1.7e^{-3}$	0.038	171.8	12.51
	C_2H_4	$8.2e^{-3}$	0.089	3.169	1.765	$7.9e^{-3}$	0.051	0.314	0.494
	C_2H_2	7.242	2.683	1788	39.21	86.98	8.818	1978	43.53
	CO_2	0.124	0.239	0.530	0.635	0.102	0.116	0.084	0.157
	CO	0.108	0.113	0.354	0.272	0.040	0.136	2.523	1.526
8	H_2	1.720	0.538	0.857	0.925	0.332	0.401	0.537	0.685
	C_2H_6	$7.3e^{-3}$	0.071	$3.3e^{-3}$	0.031	0.036	0.120	0.012	0.076
	CH_4	$6.3e^{-3}$	0.072	$3.2e^{-3}$	0.022	$9.8e^{-3}$	0.041	$3.7e^{-3}$	0.040
	C_2H_4	0.474	0.177	0.144	0.234	0.192	0.238	0.025	0.081
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.093	0.058	0.239	0.370	0.011	0.103	0.064	0.201
	CO	0.165	0.406	0.058	0.200	0.092	0.296	0.029	0.167

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
9	H_2	0.342	0.315	0.103	0.185	0.292	0.217	0.511	0.115
	C_2H_6	0.079	0.274	3.350	1.823	293.0	16.76	71.36	8.055
	CH_4	0.159	0.227	0.063	0.237	0.073	0.194	0.165	0.136
	C_2H_4	6.8e^{-3}	0.063	6.8e^{-3}	0.082	0.020	0.137	2.6e^{-3}	0.033
	C_2H_2	104.3	10.16	1422	37.67	3.047	1.732	93.72	9.626
	CO_2	5.444	1.521	2.131	1.350	4.330	1.497	8.574	1.007
	CO	5.564	2.063	3.977	1.356	2.999	1.651	1.969	0.956
10	H_2	0.230	0.478	7.387	2.640	3.4e^{-3}	0.042	8.1e^{-3}	0.062
	C_2H_6	1.8e^{-4}	0.011	0.140	0.369	1.7e^{-3}	9.2e^{-3}	25.25	5.020
	CH_4	0.014	0.115	6.3e^{-3}	0.061	3.5e^{-4}	0.015	186.7	13.18
	C_2H_4	6.8e^{-5}	8.1e^{-3}	1.2e^{-3}	7.2e^{-3}	1.2e^{-4}	2.5e^{-3}	0.024	0.149
	C_2H_2	960.6	28.08	7291	84.20	8.4e^{-3}	8.3e^{-3}	3479	58.74
	CO_2	3.107	0.956	18.78	3.126	1.545	0.329	0.521	0.720
	CO	10.78	2.800	17.88	2.872	4.317	2.064	0.461	0.671
11	H_2	72.73	7.279	177.2	8.452	54.07	5.042	0.780	0.569
	C_2H_6	8.4e^{-4}	0.023	1.677	1.269	2.9e^{-3}	0.050	0.524	0.720
	CH_4	0.221	0.358	0.165	0.374	0.067	0.102	8.6e^{-3}	0.070
	C_2H_4	0.296	0.439	0.590	0.524	0.079	0.115	0.141	0.336
	C_2H_2	265.6	12.77	967.9	29.57	29.39	3.202	59.41	6.889
	CO_2	3.696	1.688	2.518	1.425	0.882	0.923	0.078	0.263
	CO	5.948	1.592	3.239	1.388	0.425	0.592	0.732	0.416
12	H_2	0.630	0.410	0.444	0.661	0.203	0.275	0.053	0.211
	C_2H_6	8.9e^{-3}	0.067	0.464	0.666	1.822	1.310	43.25	6.530
	CH_4	0.065	0.186	0.134	0.351	0.024	0.075	344.8	18.05
	C_2H_4	0.072	0.254	0.279	0.362	0.041	0.194	0.053	0.186
	C_2H_2	9.621	3.075	383.2	19.45	55.67	7.441	1725	40.92
	CO_2	0.221	0.255	0.311	0.445	0.068	0.181	0.040	0.198
	CO	2.582	1.256	0.079	0.274	7.6e^{-3}	0.080	6.9e^{-3}	0.077

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
13	H_2	15.33	3.887	10.44	2.322	16.04	3.227	27.33	1.327
	C_2H_6	0.025	0.054	23.71	4.829	0.560	0.118	3053	54.74
	CH_4	4.579	2.107	4.512	2.123	0.183	0.119	5097	68.89
	C_2H_4	0.482	0.694	2.169	1.317	0.483	0.559	0.442	0.555
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.509	0.648	0.465	0.633	0.448	0.660	0.170	0.404
	CO	6.334	2.016	13.06	3.610	4.989	2.222	4.454	0.975
14	H_2	2.102	0.352	1.811	1.271	2.216	0.372	0.786	0.717
	C_2H_6	4.642	1.202	0.088	0.281	6.235	1.502	0.551	0.405
	CH_4	$3.8e^{-3}$	0.050	0.012	0.064	$7.9e^{-3}$	0.060	0.013	0.079
	C_2H_4	0.895	0.581	0.233	0.335	1.994	0.857	0.800	0.312
	C_2H_2	0.612	0.509	194.2	12.51	0.096	0.249	507.6	21.94
	CO_2	1.040	0.676	1.194	0.618	1.250	0.720	1.016	0.534
	CO	2.057	0.869	1.782	0.773	1.697	0.750	0.132	0.346
15	H_2	0.170	0.248	1.450	0.691	0.101	0.174	0.758	0.469
	C_2H_6	0.159	0.210	$3.1e^{-3}$	0.047	0.038	0.123	70.68	8.057
	CH_4	0.174	0.228	$7.6e^{-3}$	0.074	0.022	0.110	$5.7e^{-4}$	0.023
	C_2H_4	0.094	0.197	0.014	0.094	0.046	0.188	0.075	0.195
	C_2H_2	0.434	0.377	202.2	13.97	0.051	0.135	709.0	26.41
	CO_2	0.508	0.503	13.23	2.131	0.623	0.550	7.778	1.806
	CO	0.040	0.115	2.343	1.111	0.062	0.146	0.049	0.190
16	H_2	0.999	0.235	0.826	0.885	0.651	0.436	0.972	0.895
	C_2H_6	1.590	0.363	0.281	0.456	15.37	3.561	403.9	17.81
	CH_4	1.790	0.452	0.271	0.444	0.637	0.293	62.14	7.774
	C_2H_4	1.280	0.264	0.277	0.524	0.385	0.093	$6.0e^{-3}$	0.050
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	1.593	0.523	1.361	0.746	0.969	0.838	4.531	0.921
	CO	1.792	0.675	5.824	2.413	0.222	0.224	0.291	0.484

Trans No.	Feat	Original				Augmented			
		Valid		Test		Valid		Test	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
17	H_2	0.650	0.173	0.685	0.508	0.050	0.116	0.058	0.227
	C_2H_6	3.965	0.862	1.904	1.313	4.430	2.046	199.9	12.93
	CH_4	3.446	0.669	1.179	1.055	0.300	0.239	1.754	1.287
	C_2H_4	3.927	0.729	0.982	0.919	0.630	0.230	0.143	0.377
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	5.941	0.816	3.058	1.631	0.893	0.828	1.391	0.613
	CO	4.572	0.203	2.310	1.520	0.103	0.215	0.107	0.281
18	H_2	2.467	1.183	0.381	0.571	2.839	1.241	0.549	0.730
	C_2H_6	0.380	0.475	0.189	0.433	1.683	0.843	0.303	0.549
	CH_4	$3.9e^{-3}$	0.055	0.021	0.126	$1.2e^{-3}$	0.031	0.025	0.116
	C_2H_4	2.860	1.376	0.085	0.255	1.757	1.112	0.134	0.324
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	1.899	1.326	0.780	0.507	0.535	0.651	0.159	0.305
	CO	0.784	0.884	0.287	0.507	2.700	1.123	0.277	0.348
All	H_2	2.266	1.505	2.829	1.675	1.621	1.273	1.199	1.080
	C_2H_6	0.370	0.562	0.978	0.968	1.331	1.140	43.12	6.556
	CH_4	0.427	0.590	0.176	0.417	0.101	0.274	78.04	8.808
	C_2H_4	0.398	0.603	1.183	1.053	0.213	0.435	0.890	0.917
	C_2H_2	10.93	3.304	456.2	21.35	2.706	1.622	335.2	18.31
	CO_2	1.798	1.315	2.790	1.667	0.797	0.884	1.111	0.981
	CO	2.417	1.551	2.695	1.640	0.986	0.985	0.414	0.626

A.2.3 Testing Errors

Tables 13 and 14 show all of the Mean Squared Error results obtained for the testing errors of all experiments conducted and all baseline tests, using the original and augmented input vectors respectively.

Table 13: Testing Results Summary - Original Features

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
1	H_2	0.655	0.565	0.323	0.479	1.866	1.294	0.477	0.600	0.827	0.819	0.788	0.741	0.775	0.907	1.205	1.058
	C_2H_6	0.856	0.228	0.807	0.767	1.379	1.174	0.559	0.179	0.012	0.007	0.016	0.005	1.713	1.366	1.209	1.134
	CH_4	0.207	0.122	0.148	0.144	0.020	0.058	0.061	0.116	0.071	0.058	0.001	0.023	0.499	0.658	1.342	0.508
	C_2H_4	0.314	0.415	0.893	0.661	0.888	0.815	0.091	0.266	0.036	0.029	0.001	0.019	0.992	1.040	1.179	1.139
	C_2H_2	0.327	0.196	0.341	0.270	1309	32.65	1310	36.19	0.046	0.064	0.000	0.012	0.695	0.869	1.214	1.147
	CO_2	3.700	0.242	1.399	0.498	1.376	1.070	0.921	0.516	$46e^{+4}$	73.61	$38e^{+4}$	370.4	1.514	1.290	1.269	1.181
	CO	1.748	0.707	0.776	0.583	3.351	1.638	27.27	5.187	669.2	23.43	1303	17.95	1.261	1.177	1.272	1.182
2	H_2	0.011	0.097	0.029	0.067	24.09	4.484	0.109	0.301	0.730	0.252	0.046	0.210	0.577	0.781	1.245	0.867
	C_2H_6	9.647	3.025	10.83	3.107	13.32	3.371	17.02	3.837	0.791	0.828	0.979	0.864	0.905	0.963	1.246	1.046
	CH_4	0.096	0.300	0.082	0.287	0.209	0.412	6.859	2.615	4.643	1.498	2.148	1.369	1.004	1.037	1.175	1.048
	C_2H_4	47.62	6.143	47.02	6.107	42.64	5.730	45.36	6.162	95.37	8.546	103.6	8.750	0.333	0.570	1.223	0.647
	C_2H_2	233.7	13.25	232.9	13.27	1230	34.90	234.1	13.25	164.6	11.15	165.7	11.15	0.295	0.497	1.321	0.433
	CO_2	0.519	0.598	0.404	0.595	3.874	1.474	3.247	1.203	$21e^{+4}$	444.5	$50e^{+4}$	639.9	1.054	1.065	1.414	1.103
	CO	0.112	0.331	0.116	0.303	0.601	0.674	1.268	1.098	620.7	23.19	1186	31.50	0.981	1.027	1.432	1.093

Trans No.	**Feat**	**LS-SVM(1)**		**LS-SVM(2)**		**ϵ-SVR(2)**		**LS-SVM-UV(2)**		**ARIMA**		**E.S.**		**Naive**		**Mean**	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
3	H_2	6.0e^{-3}	0.015	7.9e^{-4}	6.1e^{-3}	9.3e^{-3}	0.096	3.3e^{-4}	0.017	5975	6.640	0.192	0.216	0.665	0.840	1.320	0.861
	C_2H_6	3.5e^{-4}	1.5e^{-3}	1.1e^{-3}	0.034	0.313	0.554	0.056	0.105	125.4	1.010	0.257	0.210	0.669	0.842	1.322	0.866
	CH_4	1.8e^{-4}	2.7e^{-3}	6.1e^{-4}	0.013	0.077	0.276	4.6e^{-3}	0.067	341.9	1.702	0.790	0.316	0.671	0.843	1.317	0.850
	C_2H_4	1.4e^{-3}	0.034	3.6e^{-4}	0.018	8.6e^{-5}	7.6e^{-3}	3.1e^{-3}	0.043	4710	5.997	1.179	0.552	0.668	0.841	1.321	0.865
	C_2H_2	0.135	0.150	2.3e^{-5}	3.5e^{-3}	2252	47.42	6.6e^{-4}	0.025	6472	6.771	0.016	0.028	0.667	0.840	1.321	0.864
	CO_2	0.874	0.723	0.540	0.687	6.127	1.126	13.37	3.294	32e^{+4}	303.1	43e^{+4}	652.6	1.156	1.119	1.199	1.122
	CO	0.483	0.669	0.410	0.640	1.435	1.147	0.596	0.768	1213	25.60	1689	41.08	1.155	1.113	1.195	1.065
4	H_2	12.74	1.379	1.620	1.235	17.70	2.730	11.71	2.347	862.1	24.01	764.8	23.94	0.530	0.757	1.228	1.122
	C_2H_6	18.81	0.646	6.310	1.129	11.87	1.789	10.64	2.849	1496	1.639	633.2	13.10	0.713	0.875	1.173	1.124
	CH_4	5.061	2.246	5.032	1.717	10.18	1.756	10.68	2.389	9225	20.70	4430	46.47	0.707	0.870	1.169	1.105
	C_2H_4	8.087	2.737	8.912	1.949	14.61	2.570	17.98	3.487	3.7e^{+4}	22.15	1.8e^{+4}	80.42	0.828	0.944	1.190	1.133
	C_2H_2	56.35	7.030	2.161	1.466	171.0	13.07	4.549	2.129	21.12	1.548	9.454	1.967	1.000	1.040	1.239	0.878
	CO_2	4.194	1.544	0.608	0.740	6.952	2.471	1.047	0.940	15e^{+4}	380.5	32e^{+4}	519.0	1.225	1.145	1.180	1.036
	CO	1.379	0.449	0.403	0.630	17.80	4.219	163.5	11.76	117.5	7.648	1974	8.335	1.020	1.042	1.123	0.954
5	H_2	0.078	0.130	0.029	0.029	0.227	0.351	0.341	0.314	5.333	0.767	2.188	0.999	0.415	0.667	1.205	0.953
	C_2H_6	3.8e^{-5}	6.0e^{-3}	3.2e^{-3}	0.011	3.8e^{-4}	6.2e^{-3}	3.3e^{-3}	0.051	2921	3.948	9.843	1.456	1.044	1.063	1.327	1.147
	CH_4	9.8e^{-4}	0.014	9.9e^{-4}	0.030	1.1e^{-3}	0.023	4.2e^{-3}	0.064	70.02	0.544	0.505	0.381	0.878	0.975	1.294	1.122
	C_2H_4	0.296	0.051	1.6e^{-3}	0.029	9.9e^{-3}	0.066	0.056	0.236	10.40	0.398	0.117	0.317	0.872	0.971	1.298	1.113
	C_2H_2	0.015	0.056	1.076	0.058	1914	43.74	0.777	0.731	0.017	0.030	0.000	0.000	0.767	0.906	1.183	1.032
	CO_2	0.105	0.301	0.117	0.338	0.484	0.688	2513	49.90	8.0e^{+4}	185.1	9.2e^{+4}	151.5	1.095	1.087	1.215	1.145
	CO	0.213	0.174	0.179	0.198	0.441	0.486	0.266	0.371	440.6	7.117	351.1	8.477	1.112	1.096	1.251	1.161

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
6	H_2	0.085	0.035	0.035	0.026	0.792	0.887	0.719	0.846	0.168	0.000	$3.8e^{-3}$	0.040	0.319	0.572	1.205	0.798
	C_2H_6	0.030	0.021	0.019	0.035	0.013	0.033	$9.0e^{-3}$	0.016	0.896	0.089	$7.7e^{-3}$	0.042	0.812	0.933	1.399	1.047
	CH_4	$2.2e^{-3}$	0.015	$1.7e^{-3}$	0.041	0.032	0.174	$7.3e^{-3}$	0.085	17.176	1.441	0.059	0.116	0.768	0.906	1.251	0.951
	C_2H_4	0.041	0.152	0.015	0.115	$5.4e^{-3}$	0.065	0.013	0.109	0.926	0.368	0.102	0.296	0.791	0.920	1.381	1.021
	C_2H_2	0.012	0.049	$7.8e^{-3}$	0.027	799.8	27.60	$2.3e^{-3}$	0.046	0.018	0.037	$9.0e^{-9}$	$5.2e^{-5}$	0.609	0.806	1.332	0.952
	CO_2	0.080	0.157	0.042	0.202	0.113	0.291	1318	36.20	$17e^{+3}$	126.9	$67e^{+4}$	130.4	0.945	1.008	1.414	1.078
	CO	0.053	0.224	0.029	0.163	0.038	0.190	$4.0e^{-3}$	0.051	271.9	7.284	553.5	2.668	0.892	0.978	1.406	1.053
7	H_2	0.026	0.052	$3.9e^{-3}$	$8.9e^{-3}$	1.017	0.989	0.033	0.129	0.909	0.128	$1.9e^{-3}$	0.019	0.195	0.428	1.161	0.655
	C_2H_6	0.011	0.027	$2.9e^{-3}$	0.053	2.126	1.379	0.017	0.109	0.942	0.089	0.038	0.087	0.513	0.743	1.308	1.005
	CH_4	$2.1e^{-3}$	0.032	$2.5e^{-3}$	0.048	0.044	0.154	0.061	0.156	1.282	0.083	0.038	0.087	0.599	0.803	1.324	0.997
	C_2H_4	0.043	0.125	$5.7e^{-3}$	0.072	3.169	1.765	0.013	0.104	0.332	0.233	0.079	0.276	0.566	0.782	1.325	1.030
	C_2H_2	0.024	0.152	0.017	0.041	1788	39.21	0.214	0.048	0.011	0.015	$2.1e^{-6}$	$6.5e^{-4}$	0.513	0.739	1.168	1.109
	CO_2	0.018	0.075	0.017	0.054	0.530	0.635	0.059	0.068	$97e^{+3}$	48.41	5087	71.17	0.641	0.833	1.357	1.078
	CO	0.028	0.086	0.010	0.099	0.354	0.272	0.059	0.198	349.3	4.727	131.4	4.729	0.670	0.852	1.364	1.096
8	H_2	1.865	0.873	1.011	0.880	0.857	0.925	1.385	0.964	7.836	2.662	24.88	4.510	0.837	0.952	1.223	1.106
	C_2H_6	$4.1e^{-3}$	$5.8e^{-3}$	$9.7e^{-3}$	0.096	$3.3e^{-3}$	0.031	0.150	0.123	4458	7.127	6.761	0.910	0.459	0.702	1.283	1.003
	CH_4	$3.8e^{-3}$	0.011	$2.1e^{-3}$	0.018	$3.2e^{-3}$	0.022	0.018	0.073	137.5	1.186	1.593	0.982	0.419	0.671	1.268	0.983
	C_2H_4	0.012	0.085	0.056	0.119	0.144	0.234	0.093	0.068	27.96	0.874	0.150	0.348	0.583	0.791	1.301	1.029
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.120	0.322	0.133	0.273	0.239	0.370	0.134	0.284	$61e^{+3}$	221.9	$80e^{+3}$	167.1	0.506	0.740	1.285	1.035
	CO	0.048	0.191	0.051	0.185	0.058	0.200	0.046	0.215	247.4	6.598	190.0	8.185	0.495	0.732	1.289	1.063

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
9	H_2	0.073	0.031	0.139	0.100	0.103	0.185	0.330	0.527	29.40	5.136	3.374	0.292	0.382	0.596	1.308	0.653
	C_2H_6	4.6e^{-3}	0.022	9.8e^{-3}	0.051	3.350	1.823	17.71	4.176	17.44	0.805	3.474	0.708	0.224	0.428	1.270	0.432
	CH_4	0.103	0.048	0.097	0.135	0.063	0.237	0.049	0.217	406.2	20.08	22.65	1.584	0.319	0.539	1.298	0.587
	C_2H_4	2.4e^{-3}	0.023	0.023	0.068	6.8e^{-3}	0.082	0.089	0.284	689.0	6.985	341.9	7.024	0.268	0.483	1.299	0.541
	C_2H_2	9.9e^{-5}	9.7e^{-3}	4.4e^{-3}	0.015	1422	37.67	0.053	0.229	2174	3.363	10.09	1.207	0.252	0.458	1.288	0.462
	CO_2	0.232	0.357	0.198	0.377	2.131	1.350	170.05	13.03	62e^{+3}	99.94	14e^{+3}	41.42	1.493	1.267	1.184	1.099
	CO	0.314	0.268	0.193	0.286	3.977	1.356	0.994	0.961	639.1	21.97	104.2	6.309	1.450	1.246	1.140	1.012
10	H_2	7.3e^{-4}	0.016	7.0e^{-4}	0.020	7.387	2.640	1.5e^{-3}	0.035	468.7	1.312	8.274	2.869	0.734	0.883	1.354	0.949
	C_2H_6	5.0e^{-6}	1.6e^{-3}	1.5e^{-4}	0.011	0.140	0.369	1.1e^{-3}	0.031	38e^{+4}	45.73	14.10	1.258	0.889	0.974	1.410	1.029
	CH_4	1.9e^{-4}	4.4e^{-3}	5.8e^{-5}	5.7e^{-3}	6.3e^{-3}	0.061	5.0e^{-3}	0.071	13e^{+3}	8.150	1.261	0.782	0.843	0.948	1.392	1.001
	C_2H_4	1.3e^{-3}	0.027	1.6e^{-4}	7.3e^{-3}	1.2e^{-3}	7.2e^{-3}	6.1e^{-3}	0.075	35e^{+4}	42.69	2.204	1.126	0.561	0.767	1.298	0.836
	C_2H_2	0.130	0.271	5.2e^{-4}	0.014	7291	84.20	2.4e^{-3}	0.045	57e^{+3}	17.55	1.339	0.617	0.962	1.014	1.437	1.065
	CO_2	0.032	0.161	0.050	0.209	18.78	3.126	7.526	2.741	1537	37.41	4340	65.91	0.796	0.922	1.210	1.133
	CO	0.231	0.146	0.041	0.200	17.88	2.872	52.28	7.230	246.2	2.646	39.69	3.890	0.532	0.746	1.125	1.009
11	H_2	0.072	0.108	0.108	0.217	177.2	8.452	0.125	0.310	34.06	2.054	28.12	2.990	1.080	1.078	1.102	1.075
	C_2H_6	1.0e^{-4}	5.0e^{-3}	3.7e^{-4}	6.8e^{-3}	1.677	1.269	0.017	0.131	1738	2.932	0.286	0.044	0.242	0.445	1.282	0.433
	CH_4	4.4e^{-4}	0.021	0.030	0.173	0.165	0.374	0.036	0.122	84.10	1.804	0.136	0.329	0.280	0.485	1.293	0.462
	C_2H_4	0.014	0.119	7.7e^{-3}	0.088	0.590	0.524	0.048	0.167	92.83	1.375	2.052	0.753	0.271	0.478	1.296	0.506
	C_2H_2	1.711	1.292	0.657	0.805	967.9	29.57	0.346	0.575	10.29	0.740	2.522	0.796	1.073	1.073	1.114	1.094
	CO_2	0.107	0.151	0.191	0.207	2.518	1.425	0.809	0.869	3103	48.02	6035	30.88	0.463	0.705	1.180	1.074
	CO	0.544	0.412	0.655	0.442	3.239	1.388	0.532	0.510	615.5	15.41	249.0	12.88	0.675	0.851	1.220	1.138

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
12	H_2	0.184	0.221	0.022	0.148	0.444	0.661	0.401	0.109	0.107	0.325	0.942	0.013	0.521	0.748	1.235	1.121
	C_2H_6	$5.8e^{-4}$	$4.5e^{-3}$	$1.1e^{-3}$	0.032	0.464	0.666	0.021	0.134	145.7	1.373	0.383	0.458	0.540	0.763	1.327	1.024
	CH_4	$4.0e^{-5}$	$5.0e^{-3}$	$1.9e^{-4}$	0.014	0.134	0.351	0.068	0.186	82.23	2.150	0.153	0.175	0.482	0.720	1.287	0.990
	C_2H_4	0.082	0.184	0.031	0.162	0.279	0.362	2.841	1.679	0.561	0.643	1.596	0.357	0.600	0.806	1.318	1.085
	C_2H_2	0.203	0.154	0.148	0.339	383.2	19.45	0.430	0.490	0.499	0.314	0.798	0.400	0.236	0.479	1.086	0.963
	CO_2	0.134	0.317	0.041	0.140	0.311	0.445	0.064	0.207	$85e^{+3}$	287.8	$12e^{+3}$	105.0	0.895	0.984	1.423	1.164
	CO	0.041	0.107	0.011	0.079	0.079	0.274	0.011	0.067	86.32	2.265	19.84	1.373	0.821	0.943	1.418	1.143
13	H_2	1.726	0.696	2.638	0.652	10.44	2.322	2867	53.17	3.391	0.262	4.990	0.277	1.743	1.367	1.144	1.096
	C_2H_6	$8.3e^{-3}$	0.040	$3.2e^{-3}$	0.055	23.71	4.829	0.225	0.275	37.20	1.094	0.959	0.370	0.585	0.792	1.276	1.007
	CH_4	1.744	0.529	0.028	0.141	4.512	2.123	4.768	1.466	17.57	0.440	1.167	0.511	0.967	1.020	1.186	1.099
	C_2H_4	0.394	0.605	0.618	0.756	2.169	1.317	0.314	0.516	1.341	0.315	1.209	0.622	0.838	0.936	1.240	1.154
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.619	0.584	0.329	0.527	0.465	0.633	0.975	0.865	7229	82.54	$48e^{+3}$	125.7	0.674	0.825	1.151	0.893
	CO	0.949	0.771	0.528	0.666	13.06	3.610	2.117	1.271	315.6	0.000	423.2	15.82	1.435	1.230	1.162	1.072
14	H_2	0.027	0.137	0.040	0.106	1.811	1.271	8.816	2.927	1.965	0.765	1.360	0.144	0.753	0.887	1.152	1.085
	C_2H_6	1.724	1.098	0.585	0.492	0.088	0.281	27.57	4.356	0.741	0.458	3.550	0.705	0.443	0.648	1.115	0.972
	CH_4	0.044	0.167	0.022	0.128	0.012	0.064	1.181	0.917	18.13	1.737	30.32	2.379	1.367	1.221	1.285	1.183
	C_2H_4	0.572	0.635	0.630	0.448	0.233	0.335	0.415	0.644	2.839	1.652	4.028	1.173	0.778	0.903	1.191	0.968
	C_2H_2	3.983	0.896	3.678	0.523	194.2	12.51	3.380	0.949	$5.8e^{-4}$	$3.7e^{-3}$	$9.1e^{-4}$	0.014	1.095	1.093	1.117	1.076
	CO_2	0.158	0.382	0.116	0.302	1.194	0.618	0.223	0.230	$45e^{+3}$	211.2	$81e^{+3}$	67.34	1.085	1.076	1.201	1.078
	CO	0.309	0.554	0.253	0.491	1.782	0.773	0.329	0.553	384.6	18.46	403.8	12.41	1.057	1.062	1.189	1.107

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
15	H_2	0.192	0.153	0.407	0.291	1.450	0.691	2.748	1.182	48.12	5.144	8.111	1.647	0.525	0.757	1.221	1.030
	C_2H_6	0.106	0.157	0.016	0.110	$3.1e^{-3}$	0.047	0.059	0.120	1220	4.391	1.970	0.904	0.259	0.520	1.158	0.954
	CH_4	0.064	0.162	0.071	0.171	$7.6e^{-3}$	0.074	0.086	0.274	41.02	1.079	0.174	0.403	0.281	0.545	1.152	0.980
	C_2H_4	0.029	0.159	0.021	0.132	0.014	0.094	0.040	0.188	2.508	0.435	0.211	0.044	0.299	0.564	1.164	0.956
	C_2H_2	0.221	0.150	0.948	0.391	202.2	13.98	2517	50.10	0.947	0.068	1.717	0.922	1.093	1.090	1.181	1.133
	CO_2	0.088	0.244	0.113	0.153	13.23	2.131	709.1	26.62	$31e^{+3}$	21.76	9379	61.01	0.769	0.916	1.233	1.047
	CO	0.060	0.094	0.061	0.145	2.343	1.111	56.43	7.509	2670	9.722	25.26	1.286	0.393	0.653	1.173	0.991
16	H_2	0.235	0.247	0.095	0.193	0.826	0.885	0.501	0.704	17.91	2.610	5.669	2.381	0.499	0.736	1.170	1.067
	C_2H_6	0.083	0.031	0.057	0.138	0.281	0.456	0.042	0.171	647.4	3.890	1.381	0.671	0.258	0.523	1.165	0.985
	CH_4	0.099	0.024	0.050	0.106	0.271	0.444	0.234	0.465	39.39	0.927	1.034	0.607	0.320	0.586	1.159	1.024
	C_2H_4	0.062	0.151	0.100	0.252	0.277	0.524	57.80	7.420	11.61	0.806	0.417	0.398	0.384	0.644	1.187	1.013
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.376	0.451	1.089	0.531	1.361	0.746	2.903	1.633	$11e^{+4}$	126.2	$62e^{+3}$	165.3	0.660	0.846	1.227	1.087
	CO	0.082	0.260	0.155	0.335	5.824	2.413	17.74	4.000	236.7	13.91	235.5	15.33	0.524	0.753	1.175	1.061
17	H_2	0.071	0.126	0.196	0.096	0.685	0.508	0.057	0.199	30.71	0.581	0.208	0.058	0.391	0.649	1.216	1.000
	C_2H_6	0.248	0.100	0.181	0.144	1.904	1.313	0.229	0.303	413.6	2.340	0.132	0.048	0.214	0.472	1.148	0.928
	CH_4	0.268	0.058	0.193	0.085	1.179	1.055	0.624	0.715	36.57	0.679	1.009	0.974	0.296	0.560	1.149	0.968
	C_2H_4	0.481	0.576	0.276	0.336	0.982	0.919	1.021	0.873	0.530	0.142	0.021	0.056	0.311	0.578	1.166	0.953
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.542	0.702	0.963	0.913	3.058	1.631	1.304	1.139	$87e^{+3}$	256.2	$60e^{+3}$	233.2	0.531	0.751	1.155	1.103
	CO	0.190	0.102	0.092	0.283	2.310	1.520	0.162	0.132	197.8	12.71	200.8	14.11	0.426	0.678	1.177	0.987

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
18	H_2	0.155	0.230	0.047	0.103	0.381	0.571	0.105	0.321	68.46	0.000	13.11	0.559	0.223	0.444	1.100	1.021
	C_2H_6	1.321	1.114	0.391	0.573	0.189	0.433	0.948	0.774	0.733	0.773	2.698	1.025	0.371	0.586	1.121	1.047
	CH_4	0.082	0.273	0.040	0.200	0.021	0.126	73.113	5.760	9.693	2.349	13.72	2.958	1.441	1.253	1.284	1.183
	C_2H_4	0.594	0.721	0.126	0.287	0.085	0.255	2.240	1.233	7.071	2.177	8.250	2.130	0.277	0.499	1.099	0.984
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.094	0.306	0.133	0.304	0.780	0.507	261.7	16.17	1511	38.37	$66e^{+3}$	114.9	0.357	0.581	1.109	1.002
	CO	0.310	0.503	0.350	0.589	0.287	0.507	0.162	0.256	513.4	2.276	2277	24.68	0.208	0.423	1.105	1.017
All	H_2	0.425	0.652	0.278	0.521	2.829	1.675	160.8	12.67	491.7	21.08	48.17	6.939	0.626	0.790	1.213	1.016
	C_2H_6	0.534	0.707	0.217	0.420	0.978	0.968	4.182	2.026	$22e^{+3}$	141.2	37.78	6.142	0.598	0.771	1.256	1.013
	CH_4	0.186	0.426	0.053	0.207	0.176	0.417	5.437	2.263	1122	32.07	250.4	15.72	0.673	0.816	1.257	0.989
	C_2H_4	0.632	0.776	0.672	0.773	1.183	1.053	7.135	2.668	$21e^{+3}$	142.0	1040	31.88	0.564	0.750	1.252	1.018
	C_2H_2	1.451	1.204	0.510	0.702	456.2	21.35	226.1	15.04	3650	57.03	10.65	3.262	0.512	0.713	0.895	0.900
	CO_2	0.495	0.680	0.331	0.535	2.790	1.667	278.0	16.67	$87e^{+3}$	295.2	$16e^{+4}$	389.7	0.878	0.936	1.246	1.113
	CO	0.348	0.586	0.224	0.472	2.695	1.640	17.99	4.239	502.7	22.05	630.9	25.09	0.842	0.917	1.235	1.104

Table 14: Testing Results Summary - Augmented Features

Trans No.	**Feat**	**LS-SVM(1)**		**LS-SVM(2)**		**ϵ-SVR(2)**		**LS-SVM-UV(2)**		**ARIMA**		**E.S.**		**Naive**		**Mean**	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
1	H_2	0.330	0.513	0.227	0.443	6.564	2.451	0.477	0.600	0.827	0.819	0.788	0.741	0.775	0.907	1.205	1.058
	C_2H_6	0.721	0.209	0.194	0.439	187.6	13.70	0.559	0.179	0.012	$7.3e^{-3}$	0.016	$5.4e^{-3}$	1.713	1.366	1.209	1.134
	CH_4	0.161	0.108	0.118	0.137	$3.9e^{-3}$	0.028	0.061	0.116	0.071	0.058	$1.3e^{-3}$	0.023	0.499	0.658	1.342	0.508
	C_2H_4	0.916	0.232	0.089	0.265	0.168	0.410	0.091	0.266	0.036	0.029	$1.0e^{-3}$	0.019	0.992	1.040	1.179	1.139
	C_2H_2	0.673	0.177	0.204	0.215	4802	64.39	1310	36.19	0.046	0.064	$3.3e^{-4}$	0.012	0.695	0.869	1.214	1.147
	CO_2	2.417	0.289	1.876	0.485	0.710	0.753	0.921	0.516	$46e^{+4}$	73.61	$38e^{+4}$	370.4	1.514	1.290	1.269	1.181
	CO	1.713	0.682	0.847	0.565	0.394	0.409	27.27	5.187	669.2	23.43	1303	17.95	1.261	1.177	1.272	1.182
2	H_2	$4.7e^{-3}$	0.063	$4.2e^{-3}$	0.063	21.19	4.555	0.109	0.301	0.730	0.252	0.046	0.210	0.577	0.781	1.245	0.867
	C_2H_6	9.661	3.015	10.61	2.984	28.67	4.503	17.02	3.837	0.791	0.828	0.979	0.864	0.905	0.963	1.246	1.046
	CH_4	0.127	0.304	0.134	0.337	4905	69.84	6.859	2.615	4.643	1.498	2.148	1.369	1.004	1.037	1.175	1.048
	C_2H_4	47.76	6.131	46.03	5.886	43.56	5.614	45.36	6.162	95.37	8.546	103.6	8.750	0.333	0.570	1.223	0.647
	C_2H_2	234.6	13.25	232.6	13.27	1072	32.29	234.1	13.25	164.6	11.15	165.7	11.15	0.295	0.497	1.321	0.433
	CO_2	0.616	0.593	0.586	0.592	0.678	0.526	3.247	1.203	$21e^{+4}$	444.5	$49e^{+4}$	639.9	1.054	1.065	1.414	1.103
	CO	0.111	0.327	0.121	0.345	0.094	0.307	1.268	1.098	620.7	23.19	1186	31.50	0.981	1.027	1.432	1.093
3	H_2	$3.2e^{-4}$	0.017	$1.3e^{-3}$	0.013	0.048	0.219	$3.3e^{-4}$	0.017	5976	6.640	0.192	0.216	0.665	0.840	1.320	0.861
	C_2H_6	$5.5e^{-4}$	0.016	$3.1e^{-5}$	$4.8e^{-3}$	10.11	3.071	0.056	0.105	125.4	1.010	0.257	0.210	0.669	0.842	1.322	0.866
	CH_4	$7.7e^{-5}$	$8.4e^{-3}$	$7.9e^{-4}$	0.027	$5.1e^{-4}$	0.015	$4.6e^{-3}$	0.067	341.9	1.702	0.790	0.316	0.671	0.843	1.317	0.850
	C_2H_4	$1.3e^{-3}$	0.032	$3.9e^{-4}$	0.012	0.011	0.106	$3.1e^{-3}$	0.043	4710	5.997	1.179	0.552	0.668	0.841	1.321	0.865
	C_2H_2	$7.0e^{-3}$	0.056	$5.6e^{-5}$	$4.3e^{-3}$	1260	34.24	$6.6e^{-4}$	0.025	6472	6.771	0.016	0.028	0.667	0.840	1.321	0.864
	CO_2	0.307	0.446	0.368	0.451	0.713	0.843	13.37	3.294	$32e^{+4}$	303.1	$43e^{+4}$	652.6	1.156	1.119	1.199	1.122
	CO	0.277	0.511	0.290	0.485	0.589	0.762	0.596	0.768	1213	25.60	1689	41.08	1.155	1.113	1.195	1.065

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
4	H_2	21.84	1.681	1.669	0.776	6.053	1.289	11.71	2.347	862.1	24.01	764.8	23.94	0.530	0.757	1.228	1.122
	C_2H_6	21.29	0.801	4.642	1.284	10.03	1.886	10.64	2.849	1496	1.639	633.2	13.10	0.713	0.875	1.173	1.124
	CH_4	3.293	1.429	4.218	1.575	8.958	1.260	10.68	2.389	9225	20.70	4430	46.47	0.707	0.870	1.169	1.105
	C_2H_4	6.553	1.370	11.68	1.998	16.11	1.907	17.98	3.487	$37e^{+3}$	22.15	$18e^{+3}$	80.42	0.828	0.944	1.190	1.133
	C_2H_2	40.33	3.801	4.074	1.984	2.304	1.214	4.549	2.129	21.12	1.548	9.454	1.967	1.000	1.040	1.239	0.878
	CO_2	0.567	0.752	0.528	0.538	0.760	0.633	1.047	0.940	$15e^{+4}$	380.5	$32e^{+4}$	519.0	1.225	1.145	1.180	1.036
	CO	0.120	0.146	0.459	0.279	0.922	0.463	163.5	11.76	117.5	7.648	1974	8.335	1.020	1.042	1.123	0.954
5	H_2	0.036	0.117	0.011	0.038	0.072	0.268	0.341	0.314	5.333	0.767	2.188	0.999	0.415	0.667	1.205	0.953
	C_2H_6	0.158	0.392	0.120	0.043	0.043	0.129	$3.3e^{-3}$	0.051	2921	3.948	9.843	1.456	1.044	1.063	1.327	1.147
	CH_4	$1.5e^{-3}$	0.013	0.029	0.166	$8.2e^{-4}$	0.027	$4.2e^{-3}$	0.064	70.02	0.544	0.505	0.381	0.878	0.975	1.294	1.122
	C_2H_4	0.119	0.048	0.016	0.037	$2.8e^{-3}$	0.051	0.056	0.236	10.40	0.398	0.117	0.317	0.872	0.971	1.298	1.113
	C_2H_2	0.010	0.093	0.208	0.015	0.046	0.076	0.777	0.731	0.017	0.030	$3.7e^{-7}$	$3.8e^{-4}$	0.767	0.906	1.183	1.032
	CO_2	0.074	0.272	0.081	0.283	0.337	0.580	2513	49.90	$80e^{+3}$	185.1	$92e^{+3}$	151.5	1.095	1.087	1.215	1.145
	CO	0.116	0.156	0.080	0.156	0.083	0.223	0.266	0.371	440.6	7.117	351.1	8.477	1.112	1.096	1.251	1.161
6	H_2	0.070	0.030	0.108	0.030	0.030	0.061	0.719	0.846	0.168	0.000	$3.8e^{-3}$	0.040	0.319	0.572	1.205	0.798
	C_2H_6	0.035	0.096	0.050	0.052	699.1	25.51	$9.0e^{-3}$	0.016	0.896	0.089	$7.7e^{-3}$	0.042	0.812	0.933	1.399	1.047
	CH_4	0.030	0.016	0.095	0.198	1465	38.25	$7.3e^{-3}$	0.085	17.18	1.441	0.059	0.116	0.768	0.906	1.251	0.951
	C_2H_4	0.335	0.176	0.014	0.113	0.058	0.151	0.013	0.109	0.926	0.368	0.102	0.296	0.791	0.920	1.381	1.021
	C_2H_2	0.439	0.600	0.026	0.034	2940	54.21	$2.3e^{-3}$	0.046	0.018	0.037	$9.0e^{-9}$	$5.2e^{-5}$	0.609	0.806	1.332	0.952
	CO_2	0.018	0.122	0.080	0.086	0.242	0.491	1318	36.20	$17e^{+3}$	126.9	$67e^{+4}$	130.4	0.945	1.008	1.414	1.078
	CO	0.079	0.118	0.014	0.082	0.268	0.516	$4.0e^{-3}$	0.051	271.9	7.284	553.5	2.668	0.892	0.978	1.406	1.053

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
7	H_2	0.015	0.069	5.8e^{-3}	0.018	2.910	1.689	0.033	0.129	0.909	0.128	1.9e^{-3}	0.019	0.195	0.428	1.161	0.655
	C_2H_6	0.063	0.230	0.024	0.018	36.24	5.682	0.017	0.109	0.942	0.089	0.038	0.087	0.513	0.743	1.308	1.005
	CH_4	2.6e^{-3}	0.023	2.8e^{-3}	0.034	171.8	12.51	0.061	0.156	1.282	0.083	0.038	0.087	0.599	0.803	1.324	0.997
	C_2H_4	0.069	0.135	0.026	0.116	0.314	0.494	0.013	0.104	0.332	0.233	0.079	0.276	0.566	0.782	1.325	1.030
	C_2H_2	0.343	0.361	5.7e^{-3}	0.047	1978	43.53	0.214	0.048	0.011	0.015	2.1e^{-6}	6.5e^{-4}	0.513	0.739	1.168	1.109
	CO_2	2.3e^{-3}	0.047	7.2e^{-3}	0.026	0.084	0.157	0.059	0.068	97e^{+3}	48.41	5087	71.17	0.641	0.833	1.357	1.078
	CO	0.082	0.036	0.015	0.074	2.523	1.526	0.059	0.198	349.3	4.727	131.4	4.729	0.670	0.852	1.364	1.096
8	H_2	1.451	0.943	0.965	0.920	0.537	0.685	1.385	0.964	7.836	2.662	24.88	4.510	0.837	0.952	1.223	1.106
	C_2H_6	0.117	0.325	0.020	0.038	0.012	0.076	0.150	0.123	4458	7.127	6.761	0.910	0.459	0.702	1.283	1.003
	CH_4	6.7e^{-3}	0.013	0.015	0.076	3.7e^{-3}	0.040	0.018	0.073	137.5	1.186	1.593	0.982	0.419	0.671	1.268	0.983
	C_2H_4	0.241	0.128	3.3e^{-3}	0.050	0.025	0.081	0.093	0.068	27.96	0.874	0.150	0.348	0.583	0.791	1.301	1.029
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.122	0.298	0.114	0.291	0.064	0.201	0.134	0.284	61e^{+3}	221.9	80e^{+3}	167.1	0.506	0.740	1.285	1.035
	CO	0.039	0.171	0.029	0.169	0.029	0.167	0.046	0.215	247.4	6.598	190.0	8.185	0.495	0.732	1.289	1.063
9	H_2	0.098	0.020	0.262	0.067	0.511	0.115	0.330	0.527	29.40	5.136	3.374	0.292	0.382	0.596	1.308	0.653
	C_2H_6	5.6e^{-3}	0.059	4.1e^{-4}	0.020	71.36	8.055	17.71	4.176	17.44	0.805	3.474	0.708	0.224	0.428	1.270	0.432
	CH_4	0.085	0.070	0.121	0.106	0.165	0.136	0.049	0.217	406.2	20.08	22.65	1.584	0.319	0.539	1.298	0.587
	C_2H_4	3.0e^{-3}	0.055	4.4e^{-3}	0.026	2.6e^{-3}	0.033	0.089	0.284	689.0	6.985	341.9	7.024	0.268	0.483	1.299	0.541
	C_2H_2	0.011	0.020	0.011	0.027	93.73	9.626	0.053	0.229	2174	3.363	10.09	1.207	0.252	0.458	1.288	0.462
	CO_2	1.807	0.494	0.791	0.358	8.574	1.007	170.1	13.03	62e^{+3}	99.94	14e^{+3}	41.42	1.493	1.267	1.184	1.099
	CO	0.738	0.331	0.853	0.334	1.969	0.956	0.994	0.961	639.1	21.97	104.2	6.309	1.450	1.246	1.140	1.012

Trans No.	**Feat**	**LS-SVM(1)**		**LS-SVM(2)**		**ϵ-SVR(2)**		**LS-SVM-UV(2)**		**ARIMA**		**E.S.**		**Naive**		**Mean**	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
10	H_2	$4.0e^{-4}$	0.020	$5.5e^{-4}$	0.021	$8.1e^{-3}$	0.062	$1.5e^{-3}$	0.035	468.7	1.312	8.274	2.869	0.734	0.883	1.354	0.949
	C_2H_6	$6.0e^{-3}$	0.077	$1.9e^{-4}$	0.010	25.25	5.020	$1.1e^{-3}$	0.031	$38e^{+5}$	45.73	14.10	1.258	0.889	0.974	1.410	1.029
	CH_4	$4.1e^{-5}$	$3.5e^{-3}$	$2.0e^{-4}$	$8.9e^{-3}$	186.7	13.18	$5.0e^{-3}$	0.071	$13e^{+3}$	8.150	1.261	0.782	0.843	0.948	1.392	1.001
	C_2H_4	$2.0e^{-5}$	$4.3e^{-3}$	$6.9e^{-5}$	$7.8e^{-3}$	0.024	0.149	$6.1e^{-3}$	0.075	$35e^{+4}$	42.69	2.204	1.126	0.561	0.767	1.298	0.836
	C_2H_2	$2.3e^{-3}$	0.046	$2.4e^{-4}$	0.011	3479	58.74	$2.4e^{-3}$	0.045	$57e^{+4}$	17.55	1.339	0.617	0.962	1.014	1.437	1.065
	CO_2	$8.9e^{-3}$	0.082	$4.7e^{-3}$	0.068	0.521	0.720	7.526	2.741	1537	37.41	4360	65.91	0.796	0.922	1.210	1.133
	CO	0.752	0.364	0.043	0.206	0.461	0.671	52.28	7.230	246.2	2.646	39.69	3.890	0.532	0.746	1.125	1.009
11	H_2	0.066	0.068	0.113	0.245	0.780	0.569	0.125	0.310	34.06	2.054	28.12	2.990	1.080	1.078	1.102	1.075
	C_2H_6	$2.3e^{-3}$	0.047	$2.4e^{-4}$	$7.6e^{-3}$	0.524	0.720	0.017	0.131	1738	2.932	0.286	0.044	0.242	0.445	1.282	0.433
	CH_4	$3.8e^{-4}$	0.018	$1.2e^{-3}$	0.033	$8.6e^{-3}$	0.070	0.036	0.122	84.10	1.804	0.136	0.329	0.280	0.485	1.293	0.462
	C_2H_4	$4.8e^{-3}$	0.014	$3.9e^{-3}$	0.043	0.141	0.336	0.048	0.167	92.83	1.375	2.052	0.753	0.271	0.478	1.296	0.506
	C_2H_2	0.633	0.685	0.829	0.521	59.41	6.889	0.346	0.575	10.29	0.740	2.522	0.796	1.073	1.073	1.114	1.094
	CO_2	0.022	0.058	0.059	0.044	0.078	0.263	0.809	0.869	3103	48.02	6035	30.88	0.463	0.705	1.180	1.074
	CO	0.577	0.419	0.634	0.396	0.732	0.416	0.532	0.510	615.5	15.41	249.0	12.88	0.675	0.851	1.220	1.138
12	H_2	0.104	0.320	0.032	0.174	0.053	0.211	0.401	0.109	0.107	0.325	0.942	0.013	0.521	0.748	1.235	1.121
	C_2H_6	$1.0e^{-3}$	$6.3e^{-3}$	$5.5e^{-3}$	0.025	43.25	6.530	0.021	0.134	145.7	1.373	0.383	0.458	0.540	0.763	1.327	1.024
	CH_4	$6.7e^{-4}$	$3.3e^{-3}$	$1.6e^{-3}$	0.039	344.8	18.05	0.068	0.186	82.23	2.150	0.153	0.175	0.482	0.720	1.287	0.990
	C_2H_4	0.078	0.072	0.010	0.097	0.053	0.186	2.841	1.679	0.561	0.643	1.596	0.357	0.600	0.806	1.318	1.085
	C_2H_2	0.410	0.517	0.077	0.202	1725	40.92	0.430	0.490	0.499	0.314	0.798	0.400	0.236	0.479	1.086	0.963
	CO_2	$8.1e^{-3}$	0.090	0.042	0.100	0.040	0.198	0.064	0.207	$84e^{+3}$	287.8	$12e^{+3}$	105.0	0.895	0.984	1.423	1.164
	CO	0.028	$7.0e^{-3}$	$1.2e^{-3}$	$9.6e^{-3}$	$6.9e^{-3}$	0.077	0.011	0.067	86.32	2.265	19.84	1.373	0.821	0.943	1.418	1.143

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
13	H_2	3.686	0.332	6.440	0.457	27.33	1.327	2867	53.17	3.391	0.262	4.990	0.277	1.743	1.367	1.144	1.096
	C_2H_6	0.013	0.074	0.055	0.023	3053	54.74	0.225	0.275	37.20	1.094	0.959	0.370	0.585	0.792	1.276	1.007
	CH_4	1.446	0.472	0.558	0.650	5097	68.90	4.768	1.466	17.57	0.440	1.167	0.511	0.967	1.020	1.186	1.099
	C_2H_4	0.821	0.836	0.424	0.651	0.442	0.555	0.314	0.516	1.341	0.315	1.209	0.622	0.838	0.936	1.240	1.154
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.698	0.419	0.802	0.337	0.170	0.404	0.975	0.865	7229	82.54	$47e^{+3}$	125.7	0.674	0.825	1.151	0.893
	CO	1.079	0.612	1.667	0.530	4.454	0.975	2.117	1.271	315.6	0.000	423.2	15.82	1.435	1.230	1.162	1.072
14	H_2	0.231	0.337	0.063	0.112	0.786	0.717	8.816	2.927	1.965	0.765	1.360	0.144	0.753	0.887	1.152	1.085
	C_2H_6	2.341	1.440	0.930	0.595	0.551	0.405	27.57	4.356	0.741	0.458	3.550	0.705	0.443	0.648	1.115	0.972
	CH_4	0.042	0.170	0.025	0.151	0.013	0.079	1.181	0.917	18.13	1.737	30.32	2.379	1.367	1.221	1.285	1.183
	C_2H_4	0.970	0.923	0.593	0.347	0.800	0.312	0.415	0.644	2.839	1.652	4.028	1.173	0.778	0.903	1.191	0.968
	C_2H_2	6.617	1.303	3.993	0.625	507.6	21.94	3.380	0.949	$5.8e^{-4}$	$3.7e^{-3}$	$9.1e^{-4}$	0.014	1.095	1.093	1.117	1.076
	CO_2	0.141	0.374	0.169	0.316	1.016	0.534	0.223	0.230	$45e^{+3}$	211.2	$81e^{+3}$	67.34	1.085	1.076	1.201	1.078
	CO	0.301	0.536	0.248	0.470	0.132	0.346	0.329	0.553	384.6	18.46	403.8	12.41	1.057	1.062	1.189	1.107
15	H_2	0.102	0.116	0.174	0.244	0.758	0.469	2.748	1.182	48.12	5.144	8.111	1.647	0.525	0.757	1.221	1.030
	C_2H_6	0.340	0.502	0.097	0.152	70.68	8.057	0.059	0.120	1220	4.391	1.970	0.904	0.259	0.520	1.158	0.954
	CH_4	$5.4e^{-3}$	0.050	$8.1e^{-3}$	0.084	$5.7e^{-4}$	0.023	0.086	0.274	41.02	1.079	0.174	0.403	0.281	0.545	1.152	0.980
	C_2H_4	0.012	0.063	0.048	0.155	0.075	0.195	0.040	0.188	2.508	0.435	0.211	0.044	0.299	0.564	1.164	0.956
	C_2H_2	20.81	4.286	$9.9e^{-3}$	0.077	709.0	26.41	2517	50.05	0.947	0.068	1.717	0.922	1.093	1.090	1.181	1.133
	CO_2	0.017	0.131	0.011	0.086	7.778	1.806	709.1	26.62	$31e^{+3}$	21.76	9379	61.01	0.769	0.916	1.233	1.047
	CO	0.019	0.134	$4.2e^{-3}$	0.065	0.049	0.190	56.43	7.509	2670	9.722	25.26	1.286	0.393	0.653	1.173	0.991

Trans No.	Feat	LS-SVM(1)		LS-SVM(2)		ϵ-SVR(2)		LS-SVM-UV(2)		ARIMA		E.S.		Naive		Mean	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
16	H_2	0.034	0.180	0.022	0.145	0.972	0.895	0.501	0.704	17.91	2.610	5.669	2.381	0.499	0.736	1.170	1.067
	C_2H_6	$1.2e^{-3}$	0.023	0.028	0.147	403.9	17.81	0.042	0.171	647.4	3.890	1.381	0.671	0.258	0.523	1.165	0.985
	CH_4	$2.1e^{-3}$	0.046	0.028	0.144	62.14	7.774	0.234	0.465	39.39	0.927	1.034	0.607	0.320	0.586	1.159	1.024
	C_2H_4	0.016	0.047	0.141	0.133	$6.0e^{-3}$	0.050	57.80	7.420	11.61	0.806	0.417	0.398	0.384	0.644	1.187	1.013
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.322	0.352	0.322	0.344	4.531	0.921	2.903	1.633	$11e^{+4}$	126.2	$62e^{+3}$	165.3	0.660	0.846	1.227	1.087
	CO	0.031	0.168	0.029	0.154	0.291	0.484	17.74	4.000	236.7	13.91	235.5	15.33	0.524	0.753	1.175	1.061
17	H_2	0.042	0.084	0.035	0.060	0.058	0.227	0.057	0.199	30.71	0.581	0.208	0.058	0.391	0.649	1.216	1.000
	C_2H_6	0.024	0.057	0.044	0.197	199.9	12.93	0.229	0.303	413.6	2.340	0.132	0.048	0.214	0.472	1.148	0.928
	CH_4	0.025	0.016	0.020	0.079	1.754	1.287	0.624	0.715	36.57	0.679	1.009	0.974	0.296	0.560	1.149	0.968
	C_2H_4	0.084	$7.9e^{-3}$	0.064	0.253	0.143	0.377	1.021	0.873	0.530	0.142	0.021	0.056	0.311	0.578	1.166	0.953
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.372	0.338	0.359	0.414	1.391	0.613	1.304	1.139	$87e^{+3}$	256.2	$60e^{+3}$	233.2	0.531	0.751	1.155	1.103
	CO	0.056	0.107	0.034	0.167	0.107	0.281	0.162	0.132	197.8	12.71	200.8	14.11	0.426	0.678	1.177	0.987
18	H_2	0.083	0.247	0.035	0.072	0.549	0.730	0.105	0.321	68.46	0.000	13.11	0.559	0.223	0.444	1.100	1.021
	C_2H_6	1.647	1.273	0.558	0.690	0.303	0.549	0.948	0.774	0.733	0.773	2.698	1.025	0.371	0.586	1.121	1.047
	CH_4	0.088	0.285	0.117	0.317	0.025	0.116	73.11	5.760	9.693	2.349	13.72	2.958	1.441	1.253	1.284	1.183
	C_2H_4	0.524	0.713	0.082	0.236	0.134	0.324	2.240	1.233	7.071	2.177	8.250	2.130	0.277	0.499	1.099	0.984
	C_2H_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CO_2	0.135	0.345	0.080	0.283	0.159	0.305	261.7	16.17	1511	38.37	$66e^{+3}$	114.9	0.357	0.581	1.109	1.002
	CO	0.331	0.570	0.270	0.440	0.277	0.348	0.162	0.256	513.4	2.276	2277	24.68	0.208	0.423	1.105	1.017

Trans No.	**Feat**	**LS-SVM(1)**		**LS-SVM(2)**		**ϵ-SVR(2)**		**LS-SVM-UV(2)**		**ARIMA**		**E.S.**		**Naive**		**Mean**	
		μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ	μ_{MSE}	σ
All	H_2	0.638	0.785	0.336	0.574	1.199	1.080	160.8	12.67	491.7	21.08	48.17	6.939	0.626	0.790	1.213	1.016
	C_2H_6	0.579	0.750	0.290	0.520	43.12	6.556	4.182	2.026	$22e^{+3}$	141.2	37.78	6.142	0.598	0.771	1.256	1.013
	CH_4	0.094	0.302	0.058	0.232	78.04	8.808	5.437	2.263	1122	32.07	250.4	15.72	0.673	0.816	1.257	0.989
	C_2H_4	0.614	0.775	0.748	0.837	0.890	0.917	7.135	2.668	$21e^{+3}$	142.0	1040	31.88	0.564	0.750	1.252	1.018
	C_2H_2	2.630	1.622	0.533	0.706	335.2	18.31	226.1	15.04	3650	57.03	10.65	3.262	0.512	0.713	0.895	0.900
	CO_2	0.377	0.604	0.349	0.584	1.111	0.981	278.0	16.67	$87e^{+3}$	295.2	$16e^{+4}$	389.7	0.878	0.936	1.246	1.113
	CO	0.313	0.559	0.263	0.513	0.414	0.626	17.99	4.239	502.7	22.05	630.9	25.09	0.842	0.917	1.235	1.104