VDMA
Fluid Power Association

22nd ISC

International Sealing Conference

Stuttgart, Germany
October 01 - 02, 2024

Sealing Technology –
Challenges accepted!

IMA
UNI STUTTGART

ISBN 978-3-8163-0768-6

Fachverband Fluidtechnik im VDMA e. V. Lyoner Str. 18
50628 Frankfurt am Main
Germany

Phone +49 69 6603-1513
E-Mail maximilian.baxmann@vdma.org
Internet www.vdma.org/fluid

Predicting Compatibility of Sealing Material with Bio-Hybrid Fuels: Development and Comparison of Machine Learning Methods

Lukas Boden, Marius Hofmeister, Faras Brumand-Poor, Linda Pleninger, Katharina Schmitz

Bio-hybrid fuels, derived from sustainable raw materials and green energies, offer a promising alternative to conventional fuels made of mineral oil. Within the cluster of excellence “The Fuel Science Center (FSC)” at RWTH Aachen bio-hybrid fuels are investigated on a holistic level, including an examination of their compatibility with sealing materials. Previous time-consuming experiments revealed that many bio-hybrid fuels show poor material compatibility with elastomer sealing materials (e.g. NBR & FKM) leading to issues such as volume expansion, hardness alteration, or chemical reactions upon immersion. These incompatibilities could result in catastrophic failures during practical applications. Due to the high number of possible fuels, fuel-mixtures and therefore fuel/seal combinations a solely empirical approach is impractical. Consequently, this work presents a hands-on framework of machine learning (ML) regression models that predicts material compatibility based on experimental results and selected fluid properties, thereby enabling the preselection of possible fuel/seal combinations for subsequent immersion tests.

1 Motivation and Introduction

In terms of global climate change, CO_2-neutral forms of energy from sources other than fossil feedstock are of immense importance to all of society. One possibility for providing climate-neutral energy for automotive applications is the use of bio-hybrid fuels, which are produced either based on biological resources or with the use of other carbon sources and renewable energy. In order to select fuels from the almost infinite number of possible molecules that meet environmental, economical, and technical requirements, the Cluster of Excellence "The Fuel Science Center (FSC)" is developing methods to identify optimal fuel candidates.

One advantage of bio-hybrid fuels over other alternative energy sources, such as fuel cells or electrical energy, is the use of existing infrastructure. If bio-hybrid fuels are used as so-called "drop-in fuels" in existing combustion engines, a worldwide network of filling stations and production facilities in the fuel industry can be utilized.

A mandatory prerequisite for the unrestricted use of “drop-in fuels” developed in the FSC is material compatibility with all system components. Static and dynamic seals are particularly at risk. Failure of seals can have far-reaching consequences as has been shown in the past with test bench failures caused by damaged seals. Previous studies and immersion tests have shown that many bio-hybrid fuels have poor material compatibility with conventional elastomer sealing materials, leading to an increase in elastomer volume of more than 200 % [1, 2]. Such large amounts of swelling lead to immediate failure in real technical applications. In addition to swelling,

other wear mechanisms such as changes in hardness or chemical reactions were observed [1, 2]. Based on these findings, material compatibility has become a important design parameter for bio-hybrid fuels within the FSC and therefore is addressed in this study.

Given the vast array of fuel candidates and blends, along with corresponding fuel/seal combinations, conducting manual immersion tests becomes impractical. In addressing this challenge, an automated test rig was developed, facilitating the simultaneous execution of 90 immersion tests. However, despite this advancement, achieving a comprehensive examination of all potential fuel/seal combinations remains unfeasible. Thus, there arises the necessity for a strategic selection of combinations to investigate further.

Machine learning (ML) has proven to be successful in the area of fluid power applications, like fault detection [3] and data-based condition monitoring [4]. Furthermore, several studies have been conducted on the discovery of new materials and prediction of fuel properties with ML algorithms [5, 6]. The ability of ML algorithms to discover patterns and trends in a vast amount of data represents a powerful tool for the identification of material compatibility and the prediction of compatibility of untested combinations, resulting in a narrowed-down fuel/seal configuration. Hereby, a pairing is valid if the hardness change and sealing material volume after the immersion test are close to the values of immersion tests with conventional fuels, such as diesel or gasoline. The configurations are not further considered if the values deviate too much.

In the following study, the process of conducting immersion tests, selecting suitable parameters, and preprocessing the resulting dataset is presented. Furthermore, the setup of the ML framework, model training, and evaluation and subsequently the comparison of different regression models are shown.

2 Methodology

In the following section the investigated bio-hybrid fuels and regression methods are introduced.

2.1 Investigated bio hybrid fuels

Different promising drop-in fuel candidates have emerged out of the FSC. Six of them are included in this work. Additionally, diesel in accordance with DIN EN 590 and gasoline according to DIN EN 228 are used as reference fuels [7, 8]. In the following, the composition of the bio-hybrid fuels is shortly introduced.

The first representative for bio-hybrid diesel fuels is OME 3-5. OME 3-5 is a mixture of polyoxymethylene dimethyl ethers with chain lengths between 3 and 5 carbon atoms. It is intended to be produced based on methanol and formaldehyde by using low carbon electricity [9]. The two HyFit fuels can be produced using Fischer Tropsch synthesis and consist of different shares of n-alcohols and n-alkanes. HyFit1 and HyFit2 differ primarily in their accumulated alcohol content of 15 and 40 %

respectively [10]. The third bio-hybrid diesel fuel is a mixture of different methyl ketones which is derived from a microbiological process [2]. Representatives of bio-hybrid Otto fuels are KEAA and EBCC blends which are mixtures of different ketones, esters, alcohols, and alkanes [11].

Most of the fuel candidates mentioned above are mixtures of pure liquids. Some pure liquids have already been tested by the FSC for their material compatibility and are therefore included in this study [2, 12]. This has two advantages. Firstly, the experimental database is expanded, making predictions by ML regression models more significant. Secondly, the foundation is provided to further investigate the material compatibility of mixtures or fuel blends and to understand the underlying patterns.

2.2 Regression models

The field of machine learning focuses on finding patterns in data using computer algorithms. Its main purpose is to utilize these patterns to perform various tasks, such as data classification or estimation of continuous values based on the given data [13]. Since the aim of this study is to predict numerical values based on a dataset that consists of both input and their corresponding target (output) variables, a supervised learning (SL) approach is selected. In the case of SL, the output is the true value, called a label, and the model should be tuned in such a way that it predicts the particular label for the provided input set. The inputs of the ML models are the fuel properties, which are listed in Table 1. Their selection process is elaborated further in section 3.2. The provided labels are the volume amount and hardness change of the elastomer specimen. This setup represents a multi-output regression problem, which can be solved by a variety of SL algorithms, listed in Table 2. This work aims to implement a framework, which integrates these algorithms and could be extended and validated with currently available and further collected data.

Table 1: Overview of selected describing parameters

Descriptor:	**Source:**	**Reference:**	**Data type and unit:**
Density	Measured and from literature	[23, 24]	Numerical value $\mathrm{kg/m^3}$
Molar mass	From literature	[23, 24]	Numerical value $\mathrm{g/mol}$
Molar volume	Engineered from density and molar mass	-	Numerical value $\mathrm{m^3/mol}$
Boiling point	From literature	[23, 24]	Numerical value $\mathrm{°C}$
Enthalpy of vaporization	From literature (at standard conditions)	[23, 24]	Numerical value $\mathrm{kJ/mol}$

Vapor pressure	From literature (at 20°C)	[23, 24]	Numerical value kPa
Oxygen content	Calculated from molar structure	[23, 24]	Numerical value %
Ring structure	From molecular structure	[23, 24]	Either 1 or 0 -
logP_{ow}	From literature	[25]	Numerical value -

Table 2: Comparison of the investigated regression models [14]

Regression model:		**Working principle:**
Linear Model	Linear Regression (polynomial Regression)	Fits a linear (polynomial) function to the data by minimizing the difference between the true and predicted values, adjusting coefficients to find the best-fitting line (or curve)
	Lasso Regression	Additional to linear regression, the absolute size of coefficients is penalized using L1 regularization
	Stochastic gradient decent	Minimizes the loss function iteratively by randomly selecting mini batches of data and updating model parameters in the direction opposite to the computed gradient
Nonlinear Models	Multi-layer Perceptron	Data is propagated through (multiple) layers of neurons, with each layer performing transformations on the input, ultimately predicting the output by minimizing the loss function
	SVR	Aims to find the hyperplane that best fits the data by maximizing the margin between the data points and the hyperplane and minimizing the loss function
	Decision Tree Regressor	Constructs a tree-like structure where data is split into feature subsets, optimizing splits to minimize the variance of predicted values within each subset, resulting in a piecewise constant approximation of the output variable

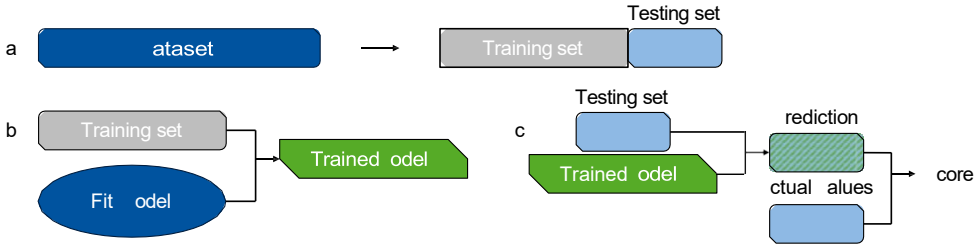

Figure 1: General process of regression tasks

The general process for solving regression tasks consists of three steps that are denoted from a) to c) in Figure 1. The first step is splitting up the dataset (a). In a second step, the model is trained (b) and lastly, the performance of the model is evaluated (c).

Splitting the dataset into a training and a test subset ensures that the subsequent training of the model is not performed on the entire dataset. The withheld test dataset is later used to evaluate the performance of the model. This also allows to assess the model's ability to generalize, which is defined as the ability to perform well on previously unseen data.

Subsequently only the training subset is considered for the fitting of the model. Based on the training input and output values, underlying patterns in the relationship between input and output are learned by the model. During this process different parameters of the models, such as weights and biases, are adjusted iteratively. If the model describes the training data well enough or if a maximum number of iterations is reached, the training process is finished, and the model can be applied to the unseen testing subset to evaluate its performance.

Based only on the testing input data, the trained model applies the learned patterns and predicts new output data. To evaluate the accuracy of this predictions, the actual testing output data are compared to the predicted values. A score is calculated using metrics, which allows the model's performance to be compared. Two commonly used metrics are introduced in Equation (2) and (3). For calculating the mean squared error (MSE), first the difference between an actual output value Y and the predicted output value Y_{pred} is computed. The difference is then squared and averaged over all n samples. A better performance is represented by a lower MSE. The coefficient of determination R^2 provides a measure of how well the regression model fits the actual data. It quantifies the proportion of the variance in the predicted values to a scale from 0 to 1, with larger values representing a better model fit.

The main benefit of observing training and testing performance is the detection of overfitting and underfitting. The former is a common weakness of SL models, where parameters are customized too specifically to a particular problem rather than learning the underlying patterns. The latter occurs when a model is insufficiently trained, often due to a lack of quality or quantity of data.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_{pred,i})^2 \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - Y_{pred,i})^2}{\sum_{i=1}^{n}(Y_i - Y_{mean})^2} = 1 - \frac{\mathrm{MSE}}{\mathrm{var(Y)}} \quad (3)$$

3 Structure and Approach

In the following section, the underlying experimental investigation as well as the preselection process of suitable parameters and machine learning methods are presented.

3.1 Immersion tests

The application of machine learning methods for predicting material compatibility in this study is founded upon experimental immersion tests conducted in accordance with ISO 1817 [15]. Standard reference elastomers (SRE) according to DIN ISO 13226 (SRE-NBR 28/SX and SRE-FKM/2X) were used as sealing materials and immersed into the fuels [16]. The size of the specimen is 25x25x2 mm³. It was ensured that the specimen was always covered with sufficient fuel in the vessels during investigation.

The specimens are removed from the vessels filled with the investigated fuel after defined intervals to evaluate the change of seal properties. The time intervals for evaluation were chosen in accordance with standard DIN 53521 [17]. For this study, only the measurements of the specimen volume and hardness at the final time interval after 672 h in relation to its initial state are considered. In the work of Hofmeister et al. the measurement techniques and devices are presented in detail. It was especially paid attention to the fact, that most investigated liquids have a high vapor pressure at standard conditions and therefore significant evaporation rates are present. [1]

In the following, the focus solely is on SRE-NBR 28/SX, since in comparison to FKM a higher number of fluids are investigated. The composition of SRE-NBR 28/SX is shown in Table 3. Figure 2 lists the change of volume and hardness of SRE-NBR for all investigated fuels and fluids.

Table 3: Composition of SRE-NBR 28/SX according to ISO 13226

Formulation of the rubber mixture	**Parts by mass**
NBR with 28 % by mass of acrylonitrile (Perbunan NT 2845)	100,0
Antidegradant TMQ (Vulkanox HS)	2,0
Zinc oxide (Zinkoxyd aktiv)	5,0
Stearic acid	1,0

Carbon black N 550 (Corax N 550)	65,0
Accelerator TBzTD	2,5
Accelerator CBS (Vulkacit CZ)	1,5
Sulphur	0,2

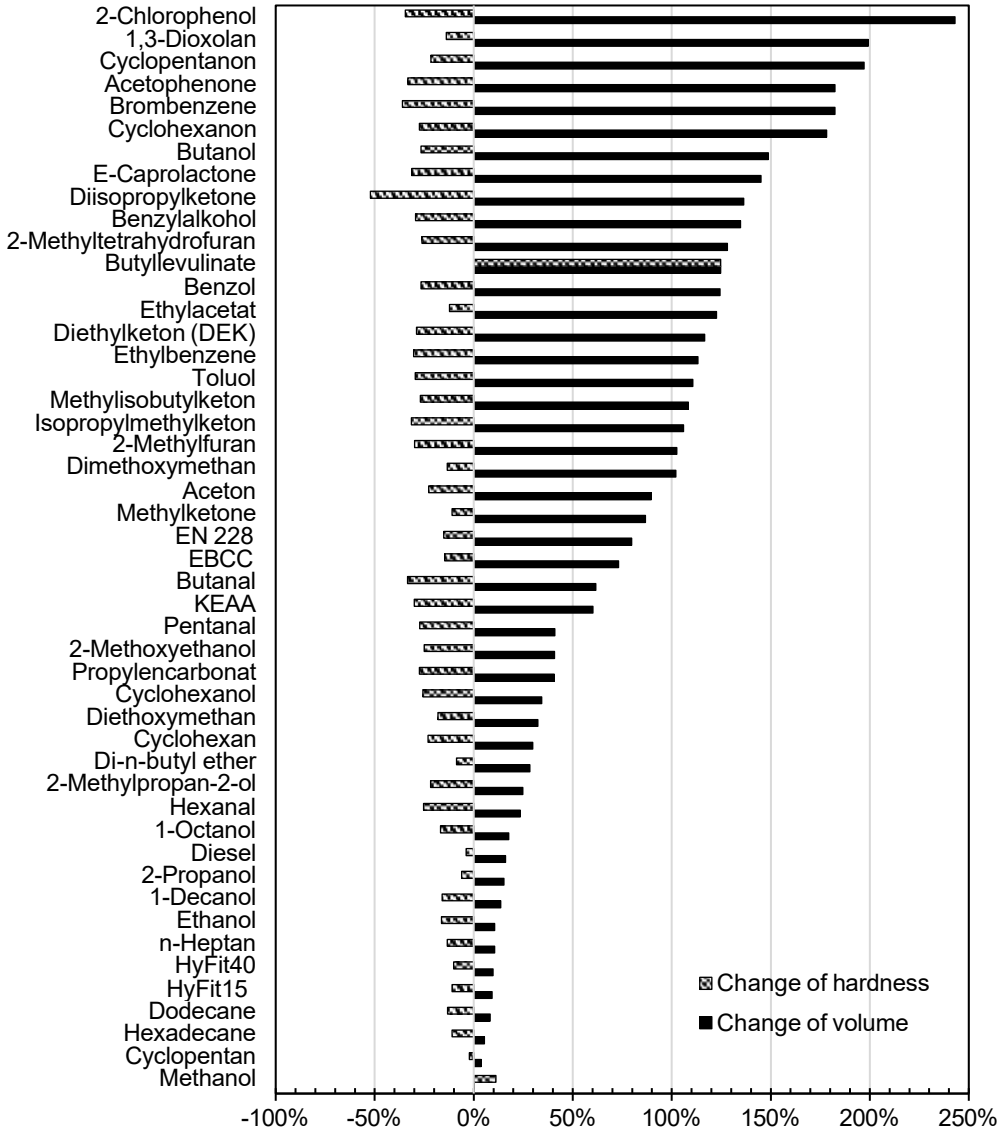

Figure 2: Results of immersion tests for SRE-NBR 28/SX

3.2 Choosing relevant parameters

The tendency of a fluid to cause polymer swelling depends largely on two factors: the polarity of the fluid and the elastomer as well as the size of the fluid molecules. According to the chemical principle “like dissolves like”, polar fluid molecules prefer to penetrate polar elastomers while nonpolar fluid molecules lead to an increased swelling of nonpolar elastomers. In terms of the steric requirement of the fluid molecules, larger molecules are more hindered to penetrate the elastomer crosslinked network. Besides fluid properties, properties of the elastomer such as the degree of crosslinking, the bulk modulus or the type and amount of additives play an important role when investigation property changes of the elastomer. Since standard reference elastomers are considered in this work, the influence of these parameters is not considered. Elastomer properties might be included as parameters in further investigations regarding technical seals.

The quantification of the polarity of molecules is the subject of numerous research projects. Kier suggests the use of a structure-based numerical index to quantify polarity [18]. Palomar et al. introduce a quantum chemical parameter for this application [19]. Furthermore, the polarity of molecules is considered in the Hansen solubility parameters to predict compatibility between different compounds [20]. Having said that, those theoretical descriptors are not commonly available and can be complex to determine especially for compound mixtures. Additionally, the weaknesses of such parameters have already been demonstrated in various studies. For example, the work of Heitzig et al. concludes that the Hansen parameters are not suitable to predict polymer swelling. One reason for this is the copolymer structure of widely used sealing materials like NBR, whose solubility would presumably have to be quantified using more than one Hansen sphere. Additionally chemical reactions, that might occur during immersion, are not considered in the theory of Hansen. [12]

For this reason, the approach presented here uses familiar physical and chemical parameters that can be found in common databases and handbooks for a wide variety of compounds. The aim was to find parameters that are both easily accessible as well as a good representation of the polarity and/or the steric requirement of the fluid-molecule. Eventually, a set of nine parameters was defined, which are presented in Table 1. The values of these parameters were taken from several open access databases and the own database of the FSC. For reasons of complexity, the focus is initially placed on the properties of the fluids while the polarity and properties of the elastomer is not taken into consideration.

The molar volume of a compound, being the ratio of the molar mass and the density, is an indicator of the steric requirement of the fluid molecule. The reason for that becomes apparent when comparing molar volumes of different C_6 organic compounds. Comparing the molar volumes of cyclohexane to hexane demonstrates the lower steric requirement of cyclic compounds compared to linear and branched molecules. Likewise, the comparison of cyclohexane to benzene shows the lower steric requirements of aromatic compounds due to their inherent planarity. This is important since compounds with a lower steric requirement show lower activation

energies for diffusion and thus show a higher tendency to penetrate the elastomer network [21].

The boiling point, the molar enthalpy of vaporization, and the vapor pressure represent the magnitude of the occurring intermolecular interactions [22]. This is closely related to the polarity of the molecule due to the ability of polar substances to form stronger intermolecular interactions such as hydrogen bonds and π-π interactions in comparison to nonpolar substances.

Oxygen content can be easily determined using molecular formula and can be an indicator for bond polarization due to the strong electronegativity of oxygen. Whether the compound is cyclic derives from the molecular structure and indicates a lower steric hindrance compared to the acyclic analogue.

The decadic logarithm of the partition coefficient in octanol/water logP_{OW} is a measure of the hydrophobicity of a substance. Since hydrophilicity correlates to a certain extent with the polarity of a substance, this parameter also provides an indication of polarity.

3.3 Application of Regression Models

The general process of application is stated above in Section 2.2. In this section, the focus is set on preprocessing the data and the framework in which the different ML models are set up, tuned, and subsequently evaluated. As mentioned before the aim of this work is to implement a framework with all necessary modules for further validation land investigation.

The framework consists of several components in order to tackle common drawbacks in data-driven algorithms. Starting with the feature scaling of the input data. This results in an equalization of the features in a common band, resolving the issue of different units and magnitudes in the input data. The scaling can be done by applying min-max normalization, converting all input data to the range 0 to 1. Another method for scaling input data is by using a standard scaler, which normalizes each feature by subtracting its mean and dividing by its standard deviation. This transformation ensures that each feature has a mean of 0 and a standard deviation of 1. [14]

In a first step, a general baseline for the prediction performance is obtained by considering all nine features introduced in the previous section. One main enhancement to ML algorithms is the reduction of the input vector to the most crucial features, resulting in faster computation and less data requirement. This is achieved by comparing correlation scores between features and only considering uncorrelated features for further use (features with a correlation score lower than a set limit). Principal component analysis (PCA) further facilitates this process by transforming the original features into a new set of orthogonal (uncorrelated) variables, thereby enabling effective dimensionality reduction. [14]

Finding the best parameter of a given model requires hyperparameter tuning, since for example the optimal parameters of a neural network can be suboptimal if the network is not deep/wide enough. This hyperparameter tuning is labor-intensive and

therefore is achieved by automated/heuristic/model-based algorithms. These ensure that not only the best parameter, furthermore the best hyperparameter are used.

The framework combines the above-mentioned steps to investigate the elastomer specimen volume change for a given liquid/seal combination. For an exemplary regression model and a scaled dataset, the framework is visualized in Figure 3. Since the number of samples in the entire dataset of this work is small ($n = 48$), statistical uncertainties are introduced during evaluation, making it difficult to compare different models. For this reason, the training and testing process is repeated on different randomly selected subsets, which is known as k-fold cross-validation [14]. This approach is visualized in section (a) in Figure 3. In this case, the dataset is divided into $k = 3$ separate subsets that do not overlap. For each fold the best performing model is determined using hyperparameter tuning (see (b)) in Figure 3. For this the HalvingGridSearchCV module is employed [14]. Given the best model and the testing set the model is then evaluated by applying the MSE and R^2 metric (see (c)). This is repeated for every fold and average scores are calculated for each model, enabling comparison between models.

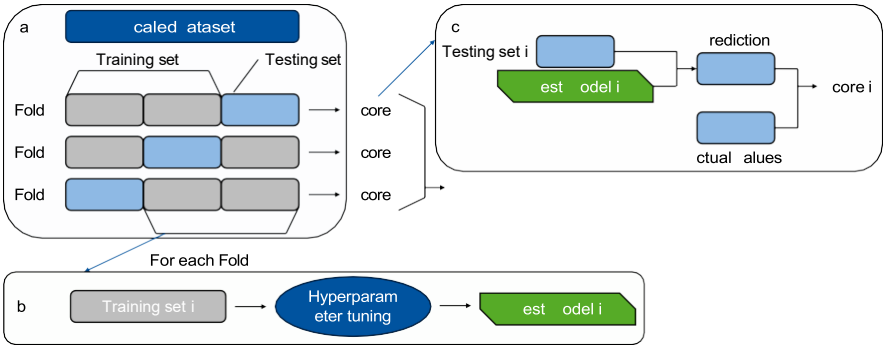

Figure 3: 3-fold cross validation and hyperparameter tuning. Redrawn from [26]

4 Results

In this section, the presented framework is evaluated by applying the MSE and R^2 score to the predictions made by the investigated models. As the data set contains only 48 samples, no statement is made about the actual performance of the investigated models. The focus of the results is on validating the hands-on approach of the framework itself. It must be noted that with each addition of new samples, the database expands, leading to more significant results.

The baseline score is obtained by utilising 5-fold cross-validation, hyperparameter tuning by HalvingGridSearchCV and normalizing all nine input parameters using the min-max scaler of the scikit learn toolbox [14]. The averaged baseline scores for the investigated models are shown in Table 4. It becomes apparent that for all models, the MSE of the training process is always lower than the prediction (testing) process. The discrepancy in R^2 values between the training and testing phases, coupled with the overall low or even negative testing R^2 values, leads to the conclusion that the

models exhibit a high degree of overfitting, failing to adequately represent the underlying patterns.

Table 4: Baseline results of normalized data using all input parameters

Model	MSE		R^2	
	Training	Testing	Training	Testing
linear	0.17	0.48	58.52%	-23.13%
poly	0.00	220.01	100.00%	-52886.43%
lasso	0.31	0.36	23.97%	7.54%
mlp	0.24	0.33	39.24%	13.11%
svr	0.13	0.32	66.71%	13.57%
tree	0.13	0.35	68.90%	13.80%

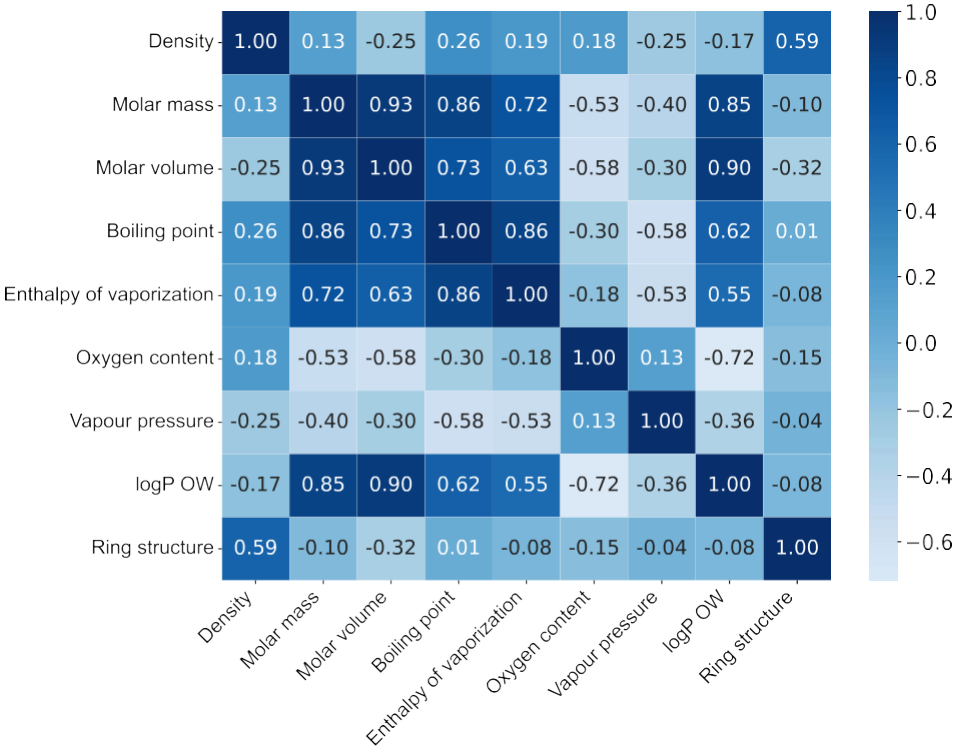

Figure 4: Correlation matrix of input parameters

When investigating the correlation matrix of the input parameters, it becomes apparent that several parameters show a high correlation score (see Figure 4). Applying the same models only on selected parameters with a correlation score lower than 0.80, leads to a small decrease in training accuracy (increase of MSE) while

positively affecting the testing accuracy to a greater degree, and thus decreasing the overfitting effect. This is also apparent in the increase in R^2 score during the testing process (see Table 5). Another benefit of reducing the number of input parameters is the accelerated computation time.

Table 5: Results of normalized input data using selected input parameters. With absolute difference to baseline ΔBL

Model	MSE				R^2			
	Train	ΔBL	Test	ΔBL	Train	ΔBL	Test	ΔBL
linear	0.19	0.02	0.33	-0.15	52.58%	-6%	13.06%	36%
poly	0.09	0.09	0.59	-219.42	76.99%	-23%	-51.02%	52835%
lasso	0.32	0.01	0.33	-0.03	20.83%	-3%	16.84%	9%
mlp	0.20	-0.05	0.30	-0.03	51.49%	12%	21.23%	8%
svr	0.23	0.10	0.32	-0.01	42.50%	-24%	19.53%	6%
tree	0.13	0.01	0.39	0.04	67.58%	-1%	-6.27%	-20%

5 Discussion

Although the overall performance of the single methods cannot be evaluated at this time due to the limited number of data samples, the results indicate the potential for a promising framework that addresses common drawbacks in data-based regression tasks and predicts the elastomers volume change based on known fluid parameters. The initial setup of the framework yields baseline performance scores, to which all future investigations and additions can be compared. It is expected that with more data samples the results become more evident. Those data samples are constantly being generated by conducting more experimental immersion tests.

It is shown that the reduction of dimensionality by only selecting uncorrelated input parameters, the effect of overfitting was reduced. Although a loss of input information was introduced, the overall performance was not affected. Since at this point only a first selection of parameters is chosen to indirectly describe the polarity and steric requirement, further, more substantial, parameters could be introduced to describe the underlying effects of elastomer swelling more accurately.

It must be noted that at this stage of development, no special attention was paid to preselecting the most suitable hyperparameter-grid for each model on which the method “hal ing grid search” was applied to determine the best hyperparameters for the given model and data. It is evident, that providing at least a preliminary selection of suitable hyperparameters would enhance each model’s performance.

6 Summary and outlook

In this work, a hands-on application of ML regression models was developed to precisely plan the further experimental investigation of the material compatibility of bio-hybrid fuels and elastomer sealing materials. Based on the results of previously conducted immersion tests and physics related fluid properties as input parameters the elastomers volume change is predicted with several regression models. Within the developed framework common drawbacks of data-based ML methods were addressed.

Due to the limited number of samples available at this stage of development, a baseline score of the different models was computed which builds the foundation in the ongoing development process. The framework allows to be easily extended by means of addition of new samples or input parameters. Therefore, further immersion tests are currently being conducted, firstly focusing solely on SRE-NBR 28/SX and different fuels or liquids. Based on the expanded database the actual performance of the models can be tuned and evaluated. Simultaneously additional fluid properties are considered as input parameters allowing for the use of an iterative selection process of few parameters that describe the underlying physics most accurately. This way, the best performing model and input parameters are selected for SRE-NBR 28/SX, enabling the extension of the investigation to other (standard-reference) elastomeric materials. Given the selected model and understanding of how fluid properties affect the material compatibility, also the influence of additives in technical seals can be investigated in the future.

7 Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – Exzellenzcluster 86 „The Fuel cience Center” ID: 390919832.

8 References

[1] Hofmeister, M.; Schmitz, K.; Laker, J.; Pischinger, S.; Fischer, M. Neue Herausforderungen an Dichtungswerkstoffe im Hinblick auf bio-hybride Kraftstoffe. *Mobility*, **2022**, *7*.

[2] Hofmeister, M.; Frische, A.; Grunewald, M.; Reddemann, M.; Gütering, C.; Blank, L.M.; Kneer, R.; Schmitz, K. Tribological properties of PTFE sealing materials with regard to bio-hybrid fuels. In: *Sealing technology - old school and cutting edge*: *International Sealing Conference : 21st ISC, Stuttgart, Germany, October 12-13, 2022*; Fachverband Fluidtechnik im VDMA e.V: Frankfurt am Main, **2022**; pp. 89–104.

[3] Makansi, F.; Schmitz, K. Simulation-Based Data Sampling for Condition Monitoring of Fluid Power Drives. *IOP Conf. Ser.: Mater. Sci. Eng.*, **2021**, *1097,* 12018.

[4] Duensing, Y.; Rodas Rivas, A.; Schmitz, K. Machine Learning for failure mode detection in mobile machinery. In: *11. Kolloquium Mobilhydraulik : Karlsruhe, 10. September 2020*. Geimer, M., Synek, P.-M., Eds.; KIT Scientific Publishing, **2020**.
[5] Okolie, J.A. Introduction of Machine Learning and artificial intelligence in biofuel technology. *Current Opinion in Green and Sustainable Chemistry*, **2024,** 100928.
[6] Li, J.; Pan, L.; Suvarna, M.; Tong, Y.W.; Wang, X. Fuel properties of hydrochar and pyrochar: Prediction and exploration with machine learning. *Applied Energy*, **2020**, *269,* 115166.
[7] DIN e.V. *DIN EN 590:2022-05: Kraftstoffe - Dieselkraftstoff - Anforderungen und Prüfverfahren;* Beuth Verlag GmbH: 10772 Berlin, **2022-05-00**.
[8] DIN e.V. *DIN EN 228:2017-08: Kraftstoffe - Unverbleite Ottokraftstoffe - Anforderungen und Prüfverfahren;* Beuth Verlag GmbH: 10772 Berlin, **2017-08**.
[9] Voelker, S.; Deutz, S.; Burre, J.; Bongartz, D.; Omari, A.; Lehrheuer, B.; Mitsos, A.; Pischinger, S.; Bardow, A.; Assen, N. von der. Blend for all or pure for few? Well-to-wheel life cycle assessment of blending electricity-based OME 3–5 with fossil diesel. *Sustainable Energy Fuels*, **2022**, *6,* 1959–1973.
[10] Bardow, A.; Voelker, S.; Groll, N.; Bachmann, M.; Mueller, L.; Neumann, M.; Kossioris, T.; Muthyala, P.; Lehrheuer, B.; Hofmeister, M.; Vorholt, A.; Schmitz, K.; Pischinger, S.; Leitner, W. Towards carbon-neutral and clean combustion with hydroformylated Fischer-Tropsch (HyFiT) fuels, **2023**.
[11] Ackermann, P.; Braun, K.E.; Burkardt, P.; Heger, S.; König, A.; Morsch, P.; Lehrheuer, B.; Surger, M.; Völker, S.; Blank, L.M.; Du, M.; Heufer, K.A.; Roß-Nickoll, M.; Viell, J.; Aßen, N. von der; Mitsos, A.; Pischinger, S.; Dahmen, M. Designed to Be Green, Economic, and Efficient: A Ketone-Ester-Alcohol-Alkane Blend for Future Spark-Ignition Engines. *ChemSusChem*, **2021**, *14,* 5254–5264.
[12] Heitzig, S.; Murrenhoff, H.; Weinebeck, A. Investigation of fluid-seal interaction and their prediction based on the Hansen Parameters. *O + P : Fluidtechnik für den Maschinen- und Anlagenbau*, **2015**, *59*.
[13] Bishop, C.M. *Pattern Recognition and Machine Learning,* 1st ed.; Springer New York: New York, NY, **2016**.
[14] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, **2011**, *12,* 2825–2830.
[15] DIN e.V. *DIN ISO 1817:2016-11: Rubber vulcanized or thermoplastic - Determination of the effect of liquids;* Beuth Verlag GmbH: 10772 Berlin, **2016-11**.
[16] DIN e.V. *DIN ISO 13226:2021-06: Elastomere - Standard-Referenz-Elastomere (SREs) zur Charakterisierung des Verhaltens von Flüssigkeiten auf Elastomere;* Beuth Verlag GmbH: 10772 Berlin, **06.2021**.
[17] DIN e.V. *DIN 53521:1987-11: Prüfung von Kautschuk und Elastomeren - Bestimmung des Verhaltens gegen Flüssigkeiten, Dämpfe und Gase;* Beuth Verlag GmbH: Berlin, **11.1987**.

[18] Kier, L.B. Quantitation of solvent polarity based on molecular structure. *Journal of Pharmaceutical Sciences*, **1981**, *70,* 930–933.
[19] Palomar, J.; Torrecilla, J.S.; Lemus, J.; Ferro, V.R.; Rodríguez, F. A COSMO-RS based guide to analyze/quantify the polarity of ionic liquids and their mixtures with organic cosolvents. *Physical chemistry chemical physics : PCCP*, **2010**, *12,* 1991–2000.
[20] Hansen, C.M. *Hansen solubility parameters: A user's handbook,* 2nd ed.; Taylor & Francis: Boca Raton, **2007**.
[21] Asfour, A.-F.A.; Saleem, M.; Kee, D. de; Harrison, B.H. Diffusion of saturated hydrocarbons in low density polyethylene (LDPE) films. *J of Applied Polymer Sci*, **1989**, *38,* 1503–1514.
[22] Hildebrand, J.H.; Scott, R.L. *The solubility of nonelectrolytes / Joel H. Hildebrand Robert L. Scott,* 3rd ed.; Dover Publ: New York, **1964**.
[23] Cheméo. Chemical and physical properties from Cheméo. https://www.chemeo.com/ (Accessed April 25, 2024).
[24] Linstrom, P. *NIST Chemistry WebBook, NIST Standard Reference Database 69*, **1997**.
[25] Wildman, S.A.; Crippen, G.M. Prediction of Physicochemical Parameters by Atomic Contributions. *J. Chem. Inf. Comput. Sci.*, **1999**, *39,* 868–873.
[26] Raschka, S. *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*, **November 13, 2018**.

9 Authors

RWTH Aachen University, Institute for Fluid Power Drives and Systems (ifas)

Campus-Boulevard 30, 52074 Aachen, Germany:

Lukas Boden, M. Sc., ORCID 0009-0003-1282-8032,
lukas.boden@ifas.rwth-aachen.de

Marius Hofmeister, M. Sc., ORCID 0000-0002-6672-5800,
marius.hofmeister@ifas.rwth-aachen.de

Faras Brumand-Poor, M. Sc. M. Sc., ORCID 0009-0006-7442-8706,
faras.brumand@ifas.rwth-aachen.de

Linda Pleninger, B. Eng., ORCID 0009-0003-9085-2572,
linda.pleninger@rwth-aachen.de

Univ.-Prof. Dr.-Ing. Katharina Schmitz, ORCID 0000-0002-1454-8267,
katharina.schmitz@ifas.rwth-aachen.de

https://doi.org/10.61319/A2FETW2F