Utrecht University

Department of Information and Computing Science

Master Thesis Computing Science

Estimating the distribution of SOH and RUL of electric vertical takeoff and landing aircraft (eVTOL) batteries using Machine Learning algorithms

Supervisor:
Dr. Mihaela Mitici

Student:
Lalitha Boddapati (5867342)

Second examiner:
Dr. Shihan Wang

December 12, 2023

Abstract

Prognostics and Health Management (PHM) involves a thorough analysis of the health condition of a machine and its components. In the field of battery management and PHM, the State of Health (SOH) and the Remaining Useful Life (RUL) are key indicators used to assess the current condition of a battery and predict how much more service life it has. Many studies focus on point estimates for the RUL and SOH of batteries in various domains. However, for the reliability and predictive maintenance planning, it is of utmost importance to determine the uncertainty associated with the predictions of SOH and RUL so that the decision-makers can make accurate informed decisions.

In this thesis, a machine learning pipeline for estimating the distribution of SOH and RUL on 21 Electric Vertical Take-off and Landing vehicle (eVTOL) batteries that are cycled under various conditions is proposed. Utilising the segments of charge and discharge phases of these batteries, 30 features are generated and an automatic feature selection is performed using the Boruta-SHAP algorithm. The pipeline estimates the distributional mean (μ) and provides uncertainty (σ) for each capacity test cycle of eVTOL batteries by using Random Forest Regression, Convolution Neural Network with Monte Carlo Dropout, and Mixture Density Network algorithms. The accuracy and sharpness of the obtained distributions are evaluated using the Continuous Ranked Probability Score (CRPS). The results show that the Random Forest Regression performs better than the other models and predicts the distribution of SOH and RUL with an average CRPS score of 0.96% and 40.49 missions/cycles, respectively.

Contents

List of Figures

List of Tables

1 Introduction

Electric Vertical Take-off and Landing vehicles (eVTOLs) are the ones that use electric power to hover, take off, and land vertically. eVTOL is one of the new technological developments in the aerospace industry. eVTOL uses the technology that has seen success in electric propulsion based on progress in motors, batteries (Lithium-ion), fuel cells, and electronic controller technologies and is fuelled by the emerging need for new aerial vehicles for urban air mobility that can enable greener and quieter flights. They are also seen as a solution to the growing traffic congestion in large cities, the traffic-related pollution, and the inter-city connectivity needs. Several companies such as Airbus, Bell, Boeing, Embraer, Joby Aviation, Kitty Hawk, Pipistrel, Volocopter, and Aurora Flight Sciences have been designing, building, and testing eVTOLs for the past few years [44][56].

Lithium-ion (Li-ion) batteries are widely used as energy storage systems, such as in electric vehicles (EVs), and hybrid electric vehicles (HEVs) because of their high energy density, low self-discharge rates, and acceptable costs. The performance of Li-ion batteries deteriorates with time and usage due to the degradation of its electrochemical constituents, resulting in capacity and power fade. This is called battery aging and is a consequence of multiple aging mechanisms that are influenced by factors such as battery chemistry, manufacturing, environmental, and operating conditions. Due to its complicated electrochemical characteristics and complex working conditions, lithium-ion batteries are highly dynamic and have non-linear features. To ensure the reliability and safety of battery systems, prognostics and health management methods become significantly important to accurately estimate the battery status so that one can avoid unexpected failures of the system and subsequent losses [35].

Prognostics and Health Management (PHM) is an engineering technique that involves thorough analysis of the health condition of a machine and its components. It employs tools from data science, statistics, and physics to detect an eventual fault in the system (anomaly detection), classify it according to its specific type (diagnostic), and forecast how long the machine will be able to work in the presence of this fault (prognostic). In the field of batteries, PHM approaches for batteries include accurate estimation of State of Charge (SOC), State of Health (SOH) and Remaining Useful Life (RUL) are of critical importance for obtaining the remaining charge, capacity, and life of batteries. These estimates enable us to plan maintenance strategies in a timely manner, conduct disposal and replacement of batteries [43][57][55]. In this thesis, the focus is on the distribution estimation of SOH and RUL of batteries, as they are considered to be the two significant health indicators of Li-ion batteries.

In general, the prognostic approaches for RUL and SOH estimation in the literature are classified as physics-based, data-driven, and hybrid techniques. Physics-based models are accurate and can depict the degradation mechanism of a battery by building a mathematical model of the electrochemical degradation mechanism. These models require detailed domain knowledge of the underlying degradation processes leading to failure, which is sometimes difficult to observe and incorporate in models. Moreover, tests related to batteries would be conducted in a stable environment, which might not be representative of real-time operating environment, as in the real life these conditions keep changing. There is huge scope for randomness that has an influence on the degradation process of batteries. For this reason, it is known that establishing a reliable mathematical model is difficult in complex and uncertain environments. On the other hand, data-driven methods do not require a thorough understanding of the physical phenomenon or expert knowledge of the capacity degradation mechanism, instead, they use information from historical data to predict the future degradation trend. Hybrid models combine the advantages of physics-based and data-driven models but they still remain too complex.

Among these approaches, data-based approaches are gaining more importance due to recent advancements in computational power, artificial intelligence, and statistical machine learning techniques. Also, the increase in the availability of battery cycle data from actual applications and the development of a wide range of sensors to measure various battery parameters have

enabled data-driven battery prognostics. Sensor data from real time measurements on batteries is used to estimate battery health and predict remaining useful life without the need for complex physical models. Hence, our focus lies in the estimation of SOH and RUL of batteries using data-driven approaches i.e. using machine learning algorithms. However, the efficiency and accuracy of predictions from such models largely depend upon the data diversity, quality, and size of the dataset used for training the data based prognostic model [34][20][40][41]. This means when fitting models in terms of supervised learning (using labeled datasets for prediction purposes), there is an uncertainty associated with the predictions.

Uncertainty refers to the lack of confidence in output of a machine learning algorithm. We need to understand what generates this uncertainty, how to quantify it, and how to reduce it so that we can make informed decisions. That is, we want to make sure that models can describe as precisely as possible, the likelihood of their outcome being wrong or outside a given range of accuracy. Uncertainty estimation is particularly important for machine learning and neural networks because the models tend to provide overconfident predictions. Incorrect overconfident predictions can be very hazardous in critical use cases such as healthcare or autonomous vehicles. For safe and informed decisions, we want the models not only to provide an output, but also provide information regarding the level of certainty of the outcomes.

Several studies are done on the health management of lithium-ion batteries for on-ground electric vehicles. The batteries in these studies are usually subjected to constant current (CC) and constant voltage (CV) charging protocol with constant discharge C-rates. Whereas, for eVTOLs, the take-offs and landings have a larger discharge rate than the cruise phase. This is the difference in the batteries that are used for on-ground electric vehicles and eVTOLS. This in fact, is expected to have a direct influence on the health of the batteries in the long run. Moreover, it affects the SOH and RUL of the batteries, which are sensitive to the long-run degradation trends of the batteries. Thus, given their safety-criticality, the health management of the batteries remains one of the challenges in the eVTOL industry [37]. Moreover, RUL and SOH estimations are essential for scheduling, planning, maintenance, and replacement procedures in order to avoid unexpected capacity drops and failure to ensure system reliability [34][30].

Many studies focused on point estimates i.e., one value for the RUL and SOH prediction in various domains. In [47], the authors have integrated a deep learning approach for RUL prediction of multiple lithium-ion batteries, by integrating autoencoder with DNN (ADNN) that is with fully connected layers and achieved RMSE - 6.66%, and accuracy - 93.34%. To achieve a better RUL prediction, the authors [20] proposed a hybrid deep neural network that combines CNN and LSTM to extract the spatio-temporal relations in multivariate time series data in order to capture nonlinear characteristics. In [22], the authors used Deep Neural Networks (DNN) approach to predict the SoH and the RUL of the lithium-ion battery. The proposed DNN algorithm was compared against other machine learning algorithms, namely, Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Artificial Neural Networks (ANN), and Linear Regression (LR). In [29], random forest regression (RFR) is used for battery capacity estimation and to evaluate the health status of different batteries under varied cycling conditions. They have achieved a root-mean-square error of less than 1.3%. In [41], battery capacity and SOH estimations are carried out using Gaussion process regression (GPR) and Support vector machine (SVM) models and predicted accuracy with low RMSE values of 0.0181 and 0.0221 for GPR and SVM based models, respectively. Further, the authors have combined multi battery data sets to improve the prediction accuracy. They showed that combining the data of multiple batteries with similar operating conditions for training a model resulted in higher prediction accuracy. In [13], Convolutional Neural Networks (CNNs) are used to obtain the RUL prognostics for turbofan engines and based on RUL prognostics, an alarm policy is implemented to trigger for maintenance tasks. In the paper [37], the authors proposed a framework for SOH and RUL prognostics for eVTOL batteries along with feature importance quantification for the estimation. They have estimated the SOH and RUL of eVTOL batteries using SVM, RFR,

Extreme Gradient Boosting (XGBoost), GPR, and Multilayer perceptron (MLP). The authors show that the discharge-related features have the highest importance when predicting battery state-of-health (SOH) and remaining useful lifetime (RUL) and it is shown that the battery SOH and RUL are well estimated using Random Forest regression and Extreme Gradient Boosting, respectively.

All the above mentioned studies predict RUL and/or SOH as a point estimate. For the reliability or predictive maintenance planning, however, it is of utmost importance to determine the uncertainty associated with the estimated SOH and RUL and point estimates are not enough, as it is risky to make decisions based on such deterministic results. In contrast to paper [37], the main goal of this thesis is to estimate the distribution of SOH and RUL of electric vehicle takeoff and landing (eVTOL) batteries using Machine Learning algorithms. That is to quantify the uncertainty associated with predictions of SOH and RUL of eVTOL batteries, in order to provide additional support for decision-making.

2 Literature Survey

The literature survey is divided into two sections. In the first section, Uncertainty Quantification(UQ) methods are briefly discussed. In the second section, the relevant papers that have implemented a few of the UQ methods and other machine learning methods for estimating the distribution are discussed. This is done to understand the various methodologies implemented and to see if they can be adopted for the eVTOL dataset used in the current thesis.

2.1 Uncertainty Quantification Methods

In general, there are two main sources of uncertainty described in the literature, namely aleatoric and epistemic uncertainties [4][1][11]. The intrinsic randomness of the data generating process that cannot be explained even when given more observations or data is an aleatoric uncertainty (also known as data uncertainty or statistical uncertainty) e.g., in sensor readings, manufacturing, and operational processes. This type of uncertainty is not the property of the model, but rather it is the result of the inherent noise of the data distribution and hence it is irreducible. We further have two types of aleatory uncertainty. Homoscedastic uncertainty denotes the variance that stays constant for all input parameters. Heteroscedastic uncertainty, on the other hand, is the variance that evolves over the lifetime of the system like in prognostics, and can potentially be predicted as a model output.

Another type of uncertainty is epistemic uncertainty (also known as knowledge uncertainty or systematic or model uncertainty) which occurs due to inadequate knowledge and data. In other words, it is due to a lack of data that is potentially available, for instance in the case of a data-rich problem, there is a collection of massive data but it may be informatively poor, increase in the diversity of a number of observations may explain this uncertainty. Note that when augmenting the model to incorporate additional features, parts of the uncertainty that previously were of aleatory type will become epistemic uncertainty. Once the additional input parameters are captured, they can be incorporated into the model calculations. In practice, however, it is difficult to gather enough data to cover the diversity of operating situations leading to the different failure modes. In literature, the most widely used approaches to estimate the uncertainty of the output of neural networks are Bayesian learning, Gaussian process, deep ensembles (DE), and Monte Carlo dropout (MCD) [4][1][11].

Bayesian deep learning (BDL): The Bayesian paradigm is a statistical/probabilistic paradigm in which a prior knowledge, modeled by a probability distribution, is updated each time a new observation, whose uncertainty is modeled by another probability distribution, is recorded. The whole idea behind the Bayesian paradigm is so called Bayes theorem that expresses the rela-

tion between the updated knowledge (the "posterior"), the prior knowledge (the "prior") and the knowledge coming from the observation (the "likelihood"). In a traditional neural network (NN), weights are assigned as a single value or point estimate, whereas in a Bayesian neural network (BNN), weights are considered to be a probability distribution. These probability distributions of network weights are used to estimate the uncertainty in weights and predictions. The epistemic uncertainty is introduced with a prior distribution over the weight space and a posterior distribution inferred from data. Heteroscedastic uncertainty is usually modeled with a parametric likelihood distribution, whose parameters are the BNN outputs.

Let us consider a model where data X is generated from a probability distribution depending on an unknown parameter θ. Assume that the prior knowledge about the parameter θ is known and can be expressed as a probability distribution $P(\theta)$. Then, when data X is observed, the prior knowledge about this parameter can be updated using the Bayes theorem, which is as follows:

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

where, $P(\theta|X)$ is the posterior probability distribution of the parameter given the observed data, $P(X|\theta)$ is the likelihood probability distribution of the parameter given a parameter value, $P(\theta)$ is the prior probability distribution of the parameter independently from any observation, $P(X)$ is the probability distribution of the observed data independently from any parameter value (marginal likelihood) or evidence or normalisation factor.

The computational difficulty with Bayes theorem is that, though, in many situations, the prior and the likelihood are explicitly known, the normalisation factor requires to be computed such that

$$p(X) = \int_{\theta} p(X|\theta)p(\theta)d\theta$$

This integral computation is not a problem for low dimensions, however, it becomes intractable in higher dimensions if we have to predict output y' for a new data point x' by integrating over all parameters θ. That is, the exact computation of the posterior distribution becomes practically infeasible and some approximation techniques have to be used in order to find solutions to problems that require to know this posterior. There are two ways to get around computing the model evidence. The first is to work with posterior ratios so the model evidence term cancels out; this is the idea behind Markov chain Monte Carlo (MCMC). The second is to optimize a bound on the evidence, the idea behind variational inference (VI) [42]. Alternatively, Monte Carlo Dropout (MCD) [15] and deep ensembles (DE) [26] have become popular for practical reasons due to their simplicity, scalability, and computational efficiency though they are not strictly Bayesian.

Markov chain Monte Carlo (MCMC): These methods provide a potential solution by simulating a large number of samples from a distribution that gradually approach the true posterior, assuming that we have a way to draw samples from a probability distribution defined up to a factor. Then, use these samples to compute various parameters such as mean and variance or even to approximate the distribution by Kernel Density Estimation. Most of the sampling-based inference algorithms are instances of Markov Chain Monte-Carlo (MCMC). Two popular MCMC methods are Gibbs sampling and Metropolis-Hastings. Although these methods guarantee to find a globally optimal solution given enough time, it is difficult to say how close they are to finding a good solution given the finite amount of time in practice. Therefore, choosing an appropriate sampling technique is itself tough and comes with huge computational requirements. [24]

Variational Inference (VI): The main idea of variational methods is to treat inference as an optimization problem. These methods try to find the best approximation for intractable probability distribution p among a given family. That is, these techniques try to solve an optimization

problem over a class of tractable distributions $\mathcal{Q}$ in order to find $q \in \mathcal{Q}$ that is most similar to p.

$$q^*(\theta) = argmin\ KL\ (q(\theta)\ ||\ p(\theta|x))$$

where, $q^*(\theta)$ is the best approximation of the posterior distribution that belongs to a class of tractable distributions $\mathcal{Q}$. To optimise the parameters to obtain the closest element to the target, we need a well defined error measure [25]. The similarity is measured in terms of the Kullback-Leibler (KL) divergence, which is also called relative entropy, in the above equation KL (.||.) denotes the Kullback-Leibler divergence given as below:

$$KL\ (q(\theta)\ ||\ p(\theta|x))\ =\ \int\ q(\theta)\ log\frac{q(\theta)}{p(\theta|x)}\ d\theta$$

However, KL divergence cannot be computed because it still depends on the integral $p(x) = \int p(x|\theta)p(\theta)d\theta$. So, instead of minimizing the quantities, we try to maximize the evidence lower bound (ELBO). The ELBO gives an approximate distribution of $q(\theta)$ for latent quantities θ that allows the data to be predicted well, i.e., leads to a high expected log likelihood, but a penalty is incurred if $q(\theta)$ is far away from the prior $p(\theta)$. To see the dependency of maximizing the ELBO, the desired KL divergence needs to be minimized (refer [25]).

To solve the tractability issue, there are two common approaches, one is the Expectation-Maximization (EM) algorithm that fixes the parameters during the expectation step and then computes when it's possible $P(\theta|X)$ to then optimize the ELBO. That is the EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. These parameter estimates are then used to determine the distribution of the latent variables in the next E step. Two, directly optimize the ELBO term by making assumptions on q.

Mean-field inference: VI concerns the choice of approximating family $\mathcal{Q}$ which includes exponential families, neural networks, Gaussian processes, latent variable models, and many other types of models. The most widely used is factorizing the variational posterior over its dimensions, $q(x) = q_1(x_1)....q_n(x_n)$. Here each $q_i(x_i)$ is a categorical distribution over a one-dimensional discrete variable, which can be described as a one-dimensional table. This choice of $\mathcal{Q}$ is called mean-field inference. The resulting cost function is as follows

$$\mathcal{L}(\mathcal{D}, \theta) = KL(q(w|\theta)||P(w)) - E_{q(w|\theta)}[logP(\mathcal{D}|w)]$$

where $q(w|\theta)$ posterior over its dimensions, $P(w)$ the prior, $w, \mathcal{D}$ are the weights and training data respectively. the first term is the prior-dependent complexity cost, while the second one is the data-dependent likelihood cost. The goal is to find a trade-off between the two costs, which is also a way of fitting the data with regularization. Unlike sampling-based methods, variational approaches will almost never find the globally optimal solution. However, we will always know if they have converged. In some cases, it is even necessary to have bounds on the accuracy. In practice, VI methods often scale better [25]. Bayes by Backprop, The local reparametrization trick (LRT), Flipout (FO), and Radial BNN (RAD) are some of the practical techniques that offer a solution to adequately solve the issue of the aforementioned intractability [9][4].

Monte Carlo Dropout (MCD): Dropout is originally a regularization technique that is used to prevent overfitting. This is done by randomly deactivating some neurons with a certain probability when training the neural network. As a regularization technique, dropout is only used during the training phase. In [15], the authors proposed MCD as a method that uses dropout at test time to obtain a distribution of possible models. To estimate epistemic uncertainties, we drop a different set of neurons on each pass and look at the distribution of results obtained.

Its effectiveness as being an approximation of the Bayesian neural network has been proved in the paper. They have shown that deep neural networks with arbitrary depth, non-linearity, and dropout layer are a mathematically equivalent approximation to the Bayesian inference in deep Gaussian process. It can be interpreted as an ensemble combination model where the final prediction is determined based on averaged network outputs. It can be regarded as VI whose variational posterior distribution of weight matrix is defined as:

$$z_{i,j} \sim Bernouilli(p)_i$$

$$W_i = M_i.diag(z)_i$$

where $z_{i,j}$ is a coefficient to control the state of neurons, and M_i are the weights of the neural network before applying dropout.

Deep ensembles (DE): In [26], an alternative to MCD for predictive uncertainty estimation was developed. Unlike MCD, which generates an ensemble from only one trained model using dropout, DE individually trains multiple models with different parameter initializations. The expectation is that their predictions should be similar, while not being exactly the same. The randomness inherent in the initializations and the training process provides different samples of trained network parameters. If the networks are optimised to minimize the mean squared error loss, this gives a measure of epistemic uncertainty. On the other hand, the data log likelihood can be optimised to estimate both aleatoric and epistemic uncertainties. A related technique is bagging, also called as 'bootstrap aggregating.' This adds another source of randomness by training each network with a different randomly drawn data with replacement from the training set. In the paper, it is observed that performance has deteriorated when bagging is used, and no improvement is seen. Moreover, DE comes at the cost of training multiple models and multiplying the number of parameters, so it is less computationally efficient than MCD. In the context of a supervised regression problem, a mean μ and a variance σ^2 are computed independently by each neural network in the ensemble. For an input x_i of the dataset, the mixture mean and variance are computed as follows:

$$\hat{\mu}(x_i) = \frac{1}{M}\sum_{m=1}^{M}\mu_m(x_i)$$

$$\hat{\sigma}^2(x_i) = \frac{1}{M}\sum_{m=1}^{M}(\sigma_m^2(x_i) + \mu_m^2(x_i)) - \mu^2(x_i)$$

Gaussian process (GP): A Gaussian Process is a collection of random variables, any finite number of which have a joint Gaussian, i.e. multivariate normal distribution. We can express a Gaussian Process (GP) using an index set T such that the collection of random variables is given by $(X_t)_{t\in T}$ where $(X_{t_1}, .., X_{t_n}) \sim \mathcal{N}(\mu, \Sigma)$, is multivariate normally distributed with parameters $\mu \in \mathbb{R}^n$ and $\Sigma = (\Sigma_{i,j}) \in \mathbb{R}^{n\times n}$ depending on $t_1, ..., t_n \in T$ in the following manner:

$$\mu_i \;=\; m(t_i),\; m \;:\; T \;\rightarrow\; \mathbb{R}$$

$$\Sigma_{i,j} \;=\; k(t_i, t_j),\; k \;:\; T \;\times\; T \;\rightarrow\; \mathbb{R}$$

where m the mean function and k the covariance function of the Gaussian Process, which is as follows

$$(X_t) \;\sim\; \mathcal{GP}\;(m(t), k(t, t^{'}))$$

The mean function can be an arbitrary real-valued function, in order to ensure that Σ is always symmetric and positive semi-definite, the covariance function needs to be a symmetric and

positive semi-definite kernel. Mean and covariance functions do not impose a restriction on the set of GPs that can be represented. The mean function of a Gaussian Process describes its general trend, while the covariance function is responsible for the characteristics of the process. More specifically, the covariance function takes influence on, for instance, the volatility in terms of amplitude and length-scale of the random function. Periodicity can also be generated with a suitable covariance function. It can have many forms as long as it follows the properties of a kernel (i.e. semi-positive definite and symmetric). Some common kernel functions include constant, linear, square exponential, and composition of multiple kernels. A popular kernel is the composition of the constant kernel with the radial basis function (RBF) kernel, which encodes for smoothness of functions (i.e. similarity of inputs in space corresponds to the similarity of outputs), which is given as below:

$$k(t, t^{'}) \ = \ \sigma^2 exp(-\frac{1}{2l^2} \ ||t - t^{'}||^2)$$

where, the hyperparameters of the kernel are signal variance σ^2, and length scale l. The form of the mean function and covariance kernel function in the GP prior is chosen and tuned during model selection. The mean function is typically constant, either zero or the mean of the training dataset [46]. Thus, Gaussian process regression (GPR) is a nonparametric method and calculates the probability distribution of parameters of a specific function, GPR calculates the probability distribution over all admissible functions that fit the data, i.e., variance around its mean prediction to describe the associated uncertainty. Since, we chose a Gaussian process prior, calculating the predictive distribution is tractable, and leads to a normal distribution that can be completely described by the mean and covariance (variances can be obtained from the diagonal of the covariance matrix).

2.2 Relevant papers for prognostics that used above mentioned methods

In [12], the distribution of RUL of turbofan engines is estimated as a probability density function (pdf) using a Convolutional Neural Network (CNN) with Monte Carlo dropout. The CNN consists of 5 convolutional layers, where the first four convolutional layers have 10 kernels, each of size 10×1 (i.e., one-dimensional kernels). The last convolutional layer has one kernel of size 3×1, combining all 10 feature maps into one feature map. This last feature map is flattened in a flatten layer, and connected to a fully connected layer. All these layers used tanh activation function. To predict the RUL using the Rectified Linear Unit (ReLU) activation function, one single neuron is attached to the fully connected layer. To obtain a probability distribution of the RUL using CNN, the authors applied dropout rate of $\rho = 0.5$ in each layer (except for the last convolutional layer before the flatten layer and the first to avoid the loss of input information) during the training phase and also during the testing phase for each test instance by randomly selecting neurons to be dropped.

When the probability density function (PDF) of the RUL is estimated, the regular MAE (Mean Absolute Error), RSME (Root Squared Mean Error), and the Mean Score are computed based on the actual RUL vs. the mean of the predicted RUL. The following are the formulas for the computation, where $\bar{y}_i$ is the mean predicted RUL of test instance i:

$$MAE \ = \ \frac{1}{N} \sum_{i=1}^{N} |\bar{y}_i - y_i|$$

$$RSME \ = \ \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\bar{y}_i - y_i)^2}$$

$$Mean \ Score \ = \ \frac{1}{N} \sum_{i=1}^{N} s_i$$

where,

$$s_i = \begin{cases} e^{-\frac{\bar{y}_i - y_i}{\gamma}} - 1, & \bar{y}_i - y_i < 0 \\ e^{-\frac{\bar{y}_i - y_i}{\delta}} - 1, & \bar{y}_i - y_i \geq 0 \end{cases}$$

where γ and δ are user defined values.

The authors have also introduced metrics that are suitable to evaluate the accuracy, sharpness, and reliability of probabilistic RUL prognostics. For accuracy and sharpness evaluation, they used Continuous Ranked Probability Score ($CPRS$) and weighted Continuous Ranked Probability Score ($CPRS^w$). To explicitly evaluate the reliability, they used $\alpha - Coverage$ and reliability score (RS). CRPS evaluates two things, one if the estimated RUL distribution is centered around the actual RUL of a component i, i.e., the accuracy of the RUL. Two, if the variance of the RUL distribution is low, i.e., the sharpness of the RUL prognostic. The weighted CRPS uses penalties when the RUL is overestimated/underestimated, depending on the type of component, the penalty is adjusted.

To evaluate the reliability of the predictions, the authors have constructed the coverage of a probabilistic RUL prognostic without assuming that this prognostic follows a specific distribution, such as the Gaussian distribution. To calculate the coverage, they constructed a credible interval around the median of the estimated RUL distribution with width α. The closer the coverage is to α, the more reliable the estimated RUL distribution is. The uncertainty is overestimated if the coverage is larger than α. Conversely, the uncertainty is underestimated if the coverage is smaller than α. If two RUL prediction methods have the same coverage for a width α, the method that provides tighter credible intervals is preferred. In other words, a higher sharpness of the RUL distributions is preferred. Note that, a higher sharpness also leads to a lower CRPS.

$$\alpha - coverage = \frac{1}{N} \sum_{i=1}^{N} \mathcal{I}(\alpha)_i$$

$$\mathcal{I}(\alpha)_i = \begin{cases} 1, & y_i \in [\hat{y}_i^{0.5-0.5\alpha}, \hat{y}_i^{0.5+0.5\alpha}] \\ 0, & otherwise \end{cases}$$

where $\alpha \in [0, 1]$ is an user-defined parameter and $\hat{y}_i^k$ is the RUL prediction of the k^{th} percentile of the estimated RUL distribution of component i.

Since the reliability evaluated by coverage metric is relative to a specific α, the authors introduced the Reliability Score (RS) in order to conduct a generic, parameter-free reliability analysis of the estimated RUL distribution. They defined a reliability curve $C(\alpha)$ based on α-Coverage, that is $C(\alpha) = \{\alpha\text{-Coverage}, \alpha \in \{0.00, 0.01, 0.02, ..., 1.00\}\}$. In general, for classification problems, the Brier Score is used to quantify the reliability of predictions, this is not directly applicable to regression problems because each test instance may fall into multiple credible intervals. To address this, The authors defined Reliability scores (RS) to quantify the reliability of the RUL prognostics as follows:

$$RS^{underestimated} = \int_0^1 \mathcal{I}\{C(\alpha) \leq \alpha\}(\alpha - C(\alpha))d\alpha$$

$$RS^{overestimated} = \int_0^1 (1 - \mathcal{I}\{C(\alpha) \leq \alpha\})(C(\alpha) - \alpha)d\alpha$$

$$RS^{total} = RS^{overestimated} + RS^{underestimated}$$

with

$$\mathcal{I}\{C(\alpha) \leq \alpha\} = \begin{cases} 1, & C(\alpha) \leq \alpha \\ 0, & Otherwise \end{cases}$$

The area RS-underestimate is computed between the ideal curve and the reliability curve $C(\alpha)$ when the reliability curve is below the ideal curve to quantify the extent to which the uncertainty associated with the probabilistic RUL prognostics is underestimated, similarly to quantify the overestimation, we calculate the area RS-overestimated between the ideal curve and the reliability curve $C(\alpha)$ when the reliability curve is above the ideal curve.

In [54], machine learning methods to predict the distributions of flight delays, and three metrics are developed to evaluate the performance of the algorithms is proposed. The machine learning algorithms selected to predict the distribution of flight delays are multilayer perceptron (MLP), LightGBM, and random forest (RF). The authors assumed that flight delays follow normal distributions. This is because, in their experiments, the kernel density curve is very close to the fitted normal distribution curve. In MLP, they had fully connected NN and used dropout to prevent overfitting. The algorithm updates the network weights through backpropagation according to the loss function. In order to get interval prediction, the quantile loss function (Quantile) is used, which estimates the conditional quantile of a given predicted value. The loss function is as follows:

$$Loss_{quantile} = \frac{1}{m}\sum_{i=1}^{m}\left(\sum_{i:y_i<\hat{y}_i}(1-\gamma)|y_i-\hat{y}_i| + \sum_{i:y_i\geq\hat{y}_i}\gamma|y_i-\hat{y}_i|\right)$$

where $\gamma \in [0,1]$ is a user defined value.

LightGBM is an ensemble method that trains a series of decision trees sequentially and uses Gradient-Based One-side Sampling (GOSS). Instead of using all sample points to calculate gradients, LightGBM calculates gradients after sampling. Exclusive Feature Building (EFB) does not use all features to obtain the best segmentation point but rather correlates some features together to reduce the feature dimension. This technique can be used for the feature section. In General, decision tree algorithms build their trees through a level-wise strategy, which does not distinguish leaf nodes in the same layer. Additionally, some leaf nodes in the same layer need not grow. To overcome such problems, LightGBM uses leaf-wise tree growth and prevents overfitting with a max depth limit. Another algorithm implemented is Random forest, which is a tree based ensemble learning algorithm, RFs consist of hundreds of decision trees, each of them built over a random extraction of the observations from the dataset and a random extraction of the features. Since not every tree uses all the features or all the observations for its construction, we can say that the trees are de-correlated and therefore less prone to over-fitting. Average/mean of all the constructed random forest outputs is considered as the final prediction.

The authors have proposed three metrics, one, is the prediction accuracy rate under a given confidence level: the algorithm predicts the mean and standard deviation of the distribution, and it generates the corresponding normal distribution curve. If the mean value of the actual delay distribution falls within this confidence interval under this confidence level, the prediction is said to be a correct prediction. Two, the average regional area under given interval level. Here, the mean value of the actual delay distribution corresponds to two points on the x-axis at a given interval level. A larger regional area indicates a more accurate prediction. Third, the Wasserstein distance is used to measure the difference between the actual delay distribution and the predicted delay distribution. The smaller the value, the more accurate the prediction. Wasserstein distance is considered instead of KL divergence because every flight follows a different normal distribution. The predicting performance is also from an entire flight schedule perspective by using the Cumulative Distribution Function (CDF) of flight delays of an entire schedule from actual (original), actual (fitted), and predicted data The authors have also tested various distribution functions to model flight delays, including Beta distribution, Erlang distribution, and Normal distribution. However, the results suggested that Normal distribution is able to capture the stochastic nature of flight delay than others.

The authors of [60] used Mixture Density Networks (MDN) and Random Forest regression (RFR) models to estimate a probability distribution for flight delays on an individual flight basis and integrated these probabilistic predictions into a probabilistic flight-to-gate assignment problem. MDNs are a combination of a neural network and a Gaussian mixture model. Given features, MDN outputs parameters for each Gaussian in the mixture: the weight α, the mean μ, and the standard deviation σ. With these parameters, the probability density function of the target variable is determined. If there are m total number of Gaussian components for a mixture, then there are 3m outputs of the MDN. The weights use a softmax activation function, and the standard deviations use an exponential activation function, while the means are unrestricted. The neural network is trained using backpropagation, i.e., the weights and biases of each node are updated using an error function, which is the negative logarithm of the likelihood that the model derived from the output of the current network gives rise to the training data. RFR is the same as the previously discussed RF, however, to obtain the distribution in the test phase, the output values of the decision trees are not averaged, but collected, and a kernel density estimation (KDE) is performed. Two settings of the KDE are the kernel type and the bandwidth. In the paper, a bandwidth of 1.5 and the probability density function are estimated from deterministic feature values without the need for the stochastic variable. Metrics used are CPRS, RMSE, MAE.

In [27][36], the distribution of RUL is achieved using Convolutional Neural Networks with Monte Carlo dropout. In [27] the method proposed is Multi-channel CNN with MC dropout, where one 1D kernel per each time-series of a feature is applied, i.e., each column of the input x is convoluted with different 1D kernels, whereas in single channel CNN with monte carlo dropout a common 1D kernel is applied for all features. The authors have implemented both. Multi-channel 1D convolutional layers are shown to be effective for multi-variate time-series, which is also the case of the C-MAPSS data set that is analysed in this study. Since an independent kernel is used for the time-series of each feature, the convolutional layers are able to learn the patterns of each feature. RMSE is used as a metric for evaluation and the quality of the estimated RUL distributions are done using calibration plots. For a perfectly calibrated model, the probability that the true RUL is less than or equal to the δ% quantile of the estimated distribution is δ%. In [5], deep Gaussian process learning is used to estimate the confidence interval of RUL predictions for N-CMAPSS (New Commercial Modular Aero-Propulsion System Simulation) dataset from NASA. They have used regular RMSE and negative log-likelihood (NLL) for evaluation. Apart from these, they have also incorporated $\alpha - \lambda$ metric that measures if the prediction accuracy of a RUL model is within a α% error at specific time instances during the relative lifetime λ of the system. They have introduced a probabilistic version of the $\alpha - \lambda$ metric to account for predictive uncertainty. [36] used RMSE and $\alpha - \mathit{coverage}$ as evaluation metrics.

2.3 Feature Selection & Feature Importance Techniques

Feature Selection (FS) is the way of identifying and selecting the most relevant subset of features out of the original features in a dataset. In other words, we get rid of noise/bias in data. The selected subset of features are used as inputs to a model. FS reduces the dimensionality of the dataset, thereby resulting in faster training time, improved accuracy, and reduced generalization error introduced by irrelevant features, and leads to a more robust models that are less prone to overfitting. However, the number of possible feature subsets grows exponentially with the increase of dimensions. Finding a globally optimal subset is usually intractable and is considered as NP-hard due to the exponentially increased time requirement [58]. Statistics is one technique which shows the relationship between the features and its importance. The misleading results obtained due to irrelevant and redundant features is not because of lack of useful information but rather because the features did not have a statistical relationship with other features. Individually any feature may be irrelevant but can be relevant when joined with other features

[16][53]. Moreover, by applying feature importance, we can have post-hoc interpretability of complex models such as a neural network or ensemble model.

Based on the interaction with the learning model, FS methods are classified into three types, namely, Filter, Wrapper, and Embedded Methods. In the *Filter method*, the features are selected based on statistical measures. Few of such measures used to understand the importance of the features include Information gain, chi-square test, Fisher score, correlation coefficient, and Variance threshold. These methods are independent of the learning algorithm and require less computational time. However, due to the lack of a specific learning algorithm guiding the feature selection process, the selected features may not be optimal for the target learning algorithms. A filter method consists of two steps. In the first step, feature importance is ranked according to some statistical measures. The feature importance evaluation process can be either univariate or multivariate. In the univariate scheme, each feature is ranked individually regardless of other features, while the multivariate scheme ranks multiple features in a batch manner. In the second step of a typical filter method, the lower ranked features are filtered out [53] [28].

The second approach is the *Wrapper method* where the feature subsets are selected based on the predictive performance of a predefined learning algorithm to evaluate the quality of selected features. Given a specific learning algorithm, a typical wrapper method performs two steps, in the first step it searches for a subset of features. In the second step, it evaluates the selected features. It repeats steps one and two until some stopping criteria are satisfied. The feature set search component first generates a subset of features, and then the learning algorithm acts as a black box to evaluate the quality of these features based on the learning performance. Depending on the accuracy measured from the previous step, the method will decide whether to add or remove a feature from the selected subset. Due to this, the wrapper methods are computationally expensive. However, these methods are considered to be more accurate than the filter method. A few examples include Recursive feature elimination, Sequential feature selection, Exhaustive feature , and Genetic algorithms [53] [28]. *Forward selection* is a greedy algorithm that starts with an empty set of features and adds features one by one until the model performance reaches a peak. *Backward elimination* starts with a complete set of features and removes features one by one until the model performance reaches a peak. *Exhaustive feature selection* one can go over all possible feature combinations and pick up the model with the highest accuracy. *Recursive feature elimination* (RFE) fits a model and removes the weakest features until the specified number of features is reached. Recursive feature elimination is a type of backward selection method and works on feature ranking.

In *embedded method*, the features are selected in the training process of the learning model, and the feature selection result can be automatically seen as output while the training process is finished. Here, the features are selected during the learning process, which is different from previous methods where FS is performed at the pre-processing level. Since FS and model learning are performed at the same time, and the features are selected during the training phase of the model, these methods are far more efficient than the wrapper methods, since they do not need to evaluate feature sets iteratively. This method avoids the training of the model each time a new feature is added. This method uses ensemble learning and hybrid learning methods for feature selection. The most widely used embedded methods are the regularization models, which involve adding a penalty term to the loss function during model training and encourages the model to select only the most important features. The penalty term can be based on L1 (LASSO) or L2 (Ridge) regularization and is used to constrain the weights of the features. Features with low weights are effectively ignored by the model, while features with high weights are considered important for making predictions. Afterwards, both the regularization model and selected feature sets are returned as the final results. In literature, a fourth category is

sometimes discussed as hybrid methods, which can be regarded as a combination of multiple feature selection algorithms (i.e., wrapper, filter, and embedded). The main goal of this method is to tackle the instability and perturbation issues of many existing feature selection algorithms. For instance, for small-sized high-dimensional data, a small perturbation in the training data may result in totally different feature selection results. By aggregating multiple selected feature subsets from different methods together, the results are more robust and it enhances the credibility of the selected features [53] [28].

There are different types of feature importance methods, one distinction is whether the feature importance measures model-specific or model-agnostic. Model-specific feature importance methods are designed to interpret the importance of features within the context of a specific machine learning model. Examples include decision tree-based models such as Random Forests, Gradient Boosting, where the Gini importance or the decrease in impurity for each feature is model-specific. In contrast, model-agnostic tools can be used on any machine learning model, no matter how complicated. Examples include Permutation Feature Importance, SHAP (SHapley Additive exPlanations), and LASSO (L1 regularization).

Another distinction is global vs local feature importance. Local measures focus on the contribution of features for a specific prediction, whereas global measures take all predictions into account, that is, it gives a score for each feature over the entire dataset. For instance, if we want to know as why a particular person is rejected from granting a loan, we would go for local measures specific to that person but not general to whole data distribution. In the following subsections, we briefly discuss feature importance methods.

Ensemble tree specific feature importance:
It is a local model-specific method and an embedded technique. If we consider the random forest algorithm, the feature importance measure is computed by taking the average reduction in impurity across all trees in the forest due to each feature. That is, features that tend to split nodes closer to the root of a tree will result in a larger importance value. This measure is the Gini impurity or the information gain/entropy for classification problems. For regression the measure of impurity is variance. It is also called as mean decrease impurity. Therefore, when training a tree, it is possible to compute how much each feature decreases the impurity. The more a feature decreases the impurity, the more important the feature is. The advantage is that all needed values are computed during the training, saves some computation time. The drawback is that it has a tendency to select features with high cardinality. Moreover, in the case of correlated features, it can select one of the features and neglects the importance of the second one which can lead to wrong conclusions [39].

Permuted Feature Importance:
The permutation based feature importance can be used to overcome drawbacks of feature importance computed with mean impurity decrease. It is a model-agnostic approach based on mean decrease in the accuracy and can be applied to any model. The idea is that after evaluating the performance of the model, you permute the values of a feature of interest and re-evaluate model performance. The observed mean decrease in the performance of the model indicates the feature importance and in general, the features which impact the performance the most are the most important ones'. This performance decrease can be compared on the test set as well as the training set. The latter says about generalizable feature importance. This method is less likely to select two correlated top features. It more accurately reflects the added value of a feature given the presence of all other possible correlated features in the model. [10]

SHAP value approach:
The Shapley value is a game-theory approach designed to have a fair payout scheme. It is a

method to assign payouts to players depending on their contribution to the total payout. In other words, the Shapley value is the average marginal contribution of a feature value for all permutations at a local level. It is based on four properties *Efficiency* - the sum of the Shapley values of all players should sum up to the total payout, *Symmetry* - two players should get the same payout if they add the same value in all team combinations, *Dummy* - a player gets a Shapley value of zero if it never improves a subteams performance when it is added, and *Additivity* - if we have two game bonuses, the combined Shapley value of a player across these two games is the sum of the individual game's Shapley values. For instance, if a random forest is trained, the prediction is an average of many decision trees. The Additivity property guarantees that for a feature value, you can calculate the Shapley value for each tree individually, average them, and get the Shapley value for the feature value for the random forest. These properties together are considered as a definition of a fair payout. Shapley value determines how important (value) each feature (player) is for predicting a target variable (game). Let N_f be the number of features, then the Shapley value of feature i is defined by

$$\phi(v) \ = \sum_{S \subseteq \{1,...,f\} \setminus \{i\}} \frac{|S|!.(N_f - |S| - 1)!}{N_f!}.(v(S \cup \{i\}) - v(S))$$

where *v(S)* is the worth of the coalition *S*. In words, it is explained as follows, for every possible sequence of features up to feature i, the added value of feature i is the difference between the worth before it was included that is *v(S)* and after that is $v(S \cup \{i\})$, that is total payout. Averaging these added values over all possible sequences of features gives the final Shapley value for feature i.

There are many existing feature importance methods that use Shapley values. One of these techniques is the SHAP (SHapley Additive exPlanations) method, which is used to explain how each feature affects the model, and allows both local and global analysis for the data set. To measure the local FI for a specific sample x and a prediction model m, the conditional expectation is used as a characteristic function that is v. Let $x = (x_1, ..., x_{N_f})$, where x_i is the feature value of feature i, then SHAP FI values can be determined as follows:

$$v_x(S) \ := \ \mathbb{E}_z[f(z)|z_i = x_i \ \forall \ i \in S \ where \ z = (z_1..., Z_{N_f})]$$

Here, function V_x is defined locally for each feature x. As this method is additive, the mean value of these marginal contributions is used as a global feature importance [31][32][45].

SHAP FI is an alternative to permutation feature importance. The difference between both importance measures is that permutation feature importance is based on the decrease in model performance. SHAP is based on the magnitude of feature contributions. The advantages of this method are that, it has a theoretical foundation in game theory and prediction is fairly distributed among the feature values. It can be used for both local and global explanations. SHAP has a fast implementation for tree-based models, which is key for its popularity because otherwise, the adoption of Shapley values as a global method is a slow computation as it requires computations for a lot of instances [38].

3 Research Questions and Relevance

The main focus of this thesis is to *estimate the distribution of SOH and RUL* of electric vehicle takeoff and landing ($eVTOL$) batteries using Machine Learning algorithms. The aim is to quantify the uncertainty by predicting a distribution mean (μ) and an associated standard deviation (σ).

The following are the questions that follow the main thesis question:

- Which features of the dataset are to be considered for the prognostics? That is one should be able to extract information about what features are important and how features interact to create powerful information.
- How to quantify feature importance so that the model is interpretable? In other words, one should have the ability to extract information about specific predictions to validate and justify as to why the model produced a certain result.
- How to estimate the distribution of SOH and RUL of (eVTOL) batteries, that is which models/methods can be implemented and are applicable?
- What are the different model performance criteria that can be used to compare different model predictions of SOH and RUL implemented in this study? In other words, how to evaluate the predictions of the model?

The relevance of above the research questions in the context of uncertainty is to understand what generates this uncertainty, how to quantify it and reduce it (if possible and/or needed), so that accurate informed decisions can be made. By feature selection and feature quantification techniques (in terms of a scoring metric that helps in identifying the relevant features), important features are selected, which in turn helps to build a robust model by reducing sensitivity to irrelevant or noisy variables. Noisy features can introduce unreliable uncertainty estimates and mislead the model, and their exclusion can lead to more reliable and stable predictions. Moreover, selecting a subset of features can make the model more interpretable by facilitating a better understanding of the model's behavior and the factors influencing predictions. It's easier to understand and explain the relationship between a few important features and the target variable, rather than trying to interpret a model with many features. The third and the fourth sub-questions are directly related to the methodology and the evaluation of the distribution estimation of SOH and RUL of eVTOL batteries.

The relevance to society is that, as discussed in the introduction eVTOLS are seen as a solution to the growing traffic congestion in large cities, the traffic-related pollution and the inter-city connectivity needs. Moreover, the health management of the batteries remains one of the challenges in the eVTOL industry and it is crucial to study its behaviour for safety purposes and to take informed decisions. In order to take informative decisions, such as for scheduling and planning, maintenance and replacement procedures for such systems, it is important to quantify the uncertainty associated with the estimated SOH and RUL values. SOH and RUL are considered to be two significant health indicators of Li-ion batteries. Moreover, in contrast to batteries used in on-ground EVs that are subjected to CC-CV charging protocol with a constant discharge C-rates, eVTOLs require larger discharge rates during the take-off and landing, which makes this study important and interesting. Also, the application of machine learning and deep learning techniques for SOH and RUL estimation is an active area of research. The lack of uncertainty quantification (UQ) in most of the current state-of-the-art machine learning methods is a major drawback for their adoption in critical decision making systems, which is also the case in prognostics of eVTOL batteries. This is especially true because of the black-box nature of deep learning models, whose predictions are often overconfident and sometimes difficult to interpret.

4 Data Description

The data-set [7] (Verison 1, Accessed on 26th February, 2023) used in this thesis is developed by Carnegie Mellon University. The authors experimentally cycled twenty-two lithium-ion batteries over several charge and discharge cycles in order to mimic the expected duty cycle of an electric aircraft. To the best of our knowledge, it is the first battery data-set generated specifically for eVTOLs. The authors performed laboratory experiments using Sony-Murata Model 18650

VTC-6 cells, that has rated capacity of 3000mAh at a nominal voltage of 3.6V. These cells are suitable for evaluation of eVTOL applications as it can sustain high power demand while providing a cell specific energy of 230 Wh/kg [6].

All battery cells are tested in an Arbin 200A cylindrical cell holder paired with a BioLogic BCS-815 modular battery cycler. Cells with a specified temperature are placed in a temperature chamber that is maintained at that specified temperature. Cell temperature is measured at the surface of the cell via a thermocouple fixed to the cell body with aluminum tape [19][6]. These batteries are used to perform short-range missions with the Vahana eVTOL. Vahana is an eVTOL designed by Acubed (Airbus) for urban air mobility. It is an all-electric, single-seat, tilt-wing vehicle with a range of 50 km. It achieves an average speed of 190 km/h during cruise and with a maximum speed of 220 km/h [2]. This eVTOL is designed for a total capacity of 5 persons with a range of 400 km [3].

4.1 Generic mission profile pattern

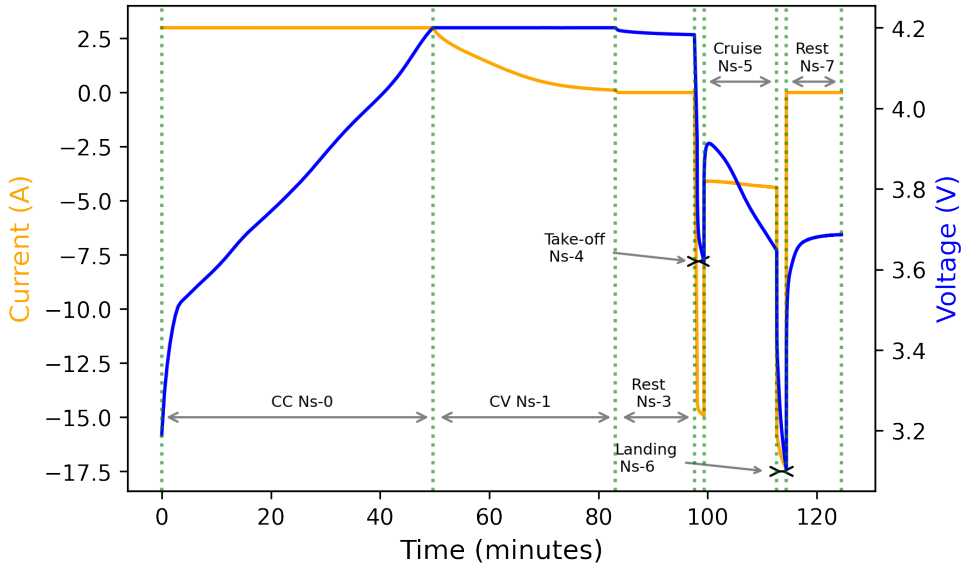
Two main aspects of this data-set compared to other battery data sets are, one, the batteries subject to different C-rates during the discharge phase of a flight. To be more specific, the take-off and landing of eVTOLs are performed at a larger C-rate than the cruise phase, which is expected from an eVTOL mission. Note that, C-rate is defined as a measure of the rate at which a battery is discharged relative to its maximum capacity. Two, each mission profile is parameterized to test different conditions such as varying the discharge power during flight, charge current, charge voltage, ambient temperature, and cruise length, which could be encountered during a mission of an eVTOL. These parameters are varied in one at a time, generating twenty-two different mission profiles, refer Table 1.

A mission profile is defined as a set of flights/cycles whose parameter specifications are kept fixed across its flights. According to the data files, all the mission profiles follow the same generic format in all cases as below:

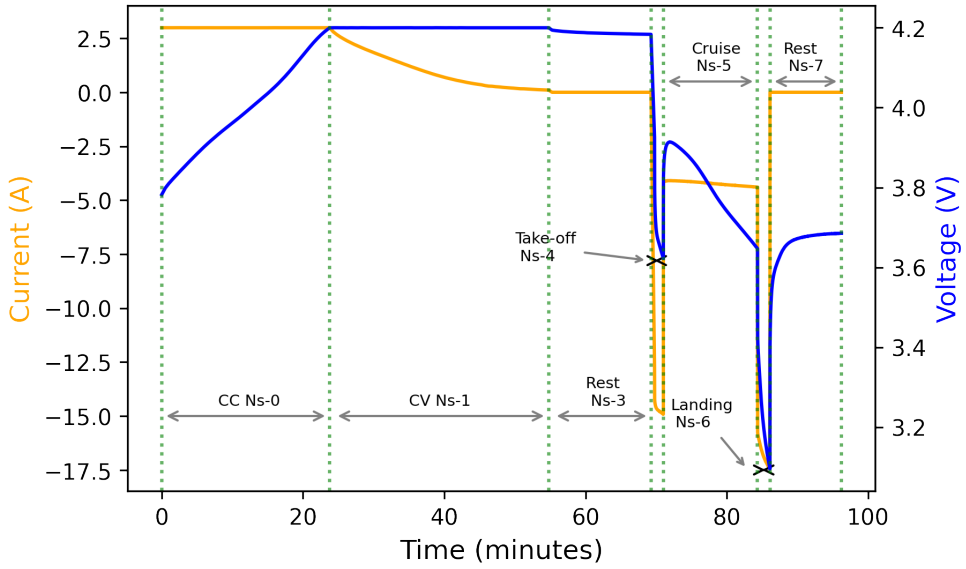
1. **Charge Phase**: the cell is charged using a constant current-constant voltage (CC-CV) charging protocol. This constitutes of mission/cycle segments (Ns variable of the data-set) *0* and *1* in the data-set that indicates *CC* battery Charging phase and *CV* battery Charging phase respectively.

2. **Discharge Phase**: The discharge phase has three segments. It starts with the *take-off* of the eVTOL, where the cell is discharged at a high constant power for a period. Followed by the *cruise* phase, where the cell is discharged at a lower constant power compared to take-off and landing. This phase has a longer period of flight duration, which ranges from 400 seconds to 1000 seconds. The discharge phase ends with an eVTOL *landing*, the cell is discharged at high constant power (same rate as take-off) for a slightly longer duration of time. These relate to mission/cycle segment (Ns variable of the data-set) *4, 5, 6* in the data set respectively.

3. **Rest**: The batteries are allowed to rest between two charge and discharge cycles of a mission. After charge phase, the battery cell rests until the cell temperature reaches 35 °C. After discharge phase, the battery rests until its temperature decreases to 27 °C or for at least 15 minutes. These rest periods (after charge and after discharge) relate to mission/cycle segment (Ns variable of the data-set) *3, 7* in the data set respectively. Sub-figures 1a and 1b of Figure 1 show generic mission profile pattern of charging, discharging and rest phases of Vahana 1. Note that '-' in 'Ns-segment number' indicates a hyphen/dash, this applicable to all plots that contains '-' unless explicitly specified.



(a) Mission 2 (capacity test-1) of VAH01



(b) Mission 3 of VAH01

Figure 1: Charging and discharging of Mission 2 and 3 of VAH01

From 1a, it can be noticed that the CC charging phase has a duration of 50 min, and the battery is charged from 3.2 V to 4.2 V with constant current of 3.0 A. Followed by the CV charging phase, that takes place with 4.2 V for 33 min. The Rest period immediately after charging has a duration of 14 min. Then, the discharge phase starts with the take-off of the eVTOL with 75 s of duration and voltage drop from 3.92 V to 3.62 V. Followed by the cruise phase with a duration of 800 s and the landing phase with 105 s of duration. During landing, the voltage drops from 3.57 V to 3.1 V. The mission ends with a rest period of 605 s, where the battery reaches a temperature of 27.3 °C. 1b is the 3^{rd} mission that immediately follows the 1^{st} capacity test (Mission 2) of VAH01. It can be seen that it follows same mission profile pattern as in 1a. However, the battery is now charged from 3.8 V to 4.2 V and has a duration of 55 min. Apparently, it has a shorter duration of the CC and CV charging period (by 28 min) compared to mission 2.

4.2 Baseline mission profiles

The mission profiles are named as Vahana (VAH01-VAH30). Table 1 summarizes the mission profiles with the parameter that is changed for that mission profile, while having the rest of the parameters same as the baseline mission profile. The baseline mission profiles (VAH01,VAH17,VAH27) have the following charge-discharge protocol, during the CC charging phase, the battery is charged with 1 C-rate and the phase ends when the battery's voltage reaches 4.2 V. Followed by CV charging phase which starts with a constant voltage of 4.2 V until the current is less than C/30. After charging phase, the battery cell is allowed to rest until the cell temperature reaches 35 °C. After this Rest period, the discharge phase starts, i.e. an eVTOL performs a flight. During the discharge phase of an eVTOL, take-off has a duration of 75 s, with a discharge power of 54W, 5 C-rate, and 1.12Wh discharge energy. Followed by the cruise phase, which has a duration of 800 s, with a discharge power of 16 W, 1.48 C-rate, and 3.55 Wh discharge energy. Later, the landing takes place with a duration of 105 s, a discharge power of 54 W, 5 C-rate, and 1.57 Wh discharge energy. Finally, the battery is allowed to rest until its temperature reaches 27 °C.

Table 1: Mission profile characteristics according to [7] (Verison 1, Accessed on 26th February, 2023)

Baseline Mission Profiles	**VAH01, VAH17, VAH27**
Mission Profiles with adjusted parameters	
Extended cruise length - 1000 sec	VAH02, VAH15, VAH22
10% power reduction during discharge (flight)	VAH05, VAH28
CC charge current reduced to C/2	VAH06, VAH24
CV charge voltage reduced to 4.0V	VAH07
Thermal chamber temperature of 20°C	VAH09, VAH25
Thermal chamber temperature of 30°C	VAH10
20% power reduction during discharge (flight)	VAH11
Short cruise length - 400 sec	VAH12
Short cruise length - 600 sec	VAH13, VAH26
CC charge current reduced to 1.5C	VAH16, VAH20
CV charge voltage reduced to 4.1V	VAH23
Thermal chamber temperature of 35°C	VAH30

4.3 Measurements

The raw cell data are collected in multiple files over the lifetime of the cell and are concatenated for each cell. The data is made available in comma-separated values (CSV) format files. There are 10 variables for each file, namely, *time_s*: Time since beginning of experiment in seconds; *Ecell_V*: Terminal voltage; *I_mA*: Cell current in milliamperes, where positive values represents charge and negative values represents discharge; *EnergyCharge_W_h*: The amount of energy supplied to the cell during charge in watt-hours; *QCharge_mA_h*: The amount of charge supplied to the cell during charge in milliampere-hours; *EnergyDischarge_W_h*: The amount of energy extracted from the cell during discharge in watt-hours; *QDischarge_mA_h*: The amount of charge extracted from the cell during discharge in milliampere-hours; *Temperature__C*: The cell surface temperature measured in °C; *cycleNumber*: Cycle number as recorded by the cell tester; *Ns*: Cycle segment [7].

5 Data Cleaning

5.1 Implementation

The code is implemented using Python 3.8 and by using Jupyter Notebooks with libraries scipy, numpy, pandas, matplotlib and scikit-learn, keras-tensorflow. The computations are carried out using an 64-bit operating system Intel(R) Core(TM) i5-9300H CPU, clocked at 2.40 GHz, and 8GB of RAM.

5.2 Defining and Identifying Capacity Tests

In this section the process of identifying capacity tests is discussed. At first, mission/cycle numbers are assigned to each mission profile as cycle numbers is not accurate in the original raw data-set (also mentioned by the authors of [7]). In the present approach, mission number is incremented by one when Ns (cycle segment) hit zero.

It is observed that a generic mission has cycle segments [0, 1, 3, 4, 5, 6, 7] which correspond to [cc charge phase, cv charge phase, rest after charge phase, take-off of eVTOL, cruise of eVTOL, landing of eVTOL , rest after discharge phase] respectively.

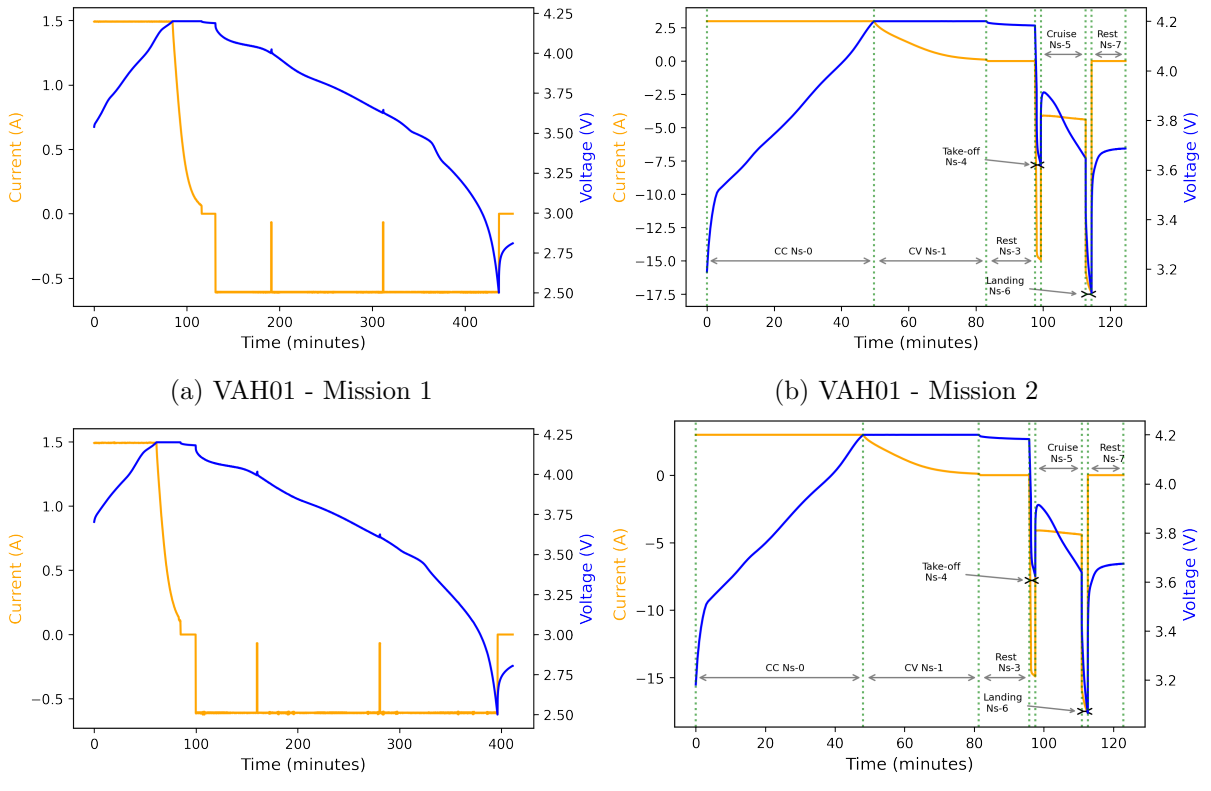
The data is further cleaned by removing missions which do not have entire sequence mission cycle segments [0,1,3,4,5,6,7] and mission numbers are recounted, so that the total number of missions for a profile is not effected (which is later required for RUL calculation). Refer Table 2 column name '#Mission' for number of missions before and after cleaning data of each mission profile.

Table 2: The total number of Missions and capacity tests for each mission profile before and after cleaning

	Profile Name	**#Missions**		**#Capacity Tests**	
		Before Cleaning	After Cleaning	Before Cleaning	After Cleaning
1	VAH01	847	847	17	17
2	VAH02	625	625	13	13
3	VAH05	2232	1544	30	29
4	VAH06	12372	960	20	19
5	VAH07	1144	273	6	5
6	VAH09	12334	805	22	21
7	VAH10	1431	1428	28	28
8	VAH11	2249	2247	44	44
9	VAH12	2349	2349	46	46
10	VAH13	1041	1041	20	20
11	VAH15	554	553	11	11
12	VAH16	559	558	11	11
13	VAH17	1002	1002	20	20
14	VAH20	611	611	12	12
15	VAH22	1294	550	11	10
16	VAH23	697	697	14	14
17	VAH24	801	801	16	16
18	VAH25	554	554	12	11
19	VAH26	1163	1162	23	21
20	VAH27	587	586	12	11
21	VAH28	1182	1180	24	23
22	VAH30	919	919	18	18

In the second step, capacity tests are identified as follows, after every 50 cycles/missions

i.e. at 51^{st} mission/cycle, the authors of [7] performed capacity tests wherein the battery cell is discharged at constant current (C/5) until its voltage drops below 2.5 V, i.e. battery charge is reduced to 0% SOC. This mission is identified to have cycle segments 1,2,3,4,5,6,7,8 in the data set. In the following mission, the battery cell is charged to 100% SOC at a charging rate of 1 C-rate (except for vahana 6,16,20,24) and a constant voltage of 4.2 V (like every other battery cell cycle with cycle segments (Ns) 0,1,3,4,5,6,7, except for vahana 23 and 7) and then eVTOL performs a flight (take-off, cruise, and landing). This special mission where the battery is charged to 100% SOC from 0% SOC and then flight is performed, is referred to as a capacity test. However, it is observed that all battery cells of mission profiles did not undergo a capacity test exactly after every 50 cycles. For example, in VAH09 capacity test 5-11 found to have taken place after every 11 cycles and 12th and 21st capacity test after 23 and 57 cycles respectively, refer 'VAH09' of Figure 19. Moreover, the mission profiles that underwent a capacity test after every 50 cycles, for instance VAH01, VAH02, VAH13, VAH15, VAH16, VAH30, VAH24, VAH20, VAH17, VAH22, VAH23, VAH27 followed a specific pattern as in Figure 2. That is the special mission which has cycle segments 0,1,3,4,5,6,7 (2b, 2d) is preceded by a mission whose cycle segments are 1,2,3,4,5,6,7,8 (2a, 2c). Therefore, in the present approach, the mission with cycle segments 0,1,3,4,5,6,7 that immediately followed a mission with cycle segments 1,2,3,4,5,6,7,8 is considered as a capacity test. Refer Table 2 column name '#Capacity Tests' for number of capacity tests identified before analysing the selected capacity tests.



(a) VAH01 - Mission 1

(b) VAH01 - Mission 2

(c) VAH01 - Mission 51

(d) VAH01 - Mission 52

Figure 2: Cycle segment pattern to identify capacity tests of a mission profile

5.3 Selection of capacity tests and vahanas for distribution estimation

To estimate the distribution of SOH and RUL of eVTOL batteries, we consider the missions and the capacity tests after cleaning the data Table 2. A few capacity tests have been removed because either they do not adhere to capacity test definition or do not follow the mission profile specific parameter conditions, Figure 20a, 20b show consolidated battery capacity (%) degrada-

tion trend of each mission profile during the capacity tests before and after cleaning the data. This may happen due to sensor measurement errors or malfunctions while recording the measurements. Moreover, we consider 21 mission profiles of the total 22 profiles. Mission profile VAH07 (see Figure 3) is not considered as it has not reached EOL threshold of 85% (see section 6.1 for SOH and RUL definitions).

In the mission profiles, VAH05, VAH07, VAH22, identified last capacity test is observed to have a QCharge greater than 0 and starting cell temperature greater than 25 °C. These conditions ideally should be QCharge equal to 0 and starting cell temperature less than or equal to 25 °C. Therefore, last capacity test is removed from the respective mission profiles. For VAH06, identified 16^{th} capacity test mission (cycle number 766/767 before cleaning) did not have entire sequence mission cycle segments [0,1,3,4,5,6,7] and therefore is removed.

It is observed that for the mission profiles VAH25 and VAH28, two capacity tests are run in a sequence. For VAH25, the consecutive capacity tests are found to run at cycles 461 and 462. For VAH28, these are found to run at cycles 256 and 257. This is also mentioned by the authors of [7] and can be seen in the plots of VAH25, VAH28 under Figure 19. In the current approach, only one capacity test is considered which has segments [0,1,3,4,5,6,7]. For VAH09, the 12^{th} capacity test (mission 262) has a battery capacity of 2.156 Ah, while the following 13^{th} capacity test mission has a battery capacity of 2.689 Ah. This is unusual, as battery capacity is expected to decrease over time. However, at 12^{th} capacity test, it is noticed that the starting cell temperature is 24.43 °C, which should be around 20 °C for VAH09. Hence, the capacity test is removed.

The capacity test 6, 11 are removed from VAH26 and the capacity test 6 is removed from VAH27. These are removed because, it is observed that the battery cell is not properly charged during CC-CV charging protocol, resulting in decreased battery capacity followed by capacity tests which have higher battery capacity.

6 Methodology

6.1 Defining the SOH and RUL of eVTOL batteries

The state of health (SOH) of a battery refers to its overall condition and performance level relative to its original design and specifications. It is a measure of how well a battery can store and deliver electrical energy compared to when it was first manufactured. In our context, the SOH of a mission profile is defined as the ratio of the maximum charge supplied during a capacity test of a mission profile to its rated capacity (maximum battery charge supplied during first capacity test of that mission profile) [37]. SOH is expressed as a percentage as below:

$$SOH^{M,c} = \frac{max(QCharge^{M,c})}{max(QCharge^{m,1})}.100\%$$

where, $QCharge^{M,c}$ is the maximum measured capacity during a capacity test of a mission profile, and $QCharge^{M,1}$ is the maximum battery capacity measured during the first capacity test (c=1) of the mission profile. This means for every Vahana, the SOH at first capacity test is always 100%.

The remaining useful lifetime (RUL) of a battery is defined as the remaining number of missions/cycles for this battery until the End of Life (EOL), given that the battery has been used for $m >= 1$ missions. Formally, the RUL of a battery, estimated after m missions under mission profile M, is defined as:

$$RUL^{M,m} = T^{M}_{EOL} - T^{M,m}$$

where, $T^{M,m}$ is the current mission/cycle number under mission profile M, and T^{M}_{EOL} is the mission/cycle number when the battery capacity drops for the first time below an EOL-threshold

under mission profile M [37]. RUL is expressed as number of cycles.

Existing studies based on experimental battery measurements set the EOL-threshold to 80% of the nominal battery capacity, however EOL-thresholds for eVTOL batteries have not yet been formally established. For eVTOL batteries, it is expected that conservative safety margins will be considered and a conservative EOL-threshold of 85% of a nominal battery capacity of an eVTOL is considered [37]. In this thesis, we also consider an EOL-threshold of 85% of the initially measured battery capacity.

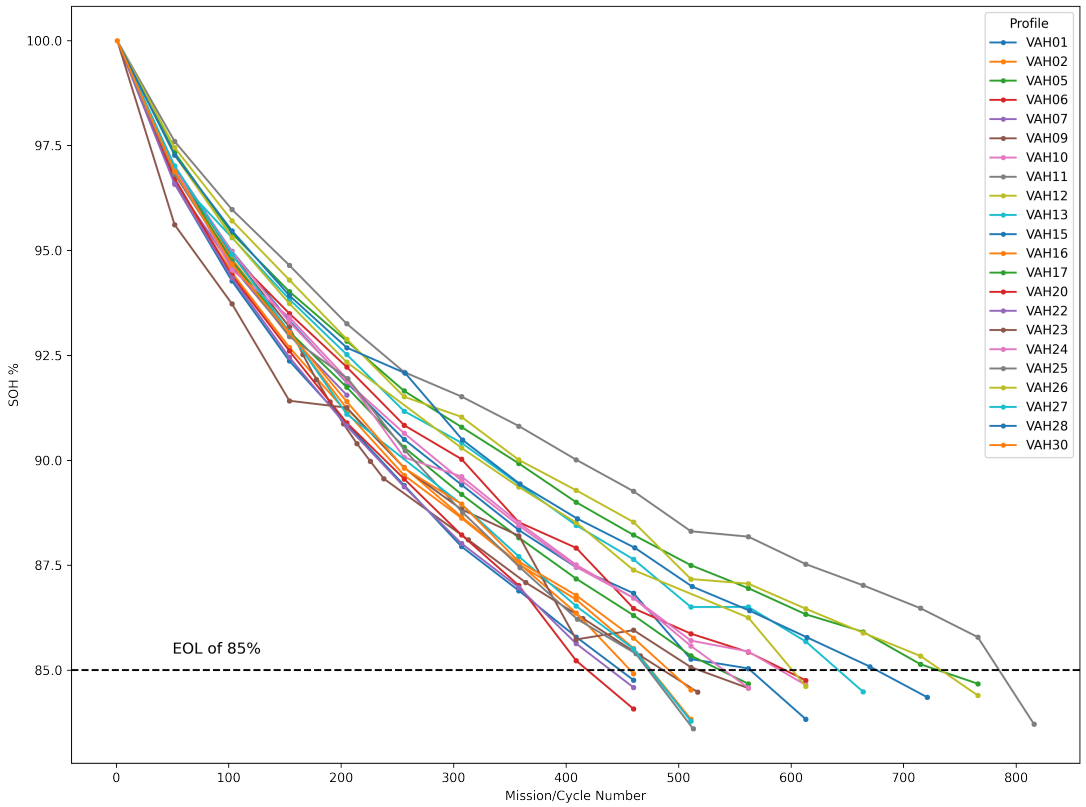


Figure 3: SOH (%) of capacity test missions vs. mission cyles with EOL=85%

The choice of the EOL-threshold has also an effect on the selection of the mission profiles. Using an EOL-threshold of 80%, not all mission profiles in the dataset will have their batteries reaching EOL. Specifically, for mission profiles VAH01, VAH02, VAH15, VAH16, VAH20, VAH22, VAH23, VAH24, VAH25, VAH27, and VAH28 the measurements stop before the battery capacity reaches 80% of the initially measured battery capacity Figure 20b. In other words, by using an EOL-threshold of 80% of the initial battery capacity, these mission profiles will not have run-to-EOL series of measurements [37]. Considering an EOL-threshold of 85% of the initially measured battery capacity, all mission profiles in dataset (except for VAH07 (see Figure 3)) reach their EOL . Table 3 shows the number of missions until each battery reaches its EOL, and the number of capacity tests until these batteries reach their EOL.

Table 3: Mission profiles and capacity tests, EOL-threshold 85% of initially measured battery capacity

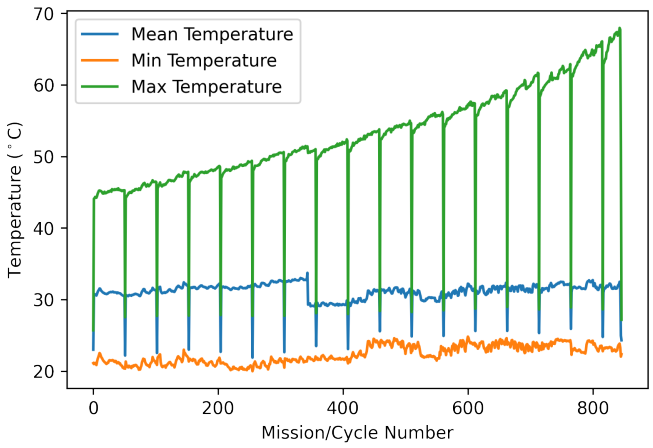
	Profile	Missions until EOL is reached	Capacity tests until EOL is reached
1	VAH01	613	13
2	VAH02	511	11
3	VAH05	766	16
4	VAH06	613	13
5	VAH09	517	16
6	VAH10	613	13
7	VAH11	816	17
8	VAH12	766	16
9	VAH13	664	14
10	VAH15	460	10
11	VAH16	460	10
12	VAH17	562	12
13	VAH20	460	10
14	VAH22	460	10
15	VAH23	562	12
16	VAH24	562	12
17	VAH25	513	11
18	VAH26	613	11
19	VAH27	511	10
20	VAH28	721	15
21	VAH30	511	11

6.2 Feature extraction for SOH and RUL distribution estimation

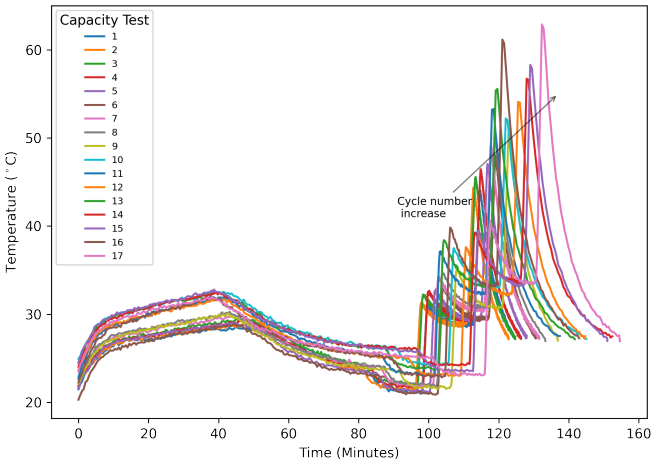
We have incorporated features akin to those presented in [37] in the current work. We explain the features as charge phase and discharge phase (as discussed in section 4.1) related features.

Charge Phase related features: From Figure 1a,1b and Figure 4c, it can be noticed that as the number of missions increases, the duration of the CC charging phase decreases. As a result, the duration of CV charging phase increases. Analyzing the charging phase of the battery, we consider features **CC duration** and **CV duration** of the capacity tests, i.e. the charge time period of mission cycle segments(Ns) 0 and 1, respectively.

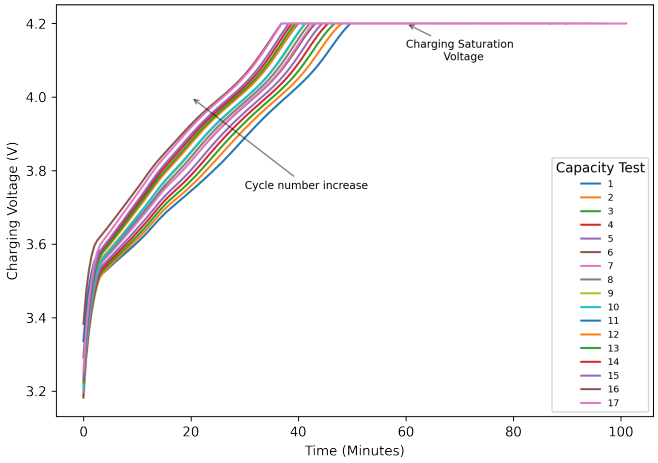
Discharge Phase related features: In the discharge phase. there are three cycle segments (Ns - 3,5,6) namely, take-off, cruise and landing of an eVTOL respectively. For each of these segments, we consider the voltage, the discharge capacity and the temperature related features. Figure 4d shows the discharge capacity vs. the discharge voltage during the capacity tests of baseline mission profile VAH01. It can be observed that as the number of missions increase, the minimum discharge voltage decreases during each flight discharge phase. Also, during take-off and landing, the discharge voltage drop is higher than during cruise. This is expected since the take-off and landing are performed at a higher C-rate. Since the discharge voltage varies for every flight discharge phase and across capacity tests, in order to capture the impact of these variations we consider voltage-related features. Therefore, we consider **Minimum, Maximum, Mean and Variance voltage for each discharge phase** i.e. take-off, cruise and landing of each capacity test of each mission profile. Moreover, voltage-related features reflect the open circuit voltage and internal resistance, which are closely related to the remaining capacity and the aging of the battery [37].



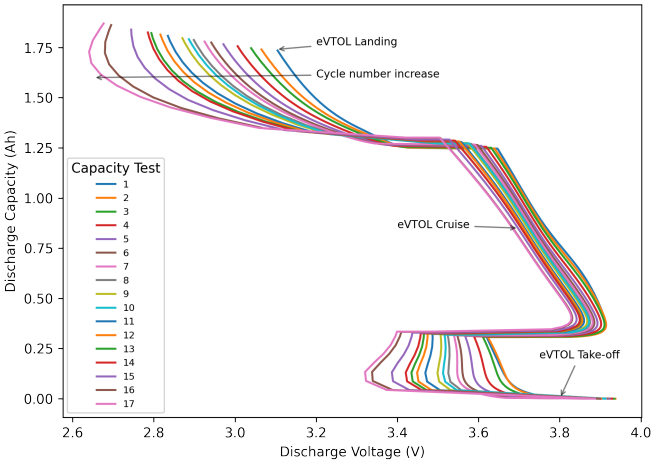
(a) Mean, Min, Max battery surface temperature of missions of VAH01



(b) Battery surface temperature of capacity tests of VAH01



(c) Charge duration of capacity tests of VAH01



(d) Discharge capacity of capacity tests of VAH01

Figure 4: Feature extraction plots for SOH and RUL distribution estimation

Similarly, for discharge capacity related features, we consider **Minimum, Maximum, Mean and Variance discharge capacity for each discharge phase** $\in$ {take-off, landing, cruise} of each capacity test of each mission profile. Figure 4d also shows that the discharge capacity increases from take-off to cruise to landing. The discharge capacity also increases as the number of missions increase. To capture these patterns, we considered discharge capacity related features. Additionally, the discharge capacity and its variation reflect the load characteristics of the battery, which directly impacts the aging of the battery [37].

Figure 4a shows the maximum, minimum, and mean battery surface temperature recorded during all cycles of mission profile VAH01, with a total of 17 capacity tests. We can see that at every capacity test, the maximum and average temperature drops abruptly. This is because the battery cell is allowed to rest at the end of a capacity test until the temperature reduces to 27°C. It can also noticed that as the mission cycles increase, its maximum temperature increases. Moreover, during a capacity test (see Figure 4b) the highest cell surface temperature is observed to reach during the landing phase of an eVTOL. The Figure 4b also shows that the temperature reaches a peak during take-off phase and then decreases during cruise phase. Moreover, when considering all the three discharge phases, the temperature increases as the number of capacity tests increase. To capture the change in battery surface temperature as more missions are performed, we consider **Maximum temperature of each discharge phase** $\in$ {take-off, landing, cruise} of capacity test of each mission profile.

Since the cruise duration varies across mission profiles, we also considered the **duration of the cruise phase** as a feature. However, in contrast to [37] the features- duration of rest period after CC-CV charging phase (840 seconds), take-off duration (75 seconds) and landing duration (105 seconds) of an eVTOL are not considered as they are constant across the capacity tests and the mission profiles. Overall, we have a total of 30 features related to charge and discharge phases.

6.3 Feature Selection and Feature Importance Quantification

Following feature extraction, the subsequent stage involves feature selection, with the idea to make the process of prediction in machine learning more accurate. This is achieved by removing irrelevant and redundant features, which otherwise would lead to a training model that can be both ineffective and inefficient. That is, models might give a lot of importance to those features based on spurious patterns detected in the training data-set, which leads to over-fitting and poor performance on test data. Moreover, when training a model, run-time, memory usage, and CPU usage grow with the number of features. By not removing redundant features and non-informative features make the model training process unnecessarily longer and costlier [49]. In the current context, though the aim is to provide uncertainity associated with predicted estimates. However, it is still important to have the predicted estimates accurate, so that the model has a better chance of quantifying uncertainty of the estimates that align with the true variability in the data.

We use Boruta-Shap feature selection algorithm, as a wrapper around Random forest regression. The algorithm is an extension of the idea introduced by the "Party On" paper [51] which determines feature importance by comparing the relevance of the real features to that of the randomised features. The choice of Random Forest Regressor (RFR) as a base model for selecting relevant features is due to it's ability to deal with unscaled data and are robust to outliers [48]. Additionally, in all most all feature selecting algorithms (see 2.3), require a human to make an arbitrary decision. That is the choice of features (in terms of number or by using threshold variance) to use in the model is left to the user, this might in-turn become another source of uncertainty introduced in the predictions. In order to enhance the realism of model predictions and to ensure accurate uncertainty estimates while mitigating the uncertainties arising from human intervention, Boruta-Shap algorithm is considered. The working of the Boruta algorithm is as follows:

First, Boruta randomly permutes rows of each feature column to construct its "shadow" version of features. This is done to effectively destroy any potential association between the feature, the target, and other features, thereby making them as irrelevant features to the task being performed. Second, the random forest regressor is trained on the whole feature set, including the new shadow features. Then the feature importance (Z-scores) is computed based on metric of choice. We use SHAP feature importance to evaluate the contribution of each variable to the prediction. Third, the maximum feature importance among the shadow features serves as a threshold for the original features. The idea is that a feature is useful only if it's capable of doing better than the best randomized feature, i.e. a hit is assigned to any feature that exceeds this threshold. Fourth, to robustify the decision making process, at the end of each run, boruta constructs a two sided T-test of equality by removing all the shadow features and repeating the process [49]. Note that a feature is accepted only if it has an importance greater than that of maximum feature importance of random features a significant number of times, otherwise it is rejected.

The hypothesis test of boruta is as below:
The process is repeated iteratively k times and each of the original features x_i have accumulated a certain number of hits h_i to its name. Boruta models our lack of knowledge about whether a feature is useful or not using the null hypothesis H_0 that there is a 0.5 probability that the

feature will get a point. After k times of repeating the process, the number of hits h_k of any real feature for which H_0 holds (i.e. when we don't know whether the feature is useful or not) will follow the binomial distribution $B(k, 0.5)$, of which the confidence bounds $[hm_k, hM_k]$ are evaluated [49]. We conclude if a feature is accepted or not as follows:

1. $h_k > hM_k$, the null hypothesis H_0 is rejected in favour of alternative hypothesis H_1 as feature is accepted.

2. $h_k < hm_k$, the null hypothesis H_0 is rejected in favour of alternative hypothesis H_2 as feature is rejected.

3. $h_k \in [hm_k, hM_k]$, the algorithm becomes inconclusive (tentative) about the feature status, i.e. we still do not know whether to accept or reject the feature.

In the case of '3', either the number of iterations could be increased until every feature is assigned with either accept or reject status. Since, it is practically not possible to run until every feature is assigned with accept or reject status, we considered to have 100 iterations on random forest regressor with 100 estimators. In case of inconclusive features, they are still considered as important features in the remainder of the steps in this thesis. The original Boruta algorithm uses permutation feature importance as it's metric of choice. However, the combination of Boruta and SHAP has proven to out perform the original Permutation Importance method in both speed, and the quality of the feature subset produced. This algorithm not only provides a better subset of features, but it can also simultaneously provide the most accurate and consistent global feature rankings which can be used for model inference too. [50][21]

6.4 Machine Learning Algorithms to estimate the distribution of SOH and RUL

After the feature selection phase, three algorithms are proposed to estimate the distribution of State-Of-Health (SOH) and Remaining Useful Life (RUL): Random Forests Regression (RFR), Convolution Neural Network with Monte Carlo Dropout (CNN-MCD) and Mixture Density Network (MDN). These belong to different classes of machine learning algorithms, one belong to tree-based machine learning algorithms and others belong to the category of neural networks, respectively. Another distinction between the chosen models is that, RFR and CNN-MCD do not explicitly assume a specific output form rather give several point estimates for the a test instance, whereas MDNs parameterize the output as a mixture of several probability distributions (typically Gaussian distributions).

6.4.1 Random Forest Regression (RFR)

Random Forest Regression is a well-established technique and is a part of ensemble learning methods. At its core, it is a decision tree model. Single decision tree, when they grow deep and complex is prone to overfitting and leads to poor generalisation on unseen data. To mitigate these issues random forest constructs multiple decision trees and aggregate their predictions. They introduce randomness during the tree-building process in two main ways: one, each tree is built on a random subset of the training data (sampling with replacement). This process introduces diversity amongst the trees, as each tree sees a slightly different subset of the data. Two, at each node of the tree, instead of considering all features to make a split, random forests consider only a random subset of features. This helps in reducing the correlation between trees and prevents one dominant feature from being consistently chosen as the top split. During prediction, the individual predictions from each tree are averaged (for regression) to produce the final output. In order to estimate the distribution of SOH and RUL of a mission profile during the test phase, the output values of decision trees are not averaged, but collected. The mean and standard deviation (σ) for each test instance is calculated from the predictions of individual trees of the

random forest. This process allows us to obtain the mean prediction across all trees, indicating the central tendency of the predictions, and the standard deviation, provides a measure of the spread or uncertainty in the predictions for that particular test instance. [60]

6.4.2 Convolution Neural Network- Monte Carlo Dropout (CNN-MCD)

Convolutional Neural Networks (CNNs) have gained significant popularity and are widely used in various applications. Inspired by the CNNs usage in various domains and applications for probabilistic predictions [36], [27], [60], [12], we implement it in this thesis as well for SOH and RUL distribution prediction. A one-dimensional convolutional neural network (1D CNN), is primarily used for processing and extracting features from sequential data, such as time series, audio signals, or text [23]. The layers in a CNN apply a set of filters (kernels) across the sequential input to detect local patterns. Each filter slides over the input and performs element-wise multiplications and summations, generating feature maps. In 1D CNNs, filters move along one direction (the temporal axis in the case of time series data) to capture temporal patterns. In our scenario, the input to the 1d CNN is [samples,input_features, window_size], where samples indicate the size of the training data, input_features are the obtained features from 6.3 and we apply a window size of 2 time steps. That is for each Vahana, capacity tests are split into window size of two, such that each sample has information regarding its one previous capacity test.

The general network architecture for both SOH and RUL distribution prediction is as follows: we use 3 convolution layers with 64, 32, 16 filters (i.e. depth of the output layer) respectively. Followed by 4 dense layers (last layer being the output layer) with filters 128, 64, 32, 1 respectively. For all convolutional layers zero padding is used to ensure the same size of the output and the chosen kernel size is 2. ReLU activation function is applied to all convolution layers followed by a dropout layer. That is dropout layer is added to each of the network layers (except for input and output layer) for RUL distribution prediction. However, for SOH distribution prediction, we do not apply dropout in the dense layers (that is after flatten layer) and ReLU is thresholded to 100 in all the layers. This is done to avoid overshooting of output above 100% of state of health. Moreover ReLU activation is not applied to the output neuron. Training of the network is done using a fixed learning rate of 0.001 with maximum epochs of 1000 and mean-squared-error (MSE) as the loss function. The weights and bias of the network are optimised using Adam optimiser.

We apply a dropout rate ρ in specified layers of the CNN, except the input layer to avoid the loss of input information. During training, dropout is used to prevent over-fitting. During testing, we perform 1000 forward passes through the neural network for each test sample. During each forward pass, different, randomly selected neurons (ρ percent) are dropped. Thus, a different SOH or RUL prediction is obtained with each forward pass. In [15] (also, see section 2.1), it is shown that a neural network with Monte Carlo dropout approximates a Bayesian neural network representing a deep Gaussian process.

6.4.3 Mixture Density Network (MDN)

Mixture Density Network (MDN) is a type of neural network architecture that is specifically designed to model complex probability distributions. They are introduced by Christopher Bishop in the early 1990s [8], primarily in the context of neural network-based approaches to modeling uncertainty in machine learning. Unlike traditional neural networks that output a single deterministic value, MDNs' primary goal is to model complex output distributions by producing a mixture of probability distributions.

They combine neural network with a Gaussian mixture models (GMMs) to generate outputs that represent a weighted sum of several Gaussian distributions. Each Gaussian component in the mixture corresponds to a different mode of the output distribution. The network architecture

of MDNs involves predicting parameters (such as mean, variance, and mixing coefficients) of multiple Gaussian distributions, enabling the model to capture diverse output behaviors. During training, MDNs learn to estimate these parameters based on the input data. The model learns to adjust these parameters to maximize the likelihood of the observed data. When making predictions using a trained MDN, the most likely value given the input data is estimated by sampling from the learned probability distribution. Formally, the weighted sum of gaussian distributions is as follows:

$$p(y_i|x_i) = \sum_{j=1}^{m} \alpha_j(x_i)\phi_j(y_i|x_i)$$

where, $p(y_i|x_i)$ is the conditional probability of y_i given x_i, m is number of components, $\alpha_j(x_i)$ is the mixing coefficient (weight) associated with the j^{th} component. $\phi_j(y_i|x_i)$ is the probability density function of the j^{th} component distribution expressed as following:

$$\phi_j(y_i|x_i) = \frac{1}{\sqrt{2\pi\sigma_j(x_i)^2}} \exp\left(-\frac{(y_i - \mu_j(x_i))^2}{2\sigma_j(x_i)^2}\right)$$

with $\mu_j(x_i)$ and $\sigma_j(x_i)$ as the mean and standard deviation of the j^{th} gaussian component. The neural network is trained using backpropagation, i.e, the network parameters, the weights and biases of each node are updated using an error function E, which is the Negative Log-Likelihood that the model derived from the output of the current network gives rise to the training data (inverse problem) [8],[60]. Formally,

$$E = \sum_{i=1}^{N_f} \left(-\ln \sum_{j=1}^{m} \alpha_j(x_i)\phi_j(y_i|x_i)\right) = \sum_{i=1}^{N_f} (-\ln p(y_i|x_i))$$

where N_f is the total number of test instances. In the current implementation, the weights use a softmax activation function, and the standard deviations use non-negative exponential linear activation function (nnELU). The means in case of SOH have ReLU activation function thresholded (saturation) to 100 and for RUL, means also have non-negative exponential linear activation function. Moreover, the input features have been scaled using min-max scaler, that shrinks the data to the interval [0,1].

6.5 Hyperparameter tuning

The hyperparameters of the models RFR, CNN-MCD, MDN are selected using grid search. Table 4 shows the considered hyperparameters and its respective range of values doe each model. The hyperparameters leading to lowest mean MSE scores on the traning data are chosen. For Random Forest Regression model, 5 fold cross validation is performed on the training data, whereas for CNN-MCD a maximum of 1000 epochs is used as the stopping criteria. In Mixture Density Network a maximum of 1000 epochs with early stopping criteria set to 50 iterations is used, that is training will stop if negative log-likelihood is no longer decreasing. This is opted in order to reduce over-fitting in case of MDNs. Early stopping is however not used in CNN-MCD because the model layers have dropout rate that already prevent the model from over-fitting. The range values of 'Batch size' for MDN is chosen based on the number of capacity tests in the SOH and RUL datasets. See Table 6 and Table 11 for optimised hyperparameters for each test Vahana of SOH and RUL, respectively.

Table 4: Hyperameter Grid for RFR, CNN-MCD, MDN

Model	Hyperparameter	Range
RFR	Number of Estimators	[100, 300, 500, 750, 850, 1000, 1500, 2000]
	Maximum Features	[0.25, 0.50, 0.75, 1.00]
CNN-MCD	Dropout Rate	[0.2, 0.3, 0.5]
	Batch Size	[4, 8, 10, 12, 14, 16, 20]
MDN	Number of Components	[1, 2, 3, 5, 6]
	Number of Neurons	[50, 100, 150, 200]
	Batch Size (SOH)	[10, 15, 20, 25, 30]
	Batch Size (RUL)	[8, 10, 12, 14, 16]

6.6 Evaluation method

When developing any machine learning model it is important to decide on how to evaluate it's final performance. To get an unbiased estimate of the model's performance, evaluation needs to be done on the data that is not used for training. The simplest way to split the data is to use the train-test split method which randomly partitions the dataset into two subsets called training and test sets, such that a predefined percentage of the entire dataset is in the training set. Then, we train our machine learning model on the training set and evaluate its performance on the test set. This makes sure that the samples used for training are not used for evaluation and vice versa.

Random partitioning has its limitations. One, when the dataset is small which is the case of eVTOL dataset, the method can be prone to high variance. Due to the random partition, the results can be entirely different for different test sets because of uneven distribution of target variable values between the training and testing sets. This can be problematic, especially in regression problems where the goal is to accurately predict continuous values and have reliable uncertainty estimates. Two, random partitioning might not ensure an even representation of different patterns or trends present in the dataset across the training and testing sets. This can be particularly challenging if the dataset has specific characteristics that are important for model learning, which is also the case of eVTOL dataset that has vahanas with varied parameters (for example varied temperature, cruise duration etc.) and we want the model to capture such variations across the mission profiles.

To deal with these issues, leave-one-out-vahana-set (LOOVS) to evaluate the performance of a machine learning models is used in the current study. In general, in this method, machine-learning models are trained n times, where n is the to dataset's size. During each iteration of model training, only one sample is used as a test set while the rest are used to train the models. In our case, n is the number of vahanas i.e. 21. This method specially in our case provides a less biased estimate of the model's performance compared to other resampling methods and uses all the available data for training and testing, making an efficient use of the data-set. It also enable the models to learn better representations. Also, note that we treat each Vahana as a test case, meaning training data-set also changes, therefore we repeat the entire process from feature selection till evaluation for the models inclusive of hyperparameter tuning.

6.7 Performance Metrics for Distribution Estimation of SOH and RUL

In this section, we discuss the metrics used to evaluate the performance of models. The metrics MAE, RMSE, CRPS, Sharpness(Sh(%)) and Standard deviation(σ) are used for performance evaluation. These metrics are computed at three levels namely, capacity test level, Vahana level and model level.

At capacity test level:
In general, MAE is defined as the mean of the total absolute difference between the actual and predicted value. RMSE is defined the square root of the average squared differences between predicted and actual outcomes. It is the standard deviation of the residuals (prediction errors). These metrics at capacity test level for models $I \in \{\text{RFR,CNN-MCD}\}$ are defined as following:

$$MAE_I(c_i) = \frac{1}{n_{ep}} \sum_{j=1}^{n_{ep}} |y_i - \hat{y}_{i,j}| \tag{1}$$

$$RMSE_I(c_i) = \sqrt{\frac{1}{n_{ep}} \sum_{j=1}^{n_{ep}} (y_i - \hat{y}_{i,j})^2} \tag{2}$$

where, n_{ep} is the number of estimators used in RFR and number of passes in CNN-MCD, y_i is the actual value of the capacity test $c_{i=1,..,z} \in \{SOH, RUL\}$, $\hat{y}_{i,j}$ is the jth point estimate of the capacity test $c_{i=1,..,z} \in \{SOH, RUL\}$. Note that n_{ep} varies for each vahana in case of RFR, however n_{ep} is 1000 for CNN-MCD.

The uncertainty (Standard Deviation) of a capacity test $c_{i=1,..,z} \in \{SOH, RUL\}$ is given by $\sigma(c_i)$

$$\sigma_I(c_i) = \sqrt{\frac{\sum (\hat{y}_{i,j} - \mu^I_{n_{ep}})^2}{n_{ep}}} \tag{3}$$

where, $\mu^I_{n_{ep}}$ is the mean of the point estimates of each capacity test $c_i \in \{SOH, RUL\}$ of $I \in \{\text{RFR,CNN-MCD}\}$.

The Sharpness $Sh_I(c_i)$ is defined as number of times the actual value falls within one standard deviation $\sigma_I(c_i)$ of the predicted point estimates of that capacity test c_i . This lead to the following expression $Sh_I(c_i)$:

$$Sh_I(c_i) = \frac{1}{n_{ep}} \sum_{j=1}^{n_{ep}} f_{c_{i,j}} \tag{4}$$

with,

$$f_{c_{i,j}} = \begin{cases} 1, & \text{if } |\hat{y}_{i,j} - y_i| < \sigma(c_i) \\ 0, & \text{if } |\hat{y}_{i,j} - y_i| \geq \sigma(c_i) \end{cases} \tag{5}$$

The Continuous Ranked Probability Score (CRPS):
In general, The Continuous Ranked Probability Score (CRPS) is a proper scoring function that compares a single ground truth value to a Cumulative Distribution Function (CDF)[14][17]. It is first introduced in the 70's [33] and measures the distance of the predictive distribution to the observed data-generating distribution.

$$\text{CRPS}(F, y) = \int_{-\infty}^{\infty} \left(F(x) - 1_{(x \geq y)}\right)^2 \, dx \tag{INT}$$

$$= \mathbb{E}_X \left[|X - y| - \frac{1}{2} \mathbb{E}_{X,X'} \left[|X - X'| \right] \right] \tag{NRG}$$

$$= \mathbb{E}_X \left[|X - y| \right] + \mathbb{E}_X[X] - 2\mathbb{E}_X[X \cdot F(X)] \tag{PWM}$$

where, y is the actual value and F the CDF of predictive distribution, X and X' are independently and identically distributed according to F, and $\mathbb{E}_A$ is the expectation according to the law of the random variables A [59].

INT is the integral form CRPS. The equation NRG is based on the energy distance and therefore called the energy form of the CRPS. The equation PWM is called the probability weighted moment (PWM) form of the CRPS, which is valid for continuous forecast CDFs [59]. Since the definition includes the CDF, the CRPS can be used for both parametric and non-parametric predictions [14]. That is, we can use it when the model predicts several point estimates for a test instance and also when the model predicts mean and standard deviation (assuming that the output follows normal distribution) of the distribution directly. If the predicted distribution is a a point-wise estimate, the CRPS reduces to the MAE and can be viewed as a generalization of the MAE into distributional predictions. For RFR and CNN-MCD, we use the NRG and PWM forms of CRPS as below:

$$\mathrm{CRPS}^I_{\mathrm{NRG}}(N, y) = \frac{1}{N} \sum_{i=1}^{N} |x_i - y| - \frac{1}{2N^2} \sum_{i,j=1}^{N} |x_i - x_j| \tag{6}$$

$$\mathrm{CRPS}^I_{\mathrm{PWM}}(N, y) = \frac{1}{N} \sum_{i=1}^{N} |x_i - y| + \hat{\beta}_0 - 2\hat{\beta}_1 \tag{7}$$

where, F is known through an N-ensemble $x_{i=1,\ldots,N}$, the members of x_i need to be sorted in ascending order to avoid the loss of generality, $\hat{\beta}_0 = \frac{1}{N} \sum_{i=1}^{N} x_i$, $\hat{\beta}_1 = \frac{1}{N(N-1)} \sum_{i=1}^{N} (i-1) \cdot x_i$ [59].

For Mixture Density Network, the output is not a point estimate(s), rather the model outputs defining parameters of the output distribution (assumed as normal distribution in this thesis) i.e., means, standard deviations and mixing coefficients (weights). Therefore, the above defined metrics are not applicable in the case of MDN. At the capacity level, for MDN we define distributional mean of a capacity test $c_{i=1,..,z} \in \{SOH, RUL\}$ as the weighted average of the component (mixtures) means as below:

$$\mu_{MDN}(c_i) = \sum_{j=1}^{k} \alpha_j(c_i) \mu_j(c_i) \tag{8}$$

where, $\alpha_j(c_i)$ and $\mu_j(c_i)$ are the weight and mean of the component j. Note that the weights sum to 1.

The standard deviation for MDN is a multimodal probability density function for a capacity test c_i [8]:

$$\sigma_{MDN}(\mathbf{c}_i) = \sqrt{\sum_{j=1}^{k} \alpha_j(\mathbf{c}_i) \left(\sigma_j(\mathbf{c}_i)^2 + (\mu_j(\mathbf{c}_i) - \mu_{MDN}(\mathbf{c}_i))^2 \right)} \tag{9}$$

where, $\alpha_j(c_i)$, $\mu_j(c_i)$, $\sigma_j(c_i)$ are the weight, mean and standard deviation of the component j respectively.

$$Sh_{\mathrm{MDN}}(c_i) = \frac{1}{N_c} \sum_{i=1}^{N_c} f_{c_{i,j}} \tag{10}$$

with,

$$f_{c_{i,j}} = \begin{cases} 1, & \text{if } |\mu_{\text{MDN}}(c_i) - y_i| < \sigma_{\text{MDN}}(c_i) \\ 0, & \text{if } |\mu_{\text{MDN}}(c_i) - y_i| \geq \sigma_{\text{MDN}}(c_i) \end{cases} \tag{11}$$

where, N_c is the number of capacity tests of the test Vahana, y_i is the actual value of the capacity test c_i.

Since the MDN model predicts the parameters μ and σ of a Normal distribution, the CRPS can be calculated with [52] [18]:

$$\text{CRPS}(\mathcal{N}(\mu, \sigma^2), y) = \sigma \left(w(2.\Phi(w) - 1) + 2.\phi(w) - \pi^{-\frac{1}{2}} \right) \tag{12}$$

where, $w = \left(\frac{y-\mu}{\sigma}\right)$, Φ is the cumulative distribution function (CDF), ϕ the probability density function (PDF) of the normal distribution and with μ, σ, y - the mean, standard deviation and actual value of capacity test c_i.

Additionally, error (residual) for all models is computed by taking the difference between actual value of the c_i and the mean of the predicted distribution of that capacity test c_i. In case of RFR and CNN-MCD, this is the mean of the point estimates for each c_i.

At Vahana level:
For models RFR and CNN-MCD, the metrics sharpness $Sh_I(c_i)$ and $CRPS$ (NRG,PWM) are averaged, i.e., average across number of capacity tests for that Vahana. Similarly for MDN, $CRPS$ is a average value across the capacity tests of a Vahana and sharpness is the equation (10). The average standard deviation for all the models is the square root of the average standard deviations (SD) across capacity tests of a Vahana. The metrics, MAE and RMSE for all the models $V \in \{RFR, \text{CNN-MCD}, MDN\}$ is as follows:

$$MAE_V(v_i) = \frac{1}{N_c} \sum_{i=1}^{N_c} |y_i - \mu_V(c_i)| \tag{13}$$

$$RMSE_V(v_i) = \sqrt{\frac{1}{N_c} \sum_{i=1}^{N_c} (y_i - \mu_V(c_i))^2} \tag{14}$$

where, $\mu_V(c_i)$ is the distribution mean of the capacity test c_i.

At Model level: The metrics are averaged for all the models considering all 21 Vahanas. However, the average standard deviation is the square root of the average standard deviations (SD) across Vahanas.

7 Results - Estimating the distribution of State-of-Health of eVTOL batteries

The Figure 5 shows the selected features and their occurrences across 21 tests performed for SOH distribution estimation. It can be observed that overall 19 features out of 30 features (63.33%) have been selected across the tests. Take-off voltage related features; duration of cc, cv ,cruise ; landing and cruise maximum temperature have been selected in all 21 tests, which are also the features that are expected to vary across the mission profiles (see 6.2). The feature selection method is indeed able to capture the variations in the features in case of SOH predictions.

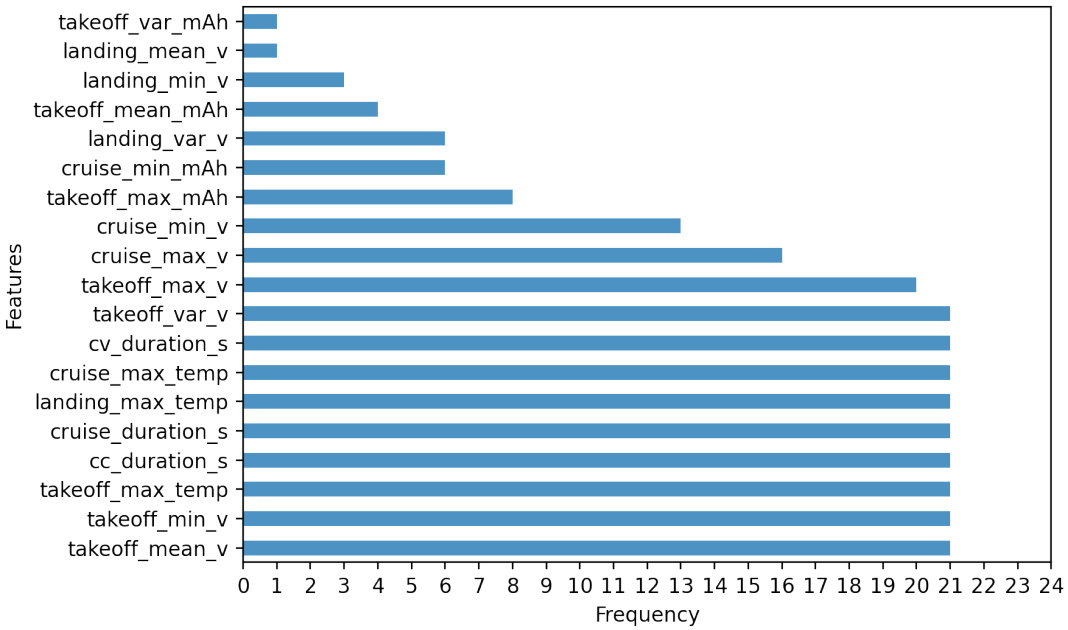


Figure 5: The selected features frequency across 21 tests for SOH distribution estimation

The Table 5 shows the overall results of the models when all 21 mission profiles are considered. We can see that RFR performs better when all the metrics are considered. We consider MAE, RMSE, Sharpness and CRPS (PWM compared to CRPS normal distribution). The comparison of CRPS in PWM form to CRPS of normal distribution is because CRPS PWM gives an unbiased estimator of the average CRPS of a random ensemble [59]. Apart from Sharpness (higher the better), lower the metric score better the model performance.

Table 5: Overall Performance Metrics of Models - SOH(%)

Model	MAE	RMSE	SD	Sh(%)	CRPS_NRG	CRPS_PWM	CRPS_Norm
RFR	**0.9416**	**1.1444**	1.5425	**0.5806**	0.6565	**0.9662**	-
CNN-MCD	1.5351	1.8469	1.0933	0.3874	1.1886	1.2754	-
MDN	1.3748	1.5681	0.6726	0.3784	-	-	1.1722

The Table 6 shows the optimised hyperparameters for the models RFR, CNN-MCD and MDN for each mission profile. The dropout rate is 0.2 across all tests performed using CNN-MCD model for SOH distribution estimation.

Table 6: Results — Hyperparameter Grid Search of each test for SOH, #Fs is number of features selected out of 30 features, ρ is dropout rate, BS is Batch Size

SOH								
Vahana	**#Fs**	**RFR**		**CNN-MCD**		**MDN**		
		Max features	**Estimators**	ρ	**BS**	**Nodes**	**Mixtures**	**BS**
VAH01	11	0.5	500	0.2	10	200	1	25
VAH02	11	0.5	750	0.2	10	200	6	15
VAH05	13	1	1500	0.2	10	50	6	10
VAH06	12	0.5	100	0.2	10	150	3	25
VAH09	12	0.75	850	0.2	16	150	1	20
VAH10	13	1	850	0.2	16	150	5	15
VAH11	13	0.5	300	0.2	16	100	3	15
VAH12	10	0.75	100	0.2	16	200	2	10
VAH13	13	0.5	100	0.2	12	150	5	30
VAH15	12	0.5	850	0.2	10	50	2	10
VAH16	14	0.5	750	0.2	14	100	6	25
VAH17	12	0.25	500	0.2	10	100	3	10
VAH20	12	0.5	100	0.2	10	50	1	10
VAH22	14	0.5	500	0.2	4	150	6	10
VAH23	11	0.25	750	0.2	10	150	3	30
VAH24	14	0.25	850	0.2	10	100	3	25
VAH25	14	0.5	300	0.2	10	100	3	10
VAH26	13	0.5	1000	0.2	20	100	2	10
VAH27	13	0.5	1500	0.2	20	150	5	15
VAH28	13	0.5	100	0.2	20	50	1	10
VAH30	17	0.5	1000	0.2	8	200	2	10

7.1 Machine Learning algorithms- SOH results

The tables Table 7, Table 8, Table 10 show MAE, RMSE, Standard deviation (SD), Sharpness (Sh(%)), CRPS (in both NRG, PWM) for RFR, CNN-MCD and CRPS Normal distribution for the MDN that are obtained SOH distribution prediction for each mission profile. The last line of the each table gives the average values of the error metrics, which shows the performance metrics at model level as discussed in 6.7.

Moreover, CRPS evaluates both accuracy and sharpness of the predicted distribution. The smaller the value of CRPS, the closer is the prediction to the true value. In an ideal case, when a perfect prediction without uncertainty (i.e., a point prediction) is obtained, CRPS equals zero [12], [60]. Therefore, for overall model assessment we consider CRPS as the performance metric.

The average values are highlighted in bold and they indicate that the particular model has performed better than others with respect to that metric. We can see that RFR is has the low MAE, RMSE, CRPS with high sharpness compared to other models. If we look at the individual error metrics RMSE and Sharpness when RFR is considered, for Vahana 11 the RMSE is 2.70% (high) and Sharpness is only 0.26% (low). Intuitively, we can expect the CRPS score to be high, this is indeed reflected in CRPS score (2.096). In case of Vahana 23, the RMSE is 3.39% (high) but sharpness is 0.54%. If we confine our evaluation to sharpness alone, the performance can be deemed satisfactory, however, CRPS metric is high with 2.02%. Also in case of Vahana 28, if RMSE- 0.96% is alone considers, the performance can be deemed satisfactory, however CRPS score is 1.43%. Therefore, we can say that the choice of the CRPS metric for assessing performance is reasonable.

Table 7: Performance Metrics of RFR- SOH (%)

Vahana	MAE	RMSE	SD	Sh(%)	CRPS_NRG	CRPS_PWM
VAH01	1.4745	1.6773	1.9114	0.4720	1.0001	1.1738
VAH02	0.8371	0.9849	1.6459	0.6634	0.4898	0.6093
VAH05	0.8488	0.9988	1.0986	0.5410	0.5779	0.6352
VAH06	0.6429	0.7885	1.5690	0.7026	0.4476	1.3251
VAH09	0.5162	0.6617	0.9847	0.6030	0.3634	0.4672
VAH10	0.7828	0.9170	1.4176	0.5077	0.6309	0.7309
VAH11	2.5404	2.7013	1.6434	0.2627	1.8124	2.0957
VAH12	0.7577	1.0073	1.4364	0.5765	0.5814	1.4013
VAH13	0.4744	0.5635	1.0787	0.6685	0.3225	1.1968
VAH15	0.3063	0.4722	1.1439	0.7015	0.2169	0.3230
VAH16	0.7760	0.9732	1.6103	0.7195	0.4270	0.5470
VAH17	0.6825	0.8336	0.9375	0.5592	0.4631	0.6379
VAH20	0.6296	0.9221	2.0130	0.7583	0.4177	1.3068
VAH22	0.4334	0.6863	1.0970	0.7072	0.2406	0.4218
VAH23	2.4711	3.3873	3.3037	0.5365	1.9041	2.0219
VAH24	0.4921	0.6672	1.6401	0.7239	0.3729	0.4774
VAH25	0.9132	1.0151	1.1742	0.4148	0.6325	0.9319
VAH26	0.6299	0.8703	1.2057	0.6240	0.4457	0.5316
VAH27	0.7435	0.8551	0.8926	0.5911	0.4362	0.4958
VAH28	0.8010	0.9559	1.4555	0.6287	0.5613	1.4281
VAH30	2.0206	2.0928	1.2792	0.2311	1.4416	1.5308
Average	**0.9416**	**1.1444**	1.5425	**0.5806**	0.6565	**0.9662**

When considering CRPS, RF regression model has better performance with average CRPS score of 0.97%. Moreover, it is observed that Vahana 11 and Vahana 23 consistently have higher CRPS scores (highlighted in blue) across all the models of SOH distribution prediction. Vahana 15 (highlighted in green) in the Table 7 has lowest CRPS score when RFR is considered. It has relatively low CRPS score in MDN and CNN-MCD models as well.

Table 9 shows the intermediate capacity test errors obtained for SOH distribution prediction when RF regression is considered. There is no specific trend or pattern in the performance metrics. This can intuitively be explained by the fact that each mission profile is different as its' values are recorded under varying environment conditions. Having said that, however in general, the predicted mean of the distribution of capacity tests for each vahana deceases as mission cycles increase without any drastic changes.

Table 8: Performance Metrics of CNN-MCD - SOH (%)

Vahana	MAE	RMSE	SD	Sh(%)	CRPS_NRG	CRPS_PWM
VAH01	1.1792	1.3408	0.8374	0.3371	0.8503	0.9367
VAH02	0.7088	0.7999	0.7601	0.4224	0.5023	0.5908
VAH05	1.1732	1.5955	1.0478	0.4421	0.8836	0.9701
VAH06	1.2273	1.5815	1.6501	0.6228	0.9105	0.9981
VAH09	1.4441	1.5883	0.9191	0.3092	1.0153	1.1043
VAH10	1.7464	2.2026	1.5080	0.3452	1.1643	1.2471
VAH11	3.2496	4.4992	0.6858	0.2442	3.0110	3.0962
VAH12	2.0871	2.4916	1.0381	0.2610	1.6402	1.7217
VAH13	0.9225	1.1411	1.0784	0.5108	0.6663	0.7533
VAH15	0.7049	0.8121	0.8470	0.5299	0.4626	0.5525
VAH16	3.2614	3.4527	0.9247	0.0589	2.7762	2.8626
VAH17	1.4525	1.6397	0.8839	0.2892	1.0597	1.1473
VAH20	3.9353	4.2570	1.1434	0.0514	3.3149	3.3996
VAH22	0.8549	1.0559	0.7112	0.4400	0.6072	0.6970
VAH23	2.3442	3.3069	1.1801	0.3428	1.9361	2.0233
VAH24	0.4872	0.6146	1.3470	0.6851	0.4210	0.5091
VAH25	0.9371	1.1393	0.7443	0.3719	0.7080	0.7969
VAH26	1.2535	1.5058	1.1983	0.4463	0.8888	0.9757
VAH27	0.9937	1.1124	0.8248	0.3850	0.7103	0.7990
VAH28	0.8814	1.0130	1.0824	0.4966	0.6096	0.6956
VAH30	1.3925	1.6354	1.7188	0.5432	0.8222	0.9073
Average	1.5351	1.8469	1.0933	0.3874	1.1886	1.2754

Capacity test across missions											
Vahana	SOH	1	5	10	15	20	25	30	35	40	45
VAH01	Actual	100.00	91.85	86.83	82.66	-	-	-	-	-	-
	Mean	95.97	90.43	86.00	81.24	-	-	-	-	-	-
	Error	4.03	1.42	0.83	1.42	-	-	-	-	-	-
	RMSE	5.27	2.92	1.45	1.85	-	-	-	-	-	-
	SD	3.39	2.55	1.19	1.18	-	-	-	-	-	-
	CRPS	2.36	1.34	0.65	1.04	-	-	-	-	-	-
VAH02	Actual	100.00	91.14	85.52	-	-	-	-	-	-	-
	Mean	98.34	92.37	85.06	-	-	-	-	-	-	-
	Error	1.66	-1.23	0.47	-	-	-	-	-	-	-
	RMSE	2.52	2.53	1.18	-	-	-	-	-	-	-
	SD	1.89	2.21	1.09	-	-	-	-	-	-	-
	CRPS	0.81	0.78	0.38	-	-	-	-	-	-	-
VAH05	Actual	100.00	92.84	88.22	85.14	82.53	79.33	-	-	-	-
	Mean	99.53	94.73	88.89	84.54	82.02	78.50	-	-	-	-
	Error	0.47	-1.89	-0.67	0.60	0.51	0.83	-	-	-	-
	RMSE	1.27	2.32	1.25	1.29	1.37	1.10	-	-	-	-
	SD	1.18	1.35	1.06	1.14	1.27	0.72	-	-	-	-
	CRPS	0.13	1.39	0.44	0.29	0.44	0.58	-	-	-	-
VAH06	Actual	100.00	92.22	86.47	83.28	-	-	-	-	-	-

Table 9 – *Continued from previous page*

Vahana	SOH	1	5	10	15	20	25	30	35	40	45
	Mean	99.79	92.37	86.67	83.30	-	-	-	-	-	-
	Error	0.21	-0.15	-0.20	-0.02	-	-	-	-	-	-
	RMSE	0.79	1.78	1.28	1.02	-	-	-	-	-	-
	SD	0.76	1.77	1.26	1.02	-	-	-	-	-	-
	CRPS	1.01	1.22	1.09	1.02	-	-	-	-	-	-
	Actual	100.00	92.52	89.98	85.35	79.76	-	-	-	-	-
	Mean	98.62	91.66	90.13	84.99	79.75	-	-	-	-	-
	Error	1.38	0.86	-0.15	0.36	0.01	-	-	-	-	-
VAH09	RMSE	2.51	1.27	0.91	0.72	1.14	-	-	-	-	-
	SD	2.10	0.93	0.89	0.62	1.14	-	-	-	-	-
	CRPS	0.50	0.69	0.31	0.39	0.38	-	-	-	-	-
	Actual	100.00	91.87	86.72	83.16	79.63	76.49	-	-	-	-
	Mean	98.98	92.37	87.09	82.98	80.31	77.94	-	-	-	-
	Error	1.02	-0.49	-0.37	0.18	-0.67	-1.45	-	-	-	-
VAH10	RMSE	2.10	1.85	1.36	1.61	0.92	1.94	-	-	-	-
	SD	1.84	1.78	1.31	1.60	0.62	1.30	-	-	-	-
	CRPS	0.34	0.66	0.50	0.59	0.46	0.97	-	-	-	-
	Actual	100.00	93.25	89.26	86.48	82.20	80.05	77.70	75.62	73.01	-
	Mean	99.42	95.07	91.64	88.77	84.78	82.53	80.02	78.26	76.39	-
	Error	0.58	-1.82	-2.38	-2.29	-2.58	-2.48	-2.32	-2.64	-3.38	-
VAH11	RMSE	1.27	2.38	2.78	2.76	2.75	2.75	2.62	3.16	4.07	-
	SD	1.13	1.54	1.44	1.55	0.94	1.19	1.22	1.74	2.28	-
	CRPS	0.45	1.44	1.92	1.86	2.41	2.10	1.94	2.35	2.43	-
	Actual	100.00	92.89	88.53	85.34	81.28	78.31	75.89	75.03	73.69	71.61
	Mean	99.78	91.90	88.40	85.12	80.98	79.55	77.86	75.89	74.52	73.31
	Error	0.22	0.99	0.13	0.21	0.30	-1.25	-1.97	-0.86	-0.84	-1.70
VAH12	RMSE	0.97	1.64	0.79	0.70	1.11	1.57	2.23	1.95	2.42	3.03
	SD	0.94	1.31	0.78	0.66	1.07	0.95	1.05	1.75	2.27	2.51
	CRPS	1.01	1.44	1.05	1.07	1.22	1.61	2.26	1.45	1.64	1.57
	Actual	100.00	92.52	87.64	84.20	75.76	-	-	-	-	-
	Mean	99.44	93.18	88.27	83.37	75.58	-	-	-	-	-
	Error	0.56	-0.66	-0.63	0.83	0.19	-	-	-	-	-
VAH13	RMSE	1.23	1.42	0.94	1.41	1.80	-	-	-	-	-
	SD	1.09	1.26	0.70	1.15	1.79	-	-	-	-	-
	CRPS	1.11	1.36	1.30	1.25	1.29	-	-	-	-	-
	Actual	100.00	90.87	84.76	-	-	-	-	-	-	-
	Mean	98.64	90.73	85.26	-	-	-	-	-	-	-
	Error	1.36	0.14	-0.50	-	-	-	-	-	-	-
VAH15	RMSE	2.56	1.03	0.97	-	-	-	-	-	-	-
	SD	2.16	1.02	0.83	-	-	-	-	-	-	-
	CRPS	0.47	0.24	0.47	-	-	-	-	-	-	-
	Actual	100.00	91.24	84.92	-	-	-	-	-	-	-
	Mean	97.71	91.75	84.37	-	-	-	-	-	-	-
	Error	2.29	-0.52	0.55	-	-	-	-	-	-	-
VAH16	RMSE	3.49	1.63	1.57	-	-	-	-	-	-	-
	SD	2.63	1.54	1.47	-	-	-	-	-	-	-

Table 9 – *Continued from previous page*

Vahana	SOH	1	5	10	15	20	25	30	35	40	45
	CRPS	1.03	0.48	0.46	-	-	-	-	-	-	-
VAH17	Actual	100.00	91.74	86.31	82.34	77.93	-	-	-	-	-
	Mean	99.74	92.02	87.03	83.44	79.17	-	-	-	-	-
	Error	0.26	-0.28	-0.73	-1.10	-1.24	-	-	-	-	-
	RMSE	0.98	0.92	1.02	1.33	1.60	-	-	-	-	-
	SD	0.95	0.88	0.72	0.75	1.01	-	-	-	-	-
	CRPS	0.22	0.51	0.68	0.97	0.92	-	-	-	-	-
VAH20	Actual	100.00	90.89	84.08	-	-	-	-	-	-	-
	Mean	97.65	90.10	83.79	-	-	-	-	-	-	-
	Error	2.35	0.79	0.28	-	-	-	-	-	-	-
	RMSE	3.51	1.64	2.45	-	-	-	-	-	-	-
	SD	2.60	1.44	2.43	-	-	-	-	-	-	-
	CRPS	1.94	1.31	1.06	-	-	-	-	-	-	-
VAH22	Actual	100.00	90.82	84.59	-	-	-	-	-	-	-
	Mean	98.46	90.61	84.49	-	-	-	-	-	-	-
	Error	1.54	0.21	0.10	-	-	-	-	-	-	-
	RMSE	2.53	0.97	0.79	-	-	-	-	-	-	-
	SD	2.01	0.95	0.78	-	-	-	-	-	-	-
	CRPS	0.73	0.33	0.35	-	-	-	-	-	-	-
VAH23	Actual	100.00	91.25	85.95	-	-	-	-	-	-	-
	Mean	90.50	89.17	87.86	-	-	-	-	-	-	-
	Error	9.50	2.08	-1.91	-	-	-	-	-	-	-
	RMSE	10.09	3.90	3.82	-	-	-	-	-	-	-
	SD	3.41	3.30	3.31	-	-	-	-	-	-	-
	CRPS	7.78	1.39	1.40	-	-	-	-	-	-	-
VAH24	Actual	100.00	91.95	86.72	82.18	-	-	-	-	-	-
	Mean	98.45	92.70	86.59	81.97	-	-	-	-	-	-
	Error	1.55	-0.75	0.13	0.21	-	-	-	-	-	-
	RMSE	2.55	2.31	1.14	1.41	-	-	-	-	-	-
	SD	2.02	2.18	1.13	1.39	-	-	-	-	-	-
	CRPS	0.66	0.53	0.39	0.45	-	-	-	-	-	-
VAH25	Actual	100.00	91.94	85.39	-	-	-	-	-	-	-
	Mean	98.79	90.64	83.97	-	-	-	-	-	-	-
	Error	1.21	1.30	1.42	-	-	-	-	-	-	-
	RMSE	2.43	1.55	1.58	-	-	-	-	-	-	-
	SD	2.10	0.85	0.68	-	-	-	-	-	-	-
	CRPS	0.60	1.30	1.35	-	-	-	-	-	-	-
VAH26	Actual	100.00	92.34	86.25	81.97	78.39	-	-	-	-	-
	Mean	97.90	92.79	85.94	81.28	76.30	-	-	-	-	-
	Error	2.10	-0.46	0.31	0.68	2.09	-	-	-	-	-
	RMSE	2.91	1.48	0.70	1.56	2.60	-	-	-	-	-
	SD	2.02	1.41	0.62	1.40	1.54	-	-	-	-	-
	CRPS	1.11	0.33	0.28	0.60	1.37	-	-	-	-	-
VAH27	Actual	100.00	91.10	83.79	-	-	-	-	-	-	-
	Mean	99.84	91.87	82.24	-	-	-	-	-	-	-
	Error	0.16	-0.77	1.55	-	-	-	-	-	-	-

Table 9 – *Continued from previous page*

Vahana	SOH	1	5	10	15	20	25	30	35	40	45
	RMSE	0.78	1.30	1.89	-	-	-	-	-	-	-
	SD	0.77	1.04	1.09	-	-	-	-	-	-	-
	CRPS	0.07	0.47	1.02	-	-	-	-	-	-	-
VAH28	Actual	100.00	92.68	87.92	84.36	81.78	-	-	-	-	-
	Mean	98.38	92.07	87.95	83.59	81.72	-	-	-	-	-
	Error	1.62	0.61	-0.03	0.76	0.06	-	-	-	-	-
	RMSE	2.86	1.43	1.15	1.50	1.14	-	-	-	-	-
	SD	2.36	1.30	1.15	1.30	1.14	-	-	-	-	-
	CRPS	1.46	1.50	1.12	1.19	1.08	-	-	-	-	-
VAH30	Actual	100.00	91.40	85.77	81.52	-	-	-	-	-	-
	Mean	98.93	93.91	87.65	83.74	-	-	-	-	-	-
	Error	1.07	-2.51	-1.88	-2.21	-	-	-	-	-	-
	RMSE	1.97	2.92	2.15	2.39	-	-	-	-	-	-
	SD	1.65	1.49	1.03	0.90	-	-	-	-	-	-
	CRPS	0.40	1.86	1.45	1.86	-	-	-	-	-	-

Table 9: Results — Distribution Estimation of SOH (%) at each capacity test using RF Regression, - implies Non-applicable, SD- Standard Deviation, CRPS is in PWM form

Table 10: Performance Metrics of MDN- SOH (%)

Vahana	MAE	RMSE	SD	Sh(%)	CRPS_Norm
VAH01	2.5731	2.7629	1.3241	0.2353	2.1044
VAH02	1.0957	1.4913	0.8366	0.3846	0.8758
VAH05	0.9175	1.0214	0.6988	0.3448	0.6666
VAH06	0.4630	0.6145	0.4277	0.5263	0.3368
VAH09	1.3834	1.4671	0.3987	0.0952	1.1706
VAH10	0.9206	1.1215	0.3503	0.1786	0.7694
VAH11	3.0210	3.2452	0.8719	0.0455	2.7242
VAH12	3.0443	3.3228	0.5775	0.0217	2.9039
VAH13	0.4841	0.6265	0.6023	0.5500	0.3411
VAH15	0.4424	0.5588	0.5277	0.7273	0.3123
VAH16	0.4292	0.5117	0.3973	0.3636	0.2988
VAH17	0.6655	0.8539	0.4522	0.4500	0.5018
VAH20	0.6017	0.7353	0.7623	0.5833	0.4432
VAH22	0.3709	0.4429	0.3701	0.7000	0.2566
VAH23	9.2524	9.9403	1.4975	0.0714	8.5068
VAH24	0.4010	0.5799	0.2947	0.5000	0.2983
VAH25	0.7177	0.7996	0.3647	0.1818	0.5576
VAH26	0.4239	0.7535	0.5440	0.8571	0.3241
VAH27	0.6467	0.7944	0.4784	0.4545	0.4905
VAH28	0.4024	0.5883	0.3774	0.5652	0.2953
VAH30	0.6146	0.6973	0.3862	0.1111	0.4383
Average	1.3748	1.5681	0.6726	0.3784	1.1722

7.2 Discussion

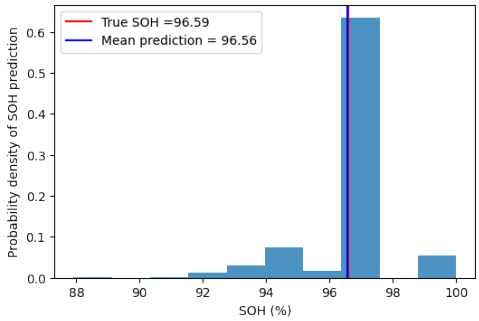
In this section, we examine the distribution of State of Health (SOH) derived from all the models. We opt to display plots for each model to clarify how the distribution is depicted across different models. We choose intermediate capacity tests of Vahana 15 and 23, as they have similar CRPS metrics (low, high) across the models.

7.2.1 Illustrations- Vahana 15

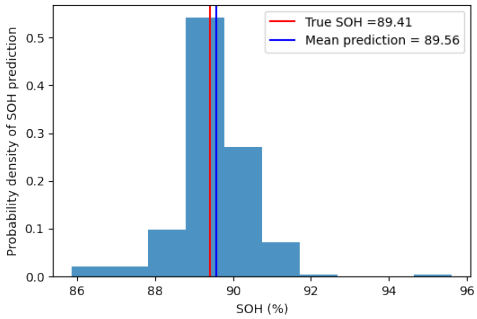
Figures 6a, 6b, 6c show the distribution of Vahana 15 at capacity tests 2, 6, 10 respectively, of the RFR model. The model slightly underestimated the mean SOH at capacity test 2 and started to over estimate as capacity tests progress. Moreover, from the standard deviation (SD) values for the respective capacity tests are 1.42 , 0.87, 0.83 (see vahana 15 in Table 17) implying a decreasing trend. This shows, in case of RFR for vahana 15, while the uncertainty decreases, the model tends to overestimate the mean SOH. That is from capacity test 6 the model tends to over estimate the SOH of the battery. (Note that this is only in the case of vahana 15 when rfr is considered).

If we look at the same vahana at the same capacity test positions with the CNN-MCD model (see plots Figures 7a, 7b, 7c), it has similar pattern, however with increasing uncertainty. That is standard deviation values at capacity test 2, 6, 10 of vahana 15 are 0.53, 0.72, 0.84, which show increasing trend. (see vahana 15 in Table 19). The CNN-MCD model also slightly underestimated mean SOH at capacity test 2 and started to over estimate from capacity test 6. This shows, in case of CNN-MCD model for vahana 15, while the uncertainty increases, the model also starts to overestimate the mean SOH.

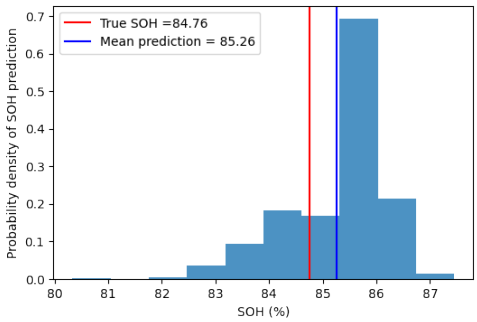
In case of MDN model, the uncertainity decreases and then increases for vahana 15 (Standard deviations are 0.59,0.47, 0.51 at capacity test 2 , 6, 10 see Table 21)), while the model has always overestimated the mean SOH for the capacity tests at 2, 6, 10 for vahana 15 (see Figures 8a, 8b, 8c)). Moreover, as expected the mean SOH for the capacity tests is centered around the predicted distribution.



(a) At Capacity test 2 of Vahana 15

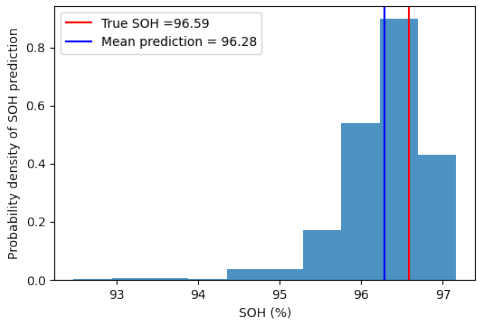


(b) At Capacity test 6 of Vahana 15

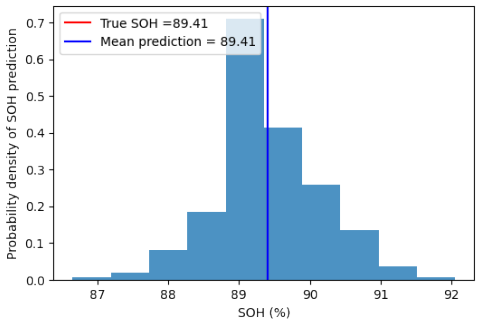


(c) At Capacity test 10 of Vahana 15

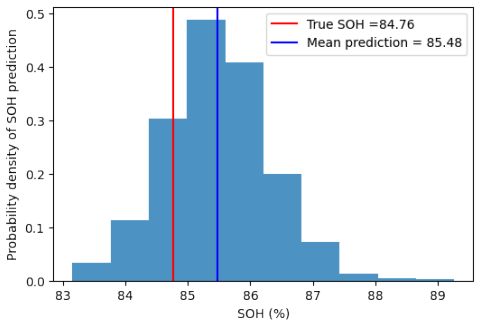
Figure 6: Capacity Test plots of RFR-SOH for Vahana 15



(a) At Capacity test 2 of Vahana 15

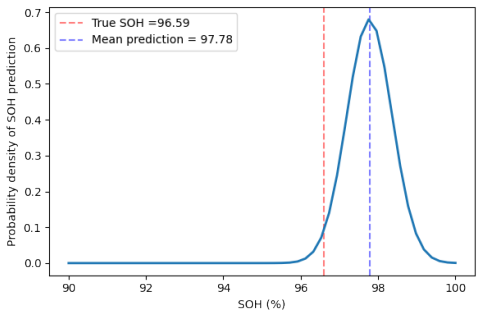


(b) At Capacity test 6 of Vahana 15

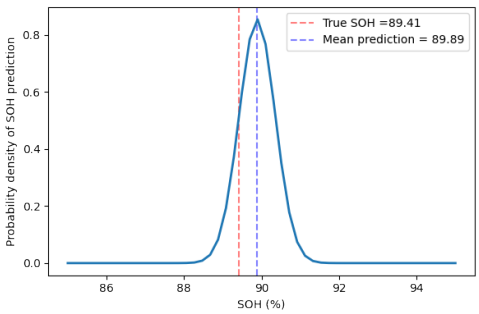


(c) At Capacity test 10 of Vahana 15

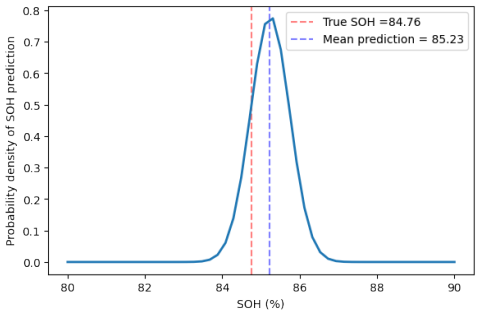
Figure 7: Capacity Test plots of CNN-MCD-SOH for Vahana 15



(a) At Capacity test 2 of Vahana 15



(b) At Capacity test 6 of Vahana 15



(c) At Capacity test 10 of Vahana 15

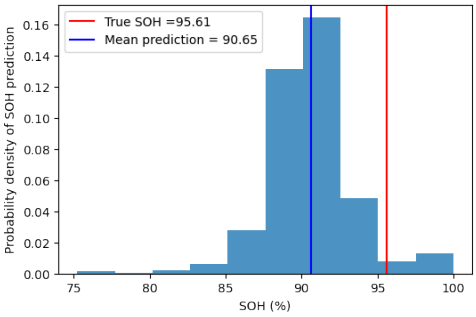
Figure 8: Capacity Test plots of MDN-SOH for Vahana 15

7.2.2 Illustrations- Vahana 23

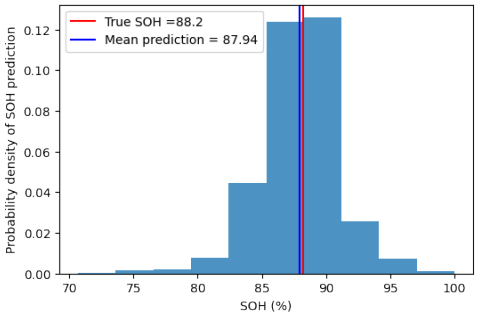
Figures 9a, 9b, 9c show the distribution of Vahana 23 at capacity tests 2, 8, 13 respectively, of the RFR model. The model underestimated the mean SOH at capacity test 2 and started to over estimate around capacity test 8. Moreover, from the uncertainty perspective the standard deviation (SD) increases as the capacity tests progress. SD values for the respective capacity tests are 2.93, 3.22, 3.53 (see vahana 23 in Table 17) implying a increasing trend. This shows, in case of RFR for vahana 23, when the uncertainty increases, the model tends to overestimate the mean SOH values for the capacity tests. That is around capacity test 8 the model tends to over estimate the SOH of the battery. (Note that this is only in the case of vahana 23 when rfr is considered).

Considering CNN-MCD model for the same vahana at the same capacity test positions (see plots Figures 10a, 10b, 10c), it has the SD decreasing (SD values- 1.11 to 1.01) (see vahana 23 in Table 19 and note that there are fluctuations in SD values). The CNN-MCD model highly underestimated mean SOH at capacity test 2 and over estimated at capacity test 13. However we cannot draw conclusions from this behaviour as there are fluctuations in Mean SOH predictions. This was also true in the case of RFR model (in case of vahana 23) but the change in mean SOH predictions is not too drastic.

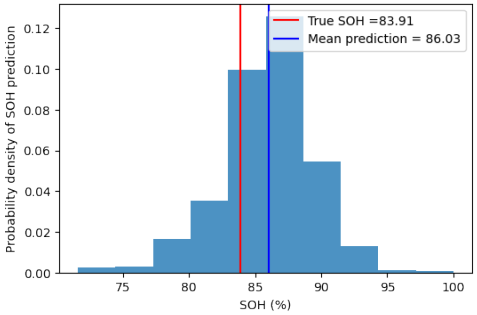
In case of MDN model, the uncertainity decreases for vahana 23 (Standard deviations are 1.62, 1.45, 1.42 at capacity tests- 2 , 8, 13 respectively) Table 21, while the model has always overestimated the mean SOH of the capacity tests at 2, 8, 13 for vahana 23 and the means are centered around the distribution as expected (see Figures 11a, 11b, 11c)).



(a) At Capacity test 2 of Vahana 23

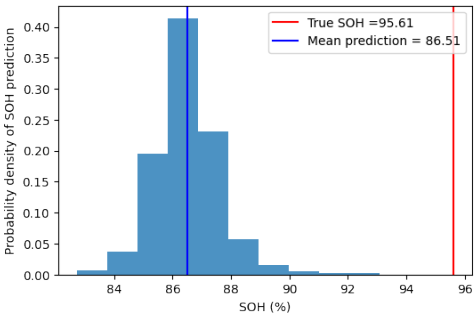


(b) At Capacity test 8 of Vahana 23

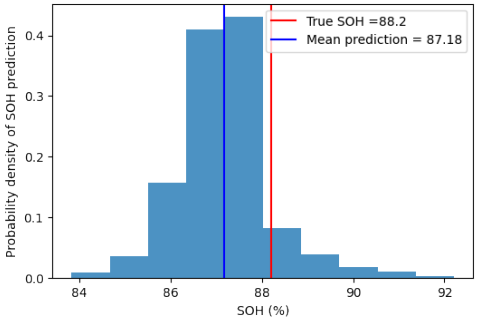


(c) At Capacity test 13 of Vahana 23

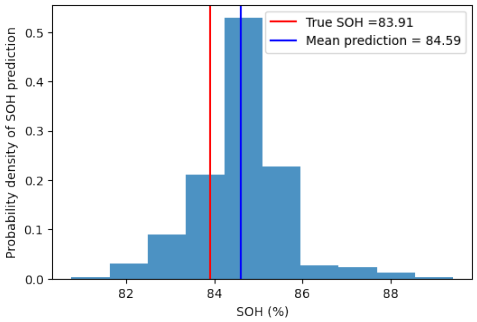
Figure 9: Capacity Test plots of RFR-SOH for Vahana 23



(a) At Capacity test 2 of Vahana 23

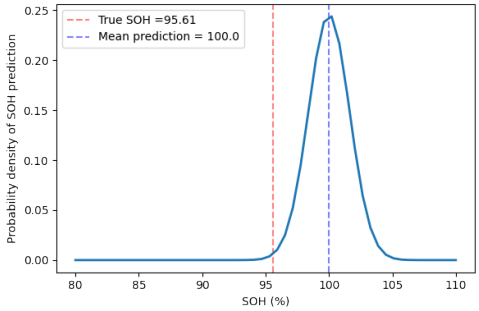


(b) At Capacity test 8 of Vahana 23

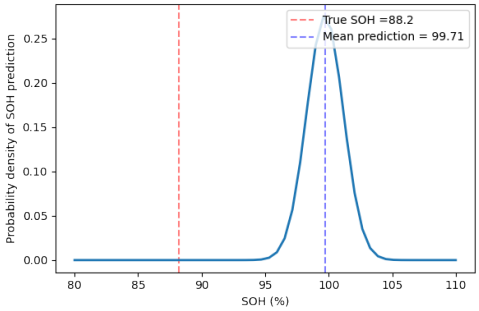


(c) At Capacity test 13 of Vahana 23

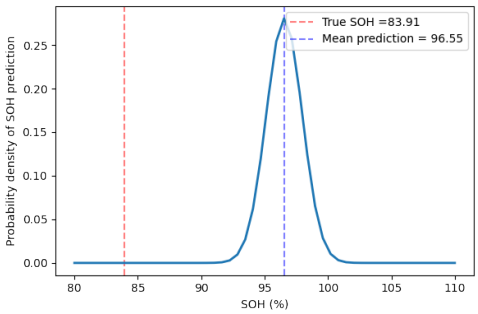
Figure 10: Capacity Test plots of CNN-MCD-SOH for Vahana 23



(a) At Capacity test 2 of Vahana 23



(b) At Capacity test 8 of Vahana 23



(c) At Capacity test 13 of Vahana 23

Figure 11: Capacity Test plots of MDN-SOH for Vahana 23

8 Results - Estimating the distribution of Remaining Useful Life of eVTOL batteries

The Figure 12 shows the selected features and their occurrences across 21 tests performed for RUL distribution estimation. It can be observed that overall 25 features out of 30 features (83.33%) have been selected when all the tests are considered. Take-off cruise and landing voltage related features; cruise phase discharge capacity variance; duration of cc, cv ,cruise ; landing maximum temperature have been selected in all 21 tests. Interestingly, though cruise duration changes across mission profiles, the feature is not selected in any of RUL test runs, despite of considering tentative features for the tests. We can also observe that there is a need for more information as the training dataset size decreases (The number of capacity tests used for RUL distribution prediction are comparatively low to those used for SOH distribution prediction). This is also seen in the feature selection phase as there were more tentative features during each test. The number of tentative features can be reduced either increasing training dataset size or by increasing iterations (which is 100 in this thesis) during the feature selection phase. Landing maximum temperature, cc duration, takeoff- mean, minimum, variance of voltage are seen as important features when both SOH and RUL are considered.

The Table 11 shows the optimised hyperparameters chosen from the provided range of values in the grid for RFR, CNN-MCD and MDN for each mission profile. It can be observed that dropout rate is also 0.2 across all RUL distribution prediction tests performed when using CNN-MCD model. Moreover, the hyperparameter- number of mixtures is one in all test cases for MDN model.

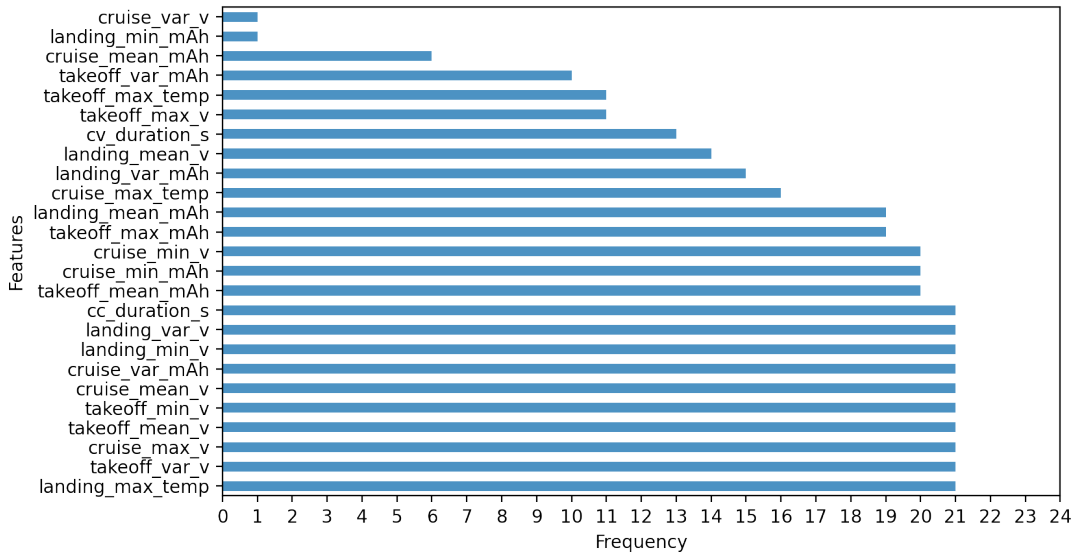


Figure 12: The selected features frequency across 21 tests for RUL distribution estimation

Table 11: Results — Hyperparameter Grid Search of each test for RUL, #Fs is number of features selected out of 30 features, ρ is dropout rate, BS is Batch Size

		RUL						
Vahana	**#Fs**	**RFR**		**CNN-MCD**		**MDN**		
		Max features	**Estimators**	ρ	**BS**	**Nodes**	**Mixtures**	**BS**
VAH01	15	0.5	2000	0.2	10	100	1	8
VAH02	19	0.25	100	0.2	8	100	1	8
VAH05	22	0.75	100	0.2	4	100	1	16
VAH06	20	0.5	100	0.2	8	200	1	16
VAH09	18	0.25	1000	0.2	14	100	1	8
VAH10	20	0.75	100	0.2	8	200	1	16
VAH11	20	0.25	850	0.2	12	200	1	14
VAH12	19	1	100	0.2	14	150	1	12
VAH13	18	0.5	100	0.2	4	200	1	16
VAH15	18	1	100	0.2	20	150	1	10
VAH16	17	0.75	100	0.3	4	100	1	10
VAH17	17	0.75	500	0.2	8	100	1	12
VAH20	18	1	500	0.2	20	150	1	16
VAH22	20	1	850	0.2	10	200	1	8
VAH23	19	1	500	0.2	14	50	1	8
VAH24	20	0.25	1500	0.2	12	150	1	12
VAH25	22	1	500	0.2	8	150	1	16
VAH26	23	0.75	750	0.2	12	150	1	10
VAH27	21	0.75	100	0.2	20	100	1	8
VAH28	20	1	500	0.2	8	200	1	16
VAH30	20	0.25	1000	0.2	12	200	1	10

The Table 12 shows the overall results of the models when all 21 mission profiles are considered. We can see that MDN performs better when MAE and RMSE are considered. When sharpness (Sh (%)) and CRPS is considered RFR performs better compared to other models. Apart from Sharpness (higher the better), lower the metric score better the model performance.

Table 12: Overall Performance Metrics of Models - RUL (#Missions)

Model	MAE	RMSE	SD	Sh(%)	CRPS_NRG	CRPS_PWM	CRPS_Norm
RFR	54.0454	60.9802	59.7341	**0.4847**	38.9977	**40.4908**	-
CNN-MCD	71.6918	81.7347	67.6708	0.4480	56.3163	56.6011	-
MDN	**52.9567**	**60.0267**	20.9756	0.3965	-	-	46.3952

8.1 Machine Learning algorithms - RUL results

The tables Table 13, Table 14, Table 15 show MAE, RMSE, Standard deviation (SD), Sharpness (Sh(%)), CRPS (in both NRG, PWM) for RFR, CNN-MCD and CRPS Normal distribution for the MDN that are obtained RUL distribution prediction for each mission profile. The last line of the each table gives the average values of the error metrics, which shows the performance metrics at model level as discussed in 6.7. Similar to SOH performance evaluation, CRPS is considered for the distribution prediction evaluation of RUL as well.

Table 13: Performance Metrics of RFR- RUL (#Missions)

Vahana	MAE	RMSE	SD	Sh(%)	CRPS_NRG	CRPS_PWM
VAH01	101.2073	110.0351	73.4708	0.2505	72.4820	72.5855
VAH02	31.8509	38.2626	74.1758	0.5336	24.6569	27.4073
VAH05	27.8556	38.6344	68.3364	0.7169	20.5625	24.2578
VAH06	44.4477	50.9569	59.3068	0.6392	26.6822	29.3904
VAH09	16.2000	18.6820	50.7039	0.7080	12.1733	12.4443
VAH10	49.0908	54.7384	74.0349	0.4708	37.7591	40.4574
VAH11	209.8720	242.0483	107.3770	0.1772	158.1851	158.8946
VAH12	74.7806	81.7272	67.0893	0.3844	51.0774	54.1858
VAH13	16.4114	18.6104	56.0259	0.6871	13.6650	16.9696
VAH15	26.9290	30.8919	41.5549	0.4770	13.8693	16.4336
VAH16	48.3740	51.8946	57.1012	0.5110	28.7391	31.5178
VAH17	32.0647	36.1719	41.8939	0.3562	20.9673	21.5147
VAH20	17.7334	19.3879	54.5948	0.7046	10.4176	10.9068
VAH22	22.8102	26.0814	42.9397	0.5540	11.0258	11.3226
VAH23	134.8928	154.8124	49.4723	0.1092	112.0264	112.6017
VAH24	17.4663	24.3962	56.6139	0.7884	11.6274	11.8149
VAH25	20.6496	29.9361	35.1964	0.5565	14.7799	15.2719
VAH26	28.6522	32.1528	47.2832	0.5262	17.8348	18.2966
VAH27	15.9630	18.4381	42.3609	0.4720	12.3101	14.8834
VAH28	48.2048	52.8723	51.9876	0.4892	29.0405	29.6745
VAH30	149.4980	149.8532	58.5781	0.0670	119.0710	119.4755
Average	54.0454	60.9802	59.7341	**0.4847**	38.9977	**40.4908**

The average values that are highlighted in bold indicate that the particular model has performed better than others when that particular metric is considered. It can be observed that the RFR model once again has low average CRPS score of 40.49 missions (see Table 13) when compared to other CNN-MCD and MDN models. It can also be noticed that MDN model has low average MAE (52.96 mission) and RMSE (60.03 missions) when compared to other RUL models and average CRPS of 46.40 missions. However, MDN model has low average sharpness

(0.39%), while RFR has average sharpness of 0.49%. Once Again, it can be affirmed that the selection of the CRPS metric for performance assessment is well-reasoned.

It is observed that Vahana 11 and Vahana 23 once again have consistent higher CRPS scores (highlighted in blue) across all the models for RUL distribution prediction when compared to other Vahanas. Additionally, Vahana 30, Vahana 12 also have high CRPS scores (highlighted in blue) in the RFR and MDN models (see Table 13, Table 15) respectively. Vahana 22 (highlighted in green) in the tables Table 13, Table 14, Table 15 has lowest CRPS score compared to other mission profiles of the models. Futhermore, Vahana 15 of MDN model has lowest CRPS scores across all mission profiles and the models considered together.

Table 16 shows the intermediate capacity test results obtained for RUL distribution prediction when RFR is considered. In general, the distributional means of Vahanas decease as mission cycles (capacity tests) increase. However, there is no specific trend or pattern in the error observed across mission profiles.

Table 14: Performance Metrics of CNN-MCD - RUL (#Missions)

Vahana	MAE	RMSE	SD	Sh(%)	CRPS_NRG	CRPS_PWM
VAH01	122.3248	138.1907	51.3755	0.1318	97.7609	97.9229
VAH02	28.3588	32.4176	53.4984	0.5272	21.0508	21.2957
VAH05	31.5145	38.8637	69.2814	0.5401	25.6767	26.0145
VAH06	36.6682	40.6970	63.1059	0.5598	25.6812	25.9352
VAH09	80.2861	90.7994	59.2668	0.3531	58.9167	59.2608
VAH10	80.8576	88.6506	44.9453	0.1865	60.0038	60.2098
VAH11	394.9753	442.2825	141.3388	0.1693	323.5092	324.2857
VAH12	102.9372	116.9828	107.2251	0.4663	70.2954	70.7094
VAH13	29.6059	39.7854	60.8891	0.5645	23.6097	23.9023
VAH15	28.7993	32.1882	45.9808	0.4899	21.3853	21.5683
VAH16	125.5945	126.6358	82.7687	0.2909	84.0356	84.3652
VAH17	48.4432	55.7212	60.3411	0.4999	33.8502	34.1057
VAH20	25.3670	27.7448	48.1857	0.5653	18.2994	18.5237
VAH22	14.1111	17.9240	44.7232	0.6170	13.5879	13.8016
VAH23	166.9776	202.4107	32.6086	0.1010	151.5445	151.6386
VAH24	26.9201	35.1902	59.3176	0.5506	23.6660	23.9127
VAH25	12.5486	17.5976	49.0488	0.6156	14.4915	14.7302
VAH26	73.7627	79.6367	68.8371	0.4067	49.9994	50.3638
VAH27	17.4350	19.6295	49.0265	0.5940	14.8795	15.1111
VAH28	29.5795	35.7770	68.1490	0.5934	23.9333	24.2609
VAH30	28.4616	37.3023	73.9334	0.5854	26.4647	26.7058
Average	71.6918	81.7347	67.6708	0.4480	56.3163	56.6011

Table 15: Performance Metrics of MDN - RUL (#Missions)

Vahana	MAE	RMSE	SD	Sh(%)	CRPS_Norm
VAH01	80.2190	88.7635	17.4119	0.0769	72.2546
VAH02	12.7260	19.2338	12.8361	0.4545	10.3034
VAH05	30.9431	34.3286	18.4245	0.1875	23.1048
VAH06	22.9620	27.5295	17.8210	0.4615	16.5702
VAH09	17.1585	26.0072	18.7861	0.7500	13.0431
VAH10	58.5645	62.2781	13.6486	0.0769	52.3123
VAH11	215.9625	254.4999	3.8126	0.0000	214.4782
VAH12	165.7678	177.0223	27.9809	0.0625	151.2610
VAH13	18.5581	22.5335	22.2280	0.5714	13.0432
VAH15	6.5096	7.8905	16.7846	1.0000	4.9539
VAH16	41.5910	42.8221	30.6337	0.2000	27.9745
VAH17	32.2688	37.1670	18.1933	0.2500	24.1730
VAH20	15.5646	17.7622	29.6510	0.9000	10.7450
VAH22	5.7982	7.8654	17.1229	0.8000	5.7786
VAH23	160.9350	185.9561	17.0125	0.0833	152.8219
VAH24	20.4222	24.6787	17.7923	0.6667	14.4973
VAH25	26.0248	32.8273	19.2959	0.4545	19.0496
VAH26	31.4220	34.8206	27.8823	0.3636	21.0288
VAH27	30.5761	32.7676	19.1407	0.1000	23.8076
VAH28	18.9275	23.0966	33.3936	0.8667	13.6349
VAH30	99.1887	100.7111	17.9730	0.0000	89.4632
Average	**52.9567**	**60.0267**	20.9756	0.3965	46.3952

Capacity test across missions										
Vahana	RUL	1	3	5	7	9	11	13	15	17
VAH01	Actual	612	510	408	306	204	102	0	-	-
	Mean	440.48	374.39	280.79	211.62	73.46	34.73	14.69	-	-
	Error	171.52	135.61	127.21	94.38	130.54	67.27	-14.69	-	-
	RMSE	200.54	164.41	154.13	133.06	138.45	79.20	29.71	-	-
	SD	103.92	92.94	87.03	93.79	46.10	41.81	25.83	-	-
	CRPS	117.57	101.54	93.64	64.82	105.55	46.35	3.55	-	-
VAH02	Actual	510	408	306	204	102	0	-	-	-
	Mean	488.36	454.19	342.04	271.25	103.53	11.75	-	-	-
	Error	21.64	-46.19	-36.04	-67.25	-1.53	-11.75	-	-	-
	RMSE	75.67	86.86	99.66	121.88	50.23	36.48	-	-	-
	SD	72.51	73.56	92.91	101.65	50.21	34.53	-	-	-
	CRPS	24.96	33.49	31.50	49.57	6.76	1.28	-	-	-
VAH05	Actual	765	663	561	459	357	255	153	51	-
	Mean	655.50	629.56	568.10	444.51	341.66	263.69	155.93	50.23	-
	Error	109.50	33.44	-7.10	14.49	15.34	-8.69	-2.93	0.77	-
	RMSE	128.34	73.83	77.52	77.00	75.94	76.25	72.14	38.84	-
	SD	66.93	65.82	77.20	75.62	74.37	75.75	72.08	38.84	-
	CRPS	80.45	24.19	21.34	18.90	18.62	19.87	21.34	2.69	-
VAH06	Actual	612	510	408	306	204	102	0	-	-

Table 16 – *Continued from previous page*

Vahana	RUL	1	3	5	7	9	11	13	15	17
VAH09	Mean	511.26	455.04	354.24	251.75	180.10	128.06	34.20	-	-
	Error	100.74	54.96	53.76	54.25	23.90	-26.06	-34.20	-	-
	RMSE	116.57	73.12	74.95	89.70	51.26	72.69	62.74	-	-
	SD	58.65	48.23	52.23	71.44	45.34	67.86	52.60	-	-
	CRPS	75.38	38.20	29.57	35.53	12.06	18.51	10.11	-	-
	Actual	516	414	351	327	303	279	153	51	-
	Mean	489.99	411.16	335.60	314.75	270.51	264.70	162.71	43.52	-
	Error	26.01	2.84	15.40	12.25	32.49	14.30	-9.71	7.48	-
	RMSE	56.24	38.11	58.37	57.00	65.92	53.52	56.63	50.50	-
	SD	49.86	38.00	56.30	55.67	57.36	51.58	55.79	49.94	-
	CRPS	11.73	5.33	14.28	15.95	15.37	15.11	9.20	12.59	-
VAH10	Actual	612	510	408	306	204	102	0	-	-
	Mean	506.60	461.21	363.24	289.72	152.70	72.06	59.82	-	-
	Error	105.40	48.79	44.76	16.28	51.30	29.94	-59.82	-	-
	RMSE	132.10	92.09	88.77	73.11	98.10	71.69	73.30	-	-
	SD	79.63	78.11	76.66	71.27	83.62	65.14	42.37	-	-
	CRPS	85.51	42.35	38.61	21.22	38.01	25.76	39.26	-	-
VAH11	Actual	815	713	611	509	407	305	203	101	0
	Mean	732.97	723.02	700.32	671.63	642.45	589.43	503.36	458.30	310.52
	Error	82.03	-10.02	-89.32	-162.63	-235.45	-284.43	-300.36	-357.30	-310.52
	RMSE	95.20	49.87	108.71	184.86	257.81	310.36	330.85	386.09	345.17
	SD	48.33	48.85	61.96	87.88	105.02	124.19	138.72	146.31	150.74
	CRPS	61.35	13.12	64.76	118.60	178.80	215.76	222.01	274.34	227.55
VAH12	Actual	765	663	561	459	357	255	153	51	-
	Mean	649.80	573.68	435.61	348.41	231.43	186.13	108.38	26.01	-
	Error	115.20	89.32	125.39	110.59	125.57	68.87	44.62	24.99	-
	RMSE	131.37	117.32	155.32	130.61	163.02	100.62	61.49	39.17	-
	SD	63.15	76.07	91.67	69.49	103.96	73.36	42.31	30.17	-
	CRPS	91.26	66.10	99.83	84.95	75.04	42.95	25.12	14.64	-
VAH13	Actual	663	561	459	357	255	153	51	-	-
	Mean	687.68	581.75	445.22	330.93	264.55	122.36	44.88	-	-
	Error	-24.68	-20.75	13.78	26.07	-9.55	30.64	6.12	-	-
	RMSE	84.16	60.48	59.04	59.25	45.02	63.16	40.80	-	-
	SD	80.46	56.81	57.40	53.20	44.00	55.23	40.34	-	-
	CRPS	41.41	19.86	16.67	27.50	10.62	18.31	4.24	-	-
VAH15	Actual	459	357	255	153	51	-	-	-	-
	Mean	466.16	399.26	278.33	153.05	80.17	-	-	-	-
	Error	-7.16	-42.26	-23.33	-0.05	-29.17	-	-	-	-
	RMSE	43.90	63.61	47.90	41.49	48.78	-	-	-	-
	SD	43.31	47.55	41.83	41.49	39.10	-	-	-	-
	CRPS	8.57	24.53	13.16	5.99	10.70	-	-	-	-
VAH16	Actual	459	357	255	153	51	-	-	-	-
	Mean	499.89	435.27	314.85	195.57	76.26	-	-	-	-
	Error	-40.89	-78.27	-59.85	-42.57	-25.26	-	-	-	-
	RMSE	76.09	97.86	90.68	67.48	47.49	-	-	-	-
	SD	64.16	58.75	68.12	52.36	40.21	-	-	-	-

Table 16 – *Continued from previous page*

Vahana	RUL	1	3	5	7	9	11	13	15	17
	CRPS	27.24	50.72	48.72	23.51	7.95	-	-	-	-
VAH17	Actual	561	459	357	255	153	51	-	-	-
	Mean	499.21	426.45	333.32	253.10	174.93	86.45	-	-	-
	Error	61.79	32.55	23.68	1.90	-21.93	-35.45	-	-	-
	RMSE	72.79	56.67	49.07	47.05	51.61	51.75	-	-	-
	SD	38.46	46.38	42.98	47.01	46.71	37.70	-	-	-
	CRPS	47.48	22.71	13.09	10.43	14.94	17.55	-	-	-
VAH20	Actual	459	357	255	153	51	-	-	-	-
	Mean	445.99	377.11	257.81	164.46	72.23	-	-	-	-
	Error	13.01	-20.11	-2.81	-11.46	-21.23	-	-	-	-
	RMSE	64.66	57.96	69.16	53.58	46.76	-	-	-	-
	SD	63.34	54.36	69.11	52.34	41.67	-	-	-	-
	CRPS	10.72	10.88	12.95	8.03	9.83	-	-	-	-
VAH22	Actual	459	357	255	153	51	-	-	-	-
	Mean	461.70	373.19	282.77	176.13	72.22	-	-	-	-
	Error	-2.70	-16.19	-27.77	-23.13	-21.22	-	-	-	-
	RMSE	56.23	52.94	49.25	46.28	42.61	-	-	-	-
	SD	56.16	50.40	40.67	40.09	36.94	-	-	-	-
	CRPS	5.22	8.88	14.86	11.52	5.69	-	-	-	-
VAH23	Actual	561	459	357	255	153	51	-	-	-
	Mean	335.97	310.09	292.53	277.02	270.50	265.91	-	-	-
	Error	225.03	148.91	64.47	-22.02	-117.50	-214.91	-	-	-
	RMSE	230.49	157.36	81.91	53.51	127.13	220.80	-	-	-
	SD	49.86	50.86	50.53	48.77	48.55	50.67	-	-	-
	CRPS	199.41	123.32	44.34	14.44	91.80	187.85	-	-	-
VAH24	Actual	561	459	357	255	153	51	-	-	-
	Mean	510.39	452.49	355.26	243.13	165.56	62.33	-	-	-
	Error	50.61	6.51	1.74	11.87	-12.56	-11.33	-	-	-
	RMSE	74.68	54.44	57.32	53.72	57.60	69.47	-	-	-
	SD	54.92	54.04	57.29	52.39	56.22	68.54	-	-	-
	CRPS	27.30	8.81	6.30	7.33	9.24	13.09	-	-	-
VAH25	Actual	512	410	307	205	103	0	-	-	-
	Mean	508.52	431.48	315.66	179.26	23.47	8.47	-	-	-
	Error	3.48	-21.48	-8.66	25.74	79.53	-8.47	-	-	-
	RMSE	35.71	52.32	28.89	45.19	84.56	21.52	-	-	-
	SD	35.54	47.71	27.57	37.14	28.72	19.78	-	-	-
	CRPS	4.73	9.30	4.95	12.32	66.73	1.32	-	-	-
VAH26	Actual	612	510	408	255	153	0	-	-	-
	Mean	601.50	539.35	442.11	275.93	197.50	30.39	-	-	-
	Error	10.50	-29.35	-34.11	-20.93	-44.50	-30.39	-	-	-
	RMSE	59.75	65.72	70.04	41.07	57.90	45.23	-	-	-
	SD	58.82	58.80	61.18	35.33	37.03	33.49	-	-	-
	CRPS	13.04	17.80	19.05	9.88	32.27	13.64	-	-	-
VAH27	Actual	510	408	306	153	51	-	-	-	-
	Mean	515.22	418.60	337.36	148.98	26.02	-	-	-	-
	Error	-5.22	-10.60	-31.36	4.02	24.98	-	-	-	-

Table 16 – *Continued from previous page*

Vahana	RUL	1	3	5	7	9	11	13	15	17
VAH28	RMSE	50.49	47.13	49.07	39.46	37.14	-	-	-	-
	SD	50.22	45.92	37.74	39.25	27.49	-	-	-	-
	CRPS	19.18	19.62	26.44	6.54	13.55	-	-	-	-
	Actual	720	618	516	413	311	209	107	0	-
	Mean	634.58	577.00	487.66	350.39	258.00	129.94	86.63	26.48	-
	Error	85.42	41.00	28.34	62.61	53.00	79.06	20.37	-26.48	-
	RMSE	103.53	68.62	59.08	86.50	73.51	91.82	53.00	43.24	-
	SD	58.50	55.02	51.84	59.69	50.94	46.71	48.93	34.18	-
	CRPS	59.23	20.78	14.13	35.93	30.47	56.25	14.94	9.89	-
VAH30	Actual	510	408	306	204	102	0	-	-	-
	Mean	644.38	559.86	470.68	344.42	261.73	147.11	-	-	-
	Error	-134.38	-151.86	-164.68	-140.42	-159.73	-147.11	-	-	-
	RMSE	152.10	159.97	174.23	152.15	169.86	157.42	-	-	-
	SD	71.26	50.29	56.91	58.59	57.78	56.04	-	-	-
	CRPS	96.83	126.15	135.90	111.85	129.79	116.63	-	-	-

Table 16: Results — Distribution Estimation of RUL (#Missions) at each capacity test using RF Regression, - implies Non-applicable, SD- Standard Deviation, CRPS is in PWM form

8.2 Discussion

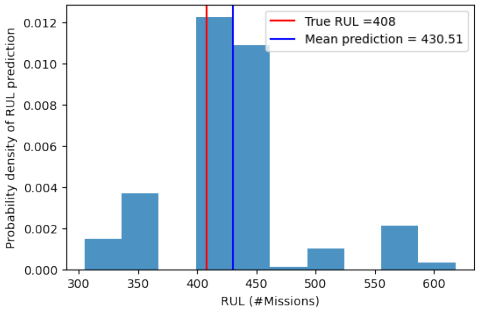
In this section, we examine the distribution of Remaining Useful Life (RUL) derived from all the models. We opt to display plots for each model to clarify how the distribution is depicted across different models. We choose intermediate capacity tests of Vahana 20 and 11, as they have similar CRPS metrics (low, high) across the models.

8.2.1 Illustrations- Vahana 20

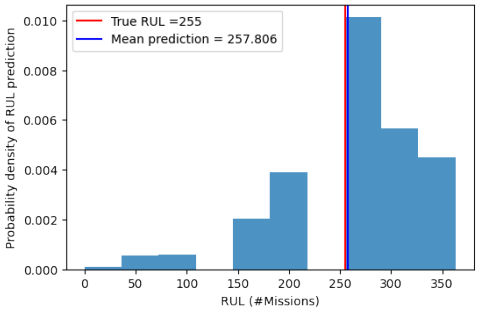
Figures 13a, 13b, 13c show the distribution of Vahana 20 at intermediate capacity tests 2, 5, 9 respectively, of the RFR model. The model consistently underestimates the mean RUL when the capacity tests 2, 5, and 9 of Vahana 20 are taken into account. This can be regarded as a positive aspect, as it aligns with the objective to prevent fatal system failures that could arise if the predictions are overestimated, particularly in situations where human safety is of high concern. The standard deviation (SD) values for the respective capacity tests are 59.04 , 69.11, 41.67 (refer vahana 20 in Table 18) implying an overall decreasing trend. However, the distributions are quite uneven.

The CNN-MCD model for the same vahana at the same capacity test positions with the CNN-MCD model, the model has underestimated the mean RUL of the distribution at capacity test 2 and started to overestimate from capacity test 5 of vahana 20. (refer Figures 14a, 14b, 14c). The uncertainty of distributions trend to decrease as the capacity tests increase (SD values at capacity tests 2, 5, 9 are 66.09, 51.24, 30.29, respectively). Moreover, the sharpness also decreases from 0.65 to 0.43 for vahana 20 (vahana 20 in Table 20).

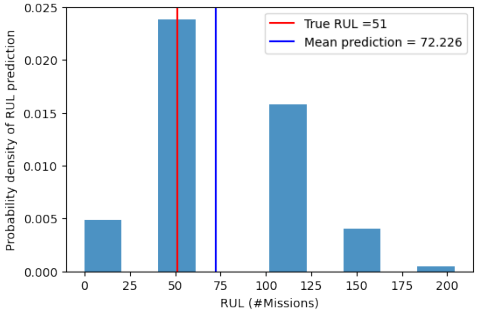
In case of MDN model, the uncertainity increases from 24.68 to 34.91 for the capacity tests (2, 5, 9) of vahana 20 (refer Table 22) and the model has always overestimated the mean RUL for the respective capacity tests of vahana 20 (refer Figures 15a, 15b, 15c).



(a) At Capacity test 2 of Vahana 20

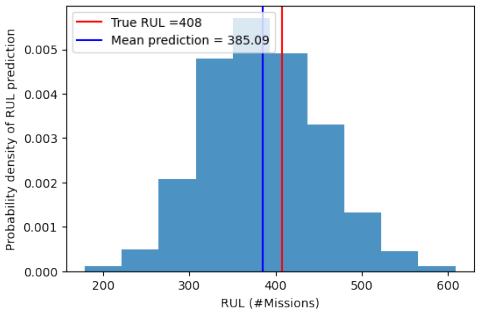


(b) At Capacity test 5 of Vahana 20

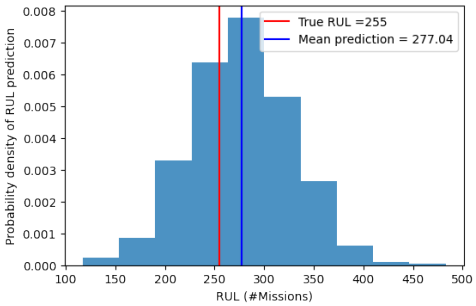


(c) At Capacity test 9 of Vahana 20

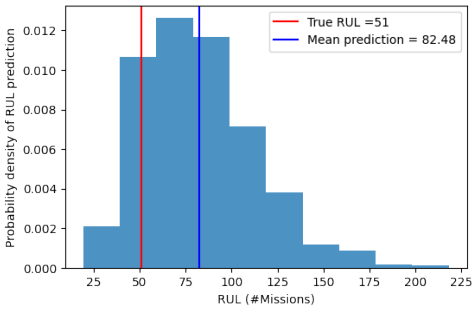
Figure 13: Capacity Test plots of RFR-RUL for Vahana 20



(a) At Capacity test 2 of Vahana 20

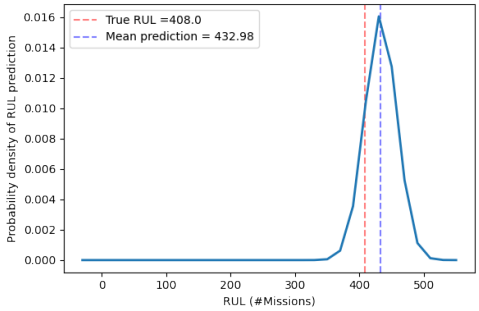


(b) At Capacity test 5 of Vahana 20

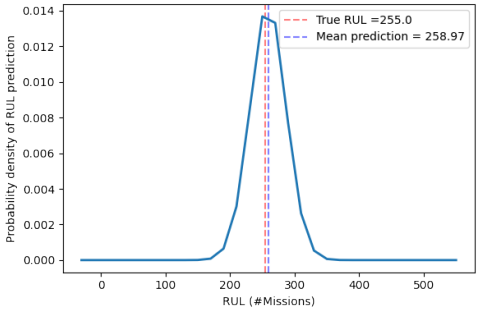


(c) At Capacity test 9 of Vahana 20

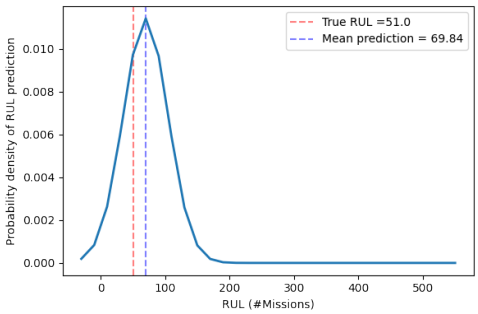
Figure 14: Capacity Test plots of CNN-MCD-RUL for Vahana 20



(a) At Capacity test 2 of Vahana 20



(b) At Capacity test 5 of Vahana 20



(c) At Capacity test 9 of Vahana 20

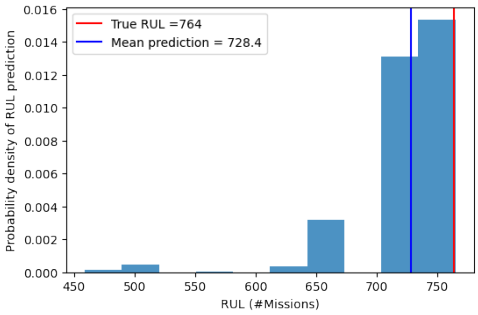
Figure 15: Capacity Test plots of MDN-RUL for Vahana 20

8.2.2 Illustrations- Vahana 11

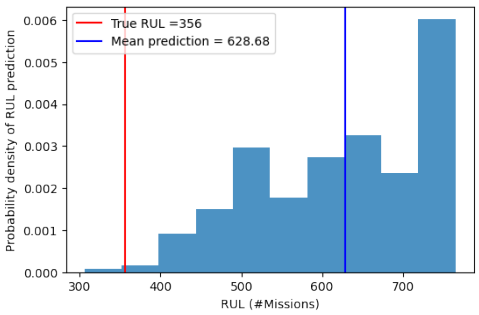
In case Vahana 11, we look at capacity tests 2, 10, 16. For RFR model, Figures 16a, 16b, 16c show the distribution of Vahana 11 at intermediate capacity tests 2, 10, 16 respectively. The error from the true value and the mean predicted value is high across the capacity tests. This is reflected in increase in uncertainty from 47.88 to 147.73 and decrease in sharpness from 0.58 to 0.02 (refer vahana 11 in Table 18).

The CNN-MCD model also exhibits the same behaviour of having high error from the true value and the mean predicted value is high across the capacity tests Figures 17a, 17b, 17c. However, the uncertainty decreases from 153.05 to 142.69 to 131.51 for the capacity tests 2, 10, 16 of the vahana 11.The sharpness of decreases from 0.64 to 0 (vahana 11 in Table 20). This is also indicated in the plots, the true value is completely out of the predicted distribution.

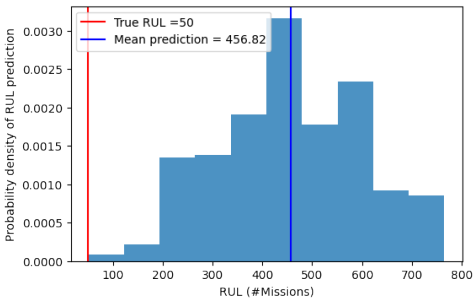
This is also the case with MDN model, the uncertainity increases from 0.19 to 0.01 for the capacity tests 2, 10, 16 of the vahana 11 (Table 22) and the true value is always out of the predicted distribution, this is also reflected in sharpness value, as it is 0 for vahana 11 (Figures 18a, 18b). The plot for capacity test 16 is not shown, as standard deviation is 0.01, the distribution collapsed to a point estimate and error is high.



(a) At Capacity test 2 of Vahana 11

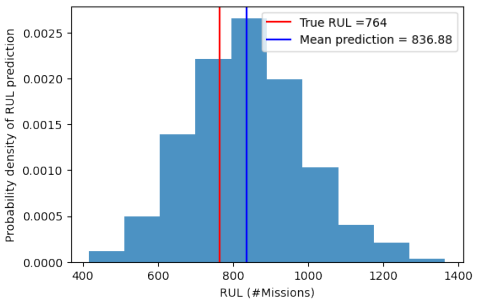


(b) At Capacity test 10 of Vahana 11

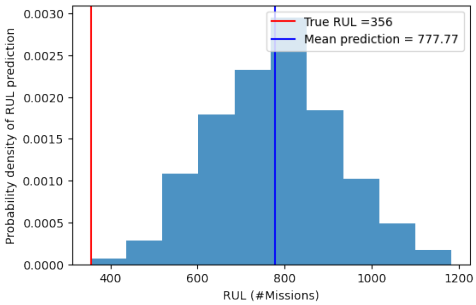


(c) At Capacity test 16 of Vahana 11

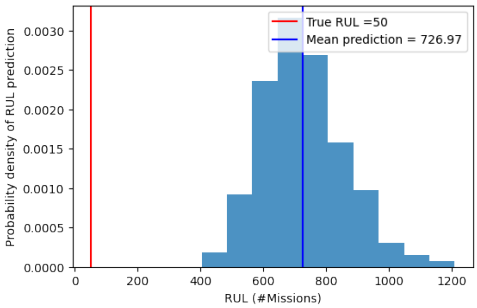
Figure 16: Capacity Test plots of RFR-RUL for Vahana 11



(a) At Capacity test 2 of Vahana 11

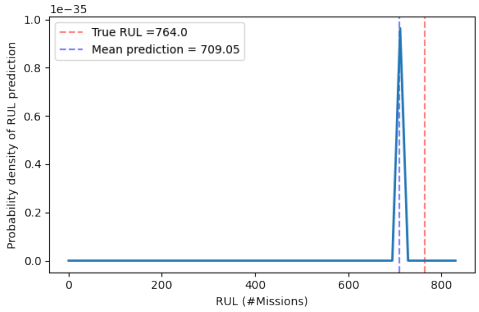


(b) At Capacity test 10 of Vahana 11

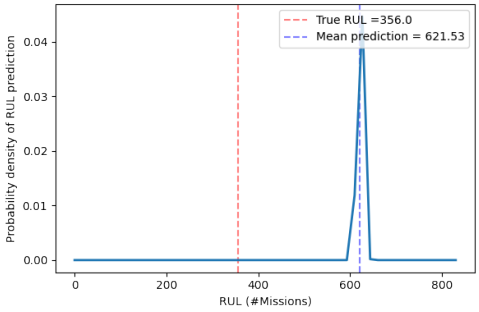


(c) At Capacity test 16 of Vahana 11

Figure 17: Capacity Test plots of CNN-MCD-RUL for Vahana 11



(a) At Capacity test 2 of Vahana 11



(b) At Capacity test 10 of Vahana 11

Figure 18: Capacity Test plots of MDN-RULfor Vahana 11

9 Conclusions and Future Work

A machine learning framework is proposed to estimate the distribution of State-of-Health (SOH) and Remaining Useful Lifetime (RUL) Electric Vertical Take-off and Landing vehicles (eVTOLs) batteries. To measure the uncertainty, two non-parametric approaches– Random Forest Regression and Convolution Neural Network with Monte Carlo dropout and one parametric method- Mixture Density Network are employed. The idea behind using Non-parametric methods is that they generate several point estimates, from which distributional mean (μ) and an associated standard deviation (σ) are derived, while the parametric approach assumes that the output follows a Normal Distribution and therefore directly gives an output mean and an associated standard deviation.

A dedicated lithium-ion batteries dataset that are cycled over several charge and discharge cycles to mimic the expected duty cycle of an electric aircraft is considered. Based on the charge and discharge cycles across the batteries, 30 features are generated. Given that these batteries operate in diverse conditions and the dataset is relatively moderate, a choice is made to treat each battery as a test case, resulting in 21 test cases for each machine learning algorithm. An automatic feature selection algorithm- Boruta SHAP is used to choose important features for each test case, with the idea of building a robust model by reducing sensitivity to irrelevant or noisy variables. This enhances the models' reliability and stability, leading to predictions that are more dependable and consistent. The selected features based on training data for each test case are then fed to the RFR, CNN-MCD, and MDN. The accuracy and sharpness of the obtained distributions from each model are evaluated using the Continuous Ranked Probability Score (CRPS). The results show that Random Forest Regression performs better than the other models and predicts the distribution of SOH and RUL with an average CRPS score of 0.96% and 40.49 missions/cycles, respectively. Moreover, the current methodology when Random forest regression is used, has a lower RMSE- 1.14% (SOH) and 60.98 missions (RUL) compared to the RMSE of the best models presented in [37] (RMSE of RFR for SOH is 1.80% and RMSE of XGBoost is 67.92 missions).

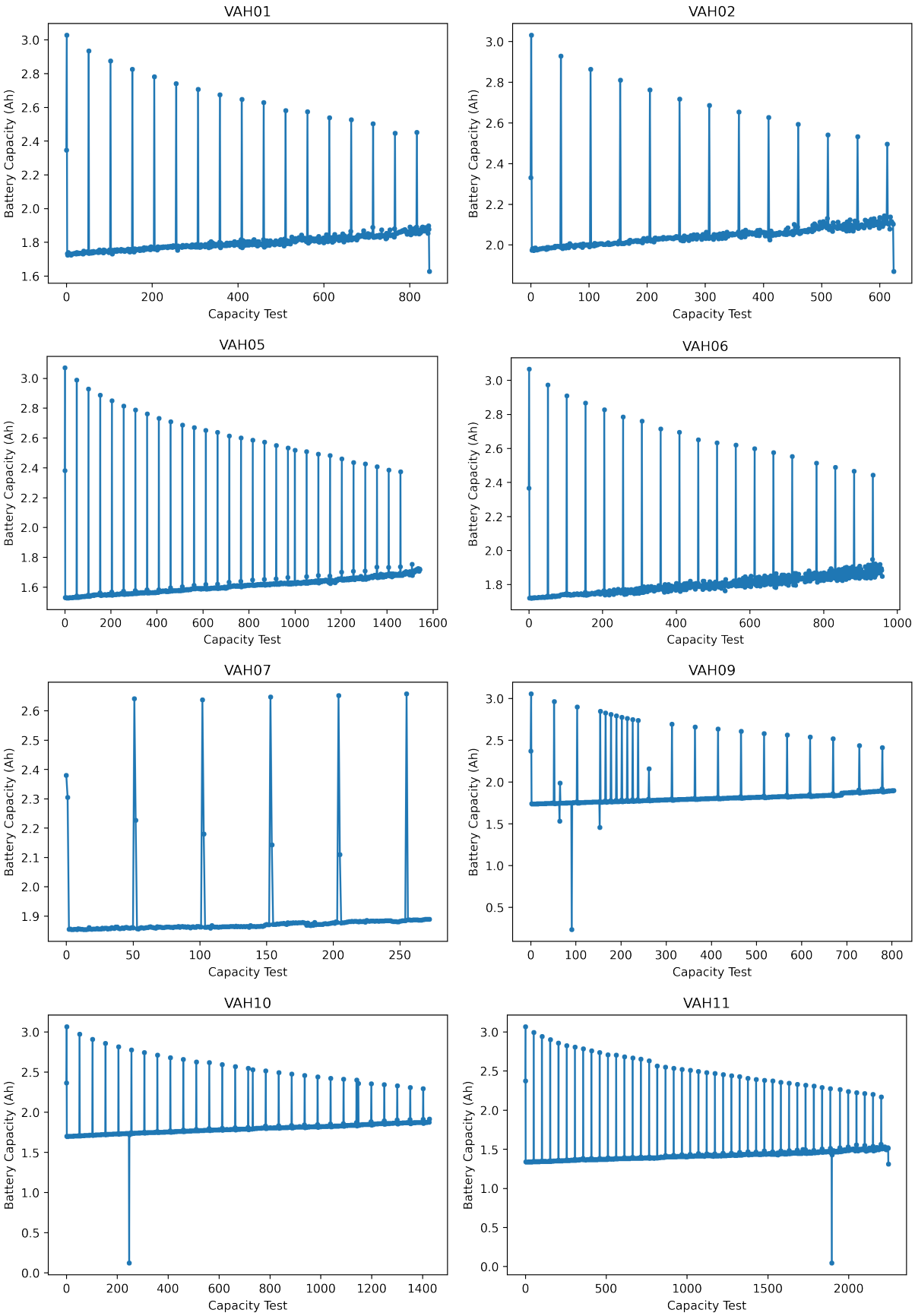
There are no specific patterns observed at the capacity test level. However, Vahana 11 and 23 consistently yield poor results both in the case of SOH and RUL distribution estimation and across machine learning models implemented in this thesis. Vahana 11 is operated under 20% power reduction during the discharge phase and Vahana 23 is operated under a reduced charge voltage of 4.1V. There is a possibility that the performance issues may be associated with the specific characteristics or behaviors of these individual Vahanas.

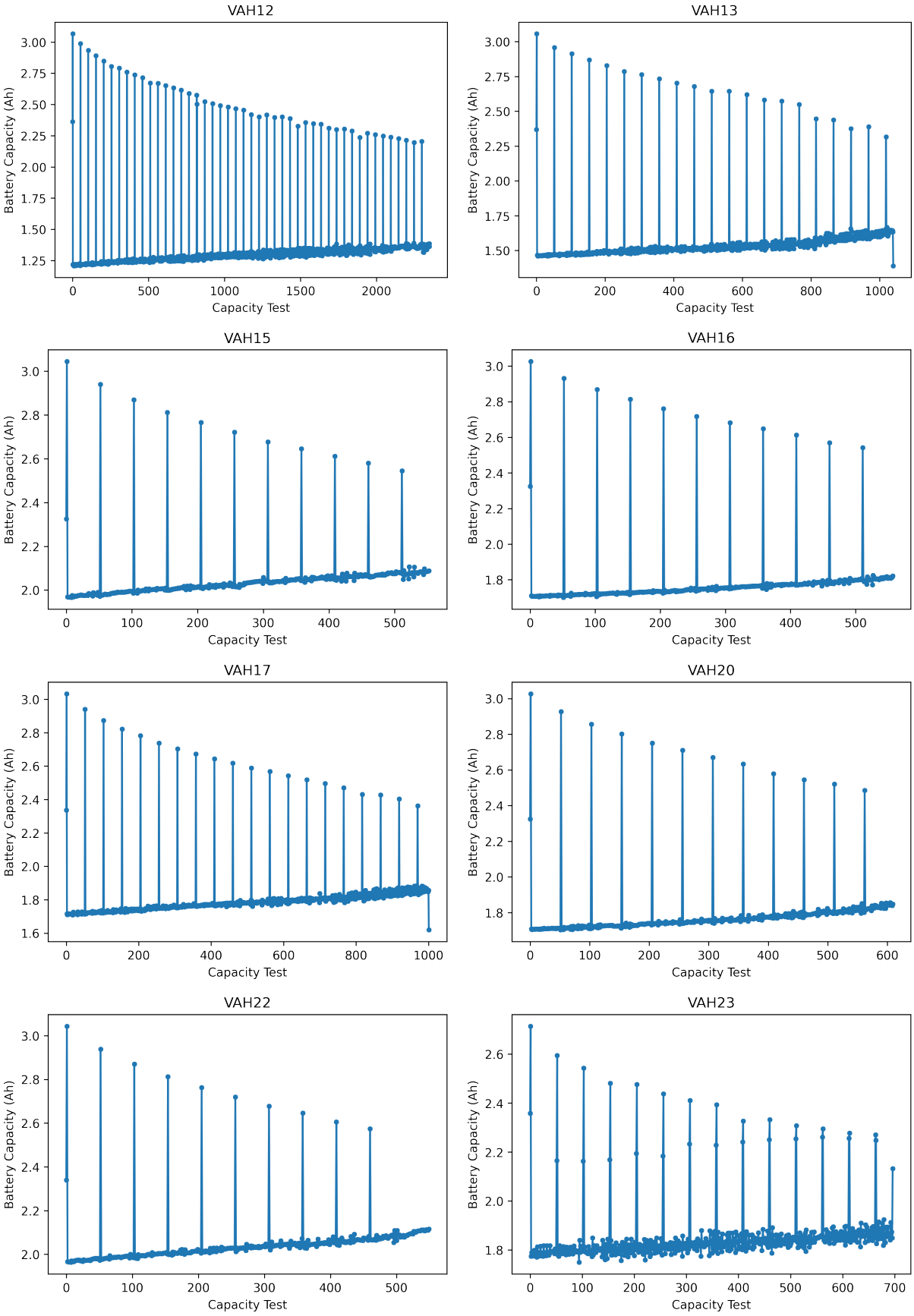
The limitations of this methodology is that it becomes computationally expensive if the data-set

size increases (in variety and a number of capacity tests), especially during the hyper-parameter tuning phase. Moreover, the methodology is based on a supervised machine learning paradigm and we have set the EOL-threshold to 85% of initial battery capacity in RUL distribution estimation. This might not be the case in real-world scenarios, there might not be enough data that has the entire life trajectory of a battery/batteries (Run-to-Failure Data). In the future, it will be beneficial to explicitly construct health indicator(s) for eVTOL batteries by using unsupervised learning model approaches or use Vahana 11 and 23 batteries operating conditions to threshold data, and subsequently, degradation models can be fitted to condition indicators extracted from sensor data, reflecting linear or exponential changes over time.

A Battery Capacity test results of all missions

In general, the peaks in the upward direction of battery capacity results across all missions are considered as capacity test missions, moreover, the selected capacity tests are further analysed and few of them are removed based on the observations made in 5.3 for later experiments done.


VAH01
VAH02
VAH05
VAH06
VAH07
VAH09
VAH10
VAH11
Battery Capacity (Ah)
Capacity Test


VAH12
VAH13
VAH15
VAH16
VAH17
VAH20
VAH22
VAH23
Battery Capacity (Ah)
Capacity Test

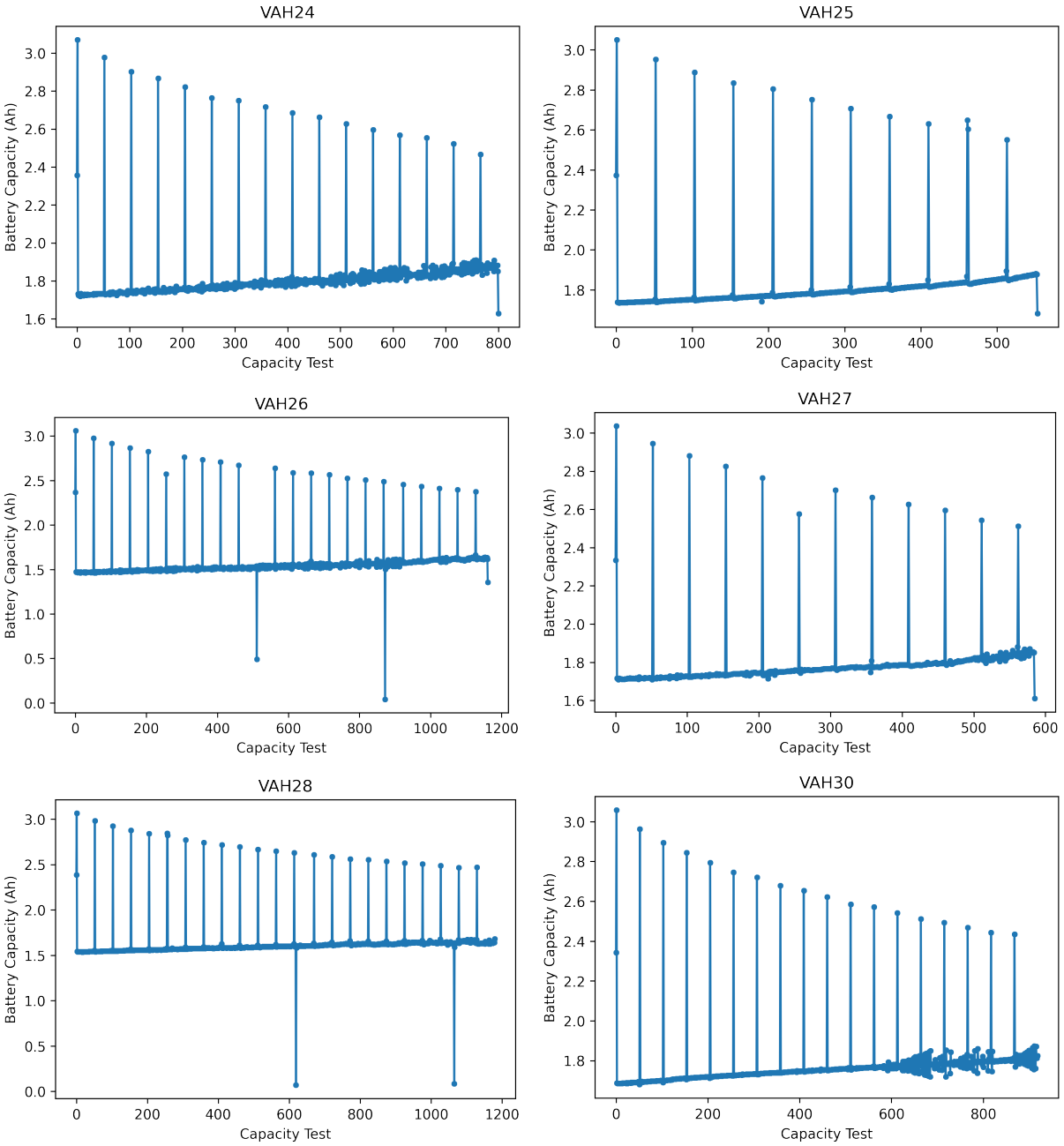
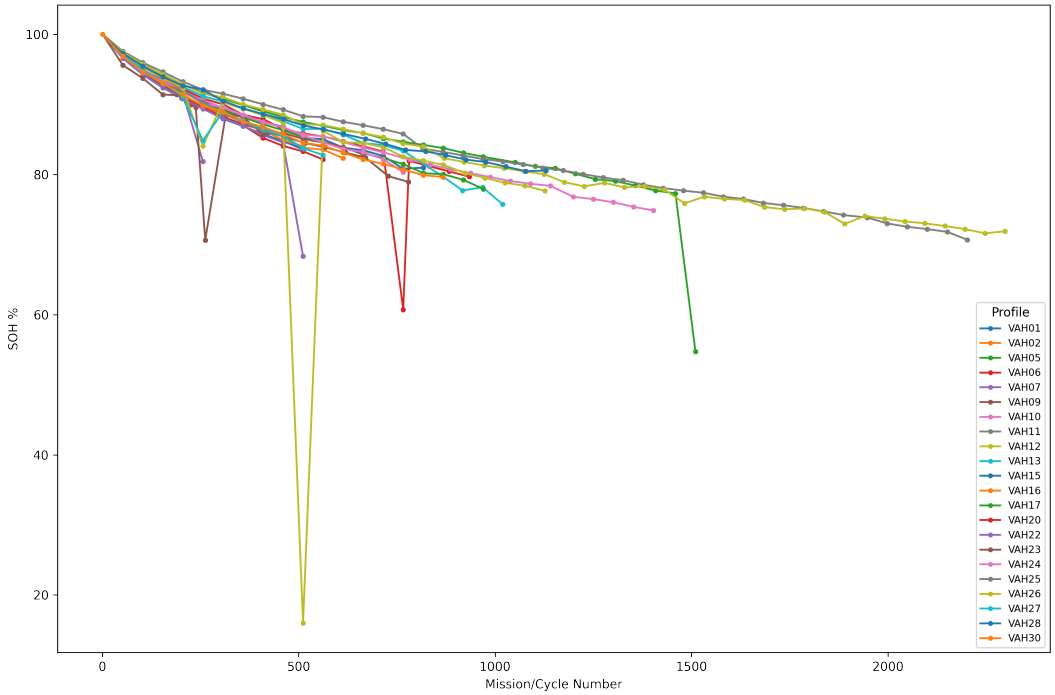
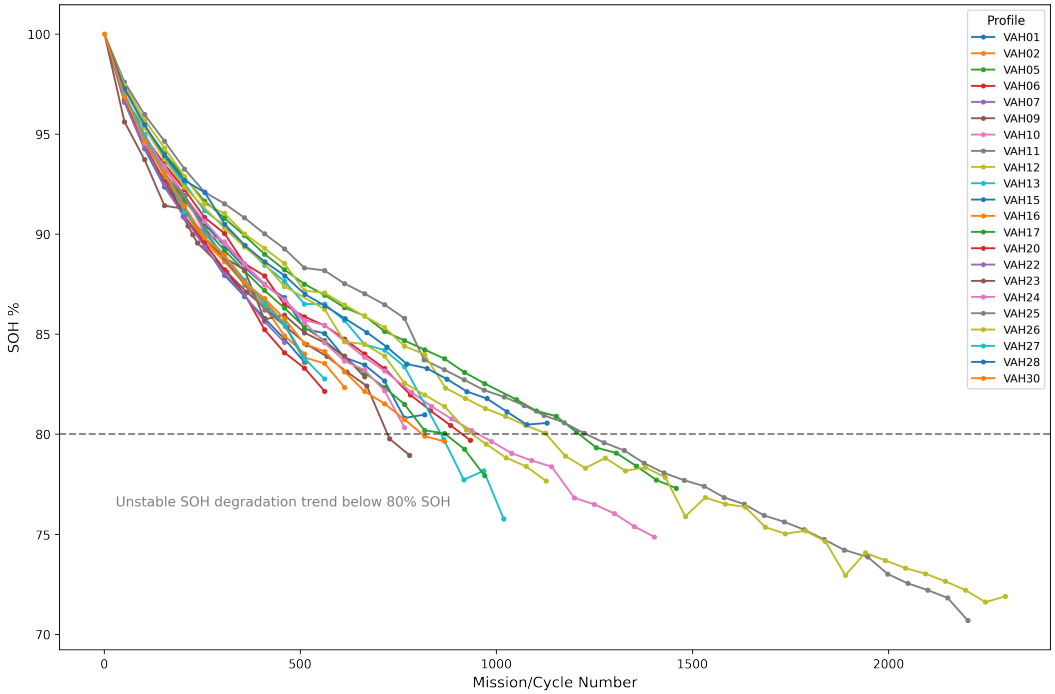


Figure 19: Battery capacity vs. mission cycle number of each mission profile

B Battery capacity (%) degradation trend



(a) SOH (%) of capacity test missions vs. mission cyles before cleaning



(b) SOH (%) of capacity test missions vs. mission cyles after cleaning

Figure 20: 20a and 20b shows consolidated battery capacity (%) degradation trend of each mission profile during the capacity tests before and after cleaning the data

C Random Forest Regression - SOH Capacity Test Results

Note that all results are rounded to 2. Mean indicates distribution mean prediction and SD is the standard deviation.

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH01	1	100	95.97	4.03	3.39	4.03	5.27	0.42	2.17	2.36
VAH01	2	96.95	94.92	2.03	2.63	2.85	3.32	0.47	1.39	1.58
VAH01	3	94.98	93.59	1.39	2.49	2.41	2.85	0.64	1.12	1.31
VAH01	4	93.33	92.8	0.53	2.9	2.3	2.95	0.71	0.81	0.99
VAH01	5	91.85	90.43	1.42	2.55	2.4	2.92	0.61	1.16	1.34
VAH01	6	90.5	88.79	1.71	2.47	2.53	3	0.52	1.36	1.54
VAH01	7	89.42	88.13	1.29	2.01	1.95	2.39	0.65	0.97	1.15
VAH01	8	88.34	87.14	1.2	0.79	1.23	1.44	0.25	0.8	0.98
VAH01	9	87.45	86.16	1.29	0.91	1.32	1.58	0.26	0.8	0.98
VAH01	10	86.83	86	0.83	1.19	1.14	1.45	0.61	0.48	0.65
VAH01	11	85.26	84.51	0.75	0.92	0.99	1.19	0.46	0.48	0.65
VAH01	12	85.04	83.86	1.18	0.93	1.29	1.51	0.26	0.79	0.96
VAH01	13	83.83	83.3	0.53	1.17	0.92	1.29	0.68	0.28	0.45
VAH01	14	83.46	82.24	1.22	1.1	1.33	1.65	0.35	0.73	0.89
VAH01	15	82.66	81.24	1.42	1.18	1.54	1.85	0.4	0.88	1.04
VAH01	16	80.81	78.89	1.92	1.29	1.97	2.3	0.34	1.26	1.42
VAH01	17	80.98	78.65	2.33	1.51	2.35	2.77	0.39	1.52	1.67
VAH02	1	100	98.34	1.66	1.89	1.66	2.52	0.53	0.68	0.81
VAH02	2	96.62	96.99	-0.37	1.69	1.17	1.73	0.7	0.31	0.44
VAH02	3	94.47	95.7	-1.23	1.9	1.64	2.26	0.61	0.64	0.76
VAH02	4	92.69	94.18	-1.49	2.18	2.07	2.64	0.51	0.88	1.01
VAH02	5	91.14	92.37	-1.23	2.21	1.83	2.53	0.75	0.66	0.78
VAH02	6	89.64	90.89	-1.25	2.15	1.7	2.49	0.67	0.61	0.73
VAH02	7	88.62	89.7	-1.08	1.89	1.57	2.17	0.79	0.62	0.74
VAH02	8	87.52	87.29	0.23	0.83	0.64	0.86	0.84	0.22	0.33
VAH02	9	86.68	86.56	0.12	1.14	0.85	1.14	0.74	0.25	0.37
VAH02	10	85.52	85.06	0.46	1.09	0.87	1.18	0.74	0.27	0.38
VAH02	11	83.83	83.69	0.14	1.2	0.9	1.21	0.67	0.23	0.35
VAH02	12	83.54	82.98	0.56	1.24	1.03	1.36	0.61	0.34	0.45
VAH02	13	82.34	81.26	1.08	1.09	1.28	1.53	0.47	0.66	0.77
VAH05	1	100	99.53	0.47	1.18	0.47	1.27	0.85	0.06	0.13
VAH05	2	97.33	97.82	-0.49	1.54	1.17	1.61	0.66	0.36	0.43
VAH05	3	95.41	96.48	-1.07	1.14	1.36	1.56	0.33	0.78	0.84
VAH05	4	94.02	95.27	-1.25	1.3	1.51	1.8	0.41	0.81	0.88
VAH05	5	92.84	94.73	-1.89	1.35	2.05	2.32	0.29	1.33	1.39
VAH05	6	91.65	92.59	-0.94	1.11	1.19	1.45	0.73	0.64	0.7
VAH05	7	90.79	91.8	-1.01	1.45	1.33	1.76	0.6	0.56	0.62
VAH05	8	89.92	91.23	-1.31	1.16	1.45	1.74	0.4	0.82	0.88
VAH05	9	89	90.05	-1.05	1.52	1.34	1.85	0.77	0.6	0.66
VAH05	10	88.22	88.89	-0.67	1.06	0.92	1.25	0.55	0.38	0.44
VAH05	11	87.5	88.08	-0.58	0.79	0.8	0.98	0.57	0.41	0.47
VAH05	12	86.95	86.95	0	1.07	0.69	1.07	0.73	0.14	0.2
VAH05	13	86.33	86.68	-0.35	1.01	0.81	1.07	0.78	0.3	0.36
VAH05	14	85.91	86.34	-0.43	1.19	1.01	1.27	0.69	0.38	0.44

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH05	15	85.14	84.54	0.6	1.14	0.84	1.29	0.73	0.24	0.29
VAH05	16	84.68	84.11	0.57	1.12	0.91	1.25	0.72	0.3	0.36
VAH05	17	84.22	83.9	0.32	1.08	0.76	1.13	0.72	0.2	0.26
VAH05	18	83.77	83.09	0.68	1.24	1.12	1.42	0.62	0.42	0.48
VAH05	19	83.09	82.88	0.21	1.22	0.99	1.23	0.61	0.3	0.35
VAH05	20	82.53	82.02	0.51	1.27	1.11	1.37	0.62	0.39	0.44
VAH05	21	81.73	80.67	1.06	0.69	1.13	1.27	0.25	0.76	0.82
VAH05	22	81.17	80.54	0.63	0.62	0.7	0.88	0.53	0.37	0.42
VAH05	23	80.89	80.63	0.26	0.81	0.68	0.85	0.58	0.26	0.31
VAH05	24	80.11	78.83	1.28	0.92	1.39	1.58	0.31	0.88	0.93
VAH05	25	79.33	78.5	0.83	0.72	0.91	1.1	0.45	0.52	0.58
VAH05	26	79.06	78.36	0.7	0.71	0.81	1	0.43	0.43	0.48
VAH05	27	78.41	77.12	1.29	0.95	1.34	1.6	0.52	0.81	0.86
VAH05	28	77.71	75.49	2.22	0.89	2.25	2.4	0.1	1.78	1.83
VAH05	29	77.3	75.34	1.96	0.85	1.99	2.14	0.12	1.52	1.57
VAH06	1	100	99.79	0.21	0.76	0.21	0.79	0.93	0.01	1.01
VAH06	2	97	97.88	-0.88	1.62	1.17	1.84	0.62	0.36	1.34
VAH06	3	94.88	95.38	-0.5	2.14	1.46	2.2	0.82	0.38	1.34
VAH06	4	93.5	94.54	-1.04	2.13	1.61	2.37	0.77	0.51	1.46
VAH06	5	92.22	92.37	-0.15	1.77	1.12	1.78	0.84	0.3	1.22
VAH06	6	90.84	91.14	-0.3	1.73	1.07	1.75	0.84	0.27	1.18
VAH06	7	90.03	90.5	-0.47	2.13	1.49	2.18	0.8	0.41	1.31
VAH06	8	88.52	89.83	-1.31	2.25	1.81	2.6	0.74	0.69	1.59
VAH06	9	87.91	88.16	-0.25	1.01	0.84	1.05	0.75	0.29	1.17
VAH06	10	86.47	86.67	-0.2	1.26	0.88	1.27	0.83	0.23	1.09
VAH06	11	85.87	86.47	-0.6	1.22	1.04	1.36	0.77	0.39	1.25
VAH06	12	85.43	83.84	1.59	1.06	1.65	1.92	0.31	1.09	1.93
VAH06	13	84.76	84.5	0.26	0.98	0.82	1.01	0.75	0.29	1.13
VAH06	14	84.01	83.66	0.35	1.03	0.89	1.08	0.66	0.33	1.17
VAH06	15	83.28	83.3	-0.02	1.02	0.73	1.02	0.77	0.19	1.02
VAH06	16	81.97	80.68	1.29	1.58	1.67	2.04	0.5	0.78	1.58
VAH06	17	81.18	80.32	0.86	1.42	1.38	1.66	0.6	0.58	1.38
VAH06	18	80.44	79.74	0.7	1.55	1.38	1.7	0.52	0.51	1.3
VAH06	19	79.69	78.45	1.24	1.88	1.97	2.25	0.53	0.91	1.7
VAH09	1	100	98.62	1.38	2.1	1.38	2.51	0.67	0.38	0.5
VAH09	2	96.98	95.29	1.69	1.18	1.77	2.06	0.25	1.15	1.26
VAH09	3	94.89	93.97	0.92	1.16	1.08	1.48	0.53	0.51	0.62
VAH09	4	93.18	92.34	0.84	1.01	0.98	1.31	0.56	0.48	0.59
VAH09	5	92.52	91.66	0.86	0.93	1.07	1.27	0.49	0.59	0.69
VAH09	6	91.93	91.48	0.45	0.93	0.79	1.03	0.61	0.31	0.42
VAH09	7	91.39	91.3	0.09	0.9	0.74	0.91	0.69	0.25	0.36
VAH09	8	90.87	90.49	0.38	0.77	0.73	0.86	0.68	0.35	0.46
VAH09	9	90.4	89.91	0.49	0.7	0.66	0.86	0.52	0.31	0.41
VAH09	10	89.98	90.13	-0.15	0.89	0.67	0.91	0.79	0.2	0.31
VAH09	11	89.56	89.86	-0.3	0.94	0.74	0.98	0.77	0.24	0.34
VAH09	12	88.11	87.63	0.48	0.68	0.74	0.82	0.48	0.38	0.48
VAH09	13	87.09	87.5	-0.41	0.84	0.72	0.94	0.76	0.28	0.38
VAH09	14	86.23	86.95	-0.72	0.78	0.87	1.06	0.44	0.45	0.55

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH09	15	85.35	84.99	0.36	0.62	0.62	0.72	0.37	0.29	0.39
VAH09	16	84.48	84.22	0.26	0.72	0.57	0.76	0.75	0.19	0.29
VAH09	17	83.89	83.57	0.32	0.74	0.57	0.81	0.67	0.17	0.27
VAH09	18	83.11	83.2	-0.09	0.62	0.5	0.63	0.66	0.16	0.26
VAH09	19	82.41	82.78	-0.37	0.92	0.8	0.99	0.61	0.32	0.42
VAH09	20	79.76	79.75	0.01	1.14	0.92	1.14	0.7	0.29	0.38
VAH09	21	78.95	78.67	0.28	1.07	0.94	1.11	0.66	0.34	0.44
VAH10	1	100	98.98	1.02	1.84	1.02	2.1	0.74	0.22	0.34
VAH10	2	97.01	96.95	0.06	1.96	1.41	1.96	0.52	0.36	0.48
VAH10	3	94.97	95.8	-0.83	2.24	2.05	2.39	0.78	0.81	0.92
VAH10	4	93.29	94.12	-0.83	2.39	2	2.53	0.71	0.67	0.79
VAH10	5	91.87	92.37	-0.5	1.78	1.52	1.85	0.54	0.55	0.66
VAH10	6	90.64	91.17	-0.53	1.51	1.4	1.59	0.54	0.58	0.69
VAH10	7	89.52	91	-1.48	1.24	1.7	1.93	0.22	1.06	1.16
VAH10	8	88.45	88.19	0.26	1.41	1.29	1.43	0.54	0.52	0.62
VAH10	9	87.46	87.65	-0.19	1.35	1.17	1.36	0.52	0.4	0.5
VAH10	10	86.72	87.09	-0.37	1.31	1.14	1.36	0.66	0.4	0.5
VAH10	11	85.7	85.23	0.47	1.29	1.18	1.38	0.56	0.47	0.57
VAH10	12	85.44	84.97	0.47	1.25	1.14	1.34	0.46	0.44	0.54
VAH10	13	84.65	84.19	0.46	1.31	1.16	1.39	0.58	0.42	0.52
VAH10	14	83.84	84.02	-0.18	1.36	1.21	1.37	0.5	0.45	0.55
VAH10	15	83.16	82.98	0.18	1.6	1.41	1.61	0.52	0.49	0.59
VAH10	16	82.1	82.44	-0.34	1.52	1.33	1.56	0.69	0.46	0.56
VAH10	17	81.39	82.08	-0.69	1.21	1.21	1.39	0.55	0.53	0.63
VAH10	18	80.76	81.84	-1.08	1.2	1.43	1.61	0.41	0.76	0.86
VAH10	19	80.19	80.61	-0.42	0.99	0.85	1.08	0.7	0.3	0.39
VAH10	20	79.63	80.31	-0.68	0.62	0.69	0.92	0.67	0.36	0.46
VAH10	21	79.05	80.27	-1.22	0.58	1.22	1.35	0.14	0.92	1.01
VAH10	22	78.68	80.02	-1.34	0.78	1.38	1.55	0.24	0.95	1.05
VAH10	23	78.38	79.51	-1.13	0.87	1.22	1.42	0.49	0.74	0.83
VAH10	24	76.81	78.55	-1.74	1.16	1.84	2.09	0.28	1.2	1.29
VAH10	25	76.49	77.94	-1.45	1.3	1.6	1.94	0.37	0.88	0.97
VAH10	26	76.03	77.41	-1.38	1.3	1.56	1.9	0.45	0.85	0.94
VAH10	27	75.38	76.6	-1.22	1.28	1.49	1.77	0.52	0.82	0.91
VAH10	28	74.87	76.31	-1.44	1.34	1.74	1.97	0.29	1.04	1.13
VAH11	1	100	99.42	0.58	1.13	0.58	1.27	0.79	0.12	0.45
VAH11	2	97.6	97.82	-0.22	1.35	0.98	1.37	0.7	0.32	0.65
VAH11	3	95.98	97	-1.02	1.22	1.33	1.59	0.33	0.74	1.06
VAH11	4	94.65	96.42	-1.77	1.51	1.94	2.33	0.45	1.14	1.46
VAH11	5	93.25	95.07	-1.82	1.54	1.97	2.38	0.41	1.12	1.44
VAH11	6	92.1	95.5	-3.4	2.45	3.44	4.19	0.44	2.07	2.39
VAH11	7	91.52	93.19	-1.67	1.56	1.8	2.29	0.56	0.98	1.29
VAH11	8	90.82	92.67	-1.85	1.46	1.92	2.36	0.42	1.14	1.44
VAH11	9	90.01	91.97	-1.96	1.54	1.99	2.49	0.37	1.16	1.47
VAH11	10	89.26	91.64	-2.38	1.44	2.4	2.78	0.24	1.62	1.92
VAH11	11	88.31	91.01	-2.7	1.79	2.73	3.24	0.34	1.76	2.07
VAH11	12	88.18	90.62	-2.44	1.69	2.45	2.97	0.35	1.54	1.84
VAH11	13	87.53	89.92	-2.39	1.83	2.41	3.01	0.45	1.43	1.73

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH11	14	87.02	89.6	-2.58	1.69	2.58	3.09	0.39	1.71	2.01
VAH11	15	86.48	88.77	-2.29	1.55	2.32	2.76	0.35	1.57	1.86
VAH11	16	85.79	88.84	-3.05	1.76	3.07	3.52	0.21	2.23	2.52
VAH11	17	83.71	87.9	-4.19	2.74	4.19	5	0.29	3.08	3.37
VAH11	18	83.23	85.38	-2.15	0.96	2.16	2.35	0.06	1.65	1.93
VAH11	19	82.72	84.95	-2.23	0.86	2.26	2.39	0.04	1.81	2.09
VAH11	20	82.2	84.78	-2.58	0.94	2.6	2.75	0.05	2.12	2.41
VAH11	21	81.86	84.28	-2.42	0.98	2.43	2.61	0.08	1.89	2.17
VAH11	22	81.43	83.54	-2.11	1.07	2.1	2.36	0.14	1.5	1.78
VAH11	23	80.94	83.32	-2.38	1.02	2.4	2.59	0.11	1.84	2.12
VAH11	24	80.58	83.31	-2.73	1.03	2.75	2.92	0.03	2.18	2.46
VAH11	25	80.05	82.53	-2.48	1.19	2.49	2.75	0.18	1.82	2.1
VAH11	26	79.57	81.53	-1.96	1.08	1.97	2.23	0.35	1.37	1.64
VAH11	27	79.19	81.65	-2.46	1.12	2.46	2.7	0.03	1.84	2.11
VAH11	28	78.55	80.39	-1.84	1.14	1.93	2.17	0.21	1.31	1.58
VAH11	29	78.06	80.27	-2.21	1.19	2.27	2.51	0.2	1.62	1.88
VAH11	30	77.7	80.02	-2.32	1.22	2.36	2.62	0.17	1.68	1.94
VAH11	31	77.4	80.37	-2.97	1.28	3	3.23	0.12	2.31	2.58
VAH11	32	76.84	79.26	-2.42	1.36	2.54	2.78	0.13	1.82	2.08
VAH11	33	76.5	78.85	-2.35	1.82	2.79	2.97	0.13	1.89	2.16
VAH11	34	75.94	77.82	-1.88	1.9	2.45	2.67	0.36	1.46	1.72
VAH11	35	75.62	78.26	-2.64	1.74	2.96	3.16	0.17	2.08	2.35
VAH11	36	75.23	78.09	-2.86	1.8	3.15	3.39	0.12	2.23	2.49
VAH11	37	74.74	77.6	-2.86	1.96	3.13	3.47	0.18	2.09	2.35
VAH11	38	74.21	77.44	-3.23	1.93	3.42	3.76	0.18	2.4	2.66
VAH11	39	73.86	77.28	-3.42	1.96	3.57	3.93	0.21	2.53	2.78
VAH11	40	73.01	76.39	-3.38	2.28	3.46	4.07	0.4	2.18	2.43
VAH11	41	72.54	76.23	-3.69	2.3	3.74	4.35	0.33	2.43	2.69
VAH11	42	72.21	76.43	-4.22	2.27	4.25	4.79	0.16	2.97	3.22
VAH11	43	71.82	76.38	-4.56	2.19	4.57	5.06	0.13	3.34	3.59
VAH11	44	70.69	75.78	-5.09	2.48	5.09	5.66	0.2	3.67	3.93
VAH12	1	100	99.78	0.22	0.94	0.22	0.97	0.94	0.01	1.01
VAH12	2	97.45	97.37	0.08	1.8	1.41	1.8	0.53	0.44	1.41
VAH12	3	95.71	94.68	1.03	1.56	1.51	1.87	0.57	0.66	1.6
VAH12	4	94.3	94.31	-0.01	1.63	1.15	1.63	0.79	0.32	1.27
VAH12	5	92.89	91.9	0.99	1.31	1.25	1.64	0.61	0.52	1.44
VAH12	6	91.51	91.35	0.16	0.74	0.57	0.76	0.82	0.19	1.1
VAH12	7	91.03	90.46	0.57	0.71	0.71	0.91	0.52	0.34	1.24
VAH12	8	90.01	89.5	0.51	0.77	0.75	0.93	0.57	0.33	1.22
VAH12	9	89.28	88.22	1.06	1.18	1.24	1.59	0.57	0.58	1.46
VAH12	10	88.53	88.4	0.13	0.78	0.59	0.79	0.58	0.16	1.05
VAH12	11	87.17	88.4	-1.23	1.81	1.48	2.19	0.82	0.63	1.52
VAH12	12	87.06	86.66	0.4	0.74	0.73	0.84	0.7	0.34	1.21
VAH12	13	86.46	86.31	0.15	0.56	0.36	0.58	0.67	0.08	0.95
VAH12	14	85.89	85.86	0.03	0.73	0.6	0.73	0.76	0.22	1.08
VAH12	15	85.34	85.12	0.22	0.66	0.59	0.7	0.69	0.22	1.07
VAH12	16	84.4	84.19	0.21	0.88	0.71	0.9	0.76	0.23	1.07
VAH12	17	83.99	84.64	-0.65	0.68	0.79	0.94	0.53	0.43	1.28

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH12	18	82.31	82.82	-0.51	0.94	0.86	1.06	0.48	0.37	1.2
VAH12	19	81.79	81.9	-0.11	1.12	0.92	1.13	0.72	0.31	1.13
VAH12	20	81.28	80.98	0.3	1.07	1.01	1.11	0.45	0.41	1.22
VAH12	21	80.89	80.81	0.08	1.02	0.91	1.02	0.68	0.35	1.16
VAH12	22	80.45	80.27	0.18	0.88	0.69	0.9	0.68	0.24	1.04
VAH12	23	80.05	80.12	-0.07	0.85	0.68	0.86	0.69	0.2	1
VAH12	24	78.91	79.52	-0.61	0.95	0.94	1.13	0.54	0.41	1.2
VAH12	25	78.31	79.55	-1.24	0.95	1.35	1.57	0.4	0.81	1.61
VAH12	26	78.81	79.11	-0.3	0.87	0.78	0.92	0.46	0.3	1.09
VAH12	27	78.17	78.31	-0.14	0.88	0.7	0.89	0.66	0.22	1
VAH12	28	78.34	79.01	-0.67	0.89	0.92	1.11	0.51	0.42	1.21
VAH12	29	77.84	78.43	-0.59	0.79	0.74	0.99	0.55	0.31	1.09
VAH12	30	75.89	77.86	-1.97	1.05	1.99	2.23	0.13	1.49	2.26
VAH12	31	76.83	77.75	-0.92	0.83	1.14	1.24	0.23	0.71	1.49
VAH12	32	76.51	77.76	-1.25	0.91	1.4	1.55	0.22	0.93	1.71
VAH12	33	76.36	77.05	-0.69	1.15	1.18	1.34	0.51	0.55	1.32
VAH12	34	75.35	76.71	-1.36	1.55	1.68	2.06	0.43	0.87	1.64
VAH12	35	75.03	75.89	-0.86	1.75	1.59	1.95	0.58	0.69	1.45
VAH12	36	75.17	75.77	-0.6	1.8	1.47	1.89	0.59	0.55	1.31
VAH12	37	74.64	75.83	-1.19	1.73	1.78	2.1	0.59	0.86	1.62
VAH12	38	72.95	76.44	-3.49	1.61	3.58	3.85	0.07	2.73	3.5
VAH12	39	74.07	74.43	-0.36	2.35	2.12	2.38	0.68	0.79	1.53
VAH12	40	73.69	74.52	-0.83	2.27	2.17	2.42	0.73	0.9	1.64
VAH12	41	73.31	74.65	-1.34	2.18	2.26	2.56	0.49	1.04	1.79
VAH12	42	73.03	74.43	-1.4	2.35	2.39	2.73	0.54	1.07	1.81
VAH12	43	72.65	73.99	-1.34	2.45	2.37	2.8	0.6	0.97	1.71
VAH12	44	72.21	73.58	-1.37	2.47	2.18	2.82	0.61	0.79	1.53
VAH12	45	71.61	73.31	-1.7	2.51	2.22	3.03	0.66	0.84	1.57
VAH12	46	71.89	73.57	-1.68	2.42	2.27	2.94	0.61	0.9	1.64
VAH13	1	100	99.44	0.56	1.09	0.56	1.23	0.79	0.11	1.11
VAH13	2	96.77	97.57	-0.8	1.51	1.22	1.71	0.75	0.49	1.47
VAH13	3	95.3	96.34	-1.04	1.83	1.57	2.11	0.65	0.59	1.55
VAH13	4	93.84	94.64	-0.8	1.19	1.1	1.43	0.69	0.55	1.49
VAH13	5	92.52	93.18	-0.66	1.26	1.01	1.42	0.8	0.43	1.36
VAH13	6	91.17	91.6	-0.43	0.69	0.62	0.81	0.62	0.24	1.16
VAH13	7	90.42	90.21	0.21	0.86	0.66	0.88	0.68	0.19	1.1
VAH13	8	89.43	89.43	0	0.75	0.54	0.75	0.55	0.15	1.05
VAH13	9	88.45	89	-0.55	0.69	0.66	0.88	0.48	0.3	1.19
VAH13	10	87.64	88.27	-0.63	0.7	0.78	0.94	0.55	0.42	1.3
VAH13	11	86.5	86.35	0.15	0.72	0.53	0.74	0.79	0.16	1.03
VAH13	12	86.51	86.29	0.22	0.6	0.36	0.64	0.79	0.09	0.95
VAH13	13	85.68	85.19	0.49	0.7	0.69	0.86	0.65	0.29	1.14
VAH13	14	84.49	84.76	-0.27	0.62	0.47	0.68	0.7	0.15	1
VAH13	15	84.2	83.37	0.83	1.15	1.05	1.41	0.64	0.42	1.25
VAH13	16	83.37	83.26	0.11	0.76	0.59	0.76	0.74	0.19	1.02
VAH13	17	79.75	78.93	0.82	0.84	0.96	1.17	0.53	0.52	1.31
VAH13	18	77.72	77.72	0	1.25	0.97	1.25	0.74	0.3	1.07
VAH13	19	78.17	77.45	0.72	1.26	0.99	1.45	0.72	0.33	1.11

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH13	20	75.76	75.58	0.18	1.79	1.55	1.8	0.51	0.54	1.29
VAH15	1	100	98.64	1.36	2.16	1.36	2.56	0.68	0.35	0.47
VAH15	2	96.59	96.56	0.03	1.42	0.78	1.42	0.79	0.14	0.25
VAH15	3	94.27	94.17	0.1	1.16	0.74	1.16	0.8	0.19	0.3
VAH15	4	92.37	92.07	0.3	1.11	0.78	1.15	0.73	0.21	0.31
VAH15	5	90.87	90.73	0.14	1.02	0.64	1.03	0.71	0.13	0.24
VAH15	6	89.41	89.56	-0.15	0.87	0.57	0.88	0.67	0.14	0.24
VAH15	7	87.95	87.92	0.03	0.99	0.71	0.99	0.67	0.19	0.29
VAH15	8	86.9	87.26	-0.36	0.77	0.69	0.85	0.67	0.3	0.41
VAH15	9	85.79	85.95	-0.16	0.68	0.51	0.7	0.68	0.16	0.26
VAH15	10	84.76	85.26	-0.5	0.83	0.82	0.97	0.61	0.37	0.47
VAH15	11	83.59	83.82	-0.23	0.75	0.58	0.79	0.71	0.2	0.3
VAH16	1	100	97.71	2.29	2.63	2.29	3.49	0.52	0.9	1.03
VAH16	2	96.86	96.25	0.61	1.82	1.22	1.92	0.65	0.3	0.43
VAH16	3	94.8	94.73	0.07	1.8	1.17	1.8	0.69	0.27	0.4
VAH16	4	93.01	92.62	0.39	1.39	0.87	1.44	0.8	0.22	0.34
VAH16	5	91.24	91.75	-0.51	1.54	1.15	1.63	0.73	0.36	0.48
VAH16	6	89.83	89.19	0.64	1.1	0.9	1.28	0.69	0.34	0.46
VAH16	7	88.65	88.28	0.37	1.09	0.84	1.16	0.77	0.29	0.41
VAH16	8	87.54	86.78	0.76	1.23	0.96	1.45	0.75	0.4	0.51
VAH16	9	86.36	84.85	1.51	1.6	1.62	2.2	0.68	0.87	0.98
VAH16	10	84.92	84.37	0.55	1.47	1.03	1.57	0.82	0.35	0.46
VAH16	11	84.02	83.2	0.82	1.45	1.05	1.67	0.8	0.4	0.51
VAH17	1	100	99.74	0.26	0.95	0.26	0.98	0.92	0.02	0.22
VAH17	2	96.95	96.76	0.19	0.94	0.49	0.96	0.85	0.11	0.3
VAH17	3	94.74	95.4	-0.66	1.62	1.1	1.75	0.7	0.28	0.47
VAH17	4	93.06	93.36	-0.3	1.02	0.75	1.07	0.67	0.2	0.39
VAH17	5	91.74	92.02	-0.28	0.88	0.77	0.92	0.78	0.32	0.51
VAH17	6	90.31	90.43	-0.12	0.84	0.68	0.85	0.81	0.25	0.43
VAH17	7	89.19	89.46	-0.27	0.75	0.59	0.8	0.69	0.22	0.4
VAH17	8	88.16	88.39	-0.23	0.72	0.53	0.76	0.79	0.17	0.34
VAH17	9	87.18	87.7	-0.52	0.7	0.68	0.87	0.56	0.3	0.48
VAH17	10	86.31	87.03	-0.72	0.72	0.88	1.02	0.46	0.5	0.68
VAH17	11	85.34	86.11	-0.77	0.62	0.88	0.99	0.33	0.56	0.73
VAH17	12	84.68	85.13	-0.45	0.61	0.61	0.76	0.38	0.3	0.47
VAH17	13	83.84	84.53	-0.69	0.76	0.81	1.03	0.57	0.39	0.56
VAH17	14	83.07	84.07	-1	0.66	1.06	1.2	0.18	0.72	0.89
VAH17	15	82.34	83.44	-1.1	0.75	1.19	1.33	0.09	0.81	0.97
VAH17	16	81.49	83.57	-2.08	1	2.15	2.31	0.15	1.63	1.8
VAH17	17	80.19	81.54	-1.35	1.32	1.52	1.89	0.5	0.79	0.95
VAH17	18	80.04	80.96	-0.92	1.08	1.16	1.42	0.51	0.55	0.72
VAH17	19	79.25	79.74	-0.49	1.13	0.99	1.23	0.74	0.36	0.52
VAH17	20	77.93	79.17	-1.24	1.01	1.31	1.6	0.51	0.76	0.92
VAH20	1	100	97.65	2.35	2.6	2.35	3.51	0.5	0.96	1.94
VAH20	2	96.72	96.04	0.68	2.14	1.65	2.25	0.68	0.49	1.45
VAH20	3	94.4	94.22	0.18	2.08	1.41	2.09	0.79	0.37	1.31
VAH20	4	92.61	92.01	0.6	1.53	1.18	1.64	0.74	0.39	1.31
VAH20	5	90.89	90.1	0.79	1.44	1.15	1.64	0.72	0.4	1.31

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH20	6	89.56	89.58	-0.02	1.17	0.87	1.17	0.67	0.24	1.14
VAH20	7	88.22	88.25	-0.03	0.89	0.69	0.89	0.7	0.21	1.1
VAH20	8	87.02	86.97	0.05	0.84	0.54	0.84	0.83	0.12	0.99
VAH20	9	85.23	85.31	-0.08	2.54	1.39	2.54	0.91	0.38	1.23
VAH20	10	84.08	83.79	0.29	2.43	1.15	2.45	0.92	0.22	1.06
VAH20	11	83.3	82.04	1.26	2.56	1.82	2.86	0.8	0.58	1.4
VAH20	12	82.15	80.92	1.23	2.6	2.01	2.88	0.84	0.65	1.46
VAH22	1	100	98.46	1.54	2.01	1.54	2.53	0.59	0.54	0.73
VAH22	2	96.58	96.52	0.06	1.17	0.61	1.17	0.8	0.1	0.29
VAH22	3	94.35	92.93	1.42	1.66	1.56	2.18	0.62	0.65	0.84
VAH22	4	92.45	92.21	0.24	0.87	0.57	0.9	0.84	0.16	0.34
VAH22	5	90.82	90.61	0.21	0.95	0.63	0.97	0.73	0.15	0.33
VAH22	6	89.37	89.49	-0.12	0.7	0.5	0.71	0.71	0.13	0.31
VAH22	7	88.02	88	0.02	0.56	0.39	0.56	0.75	0.09	0.27
VAH22	8	86.97	87.37	-0.4	0.63	0.59	0.75	0.66	0.25	0.42
VAH22	9	85.63	85.87	-0.24	0.63	0.47	0.68	0.68	0.16	0.33
VAH22	10	84.59	84.49	0.1	0.78	0.61	0.79	0.69	0.19	0.35
VAH23	1	100	90.5	9.5	3.41	9.5	10.09	0.05	7.66	7.78
VAH23	2	95.61	90.65	4.96	2.93	5.29	5.76	0.13	3.82	3.94
VAH23	3	93.73	90.33	3.4	2.82	3.88	4.42	0.29	2.43	2.55
VAH23	4	91.42	90.9	0.52	2.91	2.04	2.95	0.77	0.57	0.69
VAH23	5	91.25	89.17	2.08	3.3	3	3.9	0.62	1.28	1.39
VAH23	6	89.82	88.65	1.17	3.54	2.71	3.73	0.73	0.85	0.97
VAH23	7	88.83	88.65	0.18	3.17	2.25	3.18	0.79	0.6	0.72
VAH23	8	88.2	87.94	0.26	3.22	2.38	3.23	0.75	0.66	0.77
VAH23	9	85.73	88.97	-3.24	3.49	4.02	4.76	0.41	2.17	2.28
VAH23	10	85.95	87.86	-1.91	3.31	3.03	3.82	0.59	1.28	1.4
VAH23	11	85.07	86.72	-1.65	3.56	3.11	3.92	0.65	1.18	1.3
VAH23	12	84.57	85.96	-1.39	3.45	3.01	3.72	0.62	1.14	1.26
VAH23	13	83.91	86.03	-2.12	3.52	3.37	4.11	0.56	1.45	1.56
VAH23	14	82.86	85.08	-2.22	3.5	3.47	4.14	0.54	1.57	1.68
VAH24	1	100	98.45	1.55	2.02	1.55	2.55	0.58	0.55	0.66
VAH24	2	96.99	97.26	-0.27	1.83	1.18	1.85	0.64	0.25	0.36
VAH24	3	94.54	94.88	-0.34	1.81	1.27	1.85	0.72	0.33	0.44
VAH24	4	93.41	94	-0.59	2.06	1.43	2.14	0.85	0.4	0.51
VAH24	5	91.95	92.7	-0.75	2.18	1.49	2.31	0.8	0.42	0.53
VAH24	6	90.07	91.25	-1.18	2.15	1.47	2.45	0.77	0.48	0.59
VAH24	7	89.62	89.52	0.1	1.95	1.47	1.95	0.82	0.53	0.63
VAH24	8	88.51	88.51	0	1.52	1.01	1.52	0.85	0.31	0.41
VAH24	9	87.51	87.3	0.21	1.48	1.11	1.5	0.75	0.34	0.44
VAH24	10	86.72	86.59	0.13	1.13	0.9	1.14	0.69	0.29	0.39
VAH24	11	85.58	85.11	0.47	1.12	0.86	1.21	0.76	0.28	0.38
VAH24	12	84.58	84.57	0.01	1.13	0.82	1.13	0.8	0.23	0.33
VAH24	13	83.65	84.23	-0.58	1.09	0.95	1.23	0.56	0.37	0.47
VAH24	14	83.2	83.48	-0.28	0.95	0.7	0.99	0.76	0.2	0.3
VAH24	15	82.18	81.97	0.21	1.39	1.13	1.41	0.67	0.35	0.45
VAH24	16	80.33	81.55	-1.22	1.61	1.53	2.01	0.54	0.65	0.75
VAH25	1	100	98.79	1.21	2.1	1.21	2.43	0.72	0.27	0.6

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH25	2	96.82	96.73	0.09	1.78	1.18	1.78	0.68	0.27	0.6
VAH25	3	94.7	94.58	0.12	1.49	0.98	1.5	0.67	0.26	0.57
VAH25	4	92.95	92.06	0.89	0.97	1.03	1.31	0.49	0.52	0.82
VAH25	5	91.94	90.64	1.3	0.85	1.43	1.55	0.19	1	1.3
VAH25	6	90.23	89.18	1.05	0.95	1.19	1.41	0.49	0.73	1.03
VAH25	7	88.76	87.67	1.09	0.72	1.16	1.31	0.25	0.79	1.09
VAH25	8	87.45	86.52	0.93	0.7	1.03	1.16	0.22	0.68	0.97
VAH25	9	86.21	84.85	1.36	0.73	1.43	1.54	0.11	1.06	1.34
VAH25	10	85.39	83.97	1.42	0.68	1.44	1.58	0.08	1.07	1.35
VAH25	11	83.6	83.01	0.59	0.89	0.75	1.07	0.66	0.31	0.59
VAH26	1	100	97.9	2.1	2.02	2.1	2.91	0.44	1.01	1.11
VAH26	2	97.27	97.36	-0.09	1.53	1.09	1.53	0.68	0.33	0.42
VAH26	3	95.31	94.97	0.34	1.3	1.02	1.34	0.66	0.38	0.47
VAH26	4	93.73	93.87	-0.14	1.65	1.23	1.65	0.86	0.39	0.48
VAH26	5	92.34	92.79	-0.45	1.41	0.86	1.48	0.87	0.24	0.33
VAH26	6	90.3	89.56	0.74	0.79	0.85	1.08	0.35	0.48	0.57
VAH26	7	89.37	89.03	0.34	0.78	0.63	0.85	0.49	0.24	0.33
VAH26	8	88.51	88.22	0.29	0.84	0.68	0.89	0.66	0.23	0.32
VAH26	9	87.38	87.86	-0.48	0.74	0.67	0.88	0.68	0.27	0.36
VAH26	10	86.25	85.94	0.31	0.62	0.53	0.7	0.73	0.19	0.28
VAH26	11	84.62	84.79	-0.17	0.78	0.55	0.8	0.78	0.15	0.23
VAH26	12	84.5	83.83	0.67	1.11	0.97	1.3	0.64	0.38	0.46
VAH26	13	83.91	83.99	-0.08	0.73	0.5	0.73	0.79	0.12	0.21
VAH26	14	82.55	82.34	0.21	1.31	1.07	1.33	0.78	0.35	0.43
VAH26	15	81.97	81.28	0.69	1.4	1.29	1.56	0.52	0.52	0.6
VAH26	16	81.38	80.76	0.62	1.22	1.12	1.37	0.59	0.48	0.56
VAH26	17	80.22	80.4	-0.18	1.13	0.79	1.14	0.79	0.2	0.28
VAH26	18	79.5	79.15	0.35	0.85	0.69	0.92	0.72	0.24	0.32
VAH26	19	78.82	77.49	1.33	1.04	1.46	1.69	0.34	0.91	0.98
VAH26	20	78.39	76.3	2.09	1.54	2.14	2.6	0.34	1.29	1.37
VAH26	21	77.66	76.12	1.54	1.36	1.74	2.06	0.38	0.98	1.06
VAH27	1	100	99.84	0.16	0.77	0.16	0.78	0.96	0.01	0.07
VAH27	2	97.01	96.71	0.3	0.97	0.54	1.02	0.86	0.14	0.21
VAH27	3	94.91	94.44	0.47	0.9	0.73	1.01	0.7	0.29	0.35
VAH27	4	93.07	92.63	0.44	0.84	0.6	0.94	0.79	0.22	0.28
VAH27	5	91.1	91.87	-0.77	1.04	0.92	1.3	0.63	0.41	0.47
VAH27	6	88.95	87.57	1.38	1.04	1.44	1.73	0.5	0.9	0.96
VAH27	7	87.71	87.04	0.67	0.62	0.77	0.91	0.44	0.43	0.49
VAH27	8	86.53	86.01	0.52	0.63	0.68	0.82	0.43	0.33	0.39
VAH27	9	85.51	84.68	0.83	0.71	0.88	1.09	0.33	0.49	0.54
VAH27	10	83.79	82.24	1.55	1.09	1.55	1.89	0.24	0.97	1.02
VAH27	11	82.76	81.66	1.1	1.03	1.2	1.51	0.62	0.62	0.68
VAH28	1	100	98.38	1.62	2.36	1.62	2.86	0.62	0.47	1.46
VAH28	2	97.27	95.39	1.88	1.74	2.22	2.57	0.33	1.3	2.25
VAH28	3	95.47	94.13	1.34	1.53	1.75	2.03	0.63	0.96	1.9
VAH28	4	93.93	92.66	1.27	1.77	1.88	2.18	0.42	0.94	1.87
VAH28	5	92.68	92.07	0.61	1.3	1.23	1.43	0.74	0.58	1.5
VAH28	6	92.08	90.31	1.77	1.15	1.88	2.11	0.2	1.25	2.15

Continued on next page

Table 17 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH28	7	90.48	89.79	0.69	1.35	1.31	1.51	0.52	0.58	1.48
VAH28	8	89.43	88.84	0.59	1.78	1.36	1.87	0.81	0.54	1.43
VAH28	9	88.61	88.15	0.46	1.66	1.17	1.73	0.85	0.46	1.34
VAH28	10	87.92	87.95	-0.03	1.15	0.81	1.15	0.78	0.24	1.12
VAH28	11	86.99	86.46	0.53	1.38	1	1.48	0.83	0.33	1.19
VAH28	12	86.42	85.65	0.77	1.48	1.14	1.67	0.76	0.35	1.21
VAH28	13	85.78	85.52	0.26	1.5	1.14	1.52	0.72	0.33	1.18
VAH28	14	85.09	84.34	0.75	1.34	1.09	1.53	0.66	0.36	1.2
VAH28	15	84.36	83.59	0.77	1.3	1.05	1.5	0.73	0.35	1.19
VAH28	16	83.51	82.1	1.41	1.5	1.7	2.06	0.42	0.87	1.69
VAH28	17	83.29	82.19	1.1	1.25	1.31	1.66	0.6	0.64	1.46
VAH28	18	82.75	82.18	0.57	1.26	1	1.38	0.64	0.32	1.14
VAH28	19	82.13	81.56	0.57	1.4	1.24	1.51	0.67	0.46	1.27
VAH28	20	81.78	81.72	0.06	1.14	0.9	1.14	0.71	0.27	1.08
VAH28	21	81.12	80.32	0.8	1.17	1.12	1.42	0.6	0.46	1.26
VAH28	22	80.48	79.91	0.57	1.12	1.1	1.26	0.52	0.49	1.29
VAH28	23	80.56	80.56	0	1.23	1.06	1.23	0.7	0.37	1.18
VAH30	1	100	98.93	1.07	1.65	1.07	1.97	0.68	0.3	0.4
VAH30	2	96.89	97.55	-0.66	1.8	1.39	1.92	0.54	0.43	0.53
VAH30	3	94.67	97.22	-2.55	1.75	2.6	3.09	0.22	1.69	1.79
VAH30	4	93.05	96	-2.95	1.56	2.96	3.34	0.07	2.14	2.24
VAH30	5	91.4	93.91	-2.51	1.49	2.53	2.92	0.14	1.76	1.86
VAH30	6	89.81	92.32	-2.51	1.4	2.54	2.87	0.23	1.78	1.87
VAH30	7	88.96	91.07	-2.11	1.49	2.28	2.59	0.23	1.46	1.55
VAH30	8	87.59	89.32	-1.73	1.38	1.83	2.21	0.49	1.08	1.17
VAH30	9	86.78	89.11	-2.33	1.15	2.38	2.6	0.15	1.75	1.84
VAH30	10	85.77	87.65	-1.88	1.03	1.92	2.15	0.31	1.37	1.45
VAH30	11	84.53	87.02	-2.49	0.95	2.52	2.66	0.03	2.03	2.12
VAH30	12	84.13	86.01	-1.88	0.87	1.92	2.08	0.08	1.45	1.54
VAH30	13	83.13	85.42	-2.29	0.8	2.3	2.42	0.08	1.88	1.97
VAH30	14	82.14	84.17	-2.03	0.9	2.04	2.22	0.03	1.55	1.64
VAH30	15	81.52	83.74	-2.22	0.9	2.23	2.39	0.09	1.78	1.86
VAH30	16	80.73	82.76	-2.03	1.15	2.05	2.33	0.25	1.4	1.48
VAH30	17	79.91	81.53	-1.62	1.04	1.68	1.93	0.29	1.11	1.19
VAH30	18	79.63	81.13	-1.5	0.99	1.54	1.8	0.24	0.99	1.07

Table 17: Random Forest Regression - SOH Capacity Test Results

D Random Forest Regression - RUL Capacity Test Results

Note that all results are rounded to 2. Mean indicates distribution mean prediction and SD is the standard deviation.

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH01	1	612	440.48	171.52	103.92	173.67	200.54	0.4	117.35	117.57
VAH01	2	561	442.75	118.25	96.92	125.11	152.89	0.37	72.34	72.56
VAH01	3	510	374.39	135.61	92.94	145.72	164.41	0.16	101.36	101.54

Continued on next page

Table 18 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH01	4	459	345.72	113.28	88.59	129.23	143.81	0.15	85.01	85.18
VAH01	5	408	280.79	127.21	87.03	138.93	154.13	0.19	93.5	93.64
VAH01	6	357	230.14	126.86	87.98	139.32	154.38	0.2	93.82	93.93
VAH01	7	306	211.62	94.38	93.79	114.27	133.06	0.28	64.71	64.82
VAH01	8	255	133.55	121.45	43.92	121.61	129.15	0	98.9	98.97
VAH01	9	204	73.46	130.54	46.1	130.54	138.45	0	105.52	105.55
VAH01	10	153	88.59	64.41	46.14	65.22	79.23	0.21	40.07	40.11
VAH01	11	102	34.73	67.27	41.81	67.93	79.2	0.21	46.33	46.35
VAH01	12	51	20.76	30.24	27.98	33.26	41.2	0.35	19.81	19.82
VAH01	13	0	14.69	-14.69	25.83	14.69	29.71	0.74	3.55	3.55
VAH02	1	510	488.36	21.64	72.51	60.52	75.67	0.71	20.08	24.96
VAH02	2	459	468.42	-9.42	72.86	55.22	73.47	0.68	14.73	19.41
VAH02	3	408	454.19	-46.19	73.56	70.45	86.86	0.58	28.95	33.49
VAH02	4	357	416.71	-59.71	85.17	90.51	104.02	0.39	43.41	47.57
VAH02	5	306	342.04	-36.04	92.91	79.54	99.66	0.49	28.08	31.5
VAH02	6	255	298.76	-43.76	104.92	96.82	113.68	0.68	37.97	40.96
VAH02	7	204	271.25	-67.25	101.65	103.99	121.88	0.38	46.86	49.57
VAH02	8	153	109.68	43.32	36.4	50.52	56.58	0.13	34.27	35.37
VAH02	9	102	103.53	-1.53	50.21	31.11	50.23	0.56	5.73	6.76
VAH02	10	51	60.75	-9.75	47.76	35.25	48.75	0.39	10	10.6
VAH02	11	0	11.75	-11.75	34.53	11.75	36.48	0.88	1.16	1.28
VAH05	1	765	655.5	109.5	66.93	109.5	128.34	0.32	73.9	80.45
VAH05	2	714	650.34	63.66	71.46	69.56	95.7	0.68	31.48	37.98
VAH05	3	663	629.56	33.44	65.82	53.12	73.83	0.75	17.89	24.19
VAH05	4	612	581.53	30.47	77.63	49.37	83.4	0.8	12.26	18.07
VAH05	5	561	568.1	-7.1	77.2	54.44	77.52	0.82	15.66	21.34
VAH05	6	510	496.36	13.64	60.86	34.94	62.37	0.84	7.61	12.58
VAH05	7	459	444.51	14.49	75.62	54.25	77	0.8	14.46	18.9
VAH05	8	408	425.86	-17.86	59.06	48.56	61.71	0.87	18.54	22.8
VAH05	9	357	341.66	15.34	74.37	55.76	75.94	0.74	15.2	18.62
VAH05	10	306	295.12	10.88	75.06	55.02	75.84	0.84	16.27	19.22
VAH05	11	255	263.69	-8.69	75.75	55.41	76.25	0.79	17.24	19.87
VAH05	12	204	169.53	34.47	80.23	56.53	87.32	0.72	15.5	17.19
VAH05	13	153	155.93	-2.93	72.08	57.97	72.14	0.83	19.78	21.34
VAH05	14	102	142.38	-40.38	63.35	62.92	75.13	0.6	28.28	29.7
VAH05	15	51	50.23	0.77	38.84	17.59	38.84	0.73	2.19	2.69
VAH05	16	0	42.07	-42.07	42.96	42.07	60.13	0.34	22.75	23.17
VAH06	1	612	511.26	100.74	58.65	101.76	116.57	0.44	70.27	75.38
VAH06	2	561	487.17	73.83	52.75	73.83	90.74	0.5	44.76	49.63
VAH06	3	510	455.04	54.96	48.23	58.16	73.12	0.27	33.65	38.2
VAH06	4	459	430.86	28.14	51.15	40.16	58.38	0.8	13.66	17.97
VAH06	5	408	354.24	53.76	52.23	54.32	74.95	0.71	26.03	29.57
VAH06	6	357	304.58	52.42	66.12	63.96	84.38	0.64	27.97	31.01
VAH06	7	306	251.75	54.25	71.44	70.27	89.7	0.63	33.02	35.53
VAH06	8	255	199.79	55.21	64.17	62.09	84.65	0.73	28.96	30.95
VAH06	9	204	180.1	23.9	45.34	34.24	51.26	0.49	10.25	12.06
VAH06	10	153	136.2	16.8	75.33	62.74	77.18	0.58	21	22.36
VAH06	11	102	128.06	-26.06	67.86	54.62	72.69	0.78	17.23	18.51

Continued on next page

Table 18 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH06	12	51	47.45	3.55	56.02	37.25	56.13	0.89	10.31	10.78
VAH06	13	0	34.2	-34.2	52.6	34.2	62.74	0.85	9.77	10.11
VAH09	1	516	489.99	26.01	49.86	34.78	56.24	0.62	11.24	11.73
VAH09	2	465	459.29	5.71	43.47	26.12	43.84	0.64	6.13	6.59
VAH09	3	414	411.16	2.84	38	20.71	38.11	0.76	4.92	5.33
VAH09	4	363	356.96	6.04	50.4	30.02	50.76	0.76	6.75	7.11
VAH09	5	351	335.6	15.4	56.3	41.59	58.37	0.8	13.95	14.28
VAH09	6	339	326.06	12.94	56.92	44.37	58.38	0.77	16.48	16.81
VAH09	7	327	314.75	12.25	55.67	40.06	57	0.78	15.63	15.95
VAH09	8	315	301.12	13.88	44.03	26.49	46.17	0.8	9.06	9.36
VAH09	9	303	270.51	32.49	57.36	44.01	65.92	0.7	15.1	15.37
VAH09	10	291	269.96	21.04	50.77	41.6	54.96	0.76	15.53	15.8
VAH09	11	279	264.7	14.3	51.58	39.74	53.52	0.8	14.84	15.11
VAH09	12	204	176.34	27.66	47.28	37.46	54.78	0.45	12.97	13.14
VAH09	13	153	162.71	-9.71	55.79	38.16	56.63	0.87	9.04	9.2
VAH09	14	102	119.3	-17.3	54.3	47.19	56.99	0.92	18.69	18.81
VAH09	15	51	43.52	7.48	49.94	37.8	50.5	0.34	12.54	12.59
VAH09	16	0	34.14	-34.14	44.8	34.14	56.32	0.56	11.91	11.95
VAH10	1	612	506.6	105.4	79.63	120.94	132.1	0.2	80.44	85.51
VAH10	2	561	475.43	85.57	70.04	97.93	110.58	0.29	60.69	65.45
VAH10	3	510	461.21	48.79	78.11	80.65	92.09	0.52	37.74	42.35
VAH10	4	459	426.05	32.95	81.64	69.05	88.04	0.58	23.11	27.37
VAH10	5	408	363.24	44.76	76.66	77.64	88.77	0.49	34.98	38.61
VAH10	6	357	317.53	39.47	75.71	77.83	85.38	0.48	36.92	40.09
VAH10	7	306	289.72	16.28	71.27	56.04	73.11	0.74	18.32	21.22
VAH10	8	255	188.34	66.66	82.64	93.52	106.17	0.34	48.86	50.75
VAH10	9	204	152.7	51.3	83.62	83.38	98.1	0.43	36.48	38.01
VAH10	10	153	119.93	33.07	89.39	82.21	95.31	0.42	32.53	33.73
VAH10	11	102	72.06	29.94	65.14	60.84	71.69	0.61	25.04	25.76
VAH10	12	51	75.17	-24.17	52	45.59	57.34	0.85	17.08	17.84
VAH10	13	0	59.82	-59.82	42.37	59.82	73.3	0.17	38.67	39.26
VAH11	1	815	732.97	82.03	48.33	82.03	95.2	0	60.49	61.35
VAH11	2	764	728.4	35.6	47.88	36.54	59.67	0.58	14.56	15.42
VAH11	3	713	723.02	-10.02	48.85	35.98	49.87	0.46	12.27	13.12
VAH11	4	662	718.31	-56.31	51.25	66.62	76.14	0.21	40.73	41.58
VAH11	5	611	700.32	-89.32	61.96	96.9	108.71	0.33	63.93	64.76
VAH11	6	560	687	-127	78.03	136.92	149.05	0.21	95.77	96.58
VAH11	7	509	671.63	-162.63	87.88	165.65	184.86	0.19	117.81	118.6
VAH11	8	458	660.23	-202.23	93.07	203.62	222.62	0.15	152.61	153.38
VAH11	9	407	642.45	-235.45	105.02	236.26	257.81	0.17	178.05	178.8
VAH11	10	356	628.68	-272.68	109.84	273.04	293.98	0.12	211.31	212.05
VAH11	11	305	589.43	-284.43	124.19	285.23	310.36	0.12	215.07	215.76
VAH11	12	254	548.35	-294.35	129.68	294.58	321.65	0.13	220.46	221.11
VAH11	13	203	503.36	-300.36	138.72	300.59	330.85	0.13	221.41	222.01
VAH11	14	152	492.78	-340.78	139.72	340.88	368.31	0.06	261.09	261.67
VAH11	15	101	458.3	-357.3	146.31	357.3	386.09	0.05	273.8	274.34
VAH11	16	50	456.82	-406.82	147.73	406.82	432.81	0.02	322.6	323.14

Continued on next page

Table 18 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH11	17	0	310.52	-310.52	150.74	310.52	345.17	0.08	227.18	227.55
VAH12	1	765	649.8	115.2	63.15	118.2	131.37	0.16	84.77	91.26
VAH12	2	714	645.73	68.27	55.64	76.71	88.07	0.59	48.21	54.66
VAH12	3	663	573.68	89.32	76.07	101.48	117.32	0.35	60.36	66.1
VAH12	4	612	550.24	61.76	75.69	81.08	97.69	0.47	39.91	45.41
VAH12	5	561	435.61	125.39	91.67	142.65	155.32	0.17	95.47	99.83
VAH12	6	510	443.53	66.47	96.83	106.09	117.45	0.26	56.72	61.15
VAH12	7	459	348.41	110.59	69.49	116.71	130.61	0.19	81.46	84.95
VAH12	8	408	305.46	102.54	61.61	103.56	119.63	0.37	69.51	72.56
VAH12	9	357	231.43	125.57	103.96	130.67	163.02	0.56	72.72	75.04
VAH12	10	306	264.04	41.96	56.37	59.3	70.28	0.78	31.07	33.71
VAH12	11	255	186.13	68.87	73.36	79.07	100.62	0.63	41.09	42.95
VAH12	12	204	120.82	83.18	47.13	83.2	95.61	0.12	57.67	58.88
VAH12	13	153	108.38	44.62	42.31	46.8	61.49	0.33	24.04	25.12
VAH12	14	102	59.23	42.77	44.96	52.09	62.05	0.21	28.11	28.7
VAH12	15	51	26.01	24.99	30.17	29.07	39.17	0.44	14.38	14.64
VAH12	16	0	24.99	-24.99	26.5	24.99	36.42	0.52	11.76	12.01
VAH13	1	663	687.68	-24.68	80.46	78.64	84.16	0.45	34.53	41.41
VAH13	2	612	599.98	12.02	53.59	40.02	54.92	0.91	13.51	19.51
VAH13	3	561	581.75	-20.75	56.81	44.23	60.48	0.78	14.04	19.86
VAH13	4	510	519.91	-9.91	69.51	46.67	70.21	0.74	10.12	15.32
VAH13	5	459	445.22	13.78	57.4	41.9	59.04	0.88	12.22	16.67
VAH13	6	408	394.62	13.38	69.14	49.14	70.42	0.79	12.52	16.46
VAH13	7	357	330.93	26.07	53.2	52.29	59.25	0.81	24.19	27.5
VAH13	8	306	285.81	20.19	45.86	31.49	50.1	0.48	10.7	13.56
VAH13	9	255	264.55	-9.55	44	30.97	45.02	0.48	7.97	10.62
VAH13	10	204	204.16	-0.16	37.57	25.66	37.57	0.52	5.96	8
VAH13	11	153	122.36	30.64	55.23	47.7	63.16	0.71	17.09	18.31
VAH13	12	102	114.91	-12.91	58.43	47.59	59.84	0.82	15.18	16.33
VAH13	13	51	44.88	6.12	40.34	20.4	40.8	0.67	3.79	4.24
VAH13	14	0	29.6	-29.6	45.75	29.6	54.49	0.58	9.49	9.78
VAH15	1	459	466.16	-7.16	43.31	24.5	43.9	0.63	3.91	8.57
VAH15	2	408	437.66	-29.66	44.09	32.72	53.14	0.55	10.4	14.78
VAH15	3	357	399.26	-42.26	47.55	44.3	63.61	0.37	20.53	24.52
VAH15	4	306	327.6	-21.6	39.12	29.76	44.69	0.51	9.58	12.85
VAH15	5	255	278.33	-23.33	41.83	32.51	47.9	0.49	10.37	13.16
VAH15	6	204	227.45	-23.45	40.66	28.55	46.94	0.59	8.31	10.58
VAH15	7	153	153.05	-0.05	41.49	24.53	41.49	0.59	4.46	5.99
VAH15	8	102	140.36	-38.36	40.53	46.52	55.8	0.23	26.66	28.06
VAH15	9	51	80.17	-29.17	39.1	29.17	48.78	0.58	9.9	10.7
VAH15	10	0	54.25	-54.25	36.9	54.25	65.61	0.23	34.56	35.11
VAH16	1	459	499.89	-40.89	64.16	57.23	76.09	0.59	22.24	27.24
VAH16	2	408	473.09	-65.09	65.96	72.23	92.67	0.59	35.42	40.15
VAH16	3	357	435.27	-78.27	58.75	79.29	97.86	0.47	46.36	50.72
VAH16	4	306	361.25	-55.25	57.58	65.87	79.8	0.47	35.43	39.04
VAH16	5	255	314.85	-59.85	68.12	80.45	90.68	0.37	45.57	48.72
VAH16	6	204	223.44	-19.44	69.18	51.18	71.86	0.58	14.53	16.77
VAH16	7	153	195.57	-42.57	52.36	49.73	67.48	0.7	21.56	23.51

Continued on next page

Table 18 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH16	8	102	131.63	-29.63	41.95	35.77	51.36	0.46	13.62	14.94
VAH16	9	51	76.26	-25.26	40.21	26.3	47.49	0.64	7.19	7.95
VAH16	10	0	67.49	-67.49	42.74	67.49	79.88	0.24	45.47	46.14
VAH17	1	561	499.21	61.79	38.46	63.43	72.79	0.06	46.48	47.48
VAH17	2	510	455.02	54.98	48.46	59.1	73.29	0.28	32.75	33.66
VAH17	3	459	426.45	32.55	46.38	46.59	56.67	0.27	21.85	22.71
VAH17	4	408	367.75	40.25	45.6	46.54	60.83	0.35	23.01	23.74
VAH17	5	357	333.32	23.68	42.98	34.7	49.07	0.46	12.42	13.09
VAH17	6	306	287.95	18.05	36	25.48	40.27	0.68	7.06	7.64
VAH17	7	255	253.1	1.9	47.01	34.82	47.05	0.51	9.92	10.43
VAH17	8	204	216.67	-12.67	45.86	36.14	47.58	0.35	12.32	12.75
VAH17	9	153	174.93	-21.93	46.71	40.09	51.61	0.33	14.59	14.94
VAH17	10	102	148.62	-46.62	34.37	48.45	57.92	0.21	31.72	32.01
VAH17	11	51	86.45	-35.45	37.7	36.27	51.75	0.42	17.38	17.55
VAH17	12	0	34.89	-34.89	27.59	34.89	44.48	0.34	22.1	22.17
VAH20	1	459	445.99	13.01	63.34	43.31	64.66	0.74	9.83	10.72
VAH20	2	408	430.51	-22.51	59.04	43.73	63.19	0.85	12.86	13.72
VAH20	3	357	377.11	-20.11	54.36	38.69	57.96	0.81	10.13	10.88
VAH20	4	306	328.65	-22.65	53.56	36.35	58.15	0.8	9.17	9.83
VAH20	5	255	257.81	-2.81	69.11	49.42	69.16	0.72	12.44	12.95
VAH20	6	204	237.44	-33.44	66.02	56.97	74.01	0.58	20.82	21.29
VAH20	7	153	164.46	-11.46	52.34	35.01	53.58	0.82	7.7	8.02
VAH20	8	102	118.44	-16.44	46.29	33	49.13	0.48	8.82	9.05
VAH20	9	51	72.23	-21.23	41.67	31.43	46.76	0.49	9.69	9.83
VAH20	10	0	13.68	-13.68	27.05	13.68	30.31	0.77	2.73	2.75
VAH22	1	459	461.7	-2.7	56.16	31.44	56.23	0.84	4.67	5.22
VAH22	2	408	440.06	-32.06	48.29	39	57.96	0.48	13.37	13.88
VAH22	3	357	373.19	-16.19	50.4	34.65	52.94	0.5	8.44	8.88
VAH22	4	306	327.99	-21.99	35.33	24.94	41.62	0.6	7.86	8.25
VAH22	5	255	282.77	-27.77	40.67	35.45	49.25	0.46	14.53	14.86
VAH22	6	204	225.03	-21.03	45.95	33.87	50.54	0.48	10.23	10.49
VAH22	7	153	176.13	-23.13	40.09	31.77	46.28	0.47	11.31	11.52
VAH22	8	102	154.22	-52.22	44.19	54.98	68.41	0.21	32.6	32.78
VAH22	9	51	72.22	-21.22	36.94	22.3	42.61	0.67	5.61	5.69
VAH22	10	0	9.79	-9.79	21.74	9.79	23.84	0.82	1.64	1.65
VAH23	1	561	335.97	225.03	49.86	225.03	230.49	0	198.74	199.41
VAH23	2	510	320.67	189.33	47.98	189.53	195.31	0	164.38	165.02
VAH23	3	459	310.09	148.91	50.86	149.53	157.36	0.04	122.7	123.32
VAH23	4	408	293.64	114.36	50.02	116.26	124.82	0.03	90.04	90.63
VAH23	5	357	292.53	64.47	50.53	70.24	81.91	0.17	43.76	44.34
VAH23	6	306	281.63	24.37	48.15	39.78	53.97	0.47	13.82	14.39
VAH23	7	255	277.02	-22.02	48.77	40.18	53.51	0.41	13.89	14.44
VAH23	8	204	274.74	-70.74	48.78	72.16	85.92	0.18	45.57	46.12
VAH23	9	153	270.5	-117.5	48.55	117.7	127.13	0.01	91.26	91.8
VAH23	10	102	268.46	-166.46	50.19	166.46	173.86	0	139.01	139.54
VAH23	11	51	265.91	-214.91	50.67	214.91	220.8	0	187.31	187.85
VAH23	12	0	260.63	-260.63	49.19	260.63	265.23	0	233.85	234.37
VAH24	1	561	510.39	50.61	54.92	57	74.68	0.65	26.95	27.3

Continued on next page

Table 18 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH24	2	510	495.96	14.04	53.32	39.73	55.14	0.86	10.89	11.22
VAH24	3	459	452.49	6.51	54.04	36.87	54.44	0.86	8.51	8.81
VAH24	4	408	402.23	5.77	52.65	31.47	52.97	0.84	5.5	5.76
VAH24	5	357	355.26	1.74	57.29	34.03	57.32	0.85	6.06	6.3
VAH24	6	306	303.22	2.78	56.63	32.41	56.69	0.85	5.3	5.51
VAH24	7	255	243.13	11.87	52.39	33.34	53.72	0.85	7.16	7.33
VAH24	8	204	196.75	7.25	45.52	29.25	46.1	0.53	6.4	6.54
VAH24	9	153	165.56	-12.56	56.22	38.78	57.6	0.81	9.13	9.24
VAH24	10	102	134.75	-32.75	59.31	52.47	67.75	0.8	20.35	20.44
VAH24	11	51	62.33	-11.33	68.54	48.52	69.47	0.82	13.05	13.09
VAH24	12	0	52.4	-52.4	65.02	52.4	83.51	0.75	20.22	20.25
VAH25	1	512	508.52	3.48	35.54	18.93	35.71	0.73	3.71	4.73
VAH25	2	461	479.17	-18.17	40.58	31.73	44.46	0.49	11.28	12.24
VAH25	3	410	431.48	-21.48	47.71	32.33	52.32	0.62	8.43	9.3
VAH25	4	359	358.19	0.81	22.84	12.03	22.85	0.87	2.55	3.26
VAH25	5	307	315.66	-8.66	27.56	15.56	28.89	0.85	4.32	4.95
VAH25	6	256	256.84	-0.84	42.09	28.74	42.1	0.65	7.13	7.65
VAH25	7	205	179.26	25.74	37.14	30.65	45.19	0.5	11.96	12.32
VAH25	8	154	131.77	22.23	47.44	41.2	52.39	0.32	16.67	16.93
VAH25	9	103	23.47	79.53	28.72	80.35	84.56	0.01	66.68	66.73
VAH25	10	51	13.26	37.74	23.28	38.56	44.34	0.24	28.54	28.56
VAH25	11	0	8.47	-8.47	19.78	8.47	21.52	0.84	1.31	1.32
VAH26	1	612	601.5	10.5	58.82	44.05	59.75	0.83	12.24	13.04
VAH26	2	561	613.99	-52.99	43.55	56.54	68.59	0.18	35.15	35.97
VAH26	3	510	539.35	-29.35	58.8	49.32	65.72	0.75	17.08	17.8
VAH26	4	459	484.72	-25.72	57.18	42.72	62.7	0.81	12.9	13.55
VAH26	5	408	442.11	-34.11	61.18	51.77	70.04	0.72	18.46	19.05
VAH26	6	306	304.33	1.67	40.99	25.57	41.03	0.6	4.94	5.35
VAH26	7	255	275.93	-20.93	35.33	27.06	41.07	0.53	9.51	9.88
VAH26	8	204	224.06	-20.06	44.65	31.49	48.95	0.5	9.86	10.16
VAH26	9	153	197.5	-44.5	37.03	49.15	57.9	0.18	32.01	32.27
VAH26	10	51	95.94	-44.94	36.76	48.2	58.05	0.21	30.42	30.55
VAH26	11	0	30.39	-30.39	33.49	30.39	45.23	0.49	13.6	13.64
VAH27	1	510	515.22	-5.22	50.22	40.92	50.49	0.29	14.03	19.18
VAH27	2	459	474.4	-15.4	53.73	42.94	55.89	0.82	13.2	17.94
VAH27	3	408	418.6	-10.6	45.92	40.18	47.13	0.25	15.43	19.62
VAH27	4	357	375.7	-18.7	43.57	36.04	47.42	0.36	13.71	17.47
VAH27	5	306	337.36	-31.36	37.74	42.58	49.07	0.29	23.07	26.44
VAH27	6	204	174.96	29.04	46.34	35.22	54.69	0.47	12.14	13.89
VAH27	7	153	148.98	4.02	39.25	24.44	39.46	0.57	5.05	6.54
VAH27	8	102	91.9	10.1	46.24	35.8	47.33	0.38	11.51	12.43
VAH27	9	51	26.02	24.98	27.49	27.04	37.14	0.47	13.29	13.55
VAH27	10	0	10.21	-10.21	22.85	10.21	25.03	0.82	1.67	1.78
VAH28	1	720	634.58	85.42	58.5	87.98	103.53	0.56	57.96	59.23
VAH28	2	669	630.91	38.09	53.99	42.77	66.08	0.57	16.4	17.66
VAH28	3	618	577	41	55.02	47.48	68.62	0.5	19.63	20.78
VAH28	4	567	517.13	49.87	60.01	62.13	78.02	0.75	30.57	31.6
VAH28	5	516	487.66	28.34	51.84	38.18	59.08	0.62	13.16	14.13

Continued on next page

Table 18 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH28	6	464	401.2	62.8	50.68	66.53	80.7	0.2	41.26	42.06
VAH28	7	413	350.39	62.61	59.69	65.83	86.5	0.73	35.23	35.93
VAH28	8	362	287.66	74.34	52.66	75.88	91.1	0.18	50.05	50.63
VAH28	9	311	258	53	50.94	55.91	73.51	0.34	29.96	30.47
VAH28	10	260	226.34	33.66	65.2	53.83	73.38	0.78	18.45	18.9
VAH28	11	209	129.94	79.06	46.71	80.21	91.82	0.14	55.99	56.25
VAH28	12	158	99.48	58.52	49.04	63.55	76.35	0.25	37.52	37.72
VAH28	13	107	86.63	20.37	48.93	41.3	53	0.55	14.77	14.94
VAH28	14	51	41.48	9.52	31.44	20.27	32.85	0.61	4.85	4.93
VAH28	15	0	26.48	-26.48	34.18	26.48	43.24	0.57	9.83	9.89
VAH30	1	510	644.38	-134.38	71.26	135.19	152.1	0.17	96.19	96.83
VAH30	2	459	610.41	-151.41	61.13	151.72	163.29	0.08	120.61	121.22
VAH30	3	408	559.86	-151.86	50.29	151.86	159.97	0.01	125.59	126.15
VAH30	4	357	522.48	-165.48	58.1	165.48	175.38	0.04	133.39	133.92
VAH30	5	306	470.68	-164.68	56.91	165.39	174.23	0.02	135.43	135.9
VAH30	6	255	391.12	-136.12	53.27	136.84	146.18	0.07	108.65	109.04
VAH30	7	204	344.42	-140.42	58.59	142.06	152.15	0.08	111.51	111.85
VAH30	8	153	294.26	-141.26	61.08	142.5	153.9	0.07	110.34	110.64
VAH30	9	102	261.73	-159.73	57.78	160.14	169.86	0.04	129.53	129.79
VAH30	10	51	203.03	-152.03	57.5	152.74	162.54	0.06	122.06	122.26
VAH30	11	0	147.11	-147.11	56.04	147.11	157.42	0.08	116.49	116.63

Table 18: Random Forest Regression - RUL Capacity Test Results

E Convolution Neural Network with Monte Carlo Dropout - SOH Capacity Test Results

Note that all the results are rounded to 2 and since the input to the network has a window size of 2, we do not have results of capacity test '1' for each Vahana in the table. Mean indicates distribution mean prediction and SD is the standard deviation.

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH01	2	96.95	95.87	1.08	0.59	1.09	1.24	0.19	0.76	0.86
VAH01	3	94.98	94.55	0.43	0.9	0.79	1	0.65	0.28	0.37
VAH01	4	93.33	92.24	1.09	0.77	1.17	1.34	0.26	0.75	0.84
VAH01	5	91.85	91.14	0.71	0.71	0.83	1.01	0.49	0.42	0.51
VAH01	6	90.5	89.32	1.18	0.79	1.25	1.42	0.28	0.81	0.9
VAH01	7	89.42	87.51	1.91	0.77	1.92	2.05	0.07	1.49	1.58
VAH01	8	88.34	86.98	1.36	0.77	1.39	1.56	0.19	0.96	1.05
VAH01	9	87.45	85.41	2.04	0.81	2.06	2.19	0.06	1.62	1.7
VAH01	10	86.83	84.51	2.32	0.88	2.33	2.49	0.06	1.85	1.93
VAH01	11	85.26	84.73	0.53	0.92	0.88	1.07	0.57	0.36	0.44
VAH01	12	85.04	83.96	1.08	0.75	1.12	1.31	0.34	0.71	0.79
VAH01	13	83.83	83.08	0.75	0.83	0.92	1.12	0.5	0.46	0.54
VAH01	14	83.46	83.03	0.43	0.92	0.8	1.02	0.66	0.28	0.36
VAH01	15	82.66	80.92	1.74	0.88	1.75	1.95	0.16	1.26	1.34

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH01	16	80.81	81.01	-0.2	0.79	0.64	0.82	0.69	0.2	0.28
VAH01	17	80.98	78.97	2.01	1.16	2.05	2.32	0.24	1.39	1.47
VAH02	2	96.62	95.88	0.74	0.45	0.75	0.87	0	0.54	0.64
VAH02	3	94.47	95.25	-0.78	0.68	0.89	1.03	0.38	0.51	0.61
VAH02	4	92.69	93.96	-1.27	0.72	1.31	1.46	0.2	0.91	1.01
VAH02	5	91.14	91.91	-0.77	0.78	0.91	1.09	0.49	0.48	0.57
VAH02	6	89.64	89.51	0.13	1.03	0.82	1.03	0.69	0.24	0.33
VAH02	7	88.62	87.59	1.03	0.85	1.15	1.34	0.36	0.68	0.77
VAH02	8	87.52	86.48	1.04	0.72	1.11	1.26	0.28	0.72	0.8
VAH02	9	86.68	85.63	1.05	0.63	1.1	1.22	0.2	0.76	0.85
VAH02	10	85.52	85.1	0.42	0.72	0.64	0.83	0.65	0.27	0.35
VAH02	11	83.83	84.72	-0.89	0.89	0.97	1.26	0.53	0.5	0.59
VAH02	12	83.54	83.82	-0.28	0.56	0.39	0.63	0.74	0.12	0.2
VAH02	13	82.34	82.44	-0.1	0.9	0.8	0.91	0.54	0.28	0.37
VAH05	2	97.33	96.21	1.12	0.75	1.12	1.35	0.33	0.76	0.86
VAH05	3	95.41	95.9	-0.49	0.83	0.85	0.97	0.46	0.41	0.51
VAH05	4	94.02	94.67	-0.65	0.96	0.94	1.16	0.6	0.39	0.49
VAH05	5	92.84	93.3	-0.46	0.85	0.73	0.97	0.68	0.27	0.36
VAH05	6	91.65	92.83	-1.18	0.8	1.24	1.42	0.27	0.8	0.9
VAH05	7	90.79	91.33	-0.54	1	0.88	1.14	0.65	0.32	0.41
VAH05	8	89.92	90.16	-0.24	0.9	0.73	0.94	0.7	0.23	0.32
VAH05	9	89	90.17	-1.17	0.9	1.22	1.48	0.38	0.74	0.83
VAH05	10	88.22	88.69	-0.47	0.85	0.76	0.97	0.63	0.29	0.38
VAH05	11	87.5	87.99	-0.49	0.74	0.69	0.89	0.64	0.28	0.37
VAH05	12	86.95	87.45	-0.5	0.83	0.8	0.97	0.59	0.35	0.44
VAH05	13	86.33	86.4	-0.07	0.87	0.68	0.87	0.7	0.2	0.29
VAH05	14	85.91	86.02	-0.11	0.81	0.64	0.82	0.7	0.18	0.27
VAH05	15	85.14	85.74	-0.6	0.91	0.88	1.09	0.6	0.37	0.46
VAH05	16	84.68	84.32	0.36	0.99	0.83	1.06	0.67	0.27	0.36
VAH05	17	84.22	83.26	0.96	0.95	1.14	1.35	0.44	0.61	0.69
VAH05	18	83.77	82.96	0.81	0.83	0.98	1.15	0.46	0.52	0.6
VAH05	19	83.09	82.89	0.2	0.78	0.67	0.8	0.59	0.23	0.32
VAH05	20	82.53	82.82	-0.29	0.71	0.62	0.76	0.68	0.22	0.3
VAH05	21	81.73	82.58	-0.85	0.83	0.99	1.19	0.46	0.53	0.61
VAH05	22	81.17	82.29	-1.12	0.99	1.31	1.5	0.41	0.76	0.84
VAH05	23	80.89	82.43	-1.54	0.71	1.59	1.7	0.07	1.21	1.3
VAH05	24	80.11	82.51	-2.4	1.13	2.46	2.65	0.1	1.85	1.94
VAH05	25	79.33	82.16	-2.83	1.18	2.9	3.07	0.07	2.29	2.37
VAH05	26	79.06	82.22	-3.16	1.11	3.22	3.35	0.04	2.68	2.76
VAH05	27	78.41	81.85	-3.44	1.49	3.53	3.75	0.1	2.76	2.84
VAH05	28	77.71	81.05	-3.34	2	3.58	3.89	0.18	2.53	2.61
VAH05	29	77.3	80.77	-3.47	2.08	3.75	4.05	0.16	2.67	2.75
VAH06	2	97	95.4	1.6	1.15	1.59	1.96	0.53	1.01	1.1
VAH06	3	94.88	95.08	-0.2	1.21	1.05	1.22	0.64	0.39	0.48
VAH06	4	93.5	91.98	1.52	1.49	1.75	2.13	0.5	0.92	1.01
VAH06	5	92.22	91.03	1.19	1.32	1.43	1.78	0.5	0.69	0.79
VAH06	6	90.84	90.02	0.82	1.27	1.23	1.51	0.58	0.54	0.63
VAH06	7	90.03	89.2	0.83	2.04	1.86	2.2	0.61	0.75	0.83

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH06	8	88.52	89.54	-1.02	1.87	1.58	2.13	0.7	0.55	0.64
VAH06	9	87.91	88.26	-0.35	1.34	0.99	1.39	0.75	0.29	0.38
VAH06	10	86.47	87.88	-1.41	1.64	1.63	2.16	0.62	0.77	0.86
VAH06	11	85.87	87.03	-1.16	1.96	1.53	2.28	0.77	0.54	0.63
VAH06	12	85.43	86.01	-0.58	2.14	1.43	2.22	0.84	0.37	0.45
VAH06	13	84.76	84.28	0.48	1.95	1.47	2.01	0.76	0.58	0.67
VAH06	14	84.01	83.65	0.36	1.98	1.34	2.01	0.93	0.6	0.68
VAH06	15	83.28	83.18	0.1	1.38	0.8	1.39	0.94	0.34	0.42
VAH06	16	81.97	83.21	-1.24	1.61	1.25	2.03	0.87	0.74	0.83
VAH06	17	81.18	83.3	-2.12	1.68	2.13	2.71	0.69	1.56	1.64
VAH06	18	80.44	83.37	-2.93	1.57	2.92	3.32	0	2.33	2.42
VAH06	19	79.69	83.88	-4.19	1.61	4.19	4.49	0	3.41	3.5
VAH09	2	96.98	95.75	1.23	1.01	1.22	1.58	0.52	0.68	0.78
VAH09	3	94.89	93.77	1.12	0.82	1.2	1.39	0.32	0.75	0.84
VAH09	4	93.18	92.15	1.03	0.61	1.06	1.2	0.22	0.72	0.81
VAH09	5	92.52	91.85	0.67	0.69	0.79	0.96	0.49	0.41	0.5
VAH09	6	91.93	91.76	0.17	0.71	0.56	0.73	0.71	0.16	0.26
VAH09	7	91.39	91.68	-0.29	0.77	0.62	0.82	0.72	0.2	0.29
VAH09	8	90.87	91.47	-0.6	0.8	0.79	0.99	0.57	0.35	0.44
VAH09	9	90.4	91.44	-1.04	0.81	1.1	1.32	0.38	0.65	0.75
VAH09	10	89.98	91.28	-1.3	0.9	1.36	1.58	0.3	0.87	0.97
VAH09	11	89.56	91.14	-1.58	0.9	1.6	1.82	0.21	1.1	1.19
VAH09	12	88.11	90.41	-2.3	1.05	2.31	2.53	0.1	1.74	1.83
VAH09	13	87.09	89.08	-1.99	1.25	2.02	2.35	0.32	1.31	1.4
VAH09	14	86.23	87.99	-1.76	1.19	1.77	2.12	0.37	1.11	1.19
VAH09	15	85.35	87.11	-1.76	1.04	1.77	2.05	0.26	1.2	1.29
VAH09	16	84.48	86.28	-1.8	0.93	1.81	2.03	0.15	1.3	1.38
VAH09	17	83.89	85.57	-1.68	0.84	1.68	1.88	0.13	1.23	1.32
VAH09	18	83.11	85.02	-1.91	0.73	1.91	2.05	0.03	1.51	1.6
VAH09	19	82.41	84.28	-1.87	0.66	1.88	1.99	0.03	1.52	1.61
VAH09	20	79.76	82.55	-2.79	1.08	2.78	2.98	0.07	2.17	2.25
VAH09	21	78.95	80.95	-2	1.22	2.01	2.35	0.3	1.31	1.39
VAH10	2	97.01	90.03	6.98	3.12	6.99	7.65	0.14	5.19	5.28
VAH10	3	94.97	90.23	4.74	3.32	4.76	5.79	0.39	2.86	2.95
VAH10	4	93.29	90.59	2.7	3.02	3.09	4.05	0.59	1.38	1.47
VAH10	5	91.87	89.64	2.23	2.09	2.42	3.06	0.53	1.25	1.34
VAH10	6	90.64	88.59	2.05	1.63	2.12	2.62	0.47	1.2	1.29
VAH10	7	89.52	88.02	1.5	1.46	1.68	2.1	0.51	0.85	0.93
VAH10	8	88.45	87.08	1.37	1.19	1.55	1.81	0.37	0.88	0.97
VAH10	9	87.46	86.25	1.21	0.86	1.31	1.48	0.28	0.84	0.92
VAH10	10	86.72	85.47	1.25	0.78	1.29	1.48	0.26	0.86	0.95
VAH10	11	85.7	84.14	1.56	0.83	1.59	1.77	0.18	1.12	1.2
VAH10	12	85.44	83.11	2.33	0.95	2.34	2.52	0.06	1.8	1.89
VAH10	13	84.65	82.67	1.98	0.93	1.98	2.18	0.14	1.46	1.54
VAH10	14	83.84	81.82	2.02	0.7	2.03	2.15	0.05	1.65	1.73
VAH10	15	83.16	81.4	1.76	0.6	1.77	1.86	0.04	1.44	1.52
VAH10	16	82.1	80.99	1.11	0.5	1.12	1.22	0.11	0.84	0.92
VAH10	17	81.39	80.66	0.73	0.52	0.77	0.9	0.33	0.48	0.56

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH10	18	80.76	80.5	0.26	0.52	0.46	0.58	0.63	0.17	0.25
VAH10	19	80.19	80.36	-0.17	0.51	0.43	0.54	0.67	0.14	0.22
VAH10	20	79.63	80.25	-0.62	0.54	0.68	0.81	0.42	0.38	0.46
VAH10	21	79.05	80.1	-1.05	0.55	1.07	1.19	0.16	0.77	0.85
VAH10	22	78.68	79.87	-1.19	0.59	1.21	1.33	0.14	0.89	0.97
VAH10	23	78.38	79.36	-0.98	0.76	1.08	1.24	0.32	0.66	0.74
VAH10	24	76.81	78.6	-1.79	1.05	1.84	2.07	0.23	1.26	1.34
VAH10	25	76.49	77.58	-1.09	1.36	1.4	1.74	0.55	0.65	0.72
VAH10	26	76.03	77.11	-1.08	1.72	1.6	2.03	0.62	0.64	0.71
VAH10	27	75.38	76.9	-1.52	1.87	1.88	2.41	0.58	0.83	0.9
VAH10	28	74.87	76.72	-1.85	2.08	2.13	2.78	0.58	0.96	1.04
VAH11	2	97.6	96.49	1.11	0.52	1.11	1.23	0	0.92	1.02
VAH11	3	95.98	96.43	-0.45	0.48	0.59	0.66	0.29	0.38	0.47
VAH11	4	94.65	95.86	-1.21	0.71	1.27	1.4	0.21	0.89	0.98
VAH11	5	93.25	93.78	-0.53	0.97	0.88	1.1	0.61	0.33	0.42
VAH11	6	92.1	92.76	-0.66	0.95	0.91	1.15	0.61	0.37	0.47
VAH11	7	91.52	92.07	-0.55	0.94	0.83	1.09	0.64	0.31	0.41
VAH11	8	90.82	91.33	-0.51	0.98	0.87	1.1	0.67	0.34	0.43
VAH11	9	90.01	90.54	-0.53	0.98	0.88	1.11	0.62	0.33	0.42
VAH11	10	89.26	89.87	-0.61	1.15	0.96	1.3	0.7	0.34	0.43
VAH11	11	88.31	88.77	-0.46	0.93	0.76	1.03	0.7	0.26	0.35
VAH11	12	88.18	88.69	-0.51	1.14	0.89	1.25	0.73	0.28	0.37
VAH11	13	87.53	88.03	-0.5	0.93	0.8	1.06	0.68	0.3	0.38
VAH11	14	87.02	87.6	-0.58	0.99	0.85	1.15	0.68	0.32	0.41
VAH11	15	86.48	87.1	-0.62	0.93	0.86	1.12	0.64	0.35	0.44
VAH11	16	85.79	86.13	-0.34	0.96	0.72	1.02	0.76	0.22	0.3
VAH11	17	83.71	85.38	-1.67	0.9	1.67	1.9	0.17	1.2	1.29
VAH11	18	83.23	84.61	-1.38	0.81	1.38	1.6	0.24	0.94	1.02
VAH11	19	82.72	84.01	-1.29	0.88	1.31	1.57	0.38	0.82	0.91
VAH11	20	82.2	83.25	-1.05	0.83	1.07	1.34	0.45	0.63	0.71
VAH11	21	81.86	82.77	-0.91	0.7	0.92	1.15	0.43	0.55	0.63
VAH11	22	81.43	82.53	-1.1	0.63	1.1	1.27	0.26	0.76	0.84
VAH11	23	80.94	82.38	-1.44	0.64	1.44	1.58	0.01	1.11	1.2
VAH11	24	80.58	82.19	-1.61	0.52	1.61	1.69	0	1.34	1.43
VAH11	25	80.05	82.15	-2.1	0.46	2.1	2.15	0	1.86	1.94
VAH11	26	79.57	81.96	-2.39	0.4	2.39	2.42	0	2.19	2.27
VAH11	27	79.19	81.9	-2.71	0.4	2.71	2.74	0	2.51	2.59
VAH11	28	78.55	81.93	-3.38	0.37	3.38	3.4	0	3.19	3.27
VAH11	29	78.06	81.75	-3.69	0.33	3.69	3.7	0	3.52	3.6
VAH11	30	77.7	81.74	-4.04	0.34	4.04	4.06	0	3.87	3.95
VAH11	31	77.4	81.71	-4.31	0.36	4.31	4.33	0	4.13	4.21
VAH11	32	76.84	81.71	-4.87	0.35	4.87	4.88	0	4.7	4.78
VAH11	33	76.5	81.64	-5.14	0.36	5.14	5.15	0	4.95	5.03
VAH11	34	75.94	81.64	-5.7	0.35	5.71	5.72	0	5.53	5.61
VAH11	35	75.62	81.62	-6	0.36	6	6.01	0	5.81	5.9
VAH11	36	75.23	81.62	-6.39	0.39	6.39	6.4	0	6.2	6.28
VAH11	37	74.74	81.61	-6.87	0.34	6.88	6.88	0	6.7	6.78
VAH11	38	74.21	81.62	-7.41	0.37	7.4	7.41	0	7.22	7.3

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH11	39	73.86	81.61	-7.75	0.38	7.75	7.76	0	7.55	7.63
VAH11	40	73.01	81.57	-8.56	0.39	8.56	8.56	0	8.35	8.43
VAH11	41	72.54	81.58	-9.04	0.36	9.04	9.04	0	8.85	8.93
VAH11	42	72.21	81.54	-9.33	0.41	9.33	9.34	0	9.12	9.2
VAH11	43	71.82	81.5	-9.68	0.42	9.69	9.7	0	9.47	9.55
VAH11	44	70.69	81.42	-10.73	0.52	10.73	10.74	0	10.47	10.55
VAH12	2	97.45	95.26	2.19	1.5	2.19	2.65	0.39	1.38	1.47
VAH12	3	95.71	94.44	1.27	1.57	1.54	2.02	0.59	0.67	0.77
VAH12	4	94.3	92.72	1.58	1.42	1.74	2.12	0.45	0.97	1.07
VAH12	5	92.89	91.59	1.3	1.25	1.32	1.8	0.6	0.68	0.77
VAH12	6	91.51	89.87	1.64	1.61	1.79	2.3	0.51	0.88	0.97
VAH12	7	91.03	88.63	2.4	1.56	2.48	2.87	0.29	1.59	1.68
VAH12	8	90.01	87.55	2.46	1.22	2.51	2.75	0.15	1.83	1.91
VAH12	9	89.28	86.67	2.61	1.04	2.63	2.82	0.07	2.05	2.14
VAH12	10	88.53	86.16	2.37	0.92	2.37	2.53	0.05	1.85	1.94
VAH12	11	87.17	87.03	0.14	1.47	1.16	1.48	0.69	0.34	0.42
VAH12	12	87.06	85.87	1.19	0.8	1.22	1.44	0.34	0.78	0.86
VAH12	13	86.46	84.11	2.35	1.05	2.36	2.58	0.08	1.78	1.87
VAH12	14	85.89	83.01	2.88	1.18	2.88	3.12	0.08	2.21	2.29
VAH12	15	85.34	81.77	3.57	1	3.57	3.71	0.01	3.01	3.09
VAH12	16	84.4	81.08	3.32	0.77	3.32	3.4	0	2.91	2.99
VAH12	17	83.99	82.74	1.25	0.92	1.29	1.55	0.4	0.77	0.85
VAH12	18	82.31	83.47	-1.16	1.06	1.34	1.56	0.4	0.74	0.83
VAH12	19	81.79	80.35	1.44	0.61	1.46	1.57	0.05	1.12	1.2
VAH12	20	81.28	80.12	1.16	0.7	1.19	1.36	0.29	0.8	0.88
VAH12	21	80.89	79.88	1.01	0.73	1.04	1.25	0.37	0.63	0.71
VAH12	22	80.45	79.89	0.56	0.74	0.73	0.93	0.6	0.31	0.39
VAH12	23	80.05	79.57	0.48	0.79	0.76	0.93	0.61	0.31	0.39
VAH12	24	78.91	79.19	-0.28	0.93	0.79	0.98	0.65	0.26	0.34
VAH12	25	78.31	80.13	-1.82	0.7	1.82	1.95	0.06	1.43	1.51
VAH12	26	78.81	79.69	-0.88	0.72	0.95	1.14	0.43	0.54	0.62
VAH12	27	78.17	78.18	-0.01	0.71	0.55	0.71	0.71	0.16	0.23
VAH12	28	78.34	77.73	0.61	0.75	0.79	0.97	0.52	0.37	0.45
VAH12	29	77.84	77.6	0.24	0.8	0.67	0.84	0.66	0.22	0.3
VAH12	30	75.89	77.04	-1.15	0.87	1.15	1.44	0.46	0.69	0.77
VAH12	31	76.83	78.87	-2.04	0.59	2.04	2.12	0	1.72	1.8
VAH12	32	76.51	77.21	-0.7	0.79	0.81	1.06	0.58	0.37	0.44
VAH12	33	76.36	77.75	-1.39	0.95	1.41	1.69	0.32	0.91	0.98
VAH12	34	75.35	76.86	-1.51	0.84	1.51	1.73	0.19	1.08	1.16
VAH12	35	75.03	77.66	-2.63	0.86	2.63	2.77	0	2.17	2.25
VAH12	36	75.17	77.32	-2.15	1.07	2.15	2.4	0.14	1.59	1.66
VAH12	37	74.64	76.88	-2.24	0.98	2.24	2.44	0	1.73	1.81
VAH12	38	72.95	76.73	-3.78	0.77	3.78	3.86	0	3.41	3.48
VAH12	39	74.07	78.35	-4.28	0.84	4.28	4.36	0	3.86	3.94
VAH12	40	73.69	76.78	-3.09	1.04	3.09	3.26	0	2.56	2.64
VAH12	41	73.31	76.91	-3.6	1.07	3.6	3.76	0	3.06	3.14
VAH12	42	73.03	76.95	-3.92	1.16	3.92	4.09	0	3.32	3.4
VAH12	43	72.65	76.97	-4.32	1.29	4.32	4.51	0	3.65	3.73

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH12	44	72.21	76.99	-4.78	1.09	4.78	4.9	0	4.2	4.27
VAH12	45	71.61	76.79	-5.18	1.17	5.19	5.32	0	4.59	4.66
VAH12	46	71.89	76.85	-4.96	1.21	4.96	5.1	0	4.33	4.4
VAH13	2	96.77	96.37	0.4	0.34	0.4	0.52	0.84	0.31	0.41
VAH13	3	95.3	96.06	-0.76	0.67	0.94	1.01	0.22	0.63	0.73
VAH13	4	93.84	94.48	-0.64	1.22	1.14	1.38	0.6	0.44	0.54
VAH13	5	92.52	93.13	-0.61	1.11	1.01	1.26	0.62	0.39	0.48
VAH13	6	91.17	90.79	0.38	1.2	1.02	1.26	0.63	0.36	0.45
VAH13	7	90.42	90.47	-0.05	1.12	0.81	1.13	0.78	0.22	0.31
VAH13	8	89.43	90.12	-0.69	1.04	0.94	1.25	0.64	0.39	0.48
VAH13	9	88.45	89.67	-1.22	1.14	1.38	1.67	0.48	0.77	0.86
VAH13	10	87.64	89.09	-1.45	1.19	1.62	1.88	0.33	0.97	1.06
VAH13	11	86.5	87.42	-0.92	1.25	1.24	1.55	0.57	0.55	0.64
VAH13	12	86.51	85.46	1.05	1.33	1.36	1.69	0.57	0.63	0.71
VAH13	13	85.68	84.05	1.63	1.19	1.71	2.02	0.36	1.04	1.13
VAH13	14	84.49	84.05	0.44	1.33	1.12	1.4	0.63	0.38	0.47
VAH13	15	84.2	83.38	0.82	1.15	1.17	1.41	0.55	0.51	0.6
VAH13	16	83.37	82.3	1.07	0.93	1.25	1.42	0.33	0.75	0.83
VAH13	17	79.75	79.85	-0.1	0.91	0.69	0.92	0.73	0.19	0.27
VAH13	18	77.72	79.71	-1.99	0.95	2	2.21	0.13	1.47	1.55
VAH13	19	78.17	78.61	-0.44	1.01	0.84	1.11	0.69	0.29	0.37
VAH13	20	75.76	78.63	-2.87	0.91	2.87	3.01	0.01	2.37	2.44
VAH15	2	96.59	96.28	0.31	0.53	0.42	0.61	0.72	0.15	0.24
VAH15	3	94.27	95.38	-1.11	0.8	1.24	1.36	0.25	0.82	0.91
VAH15	4	92.37	93.66	-1.29	1.06	1.43	1.68	0.4	0.82	0.92
VAH15	5	90.87	91.14	-0.27	0.89	0.74	0.93	0.66	0.24	0.33
VAH15	6	89.41	89.41	0	0.73	0.56	0.73	0.71	0.16	0.25
VAH15	7	87.95	89.02	-1.07	0.81	1.13	1.34	0.32	0.7	0.79
VAH15	8	86.9	87.63	-0.73	0.92	0.92	1.18	0.58	0.42	0.5
VAH15	9	85.79	86.31	-0.52	0.87	0.77	1.01	0.65	0.3	0.38
VAH15	10	84.76	85.48	-0.72	0.84	0.88	1.1	0.54	0.42	0.5
VAH15	11	83.59	84.62	-1.03	0.92	1.11	1.38	0.47	0.61	0.69
VAH16	2	96.86	91.31	5.55	0.67	5.56	5.6	0	5.19	5.28
VAH16	3	94.8	91	3.8	0.96	3.8	3.92	0	3.26	3.35
VAH16	4	93.01	88.66	4.35	1.13	4.36	4.5	0.01	3.74	3.83
VAH16	5	91.24	89.54	1.7	1.31	1.86	2.14	0.34	1.12	1.21
VAH16	6	89.83	86.91	2.92	0.89	2.93	3.06	0.02	2.43	2.51
VAH16	7	88.65	86.73	1.92	0.9	1.95	2.13	0.12	1.44	1.53
VAH16	8	87.54	85.01	2.53	0.82	2.54	2.66	0.03	2.08	2.17
VAH16	9	86.36	82.42	3.94	0.78	3.94	4.02	0	3.51	3.59
VAH16	10	84.92	81.45	3.47	0.68	3.47	3.54	0	3.1	3.18
VAH16	11	84.02	81.6	2.42	0.91	2.42	2.58	0.05	1.91	1.99
VAH17	2	96.95	97.09	-0.14	0.8	0.64	0.81	0.68	0.21	0.31
VAH17	3	94.74	96.38	-1.64	0.96	1.72	1.9	0.18	1.19	1.29
VAH17	4	93.06	94.35	-1.29	0.62	1.32	1.43	0.12	1.03	1.12
VAH17	5	91.74	92.36	-0.62	0.71	0.76	0.94	0.52	0.37	0.46
VAH17	6	90.31	91.4	-1.09	0.73	1.13	1.31	0.29	0.73	0.82
VAH17	7	89.19	90.74	-1.55	0.88	1.58	1.78	0.19	1.09	1.18

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH17	8	88.16	89.99	-1.83	1.27	1.9	2.23	0.33	1.18	1.27
VAH17	9	87.18	88.69	-1.51	1.45	1.64	2.09	0.53	0.82	0.91
VAH17	10	86.31	87.21	-0.9	1.11	1	1.43	0.69	0.43	0.51
VAH17	11	85.34	86.41	-1.07	0.8	1.09	1.33	0.4	0.67	0.76
VAH17	12	84.68	85.84	-1.16	0.64	1.18	1.33	0.16	0.84	0.93
VAH17	13	83.84	85.36	-1.52	0.62	1.53	1.65	0.07	1.2	1.28
VAH17	14	83.07	84.56	-1.49	0.66	1.51	1.63	0.09	1.15	1.23
VAH17	15	82.34	83.03	-0.69	0.7	0.81	0.98	0.5	0.41	0.5
VAH17	16	81.49	83.11	-1.62	0.77	1.64	1.79	0.13	1.21	1.29
VAH17	17	80.19	83.92	-3.73	0.77	3.73	3.81	0	3.3	3.39
VAH17	18	80.04	82.54	-2.5	0.82	2.51	2.64	0.03	2.05	2.13
VAH17	19	79.25	80.26	-1.01	0.93	1.15	1.37	0.46	0.62	0.7
VAH17	20	77.93	80.15	-2.22	1.03	2.23	2.45	0.14	1.64	1.72
VAH20	2	96.72	89.27	7.45	1.08	7.44	7.52	0	6.83	6.92
VAH20	3	94.4	88.8	5.6	1.11	5.6	5.71	0	4.97	5.05
VAH20	4	92.61	87.36	5.25	1.12	5.24	5.36	0	4.62	4.71
VAH20	5	90.89	86.79	4.1	1.19	4.11	4.27	0.01	3.43	3.52
VAH20	6	89.56	85.61	3.95	1.03	3.95	4.08	0	3.37	3.46
VAH20	7	88.22	85.18	3.04	1.04	3.05	3.22	0.04	2.46	2.55
VAH20	8	87.02	84.12	2.9	1.01	2.9	3.07	0.04	2.33	2.42
VAH20	9	85.23	82.84	2.39	0.83	2.4	2.53	0.05	1.93	2.02
VAH20	10	84.08	82.04	2.04	0.65	2.04	2.14	0.03	1.69	1.77
VAH20	11	83.3	81.33	1.97	1.2	1.97	2.3	0.34	1.29	1.37
VAH20	12	82.15	77.54	4.61	1.9	4.61	4.99	0.05	3.55	3.63
VAH22	2	96.58	94.4	2.18	1.12	2.19	2.46	0.19	1.55	1.64
VAH22	3	94.35	93.35	1	0.84	1.1	1.3	0.45	0.62	0.71
VAH22	4	92.45	92.1	0.35	0.4	0.41	0.53	0.77	0.24	0.33
VAH22	5	90.82	90.43	0.39	0.75	0.73	0.84	0.5	0.32	0.41
VAH22	6	89.37	89.4	-0.03	0.42	0.3	0.42	0.81	0.08	0.17
VAH22	7	88.02	88.59	-0.57	0.61	0.69	0.84	0.49	0.34	0.43
VAH22	8	86.97	87.72	-0.75	0.46	0.77	0.88	0.25	0.52	0.6
VAH22	9	85.63	87.12	-1.49	0.45	1.5	1.55	0.02	1.26	1.35
VAH22	10	84.59	85.52	-0.93	0.95	1.09	1.33	0.48	0.55	0.63
VAH23	2	95.61	86.51	9.1	1.11	9.11	9.17	0	8.51	8.6
VAH23	3	93.73	89.37	4.36	1.19	4.37	4.52	0.02	3.73	3.82
VAH23	4	91.42	91.58	-0.16	1.4	1.09	1.41	0.71	0.31	0.41
VAH23	5	91.25	89.74	1.51	0.9	1.6	1.76	0.15	1.13	1.22
VAH23	6	89.82	86.59	3.23	1.07	3.26	3.41	0.02	2.69	2.77
VAH23	7	88.83	85.66	3.17	0.85	3.18	3.28	0.01	2.73	2.81
VAH23	8	88.2	87.18	1.02	1.01	1.24	1.44	0.44	0.71	0.8
VAH23	9	85.73	87.25	-1.52	1.85	1.71	2.4	0.66	0.72	0.81
VAH23	10	85.95	89.21	-3.26	1.31	3.26	3.51	0.02	2.59	2.68
VAH23	11	85.07	85.5	-0.43	1.03	0.79	1.11	0.76	0.26	0.34
VAH23	12	84.57	84.96	-0.39	1.04	0.82	1.11	0.75	0.28	0.36
VAH23	13	83.91	84.59	-0.68	1.01	0.98	1.22	0.61	0.44	0.53
VAH23	14	82.86	84.49	-1.63	1.23	1.73	2.04	0.32	1.07	1.15
VAH24	2	96.99	95.9	1.09	0.94	1.09	1.44	0.61	0.63	0.73
VAH24	3	94.54	95.71	-1.17	1.04	1.41	1.57	0.31	0.87	0.97

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH24	4	93.41	94.06	-0.65	1.46	1.32	1.6	0.61	0.5	0.59
VAH24	5	91.95	92.14	-0.19	1.42	1.16	1.44	0.68	0.35	0.45
VAH24	6	90.07	91.11	-1.04	1.39	1.33	1.73	0.61	0.55	0.64
VAH24	7	89.62	90.03	-0.41	1.6	1.24	1.65	0.74	0.36	0.45
VAH24	8	88.51	88.68	-0.17	1.74	1.28	1.75	0.76	0.36	0.44
VAH24	9	87.51	87.27	0.24	1.75	1.31	1.76	0.77	0.44	0.53
VAH24	10	86.72	86.25	0.47	1.64	1.28	1.7	0.75	0.51	0.59
VAH24	11	85.58	85.5	0.08	1.42	0.94	1.43	0.86	0.28	0.36
VAH24	12	84.58	84.76	-0.18	1.1	0.68	1.11	0.87	0.18	0.26
VAH24	13	83.65	84.15	-0.5	1.12	0.83	1.23	0.77	0.28	0.37
VAH24	14	83.2	83.19	0.01	1.2	0.91	1.2	0.73	0.26	0.34
VAH24	15	82.18	81.94	0.24	1.08	0.91	1.11	0.61	0.31	0.39
VAH24	16	80.33	81.2	-0.87	0.84	0.87	1.21	0.6	0.45	0.53
VAH25	2	96.82	94.34	2.48	0.77	2.48	2.6	0	2.05	2.14
VAH25	3	94.7	93.15	1.55	0.79	1.55	1.73	0.17	1.1	1.2
VAH25	4	92.95	92.01	0.94	0.72	1.01	1.18	0.37	0.6	0.7
VAH25	5	91.94	90.85	1.09	0.63	1.11	1.26	0.24	0.75	0.85
VAH25	6	90.23	89.44	0.79	0.49	0.83	0.93	0.21	0.57	0.66
VAH25	7	88.76	88.77	-0.01	0.87	0.59	0.87	0.79	0.16	0.24
VAH25	8	87.45	86.79	0.66	0.76	0.84	1	0.46	0.48	0.57
VAH25	9	86.21	85.71	0.5	0.84	0.81	0.97	0.57	0.39	0.47
VAH25	10	85.39	84.43	0.96	0.75	1.09	1.22	0.26	0.72	0.8
VAH25	11	83.6	83.19	0.41	0.74	0.65	0.85	0.65	0.25	0.34
VAH26	2	97.27	96.7	0.57	1.08	0.85	1.22	0.73	0.29	0.39
VAH26	3	95.31	96.05	-0.74	1.19	1.16	1.4	0.56	0.55	0.65
VAH26	4	93.73	94.95	-1.22	1.32	1.52	1.8	0.45	0.79	0.88
VAH26	5	92.34	93.64	-1.3	1.31	1.55	1.85	0.45	0.81	0.91
VAH26	6	90.3	90.71	-0.41	0.92	0.76	1.01	0.69	0.25	0.35
VAH26	7	89.37	90.02	-0.65	0.94	0.89	1.15	0.64	0.39	0.48
VAH26	8	88.51	89.39	-0.88	0.98	1.07	1.31	0.48	0.54	0.63
VAH26	9	87.38	89.38	-2	1.07	2.02	2.26	0.16	1.44	1.53
VAH26	10	86.25	87.18	-0.93	0.91	1.05	1.3	0.48	0.55	0.64
VAH26	11	84.62	87.58	-2.96	1.41	2.97	3.28	0.14	2.18	2.27
VAH26	12	84.5	86.31	-1.81	1.01	1.85	2.08	0.19	1.31	1.39
VAH26	13	83.91	84.27	-0.36	1.35	1.11	1.39	0.66	0.35	0.44
VAH26	14	82.55	83.82	-1.27	1.24	1.44	1.77	0.49	0.74	0.83
VAH26	15	81.97	83.36	-1.39	1.18	1.48	1.82	0.46	0.82	0.9
VAH26	16	81.38	82.85	-1.47	1.17	1.53	1.87	0.43	0.88	0.96
VAH26	17	80.22	82.5	-2.28	1.14	2.3	2.55	0.14	1.66	1.75
VAH26	18	79.5	82.62	-3.12	1.09	3.11	3.3	0.01	2.51	2.59
VAH26	19	78.82	80.35	-1.53	1.26	1.65	1.98	0.42	0.95	1.03
VAH26	20	78.39	78.44	-0.05	1.56	1.25	1.56	0.69	0.37	0.45
VAH26	21	77.66	77.82	-0.16	1.56	1.27	1.56	0.65	0.38	0.46
VAH27	2	97.01	95.3	1.71	0.55	1.72	1.8	0	1.44	1.53
VAH27	3	94.91	94.75	0.16	0.72	0.54	0.74	0.75	0.16	0.25
VAH27	4	93.07	92.77	0.3	0.92	0.76	0.97	0.67	0.24	0.34
VAH27	5	91.1	92.57	-1.47	0.89	1.51	1.72	0.26	1	1.1
VAH27	6	88.95	90.11	-1.16	0.99	1.33	1.52	0.38	0.77	0.86

Continued on next page

Table 19 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH27	7	87.71	86.83	0.88	0.84	1.02	1.22	0.44	0.55	0.63
VAH27	8	86.53	85.42	1.11	0.7	1.16	1.31	0.23	0.77	0.86
VAH27	9	85.51	84.23	1.28	0.73	1.31	1.47	0.22	0.9	0.99
VAH27	10	83.79	83.31	0.48	0.89	0.81	1.01	0.64	0.31	0.39
VAH27	11	82.76	81.36	1.4	0.9	1.47	1.67	0.26	0.96	1.04
VAH28	2	97.27	95.55	1.72	1.09	1.73	2.03	0.33	1.12	1.22
VAH28	3	95.47	94.51	0.96	1.09	1.19	1.46	0.52	0.57	0.66
VAH28	4	93.93	92.75	1.18	0.91	1.27	1.49	0.37	0.76	0.85
VAH28	5	92.68	91.8	0.88	0.81	1.01	1.2	0.42	0.56	0.65
VAH28	6	92.08	90.78	1.3	0.81	1.35	1.53	0.25	0.9	0.99
VAH28	7	90.48	89.58	0.9	0.93	1.09	1.29	0.45	0.57	0.66
VAH28	8	89.43	89.04	0.39	1.07	0.9	1.14	0.66	0.32	0.41
VAH28	9	88.61	88.45	0.16	1.03	0.81	1.05	0.7	0.24	0.33
VAH28	10	87.92	88.19	-0.27	1.01	0.79	1.04	0.7	0.24	0.32
VAH28	11	86.99	86.87	0.12	1.17	0.9	1.17	0.7	0.25	0.34
VAH28	12	86.42	85.81	0.61	1.28	1.13	1.42	0.64	0.41	0.49
VAH28	13	85.78	85.42	0.36	1.32	1.08	1.37	0.68	0.34	0.43
VAH28	14	85.09	84.31	0.78	1.24	1.2	1.46	0.56	0.51	0.59
VAH28	15	84.36	83.27	1.09	1.2	1.37	1.62	0.45	0.7	0.79
VAH28	16	83.51	82.22	1.29	1.12	1.48	1.7	0.36	0.86	0.95
VAH28	17	83.29	81.27	2.02	1.01	2.06	2.26	0.13	1.49	1.58
VAH28	18	82.75	81.5	1.25	1.05	1.41	1.63	0.36	0.83	0.91
VAH28	19	82.13	81.67	0.46	1.02	0.91	1.12	0.61	0.35	0.43
VAH28	20	81.78	80.93	0.85	1.01	1.1	1.32	0.5	0.54	0.62
VAH28	21	81.12	80.81	0.31	1	0.84	1.04	0.65	0.28	0.36
VAH28	22	80.48	79.37	1.11	1.13	1.34	1.58	0.46	0.7	0.78
VAH28	23	80.56	79.18	1.38	1.32	1.61	1.91	0.44	0.87	0.95
VAH30	2	96.89	92.57	4.32	2.08	4.32	4.79	0.09	3.16	3.25
VAH30	3	94.67	93.04	1.63	1.76	1.69	2.4	0.64	0.78	0.87
VAH30	4	93.05	91.48	1.57	2.34	2.05	2.82	0.68	0.76	0.85
VAH30	5	91.4	89.66	1.74	2.5	2.36	3.05	0.61	0.93	1.02
VAH30	6	89.81	88.43	1.38	1.94	1.79	2.38	0.66	0.73	0.82
VAH30	7	88.96	87	1.96	2.1	2.24	2.87	0.56	1.05	1.14
VAH30	8	87.59	86.22	1.37	1.9	1.69	2.34	0.67	0.66	0.74
VAH30	9	86.78	85.53	1.25	2.1	1.77	2.44	0.68	0.61	0.69
VAH30	10	85.77	84.88	0.89	2.09	1.73	2.27	0.7	0.56	0.64
VAH30	11	84.53	83.37	1.16	1.77	1.69	2.11	0.6	0.68	0.77
VAH30	12	84.13	82.28	1.85	1.34	1.91	2.28	0.38	1.15	1.23
VAH30	13	83.13	81.85	1.28	1.3	1.52	1.82	0.45	0.78	0.86
VAH30	14	82.14	81.32	0.82	1.26	1.25	1.5	0.56	0.54	0.62
VAH30	15	81.52	80.61	0.91	1.08	1.19	1.41	0.48	0.59	0.67
VAH30	16	80.73	80.01	0.72	0.81	0.89	1.08	0.5	0.46	0.54
VAH30	17	79.91	79.43	0.48	0.53	0.59	0.72	0.49	0.3	0.38
VAH30	18	79.63	79.26	0.37	0.42	0.47	0.56	0.46	0.24	0.32

Table 19: CNN Monte Carlo Dropout - SOH Capacity Test Results

F Convolution Neural Network with Monte Carlo Dropout - RUL Capacity Test Results

Note that all the results are rounded to 2 and since the input to the network has a window size of 2, we do not have results of capacity test '1' for each Vahana in the table. Mean indicates distribution mean prediction and SD is the standard deviation.

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh(%)	CRPS NRG	CRPS PWM
VAH01	2	561	350.32	210.68	68.39	210.85	221.5	0.02	172.4	172.75
VAH01	3	510	322.88	187.12	64.64	187.23	197.97	0.04	150.93	151.25
VAH01	4	459	270.22	188.78	61.05	189.03	198.41	0.03	154.88	155.15
VAH01	5	408	228.26	179.74	60.46	180.58	189.63	0.03	147.36	147.59
VAH01	6	357	190	167	60.69	168.02	177.69	0.05	134.44	134.63
VAH01	7	306	164.33	141.67	59.87	143.75	153.81	0.08	110.93	111.09
VAH01	8	255	137.69	117.31	52.68	119.02	128.6	0.11	89.92	90.06
VAH01	9	204	100.13	103.87	46.85	106.4	113.95	0.11	81.19	81.29
VAH01	10	153	83.66	69.34	41.89	73.82	81.01	0.2	51.15	51.24
VAH01	11	102	42.1	59.9	26.08	61.64	65.33	0.09	48.27	48.31
VAH01	12	51	31.34	19.66	21.78	26.15	29.34	0.34	15.87	15.9
VAH01	13	0	22.82	-22.82	16.98	22.82	28.44	0.47	15.79	15.81
VAH02	2	459	462.63	-3.63	77.58	62.5	77.66	0.68	18.54	19
VAH02	3	408	431.63	-23.63	73.87	60.31	77.56	0.67	18.83	19.26
VAH02	4	357	391.59	-34.59	67.17	60.9	75.55	0.62	22.9	23.29
VAH02	5	306	347.41	-41.41	60.11	57.55	72.99	0.6	23.79	24.14
VAH02	6	255	305.09	-50.09	54.78	59.32	74.23	0.54	28.56	28.86
VAH02	7	204	239.93	-35.93	54.76	52.05	65.5	0.61	21.22	21.46
VAH02	8	153	150.94	2.06	40.91	32.87	40.96	0.69	10.17	10.32
VAH02	9	102	56.65	45.35	25.46	47.54	52.01	0.17	33.78	33.84
VAH02	10	51	33.29	17.71	19.43	23.33	26.29	0.38	13.02	13.06
VAH02	11	0	29.18	-29.18	17.92	29.18	34.24	0.32	19.7	19.73
VAH05	2	714	632.18	81.82	97.9	104.14	127.59	0.52	48.92	49.55
VAH05	3	663	618.34	44.66	96.02	84.7	105.89	0.62	30.86	31.48
VAH05	4	612	600.34	11.66	94.43	76.55	95.15	0.66	23.06	23.66
VAH05	5	561	555.49	5.51	91.12	73.47	91.28	0.68	22.05	22.6
VAH05	6	510	523.74	-13.74	84.85	68.7	85.96	0.67	20.78	21.31
VAH05	7	459	506.68	-47.68	83.45	76.2	96.12	0.61	28.87	29.38
VAH05	8	408	404.37	3.63	72.21	57.12	72.3	0.68	16.7	17.11
VAH05	9	357	340.87	16.13	60.64	49.69	62.75	0.67	15.86	16.2
VAH05	10	306	279.71	26.29	56.1	49.81	61.96	0.61	18.2	18.48
VAH05	11	255	224.84	30.16	55.62	51.65	63.27	0.58	20.09	20.32
VAH05	12	204	158.23	45.77	53.09	59.06	70.09	0.46	29.23	29.39
VAH05	13	153	84.56	68.44	41.52	72.66	80.06	0.19	50.03	50.12
VAH05	14	102	60.64	41.36	33.08	47.66	52.96	0.28	30.06	30.12
VAH05	15	51	45.94	5.06	27.06	21.55	27.53	0.72	7.89	7.94
VAH05	16	0	30.8	-30.8	17.56	30.8	35.46	0.13	22.54	22.57
VAH06	2	561	484.57	76.43	83.13	93.93	112.92	0.48	47.06	47.54
VAH06	3	510	457.9	52.1	79.48	77.39	95.03	0.58	32.65	33.11
VAH06	4	459	419.48	39.52	76.2	69.42	85.83	0.61	26.22	26.64
VAH06	5	408	399.18	8.82	77.23	63.32	77.73	0.65	19.57	19.97
VAH06	6	357	327.44	29.56	71.86	62.75	77.7	0.62	22.31	22.64

Continued on next page

Table 20 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH06	7	306	268.58	37.42	72.86	68.27	81.91	0.58	27.47	27.74
VAH06	8	255	214.62	40.38	61.84	62.06	73.86	0.53	27.47	27.68
VAH06	9	204	154.88	49.12	51.3	61.7	71.02	0.41	33.77	33.93
VAH06	10	153	123.02	29.98	45.68	45.6	54.63	0.54	20.45	20.57
VAH06	11	102	86.16	15.84	41.5	35.28	44.43	0.63	13.24	13.33
VAH06	12	51	67.98	-16.98	35.5	28.66	39.35	0.73	9.4	9.47
VAH06	13	0	43.87	-43.87	28.17	43.87	52.14	0.35	28.56	28.61
VAH09	2	465	466.89	-1.89	72.76	57.97	72.78	0.68	16.83	17.29
VAH09	3	414	425.33	-11.33	67.62	54.8	68.56	0.69	16.64	17.06
VAH09	4	363	405.88	-42.88	65.07	61.64	77.93	0.61	25.02	25.42
VAH09	5	351	399.87	-48.87	63.45	63.86	80.09	0.58	28.14	28.54
VAH09	6	339	397.9	-58.9	64.37	69.82	87.25	0.53	33.5	33.9
VAH09	7	327	389.14	-62.14	64.29	72.21	89.41	0.51	36.16	36.55
VAH09	8	315	384.24	-69.24	61.92	76.98	92.89	0.44	42.16	42.55
VAH09	9	303	386.42	-83.42	63.23	88.45	104.68	0.4	52.69	53.07
VAH09	10	291	377.89	-86.89	60.48	90.4	105.86	0.32	56.5	56.88
VAH09	11	279	366.94	-87.94	58.53	90.5	105.64	0.32	57.46	57.83
VAH09	12	204	329.37	-125.37	56.3	125.56	137.43	0.1	93.97	94.3
VAH09	13	153	277.38	-124.38	50.99	124.39	134.43	0.07	95.75	96.02
VAH09	14	102	228.63	-126.63	45.75	126.63	134.64	0.02	101.51	101.74
VAH09	15	51	193.98	-142.98	42.16	142.98	149.06	0	119.29	119.49
VAH09	16	0	131.44	-131.44	41.58	131.44	137.86	0.01	108.13	108.27
VAH10	2	561	420.01	140.99	72.33	142.72	158.46	0.17	101.88	102.3
VAH10	3	510	391.85	118.15	63.88	120.04	134.31	0.19	84.1	84.49
VAH10	4	459	358.31	100.69	60.6	103.53	117.51	0.25	69.38	69.74
VAH10	5	408	318.91	89.09	55.88	91.99	105.16	0.26	60.47	60.78
VAH10	6	357	273.03	83.97	48.27	86.3	96.85	0.21	59.15	59.43
VAH10	7	306	207.8	98.2	42.28	98.84	106.92	0.1	75.03	75.24
VAH10	8	255	150.55	104.45	33.49	104.49	109.69	0.03	85.7	85.85
VAH10	9	204	116.61	87.39	30.21	87.68	92.47	0.04	70.85	70.97
VAH10	10	153	91.64	61.36	29.92	62.14	68.26	0.17	45.24	45.34
VAH10	11	102	58.6	43.4	21.98	44.65	48.65	0.14	32.53	32.59
VAH10	12	51	46.97	4.03	19.32	15.71	19.73	0.68	5.32	5.37
VAH10	13	0	38.58	-38.58	15.23	38.58	41.47	0.01	30.38	30.42
VAH11	2	764	836.88	-72.88	153.05	134.05	169.52	0.64	47.86	48.69
VAH11	3	713	826.2	-113.2	150.37	149.91	188.21	0.58	65.12	65.95
VAH11	4	662	819.9	-157.9	147.95	175.43	216.38	0.49	91.85	92.67
VAH11	5	611	809.3	-198.3	149.84	206.55	248.54	0.39	122.02	122.83
VAH11	6	560	799.7	-239.7	144.16	243.53	279.71	0.26	162.62	163.42
VAH11	7	509	797.11	-288.11	144.21	288.66	322.18	0.16	207.23	208.02
VAH11	8	458	777.66	-319.66	144.87	319.92	350.95	0.12	237.88	238.66
VAH11	9	407	783.07	-376.07	144.67	376.2	402.94	0.04	295.1	295.88
VAH11	10	356	777.77	-421.77	142.69	421.77	445.25	0.02	340.96	341.74
VAH11	11	305	768.89	-463.89	138.48	463.89	484.12	0	385.81	386.58
VAH11	12	254	763.86	-509.86	140.78	509.86	528.94	0	431.09	431.85
VAH11	13	203	754.34	-551.34	133.59	551.34	567.3	0	476.53	477.28
VAH11	14	152	740.95	-588.95	127.62	588.95	602.62	0	516.62	517.37
VAH11	15	101	739.53	-638.53	135.87	638.53	652.83	0	561.8	562.54

Continued on next page

Table 20 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH11	16	50	726.97	-676.97	131.51	676.97	689.62	0	603.39	604.12
VAH11	17	0	702.49	-702.49	128.49	702.49	714.14	0	630.28	630.98
VAH12	2	714	571.62	142.38	119.98	158.99	186.19	0.38	91.48	92.05
VAH12	3	663	552.41	110.59	123.57	138.89	165.83	0.48	69.45	70
VAH12	4	612	539.32	72.68	117.51	113.54	138.17	0.58	47.05	47.59
VAH12	5	561	542.4	18.6	117.04	94.99	118.51	0.69	29.15	29.69
VAH12	6	510	515.56	-5.56	119.43	93.51	119.56	0.71	27.05	27.57
VAH12	7	459	501.49	-42.49	112.08	95.43	119.87	0.66	32.23	32.73
VAH12	8	408	458.32	-50.32	108.99	95.43	120.04	0.63	33.68	34.14
VAH12	9	357	449.78	-92.78	113.51	113.02	146.61	0.59	49.87	50.32
VAH12	10	306	425.48	-119.48	110.59	130.78	162.81	0.49	68.98	69.41
VAH12	11	255	375.68	-120.68	100.61	129.5	157.12	0.44	73.08	73.46
VAH12	12	204	303.79	-99.79	95.87	111.08	138.38	0.51	57.42	57.73
VAH12	13	153	278.83	-125.83	95.59	130.43	158.02	0.41	77.24	77.52
VAH12	14	102	257.99	-155.99	93.06	156.83	181.64	0.27	105.33	105.59
VAH12	15	51	230.96	-179.96	83.64	179.96	198.44	0.12	133.16	133.39
VAH12	16	0	206.92	-206.92	85.78	206.92	224	0.04	159.25	159.46
VAH13	2	612	516.73	95.27	94.56	111.33	134.23	0.47	58.08	58.6
VAH13	3	561	508.63	52.37	83.77	80.88	98.79	0.58	33.57	34.08
VAH13	4	510	484.42	25.58	82.29	69.9	86.17	0.65	23.65	24.14
VAH13	5	459	452.23	6.77	74.98	59.91	75.29	0.66	17.63	18.08
VAH13	6	408	405.96	2.04	64.61	50.93	64.64	0.68	14.61	15.01
VAH13	7	357	337.76	19.24	57.31	49.62	60.45	0.65	17.23	17.57
VAH13	8	306	273.81	32.19	51.67	49.72	60.87	0.56	20.73	21
VAH13	9	255	257.03	-2.03	54.2	43.04	54.24	0.7	12.62	12.88
VAH13	10	204	248.95	-44.95	48.26	52.73	65.96	0.52	25.65	25.9
VAH13	11	153	145.24	7.76	42.08	34.16	42.79	0.66	10.65	10.8
VAH13	12	102	68.21	33.79	31.38	40.81	46.11	0.34	23.86	23.92
VAH13	13	51	46.42	4.58	25.93	21.06	26.33	0.68	7.01	7.06
VAH13	14	0	58.32	-58.32	31.95	58.32	66.5	0.18	41.64	41.69
VAH15	2	408	370.17	37.83	64.96	62.39	75.17	0.56	25.92	26.29
VAH15	3	357	339.41	17.59	59.8	50.58	62.33	0.65	16.91	17.24
VAH15	4	306	294.19	11.81	53.59	44.49	54.88	0.65	14.25	14.54
VAH15	5	255	230.04	24.96	52.33	46.62	57.98	0.62	17.09	17.32
VAH15	6	204	156.93	47.07	48.88	58.16	67.86	0.42	30.79	30.95
VAH15	7	153	107.95	45.05	39.45	52.18	59.89	0.36	30.25	30.36
VAH15	8	102	65.37	36.63	28.49	41.43	46.4	0.28	25.98	26.05
VAH15	9	51	47.73	3.27	23.47	18.14	23.7	0.72	5.81	5.86
VAH15	10	0	34.98	-34.98	17.69	34.98	39.2	0.14	25.47	25.5
VAH16	2	408	512.96	-104.96	105.06	119.41	148.51	0.51	60.22	60.73
VAH16	3	357	482.36	-125.36	101.16	133.72	161.08	0.42	76.71	77.19
VAH16	4	306	436.53	-130.53	98.05	135.11	163.25	0.41	80.32	80.76
VAH16	5	255	369.65	-114.65	83.24	118.34	141.68	0.38	71.56	71.93
VAH16	6	204	322.59	-118.59	74.18	119.83	139.88	0.3	78.62	78.94
VAH16	7	153	289.19	-136.19	74.22	136.49	155.1	0.22	95.22	95.51
VAH16	8	102	244.04	-142.04	69.89	142.04	158.3	0.13	103.84	104.08
VAH16	9	51	205.89	-154.89	71.36	154.89	170.54	0.1	114.92	115.13

Continued on next page

Table 20 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH16	10	0	103.14	-103.14	53.45	103.14	116.17	0.14	74.92	75.03
VAH17	2	510	439.27	70.73	82.64	91.32	108.78	0.5	45.05	45.49
VAH17	3	459	400.09	58.91	76.31	80.65	96.41	0.51	37.71	38.11
VAH17	4	408	364.65	43.35	69.48	67.57	81.9	0.56	28.62	28.99
VAH17	5	357	297.3	59.7	61.38	71.48	85.63	0.47	37.03	37.33
VAH17	6	306	286.61	19.39	62.67	53.86	65.6	0.62	18.66	18.94
VAH17	7	255	243.51	11.49	57.28	46.63	58.42	0.67	14.76	15
VAH17	8	204	213.94	-9.94	56.04	44.33	56.92	0.69	13.03	13.24
VAH17	9	153	186.32	-33.32	53.12	47.39	62.71	0.65	17.78	17.96
VAH17	10	102	149.18	-47.18	48.14	51.19	67.41	0.57	24.86	25.01
VAH17	11	51	128.23	-77.23	44.56	77.39	89.16	0.24	52.95	53.08
VAH17	12	0	101.63	-101.63	36.42	101.63	107.96	0.01	81.91	82.02
VAH20	2	408	385.09	22.91	66.09	56.61	69.95	0.65	19.36	19.74
VAH20	3	357	361.67	-4.67	59.82	48	60	0.69	14.34	14.7
VAH20	4	306	315.55	-9.55	53.79	43.16	54.63	0.7	13.02	13.34
VAH20	5	255	277.04	-22.04	51.24	44.19	55.78	0.65	15.4	15.68
VAH20	6	204	231.03	-27.03	49.66	44.74	56.54	0.63	16.68	16.91
VAH20	7	153	189.29	-36.29	46.23	46.24	58.77	0.59	20.45	20.64
VAH20	8	102	137.8	-35.8	39.38	41.52	53.22	0.55	19.69	19.82
VAH20	9	51	82.48	-31.48	30.29	33.95	43.68	0.53	17.12	17.2
VAH20	10	0	38.52	-38.52	18.99	38.52	42.95	0.09	28.65	28.69
VAH22	2	408	395.09	12.91	63.95	51.97	65.24	0.69	16.05	16.45
VAH22	3	357	365.17	-8.17	59.47	48.11	60.03	0.68	14.52	14.88
VAH22	4	306	320.52	-14.52	50.04	41.41	52.11	0.65	13.04	13.36
VAH22	5	255	280.12	-25.12	48.58	42.6	54.69	0.64	15.43	15.71
VAH22	6	204	213.92	-9.92	43.12	34.77	44.24	0.69	10.55	10.77
VAH22	7	153	150.7	2.3	36.92	29.3	36.99	0.68	8.66	8.81
VAH22	8	102	97.33	4.67	32.82	26.59	33.15	0.66	8.2	8.3
VAH22	9	51	60.34	-9.34	28.29	21.95	29.79	0.74	6.58	6.64
VAH22	10	0	40.04	-40.04	20.27	40.04	44.88	0.12	29.26	29.3
VAH23	2	510	107.13	402.87	38.28	402.87	404.68	0	382.03	382.13
VAH23	3	459	186.38	272.62	42.78	272.62	275.96	0	248.57	248.76
VAH23	4	408	205.65	202.35	45.49	202.35	207.4	0	176.72	176.92
VAH23	5	357	135.45	221.55	44.81	221.55	226.03	0	196.5	196.64
VAH23	6	306	70.16	235.84	25.62	235.84	237.23	0	221.49	221.56
VAH23	7	255	64.42	190.58	25.66	190.58	192.3	0	176.59	176.65
VAH23	8	204	59.29	144.71	25.74	144.71	146.98	0	130.81	130.87
VAH23	9	153	57.42	95.58	24.9	95.93	98.77	0.01	82.57	82.63
VAH23	10	102	73.96	28.04	33.69	37.5	43.83	0.45	18.89	18.96
VAH23	11	51	42.09	8.91	18.77	17.53	20.78	0.57	7.24	7.28
VAH23	12	0	33.72	-33.72	15.42	33.72	37.08	0.08	25.58	25.62
VAH24	2	510	497.39	12.61	84.01	68.53	84.95	0.67	21.01	21.51
VAH24	3	459	495.33	-36.33	91.56	77.14	98.5	0.67	25.61	26.11
VAH24	4	408	410.59	-2.59	70.75	56.88	70.8	0.7	16.98	17.39
VAH24	5	357	357.37	-0.37	60.26	47.86	60.26	0.68	13.94	14.3
VAH24	6	306	314.79	-8.79	61.06	49.19	61.69	0.69	14.8	15.11
VAH24	7	255	210.99	44.01	65.94	66.49	79.28	0.52	29.93	30.14
VAH24	8	204	129.39	74.61	43.78	79.13	86.5	0.19	55.59	55.72

Continued on next page

Table 20 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH24	9	153	106.58	46.42	41.08	54.03	61.99	0.36	31.57	31.68
VAH24	10	102	85.87	16.13	36.98	33.37	40.34	0.59	12.98	13.07
VAH24	11	51	59.85	-8.85	29.21	22.74	30.52	0.72	6.73	6.79
VAH24	12	0	45.42	-45.42	26.93	45.42	52.8	0.27	31.18	31.23
VAH25	2	461	468.57	-7.57	77.39	62.43	77.76	0.67	18.72	19.19
VAH25	3	410	410.68	-0.68	65.39	52.79	65.39	0.67	15.73	16.14
VAH25	4	359	355.27	3.73	58.08	47.15	58.2	0.67	14.27	14.62
VAH25	5	307	312.02	-5.02	51.14	41.42	51.39	0.69	12.52	12.83
VAH25	6	256	275.77	-19.77	45.32	38.88	49.45	0.66	13.45	13.73
VAH25	7	205	242.2	-37.2	46.05	46.35	59.2	0.57	20.67	20.91
VAH25	8	154	153.89	0.11	40.35	32.34	40.35	0.67	9.65	9.8
VAH25	9	103	84.54	18.46	31.91	31.08	36.87	0.53	13.94	14.03
VAH25	10	51	54.41	-3.41	26.47	19.31	26.69	0.77	5.49	5.54
VAH25	11	0	29.55	-29.55	16.79	29.55	33.98	0.25	20.48	20.51
VAH26	2	561	583.36	-22.36	92.83	74.83	95.49	0.69	22.79	23.37
VAH26	3	510	570.19	-60.19	95.48	89.97	112.87	0.6	36.12	36.69
VAH26	4	459	535.31	-76.31	86.22	92.17	115.14	0.54	43.7	44.24
VAH26	5	408	500.69	-92.69	80.2	101.01	122.57	0.44	55.56	56.06
VAH26	6	306	398.52	-92.52	68.62	97.1	115.19	0.38	58.45	58.85
VAH26	7	255	291.7	-36.7	52.39	50.76	63.97	0.6	21.24	21.53
VAH26	8	204	254.15	-50.15	47.56	55.91	69.12	0.5	29.27	29.53
VAH26	9	153	231.81	-78.81	48.37	80.21	92.47	0.26	53.01	53.25
VAH26	10	51	153.09	-102.09	42.22	102.11	110.48	0.06	78.54	78.69
VAH26	11	0	125.81	-125.81	44.12	125.81	133.32	0	101.3	101.43
VAH27	2	459	438.51	20.49	73.4	62.2	76.2	0.65	20.81	21.24
VAH27	3	408	411.25	-3.25	66.37	51.98	66.45	0.71	14.95	15.36
VAH27	4	357	375.42	-18.42	60.62	49.7	63.36	0.67	15.74	16.11
VAH27	5	306	331.05	-25.05	53.74	46.64	59.29	0.65	16.44	16.77
VAH27	6	204	238.52	-34.52	42.4	43.67	54.68	0.56	19.71	19.95
VAH27	7	153	144.55	8.45	42.17	34.5	43.01	0.66	11	11.15
VAH27	8	102	82.92	19.08	31.97	31.33	37.23	0.55	13.99	14.07
VAH27	9	51	42.71	8.29	23.52	19.68	24.94	0.65	7.71	7.76
VAH27	10	0	19.36	-19.36	11.21	19.36	22.37	0.23	13.57	13.58
VAH28	2	669	603.97	65.03	97.27	96.57	117	0.56	41.47	42.08
VAH28	3	618	587.9	30.1	95.78	81.7	100.4	0.64	27.72	28.31
VAH28	4	567	560.34	6.66	88.72	70.78	88.97	0.68	20.93	21.49
VAH28	5	516	541.32	-25.32	86.68	71.79	90.31	0.66	22.73	23.28
VAH28	6	464	442.34	21.66	74.6	62.77	77.68	0.66	20.98	21.42
VAH28	7	413	381.56	31.44	69.95	61.5	76.69	0.63	22.63	23.01
VAH28	8	362	333.63	28.37	59.86	53.08	66.25	0.62	19.31	19.64
VAH28	9	311	290.12	20.88	58	49.93	61.65	0.65	17.39	17.68
VAH28	10	260	260.74	-0.74	56.62	44.84	56.62	0.68	13.05	13.31
VAH28	11	209	178.45	30.55	52.07	49.15	60.37	0.59	19.76	19.94
VAH28	12	158	126.51	31.49	49.58	49.5	58.74	0.52	21.74	21.87
VAH28	13	107	111.5	-4.5	44.78	34.88	45	0.72	10.26	10.37
VAH28	14	51	92.83	-41.83	42.77	44.85	59.82	0.58	21.56	21.66
VAH28	15	0	75.53	-75.53	37.19	75.53	84.19	0.12	55.53	55.6
VAH30	2	459	426.65	32.35	85.46	73.05	91.37	0.65	25.21	25.63

Continued on next page

Table 20 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	MAE	RMSE	Sh (%)	CRPS NRG	CRPS PWM
VAH30	3	408	421.21	-13.21	92.62	74.11	93.56	0.7	22.09	22.52
VAH30	4	357	376.99	-19.99	94.53	75.56	96.62	0.7	23.18	23.56
VAH30	5	306	306.54	-0.54	96.8	76.25	96.8	0.7	21.85	22.16
VAH30	6	255	237.85	17.15	82.82	67.34	84.58	0.68	20.64	20.88
VAH30	7	204	182.17	21.83	73	60.79	76.19	0.67	19.81	20
VAH30	8	153	140.16	12.84	54.88	44.14	56.36	0.68	13.94	14.08
VAH30	9	102	123.69	-21.69	46.64	39.16	51.44	0.68	13.23	13.35
VAH30	10	51	109.85	-58.85	44.25	61.37	73.63	0.38	37.07	37.18
VAH30	11	0	86.16	-86.16	33.75	86.16	92.54	0.04	67.63	67.71

Table 20: CNN Monte Carlo Dropout - RUL Capacity Test Results

G Mixture Density Network - SOH Capacity Test Results

Note that all the results are rounded to 2. Mean indicates distribution mean prediction, SD is the standard deviation and CRPS Norm is the CRPS for normal distribution. Sharpness is 100 percent if the actual value falls within one standard deviation of distributional mean of that capacity test indicated as 1 , 0 otherwise.

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH01	1	100	98.6	1.4	2.14	1	0.85
VAH01	2	96.95	96.09	0.86	1.86	1	0.59
VAH01	3	94.98	94.44	0.54	1.85	1	0.5
VAH01	4	93.33	91.94	1.39	1.86	1	0.83
VAH01	5	91.85	89.54	2.31	1.87	0	1.45
VAH01	6	90.5	87.38	3.12	1.87	0	2.14
VAH01	7	89.42	86.06	3.35	1.9	0	2.34
VAH01	8	88.34	85.81	2.53	0.64	0	2.17
VAH01	9	87.45	84.73	2.73	0.65	0	2.36
VAH01	10	86.83	84.71	2.12	0.61	0	1.78
VAH01	11	85.26	82.52	2.74	0.67	0	2.37
VAH01	12	85.04	82.11	2.93	0.66	0	2.56
VAH01	13	83.83	80.51	3.33	0.67	0	2.95
VAH01	14	83.46	80.7	2.76	0.65	0	2.4
VAH01	15	82.66	79.3	3.35	0.67	0	2.98
VAH01	16	80.81	76.73	4.07	0.66	0	3.7
VAH01	17	80.98	76.79	4.19	0.66	0	3.81
VAH02	1	100	100	0	1.19	1	0.28
VAH02	2	96.62	98.03	-1.4	1.35	0	0.85
VAH02	3	94.47	95.12	-0.65	1.18	1	0.41
VAH02	4	92.69	92.99	-0.3	1.08	1	0.28
VAH02	5	91.14	89.51	1.64	0.99	0	1.12
VAH02	6	89.64	87.4	2.24	0.89	0	1.75
VAH02	7	88.62	84.9	3.72	0.8	0	3.27
VAH02	8	87.52	86.88	0.64	0.35	0	0.45
VAH02	9	86.68	86.56	0.12	0.33	1	0.09

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH02	10	85.52	84.63	0.89	0.36	0	0.69
VAH02	11	83.83	83.66	0.17	0.46	1	0.13
VAH02	12	83.54	82.83	0.71	0.45	0	0.48
VAH02	13	82.34	80.57	1.77	0.34	0	1.58
VAH05	1	100	99.55	0.45	1.81	1	0.47
VAH05	2	97.33	98.64	-1.31	1.35	1	0.78
VAH05	3	95.41	96.29	-0.89	0.95	1	0.53
VAH05	4	94.02	94.35	-0.33	0.65	1	0.22
VAH05	5	92.84	93.41	-0.57	0.6	1	0.34
VAH05	6	91.65	92.33	-0.67	0.61	0	0.41
VAH05	7	90.79	91.3	-0.51	0.58	1	0.3
VAH05	8	89.92	90.51	-0.58	0.54	0	0.35
VAH05	9	89	89.37	-0.37	0.57	1	0.22
VAH05	10	88.22	88.53	-0.31	0.54	1	0.19
VAH05	11	87.5	87.45	0.05	0.52	1	0.12
VAH05	12	86.95	86.39	0.56	0.53	0	0.34
VAH05	13	86.33	85.9	0.43	0.5	1	0.26
VAH05	14	85.91	85.42	0.49	0.49	0	0.3
VAH05	15	85.14	84.2	0.95	0.43	0	0.71
VAH05	16	84.68	83.96	0.72	0.5	0	0.47
VAH05	17	84.22	82.9	1.33	0.51	0	1.05
VAH05	18	83.77	82.59	1.19	0.49	0	0.91
VAH05	19	83.09	81.72	1.37	0.51	0	1.08
VAH05	20	82.53	81.31	1.22	0.5	0	0.94
VAH05	21	81.73	80.32	1.42	0.53	0	1.12
VAH05	22	81.17	79.86	1.31	0.55	0	1
VAH05	23	80.89	79.9	0.99	0.48	0	0.73
VAH05	24	80.11	78.68	1.43	0.59	0	1.1
VAH05	25	79.33	77.85	1.47	0.63	0	1.13
VAH05	26	79.06	77.82	1.23	0.6	0	0.9
VAH05	27	78.41	76.98	1.43	0.65	0	1.07
VAH05	28	77.71	76.03	1.68	0.72	0	1.28
VAH05	29	77.3	75.93	1.37	0.66	0	1
VAH06	1	100	100	0	0.75	1	0.17
VAH06	2	97	98.66	-1.67	0.58	0	1.34
VAH06	3	94.88	95.73	-0.85	0.46	0	0.6
VAH06	4	93.5	94.7	-1.2	0.42	0	0.97
VAH06	5	92.22	92.03	0.19	0.34	1	0.12
VAH06	6	90.84	90.94	-0.1	0.32	1	0.09
VAH06	7	90.03	90.14	-0.11	0.37	1	0.1
VAH06	8	88.52	88.28	0.25	0.34	1	0.15
VAH06	9	87.91	88.24	-0.33	0.31	0	0.2
VAH06	10	86.47	85.97	0.5	0.32	0	0.34
VAH06	11	85.87	85.96	-0.1	0.32	1	0.09
VAH06	12	85.43	84.86	0.57	0.38	0	0.38
VAH06	13	84.76	85.02	-0.26	0.4	1	0.16
VAH06	14	84.01	84.68	-0.67	0.42	0	0.45
VAH06	15	83.28	83.88	-0.6	0.43	0	0.39

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH06	16	81.97	81.68	0.29	0.38	1	0.17
VAH06	17	81.18	81.72	-0.55	0.45	0	0.34
VAH06	18	80.44	80.73	-0.28	0.47	1	0.18
VAH06	19	79.69	79.4	0.28	0.43	1	0.17
VAH09	1	100	98.89	1.11	0.57	0	0.8
VAH09	2	96.98	96.57	0.41	0.49	1	0.24
VAH09	3	94.89	95.3	-0.4	0.44	1	0.24
VAH09	4	93.18	94.12	-0.94	0.41	0	0.71
VAH09	5	92.52	93.76	-1.24	0.4	0	1.02
VAH09	6	91.93	93.05	-1.12	0.4	0	0.9
VAH09	7	91.39	92.93	-1.54	0.39	0	1.31
VAH09	8	90.87	92.01	-1.14	0.37	0	0.93
VAH09	9	90.4	91.79	-1.39	0.37	0	1.18
VAH09	10	89.98	92.04	-2.06	0.38	0	1.84
VAH09	11	89.56	91.48	-1.92	0.37	0	1.71
VAH09	12	88.11	89.5	-1.39	0.35	0	1.2
VAH09	13	87.09	88.68	-1.59	0.34	0	1.39
VAH09	14	86.23	87.84	-1.6	0.34	0	1.41
VAH09	15	85.35	86.7	-1.35	0.35	0	1.16
VAH09	16	84.48	86.16	-1.68	0.34	0	1.48
VAH09	17	83.89	84.99	-1.1	0.36	0	0.9
VAH09	18	83.11	84.54	-1.43	0.37	0	1.22
VAH09	19	82.41	84.75	-2.35	0.37	0	2.14
VAH09	20	79.76	81.96	-2.19	0.43	0	1.95
VAH09	21	78.95	80.05	-1.1	0.46	0	0.84
VAH10	1	100	100	0	0.37	1	0.09
VAH10	2	97.01	98.07	-1.05	0.36	0	0.85
VAH10	3	94.97	95.91	-0.94	0.34	0	0.75
VAH10	4	93.29	93.32	-0.03	0.32	1	0.07
VAH10	5	91.87	91.88	-0.01	0.31	1	0.07
VAH10	6	90.64	90.71	-0.06	0.31	1	0.08
VAH10	7	89.52	90.08	-0.55	0.33	0	0.38
VAH10	8	88.45	88.79	-0.35	0.3	0	0.22
VAH10	9	87.46	87.98	-0.52	0.31	0	0.36
VAH10	10	86.72	87.11	-0.39	0.31	0	0.24
VAH10	11	85.7	86.52	-0.82	0.32	0	0.64
VAH10	12	85.44	85.58	-0.14	0.33	1	0.1
VAH10	13	84.65	85.01	-0.37	0.33	0	0.22
VAH10	14	83.84	84.38	-0.54	0.35	0	0.36
VAH10	15	83.16	83.8	-0.64	0.35	0	0.45
VAH10	16	82.1	82.81	-0.71	0.35	0	0.52
VAH10	17	81.39	82.43	-1.05	0.39	0	0.83
VAH10	18	80.76	81.94	-1.18	0.39	0	0.96
VAH10	19	80.19	81.53	-1.34	0.4	0	1.11
VAH10	20	79.63	81.1	-1.47	0.41	0	1.24
VAH10	21	79.05	80.5	-1.45	0.41	0	1.21
VAH10	22	78.68	80.14	-1.46	0.4	0	1.23
VAH10	23	78.38	79.74	-1.37	0.39	0	1.14

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH10	24	76.81	78.35	-1.53	0.35	0	1.34
VAH10	25	76.49	78.25	-1.76	0.34	0	1.56
VAH10	26	76.03	77.79	-1.76	0.33	0	1.57
VAH10	27	75.38	77.42	-2.03	0.33	0	1.84
VAH10	28	74.87	77.14	-2.27	0.33	0	2.08
VAH11	1	100	100	0	2.59	1	0.61
VAH11	2	97.6	100	-2.4	2.59	1	1.43
VAH11	3	95.98	100	-4.02	2.27	0	2.81
VAH11	4	94.65	100	-5.35	1.94	0	4.26
VAH11	5	93.25	98.04	-4.79	1.58	0	3.9
VAH11	6	92.1	97.05	-4.94	1.24	0	4.25
VAH11	7	91.52	94.65	-3.13	0.91	0	2.62
VAH11	8	90.82	93.65	-2.84	0.67	0	2.46
VAH11	9	90.01	92.39	-2.38	0.49	0	2.1
VAH11	10	89.26	91.25	-1.98	0.4	0	1.76
VAH11	11	88.31	91.07	-2.77	0.38	0	2.55
VAH11	12	88.18	89.48	-1.3	0.35	0	1.1
VAH11	13	87.53	89.07	-1.55	0.33	0	1.36
VAH11	14	87.02	88.48	-1.45	0.31	0	1.28
VAH11	15	86.48	88.04	-1.57	0.3	0	1.39
VAH11	16	85.79	88.45	-2.66	0.3	0	2.49
VAH11	17	83.71	86.61	-2.9	0.27	0	2.75
VAH11	18	83.23	84.85	-1.62	0.26	0	1.47
VAH11	19	82.72	84.35	-1.63	0.27	0	1.48
VAH11	20	82.2	83.99	-1.79	0.26	0	1.64
VAH11	21	81.86	83.88	-2.02	0.27	0	1.87
VAH11	22	81.43	83.12	-1.68	0.26	0	1.54
VAH11	23	80.94	82.91	-1.97	0.27	0	1.82
VAH11	24	80.58	82.7	-2.13	0.26	0	1.98
VAH11	25	80.05	82.67	-2.62	0.29	0	2.46
VAH11	26	79.57	82.34	-2.76	0.33	0	2.57
VAH11	27	79.19	81.97	-2.79	0.31	0	2.61
VAH11	28	78.55	82.67	-4.12	0.41	0	3.89
VAH11	29	78.06	81.67	-3.61	0.45	0	3.36
VAH11	30	77.7	81.28	-3.58	0.45	0	3.33
VAH11	31	77.4	81.43	-4.03	0.42	0	3.79
VAH11	32	76.84	80.84	-4.01	0.45	0	3.75
VAH11	33	76.5	80.82	-4.32	0.46	0	4.06
VAH11	34	75.94	80.08	-4.14	0.46	0	3.88
VAH11	35	75.62	79.36	-3.75	0.47	0	3.48
VAH11	36	75.23	79.06	-3.83	0.48	0	3.57
VAH11	37	74.74	79.68	-4.95	0.51	0	4.66
VAH11	38	74.21	79.14	-4.93	0.53	0	4.63
VAH11	39	73.86	77.99	-4.12	0.5	0	3.84
VAH11	40	73.01	76.23	-3.22	0.52	0	2.92
VAH11	41	72.54	75.65	-3.11	0.52	0	2.81
VAH11	42	72.21	75.8	-3.6	0.5	0	3.32
VAH11	43	71.82	75.38	-3.56	0.48	0	3.29

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH11	44	70.69	73.72	-3.03	0.53	0	2.73
VAH12	1	100	100	0	2.42	1	0.57
VAH12	2	97.45	100	-2.55	1.96	0	1.62
VAH12	3	95.71	97.89	-2.18	1.05	0	1.6
VAH12	4	94.3	97.33	-3.03	1.07	0	2.43
VAH12	5	92.89	95.26	-2.37	0.84	0	1.9
VAH12	6	91.51	94.49	-2.98	0.73	0	2.57
VAH12	7	91.03	93.12	-2.08	0.67	0	1.7
VAH12	8	90.01	92.29	-2.28	0.61	0	1.94
VAH12	9	89.28	91.61	-2.33	0.42	0	2.09
VAH12	10	88.53	91.17	-2.65	0.42	0	2.41
VAH12	11	87.17	89.72	-2.55	0.37	0	2.35
VAH12	12	87.06	89.22	-2.16	0.39	0	1.94
VAH12	13	86.46	88.64	-2.18	0.34	0	1.99
VAH12	14	85.89	88.2	-2.31	0.31	0	2.14
VAH12	15	85.34	87.54	-2.21	0.28	0	2.05
VAH12	16	84.4	87.57	-3.17	0.21	0	3.05
VAH12	17	83.99	85.63	-1.64	0.25	0	1.5
VAH12	18	82.31	83.98	-1.67	0.13	0	1.59
VAH12	19	81.79	83.4	-1.61	0.12	0	1.54
VAH12	20	81.28	82.89	-1.6	0.11	0	1.54
VAH12	21	80.89	82.52	-1.63	0.1	0	1.57
VAH12	22	80.45	82.3	-1.85	0.1	0	1.8
VAH12	23	80.05	82.22	-2.18	0.1	0	2.12
VAH12	24	78.91	80.5	-1.59	0.1	0	1.53
VAH12	25	78.31	80.04	-1.73	0.11	0	1.67
VAH12	26	78.81	81.31	-2.5	0.09	0	2.45
VAH12	27	78.17	80.89	-2.73	0.1	0	2.67
VAH12	28	78.34	81.02	-2.68	0.09	0	2.63
VAH12	29	77.84	80.79	-2.95	0.09	0	2.9
VAH12	30	75.89	79.32	-3.43	0.09	0	3.38
VAH12	31	76.83	80.02	-3.19	0.09	0	3.14
VAH12	32	76.51	80.11	-3.6	0.09	0	3.55
VAH12	33	76.36	80.36	-4	0.09	0	3.95
VAH12	34	75.35	79.16	-3.81	0.09	0	3.76
VAH12	35	75.03	79.07	-4.05	0.09	0	4
VAH12	36	75.17	79.42	-4.25	0.09	0	4.2
VAH12	37	74.64	78.3	-3.66	0.11	0	3.59
VAH12	38	72.95	78.46	-5.51	0.1	0	5.45
VAH12	39	74.07	78.86	-4.8	0.09	0	4.74
VAH12	40	73.69	78.74	-5.05	0.09	0	5
VAH12	41	73.31	78.71	-5.4	0.09	0	5.35
VAH12	42	73.03	78.5	-5.47	0.09	0	5.42
VAH12	43	72.65	76.51	-3.86	0.11	0	3.8
VAH12	44	72.21	78.13	-5.92	0.1	0	5.87
VAH12	45	71.61	77.33	-5.72	0.1	0	5.67
VAH12	46	71.89	76.82	-4.92	0.11	0	4.86
VAH13	1	100	100	0	1.48	1	0.35

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH13	2	96.77	98.3	-1.53	1.05	0	1
VAH13	3	95.3	96.5	-1.2	0.77	0	0.8
VAH13	4	93.84	94.47	-0.63	0.57	0	0.38
VAH13	5	92.52	92.87	-0.35	0.5	1	0.21
VAH13	6	91.17	91.51	-0.33	0.45	1	0.2
VAH13	7	90.42	90.46	-0.04	0.43	1	0.1
VAH13	8	89.43	89.74	-0.31	0.43	1	0.19
VAH13	9	88.45	88.92	-0.46	0.43	0	0.28
VAH13	10	87.64	88.03	-0.39	0.42	1	0.24
VAH13	11	86.5	86.9	-0.4	0.43	1	0.24
VAH13	12	86.51	86.3	0.2	0.41	1	0.14
VAH13	13	85.68	85.25	0.43	0.36	0	0.27
VAH13	14	84.49	84.33	0.16	0.45	1	0.13
VAH13	15	84.2	83.66	0.54	0.35	0	0.36
VAH13	16	83.37	83.28	0.09	0.42	1	0.11
VAH13	17	79.75	79.57	0.17	0.39	1	0.12
VAH13	18	77.72	77.16	0.57	0.51	0	0.35
VAH13	19	78.17	77.52	0.65	0.37	0	0.45
VAH13	20	75.76	74.54	1.23	0.56	0	0.92
VAH15	1	100	100	0	0.57	1	0.13
VAH15	2	96.59	97.78	-1.19	0.59	0	0.87
VAH15	3	94.27	94.82	-0.55	0.59	1	0.33
VAH15	4	92.37	92.81	-0.44	0.58	1	0.26
VAH15	5	90.87	91.11	-0.24	0.52	1	0.17
VAH15	6	89.41	89.89	-0.48	0.47	0	0.29
VAH15	7	87.95	88.12	-0.17	0.46	1	0.13
VAH15	8	86.9	86.98	-0.08	0.49	1	0.12
VAH15	9	85.79	86.08	-0.29	0.5	1	0.18
VAH15	10	84.76	85.23	-0.47	0.51	1	0.28
VAH15	11	83.59	84.54	-0.94	0.51	0	0.67
VAH16	1	100	99.2	0.8	0.5	0	0.54
VAH16	2	96.86	97.39	-0.53	0.44	0	0.33
VAH16	3	94.8	94.94	-0.14	0.38	1	0.11
VAH16	4	93.01	93.03	-0.02	0.34	1	0.08
VAH16	5	91.24	91.4	-0.16	0.36	1	0.11
VAH16	6	89.83	89.8	0.03	0.3	1	0.07
VAH16	7	88.65	88.16	0.5	0.32	0	0.33
VAH16	8	87.54	87.03	0.51	0.34	0	0.34
VAH16	9	86.36	85.76	0.61	0.38	0	0.41
VAH16	10	84.92	84.1	0.82	0.45	0	0.58
VAH16	11	84.02	83.41	0.61	0.5	0	0.38
VAH17	1	100	98.4	1.6	0.65	0	1.23
VAH17	2	96.95	96.72	0.23	0.56	1	0.17
VAH17	3	94.74	95.11	-0.36	0.47	1	0.22
VAH17	4	93.06	93.13	-0.07	0.41	1	0.1
VAH17	5	91.74	91.59	0.15	0.37	1	0.11
VAH17	6	90.31	90.04	0.27	0.38	1	0.16
VAH17	7	89.19	89.11	0.08	0.38	1	0.1

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH17	8	88.16	88.2	-0.04	0.38	1	0.09
VAH17	9	87.18	87.48	-0.3	0.37	1	0.18
VAH17	10	86.31	86.76	-0.46	0.38	0	0.29
VAH17	11	85.34	86.01	-0.67	0.39	0	0.46
VAH17	12	84.68	84.99	-0.31	0.42	1	0.19
VAH17	13	83.84	84.45	-0.62	0.42	0	0.4
VAH17	14	83.07	83.65	-0.58	0.45	0	0.37
VAH17	15	82.34	83.04	-0.71	0.46	0	0.47
VAH17	16	81.49	82.98	-1.48	0.41	0	1.25
VAH17	17	80.19	81.53	-1.34	0.45	0	1.09
VAH17	18	80.04	81.05	-1.01	0.51	0	0.73
VAH17	19	79.25	80.59	-1.34	0.52	0	1.05
VAH17	20	77.93	79.62	-1.69	0.56	0	1.37
VAH20	1	100	99.12	0.88	0.96	1	0.52
VAH20	2	96.72	96.97	-0.25	0.87	1	0.23
VAH20	3	94.4	95.27	-0.87	0.8	0	0.53
VAH20	4	92.61	92.74	-0.13	0.71	1	0.18
VAH20	5	90.89	91.22	-0.32	0.66	1	0.22
VAH20	6	89.56	89.86	-0.31	0.63	1	0.2
VAH20	7	88.22	88.86	-0.63	0.59	0	0.39
VAH20	8	87.02	87.97	-0.96	0.55	0	0.66
VAH20	9	85.23	86.71	-1.49	0.58	0	1.16
VAH20	10	84.08	85.13	-1.05	0.58	0	0.74
VAH20	11	83.3	83.45	-0.15	0.93	1	0.23
VAH20	12	82.15	81.97	0.18	1.06	1	0.26
VAH22	1	100	99.27	0.73	0.48	0	0.49
VAH22	2	96.58	96.28	0.3	0.35	1	0.18
VAH22	3	94.35	93.45	0.9	0.28	0	0.74
VAH22	4	92.45	92.15	0.3	0.32	1	0.18
VAH22	5	90.82	90.33	0.49	0.32	0	0.32
VAH22	6	89.37	89.11	0.26	0.32	1	0.16
VAH22	7	88.02	87.84	0.18	0.36	1	0.12
VAH22	8	86.97	87.13	-0.15	0.34	1	0.11
VAH22	9	85.63	85.45	0.18	0.38	1	0.12
VAH22	10	84.59	84.37	0.22	0.49	1	0.15
VAH23	1	100	100	0	1.62	1	0.38
VAH23	2	95.61	100	-4.39	1.62	0	3.47
VAH23	3	93.73	100	-6.27	1.57	0	5.39
VAH23	4	91.42	100	-8.58	1.73	0	7.61
VAH23	5	91.25	100	-8.75	1.45	0	7.93
VAH23	6	89.82	100	-10.18	1.45	0	9.36
VAH23	7	88.83	96.87	-8.04	1.24	0	7.34
VAH23	8	88.2	99.71	-11.51	1.45	0	10.69
VAH23	9	85.73	99.38	-13.65	1.82	0	12.63
VAH23	10	85.95	98.9	-12.95	1.57	0	12.06
VAH23	11	85.07	97.97	-12.9	1.55	0	12.03
VAH23	12	84.57	94.24	-9.67	1.14	0	9.03
VAH23	13	83.91	96.55	-12.64	1.42	0	11.84

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH23	14	82.86	92.87	-10	1.17	0	9.34
VAH24	1	100	98.18	1.82	0.33	0	1.63
VAH24	2	96.99	96.64	0.36	0.3	0	0.22
VAH24	3	94.54	93.79	0.75	0.29	0	0.59
VAH24	4	93.41	93.07	0.34	0.26	0	0.21
VAH24	5	91.95	91.92	0.03	0.25	1	0.06
VAH24	6	90.07	90.18	-0.11	0.28	1	0.08
VAH24	7	89.62	89.56	0.06	0.25	1	0.06
VAH24	8	88.51	88.4	0.11	0.25	1	0.08
VAH24	9	87.51	87.73	-0.22	0.25	1	0.13
VAH24	10	86.72	86.85	-0.13	0.23	1	0.08
VAH24	11	85.58	85.78	-0.2	0.23	1	0.12
VAH24	12	84.58	84.12	0.46	0.31	0	0.3
VAH24	13	83.65	83.24	0.42	0.31	0	0.27
VAH24	14	83.2	82.95	0.25	0.3	1	0.15
VAH24	15	82.18	81.55	0.63	0.36	0	0.44
VAH24	16	80.33	79.8	0.53	0.43	0	0.33
VAH25	1	100	100	0	0.32	1	0.08
VAH25	2	96.82	97.53	-0.71	0.31	0	0.54
VAH25	3	94.7	94.83	-0.13	0.27	1	0.09
VAH25	4	92.95	92.41	0.54	0.23	0	0.41
VAH25	5	91.94	91.07	0.87	0.26	0	0.73
VAH25	6	90.23	89.21	1.02	0.28	0	0.86
VAH25	7	88.76	88.06	0.7	0.3	0	0.53
VAH25	8	87.45	86.5	0.95	0.38	0	0.74
VAH25	9	86.21	85.32	0.89	0.44	0	0.65
VAH25	10	85.39	84.16	1.23	0.49	0	0.95
VAH25	11	83.6	82.74	0.86	0.56	0	0.57
VAH26	1	100	96.93	3.07	0.73	0	2.66
VAH26	2	97.27	97.81	-0.54	0.75	1	0.32
VAH26	3	95.31	95.18	0.13	0.64	1	0.16
VAH26	4	93.73	93.64	0.1	0.58	1	0.14
VAH26	5	92.34	92.38	-0.04	0.54	1	0.13
VAH26	6	90.3	90.01	0.29	0.47	1	0.18
VAH26	7	89.37	89.54	-0.17	0.46	1	0.13
VAH26	8	88.51	88.82	-0.31	0.45	1	0.19
VAH26	9	87.38	87.6	-0.22	0.47	1	0.15
VAH26	10	86.25	86.38	-0.13	0.45	1	0.12
VAH26	11	84.62	84.19	0.43	0.52	1	0.26
VAH26	12	84.5	84.33	0.17	0.43	1	0.13
VAH26	13	83.91	84.25	-0.34	0.48	1	0.21
VAH26	14	82.55	82.39	0.16	0.54	1	0.14
VAH26	15	81.97	81.36	0.61	0.48	0	0.39
VAH26	16	81.38	81.03	0.35	0.48	1	0.21
VAH26	17	80.22	80.11	0.11	0.57	1	0.14
VAH26	18	79.5	79.2	0.3	0.53	1	0.19
VAH26	19	78.82	79.72	-0.89	0.58	0	0.6
VAH26	20	78.39	78.65	-0.26	0.61	1	0.19

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH26	21	77.66	77.37	0.29	0.52	1	0.18
VAH27	1	100	100	0	0.6	1	0.14
VAH27	2	97.01	97.26	-0.25	0.47	1	0.16
VAH27	3	94.91	95.43	-0.52	0.45	0	0.32
VAH27	4	93.07	92.94	0.13	0.43	1	0.12
VAH27	5	91.1	91.88	-0.78	0.45	0	0.54
VAH27	6	88.95	87.68	1.27	0.39	0	1.05
VAH27	7	87.71	87.47	0.24	0.43	1	0.15
VAH27	8	86.53	86.12	0.41	0.44	1	0.24
VAH27	9	85.51	84.27	1.23	0.41	0	1
VAH27	10	83.79	82.61	1.18	0.57	0	0.86
VAH27	11	82.76	81.65	1.11	0.57	0	0.8
VAH28	1	100	97.95	2.05	0.56	0	1.73
VAH28	2	97.27	98.02	-0.75	0.53	0	0.49
VAH28	3	95.47	95.95	-0.48	0.43	0	0.29
VAH28	4	93.93	94.2	-0.27	0.35	1	0.16
VAH28	5	92.68	93.02	-0.33	0.33	1	0.2
VAH28	6	92.08	92.04	0.04	0.32	1	0.08
VAH28	7	90.48	90.65	-0.17	0.31	1	0.11
VAH28	8	89.43	89.5	-0.06	0.33	1	0.08
VAH28	9	88.61	88.86	-0.25	0.32	1	0.15
VAH28	10	87.92	88.26	-0.34	0.31	0	0.21
VAH28	11	86.99	87.1	-0.11	0.32	1	0.09
VAH28	12	86.42	86.27	0.15	0.32	1	0.1
VAH28	13	85.78	85.79	-0.01	0.32	1	0.07
VAH28	14	85.09	84.86	0.22	0.34	1	0.13
VAH28	15	84.36	83.88	0.48	0.36	0	0.31
VAH28	16	83.51	83.93	-0.42	0.37	0	0.26
VAH28	17	83.29	82.45	0.84	0.4	0	0.62
VAH28	18	82.75	82.14	0.61	0.39	0	0.41
VAH28	19	82.13	82.13	0	0.39	1	0.09
VAH28	20	81.78	81.96	-0.18	0.37	1	0.12
VAH28	21	81.12	81	0.12	0.42	1	0.11
VAH28	22	80.48	81.25	-0.77	0.39	0	0.56
VAH28	23	80.56	81.17	-0.61	0.36	0	0.42
VAH30	1	100	98.43	1.57	0.42	0	1.33
VAH30	2	96.89	96.36	0.52	0.37	0	0.34
VAH30	3	94.67	94.16	0.51	0.33	0	0.34
VAH30	4	93.05	92.43	0.61	0.33	0	0.43
VAH30	5	91.4	91.05	0.35	0.33	0	0.21
VAH30	6	89.81	89.66	0.15	0.34	1	0.11
VAH30	7	88.96	88.16	0.8	0.34	0	0.61
VAH30	8	87.59	87.08	0.51	0.35	0	0.34
VAH30	9	86.78	86.35	0.43	0.36	0	0.27
VAH30	10	85.77	85.3	0.46	0.36	0	0.29
VAH30	11	84.53	84.62	-0.09	0.38	1	0.1
VAH30	12	84.13	83.59	0.54	0.4	0	0.35
VAH30	13	83.13	82.49	0.64	0.42	0	0.43

Continued on next page

Table 21 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH30	14	82.14	81.63	0.51	0.43	0	0.32
VAH30	15	81.52	80.7	0.82	0.44	0	0.58
VAH30	16	80.73	79.96	0.77	0.44	0	0.54
VAH30	17	79.91	79.29	0.62	0.44	0	0.41
VAH30	18	79.63	78.49	1.14	0.44	0	0.9

Table 21: Mixture Density Network - SOH Capacity Test Results

H Mixture Density Network - RUL Capacity Test Results

Note that all the results are rounded to 2. Mean indicates distribution mean prediction, SD is the standard deviation and CRPS Norm is the CRPS for normal distribution. Sharpness is 100 percent if the actual value falls within one standard deviation of distributional mean of that capacity test indicated as 1 , 0 otherwise.

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH01	1	612	561.26	50.74	11.26	0	44.38
VAH01	2	561	501.54	59.46	12.35	0	52.49
VAH01	3	510	441.29	68.71	12.21	0	61.82
VAH01	4	459	377.2	81.8	10.34	0	75.97
VAH01	5	408	320.03	87.97	8.81	0	83
VAH01	6	357	273.14	83.86	8.85	0	78.86
VAH01	7	306	241.12	64.88	9.51	0	59.51
VAH01	8	255	111.54	143.46	22.52	0	130.75
VAH01	9	204	53.58	150.42	25.48	0	136.04
VAH01	10	153	54.43	98.57	26.08	0	83.86
VAH01	11	102	0	102	22.72	0	89.18
VAH01	12	51	0	51	21.36	0	39.07
VAH01	13	0	0	0	18.73	1	4.38
VAH02	1	510	535.78	-25.78	2.47	0	24.38
VAH02	2	459	459.52	-0.52	3.04	1	0.74
VAH02	3	408	401.48	6.52	3.83	0	4.5
VAH02	4	357	346.3	10.7	5.63	0	7.64
VAH02	5	306	298.09	7.91	7.46	0	4.81
VAH02	6	255	249.96	5.04	9.65	1	3.28
VAH02	7	204	205.29	-1.29	11.23	1	2.68
VAH02	8	153	111.45	41.55	22.95	0	29.24
VAH02	9	102	99.04	2.96	24.14	1	5.79
VAH02	10	51	13.28	37.72	17.07	0	28.25
VAH02	11	0	0	0	8.65	1	2.02
VAH05	1	765	726.6	38.4	16.44	0	29.23
VAH05	2	714	671.7	42.3	16.25	0	33.18
VAH05	3	663	623.4	39.6	17.23	0	30.01
VAH05	4	612	567.51	44.49	18.06	0	34.38
VAH05	5	561	509.07	51.93	18.41	0	41.57
VAH05	6	510	473.47	36.53	18.99	0	26.21

Continued on next page

Table 22 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH05	7	459	420.66	38.34	19.03	0	27.92
VAH05	8	408	375.76	32.24	19.32	0	22.1
VAH05	9	357	334.33	22.67	19.28	0	14.06
VAH05	10	306	283.64	22.36	19.27	0	13.83
VAH05	11	255	234.03	20.97	19.01	0	12.84
VAH05	12	204	170.85	33.15	18.68	0	23.18
VAH05	13	153	134.39	18.61	18.67	1	11.21
VAH05	14	102	99.5	2.5	18.53	1	4.46
VAH05	15	51	0.03	50.97	17.78	0	40.97
VAH05	16	0	0.03	-0.03	19.43	1	4.54
VAH06	1	612	610.02	1.98	14.49	1	3.49
VAH06	2	561	554.66	6.34	13.25	1	4.28
VAH06	3	510	470.8	39.2	12.58	0	32.11
VAH06	4	459	448.99	10.01	14.55	1	6.05
VAH06	5	408	384.74	23.26	14.36	0	15.8
VAH06	6	357	340.33	16.67	14.86	0	10.24
VAH06	7	306	286.11	19.89	16.68	0	12.38
VAH06	8	255	207.35	47.65	16.24	0	38.51
VAH06	9	204	185.88	18.12	18.19	1	10.91
VAH06	10	153	102.02	50.98	18.98	0	40.31
VAH06	11	102	110.74	-8.74	21.19	1	6.37
VAH06	12	51	31.35	19.65	24.1	1	11.69
VAH06	13	0	36	-36	26.14	0	23.27
VAH09	1	516	519.62	-3.62	16.44	1	4.16
VAH09	2	465	461.68	3.32	15.85	1	3.98
VAH09	3	414	415.87	-1.87	16.35	1	3.91
VAH09	4	363	372.92	-9.92	16.87	1	6.2
VAH09	5	351	359.01	-8.01	17	1	5.45
VAH09	6	339	347.12	-8.12	16.93	1	5.48
VAH09	7	327	335.55	-8.55	17.19	1	5.68
VAH09	8	315	306.45	8.55	17.85	1	5.77
VAH09	9	303	296.79	6.21	17.81	1	5.02
VAH09	10	291	299.47	-8.47	17.8	1	5.74
VAH09	11	279	288.4	-9.4	17.97	1	6.12
VAH09	12	204	208.39	-4.39	19.82	1	5.02
VAH09	13	153	203.63	-50.63	20.99	0	38.89
VAH09	14	102	150.28	-48.28	22.28	0	35.95
VAH09	15	51	78.93	-27.93	22.98	0	17.46
VAH09	16	0	67.29	-67.29	23.85	0	53.87
VAH10	1	612	539.71	72.29	7.91	0	67.83
VAH10	2	561	490.13	70.87	8.18	0	66.26
VAH10	3	510	439.33	70.67	9.14	0	65.51
VAH10	4	459	375.74	83.26	10.06	0	77.58
VAH10	5	408	331.54	76.46	10.95	0	70.28
VAH10	6	357	288.42	68.58	11.84	0	61.9
VAH10	7	306	269.22	36.78	13.39	0	29.25
VAH10	8	255	189.8	65.2	13.97	0	57.31
VAH10	9	204	143.83	60.17	15.08	0	51.66

Continued on next page

Table 22 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH10	10	153	95.8	57.2	16.18	0	48.07
VAH10	11	102	41.69	60.31	17.44	0	50.47
VAH10	12	51	11.44	39.56	18.07	0	29.55
VAH10	13	0	0	0	18.79	1	4.39
VAH11	1	815	718.54	96.46	0.13	0	96.39
VAH11	2	764	709.05	54.95	0.19	0	54.84
VAH11	3	713	700.24	12.76	0.73	0	12.35
VAH11	4	662	695.82	-33.82	1.88	0	32.76
VAH11	5	611	662.1	-51.1	2.81	0	49.51
VAH11	6	560	669.95	-109.95	4.49	0	107.41
VAH11	7	509	650.26	-141.26	5.61	0	138.09
VAH11	8	458	653.94	-195.94	6.95	0	192.02
VAH11	9	407	651.46	-244.46	8.16	0	239.85
VAH11	10	356	621.53	-265.53	6.47	0	261.88
VAH11	11	305	587.35	-282.35	4.65	0	279.73
VAH11	12	254	565.32	-311.32	2.23	0	310.06
VAH11	13	203	531.43	-328.43	0.35	0	328.23
VAH11	14	152	518.21	-366.21	0.06	0	366.18
VAH11	15	101	494.01	-393.01	0.01	0	393.01
VAH11	16	50	486.01	-436.01	0.01	0	436
VAH11	17	0	347.81	-347.81	0	0	347.81
VAH12	1	765	586.06	178.94	32.12	0	160.82
VAH12	2	714	533.02	180.98	31.48	0	163.22
VAH12	3	663	438.55	224.45	30.64	0	207.16
VAH12	4	612	423.28	188.72	29.18	0	172.26
VAH12	5	561	316.29	244.71	29.64	0	227.99
VAH12	6	510	315.75	194.25	28.1	0	178.4
VAH12	7	459	255.85	203.15	27.75	0	187.49
VAH12	8	408	209.81	198.19	27.02	0	182.94
VAH12	9	357	139.26	217.74	27.91	0	202
VAH12	10	306	144.65	161.35	26.8	0	146.23
VAH12	11	255	80.48	174.52	26.82	0	159.39
VAH12	12	204	24.7	179.3	26.69	0	164.24
VAH12	13	153	0.01	152.99	26.37	0	138.11
VAH12	14	102	0	102	26.04	0	87.31
VAH12	15	51	0	51	25.66	0	36.97
VAH12	16	0	0	0	24.18	1	5.65
VAH13	1	663	662.35	0.65	19.5	1	4.57
VAH13	2	612	581.52	30.48	20.21	0	20.24
VAH13	3	561	550.26	10.74	21.25	1	7.09
VAH13	4	510	495.64	14.36	21.94	1	8.75
VAH13	5	459	440.98	18.02	21.91	1	10.72
VAH13	6	408	379.61	28.39	21.52	0	18.13
VAH13	7	357	324.99	32.01	19.19	0	21.94
VAH13	8	306	283.48	22.52	20.2	0	13.82
VAH13	9	255	248.31	6.69	21.64	1	5.88
VAH13	10	204	207.37	-3.37	22.69	1	5.5
VAH13	11	153	123.2	29.8	23.39	0	18.85

Continued on next page

Table 22 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH13	12	102	82.55	19.45	24.39	1	11.58
VAH13	13	51	7.73	43.27	26.28	0	29.54
VAH13	14	0	0.06	-0.06	25.69	1	6
VAH15	1	459	458.28	0.72	10.92	1	2.57
VAH15	2	408	411.84	-3.84	9.81	1	2.89
VAH15	3	357	350.42	6.58	9.67	1	3.98
VAH15	4	306	301.09	4.91	10.31	1	3.33
VAH15	5	255	251.56	3.44	13.02	1	3.4
VAH15	6	204	205.9	-1.9	16.11	1	3.85
VAH15	7	153	147.13	5.87	19.28	1	5.21
VAH15	8	102	112.35	-10.35	21.31	1	6.95
VAH15	9	51	63.12	-12.12	23.34	1	7.91
VAH15	10	0	15.36	-15.36	24.54	1	9.45
VAH16	1	459	503.12	-44.12	28.53	0	29.52
VAH16	2	408	460.8	-52.8	27.56	0	37.84
VAH16	3	357	412.41	-55.41	27.24	0	40.46
VAH16	4	306	351.23	-45.23	26.62	0	31.19
VAH16	5	255	304.86	-49.86	27.25	0	35.21
VAH16	6	204	237.1	-33.1	28.24	0	20.52
VAH16	7	153	201.3	-48.3	30.54	0	32.56
VAH16	8	102	136.28	-34.28	33.64	0	20.71
VAH16	9	51	76.2	-25.2	36.56	1	15.21
VAH16	10	0	27.61	-27.61	37.67	1	16.53
VAH17	1	561	533.9	27.1	10.5	0	21.21
VAH17	2	510	476.65	33.35	9.1	0	28.22
VAH17	3	459	429.97	29.03	9.94	0	23.44
VAH17	4	408	382.58	25.42	10.89	0	19.35
VAH17	5	357	347.01	9.99	12.54	1	5.95
VAH17	6	306	298.3	7.7	14.47	1	4.98
VAH17	7	255	263.43	-8.43	16.67	1	5.56
VAH17	8	204	229.45	-25.45	19.2	0	16.28
VAH17	9	153	198.7	-45.7	21.44	0	33.86
VAH17	10	102	155.13	-53.13	24	0	39.81
VAH17	11	51	108.81	-57.81	26.21	0	43.28
VAH17	12	0	64.1	-64.1	28.74	0	48.15
VAH20	1	459	467.04	-8.04	25.36	1	6.94
VAH20	2	408	432.98	-24.98	24.68	0	15.08
VAH20	3	357	369.34	-12.34	24.93	1	8.21
VAH20	4	306	315.71	-9.71	26.52	1	7.6
VAH20	5	255	258.97	-3.97	27.66	1	6.69
VAH20	6	204	220.98	-16.98	29.36	1	10.67
VAH20	7	153	162.32	-9.32	31.21	1	8.39
VAH20	8	102	119.1	-17.1	32.94	1	11.16
VAH20	9	51	69.84	-18.84	34.91	1	12.12
VAH20	10	0	34.38	-34.38	36.18	1	20.59
VAH22	1	459	469.93	-10.93	7.84	0	7.09
VAH22	2	408	404.82	3.18	5.36	1	1.99
VAH22	3	357	343.22	13.78	4.59	0	11.19

Continued on next page

Table 22 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH22	4	306	299.5	6.5	7.27	1	3.88
VAH22	5	255	251.53	3.47	11.25	1	3.05
VAH22	6	204	208.72	-4.72	15.1	1	4.11
VAH22	7	153	152.65	0.35	21.12	1	4.94
VAH22	8	102	116.92	-14.92	23.67	1	9.16
VAH22	9	51	50.94	0.06	27.98	1	6.54
VAH22	10	0	0.08	-0.08	24.98	1	5.84
VAH23	1	561	267.72	293.28	6.34	0	289.71
VAH23	2	510	224.89	285.11	5.59	0	281.95
VAH23	3	459	204.49	254.51	6.8	0	250.67
VAH23	4	408	160.74	247.26	11.72	0	240.65
VAH23	5	357	164.6	192.4	10	0	186.75
VAH23	6	306	127.73	178.27	14.91	0	169.86
VAH23	7	255	103.68	151.32	18.41	0	140.93
VAH23	8	204	85.98	118.02	20.86	0	106.25
VAH23	9	153	46.3	106.7	23.59	0	93.39
VAH23	10	102	41.74	60.26	25.09	0	46.25
VAH23	11	51	20.78	30.22	23.65	0	19.13
VAH23	12	0	13.88	-13.88	19.39	1	8.34
VAH24	1	561	572.17	-11.17	13.49	1	6.65
VAH24	2	510	552.04	-42.04	13.59	0	34.38
VAH24	3	459	453.37	5.63	10.97	1	3.69
VAH24	4	408	435.92	-27.92	13.25	0	20.61
VAH24	5	357	383.06	-26.06	13.72	0	18.62
VAH24	6	306	312.35	-6.35	13.55	1	4.33
VAH24	7	255	272	-17	18.23	1	10.16
VAH24	8	204	217.22	-13.22	20.14	1	8.05
VAH24	9	153	175	-22	22.08	1	13.25
VAH24	10	102	153.86	-51.86	22.41	0	39.37
VAH24	11	51	37.76	13.24	23.42	1	8.38
VAH24	12	0	8.57	-8.57	22.09	1	6.47
VAH25	1	512	538.98	-26.98	17.96	0	17.89
VAH25	2	461	481.6	-20.6	16.85	0	12.9
VAH25	3	410	422.55	-12.55	17.61	1	7.54
VAH25	4	359	346.75	12.25	18.15	1	7.42
VAH25	5	307	320.09	-13.09	19.56	1	7.94
VAH25	6	256	262.09	-6.09	20.75	1	5.56
VAH25	7	205	167.49	37.51	21.36	0	26.14
VAH25	8	154	117.29	36.71	23.03	0	24.8
VAH25	9	103	33.5	69.5	24.22	0	55.86
VAH25	10	51	0	51	18.14	0	40.79
VAH25	11	0	0	0	11.57	1	2.7
VAH26	1	612	626.01	-14.01	25.34	1	8.93
VAH26	2	561	621.04	-60.04	25.88	0	45.62
VAH26	3	510	545.54	-35.54	26.01	0	22.92
VAH26	4	459	484.71	-25.71	26.12	1	15.46
VAH26	5	408	434.97	-26.97	26.4	0	16.3
VAH26	6	306	307.28	-1.28	27.32	1	6.41

Continued on next page

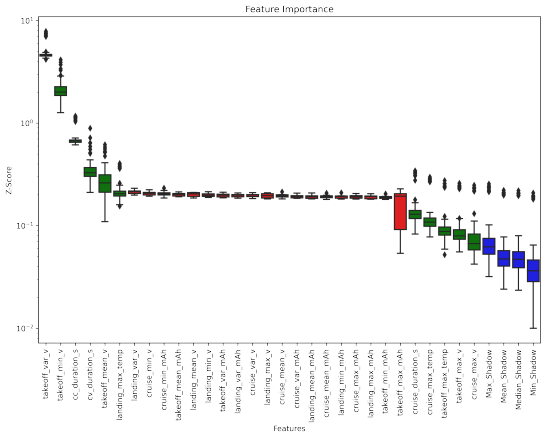
Table 22 – *Continued from previous page*

Vahana	Capacity Test	Actual	Mean	Error	SD	Sh(%)	CRPS Norm
VAH26	7	255	279.77	-24.77	27.75	1	14.76
VAH26	8	204	239.91	-35.91	28.04	0	22.75
VAH26	9	153	200.01	-47.01	29.06	0	31.91
VAH26	10	51	89.26	-38.26	29.64	0	24.29
VAH26	11	0	36.15	-36.15	34.03	0	21.98
VAH27	1	510	543.42	-33.42	12.13	0	26.6
VAH27	2	459	500.89	-41.89	12.8	0	34.67
VAH27	3	408	445.79	-37.79	12.61	0	30.68
VAH27	4	357	384.65	-27.65	11.82	0	21.06
VAH27	5	306	337.22	-31.22	11.35	0	24.84
VAH27	6	204	163.09	40.91	13.65	0	33.22
VAH27	7	153	130.43	22.57	20.67	0	13.79
VAH27	8	102	71.5	30.5	25.54	0	18.99
VAH27	9	51	11.19	39.81	25.36	0	26.77
VAH27	10	0	0	0	31.88	1	7.45
VAH28	1	720	716	4	37.83	1	9.01
VAH28	2	669	711.82	-42.82	33.83	0	27.05
VAH28	3	618	659.7	-41.7	34.59	0	26.02
VAH28	4	567	598.22	-31.22	34.11	1	18.64
VAH28	5	516	550.29	-34.29	35.7	1	20.56
VAH28	6	464	469.39	-5.39	34.05	1	8.3
VAH28	7	413	409.9	3.1	35.68	1	8.45
VAH28	8	362	344.63	17.37	33.49	1	11.34
VAH28	9	311	300.29	10.71	33.8	1	9.24
VAH28	10	260	277.95	-17.95	34.7	1	11.73
VAH28	11	209	189.38	19.62	30.78	1	12.02
VAH28	12	158	155.88	2.12	30.64	1	7.22
VAH28	13	107	128.01	-21.01	30.64	1	12.69
VAH28	14	51	75.15	-24.15	30.21	1	14.37
VAH28	15	0	8.45	-8.45	29.66	1	7.89
VAH30	1	510	635.96	-125.96	13.98	0	118.07
VAH30	2	459	583.04	-124.04	13.1	0	116.65
VAH30	3	408	524.31	-116.31	12.5	0	109.26
VAH30	4	357	465.81	-108.81	12.6	0	101.69
VAH30	5	306	412.1	-106.1	13.02	0	98.75
VAH30	6	255	351.72	-96.72	14.01	0	88.81
VAH30	7	204	287.02	-83.02	16.85	0	73.52
VAH30	8	153	226.29	-73.29	19.82	0	62.11
VAH30	9	102	193.51	-91.51	21.77	0	79.23
VAH30	10	51	129.37	-78.37	24.89	0	64.34
VAH30	11	0	86.95	-86.95	27.12	0	71.66

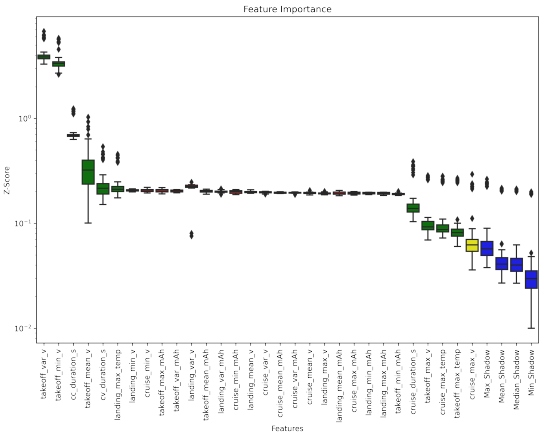
Table 22: Mixture Density Network - RUL Capacity Test Results

I SOH Feature Importance Plots

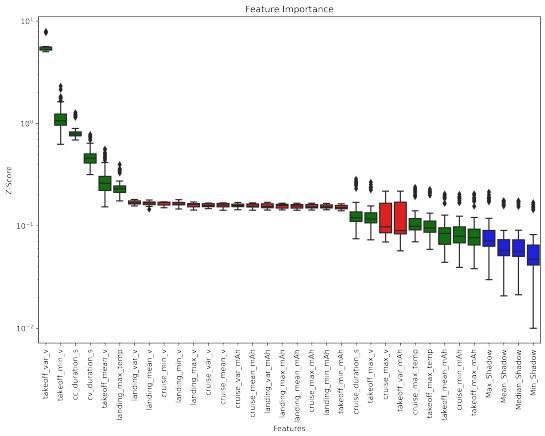
J RUL Feature Importance Plots



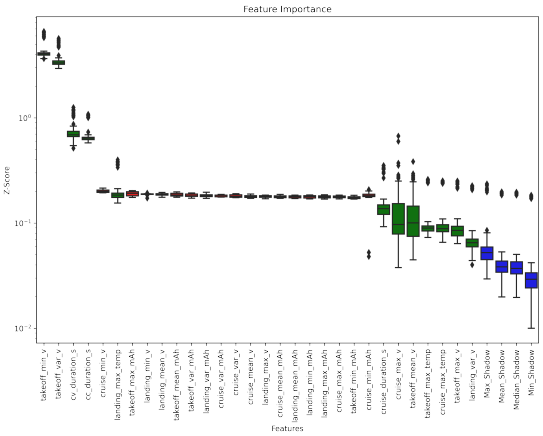
(a) VAH01



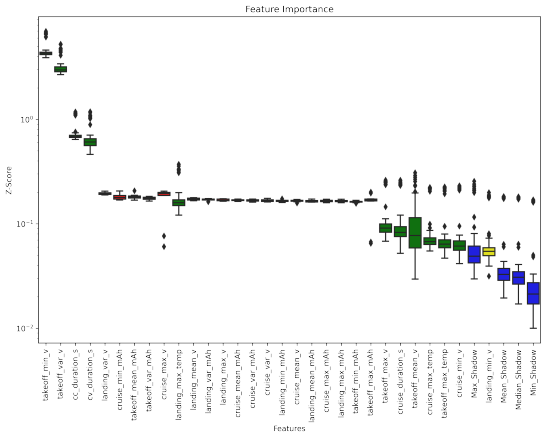
(b) VAH02



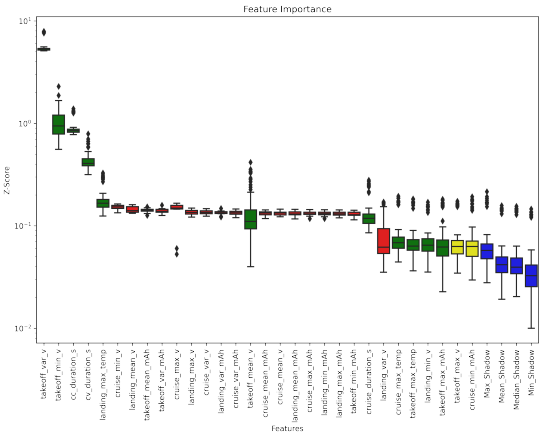
(c) VAH05



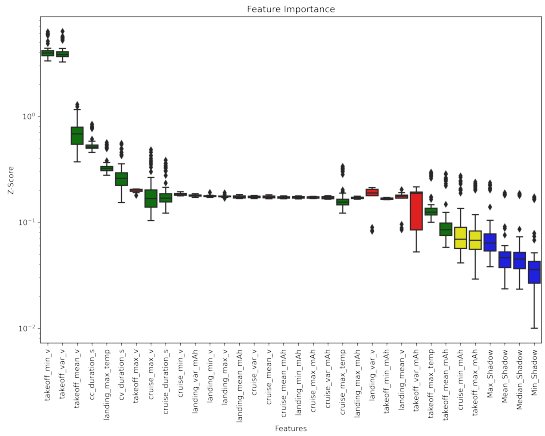
(d) VAH06



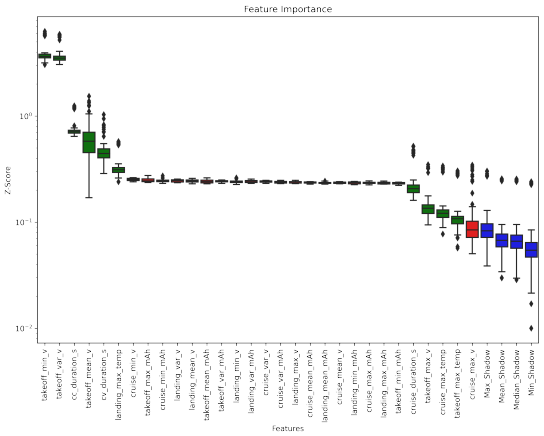
(e) VAH09



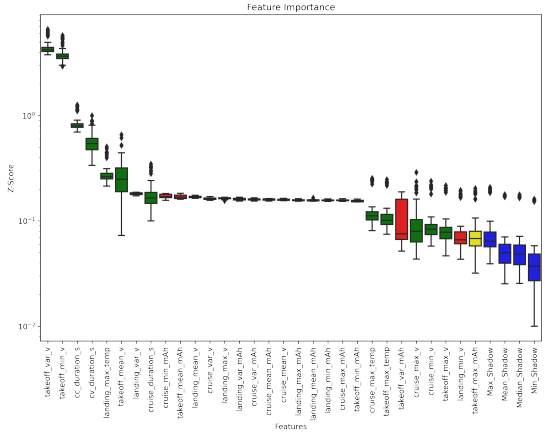
(f) VAH10



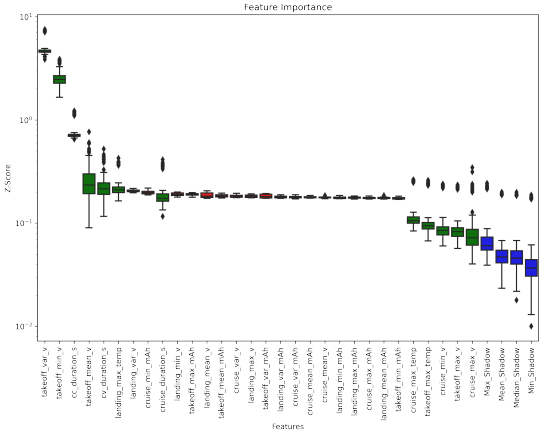
(g) VAH11



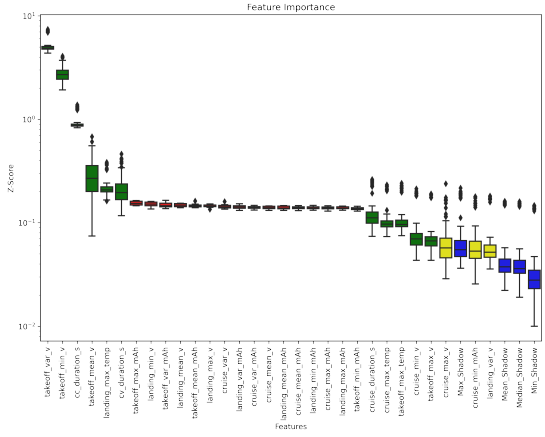
(h) VAH12



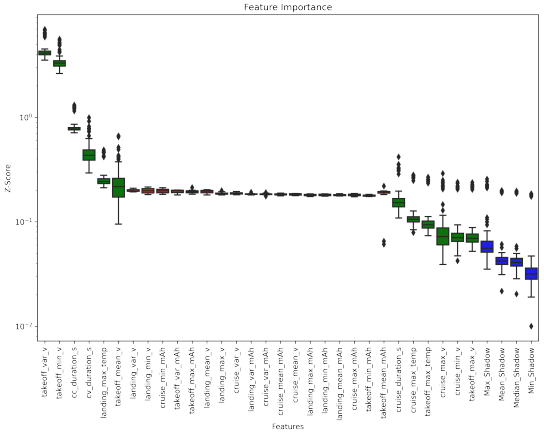
(i) VAH13



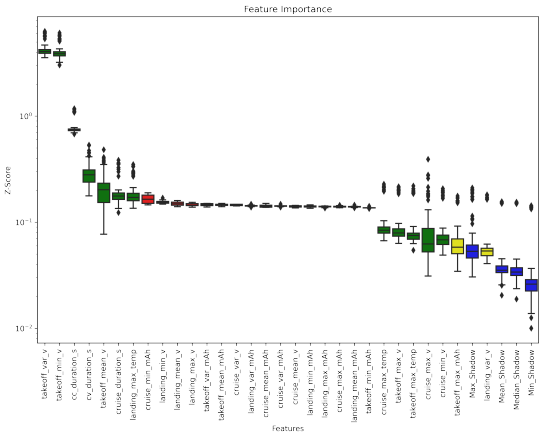
(j) VAH15



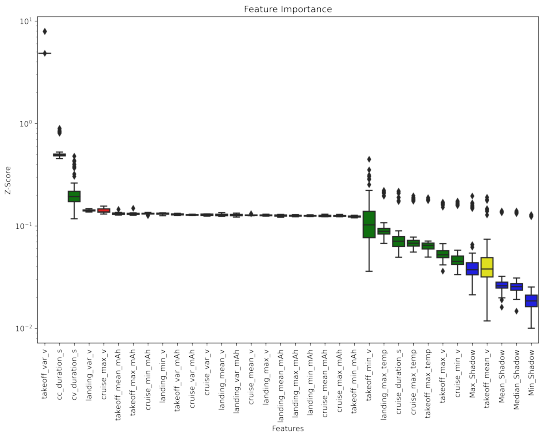
(k) VAH16



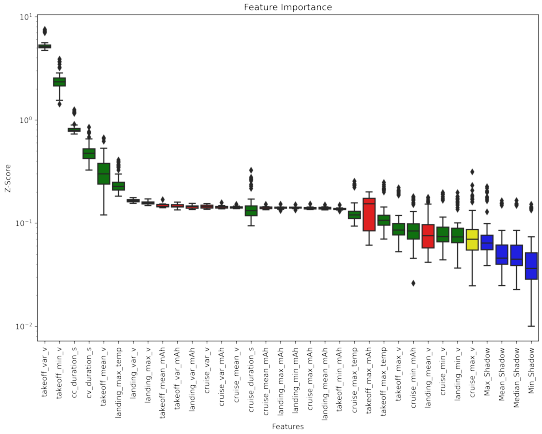
(l) VAH17



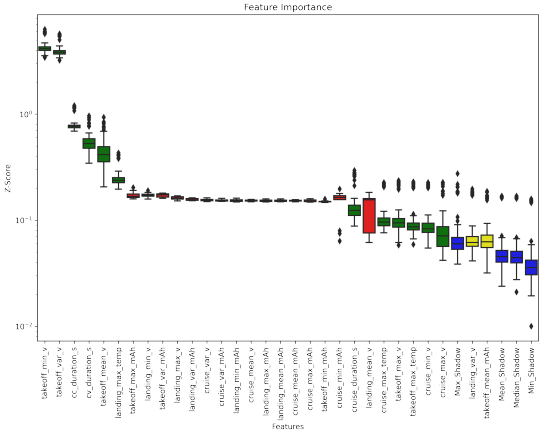
(m) VAH22



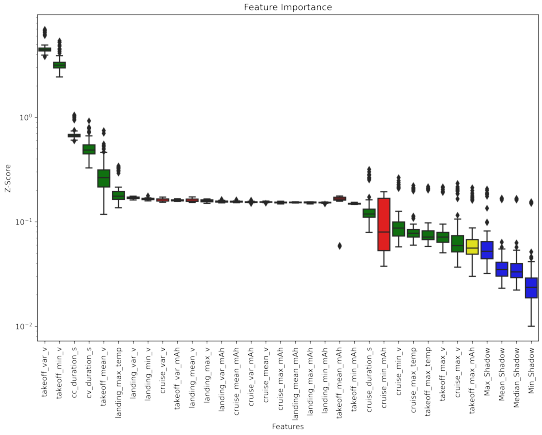
(n) VAH23



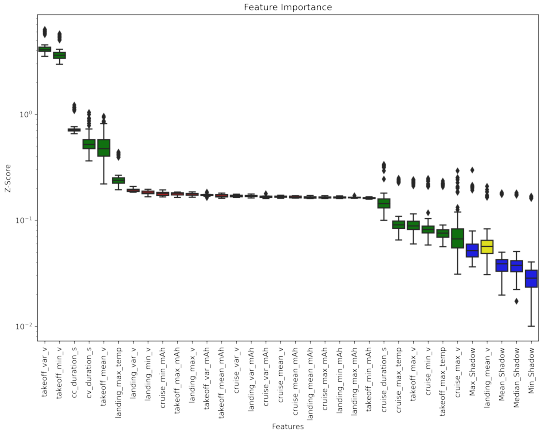
(o) VAH24



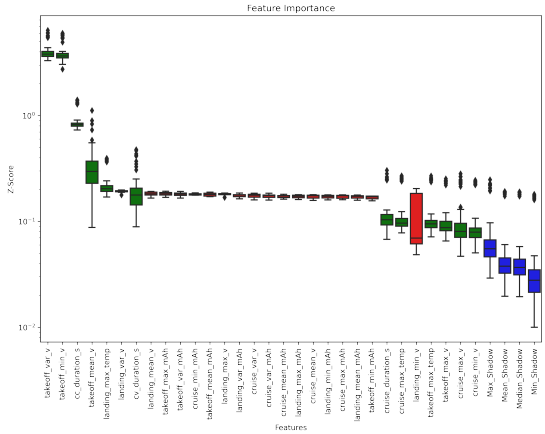
(p) VAH25



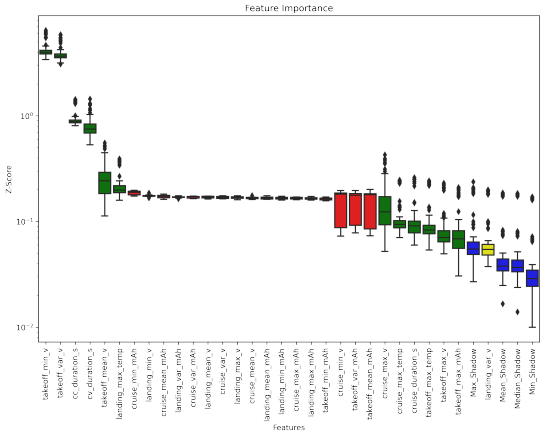
(q) VAH26



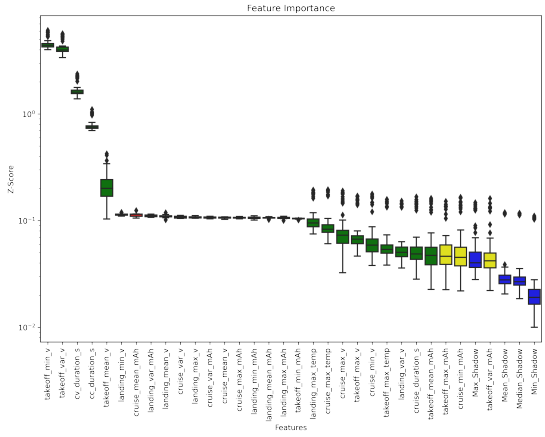
(r) VAH27



(s) VAH20

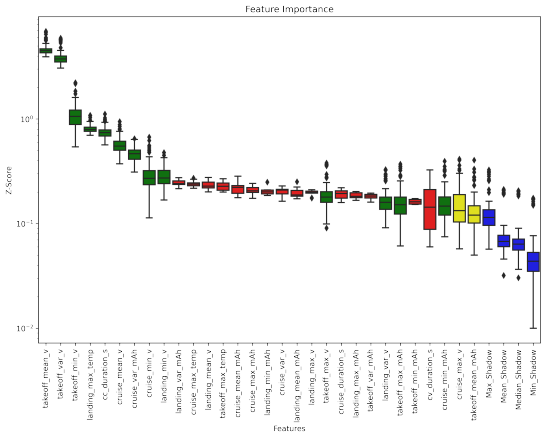


(t) VAH28

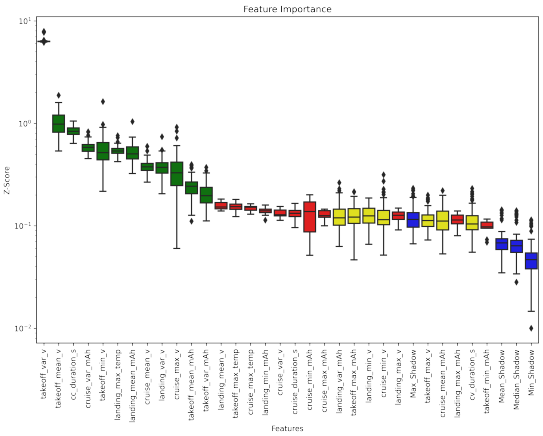


(u) VAH30

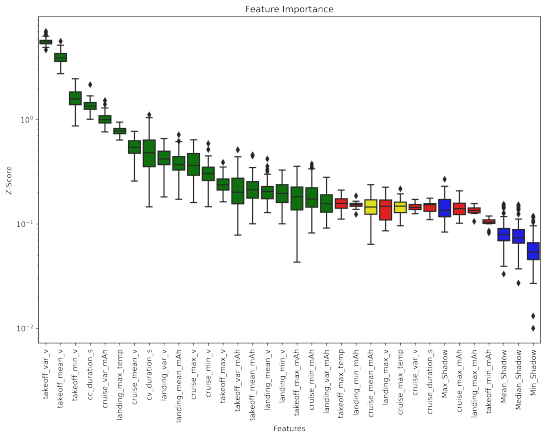
Figure 21: The figures show the feature importance for SOH of eVTOL batteries. Blue indicates shadow features, Green indicated important features, Red indicated unimportant features, while Yellow indicated tentative features.



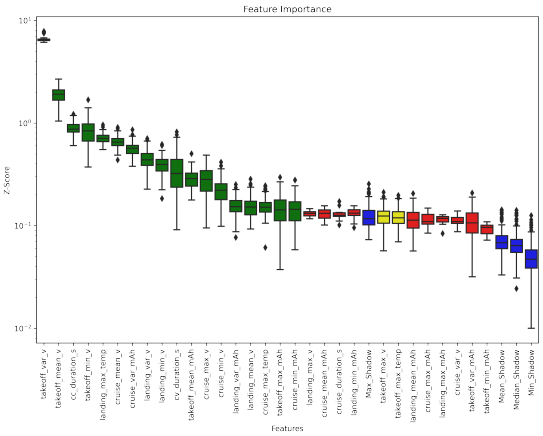
(a) VAH01



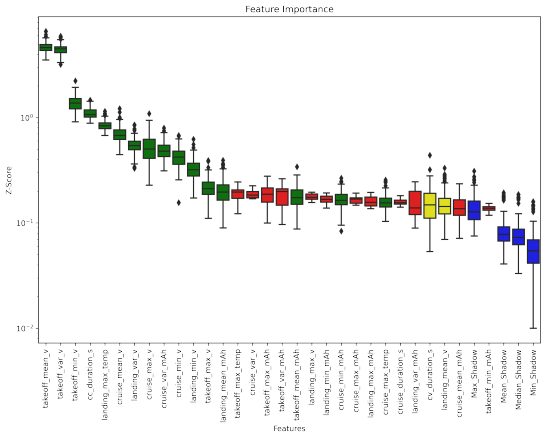
(b) VAH02



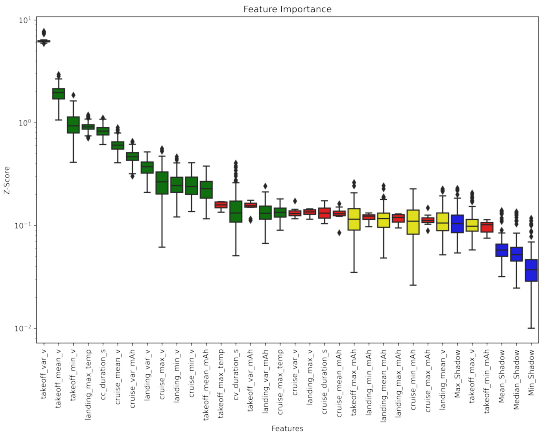
(c) VAH05



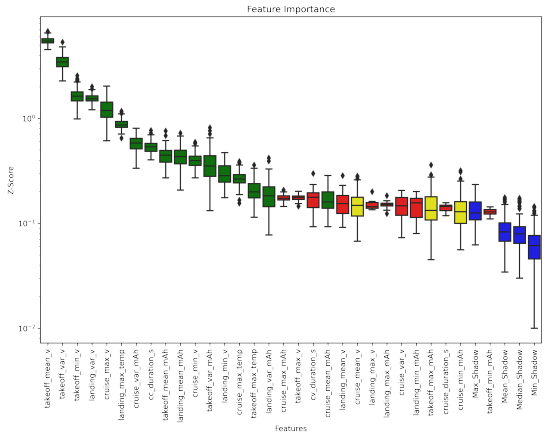
(d) VAH06



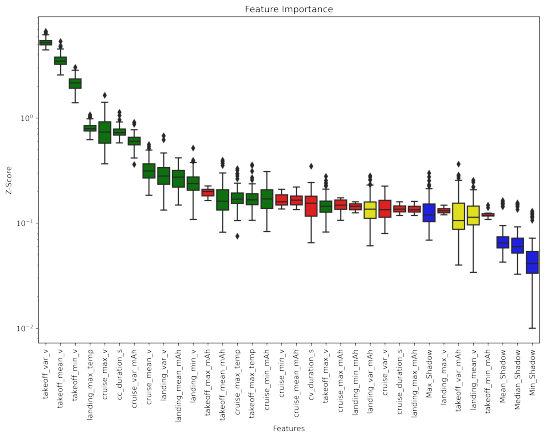
(e) VAH09



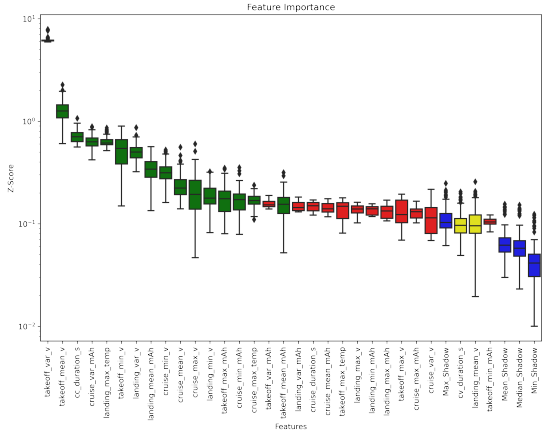
(f) VAH10



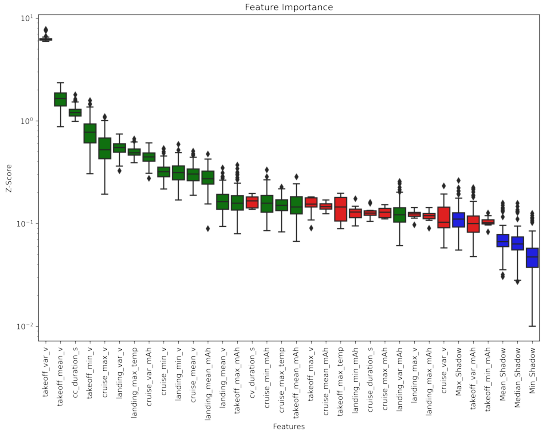
(g) VAH11



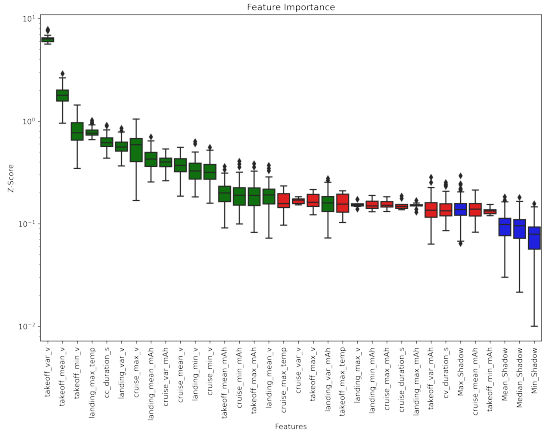
(h) VAH12



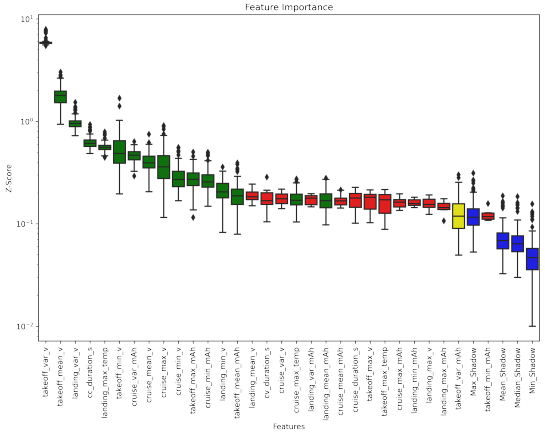
(i) VAH13



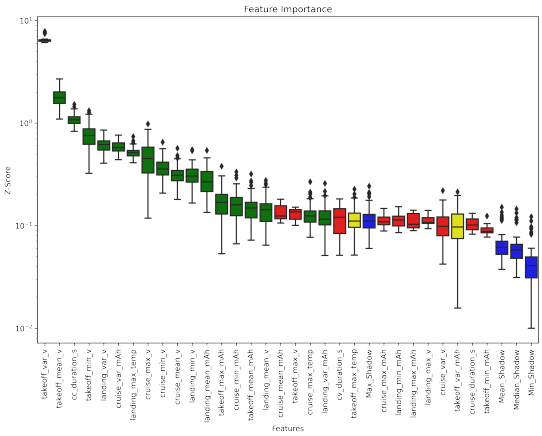
(j) VAH15



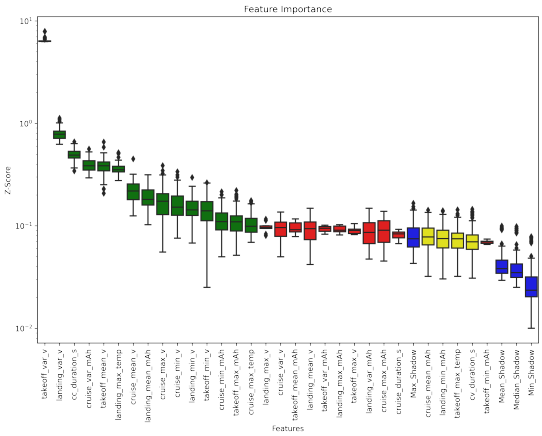
(k) VAH16



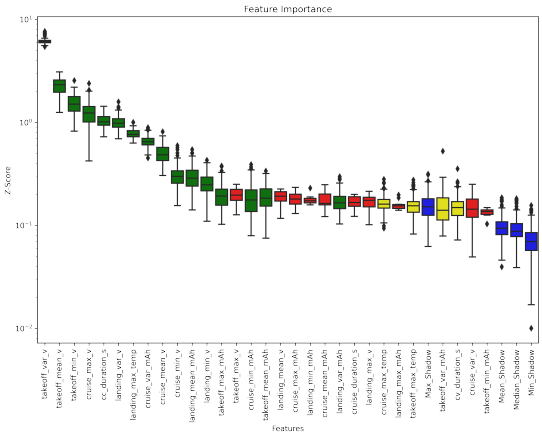
(l) VAH17



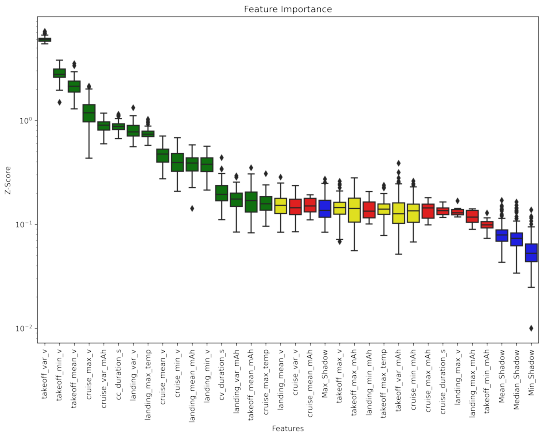
(m) VAH22



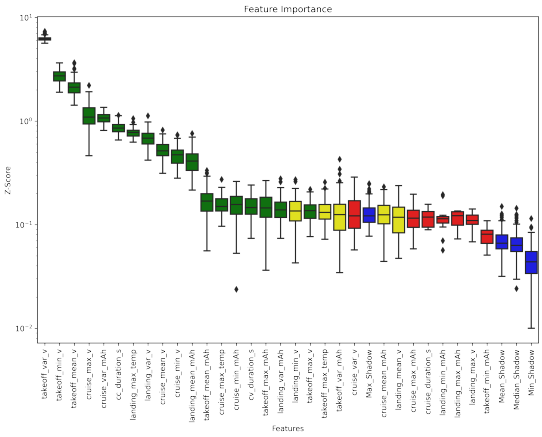
(n) VAH23



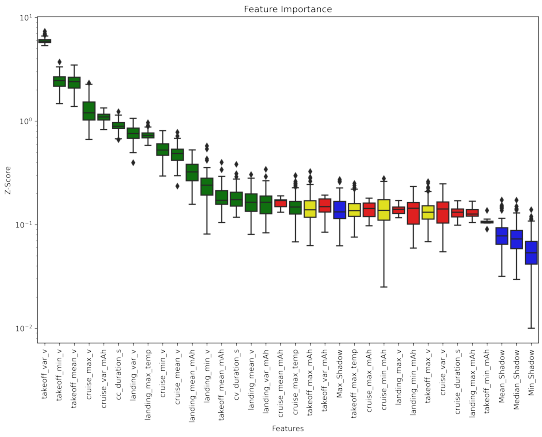
(o) VAH24



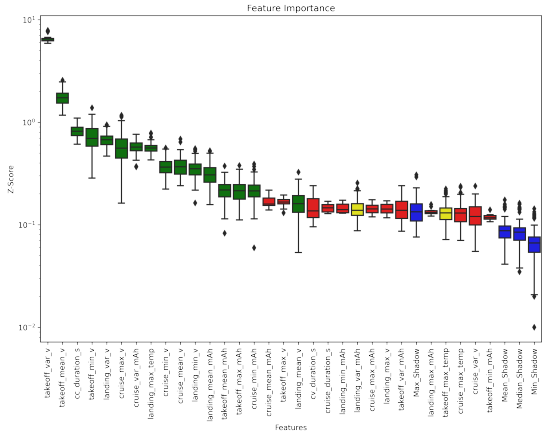
(p) VAH25



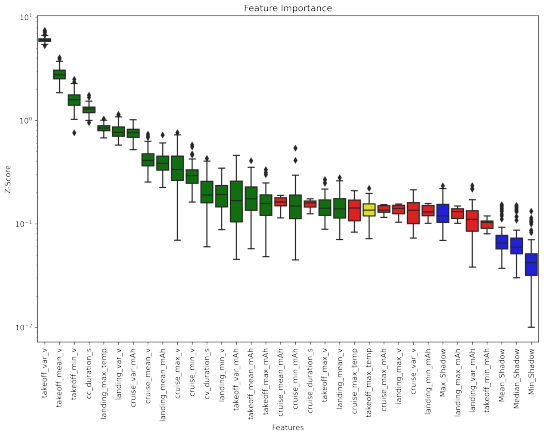
(q) VAH26



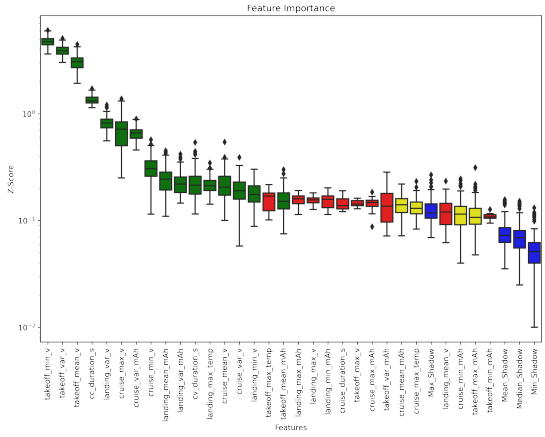
(r) VAH27



(s) VAH20



(t) VAH28



(u) VAH30

Figure 22: The figures show the feature importance for RUL of eVTOL batteries. Blue indicates shadow features, Green indicated important features, Red indicated unimportant features, while Yellow indicated tentative features.

References

[1] Moloud Abdar, Farhad Pourpanah, Sadiq Hussain, Dana Rezazadegan, Li Liu, Mohammad Ghavamzadeh, Paul Fieguth, Xiaochun Cao, Abbas Khosravi, U. Rajendra Acharya, Vladimir Makarenkov, and Saeid Nahavandi. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion*, 76:243–297, 2021.

[2] Airbus. Vahana- our single-seat evtol demonstrator. Available at https://www.airbus.com/en/innovation/low-carbon-aviation/urban-air-mobility/cityairbus-nextgen/vahana, 2019.

[3] Javier Alba-Maestre, Koen Prud'homme van Reine, Tomas Sinnige, and Saullo G. P. Castro. Preliminary propulsion and power system design of a tandem-wing long-range evtol aircraft. *Applied Sciences*, 11(23), 2021.

[4] Luis Basora, Arthur Viens, Manuel Arias Chao, and Xavier Olive. A benchmark on uncertainty quantification for deep learning prognostics, 2023.

[5] Luca Biggio, Alexander Wieland, Manuel Arias Chao, Iason Kastanis, and Olga Fink. Uncertainty-aware prognosis via deep gaussian process. *IEEE Access*, 9:123517–123527, 2021.

[6] Alexander Bills, Shashank Sripad, Leif Fredericks, Matthew Guttenberg, Devin Charles, Evan Frank, and Venkatasubramanian Viswanathan. A battery dataset for electric vertical takeoff and landing aircraft. *Scientific Data*, 10, 06 2023.

[7] Alexander Bills, Venkatasubramanian Viswanathan, Shashank Sripad, Evan Frank, Devin Charles, and William Leif Fredericks. evtol battery dataset. Available at https://kilthub.cmu.edu/articles/dataset/eVTOL_Battery_Dataset/14226830?file=26855063, 2022.

[8] Christopher M. Bishop. Mixture density networks. 1994.

[9] David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, apr 2017.

[10] L Breiman. Random forests. *Machine Learning*, 45:5–32, 10 2001.

[11] Joã o Caldeira and Brian Nord. Deeply uncertain: comparing methods of uncertainty quantification in deep learning algorithms. *Machine Learning: Science and Technology*, 2(1):015002, dec 2020.

[12] Ingeborg de Pater and Mihaela Mitici. Novel metrics to evaluate probabilistic remaining useful life prognostics with applications to turbofan engines. *PHM Society European Conference*, 7:96–109, 06 2022.

[13] Ingeborg de Pater, Arthur Reijns, and Mihaela Mitici. Alarm-based predictive maintenance scheduling for aircraft engines with imperfect remaining useful life prognostics. *Reliability Engineering System Safety*, 221:108341, 2022.

[14] Itamar Faran. Crps — a scoring function for bayesian machine learning models. Available at https://towardsdatascience.com/crps-a-scoring-function-for-bayesian-machine-learning-models-dd55a7a337a8.

[15] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning, 2016.

[16] Iffat A. Gheyas and Leslie S. Smith. Feature subset selection in large dimensionality domains. *Pattern Recognition*, 43(1):5–13, 2010.

[17] Tilmann Gneiting and Adrian Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378, 03 2007.

[18] Tilmann Gneiting, Adrian E. Raftery, Anton H. Westveld, and Tom Goldman. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, 133(5):1098 – 1118, 2005.

[19] Lérys Granado, Mohamed Ben-Marzouk, Eduard Solano Saenz, Yassine Boukal, and Sylvain Jugé. Machine learning predictions of lithium-ion battery state-of-health for evtol applications. *Journal of Power Sources*, 548:232051, 2022.

[20] Ahmet Kara. A data-driven approach based on deep neural networks for lithium-ion battery prognostics. *Neural Computing and Applications*, 33:1–14, 10 2021.

[21] Eoghan Keany. BorutaShap : A wrapper feature selection method which combines the Boruta feature selection algorithm with Shapley values., November 2020.

[22] Phattara Khumprom and Nita Yodo. A data-driven predictive prognostic model for lithium-ion batteries based on a deep learning algorithm. *Energies*, 12(4):660, Feb 2019.

[23] Serkan Kiranyaz, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J. Inman. 1d convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151:107398, 2021.

[24] Volodymyr Kuleshov and Stefano Ermon. Sampling methods. Available at https://ermongroup.github.io/cs228notes/inference/sampling/.

[25] Volodymyr Kuleshov and Stefano Ermon. Variational inference. Available at https://ermongroup.github.io/cs228notes/inference/variational/.

[26] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles, 2017.

[27] Juseong Lee and Mihaela Mitici. Deep reinforcement learning for predictive aircraft maintenance using probabilistic remaining-useful-life prognostics. *Reliability Engineering System Safety*, 230:108908, 2023.

[28] Jundong Li, Kewei Cheng, Suhang Wang, Fred Morstatter, Robert P. Trevino, Jiliang Tang, and Huan Liu. Feature selection: A data perspective. *ACM Computing Surveys*, 50(6):1–45, dec 2017.

[29] Yi Li, Changfu Zou, Maitane Berecibar, Elise Nanini-Maury, Jonathan C.-W. Chan, Peter van den Bossche, Joeri Van Mierlo, and Noshin Omar. Random forest regression for online capacity estimation of lithium-ion batteries. *Applied Energy*, 232:197–210, 2018.

[30] MS Hossain Lipu, MA Hannan, Aini Hussain, Afida Ayob, Mohamad HM Saad, Tahia F Karim, and Dickson NT How. Data-driven state of charge estimation of lithium-ion batteries: Algorithms, implementation factors, limitations and future trends. *Journal of Cleaner Production*, 277:124110, 2020.

[31] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions, 2017.

[32] Wilson E. Marcílio and Danilo M. Eler. From explanations to feature selection: assessing shap values as feature selection mechanism. In *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 340–347, 2020.

[33] James E. Matheson and Robert L. Winkler. Scoring rules for continuous probability distributions. *Management Science*, 22:1087–1096, 1976.

[34] Huixing Meng and Yan-Fu Li. A review on prognostics and health management (phm) methods of lithium-ion batteries. *Renewable and Sustainable Energy Reviews*, 116:109405, 2019.

[35] Celina Mikolajczak, Michael Kahn, Kevin White, and Richard Thomas Long. *Lithium-ion batteries hazard and use assessment*. Springer Science & Business Media, 2012.

[36] Mihaela Mitici, Ingeborg de Pater, Anne Barros, and Zhiguo Zeng. Dynamic predictive maintenance for multiple components using data-driven probabilistic rul prognostics: The case of turbofan engines. *Reliability Engineering System Safety*, 234:109199, 2023.

[37] Mihaela Mitici, Birgitte Hennink, Marilena Pavel, and Jianning Dong. Prognostics for lithium-ion batteries for electric vertical take-off and landing aircraft using data-driven machine learning. *Energy and AI*, 12:100233, 2023.

[38] Christoff Molnar. Interpretable machine learning - a guide for making black box models explainable). Available at https://christophm.github.io/interpretable-ml-book/shap.html#shap-feature-importance, 2023.

[39] Derrick Mwiti. Random forest regression: When does it fail and why? Available at https://neptune.ai/blog/random-forest-regression-when-does-it-fail-and-why, 2023.

[40] Vijay Mohan Nagulapati, Hyunjun Lee, DaWoon Jung, Boris Brigljevic, Yunseok Choi, and Hankwon Lim. Capacity estimation of batteries: Influence of training dataset size and diversity on data driven prognostic models. *Reliability Engineering System Safety*, 216:108048, 2021.

[41] Vijay Mohan Nagulapati, Hyunjun Lee, DaWoon Jung, SalaiSargunan S Paramanantham, Boris Brigljevic, Yunseok Choi, and Hankwon Lim. A novel combined multi-battery dataset based approach for enhanced prediction accuracy of data driven prognostic models in capacity estimation of lithium ion batteries. *Energy and AI*, 5:100089, 2021.

[42] Eric Thomas Nalisnick. *On Priors for Bayesian Neural Networks*. PhD thesis, UC Irvine, 2018.

[43] Man-Fai Ng, Jin Zhao, Qingyu Yan, Gareth Conduit, and Zhi Seh. Predicting the state of charge and health of batteries using data-driven machine learning. *Nature Machine Intelligence*, 2, 03 2020.

[44] Nicholas Polaczyk, Enzo Trombino, Peng Wei, and Mihaela Mitici. A review of current technology and research in urban on-demand air mobility applications. In *8th Biennial Autonomous VTOL technical meeting and 6th Annual electric VTOL Symposium*, pages 333–343, 2019.

[45] Joris Pries, Guus Berkelmans, Sandjai Bhulai, and Rob van der Mei. The berkelmans-pries feature importance method: A generic measure of informativeness of features, 2023.

[46] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, 2006.

[47] Lei Ren, Li Zhao, Sheng Hong, Shiqiang Zhao, Hao Wang, and Lin Zhang. Remaining useful life prediction for lithium-ion battery: A deep learning approach. *IEEE Access*, 6:50587–50598, 2018.

[48] Darius Roman, Saurabh Saxena, Valentin Robu, Michael Pecht, and David Flynn. Machine learning pipeline for battery state of health estimation, 2021.

[49] Yves-Laurent Kom Samo. Effective feature selection: Beyond shapley values, recursive feature elimination (rfe) and boruta. Available at https://blog.kxy.ai/effective-feature-selection/index.html, 2022.

[50] Carlos Sebastián and Carlos E. González-Guillén. A feature selection method based on shapley values robust to concept shift in regression, 2023.

[51] Carolin Strobl, Torsten Hothorn, and Achim Zeileis. Party on! a new, conditional variable importance measure for random forests available in the party package. *The R Journal*, 1:14–17, 12 2009.

[52] Maxime Taillardat, Olivier Mestre, Michaël Zamo, and Philippe Naveau. Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Monthly Weather Review*, 144(6):2375 – 2393, 2016.

[53] B. Venkatesh and J. Anuradha. A review of feature selection and its methods. *Cybernetics and Information Technologies*, 19(1):3–26, 2019.

[54] Ziming Wang, Chaohao Liao, Xu Hang, Lishuai Li, Daniel Delahaye, and Mark Hansen. Distribution prediction of strategic flight delays via machine learning methods. *Sustainability*, 14(22):15180, Nov 2022.

[55] Jingwen Wei, Guangzhong Dong, and Zonghai Chen. Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression. *IEEE Transactions on Industrial Electronics*, 65(7):5634–5643, 2018.

[56] Wikipedia. evtol. Available at https://en.wikipedia.org/wiki/EVTOL.

[57] Xiaodong Xu, Shengjin Tang, Chuanqiang Yu, Jian Xie, Xuebing Han, and Minggao Ouyang. Remaining useful life prediction of lithium-ion batteries based on wiener process under time-varying temperature condition. *Reliability Engineering System Safety*, 214:107675, 2021.

[58] Lei Yu and Huan Liu. Eficient feature selection via analysis of relevance and redundancy. *Journal of Machine Learning Research*, 5:1205–1224, 12 2004.

[59] Michaël Zamo and Philippe Naveau. Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts. *Mathematical Geosciences*, 11 2017.

[60] Micha Zoutendijk and Mihaela Mitici. Probabilistic flight delay predictions using machine learning and applications to the flight-to-gate assignment problem. *Aerospace*, 8(6):152, May 2021.