Predicting Haze Formation in Titan's Upper Atmosphere Using Artificial Neural Networks

Third Year Physics Project

Supervisor: Ravindra T. Desai
Assessor: Jonathan P. Eastwood

Ewan Saw
CID: 01521087

Word Count: 5987

Department of Physics
Blackett Laboratory
Imperial College London
19/1/2021

Abstract

In this study, we explore the viability of using artificial neural networks to predict the spatial distributions of anions densities within Titan's ionosphere. A systematic search for the optimal hyperparamaters was conducted to give a working model for predictions. This neural network was then trained on Cassini's Electron Spectrometer data before being used to produce results. The predictions made successfully reproduced key trends of Titan's ionosphere that have been identified in previous studies, namely that of the anion's dependence on photo-ionisation mechanisms, as well as the general spatial distributions of these anions. Despite there being unexpected behaviour in the results, these are likely due to problems within the data-set. Overall, this study manages to show the ability of neural networks, despite some minor disadvantages. This promises high implications for future investigations of Titan.

CONTENTS

I. Introduction

TITAN is the largest moon of Saturn and the second largest natural satellite of the Solar System. Its extended atmosphere is rich with complex organic chemistry[13] and is densely packed with molecular methane, nitrogen and hydrogen. [12][5] It is known that a thick layer of photo-chemical haze envelops the moon. These aerosol-type particles obscure surface observations of Titan as it blocks visible light from the sun and other sources. The composition of this thick orange haze was studied during the *Cassini-Huygens* mission where its production mechanisms were explored.[13] However, as it stands, we are yet to fully understand the processes under which these naturally occurring organic compounds are formed.

Based on data collected from 16 flybys, the *Cassini* spacecraft's Electron Spectrometer (ELS), which is an instrument of the *Cassini* Plasma Spectrometer (CAPS), detected negative ions (anions) of masses up to $10,000$ amu/q within Titan's ionosphere, at altitudes of $950 - 1400$ km.[3][18] This discovery was unexpected, but it is believed that these anions could be precursors of the aerosol-particles forming down below.[15][10]

Of all the anions detected, only two distinct species have been identified from the CAPS-ELS data, i.e. the two lightest mass groups.[13] Due to the energy resolution of the ELS, precise identifications of the heavier species have been limited. Given this, the speed at which we are approaching a full picture of Titan's chemistry is severely restricted, especially considering the complexity of its atmosphere.

As such, we resort to finding other tools that could aid us in understanding Titan's chemistry. In this study, we use this as a motivation to explore an alternate method which could help speed up this process - Artificial Neural Networks (ANNs). In this investigation, we wish to determine the viability of using neural networks to gain further insight into the chemistry of Titan's upper atmosphere. By constructing an ANN that is able to train on the sparse sets of CAPS-ELS data through deep learning, we aim to accurately predict the spatial distributions of the low mass anion densities and produce global heat-maps of the moon's ionosphere.

A successful step in this direction would promise the possibility of developing a more complex neural network; one that could train on existing measurements to predict the distributions of the higher mass ion groups. This in turn would allow us to predict the formation of Haze on Titan, which could be achieved by feeding such a neural network with our current knowledge of Titan's chemical models. However, such a step will be left for a future study.

II. Theory

A. Titan's Upper Atmosphere

As mentioned in section I, *Cassini's* spacecraft discovered negative ions within Titan's ionosphere, where two low-mass anion species have recently been identified. These correspond to carbon chain atoms CN^- and/or C_2H^- for the lightest species, and C_3N^- and/or C_4H^- for the second lightest.[13] It was found that it was not possible to further resolve the differences between these nitrile and hydrocarbon compounds from the ELS data, hence we treat their trends as combined in the context of this study.[13]

The spatial distributions of these anions were previously explored, and key features of the CAPS-ELS data have been highlighted. It was concluded that the maximum altitudes at which negative ions of a given mass were found decreased with an increasing mass.[19] Thus, we expect the ionosphere to be denser at lower altitudes. It was also found that the low mass anions have the highest densities on the day-side; close to midday and around the equator, and low densities on the night-side and close to the poles. Conversely, the high mass anions reach their highest densities on the night-side.[20][21] These findings present the key features we would need to reproduce with our predicted maps. Successfully doing so would aid in validating our model.

1) Production and Loss Routes of Low Mass Anions: It was found that these low mass anion groups in the ionosphere are produced through dissociative electron attachment to, or de-protonation of their respective neutral parent species.[17][8] In the case of CN^- and C_3N^-, they are produced by

$$\mathrm{HCN} + \mathrm{e}^- \longrightarrow \mathrm{H} + \mathrm{CN}^-, \tag{1}$$

$$\mathrm{HC_3N} + \mathrm{CN}^- \longrightarrow \mathrm{HCN} + \mathrm{C_3N}^-. \tag{2}$$

These reactions are favourable given the abundances of HCN and HC_3N in the atmosphere. The same mechanisms apply to the hydrocarbons,[13]

$$\mathrm{C_2H_2} + \mathrm{e}^- \longrightarrow \mathrm{H} + \mathrm{C_2H}^-, \tag{3}$$

$$\mathrm{C_4H_2} + \mathrm{C_2H}^- \longrightarrow \mathrm{C_2H_2} + \mathrm{C_4H}^-. \tag{4}$$

Again, the abundances of the parent species C_2H_2 and C_4H_2 seem to imply that these reactions proceed rapidly, with the only limiting reagent being the availability of free electrons, as shown in reactions 1 and 3. The main loss processes of these anion species were also found to be due to associative detachment with neutral radicals.[17]

A recent investigation into the trends exhibited by both mass groups (CN^-/C_2H^- and C_3N^-/C_4H^-) showed that they displayed a high degree of similarity, such that their properties could be discussed together.[15] Hence for simplicity, this study will only investigate and predict the spatial distributions of the CN^-/C_2H^- group.

2) Electron Density Dependence: It was previously measured that Titan's thermal ionospheric plasma is cold ~ 0.1 eV.[24] Despite this, photo-ionisation processes occurring on Titan's day side produce large populations of supra-thermal electrons.[27][9] [4] Given the production mechanisms described in the previous section, one would expect there to be a correlation between these two events; the high production of electrons favouring the forward reactions of reactions 1 to 4. In fact, it was found from the ELS data that the low-mass anion densities in the ionosphere are approximately *linearly* proportional to the electron densities on the day-side of Titan. However, no obvious dependence was identified on the night-side.[15]

The former relation is in agreement with the production channels of the mass groups outlined in section II-A1, and highlights another key feature to expect from the predicted ion density heat maps produced by our neural network.

3) Formation of Heavier Anion Mass Groups: It is suggested that the heavier anion mass groups play a large role in the formation of organic aerosol particles found on Titan.[3][18] Hence, it is useful to understand the conditions under which they are formed. It was found that the heaviest anions are most likely to form at high latitudes and winter conditions, where there is low solar flux.[14] As such, greater densities of high mass anions are expected near the north pole and at lower altitudes.

Despite our study being limited to only predicting the distributions of the low mass anions, we attempt to relate our results to the traits of the higher mass anions in section IV-D.

B. Artificial Neural Networks (ANNs)

An ANN is a type of computational model, designed to simulate the processes under which animal and human brains analyse and process information. They are composed of various numbers of layers and nodes, as shown in Fig. 2.

To begin, we need to determine the type of problem at hand. Issues that employ deep learning methods are grouped as either *classification* or *regression*-type problems.[28] In the context of this study, we wish to predict values of a continuous variable - anion density. Hence, we need to construct a regression-type model. In this subsection, we outline the general architecture of a neural network and introduce the various hyperparameters that need to be considered.

1) Nodes & Layers: The main constituents of a neural network are its artificial neurons, often referred to as nodes. The basic function of a node is outlined in Fig. 1. The operation performed by a typical node can be

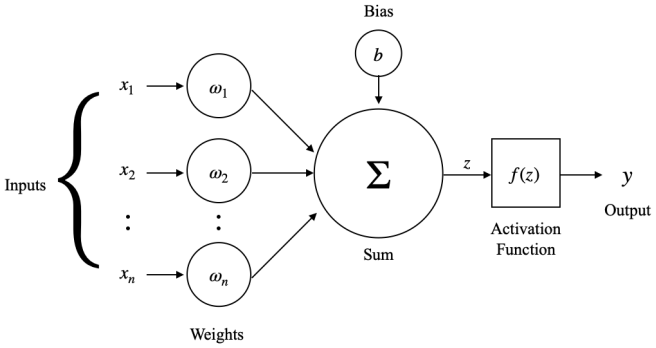


Fig. 1. Simple model of an artificial neuron. It is shown that the neuron takes a given number of inputs x_i, where $i = 1, ..., n$. Each input is multiplied by an associated weight w_i and summed. A bias b may also be added to the overall sum and acts as a constant which delays the triggering of the activation function. This is then passed through the function $f(z)$ to give an output of y (reproduced from [1]).

summarised as

$$y = f\left(\sum_i^n \omega_i x_i + b\right), \tag{5}$$

with f as its activation function.

Large numbers of these nodes are then grouped together to form layers of artificial neurons. These layers are then stacked according to Fig. 2 to form the basic structure of a neural network. It is technically possible to approximate any function using a model consisting of only 1 hidden layer, given a sufficiently large number of nodes. However, this is likely to fail in terms of efficiency and in most cases, multiple layers are required.[32] In constructing an ANN, we specify these two main hyperparameters; the number of nodes and layers. Systematic experimentation with different combinations of these values would inform us of the optimal configurations which pertain to our problem.

2) Activation Function: The choice of activation function f which acts on the sum performed by a given node defines its 'activation' (see Fig. 1). This weighted sum denoted by z is usually referred to as the summed activation of a node.[31] Typically, non-linear activation functions are preferred, as a linear activation function would restrict our model to a linear solution space. This

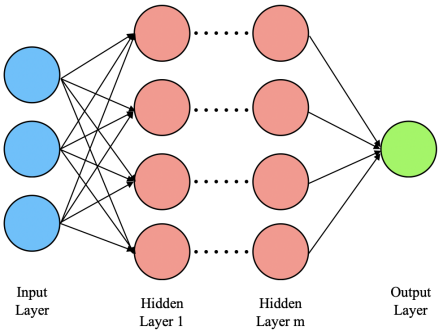


Fig. 2. Architecture of an artificial neural network. This is composed of an input layer, m number of hidden layers, and an output layer. Data is fed into the neural network through the input layer, and all the computation occurs within the intermediate layers where all nodes interact, before producing a result which is passed onto the output layer.

can be easily shown by considering the linearity of two functions, f and g,

$$h(z) = f(z) + g(z), \tag{6}$$

such that we obtain a linear solution, h. As such, a popular choice for f is the sigmoid function,[31]

$$S(x) = \frac{1}{1 + e^{-x}}, \tag{7}$$

which outputs a value within the range $0 \leq S(z) \leq 1$. This gives a range of probabilities for the 'activation' of a node. However, equation 7 works best for *classification*-type problems given its range of 0 to 1.

A more appropriate choice for our regression-type problem would be the Scaled Exponential Linear Unit (SELU). This is given by

$$SELU(x) = \lambda \begin{cases} x, & \text{if } x > 0 \\ \alpha e^x - \alpha, & \text{if } x \leq 0 \end{cases}, \tag{8}$$

where α and λ are constants determined through the training process. Seeing that the range of this function extends to infinity, it is useful in predictions of continuous quantities. A variety of other activation functions with a similar form to equation 8 also exist for the construction of ANNs. As such, we explore these further options in section III-C.[33]

3) Loss Function: In mathematical optimisation and decision theory, the 'loss function' (sometimes called the 'cost function') typically denotes the function of a problem we wish to minimise.[29] In the context of ANNs, this is usually defined by a metric of error which we also want to minimise, as this quantifies the distance between the true and predicted values of a target.[36]

The most common choices of loss functions in regression problems are the mean-sqaured error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2, \tag{9}$$

and the mean-absolute error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|, \tag{10}$$

where $\hat{y}_i$ is the prediction of the true value y_i. The main distinction between using MSE or MAE as loss functions is that the former applies more weight to outliers in the results (given the squared difference). In the context of Titan, we do not know *a priori* if we wish to penalise our model for any extreme predictions, thus, both functions are considered in the construction of our ANN.

4) Optimisers: Optimisers act to update the parameters of our model throughout each iteration of the training process. This process is also known as 'back propagation', as the algorithm propagates backwards to update the weights of the nodes. These aim to minimise the output of our chosen loss function; 'optimising' the weights and biases to converge on the global minimum. Implementation of different optimisation algorithms vary in complexity and are provided by most deep learning libraries, however the core concepts of each method remain the same.

Given a loss function $J(\theta)$, where θ defines our model parameters, an optimisation algorithm will attempt to converge on the values of the global minimaum $J(\theta^*)$ using gradient descent. A direct implementation of this involves taking small-iterative steps in the direction opposite to the gradient $\nabla_\theta J(\theta)$, where

$$\theta_{n+1} = \theta_n - \eta \cdot \nabla_\theta J(\theta_n). \tag{11}$$

Here, the step-size η can be interpreted as the 'learning rate' of our model. This method is iteratively applied until we approximate θ^* to a sufficient degree of accuracy. An extra hyperparameter we define in this 'training' process is the number of epochs, this denotes the number of times the optimisation algorithm sees the complete set of data.

Unfortunately, the method described by equation 11 suffers in poor memory usage and overall efficiency. It can also get stuck on a local minimum and not converge onto the true value of θ^*. Thus, we improve on this method using mini-batch stochastic gradient descent optimisers which can escape local minima. These approximate the $\nabla_\theta J(\theta)$ term differently and are explored in section III of constructing our ANN. [34][36]

III. METHOD

This section outlines the main steps taken in constructing a working ANN, by applying the theories defined in section II-B. Before proceeding, we needed to 'prepare' the data-set for our model to train on.

A. CAPS-ELS Data-set

The CAPS-ELS data-set that we used was obtained from a total of 28 Titan flybys. These readings were taken from 2004 October 26 to 2012 May 22 and their trends were previously discussed.[15]

1) MLP Detection Efficiency: The anion count rates, R_c, were then converted into their respective number densities, n_{ni}, using the ion current approximation[3][18]

$$n_{ni} = \frac{R_c}{v_{sc} A_F \varepsilon}. \tag{12}$$

Here, v_{sc} denotes the *Cassini* spacecraft's velocity in Titan's reference frame at the time of reading, $A_F = 0.33$ cm^2 is the effective area of acceptance[22], and ε is a dimensionless value which refers to the Microchannel Plate (MCP) anion detection efficiency. As the discovery of negative ions was unexpected, the response of the CAPS-ELS was not calibrated for the detection of anions; it was designed to measure electrons. Hence, the value of ε can only be estimated using nominal ion efficiencies and is found to be energy dependent.[15]

The values of ε for the MCP detection efficiency used in previous studies have varied depending on different mass groups, with values of $\varepsilon = 0.05, 0.25$ and 0.5 being used. The smallest efficiency of $\varepsilon = 0.05$ has been found to be the best estimate corresponding to larger anion species.[25] Studies have suggested that a value of $\varepsilon = 0.25$ may be more accurate for lower mass groups,[35] thus, this value was used throughout this study to normalise the anion densities, i.e. $n_{ni} \cdot \varepsilon$. The error introduced by the MCP detection efficiency is systematic, hence, this error can be overlooked as we are more concerned with the relative magnitude of distributions, as opposed to the absolute distributions of anions.[15]

2) Omission of Incomplete Data Points: Each measurement of the anion densities in the data-set is associated with 5 other variables; the electron density, electron temperature, and the 3 independent spatial coordinates (latitude, longitude and altitude). Of the 5 variables, the 2 non-spatial parameters were not recorded in some of the readings, likely due to the ELS's actuator not pointing in the correct position at the time of measurement.[15] For consistency, these incomplete data points have been omitted from this study and only the remaining readings were used to train our neural network.

3) Periodicity of Data: The spatial coordinates of the data have latitudes defined by $-90° \leq \phi \leq 90°$ (subtending the equator from the south to the north pole), longitudes $-180° \leq \theta \leq 180°$ (where $\theta = 0$ represents mid-day), and altitudes $950 \leq r \leq 1400$ km. We needed to ensure that our model was able to recognise the periodic nature of these spherical polar variables. For example, given a predicting function $P(r, \theta, \phi)$, we would require $P(r, \theta, \phi) = P(r, \theta + 360°, \phi)$.

To achieve this, copies of the data-set were shifted longitudinally and 'stitched' together to pad out the training data. Two 'stitches' were made; one was shifted by $-360°$ in the θ-direction and the other by $+360°$. Given this, the training data now exhibits this 'longitudinal periodicity' for the neural network to pick up on.

B. Constructing the ANN

Following the details of section II-B, we constructed a simple ANN that was able to read the CAPS-ELS data-set and make initial predictions. The initial choice of hyperparameters (numbers of nodes & layers etc.) were arbitrary, as there was no preceding indication of an optimal combination to use. Our ANN was constructed using Keras, a deep learning API written for Python.

To refine these choices we needed to evaluate the skill of our model for a given set of hyperparameters and systematically vary them to converge on an optimal level of performance. It is also possible to attempt a grid-search; an algorithm which cycles through all possible choices to find the optimal combinations of hyper-parameters. However, this was found to be very computationally heavy. A logical approach in choosing these parameters was concluded to be the most efficient.

1) Evaluating the Performance of the Model: The basis of our evaluation method lies within the use of an '$80/20$ train-test split'. This involves selecting a random 80% set of the CAPS-ELS data for the ANN to 'train' on, then using the resulting model to predict the values of anion densities at positions of the remaining 20% (see Fig. 4).[26] This gives an estimate on the performance of our model.

In this investigation, we considered 2 metrics of error to evaluate the ANN's performance. These were the RMS error,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}, \tag{13}$$

and the mean absolute percentage error (MAPE),

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right|, \tag{14}$$

where $\hat{y}_i$ and y_i are the predicted and true values respectively. Note that we use these in a different context to the loss functions, equations 9 and 10, which the optimisation algorithm seeks to minimise.

Using only a single metric of error limits our understanding of the performance of our model. Thus, we chose the error metrics of RMSE to evaluate the standard uncertainties of our predictions, and MAPE to determine their relative errors. We also noted that MAPE is likely to fail when evaluating predictions at low values of anion densities, i.e. when calculating the percentage error on a data point close to zero, which blows up the term and in some cases, exceeds 100%. Desipte this, MAPE is able to penalise a model for predicting extreme values.

We also had to consider the fact that the optimisation algorithms used by our ANN were stochastic, such that two models with identical hyperparameters would give slightly different predictions, even when trained on the same set of data. In order to to address this, we chose to repeat the training routine of our model 5 times. By wrapping this into a loop, we averaged the values of MAPE and RMSE to give an overall estimate of these errors, as well as their standard deviations. An important step was ensuring that we fixed the random seed of the 80/20 test-split, as we needed to ensure that we were training and testing the same sets of data over each cycle.[23]

C. Parameter Choices & Their Errors

We systematically experimented with different hyperparameters, evaluating the performance at each step to narrow down the optimal choices. This is summarised in Fig. 3. Note that the high values of MAPE are due to mis-predictions of small values, where division by a small number (in equation 14) dominates the value of the error. However, this section is mainly concerned with the relative improvements made through configuring our model, so this can be overlooked. After identifying the best hyperparameters, we present the final uncertainty of our setup in section III-D.

The main trend exhibited by subplots (a) - (e) in Fig. 3 is that in almost all cases, the model performs significantly better when trained using the electron data (and the few cases which do not exhibit this behaviour are credited to the high variance of the readings). This suggests that the anion densities do display a dependence on the electron densities, which is in agreement with the findings discussed in section II-A2. The overall significance this has on the performance of our ANN is outlined in section III-D.

Nodes: Fig.3(a) The general trend shown is that errors decrease with an increasing number of nodes up to ~ 200 nodes. An extreme evaluation of 1000 nodes was performed and this showed a decrease in performance. This phenomenon is likely due to overfitting. The number of function evaluations for our model should scale proportionally with the number of nodes present, hence we chose 200 nodes as our ideal hyperparameter with respect to performance and efficiency.

Layers: Fig.3(b) A similar trend is shown, where increasing the number of layers improves the performance of the model. Note that a higher number of layers also decreases the variance of predictions, giving more consistent results. A final setting of 6 functioning layers was chosen, as we are again limited by computational time.

Activation Function: Fig.3(c) Four different activation functions were considered: the Rectified Linear unit (ReLU), the Softmax function, SELU (equation 8) and the Sigmoid function (equation 7). ReLU is another function commonly used for regression-type problems and is given by

$$ReLU(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}. \tag{15}$$

The Softmax function is known to be a generalisation of the Sigmoid function to multiple dimensions, thus it is still limited to a range of 0 to 1. SELU was found to perform the best, followed by ReLU. As expected, the remaining classification-type activation functions gave higher errors. Despite it giving a slightly higher variance in predictions, SELU still performed better ReLU, thus this was chosen for our activation function.[28]

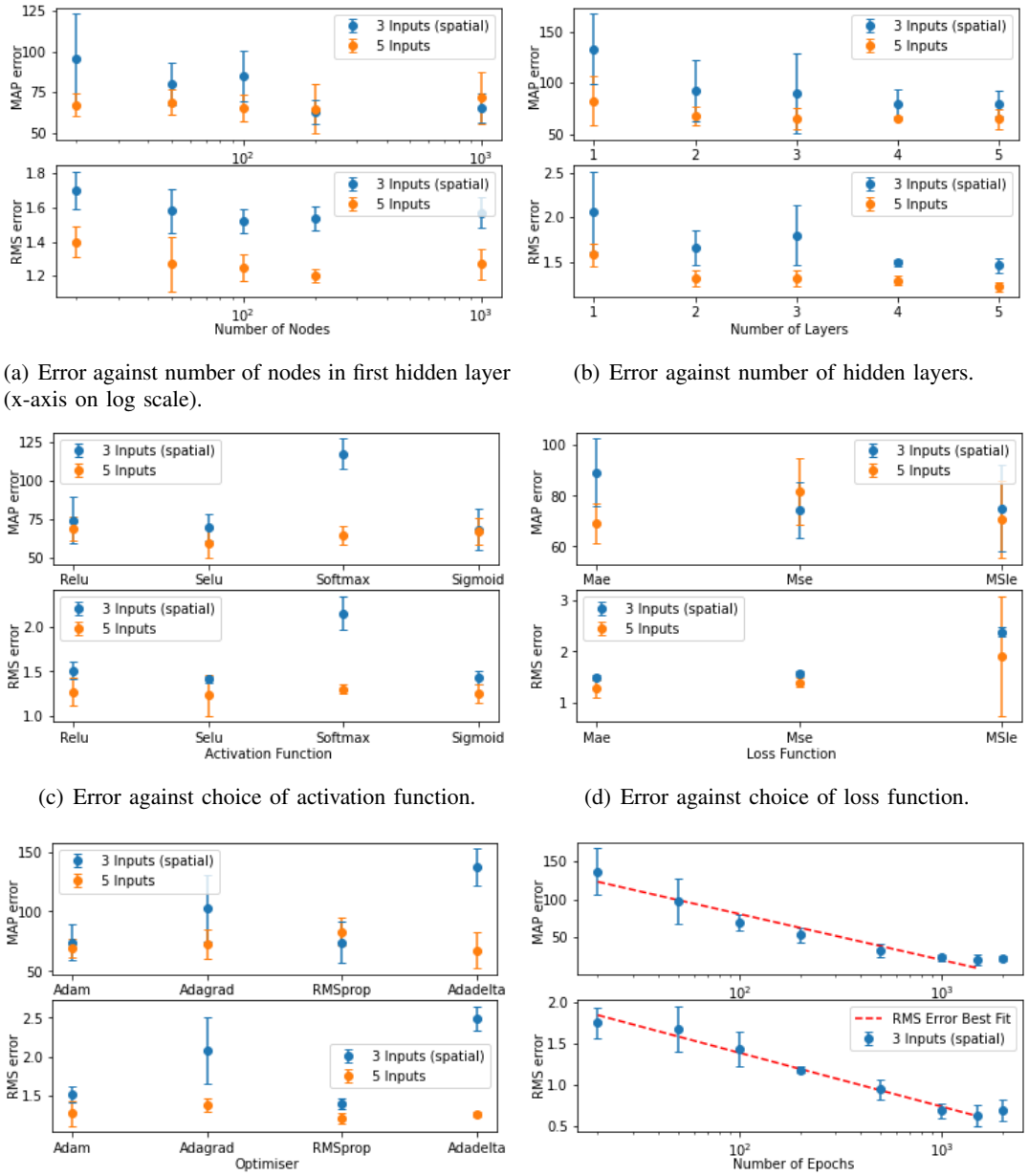
Loss Function: Fig.3(d) We considered MAE (equation 10), MSE (equation 9) and the mean-squared-logarithmic error (MSLE),

$$MSLE = \frac{1}{n}\sum_{i=1}^{n}(log(\hat{y}_i + 1) - log(y_i + 1))^2, \tag{16}$$

where $\hat{y}$ is the predicted value. Using MSLE as a loss function focuses on minimising the percentage error of predictions,[30] this is reflected in the MAPE evaluation of MSLE, Fig. 3(d), though it showed a high degree of variance. Overall, MAE as a loss function performed consistently well in terms of percentage and RMS error, hence this was used.

Optimiser: Fig.3(e) The different choices of optimisers hinted at in section II-B4 are generally of a form similar to equation 11,[34]

$$\theta_{n+1} = \theta_n - \eta \cdot \bar{g_n}, \tag{17}$$



(a) Error against number of nodes in first hidden layer (x-axis on log scale).

(b) Error against number of hidden layers.

(c) Error against choice of activation function.

(d) Error against choice of loss function.

(e) Error against choice of optimisation algorithm.

(f) Error against number of epochs (both axes on log-scales). A line of best fit is also plotted.

Fig. 3. Analysis of ANN performance with regards to different choices of hyperparameters. For every case, the base model mentioned in section III-B was used, whilst only varying the respective hyperparameter. Each combination in (a) - (e) was analysed twice; training with the electron data (all 5 inputs, including electron density and electron temperature), and training on spatial data (3 inputs). The implications of each subplot are discussed in section III-C.

where $\bar{g_n}$ is some algorithmic approximation of the gradient at that point. By performing the same analysis on a combination of the most popular optimisers, we settled on the 'RMSprop' algorithm. This is given by

$$\theta_{n+1} = \theta_n - \frac{\eta}{\sqrt{E[g^2]_n + \epsilon}} g_n, \tag{18}$$

where,

$$E[g^2]_n = 0.9E[g^2]_{n-1} + 0.1g_n^2. \tag{19}$$

Here, ϵ is a smoothing term $\sim \mathcal{O}(10^{-8})$ that avoids division by zero, and $g_n = \nabla_\theta J(\theta_n)$, the gradient at the n^{th} step. Despite its poorer performance with respect to MAPE, RMSprop was chosen given its significantly lower RMS error and small variance.

Number of Epochs: Fig.3(f) Considering the log-scales on the line of best fit, we find that the errors initially decay exponentially with an increasing number of epochs. Past $\sim$ 1500 epochs, the error deviates from

the initial trend and begins to increase. This suggests that training past this point does not benefit the model, and that the ANN begins to overfit. As such, the final hyperparameter of 1500 epochs was chosen.

D. Final Setup & Validations

A final evaluation of skill was performed on our model using the finalised hyperparameters from section III-C. Fig. 4 shows an example result of this process. We

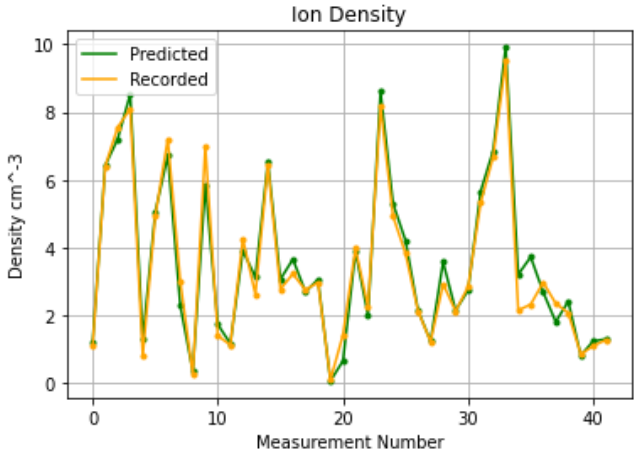


Fig. 4. Example 80/20 train-test split evaluation using the final parameters of our model. The ANN was trained on a randomly chosen 80% set of our input data (with only the 3 spatial inputs). The model's predictions (green) were then tested against the remaining 20% of data (yellow). This was predicted over 5 cycles with a prediction being made each run. The average root-mean-square error of predictions was found to be 0.6215 ± 0.1261 cm^{-3}.

evaluated the performance of the model twice, using 3 and 5 inputs, to investigate the effect that electron data has on the anion density predictions. This is shown in Table I. As expected, the model that trained with the available electron data had lower average errors, agreeing with the production routes suggested by previous studies. We treated this as a key validation for our ANN as it reproduces this electron density dependence.

TABLE I
PERFORMANCE EVALUATIONS USING FINALISED MODEL

	RMSE (cm^{-3})	MAPE(%)
3 Inputs	0.6215 ± 0.1261	21.38 ± 2.74
5 Inputs	0.5404 ± 0.2247	19.21 ± 6.39

Interestingly, despite the lower error, we find that the uncertainty of the ANN performance is higher for 5 inputs, indicating more inconsistent predictions between training cycles. In section II-A2, it was suggested that these anions only displayed electron dependence on the day-side of Titan, with no obvious correlation on the night-side. As such, it is likely that the electron data misled the model in predicting densities situated on the night-side of Titan, whilst improving the predictions on the day-side. The anion densities in the CAPS-ELS data set ranged from 0 to 10 cm^{-3}. Given this, an error of 0.6215 ± 0.1261 cm^{-3} gives us a good measure of confidence in our predictions.

E. Generating Heat Maps

With a trained ANN, we were able to begin making predictions of the low mass anion densities. Note that although our predictions improved with electron density, we were unable to make predictions using a model that used all 5 inputs, as we did not have electron data covering the entire ionosphere. Thus, we were restricted to only training our model on the 3 spatial variables of the data-set.

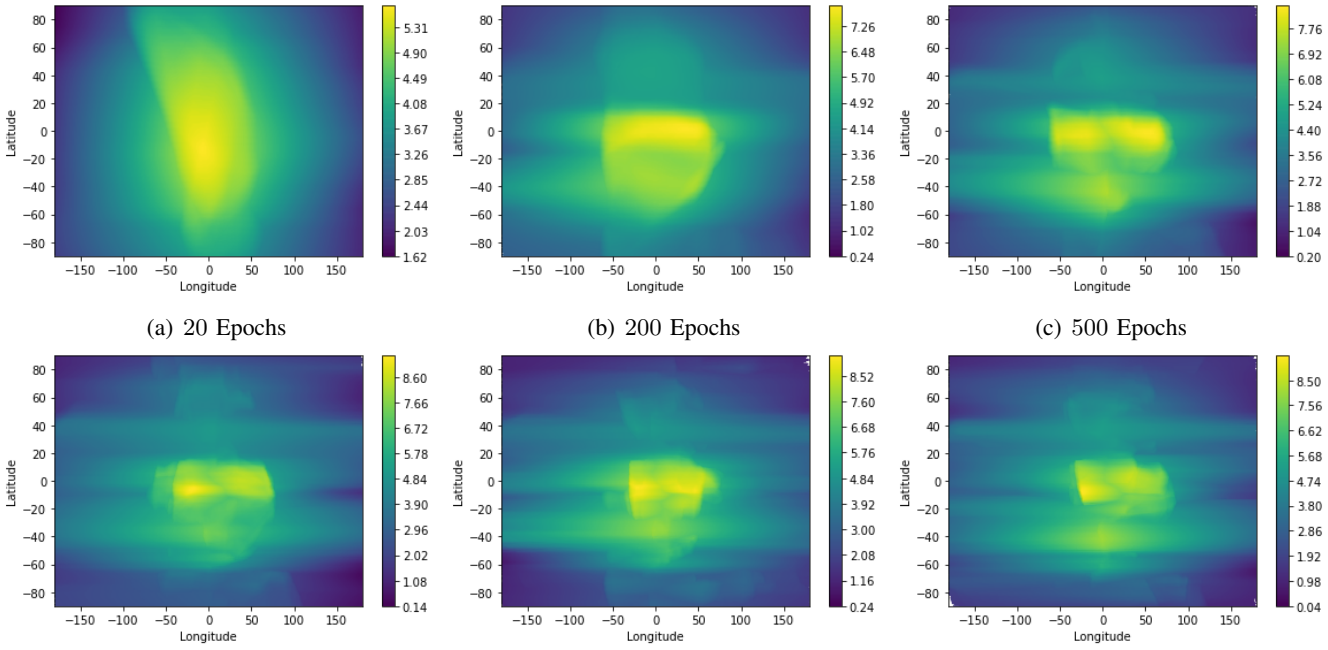
We designed our model such that it could predict two different 2D cross-sections of Titan's ionosphere; one along fixed altitudes (see Fig. 6), and another along fixed longitudes (see Fig. 7). For each case, the algorithm generates a 2-dimensional grid of points corresponding to each coordinate. It then loops through the entire grid to make a prediction at each point in space. Ideally, we would want to make a continuous prediction along the entire 3D atmosphere in spherical polars. However, we were computationally limited to making these predictions at discrete cross-sections and projecting this onto a 2D space. The latitude vs. longitude maps were discretised into 181×361 points, and the altitude vs. latitude maps were discretised into 181×161 points. For the latter plots, the altitude range chosen was $950 - 1350$ km, with a step size of 2.5 km. A visualisation of the results produced at different stages of the training process are shown in Fig. 5.

IV. RESULTS & DISCUSSION

Using our final setup, we trained our model on the full set of CAPS-ELS data, and predicted spatial distribution maps of the low mass anion densities throughout Titan's ionosphere. These predictions were made at various fixed altitudes (Fig. 6) and at different local times of Titan (Fig. 7). Each subplot was produced by making predictions at the fixed cross-section, averaged over 5 repeated training cycles. This was done so that the error evaluations of the model presented in Table I would also apply the these predictions.

A. General Trends

Referring to Figs. 6 & 7, we find that some key features of the CAPS-ELS data discussed in section



(a) 20 Epochs (b) 200 Epochs (c) 500 Epochs

(d) 1000 Epochs (e) 1500 Epochs (f) 2000 Epochs

Fig. 5. Latitude against longitude predictions of anion densities [units of normalised number density ($cm^{-3} \cdot \varepsilon$)] at a fixed altitude of 1000km. This was produced using an increasing number of epochs. The sequence of subplots show that the model continually adds layers of complexity and structure to its predictions the more it trains on the CAPS-ELS data-set. Past $\sim$ 1500 epochs, we expect the ANN to have overfitted based on discussions in section III-C

II-A have been reproduced, namely that of the denser ionospheres near the equator and on the day-side. It also shows that there are much lower densities close to the poles and on the night-side. In a previous study, it was said to be unclear if this same trend carried on to higher altitudes, as they were limited by the coverage of the data-set.[15] Our predictions seemed to confirm this assumption, given the prominent centralised peak shown in Fig. 7(c). This distinct relation readily supports the suggestion that the production routes of the low mass anions have a dependency on photo-ionisation, given the changes in solar flux when moving between these regions.

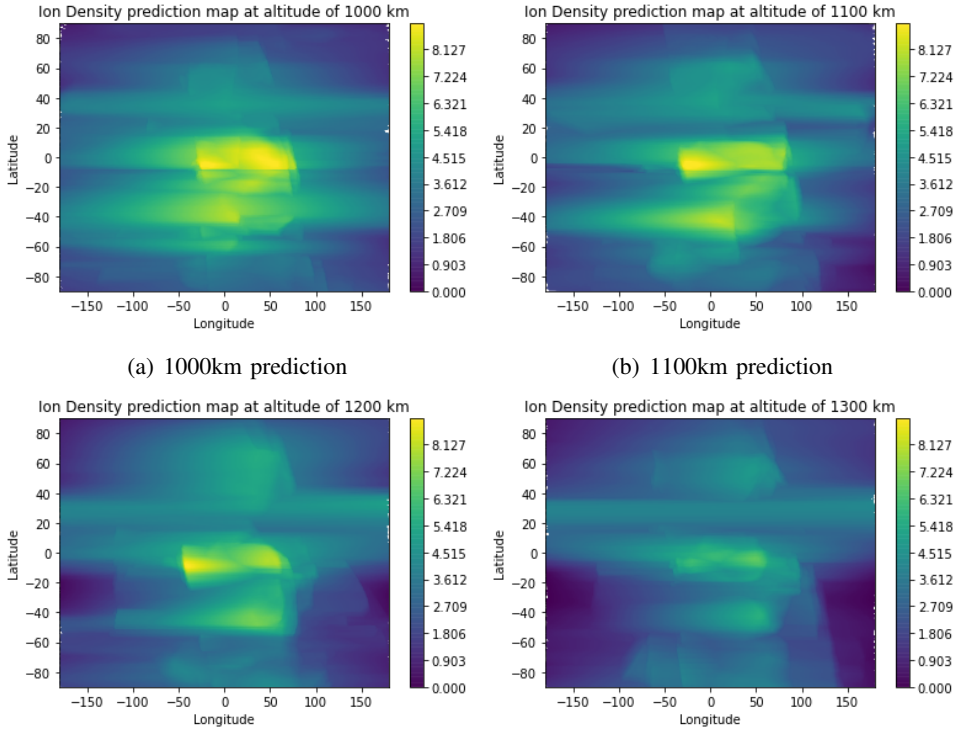
It was previously found that for the the low mass anions, their densities decreased at a slower rate with increasing altitude in comparison to heavier anions.[19] Although the densities in our figures do show that they still gradually decreased with altitude, this helps explain the small variations between altitude layers on the values of the peaks in Fig. 7(a) and (c), which extend up to such high altitudes with a seemingly constant anion density.

It is known that the coverage of the ionosphere from the ELS data is biased towards certain regions; majority of the data was taken within altitudes of 950–1050km, with more readings taken at higher latitudes and at night. We note that this implies that a blanket evaluation of the performance of our ANN cannot truly be reflected within our predictions, i.e. we cannot simply assume that a prediction made at a high altitude on the day-side (where data coverage was sparse) has an estimated error of 0.6215 ± 0.1261 cm^{-3} as quoted in Table I, but rather, it is likely to be much higher. Conversely, predictions in regions of high data coverage are likely to have much lower errors than the value stated above. Nonetheless, these evaluations are still useful as a rough gauge of errors on the intermediate predictions, close to regions with high data coverage.

B. Dusk to Dawn Evolution

The evolution of anion densities from dusk (Fig. 7(a)) to dawn (Fig. 7(b)), displays some asymmetry. Considering they are both equidistant from 0°longitude, as well as the fact that both regions are exposed to similarly attenuated levels of solar flux, one would expect the densities at both of these regions to be similar. However, this seems to not be the case. It is unlikely that this deviation is due to an error in our prediction, as we note that this discrepancy is still present within low altitudes near the night-side, where the data coverage is highest, and hence so is the confidence of our predictions.

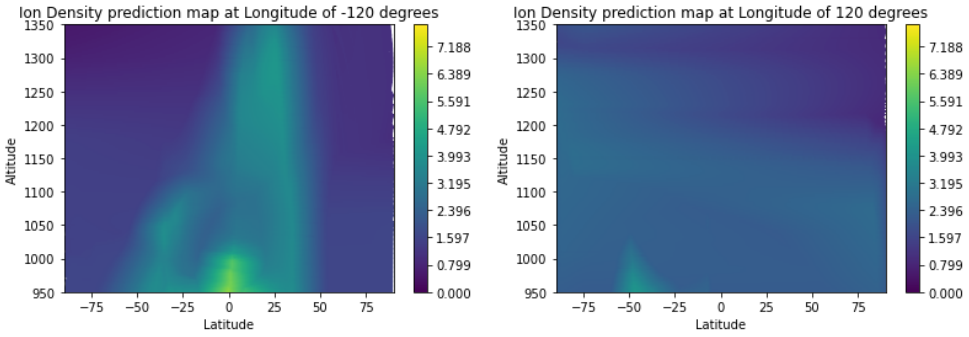
From previous studies, it was suggested that the low mass anions present on the night-side were either locally produced through an additional production mechanism,



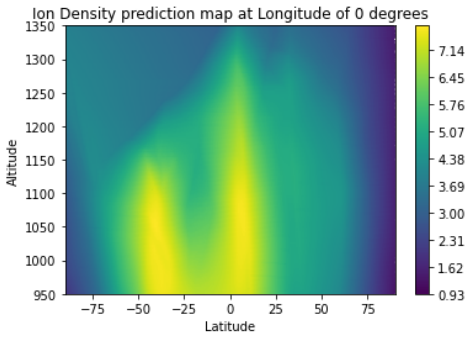
(a) 1000km prediction (b) 1100km prediction

(c) 1200km prediction (d) 1300km prediction

Fig. 6. Latitude against longitude predictions of anion densities [units of normalised number density ($cm^{-3} \cdot \varepsilon$)] at increasing altitudes, from 1000km to 1300km at 100km increments. All subplots exhibit anion densities which are the highest on the day-side (longitude of -90°to 90°), centred just south of the equator, with low-to-zero anion densities deep into the night-side and near the poles. The sequence of subplots show that the anion densities decrease with an increasing altitude. These also show horizontal streaks of intermediate anion densities, along latitudes of 0°to 40°.



(a) Predicted densities at dawn (local time: 04:00) (b) Predicted densities at dusk (local time: 20:00)



(c) Predicted densities at midday (local time: 12:00)

Fig. 7. Altitude against latitude predictions of anion densities [units of normalised number density ($cm^{-3} \cdot \varepsilon$)] at longitudes of -120°, 0°, and 120°; corresponding to dusk, midday, and dawn respectively. Each plot shows its highest densities to be at the lowest altitude of 950km, which diminishes at higher altitudes. The predicted densities at dawn (a), show a faint vertical streak of a constant density along $\sim 35°$ latitude. It also shows presence of anions at lower altitudes, south of the equator. The prediction at midday (c) resembles the densities at dawn (a) but with much higher intensities. It also shows a very distinct splitting of peaks, centred at $\sim -37°$ and $\sim 10°$ latitudes; the former peak diminishes at around 1150km altitude, whilst the latter peak extends throughout the ionosphere. The densities at dusk (b) show no other distinct features apart from a small peak at $\sim -45°$ latitude. It instead has a general 'cloud' of low densities at $\sim 3cm^{-3}$, but this fades at high enough altitudes near the north pole.

or were remnants of the higher anion densities which remained from the day,[2][6][16] such as being transported from day to night.[11][7] With these suggestions, we can speculate how these processes relate to our subplots in Fig. 7.

We know that the peak shape of Fig. 7(c) is mainly credited to the production mechanisms outlined in section II-A1. By drawing similarities from this shape to the one in Fig. 7(a), it is possible that this peak at dawn is also formed from the same photo-ionisation processes (but at much lower quantities), along with any additional production mechanisms that only occur in the night-side. Fig. 7(c) may then be the precursor to the peak shown at 0°longitude, and this would represent the gradual increase in anion density as the region shifts from dawn to midday.

As the ionosphere moves from midday to dawn, higher attenuation of solar flux occurs, which is likely to inhibit the production mechanisms of the low mass anions. The small peak shown at $\sim -45°$ latitude of Fig. 7(b) appears to be a remnant of the left most peak from the day, as shown in Fig. 7(c). It is possible that some of the remaining anions may have dissipated in transporting from day to night. This could also be the cause of the 'cloudy' distributions at this longitude, as it seems that despite there being a smaller peak, the surrounding atmosphere has an overall higher anion density.

C. Splitting of Peaks

An unexpected feature obtained from our predictions was the distinct splitting of peaks in Fig. 7(c). In a previous study[15], it was noticed from the CAPS-ELS data that there were these low *and* high anion densities near the equator on the day-side, especially at higher altitudes. This large variation over such small ranges was found to be unusual and was said to be an error which arose from binning together large altitude ranges. However, our ANN has seemed to pick up on this feature and interpreted it as a real and distinct phenomenon.

It is likely that this is just due to the changes of Titan's ionosphere throughout the different flybys, as these measurements were taken over the span of 8 years (see section III-A). However, following the sequence of events we speculated in section IV-B, we also have reason to believe that this split peak is real, as the latitudinal positions of these peaks line up as they evolve from dusk to dawn.

We also note that these peaks line up with the horizontal streaks of intermediate anion densities shown in Fig. 6. As for now, it is unclear why this splitting of peaks occurs. This could perhaps be credited to particular loss mechanisms that occur on the day-side of Titan, or simply the transport of anions from day to night.

D. Relation to High-Mass Anions

Current findings show that the heaviest anions were detected at night near the north pole in winter conditions, and it is suggested that these low-mass anions aggregate together to form the heavier anions. This could possibly explain the distinct lack of low mass anions in the north pole, shown in Figs. 6 and 7, as the formation of these high-mass anions initiates the loss mechanisms of the CN^-/C_2H^- groups. This could also explain the slight shifts of these peaks towards lower latitudes. However, these exact processes and loss mechanisms are not entirely known.[17] Despite this, there is still strong evidence of a correlation between the low and high mass anions[13] and our predictions appear to be consistent with these suggestions.

E. Further Discussion

So far, we have managed to successfully produce predicted maps of Titan's ionosphere along with a rough estimation of their uncertainties. However, the method we have undertaken, especially in analysing these results, is very vulnerable to a common criticism that can be made about most deep learning approaches. That is, although ANN's have been hugely efficient in picking up key features from the data and filling in gaps between these sparse sets of information, they do not perform much more in explaining the chemistry or reasons behind certain predictions. These neural networks perform very much like 'black-box' models, and attempts to gain further insight into the solutions provided rely heavily on speculation, as we have done. It is always possible to further improve our ANN by taking into account different types of layers and making extra configurations. Even though this is likely to give stronger predictions, it still may not escape the issue at hand.

Another obvious criticism is that the performance of the ANN is always going to be limited by the quality and size of the training-set, hence there may be a theoretical limit to how accurate we can possibly predict Titan's atmosphere using existing measurements. Despite these pitfalls, the use of ANNs in this study have shown its strengths in efficiency, and through its predictions, we were able to more easily visualise the highlighted trends exhibited by the data that were invaluable for analysis and discussion.

V. Conclusion

By building off the works of countless studies conducted on the upper atmosphere of Titan, we were

pushed to explore the viability of using ANNs to accelerate this area of research. In constructing our own neural network, we conducted a systematic search for the optimal hyperparamaters which pertained to our problem. Ultimately, we were successful in using our model to produce the anion density maps of Titan's ionosphere, and have managed to quantify the performance of this model, as shown in Table I. Through this, our predictions managed to reproduce some key trends that have been previously identified as part of Titan's ionosphere. These served as useful validations for our model.

With some prior knowledge on the chemistry within Titan's upper atmosphere, we were able to validate the anion density's dependence on photo-ionisation from our predictions, which has been suggested in previous studies. However, these predictions also displayed some unexpected behaviour, particularly the splitting of density peaks near the equator. This was possibly due to the fact that the CAPS-ELS data-set was recorded over different times and conditions, which may have misled the ANN's training process. Nonetheless, these trends were analysed and possible implications were discussed.

Such an issue highlighted the main pitfalls of relying on an ANN to solve a problem, in that we often get little insight from the results they produce. However, these predictions still proved to be widely useful in the visualisation and emphasis of certain trends, which heavily contributed to our discussions on Titan's ionosphere.

Overall, the use of ANNs in this study has proven to be invaluable in aiding the discussions of Titan's upper atmosphere. Its ability to pick up on, and reproduce key trends in data with arguably low amounts of effort greatly outweighs its few pitfalls. Given this, the next obvious step would be to improve on and generalise this ANN to account for the distributions of higher mass anions. Hopefully, this step would allow us to predict the distributions of other anions throughout the atmosphere, and by feeding this into working chemical model of the ionosphere, we can truly begin predicting the formation of Titan's organic Haze.

VI. Acknowledgements

This study was only possible with the aid of my project partner, under the supervision of Dr R.T. Desai.

Declaration of Work Undertaken

All project work was split evenly between both partners.

References

[1] Jayesh Bapu Ahire. *The Artificial Neural Networks Handbook: Part 4*. https://medium.com/@jayeshbahire/the-artificial-neural-networks-handbook-part-4-d2087d1f583e. 2018.

[2] Agren et al. "On magnetospheric electron impact ionisation and dynamics in Titan's ram-side and polar ionosphere – a Cassini case study". In: *AnGeo* (2007).

[3] Coates et al. "Discovery of heavy negative ions in Titan's ionosphere". In: *GeoRL* (2007).

[4] Coates et al. "Ionospheric photoelectrons: Comparing Venus, Earth, Mars and Titan". In: *P&SS* (2011).

[5] Coustenis et al. "The composition of Titan's stratosphere from Cassini/CIRS mid-infrared spectra". In: *Icar* (2007).

[6] Cravens et al. "Model-data comparisons for Titan's nightside ionosphere". In: *Icar* (2009).

[7] Cui et al. "Ion transport in Titan's upper atmosphere". In: *JGRE* (2010).

[8] Dobrijevic et al. "1D-coupled photochemical model of neutrals, cations and anions in the atmosphere of Titan". In: *Icar* (2016).

[9] Galand et al. "Electron temperature of Titan's sunlit ionosphere". In: *GeoRL* (2006).

[10] Linder et al. "The Cassini CAPS Electron Spectrometer". In: *GMS* (1998).

[11] Müller-Wodarg et al. "Horizontal structures and dynamics of Titan's thermosphere". In: *JGRE* (2018).

[12] Niemann et al. "The abundances of constituents of Titan's atmosphere from the GCMS instrument on the Huygens probe". In: *Nature* (2005).

[13] R. T. Desai et al. "Carbon Chain Anions and the Growth of Complex Organic Molecules in Titan's Ionosphere". In: *The Astrophysical Journal Letters* (July 2017).

[14] Shebanits et al. "Titan's ionosphere: A survey of solar EUV influences". In: *JGRA* (2017).

[15] Teodora Mihailescu et al. "Spatial Variations of Low-mass Negative Ions in Titan's Upper Atmosphere". In: *The Planetary Science Journal* (Sept. 2020).

[16] Vigren et al. "Ionization balance in Titan's nightside ionosphere". In: *Icar* (2015).

[17] Vuitton et al. "Negative ion chemistry in Titan's upper atmosphere". In: *P&SS* (2009).

[18] Waite et al. "The Process of Tholin Formation in Titan's Upper Atmosphere". In: *Sci* (2007).

[19] Wellbrock et al. “Cassini CAPS-ELS observations of negative ions in Titan’s ionosphere: Trends of density with altitude”. In: *GeoRL* (2013).

[20] Wellbrock et al. “Density trends of negative ions at Titan”. In: *AGUFM* (2012).

[21] Wellbrock et al. “Heavy negative ion growth in Titan’s polar winter”. In: *MNRAS* (2019).

[22] Young et al. “Cassini Plasma Spectrometer Investigation”. In: *SSRv* (2004).

[23] Jason Brownlee. *How to Evaluate the Skill of Deep Learning Models*. https://machinelearningmastery.com/evaluate-skill-deep-learning-models/.

[24] F.J. Crary et al. “Heavy ions, temperatures and winds in Titan’s ionosphere: Combined Cassini CAPS and INMS observations”. In: *Planetary and Space Science* 57.14 (2009), pp. 1847–1856. ISSN: 0032-0633. DOI: https://doi.org/10.1016/j.pss.2009.09.006. URL: http://www.sciencedirect.com/science/article/pii/S0032063309002682.

[25] G. Fraser. “The ion detection efficiency of microchannel plates (MCPs)”. In: *IJMSp* (2002).

[26] Vahid Gholami et al. “Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers”. In: *Journal of Hydrology* 529 (Sept. 2015), pp. 1060–1069. DOI: 10.1016/j.jhydrol.2015.09.028.

[27] S. A. Haider. “Some molecular nitrogen emission from Titan-solar EUV interaction”. In: *JGR* (1986).

[28] Casper Hansen. *Activation Functions Explained - GELU, SELU, ELU, ReLU and more*. https://mlfromscratch.com/activation-functions-explained/. 2019.

[29] Paul Dauncey Mark Scott. *Computational Physics Course Notes*. Imperial College London. 2020.

[30] *Mean squared logarithmic error (MSLE)*. https://peltarion.com/knowledge-center/documentation/modeling-view/build-an-ai-model/loss-functions/mean-squared-logarithmic-error-(msle).

[31] Gavril Ognjanovski. *Everything you need to know about Neural Networks and Backpropagation — Machine Learning Easy and Fun*. https://towardsdatascience.com/everything-you-need-to-know-about-neural-networks-and-backpropagation-machine-learning-made-easy-e5285bc2be3a. 2019.

[32] Russell D. Reed. *Neural Smithing: Supervised Learning in Feedforward Artificial Neural*. 1999.

[33] Stacey Ronaghan. *Deep Learning: Which Loss and Activation Functions should I use?* https://towardsdatascience.com/deep-learning-which-loss-and-activation-functions-should-i-use-ac02f1c56aa8. 2018.

[34] Sebastian Ruder. “An overview of gradient descent optimization algorithms”. In: *Insight Centre for Data Analytics* (2016).

[35] & Peko B. Stephen T. “Absolute calibration of a multichannel plate detector for low energy O ions”. In: *RScI* (2000).

[36] Aston Zhang et al. *Dive into Deep Learning*. https://d2l.ai. 2020.