Machine Learning for Clutter Loss Prediction in Aggregate Interference Assessments

Masoud Farshchian*§
The MITRE Corporation
Bedford, Massachusetts

Christopher J. Blackburn*§
The MITRE Corporation
McLean, Virginia

Harris Z. Zebrowitz
The MITRE Corporation
McLean, Virginia

Amy Baker
The MITRE Corporation
Bedford, Massachusetts

Anthony Rennier
Foundry Corporation
Severn, Maryland

Michael W. Bowman
Defense Spectrum Organization DISA
Annapolis, MD

***Abstract*—The Advanced Wireless Services (AWS-3) Spectrum Sharing Test and Demonstration (SSTD) Program has been investigating refinements to propagation models within the context of AWS-3 licensee coordination requests. A variety of methods have been developed to derive enhanced clutter models based on data from multiple sources, including but not limited to field measurements, empirical models, and simulations. Building on previous work, we use machine learning methods to create computationally tractable clutter loss models for the purpose of spectrum sharing and improving the methodology for prediction of interference into Department of Defense (DoD) incumbent receivers. We introduce three machine learning techniques for predicting clutter loss, a key ingredient in aggregate interference modeling. These machine learning methods process available nation-wide topographical and morphological data as part of their own input features. We show that they outperform existing approaches for predicting clutter loss by a non-trivial margin. Our results indicate machine learning based approaches could offer a high-fidelity standard for predicting clutter loss in computationally tractable evaluation of deployment scenarios.**

***Index Terms*—Machine Learning, Statistical Modeling, Aggregate Interference, Advanced Wireless Services 3 (AWS-3), Spectrum Sharing, Long Term Evolution (LTE), Propagation, Clutter Loss**

I. INTRODUCTION

For policymakers in the United States, solving the spectrum sharing challenge is pivotal for bolstering national security and economic competitiveness. Currently, there is need for innovative spectrum sharing methods so that commercial operators in the United States can access desirable mid-band frequencies. A key implementation challenge is how to ensure spectrum sharing activities do not interfere with the continuity of United States national security operations. Rising to this challenge, Defense Information Systems Agency (DISA) Defense Spectrum Organization (DSO) established the Spectrum Sharing Test and Demonstration (SSTD) Program to facilitate and operationalize spectrum sharing between commercial deployments and DoD incumbent receivers [5].

The SSTD Program is responsible for recommending improvements in several key areas to ensure spectrum sharing activities. One high priority area for the SSTD Program is to develop new computationally tractable models to assess aggregate interference into DoD systems caused by commercial deployments. In this area, the SSTD Program has introduced computational advancements to the input parameters used in a standard link budget equation employed to assess how various factors influence aggregate interference at DoD incumbent receivers. Formally, the link budget equation for aggregate interference is expressed as [5]:

$$E\left(\iota_{total}\right) = \sum_{i=1}^{N} E\left(F_i\right) E(l_{cl,i}^{-1}) \quad (1)$$

where $E\left(\cdot\right)$ is the expectation operator, i indexes interference sources (e.g., User Equipments (UEs)), F_i is a function that groups multiple factors influencing interference, and $l_{cl,i}$ is the clutter loss factor.[1]

Figure 1 illustrates a sector where aggregate interference from transmitter handsets (UEs) impact a DoD incumbent receiver. The clutter loss factor of the link budget equation is the additional propagation loss due to foliage and buildings. Generally, the clutter loss distribution (CLD), which is the distribution of the clutter loss values in a sector, depends on many factors, including but not limited to operating frequency, transmitter height, and the structure of the scene. The complex mechanisms behind clutter loss statistics–such as reflections, diffractions, refractions, scattering, free space loss–are usually

§Masoud and Christopher were equally responsible for designing the algorithms and drafting the text.

[1]To be precise, we simplify the link budget equation to highlight our emphasis on the clutter loss factor. Dropping the subscript for notational simplicity, we define

$$F = \frac{nl \times g_r(\theta,\phi) \times eirp(P_{tx})}{l_p \times l_{pol} \times l_s \times fdr(\Delta f)}$$

where nl is the load factor, $eirp$ is the modeled UE transmitter effective isotropic radiated power in mW, l_p is the interference path propagation loss factor between a modeled UE and DoD incumbent receiver,$fdr(\Delta f)$ is the frequency dependent rejection factor of the victim DoD incumbent receiver, $g_r(\theta,\phi)$ is the DoD antenna incumbent receiver gain factor in the direction of the transmitted, l_{pol} is the DoD incumbent receiver antenna polarization mismatch loss factor, and l_s is the DoD incumbent receiver system loss factor. Furthermore, unlike previous work, we separate the clutter loss factor from the path loss factor.

modeled using empirical techniques or electromagnetic simulation.

Both empirical and simulation-based approaches for obtaining the clutter loss statistics have their advantages and disadvantages. Empirical models, for example, are easier to implement and more robust to an environment's details but are less accurate than electromagnetic simulation models [2]. In contrast, electromagnetic simulations, such as deterministic ray-tracing models, are more accurate but require significant computational resources and detailed environmental data to execute, both of which may be limited in real-world evaluation of deployment scenarios. Safe and reliable spectrum sharing necessitates computationally tractable aggregate interference assessments, and accurate clutter loss prediction. Given these limitations, new computationally tractable methods are needed for accurately estimating clutter loss of aggregate interference assessments.

The link budget equation sets the stage for using statistical methods to predict aggregate interference into DoD incumbent receivers within computationally tractable evaluation of deployment scenarios. Departing from the traditional empirical and simulation-based methods, we developed computationally tractable machine learning methods to predict the CLD, and also the equivalent clutter factor $E(l_{cl}^{-1})$, in the link budget equation. The inverse (negative) of the equivalent clutter factor is the equivalent clutter loss (ECL) when ECL is represented in linear (dB) units. The ECL represents the expected impact of the clutter to aggregate interference into DoD systems.

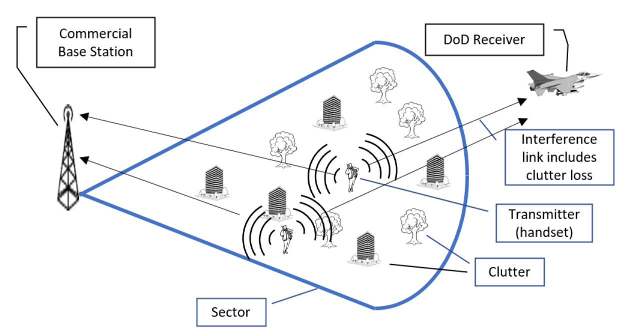


Fig. 1: Diagram illustrating aggregate interference of a sector and its impact on a DoD incumbent receiver. The clutter loss factor is due to the additional propagation loss from foliage and buildings.

We introduce new machine learning based approaches for predicting clutter loss statistics, and specifically, we develop three alternative methodologies for the community to select from based on their own deployment evaluation criteria. These three methods use available nation-wide morphological and topographical features. Our first method uses a combination of supervised and unsupervised learning techniques to derive a new category model to predict ECL for new sites. Previous work in this area has utilized subject matter expertise to handcraft decision trees for estimating clutter loss for new site interference assessments [3]. These handcrafted decision trees group sites according to their morphological details, where each group is referred to as a *category*. Each category represents a different spatial morphology, ranging from dense urban environments to barren landscapes. The handcrafted categories are used to construct CLDs which are used to estimate clutter loss for any sites that belong to the category.

Our category model which is denoted as neural network Gaussian mixture models (NN-GMM) is inspired by recent advances in using deep neural networks for learning feature transformations that are amenable for clustering analysis [8]. Specifically, we train a neural network to learn a feature transformation that projects our inputs into a lower dimensional space that is highly correlated with our target value–ECL. Once we discover this feature transformation, we project our inputs using the encoder portion of the network [6] and cluster the projection using Gaussian Mixture Modeling [4] to discover new categories among our transformed features. We show our approach outperforms the benchmark, handcrafted category models.

Our second and third methods both rely on deep neural networks (DNN) to predict ECL and the CLD.[2] Our second method (NN-ECL) uses the trained DNN from our category model to predict ECL at the site-level. We demonstrate this method tends to outperform our other DNN-based methods for predicting ECL at the site-level. Our third approach (NN-CL) is designed to estimate the clutter loss distribution at the site-level. Unlike the other approaches, this method starts by building a network to predict the clutter loss factor's distribution, rather than the mean value directly. Using the estimated clutter loss distributions, we can compute the ECL value by transforming the distributions into a linear transmission gain and taking the mean of the transformed distribution. Despite being lightly optimized, we find both DNN methods outperform the benchmark category models, including our encoder-based clustering algorithm, by a non-trivial margin.

A comparison of performance across the models shows that the machine learning methods result in meaningful improvements over the benchmark category models. Our machine learning based category model, NN-GMM, reduces the prediction error by 31% relative to the benchmark category models. We also demonstrate our DNN regression methods, NN-ECL and NN-CL, can reduce prediction errors between 51-84% relative to the handcrafted category model. Further, the machine learning models result in fewer outliers in the error distribution, implying these models are less likely to generate non-trivial mistakes in computationally tractable aggregate interference assessments. Our initial results suggest computationally efficient machine learning methods can offer high fidelity predictions of aggregate interference assessments.

The organization of the paper is as follows. In Section II, we review the prior work on data-driven and machine learning clutter loss modeling techniques. Section III provides a background on the data used in the analysis. We survey

[2]In many respects, the DNNs used in both methods are lightly fine-tuned and can be reasonably considered as "off-the-shelf".

the data generating process and a subset of features used in the analysis. Section IV discusses three new machine learning approaches for clutter loss prediction. The performance results of these new techniques and the base-line techniques are presented in Section V. Section VI concludes the paper with future avenues of research.

II. Prior Work

In this section, we briefly review some of the p rior work in the area of machine learning and clutter loss prediction. We start by reviewing various studies that also use machine learning methods for predicting clutter loss. We close the section by presenting a detailed review of the CASY-19 and CASY-21 [3] category models used as benchmarks in our study. We elect to use these category models as benchmarks due to their deployment on the same data used within our study.

We briefly note some of the relevant machine learning techniques that have been used for path-loss modeling, and we recommend [10] for an extensive review of work in this area. One of the earliest results of using machine learning [1], was an improved path-loss prediction model for urban environments based on a radial basis function neural network in which the height of the receiving antenna and the transmitting antenna and the distance between are used as inputs.In [15], a neuro-fuzzy inference system and a generalized regression neural network were used to predict path-loss of base stations with significant prediction accuracy relative to the empirical Hata [2] and Walfisch-Ikegami [12] models. Some of the features utilized were results from other base-stations, distance, frequency, and transmitter height. The authors of [11] proposed a semi-empirical artificial neural network predictor using the Walfisch-Ikegami empirical model as an expert input to predict base-station path-loss. The artificial neural network resulted in a significant improvement for path-loss predictions for the 800 MHz to 1800 MHz bands. In [13], the authors proposed a machine learning framework for modeling path-loss using a combination of three key techniques: artificial neural networks, Gaussian process, and principle component analysis (PCA). The artificial neural network, w hich outperformed empirical models, was optimized to perform both path-loss prediction and shadowing prediction for the 450, 1450 and 2300 MHz bands. The authors in [14] presented FadeNet which can predict large-scale fading from a base station to each location in its coverage area based on convolutional neural networks. FadeNet uses height information of buildings, foliage, terrain, etc., of a target site as input, to generate the large-scale fading prediction as an output. FadeNet was trained using labeled data obtained via computationally-expensive deterministic ray-tracing. These results suggest that combining expert-features, morphological and geological features can help improve the estimation of propagation loss factors.

Prior SSTD program work on data-driven clutter loss estimation involves using handcrafted decision trees for creating category models [3]. These category models, coined CASY-19 and CASY-21 (jointly referred to as CASY), are designed to categorize a large number of site locations into one of seven clutter morphology categories (Dense-Urban, Urban, Suburban-Forested, Suburban, Barren, Rural, Rural-Forested). The first step in the CASY-19 clutter category model is to extract the land cover statistics from the US Geological Survey's National Land Cover Database (NLCD) within a radius of 300 meters of a site location.

Note that the NLCD provides the type of land cover data, not the clutter height information. The extracted NLCD land cover statistics for the sector location are then fed into the CASY decision tree algorithm, which will determine the corresponding CASY clutter morphology. In the CASY algorithm, the percentage of an NLCD value that exists within a radius of 300 m of a sector location is used as input feature. The CASY-19 and CASY-21 algorithms are described in [3].

The CASY-21 clutter category model is based on the CASY-19 algorithm with further refinement based on the clutter height profile of a sector. The first step is to run the CASY-19 algorithm to determine the CASY-19 clutter categories for a given sector location. The second step is to calculate the average clutter height for a given sector location coordination request (CR) using the 5 meter resolution nationwide Digital Surface Model (DSM) and Digital Terrain Model (DTM) data. The average clutter height for a sector is computed by taking the average of the difference between the DSM pixel heights subtracted by the DTM pixel height for a 300 meter sector radius. The average clutter height is then used to refine all CASY-19 categories with the exception of barren into subcategories of "flat" and "non-flat". For each of the CASY-19 categories (except for barren), there is an associated unique threshold which was calculated based on the category type. If the average clutter is above the threshold, the sector is assigned as a flat subcategory (e.g., Urban-Flat). If it is below the threshold, the sector is associated with the non-flat subcategory (e.g., Urban Non-Flat). Consequently, the CASY-21 algorithm has 13 categories and provides further improvement for ECL prediction relative to the CASY-19 algorithm as shown below.

III. Data

The clutter statistics (CLDs and their associated ECL values) generated for use in this analysis were computed using the Terrain Integrated Rough Earth Model (TIREM) [16] plus Light Detection and Ranging (LiDAR) technique. These statistics were obtained from 4,052 tower locations for different morphologies including Dense-Urban, Urban, Suburban, Suburban-Forested, Rural, Rural-Forested and Barren. The UEs were placed outdoors for the generation of this set of data. Clutter loss statistics for these morphologies were calculated by subtracting the LiDAR layer with clutter path-loss from the bare earth layer path-loss for each path. The measurements of the CLDs were generated at a frequency of 1755 MHz, and at elevation angles of 1.5, 3, 5, 7, 10, 20, 30, 50 and 70 degrees, respectively. In order to obtain the CLDs used in our data, 200 UEs were generated inside the 300 meter radius at random locations, and clutter loss was measured for all the

200 UEs at at 16 radials. Figure 2 illustrates an example of clutter path rays emanating from UEs in a dense urban setting whose values are measured at 16 radials that are 2km from a cell center with a radius of 300 meters.

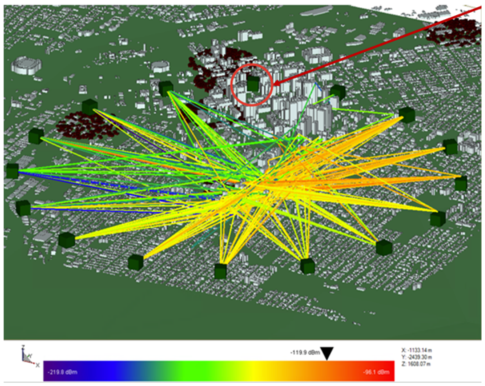


Fig. 2: An illustrative example of clutter loss statistics measured at 16 different angles in an urban environment

In order to perform ECL and clutter loss distribution prediction, a rich dataset of 114 obtainable features including the clutter height distribution, the percent of azimuth paths blocked at 10 degrees interval, elevation angles, and National Land Cover Database (NLCD) pixel percentage of a defined Lat/Long area were provided. Additional features on NLCD percentage over clutter heights of 1.5, 3 and 5 meters were also provided. Height information such as the average clutter height and clutter height standard deviation (std) of the sector, and the probability of clutter height above 1.5, 3 and 5 meters are also provided. The ECL and CLD target variables were calculated using the aforementioned electromagnetic modeling tools. The goal of the ML techniques is to find useful and hidden correlations between the training input features in order to learn new category models, and accurately predict ECL and CLD.

IV. Machine Learning Methods for Clutter Loss

In this section, we describe the details behind our three machine learning based methods for predicting clutter loss. For model training and evaluation, we utilize 70 percent of the data for training and fine-tuning and reserve 30 percent of the data for testing. All results presented herein are based on performance on the test dataset.

For optimizing the DNNs, we use the adaptive momentum (Adam) optimizer first introduced by Kingma [7]. Unlike regular stochastic gradient descent which uses a fixed learning rate, the Adam optimizer uses adaptive learning rates where the learning rate of each parameter is adapted in each iteration by scaling the initial step-size by exponential moving averages of the first and second moments of the gradient. We use a step-size of 0.001, a value of 0.9 for the decay rate of the first moment estimates (β_1), and 0.99 for the decay rate of the second moment estimates (β_2). For each DNN, we minimize the mean-squared error loss function as the optimization criteria, and we train the network using a batch size of 32 across 500 training cycles (epochs).

A. A DNN for Predicting Equivalent Clutter Loss

Our first DNN architecture (NN-ECL) is designed to predict site-level ECL values as the primary output of the model. As discussed below, a secondary use of the NN-ECL architecture is for discovering non-linear transformations of our feature set that are more correlated with site-level ECL values. The NN-ECL architecture is lightly fine-tuned based on performance on the training dataset, and our architectural choices can likely be improved on through automated, hyperparameter optimization algorithms. In addition to the input layer and prediction head, the NN-ECL architecture consists of 3 hidden layers with each layer size set to 150, 50, and 10, respectively. We use the Rectified Linear Unit (ReLU) activation function for each hidden layer, and a linear activation function for the prediction. We visualize the NN-ECL architecture in Figure 3. The loss function for the NN-ECL is given as

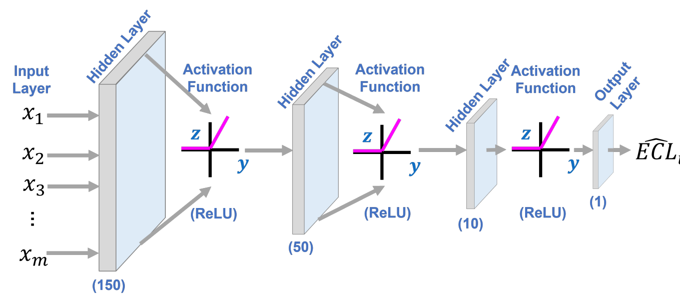


Fig. 3: DNN Architecture for the NN-ECL model. The model consists of three hidden layers in addition to the input (pass through) layer and the prediction head. Each hidden layer is transformed using the ReLU activation function. The output of the NN-ECL is site-level ECL predictions.

$$\mathcal{L} = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \left(ECL_i - E\hat{C}L_i \right)^2$$

where i corresponds to a site-level measurement and N_{train} is the number of measurements in the training dataset.

B. A Clustering Algorithm for Learning Representative ECL Categories

Our second approach is inspired by previous category modeling approaches, namely CASY-19 and CASY-21, where learned categories generate distributions of ECL values for new site interference assessments. The objective of category modeling is to learn a set of K categories where each category represents a latent partitioning of the feature space $\mathbf{X}$ such that each partition k generates a distribution F_k of ECL values. Once the latent categories are discovered, the mean of the distribution F_k is used as a prediction for new sites that fall

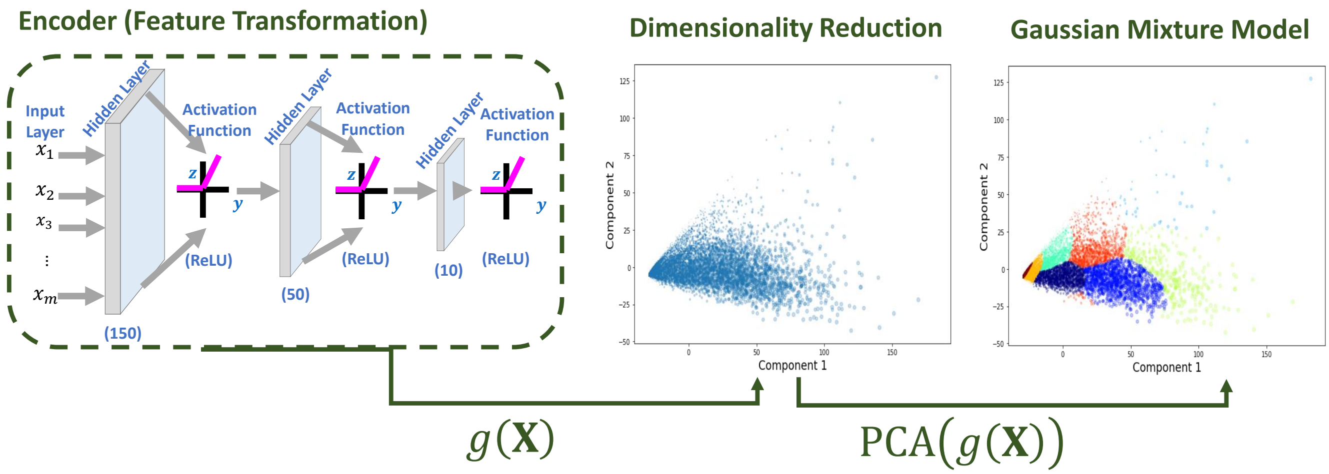


Fig. 4: Algorithm for the NN-GMM model. The algorithm consists of learning the encoder portion of the NN-ECL architecture, using the learned encoder to project the features into a lower dimensional space that is more correlated with ECL, and clustering the features using GMM to obtain representative categories.

into the same category. Formally, we express the category loss function as follows

$$\mathcal{L} = \frac{1}{K} \sum_{k=1}^{K} \sum_{i=1}^{N_k} \frac{|ECL_{ik} - E\hat{C}L_k|}{N_k} \tag{2}$$

where the inner term represents the Mean Absolute Error (MAE) within a category, and the mean is taken across all categories to obtain a representative MAE for the category model.

The task of category modeling is to minimize the loss expressed in (2). However, unsupervised clustering algorithms are not designed to minimize this function directly, and in most cases, the clustering algorithm needs to be adapted to fit the category modeling task. To adapt clustering algorithms for the category modeling task, we utilize the *encoder* portion of the NN-ECL model to transform the inputs into a space that is more correlated with site-level ECL values. The learned feature transformation $g(\mathbf{X})$ is then used as an input into standard clustering algorithms. Here we apply GMM after using a PCA reduction of the learned feature transform features. This reduced data feature is denoted by V where $V = PCA(g(\mathbf{X}))$ and V has a size of $N \times M$ where N is the number of observed data and M is the number of principal components retained from the PCA algorithm. We denote by V_n as the nth transform sample of our data from the N observed dataset. The GMM model is described in more detail in [9], and here we provide a brief summary. The GMM model can be written as a linear sum of K Gaussian variables, and has the probability density form of:

$$p(V|\Theta) = \sum_{k=1}^{K} \pi_k \mathcal{N}(V|\mu_k, \Sigma_k). \tag{3}$$

For unsupervised learning, K is the number of clusters used to fit the GMM probability density function into the transformed featured data. It should be noted that GMMs are

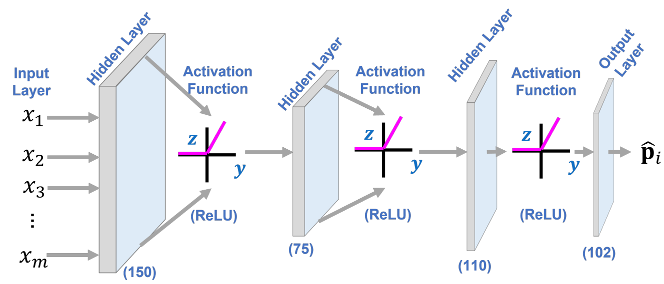


Fig. 5: Architecture for NN-CL model. The NN-CL model has three hidden layers plus a prediction head (output layer). The output layer is transformed using the softmax activation function to ensure the estimates conform with the restrictions on probability distributions.

considered to be a universal approximator for any continuous probabilistic distribution [9]. Here Θ represents the complete set of mixing parameters which are the mixing coefficients π_k, means μ_k and covariance matrices Σ_k of the kth component of K.

Considering the whole matrix data V, assuming that the data points are drawn independently from the distribution, we can express the likelihood function of Equation (3) as:

$$\ln p\,(V|\pi, \mu, \Sigma) = \sum_{n=1}^{N} \ln \sum_{k=1}^{K} \pi_k \mathcal{N}(x_n|\mu_k, \Sigma_k). \tag{4}$$

We use the Expectation Maximization algorithm to maximize this function [9].

C. A DNN for Predicting Clutter Loss Distributions

Our second DNN architecture (NN-CL) is designed to predict site-level clutter loss distributions. The architecture is illustrated in Figure 5. The network is architecturally similar to the NN-ECL model where we use 3 hidden layers with sizes of 150, 75, and 110. Each hidden layer is again transformed using ReLU activation functions. However, the most important difference is the output of the NN-CL model is an array of probability estimates, denoted as $\hat{\mathbf{p}}_i$, which correspond to the estimated site-level clutter loss distribution. Each element of the array $\hat{p}_{ib} = \hat{p}(ECL_{ib} = ECL_b)$ corresponds to the discretized, probability density function estimates, where b represents a discretized bin ranging from -1 to 100 dB in one dB increments.

Since we are estimating a probability distribution, we force the NN-CL model to ensure the conditions $0 \leq \hat{p}_{ib} \leq 1$ and $\sum_b \hat{p}_{ib} = 1$ are met. To force these restrictions on the model, we use the softmax activation function on the prediction head such that each binned probability estimate is constructed as follows

$$\hat{p}_{ib} = \frac{\exp(x_{ib})}{\sum_b \exp(x_{ib})}$$

where x_i corresponds to raw output of the final layer in the neural network. The softmax activation is applied to this raw output to ensure each estimate meets the restrictions described above. Lastly, we optimize the NN-CL using the follow loss function

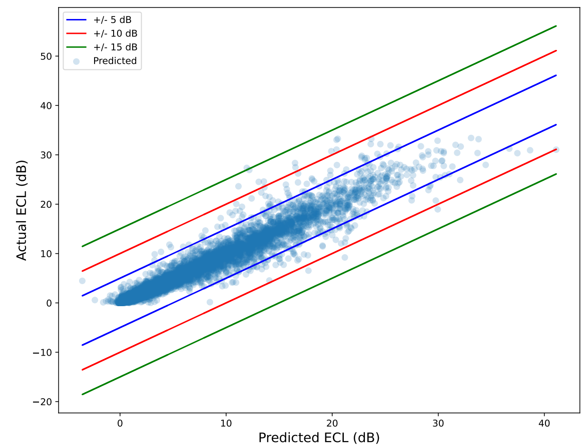
$$\mathcal{L} = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \sum_{b=-1}^{100} (p_{ib} - \hat{p}_{ib})^2$$

V. Results

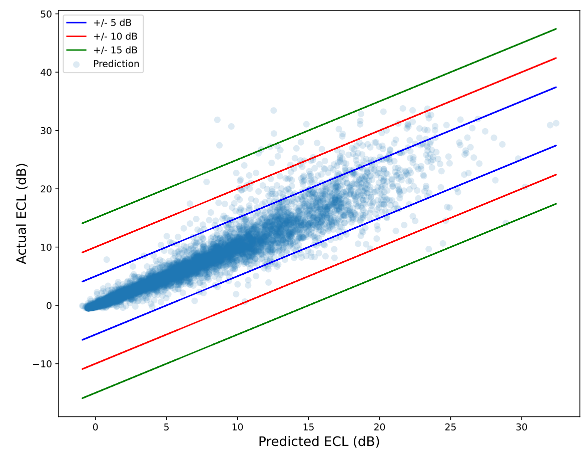
In this section, we present the results from our machine learning methods, demonstrating their out-of-sample prediction performance and comparing them against the benchmark category modeling algorithms, CASY-19 and CASY-21.

A. DNN Performance

Figure 6 visualizes the out-of-sample performance of the trained NN-ECL and NN-CL models. Each model's ECL predictions are compared to the actual ECL values at the site-level within the test dataset. We find the MAE for both models to be less than 1 dB on average, with the NN-ECL model (0.69 dB) slightly outperforming the NN-CL model (0.83 dB). We also evaluate performance at lower elevation angles, namely at the 1.5-degree elevation angle where clutter loss is most severe, and find the MAE from the NN-ECL model (2.23 dB) slightly outperforms the NN-CL model (2.55 dB). However, we also use a weighted MAE (wMAE) that accounts for how lower ECL values are more detrimental to DoD systems in deployment.[3] Using the wMAE, we find the NN-CL model outperforms NN-ECL at the 1.5-degree elevation angle, where the wMAEs are given by 1.36 dB and 1.66 dB, respectively.



(a) NN-ECL Test Performance



(b) NN-CL Test Performance

Fig. 6: Out-of-sample ECL Predictions from the DNN models

Lastly, we provide additional visualizations to summarize the out-of-sample performance of the NN-CL clutter loss distribution estimates in Figures 7 and 8. In Figure 7, the y-axis corresponds to the observed probability densities (p_{ib}) and the x-axis corresponds to the predicted probabilities ($\hat{p}_{ib}$) across all sites in our test dataset. We provide a 45-degree line for reference. As an example, Figure 8 compares the actual clutter loss distribution versus the clutter loss distribution predicted by the NN-CL model for a random site in our data. We note the NN-CL model's performance at the spike at 0 dB in the clutter loss distribution.

B. Feature Transformation using NN-ECL

As discussed above, we use the NN-ECL as an encoder to learn improved representations of our features to align GMM clustering with the task represented in equation 2. To

[3]We use a weighting formula that penalizes errors that occur in the lower range of ECL values. Specifically, the weights are given by

$$w_i = 10^{\frac{-ECL_i}{10}} / \sum_i 10^{\frac{-ECL_i}{10}}$$

and thus, the wMAE is

$$wMAE = \frac{1}{N} \sum_i w_i |ECL_i - \hat{ECL}_i|$$

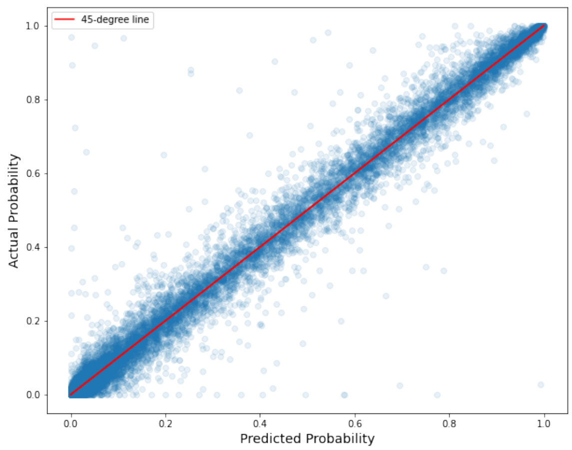


Fig. 7: NN-CL Clutter Loss Distribution Estimates

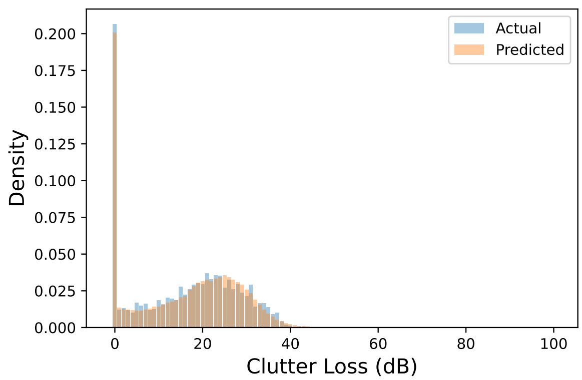


Fig. 8: NN-CL Clutter Loss Distribution Example Comparison

TABLE I: Correlation Comparison

Raw Inputs		Encoder	
Feature	Correlation	Feature	Correlation
% Azimuth Blocked	0.73	Feature 4	0.90
Elevation Angle	0.58	Feature 8	0.87
Height % >1.5 m	0.30	Feature 2	0.78

demonstrate this, we compare the top 3 features correlated with ECL in Table I. We show both the raw inputs and compare them to the transformed features using the NN-ECL encoder. We note the encoder-transformed features are computed using features in the test dataset. Our results show the DNN encoder approach greatly improves the correlation between our feature set and the target value, ECL.

C. Category Model Comparison

The improved correlation introduced by the feature transformation step suggests clustering algorithms may improve on existing category models with respect to the category loss function introduced in equation 2. To demonstrate this, we compare the benchmark category algorithms with the performance of the NN-GMM algorithm

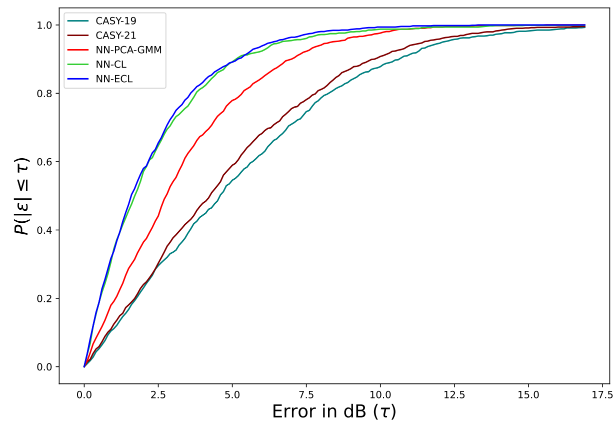


Fig. 9: Error CDF for Comparing Performance

using the standard deviation of the ECL category distributions. We retained the first three principle components from PCA and used ten clusters (categories) for the GMM. In particular, Figure 10 shows the percent of categories with a standard deviation below a fixed dB threshold for the training data set for the CASY-19, CASY-21 and the NN-GMM algorithm. We can see that all categories of the NN-GMM have a standard deviation (STD) which is less than 4.75 dB while for the CASY and CASY-21, all categories have a standard deviation which is less than 8.75 dB. This shows that the NN-GMM produces categories that are less dispersed around the mean, and thus lead to a smaller prediction error when using the category prediction approach.

D. Performance Comparison across Models

The objectives of the algorithms reviewed in this paper are fundamentally the same. That is, they are all designed to predict ECL for new site interference assessments. Given the objectives are the same, we can compare each algorithm directly using standard loss metrics, such as MAE and wMAE, to determine what algorithms yield the highest fidelity predictions. In Figure 9, we plot the cumulative distribution function (CDF) of the (absolute) prediction error on the test dataset using all models. We note the plot is restricted to include only predictions for 1.5-degree elevation angle measurements. In the figure, better model performance is associated with CDFs shifted leftward, implying a higher fraction of sites have smaller prediction errors. Overall, the figure shows the extent to which the DNN-based approaches, including the NN-GMM category, outperform the benchmark algorithms.

We summarize each algorithm's performance across a range of metrics in Table II. In particular, we compare each algorithm's performance across MAE, wMAE, and two metrics that provide some additional details about the error distribution. Within all elevation angles, the DNN models, NN-CL and NN-ECL outperform the category models across each metric.

TABLE II: Model Comparison

	All Elevation Angles			1.5 Degree Elevation Angle		
	MAE (wMAE)	% ≤ 3 dB	% ≥ 7 dB	MAE (wMAE)	% ≤ 3 dB	% ≥ 7 dB
CASY-19	4.36 (4.39)	37.55	12.27	5.24 (6.54)	33.51	29.04
CASY-21	4.28 (4.24)	39.78	12.66	4.84 (5.01)	37.72	24.47
NN-GMM	–	–	–	3.33 (3.33)	54.16	9.99
NN-CL	0.91 (0.55)	94.65	0.87	2.35 (1.54)	71.95	4.56
NN-ECL	0.68 (0.34)	96.01	0.46	2.23 (1.81)	73.53	3.59

Notes: All values are estimated using the test dataset. We note the NN-GMM model is only trained on the 1.5 degree elevation angle.

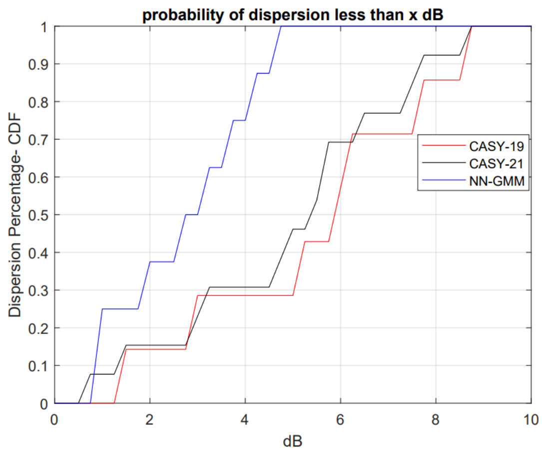


Fig. 10: Percentage of Categories with std below a fixed (dB) threshold

VI. CONCLUSION

Under the AWS-3 SSTD Program, we developed three new ML methods that improve upon the CASY-19 and CASY-21 algorithms while using the CASY-19 and CASY-21 as a subset of their feature set. Reducing the errors in the clutter loss model improves the performance of the aggregate interference model which enables greater spectrum sharing in the AWS-3 band. The NN-GMM method category method is based on a combined supervised-unsupervised learning method and its categories are more tightly dispersed than the CASY-19 and CASY-21 algorithm. It also outperforms those algorithms with respect to MAE and wMAE. The NN-CL and NN-ECL which are single and multi-output supervised learning regression methods were also presented. These methods vastly improve upon category models in terms of their ECL predictive capability. The NN-CL method is shown to have the best results in terms of both wMAE while also producing the actual CLD of a sector. This suggests that it can be possible to predict the CLD of the majority of sectors with a fair degree of accuracy that is comparable to the more time-consuming EM software tools and real-world measurements. This increase in accuracy can lead to better protection of the DoD incumbent receivers while allowing more licensees in the AWS-3 band. Further research and development avenues include optimization of the neural network, having a weighted cost function, feature reduction to remove unnecessary features for the deployment scenarios, training on additional topographical features (e.g., two-dimensional clutter height and NLCD profiles), crossvalidation on wider simulation and measurement datasets, and improving the fidelity of electromagnetic modeling used to generate the training and test data. Additionally, further research of applying the CLD and channel loss factor estimates to different applications can be explored. Some of these applications could include improving wireless communications capacity, improving radar detection algorithms and adaptive waveform design.

REFERENCES

[1] P.-R. Chang and W.-H. Yang, "Environment-adaptation mobile radio prop agation prediction using radial basis function neural networks,", IEEE Trans. Veh. Technol., vol. 46, no. 1, pp. 155–160, 1997

[2] Adawi, Nadia S. "Coverage prediction for cellular mobile radio system operating in the 800/900 MHz frequency band." IEEE Trans. on Vehicular Technology 37.1 (1988): 3-72.

[3] Yao, S. "SSTD FY21 Clutter Modeling Improvements Recommendation", Applied Physics Laboratory Propagation Working Group Report, 2021.

[4] Marsland, Stephen. Machine learning: an algorithmic perspective. Chapman and Hall/CRC, 2011.

[5] Rennier, Anthony, et al. "Spectrum Sharing Test and Demonstration Program: Key Findings and Implications." IEEE Military Communications Conference (MILCOM), 2019.

[6] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

[7] Kingma, D and Ba, J . (2015) Adam: A method for Stochastic Optimization. Available at: https://arxiv.org/pdf/1412.6980.pdf

[8] Xie, Junyuan, Ross Girshick, and Ali Farhadi. "Unsupervised deep embedding for clustering analysis." International conference on machine learning. PMLR, 2016.

[9] C. M. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., 2006

[10] Seretis, Aristeidis, and Costas D. Sarris. "An overview of machine learning techniques for radiowave propagation modeling." IEEE Transactions on Antennas and Propagation (2021)

[11] Cheerla, S. "A semi empirical path loss model by using artificial neural networks (Ann)." Int. J. Innov. Technol. Explor. Eng. 8.7 (2019): 835-841.

[12] Walfisch, Joram, and Henry L. Bertoni. "A theoretical model of UHF propagation in urban environments." IEEE Transactions on antennas and propagation 36.12 (1988): 1788-1796.

[13] Jo, Han-Shin, et al. "Path loss prediction based on machine learning techniques: Principal component analysis, artificial neural network, and Gaussian process." Sensors 20.7 (2020): 1927.

[14] Ratnam, Vishnu V., et al. "FadeNet: Deep learning-based mm-wave large-scale channel fading prediction and its applications." IEEE Access 9 (2020): 3278-3290.

[15] Abraham, Deme C. "Soft Computing Techniques for Path Loss Estimation at 1800MHz in a Tropical Metropolitan Environment: Case Study of Abuja-Nigeria." Soft Computing 4.10 (2017).

[16] Eppink, David, and Wolf Kuebler. TIREM/SEM Handbook. Electromagnetic Compatibility Analysis Center Report, 1994.