Model Standardization and Operationalization with a Development and Management Platform ("DMP")

Dan McCloskey
US Government Division
Palantir Technologies Inc.
Denver, CO
dmccloskey@palantir.com

***Abstract*—No industry standard currently exists for evaluating, comparing, and selecting the most appropriate model for analyzing spectrum interferences. Instead, spectrum models use a range of different assumptions, data inputs, and interference mechanics. Palantir Technologies discusses how implementing a Model Development and Management Platform can promote the standardization of model development and validation, as well as support regulators' need to make operational decisions on spectrum management.**

Keywords—spectrum management, spectrum model management, model development, interference analysis, model deconfliction, model training, cloud, commercial software

I. Introduction

The proliferation of useful data and open-source capabilities for model development offer great promise for improving the analysis and simulation of electromagnetic propagation and interference. However, because much of this data and software is available via open sources, and there are no central hubs or industry standards for collaboration, there now exists an ecosystem of spectrum models that use different assumptions, input data types, and propagation and interference mechanics. There is no standard method for evaluating, comparing, and making operational decisions on how to select the best models for a given circumstance. The different outputs produced from different models can lead to difficulties credibly analyzing or advocating for a specific regulatory or operational course of action.

To foster community and stakeholder trust in models used for critical analyses, regulators should consider implementing a standardized environment for model development, training, validation, and management. Commercial software such as a Model Development and Management Platform ("DMP") exists today that can be rapidly fielded to meet this need.

Palantir Technologies ("Palantir") presents this abstract to detail how a DMP could be leveraged to:

- **Standardize approaches to modeling** for interference analyses and other spectrum-related data initiatives;
- **Increase transparency** into how models were developed and the data used;
- **Deconflict competing models and identify the most useful approaches** for a given analysis/scenario, while adapting over time; and
- **Empower model developers in the community** with tools for enhancing the accuracy of their models, gaining insight into their performance, and more rapidly developing, training, testing/evaluating, and deploying them.

We will summarize our lessons learned deploying DMP solutions that centralize, standardize, and augment modeling efforts in industries such as defense, intelligence, research, and energy, among others.

II. Features of a Model Development and Management Platform

Standardized approaches to modeling require the integrated management and validation of both data and models. As a commercial software company specializing in data integration and data-driven decision support, Palantir has built a DMP that addresses each of these challenges in a single, secure, collaborative ecosystem. The DMP enables data scientists, model developers, and operational users to collaborate securely on model development, evaluation, deployment, and performance monitoring. This abstract outlines the features of the DMP that we have found to be critical for meeting objectives of standardizing and augmenting approaches to modeling.

An enriched and model-ready data foundation for transparent, consistent model development and operationalization. When models are developed using different data foundations, it can be difficult to make effective comparisons or determine their relative utility in a given context. Ensuring the consistency and accuracy of the data used to inform model development, training, and validation is critical to the utility of the models themselves. To effectively compare models against one another, there needs to be transparency and standardization of the data underpinning them. We have consistently found that the best model approaches are

constructed using variegated and often sizable data, which is well-organized, cleaned, labeled, and split into training, validation, and testing datasets. The DMP integrates multiple data streams and provides out-of-the-box connectors to integrate virtually any type of data into a unified platform, e.g., real-time sensor data, imagery data, video data, acoustic data, historical data, and tracks data. Users can then utilize DMP's version-controlled development environment and open APIs to rapidly iterate on top of the integrated data asset.

A version-controlled and provisioned environment for experts to build, train, evaluate, debug, and deploy models securely and rapidly. Models can be seamlessly deployed to the DMP through containers or raw binaries. Ingested models automatically run against curated subsets of data for developer-facing testing and evaluation, enabling a diverse field of contributors to continually identify, characterize, and visualize model performance in real world or fabricated scenarios. Models use training datasets derived from data integrated in the DMP (such as all-source data, metadata, third-party services, and/or model outputs) and can be analyzed alongside the data for rapid iteration and experimentation. Models can be trained programmatically or surfaced along with important context (like user-added feedback) for a human to review in closer detail as the context demands. Model developers can also use the platform for debugging and analysis on both intermediate and flagship models.

A Testing & Evaluation ("T&E") environment to improve model performance. Users build confidence in models when they can understand a model's behavior, assess its suitability, and make critical, immediate adjustments both during T&E and in the field. The DMP stores all model versions with user and model metadata so users can interrogate causation and model performance. T&E can be performed on any data in the platform—from historical data to synthesized inputs like simulating inaccurate metadata. Users can branch production data and code to improve, test, and debug models using real world data in a staging environment before models move to production. This protects data pipelines, promotes model reliability, accelerates model deployment, and avoids unnecessary downtime. Users can visualize the impact of proposed changes and conduct counterfactual analysis through the entire downstream data ecosystem (on reports, data analytics, and more).

The DMP provides configurable visualization capabilities to display metrics and results of algorithms, including the following:

- **Interactive accuracy plots and curves:** Visualize the results of models with different parameters to understand how those parameters affect model accuracy.
- **Comparison of different versions of the same initial algorithm:** Compare the outputs of different algorithm branches, for example, to assist them in selecting models for continued development.
- **Visualization of algorithm output:** Visualize metrics on model speed and accuracy.
- **Side-by-side comparisons of different algorithms:** In addition to comparing different versions of the same algorithm, users can compare different algorithms. For example, users can run the models against the same evaluation data and assess results.

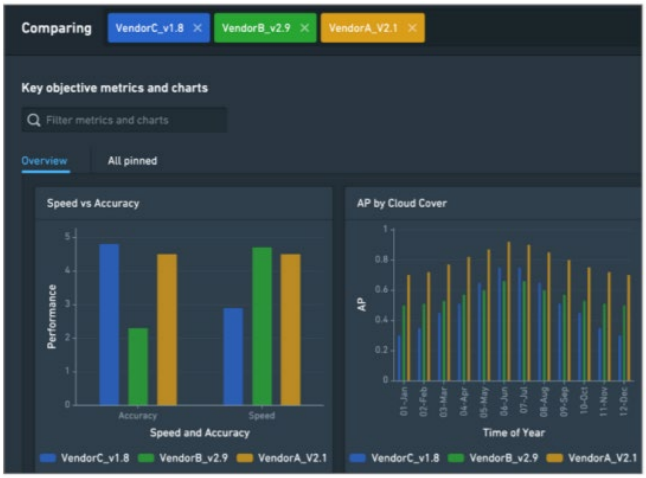

Fig. 1. Model Comparison. Users can measure, monitor, and compare hundreds of models side-by-side to find the most appropriate algorithm. All data notional.

To answer questions such as these, the DMP provides Insight Reports and other integrated analytic tooling. The DMP supports computation of model results (e.g., metrics, diagnostics) for each model and model version, broken out by arbitrary data segments. Subsequent ad-hoc or systematic analyses (either code-based or via point-and-click) can leverage this data to:

- Identify the best model for a certain segment
- Monitor model performance over time for each individual segment
- Identify segments with poor model performance overall (which might require additional training data)
- Identify model variance and reliability on those segments

Users can perform analyses to, for example: identify possible effective ensembles of models (e.g., models that complement each other) for a given operating region; or cluster models with similar overall performance profiles, using readily available or bespoke clustering logic. Analytic outputs can be surfaced as dynamically updating reports or dashboards to easily visualize algorithm performance.

Model management and versioning. The DMP includes Git-like versioning and branching for model management. Metadata about every model training cycle (including a record

of the data on which the model is trained) is logged, as are all model versions. All information is stored securely in Palantir and is subject to strict access controls. The DMP can capture model metadata, including a model's description, format, architecture, parameters, hyperparameters, target application problems or subsets, key training data, feature information, author information, and other configurable metadata. Metadata is available via user interfaces (e.g., model summary view), and through platform workflows such as discovery within the model repository. Metadata can be declared programmatically or manually by model developers, as well as (where possible) extracted programmatically within the platform. Programmatic and UI-based model management capabilities allow full versioning of models. This is useful for automatic or conditional deployment, model monitoring (e.g., performance, distributions, invariants), model comparison, and the ability to roll back to previous versions. The DMP supports both programmatic and UI mechanisms for tracking and comparing any model metadata or performance attributes across model versions and branches, as well as between different (comparable) models.

Interoperability. The DMP is a fully open, interoperable platform. It seamlessly connects through robust, open APIs with legacy or future systems, tools, and models (whether customer-developed, third-party, or Open Source), supporting advancements as technologies evolve. The platform is designed with an open architecture and open APIs to allow for the integration of tools at any step of the model development process, including pre-processing, transformation, T&E, or model training tools. Each of these tools can integrate with the DMP's data versioning capabilities, ensuring that all relevant activity and metadata is tracked.