Project Report
ATC-430

The Offshore Precipitation Capability

M. Veillette
H. Iskenderian
M. Wolfson
C. Mattioli
E. Hassey
P. Lamey

16 September 2016

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

Prepared for the Federal Aviation Administration,
Washington, DC 20591

This document is available to the public through
the National Technical Information Service,
Springfield, Virginia 22161

This document is disseminated under the sponsorship of the Department of Transportation, Federal Aviation Administration (FAA), in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

This material is based upon work supported by the Federal Aviation Administration (FAA) under Air Force Contract No. FA8721-05-C-0002 and FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the FAA.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
ATC-430		

4. Title and Subtitle	5. Report Date
The Offshore Precipitation Capability	16 September 2016
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
M. Veillette, H. Iskenderian, M. Wolfson, C. Mattioli, E. Hassey, P. Lamey	ATC-430

9. Performing Organization Name and Address	10. Work Unit No.
MIT Lincoln Laboratory 244 Wood Street Lexington, MA 02420-9108	(TRAIS)
	11. Contract or Grant No. FA8721-05-C-0002 and/or FA8702-15-D-0001

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Department of Transportation Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	Project Report
	14. Sponsoring Agency Code

15. Supplementary Notes

This report is based on studies performed at Lincoln Laboratory, a federally funded research and development center operated by Massachusetts Institute of Technology, under Air Force Contract FA8721-05-C-0002 and FA8702-15-D-0001.

16. Abstract

In this work, machine learning and image processing methods are used to estimate radar-like precipitation intensity and echo top heights beyond the range of weather radar. The technology, called the Offshore Precipitation Capability (OPC), combines global lightning data with existing radar mosaics, five Geostationary Operational Environmental Satellite (GOES) channels, and several fields from the Rapid Refresh (RAP) 13 km numerical weather prediction model to create precipitation and echo top fields similar to those provided by existing Federal Aviation Administration (FAA) weather systems. Preprocessing and feature extraction methods are described to construct inputs for model training. A variety of machine learning algorithms are investigated to identify which provides the most accuracy. Output from the machine learning model is blended with existing radar mosaics to create weather radar-like analyses that extend into offshore regions. The resulting fields are validated using land radars and satellite precipitation measurements provided by the National Aeronautics and Space Administration (NASA) Global Precipitation Measurement Mission (GPM) core observatory satellite. This capability is initially being developed for the Miami Oceanic airspace with the goal of providing improved situational awareness for offshore air traffic control.

17. Key Words	18. Distribution Statement
	This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	38	

FORM DOT F 1700.7 (8-72) Reproduction of completed page authorized

This page intentionally left blank.

ABSTRACT

In this work, machine learning and image processing methods are used to estimate radar-like precipitation intensity and echo top heights beyond the range of weather radar. The technology, called the Offshore Precipitation Capability (OPC), combines global lightning data with existing radar mosaics, five Geostationary Operational Environmental Satellite (GOES) channels, and several fields from the Rapid Refresh (RAP) 13 km numerical weather prediction model to create precipitation and echo top fields similar to those provided by existing Federal Aviation Administration (FAA) weather systems. Preprocessing and feature extraction methods are described to construct inputs for model training. A variety of machine learning algorithms are investigated to identify which provides the most accuracy. Output from the machine learning model is blended with existing radar mosaics to create weather radar-like analyses that extend into offshore regions. The resulting fields are validated using land radars and satellite precipitation measurements provided by the National Aeronautics and Space Administration (NASA) Global Precipitation Measurement Mission (GPM) core observatory satellite. This capability is initially being developed for the Miami Oceanic airspace with the goal of providing improved situational awareness for offshore air traffic control.

This page intentionally left blank.

ACKNOWLEDGMENTS

We are especially appreciative of the support provided by Randy Bass and Steve Abelman of the FAA Aviation Weather Research Program (AWRP) for funding work to develop the Offshore Precipitation Capability. We'd also like to thank Sean Bodkin, Leonard Story and James Karlovich of FAA AJV-7 for their support. Finally, we wish to thank Peter Erickson, Richard Ferris and Joe Venuti of MIT Lincoln Laboratory for their contributions to this work.

This page intentionally left blank.

TABLE OF CONTENTS

This page intentionally left blank.

LIST OF ILLUSTRATIONS

This page intentionally left blank.

LIST OF TABLES

This page intentionally left blank.

1. INTRODUCTION

Current weather processing systems used in Air Traffic Management (ATM) rely heavily on the land-based Next Generation Weather Radar (NEXRAD), and Terminal Doppler Weather Radar (TDWR) radar networks for the detection of aviation-impacting weather. Systems such as the Corridor Integrated Weather System (CIWS)[1], and the Weather and Radar Processor (WARP) [2] provide rapidly updating Contiguous United States (CONUS)-wide analyses of precipitation intensity and storm height derived from volume scans of overlapping radars. However, the National Airspace (NAS) covers areas both inside and outside of land-based radar range leaving many offshore and oceanic controllers without timely weather information required for proper air traffic management. Figure 1 demonstrates an example of this short-coming by showing how a storm outside of radar range is not depicted whatsoever in the radar mosaic, despite the presence of lightning flashes observed offshore. This lack of adequate situational awareness may be detrimental to both passenger safety and can lead to inefficiencies in the NAS. This shortcoming resulted in an Air Traffic Organization Corrective Action Report (CAR-2011-023) to highlight the safety concerns of limited or no NEXRAD coverage in the Miami offshore airspace and called for a solution [3].

The Offshore Precipitation Capability (OPC) is a system designed to address this problem. The goal of OPC is to provide reliable and timely depictions of offshore precipitation and storm height fields commonly used by ATM in regions beyond weather radar coverage. OPC works by fusing non-radar data in regions where radar data is diminished or non-existent. These non-radar data used to estimate radar-like fields include global lightning detections (cloud-to-ground and inter-cloud), geostationary satellite (visible (VIS) & infrared (IR) channels) and Numerical Weather Prediction model output (Temperature, Pressure, Humidity, etc.).

These data sources are brought together in a supervised machine learning framework that allows for the estimation of fields normally derived through weather radar measurements. A machine learning algorithm is trained in regions where both NEXRAD data and these non-radar data are available. This trained algorithm is able to estimate values of radar-derived fields from the current conditions observed in non-radar data. OPC generates two fields that are of interest to air traffic operations. The first is Vertically Integrated Liquid (VIL), which is a field that represents the amount of liquid water in a vertical column of the atmosphere and serves as a proxy for precipitation intensity. The units of VIL are kg/m^2, however it is common to bin VIL into 6 intensity levels as shown in Figure 1. The other field generated by OPC is Echo Tops (ET), which represents the height of a storm, defined here as the height of the 18 dBz radar echo. Once these fields are generated from the machine learning model, the results are blended with existing radar mosaics to create a seamless radar mosaic that extends into offshore and oceanic regions.

Methods for estimating precipitation by merging multiple datasets have existed for decades and are well studied [4], [5]. Multi-channel rainfall estimates derived from Visible (VIS) and Infrared (IR) satellite imagery are common, as geostationary (GEO) satellites provide imagery on a global scale fairly

regularly [6] [7],and [8]. Other methods combine VIS and IR imagery from a GEO satellite with the limited coverage of passive microwave (PMV) sounders that are capable of more accurate precipitation estimates, for instance Climate Prediction Center morphing method (CMORPH) [9], and Integrated Multi-satellitE Retrievals for GPM (IMERG) [10]. Lightning flash density has been used to provide a surprisingly accurate proxy for radar reflectivity [11], [12]. This is particularly enticing for OPC due to the availability of global lightning detection networks [13] [14], as well as the future availability of the Geostationary Lightning Mapper (GLM) on the GOES-R satellite [15].

Not surprisingly, improved results are obtained by combining data across multiple sensors, for example, combing lightning with satellite data [16]. OPC falls into the family of these multi-sensor techniques by combining satellite, lightning, numerical model, and radar data into one precipitation-related product. However, OPC also sets itself apart in numerous ways. Namely, OPC focuses on fields which are primarily of interest to ATM. A rapid updating capability was developed to make use of the freshest data possible in order to provide timely data on par with existing radar-based systems. OPC also employs novel feature extraction and non-linear regression models to combine highly heterogeneous input datasets.

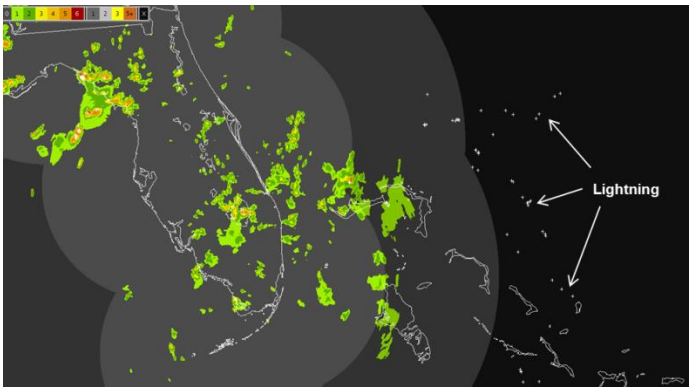

Figure 1. VIL with lightning flashes observed to the east (white markers). In this case, NEXRAD range (indicated with gray shading) does not extend far enough to properly depict these storms.

An example of OPC being applied in the Gulf of Mexico is shown in Figure 2. The large cloud seen off the coast of Texas and Louisiana in the left panel indicates that a potential storm exists, but the limited range of NEXRAD is not able to depict the storm intensity. The right panel shows the same case with the OPC algorithm applied. Using other types of data like lightning and satellite yields a detailed picture of storm structure and intensity that would be valuable for controllers directing air traffic through this region.

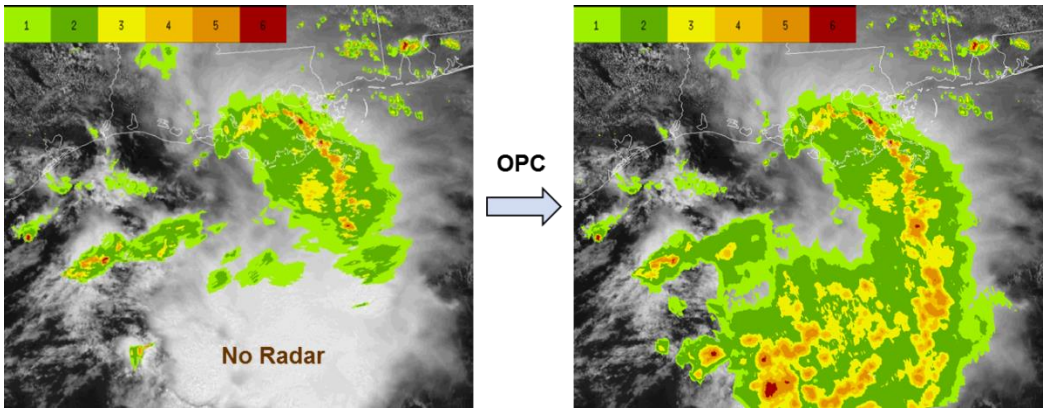

Figure 2. Example of the OPC algorithm. The satellite in the left panel shows a potential storm exists over the Gulf of Mexico; however the limited range of NEXRAD is unable to depict details of storm intensity. On the right, the OPC algorithm combined a number of lightning, satellite and model indicators to estimate precipitation intensity over the region.

This report is organized as follows. Section 2 describes details of the OPC development, including input data sources, preprocessing, and model training. Section 3 contains initial validation results, where OPC was compared to land radar, and spaceborne radar.

This page intentionally left blank.

2. DEVELOPMENT OF THE OPC MODEL

The OPC algorithm is based upon a supervised machine learning approach. In this approach, a regression model is trained using data gathered over regions with full radar coverage. This dataset consists of training examples that contain the field(s) of interest (VIL, ET), as well as features extracted from non-radar data. For optimal results, it is important that this training set contain a representative sample of conditions that will eventually be modeled by OPC.

This section will provide details about the development of OPC. It will detail what non-radar data were used, what preprocessing steps were applied to the data, what features are extracted from the data, and the machine learning algorithm used to create the model

2.1 INPUT DATA

Inputs for OPC come from three major sources: lightning, geostationary satellite, and numerical weather prediction models. Each source of data is explained in more detail below.

2.1.1 Lightning

Lightning is a strong indicator of convective activity. Earth Network's Total Lightning Network provides historical and real-time lightning data for OPC training and application. The data contains latitude/longitude, timestamp, amplitude, multiplicity, and height information of all detected flashes. This data also contains a cloud-to-ground/inter-cloud delimiter, though only cloud-to-ground was used in this work.

2.1.2 Geostationary Satellite

Satellite data for OPC comes from the GOES-13 (GOES-E) satellite. All imager channels provided by the satellite are utilized, including the 1 km VIS (0.65 μm) and all four 4-km IR channels (3.9 μm, 6.7 μm, 10.7 μm, and 13.3 μm). Data arrival times depend on the current GOES-E schedule set by the Office of Satellite and Product Operations (OSPO). There are different sectors scanned by the GOES-E satellites: CONUS, Northern Hemisphere, Southern Hemisphere, and Full Disk. During routine operations, CONUS scans arrive roughly every 10–15 minutes whereas the larger Northern Hemisphere and Full Fisk scans take longer to arrive (30–45 minutes).

2.1.3 Numerical Weather Prediction Models

Numerical weather models are capable of simulating a large number of meteorological quantities in regions not covered by traditional sensors, including 3D sounding of temperature, pressure and humidity, as well as relevant 2D fields such as simulated radar reflectivity, rain rate, and convective cloud top height. The National Oceanic and Atmospheric Administration's (NOAA) 13 km Rapid Refresh (RAP)

model[1] was chosen for the initial version of OPC because of its relatively high resolution, hourly update and large spatial coverage. A portion of the full RAP domain (Grid 83), which covers the Atlantic ocean from south of Puerto Rico up to Southern Canada, are ingested and archived for training and simulation. Model fields used in OPC include 2 hour forecasts of Temperature, Pressure, Humidity, Composite Reflectivity, Precipitation Rate, Precipitable Water, Mean Sea Level Pressure, and Convective Available Potential Energy (CAPE). The 2 hour lead time was chosen to account for forecast latency.

2.2 DATA PREPROCESSING AND FEATURE EXTRACTION

Prior to extracting features for the machine learning model, a number of preprocessing steps are necessary. Preprocessing steps include visible satellite normalization, computation of cloud top height, parallax correction of satellite data, and binning and filtering of lightning data at multiple spatial and temporal scales. These steps are applied both during training and in real time application.

Feature extraction translates the set of input satellite, lightning and model images into a dataset containing a fixed number of columns. Each row of this dataset represents a vector of features sampled at a particular geographic location, and each column represents a particular image feature. Features are obtained by applying a function to all pixels within a kernel centered at each point of interest. A number of different feature functions and kernel sizes are used in OPC feature extraction. Many features are classified as "1st order statistics" which are quantities that don't depend on how data are ordered within the kernel, such as mean, median, standard deviation, or a certain percentile of the data. Higher order feature functions are also used to capture quantities like image texture. For these higher order features, the magnitude of the gradient of underlying image $|\nabla I|^2 = {I_x}^2 + {I_y}^2$, is applied and 1st order statistics are computed using this transformed image.

Kernel sizes used in feature extraction range from 1 to 13 km. The set of features used in OPC are summarized in Table 1. For satellite images affected by parallax, displacement vectors are applied to reposition kernel centers prior to extracting features.

[1] http://rapidrefresh.noaa.gov

TABLE 1

Image Features used for OPC Model Training

Image Type	Kernel sizes (km)	Image Features	Gradient Features
5 Min Lightning	1, 5, 10	mean	-
10 Min Lightning	1, 5, 10	mean	-
20 Min Lightning	1, 5, 10	mean	-
VIS	1, 4, 10	max, min, std, range	median, 95th percentile
IR 3.9	4, 12	max, min, median, std, range	median, 95th percentile
IR 6.7	4, 12	max, min, median, std, range	median, 95th percentile
IR 10.7	4, 12	max, min, std, range	median, 95th percentile
IR 13.3	4, 12	max, min, std, range	median, 95th percentile
Cloud Top Height	4, 12	max, min, std, range	median, 95th percentile
Solar Zenith Angle	1	mean	-
Filtered Lightning	1	mean	-
RAP Comp Refl	13	max, min, median, mean	-
RAP Precip. Water	13	max, min, median, mean	-
RAP Precip. Rate	13	max, min, median, mean	-
RAP Mean Sea Level Pressure	13	mean	-
RAP CAPE	13	mean	-

Training data for OPC covers a year of data between May 2014 and May 2015 and covers the southeast domain shown in Figure 3. Radar mosaics from the Corridor Integrated Weather Service (CIWS) [17], [1] are used for truth. Radar mosaics with time stamps within 2.5 minutes of GOES scans were sampled at approximately one hour increments. Locations were randomly sampled in three data regimes: (a) No VIL (b) VIL Levels 1 & 2, and (c) VIL Levels 3 and higher. At each sample location, values of the radar mosaics were saved along with the location information and time stamp. For training, VIL is converted to *digital VIL*, in which a logarithmic transform is applied to larger VIL values in order to balance the distribution (i.e., lessen the long tail). No transformation was applied to Echo Tops data.

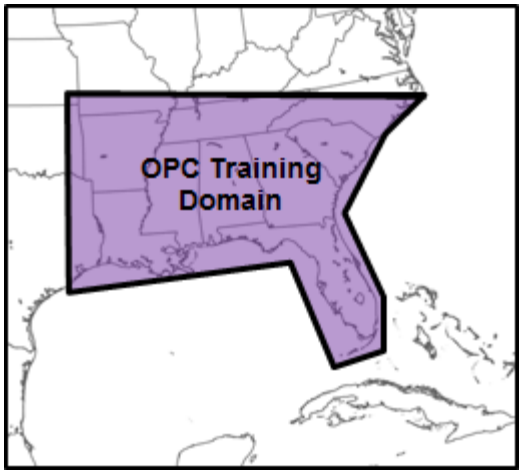

Figure 3. OPC Training Domain.

For each point in the target dataset, lightning, satellite and model data were searched for a matching time stamp. If a time stamp that was sufficiently close could not be found for each input image type, this point was not used for training. If data matching this time stamp was found, input data underwent preprocessing and feature extraction. These features were paired with the observed radar fields and saved in the training database.

Since VIS features are not valid during night time hours, the dataset was split into two pieces based on the solar zenith angle. Rows with solar zenith angle less than 60^o were used for day-time modelling. Rows with solar zenith angle greater than 60^o had the visible features removed, and were used for night-time training. Finally, the resulting datasets were scrubbed of rows with any missing values, and downsampled such that they contained equal distributions of the three VIL regimes listed above (No VIL, low VIL, and high VIL). Finally, data was split into monthly training files, each containing 45K rows for both day and night time training.

2.3 MODEL TRAINING

A number of machine learning algorithms were tested for training the OPC model. The families of algorithms considered were

- *Linear Models:* Models which use a linear combination of input features for predictions. Coefficients in the linear model are fit to minimize mean squared error plus a regularization term. Two types of regularization were investigated: *Ridge Regression* (L2 regularization), and *Lasso Regression* (L1 regularization), [18].

- *Artificial Neural Networks (ANN):* These are non-linear models made up of groups of nodes connected through transfer functions whose parameters are fit from training data. The input features make up the set of input nodes, and these are connected to a number of hidden layers of nodes that feed forward into the final output. The parameters of the transfer functions are tuned through a method called back propagation.

- *Ensemble Learning Models:* This family of models trains a number of "weak learners" (in this case, regression trees) and combines them to make a prediction which is, in general, more accurate than any individual learner. Algorithms tested in this family include Random Forest Regression (RFR), Extra Trees Regression (ETR), and Gradient Boosting Regression (GBR) [19], [20], [21].

Prior to training, columns of the input dataset were normalized by applying a z-score transformation, and a grid search was performed to choose optimal hyper-parameters. In addition, lightning based features underwent a log-transformation.

For each model, a grid search was performed to choose optimal hyper-parameters. A 10-fold cross-validation was run to measure the coefficient of determination (r^2) for both day and night time datasets, and for digital VIL, Echo Tops, and Composite Reflectivity datasets. To avoid over-fitting during cross validation, the dataset was sorted by time stamp, and train/test datasets contained only segments of consecutive days. Only summertime data (June–August) for VIL and Echo Tops was used for model selection. The results of this series of cross validations for model selection are shown in Table 2.

TABLE 2

Cross Validation Performance of the Machine Learning Algorithms Considered

In each cell, the numbers are RMSE/r^2. For VIL's RMSE, the units are in digital VIL and for echo tops they are in units of kft.

	Digital VIL		Echo Tops (kft)	
	Day	Night	Day	Night
Ridge Regression	42.30/0.6523	48.17/0.5415	6.891/0.7944	7.939/0.7164
Lasso Regression	42.31/0.6521	48.06/0.5435	7.118/0.7806	8.348/0.6863
Neural Networks	39.15/0.7011	44.52/0.6092	6.276/0.8294	**7.172/0.7686**
Random Forest	38.93/0.7054	44.50/0.6086	6.232/0.8326	7.273/0.7595
Extra Trees Regression	38.81/0.7072	**44.39/0.6105**	6.190/0.8349	7.231/0.7622
Gradient Boosting	**38.79/0.7075**	44.48/0.6089	**6.180/0.8354**	7.256/0.7606

With the set of features chosen, GBR outperforms the other models during day time, but by only a small margin (~0.3%) over other ensemble models, and a slightly wider margin over neural networks (~1%). ANN and ETR outperformed other models for night time data, but again, only by a small amount. The linear models (Ridge and Lasso) were outperformed by all other algorithms, which suggest that non-linear models are better suited for prediction based on these heterogeneous features. With the exception of night time echo top data, for which ANNs offer a slight improvement, ensemble methods show the best potential for OPC modeling. Despite the minuscule improvement provided by GBR and ETR, the Random Forest was chosen as it was most easily implemented in a real-time processing system.

2.4 RADAR MOSAIC STITCHING

As a final step, the output of OPC is "stitched" or merged with radar mosaics over land or wherever valid radar data is available. Since OPC is ultimately an estimation of radar mosaics, using true radar is almost always preferred when and where it is available. To stitch OPC with radar, a weight field is derived from a weighting function, $w(d)$, that depends on the distance in kilometers, d, to a closest radar. Near a radar, where d is small, $w(d)$ is at or near 0. As d increases towards the maximum range of radar data, $w(d)$ transitions smoothly towards 1. The weight computed at each pixel is then used in a linear combination of the OPC and CIWS data. This combination results in the stitched output. A demonstration of this process can be seen in Figure 4.

$$OPC_{stitched} = w(d)OPC + (1 - w(d))CIWS$$

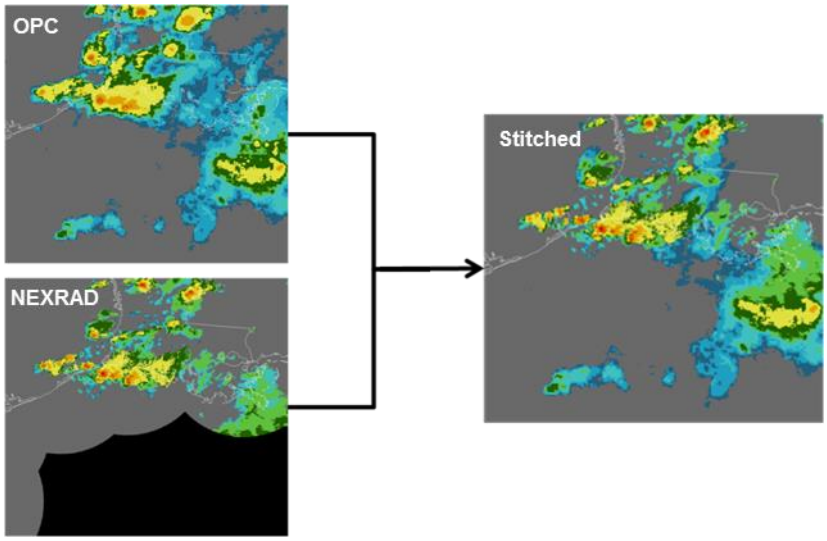

Figure 4. Example of OPC radar stitching process. A weighting function is applied to combine OPC (top left) and a radar mosaic (bottom left). The result of the weighted combination is pictured on the right.

3. OPC MODEL ASSESSMENT

This section will look at the reliability of OPC VIL & ET in regions near and around the southeast U.S. where OPC was trained. Two datasets are used for OPC validation. The first is NEXRAD VIL and ET produced by CIWS. While NEXRAD data isn't available over ocean, these data are abundant (CIWS generates a new mosaic every 2.5 minutes) and also of high quality. A comparison of NEXRAD VIL and ET to their OPC counterparts is shown in Figure 5. Note that throughout this section, OPC is *not* stitched with NEXRAD.

To validate OPC outside the range of NEXRAD, we utilize spaceborne radar onboard NASA's Global Precipitation Measurement Mission's (GPM) Core Observatory Satellite [5]. This satellite, launched in February of 2014, is a low earth orbiting satellite equipped with Dual-frequency Precipitation Radar (DPR) with Ku (13.6 GHz) and Ka (35.5 GHz) bands. As such, this radar provides validation data over ocean, however its availability is much less than NEXRAD's. An overpass of the core observatory provides a 3D swath of radar reflectivity that is 245 km wide for the Ku band, and 125 km wide for the Ka band. The horizontal resolution of the data is 5km, with a vertical resolution of 500m. Reflectivity data obtained from this satellite is mapped to VIL and ET, and these fields are compared to their counterparts in OPC. An important fact to consider is that the high frequency Ka band on this radar suffers from attenuation in deep convective storms [22]. This should be kept in mind when treating the DPR as "truth". A comparison between GPM VIL and ET, to OPC over the Caribbean is shown in Figure 6.

One qualitative difference between the two fields is that, in general, OPC appears more "dilated", or softer than radar, and does not match the same 1 km storm resolution of NEXRAD radar mosaics. This can be seen by noting that the coverage of Level 1 & 2 VIL, as well as coverage of lower ET appears higher in OPC than in corresponding truth (NEXRAD or GPM). This is mainly a result of two factors: larger kernel sizes in the feature extraction, and the low resolution nature of satellite and model data. Despite this difference, OPC is able to capture the shape and character of many of the severe storms.

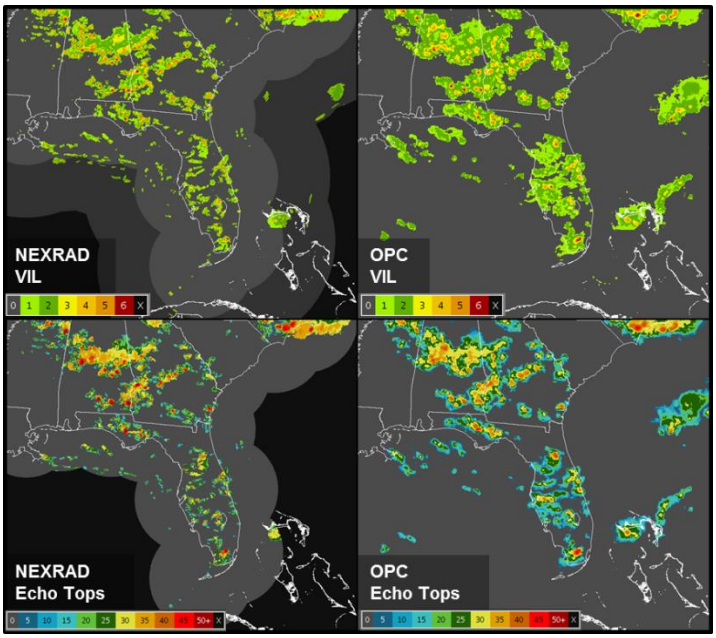

Figure 5. Comparison of NEXRAD (Left), and OPC (Right) VIL and ET.

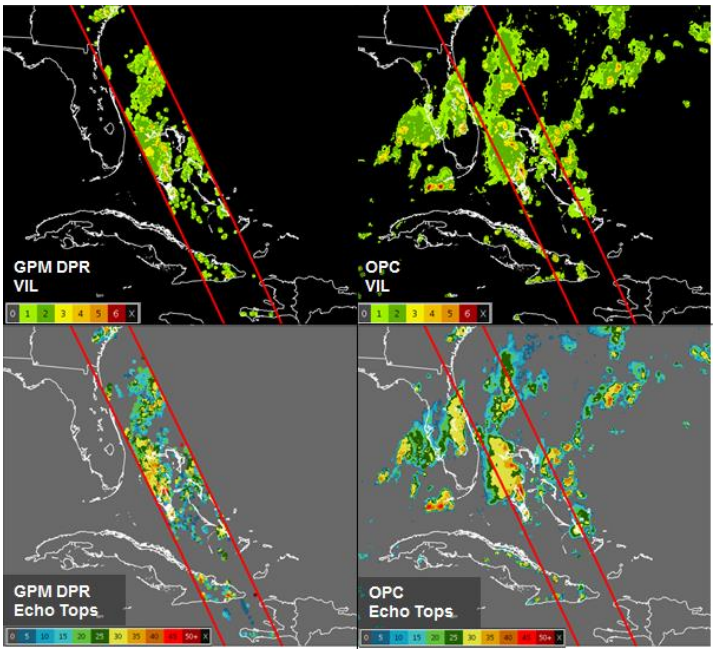

Figure 6. Comparison of GPM DPR, and OPC VIL and ET.

OPC was run for the summer of 2015 over the regions shows in Figure 7. Data generated within the southeast region will be assessed using NEXRAD (CIWS), whereas data over the Gulf and Caribbean regions will use GPM DPR. For the NEXRAD comparison, geographic points were randomly selected for OPC VIL & ET images generated every 30 minutes over the months of June, July, and August, 2015. For GPM DPR, 20 overpasses were selected within both the Gulf and Caribbean regions and points were selected in the portion of the overpass that intersected the scoring regions. These overpasses were selected based on weather content (a mixture of high and low impact cases) and data quality.

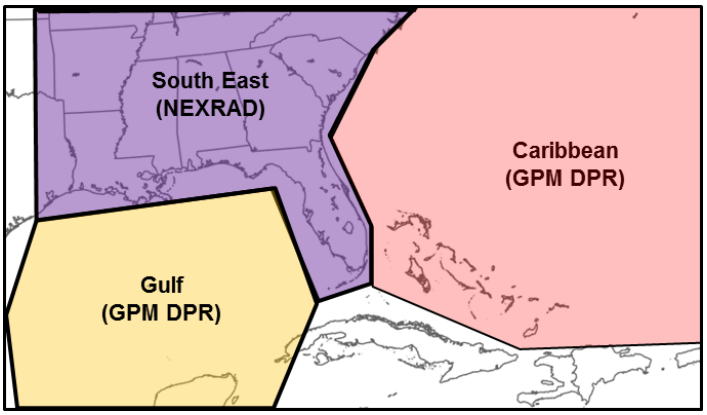

Figure 7. OPC scoring regions. Over land (southeast domain), NEXRAD data taken from CIWS will be used as validation data. Offshore (Gulf & Caribbean), VIL and ET are measured from GPM DPR.

Figure 8 shows the "pixel-by-pixel" comparison between NEXRAD, and OPC VIL & ET in the southeast domain. The y-axis in both of these plots show the average value of VIL/ET from NEXRAD conditioned on OPC values along the x-axis. Ideally, curves should lie along the dashed diagonal line, which would imply that OPC is (conditionally) unbiased. Curves are split by day (10UTC–22UTC) and night to show differences in performance when visible satellite is or isn't included in the set of input features.

All curves are monotone increasing, which suggests that OPC values are coupled to observations (i.e., observing high/low OPC implies high/low NEXRAD). For both VIL and ET, OPC is conditionally biased high in the "mid-ranges" (i.e., when OPC is showing VIL Level 1 & 2 or ET less than 30–35 kft). This bias is worse at night, especially in OPC ET. An example of this high bias in these ranges can be observed in Figure 5. In this example, the storms along the Florida Peninsula show gaps in the low level NEXRAD VIL and ET. In OPC, many of these gaps appear filled with VIL level 1 and 2.

The results improve for high values of OPC VIL (≥ Level 3) & ET (≥ 30 kft). Both fields show a low conditional bias in both day and night time, as all curves converge near the diagonal line. This is also

consistent with the example shown in Figure 5 – nearly all storm cores visible in OPC are associated to a storm core in NEXRAD, although OPC storms appear larger, as discussed previously.

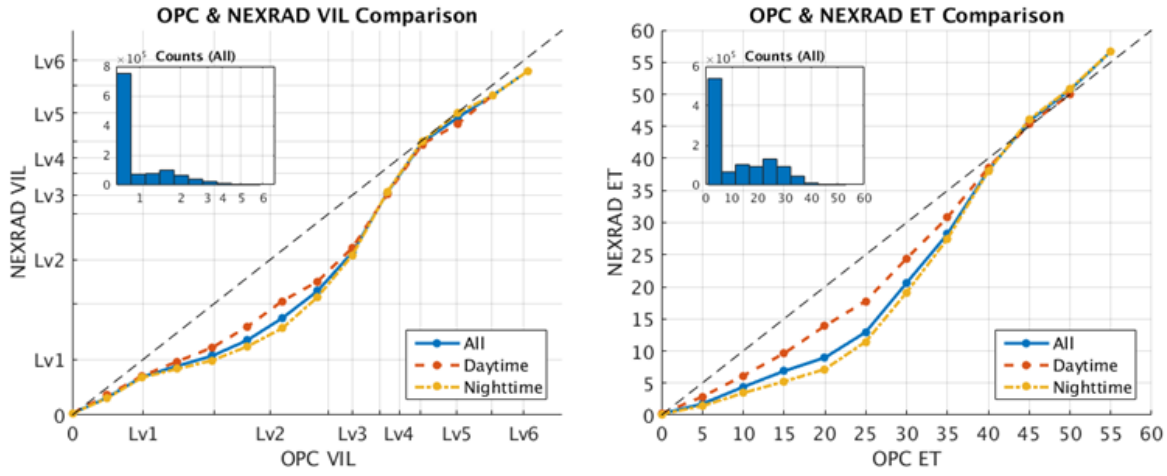

Figure 8. Comparison between OPC and NEXRAD (CIWS), VIL(left) and ET (right) for the southeast domain. The y-axis in both plots shows the average value of the NEXRAD VIL/ET conditioned on OPC taking the values along the x-axis.

Figure 9 shows a similar comparison between GPM DPR, and OPC VIL and ET for the Caribbean and Gulf regions (no day/night split was performed on these datasets due to a limited number of passes). Similar to NEXRAD, the OPC VIL and ET behave similarly to the fields derived from the DPR, as each curve is monotone increasing. Both OPC fields show a high conditional bias relative to GPM. Unlike NEXRAD, this high bias persists for the "high range" of VIL and ET. In the VIL especially, the curves remain below the diagonal even for high OPC showing Level 5 and 6. The bias appears to be slightly better for the Gulf domain, aside from OPC predictions of 50kft (which is likely a statistical anomaly due to small sample size).

There are two possibilities that might explain the high bias observed offshore that's not present in the NEXRAD comparison. First, weaker vertical motion in the storms associated with weaker surface forcing and instability observed in oceanic storms, may cause their ET and VIL to not reach as high as over land (where OPC was trained). Better features might need to be added to OPC to alleviate this issue. The second factor contributing to high bias offshore is the attenuation of the DPR. We suspect this is especially true for the low DPR VIL values observed in areas of high OPC VIL.

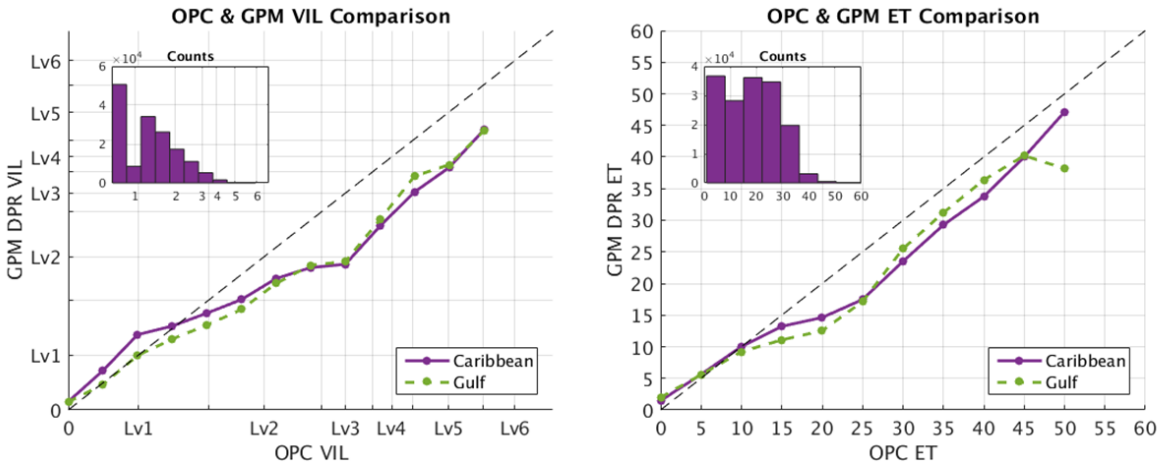

Figure 9. Comparison between OPC and GPM DPR, VIL(left) and ET (right) for the Caribbean and Gulf domains. The y-axis in both plots shows the average value of the DPR VIL/ET conditioned on OPC taking the values along the x-axis.

3.1 DETECTION OF AVIATION IMPACTING WEATHER

In addition to comparisons in the previous section that looked at OPC performance across the entire range of VIL & ET values, this section studies the performance of OPC in depicting aviation impacting weather. "Aviation Impacting" is defined here as VIL Level 3 or greater and Echo Tops ≥ 30 kft. OPC images are compared to NEXRAD images and scored based on the following contingency table:

TABLE 3

Contingency Table used for OPC Assessment

NEXRAD VIL \ OPC VIL	VIL < Lv3	VIL ≥ Lv3
VIL < Lv3	Correct Rejection	False Alarm
VIL ≥ Lv3	Miss	Hit

NEXRAD ET \ OPC ET	ET < 30 kft	ET ≥ 30 kft
ET < 30 kft	Correct Rejection	False Alarm
ET ≥ 30 kft	Miss	Hit

To account for differences in resolution, a 5 km search radius was applied to look for matching OPC pixels. This is done to prevent penalizing OPC for creating storms that are more dilated than those seen in observations.

Once all pixels are classified, the following statistics Probability Of Detection (POD), BIAS, and Critical Success Index (CSI) are computed within the scoring domain. These statistics are defined as

$$\boldsymbol{POD} = \frac{\#Hits}{\#Hits + \#Misses} \quad \boldsymbol{BIAS} = \frac{\#OPC}{\#Truth} = \frac{\#Hits + \#FAs}{\#Hits + \#Misses} \quad \boldsymbol{CSI} = \frac{\#Hits}{\#Hits + \#Misses + \#FAs}$$

These statistics were computed within the southeast domain in Figure 7 for June, July and August of 2015 and the results are shown in Figure 10. The points in these plots represent individual OPC images sampled at least 15 minutes apart (images without enough pixels exceeding a threshold are not plotted). The orange curve in each plot shows a 4-day moving average. The average (POD, BIAS, CSI) for the whole summer is (0.6272, 0.8727, 0.5026) for OPC VIL, and (0.7201,1.0323,0.559) for OPC ET.

Overall, scores are better for OPC ET than for VIL, which is not surprising given that ET is a function of the top of the storm, in contrast to VIL which depends on an entire column of the atmosphere. There is a noticeable diurnal pattern in the OPC scores, where the skill drops overnight. This drop is due to the loss of the VIS channel, and due to the nature of convection in the southeast domain (afternoon thunderstorms which generate a lot of lightning are generally well detected by OPC). Periods where the average performance drops (e.g., around June 17th) are typically a result of heavy precipitation without a lot of lightning. Better satellite features (e.g., GOES-R) will help alleviate both of these issues.

At first glance it seems contradictory that the results in the previous section show a high bias in OPC VIL (Figure 8) whereas the bias in these results is lower on average. The difference can be explained by the 5km search radius used here to classify OPC pixels as hit or false alarm. This has the effect of removing the "dilated" pixels from the bias computation. This helps confirm that OPC creates storms that are larger than observed storms, but rarely creates any aviation impacting storms that aren't there (low false alarm rate).

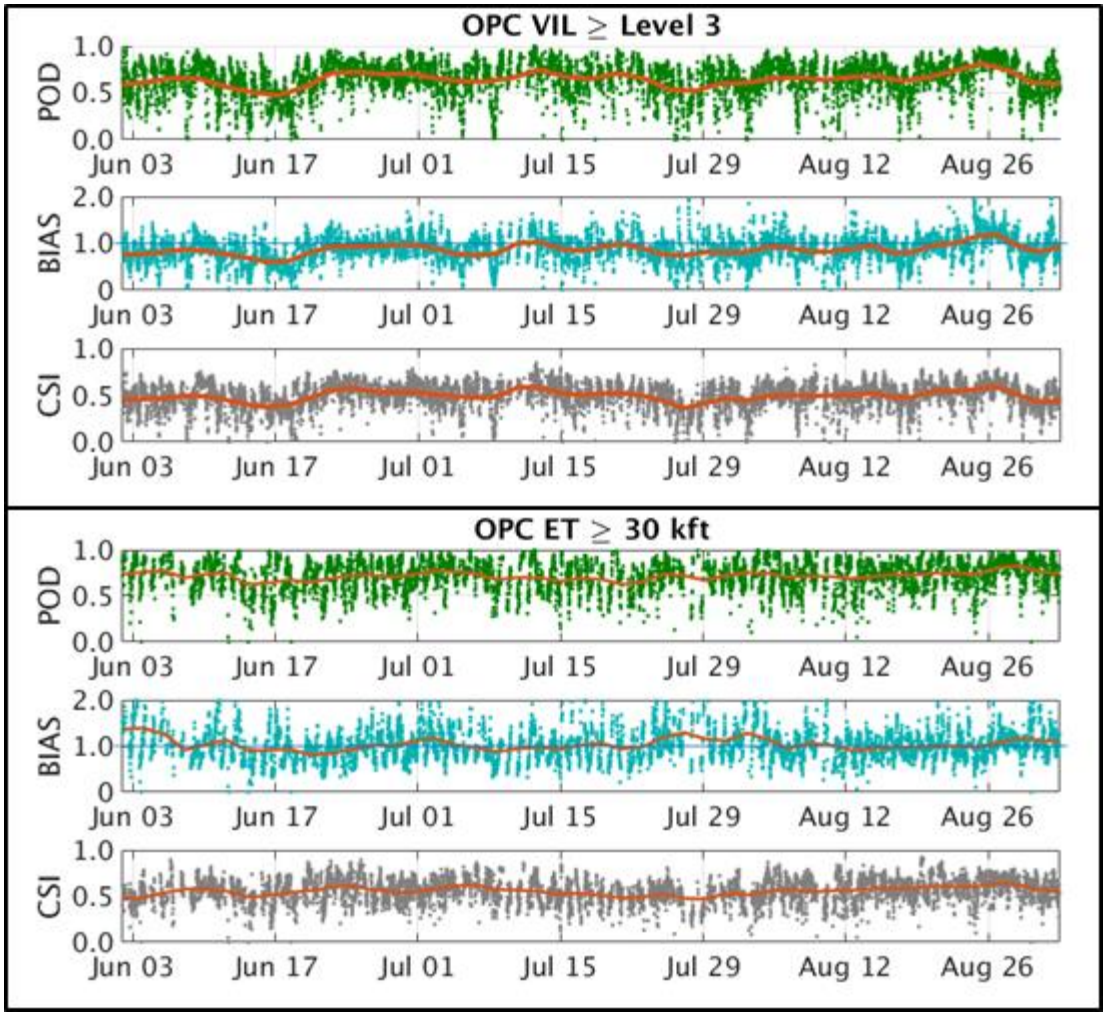

Figure 10. OPC performance for detection VIL ≥Level 3(top) and ET≥30kft (bottom) for the summer of 2015. The points in each plot represent OPC images sampled 15 minutes apart (images without enough impacting weather are not plotted). The orange line in each plot represents a 4-day moving average.

This page intentionally left blank.

4. CONCLUSION

The Offshore Precipitation Capability (OPC) is a system that estimates weather radar-like fields used by aviation in areas without weather radar coverage. This capability has the potential to extend the coverage of current weather radar mosaics used in today's ATC to offshore and oceanic areas. In particular, OPC can benefit Air Traffic Controllers by improving the safety and efficiency of the National Airspace.

OPC works by estimating radar-derived fields, such as VIL and Echo Tops, by combining a number of non-radar data that are available in offshore and oceanic areas. These non-radar data include lightning flashes, geostationary satellite data, and output from numerical weather prediction models. Features extracted from the data are combined in a machine learning model to estimate VIL, and Echo Tops in regions outside weather radar coverage. Output from OPC is merged with existing radar in regions with radar coverage to create a seamless radar mosaic.

OPC performance was studied over the summer of 2015. OPC outputs were compared to NEXRAD radar in the southeast U.S., and to GPM DPR radar over the Caribbean and Gulf of Mexico. Pixel-to-pixel comparisons showed that OPC is generally higher than NEXRAD- and DPR-based fields. The high bias was mainly observed in the mid-ranges of VIL and ET. Bias over ocean appeared to be higher, although some of this might be caused by attenuation in the DPR.

The ability of OPC to depict aviation impact storms was also assessed. For this a 5 km search radius was applied for matching OPC storms to observations. The results here showed that OPC had a POD of 63% for VIL $\geq$ Level 3, and 72% for ET $\geq$ 30fkt. In contrast to pixel-to-pixel scores, the bias for VIL $\geq$ Level 3 storms was slightly low (0.87), which was the effect of the 5km search radius applied when scoring. The bias ET $\geq$ 30fkt was low (1.03), suggesting that OPC is better at depicting impacting storms with high tops.

Future work for OPC involves a third party assessment of OPC that will contain a much larger sample than was used in this report. Feedback from this assessment will be used to further improve the OPC model. Additional OPC improvements include expansion of the OPC domain to cover the expanded CONUS domain, and creation of Composite Reflectivity in selectable layers of the atmosphere and Echo Tops at selectable reflectivity thresholds. This work is planned for FY2016 and beyond.

This page intentionally left blank.

GLOSSARY

ANN	Artificial Neural Networks
ATM	Air Traffic Management
CAPE	Convective Available Potential Energy
CIWS	Corridor Integrated Weather System
CMORPH	Climate Prediction Center morphing method
CONUS	Contiguous United States
CSI	Critical Success Index
DPR	Dual-frequency Precipitation Radar
ET	Echo Tops
ETR	Extra Trees Regression
FAA	Federal Aviation Administration
GBR	Gradient Boosting Regression
GEO	geostationary
GLM	Geostationary Lightning Mapper
GOES	Geostationary Operational Environmental Satellite
GPM	Global Precipitation Measurement Mission
IMERG	Integrated Multi-satellitE Retrievals for GPM
IR	infrared
NAS	National Airspace
NASA	National Aeronautics and Space Administration
NEXRAD	Next Generation Weather Radar
NOAA	National Oceanic and Atmospheric Administration
OPC	Offshore Precipitation Capability
OSPO	Office of Satellite and Product Operations
PMV	passive microwave
POD	Probability Of Detection
RAP	Rapid Refresh
RFR	Random Forest Regression
TDWR	Terminal Doppler Weather Radar
VIL	Vertically Integrated Liquid
VIS	visible
WARP	Weather and Radar Processor

This page intentionally left blank.

REFERENCES

1. D. Klingle-Wilson, and J. Evans, “Description of the Corridor Integrated Weather System (CIWS) Weather Products.” Lexington, MA: Massachusetts Institute Technology Lincoln Laboratory, Project Report ATC, 2005.

2. B. Deans, T. Hicks, R. Graff, and S. Walden, "FAA's Weather and Radar Processor (WARP) convective storm demonstration." 9th Conference on Aviation, Range, and Aerospace Meteorology, Orlando, FL, 2000.

3. L. Mack, and S. Hansen, "Friends and Partners in Aviation Weather." FPAW Summer Meeting. Washington, D.C, 2013.

4. Chris Kidd, and George Huffman, "Global precipitation measurement." Meteorological Applications 18.3 (2011): 334-353.

5. Francisco J. Tapiador, Francis J. Turk, Walt Petersen, Arthur Y. Hou, Eduardo García-Ortega, Luiz AT Machado, Carlos F. Angelis, et al. "Global precipitation measurement: Methods, datasets and applications." Atmospheric Research 104 (2012): 70-97.

6. M. B. Ba and A. Gruber, "GOES multispectral rainfall algorithm (GMSRA)." Journal of Applied Meteorology 40.8 (2001): 1500-1514.

7. A. Behrangi, K.L. Hsu, B. Imam, S. Sorooshian, and R. J. Kuligowski, "Evaluating the utility of multispectral information in delineating the areal extent of precipitation." Journal of Hydrometeorology 10.3 (2009): 684-700.

8. K. L. Hsu, H. V. Gupta, X. Gao, and S. Sorooshian, "Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation." Water Resources Research 35.5 (1999): 1605-1618.

9. Robert J. Joyce, John E. Janowiak, Phillip A. Arkin, and Pingping Xie, "CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution." Journal of Hydrometeorology (2004).

10. George J. Huffman, David T. Bolvin, Daniel Braithwaite, Kuolin Hsu, Robert Joyce, Chris Kidd, Soroosh Sorooshian, and Pingping Xie, "Heading Toward Launch with the Integrated Multi-Satellite Retrievals for GPM (IMERG)." 2012.

11. H. Iskenderian, "Cloud-to-ground lightning as a proxy for nowcasts of VIL and echo tops." 13th Conference on Aviation, Range and Aerospace Meteorology,. 2008.

12. J. Anderson, C. C. Liu, E. Novakovskaia, and S. Heckman, "Total Lightning and Radar Reflectivity: A Study and Application." Earth Networks, n.d.

13. Earth Networks Total Lightning Network. n.d. Earth Networks. 4 June 2015. <https://www.earthnetworks.com/Products/TotalLightningNetwork.aspx>.

14. Vaisala Global Lightning Dataset GLD360. n.d. Vaisala. 4 June 2015. <http://www.vaisala.com/en/products/thunderstormandlightningdetectionsystems/Pages/GLD360.aspx>.

15. Steven J. Goodman, Richard J. Blakeslee, William J. Koshak, Douglas Mach, Jeffrey Bailey, Dennis Buechler, Larry Carey, et al., "The GOES-R Geostationary Lightning Mapper (GLM)." Atmospheric Research 125 (2013): 34-49.

16. M. Grecu, E. N. Anagnostou, and R. F. Adler, "Assessment of the use of lightning information in satellite infrared rainfall estimation." Journal of Hydrometeorology 1.3 (2000): 211-221.

17. Jim Evans, et al, "The Corridor Integrated Weather System (CIWS)." 10th Conference on Aviation, Range, and Aerospace Meteorology, 2001.

18. Christopher M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

19. Leo Breiman, "Random forests." Machine learning 45.1 (2001): 5-32.

20. Pierre Geurts, Damien Ernst, and Louis Wehenkel, "Extremely randomized trees." Machine learning 63.1 (2006): 3-42.

21. Jerome H. Friedman, "Stochastic gradient boosting." Computational Statistics & Data Analysis 38.4 (2002): 367-378.

22. Takuji Kubota, et al., "Evaluation of precipitation estimates by at-launch codes of GPM/DPR algorithms using synthetic data from TRMM/PR observations." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7.9 (2014): 3931-3944.