

UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

EXPLORING THE CLASSIFICATION OF WILDLAND FIRE RELATED SMOKE
USING WSR-88DS

A THESIS
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
Degree of
MASTER OF SCIENCE

By
EMMA MILLER
Norman, Oklahoma
2025

EXPLORING THE CLASSIFICATION OF WILDLAND FIRE RELATED SMOKE
USING WSR-88DS

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. David Schwartzman, Chair

Dr. Dušan Zrnić

Dr. Aaron Hill

This thesis is dedicated to my parents, John and Rhea Miller, and to my grandparents. Who I am now is because of their never ending support for my endeavors in my personal and professional lives. My successes in the past, present, and the rest of my life are due to their sacrifices, and my gratitude will always be boundless.

Acknowledgments

I would like to extend heartfelt gratitude to my advisor Dr. David Schwartzman, for the dedication and support in my endeavors for this work. Having the freedom to explore this topic and dive into something new has helped to improve my skills as a scientist and researcher. This work has given me a new passion for my career, and it would not have been possible without his encouragement to step out onto a new branch and try something different. I'm very fortunate to have worked with him, and other committee members, Drs. Dušan Zrnić and Aaron Hill on this work. Their insight and knowledge have helped to mold this work into something exciting moving forward. I hope to continue in my professional career with the skills learned from them, and provide new, innovative thoughts to my field.

My thanks and gratitude also extends to my family; the backbone of my professional career comes from their endless support. Many difficulties and challenges, personally and professionally, can come from continuing one's education, and their encouragement has kept me going. They are my biggest cheerleaders and I can never fully express all of my appreciation for their support. I also want to express my gratitude to my friends and peers in the School of Meteorology and Advanced Radar Research Center. Having a supporting community makes challenges easier to overcome, and the joy of new experiences is made better with friendship.

Table of Contents

Acknowledgments	v
Table of Contents	vii
List of Tables	viii
List of Figures	xvii
Abstract	xviii
1 Introduction	1
1.1 Background and Motivation	1
1.2 Scope of the Thesis and Outline	6
2 Review of Fuzzy Logic Classification and Machine Learning Classification Models	8
2.1 Current Fuzzy Logic Classification Methodology	8
2.2 Machine Learning Classification Models	15
2.2.1 SGDClassifier and Decision Trees	16
2.2.2 Random Forest	19
3 Detection and Classification of Plumes	21
3.1 Statistical Analysis of Smoke Plumes	21
3.1.1 Data	21

3.1.2	Statistical Analysis and Plume Detection	22
3.2	Implementation for Fuzzy Logic Classification	30
3.3	Detection and Replacement Algorithm (DRA)	35
3.3.1	Accuracy Testing	38
3.3.2	Limitations and Further Applications	41
4	Machine Learning Models for Smoke Classification	45
4.1	Building the Dataset	45
4.2	Hyperparameter Selection and Training	47
4.2.1	Decision Tree Models	50
4.2.2	Random Forest	56
4.3	Current Limitations and Suggestions	58
5	Application of Classification Algorithms and Models to Select Cases	60
5.1	Case Studies	60
5.1.1	Palisades and Eaton Fires, California	60
5.1.2	March 14th, 2025 Wildfires, Oklahoma	71
5.1.3	Snake Trail Fire, Minnesota	81
5.1.4	Smokehouse Creek Fire, Texas	90
5.1.5	Labor Day 2020 Fires, Oregon and Washington	103
5.2	Overall results	115
6	Conclusions and Future Work	117
6.1	Summary and Conclusions	117
6.2	Considerations for Future Work	121
	Bibliography	123

List of Tables

2.1	Parameters of the membership function for the hydrometeor classification algorithm as outline in Park et. al (2009). The functions $f_1 - f_3$ are defined in Equation 2.1, 2.2, and 2.3.	11
2.2	Weights used in the current hydrometeor classification algorithm as in [9]. .	12
3.1	Parameters of the membership function for the hydrometeor classification algorithm as outlined in Park et. al (2009) with the addition of a smoke class derived from statistical analysis. The functions $f_1 - f_3$ are defined in Equation 2.1, 2.2, and 2.3.	32
3.2	Matrix of weights as outline in Park et. al (2009) and [33] with the addition of smoke weights derived from statistical analysis.	33
4.1	Feature scenarios with different included features for comparison in model training and testing.	47
4.2	The hyperparameter grid options and the final selected value after using GridSearchCV, with the final values being used in new decision tree and random forest models.	49
4.3	Feature scenarios with different included features for comparison in model training and testing.	50

List of Figures

1.1	Yearly number and acres burned via wildfires as reported by the National Interagency Fire Center since 1983. The blue bars are the number of wildfires reported and the orange graph is the total acres burned during the year.	2
2.1	Example of a trapezoidal membership function for some polarimetric variable.	10
2.2	An example of the current output of the hybrid hydrometeor classification algorithm for the Palisades fire from the KSOX radar on 01/07/2025 at 22:39:36 UTC. The red circle indicates where there is a known wildfire and smoke plume.	14
2.3	Chart for showing the intersections of explainability and model complexity for common machine learning models. Those with high human understanding and lower model complexity are beneficial for confirming the model is learning correct physical relationships (from Flora et. al 2024 [28]).	16
2.4	An example hinge loss case that uses pluses and minuses to indicate samples with regards to the decision boundary (orange line). Samples that would be missclassified are noted in red.	17
3.1	Radars used for analysis of wildfires (red stars) and prescribed burns (green stars) from across the country, using the NEXRAD coverage map as reference. Multiple radar locations were used for more than one fire.	22

3.2	The smoke detection output on a.) reflectivity b.) correlation coefficient and c.) differential reflectivity for the March 1st, 2017 wildfire via the KVNIX radar on 03/01/2017 at 18:35:09 UTC.	24
3.3	Histogram of Z from the 0.5° elevation angle in two cases for a.) a fire in Jackson County, Oklahoma from the KFDR radar at 17:55:50 UTC on March 1st, 2017, with the corresponding Z field with the black contour indicating the area of smoke detection below and b.) the Buffalo fire from the KVNIX radar at 22:14:24 UTC on February 18th, 2016 and the Z field below to show the left-skewed and bimodal distributions respectively. .	26
3.4	Histogram from the 0.5° elevation angle ρ_{hv} in two cases for a.) the Snake Trail fire from the KMPX radar at 20:12:41 UTC on March 3rd, 2024 and b.) the Bolt Creek fire from the KATX radar at 16:43:09 UTC on September 10th, 2022 with the red line indicating the mean of the distribution.	27
3.5	The smoke detection output in black and the active fire contour in red that highlights areas of lowest Z_{DR} , highest Z , and low ρ_{hv} the same as everywhere else in the smoke plume for 05:21:03 UTC on September 8th, 2020 from the KRTX (Portland, OR) radar.	28
3.6	The detection of smoke with a.) no addition of the texture of reflectivity and b.) with texture of reflectivity added to the detection algorithm for the Beachie Creek fire via the KRTX radar on 2020/09/08 at 05:21:03 UTC. . .	30
3.7	Smoke observations in fields of a.) the reflectivity field for the time of interest for the Smokehouse Creek fire, and b.) the HCA output with the inclusion of smoke (black) for a case with multiple wildfires on 02/27/2024 at 20:39:51Z from the KAMA (Amarillo, TX) radar. The dark gray denotes areas that are classified as biological. There is a noted overlap of areas still classified as BS in the smoke plume, and areas of clutter or BS now being classified as smoke.	34

3.8	Smoke observations in fields of a.) the reflectivity field for the time of interest for multiple wildfires (red circles), and b.) the HCA output with the inclusion of smoke (black) on 09/08/2020 at 20:08:14Z from the KRTX (Portland, OR) radar. The dark gray denotes areas that are classified as biological. There is a noted overlap of areas still classified as BS in the smoke plume, and areas of clutter or BS now being classified as smoke. . . .	35
3.9	An example of the workflow for the the detection and replacement of values for smoke using the hybrid hydrometeor classification algorithm where a.) the original HHCA output b.) the original HHCA output with the smoke area detected, and c.) the final output of the algorithm where smoke is black on the colorbar. This figure shows the Palisades and Eaton fires in Los Angeles on January 8th, 2025 at 03:59:37 UTC.	38
3.10	An example of the the shapefile overlay on top of the final output of detection and replacement algorithm for the Smokehouse Creek fire on February 27th, 2024 at 23:25:12 UTC as seen from the KAMA radar.	40
3.11	An example of other biological scatter cases when passed through the detection algorithm for a.) chaff case from KBYX (Key West, FL) at 17:48:53 on August 8th, 2016, b.) another chaff case but from KDOX (Dover AFB, DE) at 03:59:00 UTC on July 6th, 2016, and c.) a case with birds from the KHTX (Huntsville, AL) radar at 11:15:10 UTC on August 11th, 2015. . . .	42
4.1	The specific importances a.) variable b.) training permutation, and c.) testing permutation from the DT1 models that uses all variables and hyperparameters from the GridSearchCV selection.	52
4.2	Example decision tree that is used to decide if a grid point should be classified as smoke as in the detection and replacement algorithm.	54

4.3	The visualization of the a.) DT1 and b.) DT2 model decision trees. The difference in interpretability is evident with the DT2 model, since it has clear splits which can be traced back and compared to the physical model in Figure 4.2.	54
4.4	As in Figure 4.1, but for the DT2 model.	56
4.5	As in Figure 4.1, but for the random forest model.	58
5.1	Final fire perimeters of the Palisades and Eaton fires from the National Interagency Fire Center WFIGS Fire Perimeters	61
5.2	Z (left column), ρ_{hv} (middle column), and Z_{DR} (right column) for selected times a.) 23:20:52 UTC on January 7th, 2025, b.) 15:19:12 UTC on January 8th, 2025, and c.) 19:24:40 UTC on January 9th, 2025 from the KSOX radar. The black outlines indicate the smoke plume or plumes for each time step based on polarimetric variables. In a.) the red circle is an area of incorrect classification as seen in the red oval in Figure 5.3, and the green oval indicates areas of mountain returns from the Santa Ana mountains. . . .	63
5.3	Examples of the detection and replacement algorithm (right column) when compared to the original HCA output (left columns) for selected times a.) 23:20:52 UTC on January 7th, 2025, b.) 15:19:12 UTC on January 8th, 2025, and c.) 19:24:40 UTC on January 9th, 2025 from the KSOX radar. The red oval in a.) indicates an area of incorrect classification likely driven by the presence of other biological scatterers.	64
5.4	ROC curves for the detection and replacement algorithm for selected times on a.) January 7th, b.) January 8th, and c.) January 9th.	65

5.5	Selected metrics for model comparisons for the DT1 model (top row), the DT2 model (middle row), and the random forest model (bottom row). The left column is the confusion matrix output for each model on the selected data with 150 being the positive label for smoke and 10 indicating biological scatter. The right column is the ROC curve based on calculated probabilities for each model. The red dashed line indicates where performance would be the same as a random classifier. The metrics for this figure are for January 7th, 2025 from the KSOX (Santa Ana Mountains) radar. . . .	68
5.6	As in Figure 5.5, but for January 8th, 2025.	69
5.7	As in Figure 5.5, but for January 9th, 2025.	70
5.8	Machine learning predictions for a selected time on January 9th, 2025. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA. . . .	71
5.9	The Rapid X-Band Polarimetric Radar (RaXPoI) scanning a wildfire on March 14th, 2025 outside Norman, OK at Lake Thunderbird. The orange colored haze is dust and smoke driven by the high winds.	72
5.10	The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 23:23:52 UTC from the KVNXX radar and b.) 22:09:20 UTC from KTLX. . . .	74
5.11	The final output of the detection and replacement algorithm for a.) 23:23:52 UTC from the KVNXX radar and b.) 22:09:20 UTC from KTLX. The red ovals indicate areas of missed classification of smoke.	75
5.12	ROC curves for the detection and replacement algorithm for a.) KVNXX and b.) KTLX for the March 14th, 2025 wildfires.	75
5.13	As in Figure 5.5 but for the wildfires that took place on March 14th, 2025 from the KVNXX radar.	78
5.14	As in Figure 5.5 but for the wildfires that took place on March 14th, 2025 from the KTLX radar.	79

5.15	Machine learning predictions for 23:23:52 UTC on March 14th, 2025 from the KVNX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.	80
5.16	Machine learning predictions for 22:09:20 UTC on March 14th, 2025 from the KTLX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.	81
5.17	The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 20:12:41 UTC, b.) 20:46:52 UTC, and c.) 21:20:21 from the KMPX (Minneapolis, MN) radar.	83
5.18	The final output of the detection and replacement algorithm for a.) 20:12:41 UTC, b.) 20:46:52 UTC, and c.) 21:20:21 UTC from the KMPX radar on March 3rd, 2024.	84
5.20	ROC curve for the detection and replacement algorithm for the Snake Trail fire on March 3rd, 2024.	85
5.19	Selected times a.) 20:12:41Z, b.) 20:46:52Z, c.) 21:03:37Z, and d.) 21:20:21Z for the smoke plume size ticker addition to the detection and replacement algorithm. The dark blue line in b.) indicates the position and passage of the cold front through the fire and smoke plume. There is a noted reinvigoration of the fire and smoke production between b.) and d.).	86
5.21	As in Figure 5.5 but for Snake Trail fire on March 3rd, 2024 from the KMPX radar.	88
5.22	Machine learning predictions for 20:46:52 UTC on March 3rd, 2024 from the KMPX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.	89

5.23	Machine learning predictions for 21:20:21 UTC on March 3rd, 2024 from the KMPX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.	90
5.24	Final fire perimeters of Smokehouse Creek and other fires from the same time period across the Texas and Oklahoma Panhandles.	91
5.25	The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 20:31:46 UTC on 2/26/24, b.) 23:25:12 UTC on 2/27/24, c.) 00:29:46 UTC on 2/28/24, and d.) 02:17:05 UTC on 2/28/24 from the KAMA (Amarillo, TX) radar. The red star indicates the location of Amarillo, TX. The blue line indicates the approximate location of cold front, and the red ovals indicate the smoke area associated with the Windy Deuce fire.	93
5.26	The final output of the detection and replacement algorithm for a.) 20:31:46 UTC on 2/26/24, b.) 23:25:12 UTC on 2/27/24, c.) 00:29:46 UTC on 2/28/24, and d.) 02:17:05 UTC on 2/28/24 from the KAMA (Amarillo, TX) radar. The red ovals indicate the smoke area associated with the Windy Deuce fire.	95
5.27	ROC curves for the detection and replacement algorithm for selected times on a.) February 26th, b.) February 27th, and c.) February 28th.	96
5.28	As in Figure 5.5 but for Smokehouse Creek and other fires on February 26th, 2024 from the KAMA radar.	98
5.29	As in Figure 5.5 but for Smokehouse Creek and other fires on February 27th, 2024 from the KAMA radar.	99
5.30	As in Figure 5.5 but for Smokehouse Creek and other fires on February 28th, 2024 from the KAMA radar.	100

5.31	Machine learning predictions for 23:25:12 UTC on February 27th, 2024 from the KAMA radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.	102
5.32	Machine learning predictions for 02:17:05 UTC on February 28th, 2024 from the KAMA radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.	103
5.33	Final fire perimeters for Labor Day 2020 fires across the Cascades in Oregon and Washington.	104
5.34	The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 23:46:38 UTC on 9/7/20, b.) 23:56:57 UTC on 9/8/20, and c.) 03:34:21 UTC on 9/9/20 from the KRTX (Portland, OR) radar. Red ovals indicate areas of returns associated with topography, and the black outlines show where smoke plumes are located at.	106
5.35	The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 23:46:38 UTC on 9/7/20, b.) 23:56:57 UTC on 9/8/20, and c.) 03:34:21 UTC on 9/9/20 from the KRTX (Portland, OR) radar.	108
5.36	ROC curves for the detection and replacement algorithm for selected times on a.) September 8th and b.) September 9th.	109
5.37	As in Figure 5.5 but for fires in Oregon and Washington on September 8th, 2020 from the KRTX radar.	111
5.38	As in Figure 5.5 but for fires in Oregon and Washington on September 9th, 2020 from the KRTX radar.	112

5.39	Machine learning predictions for a.) 23:56:57 UTC on September 8th, 2020 and b.) 03:34:21 UTC on September 9th, 2020 from the KRTX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA. The red ovals indicate areas with incorrect classifications of smoke by both the DRA and the machine learning models.	114
------	---	-----

Abstract

In this thesis, we present multiple methods for classifying wildland fire-related (both wildfires and prescribed burns) particles on the WSR-88D radar. The word smoke in this work describes particles of sizes large enough to be sensed using the S-band WSR-88D system. The first method includes a new smoke class for the operational HCA method, the fuzzy logic classification algorithm. The basis of this new class lies in statistical analysis of key polarimetric variables for areas of smoke. Because of persistent challenges with implementing a new class with this method, two others were explored as well. A new detection and replacement algorithm was built based off the statistical analysis of areas of smoke from various wildfires and prescribed burns across the country. Areas detected as smoke are assigned a new value in the original HCA output from the WSR-88D, effectively creating a new separate class while maintaining the quality of the HCA output from the operational algorithm. The output from this algorithm also provides truth data necessary for machine learning models. A new dataset containing polarimetric variables, and the old and new HCA outputs, were used as inputs into tree-based machine learning models. This work shows the potential of machine learning for classification with further development. The detection and replacement algorithm performed the best out of these methods and shows promise for further applications with continued development. Through proposed future work, the detection and replacement algorithm will provide additional situational awareness for wildfires and additional hazards related to smoke.

Chapter 1

Introduction

1.1 Background and Motivation

In a changing climate, the three ingredients necessary for wildfires (fuels, oxygen, and ignition) are expected to be influenced in a myriad of ways. With regards to the fuels necessary for fire for example, changing precipitation patterns are expected to influence changes in vegetation resilience and drought length and severity. This can lead to areas experiencing prolonged droughts and having increased dry fuel loads available for wildfires. Since 1983, the National Interagency Coordination Center has collected data on the number of wildfires and the acres burned per year [1]. Figure 1.1 shows the number of fires and acres burned, per year. The yearly number of wildfires has remained fairly constant since recordkeeping began by this entity. However, the average acres burned per year has increased in the past twenty five years, nearly doubling, from an average of approximately 3 million acres prior to 2000 to nearly 7 million since [1]. So with approximately the same number of wildfires on average, the average acres burned has increased dramatically. This can be attributed to factors related to climate change, as well as, but not limited to, changes in population density, and intense fire suppression in the mid 20th century [2], [3], that has lead to build up of fuel. In addition to wildfires on the landscape, numerous agencies use prescribed fires for resource management and to improve the health of our land [4]. For example,

prescribed burning combats the build up of fuels and supports wildfire hazard abatement. Removing dense fuels via manual thinning or prescribed fire can alter and reduce the intensity and spatial spread of wildfires that may occur in the area, though this is dependent on the ecosystems and fuels being treated [5], [3]. Thus, a balance of fire suppression and using controlled prescribed burning is imperative for optimal resource management, hazard reduction, and improved safety for increasing populations living in fire-prone areas. The scope of this work investigates both wildfires and prescribed burns; these will be described together hereafter using the term “wildland fires”.

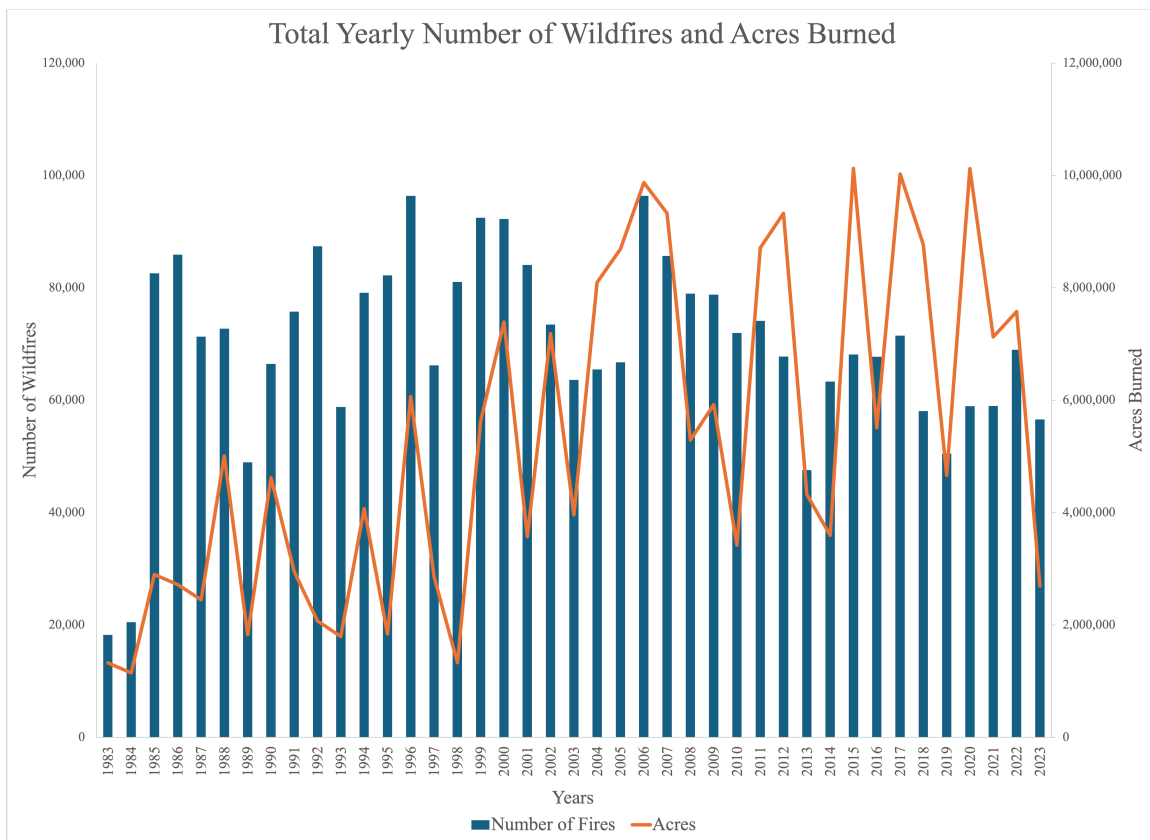


Figure 1.1: Yearly number and acres burned via wildfires as reported by the National Inter-agency Fire Center since 1983. The blue bars are the number of wildfires reported and the orange graph is the total acres burned during the year.

Due to their often remote nature, in-situ observations of smoke and fire are difficult to

collect. Remote sensing systems, like weather radars and satellites, are the most viable options to monitor fires and their evolution continuously. Satellites can monitor smoke plumes and the returned intensity of heat from fires across the country, but are often limited to being used during the day and in clear conditions. Radars can be operated 24/7 and can provide full volumetric scans every 5 minutes, but scan over a fixed area only [6]. This area however, is typically collocated with population centers or are near U.S. National Weather Service (NWS) offices nationwide. Currently the NWS operates 159 radars in their network of Weather Surveillance Radar 1988 Doppler (WSR-88D), that cover a vast majority of the central and eastern portions of the country with overlapping coverage [7]. The S-band (10-cm wavelength) frequency systems can observe fires at distances exceeding 100 kilometers and are sensitive to larger particles associated with fire, which are often called pyrometeors [8], or biomass, making them preferred candidates for smoke monitoring [8]. Dual-polarization weather radars can aid in detection and classification of various scatterers, as used the currently operational hydrometeor classification algorithm (HCA) for NEXRAD radars [9]. With these radar systems available, understanding the nuances of fire, smoke, and pyrometeor presentation on radar data is also feasible and can be used for tactical support for firefighting, communications with the public about health and safety threats, and for potentially anticipating and tracking changes in fire behavior.

Currently, an HCA is used operationally to determine what hydrometeor type is most likely to be present in each resolution volume, as in [9]. A resolution volume is the smallest volume space that can be resolved given the characteristics of the pulse(s) used, like pulse length, receiver filter, and the beamwidth of the radar. All of the scatterers in this volume influence the final value for polarimetric variables. The operational HCA uses fuzzy logic membership functions and weights to determine a possible class for the scatterers present based on values of polarimetric radar variables and textures of certain fields such as reflectivity and differential phase [9], to be explained further in Chapter 2. Smoke returns

produce distinctive values that are separate from meteorological scatterers such as low correlation coefficient and differential reflectivity typically larger than 0 dB. However, fields like reflectivity and local textures of differential phase and reflectivity may have similar values and will result in overlapping membership functions. This is especially true for biological scatterers and ground clutter. It would thus be beneficial for a separate class to be used for classification of smoke to mitigate misclassifications between the typical classification of biological scatterers and an active fire and smoke plume class, which are neither meteorological nor biological. There are numerous studies that investigate the separation of non-meteorological versus meteorological scatterers using fuzzy logic as well as machine learning (ML) models [10], [11], [12]. A similar misclassification issue was observed due to the overlap of membership functions for certain classes of liquid and frozen hydrometeors [9]. The original algorithm was improved upon by accounting for the melting layer in the resolution volume, so that the liquid and frozen hydrometeor classes could be divided (liquid below melting layer and frozen above) [9].

Changes in population density and urban expansion also impact the need for improved smoke and pyrometeor detection via the current operational system. Urban expansion and the movement of populations to suburban areas means that people are more frequently living in areas that have been previously undeveloped. The wildlife-urban interface (WUI) is defined by the U.S. Fire Administration as a zone of transition between unoccupied land and human development [13]. Communities in these areas are often at a higher risk of more severe wildfire hazards as this is the intersection of undeveloped land with fuels for fires and human development and the associated structures. Because of this intersection, the threat to life and property is inherently elevated if proper management of surrounding fuels does not occur. Studies on WUIs and wildfires have shown that WUI fires will likely increase in the future, and societal impacts will be larger [14], [15]. In addition to the immediate hazards from wildfires, smoke and ash particulates that are small, can be hazardous to

health. Air quality monitoring and public health agencies have also shown that levels of smoke and associated particulates such as $PM_{2.5}$ and carbon monoxide are detrimental to health especially in vulnerable groups [16–18]. The air pollutants associated with wildfires have been shown to have short term physical effects as well as potential long-term physical and mental health effects, although the studies and data for this are sparse [19]. Thus, early detection for the sake of public health and also to aid in timely fire suppression is of increasing importance. NEXRAD radars are typically positioned close to populated areas and can provide advanced warnings for smoke and fire for those living nearby or within the radar’s coverage area.

Artificial intelligence and machine learning are currently being applied to similar classification problems within the field of radar meteorology. For example, a similar challenge is present for separating birds and bugs from other biological scatterers (BS). This was addressed by using a ridge classifier and decision tree classifiers in [20], to separate birds from insects amongst the other biological scatterer fields as set in the current WSR-88D HCA. They found that the ridge classifier performed best and used just polarimetric variables for cases with different bird and insect patterns [20]. In aeroecology, radars have become increasingly useful in bird studies, for example. Further discrimination between birds and other biological scatterers is thus advantageous for numerous science fields. In [21], a random forest model was developed to predict areas of specific bird and bat species, based on polarimetric returns. They found the greatest distinction between meteorological and biological scatterers. Within the biological scatterers, further distinction could be found between arthropods and vertebrates. They noted that further work would be necessary, especially as scattering properties of flying animals are lesser known than that of hydrometeors. The same is true of ash and pyrometeors associated with wildland fires. Machine learning and convolutional neural networks have been applied to the task of dealiasing Doppler velocities, which was a task previously performed manually or via an algorithm

[22]. Previous works have shown that machine learning is a useful tool for classification problems that occur within the field of radar meteorology, and this work builds on that by expanding the use of different ML models. Other misclassification issues are still present with regards to clutter, chaff, and other non-meteorological scatterers, and are being addressed via algorithms and other ML models [23], [11], [10]. Evaluating if ML can or cannot tackle smoke classification will help provide a new solution to an existing problem for other biological scatterers or ground clutter/anomalous propagation (GC/AP) scatterers. If ML modeling is successful for the smoke case, accurate identification on radar can help with resource allocation for fire suppression, better messaging to the public about impacts for smoke, and the data could eventually be used as truth or label data for smoke modeling data assimilation in the future.

1.2 Scope of the Thesis and Outline

The main contributions of this work are providing context and exploring methods for detecting and classifying wildland fire smoke and pyrometeors using the current operational NEXRAD system. The scope includes statistical analysis of smoke plumes which provides some characterization of scatterers using dual-polarization variables, in addition to three methods of classifying smoke. First, new smoke membership functions were created for an added smoke class to the legacy fuzzy logic classification method; second, empirically derived range values are used as a base for a detection algorithm, which is applied to the original legacy HCA product; and three, classification is done via various machine learning models that are suited for binary classification tasks. Additionally, different testing methods will be used to determine the accuracy of the various methods for comparisons.

This thesis is organized as follows. Chapter 2 explores the current fuzzy logic classification method and implementation on the WSR-88D system. This section highlights the

main motivating challenge of the biological scatterers class and how we can create a new smoke class for the current operational system. This chapter will also set the foundation for the different machine learning models explored for this thesis, breaking down how the training, testing, and mathematics apply to the new dataset created for this work. Chapter 3 presents the statistical analysis necessary for the fuzzy logic classifier and the results from testing several cases. After this, the implementation of the new membership function will be presented with results. The poor results of this method were the basis for the next model to be presented in this chapter. The detection and replacement algorithm will be explained with examples from a selected wildfire case, and finally the chosen method, receiver operating characteristic (ROC) curves for testing the accuracy of the detection will be explained. Chapter 4 discusses the machine learning component of this thesis, starting with the creation of a new dataset for the purpose of smoke classification. The training, testing, and hyperparameter selection for each model will be discussed next, and metrics for the best options for machine learning models will be presented. Chapter 5 is a synthesis of all of the methods for selected cases from across the country. A general overview of the fire will be provided with geographical and meteorological context, if necessary. The detection and replacement algorithm and the machine learning models will be the focus of discussion for this chapter. Finally, Chapter 6 will summarize the findings and contributions of this work. Suggestions for further work and applications will also be provided, as well as the final outcome for which method performed the best out of the options presented.

Chapter 2

Review of Fuzzy Logic Classification and Machine Learning Classification Models

2.1 Current Fuzzy Logic Classification Methodology

In the operational WSR-88D network, the classification of radar returns is determined using a fuzzy logic classification algorithm. This method is used across numerous different scientific fields to determine a final outcome based on fuzzy boundaries for different variable attributes. Fuzzy logic classification has three main steps, which are fuzzification, aggregation, and defuzzification. They are discussed below for applications to weather radar. Operationally, there are 14 classes in the hybrid hydrometeor classification algorithm, and 10 in the original HCA [9]. The classes from [9] are: 1) ground clutter and anomalous propagation (GC/AP); 2) biological scatterers (BS); 3) dry aggregated snow (DS); 4) wet snow (WS); 5) crystals of various orientations (CR); 6) graupel (GR); 7) “big drops” (BD); 8) light and moderate rain (RA); 9) heavy rain (HR); and 10) a mixture of rain and hail (RH). This method is discussed in later sections to highlight the process of fuzzy logic classification. However, the data used in this work is produced by the hybrid hydrometeor classification algorithm, which also uses the Hail Size Discrimination algorithm [24] to improve the hail and rain classes further, bringing the total number of classes up to 14. For this study, five radar variables are used as inputs for discrimination between the 14 classes.

These include 1) radar reflectivity factor Z from the horizontal polarization, 2) differential reflectivity Z_{DR} , 3) correlation coefficient of the horizontal and vertical polarization returns ρ_{hv} , 4) a texture parameter of the Z field, $SD(Z)$, and 5) a texture parameter of the differential phase, Φ_{DP} , field $SD(\Phi_{DP})$. The texture parameters are calculated as described in [9] and [25], specifically using the texture along ray function in the Python Py-ART module [26]. A 1-km running average window was used for $SD(Z)$ and a 2-km window was used for $SD(\Phi_{DP})$. In [25] and [9] it is explained that the variable, LKdp, or a logarithmic scale version of the variable K_{DP} , specific differential phase, is also used. K_{DP} is a difference between propagation constants for horizontally and vertically polarized waves [27], and can be directly calculated by using differential phase measurements at two different range locations. However, [25] notes that LKdp does not have a diagnostic value for biological scatterers, based on the weights in the BS class, and is neglected for this study.

The first step of the fuzzy logic classification process is fuzzification. During this step, membership functions for each class for various dual-polarization variables are used. Membership functions are found by applying typical range values for the different polarimetric variables for each radar echo class. Fuzzy logic classification is applied at each resolution volume, and thus precise values for the polarimetric variables of interest are used in the membership functions. The main purpose is to place the precise values into fuzzy subsets for the different class membership functions. Membership functions have a trapezoidal shape with a maximum value of 1 and a minimum of 0, with four parameters that are noted as x_1 , x_2 , x_3 , and x_4 as shown in Figure 2.1. Park et. al [9] explains that the trapezoidal membership function approximates the shapes of the probability density functions well enough for both symmetric and asymmetric distributions. Thus, they are used for this work as well, with distributions for ρ_{hv} and/or Z_{DR} being typically asymmetric for smoke as well. Table 2.1 shows the parameter values for the classes and polarimetric variables of interest for this work before smoke was investigated, specifically for the hybrid hydrometeor

classification used in the NEXRAD network (for S-band).

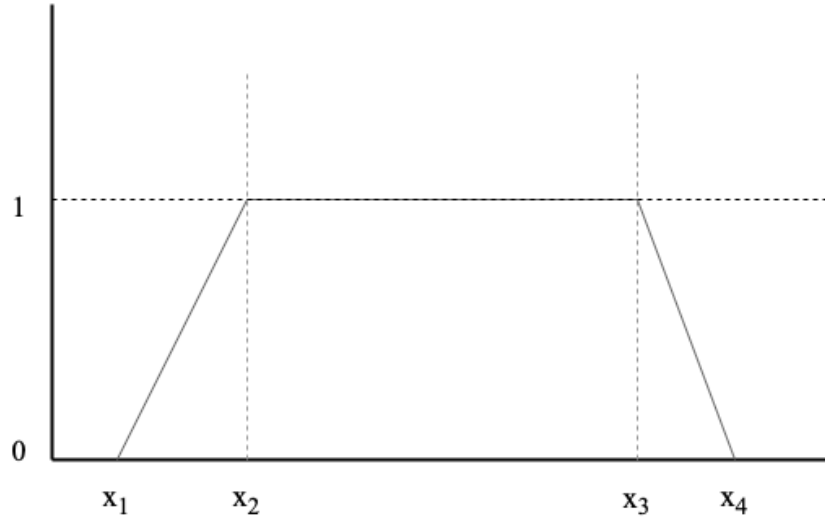


Figure 2.1: Example of a trapezoidal membership function for some polarimetric variable.

For Z_{DR} , some of the classes also include a dependence on reflectivity to improve the accuracy. The dependency was chosen for a nonrectangular area in a Z - Z_{DR} space where pairs of values correspond to a particular rain class [9]. These functions are:

$$f_1(Z) = -0.50 + 2.50 \times 10^{-3}Z + 7.50 \times 10^{-4}Z^2, \quad (2.1)$$

$$f_2(Z) = 0.68 - 4.81 \times 10^{-2}Z + 2.92 \times 10^{-3}Z^2, \quad (2.2)$$

$$f_3(Z) = 1.42 + 6.67 \times 10^{-2}Z + 4.85 \times 10^{-4}Z^2 \quad (2.3)$$

The second step of the fuzzy logic classification methodology is the aggregation of the values obtained after applying the membership functions. This can be accomplished by using simple to complex cost functions that perform summation, multiplication, or some combination of both. The current operational HCA uses an additive aggregation scheme, as [9] notes that hybrid schemes are vulnerable to biases in reflectivity caused by miscalculation and/or attenuation. As described in [25], the simplest aggregation scheme might

	GC/AP	BS	DS	WS	CR	GR	BD	RA	HR	RH
$P[Z(\text{dBZ})]$										
x_1	15	5	5	25	0	25	20	5	40	45
x_2	20	10	10	30	5	35	25	10	45	50
x_3	70	20	35	40	20	50	45	45	55	75
x_4	80	30	40	50	25	55	50	50	60	80
$P[Z_{DR}(\text{dB})]$										
x_1	-4	0	-0.3	0.5	0.1	-0.3	$f_2-0.3$	$f_1-0.3$	$f_1-0.3$	-0.3
x_2	-2	2	0	1.0	0.4	0	f_2	f_1	f_1	0
x_3	1	10	0.3	2.0	3.0	f_1	f_3	f_2	f_2	f_1
x_4	2	12	0.6	3.0	3.3	$f_1+0.3$	$f_3+1.0$	$f_2+0.5$	$f_2+0.5$	$f_1+0.5$
$P(\rho_{HV})$										
x_1	0.5	0.3	0.95	0.88	0.95	0.90	0.92	0.95	0.92	0.85
x_2	0.6	0.5	0.98	0.92	0.98	0.97	0.95	0.97	0.95	0.90
x_3	0.9	0.8	1.0	0.95	1.0	1.0	1.0	1.0	1.0	1.0
x_4	0.95	0.83	1.01	0.985	1.01	1.01	1.01	1.01	1.01	1.01
$P[SD(Z)(\text{dB})]$										
x_1	2	1	0	0	0	0	0	0	0	0
x_2	4	2	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
x_3	10	4	3	3	3	3	3	3	3	3
x_4	15	7	6	6	6	6	6	6	6	6
$P[SD(\Phi_{DP})(\text{deg})]$										
x_1	30	8	0	0	0	0	0	0	0	0
x_2	40	10	1	1	1	1	1	1	1	1
x_3	50	40	15	15	15	15	15	15	15	15
x_4	60	60	30	30	30	30	30	30	30	30

Table 2.1: Parameters of the membership function for the hydrometeor classification algorithm as outline in Park et. al (2009). The functions $f_1 - f_3$ are defined in Equation 2.1, 2.2, and 2.3.

look something like Equation 2.4 for the i -th class, where $P^{(i)}V_j$ is the characterizing distributions of V_j dual-polarization variables, that is,

$$A_i = \sum_j P^{(i)}(V_j) \quad (2.4)$$

However, the addition of weights and a confidence vector or index can help to strengthen the classification based on the membership functions. A matrix of weights is used to represent the discriminating efficiency of each variable with respect to a particular class [9]. For example, the low correlation coefficient values will likely increase the potential for a correct classification of biological scatter (BS) or ground clutter/anomalous propagation (GC/AP), when compared to another hydrometeor class. A vector or matrix of weights can be used, with values ranging from 0 to 1. A 2-D matrix is used operationally, as the weights used depended on both the polarimetric variable and the class. An example of the weights used are included in Table 2.2.

Class	Z	Z_{DR}	ρ_{HV}	$SD(Z)$	$SD(\Phi_{DP})$
GC/AP	0.2	0.4	1.0	0.6	0.8
BS	0.4	0.6	1.0	0.8	0.8
DS	1.0	0.8	0.6	0.2	0.2
WS	0.6	0.8	1.0	0.2	0.2
CR	1.0	0.6	0.4	0.2	0.2
GR	0.8	1.0	0.4	0.2	0.2
BD	0.8	1.0	0.6	0.2	0.2
RA	1.0	0.8	0.6	0.2	0.2
HR	1.0	0.8	0.6	0.2	0.2
RH	1.0	0.8	0.6	0.2	0.2

Table 2.2: Weights used in the current hydrometeor classification algorithm as in [9].

The confidence vector or index helps to improve the relationship that occurs between the quality of radar echo classification and the measurements from the radars themselves. There are six main factors that are included in this step that are related to the quality of the measured radar variables. They include (1) radar miscalibration, (2) attenuation, (3) nonuniform beam filling, (4) partial beam blockage, (5) the magnitude of ρ_{hv} , and (6) the receiver noise [9]. Park et al [9] explains in depth the necessary variables and methods needed to include the confidence vector in the aggregation step. The confidence vector is used to reduce the weight if an area falls within a compromised area of the returns. The

confidence vector is:

$$Q = (Q_Z, Q_{Z_{DR}}, Q_{\rho_{HV}}, Q_{SD(Z)}, Q_{SD(\phi_{DP})}), \quad (2.5)$$

and each variable quality component is defined as the multiplication of all the factors from the different measurement errors. That is,

$$Q_j = F_{attenuation} F_{ndf} F_{rhv} F_{snr} F_{pbb}. \quad (2.6)$$

So, with the addition of the weight matrix and the confidence vector, the final equation that is used in the aggregation step is:

$$A_i = \frac{\sum_j W_{ij} Q_j P^i(V_j)}{\sum_j Q_j W_{ij}}. \quad (2.7)$$

The final step is defuzzification, where all aggregation values (A_i from above) are ranked and the maximum value from the cost function for each class is selected as the most likely classification.

It is important to note the current limitations for classification of smoke using the WSR-88D system. Figure 2.2 shows an example of a wildfire in the hybrid hydrometeor classification field from the KSOX radar on January 7th, 2025. The current class for the smoke related to the wildfire (indicated in the red circle) is BS. This class includes a plethora of other scatterers like birds, bats, insects, chaff, and sea clutter. So in large BS classified areas, it is difficult to highlight the areas of smoke in comparison to other returns and provide information to the public about the areal coverage of smoke. Improvements to the BS classification and the detection of chaff, for example, have been explored using a variety of techniques as in [11], [10]. The importance of having a smoke class drives the need for improving the current classification algorithm and developing new techniques.

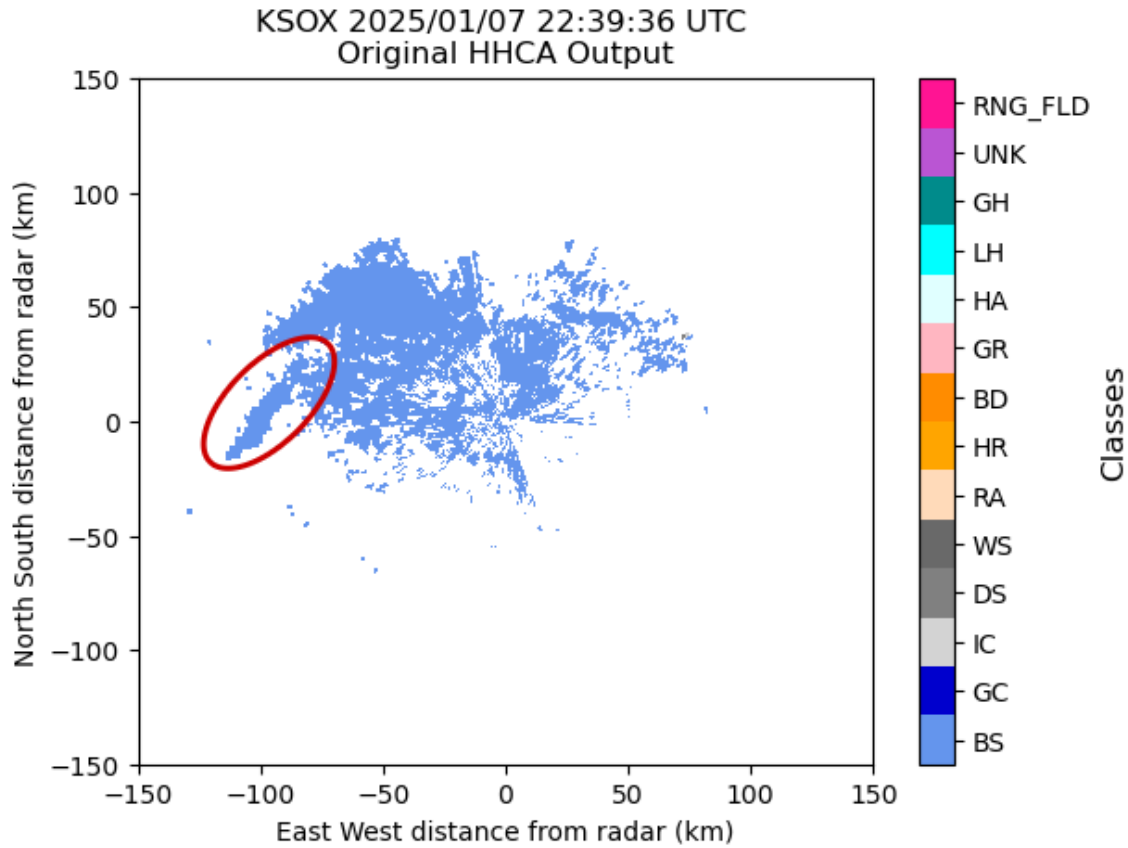


Figure 2.2: An example of the current output of the hybrid hydrometeor classification algorithm for the Palisades fire from the KSOX radar on 01/07/2025 at 22:39:36 UTC. The red circle indicates where there is a known wildfire and smoke plume.

Fuzzy logic has been the main choice of classification using dual-polarization data because it works well in domains where there is overlap between polarimetric variables, the relative ease of implementation, and easy addition of confidence vectors [25]. There has been much work in improving the current HCA and the implementation in the national network. However, there have been other methods proposed that revolve around machine learning and artificial intelligence. These models require labeled inputs, which are the verification or truth of the class at each resolution volume. Datasets used in machine learning models need to be quite large and they require intensive human inputs for the numerous classes necessary for classification of hydrometeors and other scatterers with weather

radars. This has limited the usage of machine learning in the past, in addition to more limited statistical information for all of the scatterers of interest. However, the introduction of smoke classification provides a new verification dataset, given that the statistics-based algorithm(s) perform with reasonable accuracy.

2.2 Machine Learning Classification Models

For this work, we are testing the viability of implementing machine learning for the classification of smoke using the most explainable methods rooted in previous literature. Human understandability varies as a function of model complexity (Figure 2.3), where more complex models typically have lower human understandability and need more post-hoc explainability methods. Based on Figure 2.3, choosing classification models like random forest and those that are less complex means that we can use the current post-hoc methods to increase human explainability for model output. The three main models for exploration in this work are support vector machine (SVM) via the `SGDClassifier` model (stochastic gradient descent classifier), decision trees of varying depth, and random forest. Thus, we investigate a wide variety of model complexity, while still falling in the preferred range of explainability.

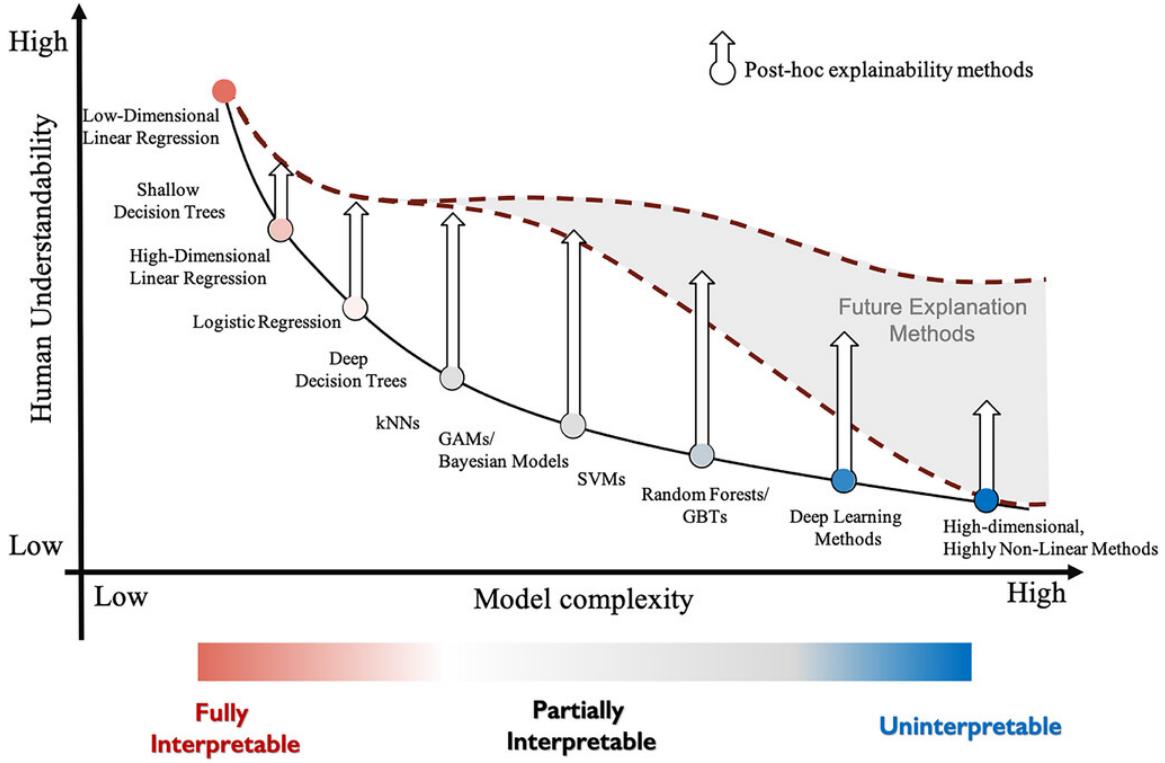


Figure 2.3: Chart for showing the intersections of explainability and model complexity for common machine learning models. Those with high human understanding and lower model complexity are beneficial for confirming the model is learning correct physical relationships (from Flora et. al 2024 [28]).

2.2.1 SGDClassifier and Decision Trees

SGDClassifier

To start off testing different machine learning models for this classification beyond the explainability/model complexity, methods used in Jatau et. al [20] were explored as a base and to provide comparisons to if necessary. In their article, they explored classification of birds and insects, which also fall under the BS classification umbrella. Their input features were normalized dual-polarization variables and their textures, in addition to a sector and range interval analysis. Similar variables were tested for this study, and were used as the features. Their classification methods included the SGDClassifier and DecisionTreeClas-

sifier from scikit-learn [29]. The first method tested was the `SGDClassifier` class, which is an estimator that uses stochastic gradient descent (SGD) and regularized linear models set by the user. For this specific case, the hinge loss function was selected. The hinge loss is used typically for binary classification, and is represented by:

$$L(f(x), y) = \max(0, 1 - y \cdot \hat{y}), \quad (2.8)$$

where L is the loss function, the $f(x)$ is the predicted value, $\hat{y}$ is the predicted score or output from the model, and y is the actual label (either -1 or 1). The hinge loss optimizes the distance or margin from data and a decision boundary. The maximum in the function helps to ensure and keep the largest margin between data points and the boundary while penalizing misclassifications. Incorrect classifications are when the predictions are on the wrong side of the decision boundary, where -1 is classed as 1 or vice versa. It also penalizes correct classifications that are close to the decision boundary (Figure 2.4).

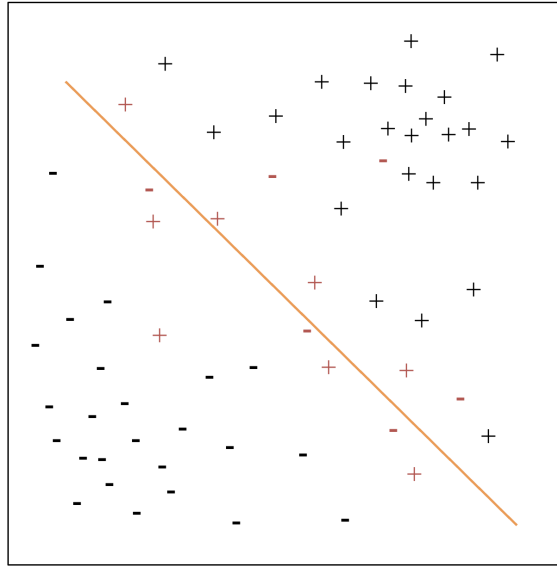


Figure 2.4: An example hinge loss case that uses pluses and minuses to indicate samples with regards to the decision boundary (orange line). Samples that would be misclassified are noted in red.

Additionally, the `SGDClassifier` uses a regularization term or a penalty, to help with overfitting. A regularization term helps generalize and reduce the complexity of the model. The default and recommended regularization term for `SGDClassifier` is the L2 regularization:

$$Loss = \frac{1}{n} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^m w_i^2, \quad (2.9)$$

where y_i is the actual target value, $\hat{y}_i$ is the predicted target value, N is the number of data points, m is the number of features, λ is the regularization parameter for the strength of the regularization, and w_i are the coefficients of the features.

The L2 or ridge classifier applies a penalty of the squared magnitude of the weights and is recommended for linear SVMs. The L1 and elastic net regularization terms are also options for this sklearn class. Both were tested and had little difference on the accuracy of the model, so the L2 default was kept. `SGDClassifier` uses the stochastic gradient descent optimizer. IT has updates for the different model parameters iteratively using each training sample for a gradient descent. In comparison, the regular gradient descent uses all of the samples to compute the gradient and update the model parameters. The training data is shuffled after each epoch which is suggested in the documentation for `SGDClassifier`. `SGDClassifier` has an early stopping and adaptive learning rate option as well, to help improve overfitting and the speed at which the global minimum is found.

Decision Trees

A second method tested in Jatau et. al [20] was decision trees. Decision trees use recursive partitioning of data based on values of the selected attributes, and split and group data together based on an internal decision, which creates branches. If the value of a specific feature j is below a threshold then, the left branch of the tree is followed, otherwise it

follows the right [30]. Further splits can occur until a stopping criteria is met, and at the end we reach the leaf nodes or the decisions for classification. The splitting criteria, or how “good” the split for a nodes is, decides how splitting is completed. The Gini impurity was used for decision trees in this work, which will be explained further in Chapter 4. The equation for the Gini impurity is given by:

$$I_G = \sum_1^C p_j(1 - p_j) = 1 - \sum_1^C p_j^2, \quad (2.10)$$

where p_j is the proportion of samples that belong to a class, C , for a specific node.

This index goes from 0 to 0.5 with 0.5 being the most impure, and is the likelihood that a randomly selected sample would be misclassified by that node. A node is where a splitting decision is made, and can be a root, decision or parent, or leaf node. The root node is the base of the tree, with the decision nodes branching off of that. Leaf nodes are the final, terminal nodes in the tree, where the predictions occur. The final class prediction at the leaf nodes are based on the majority class of the training data that is split to that particular leaf, via the previous decision splits. The more decision nodes there are, the deeper the tree becomes. Decision trees are prone to overfitting and with too much depth, can learn relationships that are specific to the training data. This means that the model will not be generalized to be used on the testing or validation datasets. There are numerous methods for tackling overfitting which include pruning, limiting tree depth, feature engineering, setting minimum number of samples in each leaf node, and cross validation.

2.2.2 Random Forest

Decision trees are considered to be weak learners, and are inherently simple and prone to overfitting. Trying complicated and computationally expensive methods like deep neural networks could yield better results, given a sufficient amount of labeled data. However, en-

semble methods are a good way to tackle the weak learner aspect of decision trees. Whereas the decision tree uses one single tree to describe the patterns and relationships in the training data, the random forest uses numerous low-accuracy trees that are shallow, which are then combined. Random forests use bagging to create subset copies of the training data, using random selections to create different, unique datasets for trees to be trained on. It can help to reduce the variance, and make the shallower trees less correlated and more generalized overall. Sklearn's `RandomForestClassifier` was used for this project, which builds on the `DecisionTreeClassifier` mentioned above, but for a specific number of trees. Bootstrapping aggregation (another name for bagging and a hyperparameter that can be specified) is used to construct a number of subsets of training data as discussed above. A decision tree model is then created based on each of unique datasets. When a random forest is given a classification prediction task, each decision tree is used to make a final class decision, as discussed in Section 2.2.1. The final results from all of the decision trees are combined via averaging or majority voting to have a final value or prediction. To help with hyperparameter selection, `GridSearchCV` was used to find the the number of trees, the minimum samples for a leaf, the minimum samples for a split, and the criterion to determine the quality of the split. This is explored further in Chapter 4.

Chapter 3

Detection and Classification of Plumes

3.1 Statistical Analysis of Smoke Plumes

3.1.1 Data

The data for this work comes from the WSR-88D data archive, accessed via the National Centers for Environmental Information. Initial cases selected come from the literature that discusses the fire of interest in depth, ensuring that the signatures analyzed come from a confirmed wildfire. Figure 3.1 shows the radars used for the cases studied in this work with red stars indicating wildfires and green stars indicating what is likely a prescribed burn. It should be noted that there are radars like the operational KTLX WSR-88D in Oklahoma City, where multiple fire cases from different years were obtained. During the 2024 and 2025 spring burn windows, a few fires were also noted via satellite imagery and were cross-referenced to the nearest WSR-88D radar. Additionally, wildfires were brought to the attention of the author by colleagues and peers and were added to the list if sufficiently close to a WSR-88D (indicated with red stars). Having both wildfires and prescribed burns ensures that a wide variety of topography, fuels, and seasons are used to provide the most general statistical overview of smoke plumes. A robust set of unique fires will give better results for new cases that are added to the list in the future.

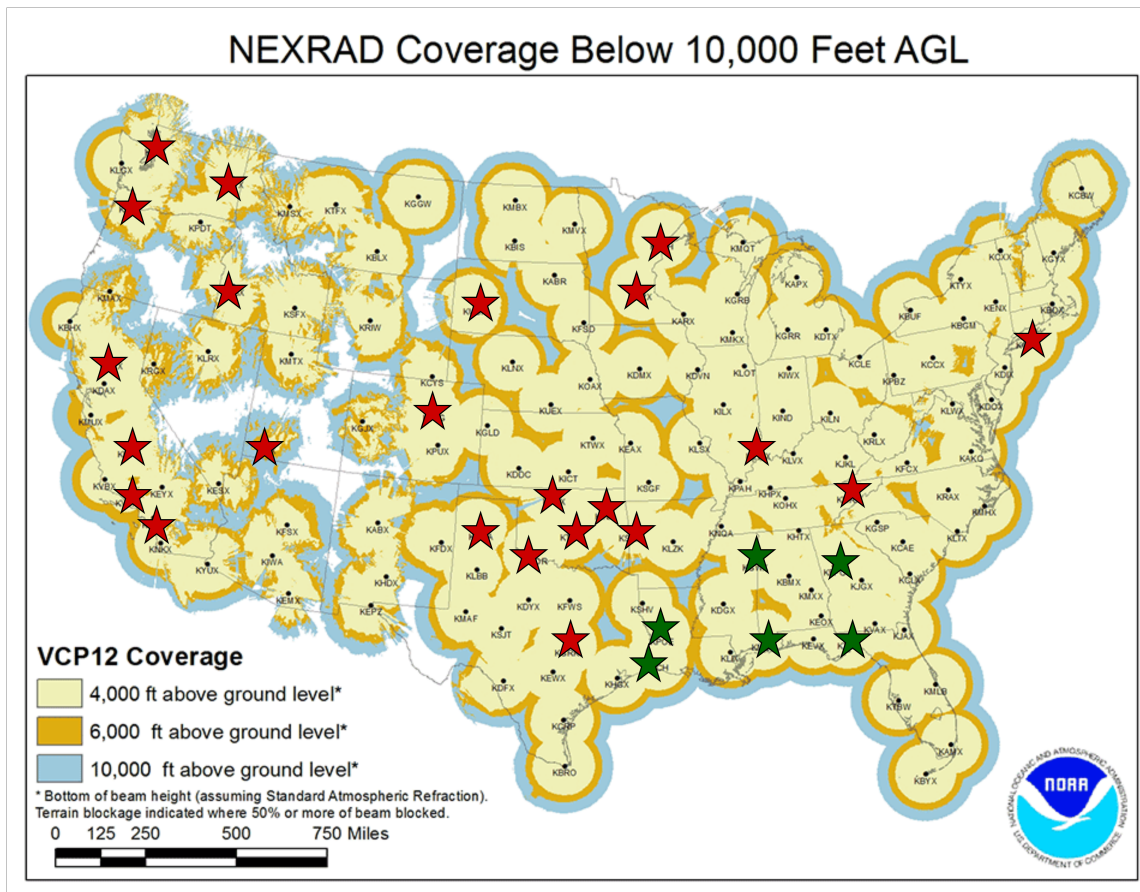


Figure 3.1: Radars used for analysis of wildfires (red stars) and prescribed burns (green stars) from across the country, using the NEXRAD coverage map as reference. Multiple radar locations were used for more than one fire.

3.1.2 Statistical Analysis and Plume Detection

The backbone of this work lies in the algorithm that was developed for detecting smoke plumes. It is driven by recursive improvements made to a series of thresholds that are based in the statistical analysis of smoke observed with S-band radars. The initial contouring algorithm uses a complex mask to remove all areas that are deemed to be “not smoke” As noted in previous literature, the correlation coefficient values of smoke are much lower than that of hydrometeors [31], [8], [32]. The initial threshold for correlation coefficient was set to 0.8, which is one of the upper bounds for the BS class membership function. This

separates the hydrometeors from the non-hydrometeors, where everything with a ρ_{hv} value greater than 0.8 is masked. Thresholds for reflectivity and differential reflectivity are also used in the detection algorithm. A reflectivity mask of Z greater than 1 dBZ is typically used to remove any erroneous weak echoes that could be included in the smoke plume. This will reduce where the detection boundary will likely be for the smoke plume, which can have very low reflectivities where there are smaller, less concentrated particles. There is a trade off as well, such that mountainous terrain and ground clutter, which can have similar polarimetric signatures are correctly masked. More on this will be discussed later, as the textures of Z and Φ_{DP} are added. Initially, Z_{DR} had a wide range that corresponded to the membership function for the BS, but also including negative values down to -3 dB. This was to ensure, initially, that Z_{DR} would not have too great an impact on detecting smoke.

As mentioned previously, the thresholds used for detection were improved recursively, typically as more fires were added to the case list. A case with limited complex terrain and a clear smoke plume was tested first for the detection algorithm so that it was easier to delineate between the smoke plume and the surrounding environment. A wildfire in Jackson County occurred on March 1, 2017 in southwestern OK, and one time step was selected to improve the initial thresholds that were based on previous literature. To accurately select the fire, the ρ_{hv} threshold was lowered to 0.6, and the Z_{DR} range was limited to encompass -2 to 8 dB. This helped to remove areas that were detected as smoke, but were actually a dense network of wind turbines. Once a sufficiently accurate contour was found based on human interpretation and the areal extent of ρ_{hv} , it was used as a mask to select all of the data that fell within the smoke contour. Figure 3.2 shows an example of the smoke contour after iteratively improving the thresholds. Histograms were used to determine statistical information about the distribution of each polarimetric variable for the area inside final contour. They showed the distributions and important values like mean, median, and range that could be tabulated for each time step, for each fire. These steps were repeated

for the fire cases shown in Figure 3.1 at a frequency of fifteen minute time steps during the times of interest. A size threshold is also built into the detection algorithm, which was also improved based on the size of the smoke plumes seen in the case list. The value for this is not necessarily tied to areal extent directly from the radar data itself, but rather the comparative size of the contours. A value of 80 km^2 was used to select the largest contour(s) to be included in the final output.

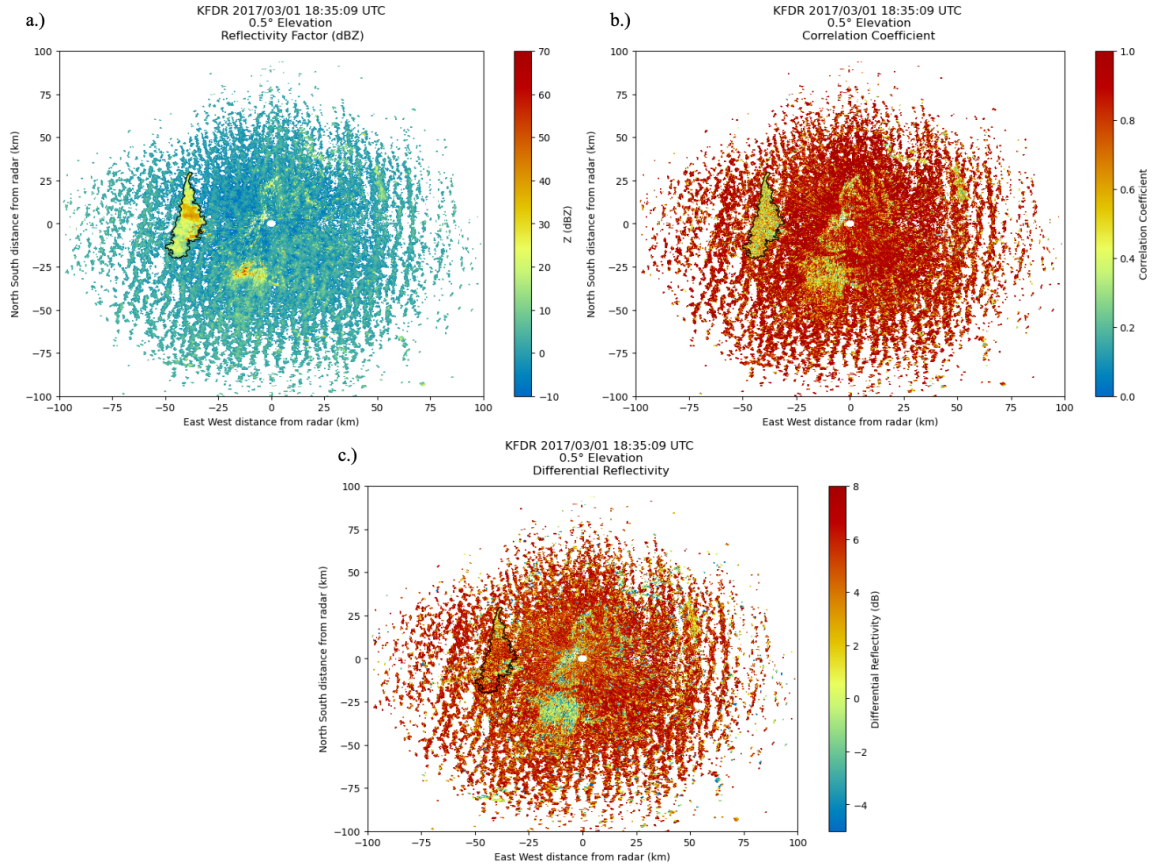


Figure 3.2: The smoke detection output on a.) reflectivity b.) correlation coefficient and c.) differential reflectivity for the March 1st, 2017 wildfire via the KVNx radar on 03/01/2017 at 18:35:09 UTC.

As more cases were added, the detection algorithm thresholds were qualitatively improved for a more general polarimetric overview of the smoke plumes. While this process was widely qualitative, which could be prone to subjectivity, the wide variety of cases con-

sidered helped to improve derived thresholds. The thresholds discussed in the paragraph above remained widely the same, though in more mountainous regions where there were dense fuels being burned, a lower ρ_{hv} threshold of 0.4 or 0.5 and a higher Z threshold could be used to improve incorrect detections. After testing around 25 fires, some patterns in the statistical values of each variable were noted. Reflectivity typically had a normal distribution with the mean anywhere from 10 to 18 dBZ. There were cases, however, where the fire was particularly intense for periods of time and a large concentration of particles were being lofted into the air or there were a larger amount of large particles. During these times, the distribution tended to be left skewed or had a bimodal distribution (Figure 3.3a and b respectively). By comparing the radar output at each time to the frequency histogram, the low frequency, high reflectivity areas represent the highest volume of particles typically over the active portions of the fire. The rest of the distribution characterizes the lofted ash and pyrometeor particles that are blown in the background flow.

Correlation coefficient and differential reflectivity distribution profiles show right and left skewed distributions respectively. Correlation coefficient distributions followed two different distribution types. One had a very right skewed distribution with the highest frequency values being 0.2-0.4 as seen in Figure 3.4a, and the other had a more normal distribution with a frequency peak around 0.5 (Figure 3.4b). This is supported by the findings in [31], where ρ_{hv} values peaked in frequency near either 0.2-0.3 or 0.5. The detection algorithm performed better for fires with timber or timber litter as their fuel type when the correlation coefficient threshold was lowered to 0.4 or 0.5. The typical range values and the peak in frequency supported the decision to have the ρ_{hv} threshold to be set at 0.6, where all values above this are not considered to be smoke. Differential reflectivity had a slightly left skewed distribution consistently, though the mean varied depending on time and size of the plume and could be anywhere between 2 and 5 dB. This indicates the vast majority of the plume had Z_{DR} values that are positive and have a more horizontal

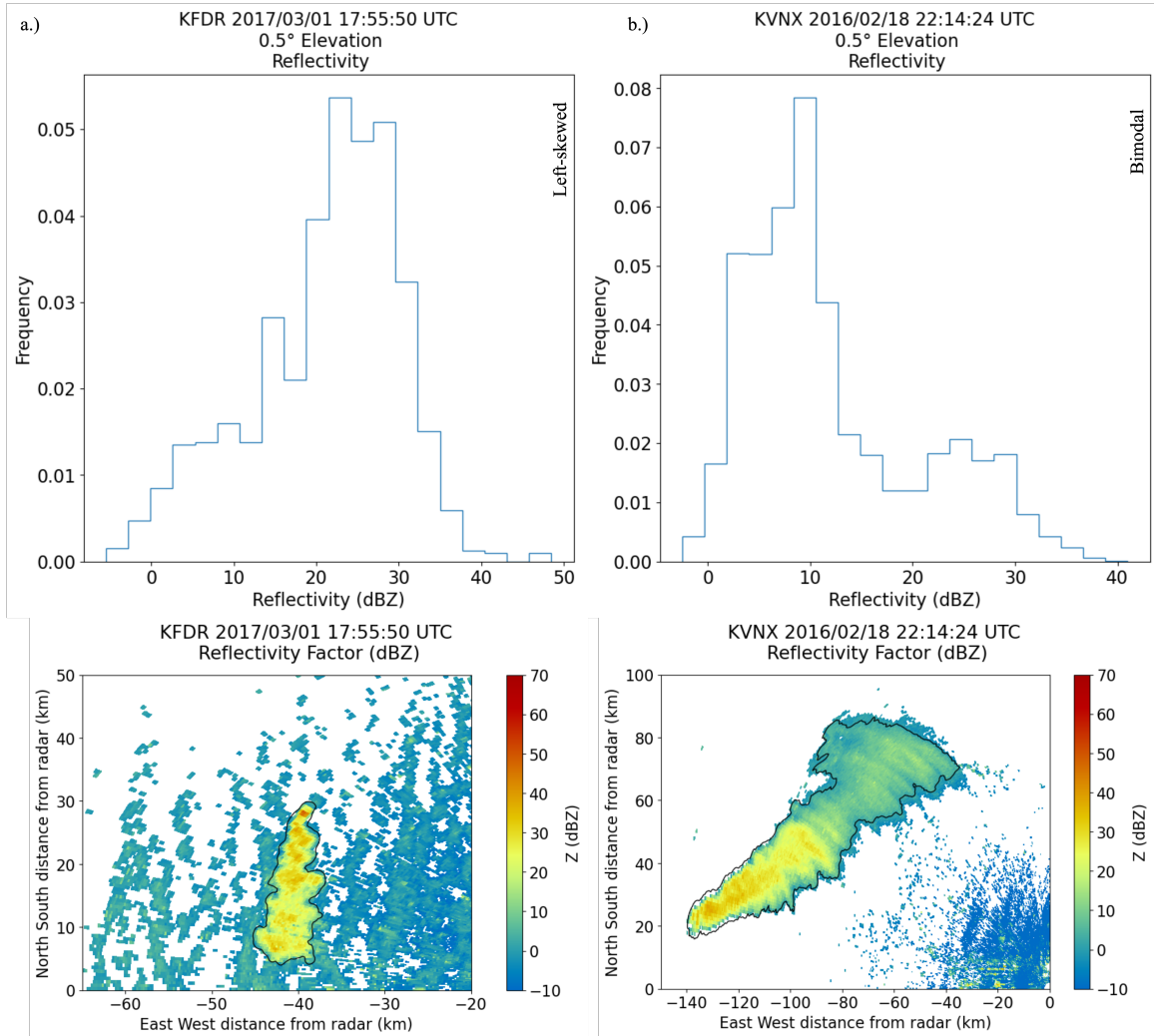


Figure 3.3: Histogram of Z from the 0.5° elevation angle in two cases for a.) a fire in Jackson County, Oklahoma from the KFDR radar at 17:55:50 UTC on March 1st, 2017, with the corresponding Z field with the black contour indicating the area of smoke detection below and b.) the Buffalo fire from the KVNx radar at 22:14:24 UTC on February 18th, 2016 and the Z field below to show the left-skewed and bimodal distributions respectively.

orientation. In terms of the Z_{DR} distribution, it also exhibited a pattern where Z_{DR} increased with distance away from the active fire. Specifically, Z_{DR} trended towards 5-8 dB in the smoke plume furthest from the fire. This could perhaps be due to preferential orientation in the background wind and the shape of the pyrometeor itself or the condensation of water vapor onto the surface of ash particles. There is also the factor of beam height for some

cases, as particles at further ranges will be higher up and may be oriented by the background wind field. Negative values and those close to zero were found to be near the active fire and the turbulent updraft, and could indicate on the radar image that that is where the active fire is. Figure 3.5 shows an example of an active fire contour in red where the lower Z_{DR} is used to indicate the fire location. There are, however, cases where there are lower Z_{DR} values away from the fire's location. Thus, this threshold would not have general use in a wide variety of cases. Perhaps this could be overcome with the assistance of integrated satellite data, where variables like the fire returned power (FRP) can help to characterize the amount of biomass burning, fire size, and temperature.

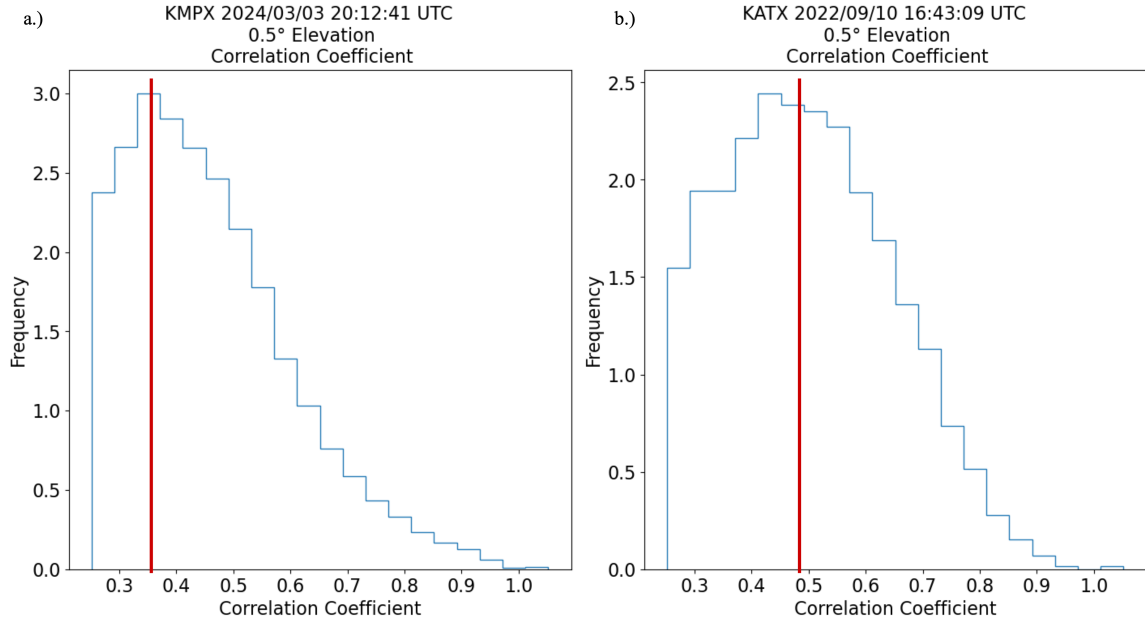


Figure 3.4: Histogram from the 0.5° elevation angle ρ_{hv} in two cases for a.) the Snake Trail fire from the KMPX radar at 20:12:41 UTC on March 3rd, 2024 and b.) the Bolt Creek fire from the KATX radar at 16:43:09 UTC on September 10th, 2022 with the red line indicating the mean of the distribution.

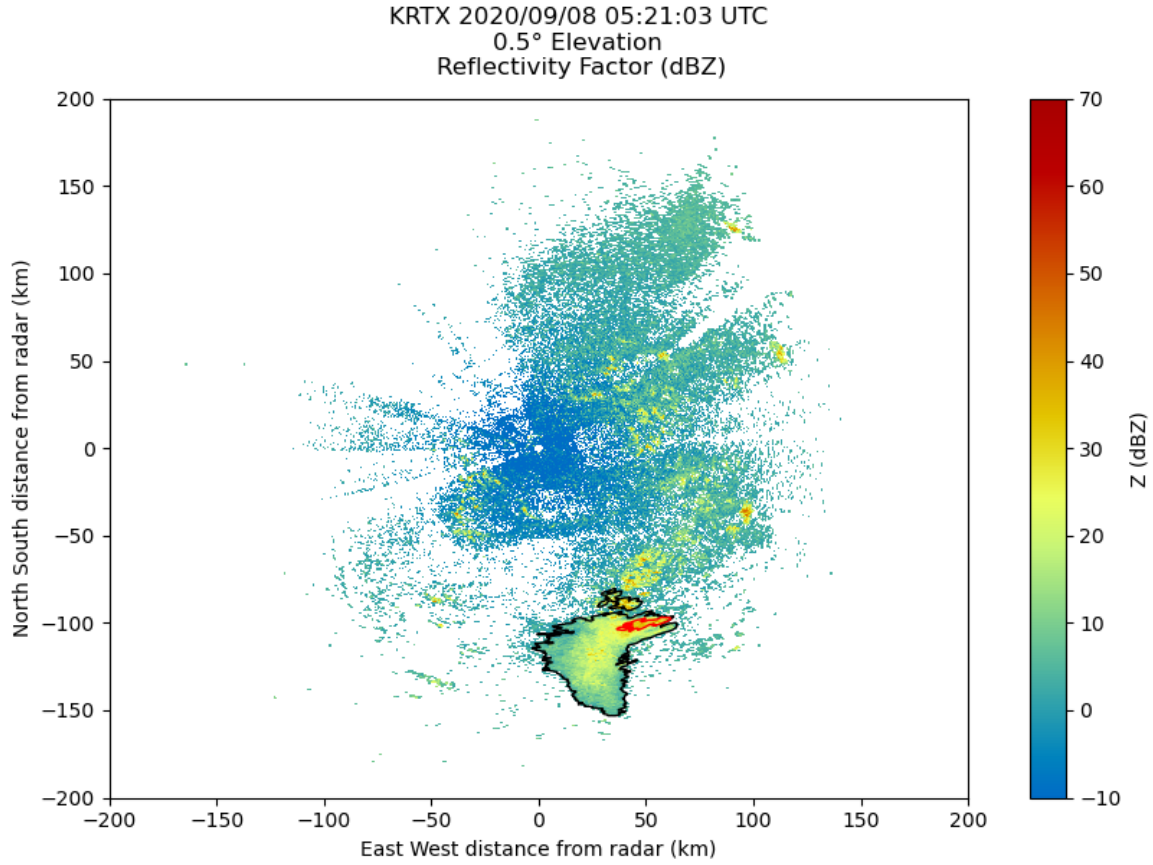


Figure 3.5: The smoke detection output in black and the active fire contour in red that highlights areas of lowest Z_{DR} , highest Z , and low ρ_{hv} the same as everywhere else in the smoke plume for 05:21:03 UTC on September 8th, 2020 from the KRTX (Portland, OR) radar.

Further improvements for the detection algorithm could be made with the addition of texture fields. The textures of reflectivity and differential phase are both used in the operational HCA product, as they are important fields for separating nonmeteorological classes like GC/AP and BS from the weather-related classes [33]. They typically have much higher values, so analyses of these fields were important for improving the algorithm for the removal of clutter and differentiating between smoke and other biological scatterers. To calculate these fields, the texture along ray function was used from the Py-ART module [26], and the window sizes that are noted in [33] were used for the standard deviation calculation.

The window size for $SD(Z)$ is 1 km and $SD(\Phi_{DP})$ has a window size of 2 km. The “SD” for these variables indicates that it is the standard deviation using the specified window along the ray, or radial. Once these variables were characterized by looking at 3-4 cases, an initial threshold of $SD(Z)$ greater than 6 dBZ is used to remove likely non-smoke targets. This threshold was beneficial in removing areas that were incorrectly detected as smoke where there is complex topography, which were determined via comparisons to noted active fires at the time and the radar output for each time step. Mountains will often produce high reflectivity and a low correlation coefficient which will look like smoke to the detection algorithm without the addition of the texture field for reflectivity. Figure 3.6 shows how the addition of $SD(Z)$ improves the detection for a case along the Cascade mountains in Oregon. Though it doesn’t entirely correct the mountain crest detection, it does significantly improve the overall areas of misclassifications. With analysis, the $SD(\Phi_{DP})$ variable proved to be extremely variable in a case to case basis. For some cases, values for $SD(\Phi_{DP})$ were around 10° - 80° and for most others ranged from 100° - 150° . It was difficult to select a threshold or range of typical values to be used in the detection algorithm, and was excluded from consideration at this time. Perhaps further investigation into this field, will yield a clearer result based on looking into fuel types, distance and orientation the radar, and so on.

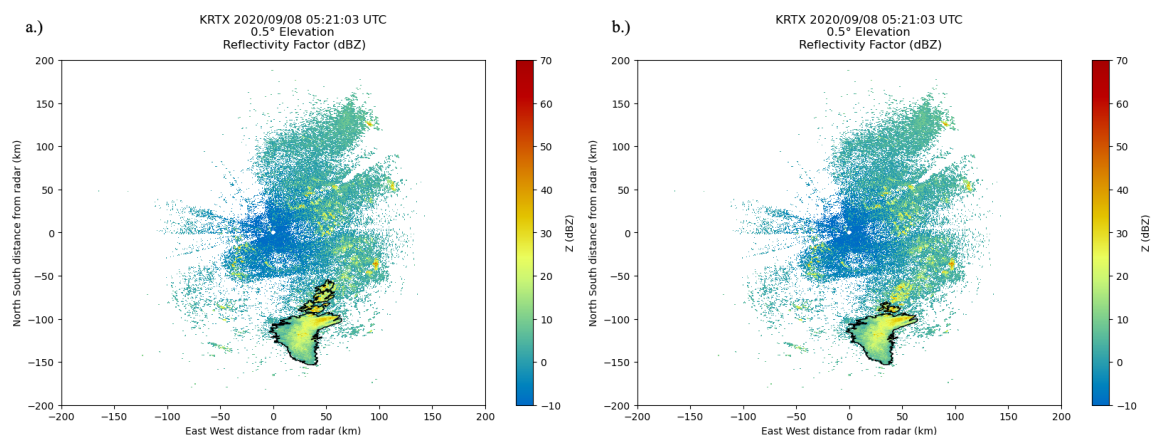


Figure 3.6: The detection of smoke with a.) no addition of the texture of reflectivity and b.) with texture of reflectivity added to the detection algorithm for the Beachie Creek fire via the KRTX radar on 2020/09/08 at 05:21:03 UTC.

3.2 Implementation for Fuzzy Logic Classification

One of the main challenges with using a traditional fuzzy logic method for smoke classification is the overlap with the biological scatterers category. Currently, smoke plumes are classified as biological scatterers (as in Figure 2.2) which typically describes insects, birds, chaff, etc. Table 3.1 shows the smoke membership function parameters as well as those related to the hydrometeor classification algorithm by [9] for reflectivity, correlation coefficient, differential reflectivity, $SD(Z)$, and $SD(\Phi_{DP})$, and Table 3.2 shows the weights. Table 3.1 shows that the biological scatterers and the initial membership functions for smoke overlap in some capacity with every dual-polarization variable. The statistical analysis portion yielded values to create the membership functions for a new class in a more traditional fuzzy logic hydrometeor classification algorithm. Correlation coefficient and differential reflectivity are the most relevant dual-polarization variables to determine membership function parameters for accurate smoke classification. These fields present distinct values when compared to precipitation. Differential reflectivity posed issues for classification of smoke, as the range from -2 to 7 dB encompasses many meteorological classes [25]. Because of the

overlap between dual-polarization variables, we tested two additional texture variables to separate smoke from the biological scatterers as described in the previous section. For the fuzzy logic portion of this project, both $SD(Z)$ and $SD(\Phi_{DP})$ are included as variables for a new smoke class. The general membership function for $SD(\Phi_{DP})$ is based on the greater 90° threshold in simple classification algorithm in [31], as the statistical methods portion showed a wide variety of values. The inclusion of this field provided limited improvement for classification, but was included to provide as similar results as possible to the current HCA implementation in the WSR-88D network. It thus was given a lower weight as shown in Table 3.2. It should also be noted that the output from this HCA implementation has a more limited melting layer algorithm, and so only cases with smoke and ground clutter will be shown for the sake of clarity.

	GC/AP	BS	SMK	DS	WS	CR	GR	BD	RA	HR	RH
$P[Z(\text{dBZ})]$											
x_1	15	5	5	5	25	0	25	20	5	40	45
x_2	20	10	15	10	30	5	35	25	10	45	50
x_3	70	20	40	35	40	20	50	45	45	55	75
x_4	80	30	50	40	50	25	55	50	50	60	80
$P[Z_{DR}(\text{dB})]$											
x_1	-4	0	-2	-0.3	0.5	0.1	-0.3	$f_2-0.3$	$f_1-0.3$	$f_1-0.3$	-0.3
x_2	-2	2	0	0	1.0	0.4	0	f_2	f_1	f_1	0
x_3	1	10	5	0.3	2.0	3.0	f_1	f_3	f_2	f_2	f_1
x_4	2	12	7	0.6	3.0	3.3	$f_1+0.3$	$f_3+1.0$	$f_2+0.5$	$f_2+0.5$	$f_1+0.5$
$P(\rho_{HV})$											
x_1	0.5	0.3	0.25	0.95	0.88	0.95	0.90	0.92	0.95	0.92	0.85
x_2	0.6	0.5	0.3	0.98	0.92	0.98	0.97	0.95	0.97	0.95	0.90
x_3	0.9	0.8	0.6	1.0	0.95	1.0	1.0	1.0	1.0	1.0	1.0
x_4	0.95	0.83	0.65	1.01	0.985	1.01	1.01	1.01	1.01	1.01	1.01
$P[\text{SD}(Z) (\text{dB})]$											
x_1	2	1	1	0	0	0	0	0	0	0	0
x_2	4	2	2	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
x_3	10	4	5	3	3	3	3	3	3	3	3
x_4	15	7	6	6	6	6	6	6	6	6	6
$P[\text{SD}(\Phi_{DP}) (\text{deg})]$											
x_1	30	8	90	0	0	0	0	0	0	0	0
x_2	40	10	110	1	1	1	1	1	1	1	1
x_3	50	40	140	15	15	15	15	15	15	15	15
x_4	60	60	155	30	30	30	30	30	30	30	30

Table 3.1: Parameters of the membership function for the hydrometeor classification algorithm as outlined in Park et. al (2009) with the addition of a smoke class derived from statistical analysis. The functions $f_1 - f_3$ are defined in Equation 2.1, 2.2, and 2.3.

Class	Z	Z_{DR}	ρ_{HV}	$SD(Z)$	$SD(\Phi_{DP})$
GC/AP	0.2	0.4	1.0	0.6	0.8
BS	0.4	0.6	1.0	0.8	0.8
SMK	0.2	0.6	1.0	0.8	0.2
DS	1.0	0.8	0.6	0.2	0.2
WS	0.6	0.8	1.0	0.2	0.2
CR	1.0	0.6	0.4	0.2	0.2
GR	0.8	1.0	0.4	0.2	0.2
BD	0.8	1.0	0.6	0.2	0.2
RA	1.0	0.8	0.6	0.2	0.2
HR	1.0	0.8	0.6	0.2	0.2
RH	1.0	0.8	0.6	0.2	0.2

Table 3.2: Matrix of weights as outline in Park et. al (2009) and [33] with the addition of smoke weights derived from statistical analysis.

Figure 3.7 shows an example of the HCA implementation for one of the two cases that will be discussed in this section, as well as in Chapter 5. There are issues of classification between the new smoke class and other biological scatterers, based on an objective first glance. The HCA did a sufficient job at accurately classifying the smoke plumes from the numerous fire fronts. This is likely driven by the lack of other scatterers in the scan, like other biological scatters (insects for example), and more limited amount of clutter around Amarillo, TX and a region of low reflectivity scatterers to the south. However, when considering these areas of clutter and other biological scatterers to the south, there are now areas that are incorrectly classified as smoke. So, while the smoke is correctly classified in most cases, new areas with an incorrect class driven by the inclusion of the smoke class appeared. In short, this is creating the opposite classification of what we are attempting to solve.

The same issue occurs in Figure 3.8b. In this particular case, there is a decent amount of clutter and returns from complex terrain. These features will often have high reflectivity and lower correlation coefficient returns, which can be similar in value to those produced associated with smoke. So we see areas of ground clutter and other biological scatter being

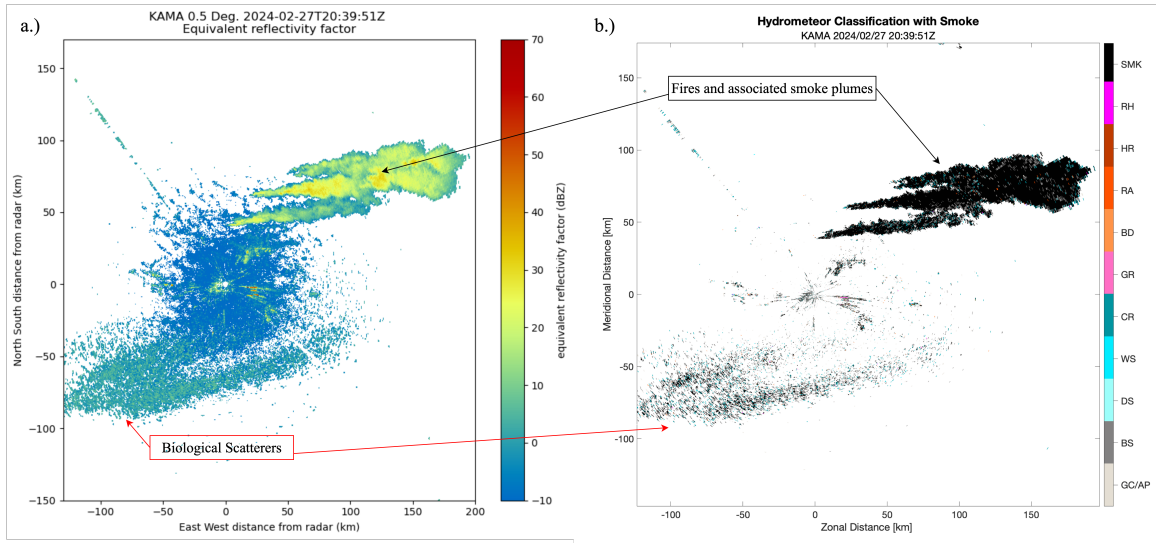


Figure 3.7: Smoke observations in fields of a.) the reflectivity field for the time of interest for the Smokehouse Creek fire, and b.) the HCA output with the inclusion of smoke (black) for a case with multiple wildfires on 02/27/2024 at 20:39:51Z from the KAMA (Amarillo, TX) radar. The dark gray denotes areas that are classified as biological. There is a noted overlap of areas still classified as BS in the smoke plume, and areas of clutter or BS now being classified as smoke.

classified as smoke now, similar to the previous case. In addition to the poor performance in areas that are not smoke, the fuzzy logic classification algorithm performed worse in the smoke plume regions as well. This is likely driven by the wildfires having a lower mean ρ_{hv} value and other influencing factors like terrain in this particular case.

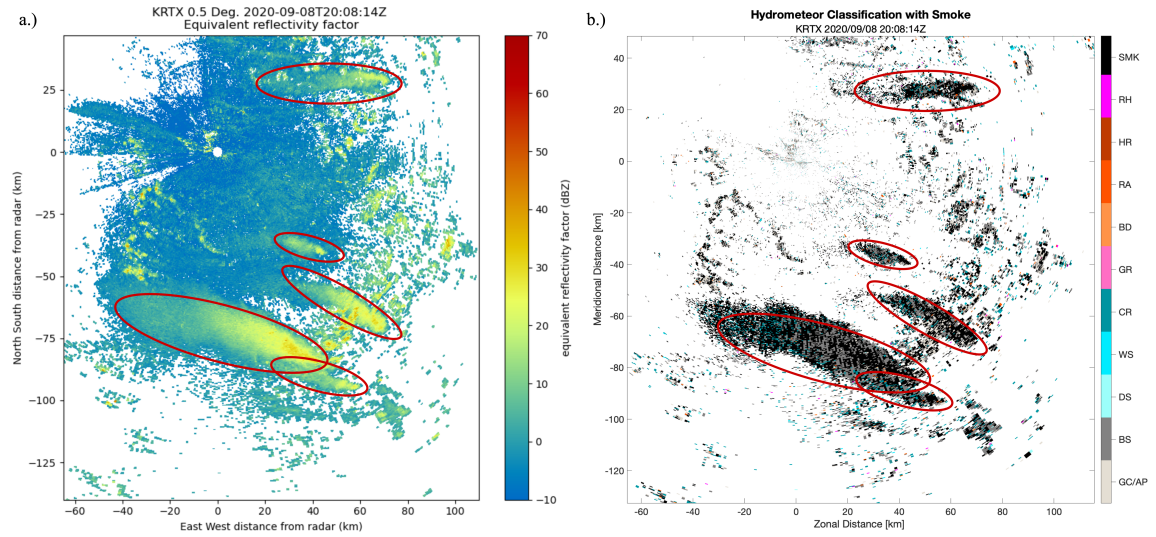


Figure 3.8: Smoke observations in fields of a.) the reflectivity field for the time of interest for multiple wildfires (red circles), and b.) the HCA output with the inclusion of smoke (black) on 09/08/2020 at 20:08:14Z from the KRTX (Portland, OR) radar. The dark gray denotes areas that are classified as biological. There is a noted overlap of areas still classified as BS in the smoke plume, and areas of clutter or BS now being classified as smoke.

3.3 Detection and Replacement Algorithm (DRA)

While the smoke class in the fuzzy logic performed sub-optimally, the detection algorithm discussed previously performed well at detecting areas of smoke using dual-polarization variables prompting further exploration of this algorithm and its application to classification.

The detection portion of this algorithm is developed using the threshold-based algorithm discussed further in Section 3.1.2. The detection algorithm needed for the statistical analysis is done using the native “C/FRadial” format. However, an important input needed for the detection and replacement algorithm is the Level-III data, specifically the hybrid hydrometeor classification algorithm (HHCA). Level-II and Level-III data are specific descriptors for various levels of radar data processing. Level-II describes the moment data that come from the original I-Q data, and are the “classic” polarimetric variables like Z or ρ_{hv} .

Level-III describes derived products from the Level-II data, like the hybrid hydrometeor classification algorithm that is needed for this work. The hybrid hydrometeor classification algorithm data is available with 16 levels and a range of 230 km. The Level-II data have a different maximum range value making masking and working with both sets of data in tandem very difficult. To work around this problem, both sets of data were regridded to a common grid. The grid has a resolution of 500 m with the x and y -limits going to 200 kilometers, and uses the nearest neighbor method is used for regridding. Both the resolution and limit values can be changed based on the case of interest, but increasing the range of the grid or improving the resolution came at the expense of run time on a local computer. The Level-II data could be gridded using the map gates to grid algorithm and the nearest neighbor method, which is typically faster. However, due to the nature of the values for the Level-III data (0 to 150 increments of 10) using this algorithm and the nearest neighbor weighting function meant that much of the data were lost. The Py-ART `map_to_grid` gridding algorithm is used for the Level-III data to combat this.

The statistical analysis helped to recursively improve the detection thresholds for the fuzzy logic HCA portion of this project. The initial detection algorithm was adapted to ingest the gridded data as opposed to the native polar coordinates. The final output indicates where the smoke plume and fire are likely located, as shown using the black contour in Figure 3.5, for example. For this work, the first, lowest level of the grid was investigated, though the algorithm continues to detect smoke at higher levels. To take the detection a step further and improve the visual aspect for clarity, the replacement portion of the algorithm was created. All of the points that are included inside of the contour(s) can be selected using the `Path` and `PathPatch` method from the `matplotlib` package in Python [34], and a new mask can be made where smoke is being detected. Once this is complete, the original data array and mask are called separately from the Level-III file. The smoke mask can be applied to the original data array in tandem to replace all of the values where the

smoke mask is indicating smoke with a new unique numeric value. This value can then be attributed to a smoke (SMK) class and plotted using the new data array and the original HHCA mask. Thus we retain all of the information from the original data that might be hydrometeor or ground clutter related classes, but now include smoke. Figure 3.9 shows an example of how this works for a single time step. This algorithm was then automated, and simply needs Level-II and Level-III data with the same time steps between files. It also runs for times where there is no smoke in the scan; the detection portion of the algorithm will skip to the next time.

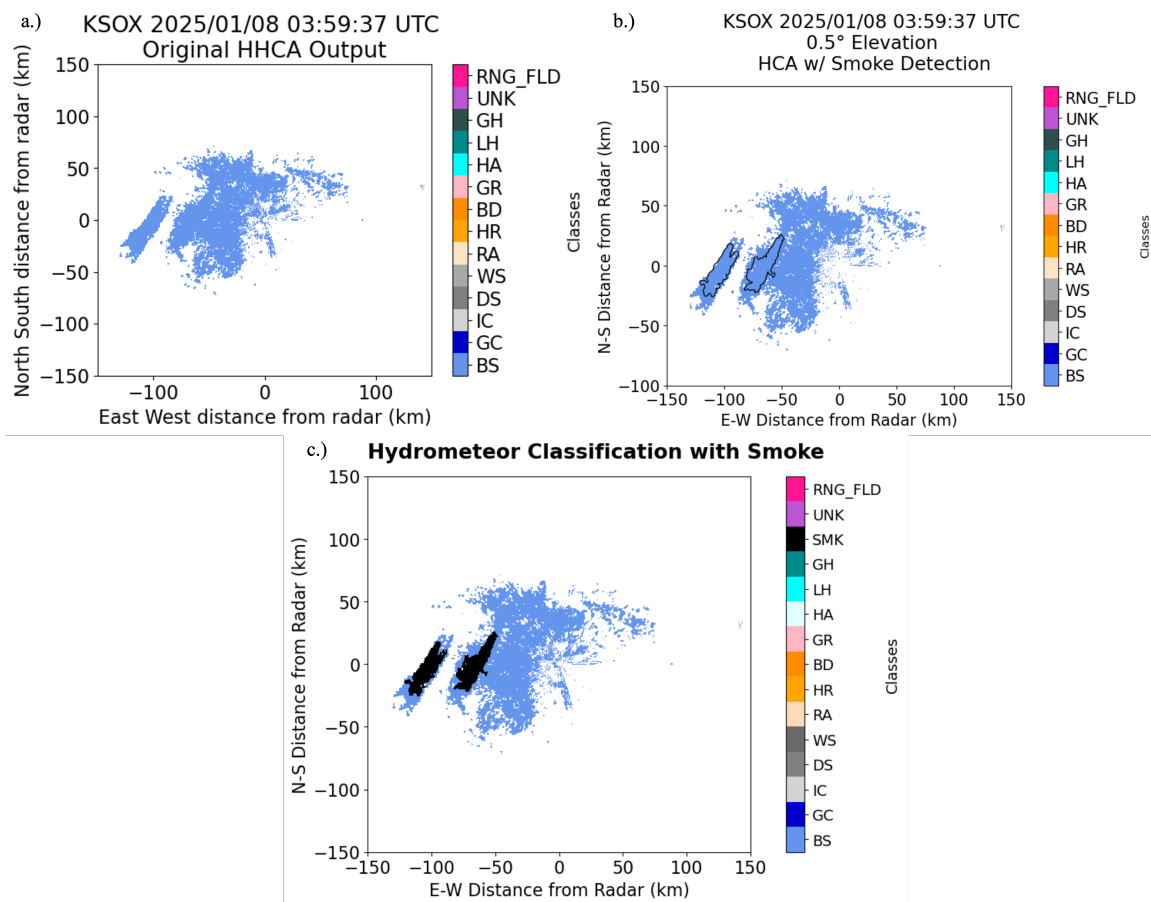


Figure 3.9: An example of the workflow for the detection and replacement of values for smoke using the hybrid hydrometeor classification algorithm where a.) the original HHCA output b.) the original HHCA output with the smoke area detected, and c.) the final output of the algorithm where smoke is black on the colorbar. This figure shows the Palisades and Eaton fires in Los Angeles on January 8th, 2025 at 03:59:37 UTC.

3.3.1 Accuracy Testing

Typically, the accuracy of an algorithm can be tested using the counts of the true positives, false positives, true negatives, and false negatives. However, the final value can be biased by one of the classes in a binary classification problem having significantly more counts. To quantify the accuracy of the detection and replacement algorithm, we use the receiver operating characteristic curve, which use the true positive rates (TPR) and false positive

rates (FPR) and are not influenced by having an imbalance of the counts of the two classes.

To quantify areas that are correctly or incorrectly classified as smoke, manual annotations of the fire and smoke plume are created. The GIS software, QGIS, was used to create shapefiles containing a series of contours or a single contour, where smoke was determined to be by a human user. This was repeated for all of the times of interest for the fires explained in depth in Chapter 5. The native latitude-longitude format from QGIS for evaluation was to a Cartesian grid used in the detection and replacement algorithm. The metadata from the NEXRAD file contains the latitude and longitude of the radar of interest, and was used as a reference for the shapefile latitude-longitude format to convert to a Cartesian grid. This was done using the `geographic_to_cartesian_aeqd` function from the Py-ART package in Python [26], which uses an azimuthal equidistant (aeqd) map projection. The shapefile contents were then overlaid onto the smoke HCA output to get the counts for each class inside and outside of the contour(s). Using the counts, a confusion matrix was used to visualize the necessary values needed for the ROC curves. To create a ROC curve, a variable or threshold in the algorithm is varied to test the changes in output and the true positive rates and false positive rates are calculated using the following equations:

$$TPR = \frac{TP}{TP + FN}, \quad (3.1)$$

and,

$$FPR = \frac{FP}{TP + FN}. \quad (3.2)$$

For this application, the ρ_{hv} variable threshold was chosen due to its importance for distinguishing smoke from other meteorological classes. A loop was created to run the detection and replacement algorithm eight times for each time step (values of 0.3 to 1.0

in increasing increments of 0.1), each with a different ρ_{hv} value as the threshold for the detection algorithm. For each iteration of the detection algorithm, a combination of true positive rate and false positive rate were calculated and were used to create the final ROC curve. The TPRs and FPRs at each ρ_{hv} threshold increment were also averaged to create an average ROC curve for the fire case of interest. Some time steps will perform better depending on various factors like plume size, VCP selections and radar sensitivity, etc. and so seeing the contributions to the final average was important.

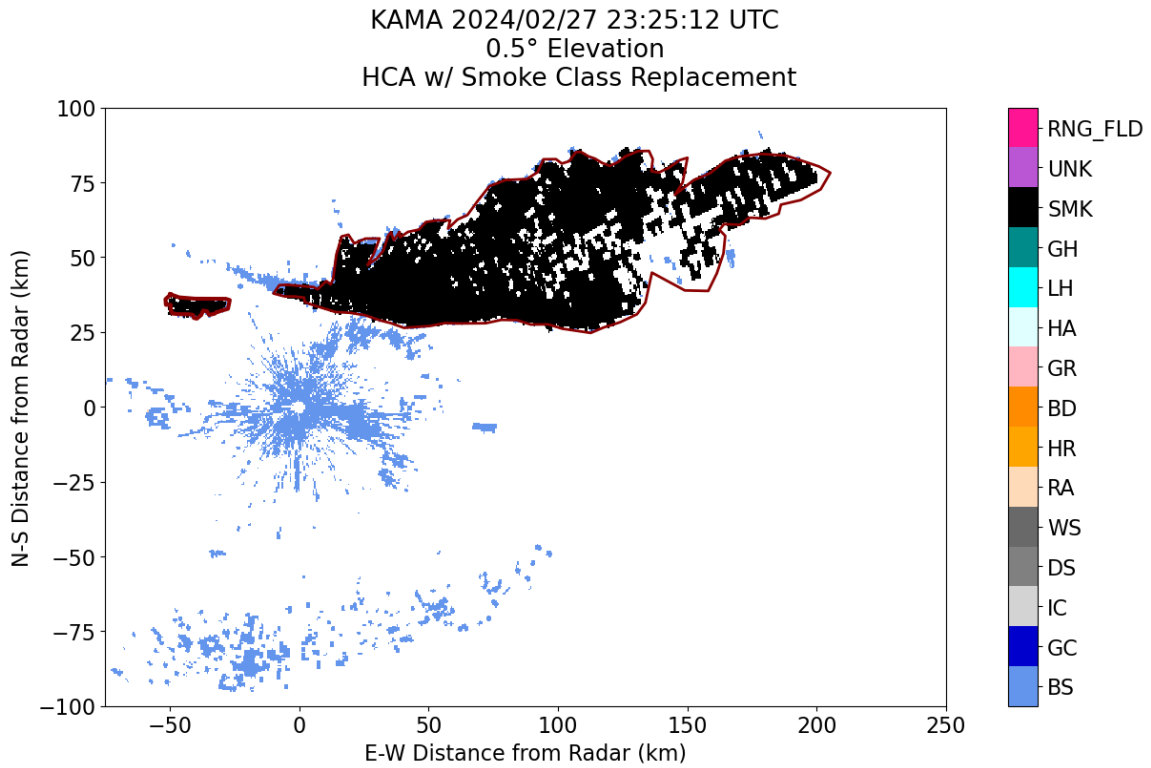


Figure 3.10: An example of the the shapefile overlay on top of the final output of detection and replacement algorithm for the Smokehouse Creek fire on February 27th, 2024 at 23:25:12 UTC as seen from the KAMA radar.

3.3.2 Limitations and Further Applications

Cases with other biological scatterers needed to be tested to ensure that areas of other biological scatterers like birds, insects, or chaff were not being incorrectly classified as smoke. To do this, cases from [20] and [10] were selected to be passed through the detection and replacement algorithm. Figure 3.11 shows an example output from three cases (two chaff, one bird) at selected times. The black contours indicate that there is still some overlap in detecting other biological scatterers as smoke using this method. There is still overlap in the important dual-polarization fields like Z , Z_{DR} , and ρ_{hv} . However, the field of $SD(\Phi_{DP})$ might be beneficial to add with further investigation. The values of $SD(\Phi_{DP})$ were found to be below 80° for the birds case as well as an insect cases which was investigated but not shown in Figure 3.11. The insects case was from the KBMX (Birmingham, AL) radar at 00:00 UTC on April 19th, 2016. Because insects and birds are typically much wider in the horizontal than vertical, and can move regardless of the wind speed and direction (though less so for insects), there are large enough differences in the dual-polarization variables when compared to smoke. Thus, the detection algorithm performed better determining that these areas were not smoke. Chaff, however, can look more similar to smoke in terms of particle sizes and aspect ratios, and the detection algorithm performed poorly, such that it detected these areas of chaff as smoke. The values of $SD(\Phi_{DP})$ were typically higher for chaff around 100° - 150° which is fairly similar to smoke in most cases. Thus, a threshold for $SD(\Phi_{DP})$ would not be as effective for eliminating chaff as an incorrect detection.

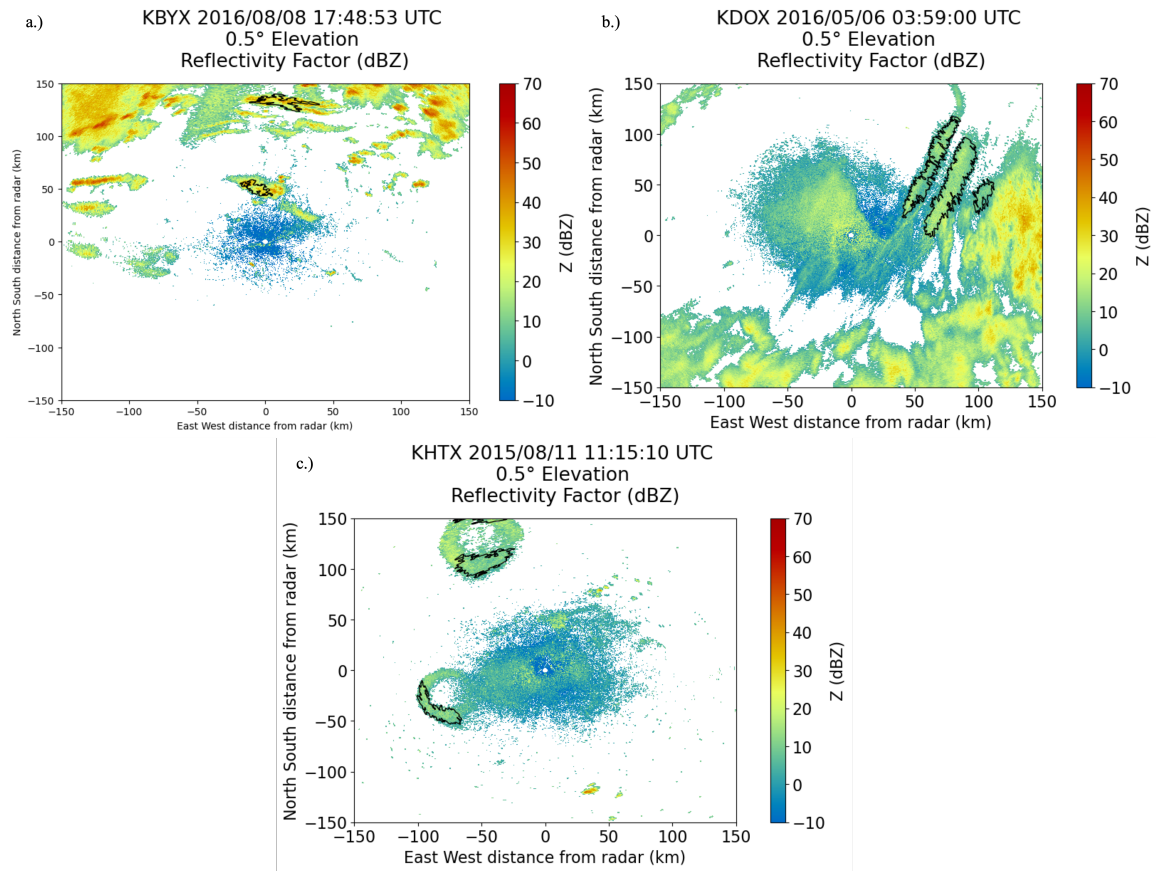


Figure 3.11: An example of other biological scatter cases when passed through the detection algorithm for a.) chaff case from KBYX (Key West, FL) at 17:48:53 on August 8th, 2016, b.) another chaff case but from KDOX (Dover AFB, DE) at 03:59:00 UTC on July 6th, 2016, and c.) a case with birds from the KHTX (Huntsville, AL) radar at 11:15:10 UTC on August 11th, 2015.

To combat the above issue and improve the algorithm for operational use, one of the methods to mitigate issues with the other biological scatterers would be adding a machine learning component. While the dual-polarization distributions might look similar between chaff, birds, and smoke, their shape and structure on radar is very different. For example, smoke will come from a single point and blow down wind, whereas chaff is transient and moves entirely with the background flow. Birds are typically seen en mass when they are leaving roosts which typically has a circular shape on radar. A complex machine learning model that detects the shapes and motion of the target would be able to separate these

three biological scatterers and maintain a strict smoke class. Additionally, adding satellite observations might aid in further improvements for the incorrect classification of other biological scatterers as smoke. Other algorithms and products like the newly unveiled NOAA Next Generation Fire System (NGFS) use GOES satellite imagery to indicate areas with fire-related returns. There are many channels included in satellite imagery like Fire Temperature RGB, shortwave infrared, and microphysics RGB that can help to support fire detection. Other biological scatterers like birds, chaff, or insects will not appear in these channels, so the lack of positive tag of a fire in either of the previously listed channels could indicate that the smoke class should be disregarded. This is also beyond the scope of this work currently, but with further improvement this could also have further implications for accurate sub classification of the different biological scatters. An additional suggestion would be the integration or combination of this detection and replacement algorithm with other models created to classify an inanimate class, or birds and insects as in [11] and [20]. These already include some aspects of machine learning inputs and the combination of multiple methods may provide additional skill improvements.

There are opportunities for the expansion of the detection and replacement algorithm to explore modeling and verification for other datasets. It is possible to show the change in the areal smoke coverage by comparing the old and new frame grid point (or pixel) counts for the smoke class. If the number of smoke grid points has increased in the new frame when compared to the old, then the smoke plume is likely growing in size. The reverse is true if the smoke plume is decreasing in size. A further concept that could be added is modeling the change in smoke coverage for the next hour or two based on previous data; this would further enhance the application for air quality, emergency managers, and the public. Examples of this concepts are included on a case by case basis in Chapter 5. Any indications of what will possibly happen can only help in improving communications and situational awareness. Further, this algorithm can be used to verify areas of smoke in

comparison to modeled data for smoke. Further development and analysis will be needed, but the possibilities extending from the detection and replacement algorithm could have promising results for improving modeling and also creating datasets to be used in machine learning.

Chapter 4

Machine Learning Models for Smoke Classification

4.1 Building the Dataset

The most challenging hurdle when using machine learning for the classification problem presented in this work was the lack of a dataset including smoke as a class. The smoke DRA discussed in the previous chapter provides the necessary labeled data or “truth” for indicating if there is or is not smoke at each grid point. To create a dataset, the smoke detection algorithm was used to generate a new file at each time step for each fire of interest. The DRA provides the labeled data for each time file, by adding a new smoke class to the original HCA where the smoke thresholds were met. To keep variables in a similar format for each time step, the final gridded variables were assembled into a netCDF file. A total of twenty variables are contained in each file: the x and y grid (801 x 801), masked reflectivity, differential reflectivity, correlation coefficient, texture of reflectivity, texture of differential phase, the original NEXRAD HCA output, and the smoke DRA output.

Once files were created for all time steps for fires, they could be combined to create a dataset that includes the gridded data for numerous fires. The time and date were added as new variables to distinguish which data were selected for the test/train split, but were not included as features for training. With a large enough dataset, the date and time could influence the performance of machine learning algorithms because fire weather, like severe

storm conditions, can be maximized at specific times of day and periods of the year. Machine learning models could identify these spatiotemporal relationships moving forward. However, with this smaller dataset, the date and time are not included to avoid the ML tools learning these relationships based on restricted cases. Once all of the included datasets have been put into one final xarray dataset, it can be used for the separations of labeled data and features. All of the radar-derived variables (reflectivity, differential reflectivity, correlation coefficient, texture of reflectivity and differential phase, and the two HCA options) are masked to create nans where there are no returns. To select data that were not masked, all rows that contain an HCA value of 10 (BS) are selected. As mentioned previously, smoke is currently classed as BS, with a value of 10. In essence, this turns a complicated multi-classification task into a binary classification. For example, in complex cases where there is weather in the same scan, which have values that range from 20-140 in increments of 10, the ML tools may be confused when learning to make the correct decision between smoke or keeping the class as BS while also having to predict hydrometeor classes. Once all rows with the BS class are selected, the labels and features data frames can be made. The label, or truth, is the smoke class replacement algorithm output which uses a value of 150 to denote smoke. Thus, the two classes will be binary, with 10 or 150 and 1 or -1 for the SVM. Because the original values from the data are so far apart in numeric values, using these in the SGDClassifier with the SVM option might influence the final output and how the model selected the correct hyperplane and split between the data. The features are reflectivity, differential reflectivity, correlation coefficient, and the texture parameters of reflectivity and differential phase, where the class was originally BS. The five-feature data frames often contains nans, and are filled with the mean of the variable, respectively, and normalized using the training data as all the variables have widely different dynamic ranges. To test the effectiveness of including only certain features, four different feature selected datasets were created. They are as follows in Table 4.1.

Scenario	Selected Features
1	$Z, Z_{\text{DR}}, \rho_{\text{hv}}$
2	Scenario 1 + $\text{SD}(Z)$
3	Scenario 1 + $\text{SD}(\Phi_{\text{DP}})$
4	All Features

Table 4.1: Feature scenarios with different included features for comparison in model training and testing.

4.2 Hyperparameter Selection and Training

For the SGDClassifier, first attempts at fitting and predicting smoke yielded relatively poor results. The accuracy for the model only reached slightly higher than 55%, which means the performance is only slightly better than that of a random classifier. Attempts were made to improve this model, as Jatau et. al [20] found this to be the best model for classifying birds and insects in their study. Specific model hyperparameters were tested by looping through a set of options to determine the best combinations of maximum iterations and the initial learning rate, with early stopping and an adaptive learning rate being used to prevent overfitting. Given an initial learning rate that was either 0.001, 0.01, or 0.1, the maximum iterations were increased from 1000, 100,000, and 1,000,000 to see the influence on model accuracy. The max iterations were found to have a limited influence on the accuracy of the model for predictions. The middle value of 100,000 iterations was used, and a starting learning rate of 0.01 was found to be the best for improving the model. The L2 penalty was used, as was the balanced class weight, since the dataset is likely imbalanced with more biological scatter grid points being present. Given the initial low performance, the SGDClassifier model was neglected for the rest of this work, though it should be noted that with different data selection techniques and perhaps non-gridded data, this model could be improved.

Other machine learning models also have numerous “knobs” which correspond to hy-

hyperparameters that can be determined by the computer or can be set by the user. There are numerous methods to help determine the best value for each hyperparameter based on the size of the dataset, one being, iteratively looping through a range of values for each hyperparameter in the training dataset, and seeing how performance metrics improve or worsen. For the deep decision tree and random forest models, the GridSearchCV function was used to determine the best hyperparameters for model performance. This method does an exhaustive search for the best hyperparameters given a user-set input grid of values to test for each hyperparameter. Given that this dataset is not too large, this brute force method is possible without significant impact to computing time and power. In this case, the initial GridSearchCV is run on the dataset that includes all variables, with a grid option being set for the maximum number of features that are suggested to be included. The grids are shown in Table 4.2 for both the decision tree and random forest models, with the final selections made by GridSearchCV included as well. In addition to general hyperparameter tuning, differing splits for testing and training data can be tested by changing the ratio of amount of data reserved for testing or training. The 80/20 random split is typically used for testing and training data, and was selected in this case over a 90/10 or a 70/30 split which had negligible impacts on the overall accuracy of the models.

Hyperparameter	Options	Selected Value
Decision Tree		
Criterion for Splitting	Entropy, Gini	Gini
Max Depth	5, 10, 25, 50, 100	10
Max Number of Features	1, 2, 3, 4, 5	4
Max Number of Leaf Nodes	1, 5, 10, 20, 50, 100	100
Minimum Samples per Leaf	1, 5, 10, 20, 50, 100	20
Minimum Samples per Split	2, 5, 10, 20, 50, 100	2
Random Forest		
Number of Estimators (Trees)	5, 10, 25, 50, 100	100
Criterion for Splitting	Entropy, Gini	Entropy
Max Samples for Bootstrapping	10, 20, 50, 100, 300	300
Max Number of Features	1, 2, 3, 4, 5	5
Minimum Samples per Leaf	5, 10, 20, 50, 100	5
Minimum Samples per Split	5, 10, 20, 50, 100	5

Table 4.2: The hyperparameter grid options and the final selected value after using GridSearchCV, with the final values being used in new decision tree and random forest models.

It's often best to let the computer select the function for node purity for the model. The GridSearchCV selected the entropy function as the best function to measure the quality of node splits for the random forest model. The entropy function is defined as the following:

$$I_E = - \sum_1^C p_j \log_2(p_j). \quad (4.1)$$

This is different than what was used for the decision tree section. In classification problems,

entropy measures the disorder, randomness, or uncertainty in the samples for the nodes. And like with the Gini impurity, we want to minimize the entropy in each node. For a two class classification, the entropy will span $[0,1]$, where 1 is the most impure value for a node and 0 is pure.

4.2.1 Decision Tree Models

The GridSearchCV decision tree model (DT1), with the above hyperparameter selections indicated that the maximum number of features to be used was 4. To test which combinations worked best in improving the accuracy of the DT1 model, all four training and testing datasets were used to fit and predict values to compute an area under the curve (AUC) value. AUC is continued to be used as there is typically a disproportionate amount of BS class grid points when compared to smoke.

Scenario	Selected Features	DT GridSearchCV AUC	Simple DT AUC
1	Z, Z_{DR}, ρ_{hv}	83.4%	77.2%
2	Scenario 1 + $SD(Z)$	85.3%	79.2%
3	Scenario 1 + $SD(\Phi_{DP})$	84.4%	77.0%
4	All Features	86.7%	79.5%

Table 4.3: Feature scenarios with different included features for comparison in model training and testing.

Scenario 4, which has all features included, performed the best, with Scenario 2 coming in second. Both had an accuracy near 86%. We can use post-hoc techniques to look at the explainability of both decision tree and random forest models, specifically in this case the variable and permutation importances. By using these, we can investigate and confirm that the model is trained in a way that follows real-world logic and science. Feature importance shows how the inputs effect the model performance. The permutation importance is a form of feature importance. Feature importance shows which features are the most useful for completing the machine learning task, like in this case predicting if a grid point is smoke

or biological scatter. Single-pass permutation importance is a measure that shows the features' importances after shuffling the values of each single feature in the training datasets and observing the changes in the model's performance. By shuffling, the connection between the feature and the target prediction can be broken, and we can determine how much the model relies on that particular feature by how the skill changes. An important note for this is that the permutation importances are dependent on how good or bad a model is; permutation importance does not show the actual predictive nature of a feature by itself, more so how important a feature is for the model [29]. Additionally, correlation between features can also impact importance and explainability. Having correlated features can impact both feature importance and relevance, as this can lead to shared relevance. Shared relevance makes it difficult to determine the impact of each individual feature. Because there was a more limited cause for concern for potential correlated features, single-pass permutation importances were calculated. Perhaps grouping features that could be correlated or have a potential relationship, like Z and Z_{DR} , would be beneficial if moving forward with machine learning. Figure 4.1 shows the variable relevance as well as the permutation importances for both the training and testing datasets for the Scenario 4 case. For this model, we see that ρ_{hv} , $SD(Z)$, and Z have the highest variable importances. This is a promising result, as we learned in the detection and replacement algorithm that ρ_{hv} is very important for separating smoke from all meteorological and $SD(Z)$ is important for improving the initial detection algorithm to remove terrain and other clutter returns.

We can also compare the results of the feature importance and relevance. In this case in particular, the inclusion of the $SD(\Phi_{DP})$ variable shows that for the DT1 model, it has a low importance when looking at feature relevance, but has the highest importance when using permutation on both the training and testing datasets. The high permutation importance indicates that the model performance decreases when the feature values are shuffled for that particular feature. So in the case of $SD(\Phi_{DP})$, when considering the low feature importance

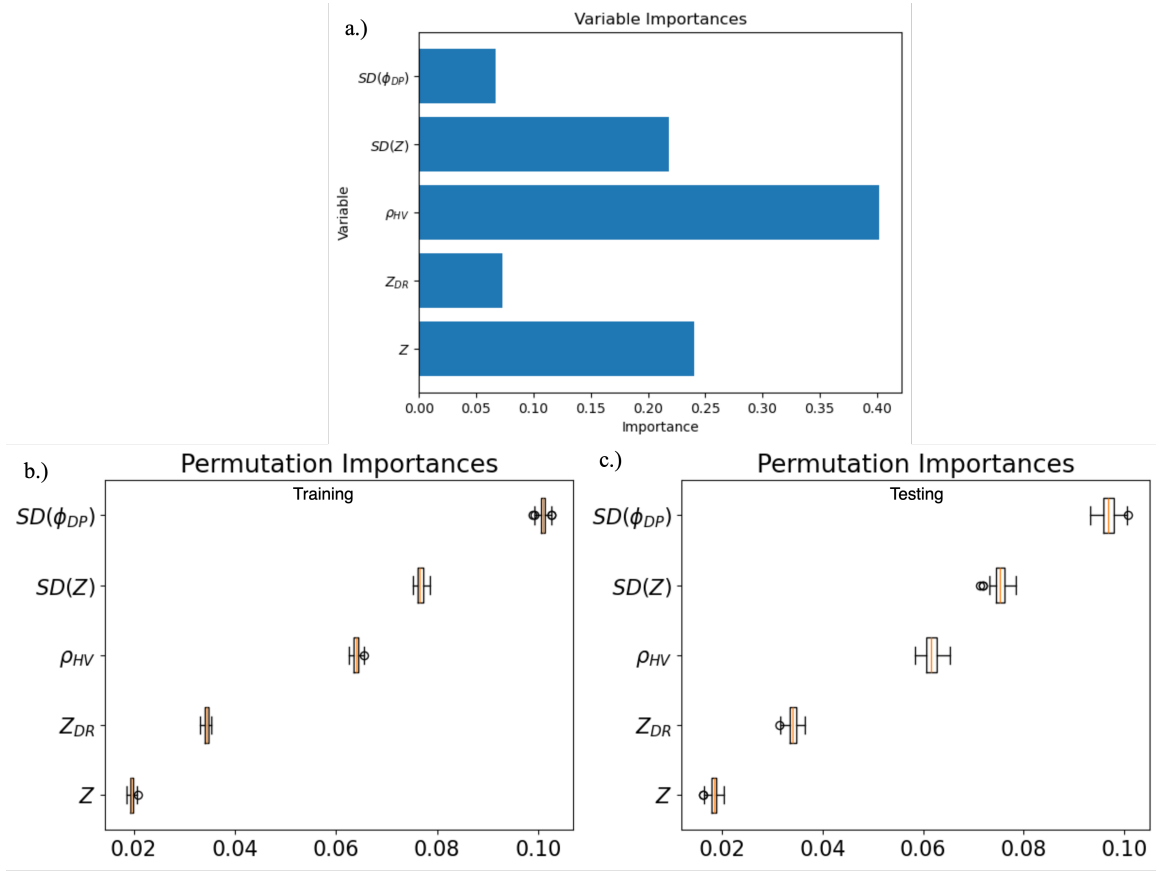


Figure 4.1: The specific importances a.) variable b.) training permutation, and c.) testing permutation from the DT1 models that uses all variables and hyperparameters from the GridSearchCV selection.

and high permutation importance, and the general lack of statistical information about this variable from the other portions of this thesis, it might be beneficial for this variable to be removed. This variable was not included in the detection algorithm because no clear threshold could be determined. Further work with this variable might provide additional insight on how to include it in the DRA as well as the relationship seen in the explainability of the machine learning models. With regards to the other variables, an additional important note for permutation importance is that feature correlations can be carried through permutations, which can result in the final importance for both variables being reduced. There is likely some correlation between the main dual-polarization variables (i.e., reflectivity, differen-

tial reflectivity, and correlation coefficient) driven by physical relationships between drop sizes, orientations, and concentrations. However, it does make the most sense that Z , for example, is the least impactful for decreasing model skill when the feature values are shuffled as there is a wide variety of values for smoke, and the range for biological scatterers, though typically lower in magnitude, is included.

An additional decision tree model was created with specific user input to create an intentionally shallow tree. Figure 4.2 shows the feature decisions of the tree nodes for the detection for smoke, if the thresholds of the DRA were shown in tree form. This tree is very shallow in terms of the number of levels in the tree to predict the final class at one grid point. To emulate this shallowness for a new decision tree method (DT2), the maximum number of final leaf nodes was limited to 5, with the minimum number of samples needed per leaf to 10. This could have been set higher given the large number of samples that needed to be grouped into just 5 leaf nodes. The intention was to highlight the most important variables needed for splitting in a shallow tree in comparison to the DT1 model. This shallow tree should also be interpretable; we should be able to see the changes in impurity and which variables are used in the sequence of splits. The final trees for each the DT1 and DT2 models are included in Figure 4.3, which shows the differences in human-interpretability.

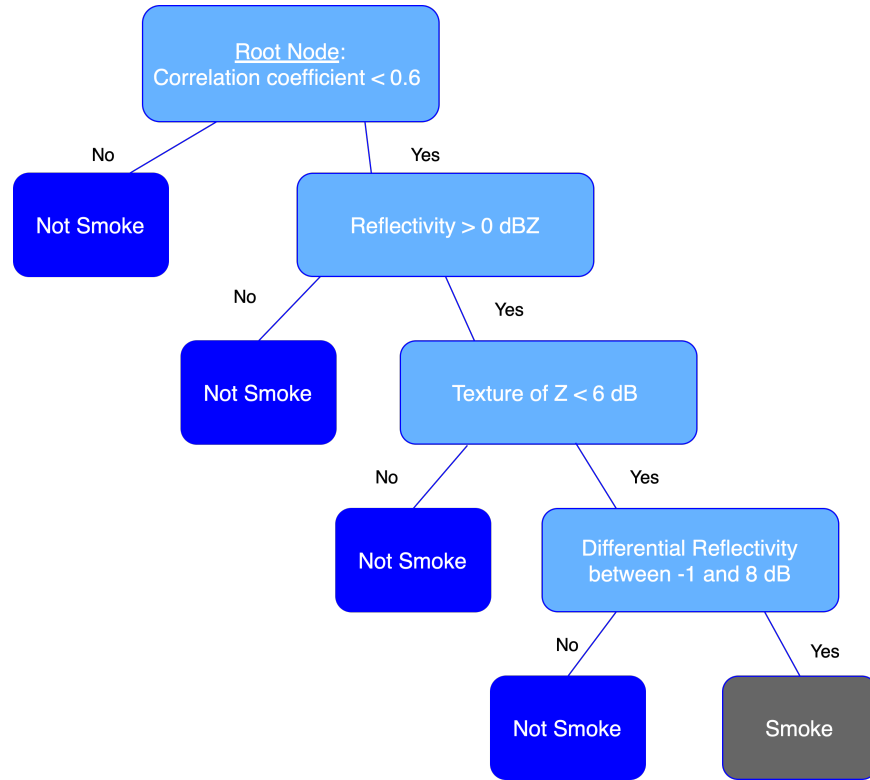


Figure 4.2: Example decision tree that is used to decide if a grid point should be classified as smoke as in the detection and replacement algorithm.

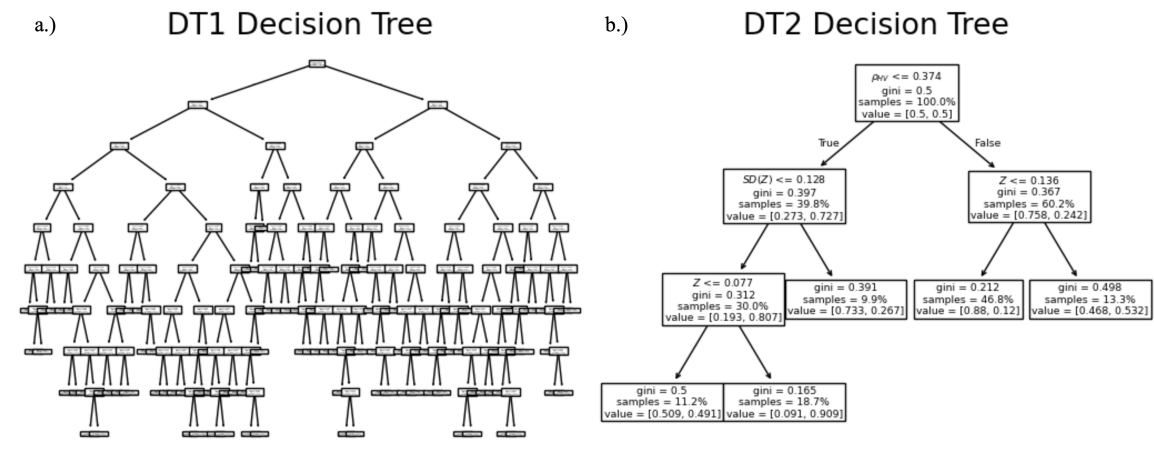


Figure 4.3: The visualization of the a.) DT1 and b.) DT2 model decision trees. The difference in interpretability is evident with the DT2 model, since it has clear splits which can be traced back and compared to the physical model in Figure 4.2.

The same methods shown above for DT1 can be employed for investigating the explain-

ability of this model. Figure 4.4 shows the same charts as in Figure 4.1 but for the DT2 model. One of the main differences is the exclusion of the $SD(\Phi_{DP})$ and Z_{DR} variables in terms of the importances based on feature relevance and permutation importances. When looking back at the decision tree for this model in Figure 4.3b, we can see what variables are used for the splitting criteria, and neither $SD(\Phi_{DP})$ or Z_{DR} are used in this model. Thus, they have a value of 0 in the importance chart for feature relevance in Figure 4.4. $SD(\Phi_{DP})$ and Z_{DR} had some of the lower feature importance when looking at feature relevance, and are thus, likely to be the first features to be neglected in DT2 based on the rankings for DT1. The permutation importances show relative ranking of the features to be fairly similar to those for DT1. The relative placements for permutations show $SD(\Phi_{DP})$ as the highest ranked feature with $SD(Z)$, ρ_{hv} next, and Z and Z_{DR} with values near zero. This indicates that the model performance is not influenced heavily by the shuffling of Z and Z_{DR} values in the training and testing data. It should be noted that these explainability techniques would change with an increase in the number of final leaf nodes, if the tree depth was increased for added generality or if more data was added, for example.

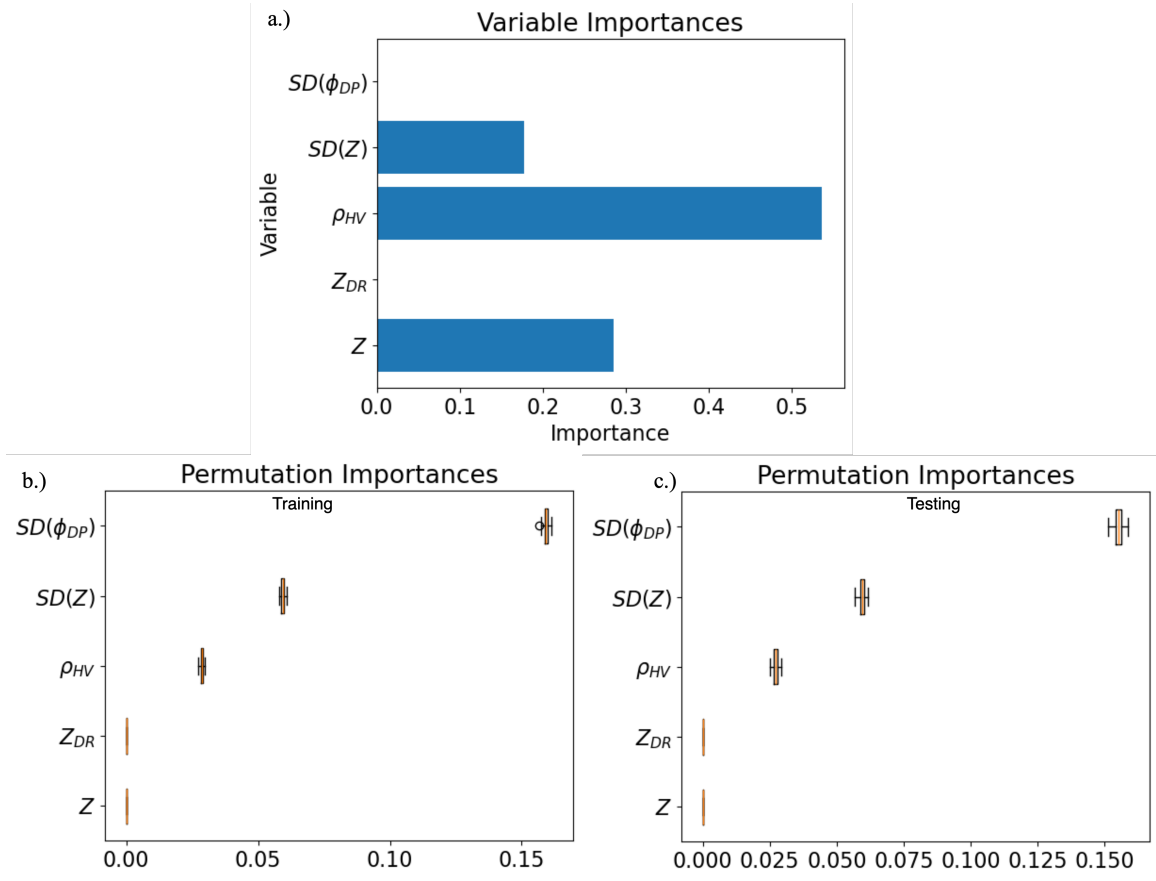


Figure 4.4: As in Figure 4.1, but for the DT2 model.

4.2.2 Random Forest

The random forest model in this work was tuned using GridSearchCV to select the best hyperparameters as shown in previous sections of this chapter. Figure 2.3 shows that random forest models have higher complexity and lower understandability, with more post-hoc methods needing to be employed. Unlike the decision trees in the previous section, the random forest model cannot be plotted to see how the model is choosing to split at nodes, and what features are being used. We can still use variable and permutation importances, however, to look at what features are important for creating the model and making predictions. Figure 4.5 shows the outputs for these as in Figures 4.1 and 4.4. The GridSearchCV

method indicated that all including all five features in the model was the best option for model performance and thus all were used in training. The importances bear some similarities and differences to that from the decision tree models. For the random forest, ρ_{hv} and Z are still the most “important” features for this model, and Z_{DR} is still last. The texture features show the most difference. $\text{SD}(Z)$ was shown to be relatively important in the previous models with the final values being higher than that of $\text{SD}(\Phi_{\text{DP}})$. In the random forest case however, they are relatively equal in value, and also have the highest permutation importance values for both the training and testing datasets (as in the decision tree models). Correlation coefficient also has a relatively high permutation importance for this model, meaning that shuffling the values for this feature can impact the model more than that of Z , for example. Overall, this model performs more like the DT1 model, though it is likely less susceptible to overfitting and has more generality as the ensemble method boosts the weak learner aspects of the decision trees to make a more accurate model. The ranking of features for feature relevance does reflect the example decision tree for the detection and replacement algorithm well, as seen in Figure 4.2. The first main splits are using ρ_{hv} and Z which is similar to the relative importances shown in Figure 4.5a, for example.

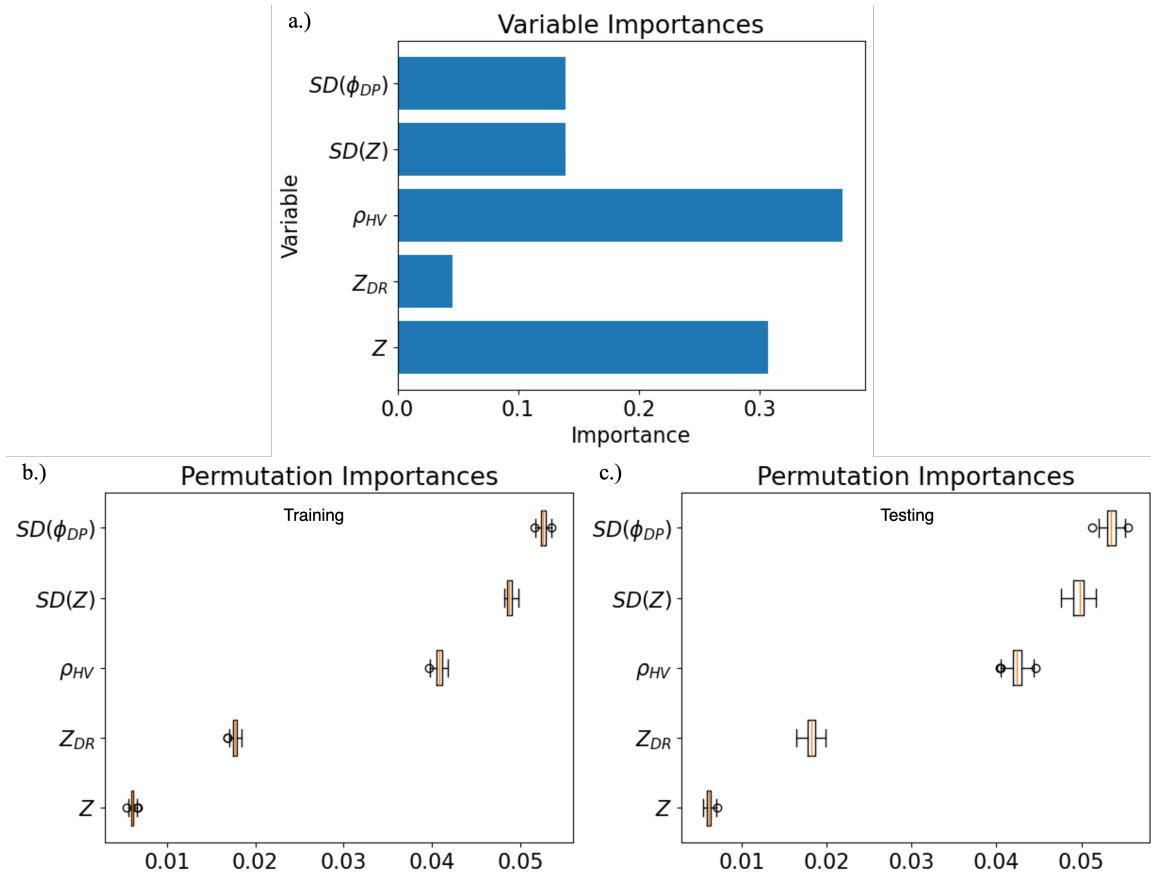


Figure 4.5: As in Figure 4.1, but for the random forest model.

4.3 Current Limitations and Suggestions

The main goal of exploring the use of machine learning methods for the classification of smoke using radar was to show an additional possible avenue for future work. There are numerous ways to approach the problem of the lack of training datasets and how the data should be selected for training models. For example, the data used as input for these models are simply grid points, with no additional information about the location relative to the radar, height above ground level, or date and time provided. This means that the models are learning simply based on the single values of all features at each point, creating a very general pathway for the models to learn. There are more complicated ways to partition

out data for use, using digital image processing techniques (blob coloring for minor region removal) as in [20], for example. Perhaps a combination of blob coloring in addition to erosion and dilation to cover all portions of the plume out to the edges would provide benefits in the DRA, and by extension, the machine learning models. Further investigation of data and how it is input into models would be beneficial for creating new, best-practice frameworks for generating datasets that include smoke or a subclassification for biological scatterers moving forward.

Chapter 5

Application of Classification Algorithms and Models to Select Cases

5.1 Case Studies

While a significant portion of the 25-30 cases explored in this work exhibit interesting interactions with various weather phenomena and locations across the country, a more limited number were selected to highlight the successes and limitations of the classification algorithms and models introduced and discussed in previous chapters. The following five events were selected for their high impact on the public, firefighters, the local environment and ecology, or because of the meteorological or topographical features of interest.

5.1.1 Palisades and Eaton Fires, California

Due to the recent nature of these fires and the impacts on the general public, this event was of high importance for testing. The Palisades and Eaton fires started under extreme fire weather conditions. During the evening hours of January 7th, 2025, strong Santa Ana winds blew across the mountains and into the Los Angeles area. A Particularly Dangerous Situation (PDS) red flag warning was issued by the NWS Los Angeles office, messaging for wind gusts from 80-100 miles per hour, and relative humidity values between 10-20% per a NWS Los Angeles Red Flag Warning (RFW) product. The Palisades fire burned 23,707 acres of homes and wildland areas, and the Eaton fire burned 14,021 acres via CALFire.

The final fire perimeters are shown in Figure 5.1. Unlike “typical” wildfires, these fires also burned fuels that are anthropogenic in nature. As opposed to nearly all fuels being some combination of grasses, timber, and litter in most wildfires, fires also burned a significant amount of structures and vehicles. Scatterers of different origin brings further complexity to the radar returns, as well as health impacts to those in the area. Both the KVTX (Los Angeles) and KSOX (Santa Ana) radars were able to detect these fires. However, there were significant issues with the location of KVTX and the fires relative to the local topography, and as such, the results will not be discussed in this work.

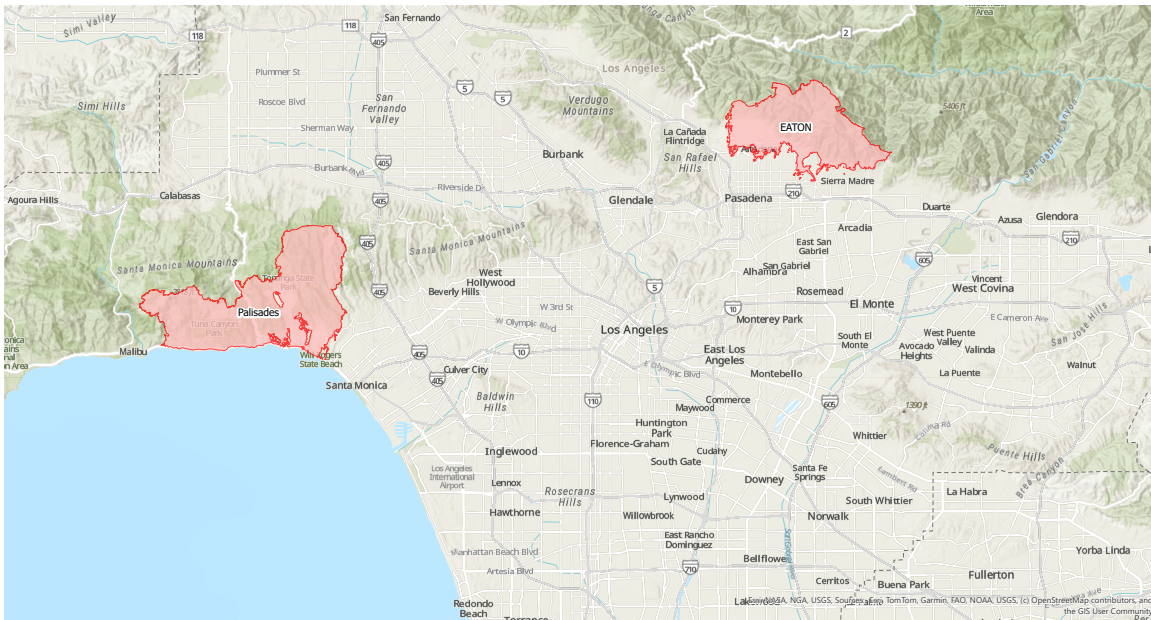


Figure 5.1: Final fire perimeters of the Palisades and Eaton fires from the National Inter-agency Fire Center WFIGS Fire Perimeters

Detection and Replacement

The Palisades fire was first detected on the KSOX radar at 19:21:21 UTC on January 7th, 2025. The smoke plume was relatively small in areal extent during this time until stronger winds increased in the evening hours. The plume had a higher ρ_{hv} value throughout and a high Z_{DR} , around 5-6 dB. The smoke plume was correctly identified using the DRA during

the entire period up until the end of the day. However, there was an area of very low ρ_{hv} in the more mountainous region to the north of the Palisades fire indicated by the red oval (north of the smoke plume in Figure 5.2a). This area was not removed with the addition of the SD(Z) filter, likely because the change in Z along the radial is not as sharp as it can be in other mountainous areas (see green oval in Figure 5.2) or there were other biological scatterers like insects present. This area was persistent during the evening and overnight hours into January 8th. There is improvement during the day time hours into January 8th, which can be seen in Figure 5.2b. The Eaton fire was first detected on January 8th, around 03 UTC. The plume for this fire at the time had a lower ρ_{hv} value at the selected times than the times on the 7th and 9th. There was also an area of much lower Z_{DR} within the plume, around 0-2 dB when compared to the outer edges of the plume. This area corresponds to the areas of highest Z values as well. As the Santa Ana wind event progressed and the dewpoints decreased during the day, the fire spread and fire behavior continued to increase. Both fires continued to be detected during the entire day of January 8th. In the overnight hours into January 9th, there was less smoke production as fire spread and behavior was likely more limited. There is thus, a period where there is no smoke detected by KSOX and the smoke produced by the fires were not seen again until just after 15 UTC. Figure 5.3 shows comparisons of the original HCA output and the new HCA output that includes smoke for selected time steps on all three days.

At all three time steps, even with the larger area of misclassified smoke in Figure 5.3a, the actual plumes are more easily identified for both the Palisades and Eaton fires. The smoke plumes had relatively low ρ_{hv} values, which is likely driven by the wide variety of particle size, shapes, and compositions included in the plume. This could have been anthropogenic or biological in nature. For this particular case, the DRA output could be improved upon by lowering the ρ_{hv} threshold in the detection stage (Figure 5.3).

As described in Section 3.3.1, we compute the ROC curves for randomly selected times

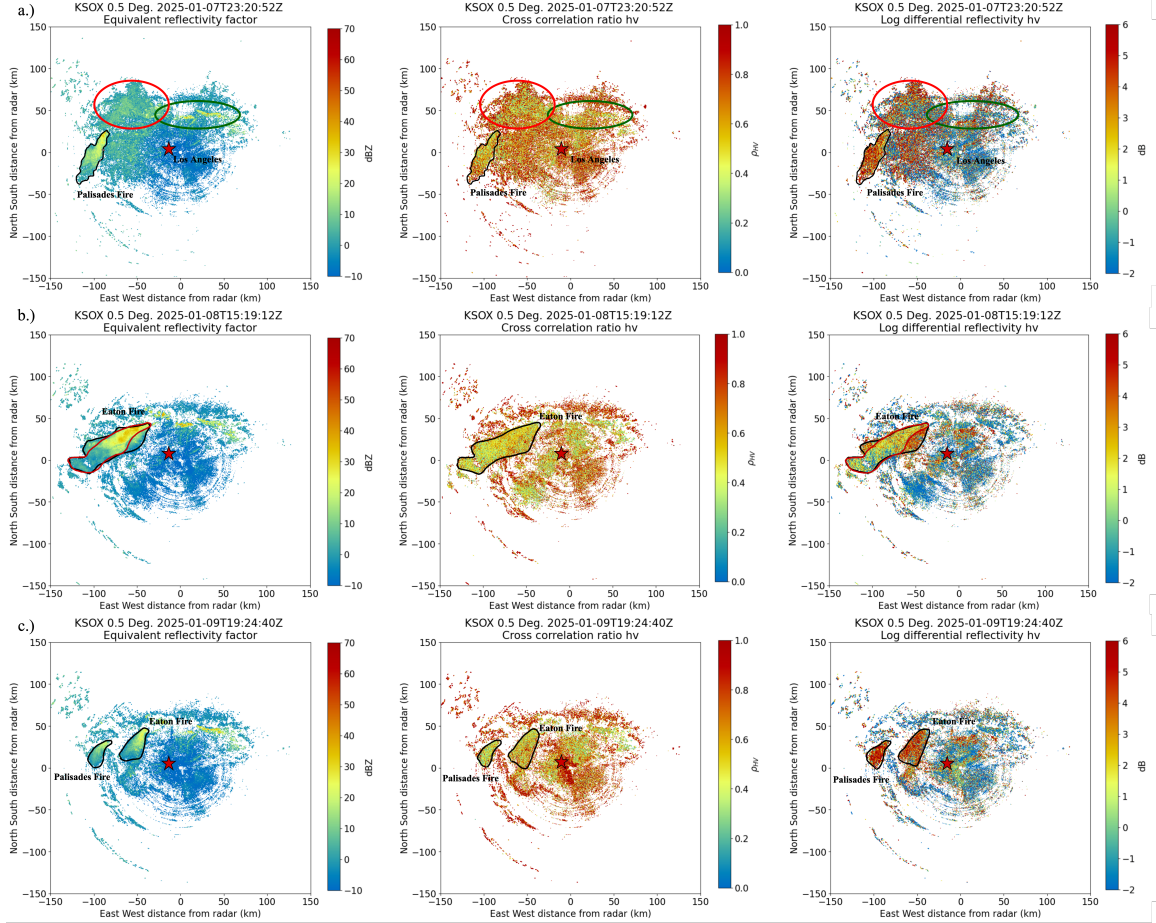


Figure 5.2: Z (left column), ρ_{hv} (middle column), and Z_{DR} (right column) for selected times a.) 23:20:52 UTC on January 7th, 2025, b.) 15:19:12 UTC on January 8th, 2025, and c.) 19:24:40 UTC on January 9th, 2025 from the KSOX radar. The black outlines indicate the smoke plume or plumes for each time step based on polarimetric variables. In a.) the red circle is an area of incorrect classification as seen in the red oval in Figure 5.3, and the green oval indicates areas of mountain returns from the Santa Ana mountains.

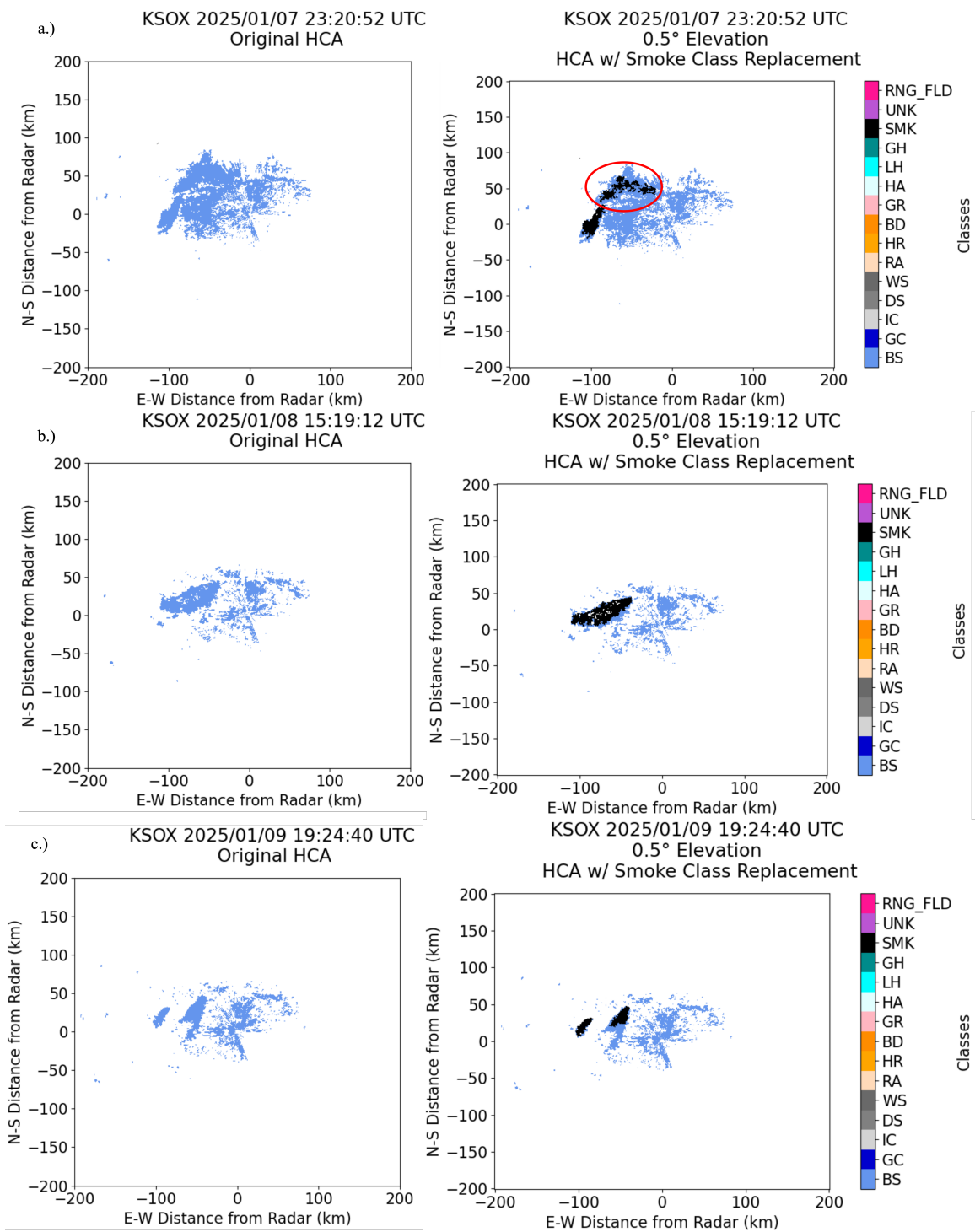


Figure 5.3: Examples of the detection and replacement algorithm (right column) when compared to the original HCA output (left columns) for selected times a.) 23:20:52 UTC on January 7th, 2025, b.) 15:19:12 UTC on January 8th, 2025, and c.) 19:24:40 UTC on January 9th, 2025 from the KSOX radar. The red oval in a.) indicates an area of incorrect classification likely driven by the presence of other biological scatterers.

that are approximately 30 minutes apart for each day of the event (Figure 5.4). The area under the curve (AUC) shows that the algorithm has skill in predicting the smoke class, and is better than a random classifier which is denoted by the red dashed line. Some of the times showed a better performance on January 7th, and by looking at the respective times for the individual curves, we know that these were the earlier times with less false detections over the mountains. The later times are slightly lowering the average in this case, but this more accurately quantifies how the algorithm performs in a more complex scenario where there is clutter in the smoke region.

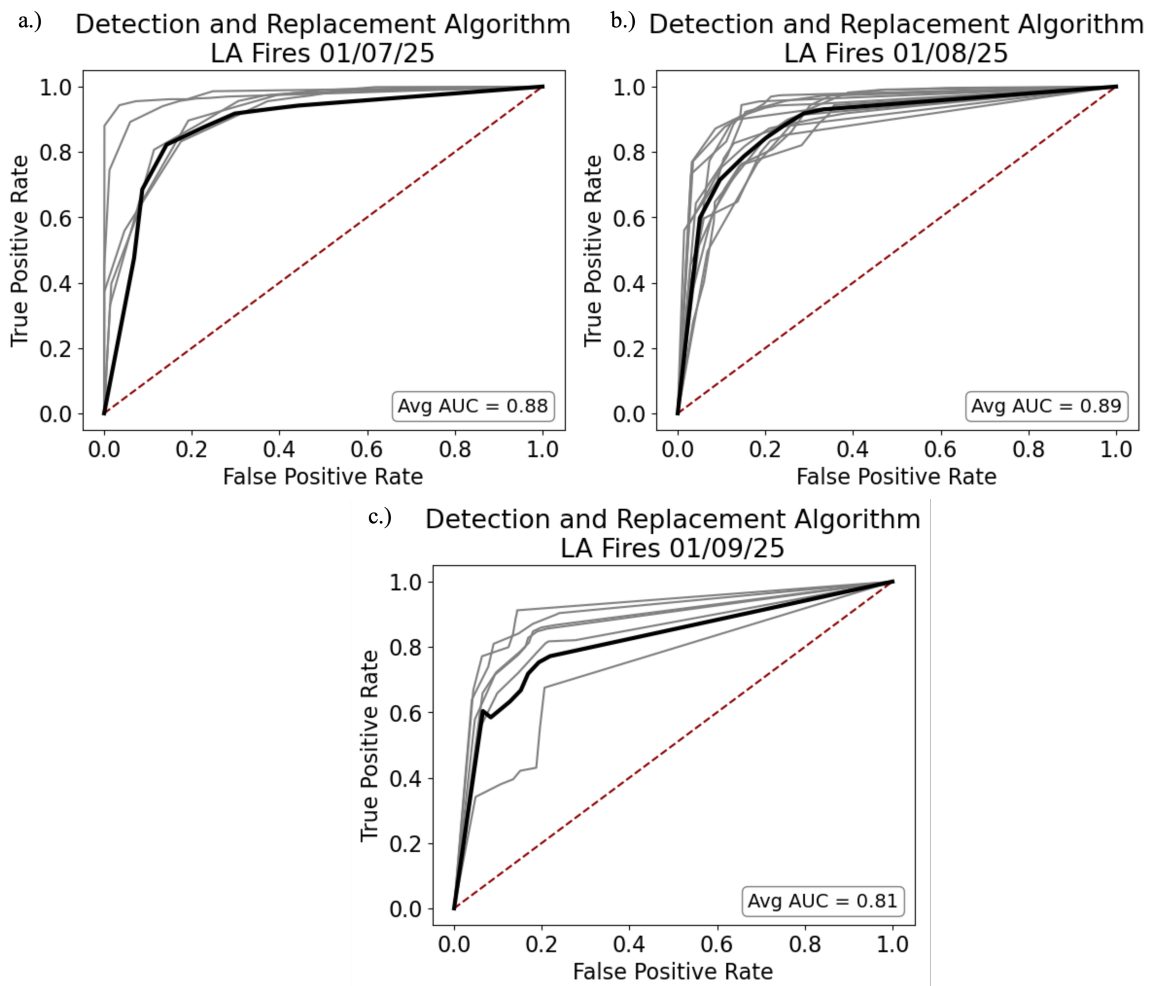


Figure 5.4: ROC curves for the detection and replacement algorithm for selected times on a.) January 7th, b.) January 8th, and c.) January 9th.

Machine Learning Models

This case was not included in the training or testing dataset for the machine learning models outlined in Chapter 4. To tie the detection and replacement algorithm together with the machine learning models, confusion matrices and ROC curves will be used to show the predictive power for this case and for each model. It should be noted that further metrics should be explored for model comparison, as AUC from the ROC curves may not be enough to fully quantify the models' performance. Additional metrics and explainability methods would be beneficial will continued work on the machine learning models. Example outputs will also be included for selected times. When using the metrics from scikit-learn packages, the confusion matrix is made using the predictions for each model, with the values in the colorbars indicating the count of grid points with that value on the lowest grid layer. The input data are the normalized radar variables, for all of the time steps for the fires per day or for the full duration of the event. The ROC curves are made using the probabilities for the same dataset and models.

Given Figures 5.5, 5.6, and 5.7 we can compare the final outputs and compare which models performed the best for this case. The detection and replacement algorithm performed the best for the January 7th date. It had the highest AUC value (for the average ROC curve in bold black), while the DT1 model had the best AUC value from the machine learning models. We can look at the counts in the confusion matrices for all three models to see if there is relative over or under prediction of smoke. Given that we are attempting to predict smoke, it is set as the “positive” value in the confusion matrix, such that the bottom right square shows the true positives. Given that the dataset will *typically* have many more biological scatterers pixels than smoke, which is true for this case, it will have the highest counts. Of more interest are the incorrect classification boxes which are the bottom left and top right. Figure 5.5 shows that the decision trees are likely over predicting smoke, and in

Figure 5.8 we can see that the models are classifying some mountain returns as smoke. The random forest model shows more counts where the model incorrectly predicted a value of 10, which is biological, where it should be smoke. Thus the random forest is under predicting smoke. The DT1 model has the highest AUC and a more balanced confusion matrix for this case, and the other days show the same pattern to some degree.

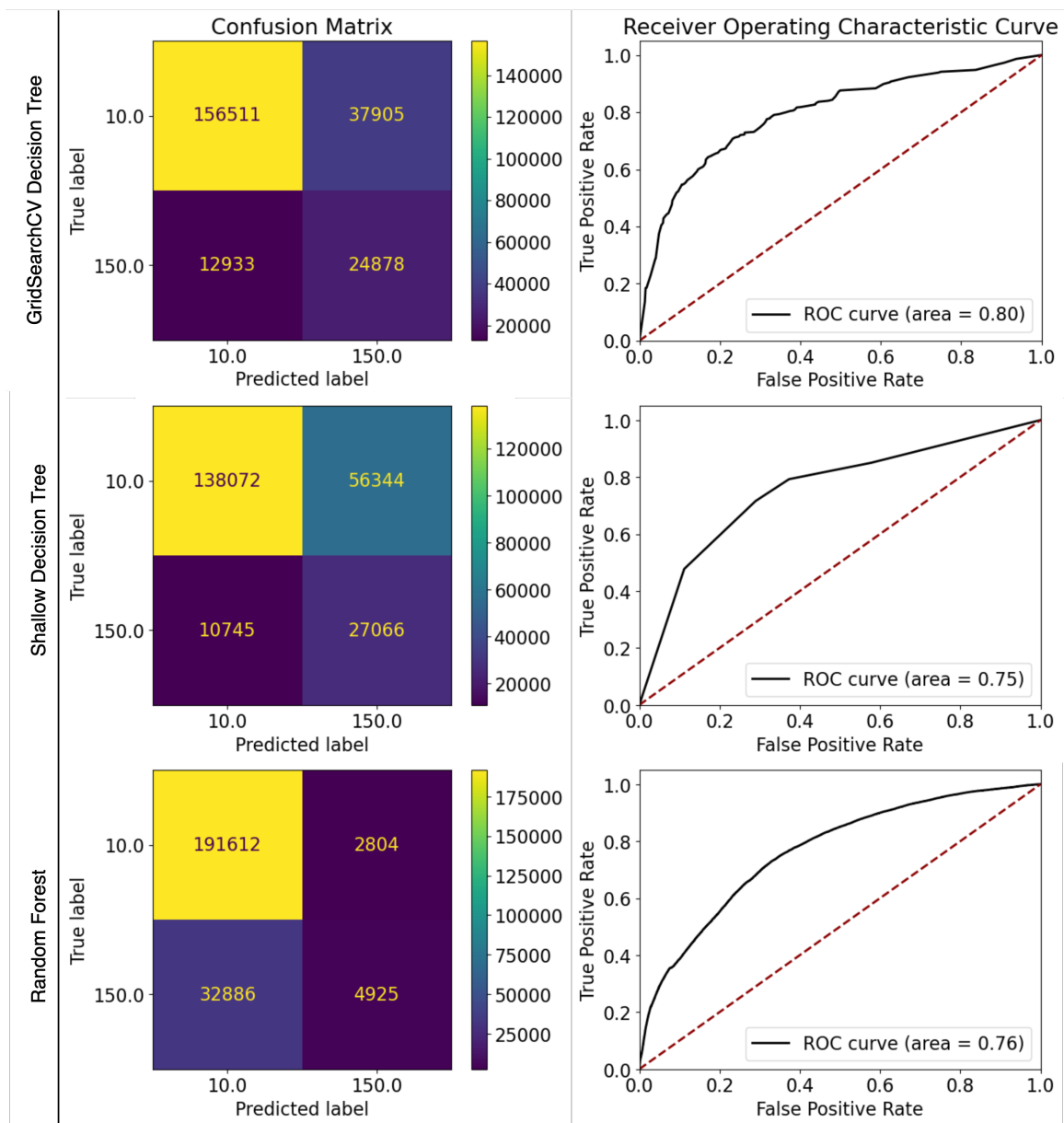


Figure 5.5: Selected metrics for model comparisons for the DT1 model (top row), the DT2 model (middle row), and the random forest model (bottom row). The left column is the confusion matrix output for each model on the selected data with 150 being the positive label for smoke and 10 indicating biological scatter. The right column is the ROC curve based on calculated probabilities for each model. The red dashed line indicates where performance would be the same as a random classifier. The metrics for this figure are for January 7th, 2025 from the KSOX (Santa Ana Mountains) radar.

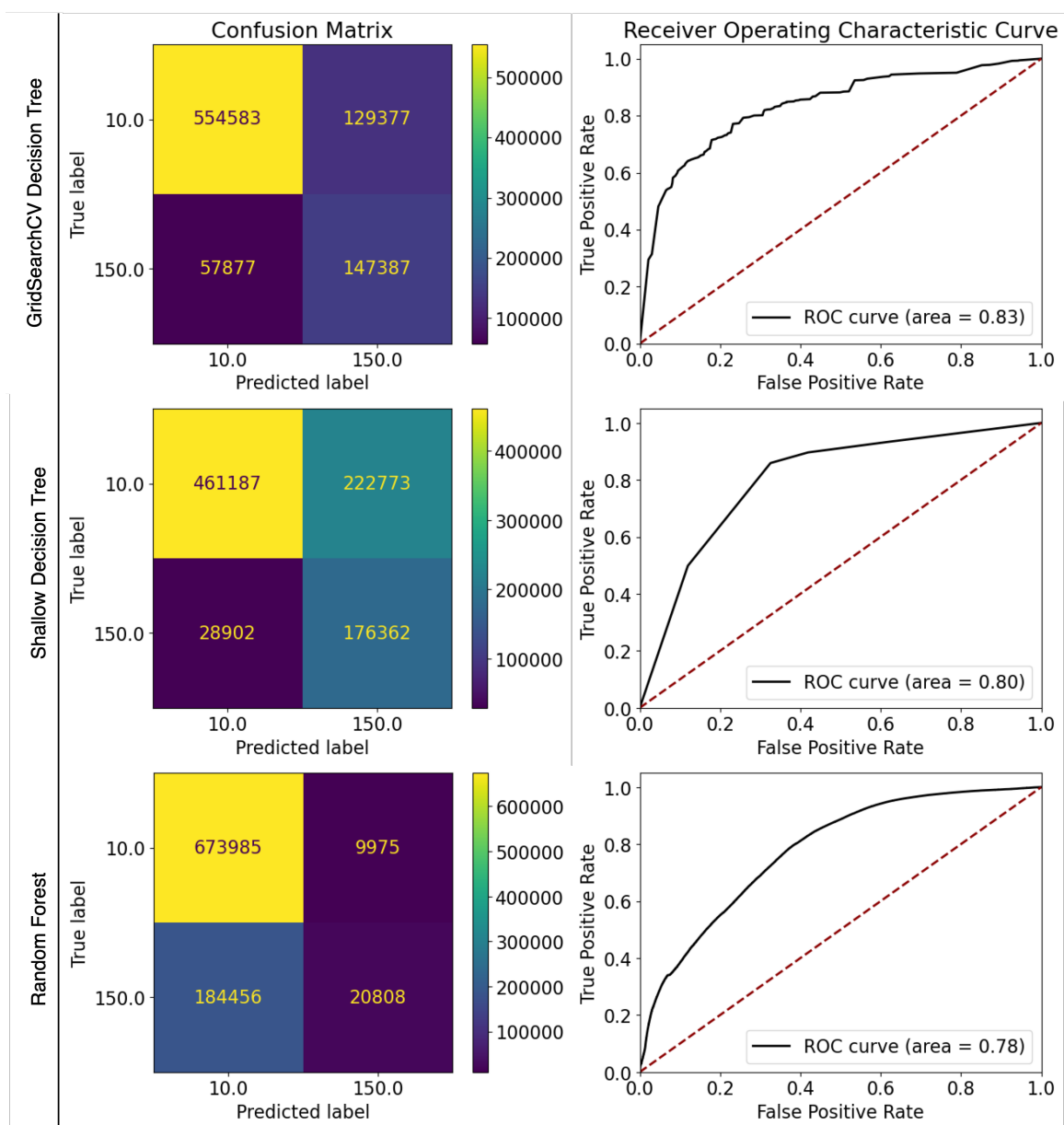


Figure 5.6: As in Figure 5.5, but for January 8th, 2025.

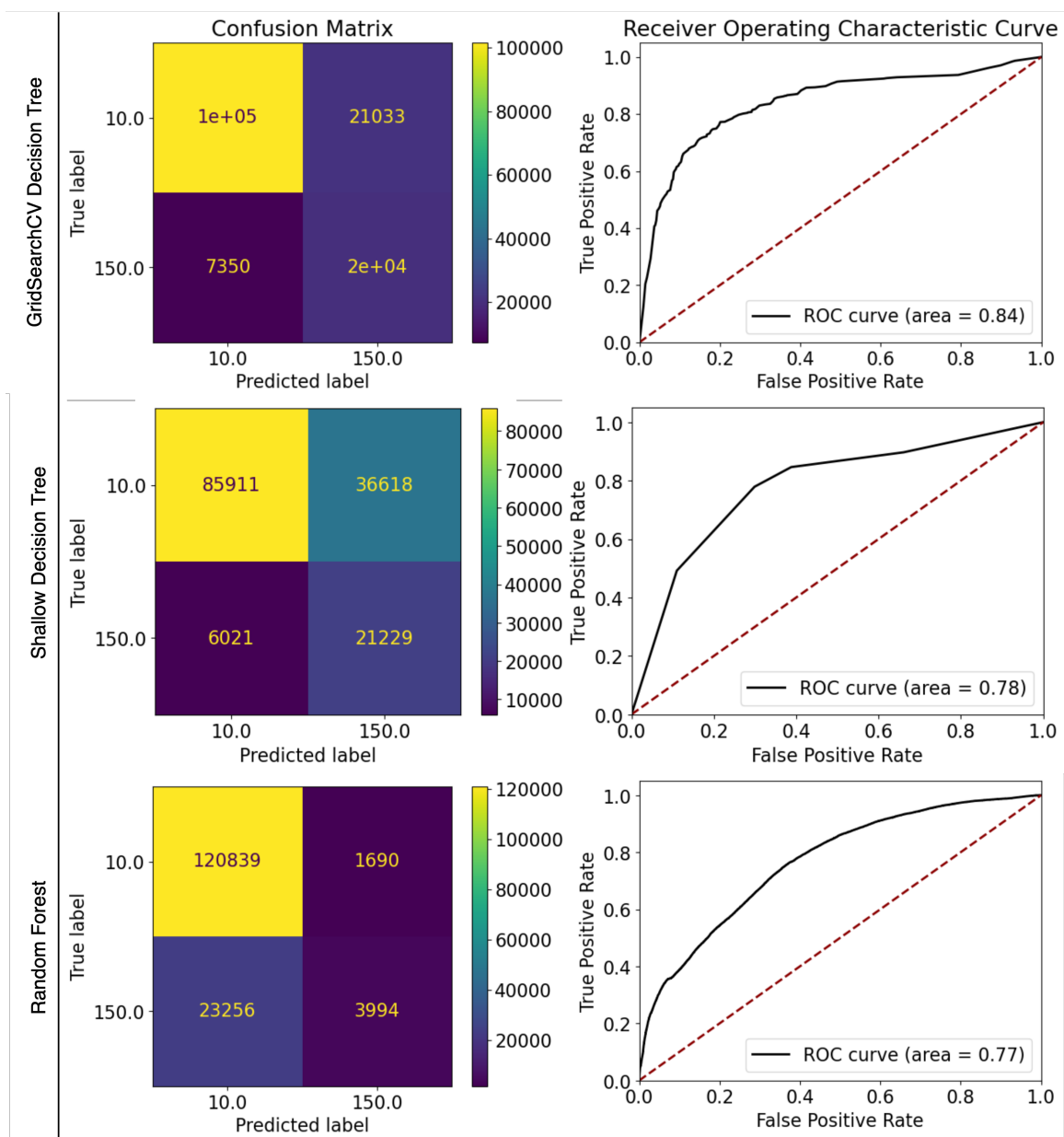


Figure 5.7: As in Figure 5.5, but for January 9th, 2025.

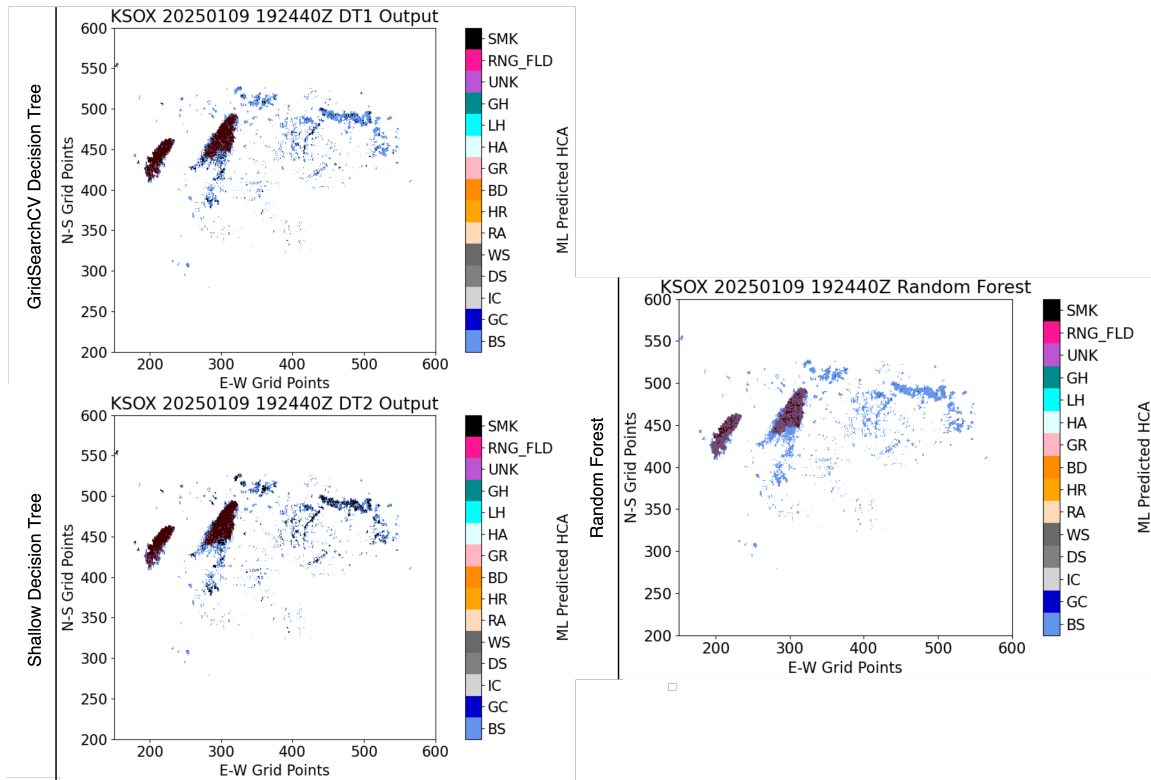


Figure 5.8: Machine learning predictions for a selected time on January 9th, 2025. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.

5.1.2 March 14th, 2025 Wildfires, Oklahoma

During the spring fire season across the southern Great Plains, an extremely strong wind event drove extreme fire weather. This event was driven by a strong low pressure system that formed and propagated off the Rocky Mountains with an associated strong upper-level trough and jet moving through the Oklahoma and Texas Panhandles. The 12Z SPC Fire Weather Forecast indicated that winds up to 50 knots were being recorded in the lowest kilometer of the atmosphere. Potential RH values in 10-20% range were also noted. In addition to fire and smoke concerns, the strong winds over dry areas with limited vegetation in the Texas Panhandle stirred up dust which spread into Oklahoma. Figure 5.9 shows the



Figure 5.9: The Rapid X-Band Polarimetric Radar (RaXPol) scanning a wildfire on March 14th, 2025 outside Norman, OK at Lake Thunderbird. The orange colored haze is dust and smoke driven by the high winds.

haze over Norman, OK, which was a mix of dust and smoke from nearby fires. By 11 am local time, fires began to pop up in north western OK and in the central Texas Panhandle. As the day progressed, a complex of fires burned northeast of Oklahoma City. Around 400 homes were burned, and there were a large number of injuries and some fatalities from fires and also from high wind hazards [35]. Given their locations, three radars were able to scan most of the wildfires on this day. These include KAMA (Amarillo, TX), KVNIX (Vance AFB, OK), and KTLX (Twin Lakes, OK). For the sake of brevity, only the wildfires in Oklahoma will be studied, which were scanned by KVNIX and KTLX.

Detection and Replacement

The first fire in northwest Oklahoma was seen on the KVNx radar around 1630 UTC. With the strong winds associated with this event, the smoke plume spread to the northeast quickly and the largest fire persisted well into the overnight hours. During the daylight hours, more fires cropped up nearby and a more dense smoke plume persisted over KVNx until a wind shift took place in the evening hours. The first fire closer to KTLX can also be seen, to the southeast of the KVNx radar. Overall, the detection and replacement algorithm picked up on the original and other smoke plumes well, though perhaps clutter and other interference around the radar influenced the ρ_{hv} values and increased them enough to have that area occasionally misclassified. There was also a fire that started fairly close to the radar, which was not correctly classified, and this concept of interference from clutter can likely be applied for that case. The KVNx radar also shows precipitation to the north of the radar, which is kept as the original class as described in Chapter 3. There are a few misses with regards to individual smoke plumes for both KVNx and KTLX. These are typically tied to the size threshold employed and the ρ_{hv} around the radar. There were numerous, small wildfires that started and were controlled or died relatively quickly, and these fires were the most missed by the algorithm. A few selected times will be discussed further, Figure 5.10 shows the main dual-polarization fields for reference.

With regards to the radar variables themselves, these fires show the typical values that were discussed previously. There are relatively high Z values, especially near the active fire, which is expected. The values of ρ_{hv} are typically between 0.4 and 0.6, and are similar in values between fires and radars. Z_{DR} exhibited the increasing trend in value downwind of the fire, with values closer to 1 dB in some regions, while the end of the plumes were around 6-8 dB. With these thoughts in mind, we can analyze the output of the detection and replacement algorithm for these time steps as well. Figure 5.11 shows the output and

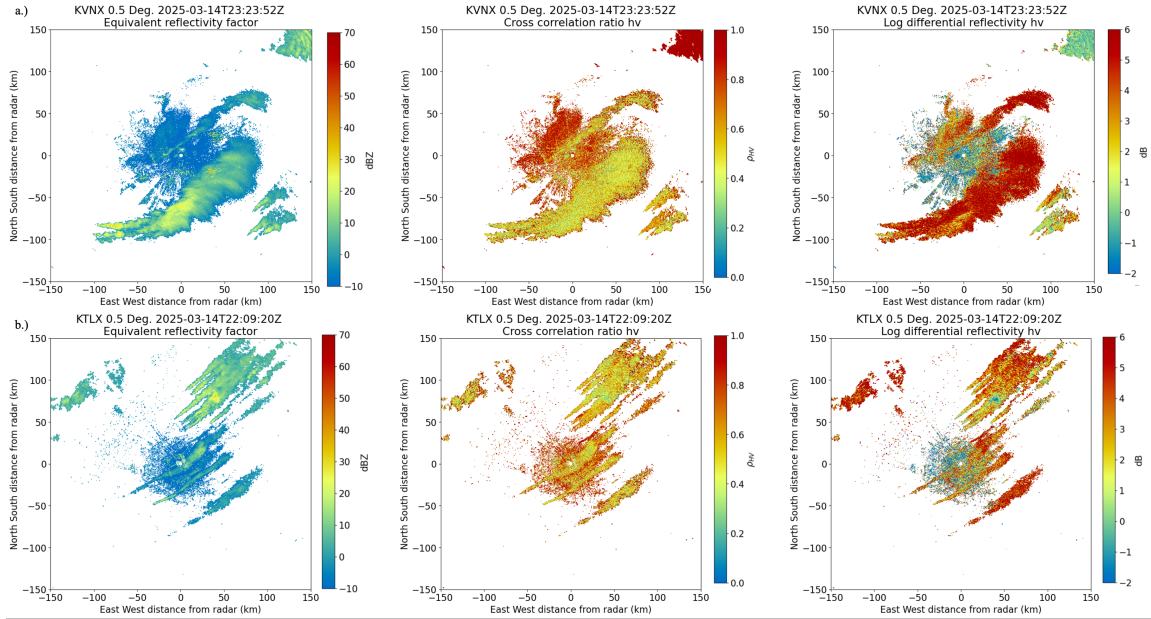


Figure 5.10: The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 23:23:52 UTC from the KVNx radar and b.) 22:09:20 UTC from KTLX.

includes some areas of missed classifications of smoke in the red ovals. These fires and portions of the smoke plumes are relatively close to the radar, and while they look distinct in the Z field, the values of ρ_{hv} are closer to the noise and other returns seen near the radar. The size threshold is also likely impacting detection, as the smoke plumes were often long, but quite narrow which could reduce the relative area to be too small for the algorithm.

We can look at the ROC curves for selected times from both the KTLX and KVNx radars to evaluate the skill of the detection and replacement algorithm for the case in Figure 5.12. For the KVNx case, there is decent agreement between the different time steps, where most of the ROC curves are centered around the average. There were fewer small fires to miss from this radar's location and so the average AUC value is much higher than the KTLX case in Figure 5.12b. For the KTLX case, there were more misses as indicated for example in Figure 5.11b near the radar. With further analysis in the contours themselves, it was noted that the size threshold was restricting other sections of these fires and

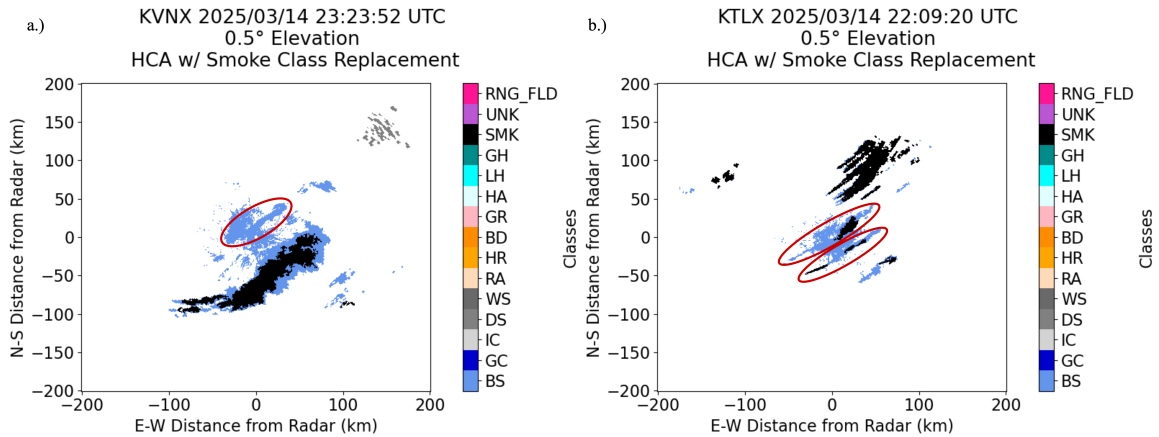


Figure 5.11: The final output of the detection and replacement algorithm for a.) 23:23:52 UTC from the KVNx radar and b.) 22:09:20 UTC from KTLX. The red ovals indicate areas of missed classification of smoke.

smoke plumes. There is a tradeoff for selecting the size threshold, where having it too small allows small sections that could be incorrectly classified as smoke but it could also allow for more correct classification in this case. The suppression of the incorrect regions is more important, but there is still some information provided by a partially correct classification.

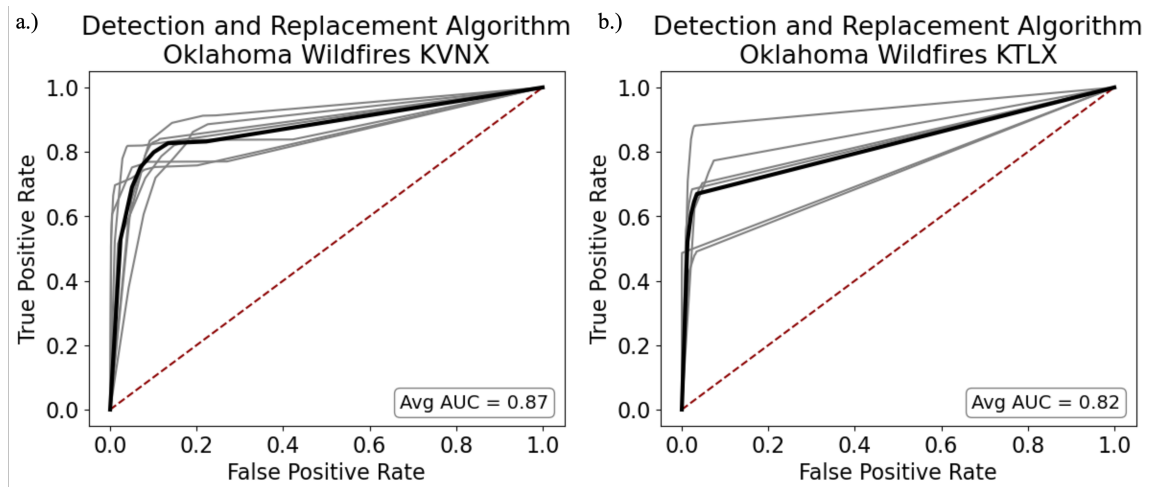


Figure 5.12: ROC curves for the detection and replacement algorithm for a.) KVNx and b.) KTLX for the March 14th, 2025 wildfires.

Machine Learning Models

Like the Palisade and Eaton fire cases, this date was not included at the time of the creation of the machine learning models. The metrics should once again provide information about the models performance, especially for over and under predicting smoke detections. As in the previous section, the confusion matrices in Figures 5.13 and 5.14 show over prediction for the DT2 model and under prediction for the random forest model. Given that the number of pixel counts are the same between each model for each figure, we can see that the distribution of smoke and biological scatterers changes from model to model. The random forest model in both figures shows significantly higher counts in the bottom left square, which is where there should be a smoke classification and is instead classed as biological scatterers. Thus, the random forest is consistently under predicting smoke for this case on both radars. The decision trees exhibit similar behavior to the previous case for KVNIX, where the DT1 model is not as prone to over predicting, in areas of wind turbines in particular for this radar. It has the highest AUC value as a result. The DT2 model has a higher count in the top right square which indicates a false positive, where the value should have corresponded to biological scatterers and not smoke. All of the models performed worse on the KTLX case when compared to the KVNIX case. There was more under prediction of smoke in all cases, which is likely driven by the misses of the small fires in the original detection and replacement algorithm, and then again in the machine learning models.

When looking at the predictions for the two selected times for the case, it is obvious that the random forest model is under predicting smoke. In both cases, there are some areas that are liberally classified as smoke, especially where the smoke plumes are the largest. This could be driven by larger areas of lower ρ_{hv} . With regards to the decision tree outputs, the over prediction is more obvious in the KVNIX case in Figure 5.15. There are areas of wind turbines around the KVNIX radar, which can be seen better in Figure 5.10a. These

are suppressed in most cases in the detection and replacement algorithm using the Z_{DR} threshold. In the case of DT2 in particular, this variable is not included in the final tree and thus might account for the over prediction in this case. The DT1 model does have a better output with regard to this particular issue and is likely why the AUC is higher. The decision tree models do pick up on the smaller plume nearest to the radar, seen in the red oval in Figure 5.10a. Because the predictions are simply driven by the dual-polarization variables and there is no size threshold, there is some improvement for this issue. The same is true for the KTLX case, though areas of clutter look to be classified as smoke in this case, likely for the same reason as with the wind turbines.

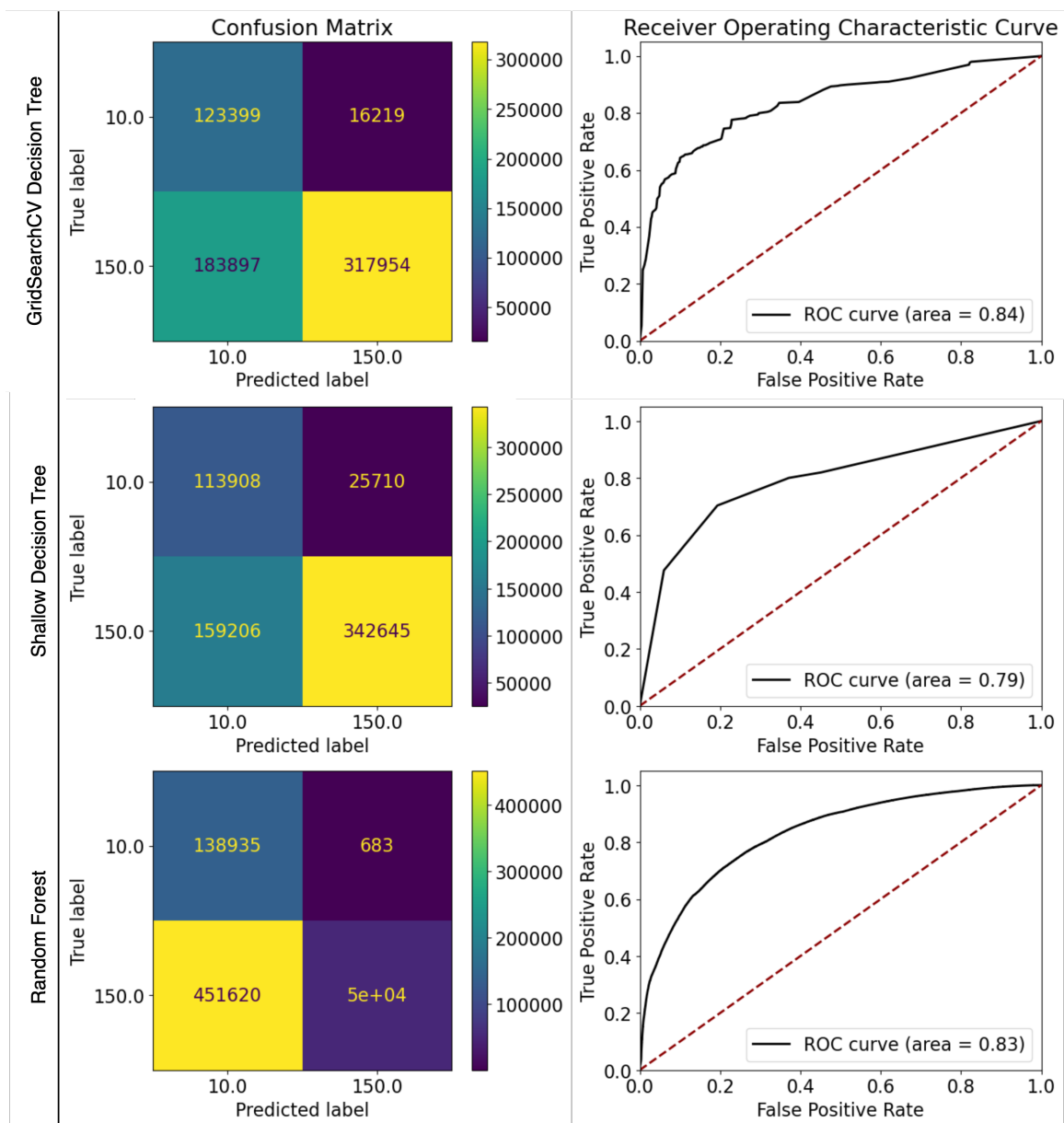


Figure 5.13: As in Figure 5.5 but for the wildfires that took place on March 14th, 2025 from the KVNx radar.

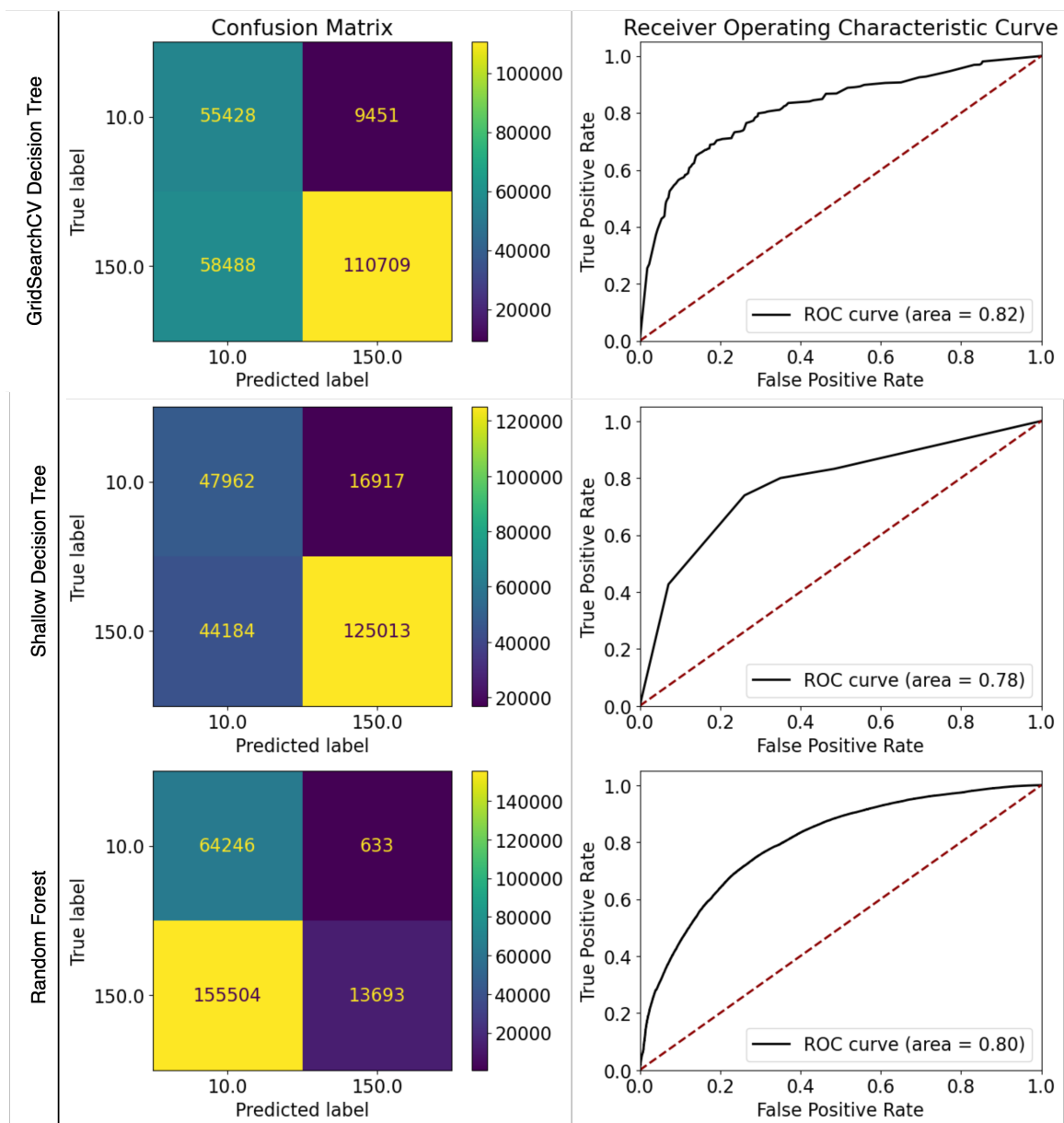


Figure 5.14: As in Figure 5.5 but for the wildfires that took place on March 14th, 2025 from the KTLX radar.

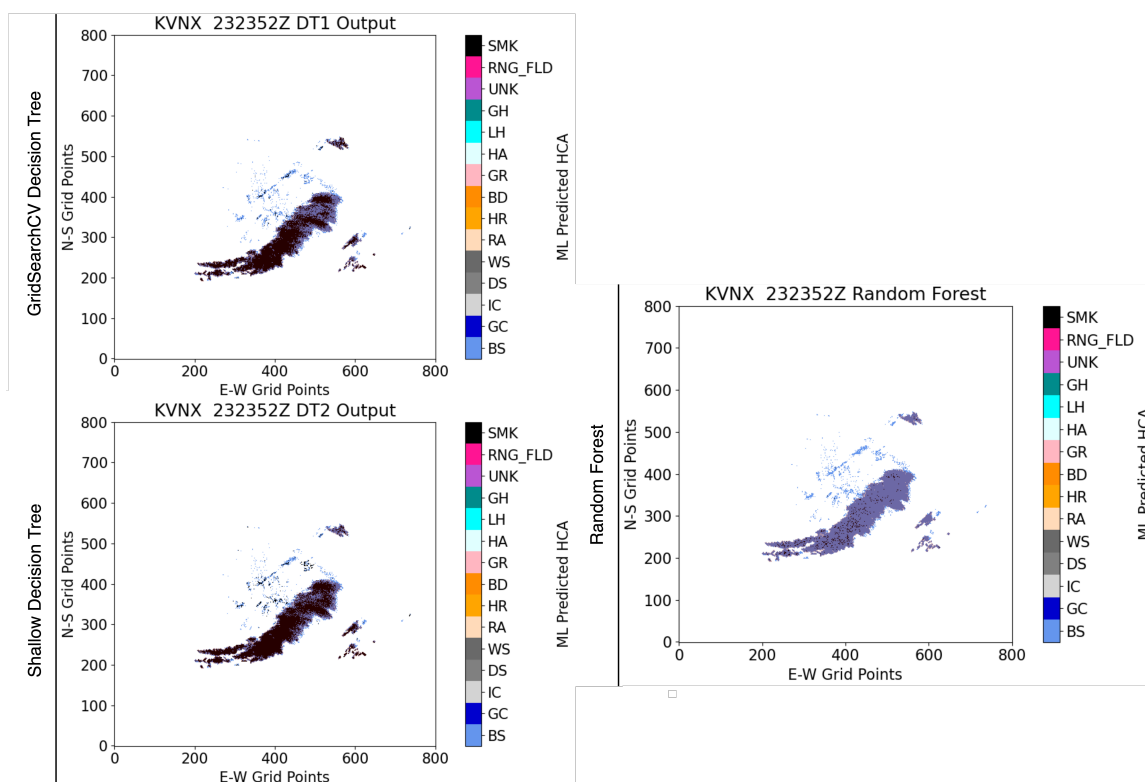


Figure 5.15: Machine learning predictions for 23:23:52 UTC on March 14th, 2025 from the KVN radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.

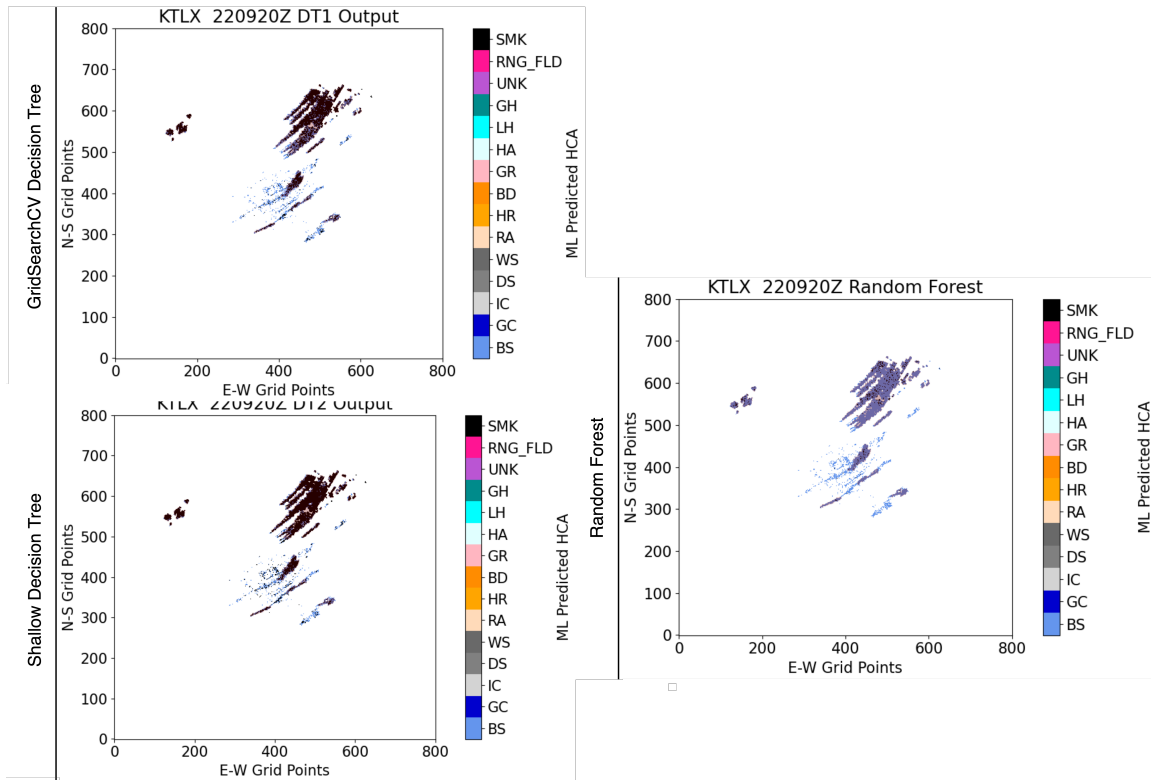


Figure 5.16: Machine learning predictions for 22:09:20 UTC on March 14th, 2025 from the KTLX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.

5.1.3 Snake Trail Fire, Minnesota

The Snake Trail fire burned around 2,000 acres in south central Minnesota on March 3rd, 2024. It took place under fire conditions that were more limited in terms of extremes, but the lack of snow in Minnesota over the prior winter lead to plentiful dry fuels across the southern portions of the state. This fire's fuel type was primarily tall grasses, which is typical of the area. This fire presents a particularly interesting case because there were also convective storms that occurred in the vicinity of the smoke plume and the radar, possibly aided by the cold front that passed through the area. This area has simpler topography and as such the beginnings of this fire were more easily noted in the polarimetric variables.

Strong southerly winds pushed the smoke plume from this fire to the north, close to Minneapolis until the cold front passage later, which changed the direction of smoke transit.

Detection and Replacement

Figure 5.17 shows the important dual-polarization variables for three selected time steps that will be discussed further in this section. The Snake Trail fire was first visible on the KMPX radar 18:19:51 UTC, and shortly after a strong plume with Z values exceeding 30 dBZ were seen. Strong southerly winds pushed the plume to the north over the metro area of Minneapolis. The plume exhibited low ρ_{hv} values between 0.4 and 0.6, and was quite distinct from regions of clutter and other scatterers near the radar. Z_{DR} exhibited low, negative values close to 0 dB, nearest to the fire as seen in all time steps. However, the further from the active fire, the higher the Z_{DR} values. This relationship was seen in the previous cases, but it is more prevalent for this case. Perhaps strong winds made it more visible with grass and other pyrometeor particles being oriented preferentially in the horizontal plane, leading to higher Z_{DR} values. In Figure 5.17c, we can see that there are areas of convective precipitation to the east of the smoke plume, as indicated by the black box. This is indicated by much higher Z , ρ_{hv} near or at 1.0, and Z_{DR} around 0 dB, which indicates primarily spherical hydrometeors like rain.

This is of interest for the detection and replacement algorithm such that we can see areas where hydrometeors directly overlap with the smoke plume. Given that the ρ_{hv} threshold is below the typical values of hydrometeors, those areas should be easily excluded from the smoke class. Figure 5.18 shows the detection and replacement algorithm output for the selected times. All three show that the algorithm does a good job picking up on the regions of smoke in comparison to the areas of hydrometeors and other scatterers. In fact, the algorithm detects when the smoke plume is separated initially by the passage of a cold front, which changes the direction of smoke propagation towards the east. Additionally

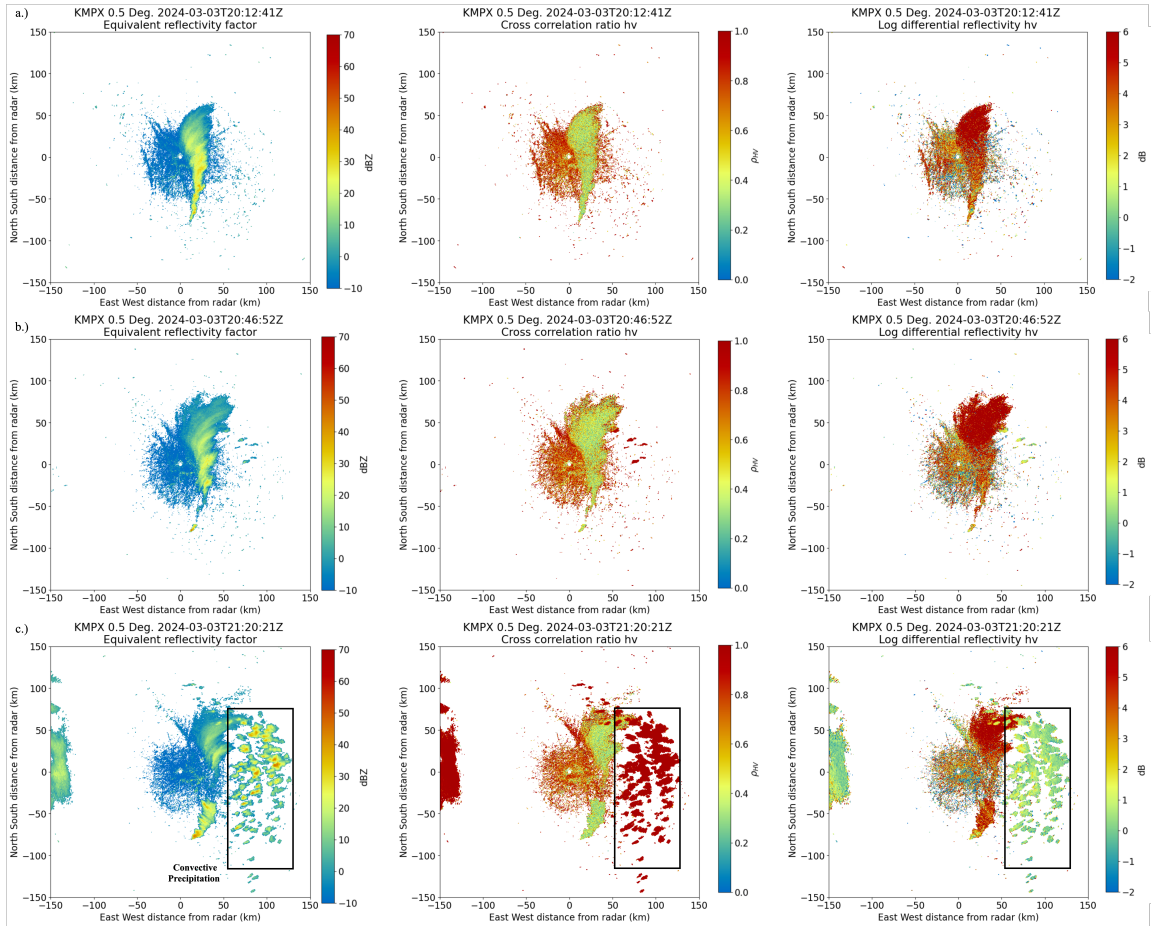


Figure 5.17: The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 20:12:41 UTC, b.) 20:46:52 UTC, and c.) 21:20:21 from the KMPX (Minneapolis, MN) radar.

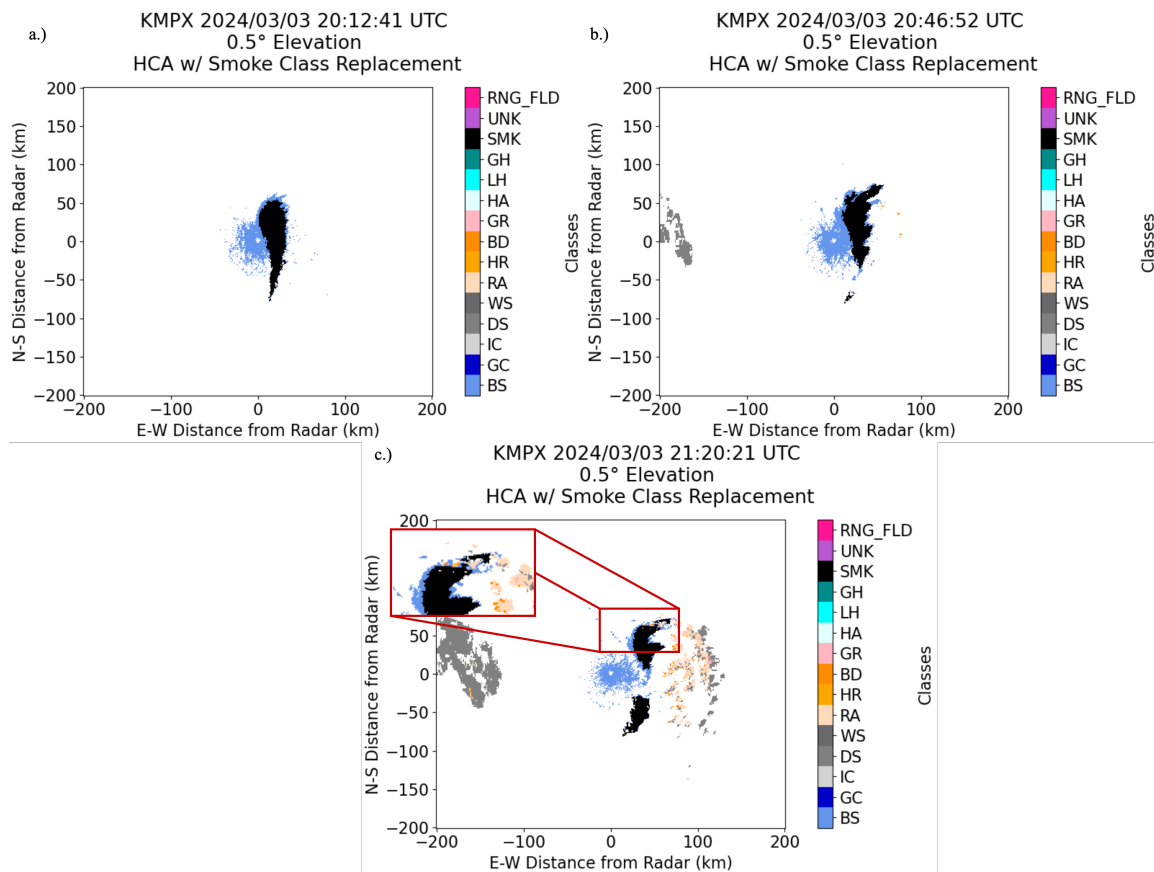


Figure 5.18: The final output of the detection and replacement algorithm for a.) 20:12:41 UTC, b.) 20:46:52 UTC, and c.) 21:20:21 UTC from the KMPX radar on March 3rd, 2024.

we can look at where there is overlap with convection cells and the far edge of the smoke plume. Figure 5.18c shows an close up look at this region in the inlay, which shows the smoke detection algorithm does a decent job of delineating between regions of smoke and pyrometeors.

Another application previously discussed was the use of change in number of smoke pixels between frames to indicate if the smoke plume is increasing or decreasing in size. Figure 5.19 shows how the smoke coverage ticker being included can provide some additional “first glance” information. This case has two to three distinct periods. In the pre-frontal time period, the plume is cohesive and moves to the north with southerly winds. During and after the frontal passage, the plume is more dispersed and is typically in two

distinct groups. With the change in wind direction, it is likely that the fire front changed and moved into new, available fuel. In Figures 5.19b and c the ticker indicates that the plume size is decreasing with the dispersal of the strongest region of smoke. However, with the change of fuel availability, there is a reinvigoration of smoke production nearest the fire and this change is highlighted by the change to increasing in Figure 5.19d. This is a great example of practical applications beyond simply adding a smoke class that could provide more information for the public, firefighters, communication, and more.

Finally, we can look at the ROC curve for this case, which includes one file per hour for the duration of the fire. The overall average AUC value is fairly similar to other cases presented in this thesis work, around 0.86. There are decently high individual ROC curves, but there is one time step that is lower and brings the average down.

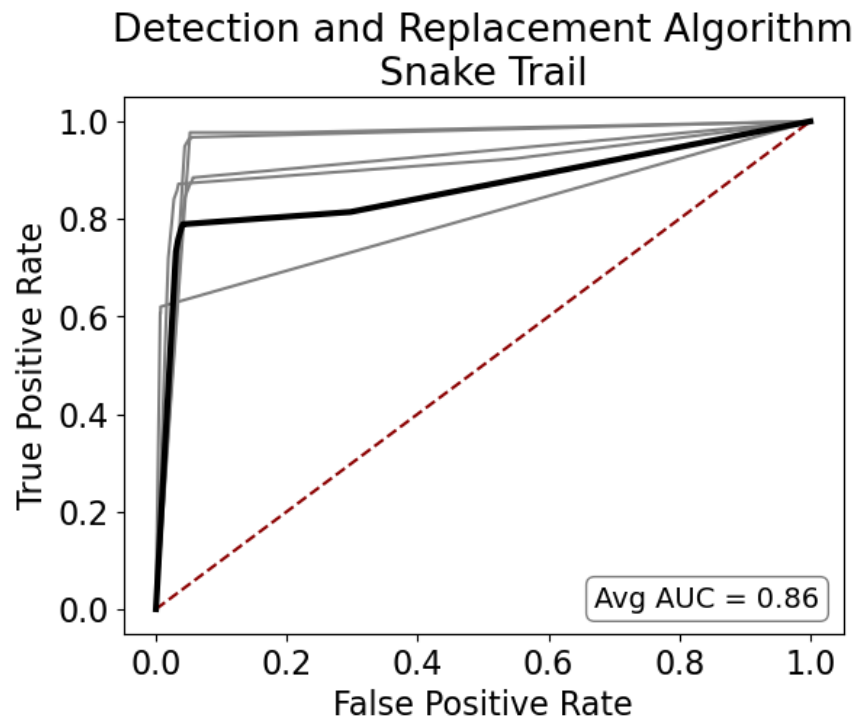


Figure 5.20: ROC curve for the detection and replacement algorithm for the Snake Trail fire on March 3rd, 2024.

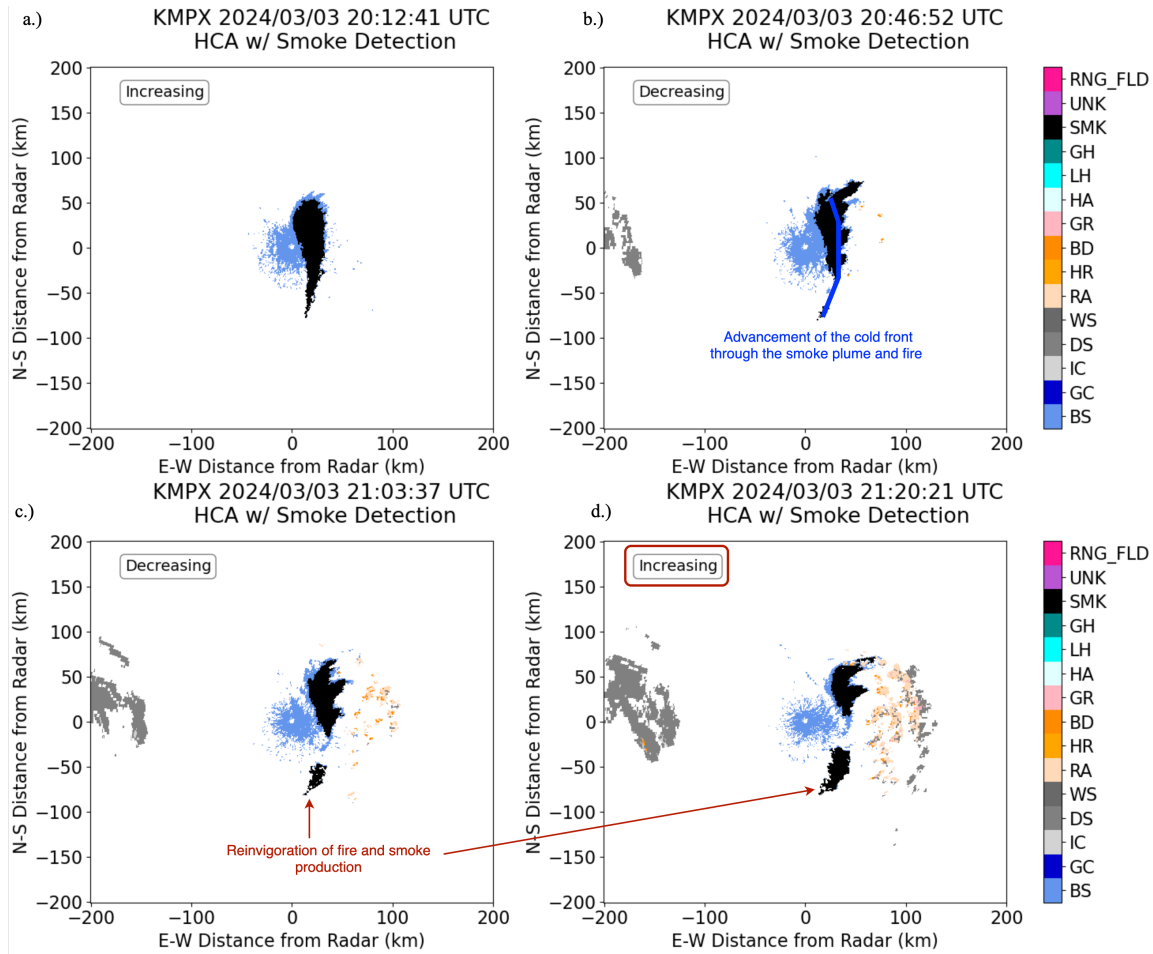


Figure 5.19: Selected times a.) 20:12:41Z, b.) 20:46:52Z, c.) 21:03:37Z, and d.) 21:20:21Z for the smoke plume size ticker addition to the detection and replacement algorithm. The dark blue line in b.) indicates the position and passage of the cold front through the fire and smoke plume. There is a noted reinvigoration of the fire and smoke production between b.) and d.).

Machine Learning Models

Unlike the previous two cases, data from this event was included in the testing and training portion of this model. However, the results for this section do not include the time steps that were randomly selected to be included in the machine learning models, and the times within 15 to 30 minutes were also excluded such that all of the individual scans were different enough from the included time. Figure 5.21 shows the final confusion matrices and ROC curves for each of the models. Unlike the previous two cases, the DT1 and random forest models perform slightly better than the detection and replacement algorithm. The same trend of over predicting for DT2 and under predicting for the random forest is still prevalent.

When looking at the machine learning predictions for two of the selected time steps, it's more obvious that the models are performing better than in the previous two cases. Figure 5.22 and 5.23 show the predictions for two of the selected time steps, and the main pattern to note is that compared to the previous two cases, the random forest has much better predictive skill. The AUC value is higher, and the number of correct predictions is much higher when comparing to the previous cases. The random forest model, notably, has very few incorrect classifications of smoke. There is still an under prediction bias, but there is also a trade off for fewer incorrect classifications of smoke where there is actually a biological scatterer class. The DT1 model has a better combination than the random forest and DT2 models, where there is a lower number of points being under predicted but slightly more being incorrectly classified as smoke. DT1 still does also have a better AUC value. For both the case where there was a cohesive plume and after the cold front separated it, the ML models have good predictive skill for showing the areas of smoke when compared to the output in Figure 5.18. It should also be noted that in Figure 5.18 the hydrometeor classification is included, but with the original BS selection for the machine

learning algorithms; they are not included in Figures 5.22 and 5.23.

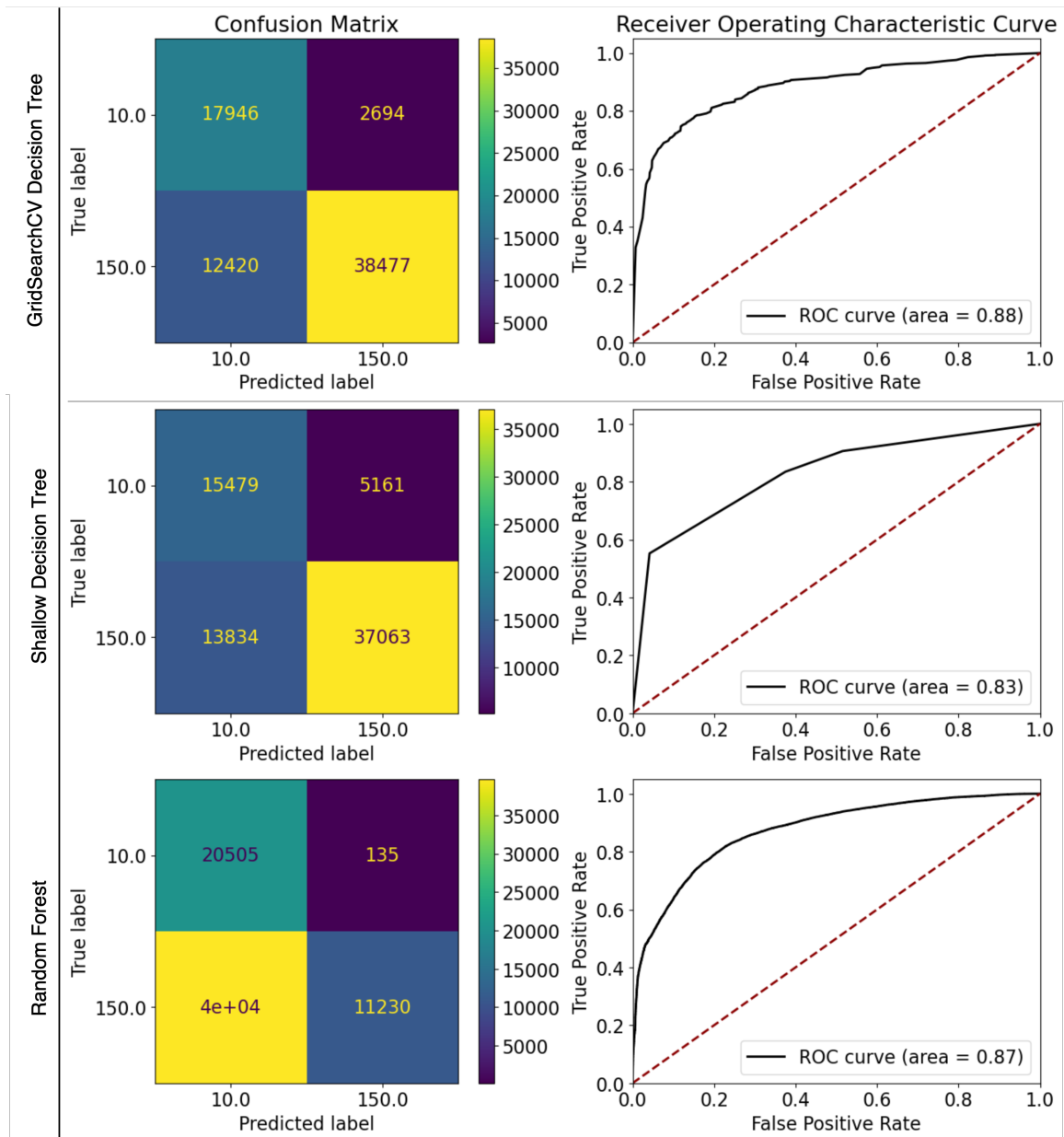


Figure 5.21: As in Figure 5.5 but for Snake Trail fire on March 3rd, 2024 from the KMPX radar.

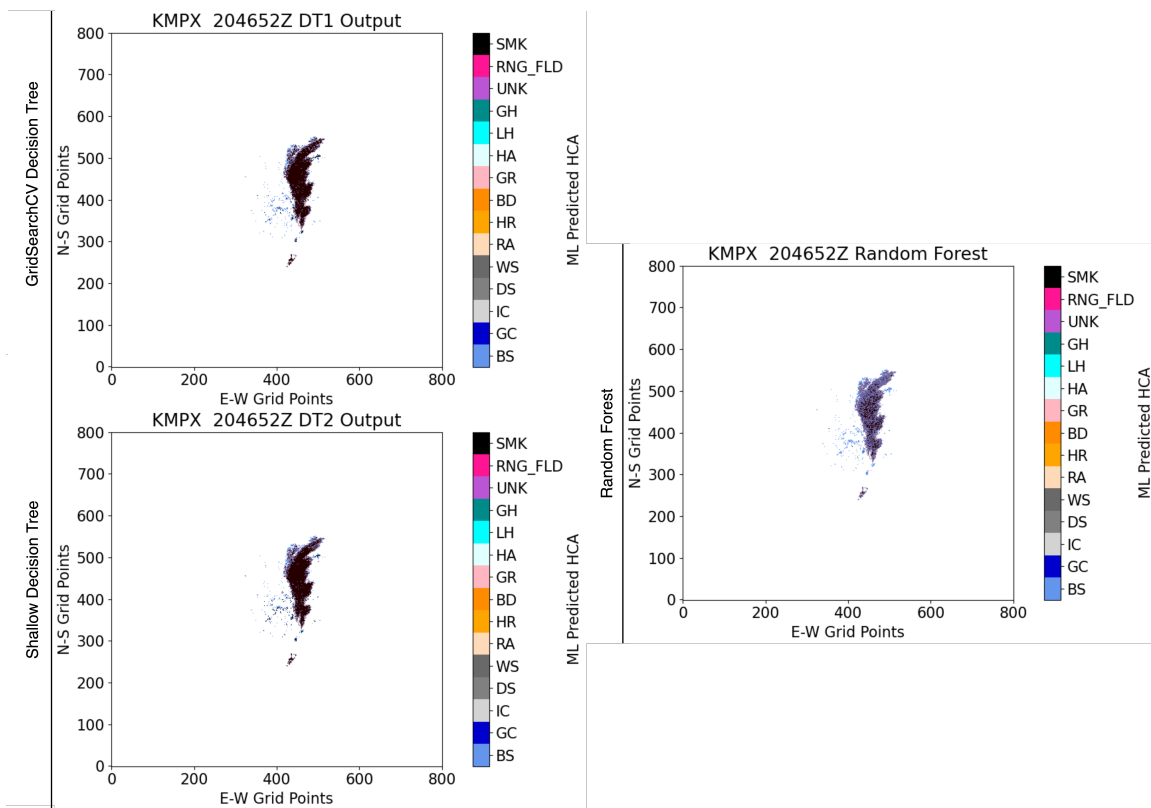


Figure 5.22: Machine learning predictions for 20:46:52 UTC on March 3rd, 2024 from the KMPX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.

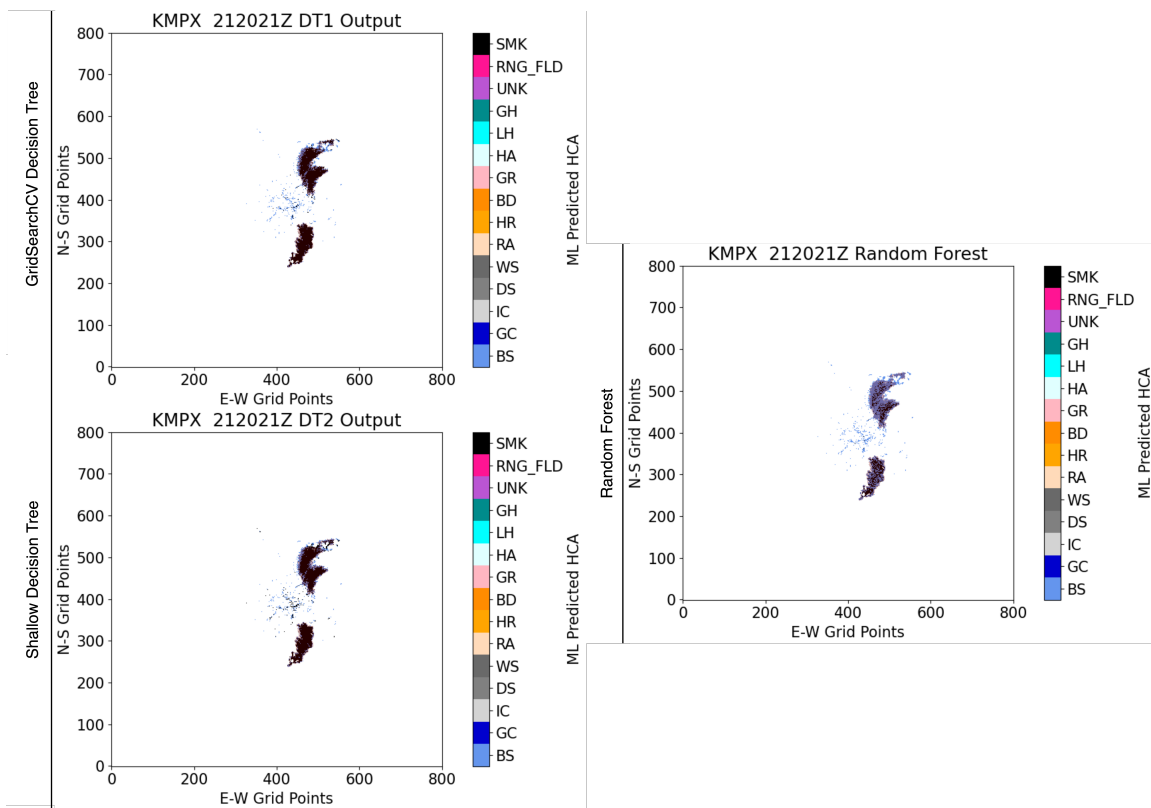


Figure 5.23: Machine learning predictions for 21:20:21 UTC on March 3rd, 2024 from the KMPX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.

5.1.4 Smokehouse Creek Fire, Texas

The Smokehouse Creek fire burned across the Texas Panhandle, just northeast of Amarillo. It was the largest wildfire in Texas history, burning approximately 1 million acres, and forcing evacuations across the eastern portion of the Texas Panhandle. The combination of dry fuels, low relative humidity, and very strong winds aided in its unprecedented growth. The fuels for this fire were classified as grasses of different heights. It took three weeks for firefighters to contain the full perimeter of the fire after it spread south following a cold front passage on February 27th, 2024. During this time, four different radars sampled the active fire and the smoke plume that extended into Oklahoma. During some

of the most widespread burning and smoke production, the WSR-88D's KAMA (Amarillo, TX), KVNK (Vance AFB, OK), KTLX (Twin Lakes/OKC, OK), and KOUN (Norman, OK) observed different portions of the active fire region (KAMA) and the smoke plume downstream (KVNK, KTLX, KOUN). There were also numerous other fires that ignited and burned in the vicinity as seen in Figure 5.24.

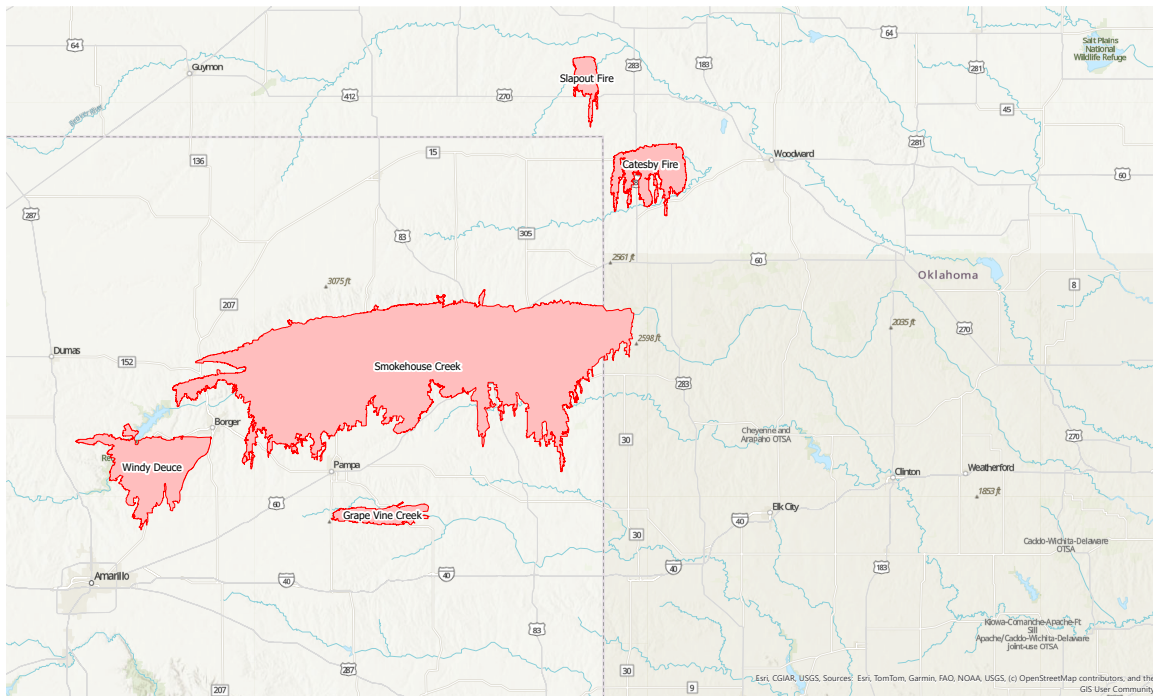


Figure 5.24: Final fire perimeters of Smokehouse Creek and other fires from the same time period across the Texas and Oklahoma Panhandles.

For this case, three days worth of data were tested with all the algorithms. Fires were first seen on the KAMA radar by 1915 UTC on February 26th, 2024 north of Amarillo, TX. This is when there were peak hours of the gustiest winds and driest air which aided in fire growth after ignition. Smoke was present on KAMA until 03 UTC on February 27th. Fire behavior decreased in the overnight hours, when winds were more limited and there was slight dewpoint and relative humidity recovery. By 16 UTC on the 27th, the fires were visible on radar again, the strong winds returned, and the critical fire weather began to peak

in the afternoon. The cold front passage starting around 2230 UTC across the fires changed the fire spread direction, and shifting winds blew smoke to the south, over Amarillo, TX through the early hours on February 28th.

Detection and Replacement

Figure 5.25 shows radar output for four times of interest for this fire over the three day period. The passage of the cold front is evident starting in 5.25b, indicated with the blue line, and then after the cold-front passage in 5.25c and d. The values of ρ_{hv} are fairly low for this case, typically around or below 0.4. During the peak smoke production, Z values exceeded 30 dBZ in a few different areas where the fire reached more fuels as in 5.25b. The Z_{DR} values for this case were often a bit harder to correlate to the microphysical aspects of pyrometeors. When the fires were stronger, the Z_{DR} was typically lower, between 0 and 2 dB, as in 5.25a and b. Further downstream, and when there was typically less turbulent mixing, the values ranged more towards 3-6 dB. The negative values of Z_{DR} near the radar are associated with wind turbines located around Amarillo, TX. Figure 5.25 also shows that there is not a correlation between ρ_{hv} and Z_{DR} . The values of Z_{DR} typically vary based on distance from the fire, while ρ_{hv} is a more consistent value throughout.

Figure 5.26 shows select detection and replacement algorithm outputs from this case over the three day period. The algorithm performed very well for this case, which is to be expected when compared to other cases. There is limited topography in this area and the smoke plume is more easily separated from clutter and other biological scatterers. The area of biological scatterers to the south of the large smoke plume in Figure 5.26b is correctly classified using this replacement algorithm, which is an improvement over the fuzzy logic classifier presented in Chapter 3. However at some times, there was an area of much lower ρ_{hv} and the detection and replacement algorithm erroneously classifies smoke. This feature is transient and does not persist through all times. The algorithm was able to detect

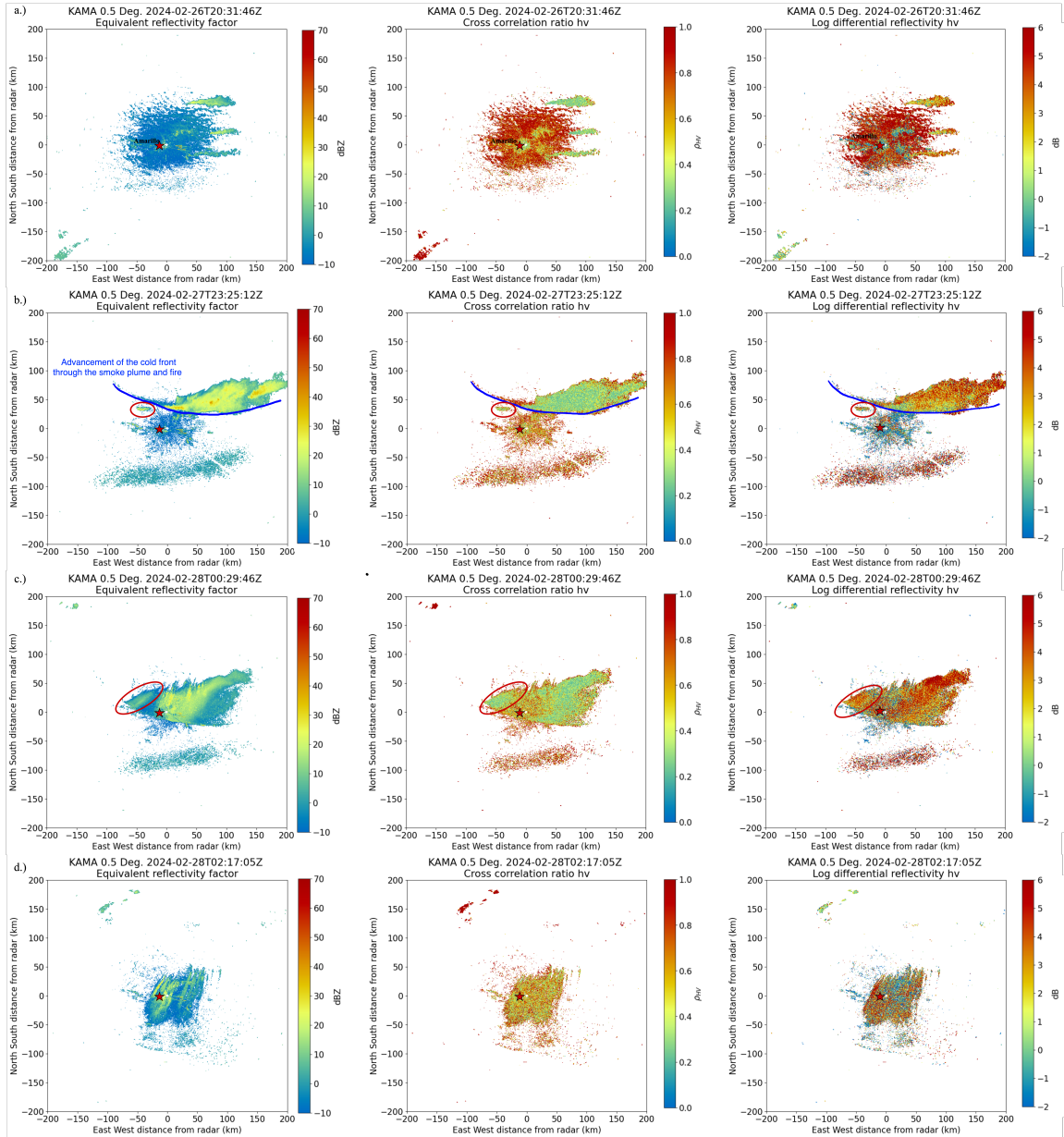


Figure 5.25: The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 20:31:46 UTC on 2/26/24, b.) 23:25:12 UTC on 2/27/24, c.) 00:29:46 UTC on 2/28/24, and d.) 02:17:05 UTC on 2/28/24 from the KAMA (Amarillo, TX) radar. The red star indicates the location of Amarillo, TX. The blue line indicates the approximate location of cold front, and the red ovals indicate the smoke area associated with the Windy Deuce fire.

multiple fires and their smoke plumes simultaneously during the 26-27th (Figure 5.26a) and the remaining smoke plumes after the wind shift on the 27th (Figure 5.26b-d). Most interestingly, in Figure 5.26c we can see that there are two separate areas of detected smoke embedded in the cold front. A fire started to the west of Smokehouse Creek, which was later named Windy Deuce (see red ovals in Figures 5.25b-c and 5.26b-c), just prior to the cold front passage. The algorithm was able to distinguish between the smoke and an area that is just classified as BS which is likely dust and/or bugs that were stirred up by strong winds. This is a promising instance of the algorithm performing well in a complex case where the fuzzy logic classifier would struggle.

Beyond the visual comparisons between the areas of very low ρ_{hv} in 5.25 and the smoke class in 5.26, we can also look at the ROC curves for each day, with randomly selected times used to make up the individual ROC curves. Each day's average AUC is above 0.80 but the differences between the first two days and the last reflects the difficulty of separating the smoke from the surrounding clutter and biological scatterer fields. During the first two days of fire weather, the smoke plumes were typically displaced from the clutter around the city and there was less potential for misclassifications. They were also typically distinct with regards to having very low ρ_{hv} . During the 2/28 period, the wind shift associated with the cold front passage changed the smoke direction to be over the city of Amarillo and the wind turbines surrounding it. This made it more difficult for the detection algorithm to determine the correct areas of smoke. During this time, the smoke production was also lower, and the general Z values were much lower, which could have made it more difficult for the algorithm to detect the portions of the plume that were farther away from the fire. The individual ROC curves in 5.27c still show that most times would have a higher AUC value than the average.

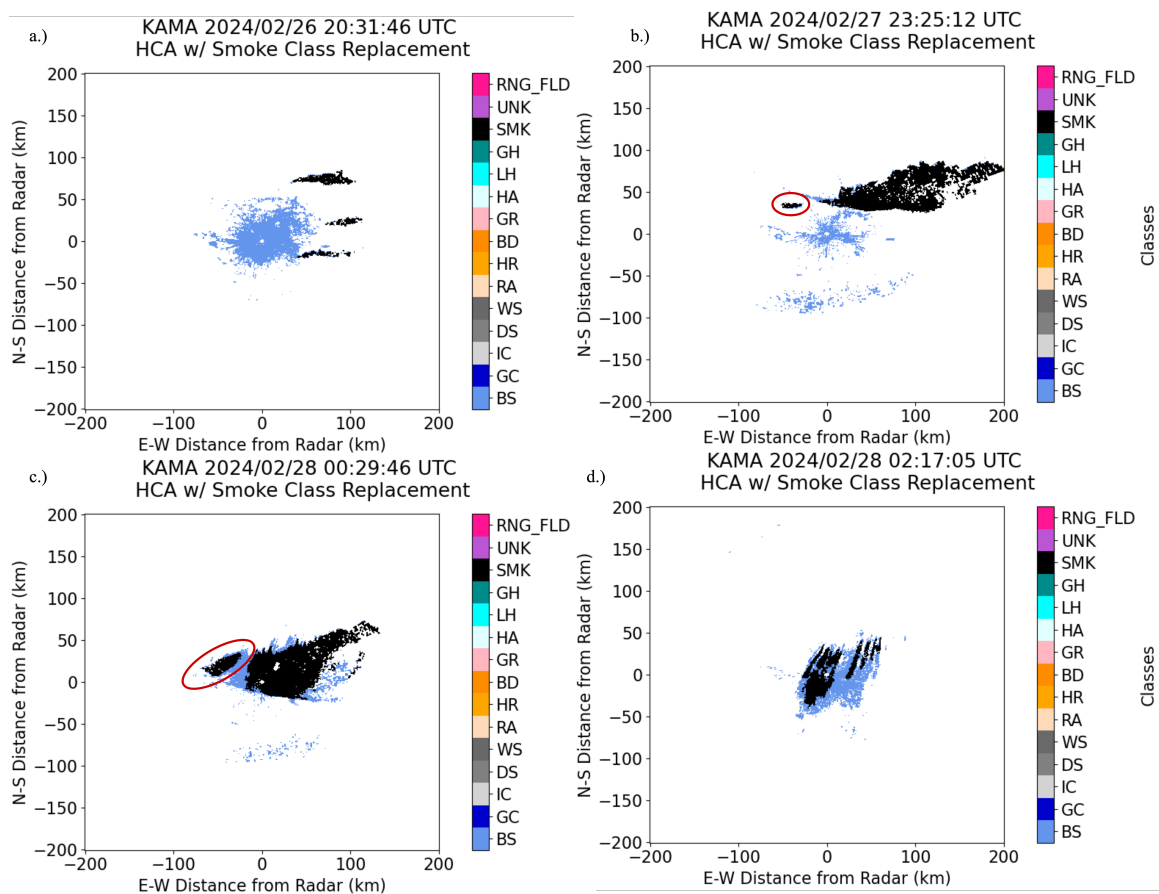


Figure 5.26: The final output of the detection and replacement algorithm for a.) 20:31:46 UTC on 2/26/24, b.) 23:25:12 UTC on 2/27/24, c.) 00:29:46 UTC on 2/28/24, and d.) 02:17:05 UTC on 2/28/24 from the KAMA (Amarillo, TX) radar. The red ovals indicate the smoke area associated with the Windy Deuce fire.

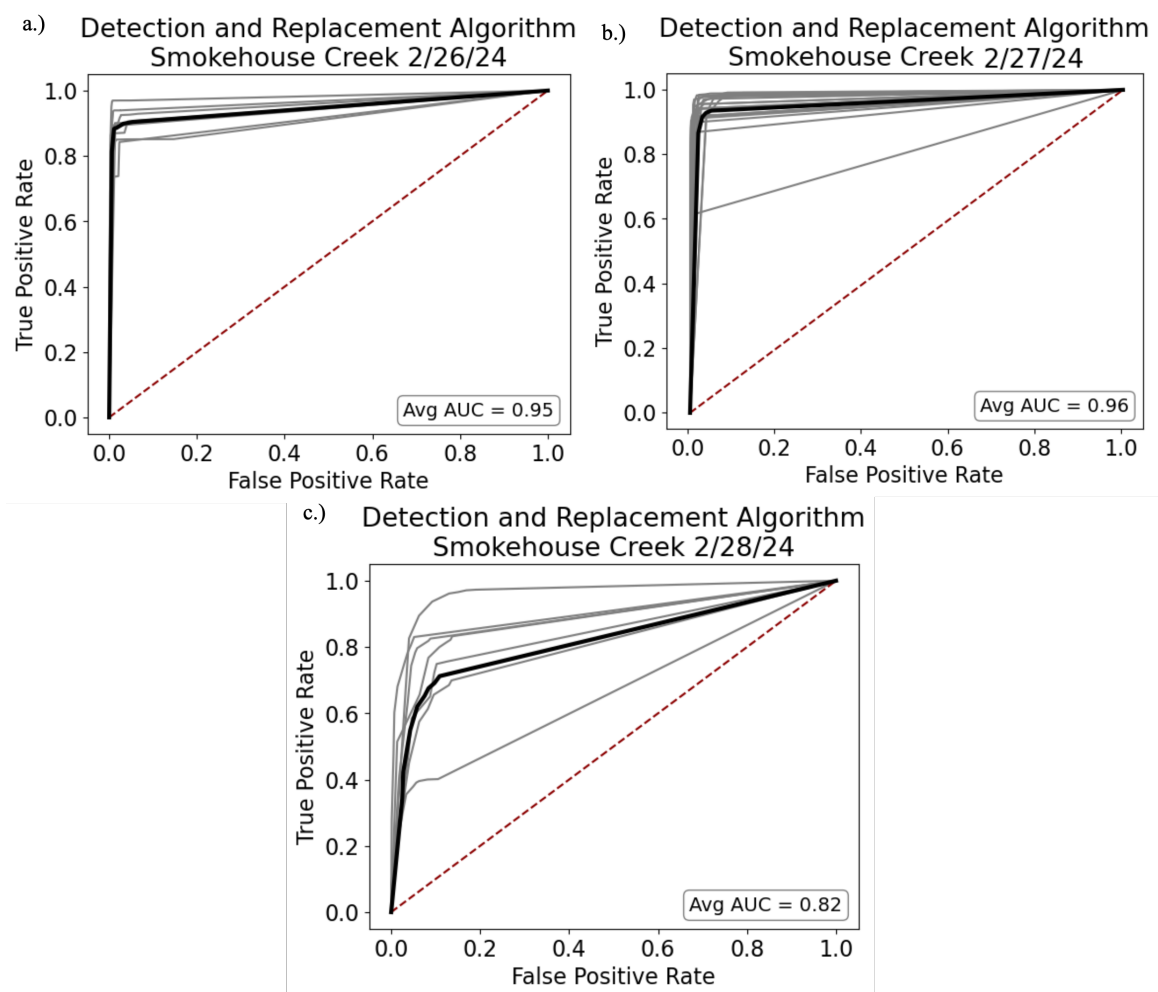


Figure 5.27: ROC curves for the detection and replacement algorithm for selected times on a.) February 26th, b.) February 27th, and c.) February 28th.

Machine Learning Models

For this case, some scans from 2/26 and 2/27 were included in the machine learning training and testing, but times from 2/28 were not. Upon plotting the 2/28 scans for the detection and replacement algorithm, it was decided that the more limited skill of that algorithm might also provide negative impacts to the overall predictive skill of all the models. Nonetheless, these time steps will be analyzed in this section.

When looking at the metrics for 2/26 and 2/27 in Figures 5.28 and 5.29, the AUC values from the ROC curves are very high for the random forest and DT1 models. However the confusion matrices show different reasons for why they might have such a high AUC value. The DT1 model has the higher of the two and the confusion matrix shows a decent balance for correctly classifying the areas of biological scatterers and smoke. While DT1 does have more false positives, it has way fewer false negatives and is not as prone to under predicting as the random forest. This follows the trend that has been seen in previous cases. The random forest is better at predicting smoke on 2/26 and 2/27 and this is likely driven by some of the scans being included in the testing and training of the models. The metrics for 2/28 in 5.30 show much lower AUC values for each model. All three machine learning models were prone to under predicting, which is similar to the outputs from the detection and replacement algorithm. Because there was no inclusion of data at times after the cold front passed, it makes sense that the skill is lower for these times. There is still agreement that under the same conditions, the detection and replacement algorithm and the machine learning models struggle when there is overlap between smoke, clutter and, biological scatterers.

Finally, we can look at some predictions for the Smokehouse Creek fire in Figures 5.31 and 5.32. As previously mentioned, neither of these scans were included in the training and testing of the models, though files from 2/27 were included. When comparing the model

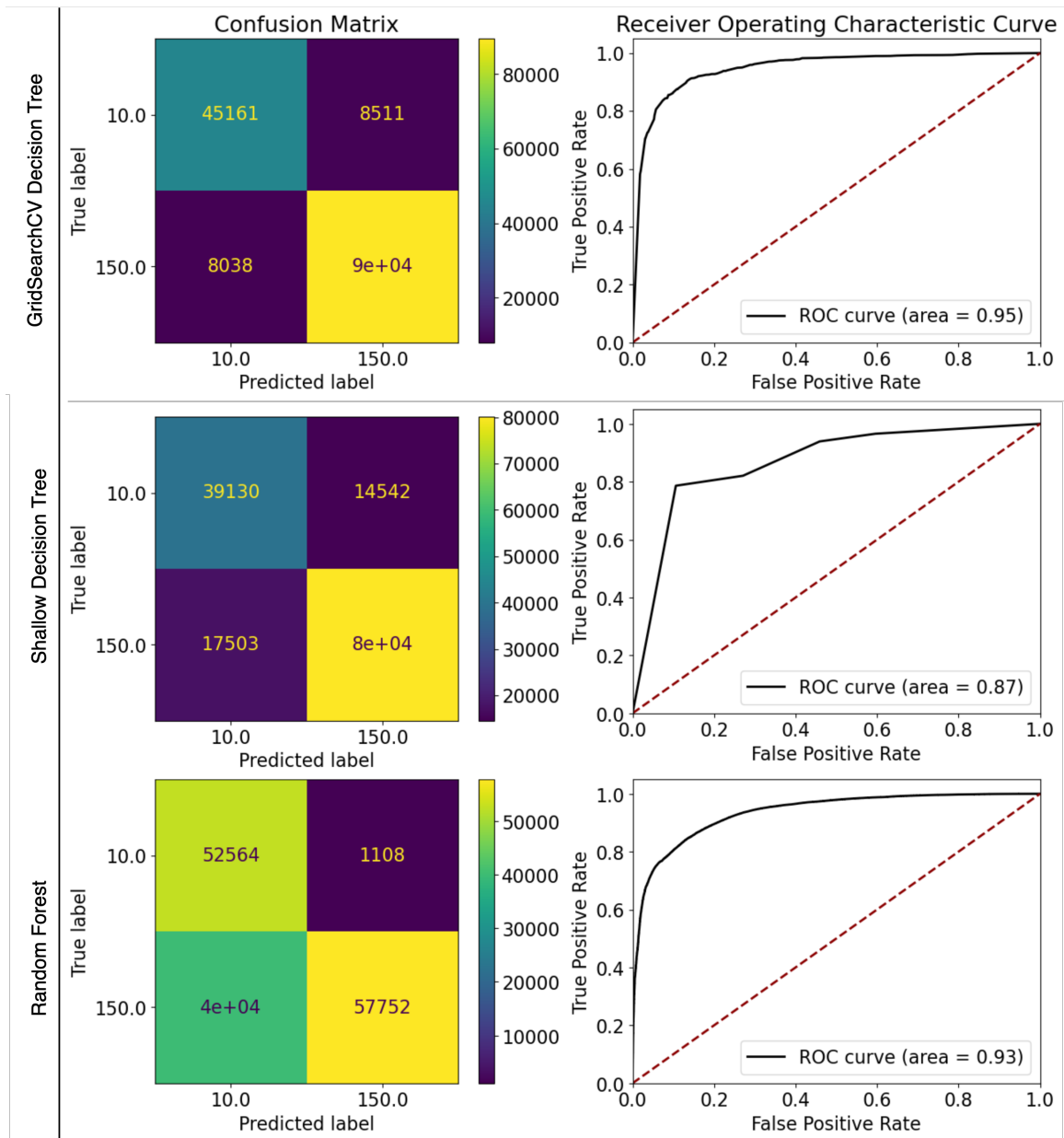


Figure 5.28: As in Figure 5.5 but for Smokehouse Creek and other fires on February 26th, 2024 from the KAMA radar.

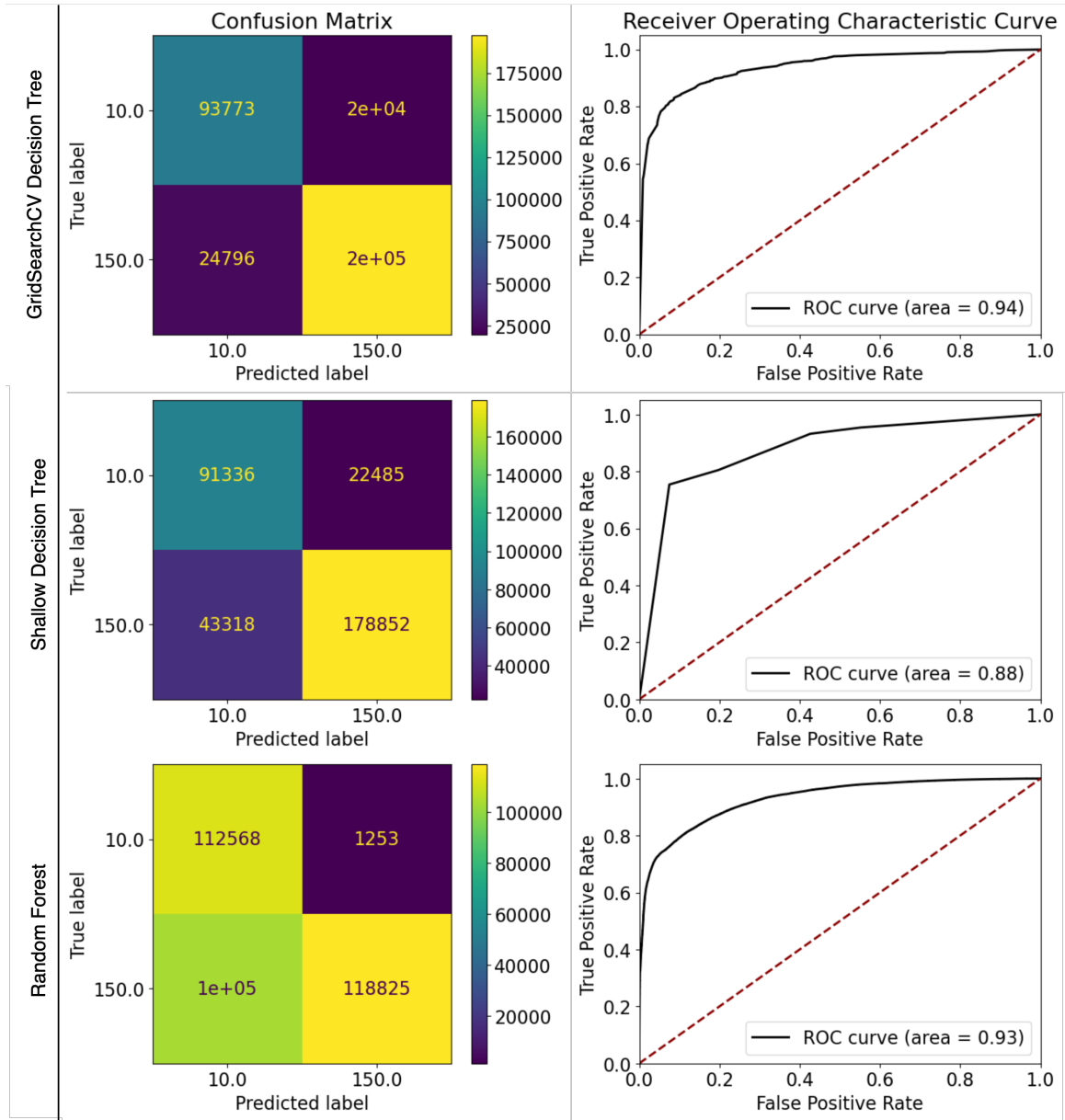


Figure 5.29: As in Figure 5.5 but for Smokehouse Creek and other fires on February 27th, 2024 from the KAMA radar.

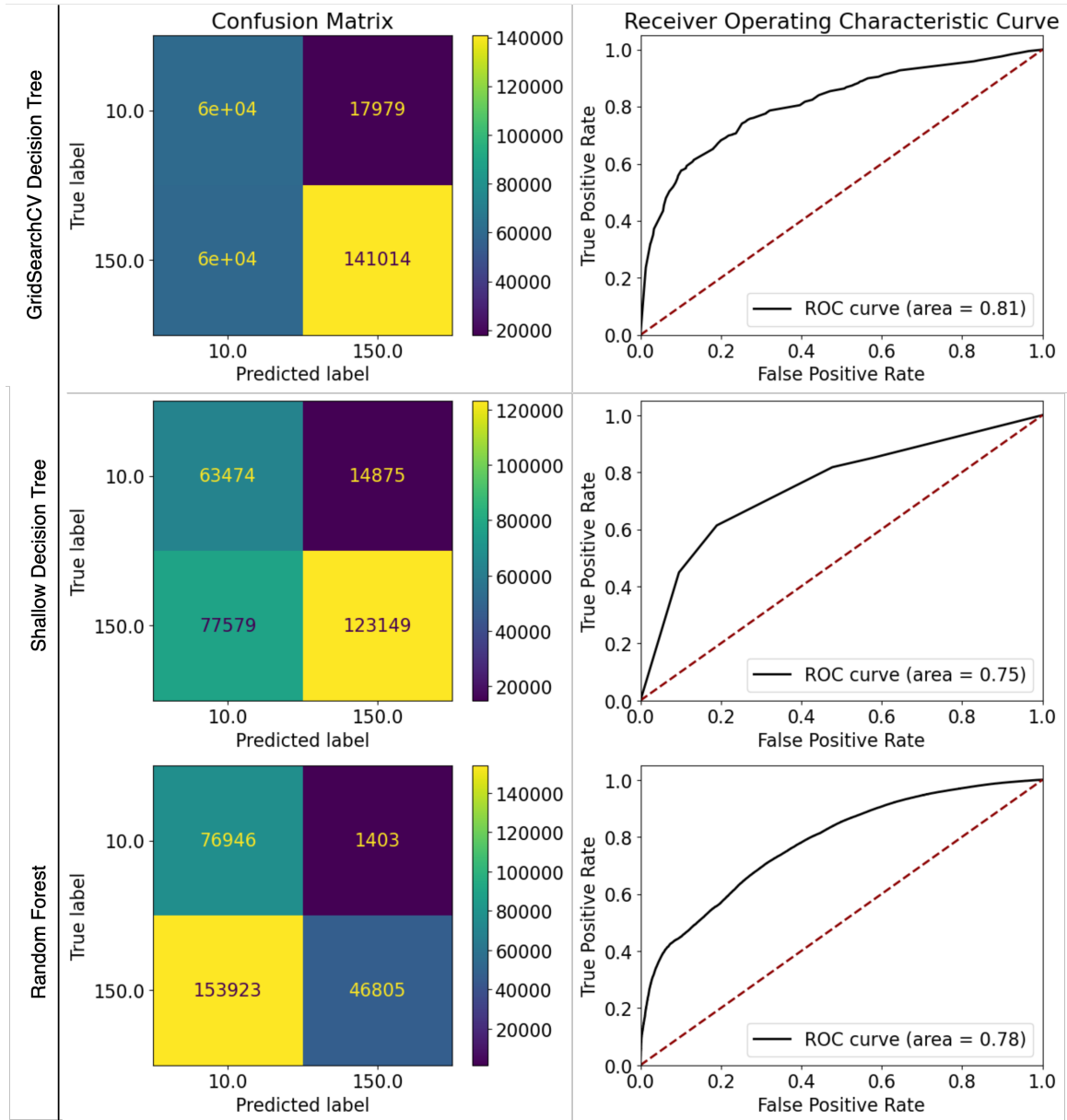


Figure 5.30: As in Figure 5.5 but for Smokehouse Creek and other fires on February 28th, 2024 from the KAMA radar.

outputs in Figure 5.31, the actual smoke plume from the Smokehouse Creek fire and the start of the Windy Deuce smoke plume are correctly classified and look very similar to the output in 5.25b. The subtle differences for the decision tree models are in areas of clutter along roads and where there are wind turbines incorrectly classified as smoke. This problem is typically resolved in the detection and replacement algorithm with a Z_{DR} threshold. With the DT2 model especially, there is little to no influence of Z_{DR} for predictions, which we can see in Figure 4.4. For the random forest model, there is still a decent amount of correct predictions, with less under prediction, and better classification of the clutter and biological scatterer. The random forest does the best at delineating between the actual smoke and areas that might look similar in fields like ρ_{hv} . The detections in Figure 5.32 are actually better than those from the detection and replacement algorithm with regards to spatial coverage. Detections typically cover nearly all of the smoke plumes, even the areas south and east of Amarillo, TX that are missed by the detection and replacement algorithm. Perhaps a lesser weight to Z or a different “threshold” would help detect the areas of weaker returned Z values. However, there are still issues with incorrect classification of clutter (as smoke) as discussed in Figure 5.31. This lowered the AUC farther when compared to the correct classification areas from the detection and replacement algorithm.

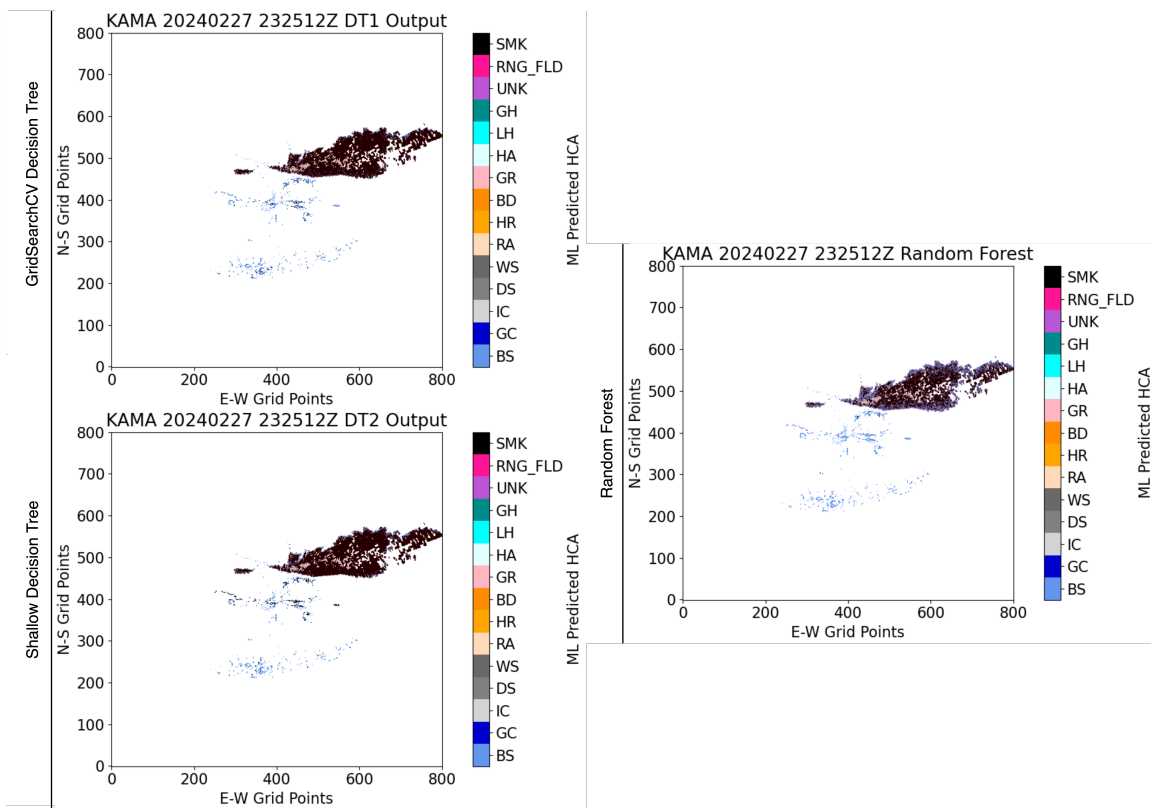


Figure 5.31: Machine learning predictions for 23:25:12 UTC on February 27th, 2024 from the KAMA radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.

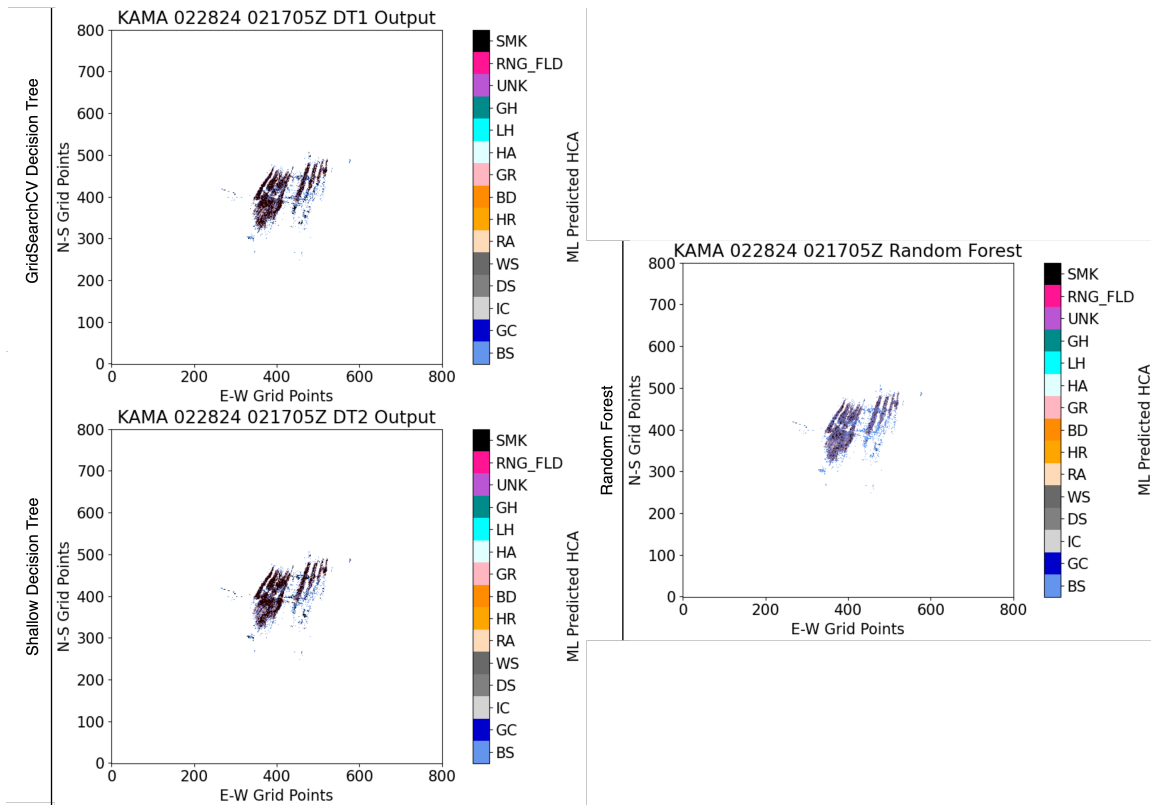


Figure 5.32: Machine learning predictions for 02:17:05 UTC on February 28th, 2024 from the KAMA radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA.

5.1.5 Labor Day 2020 Fires, Oregon and Washington

Over the Labor Day weekend period in 2020, September 7th to the 9th, fire weather conditions aligned to support extremely large fires on the west side of the Cascades in Oregon. A late summer drought, a strong east wind event across the Cascades, and ignition sources lead to fires in forested areas that had seen limited fire activity in recent years. From September 7-9th, numerous fires burned almost as much area on the west side of the Cascades, approximately 292,500 hectares, within a 48-hr period as in the prior three decades combined (406,128 hectares) [36]. The five largest fires, focused along the Oregon Cascades reached 50,000 hectares [36]. Figure 5.33 shows four of the largest fires in Oregon



Figure 5.33: Final fire perimeters for Labor Day 2020 fires across the Cascades in Oregon and Washington.

with Archie Creek burning farther south. Between an estimated 6 and 7 US billion dollars worth of property were destroyed by fire [37]. The Riverside fire burned 138,151 acres in the Mt. Hood National Forest and was started at the Riverside Campground along the Clackamas River (Figure 5.33). The Portland WSR-88D radar (KRTX) was approximately 113 km away from the Riverside fire and around 81 km the Big Hollow fire along the Cascades in Washington. The other fires were located further away, generally along the west side of the Cascades.

Detection and Replacement

Figure 5.34 shows the radar output for three selected times, one from each day of the fire(s). This case is fairly complicated with regards to topography and changes in targets throughout each day. With onset of strong easterly winds on 9/7, two preexisting fires, Beachie Creek and Lionshead, experienced an increase in fire behavior and smoke productions and were seen on radar (Figure 5.34a). During this time there was also a decent amount of clear air returns around the radar that has lower ρ_{hv} . However, the fires' ρ_{hv} values were below 0.4 throughout most of the plume, and were easily distinguished. With stronger, turbulent updrafts over the fires, we can see that the Z_{DR} values are much lower, between -1 and 1 dB. The rest of the clutter and clear air returns had a much higher typical Z_{DR} values. On 9/8 the smoke production fluctuated with the strength of the easterly winds. In the afternoon and evening hours on 9/8, daytime heating and mixing in the boundary layer helped lower dewpoints, increase temperatures, and eventually strengthen the winds. Figure 5.34b and c show the smoke plumes during peak times of fire behavior and smoke production. Z was typically greater than 15 dBZ, with areas closer to the fires up to 30-40 dBZ. Heavy timber fuels, strong winds, low dewpoints, and steep topography supported extreme fire behavior. The production of smoke and the continued presence of low Z_{DR} values supports this.

Now we can look at the detection and replacement algorithm output for the selected times in Figure 5.35. For the first time in 5.35a, the two smoke plumes associated with the Beachie Creek and Lionshead fires are classified to the south. There is clear delineation between the noise and clutter around the radar, portions of the Cascade crest to the north, and the new smoke plumes. A new detection of these plumes could have indicated the potential for new fires and an increase in fire behavior with the onset and strengthening of the east winds. For the second time step in 5.35b, the smoke plumes are significantly larger and now the Big Hollow fire in Washington and the Riverside and Holiday Farm fires in

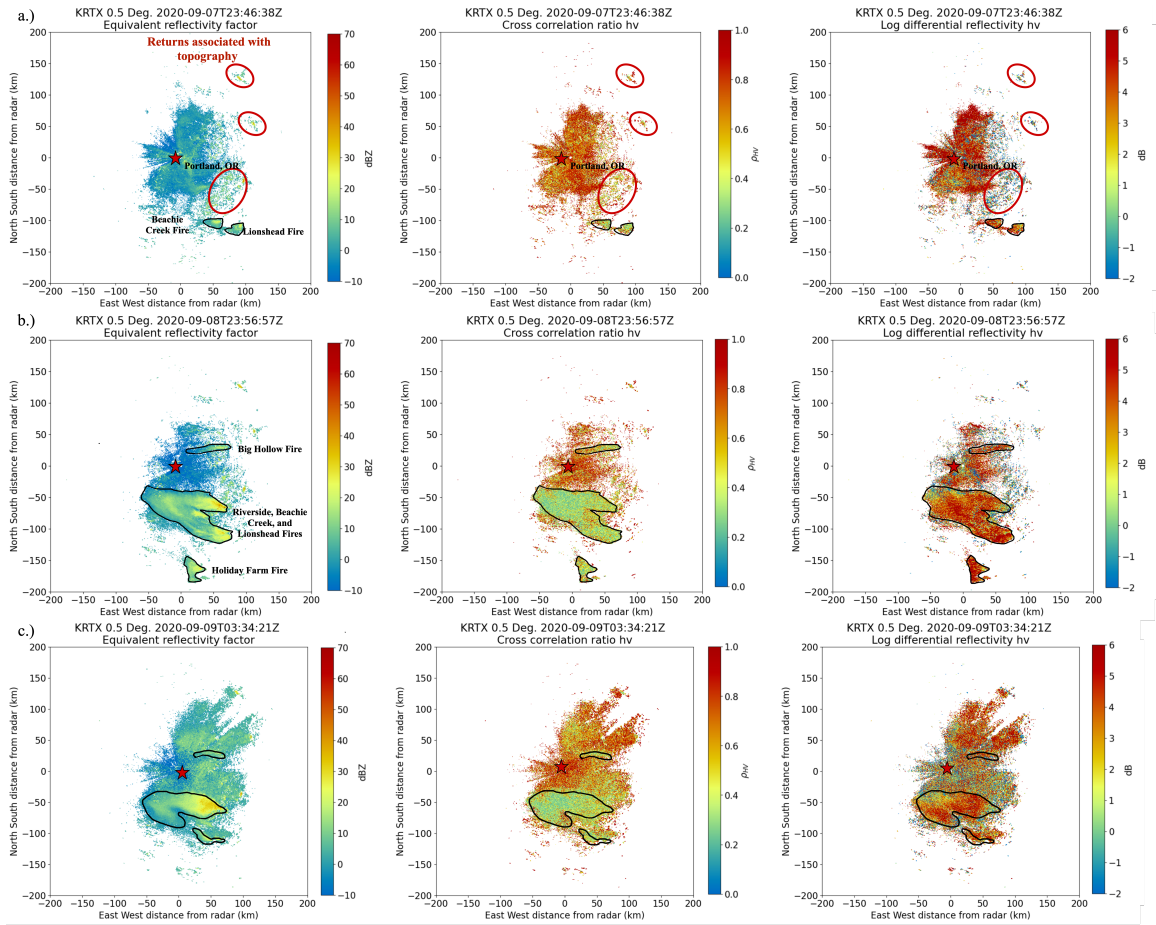


Figure 5.34: The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 23:46:38 UTC on 9/7/20, b.) 23:56:57 UTC on 9/8/20, and c.) 03:34:21 UTC on 9/9/20 from the KRTX (Portland, OR) radar. Red ovals indicate areas of returns associated with topography, and the black outlines show where smoke plumes are located at.

Oregon are visible and classified correctly as smoke. For this time, only a portion of the plume from the Big Hollow fire is missing a smoke classification. This miss is likely driven by slightly higher ρ_{hv} and lower Z values, with the plume being more diffuse downstream. Even at nearly 200 km south of KTRX, the Holiday Farm smoke plume is clearly classified as smoke. During the day on 9/8 the detection and replacement algorithm performed fairly well, with an exception of areas along the Cascade crest as seen in 5.35c. It is likely that the $SD(Z)$ threshold was not able to suppress all of the areas associated with the mountains, and that there was noise or biological scatterers in the area that lowered the ρ_{hv} as well. Further development of the detection and replacement algorithm to include satellite confirmation or topographic information would be able to suppress these areas further. In the overnight hours of 9/8 and into 9/9 there was an increase in other scatterers with low ρ_{hv} , likely some other form of biological scatterer or an increase in clear air returns. Figure 5.35d shows areas where there were incorrect classifications of smoke that due to low ρ_{hv} , not associated with the fires. The smoke plumes remaining at that time were still accurately classified, but these areas will lower the overall accuracy for the 9/9 case. The addition of other observations via satellite would help to remove these areas as well.

We can finally look at the ROC curves for these two days and because fires burned all day on 9/8 there are more individual ROC curves in light gray. The average AUC values for both days are fairly high, having values over 0.9. On 9/8, there were individual ROC curves with values higher than the average, which are some of the other times where there were less incorrect classifications in the areas of more complex topography. While the average AUC was high for this case, there were still some times with decreased performance, as in Figure 5.35c. For 9/9 in Figure 5.36b, the individual ROC curves with better AUC than average are also the earlier times, before the areas of incorrect smoke classification seen in Figure 5.35d begin to occur. Even with these areas lowering the AUC to below the average, the detection and replacement algorithm still works well for this case. The addition of the

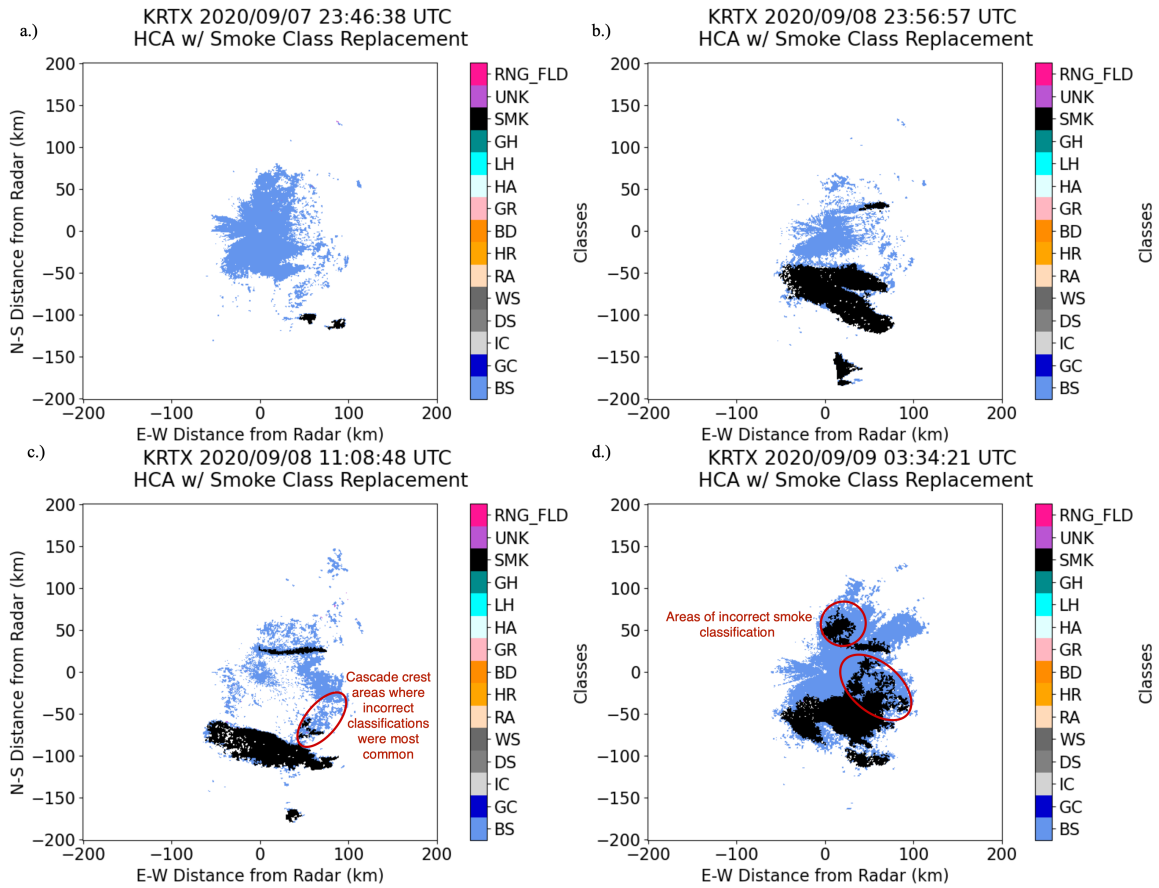


Figure 5.35: The Z (left column), ρ_{hv} (center column), and Z_{DR} (right column) for a.) 23:46:38 UTC on 9/7/20, b.) 23:56:57 UTC on 9/8/20, and c.) 03:34:21 UTC on 9/9/20 from the KRTX (Portland, OR) radar.

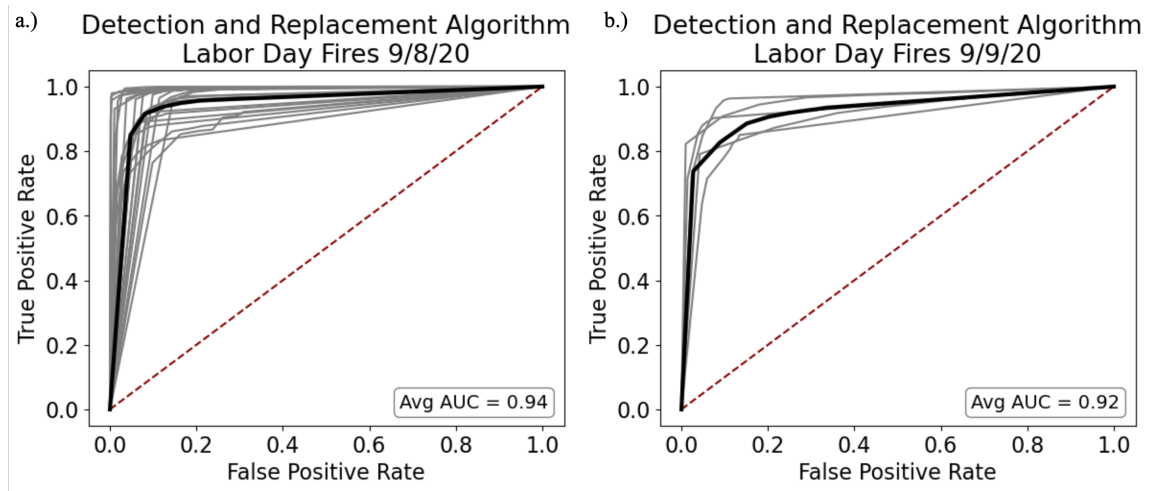


Figure 5.36: ROC curves for the detection and replacement algorithm for selected times on a.) September 8th and b.) September 9th.

$SD(Z)$ threshold improved detections for this case significantly.

Machine Learning Models

Like the first two cases discussed in this chapter, this case was not included in the training and testing of the machine learning models. In the earliest stages of training the models, there were still many incorrect classification issues with this case revolving around areas of complex topography and other clutter being classified as smoke. With a lower true positive rate initially, it was omitted so that there would be no learned bias towards classifying mountain clutter as smoke. Now, we can see how the machine learning models do when predicting smoke in a case with complex topography.

Figures 5.37 and 5.38 show the confusion matrices and ROC curves for the machine learning models on each respective day. These metrics for 9/8 cover scans for the entire day, and thus have a much higher count for each of the classes when compared to other cases. The decision tree models in the first two rows of Figure 5.37 show that there were many more counts where smoke was predicted but the true class was biological scatter.

This continues to follow the general pattern of over prediction of smoke for DT1, but more so for DT2. The random forest model performed as good as DT1 when comparing AUC values, but the confusion matrix shows a different reason for why that might be the case. While there were fewer instances of areas of smoke being incorrectly classed as biological scatter in DT1, there are significantly more instances (an order of magnitude) for areas of biological scatterers being incorrectly classified as smoke. So the random forest, while under predicting the areas of smoke, is still providing less false classifications of smoke where there is none present. For 9/9, the DT2 model has a significantly lower AUC, and the confusion matrix shows that the model is over predicting smoke and a good portion of the areas that should be classified as biological scatterers were incorrectly classified as smoke. By looking at column b in Figure 5.39 we can see that portions of lower ρ_{hv} associated with the mountains and other scatterers are now being classified as smoke. The plumes from the actual fires are well identified, but the over predicting in DT2 and in DT1 in a slightly lesser capacity make the machine learning models less useful. The random forest has the highest AUC, and though does under predict smoke in the plume areas, it has way fewer incorrect classifications of smoke. Perhaps adding a few good times from this case into the training and testing of the models at a later time might improve classifications in areas of complex terrain and when there are large numbers of other scatterers with low ρ_{hv} present.

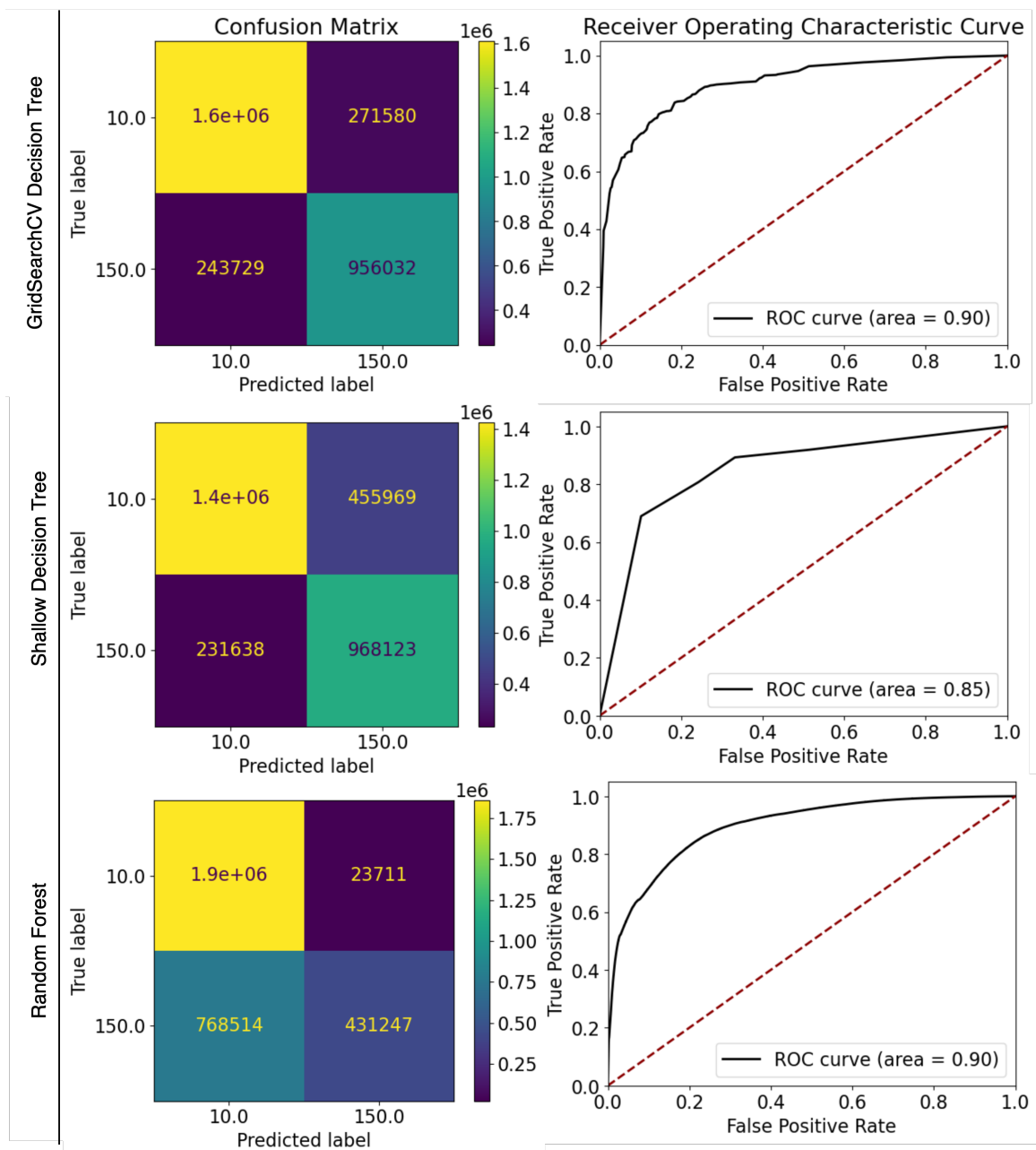


Figure 5.37: As in Figure 5.5 but for fires in Oregon and Washington on September 8th, 2020 from the KRTX radar.

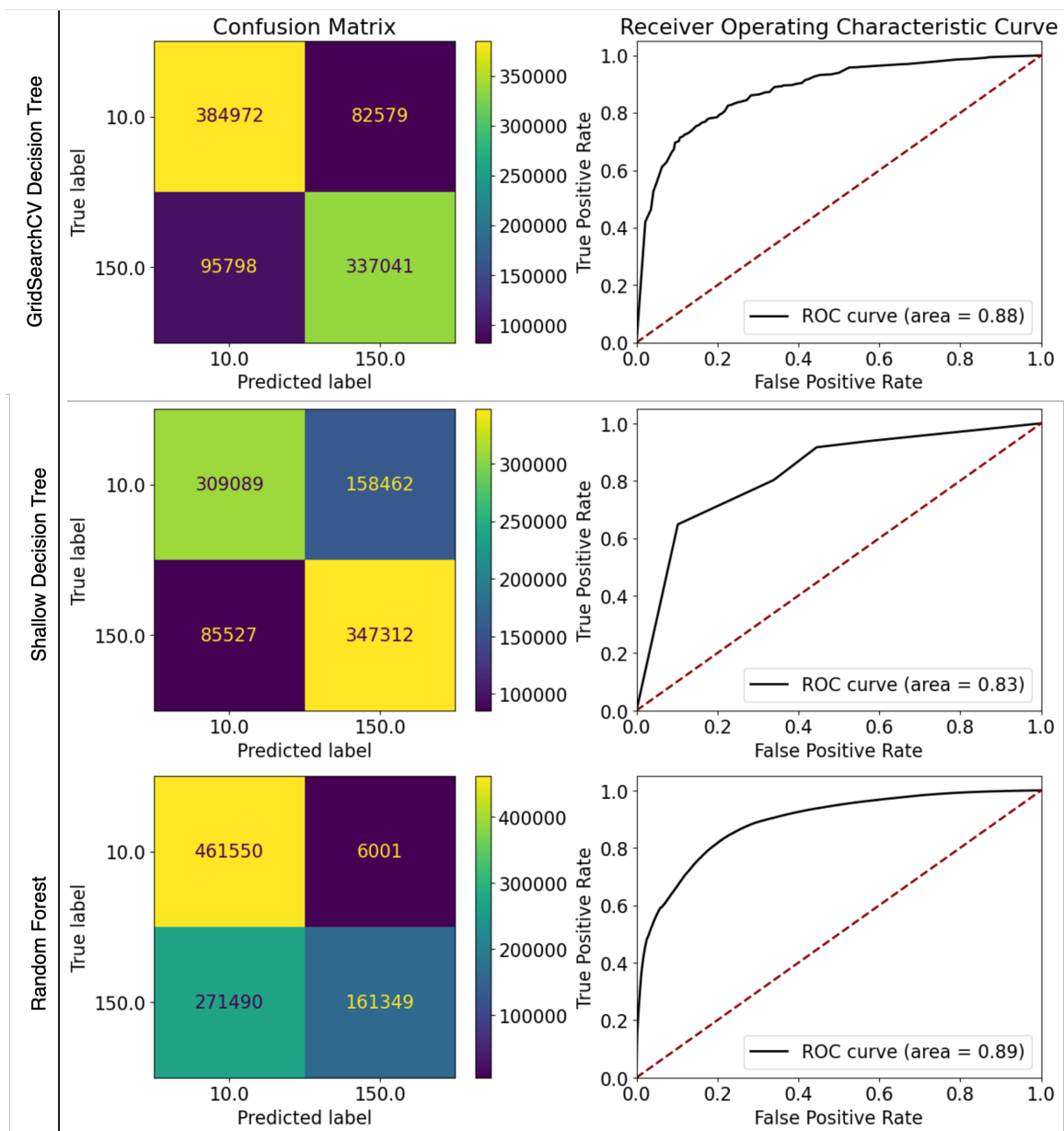


Figure 5.38: As in Figure 5.5 but for fires in Oregon and Washington on September 9th, 2020 from the KRTX radar.

Finally, we can compare some predictions from observations at specific times to the output from the detection and replacement algorithm. It is easy to see in the decision tree rows that smoke is being over predicted. In the case of the earlier time step in Figure 5.39a, areas of mountainous clutter are being classified as smoke, especially for the DT2 model.

This was improved on in the detection and replacement algorithm using $SD(Z)$, which shows limited areas of incorrect classification of the mountain areas in 5.35b, though there were still some areas seen at other times (5.35c). The random forest predictions for 5.39a show most portions of the smoke plumes being predicted correctly as smoke, though the density of predictions is still lower. The areas of incorrect predictions where there was clutter are not as prevalent. This concept also extends to the example in Figure 5.39b example, where areas of lowered ρ_{hv} that were incorrectly classified in the detection and replacement algorithm. This is suppressed in the random forest model especially when compared to the areas in the DT1 and DT2 models for the same time.

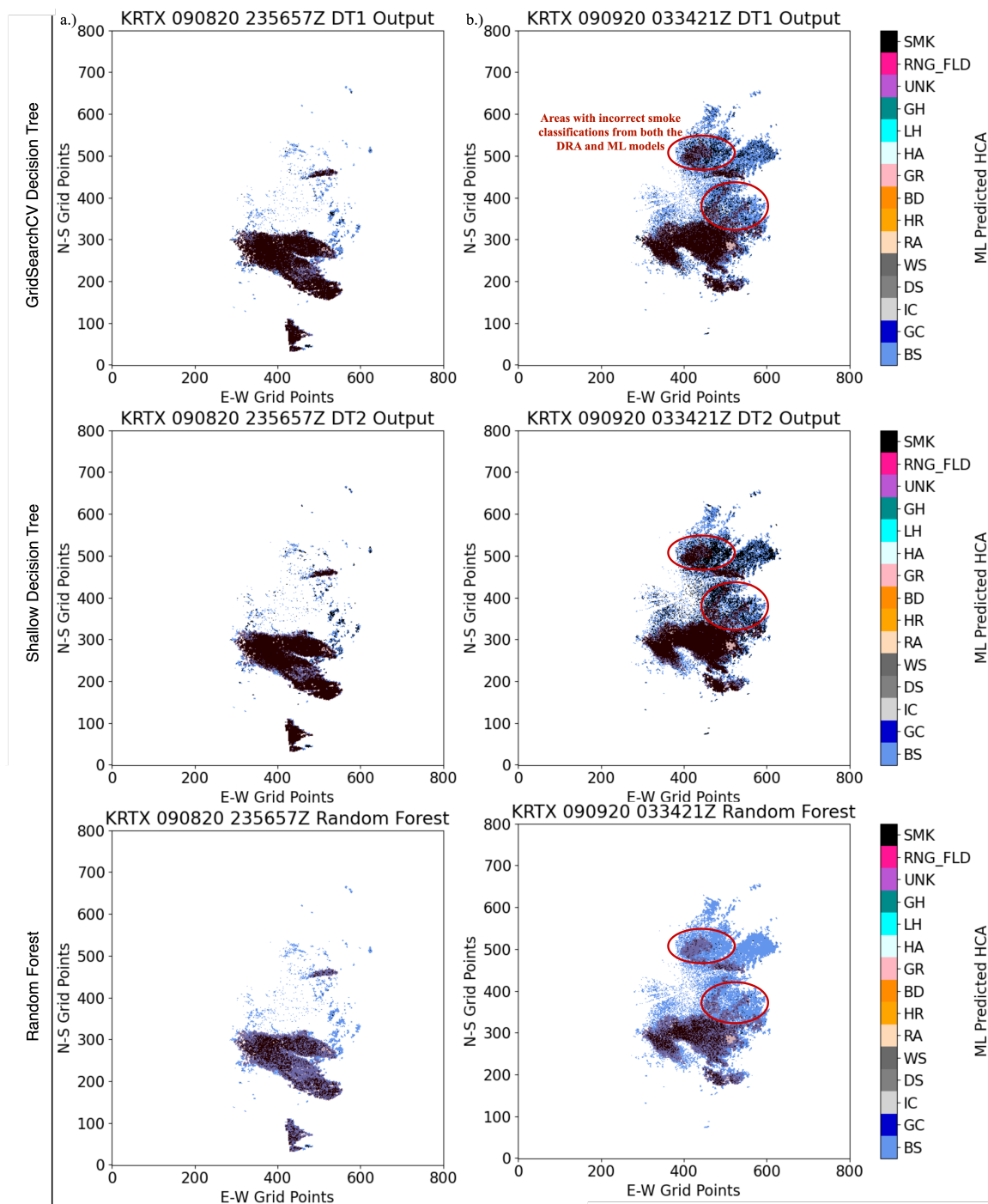


Figure 5.39: Machine learning predictions for a.) 23:56:57 UTC on September 8th, 2020 and b.) 03:34:21 UTC on September 9th, 2020 from the KRTX radar. The black color indicates where a grid point is predicted to be smoke, and the red filled contour indicates that area classified as smoke from the DRA. The red ovals indicate areas with incorrect classifications of smoke by both the DRA and the machine learning models.

5.2 Overall results

When considering the testing of all four models on a diverse set of cases from across the country, we can now provide some insight to which algorithm or models perform the best, and how they can be improved further. Between all four options, the detection and replacement algorithm consistently had the highest AUC values and was the best at correctly separating smoke from areas of other scatterers. It also works when hydrometeors were present and overall provides the best option for integration in operational settings. Further work to improve areas of biological scatterers like birds, insects, and chaff being included will help to further improve the overall accuracy. The inclusion of satellite data or the use of further machine learning for shape detection should be good options to tackle these issues. Other biological scatterers will have little to any returns in important fire bands, like Fire Temperature RGB or shortwave radiation. Some biological scatters have unique shapes to help further delineate smoke from things like birds. Various bird or bat species will often have ring-shaped returns on WSR-88Ds as they leave their roosts or caves. Adding more cases will also help refine the process and improve the thresholds, especially if any new relationships can be found in other variables, such as differential phase and associated products or fuel types. The integration of other detection and modeling tools like NOAA's Next Generation Fire System (NGFS), which already leverages artificial intelligence and machine learning, could increase the usability of this algorithm for operational use for both firefighters and meteorologists.

For the machine learning models, the DT1 model that used GridSearchCV to determine the best hyperparameters typically performed the best overall. It was often prone to predicting smoke where there was actually clutter, noise, or biological scatterers, but did so to a lesser extent than the simple DT2 model and was not as prone to under predicting like the random forest model. The use of all of the features with a more complex decision tree was

likely beneficial when comparing to the less complex decision model. The random forest model overall had the fewest number of false positives in the areas that the decision trees struggled with. There were issues however with under predicting smoke in some areas of the plume, which appeared as lower density predictions in the plume. Perhaps more data as input for this model might be more beneficial for having more accurate predictions. Moving forward, adding more cases and data to these models will be beneficial for improving their generalizability.

Chapter 6

Conclusions and Future Work

6.1 Summary and Conclusions

In this thesis, we have explored three different methods for smoke classification which includes updates to the legacy HCA fuzzy logic method, machine learning models, and a new hybrid algorithm that uses the current output from the HCA in addition to detections made from polarimetric information. To set the stage for new methods, we started with improving the current operational HCA to include a new smoke class. To create the necessary basis for a new class, statistical analysis was completed on 25-30 cases from across the country. By having geographically varying cases, a wider sample of fuels, topographical influences, and clutter were included to create more generalized information about the presentation of smoke plumes on the WSR-88Ds. Results from the statistical analysis were used to create membership function values for the fuzzy logic hydrometeor classification algorithm. Generally, smoke plumes showed low ρ_{hv} values, typically at or below 0.6. The values of Z were similar to what might be seen in classic weak convective cells, though in periods of intense fire growth and extreme fire behavior, Z values could reach up to 40-50 dBZ. The distribution of Z_{DR} values were often found to be related to the distance from the active fire. Nearest the fire, where there is often turbulent mixing of particulates, Z_{DR} values tended to be lower, between -1 and 1 dB. However, further downstream as the pyrometeors moved

with the wind, the values of Z_{DR} increase to be about to 6-8 dB. There is likely some link to preferential orientation of pyrometeors with the environmental wind to be in the horizontal, which gives higher, positive Z_{DR} values. For each of the cases, the range of values and the mean for each were recorded and used to create a set of membership functions centered around the mean. The detection algorithm was first created to highlight areas of smoke and create frequency density functions by using specific thresholds for each of the main dual-polarization variables to determine the aforementioned statistical information.

The new values for the smoke class were added into the the current fuzzy logic method, with the other classes as in Park et. al [9]. When running the fuzzy logic classification method on cases, comparing areas with known smoke to the classes provided, showed issues with distinguishing between smoke and biological scatterers. There is overlap between the membership function values for many of the variables like ρ_{hv} for these two classes (Table 3.1), which is why smoke is currently combined in the biological scatterers. Because adding this class gave limited information on what was smoke and what should remain as biological scatterers, a new method based on detection algorithm was explored next. The basis of the detection algorithm is in the dual-polarization variables for areas of smoke. Thresholds could be explored to contour and highlight areas of smoke using the lowest levels of the polar coordinate based Level-II data. Given that this algorithm was used and was improved upon for doing statistical analysis, it worked well at determining areas of smoke. To take it a step further and provide a new class, and to improve the visual aspect of classification, the replacement portion was developed. The contour or contours created in the detection step could be isolated with masking. By using gridding, the Level-II and Level-III data could be selected using the same mask. Given the created smoke mask, those areas in the original HCA output could have a new value for the smoke class by replacing the original numeric value. The addition of the $SD(Z)$ threshold helped to remove areas of clutter from complex topography that was plaguing areas where there was steep terrain

typically in the western portions of the United States.

Machine learning was the next extension for investigation for classification methods. Machine learning models are well suited to classification problems such as what is discussed in this work, but lack of “truth” data has hindered the widespread usage in radar. Given that the detection and replacement algorithm worked well in most cases, this data could be used as the truth data needed in machine learning models. To combat the more limited cases, data was used at the grid point level, which increased the order of magnitude of possible data significantly. New files containing the original HCA and polarimetric data, and the new HCA output including smoke were created for each time step of interest for the cases. These could be ingested in the machine learning models, and randomly selected, but quality controlled files were used for the training and testing data. From the cases selected, the randomly selected time needed to have a good output from the detection and replacement algorithm such that there was limited bias for training by having poor quality label data. From there, a limit was set to only use data that originally had a biological scatterer class, with a numeric value of 10. This limited classification to only be between two classes, making it binary as opposed to multiclass. The polarimetric data was then standardized to split into training and testing data using an 80/20 split. Given that the detection algorithm works similarly to a decision tree, this model and random forest were selected for investigation. Both methods had an initial model built that was prone to overfitting, with model accuracy scores close to 100%. To tune hyperparameters, GridSearchCV was used to extensively search for the best combinations of values. The two final tuned models were used to make predictions for smoke and to test the accuracy and skill. In addition, a simple decision tree model was built to emulate the detection algorithm. This tree had limited leaf nodes, and it chose to omit the two least important features, $SD(\Phi_{DP})$ and Z_{DR} , which were shown in the feature importances of the DT1 model (Figure 4.1).

Five cases were selected to show how the two main models of discussion, the detection

and replacement algorithm and the machine learning models, worked on fires with varying fuel types and geographical location. Cases with limited topography like Smokehouse Creek, Snake Trail, and the March 14th, 2025 Oklahoma wildfires performed the best with the detection and replacement algorithm. There was more limited influence from clutter and returns from mountainous topography in these cases. In places like the Pacific Northwest and Southern California, the mountains occasionally introduced areas of lowered ρ_{hv} which could not be suppressed by the $SD(Z)$ threshold. However, the detection and replacement algorithm seemed to perform best in all cases, as the machine learning models had some issues with over and under predicting areas of smoke. The decision tree models were often prone to over predicting smoke, and areas of clutter, clear air returns, and other biological scatterers were often included as smoke. The random forest model was prone to under predicting smoke, with the density of smoke grid points being much less than the decision tree methods. The tradeoff with this, however, is that there were fewer incorrect classifications of smoke that were meant to remain in the biological scatter class. It is likely that adding more cases to the training dataset and continuing to improve the data selection for input into the machine learning models will continue to improve on the results found in this study.

A goal for this work was to show options and viable pathways for improving classification of smoke and other biological scatterers, and the detection and replacement algorithm has the most upfront potential for integration for operational work. There are numerous pathways for improvement and options for its use in modeling and observational work. The machine learning aspect of this project is by no means a robust method; in fact, there are plenty more ways to improve what is presented in this work and better ways to handle the input data. The hope was to show its potential in the simplest way, such that more people might attempt to apply machine learning and more complex models like deep learning or better random forests to classification problems in the radar world.

6.2 Considerations for Future Work

In the previous section, the detection and replacement algorithm was noted for its potential for further use in the operational and modeling communities. One of the first improvements that should be explored is integrating satellite data into the algorithm. Questions of satellites' use have been prompted to the author on many occasions already. The inclusion of the important channels for fire detection from GOES and like shortwave infrared, fire temperature, GeoColor, Day Fire, and so on will be beneficial to improving the separation of smoke from biological scatter. Satellite use would be two fold. The first would be to quality control the inclusion of smoke in a final HCA output. For example, if there was not smoke or fire detected via the microphysics RGB and the fire temperature channels, then the subprocesses for the smoke HCA would not be run. This would remove areas of incorrectly classified smoke like were there is lowered ρ_{hv} for insects, birds, and chaff. Another method would be to improve the active fire contour. While it works in some cases, the inclusion of the fire temperature channel might be able to improve the accuracy of this. The inclusion of alerts from the new NOAA Next Generation Fire System could also help to improve the quality of smoke classification moving forward. Triggering the radar in the NWS WFO to run the smoke HCA if a certain rank of the alert is met, is an interesting option for example.

With regards to modeling, an improved detection and replacement algorithm also provides needed truth data for creating and improving smoke models. As noted in [33], the lack of accurate data with correct classifications makes it difficult for more complex machine learning models to be used. By making cases with smoke data available, there can be overall improvement of smoke modeling like in the Warn on Forecast (WoFS) smoke product [38]. Beyond modeling, in this manner, the inclusion of the detection and replacement algorithm in the Multi-Radar Multi-Sensor (MRMS) would add additional information for

the public and other users. For more operational uses, the inclusion of the smoke plume size change ticker adds information about how the smoke plume is changing. Modeling on how the size is anticipated to change could help to provide information to the public and emergency managers about smoke incursions into populated areas and roadways. There are numerous more avenues for expansion for the detection and replacement algorithm and hopefully with time it can be expanded for use by the public, firefighters, meteorologists, and emergency managers.

Bibliography

- [1] N. I. F. Center. “Wildfires and acres.” (2023), [Online]. Available: <https://www.nifc.gov/fire-information/statistics/wildfires>. (accessed: 9.30.2024).
- [2] J. Minor and G. A. Boyce, “Smokey bear and the pyropolitics of united states forest governance,” *Politcal Geography*, vol. 62, Pages 79–93, 2018. DOI: <https://doi.org/10.1016/j.polgeo.2017.10.005>.
- [3] K. C. Ryan, E. E. Knapp, and J. M. Varner, “Prescribed fire in north american forests and woodlands: History, current practice, and challenges,” *Frontiers in Ecology and the Environment*, vol. 11, no. s1, Pages e15–e24, 2013. DOI: <https://doi.org/10.1890/120329>.
- [4] L. N. Kobziar, D. Godwin, L. Taylor, and A. C. Watts, “Perspectives on trends, effectiveness, and impediments to prescribed burning in the southern u.s.,” *Forests*, vol. 6, no. 3, Pages 561–580, 2015. DOI: <https://doi.org/10.3390/f6030561>.
- [5] P. M. Fernandes, “Empirical support for the use of prescribed burning as a fuel treatment,” *Fire Science and Management*, vol. 1, Pages 118–127, 2015. DOI: <https://doi.org/10.1007/s40725-015-0010-z>.
- [6] N. W. Service. “Volume coverage patterns.” (2023), [Online]. Available: https://www.noaa.gov/jetstream/vcp_max.
- [7] N. C. for Environmental Information. “Next generation weather radar (nexrad).” (2024), [Online]. Available: <https://www.ncei.noaa.gov/products/radar/next-generation-weather-radar>. (accessed: 9.3.2024).
- [8] N. McCarthy, A. Guyot, A. Dowdy, and H. McGowan, “Wildfire and weather radar: A review,” *JGR Atmospheres*, vol. 124, no. 1, Pages 266–286, 2019. DOI: <https://doi.org/10.1029/2018JD029285>.
- [9] H. S. Park, A. V. Ryzhkov, D. S. Zrnić, and K.-E. Kim, “The hydrometeor classification algorithm for the polarimetric wsr-88d: Description and application to an

- mcs,” *Weather and Forecasting*, vol. 24, no. 3, Pages 730–748, 2009. DOI: <https://doi.org/10.1175/2008WAF2222205.1>.
- [10] J. M. Kurdzo, “The wsr-88d chaff detection algorithm,” Lincoln Lab at MIT, Tech. Rep., 2019.
 - [11] J. M. Kurdzo, B. J. Bennett, D. J. Smalley, and M. F. Donovan, “The wsr-88d inanimate hydrometeor class,” *Journal of Applied Meteorology and Climatology*, vol. 59, no. 5, Pages 841–858, 2020. DOI: <https://doi.org/10.1175/JAMC-D-19-0271.1>.
 - [12] S. Gauthreaux and R. Diehl, “Discrimination of biological scatterers in polarimetric weather radar data: Opportunities and challenges,” *Remote Sensing*, vol. 12, no. 3, p. 545, 2020. DOI: <https://doi.org/10.3390/rs12030545>.
 - [13] U. F. Administration. “What is the wui?” (2022), [Online]. Available: <https://www.usfa.fema.gov/wui/what-is-the-wui.html>.
 - [14] W. Tang, C. He, L. Emmons, and J. Zhang, “Global expansion of wildland-urban interface (wui) and wui fires: Insights from a multiyear worldwide unified database (wuwui),” *Environmental Research Letters*, vol. 19, no. 4, 2024. DOI: 10.1088/1748-9326/ad31da.
 - [15] A. M. Rad *et al.*, “Human and infrastructure exposure to large wildfires in the united states,” *Nature Sustainability*, pp. 1343–1351, 2023. DOI: <https://doi.org/10.1038/s41893-023-01163-z>.
 - [16] R. Aguilera, T. Corringham, A. Gershunov, and T. Benmarhnia, “Wildfire smoke impacts respiratory health more than fine particles from other sources: Observational evidence from southern california,” *Nature Communications*, 2021. DOI: <https://doi.org/10.1038/s41467-021-21708-0>.
 - [17] J. C. Liu *et al.*, “Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties,” *Epidemiology*, vol. 28, 1 2017. DOI: 10.1097/EDE.0000000000000556.
 - [18] H. Chen, J. M. Samet, P. A. Bromberg, and H. Tong, “Cardiovascular health impacts of wildfire smoke exposure,” *Particle and Fibre Toxicology*, vol. 18, 2021. DOI: <https://doi.org/10.1186/s12989-020-00394-8>.
 - [19] R. Xu *et al.*, “Wildfires, global climate change, and human health,” *The New England Journal of Medicine*, vol. 383, no. 22, Pages 2173–2181, 2020. DOI: 10.1056/NEJMSr2028985.

- [20] P. Jatau, V. Melnikov, and T.-Y. Yu, “A machine learning approach for classifying bird and insect radar echoes with s-band polarimetric weather radar,” *Journal of Atmospheric and Oceanic Technology*, vol. 38, no. 10, Pages 1797–1812, 2021. DOI: <https://doi.org/10.1175/JTECH-D-20-0180.1>.
- [21] S. Gauthreaux and R. Diehl, “Discrimination of biological scatterers in polarimetric weather radar data: Opportunities and challenges,” *Remote Sensing*, vol. 12, no. 3, 2020. DOI: <https://doi.org/10.3390/rs12030545>.
- [22] H. Kim and B. Cheong, “Robust velocity dealiasing for weather radar based on convolutional neural networks,” *Remote Sensing*, vol. 15, no. 3, Page 802, 2023. DOI: <https://doi.org/10.3390/atmos11040363>.
- [23] J. M. Krause, “A simple algorithm to discriminate between meteorological and non-meteorological radar echoes,” *Journal of Atmospheric and Oceanic Technology*, vol. 33, no. 9, Pages 1875–1885, 2016. DOI: <https://doi.org/10.1175/JTECH-D-15-0239.1>.
- [24] K. L. Ortega, J. M. Krause, and A. V. Ryzhkov, “Polarimetric radar characteristics of melting hail. part iii: Validation of the algorithm for hail size discrimination,” *Journal of Applied Meteorology and Climatology*, vol. 55, no. 4, 2016. DOI: <https://doi.org/10.1175/JAMC-D-15-0203.1>.
- [25] A. V. Ryzhkov and D. S. Zrnić, *Radar Polarimetry for Weather Observations*. Springer, 2019.
- [26] J. Helmus and S. Collis, “The python arm radar toolkit (py-art), a library for working with weather radar data in the python programming language,” *Journal of Open Research Software*, 2016. DOI: <http://doi.org/10.5334/jors.119>.
- [27] R. J. Doviak and D. S. Zrnic, *Doppler Radar and Weather Observations*. Academic Press, 1993.
- [28] M. L. Flora, C. K. Potvin, A. McGovern, and S. Handler, “A machine learning explainability tutorial for atmospheric sciences,” *AMS Artificial Intelligence for the Earth Systems*, vol. 3, no. 1, 2024. DOI: <https://doi.org/10.1175/AIES-D-23-0018.1>.
- [29] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [30] A. Burkov, *The hundred-page machine learning book*. Andriy Burkov Quebec City, QC, Canada, 2019, vol. 1.

- [31] D. Zrnica, P. Zhang, V. Melnikov, and D. Mirkovic, "Of fire and smoke plumes, polarimetric radar characteristics," *Atmosphere*, vol. 11, no. 4, Page 363, 2020. DOI: <https://doi.org/10.3390/atmos11040363>.
- [32] V. M. Melnikov, D. S. Zrnica, and R. M. Rabin, "Polarimetric radar properties of smoke plumes: A models," *JGR Atmospheres*, vol. 114, no. D21, 2009. DOI: <https://doi.org/10.1029/2009JD012647>.
- [33] A. V. Ryzhkov and D. S. Zrnica, *Radar polarimetry for weather observations*. Springer, 2019, vol. 486.
- [34] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, Pages 90–95, 2007. DOI: 10.1109/MCSE.2007.55.
- [35] K. 5. News. "4 deaths confirmed, at least 200 injured in wildfires across oklahoma." (2025), [Online]. Available: <https://www.koco.com/article/oklahoma-fires-fire-evacuation-stillwater-leedy/64194180>. (accessed: 5.13.2025).
- [36] M. J. Reilly *et al.*, "Cascadia burning: The historic, but not historically unprecedented, 2020 wildfires in the pacific northwest, usa," *ESA Ecosphere*, vol. 13, 6 2022. DOI: <https://doi.org/10.1002/ecs2.4070>.
- [37] O. O. of Economic Analysis, *Wildfires: Impacts in the forecast*, 2020.
- [38] T. Jones, R. Ahmadov, E. J. amd Gabriel Pereira, S. Freitas, and G. Grell, "Prototype of a warn-on-forecast system for smoke (wofs-smoke)," *Weather and Forecasting*, vol. 37, 7 2022. DOI: <https://doi.org/10.1175/WAF-D-21-0143.1>.