The University of Nevada, Reno

# Examining Atmospheric and Ecological Drivers of Wildfires, Modeling Wildfire Occurrence in the Southwest United States, and Using Atmospheric Sounding Observations to Verify National Weather Service Spot Forecasts

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Atmospheric Sciences

by

Nicholas J. Nauslar

Dr. Timothy J. Brown/Dissertation Advisor

December, 2015

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

**NICHOLAS J. NAUSLAR**

Entitled

**Examining Atmospheric And Ecological Drivers Of Wildfires, Modeling Wildfire Occurrence In The Southwest United States, And Using Atmospheric Sounding Observations To Verify National Weather Service Spot Forecasts**

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Timothy J. Brown, Advisor

Michael L. Kaplan, Committee Member

John F. Mejia, Committee Member

John D. Horel, Committee Member

Peter J. Weisberg, Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

December, 2015

**Abstract**

This dissertation is comprised of three different papers that all pertain to wildland fire applications. The first paper performs a verification analysis on mixing height, transport winds, and Haines Index from National Weather Service spot forecasts across the United States. The final two papers, which are closely related, examine atmospheric and ecological drivers of wildfire for the Southwest Area (SWA) (Arizona, New Mexico, west Texas, and Oklahoma panhandle) to better equip operational fire meteorologists and managers to make informed decisions on wildfire potential in this region.

The verification analysis here utilizes NWS spot forecasts of mixing height, transport winds and Haines Index from 2009-2013 issued for a location within 50 km of an upper sounding location and valid for the day of the fire event. Mixing height was calculated from the 0000 UTC sounding via the Stull, Holzworth, and Richardson methods. Transport wind speeds were determined by averaging the wind speed through the boundary layer as determined by the three mixing height methods from the 0000 UTC sounding. Haines Index was calculated at low, mid, and high elevation based on the elevation of the sounding and spot forecast locations. Mixing height forecasts exhibited large mean absolute errors and biased towards over forecasting. Forecasts of transport wind speeds and Haines Index outperformed mixing height forecasts with smaller errors relative to their respective means.

The rainfall and lightning associated with the North American Monsoon (NAM) can vary greatly intra- and inter-annually and has a large impact on wildfire activity across the SWA by igniting or suppressing wildfires. NAM onset thresholds and subsequent dates are determined for the SWA and each Predictive Service Area (PSA),

which are sub-regions used by operational fire meteorologists to predict wildfire potential within the SWA, April through September from 1995-2013. Various wildfire activity thresholds using the number of wildfires and large wildfires identified days or time periods with increased wildfire activity for each PSA and the SWA. Self-organizing maps utilizing 500 and 700 hPa geopotential heights and precipitable water were implemented to identify atmospheric patterns contributing to the NAM onset and busy days/periods for each PSA and the SWA. Resulting SOM map types also showed the transition to, during, and from the NAM. Northward and eastward displacements of the subtropical ridge (i.e., four-corners high) over the SWA were associated with NAM onset, and a suppressed subtropical ridge and breakdown of the subtropical ridge map types over the SWA were associated with increased wildfire activity.

We implemented boosted regression trees (BRT) to model wildfire occurrence for all and large wildfires for different wildfire types (i.e., lightning, human) across the SWA by PSA. BRT models for all wildfires demonstrated relatively small mean and mean absolute errors and showed better predictability on days with wildfires. Cross-validated accuracy assessments for large wildfires demonstrated the ability to discriminate between large wildfire and non-large wildfire days across all wildfire types. Measurements describing fuel conditions (i.e., 100 and 1000-hour dead fuel moisture, energy release component) were the most important predictors when considering all wildfire types and sizes. However, a combination of fuels and atmospheric predictors (i.e., lightning, temperature) proved most predictive for large wildfire occurrence, and the number of relevant predictors increases for large wildfires indicating more conditions need to align to support large wildfires.

This dissertation is dedicated to my wife, Nathalie and my parents, Tim and Chris.

**Acknowledgements**

I would like to recognize my dissertation committee consisting of Drs. Timothy J. Brown, Michael L. Kaplan, John D. Horel, John F. Mejia, and Peter J. Weisberg. Each committee member brought his own expertise and was very accommodating and flexible throughout the entire process. I would like to personally recognize Dr. Timothy J. Brown who took a chance on me as a graduate student seven years ago and has been an excellent advisor and mentor to me affording me opportunities to learn and gain experience in and outside of the research community. I would also like to thank Dr. Michael L. Kaplan whose door is always open to any graduate student and works tirelessly to help you learn. There are several people who I would like to recognize that helped me gain experience in wildland fire. Don Scronek gave me my first seasonal job with the Bureau of Land Management and supported my efforts to forecast fire weather and gain firefighting experience (special thanks to Sam Dearstyne and Justin Conrad) and remains a close friend and mentor today. Tim Mathewson gave me the opportunity to work for Predictive Services (also thanks to Gina McGuire, John Saltenberger, and Terry Marsha) and his mentoring and training made me a better meteorologist. Finally, I would like to thank my family and friends, especially my loving wife Nathalie who keeps me centered, and my parents Tim and Chris who sacrificed much to make sure I had the tools to succeed in life.

# Chapter 1.

## Dissertation Introduction

This dissertation is comprised of three different papers that all pertain to wildfire applications. The three papers reside under the large fire weather and climate umbrella, which represents the emphasis of my dissertation in atmospheric sciences. More specifically, all three papers exemplify applied wildfire and atmospheric research with results relating directly to the operational fire weather and management communities. The first paper performs a verification analysis on mixing height, transport winds, and Haines Index from National Weather Service spot forecasts across the United States. The final two papers are closely related, which examine atmospheric and ecological drivers of wildfire for the Southwest Area (SWA) (Arizona, New Mexico, west Texas, and Oklahoma panhandle) to better equip operational fire meteorologists and fire managers to make informed decisions on wildfire potential in this region.

The subsequent three chapters describe each of the three papers submitted as the author's dissertation: 1) Verification of National Weather Service spot forecasts using atmospheric sounding observations; 2) Examining the North American Monsoon's impact on wildfire activity in the southwest United States; and 3) Using boosted regression trees to model and predict wildfires in the southwest United States. A brief summary, conclusions, and recommendations chapter follows these three chapters detailing the most pertinent findings and recommendations from the dissertation.

**Chapter 2.**

# Verification of National Weather Service Spot Forecasts Using Atmospheric Sounding Observations

Nicholas J. Nauslar and Timothy J. Brown

Desert Research Institute, Reno, Nevada

John D. Horel

University of Utah, Salt Lake City, Utah

## ABSTRACT

Fire management officials request spot forecasts from National Weather Service (NWS) Weather Forecast Offices (WFOs) to provide detailed guidance of atmospheric conditions in the vicinity of prescribed and wildland fires. Verifying spot forecasts represents an integral component of the forecast process and helps assess and improve the accuracy of forecasts. The verification analysis here utilizes NWS spot forecasts of mixing height, transport winds and Haines Index from 2009-2013 issued for a location within 50 km of an upper sounding location and valid for the day of the fire event. Mixing height was calculated from the 0000 UTC sounding via the Stull, Holzworth, and Richardson methods. Transport wind speeds were determined by averaging the wind speed through the boundary layer as determined by the three mixing height methods from the 0000 UTC sounding. Haines Index was calculated at low, mid, and high elevation based on the elevation of the sounding and spot forecast locations. Forecast statistics were calculated for each lower atmospheric variable by region including mean error and

mean absolute error. Mixing height forecasts exhibited large mean absolute errors and biased towards over forecasting. Forecasts of transport wind speeds and Haines Index outperformed mixing height forecasts with smaller errors relative to their respective means. Based on these results and the methodology, recommendations are provided to improve spot forecasts and the verification process.

## 1. Introduction

A 2007 report entitled "National Wildland Fire: A Summary of User Needs and Issues" from the Office of the Federal Coordinator for Meteorological Services and Supporting Research (OFCM) emphasized a number of improvements that are needed including "the fire community establishing accuracy requirements for fire weather products and services to enable the provider community to focus improvement efforts where most beneficial" (OFCM 2007). OFCM (2011) updated the responses to these findings, but indicated that the original findings had still not been adequately addressed. A 2008 National Oceanic and Atmospheric Administration (NOAA) report entitled, "Fire Weather Research: A Burning Agenda for NOAA," echoed similar sentiments by identifying the need of improving and conducting a more thorough forecast verification for wildland fire incidents (NOAA SAB 2008; Lammers and Horel 2014).

National Weather Service (NWS) forecasters at Weather Forecast Offices (WFOs) issue spot forecasts in response to requests from fire and emergency managers. These forecasts are typically for prescribed fire, wildfire, hazardous material, and search and rescue incidents, and provide detailed guidance for atmospheric conditions in the vicinity of these incidents. The NWS issues approximately 20,000 spot forecasts each

year for prescribed fires and wildfires, which comprise the vast majority of the spot forecasts. The NWS issues prescribed fire spot forecasts nearly twice as often as wildfire spot forecasts (Lammers and Horel 2014).

Brier and Allen (1951), Joliffe and Stephenson (2003), and Wilks (2011) identify and demonstrate appropriate verification techniques to assess forecast performance and understand sources of error to improve future forecasts. Brown and Murphy (1987) detail a fire weather forecast evaluation process that identifies biases based on forecasters' perceived consequences of under forecasting key fire weather variables and the difficulties of quantifying uncertainty in their forecasts.

Lammers and Horel (2014) examined spot forecasts from April 2009 through November 2013 and evaluated spot forecasts of surface temperature, moisture, and wind by using surface observations from the closest surface station (i.e., remote automated weather stations (RAWS)) and the National Digital Forecast Database (NDFD). Spot forecasts demonstrated higher skill than NDFD output especially for maximum temperature while the smallest improvement was associated with maximum wind speed. Our paper expands on this previous work by evaluating spot forecasts of mixing height (MH), transport winds (TWs), and Haines Index (HI).

- There were two primary objectives for the analysis: Attempt to objectively verify MH, TWs, and HI in spot forecasts
- Demonstrate and review the spot verification process, and offer relevant recommendations to improve spot forecasts

Section 2 lists the data utilized in this study, and section 3 describes the methods implemented to perform the spot forecast verification. Section 4 presents the results, and

sections 5 and 6 discuss the results and draw conclusions from the current research including noting important caveats and offering recommendations.

## 2. Data

From the period 2009-2013, 89,052 NWS spot forecasts were initially gathered for analysis (Table 1). Spot forecast requests contain the date, NWS WFO, the incident's name, latitude, longitude, elevation, the forecast parameters needed, and the option to select 'today', 'tonight', and "tomorrow" for when those forecasts should be valid (Fig. 1). The spot forecasts contain a short narrative of the weather forecast then list values for each forecast element and when they are valid (Fig. 2). For example, a spot forecast might contain three different forecasts: 'today', 'tonight', and 'tomorrow' forecasts.

Many definitions and names describe the planetary boundary layer (PBL) including boundary layer (BL), mixed layer (ML), and atmospheric boundary layer (ABL). Stull (2000) describes the BL as the shallow layer near the surface where the diurnal variation of sensible and latent heat fluxes exists between the surface and atmosphere. Wallace and Hobbs (2006) defines the BL as the layer most affected by the Earth's surface, which is separated from the rest of troposphere due to the effects of turbulence and static stability. The BL undergoes diurnal variation, but the variation deviates depending on a number of factors including: 1) season; 2) terrain; 3) synoptic conditions; and 4) land-surface type. Typically, the shallowest BL occurs just before sunrise, and as radiation flux increases the BL builds throughout the day peaking in height during the afternoon (Fig. 3; adapted from Stull 2000).

A general definition of MH is the top of the BL or ML. However, there is no universally accepted definition or criteria for its determination for two reasons: 1) various processes, such as turbulence, radiation, baroclinicity, advection, divergence, convergence, and vertical motions, contribute to the structure of the BL; and 2) most definitions or criteria are constructed on available data measured through various instruments and techniques (Beyrich 1997). Within the BL or ML the mean wind speed and direction are defined as the TWs (Miller (1967); NWS (2014); AirFire (2014). AirFire (2014) notes that some state and local agencies vary their definition of TWs including using a weighted mean for wind speed and direction through the BL. The HI quantifies dry, unstable air present in the BL, which is associated with extreme fire behavior and large fires (Haines 1988; Werth et al. 2011). Daily fire weather forecasts utilize the HI for determining fire potential especially for plume-dominated wildfires (Haines 1988; Werth and Ochoa 1993; Potter et al. 2008; Werth et al. 2011).

Atmospheric sounding data valid at 0000 UTC were selected on fire days and nearest to fire locations from 2009-2013 (University of Wyoming 2015). Variables retrieved from the sounding included unit identifier, latitude, longitude, elevation, potential temperature ($\Theta$), virtual potential temperature ($\Theta_v$), wind speed, pressure, and height. These variables were necessary to calculate the three BL elements examined for comparison to spot forecasts.

## 3. Methods

### A. Spot Forecasts

Spot forecast requests placed within 50 km of an atmospheric sounding location were organized and saved by year. Lammers and Horel (2014) chose 50 km as an appropriate distance when comparing spot forecasts and surface station observations. This verification only used the 'today' forecasts from spot forecasts containing MH, TWs, and HI. In summary, only spot forecasts issued and valid on the same day and within 50 km of an atmospheric sounding were considered.

Each spot forecast request corresponds to a spot forecast. The spot forecast request contains the date, incident name, NWS WFO, and other metadata that is used to find the corresponding spot forecast, which contains all of the forecast data that is verified. Multiple spot forecast requests and corresponding spot forecasts could be associated with the same incident. Spot forecasts are requested daily for many wildfires until an incident meteorologist arrives to handle forecast responsibilities. In addition, burn bosses for prescribed fires often request consecutive daily spot forecasts to ascertain if the weather conditions permit lighting the prescribed fire. If the spot forecasts for the same incident transpired on different days, then the spot forecasts were included in the analysis. If NWS WFOs issued more than one spot forecast on a particular day for an incident, only the most recent spot forecast was used in the analysis.

**B. Atmospheric Soundings**

NWS WFOs launch the 0000 UTC atmospheric sounding around 2300-2315 UTC. Table 2 shows the local times for launching the rawinsonde of the 0000 UTC atmospheric sounding. The -102.5° longitude line, which roughly parallels the Central and Mountain Time Zone border, is used to separate spot forecasts and their

corresponding atmospheric soundings into west and east categories. The east category was then split into an east warm season (EWS) category spanning April through October and an east cold season (ECS) category containing the remaining months. 0000 UTC occurs during the afternoon or early evening throughout the western United States indicating the nocturnal BL should not affect the MH calculation. 2300-0000 UTC occurs in the early evening for most of the central and eastern United States during the warm season (Table 2). However, during the cold season, most of the central and eastern United States are near sunset or after when 2300 or 0000 UTC passes (Table 2). By splitting into west and east categories and the east category into ECS and EWS categories, it mitigates some of the potential spot forecast verification problems with the nocturnal BL and isolates potential inapplicable comparisons.

**C. Parsing Spot Forecasts and Atmospheric Soundings**

Lammers and Horel (2014) analyzed spot forecasts as a 'natural language' problem. Spot forecast formats can vary by NWS WFO (Figs. 4-5). The table format provides time as a header with forecast values of elements requested valid at each time (Fig. 4). The non-table format provides the forecast of each requested element with a numeric value sometimes accompanied by a validation time (Fig. 5). Spot forecasts occasionally include language that narrates the progression of MH, TWs, or HI values (Fig. 6). Some of these short narrations include specific times and others use more ambiguous language including, 'then', 'becoming', 'otherwise', or 'later', creating a nebulous definition of time. Other forecasts include only one value or one range of values. The differences in the specificity of the forecast values for the requested BL

elements create challenges in pattern recognition to isolate the appropriate value or range of values in the forecast and build a database.

Text and numerical values associated with MH, TWs, or HI were extracted from spot forecasts that were issued and valid on the same day (i.e. the 'TODAY' forecast) within 50 km of an atmospheric sounding. The timing of the 0000 UTC atmospheric sounding provided guidance on selecting the appropriate forecast values from the corresponding spot forecasts. If only one numerical value or range of values existed, it was chosen as the verifiable forecast. However, since much of the data included text, measures were taken to select the latest forecast numerical value of the specified variables by interpreting key words and phrases such as 'then', 'becoming', 'late in the afternoon', and 'early in the evening'. This ensures the forecast numerical value chosen should coincide with the 0000 UTC sounding since the 'today' portion of the spot forecast usually is valid until dark (sunset). If multiple numerical values existed with any of the 'afternoon' or 'evening' key words or phrases, the lowest value and the highest value were joined to form a forecast range. This method of parsing through spot forecasts and extracting values could lead to potential errors and biases. However, the method was consistently applied to all spot forecasts with many precautions, including noting very low or high values, to protect against obtaining erroneous or incorrect data.

Atmospheric soundings and their corresponding spot forecasts were removed if the data was missing for any of the necessary variables below 200 hPa. Conversely, spot forecasts missing pertinent information, such as elevation, variables' units, or numerical forecast values, were removed along with their corresponding atmospheric soundings. Figs. 7-9 show the number of spot forecasts that exist within 50 km of an atmospheric

sounding location by year, variable, and category, and that remain before and after quality control measures. Of all available spot forecasts 6.5% occur within 50 km of an atmospheric sounding location, and 80%, 85%, and 89% of all MH, TWs, and HI spot forecasts, respectively, within the 50 km distance threshold were analyzed after the quality control measures.

**D. Calculating Mixing Height, Transport Winds, and Haines Index**

The Holzworth, Stull, and Richardson methods of determining MH were calculated for the analysis. The Holzworth method defines MH as the height where Θ exceeds the surface Θ (Holzworth 1967). A rigorous application of the Stull method explores the entire vertical atmospheric profile to identify all areas of instability using $\Theta_v$ (Stull 1988, 1991). However, for verification analysis MH was determined by finding the height where $\Theta_v$ exceeded the surface $\Theta_v$ (Fearon 2000; Fearon et al. 2015; accepted). The Richardson method involves calculating the Bulk Richardson number at each level sampled within the atmospheric sounding until reaching a certain threshold delineating turbulent and laminar flow (Richardson 1920; Stull 2000). A traditional critical threshold of the Bulk Richardson Number is 0.25, but other BL studies have shown .505, which was utilized for this verification, as an appropriate threshold (Lee et al. 2008).

A temperature of 0.5°C was added to surface Θ and $\Theta_v$ to represent surface heating (Fearon 2000). The height closest to the elevation of the spot forecast within the atmospheric sounding was identified as the new surface. More than 95% of the analyzed spot forecasts resided at or above the elevations of the sounding sites. None of the remaining spot forecast elevations were more than 350m lower than the base elevation of

the corresponding atmospheric sounding. If the spot forecast specified the MH as AGL, then the height of that level minus the surface height produced the MH. If the spot forecast specified MSL, then the height of the identified level was used as the MH. This process was performed for each MH method and produced three MHs for each atmospheric sounding to be compared to the spot forecast.

TW speeds were calculated by averaging each atmospheric sounding level's wind speed that existed at and below the MH. This calculation created three TW speeds for each sounding due to the three different methods of determining MH. Wind direction was not considered due to the imprecise and nebulous language associated with wind direction in spot forecasts. Lammers and Horel (2014) cited similar reasoning for examining wind speed and not wind direction.

HI is calculated following Haines (1988). The low elevation (less than 200 m MSL) HI combines the temperature difference between 950 hPa and 850 hPa and the dewpoint depression at 850 hPa. For mid-elevation (200-1000 m MSL) this changes to the temperature difference between 850 hPa and 700 hPa and dewpoint depression at 850 hPa. For high elevation sites (greater than 1000 m MSL), HI combines the temperature difference between 700 hPa and 500 hPa and dewpoint depression at 700 hPa (Haines 1988). These values are associated with coefficients, ranging from one to three, relating to the temperature difference and dewpoint depression (Table 3). The results from this calculation yield values from two to six with higher values representing a drier, more unstable lower atmosphere (Haines 1988).

**E. Comparing Spot Forecasts and Atmospheric Sounding Data**

The calculated variables from atmospheric soundings were directly compared to the spot forecast numerical values. If there was one forecast value, mean errors (MEs) and mean absolute errors (MAEs) were calculated for each of the applicable calculated variables. If the spot forecast issued a forecast range and an atmospheric sounding calculated variable value occurred between the lower and upper bound of that forecast range, the ME and MAE was zero. If the calculated variable did not occur within the forecast range, then it was compared to the closest value, either the lower or upper bound, with ME and MAE calculated. Consideration was given to implementing an acceptable error range for single value forecasts, such as plus or minus five percent of the forecast value, which could have alleviated some of the bias towards range forecasts.

## 4. Analysis

### A. Mixing Height

Table 4 displays the mean forecast MH, the mean calculated MH using all three methods, and the number of spot forecasts that used one MH or a range of two MHs for all three regions. The EWS had the highest mean calculated Holzworth and Stull MHs using the Holzworth and Stull methods, the highest mean forecast MHs were in the West, and the ECS had the lowest mean forecast and calculated MHs. The West usually issued forecasts with one MH instead of a MH range (62.1%) compared to the ECS (48.3%) and EWS (35.6%). The Stull method consistently had the highest mean MH, and the Richardson method yielded the lowest mean MH in all three regions. The Holzworth method mean MH was in between the mean values of the Richardson and Stull methods, but its mean was closer to the Stull method's mean. The ECS had the largest ME for Stull

and Holzworth methods, and the EWS had the largest ME using the Richardson method; all indicate over forecasting (positive ME) (Table 5). The West had the largest MAE for the Holzworth and Stull methods, but the EWS had the highest MAE for the Richardson method.

**B. Transport Wind Speed**

Table 6 displays the mean TW speed using each method, the mean TW speed forecast, and the number of forecasts that issued one TW speed and two TW speeds (i.e., 10-15 MPH). The West issues more TW speed forecasts with one wind speed (40.7%) than ECS (19.4%) and EWS (18.3%) with all of these percentages lower than their respective MH percentages. The mean TW speed forecasts for all three regions are very close (< 1 $ms^{-1}$). The West also has the highest mean wind speeds for each method with the ECS having the lowest for each method. The EWS has the lowest MAE and ME for all three methods (Table 7). The West has a higher ME than EWS and has the largest MAE for all three methods. The ECS has the highest ME for all three methods. The West, ECS, and EWS all demonstrated an over forecasting bias (positive ME) with the strongest bias associated with the ECS.

**C. Haines Index**

Table 8 shows the results for low, mid, and high elevation HI calculations. HI is requested the least among the three variables at a rate about half of MH and TWs. Most HI forecasts include one value, especially in the West, where all but one forecast had one value. The ECS has the highest mean HI, which is surprising that the cold season would

be associated with the highest HI or an unstable, dry lower atmosphere. The West has the largest MAE and ME with ME indicating under forecasting (negative ME) by nearly one category (Table 8). The ECS exhibited the least amount of bias with the lowest ME, and the EWS has the lowest MAE.

**5. Discussion**

Of all three BL variables examined, MH forecasts performed the worst, and the West, ECS, and EWS all exhibited over forecasting (positive ME) of MH. All three methods of forecast MHs exhibited MEs of at least 30% of their respective mean MHs with most of the MH MEs greater than 50%. Additionally, the MAEs of all three MH methods were at least 45% of their respective MH means with several near or larger than their means (Tables 4-5). The large magnitude of the errors was a surprise, but verifying MH forecasts is difficult for a number of reasons: 1) multiple definitions of MH; 2) multiple methods of calculating MH and no information on what approach forecasters are most likely to use; and 3) the potentially large distance and terrain variability (as much as 50 km) from atmospheric sounding locations and the spot forecast locations). Fearon et al. (2015; accepted) shows that Holzworth, Stull, and Richardson methods underestimate MH, which could also help explain the over forecasting and contribute to the large MAEs. TW speeds usually exhibited smaller errors relative to their means than MH (Table 7), which could be due to averaging the wind speeds through the calculated ML. Additionally, NWS forecasters use a range of values more often in TW speed forecasts than in MH forecasts. HI is less sensitive to the issues raised above and thus performed nearly as well or better by most measures than TWs and MH (Table 8).

The inconsistency and ambiguity of the language used in many of the spot forecasts among the NWS WFOs creates challenges for verification and leaves the accuracy of the spot forecasts vulnerable to different interpretations. The verification process treats a spot forecast with one MH and a spot forecast that describes multiple MHs during the course of a day equally. Certain methods implemented could create an accuracy bias towards forecasts that issued a range of values or multiple values valid at different times during the day.

Any forecast verification must define accuracy. The spot forecast user requires accuracy for the requested variables, but the precision of those variables remains dependent on the particular needs of the user (OFCM 2007, 2011). Penalizing a spot forecast with one value or one value valid in the afternoon by not including a certain level of acceptable inaccuracy, such as plus or minus five percent of the total forecast, may not be appropriate. However, no current standards exist for determining acceptable error in spot forecasts or specifying what the contents of each spot forecast variable should include as noted in OFCM (2007, 2011). The Joint Fire Science Program (JFSP) has recently funded projects to examine weather data in the context of decision-making, of which the research outcomes could be relevant to improve spot forecasts.

## 6. Summary and Recommendations

The results of this study demonstrate that TWs and HI spot forecasts exhibit relatively small MEs and MAEs compared to the relatively large MH spot forecasts errors. The MH result in particular underlines issues in MH forecast consistency and methods, and subsequent limitations for verification. While these results are informative

for understanding the current state of MH, TWs, and HI in spot forecasts and for potential improvements of forecasts and verification methods, they are not absolute. For example, Fearon et al. (2015; accepted) highlights the challenges of MH calculation and forecasting, and hence the verification of this type of forecast. These challenges include the spatial and temporal representativeness of the atmospheric soundings to a specific location and validation time, and using independent sources (i.e., satellite optical depth data) to generate MH. With atmospheric soundings spread across four time zones, at various latitudes, and occurring throughout the year, 0000 UTC atmospheric sounding data may not be representative of the day's MH, TWs, or HI due to the increasing influence of the nocturnal BL. Additionally, complex terrain causes significantly different BL conditions over short distances. One approach to address the representativeness of 0000 UTC atmospheric soundings would be to use vertical profiles at nearby grid points from analyses from operational numerical models (e.g., North American Mesoscale Model, Rapid Refresh (RR), or High-Resolution Rapid Refresh (HRRR) (ESRL RR 2015)). However, using such a model-based verification approach could focus on the degree to which the forecaster deviated from model output since these models are used operationally.

OFCM (2007, 2011) and NOAA SAB (2008) make similar recommendations regarding accuracy requirements, improvements, and verification statistics for both surface and upper air elements in spot forecasts. Lammers and Horel (2014) and this paper represent the beginning of a process to address these recommendations and demonstrate that spot forecast verification necessitates a more nuanced approach than just aggregating statistics. Echoing Lammers and Horel (2014) principal recommendation,

forecasters and end users should develop a framework that allows flexibility in deciding how and what to verify from spot forecasts. Additionally, the results from this spot forecast verification highlights some appropriate recommendations for any future operational attempt to evaluate spot forecasts, such as planned by the NWS Performance Branch:

1) The consistency of the information provided by the spot forecasts needs to be improved. Lammers and Horel (2014) echoes this sentiment by recommending to 'isolate quantitative numerical values separately from qualitative alphabetical descriptors.' With the understanding that different regions have different climates and user needs, the forecast values and the description of these values for each variable requested still needs standardization. Lammers and Horel (2014) discusses the importance of the qualitative information in spot forecasts for fire managers, but also recommends extracting basic forecast numerical information from the spot forecast for verification purposes.

2) A framework for verification of spot forecasts needs to be developed and implemented. Without separating numerical content or increasing the standardization of spot forecasts, any verification method implemented will encounter the inconsistencies and ambiguity in spot forecasts, which will mitigate the verification's potential positive impact. Developing a framework for spot forecast verification allows forecasters to aggregate data and evaluate spot forecasts quickly. Lammers and Horel (2014) endorses this recommendation and notes forecasters evaluating spot forecasts with local knowledge would be an improvement over 'depending on bulk statistical metrics accumulated on national scales'. More robust tools are also needed to evaluate spot

forecasts including those in areas of complex terrain and not near weather stations or sounding locations.

3) Assemble a sizable sample of focused prescribed fire and wildfire case studies to evaluate and verify forecasts. Examining the forecasts made during these prescribed burns and wildfires provide insight into possible sources of consistent errors that may lead to improving forecasts. Consistent sources of errors could include utilizing only one MH method or recognizing the limitations of forecasts in complex terrain. These sources of errors could help standardize specific methods for determining which lower atmospheric variables are more useful in different regions or during different seasons.

4) Establish accuracy thresholds or requirements for spot forecasts. This would engage the user community and provide an opportunity for NWS forecasters and users to communicate concerning spot forecast performance. Current JFSP funded work is examining aspects of accuracy concerning weather data in the context of management decision-making, which could help address this issue.

*Acknowledgements*. We would like to thank Matt Lammers for providing the NWS spot forecasts (via Virgil Middendorf, NWS Billings WFO), sharing code, and providing guidance throughout the research. We also thank Drs. Michael L. Kaplan, John F. Mejia, and Peter J. Weisberg for reviewing the manuscript before submission, and Mathew G. Fearon for his expertise in Python and mixing height determination. Funding of this project was provided by the Joint Fire Science Program under grant 12-1-05-03.

REFERENCES

AirFire, USDA Forest Service, 2015: Glossary. [Available online at http://www.airfire.org/projects/wfdss-aq/help/glossary/.]

Beyrich, F., 1997: MH estimation from sodar data — A critical discussion. *Atmospheric Environment*, **31,** 3941-3953.

Brier, G. W., and R. A. Allen, 1951: Verification of weather forecasts. Compendium of Meteorology, T. F. Malone, Ed., Amer. Meteor. Soc., 841□848.

Brown, G. W., and A. H. Murphy, 1987: Quantification of uncertianity in fire-weather forecasts: Some results of operational and experimental forecasting programs. *Wea. Forecasting*, **2**, 190-205.

Doran, J. C., J. C. Fast, and J.D. Horel, 2002: The VTMX 2000 Campaign. *Bull. Amer. Meteor*. Soc., **83**, 537-551.

Earth System Research Laboratory Rapid Refresh (ESRL RR), 2015: Rapid Refresh. [Available online at http://rapidrefresh.noaa.gov.]

Favors, J. E., and J. T. Abatzoglou, 2012: Regional Surges of Monsoonal Moisture into the Southwestern United States. *Mon. Wea. Rev.*, **141,** 182-191.

Fearon, M.G., 2000: The use of nonlocal static stability to determine mixing height from NCEP Eta model output over the western United States. Pro Quest, UMI Dissertations Publishing.

Fearon, M. G., T. J. Brown, and G. M. Curcio, 2015: Establishing a national standard for methodology for operational mixing height determination. (submitted) *J. Operational Meteor.*

Haines, D.A. 1988: A lower atmospheric severity index for wildland fires. *National Wea. Digest*, **13(2)**: 23–27.

Holzworth, G.C., 1967: Mixing depths, wind speeds and air pollution potential for selected locations in the United States. *J. Appl. Meteor.*, **6**, 1039-1044.

Joliffe, I. T., and D. B. Stephenson, 2003: Forecast Verification: A Practitioner's Guide in Atmospheric Science. Wiley & Sons, 240 pp.

Lammers, M. R., and J. D. Horel, 2014: Verification of National Weather Service spot forecasts using surface observations. *J. Operational Meteor.*, **2 (20)**, 246 264.

Lareau, N. P., E. Crosman, C. D. Whiteman, J. D. Horel, S. W. Hoch, W. O. J. Brown, and T.W. Horst, 2013: The Persistent Cold-Air Pool Study. *Bull. Amer. Meteor. Soc.*, **94**, 51–63.

Lee, P., and Coauthors, 2008: Impact of consistent boundary layer mixing approaches between NAM and CMAQ. *Environ. Fluid Mech*., **9**, 23–42.

Maddox, R. A., D. M. McCollum, and K. W. Howard, 1995: Large-Scale Patterns Associated with Severe Summertime Thunderstorms over Central Arizona. *Wea. Forecasting*, **10,** 763-778.

Miller, M. E., 1967: Forecasting afternoon mixing depths and TW speeds. *Mon. Wea. Rev.*, **95**, 35–44.

National Weather Service (NWS), 2015: Transport Wind (TW) (kts) (definitions). [Available online at http://graphical.weather.gov/definitions/defineTransWind.html.]

National Oceanic and Atmospheric Administration (NOAA) SAB, 2008: Fire weather research: a burning agenda for NOAA. National Oceanographic and Atmospheric Administration. [Available online at

www.sab.noaa.gov/Reports/2008/FWRWGreportFINA

LfromSABtoNOAA_11_03_08.pdf.]

OFCM (Office of the Federal Coordinator for Meteorological Services and Supporting Research), 2007: National Wildland Fire Weather: A Summary of User Needs and Issues. 3 July 2007. Silver Spring, Maryland.

——, 2011: Wildland Fire Weather: Multi-Agency Portfolio of Current and In-development Capabilities to Support User Needs. May 2011. Silver Spring, Maryland.

Orlanski, I. 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc,* **56** (5): 527–530.

Potter, B. E., J. A. Winkler, D. F. Wilhelm, and R. P. Shadbolt, 2008: Computing the low-elevation variant of the HI for fire weather forecasts. *Wea. Forecasting*, **23**, 159–167.

Richardson, L.F., 1920: The supply of energy and from and to atmospheric eddies. *Proc. R. Soc. Lond. A*, **97**, 686, 354-373.

Stull, R.B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, 666 pp.

——, 1991: Static stability-an update. *Bull. Amer. Meteor*., **72**, 1521-1529.

——, 2000: Meteorology for Scientists and Engineers, 2nd Edition. Pacific Grove, CA: Brooks/Cole. 502 pp.

University of Wyoming, College of Engineering, Dept. of Atmospheric Sciences, 2015: Atmospheric Soundings. http://weather.uwyo.edu/upperair/sounding.html.

Wallace, J. M., and P. V. Hobbs, 2006. Atmospheric Science: An Introductory Survey 2nd Edition. San Diego, CA: Academic Press. 483 pp.

Werth, P., and R. Ochoa, 1993: The evaluation of Idaho wildfire growth using the HI. *Wea. Forecasting*, **8**, 223–234, doi:10.1175/1520-0434(1993)008<0223:TEOIWG>2.0.CO;2.

Werth, P.A., B.E. Potter, C.B. Clements, M.A. Finney, S.L. Goodrick, M.E. Alexander, M.G. Cruz, J.M. Forthofer, S.S. McAllister. 2011. Synthesis of knowledge of extreme fire behavior: Volume 1 for fire managers. USDA Forest Service, Pacific Northwest Research Station, Portland, Oregon. General Technical Report PNW-GTR-854. 144 p

Wilks, D. S., 2011: Statistical Methods in the Atmospheric Sciences. Elsevier, 676 pp.

# TABLES AND FIGURES

**Table 1.** Total number of spot forecasts issued each year and number of spot forecasts issued within 50 km of an atmospheric sounding.

| | 2009 | 2010 | 2011 | 2012 (thru May) | 2013 (thru Aug.) | Total |
|---|---|---|---|---|---|---|
| **All** | 22,077 | 20,846 | 21,678 | 9,918 | 14,533 | 89,052 |
| **Within 50km** | 1,365 | 1,347 | 1,364 | 628 | 1063 | 5,767 |

**Table 2.** Atmospheric sounding launch time (2300 UTC) and local time for each time zone for standard and daylight time.

| | 2300 UTC Standard Time (Launch time) | 2300 UTC Daylight Savings (Launch time) |
|---|---|---|
| **Eastern** | 1900 | 1800 |
| **Central** | 1800 | 1700 |
| **Mountain** | 1700 | 1600 |
| **Pacific** | 1600 | 1500 |

**Table 3.** Temperature (T) and dewpoint (Td) differences and their reference values for all three elevations. Both reference values are added to calculate Haines Index (HI) (Haines 1988).

| | **Stability Term** | **Moisture Term** |
|---|---|---|
| **Low Elevation** | 950T-850T | 850T-850Td |
| **< 200 m** | 1: < 3°C | 1: < 5°C |
| | 2: 4-7°C | 2: 6-9°C |
| | 3: > 8°C | 3: > 10°C |
| **Mid Elevation** | 850T-700T | 850T-850Td |
| **200-1000 m** | 1: < 5°C | 1: < 5°C |
| | 2: 6-10°C | 2: 6-12°C |
| | 3: > 11°C | 3: > 13°C |
| **High Elevation** | 700T-500T | 700T-700Td |
| **> 1000 m** | 1: < 17°C | 1: < 14°C |
| | 2: 18-21°C | 2: 15-20°C |
| | 3: > 22°C | 3: > 21°C |
| **Sum both terms to calculate Haines Index** | | |

**Table 4.** Mean mixing height (MH) for spot forecasts using one forecast value or a range and each method. The number of spot forecasts utilizing one forecast value or a range is totaled for MH.

| | **Mean Forecast: 1 MH** | **Mean Forecast: MH Range** | **Mean Holzworth MH** | **Mean Stull MH** | **Mean Richardson MH** | **Number of 1 MH Forecasts** | **Number of MH Range Forecasts** |
|---|---|---|---|---|---|---|---|
| **West** | 1938 m | 1466-1822 m | 843 m | 885 m | 703 m | 666 | 407 |
| **East Cold Season** | 1287 m | 789-1144 m | 557 m | 603 m | 465 m | 331 | 355 |
| **East Warm Season** | 1578 m | 1319-1687 m | 1023 m | 1107 m | 623 m | 207 | 374 |

**Table 5.** Calculating mean error (ME) and mean absolute error (MAE) of mixing height (MH) for each method.

| | ME (Spot – Holzworth) | ME (Spot – Stull) | ME (Spot – Richardson) | MAE (Spot versus Holzworth) | MAE (Spot versus Stull) | MAE (Spot versus Richardson) |
|---|---|---|---|---|---|---|
| **West** | 370 m | 341 m | 562 m | 618 m | 618 m | 659 m |
| **East Cold Season** | 465 m | 426 m | 549 m | 553 m | 535 m | 583 m |
| **East Warm Season** | 406 m | 340 m | 752 m | 529 m | 503 m | 786 m |

**Table 6.** Mean transport wind (TW) speed for spot forecasts using one forecast value or a range and each method. The number of spot forecasts utilizing one forecast value or a range is totaled for TW speed.

| | Mean Forecast: 1 TW Speed | Mean Forecast: TW Speed Range | Mean Holzworth TW Speed | Mean Stull TW Speed | Mean Richardson TW Speed | Number of 1 TW Speed Forecasts | Number of TW Speed Range Forecasts |
|---|---|---|---|---|---|---|---|
| **West** | 7.18 $ms^{-1}$ | 5.07-7.51 $ms^{-1}$ | 6.06 $ms^{-1}$ | 6.18 $ms^{-1}$ | 5.56 $ms^{-1}$ | 424 | 618 |
| **East Cold Season** | 6.18 $ms^{-1}$ | 5.37-7.52 $ms^{-1}$ | 4.33 $ms^{-1}$ | 4.51 $ms^{-1}$ | 4.75 $ms^{-1}$ | 150 | 623 |
| **East Warm Season** | 6.22 $ms^{-1}$ | 4.95-7.09 $ms^{-1}$ | 5.48 $ms^{-1}$ | 5.64 $ms^{-1}$ | 5.20 $ms^{-1}$ | 97 | 529 |

**Table 7.** Calculating mean error (ME) and mean absolute error (MAE) of transport wind (TW) speed for each method.

| | ME (Spot – Holzworth) | ME (Spot – Stull) | ME (Spot – Richardson) | MAE (Spot versus Holzworth) | MAE (Spot versus Stull) | MAE (Spot versus Richardson) |
|---|---|---|---|---|---|---|
| **West** | 0.24 $ms^{-1}$ | 0.16 $ms^{-1}$ | 0.64 $ms^{-1}$ | 2.78 $ms^{-1}$ | 2.52 $ms^{-1}$ | 2.54 $ms^{-1}$ |
| **East Cold Season** | 1.36 $ms^{-1}$ | 1.13 $ms^{-1}$ | 0.93 $ms^{-1}$ | 2.36 $ms^{-1}$ | 2.18 $ms^{-1}$ | 1.92 $ms^{-1}$ |
| **East Warm Season** | 0.04 $ms^{-1}$ | -0.11 $ms^{-1}$ | 0.32 $ms^{-1}$ | 1.88 $ms^{-1}$ | 1.76 $ms^{-1}$ | 1.62 $ms^{-1}$ |

**Table 8.** Mean Haines Index (HI) for spot forecasts using one forecast value or a range and the total number of each type of forecast. Calculating mean error (ME) and mean absolute error (MAE) of HI.

| | ME HI | MAE HI | Mean HI | Mean of 1 HI Forecast | Mean of HI Range Forecast | Number of 1 HI Forecasts | Number of HI Range Forecasts |
|---|---|---|---|---|---|---|---|
| **West** | -0.89 | 1.07 | 4.22 | 3.33 | 3.0-4.0 | 455 | 1 |
| **East Cold Season** | -0.10 | 0.78 | 4.33 | 4.29 | 3.40-4.61 | 178 | 137 |
| **East Warm Season** | -0.20 | 0.47 | 4.27 | 4.03 | 3.20-4.54 | 196 | 105 |

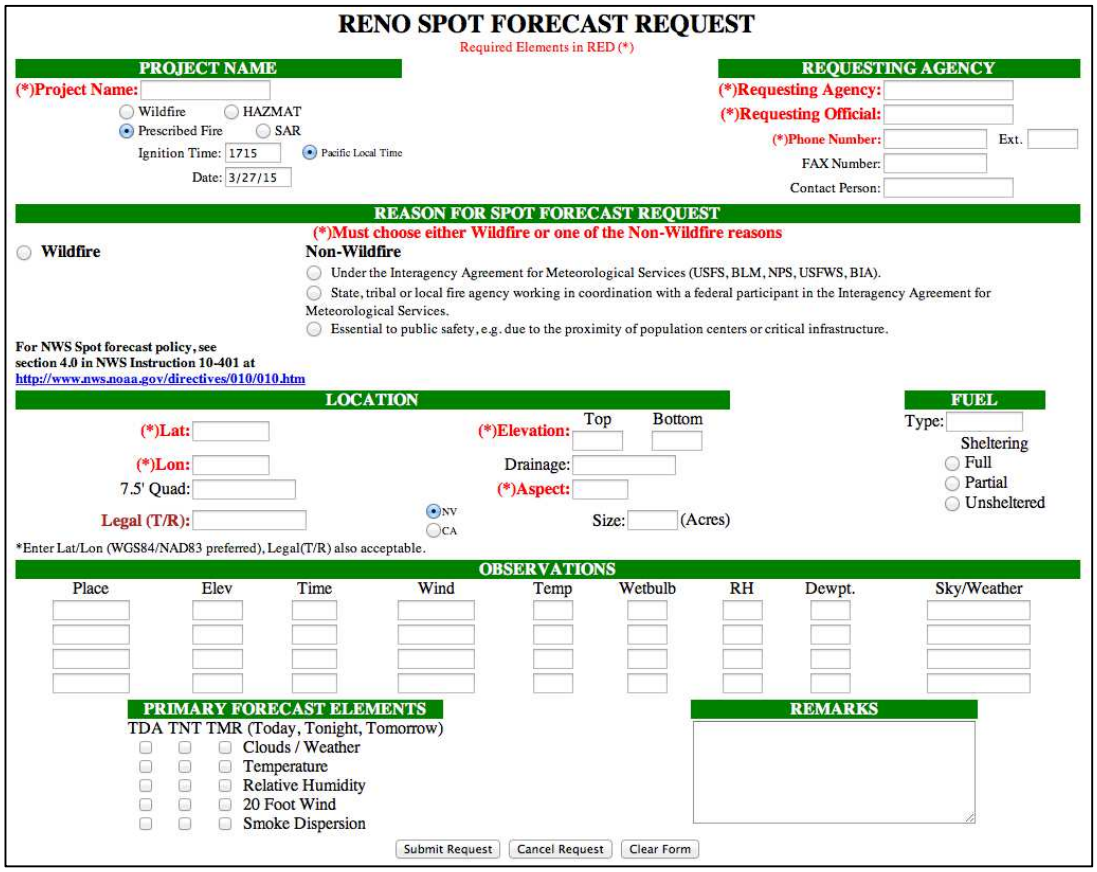


**Figure 1.** An example of the spot forecast request form via NWS WFO Reno.

```
FNUS75 KGJT 210312
FWSGJT

SPOT FORECAST FOR FRANKS BENCH...GRAND JUNCTION DISPATCH
NATIONAL WEATHER SERVICE GRAND JUNCTION CO
912 PM MDT TUE AUG 20 2013

FORECAST IS BASED ON REQUEST TIME OF 1850 MDT ON AUGUST 20.
IF CONDITIONS BECOME UNREPRESENTATIVE...CONTACT THE NATIONAL WEATHER
SERVICE.

.DISCUSSION...A PLUME OF SUB TROPICAL MOISTURE SITTING OVER THE
REGION WILL RESULT IN ANOTHER ROUND OF SHOWERS AND THUNDERSTORMS
WEDNESDAY AFTERNOON AND EVENING. STORMS WILL DIMINISH LATE IN THE
EVENING...HOWEVER A FEW SHOWERS WILL CONTINUE TO BE POSSIBLE
OVERNIGHT. STORMS WILL BE SLOW MOVING WITH HEAVY RAIN POSSIBLE IN
THE AFTERNOON AND EVENING.

.WEDNESDAY (UNTIL SUNSET)...

SKY/WEATHER.........PARTLY CLOUDY (50-60%)...BECOMING MOSTLY
                    CLOUDY (70-80%) IN THE AFTERNOON. AN ISOLATED
                    SHOWER OR THUNDERSTORM POSSIBLE IN THE MORNING...
                    THEN SCATTERED SHOWERS AND THUNDERSTORMS
                    DEVELOPING IN THE AFTERNOON.
MAX TEMPERATURE.....80-85 F.
MIN HUMIDITY........28-33%.
20-FOOT WINDS.......DOWNSLOPE/DOWNVALLEY  3-6 MPH UNTIL 1000...THEN
                    LIGHT AND VARIABLE. BECOMING UPSLOPE/UPVALLEY
                    4-8 MPH AFTER 1200. GUSTY AND ERRATIC WINDS
                    EXPECTED NEAR THUNDERSTORMS. BY LATE AFTERNOON
                    THUNDERSTORM OUTFLOW WINDS MAY DOMINATE
                    SURFACE AND 20-FOOT WINDS WITH GUSTS TO
                    35 MPH.
SMOKE DISPERSAL.....POOR UNTIL 1100...THEN BECOMING GOOD IN THE
                    AFTERNOON.
LAL.................2 UNTIL 1200...THEN 4.
MIXING HEIGHT.......INCREASING TO 11000 FT AGL IN THE AFTERNOON.
TRANSPORT WINDS.....WEST 5 MPH. VARIABLE THUNDERSTORM OUTFLOWS MAY
                    DOMINATE TRANSPORT WINDS AT TIMES.
```

**Figure 2.** Example of a spot forecast via NWS WFO Grand Junction.

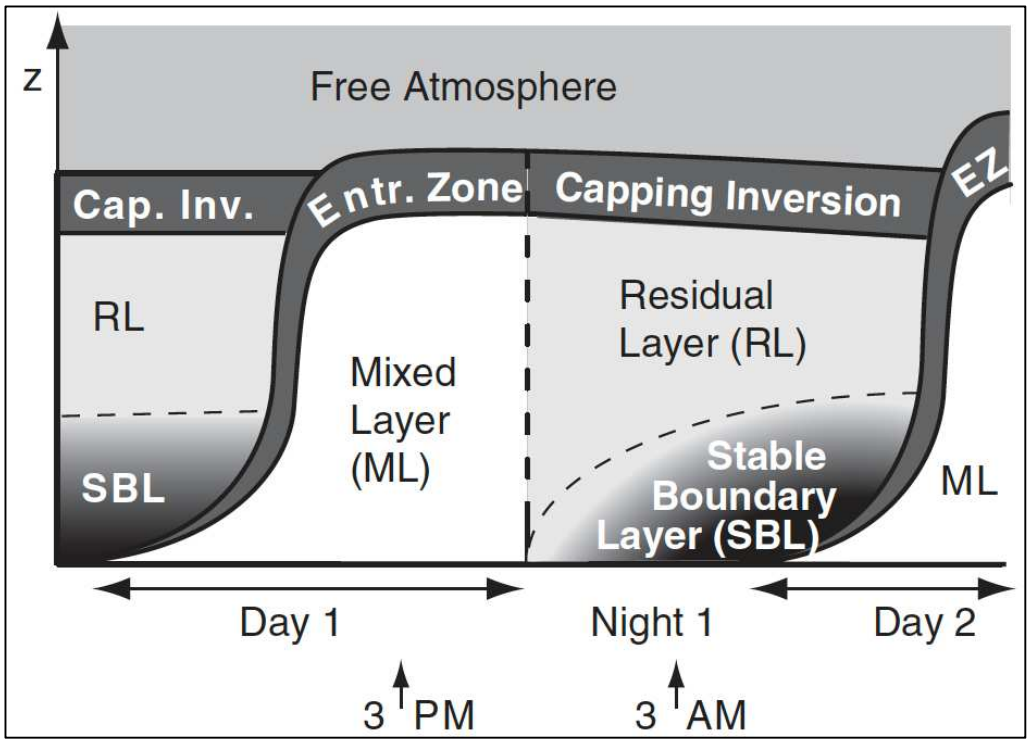


**Figure 3**. Diurnal boundary layer transition vertical profile from Stull (2000).

| .TODAY... | | | | | | |
|---|---|---|---|---|---|---|
| TIME (EDT) | 6 AM | 8 AM | 10 AM | NOON | 2 PM | 4 PM |
| SKY..... | MCLDY | MCLDY | MCLDY | MCLDY | PCLDY | PCLDY |
| WEATHER TYPE.... | RNSHWR | RNSHWR | RNSHWR | RNSHWR | RNSHWR | RNSHWR |
| CHC PRECIP (%).. | 20 | 20 | 20 | 20 | 20 | 20 |
| TEMP (F)..... | 84 | 84 | 86 | 87 | 88 | 87 |
| RH (%)..... | 74 | 73 | 69 | 67 | 66 | 67 |
| 20 FT WIND..... | S 5 | S 6 | S 6 | S 5 | S 4 | S 4 |
| 20 FT WIND GUST. | 10 | 10 | 10 | 10 | 5 | 5 |
| EYE LEVEL WIND.. | S 5 | S 6 | S 6 | S 5 | S 4 | S 4 |
| EYE LVL WND GST. | 10 | 10 | 10 | 10 | 5 | 5 |
| MIX HGT (FT).... | 700 | 1000 | 2700 | 4500 | 4500 | 4500 |
| TRANSPORT WIND.. | S 14 | S 15 | S 15 | S 14 | S 14 | S 14 |
| DISPERSION..... | 12 | 14 | 18 | 47 | 44 | 44 |
| LAL..... | 2 | 2 | 2 | 2 | 2 | 2 |

**Figure 4.** Example of a spot forecast in a table via NWS WFO Key West.

```
.TODAY...
SKY/WEATHER.........PARTLY CLOUDY(30-40%) UNTIL 1000...THEN MOSTLY
                    SUNNY UNTIL 1500...THEN PARTLY CLOUDY(50-60%).
MAX TEMPERATURE.....67-71.
MIN HUMIDITY........16-23%.
20-FOOT WINDS.......WEST WINDS 5-13 MPH UNTIL 0900...THEN SOUTH
                    AROUND 6 MPH UNTIL 1200...THEN SOUTHEAST
                    5-10 MPH.
TRANSPORT WINDS.....SOUTHWEST 8-14 MPH.
MIXING HEIGHT.......BELOW 1000 FT AGL UNTIL 1000. 6400-7400 FT AGL
                    AFTER 1200.
SMOKE DISPERSAL.....POOR UNTIL 1000...THEN FAIR UNTIL 1200...THEN
                    GOOD UNTIL 1700...THEN FAIR.
```

**Figure 5.** Example of a spot forecast using a mix of text and numerical values via NWS WFO Albuquerque.

```
.TODAY...

SKY/WEATHER.........MOSTLY SUNNY (20-30 PERCENT).
                    TEMPERATURE.........42 AT IGNITION...MAX 60.
                    MIN HUMIDITY........65 PERCENT AT IGNITION...MIN 25 PERCENT.
WIND (20 FT)........WINDS NORTH AT 1 MPH AT IGNITION...OTHERWISE
                    WEST WINDS 5 TO 6 MPH EARLY IN THE MORNING BECOMING
                    LIGHT.
CHANCE OF PCPN......0 PERCENT.
LAL.................1.
HAINES INDEX........3 OR VERY LOW POTENTIAL FOR LARGE PLUME DOMINATED
                    FIRE GROWTH.
MIXING HEIGHT.......2500 FT AGL AT IGNITION...INCREASING TO 2500-3500
                    FT AGL LATE IN THE MORNING...THEN INCREASING TO
                    4500-5500 FT AGL EARLY IN THE AFTERNOON INCREASING
                    TO 5500-6500 FT AGL LATE IN THE AFTERNOON DECREASING
                    TO 4000 FT AGL BY 1800 HOURS CST.
MIXING WINDS........WEST 5 TO 6 MPH DECREASING TO 1 TO 3 MPH LATE IN THE
                    MORNING.
```

**Figure 6.** Example of a spot forecast with a more descriptive narrative using multiple forecasts valid at different times of the day via NWS WFO Amarillo.

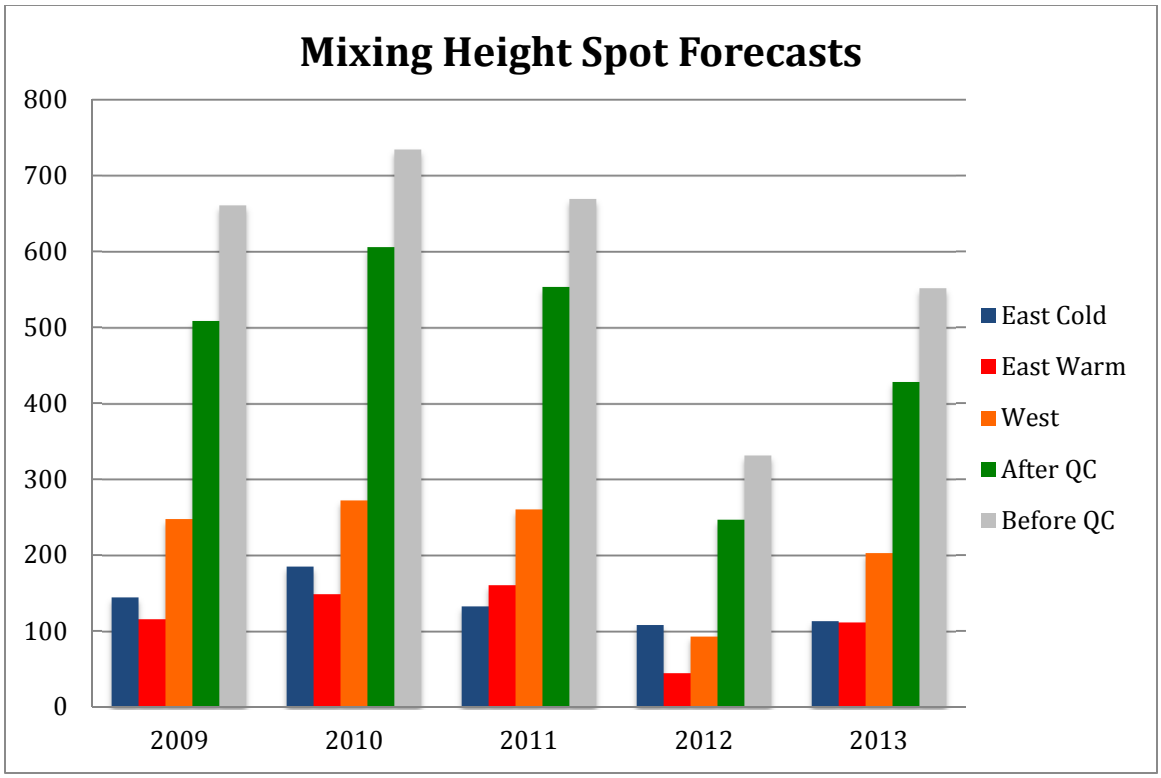


**Figure 7.** Number of spot forecasts with MH forecasts evaluated by year and category. Also shown is the number of spot forecasts before and after quality control measures.

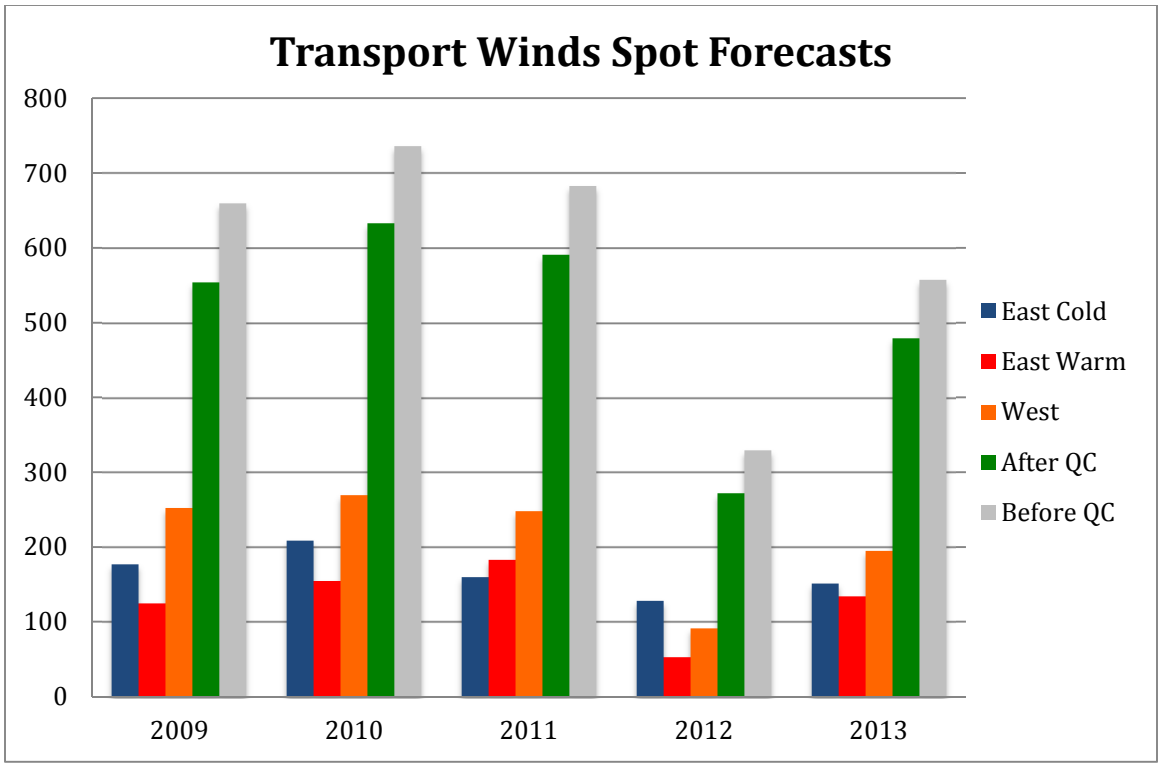


**Figure 8.** Number of spot forecasts with TWs forecasts evaluated by year and category. Also shown is the number of spot forecasts before and after quality control measures.

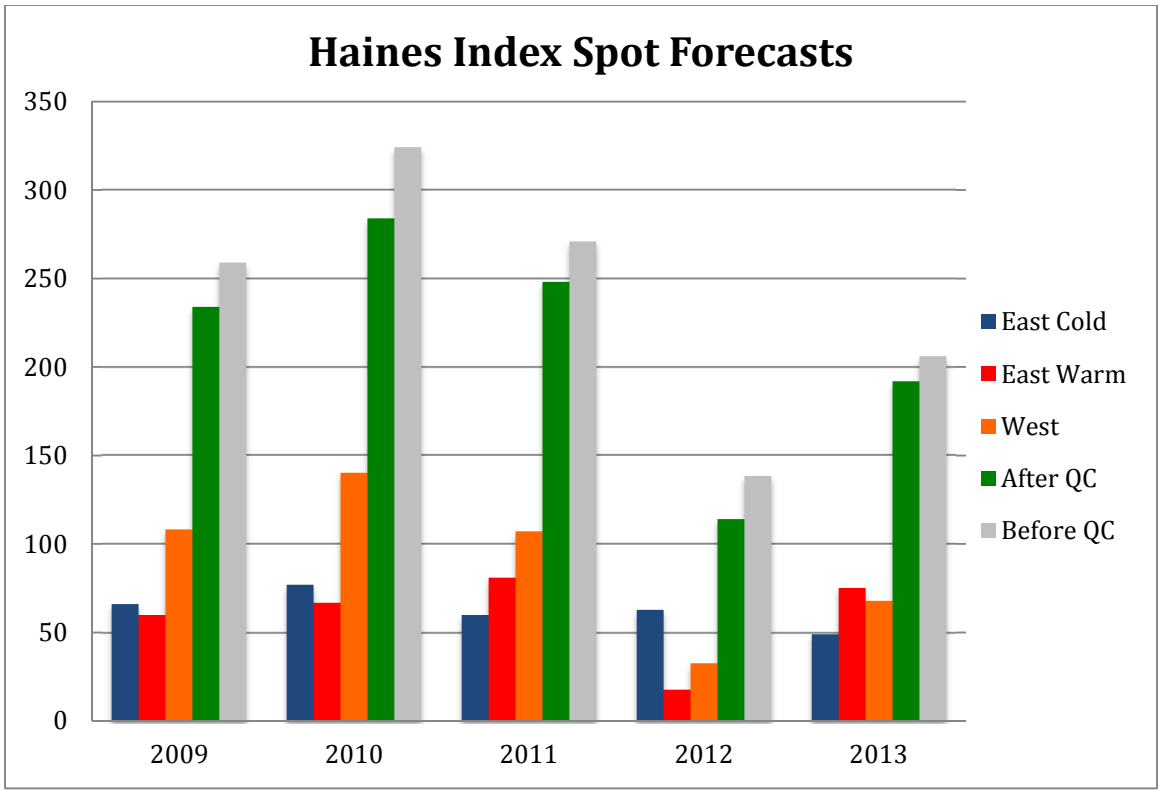


**Figure 9.** Number of spot forecasts with HI forecasts evaluated by year and category. Also shown is the number of spot forecasts before and after quality control measures.

**Chapter 3.**

# Examining the North American Monsoon's impact on wildfire activity in the southwest United States

Nicholas J. Nauslar[12], Timothy J. Brown[1], Michael L. Kaplan[1], and John F. Mejia[1]

[1] Division of Atmospheric Sciences, Desert Research Institute, Reno, NV

[2]Cooperative Institute for Mesoscale Meteorological Studies/NOAA NWS Storm Prediction Center, Norman, OK

*Corresponding author address: Nicholas J. Nauslar, 120 David L. Boren Blvd., Suite 2100*

*Norman, OK 73072-7304*

*Email: nauslar@dri.edu*

Keywords: fire weather, self-organizing maps, dry thunderstorms, critical fire weather patterns, synoptic classification, monsoon onset, fire management

**Abstract**

The North American Monsoon (NAM) is an annual climate system phenomenon that develops over the Sierra Madre Occidental in western Mexico and spreads northward into the southwest United States (Arizona and New Mexico) from June through September bringing large quantities of rainfall and lightning. The rainfall and lightning associated with the NAM can vary greatly intra- and inter-annually and has a large impact on wildfire activity in the southwest United States by igniting or suppressing wildfires. The timing of the NAM onset can lengthen or shorten the wildfire season in the southwest United States. NAM onset thresholds and subsequent dates are determined for the Southwest Area (SWA)(Arizona, New Mexico, west Texas, and Oklahoma panhandle) and each Predictive Service Area (PSA) within the SWA April through September from 1995-2013. Various 'busy' wildfire activity thresholds based on the number of wildfires especially large wildfires identified days or time periods with increased wildfire activity for each PSA and the SWA. Self-organizing maps utilizing 500 and 700 hPa geopotential heights and precipitable water were implemented to identify atmospheric patterns contributing to the NAM onset and busy days/periods for each PSA and the SWA. Resulting SOM map types also showed the transition to, during, and from the NAM. Northward and eastward displacements of the subtropical ridge (i.e., four-corners high) over the SWA were associated with NAM onset, and a suppressed subtropical ridge and breakdown of the subtropical ridge map types over the SWA were associated with increased wildfire activity.

## 1. Introduction

The North America Monsoon (NAM) affects much of North America with the largest impact in Mexico and the arid regions of the southwest United States (SWUS)(Figure 1) and is characterized as a large-scale atmospheric circulation that produces a distinct increase of warm-season (June through September) precipitation over North America (Adams and Comrie 1997; Grantz et al. 2007). The NAM core region is centered over the Sierra Madre Occidental (SMO) in western Mexico (Douglas et al. 1993; Barlow et al. 1998), but it stretches into the SWUS and often further north into the western United States (WUS) with its effects felt hemisphere wide (Reiter and Tang 1984; Tang and Reiter 1984; Douglas et al. 1993; Higgins et al. 1999; Lo and Clark 2002; Hawkins et al. 2002). The variability of NAM in the SWUS and WUS is large, sometimes larger than the mean warm season rainfall, and is modulated by intraseasonal transient features including, tropical easterly waves (TEWs), tropical storms/cyclones, and transient/tropical upper tropospheric troughs (TUTTs) (Higgins et al. 1998; Adams and Stensrud 2007; Abatzoglou and Brown 2009; Hendon et al. 2011; Means 2012; Favors and Abatzoglou 2012; Mejia et al. 2015; Seastrand et al. 2015).

Higgins et al. (1997) describes the NAM in three phases: development, mature, and decay. The development of the NAM is characterized by a transition of generally westerly flow associated with a cold season circulation shifting to a southerly and easterly flow as the subtropical ridge develops over Mexico and the SWUS (Adams and Comrie 1997; Higgins et al. 1997; Grantz et al. 2007). The shift in winds is in response to the surface heating over the continent, especially over the elevated terrain in the NAM region. The land-ocean temperature gradient increases, and the air flows towards the

continent increasing convergence and enhancing the mountain-valley circulation over the SMO (Douglas et al. 1993; Vera et al. 2006; Grantz et al. 2007; Gochis and Higgins 2007). With the shift in winds and development of the subtropical ridge, a noticeable precipitation dipole emerges between the Plains (central U.S.) and NAM region, in particular over the SWUS and northwest Mexico (NWMEX) (Higgins et al. 1997; Higgins et al. 1999; Mitchell et al. 2002; Adams and Stensrud 2007; Stensrud 2013). A northward/eastward (southward/westward) displacement of the subtropical ridge coincides with anomalous wet (dry) NAM conditions for NWMEX and SWUS (Carleton et al. 1990; Higgins et al. 1999; Cerezo-Mota et al. 2011).

When the diurnal cycle of rainfall begins increasing, the NAM onset is identified across southern Mexico, which extends northward rapidly to the SMO and into the SWUS (Douglas et al. 1993; Stensrud et al. 1995; Higgins et al. 1997). The diurnal pattern of precipitation forms earliest and most frequently at the highest elevations over the SMO early in the afternoon. Over lower elevations, the precipitation is less frequent but of higher intensity (Gebremichael et al. 2007; Gochis and Higgins 2007; Nesbitt et al. 2008). The heaviest precipitation occurs in July and August during the mature stage the NAM and then waning in September. Some areas in the NAM core region receive 50-80% of their annual rainfall during the NAM (Carleton et al. 1990; Douglas et al. 1993; Higgins et al. 1997). The decay phase of the NAM is more gradual than the onset and represents the return to a cold season circulation. This transition is noted by a return of the westerlies from a diurnally driven, warm season circulation with easterly and southerly flow over the NAM region (Higgins et al. 1997; Barlow et al. 1998; Vera et al. 2006; Cerezo-Mota et al. 2011; Means 2012).

Wildfire occurrence increases steadily through April in the SWUS, but the trend accelerates in May and June before decreasing in July and through September (Brandt 2006). The peak SWUS wildfire season occurs in June and early July due to decreasing fuel moisture, increasing temperatures, and enhanced ignition potential coinciding with drier thunderstorms (Watson et al. 1994; Westerling et al. 2003). The number of human caused wildfires decreases significantly after onset (less than 25% across fuel types) possibly due to diminished dry and windy conditions seen in the transition to the NAM and less people being active during the NAM due to rain and lightning concerns (Mohrle 2003). 57% of natural wildfires occur after onset, but only 37% of acres burned (Mohrle 2003). The NAM increases cloud-to-ground lightning, which increases wildfire ignitions, but with the corresponding increased atmospheric and fuel moisture, most wildfires remain small (Evett et al 2008; Dowdy and Mills 2012). 55% of large wildfires (40 hectares) and nearly 50% of area burned from large wildfires occur after onset (Mohrle 2003).

Dry thunderstorms contribute to large wildfire outbreaks across the WUS, but remain without a precise definition due to the variability across climates and ecosystems that would be considered ‘dry’ (Wallmann et al. 2010; Nauslar et al. 2013). Dry thunderstorms consist of higher cloud bases and limited precipitation, and form on the edges of the deepest NAM moisture and strongest dynamics associated with an upper tropospheric trough (Wallmann et al. 2010; Nauslar et al. 2013). Hall (2007) demonstrated wildfire occurrence abruptly decreases approximately midway through the climatological peak of lightning and precipitation in Arizona and New Mexico, which is

coincident with the NAM, and precipitation thresholds associated with wildfire ignitions were positively correlated with elevation.

While an explicit relationship exists between weather and wildfires, this relationship can be complex and nonlinear. Synoptic composites are often used to examine the relationships between the dominant atmospheric circulation systems and ecological factors (Hewitson and Crane 2002). Rodriguez-Iturbe et al. (1998) showed a highly interactive nonlinear soil-atmosphere system that displayed self-organizing features (Cavazos et al. 2002). Several studies have implemented self-organizing maps (SOMs)(Kohonen 2001) to show atmospheric circulation patterns associated with different phenomena (Cavazos et al. 2002; Hewitson and Crane 2002; Crimmins 2006; Reusch et al. 2007; Reusch 2010). Using SOMs, Cavazos et al. (2002) showed that the NAM in southeast Arizona was dominated by three wet modes, and Crimmins (2006) identified different atmospheric patterns associated with extreme wildfire weather conditions utilizing SOMs. SOMs are utilized to identify atmospheric circulation patterns due to their propensity for visualizing complex distribution of synoptic states (Kohonen 2001; Hewitson and Crane 2002). Hence, we argue that SOMs can help examine the intricate nonlinear relationships between the NAM main atmospheric circulation and critical fire weather patterns leading to significant wildfire episodes. By identifying atmospheric patterns pertinent to wildfire activity in the Southwest Area (SWA)(Arizona, New Mexico, west Texas, and Oklahoma panhandle)(Figure 2), we can aid operational fire weather forecasters in better prediction of wildfire activity specifically during times of increased wildfire activity, which dramatically impact local, regional, and national wildfire suppression resources.

Objectives for this analysis include: 1) determining the climatological atmospheric patterns via SOMs and wildfire occurrence characteristics in the SWA during the transition to, during, and transition from the NAM; 2) defining and determining the NAM onset within the SWA (Figures 1-2); and 3) use SOMs to identify atmospheric patterns that correspond to increased wildfire activity.

**2. Data**

Data were collected for April through September from 1992 to 2013 unless otherwise specified for the SWA. Figure 2 shows the different predictive service areas (PSAs) that comprise the Southwest Geographical Area Coordination Center (SWCC) region of responsibility (SWA) as determined by the Southwest Predictive Services (PS). Southwest PS investigates and provides forecast wildfire potential for sub regions (i.e., PSAs) with similar historical wildfire occurrences, fuel types, fuel indices, and weather characteristics.

Surface gridded data from the University of Idaho METDATA (Abatzoglou 2013) were implemented. METDATA consists of daily, 4 km grid size precipitation, minimum and maximum RH (RHmin; RHmax), specific humidity (SPH), minimum and maximum temperature (Tmin; Tmax), and wind velocity. Vapor pressure deficit (VPD) was defined as the difference between saturated vapor pressure and ambient vapor pressure. Saturated vapor pressure and ambient vapor pressure were calculated using Tmin, Tmax, RHmin, and RHmax. METDATA also provides calculated fire danger indices of burning index (BI), energy release component (ERC), 100-hour, and 1000-hour dead fuel moisture

(FM100; FM1000) using fuel model G (Deeming et al. 1978; Cohen and Deeming 1985; Abatzoglou 2013). All of the data were averaged for each PSA across the SWA.

Upper air observations were obtained from the National Center for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR; Mesinger et al. 2006) and downloaded from the Earth System Research Library (ESRL) Physical Sciences Division (PSD). Daily winds, geopotential heights, temperature, and SPH were acquired at various levels from 850 to 200 hPa. Additionally, daily values of NARR precipitable water (PWat) and integrated water vapor flux (IWVF) are utilized. Studies have shown the NARR dataset has limitations and bias results in the NAM region for moisture variables and near surface variables (Dominguez et al. 2008; Becker et al. 2009; Cerezo-Mota et al. 2011). However, the NARR offers a higher resolution dataset than the NCEP/NCAR reanalysis (Kalnay et al. 1996) that spans the study's time period, and numerous studies have shown NARR representative when examining atmospheric circulation and moisture across the study region (i.e., Dominguez and Kumar 2008; Dominguez et al. 2008; Turrent and Cavazos 2012; Doubler et al. 2015; Radhakrishna et al. 2015). All daily NARR data were organized into a 175x120 grid with full coverage spanning from 85°W to 131°W longitude and 18°N to 49.5°N latitude (Figure 1; partial coverage).

Wildfire data were obtained from the Fire Program Analysis (FPA) quality controlled wildfire database (Short 2015). Wildfires were separated by type (all, lightning, human) and size (all, large) for each PSA. Large wildfire selection criteria were established from large wildfire thresholds determined by Southwest PS (Table 1). Also included in the analysis was the proprietary National Lightning Detection Network®

(NLDN) data. Due to a major upgrade occurring after 1994, in this study only NLDN data from 1995-2013 were used.

The Madden-Julian Oscillation (MJO) (Madden and Julian 1971, 1972) and Multivariate ENSO Index (MEI) (Wolter and Timlin 2011) were used to examine possible teleconnections. MJO real-time multivariate MJO series 1 (RMM1) and 2 (RMM2) and MJO phase are used to represent MJO strength and location (Wheeler and Hendon 2004; CPC 2014). MEI includes sea-level pressure, *u* and *v* surface wind components, SST, surface air temperature, and total cloudiness fraction (ESRL-PSD 2014). MEI is generated via cluster analysis and calculated from the first unrotated principal component of the six parameters (Wolter 1987; Wolter and Timlin 1993).

## 3. Methods

### 3.1 The North American Monsoon Onset

The NAM onset is characterized by an increase of atmospheric moisture and precipitation associated with the establishment of the subtropical ridge over western Mexico and the SWUS (Adams and Comrie 1997; Higgins et al. 1997; Grantz et al. 2007). The Tucson National Weather Service Weather Forecast Office (NWS WFO) has utilized a dewpoint temperature of 54°F at the Tucson airport as its NAM onset threshold (NWS WFO Tucson 2015). However, a universally accepted NAM onset definition does not exist for the SWUS.

Without a definitive NAM onset threshold, several atmospheric and fuels variables were examined at various time scales to capture the onset signal for each PSA. Some of the onset techniques included analyzing their change over a time period and

attempting to fit polynomial functions to variable fluctuations. The predictors selected for determining onset were the daily averaged PSA value of SPH, VPD, lightning strikes (LS), and precipitation. These variables were chosen due to their predictive potential and specific representation of atmospheric moisture conditions. The primary technique for identifying onset used six consecutive days of greater than or equal to the 22-year median PSA value of SPH and VPD and six consecutive days of at least 100 lightning strikes occurring at the same time (Onset1). Other combinations of variables were evaluated that substituted precipitation for VPD (Onset2) or SPH (Onset3). Six consecutive days was chosen after examining time periods of five to ten consecutive days, all of which are within the synoptic timescale. The NAM onset date for each PSA was determined by taking the first instance that met the Onset1, Onset2, or Onset3 requirements after 1 June. We constrained the onset after 1 June to discriminate the influences of spring storms. For each PSA, the $10^{th}$ percentile, $90^{th}$ percentile, median, inter quartile range (IQR), and +/- 1 median absolute deviation (MAD) onset dates were calculated.

### 3.2 Increased Wildfire Activity

The FPA wildfire occurrence data was organized by PSA and separated into categories based on size and type. If a particular day exceeded the $95^{th}$ ($99^{th}$) percentile for number of wildfires and large wildfires or at least 2 (3) large wildfires occurred for a particular PSA, it was considered a 'busy day' and exceeded the Busy95_PSA (Busy99_PSA) threshold. If a particular day exceeded the $95^{th}$ ($99^{th}$) percentile for number of large wildfires for the SWA, it was considered a 'busy day' and exceeded the Busy95_SW (Busy99_SW) threshold. Often, increases of wildfire activity occur over

multiple days, which can stress local and regional wildfire suppression resources. Time periods of two, three, seven, and ten days that exceeded the 95$^{th}$, 99$^{th}$, and 99.5$^{th}$ percentiles for large wildfires for the entire SWUS over those respective time periods were identified as 'busy events'. The busy days and events were examined primarily to determine: 1) how wildfire activity varied across PSAs; and 2) to provide dates to examine atmospheric and fuel conditions during these days/events.

### 3.3 Self-Organizing Maps

SOMs are an objective analysis method based on neural networks that extract generalized patterns from grid data defined over a region. SOMs capture the most significant atmospheric circulation patterns, and since the analysis is performed over daily time increments, then such patterns include those modes related to synoptic scale variations (Cavazos et al. 2002; Reusch et al. 2007). The SOMs divide a continuum of atmospheric patterns into a small number of categories, which are spatially organized by similarity with similar patterns closer to each other and dissimilar patterns further away from each other on the user-defined grid (Hewitson and Crane 2002; Reusch 2010). By identifying atmospheric circulation patterns, SOMs allows a frequency analysis of each pattern, preferred transitions, and the ability to associate patterns with events (Cavazos et al. 2002; Reusch et al. 2007). SOMs are able to process multivariate, multidimensional input on the same grid by creating a spatially organized set of generalized patterns of variability from the input data (Reusch et al. 2007; Reusch 2010). Kohonen (2001), Hewitson and Crane (2002), Reusch et al. (2007), and Reusch (2010) provide further

detailed explanations of SOMs and the implementation of SOMs for meteorological and climatological applications.

SOMs provide an alternative (Reusch et al. 2005) to principal component analysis (PCA), canonical correlation analysis (CCA), and clustering (k-means and hierarchical) that Reusch et al. (2007) argues is more robust, accurate, and able to provide better visualization of structure in large, nonlinear datasets. One of the main advantages of implementing SOMs over other multivariate analysis approaches is the ability of SOMs to accommodate nonlinear relationships in the data. For example, Cavazos et al. (2000) utilized SOMs to research circulation and humidity patterns associated with extreme precipitation events in the Balkans, and the results were consistent with PCA and CCA while offering new results.

The input data consisted of standardized values of NARR 500 hPa geopotential heights, 700 hPa geopotential heights, and PWat on a 175x120 grid (85°W-131°W; 18°N-49.5°N) for 183 days (1 April – 30 September) over 22 years (1992-2013). SOMs require standardization of data for multiple variables having varying scales (Reusch et al 2007; Reusch 2010). The full input data was included for training since the primary goal of SOMs is to use generalized patterns determined from the full dataset as recommended by Reusch (2010). Grid sizes of 4x3 (rows by columns), 5x3, and 4x5 were examined at 500, 1000, 2000, and 4000 iterations or until a stable solution was reached. Examining Sammon maps of multidimensional data on the different SOMs grids and matching the input data to the final reference nodes of these grids help determine if a solution is acceptable (Sammon 1969; Reusch et al. 2007). Plots of neighbor node distances and the number of sample hits for each node are also helpful in determining if a SOMs solution is

acceptable. After a stable solution was reached, each day was classified with a node as a best match, which was composited into a map type (Reusch 2010). After each day was classified, a frequency analysis was applied to all days, busy days/events, and onset days where applicable to determine if individual map types (MTs) or MT transitions were more associated with certain types of events.

## 4. Results

### 4.1 Climatological Wildfire Occurrence and Map Patterns

Table 2 shows the average number of wildfires per year by type and size for the SWUS. Annually, 51.5% of wildfires are human caused, but 57.2% of all large wildfires are lightning caused (Table 2). Peak fuel dryness and wildfire activity (number of wildfires and number of large wildfires) coincide from mid June into early July depending on location for the SWA (Table 3). Similar to Brandt (2006), the number of wildfires was found to increase into July before decreasing late in July through September. July has nearly twice as many lightning wildfires as June, but June has 100 more large lightning wildfires than July (Table 3). This confirms results from Mohrle (2003) that showed 57% of natural wildfires but only 37% of total area burned from natural wildfires occurred after onset. Table 3 also confirms Mohrle (2003) results, which showed that human-related wildfires tend to drop after NAM onset specifically among the PSAs residing under the strongest NAM influence.

The 4x5 grid was chosen after examining the Sammon maps, neighbor node distances, and composited MTs for all three examined grid configurations. A stable solution was reached after 2000 iterations on the 4x5 SOMs grid, but the 4000 iteration

SOMs analysis was utilized due to slightly better coverage from the Sammon map depiction (not shown). The 5x3 SOMs grid MTs represented most of the MTs from the 4x5 grid, but two map patterns from the 4x5 SOMs grid that corresponded strongly to busy and onset days were not resolved on the 5x3 grid. Figures 3-5 display the resulting twenty composite MTs from the SOMs analysis. Map Type 1 (MT {1}, i.e. Figure 3 {1}) for 500 hPa geopotential heights directly corresponds to MT {1} (i.e., Figures 4-5 {1}) for both 700 hPa geopotential heights and PWat since all three variables were considered during the SOMs analysis. The SOMs nodes provide a general progression of increasing geopotential heights and PWat (Figure 3-5) {1-20}.

Ascending numbered MTs demonstrate a transition from a four-corners trough to a four corners ridge from April to July and then a transition from a four-corners ridge back to a four-corners trough begins in September descending back through the MTs (Figure 3)(Table 5). April is dominated by MTs associated with troughs approaching or passing over the SWUS, and stronger height gradients with southwest or westerly flow aloft and drier conditions are associated with these MTs (Figure 3){1-6}. The most prevalent MTs in May exhibit a building subtropical ridge from the south and east over Mexico and the Gulf of Mexico (Figure 3 {7-10}). Drier conditions remain over most of the SWA, but increases of PWat develop on the western and eastern coasts of Mexico (Figure 4 {7-10}). Geopotential heights continue to increase and the fluctuation between troughing and ridging maximizes during June (Figure 3 {8, 11-13}). The subtropical ridge is still relatively weak in June over the SWUS with the ridge center located over western or northwestern Mexico, but moisture continues to increase in extent and magnitude across Mexico and the SWUS (Figure 3-5 {11-13}). By July, the transition to

a ridge dominated atmospheric pattern is complete, and continues in place through August (Figure 3 {15-16, 18-20}). Easterlies shift northward bringing ample moisture across much of Mexico and into the SWUS (Figures 3-5 {15-16, 18-20}). In September, the center of ridge begins to migrate south with ridging and ridge breakdown patterns prevailing (Figure 3 {12-13, 17}).

The three most prevalent MTs (Figure 3 {i.e., 16, 19, 20}) show the SWUS under the influence of the subtropical ridge, which is expected during the NAM (Adams and Comrie 1997; Higgins et al. 1997). Trough thinning and Rossby wave breaking (RWB) (McIntyre and Palmer 1985) are evident for multiple MTs at 700 hPa, in particular Figure 4 {17-18, 20}, over the southeastern United States and Gulf of Mexico. Abatzoglou and Magnusdottir (2006) and Johnson et al. (2007) show RWB as precursor for increased moisture across the SWUS during the NAM. PWat increases in magnitude and northward extent during ridging confirming prior research (Figures 3-5){i.e., 15-16, 18-20}(Carleton et al. 1990; Higgins et al. 1999; Cerezo-Mota et al. 2011). The northward extension of the subtropical ridge brings easterly flow further north, which allows TEWs and TUTTs to directly affect the SWUS via increased atmospheric moisture and lightning (Figures 4-5 {13-20}).

MTs {16, 20} have the highest median lightning decile (Table 7), and MTs {16, 19-20} are most associated with days exceeding the 70$^{th}$ percentile of lightning across the SWUS. These MTs show ridging with the center of the subtropical ridge far enough north to transport larger magnitudes of PWat across the SWUS (Figures 3-5) {16, 19-20}. Abatzoglou and Brown (2009) demonstrated that lightning increased across the SWUS and Intermountain West during MJO phases 6-8. However, MJO phases 1 (17.7%) and 2

(17.1%) are most associated with days exceeding the 70th percentile of lightning across the SWUS. Coincidentally, MJO phases 1 and 2 are most associated with MTs demonstrating a ridge or ridge breakdown. However, ridge breakdown and troughing are associated with MJO phases 6-8 confirming findings from Abatzolgou and Brown (2009).

**4.2 Onset**

Tables 5-6 provide the median onset date, IQR, MAD, 10th percentile onset date, and 90th percentile onset date by PSA for the three onset methods implemented. Onset1 (SPH, VPD, and lightning) and Onset2 (VPD, precipitation, and lightning) generally yield the same results with some variation across eastern New Mexico and west Texas. Onset3 (SPH, precipitation, and lightning) unilaterally shows earlier median onset dates for every PSA except PSA SW14N, and is less variable than Onset1 and Onset2. Most onset dates for all the methods occur in early to mid July with a few instances of late June onset dates in some of the eastern New Mexico and west Texas PSAs. Depending on the PSA and method, the IQR can range from 10 to 53 days although most IQRs total less than 35 days. The MADs are generally between one to two weeks with a few exceptions totaling around three weeks, and the 10th and 90th percentile onset dates are generally separated by three to six weeks. Adjacent PSAs tended to have similar onset dates and occasionally the same onset date. Generally, onset progresses from east to west and south to north across the SWUS, but with large interannual variability.

Ridging including a northward displacement of the subtropical ridge, southerly and easterly flow, and increased moisture are associated with onset days (Figures 3-5)

{15-16, 18-20} (Table 7). These MTs are most associated with increased LS and have at least a median LS decile of 6 (Table 7). MTs {15-16, 18-20} total 28.7% of all days examined, but occur on 75.0% of onset days (Table 7). MTs {1-7} are never associated with onset for a PSA, and these MTs generally depict troughing and drier conditions over the SWA with small amounts of LS (Table 7)(Figures 3-5 {1-7}). These MTs also predominantly occur in April and May and rarely in June (Table 7). MJO phase 1 is most associated with onset accounting for 34.7% of all onset days among the PSAs while MJO phase 3 is least associated with onset (7.4%) (not shown).

### 4.3 Increased Fire Activity

There were 298 and 292 instances that exceeded the Busy95_PSA and Busy99_PSA thresholds, respectively (Table 8). 62.4% and 55.0% of days that exceeded the Busy95_PSA and Busy99_PSA thresholds, respectively, transpired before 1 July. 85.2% and 82.5% of days that exceeded the Busy95_PSA or Busy99_PSA daily thresholds for all wildfires occurred before the respective PSA's onset. 219 days exceeded the Busy95_SW threshold and 57 days exceeded Busy99_SW threshold with 66.2% and 77.2% of those days occurring before 1 July (Table 8). 95.4% of days that meet at least of one the busy event thresholds occur before 1 August including 75.6% before 1 July (Table 8). PSAs SW06S, SW08, and SW12 contributed the most often to exceeding the Busy95_SW and Busy95_SW thresholds with PSAs SW05 and SW06N closely trailing in their contributions (Table 9). On average, four PSAs had at least one large wildfire on days exceeding the Busy99_SW threshold, and two PSAs had at least one large wildfire on days exceeding the Busy95_SW threshold.

MTs {11-12, 16, 19-20} are most associated with days exceeding either the Busy95_SW or Busy99_SW threshold (Table 7). These MTs reflect two different atmospheric patterns: ridging {12, 16, 20} and ridge breakdown {11, 19}. MTs {4, 8, 11-12} have larger percentages associated with days exceeding either Busy_SW threshold than the occurrence percentage for all examined and onset days. MTs {16, 19-20} have larger onset occurrence percentages than percentages associated with days exceeding either Busy_SW threshold (Table 7). Several MTs {1, 17-18} are rarely associated (< 1%) with either Busy_SW threshold and those three MTs are relatively unique in comparison to each other (Table 7). MT {1} displays zonal, dry flow, MT {17} has ridging with the axis just inland from the Pacific Coast, and MT {18} shows a slightly eastward displaced four corners ridge.

**5. Discussion**

The amount of wildfires ignited is not always problematic for wildfire management. July has the most wildfires of any month, but due to the onset of the NAM many of these do not become large. On average, 1.9% of wildfires in July become large, but 3.8% of wildfires in June become large. Additionally, 3.5% of wildfires ignited in April through June become large while only 1.9% of wildfires become large in July through September (Table 3). Wildfires April through May coincide with increasing atmospheric and fuel dryness across the SWUS with dry and breezy MTs prevailing {1-8} (Mohrle 2003; Brandt 2006). June represents the transition from the dry and breezy spring to the NAM across the SWUS. During this transition moisture plumes traverse north and increase chances of thunderstorms. The moisture plumes in June are either

elevated or move quickly in a progressive atmospheric pattern reminiscent of dry thunderstorm conditions (Hall 2007; Wallmann et al. 2011; Nauslar et al. 2013). These initial northward ridge extensions and subsequent moisture intrusions into the SWUS could be deemed 'pre-monsoonal'. These moisture plumes coincide with peak fuel dryness helping produce the most active part of the wildfire season. The term 'Firesoon' has been used to describe the effects of these moisture plumes due to their propensity for triggering thunderstorms that ignite wildfires (Brown 2002). Not coincidentally, 75.6% of busy events occur before July and 83.9% of all busy days occur before onset (Table 8).

Busy days and events across the SWA are generally associated with suppressed ridging (subtropical ridge center in western/northwestern Mexico), ridge breakdown, or enhanced ridging (center of subtropical ridge in SWUS). The suppressed ridge pattern (Figure 3 {4, 8, 11}) has southerly flow over Arizona, but the deeper moisture is confined further south in Mexico (Figure 5 {4, 8, 11-12}. However, even a slight increase (1-2 g/kg) of SPH at 700 hPa over the SWUS provides sufficient moisture along with a weak disturbance for triggering at least isolated thunderstorms (Figure 6 {4, 8, 11-12})(Table 7). Below 700 hPa remains relatively dry creating an environment for dry thunderstorms (Hall 2007; Wallmann et al. 2011; Nauslar et al. 2013). Strong ridging brings ample moisture to the SWUS and helps trigger thunderstorms across most areas (Figures 3 and 5 {16, 20})(Table 7). Additionally, with increased moisture, thunderstorms are more likely to affect a larger elevation range, which would bring enhanced ignition potential to areas that experience less thunderstorm activity than preferred terrain features (i.e., Mogollon Rim)(Gebremichael et al. 2007; Gochis and Higgins 2007; Nesbitt et al. 2008). Strong ridging is more likely to induce busy days and events if it occurs earlier in the season

(i.e., before onset) where peak fuel dryness could offset the increase of atmospheric moisture. A breakdown of the subtropical ridge (Figure 3 {12, 19}) whether suppressed or enhanced relative to the SWUS is associated with busy days and events and represents a known critical fire weather pattern (Werth et al. 2011). This pattern is known to produce dry thunderstorms, but more importantly, it provides a pattern change from hot and relatively dry conditions under the influence of the subtropical ridge to increasing moisture and decreasing geopotential heights with an eastward shift of the ridge axis helping generate thunderstorms. As the subtropical ridge breaks down, drier and windy conditions usually develop with the passage of the trough allowing wildfires to grow in more favorable fire weather conditions.

While associating MTs with busy days or events is important, the atmospheric patterns that precede and follow those days are equally important. Four general atmospheric pattern progressions emerge when examining the evolution of MTs around busy days and events for the SWA: 1) zonal or southwest flow preceding ridging; 2) zonal or southwest flow transitioning into ridging followed by a return to zonal or southwest flow; 3) persistent ridging followed by zonal or southwest flow; and 4) fluctuation between suppressed and amplified ridging over the SWA with the ridge axis exhibiting east-west movement (Figure 7). Fuel curing takes place during some combination of hot, dry, and windy conditions that precedes, follows, or precedes and follows lightning associated with ridging. The fuel curing promotes receptive fuel beds for wildfire ignition and spread. However, these MTs and progressions of MTs are not unique to busy days and events. The time of the year MTs and progression of MTs occur is important. If they occur before onset, it is more likely a busy day or event transpires

since it coincides with drier fuels. Wildfires including large wildfires occur across a wide spectrum of fuel and weather conditions. However, median BI and ERC during busy days are greater than the overall medians for BI and ERC, and median FM100 and FM1000 during busy days are less than the overall medians for FM100 and FM1000 for all PSAs (not shown).

Holdovers are wildfires that are ignited but remain small (i.e., single tree) until dry, windy, and unstable conditions develop that promote wildfire growth. Holdovers frequently appear during dry, windy, and hot/unstable conditions (i.e., Figures 4-5 {11}) that follow days of lightning and increased atmospheric moisture. Several days may pass before a holdover is detected, and this demonstrates a challenge when using the FPA wildfire occurrence database. A wildfire's discovery date and a wildfire's ignition date may not match, and holdovers are difficult to quantify and predict (Anderson 2002). Wildfires occasionally exceed large wildfire thresholds and continue to grow after the discovery date. This poses a problem for analyzing busy days or events since large wildfire growth days do not always coincide with the discovery date. Holdovers and large wildfire growth after the discovery date introduce errors into the analysis. However, utilizing the busy day and event thresholds mitigate some of these issues since it considers relatively rare events (< 5% of occurrence) and deemphasizes total acres burned.

The NAM onset is easy to distinguish but difficult to define and objectively quantify. The methods implemented in this analysis produce realistic and meaningful annual NAM onset dates by PSA. The results demonstrate that the NAM onset does not happen at one time for the entire SWUS. Neighboring PSAs may experience onset on the

same date, but this does not happen every year. The intraseasonal and interannual variability associated with the NAM explains the inconsistent nature of onset across the SWUS. MTs {15-16, 18-20} are most associated with the NAM onset, and these MTs illustrate ridging with moist, southerly flow over most of the SWUS and a northward shift of moist easterlies over northern Mexico producing increased LS across the SWA (Figures 3-5 {15-16, 18-20})(Table 7). MT {17} is curiously unrelated to NAM onset (1.7%), but the ridge axis is shifted west inducing drier west-northwest flow over most of the SWUS confirming prior research (Carleton et al. 1990; Higgins et al. 1999; Cerezo-Mota et al. 2011).

**6. Conclusions**

We demonstrate that NAM-related weather systems modulate wildfire occurrence and spread patterns. Several MTs or MT progressions are associated with busy days or events (Figure 7):

- MTs {4, 11-12, 16, 19-20} (Table 7)
- Zonal or southwest flow preceding ridging
- Zonal or southwest flow transitioning into ridging followed by a return to zonal or southwest flow
- Persistent ridging followed by zonal or southwest flow
- Fluctuation between suppressed and amplified ridging over the SWUS with the ridge axis exhibiting east-west movement

Additionally, we identified several MTs that were most associated with the NAM onset (MTs {15-16, 18-20})(Table 7). These MTs exhibited a northward extension of the

subtropical ridge that shifts easterlies northward and increases moisture and lightning (Carleton et al. 1990; Higgins et al. 1999; Cerezo-Mota et al. 2011) across the SWA via TEWs (Adams and Stensrud 2007), TUTTs (Newman and Johnson 2012), and antecedent RWB (Abatzoglou and Magnusdottir 2006; Johnson et al. 2007). The MTs associated with increased wildfire activity and onset mirror previous results from SOM analysis in the region (i.e., Cavazos et al. 2002; Crimmins 2006).

Numerous studies (i.e., Swetnam and Beatencourt 1990, 1998; Westerling et al. 2003; Crimmins and Comrie 2005), researched climatological predictors for wildfire activity in the SWUS. Westerling et al. (2003) and Crimmins and Comrie (2005) identified fine fuel production as important for total acres burned. MEI was exploratory analyzed with no meaningful relationships identified. This analysis focused on intraseasonal variability associated with the NAM including the transition to and from it. The MJO more directly corresponds to intraseasonal variability due to its intrinsic temporal frequency, and that is why it was investigated more thoroughly than MEI.

Future work could use the Monitoring Trends in Burn Severity (MTBS) database to calculate area burned on a daily timescale and apply the same rigorous atmospheric pattern and fuel indices analysis. This would mitigate issues with holdovers and large wildfire growth, but would be confined to examining only larger wildfires due to the spatial resolution and accuracy of MTBS (Kolden et al. 2012). Utilizing both discovery and control dates for wildfires could provide a proxy for determining wildfire season end. However, control dates are not solely based on atmospheric and fuel conditions. Wildfire management policy plays an important role and could introduce error and bias when using control dates.

The results of this work provide decision support information and improve understanding of atmospheric processes associated with NAM and their impact on wildfire activity. The improved understanding benefits operational fire meteorologists and managers with the identification of atmospheric patterns associated with increased wildfire activity, which improves planning and logistical strategies (Ray et al. 2007). A method for determining the NAM onset by PSA is established with realistic results that correspond to atmospheric patterns that promote increased moisture and lightning. We also provide an atmospheric circulation analysis of the transition to and from the NAM across the SWA demonstrating intra- and inter-annual variability. Ridging prevails during the NAM, but interactions between approaching and passing troughs with the subtropical ridge drive much of the intraseasonal variability for the atmosphere and wildfire activity in the SWA.

*Acknowledgements*

We would like to thank the USDA Forest Service for providing the funding (10-CS-11130206-047), and Vaisala for allowing the use of lightning data for wildfire applications. We would also like to thank Chuck Maxwell, Southwest PS Program Manager, John Saltenberger and Terry Marsha, Northwest PS, for their input and advice for the paper. Ben Hatchett helped create some of the figures and offered revisions early in the writing process. Finally, we would like to thank David Reusch and Michael Crimmins for their expertise and help on SOMs, John Abatzolgou for helping obtain and

organize the METDATA, and Andrew Joros for helping obtain and organize the NARR dataset.

**References**

Abatzoglou J. T., and G. Magnusdottir, 2006: Planetary wave breaking and nonlinear feflection: Seasonal cycle and interannual variability. *J. Climate*, **19**, 6139-6152.

Abatzoglou, J. T., and T. J. Brown, 2009: Influence of the Madden–Julian Oscillation on summertime cloud-to-ground lightning activity over the continental United States. *Mon. Wea. Rev.*, **137,** 3596-3601.

Abatzoglou J. T., 2013: Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.*, **33**, 121-131.

Adams, D. K., and A. C. Comrie, 1997: The North American Monsoon. *Bull. Amer. Meteor. Soc.*, **78,** 2197-2213.

Adams, J. L., and D. J. Stensrud, 2007: Impact of tropical easterly waves on the North American Monsoon. *J. Climate*, **20,** 1219-1238.

Anderson K, 2002: A model to predict lightning-caused fire occurrences. *Int. J. Wildland Fire*, **11**, 163-172.

Barlow, M., S. Nigam, and E. H. Berbery, 1998: Evolution of the North American Monsoon system. *J. Climate*, **11,** 2238-2257.

Becker, E. J., E. H. Berbery, and R. W. Higgins, 2009: Understanding the Characteristics of Daily Precipitation over the United States Using the North American Regional Reanalysis. *J. Climate*, **22**, 6268–6286.

Brandt, R. R., 2006: The North American Monsoon system in southern Arizona. Dissertation, Department of Geography and Regional Development, University of Arizona, 130 pp.

Brown, T. J., 2002: The North American "Firesoon". In report on research opportunities for climate and society interactions in the North American Monsoon region, Workshop on Applications and Human Dimensions of Monsoon Research, 18-20 June 2001, Tucson, AZ, A.J Ray and R.S. Webb (Eds.), Report for NOAA OAR, 64 pp.

Carleton, A. M., 1986: Synoptic-dynamic character of “bursts” and “breaks” in the southwest U.S. summer precipitation singularity. *J. Climatol.,* **6,** 605–623.

Carleton, A. M., D. A. Carpenter, and P. J. Weser, 1990: Mechanisms of interannual variability of the southwest United States summer rainfall maximum. *J. Climate*, **3,** 999-1015.

Cavazos, T., 2000: Using Self-Organizing Maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans. *J. Climate*, **13**, 1718–1732.

Cavazos, T., A. C. Comrie, and D. M. Liverman, 2002: Intraseasonal variability associated with wet monsoons in southeast Arizona. *J. Climate*, **15,** 2477-2490.

Cerezo-Mota, R., M. Allen, and R. Jones, 2011: Mechanisms controlling precipitation in the northern portion of the North American Monsoon. *J. Climate*, **24,** 2771-2783.

Cohen, J.D. and J.E. Deeming, 1985: The National Fire Danger Rating System: Basic equations. USDA Forest Service, Pacific Southwest Forest and Range Experiment Station, GTR PSW-82, 16 pp.

Crimmins, M. A., and A. C. Comrie, 2005: Interactions between antecedent climate and wildfire variability across southeastern Arizona. *Int. J. Wild- land Fire*, **13**, 455-466.

Crimmins, M. A., 2006: Synoptic climatology of extreme fire-weather conditions across the southwest United States. *Int. J. Climatol*., **26,** 1001–1016.

Deeming J. E., R. E. Burgan, and J. D. Cohen, 1978. The National Fire-Danger Rating System – 1978. Revisions to the 1978 National Fire-Danger Rating System. General Technical Report INT-39. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station; 63.

Dominguez, F. and P. Kumar, 2008: Precipitation Recycling Variability and Ecoclimatological Stability—A Study Using NARR Data. Part I: Central U.S. Plains Ecoregion. *J. Climate*, **21**, 5165–5186.

Dominguez, F., P. Kumar, and E. R. Vivoni, 2008: Precipitation recycling variability and ecoclimatological stability—a study using NARR data. Part II: North American Monsoon region. *Journal of Climate*, **21,** 5187-5203.

Doubler, D.L., J. A. Winkler, X. Bian, C. K. Walters, and S. Zhong, 2015: An NARR-derived climatology of southerly and northerly low-level jets over North America and coastal environments. *J. Appl. Meteor. Climatol.*, **54**, 1596–1619.

Douglas, M.W., R.A. Maddox, K. Howard, and S. Reyes, 1993: The Mexican monsoon. *J. Climate*, **6**, 1665- 1677.

Douglas, M. W., 1995: The summertime low-level jet over the Gulf of California. *Mon. Wea. Rev.*, **123,** 2334-2347.

Dowdy, A. J. and G. A. Mills, 2012: Atmospheric and fuel moisture characteristics associated with lightning-attributed fires. *J. Appl. Meteor. Climatol*., **51**, 2025–2037.

Earth System Research Laboratory Physical Sciences Division (ESRL PSD), 2014. Multivariate ENSO Index (MEI). http://www.esrl.noaa.gov/psd/enso/mei/index.html.

Evett, R.R., Mohrle, C.R., Hall, B.L., Brown, T.J., and S.L. Stephens, 2008: The effect of monsoonal atmospheric moisture on lightning wildfire ignitions in southwestern North America. *Agric. Forest Meteorol.,* **148***,* 1478-1487.

Favors, J. E., and J. T. Abatzoglou, 2012: Regional surges of monsoonal moisture into the southwestern United States. *Mon. Wea. Rev.*, **141,** 182-191.

Gebremichael, M., E. R. Vivoni, C. J. Watts, and J. C. Rodríguez, 2007: Submesoscale spatiotemporal variability of North American monsoon rainfall over complex terrain. *J. Climate*, **20**, 1751-1773.

Gochis, D. J., and R. W. Higgins, 2007: The path to improving predictions of the North American Monsoon. *Newsletter, U.S. CLIVAR Variations*, **5**.

Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona, 2007: Seasonal shifts in the North American Monsoon. *J. Climate*, **20,** 1923-1935.

Guttman, N. B., 1999: Accepting the Standandardized Precipitation Index: A calculation algorithm. *J. Amer. Water Resour. Assoc*., **35**, 311-322, https://climatedataguide.ucar.edu/climate-data/standardized-precipitation-index-spi#sthash.HbPjQOmi.dpuf.

Hall, B. L., 2007: Precipitation associated with lightning-ignited wildfires in Arizona and New Mexico. *Int. J. Wildland Fire*, **16,** 242-254.

Hawkins, T. W., A. W. Ellis, J. A. Skindlov, and D. Reigle, 2002: Intra-annual analysis of the North American snow cover–monsoon teleconnection: Seasonal forecasting utility. *J. Climate*, **15,** 1743-1753.

Hendon, H. H., K. R. Sperber, D. E. Waliser, M. C. Wheeler, 2011: Modeling monsoon intraseasonal variability: From theory to operational forecasting. *Bull. Amer. Meteor. Soc.*, **92**, ES32–ES35.

Hewitson, B. C., Crane, R. G., 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Research*, **22**, 13–26.

Higgins, R. W., Y. Yao, and X. L. Wang, 1997: Influence of the North American Monsoon system on the U.S. summer precipitation regime. *J. Climate*, **10,** 2600-2622.

Higgins, R. W., K. C. Mo, and Y. Yao, 1998: Interannual variability of the U.S. summer precipitation regime with emphasis on the Southwestern Monsoon. *J. Climate*, **11,** 2582-2606.

Higgins, R. W., Y. Chen, and A. V. Douglas, 1999: Interannual variability of the North American warm season precipitation regime. *J. Climate*, **12,** 653-680.

Johnson, R. H., P. E. Ciesielski, B. D. McNoldy, P. J. Rogers, and R. K. Taft, 2007: Multiscale variability of the flow during the North American Monsoon Experiment. *J. Climate*, **20,** 1628-1648.

Kalnay, E. and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc*., **77**, 437-470.

Kohonen, T., 2001: *Self-Organizing Maps*. Springer-Verlag, 503 pp.

Kolden, C. A., J. A. Lutz, C. H. Key, J. T. Kane, and J. W. van Wagtendonk, 2012: Mapped versus actual burned area within wildfire perimeters: Characterizing the unburned *Forest Ecol. Manage.*, **286**, 38–47.

Lo, F., and M. P. Clark, 2002: Relationships between spring snow mass and summer precipitation in the southwestern United States associated with the North American Monsoon system. *J. Climate*, **15,** 1378-1385.

Madden R.A., Julian P.R., 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702-708.

——, and ——, 1972: Description of global-scale circulation cells in the tropics with a 40-50 day period. *J. Atmos. Sci* ., **29**, 1109-1123.

Means, J. D., 2012: GPS Precipitable Water as a Diagnostic of the North American Monsoon in California and Nevada. *J. Climate*, **26,** 1432-1444.

Mejia, J. F., M. W. Douglas, M. W., and P. J. Lamb, 2015: Observational investigation of relationships between moisture surges and mesoscale to large-scale convection during the North American Monsoon. *Int. J. Climatol*.

Mesinger F., and Coauthors, 2006: North American Regional Reanalysis. Bulletin of the A.M.S., **87**, 343-360.

Mitchell, D. L., D. Ivanova, R. Rabin, T. J. Brown, and K. Redmond, 2002: Gulf of California sea surface temperatures and the North American Monsoon: Mechanistic implications from observations. *J. Climate*, **15,** 2261-2281.

Mohrle, C. R., 2003: The Southwest Monsoon and relation to fire occurence. M.S. thesis, Department of Physics/Division of Atmospheric Sciences (DRI), University of Nevada, Reno, 97 pp. Online at http://www.cefa.dri.edu/Publications/charliethesis.pdf.

National Weather Service Weather Forecast Office Tucson, 2015.: Monsoon Information Index, 12 November 2015. [Available online at: http://www.wrh.noaa.gov/twc/monsoon/monsoon_info.php].

Nauslar, N. J., M. L. Kaplan, J. Wallmann, and T. J. Brown, 2013: A forecast procedure for dry thunderstorms. *J. Operational Meteor.*, **1**, 200–214.

Nesbitt, S. W., D. J. Gochis, and T. J. Lang, 2008: The diurnal cycle of clouds and precipitation along the Sierra Madre Occidental observed during NAME-2004: Implications for warm season precipitation estimation in complex terrain. *J. Hydrometeor.*, **9,** 728-743.

Newman, A. and R. H. Johnson, 2012: Mechanisms for Precipitation Enhancement in a North American Monsoon Upper-Tropospheric Trough. *J. Atmos. Sci.*, **69**, 1775–1792.

Radhakrishna, B., F. Fabry, J.J. Braun, and T. Van Hove, 2015: Precipitable water from GPS over the continental United States: Diurnal cycle, intercomparisons with NARR, and link with convective initiation. *J. Climate*, **28**, 2584–2599.

Ray, A. J., G. M. Garfin, M. Wilder, M. Vásquez-León, M. Lenart, and A. C. Comrie, 2007: Applications of monsoon research: Opportunities to inform decision making and reduce regional vulnerability, *J. Climate*, **20**, 9, 1608-1627.

Reiter, E. R., and M. Tang, 1984: Plateau effects on diurnal circulation patterns. *Mon. Wea. Rev.,* **112,** 638–651.

Reusch, D. B., R. B. Alley, and B. C. Hewitson, 2005: Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. *Polar Geogr.*, **29**, 227 – 251.

Reusch, D. B., R. B. Alley, and B. C. Hewitson, 2007: North Atlantic climate variability from a self-organizing map perspective, *J. Geophys. Res.*, **112**, D02104.

Reusch, D. B., 2011: Nonlinear climatology and paleoclimatology: Capturing patterns of variability and change with self-organizing maps, *Physics and Chemistry of the Earth*, **35**, 329-340.

Rodriguez-Iturbe, I., P. D'Odorico, and A. Rinaldo, 1998: Possible self-organizing dynamics for land–atmosphere interaction. *J. Geophys. Res.,* **103D,** 23 071–23 077.

Sammon, J. W., 1969. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, **18**, 401–409.

Seastrand, S., Y. Serra, C. Castro, and E. Ritchie, 2015: The dominant synoptic-scale modes of North American monsoon precipitation. *Int. J. Climatol*., **35**, 2019–2032.

Short, K. C., 2015: Spatial wildfire occurrence data for the United States, 1992-2013 [FPA_FOD_20150323]. 3rd Edition. Fort Collins, CO: Forest Service Research Data Archive.

Stensrud, D. J., R. L. Gall, and M. K. Nordquist, 1997: Surges over the Gulf of California during the Mexican Monsoon. *Mon. Wea. Rev.*, **125,** 417-437.

Stensrud, D. J., 2013: Upscale effects of deep convection during the North American Monsoon. *J. Atmos. Sci*., **70**, 2681–2695.

Swetnam, T. W., and J. L. Betancourt, 1990: Fire–Southern Oscillation relations in the southwestern United States. *Science,* **249,** 1017–1020.

——, and ——, 1998: Mesoscale disturbance and ecological response to decadal climatic variability in the American southwest. *J. Climate,* **11,** 3128–3147.

Tang, M., and E. R. Reiter, 1984: Plateau monsoons of the Northern Hemisphere: A comparison between North America and Tibet. *Mon. Wea. Rev*., **112**, 617-637.

Turrent, C. and T. Cavazos, 2012: A numerical investigation of wet and dry onset modes in the North American Monsoon core region. Part I: A regional mechanism for interannual variability. *J. Climate,* **25**, 3953–3969.

Vera, C., and Coauthors, 2006: Toward a unified view of the American Monsoon systems. *J. Climate*, **19,** 4977-5000.

Wallmann, J., R. Milne, C. Smallcomb, and M. Mehle, 2010. Using the 21 June 2008 California lightning outbreak to improve dry lightning forecast procedures. *Wea. Forecasting*, **25**, 1447-1462.

Watson, A. I., R. L. Holle, and R. E. López, 1994: Cloud-to-ground lightning and upper-air patterns during bursts and breaks in the Southwest Monsoon. *Mon. Wea. Rev.,* **122,** 1726-1739.

Werth, P.A., Potter, B.E., Clements, C.B., Finney, M.A., Goodrick, S.L., Alexander, M.E., Cruz, M.G., Forthofer, J.A., McAllister, S.S., 2011. Synthesis of Knowledge of Extreme Fire Behavior: Volume I for Fire Managers. Gen. Tech. Rep. PNW-GTR-854. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Westerling AL, Brown TJ, Gershunov A., Cayan DR, Dettinger MD (2003) Climate and Wildfire in the Western United States. Bull. Amer. Met. Soc., 84(5) 595-604

Wheeler, M. and H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev*., **132**, 1917-1932.

Wolter, K., 1987: The Southern Oscillation in surface circulation and climate over the tropical Atlantic, Eastern Pacific, and Indian Oceans as captured by cluster analysis. *J. Climate Appl. Meteor.*, **26**, 540-558.

Wolter, K., and M.S. Timlin, 1993: Monitoring ENSO in COADS with a seasonally adjusted principal component index. *Proc. of the 17th Climate Diagnostics Workshop*, Norman, OK, NOAA/NMC/CAC, NSSL, Oklahoma Clim. Survey, CIMMS and the School of Meteor., Univ. of Oklahoma, 52-57.

——, and ——, 2011: El Niño/Southern Oscillation behavior since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext). *Int. J. Climatol.*, **31**, 14pp.

**Figures**

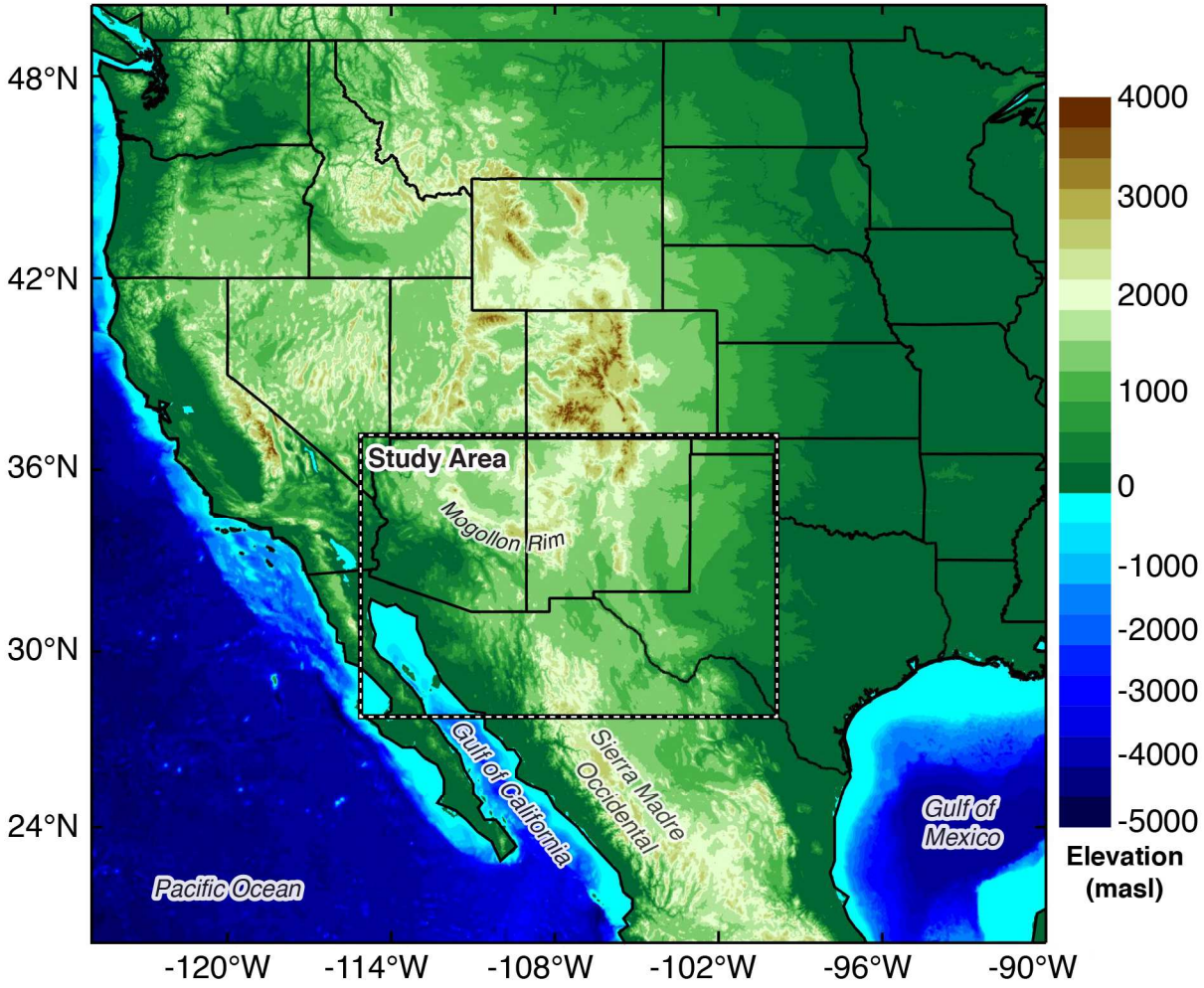


Figure 1. Elevation map of study area (m) with key geographic features labeled.

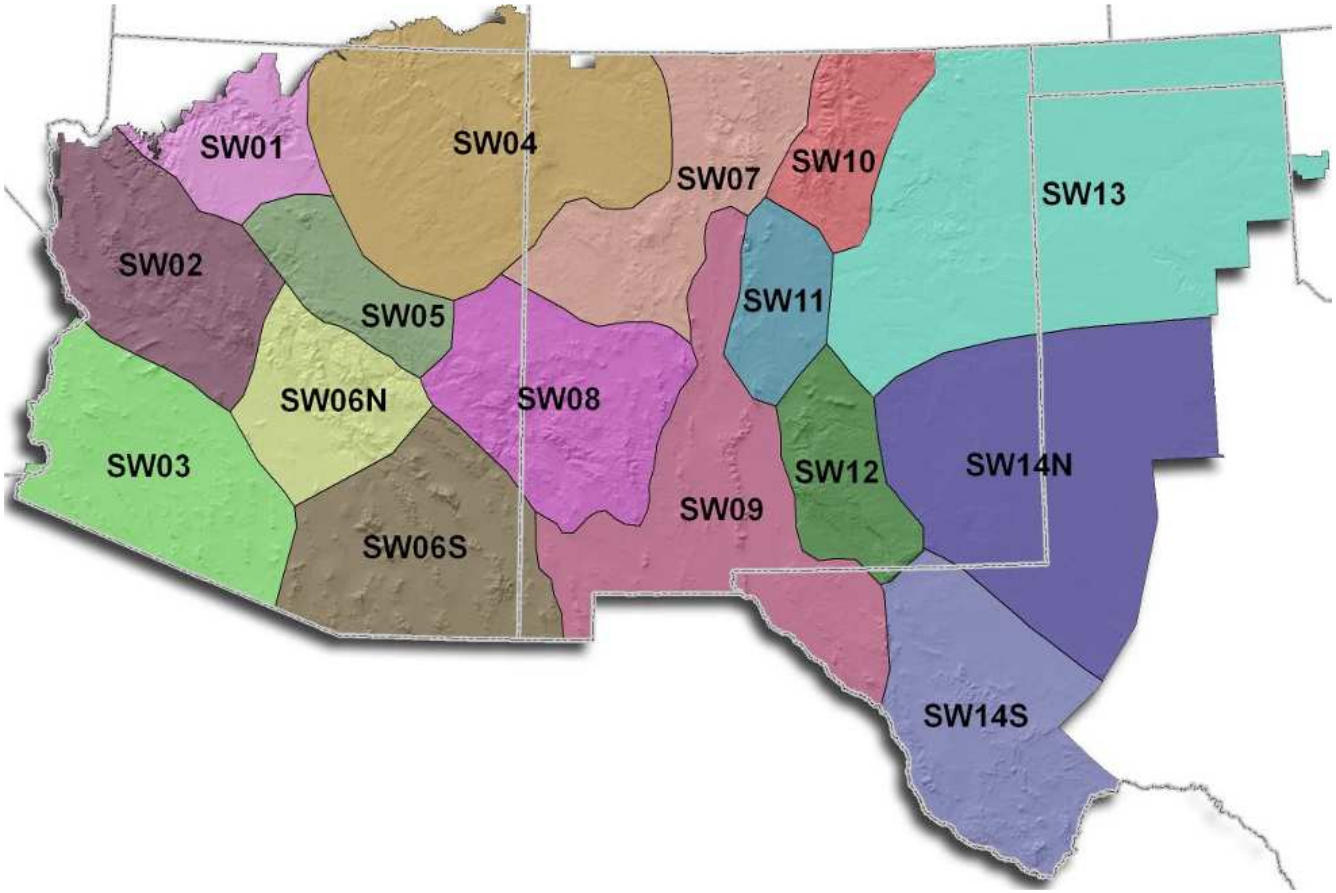


Figure 2. Study area showing the Southwest Area Predictive Service Areas (PSAs). SW01 is also known as PSA 1.

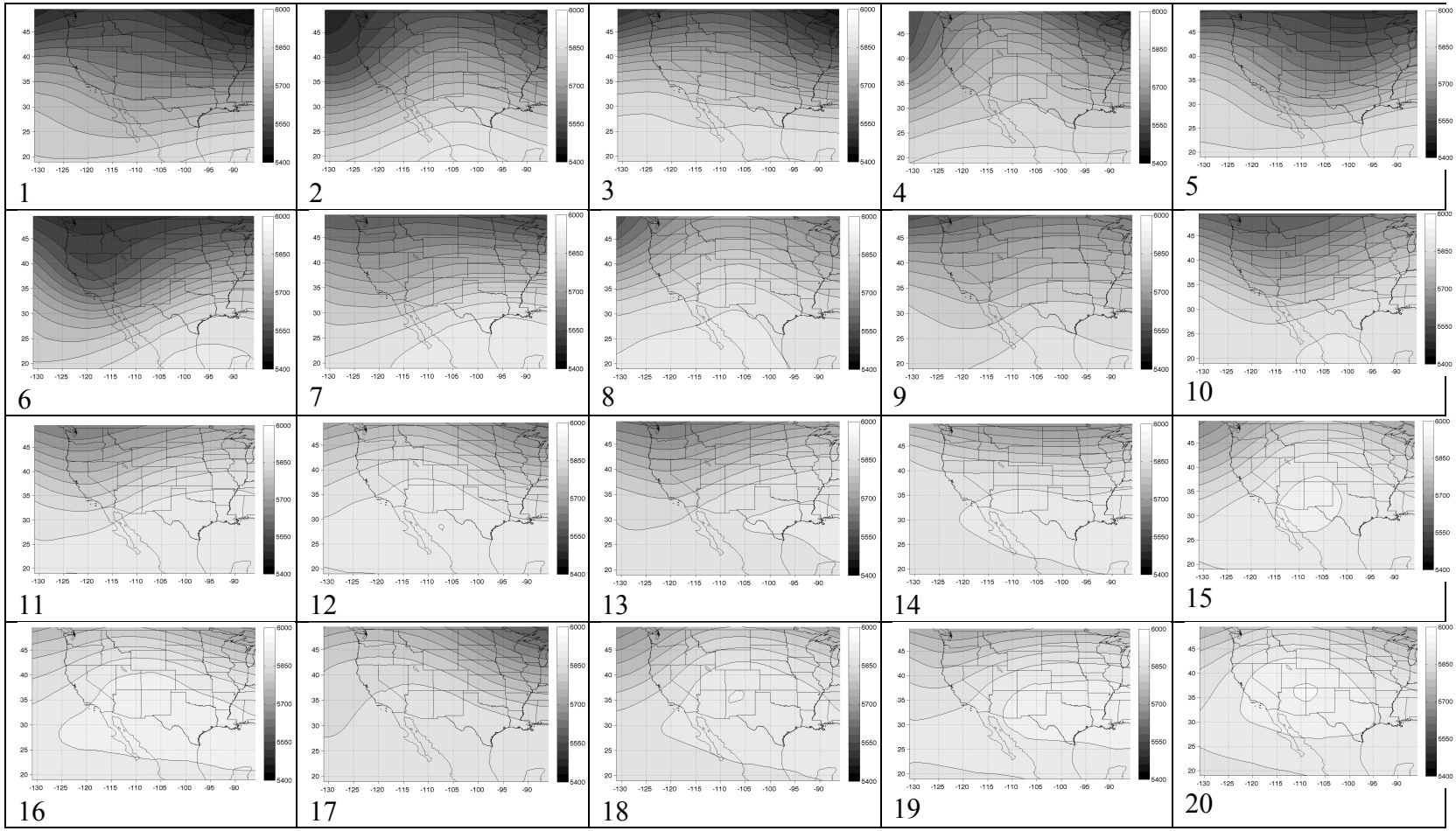


Figure 3. 20 SOM reference nodes {MTs} in the 4x5 SOM grid arrangement for 500 hPa geopotential heights (m), contoured and color shaded every 30 m

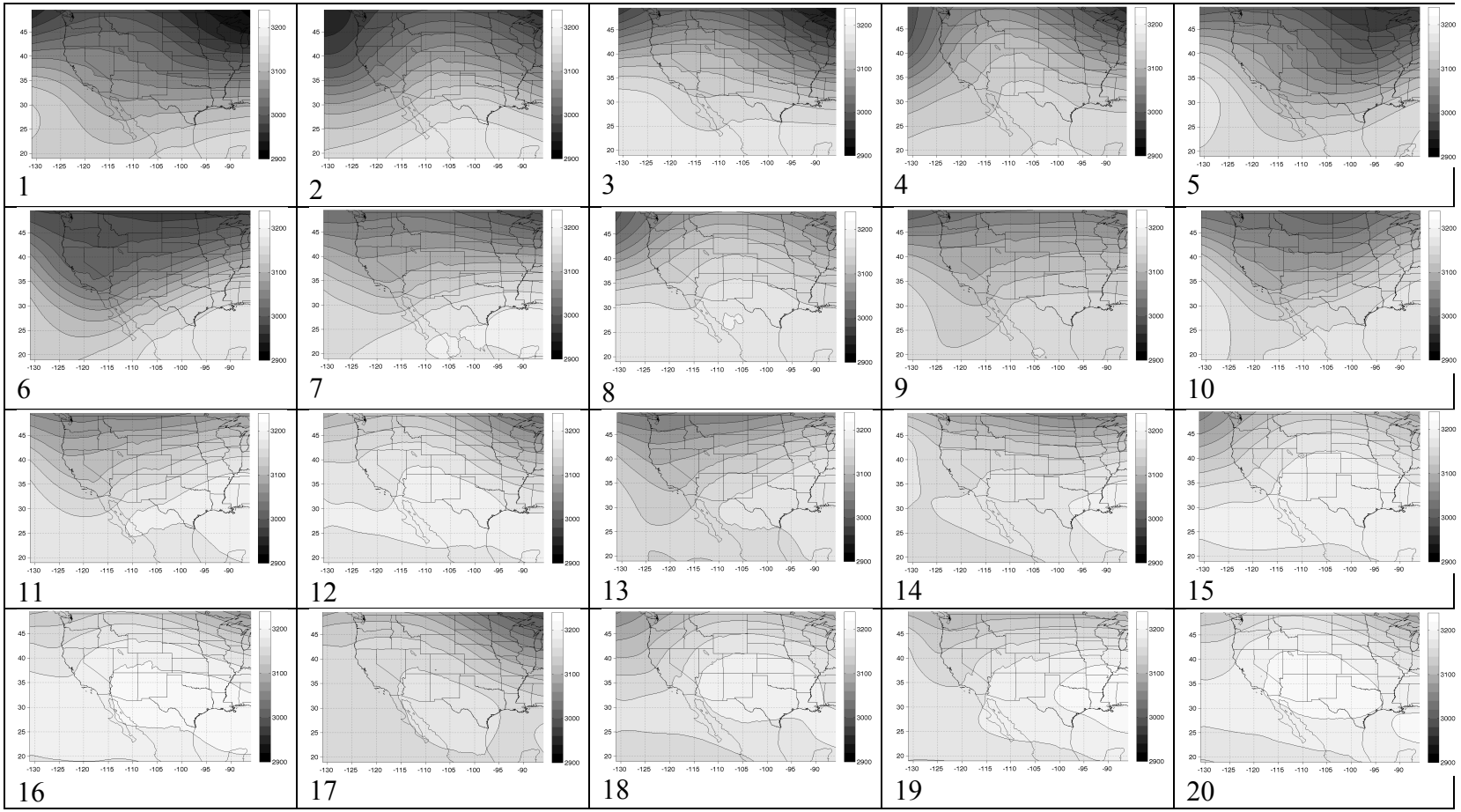


Figure 4. 20 SOM reference nodes {MTs} in the 4x5 SOM grid arrangement for 700 hPa geopotential heights (m), contoured and color shaded every 20 m

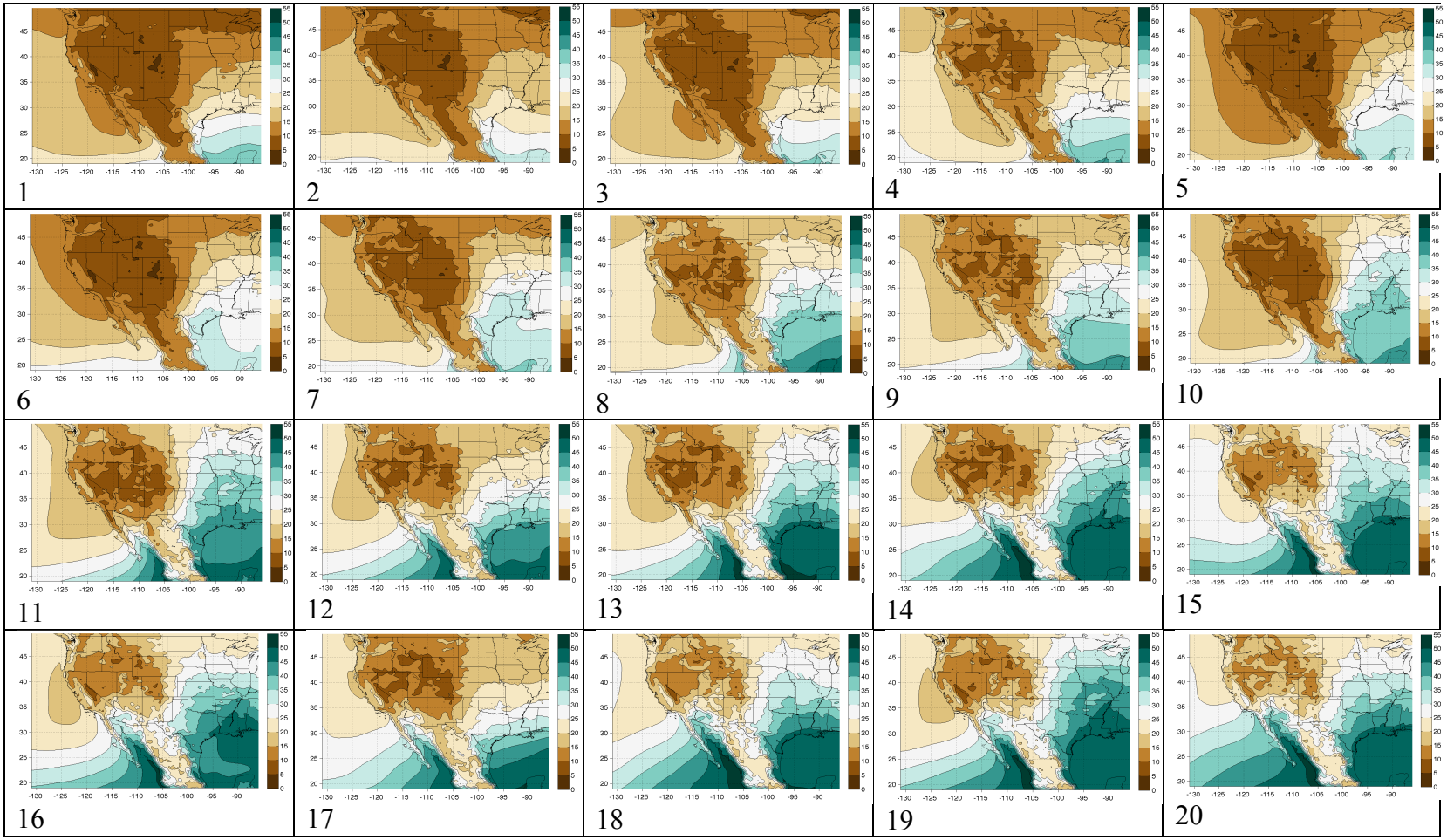


Figure 5. 20 SOM reference nodes {MTs} in the 4x5 SOM grid arrangement for precipitable water (mm), contoured and color shaded every 5 mm

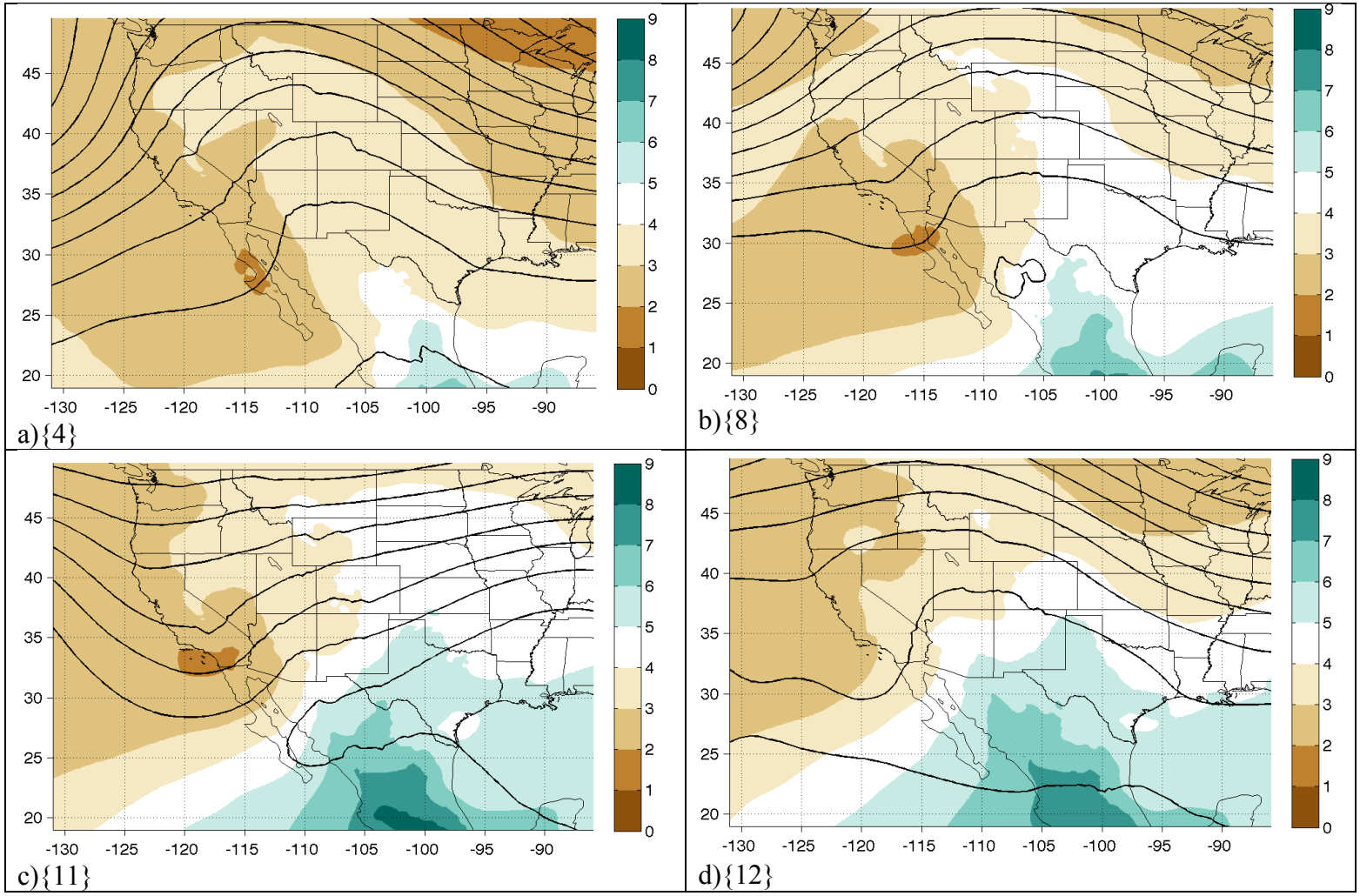


Figure 6. 700 hPa geopotential heights contoured every 20 m (black, solid lines), and specific humidity color shaded every 1 g/kg for MTs {4, 8, 11-12}.

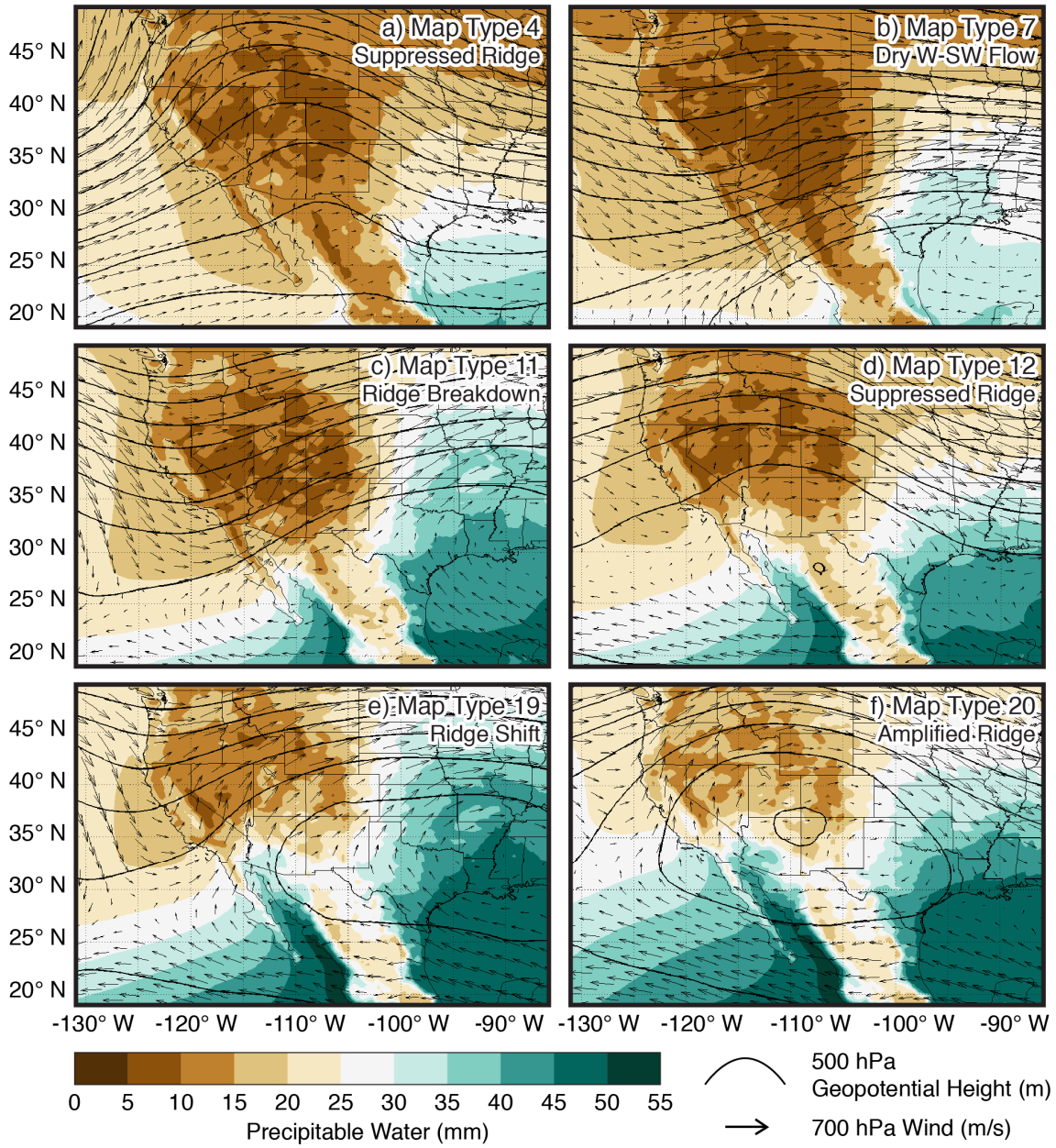


Figure 7. MTs represented in MT progressions associated with increased wildfire activity.

**Tables**

| PSA | Large Wildfire Threshold (Acres) |
|---|---|
| SW01 | 100 |
| SW02 | 300 |
| SW03 | 300 |
| SW04 | 100 |
| SW05 | 50 |
| SW06S | 100 |
| SW07 | 100 |
| SW08 | 300 |
| SW09 | 300 |
| SW10 | 100 |
| SW11 | 100 |
| SW12 | 100 |
| SW13 | 2000 |
| SW14S | 200 |
| SW14N | 2000 |
| SW06N | 500 |

Table 1. Large wildfire threshold for each PSA in the SWA as determined by Southwest PS.

| | |
|---|---|
| **Total Wildfires Per Year** | 4064 |
| **Total Large Wildfires Per Year** | 108 |
| **Lightning Caused Wildfires Per Year** | 1972 |
| **Lightning Caused Wildfires Per Year** | 62 |
| **Human Caused Wildfires Per Year** | 2092 |
| **Human Caused Large Wildfires Per Year** | 46 |

Table 2. Average number of wildfires per year for the SWA by type and size for 1992-2013

| | Wildfires | Large Wildfires | Lightning Caused Wildfires | Lightning Caused Large Wildfires | Human Wildfires | Human Caused Large Wildfires |
|---|---|---|---|---|---|---|
| **April** | 8318 | 300 | 744 | 67 | 7574 | 233 |
| **May** | 14789 | 417 | 3348 | 151 | 11441 | 266 |
| **June** | 21443 | 822 | 8774 | 476 | 12669 | 346 |
| **July** | 25687 | 496 | 17367 | 375 | 8320 | 121 |
| **August** | 13899 | 267 | 10545 | 229 | 3354 | 38 |
| **September** | 5271 | 84 | 2609 | 67 | 2662 | 17 |

Table 3. Total SWA wildfires by type and month for 1992-2013

| | April | May | June | July | August | September |
|---|---|---|---|---|---|---|
| **MT1** | 134 | 1 | 0 | 0 | 0 | 0 |
| **MT2** | 104 | 18 | 0 | 0 | 0 | 0 |
| **MT3** | 98 | 43 | 0 | 0 | 0 | 0 |
| **MT4** | 54 | 64 | 2 | 0 | 0 | 0 |
| **MT5** | 103 | 36 | 0 | 0 | 0 | 0 |
| **MT6** | 107 | 29 | 0 | 0 | 0 | 0 |
| **MT7** | 35 | 137 | 15 | 0 | 0 | 3 |
| **MT8** | 1 | 93 | 79 | 0 | 0 | 6 |
| **MT9** | 16 | 124 | 33 | 0 | 0 | 13 |
| **MT10** | 8 | 106 | 51 | 0 | 0 | 9 |
| **MT11** | 0 | 7 | 164 | 19 | 8 | 30 |
| **MT12** | 0 | 4 | 76 | 17 | 27 | 83 |
| **MT13** | 0 | 4 | 88 | 12 | 7 | 126 |
| **MT14** | 0 | 0 | 29 | 55 | 70 | 76 |
| **MT15** | 0 | 0 | 34 | 83 | 86 | 36 |
| **MT16** | 0 | 0 | 34 | 146 | 97 | 8 |
| **MT17** | 0 | 16 | 29 | 1 | 0 | 160 |
| **MT18** | 0 | 0 | 14 | 52 | 110 | 78 |
| **MT19** | 0 | 0 | 8 | 170 | 145 | 19 |
| **MT20** | 0 | 0 | 4 | 127 | 132 | 13 |

Table 4. Displaying map type (MT) occurrence by month.

| | SW1 | SW2 | SW3 | SW4 | SW5 | SW6S | SW7 | SW8 | SW9 | SW10 | SW11 | SW12 | SW13 | SW14S | SW14N | SW6N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Onset1 (SPH, VPD, Lightning)** | 24 Jul | 14 Jul | 30 Jul | 20 Jul | 18 Jul | 8 Jul | 15 Jul | 11 Jul | 6 Jul | 13 Jul | 8 Jul | 12 Jul | 30 Jun | 26 Jun | 27 Jul | 18 Jul |
| **Onset2 (VPD, Precip, Lightning)** | 24 Jul | 14 Jul | 30 Jul | 20 Jul | 18 Jul | 9 Jul | 15 Jul | 11 Jul | 6 Jul | 17 Jul | 13 Jul | 14 Jul | 4 Jul | 2 Jul | 2 Jul | 18 Jul |
| **Onset3 (SPH, Precip, Lightning)** | 18 Jul | 11 Jul | 13 Jul | 13 Jul | 8 Jul | 2 Jul | 5 Jul | 1 Jul | 2 Jul | 29 Jun | 6 Jul | 8 Jul | 15 Jun | 22 Jun | 6 Jul | 8 Jul |

Table 5. PSA median onset date for the three different onset methods from 1995-2013

| | SW1 | SW2 | SW3 | SW4 | SW5 | SW6S | SW7 | SW8 | SW9 | SW10 | SW11 | SW12 | SW13 | SW14 | SW15 | SW6N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Onset1** | 206 | 196 | 212 | 202 | 200 | 190 | 197 | 193 | 188 | 195 | 190 | 194 | 182 | 178 | 209 | 200 |
| **IQR** | 35 | 25 | 32 | 19 | 19 | 10 | 21 | 15 | 17 | 29 | 25 | 26 | 21 | 34 | 52 | 19 |
| **MAD** | 17 | 10 | 14 | 10 | 10 | 6 | 11 | 8 | 7 | 13 | 13 | 14 | 11 | 11 | 19 | 11 |
| **90th** | 234 | 224 | 243 | 225 | 220 | 202 | 216 | 206 | 212 | 210 | 214 | 217 | 205 | 217 | 233 | 216 |
| **10th** | 186 | 185 | 186 | 186 | 185 | 181 | 183 | 180 | 172 | 168 | 163 | 170 | 156 | 166 | 157 | 182 |
| | | | | | | | | | | | | | | | | |
| **Onset2** | 206 | 196 | 212 | 202 | 200 | 191 | 197 | 193 | 188 | 199 | 195 | 196 | 186 | 184 | 184 | 200 |
| **IQR** | 35 | 25 | 32 | 19 | 19 | 10 | 25 | 16 | 17 | 28 | 20 | 36 | 34 | 39 | 53 | 17 |
| **MAD** | 17 | 10 | 14 | 10 | 10 | 6 | 12 | 8 | 7 | 14 | 10 | 18 | 19 | 16 | 26 | 5 |
| **90th** | 234 | 224 | 243 | 225 | 220 | 202 | 222 | 207 | 212 | 216 | 216 | 225 | 222 | 227 | 223 | 216 |
| **10th** | 186 | 185 | 186 | 186 | 185 | 181 | 183 | 180 | 164 | 177 | 166 | 166 | 157 | 166 | 154 | 182 |
| | | | | | | | | | | | | | | | | |
| **Onset3** | 200 | 193 | 195 | 195 | 190 | 184 | 187 | 183 | 184 | 181 | 188 | 190 | 167 | 174 | 188 | 190 |
| **IQR** | 20 | 16 | 20 | 18 | 15 | 10 | 16 | 14 | 11 | 21 | 18 | 32 | 22 | 40 | 37 | 15 |
| **MAD** | 12 | 9 | 9 | 9 | 7 | 7 | 8 | 8 | 7 | 6 | 9 | 13 | 8 | 15 | 20 | 8 |
| **90th** | 220 | 213 | 229 | 213 | 199 | 197 | 196 | 197 | 206 | 201 | 200 | 223 | 193 | 214 | 219 | 207 |
| **10th** | 182 | 183 | 179 | 181 | 181 | 173 | 175 | 170 | 168 | 173 | 170 | 169 | 157 | 158 | 157 | 173 |

Table 6. IQR (number of days), MAD (number of days), 10th percentile onset date, and 90th percentile onset day-of-year (DOY) by PSA for the three onset methods from 1995-2013

| | Count | All Days | Busy95_SW | Busy99_SW | Onset1 | Median LS Decile |
|---|---|---|---|---|---|---|
| **MT1** | 135 | 3.4% | 0.5% | 0.0% | 0.0% | 1 |
| **MT2** | 122 | 3.0% | 2.3% | 3.5% | 0.0% | 1 |
| **MT3** | 141 | 3.5% | 2.7% | 1.8% | 0.0% | 1 |
| **MT4** | 120 | 3.0% | 6.4% | 8.8% | 0.0% | 3 |
| **MT5** | 139 | 3.5% | 2.3% | 1.8% | 0.0% | 1 |
| **MT6** | 136 | 3.4% | 2.7% | 3.5% | 0.0% | 1 |
| **MT7** | 190 | 4.7% | 4.1% | 3.5% | 0.0% | 2 |
| **MT8** | 179 | 4.4% | 5.5% | 5.3% | 1.8% | 4 |
| **MT9** | 186 | 4.6% | 3.2% | 3.5% | 0.7% | 4 |
| **MT10** | 174 | 4.3% | 2.7% | 1.8% | 0.0% | 2 |
| **MT11** | 228 | 5.7% | 11.0% | 14.0% | 2.8% | 4 |
| **MT12** | 207 | 5.1% | 8.2% | 10.5% | 3.9% | 5 |
| **MT13** | 237 | 5.9% | 2.7% | 5.3% | 8.5% | 4 |
| **MT14** | 230 | 5.7% | 2.7% | 5.3% | 6.0% | 5 |
| **MT15** | 239 | 5.9% | 5.0% | 1.8% | 9.9% | 6 |
| **MT16** | 285 | 7.1% | 12.8% | 7.0% | 19.0% | 8 |
| **MT17** | 206 | 5.1% | 2.3% | 0.0% | 1.4% | 3 |
| **MT18** | 254 | 6.3% | 3.2% | 0.0% | 12.3% | 6 |
| **MT19** | 342 | 8.5% | 9.6% | 10.5% | 18.7% | 7 |
| **MT20** | 276 | 6.9% | 10.0% | 12.3% | 15.1% | 8 |

Table 7. Showing the percentage of occurrence for each map type (MT) for all days, Busy95_SW, Busy99_SW, and Onset1 days and the median LS decile (Dec) for each MT.

| | Count | % Before Onset | % Before 1 July | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|
| **Busy95_PSA** | 298 | 85.23% | 62.4% | 28 | 40 | 118 | 76 | 33 | 3 |
| **Busy99_PSA** | 292 | 82.53% | 55.1% | 25 | 31 | 105 | 98 | 30 | 3 |
| **Busy95_SW** | 219 | | 66.2% | 23 | 31 | 91 | 51 | 22 | 1 |
| **Busy99_SW** | 57 | | 77.2% | 6 | 10 | 28 | 11 | 2 | 0 |
| **Event_ 2_95** | 108 | | 70.4% | 13 | 19 | 44 | 24 | 8 | 0 |
| **Event_3_95** | 83 | | 72.3% | 8 | 17 | 35 | 17 | 6 | 0 |
| **Event_7_95** | 49 | | 73.5% | 6 | 7 | 23 | 11 | 2 | 0 |
| **Event_10_95** | 36 | | 77.8% | 3 | 8 | 17 | 7 | 1 | 0 |
| **Event_2_99** | 27 | | 81.5% | 2 | 5 | 15 | 5 | 0 | 0 |
| **Event_3_99** | 17 | | 88.2% | 1 | 4 | 10 | 2 | 0 | 0 |
| **Event_7_99** | 11 | | 81.8% | 0 | 2 | 7 | 2 | 0 | 0 |
| **Event_10_99** | 10 | | 80.0% | 0 | 2 | 6 | 2 | 0 | 0 |
| **Event_2_99.5** | 11 | | 90.9% | 1 | 3 | 6 | 1 | 0 | 0 |
| **Event_3_99.5** | 9 | | 88.9% | 1 | 2 | 5 | 1 | 0 | 0 |
| **Event_7_99.5** | 6 | | 83.3% | 0 | 1 | 4 | 1 | 0 | 0 |
| **Event_10_99.5** | 6 | | 83.3% | 0 | 2 | 3 | 1 | 0 | 0 |

Table 8. Displaying the number (Count) of occurrences for each type of event from 1992-2013. Also shows the percentage of events that transpired before onset (if applicable) and before 1 July. Shows number of occurrences for each type of event by month.

| | Busy99_SW | | Busy95_SW | |
|---|---|---|---|---|
| **SW1** | 7 | 12.3% | 24 | 11.0% |
| **SW2** | 20 | 35.1% | 46 | 21.0% |
| **SW3** | 4 | 7.0% | 14 | 6.4% |
| **SW4** | 6 | 10.5% | 16 | 7.3% |
| **SW5** | 14 | 24.6% | 42 | 19.2% |
| **SW6S** | 39 | 68.4% | 120 | 54.8% |
| **SW7** | 6 | 10.5% | 22 | 10.0% |
| **SW8** | 23 | 40.4% | 70 | 32.0% |
| **SW9** | 12 | 21.1% | 25 | 11.4% |
| **SW10** | 8 | 14.0% | 27 | 12.3% |
| **SW11** | 9 | 15.8% | 18 | 8.2% |
| **SW12** | 28 | 49.1% | 61 | 27.9% |
| **SW13** | 11 | 19.3% | 31 | 14.2% |
| **SW14S** | 13 | 22.8% | 38 | 17.4% |
| **SW14N** | 13 | 22.8% | 31 | 14.2% |
| **SW6N** | 15 | 26.3% | 47 | 21.5% |

Table 9. Displays the number of times a PSA had at least one large wildfire when the Busy99_PSA or Busy95_PSA threshold was exceeded. Also shows the percentage representing the proportion for each PSA when having at least one large wildfire when the Busy99_SW or Busy95_SW threshold was exceeded.

**Chapter 4.**

# Using boosted regression trees to model and predict wildfires in the southwest United States

Nicholas J. Nauslar[12], Timothy J. Brown[1], and Peter J. Weisberg[3]

[1] Division of Atmospheric Sciences, Desert Research Institute, Reno, NV

[2]Cooperative Institute for Mesoscale Meteorological Studies/NOAA NWS Storm Prediction Center, Norman, OK

[3] Department of Natural Resources and Environmental Science, University of Nevada, Reno, Reno, NV

*Corresponding author address: Nicholas J. Nauslar, 120 David L. Boren Blvd., Suite 2100*

*Norman, OK 73072-7304*

*Email: nauslar@dri.edu*

Keywords: fire weather, wildfire drivers, large wildfire occurrence, fire management, lightning caused wildfires, human caused wildfires

**Abstract**

Understanding the drivers of and modeling wildfire occurrence, and especially large wildfire occurrence, can lead to predictive modeling in support of wildfire suppression and management. We implemented boosted regression trees (BRT) to model wildfire occurrence for wildfires of different types (i.e., lightning, human) across the Southwest Area (Arizona, New Mexico, west Texas, and Oklahoma panhandle) by sub-regions known as Predictive Service Areas (PSAs), which are used by operational fire meteorologists to predict wildfire potential. BRT models for all wildfires demonstrated relatively small mean and mean absolute errors and showed better predictability on days with wildfires. Cross-validated accuracy assessments for large wildfires demonstrated the ability to discriminate between large wildfire and non-large wildfire days. Measurements describing fuel conditions (100 and 1000-hour dead fuel moisture, energy release component) were the most important predictors when considering all wildfire types and sizes. However, a combination of fuels and atmospheric predictors (i.e., lightning, temperature) proved most predictive for large wildfire occurrence, and the number of relevant predictors increases for large wildfires indicating more conditions need to align to support large wildfires. These models can be implemented daily, providing guidance for wildfire occurrence by PSA and allowing wildfire management to make logistical decisions with suppression resources accordingly.

**Introduction**

Wildfires exist at the nexus of fuels, climate, weather, and topographic interactions. Wildfire occurrence necessitates an ignition in receptive and ample fuels under a combination of hot, dry, and windy conditions with sufficient slope. All of these elements work together to promote wildfire spread. Wildfire frequency increases in areas bounded by various climate extremes representing a complex ecological middle ground that develops a wildfire return interval, which consequently helps drive wildfire occurrence rates (Parisien and Moritz 2009). Wildfire occurrence is difficult to predict due to the complex interaction of fuels and atmospheric conditions on multiple temporal and spatial scales (Bessie and Johnson 1995; Westerling et al. 2003; Stephens 2005).

Wildfire regimes can be located along a gradient from "climate limited" to "fuel limited", which can lead to varying importance of wildfire drivers among different wildfire regimes (Littell et al. 2009). Fuel limited arid and semiarid deserts and rangelands necessitate continuous fuels to support fire, especially large fires (Abatzoglou and Kolden 2013). In the southwest United States (Arizona and New Mexico)(SWUS), area burned, especially in fuel limited dry shrub and grass-lands, depend strongly on fuel accumulation, which is modulated in part by antecedent wet periods that occur several months to more than a year earlier (Swetnam and Betancourt 1990; Westerling et al. 2003; Crimmins and Comrie 2005). However, wet antecedent conditions leading to increased fine fuel production do not always yield increased wildfire activity implying that antecedent climate cannot alone explain wildfire occurrence (Swetnam and Betancourt 1998; McKenzie et al. 2004; Morgan et al. 2008). Climate (i.e., Gedalof et al. 2005; Trouet et al. 2006; Morton et al. 2013) and weather (i.e., Flannigan and Harrington

1988; Bessie and Johnson 1995; Moritz et al. 2010) during the wildfire season play an important role for determining area burned (Abatzoglou and Kolden 2013).

Numerous studies have linked historical wildfire occurrence with weather, topography, and fuels (i.e., Heyerdahl et al. 2001, Rollins et al. 2002; Moritz 2003; Mermoz et al. 2005; Collins et al. 2007). Fuel availability (Westerling et al. 2003; Crimmins and Comrie 2005; Dennison et al. 2014), fuel moisture (Viegas et al. 1992; Meisner et al. 1993; Dennison and Moritz 2009), and weather, specifically hot, dry, and windy conditions (Parisien and Moritz 2009) or lightning (Fuquay et al. 1979; Latham and Schlieter 1989; Flannigan and Wotton 1991; Latham and Williams 2001; Anderson 2002) are all drivers of wildfire occurrence and in combination greatly influence wildfire ignition and spread potential (Wierzchowski et al. 2002; Moritz et al. 2005; Evett et al. 2008; Ordoñez et al. 2012). Ignitions are a limiting factor of wildfire occurrence, which make lightning prediction and human ignition patterns essential, and drivers between lightning and human caused fires can be distinct (Yang et al. 2007; Syphard et al. 2008; Aldersley et al. 2011; Argañaraz et al. 2015; Yang et al. 2015). Areas of increased human activity exhibit increased human caused wildfire occurrence rates and probabilities (Parisien et al. 2011; Parisien et al. 2012; Yang et al. 2015), but that relationship can be nonlinear where peak human caused wildfire activity is associated with intermediate human influence (Syphard 2007; Parisien et al. 2012). The effect of lightning on ignitions depends on coincident fuel conditions and atmospheric moisture (dewpoint, precipitation) (Wierzchowski et al. 2002; Evett et al. 2008; Parisien et al. 2011; Ordoñez et al. 2012; Parisien et al. 2012; Yang et al. 2015).

Wildfire occurrence varies greatly across the western United States (WUS) (Littell et al. 2009; Parisien and Moritz 2009; Finney et al. 2011) including across and within biomes (Schoennagel et al. 2004). The complexity of relationships among drivers of wildfire occurrence also varies greatly across the WUS (Hardy et al. 2001; Littell et al. 2009; Parisien et al. 2012). Therefore wildfire occurrence models tend to be more predictive when the study area is narrowed to a specific region (i.e., Littell et al. 2009; Stavros et al. 2014). Predictive Service Areas (PSAs) are sub regions within Geographical Area Coordination Centers (GACCs) and with similar historical wildfire occurrences, fuel types, fuel indices, and weather characteristics, and Predictive Services (PS) meteorologists quantify and assess wildfire potential by PSAs and wildfire management utilizes this information for decision-making (i.e., wildfire suppression resource allocation). Understanding and modeling wildfire occurrence especially large wildfire occurrence is important in the near term for wildfire suppression and management (Kolden and Brown 2010; Abatzolgou and Kolden 2011; Owen et al. 2012; Stavros et al. 2014) and long term due to climate change (Flannigan et al. 2009; Littell et al. 2010; Coumou and Rhamstorf 2012; Barbero et al. 2015). Other studies examined wildfire occurrence using areas bounded by eco-province or region (Littell et al. 2009; Barbero et al. 2015) and by GACC (Abatzolgou and Kolden 2013; Stavros 2014).

Wildfire suppression protects life and property across the United States costing more than $1 billion annually and continues to rise (NIFC 2015). With the wildland urban interface (WUI) increasing each year, strategies and tactics must reflect the rising priorities at risk during wildfires (Mell et al. 2010; USDA FS 2011). Strategies to suppress or to allow wildfires for resource benefit (previously wildland wildfire use,

monitor, etc.) change interannually and even during the duration of the wildfire (Pyne 2010). The effect on wildfire regimes due to increased ignitions and varying policies of wildfire suppression have affected wildfire frequency, intensity, and severity (Keeley et al. 1999; Syphard et al. 2007; Archibald et al. 2012; Argañaraz et al. 2015; Hantson et al. 2015). Wildfire drivers and probabilities are often unknown in many areas (Collins et al. 2010; Parisien et al. 2012), and this knowledge gap affects land management policies (Schoennagel and Nelson 2011) and allocation of wildfire suppression resources (Kolden and Brown 2010; Owen et al. 2012).

A fundamental approach to determining wildfire occurrence involves understanding how wildfire drivers interact spatially and temporally and how those interactions change under varying conditions (i.e., Argañaraz et al. 2015). Understanding these relationships bears importance due to a changing climate that could increase large wildfire occurrence (Barbero et al. 2015). Boosted regression trees (BRT) have been implemented to understand environmental drivers of wildfires (Parisien and Moritz 2009; Liu et al. 2013; Rodrigues and de la Riva 2014; Argañaraz et al. 2015) and to identify variable combinations and thresholds that optimally estimate the amount of burned area (Archibald et al. 2009; Addersly et al. 2011; Wu et al. 2014). Regression tree analysis provides predictive, dichotomous decision trees that split data iteratively into increasingly homogeneous groups and allows for nonlinear functional relationships (De'ath and Fabricius 2000; Weisberg et al. 2013). BRT models use boosting to combine hundreds to thousands of tree models to adaptively optimize predictive performance while providing relatively simple methods to discern relationships and contributions of each predictor (Elith et al. 2006; Leathwick et al. 2006, 2008). BRT models are able to select relevant

predictors, fit accurate functions, automatically identify and model interactions, and accurately predict various types of responses (Friedman 2002; Elith et al. 2008).

We implement BRT to identify important predictors and predictor interactions of wildfire occurrence including large wildfire occurrence for different wildfire types (i.e. lightning and human). We develop these BRT models to provide accurate and robust predictions of wildfire occurrence for each of the sixteen PSAs that comprise the Southwest GACC (SWCC) region of responsibility (Southwest Area (SWA)) (Figure 1). By building a BRT model for each wildfire type and size for every PSA, we note differences in environmental drivers and provide a predictive tool for wildfire occurrence on a resolvable scale that directly corresponds to operational forecast areas.

**Data**

Surface 4 km gridded daily precipitation, minimum and maximum relative humidity (RHmin; RHmax), specific humidity (SPH), minimum and maximum temperature (Tmin; Tmax), and wind velocity are utilized from the University of Idaho METDATA (Table 1)(Abatzoglou 2013). Vapor pressure deficit (VPD) was defined as the difference between saturated vapor pressure and ambient vapor pressure. Saturated vapor pressure and ambient vapor pressure were calculated using Tmin, Tmax, RHmin, and RHmax. The University of Idaho also archives gridded 4 km calculated fire danger indices of burning index (BI), energy release component (ERC), and 100-hour and 1000-hour dead fuel moisture (FM100; FM1000) calculated using fuel model G (Deeming et al. 1977; Cohen and Deeming 1985)(Table 1). All of the daily data were averaged across all grid points within each PSA from 1995-2013. The Southwest PS determined PSA

boundaries through a spatial analysis of wildfire activity, fuel type, and fuel index (i.e., ERC) similarities.

Wildfire data from 1995-2013 were obtained from the wildfire Program Analysis (FPA) quality controlled wildfire database (Short 2015). Wildfires were separated by type (lightning or human caused) and size (large or all wildfires) for each PSA. Large wildfire criteria were established from large wildfire thresholds determined by Southwest PS (Table 2). FPA wildfire occurrence data and proprietary National Lightning Detection Network® (NLDN) data were totaled daily and organized by PSA and across the entire SWA. Due to a major upgrade occurring after 1994, the NLDN dataset for this analysis only includes 1995 to 2013.

**Methods**

Data preparation is minimized when implementing BRTs since no predictor transformations are necessary and outliers and missing values are not problematic (De'ath 2007; Elith et al. 2008). Training and validation datasets should be separated to minimize bias towards the utilized data in the BRT model and so that accuracy is not overestimated (De'ath 2007; Parisien and Moritz 2009). However, De'ath (2007) advocates using all of the data to train a BRT model to identify relationships, find interactions, and make predictions on new data. Elith et al. (2008) recommends using different subsets of data for training and validation and cross-validation (CV) output to find the best parameters for a BRT model. The size and type of data set will impact the ability of CV output to accurately estimate the robustness of the BRT model (Elith et al. 2008). During the training process, highly correlated variables (i.e., $|r| > 0.8$) should be

removed and a distribution should be determined that best fits the training data (i.e., Bernoulli, Gaussian, Poisson) (Parisien and Moritz 2009; Argaranarz et al. 2015). One way to determine which highly correlated predictors to retain is to use all of the predictors in a BRT model and keep the predictors with the highest relative influence (Table 1). Predictor relative influence is determined by how often a predictor is used for splitting, which then is weighted by the squared improvement of the BRT model, and finally averaged over all of the trees (Elith et al. 2008).

The BRT process begins with fitting a regression tree that maximally reduces the loss function with subsequent iterations focusing on the variation of the response not yet captured by the model (Elith et al. 2008). For example, the second tree is fitted to the residuals of the first tree; then the model is updated with residuals from the third tree being fitted to the residuals of the two-tree model (Elith et al. 2008). The BRT process is stochastic, which improves predictive performance by reducing the final model's variance due in part to emphasizing the hardest observations to predict during the iterative process (Friedman 2001; Elith et al. 2008). The BRT process is stagewise, which leaves existing trees unaffected as the model continues. Only the fitted value of each observation in the training data is re-calculated for each new tree. The final BRT model is a linear combination of hundreds to thousands of trees and can be considered as a regression model with each term a tree (Elith et al. 2008). The trees could have very different predictors and splits, unless a random seed is initialized, a BRT model will be different yet similar each time the training process occurs (Elith et al. 2008).

Several adjustable parameters affect a BRT model's output. The learning rate determines the contribution of each tree, tree complexity controls interactions within each

tree, and both determine how many trees are fitted (Elith et al. 2008). The bagging fraction controls how much of the training data is randomly selected without replacement for each tree and values commonly used are between 0.5 and 0.75 (Elith et al. 2008). Overfitting often plagues BRT models, but adjusting the learning rate, tree complexity, and bagging fraction can ameliorate this issue (Hastie et al. 2001; Elith et al. 2008). Decreasing the learning rate and thus increasing the number of trees can specifically reduce overfitting. This may increase the bias, but it also drastically reduces the variance (De'ath 2007). Traditional regression methods control overfitting and improve prediction by reducing the number of terms, but the BRT process increases the number of terms and thus reducing their contribution using shrinkage (Friedman 2001). BRT also controls overfitting and improves prediction by introducing randomness via the bagging fraction, and despite the potential for overfitting, BRT regularly outperforms other regression techniques (i.e., GLM, GAM) when validated against independent data (Friedman 2002; Elith et al. 2008).

We implemented the following process for building and training a BRT model to predict wildfire occurrence in the SWUS: 1) use daily PSA values of fuels and atmospheric variables as predictors of lightning wildfires, human wildfires, all wildfires, large lightning wildfires, large human wildfires, and all large wildfires; 2) include all predictors in the training of a BRT model for a particular wildfire type in a PSA to decide which highly correlated predictors ($|r|>0.8$) were most influential and thus would be retained as a predictor; 3) systematically toggle learning rate (0.0001-0.05), tree complexity (3-7), bagging fraction (0.5-0.75), and tenfold CV to generate enough trees (> 1000) that minimized overfitting and demonstrated predictive power through CV output

(i.e., area under the curve (AUC), CV correlation); 4) implement tree complexity, learning rate, and bagging fraction values with guidance from previous steps to generate a BRT model using the entire dataset (AllYears method) to identify important predictors, relevant interactions, and the best predictive model for each wildfire type in a particular PSA to test on new data; and 5) training a BRT model on 18 of 19 years of data, repeating this process 19 times to leave each year out for validation (LeaveOne method). In total, there are six BRT models (all lightning caused wildfires, all human caused wildfires, all wildfires, lightning caused large wildfires, human caused large wildfires, and all large wildfires) for each PSA. Mean error (ME) and mean absolute error (MAE) are calculated for steps four and five for all days and only days with wildfires to provide another evaluation metric of the predictive performance for the BRT models.

To model the number of wildfires on a given day in a particular PSA, a Poisson distribution was utilized due to its ability to handle many zeros. Due to the infrequency of large wildfires, a Bernoulli distribution was utilized for modeling the probability distribution of this response variable. Since the Bernoulli distribution requires a binary response, all daily totaled large wildfires were reduced to 0 and 1 with double weight given to days with multiple large wildfires. All of the BRT modeling was done in R (gbm and dismo packages; Elith and Leathwick 2015).

## Results

### *All Fires*

Table 3 displays mean number of daily wildfires, MAE, and ME from the BRT model output for each PSA predicting the number of lightning and human caused

wildfires using all days. No ME in Table 3 has a larger magnitude than 0.021 for any wildfire type and all MEs are negative. Human caused wildfires exhibit the largest MAE relative to the daily wildfires mean, and across all wildfire types, the MAE is relatively larger to the daily wildfires mean when the daily wildfires mean is smaller (Table 3). Table 4 displays the same information as Table 3 but only considers days when at least one wildfire occurred. Similar to Table 3, MEs across all wildfire types and PSAs are negative including some MEs with identical magnitudes as their corresponding MAEs (Table 4). Relative to its mean daily wildfire occurrence, all wildfires have the smallest magnitude of MAE and ME while human caused wildfires have the largest MAE and ME (Table 4). MAEs for fire days (Table 4) are smaller than MAEs for all days (Table 3) when comparing to their respective means of daily wildfires, but the converse is true when examining MEs.

Tables 5-7 display the mean number of daily wildfires, MAEs, and MEs for each wildfire type across all PSAs using the AllYears and LeaveOne methods. Overall, the MAEs in Tables 5-7 are slightly larger than the MAEs in Tables 3-4, but are relatively close in magnitude (i.e., 0.887 to 0.927). The MEs have similar magnitudes when comparing the BRT models using the AllYears and LeaveOne methods, but MEs are positive when using all days for the LeaveOne method (Tables 5-7) and negative when using the AllYears method (Tables 3-4). MEs from the LeaveOne method during wildfire days are negative, which is similar to the MEs from using the AllYears method, but they are slightly larger in magnitude.

Lightning strikes (LS) and fuels (FM1000, FM100, ERC) are the most important predictors for all lightning caused wildfires when considering the median relative

influence and the number of times each predictor is used in a PSA's BRT all lightning caused wildfires model (Table 8). LS and a fuels variable combine to exceed more than half of the relative influence across the PSAs for all lightning caused wildfires. Only PSAs 4 and 11 have relative influence for LS below 30%, and six PSAs do not have LS as its leading predictor with FM1000 (three times), FM100 (once), ERC (once), and Tmin (once) as the leading predictors in those PSAs (not shown). Temperature (Tmax, Tmin) appears to be a secondary predictor with median relative influence of 12-14% and Tmax or Tmin is in each BRT model for all lightning caused wildfires (Table 8). The rest of the variables have median relative influences below 7% and VPD is not utilized in any BRT model due to its correlation to several variables and exhibiting less relative influence to these correlated variables (Table 8; Table 1). Interactions were most frequent and strongest with LS for all lightning fires (Table 9). The most frequent interactions paired LS with a fuels variable (FM1000, ERC, BI) with break points located at lower (higher) FM1000 (ERC, BI) and higher LS although the fitted values did flatten after a certain LS threshold (Figure 2).

Fuels (FM1000, FM100, ERC) are the most important predictor for all human caused wildfires (Table 8). ERC accounts for a median relative influence of 43.7% the six times it is utilized in a BRT all human fires model, and FM100 and FM1000, with median relative influences of 32.1% and 18.4% respectively, are each utilized in six BRT all human caused fires models including in two of the same models. DOY also proved important with median relative influence of 13.2% and several remaining variables had median relative influence 7-11% (Tmax, Tmin, RHmax, RHmin, SPH, WndSpd, BI) (Table 8). LS, precipitation, and VPD are the least important although VPD did have a

relative influence of 21.4% the one time it was utilized. PSAs clustered along three main themes for important predictors: DOY and fuels (5, 6S, 6N, 7, 8, 9, 12), fuels (PSAs 1, 2, 4, 13, 14S, 14N), and fuels and RH/VPD (PSAs 3, 10). Interactions were slightly less frequent and there were fewer interactions between the same two variables than for all lightning caused fires (Table 10). DOY, fuels (ERC, FM100, FM1000) and wildfire spread (BI, WndSpd) variables were paired most frequently (Table 10)(Figure 3).

FM1000/ERC (24.9%; 13 counts/29.9%; 3 counts) and LS (21.9%; 16 counts) were the most important predictors for the all fires BRT models with the highest relative importance followed by Tmax/Tmin (14.8%; 9 counts/10.9%; 7 counts) (Table 8). Similar to the lightning and human all fires models, the remaining predictors contributed much less (3.8%-8.7%) than the most important predictors. Nine of the PSAs (PSAs 1, 2, 5, 6S, 7, 8, 10, 12, 14S) demonstrated lightning and fuels as the most important predictors with the remaining PSAs clustering along fuels and temperature (PSAs 4, 11), fuels (PSAs 13, 14N), or fuels and DOY (PSAs 6N, 9) (not shown). FM1000 and LS is the most frequent interaction for the all fires models, and DOY and LS are paired with nearly every variable at least once (Table 11)(Figure 4). LS, FM1000, DOY, and temperature (Tmax/Tmin) are represented in the most frequent interactions.

*Large Fires*

Table 12 displays the CV AUC for each BRT model including the median CV AUC from the BRT models using the LeaveOne method. All CV AUC values are larger than 0.7 except for PSA 2 human caused large wildfires, PSA 7 all large wildfires, and PSA 11 large human wildfires BRT models. The PSA 2 lightning caused large wildfires

BRT model has a CV AUC value exceeding 0.9 using both methods. The median CV AUC values for both methods and each large wildfire type are greater than 0.75 with the CV AUC from the large lightning wildfires BRT model slightly exceeding the CV AUC from the human caused large wildfires BRT model (0.797 to 0.789). The CV AUC values across all large wildfire types for the LeaveOne method are very similar to the CV AUC values using the AllYears method.

Table 13 shows the probabilities associated with large wildfire days, days without large wildfires, and daily occurrence of large wildfires by type and PSA using AllYears method. Probabilities on days with large wildfires are much higher than days without large wildfires (Table 13). The differences in median probabilities of large wildfire and non-large wildfire days are 0.161, 0.182, and 0.121 for all large wildfires, lightning caused large wildfires, and human caused large wildfires respectively. Additionally, the probabilities on days without large wildfires are slightly lower than the mean daily probabilities for large wildfires. Table 14 shows the probabilities associated with large wildfire days and days without large wildfires using the LeaveOne method. The probabilities are lower and closer in magnitude for large wildfire and non-large wildfire days, and remain higher for large wildfire days compared to non-large wildfire days with a much smaller difference.

Temperature (Tmax, Tmin), LS, and FM100/FM1000 are the most important predictors for lightning caused large wildfires (Table 15). The remaining predictors have a median relative influence between 6.53% and 8.81%. The PSAs form roughly three separate groups for predictors of lightning caused large wildfires (in descending importance): temperature, lightning, and fuels (PSAs 1, 4, 6S, 6N, 7, 8, 14S), lightning,

SPH/precipitation, and fuels (PSAs 2, 3, 5, 12, 13). The remaining four PSAs represent a transition between the two main groups (i.e., PSA 9) or are comparatively unique with its predictors for lightning caused large wildfires. Tmax is involved with the four most frequent interactions for lightning caused large wildfires, but LS is still involved with the strongest interactions especially with FM1000/FM100 and BI (Table 16; Figure 5). BI also interacts frequently and its interactions often generally depict how most thresholds for interactions are at the upper or lower bounds of a particular interactive pair with a sharp increase of fitted values.

Fuels (FM100, FM1000, ERC, BI) and RH are the most important predictors (counts and median relative influence) for human caused large fires (Table 15). BI represents a combination of fuel and spread components and could be classified as either, and is more sensitive to changes in weather than other fuels variables in this analysis (Cohen and Deeming 1985). Precipitation, WndSpd, and Tmax also show some predictive potential for human caused large wildfires. The PSAs do not form larger groups as they did for lightning caused large fires, but several smaller groups consisting of two to four PSAs. These groups have either one dominant predictor (fuels or spread) or some sort of combination of the three main predictors (fuels and RH) (Figure 6). Many of the interactions for human caused large wildfires are very small and occur across narrow margins between the interactive pairs (Table 17; Figure 6). BI is included with the most frequent interactions and the few stronger interactions within the human caused large wildfires analysis.

Fuels and temperature are the two most important predictors for all large wildfires with LS and RHmin trailing closely (Table 15). It appears the blending of the most

important predictors from large lightning and human wildfires occurs with most PSAs. Two groups appear among the PSAs with temperature (PSAs 1, 4, 5, 6N, 7, 8) and fuels (PSAs 6S, 9, 10, 12, 13, 14S, 14N) as the main predictors with a smaller group that emphasizes lightning over fuels and temperature (PSAs 2, 3)(not shown). However, the secondary predictors are important too, and usually represent at least one of the other main predictors. The number of interactions is spread across more pairs for all large fires than human or lightning large wildfires (Tables 16-18). The strongest interactions involve fuels and either a spread (i.e., BI) or ignition (i.e., LS) component (Figure 7).

**Discussion**

The BRT models for predicting the number of wildfires demonstrated predictive power and robustness with relatively small MAE and MEs, however, these BRT models did consistently under forecast (negative ME) the number of daily wildfires (Tables 3-7). During days when wildfires occurred, lower MAEs were observed, but the magnitude of MEs increased respective to the mean (Tables 4-7). BRT models performed as well and consistently better on days with wildfires demonstrating predictability during days when they are needed most (Tables 4-7, 13-14). BRT models predicted each wildfire type reasonably well and their respective MAEs and MEs were relatively similar indicating that an all wildfires model would be sufficient. We examined rounding all model output to the nearest integer or just rounding values below one to the nearest integer for four different PSAs (PSAs 2, 5, 8, 11). Due to the amount of days without fires, this reduced the ME since all BRT model output was greater than zero, but the rounding increased MAE for the three PSAs examined.

BRT models for large wildfires had good CV AUC values (mostly 0.7-0.8), generally demonstrating the ability to discriminate between large wildfire and non-large wildfire days across all wildfire types indicating that an all large wildfires model would be sufficient (Table 12). However, attempts were made to determine a probability threshold for large wildfires. Some of these methods included using the mean probability of a large wildfire plus twice the standard deviation of large wildfires and using various thresholds for specificity and sensitivity (Freeman and Moisen 2008). When these methods were implemented using the AllYears method, some success was found. However, when applying the same thresholds to the LeaveOne method, the success rate precipitously dropped demonstrating overfitting and a lack of robustness. Wildfire management and operational forecasters would rather not miss days with large wildfires and have a higher false alarm rate. However, most thresholds using the LeaveOne method consistently had nearly as many or more days without large fires as days with large wildfires above the specified large wildfire threshold.

Fuels (FM1000, FM100, ERC) are the most important predictor(s) when considering lightning, human, or both types of wildfires confirming prior findings (i.e., Parisien et al. 2011; Riley et al. 2013; Stavros et al. 2014; Barbero et al. 2015). LS is also very important and the leading predictor for lightning wildfires and second most important predictor for all wildfires. Human fires are mostly dependent on fuels, and DOY is the leading non-fuels predictor. When examining interaction plots with DOY, a noticeable decrease occurs near the beginning of July (Figures 3b, 4b, 8), which coincides with the mean North American Monsoon (NAM) (Douglas et al. 1993; Adams and Comrie 1997; Barlow et al. 1998; Higgins et al. 1999) onset (Mohrle 2003).

More predictors become important when examining large wildfires, and the relative influence of predictors can change when comparing all and large wildfires of a particular wildfire type. A combination of fuels (i.e. FM1000, ERC) and atmospheric predictors (i.e., LS, Tmax) drive large wildfire occurrence (Abatzoglou and Kolden 2013; Riley et al. 2013; Stavros et al. 2014). The increase in the number of predictors and the smoothing of relative influence among those predictors for large wildfires compared to all wildfires indicates more conditions need to align to support and sustain large wildfires (Tables 8, 15). For example, lightning and fuels are overwhelmingly the two most important predictors across the PSAs for all lightning fires, but for lightning caused large wildfires, temperature also becomes very important along with atmospheric moisture predictors (i.e. SPH, precipitation) (Tables 8, 15). For all human caused fires, fuels and DOY are the most important predictors, but RHmin and BI or WndSpd increase in importance for human caused large wildfires (Tables 8, 15). While dry fuels and/or LS are usually sufficient to drive wildfire occurrence, predictors that promote wildfire growth via atmospheric dryness (i.e., RHmin), instability (i.e., Tmax), and spread (i.e., BI, WndSpd) become more necessary for large wildfires. Fuels variables (FM1000, FM100, ERC, BI) take into account previous and current atmospheric conditions (Cohen and Deeming 1985; Riley et al. 2013) and thus can be viewed as further evidence of the importance of weather and short-term climate (days to weeks) variations being important for wildfire occurrence, especially large wildfire occurrence. The importance of these fuels variables also indicate the importance of dead fuel moisture, and dead fuels are the primary carriers of surface fires and help fires transition into crown fires (Van Wagner

1977). This transition from a surface fire to an accompanying crown fire increases the likelihood of larger, more intense wildfires.

Predictor interactions exhibit thresholds that drive wildfire occurrence. Wildfire occurrence increases as LS and BI increase, but usually peaks somewhere between 2000 and 4000 LS and greater than a BI of 75-100 before leveling off (i.e., Figures 2d, 7a). As FM100/FM1000 decreases, wildfire occurrence increases initially below 10% and especially below 5% (Figures 2a-b, 3b, 4a, 5c-d, 8). The co-varying of the interactive pair of predictors illuminates multiple important critical fire weather conditions including ignition and spread (i.e., LS and BI), dry thunderstorms (i.e., LS and precipitation)(Figure 9)(Nauslar et al. 2013) and a combination of hot, dry, and windy conditions (i.e., Tmax, RHmin, WndSpd)(Parisien and Moritz 2009) (Tables 9-11, 16-18).

The NAM modulates most of the weather conditions during the SWA wildfire season (Mohrle 2003; Evett et al. 2008), and the NAM signal is apparent when examining DOY interactions with atmospheric and fuels predictors (Figures 3b, 4b, 8). A precipitous drop of fitted values occurs around and after DOY 100 (9 July), which is around the first week of July. Once the NAM onset begins, atmospheric moisture begins to stifle wildfire occurrence, especially large wildfire occurrence. A noticeable peak of fitted values occurs in many of these DOY interactions during the days and weeks leading up to the climatological NAM onset DOY (Figures 3b, 4b, 8). This timing coincides with peak fuel dryness helping produce the most active part of the wildfire season. The term ‘Firesoon’ has been used to describe the time period just before the NAM onset noting the peak fuel dryness coinciding with the beginning surges of NAM moisture triggering lightning ignited wildfires (Brown 2002). Some interactions involving large wildfires consistently

show multiple smaller, narrower thresholds of increased fitted values in the same interaction, which we attribute to overfitting due to the lack wildfire occurrence in that particular PSA.

PSAs naturally clustered when examining wildfire type and size, but those clusters were not consistent across wildfire types or sizes thwarting attempts to group PSAs across all wildfire types and sizes. PSAs would generally cluster along topographic features (i.e., Mogollon Rim), elevation, and location (both latitude and longitude), all of which also affect climate and subsequently fuel type, continuity, and loading. PSAs also tended to cluster along eco-provinces or regions (Littell et al. 2009; Barbero et al. 2015). Usually two main clusters of PSAs would emerge with similar important predictors and interactions with the remaining PSAs combining to form one or two more unique, smaller groups or representing a transition between the two main groups.

Weather and short-term climate drive wildfire occurrence across the SWUS with low fuel moisture (Viegas et al. 1992; Meisner et al. 1993; Dennison and Moritz 2009; Riley et al. 2013), lightning (Fuquay et al. 1979; Latham and Schlieter 1989; Flannigan and Wotton 1991; Latham and Williams 2001; Anderson 2002), and hot, dry, and/or windy atmospheric conditions (Parisien and Moritz 2009) specifically driving wildfire occurrence especially large wildfires (Wierzchowski et al. 2002; Moritz et al. 2005; Evett et al. 2008; Ordoñez et al. 2012). A component missing from this analysis is accurately quantifying fine fuel loading, which is important for wildfire activity and area burned in the SWUS (Westerling et al. 2003). Antecedent precipitation has been shown to modulate fine fuel loading, but that signal is not consistent particularly in the SWUS (Swetnam and Beatencourt 1998). Using the discovery date confines the conditions examined to the date

of discovery, which may not coincide with large wildfire growth. Holdover wildfires also pose a problem since the discovery date most likely does not coincide with its actual ignition date. Stavros et al. (2014) examined conditions during weeks pre and post discovery, which were included in the model for very large wildfires for the SWCC. We did not using any time-lagged variables, deciding to focus on the conditions during the discovery date while capturing antecedent conditions inherently in some of the variables utilized (i.e., ERC).

**Conclusions**

We analyzed wildfire occurrence by type (lightning, human, all) and size (all, large) for each PSA resulting in six distinct BRT models for each PSA. We were able to examine different drivers of wildfire by type, size, and location while recognizing similarities and differences and develop robust, predictive BRT models across the various categories. The BRT analysis of wildfire activity in the SWA produced several key findings:

- Median CV AUC values between 0.75 and 0.8 demonstrate the ability of BRT models to distinguish between large and non large wildfire days
- Relatively small ME and MAE for all wildfires (all types) from BRT output including improved performance (smaller MAE) on days with wildfires
- Fuels (FM1000, FM100, ERC) are the most important predictor(s) with when considering all wildfires regardless of type, although the number lightning strikes also proved important

- More predictors become important when examining large wildfires with a combination of fuels (i.e. FM1000, ERC) and atmospheric predictors (i.e., LS, Tmax) driving large wildfire occurrence indicating more conditions need to align to support and sustain large wildfires
- Interactions demonstrate multivariate thresholds for wildfire occurrence (i.e., LS and FM1000) and important critical fire weather conditions including ignition and spread (i.e., LS and BI), dry thunderstorms (i.e., LS and precipitation), and a combination of hot, dry, and windy conditions (i.e., Tmax, RHmin, WndSpd)
- The NAM signal is apparent in interaction plots of DOY and atmospheric (i.e., LS) or fuels (i.e., FM100) variables

We modeled wildfire occurrence by PSA mirroring the same forecast areas that operational PS meteorologists forecast with. These models can be directly applied to the SWA PSAs and implemented daily with output providing guidance for number of wildfires and the probability of a large wildfire. This allows wildfire management to identify more specific areas of increased wildfire potential and make logistical decisions with suppression resources accordingly. Differences emerged among the PSAs across the various wildfire types and sizes with groups of PSAs clustering together. However, these groups of PSAs were inconsistent when examining different wildfire types and sizes thus reinforcing the decision to split the SWA into PSAs and model by wildfire type and size.

The BRT models for both large and number of wildfires show predictive potential and demonstrate robustness across all wildfire types and PSAs. The large wildfires LeaveOne method results give pause due to the lower probabilities and poor performance

using any of the large wildfire probability thresholds. However, important predictors and interactions are identified that drive wildfire occurrence for different wildfire types and sizes across the SWA. Additionally, important thresholds are contained within the results that will help better understand the confluence of conditions necessary for wildfire ignition and spread.

*Acknowledgements*

We would like to thank the USDA Forest Service for providing the funding (10-CS-11130206-047), and Vaisala for allowing the use of lightning data for wildfire applications. We would also like to thank Chuck Maxwell, Southwest PS Program Manager, for his input and preliminary analysis of the research, and John Abatzolgou for helping obtain and organize the METDATA. Finally, we would like to thank Patrick Marsh, Matthew Fearon, and Andrew Joros for their programming help.

**References**

Abatzoglou JT, Brown TJ (2009) Influence of the Madden–Julian Oscillation on summertime cloud-to-ground lightning activity over the Continental United States. *Mon. Wea. Rev.* **137,** 3596-3601.

Abatzoglou JT (2013) Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.* **33**, 121-131.

Abatzoglou JT, Kolden CA (2013) Relationships between climate and macroscale area burned in the western United States. *Int. J. Wildland Fire* **22**, 1003–1020.

Aldersley A, Murray SJ, Cornell SE (2011) Global and regional analysis of climate and human drivers of wildfire. *Sci. Total Environ*. **409**, 3472–3481.

Anderson K (2002) A model to predict lightning-caused fire occurrences. *Int. J. Wildland Fire*, **11**, 163-172.

Archibald S, Roy DP, Van Wilgen BW and Scholes RJ (2009) What limits wildfire? An examination of drivers of burnt area in Southern Africa. *Global Change Biology* **15**, 613–630.

Archibald S, Staver AC, Levin SA (2012) Evolution of human-driven fire regimes in Africa. *Proc. Natl. Acad. Sci.* **109**, 847–852.

Argañaraz JP, Pizarro GG, Zak M, Landi MA, Bellis LM (2015) Human and biophysical drivers of wildfires in Semiarid Chaco mountains of Central Argentina, *Sci. Total Environ.* **520**, 1-12.

Barbero R, Abatzoglou JT, Larkin NK, Kolden CA, Stocks B (2015) Climate change presents increased potential for very large wildfires in the contiguous United States. *Int. J. Wildland Fire* **24**, 892-899.

Bessie WC, Johnson EA (1995) The relative importance of fuels and weather on fire behavior in sub-alpine forests. *Ecology* **76**, 747–762.

Brown TJ (2002) The North American "Firesoon". In report on research opportunities for climate and society interactions in the North American Monsoon region, Workshop on Applications and Human Dimensions of Monsoon Research, 18-20 June 2001, Tucson, AZ, A.J Ray and R.S. Webb (Eds.), Report for NOAA OAR, 64 pp.

Climate Prediction Center, 2014: CPC - Climate Weather Linkage: Madden-Julian Oscillation. http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/mjo.shtml

Cohen JD, Deeming JE (1985) The National Fire Danger Rating System: Basic equations. USDA Forest Service, Pacific Southwest Forest and Range Experiment Station, GTR PSW-82, 16 pp.

Collins BM, Kelly NM, van Wagtendonk JW, and SL Stephens (2007) Spatial patterns of large natural wildfires in Sierra Nevada wilderness areas. *Landscape Ecology* **22**, 545–557.

Collins BM, Stephens SL, Moghaddas JJ, Battles J (2010) Challenges and approaches in planning fuel treatments across fire-excluded forested landscapes. *Journal of Forestry* **108**, 24–31.

Coumou D, Rahmstorf S (2012) A decade of weather extremes. *Nature Climate Change* **2**, 491–496.

Crimmins MC, Comrie AC (2005) Interactions between antecedent climate and wildfire variability across southeastern Arizona. *Int. J. Wildland Fire* **13 (4)**, 455-466.

De'ath G (2007) Boosted trees for ecological modeling and prediction. *Ecology* **88**, 243–251.

Deeming JE, Burgan RE and Cohen JD (1977) The national fire- danger rating system—1978 (Ogden, UT: USDA Forest Service Intermountain Forest and Range Experiment station) Gen. Tech. Rep. INT-39

Dennison PE, Brewer SC, Arnold JD and Moritz MA (2014) Large wildfire trends in the western United States, 1984-2011. *Geophys. Res. Lett.* **41**, 2928–33.

Dennison PE and Moritz MA (2009) Critical live fuel moisture in chaparral ecosystems: a threshold for fire activity and its relationship to antecedent precipitation. *Int. J. Wildland Fire* **18**, 1021–27.

Dissing D, Verbyla DL (2003) Spatial patterns of lightning strikes in interior Alaska and their relations to elevation and vegetation. *Can. J. For. Res*. **33**, 770–782.

Earth System Research Laboratory Physical Sciences Division (ESRL PSD), 2014. Multivariate ENSO Index (MEI). http://www.esrl.noaa.gov/psd/enso/mei/index.html

Elith J, Graham CH, Anderson RP, et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151.

Elith J, Leathwick JR, Hastie T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology* **77**, 802–813.

Elith J, Leathwick JR (2015) Boosted regression trees for ecological modeling. https://cran.r-project.org/web/packages/dismo/vignettes/brt.pdf

Evett RR, Mohrle CR, Hall BL, Brown TJ, Stephens SL (2008) The effect of monsoonal atmospheric moisture on lightning wildfire ignitions in southwestern North America. *Agric. Forest Meteorol.* **148***,* 1478-1487.

Finney MA, McHugh CW, Grenfell IC, Riley KL, Short KC (2011) A simulation of probabilistic wildfire risk components for the continental United States. *Stochastic Environmental Research and Risk Assessment* **25**, 973–1000.

Flannigan MD, Harrington JB (1988) A study of the relation of meteorological variables to monthly provincial area burned by wildfire in Canada (1953–80). *Journal of Applied Meteorology* **27**, 441–452.

Flannigan MD, Wotton BM (1990) Lightning-ignited forest wildfires in northwestern Ontario. *Can. J. For. Res*. **21**, 277-287.

Freeman EA, Moisen GG (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling* **217**, 48-58

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Annals of Statistics* **29**, 1189–1232.

Friedman JH (2002) Stochastic gradient boosting. *Computational Statistics and Data Analysis* **38**, 367–378.

Fuquay DM, Baughman RG, Latham DJ (1979) A model for predicting lightning wildfire ignition in wildland fuels. *USDA Forest Service Research Paper* INT-217, Ogden, UT.

Gedalof Z, Peterson D, Mantua N (2005) Atmospheric, climatic, and ecological controls on extreme wildfire years in the northwestern United States. *Ecological Applications* **15**, 154–174.

Hantson S, Pueyo S, Chuvieco E (2015) Global wildfire size distribution is driven by human impact and climate. *Glob. Ecol. Biogeogr.* **24**, 77–86.

Hardy CC, Schmidt KM, Menakis JP, Sampson RN (2001) Spatial data for national fire planning and fuel management. *International Journal of Wildland Fire* **10**, 353–372.

Hastie T, Tibshirani R, Friedman JH (2001) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York.

Heyerdahl EK, Braubaker LB, Agee JK, (2001) Spatial controls of historical wildfire regimes: a multiscale example from the interior west, USA. *Ecology* **82**, 660–678.

Keeley JE, Fotheringham CJ, Morais M (1999) Reexamining wildfire suppression impacts on brushland wildfire regimes. *Science* **284**, 1829–1832.

Kolden CA, Brown TJ (2010) Beyond wildfire: perspectives of climate, managed fire and policy in the USA. *International Journal of Wildland Fire* **19**(3), 364–373.

Latham DJ, Schlieter JA (1989) Ignition Probabilities of Wildland Fuels Based on Simulated Lightning Discharges. *USDA For. Serv. Intermt. Res. Stn. Res. Pap*. INT-411. Ogden, UT.

Latham D, Williams E (2001) Lightning and forest wildfires. Chapter 11, Forest wildfires: Behavior and ecological effects. *Academic Press*, 375- 418.

Leathwick JR, Elith J, Francis MP, Hastie T, Taylor P (2006) Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecology Progress Series*, **321**, 267–281.

Leathwick JR, Elith J, Chadderton WL, Rowe D, Hastie T (2008) Dispersal, disturbance, and the contrasting biogeographies of New Zealand's diadromous and non-diadromous fish species. *Journal of Biogeography* **35**, 1481–1497.

Liu Z, Yang J, He HS (2013) Identifying the Threshold of dominant controls on wildfire spread in a boreal forest landscape of northeast China. PLoS ONE 8(1): e55618.

Littell JS, McKenzie D, Peterson DL, Westerling AL (2009) Climate and wildfire area burned in western US ecoprovinces, 1916–2003. *Ecological Applications* **19**, 1003–1021.

Littell JS, Oneil EE, McKenzie D, Hicke JA, Lutz JA, Norheim RA, Elsner MM (2010) Forest ecosystems, disturbance, and climatic change in Washington State, USA. *Climatic Change* **102**, 129–158.

McKenzie D, Gedalof ZM, Peterson DL, Mote P (2004) Climatic change, wildfire, and conservation. *Conservation Biology* **18**, 890–902.

Meisner BN, Chase RA, McCutchan MH, Mees R, Benoit JW, Ly B, Albright D, Strauss D, Ferryman T (1993) A lightning wildfire ignition assessment model. In 12th Conference on wildfire and Forest Meteorology. Society of American Foresters, Jekyll Island, GA. pp. 172–178.

Mell WE, Manzello SL, Maranghides A, Butry D, Rehm RG (2010) The wildland-urban interface wildfire problem - current approaches and research needs, *Int. J. Wildland Fire* **19**, 238-251.

Mermoz M, Kitzberger T, Veblen TT (2005) Landscape influences on occurrence and spread of wildfires in Patagonian forests and shrublands. *Ecology* **86**, 2705–2715.

Mohrle CR (2003) The Southwest Monsoon and relation to fire occurence. M.S. thesis, Department of Physics/Division of Atmospheric Sciences (DRI), University of Nevada, Reno, 97 pp. Online at http://www.cefa.dri.edu/Publications/charliethesis.pdf.

Morgan P, Heyerdahl EK, Gibson CE (2008) Multi-season climate syn- chronized forest fires throughout the 20th century, northern Rockies, USA. *Ecology* **89**, 717–728.

Moritz MA (2003) Spatiotemporal analysis of controls on shrubland wildfire regimes: age dependency and wildfire hazard. *Ecology* **84**, 351–361.

Moritz MA, Moody TJ, Krawchuk MA, Hughes M, Hall A (2010) Spatial variation in extreme winds predicts large wildfire locations in chaparral ecosystems. *Geophysical Research Letters* **37**, L04801.

Morton DC, Collatz GJ, Wang D, Randerson JT, Giglio L, Chen Y (2013) Satellite-based assessment of climate controls on US burned area. *Biogeosciences* **10**, 247–260.

Ordóñez C, Saavedra A, Rodríguez-Pérez JR, Castedo-Dorado F, Covián E (2012) Using model-based geostatistics to predict lightning-caused wildfires, *Environmental Modelling & Software* **29**, 44-50.

Owen G, McLeod JD, Kolden CA, Ferguson DB, Brown TJ (2012) Wildfire management and forecasting fire potential: the roles of climate information and social networks in the Southwest US. *Weather, Climate, and Society* **4**, 90–102.

National Interagency Fire Center (NIFC) (2015) Federal firefighting costs (suppression only. https://www.nifc.gov/fireInfo/fireInfo_documents/SuppCosts.pdf

Nauslar NJ, Kaplan ML, Wallmann J, Brown TJ (2013) A forecast procedure for dry thunderstorms. *J. Operational Meteor.* **1**, 200–214.

Parisien MA, Moritz MA (2009) Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecological Monographs* **79**, 127-154.

Parisien M-A, Parks SA, Miller C, Krawchuk MA, Heathcott M, Moritz MA (2011) Contributions of ignitions, fuels, and weather to the spatial patterns of burn probability of a boreal landscape. *Ecosystems* **14**, 1141–1155.

Parisien M-A, Snetsinger S, Greenberg JA, Nelson CR, Schoennagel T, Dobrowski SZ, Moritz MA (2012) Spatial variability in wildfire probability across the western United States. *International Journal of Wildland Fire* **21**, 313–327.

Pyne SJ (2010) 'America's wildfires. A Historical Context for Policy and Practice.' (Durham, NC: Forest History Society, 2010). 93 pp.

Riley KL, Abatzoglou JT, Grenfell IC, Klene AE, Heinsch FA (2013) The relationship of large fire occurrence with drought and fire danger indices in the western USA, 1984–2008: the role of temporal scale. *Int. J. Wildland Fire* **2013**, *22*, 894–909

Rodrigues M, de la Riva J (2014) An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environmental Modelling & Software* **57**, 192-201.

Rollins MG, Morgan P, Swetnam T (2002) Landscape scale controls over 20th century wildfire occurrence in two large Rocky Mountain (USA) wilderness areas. *Landscape Ecology* **17**, 539–557.

Schoennagel T, Veblen TT, Romme WH (2004) The interaction of fire, fuels and climate across Rocky Mountain forests. *Bioscience* **54**, 661–676.

Schoennagel T, Nelson CR (2011) Restoration relevance of National Fire Plan treatments across the western US. *Frontiers in Ecology and the Environment* **9**, 271–277.

Short KC (2015) Spatial wildfire occurrence data for the United States, 1992-2013 [FPA_FOD_20150323]. 3rd Edition. Fort Collins, CO: Forest Service Research Data Archive.

Stavros NE, Abatzoglou JT, Larkin NK, McKenzie D, Steel EA (2014) Climate and very large wildland fires in the contiguous western USA. *Int. J. Wildland Fire* **23**, 899–914.

Stephens SL (2005) Forest fire causes and extent on United States Forest Service lands. *Int. J. of Wildland Fire* **14**, 213–222.

Swetnam TW, Betancourt JL (1990) Wildfire–Southern Oscillation relations in the southwestern United States. *Science* **249,** 1017–1020.

——, —— (1998) Mesoscale disturbance and ecologi- cal response to decadal climatic variability in the American southwest. *J. Climate* **11,** 3128–3147.

Syphard AD, Radeloff VC, Keeley JE, Hawbaker TJ, Clayton MK, Stewart SI, Hammer RB, (2007) Human influence on California wildfire regimes. *Ecol. Appl.* **17**, 1388–1402.

Syphard AD, Radeloff VC, Keuler NS, Taylor RS, Hawbaker TJ, Stewart SI, Clayton MK (2008) Predicting spatial patterns of fire on a southern California landscape. *Int. J. Wildland Fire* **17**, 602–613.

Trouet V, Taylor AH, Carleton AM, Skinner CN (2006) Fire–climate interactions in forests of the American Pacific coast. *Geophysical Research Letters* **33**, L18704.

USDA Forest Service (2011) A national cohesive wildland fire management strategy. USDA Forest Service, Wildland Fire Leadership Council, Cohesive Strategy Oversight Committee, http://www.fs.fed.us/rm/pubs_other/rmrs_2011_usda_fs001.pdf.

Van Wagner CE (1977) Conditions for the start and spread of crown fire. *Canadian Journal of Forest Research* **7**, 23–34.

Viegas DX, Viegas MTSP, Ferreira AD (1992) Moisture content of fine forest fuels and wildfire occurrence in central Portugal. *Int. J. Wildland Fire* **2**, *69–86*.

Weisberg PJ, Shandra O, Becker ME (2013)Landscape influences on recent timberline shifts in the Carpathian Mountains: Abiotic influences modulate effects of land-use change. *Arctic, Antarctic, and Alpine Research,* **45**, 404–414.

Westerling AL, Brown TJ, Gershunov A., Cayan DR, Dettinger MD (2003) Climate and Wildfire in the Western United States. *Bull. Amer. Met. Soc*., **84**, 595-604.

Wu Z, He HS, Yang J, Liu Z, Liang Y (2014) Relative effects of climatic and local factors on wildfire occurrence in boreal forest landscapes of northeastern China. *Sci. Total Environ.* **493**, 472–480.

Yang J, He HS, Shifley SR, Gustafson EJ (2007) Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. *Forest Science* **53**, 1–15.

Yang J, Weisberg PJ. Dilts TE, Loudermilk EL, Scheller RM, Stanton A, Skinner C (2015) Predicting wildfire occurrence distribution with spatial point process models and its uncertainty assessment: a case study in the Lake Tahoe Basin, USA. *Int. J. Wildland Fire*. **24**, 380-390

**Figures**

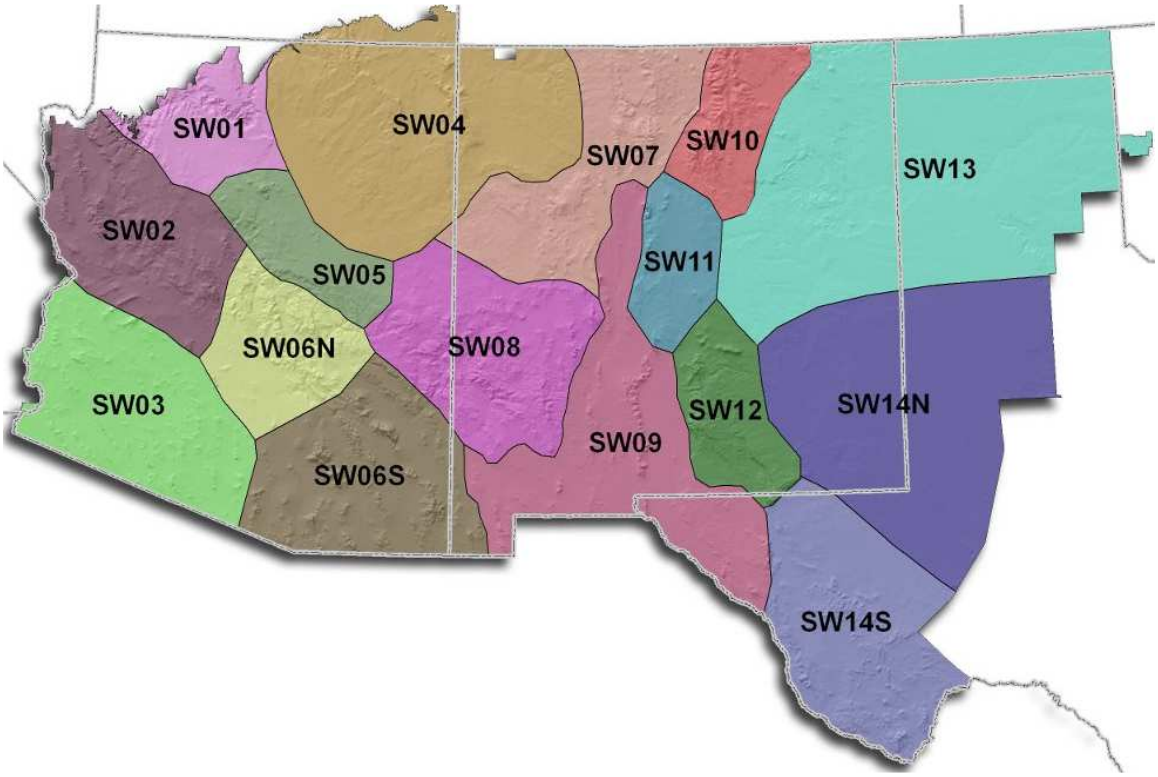


Figure 1. Study area: Southwest Area Predictive Service Areas (PSAs; i.e., PSA 1 is SW01). Also referenced to as the Southwest Area (SWA) in text

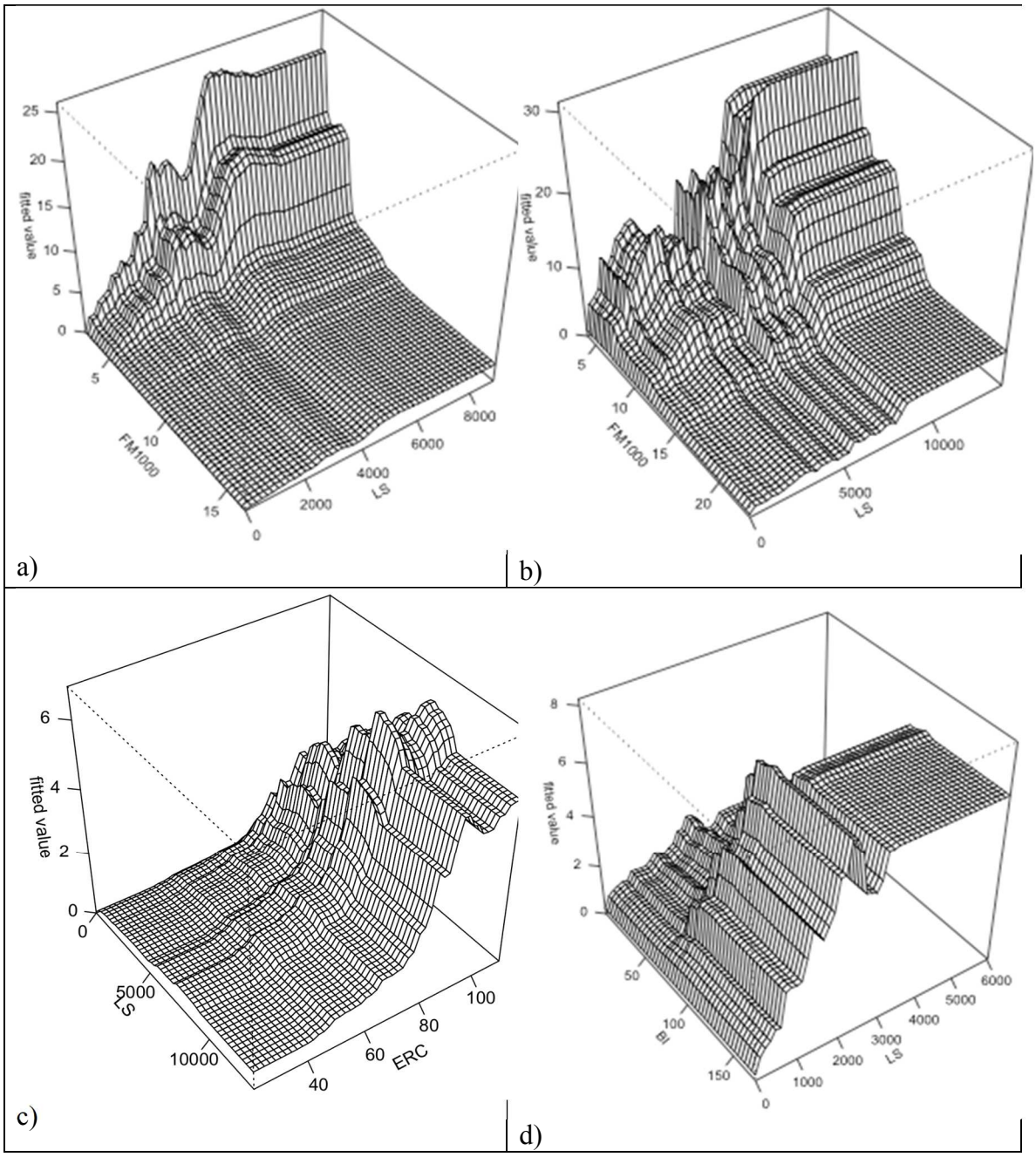


Figure 2. a) PSA 2 all lightning wildfires interaction plot for LS and FM1000 with the fitted value (number of fires) on the vertical axis and FM1000 (%) and LS (number of) on the horizontal axes; b) PSA 8 all lightning wildfires interaction plot for LS and FM1000 with the fitted value (number of fires) on the vertical axis and FM1000 (%) and LS

(number of) on the horizontal axes; c) PSA 6S all lightning wildfires interaction plot for LS and ERC with the fitted value (number of fires) on the vertical axis and ERC (BTU $ft^{-2}$) and LS (number of) on the horizontal axes; d) PSA 1 all lightning wildfires interaction plot for LS and BI with the fitted value (number of fires) on the vertical axis and BI and LS (number of) on the horizontal axes.

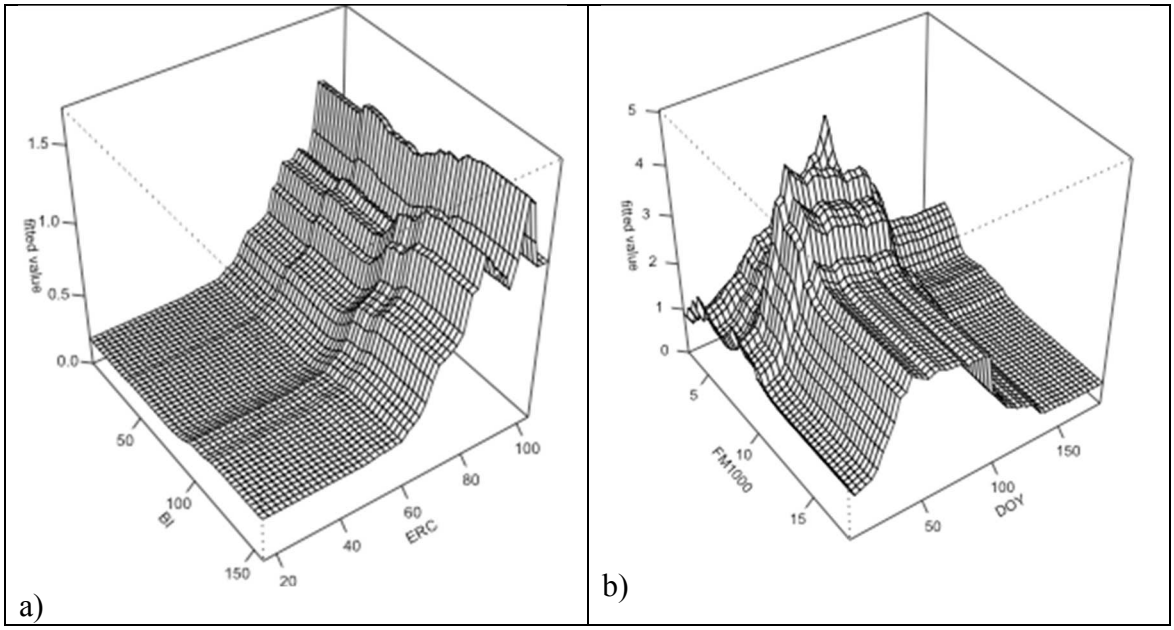


Figure 3. PSA 4 all human wildfires interaction plot for BI and ERC with the fitted value (number of fires) on the vertical axis and ERC (BTU $ft^{-2}$) and BI on the horizontal axes; b) PSA 6N all human wildfires interaction plot for DOY and FM1000 with the fitted value (number of fires) on the vertical axis and FM1000 (%) and DOY on the horizontal axes.

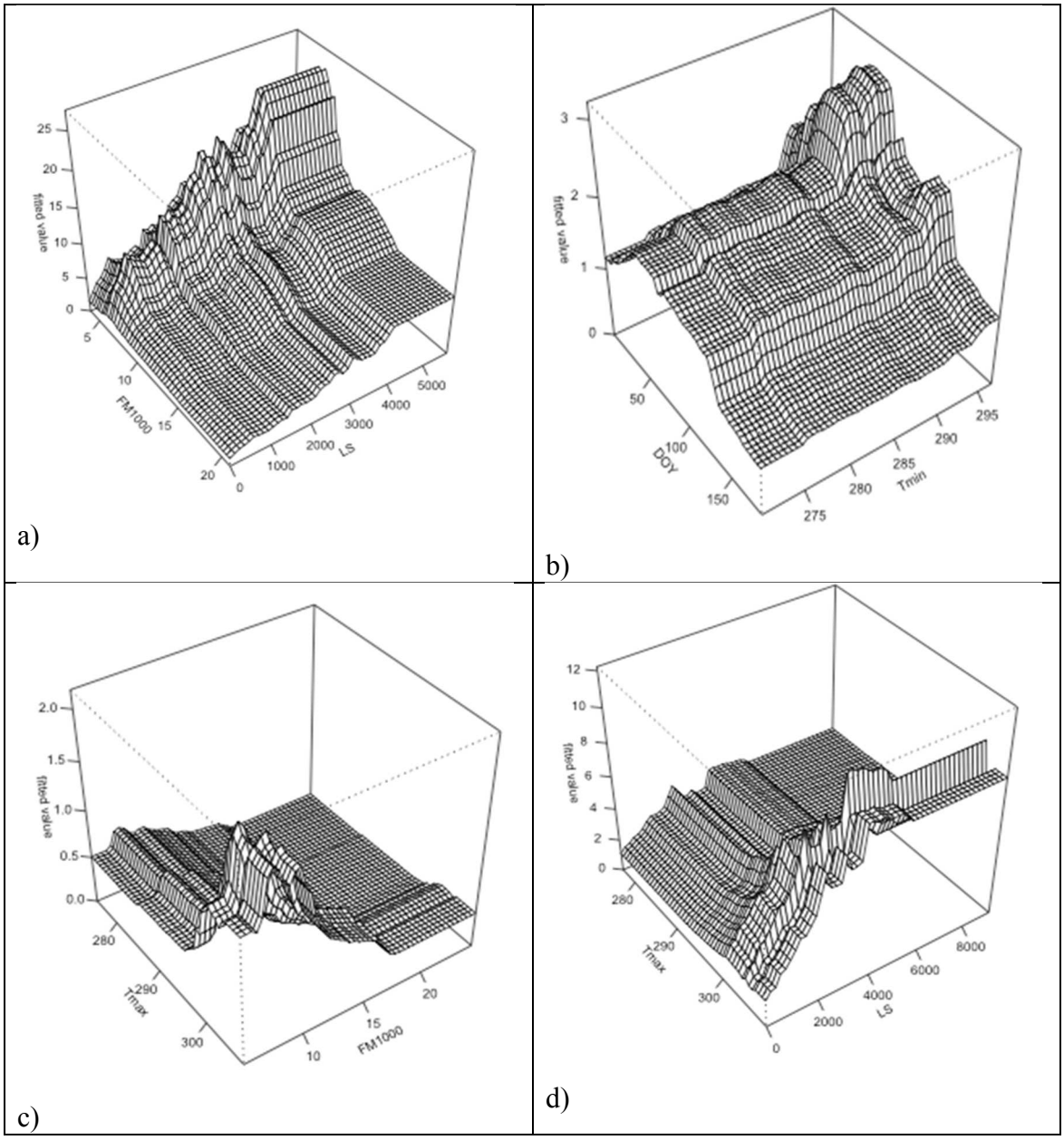


Figure 4. a) PSA 5 all wildfires interaction plot for LS and FM1000 with the fitted value (number of fires) on the vertical axis and FM1000 (%) and LS (number of) on the horizontal axes; b) PSA 6S all wildfires interaction plot for DOY and Tmin with the fitted value (number of fires) on the vertical axis and DOY and Tmin (Kelvin) on the horizontal axes c) PSA 10 all wildfires interaction plot for Tmax and FM1000 with the

fitted value (number of fires) on the vertical axis and FM1000 (%) and Tmax (Kelvin) on the horizontal axes; d) PSA 10 all wildfires interaction plot for Tmax and FM1000 with the fitted value (number of fires) on the vertical axis and Tmax (Kelvin) and LS (number of) on the horizontal axes.

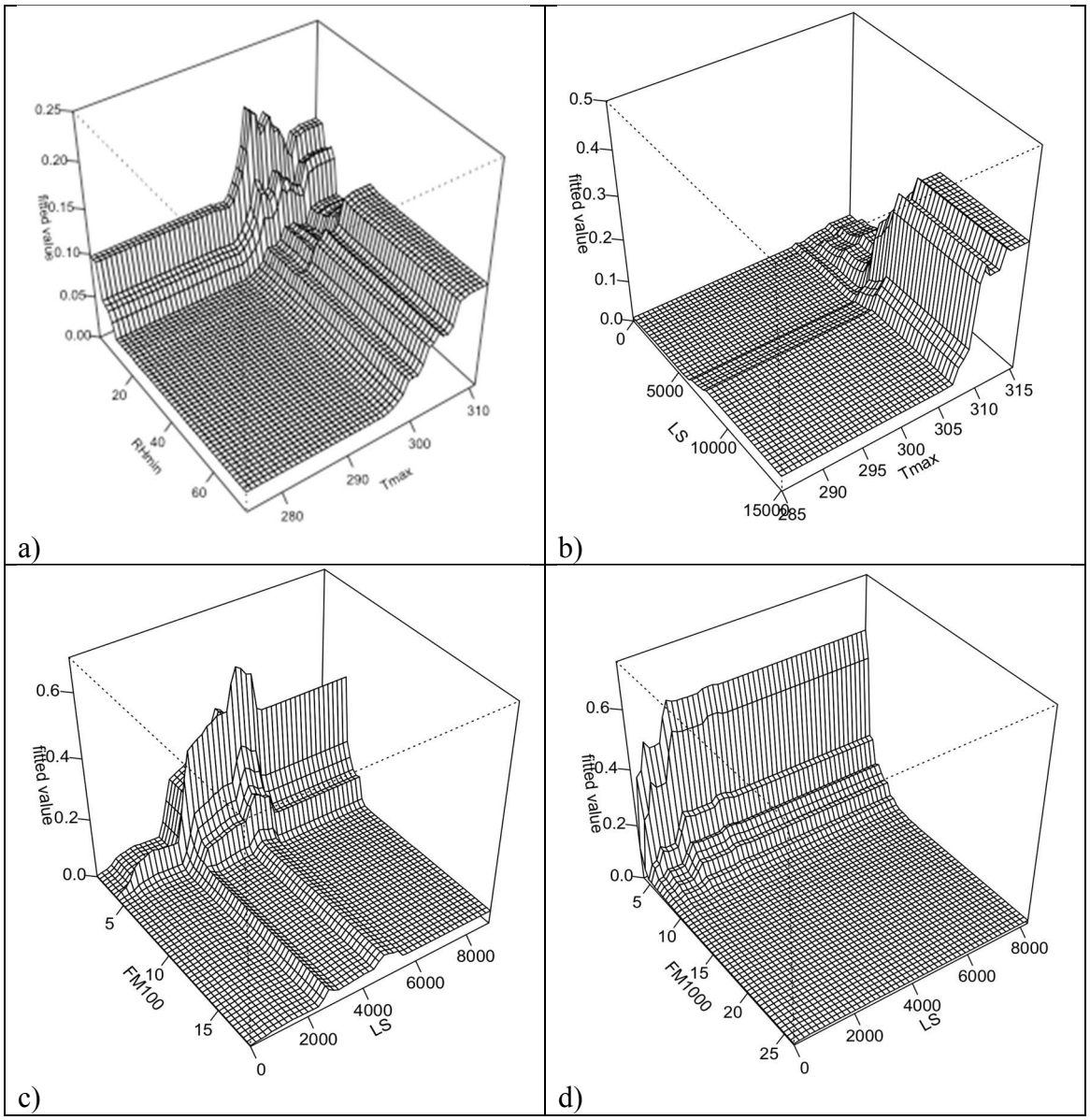


Figure 5. PSA 5 lightning caused large wildfires interaction plot for Tmax and RHmin with the fitted value (number of fires) on the vertical axis and RHmin (%) and Tmax

(Kelvin) on the horizontal axes; b) PSA 14S lightning caused large wildfires interaction plot for LS and Tmax with the fitted value (number of fires) on the vertical axis and Tmax (Kelvin) and LS (number of) on the horizontal axes; c) PSA 2 lightning caused large wildfires interaction plot for LS and FM100 with the fitted value (number of fires) on the vertical axis and FM100 (%) and LS (number of) on the horizontal axes; d) PSA 12 lightning caused large wildfires interaction plot for LS and FM1000 with the fitted value (number of fires) on the vertical axis and FM1000 (%) and LS (number of) on the horizontal axes.

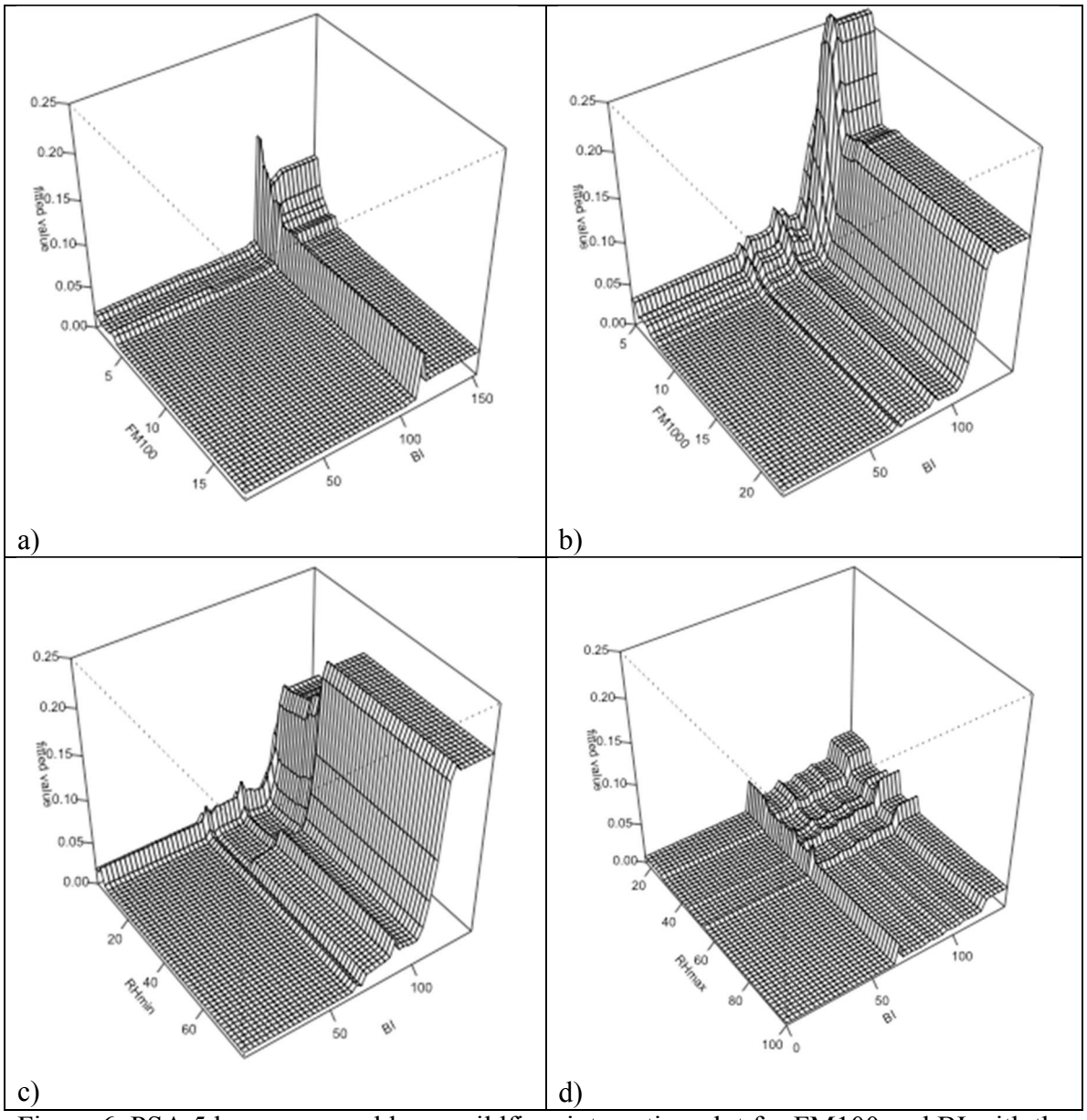


Figure 6. PSA 5 human caused large wildfires interaction plot for FM100 and BI with the fitted value (number of fires) on the vertical axis and FM100 (%) and BI on the horizontal axes; b) PSA 14N human caused large wildfires interaction plot for FM1000 and BI with the fitted value (number of fires) on the vertical axis and FM1000 (%) and BI on the horizontal axes; c) PSA 14N human caused large wildfires interaction plot for RHmin and BI with the fitted value (number of fires) on the vertical axis and RHmin (%)

and BI on the horizontal axes; d) PSA 10 human caused large wildfires interaction plot for BI and RHmax with the fitted value (number of fires) on the vertical axis and RHmax (%) and BI on the horizontal axes.

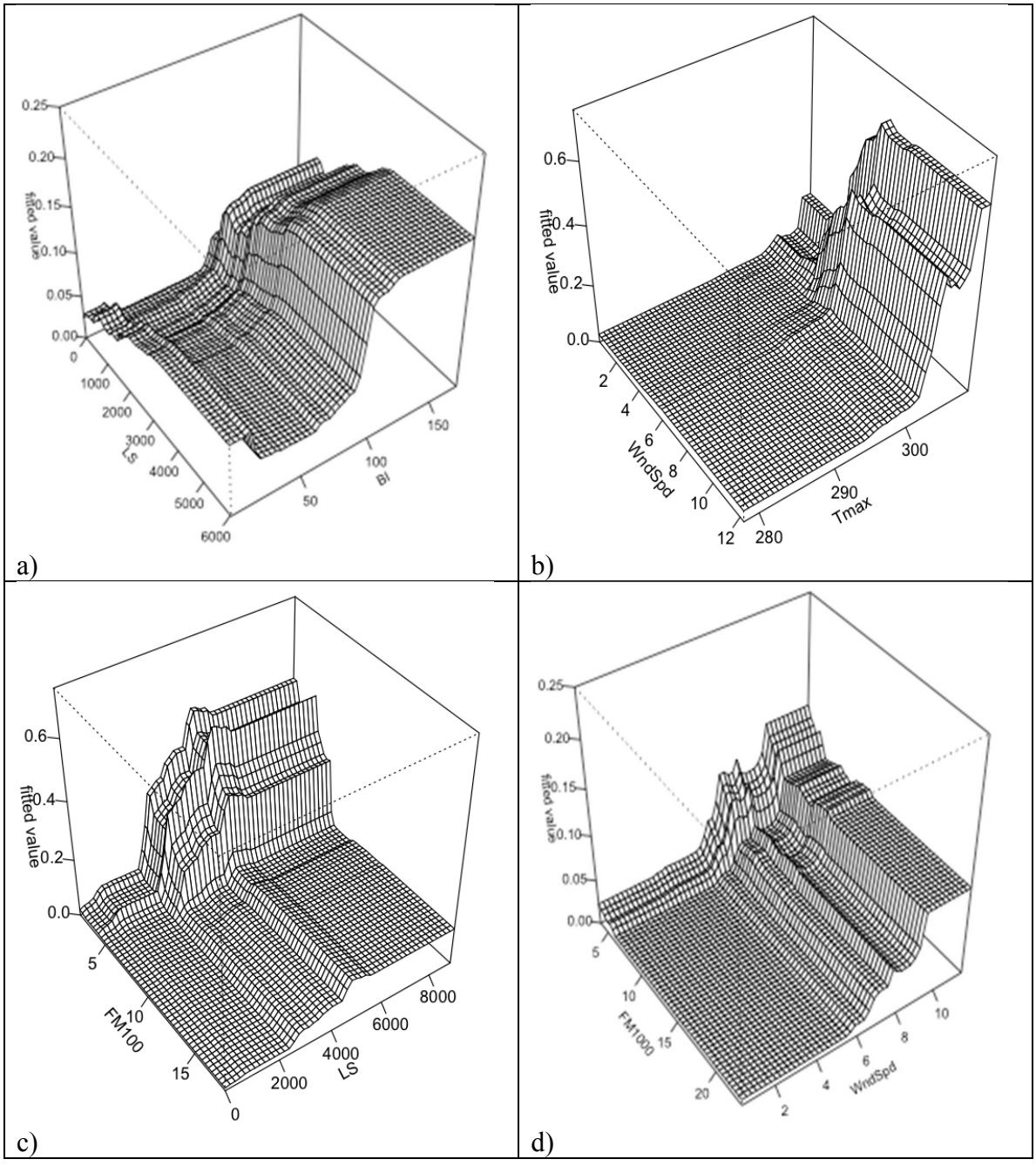

Figure 7. PSA 1 all large wildfires interaction plot for LS and BI with the fitted value (number of fires) on the vertical axis and LS (number of) and BI on the horizontal axes; b) PSA 8 all large wildfires interaction plot for WndSpd and Tmax with the fitted value (number of fires) on the vertical axis and WndSpd ($ms^{-1}$) and Tmax (Kelvin) on the horizontal axes; c) PSA 14S all large wildfires interaction plot for FM100 and LS with the fitted value (number of fires) on the vertical axis and FM100 (%) and LS (number of) on the horizontal axes; d) PSA 11 all large wildfires interaction plot for FM1000 and WndSpd with the fitted value (number of fires) on the vertical axis and FM1000 (%) and WndSpd on the horizontal axes.

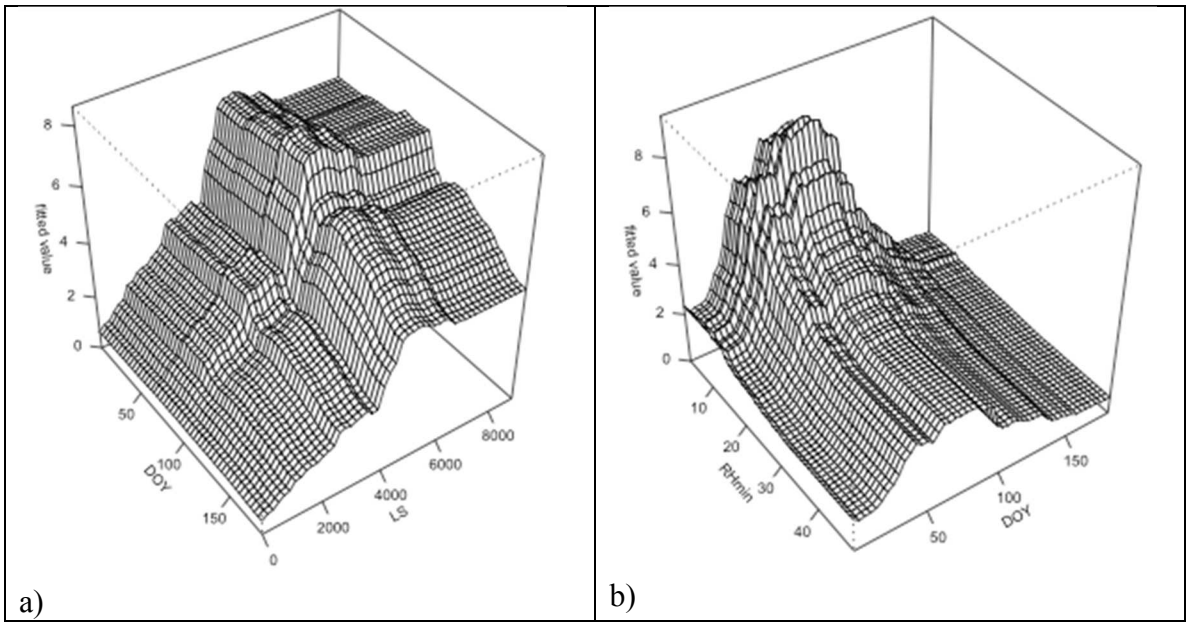


Figure 8. a) PSA 5 all wildfires interaction plot for LS and DOY with the fitted value (number of fires) on the vertical axis and LS (number of) and DOY on the horizontal axes; b) PSA 6N all human caused wildfires interaction plot for RHmin and DOY with the fitted value (number of fires) on the vertical axis and RHmin (%) and DOY on the horizontal axes.

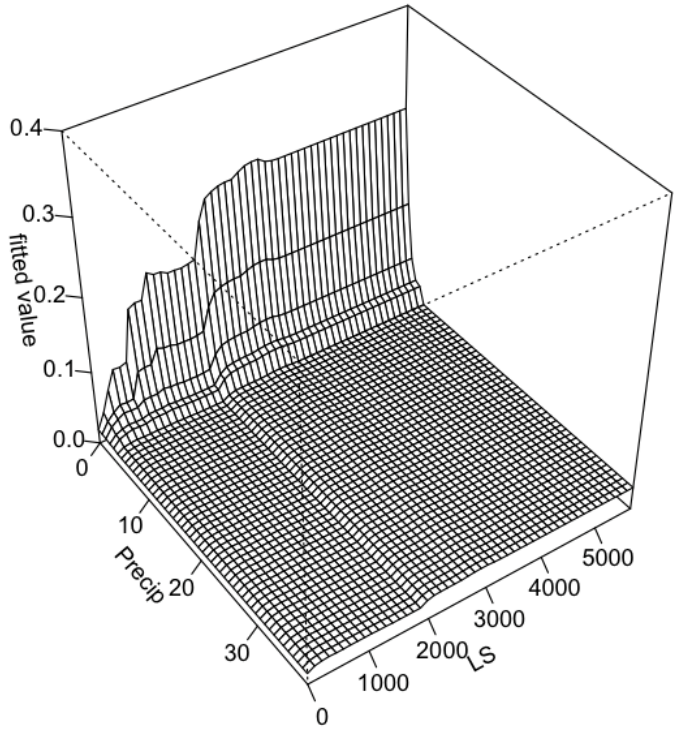


Figure 9. PSA 5 lightning caused wildfires interaction plot for LS and precipitation with the fitted value (number of fires) on the vertical axis and LS (number of) and precipitation (mm) on the horizontal axes

**Tables**

| **Variables** | **Units** | **Correlated frequently with** |
|---|---|---|
| Precipitation | mm | |
| RHmax | % | Rhmin, FM100, ERC, VPD |
| Rhmin | % | Rhmax, FM100, ERC, VPD |
| Tmax | Kelvin | Tmin, VPD |
| Tmin | Kelvin | Tmax, VPD |
| SPH | g/Kg | |
| VPD | kPa | Rhmax, Rhmin, FM100, Tmax, Tmin |
| ERC | BTU $ft^{-2}$ | FM100, FM1000, VPD |
| BI | unitless | |
| FM100 | % | ERC, FM1000, Rhmax, Rhmin, VPD |
| FM1000 | % | FM100, ERC |
| WndSpd | $ms^{-1}$ | |

Table 1. List of variables and their units used for this anaylsis.

| PSA | Large Wildfire Threshold (Acres) |
|---|---|
| SW01 | 100 |
| SW02 | 300 |
| SW03 | 300 |
| SW04 | 100 |
| SW05 | 50 |
| SW06S | 100 |
| SW07 | 100 |
| SW08 | 300 |
| SW09 | 300 |
| SW10 | 100 |
| SW11 | 100 |
| SW12 | 100 |
| SW13 | 2000 |
| SW14 | 200 |
| SW14N | 2000 |
| SW06N | 500 |

Table 2. Large wildfire threshold for each PSA in the SWA as determined by Southwest PS.

| All Days | All Wildfires | | | Lightning Wildfires | | | Human Wildfires | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Mean** | **MAE** | **ME** | **Mean** | **MAE** | **ME** | **Mean** | **MAE** | **ME** |
| **PSA 1** | 0.793 | 0.723 | -0.013 | 0.647 | 0.541 | -0.012 | 0.146 | 0.228 | -0.011 |
| **PSA 2** | 1.165 | 0.891 | -0.011 | 0.487 | 0.357 | -0.009 | 0.678 | 0.662 | -0.014 |
| **PSA 3** | 0.553 | 0.636 | -0.014 | 0.028 | 0.042 | -0.010 | 0.525 | 0.609 | -0.014 |
| **PSA 4** | 1.343 | 1.024 | -0.011 | 0.739 | 0.653 | -0.015 | 0.603 | 0.573 | -0.011 |
| **PSA 5** | 3.447 | 2.010 | -0.022 | 2.482 | 1.439 | -0.020 | 0.965 | 0.840 | -0.012 |
| **PSA 6S** | 1.546 | 1.063 | -0.010 | 0.547 | 0.469 | -0.010 | 0.999 | 0.796 | -0.012 |
| **PSA 6N** | 2.889 | 1.670 | -0.008 | 0.565 | 0.479 | -0.007 | 2.324 | 1.316 | -0.014 |
| **PSA 7** | 2.311 | 1.472 | -0.015 | 1.471 | 1.017 | -0.013 | 0.840 | 0.773 | -0.008 |
| **PSA 8** | 2.611 | 1.459 | -0.021 | 2.102 | 1.142 | -0.015 | 0.508 | 0.511 | -0.009 |
| **PSA 9** | 0.413 | 0.455 | -0.011 | 0.054 | 0.078 | -0.011 | 0.359 | 0.422 | -0.012 |
| **PSA 10** | 0.706 | 0.675 | -0.011 | 0.427 | 0.442 | -0.010 | 0.279 | 0.361 | -0.013 |
| **PSA 11** | 0.276 | 0.364 | -0.009 | 0.105 | 0.149 | -0.010 | 0.170 | 0.255 | -0.012 |
| **PSA 12** | 0.586 | 0.627 | -0.013 | 0.326 | 0.372 | -0.019 | 0.260 | 0.347 | -0.013 |
| **PSA 13** | 1.086 | 0.882 | -0.014 | 0.322 | 0.348 | -0.015 | 0.764 | 0.732 | -0.012 |
| **PSA 14S** | 0.192 | 0.266 | -0.015 | 0.129 | 0.179 | -0.017 | 0.062 | 0.101 | -0.011 |
| **PSA 14N** | 1.227 | 0.978 | -0.012 | 0.271 | 0.362 | -0.020 | 0.955 | 0.787 | -0.009 |
| | | | | | | | | | |
| **Median** | 1.126 | 0.887 | -0.013 | 0.457 | 0.407 | -0.013 | 0.564 | 0.591 | -0.012 |

Table 3. Displaying mean number of daily wildfires, median absolute error (MAE), and mean error (ME) for each PSA and wildfire type for AllYears method.

| **Fire Days** | **All Wildfires** | | | **Lightning Wildfires** | | | **Human Wildfires** | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | MAE | ME | Mean | MAE | ME | Mean | MAE | ME |
| **PSA 1** | 2.245 | 1.169 | -0.917 | 2.574 | 1.231 | -0.967 | 1.134 | 0.932 | -0.932 |
| **PSA 2** | 2.165 | 0.997 | -0.678 | 2.475 | 1.073 | -0.787 | 1.611 | 0.841 | -0.764 |
| **PSA 3** | 1.483 | 0.873 | -0.871 | 1.213 | 1.127 | -1.127 | 1.479 | 0.879 | -0.876 |
| **PSA 4** | 2.536 | 1.232 | -0.723 | 2.737 | 1.407 | -1.064 | 1.689 | 0.870 | -0.764 |
| **PSA 5** | 4.694 | 2.172 | -0.595 | 5.843 | 2.563 | -0.871 | 2.021 | 1.012 | -0.773 |
| **PSA 6S** | 2.524 | 1.163 | -0.590 | 2.400 | 1.207 | -0.889 | 2.053 | 1.014 | -0.647 |
| **PSA 6N** | 2.524 | 1.776 | -0.490 | 2.536 | 1.280 | -0.901 | 3.503 | 1.479 | -0.525 |
| **PSA 7** | 3.381 | 1.552 | -0.622 | 3.478 | 1.603 | -0.832 | 1.871 | 0.960 | -0.781 |
| **PSA 8** | 4.176 | 1.706 | -0.661 | 4.860 | 1.862 | -0.811 | 1.628 | 0.860 | -0.804 |
| **PSA 9** | 1.551 | 0.884 | -0.866 | 1.227 | 1.012 | -1.012 | 1.534 | 0.933 | -0.923 |
| **PSA 10** | 1.904 | 1.038 | -0.814 | 1.842 | 1.074 | -0.875 | 1.531 | 1.038 | -1.020 |
| **PSA 11** | 1.347 | 0.915 | -0.911 | 1.321 | 1.003 | -0.997 | 1.238 | 0.972 | -0.972 |
| **PSA 12** | 1.764 | 1.015 | -0.910 | 1.892 | 1.185 | -1.082 | 1.341 | 0.931 | -0.931 |
| **PSA 13** | 2.617 | 1.253 | -0.905 | 2.261 | 1.336 | -1.218 | 2.164 | 1.159 | -0.947 |
| **PSA 14S** | 1.597 | 1.186 | -1.157 | 1.667 | 1.272 | -1.246 | 1.241 | 1.121 | -1.121 |
| **PSA 14N** | 2.764 | 1.293 | -0.938 | 2.195 | 1.602 | -1.487 | 2.394 | 1.114 | -0.882 |
| | | | | | | | | | |
| **Median** | 2.385 | 1.177 | -0.840 | 2.330 | 1.251 | -0.982 | 1.620 | 0.966 | -0.879 |

Table 4. Displaying mean number of daily wildfires, median absolute error (MAE), and mean error (ME) for each PSA and wildfire type only for days with wildfires using the AllYears method.

| | All Wildfires | | | | | |
|---|---|---|---|---|---|---|
| | **Mean** | **MAE** | **ME** | **Mean FD** | **MAE_FD** | **ME_FD** |
| **PSA 1** | 0.760 | 0.819 | 0.014 | 2.245 | 1.359 | -0.949 |
| **PSA 2** | 1.195 | 0.947 | 0.134 | 2.165 | 0.964 | -0.546 |
| **PSA 3** | 0.549 | 0.598 | 0.052 | 1.483 | 0.769 | -0.769 |
| **PSA 4** | 1.321 | 1.115 | 0.026 | 2.536 | 1.268 | -0.663 |
| **PSA 5** | 3.226 | 2.311 | 0.152 | 4.694 | 2.546 | -0.440 |
| **PSA 6S** | 3.199 | 2.453 | 0.055 | 2.524 | 2.523 | -0.594 |
| **PSA 6N** | 2.782 | 1.809 | 0.481 | 2.524 | 1.829 | 0.182 |
| **PSA 7** | 2.149 | 1.656 | 0.103 | 3.381 | 1.766 | -0.603 |
| **PSA 8** | 2.303 | 1.980 | 0.016 | 4.176 | 2.554 | -0.603 |
| **PSA 9** | 0.354 | 0.487 | -0.024 | 1.551 | 0.918 | -0.918 |
| **PSA 10** | 0.706 | 0.762 | 0.058 | 1.904 | 1.055 | -0.842 |
| **PSA 11** | 0.256 | 0.376 | 0.006 | 1.347 | 0.931 | -0.929 |
| **PSA 12** | 0.552 | 0.628 | -0.061 | 1.764 | 1.127 | -0.966 |
| **PSA 13** | 0.913 | 0.906 | 0.177 | 2.617 | 1.048 | -0.888 |
| **PSA 14S** | 0.164 | 0.261 | 0.016 | 1.597 | 1.173 | -1.173 |
| **PSA 14N** | 1.018 | 1.000 | 0.128 | 2.764 | 1.009 | -0.898 |
| | | | | | | |
| **Median** | 0.966 | 0.926 | 0.053 | 2.385 | 1.150 | -0.805 |

Table 5. Displaying the mean number of daily fires, MAEs, and MEs for all wildfires across all PSAs for all days and days with wildfires (FD) using the LeaveOne method.

| | Human Wildfires | | | | | |
|---|---|---|---|---|---|---|
| | **Mean** | **MAE** | **ME** | **Mean FD** | **MAE_FD** | **ME_FD** |
| **PSA 1** | 0.135 | 0.222 | 0.040 | 1.134 | 0.989 | -0.989 |
| **PSA 2** | 0.694 | 0.672 | 0.090 | 1.611 | 0.745 | -0.660 |
| **PSA 3** | 0.517 | 0.592 | 0.035 | 1.479 | 0.786 | -0.786 |
| **PSA 4** | 0.585 | 0.612 | -0.039 | 1.689 | 0.825 | -0.688 |
| **PSA 5** | 0.939 | 0.878 | 0.178 | 2.021 | 0.880 | -0.575 |
| **PSA 6S** | 0.971 | 0.797 | 0.098 | 2.053 | 0.926 | -0.503 |
| **PSA 6N** | 2.039 | 1.367 | 0.228 | 3.503 | 1.500 | -0.123 |
| **PSA 7** | 0.810 | 0.794 | 0.054 | 1.871 | 0.906 | -0.687 |
| **PSA 8** | 0.513 | 0.520 | 0.095 | 1.628 | 0.799 | -0.734 |
| **PSA 9** | 0.342 | 0.445 | -0.030 | 1.534 | 0.973 | -0.948 |
| **PSA 10** | 0.525 | 0.584 | 0.041 | 1.531 | 0.903 | -0.810 |
| **PSA 11** | 0.151 | 0.247 | -0.007 | 1.238 | 0.983 | -0.983 |
| **PSA 12** | 0.246 | 0.351 | 0.026 | 1.341 | 0.883 | -0.883 |
| **PSA 13** | 0.606 | 0.684 | 0.092 | 2.164 | 1.106 | -0.900 |
| **PSA 14S** | 0.047 | 0.089 | 0.014 | 1.241 | 1.031 | -1.031 |
| **PSA 14N** | 0.767 | 0.834 | 0.019 | 2.394 | 1.085 | -0.937 |
| | | | | | | |
| **Median** | 0.555 | 0.602 | 0.041 | 1.620 | 0.916 | -0.798 |

Table 6. Displaying the mean number of daily wildfires, MAEs, and MEs for human caused wildfires across all PSAs for all days and days with wildfires (FD) using the LeaveOne method.

| | Lightning Wildfires | | | | | |
|---|---|---|---|---|---|---|
| | **Mean** | **MAE** | **ME** | **Mean FD** | **MAE_FD** | **ME_FD** |
| **PSA 1** | 0.589 | 0.689 | 0.051 | 2.574 | 1.705 | -1.181 |
| **PSA 2** | 0.434 | 0.458 | 0.037 | 2.475 | 1.396 | -0.861 |
| **PSA 3** | 0.018 | 0.038 | -0.004 | 1.213 | 0.983 | -0.983 |
| **PSA 4** | 0.682 | 0.784 | -0.043 | 2.737 | 1.732 | -0.978 |
| **PSA 5** | 2.227 | 1.837 | -0.075 | 5.843 | 3.469 | -1.153 |
| **PSA 6S** | 0.506 | 0.528 | 0.020 | 2.400 | 1.382 | -0.923 |
| **PSA 6N** | 0.514 | 0.588 | -0.046 | 2.536 | 1.508 | -1.198 |
| **PSA 7** | 1.353 | 1.216 | -0.027 | 3.478 | 1.885 | -0.903 |
| **PSA 8** | 1.782 | 1.761 | -0.006 | 4.860 | 2.732 | -0.796 |
| **PSA 9** | 0.039 | 0.074 | -0.009 | 1.227 | 1.078 | -1.078 |
| **PSA 10** | 0.376 | 0.450 | 0.014 | 1.842 | 1.117 | -0.819 |
| **PSA 11** | 0.086 | 0.154 | 0.010 | 1.321 | 0.989 | -0.985 |
| **PSA 12** | 0.285 | 0.393 | -0.018 | 1.892 | 1.345 | -1.218 |
| **PSA 13** | 0.209 | 0.295 | 0.090 | 2.261 | 1.180 | -1.078 |
| **PSA 14S** | 0.112 | 0.158 | 0.016 | 1.667 | 1.358 | -1.358 |
| **PSA 14N** | 0.201 | 0.335 | 0.041 | 2.195 | 1.269 | -1.231 |
| | | | | | | |
| **Median** | 0.405 | 0.454 | 0.003 | 2.330 | 1.370 | -1.032 |

Table 7. Displaying the mean number of daily wildfires, MAEs, and MEs for lightning caused wildfires across all PSAs for all days and days with wildfires (FD) using the LeaveOne method.

| | All Lightning Wildfires | | All Human Wildfires | | All Wildfires | |
|---|---|---|---|---|---|---|
| | **Count** | **Median** | **Count** | **Median** | **Count** | **Median** |
| **FM1000** | 12 | 24.2% | 6 | 18.4% | 13 | 25.0% |
| **BI** | 15 | 5.3% | 14 | 9.6% | 15 | 6.3% |
| **RHmax** | 3 | 5.1% | 2 | 10.4% | 1 | 6.4% |
| **RHmin** | 11 | 6.9% | 5 | 10.9% | 13 | 8.7% |
| **Tmax** | 8 | 13.8% | 12 | 11.7% | 9 | 14.8% |
| **Tmin** | 8 | 12.4% | 4 | 7.2% | 7 | 10.9% |
| **WndSpd** | 16 | 3.9% | 16 | 7.5% | 16 | 4.9% |
| **Precip.** | 16 | 5.0% | 16 | 2.7% | 15 | 3.8% |
| **SPH** | 13 | 5.3% | 16 | 8.6% | 14 | 5.5% |
| **DOY** | 16 | 4.5% | 16 | 13.2% | 16 | 7.7% |
| **LS** | 16 | 28.4% | 16 | 2.5% | 16 | 21.9% |
| **ERC** | 3 | 24.7% | 6 | 43.7% | 3 | 29.9% |
| **FM100** | 2 | 26.4% | 6 | 32.1% | 2 | 13.0% |
| **VPD** | 0 | | 1 | 21.4% | 0 | |

Table 8. Contains the number of times (count) each variable was used in a BRT model and the median relative influence by type when it was used for all wildfires.

| Interaction | Count | PSA(s) |
|---|---|---|
| FM1000, LS | 11 | PSA10, PSA11, PSA12, PSA13, PSA14S, PSA1, PSA2, PSA3, PSA5, PSA7, PSA8 |
| BI, LS | 4 | PSA14S, PSA1, PSA2, PSA9 |
| FM1000, Tmax | 4 | PSA10, PSA12, PSA14S, PSA8 |
| LS, Tmax | 4 | PSA14S, PSA14N, PSA8, PSA9 |
| FM1000, Precip | 4 | PSA10, PSA11, PSA2, PSA7 |
| LS, Precip | 3 | PSA3, PSA5, PSA9 |
| ERC, LS | 3 | PSA6N, PSA4, PSA6S |
| BI, FM1000 | 2 | PSA10, PSA12 |
| ERC, Tmin | 2 | PSA4, PSA6S |
| ERC, Precip | 2 | PSA6N, PSA4 |
| LS, RHmin | 2 | PSA7, PSA9 |
| LS, WndSpd | 2 | PSA3, PSA5 |
| Tmin, WndSpd | 2 | PSA1, PSA4 |
| FM1000, SPH | 2 | PSA11, PSA13 |
| LS, SPH | 1 | PSA8 |
| RHmax, Tmax | 1 | PSA5 |
| BI, RHmax | 1 | PSA14N |
| FM100, FM1000 | 1 | PSA13 |
| FM100, LS | 1 | PSA14N |
| BI, Tmin | 1 | PSA3 |
| DOY, Tmin | 1 | PSA13 |
| FM100, Tmax | 1 | PSA14N |
| RHmin, Tmin | 1 | PSA1 |
| FM1000, Tmin | 1 | PSA11 |
| LS, Tmin | 1 | PSA6N |
| DOY, LS | 1 | PSA6N |
| FM1000, WndSpd | 1 | PSA2 |
| RHmin, Tmax | 1 | PSA7 |

Table 9. Contains the pair of variables interacting and how often that occurs for all lightning caused wildfires across all PSAs.

| Interaction | Count | PSA(s) |
|---|---|---|
| DOY, ERC | 5 | PSA13, PSA4, PSA6S, PSA7, PSA8 |
| ERC, WndSpd | 4 | PSA11, PSA13, PSA4, PSA8 |
| DOY, FM100 | 3 | PSA12, PSA5, PSA9 |
| ERC, Precip | 3 | PSA11, PSA13, PSA7 |
| BI, WndSpd | 3 | PSA6N, PSA1, PSA9 |
| FM100, WndSpd | 2 | PSA12, PSA9 |
| ERC, Tmax | 2 | PSA11, PSA6S |
| FM100, FM1000 | 2 | PSA1, PSA2 |
| BI, FM100 | 2 | PSA14S, PSA1 |
| FM1000, LS | 2 | PSA10, PSA2 |
| BI, RHmin | 2 | PSA10, PSA14N |
| DOY, Tmin | 2 | PSA6N, PSA3 |
| DOY, FM1000 | 2 | PSA14N, PSA6N |
| BI, ERC | 2 | PSA13, PSA4 |
| FM1000, RHmin | 1 | PSA14N |
| DOY, WndSpd | 1 | PSA12 |
| BI, RHmax | 1 | PSA4 |
| Tmax, WndSpd | 1 | PSA5 |
| ERC, RHmin | 1 | PSA7 |
| BI, Precip | 1 | PSA1 |
| BI, Tmax | 1 | PSA14S |
| LS, Precip | 1 | PSA2 |
| Precip, RHmin | 1 | PSA14N |
| Precip, Tmax | 1 | PSA11 |
| BI, LS | 1 | PSA14S |
| FM100, SPH | 1 | PSA2 |
| FM1000, VPD | 1 | PSA3 |
| BI, DOY | 1 | PSA7 |
| FM100, Precip | 1 | PSA5 |
| DOY, VPD | 1 | PSA3 |
| RHmin, Tmax | 1 | PSA10 |
| DOY, RHmin | 1 | PSA6N |
| FM1000, Precip | 1 | PSA3 |
| DOY, Tmax | 1 | PSA10 |

Table 10. Contains the pair of variables interacting and how often that occurs for all human caused wildfires across all PSAs

| Interaction | Count | PSA(s) |
|---|---|---|
| FM1000, LS | 8 | PSA10, PSA12, PSA13, PSA1, PSA2, PSA5, PSA7, PSA8 |
| FM1000, Tmax | 4 | PSA10, PSA11, PSA14S, PSA8 |
| DOY, Tmin | 3 | PSA3, PSA4, PSA6S |
| BI, LS | 3 | PSA12, PSA1, PSA2 |
| DOY, FM1000 | 3 | PSA14S, PSA14N, PSA3 |
| LS, SPH | 2 | PSA2, PSA8 |
| ERC, Tmin | 2 | PSA6N, PSA4 |
| BI, RHmin | 2 | PSA13, PSA14N |
| LS, WndSpd | 2 | PSA5, PSA8 |
| DOY, Precip | 2 | PSA14N, PSA6N |
| ERC, LS | 2 | PSA4, PSA6S |
| LS, Tmax | 2 | PSA14S, PSA7 |
| RHmin, Tmax | 2 | PSA10, PSA7 |
| FM1000, Precip | 2 | PSA11, PSA3 |
| DOY, ERC | 2 | PSA6N, PSA4 |
| FM1000, RHmin | 1 | PSA3 |
| DOY, WndSpd | 1 | PSA9 |
| Tmax, WndSpd | 1 | PSA11 |
| BI, FM1000 | 1 | PSA10 |
| FM100, FM1000 | 1 | PSA13 |
| LS, RHmax | 1 | PSA5 |
| BI, Tmax | 1 | PSA14S |
| LS, RHmin | 1 | PSA7 |
| Precip, RHmin | 1 | PSA13 |
| RHmin, Tmin | 1 | PSA1 |
| Precip, Tmax | 1 | PSA11 |
| BI, WndSpd | 1 | PSA9 |
| Tmin, WndSpd | 1 | PSA1 |
| Precip, WndSpd | 1 | PSA12 |
| RHmax, Tmax | 1 | PSA5 |
| BI, DOY | 1 | PSA14N |
| DOY, LS | 1 | PSA2 |
| FM1000, WndSpd | 1 | PSA9 |
| DOY, RHmin | 1 | PSA6N |
| DOY, Tmax | 1 | PSA12 |

Table 11. Contains the pair of variables interacting and how often that occurs for all wildfires across all PSAs.

| | All Large Wildfires | | Lightning Caused Large Wildfires | | Human Caused Large Wildfires | |
|---|---|---|---|---|---|---|
| | **CV AUC** | **Leave One Out Median CV AUC** | **CV AUC** | **Leave One Out Median CV AUC** | **CV AUC** | **Leave One Out Median CV AUC** |
| **PSA 1** | 0.76 | 0.7252 | 0.774 | 0.7625 | 0.741 | 0.706 |
| **PSA 2** | 0.79 | 0.7755 | 0.914 | 0.9052 | 0.639 | 0.63665 |
| **PSA 3** | 0.725 | 0.701 | 0.85 | 0.8436 | 0.734 | 0.722 |
| **PSA 4** | 0.749 | 0.7574 | 0.795 | 0.7766 | 0.786 | 0.7865 |
| **PSA 5** | 0.717 | 0.7054 | 0.759 | 0.7563 | 0.828 | 0.8279 |
| **PSA 6S** | 0.758 | 0.7 | 0.851 | 0.8467 | 0.735 | 0.768 |
| **PSA 6N** | 0.747 | 0.7562 | 0.886 | 0.8818 | 0.726 | 0.7329 |
| **PSA 7** | 0.674 | 0.6861 | 0.721 | 0.7032 | 0.79 | 0.7972 |
| **PSA 8** | 0.724 | 0.7272 | 0.749 | 0.745 | 0.798 | 0.7946 |
| **PSA 9** | 0.702 | 0.70825 | 0.751 | 0.751 | 0.795 | 0.7919 |
| **PSA 10** | 0.762 | 0.7575 | 0.713 | 0.7318 | 0.863 | 0.8586 |
| **PSA 11** | 0.753 | 0.7431 | 0.798 | 0.8084 | 0.62 | 0.609 |
| **PSA 12** | 0.781 | 0.7835 | 0.857 | 0.864 | 0.787 | 0.7884 |
| **PSA 13** | 0.81 | 0.7939 | 0.827 | 0.8235 | 0.809 | 0.7972 |
| **PSA 14S** | 0.788 | 0.8053 | 0.793 | 0.7929 | 0.824 | 0.8141 |
| **PSA 14N** | 0.845 | 0.8468 | 0.823 | 0.7995 | 0.853 | 0.8559 |
| | | | | | | |
| **Median** | 0.756 | 0.750 | 0.797 | 0.796 | 0.789 | 0.790 |

Table 12. CV AUC values for all wildfire types across all PSAs using all data and the LeaveOne method.

| | All Large Wildfires | | | Lightning Caused Large Wildfires | | | Human Caused Large Wildfires | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fire Probability | No Fire Probability | Mean Probability | Fire Probability | No Fire Probability | Mean Probability | Fire Probability | No Fire Probability | Mean Probability |
| **PSA 1** | 0.159 | 0.02 | 0.024 | 0.159 | 0.02 | 0.024 | 0.039 | 0.004 | 0.004 |
| **PSA 2** | 0.225 | 0.019 | 0.023 | 0.302 | 0.009 | 0.014 | 0.044 | 0.01 | 0.011 |
| **PSA 3** | 0.132 | 0.012 | 0.123 | 0.197 | 0.002 | 0.003 | 0.125 | 0.009 | 0.01 |
| **PSA 4** | 0.125 | 0.012 | 0.013 | 0.147 | 0.004 | 0.005 | 0.084 | 0.008 | 0.009 |
| **PSA 5** | 0.171 | 0.045 | 0.049 | 0.143 | 0.03 | 0.034 | 0.247 | 0.011 | 0.015 |
| **PSA 6S** | 0.219 | 0.087 | 0.092 | 0.27 | 0.034 | 0.04 | 0.134 | 0.053 | 0.055 |
| **PSA 6N** | 0.193 | 0.022 | 0.025 | 0.266 | 0.012 | 0.015 | 0.057 | 0.01 | 0.01 |
| **PSA 7** | 0.081 | 0.022 | 0.024 | 0.115 | 0.015 | 0.017 | 0.112 | 0.007 | 0.009 |
| **PSA 8** | 0.197 | 0.045 | 0.049 | 0.192 | 0.037 | 0.04 | 0.146 | 0.008 | 0.01 |
| **PSA 9** | 0.078 | 0.014 | 0.015 | 0.167 | 0.004 | 0.005 | 0.144 | 0.009 | 0.01 |
| **PSA 10** | 0.156 | 0.022 | 0.026 | 0.108 | 0.014 | 0.016 | 0.231 | 0.007 | 0.01 |
| **PSA 11** | 0.135 | 0.01 | 0.012 | 0.229 | 0.004 | 0.006 | 0.026 | 0.006 | 0.006 |
| **PSA 12** | 0.203 | 0.029 | 0.033 | 0.258 | 0.014 | 0.016 | 0.181 | 0.014 | 0.017 |
| **PSA 13** | 0.217 | 0.014 | 0.017 | 0.202 | 0.007 | 0.008 | 0.173 | 0.008 | 0.009 |
| **PSA 14S** | 0.211 | 0.025 | 0.028 | 0.208 | 0.018 | 0.021 | 0.094 | 0.007 | 0.008 |
| **PSA 14N** | 0.277 | 0.017 | 0.022 | 0.128 | 0.004 | 0.005 | 0.275 | 0.013 | 0.018 |
| | | | | | | | | | |
| **Median** | 0.182 | 0.021 | 0.025 | 0.195 | 0.013 | 0.016 | 0.130 | 0.009 | 0.010 |

Table 13. Displaying the probabilities associated with large wildfire days (Fire Probability), days without large wildfires (No Fire Probability), and daily occurrence of large wildfires by type and PSA using the AllYears method.

| | **All Large Wildfires** | | **Lightning Caused Large Wildfires** | | **Human Caused Large Wildfires** | |
|---|---|---|---|---|---|---|
| | Fire Probability | No Fire Probability | Fire Probability | No Fire Probability | Fire Probability | No Fire Probability |
| **PSA 1** | 0.032 | 0.020 | 0.030 | 0.014 | 0.004 | 0.003 |
| **PSA 2** | 0.037 | 0.014 | 0.034 | 0.003 | 0.010 | 0.009 |
| **PSA 3** | 0.013 | 0.010 | 0.031 | 0.001 | 0.013 | 0.009 |
| **PSA 4** | 0.013 | 0.007 | 0.004 | 0.003 | 0.010 | 0.006 |
| **PSA 5** | 0.055 | 0.038 | 0.045 | 0.021 | 0.024 | 0.005 |
| **PSA 6S** | 0.058 | 0.040 | 0.119 | 0.016 | 0.055 | 0.027 |
| **PSA 6N** | 0.046 | 0.017 | 0.056 | 0.005 | 0.009 | 0.008 |
| **PSA 7** | 0.028 | 0.018 | 0.016 | 0.011 | 0.010 | 0.005 |
| **PSA 8** | 0.058 | 0.030 | 0.054 | 0.022 | 0.011 | 0.005 |
| **PSA 9** | 0.015 | 0.011 | 0.003 | 0.003 | 0.013 | 0.006 |
| **PSA 10** | 0.025 | 0.017 | 0.013 | 0.011 | 0.013 | 0.004 |
| **PSA 11** | 0.011 | 0.007 | 0.020 | 0.003 | 0.026 | 0.006 |
| **PSA 12** | 0.038 | 0.019 | 0.020 | 0.007 | 0.018 | 0.010 |
| **PSA 13** | 0.014 | 0.008 | 0.007 | 0.005 | 0.007 | 0.005 |
| **PSA 14S** | 0.045 | 0.015 | 0.030 | 0.012 | 0.007 | 0.005 |
| **PSA 14N** | 0.020 | 0.009 | 0.004 | 0.003 | 0.010 | 0.007 |
| | | | | | | |
| **Median** | 0.030 | 0.016 | 0.025 | 0.006 | 0.010 | 0.006 |

Table 14. Displaying the probabilities associated with large wildfire days (Fire Probability) and days without large wildfires (No Fire Probability) by type and PSA using the LeaveOne method in the BRT models.

| | Lightning Caused Large Wildfires | | Human Caused Large Wildfires | | All Large Wildfires | |
|---|---|---|---|---|---|---|
| | **Count** | **Median** | **Count** | **Median** | **Count** | **Median** |
| **FM1000** | 11 | 16.4% | 11 | 15.3% | 14 | 15.7% |
| **BI** | 16 | 7.4% | 16 | 17.7% | 14 | 10.0% |
| **RHmax** | 1 | 7.5% | 1 | 16.7% | 1 | 8.3% |
| **RHmin** | 8 | 8.8% | 8 | 25.1% | 12 | 12.6% |
| **Tmax** | 10 | 21.2% | 11 | 12.9% | 11 | 16.3% |
| **Tmin** | 6 | 15.1% | 5 | 10.9% | 5 | 16.4% |
| **WndSpd** | 16 | 7.3% | 16 | 11.3% | 16 | 9.3% |
| **Precip.** | 16 | 8.3% | 16 | 2.7% | 16 | 7.6% |
| **SPH** | 14 | 6.5% | 15 | 11.6% | 13 | 9.0% |
| **DOY** | 16 | 6.8% | 16 | 8.1% | 16 | 8.7% |
| **LS** | 16 | 17.5% | 16 | 2.2% | 16 | 12.4% |
| **ERC** | 0 | | 1 | 19.2% | 2 | 19.7% |
| **FM100** | 6 | 14.5% | 2 | 16.6% | 2 | 16.9% |
| **VPD** | 0 | | 4 | 25.3% | 0 | |

Table 15. Contains the number of times (count) each variable was used in a BRT model and the median relative influence by type when it was used for large wildfires.

| Interaction | Count | PSA(s) |
|---|---|---|
| FM1000, Tmax | 5 | PSA10, PSA14S PSA6N, PSA4, PSA6S |
| LS, Tmax | 5 | PSA14S PSA4, PSA6S, PSA8, PSA9 |
| Tmax, WndSpd | 5 | PSA10, PSA14S PSA6N, PSA7, PSA8 |
| BI, Tmax | 4 | PSA6N, PSA4, PSA7, PSA9 |
| BI, FM1000 | 3 | PSA10, PSA11, PSA13 |
| BI, LS | 3 | PSA14N, PSA1, PSA2 |
| BI, RHmin | 3 | PSA10, PSA1, PSA4 |
| FM100, LS | 3 | PSA14N, PSA2, PSA9 |
| DOY, FM1000 | 2 | PSA13, PSA6S |
| DOY, LS | 2 | PSA1, PSA3 |
| FM1000, LS | 2 | PSA12, PSA14 |
| FM1000, Precip | 2 | PSA12, PSA5 |
| LS, SPH | 2 | PSA3, PSA5 |
| LS, WndSpd | 2 | PSA2, PSA3 |
| Tmin, WndSpd | 2 | PSA1, PSA5 |
| BI, Precip | 1 | PSA14N |
| DOY, WndSpd | 1 | PSA11 |
| FM100, Tmax | 1 | PSA7 |
| FM1000, RHmax | 1 | PSA13 |
| FM1000, SPH | 1 | PSA13 |
| FM1000, Tmin | 1 | PSA11 |
| FM1000, WndSpd | 1 | PSA11 |
| LS, Tmin | 1 | PSA2 |
| Precip, Tmin | 1 | PSA12 |
| RHmin, Tmin | 1 | PSA5 |
| SPH, Tmax | 1 | PSA6N |

Table 16. Contains the pair of variables interacting and how often that occurs for lightning caused large wildfires across all PSAs.

| Interaction | Count | PSA(s) |
|---|---|---|
| BI, Tmax | 5 | PSA10, PSA14S, PSA6N, PSA1, PSA2 |
| BI, FM1000 | 4 | PSA14N, PSA2, PSA4, PSA7 |
| RHmin, WndSpd | 4 | PSA14S, PSA3, PSA4, PSA5 |
| BI, RHmin | 3 | PSA14S, PSA14N, PSA7 |
| BI, LS | 3 | PSA11, PSA14N, PSA5 |
| BI, WndSpd | 3 | PSA11, PSA12, PSA6S |
| DOY, Tmax | 3 | PSA1, PSA3, PSA9 |
| FM100, WndSpd | 2 | PSA1, PSA8 |
| BI, Precip | 2 | PSA2, PSA6S |
| BI, FM100 | 2 | PSA2, PSA8 |
| LS, WndSpd | 2 | PSA11, PSA5 |
| Precip, WndSpd | 2 | PSA13, PSA9 |
| FM1000, WndSpd | 2 | PSA10, PSA13 |
| ERC, Tmax | 1 | PSA5 |
| FM1000, RHmin | 1 | PSA7 |
| FM1000, VPD | 1 | PSA6N |
| BI, RHmax | 1 | PSA10 |
| Tmax, WndSpd | 1 | PSA12 |
| FM1000, RHmax | 1 | PSA10 |
| BI, Tmin | 1 | PSA14N |
| FM100, Tmax | 1 | PSA1 |
| Precip, RHmin | 1 | PSA4 |
| ERC, WndSpd | 1 | PSA3 |
| Precip, Tmax | 1 | PSA9 |
| VPD, WndSpd | 1 | PSA6N |
| FM1000, Tmax | 1 | PSA13 |
| DOY, FM1000 | 1 | PSA13 |
| FM1000, Tmin | 1 | PSA11 |
| DOY, WndSpd | 1 | PSA12 |
| Tmax, VPD | 1 | PSA6N |
| RHmin, Tmax | 1 | PSA14S |
| DOY, RHmin | 1 | PSA4 |
| FM1000, Precip | 1 | PSA7 |
| RHmin, SPH | 1 | PSA3 |

Table 17. Contains the pair of variables interacting and how often that occurs for human caused large wildfires across all PSAs.

| Interaction | Count | PSA(s) |
|---|---|---|
| Tmax, WndSpd | 4 | PSA4, PSA5, PSA7, PSA8 |
| BI, FM1000 | 4 | PSA10, PSA11, PSA14N, PSA6N |
| BI, RHmin | 4 | PSA10, PSA11, PSA14N, PSA1 |
| BI, LS | 4 | PSA14N, PSA1, PSA2, PSA9 |
| LS, Tmax | 4 | PSA14S, PSA2, PSA3, PSA8 |
| DOY, FM1000 | 3 | PSA10, PSA6N, PSA6S |
| FM1000, WndSpd | 3 | PSA11, PSA13, PSA8 |
| BI, Tmax | 2 | PSA2, PSA7 |
| ERC, WndSpd | 2 | PSA14S, PSA3 |
| Precip, Tmax | 2 | PSA3, PSA7 |
| FM1000, Tmax | 2 | PSA10, PSA4 |
| RHmin, Tmax | 2 | PSA5, PSA7 |
| DOY, RHmin | 2 | PSA4, PSA9 |
| DOY, Tmax | 2 | PSA6S, PSA9 |
| FM100, WndSpd | 1 | PSA13 |
| ERC, Tmax | 1 | PSA14S |
| DOY, WndSpd | 1 | PSA12 |
| LS, SPH | 1 | PSA5 |
| FM100, FM1000 | 1 | PSA13 |
| FM1000, LS | 1 | PSA12 |
| RHmin, WndSpd | 1 | PSA4 |
| LS, RHmin | 1 | PSA3 |
| FM100, LS | 1 | PSA2 |
| BI, Tmin | 1 | PSA14N |
| RHmin, Tmin | 1 | PSA6N |
| BI, WndSpd | 1 | PSA12 |
| FM1000, Tmin | 1 | PSA11 |
| Tmin, WndSpd | 1 | PSA1 |
| BI, DOY | 1 | PSA5 |
| FM100, Precip | 1 | PSA13 |
| DOY, LS | 1 | PSA1 |
| Precip, Tmin | 1 | PSA6N |

Table 18. Contains the pair of variables interacting and how often that occurs for all large wildfires across all PSAs.

## Chapter 5.

## Dissertation Summary, Conclusions, and Recommendations

Chapter 2 examines the verification of spot forecasts and accepts that this analysis represents the beginning of addressing accuracy requirements, improvements, and verification statistics for both surface and upper air elements (OFCM 2007, 2011 and NOAA SAB 2008). Echoing Lammers and Horel (2014) spot forecast verification necessitates a more nuanced approach than just aggregating statistics. Forecasters and end users should develop a framework that allows flexibility in deciding how and what to verify from spot forecasts. The consistency of the information provided by the spot forecasts needs to be improved, which will help in determining acceptable accuracy thresholds or requirements for spot forecasts.

Chapter 3 examines the relationship of atmospheric circulation to NAM onset and increased wildfire activity across the SWA using SOM. Resulting SOM map types also showed the transition to, during, and from the NAM. Northward and eastward displacements of the subtropical ridge (i.e., four-corners high) over the SWA were associated with NAM onset (Carleton et al. 1990; Higgins et al. 1999; Cerezo-Mota et al. 2011), and a suppressed subtropical ridge and breakdown of the subtropical ridge map types over the SWA were associated with increased wildfire activity (Werth et al. 2011). Four general atmospheric pattern progressions emerge when examining the evolution of MTs around busy days and events for the SWA: 1) zonal or southwest flow preceding ridging; 2) zonal or southwest flow transitioning into ridging followed by a return to zonal or southwest flow; 3) persistent ridging followed by zonal or southwest flow; and

4) fluctuation between suppressed and amplified ridging over the SWA with the ridge axis exhibiting east-west movement.

Chapter 4 used BRT to model wildfire occurrence by type and size for each PSA in the SWA. The BRT models show predictive potential and demonstrate robustness across all wildfire types and PSAs including on days with wildfires. Fuels (FM1000, FM100, ERC) are the most important predictor(s) when considering all wildfires regardless of type with lightning demonstrating strong predictive influence. More predictors become important when examining large wildfires with a combination of fuels (i.e. FM1000, ERC) and atmospheric predictors (i.e., LS, Tmax) driving large wildfire occurrence indicating more conditions need to align to support and sustain large wildfires. Predictor interactions demonstrate thresholds for wildfire occurrence (i.e., LS and FM1000) and important critical fire weather conditions including ignition and spread (i.e., LS and BI), dry thunderstorms (i.e., LS and precipitation), and a combination of hot, dry, and windy conditions (i.e., Tmax, RHmin, WndSpd).

While this dissertation follows a non-traditional format, each of the three chapters provides original and valid contributions to the field of atmospheric sciences. Additionally, all of the research here is directly applicable to operational fire meteorologists and mangers. The results of chapters 3 and 4 provide decision support information and improve understanding of atmospheric processes associated with NAM and their impact on wildfire activity. The improved understanding benefits operational fire meteorologists and managers with the identification of atmospheric patterns associated with increased wildfire activity and daily PSA wildfire occurrence modeling output, which improves planning and logistical strategies. Chapter 2 provides an

exploratory verification of spot forecasts and makes recommendations that should improve spot forecasts and spot forecast verification.