USFS-CSU Joint Venture Agreement Phase 2 (2019-2021): Developing a Gridded Model for Probabilistic Forecasting of Wildland-Fire Ignitions Across the Lower 48 States

Final Report

Ned Nikolov, Ph.D.
E-mail: ned_nikolov@firenet.gov

Phillip Bothwell, PhD.
John Snook, Ph.D.

January, 2022

1. Introduction

The National Predictive Services (NPS) asked the USFS Rocky Mountain Center for Fire-Weather Intelligence (RMC) as a part of the *Fire, Fuel, and Smoke Science Program* (FFS) at the USFS Rocky Mountain Research Station (RMRS) to assist with the development of a system of statistical models for predicting the ignition probability (chance of start) & growth potential of wildfires on a national grid using NWS meteorological forecast data as input. The development of this gridded system of predictive equations was envisioned to proceed in 3 stages (phases).

Phase 1 (Oct. 2017 – Dec. 2018) developed and partially validated a high-resolution spatial model for forecasting the probabilities of cloud-to-ground lightning flashes as a function of 3-D fields of atmospheric parameters stretching from the surface to the tropopause. Lightning is known to account on average for 60% of the wildland-fire area burned annually in the USA. Hence a lightning-forecast model is a prerequisite to developing a robust fire-ignition prediction application.

Phase 2 (2019-2021) completed the verification of the lightning forecast model using independent meteorological data from NARR and GFS for 2017 and 2018. Also developed were statistical (logistic) equations for predicting the probabilities of ignition of lightning-caused, human-caused and any-cause wildfires on a uniform grid across the lower 48 states. The ignition-probability forecast model was validated against independent fire-occurrence data from 2017 and 2018 as well. This addressed the need of Predictive Services and the fire-management community for forecasting the chance of wildfire start at high spatial resolution across Conterminous USA (ConUS) going out to 7-10 days based on long-term climatology and current forecast weather conditions.

The output from Phase 2 of this project is a set of monthly logistic equations capable of predicting daily probabilities of one or more wildfire ignitions due to either lightning, human factors, or any-cause on a national 20-km grid up to 10 days in advance using NWS numerical weather forecasts as input. Real-time operational wildfire-ignition forecasts are now available at the RMC Website (click on "*7-Day Wildfire Start Forecast*" under the *Forecast Applications* Panel).

2. Method

This section describes the datasets and computational procedures utilized to develop and test the system of predictive wildfire-ignition equations for the lower 48 states.

2.1. Gridded Data Sets

Three 25-year long grided records of weather reanalysis fields, observed lightning flashes and wildfire occurrences over ConUS were combined to produce *daily* datasets (from 12:00 to 12:00 UTC) of 20-km horizontal resolution. The analysis was limited to the lower 48 states (Fig. 1). Grid cells listed as water, lakes, or ocean were excluded from the model development.

a. The NOAA North American Regional Reanalysis (NARR) containing 3-D fields from 00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00, and 21:00 UTC for the period January 1990 – December 2018 were previously downloaded in GRIB2 format and archived on the RMC machines.

b. U.S lightning data for the lower 48 states were obtained via BLM and gridded at 3-hour increments for the period from January 1990 through December 2018.

c. Global Forecast System (GFS) 3-h forecast data fields in GRIB2 format (0.25 x 0.25 degree through 10 days) were archived for the peak fire season (May – September) of 2018 to verify the lightning forecast equations. GFS forecast fields were used to produce real-time lightning forecasts beginning in mid-summer 2019 and continuing to the present (visit the RMC Website for a link to the experimental lightning forecast). Additionally, GFS forecast data were also utilized to produce experimental wildfire ignition-probability forecasts beginning in June of 2021.

d. Wildfire Occurrence Data from 1992 to 2018 provided by Karen Short's database.

e. Monthly Leaf Area Index (LAI) from WRF, NOAA's Evaporative Demand Drought Index (EDDI), digital maps of vegetation cover types and percent vegetation cover.

Figure 2 shows a high-resolution map of vegetation types across the ConUS used in our analysis. Two different fuel maps containing 13 and 17 land-cover categories, respectively were included in the analysis along with the percent vegetation cover in each grid cell to provide a comprehensive fire-environment dataset. However, the combined contribution of these vegetation layers to the prediction scheme turned out to be very small.

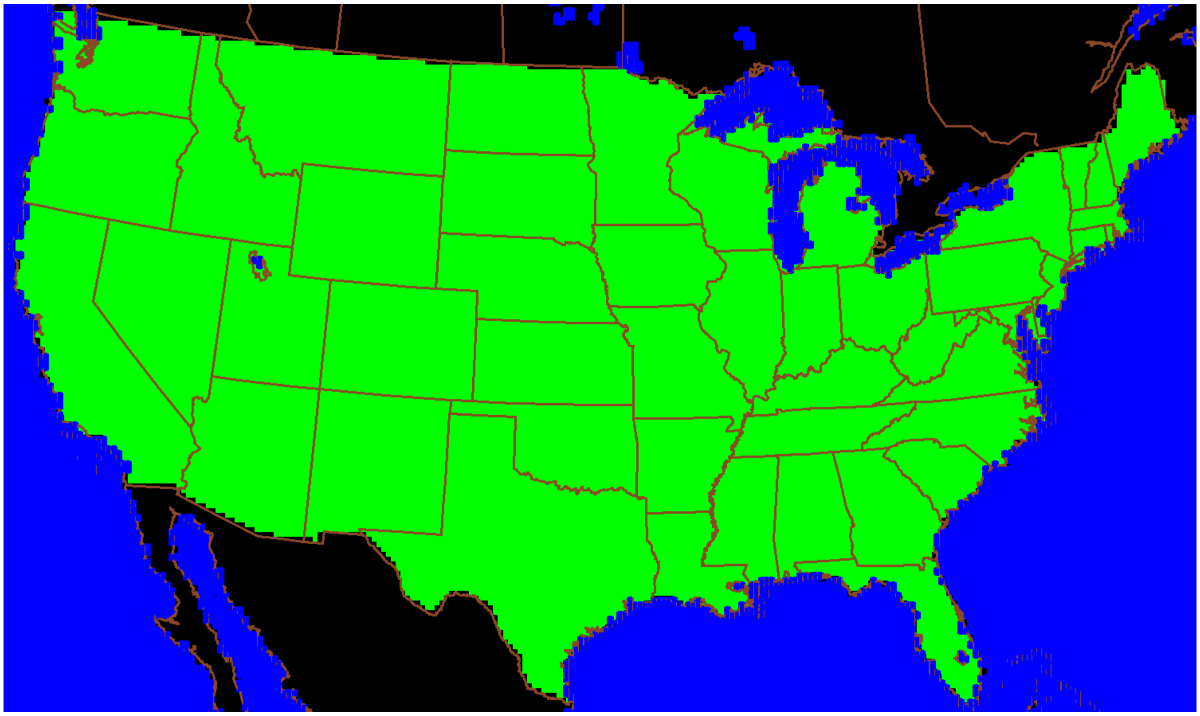


Figure 1. Full domain of NARR and GFS 20-km resolution data layers, and datasets used to develop and test the predictive wildfire-ignition equations.

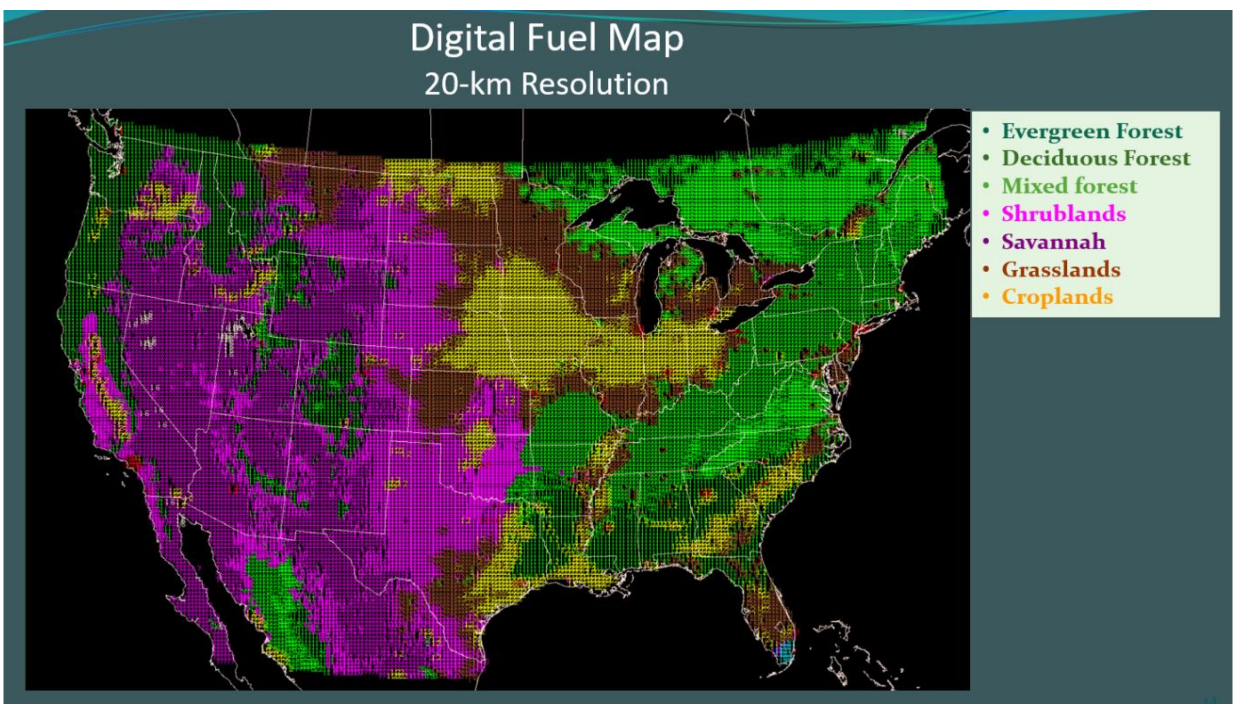


Figure 2. Gridded fuel types across the U.S. (20 km resolution).

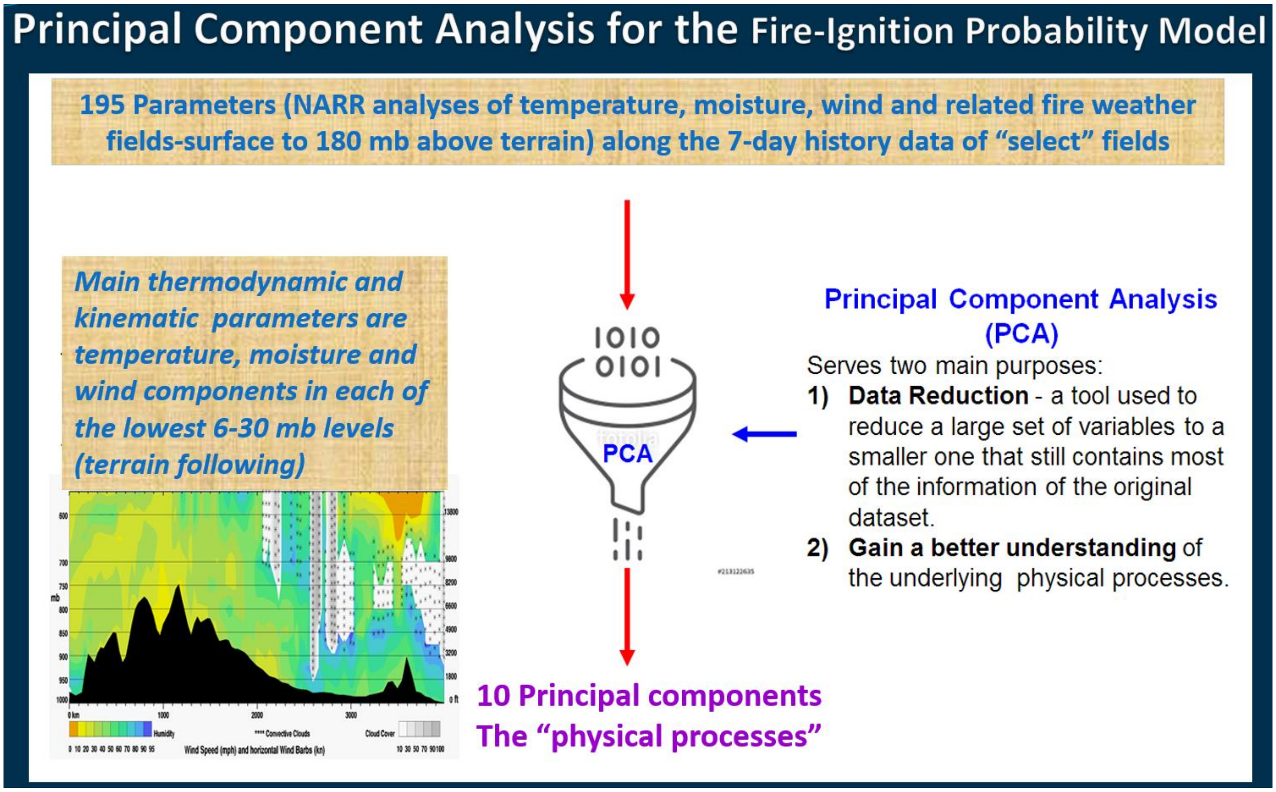


Figure 3. Schematic representation of the Principal Component Analysis (PCA) employed to identify significant predictors of wildfire ignitions. Meteorological data for the period 1992 – 2018 were obtained from the lowest terrain following 180 mb layer of the NARR data set.

2.2.Statistical Approach and Computational Procedures for Ignition Probabilities

We employed a statistical method that is similar to the one used in the development of the lightning forecast model (see our 2019 Report). It involved the application of Principal Component Analysis (PCA) with orthogonal rotation used to reduce the initial set of 195 meteorological driving variables from the lower atmosphere of the NARR dataset to a smaller subset of statistically significant fire-ignition predictors. Figure 3 schematically illustrates this PCA procedure. The reduced set of predictors was then subjected to logistic regression analyses to yield equations for computing the probabilities of one or more wildfire ignitions.

The resulting predictive equations can be applied not just to NARR fields but the output from any numerical weather prediction model as well, such as GFS and WRF to forecast ignition probabilities. The "R" statistical package, an open-source software, was utilized to perform PCA and the logistic regressions required to derive the final set of equations. Resampling the original datasets to a common national grid of 20-km horizontal resolution was done using the GEMPAK (GEneral Meteorology PAcKage) software jointly developed by NASA and Unidata.

Data processing went through the following procedure:

- Rescale original NARR fire-weather parameters (e.g., temperature, moisture, wind speed/direction etc.) to a uniform 20-km resolution grid covering the lower atmosphere (from the surface to 180 mb above the surface, terrain following) using GEMPAK.

- Aggregate the at 3-hour NARR fire-weather meteorological fields (such as maximum temperature, minimum humidity, max wind speed etc.) covering a period of 29 years (i.e., from 1990 through 2018) to 24-hour *daily totals* representing the time from 12:00 to 12:00 UTC.

- Generate a *daily* lightning climatology for the 1990 - 2016 period using NARR data and Bothwell's Cloud-to-Ground Lightning Model, and a fire-ignition climatology for the 1992 – 2016 period covering the lower 48 states using Karen Short's updated Fire-Occurrence dataset.

- Incorporate daily Evaporative Demand Drought Index (EDDI), monthly Leaf Area Index (LAI) and vegetation maps into the NARR-based predictor dataset.

- Perform PCA with varimax/orthogonal rotation on NARR data using 195 potential meteorological drivers (see Table A1 in Appendix for a full list of initial predictors) with added fire and lightning climatological records to identify a subset of the best wildfire-ignition predictors. This resulted in 10 Principal Components (PCs) listed in Table 1 and Table A3 (in Appendix) used as final predictors in each 24-hour period. Table A2 (in Appendix) shows an example of PC members and their PC loadings for all ignitions between 1992 and 2016 in the month of August. The color coding in Table A2 corresponds to the "grouped" fields of PCs in Table A3. In order to better capture the fire environment, the predictors for ignition probabilities were extracted from the NARR's lowest layer of the model atmosphere (surface to 180 mb above the surface, terrain-following). The 10 PCs (predictors) explained nearly 80% of the total variance of wildfire occurrences in every 24-hour period. The low-level atmospheric moisture fields emerged as the strongest predictor among the 10 PCs explaining 24% of the observed variance.

- Run "R" logistic regression scripts using the 10 Principal Components (PCs) to derive separate predictive equations for calculating probabilities of one or more *lightning-caused* ignition, *human-caused* ignition and *any-cause* ignitions that are representative for a 24-hour period (12:00 to 12:00 UTC each day) in every month during the active fire season (May through September). We found that monthly predictive equations of ignition probabilities better captured changing environmental conditions during the course of a fire season than a set of seasonal equations. Table A4 (in Appendix) shows a sample set of logistic regression coefficients for each of the 10 PCs. Due to scarcity of wildfire ignitions in the winter and fall, the derivation of monthly logistic equations was limited to the active fire season only. Ignition forecasts for the period from October through December employ equations developed for September, while the May equations are applied to the period January – April.

- A typical month is described by three equations, one for each wildfire ignition cause: lightning, human, and any cause. Although regional equations might theoretically be more accurate in describing the fire environment, the relatively small numbers of reported fires per grid cell and in total forced us to combine and use all data for the lower 48 states to derive 3 ignition equations for every month during the active fire season (May – September).

- When run operationally using the Perfect Prognosis method, the ignition-prediction equations are combined with 3-D meteorological fields from GFS as well as lightning and fire-ignition climatologies to produce *daily* ignition probability forecasts (12:00 to 12:00 UTC) at 20-km resolution out to 7 days across ConUS.

Table 1. Top 10 Principal Components (drivers) of daily wildfire-ignition probabilities yielded by PCA. A 24-h day refers to a period from 12:00 to 12:00 UTC.

Parameter	Explained Variance (%)
Lower atmospheric moisture and Haines Index	24
Lower atmospheric temperature	13
Soil Temperature, Leaf Area Index, Fosberg Index, and Hot-Dry-Windy Index	8
V wind component	8
U wind component	7
Lightning climatology and 7-day lightning probability	5
Soil moisture, temperatures, and vegetation	5
Climatology of human-caused fires and all wildfires	4
Climatology of lightning-caused wildfires	3
Evaporative Demand Drought Index (EDDI)	2

3. Results

3.1. Lightning Forecast Verification using Independent NARR and GFS data:

This section presents verification results (not included in our 2019 Report) for the lightning forecast model using independent data fields from NARR and GFS as drivers.

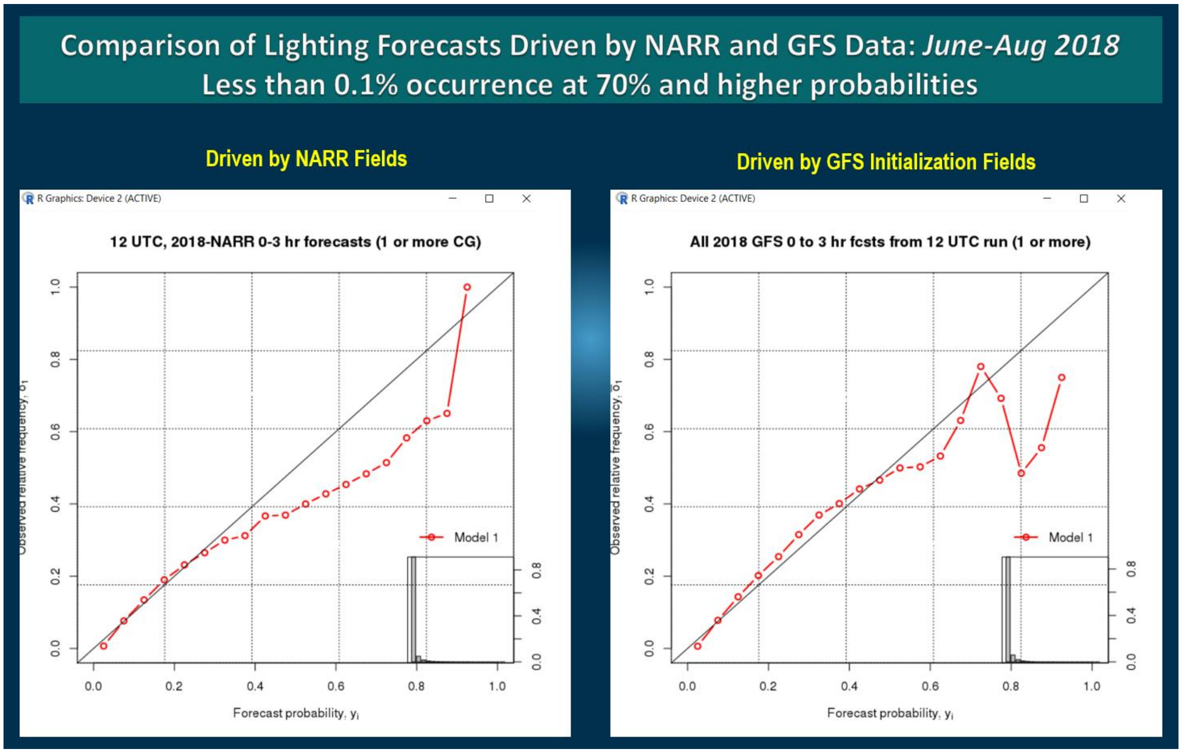


Figure 4. [Reliability diagrams]() for NARR-based (left) and GFS-based (right) lightning forecasts for June, July and August of 2018 (3-hour forecasts based on 12 UTC data).

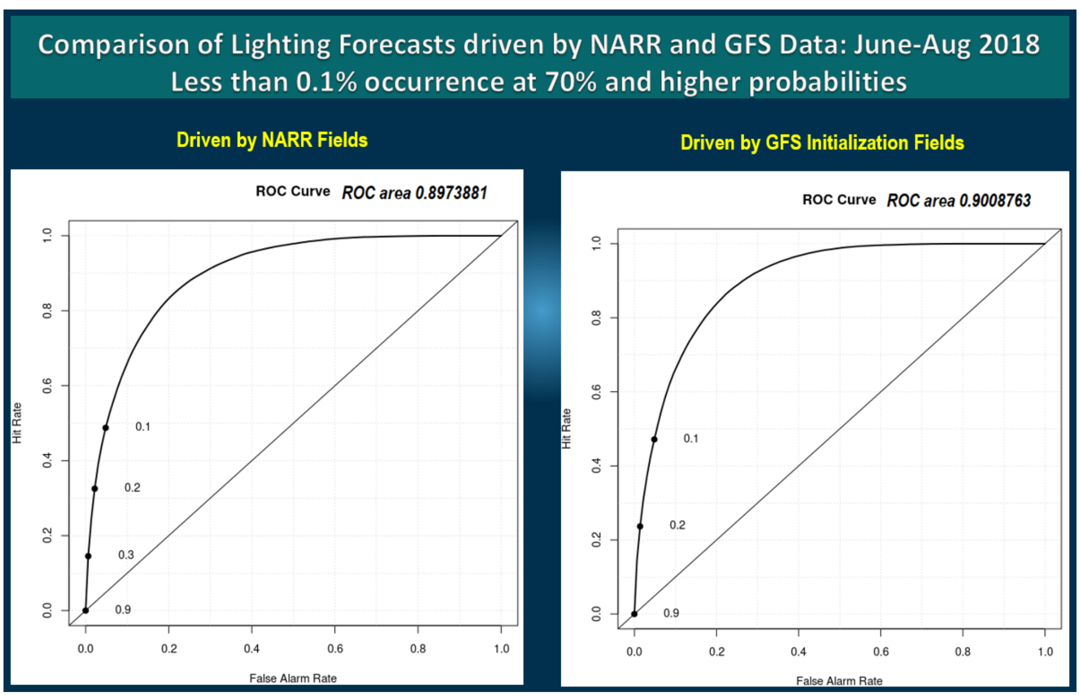


Figure 5. [Receiver Operating Characteristic]() (ROC) curves with Area Under the Curve (AUC) covering the same time period as in Fig. 4.

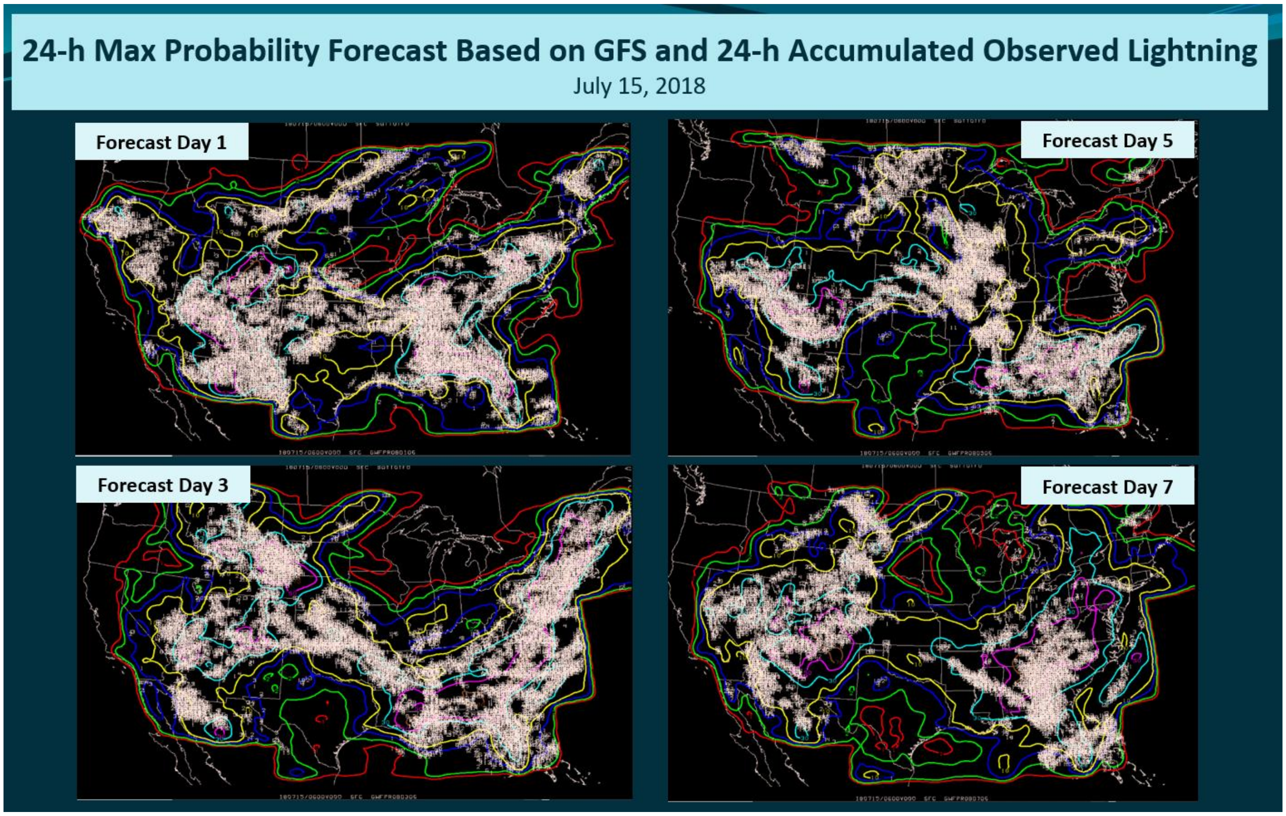


Figure 6. Maximum probability of lightning (color contours) predicted by the model for forecast days 1, 3, 5, and 7 overlayed by observed lightning flashes (numbers in white) over a 24-h period (12:00 to 12:00 UTC) for July 15, 2018**.** Contours delineate: 1% red, 2% green, 5% blue, 10% yellow, 30% cyan, 50% magenta 50%, 70% brown.

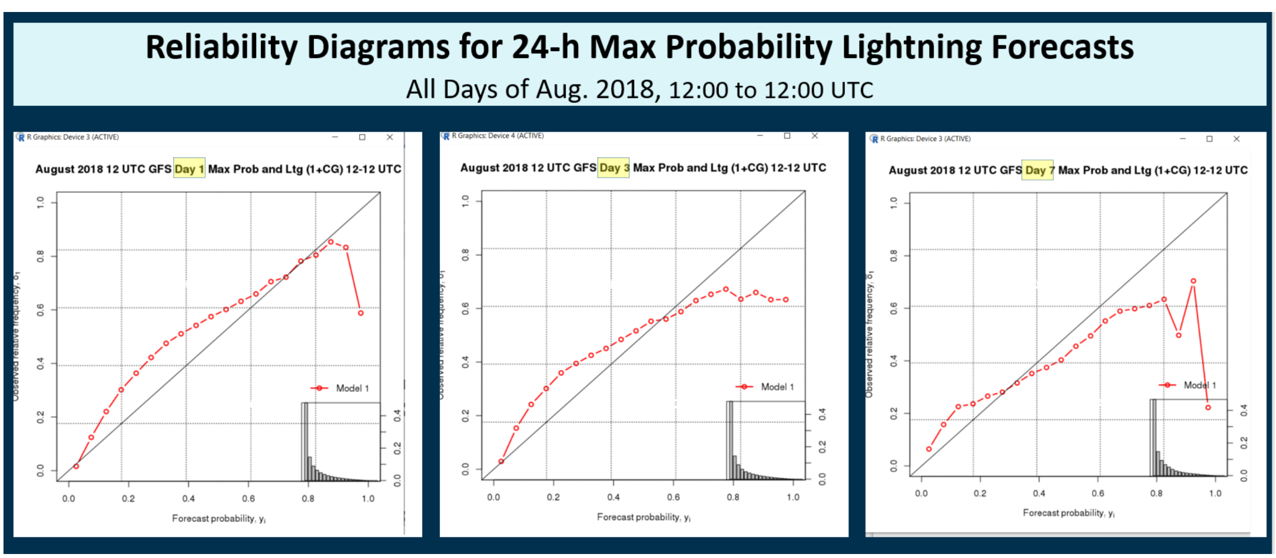


Figure 7. Reliability diagrams of GFS-based maximum-probability daily lightning forecast for days 1,3, and 7.

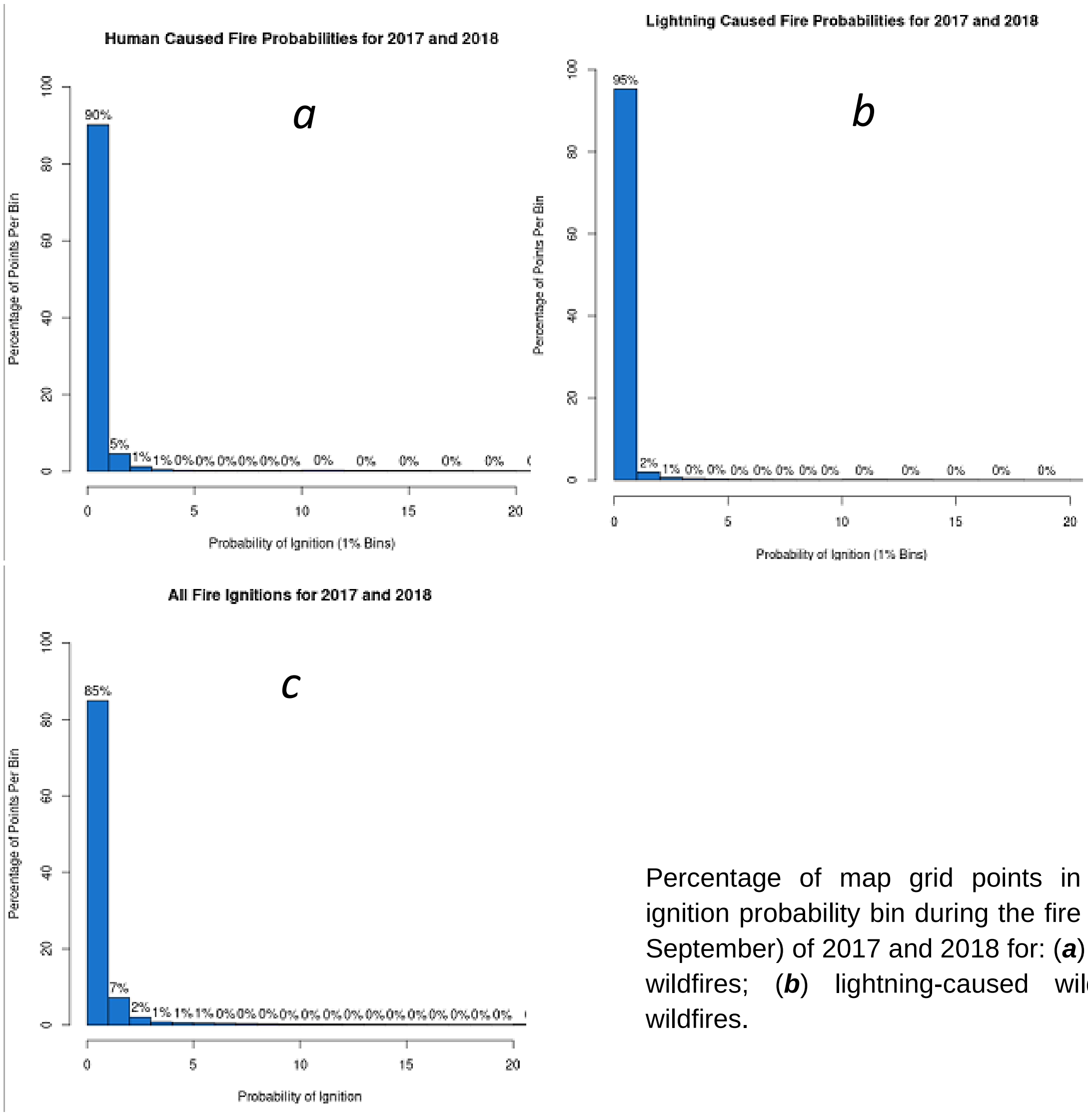
Figure 4 compares side-by-side Reliability Diagrams of all NARR- and GFS-driven 3-h forecasts for June, July and August of 2018 employing independent lightning data. Figure 5 shows the Receiver Operating Characteristic (ROC) diagrams and corresponding Area Under the Curve (AUC) for the same data. Both Figures indicate nearly *outstanding skill* of the lightning prediction model with the GFS-driven forecasts being slightly better than those driven by NARR data. The better score of GFS runs might be due to recent updates implemented to this weather-forecast model.

Figure 6 shows a 24-h accumulation of observed lightning flashes overlayered on contours of daily maximum probability predicted by the model for forecast days 1, 3, 5 and 7 using GFS fields as drivers. Figure 7 displays the corresponding reliability diagrams for forecast days 1, 3 and 7.

3.2. Wildlife-Ignition Probability Forecasts Driven by 2017 and 2018 Independent NARR data

The histograms depicted in Figure 8 (*a* through *c*) show percentages of grid points in each ignition-probability bin for lightning-caused wildfire ignitions, human-caused fire ignitions and any-cause wildfire ignitions. Most grid points (map pixels) have ignition probabilities below 1.0%, which generally is plotted as "*no ignitions"* on the operational product maps.

It is important to note that the probability of any fire ignition is much lower than that for cloud-to-ground lightning flashes. The lightning dataset contains approximately 25 million flashes per year. For wildfires, the *entire* time period from 1992 to 2018 (27 years) had between 2.0 and 2.5 million starts averaging less than 100,000 fire ignitions per year. This occurrence is several orders of magnitude smaller than lightning. Thus, the wildfire-ignition forecast model is essentially predicting *rare events*. As a result, most NARR-driven ignition probabilities are below 10%. To further illustrate the low range of ignition probabilities, Fig. 9 displays an example of the climatological (historical) probabilities of fire ignition (as percent chance) for July 15th over the period 1992 – 2016.



Percentage of map grid points in each forecast ignition probability bin during the fire season (May – September) of 2017 and 2018 for: (***a***) human-caused wildfires; (***b***) lightning-caused wildfires; (***c***) all wildfires.

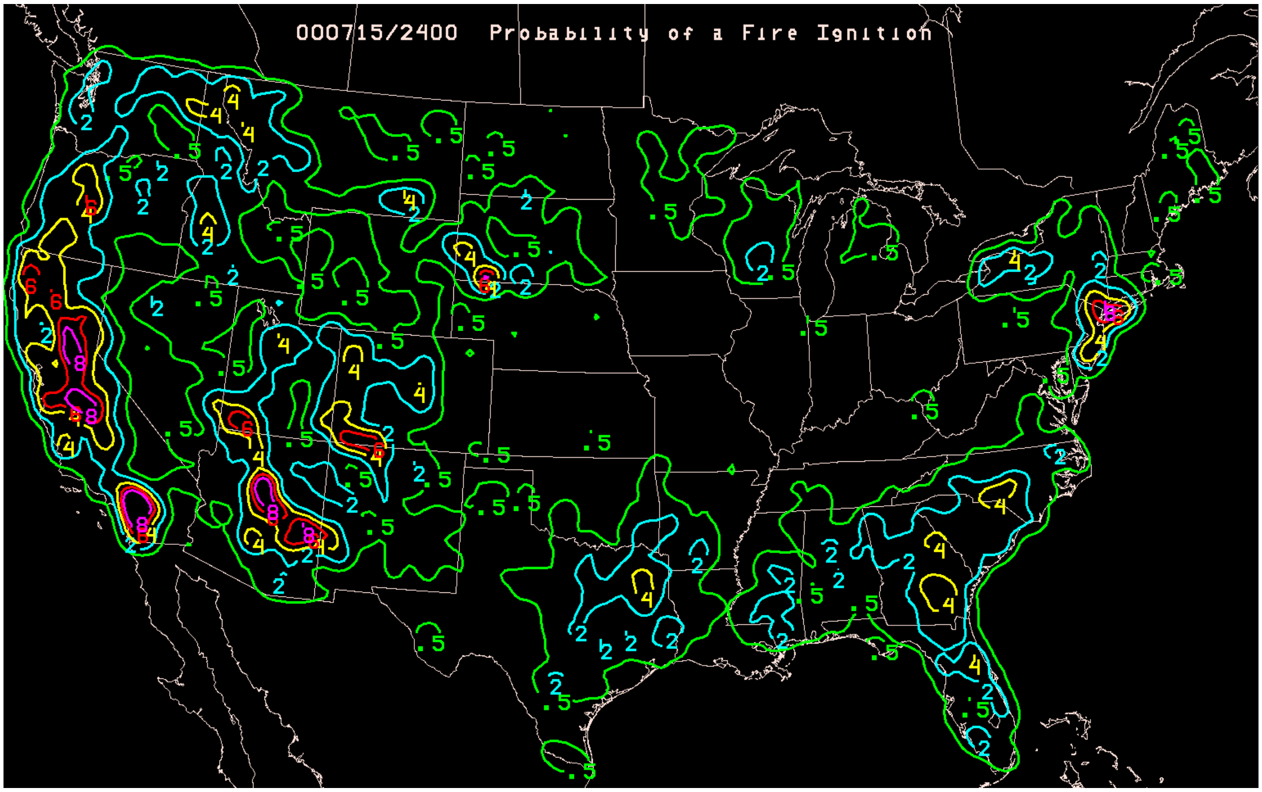


Figure 9. Sample climatological (historical) probability of wildfire ignition by all causes for July 15 over the period 1992 – 2016. Color contours show the chance for ignition: 0.5% green, 2% cyan, 4% yellow, 6% red, 8% magenta.

Figure 10 (*a* through *k*) shows the progression of forecast ignition probabilities driven by independent NARR data in 15-day increments from May through September in 2017. The forecast contours are: 1% green, 10% blue, 20% yellow and >= 30% red. Overlayed on the forecasts (in white) are the number of observed ignitions in each 20-km grid cell.

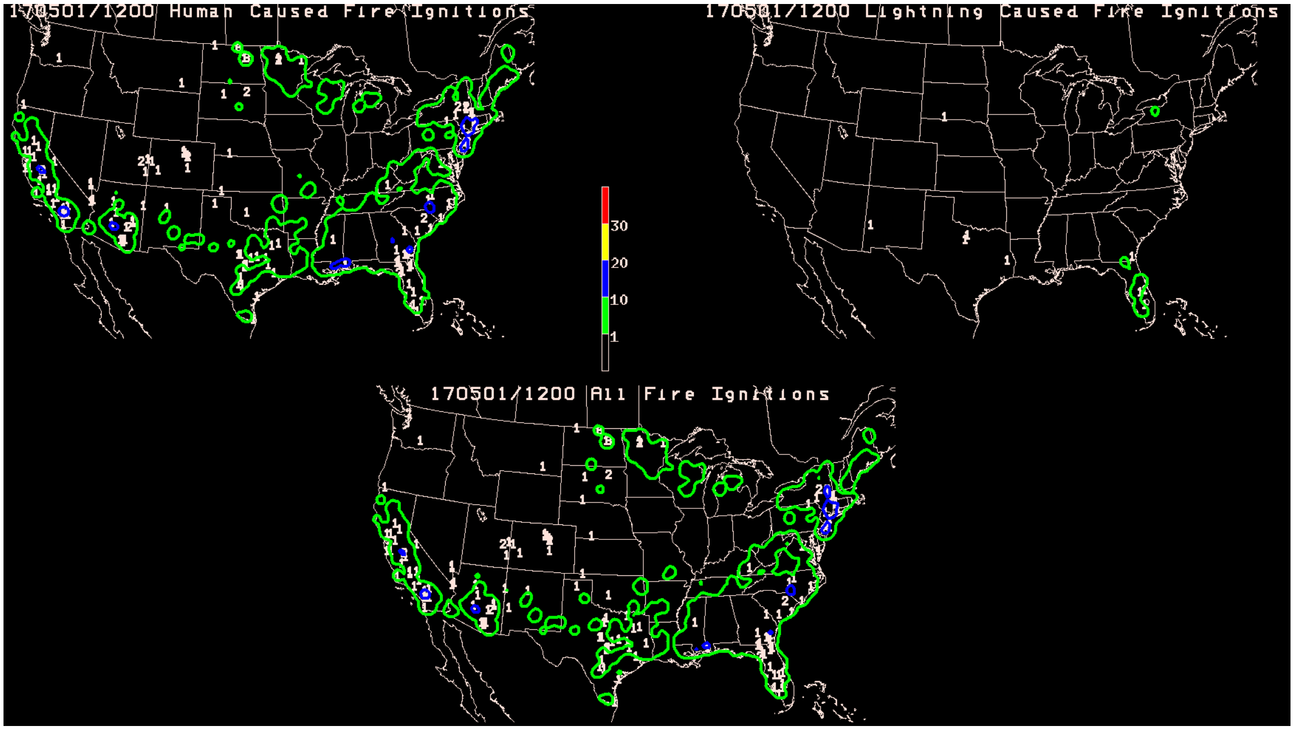


Figure 10. (*a*) Forecast ignition probabilities (color contours) and observed wildfire ignitions (white numbers) on May 1, 2017. Upper left plot: *human-caused fires*; Upper-right plot: *lightning-caused fires*; Center-lower plot: *all fires*.

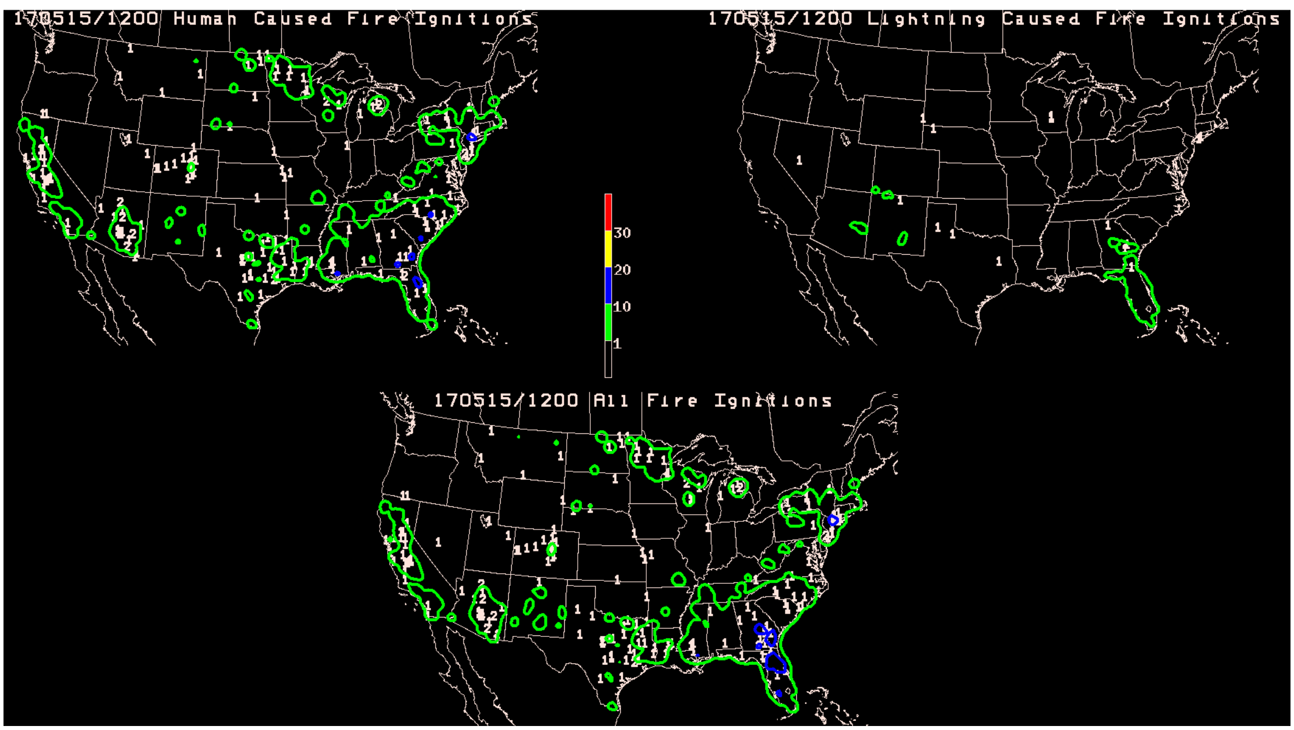


Figure 10. (*b*) The same as Fig. 10a but for May 15, 2017.

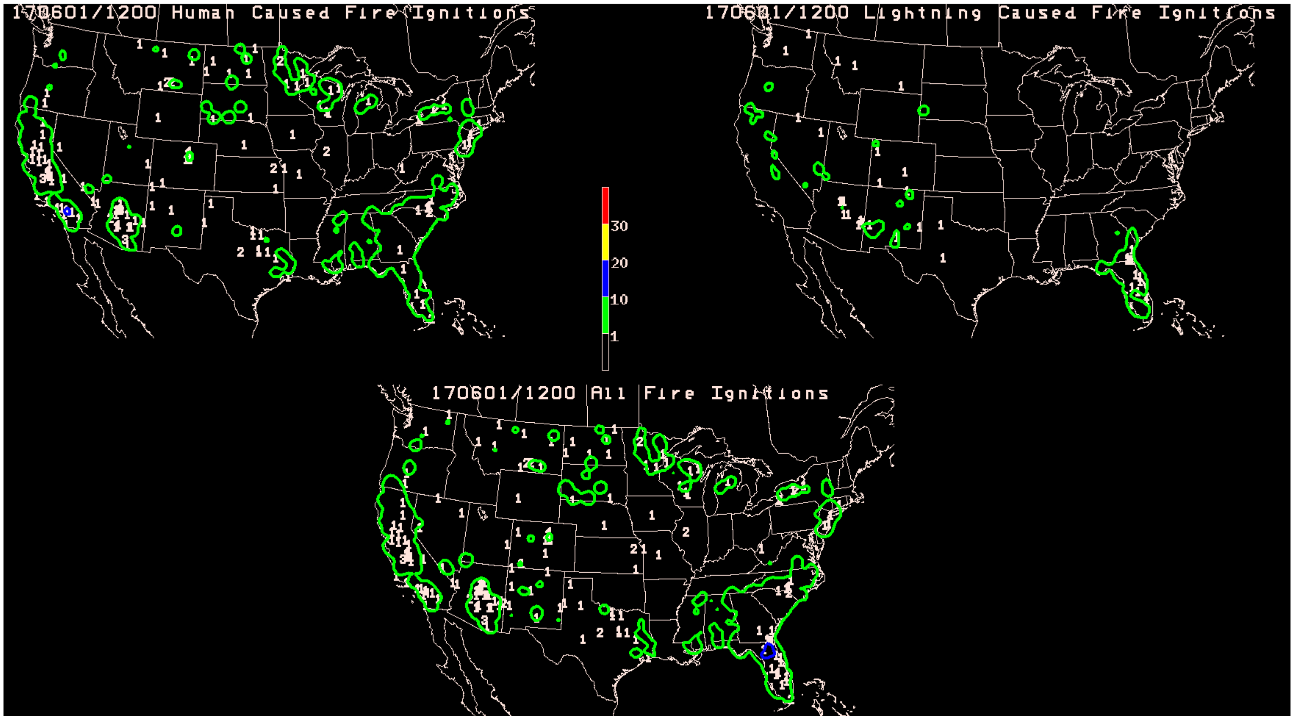


Figure 10. (*c*) The same as Fig. 10a but for June 01, 2017.

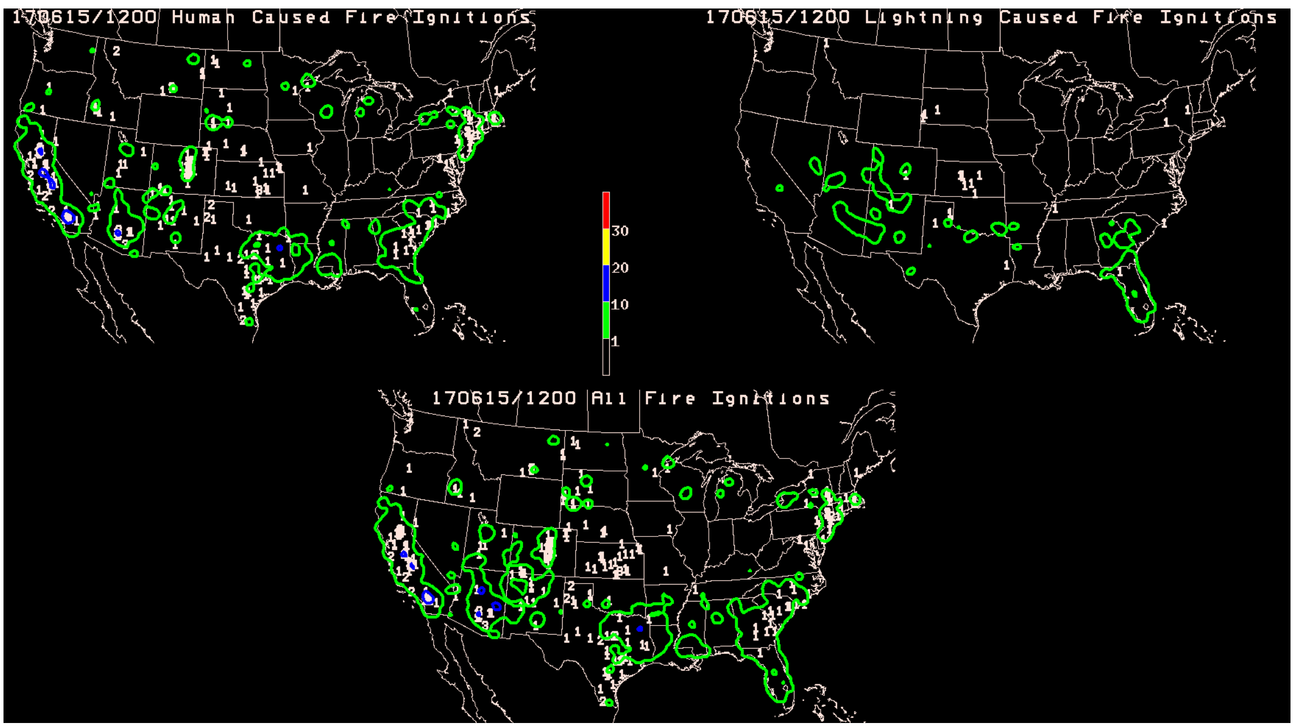


Figure 10. (*d*) The same as Fig. 10a but for June 15, 2017.

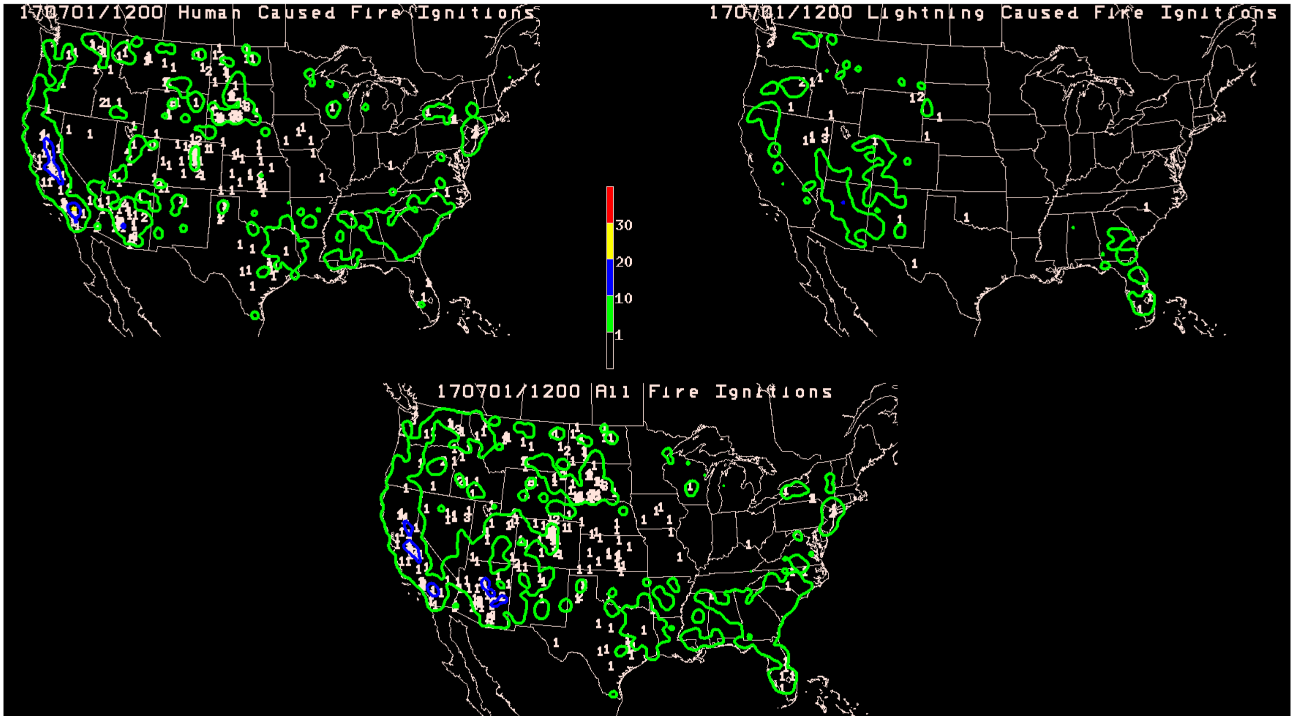


Figure 10. (*e*) The same as Fig. 10a but for July 01, 2017.

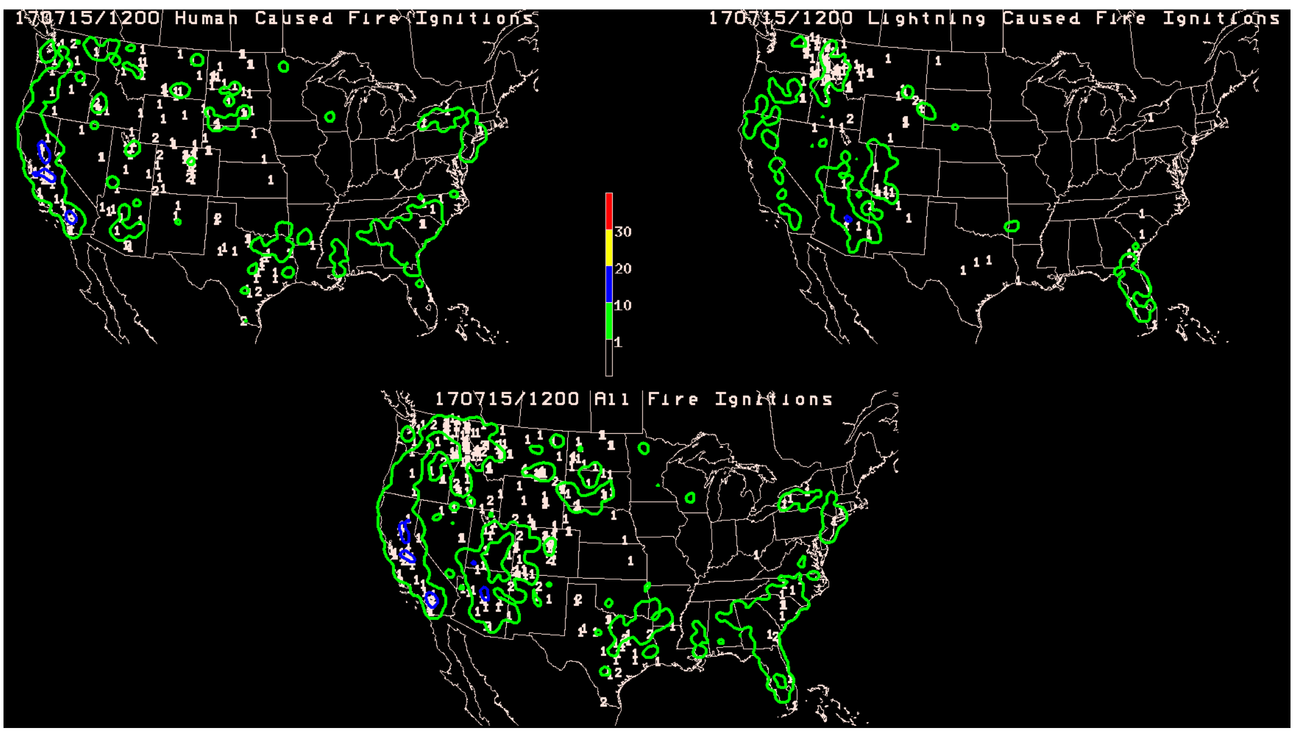


Figure 10. (*f*) The same as Fig. 10a but for July 15, 2017.

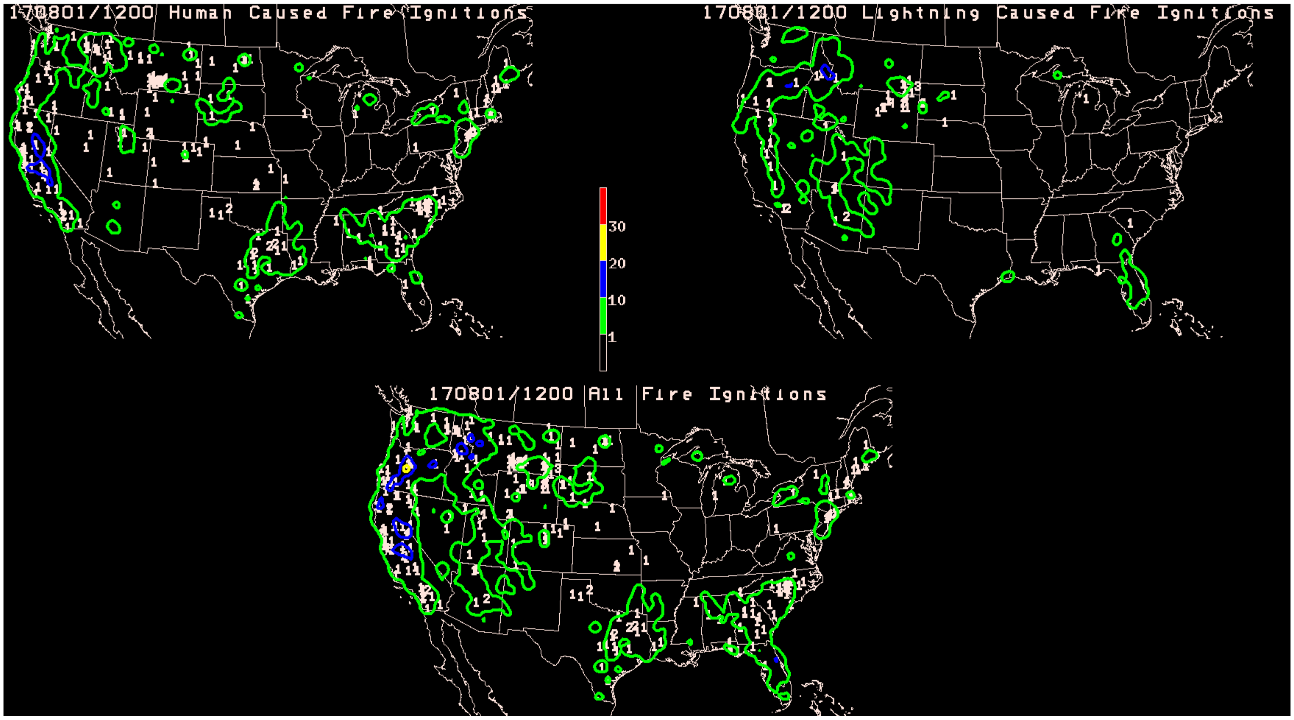


Figure 10. (*g*) The same as Fig. 10a but for August 01, 2017.

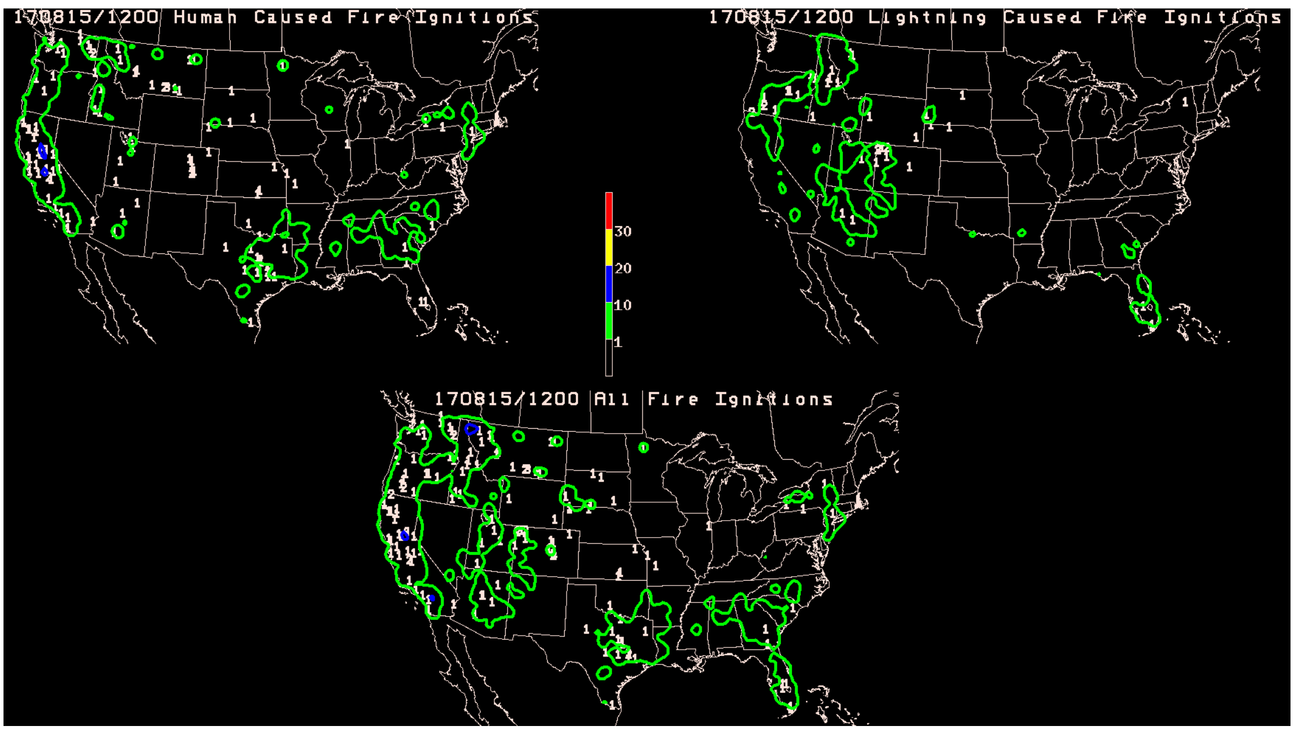


Figure 10. (*h*) The same as Fig. 10a but for August 15, 2017.

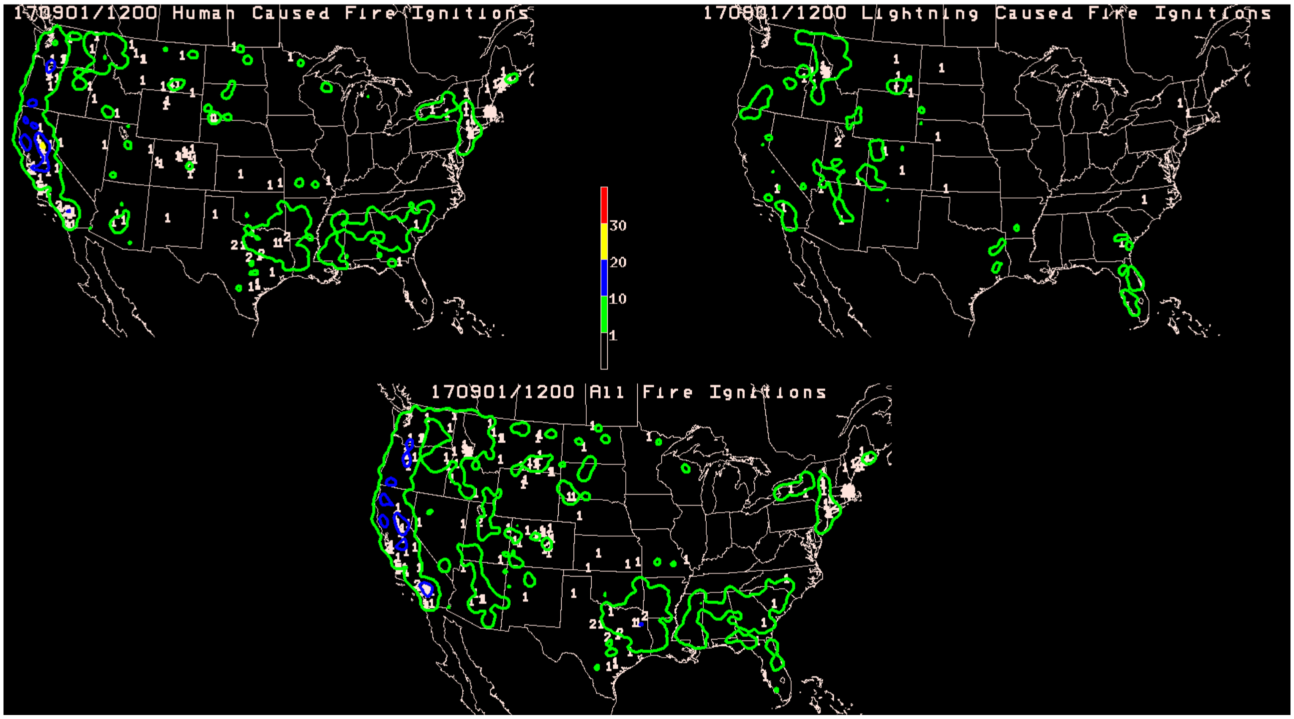


Figure 10. (*i*) The same as Fig. 10a but for September 01, 2017.

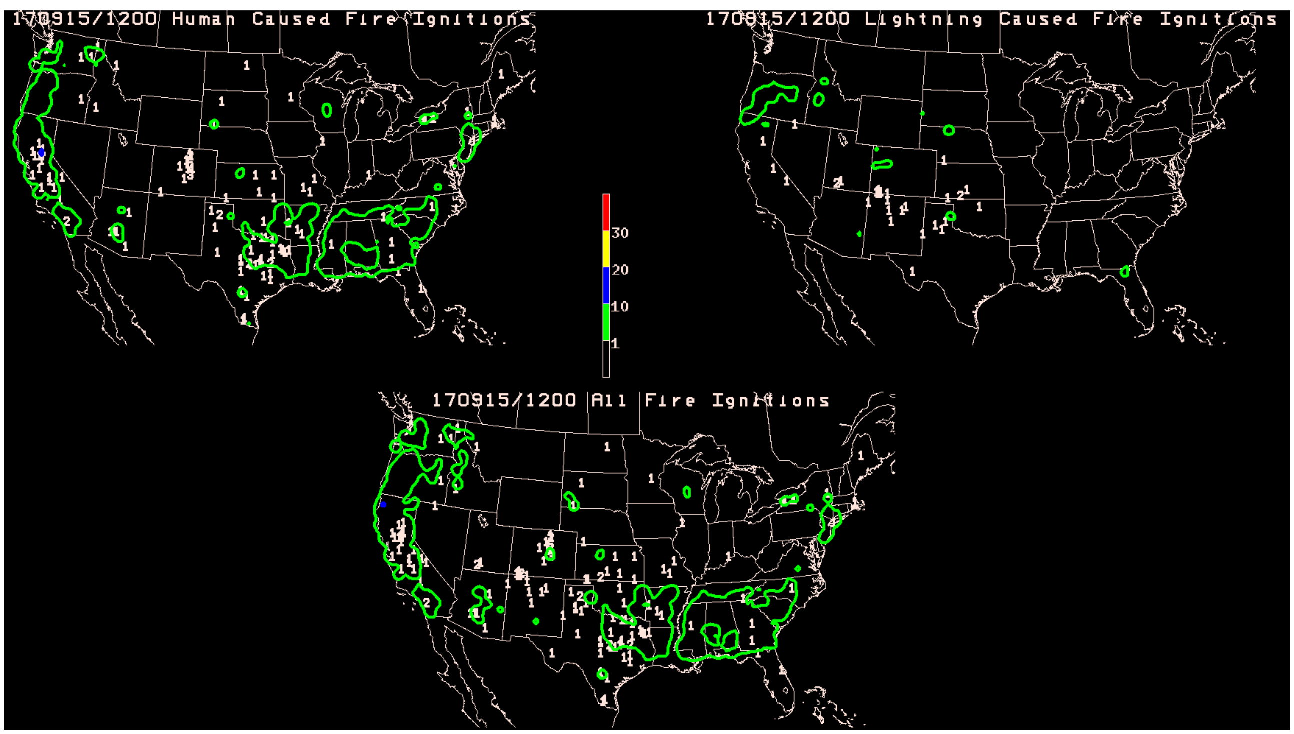


Figure 10. (*j*) The same as Fig. 10a but for September 15, 2017.

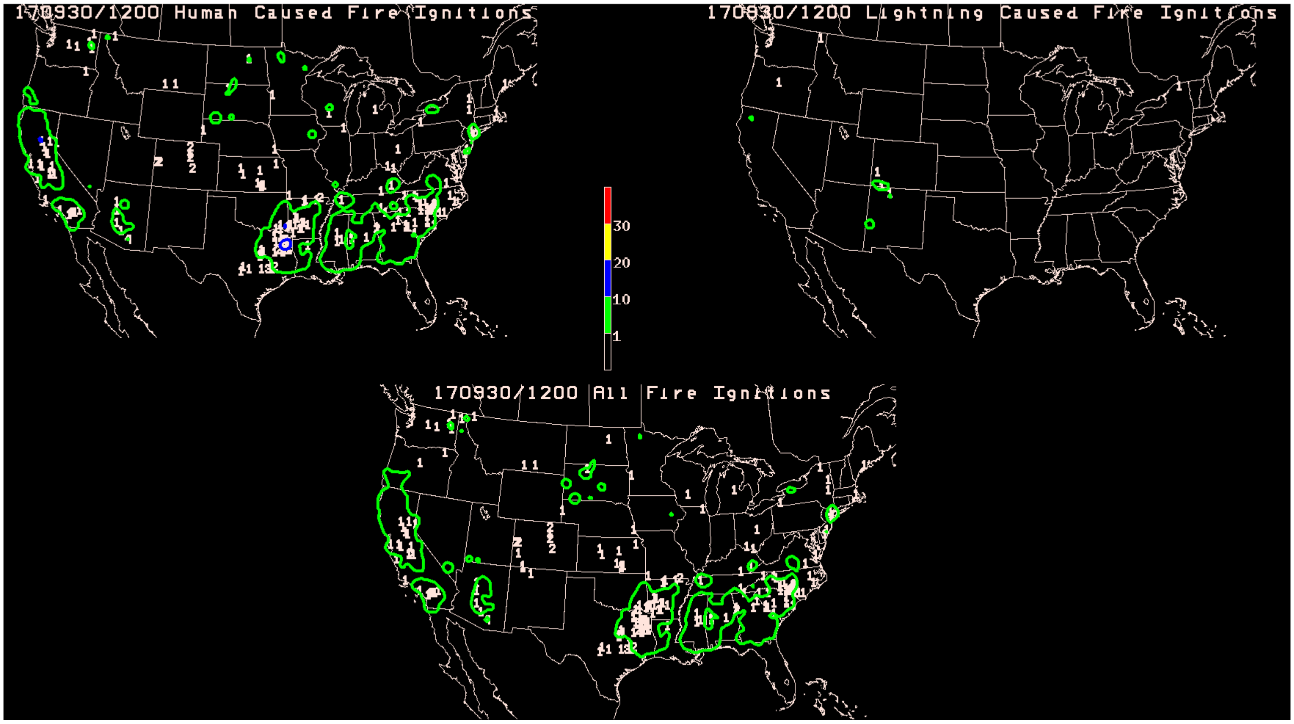


Figure 10. (*k*) The same as Fig. 10a but for September 30, 2017.

Figure 11 (*a* through *k*) shows the progression of forecast ignition probabilities driven by independent NARR data in 15-day increments from May through September in 2018. As before, the forecast contours are: 1% green, 10% blue, 20% yellow, and >= 30% red. Overlayed on the forecasts (in white) are the number of observed ignitions in each 20-km grid cell.

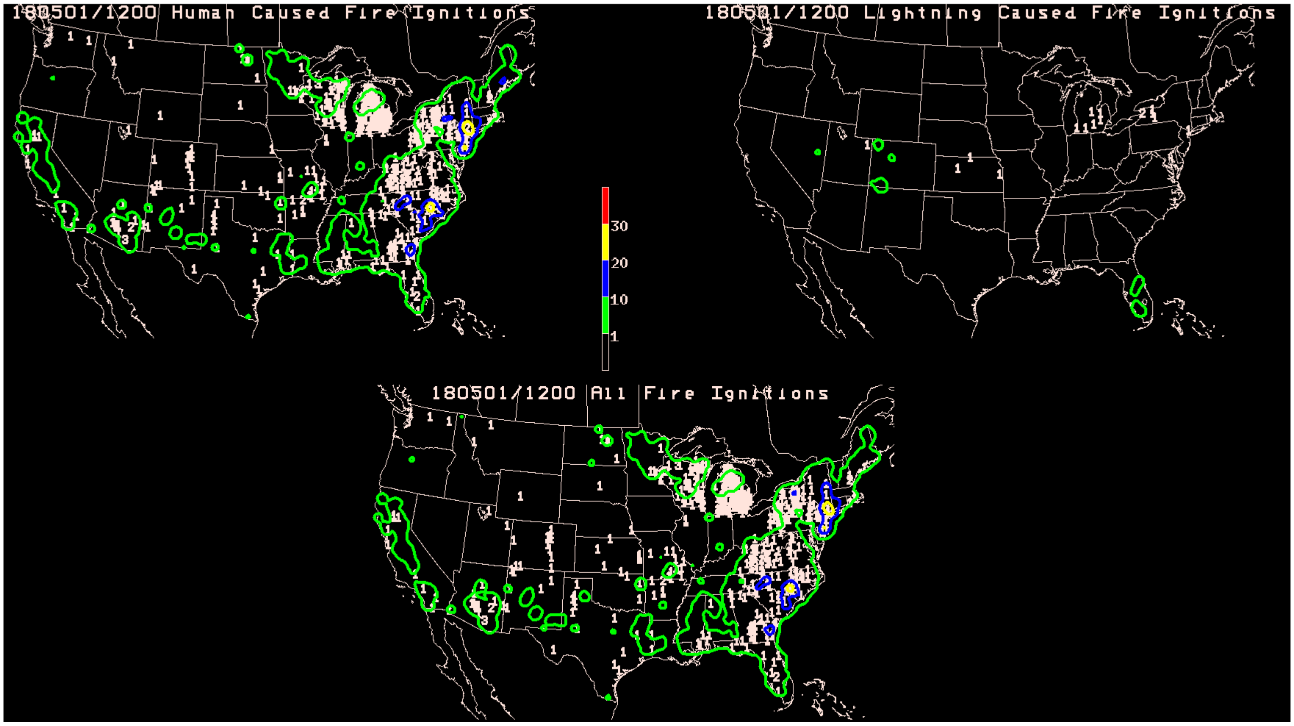


Figure 11. (*a*) Forecast ignition probabilities (color contours) and observed wildfire ignitions (white numbers) on May 1, 2018. Upper left plot: *human-caused fires*; Upper-right plot: *lightning-caused fires*; Center-lower plot: *all fires*.

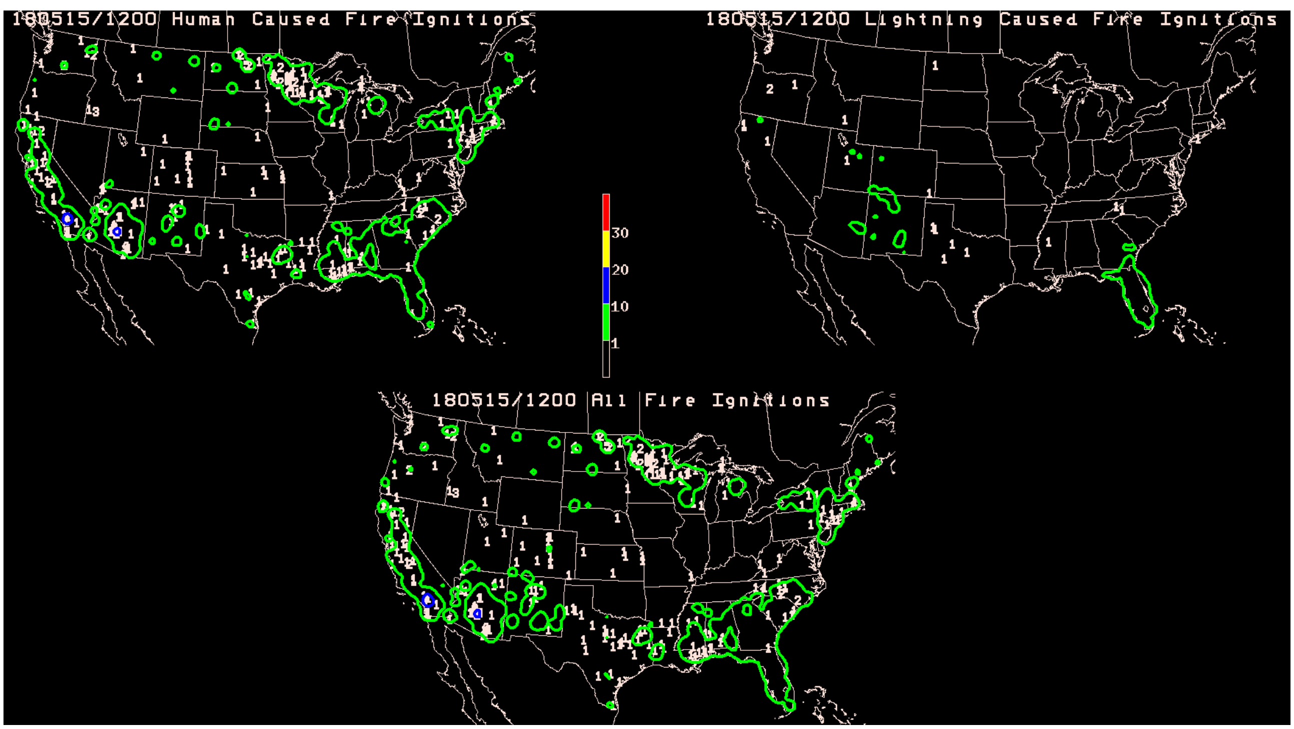


Figure 11. (*b*) The same as Fig. 11a but for May 15, 2018.

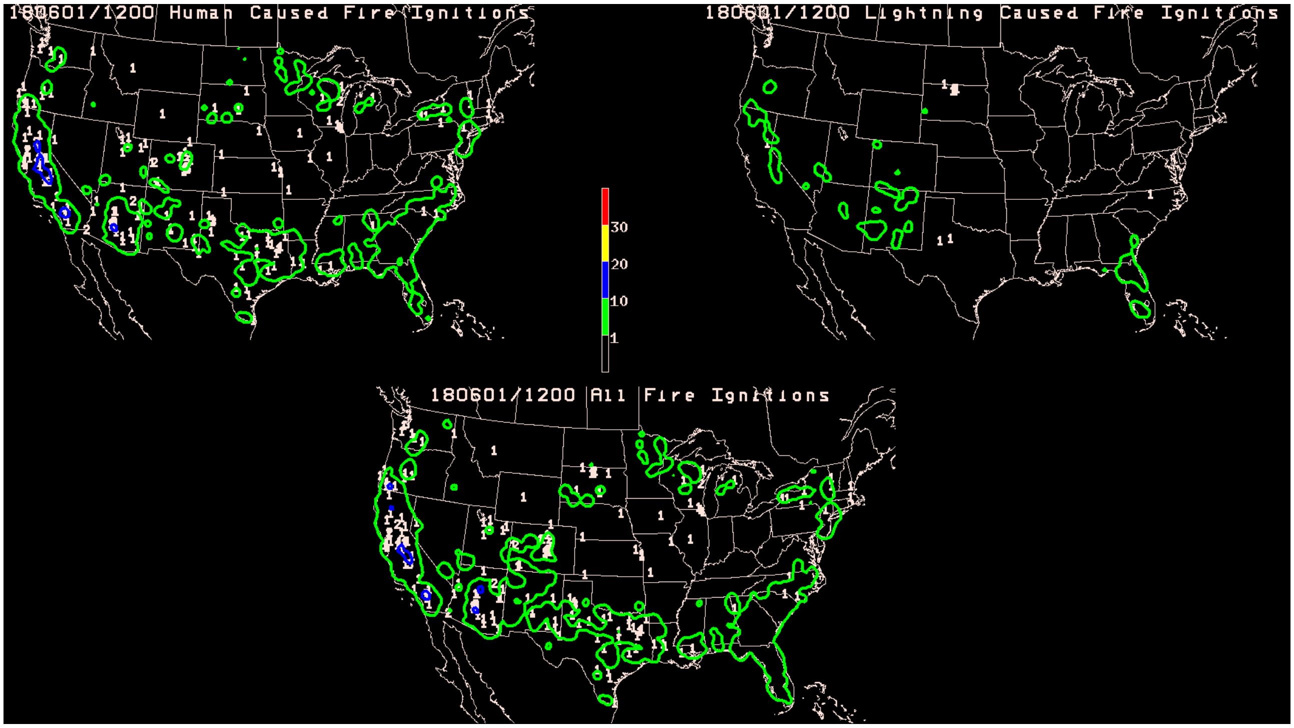


Figure 11. (*c*) The same as Fig. 11a but for June 01, 2018.

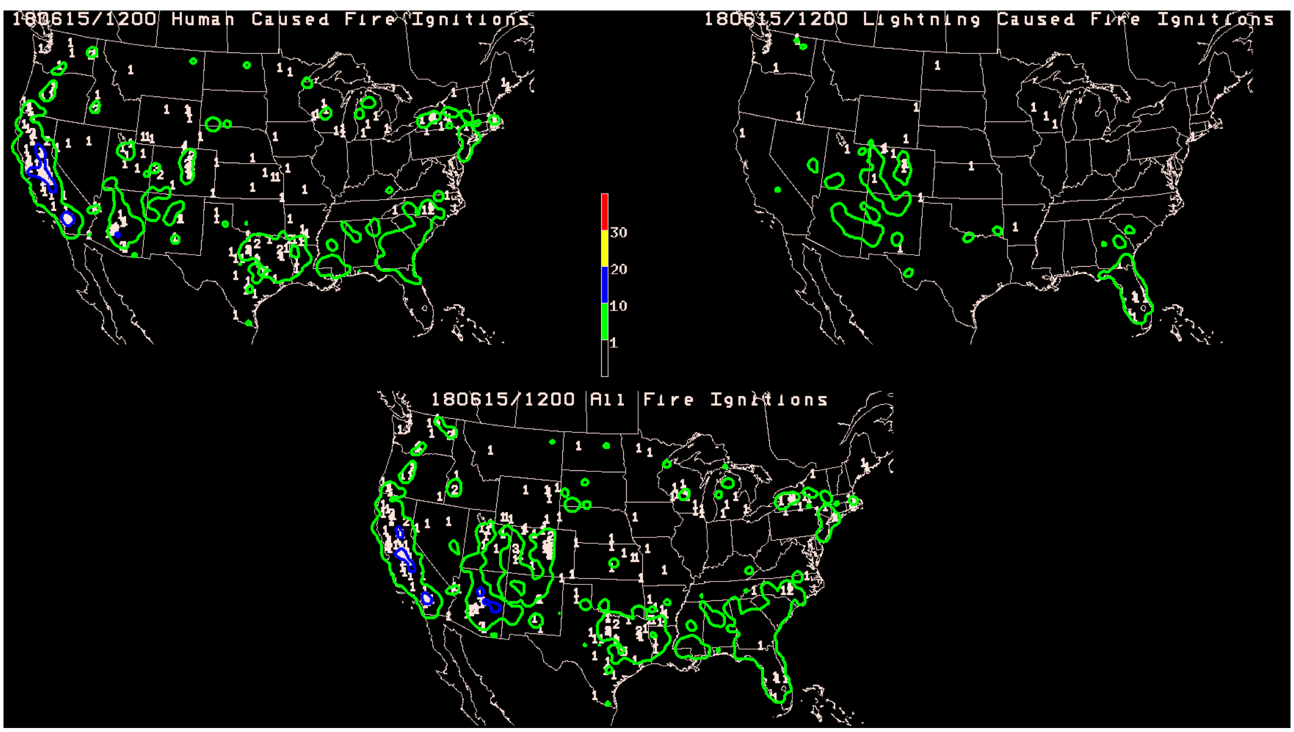


Figure 11. (*d*) The same as Fig. 11a but for June 15, 2018

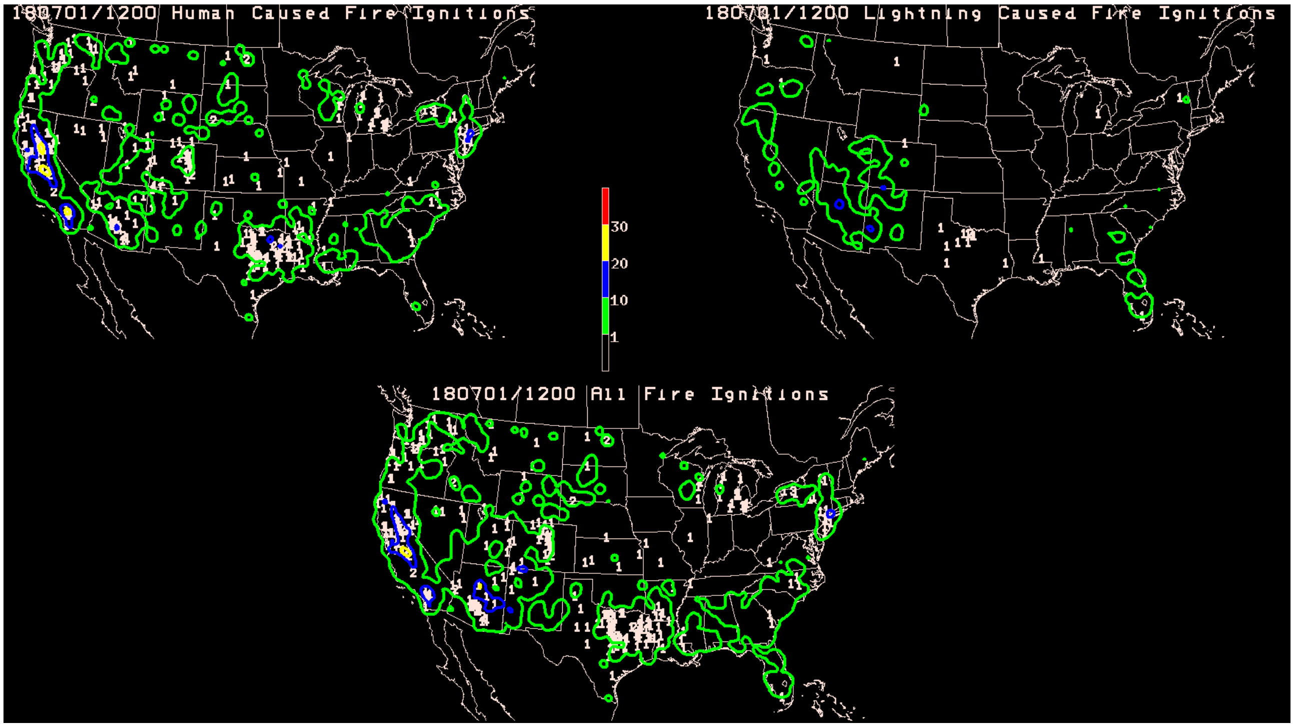


Figure 11. (*e*) The same as Fig. 11a but for July 01, 2018.

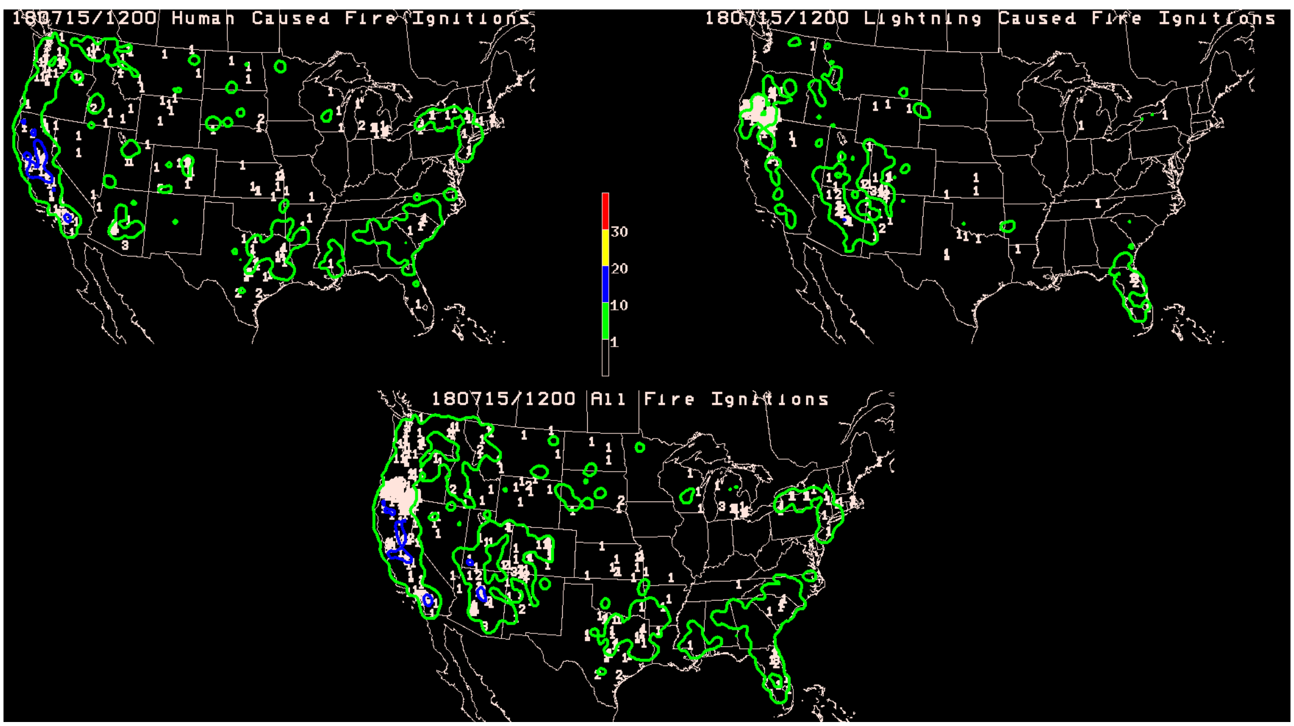


Figure 11. (*f*) The same as Fig. 11a but for July 15, 2018.

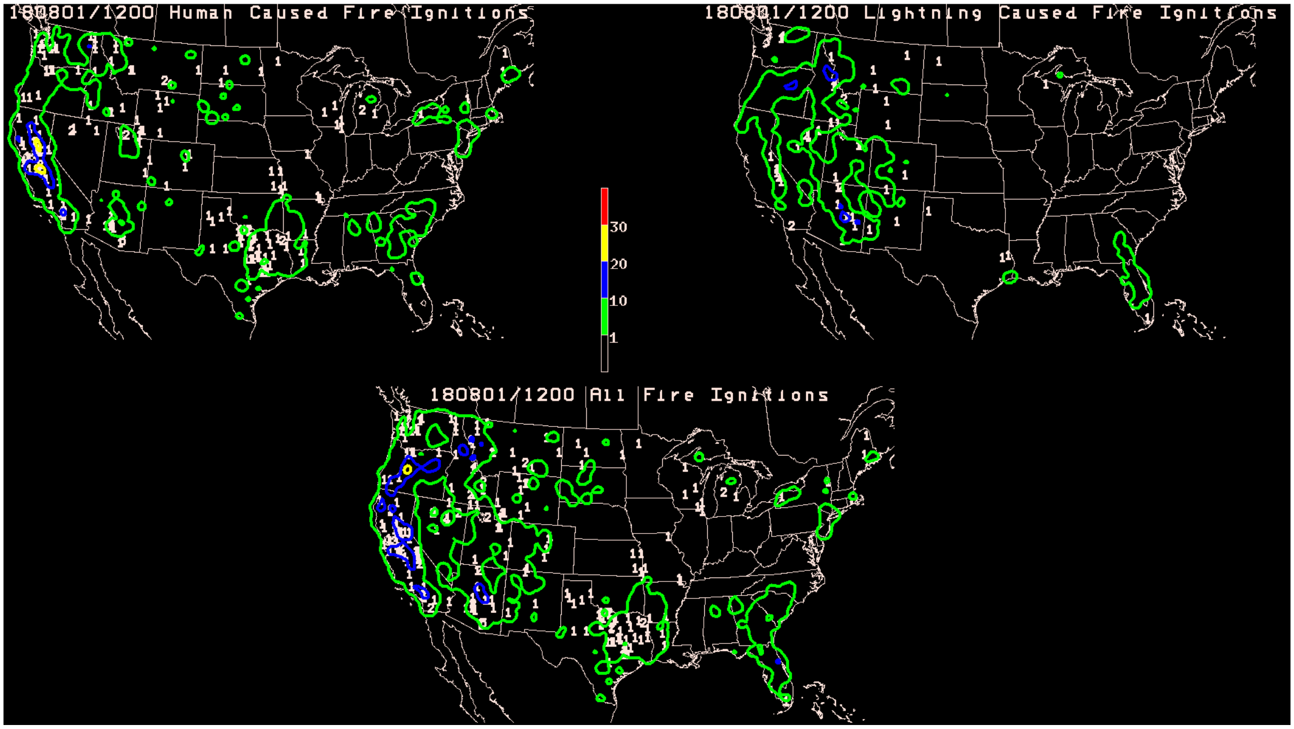


Figure 11. (*g*) The same as Fig. 11a but for August 01, 2018.

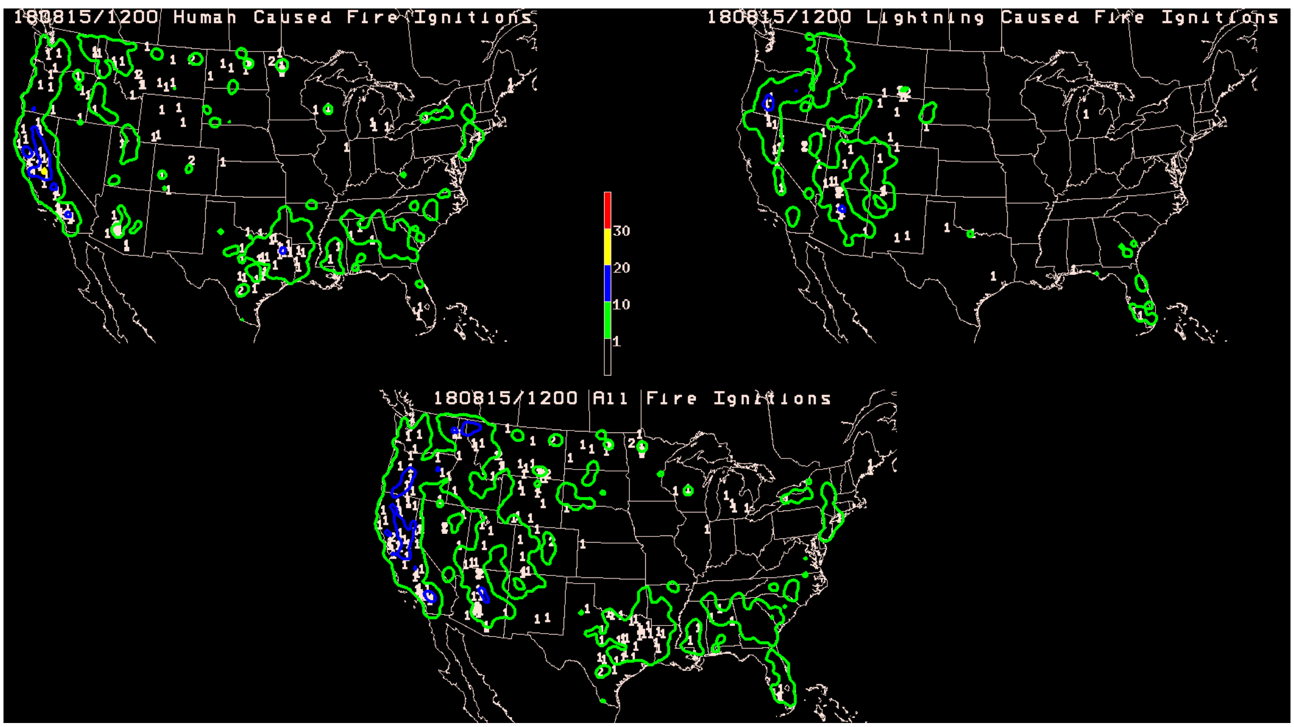


Figure 11. (*h*) The same as Fig. 11a but for August 15, 2018.

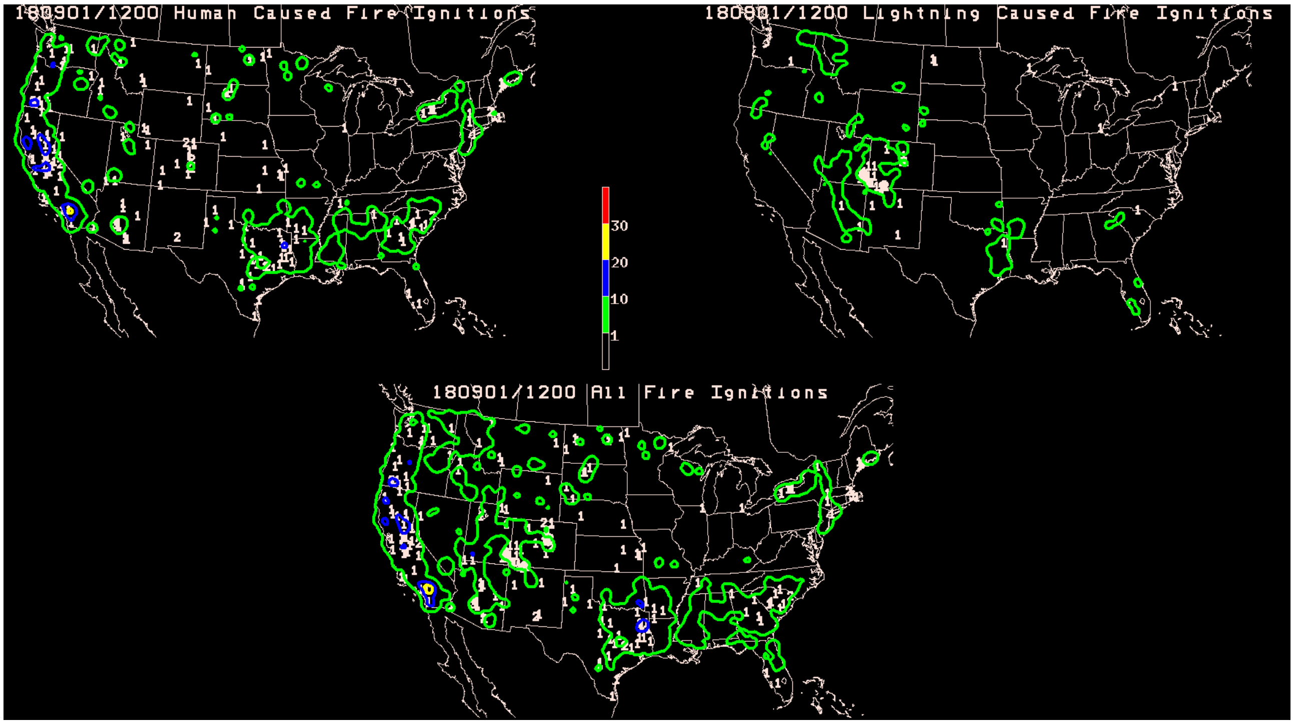


Figure 11. (*i*) The same as Fig. 11a but for September 01, 2018.

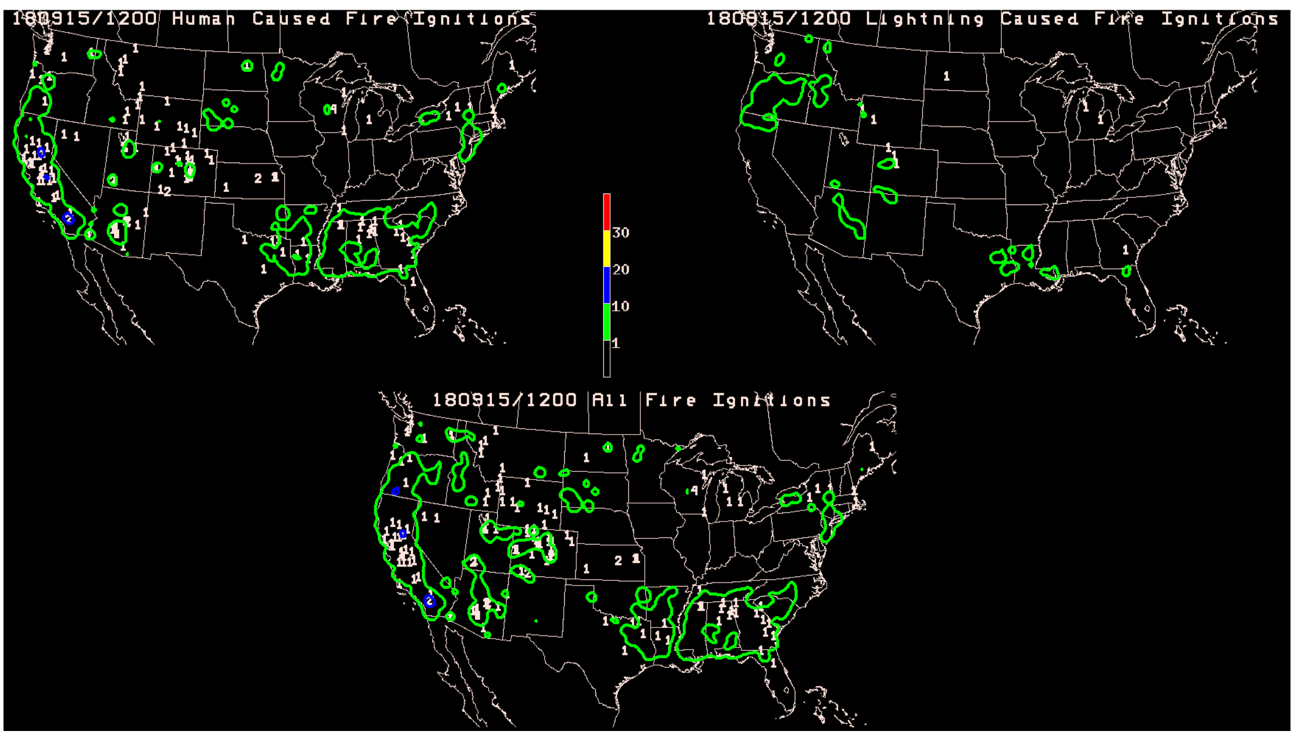


Figure 11. (j) The same as Fig. 11a but for September 15, 2018.

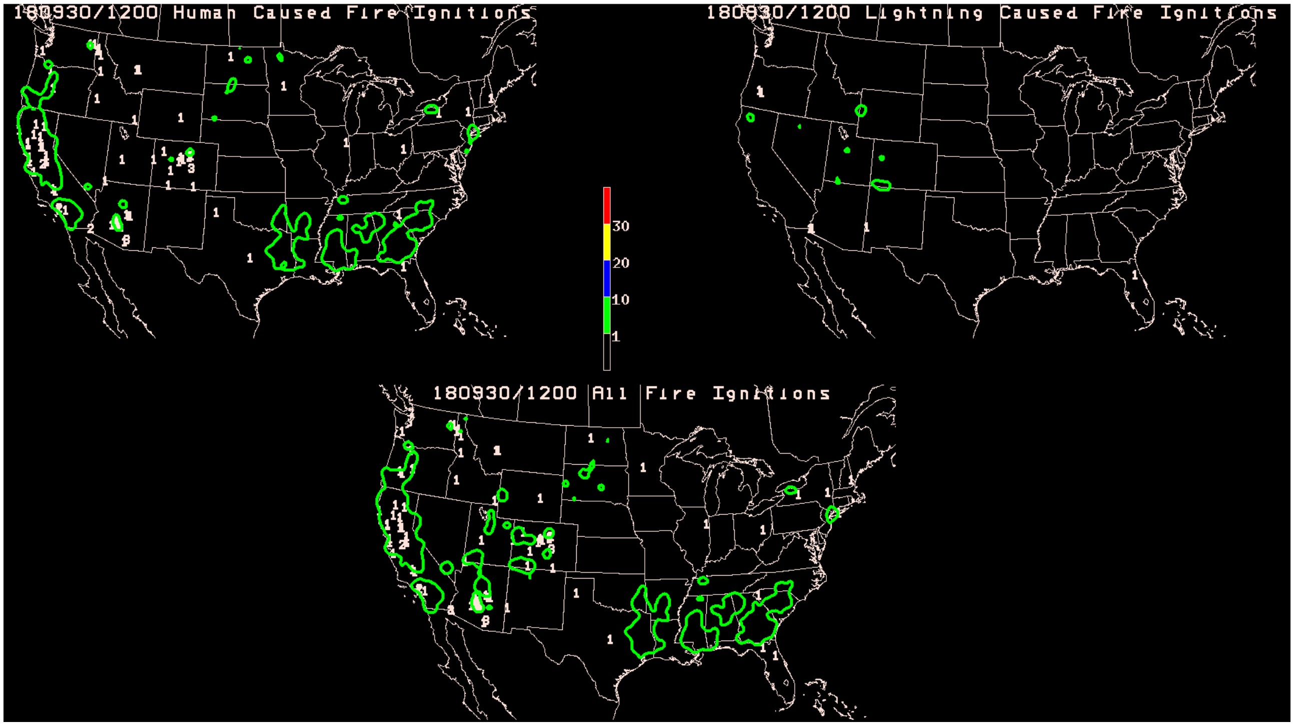


Figure 11. (*k*) The same as Fig. 11a but for September 30, 2018.

3.3. Verification of Wildfire-Ignition Forecasts Driven by 2017 and 2018 Independent NARR Data

Since wildfire-ignitions are inherently rare events, the forecast ignition probabilities are much lower than probabilities of cloud-to-ground lightning. Consequently, most ignition forecasts driven by NARR data yielded probabilities of 10% or less, which is within the magnitude of the climatological (historical) record. Numerous statistical metrics have been developed to evaluate spatially explicit forecasts. Among the most popular ones is the *Receiver Operating Characteristics* (ROC) curve, which quantifies the relationship between the Hit Rate or Probability of Detection (POD) and the False Alarm Rate or Probability of False Detection (POFD). The larger the Area Under the Curve (AUC), the higher the skill of the forecast model is. Typically, a spatially explicit model is viewed as having an *acceptable* (*good*) skill if $0.7 \leq \text{AUC} < 0.8$, an *excellent* skill if $0.8 \leq \text{AUC} \leq 0.9$, and an *outstanding* skill if $\text{AUC} > 0.9$.

Figures 12 and 13 below show the ROC curves for forecast probabilities of ignition due to human- and lightning-caused wildfires, respectively during the 2017 and 2018 fire seasons. Lightning-ignited fire forecasts assume lightning as the *only* cause. Forecasting human-caused wildfire ignitions is problematic and less certain due to unpredictable human behavior. The AUCs indicate a good (acceptable) skill with lightning-ignited forecasts showing the highest AUC of 0.7702292.

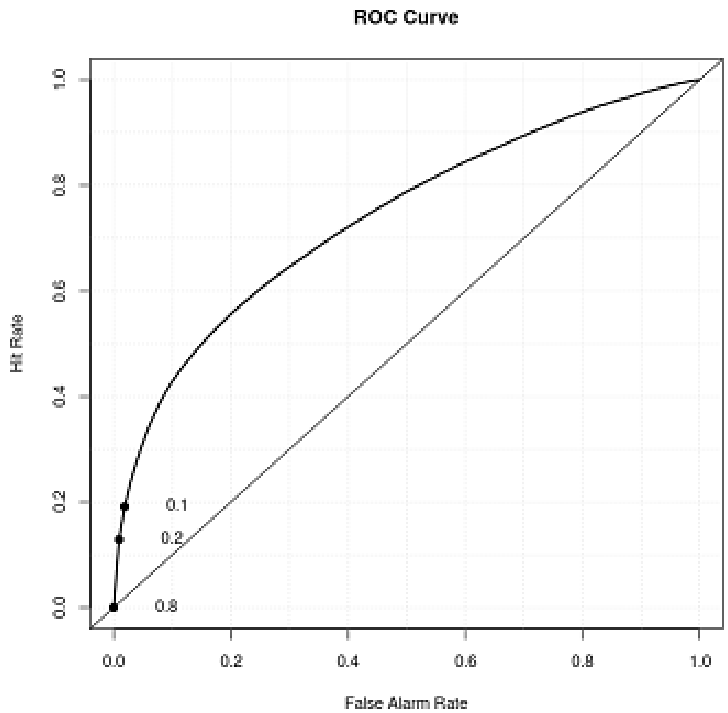


Figure 12. ROC curve for forecast wildfire-ignition probabilities due to humans in 2017 and 2018. The Area Under the Curve (AUC) is 0.7379914.

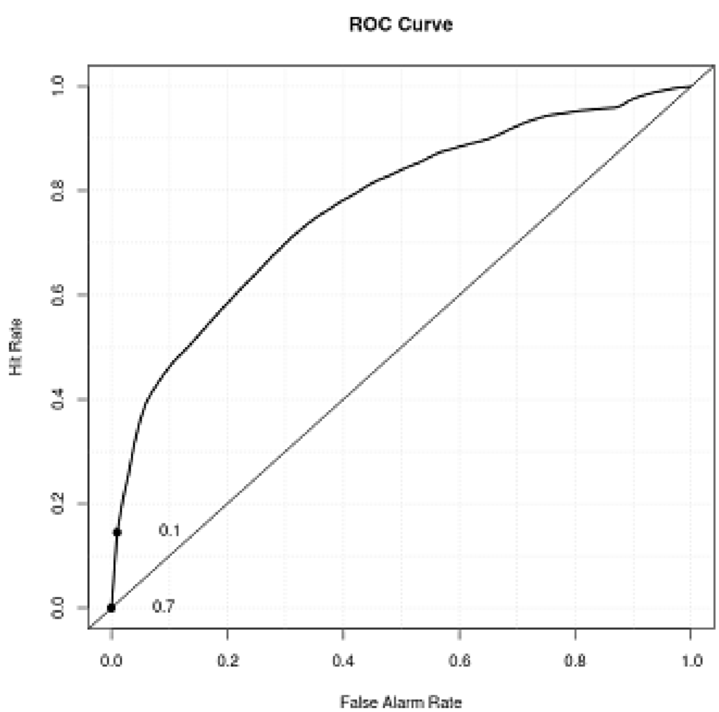


Figure 13. ROC curve for forecast wildfire-ignition probabilities due to lightning in 2017 and 2018. AUC = 0.7702292.

Figure 14 illustrates the frequency distribution of forecast ignition probabilities due to any cause (upper panels). One can see that there are hardly any grid cells with a forecast probability greater than 10%. The ROC curve (lower left panel) is within the *acceptable-skill* score ($0.7 \leq AUC < 0.8$). Since very few grid cells had forecast probabilities above 10%, the Reliability Diagram in lower right panel of Fig. 14 only shows the 0% - 10% (0.0 – 0.1) probability range. It indicates a slight over-forecasting bias with increasing probabilities beyond 5% (0.05 on the graph).

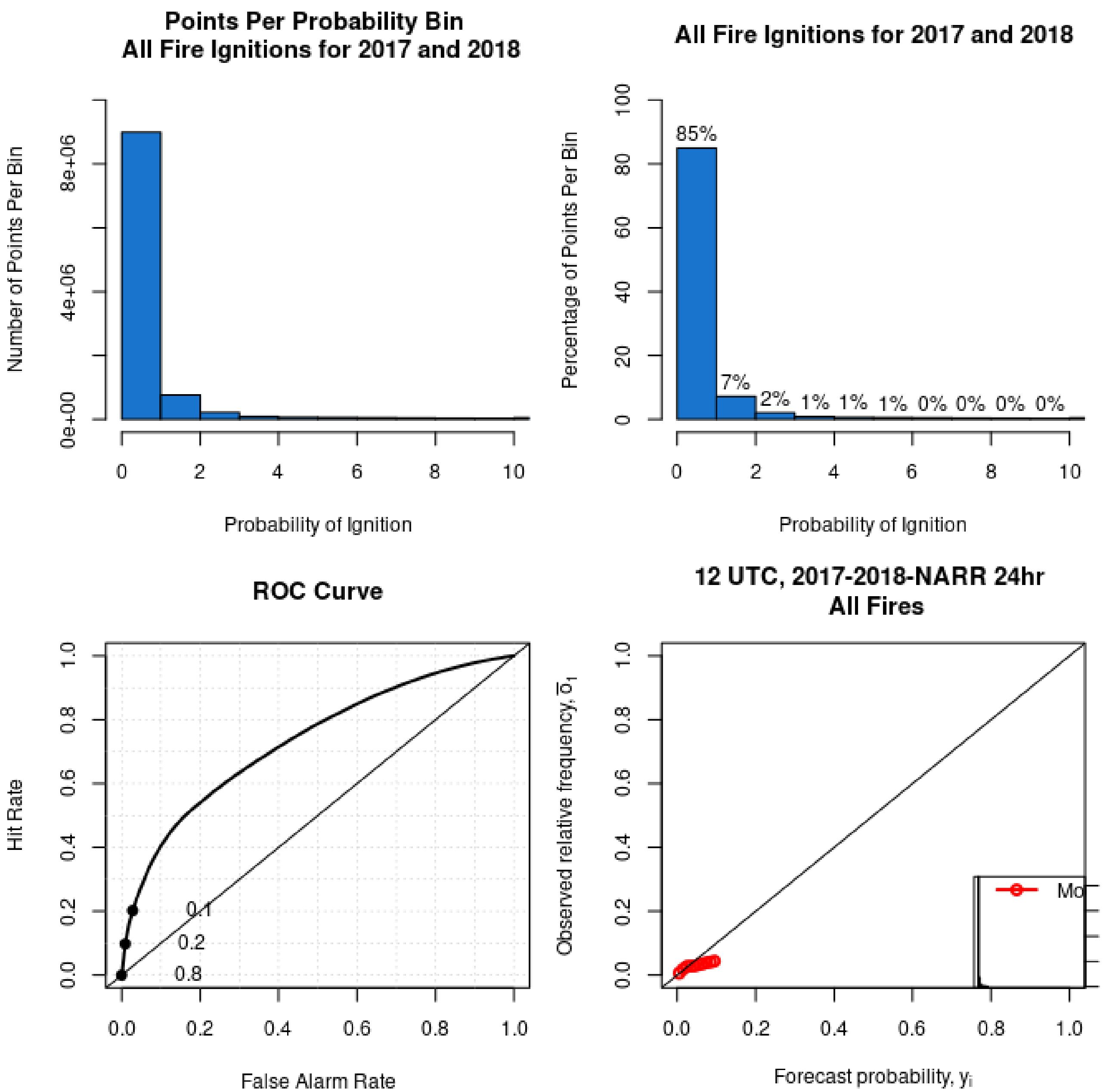
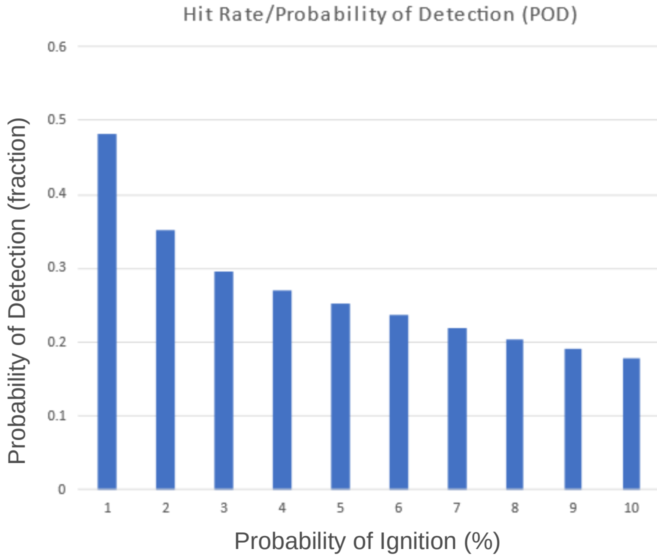


Figure 14. Forecasts of *all* fire ignitions in 2017-2018 using NARR data as a driver. *Upper left:* Occurrence histogram of forecast probabilities across ConUS; *Upper right:* Percentage of grid points in each probability interval (bin); *Lower left*: ROC diagram of all wildfire ignitions (AUC = 0.7329209). *Lower right*: Reliability Diagram for forecast probabilities in the range 1% - 10% for all wildfire ignition.



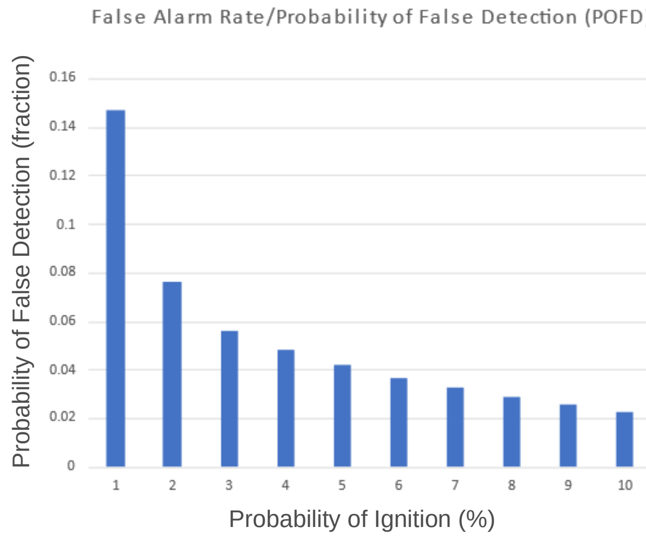


Figure 15. Histograms of the data from the X and Y axis of the ROC diagram in Fig. 14 (1% - 10% ignition probabilities). *Left panel*: Hit Rate (Probability of Detection, POD); *Right panel*: False Alarm Rate (Probability of false detection, POFD).

Figure 15 depicts occurrence histograms for the parameters forming the axes of the ROC diagram in Fig. 14 - the Hit Rate/Probability of Detection (POD) and the False Alarm Rate/Probability of False Detection (POFD). Note that the POD occurrence frequencies shown on the left panel of Fig. 15 are much higher than the POFD occurrence frequencies shown on the right panel. This indicates that the ignition-probability forecasts have value.

Since wildfire ignitions are rare events compared to atmospheric phenomena such as precipitation, for example, a useful metrics to assess the skill of the ignition prediction model is the *Extremal Dependence Index* (EDI), which is a non-degenerating measure for the quality of deterministic forecasts of rare binary events (Ferro & Stephenson 2011). If POD is higher than POFD, then EDI is positive. Conversely, if POD is lower than POFD, then EDI is negative. If POD equals POFD, EDI = 0.0. Figure 16 depicts calculated EDI values for binned forecast ignition probabilities from 0% to 10% generated for 2017 and 2018. Even though the model has a slight tendency to overpredict chances of ignition, all EDI values are greater than 0.35, which indicates a good model skill.

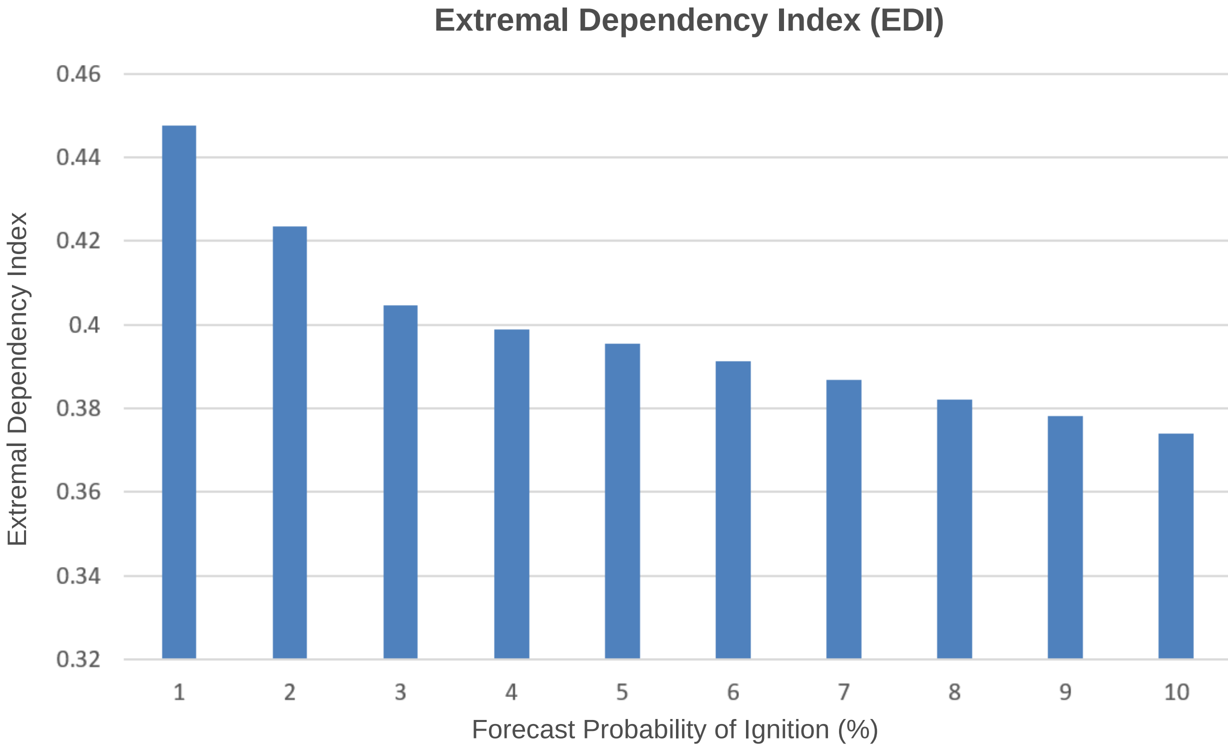


Figure 16. Extremal Dependency Index (EDI) for all-cause wildfire-ignition forecast probabilities between 1% and 10% driven by independent NARR data in 2017 and 2018.

We used two additional metrics to assess the model performance: the *Brier Score*, which is the mean squared error of the forecast; and the *Brier Skill Score* (BSS), a measure of how well the forecasts compare to climatology. The closer the *Brier Score* is to zero, the better the forecast. A value of zero for the *Brier Skill Score* means that the forecasts are indistinguishable from historical climatology. For the 2017 - 2018 NARR-driven ignition probability forecasts, the *Brier Score* is **0.008307** and the *Brier Skill Score* is **0.001603**. It should be noted that the *Brier Score* becomes inadequate for very rare events (depending on the sample size), because it does not sufficiently discriminate between small changes in weather forecasts that can significantly impact the occurrence of rare events.

We do not have the necessary data yet to evaluate GFS-driven ignition probability forecasts, which became operational on the RMC Website in August of 2021. Based on a limited amount of data from the 2021 fire season, the GFS-based probability bins indicate a slightly higher occurrence frequencies compared to NARR-based probabilities (Fig. 19).

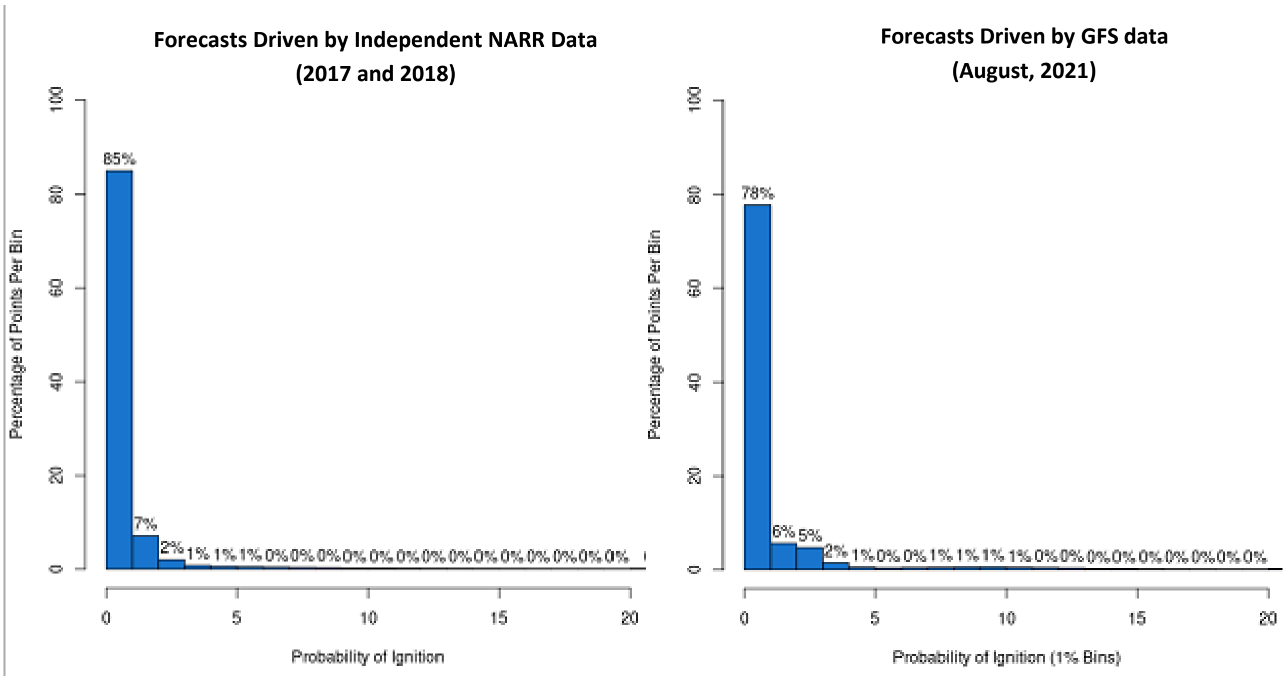


Figure 19. Frequency of occurrence of all-case wildlife-ignition forecast probabilities. *Left Panel*: NARR-driven model runs in 2017 and 2018; *Right Panel:* GFS-driven model runs in August of 2021.

4. Conclusion

The new wildfire-ignition forecast model demonstrated acceptable to very good skill as an operational tool based on two years of independent meteorological and fire-occurrence data from 2017 and 2018. The model performance will continue to be evaluated over the next few years as more independent data become available. The model currently runs twice per day using GFS forecast meteorological fields as input and 7-day predictions are posted in a form of animated (looping) maps on the RMC Website (see link "7-Day Wildfire-Start Forecast") along with links to the actual forecast data files available for download by users in a text (CSV) format.

This is the first and only National model developed by the USDA that provides *probabilistic forecasts of wildfire ignitions* on a uniform grid across the lower 48 states. As such, the model meets the legal requirements set forth in SEC. 1114 (i) of the 2019 Conservation, Management, and Recreation Act (a.k.a. the Dingell Act) to develop a system for "*Predicting Where Wildfires Will Start*". The RMC team is currently working on extending this model to cover the state of Alaska as well.

5. Acknowledgements

Funding for this project was provided by the National Fire Plan through the USDA Forest Service and the DOI Bureau of Land Management, and by the USFS Fire & Aviation Management through Peter Lahm, a USFS Manager of the Wildland Fire Air Quality Response Program. We express our gratitude to the former National Program Manager of Predictive Services, Edward (Ed) Delgado (BLM retired) for his vision about launching this project and arranging initial funding for it.

6. Appendix

Table A1. Complete list of 195 variables used in the Principal Component Analysis to derive a subset of key predictors for the wildfire-ignition probability model. Shown are parameter abbreviations, descriptions and vertical levels.

Variable Descriptions

NAME	Description/Level (PDLY is Pressure Difference Layer)
AFIRE	All fires for 24 hour period (12-12 UTC) per grid cell-SFC
AFREQ	Frequency per grid cell per 24 hours-SFC
AHUN	Probability of one hundred or more CG per grid cell (24hrs)-SFC
AONE	Probability of one or more CG per grid cell (24hrs)-SFC
AREA	Area of each grid cell (24hrs)-SFC
ATEN	Probability of ten or more CG per grid cell (24hrs)-SFC
ATHR	Probability of three or more CG per grid cell (24hrs)-SFC
ATHRT	Probability of thirty or more CG per grid cell (24hrs)-SFC
ATPCT	Percent of 24 hrs CG in grid cell-SFC
AVFL	Average number of flashes per 24 hours-SFC
AVPFL	Average number of positive flashes per 24 hours-SFC
TOTFD1	Total CG strikes for current day (12-12 UTC)-SFC
LTOTF	Total lightning ignitions (12-12 UTC) per grid cell-SFC
HTOTF	Total human caused ignitions (12-12 UTC) per grid cell-SFC
ATOTF	Total of all ignitions (12-12 UTC) per grid cell-SFC
LTOTFB	Number of years with one or more lightning ignitions per grid cell-SFC
HTOTFB	Number of years with one or more human caused ignitions per cell-SFC
ATOTFB	Number of years with one or more all ignitions per grid cell-SFC
LAVGF	Yearly Average Lightning ignitions (Over 25 years - 92-16)-SFC
HAVGF	Yearly average Human caused ignitions (over 25 years)-SFC
AAVGF	Yearly average All fire ignitions (Over 25 years)-SFC
LAVGFB	Average number of years with **one** or more lightning fire ignitions-SFC
HAVGFB	Average number of years with **one** or more Human fire ignitions-SFC
AAVGFB	Average number of years with **one** or more All fire ignitions-SFC
SIZE	average size of all fires in 24 hour period-SFC
CLAVGD1	Average Low Cloud Layer
CNAVGD1	Average Plant Canopy Surface Water [kg/m^2]
CNMIND1	Most Negative Convective Inhibition (CIN)-180:0 PDLY
CNMIND1	Most Negative Convective Inhibition (CIN)-SFC
CPMAXD1	Max Convective Available Potential Energy (CAPE)-180:PDLY
CPMAXD1	Max Convective Available Potential Energy (CAPE)-SFC
DAVGD1	24 hour average dewpoint (Deg C)-HGHT 2M
DAVGD1	24 hour average dewpoint-180:180 PDLY
DAVGD1	24 hour average dewpoint-150:120 PDLY
DAVGD1	24 hour average dewpoint-120:90 PDLY
DAVGD1	24 hour average dewpoint-90:60 PDLY
DAVGD1	24 hour average dewpoint-60:30 PDLY
DAVGD1	24 hour average dewpoint-30:0 PDLY
DMIND1	24 hour minimum dewpoint-HGHT 2M
DMIND1	24 hour minimum dewpoint-HGHT 10M
DMIND1	24 hour minimum dewpoint-180:150 PDLY
DMIND1	24 hour minimum dewpoint-150:120 PDLY
DMIND1	24 hour minimum dewpoint-120:90 PDLY
DMIND1	24 hour minimum dewpoint-90:60 PDLY
DMIND1	24 hour minimum dewpoint-60:30 PDLY
DMIND1	24 hour minimum dewpoint-30:0 PDLY
DSMAXD1	Downward short-wave radiation (W/m**2)-SFC
FAVGD1	24 hour average Fosberg FWI Current Day-HGHT 10M
FMAXM0	24 hour maximum Fosberg FWI Current Day-HGHT 10M
FMAXM1	24 hour maximum Fosberg FWI Day minus 1-HGHT 10M
FMAXM2	24 hour maximum Fosberg FWI Day minus 2-HGHT 10M
FMAXM3	24 hour maximum Fosberg FWI Day minus 3-HGHT 10M
FMAXM4	24 hour maximum Fosberg FWI Day minus 4-HGHT 10M
FMAXM5	24 hour maximum Fosberg FWI Day minus 5-HGHT 10M
FMAXM6	24 hour maximum Fosberg FWI Day minus 6-HGHT 10M
FMAXM7	24 hour maximum Fosberg FWI Day minus 7-HGHT 10M
LFIRE	Lightning caused fires for 24 hour (12-12 UTC) per grid cell-SFC
HFIRE	Human caused fires for 24 hour (12-12 UTC) per grid cell-SFC
HAVGD1	24 hour average Hot-Dry/Windy Index Current Day-HGHT 10M
HMAXM0	24 hour maximum Hot-Dry/Windy Index Current Day-HGHT 10M
HMAXM1	24 hour maximum Hot-Dry/Windy Index Day minus 1-HGHT 10M
HMAXM2	24 hour maximum Hot-Dry/Windy Index Day minus 2-HGHT 10M
HMAXM3	24 hour maximum Hot-Dry/Windy Index Day minus 3-HGHT 10M
HMAXM4	24 hour maximum Hot-Dry/Windy Index Day minus 4-HGHT 10M
HMAXM5	24 hour maximum Hot-Dry/Windy Index Day minus 5-HGHT 10M
HMAXM6	24 hour maximum Hot-Dry/Windy Index Day minus 6-HGHT 10M

HMAXM7	24 hour maximum Hot-Dry/Windy Index Day minus 7-HGHT 10M
IGPT	I grid points (only for grid calculations)-SFC
JGPT	J grid points (only for grid calculations)-SFC
LAIAVE	LAI average (Leaf Area Index)-SFC
NLAIAVE	
Integer LAI	LAI average-SFC
LAND	Mask for land vs water-SFC
LATD	Latitude (degrees) for grid points-SFC
LEMAXD1	Max Haines Index (Low)-T950-T850 PRES
MEMAXD1	Max Haines Index (Middle)-T700-T500 PRES
HEMAXD1	Max Haines Index (High)-T700-T500
HPMAXD1	Height of PBL (meters)-SFC
LIMAX1	Lifted Index max at 21 UTC (C)-SFC
LIMAXD1	Lifted Index max at 21 UTC (C)-BEST 4 LAYER
LOND	Longitude (degrees) for grid points-SFC
MASK	Mask for lower 48 states-SFC
MAVGD1	24 hour average mixing ratio-HGHT 2M
MAVGD1	24 hour average mixing ratio-HGHT 10M
MAVGD1	24 hour average mixing ratio-180:150 PDLY
MAVGD1	24 hour average mixing ratio-150:120 PDLY
MAVGD1	24 hour average mixing ratio-120:90 PDLY
MAVGD1	24 hour average mixing ratio-90:60 PDLY
MAVGD1	24 hour average mixing ratio-60:30 PDLY
MAVGD1	24 hour average mixing ratio-30:00 PDLY
MMIND1	24 hour minimum mixing ratio-HGHT 2M
MMIND1	24 hour minimum mixing ratio-HGHT 10M
MMIND1	24 hour minimum mixing ratio-180:150 PDLY
MMIND1	24 hour minimum mixing ratio-150:120 PDLY
MMIND1	24 hour minimum mixing ratio-120:90 PDLY
MMIND1	24 hour minimum mixing ratio-90:60 PDLY
MMIND1	24 hour minimum mixing ratio-60:30 PDLY
MMIND1	24 hour minimum mixing ratio-30:00 PDLY
WK01	EDDI for 1 week-SFC
WK02	EDDI for 2 weeks-SFC
MN01	EDDI for 1 month-SFC
MN03	EDDI for 3 months-SFC
MN12	EDDI for 12 months-SFC
P24ID1	24 hour precipitation (inches)-SFC
PEMAXD1	Potential evaporation-SFC
PMAVGD1	Surface pressure (PMSL)-SFC
PRMAXD1	Max precip rate-SFC
PROBM0	24 ltg prob for Day 0-"current day"-SFC
PROBM1	24 ltg prob for Day minus 1-SFC
PROBM2	24 ltg prob for Day minus 2-SFC
PROBM3	24 ltg prob for Day minus 3-SFC
PROBM4	24 ltg prob for Day minus 4-SFC
PROBM5	24 ltg prob for Day minus 5-SFC
PROBM6	24 ltg prob for Day minus 6-SFC
PROBM7	24 ltg prob for Day minus 7-SFC
PWMIND1	24 hour minimum PW -SFC
RAVGD1	24 hour average humidity-HGHT 2M
RAVGD1	24 hour average humidity-HGHT 10M
RAVGD1	24 hour average humidity-180:150 PDLY
RAVGD1	24 hour average humidity-150:120 PDLY
RAVGD1	24 hour average humidity-120:90 PDLY
RAVGD1	24 hour average humidity-90:60 PDLY
RAVGD1	24 hour average humidity-60:30 PDLY
RAVGD1	24 hour average humidity-30:00 PDLY
RMIND1	24 hour minimum humidity-HGHT 2M
RMIND1	24 hour minimum humidity-HGHT 10M
RMIND1	24 hour minimum humidity-180:150 PDLY
RMIND1	24 hour minimum humidity-150:120 PDLY
RMIND1	24 hour minimum humidity-120:90 PDLY
RMIND1	24 hour minimum humidity-90:60 PDLY
RMIND1	24 hour minimum humidity-60:30 PDLY
RMIND1	24 hour minimum humidity-30:00 PDLY
SNMAXD1	Snow maximum depth at 21 UTC-SFC
SOMINM0	Liquid volumetric soil moisture Current Day-0:10 DPTH
SOMINM1	Liquid volumetric soil moisture Day minus 1-0:10 DPTH
SOMINM2	Liquid volumetric soil moisture Day minus 2-0:10 DPTH
SOMINM3	Liquid volumetric soil moisture Day minus 3-0:10 DPTH
SOMINM4	Liquid volumetric soil moisture Day minus 4-0:10 DPTH
SOMINM5	Liquid volumetric soil moisture Day minus 5-0:10 DPTH
SOMINM6	Liquid volumetric soil moisture Day minus 6-0:10 DPTH
SOMINM7	Liquid volumetric soil moisture Day minus 7-0:10 DPTH
SOMIND1	Liquid volumetric soil moisture Current Day-10:40 DPTH
SOMIND1	Liquid volumetric soil moisture Current Day-40:100 DPTH
SOMIND1	Liquid volumetric soil moisture Current Day-100:200 DPTH
STMAXM0	
Soil Temp	Soil Temp Current Day-0:10 DPTH
STMAXM1	Soil Temp Day minus 1-0:10 DPTH
STMAXM2	Soil Temp Day minus 2-0:10 DPTH
STMAXM3	Soil Temp Day minus 3-0:10 DPTH
STMAXM4	Soil Temp Day minus 4-0:10 DPTH
STMAXM5	Soil Temp Day minus 5-0:10 DPTH
STMAXM6	Soil Temp Day minus 6-0:10 DPTH
STMAXM7	Soil Temp Day minus 7-0:10 DPTH
STMAXD1	Soil Temp Current Day-10:40 DPTH
STMAXD1	Soil Temp Current Day-40:100 DPTH
STMAXD1	Soil Temp Current Day-100:200 DPTH
TAVGD1	24 hour average temperature (Deg F)-HGHT 2M
TAVGD1	24 hour average temperature (Deg C)-HGHT 10M
TAVGD1	24 hour average temperature-HGHT 30M
TAVGD1	24 hour average temperature-180:150 PDLY
TAVGD1	24 hour average temperature-150:120 PDLY
TAVGD1	24 hour average temperature-120:90 PDLY
TAVGD1	24 hour average temperature-90:60 PDLY
TAVGD1	24 hour average temperature-60:30 PDLY
TAVGD1	24 hour average temperature-30:00 PDLY
TCAVGD1	Total cloud cover (%)-SFC
TMAXD1	24 hour maximum temperature (Deg F)-HGHT 2M
TMAXD1	24 hour maximum temperature (Deg C)-HGHT 10M
TMAXD1	24 hour maximum temperature-HGHT 30M
TMAXD1	24 hour maximum temperature-180:150 PDLY
TMAXD1	24 hour maximum temperature-150:120 PDLY
TMAXD1	24 hour maximum temperature-120:90 PDLY
TMAXD1	24 hour maximum temperature-90:60 PDLY
TMAXD1	24 hour maximum temperature-60:30 PDLY
TMAXD1	24 hour maximum temperature-30:00 PDLY
UAVGD1	24 hour average U velocity (W-E) (MPS)-HGHT 10M
UAVGD1	24 hour average U velocity (W-E)-HGHT 30M
UAVGD1	24 hour average U velocity (W-E)-180:150 PDLY
UAVGD1	24 hour average U velocity (W-E)-150:120 PDLY
UAVGD1	24 hour average U velocity (W-E)-120:90 PDLY
UAVGD1	24 hour average U velocity (W-E)-90:60 PDLY
UAVGD1	24 hour average U velocity (W-E)-60:30 PDLY
UAVGD1	24 hour average U velocity (W-E)-30:00 PDLY
UMAXD1	24 hour maximum U velocity (W-E)-HGHT 10M
UMAXD1	24 hour maximum U velocity (W-E)-HGHT 30M
UMAXD1	24 hour maximum U velocity (W-E)-180:150 PDLY
UMAXD1	24 hour maximum U velocity (W-E)-150:120 PDLY
UMAXD1	24 hour maximum U velocity (W-E)-120:90 PDLY
UMAXD1	24 hour maximum U velocity (W-E)-90:60 PDLY
UMAXD1	24 hour maximum U velocity (W-E)-60:30 PDLY
UMAXD1	24 hour maximum U velocity (W-E)-30:00 PDLY
VAVGD1	24 hour average V velocity (S-N)-HGHT 10M
VAVGD1	24 hour average V velocity (S-N)-HGHT 30M
VAVGD1	24 hour average V velocity (S-N)-180:150 PDLY
VAVGD1	24 hour average V velocity (S-N)-150-120 PDLY
VAVGD1	24 hour average V velocity (S-N)-120:90 PDLY
VAVGD1	24 hour average V velocity (S-N)-90:60 PDLY
VAVGD1	24 hour average V velocity (S-N)-60:30 PDLY
VAVGD1	24 hour average V velocity (S-N)-30:00 PDLY
VMAXD1	24 hour maximum V velocity (S-N)-HGHT 10M
VMAXD1	24 hour maximum V velocity (S-N)-HGHT 30M
VMAXD1	24 hour maximum V velocity (S-N)-180:150 PDLY
VMAXD1	24 hour maximum V velocity (S-N)-150:120 PDLY
VMAXD1	24 hour maximum V velocity (S-N)-120:90 PDLY
VMAXD1	24 hour maximum V velocity (S-N)-90:60 PDLY
VMAXD1	24 hour maximum V velocity (S-N)-60:30 PDLY
VMAXD1	24 hour maximum V velocity (S-N)-30:00 PDLY
VEG	Vegetation percent (static)-SFC
VGTYP	Vegetation type (static)-13 types-SFC
USVGTYP	Vegetation type (static)-13 types-SFC
VFSL	HI Res Vegetation type (static)-17 types-SFC

Table A2. Example of top Principal Components (PCs) for all fires in August covering the period 1992 - 2016. Color shading identifies individual members of each PC. The colored PCs are shown in Table A3

PC Loadings: (>= 0.4)

	[,1]	[,2]	[,3]	[,4]	[,5]		] [,10]
STMAXM0		0.7				0.6	
SOMINM0				0.8			
STMAXM1		0.8				0.5	
SOMINM1				0.9			
STMAXM2		0.8					
SOMINM2				0.9			
STMAXM3		0.8					
SOMINM3				0.9			
STMAXM4		0.8					
SOMINM4				0.9			
STMAXM5		0.7					
SOMINM5				0.9			
STMAXM6		0.7					
SOMINM6				0.9			
STMAXM7		0.7					
SOMINM7				0.9			
STMAXD1		0.7				0.5	
SOMIND1				0.9			
STMAXD1.1		0.7				0.5	
SOMIND1.1				0.9			
STMAXD1.2		0.6				0.5	
SOMIND1.2				0.6			
TMAXD1						0.9	
TAVGD1						0.9	
UMAXD1			-0.9				
UAVGD1			-0.9				
VMAXD1					0.9		
VAVGD1					0.9		
HMAXM0		0.5					
HMAXM1		0.6					
HMAXM2		0.6					
HMAXM3		0.6					
HMAXM4		0.6					
FMAXM1		0.4					
FMAXM2		0.5					
FMAXM3		0.5					
FMAXM4		0.5					
RMIND1	-0.8	-0.5					
RAVGD1	-0.9	-0.4					
DMIND1	-0.9					0.4	
DAVGD1	-0.9					0.4	
MMIND1	-0.8					0.4	
MAVGD1	-0.8					0.4	
CPMAXD1	-0.6					0.5	
LIMAXD1	0.5						
LEMAXD1	0.7						
MEMAXD1	0.8						
HEMAXD1	0.8						
HPMAXD1	0.6	0.5					
DSMAXD1	0.8						
PWMIND1	-0.8					0.4	

TCAVGD1	-0.6									
CLAVGD1	-0.7									
PEMAXD1	0.7	0.5								
P24ID1	-0.4									

NLAIAVE	-0.5	-0.5								
LAIAVE	-0.5	-0.5								
LIMAXD1.1	0.5									
CPMAXD1.1	-0.5							0.5		

VEG	-0.5	-0.4		0.5						
LTOTF										0.9
										0.6
LTOTFB										0.9
										0.5
										0.6
LAVGF										0.9
										0.5
LAVGFB										0.9

	[,1]	[,2]	[,3]	[,4]	[,5]					[,10]	
SS loadings	46.6	16.2	13.9	10.1	14.9	6.9	10.3	25.8	4.3	5.2	Cumulative
Proportion Var	0.24	0.09	0.07	0.05	0.08	0.04	0.05	0.13	0.02	0.03	Variance
Cumulative Var	0.24	0.32	0.39	0.45	0.52	0.56	0.61	0.74	0.76	0.790	<- explained

Table A3. Top Principal Components for Wildfire-Ignition Probabilities

Parameter	Explained Variance (%)
1. Lower atmospheric moisture/Haines Index	24
2. Soil Temp/LAI[1] and Fosberg Index/HDWI[2]	8
3. U wind component	7
4. Soil moisture temperatures-Vegetation	5
5. V wind component	8
6.	
7.	
8. Lower atmospheric temperature	13
9.	
10. Climatology of lightning caused wildfires	3

[1] Monthly Vegetation Leaf Area Index (LAI)

[2] Hot-Dry-Windy Index (HDWI)

Table A4. Example of logistic-equation coefficients calculated for the 10 PCs applicable to all wildfire ignitions in August.

```
Coefficients:
(Intercept)         ONE         TWO       THREE        FOUR        FIVE
  -5.500657    0.170232    0.008117    0.009060   -0.157519   -0.028827
        SIX       SEVEN       EIGHT        NINE         TEN
   1.122569   -0.005465    0.155062   -0.252842    0.787008
```