EPA Public Access

Author manuscript

J Expo Sci Environ Epidemiol. Author manuscript; available in PMC 2021 January 27.

About author manuscripts

Submit a manuscript

Published in final edited form as:

J Expo Sci Environ Epidemiol. 2021 February ; 31(1): 170–176. doi:10.1038/s41370-020-0246-y.

A deep learning approach to identify smoke plumes in satellite imagery in near real-time for health risk communication

Alexandra Larsen[†], Ivan Hanigan^{‡,¶,®}, Brian J. Reich[§], Yi Qin[†], Martin Cope[†], Geoffrey Morgan^{‡,⊥}, Ana G. Rappold[#]

[†] Department of Biostatistics and Bioinformatics, Duke University, NC, USA

[‡] The University of Sydney, University Centre for Rural Health, School of Public Health, Sydney, NSW, Australia

[¶] Centre for Air pollution, energy and health Research (CAR), Woolcock Institute of Medical Research, Sydney, NSW, Australia

[§] Department of Statistics, North Carolina State University, NC, USA

[†] Oceans and Atmosphere Research, Commonwealth Science and Industrial Research Organisation, Victoria, Australia

[⊥] Centre for Air pollution, Energy and Health Research (CAR), Woolcock Institute of Medical Research, Sydney, NSW, Australia

[#] U.S. Environmental Protection Agency, Center for Public Health and Environmental Assessment, Office of Research and Development, Research Triangle Park, NC, USA

[®] Centre for Research and Action in Public Health, University of Canberra, ACT, Australia

Abstract

Background: Wildland fire (wildfire; bushfire) pollution contributes to poor air quality, a risk factor for premature death. The frequency and intensity of wildfires are expected to increase; improved tools for estimating exposure to fire smoke are vital. New generation satellite-based sensors produce high-resolution spectral images, providing real-time information of surface features during wildfire episodes. Because of the vast size of such data, new automated methods for processing information are required.

Objective: We present a deep fully convolutional neural network (FCN) for predicting fire smoke in satellite imagery in near real time (NRT).

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Ana.Rappold@epa.gov.

Disclaimer and Conflicts of Interest

The research described in this article has been reviewed by the Center for Public Health and Environmental Assessment, U.S. Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflect the views and the policies of the Agency, nor does mention of trade names of commercial products constitute endorsement or recommendation for use. The authors declare that they have no conflict of interest.

Method: The FCN identifies fire smoke using output from operational smoke identification methods as training data, leveraging validated smoke products in a framework that can be operationalized in NRT. We demonstrate this for a fire episode in Australia; the algorithm is applicable to any geographic region.

Results: The algorithm has high classification accuracy (99.5% of pixels correctly classified on average) and precision (average intersection over union = 57.6%).

Significance: The FCN algorithm has high potential as an exposure assessment tool, capable of providing critical information to fire managers, health and environmental agencies and the general public to prevent the health risks associated with exposure to hazardous smoke from wildland fires in NRT.

Keywords

Wildfire smoke; remote sensing; health risk communication; artificial intelligence; fully convolutional neural network

Introduction

As a source of air pollution, wildland fires (wildfires; bushfires) are an emerging key issue in public health. Smoke is composed of hazardous gases and particles, including particulate matter, ozone, carbon monoxide, polycyclic aromatic compounds, and nitrogen dioxide, all of which have previously been demonstrated detrimental to health^{1–4}. Annually, air pollution from wildfires is estimated to contribute 340,000 deaths globally⁵ and 2,000–4,000 within continental U.S. alone⁶. It is therefore desirable to effectively manage smoke related health and air quality impacts.

Air quality impacts of wildfires are currently communicated using a combination of air pollution measurements at sparsely distributed monitoring networks, simulated forecasts of air quality, and satellite imagery^{7,8}. New generation satellite platforms emerged in recent years as a rich source of data with geostationary satellites that take images in 10 minute intervals at a 1–2 km spatial resolution, and polar orbiting satellites that observe features below one square meter resolution. The large imaging area and high spatial resolution of new satellite observations may present a challenge when considering scaling up of the operational capacities of existing platforms. Consequently, the new data may remain underutilized for communicating smoke impacts for public health purposes. New approaches need to be developed to increase operational capacity of existing applications⁹.

Artificial intelligence (AI) algorithms efficiently process vast quantities of data and have been successfully utilized in environmental applications. For example, Mazzoni and colleagues¹⁰ developed a support vector machine (SVM) to identify smoke plumes attached to fire hotspots. However, SVM approach requires time-consuming feature engineering or pre-selecting the most important descriptive features of the data before feeding it into the algorithm. Another class of algorithms called Fully Convolutional Networks (FCN) have also been applied in a variety remote sensing problems with less emphasis on feature selection and thus lower computational burden^{11–14}; including road segmentation¹⁵, weed

mapping for agriculture¹⁶, sea-land segmentation¹⁷, and ground classification¹⁸, but have not been used in smoke and fire detection problems.

Here, we present a working example of a deep FCN algorithm capable of predicting fire smoke presence in high-resolution satellite imagery in near real time (NRT). The FCN we develop conducts pixel-wise classification of smoke plumes in images by referring to the output of currently operational methods for smoke identification as the training data. In this way, our approach leverages current and validated smoke identification products in a format that can be operationalized in NRT utilizing large and rapid data from new generation satellites. We demonstrate this concept based on a fire in Northern Territory of Australia, but the algorithm is applicable to any geographic region. This new information could be used by fire managers, health and environmental agencies and the general public to better manage the health risks for fire smoke.

Materials and Methods

We illustrate the performance of a FCN approach to smoke detection from satellite imagery in a case study example from the Northern Territory (NT) of Australia which was impacted by a large wildfire smoke plume between September 1 and September 21, 2015.

Input Variables: Satellite Image Bands

The input variables for the FCN include satellite imagery measured by the Advanced Himawari Imager (AHI) onboard the Himawari-8 satellite and fire radiative power (FRP) of burn-hotspots. The Himawari-8 records images at a 2 km resolution in 10-minute intervals over 16 spectral bands¹⁹. We included six spectral bands from Himawari-8: red/green/blue (RGB) bands from the visible spectra, near infrared band, shortwave infrared band and top of atmosphere temperature band from the non-visible spectra. FRP of hotspots was used as a seventh channel in the input array. The hotspots are detected twice daily at a 375m resolution²⁰ by the NOAA Visible Infrared Imaging Radiometer Suite satellite.

Satellite spectral bands provide a visual image of smoke as seen in Figure 1, however, for public health messaging and for analysis of wildfire smoke exposure we need to have a pixel wise classification of smoke status. On a smaller scale, this can be done using visual inspection of smoke plumes (eg. Hazard Mapping System in North America) or by using analytical methods such as described in Qin et al. (2015) and Qin et al. (2019). Both have good performance and have their advantages. Their common challenge, is that they cannot be easily scaled up and made operational to accommodate the large data from the new generation satellites. In our example, we use an analytical cloud-masking algorithm described in Qin et al. (2015) as a ground truth or target variable.

Target Variable

In machine learning, the target variable is used to train the algorithm to make pixel-wise classification of smoke plumes from input variables. We use output from a deterministic cloud-masking algorithm^{9,21} consisting of analytically-derived aerosol types (cloud, smoke, dust, fog, absorptive, etc.) to create a target data set of images where each pixel is labelled "smoke" or "non-smoke". Qin et al.'s (2019) cloud-masking algorithm takes the

first five bands of satellite imagery (visible, near-infrared, shortwave infrared) measured by the Advanced Himawari Imager (AHI) onboard the Himawari-8 satellite to produce aerosol optical depth (AOD) fields on a $2 \times 2 \text{ km}^2$ resolution grid covering Australia. The AOD fields are derived for 10-minute increments. The algorithm takes $< 1 \text{ ms}$ to process one pixel per time point, or about 30 minutes on one CPU (2.6 GHz) to process the continental land surface. Details on the method are given in Qin et al (2019). Details on pre-processing data are given in the Supplemental Materials and data are available from the authors upon request. Figure 1 shows the study area overlaid with input data (RGB image + hotspots) and analytically derived target data.

FCN Model Architecture and Training

The FCN model ‘learns’ a pixel’s class by finding optimal values for the model parameters through minimizing the prediction error against the target data set. The FCN model consists of a series of operations, referred to as ‘layers’, where the output of each layer becomes the input for the next layer. The number of layers defines the depth of the algorithm.

The first group or ‘block’ of layers in our model is the encoding block, where each layer successively distills local features in the images that are indicative of smoke plumes. These layers include convolution, which conducts kernel smoothing, and max pooling (MP), which computes local spatial maxima. Nonlinear functions (restricted linear unit; ReLU) are also included, giving the algorithm the benefit of being able to learn non-linear relationships between the input and target data. After three applications of each encoding layer, the resulting output of the encoding block is a lower dimensional representation of the data.

The decoding block projects the coarse output from the encoding block back to the original spatial scale of the data (Figure 2). The layers in the encoding block include transpose convolution, ReLU and batch normalization, which refers to normalizing the values in batch of images to encourage stable estimation. In addition, the decoding block includes skip connections¹¹, where output from each layer in the encoding block is added to output in the corresponding and symmetric output of corresponding dimensions in the decoding block. The data is added before the other encoding operations are applied. Figure 2 illustrates the skip connections as well as the output from the other layers in the decoding block.

The last group of operations in the FCN model is the prediction block, which contains two layers. The first is a convolution that brings the dimensions of the output back to the original image. The sigmoid function then transforms predictions at each pixel to a scale of 0 to 1, giving us the probability of each pixel being smoke or not. We classified the site as smoke if the probability exceeds a threshold $t = 0.5$, a value chosen through sensitivity testing.

We used stochastic batch gradient descent to optimize the model parameters with respect to the binary cross entropy loss function, commonly used for object segmentation. The kernels in both the convolution and transpose convolution layers are composed of unknown parameters that must be learned, while the rest of the layers (ReLU, MP, batch normalization and the sigmoid function) are non-parametric. Further details on the modeling settings, parameters and layers of the FCN model are provided in the Supplemental Materials.

Model Performance

The model was run for 20 train-test cycles, or ‘epochs’, each time training on 70% of the 975 images in the data and evaluating model performance on a 30% hold-out set¹¹. Images were randomly assigned to either the train or test set. Classification accuracy (A) is the percentage of correctly classified pixels out of the total pixels in each image. Intersection over union (IoU; J), evaluates precision by calculating the overlap between the prediction and target variables. IoU can be broken down into the true positive IoU (I_{TP}) and the true negative IoU (I_{TN}), which describes the overlap between the target variable and the predicted smoke and non-smoke pixels respectively. We also report true positive and negative IoUs weighted by the average number of smoke and non-smoke pixels in a batch of images, respectively (I_{TP}^w, I_{TN}^w). We report weighted IoUs since in many cases, the images have a much larger proportion of non-smoke pixels to smoke pixels; unweighted IoUs do not account for this imbalance. Mathematical details for performance metrics are the Supplemental Materials.

A naive way to predict classification of each pixel in the satellite image would be to use logistic regression with target variable as response and six spectral bands and hotspots as predictor variables. Therefore, to further evaluate the FCN’s performance, we compare the algorithm to the following benchmark logistic regression (LR) model:

$$\text{logit}(E[Y_t]) = \beta_0 + \beta_1 B1_t + \beta_2 B2_t + \beta_3 B3_t + \beta_4 B4_t + \beta_5 B5_t + \beta_6 temp_t + \beta_7 FRP_t$$

for $t = 1, \dots, m$ where m is the number of images. Each covariate and the response are N-dimensional vectors, where N is the number of pixels on time point t . We treat all outcomes as independent. We conducted 5-fold cross-validation with the LR model and calculated A, IoU (TP, TN) to be able to compare with the FCN performance.

Results

Overall, the FCN outperforms the LR model. The average classification accuracy measured in the test set was 99.5% for the FCN, in comparison to 57.1% for the LR model. FCN classification also outperformed LR with respect to the mean IoU where LR did not achieve an acceptable 50% (IoU for FCN=57.6%, LR=31.1%). On average, 15.8% of the smoke pixels predicted by FCN overlapped with the target values, as compared to 6.4% for the LR model. The performance of the FCN algorithm varies depending on the number of smoke pixels present or equivalently the size of the plume. Most misclassification occurs along the edges of the plume so that the same absolute count of misclassified pixels in a small plume yield smaller I_{TP} . When we normalized images by the number of smoke pixels in an image the average I_{TP} increases to $I_{TP}^w = 22.8\%$ (Supplemental Materials, section on Model Performance). The FCN predicted 99.5% of the non-smoke pixels accurately (I_{TN} for LR=55.8%), however this high value is largely reflective of the large number of non-smoke pixels; typically there were many more non-smoke than smoke pixels in a given image.

Figure 3 shows the FCN model inputs, output fields and diagnostics for a single images of plume. Evidence of the plume is visible in most of the input variables in both cases.

Although the prediction misses some of the scattered smoke pixels, the plume position and extent are captured well. Non-smoke pixels are interpolated within the interior of the predicted plume. Additional time slices are available in the Supplemental Materials.

FCN also provided computational advantage over the cloud masking algorithm. Specifically, the cloud masking algorithm takes about 1ms to process one pixel per time point, or about 20 seconds to process the 161×105 pixel image on one 2.6 GHz CPU. In comparison, FCN took s on a 2.3 GHz processor.

Finally, we calculated the performance metrics over each of the 20 epochs for varying ratios of data in the train and test sets, i.e. train:test ratios (displayed in the Supplemental Materials). This allows us to gauge the amount of training needed to reach acceptable performance rates. We found that the loss decreases consistently, but at a faster rate for the higher train:test ratios. The average classification accuracy at each epoch trends upwards the fastest for the higher train:test ratios. Accuracy not only increases more slowly than the lower train:test ratio analyses, but also plateaus after the first few epochs. These metrics suggest that with a moderate amount of training data, acceptable accuracy can be reached with few epochs thus increasing performance speed.

Conclusions and Discussion

In this manuscript, we provide a proof of concept framework for automating smoke detection in near-real time using a deep learning algorithm and a large volume of spatially and temporally resolved data streamed from a geostationary satellite. The approach relies on the use of existing validated approaches to smoke detection to train the deep learning algorithm but adds capacity to process additional data in near real time. Existing smoke detection approaches carry substantial scientific background and good performance, however the computational burden related to processing high temporal data resolutions, limits the capacity for NRT operational utility. In the proposed concept, the current scientific knowledge-based algorithms are used to train a deep learning algorithm to detect smoke in near-real time based on direct inputs from the visible and near infrared channels of satellite imagery. The proposed framework is relatively simple to deploy in the presence of current methods and shows promising results in terms of being able to detect smoke pixels in satellite imagery accurately and timely.

If operational, the FCN model coupled with analytic algorithms, such as the one used to generate the target variable data for this project, provides smoke plume envelopes for public health risk communication and can provide automated alerts to fire control centers. The product could benefit rural areas with sensitive Agriculture (vineyards, livestock, etc.) that might not otherwise have local means for smoke detection. Urban areas would also benefit because of the denser populations and more people to protect from the harmful effects of smoke pollution exposure during intense smoke episodes^{22,23}. The resulting smoke plume predictions in conjunction with near real-time modelled and measured particulate data, including increasingly available data from dense low cost sensor networks, can provide an integrated picture of smoke exposure across an airshed. The model could also be used to calibrate chemical transport models and improve the ability to forecast smoke in NRL.

In Australia, one scenario for operationalizing the algorithm may be that as a Himawari-8 scan becomes available, the FCN model would do a rapid assessment of the presence of smoke plumes for the entire region of interest (i.e., Australia; CONUS if using GOES). This information (plume locations) would then be available for Regional Control Centre operations for developing contours of smoke exposure and informing chemical transport modeling framework. In the U.S., the National Oceanic and Atmospheric Administration (NOAA) has developed a nationwide smoke detection tool called, the Hazard Mapping System (HMS) Fire and Smoke product. The HMS provides the geospatial extent of smoke plumes by manually inspecting visible bands of satellite imagery. All HMS smoke plumes are published daily as shape files on the NOAA website approximately twice a day and are used to communicate smoke plume locations to researchers and the public. Currently, the HMS incorporates data from seven NOAA and National Aeronautics and Space Administration (NASA) environmental satellites, recording multiple plumes through-out each day. This system can also potentially be leveraged to generate the target variable data and to train the AI algorithm to utilize data made recently available by the newest generation geostationary satellites which like Himawari provide enormous amounts of data in 10 minute intervals. However, we do not have a good estimate on how much HMS data would be needed to train the algorithm.

Although the proposed framework has reasonable performance, the FCN model can be enhanced. First, the model assumes stationary over time. In the short time window we consider this may be reasonable, but if the method is used operationally over a long period of time, the model may need to be periodically refit to account for seasonality and changes in the environment and technology. To determine the frequency of retraining the FCN, the user must balance having a sufficiently large training set to ensure the model fits well, and sufficiently frequent retaining to allow the model to adapt to the aforementioned changes. The algorithm can be enhanced by adding temporal dependencies or relaxing assumptions of stationary over time. Adding temporal variability in the smoke plumes would enhance our predictive model but is beyond the scope of this paper and better suited for the computer science field. This limitation highlights a greater need within the field of environmental epidemiology for increased collaborations with computer scientists in order to fully take advantage of deep learning for large scale environmental data analysis.

Although the FCN algorithm has high classification accuracy in comparison to the logistic regression, other measures of model performance are affected by miss-classification on the fringes of smoke plumes. Namely, the same number of misclassified pixels will yield small I_{TP} when the smoke plumes are small and will increase as the size of the smoke plume increases. Additionally, model performance may also depend on the hyper-parameters in the model such as depth, kernel size, type of layer operations etc. Here we did not focus on optimizing model performance which is better suited for implementation stage. Instead we focused on providing a proof of concept. Finally, our model is implemented on a standard performance personal laptop. We expect that many features can be optimized and performance can be enhanced if implemented in professional grade environment. For this purpose, we provide data and code at github.com/aelarsen/code_smoke_detection_github and github.com/ivanhanigan/Smoke_plume_hotspots_ML_data.

In summary, we illustrate a proof of concept algorithm for leveraging artificial intelligence together with scientifically rigorous methods to create operational solutions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research is being supported by funding from the Joint Fire Science Program. Ivan Hannigan was supported by funding from The Centre for Air Pollution, Energy and Health Research (www.car-cre.org.au, an Australian National Health and Medical Research Council funded Centre for Research Excellence, APP1030259), and by funding from the United States Department of the Interior and the United States Fire Service through the Joint Fire Science Program (ID: 14-1-04-9).

References

- (1). Naeher LP, Brauer M, Lipsett M, Zelikoff JT, Simpson CD, Koenig JQ, et al. Woodsmoke Health Effects: A Review. *Inhalation Toxicology* 2007, 19, 67–106.
- (2). Wettstein ZS, Hoshiko S, Fahimi J, Harrison RJ, Cascio WE, Rappold AG. Cardiovascular and Cerebrovascular Emergency Department Visits Associated With Wildfire Smoke Exposure in California in 2015. *Journal of the American Heart Association* 2018, 7, e007492.
- (3). Rappold AG, Stone SL, Cascio WE, Neas LM, Kilaru VJ, Carraway MS, et al. Peat bog wildfire smoke exposure in rural North Carolina is associated with cardiopulmonary emergency department visits assessed through syndromic surveillance. *Environmental Health Perspectives* 2011, 119, 1415–1420. [PubMed: 21705297]
- (4). Reid CE, Brauer M, Johnston FH, Jerrett M, Balmes JR, Elliott CT. Critical Review of Health Impacts of Wildfire Smoke Exposure. *Environmental Health Perspectives* 2016, 124, 1334–1343. [PubMed: 27082891]
- (5). Johnston FH, Henderson SB, Chen Y, Randerson JT, Marlier M, DeFries RS, et al. Estimated global mortality attributable to smoke from landscape fires. *Environmental Health Perspectives* 2012, 120, 695–701. [PubMed: 22456494]
- (6). Fann N, Alman B, Broome RA, Morgan GG, Johnston FH, Pouliot G, et al. The health impacts and economic value of wildland fire episodes in the. *Science of the Total Environment* 2018, 610–611, 802–809.
- (7). U.S. EPA, AirNow. <https://airnow.gov/>.
- (8). EPA Victoria, EPA AirWatch — Environment Protection Authority Victoria. <https://www.epa.vic.gov.au/EPAirWatch/>.
- (9). Qin Y, Steven ADL, Schroeder T, McVicar TR, Huang J, Cope M, et al. Cloud Cover in the Australian Region: Development and Validation of a Cloud Masking, Classification and Optical Depth Retrieval Algorithm for the Advanced Himawari Imager. *Frontiers in Environmental Science* 2019, 7, 20.
- (10). Mazzoni D, Garay MJ, Davies R, Nelson D. An operational MISR pixel classifier using support vector machines. *Remote Sensing of Environment* 2007, 107, 149–158.
- (11). Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation. 2015,
- (12). Iglovikov V, Accord T, Mushinskiy S, Aerostate VO. Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition.
- (13). Marmanis D, Wegner JD, Galliani S, Schindler K, Datcu M, Stilla U. SEMANTIC SEGMENTATION OF AERIAL IMAGES WITH AN ENSEMBLE OF CNNs; 2016.
- (14). Zhang L, Zhang L, Kumar V. Deep learning for Remote Sensing Data. *IEEE Geoscience and Remote Sensing Magazine* 2016, 4, 22–40.
- (15). Henry C, Majid Azimi S, Merkle N. Road Segmentation in SAR Satellite Images with Deep Fully-Convolutional Neural Networks.

- (16). Huang H, Deng J, Lan Y, Yang A, Deng X, Zhang L. A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery; 2018.
- (17). Li R, Liu W, Yang L, Sun S, Hu W, Zhang F, et al. DeepUNet: A Deep Fully Convolutional Network for Pixel-level Sea-Land Segmentation. 2017,
- (18). Rizaldy A, Persello C, Gevaert CM, Oude Elberink SJ. FULLY CONVOLUTIONAL NETWORKS FOR GROUND CLASSIFICATION FROM LIDAR POINT CLOUDS. 2018,
- (19). MSC JMA, JMA/MS: Himawari-8/9 Imager (AHI). http://www.data.jma.go.jp/mscweb/en/himawari89/space_segment/spsg_ahi.html.
- (20). NASA, Active Fire Data — Earthdata. <https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms/active-fire-data>.
- (21). Qin Y, Mitchell R, Forgan BW. Characterizing the Aerosol and Surface Reflectance Over Australia Using AATSR. IEEE Transactions on Geoscience and Remote Sensing 2015, 53, 6163–6182.
- (22). Broome RA, Johnston FH, Horsley J, Morgan GG. A rapid assessment of the impact of hazard reduction burning around Sydney, May 2016. The Medical Journal of Australia 2016, 205, 407–408. [PubMed: 27809737]
- (23). Horsley JA, Broome RA, Johnston FH, Cope M, Morgan GG. Health burden associated with fire smoke in Sydney, 2001–2013. Medical Journal of Australia 2018, 208, 309–310.

Figure 1: Satellite Imagery and Target Data.

Raw data from the Himiware-8 satellite on 2015-09-11 0650 UTC over the Northern Territory of Australia on a 161×105 pixel grid (left). Smoke classification from a cloud-masking algorithm^{9,21} and hotspot locations from the NOAA VIIRS satellite (right).

Figure 2: Deep fully convolutional network architecture.

Outputs from each layer are represented as 3-dimensional grey boxes with depth dimension given by the numbers above each box. For full dimension (width and height) please see Supplemental Table 1. The output from each layer is an input to the next. Layer operations are represented with arrows (convolution (CONV), Max Pooling (MP), transpose convolution (TCONV), and batch normalization (BN)). The last convolution and transpose convolution are followed by restricted linear unit (not shown). Layers on the left and right halves of the network encompass the encoding and decoding blocks, respectively. Blue arrows denote skip connections.

Figure 3: FCN Input and Output.

The *ground truth* and *prediction* panels show smoke (red) and non-smoke (blue) pixels from the target data and the FCN model predictions, respectively. The *final conv.* panel is a visualization of the output from the last convolution layer in the prediction block. *Red*, *blue*, *green*, *infrared1* (near infrared) and *infrared2* (shortwave infrared) and *temperature* are the six spectral bands from Himawari-8. *Hotspots* are the average FRP detected at each pixel. For each panel, we report accuracy (A), intersection over union (IoU; I), true positive/negative IoU (I_{TP} , I_{TN}) and weighted true positive/negative IoU (I_{TP}^w , I_{TN}^w).