Southwestern U.S. and Northwestern Mexico Dust Storm Modeling Trough Moderate Resolution Imaging Spectroradiometer Data: A Machine Learning Perspective.

Pablo Rivas-Perea[a], James C. Tilton[b] and Jose Gerardo Rosiles[a]
[a]Department of Electrical and Computer Engineering,
The University of Texas El Paso, El Paso, TX 79968 USA
[b]Computational and Information Sciences and Technology Office,
NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA

ABSTRACT

Dust storms are subject to study since they are correlated to an increase in mortality rates due to respiratory illness, especially in the southwestern U.S. With the aim of providing better tools in the understanding of dust storms, we present models for detection of dust storms from MODIS Terra Level 1B radiances, which can be applied in near real time, in contrat to those models that are based on MODIS Aerosol Optical Thickness products that are produced two days after reception.

In this paper we present a collection of events used to construct a database that we used to model probabilistic and soft computing classification methods. Then we compare the probabilistic models: Maximum Likelihood, MLE, and Maximum *a posteriori*, MAP, against the soft computing models: Feedforward Backpropagation Neural Network, FFNN, and Probabilistic Neural Network, PNN. The results showed that the soft computing-based models perform better in classifying dust, and discriminating from other signatures, such as clouds, smoke, etc.

Keywords: Neural Networks, Probabilistic Modeling, Dust Storm, Image Processing, Pattern Recognition, Remote Sensing

1. INTRODUCTION

Dust storms are a major cause of several physical, environmental and economical hazards. Air pollution from dust storms is a significant health hazard for people with respiratory diseases and can adversely impact urban areas[1] as shown in Figure 1 and Figure 2. There is a direct correlation between exposure to high-levels of airborne particle concentrations (aerosols) and the increase in mortality rate from cardiovascular, respiratory illness and lung cancer. This situation is major concern for health and safety agencies as well as for the environmental[2] and geological science agencies.[3]

Therefore, timely warnings of dust storms need to be fully functional in populated regions for health concerns and traffic control.[5] However, in spite of the fact that several methods for detecting dust storms exist, there are still open questions in the detection process and in dust storm feature extraction.[8] Furthermore, dust storm are still considered as an open problem in analysis and modeling, since a single dust storm can travel large distances. That is the case of many Saharan dust storms whose aerosols are spread along the globe.

In the literature we can find the dust storm problem addressed from different perspectives. The geophysical perspective[9] demonstrated that it is possible to detect dust aerosols using satellite infrared bands.[4] There are also studies that show the transport of the dust storms,[5] which cover large parts of the earth,[6] by observing the origins and causes of dust storms.[7] Recently, The University of Texas El Paso (UTEP) has made significant

Further author information: (Send correspondence to J.G.R.)
P.R.P.: E-mail: privas@miners.utep.edu, Telephone: 1 915 504 9808
J.C.T.: E-mail: James.C.Tilton@nasa.gov, Telephone: 1 301 286 9510
J.G.R.: E-mail: grosiles@utep.edu, Telephone: 1 915 747 7956, Address: The University of Texas at El Paso, College of Engineering, 500 West University Avenue, El Paso, Texas 79968 USA

contributions in the latter topic by Rivera *et al.*,[10] Lee *et al.*,[11] and Swapna.[12] However, these findings need to be improved in order to produce an adequate insight into the dust storm analysis problem. Moreover, dust storm detection is still an open problem in rapid response systems, which require to minimize the processing time, and be able to produce results within moderate and high resolution imagery.

In remote sensing exist different approaches that are specialized in detection[14] and classification[13] tasks of multispectral data,[15] however, there are no specialized classification systems that use machine learning approaches to model dust storms. In literature exist detection methods based on principal components such as the one presented by Hillger *et al.*[16] and Agarwal *et al.*,[17] improve the visualization of dust storms, however, such methods show other objects besides the dust aerosols. Therefore, there is still a need to develop more accurate detection methods.

In this paper we will present five novel models for dust storms that are able to perform to high level of accuracy, and are suitable for real-time applications. This models are based on 8 different methods of feature extraction on the Moderate Resolution Imaging Spectroradiometer (MODIS) data. Such models are: Probabilistic model, Maximum Likelihood Estimation model, Maximum *a posteriori* model, Feedforward Neural Network model, and finally the Probabilistic Neural Network model. The different features extracted from multispectral MODIS data vary between the selection of visual bands and near infrared reflectances. When the models are compared, the neural approach show the best numerical results compared to ground truths from examples found in the literature. Furthermore, the probabilistic models show information not evident in the ground truth giving the ability to find non-trivial dust information.

In Section 2, the formation of a database of events is described. The spectral analysis of dust storms is introduced in Section 3, while in Section 4, the proposed models for dust storms are explained. In Section 5, the design of experiments is presented followed by a discussion of the results and findings. Finally, conclusions are drawn in Section 6.

2. EVENTS DATABASE: DATA COLLECTION

2.1. Dust Storms

We have collected a database of dust storm events using the alerts record from the National Weather Service in Santa Teresa, New Mexico , as well as other events reported by Rivera,[10] Swapna,[12] Britt *et al.*,[6] Lee *et al.*[11], and Hillger *et al.*.[16] In Table 1 is presented a list of different events which include Dust Storm (DS) events. Table 1 show the date of the events, the classification, the time of the observation (in MST), and the scan time of MODIS Terra (in UCT). In Table 2 is presented the downloaded* granule information corresponding to Table 1. The classification code of the events is a standard used in weather forecasts services, except by $C0$ which will be used to denote the background (land, sea, etc.) in the samples.

2.2. Other Events

2.2.1. Wind Blowing with Dust (BLDU)

This kind of events are very similar to a dust storm, however they were not considered as large or strong as a dust storm. This events are characterized by dust transported by the wind, with low concentration of dust aerosols. These events are also shown in Table 1.

2.2.2. Smoke (SM)

Also in Table 1 appear a few events showing smoke. This events were included in the collection since their visual and spatial properties are very similar to the dust, and we want our models to discriminate smoke signatures from dust signatures. Essentially, the smoke is desired to be considered as background.

*The granules were downloaded using NASA's Warehouse Inventory Search Tool, WIST, search tool and directly from the FTP archives.
WIST: https://wist.echo.nasa.gov/~wist/api/imswelcome/
FTP: ftp://ladsftp.nascom.nasa.gov/allData/5/

2.2.3. Validation

The validation collection contain dust storm, blowing dust events, as well as regular days with no dust events. This collection was created with the purpose of validate the results with samples not considered in the design of the classification method and either in the estimation of the probability density functions.

3. DUST STORM SPECTRAL ANALYSIS

It is important to address the selection of the spectral bands that may be relevant to the study of dust storms. Mainly because some of the spectral bands can introduce noise to the model and cause a dramatic drop in the performance of the model. In this section we justify the usage of solar reflective bands 1, 3, and 4, as well as the thermal radiance emissive bands 20, 29, 31, and 32.

3.1. Selection of Spectral Bands for Analysis and Modeling

The basis of MODIS products are MODIS level 1B data. From level 1B bands we can extract all the bands needed for analysis and modeling. It is intuitive that bands $B1, 3$, and 4 can be utilized effectively for visual assessment of the dust storms since we can map such bands to an image of RGB composite ($R = B1, G = B4, B = B3$) and obtain a true color image of MODIS data. This leads to an appropriate visual inspection of the data. However, Hao *et al.*[19] have demonstrated that bands $B20, 29, 31$, and 32 can be utilized effectively for visual enhancement of dust storms. Furthermore, Miller *et al.*[20] and Liu *et al.*,[21] proposed an algorithm for dust storms visual representation where the dust storms appear in a different color than the background.

In previous work from Ackerman *et al.*[22] it is shown that the band math approach provide with a visual enhancement of the dust storms. This approach proposes that bands $B32$ and $B31$ should be subtracted to provide a good contrast of the image containing dust storms.

Based on the previously cited research work we have selected seven bands for our analysis and modeling: $B1, 3, 4, 20, 29, 31$, and 32. It is also considered the band math approach subtracting bands $B32 - 31$ providing with an additional "derived" band for a total of eight bands for our study. As an example, we selected a random dust storm event from the collection and generated images from the bands we selected for study. In Figure 3 is shown a true color image in an RGB composite from bands $B1, 3, 4$, intuitively we can see the dust storm with the color corresponding to the human eye visual spectrum; while in Figure 4 is shown a false color image from $B20, 31, 32$, into an RGB composite, displaying the dust regions with a near pink color. The band subtraction approach is exemplified in Figure 5 subtracting bands $B32 - 31$ and displayed in grayscale to emphasize how lighter colors (near white) are associated with dust information. Note that the displayed images have a correction based on unusable pixel data and the uncertainty index which is explained in the following section.

3.1.1. Considering Uncertainty in Data

Unusable Data MODIS level 1B raw data is stored in 16 bit range. Typically all the data is stored in the first 15 most significant bits (on the range $[0, 32767]$). However, if for any reason the detector is unable to read data, the pixel will contain a value over this range, and will be typically 65535. Such data must not be used in the design of our models since it will bring uncertainty and bias to the final model. Therefore, instead of ignoring those regions in the modeling, we propose to use a series of median and mean filtering preceded of decimation filters in order to estimate the information that is lost. Then we utilize this approximation to also recover data using the uncertainty index.

Uncertainty Index The Uncertainty Index (UI) is an 8-bit matrix structure (of unsigned integers) associated to the bands being considered. The four least significant bits represent MODIS Characterization and Support Team (MCST)'s best estimate of the uncertainty in each measurement of reflectance for the reflective Solar bands, and radiance for the thermal emissive bands. The indices take on values of 0 through 15. The value of 15 is reserved for representing uncertainties of those data that cannot be calibrated, or when the calculated index is higher than 15. The UI is computed in accordance to the measured percent uncertainty as follows

$$UI^{(Bm)} = \ell^{(Bm)} \ln \left(\frac{ui_{\mathbf{n}}^{(Bm)}}{sui^{(Bm)}} \right), \tag{1}$$

where Bm is the $m-$band and is the least rapidly varying index, $\mathbf{n} = [n_1, n_2]$ for n_1 denoting the number of the row in the captured image, and n_2 is the number of the column and the most rapidly varying index; ui is the uncertainty index in percent, and sui is the specified uncertainty; ln denotes the natural logarithm, and ℓ is a scaling factor.

The percent uncertainty ui can be recovered from the uncertainty index UI as follows

$$ui_{\mathbf{n}}^{(Bm)} = sui^{(Bm)} e^{\frac{UI_{\mathbf{n}}^{(Bm)}}{\ell^{(Bm)}}}, \tag{2}$$

$$ci_{\mathbf{n}}^{(Bm)} = 1 - \frac{ui_{\mathbf{n}}^{(Bm)}}{100}, \tag{3}$$

where e denotes the exponential (also denoted as "exp"), and $ci_{\mathbf{n}}^{(Bm)}$ is simply the rational expression of the $ui_{\mathbf{n}}^{(Bm)}$ but in the range $[0, 1]$ with the opposite meaning as a certainty index.

The scaling factor ℓ and specified uncertainty sui values come from band-dependent look up tables and are attached as attributes to each uncertainty index within the associated MODIS Level 1B file. The attributes are provided for convenience to those downstream users who convert the uncertainty index to percent uncertainty, given that the values of ℓ and sui are subject to change.

In our models we choose to diminish the effect of the uncertainty by performing a weighted average between the actual sensed data, and an approximation to the unknown true value. The weighting is based in the rational certainty index $ci_{\mathbf{n}}^{(Bm)}$ defined in (3).

This process has a two folded purpose: first, to produce accurate models from as close to true data as possible, as well as to test the robustness of such models to operate under uncertainty. The named process of estimation of true data based on uncertainty and unusable data is presented in Figure 6. An example of an input and output random event band is shown in Figure 7 which contain unusable pixels and uncertainty.

Quantitative Analysis of the Proposed Method The performance of the proposed method for estimation of true spectral data is evaluated with different metrics by first using two synthetic bands, and second, using real life spectral bands. In the first test, the synthetic bands were degraded by setting an arbitrary matrix as multiplicative noise associated to the uncertainty index of a real life spectral band; in other words, we take the $ci_{\mathbf{n}}^{(Bm)}$ of a random spectral band, and use it as a multiplicative noise. Then, the images were contaminated with non usable pixels. The performance metrics follow with a brief description.

Normalized Mean Squared Error (NMSE). The NMSE is given in percentage, and is described as[23]

$$NMSE = 100 \left(\frac{\mathrm{var}\left(\widehat{B}m - Bm\right)}{\mathrm{var}\left(Bm\right)} \right) \quad \%, \tag{4}$$

where var$(\cdot)$ is the variance, Bm is the original spectral band, and $\widehat{B}m$ is either a degradation or an approximation to Bm.

Signal to Noise Ratio (SNR). The SNR of an image can be interpreted as how much of the original spectral band we have. The higher the number, the better. The SNR is given in dB, and is defined as[23]

$$SNR = 10 \log \left(\frac{NMSE \text{ in } \%}{100} \right)^{-1} \quad \text{dB}, \tag{5}$$

Peak Signal-to-Noise Ratio (PSNR) The PSNR is also used to measure the difference between two spectral bands, and it is very common because in many applications it provides with a quantitative measure that represent better the amount of likelihood between signals when compared with other measures. The PSNR is also given in dB and it is defined as

$$PSNR = 20\log_{10}\left(\frac{b}{\sqrt[2]{\epsilon^2}}\right)^{-1} \quad \text{dB}, \tag{6}$$

where b is the largest possible value of the signals (typically 255 or 65535 for 8 and 16 bit images respectively), and ϵ is the mean of the difference between the original image and an approximation or degradation: $\epsilon = \frac{1}{\mathbf{n}}\sum_{\mathbf{n}} Bm_{\mathbf{n}} - \widehat{B}m_{\mathbf{n}}$, where $\mathbf{n} = [n_1, n_2]$ are the spatial coordinates of the spectral band.

Weighted Signal-to-Noise Ratio (WPSNR). The WPSNR is other useful variant of the SNR. It uses a contrast sensitivity function (CSF)[25] to weight the spatial frequency of a spectral band error. It is given in dB as

$$WPSNR = 20\log_{10}\left(\frac{1}{\sqrt[2]{\varepsilon^2}}\right) \quad \text{dB}, \tag{7}$$

where ε is the mean given by $\varepsilon = \frac{1}{\mathbf{n}}\sum_{\mathbf{n}} \tau_{\mathbf{n}}$, $\tau_{\mathbf{n}} = CSF\left\{Bm_{\mathbf{n}} - \widehat{B}m_{\mathbf{n}}\right\}$, and $CSF\{\cdot\}$ denotes the transformation (filtering) given by the sensitivity function CSF.[24]

Using all he criteria specified previously, we present the results over synthetic and real spectral data. In Table 3 we observe the performance metrics compared against non restored data; clearly the proposed estimation of true data, performs a very good restoration. The comparison against different methodologies remain pending.

4. DUST STORM MODELING FOR CLASSIFICATION

The classification methodologies and pattern recognition techniques for remotely sensed data, involve the usage of probabilistic methods. The reason for this is they reliability, and also the fact that the expected output result can be determined intuitively. In the other hand, soft computing methods such as the neural networks, are used as a black boxes with no direct control of the model, unless supervised learning techniques are carefully chosen. However, soft computing techniques have been evolving and particularly the fuzzy logic and neural networks field. In many applications, neural networks can even perform better in comparison with the probabilistic methodologies for classification of remotely sensed data.[26]

In this section we describe five models for classification. First, we consider a simple probabilistic method based on the individual probabilities as a function of two random variables. Second, a probabilistic method based on the Maximum Likelihood assuming no independence of random variables. Third, a probabilistic method based on the Maximum a Posteriori classification (MAP). Fourth, a soft computing method based on a four-layered feed forward backpropagation neural network with feature dependent architecture. And finally the fifth model is a soft computing method based on a neuro-fuzzy classifier, which based on a possibilistic neural network.

4.1. Probabilistic Modeling as a Function of Two Random Variables

In this section we describe a simple probabilistic formulation and parameter estimation for a model of dust storms provided MODIS data. We start with the following definitions.

Definition 4.1. *Let X be a discrete random variable associated to the universe Γ of values for hyperspectral remote sensing data.* $\Gamma \in \Re$.

Definition 4.2. *Let $X^{(Bm)}$ be the a random variable associated with the values of the m-th spectral band of MODIS.*

Definition 4.3. *Let $X_{\mathbf{n}}^{(Bm)}$ be the random variable associated with the $\mathbf{n}$-th pixel of the m-th spectral band of MODIS.*

Definition 4.4. *Let $f_{X_{\mathbf{n}}^{(Bm)}}(X_{\mathbf{n}}^{(Bm)} = \mathbf{x})$ denote the probability density function of the $\mathbf{n}$-th pixel of the m-th spectral band of MODIS to have a value equal to x.*

In this probabilistic classification method we are interested on displaying the probability of the presence of a Dust Storm given MODIS data, based in the spectral band subtraction $B32-31$. Thus we are interested in the modeling of $f_{X_{\mathbf{n}}^{(B32-31)}}(X_{\mathbf{n}}^{(B32-31)} = x)$, which could be modeled assuming a Gaussian distribution as follows

$$f_{X_{\mathbf{n}}^{(B32-31)}}(X_{\mathbf{n}}^{(B32-31)} = x) = \frac{1}{\sqrt{2\pi\sigma^2_{X^{(B32-31)}}}} e^{-\frac{1}{2}\left(\frac{x-\mu_{X^{(B32-31)}}}{\sigma^2_{X^{(B32-31)}}}\right)} \tag{8}$$

where $\mu_{X^{(B32-31)}}$ is the expected value of the random variable $X^{(B32-31)}$ and $\sigma^2_{X^{(B32-31)}}$ is the variance associated with the random variable. MODIS band subtraction $B32-31$ is assumed to produce a random variable $X^{(B32-31)}$.

The PDF $f_{X_{\mathbf{n}}^{(B32-31)}}(X_{\mathbf{n}}^{(B32-31)} = x)$, indeed, is theoretically defined as a function of two random variables $g\left(X_{\mathbf{n}}^{(Bm)}\right)$, more specifically, a difference of two random variables, as follows

$$X_{\mathbf{n}}^{(B32-31)} = g\left(X_{\mathbf{n}}^{(Bm)}\right) \tag{9}$$

$$g\left(X_{\mathbf{n}}^{(Bm)}\right) = X_{\mathbf{n}}^{(B32)} - X_{\mathbf{n}}^{(B31)} \tag{10}$$

with mean $\mu_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ and variance $\sigma^2_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ defined by

$$\begin{aligned} \mu_{g\left(X_{\mathbf{n}}^{(Bm)}\right)} &= E\left[X_{\mathbf{n}}^{(B32)} - X_{\mathbf{n}}^{(B31)}\right] \\ &= \int_{X_{\mathbf{n}}^{(B32)}} \int_{X_{\mathbf{n}}^{(B31)}} \left(X_{\mathbf{n}}^{(B32)} - X_{\mathbf{n}}^{(B31)}\right) f_{X_{\mathbf{n}}^{(B32)}, X_{\mathbf{n}}^{(B31)}}(X_{\mathbf{n}}^{(B32)} = x, X_{\mathbf{n}}^{(B31)} = x), \end{aligned} \tag{11}$$

$$\begin{aligned} \sigma^2_{g\left(X_{\mathbf{n}}^{(Bm)}\right)} &= \mathrm{var}\left[X_{\mathbf{n}}^{(B32)} - X_{\mathbf{n}}^{(B31)}\right] \\ &= E\left[\left(X_{\mathbf{n}}^{(B32)} - X_{\mathbf{n}}^{(B31)} - E\left[X_{\mathbf{n}}^{(B32)} - X_{\mathbf{n}}^{(B31)}\right]\right)^2\right], \end{aligned} \tag{12}$$

where $E[\cdot]$ denotes the expected value, $\mathrm{var}[\cdot]$ is used to denote the variance, and $f_{X_{\mathbf{n}}^{(B32)}, X_{\mathbf{n}}^{(B31)}}(X_{\mathbf{n}}^{(B32)} = x, X_{\mathbf{n}}^{(B31)} = x)$ is the joint probability density function. This three parameters are unknown, moreover, we cannot assume independence in the bands since they are highly correlated as in most remote sensing applications.[27] Therefore, the attempt of showing proof of independence will not be addressed in this paper.

In spite of the fact that the latter parameters are unknown, they can be estimated by observation of the data. The only important variable to observe at this point is the function $g\left(X_{\mathbf{n}}^{(Bm)}\right)$ from (10) used in (8), and the estimation based on such observations is introduced in the following sub-section.

4.1.1. Estimation of $\mu_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ and $\sigma^2_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$.

In our experiments we to obtain approximations to the true $\mu_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ and $\sigma^2_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$, which from here will be referred to as $\widehat{\mu}_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ and $\widehat{\sigma}^2_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ respectively. Such approximations are estimated over the number of events selected for modeling and design (see Section 2). In the approximation process we observe several samples (pixels), in the order of millions since 23 out of 31 events were selected for modeling, which lead to a more accurate estimation of the required parameters. The dust regions were extracted manually and frequency histograms were computed in each one. The manual segmentation process was entirely based on the reported events referenced in Table 1. We have approximated the parameters in three different ways: Global, Eventual, and Averaged.

Estimation Based on Global Histogram This methodology consists on computing the histogram for all the events and from the total frequency observed we have estimated the sample mean $\widehat{\mu}^{(\text{global})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}$, and the standard deviation $\widehat{\sigma}^{(\text{global})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}$.

Estimation Based on Eventual Histogram The histogram was computed individually for all the events, and the sample mean $\widehat{\mu}^{(\text{event }\#)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}$, and the standard deviation $\widehat{\sigma}^{(\text{event }\#)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}$ was estimated for each one. Having a vector of sample means

$$\widehat{\boldsymbol{\mu}}^{(\text{events})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} = \left[\widehat{\mu}^{(\text{event }1)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}, ..., \widehat{\mu}^{(\text{event }k)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}\right], \tag{13}$$

as well as a vector of standard deviations

$$\widehat{\boldsymbol{\sigma}}^{(\text{events})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} = \left[\widehat{\sigma}^{(\text{event }1)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}, ..., \widehat{\sigma}^{(\text{event }k)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}\right], \tag{14}$$

and the final estimation of $\widehat{\mu}^{(\text{eventual})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}$ is defined as

$$\widehat{\mu}^{(\text{eventual})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} = \frac{1}{k}\sum_{l=1}^{k}\widehat{\mu}^{(\text{event }l)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} \tag{15}$$

and the final estimation of $\widehat{\sigma}^{(\text{eventual})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}$ is

$$\widehat{\sigma}^{(\text{eventual})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} = \frac{1}{k}\sum_{l=1}^{k}\widehat{\sigma}^{(\text{event }l)}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}. \tag{16}$$

Estimation Based on Average of Estimations This is a combination of the previous estimation defined as the non-biased average between the two previous estimations, which formally is:

$$\widehat{\mu}^{(\text{average})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} = \frac{\widehat{\mu}^{(\text{global})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} + \widehat{\mu}^{(\text{eventual})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}}{2}, \tag{17}$$

$$\widehat{\sigma}^{(\text{average})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} = \frac{\widehat{\sigma}^{(\text{global})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)} + \widehat{\sigma}^{(\text{eventual})}_{g\left(X^{(Bm)}_{\mathbf{n}}\right)}}{2}. \tag{18}$$

After the completion of the process we obtain the total histograms an final estimations. In Figure 8, we observe the estimation over 16-bit data, while the estimation for 8-bit data is shown in Figure 9. The estimation based on the recovered radiances is shown in Figure 10. In these figures, the total histogram (in blue) is shown in percent, while the different methods of estimation are shown in the following colors: Global in red, Local in green, and Averaged in Pink.

This classification method was designed such that the output $f_{X^{(B32-31)}_{\mathbf{n}}}(X^{(B32-31)}_{\mathbf{n}} = x)$ can be mapped to a visual representation of the probability of a dust event, and not for segmentation. However, the segmentation or classification between background and Dust Storm can be achieved by thresholding the output as desired, using the following decision rule

$$\begin{array}{lll} x \in DS & \text{if} & f_{X^{(B32-31)}_{\mathbf{n}}}(X^{(B32-31)}_{\mathbf{n}} = x) \quad > \quad \tau \quad , \quad 0 \leq \tau \leq 1 \\ x \in C0 & & \text{otherwise} \end{array} \tag{19}$$

where τ is the threshold, DS denotes the dust storm class, and $C0$ is the background class. This threshold can be estimated according to the confidence interval required by the final application, or also, can be tuned experimentally by the end user.

4.2. Probabilistic Modeling Based on the Multivariate Maximum Likelihood Classifier

DEFINITION 4.5. *Let $f_{C_j}(C_j = c)$ be the probability mass function of the j-th class C to occur with a value of c.*

DEFINITION 4.6. *Let $f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right)$ be the conditional probability density function of the $\mathbf{n}-th$ pixel of the m-th spectral band to have a value of x given the probability that the j-th class occur with a value of c. This might be referred to as the "a priori" probability.*

The Maximum Likelihood Classifier[28] (Maximum Likelihood Estimator, MLE), is based on the prior probabilities assuming that the posterior probabilities are unknown. The MLE is an accepted measure for classification and analysis of remotely sensed data.[29] Therefore we will model an MLE classifier specialized on classification of dust and background. In our model will use four classes, where $C_j = c$ should take the following values

$$c = \begin{cases} 1 = C0 & , \text{ for land/sea background} \\ 2 = SM & , \text{ for smoke} \\ 3 = BLDU & , \text{ for blowing dust} \\ 4 = DS & , \text{ for dust storm} \end{cases} , \tag{20}$$

therefore $C_j = c$ is defined over $j = 1, 2, 3, 4$ as described above ($C_1 = c = C0$, $C_2 = c = SM$, etc.), and the probability density functions defined above will be estimated from the collected samples.

The MLE can be derived from Bayes theorem which states that

$$f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right) = \frac{f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right)}{f_{X_{\mathbf{n}}^{(Bm)}}\left(X_{\mathbf{n}}^{(Bm)} = x\right)} \tag{21}$$

where $f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right)$ is called the posterior probability. Since we are interested on finding the maximum likelihood between the observed data and the prior probability for all classes, we can state a decision rule as follows

$$\begin{aligned} x \in C_j \text{ if, } \quad & \frac{f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right)}{f_{X_{\mathbf{n}}^{(Bm)}}\left(X_{\mathbf{n}}^{(Bm)} = x\right)} > \\ & \frac{f_{X_{\mathbf{n}}^{(Bm)}|C_i}\left(X_{\mathbf{n}}^{(Bm)} = x|C_i = c\right) f_{C_i}\left(C_i = c\right)}{f_{X_{\mathbf{n}}^{(Bm)}}\left(X_{\mathbf{n}}^{(Bm)} = x\right)}, \quad \forall j \neq i \end{aligned} \tag{22}$$

which can be simplified by removing the common factor $f_{X_{\mathbf{n}}^{(Bm)}}\left(X_{\mathbf{n}}^{(Bm)} = x\right)$, to the following

$$\begin{aligned} x \in C_j \text{ if, } \quad & f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right) > \\ & f_{X_{\mathbf{n}}^{(Bm)}|C_i}\left(X_{\mathbf{n}}^{(Bm)} = x|C_i = c\right) f_{C_i}\left(C_i = c\right), \quad \forall j \neq i \end{aligned} \tag{23}$$

which will be modeled under the assumption that the prior probability is normally distributed. For convenience we can rewrite the above terms by defining

$$\psi_{k,\mathbf{n}}^{(Bm)}(x) = f_{X_{\mathbf{n}}^{(Bm)}|C_k}\left(X_{\mathbf{n}}^{(Bm)} = x|C_k = c\right) f_{C_k}\left(C_k = c\right) \tag{24}$$

that allows restating the decision rule as follows

$$x \in C_j \quad \text{if} \quad \psi_{j,\mathbf{n}}^{(Bm)}(x) > \psi_{i,\mathbf{n}}^{(Bm)}(x) \quad \forall j \neq i \tag{25}$$

which is more convenient to handle. In MLE, the functions $\psi_{k,\mathbf{n}}^{(Bm)}(x)$ are commonly referred to as *discriminant functions*.

The discriminant functions $\psi_{k,\mathbf{n}}^{(Bm)}(x)$ can be reduced because of the gaussianity assumption of the prior PDF as follows

$$\begin{aligned}
\psi_{k,\mathbf{n}}^{(Bm)}(x) &= f_{X_{\mathbf{n}}^{(Bm)}|C_k}\left(X_{\mathbf{n}}^{(Bm)} = x|C_k = c\right) f_{C_k}\left(C_k = c\right) && (26)\\
&= \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det\left(\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)}} e^{-\frac{1}{2}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)^T \mathbf{\Sigma}^{-1}_{X_{\mathbf{n}}^{(Bm)}|C_k}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)} f_{C_k}\left(C_k = c\right)\\
&= \frac{1}{(2\pi)^{\frac{d}{2}}\det\left(\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)^{\frac{1}{2}}} e^{-\frac{1}{2}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)^T \mathbf{\Sigma}^{-1}_{X_{\mathbf{n}}^{(Bm)}|C_k}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)} f_{C_k}\left(C_k = c\right) && (27)\\
&= -\frac{d}{2}\ln(2\pi) - \frac{1}{2}\det\left(\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right) && (28)\\
&\quad -\frac{1}{2}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)^T \mathbf{\Sigma}^{-1}_{X_{\mathbf{n}}^{(Bm)}|C_k}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right) + \ln\left(f_{C_k}\left(C_k = c\right)\right)
\end{aligned}$$

where d is the dimension size of x which is related to the dimension of the features, $\det(\cdot)$ is the determinant function, $\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_k}$ is the covariance matrix of the prior PDF, $\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}$ denotes the mean vector, $(\cdot)^T$ denotes the transpose operation, $\mathbf{\Sigma}^{-1}_{X_{\mathbf{n}}^{(Bm)}|C_k}$ is the inverse of the covariance matrix $\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_k}$, and $\ln(\cdot)$ is the natural logarithm.

It is said[30] that the factor $-\frac{d}{2}\ln(2\pi)$ adds no discriminant information to the classification, and often its removed. Also when the uncertainty is high for the PMF $f_{C_k}(C_k = c)$ it is recommended to be removed from (28), leading to a commonly used simpler discriminant function

$$\psi_{k,\mathbf{n}}^{(Bm)}(x) = -\det\left(\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right) - \left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)^T \mathbf{\Sigma}^{-1}_{X_{\mathbf{n}}^{(Bm)}|C_k}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right) \quad (29)$$

where the common factor was $\frac{1}{2}$ was eliminated as well. However in this model we preserved all the factors since there is a high amount of sample data to perform an adequate estimation of the PMF $f_{C_k}(C_k = c)$. So the final discriminant function utilized in our MLE model (25) is

$$\begin{aligned}
\psi_{k,\mathbf{n}}^{(Bm)}(x) &= \ln\left(f_{C_k}\left(C_k = c\right)\right) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\det\left(\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right) && (30)\\
&\quad -\frac{1}{2}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right)^T \mathbf{\Sigma}^{-1}_{X_{\mathbf{n}}^{(Bm)}|C_k}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_k}\right).
\end{aligned}$$

In the following sections we explain the features selected for modeling dust storms using MLE, and the we explain the estimation of the prior PDF's $f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right)$ and $f_{C_j}(C_j = c)$, as well as the parameters that define them in order to be used in the discriminant function $\psi_{k,\mathbf{n}}^{(Bm)}(x)$.

4.2.1. Features and Events Selection

The selection of features is a very important step in classification since the performance is proportional to the relevancy of the data.[26] As explained before (see Section 3.1) in our study we will model dust storms based on certain bands only: $B1, 3, 4, 20, 29, 31, 32$. We have defined different feature vectors depending on the final application. In the following paragraphs we explain these features and their intended target application.

Features for Offline Radiance Intensity Analysis. We use the term "Offline" to refer to applications that do not require simpler nor faster processing algorithms, such as real time applications, rapid response applications, etc. In this category we work directly with the 16 bit MODIS data with no radiance recovering. This means that instead of having the original scale (W/m^2/μm/sr), we will work directly with the raw intensity

values. For modeling and design of the classification methods, all the intensities were processed to decrease uncertainty by using the method proposed in Section 3.1. The feature vector $F^{(1)}$ is constructed as follows

$$F^{(1)} = [B20, B29, B31, B32, (B32 - 31)], \tag{31}$$

where $F^{(m)}$ will denote just a label to distinguish between feature vectors, and does not represent any mathematical operation (i.e. $F^{(2)}$ denotes the feature vector #2, not the square of F). The feature $B32 - 31$ denotes the subtraction of band $B31$ from band $B32$. In Figure 11 is shown a scatter plot from two elements of $F^{(1)}$. This plot demonstrate the non-triviality of the dust storm problem.

Features for Offline Recovered Radiance Analysis. We use the term "Recovered" to refer 16 bit MODIS data that is recovered to its original scale (W/m^2/μm/sr). The recovery process is given by the following equation

$$L_{\mathbf{n}}^{(Bm)} = rs^{(Bm)}(si_{\mathbf{n}}^{(Bm)} - ro^{(Bm)}) \tag{32}$$

where $L_{\mathbf{n}}^{(Bm)}$ denotes the recovered radiances, $rs^{(Bm)}$ are the radiance scales, $ro^{(Bm)}$ are the radiance offsets, and $si_{\mathbf{n}}^{(Bm)}$ are the scaled intensities, or raw data (used in $F^{(1)}$). After recovering process, the radiances were processed to decrease uncertainty. The feature vector $F^{(2)}$ is constructed as follows

$$F^{(2)} = \left[L^{B20}, L^{B29}, L^{B31}, L^{B32}, \left(L^{B32} - L^{B31}\right)\right]. \tag{33}$$

The feature $L^{B32} - L^{B31}$ defines that the subtraction of band $B31$ from band $B32$ is performed after the recovery process. In Figure 12 is shown a scatter plot of two elements of $F^{(2)}$.

Features for Online Visual Band Applications. We use the term "Online Visual Assessment" to refer to applications that will work only on the visual spectral bands, that is $B1, 3, 4$, and are required to produce a rapid result. In this category we work with scaled to 8 bit MODIS data with no radiance recovering. For modeling, all the intensities were processed to decrease uncertainty. The feature vector $F^{(3)}$ is constructed as follows

$$F^{(3)} = \left[\widehat{B}1, \widehat{B}3, \widehat{B}4\right], \tag{34}$$

where $\widehat{B}$ points out that the original 16-bit data has been downscaled to 8-bit data. The scatter plot of Figure 13 includes two columns of $F^{(3)}$.

Features for Online Assessment Applications. This features are modeled in the same way as the previous with the difference that we don't work with the visual spectral band, instead we work with downscaled non-recovered raw radiances, defining the feature vector as

$$F^{(4)=} \left[\widehat{B}31, \widehat{B}32, \left(\widehat{B}32 - \widehat{B}31\right)\right], \tag{35}$$

which is plotted in Figure 14 to observe the sparsity of the data.

Selection of Events for Modeling. The events referenced in Section 2.1 show an indication of which events are utilized for modeling and which are utilized for validation. The information is shown in Table 1 and in Table 2, having that the events noted with (*) are left out for validation. Note that in spite of the fact that many of the events are selected for modeling, not all the pixels are actually used to construct the final model. The pixel-based selection will be explained in detail later at Section 5.1.

4.2.2. Estimation of Probability Density Functions

Typically in remote sensing image processing and classification, the problem of small sample size does not exist, instead, the problem of large sample size, excessive dimensions and relevant samples selection are the most common issues. In our models we face the problem of excessive sample size, which could be translated in large amounts of time for the estimation of parameters. This problem is also experienced when neural networks and other sophisticated learning-based methodologies are used. However, it is relatively easy and less time-consuming if we estimate the densities under Gaussian assumptions since the estimation of sample means, variances, and covariances require less processing time if compared to the training phase of a neural network based on backpropagation.

In MLE, the estimation of the densities is performed by histogram observation on the feature vectors described in Section 4.2.1 and for each one of the events selected for modeling, as described in Section 2.1. The histograms for all samples were added to get one final histogram for each band and perform estimation over such histograms by observing the distribution. In this section we explain the estimation process for the *a priori* probabilities $f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right)$ and $f_{C_j}\left(C_j = c\right)$ required for the MLE method.

Estimation of the *a priori* Conditional Probability Density Function $f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right)$. It is useful to know, in a general sense, the uni-variate PDF of $f_{X_n^{(Bm)}}(X_n^{(Bm)} = x)$, aiming to understand if Gaussian assumptions are appropriate to model the spectral bands we are studying. With this purpose we estimate simple uni-variate PDF's and present them in Figures 15-30. In these figures are shown all the PDF's, as well as the estimation of parameters under the assumption of gaussianity. Such parameters are the sample mean μ and standard deviation σ, (variance σ^2) for the uni-variate normal (or unidimensional Gaussian). The unidimensional Gaussian probability density function is given by

$$f_{X_n^{(Bm)}}(X_n^{(Bm)} = x) = \frac{1}{\sqrt{2\pi\sigma^2_{X^{(Bm)}}}} e^{-\left(\frac{x-\mu_{X^{(Bm)}}}{\sigma^2}\right)} = N(\mu_{X^{(Bm)}}, \sigma_{X^{(Bm)}}), \tag{36}$$

where $N(\mu_{X^{(Bm)}}, \sigma_{X^{(Bm)}})$ is a short form to denote a Gaussian distribution. Since we are approximating the true values, we will use $\widehat{\mu}_{X^{(Bm)}}$, and $\widehat{\sigma}_{X^{(Bm)}}$, in the following models. The approximation of parameters will be more accurate as we have as many samples as possible, in other words, the approximation converges to the true parameters when we have an infinite number of samples. In the estimation of $f_{X_n^{(Bm)}}(X_n^{(Bm)} = x)$ we have utilized ≈75 millions of sample pixels. From the figures mentioned above, it is clear that Gaussian assumptions are appropriate for our models. In Figure 32, we show the histogram of $B20$ raw intensities, as well as different probability density functions and how they fit the band; once again, it is clear that a Gaussian can be acceptable.

After justifying the Gaussian assumption, we can proceed to the estimation of the conditional prior probability density function $f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right)$. Here, we follow a different process since the goal is to estimate a set of parameters for each class. This parameters are: the vector of expected values $\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_j}$ and the covariance matrix $\boldsymbol{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_j}$. Since we are dealing with estimations we use the notation $\widehat{\boldsymbol{\mu}}_{X_{\mathbf{n}}^{(Bm)}|C_j}$ and $\widehat{\boldsymbol{\Sigma}}_{X_{\mathbf{n}}^{(Bm)}|C_j}$ to refer to the estimated parameters.

Then we can define the set of parameters to estimate as

$$\mathbf{E}_k = \left\{ \widehat{\boldsymbol{\mu}}^T_{F^{(1)}|C_1} \quad , \quad \ldots \quad , \quad \widehat{\boldsymbol{\mu}}^T_{F^{(1)}|C_j} \quad , \quad \widehat{\boldsymbol{\mu}}^T_{F^{(2)}|C_1} \quad , \quad \ldots \quad , \quad \widehat{\boldsymbol{\mu}}^T_{F^{(4)}|C_j} \right\} \tag{37}$$

and

$$\mathbf{S}_k = \left\{ \widehat{\boldsymbol{\Sigma}}_{F^{(1)}|C_1} \quad , \quad \ldots \quad , \quad \widehat{\boldsymbol{\Sigma}}_{F^{(1)}|C_j} \quad , \quad \widehat{\boldsymbol{\Sigma}}_{F^{(2)}|C_1} \quad , \quad \ldots \quad , \quad \widehat{\boldsymbol{\Sigma}}_{F^{(4)}|C_j} \right\} \tag{38}$$

where $\mathbf{E}_k$ can be defined as either a matrix of expected values of size $[k \times 4]$, or as a set of expected values containing 4 vectors associated with the estimated expected values; in the other hand $\mathbf{S}_k$ is a multidimensional set of covariance matrices of different sizes depending on the size of the feature vector. For instance, the first

element of the set is the covariance matrix $\widehat{\mathbf{\Sigma}}_{F^{(1)}|C_1}$ of size $[5 \times 5]$, while the last element of the set $\mathbf{S}_k$ is the covariance matrix $\widehat{\mathbf{\Sigma}}_{F^{(4)}|C_k}$ of size $[4 \times 4]$.

For the estimation, we have manually segmented the sample images containing dust storms, blowing dust, and smoke, such that we can utilize the region of the segmented image as a mask. This mask is associated with a particular class C_j. Then having all the masks and the associated classes, we extract and store the subset of pixels associated to the $j-$th class. Then, we compute the sample means of the $l-$th feature vector for the $j-$th class as follows

$$\widehat{\boldsymbol{\mu}}_{F^{(l)}|C_j} = \left[\begin{array}{ccc} \widehat{\mu}_{F^{(l,1)}|C_j} & \cdots & \widehat{\mu}_{F^{(l,m)}|C_j} \end{array} \right] \tag{39}$$

where $\widehat{\mu}_{F^{(l,m)}|C_j}$ is the sample mean of $m-$th element for the $l-$th feature vector of the $j-$th class and is formulated as follows

$$\widehat{\mu}_{F^{(l,m)}|C_j} = \frac{1}{n_{F^{(l)}|C_j}} \sum_{r=1}^{n_{F^{(l)}|C_j}} F_r^{(l,m)}|C_j \tag{40}$$

where $n_{F^{(l)}|C_j}$ is the number of samples. Then we compute the covariance matrices over all the subset of pixels of the $l-$th feature vector associated to the $j-$th class, as follows

$$\widehat{\mathbf{\Sigma}}_{F^{(l)}|C_j} = \frac{1}{n_{F^{(l)}|C_j}} \sum_{r=1}^{n_{F^{(l)}|C_j}} \left(F^{(l)}|C_j - \widehat{\boldsymbol{\mu}}_{F^{(l)}|C_j}\right) \left(F^{(l)}|C_j - \widehat{\boldsymbol{\mu}}_{F^{(l)}|C_j}\right)^T \tag{41}$$

where the covariance estimation problem is clearly not ill-posed since we have a large number of data samples available for modeling as in most remote sensing applications. Figure 31 present a graphical explanation of the process of parameter estimation for the multivariate Gaussian $f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right)$.

A Note on the Masks. The masks were created by the author (P.R.P) manually by visual inspection of the spectral bands $B1, B3, B4$, and $B32 - B32$. In all the cases the information was confirmed against papers published and other research publications as detailed in Section 2. The main objective of the masks is to contain the pixels that best represent information of $j-$th class (i.e. the dust storm), and as less as possible of everything else (i.e. background, clouds, smoke), such that the algorithms can generalize and detect other regions of the $j-$th class but with less probability, for instance: if the main class is $j = DS$, then the algorithm is desired to be able to detect both the dust storm with high probability, and the dust storm transport with less probability.

Estimation of the *a priori* Probability Mass Function $f_{C_j}(C_j = c)$. With the purpose of estimation of the PMF $f_{C_j}(C_j = c)$, we have utilized the masks described above, and counted the frequency of pixels associated with each class. Thus, the PMF can be denoted as follows

$$f_{C_j}(C_j = c) = \begin{cases} \frac{\#\text{samples} \in Cj=c}{\#\text{samples} \forall C_j} & , \text{ for } f_{C_j}(C_j = c) \\ 0 & , \text{ otherwise} \end{cases} \tag{42}$$

where intuitively, being a class specific remote sensing application, the PMF is far from being uniform. The estimated PMF is shown in Figure 33, and we also provide a logarithmic plot in Figure 34 in order to have a better appreciation of the PMF estimation.

4.3. Probabilistic Modeling Based on the Maximum *a Posteriori* Classifier

Definition 4.7. *Let $f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right)$ be the conditional probability density function of the j-th class to occur with a value of c, given the probability that the $\mathbf{n}-$th pixel of the m-th spectral band has a value of x. Also this is commonly referred to as the posterior or "a posteriori" probability.*

The Maximum *a posteriori* (MAP) approach is different from the MLE approach in that MAP estimates the posterior probability while MLE deals with the prior probability. This can be explained by reformulating the decision rule of MLE

$$x \in C_j \text{ if, } \quad \frac{f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right)}{f_{X_{\mathbf{n}}^{(Bm)}}\left(X_{\mathbf{n}}^{(Bm)} = x\right)} > \frac{f_{X_{\mathbf{n}}^{(Bm)}|C_i}\left(X_{\mathbf{n}}^{(Bm)} = x|C_i = c\right) f_{C_i}\left(C_i = c\right)}{f_{X_{\mathbf{n}}^{(Bm)}}\left(X_{\mathbf{n}}^{(Bm)} = x\right)}, \quad \forall j \neq i, \tag{43}$$

which can be restated provided that

$$\frac{f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right)}{f_{X_{\mathbf{n}}^{(Bm)}}\left(X_{\mathbf{n}}^{(Bm)} = x\right)} = f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right), \tag{44}$$

as follows

$$x \in C_j \text{ if, } \quad f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right) > f_{C_i|X_{\mathbf{n}}^{(Bm)}}\left(C_i = c|X_{\mathbf{n}}^{(Bm)} = x\right), \quad \forall j \neq i \ , \tag{45}$$

or equivalently

$$x \in C_j \text{ if, } \quad f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right) = \max_{1 \leq i \leq c} f_{C_i|X_{\mathbf{n}}^{(Bm)}}\left(C_i = c|X_{\mathbf{n}}^{(Bm)} = x\right). \tag{46}$$

Then from Bayes theorem we can obtain the posterior probability based on the priors as

$$f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right) = \frac{f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right)}{\sum_{i=1}^{c} f_{X_{\mathbf{n}}^{(Bm)}|C_i}\left(X_{\mathbf{n}}^{(Bm)} = x|C_i = c\right) f_{C_i}\left(C_i = c\right)}, \tag{47}$$

which can be used to reformulate the decision rule by recalling the usage of the discriminant functions

$$\psi_{j,\mathbf{n}}^{(Bm)}(x) = \frac{f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right)}{\sum_{i=1}^{c} f_{X_{\mathbf{n}}^{(Bm)}|C_i}\left(X_{\mathbf{n}}^{(Bm)} = x|C_i = c\right) f_{C_i}\left(C_i = c\right)} \tag{48}$$

where

$$f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)} = x|C_j = c\right) f_{C_j}\left(C_j = c\right) = \frac{1}{(2\pi)^{\frac{d}{2}} \det\left(\mathbf{\Sigma}_{X_{\mathbf{n}}^{(Bm)}|C_j}\right)^{\frac{1}{2}}} \times e^{-\frac{1}{2}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_j}\right)^T \mathbf{\Sigma}^{-1}_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(x-\boldsymbol{\mu}_{X_{\mathbf{n}}^{(Bm)}|C_j}\right)} f_{C_j}\left(C_j = c\right), \tag{49}$$

leading to the decision rule

$$x \in C_j \text{ if, } \quad \psi_{j,\mathbf{n}}^{(Bm)}(x) = \max_{1 \leq i \leq c} \psi_{i,\mathbf{n}}^{(Bm)}(x). \tag{50}$$

4.3.1. Features and Parameter Estimation

We constructed four different models based on the four feature vectors described in the MLE classification method: $F^{(1)}, ..., F^{(4)}$. Therefore the estimation of the parameters is the same. For details see Section 4.2.

4.4. Soft Computing Modeling Based on Multilayered Feedforward Backpropagation Neural Networks

In the pattern recognition field, the term Soft Computing is a new term involving a broad of areas such as neural networks, fuzzy systems, support vector machines, etc., as well as most computational intelligence techniques.[31] Multilayered Feedforward Neural Network (FFNN) are of particular interest in pattern recognition and classification applications because they can approximate any square-integrable function to any desired degree of accuracy, and can exactly implement any arbitrary finite training set.[32] There exist many remote sensing data classification problems that have been successfully solved using neural networks,[14] besides dust storms. Therefore we have designed a FFNN to model a dust storm by approximating the probability density function $f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right)$. One of the major advantages of FFNN is that it can overcome the estimated $f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right)$ when it is poorly posed. This ability is well known as "generalization". And it is due mainly because in the approximation of the true $f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right)$, we often assume Gaussian distributions, while a FFNN aims to fit the true density $f_{C_j|X_{\mathbf{n}}^{(Bm)}}\left(C_j = c|X_{\mathbf{n}}^{(Bm)} = x\right)$, however if the PDF is actually Gaussian, the FFNN can be compared with those models under such assumption. If the density represents non-linearly separable classes, the FFNN will try to minimize the error of the separation hyperplane.

A simple FFNN contains an input layer and an output layer, separated by l layers (know as the hidden layer) of neuron units. Given an input sample clamped to the input layer, the other units of the network compute their values according to the activity of the units that they are connected to in the previous layers. In this model we consider the particular topology where the input layer is fully connected to the first hidden layer, which is fully connected to the second layer and so on up to the output layer.

Given an input $x \in X_{\mathbf{n}}^{(Bm)}$, the value of the $j-$th unit in the $i-$th layer is denoted $h_j^i(x)$, with $i = 0$ referring to the input layer, $i = l+1$ referring to the output layer. We refer to the size of a layer as $\left|\mathbf{h}^i(x)\right|$. The default activation level is determined by the internal bias b_j^i of that unit. The set of weights W_{jk}^i between $h_k^{i-1}(x)$ to in layer $i-1$ and unit $h_j^{i-1}(x)$ in layer i determines the activation of unit $h_j^i(x)$ as follows:

$$h_j^i(x) = \Phi\left(a_j^i(x)\right), \tag{51}$$

where

$$\begin{aligned} a_j^i(x) &= \sum_k W_{jk}^i h_k^{i-1}(x) + b_j^i, \\ \forall i &\in \{1, ..., l\}, \end{aligned}$$

with

$$h^0(x) = x,$$

where $\Phi =$sigm$(\cdot)$ is the sigmoid activation function given by

$$\text{sigm}\,(a) = \frac{1}{1+e^{-a}},$$

which could be replaced by any desired activation function. Given the last hidden layer, the output layer is computed similarly by

$$\begin{aligned} \mathbf{o}(x) &= \mathbf{h}^{l+1}(x), \\ \mathbf{h}^{l+1}(x) &= \mathbf{\Phi}\left(\mathbf{a}^{l+1}(x)\right), \end{aligned} \tag{52}$$

where

$$\mathbf{a}^{l+1}(x) = \mathbf{W}^{l+1}\mathbf{h}^l(x) + \mathbf{b}^{l+1},$$

where the activation function $\mathbf{\Phi}$ depends on the (supervised) task the network must achieve. Typically, it will be the *signum function* defined as

$$\mathbf{\Phi} = \operatorname{sgn}(\mathbf{a}) = \begin{cases} +1 & \text{if} \quad \mathbf{a} > \mathbf{0} \\ -1 & \text{if} \quad \mathbf{a} \leq \mathbf{0} \end{cases}$$

for a simple classification problem.[33]

When an input sample x is presented to the network, the application of (51) at each layer will generate a pattern of activity in the different layers of the neural network.

In Figure 35 is illustrated a basic multilayered FFNN network and its parameters. Intuitively, we would like the activity of the first layer neurons to correspond to low-level features of the input (e.g., edges, orientations, intensity differences, etc.) and to higher level abstractions (e.g., detection of multispectral shapes, correlation, PDF approximation, etc.) in the last hidden layers.[34]

4.4.1. Features and Events Selection

To model the Dust Storms with FFNN we will use the same events utilized in previous classification methods. We will also use the same feature vectors $F^{(1)}, ..., F^{(4)}$, to design four FFNN models. Additionally, we introduce four more feature vectors with a two folded motivation: first, de-correlate the training data; and second, reduce the dimensionality preserving as most discriminant information as possible. For this purpose we use the Karhunen-Loeve Transformation (KLT) (which is analogous to the Principal Component Analysis, PCA). In our previous work,[35] we have reported that the combination of KLT features in classification, add discriminant capabilities to the neural-based classifiers.

Karhunen-Loeve Transformation of $F^{(1)}, ..., F^{(4)}$. The KLT is computed on the already defined feature vectors $F^{(1)}, ..., F^{(4)}$ to obtain a new set of feature vectors. This relationship is described as a function of the original feature vectors as follows

$$\begin{aligned} F^{(5)} &= \Upsilon\left(F^{(1)}\right), \\ &\vdots \\ F^{(8)} &= \Upsilon\left(F^{(4)}\right), \end{aligned}$$

where $\Upsilon(\cdot)$ is a transformation function

$$\Upsilon\left(F^{(l)}\right) = F^{(l)}\widehat{\mathbf{T}}_{k-KLT}, \tag{53}$$

having that $\widehat{\mathbf{T}}_{k-KLT}$ is a projection matrix based on the $k-$th columns of the orthogonal matrix $\mathbf{T}_{KLT}$, obtained trough the KLT. The process of obtaining $\mathbf{T}_{KLT}$ is dependent of the set of input vectors $F^{(l)}$.

First, we estimate the sample mean of the $m-$th column of the $l-$th feature vector

$$\mu_{F^{(l,m)}} = \frac{1}{n_{F^{(l)}}} \sum_{r=1}^{n_{F^{(l)}}} F_r^{(l,m)}, \tag{54}$$

where $n_{F^{(l)}}$ is the number of samples available for modeling within $F^{(l)}$, and we perform this for all the columns of the feature vector to construct a vector of sample means

$$\boldsymbol{\mu}_{F^{(l)}} = \left[\mu_{F^{(l,1)}}, ..., \mu_{F^{(l,m)}}\right]. \tag{55}$$

Then we compute the covariance matrix $\Sigma_{F^{(l)}}$ of $F^{(l)}$ as follows

$$\mathbf{\Sigma}_{F^{(l)}} = \frac{1}{n_{F^{(l)}}} \sum_{r=1}^{n_{F^{(l)}}} \left(F^{(l)} - \widehat{\boldsymbol{\mu}}_{F^{(l)}}\right)\left(F^{(l)} - \widehat{\boldsymbol{\mu}}_{F^{(l)}}\right)^T. \tag{56}$$

Then the eigenvalues of $\mathbf{\Sigma}_{F^{(l)}}$ are extracted and defined as $\lambda_1, \lambda_2, ..., \lambda_{n_{F^{(l)}}}$, where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{n_{F^{(l)}}}$. As well as the eigenvectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, ..., \boldsymbol{\phi}_{n_{F^{(l)}}}$ associated to $\lambda_1, \lambda_2, ..., \lambda_{n_{F^{(l)}}}$. Then, the projection matrix $\mathbf{T}_{KLT}$, is defined as

$$\mathbf{T}_{KLT} = \left[\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, ..., \boldsymbol{\phi}_{n_{F^{(l)}}}\right] \tag{57}$$

where the eigenvectors $\boldsymbol{\phi}_{n_{F^{(l)}}}$ are column vectors, and $\mathbf{T}_{KLT}$ is orthogonal such that $\mathbf{I} = \mathbf{T}_{KLT}\mathbf{T}_{KLT}^T$, where $\mathbf{I}$ is the identity matrix. The relevance of KLT lies on the fact that if we perform the projection

$$\widehat{F}^{(l)} = F^{(l)}\mathbf{T}_{KLT}, \tag{58}$$

it will produce a feature vector $\widehat{F}^{(l)}$ projected in a new eigen-space where the data is not correlated leading to more discriminant features. Furthermore, if we select only the $k-$th eigenvectors associated with the $k-$th largest eigenvalues, we can reduce the dimensionality of the feature space by projecting the data into a new sub-space. There are several methods that deal with the appropriate selection of the number of eigenvectors to keep, and in our models our goal is to keep those eigenvectors whose associated eigenvalues magnitude (energy) adds up to 99.99% of the total energy.

Our experiments in projecting all the features, showed that in average, 99.99% of the energy is concentrated in the first 2 eigenvalues, therefore, we decided to preserve only the 2 eigenvectors associated with the 2 largest eigenvalues, that is

$$\begin{aligned} \widehat{\mathbf{T}}_{k-KLT} &= \left[\boldsymbol{\phi}_{n_{F^{(l)}}-1}, \boldsymbol{\phi}_{n_{F^{(l)}}}\right], \\ \text{for } k &= 2, \text{ and } 1 \leq k \leq n_{F^{(l)}} \end{aligned} \tag{59}$$

leading to the final definition of the transformation function

$$\widehat{F}^{(l)} = \Upsilon\left(F^{(l)}\right) = F^{(l)}\widehat{\mathbf{T}}_{k-KLT} \tag{60}$$

$$= F^{(l)}\left[\boldsymbol{\phi}_{n_{F^{(l)}}-1}, \boldsymbol{\phi}_{n_{F^{(l)}}}\right], \tag{61}$$

where the new $\widehat{F}^{(l)}$ will be of size $[n_{F^{(l)}} \times 2]$.

The final set of vectors to use is $F^{(1)}$, $F^{(2)}$, $F^{(3)}$, $F^{(4)}$, $\Upsilon\left(F^{(1)}\right)$, $\Upsilon\left(F^{(2)}\right)$, $\Upsilon\left(F^{(3)}\right)$, $\Upsilon\left(F^{(4)}\right)$, but from now will be addressed as $F^{(1)}$, $F^{(2)}$, $F^{(3)}$, $F^{(4)}$, $F^{(5)}$, $F^{(6)}$, $F^{(7)}$, $F^{(8)}$. In Figures 36-39 are shown scatter plots of some elements of the feature vectors aimed to emphasize the complex distribution and overlapping of the data.

4.5. Soft Computing Modeling Based on Neurofuzzy Classification

The term "neurofuzzy" refers to those systems that involve a combination (hybrid systems) of fuzzy logic theory (fuzzy sets, fuzzy probabilities, etc.) and neural networks (supervised/non–supervised neural networks, feedforward, recurrent, etc.). It has been demonstrated that a random variable with a Gaussian distribution is analog and equivalent to linguistic variable with a radial basis (RBF) membership functions. Therefore we will utilize the properties of a probabilistic neural network to show the combination of both a fuzzy system combined in an architecture of a feedforward neural network with a fixed number of layers.

4.5.1. Probabilistic Neural Network

The Probabilistic Neural Network (PNN) is a supervised neural network widely used in pattern recognition applications.[36] One of the main advantages is that it does not require training. The original PNN was proposed by Specht and its results are often compared with common backpropagation networks. Its results are shown to always converge to the Bayesian optimal solution.[37] Indeed, the PNN is inspired in the Bayesian classification an classical estimation for probability density functions. The basic operation of the PNN is to estimate the PDF's of the features assuming Gaussian distributions. Then a Bayesian-based decision rule is performed.

The general architecture of the PNN is composed of four layers, see Figure 40. The first layer is an input layer receiving the features $F^{(l)}$. The second layer contains exponential functions $\varphi(\cdot)$ in each nodes, and the number of nodes correspond to the k number of samples available for training for the $j-$th class. This nodes are called pattern units and are fully connected to the input nodes, and we refer to them as $\nu_{jk}^{F^{(l)}}$. The output of the pattern layer is denoted by

$$\varphi_{jk}\left(F^{(l)}\right) = \frac{1}{(2\pi)^{\frac{d}{2}}\,\sigma^d} e^{-\frac{1}{2\sigma^2}\left(F^{(l)}-\nu_{jk}^{F^{(l)}}\right)^T\left(F^{(l)}-\nu_{jk}^{F^{(l)}}\right)}. \tag{62}$$

The third layer contains summation units needed to complete the probability estimation. There are as many summation units as classes (size of j). The $j-$th summation unit denoted as ϖ_j, receives input only from those pattern units belonging to the $j-$th class. This layer indeed computes the maximum likelihood of $F^{(l)}$ being classified into C_j, by averaging and summarizing the output of all neurons that belong to the same class as

$$\Omega_j\left(\varphi_{jk}\left(F^{(l)}\right)\right) = \frac{1}{(2\pi)^{\frac{d}{2}}\,\sigma^d}\frac{1}{N_j}\sum_{i=1}^{N_j} e^{-\frac{1}{2\sigma^2}\left(\varphi_{ik}\left(F^{(l)}\right)-\varpi_i\right)^T\left(\varphi_{ik}\left(F^{(l)}\right)-\varpi_i\right)}. \tag{63}$$

The last layer is the decision layer. It classifies the pattern of $F^{(l)}$ according to the Bayesian decision rule given by

$$F^{(l)} \in C_j \text{ if,} \quad C_j\left(\Omega_j\left(\varphi_{jk}\left(F^{(l)}\right)\right)\right) = \max_{1\le i\le j}\Omega_i\left(\varphi_{ik}\left(F^{(l)}\right)\right). \tag{64}$$

Thus, the maximum of the summation node outputs can be expressed as a function of $C_j(\cdot)$ characterizing the output of this layer.

Estimation of the Spread Parameter σ. In the estimation of the spread parameter σ we follow the method proposed by Srinivasan *et al.*,[37] which requires a phase of pre-normalization of the data consisting on subtracting the mean $\boldsymbol{\mu}_{F^{(l)}}$ from the training feature vector $F^{(l)}$, and also dividing it by its standard deviation $\boldsymbol{\sigma}_{F^{(l)}}$. This is formally defined as

$$\widehat{F}^{(l)} = \frac{F^{(l)} - \boldsymbol{\mu}_{F^{(l)}}}{\boldsymbol{\sigma}_{F^{(l)}}}, \tag{65}$$

and then we can follow 3 steps to find an appropriate value for σ. First, we estimate the variance for all the features available for each class using $\widehat{F}^{(l)}$. Second, find the absolute difference between the smallest two variances obtained in the previous step. Third, set the value representing the differences obtained in the second step into σ.

4.5.2. Selection of Features and Events

The features selected for modeling the Dust Storm using PNN's are the first four feature sets: $F^{(1)}$, ..., $F^{(4)}$, as well as the reduced features $F^{(5)}$, ..., $F^{(8)}$. And the events were the same as in previous classification methods.

5. EXPERIMENTS AND DISCUSSION

In this section we explain the training and testing methodology as well as the performance criteria. We also present the results based on the performance and show visual results of the classification.

5.1. Training Methodology

As established in the models for classification of Dust Storms, we have millions of data points[†] available for the design of the classification methods. And precisely this an issue in the modeling of the PNN and FFNN. This is because the PNN creates a pattern layer containing a number of neuron units which is directly proportional to the number of training samples times the number of classes, which is an enormous quantity of data to be processed. While in the FFNN, since the backpropagation training methods consider all the training data to fit a multidimensional curve (or function) that separates the data into the different classes, it needs to propagate the error across the layers of the network by updating the weights and biases. This number of times this operation is performed is directly proportional to the number of neurons interconnected times the number of samples for training times the number of adaptations required to meet the stop/continue condition.

For this reason we decided to limit the number of samples selected for training in the case of FFNN and PNN. We based our reduction method on the criteria that establishes that the number of samples required for training the networks must be at least 3 times the number of bands used as features.[14] Therefore, in the PNN design, we decided to use at least 3 times the size of the feature vector $F^{(l)}$, for $l = 1...4$, and 6 times the size $F^{(l)}$, for $l = 5...8$. In the case of the FFNN, we utilized 500 times the size of $F^{(l)}$, for $l = 1...4$, and 1000 times its size for $l = 5...8$. We define this selection as follows

$$\widetilde{F}^{(l)} = \begin{bmatrix} F_{k_1}^{(1,l)} & \cdots & F_{k_1}^{(m,l)} \\ F_{k_2}^{(1,l)} & \cdots & F_{k_2}^{(m,l)} \\ \vdots & \ddots & \vdots \\ F_{k_i}^{(1,l)} & \cdots & F_{k_i}^{(m,l)} \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{k} &= \left[\text{ uniform}\,(1,r)_1\,, \quad \text{uniform}\,(1,r)_2\,, \quad \cdots \quad \text{uniform}\,(1,r)_i \ \right], \\ i &= m \times q, \end{aligned}$$

and m is number of columns in $F^{(l)}$, uniform(a, b) denotes a random number between a and b, r is the total number of samples available for training (number of rows in $F^{(l)}$), i is the total number of samples to consider in training, and q is the desired proportion to preserve. In our case q vary between $q = 3, 6, 500, 1000$.

Considering this, we can show the final selection of training samples in each Dust Storm classification model. Table 4 shows the features and samples utilized to design (or train) the classification methods, as well as the required preprocessing of the data (PNN case).

A Note in the Training of the FFNN. The neuron units in the $i-$th layer of the FFNN have hyperbolic tangent sigmoid transfer functions (also known as *tangsig*), and at the output layer, the neuron unit has a linear transfer function (*purelin*). The backpropagation method used to update the weights and biases is the Levenberg-Marquardt optimization method (*trainlm*). Also as a learning function we used the gradient descent with momentum weight and bias learning function (*learngdm*). The stop conditions for the FFNN are either: 1) 100 epochs, 2) Performance=0, 3)Validation failures=5, and 4)Minimum performance gradient=1×10^{-10}. The performance metric is the mean squared error (MSE). An internal set of training testing and validation was randomly selected to internally evaluate the generalization ability of the network.

[†]Here a data point is a single element of a feature vector $F^{(l)}$. For instance, if the feature vector

$$F^{(l)} = \begin{bmatrix} n_{1,1} & n_{1,2} & \cdots & n_{1,r} \\ n_{2,1} & n_{2,2} & \cdots & n_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ n_{k,1} & n_{k,2} & \cdots & n_{k,r} \end{bmatrix},$$

then $\mathbf{n} = [n_{1,1}, n_{1,2}, ..., n_{1,r}]$, is considered a data point.

5.2. Testing Methodology

For testing purposes we select all the features for all the methods instead of a reduced subset of the available data. Table 5 shows this configuration and introduces an optional classification set. That is to classify to two classes instead of four. This is performed with no modification of the architecture or models but rather in a mapping post-classification. Such mapping can be defined as a function of the classification output as follows

$$F^{(l)} \in C0 \text{ if } \begin{cases} F^{(l)} \in C0, & \text{or} \\ F^{(l)} \in SM, & \text{and} \end{cases} \tag{66}$$

$$F^{(l)} \in DU \text{ if } \begin{cases} F^{(l)} \in BLDU, & \text{or} \\ F^{(l)} \in DS, & \end{cases} \tag{67}$$

which will allow the system to classify only between background $C0$, and a new class name dust DU.

5.3. Performance Metric

Given the models for classification of dust storms, and provided an input to them, there are four possible outcomes. Suppose we know the input is dust storm DS and it is classified as DS, it is counted as a true positive (TP); if it is classified as maybe $C0$, it is counted as a false negative (FN). If the input is not DS and it is not classified as DS, it is counted as a true negative (TN); if it is classified as DS, it is counted as a false positive (FP). Given a classifier, the number of j classes, and a test set, a j-by-j confusion matrix (also called a contingency table) can be constructed representing the dispositions of the test set. The numbers along the major diagonal represent the correct decisions made, and the numbers off this diagonal represent the errors (the confusion) between the various classes. This matrix forms the basis for many common metrics, such as

$$\begin{aligned}
\text{fp rate} &\approx \frac{\sum FP}{\sum FP + TN}, \\
\text{tp rate} &\approx \frac{\sum TP}{\sum TP + FN}, \\
\text{precision} &= \frac{\sum TP}{\sum TP + FP}, \\
\text{accuracy} &= \frac{\sum TP + TN}{\sum TP + FN + FP + TN}, \\
\text{recall} &= \frac{\sum TP}{\sum TP + FN}, \\
F - \text{measure} &= \frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}}, \\
\text{sensitivity} &= \text{recall} \\
\text{specificity} &= \frac{\sum TN}{\sum FP + TN}, \\
&= 1 - \text{fp rate}, \\
\text{Positive Predictive Value} &= \text{precision}.
\end{aligned}$$

Receiver Operating Characteristics. The Receiver Operating Characteristics[38] (ROC) graphs are two-dimensional graphs in which TP rate is plotted on the y axis and FP rate is plotted on the x axis. A ROC graph shows the relative trade-offs between benefits (true positives) and costs (false positives). The classification

methods we have modeled, will be treated as discrete classifier that outputs only a class label. Each discrete classifier produces an (fp rate, tp rate) pair corresponding to a single point in ROC space. Several aspects in the ROC space are important to note. The lower left point (0; 0) represents the strategy of never issuing a positive classification; such a classifier commits no false positive errors but also gains no true positives. The opposite strategy, of unconditionally issuing positive classifications, is represented by the upper right point (1; 1). The point (0; 1) represents perfect classification. Informally, one point in ROC space is better than another if it is to the northwest (tp rate is higher, fp rate is lower, or both) of the first. Classifiers appearing on the left hand-side of an ROC graph, near the x axis, may be thought of as "conservative": they make positive classifications only with strong evidence so they make few false positive errors, but they often have low true positive rates as well. Classifiers on the upper right-hand side of an ROC graph may be thought of as "liberal": they make positive classifications with weak evidence so they classify nearly all positives correctly, but they often have high false positive rates. The diagonal line $y = x$ represents the strategy of randomly guessing a class.

To compare classifiers we may want to reduce ROC performance to a single scalar value representing expected performance. A common method is to calculate the area under the ROC curve, abbreviated AUC. Since the AUC is a portion of the area of the unit square, its value will always be between 0 and 1.0. However, because random guessing produces the diagonal line between (0; 0) and (1; 1), which has an area of 0.5, no realistic classifier should have an AUC less than 0.5. The AUC has an important statistical property: the AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance.

Although ROC curves may be used to evaluate classifiers, care should be taken when using them to make conclusions about classifier superiority. Some researchers have assumed that an ROC graph may be used to select the best classifiers simply by graphing them in ROC space and seeing which ones dominate. This is misleading; it is analogous to taking the maximum of a set of accuracy figures from a single test set. Without a measure of variance we cannot compare the classifiers. Averaging ROC curves is easy if the original instances are available. Therefore, in our models we have computed the AUC in each granule of information, $T1; T2; ...; Tn$, which are generated from independent events. Then, we can simply merge sort the instances together by their assigned scores into one large test set TM. We then plot a ROC with the result. However, the primary reason for using multiple test sets is to derive a measure of variance, which in our case will be estimated from all the test set.

5.4. Results

5.4.1. Testing Models on MODIS Level 1B Data: Quantitative Results

The numerical results are totalized and averaged to produce Table 6, where it is clear that the soft computing methods have the best performance metrics. However, as pointed out previously, the best performance metric is the AUC, thus, in general, the FFNN reports the best performance. Indeed, if the problem is simplified to the classification of two classes, the FFNN still reports the highest AUC.

Then, the ROC's were computed over test sets, and the average results along with the standard deviation are shown in Figures 41-48. From this plots it is clear how the soft computing models FFNN and PNN perform better than the probabilistic-based models. The same conclusions can be derived when the problem is simplified to a problem with two classes, and the corresponding ROC's are shown in Figures 49-56.

From the overall ROC's we have estimated the total averaged AUC, which is considered to be the most accurate metric of performance. If the AUC value is high it is better. In Figure 57 and Table 8, the results show that the highest area under the ROC is reported by the FFNN for the feature vector $F^{(1)}$. The result is given when only two classes are considered, see also Figure 58 and Table 9.

To finalize with the quantitative analysis of performance, we introduce the data in Table 10, Table 11, and Table 12. This tables show the average proccessing time required by the model in order to produce an output given an image, given a pixel, and given a satellite scan respectively. From this data we can observe that the probabilistic methods are faster in comparison to the soft computing models as shown in Figure 59. This property makes them suitable for near real time applications.

5.4.2. Qualitative Visual Analysis

The results of our algorithms are displayed for visual assessment of the outputs. Here we present five different kinds of figures, the first is a true color image re-projected using the traditional Mercator approach. Follows a second graphic that shows the false color image of the same event, then, a third graphic show the probability of presence of dust storm with a custom made color-scale. The fourth graphic is a low-resolution (decimated) version of the third, and it is smoothed and displayed with a different color-scale (note that in spite they look different, they are derived from the same high resolution result). And the final graphic is a segmentation image, showing the actual mapping of the result: Red maps to DS, Green maps to $BLDU$, Blue maps to SM, and Black is mapped to background ($C0$). In the case of the Probabilistic model designed for visual assessment of probability only (Section 4.1) the latter kind of figures is not available.

From Figures 60-490 are shown all the results organized by event class, then by model, then by kind of visualization. In the Appendix B, we include true color images of some events which contain annotations indicating some of the dust sources.

6. CONCLUSION

The problem of dust storm detection has been addressed in this paper. First, we constructed a database of events from satellite observations of MODIS Terra satellite. This database was used to model the events trough a phase of selection of regions of interest. Then, we have modeled probabilistic approaches for dust storm detection and classification, and the parameters were estimated from selected samples in the database. Indeed, these models are specialized on measuring the probability of the presence of dust storm data given MODIS Level 1B data.

Novel techniques in Soft Computing were utilized to design neural architectures to model dust storms. To the best of the knowledge of the authors, the presented models are the first in its kind that can actually perform classification of dust storms pixels based on soft computing methods. We compared the probabilistic models such as Maximum Likelihood, MLE, and Maximum *a posteriori*, MAP, against the soft computing models such as Feedforward Backpropagation Neural Network, FFNN, and Probabilistic Neural Network, PNN, having that the latter report a strong ability in inferring the relationship between spectral bands to classify dust, and discriminate from other signatures, such as clouds, smoke, etc.

Moreover, the proposed probabilistic models are suitable for near real-time applications, such as direct broadcast, rapid response analysis, emergency alerts, etc. The probabilistic models are suitable for fast prototyping due to they simplicity, besides, the theory behind is easy to understand.

The work reported in this document is suitable for the study of the dust storm problem since the algorithms can show the dust presence to a resolution of 1km, which is an improvement over the methods based on the Aerosol Optical Thickness index (AOT) which lack of resolution. Besides, the AOT MODIS product is generated after two days of the satellite pass, increasing the response time in the analysis and study of the dust storms.

6.1. Future Work

The work reported in this document is part of an initiative from UTEP dedicated to develop better tools for the analysis of the dust storms. In the next stage of this work, we will detect the sources, orientation, and severity of the dust storms, aided with the output of the models presented in this paper.

An improvement in the process of classification on the neural networks, particularly in the PNN, is required. For this case we will use the tool named Recursive Hierarchical Segmentation (RHSEG) designed and developed by the author (J.C.T.). This tool provides multiple segmented regions based on they similarity, and outputs their statistical data (mean and standard deviation). Therefore, the output of RHSEG can be used to match the best n candidates closer to the values of $\mu_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ and $\sigma^2_{g\left(X_{\mathbf{n}}^{(Bm)}\right)}$ modeled in Section 4.1.1. This will decrease dramatically the processing time of the PNN classification up to a 95% less, as measured in preliminary laboratory tests.

A near real time application follows this work. The MLE model will be implemented in the Simulcast Viewer (client) application provided by NASA's Direct Readout Laboratory (DRL). The aim of this implementation is to test the ability to perform in near real time, as well as its application in rapid response systems.

ACKNOWLEDGMENTS

The authors want to thank the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC), the University of Maryland (UMBC) Graduate Student Summer Program (GSSP), and the Goddard Earth Science and Technology center (GEST), for supporting this work, and providing the necessary tools for its completeness. The author (P.R.P.) also wants to remark the mentoring work of Dr. James C. Tilton from NASA, and also thank the National Council of Science and Technology of Mexico, CONACYT, for partially support this research under grant 193324 / 303732.

The author (P.R.P.) also wants to acknowledge the insights of other scientists and engineers who collaborated in any way to this project: Patrick (Pat) Coronado, NASA's GSFC DRL and Computational and Information Sciences and Technology Office (CISTO). Jeffrey (Jeff) E. Schmaltz, (Sigma Space Corporation) MODIS Rapid Response Team NASA GSFC. Dr. Thomas Gill, Geological Sciences, Environmental Science, The University of Texas El Paso.

REFERENCES

1. Rivera Rivera, N.I., Gebhart, K.A., Gill, T.E., Hand, J.L., Novlan, D.J., and Fitzgerald, R.M., 2009. Analysis of air transport patterns bringing dust storms to El Paso, Texas. Preprints of the Symposium on Urban High Impact Weather, American Meteorological Society Annual Meeting, January 2009, Phoenix., JP2.6.
2. Loomer, S.A., "Academic research opportunities at the National Geospatial-Intelligence Agency (NGA)", Proceedings of SPIE, vol.6233, pp. 623324, 2006.
3. Gail, W.B., "Remote sensing in the coming decade: the vision and the reality", Journal of Applied Remote Sensing, vol.1, pp. 012505, 2007.
4. Chen, J. and Lin, M. and Lee, H., "Mineral dust emission estimated from satellite data", Proc. SPIE, vol. 6299, pp. 62990, 2006.
5. El-Askary, H.; Kafatos, M.; Xue Liu; El-Ghazawi, T., "Introducing new approaches for dust storms detection using remote sensing technology," Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International , vol.4, no., pp. 2439-2441 vol.4, 21-25 July 2003.
6. Anna-Britt Mahler, Kurt Thome, Dazhong Yin, and William A. Sprigg, "Dust transport model validation using satellite- and ground-based methods in the southwestern United States", Proc. SPIE 6299, 62990L (2006), DOI:10.1117/12.679868
7. Prospero, J.M., P. Ginoux, O.Torres, S.E. Nicholson, and T.E. Gill, "Environmental characterization of global sources of atmospheric soil dust identified with the Nimbus 7 Total Ozone Mapping Spectrometer (TOMS) absorbing aerosol product", Rev.Geophys., 40 (1), art. no. 1002, 2002.
8. Khazenie, N.; Lee, T.F., "Identification Of Aerosol Features Such As Smoke And Dust, In NOAA-AVHRR Data Using Spatial Textures," Geoscience and Remote Sensing Symposium, 1992. IGARSS '92. International , vol., no., pp.726-730, 26-29 May 1992.
9. Ackerman, S. A. (1997), Remote sensing aerosols using satellite infrared observations, J. Geophys. Res., 102(D14), 17,069–17,079.
10. Nancy I. Rivera Rivera, "Detection and characterization of dust source areas in the Chihuahuan desert, Southwestern North America" Masters Thesis, University of Texas at El Paso, December 2006.
11. Lee, J.A., Gill, T.E., Mulligan, K.R., Dominguez, M.A., and Perez, A.E., 2009. Land use/ land cover and point sources of the 15 December 2003 dust storm in southwestern North America. Geomorphology 105: 18-27.
12. Swapna Janugani, "Directional analysis and filtering for dust storm detection in NOAA-AVHRR imagery" Masters Thesis, University of Texas at El Paso, January 2009.
13. J. L. Kast, P. H. Swain and B. J. Davis, ECHO User's Multiband Satellite Photography Using Digital Tech Guide, LARS Publication 083077, Purdue University.
14. Kanellopoulos, G. G. Wilkinson, J. Austin, Neurocomputation in Remote Sensing Data. Analysis, Springer-Verlag, 1997.

15. B. Tso and P. M. Mather, Classification Methods for Remotely Sensed Data. London and New York: Taylor & Francis New York, 2001.
16. Hillger, D.W., and G.P. Ellrod, 2003: Detection of important atmospheric and surface features by employing Principal Component Image transformation of GOES imagery, J. Appl. Meteor., 42:5, 611-629.
17. Agarwal, A.; El-Askary, H.M.; El-Ghazawi, T.; Kafatos, M.; Le-Moigne, J., "Hierarchical PCA Techniques for Fusing Spatial and Spectral Observations With Application to MISR and Monitoring Dust Storms," Geoscience and Remote Sensing Letters, IEEE , vol.4, no.4, pp.678-682, Oct. 2007.
18. Internally maintained database of dust events. National Weather Service in Santa Teresa, New Mexico.
19. Xianjun Hao, John J. Qu, 2007, Saharan dust storm detection using MODIS thermal infrared bands. Journal of Applied Remote Sensing,Vol. 1, 013510.
20. S. D. Miller, "A consolidated technique for enhancing desert dust storms with. MODIS," Geophys. Res. Lett. 32(20), 2071, 2003.
21. Liu, San-Chao; Liu, Qinhuo; Gao, Maofang; Chen, Liangfu. Detection of dust storms by using daytime and nighttime multi-spectral MODIS images. International Geoscience and Remote Sensing Symposium (IGARSS), 2006 IEEE International Geoscience and Remote Sensing Symposium, IGARSS, 2006, p 294-296.
22. Ackerman, S.A., Strabala, K.I., Menzel, W.P., Frey, R., Moeller, C., Gumley, L., Baum, B.A., Seeman, S.W., & Zhang, H. (2002). Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35). In ATBD Ref.ATBD-MOD-06, version 4.0, 115 pp. Greenbelt, Maryland, NASA Goddard Space Flight Center.
23. J. S. Lim, Two-Dimensional Signal and Image Processing. Englewood Cliffs, N.J.: Prentice Hall, 1990, pp. 694.
24. M. Miyahara, K. Kotani and V. R. Algazi, "Objective picture quality scale (PQS) for image coding," *Communications, IEEE Transactions on,* vol. 46, pp. 1215-1226, 1998.
25. R. Liu and T. Tan, *Theoretical framework for watermark capacity and energy estimation*, SPIE Proceedings Series 4675, Conference on Security and Watermarking of Multimedia Contents, pp.433-440. 2002.
26. Depenau, J., 1997, Neural networks for classification of ice type concentration from ERS-1 SAR images. In: Neurocomputation in Remote Sensing Data Analysis, edited by I. Kanellopoulos, G. Wilkinson, F. Roli and J. Austin (Berlin: Springer), pp. 71–78.
27. Gibson, P.J. and Keating, J. (2000) Interactive Introductory Remote Sensing: Principles and Concepts Website 2000.
28. Chen, Chi Hau and Peter Ho, Pei-Gee, "Statistical pattern recognition in remote sensing", Pattern Recogn., vol. 41, no 9, 2008, issn 0031-3203, pp 2731–2741, doi http://dx.doi.org/10.1016/j.patcog.2008.04.013, Elsevier Science Inc., New York, NY, USA.
29. Aksoy, S.; Koperski, K.; Tusk, C.; Marchisio, G.; Tilton, J.C., "Learning bayesian classifiers for scene classification with a visual grammar," Geoscience and Remote Sensing, IEEE Transactions on , vol.43, no.3, pp. 581-589, March 2005.
30. John A. Richards and Xiuping Jia, Remote Sensing Digital Image Analysis, Springer, 2006.
31. Borengasser, M., Hungate, W. S., "Hyperspectral Remote Sensing. Principles and Applications", CRC Press, 2008.
32. X.H. Yu, "Can backpropagation error surface not have local minima," Neural Networks, IEEE Transactions on , vol.3, no.6, pp.1019-1021, Nov 1992.
33. S. Haykin, "Neural Networks and Learning Machines," ISBN: 9780131471399. Prentice Hall (November 14, 2008).
34. Hugo Larochelle, Yoshua Bengio, Jérôme Louradour and Pascal Lamblin, "Exploring Strategies for Training Deep Neural Networks", Journal of Machine Learning Research, 10(Jan): 1–40, 2009.
35. M. I. Chacon M., P. Rivas P., "Fusion of Fuzzy FFL-KLT and PCNN Features on the Face Recognition Problem", Dynamics of Continuous, Discrete & Impulsive Systems Journal, Series A: Mathematical Analysis, a Special Issue on Advances in Neural Networks-Theory and Applications, Part I. Watan Press - Waterloo. August 2007.

36. Foody, G.M. Thematic mapping from remotely sensed data with neural networks: MLP, RBF and PNN based approaches, Journal of Geographical Systems 3: pp. 217-232. 2001.
37. Ramakrishnan, S. and Selvan, S. 2007. Image texture classification using wavelet based curve fitting and probabilistic neural network. Int. J. Imaging Syst. Technol. 17, 4 (Dec. 2007), 266-275. DOI= http://dx.doi.org/10.1002/ima.v17:4.
38. Tom Fawcett. ROC graphs: Notes and practical considerations for data mining researchers. Technical Report HPL-2003-4, HP Laboratories, Palo Alto, CA, January 2003. Updated March 2004.
39. Z. Xu, K. Huang, J. Zhu, I. King and M. R. Lyu, "A novel kernel-based maximum a posteriori classification method," Journal of Neural Networks, vol. In Press, Accepted Manuscript, pp. -, 2009.

APPENDIX A. TABLES AND FIGURES

Figure 1. Dust storms hitting El Paso, TX, urban areas on June 08 2004 arround 6pm. Below, is a satellite image of the el paso area a few hours before the storm.

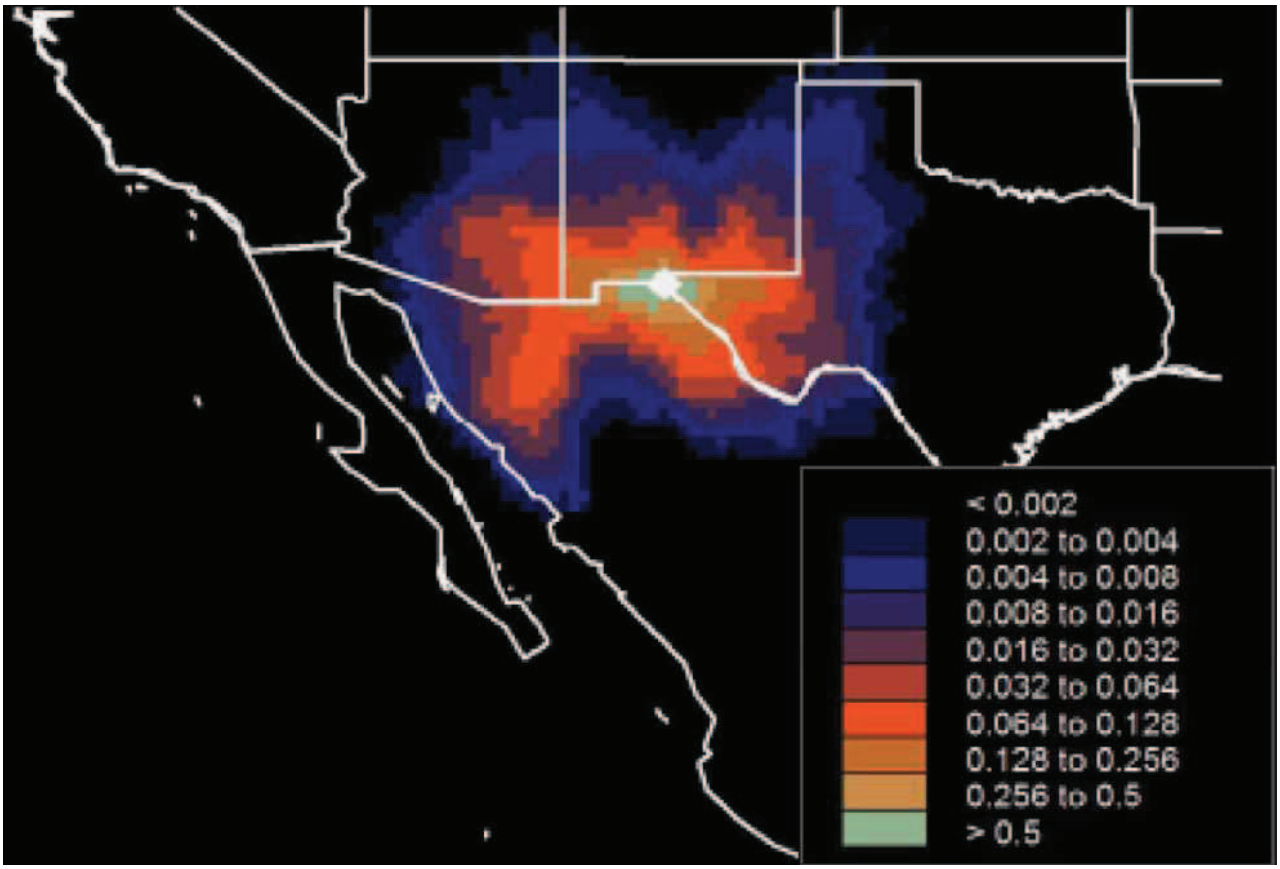

Figure 2. Observation of dust residence time on the southwestern US and north of mexico. The graphic shows the residence time probability for all months from 2001 to 2005. The residence time probability $f_{X_1X_2}(x_1 = i, x_2 = j)$ is defined as the number of back trajectory endpoints in a given grid cell (in this case, 0.25 deg latitude by 0.25 deg longitude) over a specified time interval, and it is formally defined as $f_{X_1X_2}(x_1 = i, x_2 = j) = \frac{1}{N}\sum_{t=1}^{T} x_{1t}x_{2t}$, where $x_{1t}x_{2t}$ is the number of endpoints falling in a grid cell at longitude $x_1 = i$ and latitude $x_2 = j$ before the trajectory arrived at the receptor during measurement period t, T is the total number of time periods and N is the total number of endpoints troughout T. More details in (Nancy *et al.*, 2009).[1]

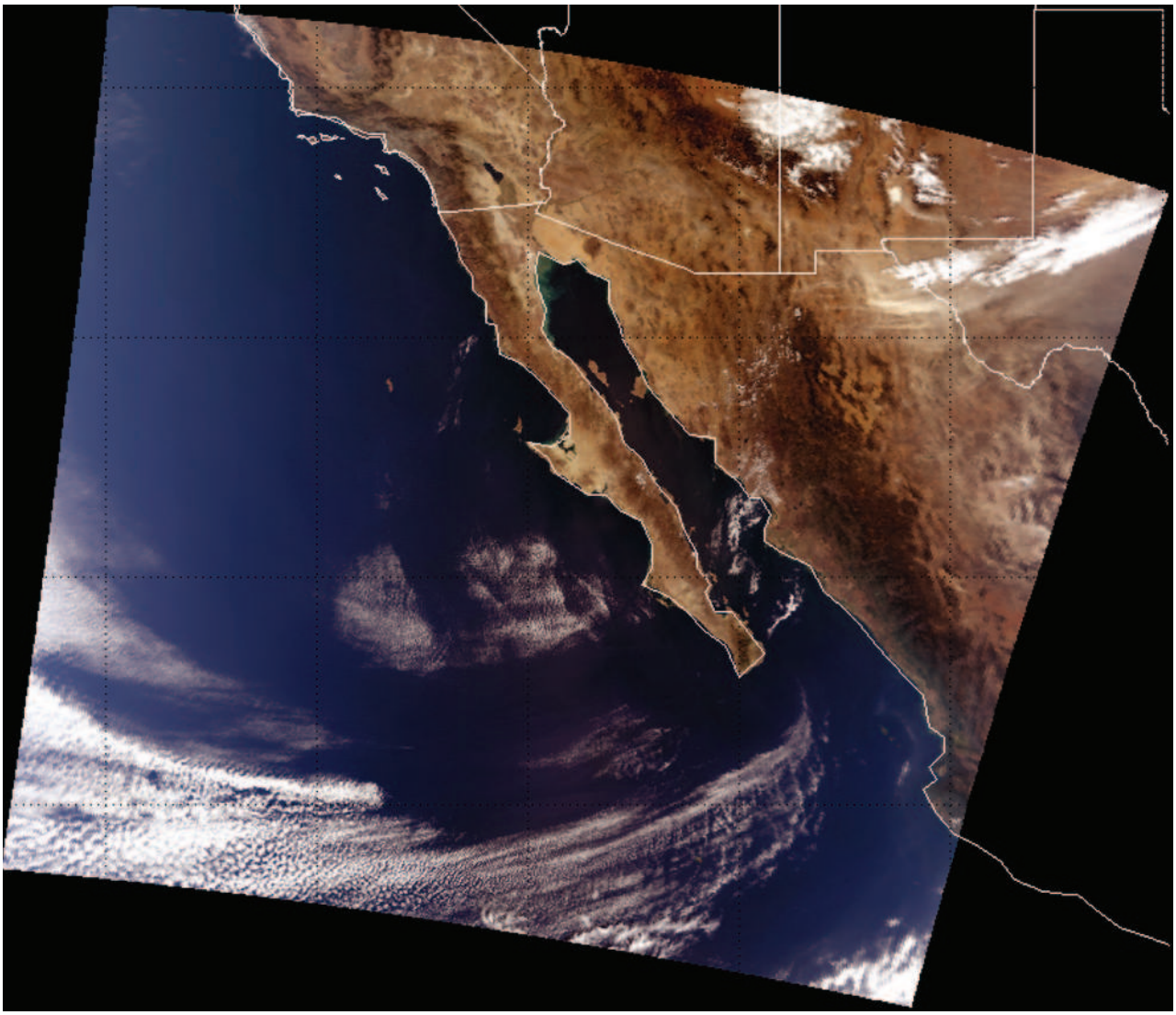

Figure 3. True color example from a random dust storm event. MODIS multiespectral bands are assigned to the RGB composite as follows: R=$B1$, G=$B4$, B=$B3$. In this example, the non-usable pixel elements were approximated to produce an adecuate picture.

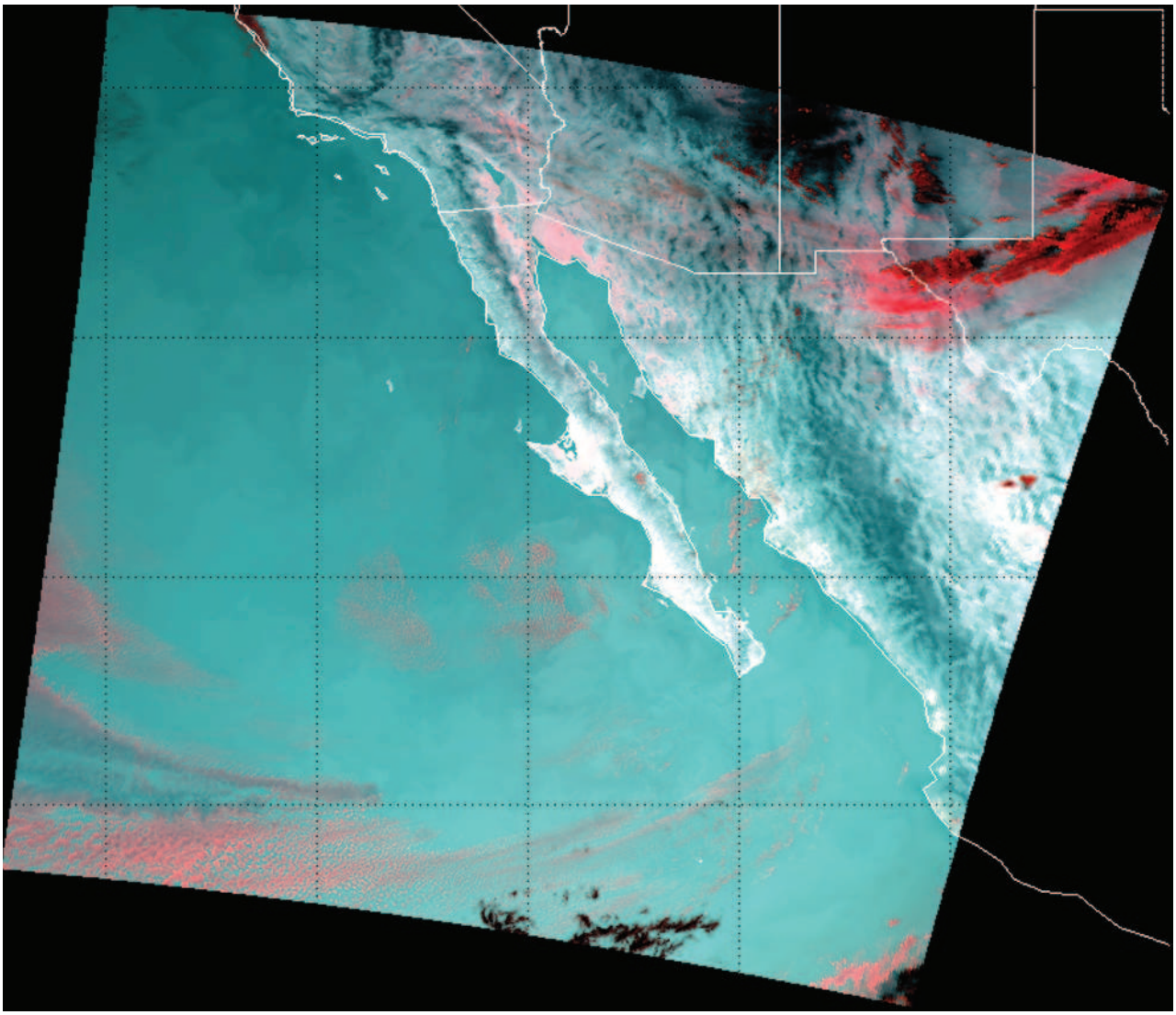

Figure 4. A false color image of the same event in Figure 3. The dust regions in this visualization look in pink color. In this case the false color is generated by mapping the RGB composite to MODIS multiespectral bands as follows: R=$B20$, G=$B31$, B=$B32$.

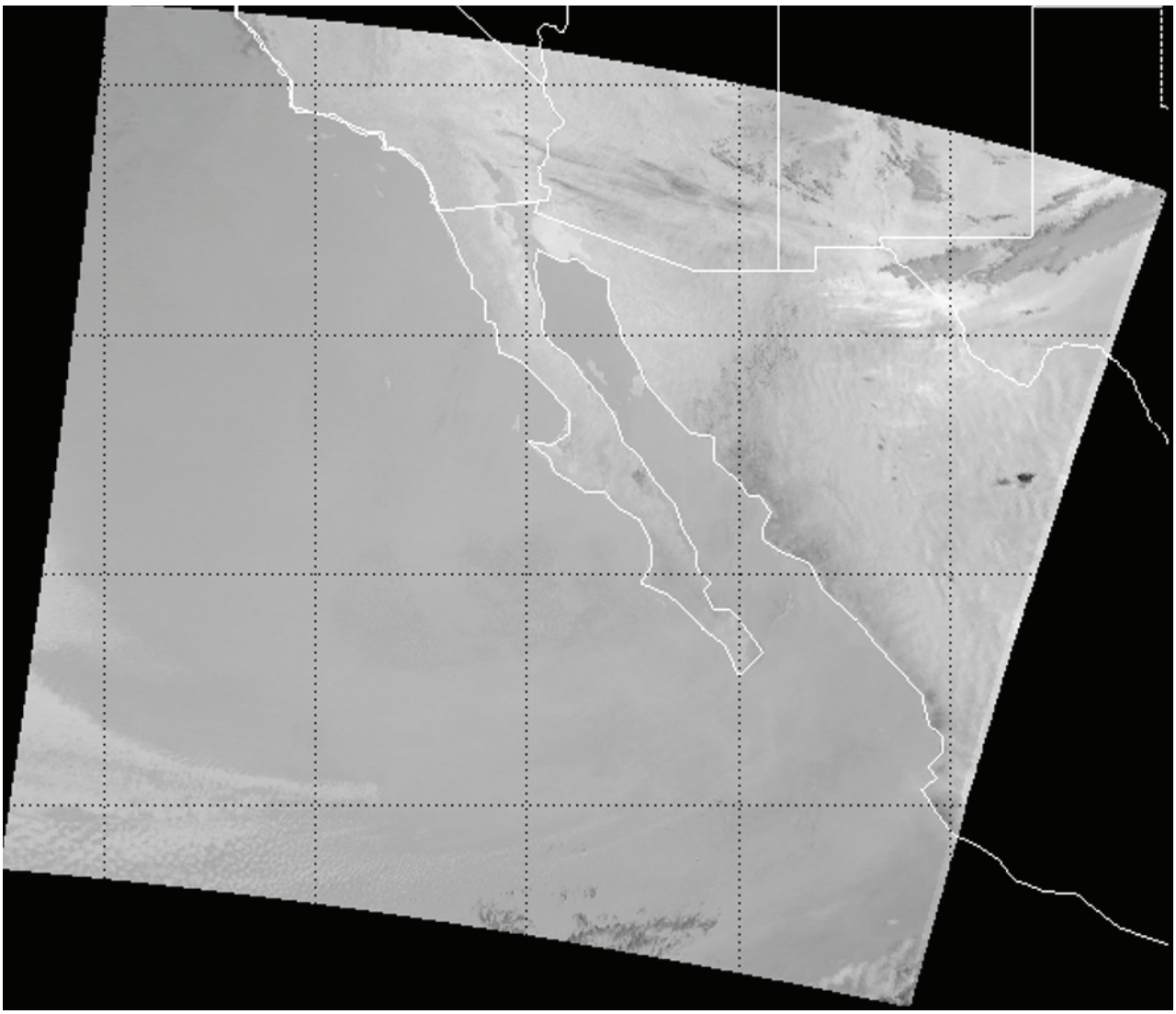

Figure 5. Example of a band subtraction grayscale result for the event in Figure 3. The bands subtracted are $B32 - B31$, the results are displayed in grayscale. The dust regions are shown in near white color.

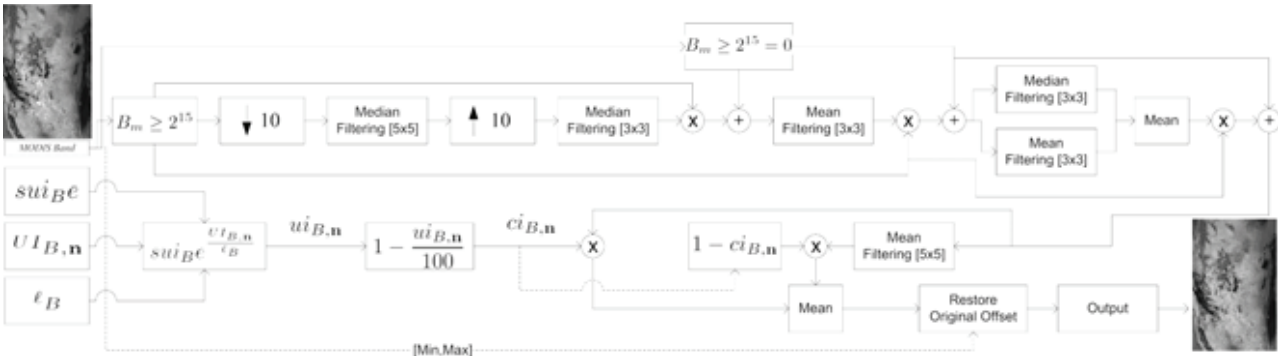

Figure 6. Process of estimation of true data based on unusable pixel data and uncertainty indexes.

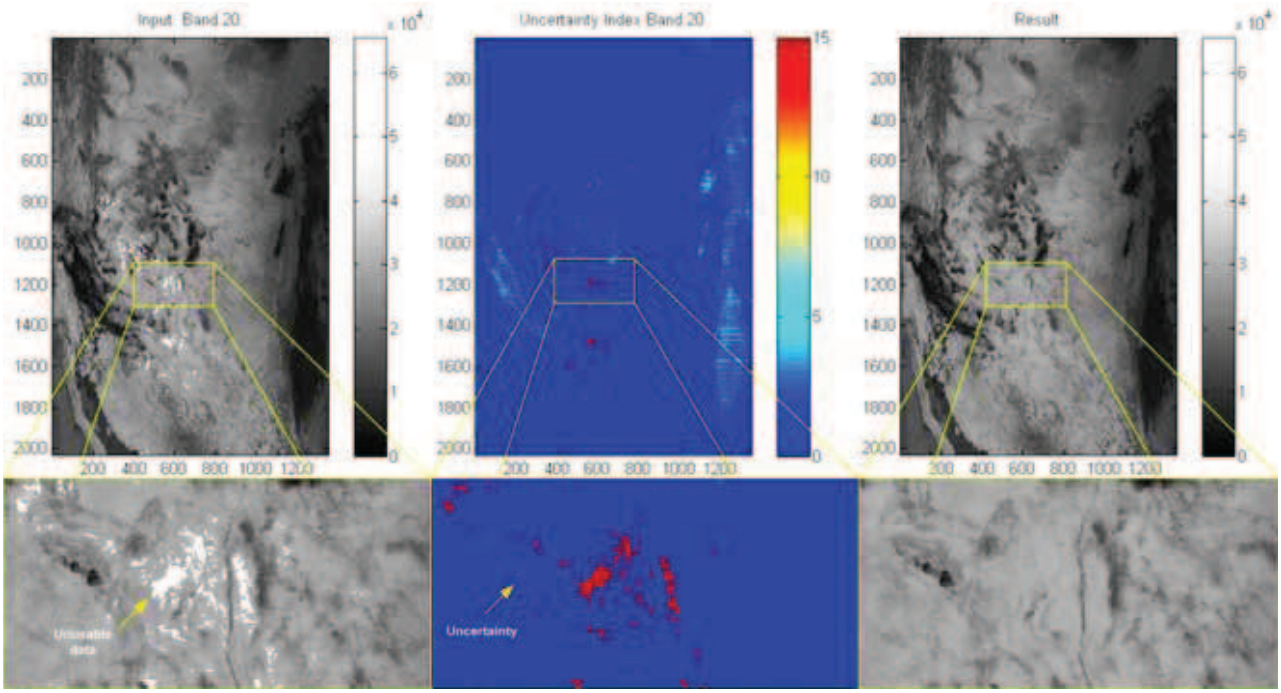

Figure 7. On the left, input data with high uncertainty and unusable data. On the middle, the given uncertainty index $UI_{B,\mathbf{n}}$. And in the right, its the result of the proposed method.

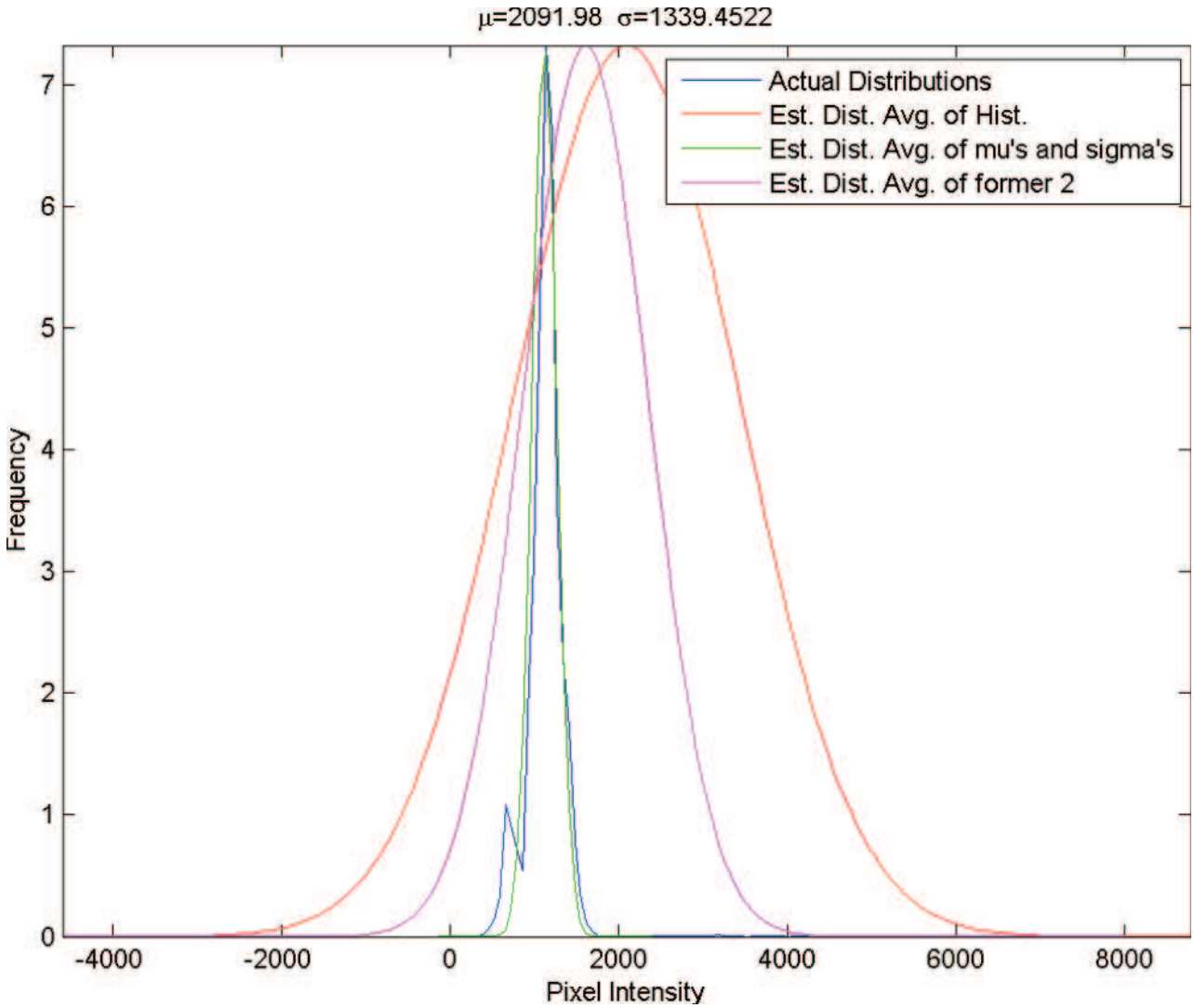

Figure 8. Total histogram in percent, and the estimated parameters using three methods. MODIS 16-bit image data for band subtraction $B32 - 31$.

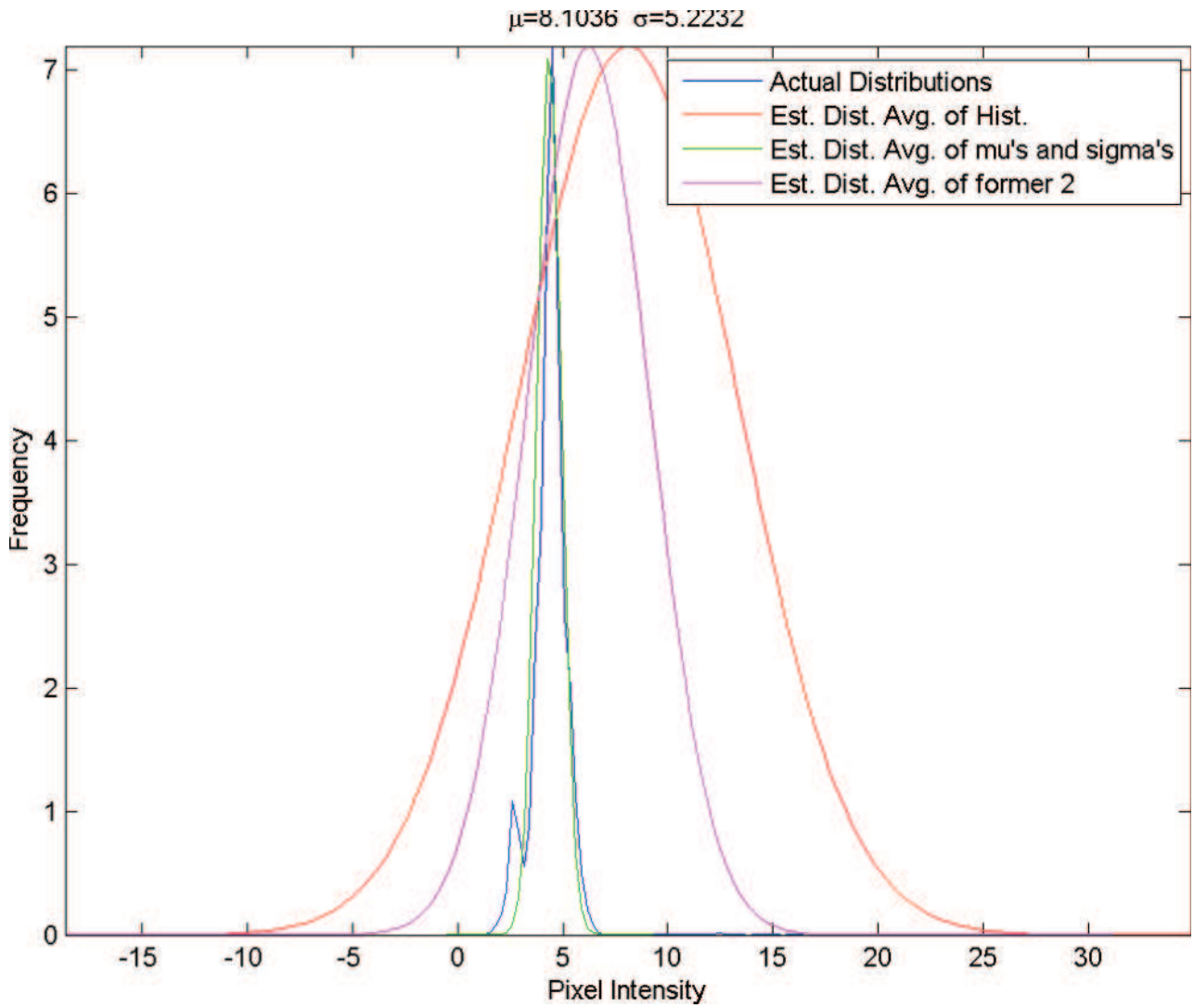

Figure 9. Total histogram in percent, and the estimated parameters using three parameter estimation methods. MODIS 8-bit image data for band subtraction $B32 - 31$.

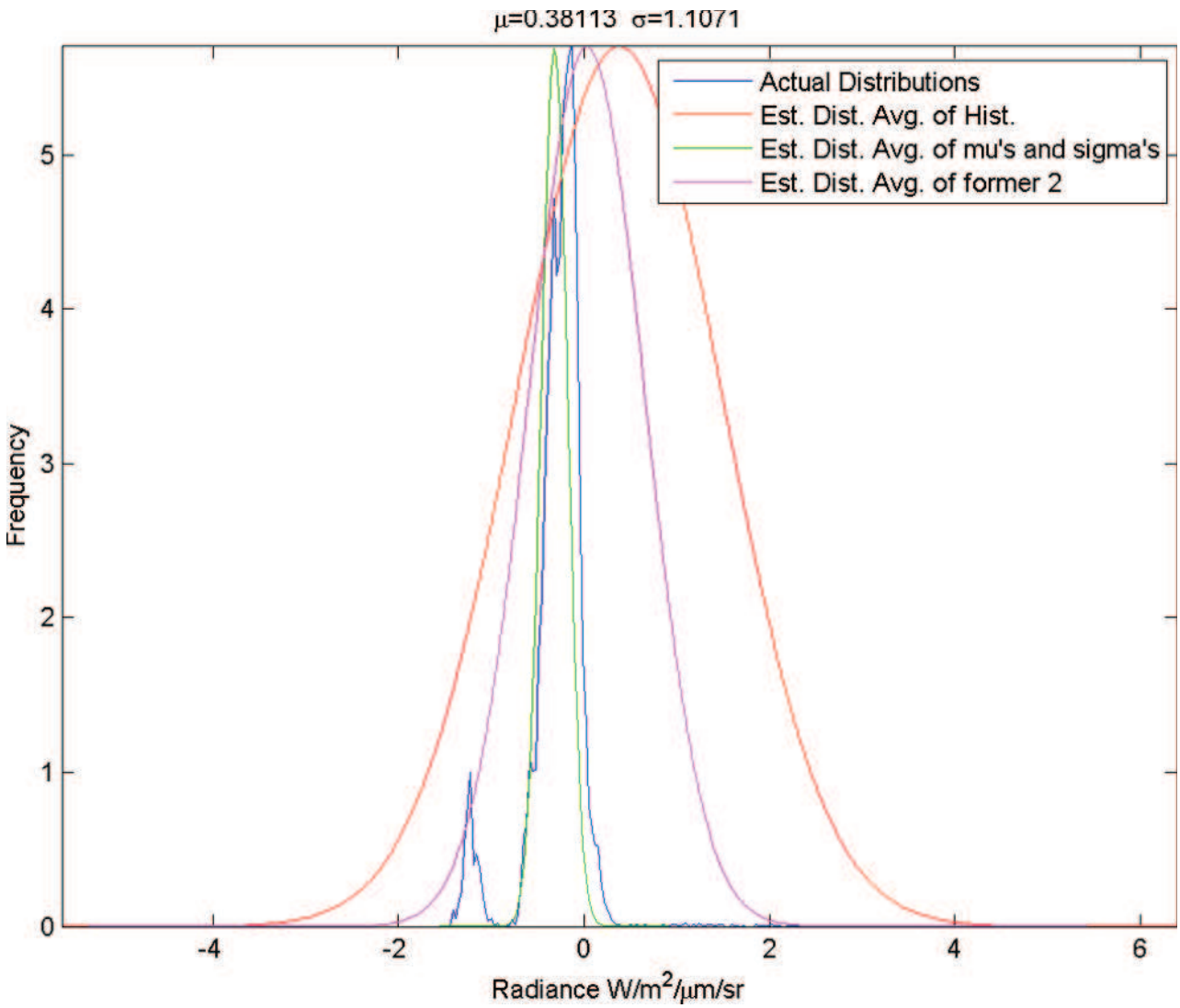

Figure 10. Total histogram in percent, using the three parameter estimation methods. These are MODIS recovered radiances from the band subtraction $B32 - 31$.

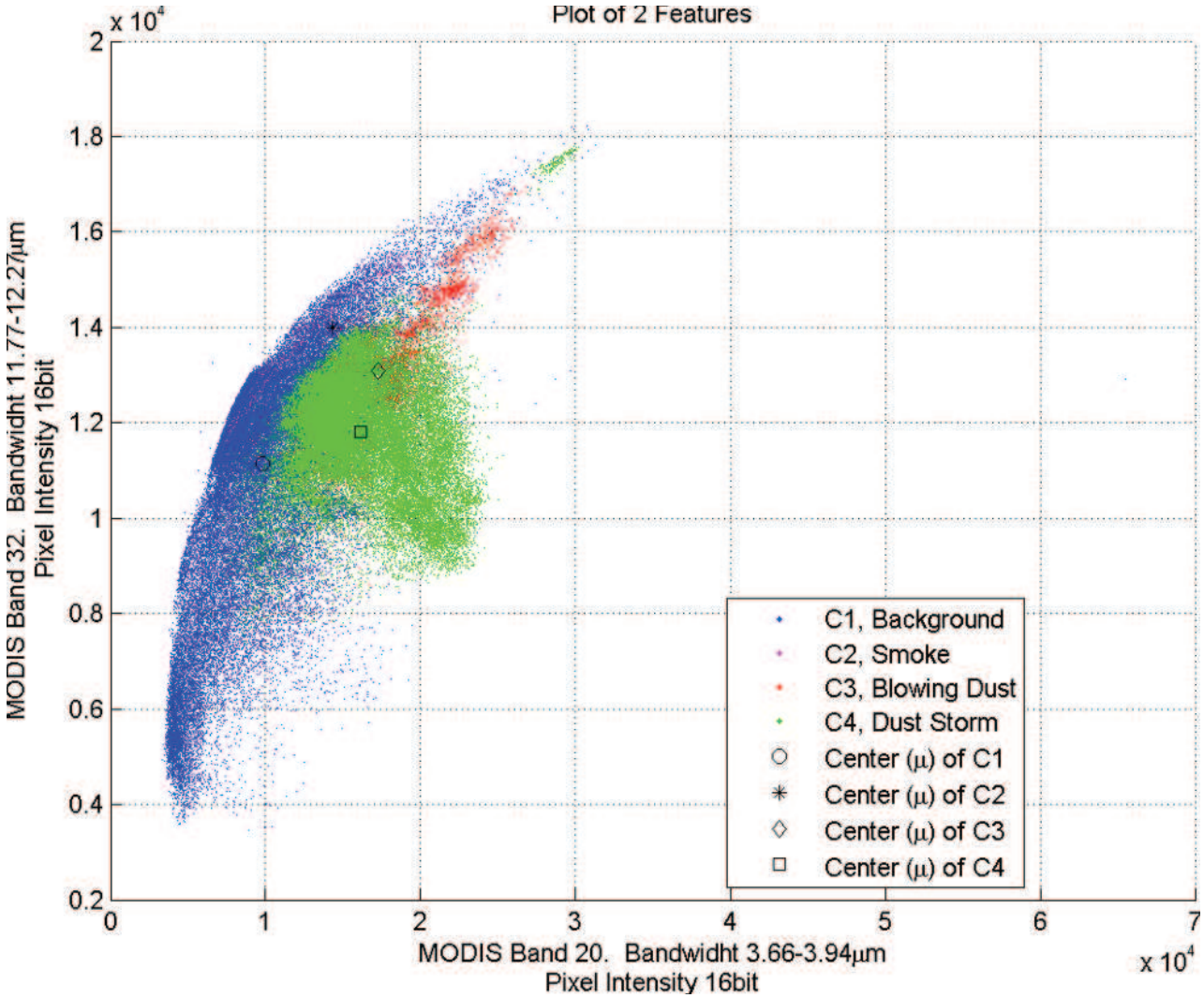

Figure 11. Distribution of feature vector $F^{(1)}$ when two of its components are ploted. In this case the x-axis correspond to $B20$, while the y-axis correspond to $B32$, raw intensities.

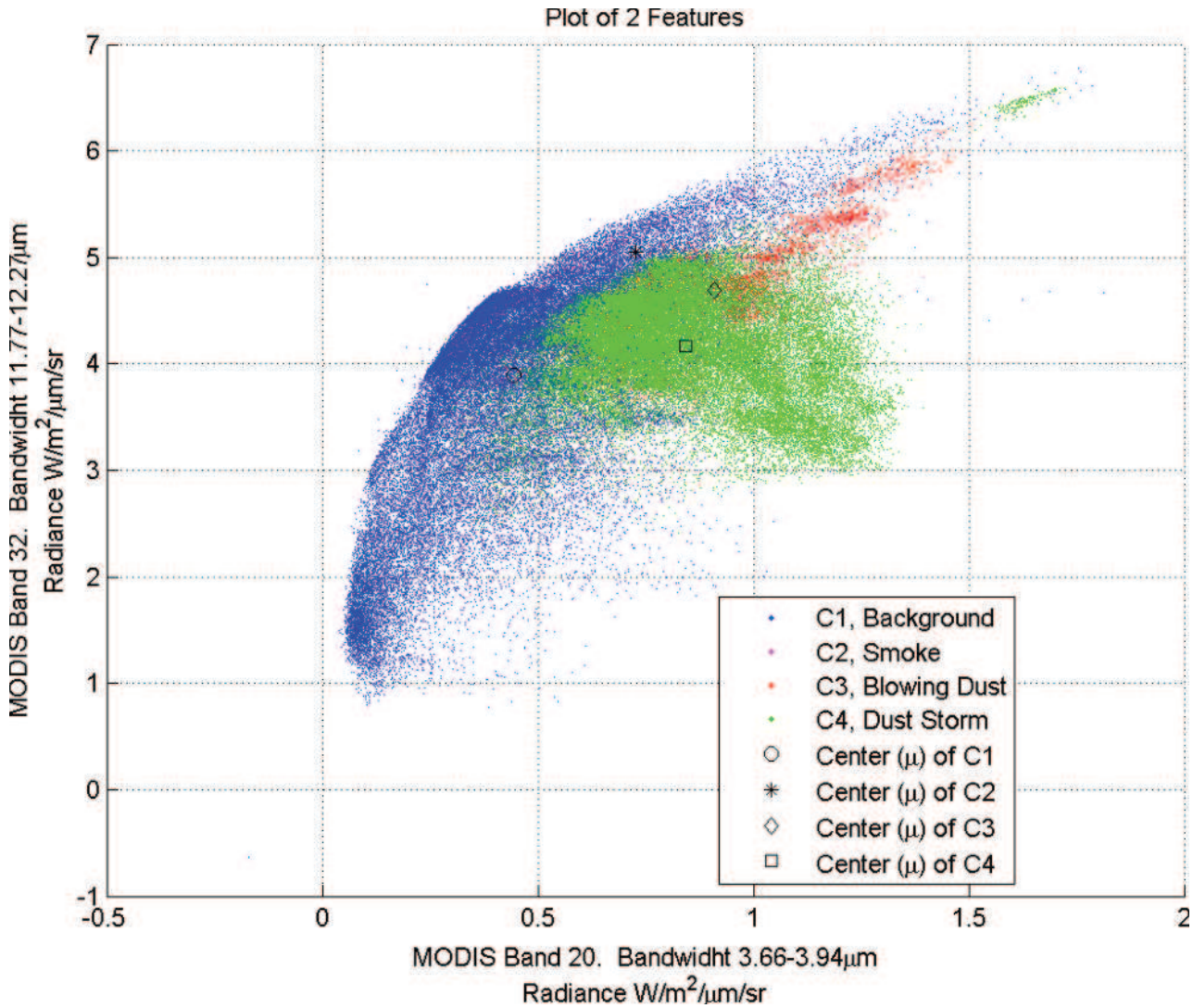

Figure 12. Distribution of feature vector $F^{(2)}$ when two of its components are ploted. In this case the x-axis correspond to $B20$, while the y-axis correspond to $B32$, recovered radiances.

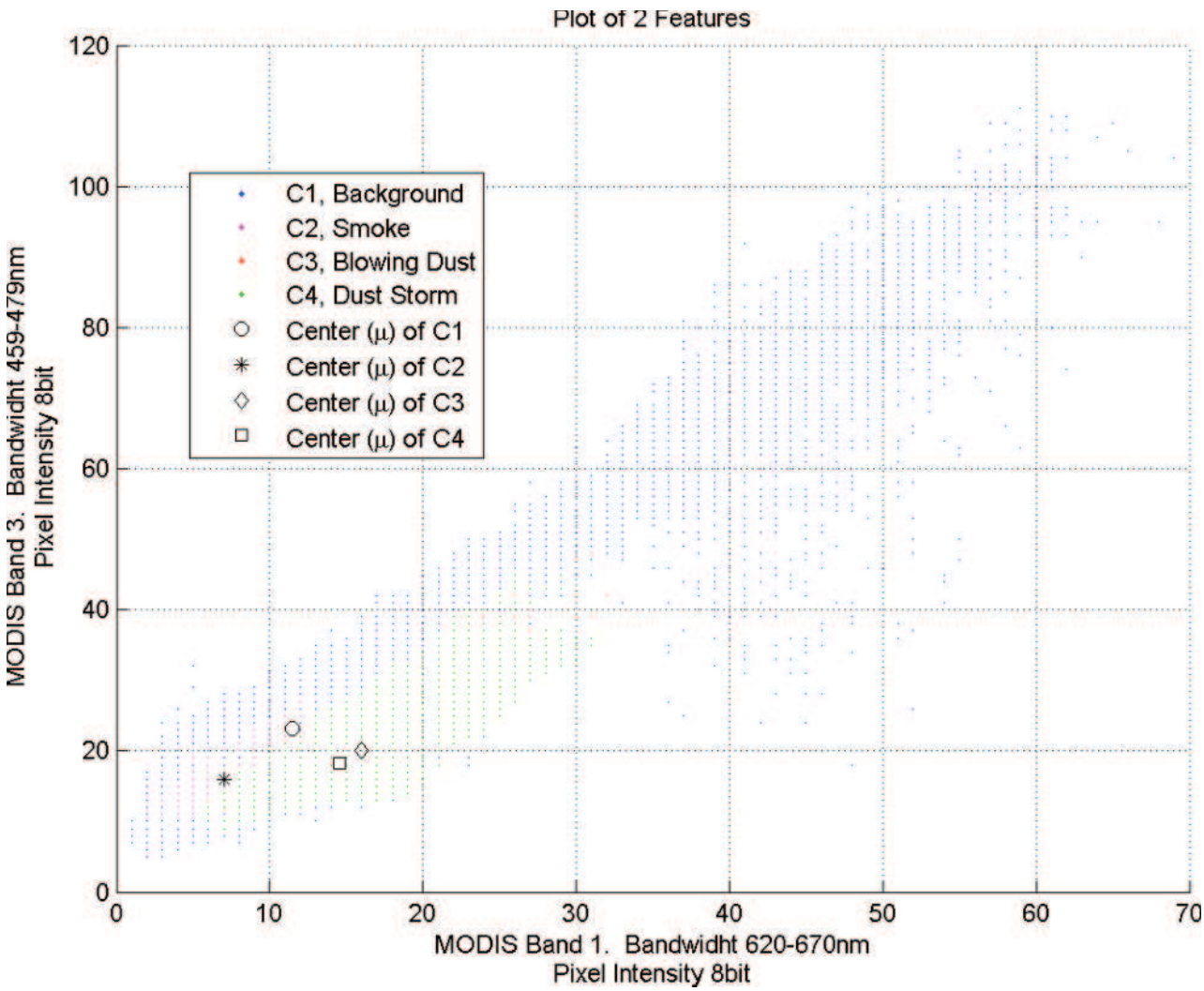

Figure 13. Distribution of feature vector $F^{(3)}$ when two of its components are ploted. In this case the x-axis correspond to $B1$, while the y-axis correspond to $B3$, raw 8-bit intensities.

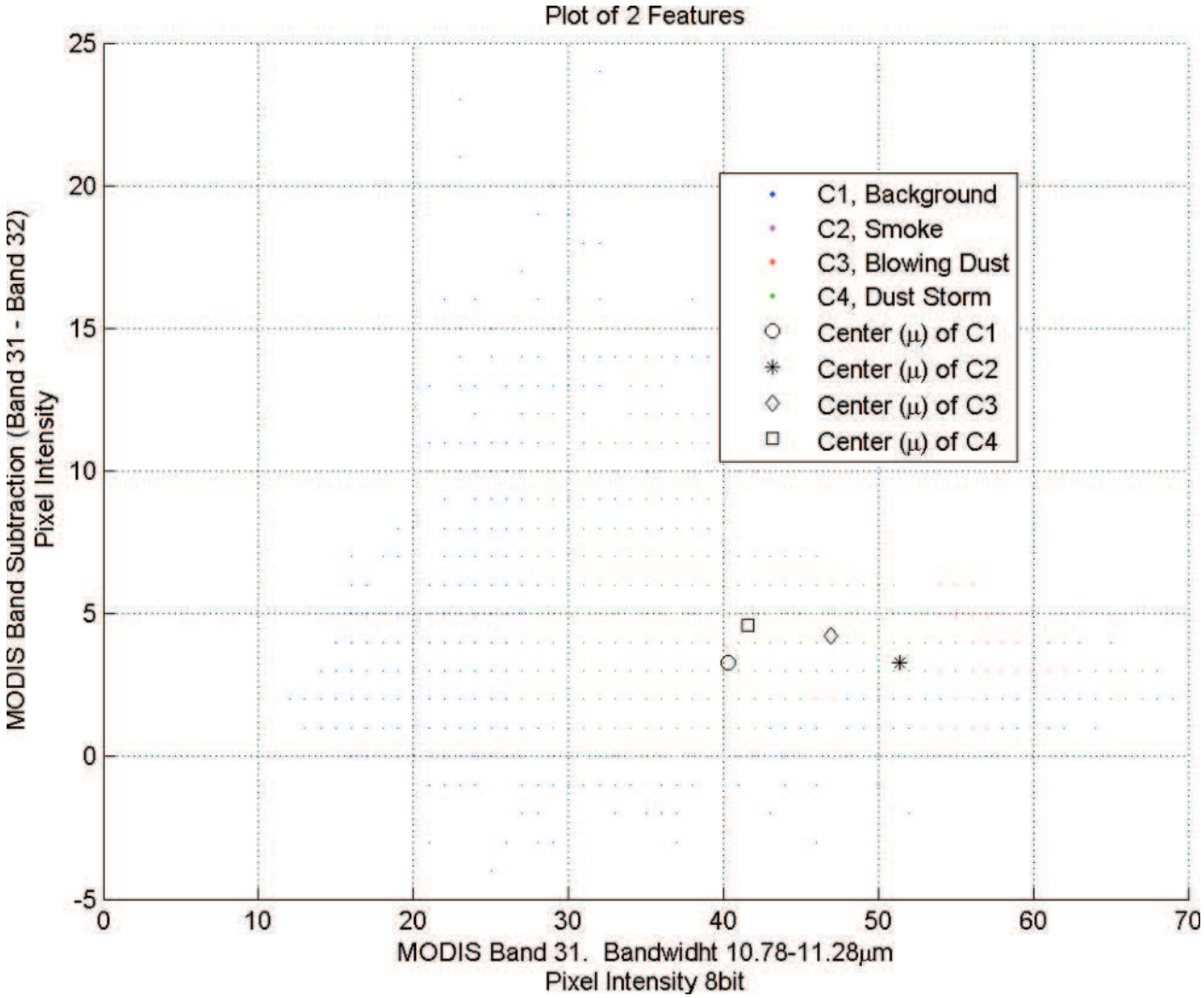

Figure 14. Distribution of feature vector $F^{(4)}$ when two of its components are ploted. In this case the x-axis correspond to $B31$, while the y-axis correspond to $B32 - 31$, raw 8-bit intensities.

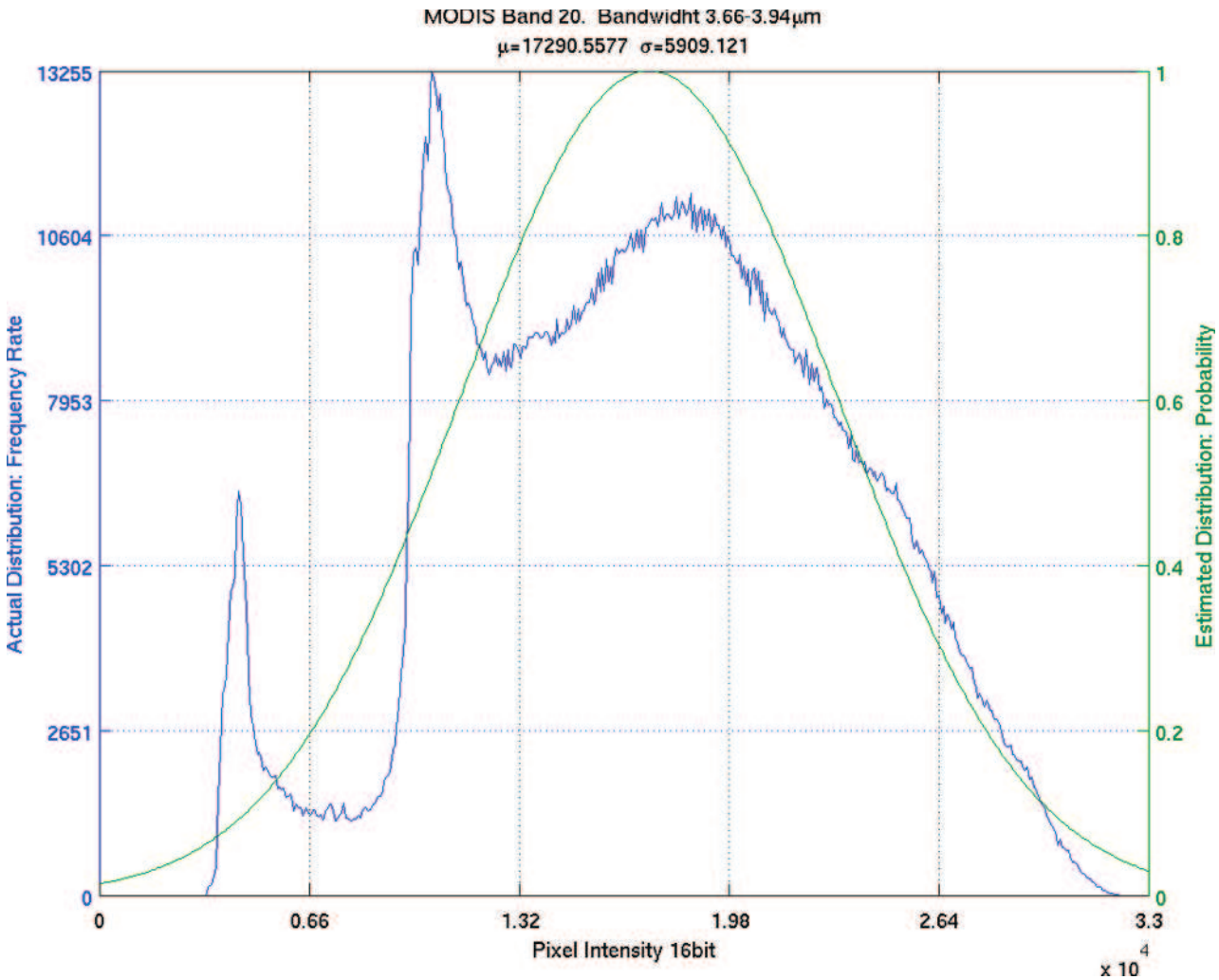

Figure 15. Histogram and estimated distribution of $f_{X_n^{(B20_{I16})}}(X_n^{(B20_{I16})} = x)$.

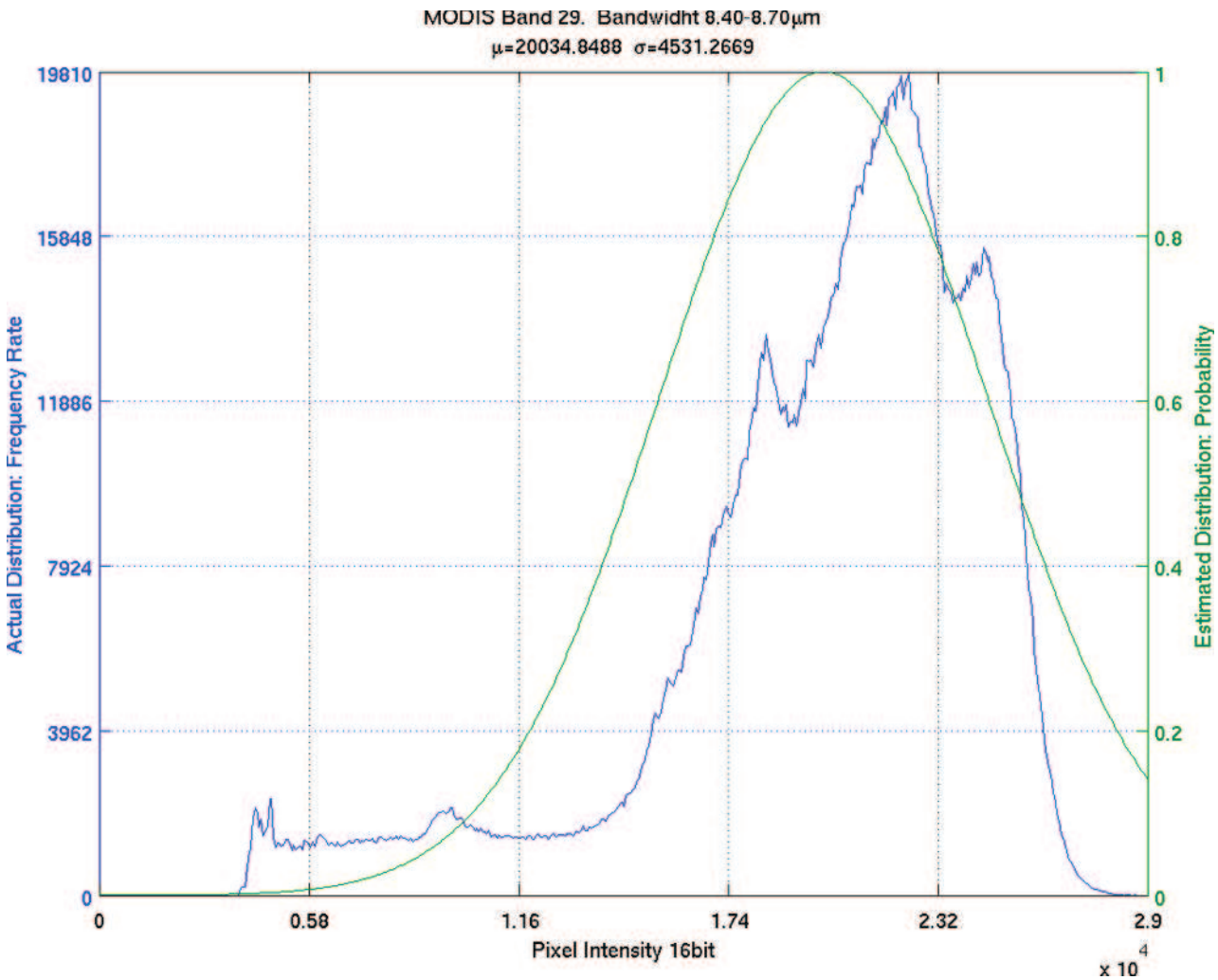

Figure 16. Histogram and estimated distribution of $f_{X_n^{(B29_{I16})}}(X_n^{(B29_{I16})} = x)$.

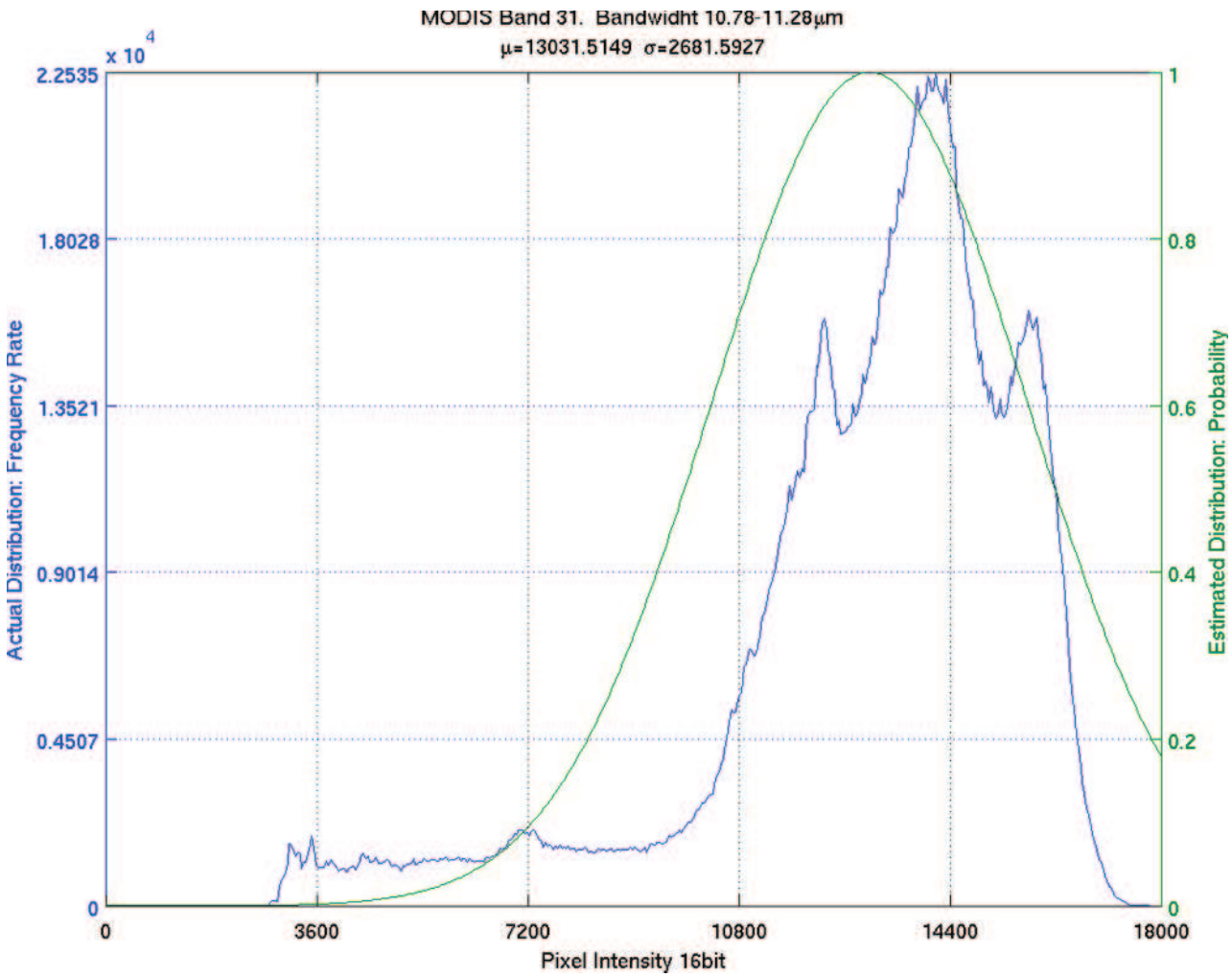

Figure 17. Histogram and estimated distribution of $f_{X_n^{(B31_{I16})}}(X_n^{(B31_{I16})} = x)$.

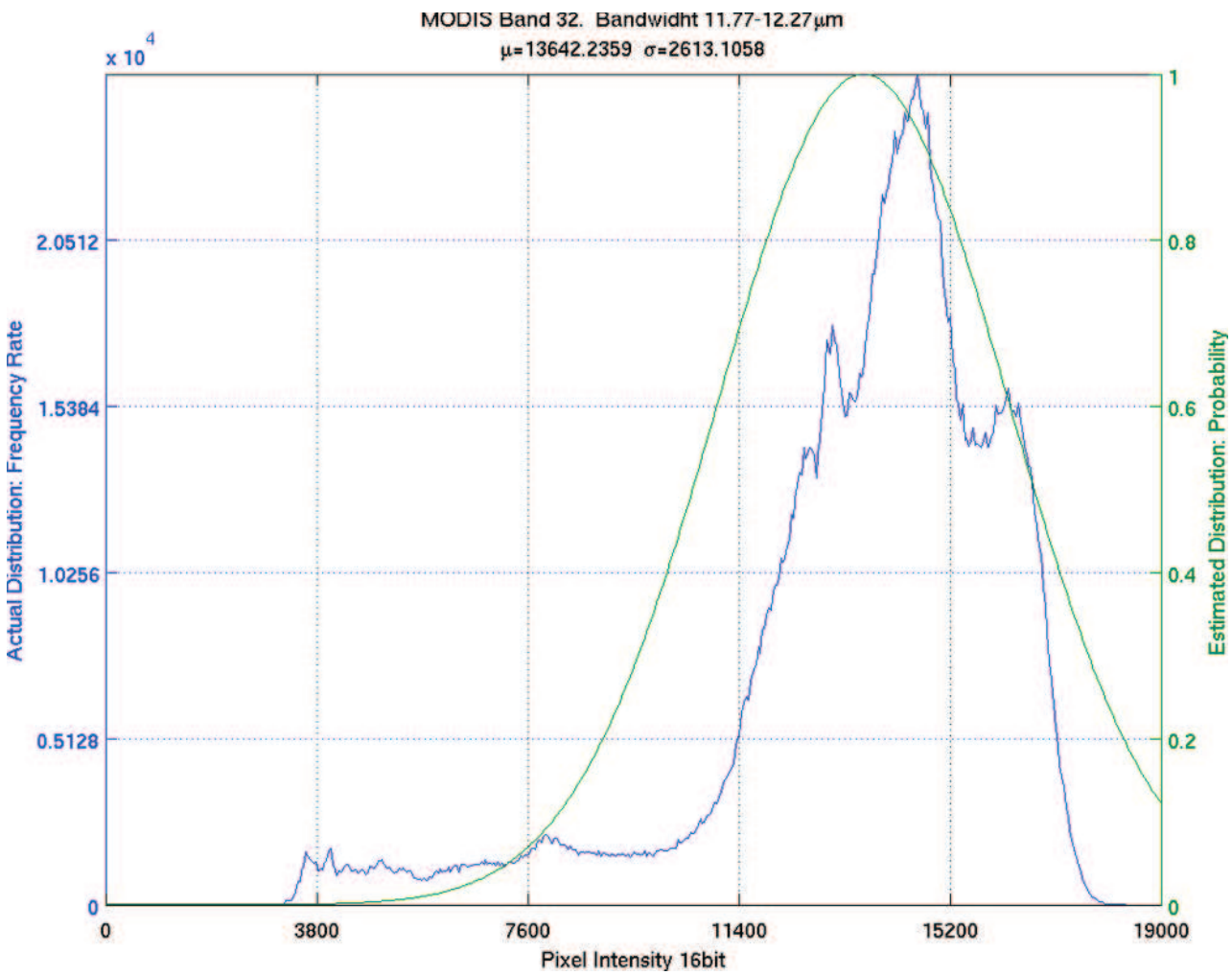

Figure 18. Histogram and estimated distribution of $f_{X_n^{(B32_{I16})}}(X_n^{(B32_{I16})} = x)$.

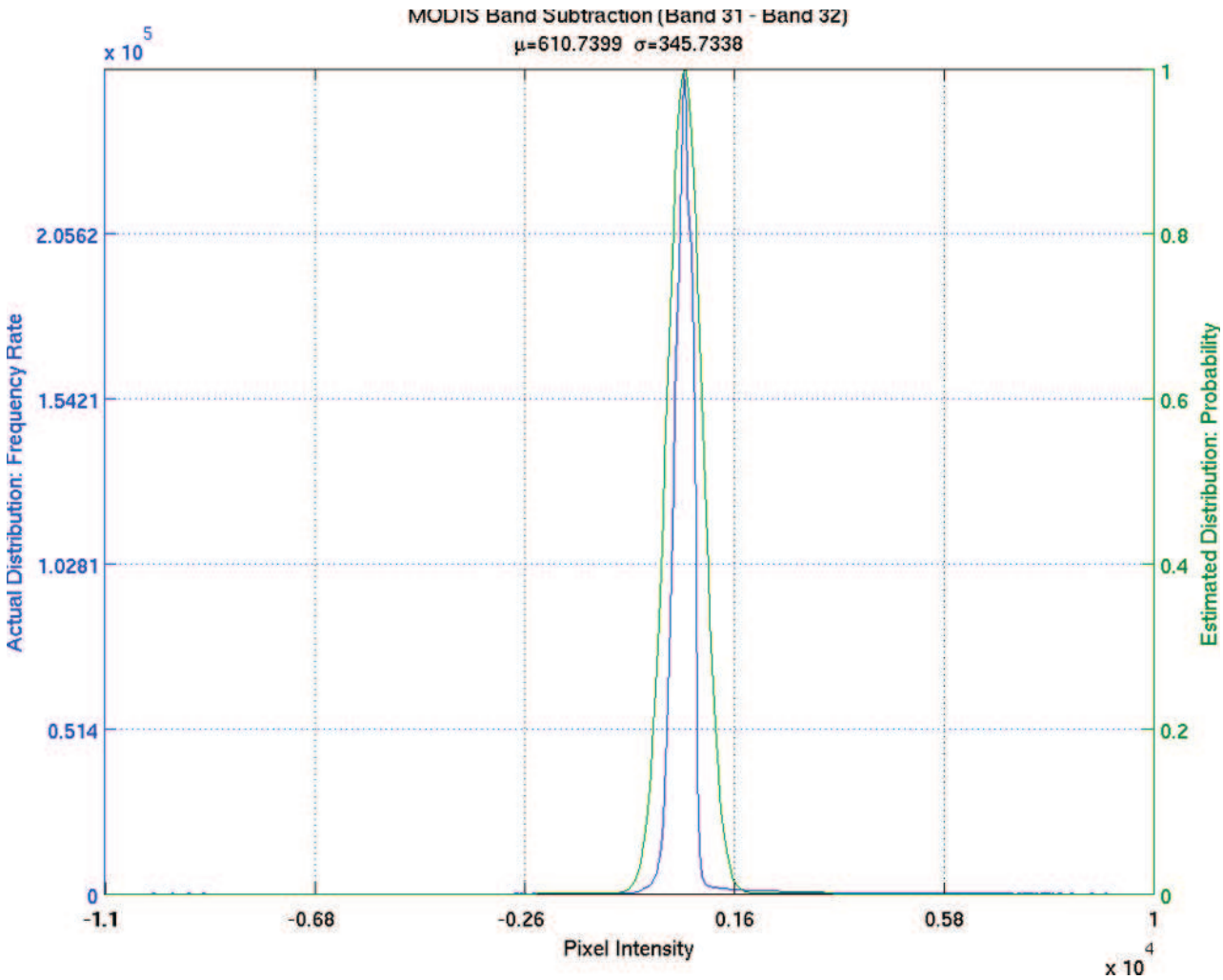

Figure 19. Histogram and estimated distribution of $f_{X_n^{(B32-31_{I16})}}(X_n^{(B32-31_{I16})} = x)$.

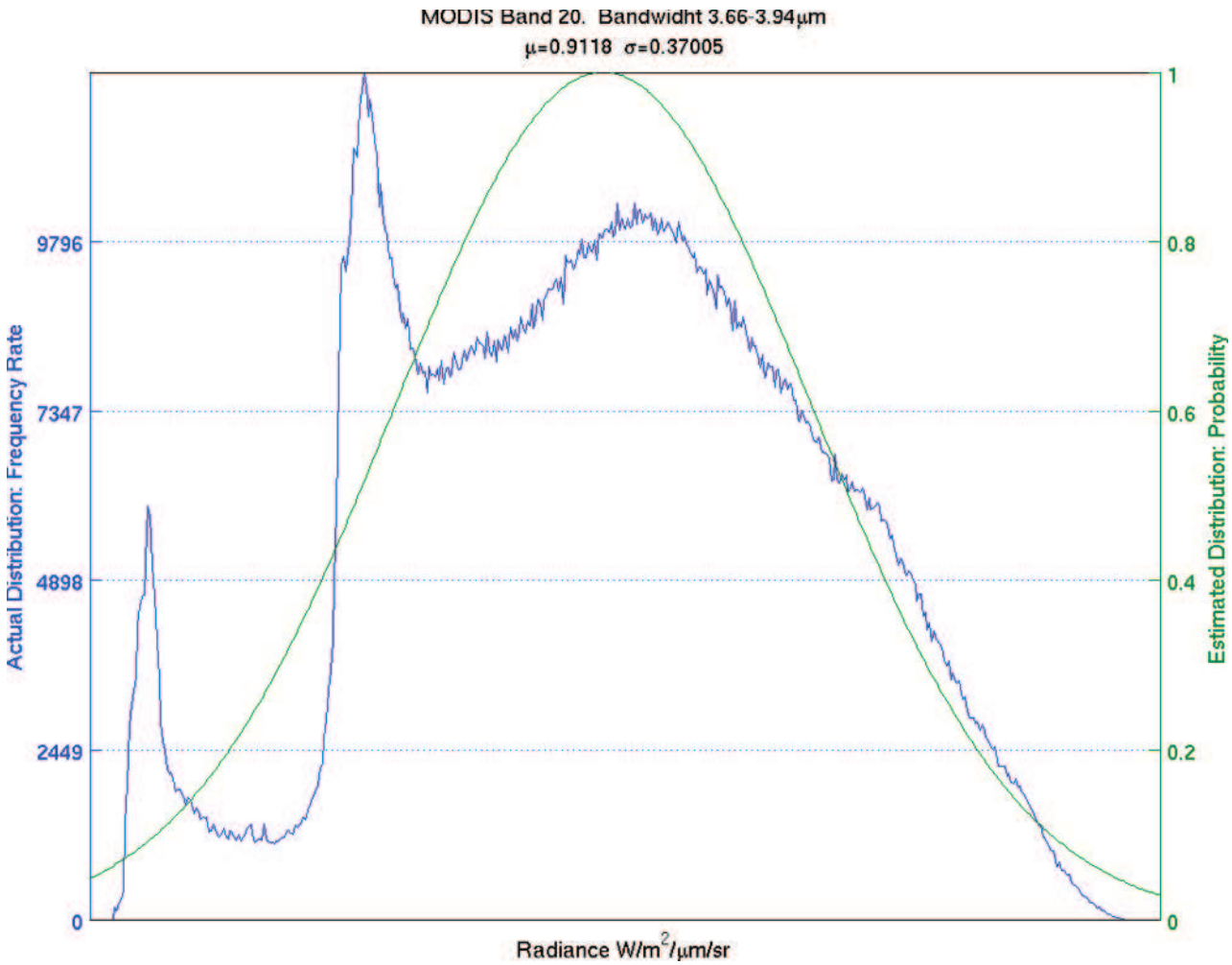

Figure 20. Histogram and estimated distribution of $f_{X_n^{(B20)}}(X_n^{(B20)} = x)$.

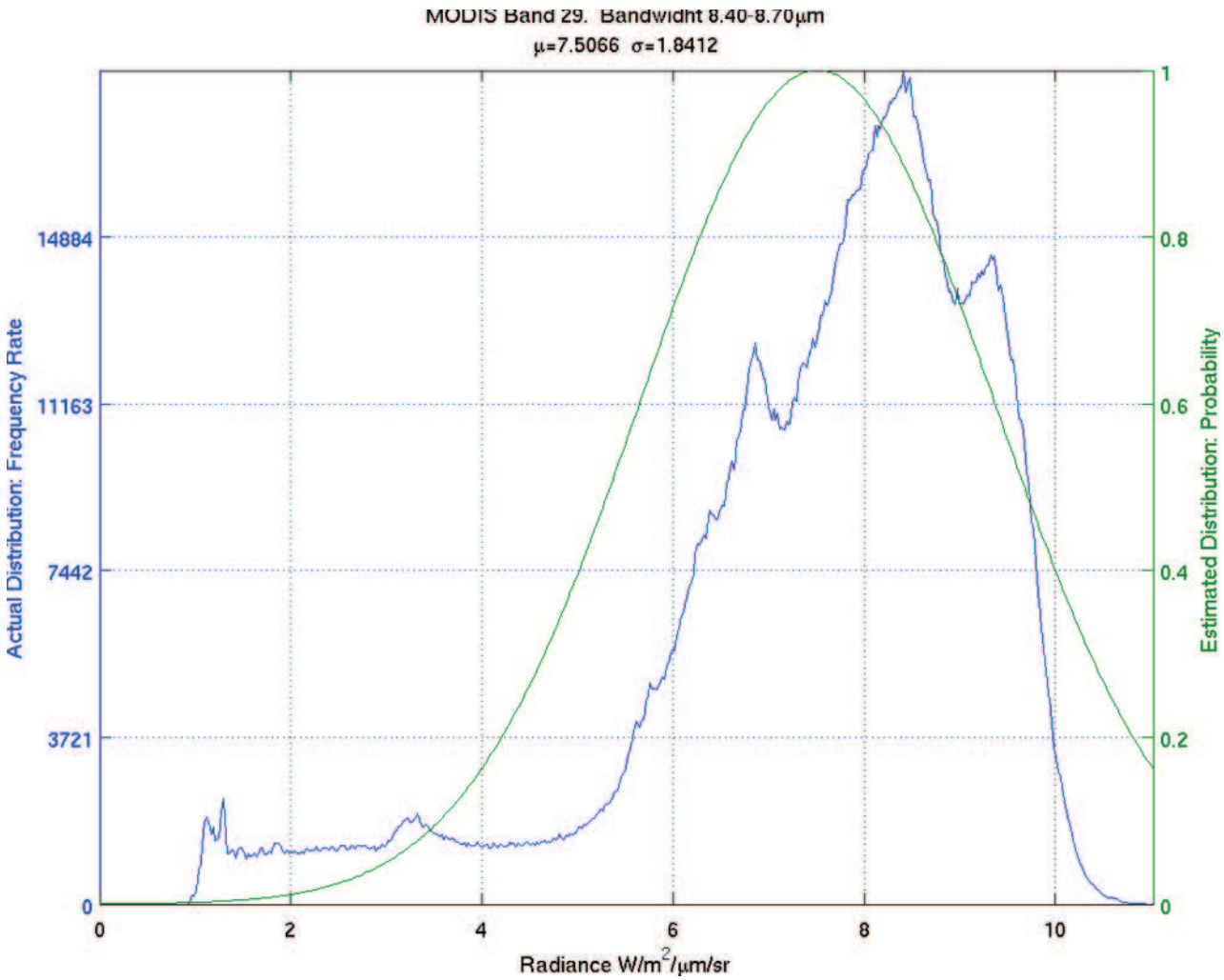

Figure 21. Histogram and estimated distribution of $f_{X_n^{(B29)}}(X_n^{(B29)} = x)$.

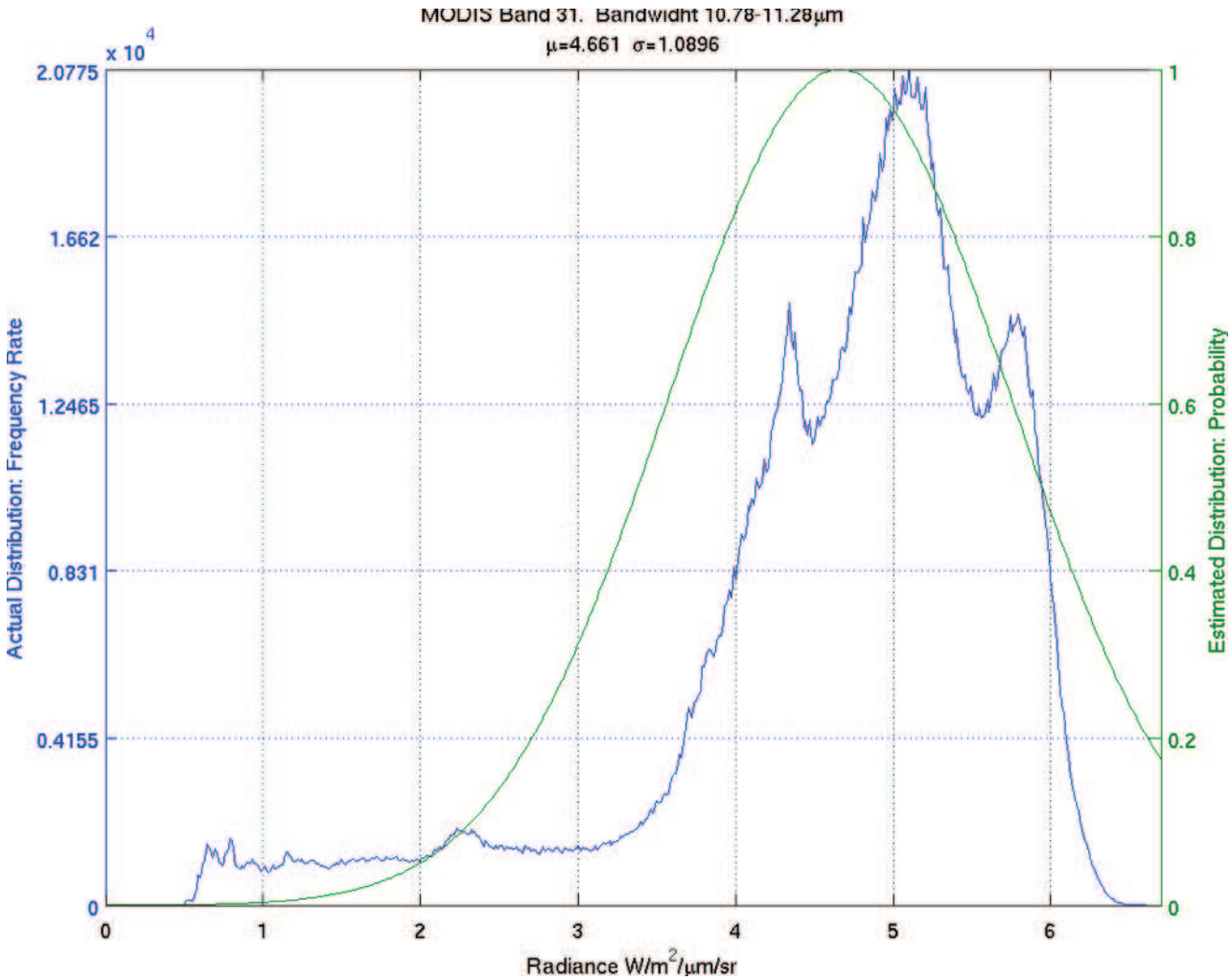

Figure 22. Histogram and estimated distribution of $f_{X_n^{(B31)}}(X_n^{(B31)} = x)$.

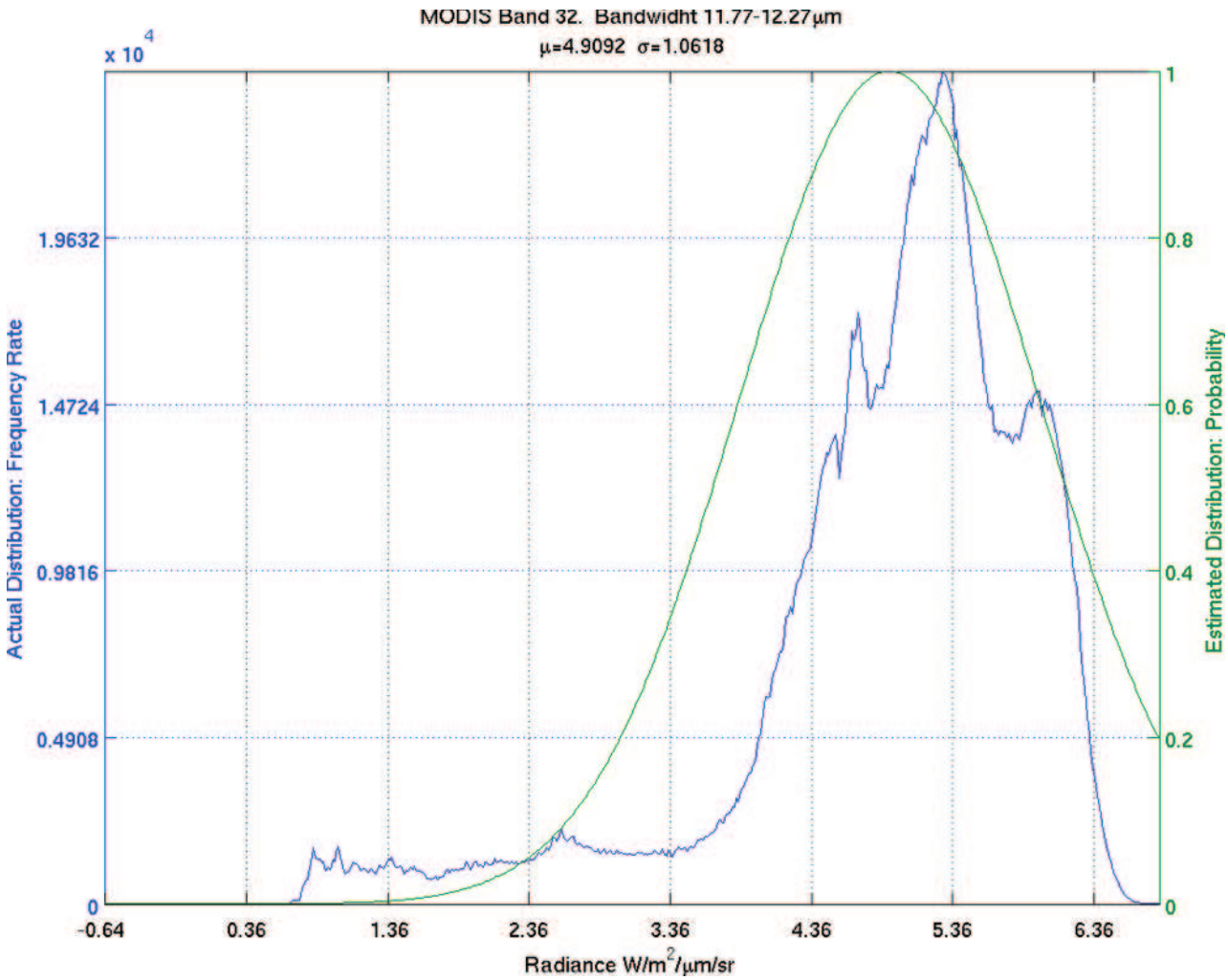

Figure 23. Histogram and estimated distribution of $f_{X_n^{(B32)}}(X_n^{(B32)} = x)$.

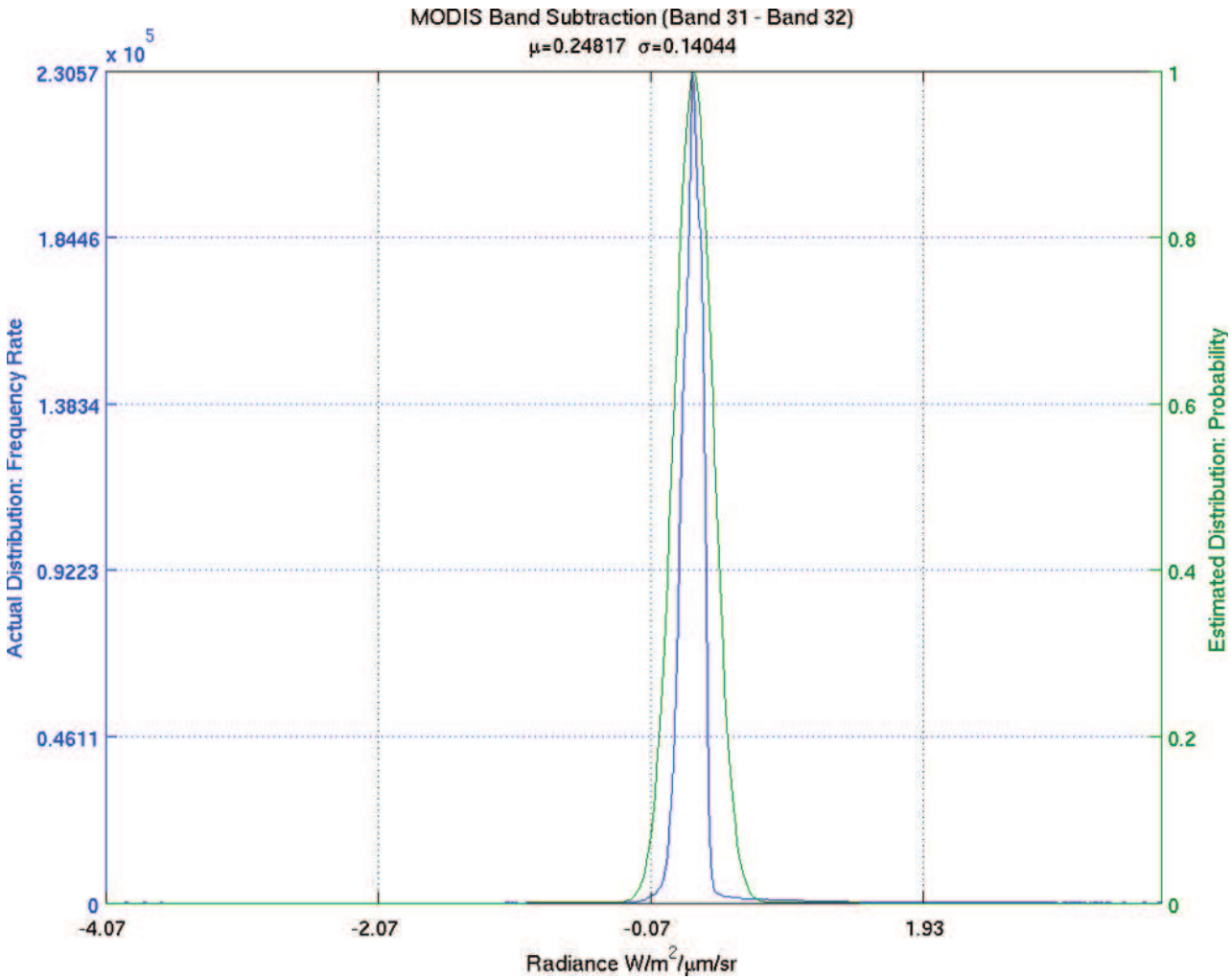

Figure 24. Histogram and estimated distribution of $f_{X_n^{(B32-31)}}(X_n^{(B32-31)} = x)$.

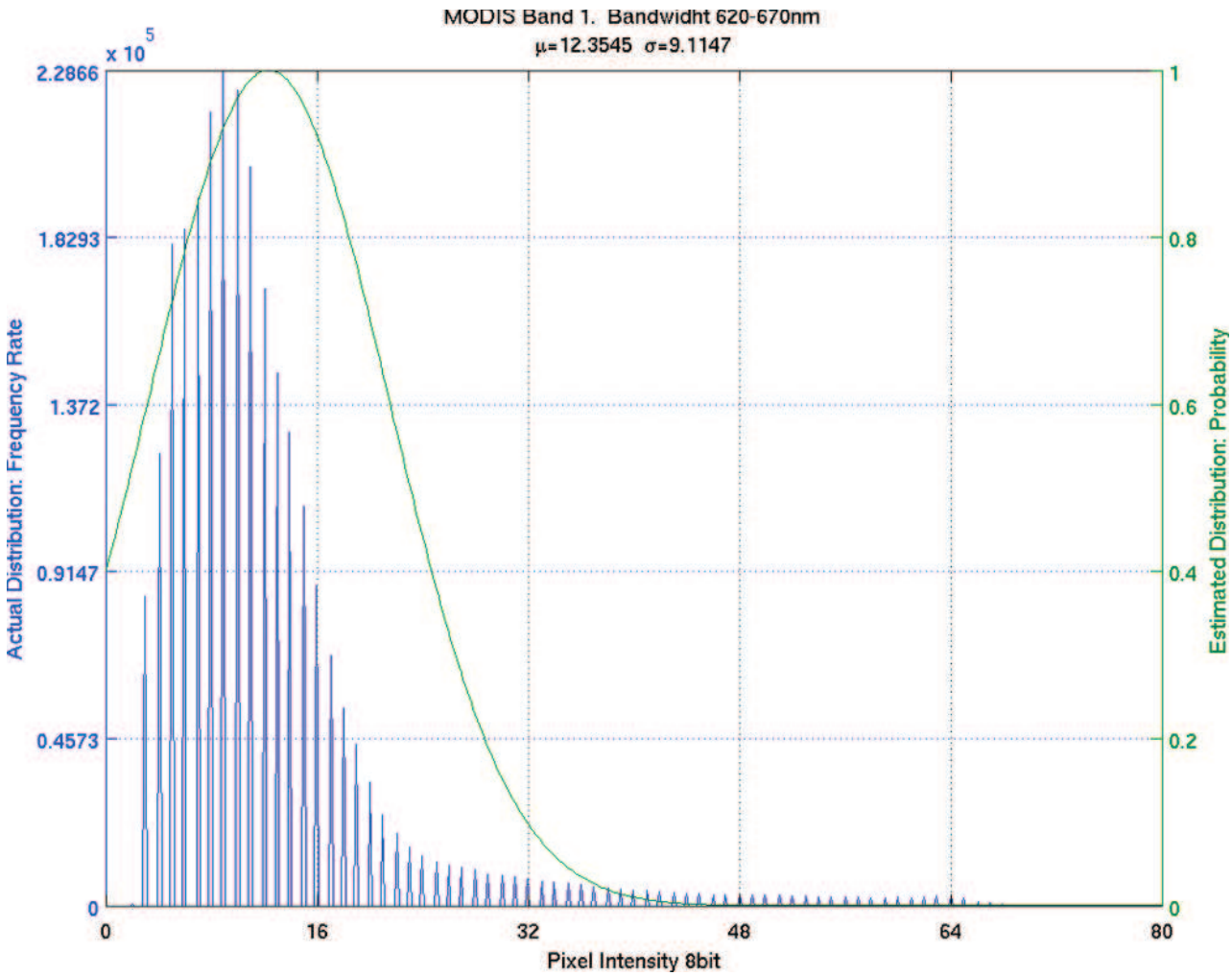

Figure 25. Histogram and estimated distribution of $f_{X_n^{(B1_{I8})}}(X_n^{(B1_{I8})} = x)$.

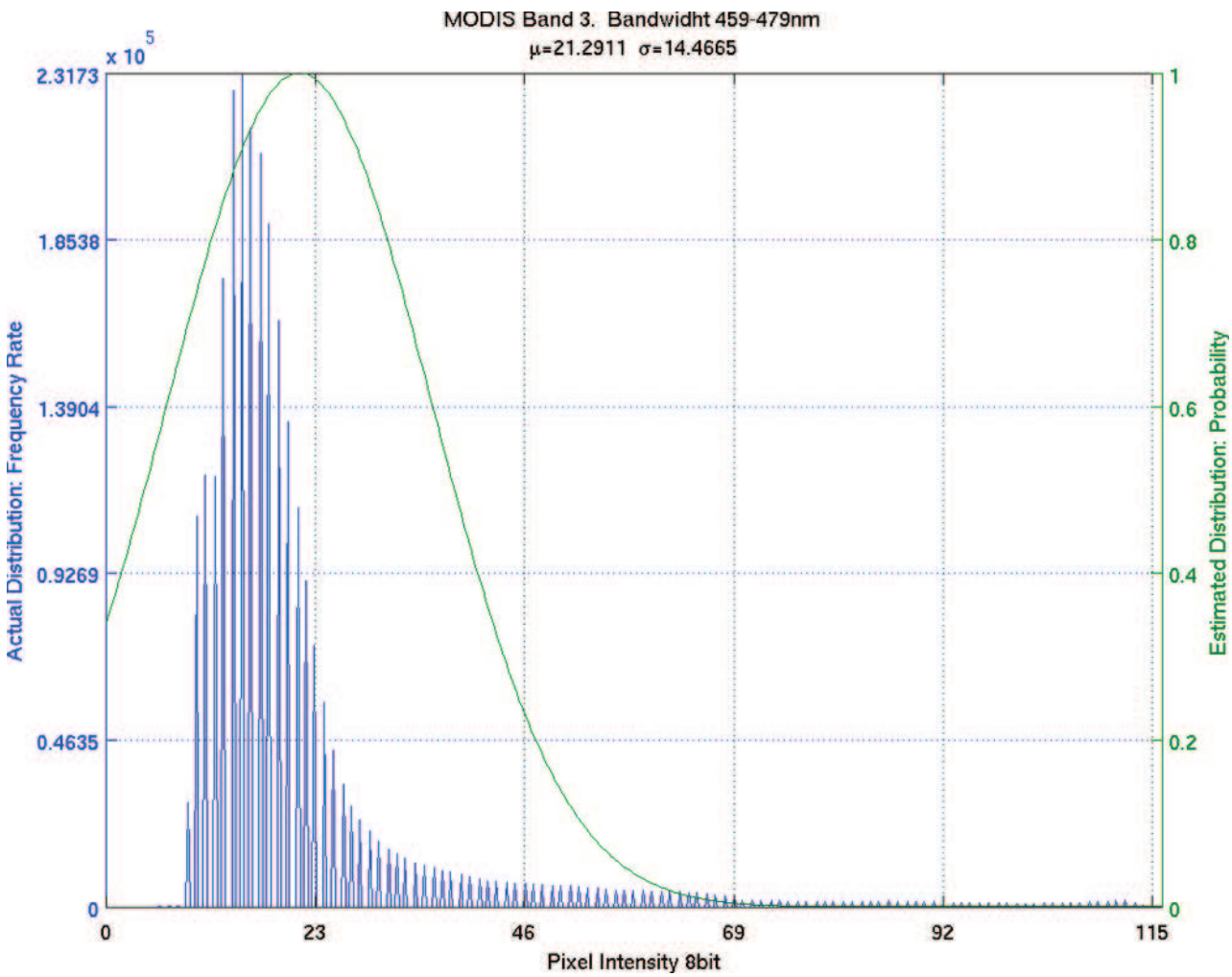

Figure 26. Histogram and estimated distribution of $f_{X_n^{(B3_{I8})}}(X_n^{(B3_{I8})} = x)$.

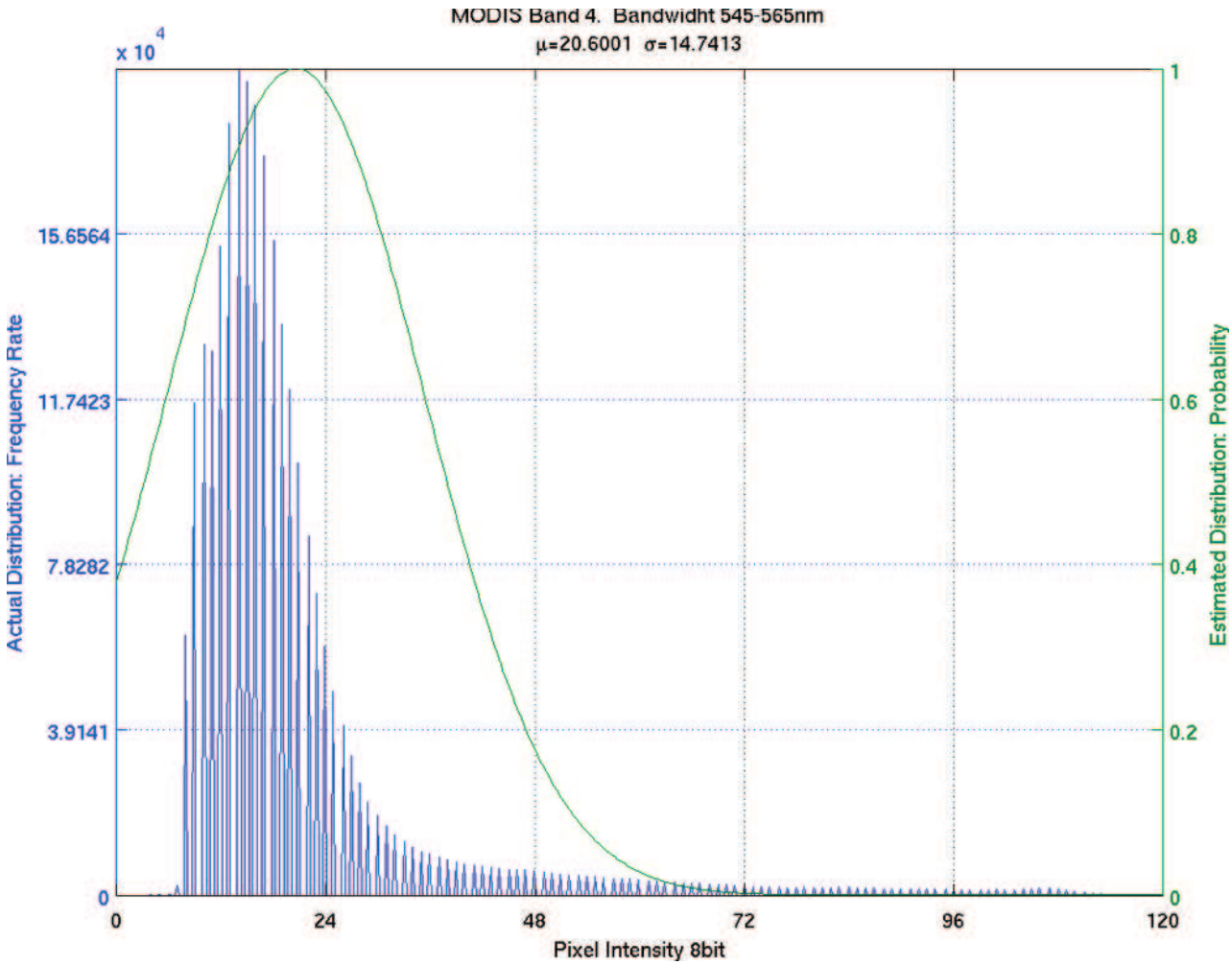

Figure 27. Histogram and estimated distribution of $f_{X_n^{(B4_{I8})}}(X_n^{(B4_{I8})} = x)$.

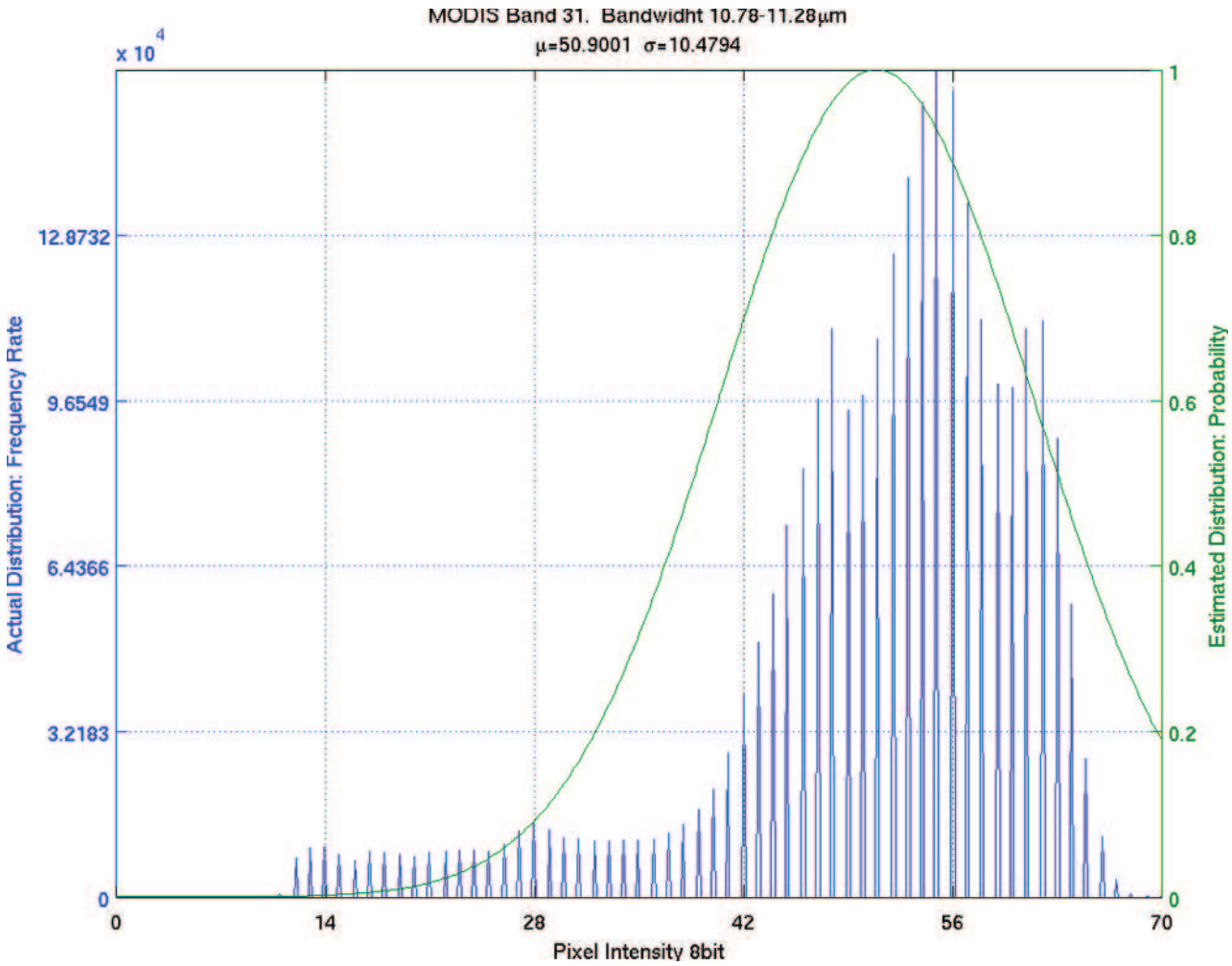

Figure 28. Histogram and estimated distribution of $f_{X_n^{(B31_{I8})}}(X_n^{(B31_{I8})} = x)$.

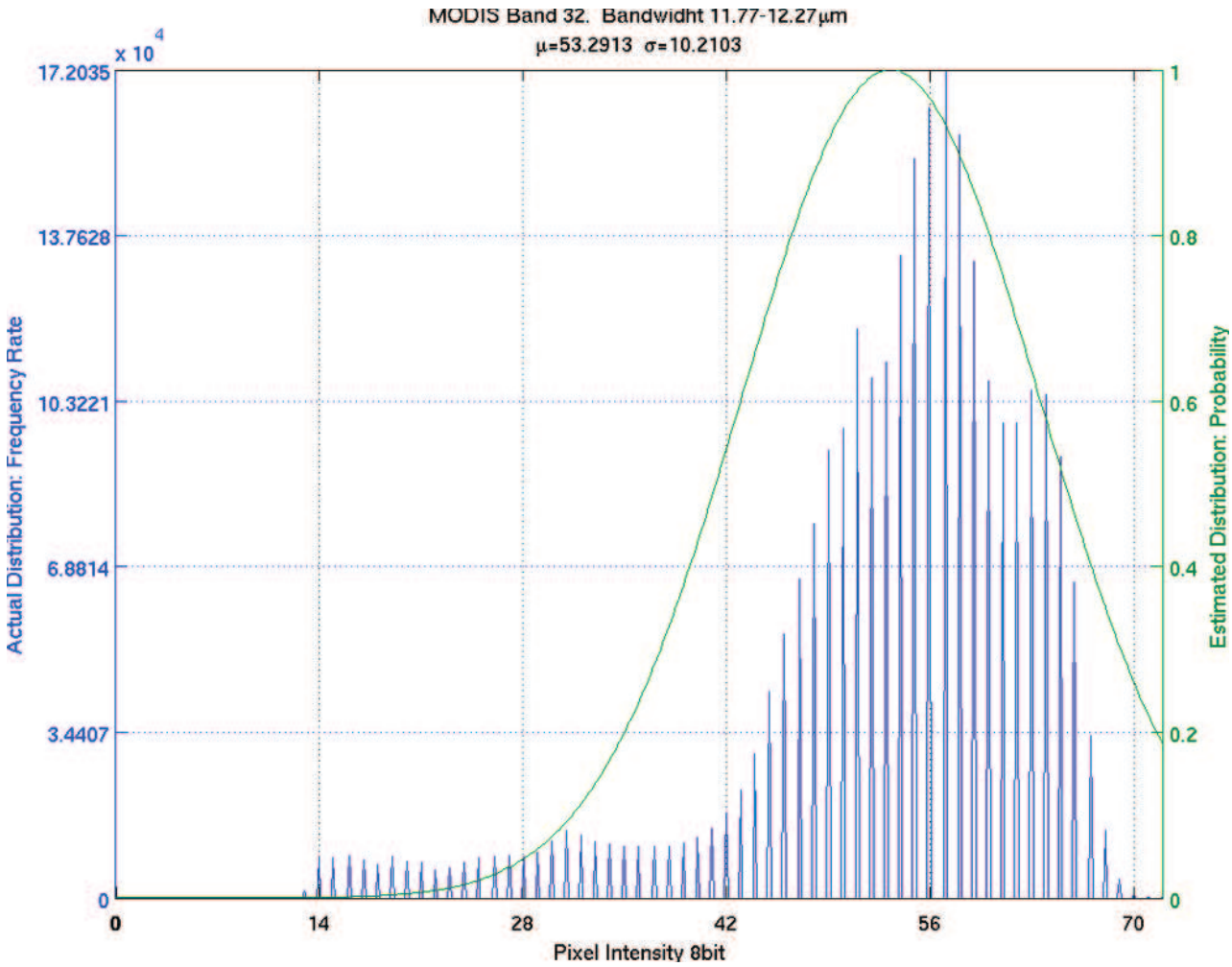

Figure 29. Histogram and estimated distribution of $f_{X_n^{(B32_{I8})}}(X_n^{(B32_{I8})} = x)$.

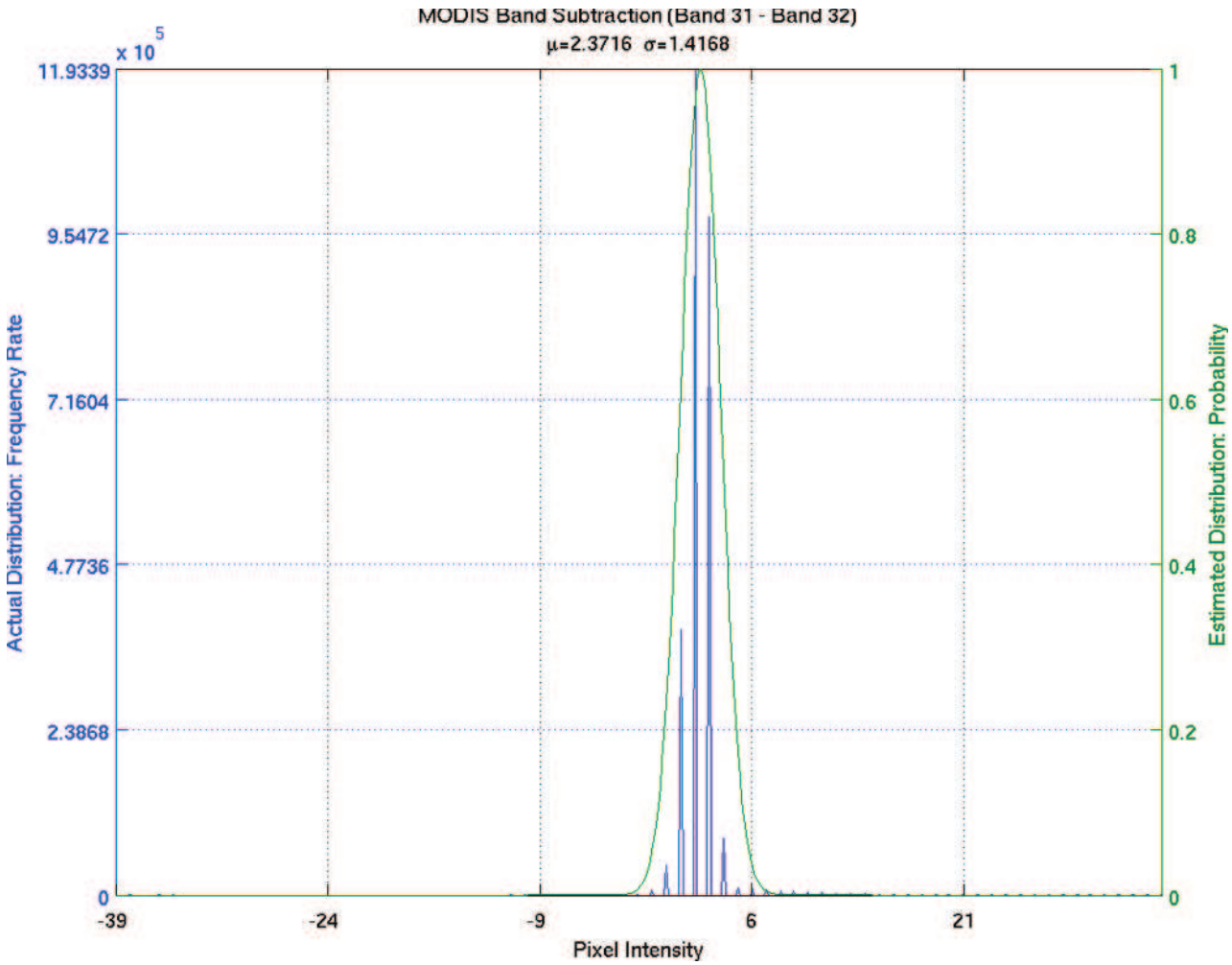

Figure 30. Histogram and estimated distribution of $f_{X_n^{(B32-31_{I8})}}(X_n^{(B32-31_{I8})} = x)$.

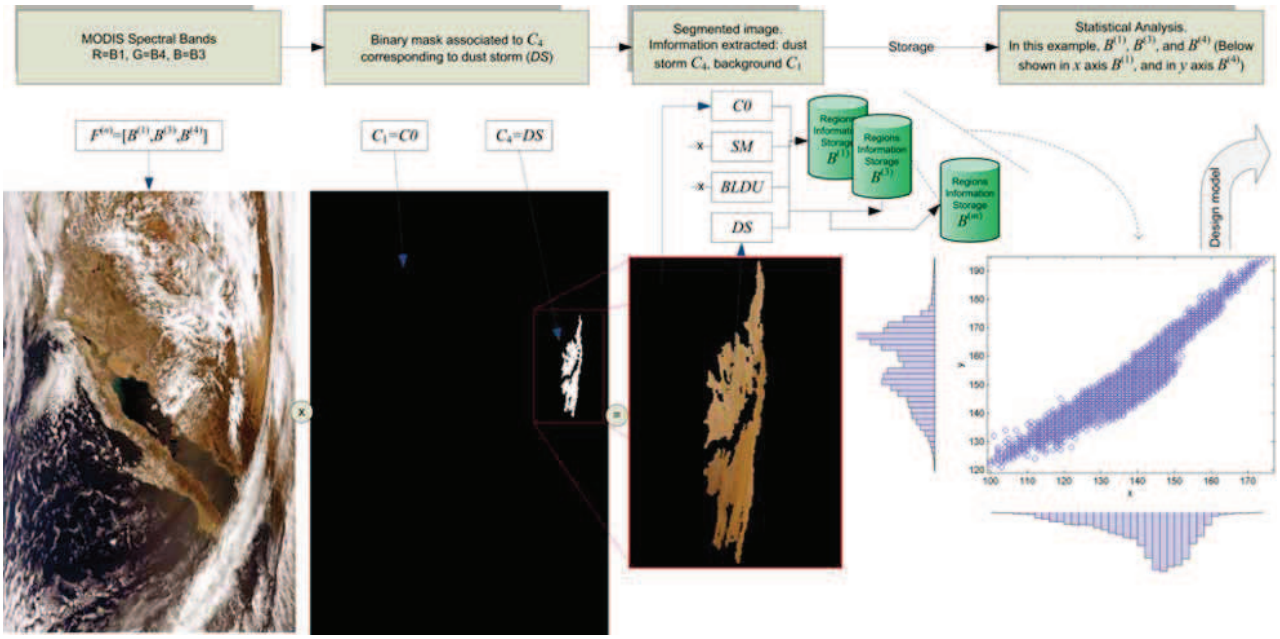

Figure 31. Process of estimation of $f_{X_{\mathbf{n}}^{(Bm)}|C_j}\left(X_{\mathbf{n}}^{(Bm)}=x|C_j=c\right)$ through segmentation (using masks). Regions obtained for each band after segmentation are analyzed for the final estimation of parameters. The mask is associated with a particular class C_j such that we can associate the segmented data to the $j-$th class. Then, the data is stored in a set that contains all the samples for the same class, and the process is repeated until the data collection and segmentation is finished for all the events selected for modeling. A statistical analysis is followed, in which the mean vectors and the covariance matrices are computed for each class features. Note that the $\log_{10}$ was computed for B1, B4, and B3 only for visualization purposes in this example. The current analysis consider the data as in $F^{(3)}$ (without the logarithm).

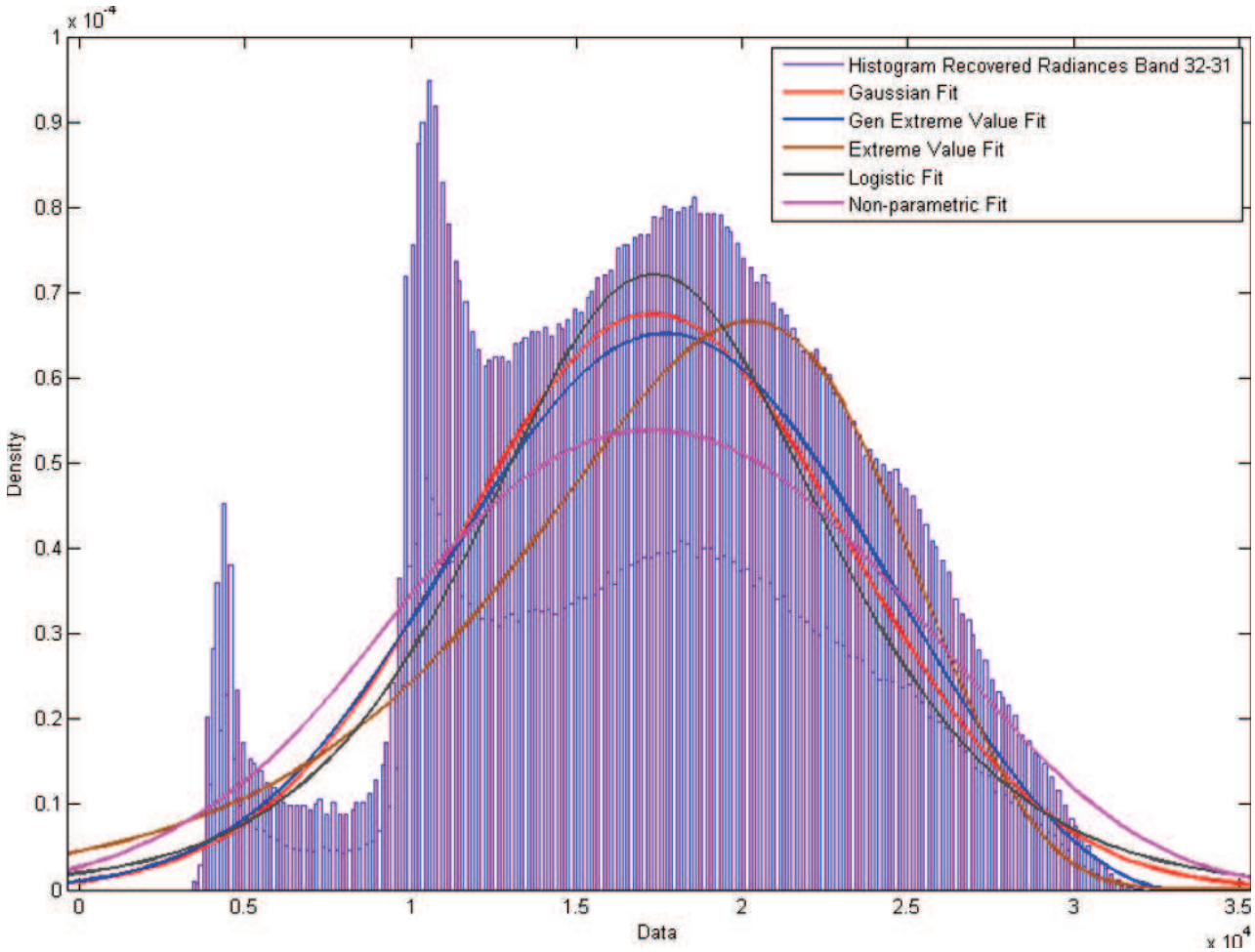

Figure 32. A comparison between different distribution fitting and estimation. Here the x-axis (Data) is the raw intensity of $B20$.

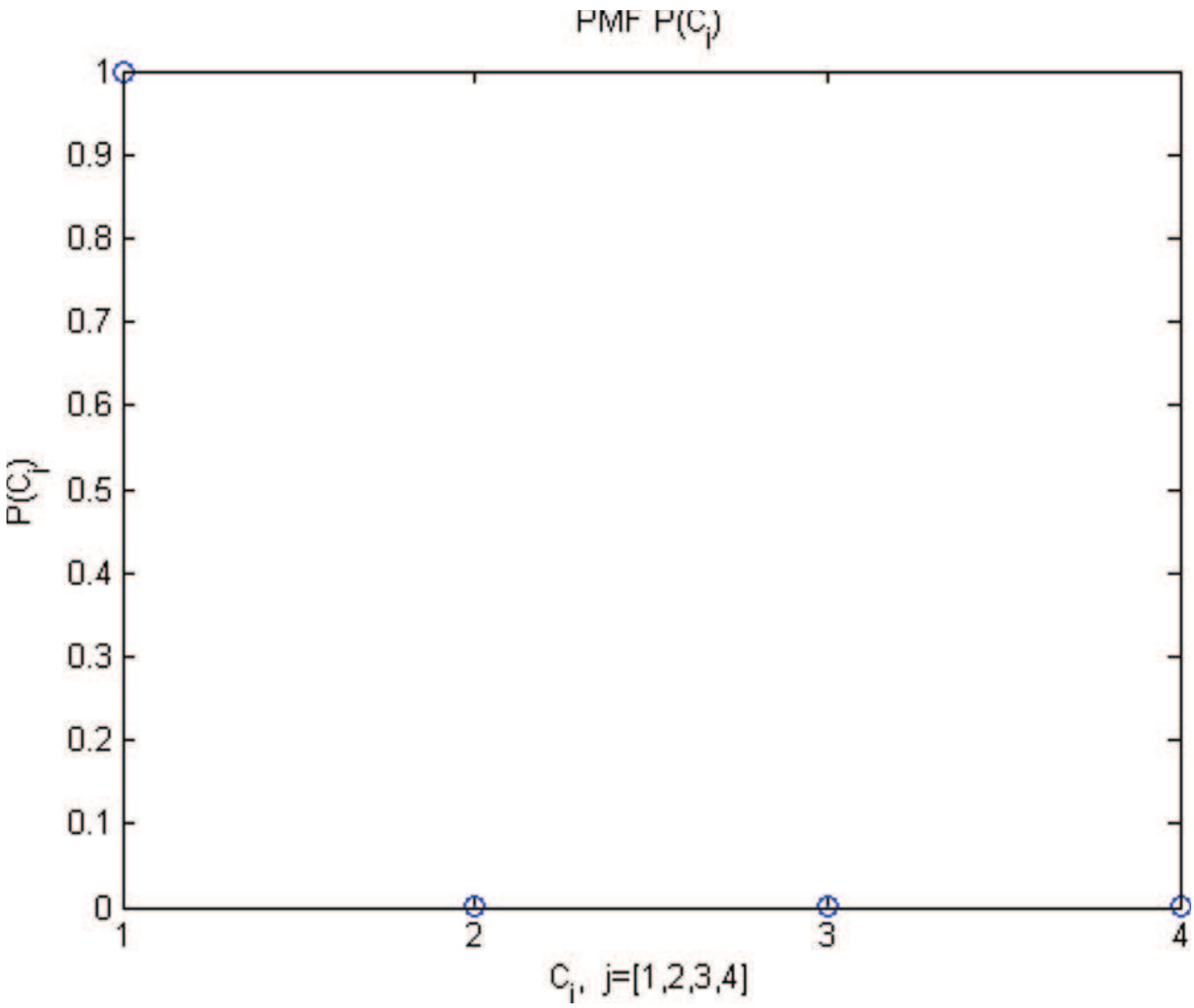

Figure 33. Estimated probability mass function of $f_{C_j}(C_j = c)$.

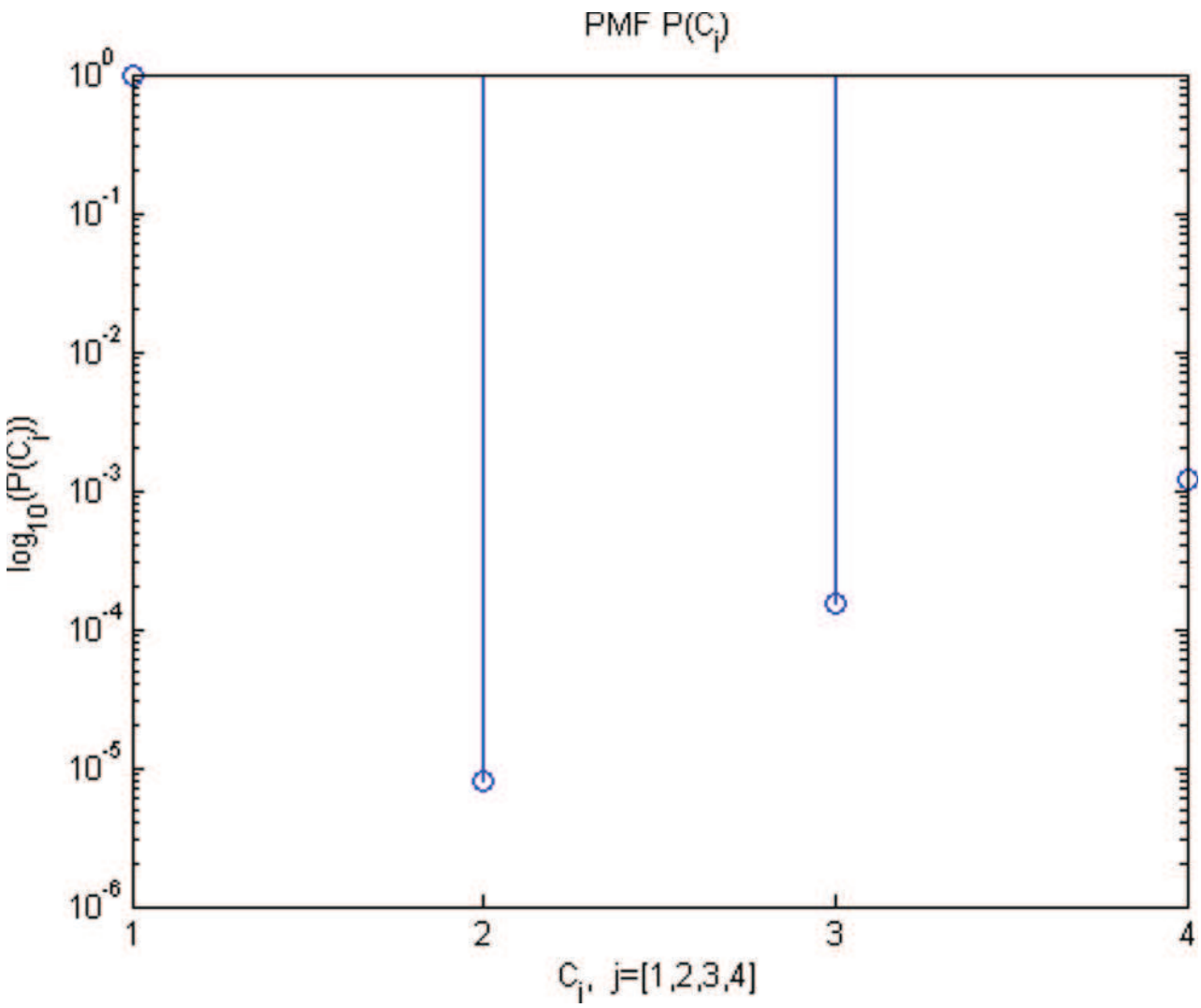

Figure 34. Logarithm of estimated probability mass function, $\log_{10}\left(f_{C_j}(C_j = c)\right)$.

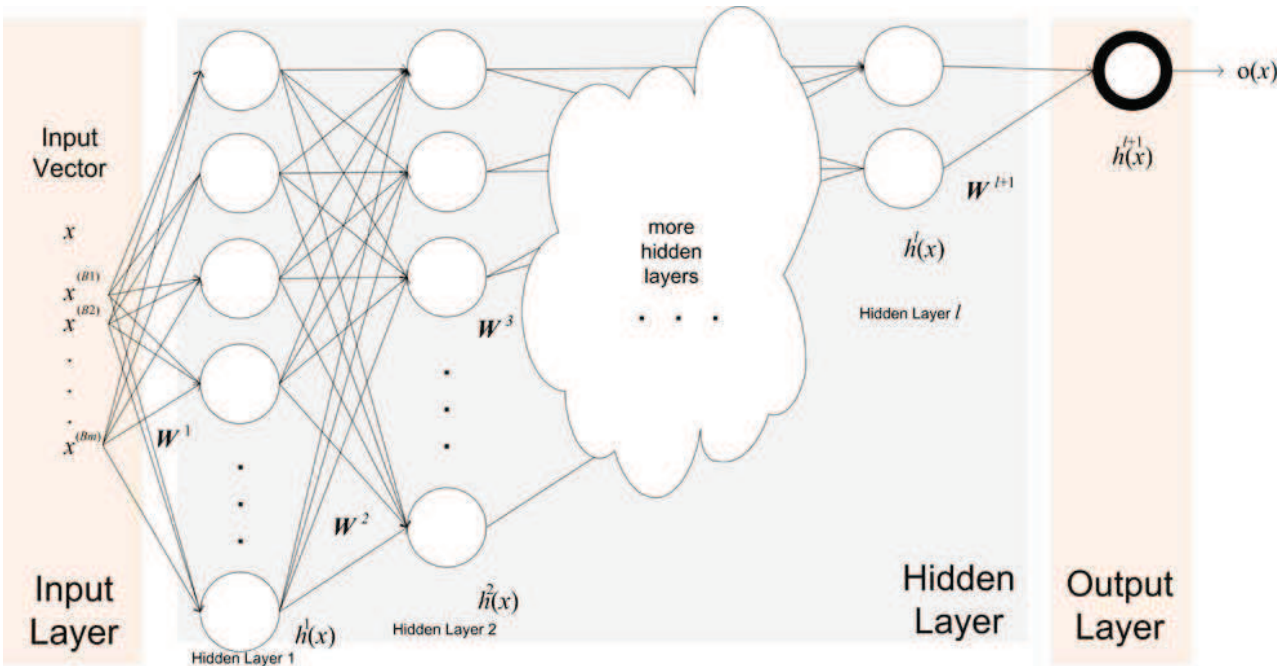

Figure 35. General architecture, parameters and information flow for a Multilayered Feedforward Backpropagation Neural Network.

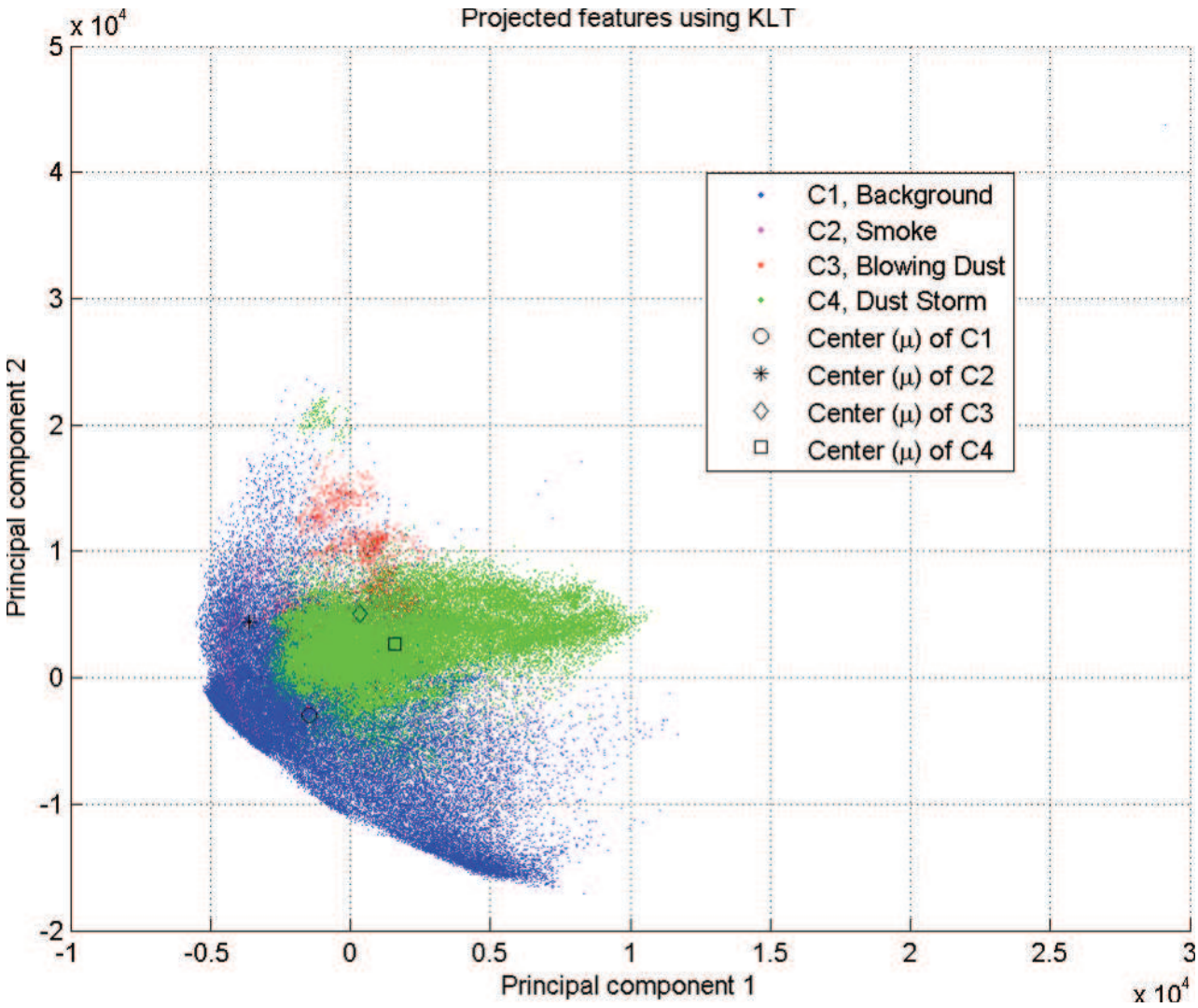

Figure 36. Distribution of feature vector $F^{(5)}$ when two of its principal components are ploted.

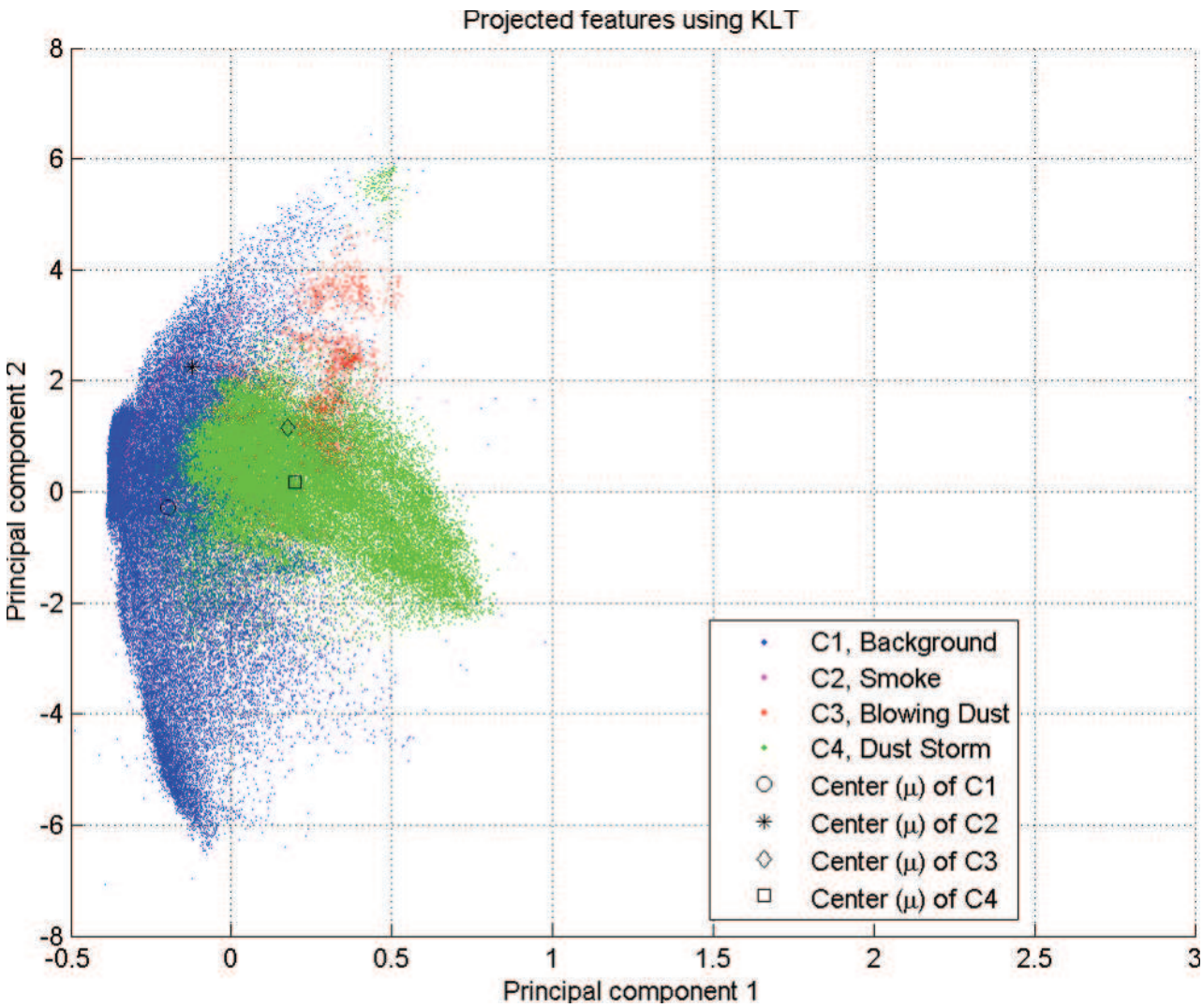

Figure 37. Distribution of feature vector $F^{(6)}$ when two of its principal components are ploted.

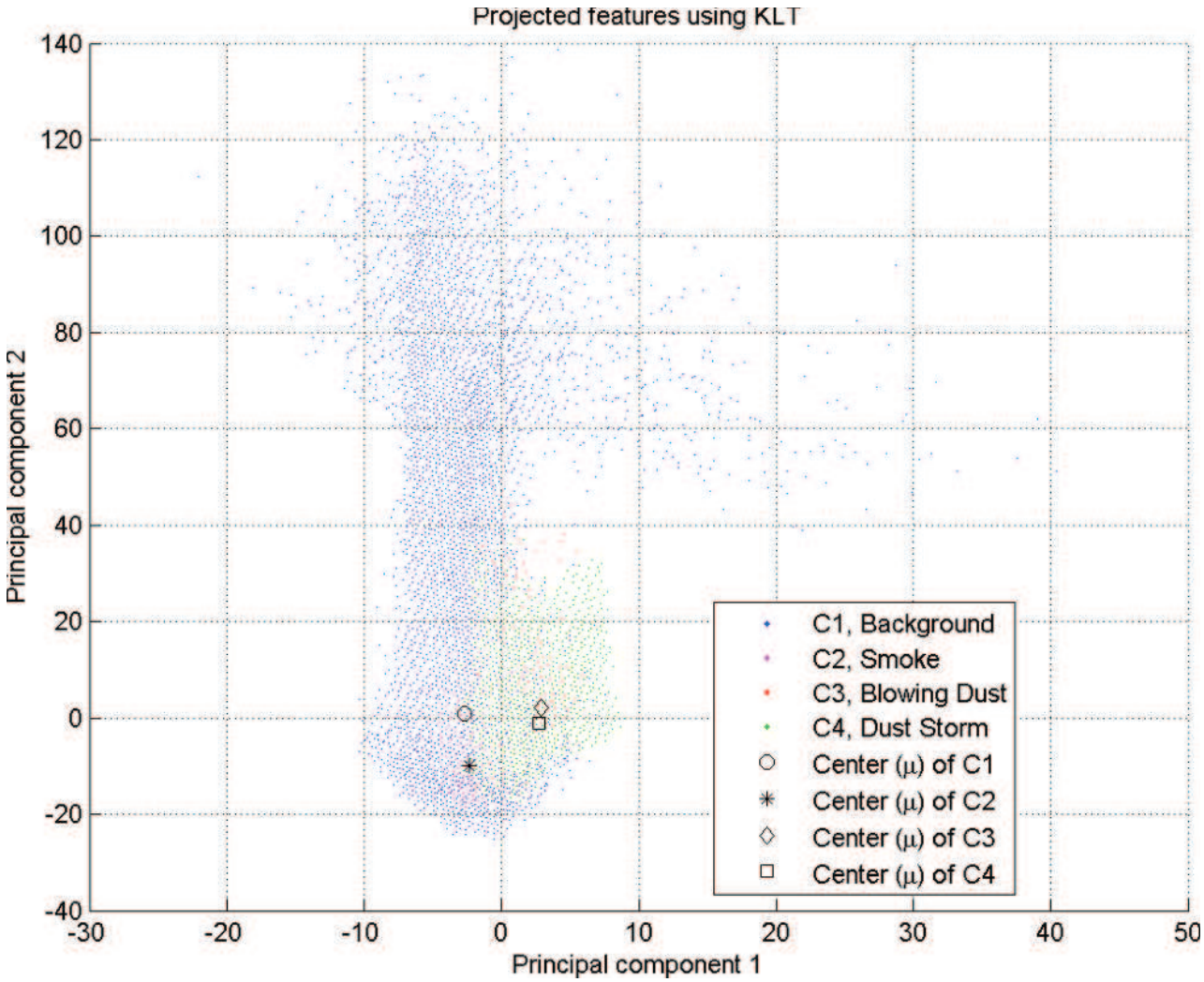

Figure 38. Distribution of feature vector $F^{(7)}$ when two of its principal components are ploted.

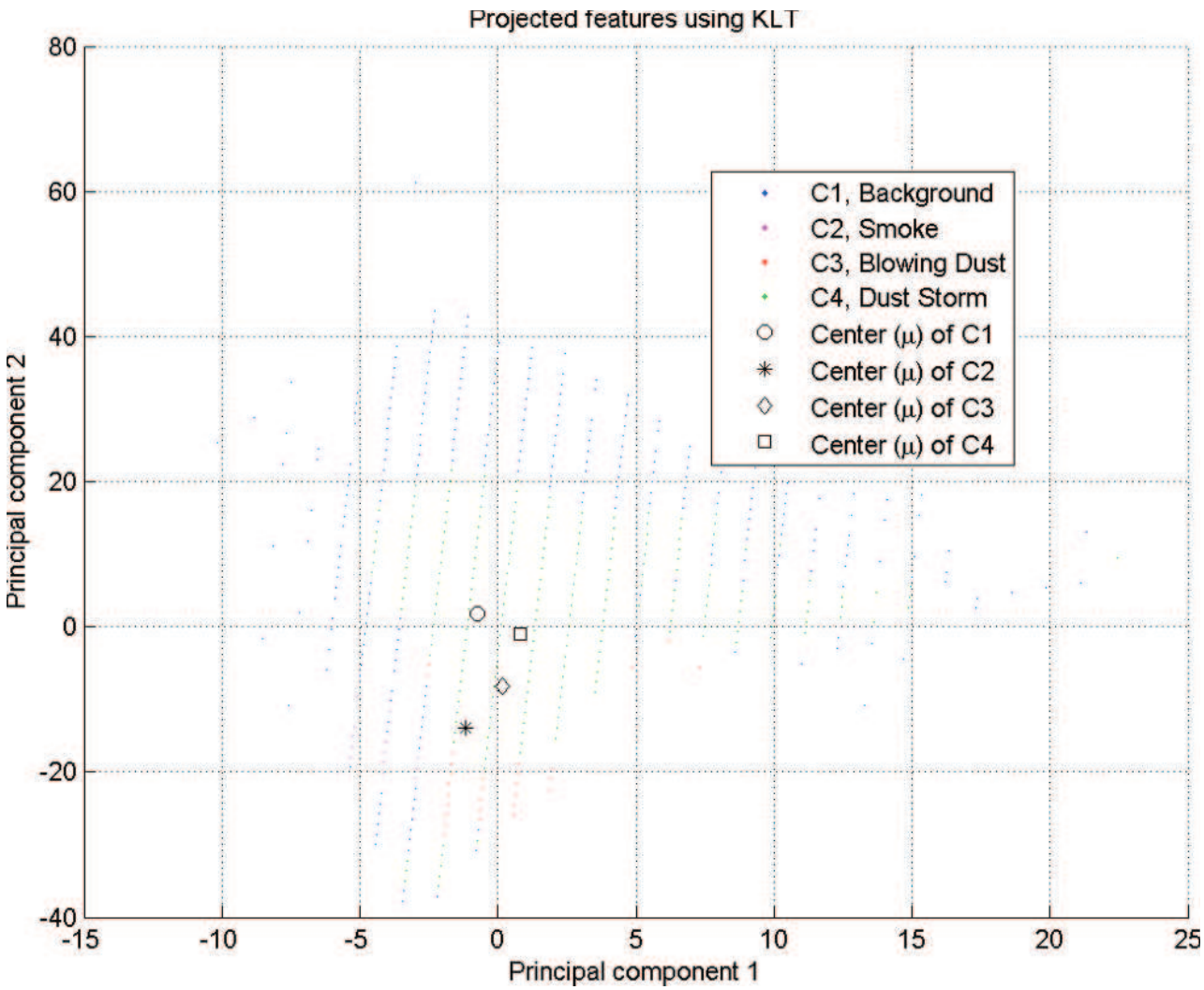

Figure 39. Distribution of feature vector $F^{(8)}$ when two of its principal components are ploted.

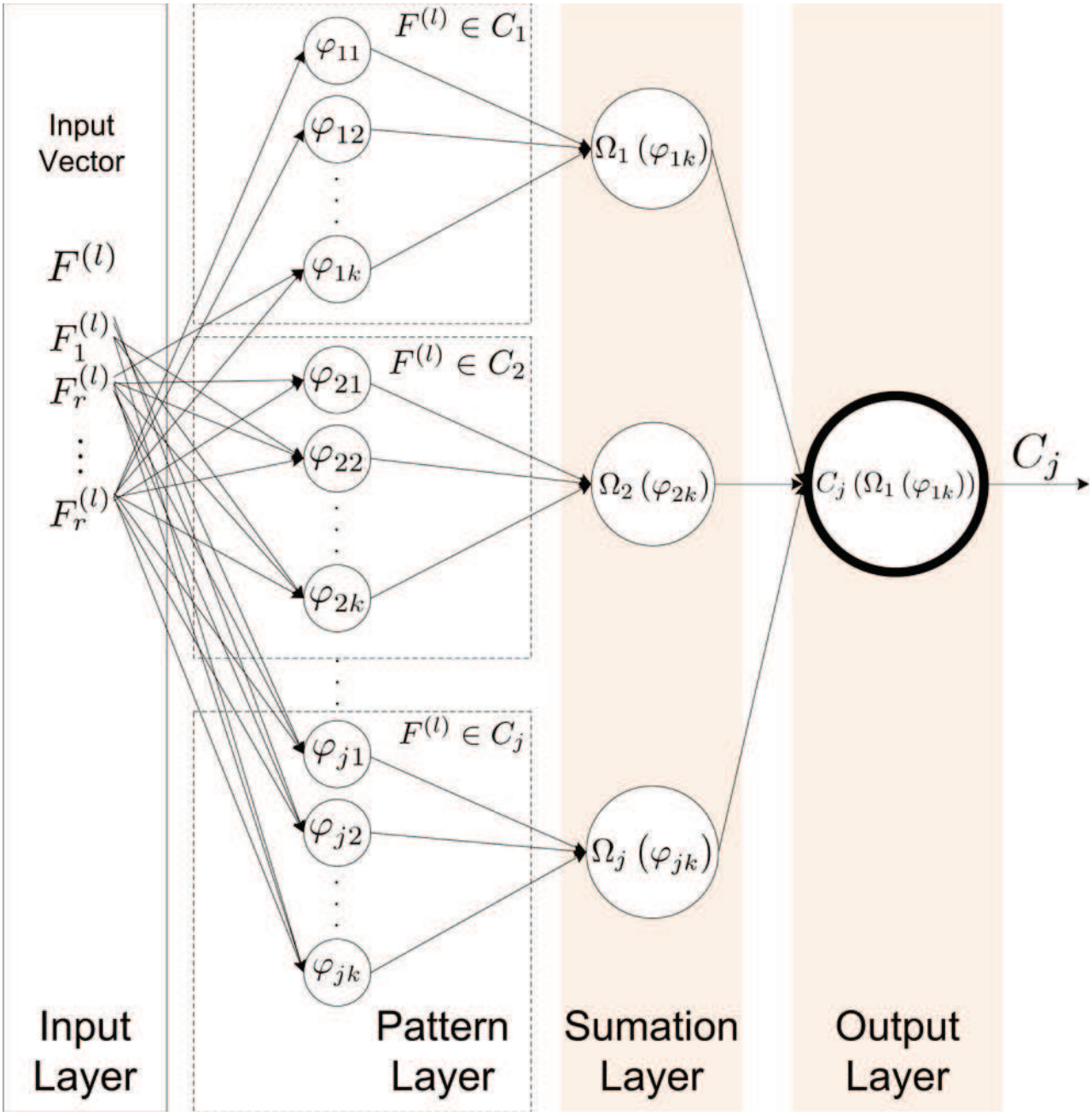

Figure 40. Diagram of the architecture of a Probabilistic Neural Network.

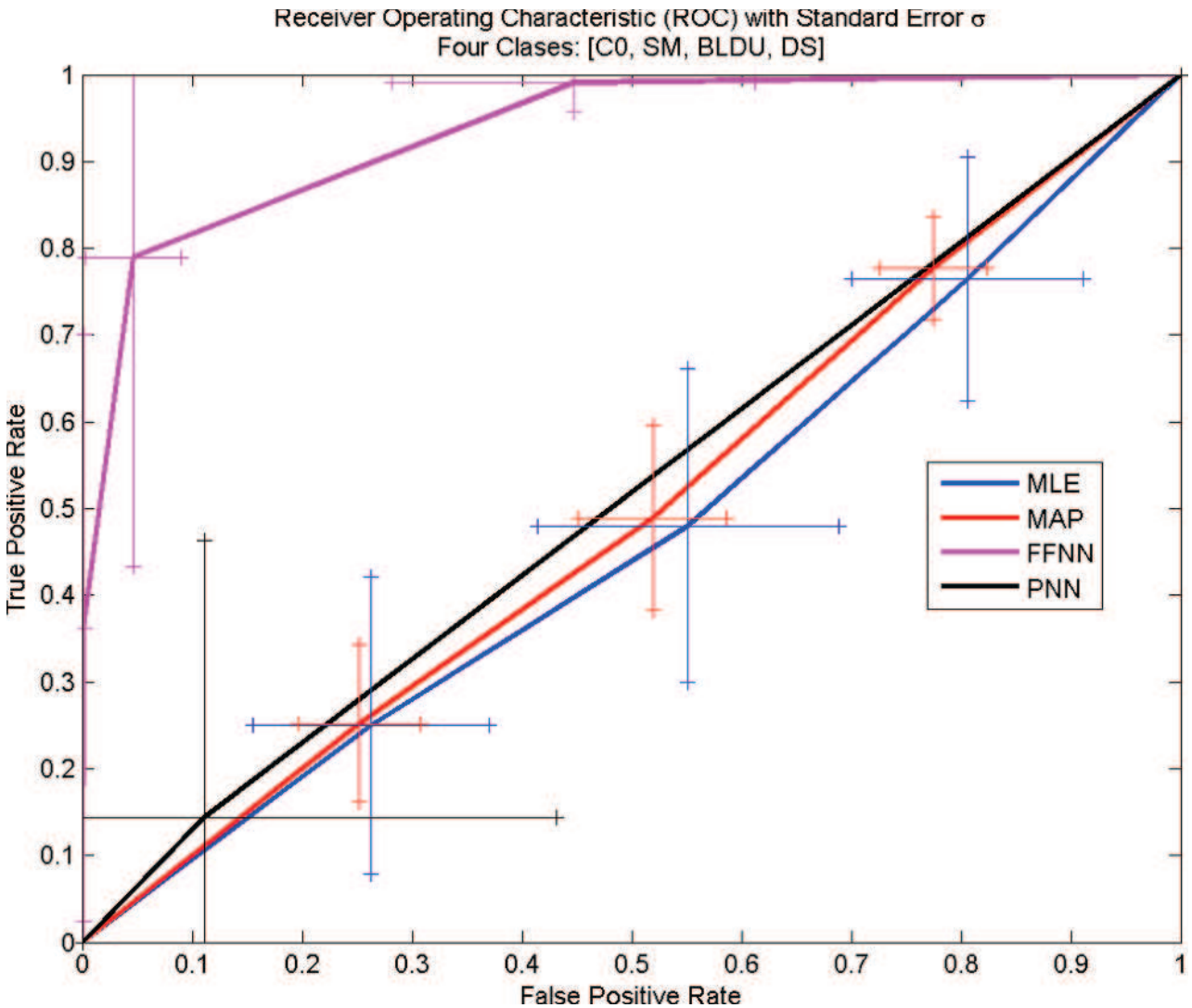

Figure 41. Averaged ROC's for the feature vector $F^{(1)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

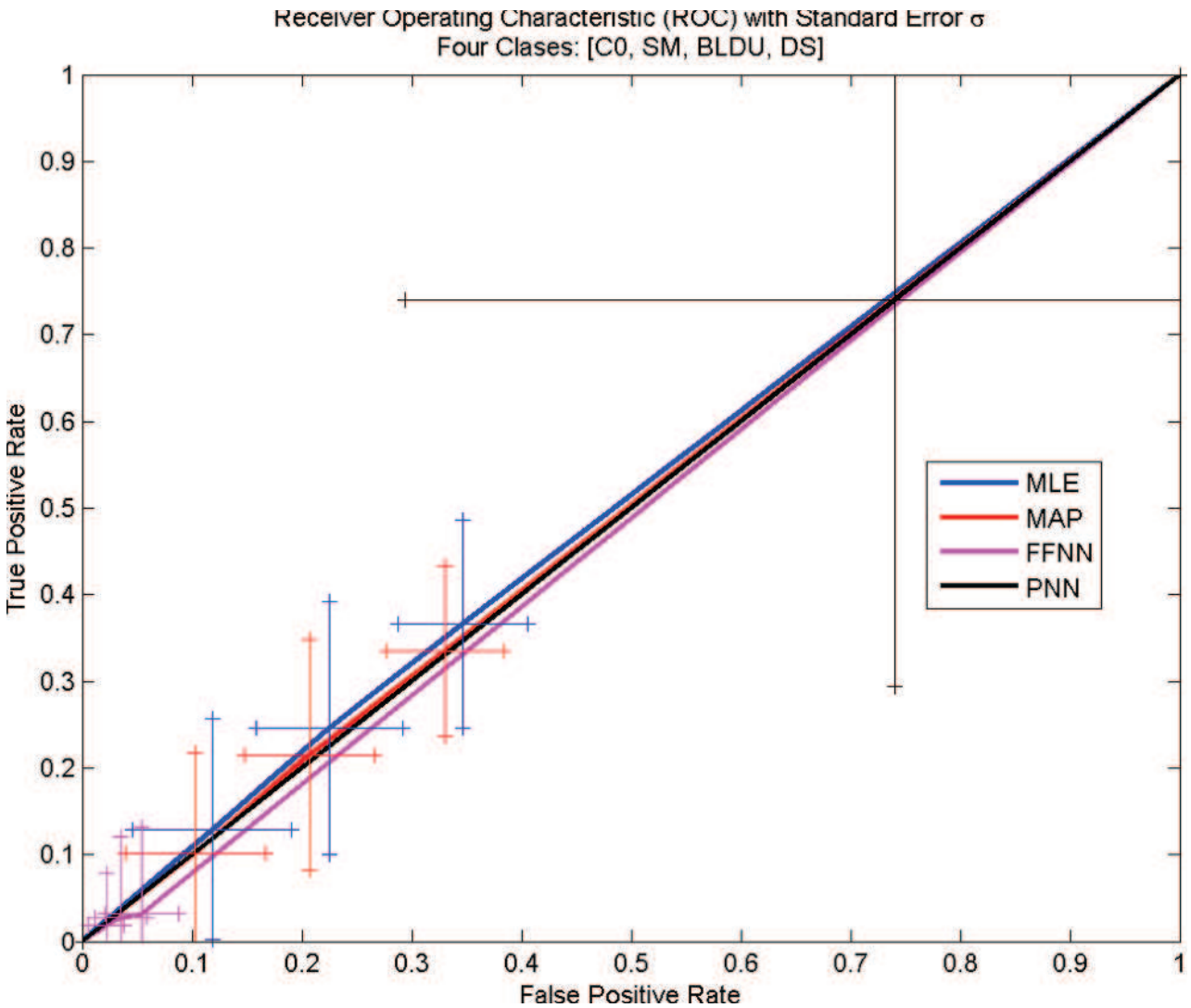

Figure 42. Averaged ROC's for the feature vector $F^{(2)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

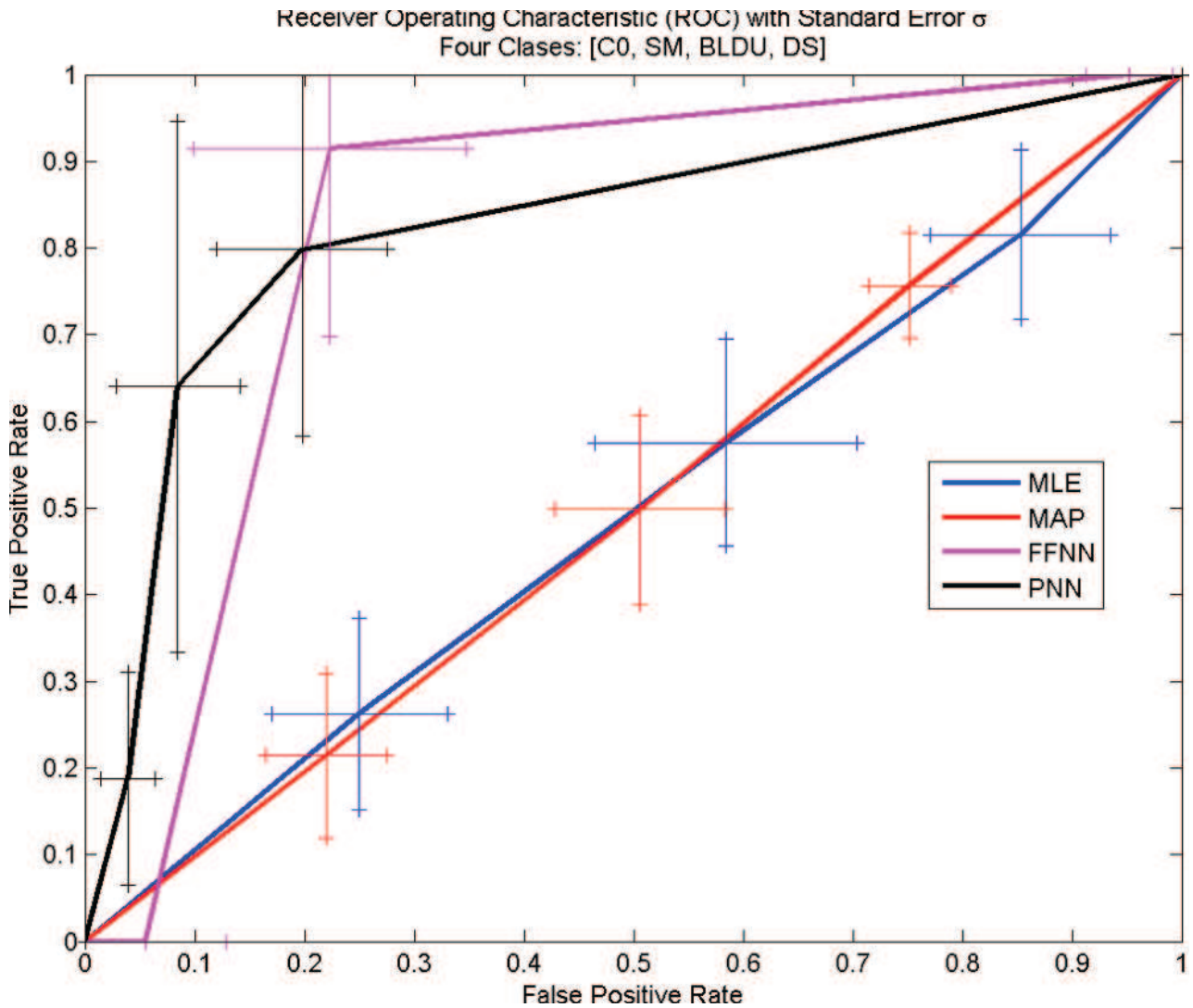

Figure 43. Averaged ROC's for the feature vector $F^{(3)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

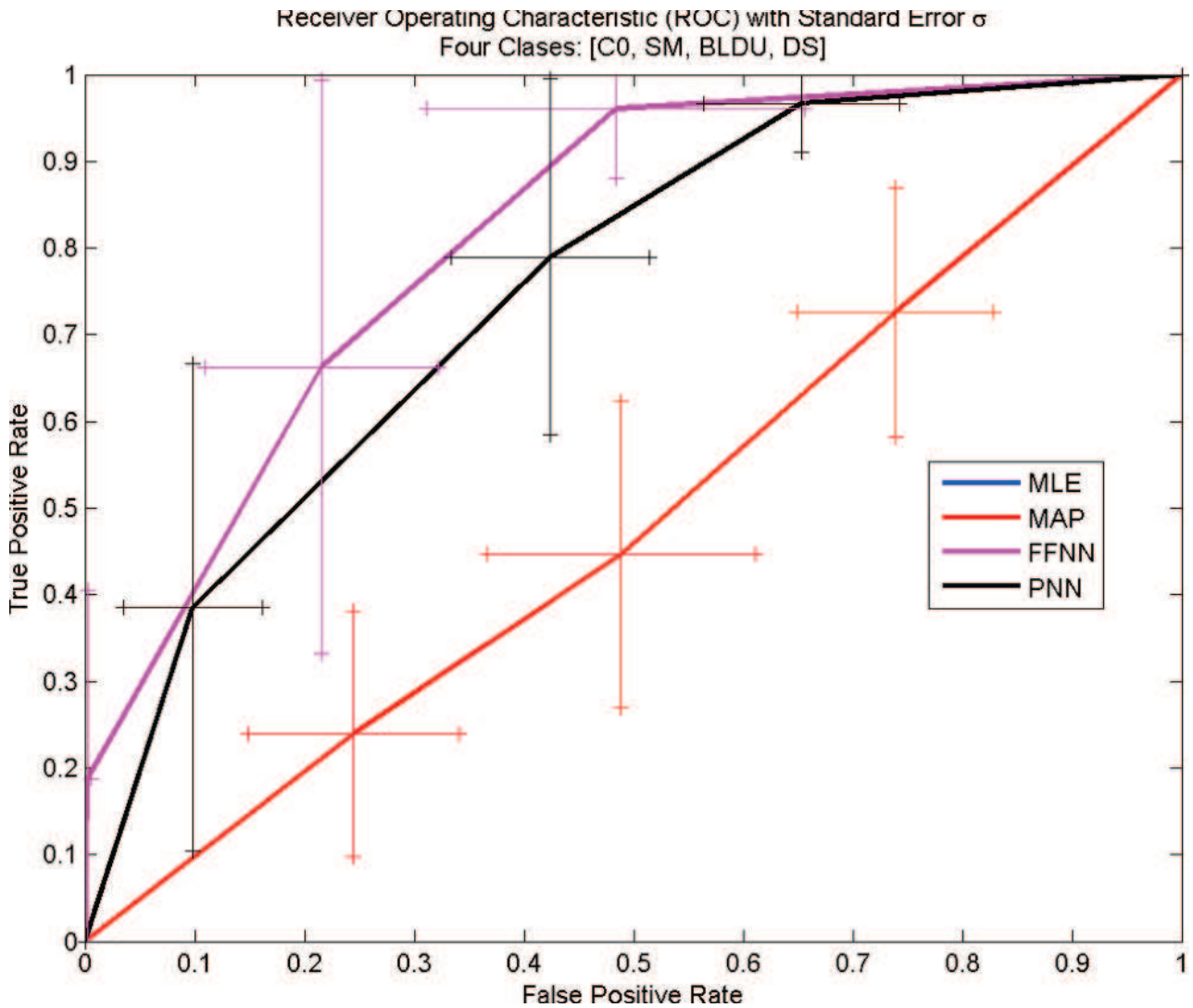

Figure 44. Averaged ROC's for the feature vector $F^{(4)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

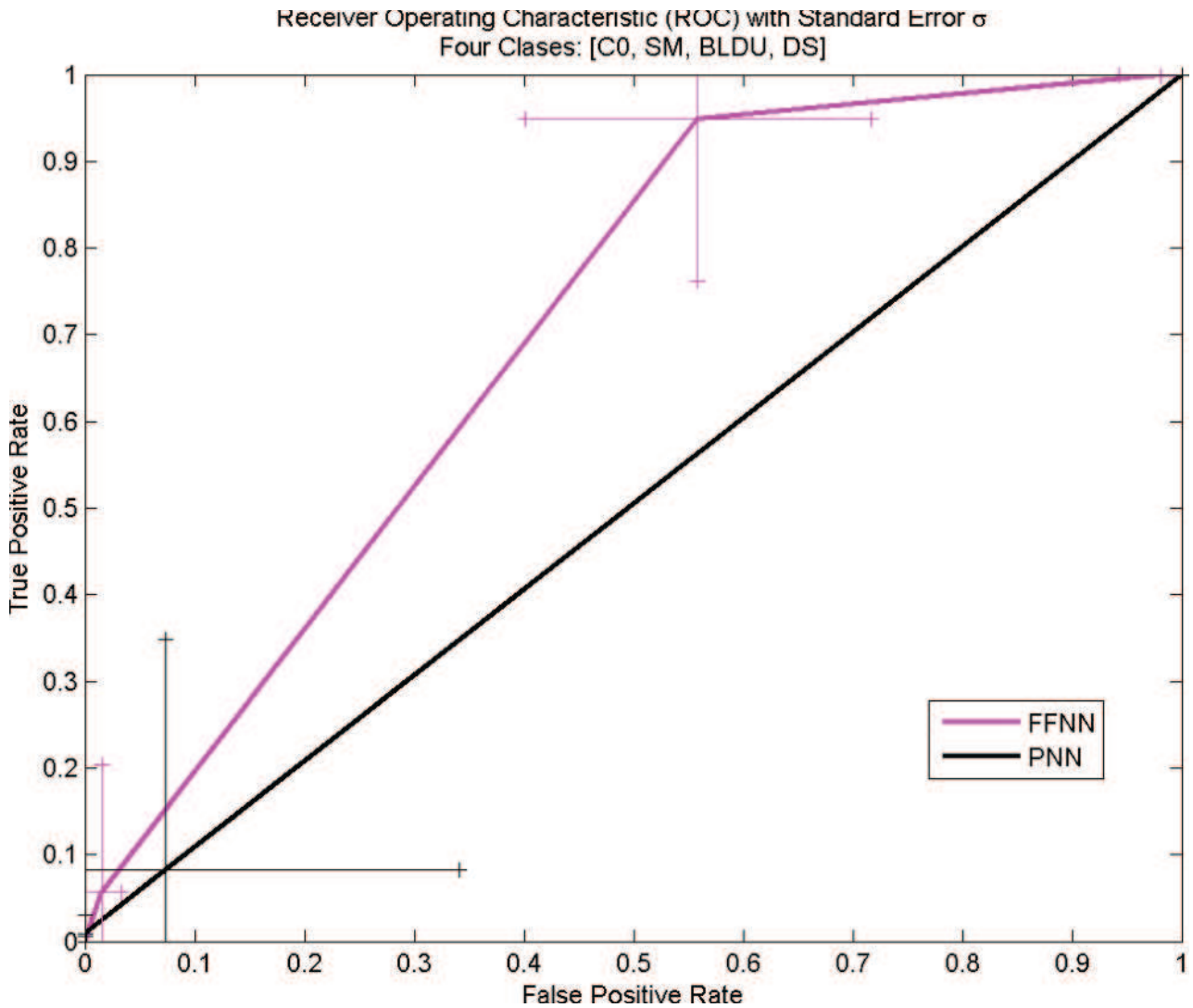

Figure 45. Averaged ROC's for the feature vector $F^{(5)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

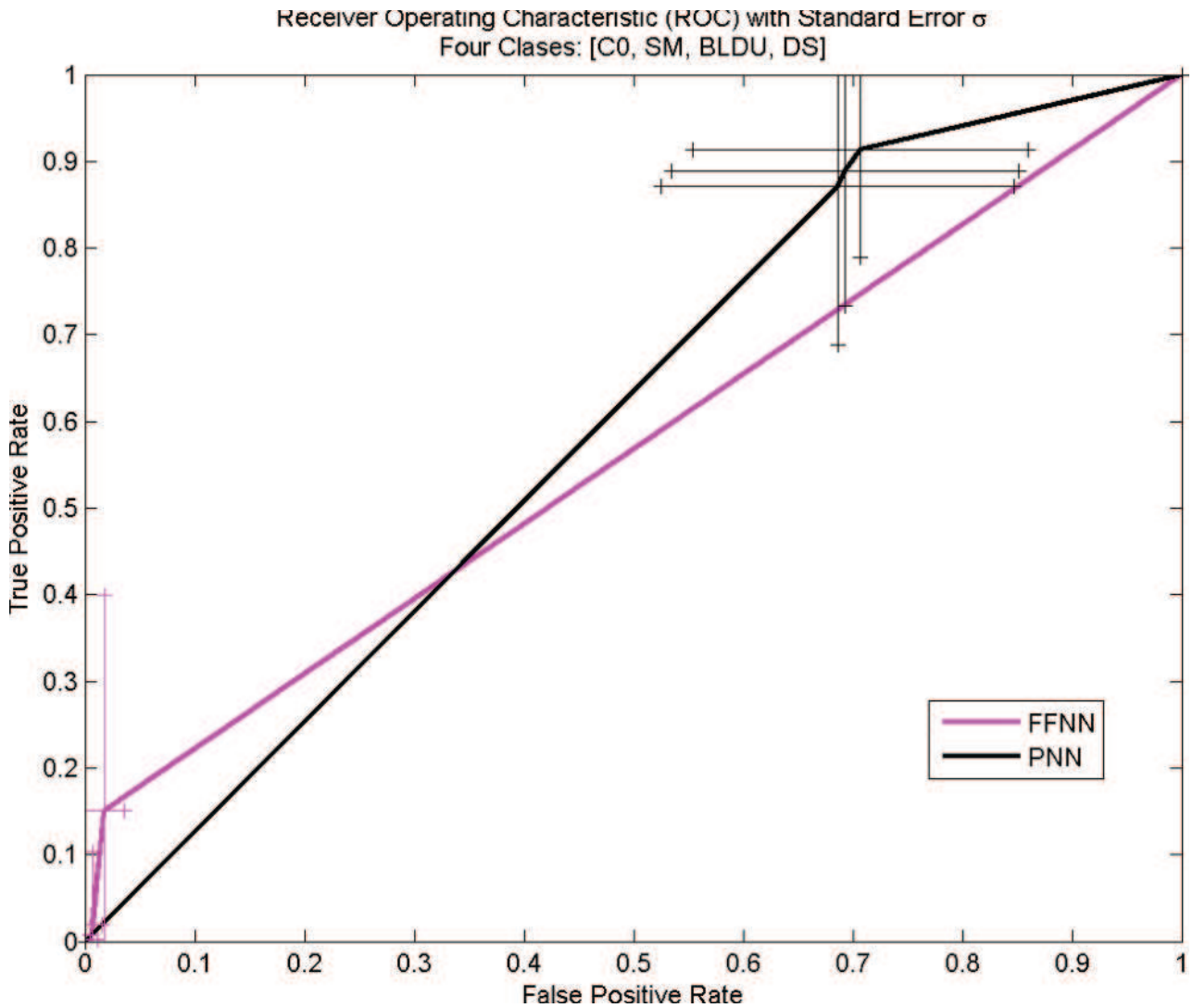

Figure 46. Averaged ROC's for the feature vector $F^{(6)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

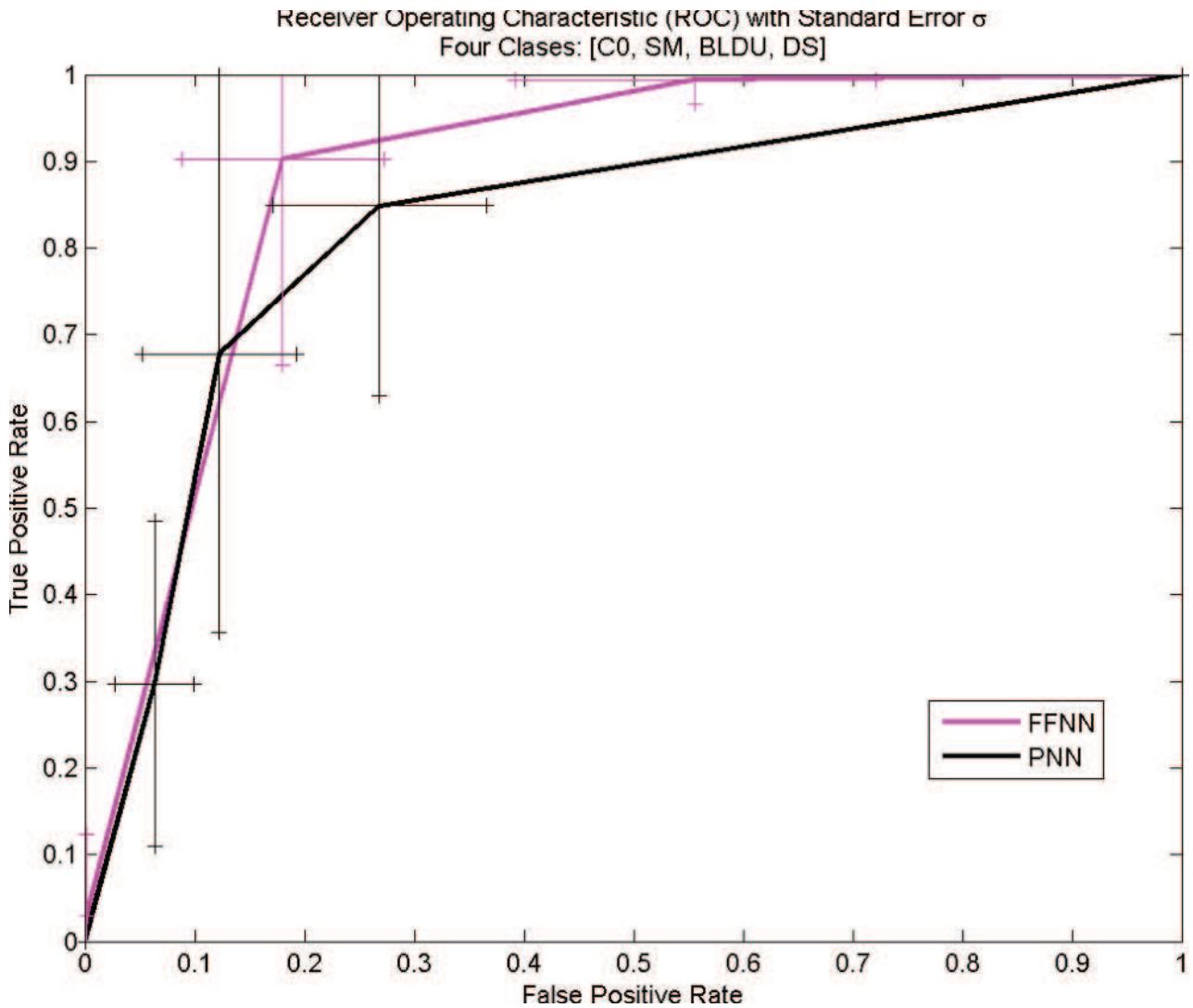

Figure 47. Averaged ROC's for the feature vector $F^{(7)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

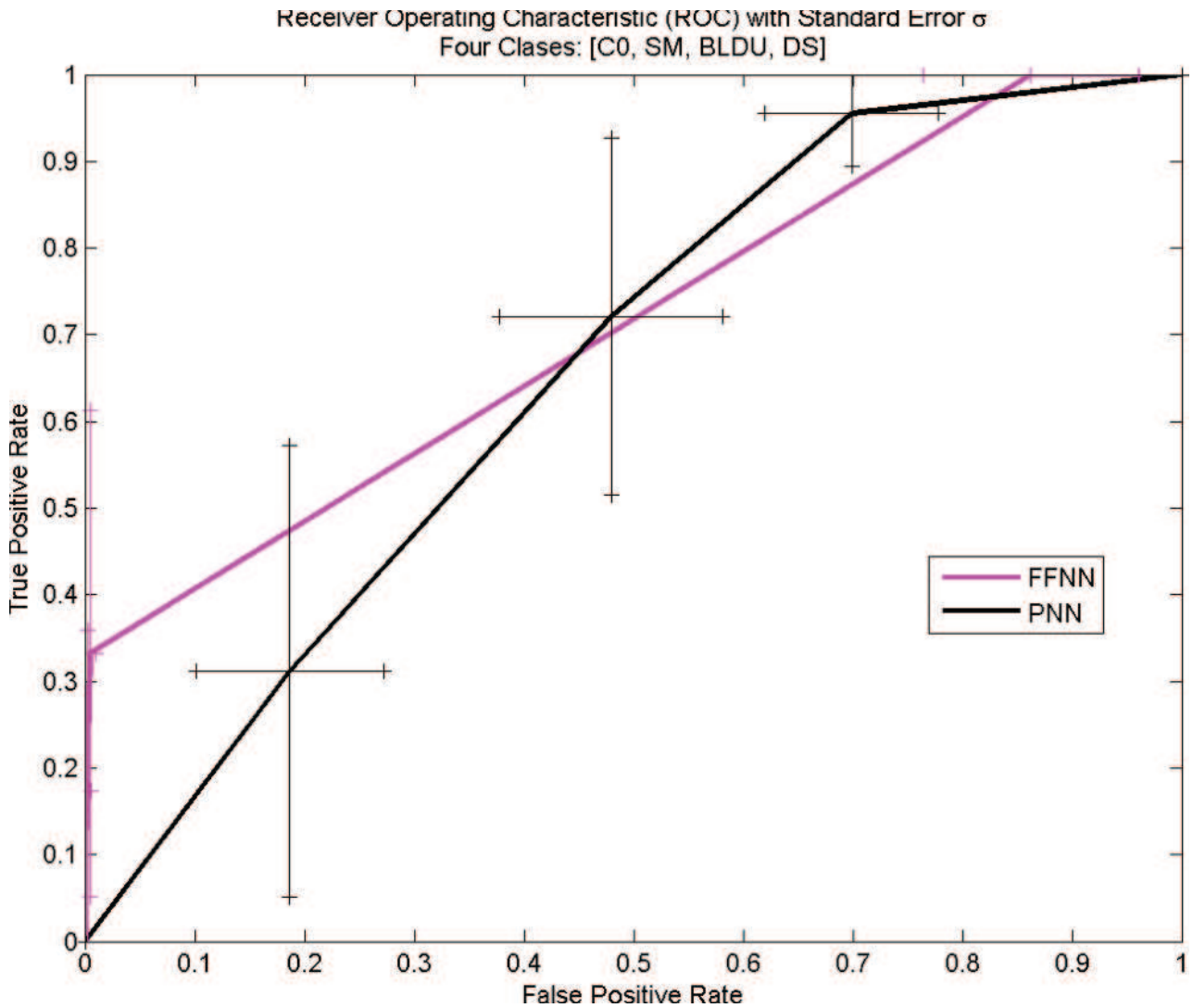

Figure 48. Averaged ROC's for the feature vector $F^{(8)}$. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

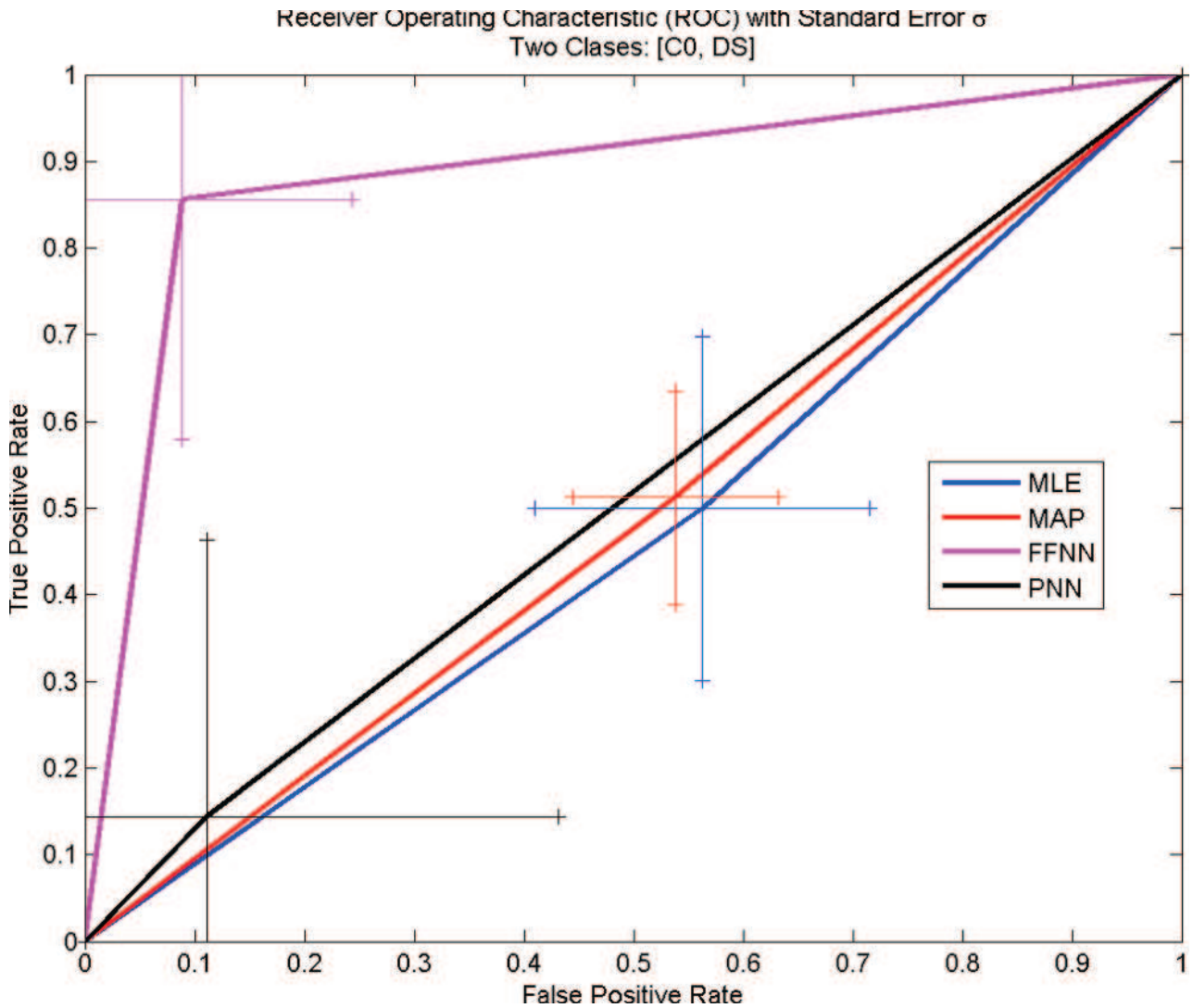

Figure 49. Averaged ROC's for the feature vector $F^{(1)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

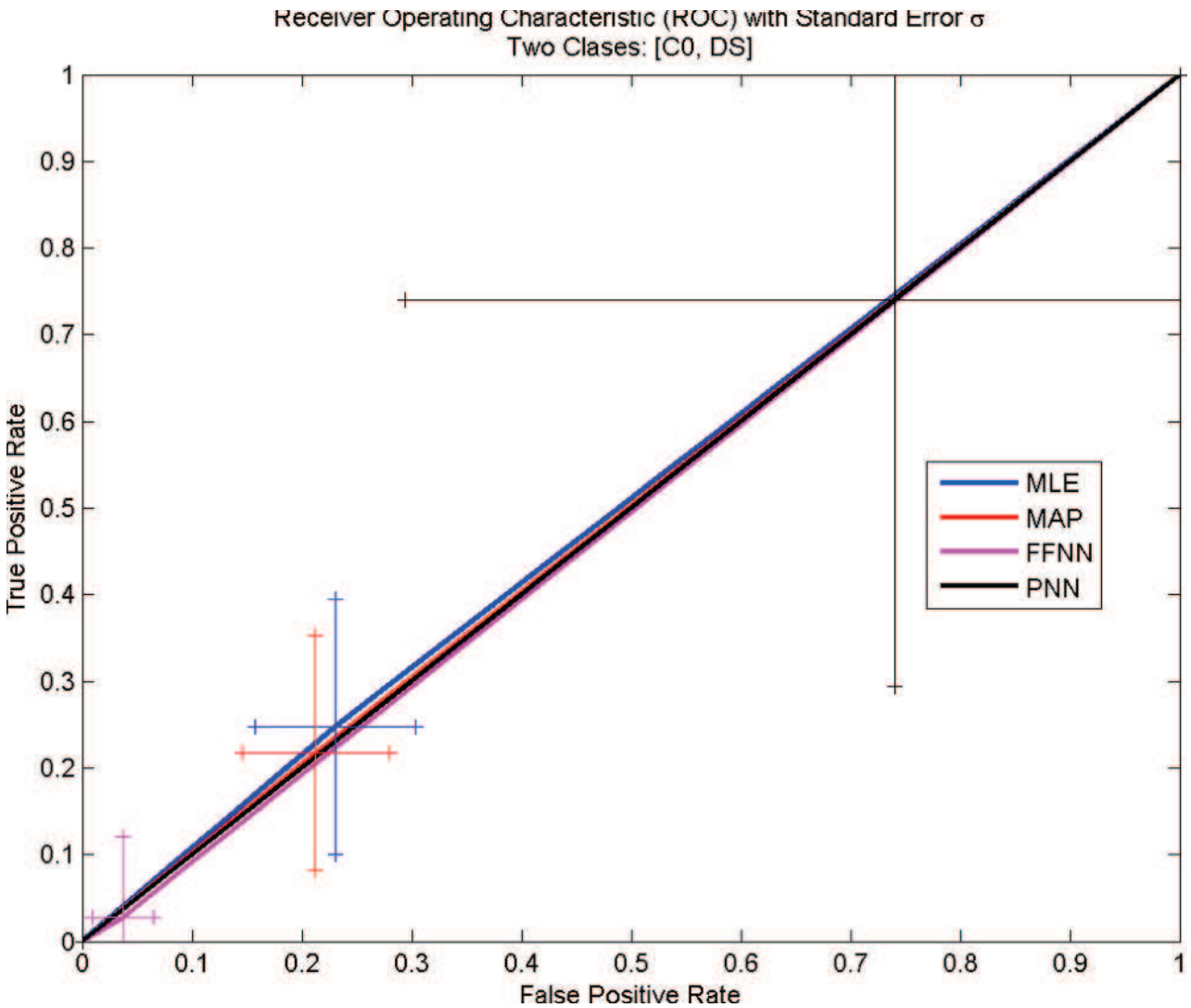

Figure 50. Averaged ROC's for the feature vector $F^{(2)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

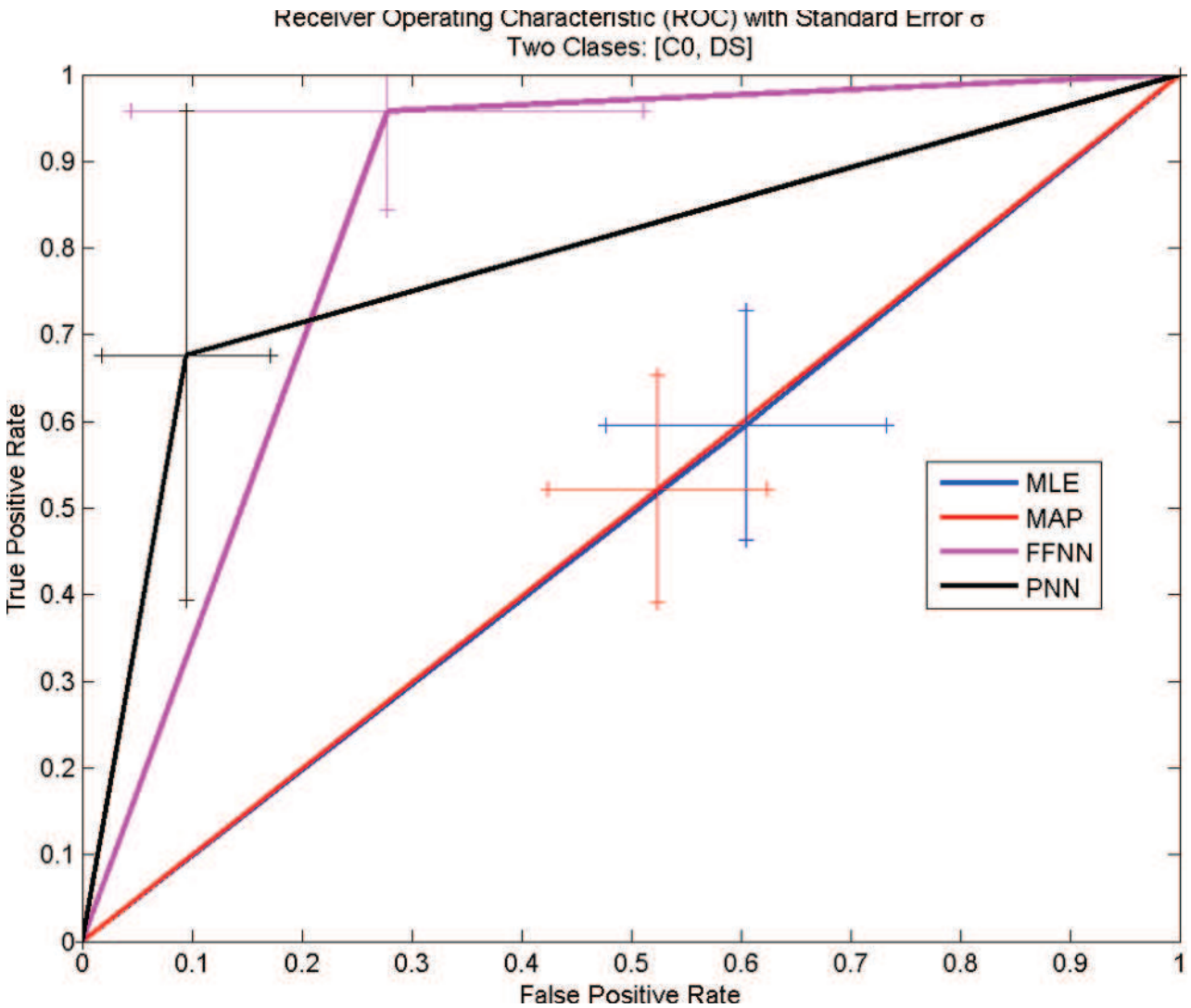

Figure 51. Averaged ROC's for the feature vector $F^{(3)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

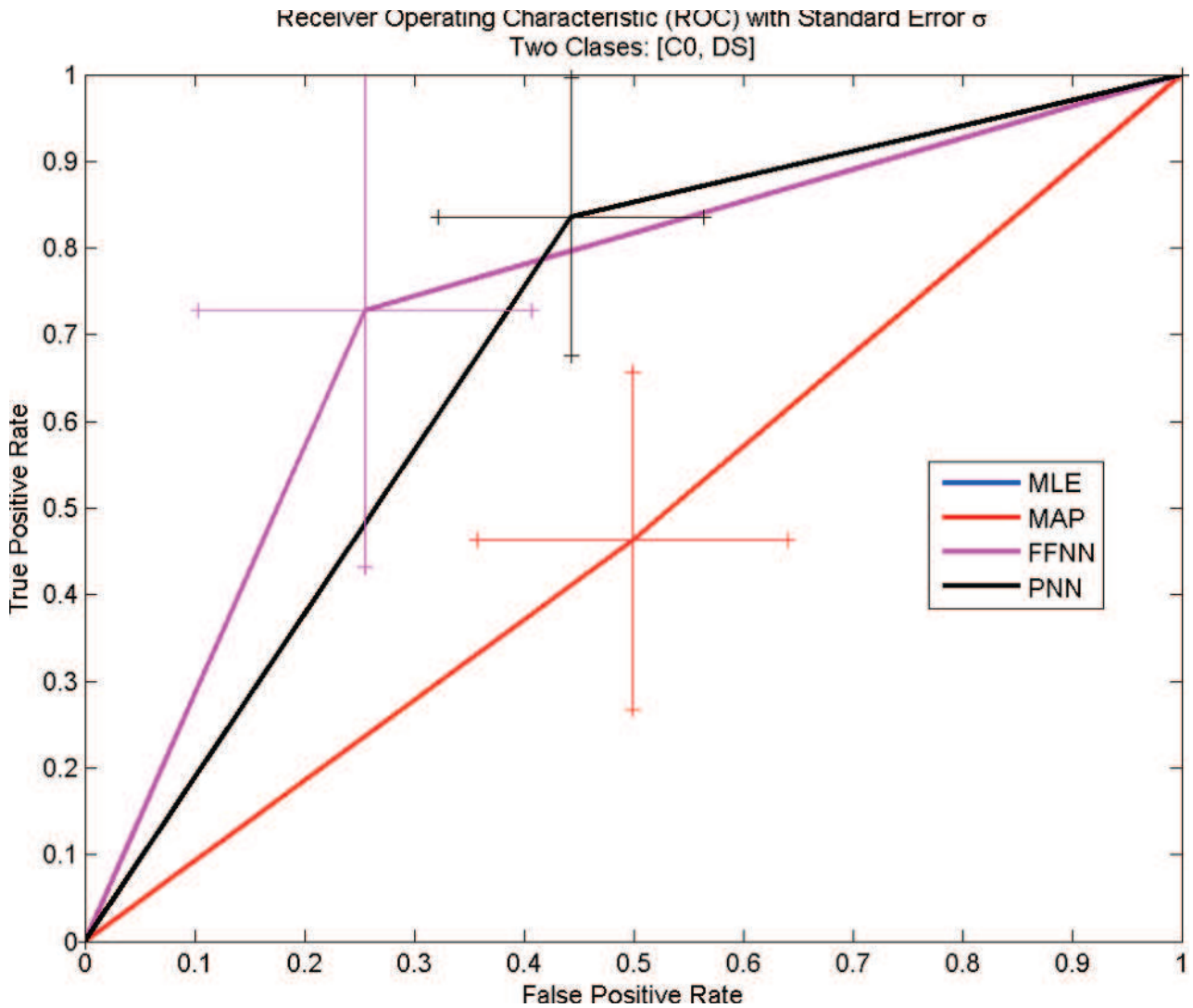

Figure 52. Averaged ROC's for the feature vector $F^{(4)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

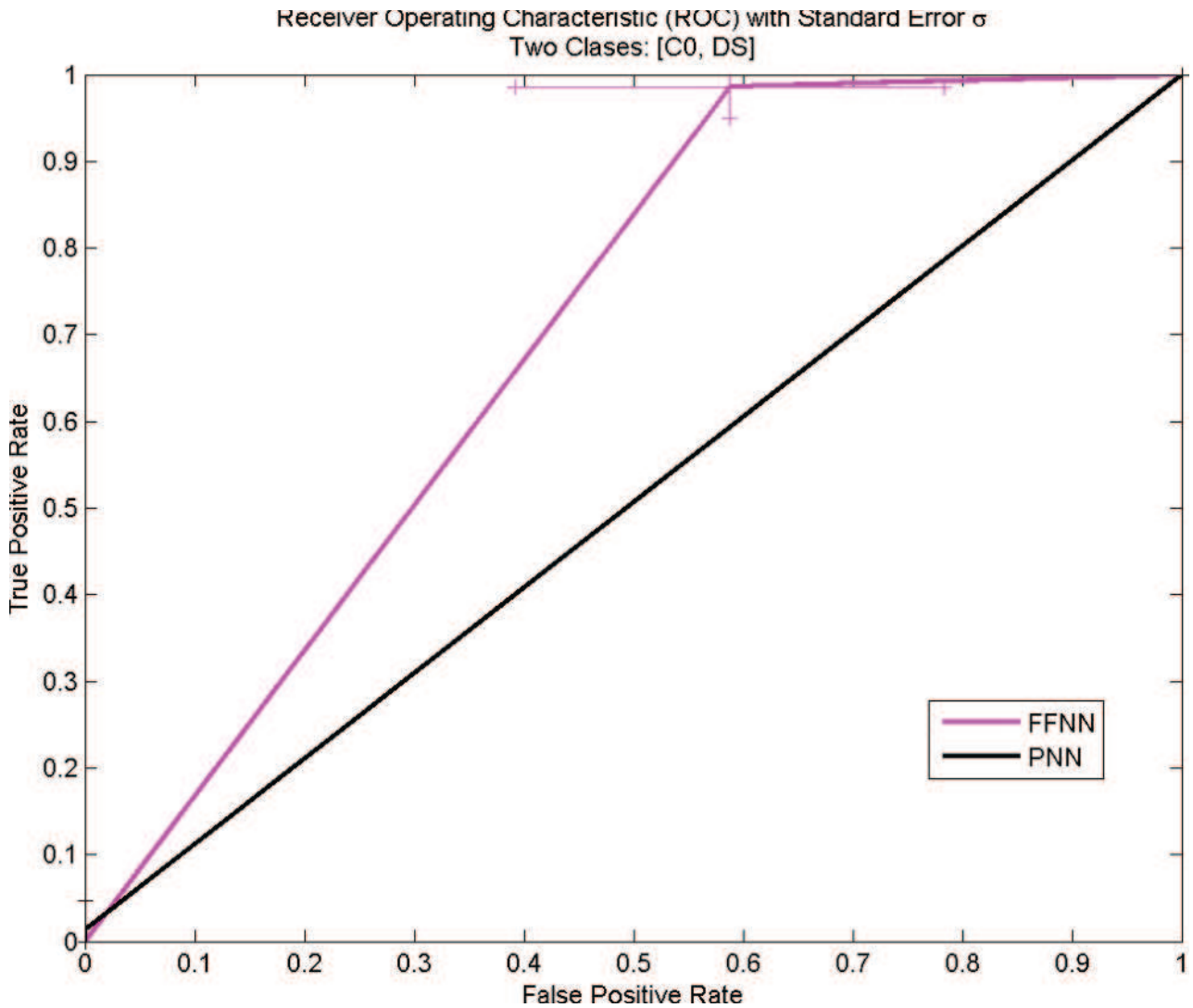

Figure 53. Averaged ROC's for the feature vector $F^{(5)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

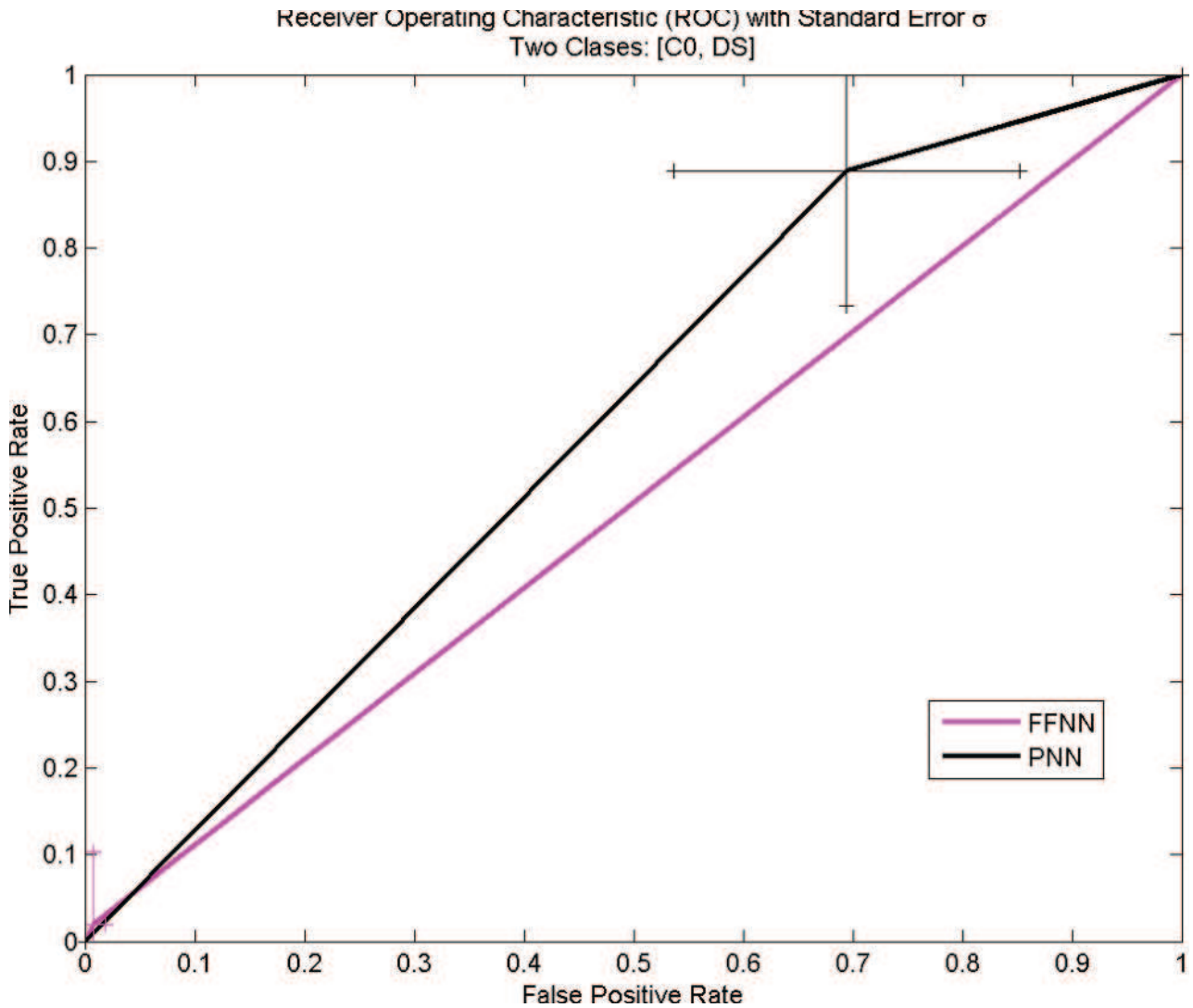

Figure 54. Averaged ROC's for the feature vector $F^{(6)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

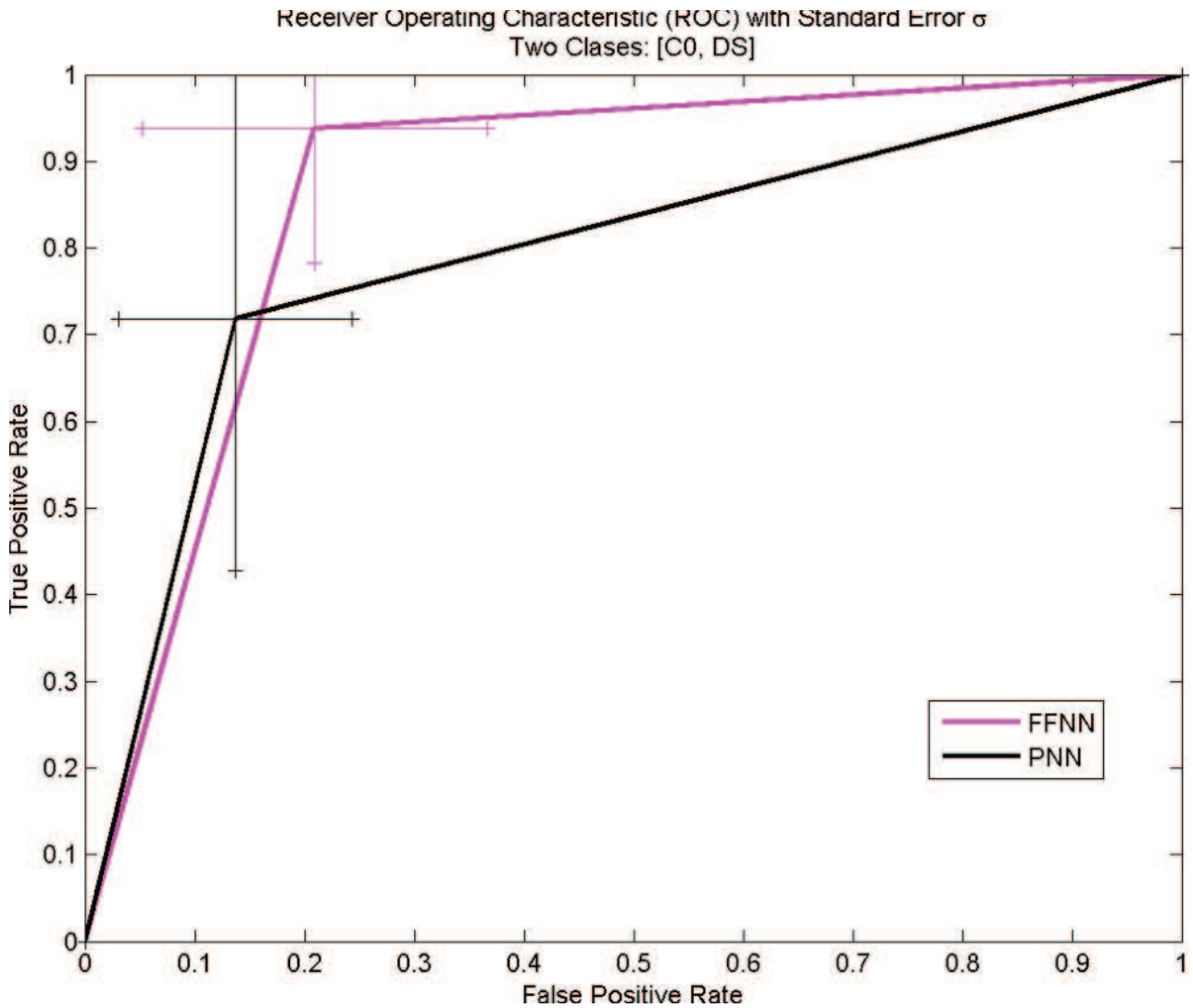

Figure 55. Averaged ROC's for the feature vector $F^{(7)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

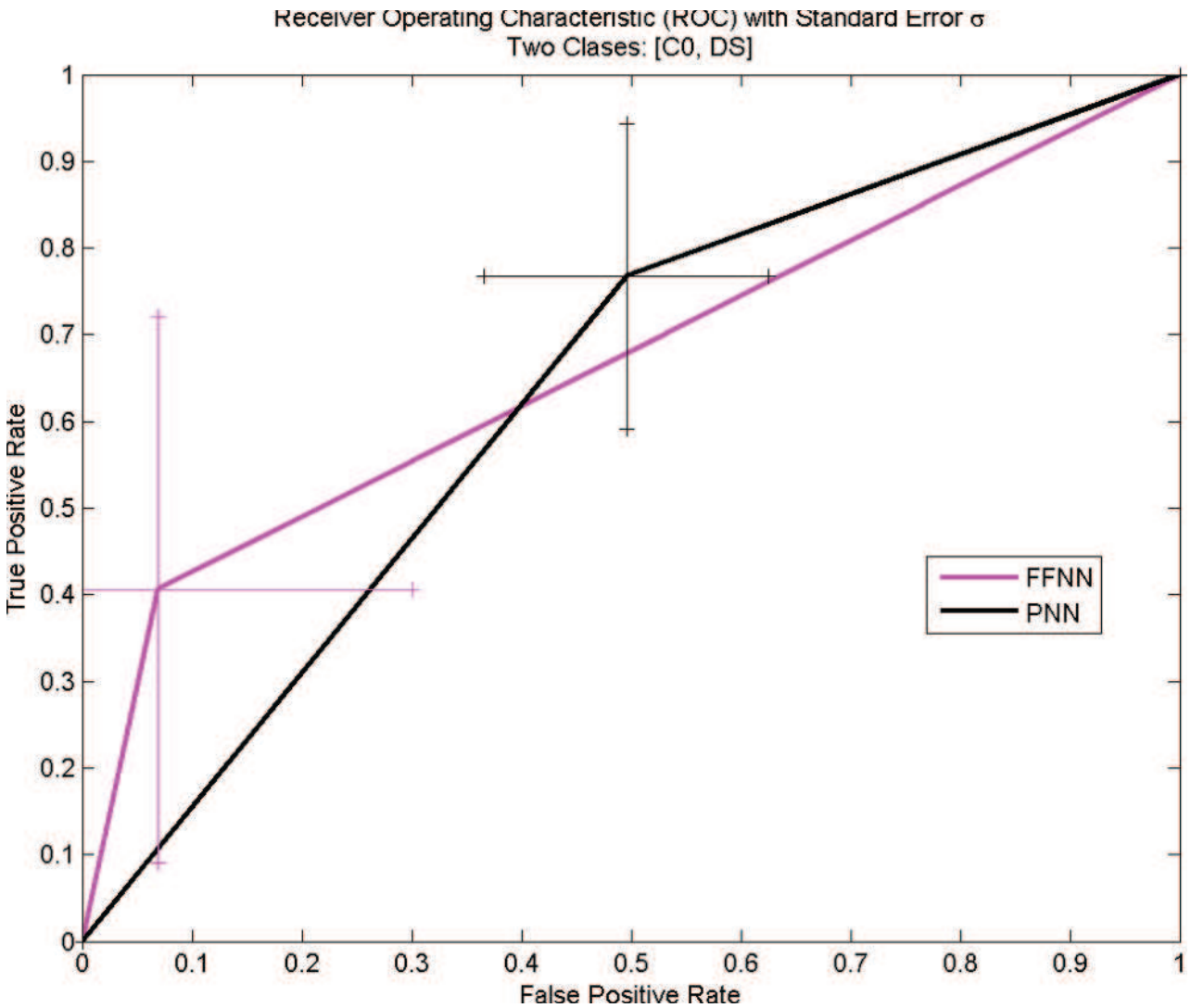

Figure 56. Averaged ROC's for the feature vector $F^{(8)}$ simplifying the problem into two classes. This plot show a comparison between the four classification methods. The standard deviation is included to see the stability associated.

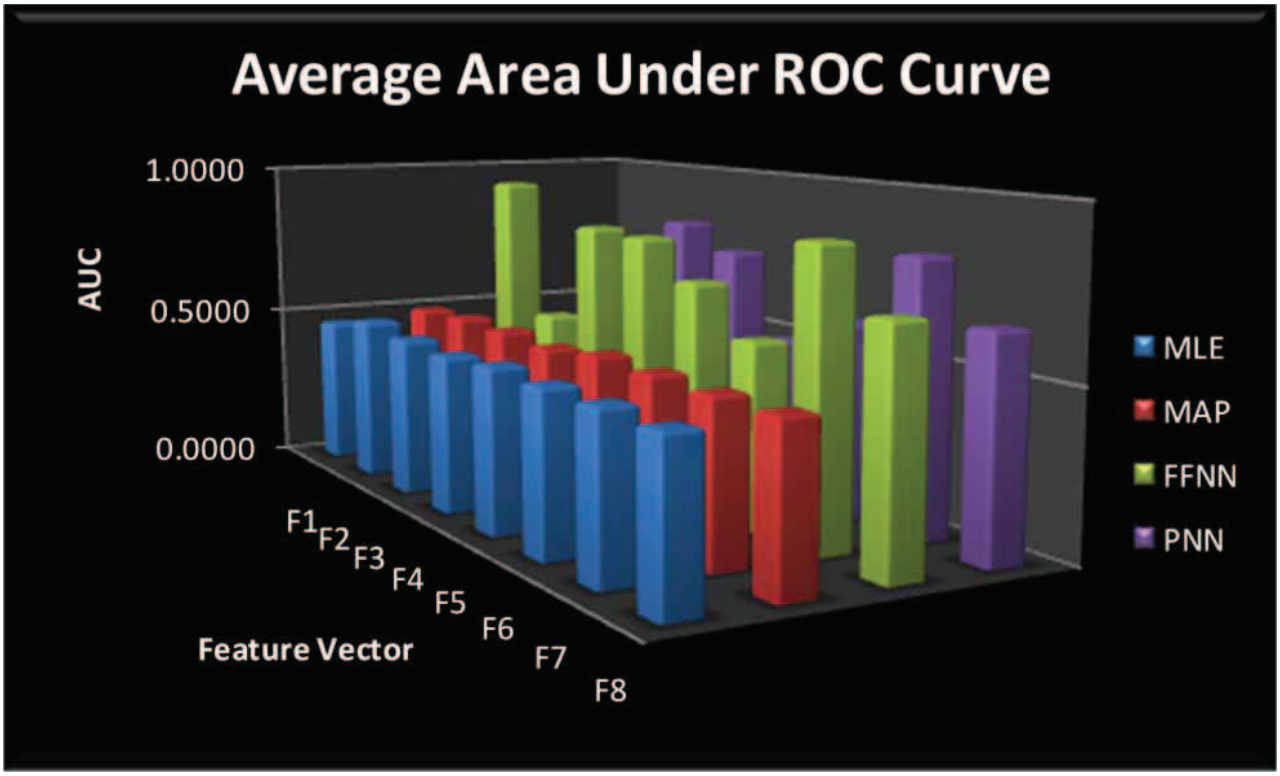

Figure 57. Averaged total area under the ROC curve. Clearly, FFNN performs better than the others.

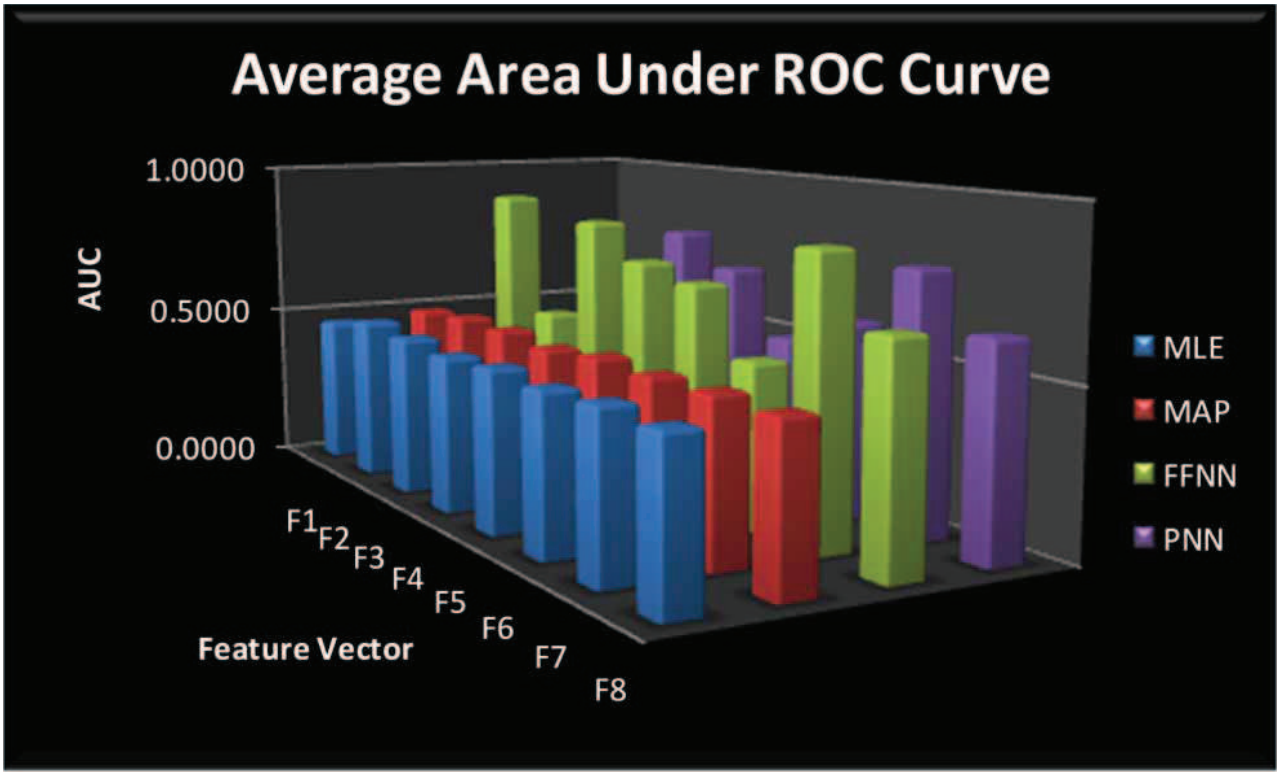

Figure 58. Averaged total area under the ROC curve, for the classification of two classes only. Clearly, FFNN performs better than the others.

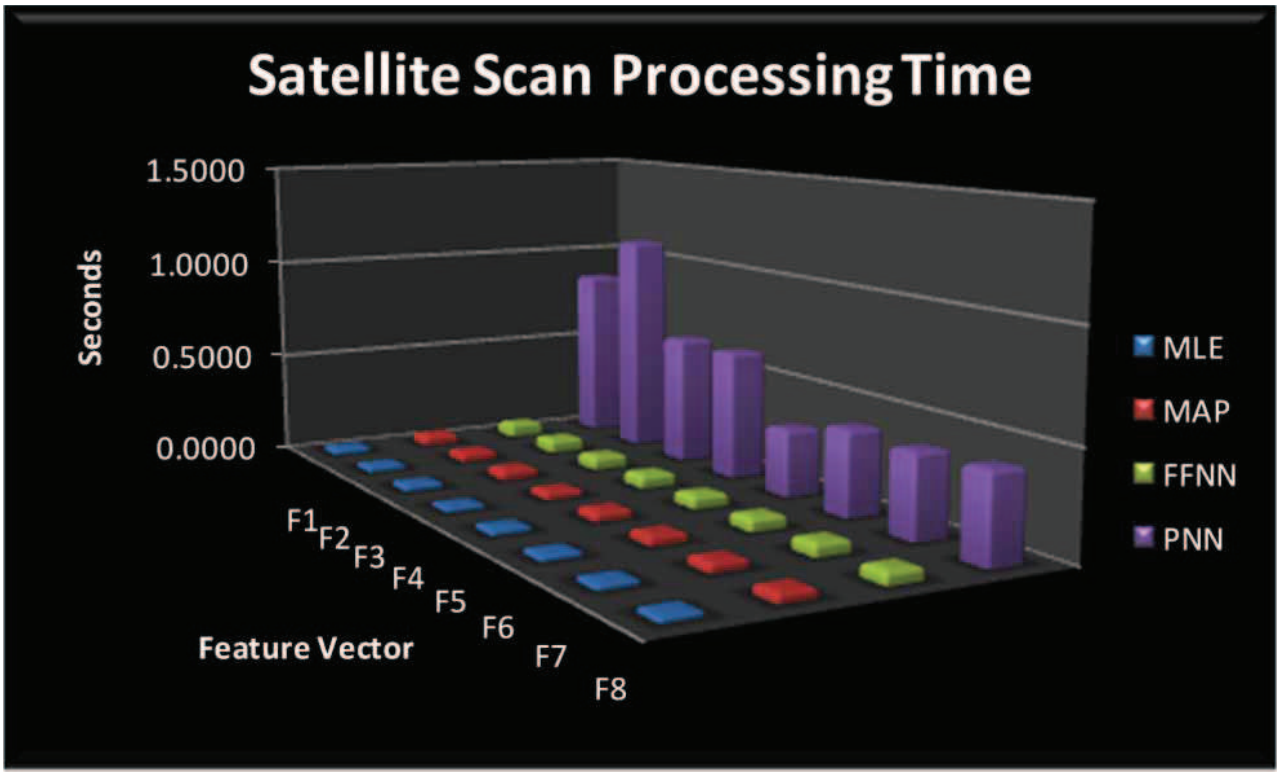

Figure 59. Processing time required to produce an output givem a satellite scan (in the case of MODIS terra multiespectral bands, a scan will be of size $[5 \times 1354]$ in the case of $F^{(1)}$). It is evident and intuitive that the probabilistic methods produce faster outputs.

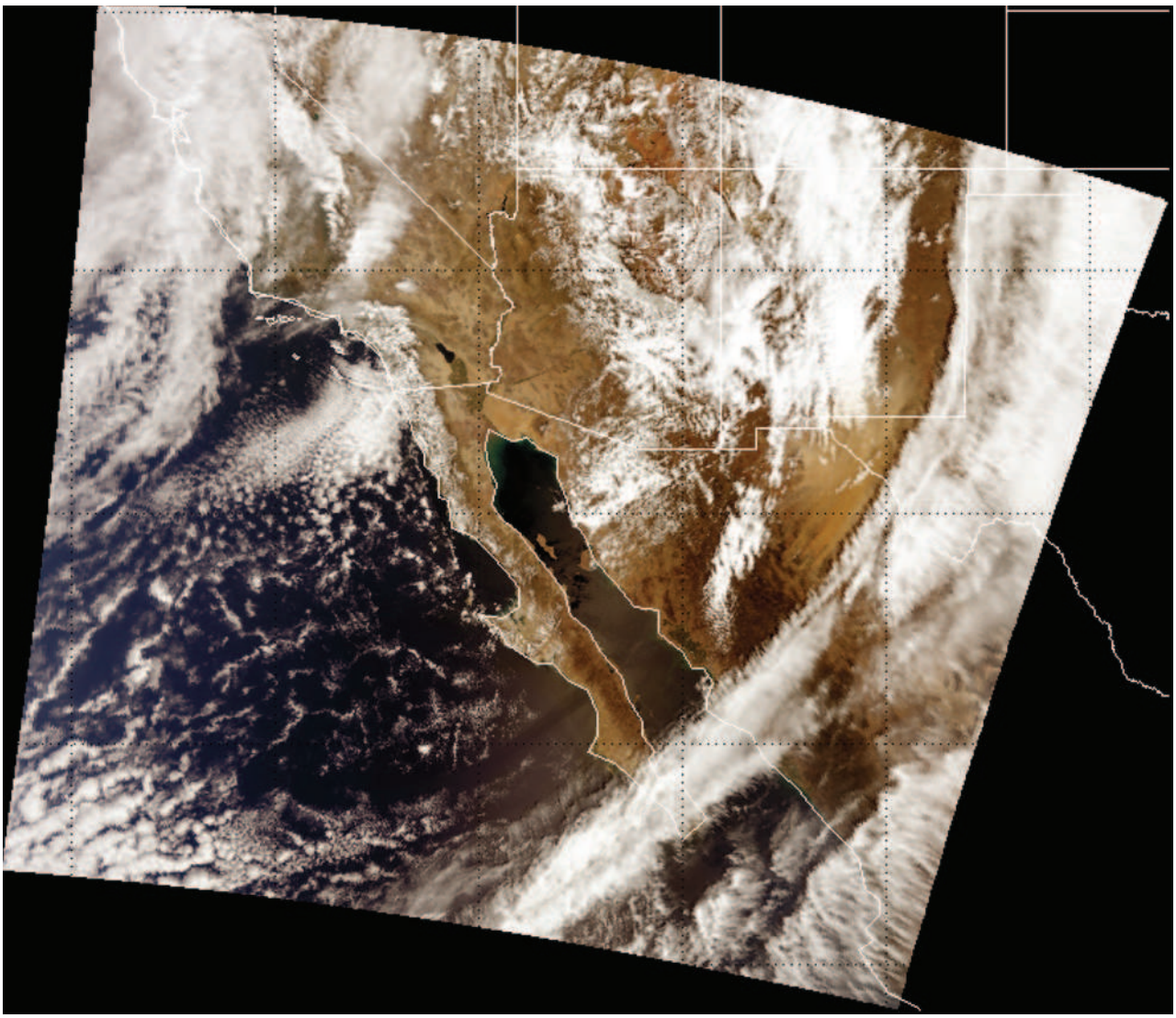

Figure 60. True color image for the event of 2001-096 at 18:30 UTC.

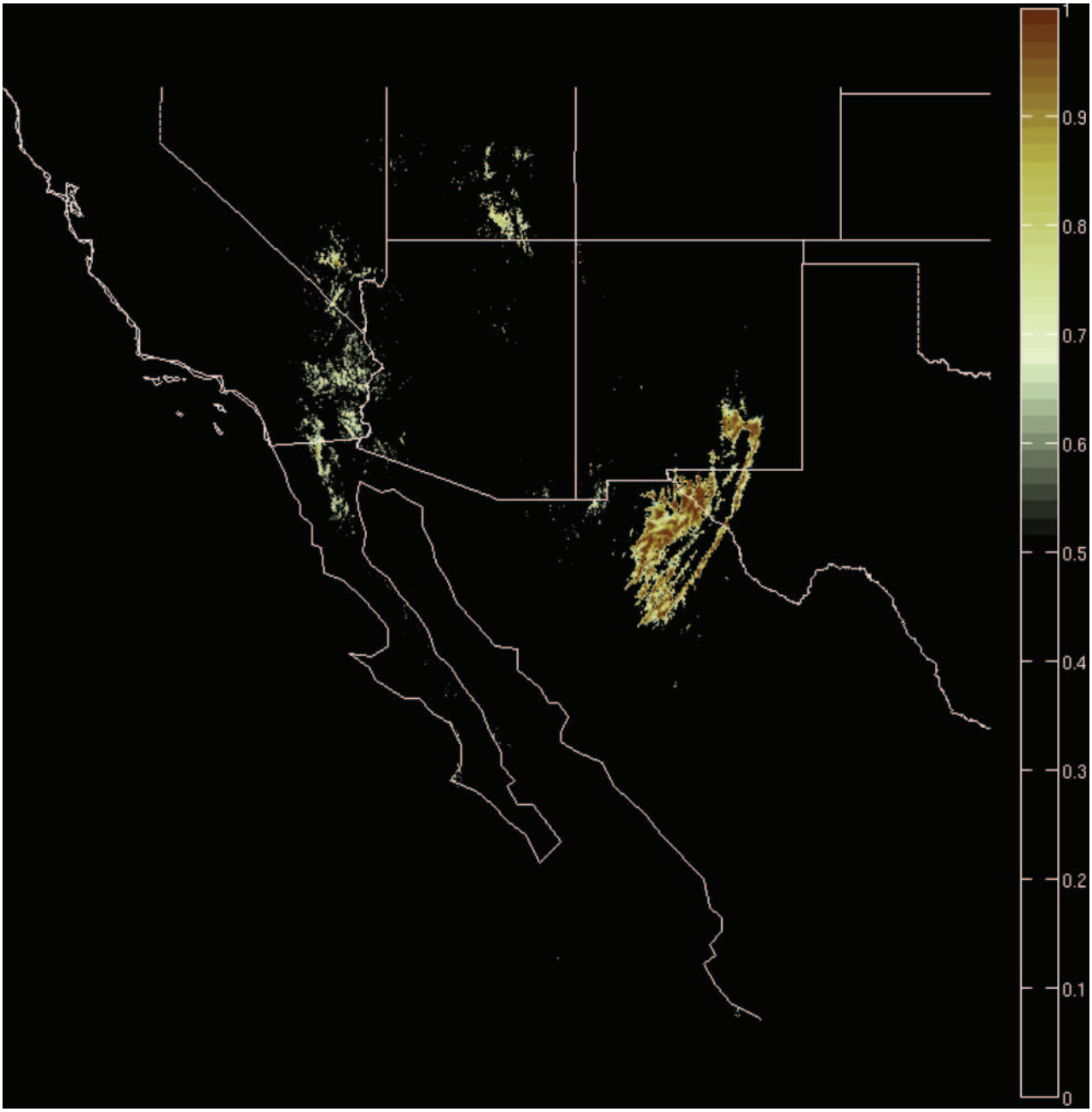

Figure 61. Result of the probabilistic display model (PD) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

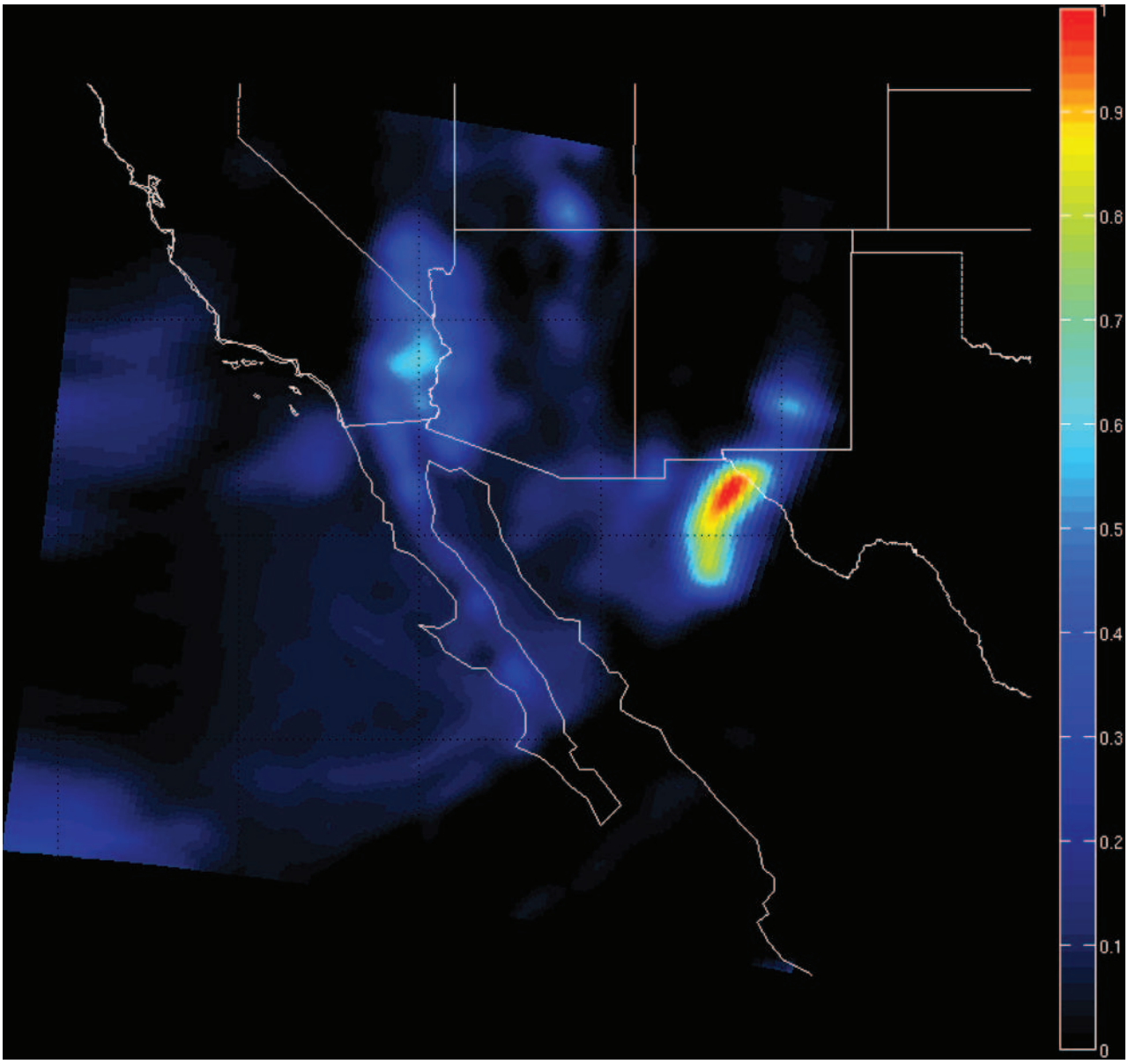

Figure 62. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

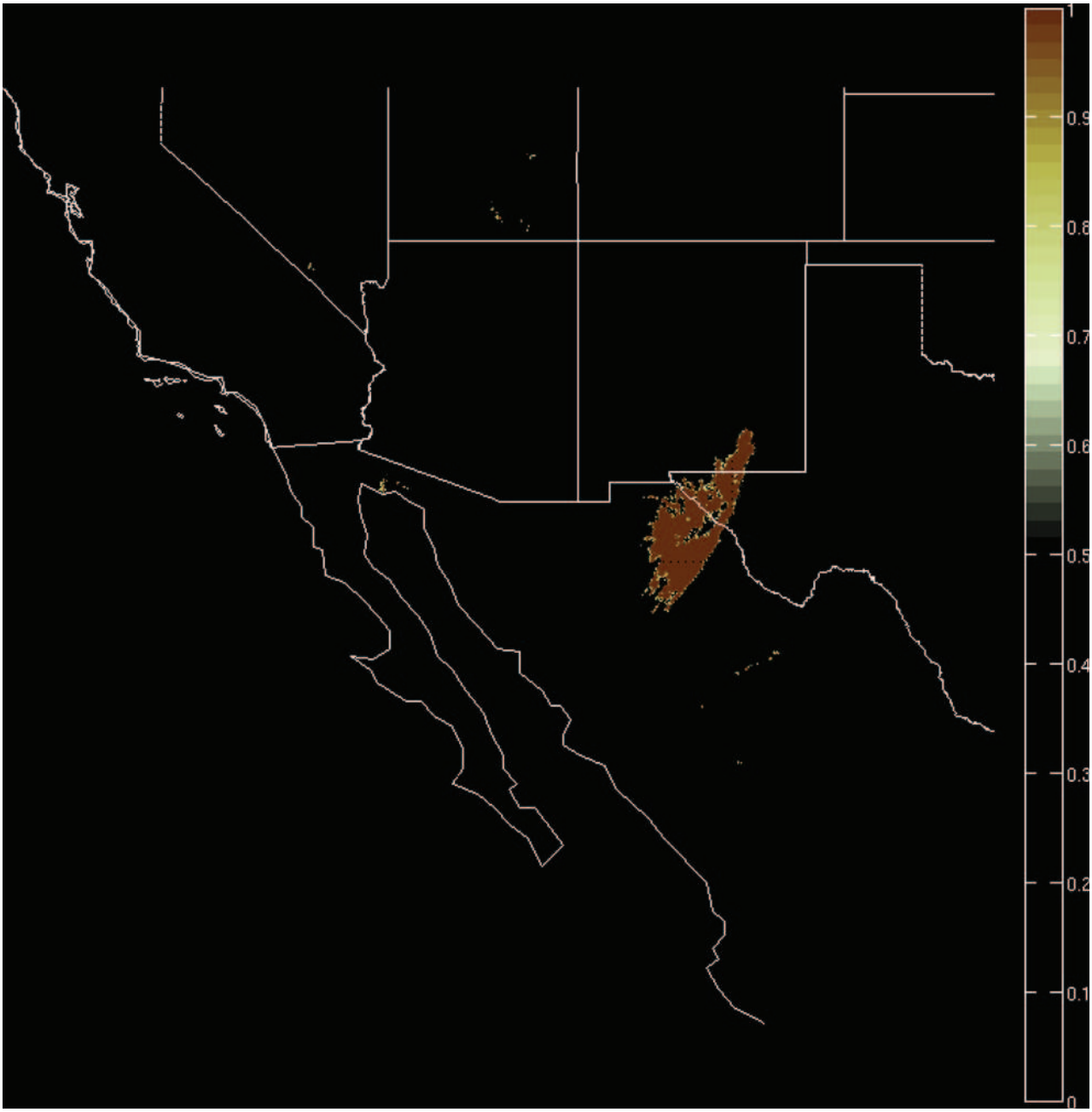

Figure 63. Result of the Maximum Likelihood model (MLE) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

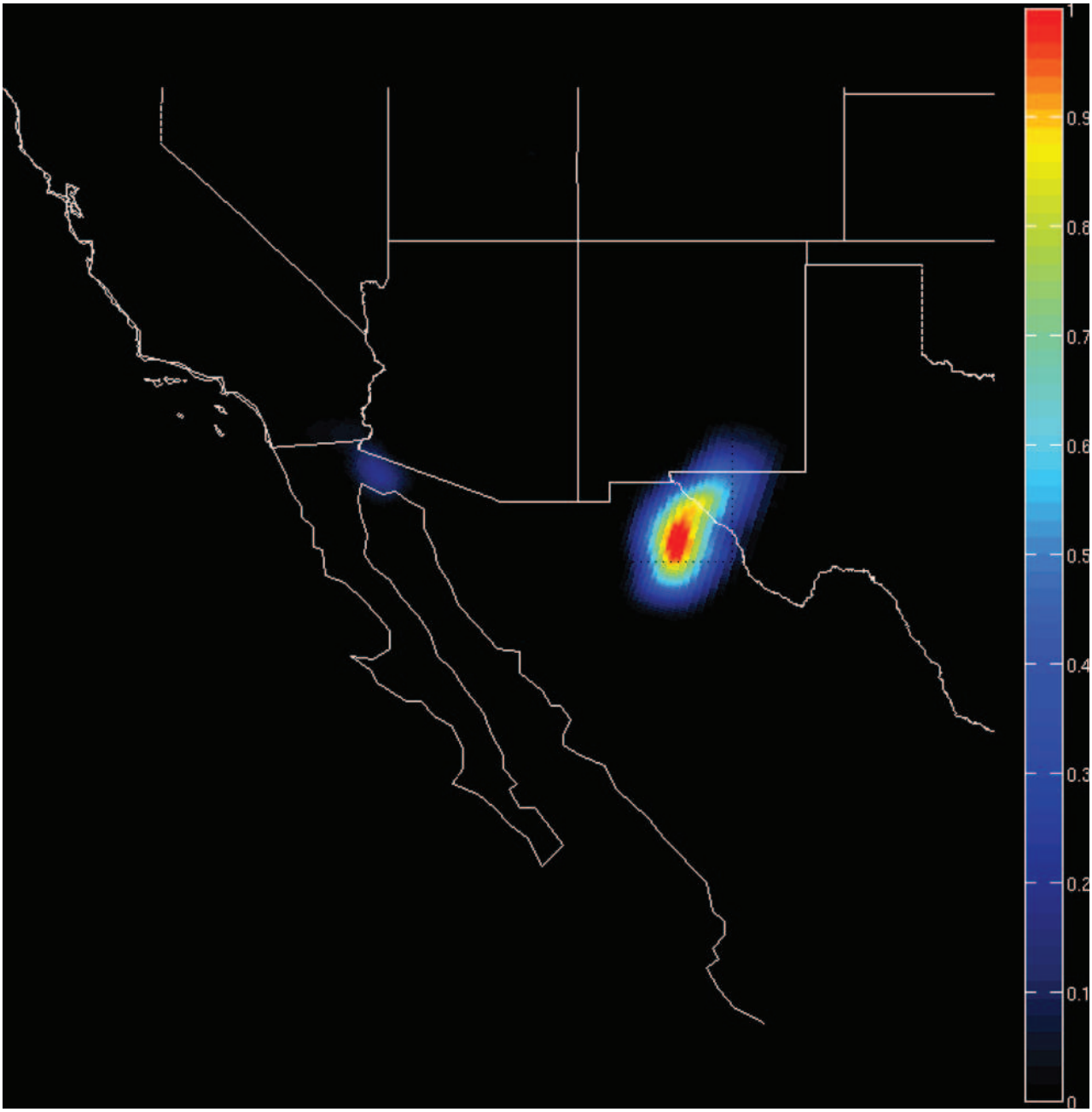

Figure 64. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

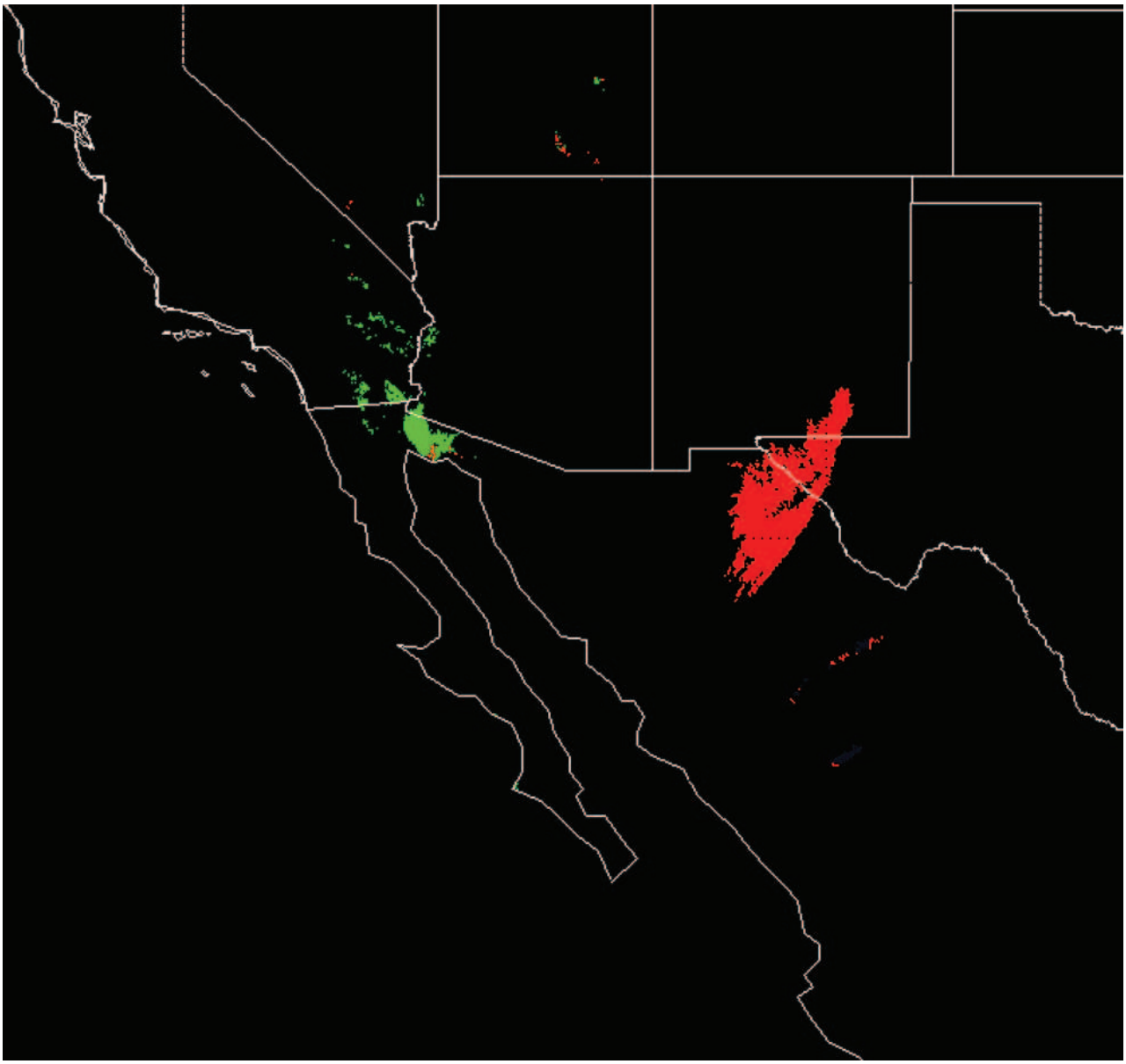

Figure 65. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2001-096 at 18:30 UTC.

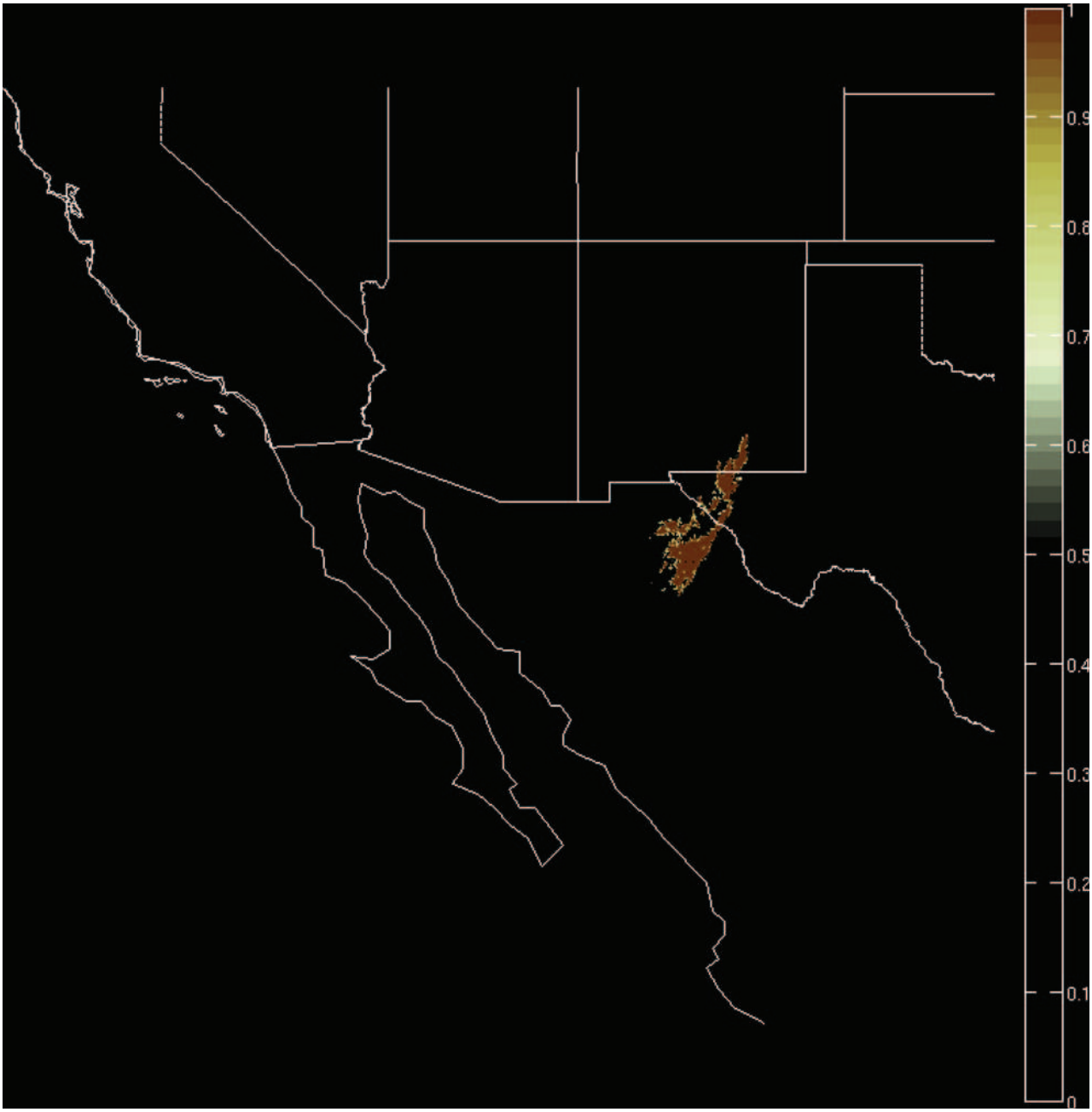

Figure 66. Result of the Maximum a posteriori model (MAP) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

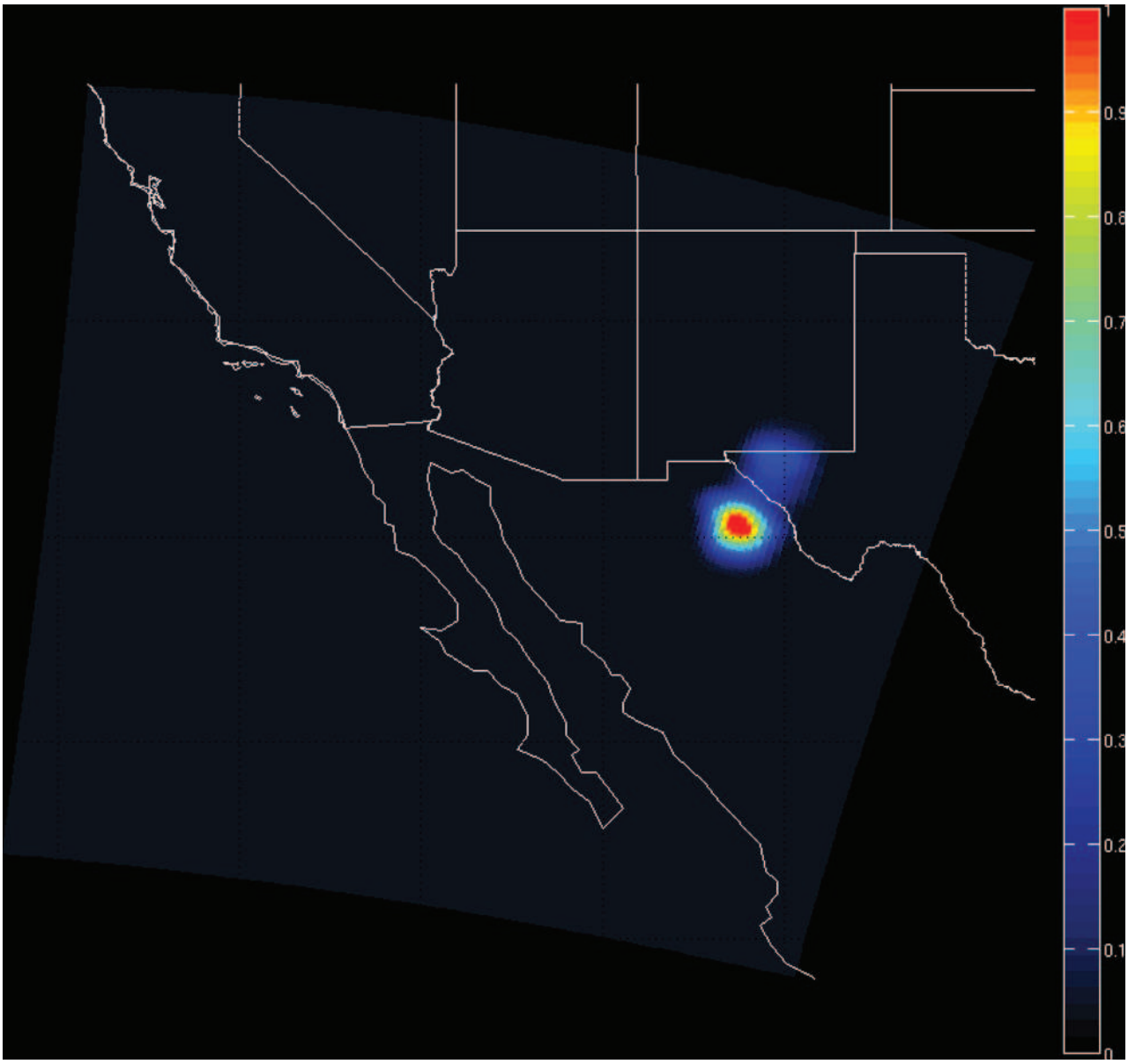

Figure 67. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

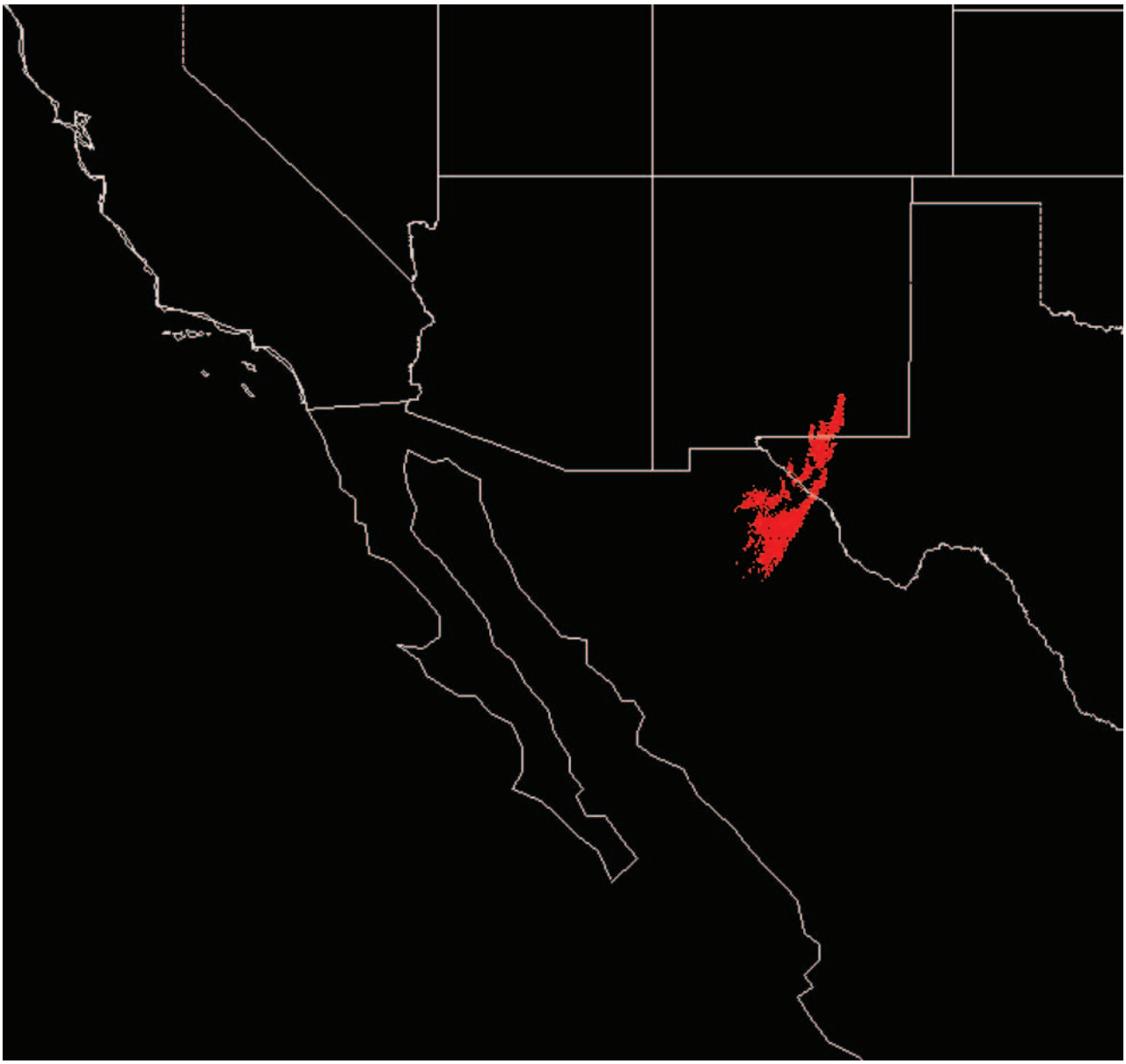

Figure 68. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2001-096 at 18:30 UTC.

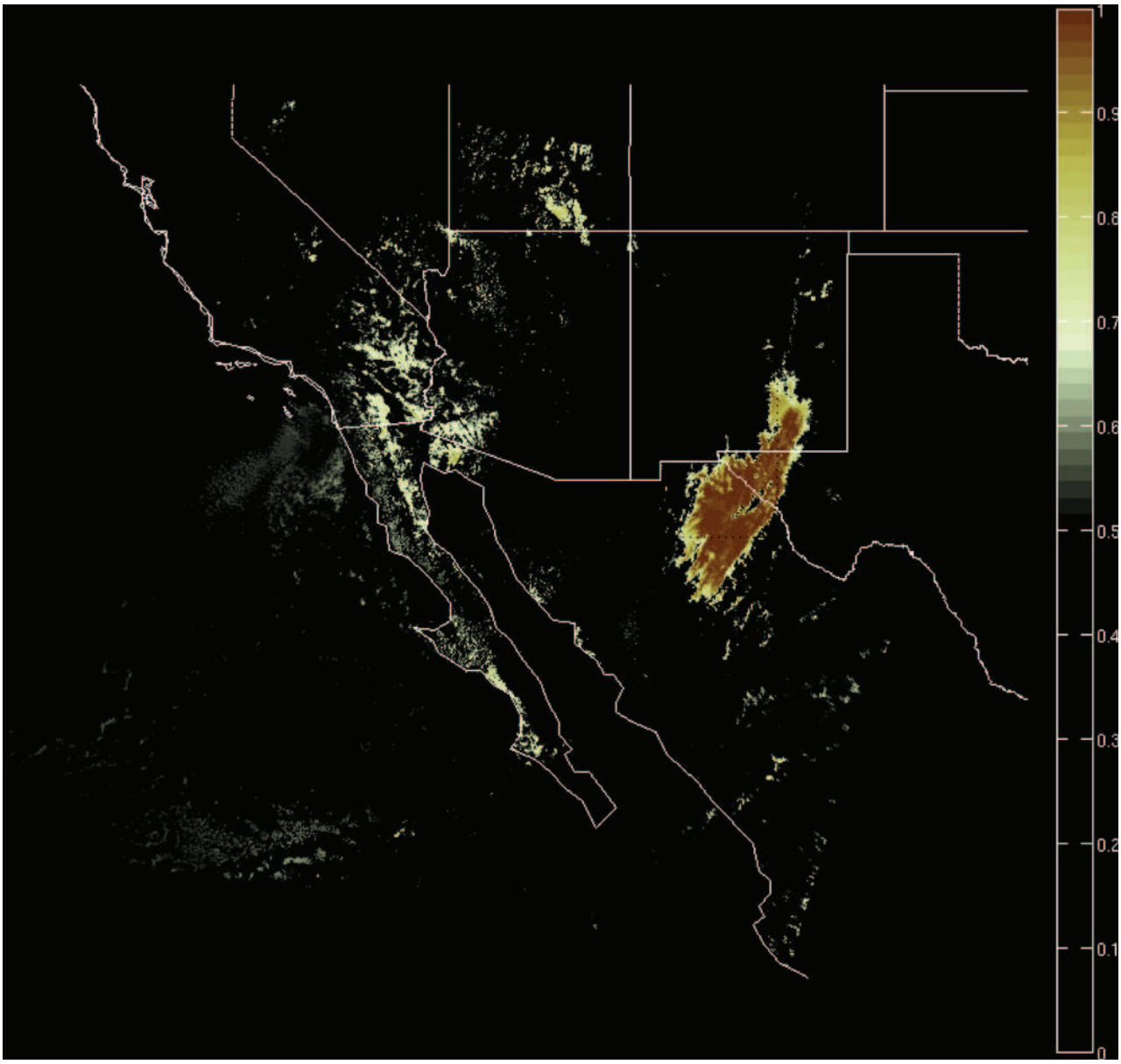

Figure 69. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

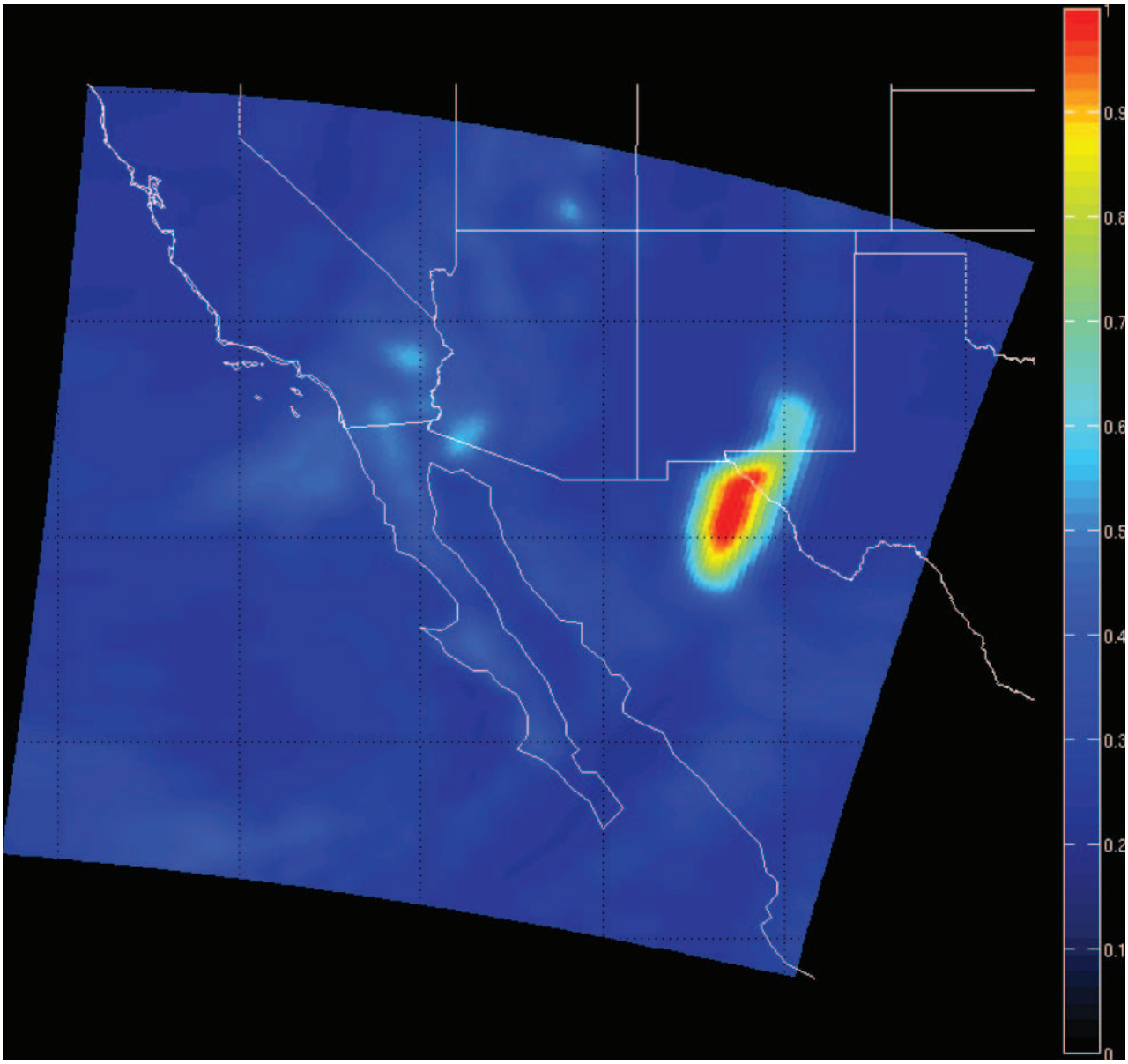

Figure 70. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

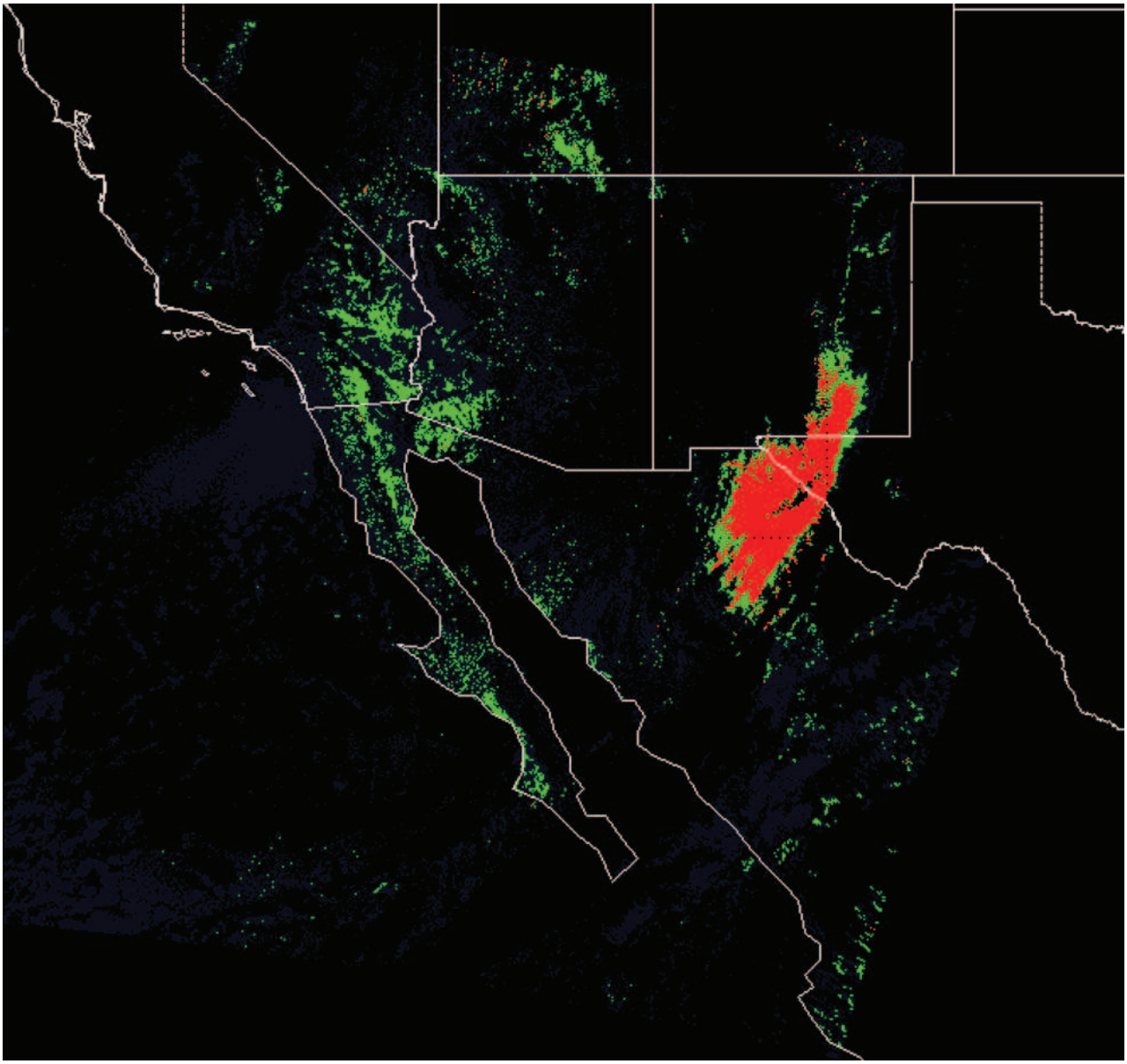

Figure 71. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-096 at 18:30 UTC.

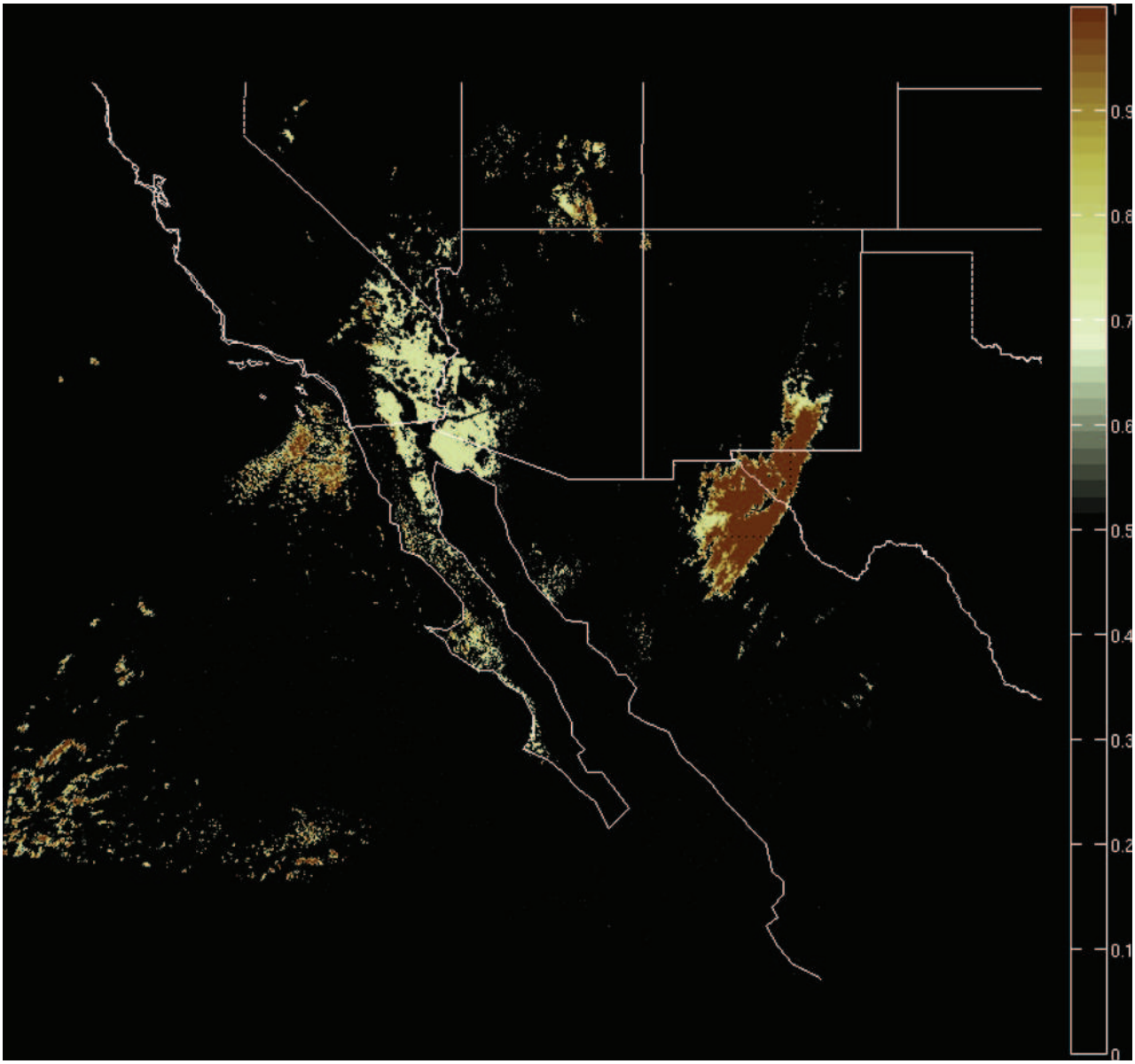

Figure 72. Result of the Probabilistic Neural Network (PNN) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

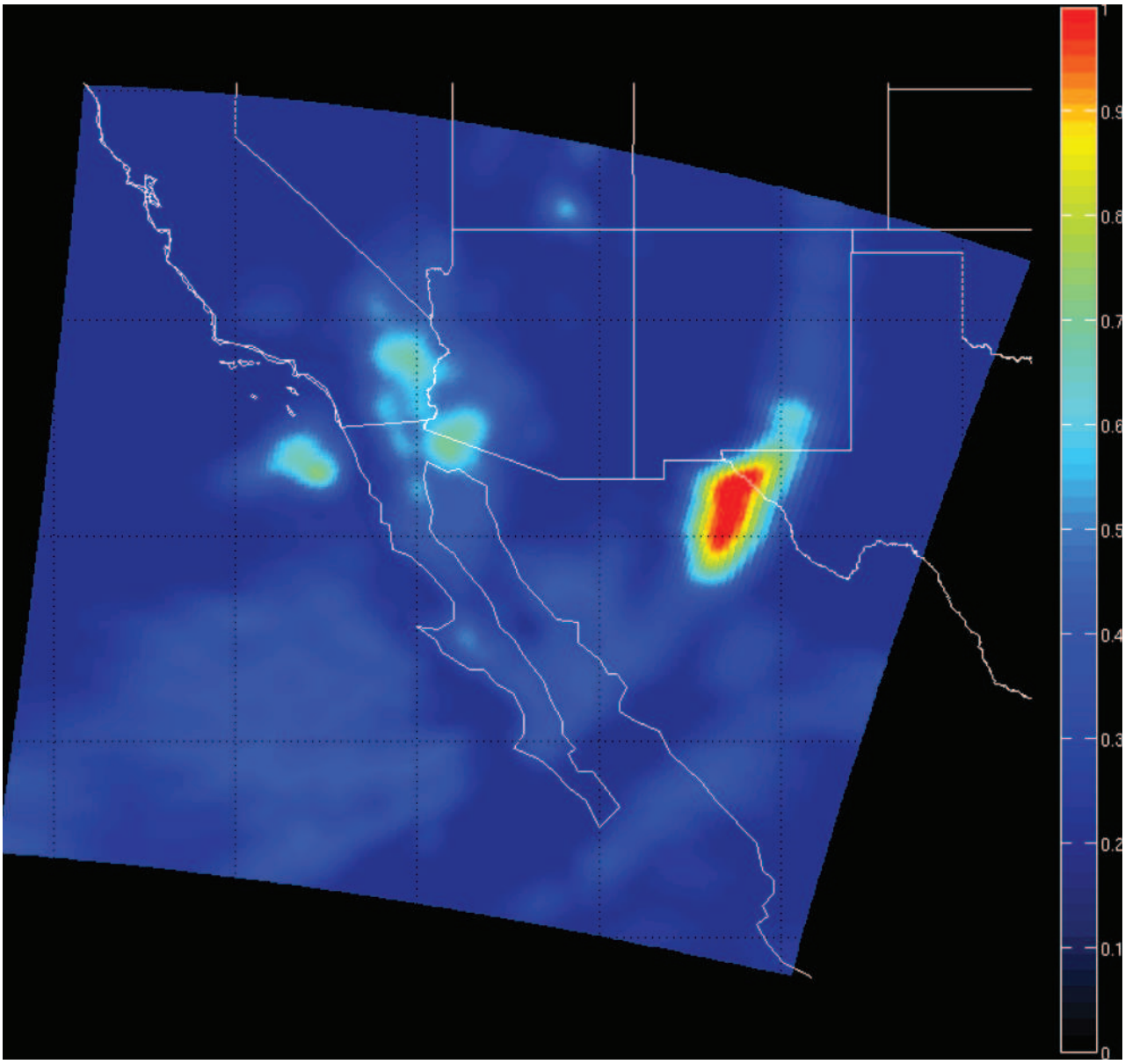

Figure 73. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2001-096 at 18:30 UTC. The colorbar shows the index of probability.

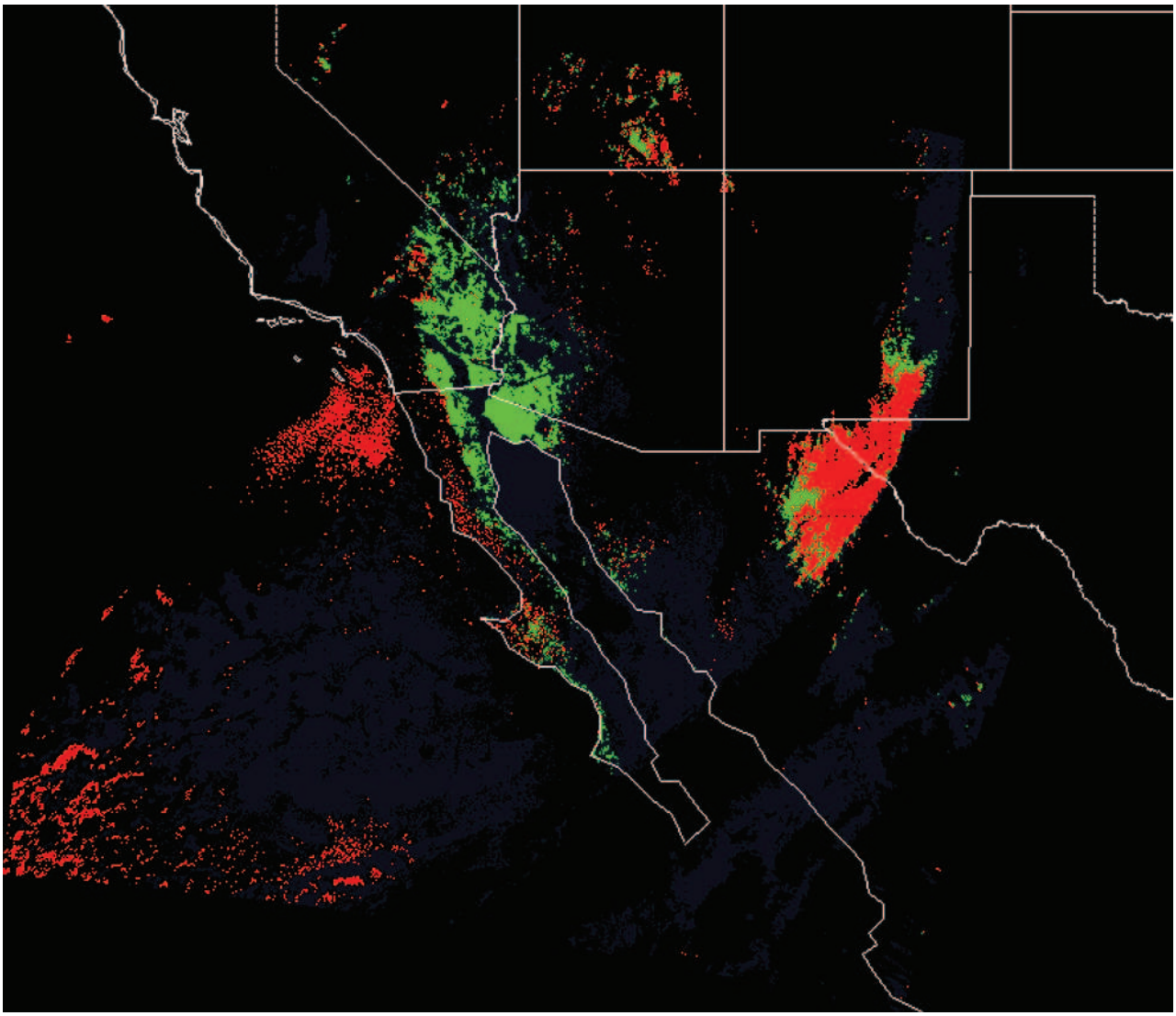

Figure 74. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2001-096 at 18:30 UTC.

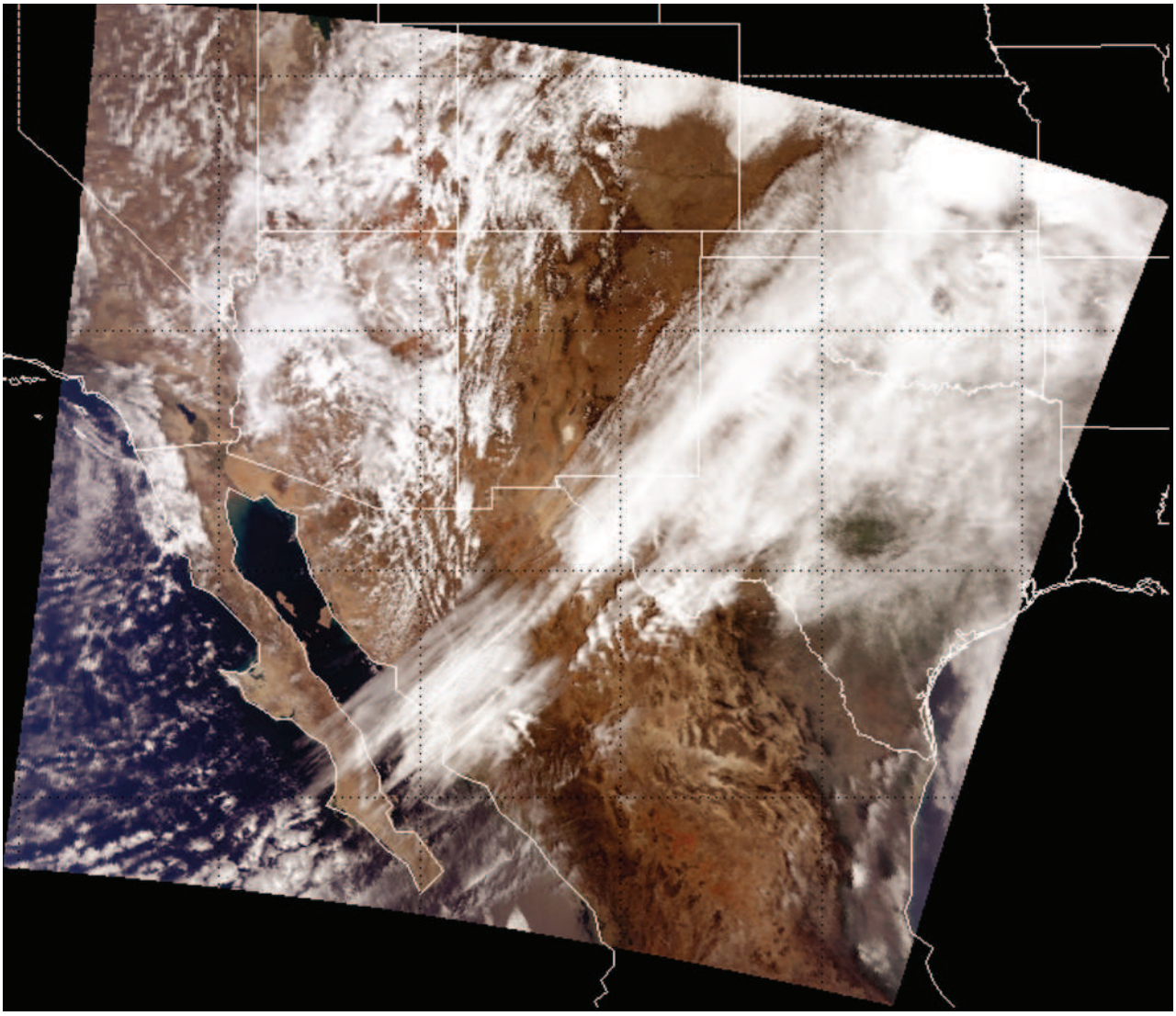

Figure 75. True color image for the event of 2001-100 at 18:05 UTC.

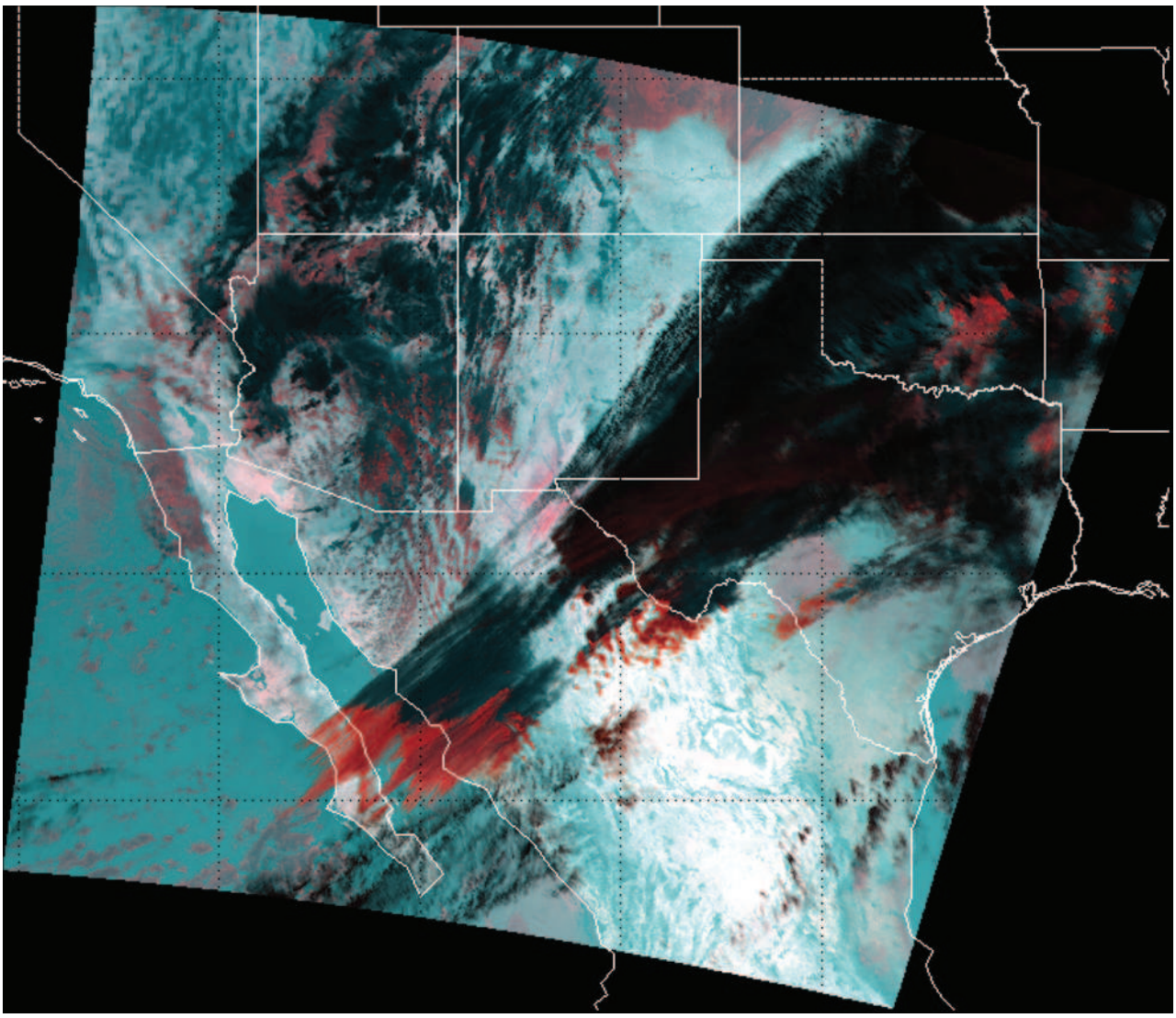

Figure 76. False color image for the event of 2001-100 at 18:05 UTC.

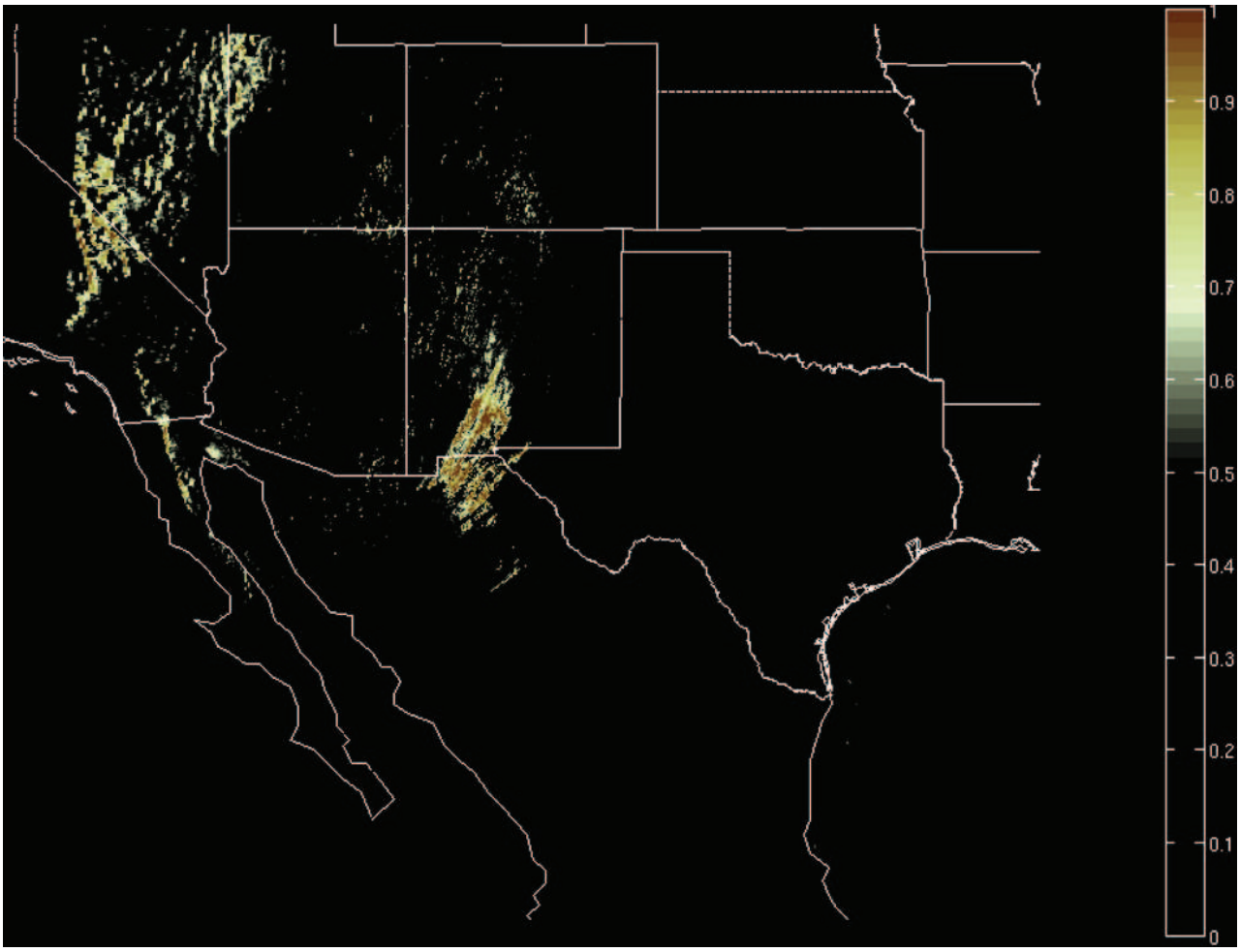

Figure 77. Result of the probabilistic display model (PD) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

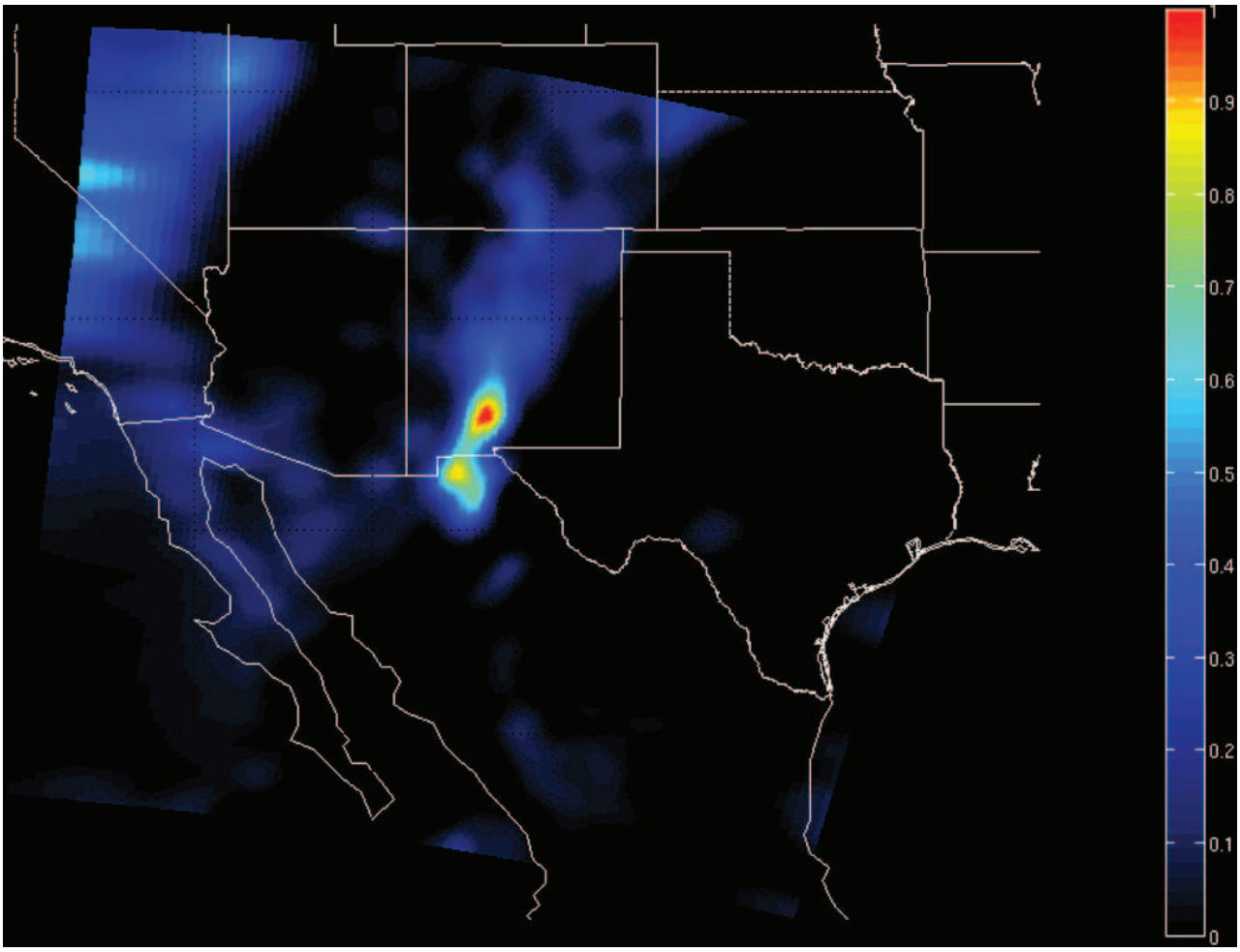

Figure 78. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

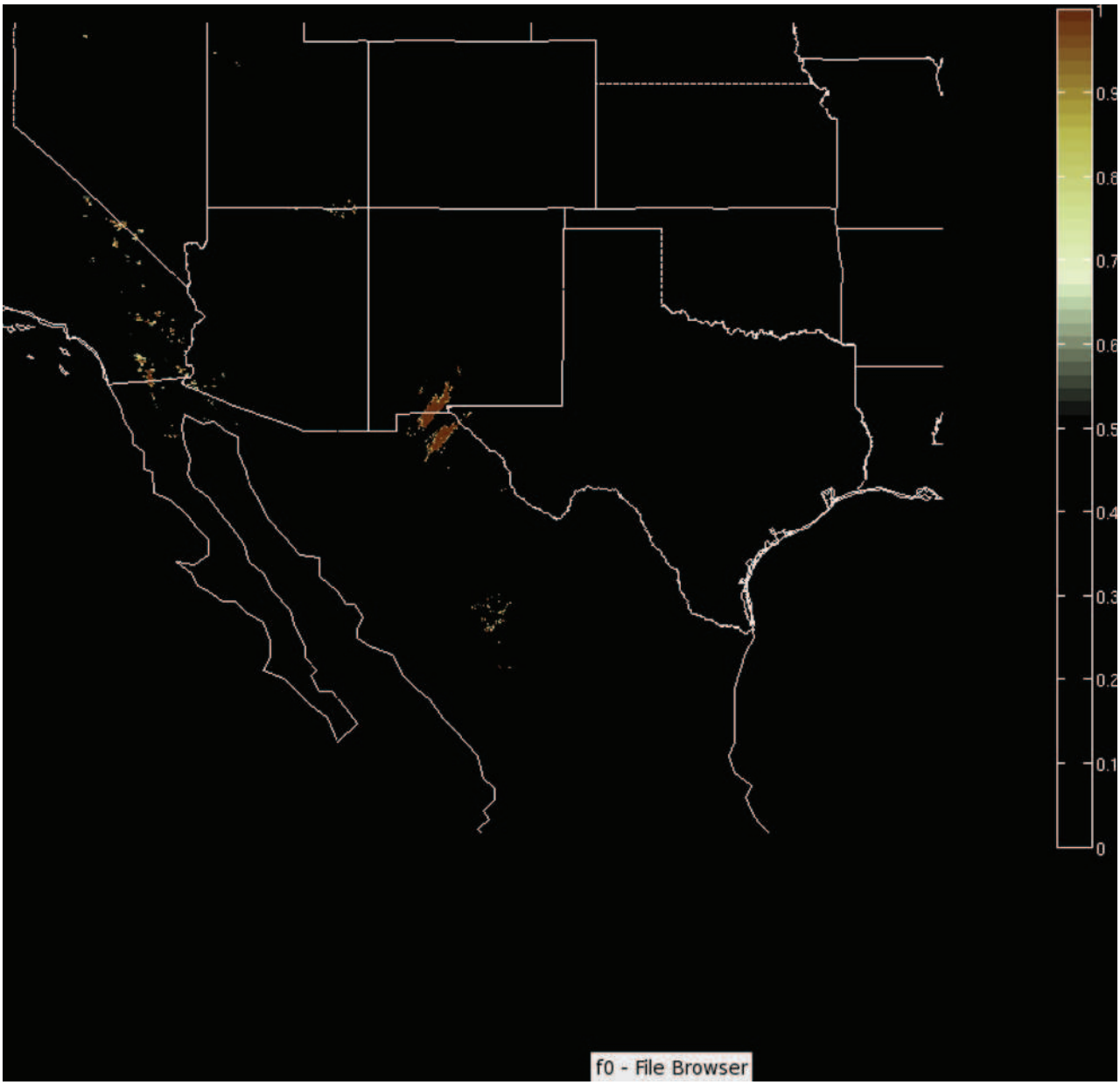

Figure 79. Result of the Maximum Likelihood model (MLE) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

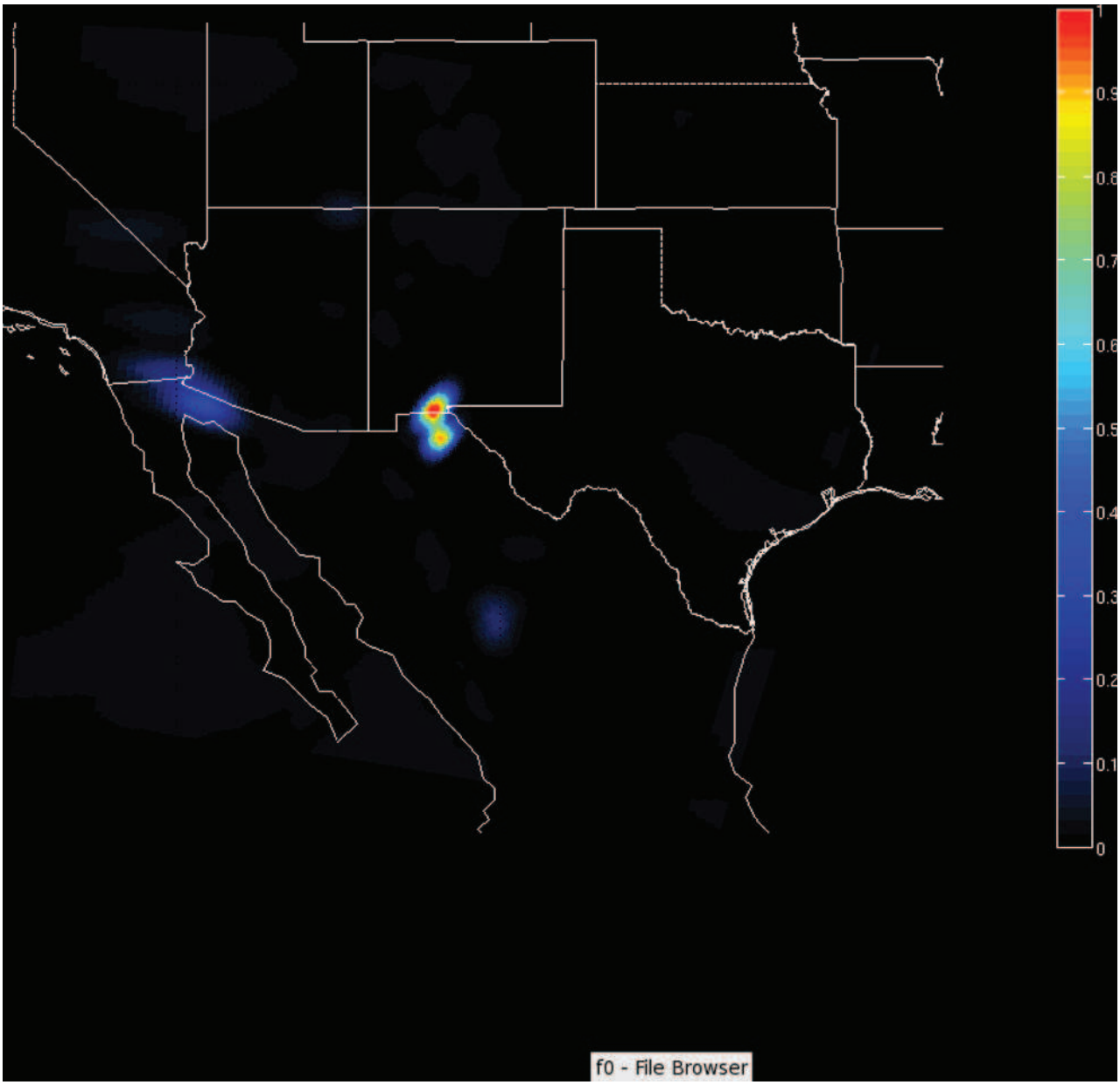

Figure 80. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

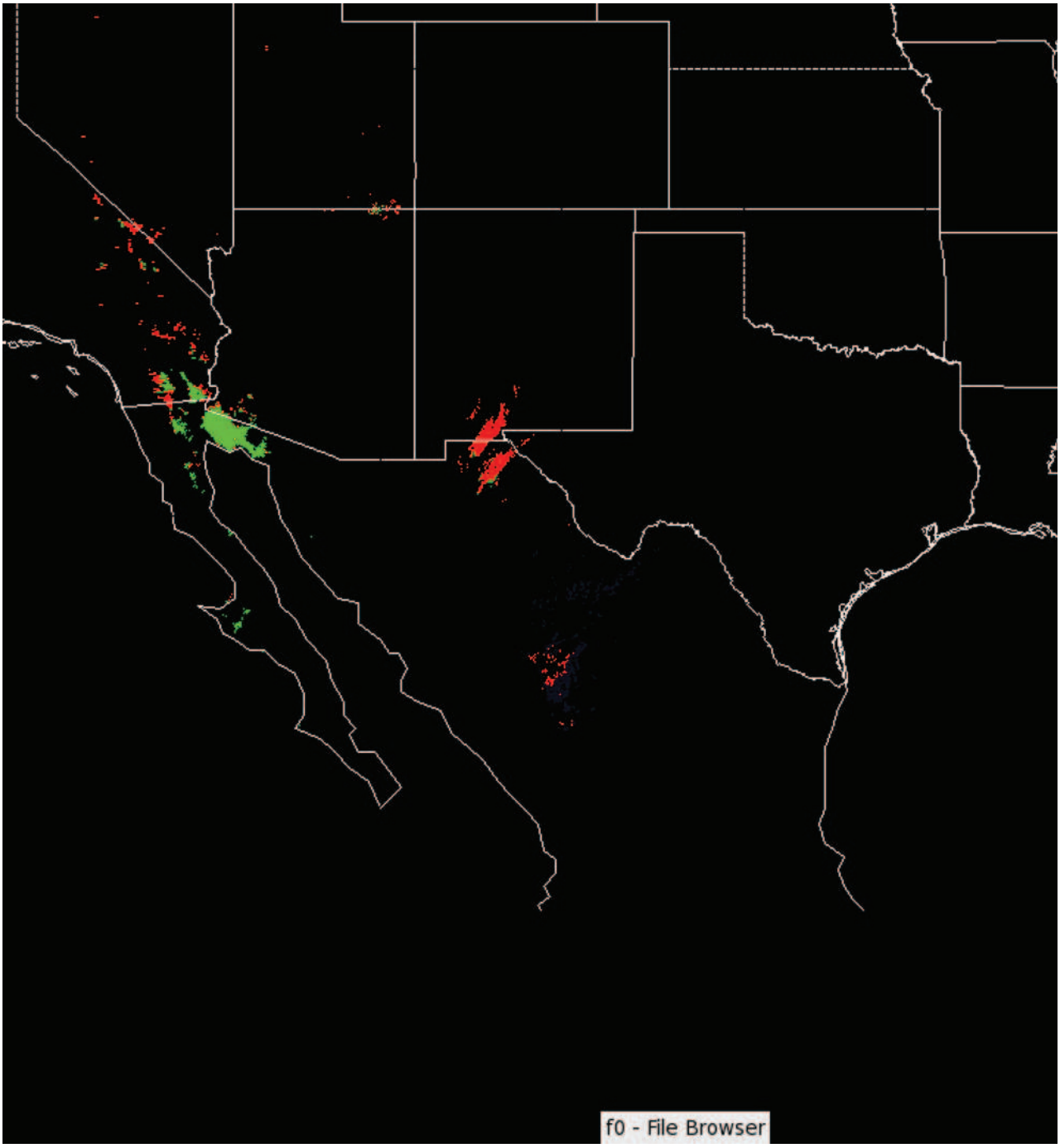

Figure 81. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2001-100 at 18:05 UTC.

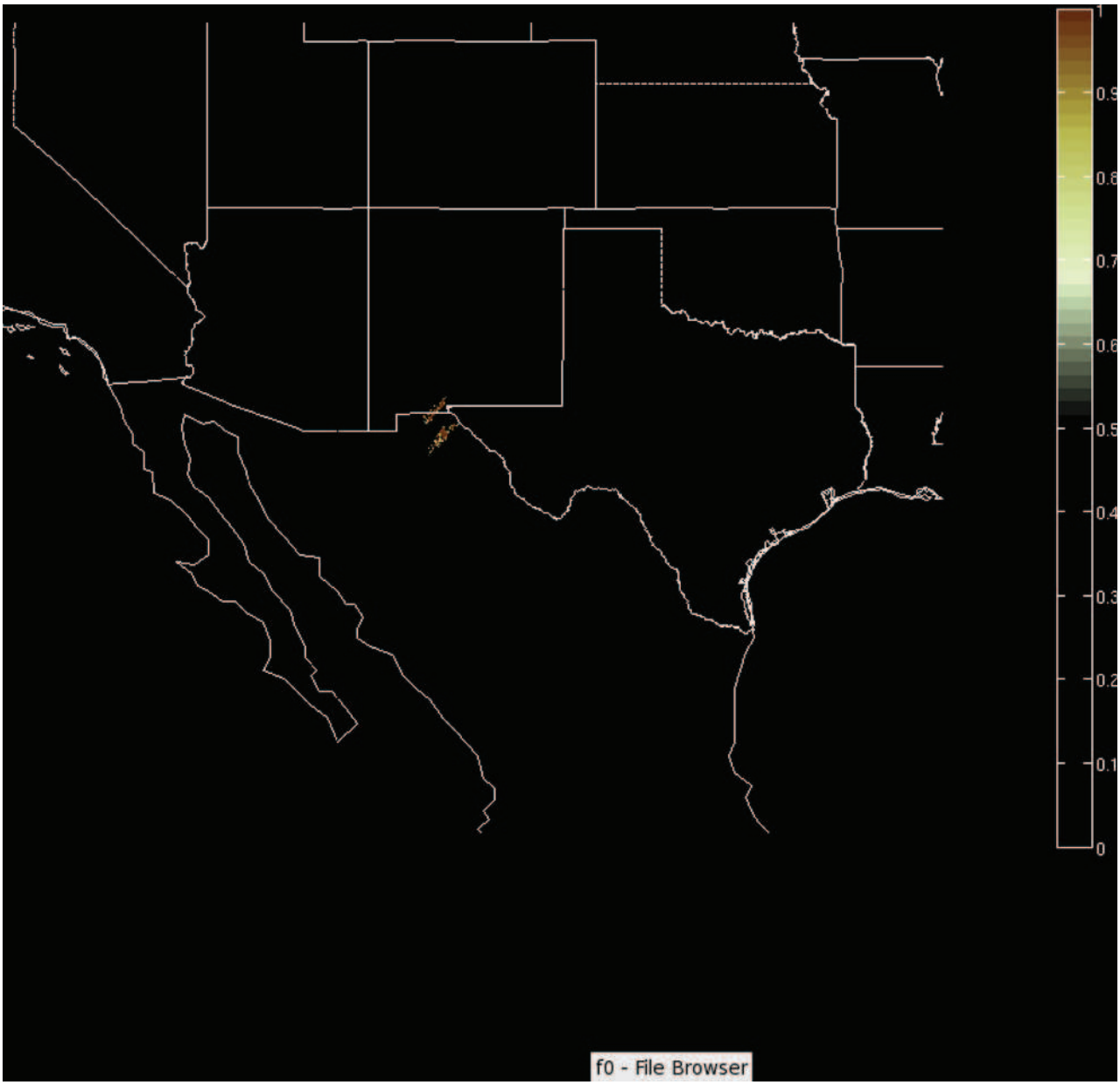

Figure 82. Result of the Maximum a posteriori model (MAP) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

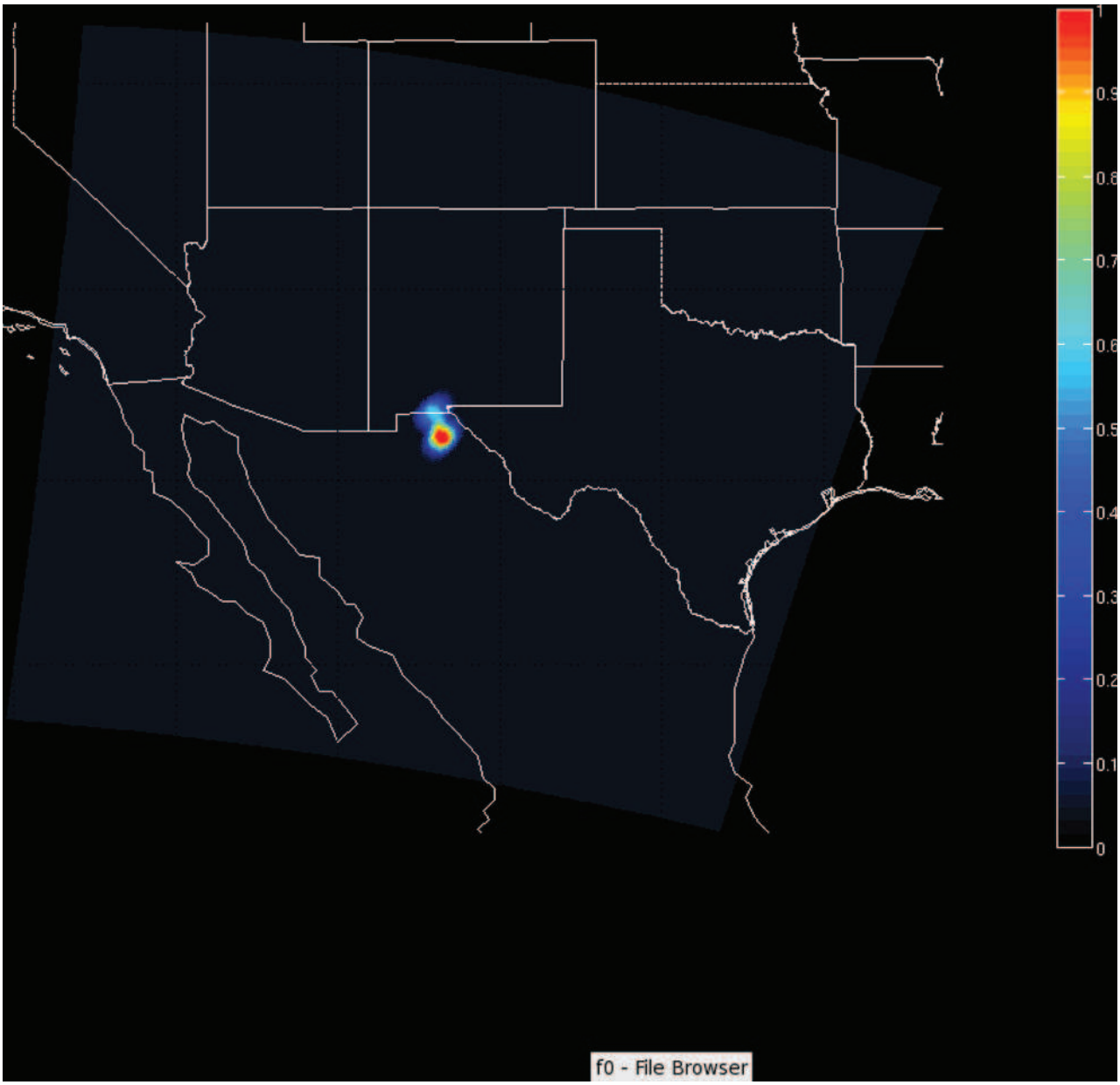

Figure 83. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

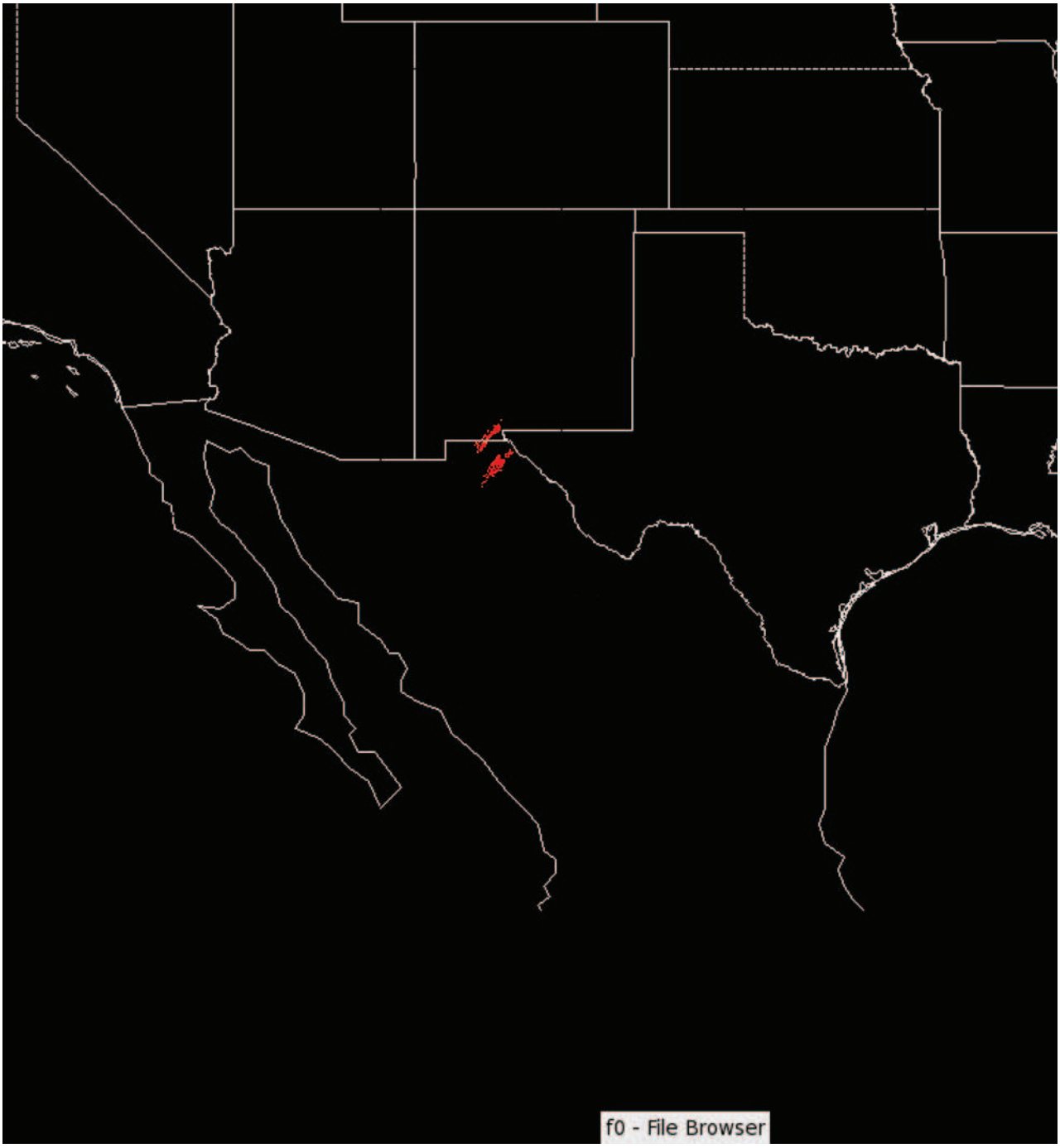

Figure 84. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2001-100 at 18:05 UTC.

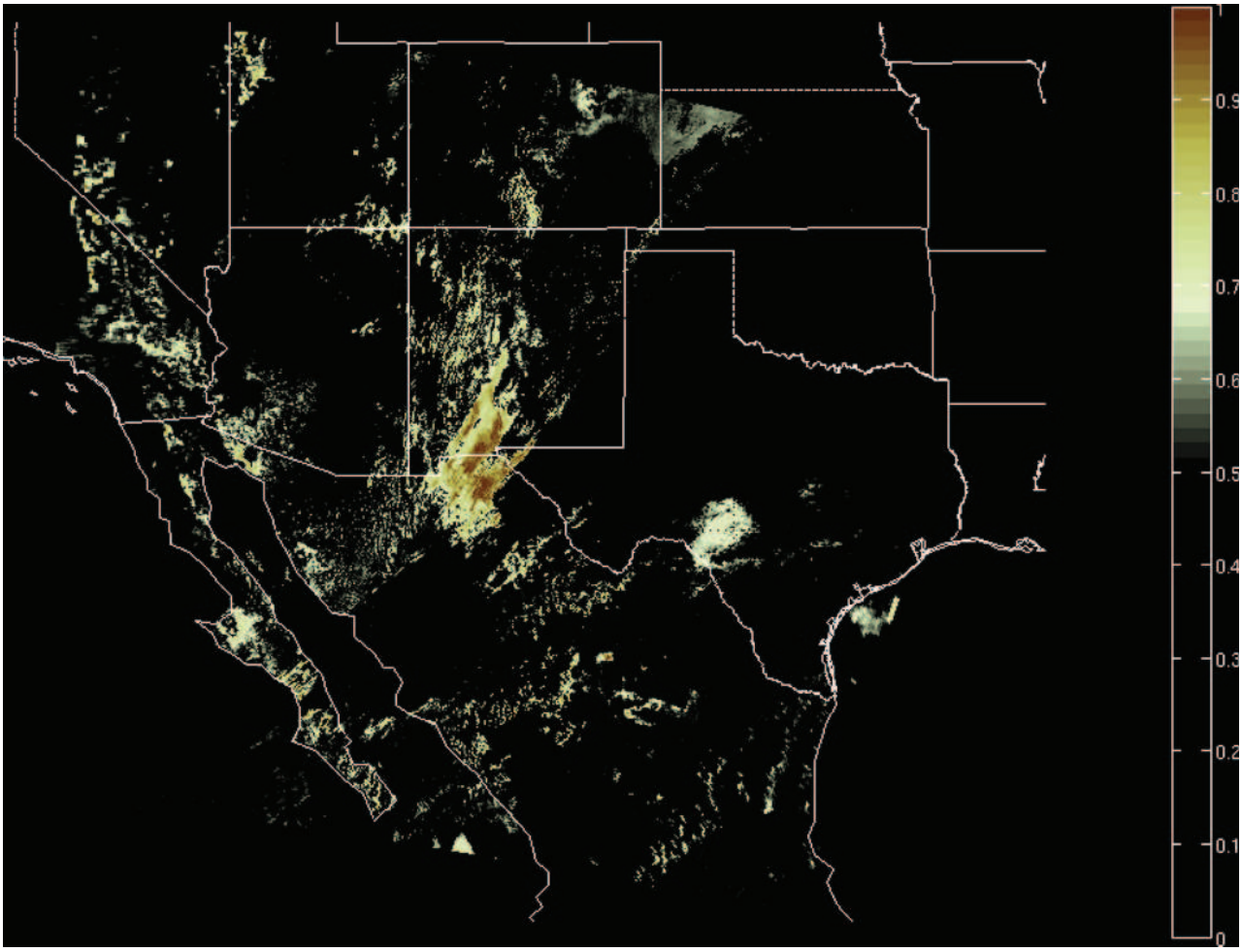

Figure 85. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

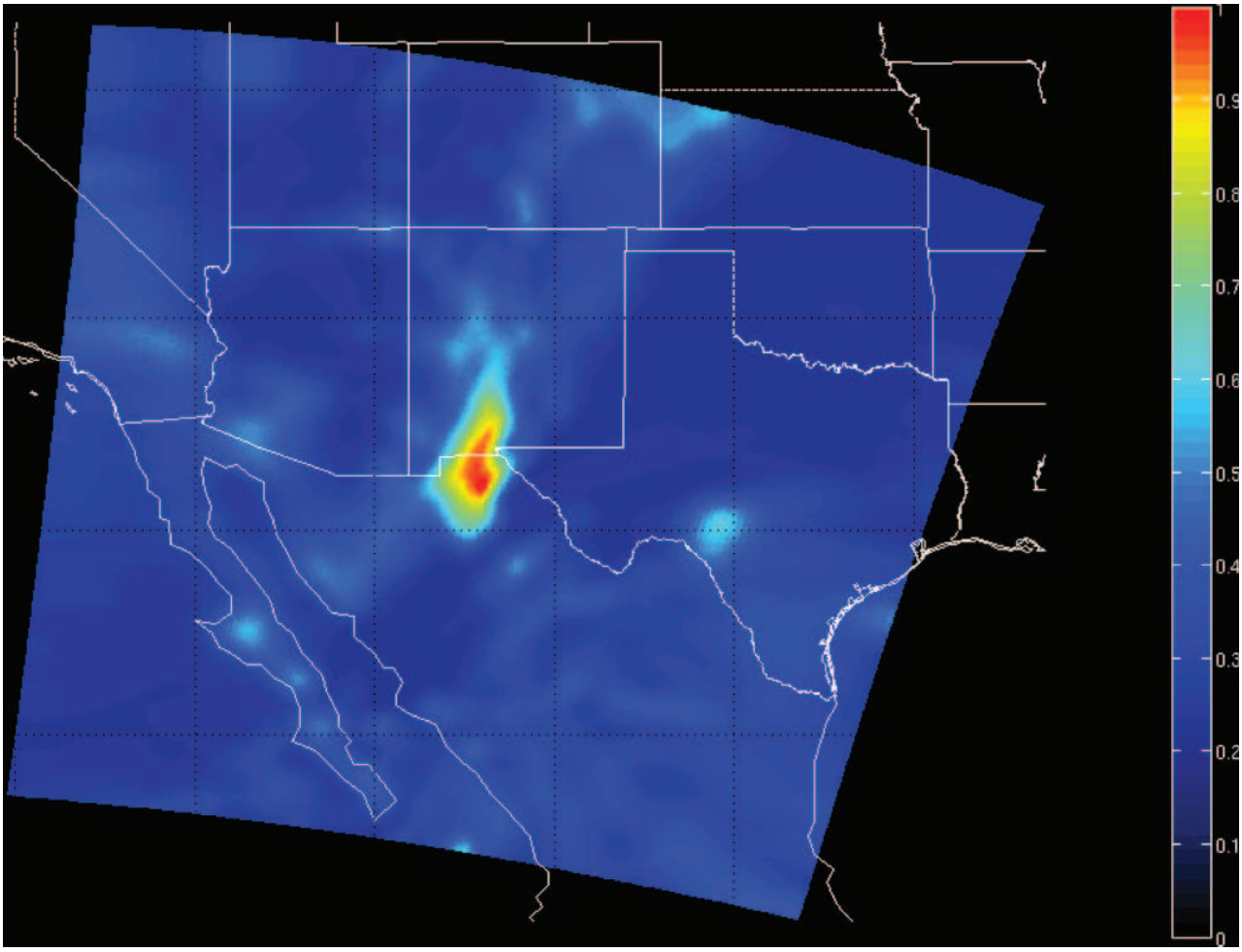

Figure 86. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

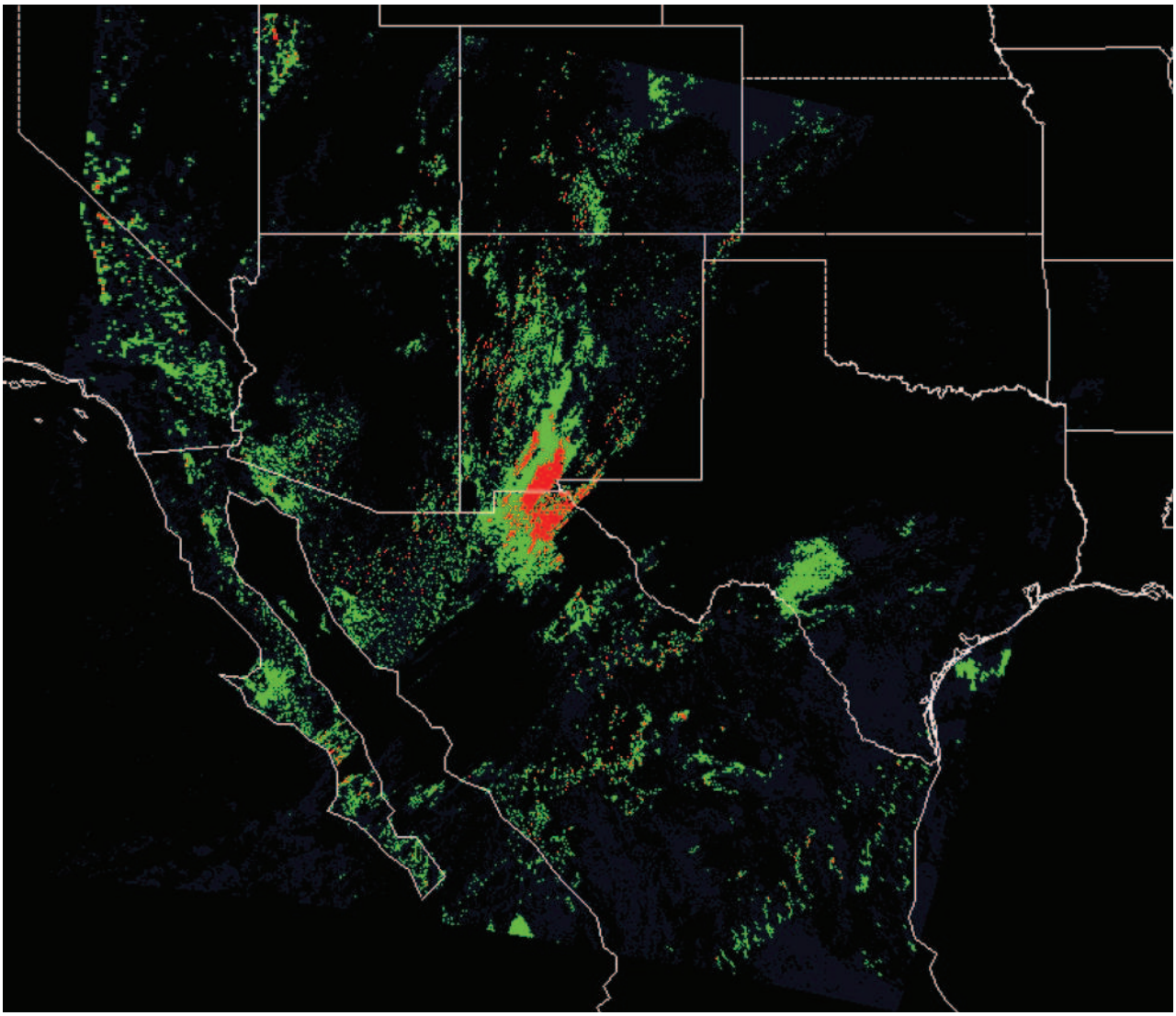

Figure 87. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-100 at 18:05 UTC.

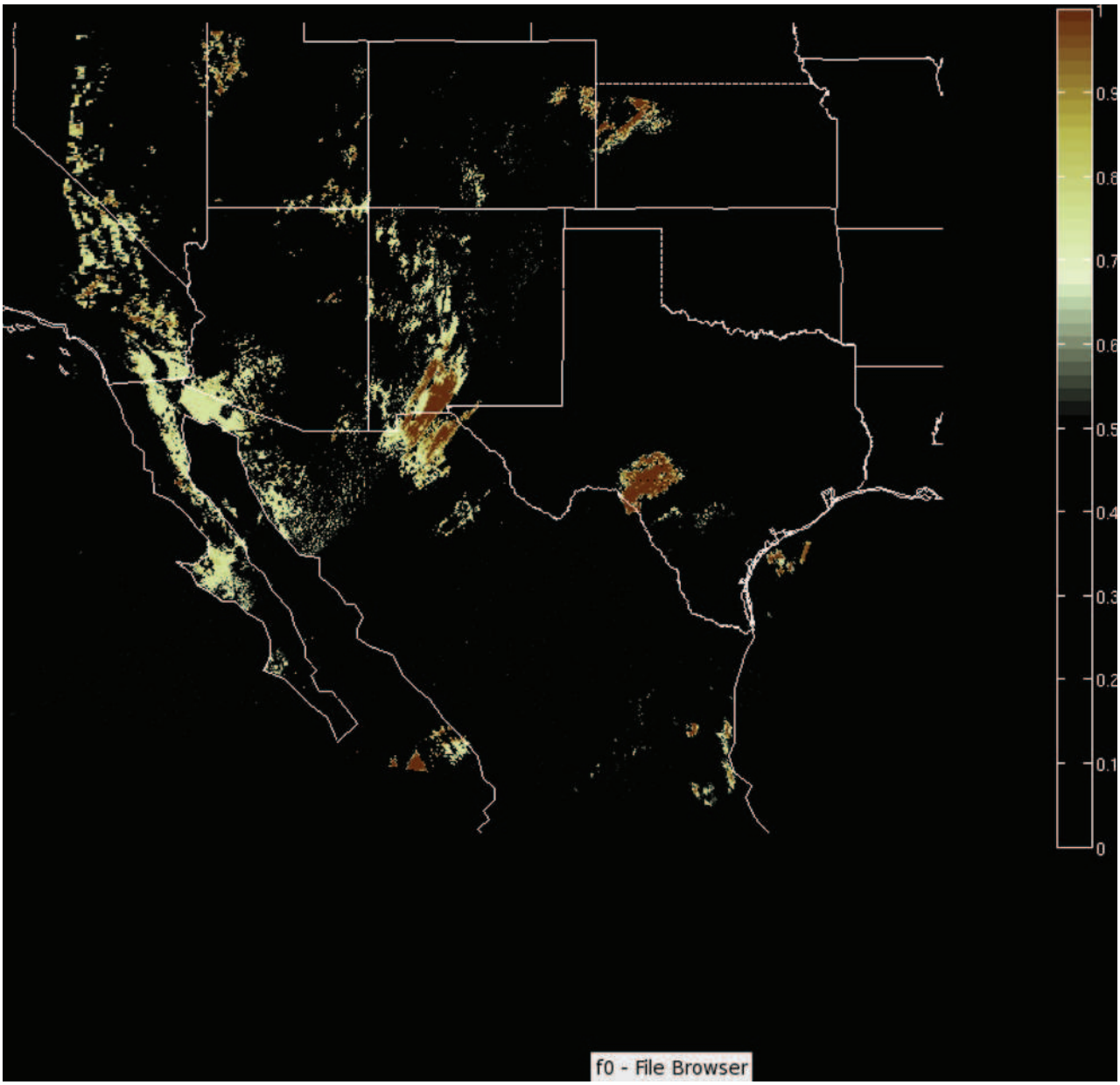

Figure 88. Result of the Probabilistic Neural Network (PNN) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

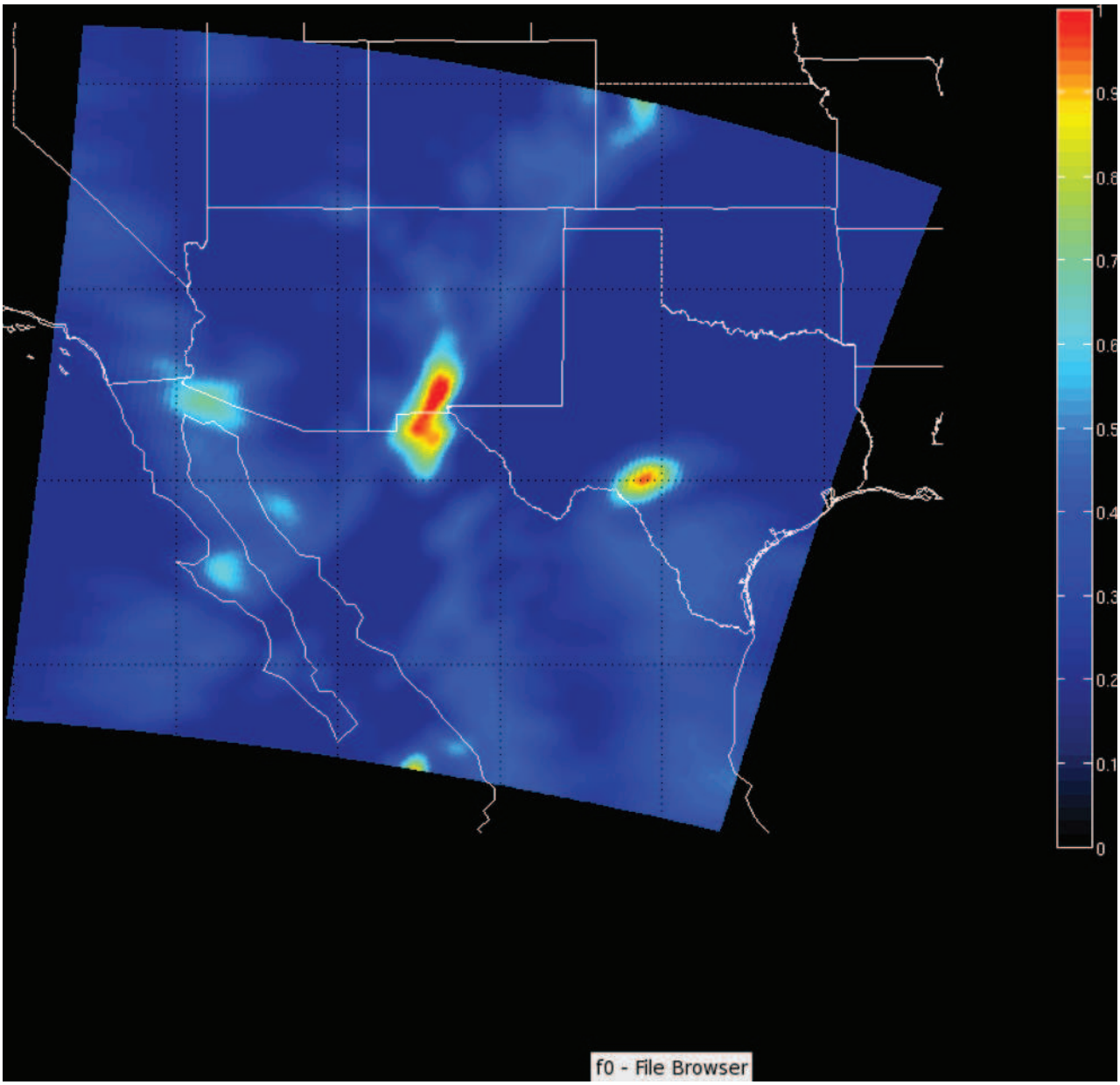

Figure 89. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2001-100 at 18:05 UTC. The colorbar shows the index of probability.

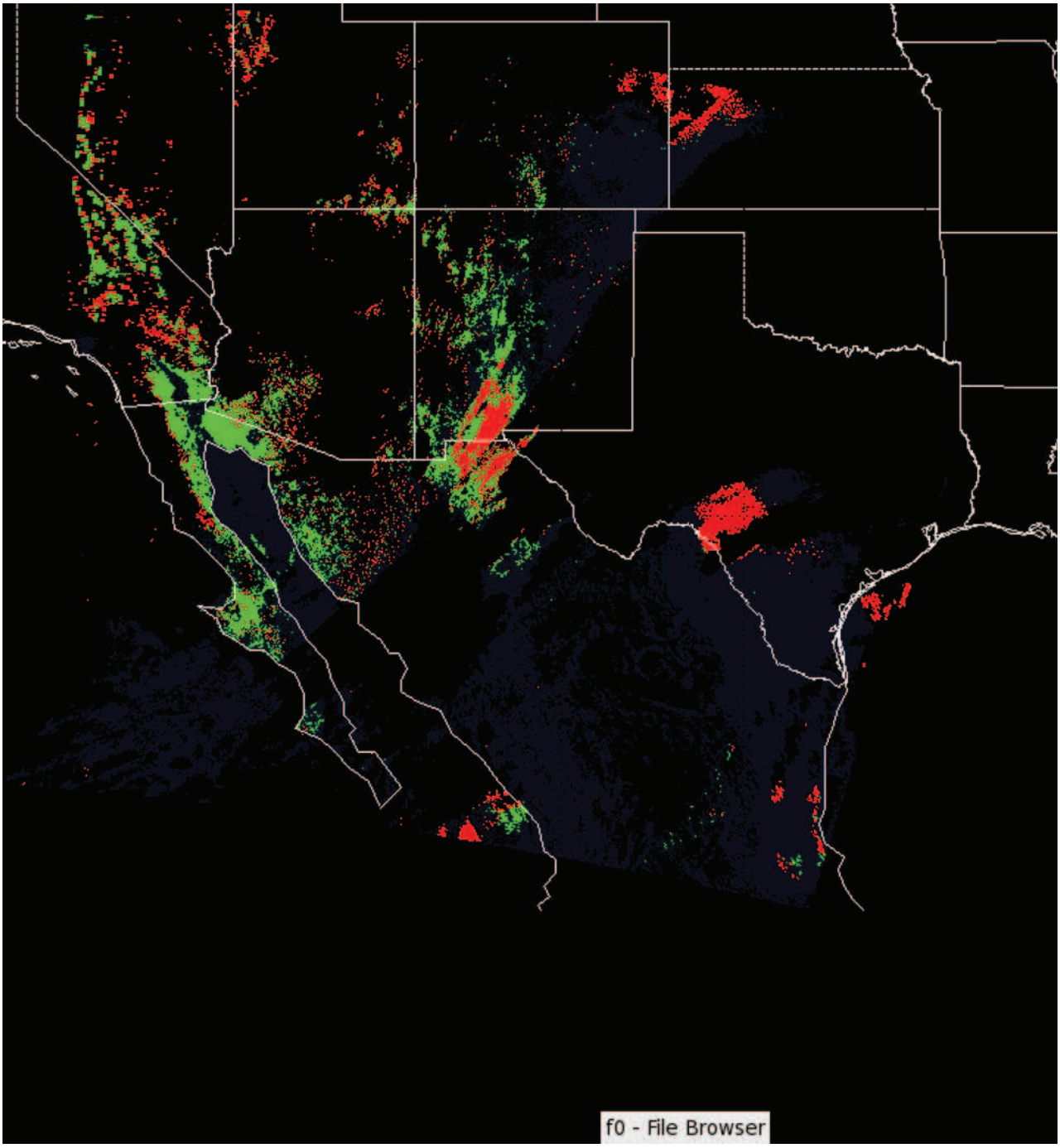

Figure 90. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2001-100 at 18:05 UTC.

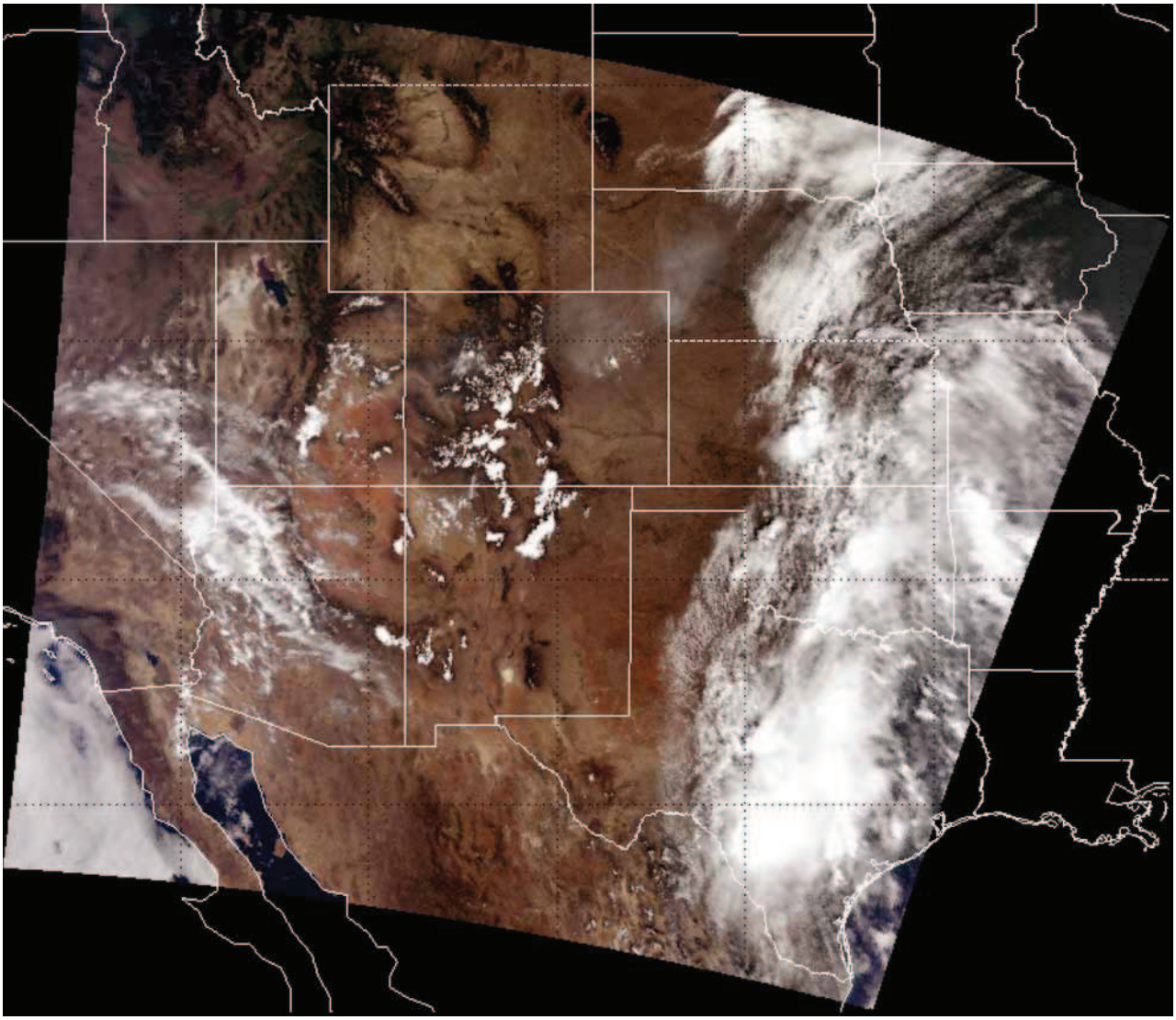

Figure 91. True color image for the event of 2002-183 at 17:55 UTC.

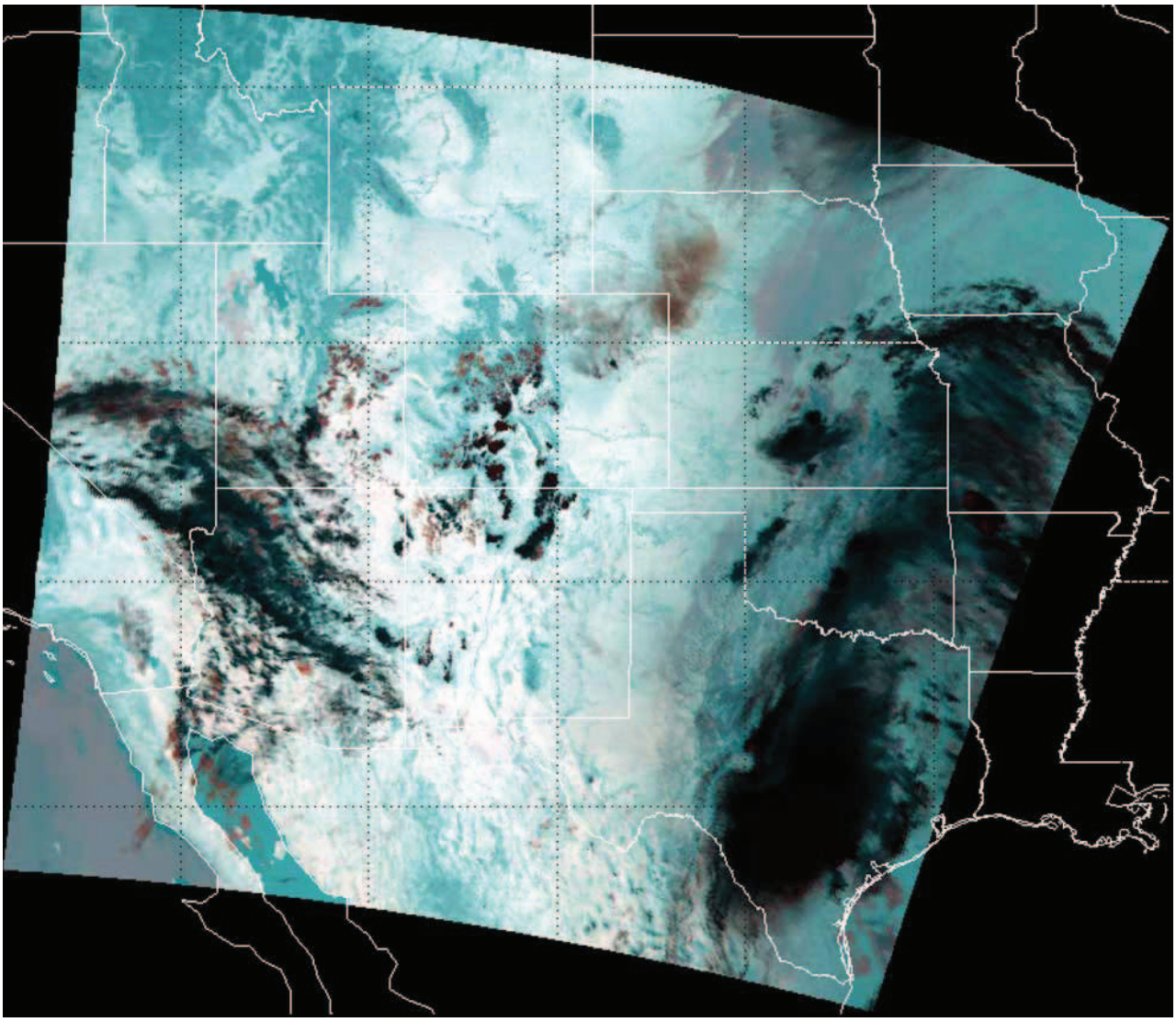

Figure 92. False color image for the event of 2002-183 at 17:55 UTC.

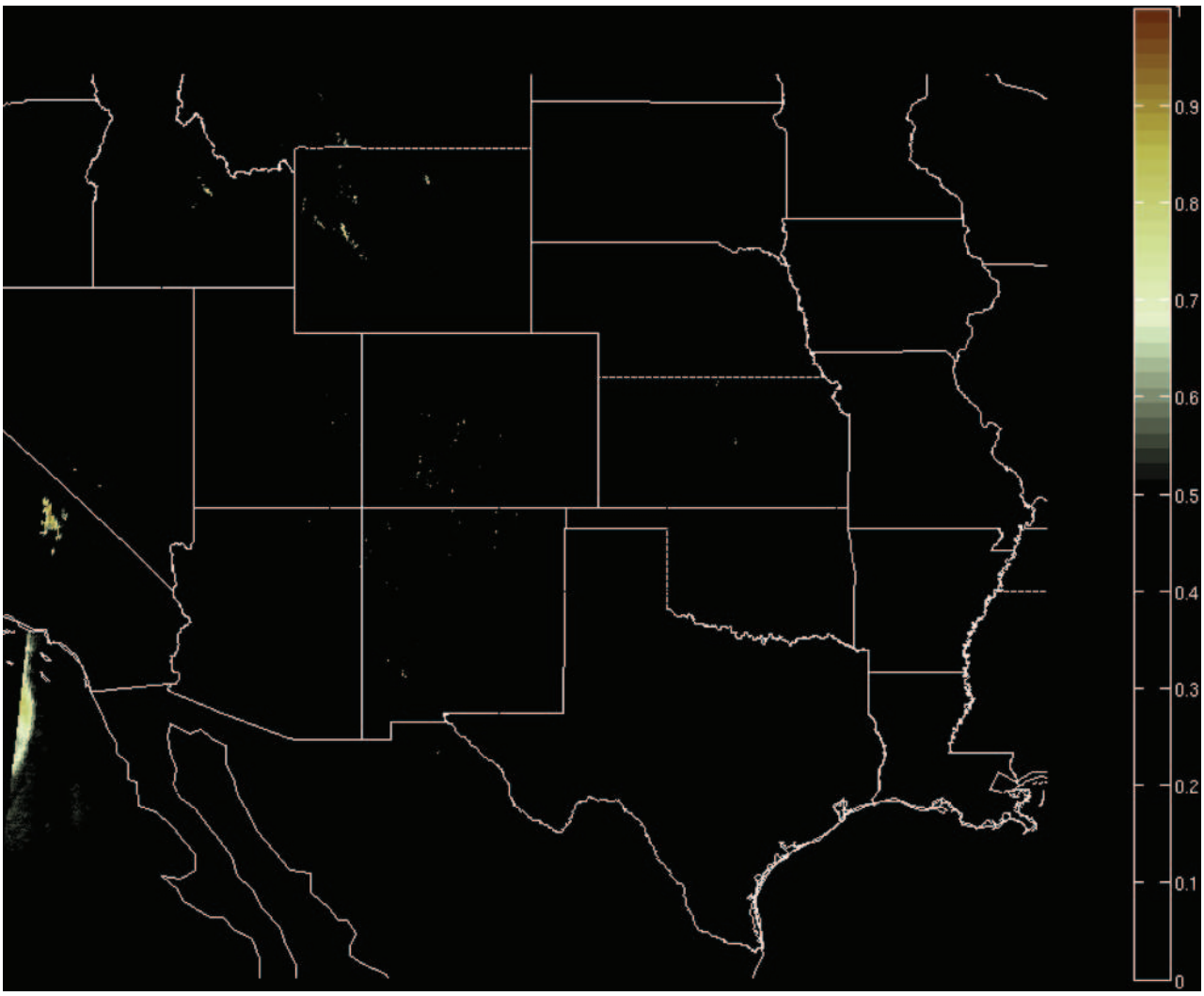

Figure 93. Result of the probabilistic display model (PD) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

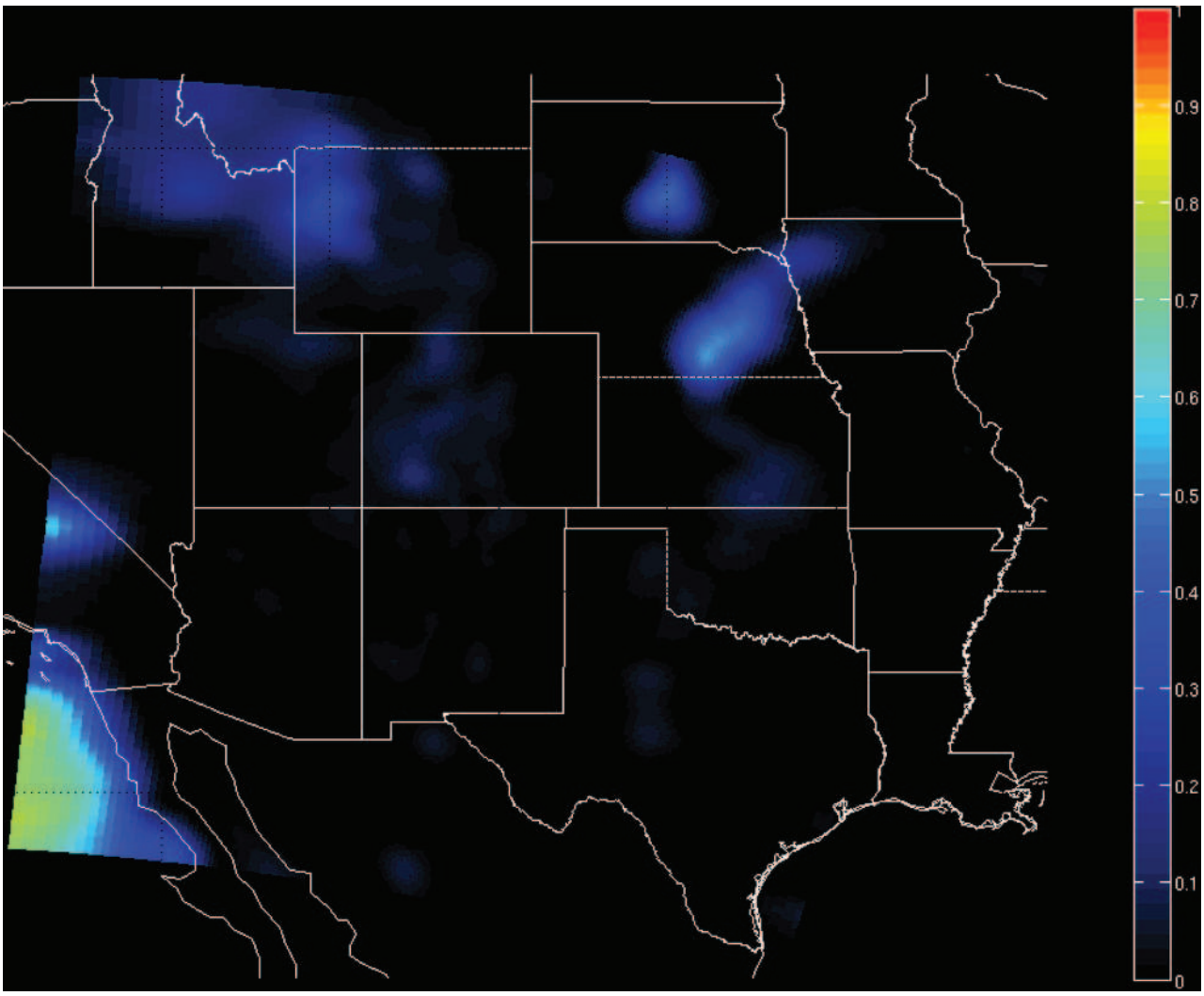

Figure 94. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

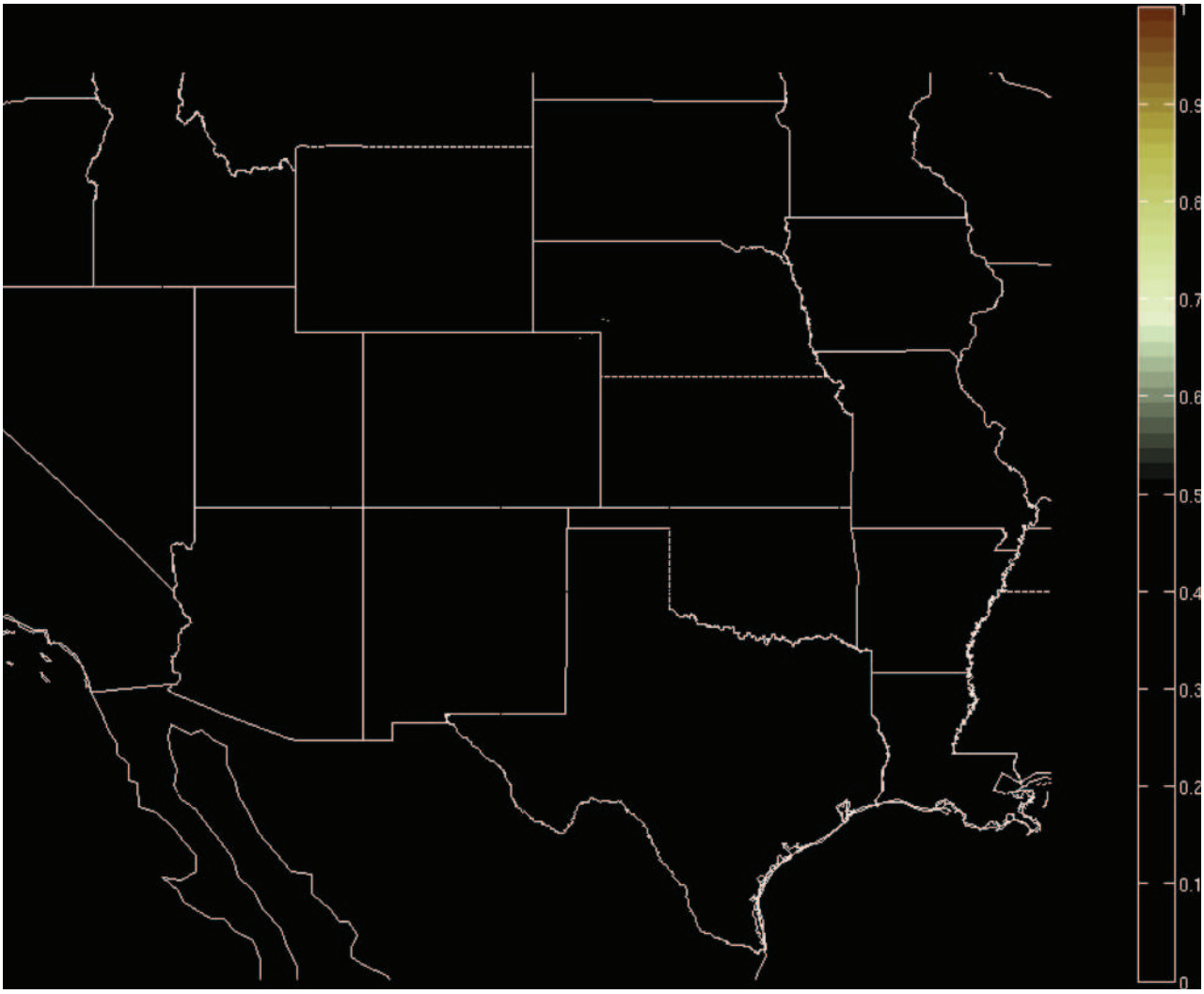

Figure 95. Result of the Maximum Likelihood model (MLE) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

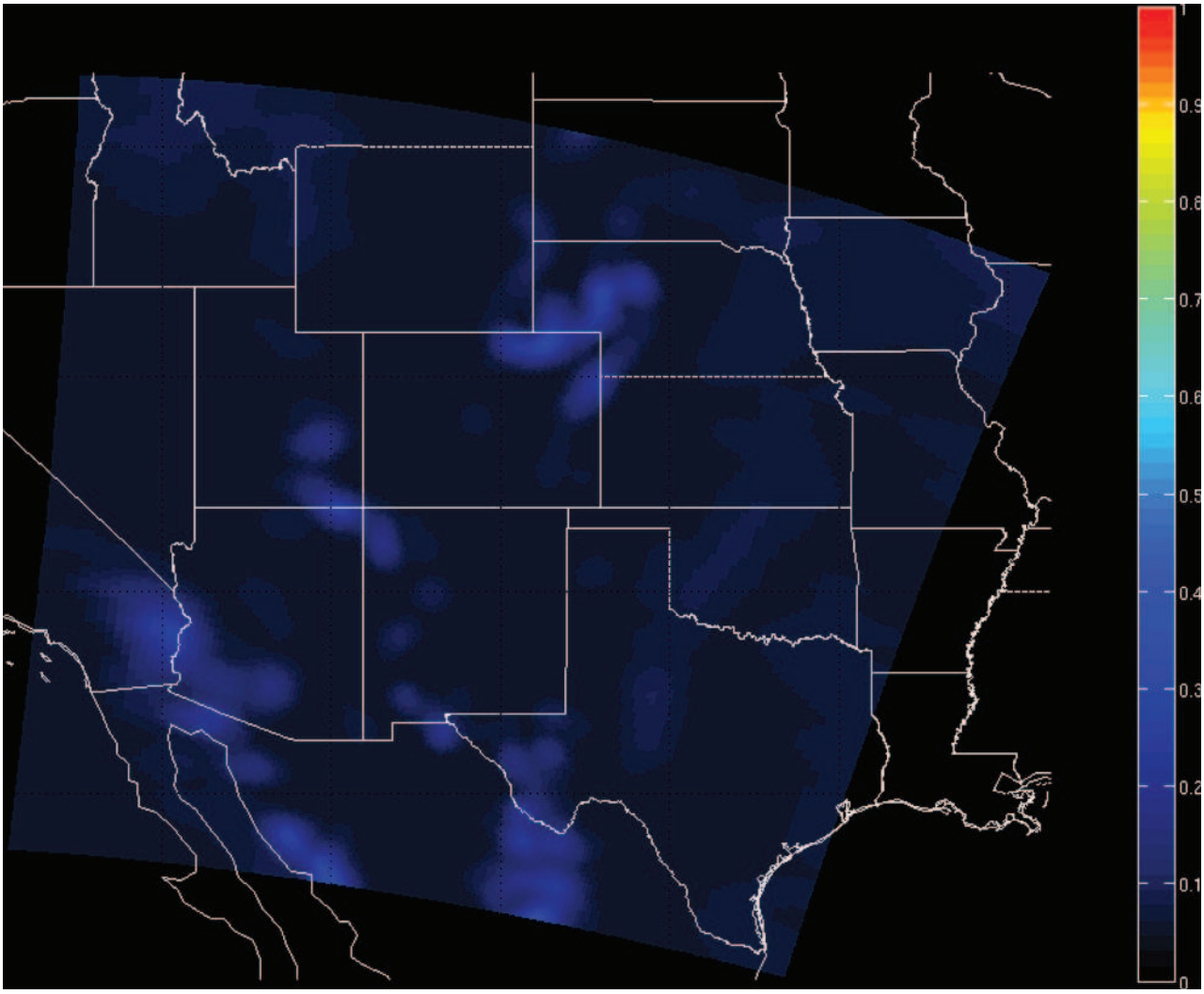

Figure 96. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

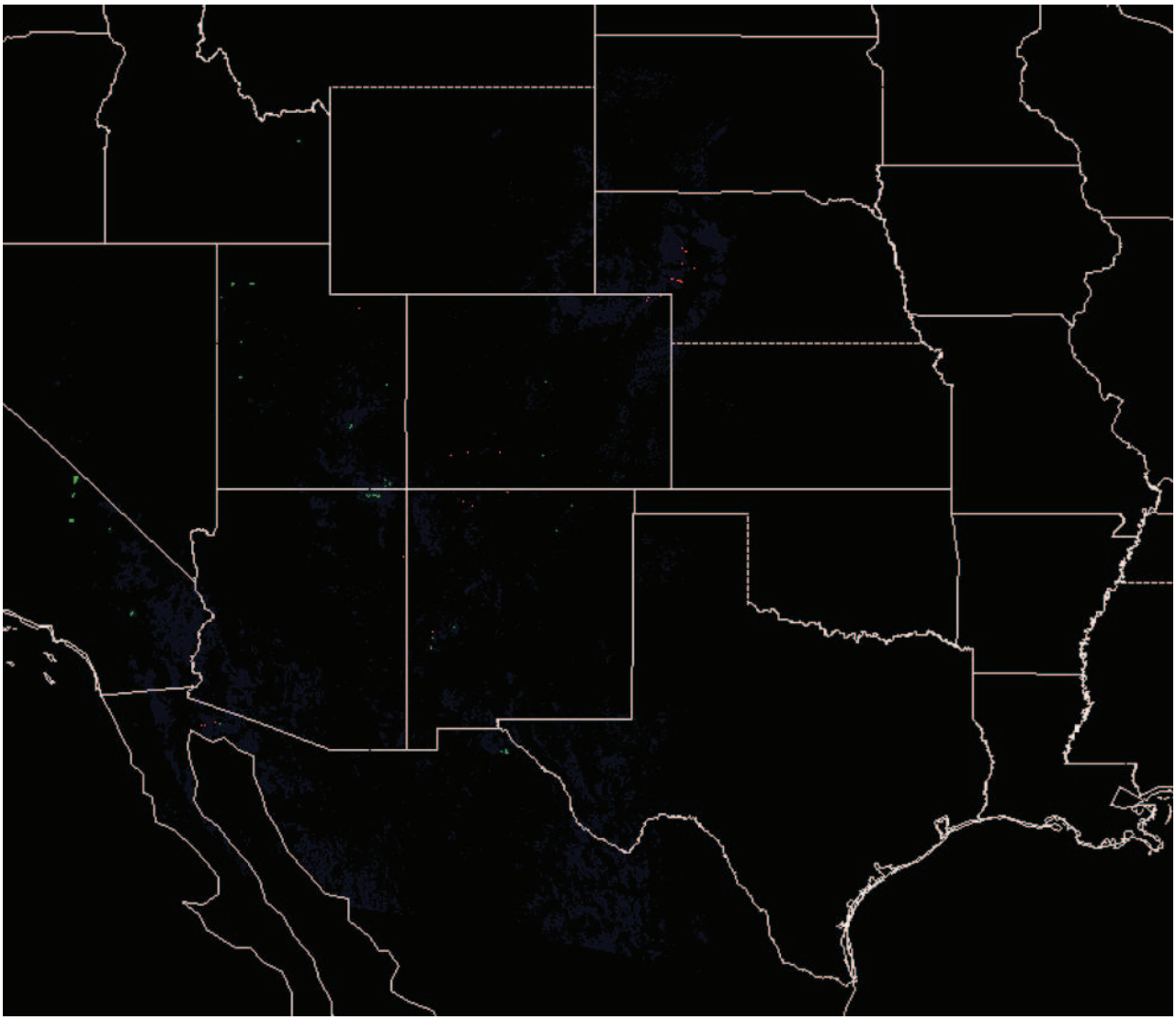

Figure 97. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2002-183 at 17:55 UTC.

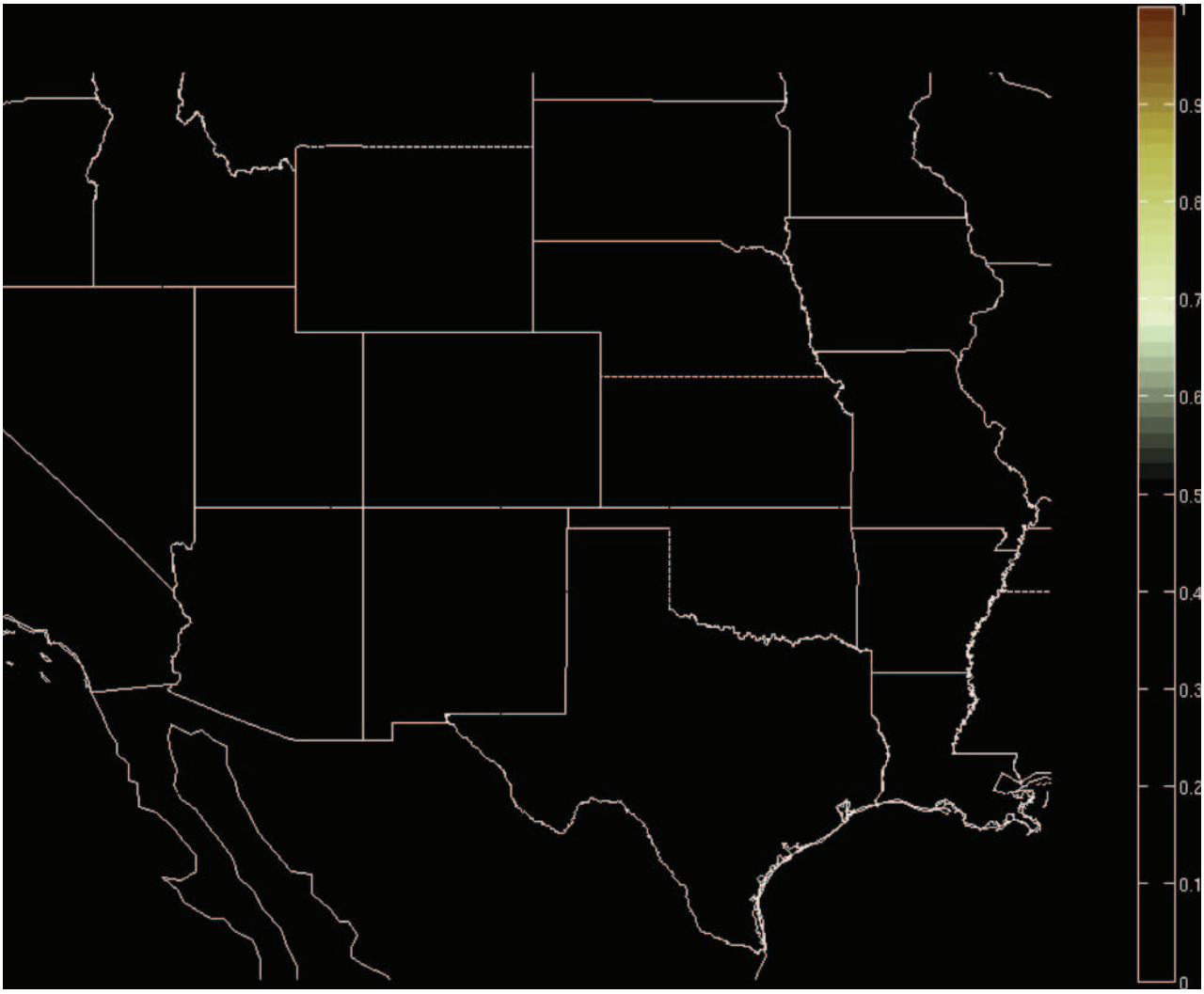

Figure 98. Result of the Maximum a posteriori model (MAP) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

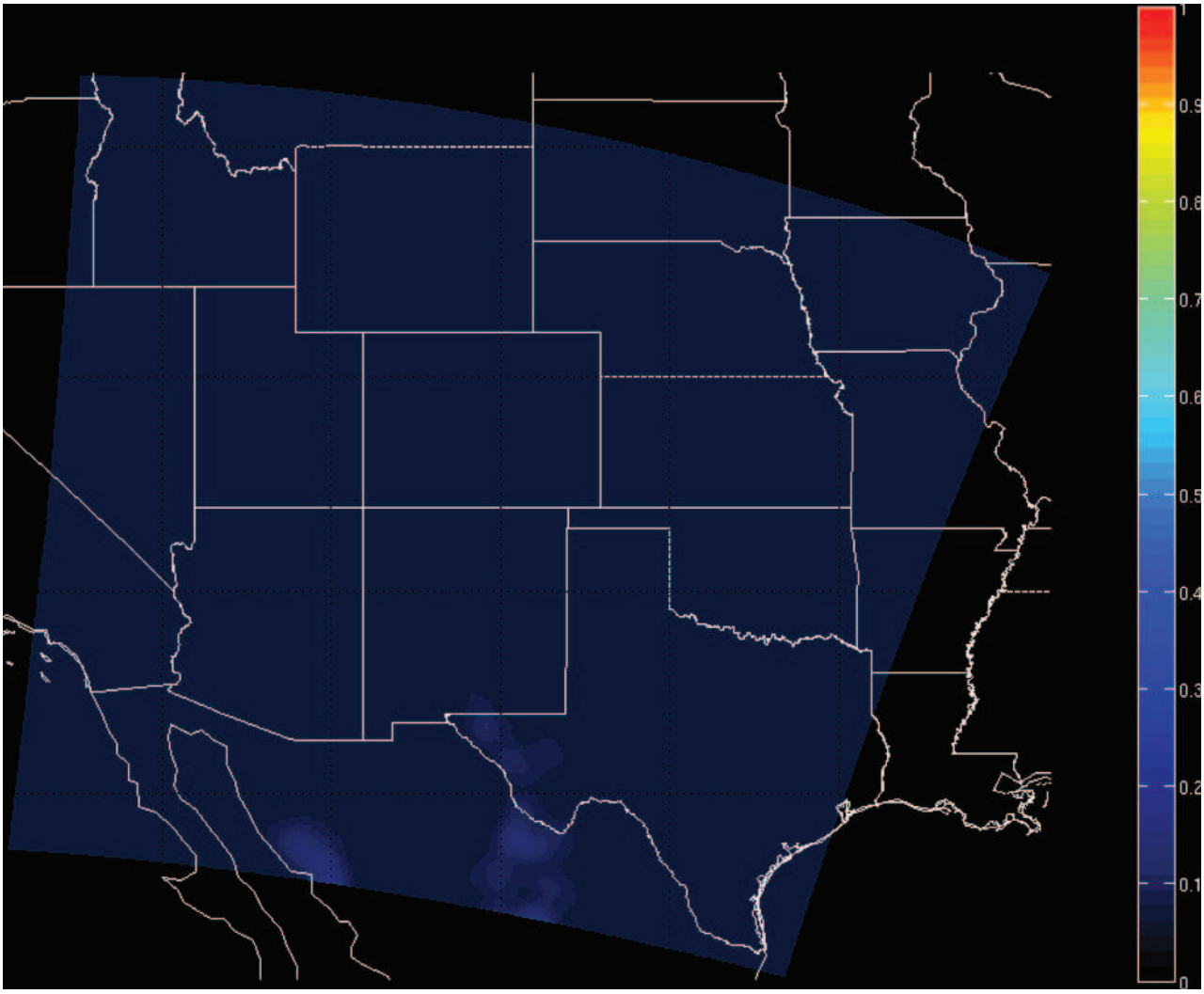

Figure 99. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

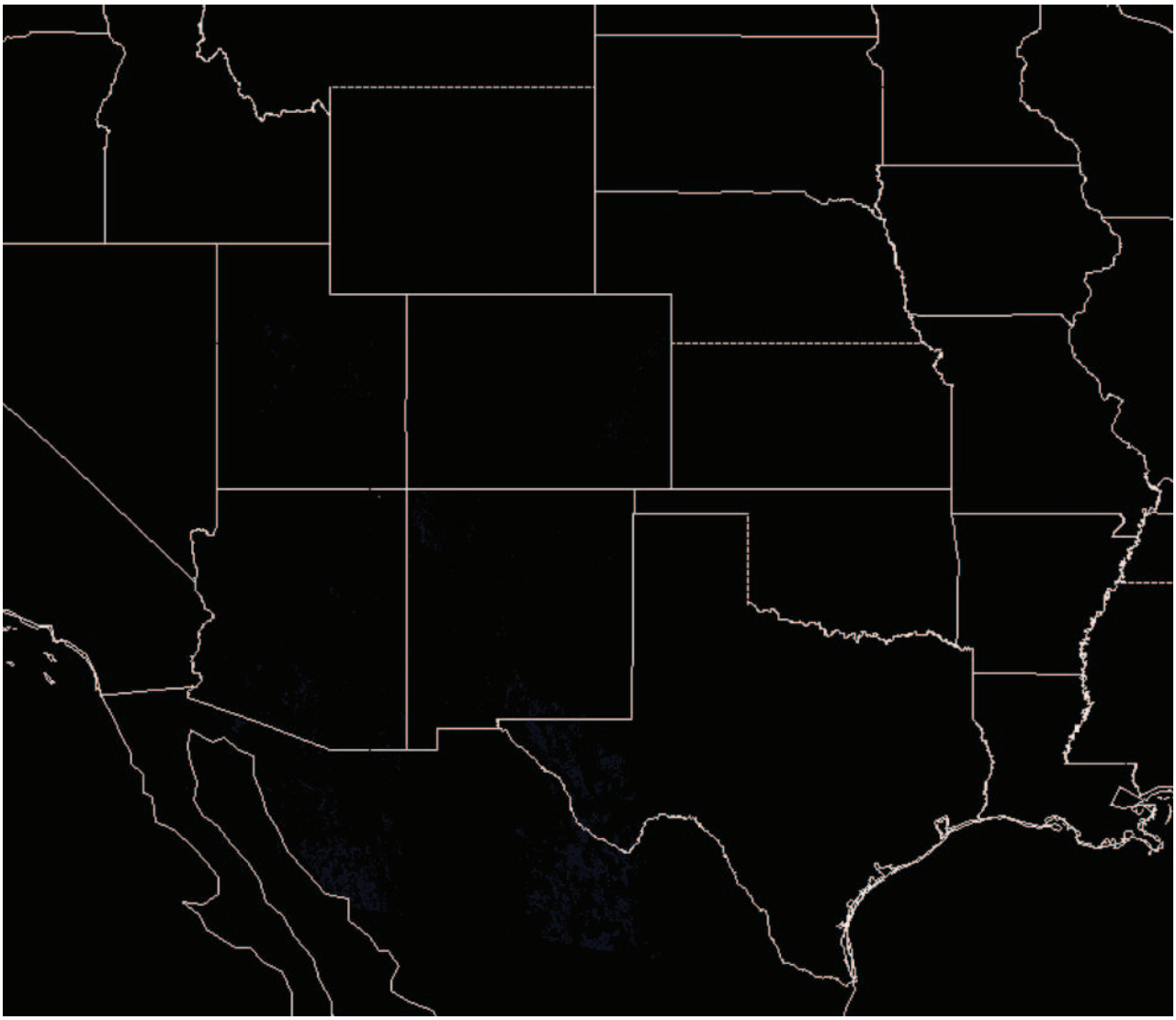

Figure 100. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2002-183 at 17:55 UTC.

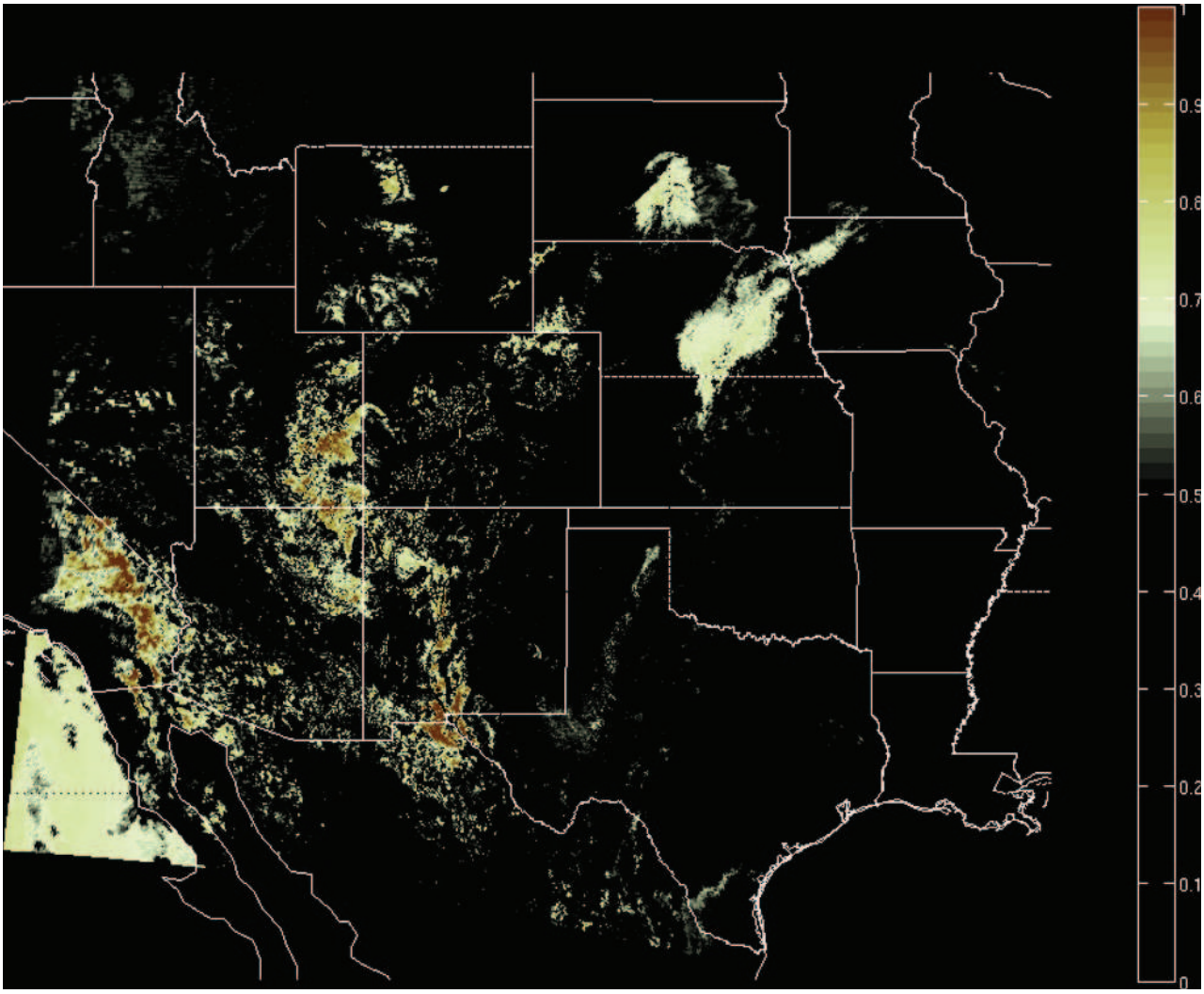

Figure 101. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

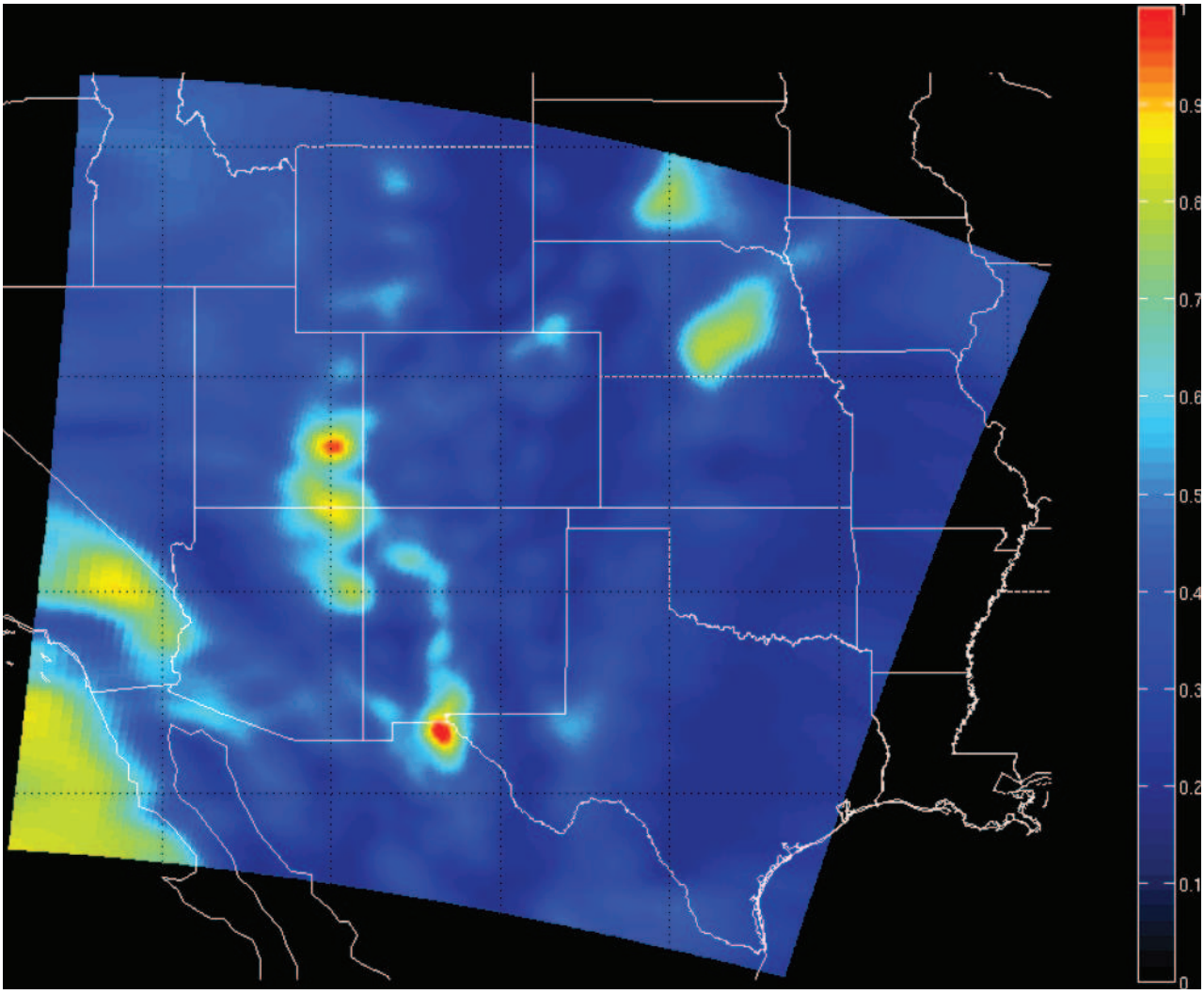

Figure 102. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

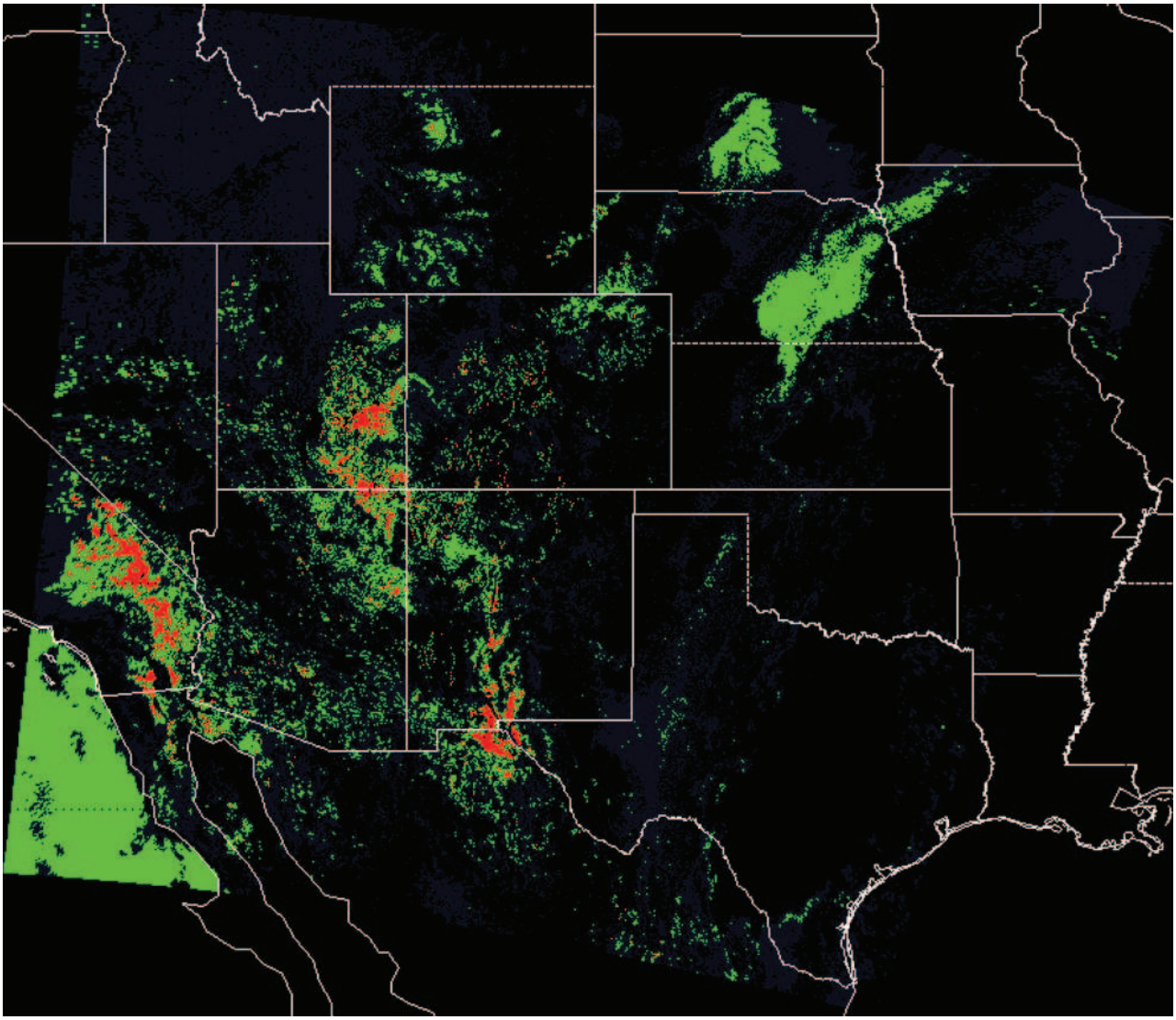

Figure 103. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-183 at 17:55 UTC.

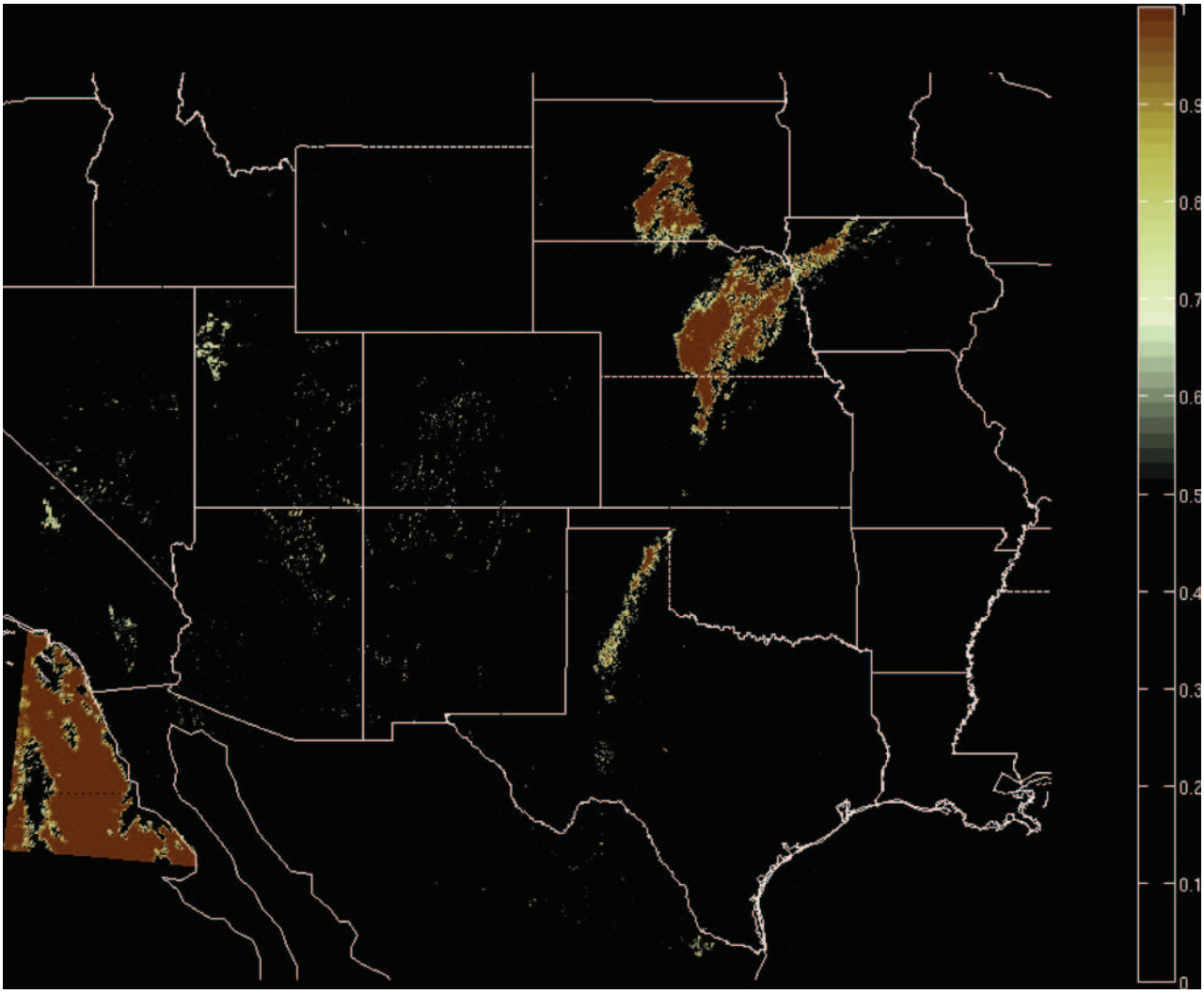

Figure 104. Result of the Probabilistic Neural Network (PNN) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

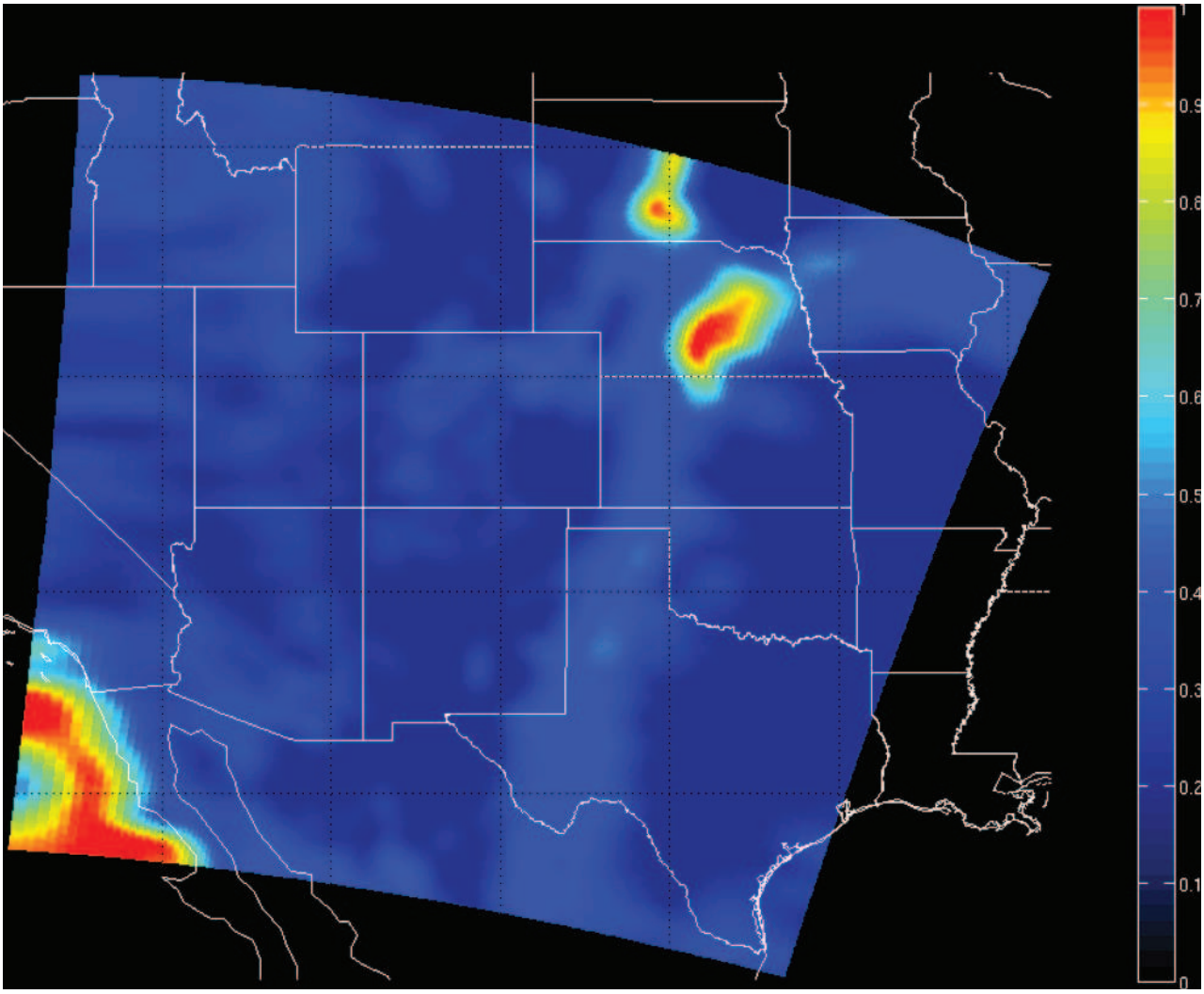

Figure 105. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2002-183 at 17:55 UTC. The colorbar shows the index of probability.

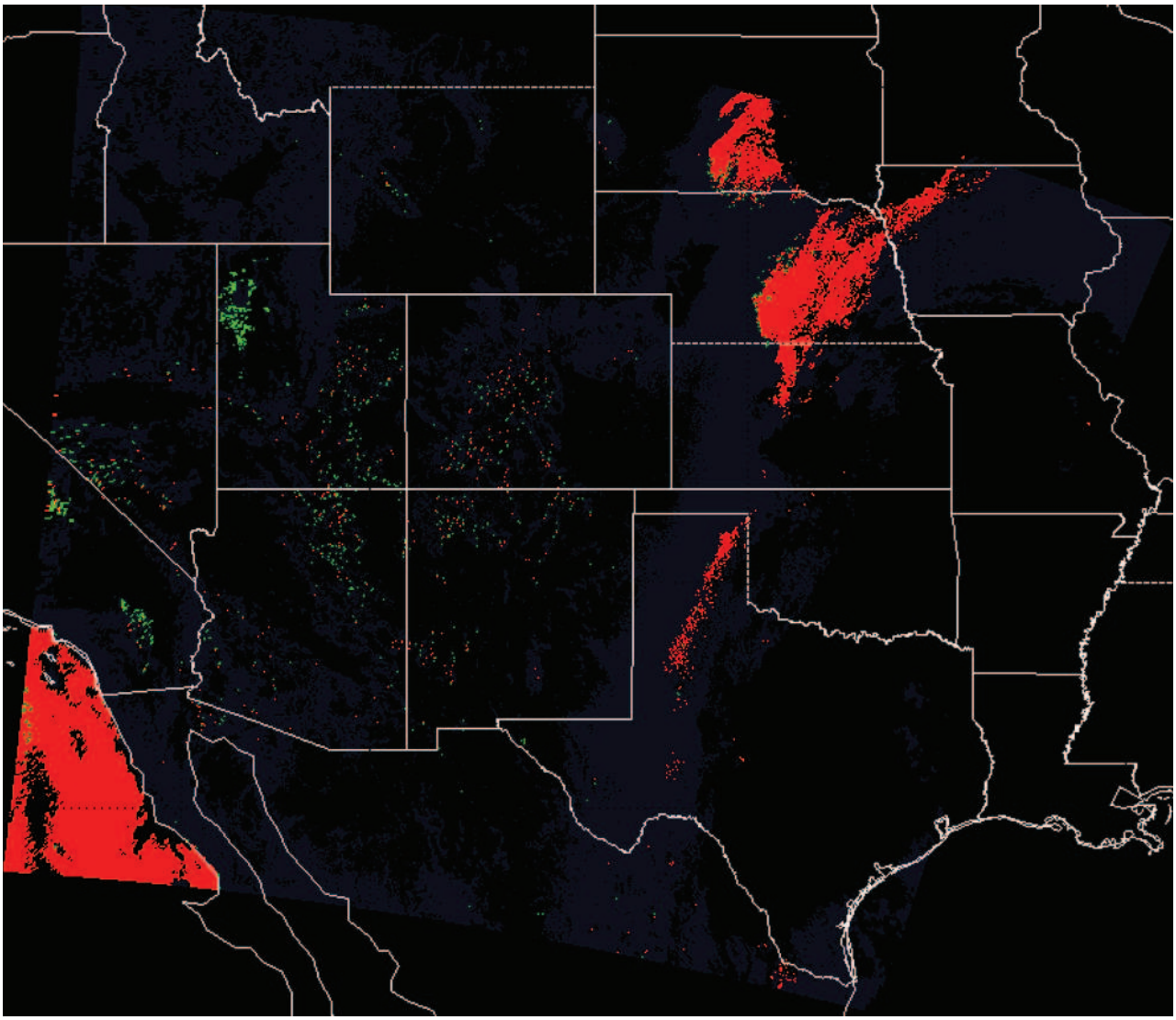

Figure 106. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2002-183 at 17:55 UTC.

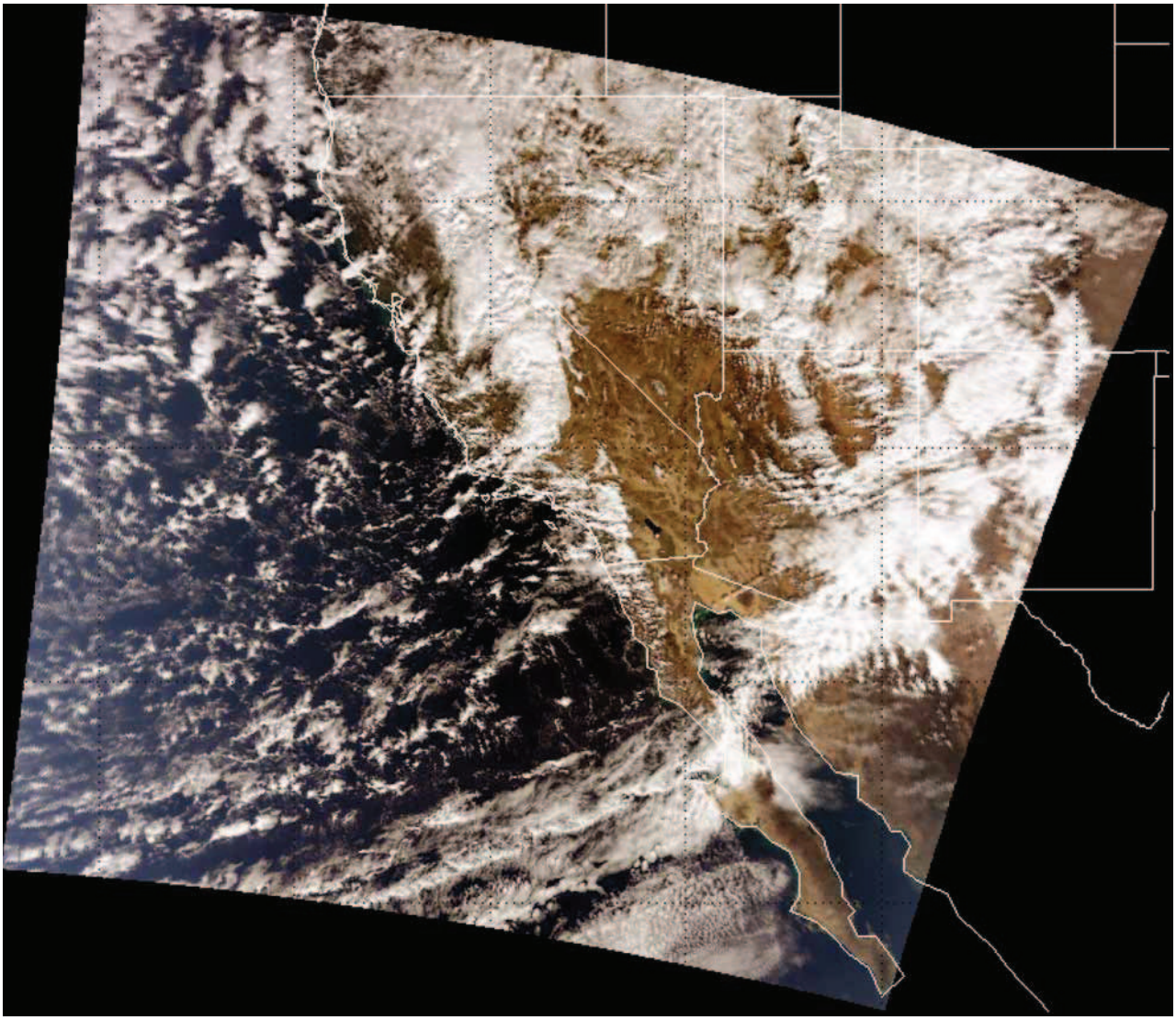

Figure 107. True color image for the event of 2002-351 at 18:45 UTC.

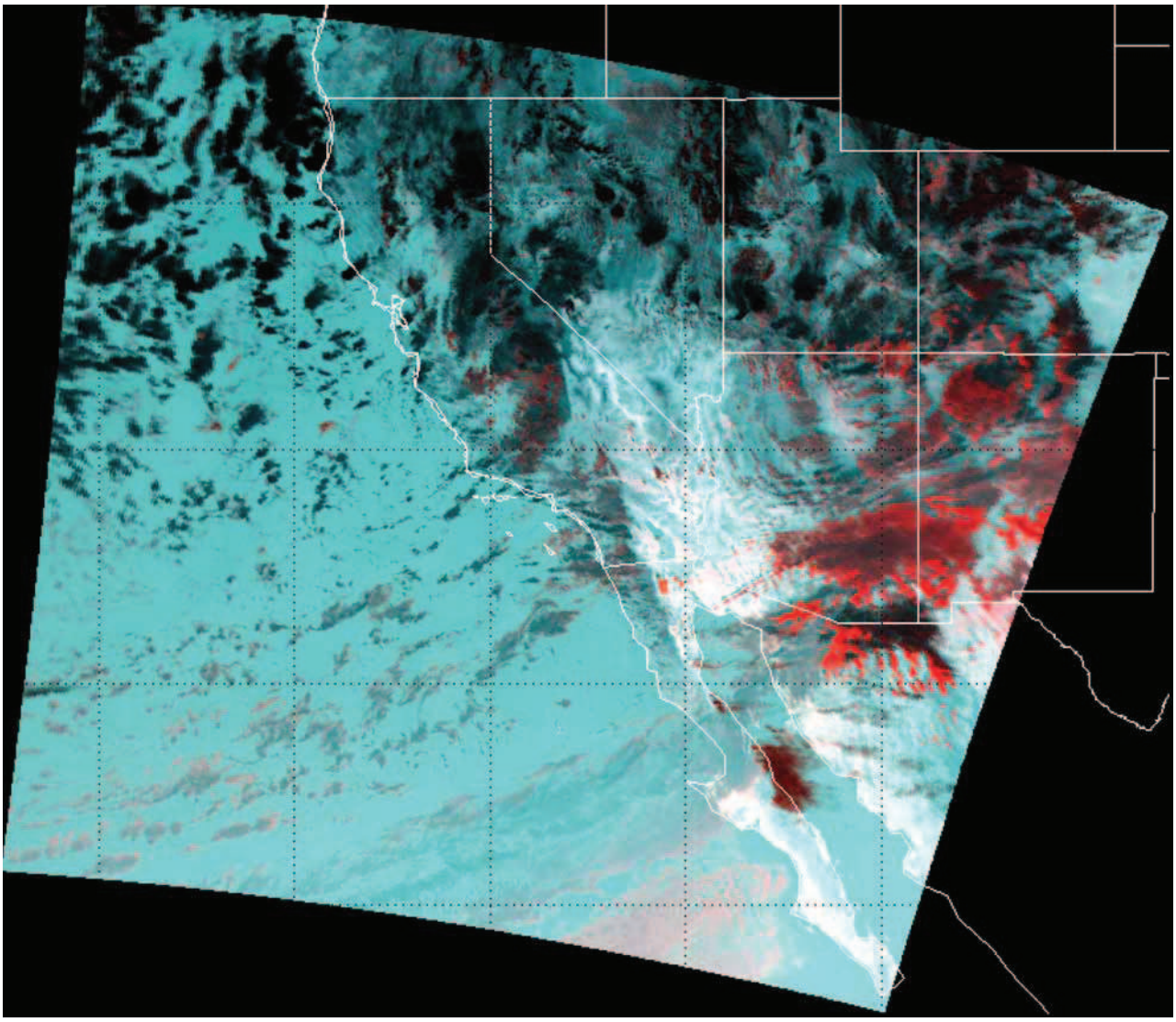

Figure 108. False color image for the event of 2002-351 at 18:45 UTC.

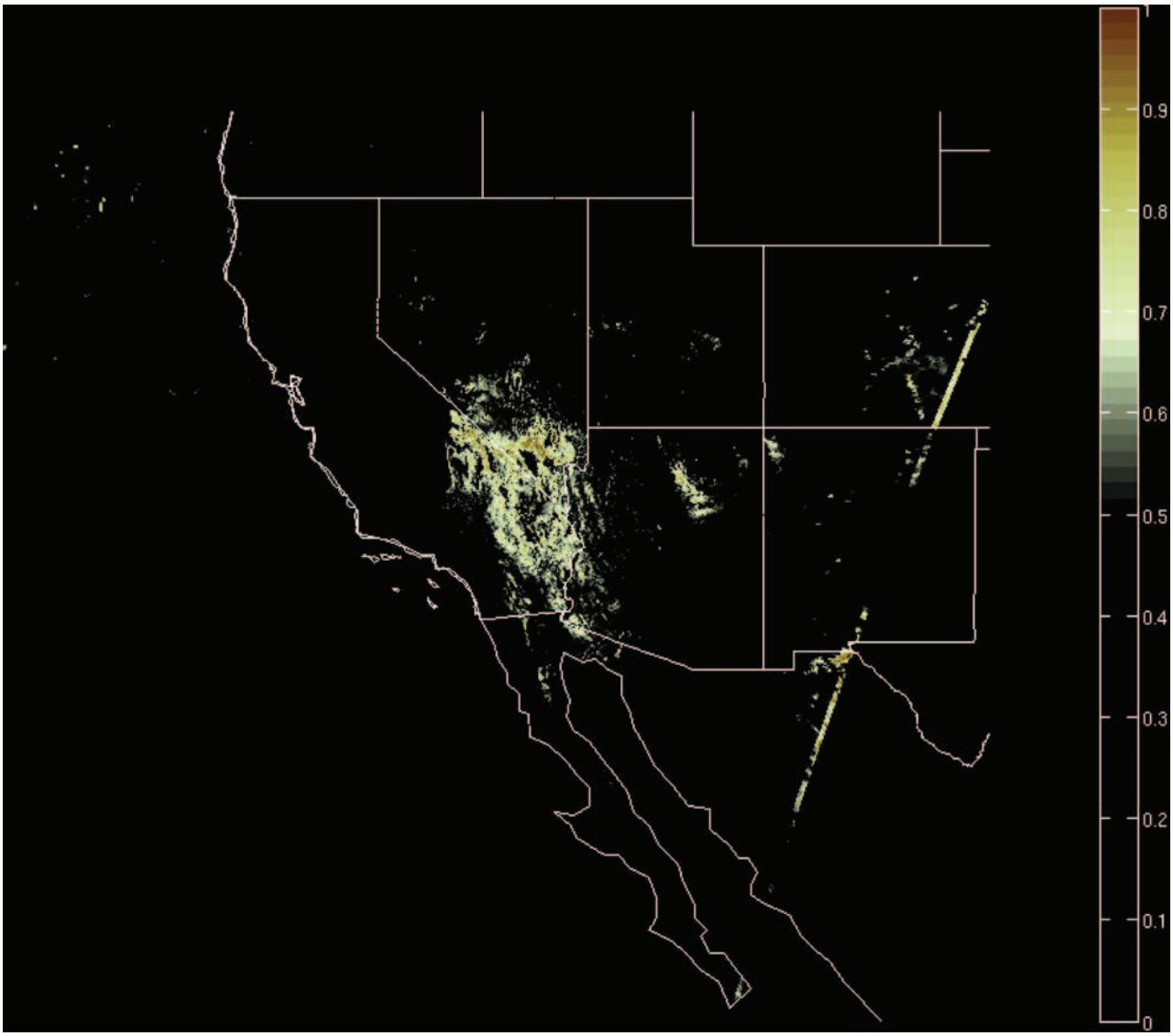

Figure 109. Result of the probabilistic display model (PD) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

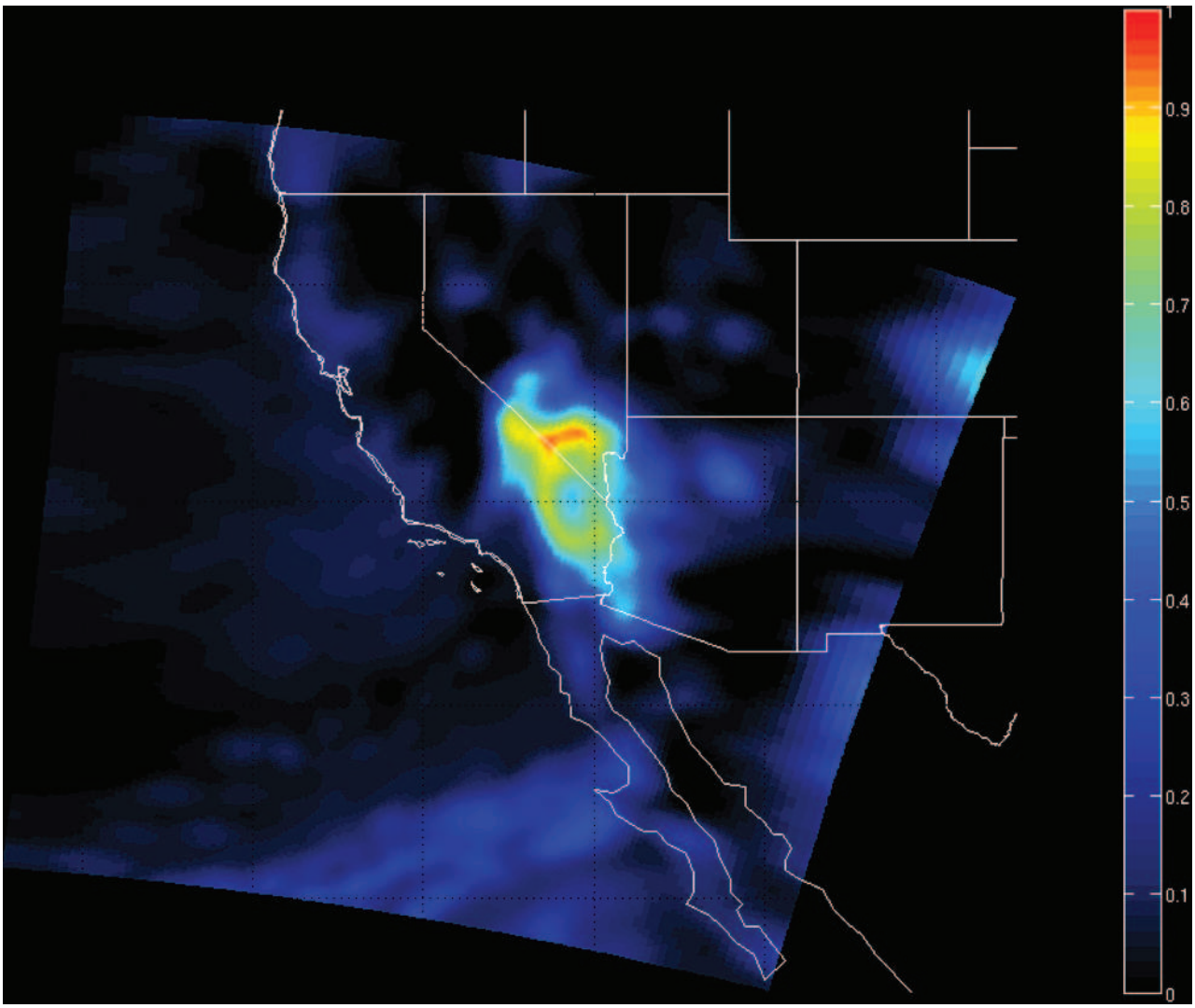

Figure 110. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

Figure 111. Result of the Maximum Likelihood model (MLE) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

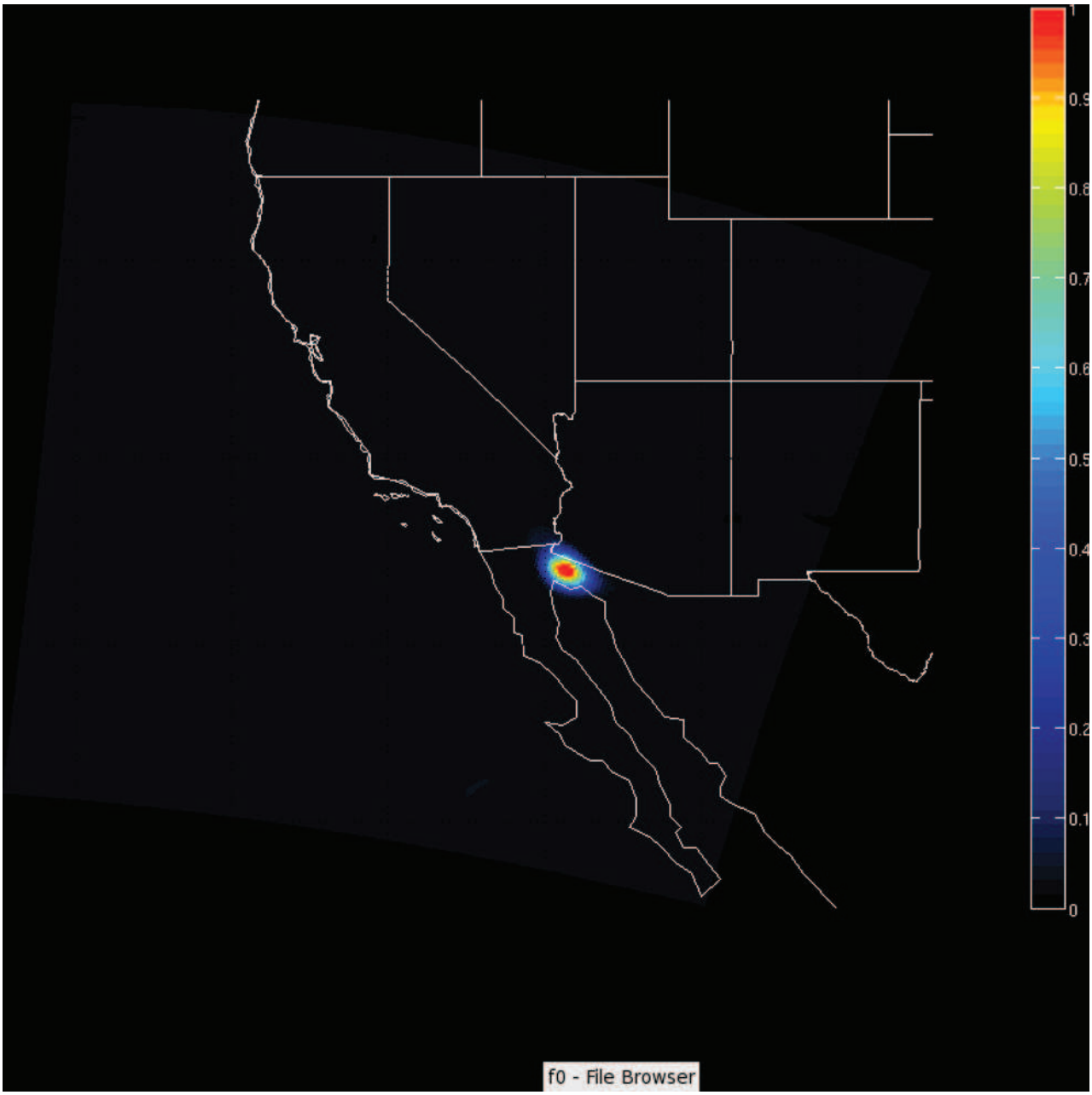

Figure 112. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

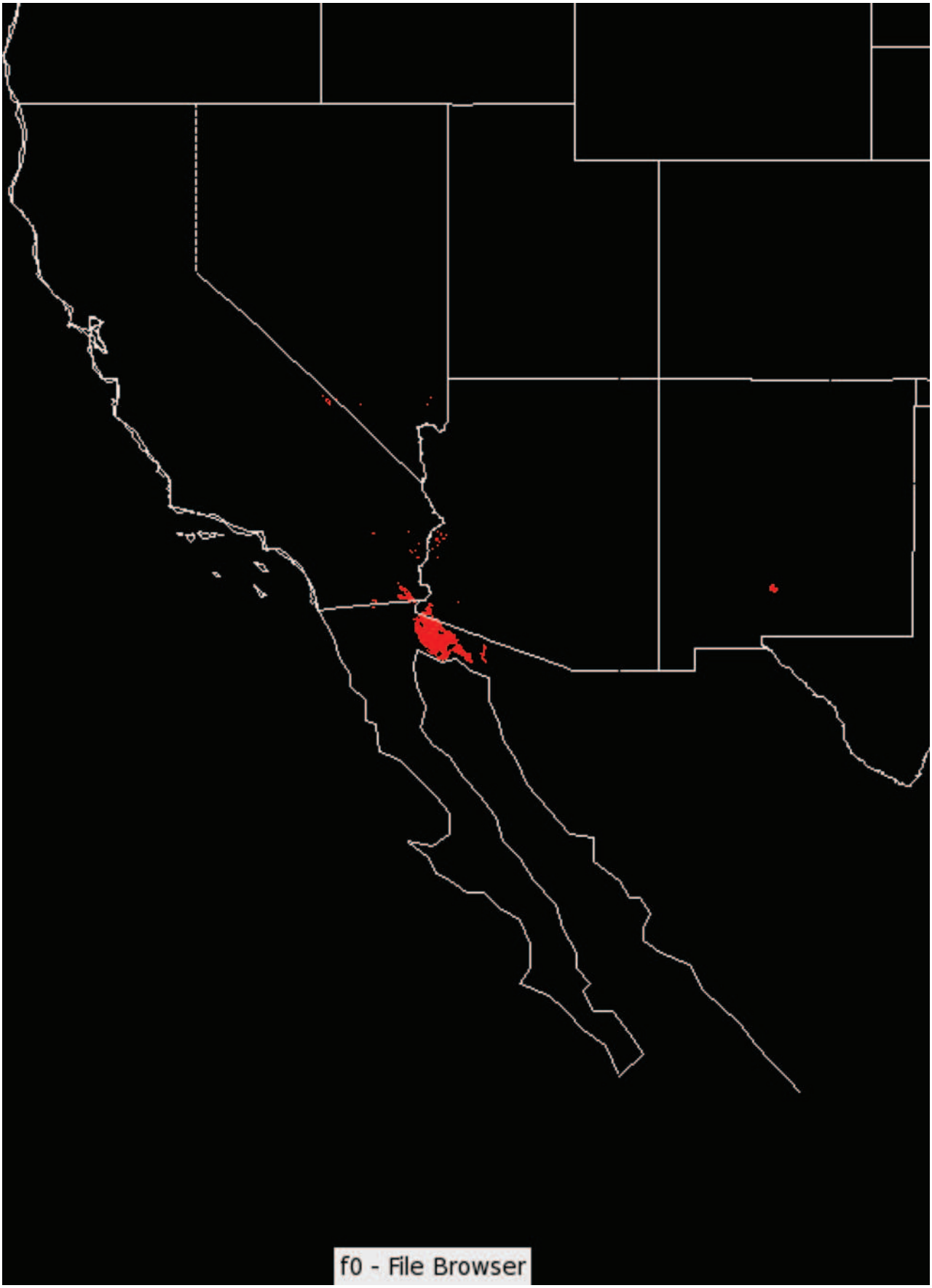

Figure 113. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2002-351 at 18:45 UTC.

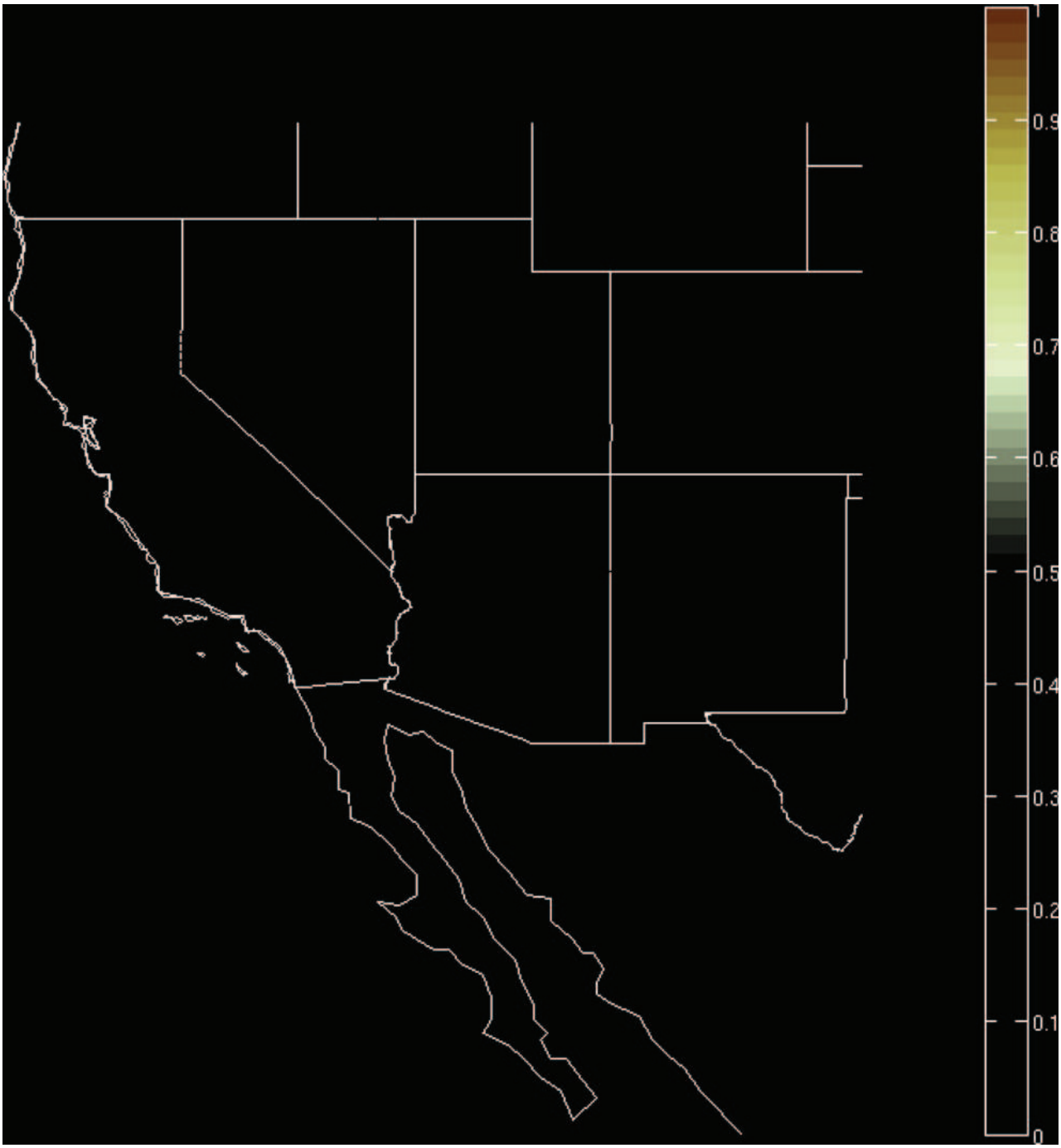

Figure 114. Result of the Maximum a posteriori model (MAP) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

Figure 115. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

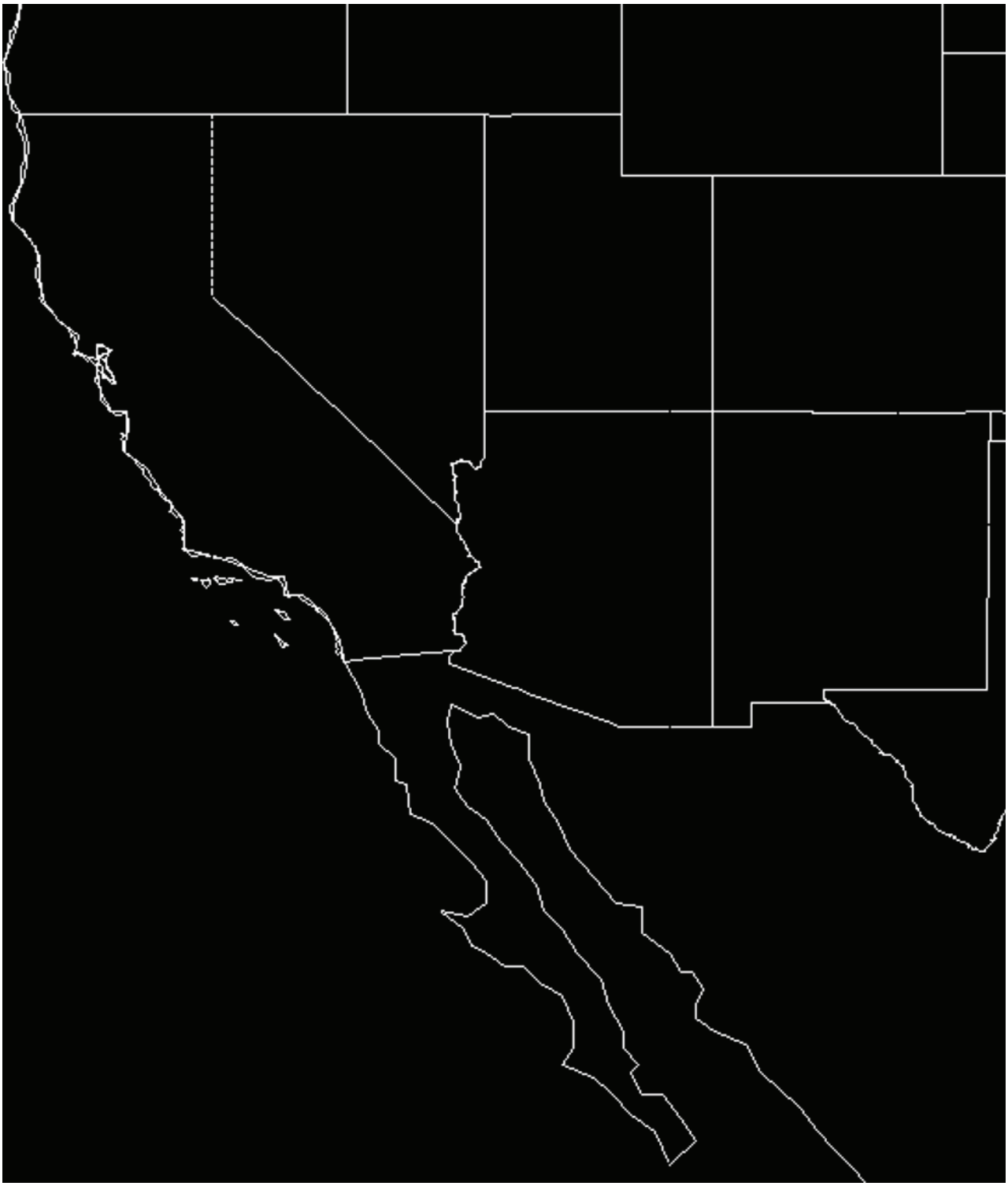

Figure 116. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2002-351 at 18:45 UTC.

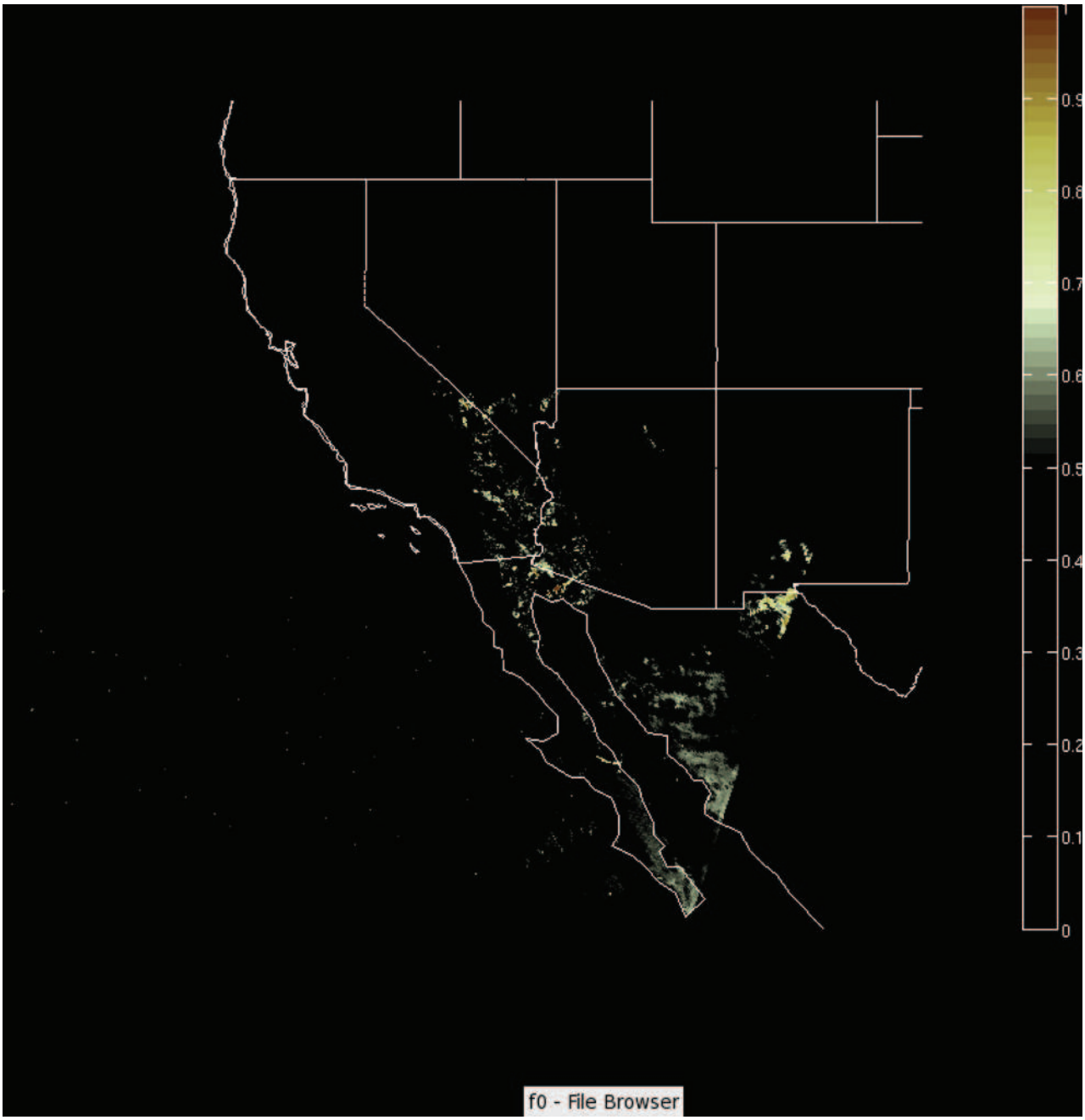

Figure 117. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

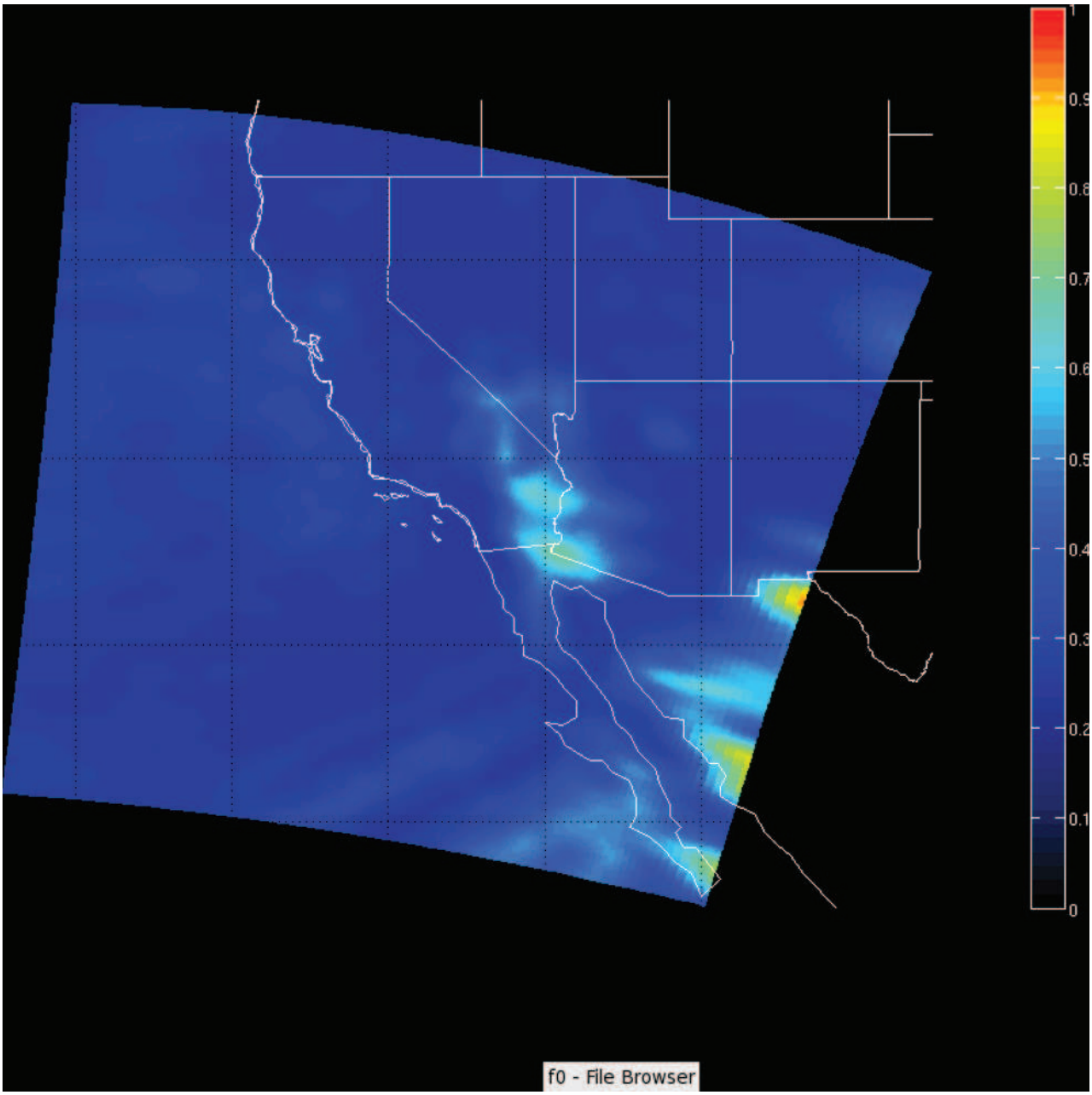

Figure 118. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

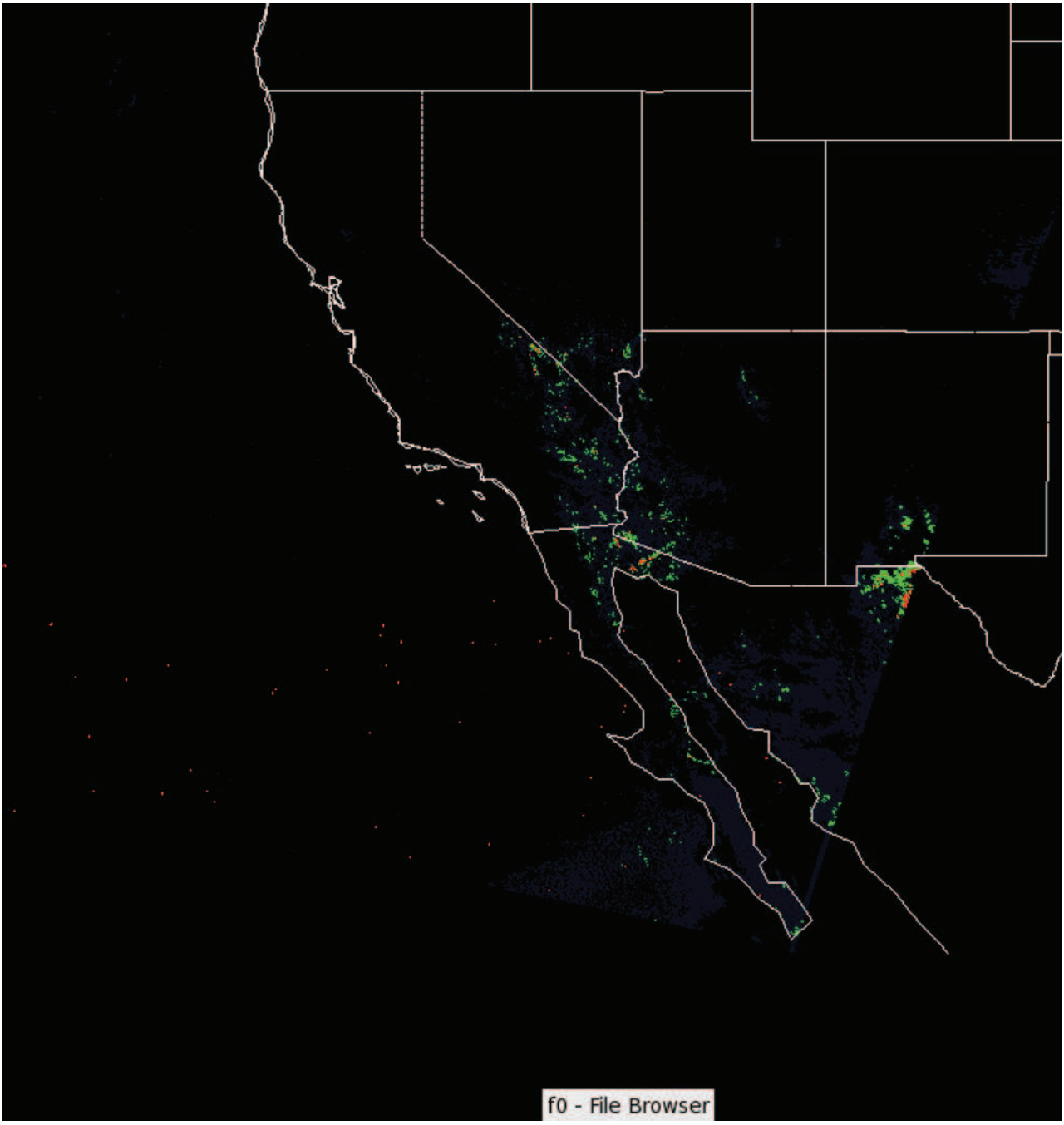

Figure 119. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-351 at 18:45 UTC.

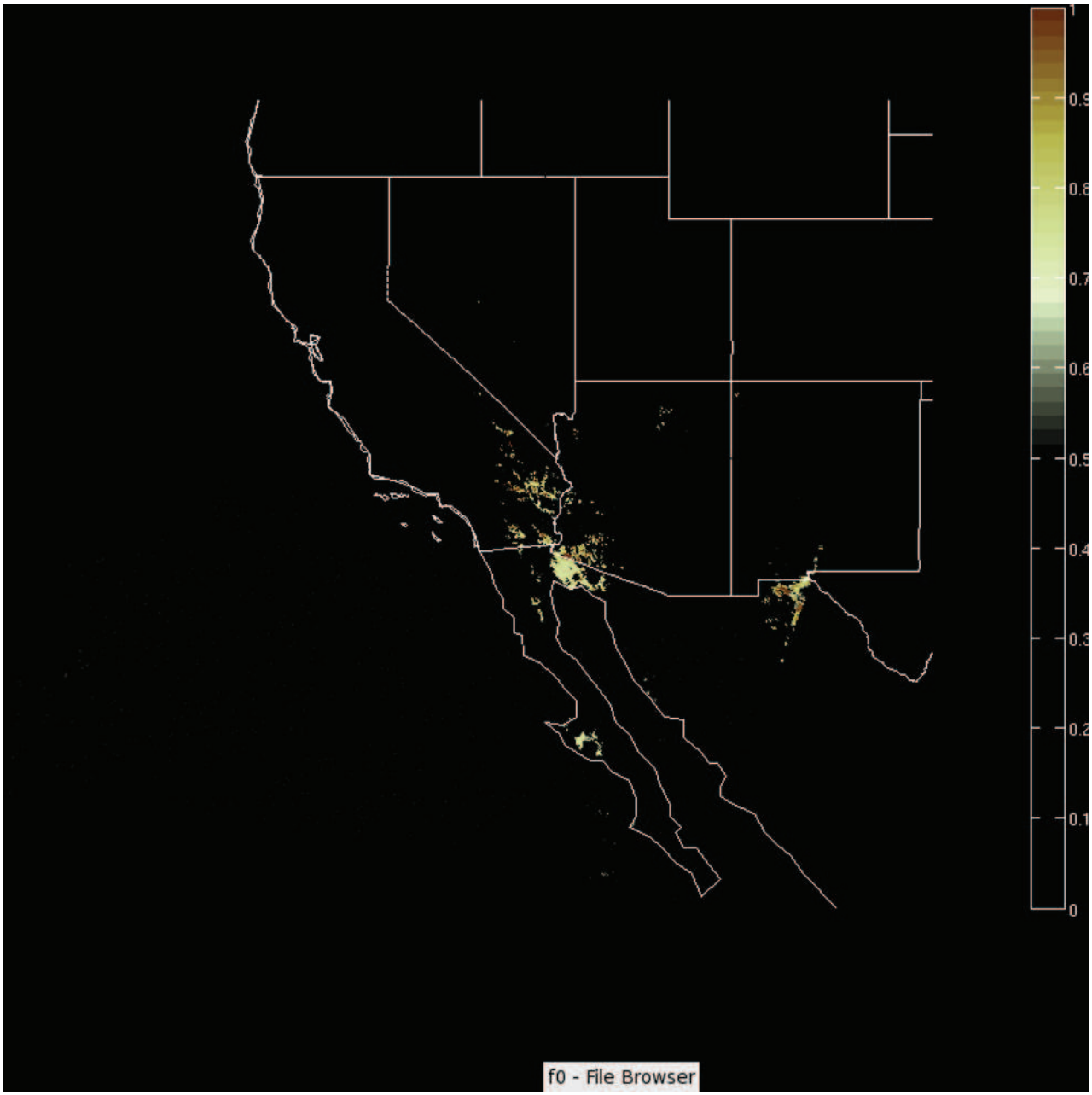

Figure 120. Result of the Probabilistic Neural Network (PNN) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

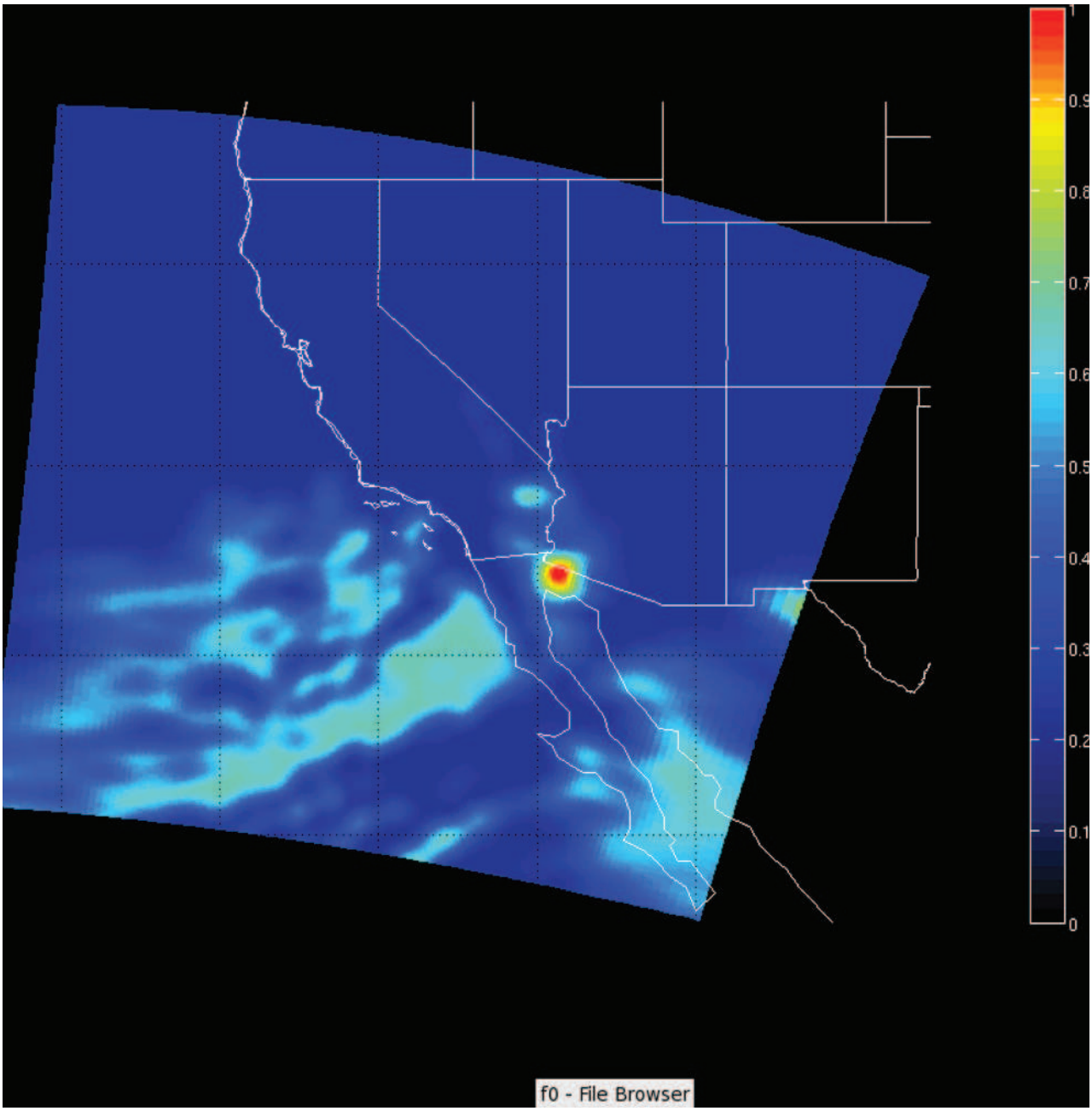

Figure 121. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2002-351 at 18:45 UTC. The colorbar shows the index of probability.

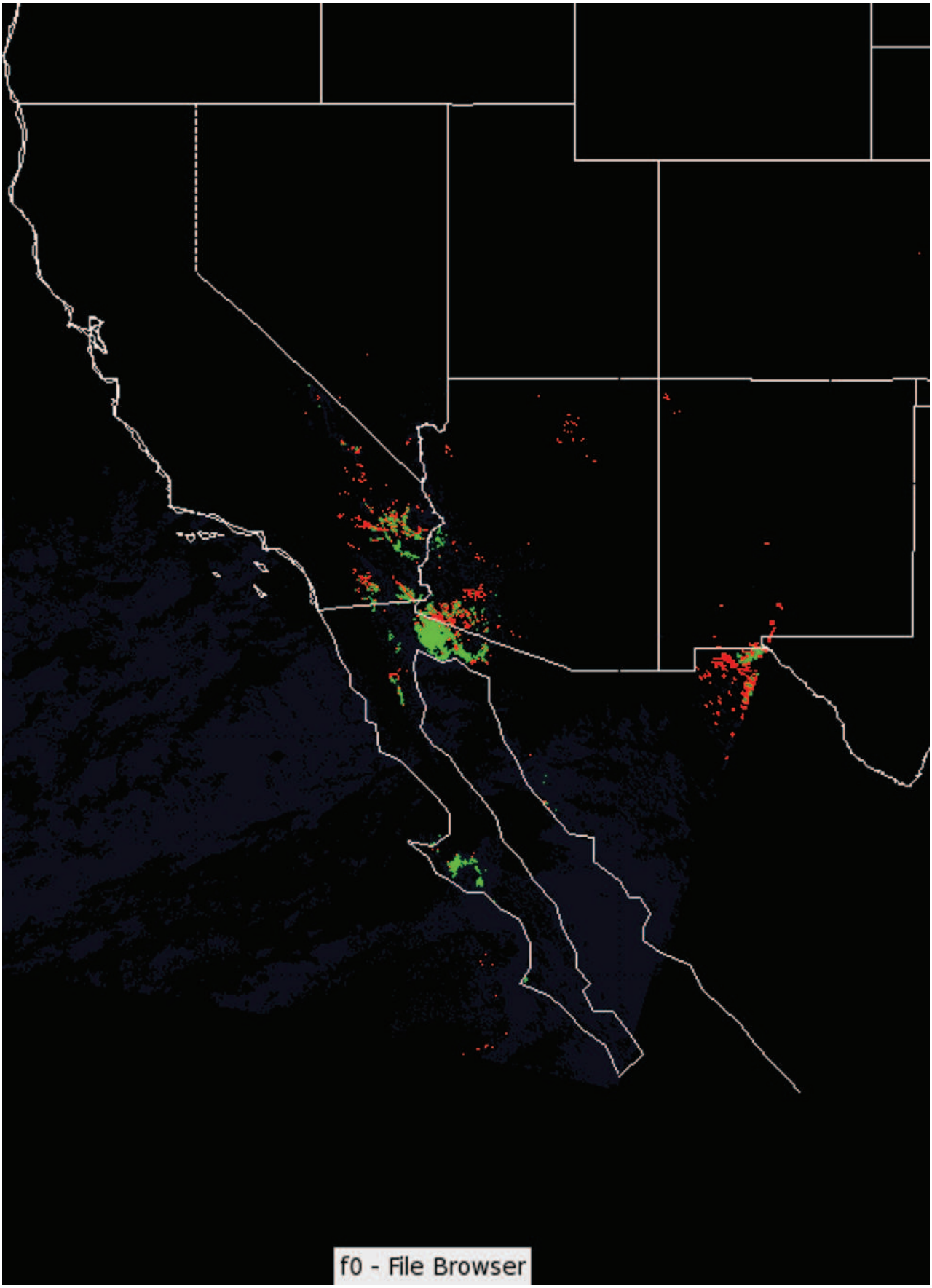

Figure 122. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2002-351 at 18:45 UTC.

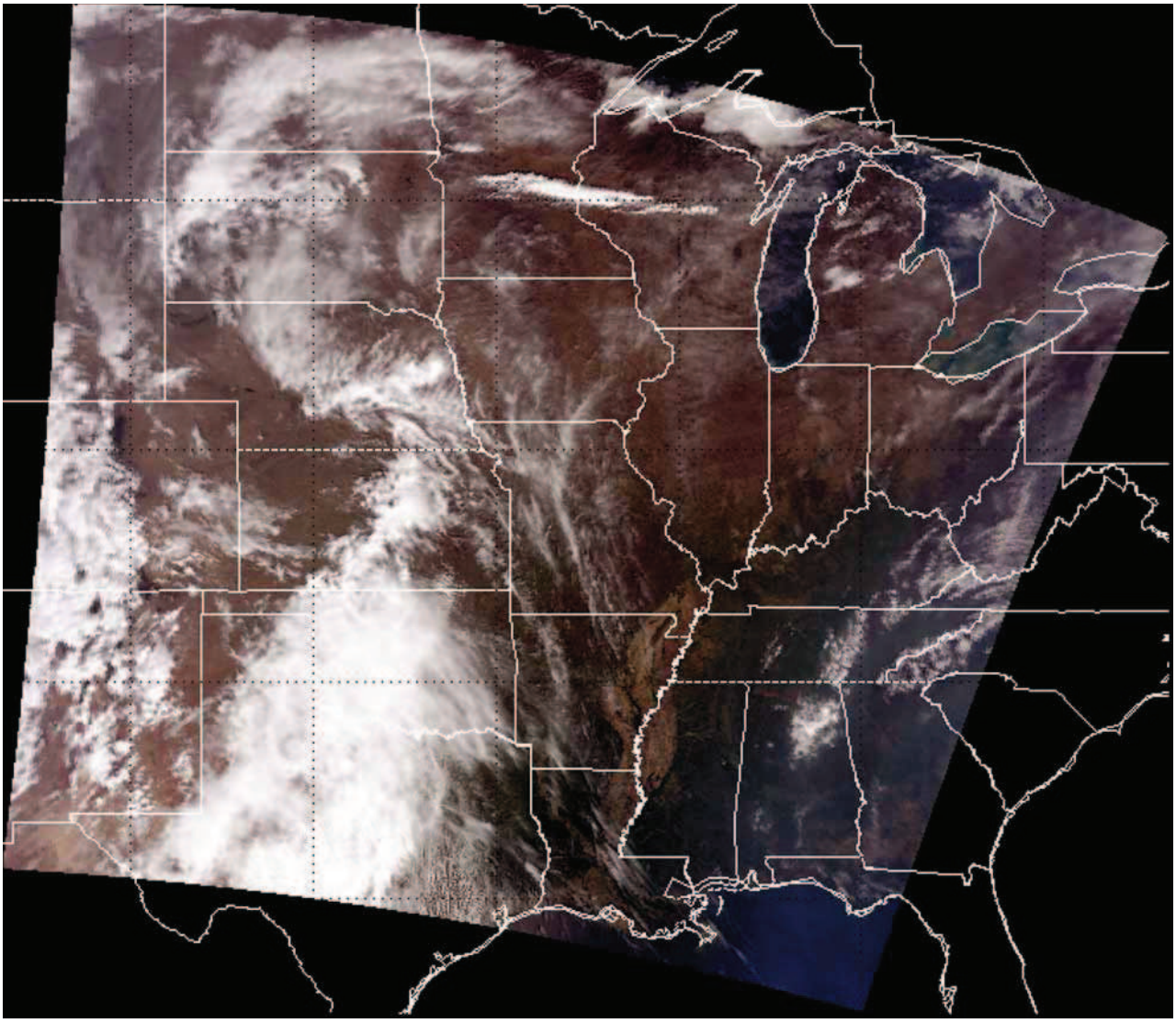

Figure 123. True color image for the event of 2003-105 at 17:10 UTC.

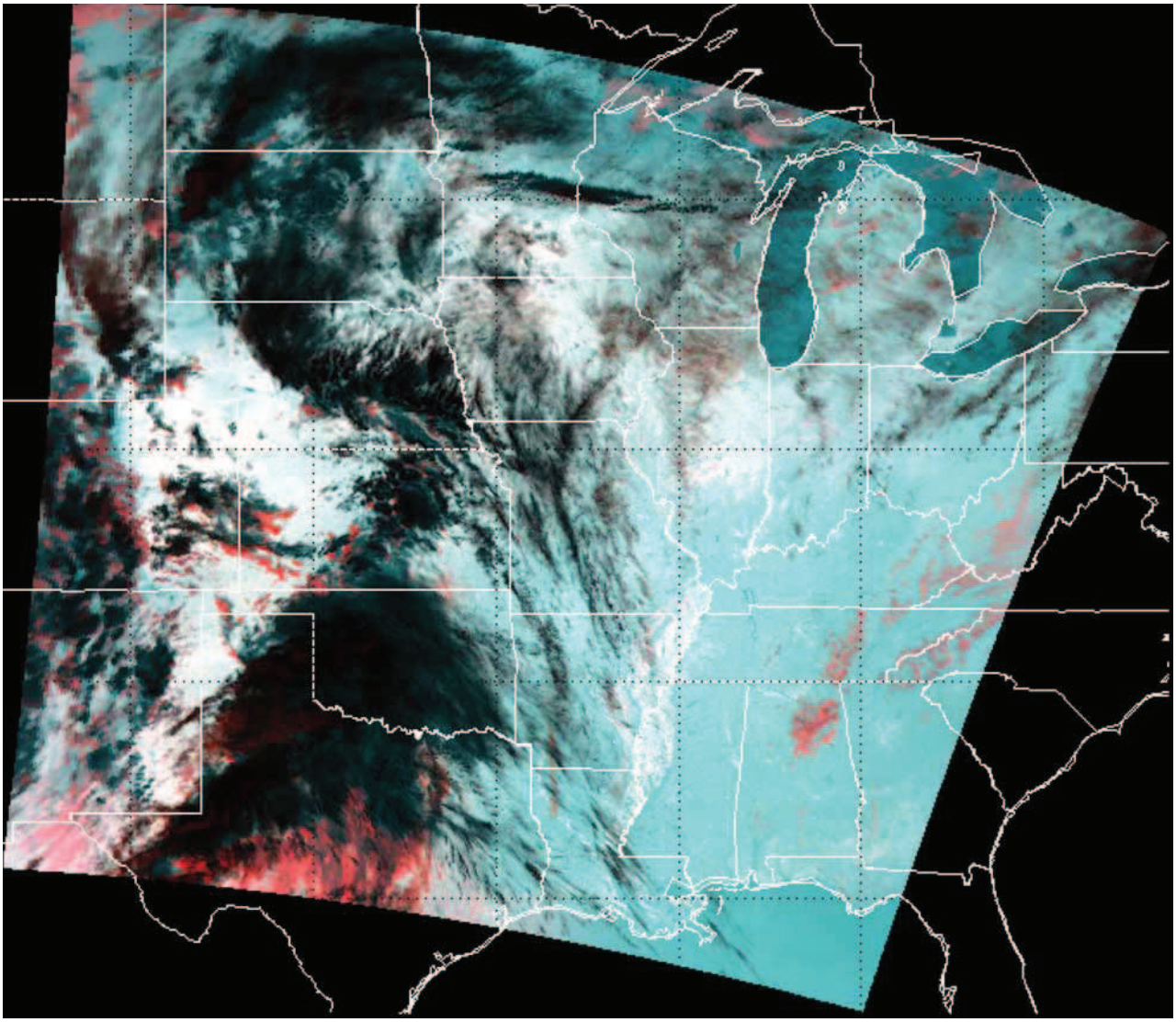

Figure 124. False color image for the event of 2003-105 at 17:10 UTC.

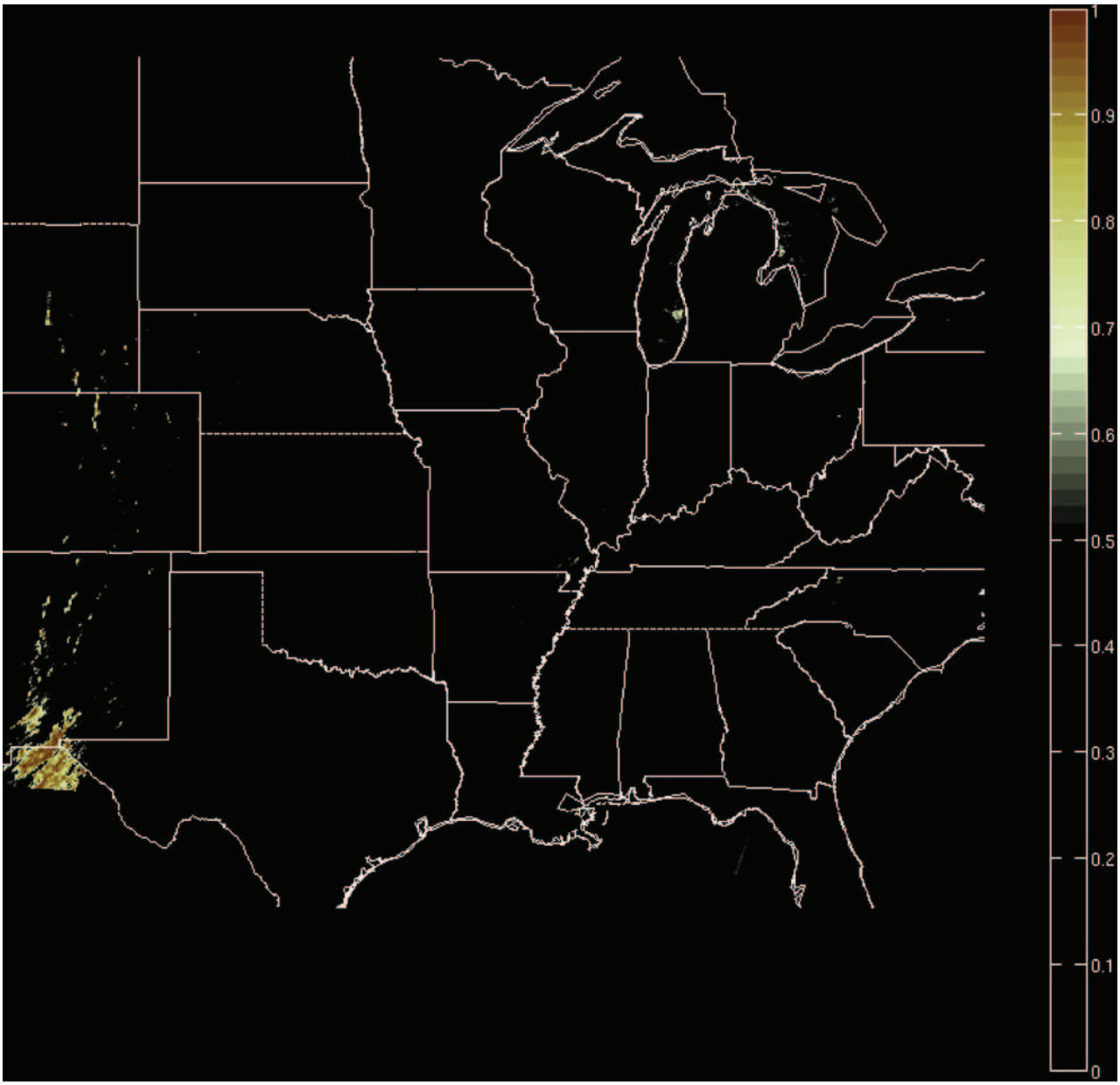

Figure 125. Result of the probabilistic display model (PD) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

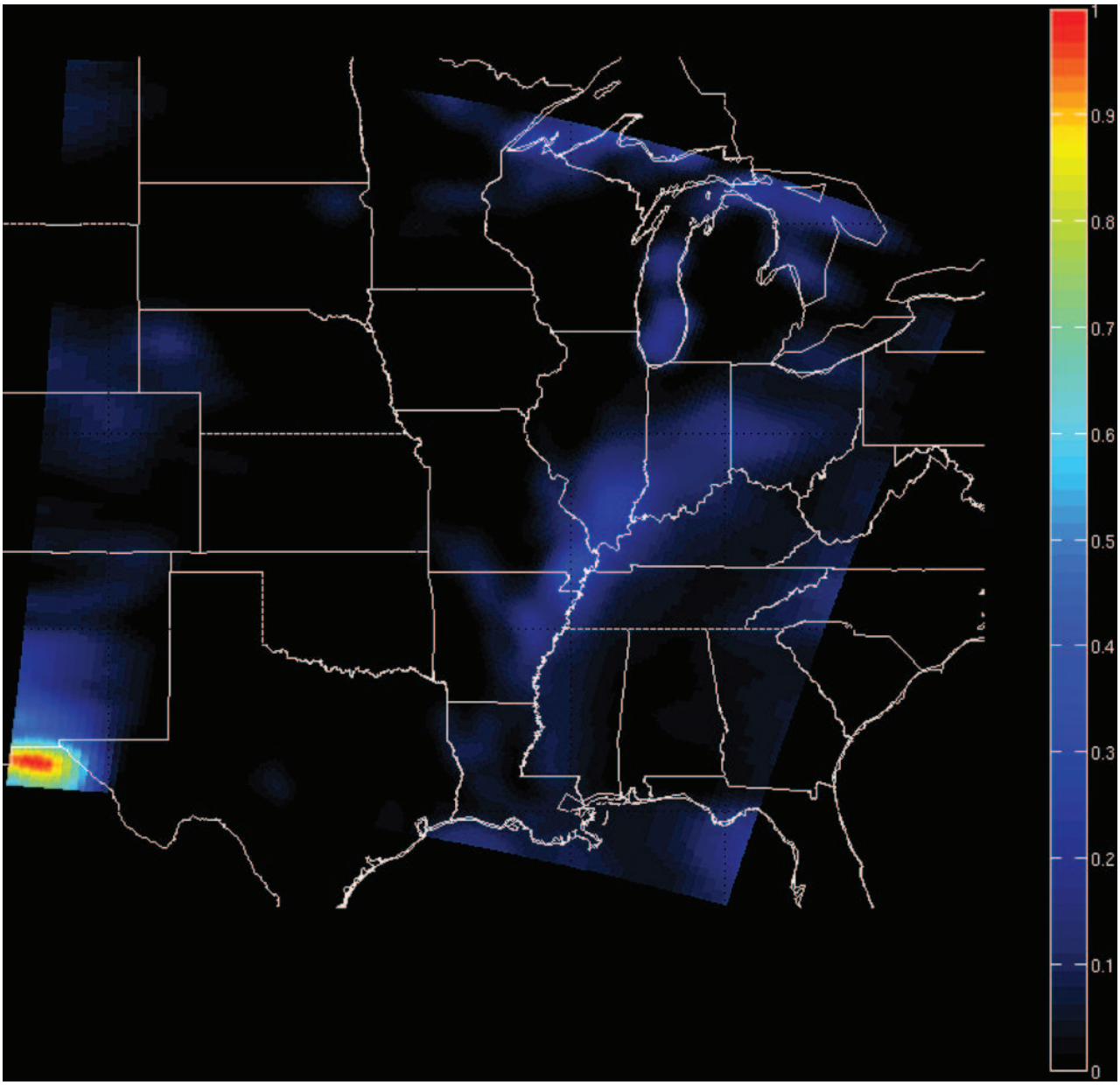

Figure 126. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

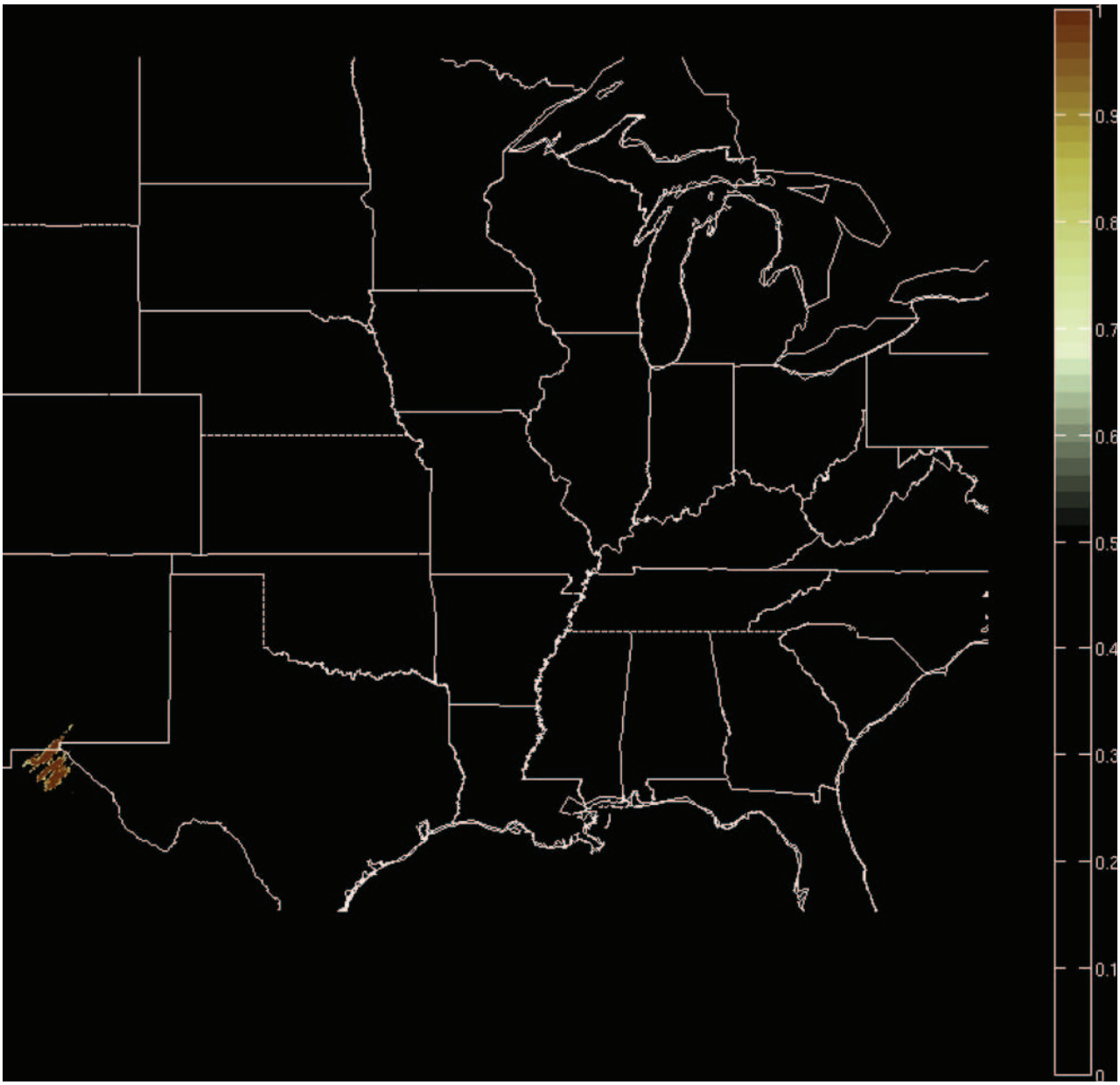

Figure 127. Result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

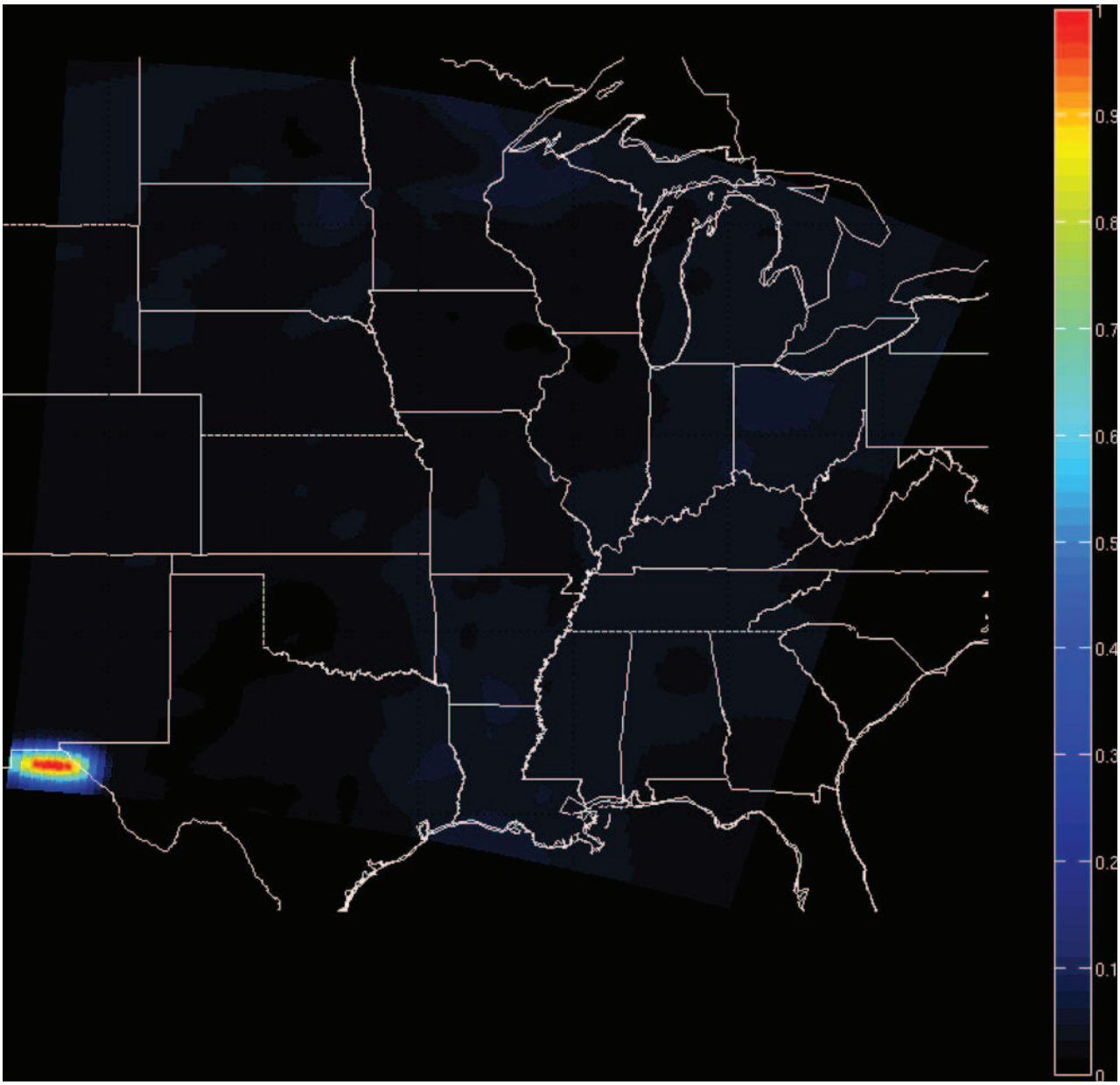

Figure 128. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

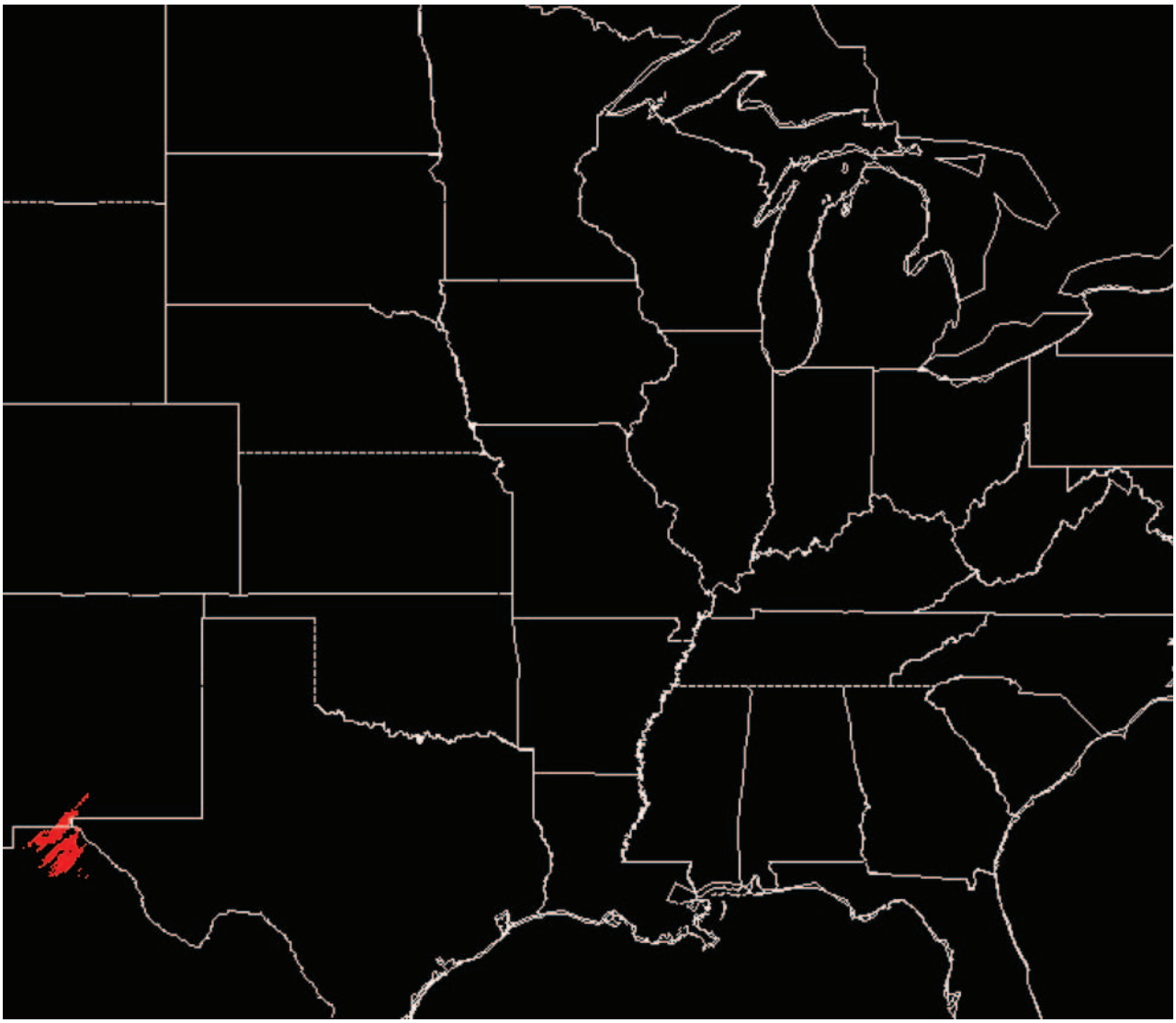

Figure 129. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 17:10 UTC.

Figure 130. Result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

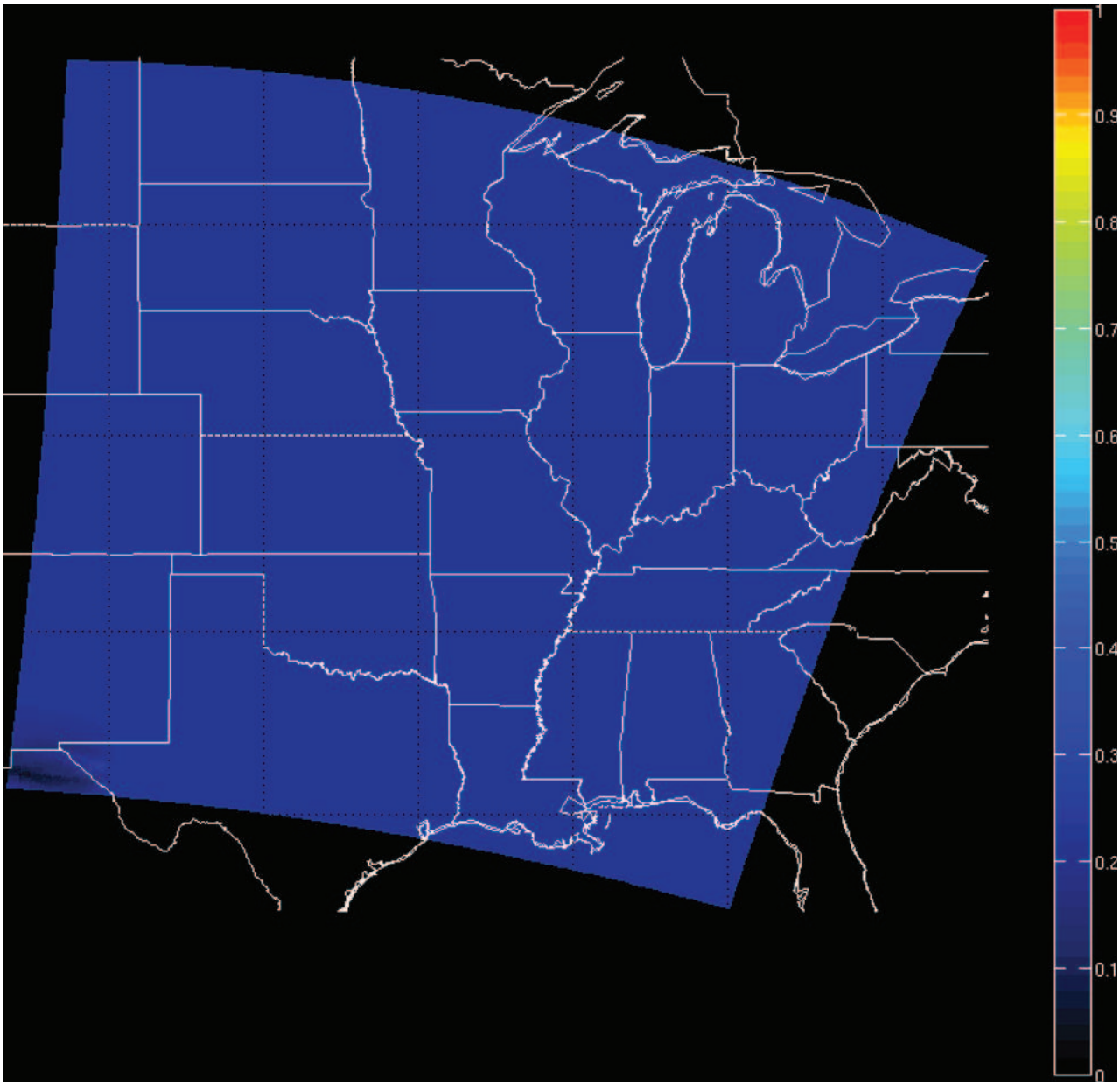

Figure 131. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

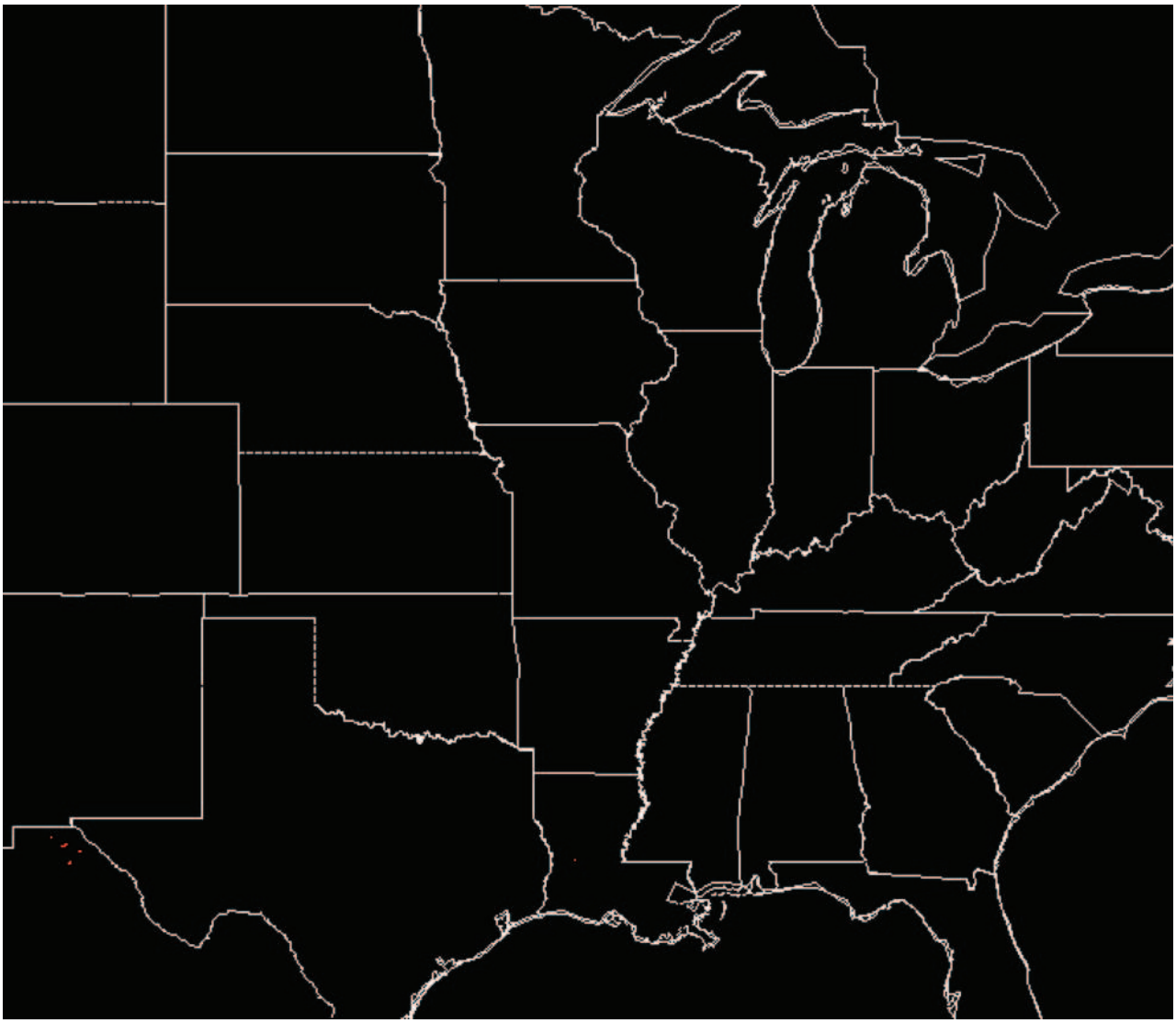

Figure 132. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 17:10 UTC.

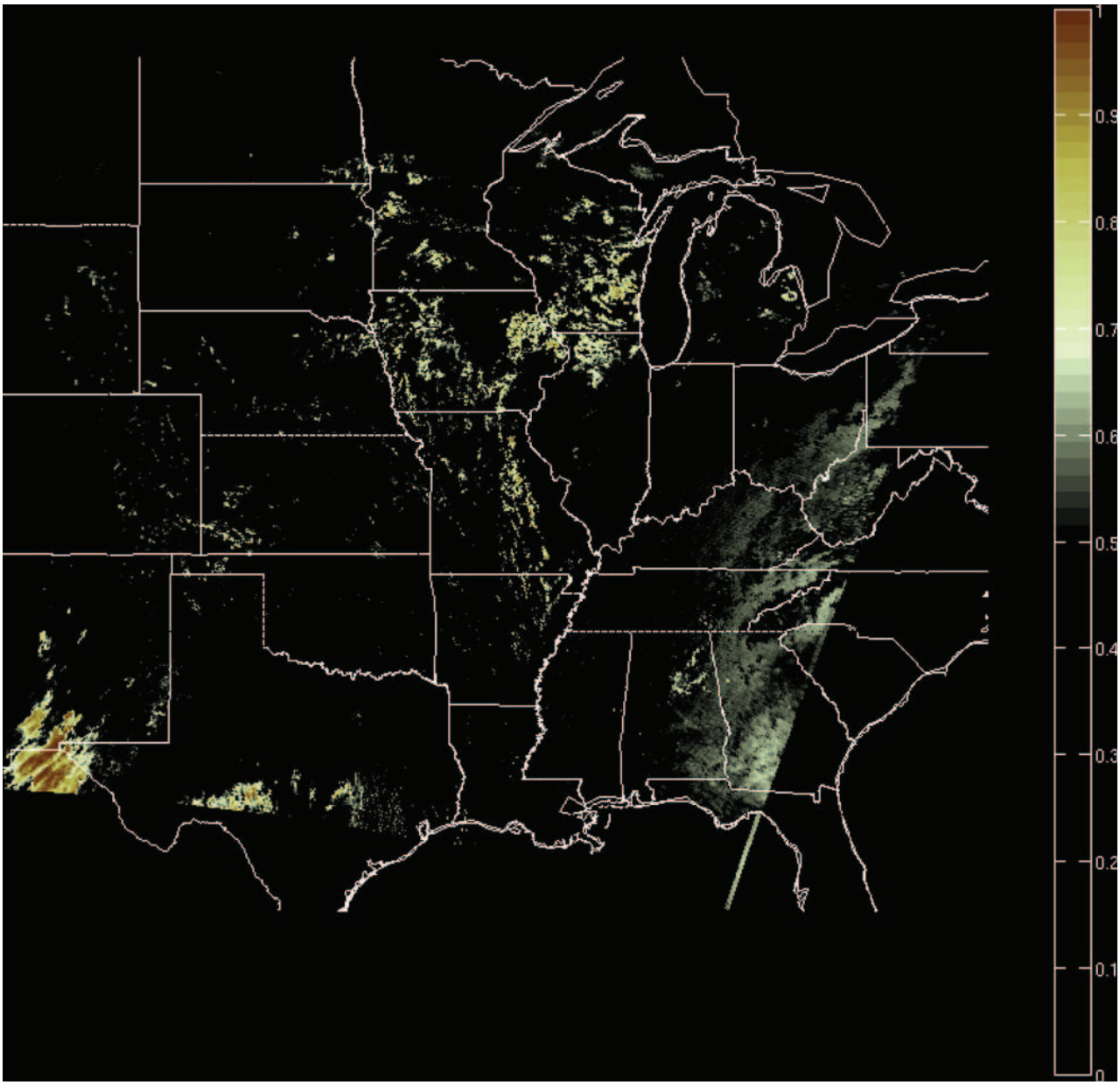

Figure 133. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

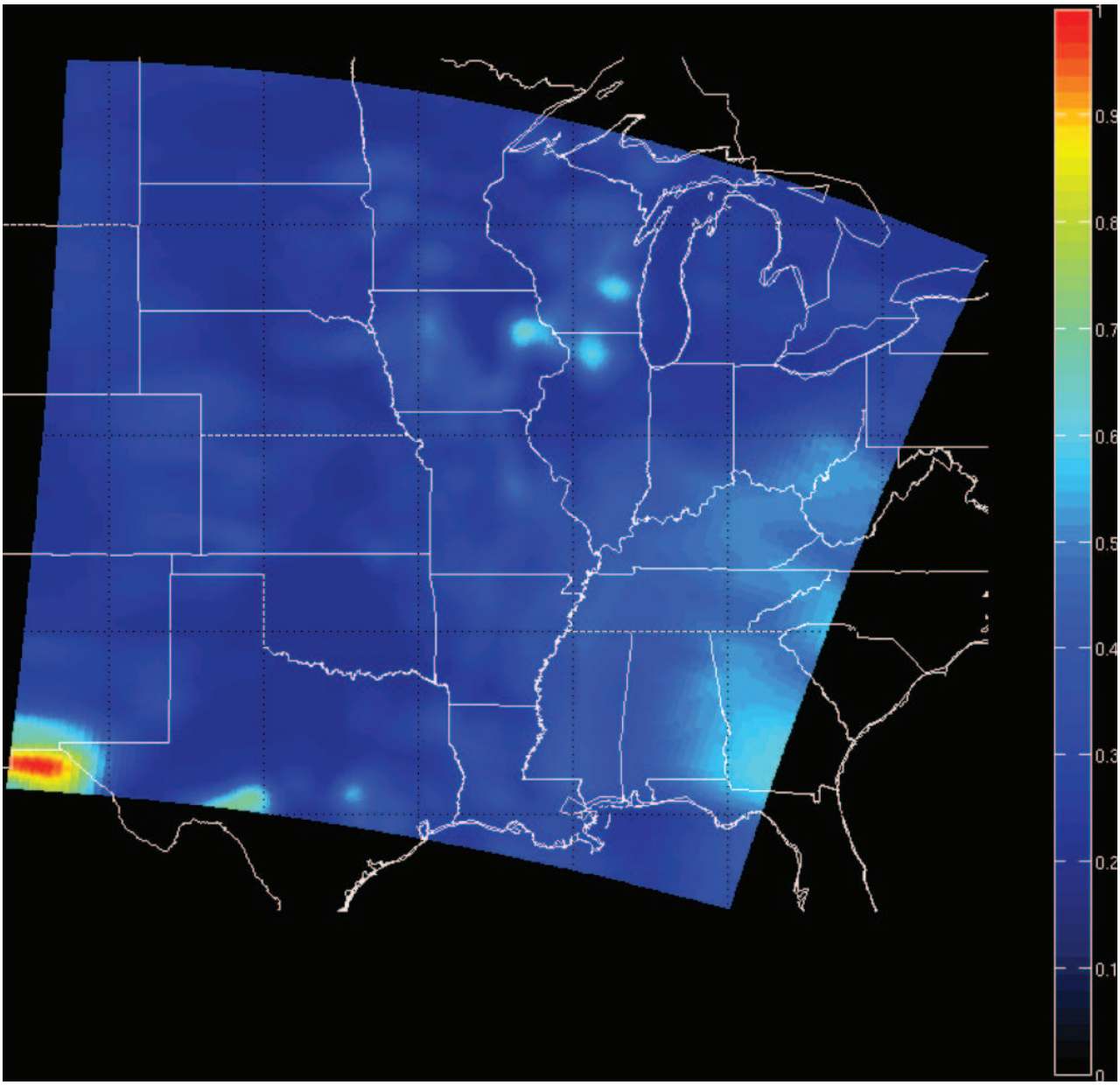

Figure 134. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

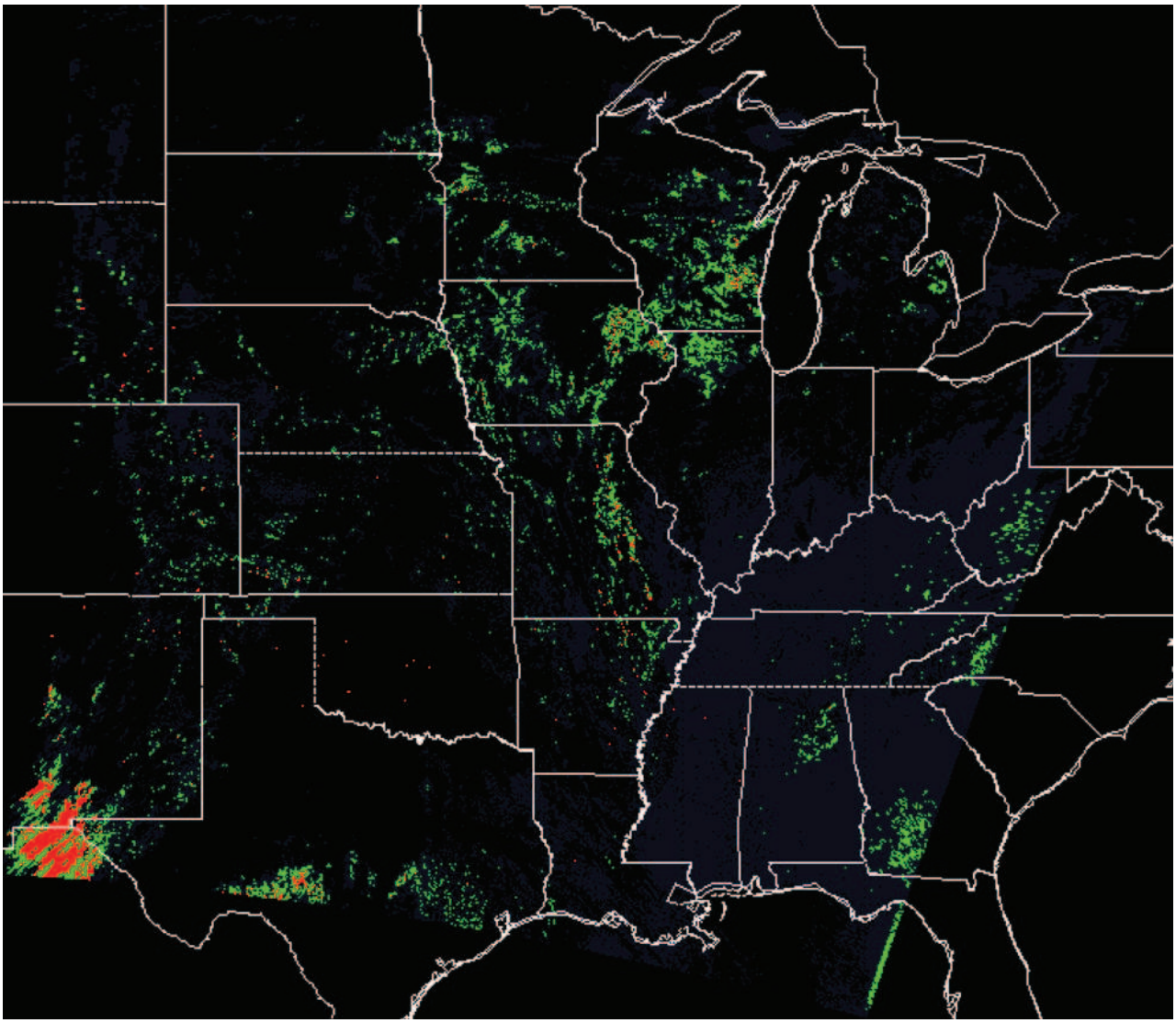

Figure 135. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 17:10 UTC.

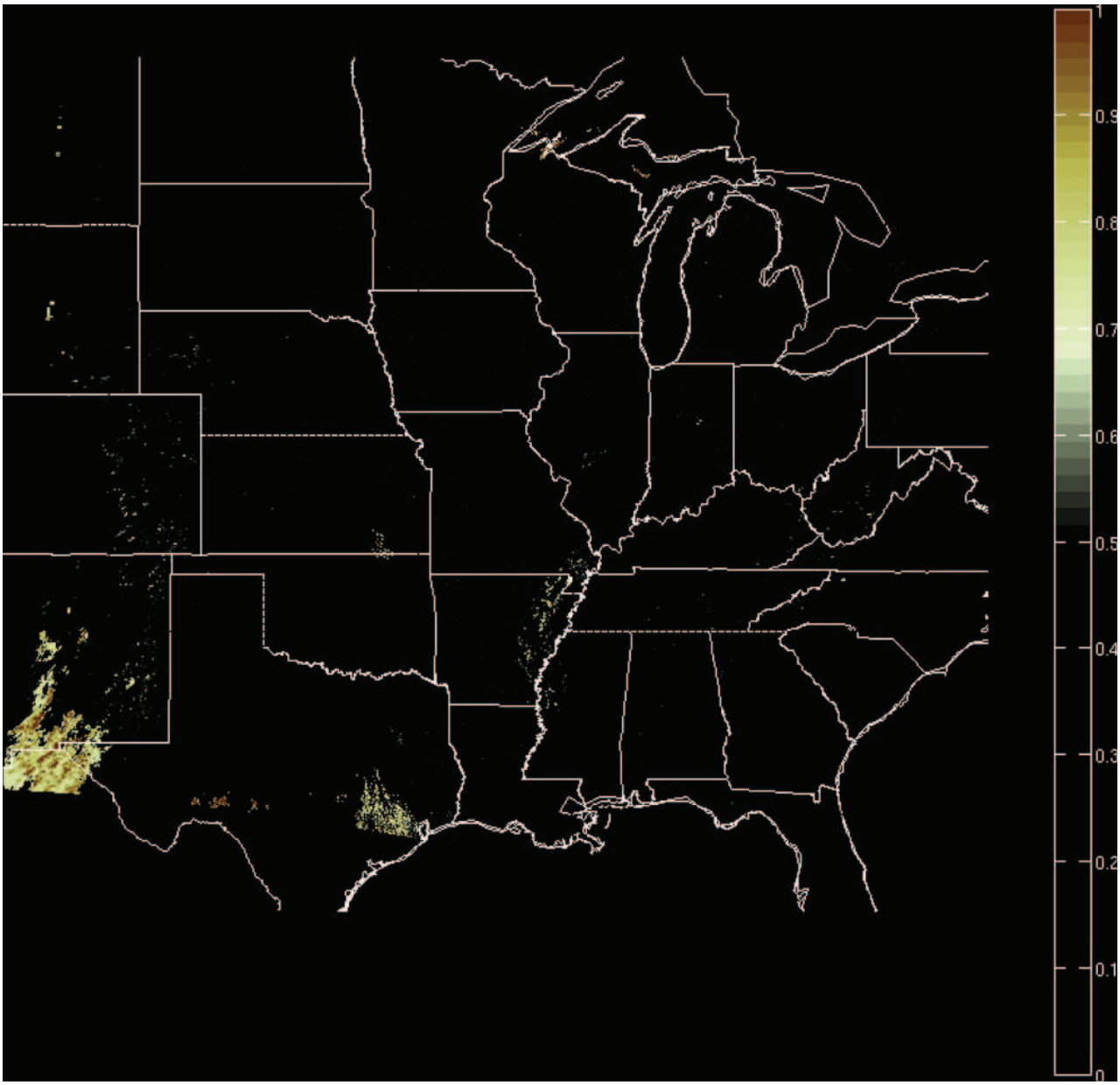

Figure 136. Result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

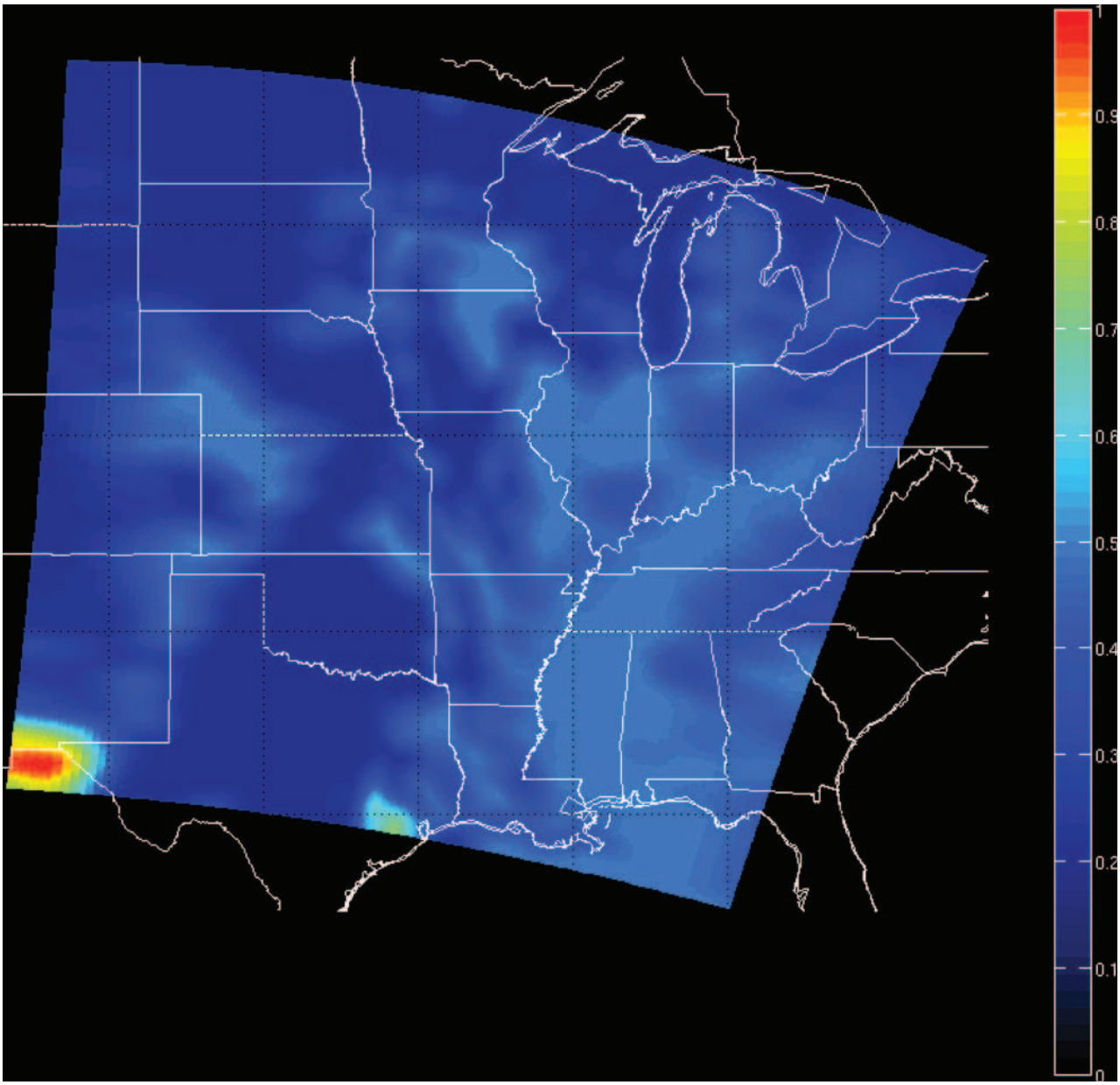

Figure 137. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 17:10 UTC. The colorbar shows the index of probability.

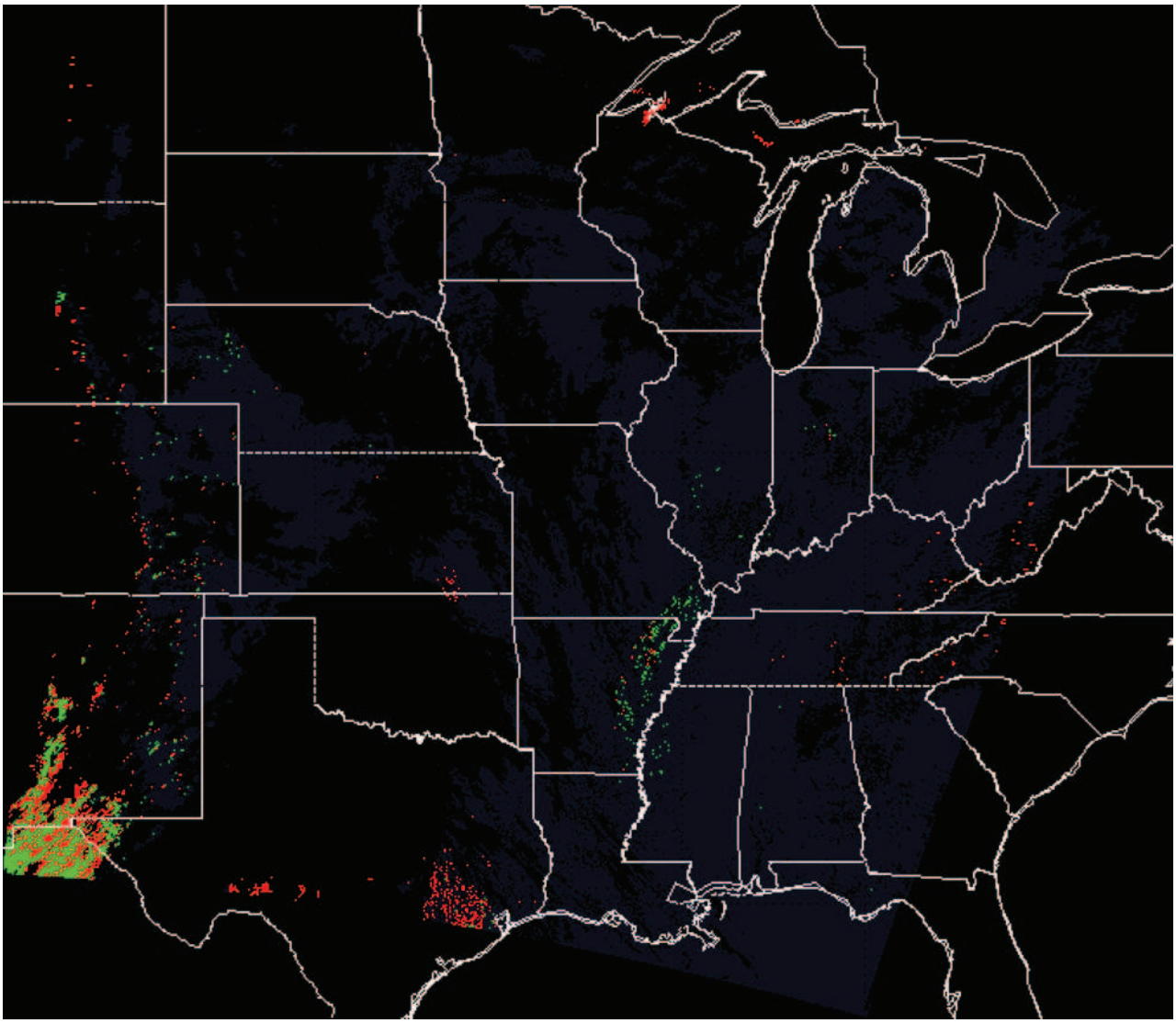

Figure 138. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 17:10 UTC.

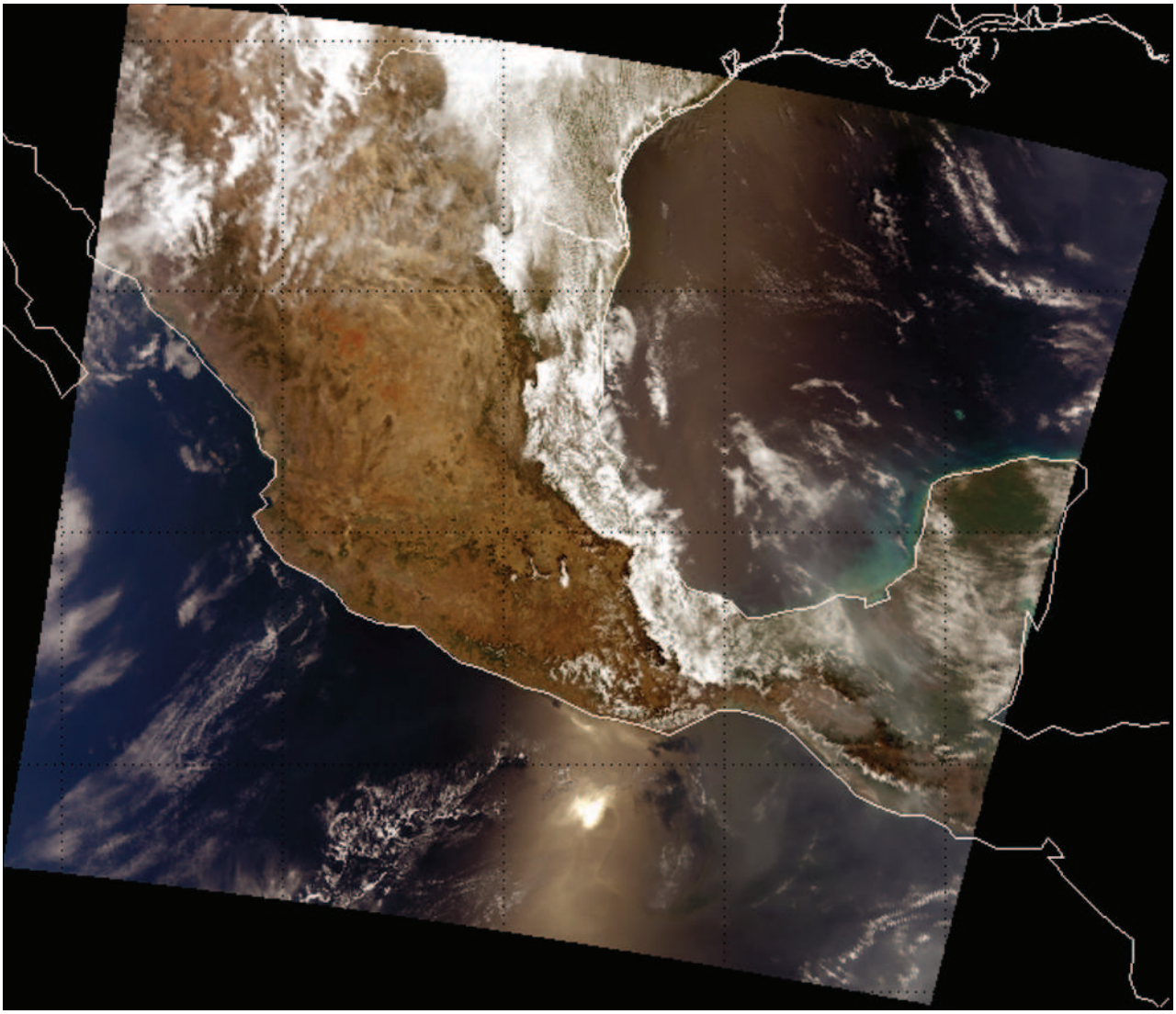

Figure 139. True color image for the event of 2003-105 at 17:15 UTC.

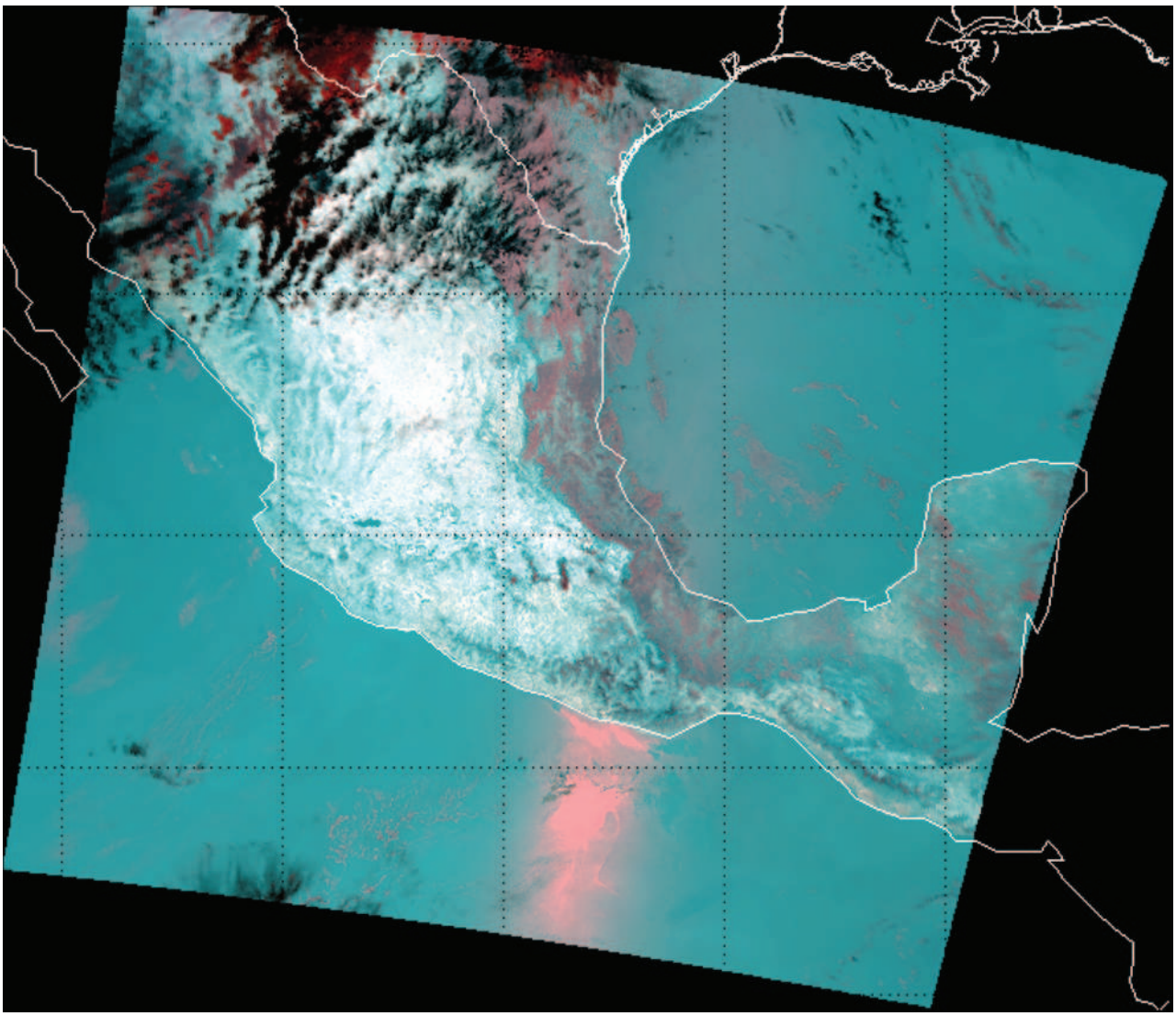

Figure 140. False color image for the event of 2003-105 at 17:15 UTC.

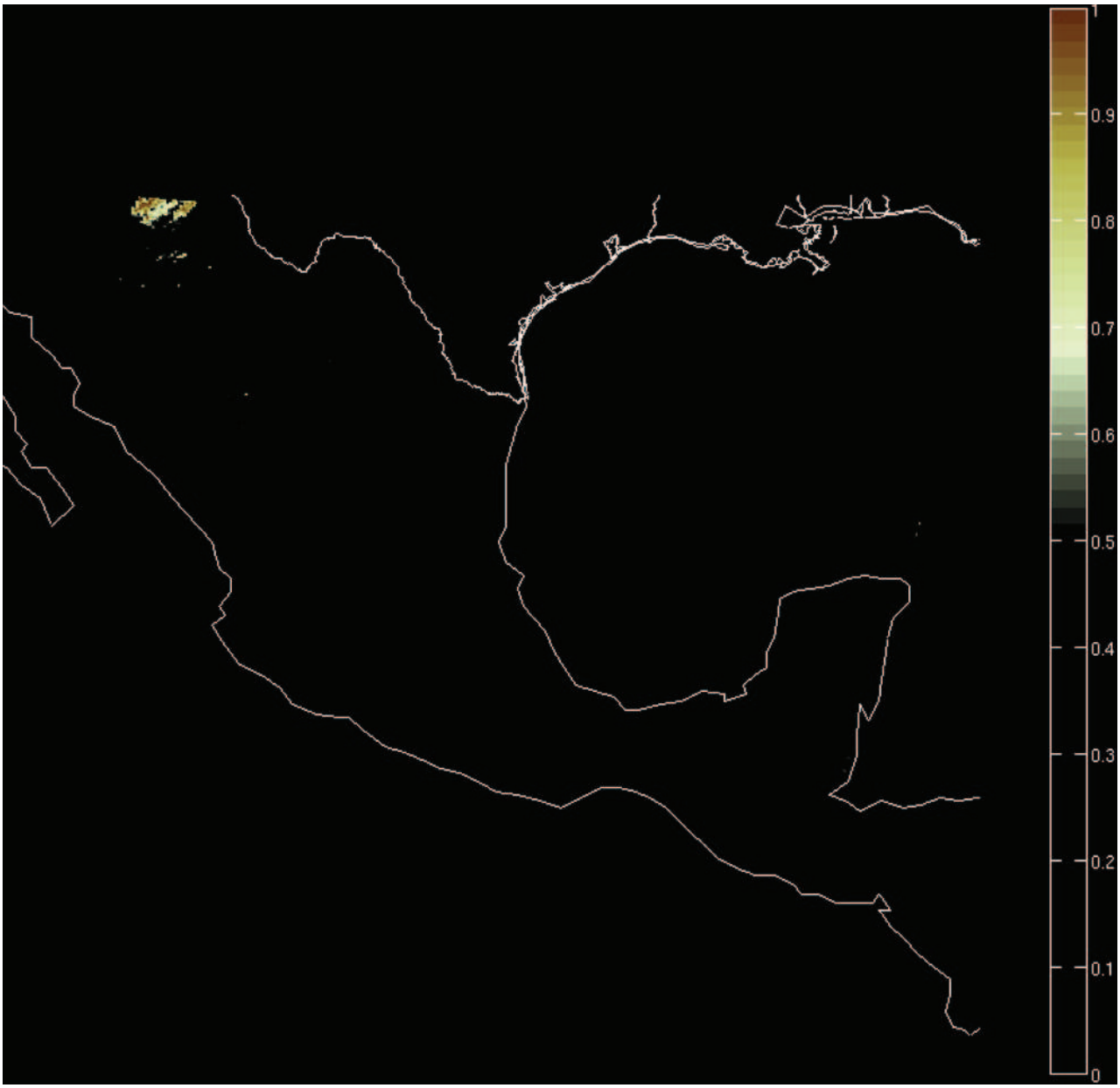

Figure 141. Result of the probabilistic display model (PD) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

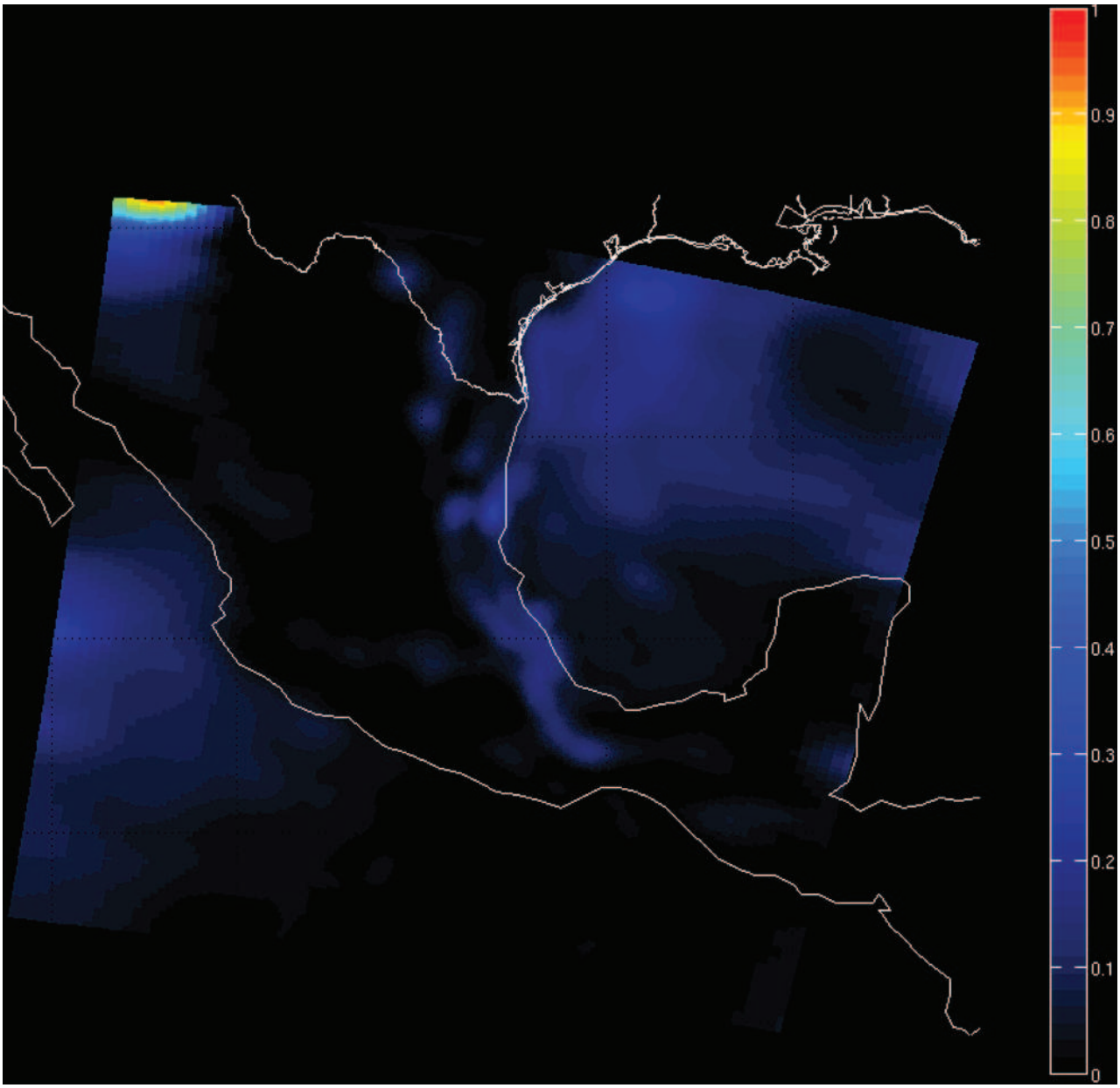

Figure 142. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

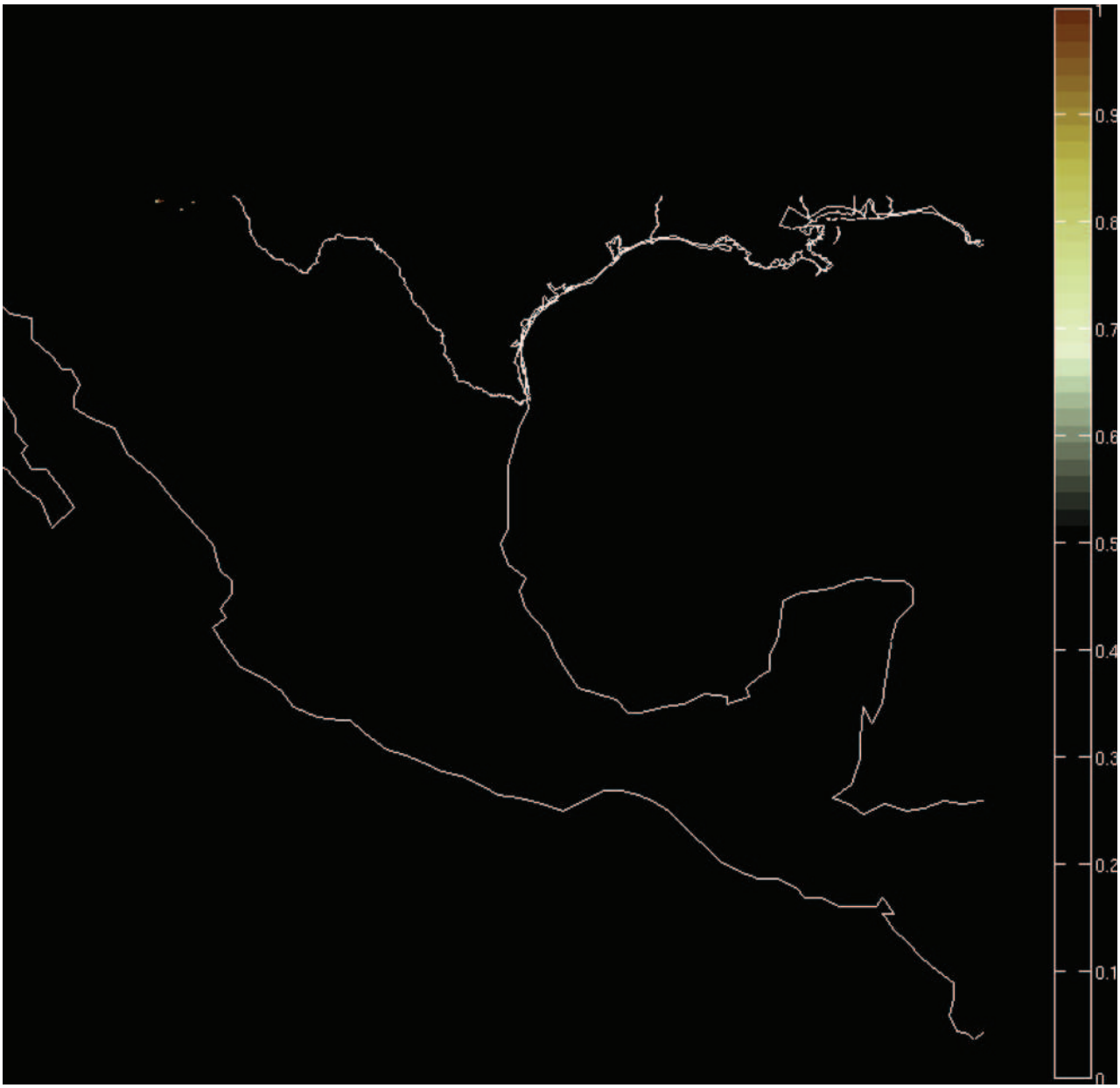

Figure 143. Result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

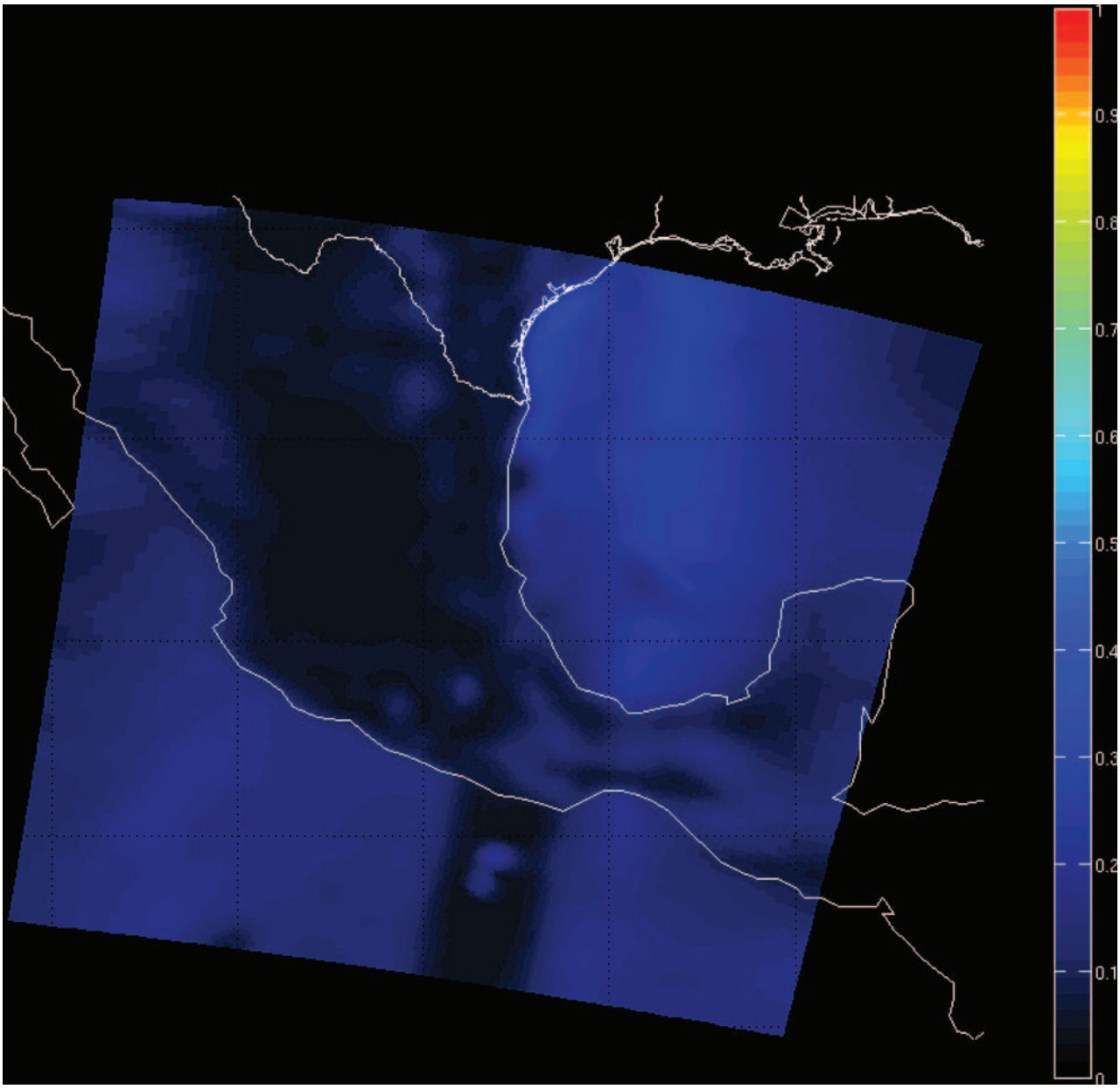

Figure 144. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

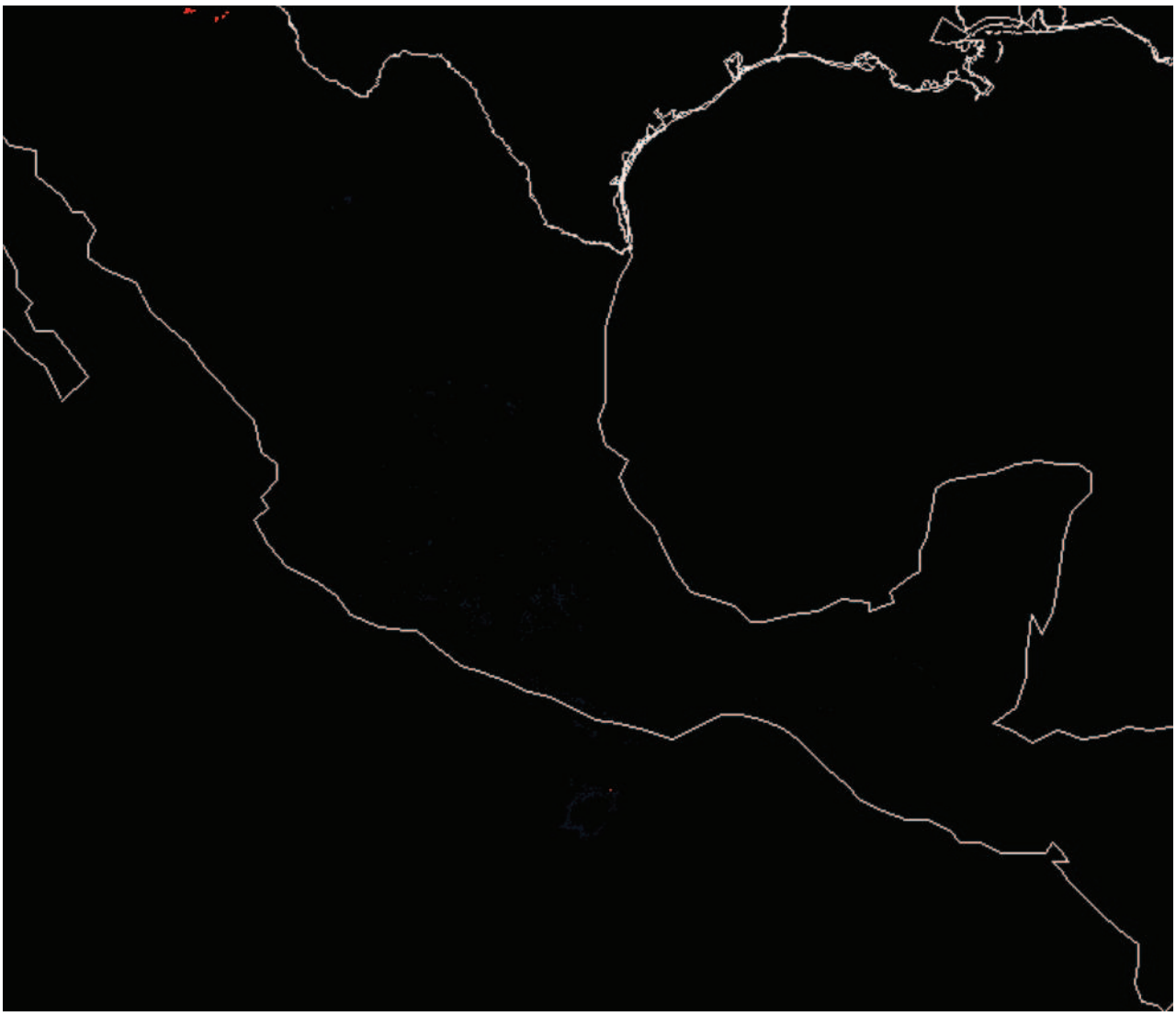

Figure 145. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 17:15 UTC.

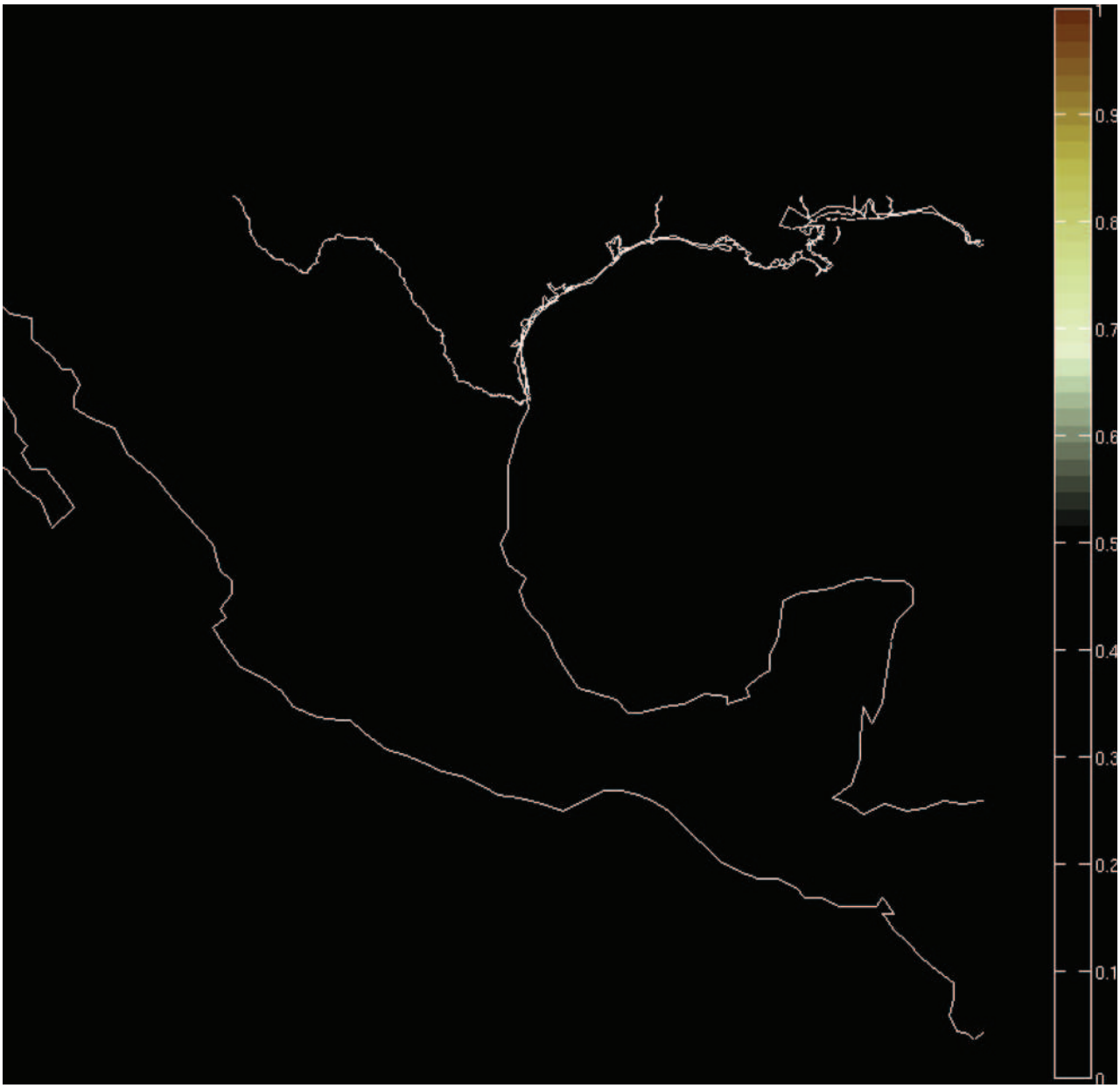

Figure 146. Result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

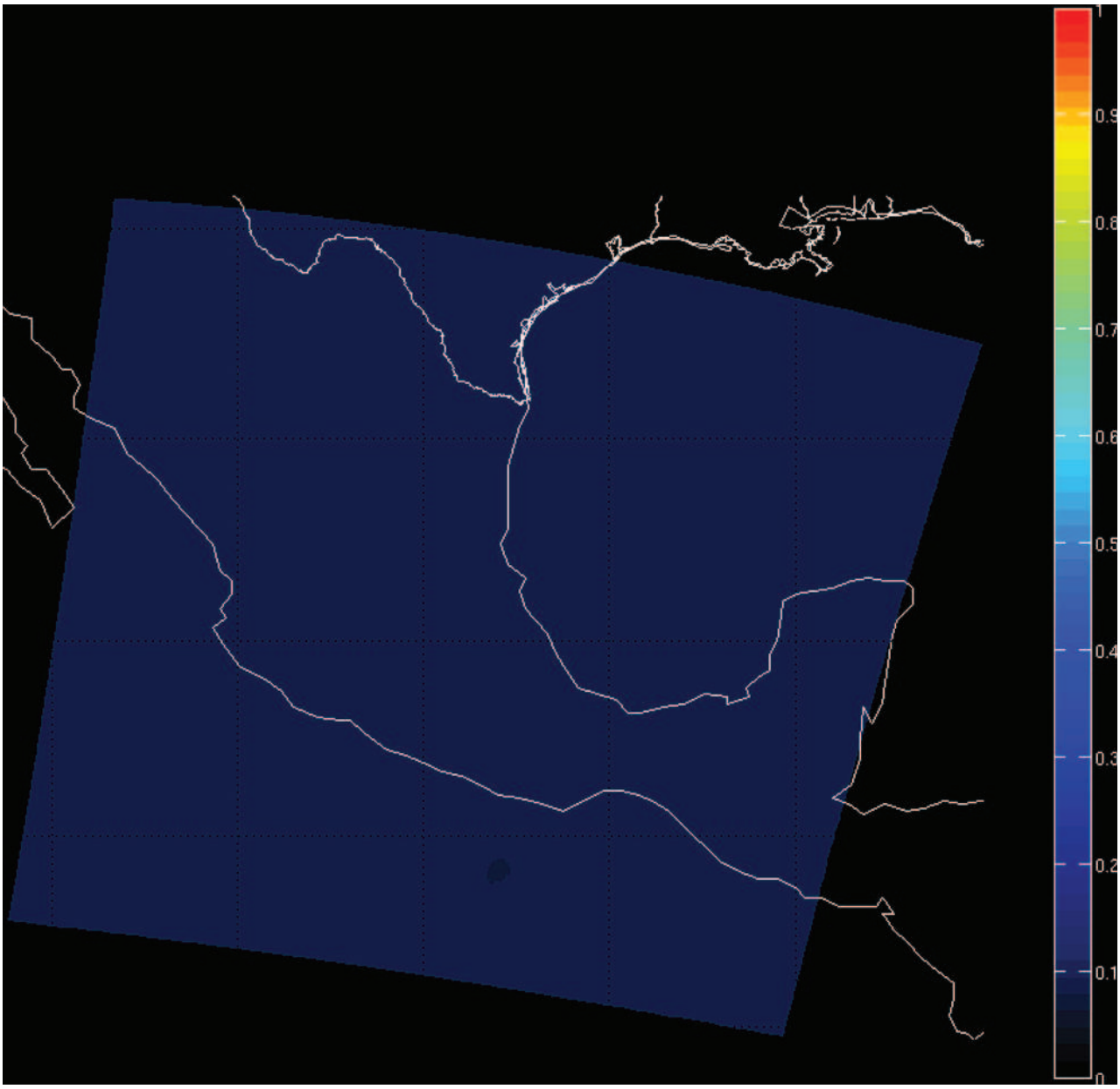

Figure 147. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

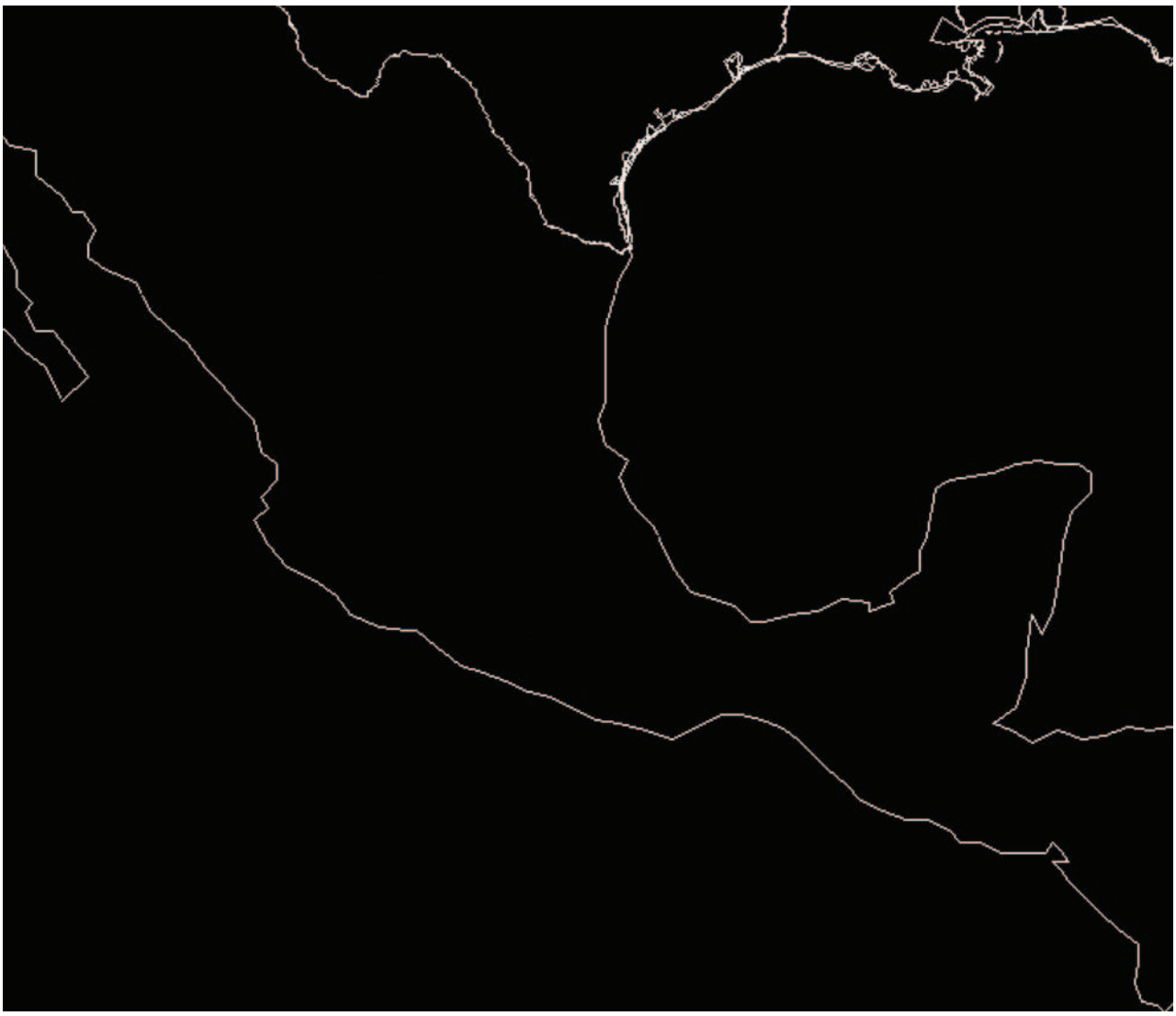

Figure 148. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 17:15 UTC.

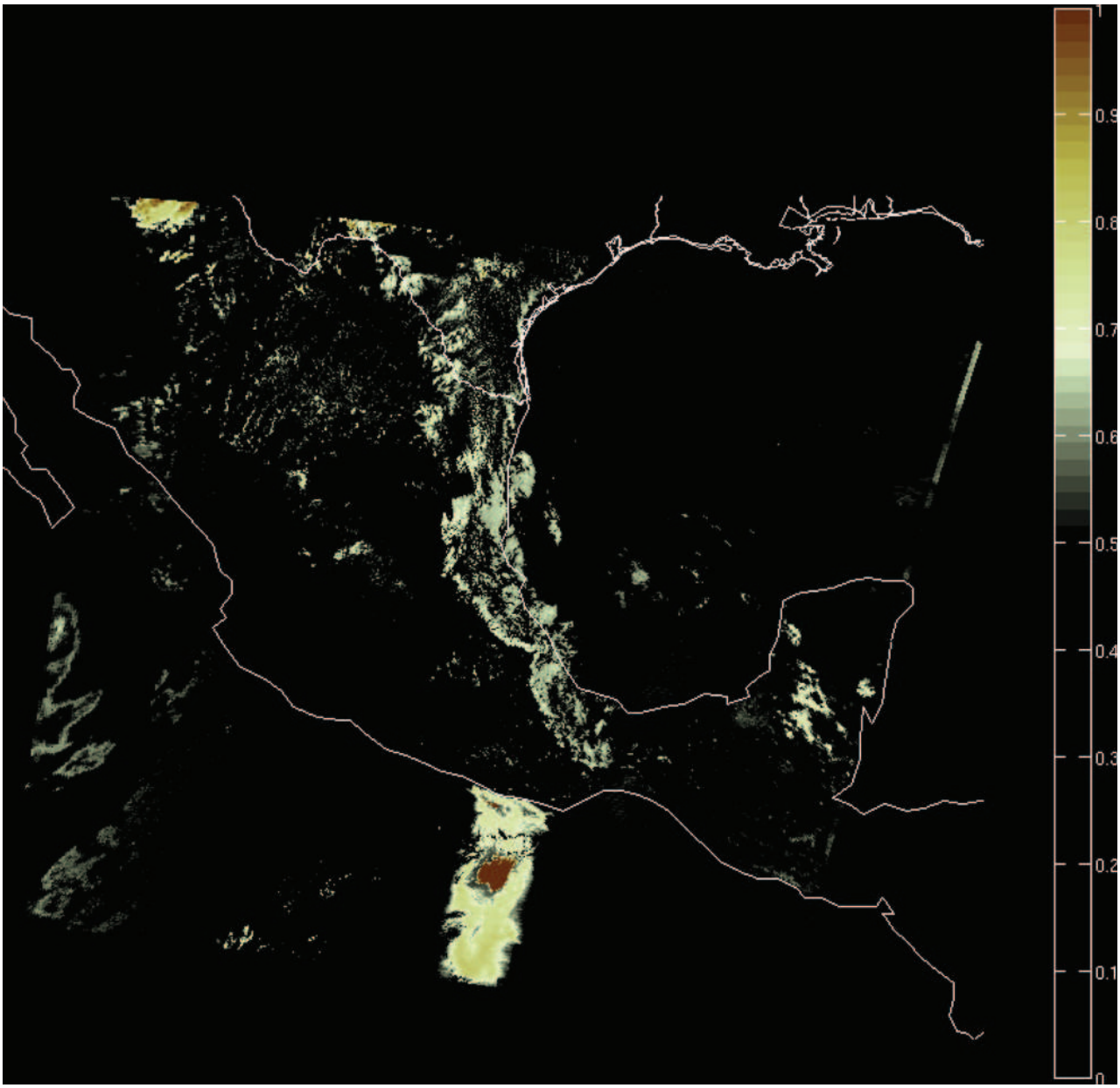

Figure 149. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

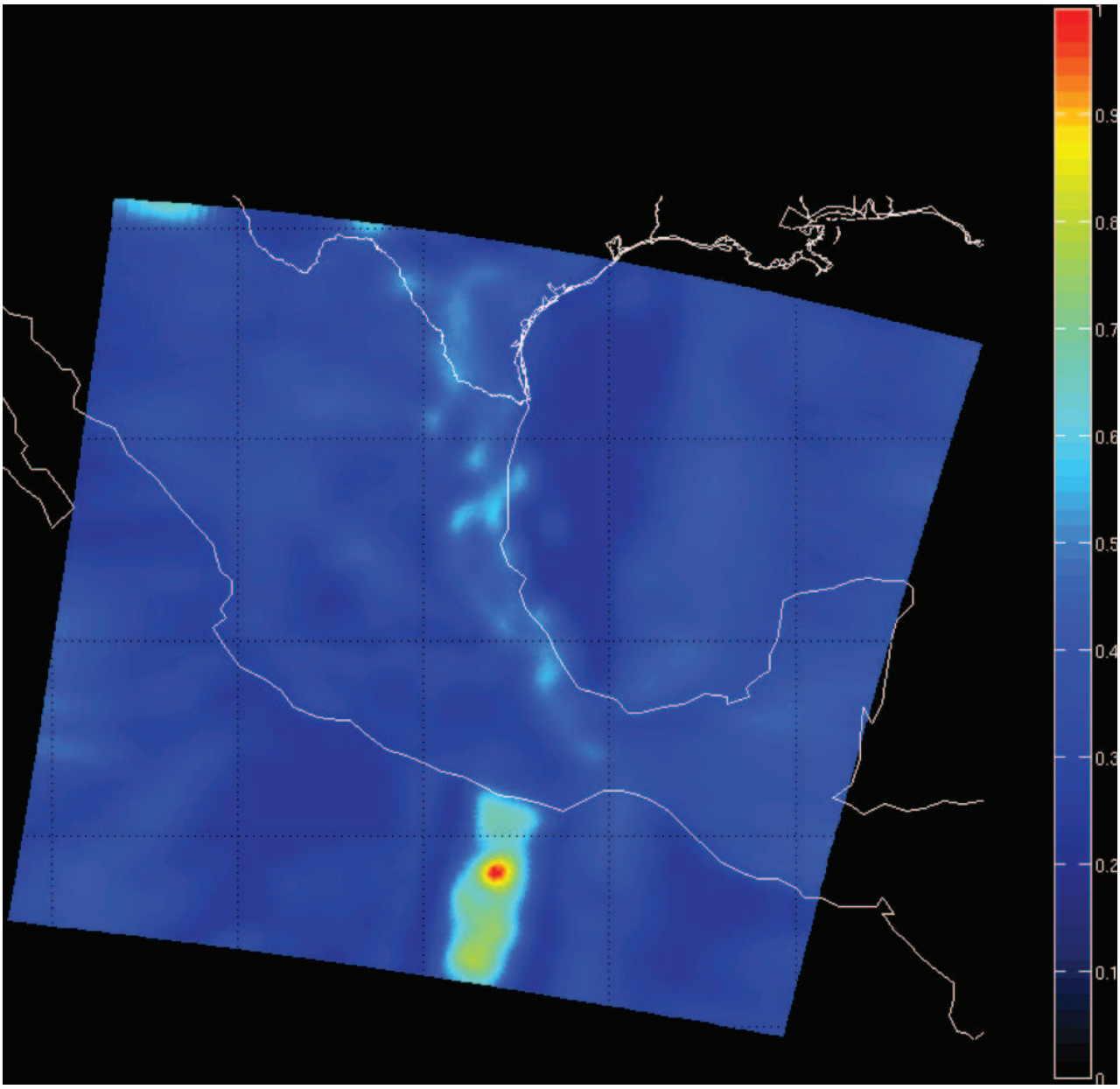

Figure 150. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

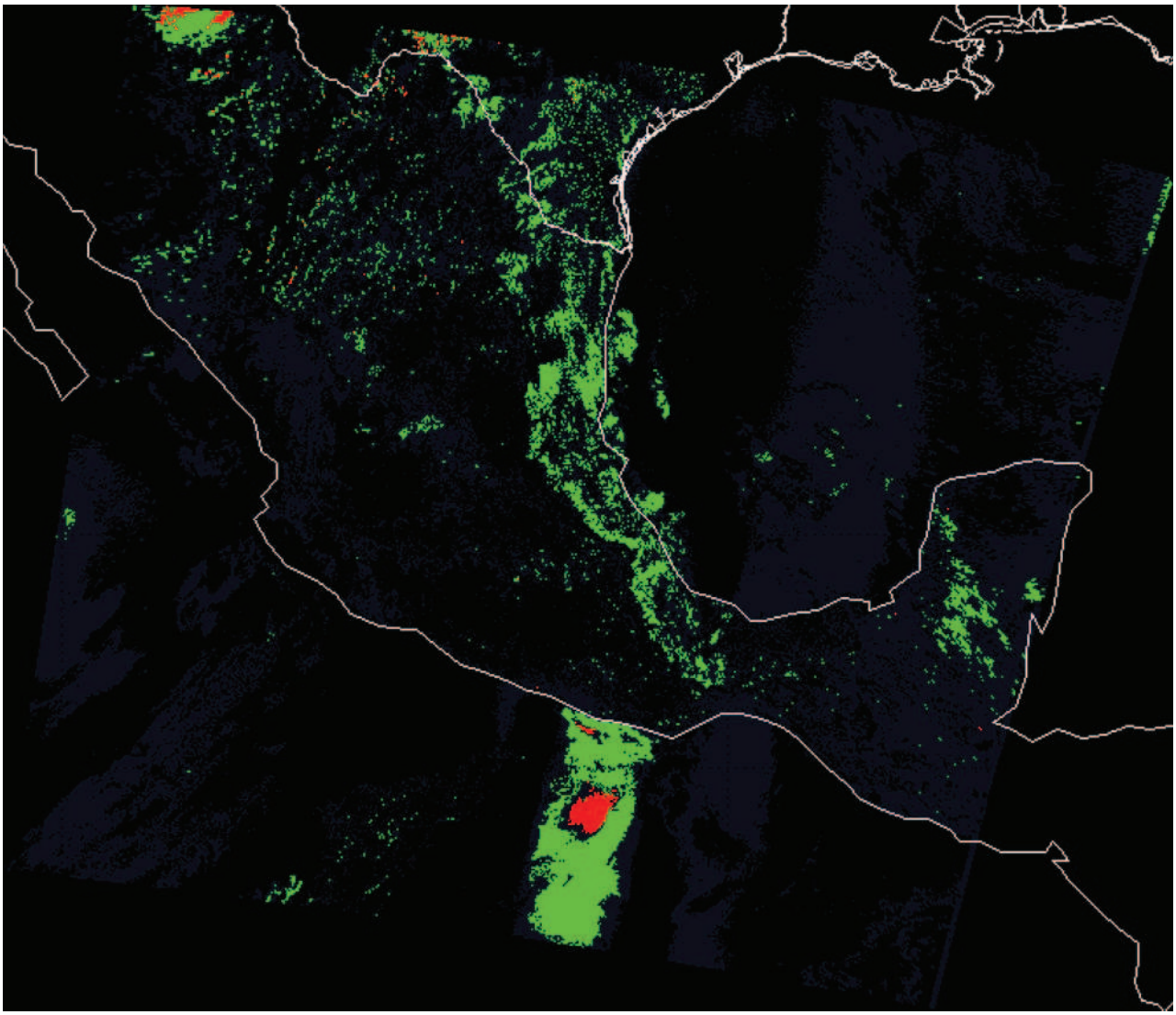

Figure 151. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 17:15 UTC.

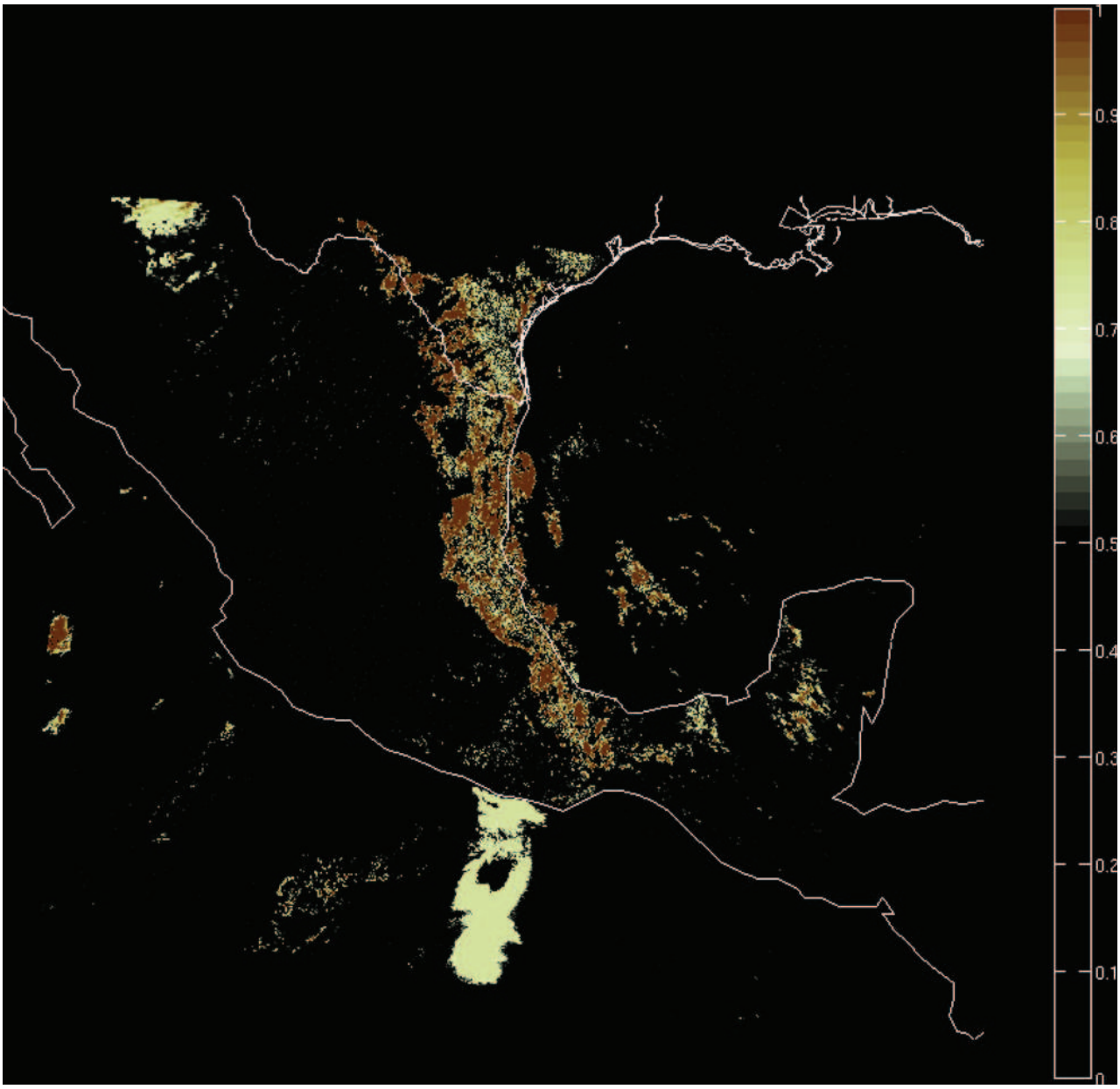

Figure 152. Result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

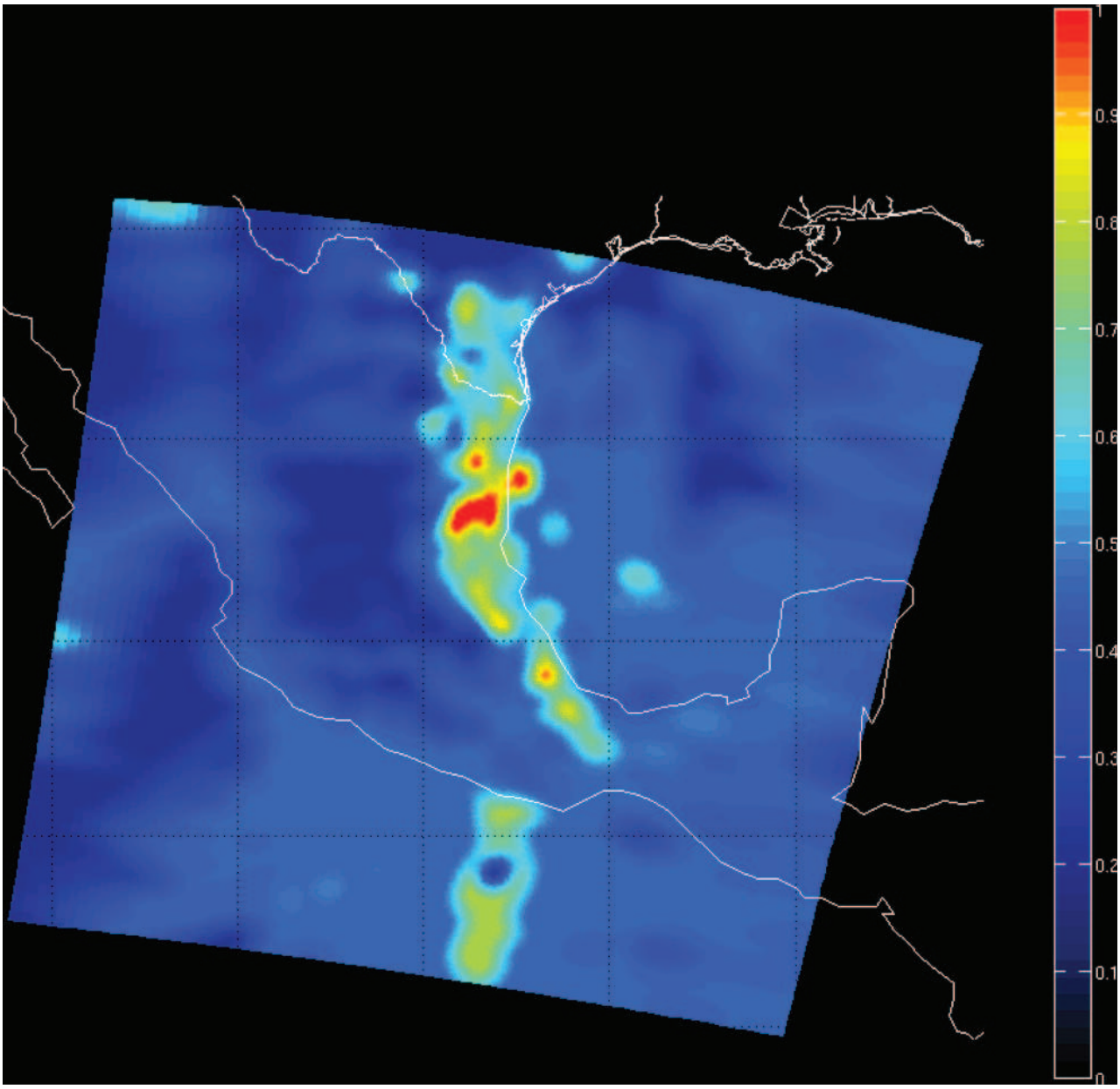

Figure 153. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 17:15 UTC. The colorbar shows the index of probability.

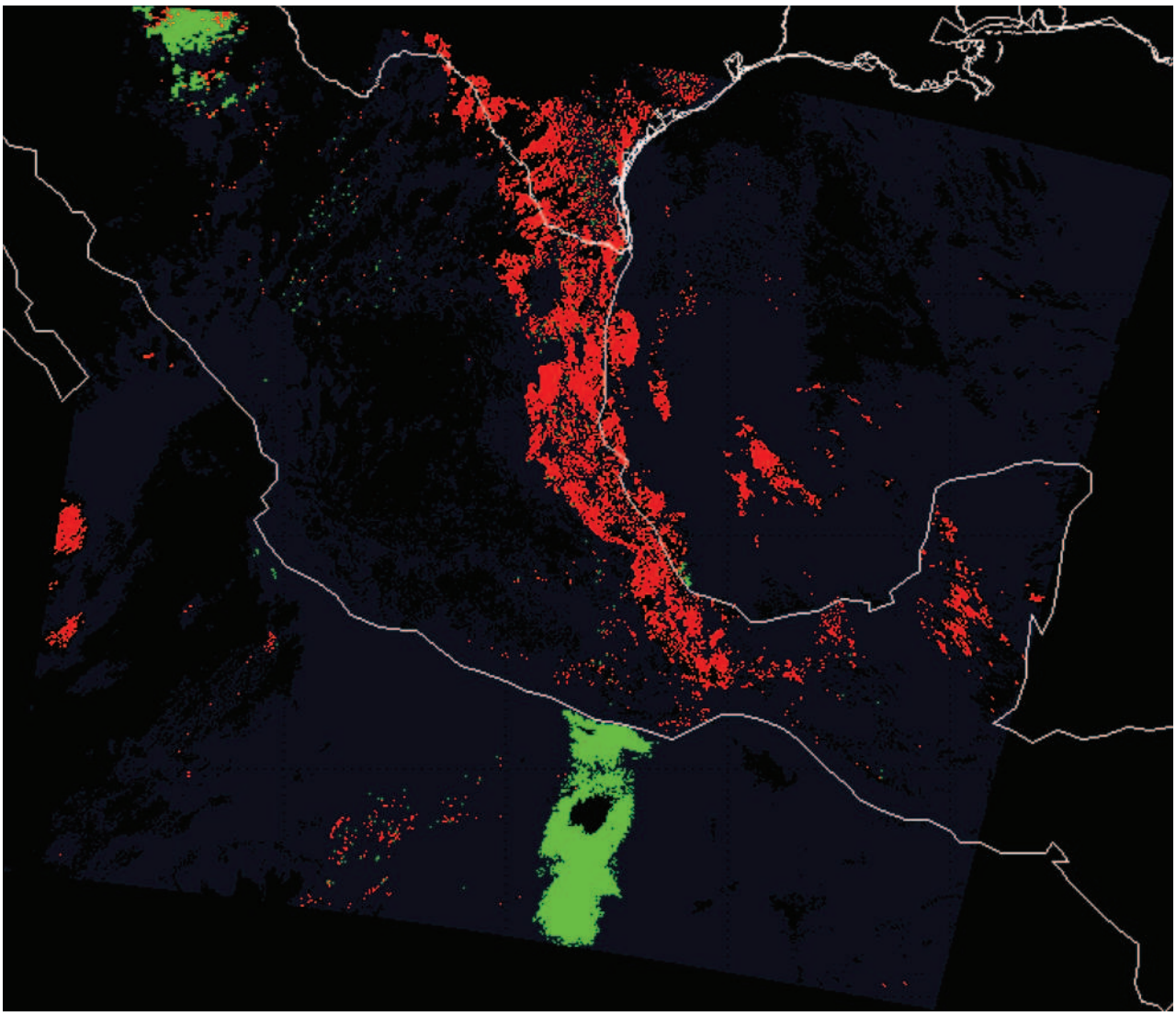

Figure 154. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 17:15 UTC.

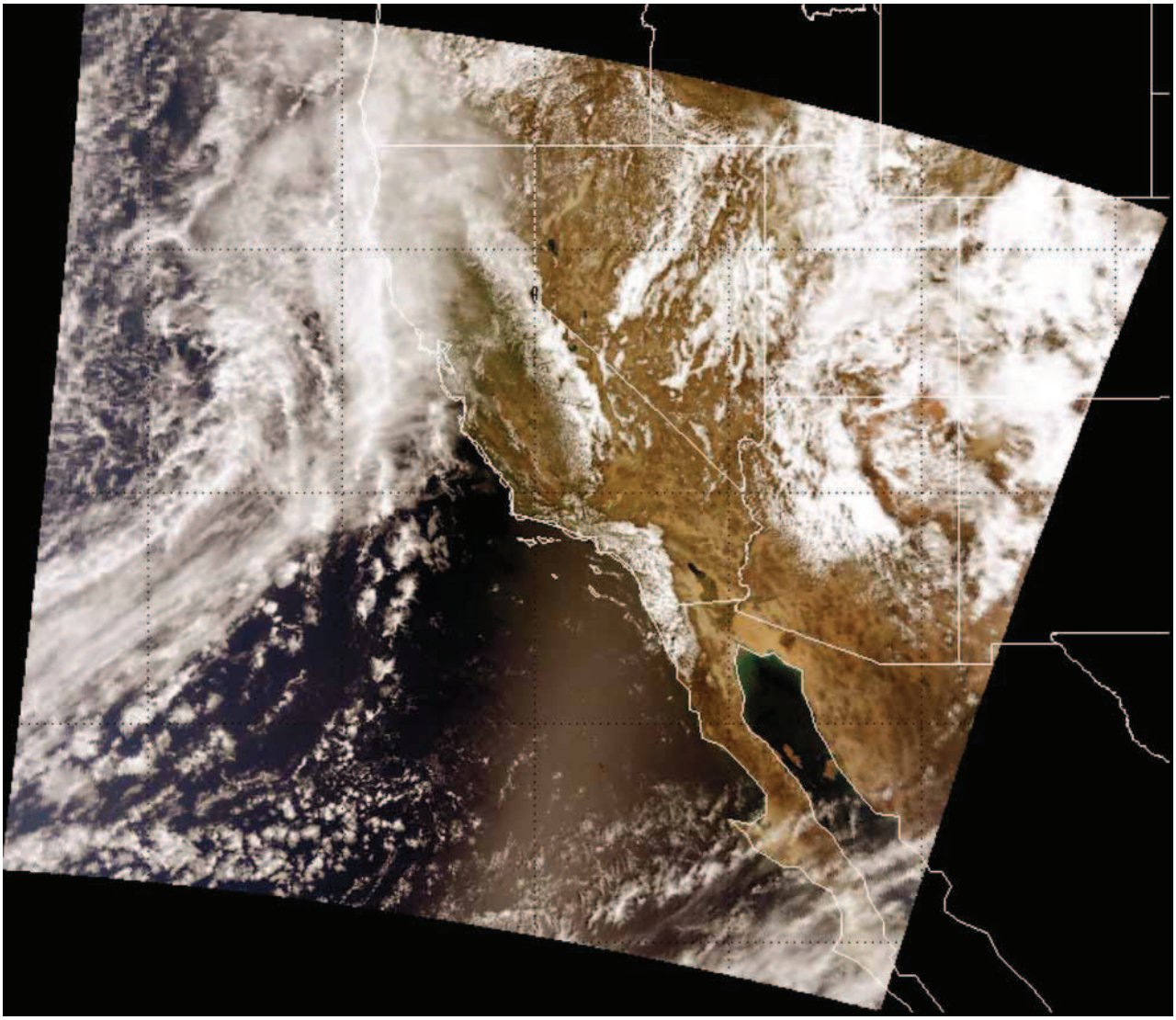

Figure 155. True color image for the event of 2003-105 at 18:50 UTC.

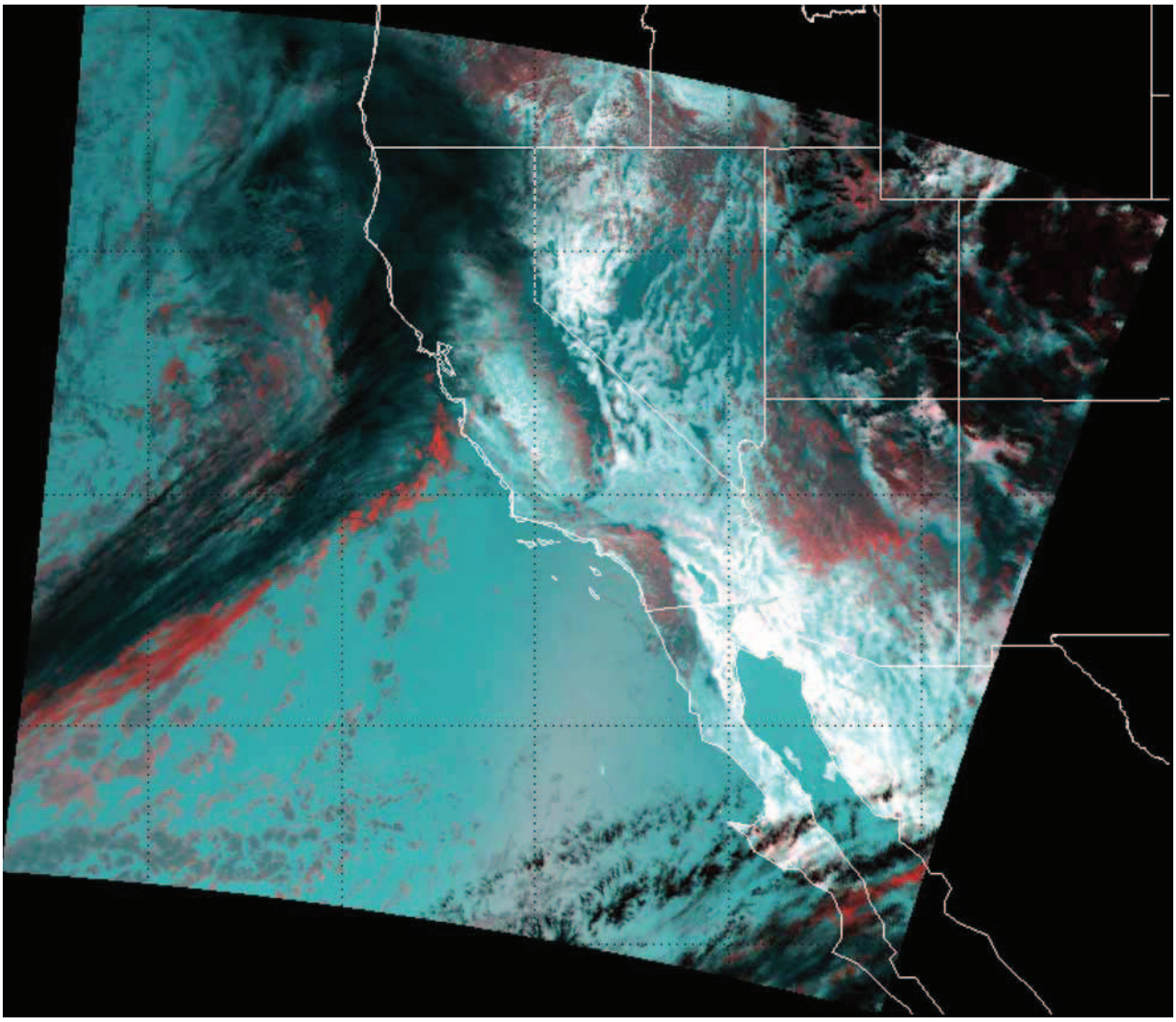

Figure 156. False color image for the event of 2003-105 at 18:50 UTC.

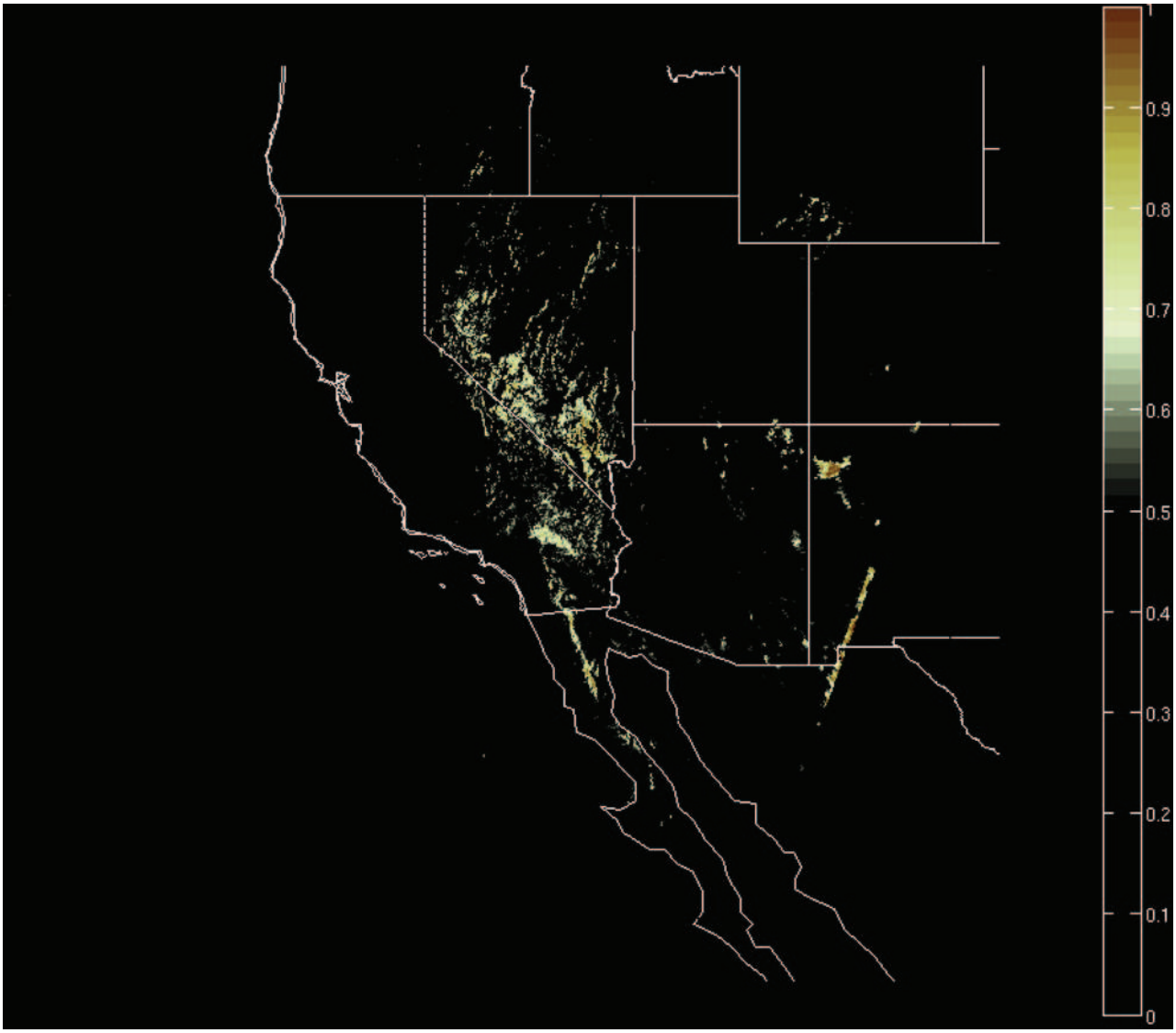

Figure 157. Result of the probabilistic display model (PD) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

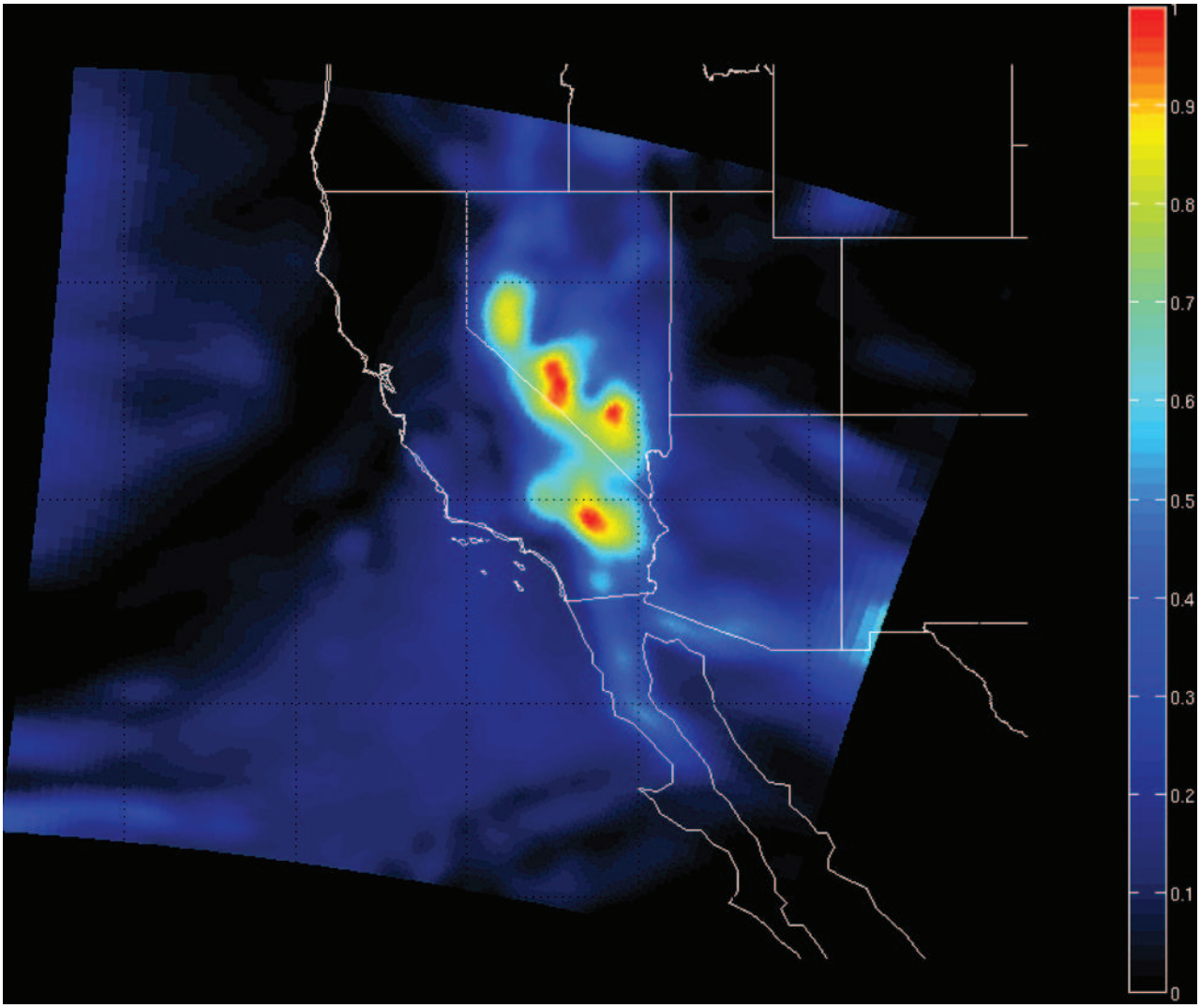

Figure 158. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

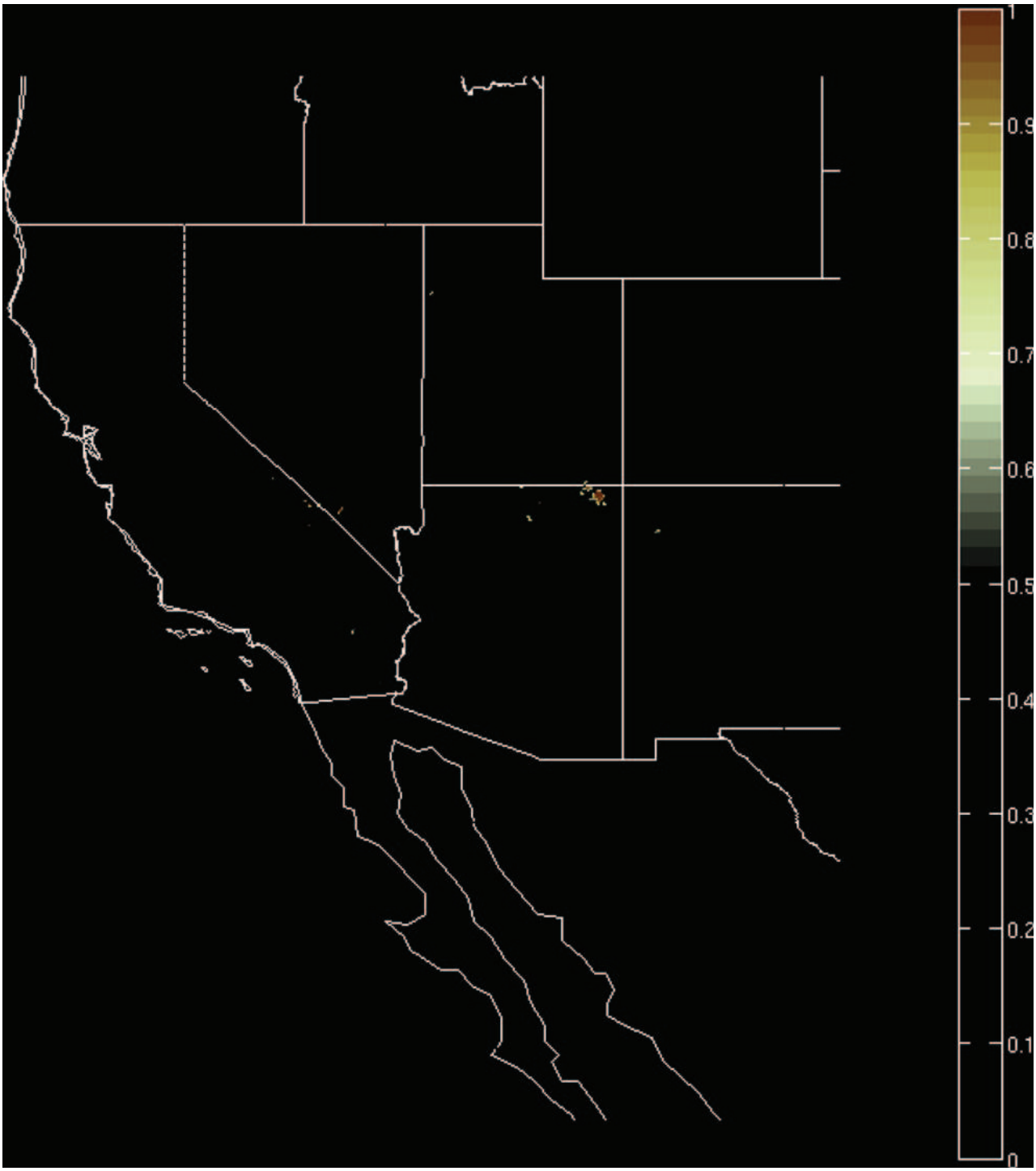

Figure 159. Result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

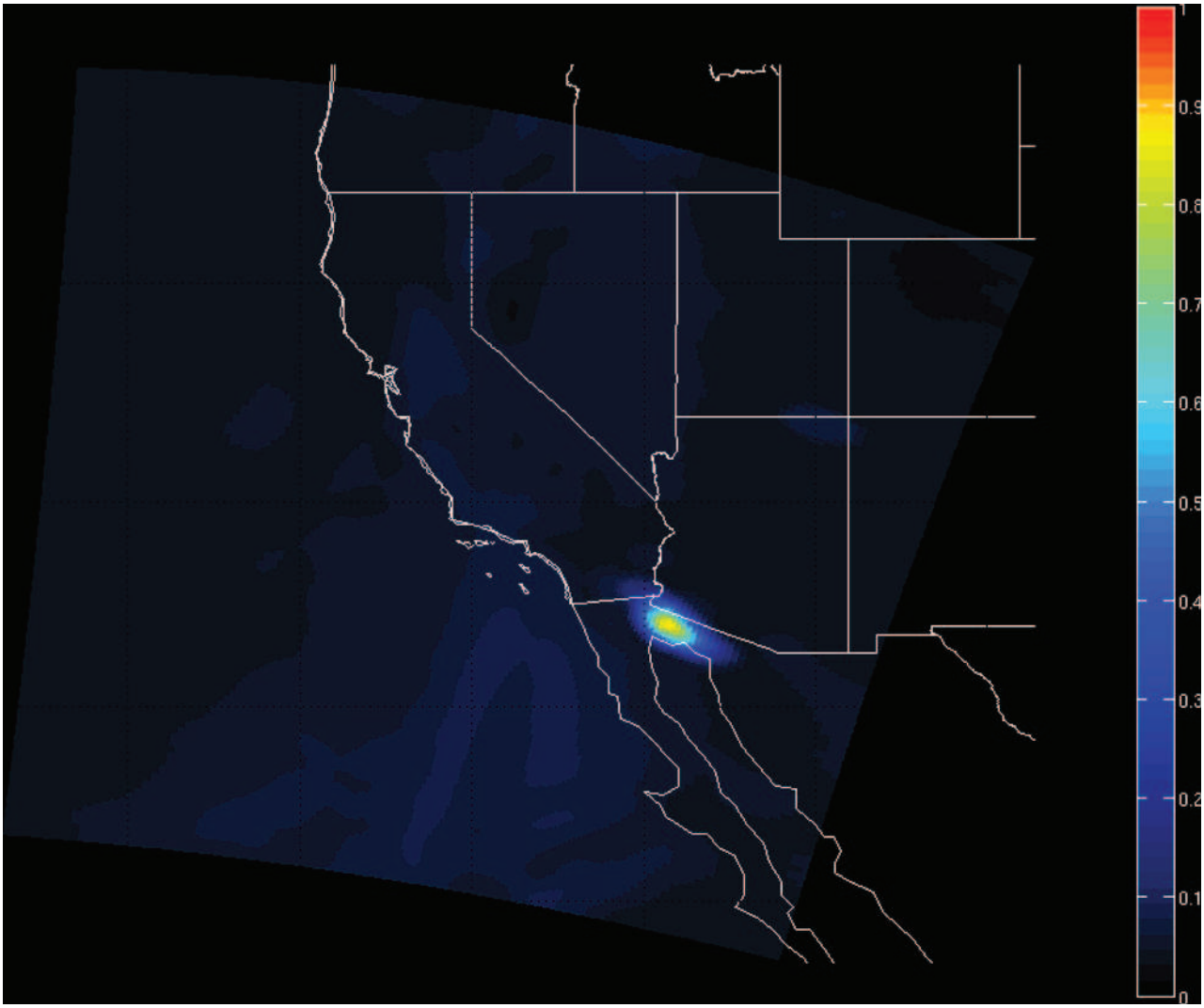

Figure 160. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

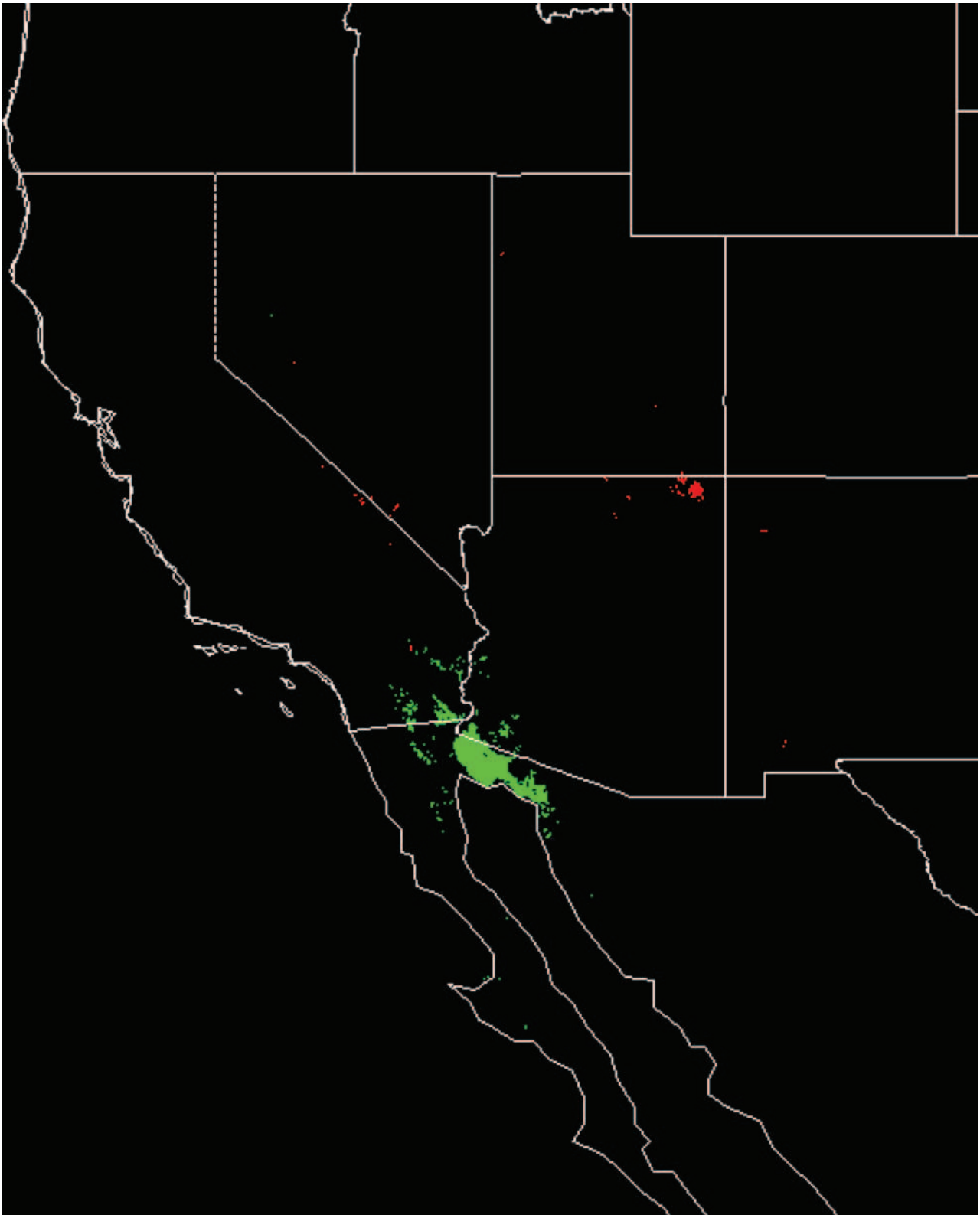

Figure 161. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-105 at 18:50 UTC.

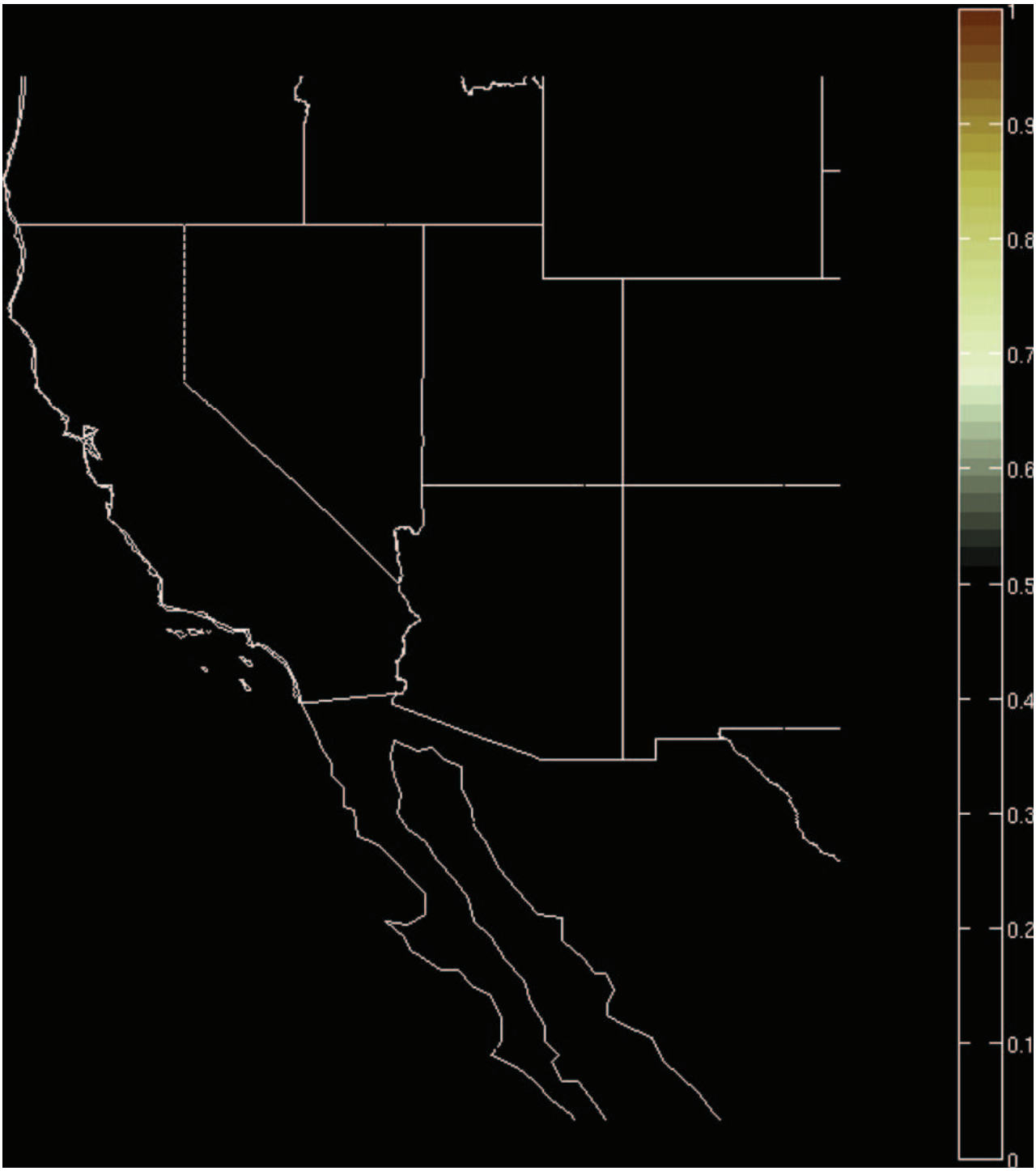

Figure 162. Result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

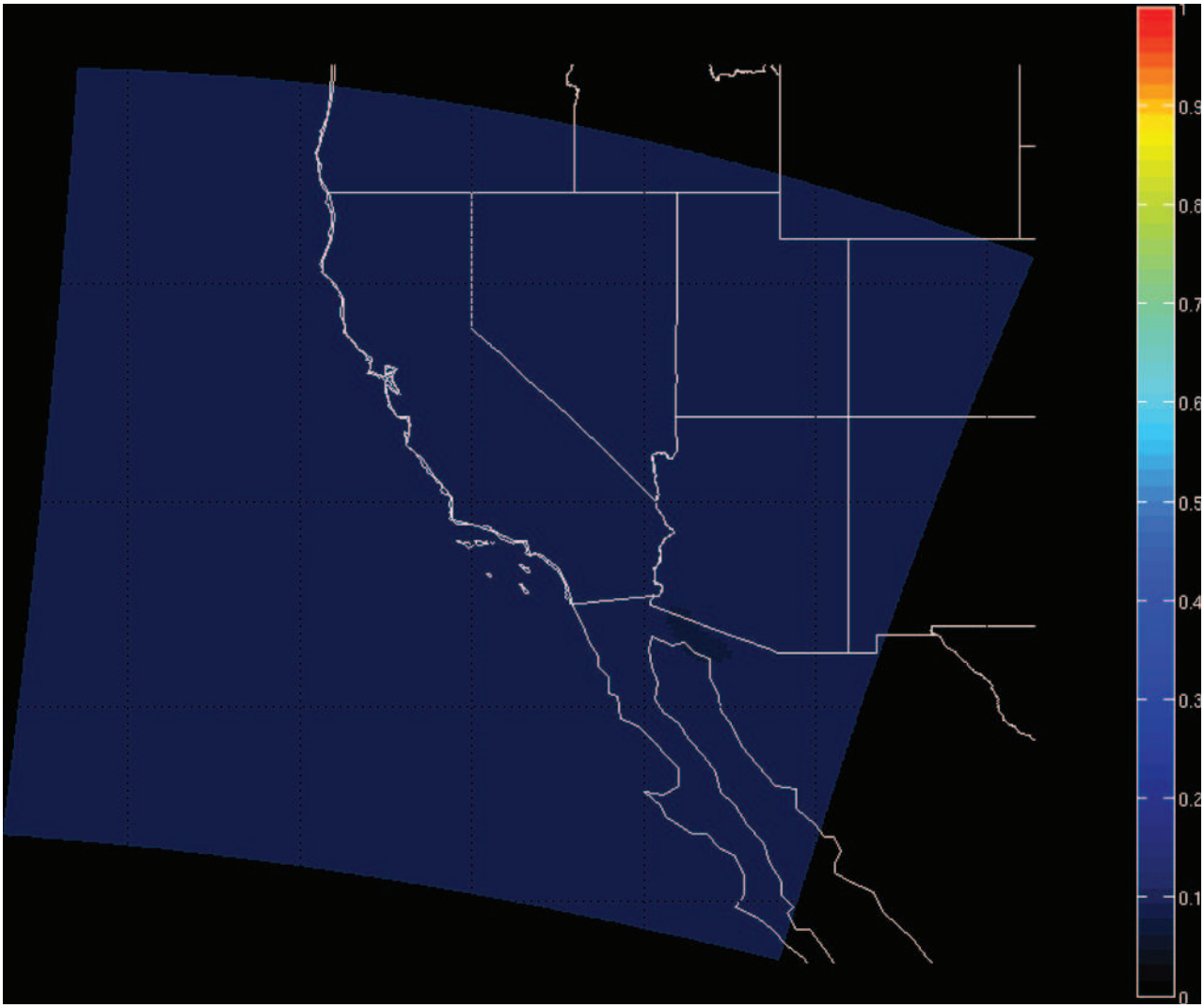

Figure 163. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

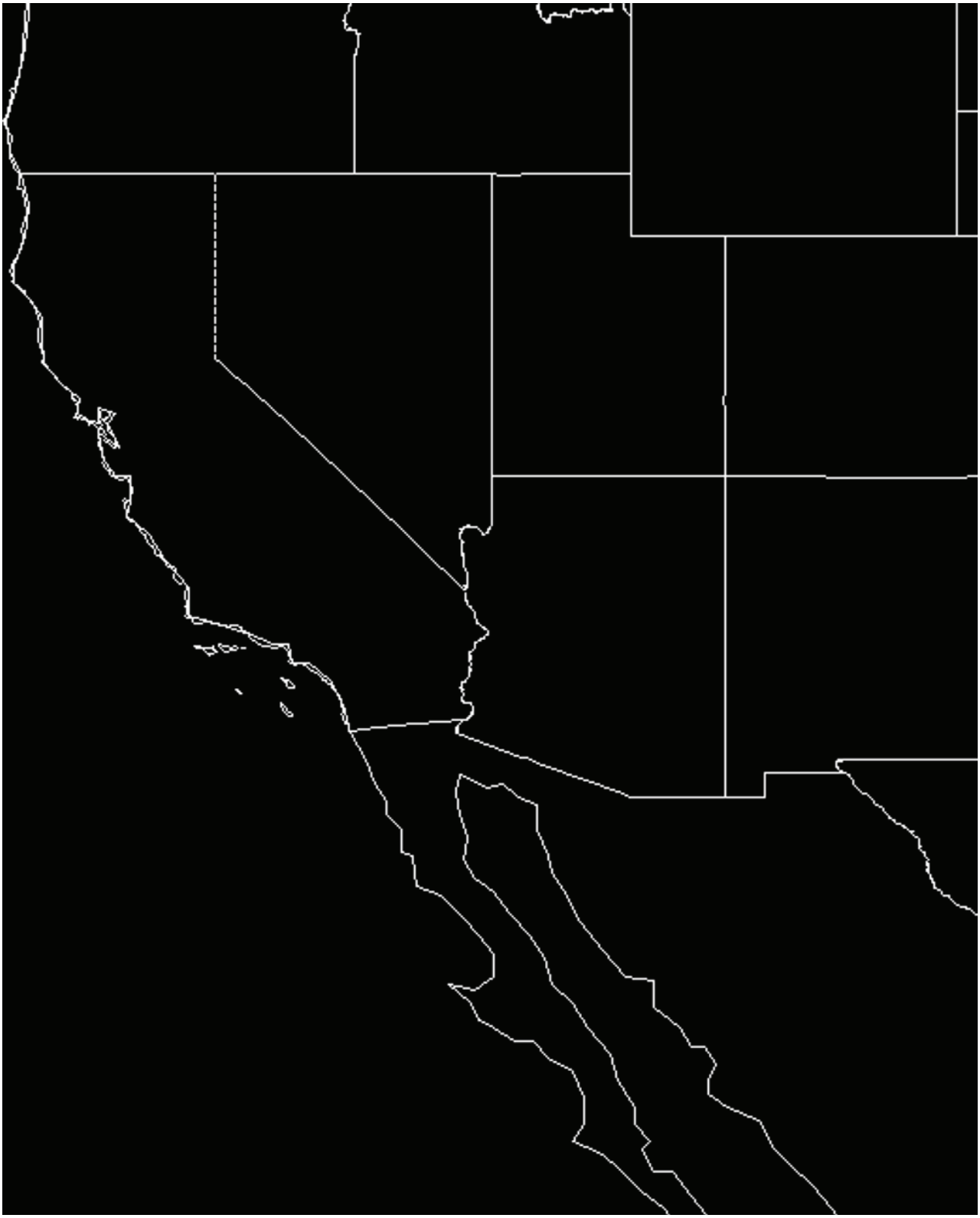

Figure 164. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-105 at 18:50 UTC.

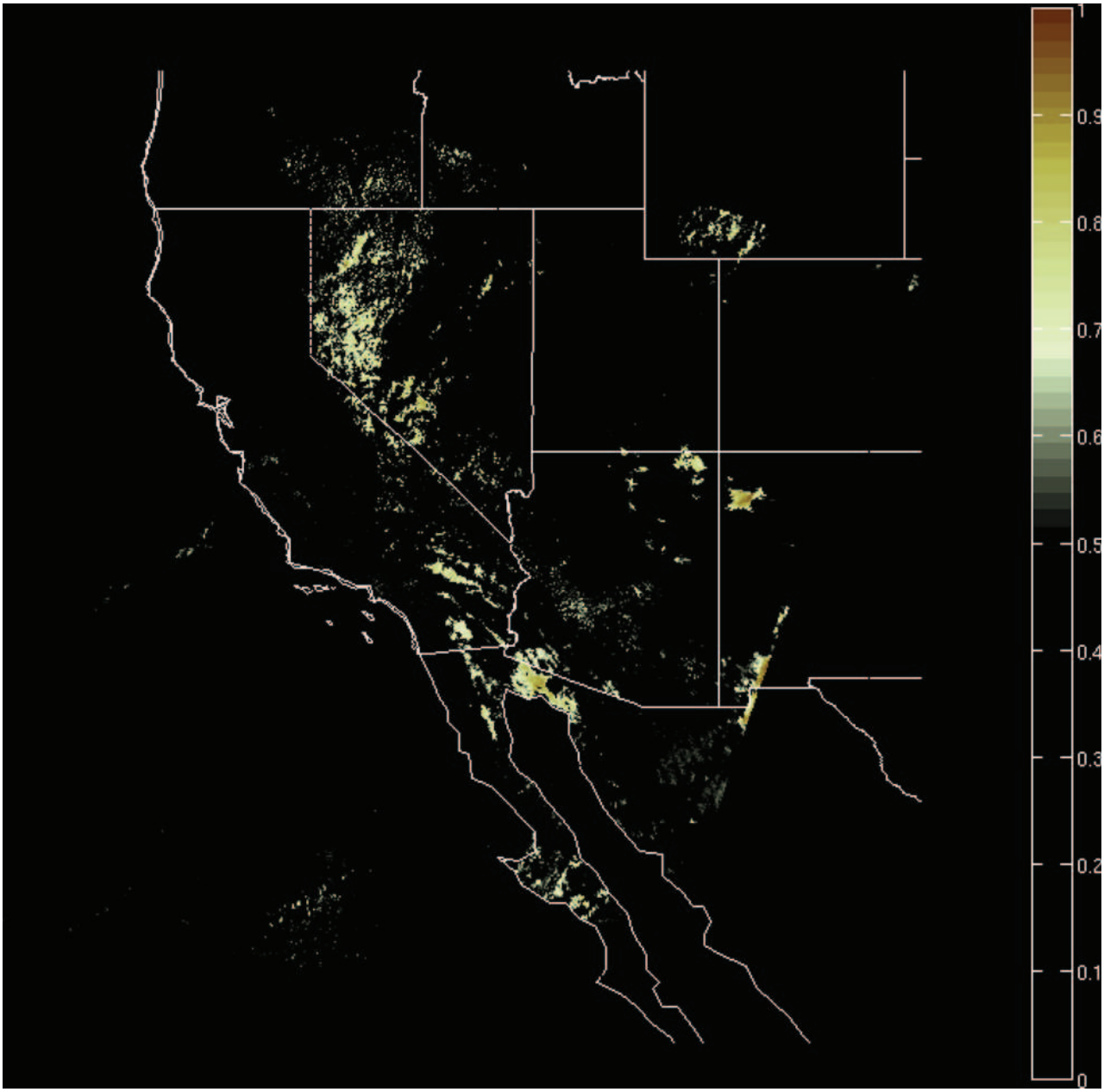

Figure 165. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

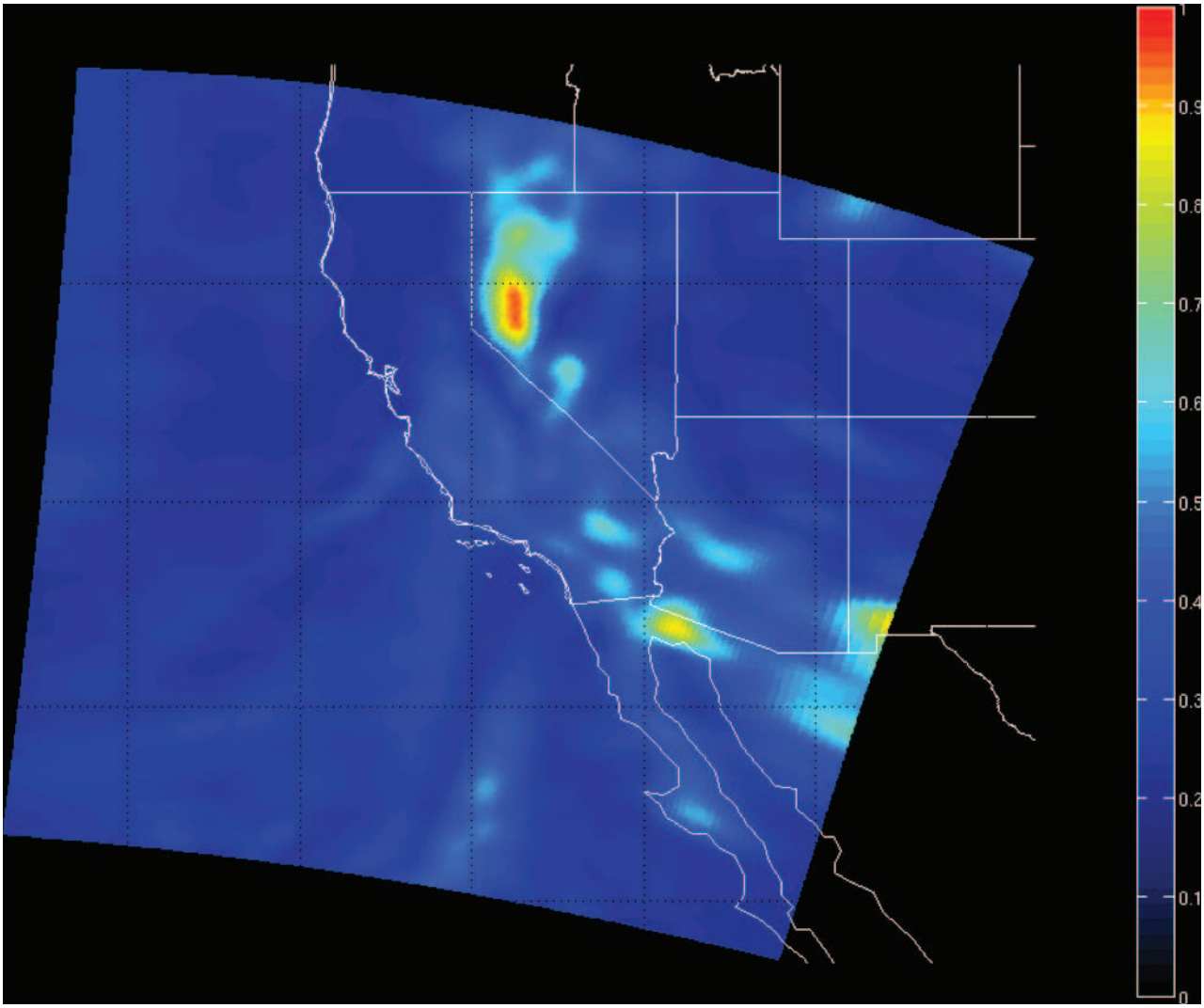

Figure 166. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

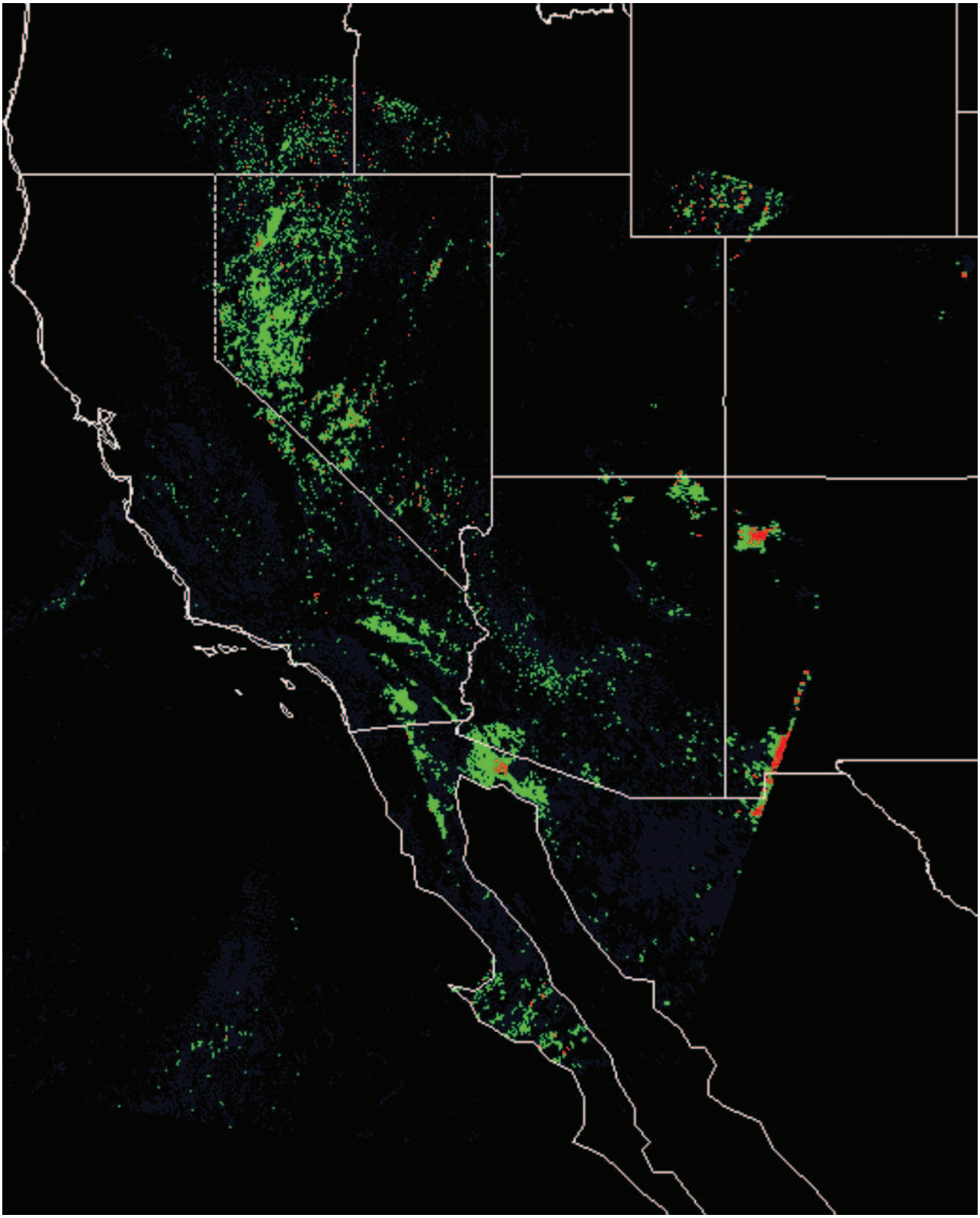

Figure 167. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-105 at 18:50 UTC.

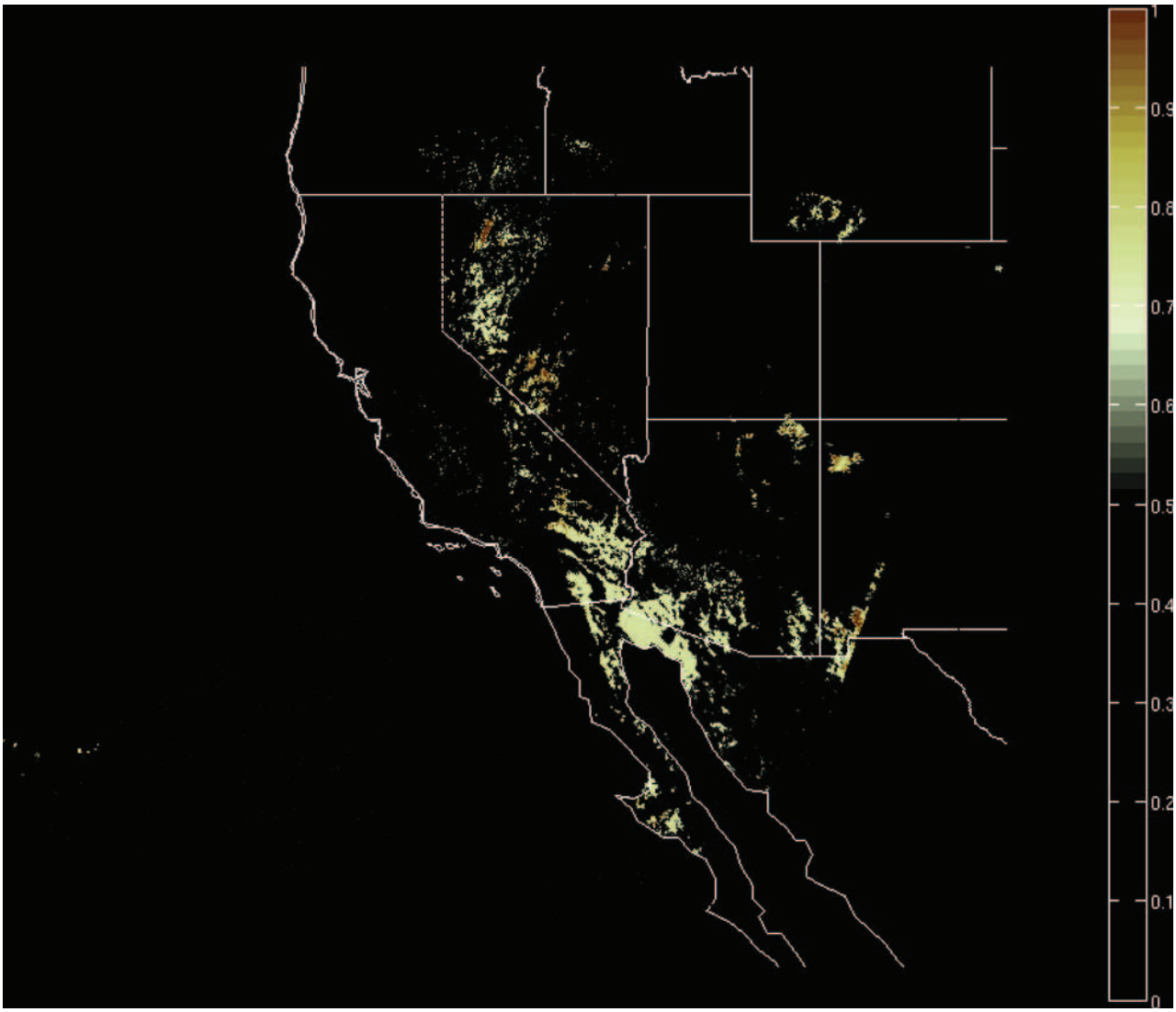

Figure 168. Result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

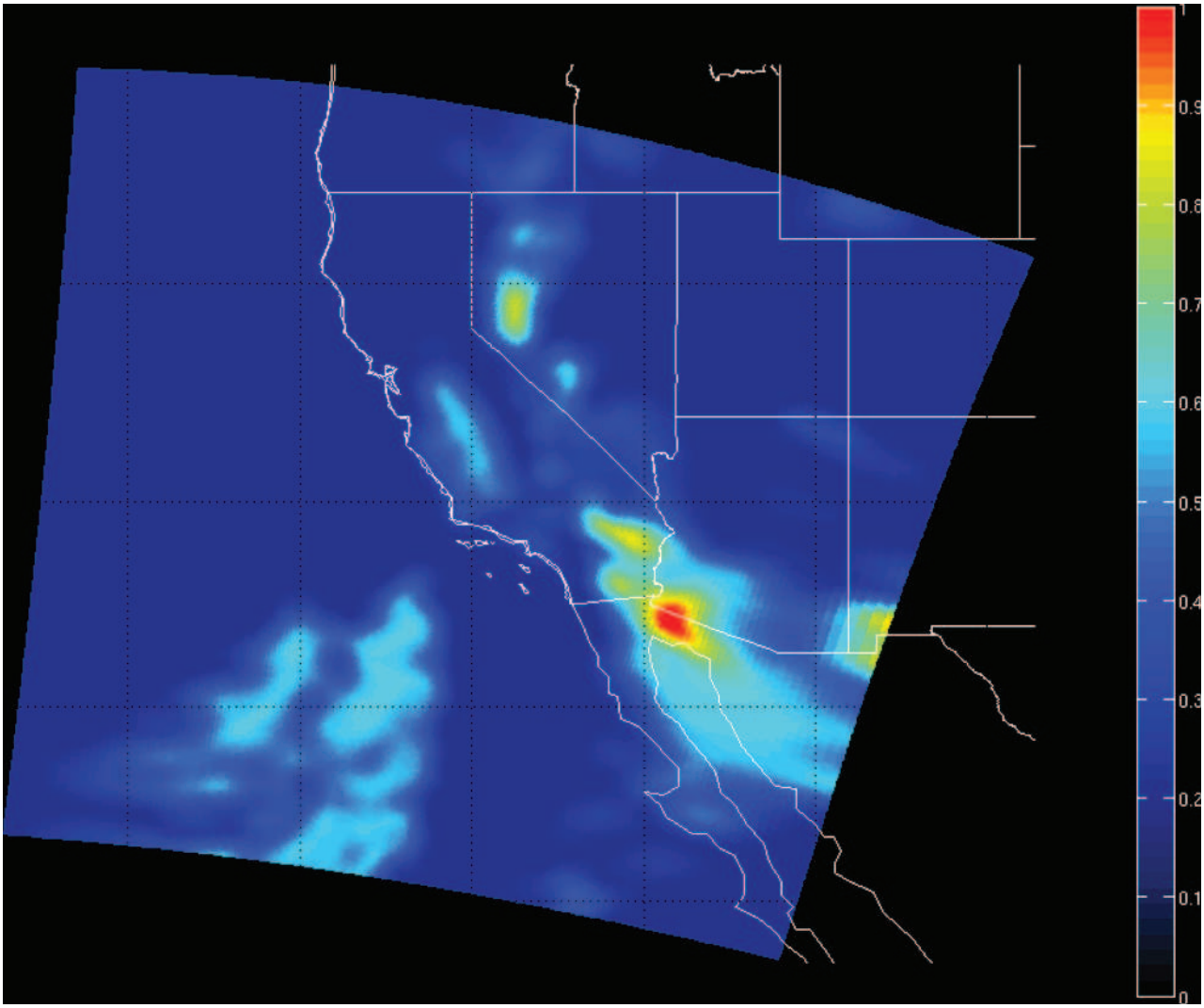

Figure 169. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 18:50 UTC. The colorbar shows the index of probability.

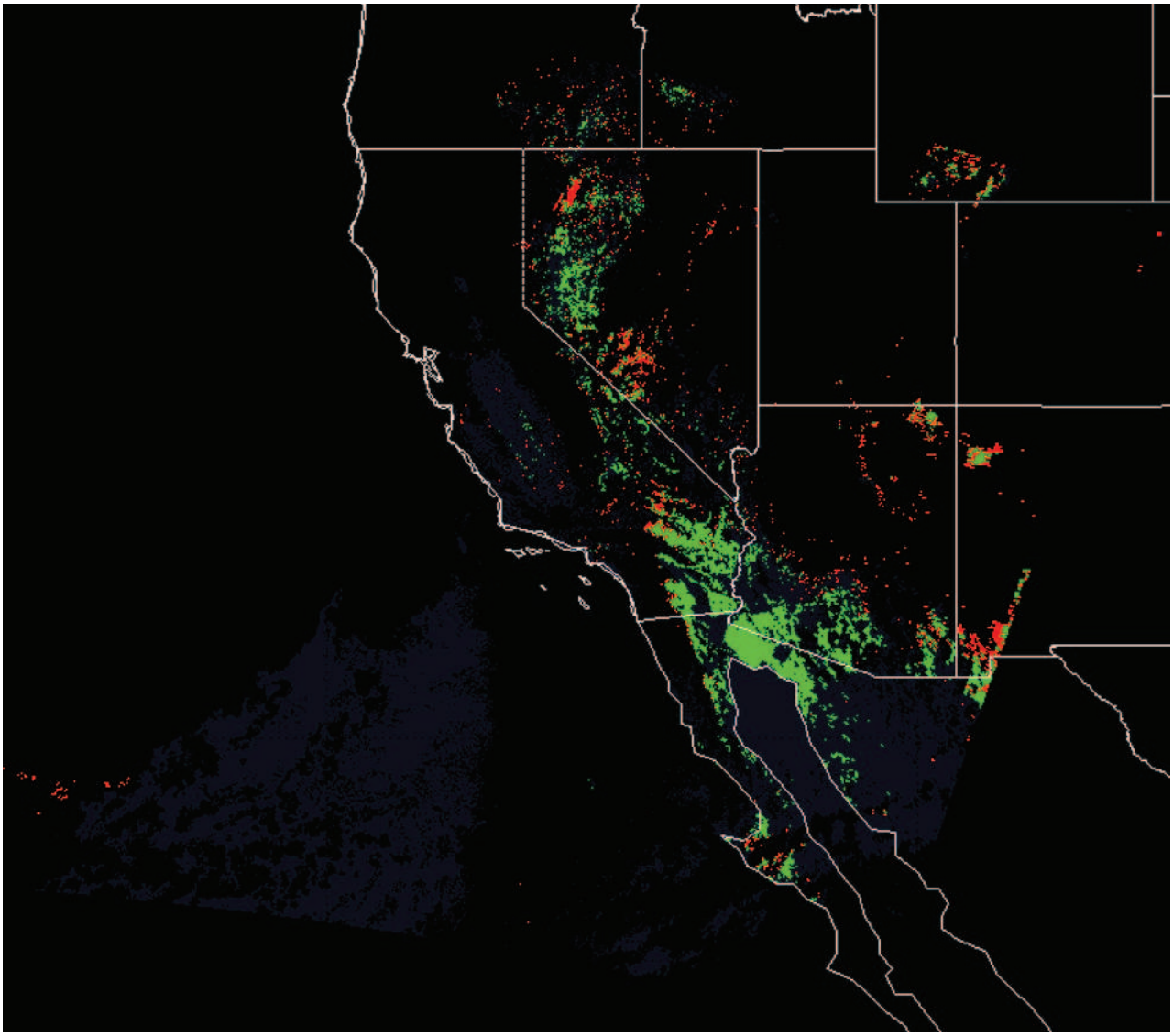

Figure 170. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-105 at 18:50 UTC.

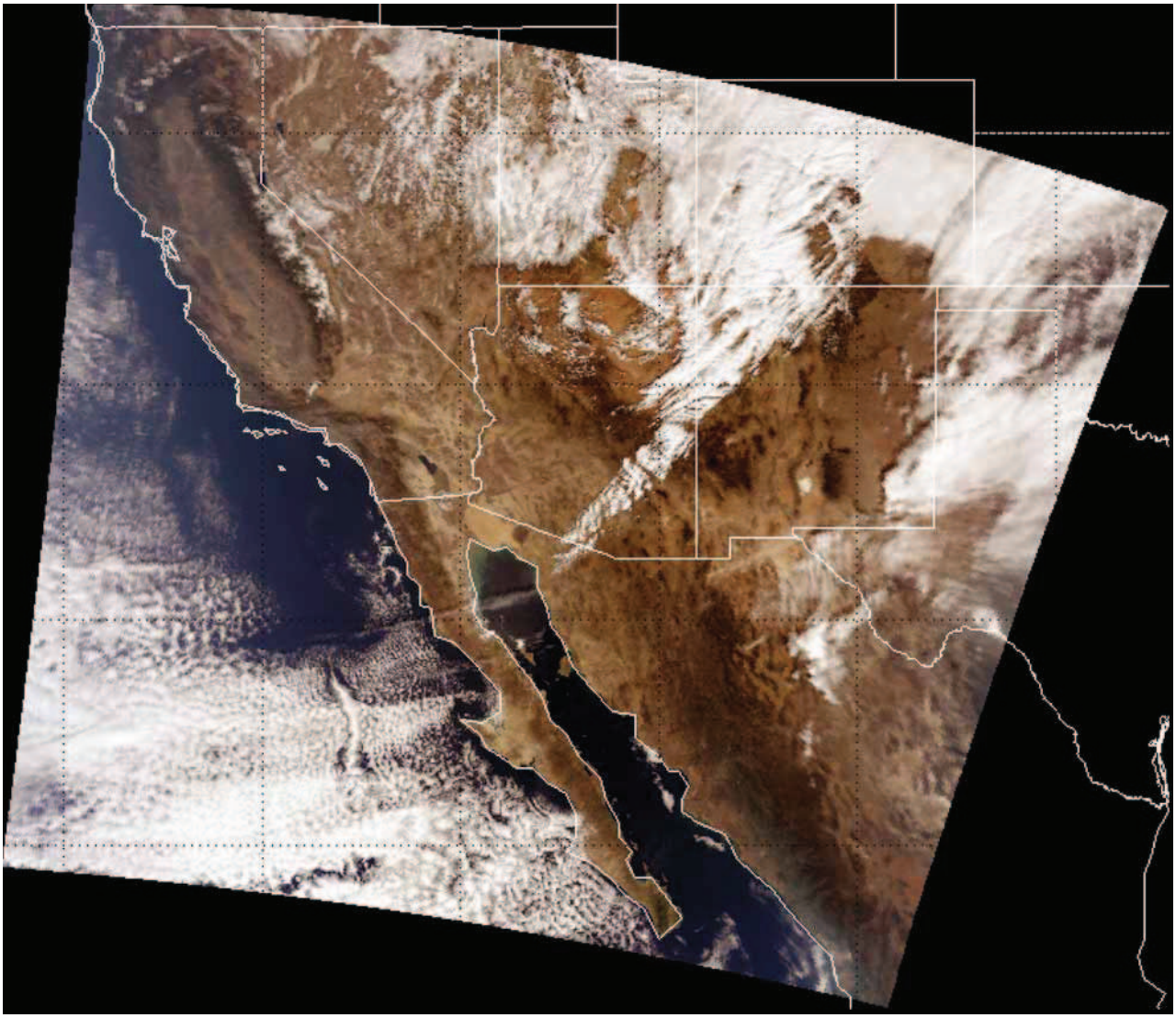

Figure 171. True color image for the event of 2003-326 at 18:20 UTC.

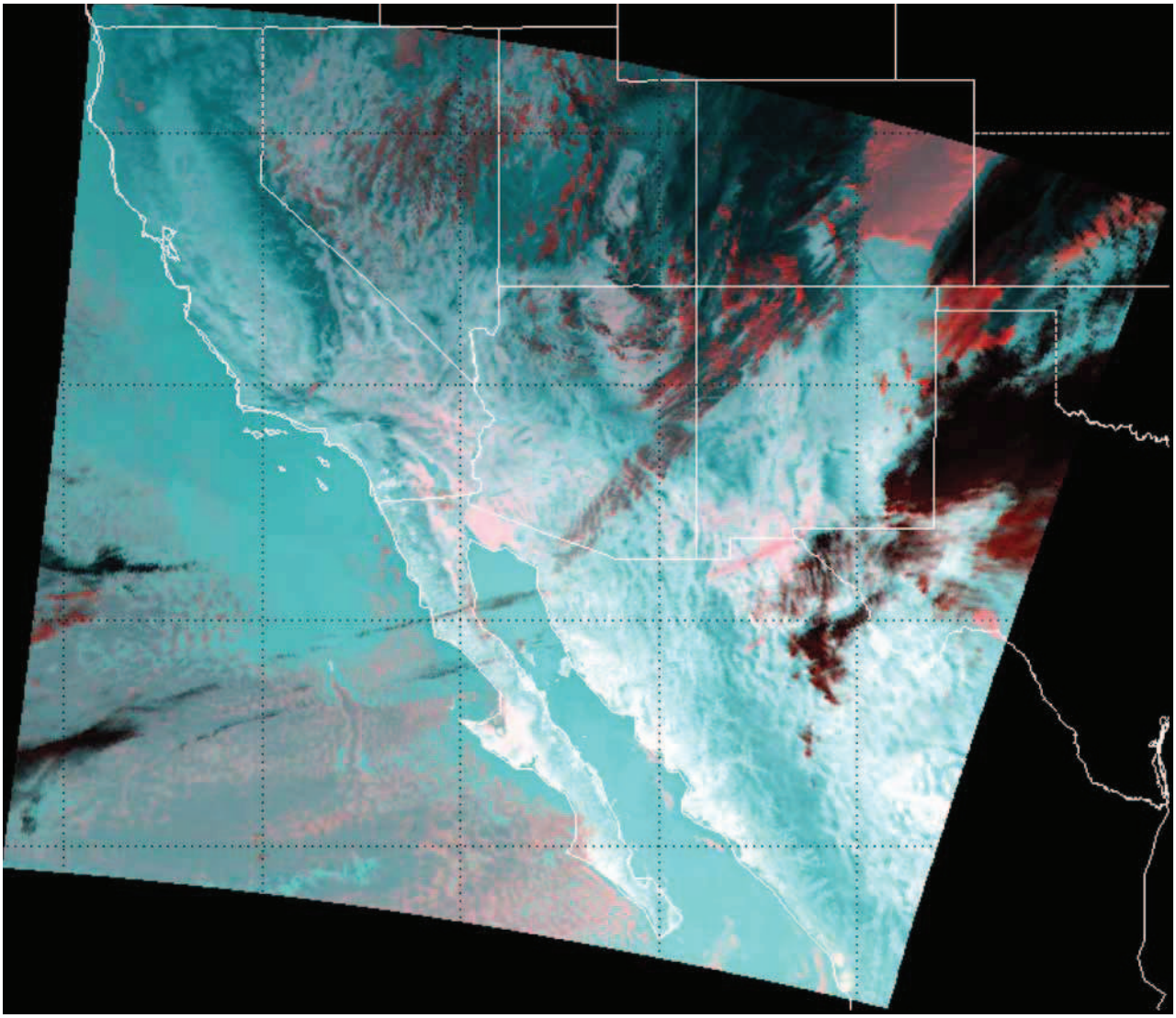

Figure 172. False color image for the event of 2003-326 at 18:20 UTC.

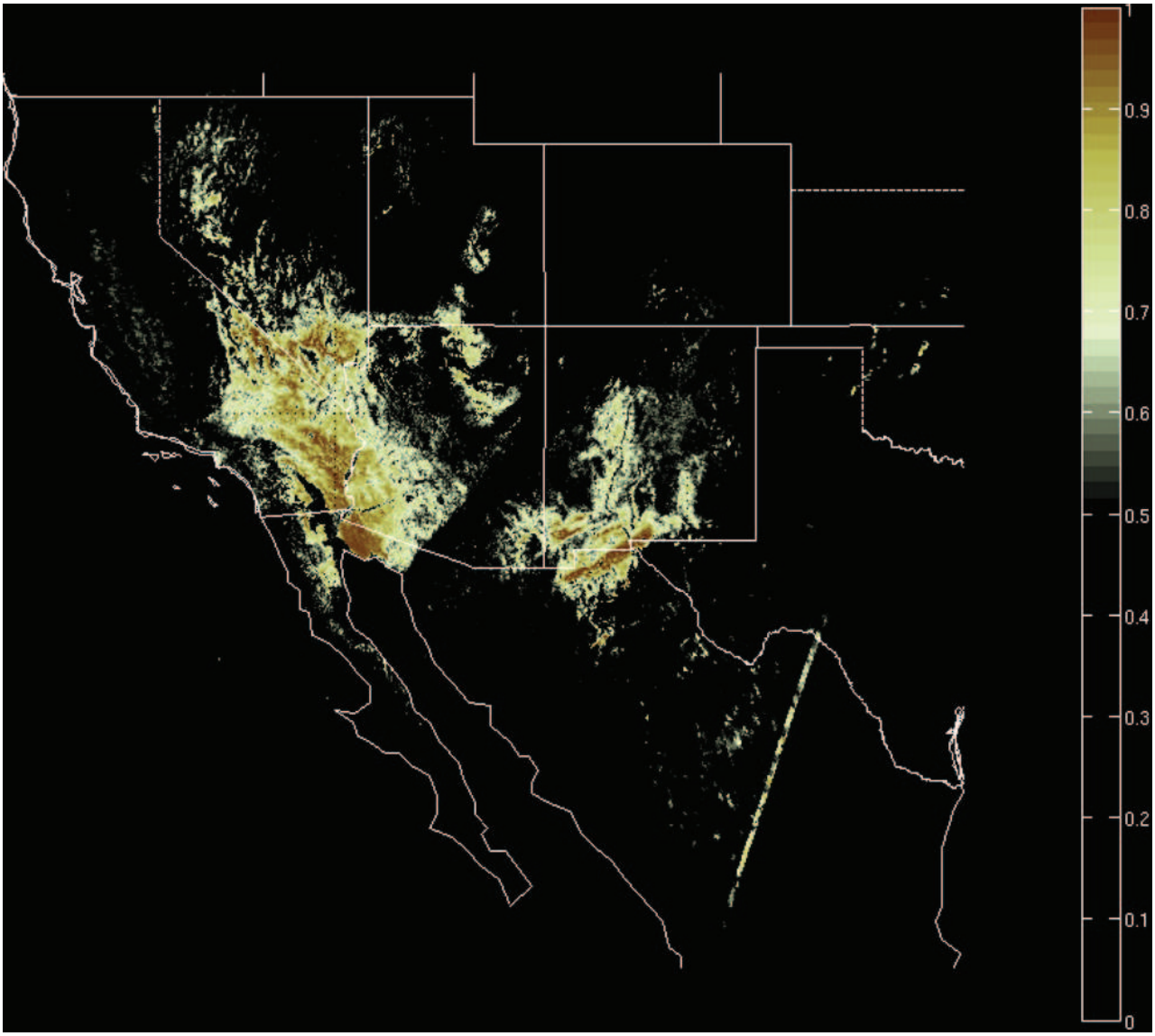

Figure 173. Result of the probabilistic display model (PD) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

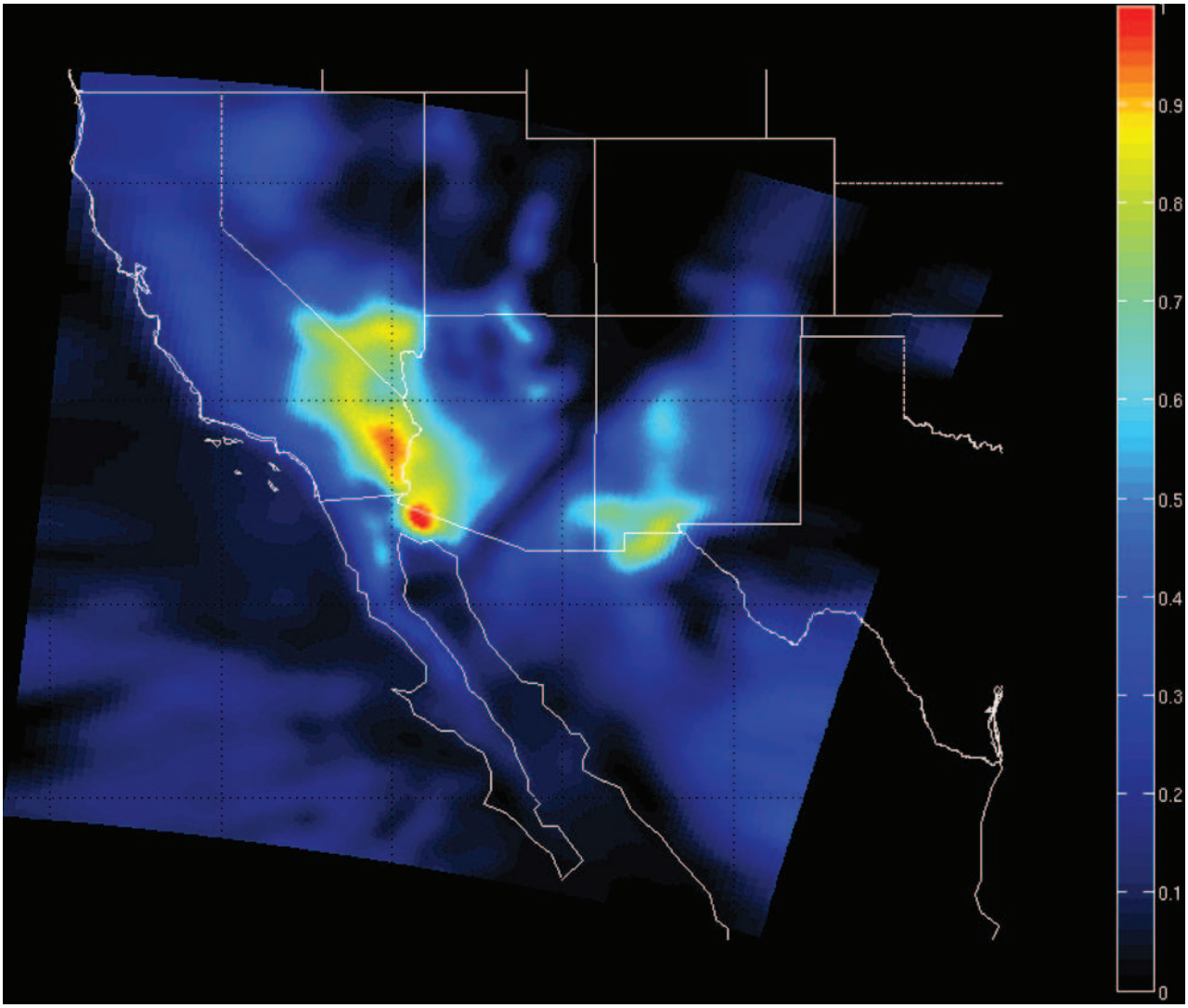

Figure 174. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

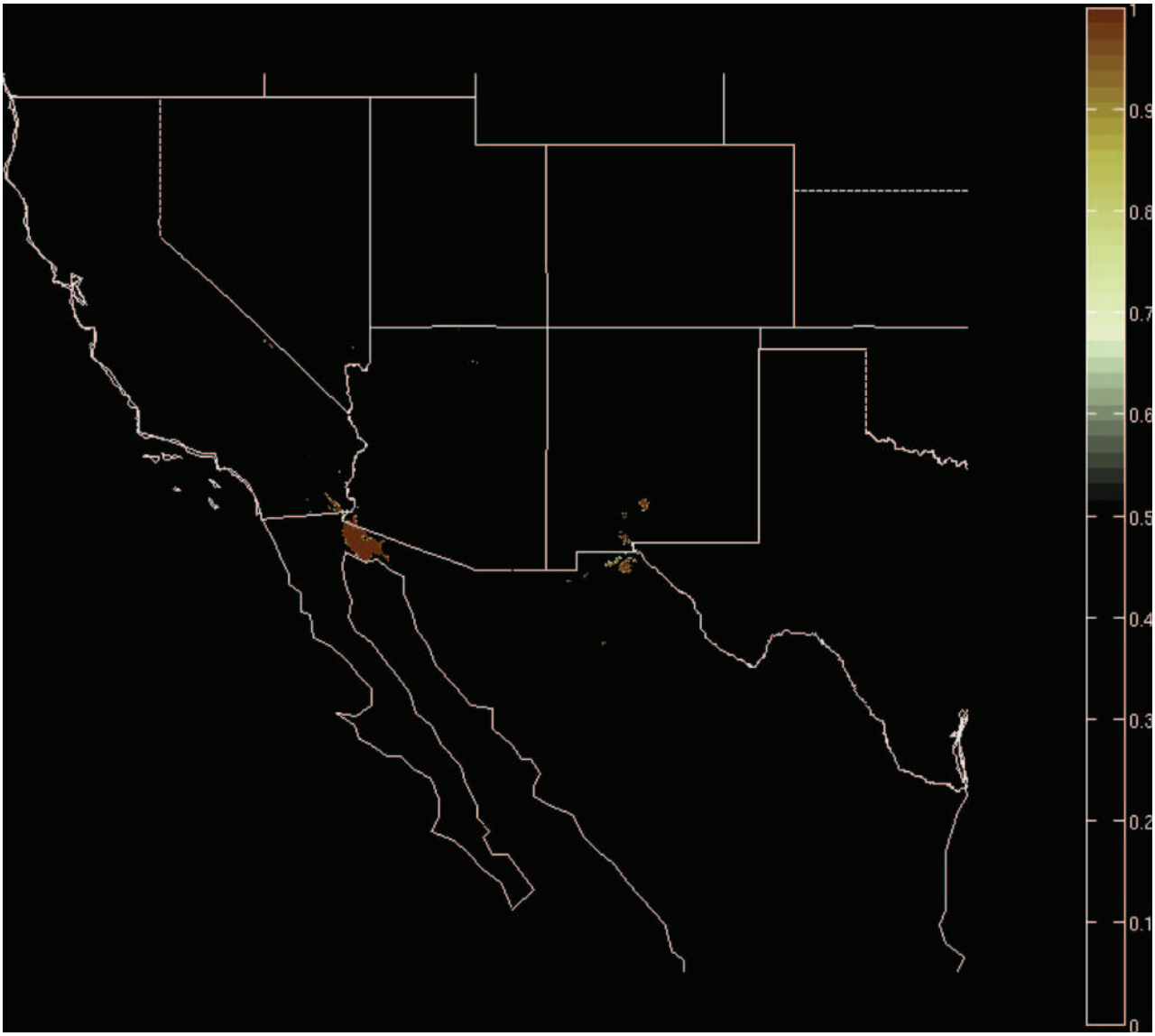

Figure 175. Result of the Maximum Likelihood model (MLE) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

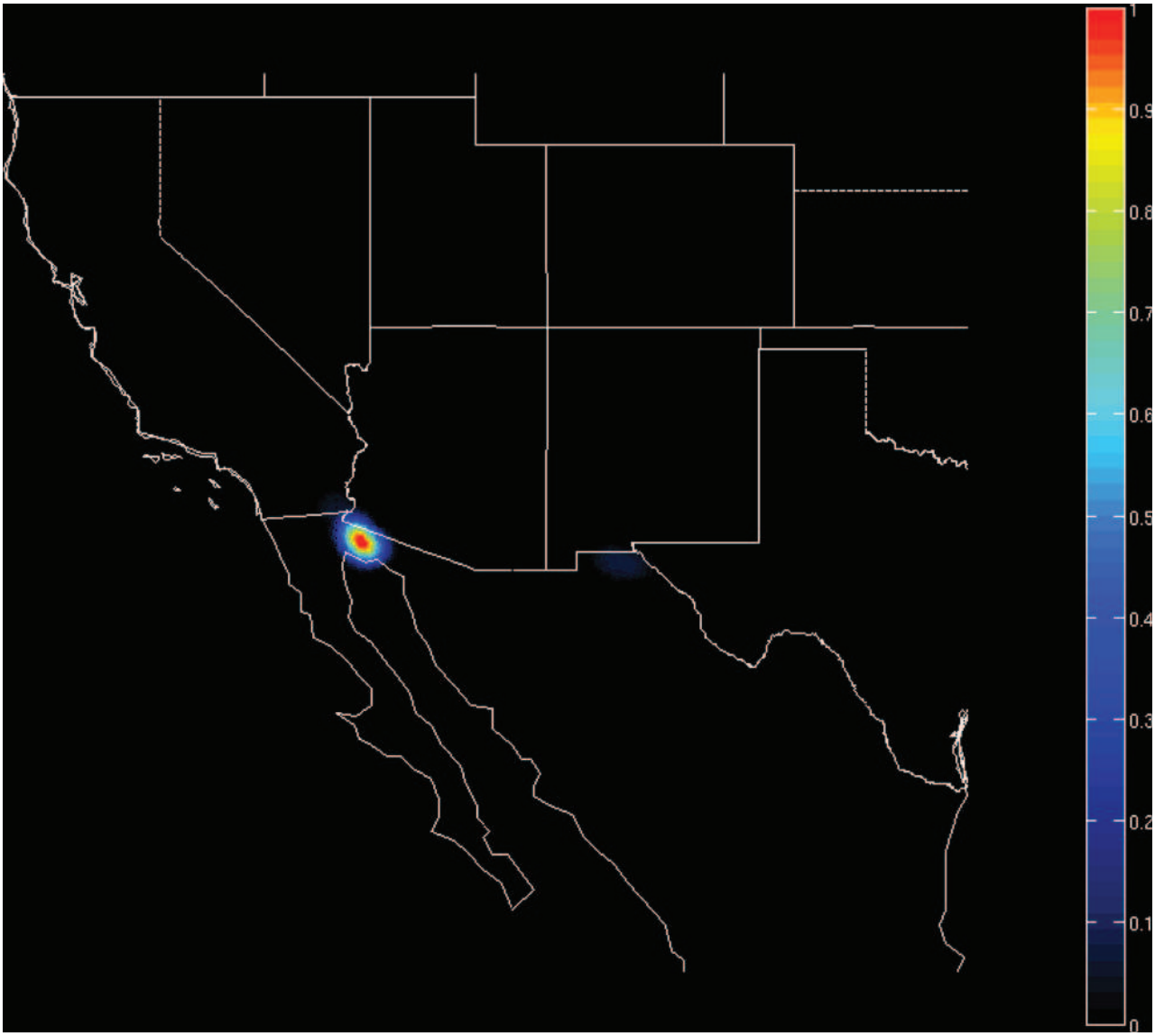

Figure 176. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

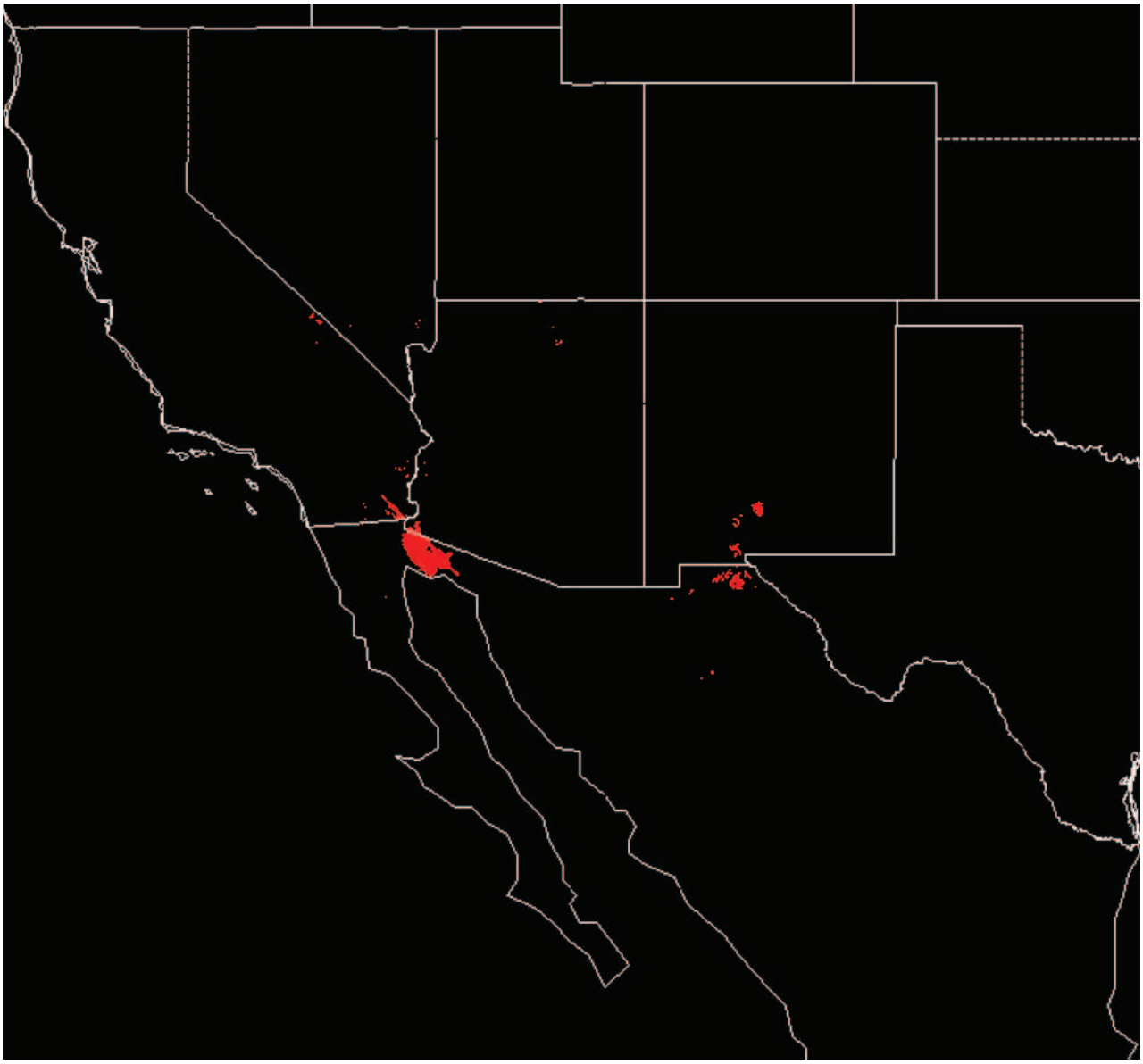

Figure 177. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-326 at 18:20 UTC.

Figure 178. Result of the Maximum a posteriori model (MAP) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

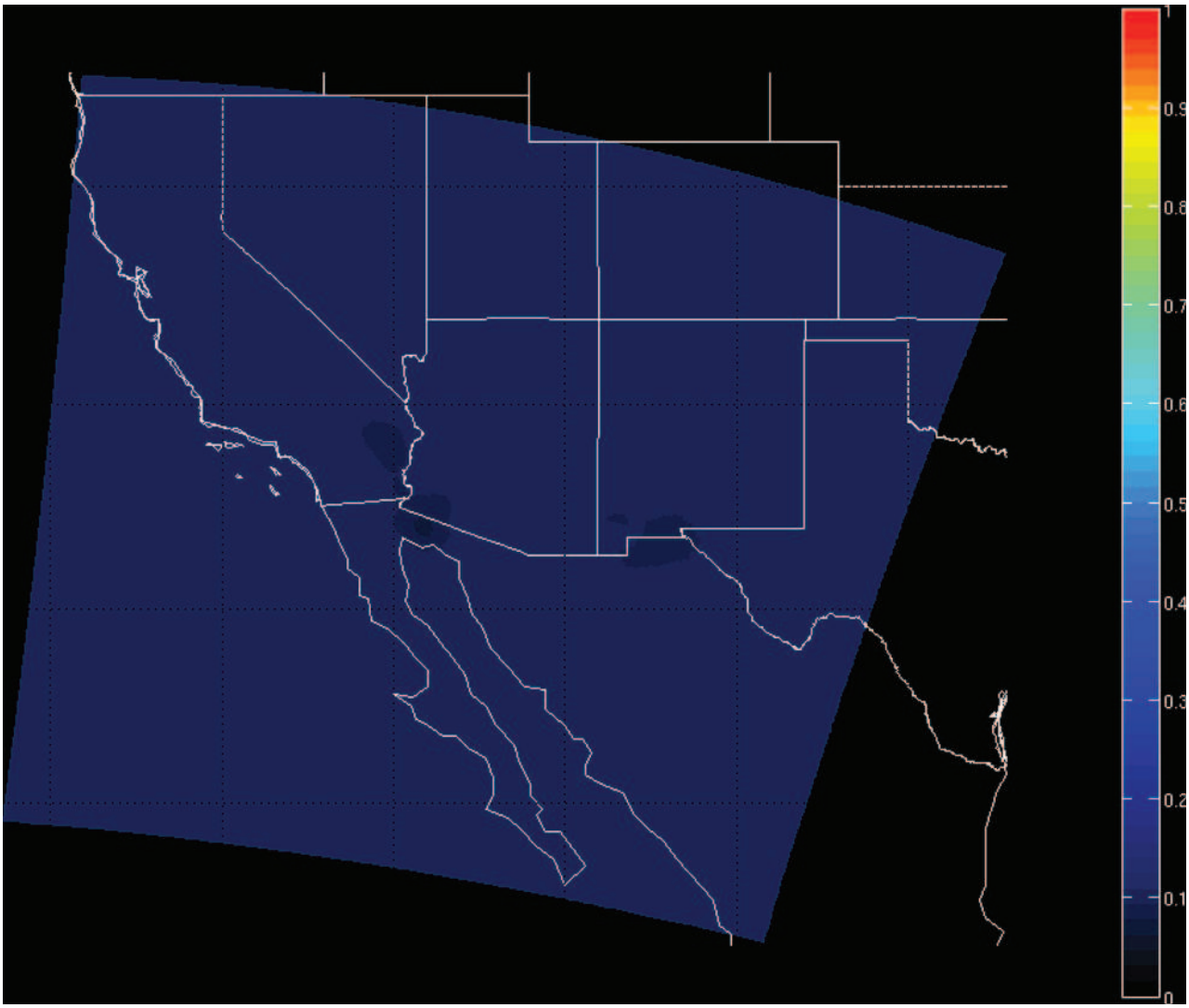

Figure 179. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

Figure 180. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-326 at 18:20 UTC.

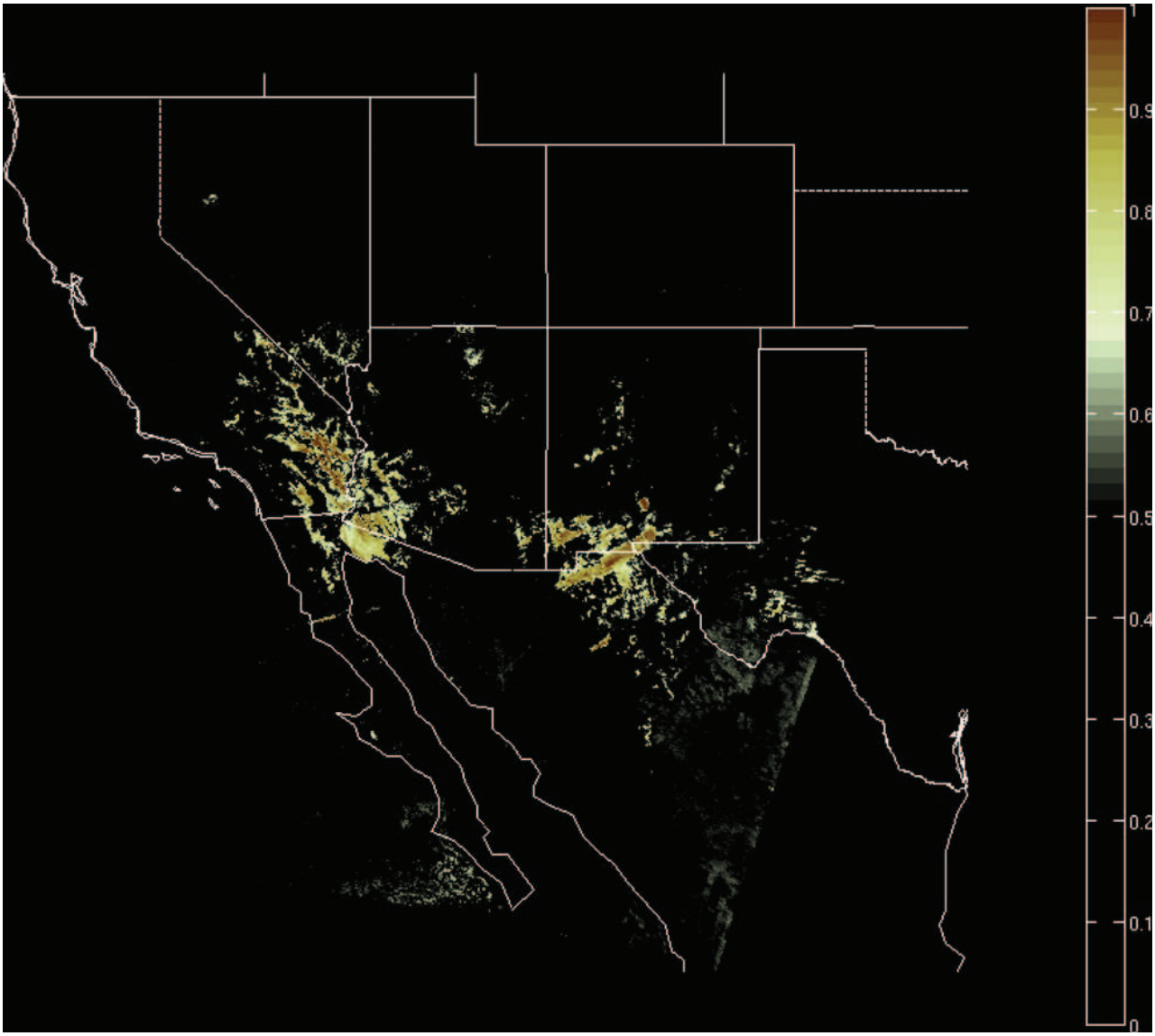

Figure 181. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

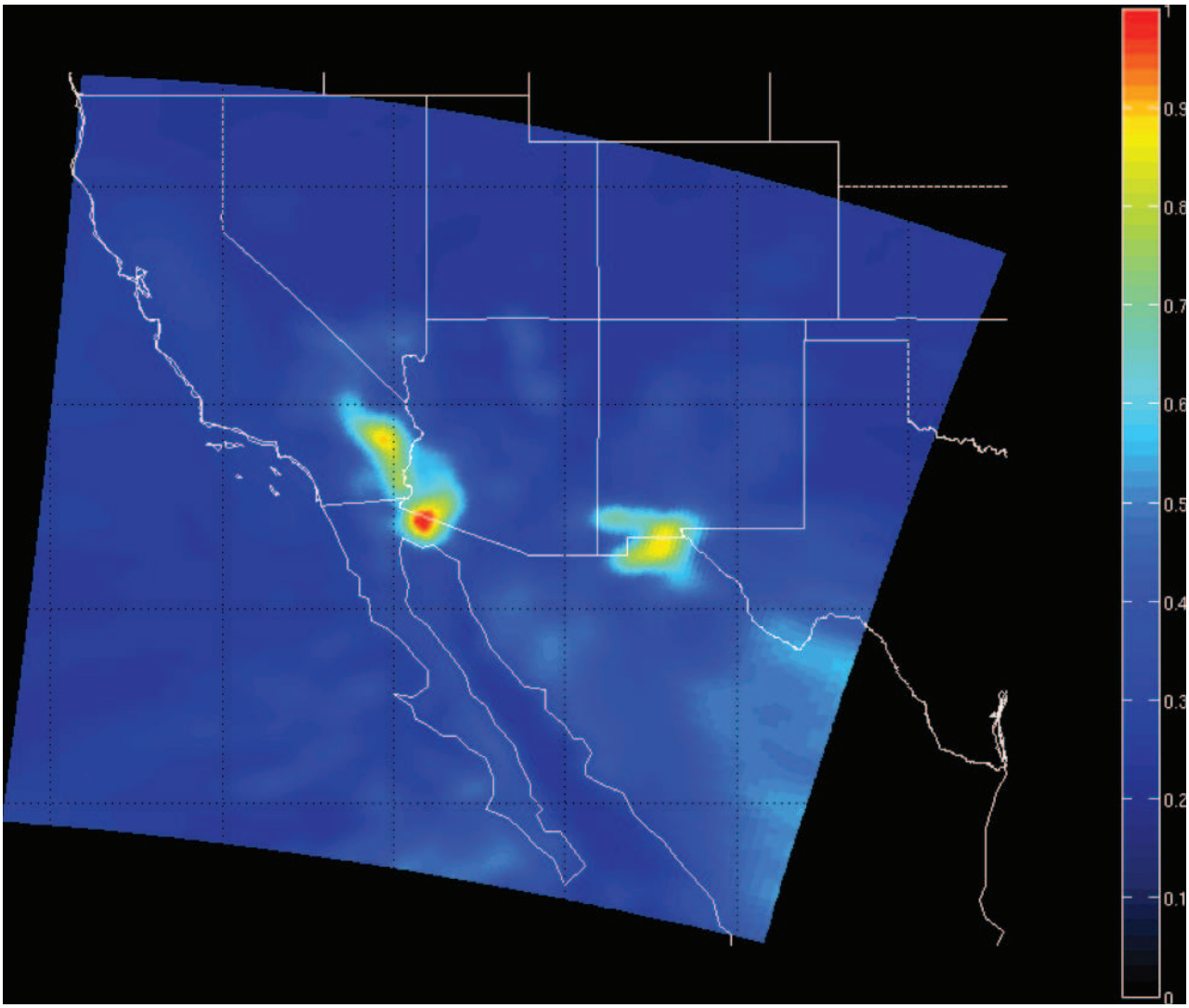

Figure 182. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

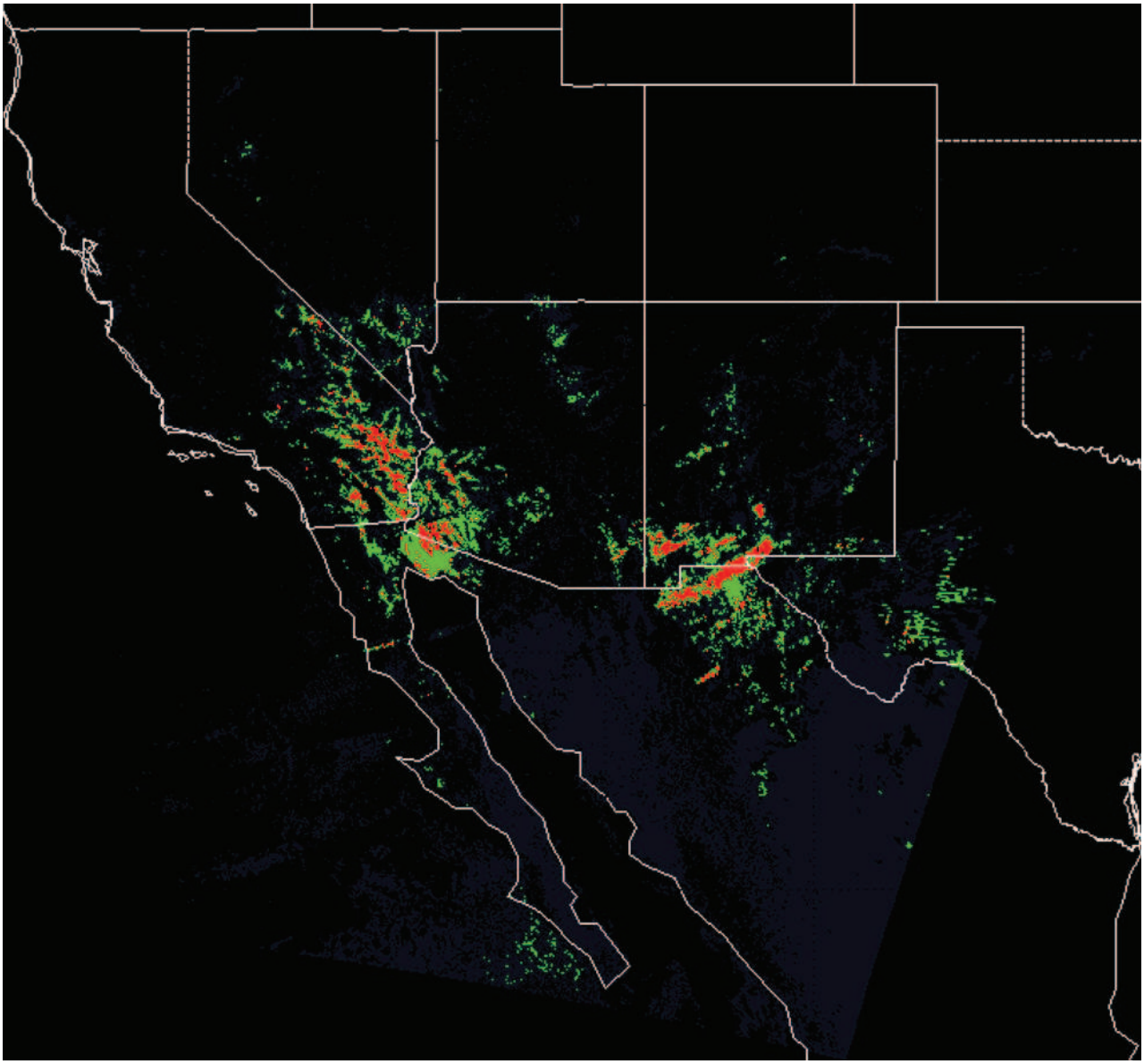

Figure 183. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-326 at 18:20 UTC.

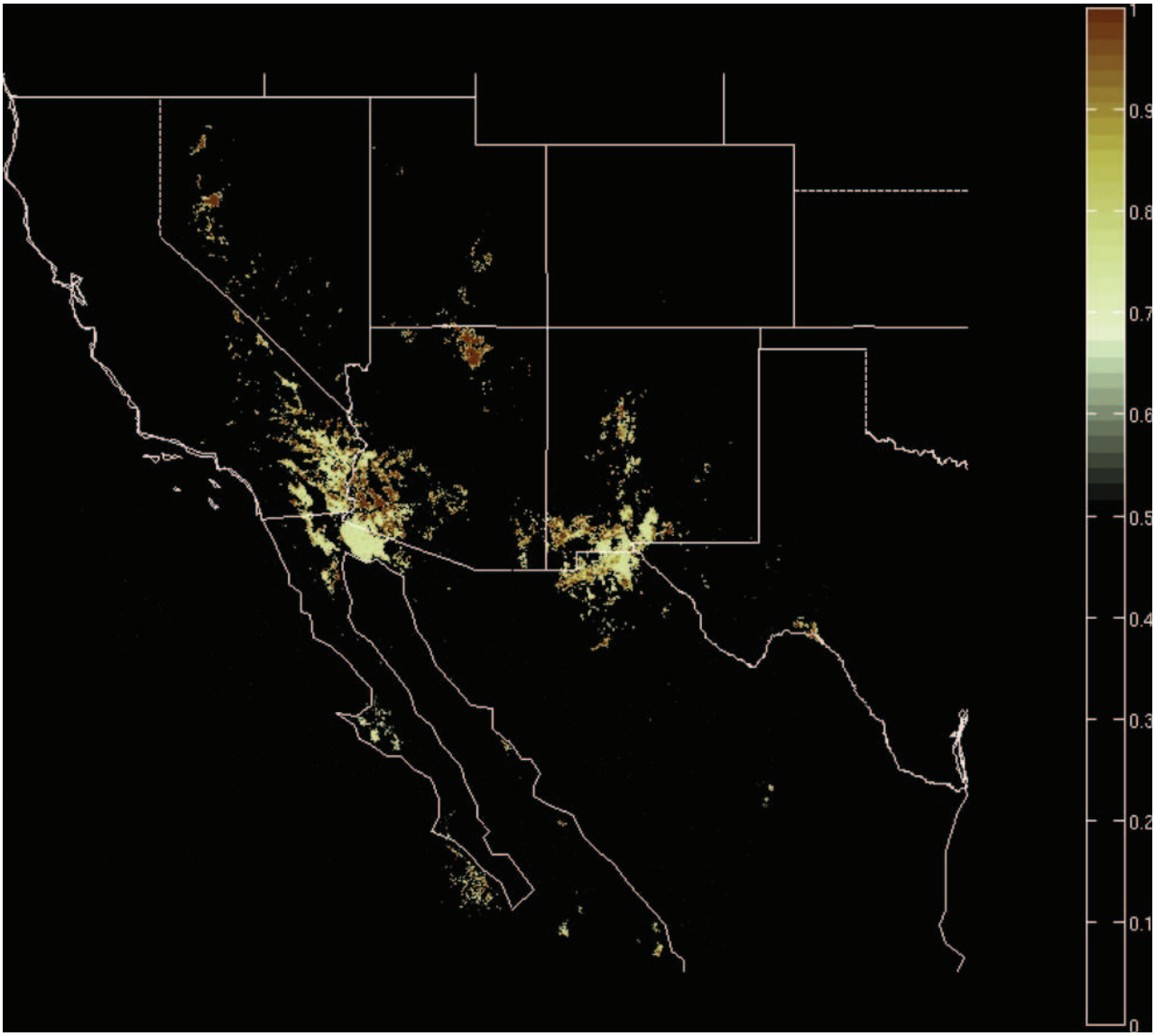

Figure 184. Result of the Probabilistic Neural Network (PNN) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

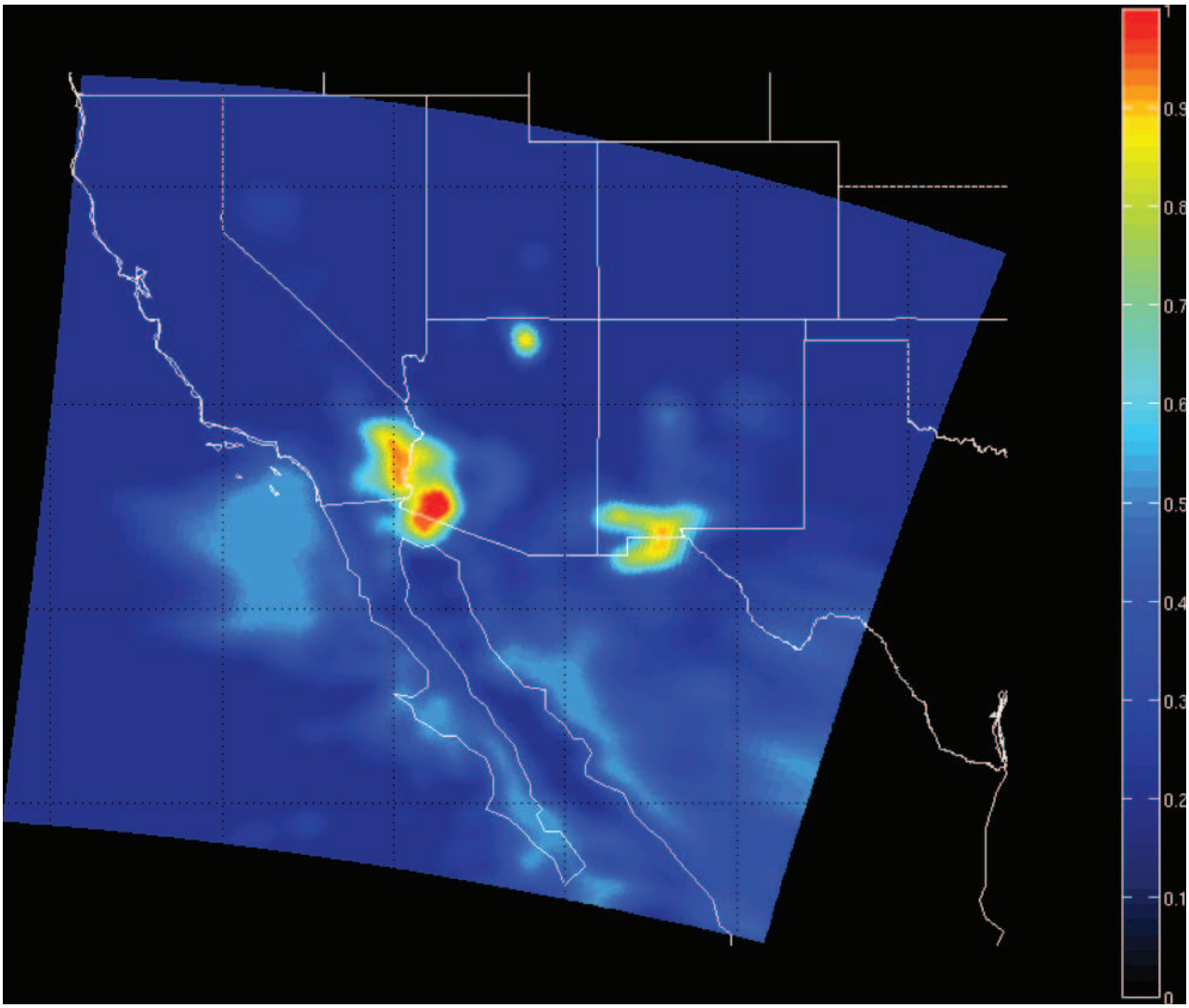

Figure 185. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-326 at 18:20 UTC. The colorbar shows the index of probability.

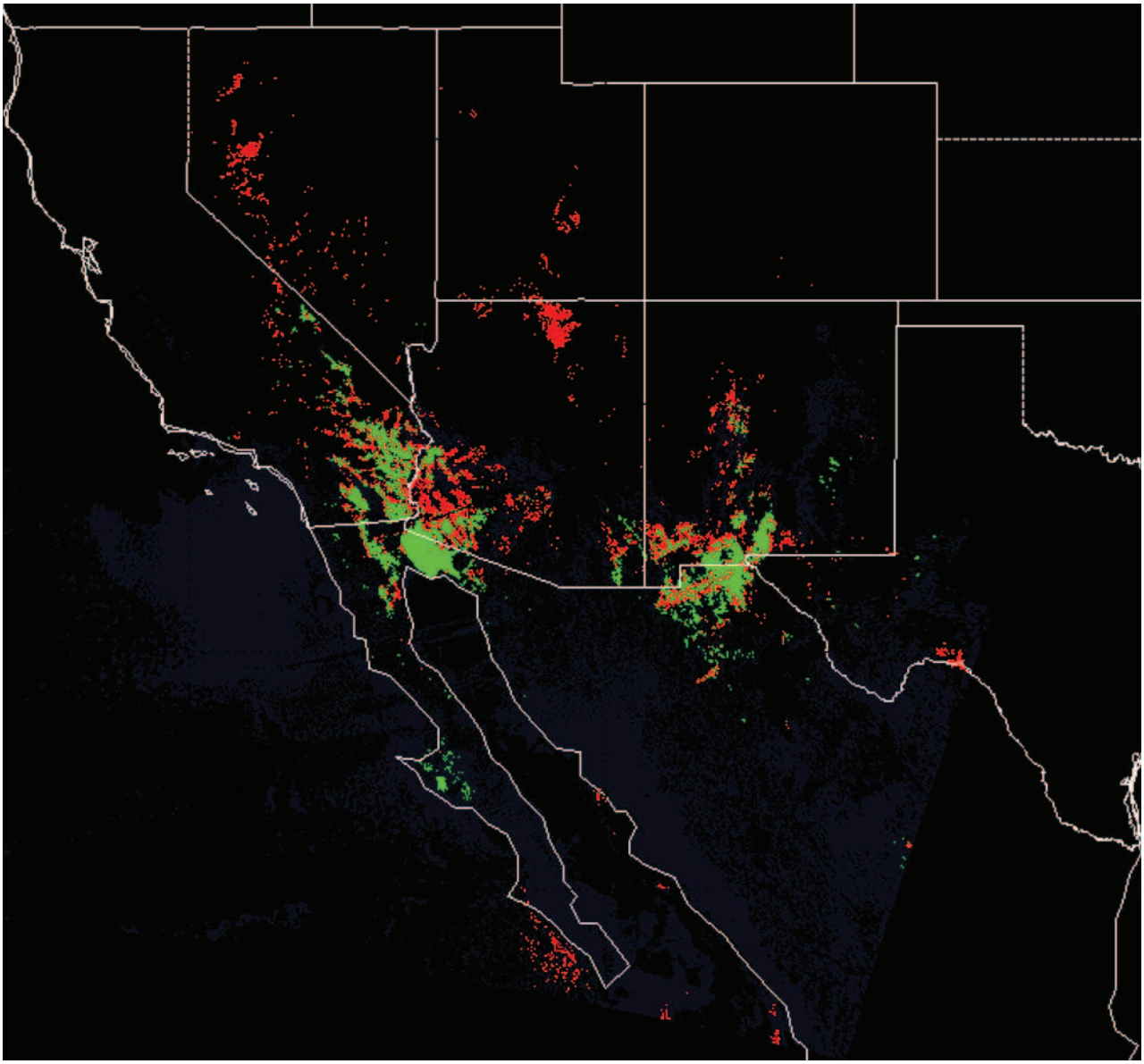

Figure 186. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-326 at 18:20 UTC.

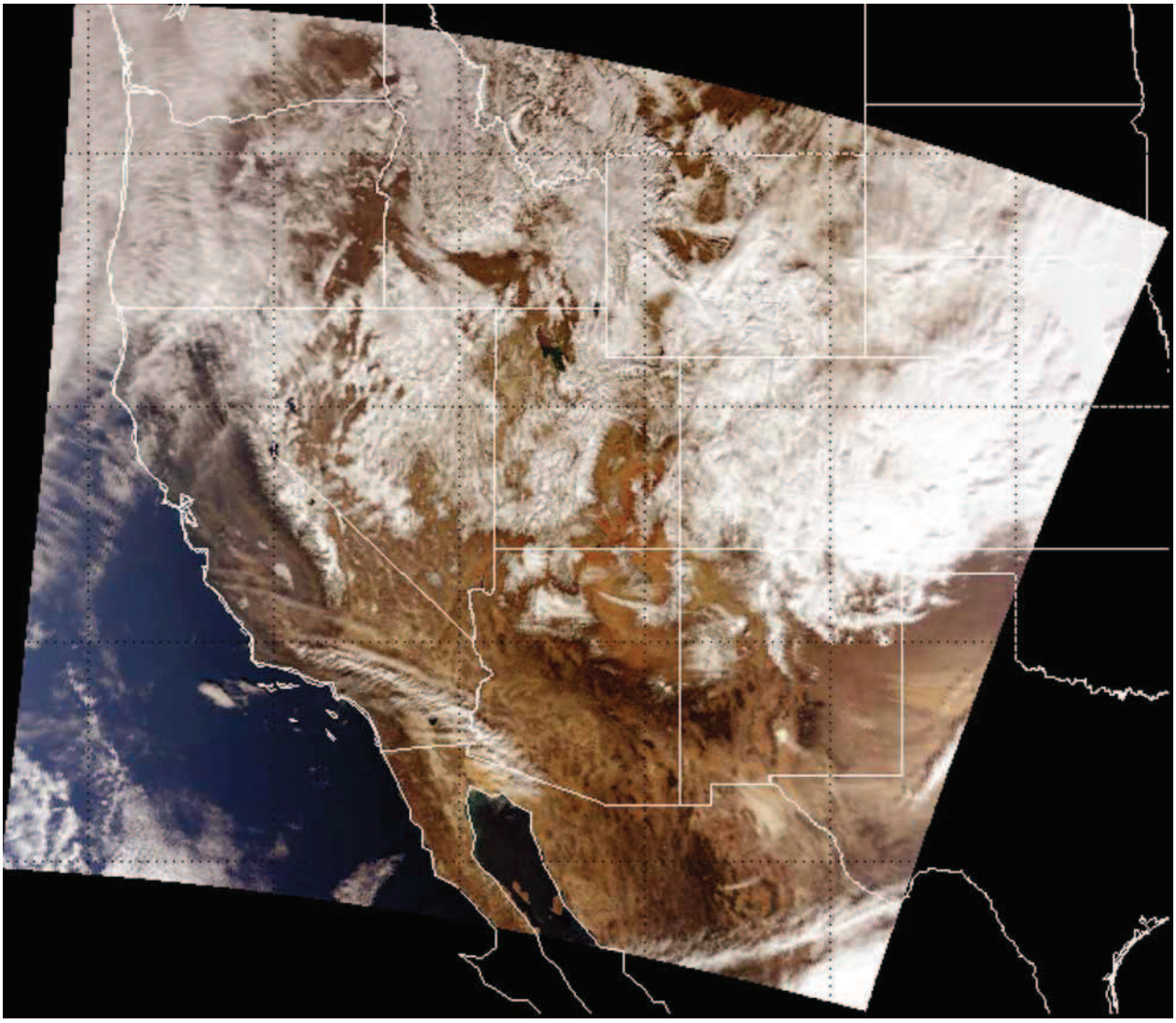

Figure 187. True color image for the event of 2003-349 at 18:25 UTC.

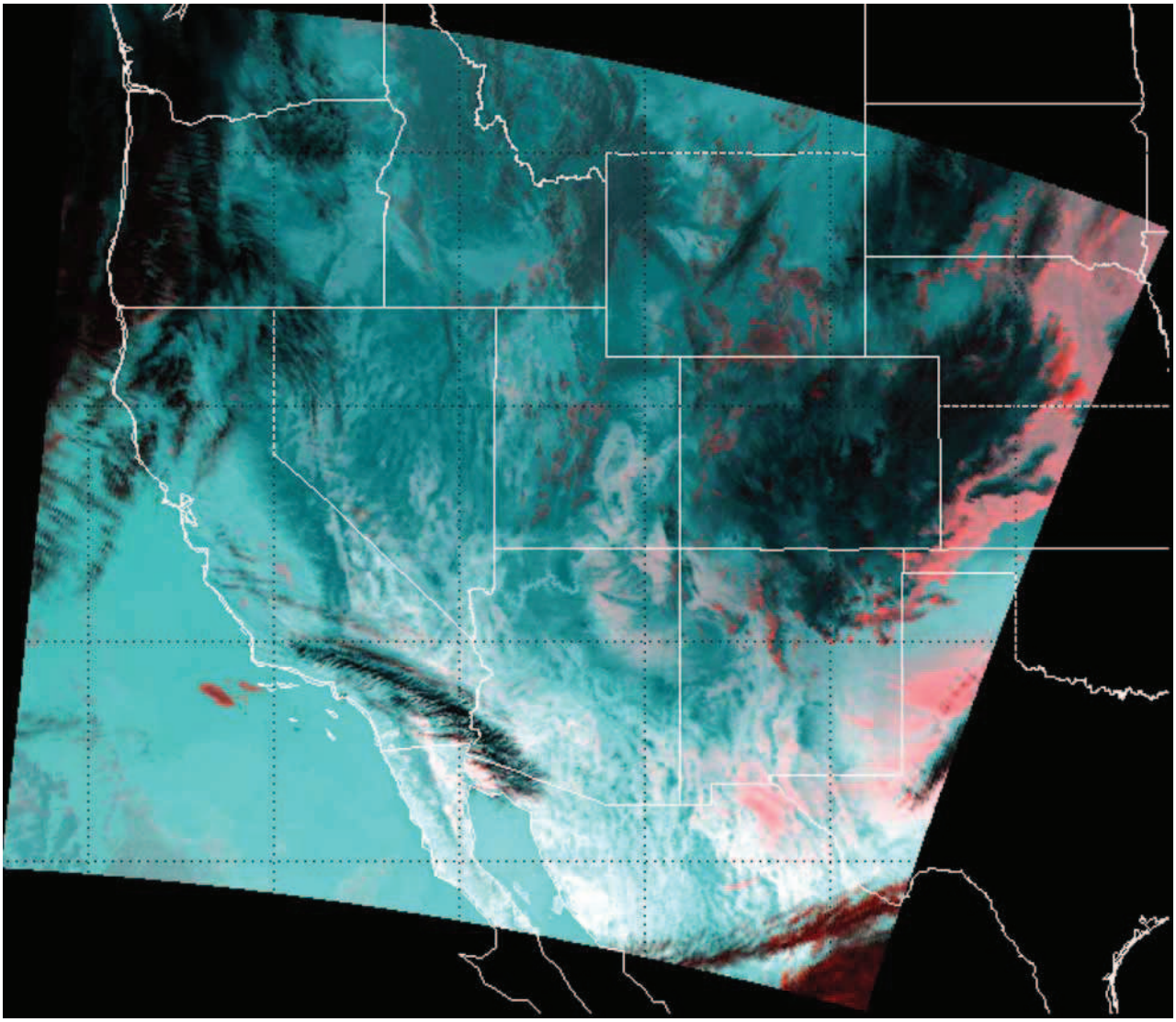

Figure 188. False color image for the event of 2003-349 at 18:25 UTC.

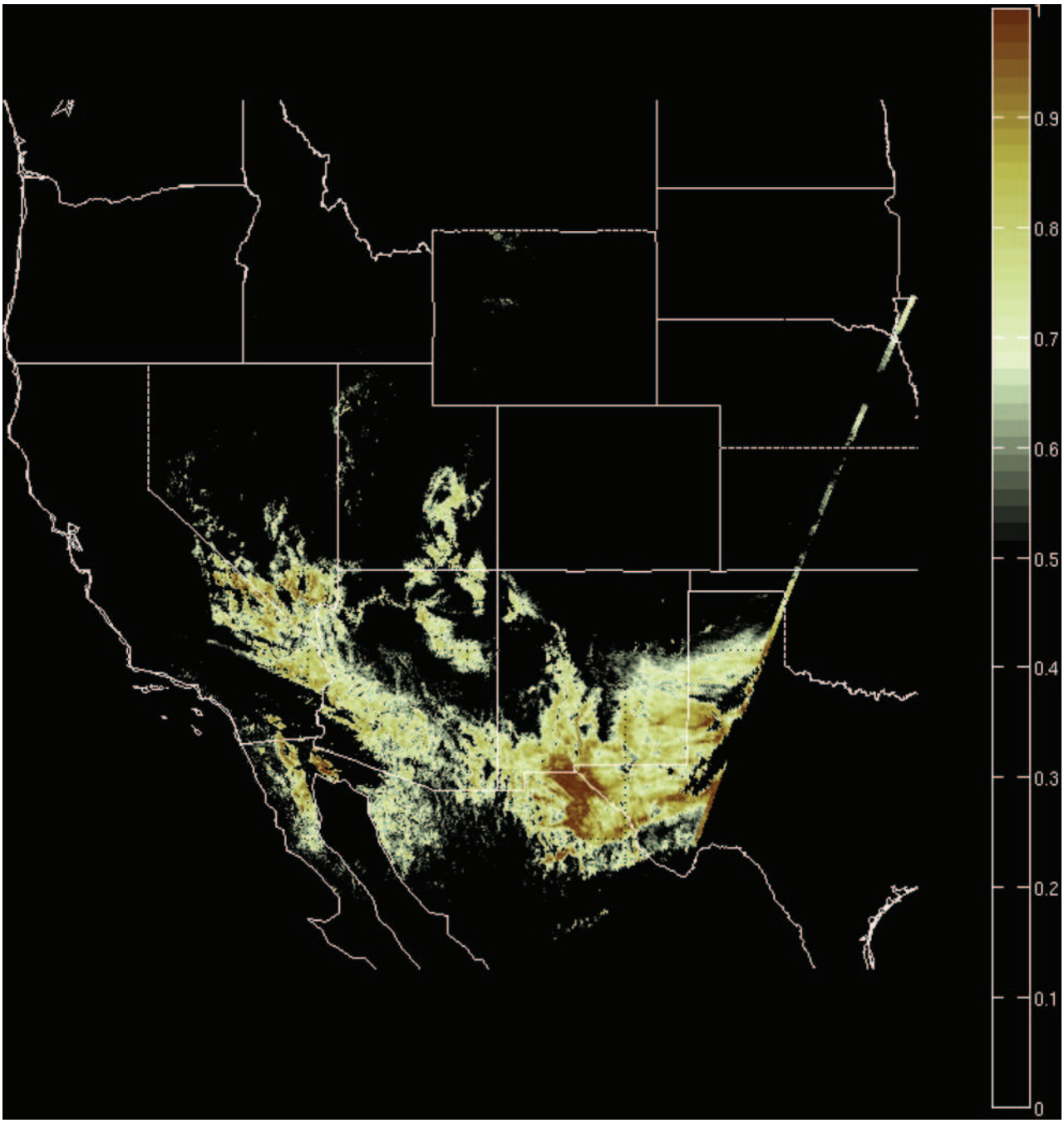

Figure 189. Result of the probabilistic display model (PD) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

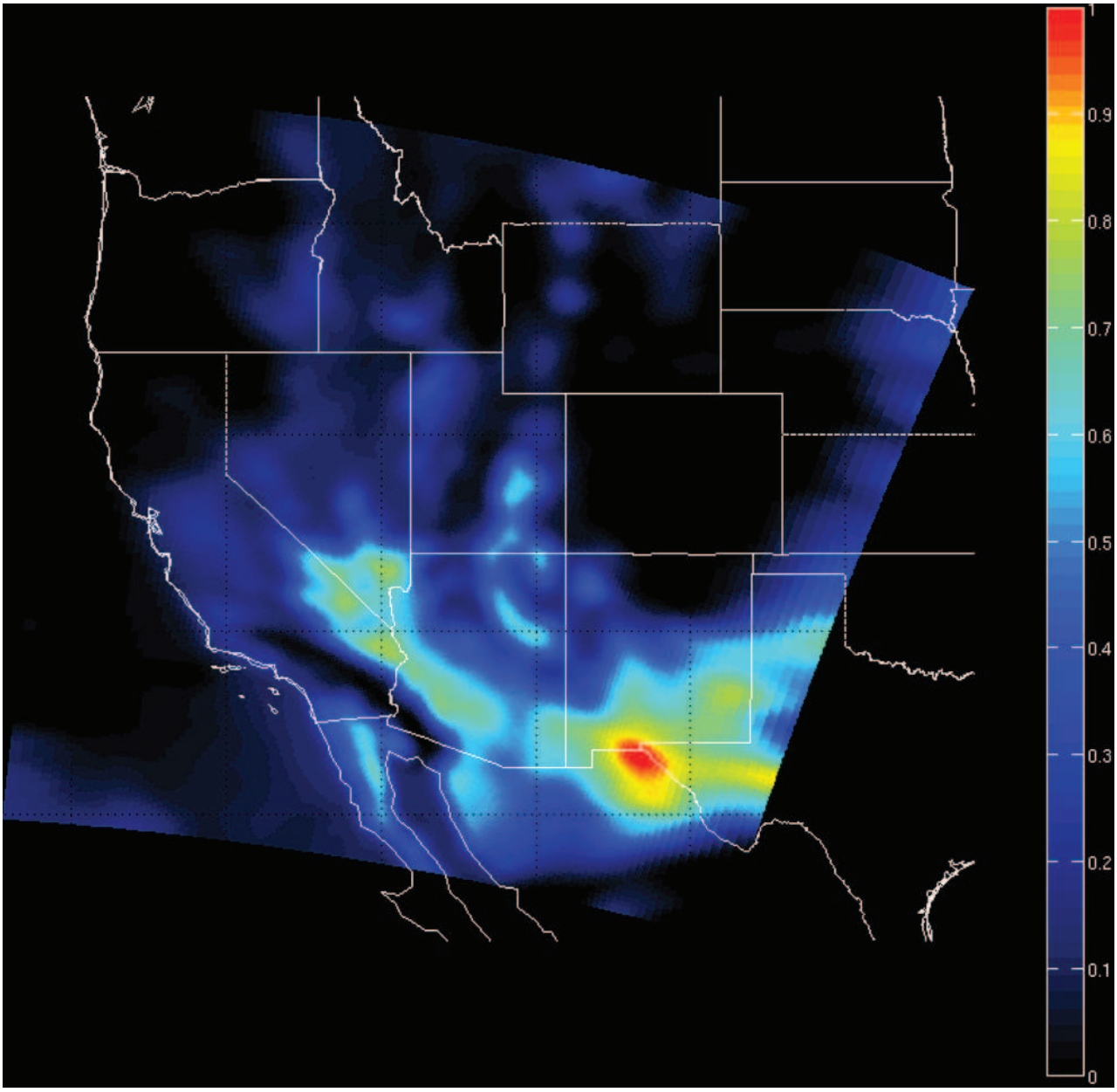

Figure 190. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

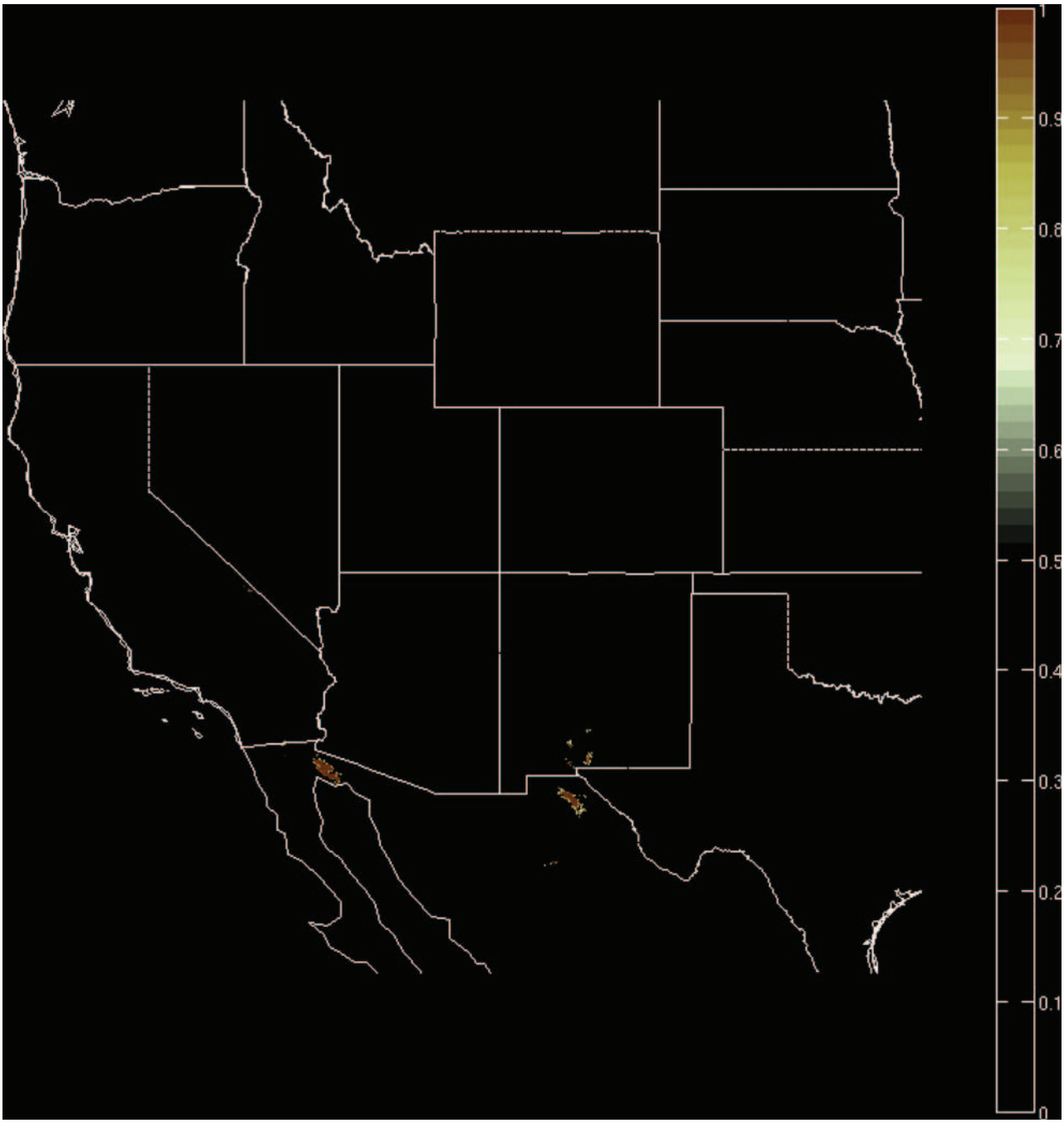

Figure 191. Result of the Maximum Likelihood model (MLE) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

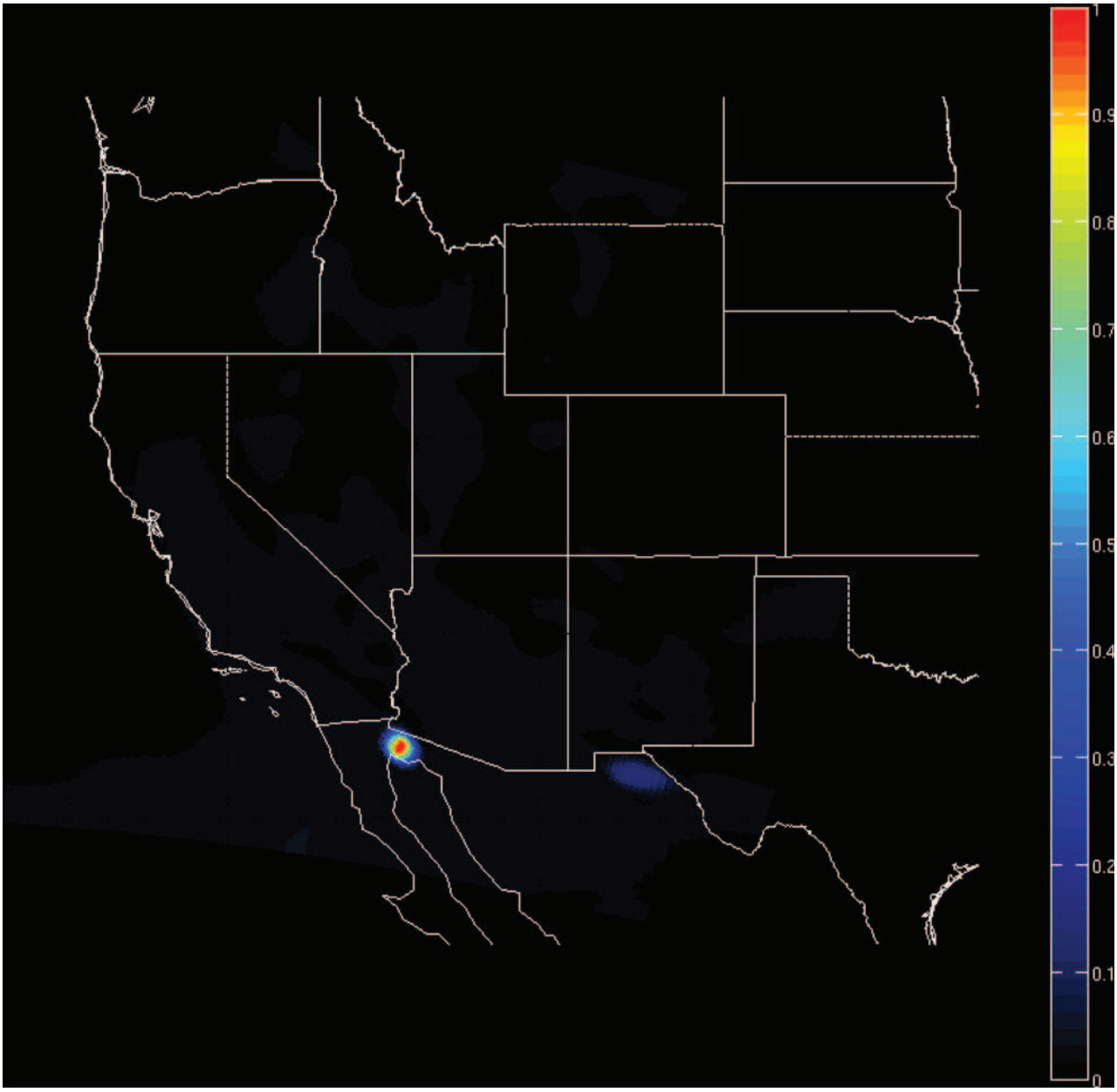

Figure 192. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

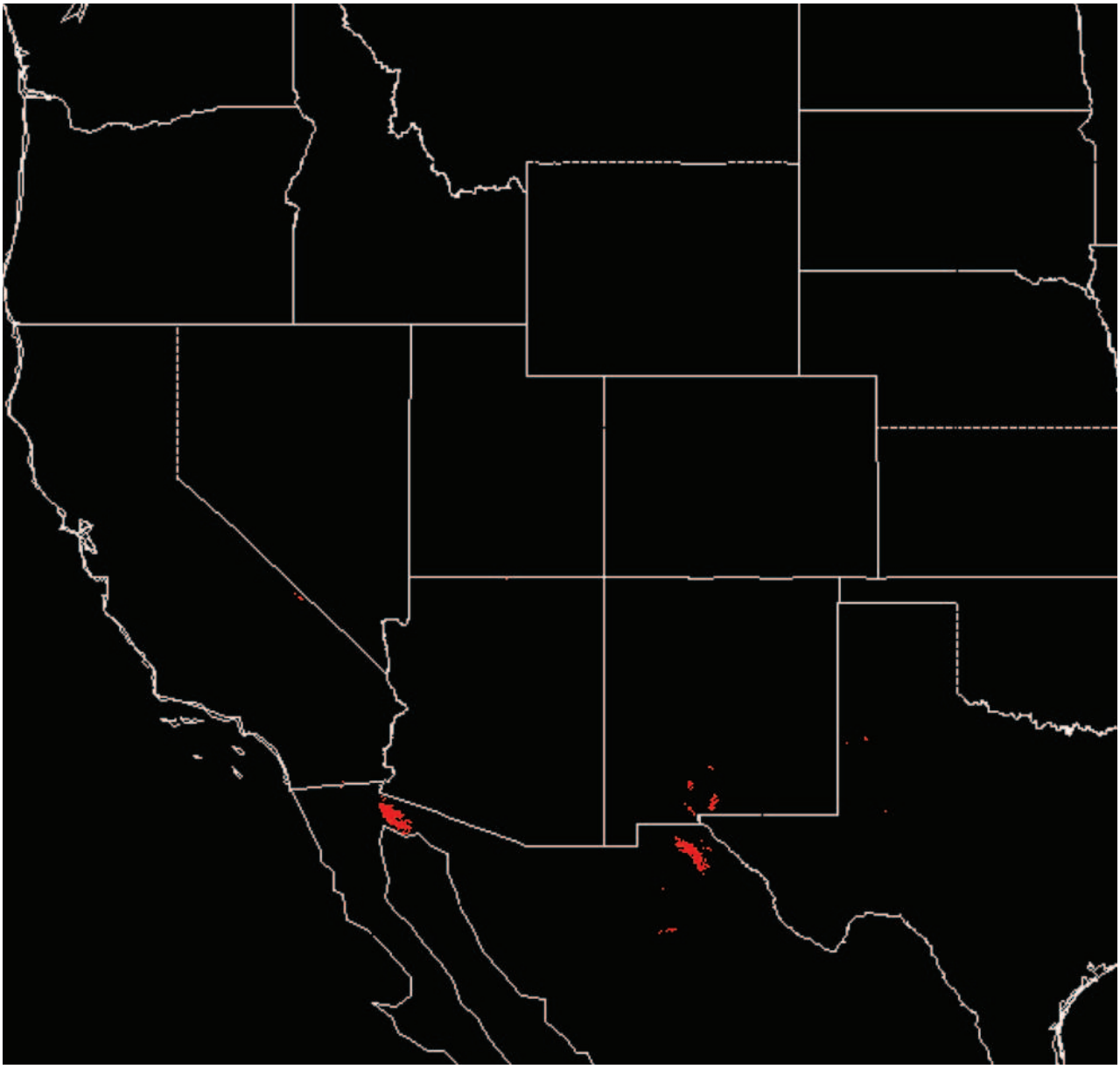

Figure 193. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-349 at 18:25 UTC.

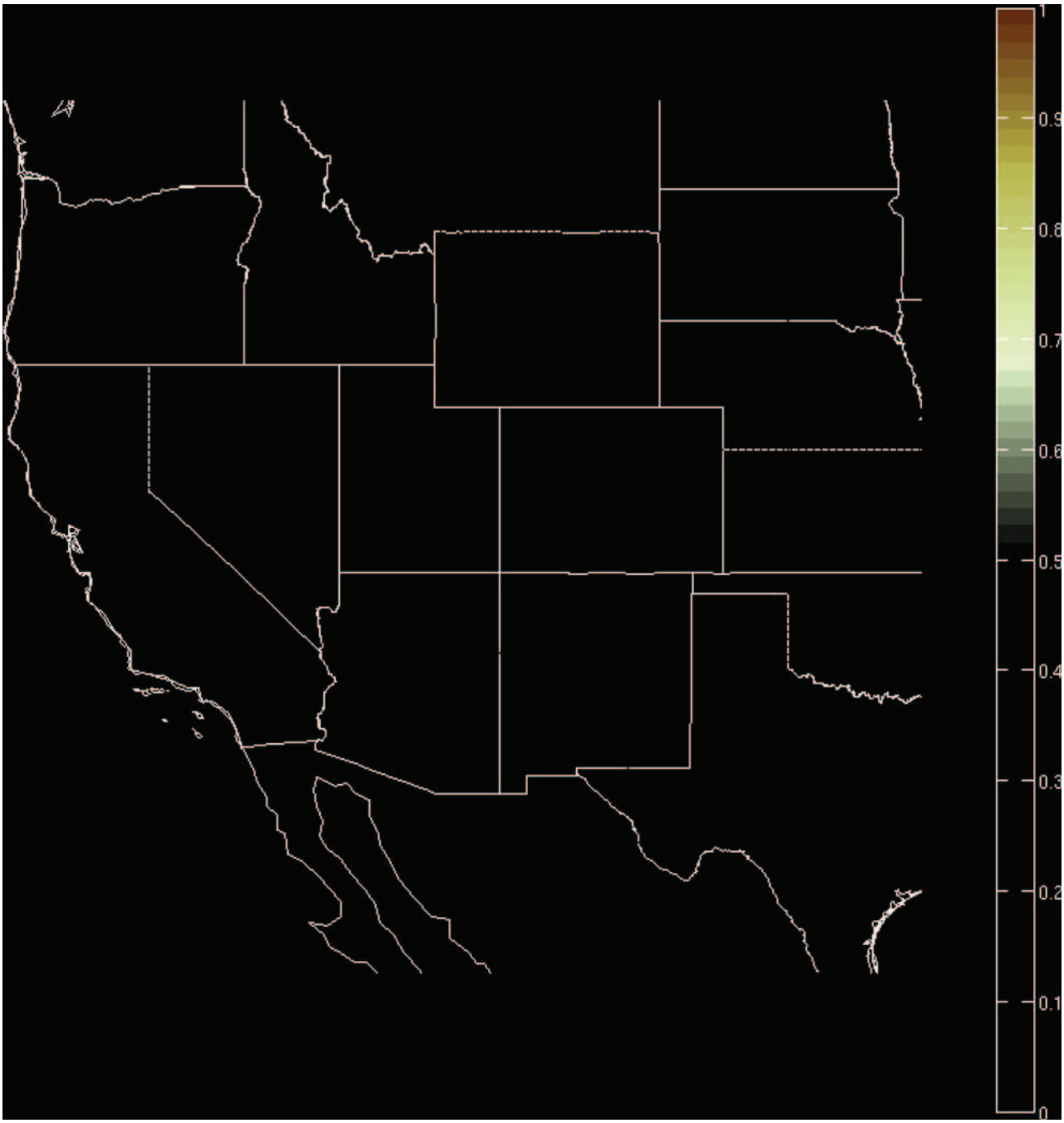

Figure 194. Result of the Maximum a posteriori model (MAP) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

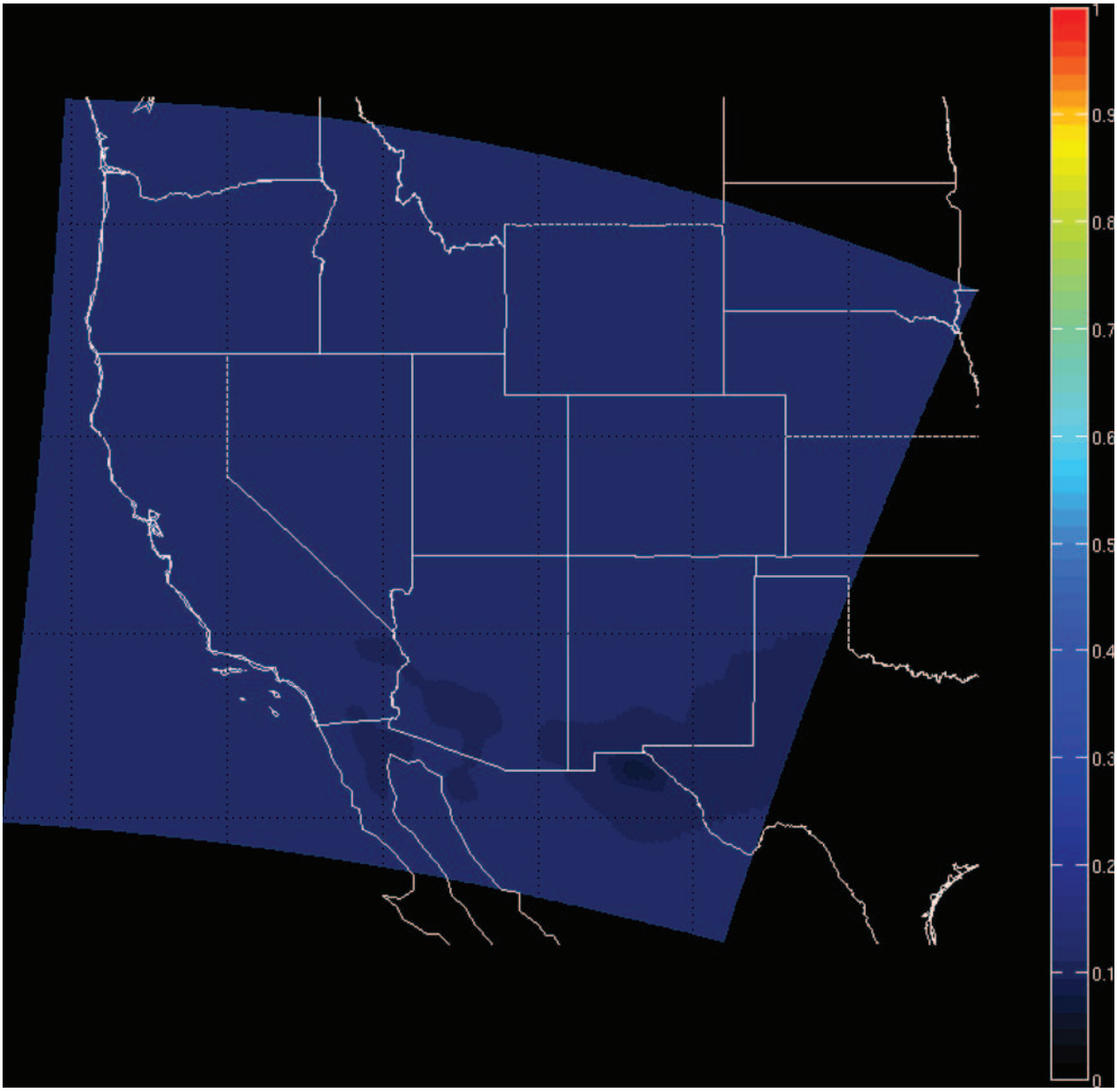

Figure 195. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

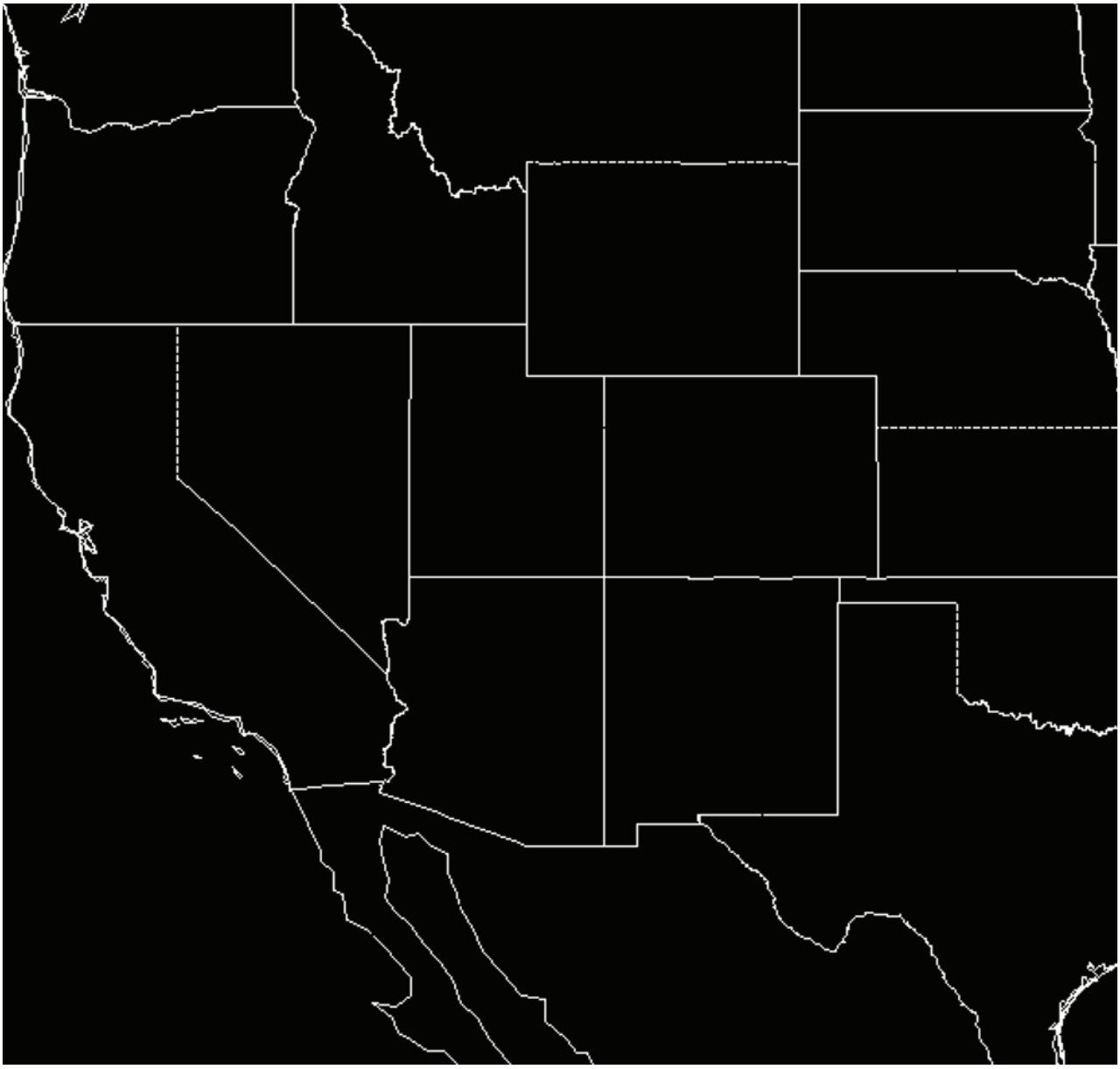

Figure 196. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-349 at 18:25 UTC.

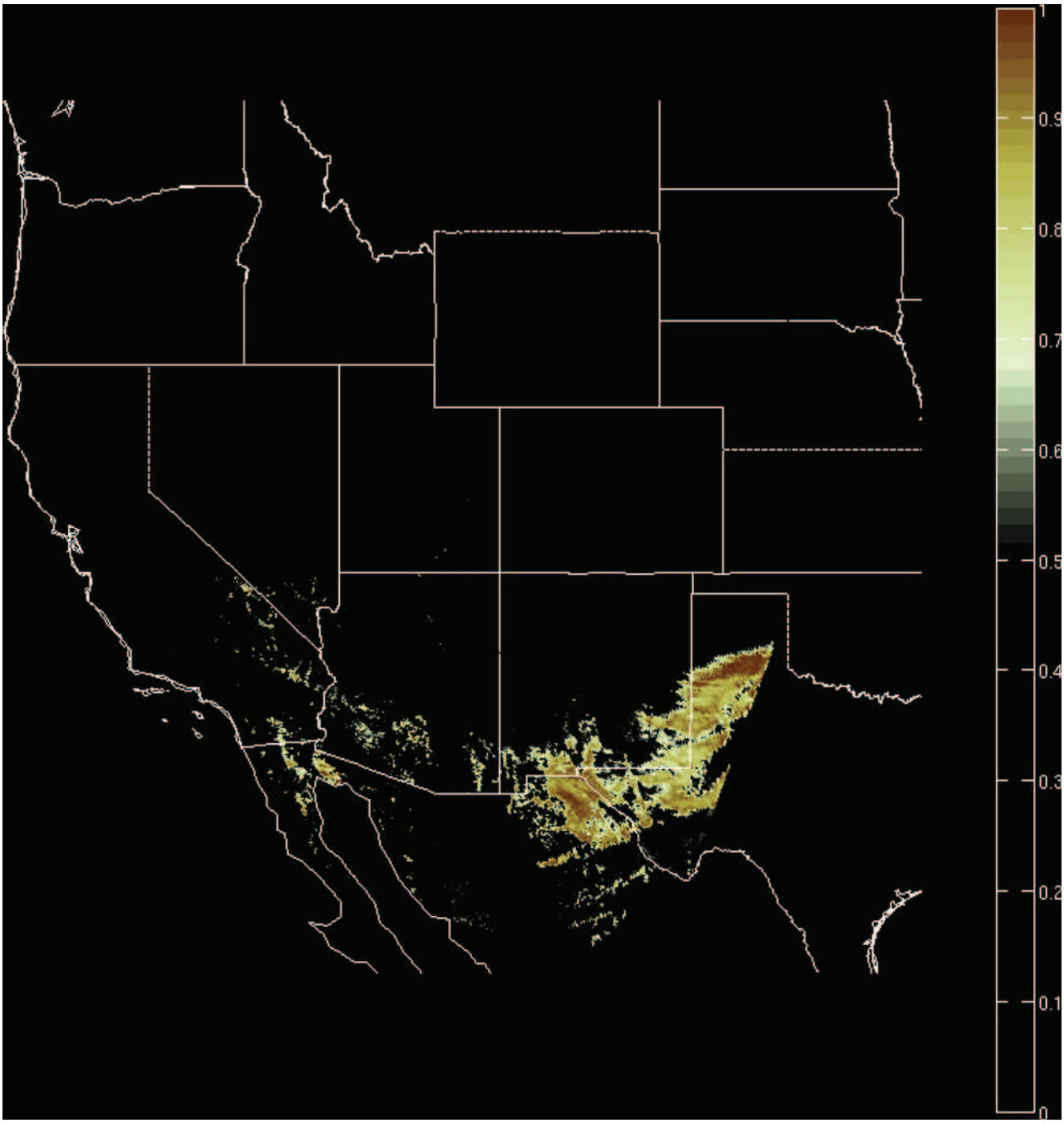

Figure 197. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

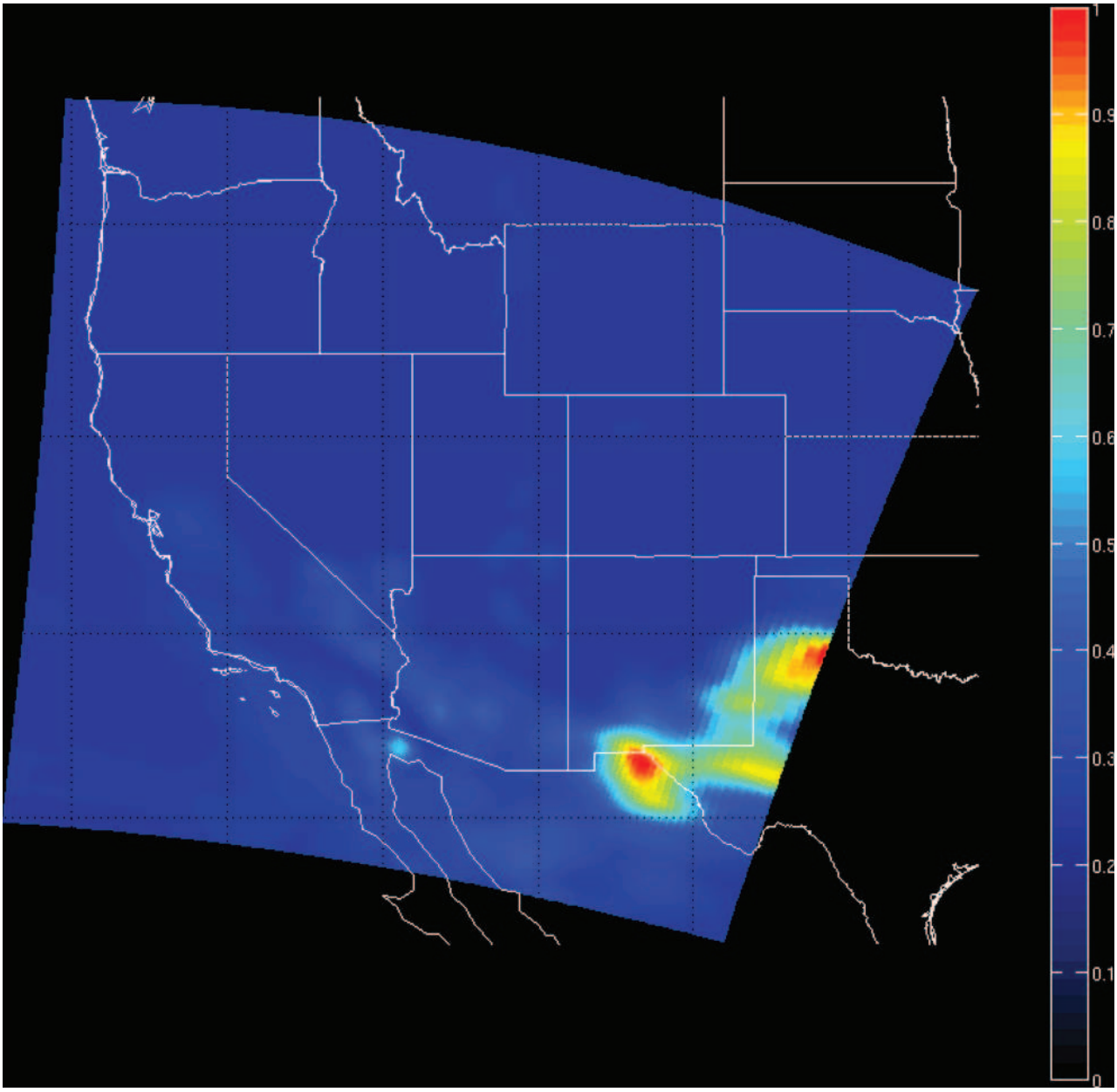

Figure 198. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

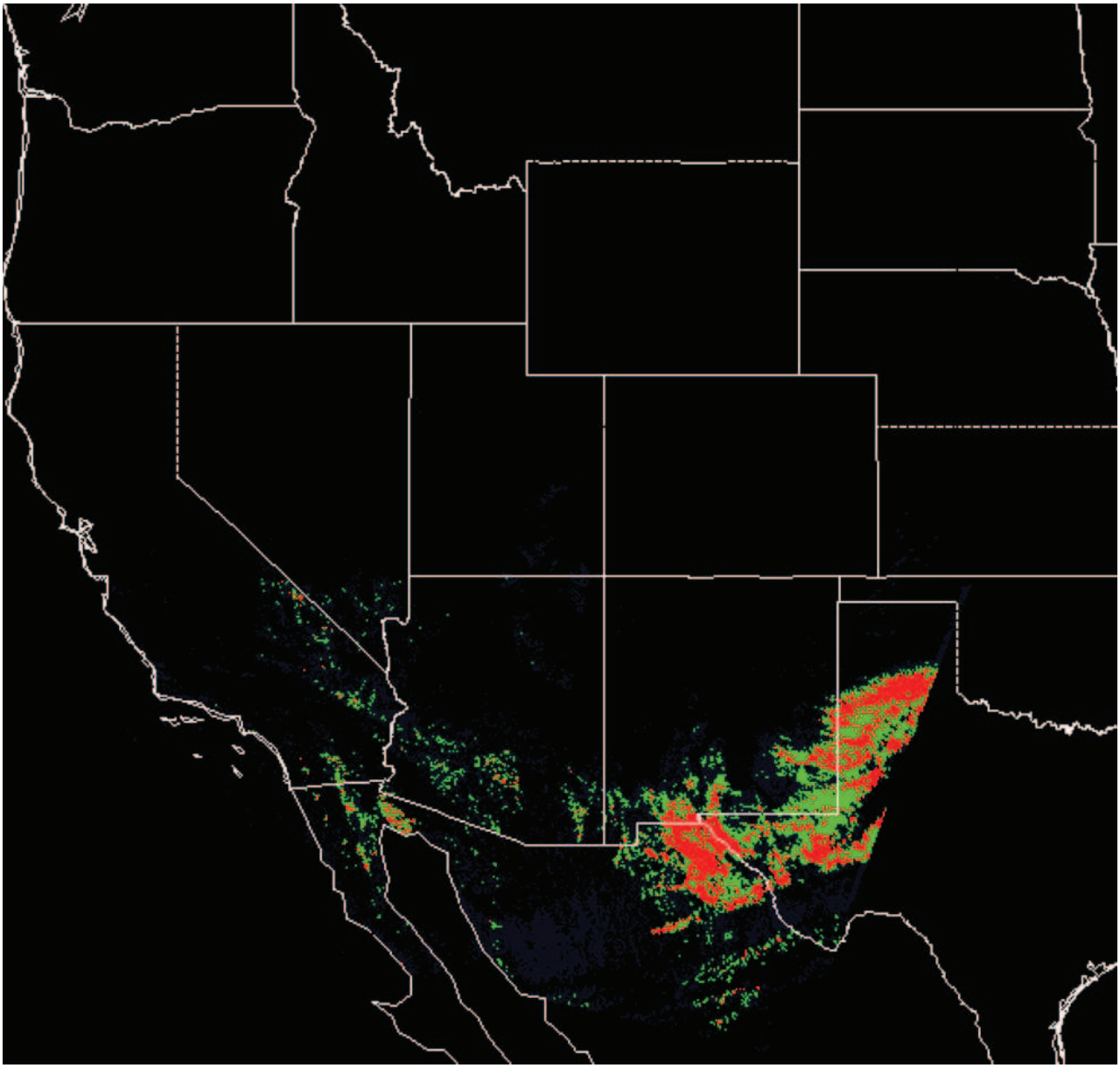

Figure 199. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-349 at 18:25 UTC.

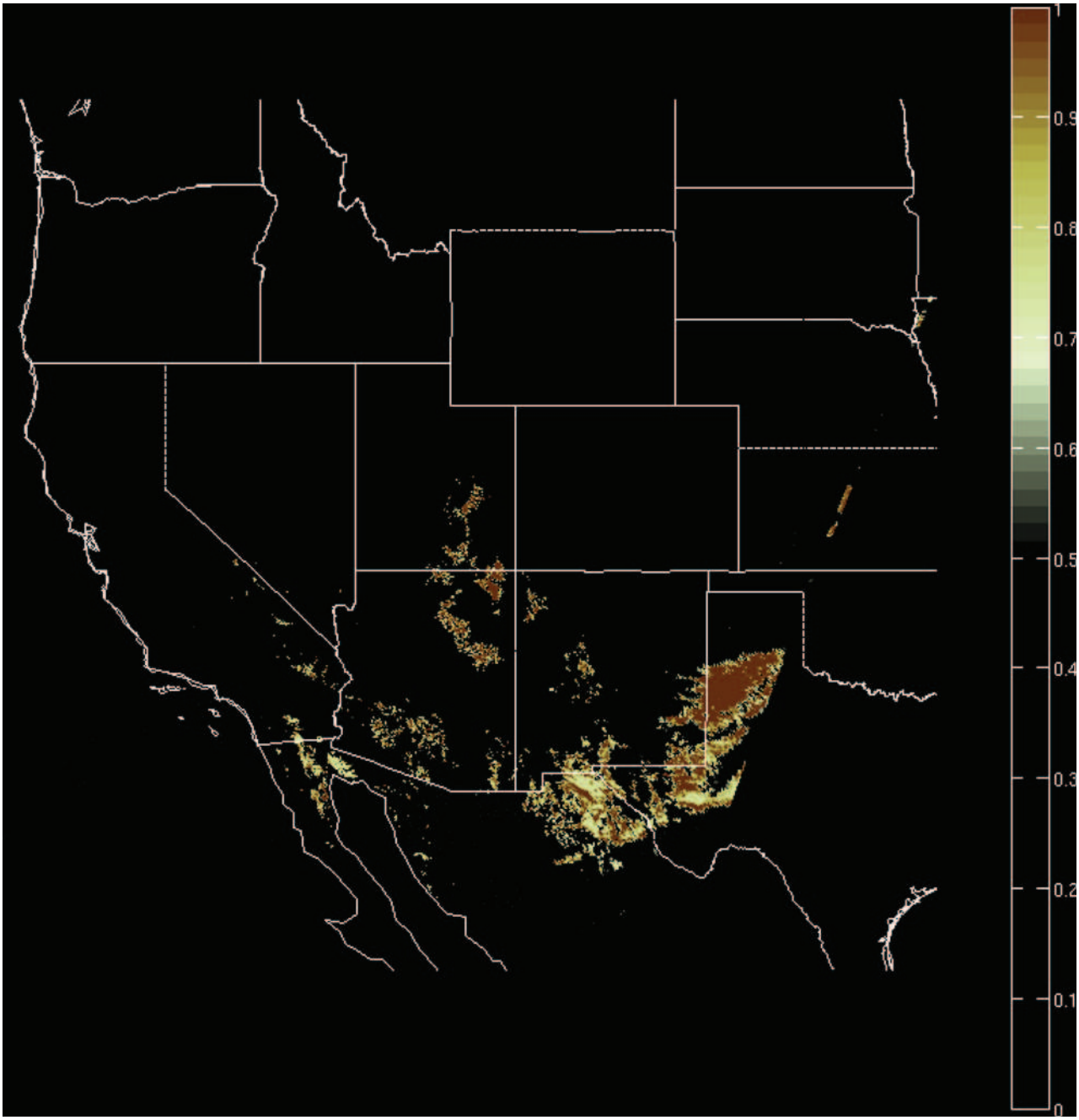

Figure 200. Result of the Probabilistic Neural Network (PNN) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

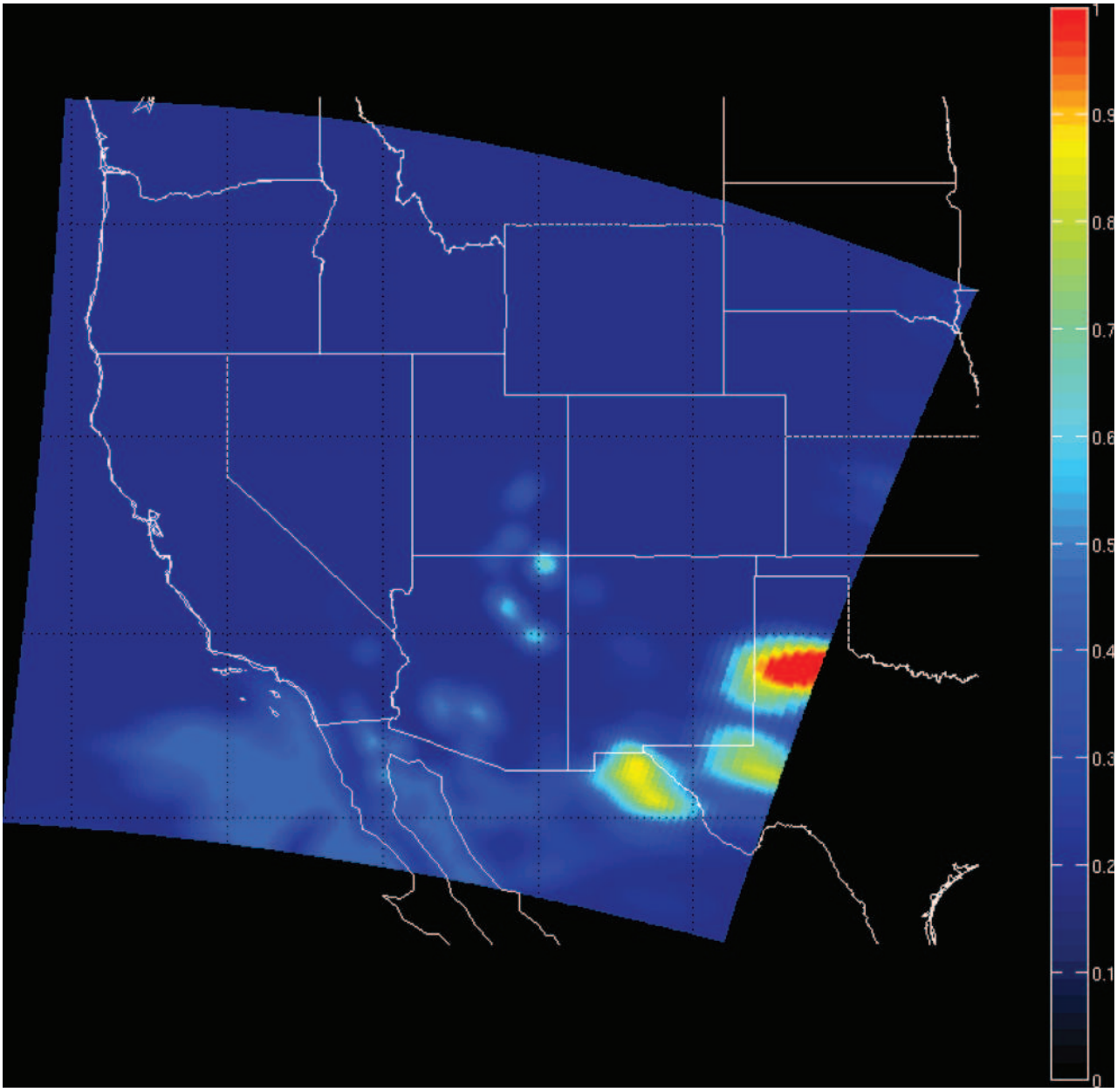

Figure 201. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-349 at 18:25 UTC. The colorbar shows the index of probability.

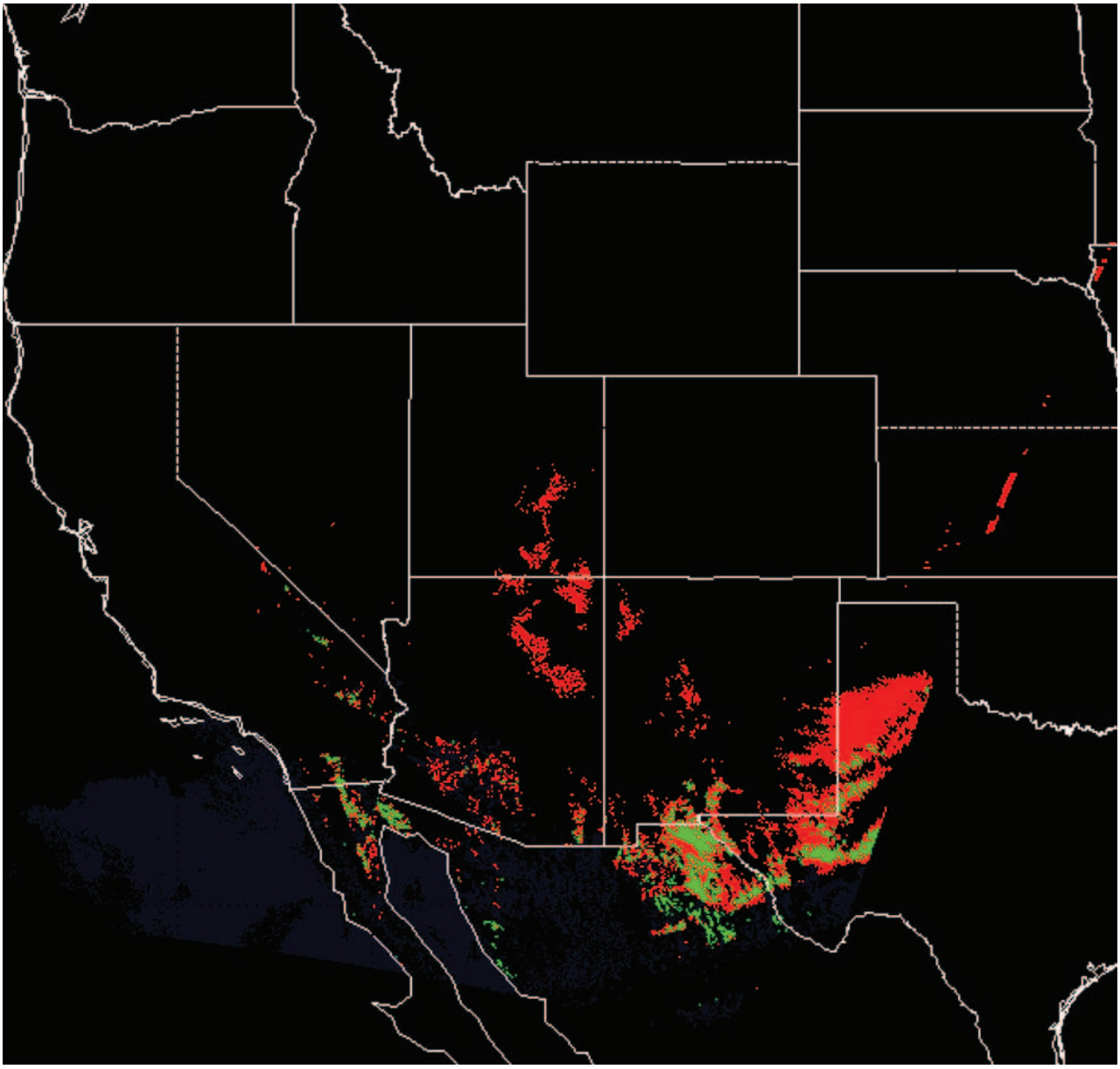

Figure 202. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-349 at 18:25 UTC.

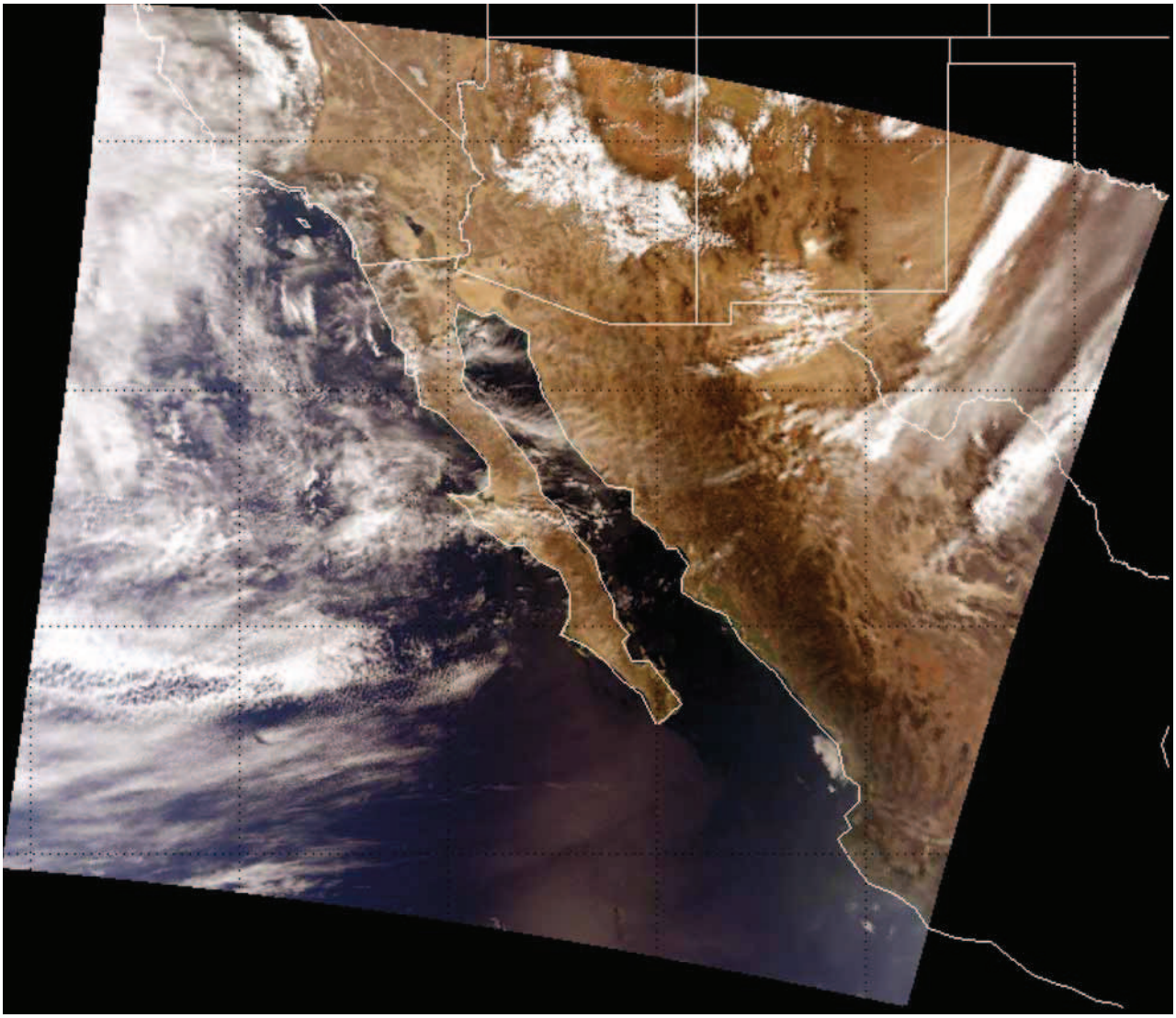

Figure 203. True color image for the event of 2004-050 at 18:15 UTC.

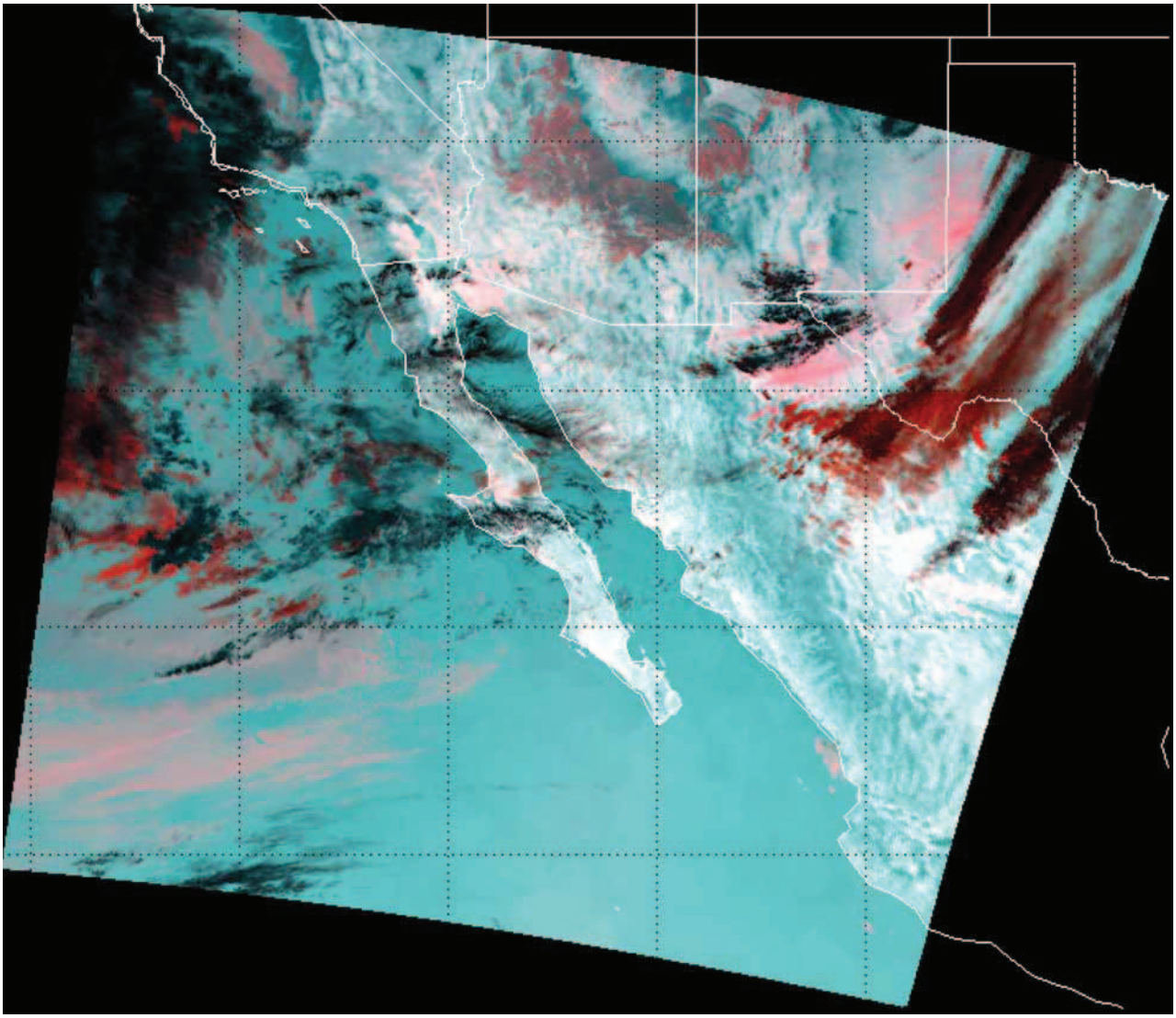

Figure 204. False color image for the event of 2004-050 at 18:15 UTC.

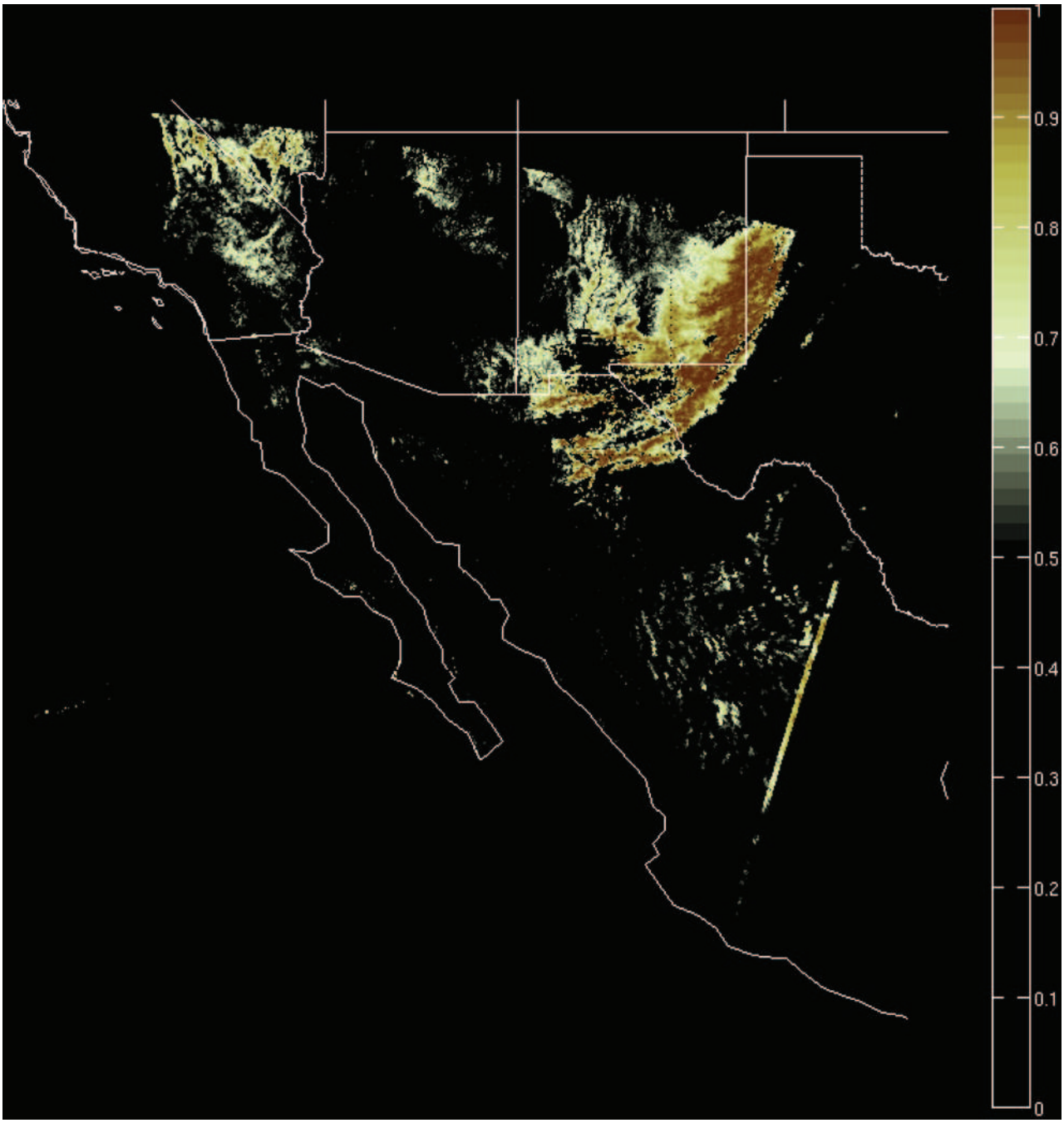

Figure 205. Result of the probabilistic display model (PD) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

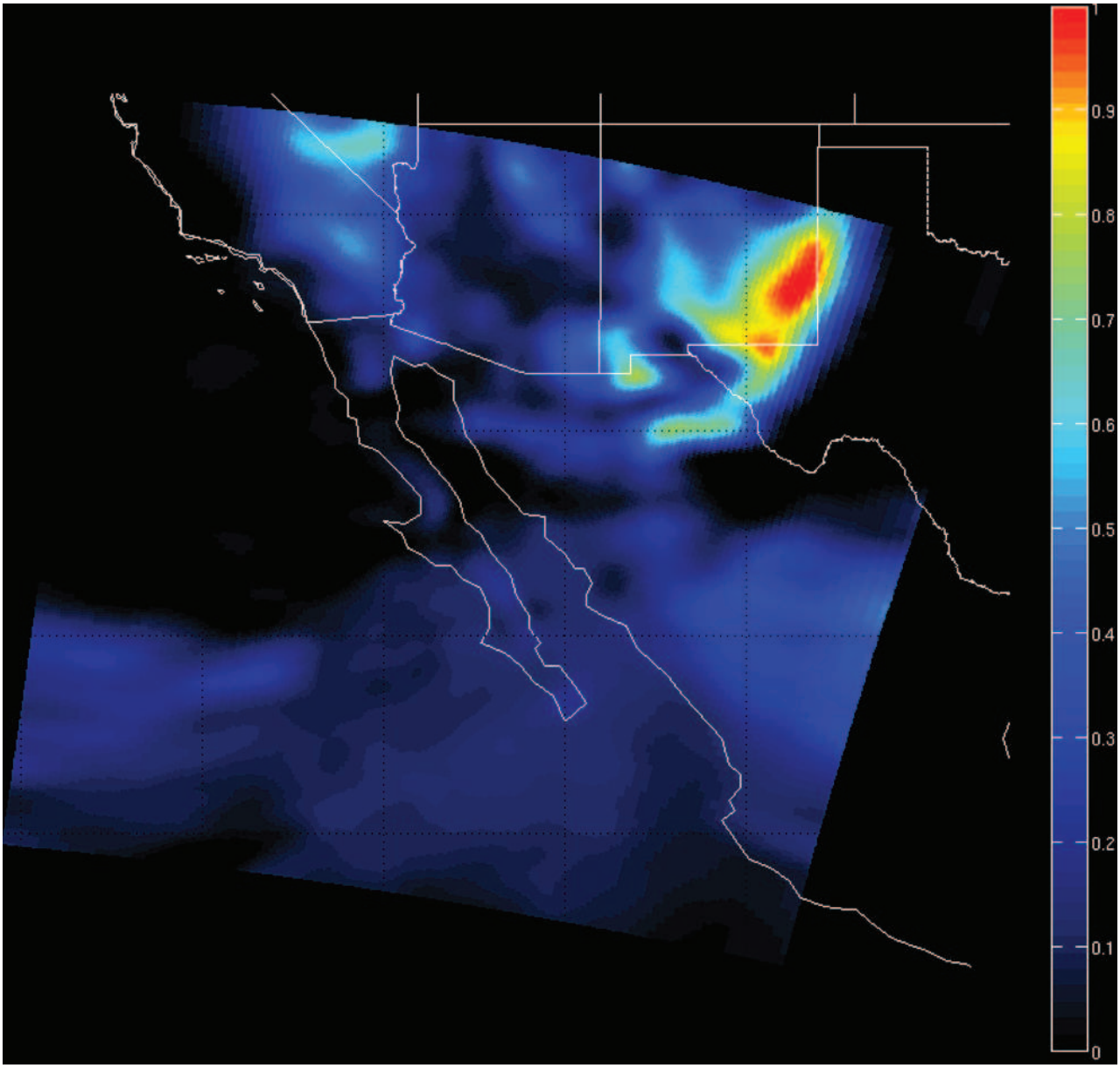

Figure 206. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

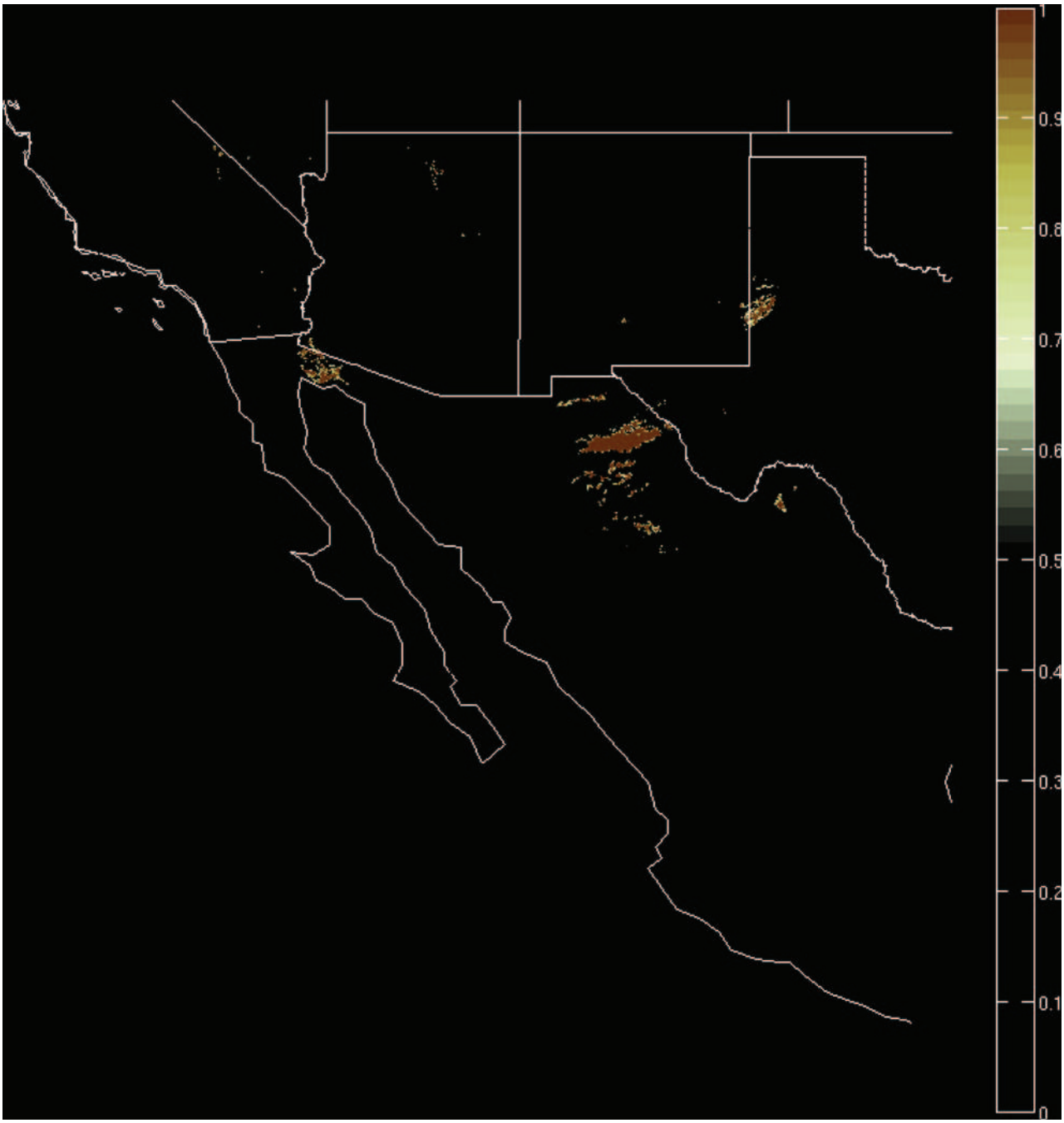

Figure 207. Result of the Maximum Likelihood model (MLE) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

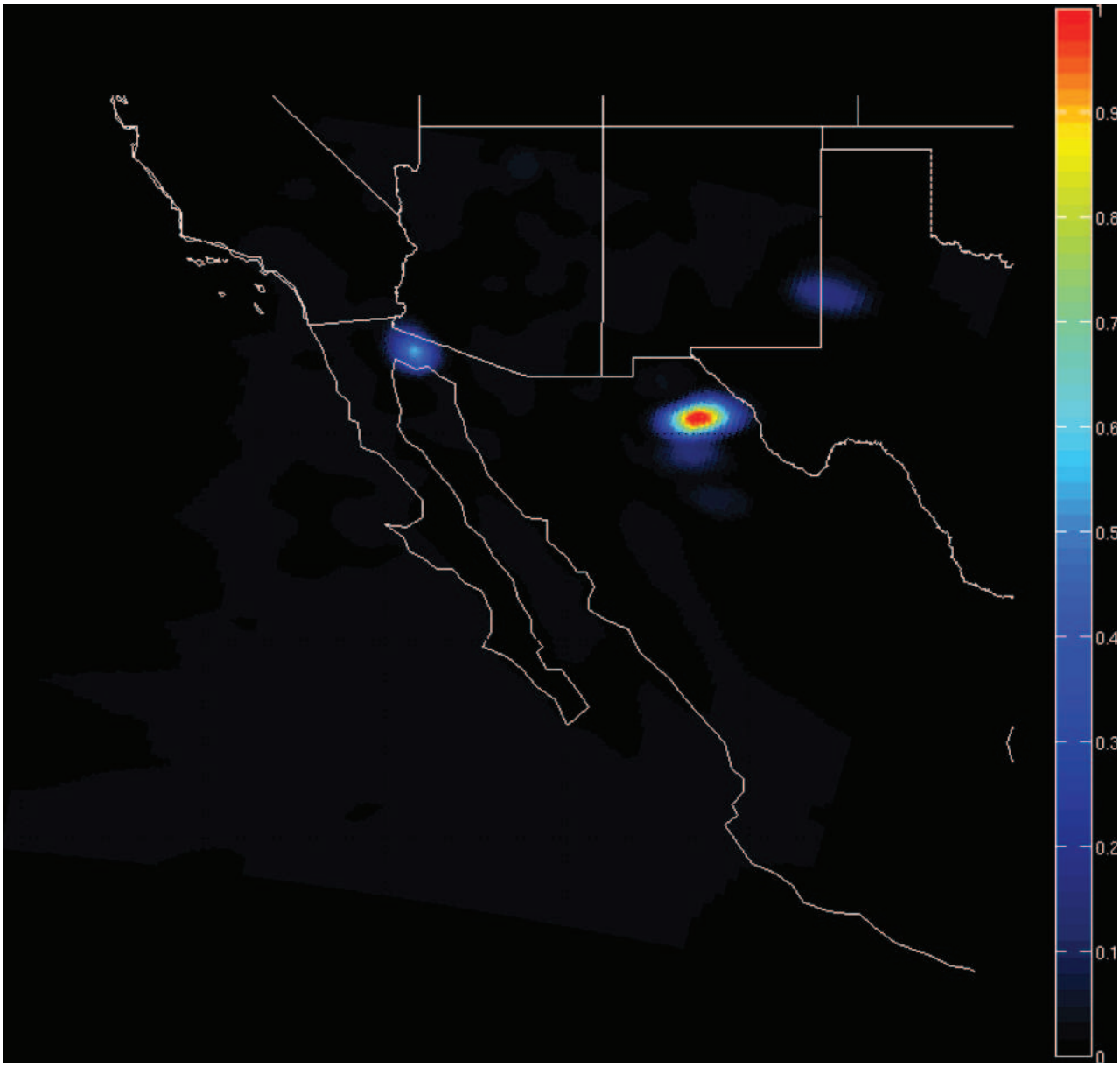

Figure 208. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

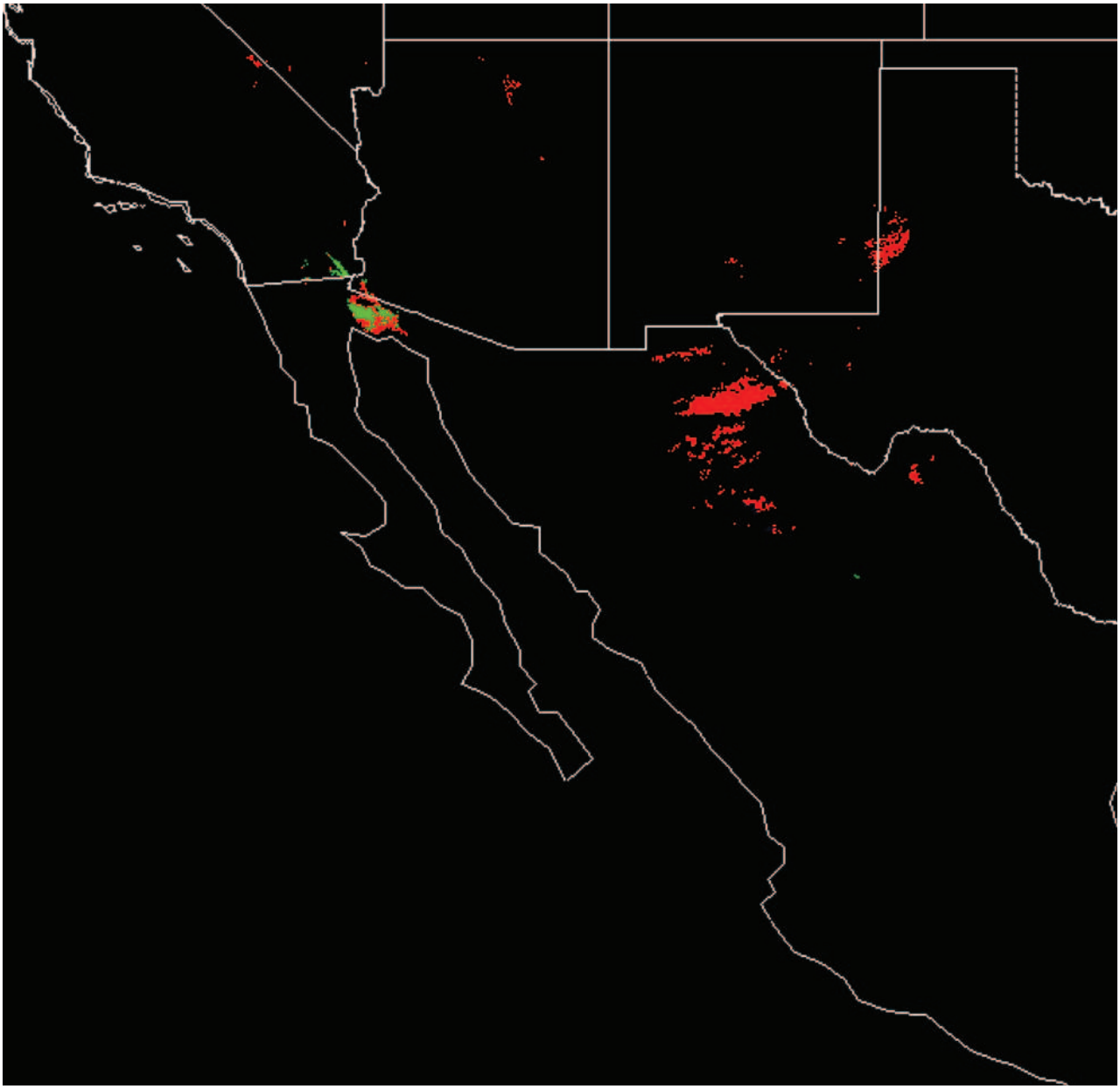

Figure 209. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2004-050 at 18:15 UTC.

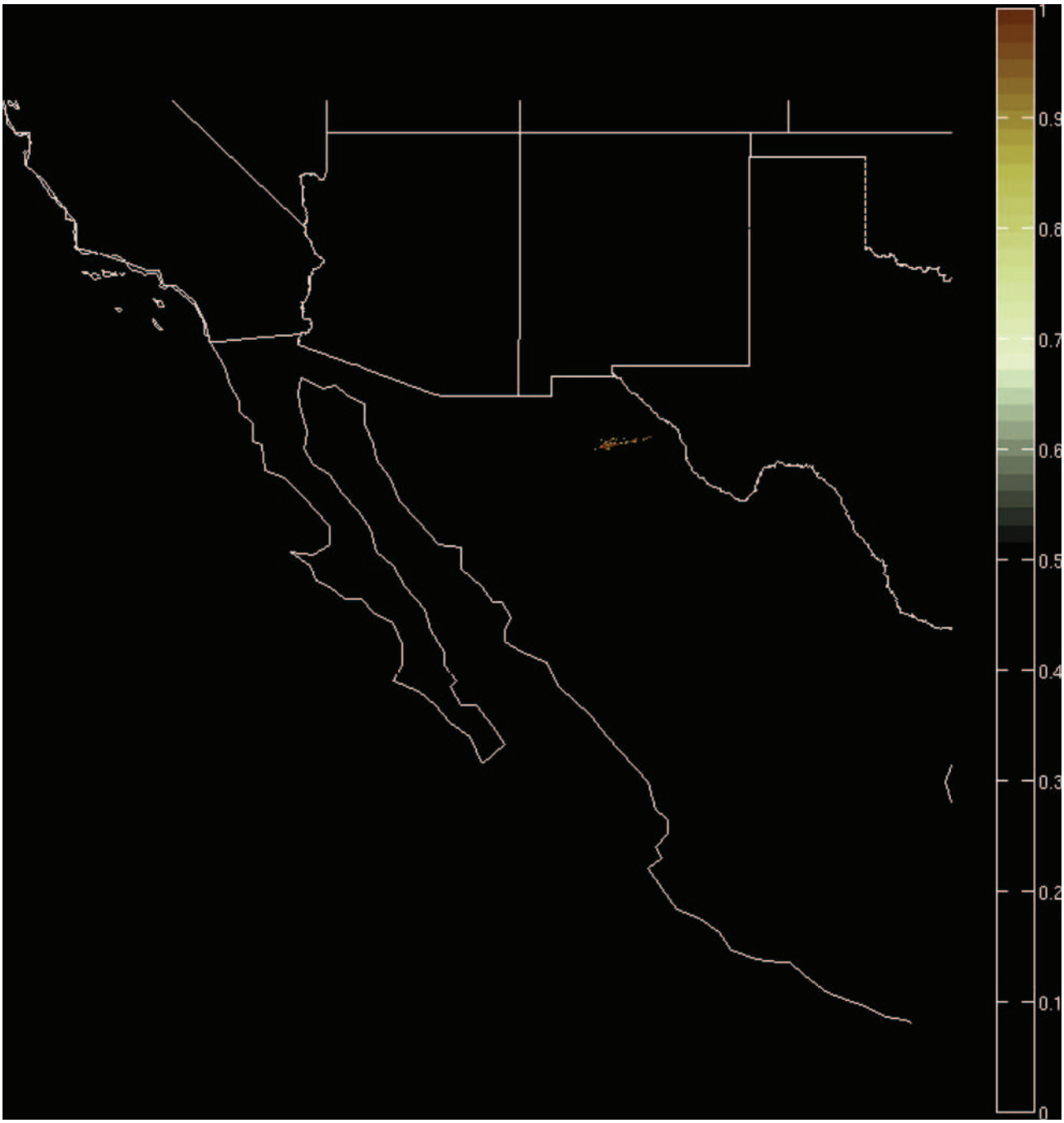

Figure 210. Result of the Maximum a posteriori model (MAP) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

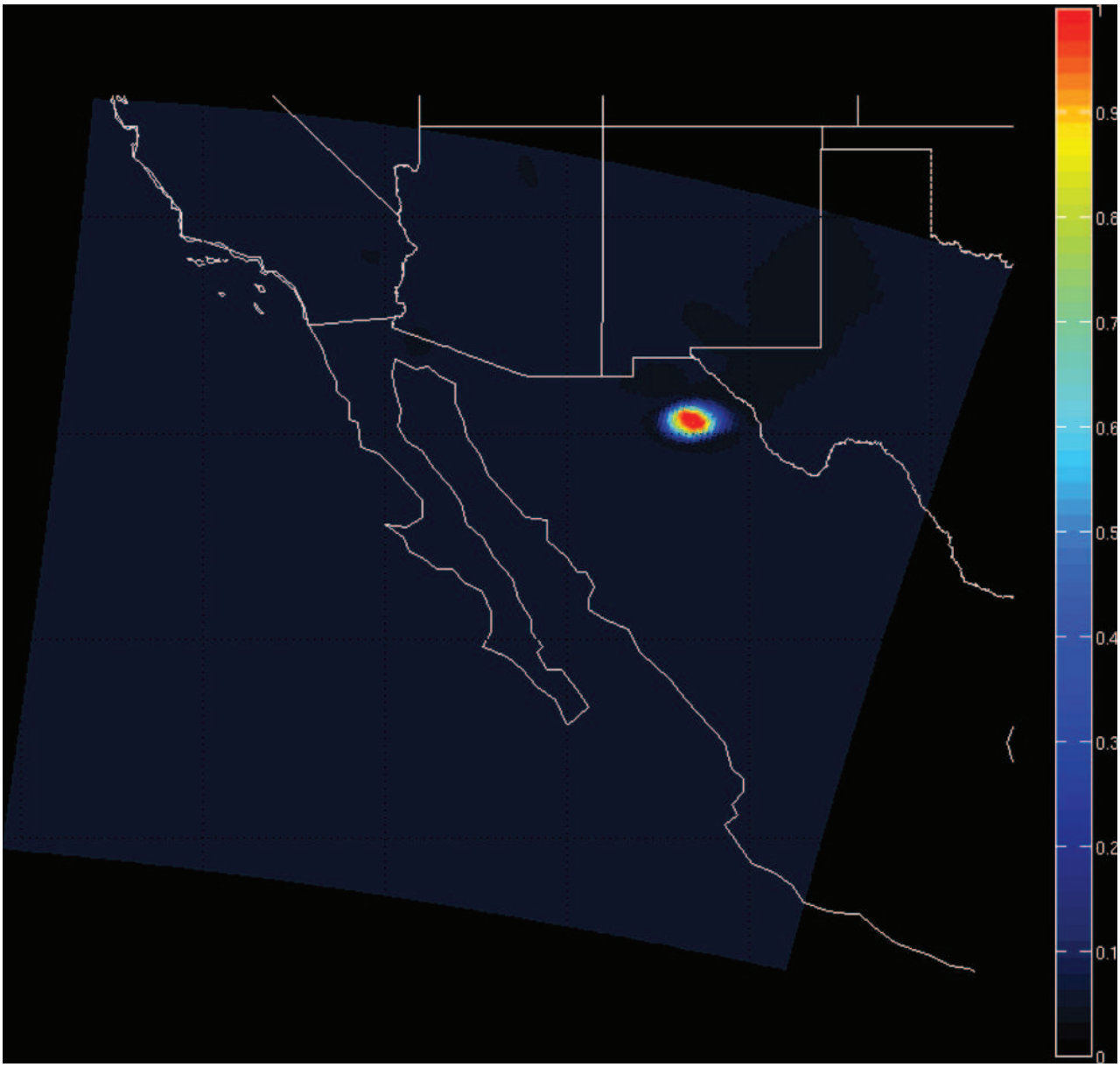

Figure 211. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

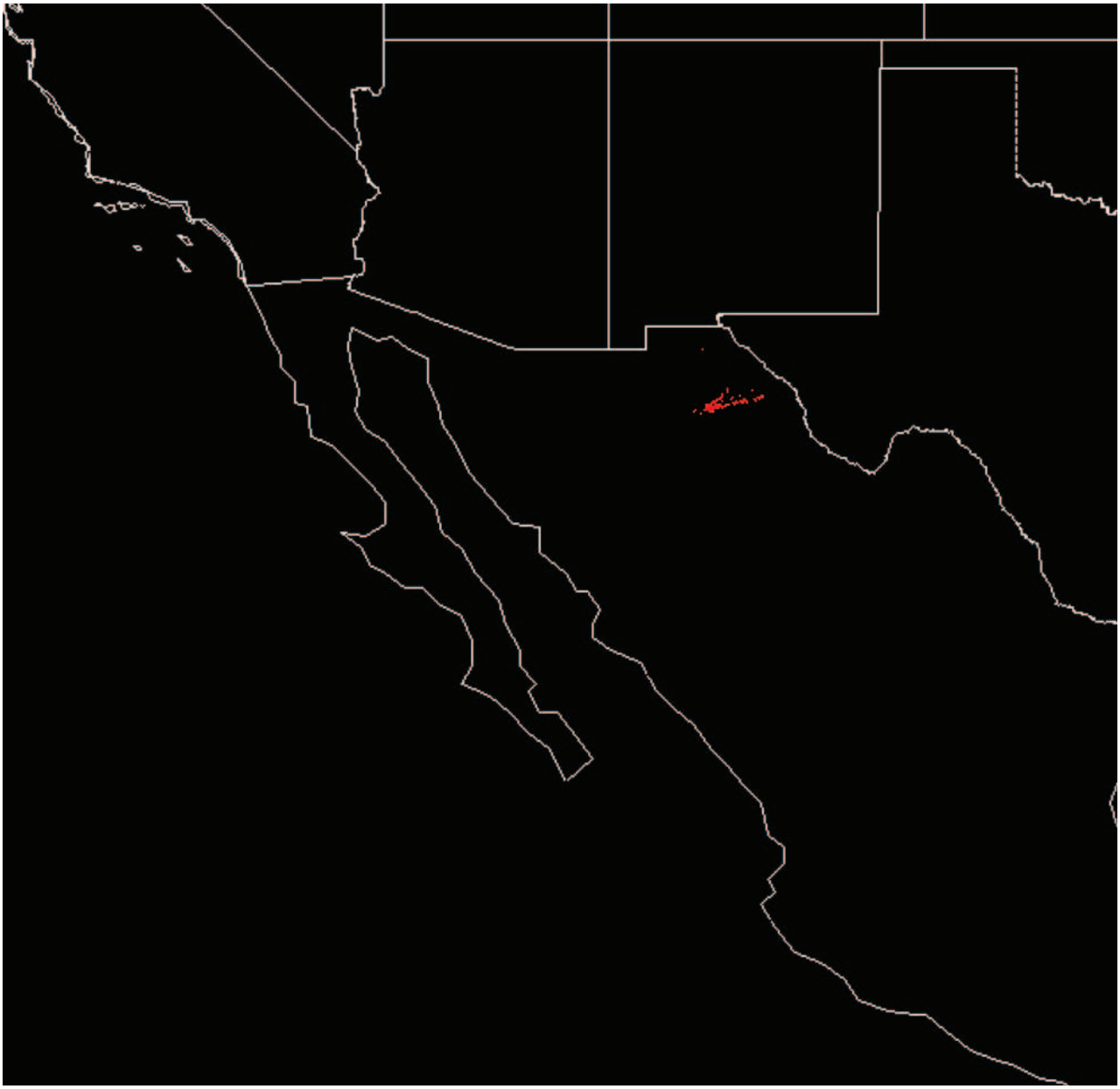

Figure 212. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2004-050 at 18:15 UTC.

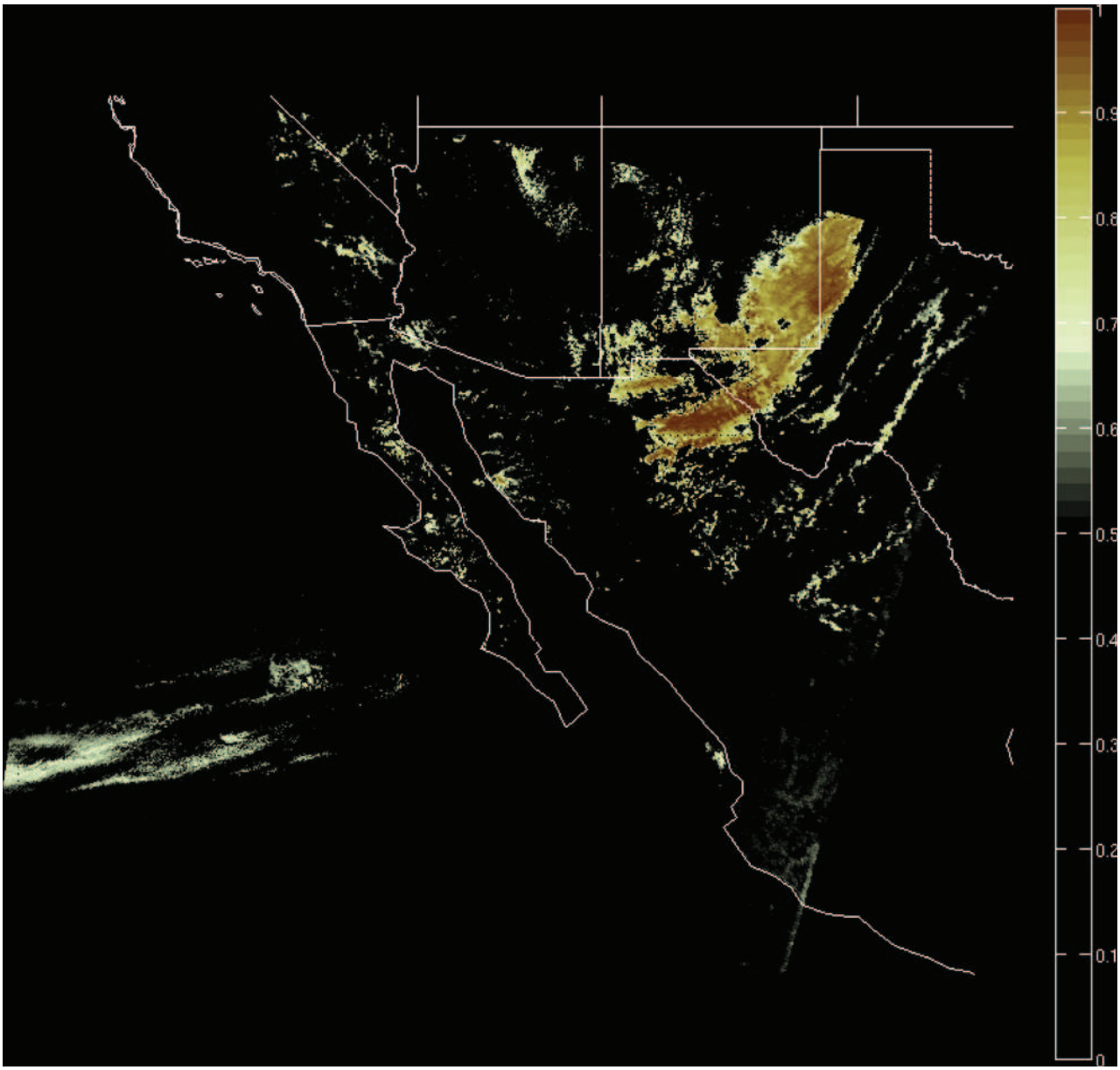

Figure 213. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

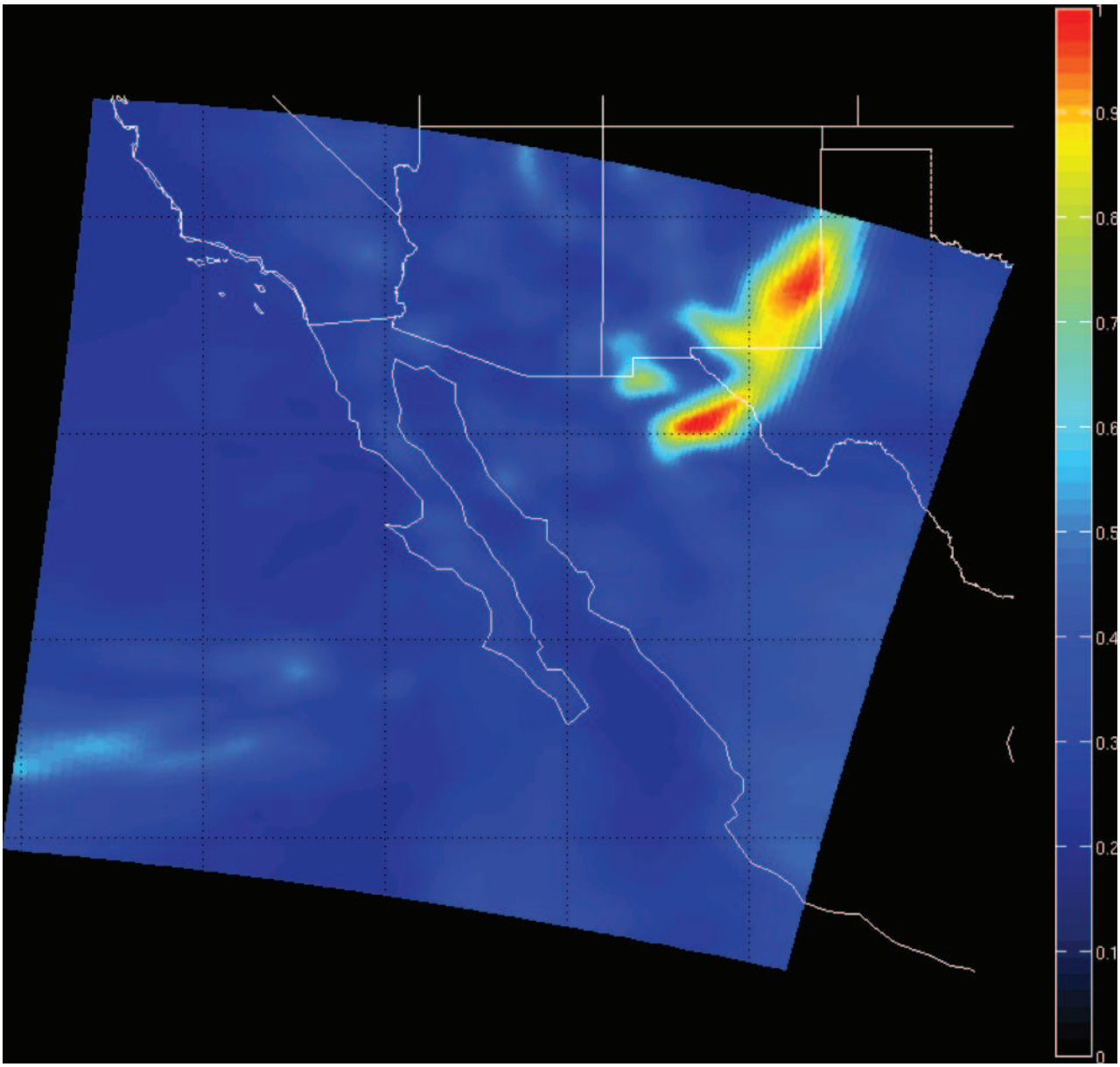

Figure 214. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

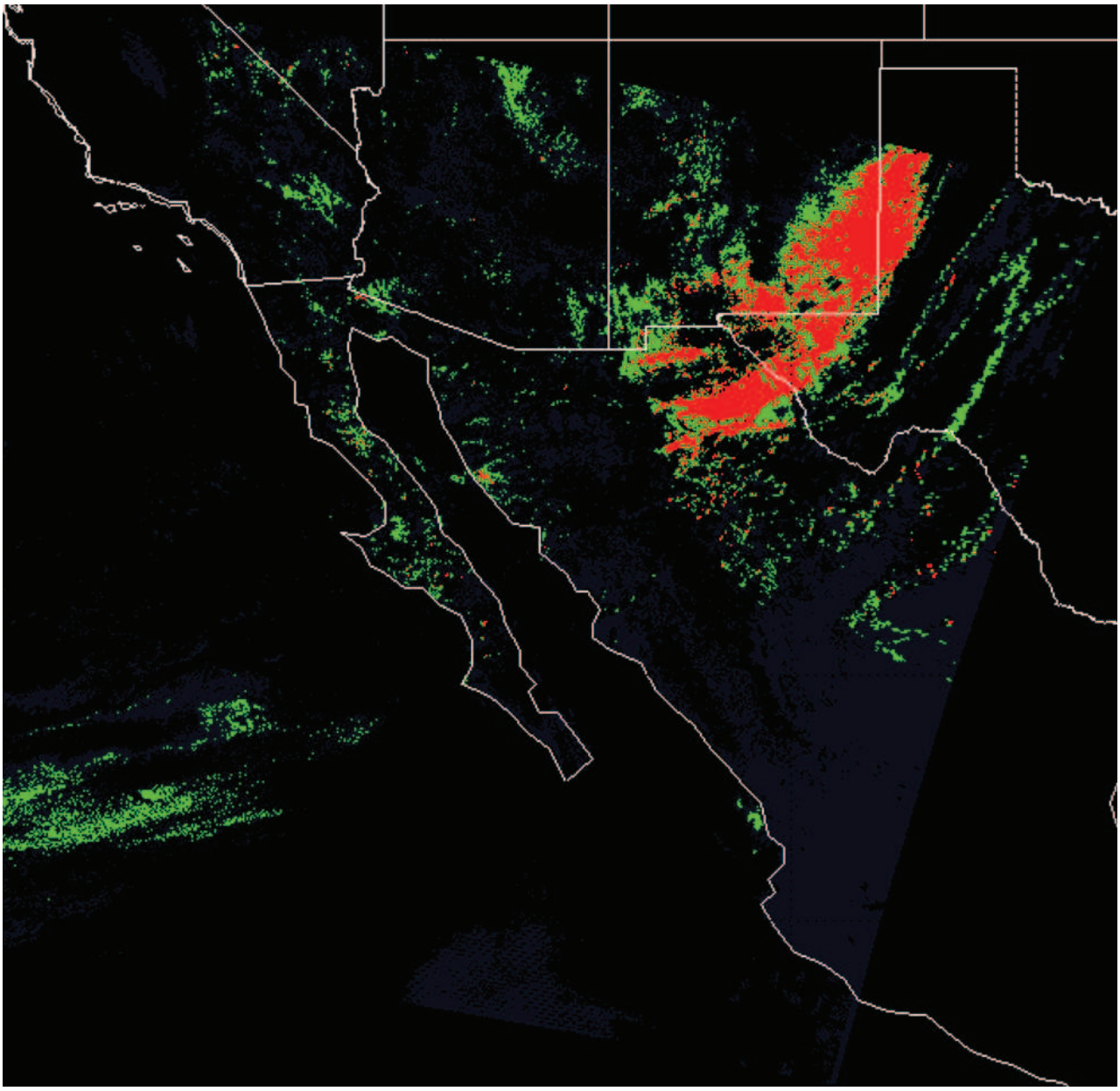

Figure 215. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2004-050 at 18:15 UTC.

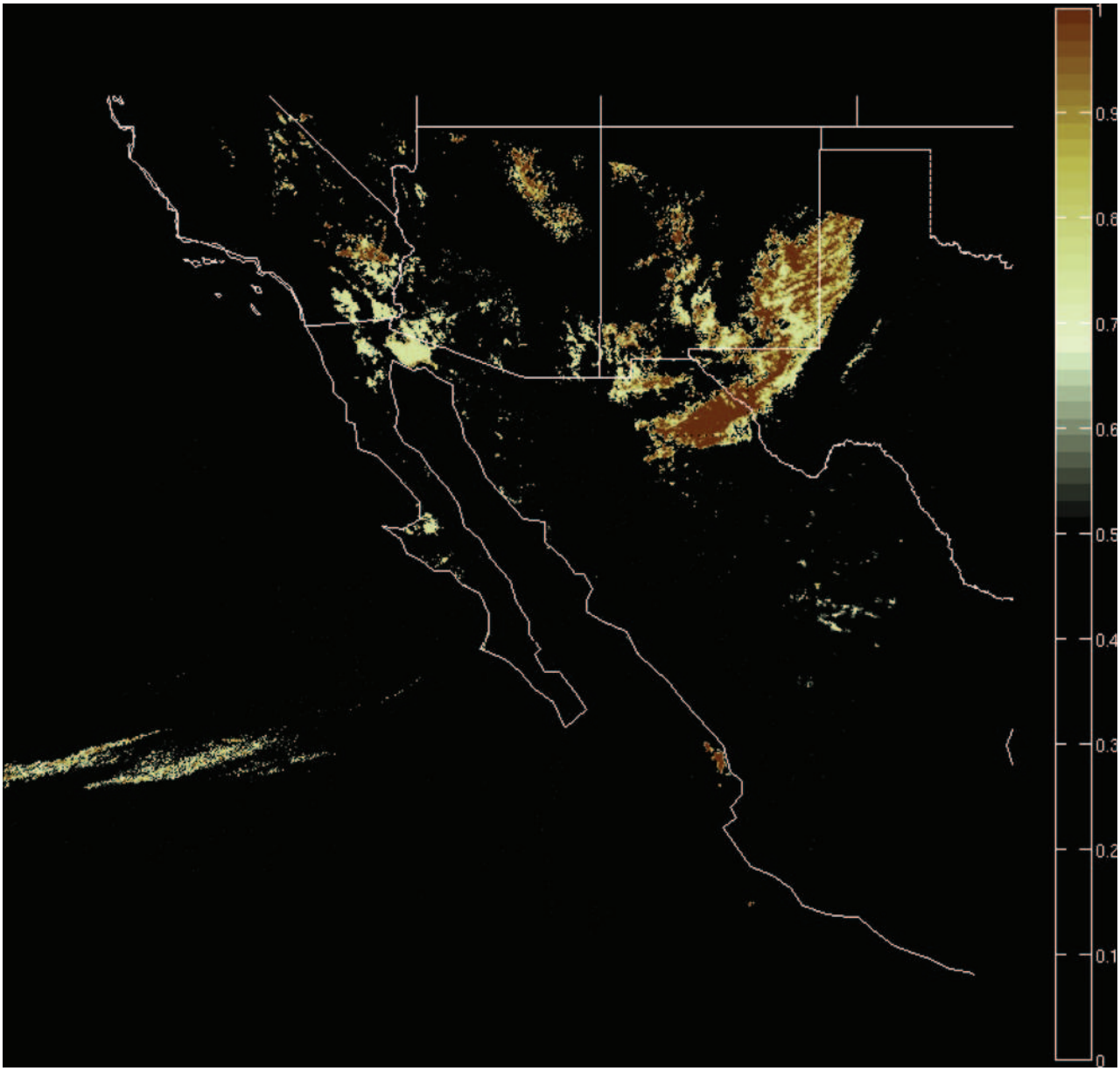

Figure 216. Result of the Probabilistic Neural Network (PNN) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

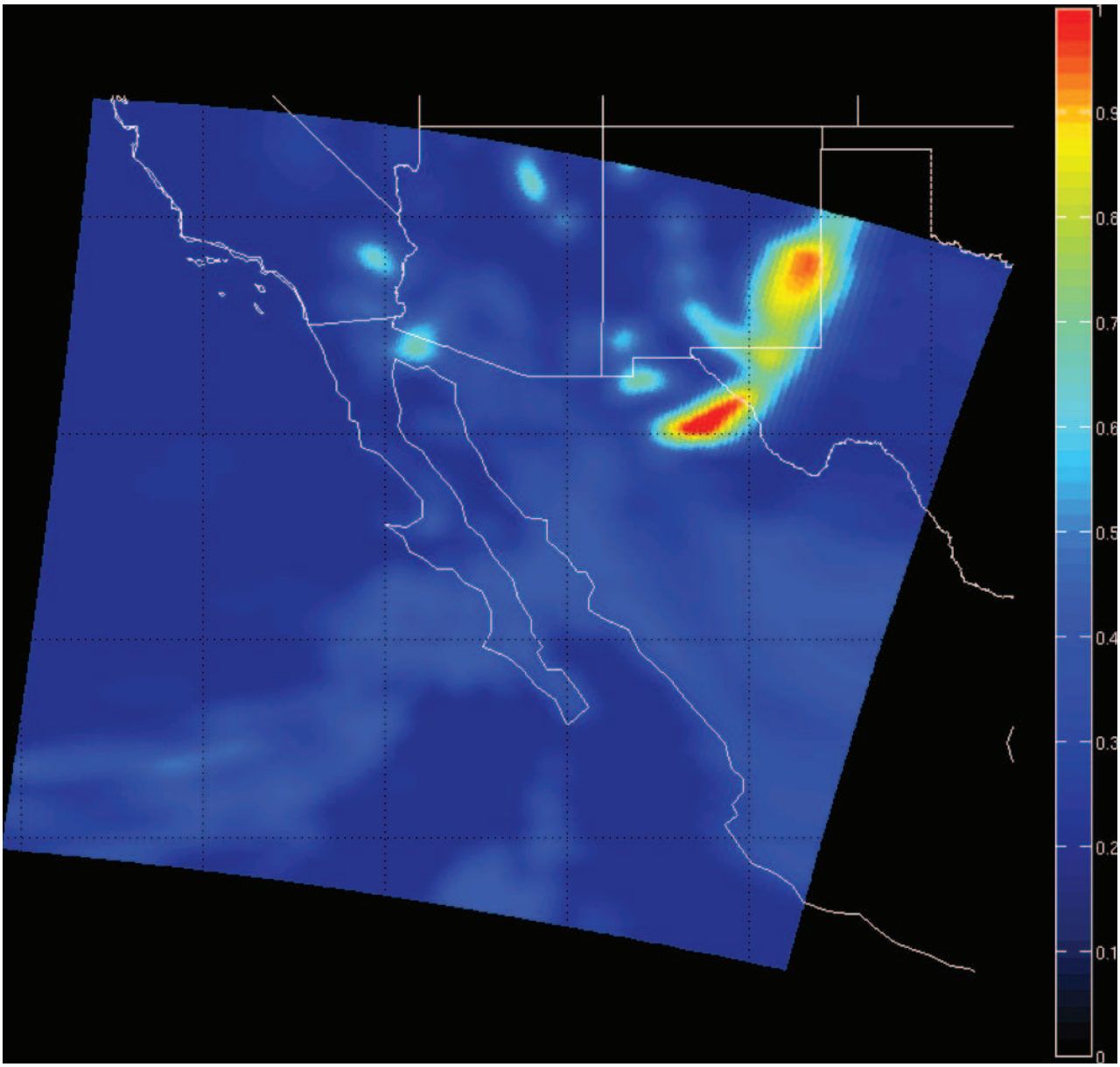

Figure 217. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2004-050 at 18:15 UTC. The colorbar shows the index of probability.

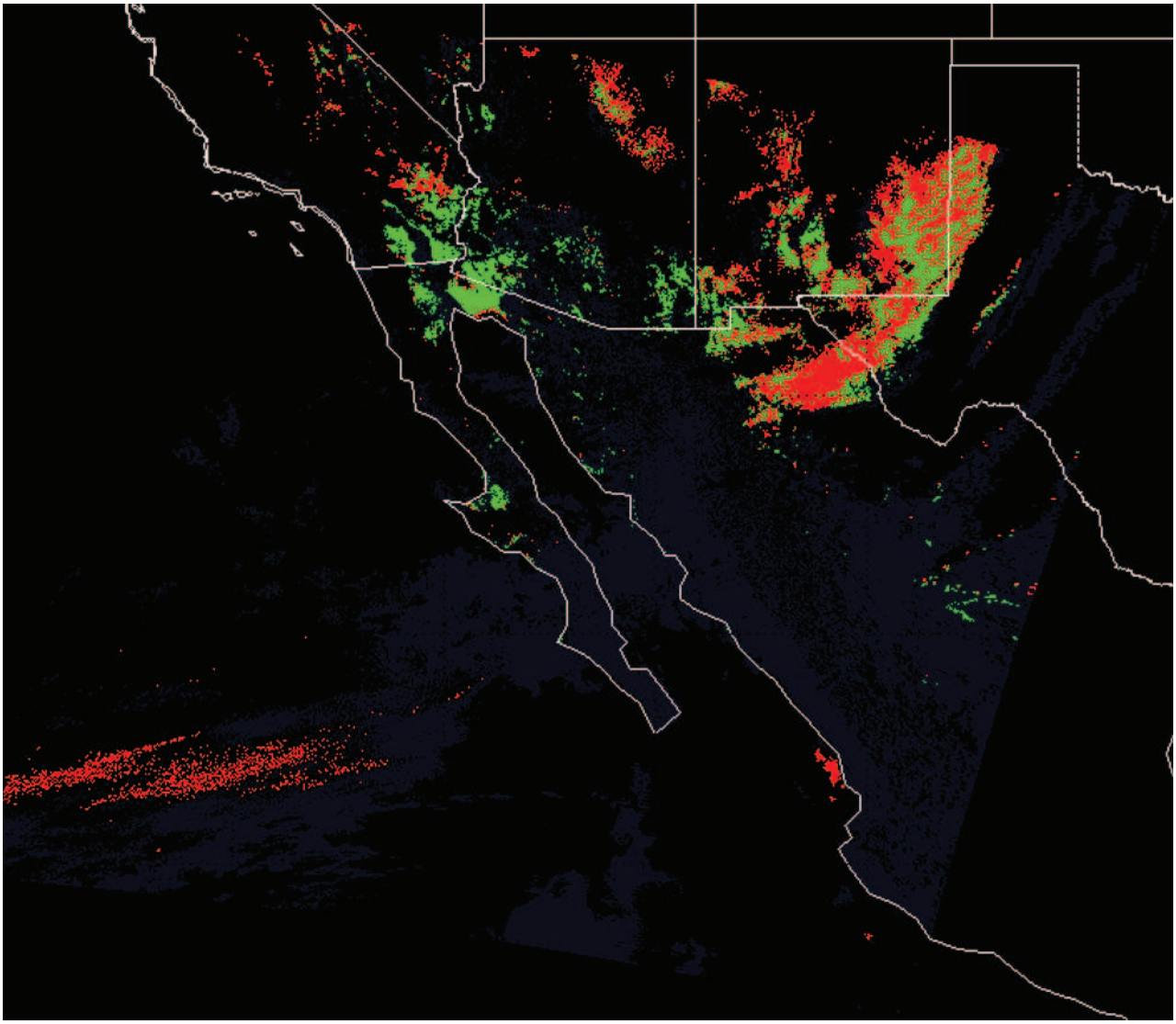

Figure 218. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2004-050 at 18:15 UTC.

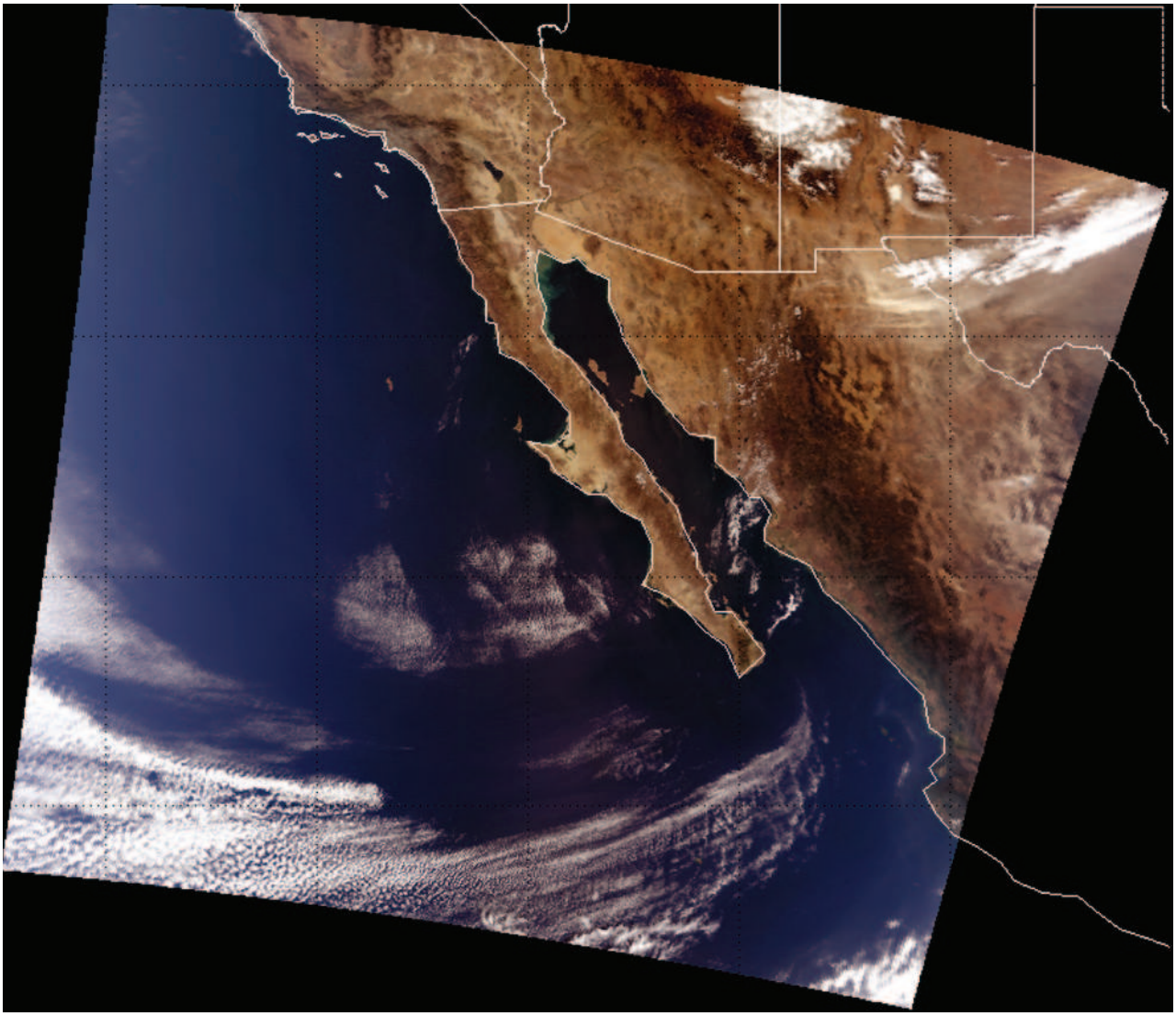

Figure 219. True color image for the event of 2005-331 at 18:20 UTC.

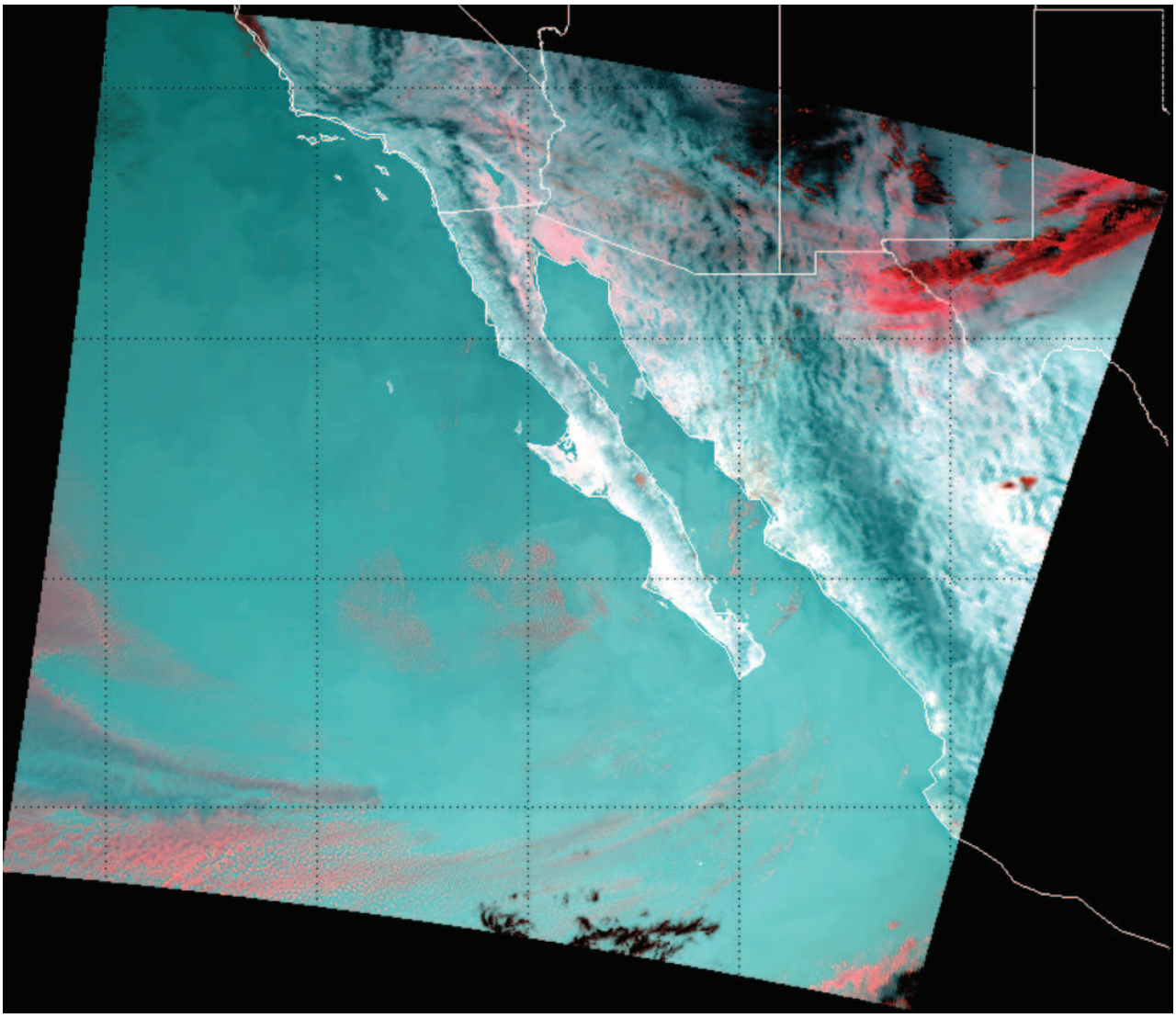

Figure 220. False color image for the event of 2005-331 at 18:20 UTC.

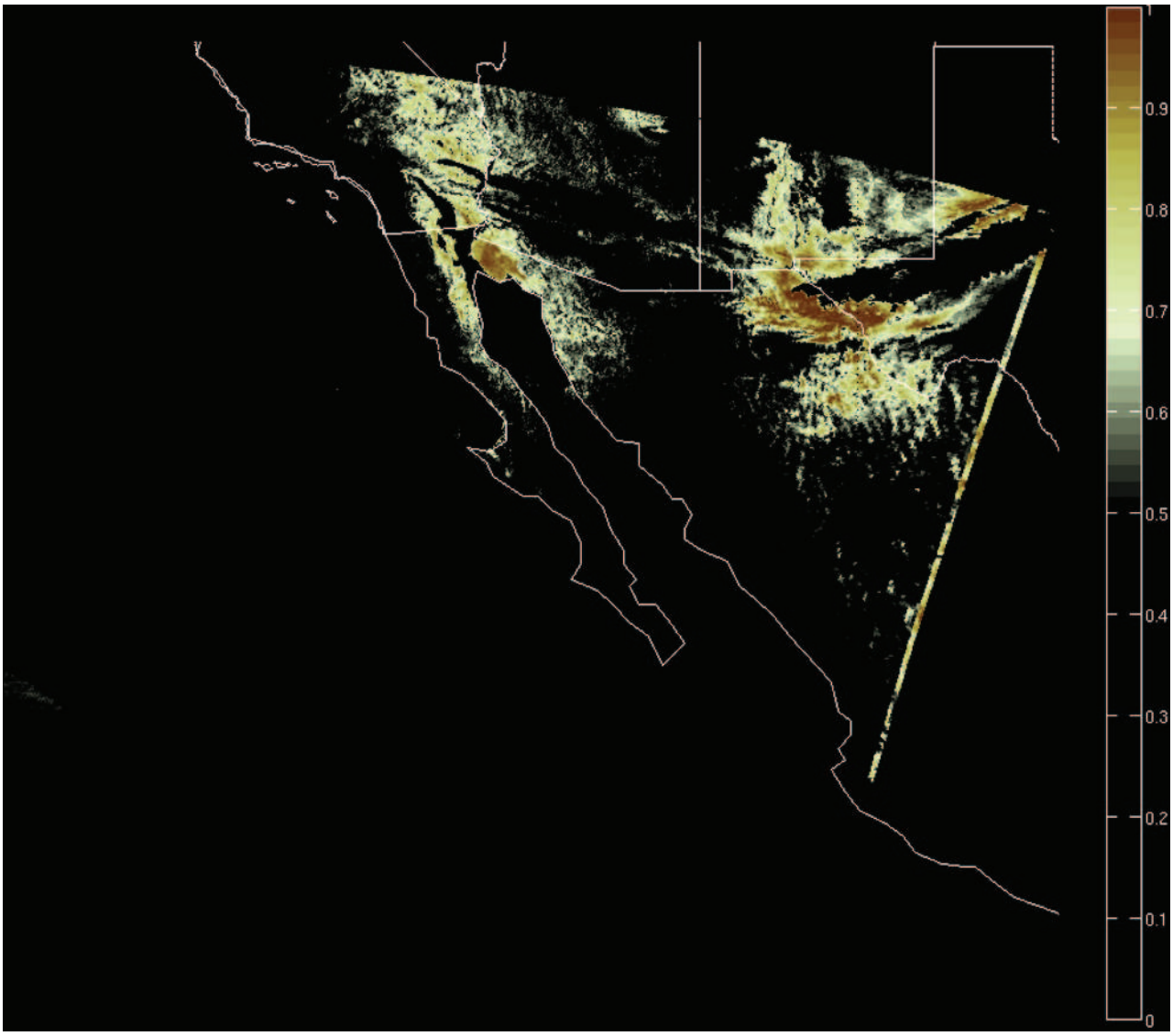

Figure 221. Result of the probabilistic display model (PD) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

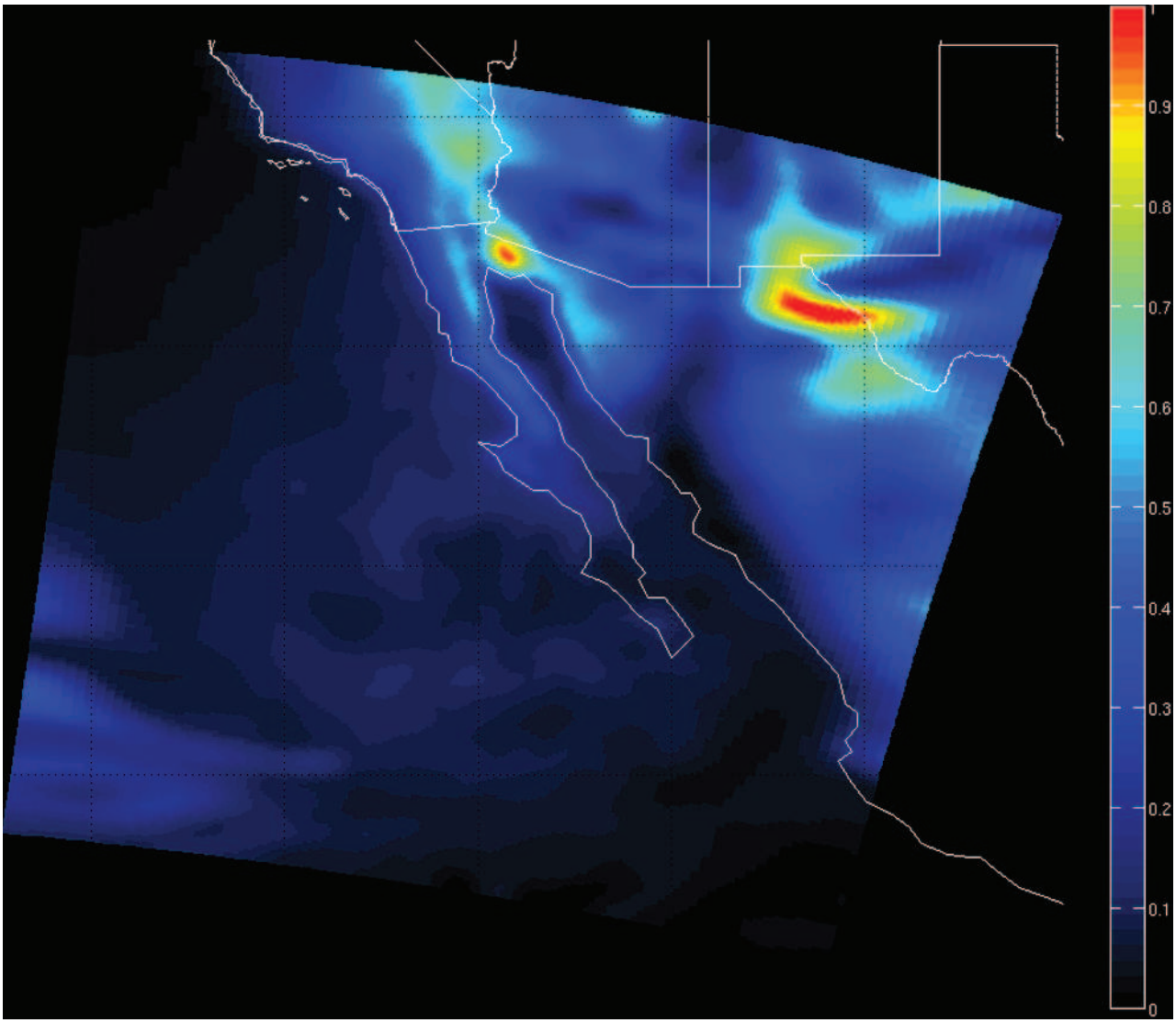

Figure 222. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

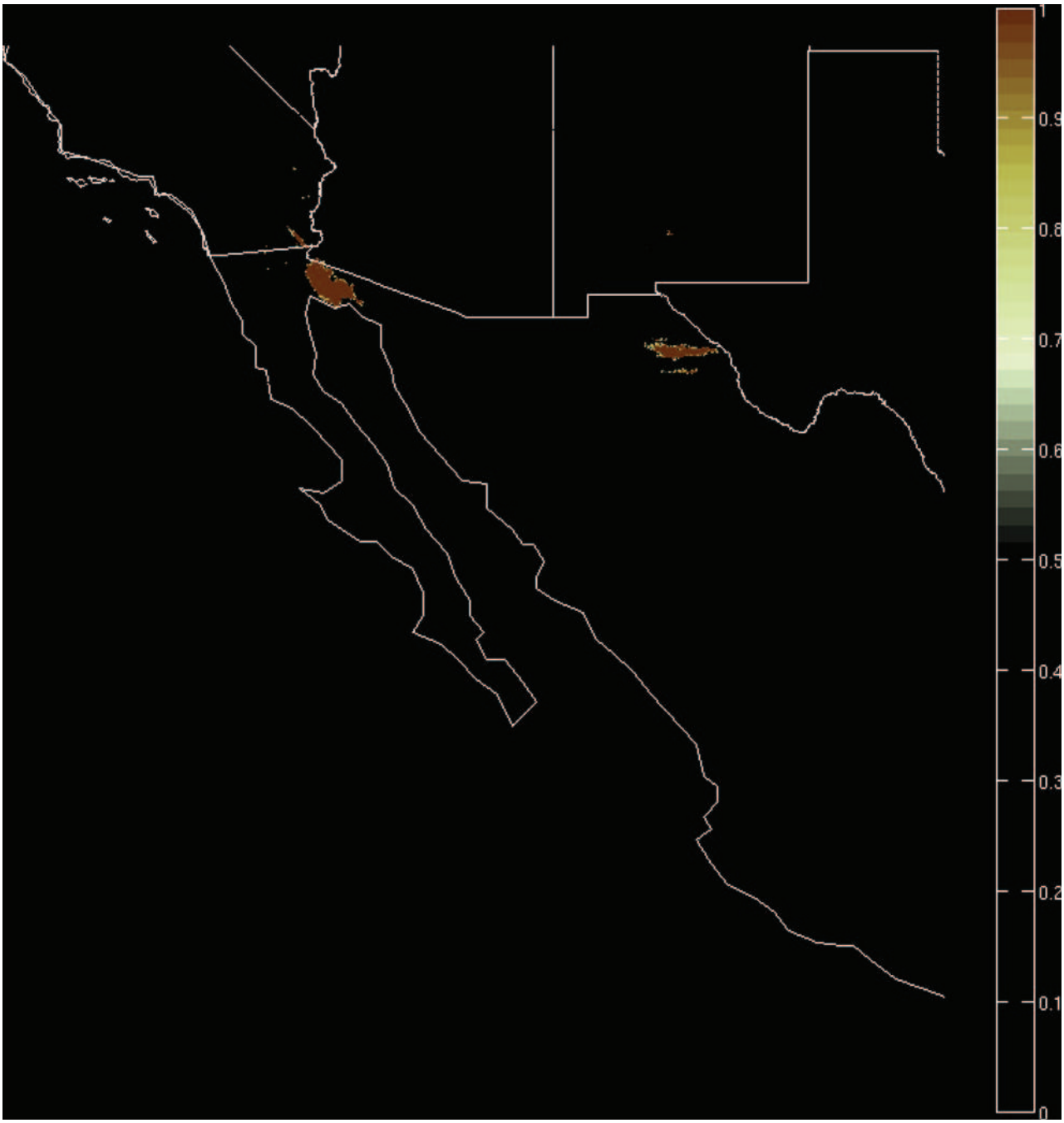

Figure 223. Result of the Maximum Likelihood model (MLE) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

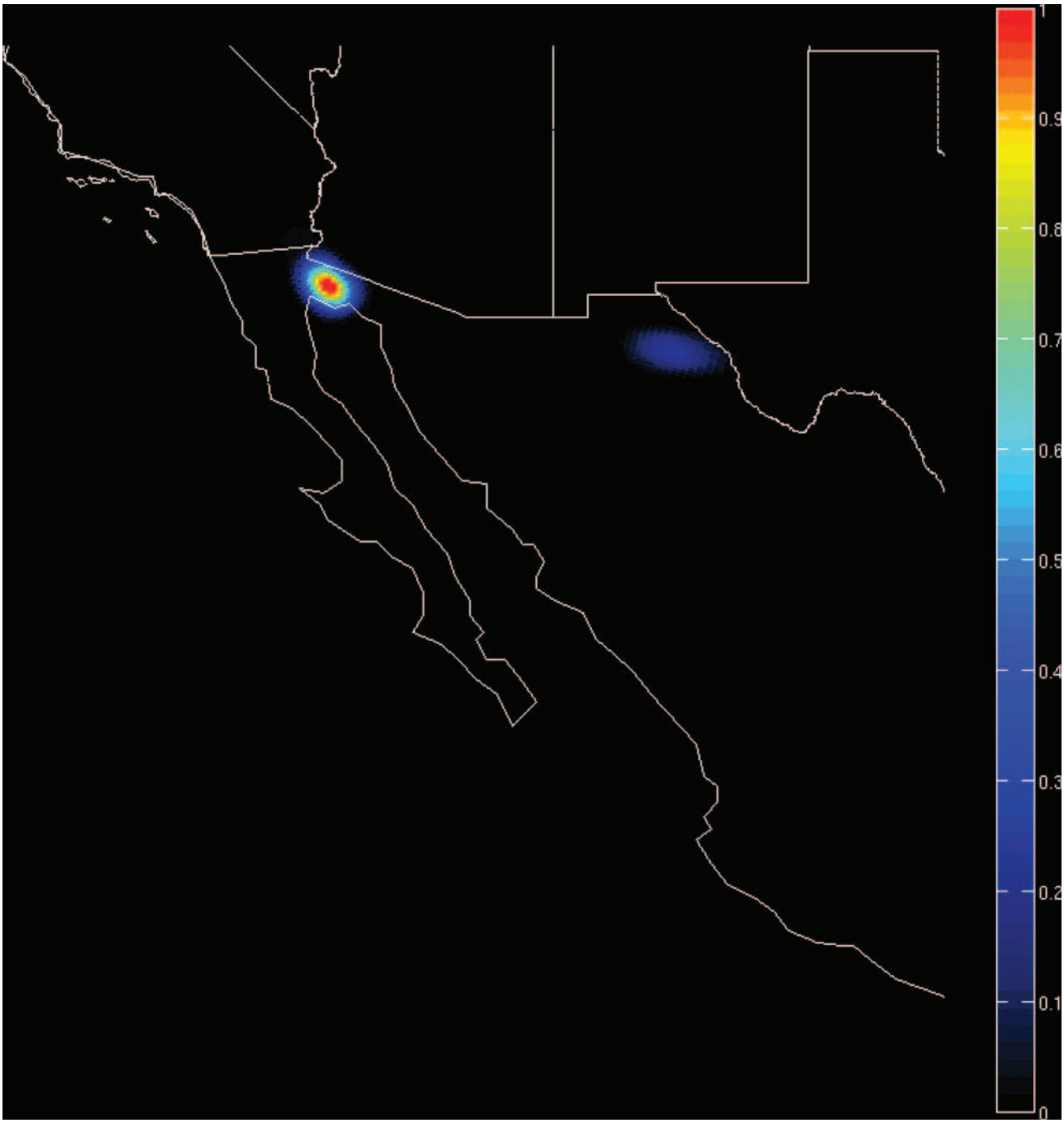

Figure 224. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

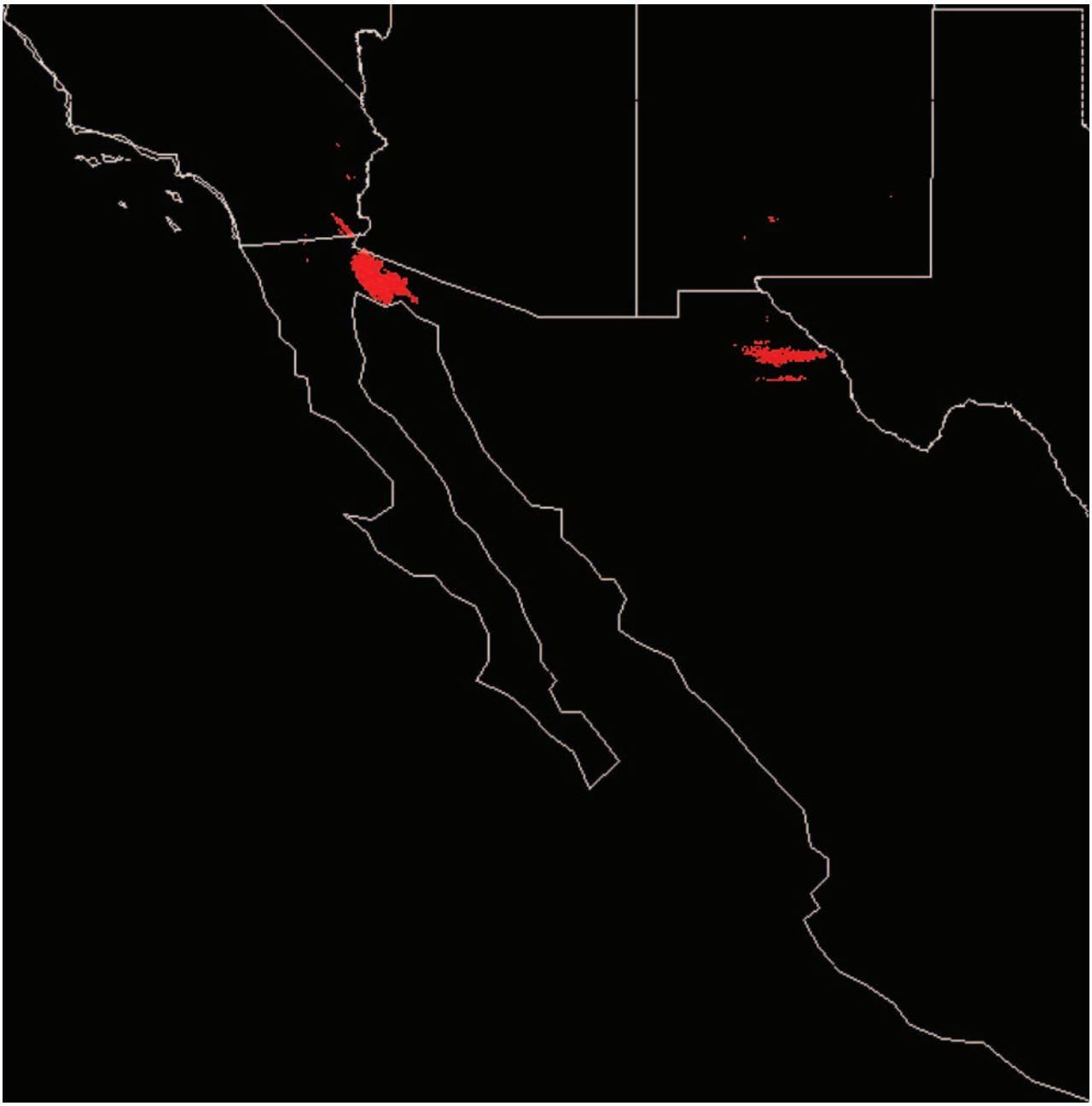

Figure 225. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2005-331 at 18:20 UTC.

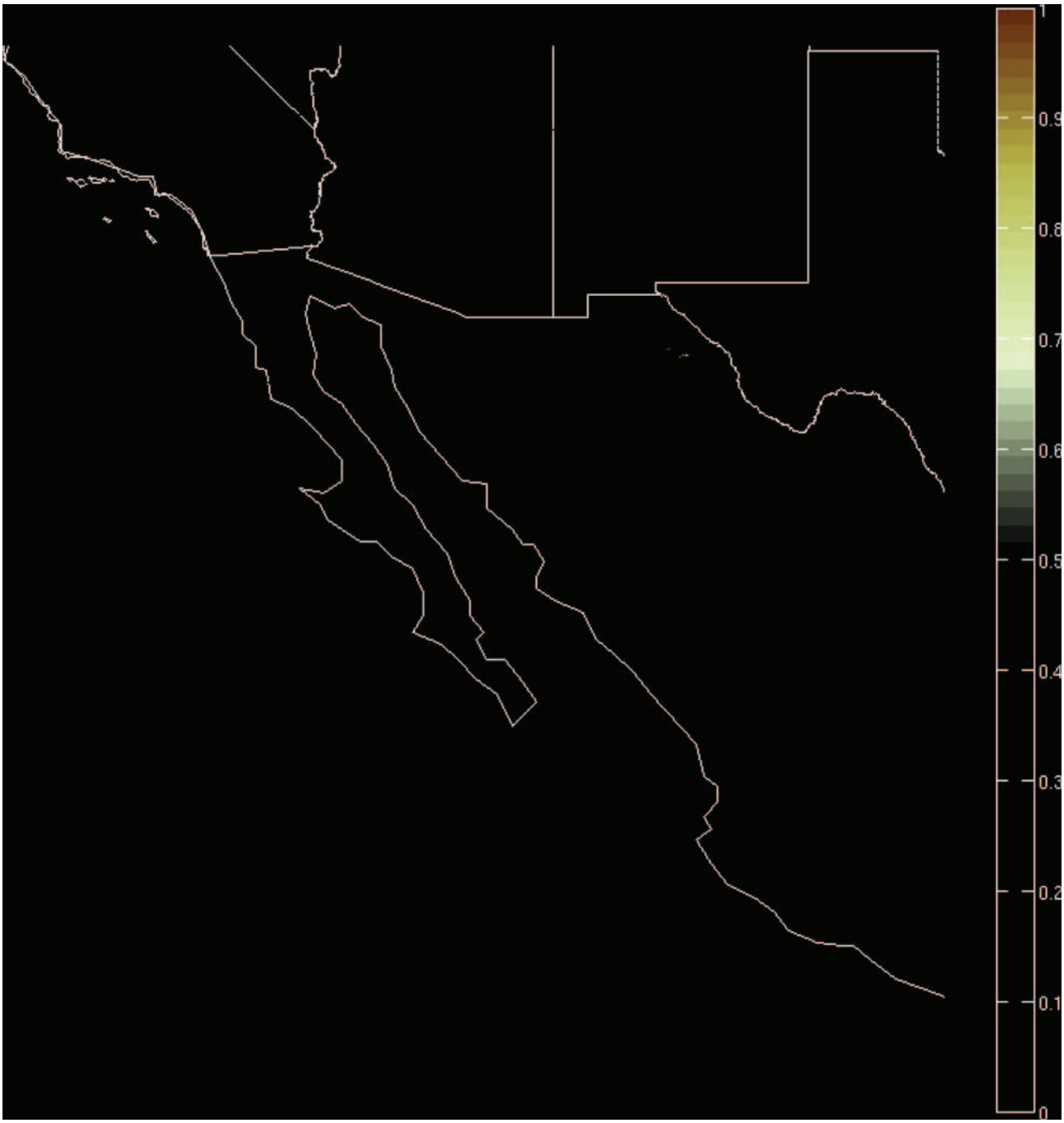

Figure 226. Result of the Maximum a posteriori model (MAP) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

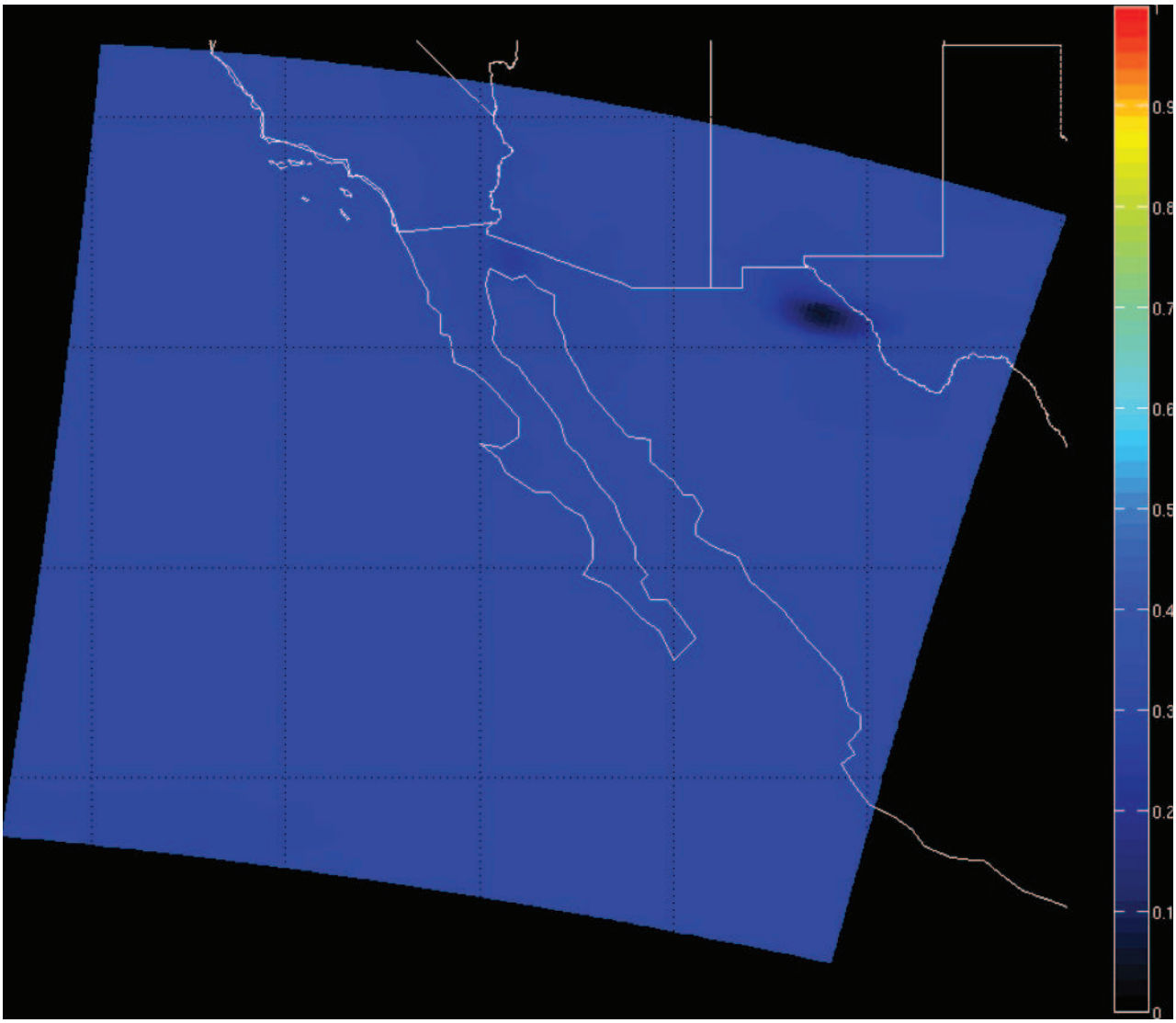

Figure 227. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

Figure 228. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2005-331 at 18:20 UTC.

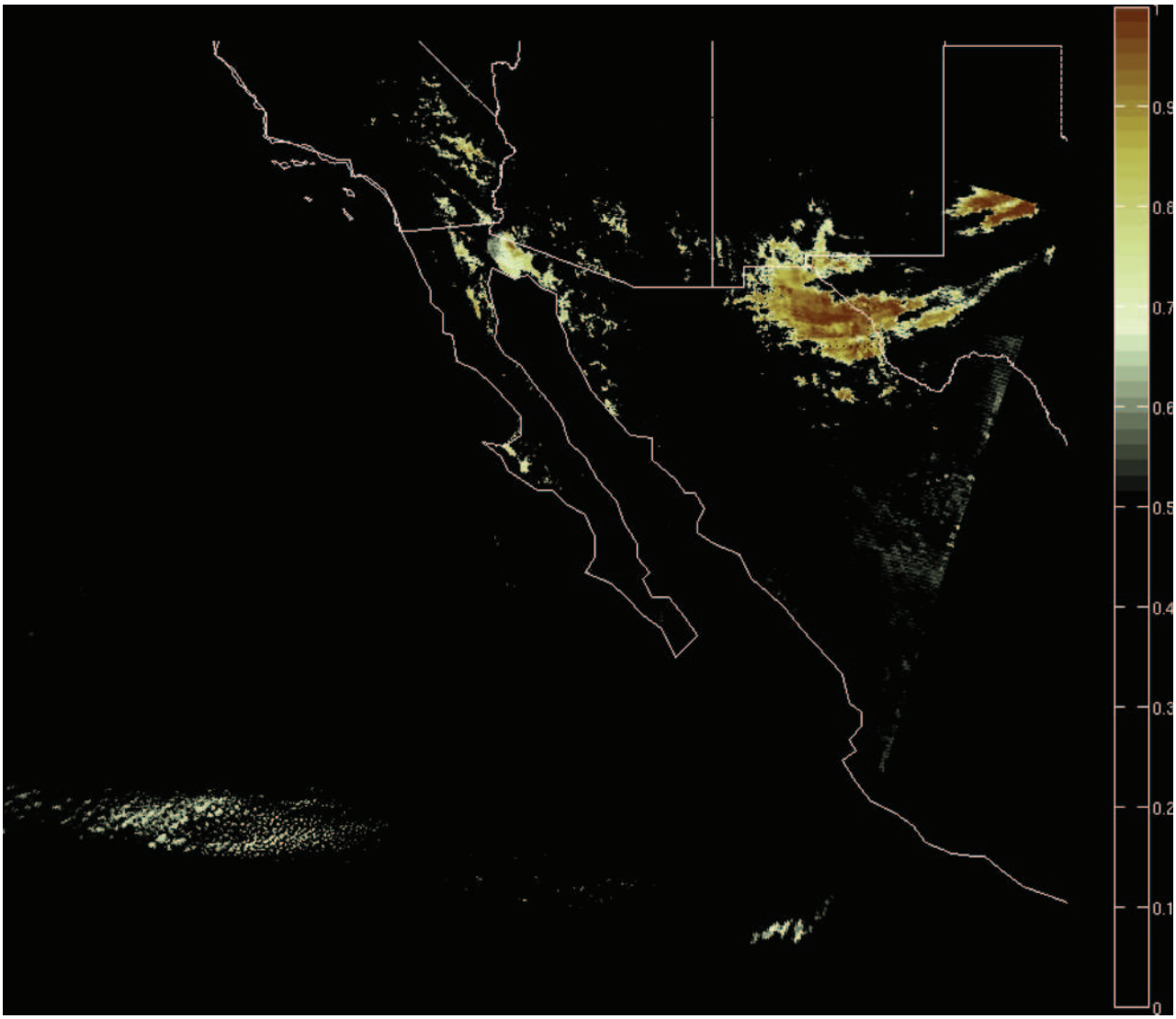

Figure 229. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

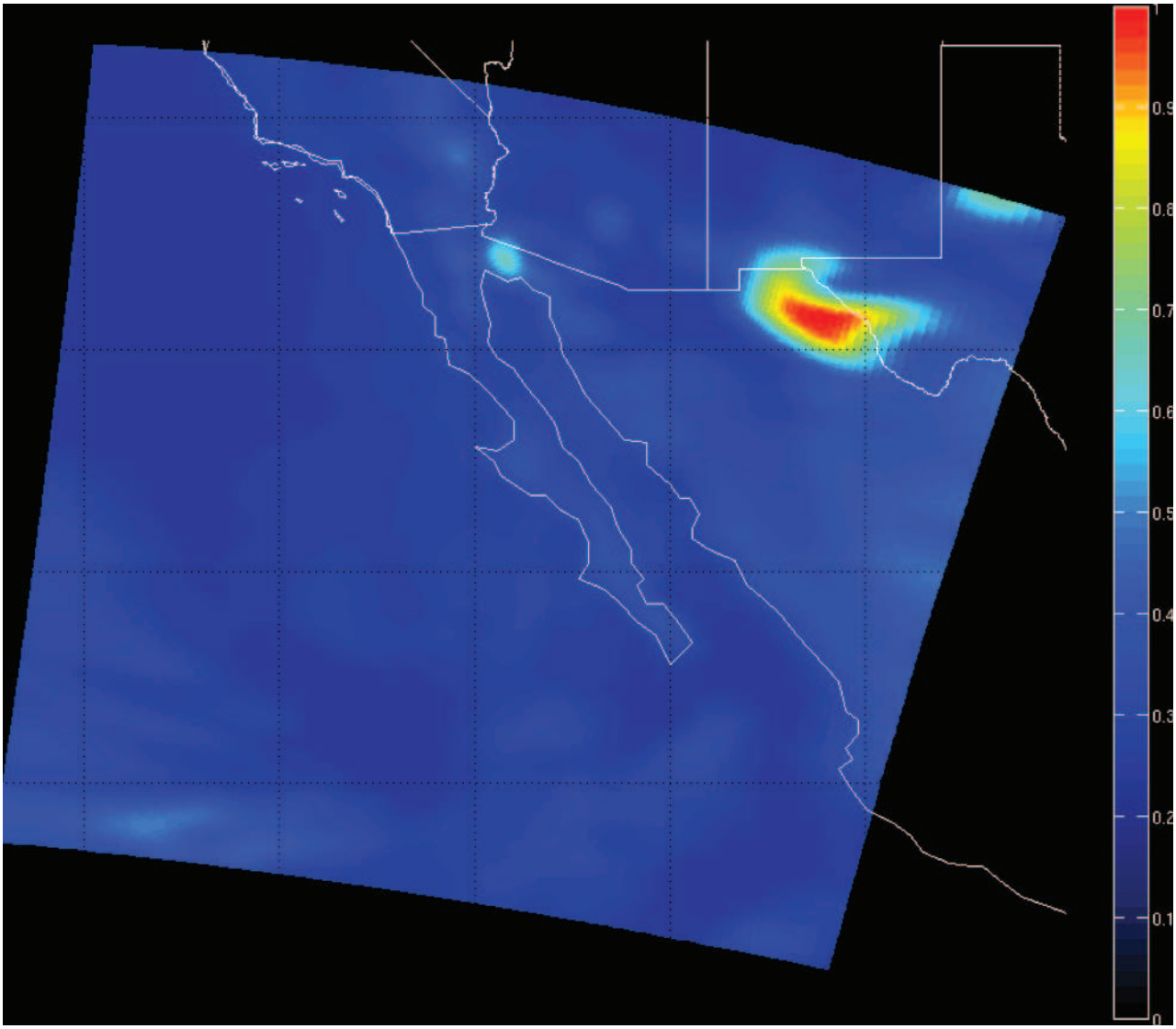

Figure 230. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

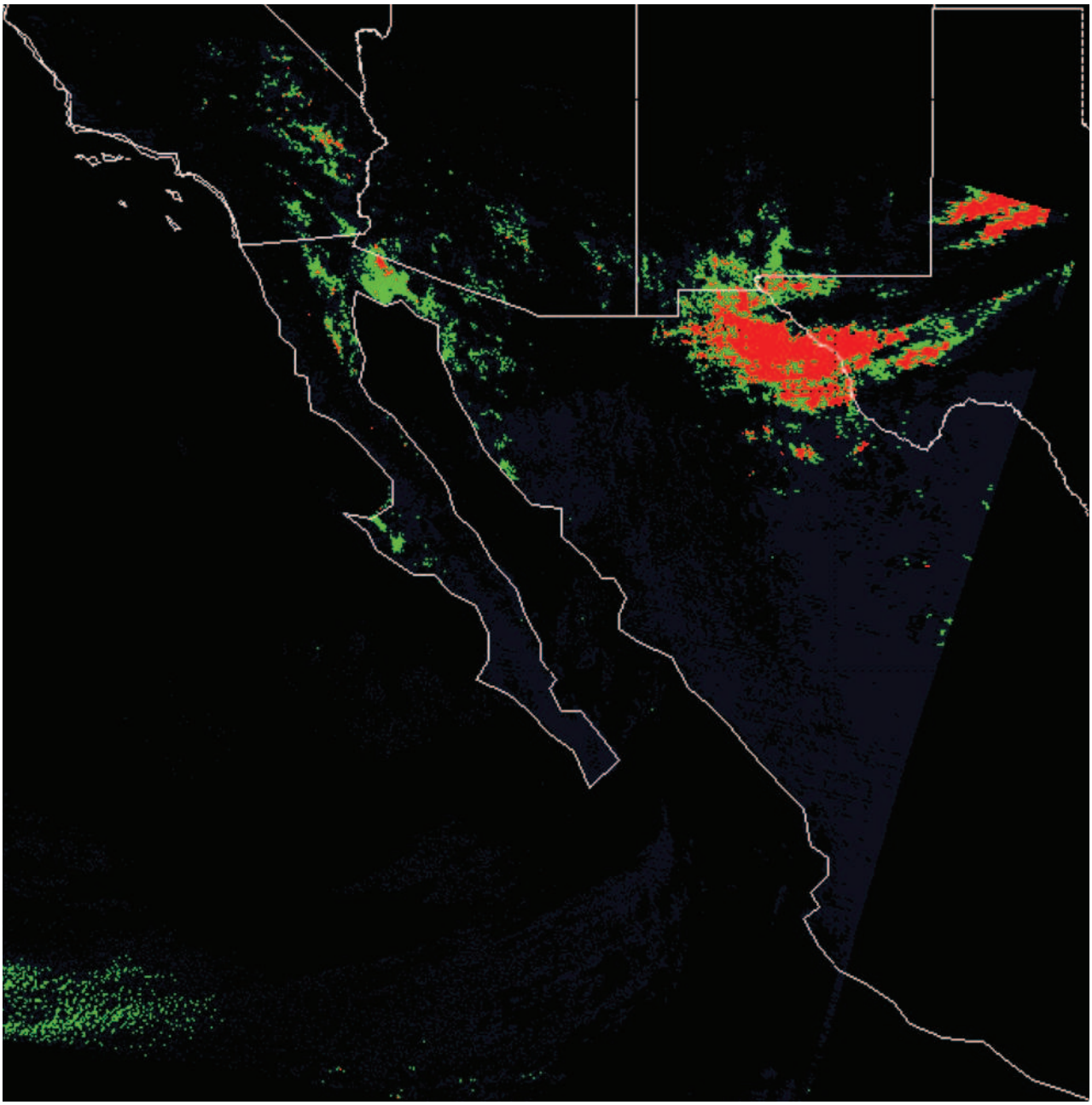

Figure 231. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2005-331 at 18:20 UTC.

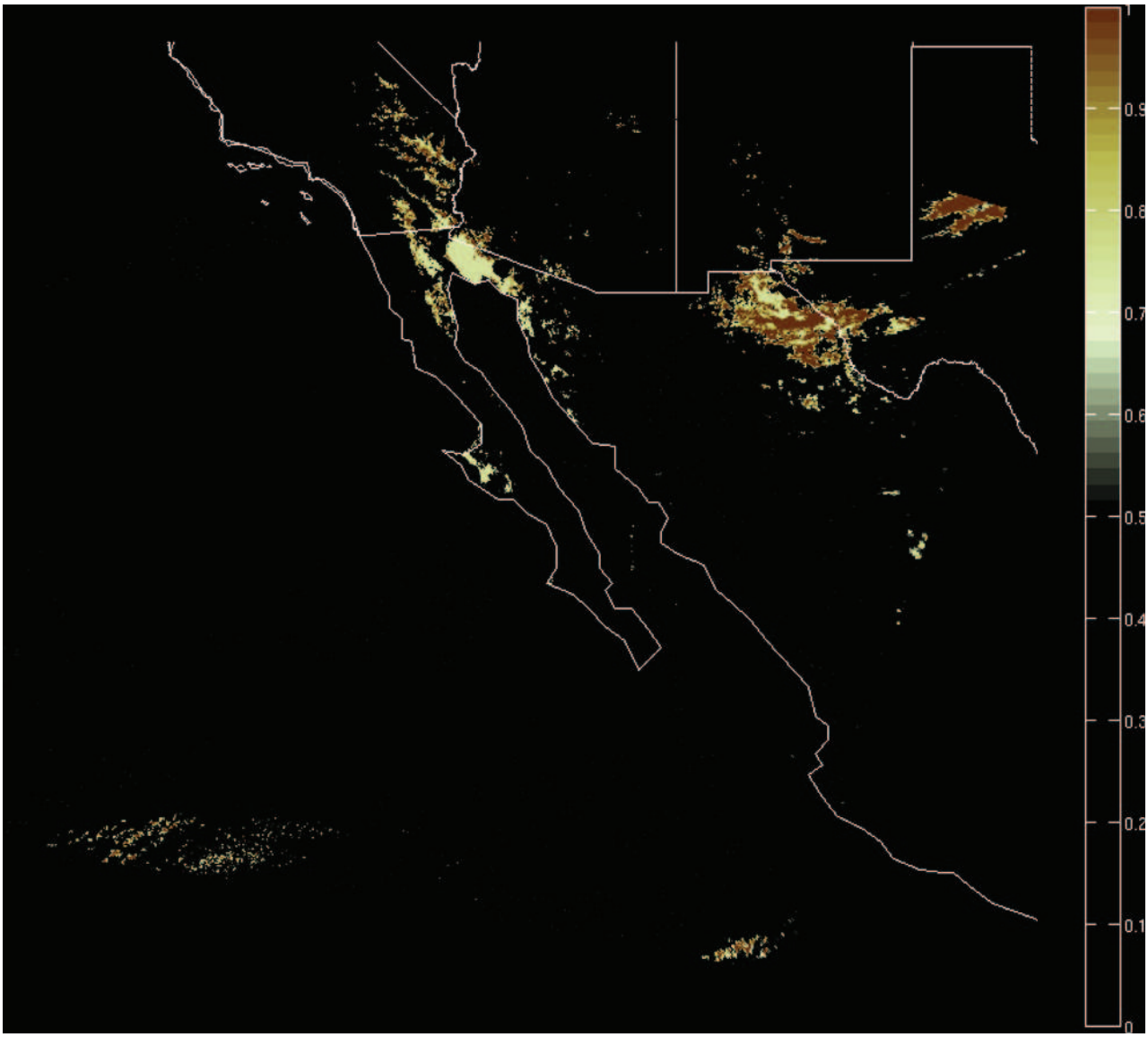

Figure 232. Result of the Probabilistic Neural Network (PNN) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

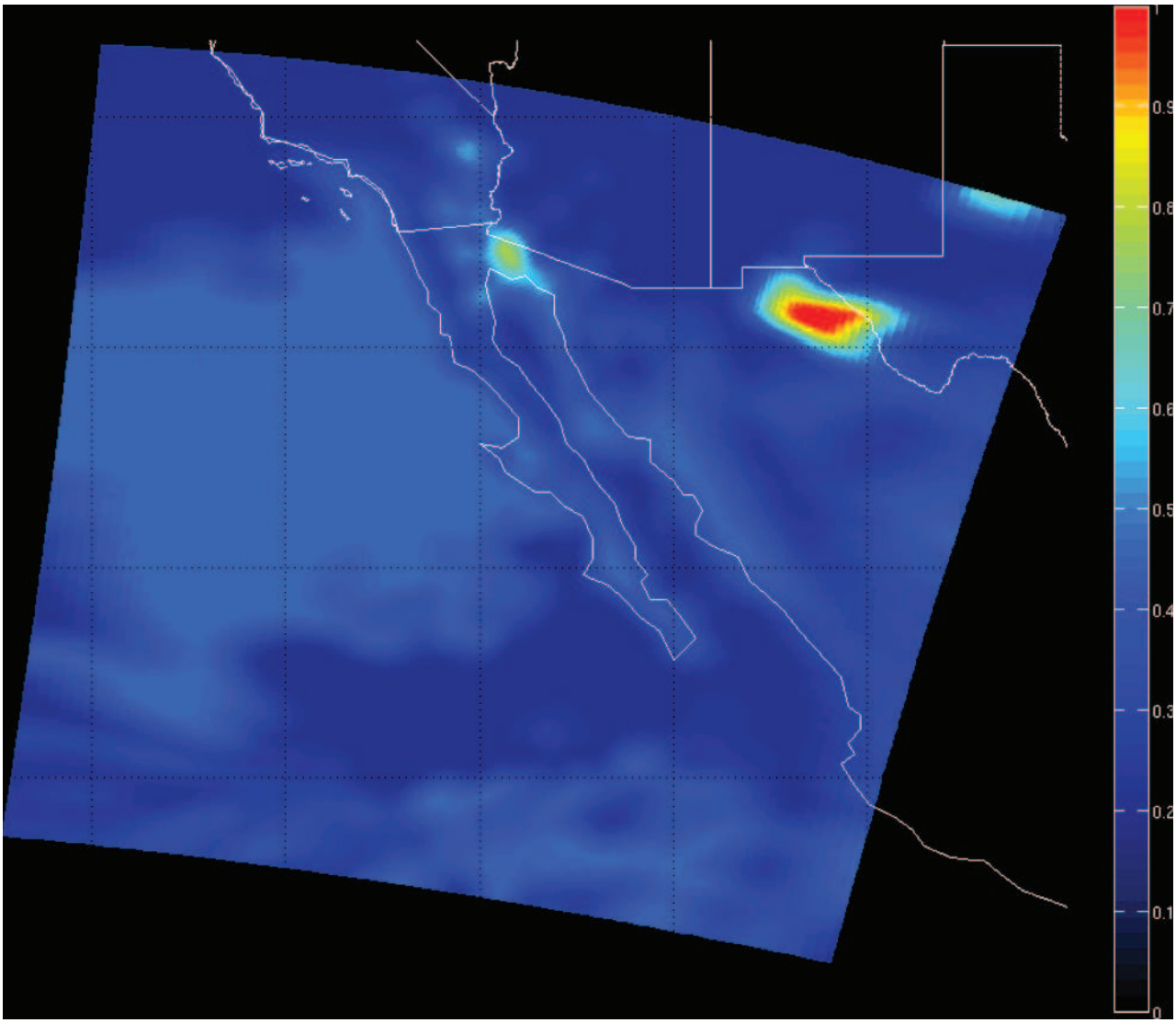

Figure 233. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2005-331 at 18:20 UTC. The colorbar shows the index of probability.

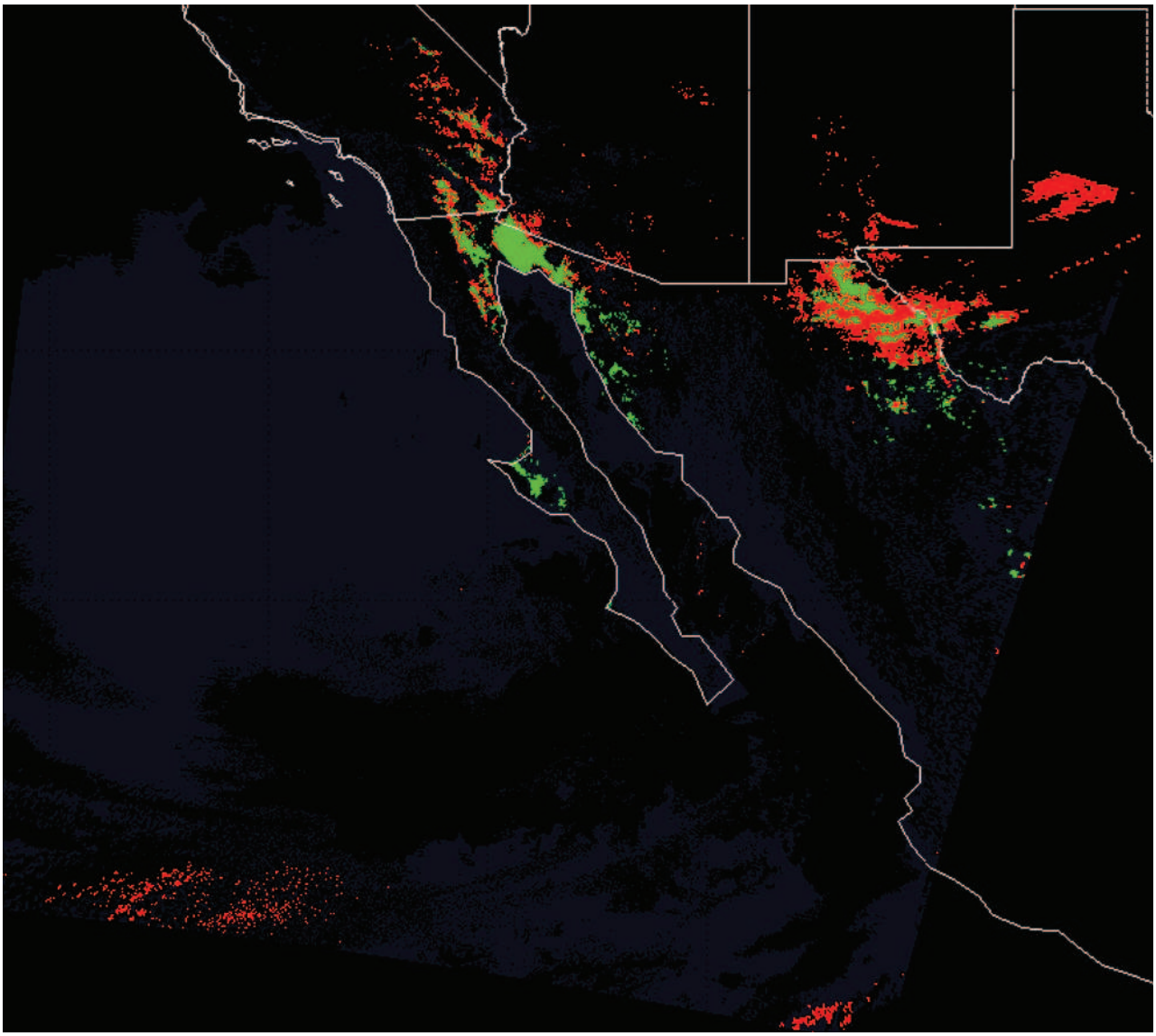

Figure 234. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2005-331 at 18:20 UTC.

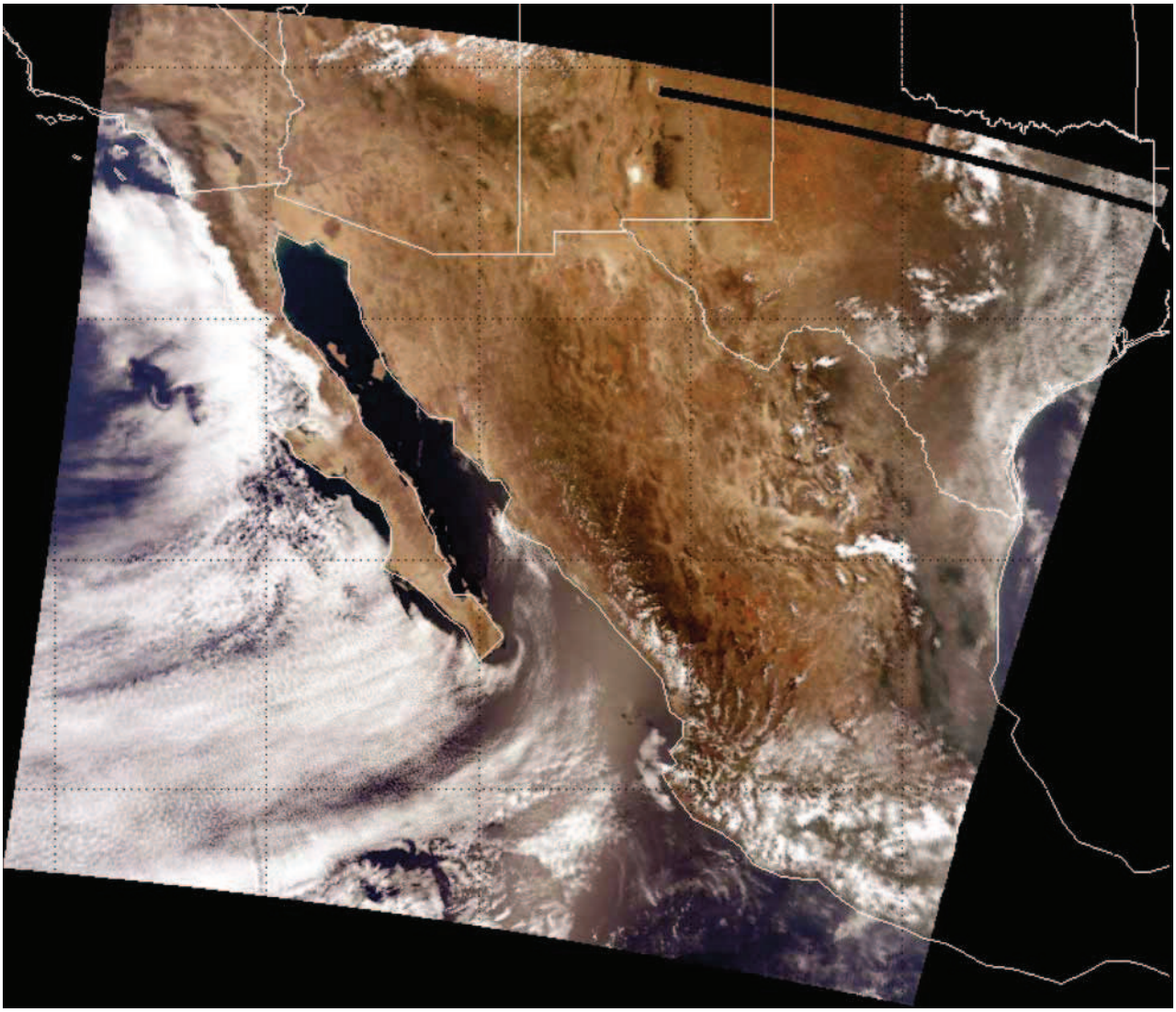

Figure 235. True color image for the event of 2001-164 at 18:05 UTC.

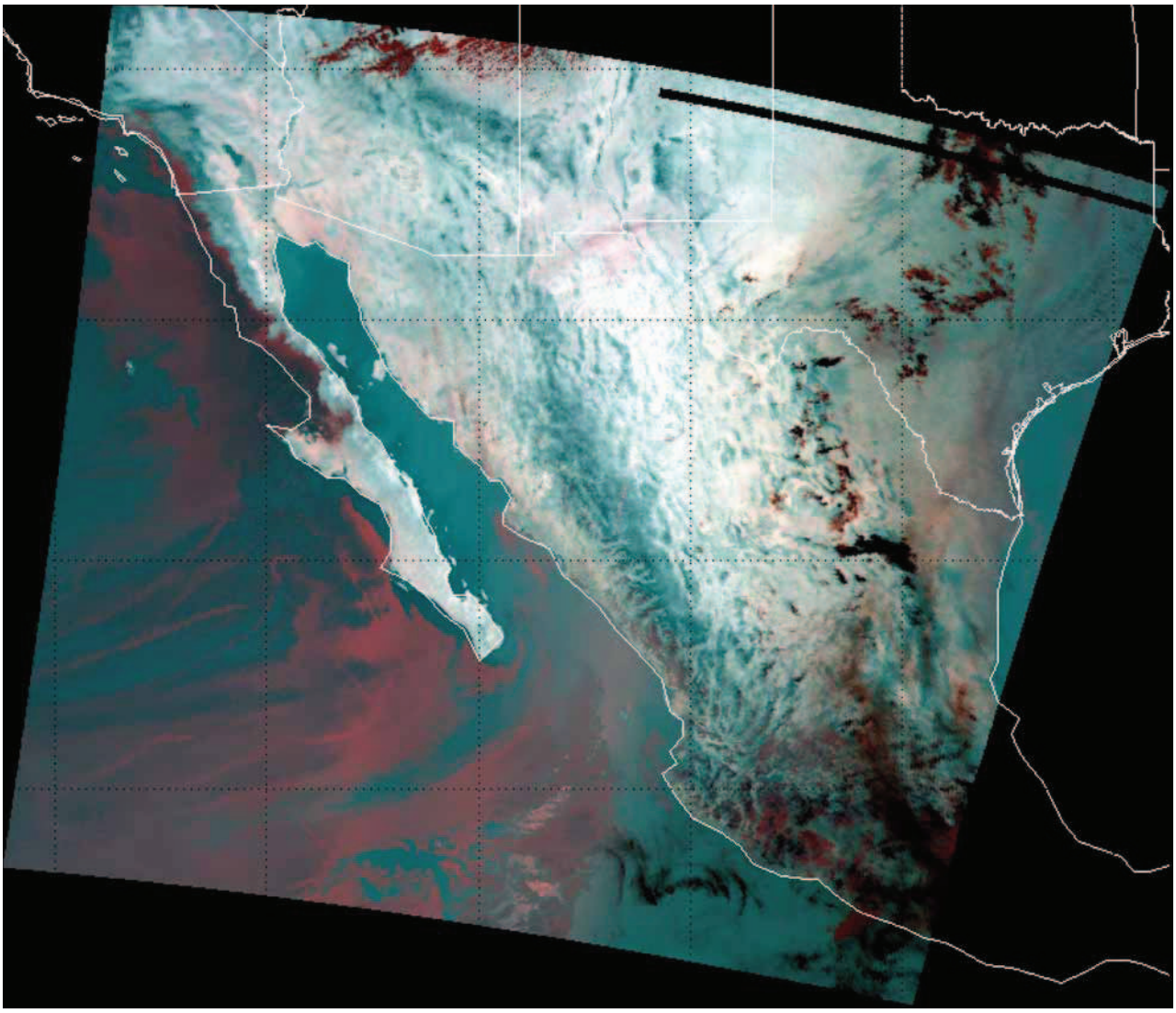

Figure 236. False color image for the event of 2001-164 at 18:05 UTC.

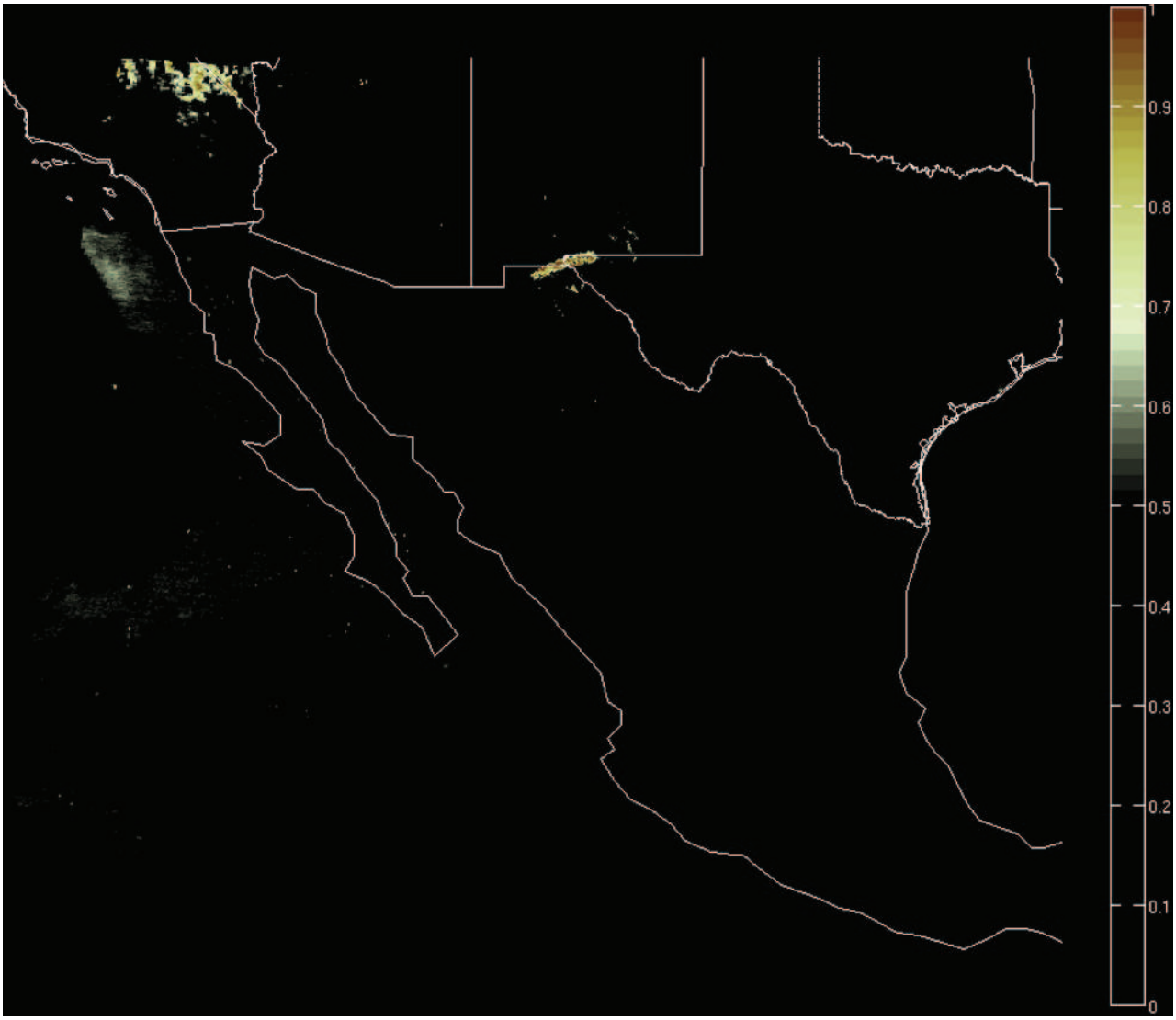

Figure 237. Result of the probabilistic display model (PD) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

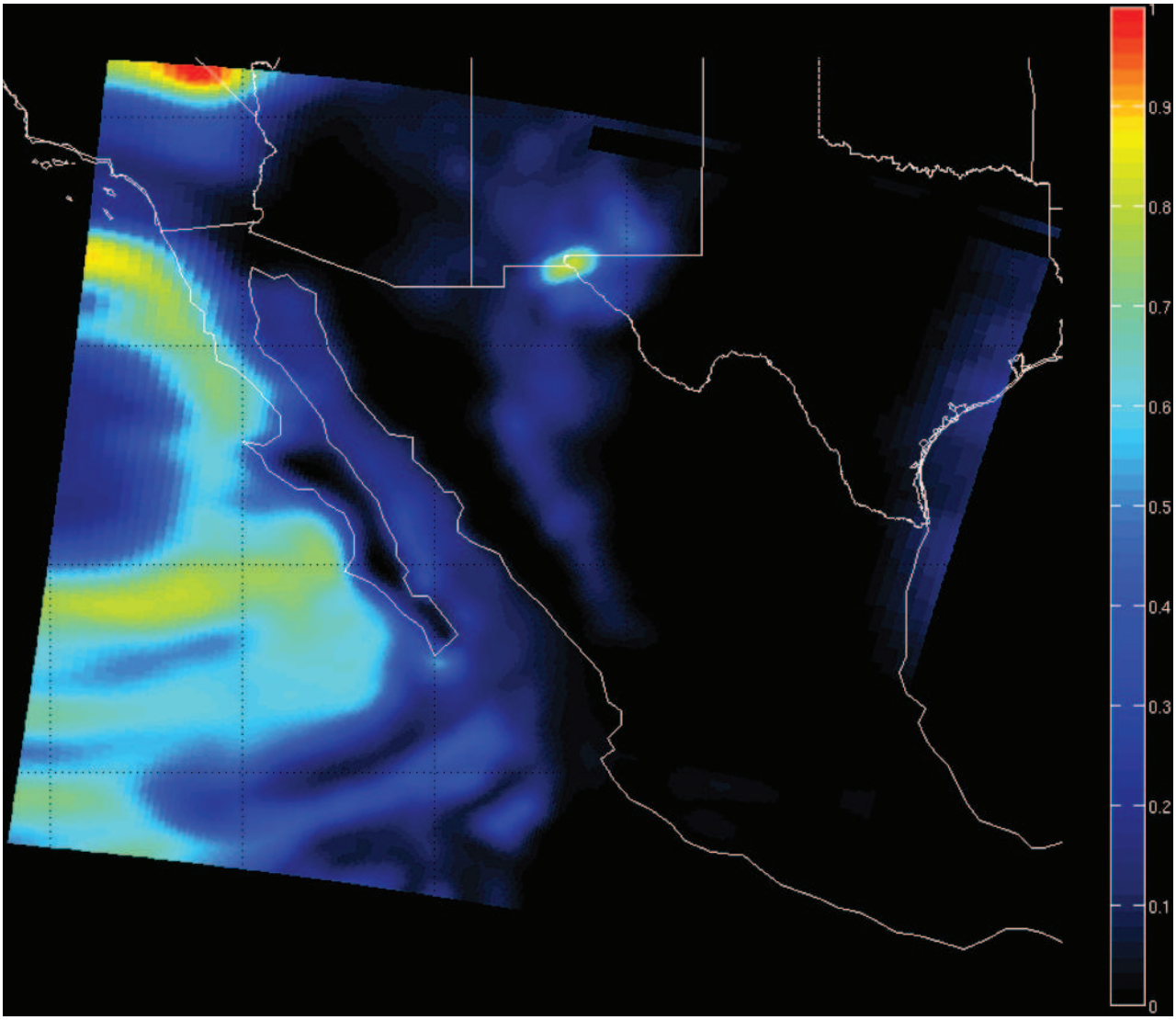

Figure 238. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

Figure 239. Result of the Maximum Likelihood model (MLE) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

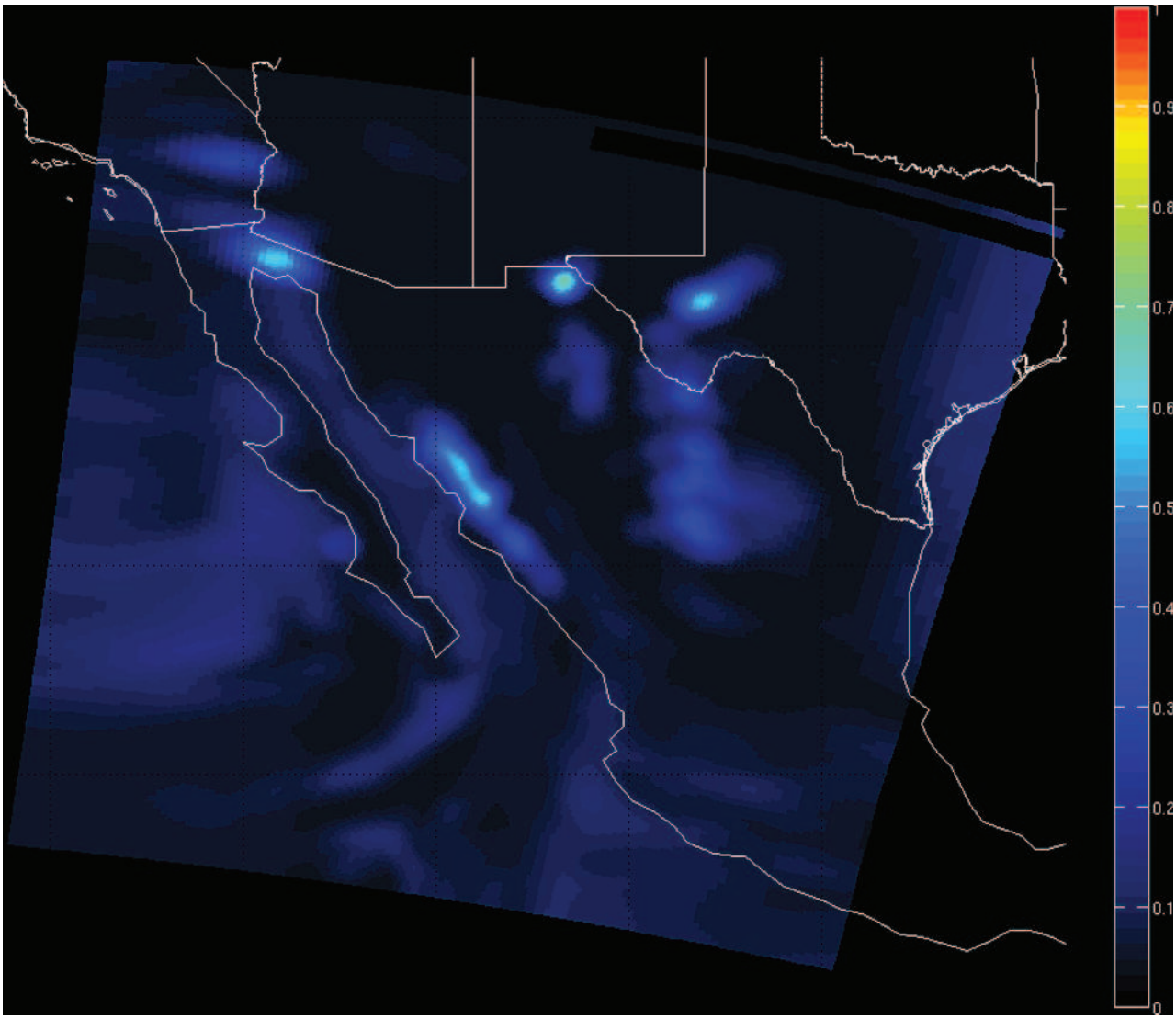

Figure 240. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

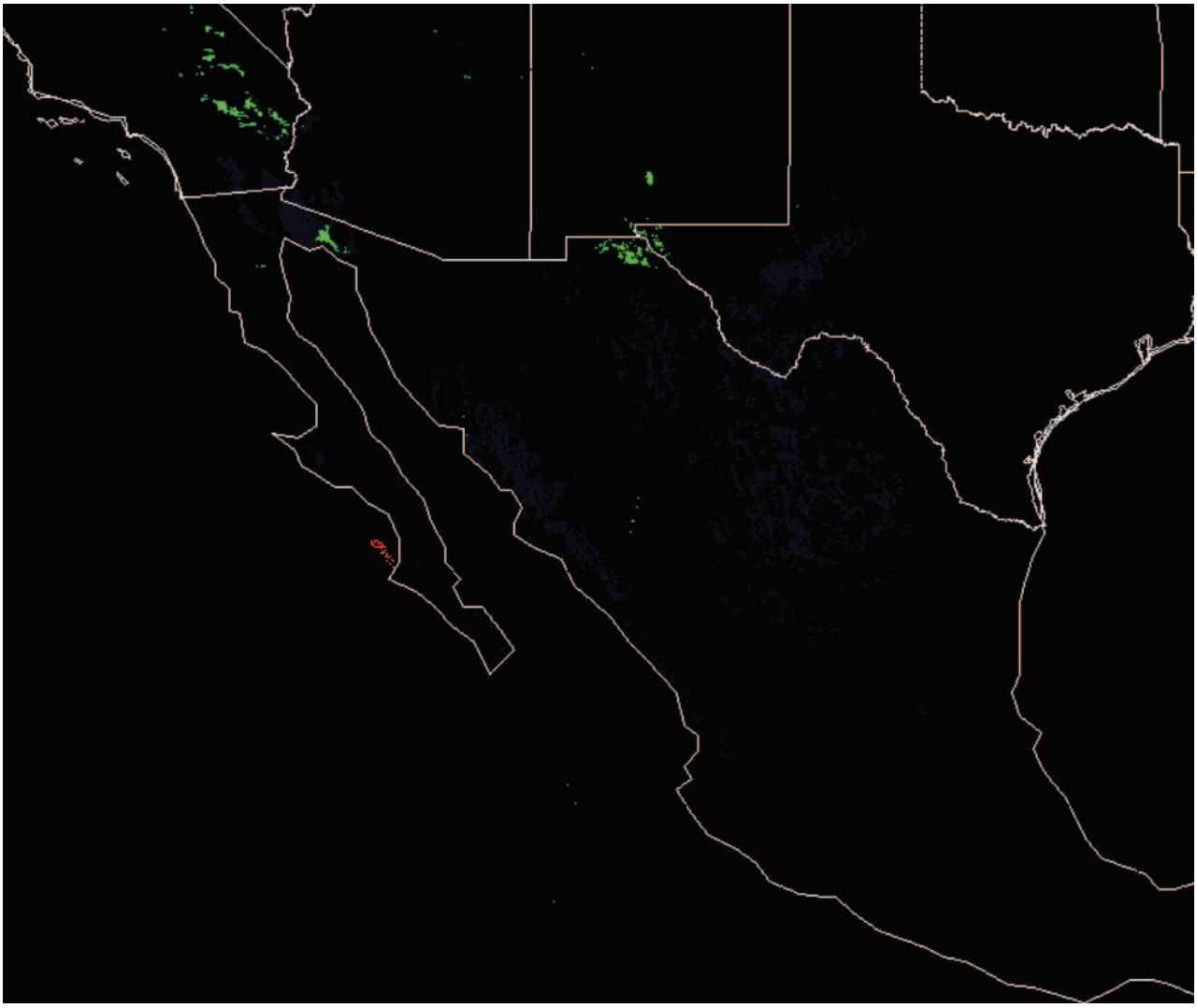

Figure 241. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2001-164 at 18:05 UTC.

Figure 242. Result of the Maximum a posteriori model (MAP) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

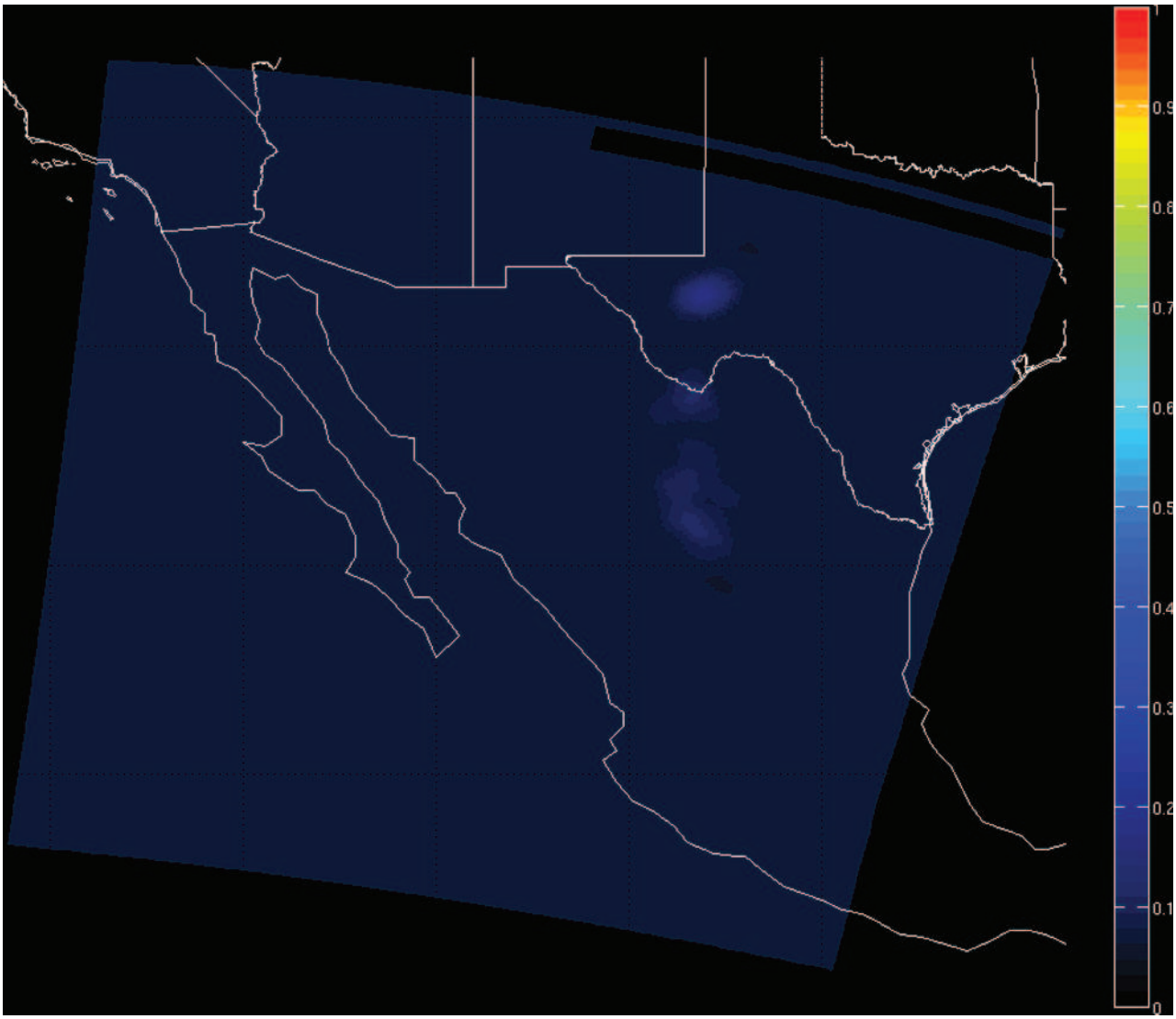

Figure 243. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

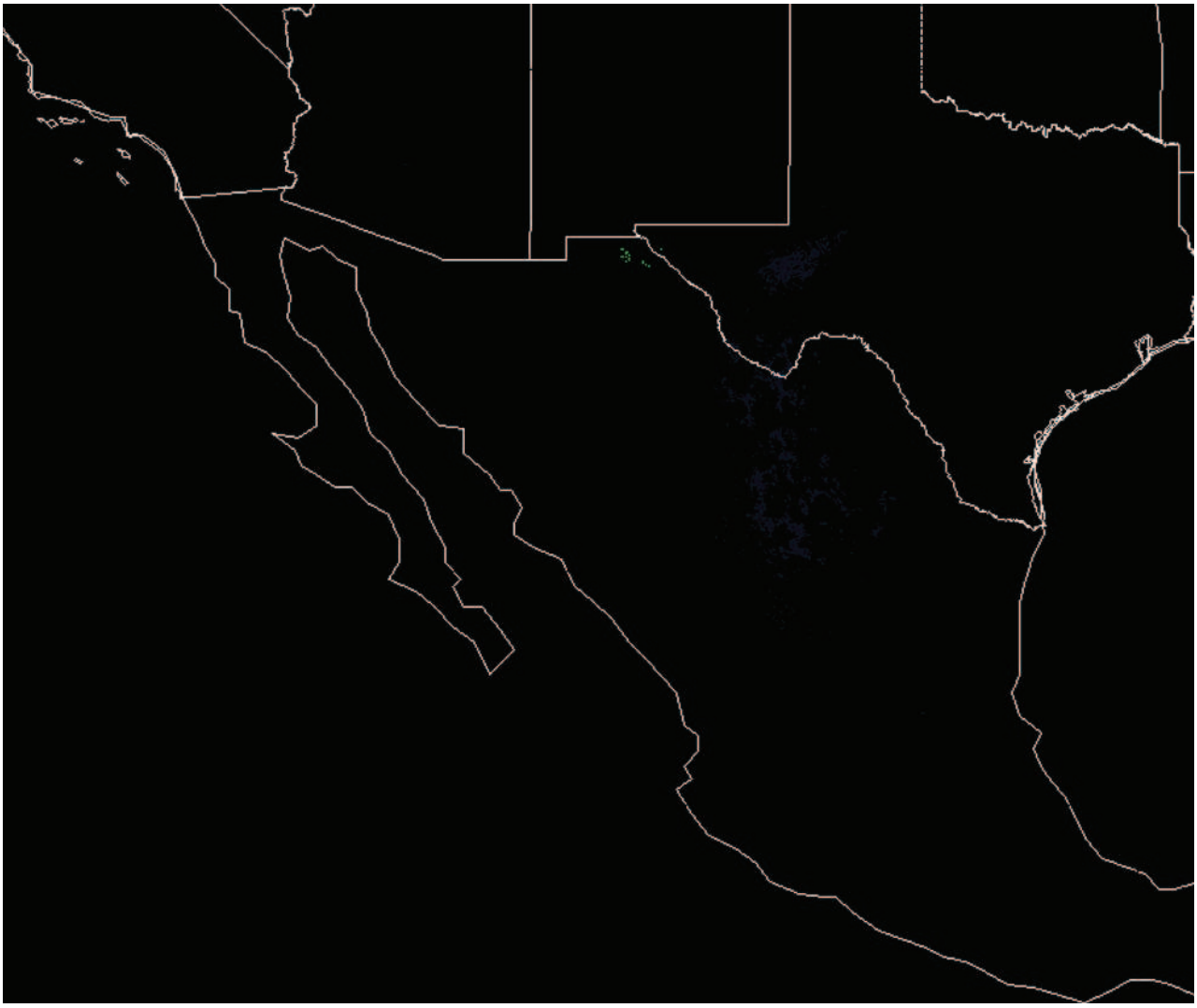

Figure 244. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2001-164 at 18:05 UTC.

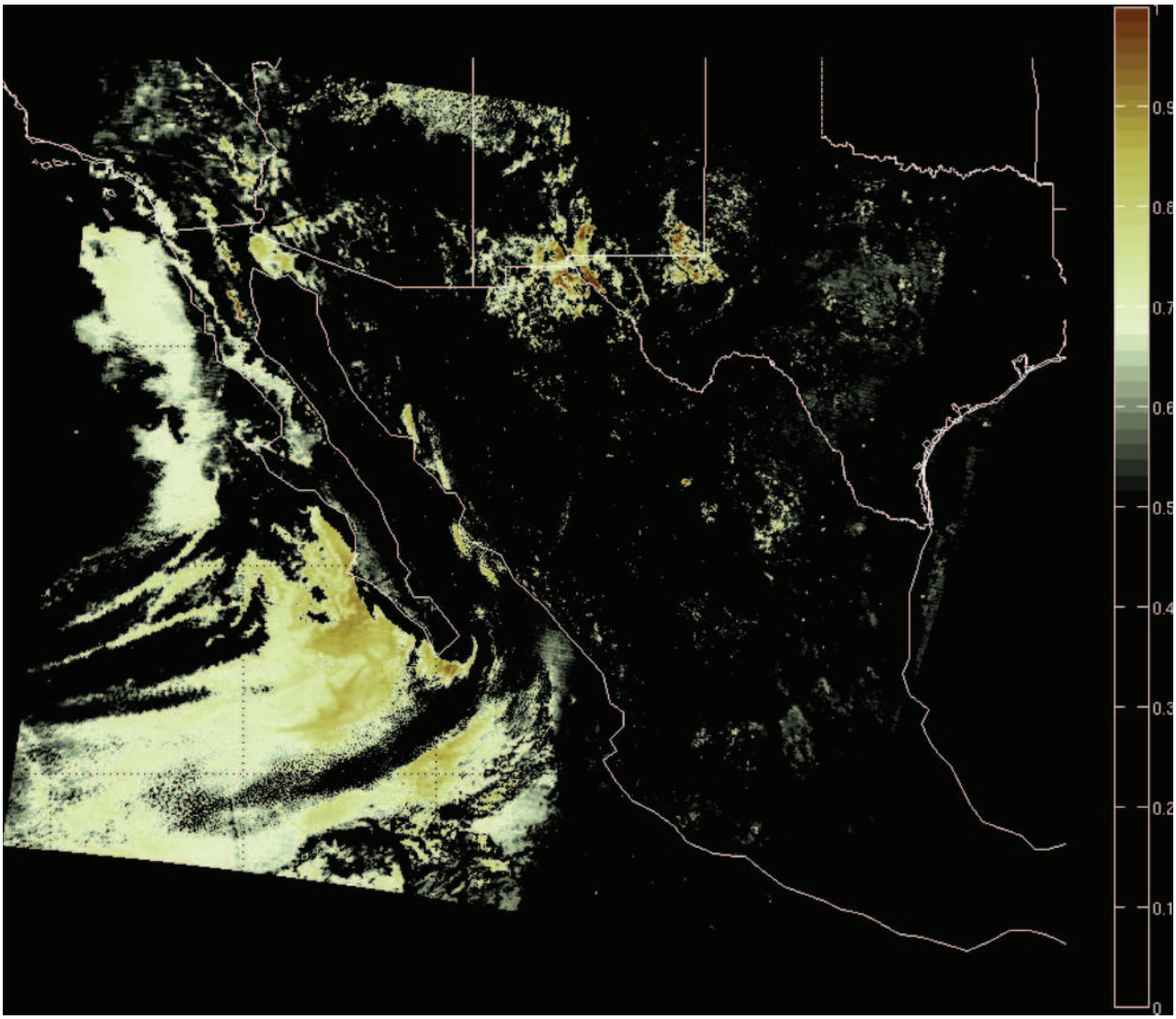

Figure 245. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

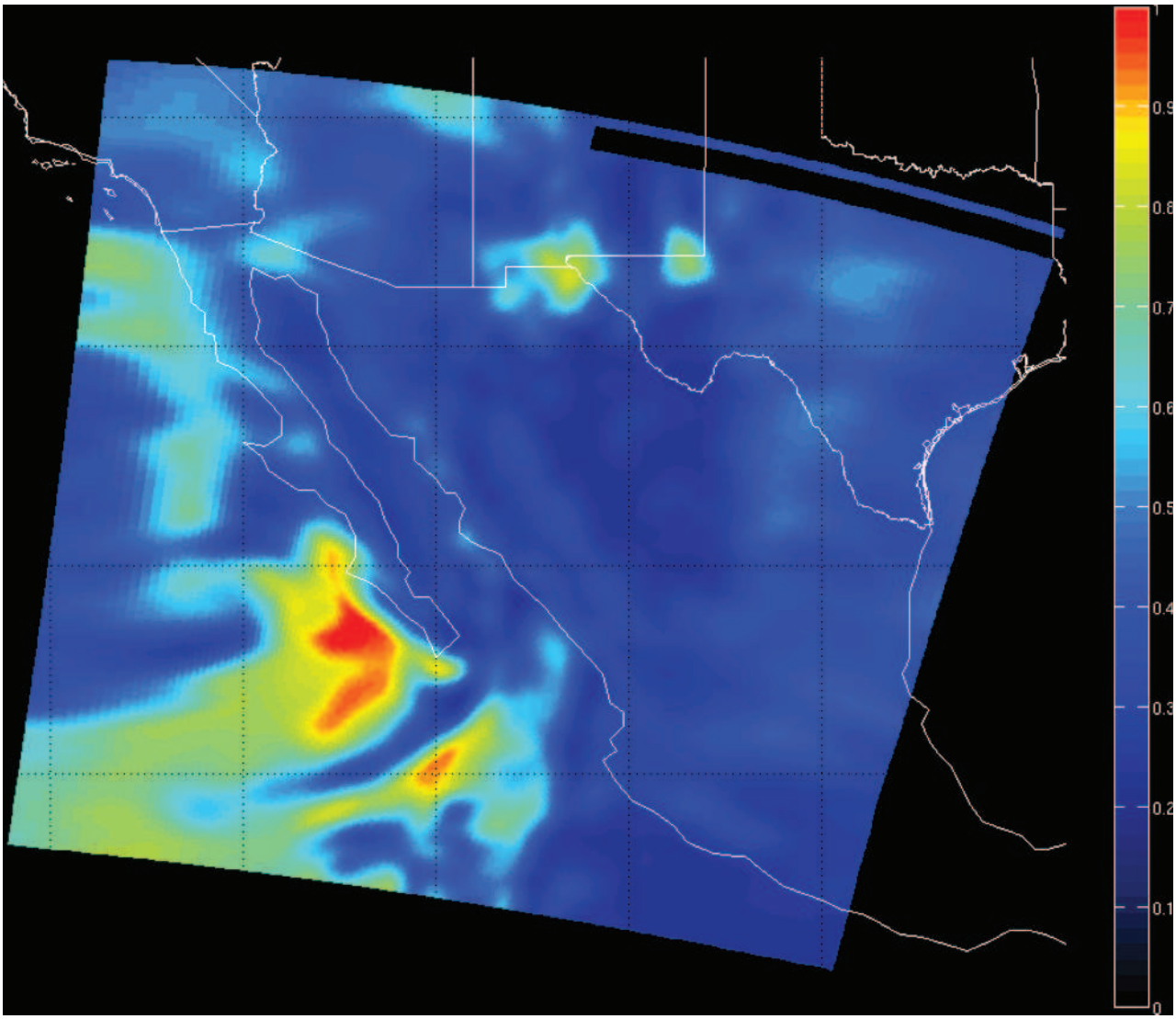

Figure 246. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

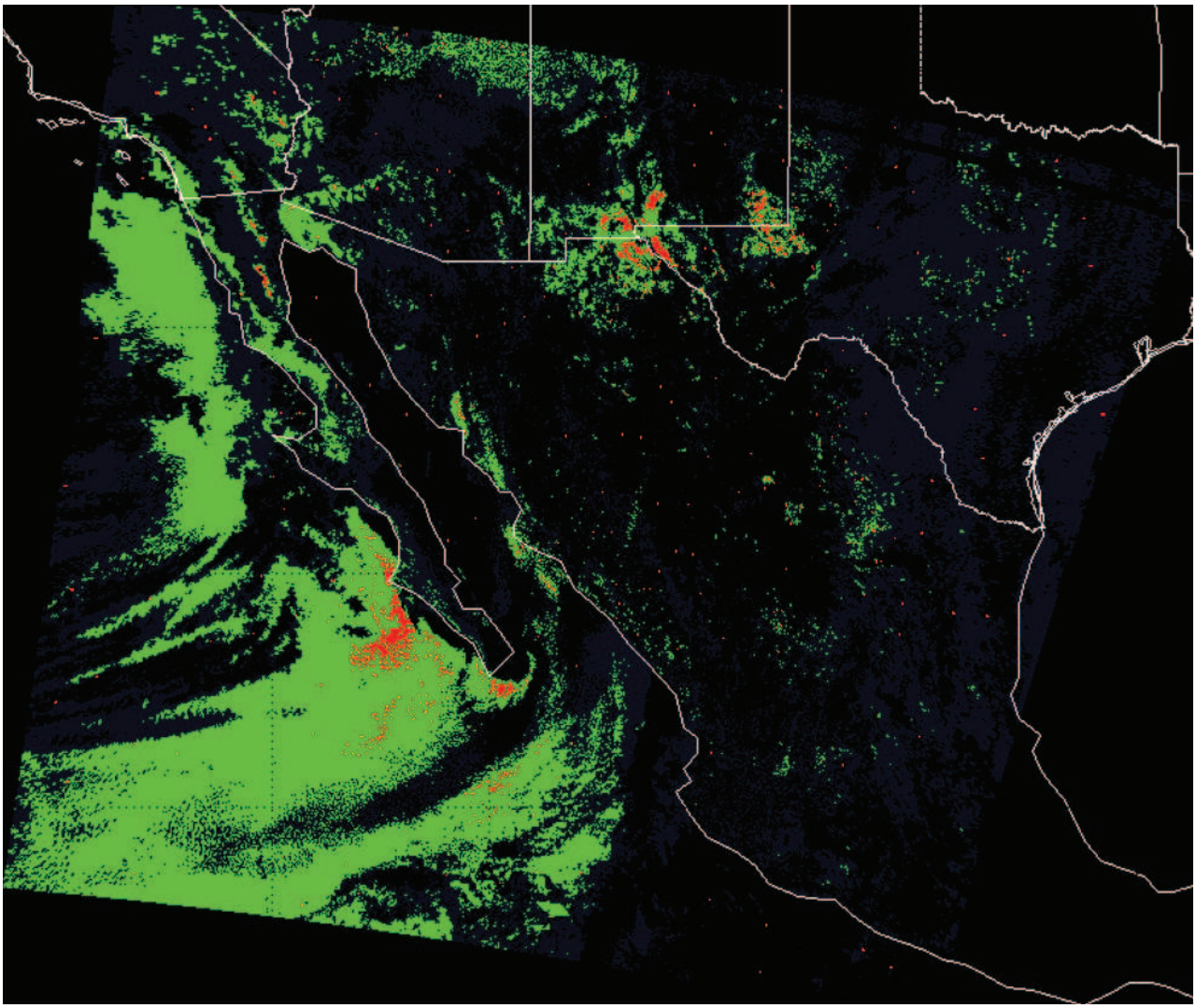

Figure 247. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-164 at 18:05 UTC.

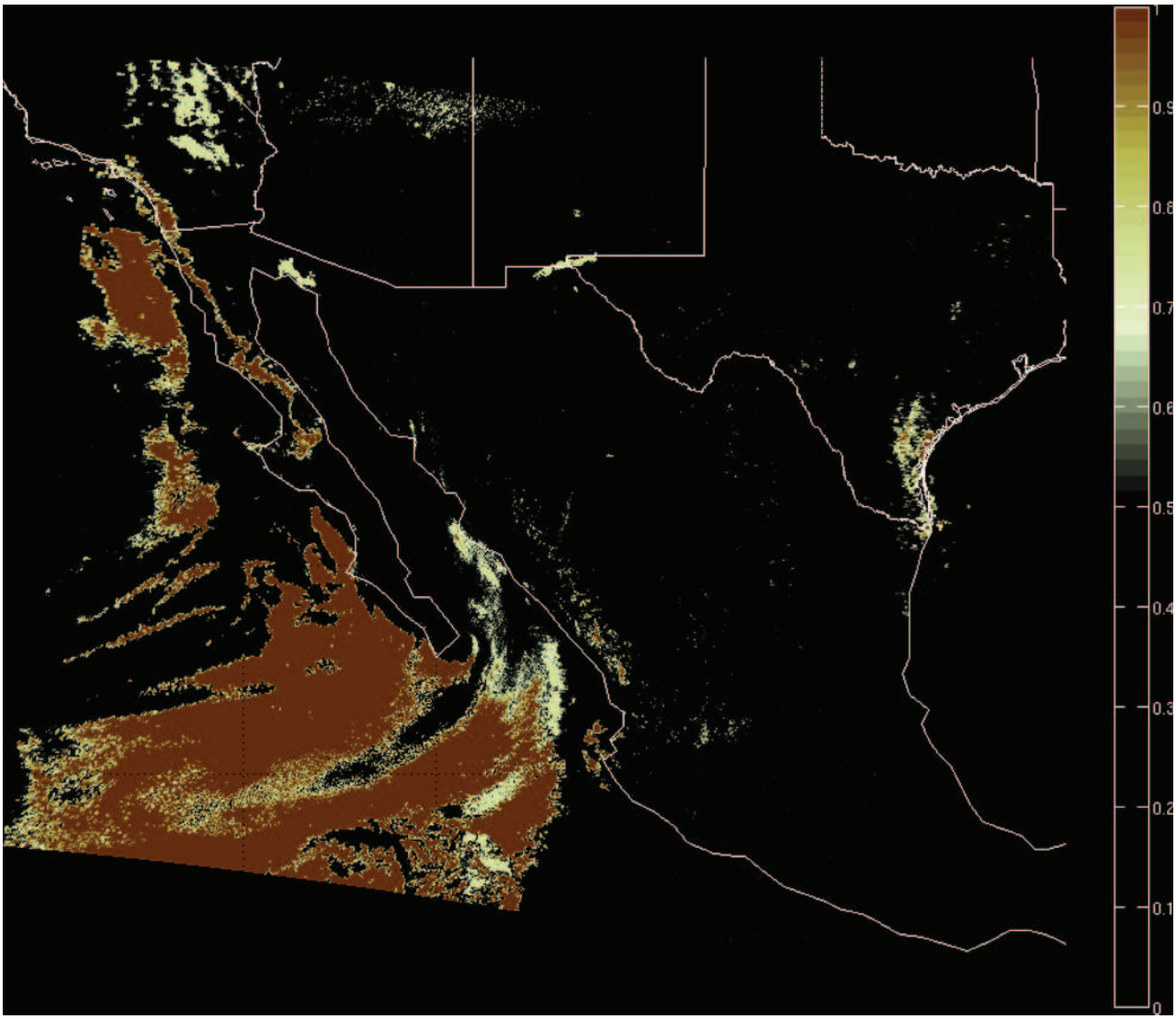

Figure 248. Result of the Probabilistic Neural Network (PNN) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

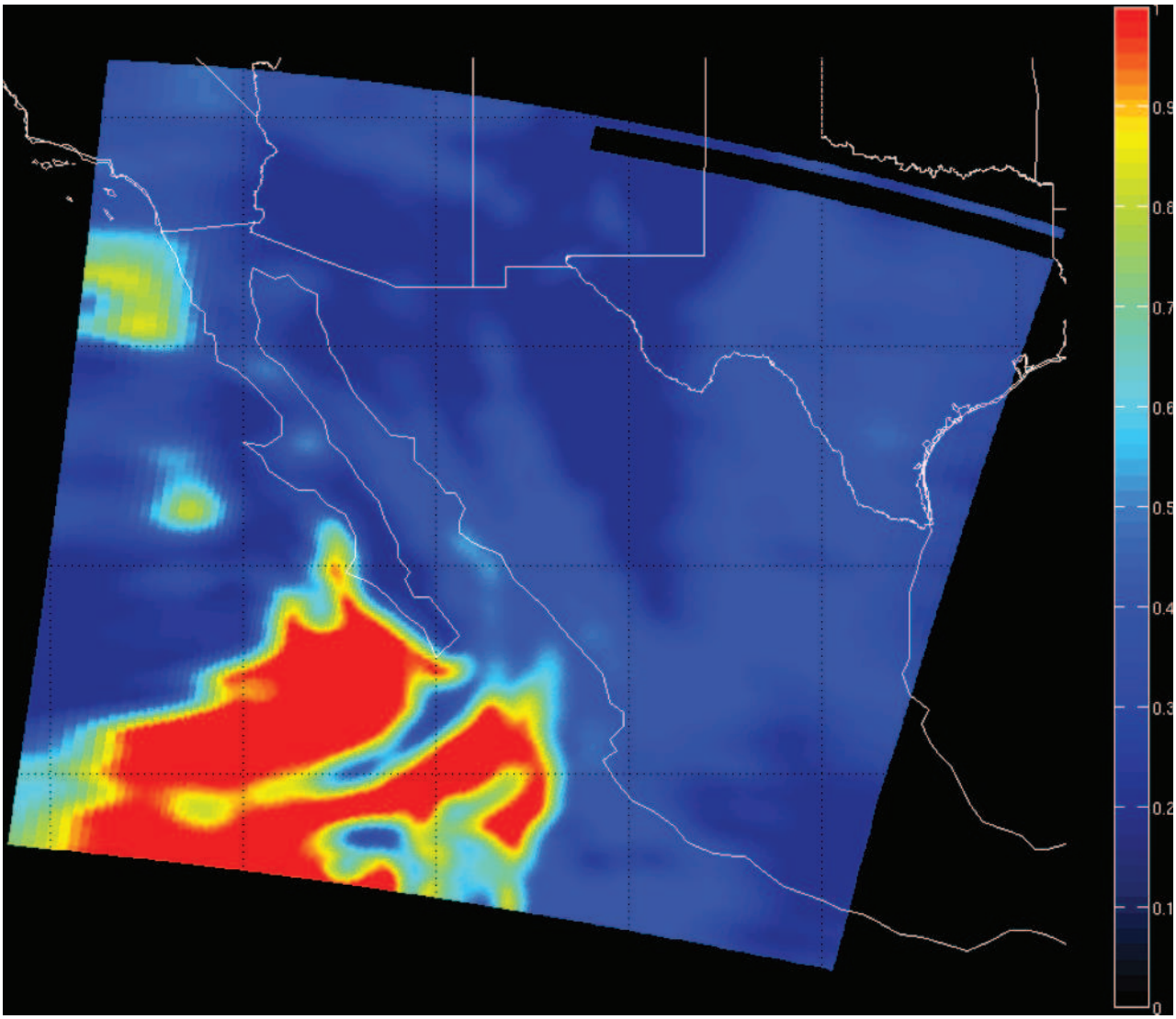

Figure 249. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2001-164 at 18:05 UTC. The colorbar shows the index of probability.

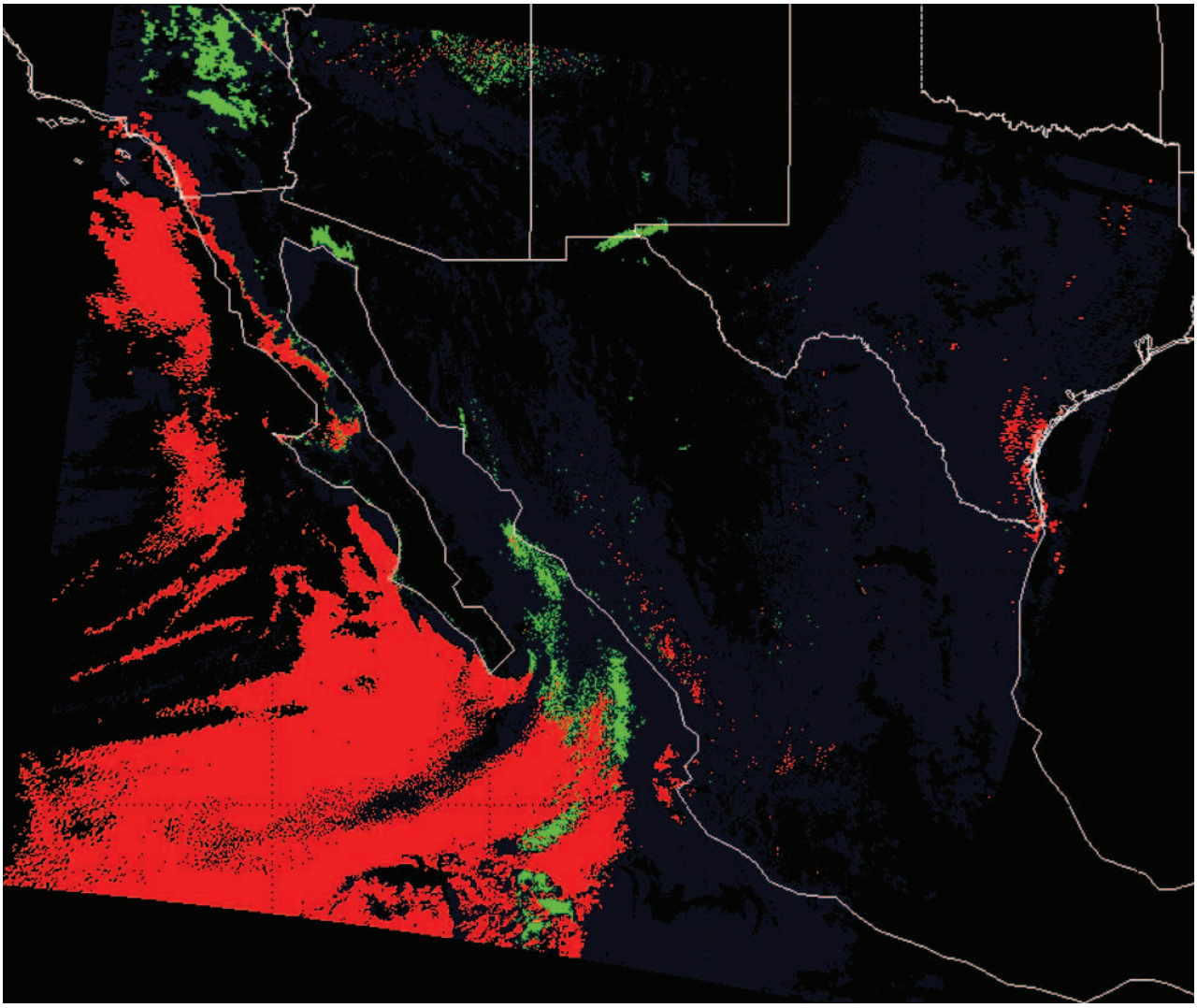

Figure 250. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2001-164 at 18:05 UTC.

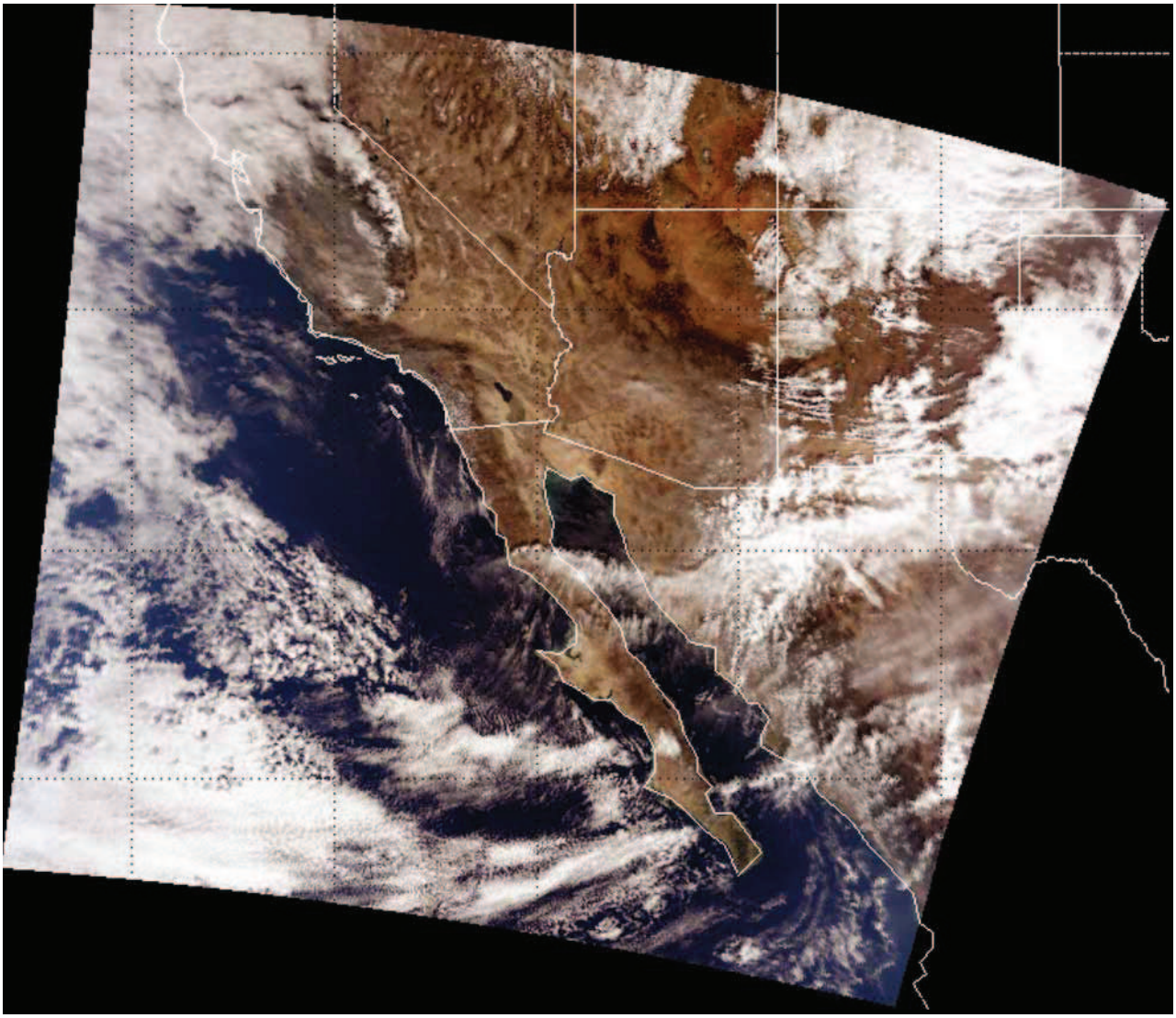

Figure 251. True color image for the event of 2001-327 at 18:30 UTC.

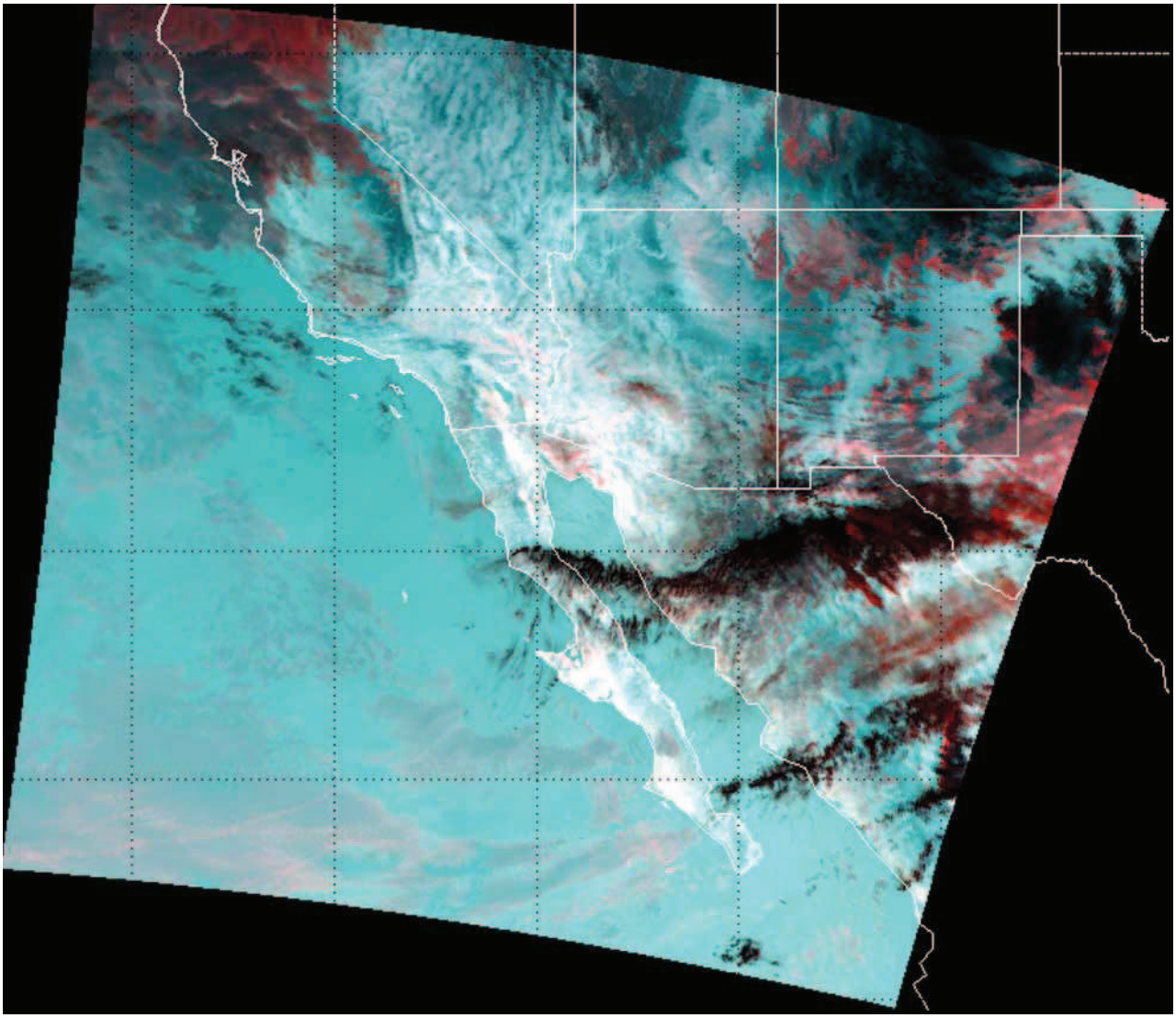

Figure 252. False color image for the event of 2001-327 at 18:30 UTC.

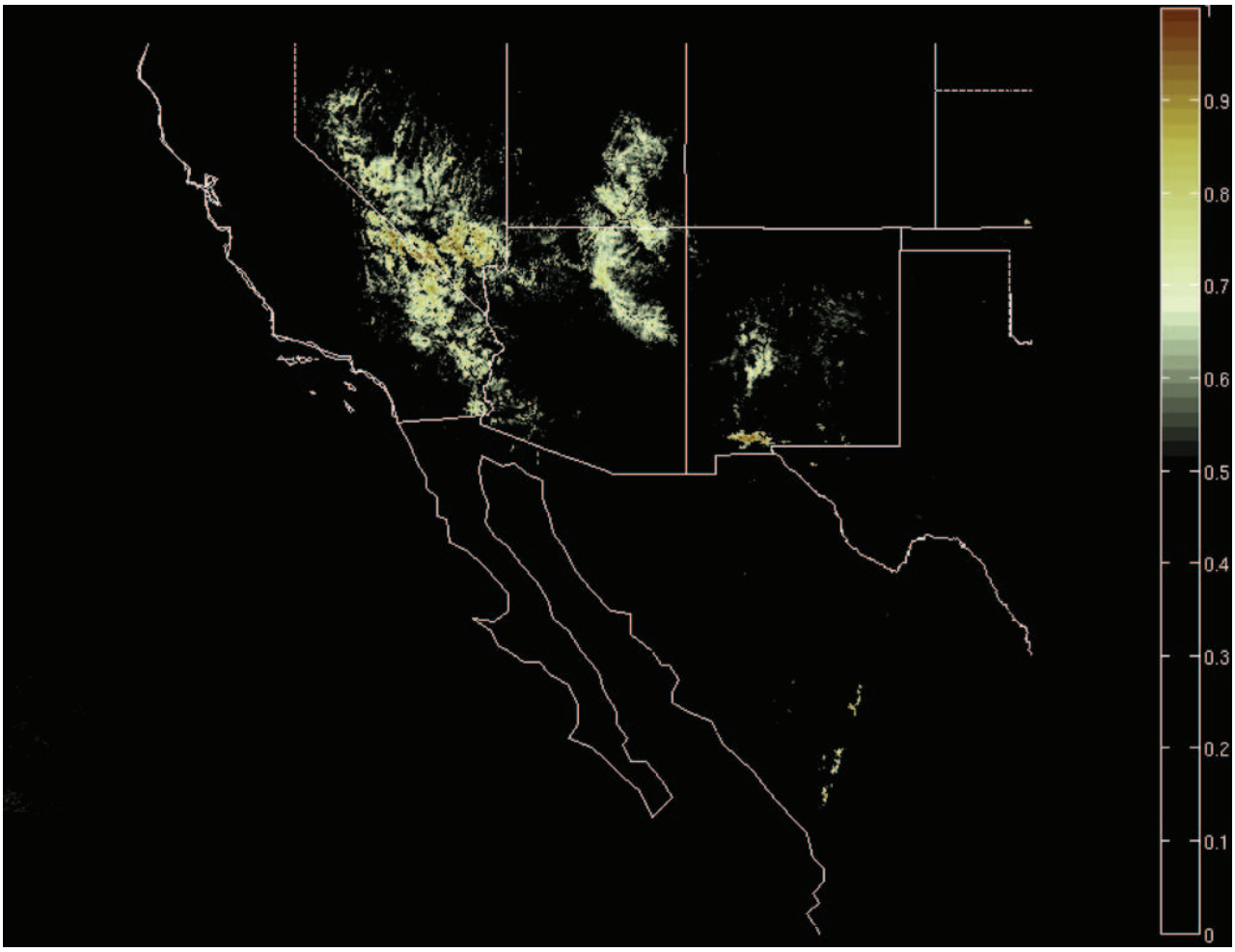

Figure 253. Result of the probabilistic display model (PD) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

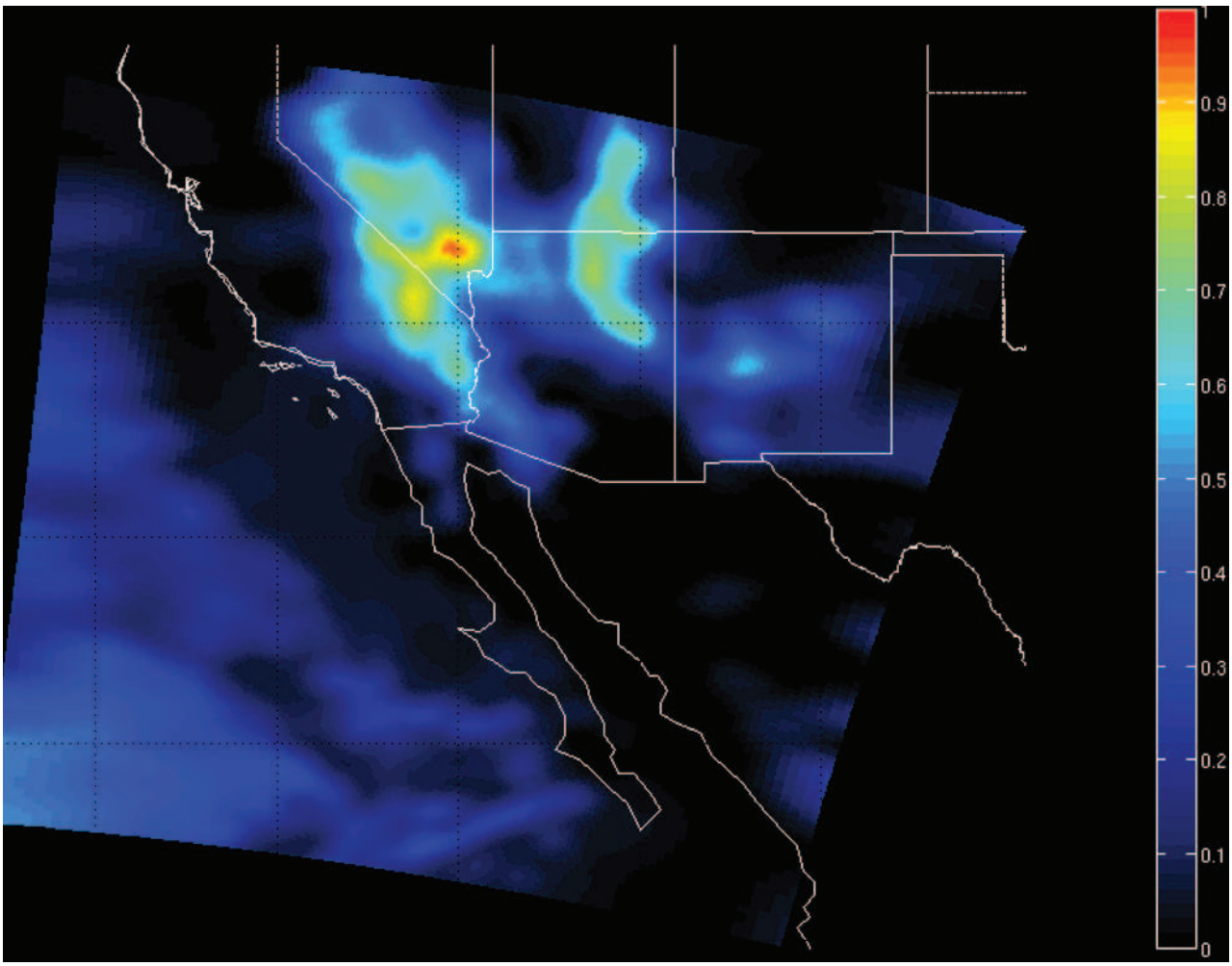

Figure 254. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

Figure 255. Result of the Maximum Likelihood model (MLE) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

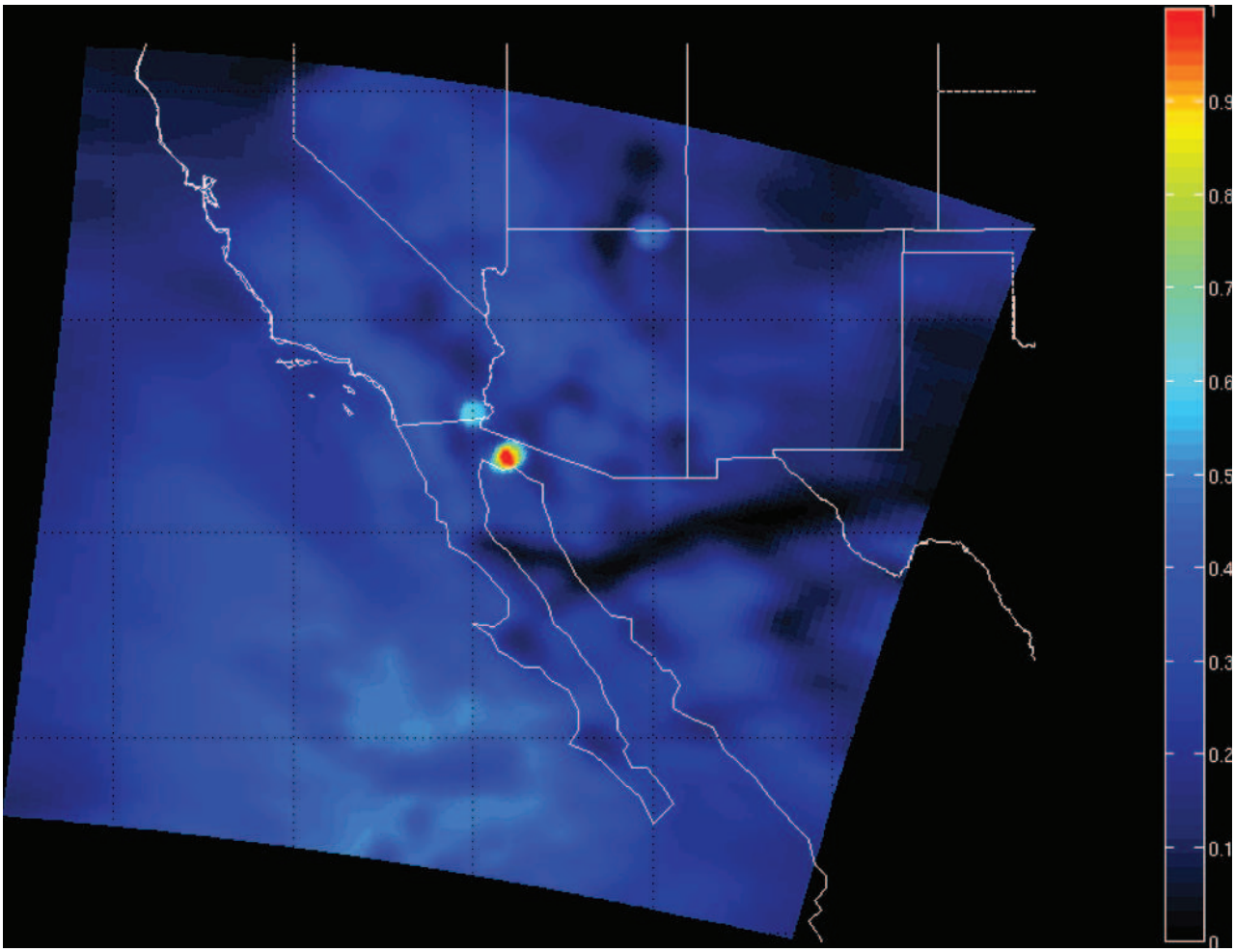

Figure 256. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

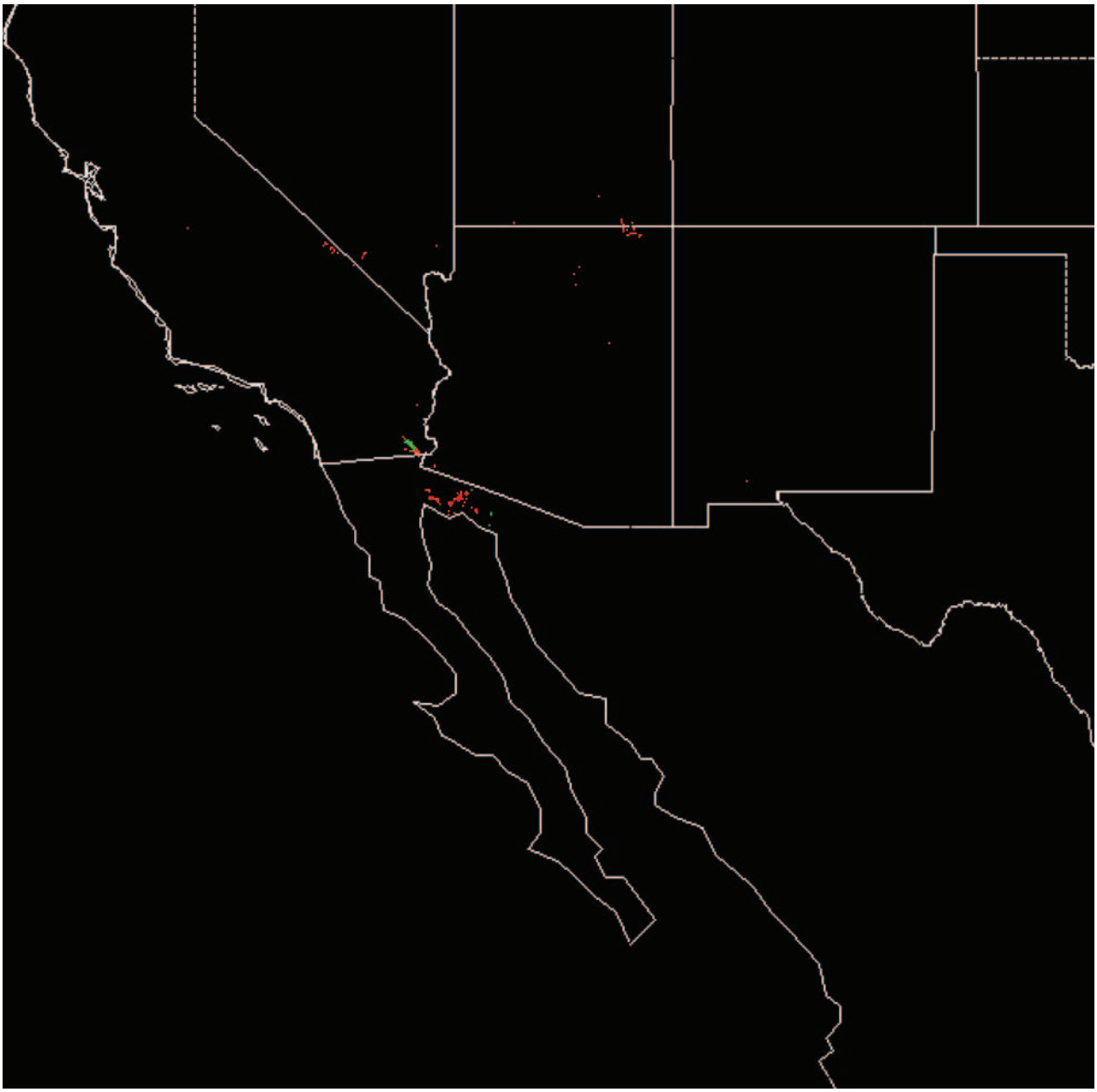

Figure 257. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2001-327 at 18:30 UTC.

Figure 258. Result of the Maximum a posteriori model (MAP) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

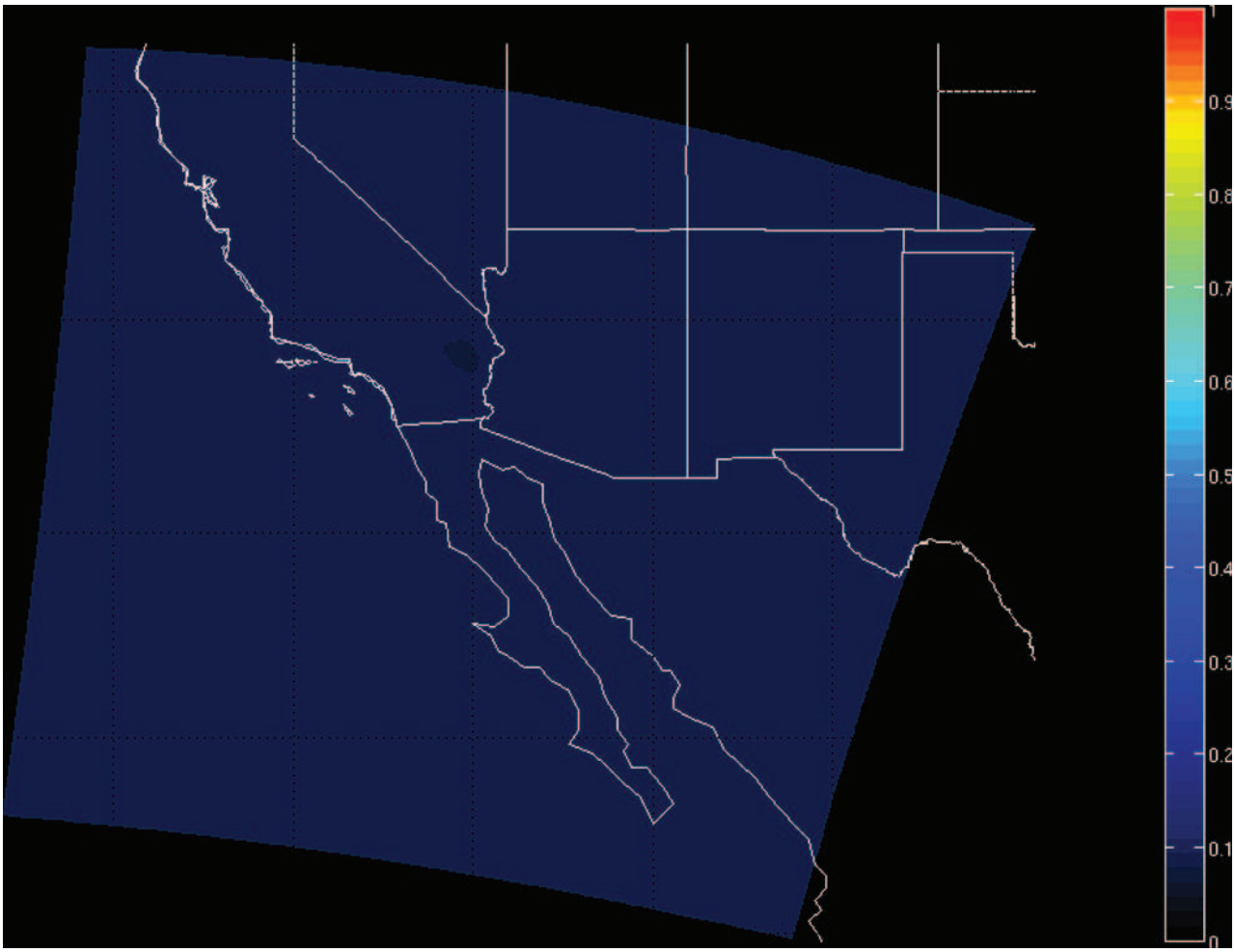

Figure 259. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

Figure 260. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2001-327 at 18:30 UTC.

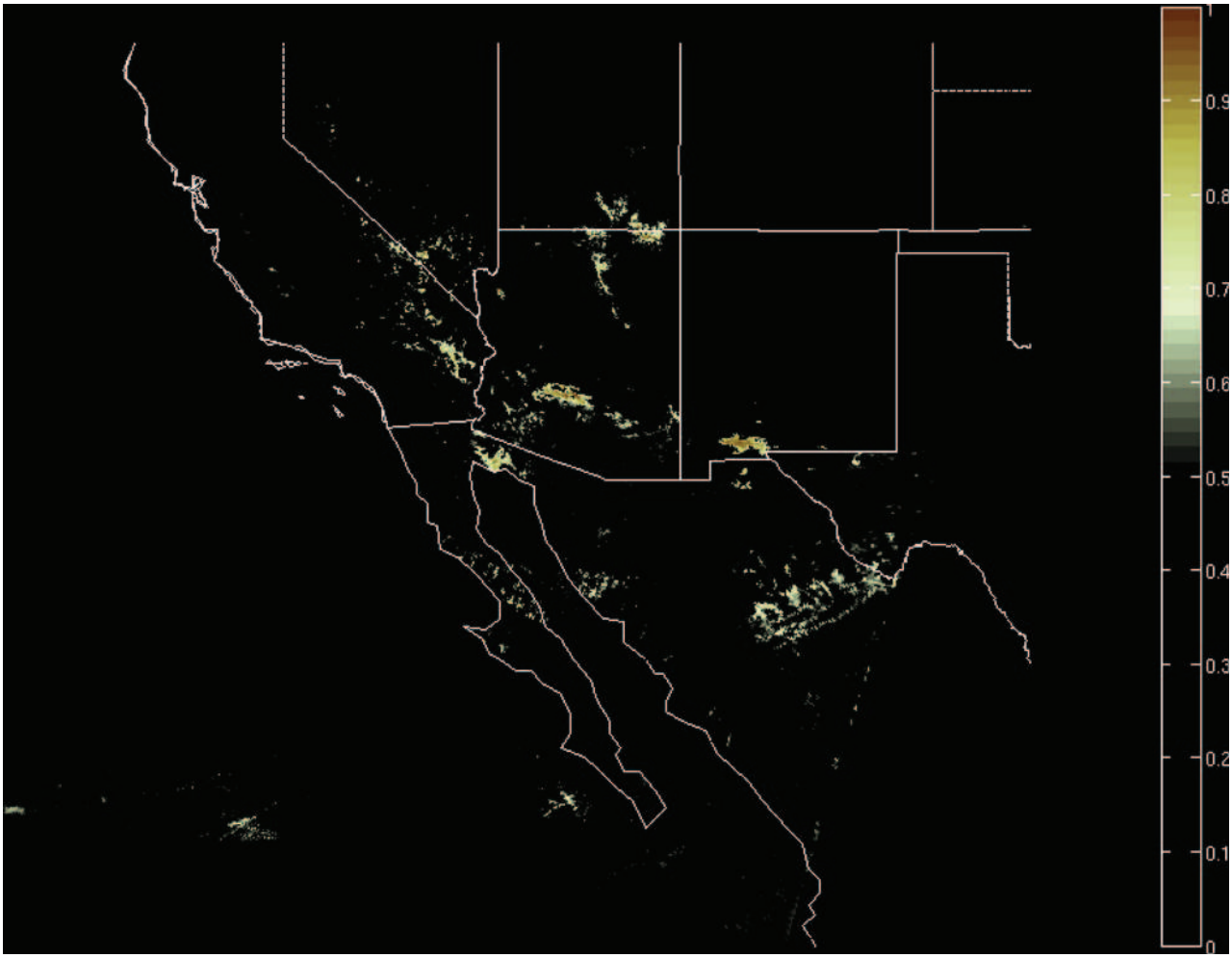

Figure 261. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

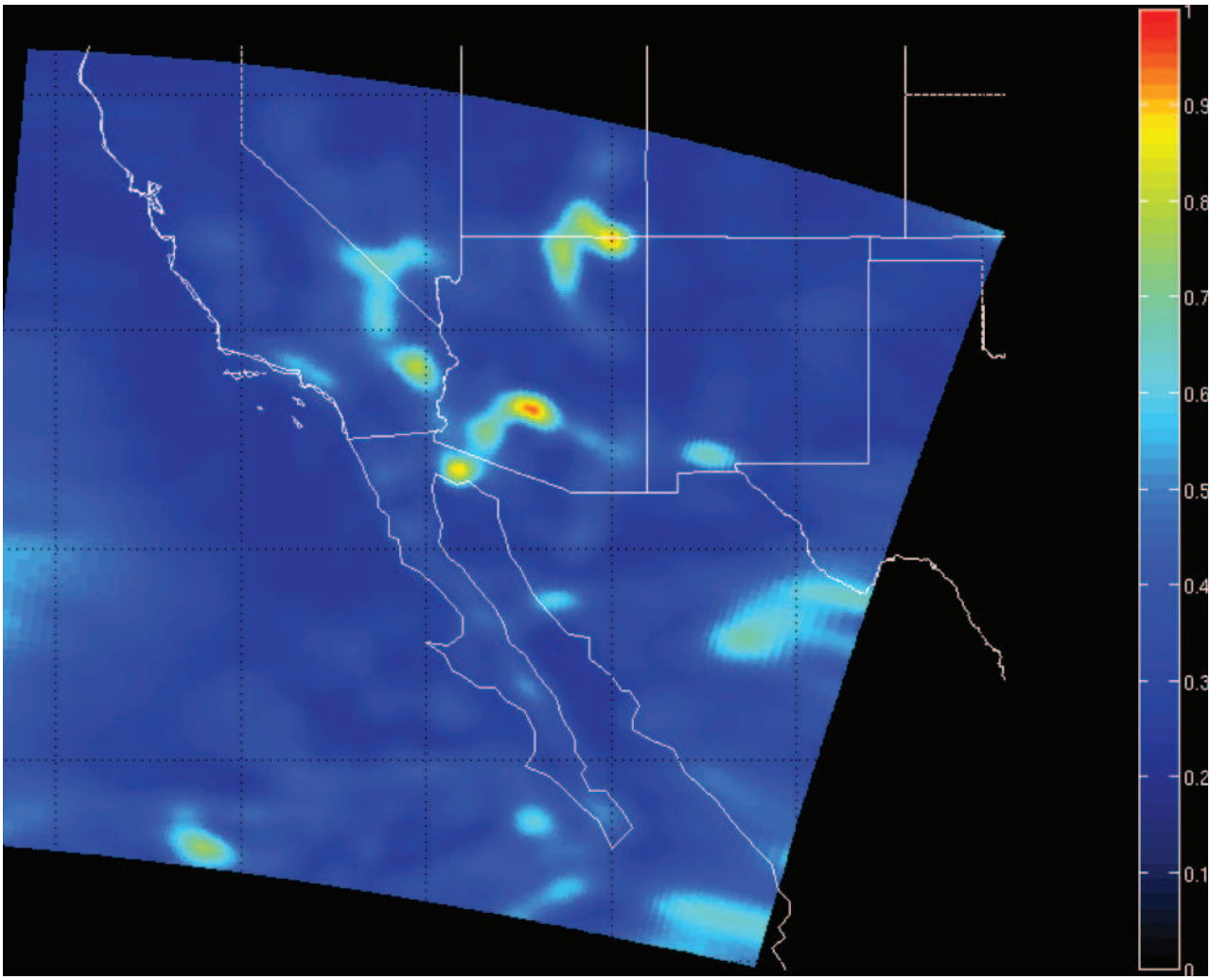

Figure 262. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

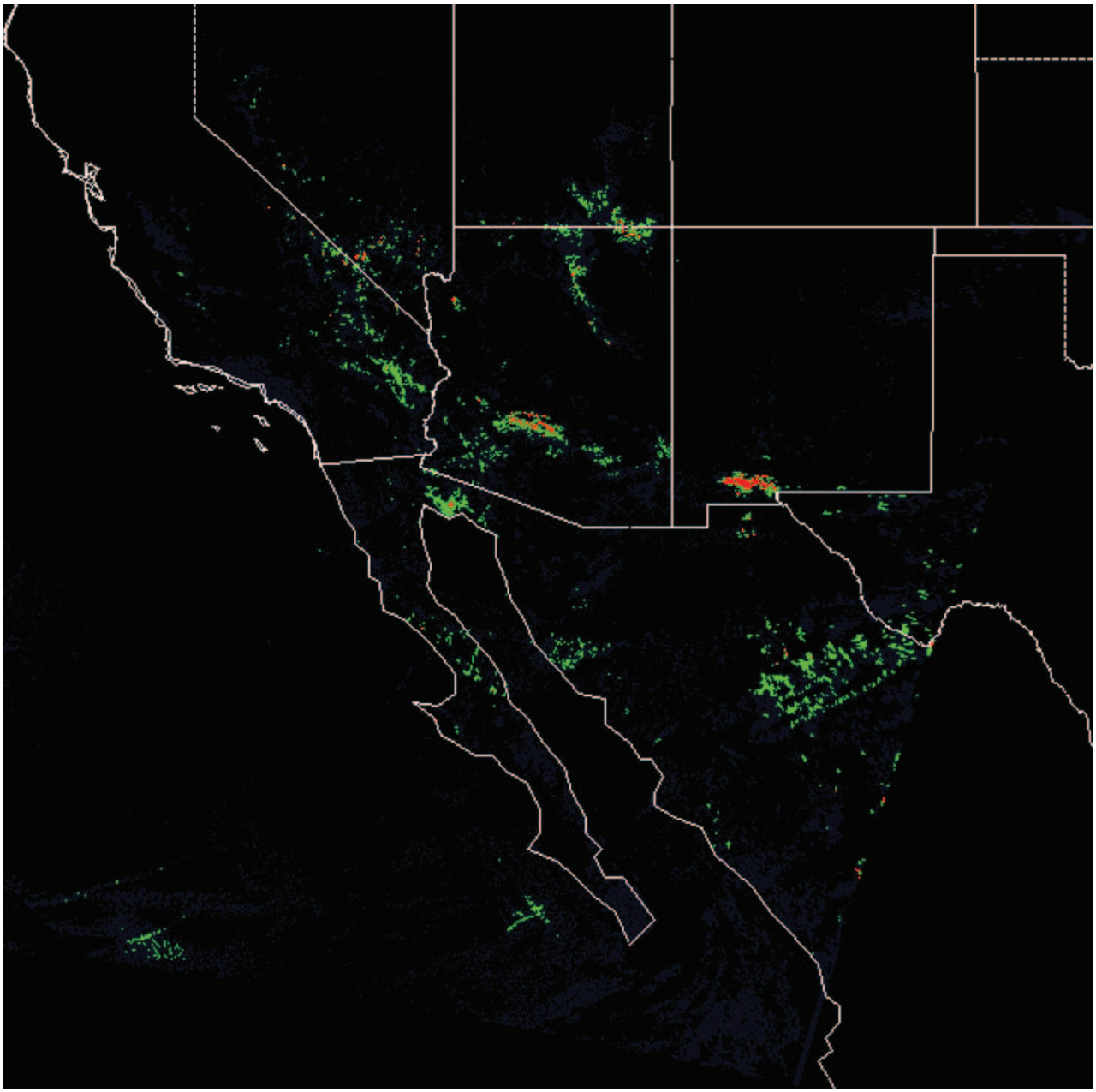

Figure 263. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2001-327 at 18:30 UTC.

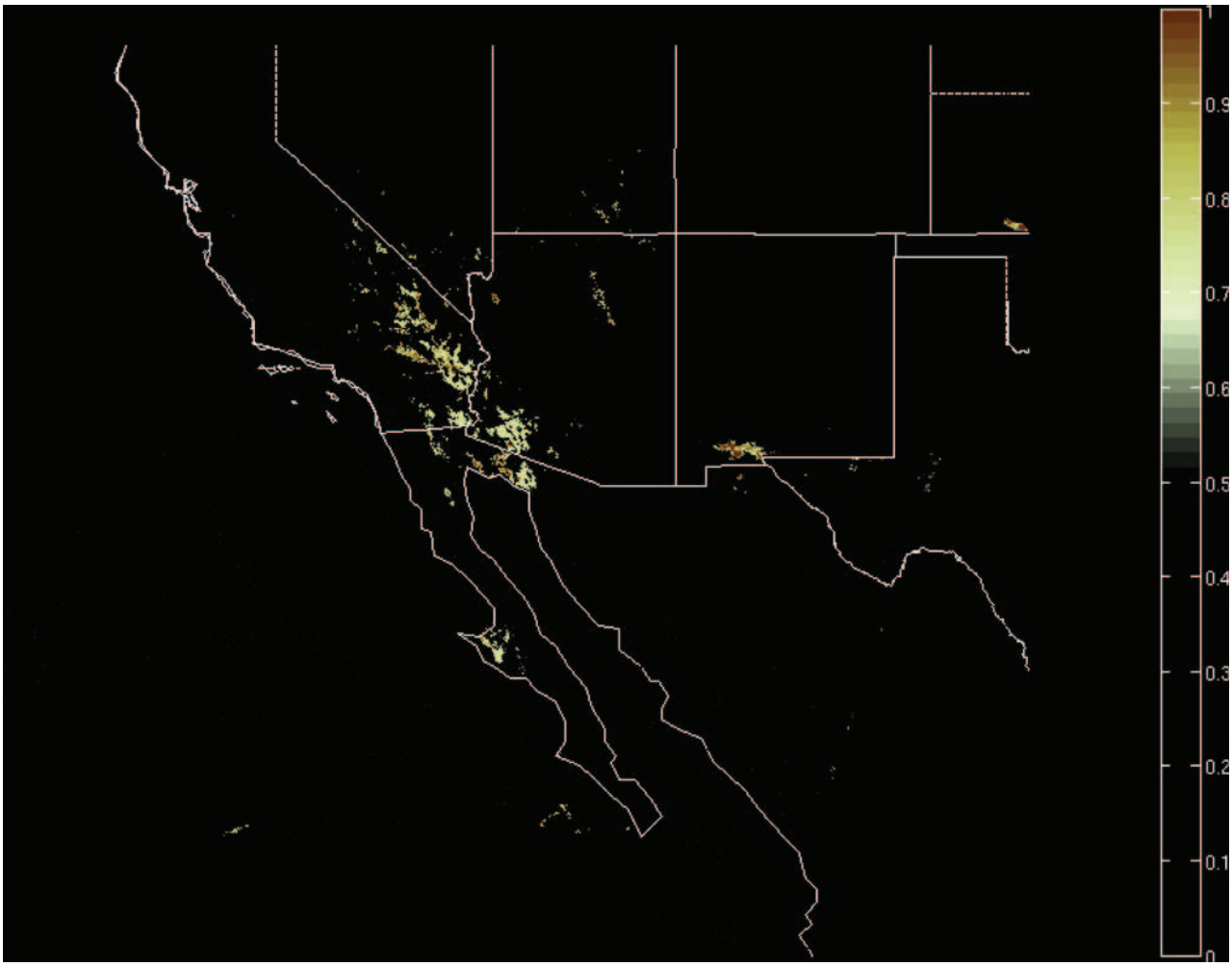

Figure 264. Result of the Probabilistic Neural Network (PNN) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

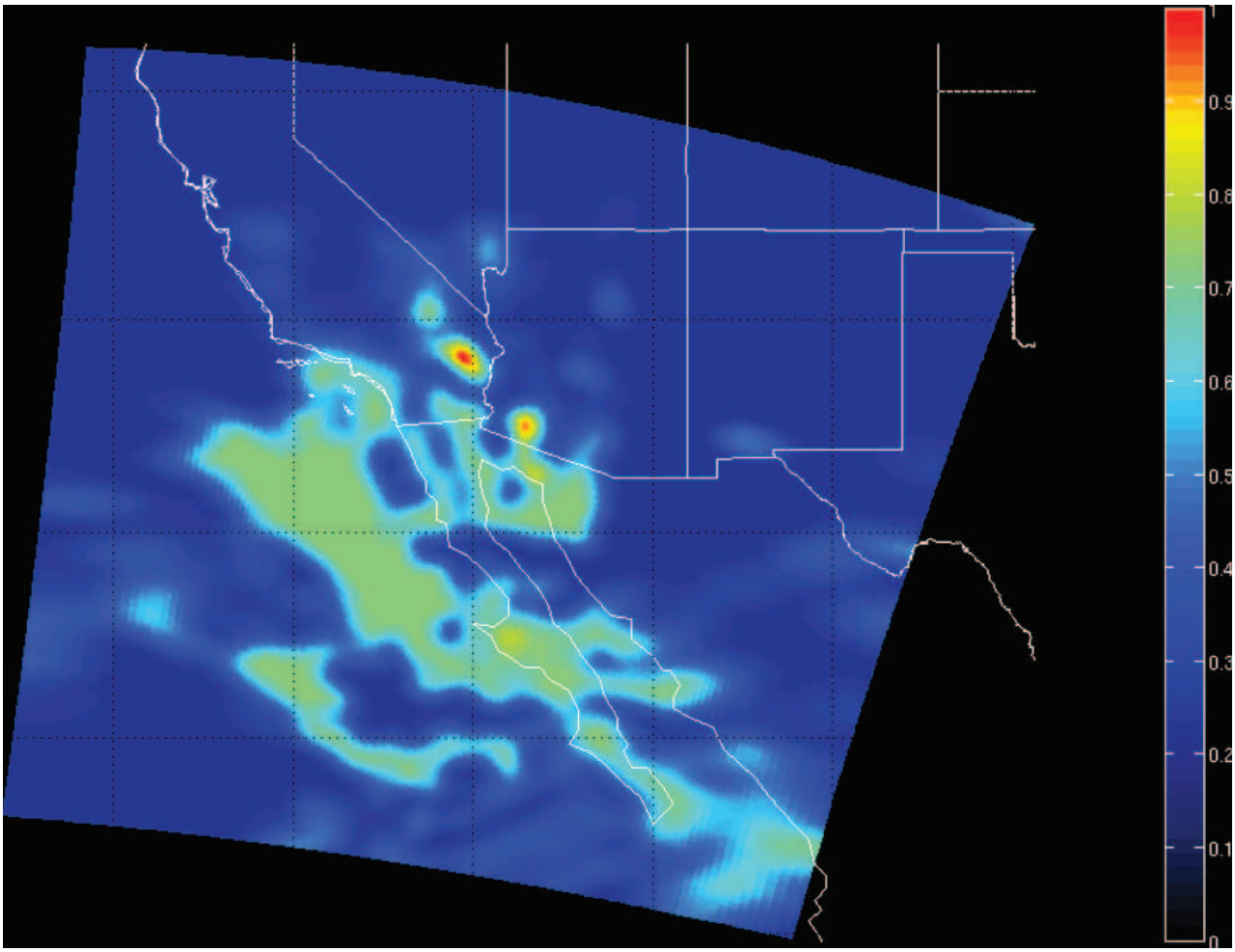

Figure 265. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2001-327 at 18:30 UTC. The colorbar shows the index of probability.

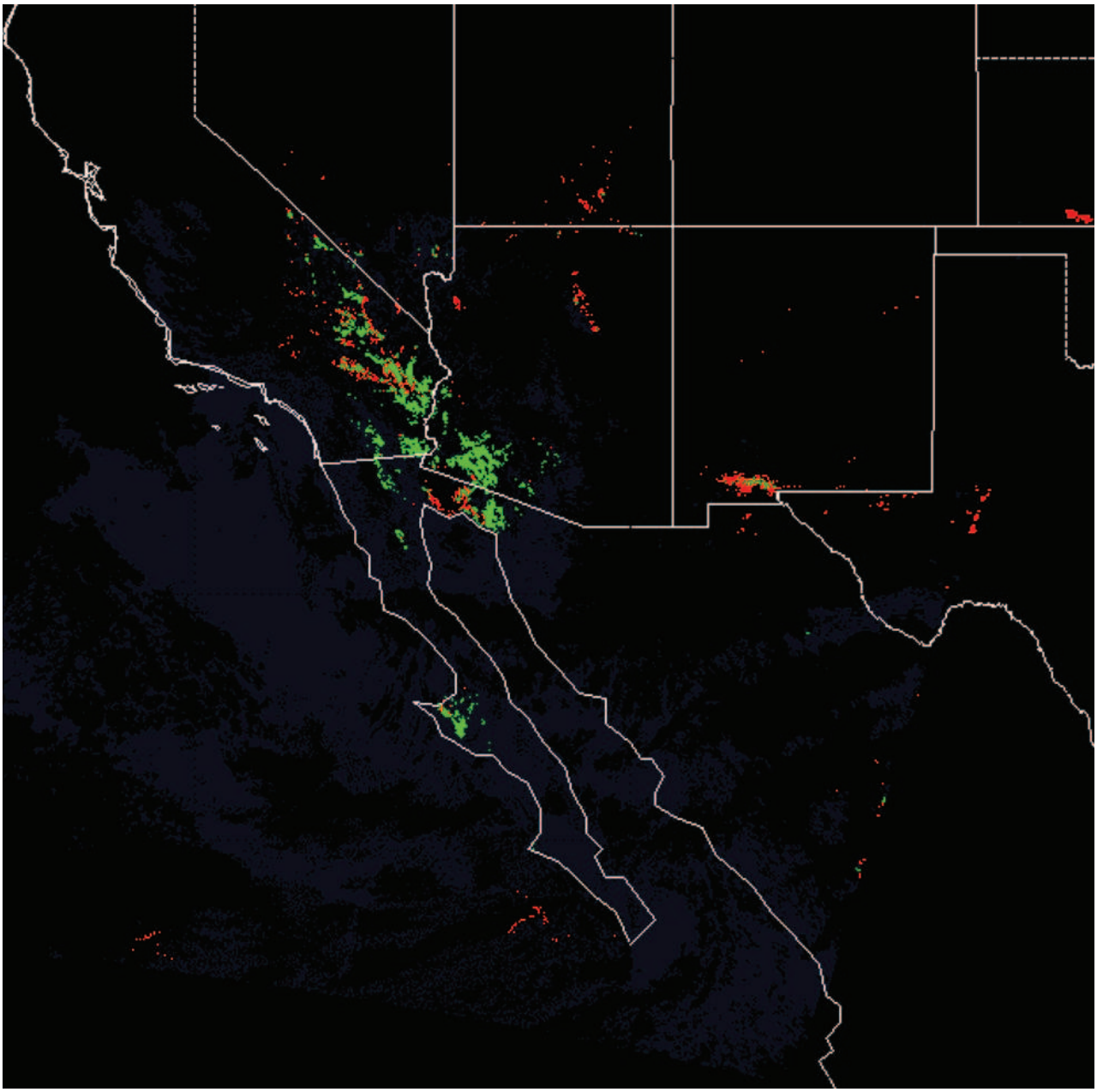

Figure 266. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2001-327 at 18:30 UTC.

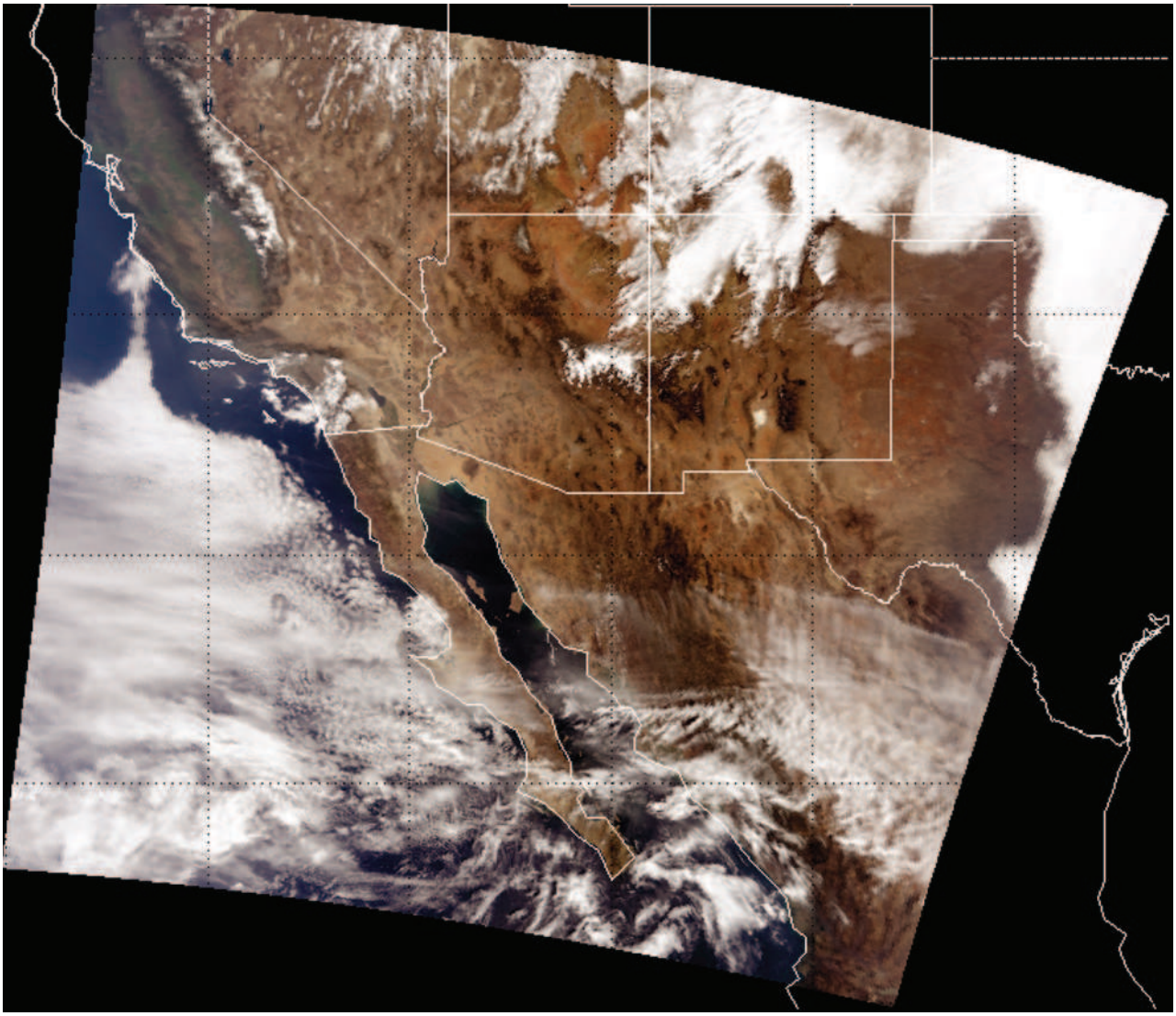

Figure 267. True color image for the event of 2002-060 at 18:15 UTC.

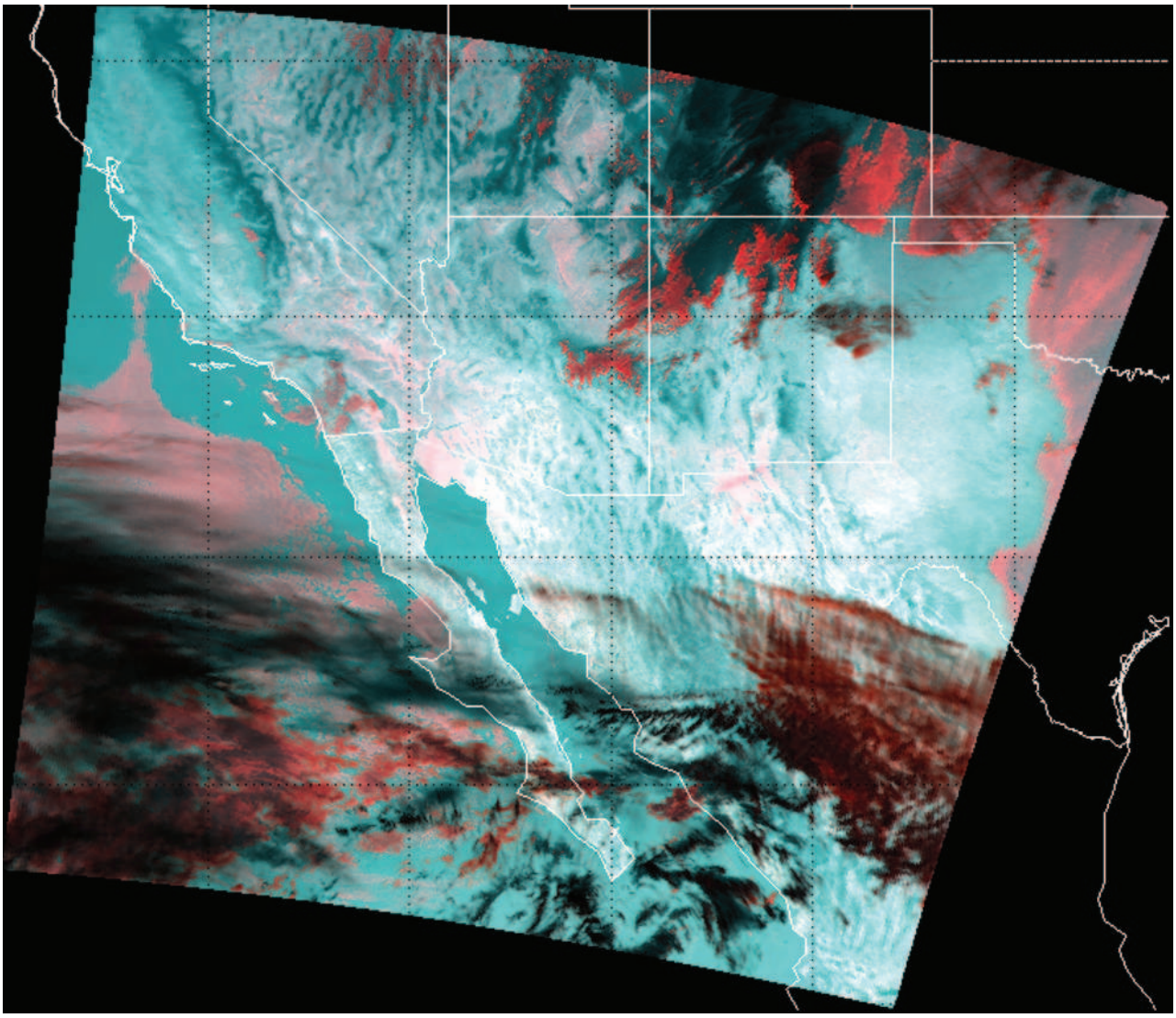

Figure 268. False color image for the event of 2002-060 at 18:15 UTC.

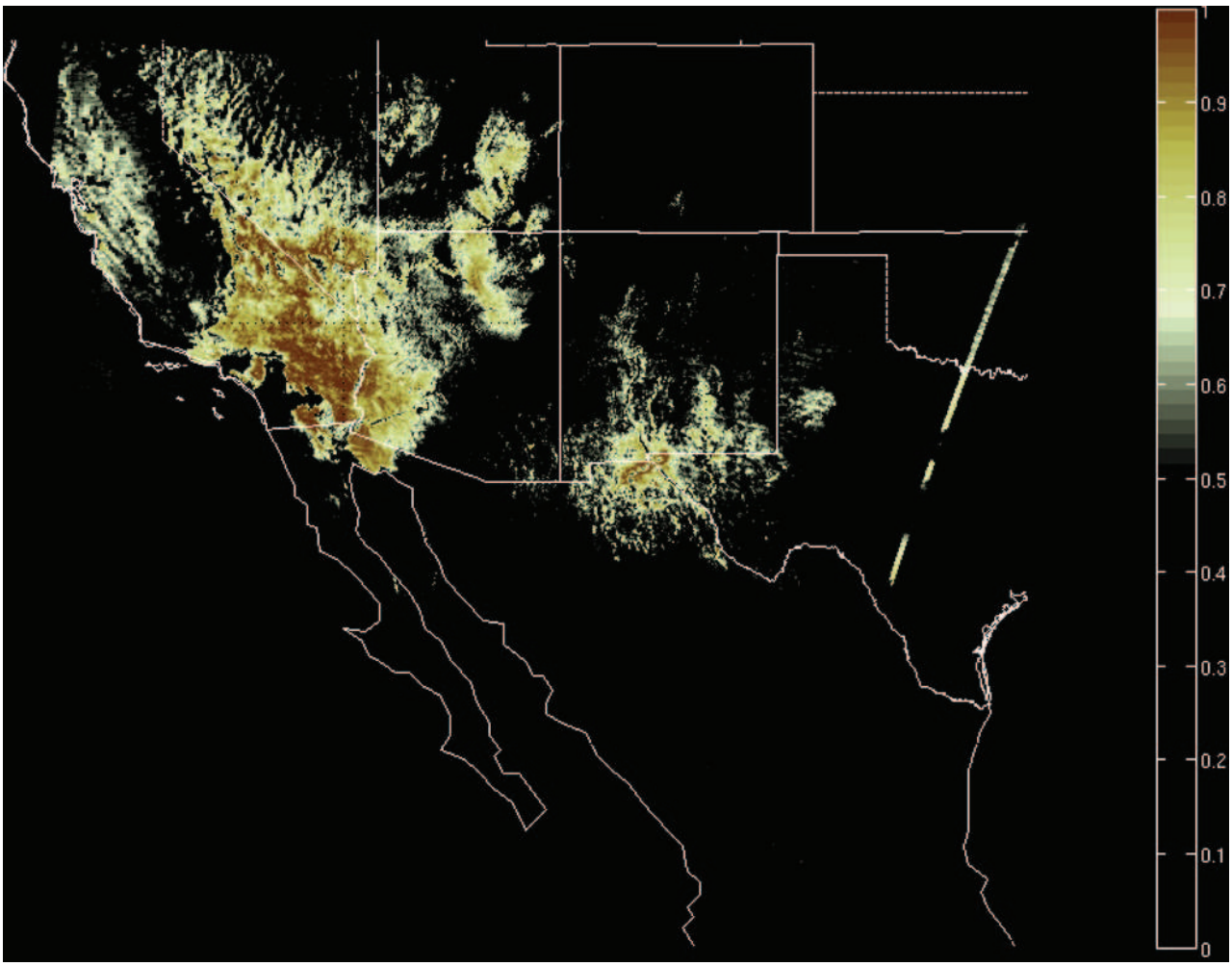

Figure 269. Result of the probabilistic display model (PD) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

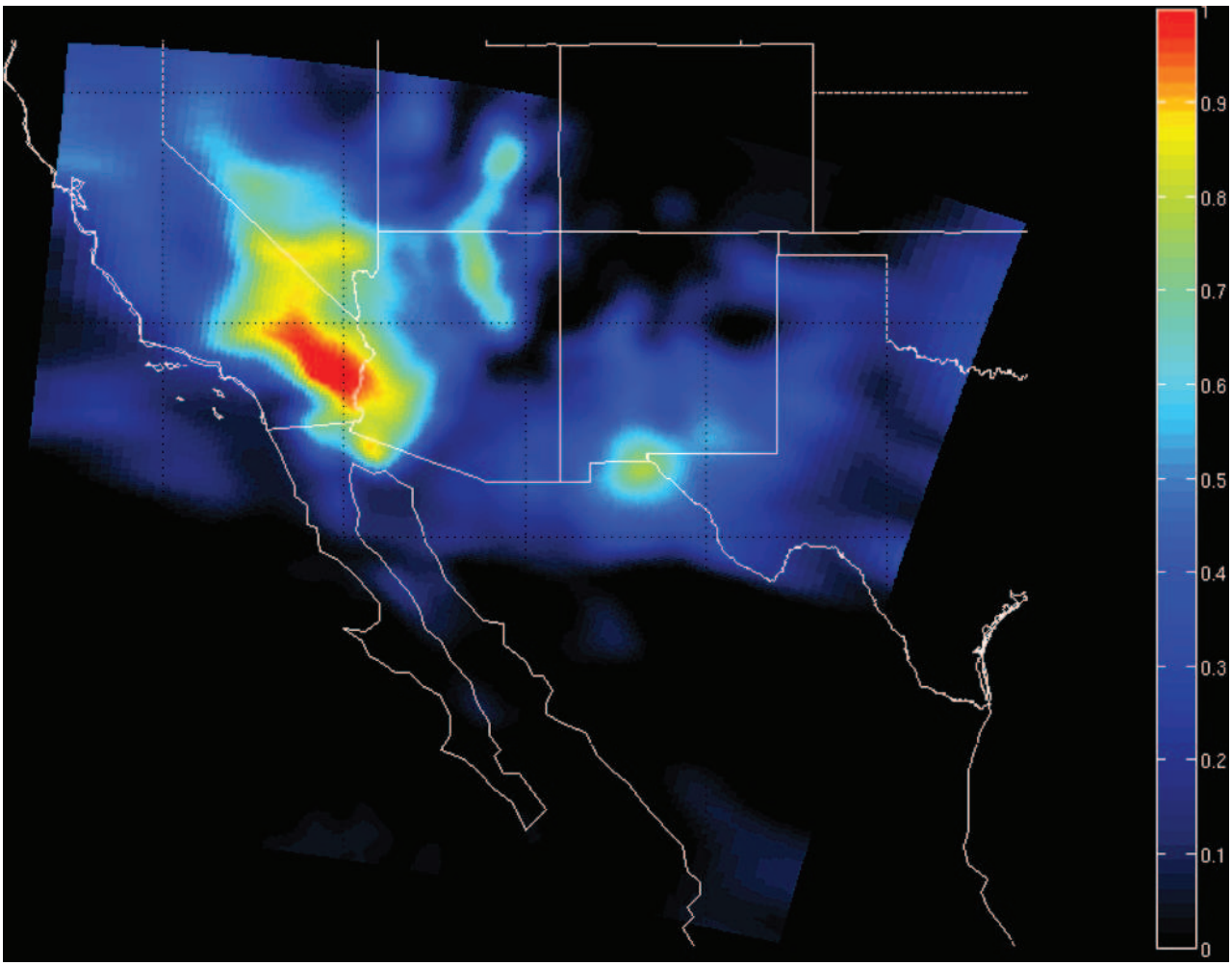

Figure 270. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

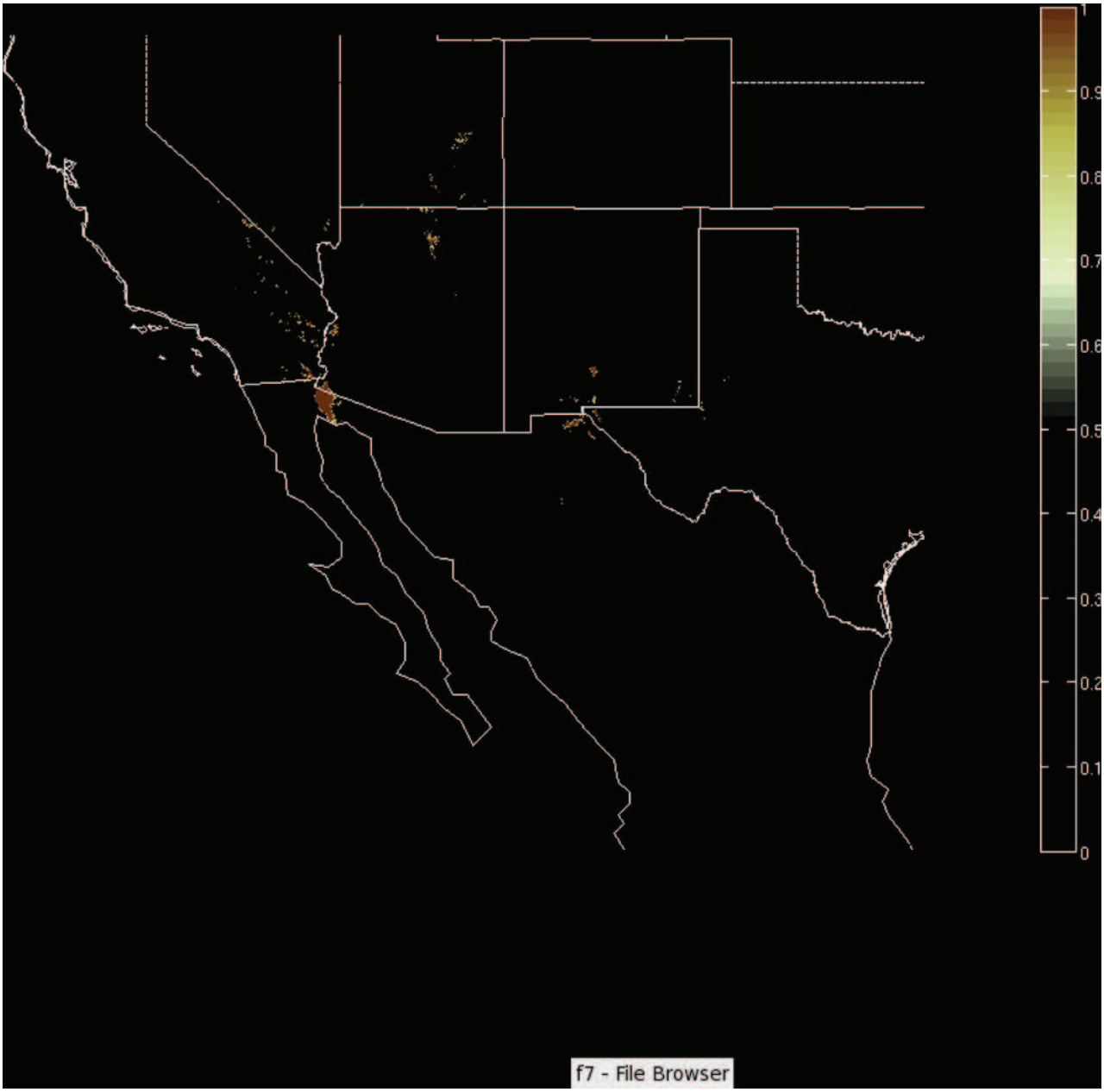

Figure 271. Result of the Maximum Likelihood model (MLE) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

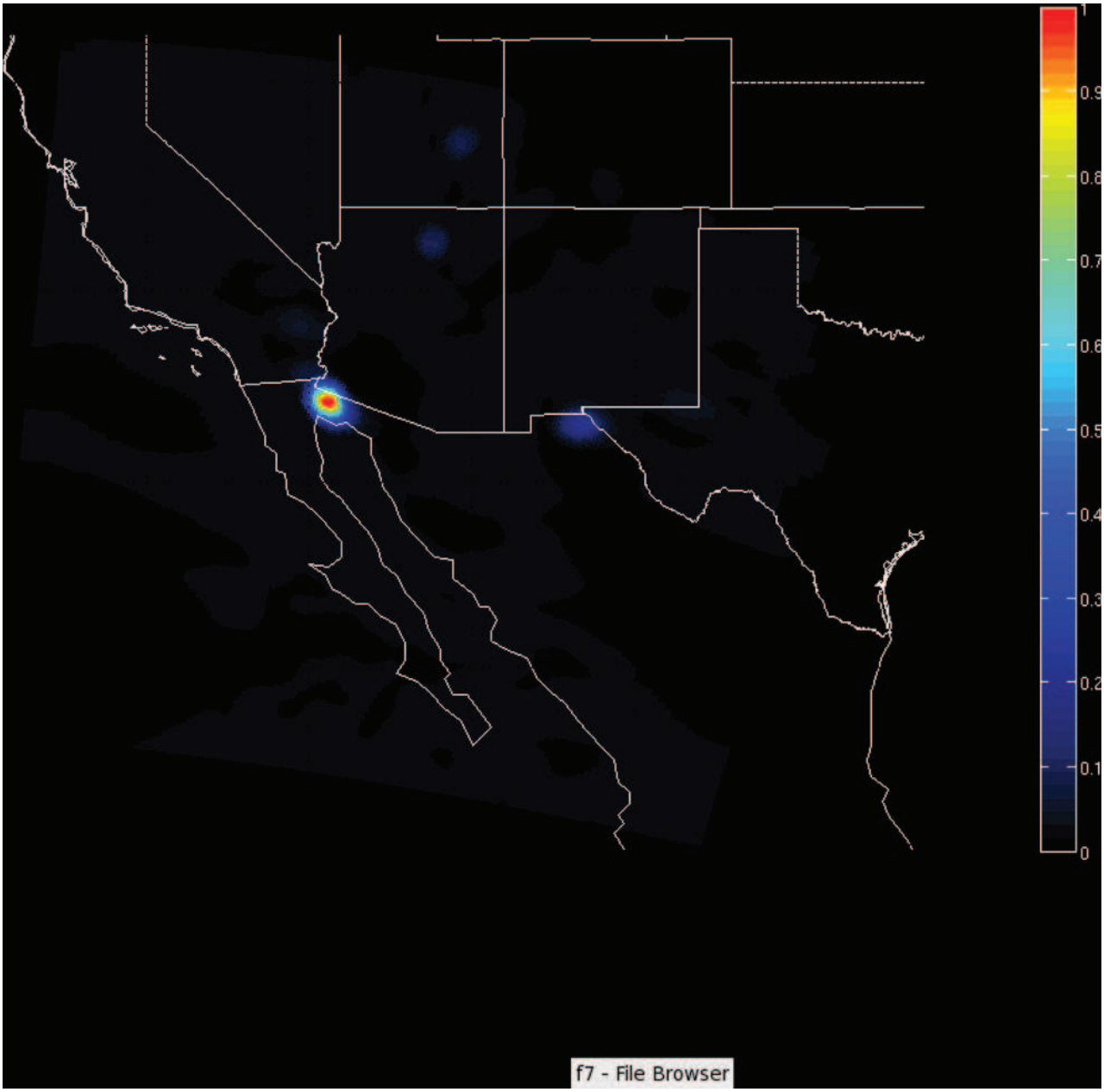

Figure 272. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

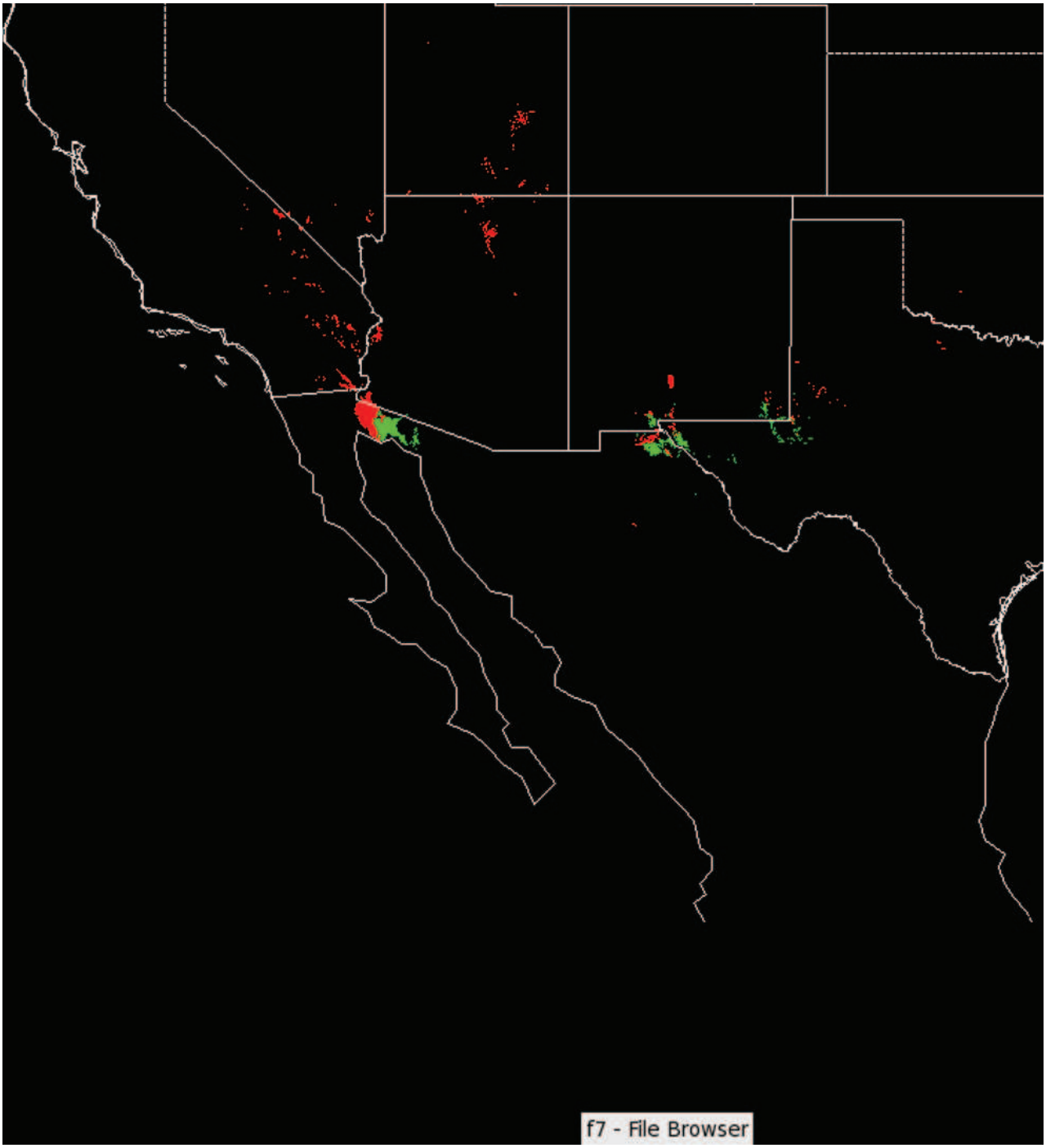

Figure 273. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2002-060 at 18:15 UTC.

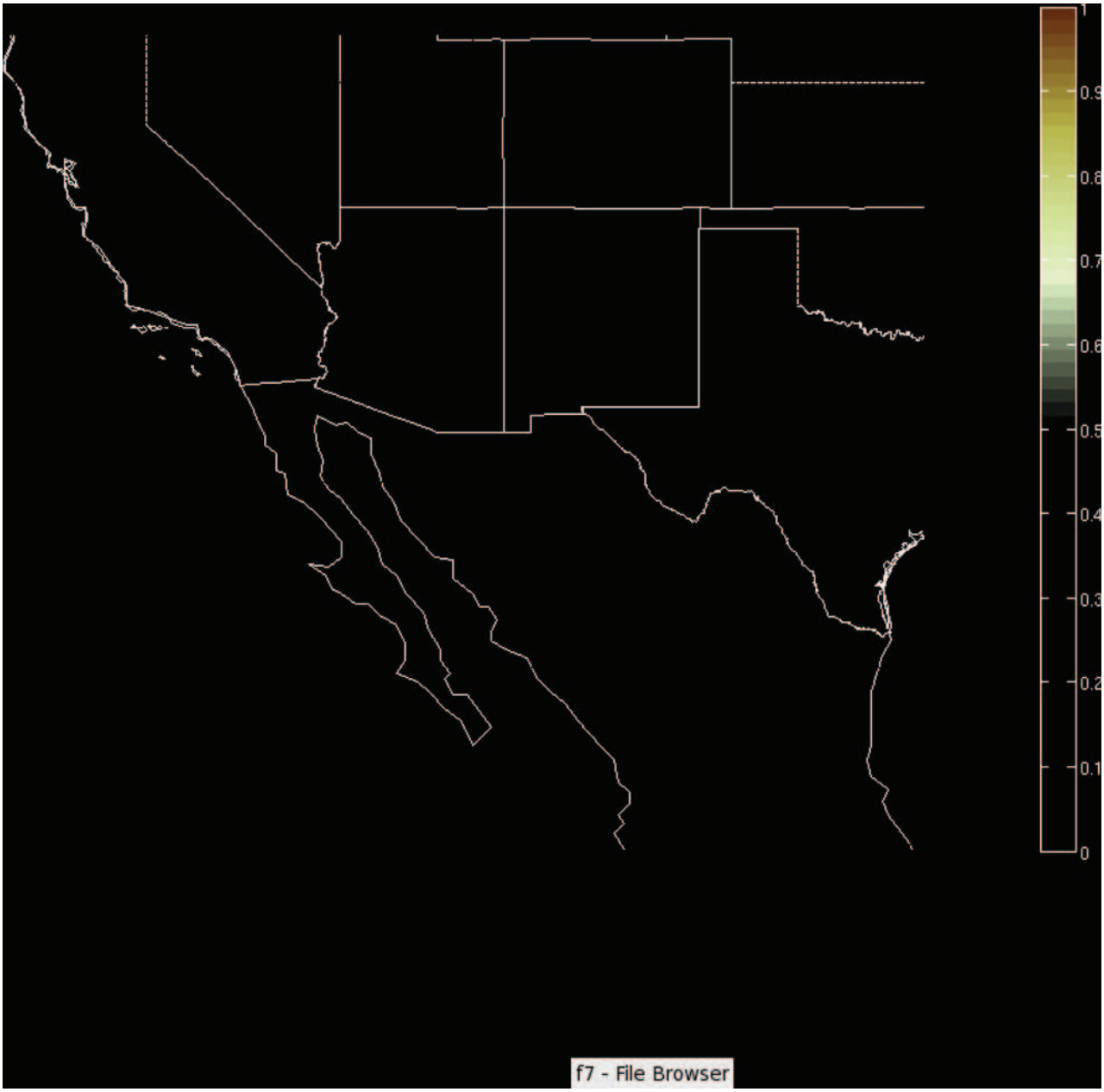

Figure 274. Result of the Maximum a posteriori model (MAP) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

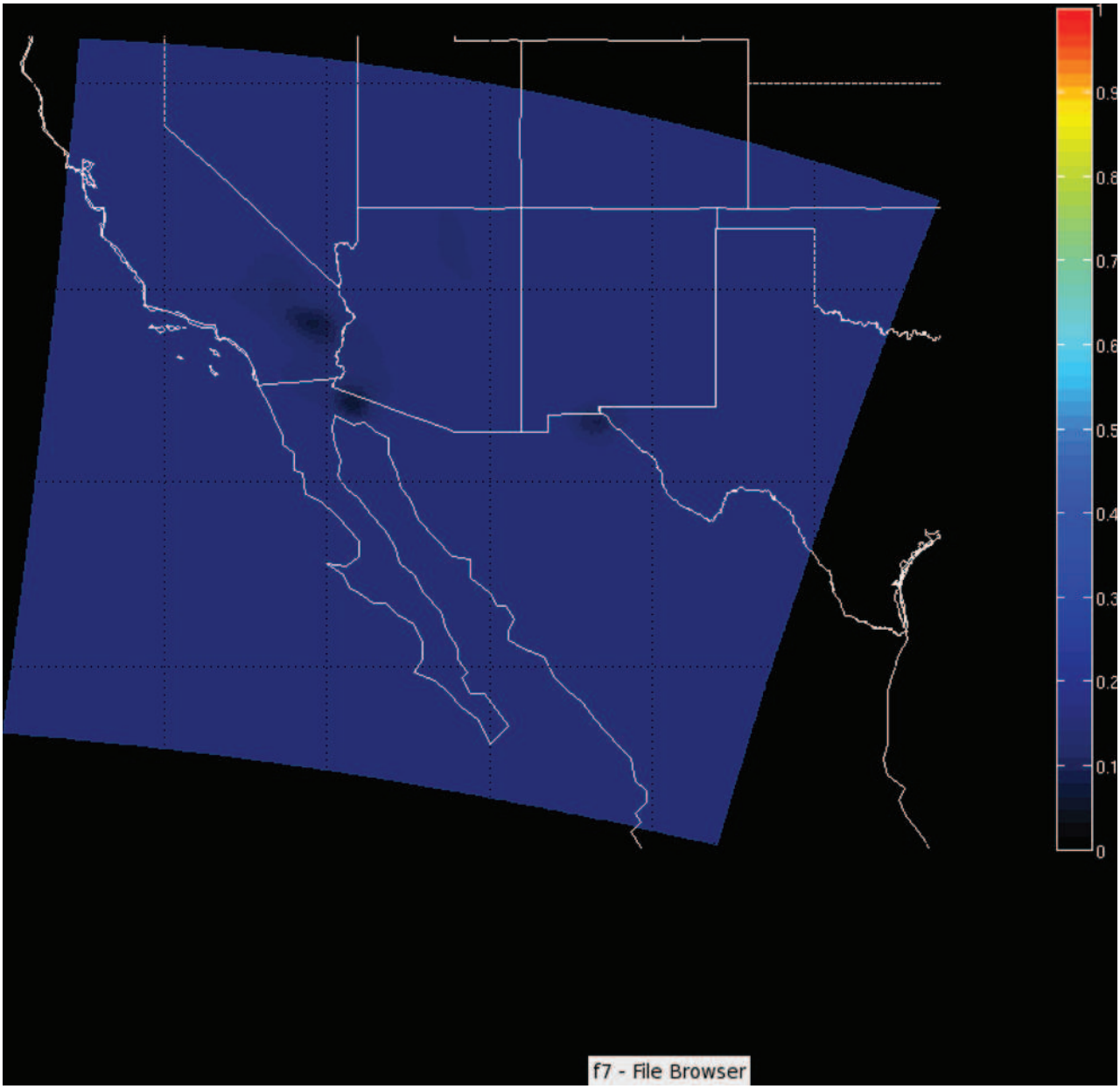

Figure 275. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

Figure 276. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2002-060 at 18:15 UTC.

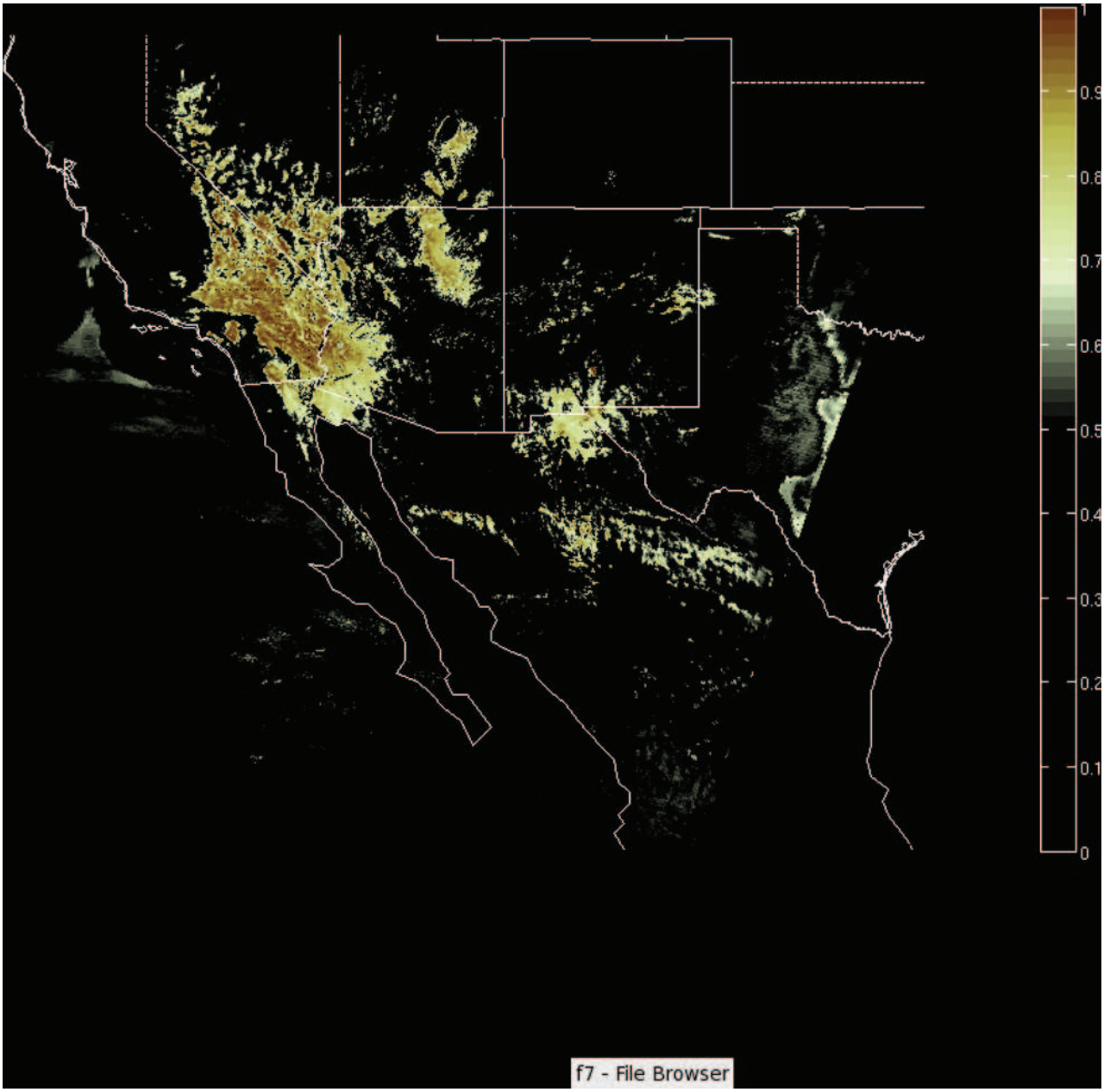

Figure 277. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

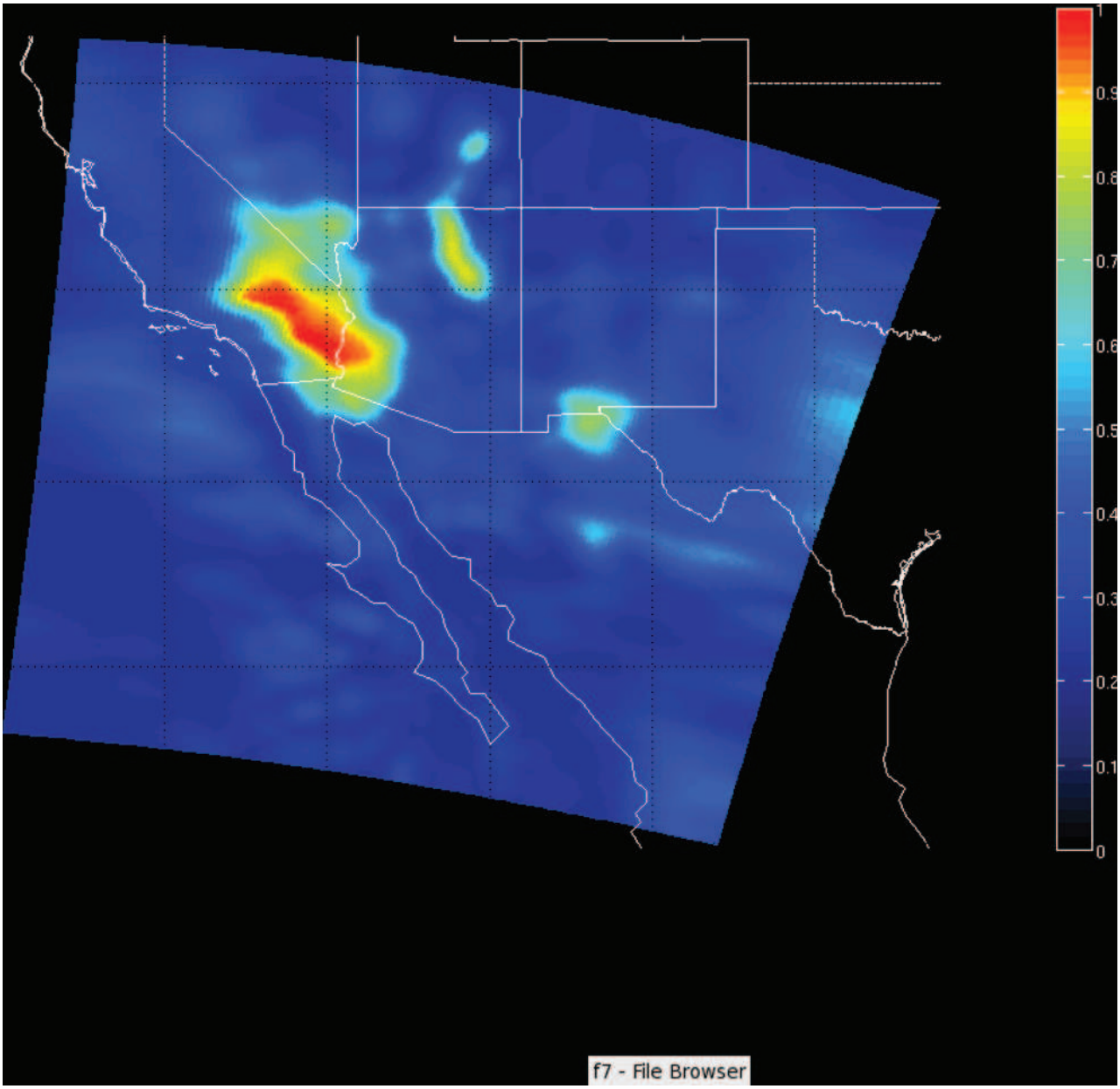

Figure 278. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

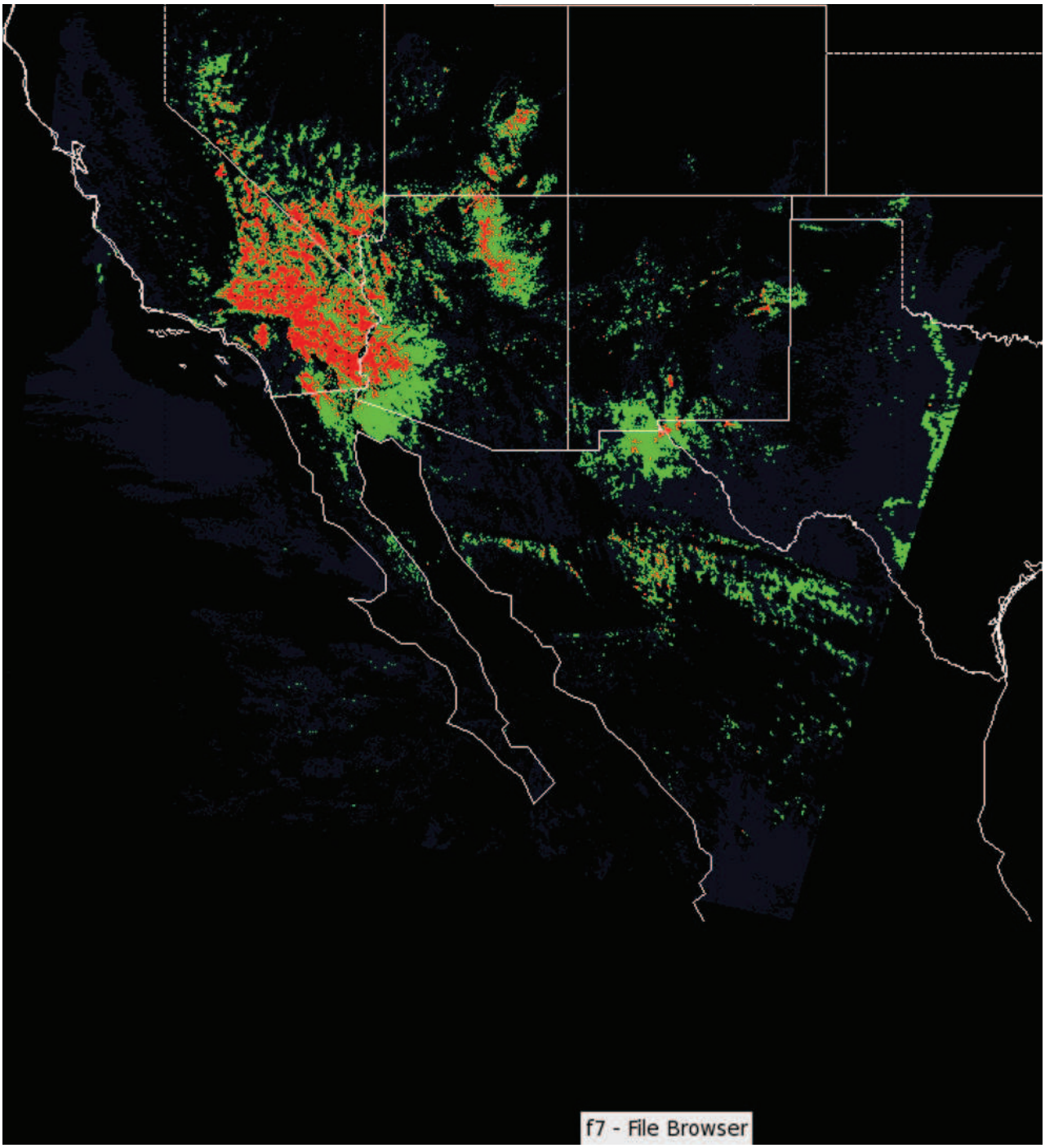

Figure 279. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-060 at 18:15 UTC.

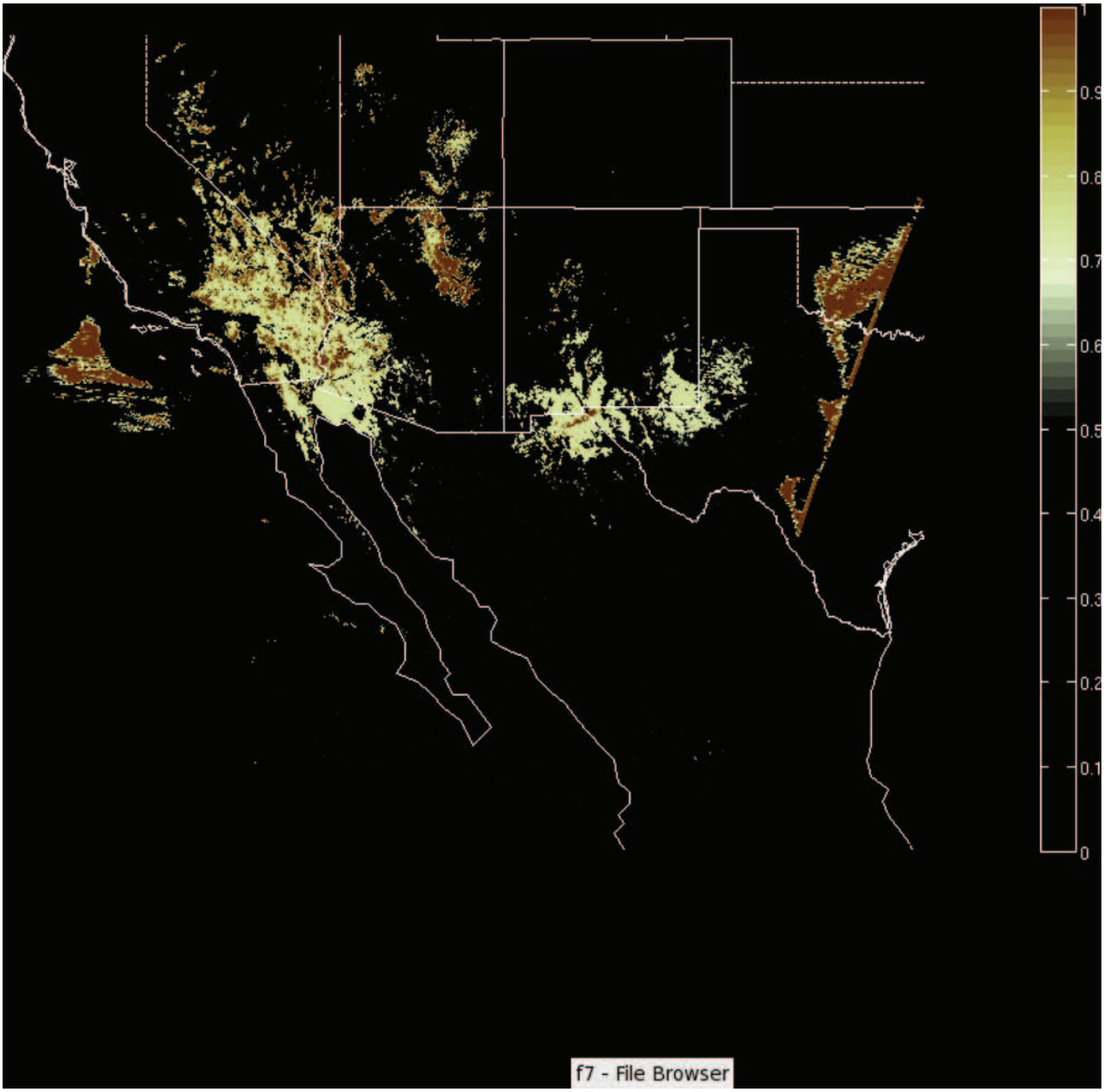

Figure 280. Result of the Probabilistic Neural Network (PNN) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

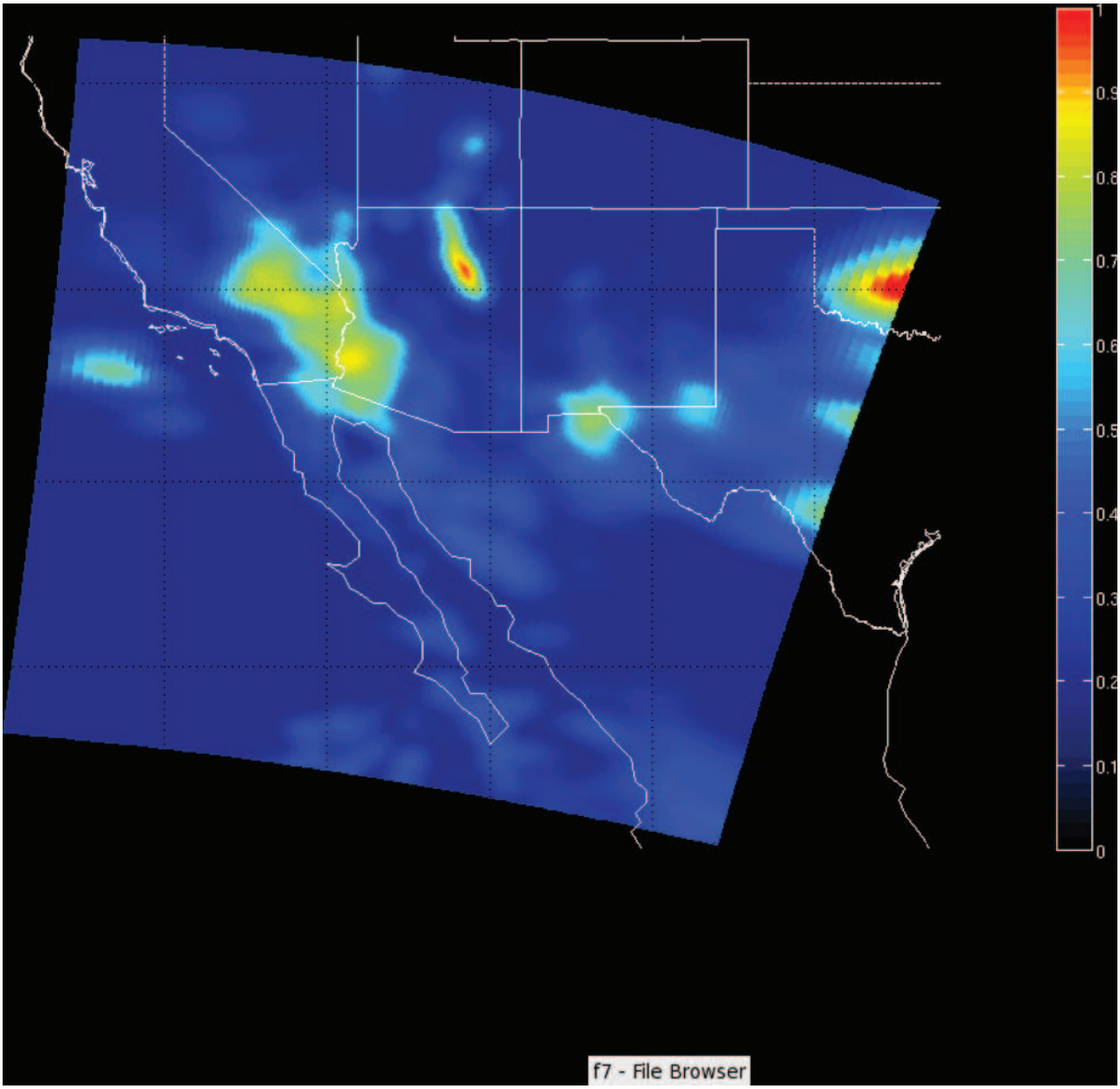

Figure 281. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2002-060 at 18:15 UTC. The colorbar shows the index of probability.

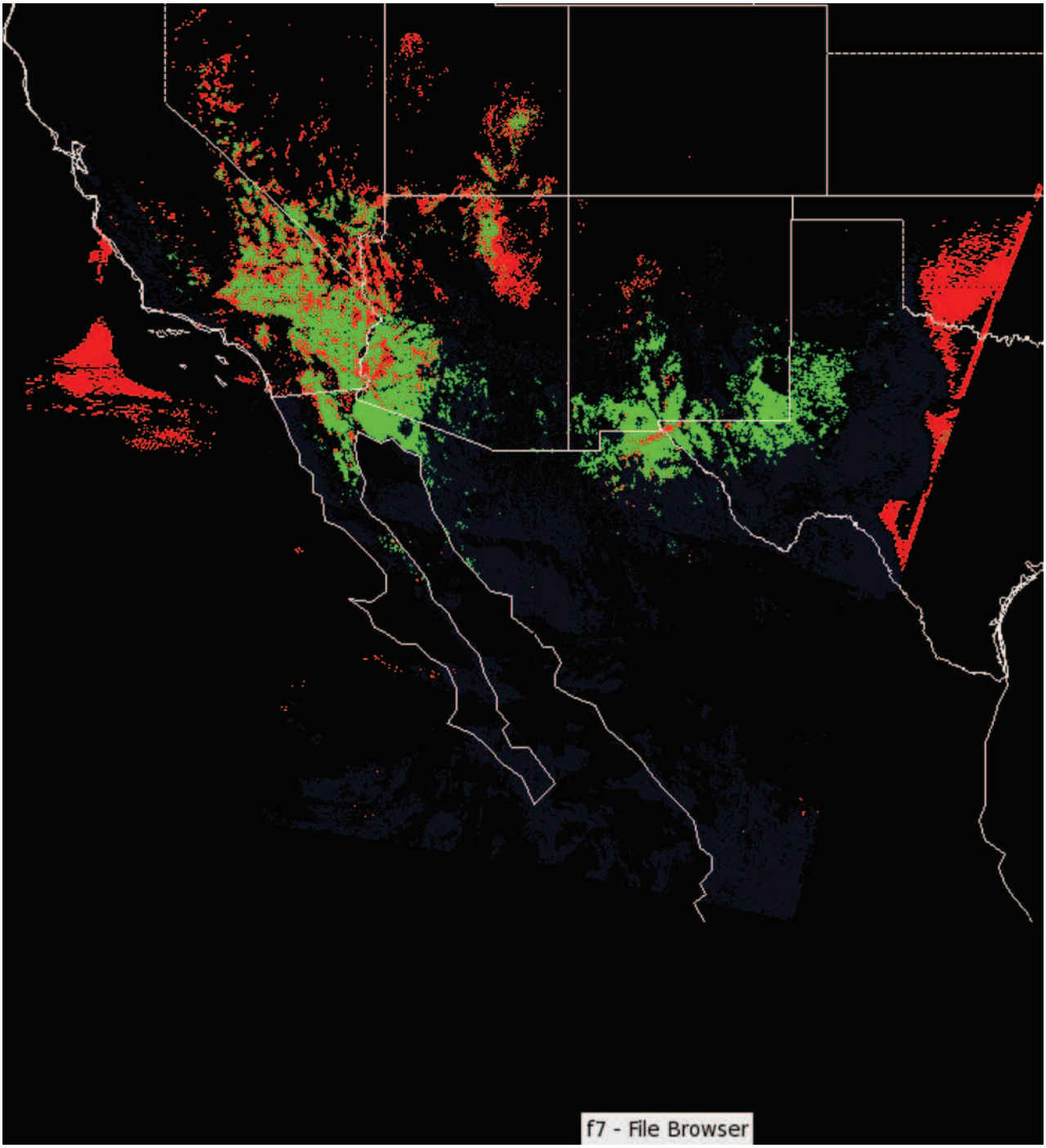

Figure 282. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2002-060 at 18:15 UTC.

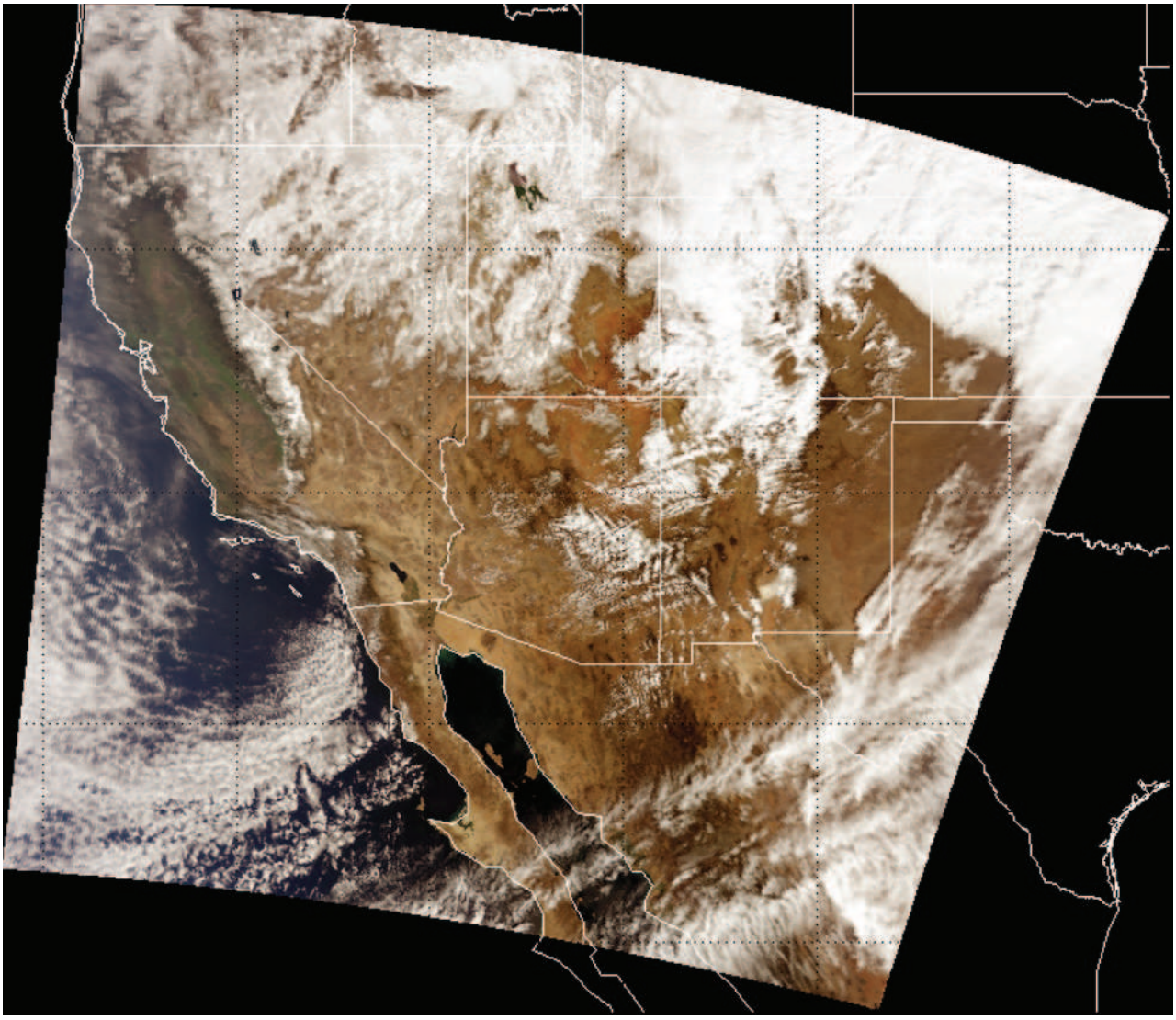

Figure 283. True color image for the event of 2002-067 at 18:20 UTC.

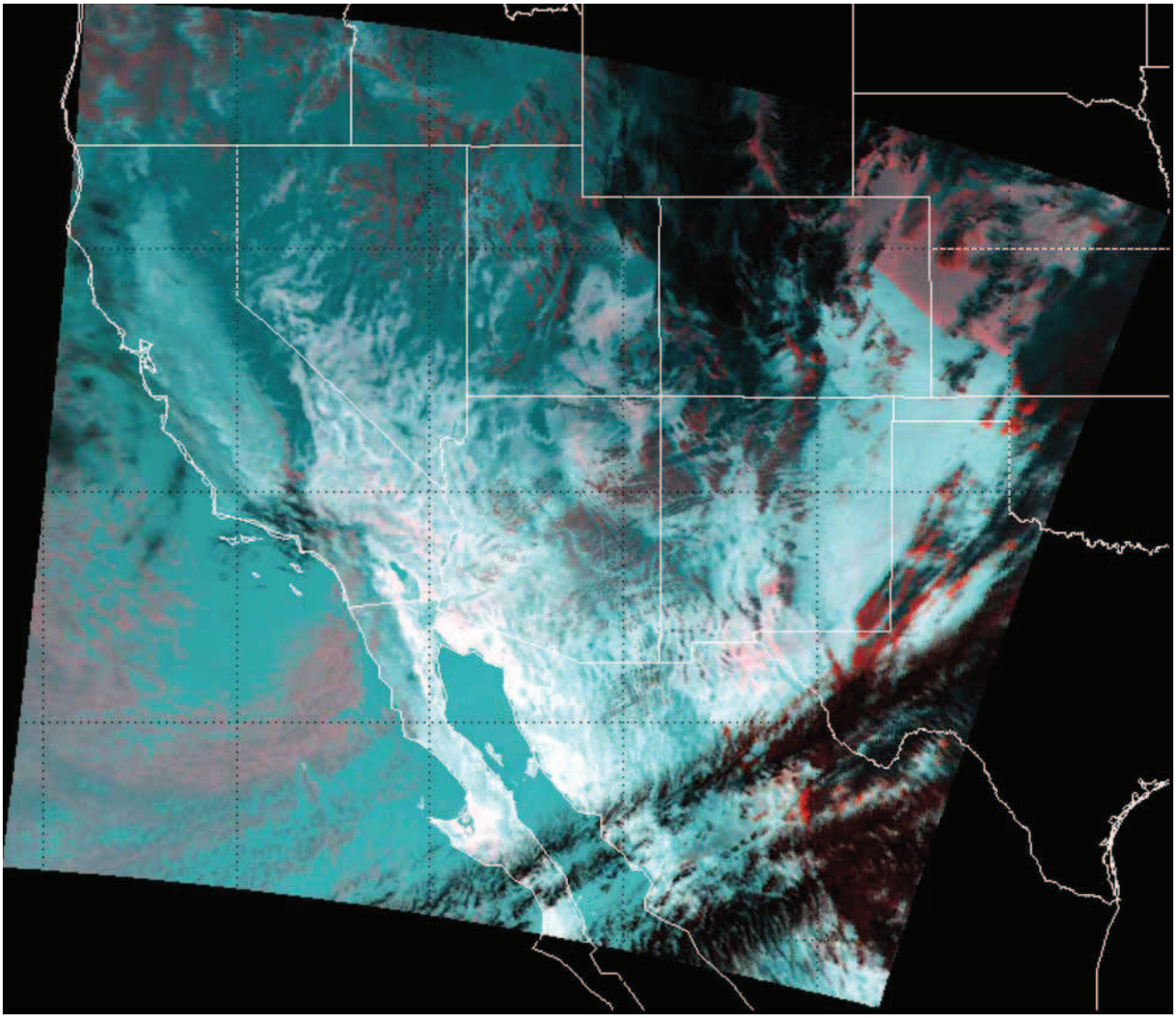

Figure 284. False color image for the event of 2002-067 at 18:20 UTC.

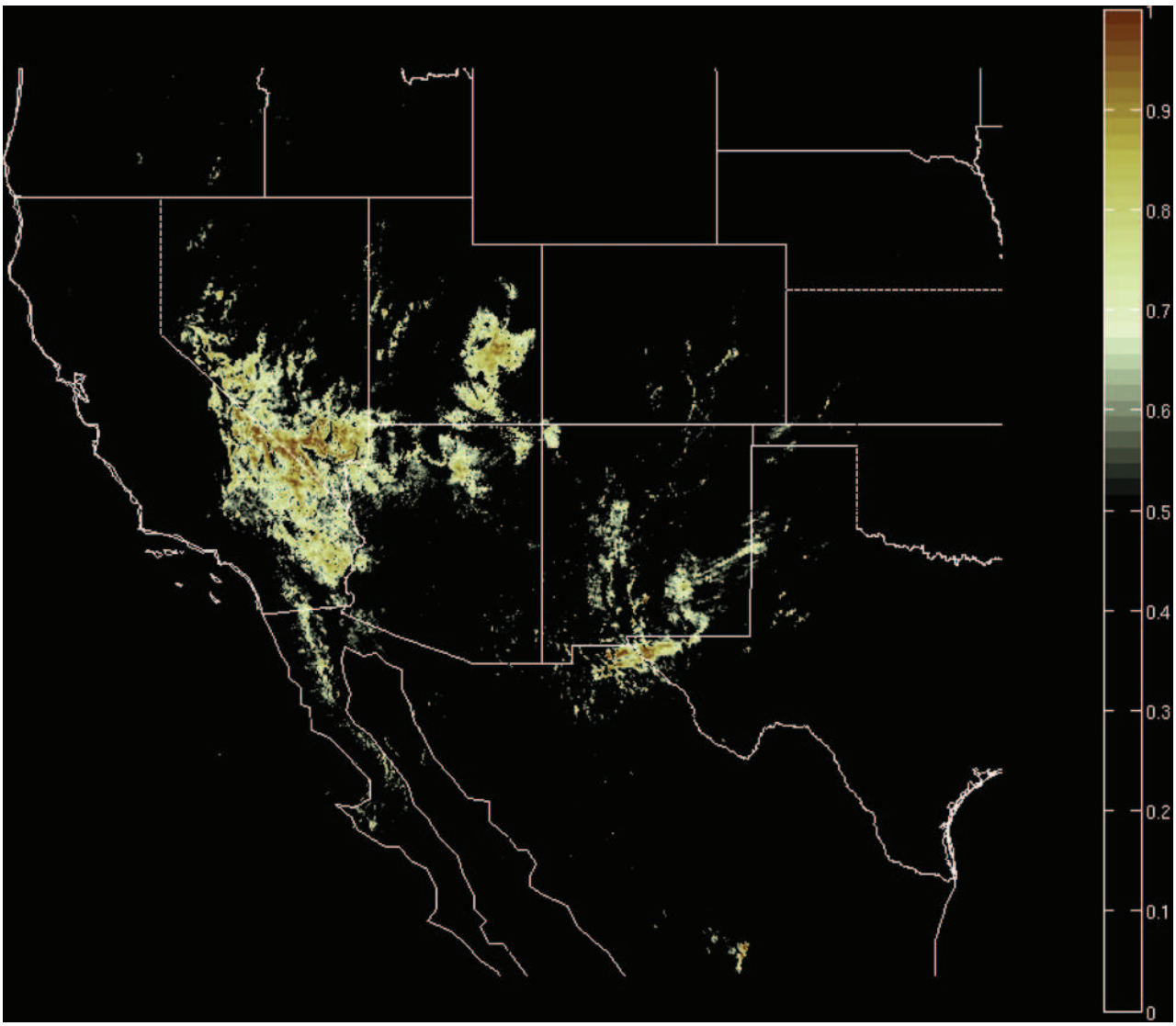

Figure 285. Result of the probabilistic display model (PD) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

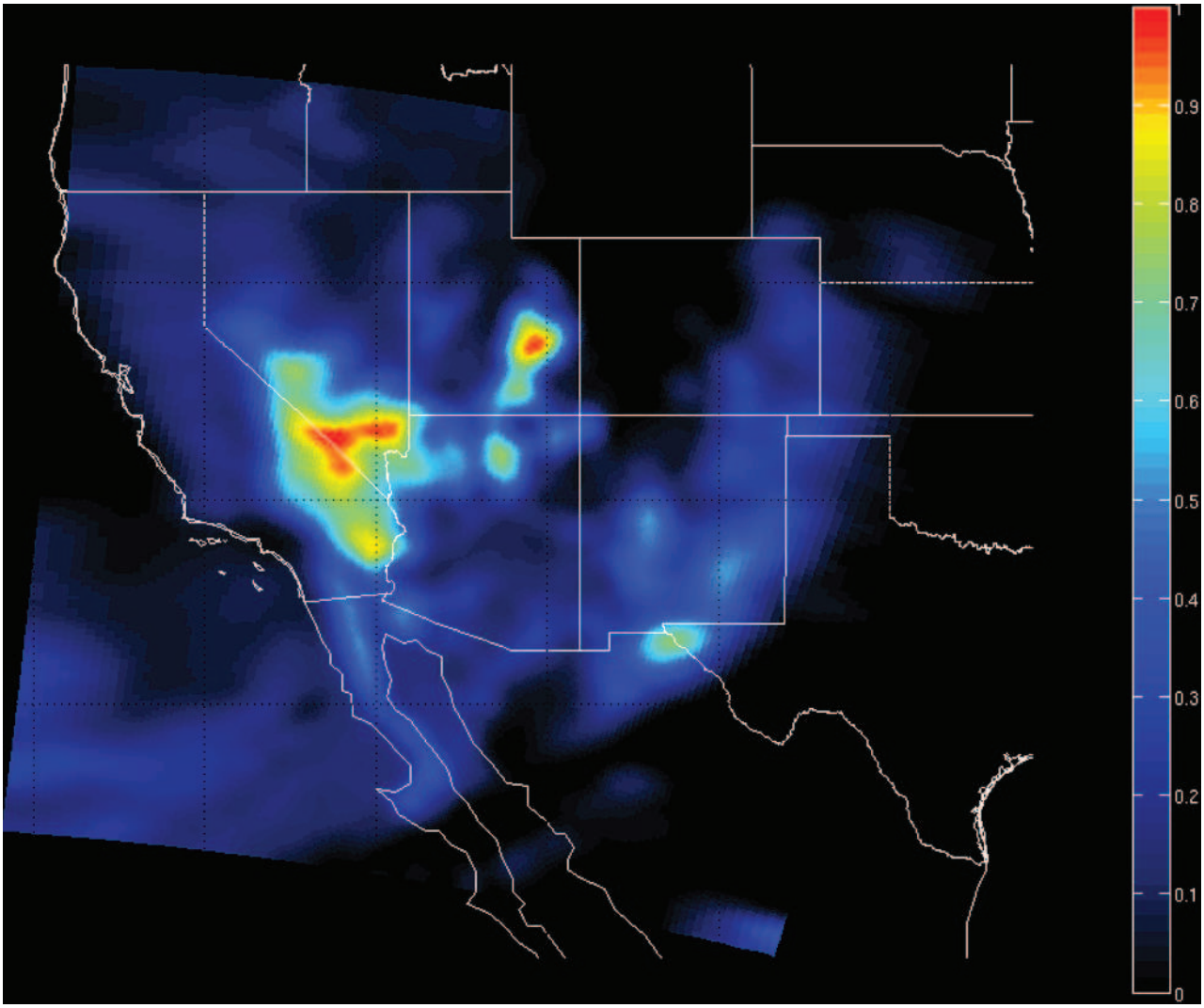

Figure 286. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

Figure 287. Result of the Maximum Likelihood model (MLE) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

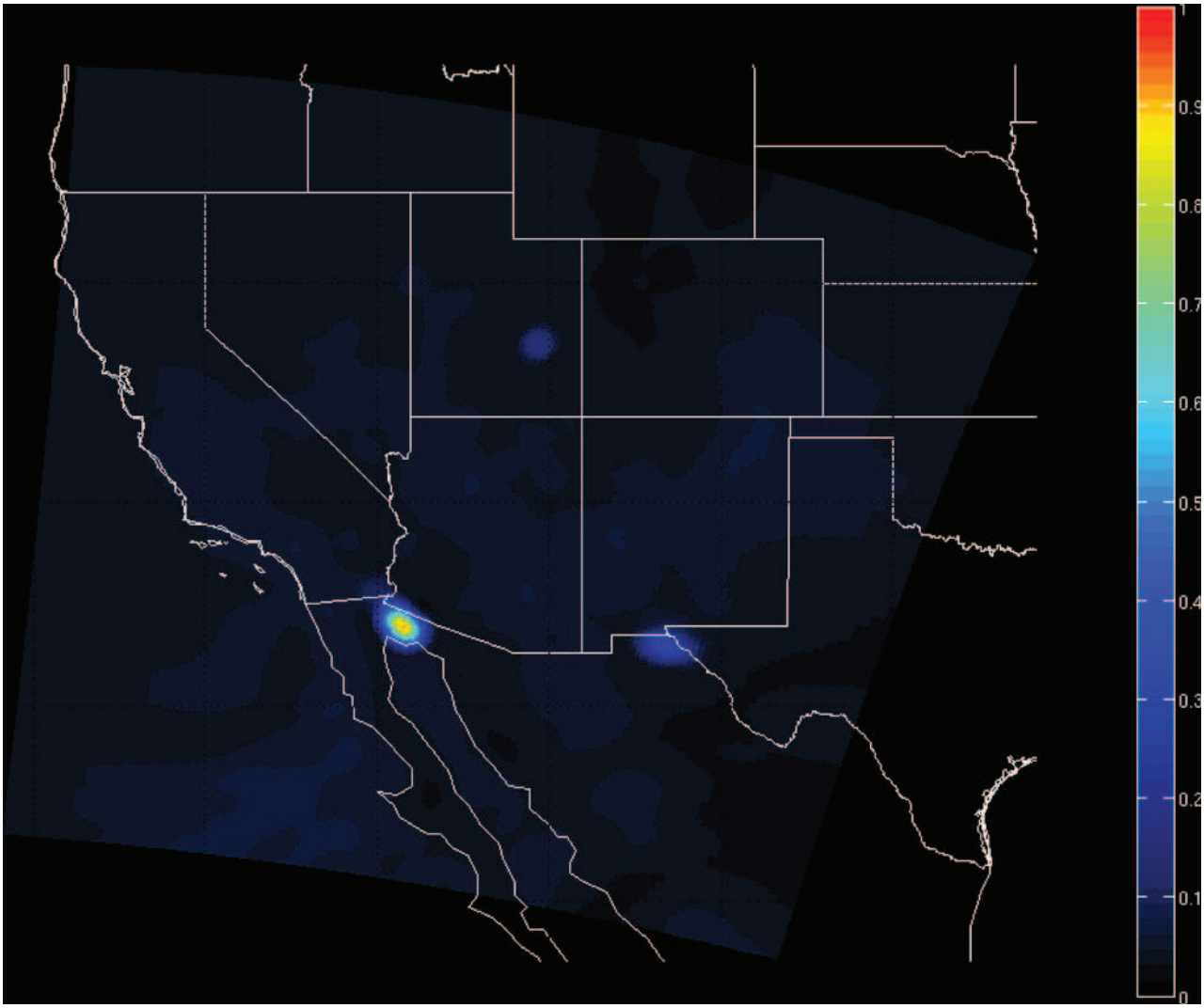

Figure 288. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

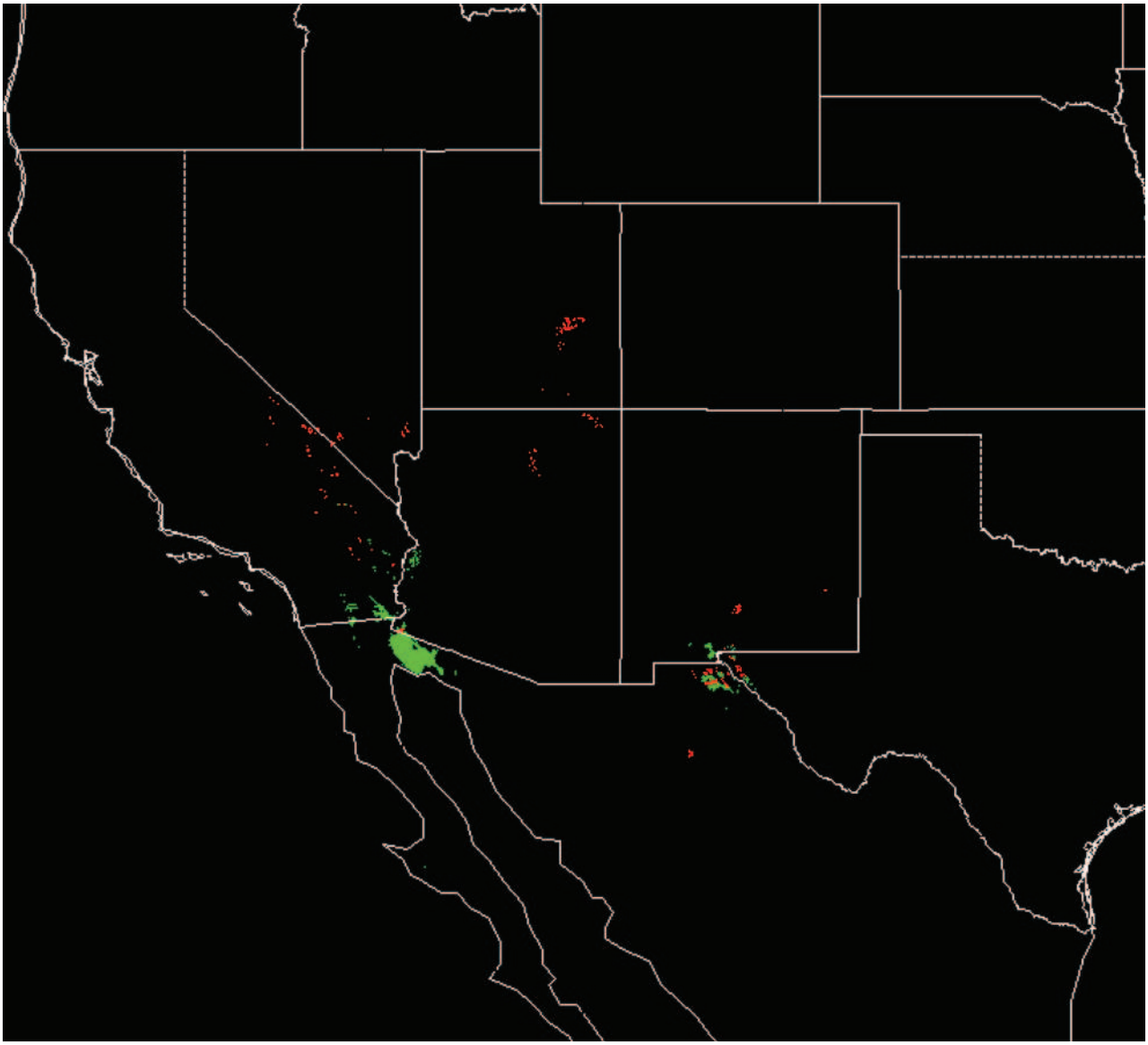

Figure 289. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2002-067 at 18:20 UTC.

Figure 290. Result of the Maximum a posteriori model (MAP) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

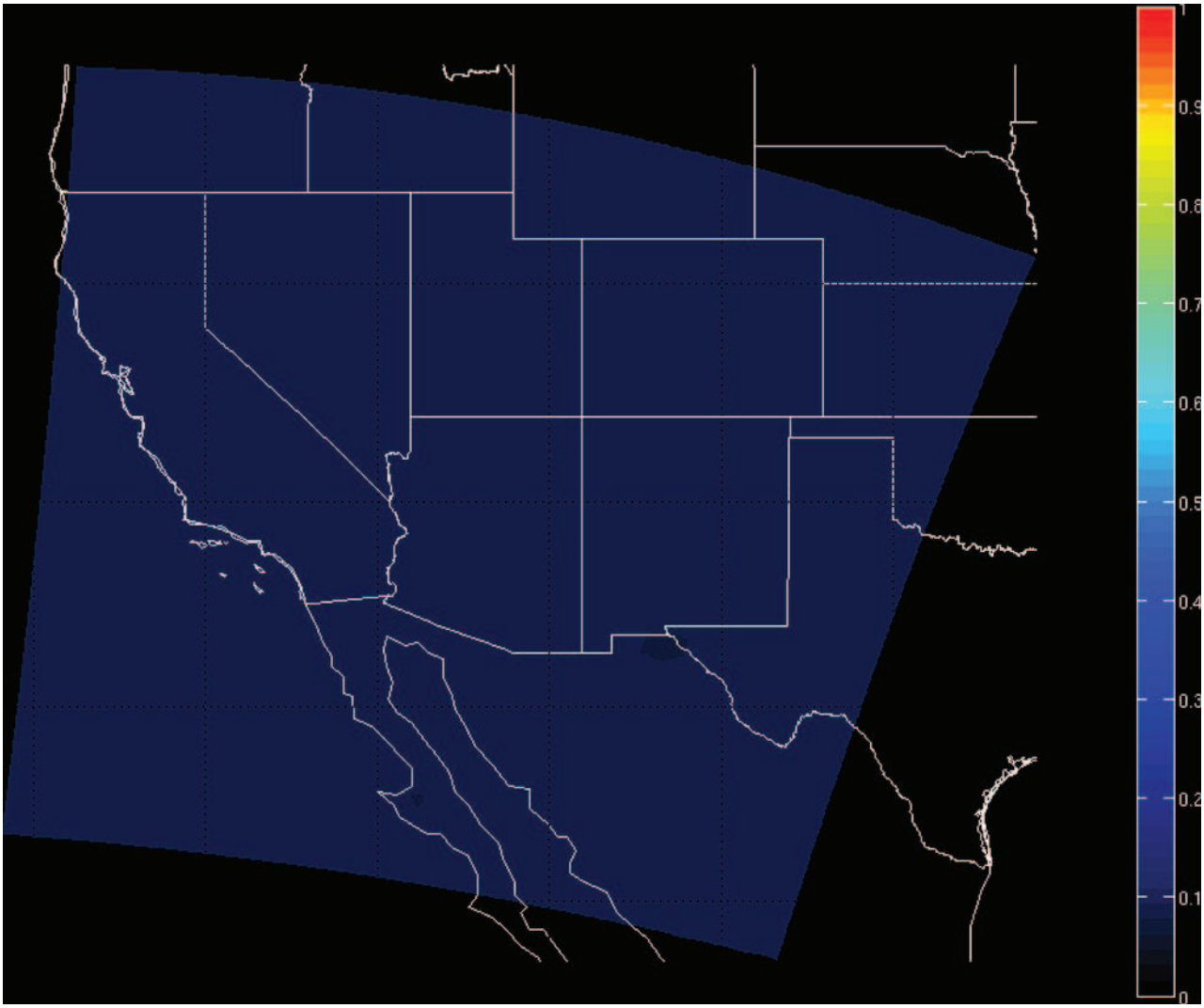

Figure 291. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

Figure 292. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2002-067 at 18:20 UTC.

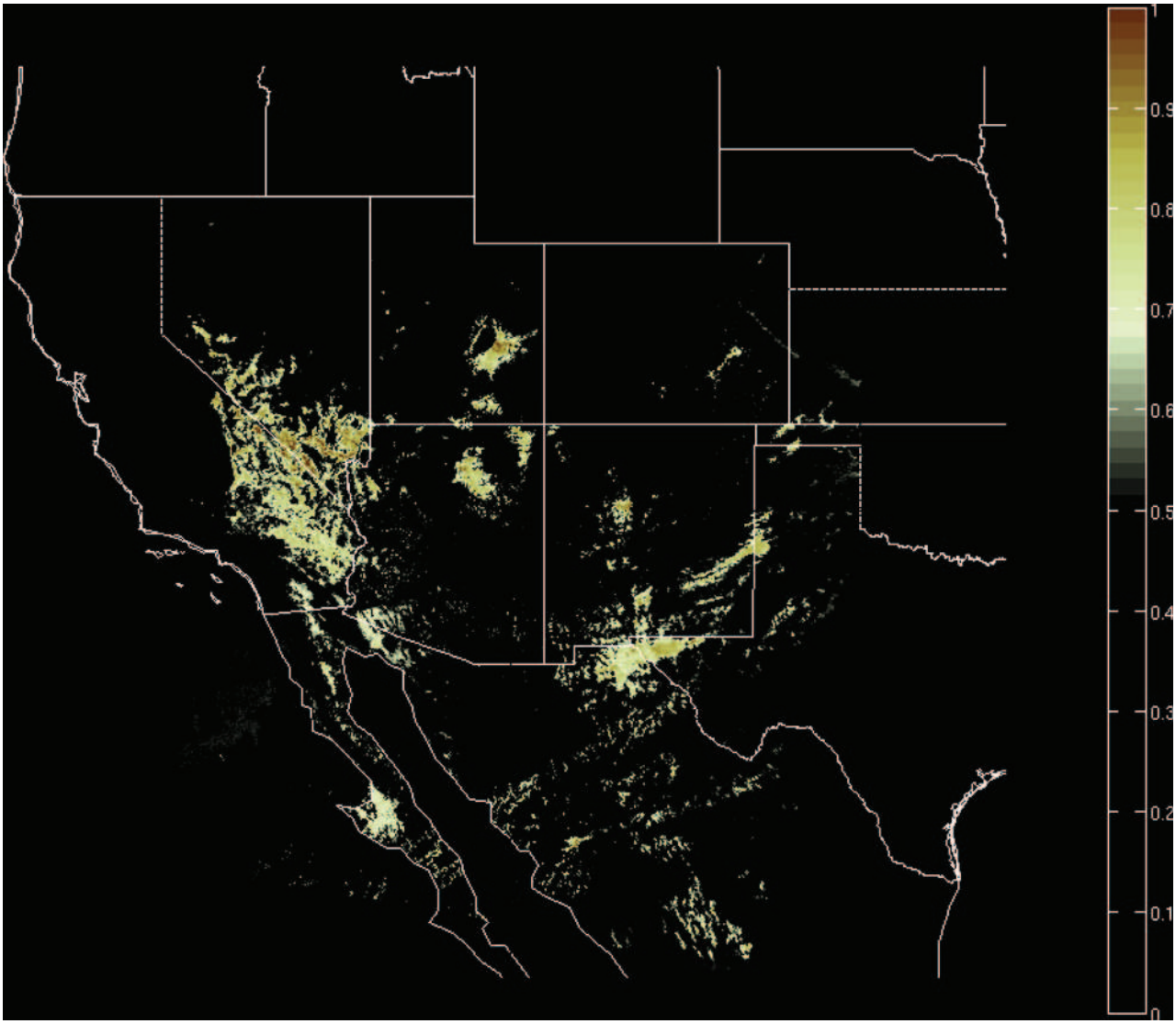

Figure 293. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

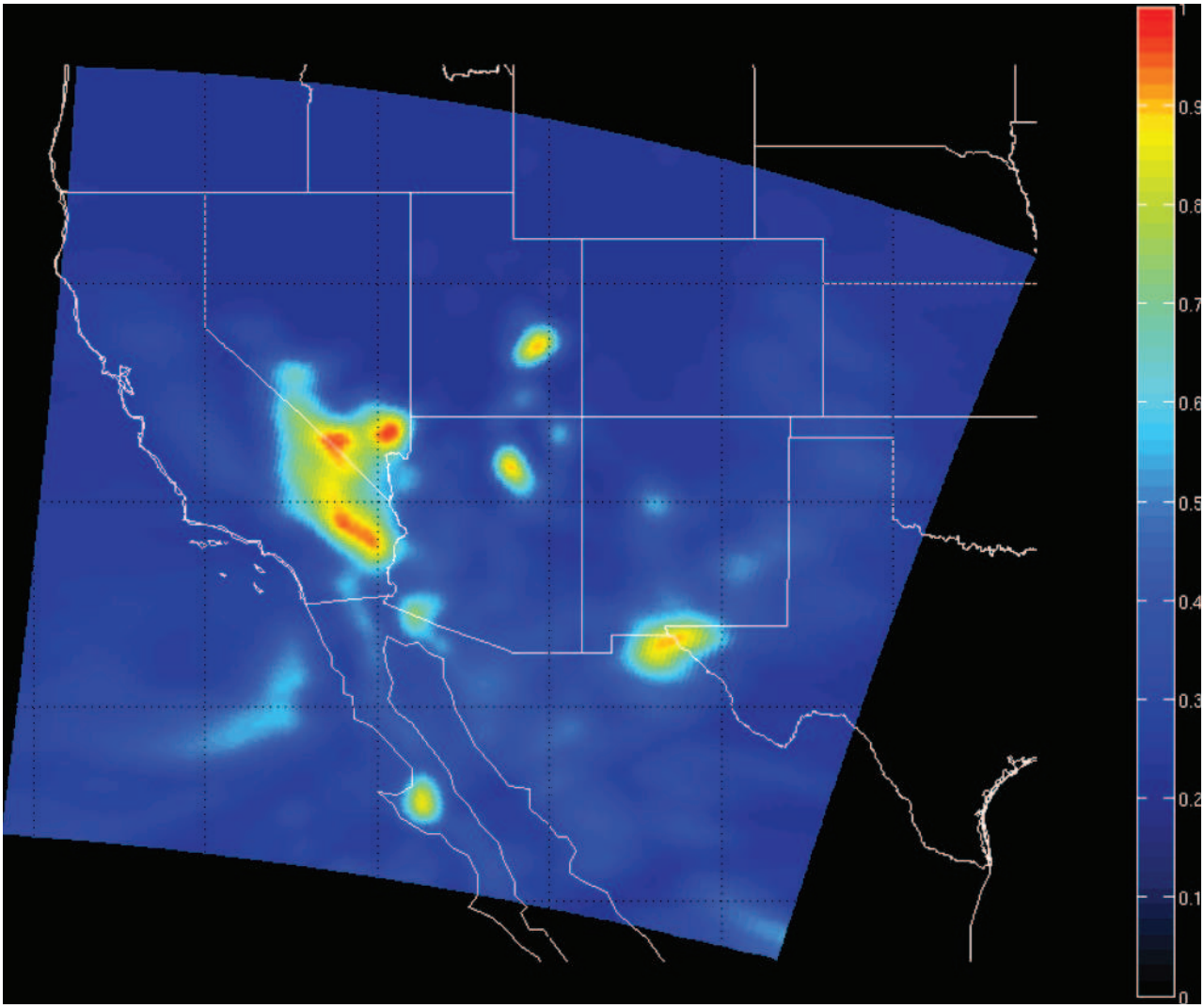

Figure 294. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

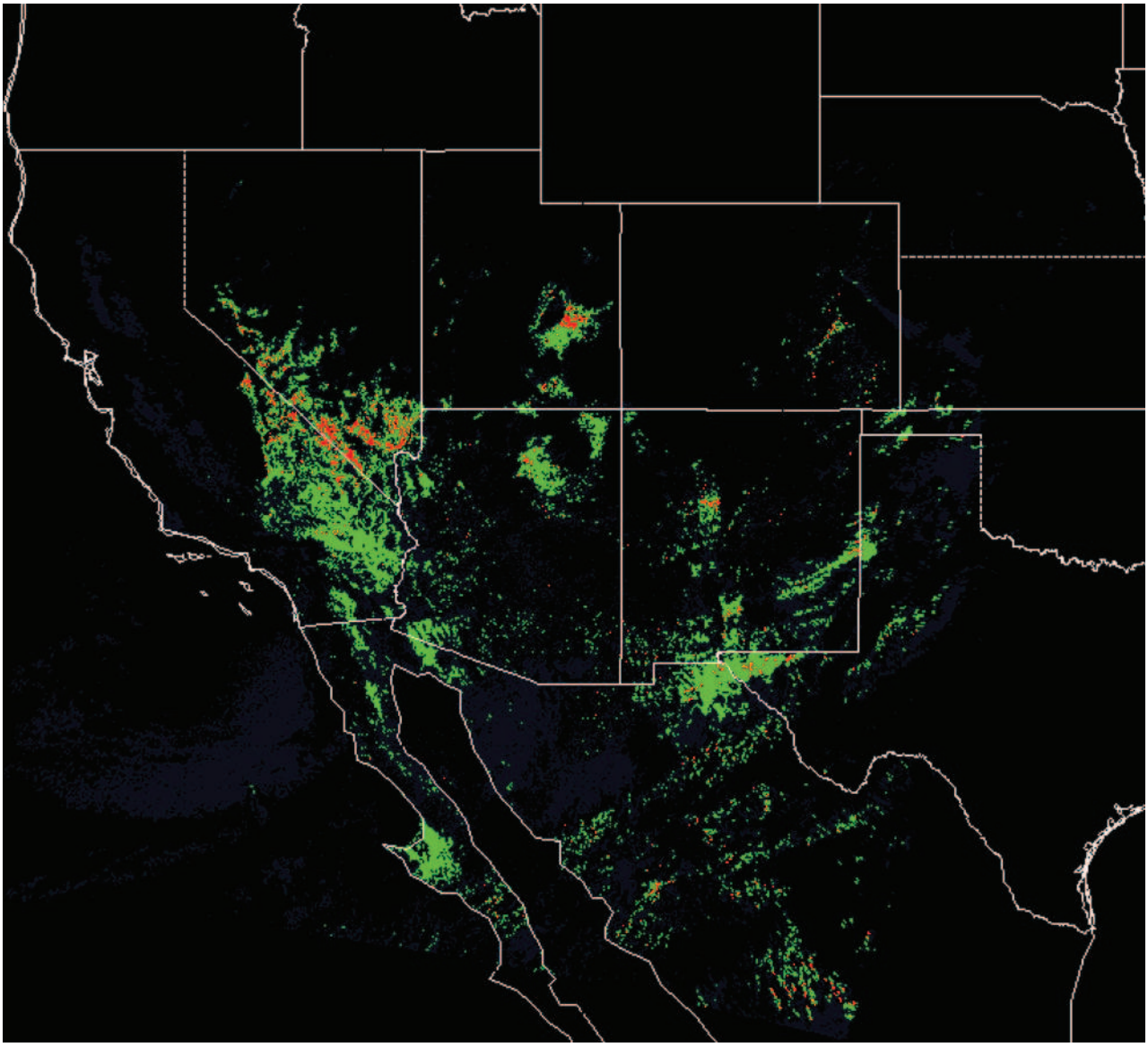

Figure 295. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-067 at 18:20 UTC.

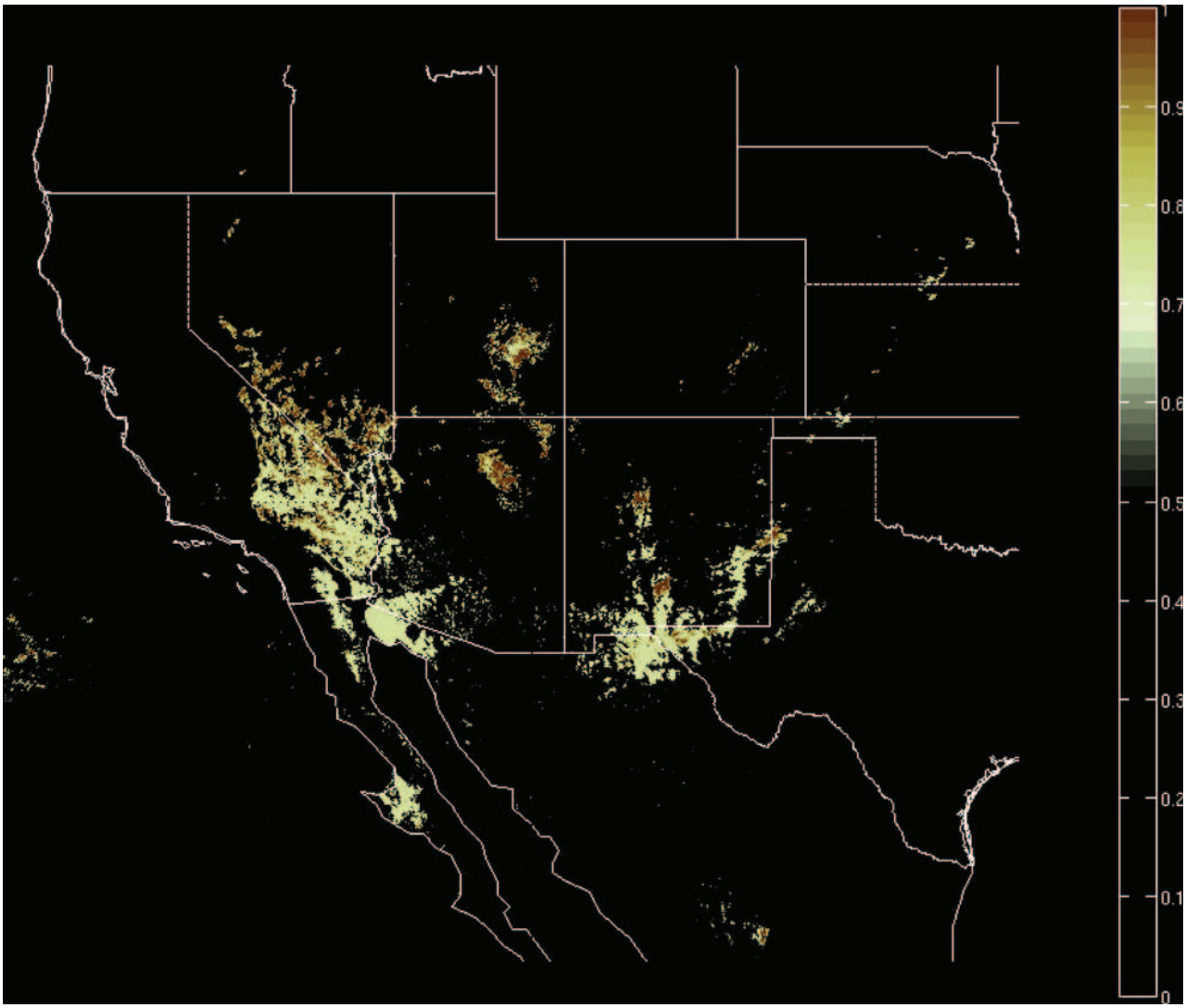

Figure 296. Result of the Probabilistic Neural Network (PNN) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

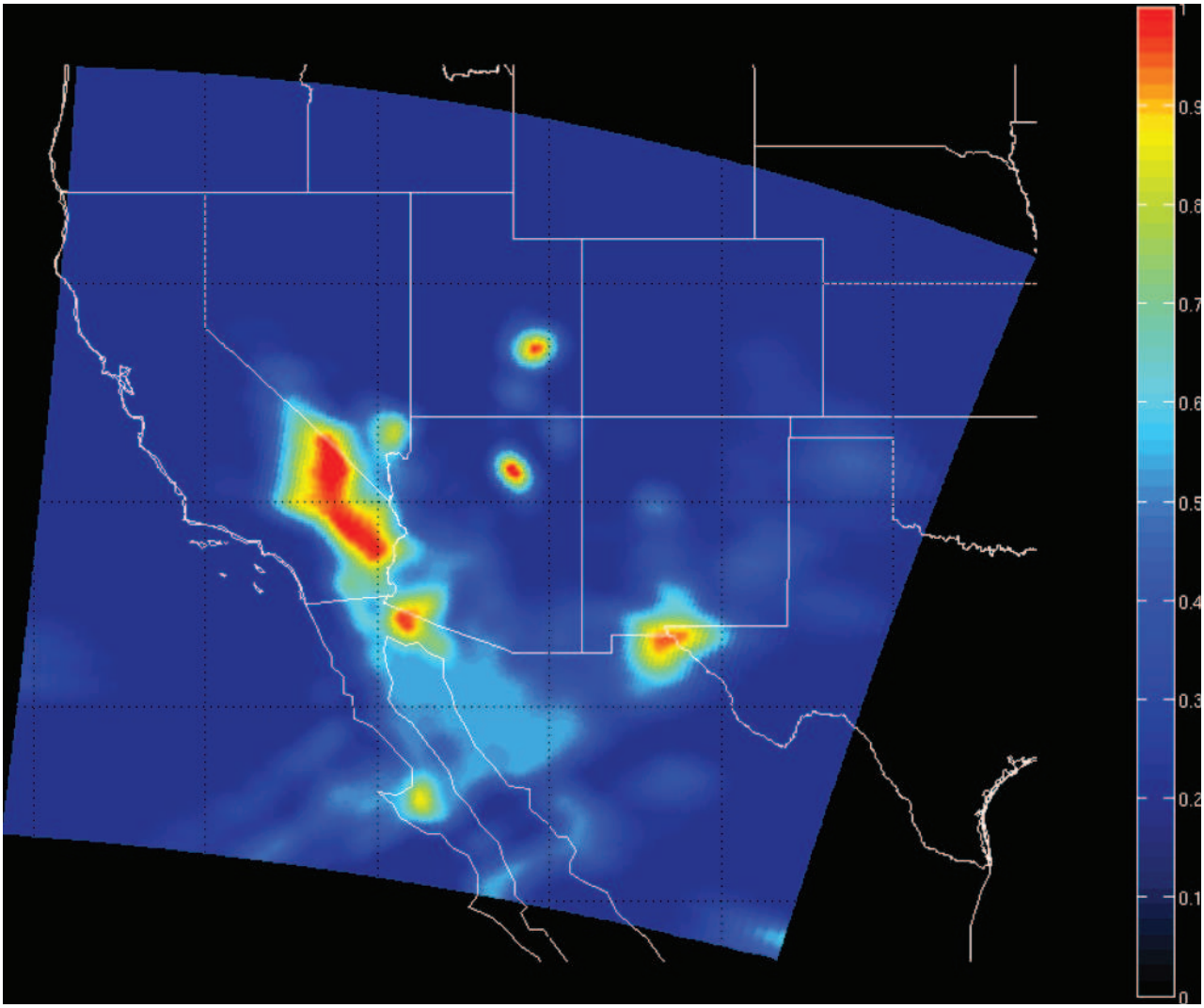

Figure 297. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2002-067 at 18:20 UTC. The colorbar shows the index of probability.

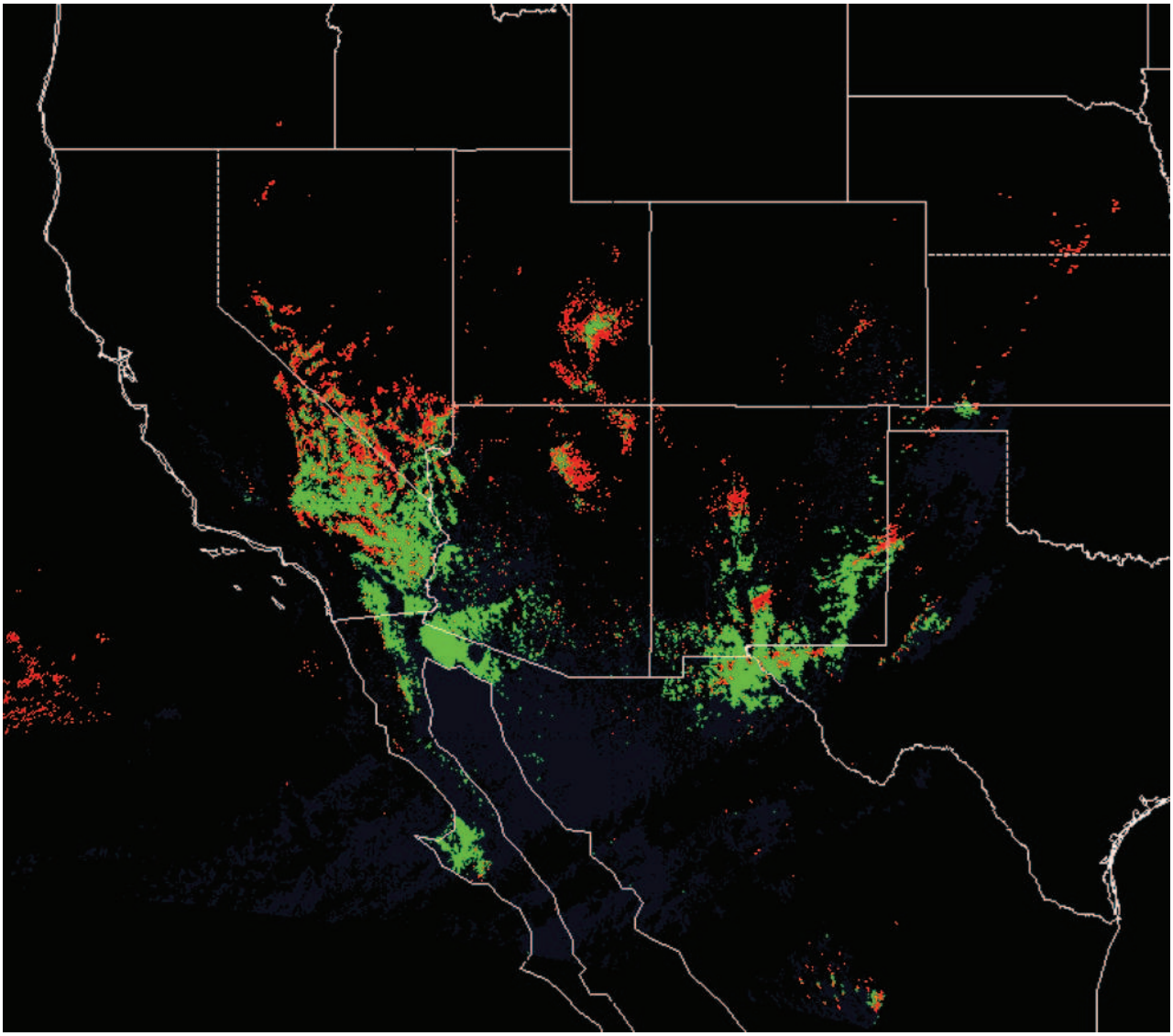

Figure 298. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2002-067 at 18:20 UTC.

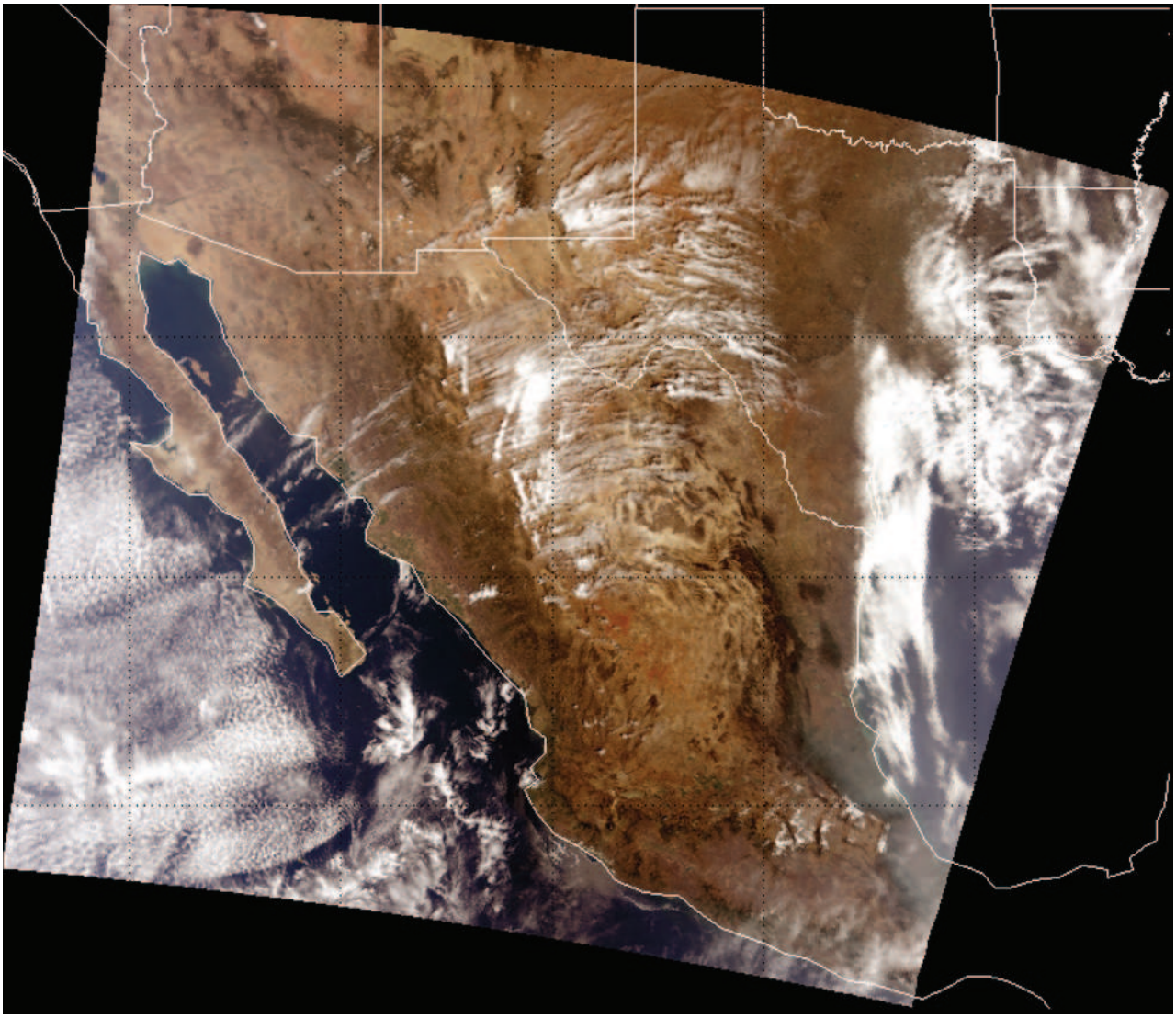

Figure 299. True color image for the event of 2002-073 at 17:45 UTC.

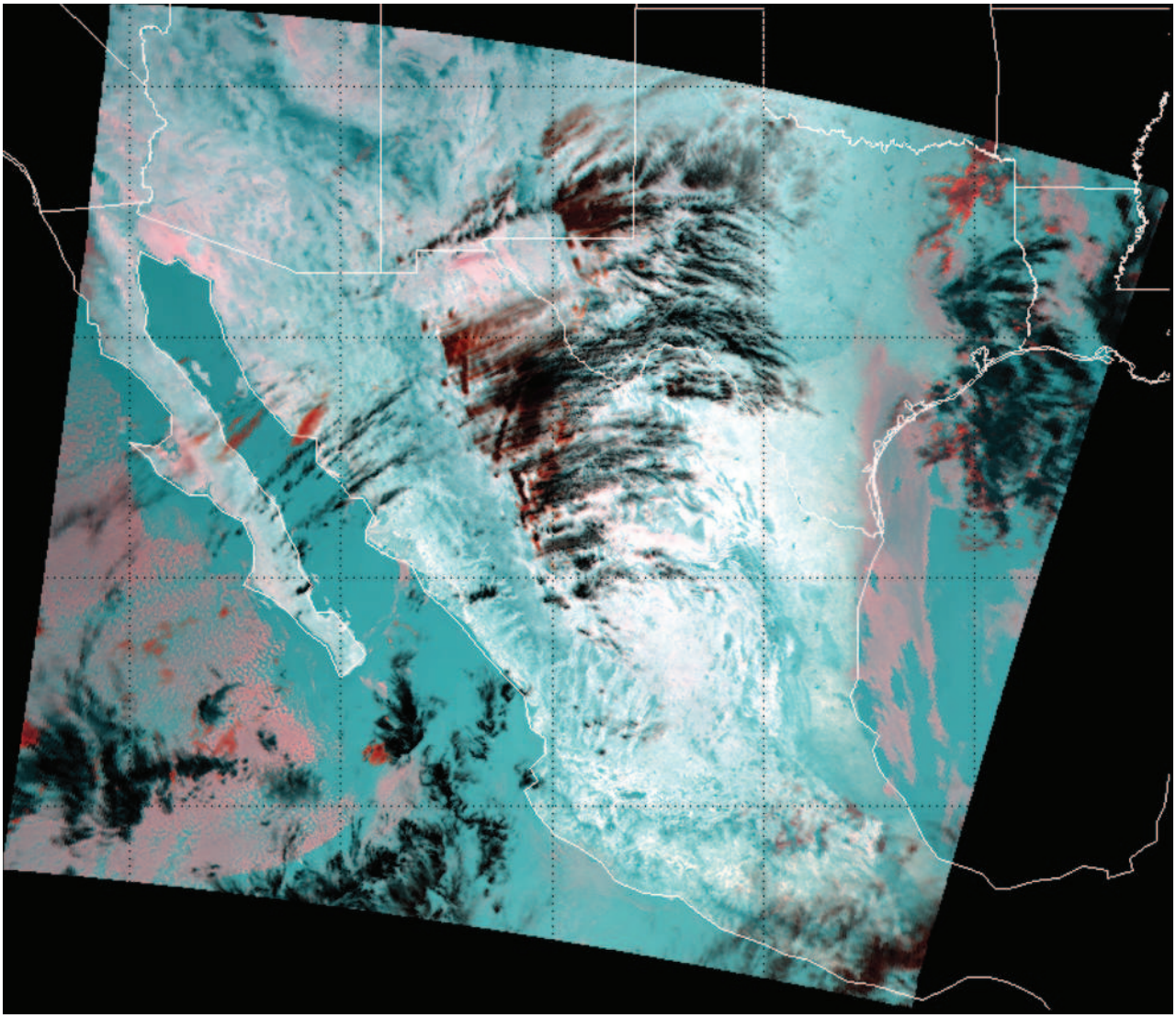

Figure 300. False color image for the event of 2002-073 at 17:45 UTC.

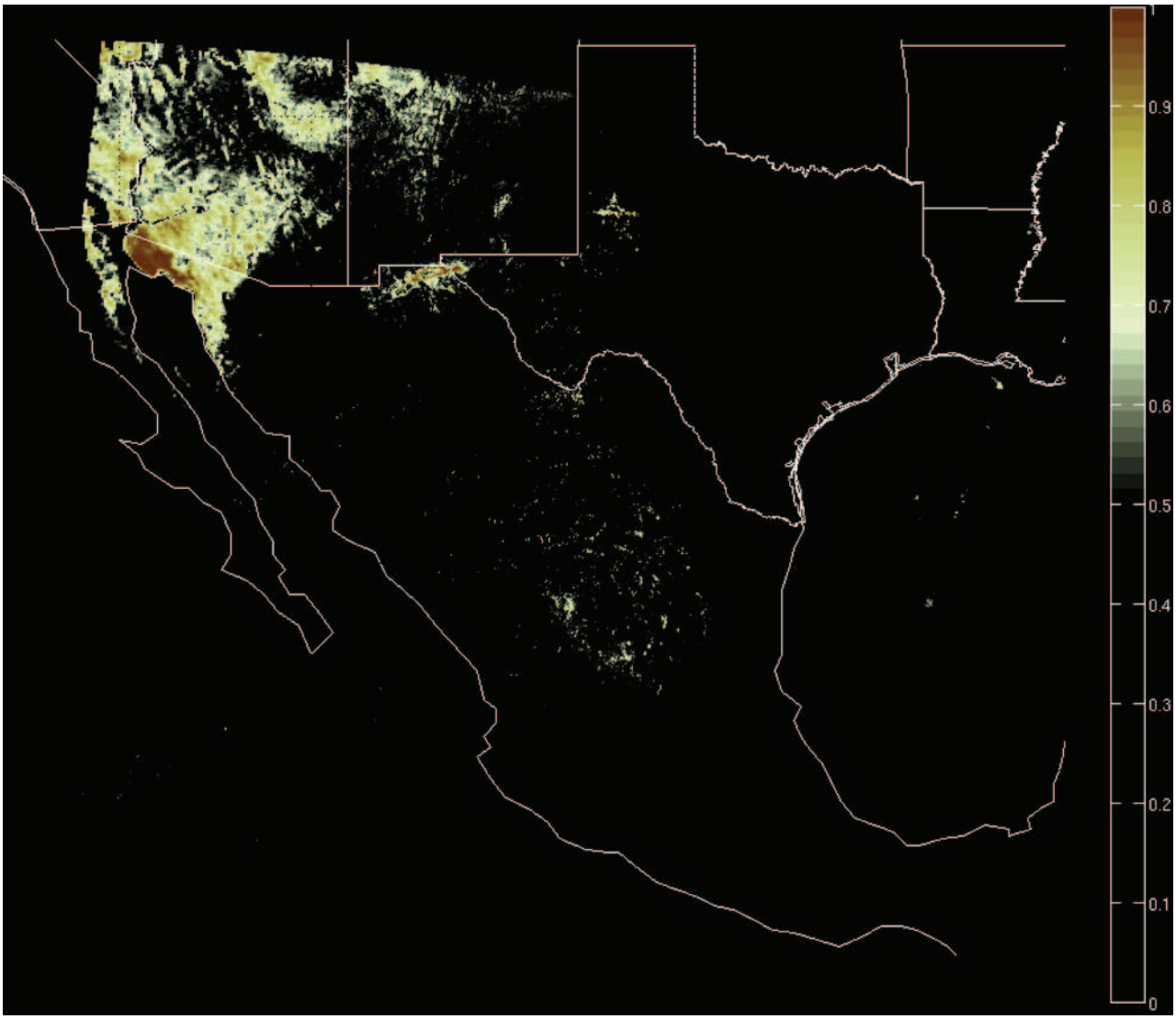

Figure 301. Result of the probabilistic display model (PD) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

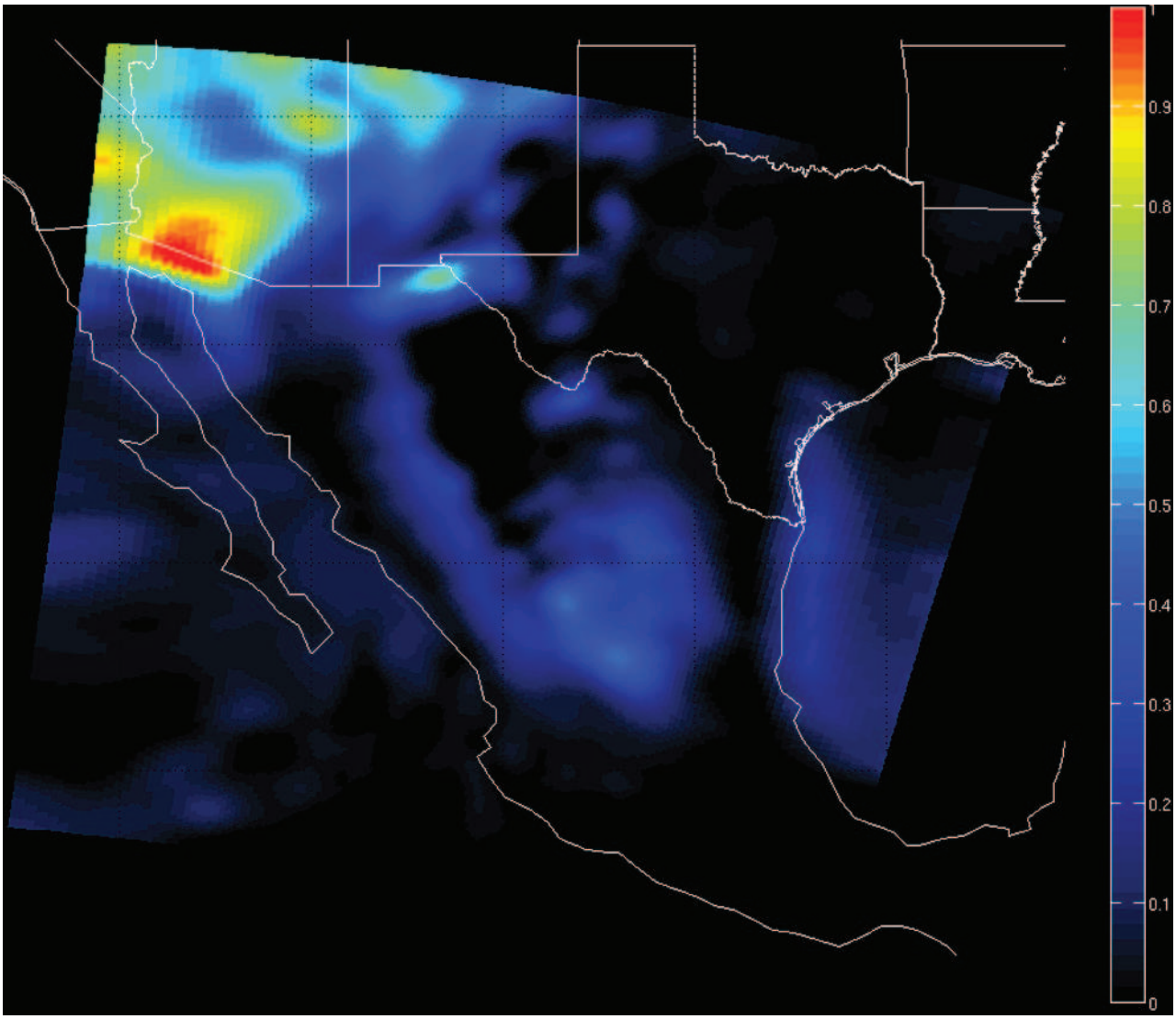

Figure 302. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

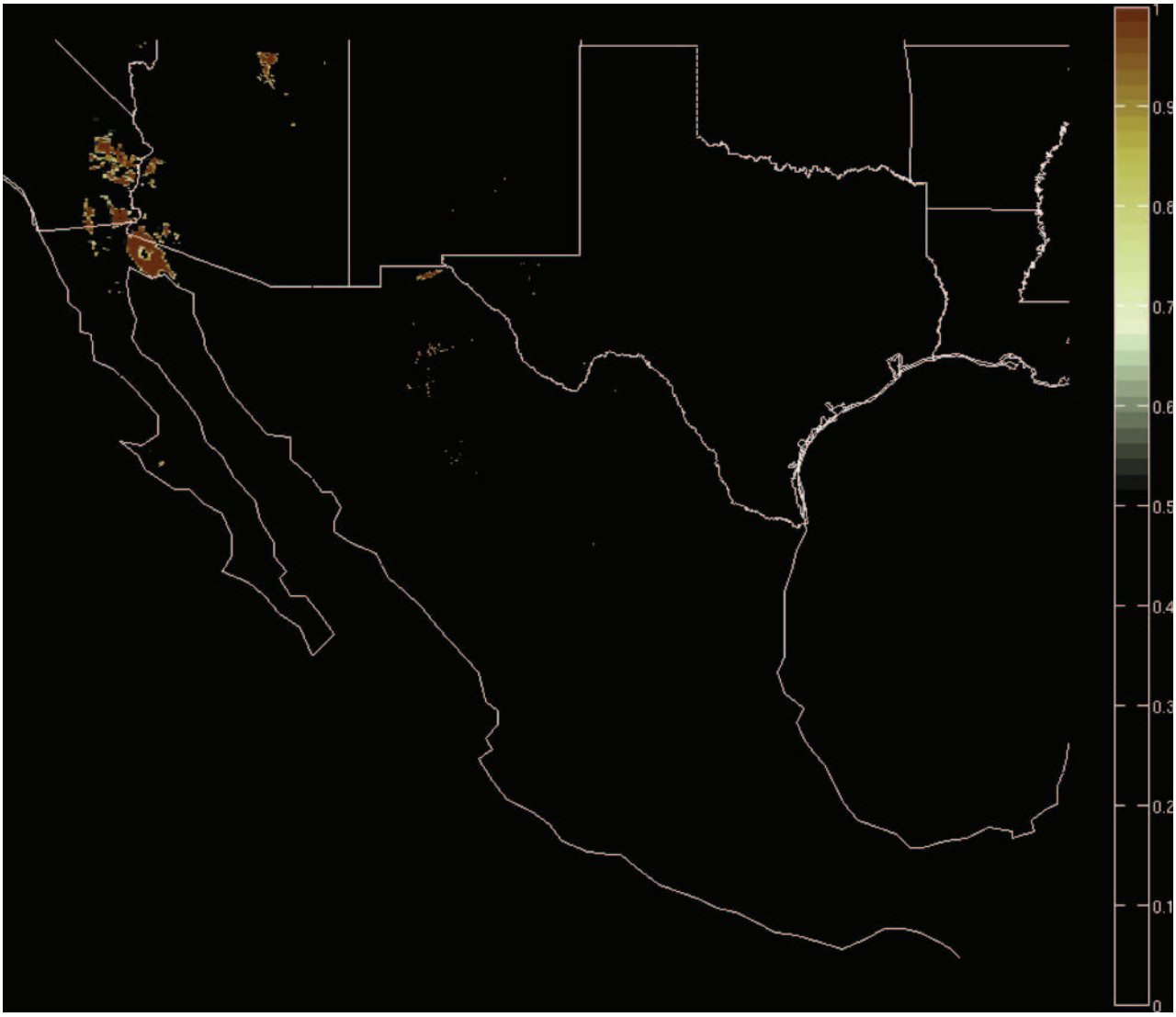

Figure 303. Result of the Maximum Likelihood model (MLE) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

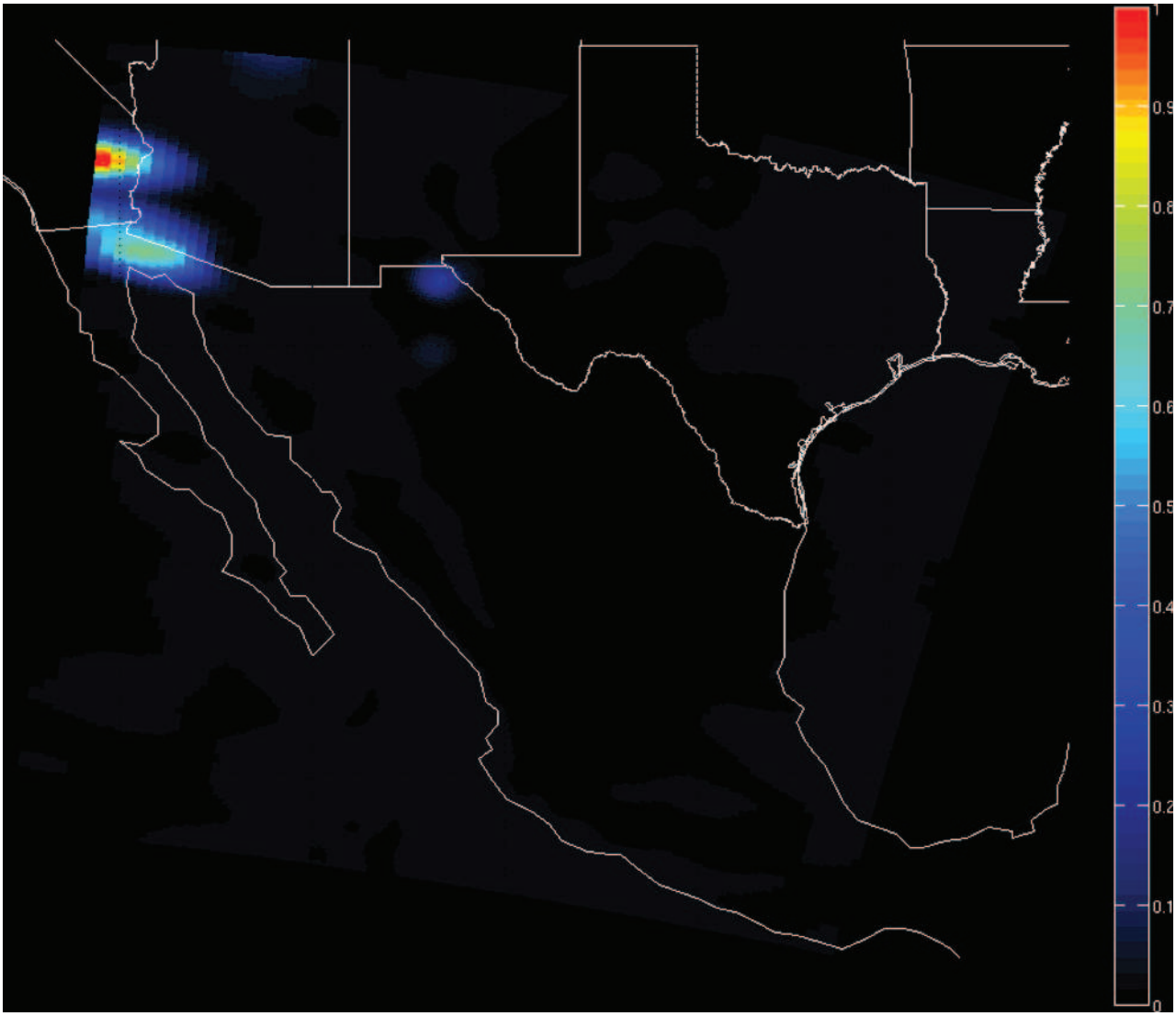

Figure 304. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

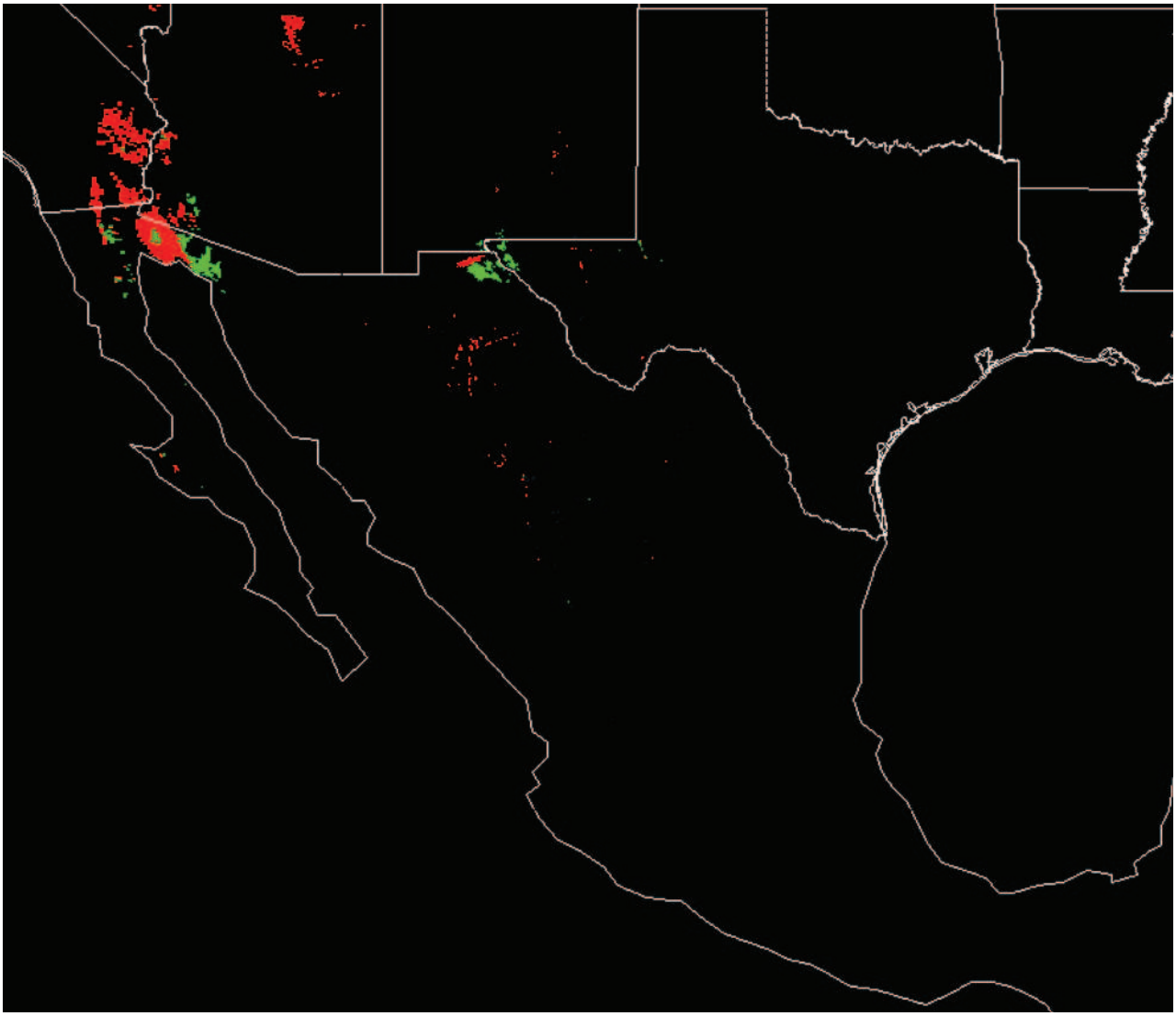

Figure 305. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2002-073 at 17:45 UTC.

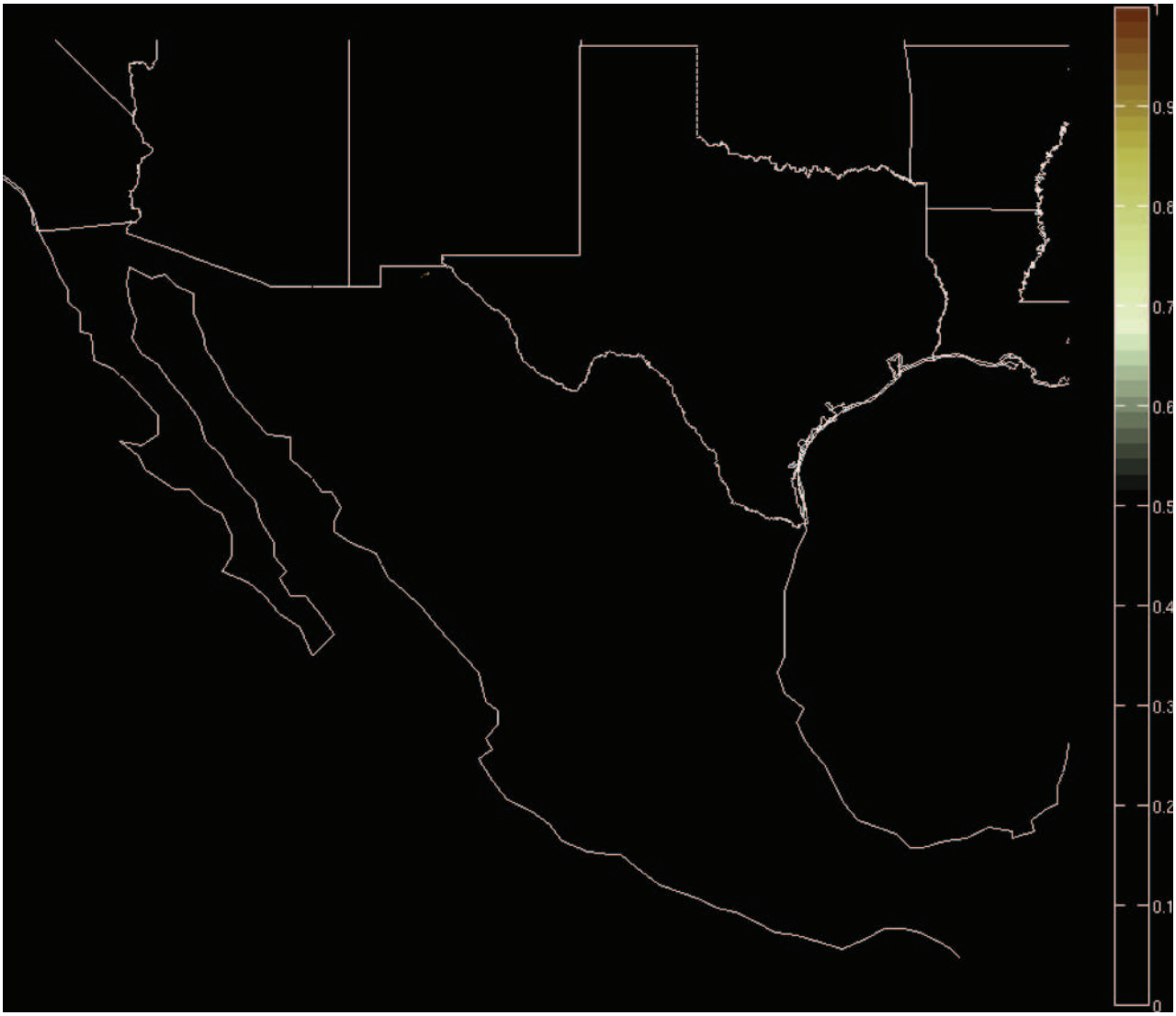

Figure 306. Result of the Maximum a posteriori model (MAP) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

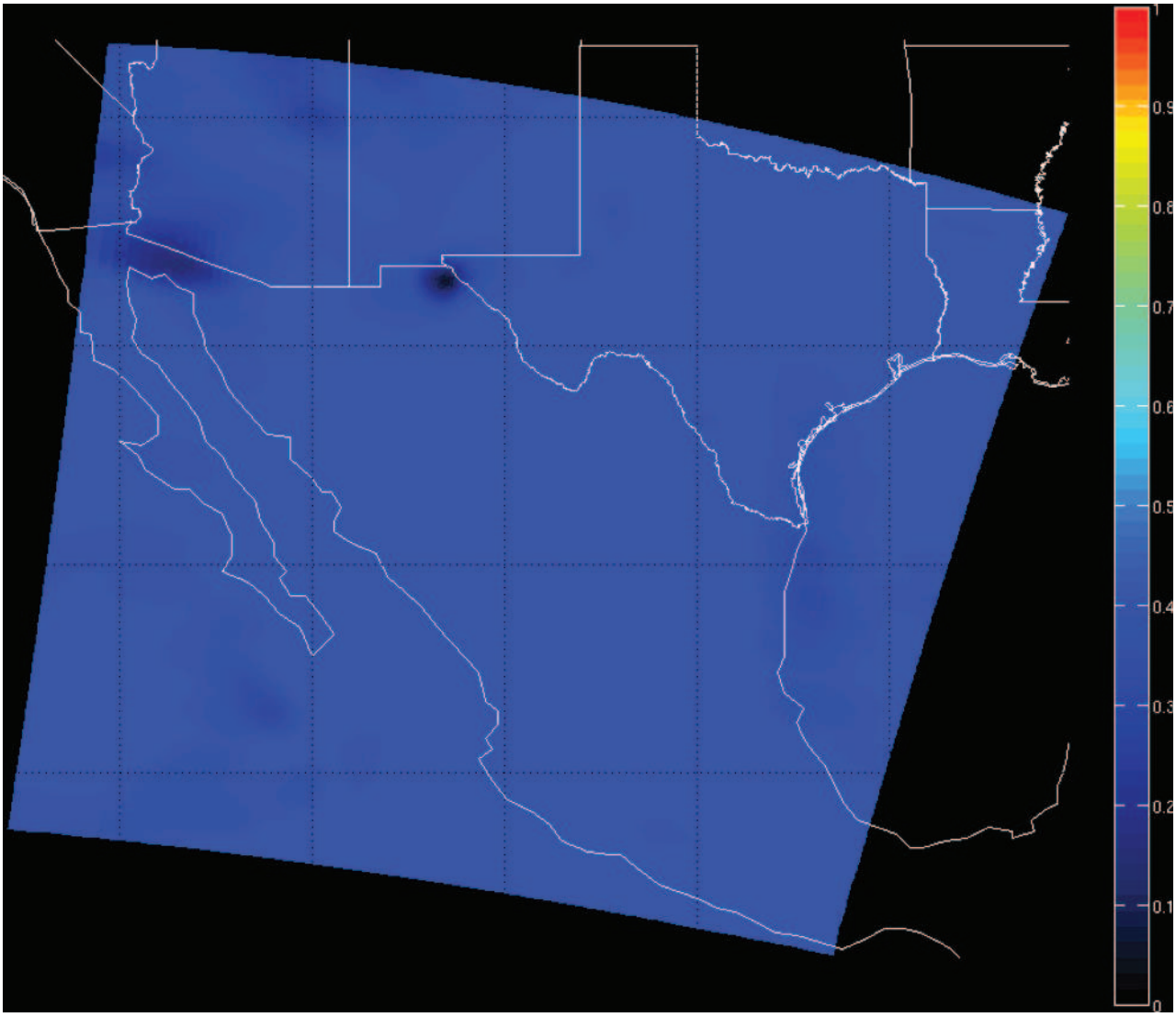

Figure 307. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

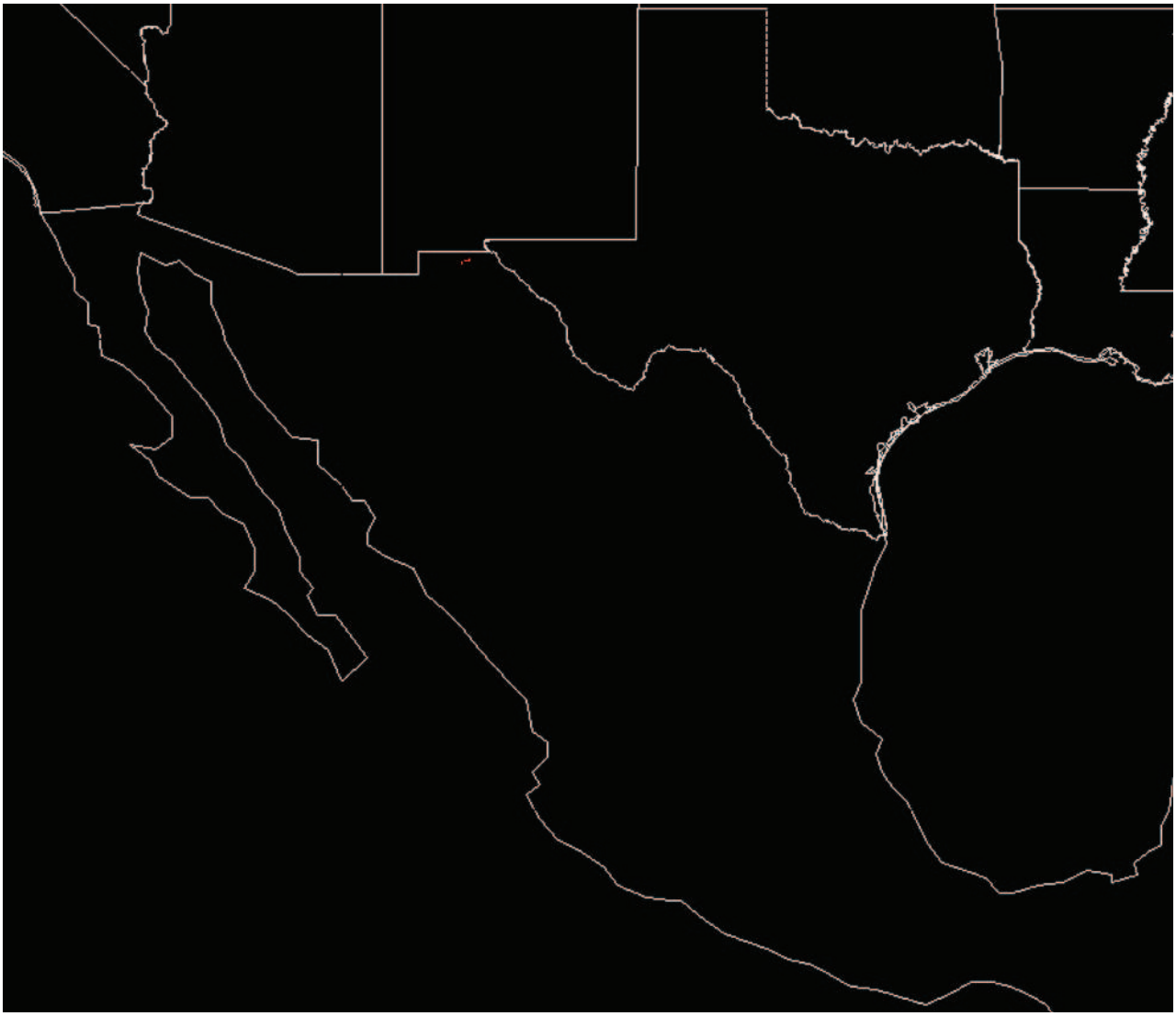

Figure 308. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2002-073 at 17:45 UTC.

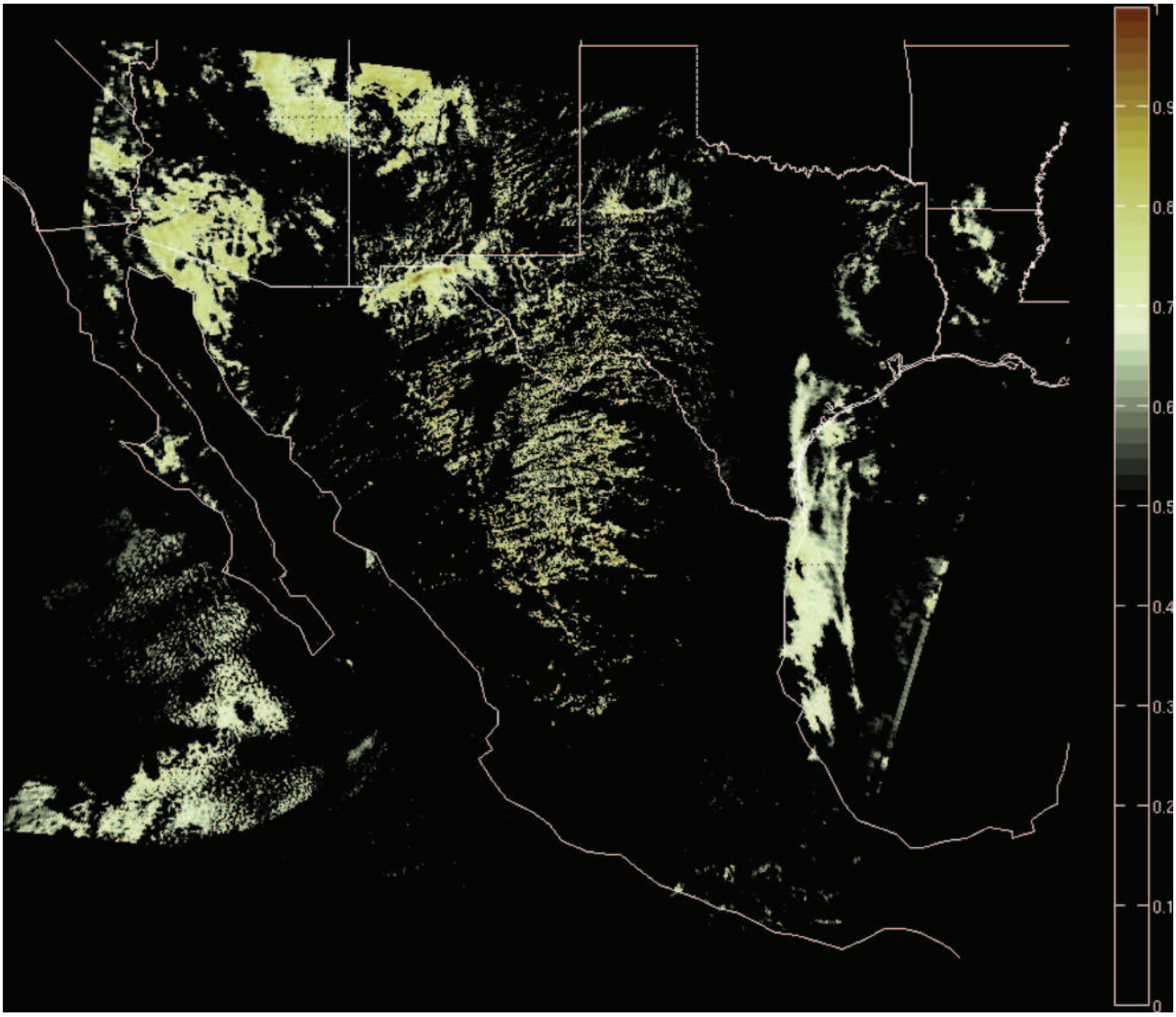

Figure 309. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

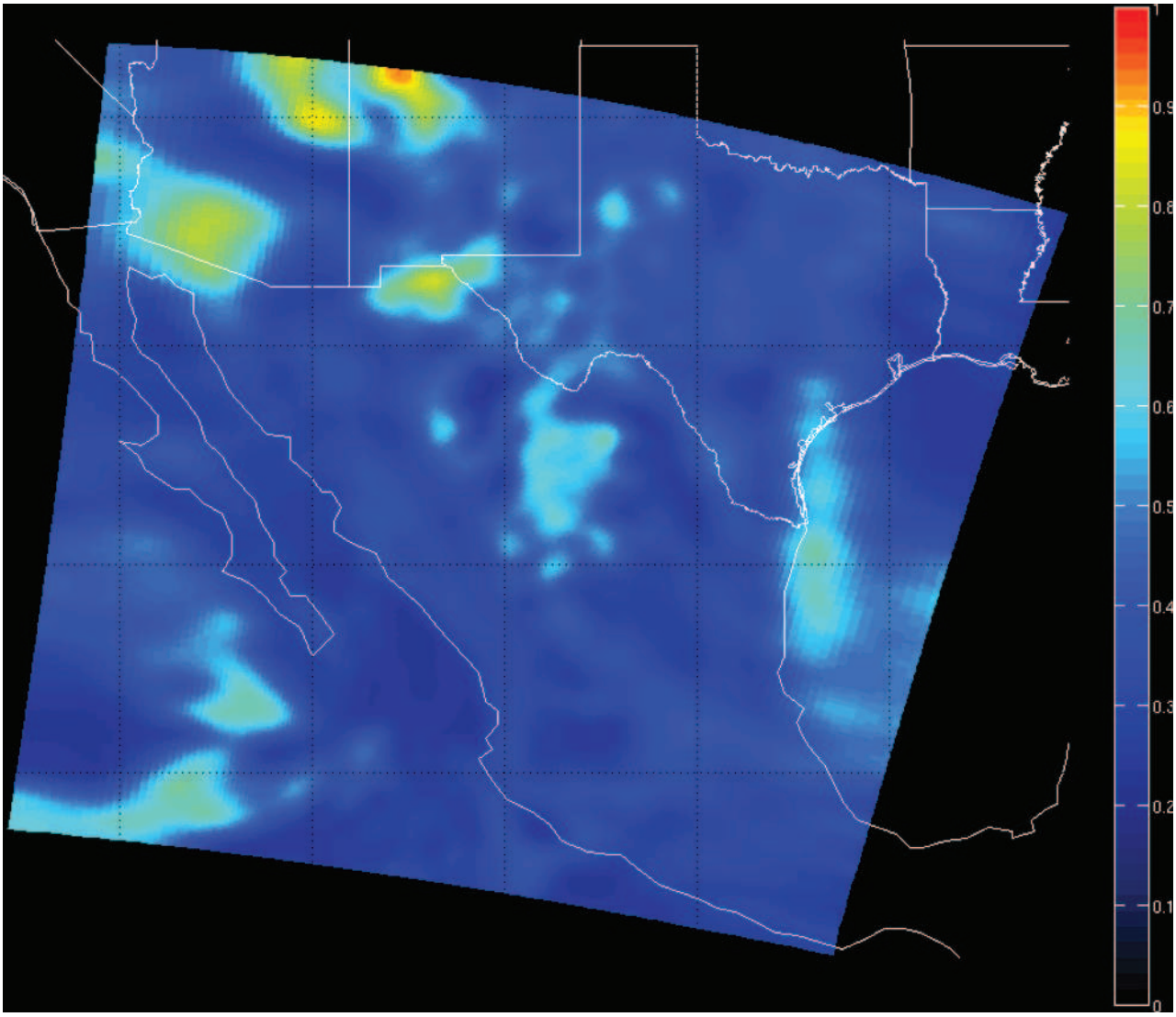

Figure 310. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

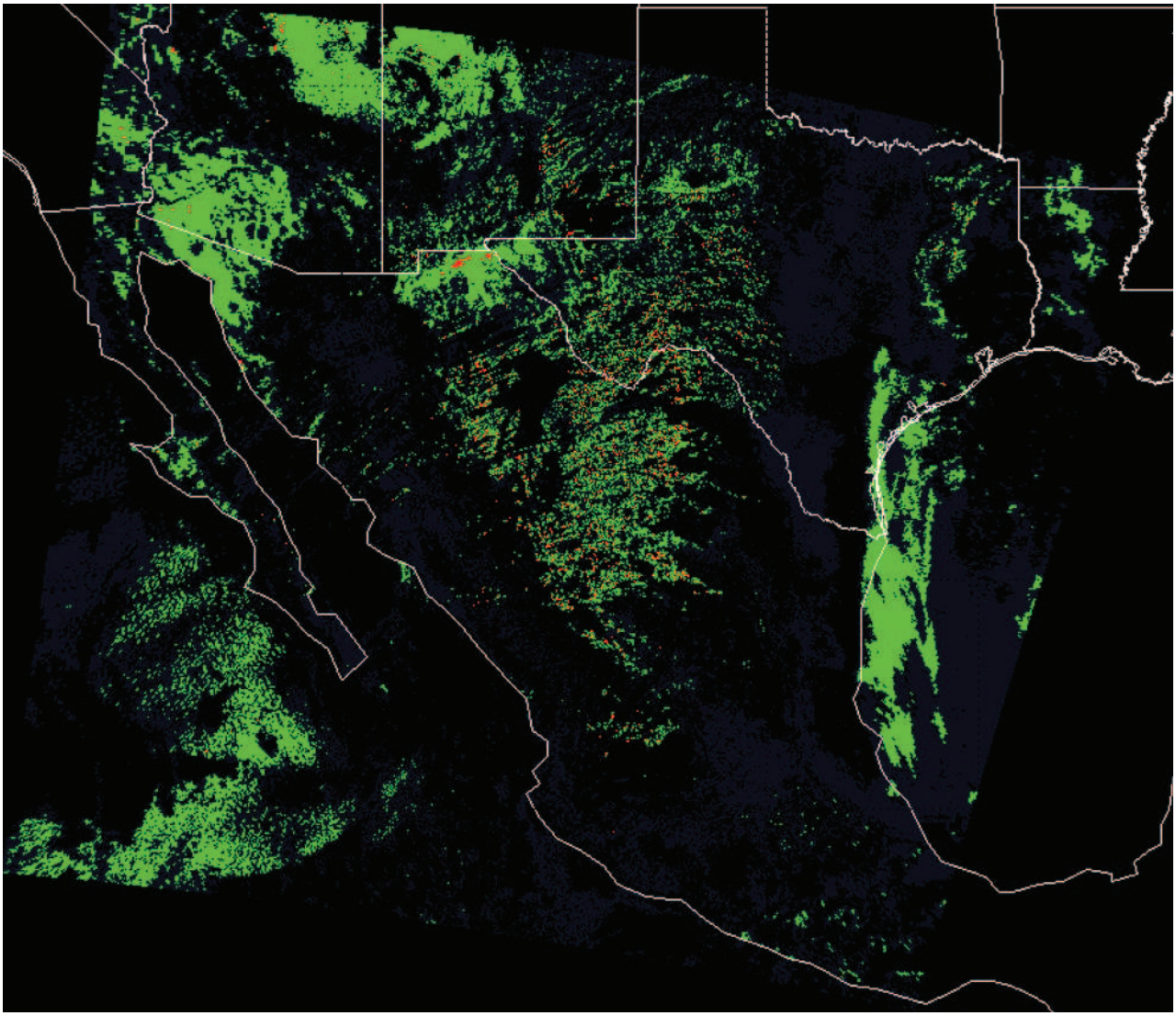

Figure 311. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-073 at 17:45 UTC.

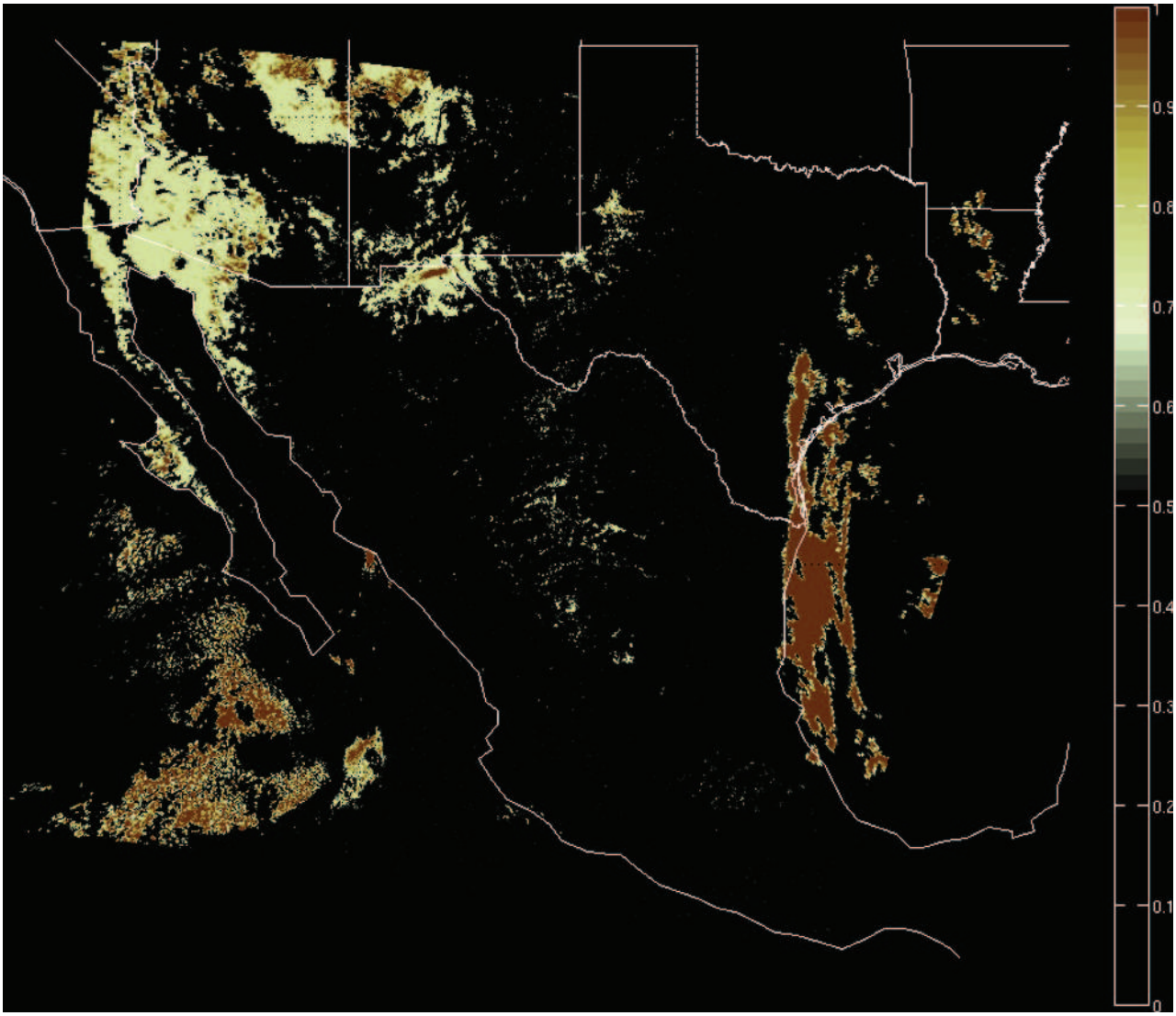

Figure 312. Result of the Probabilistic Neural Network (PNN) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

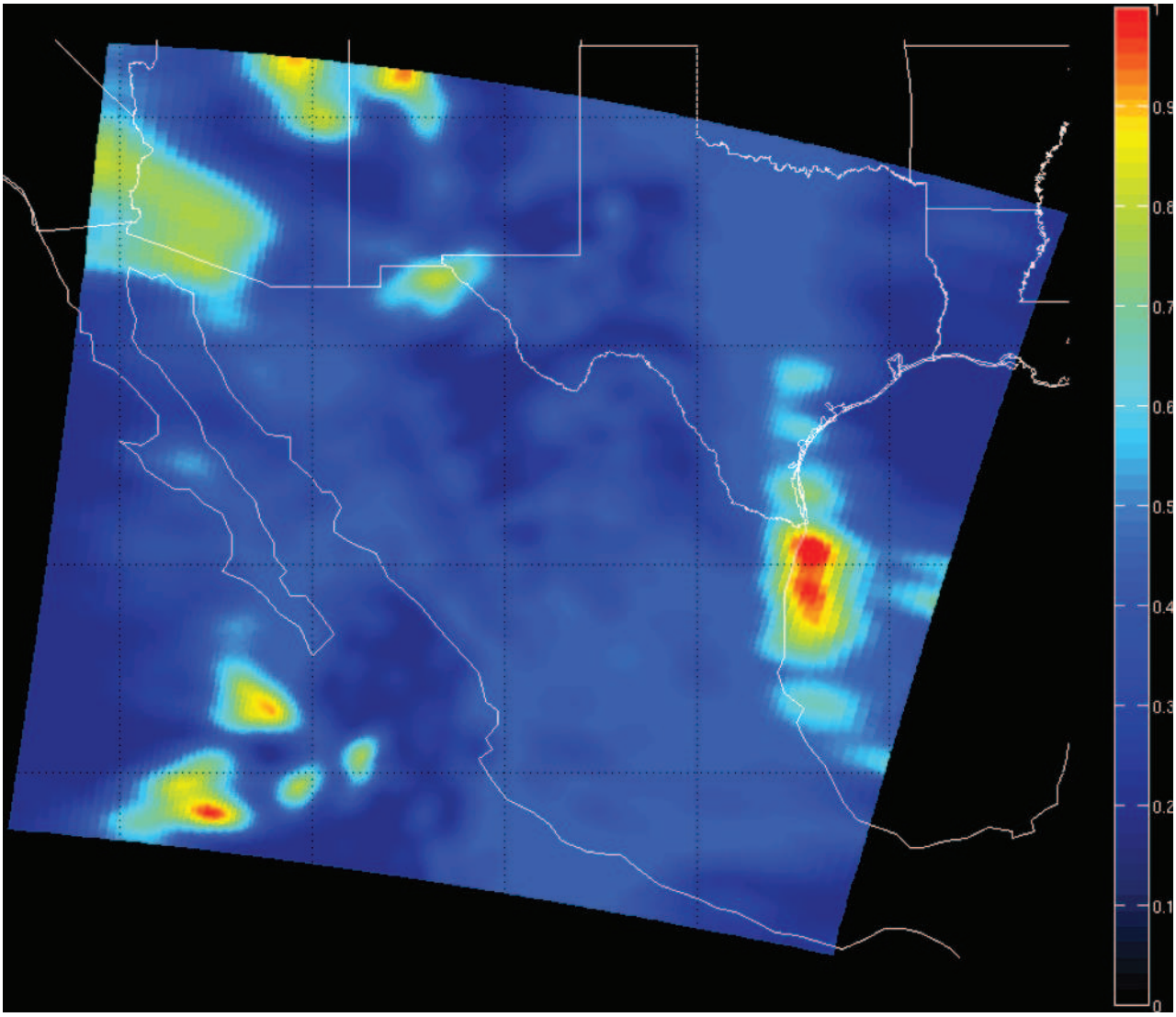

Figure 313. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2002-073 at 17:45 UTC. The colorbar shows the index of probability.

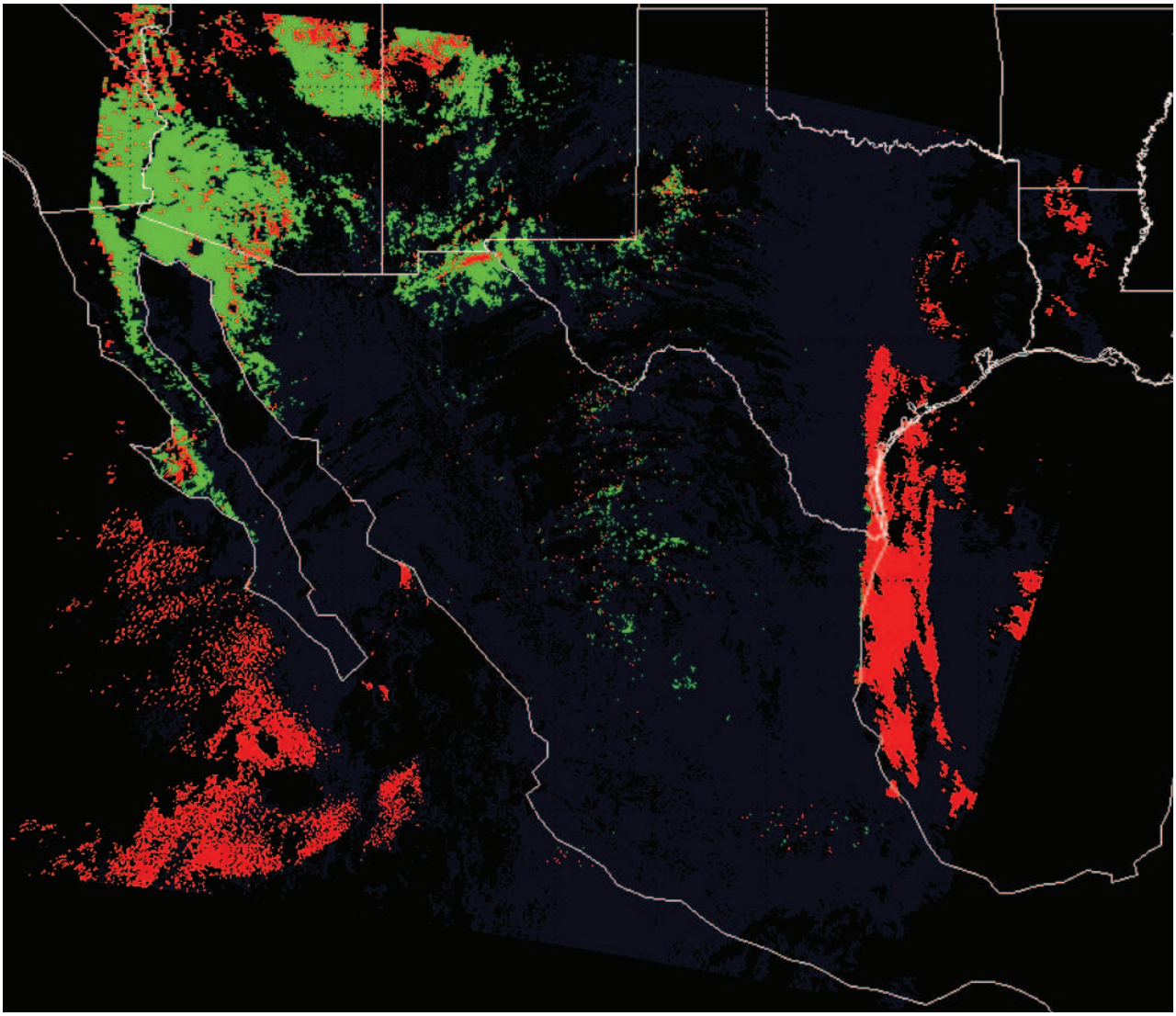

Figure 314. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2002-073 at 17:45 UTC.

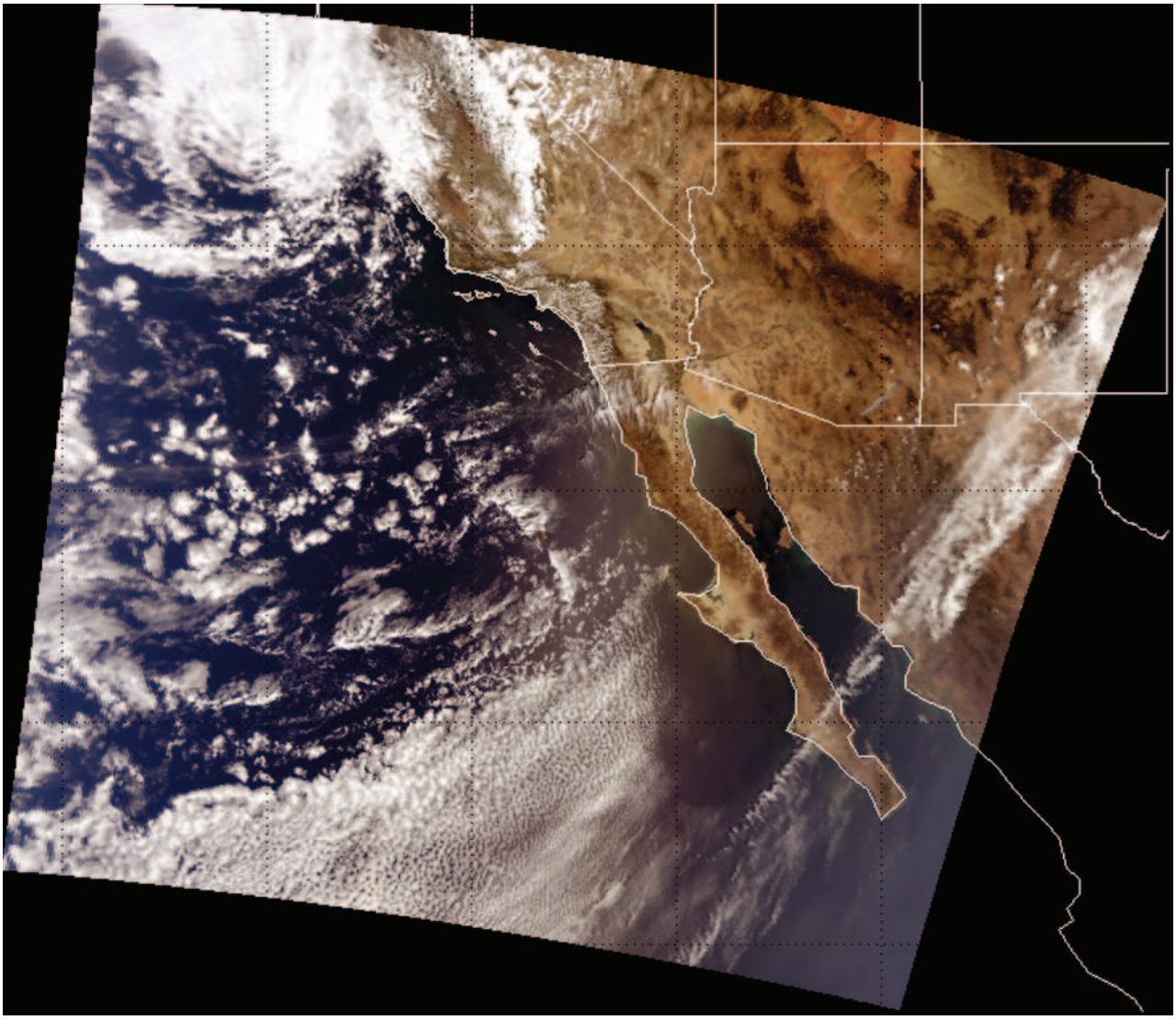

Figure 315. True color image for the event of 2002-120 at 18:40 UTC.

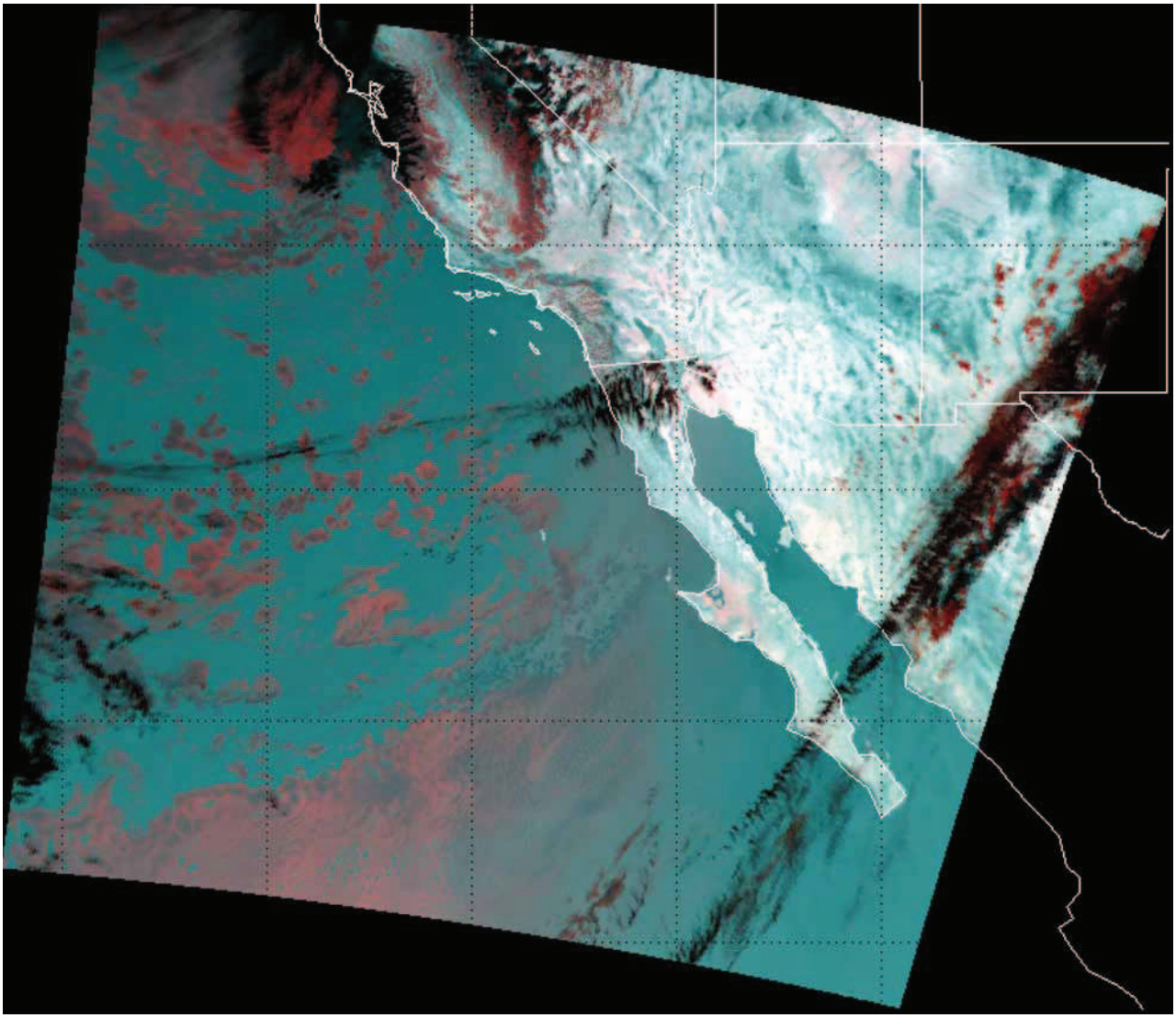

Figure 316. False color image for the event of 2002-120 at 18:40 UTC.

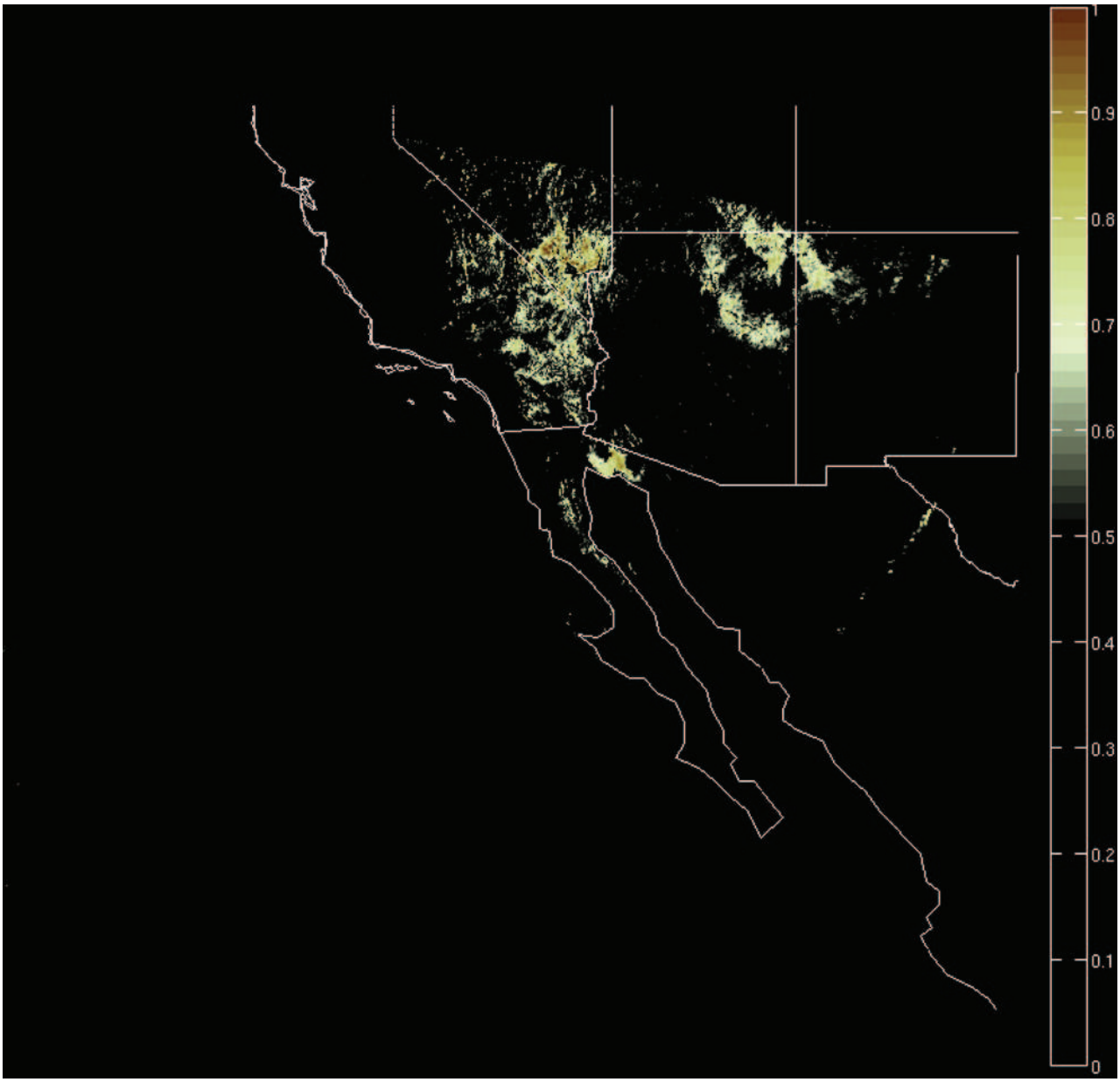

Figure 317. Result of the probabilistic display model (PD) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

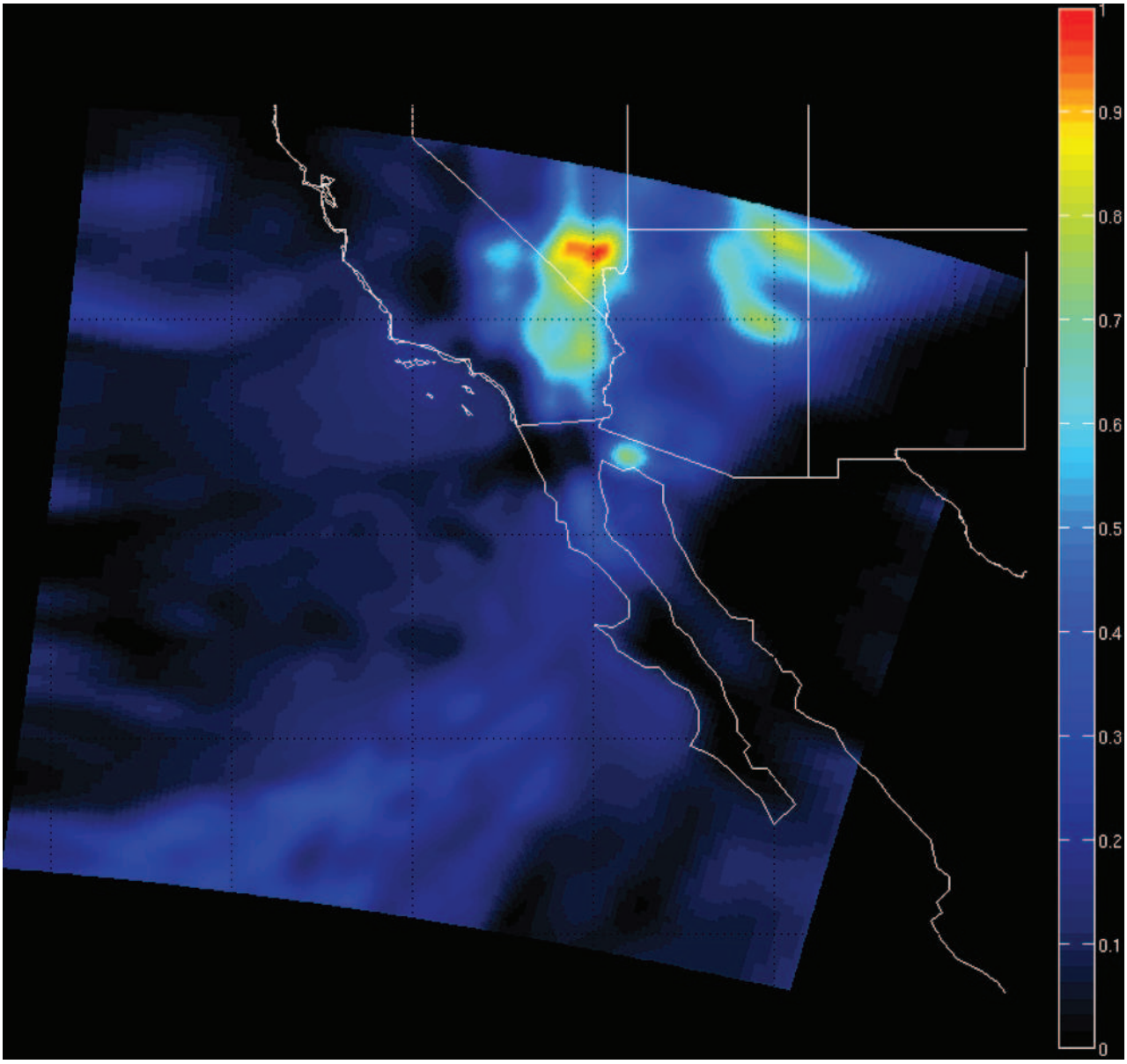

Figure 318. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

Figure 319. Result of the Maximum Likelihood model (MLE) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

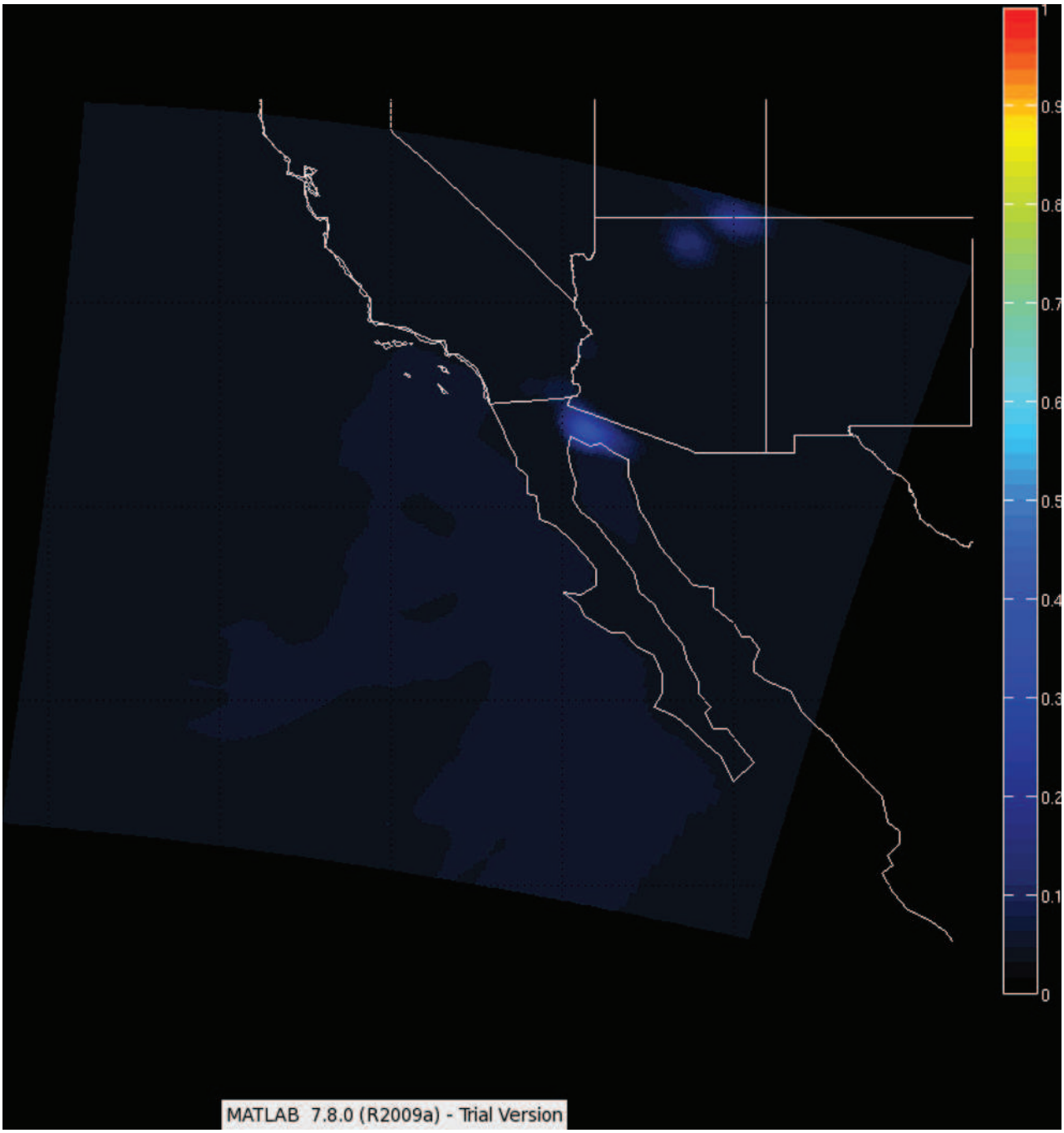

Figure 320. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

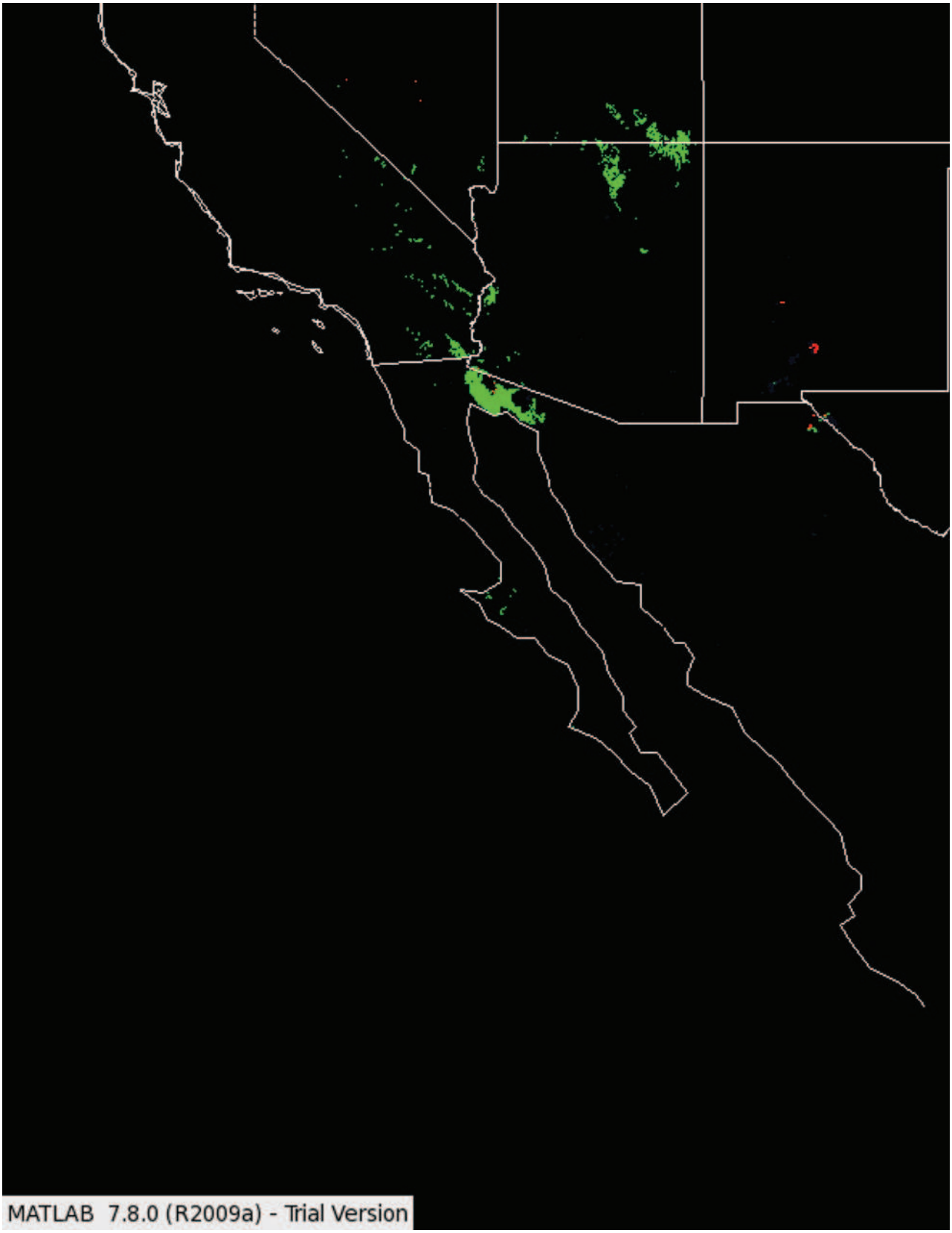

Figure 321. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2002-120 at 18:40 UTC.

Figure 322. Result of the Maximum a posteriori model (MAP) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

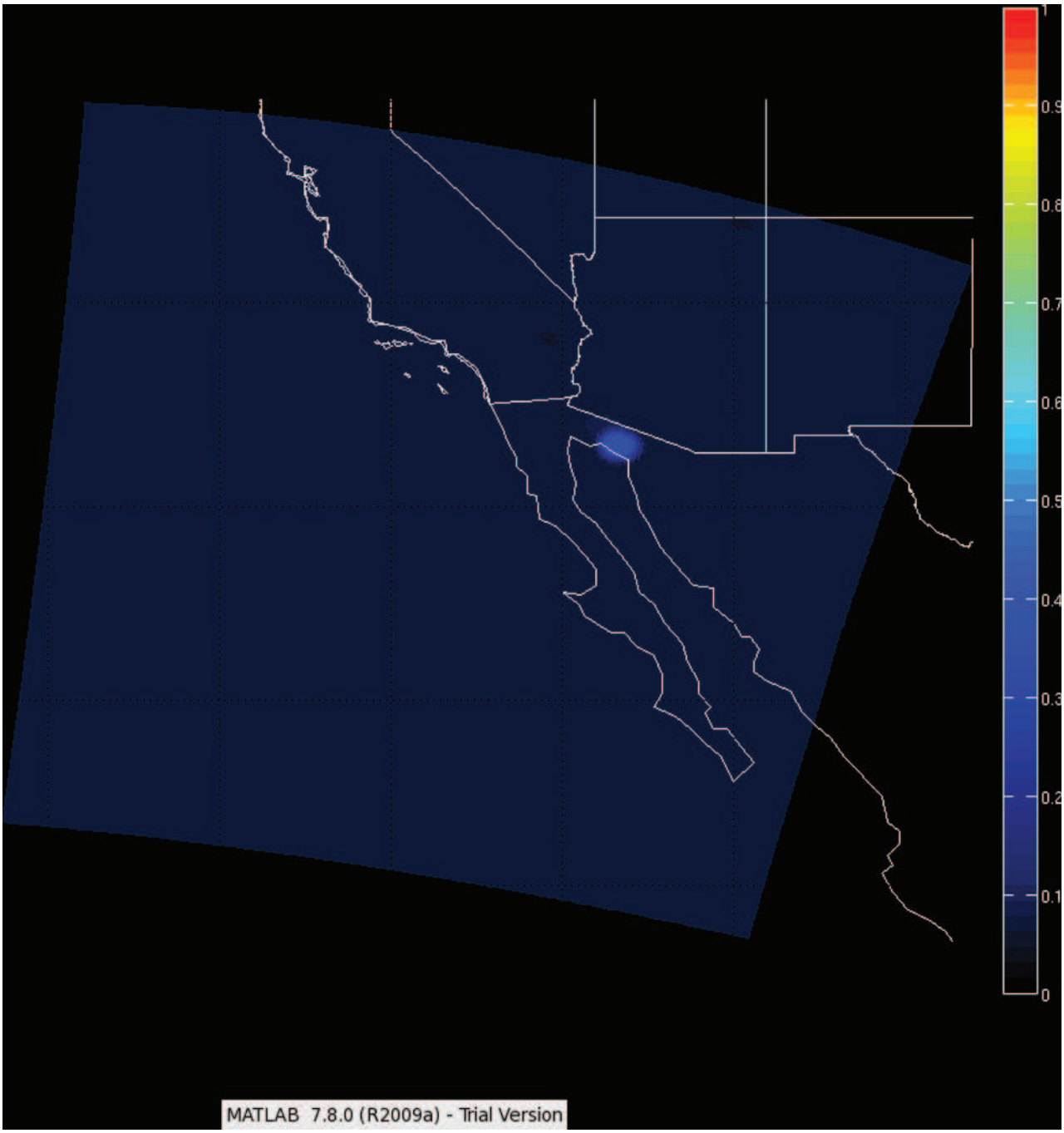

Figure 323. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

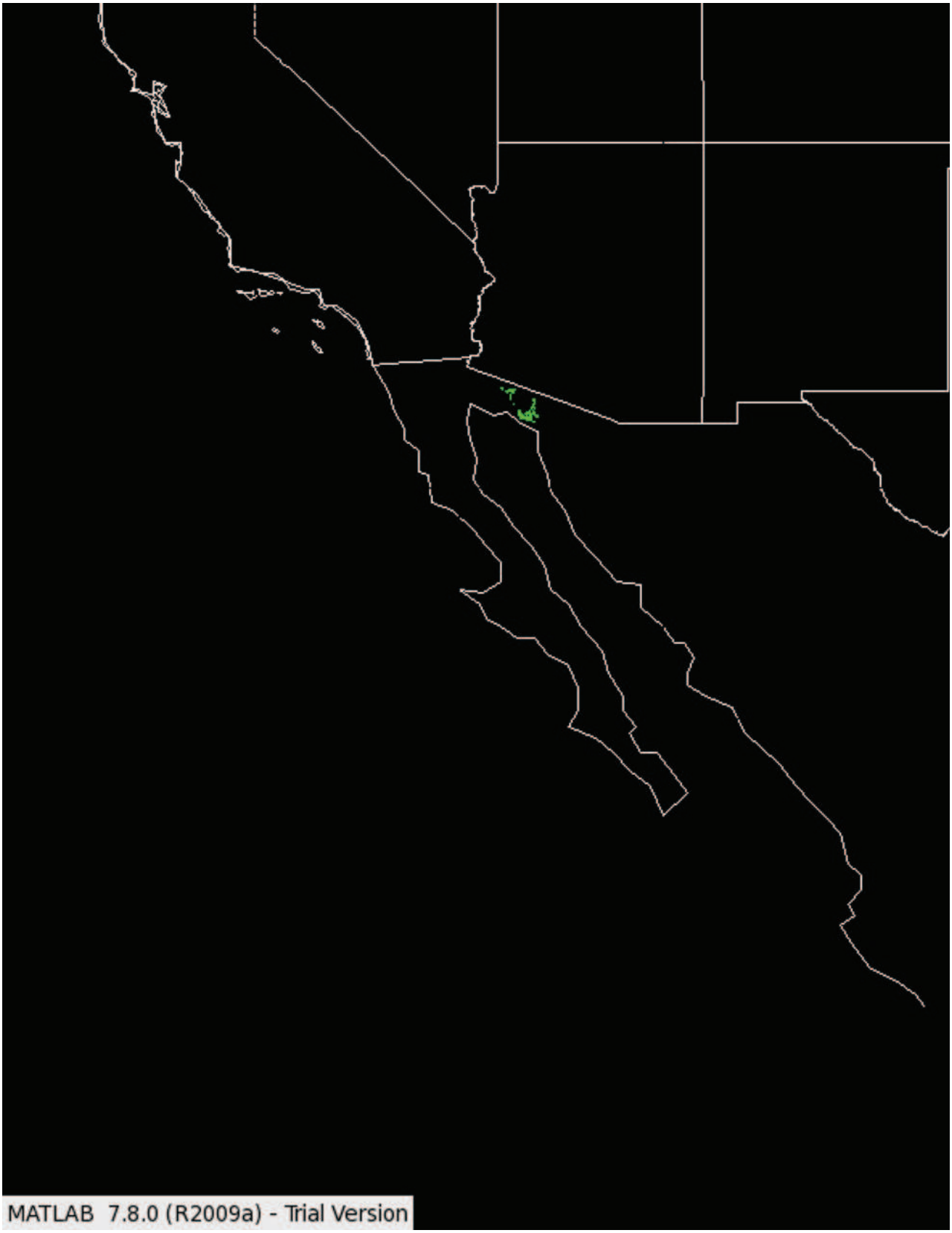

Figure 324. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2002-120 at 18:40 UTC.

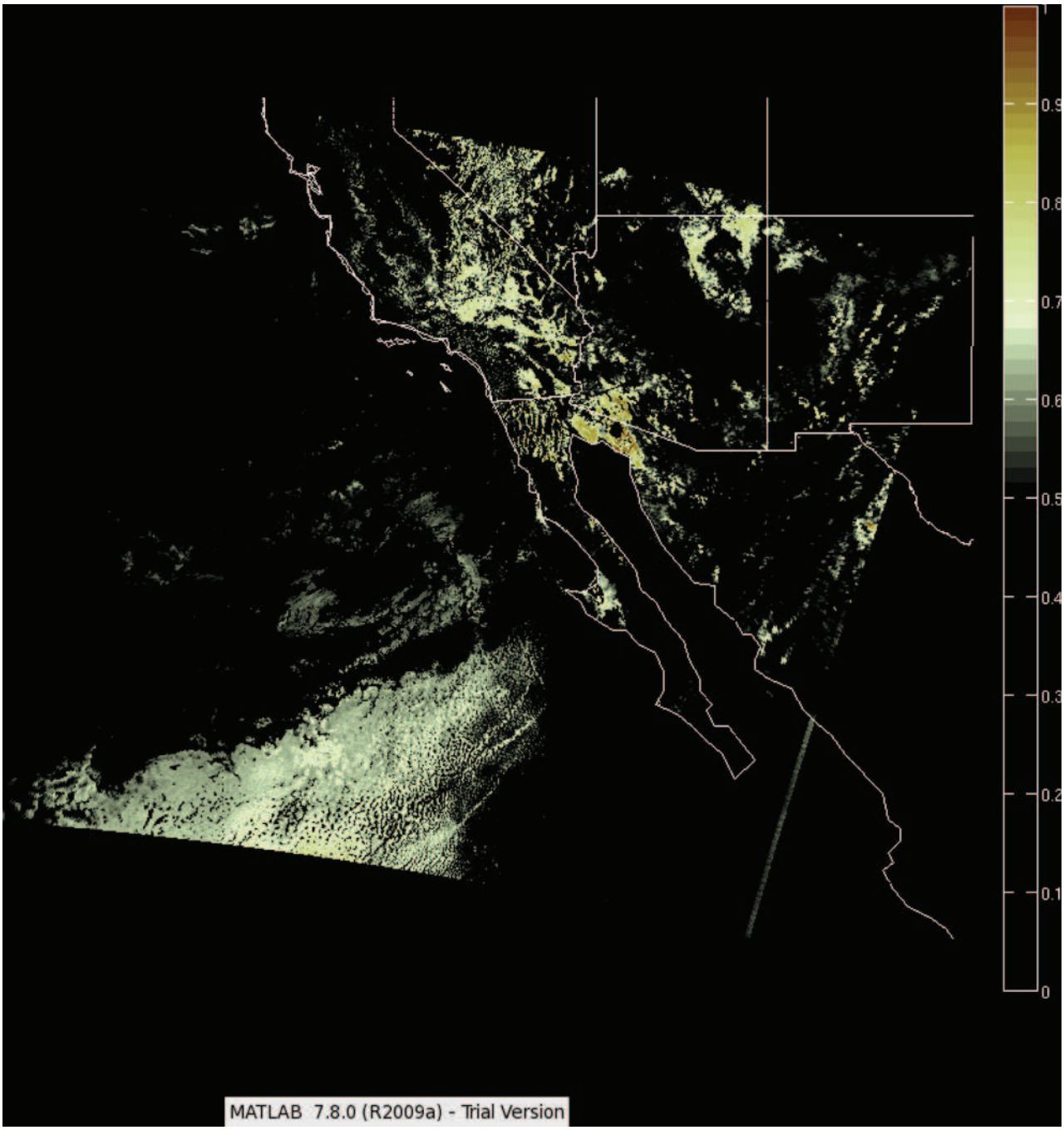

Figure 325. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

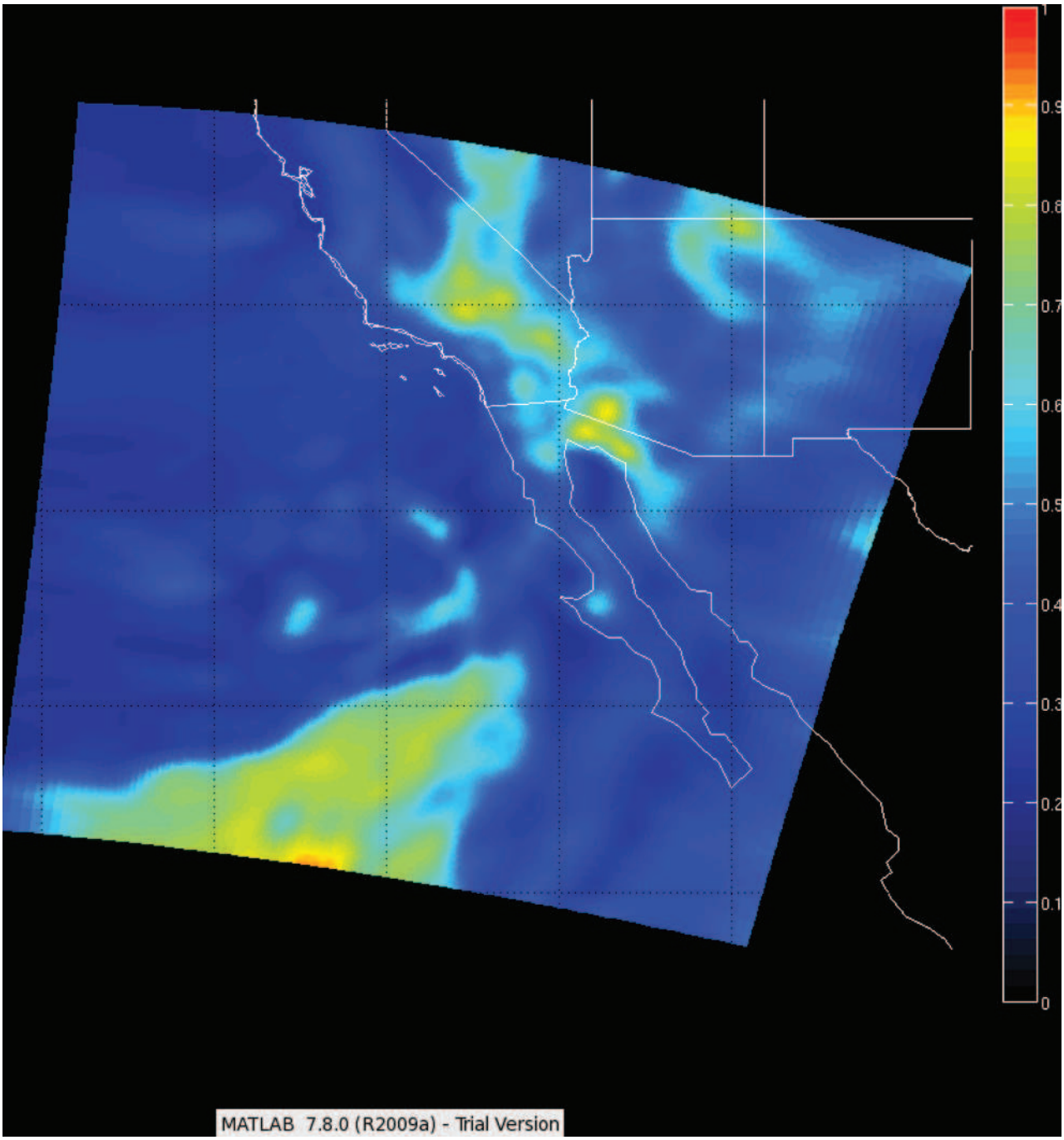

Figure 326. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

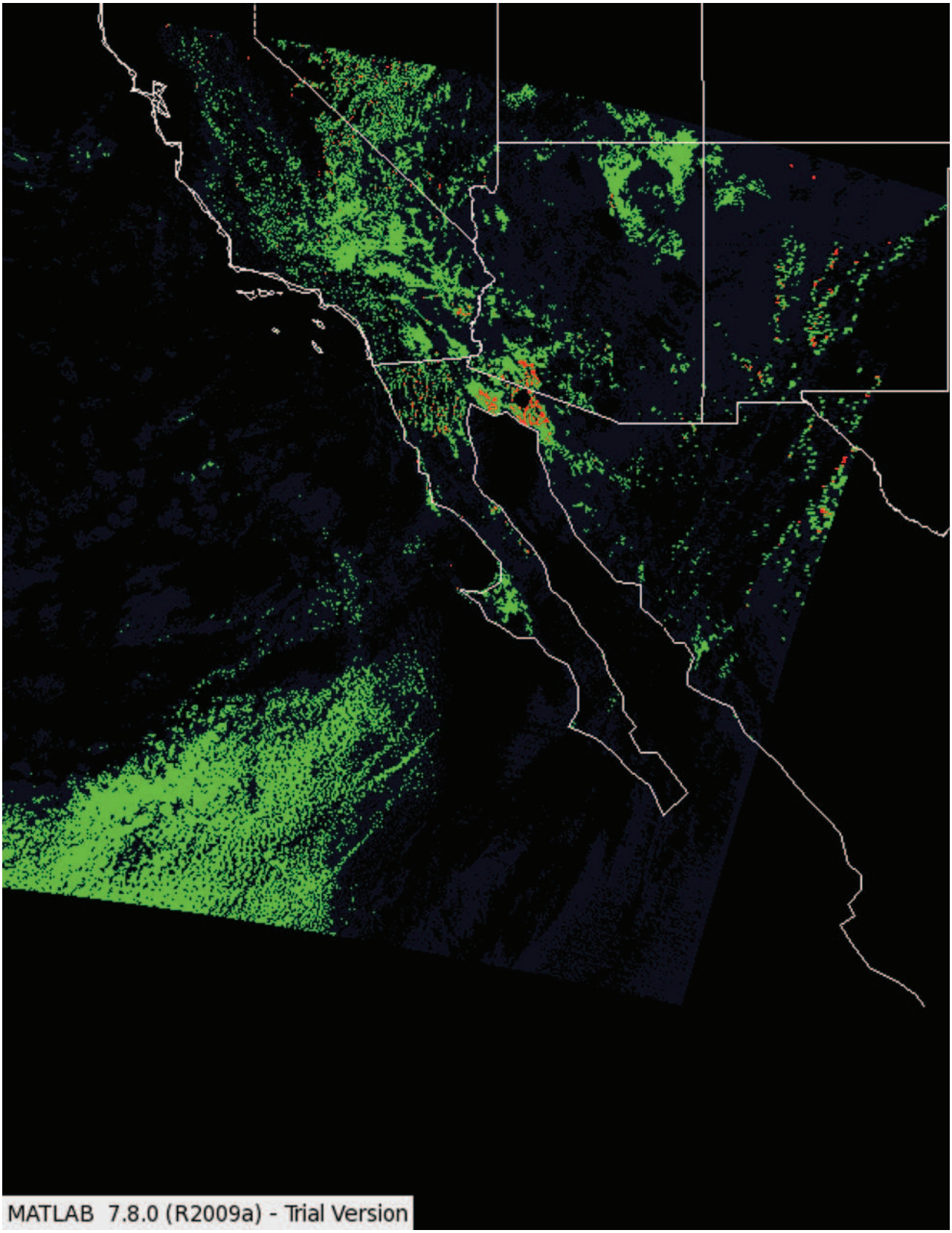

Figure 327. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2002-120 at 18:40 UTC.

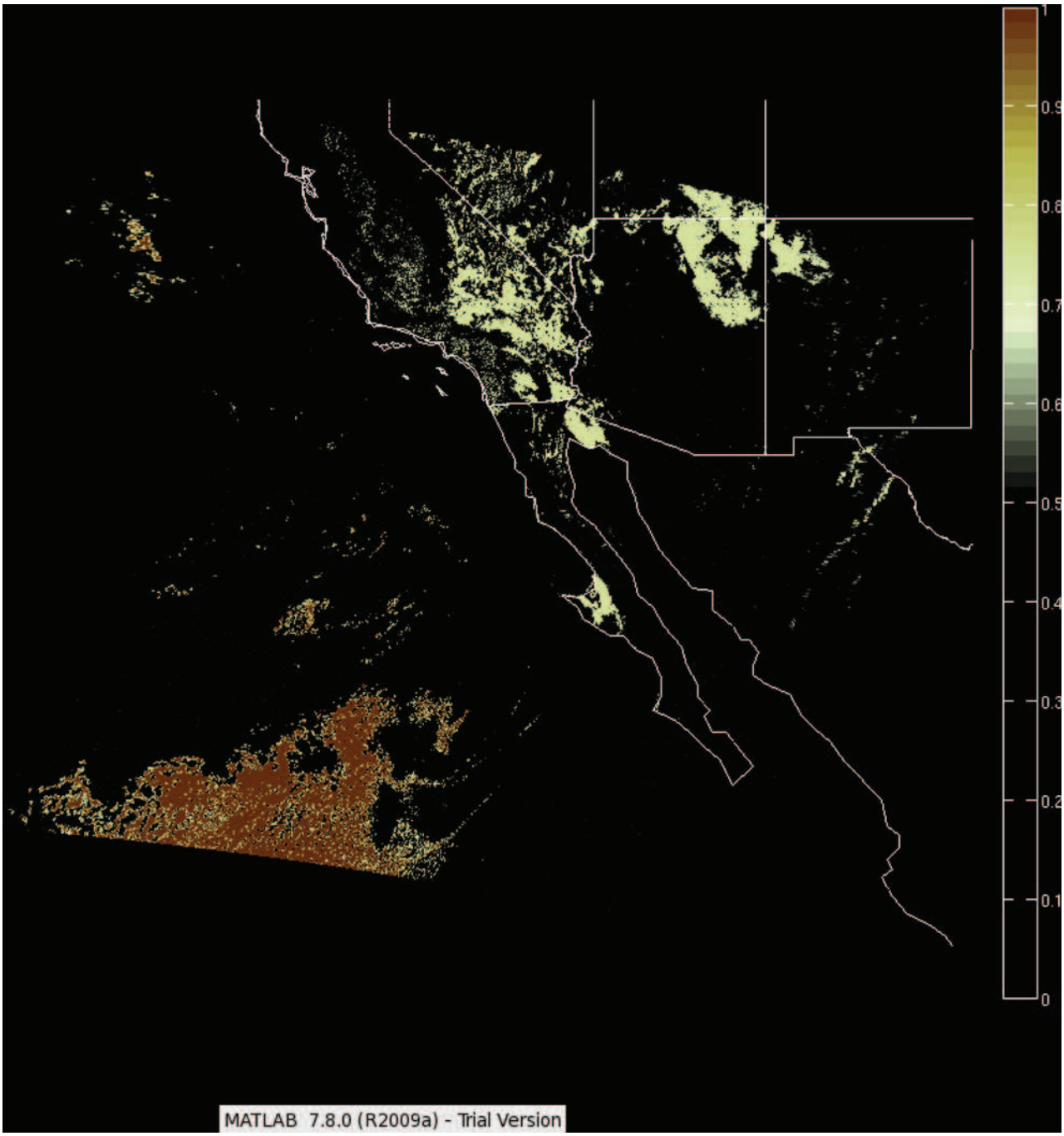

Figure 328. Result of the Probabilistic Neural Network (PNN) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

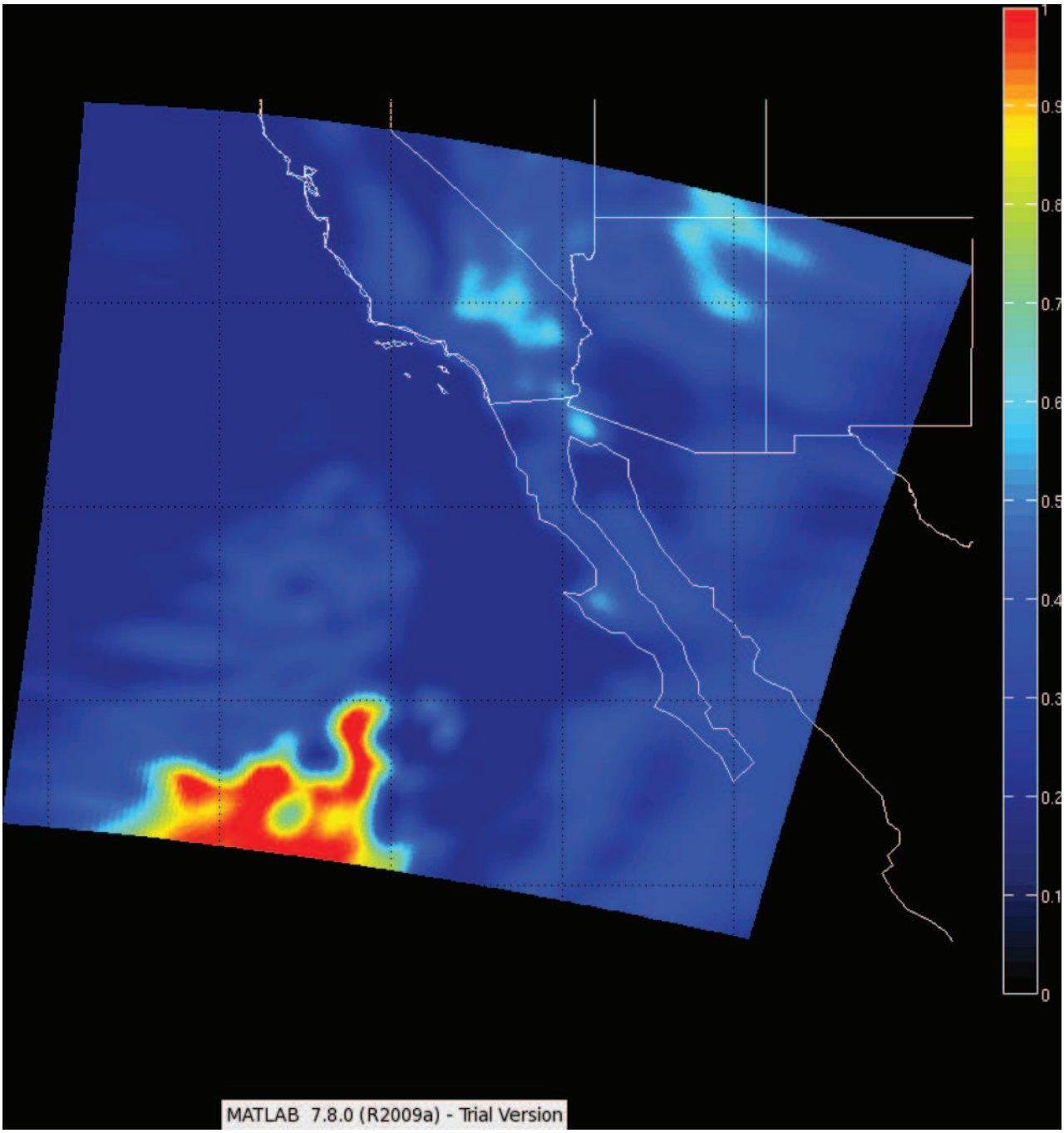

Figure 329. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2002-120 at 18:40 UTC. The colorbar shows the index of probability.

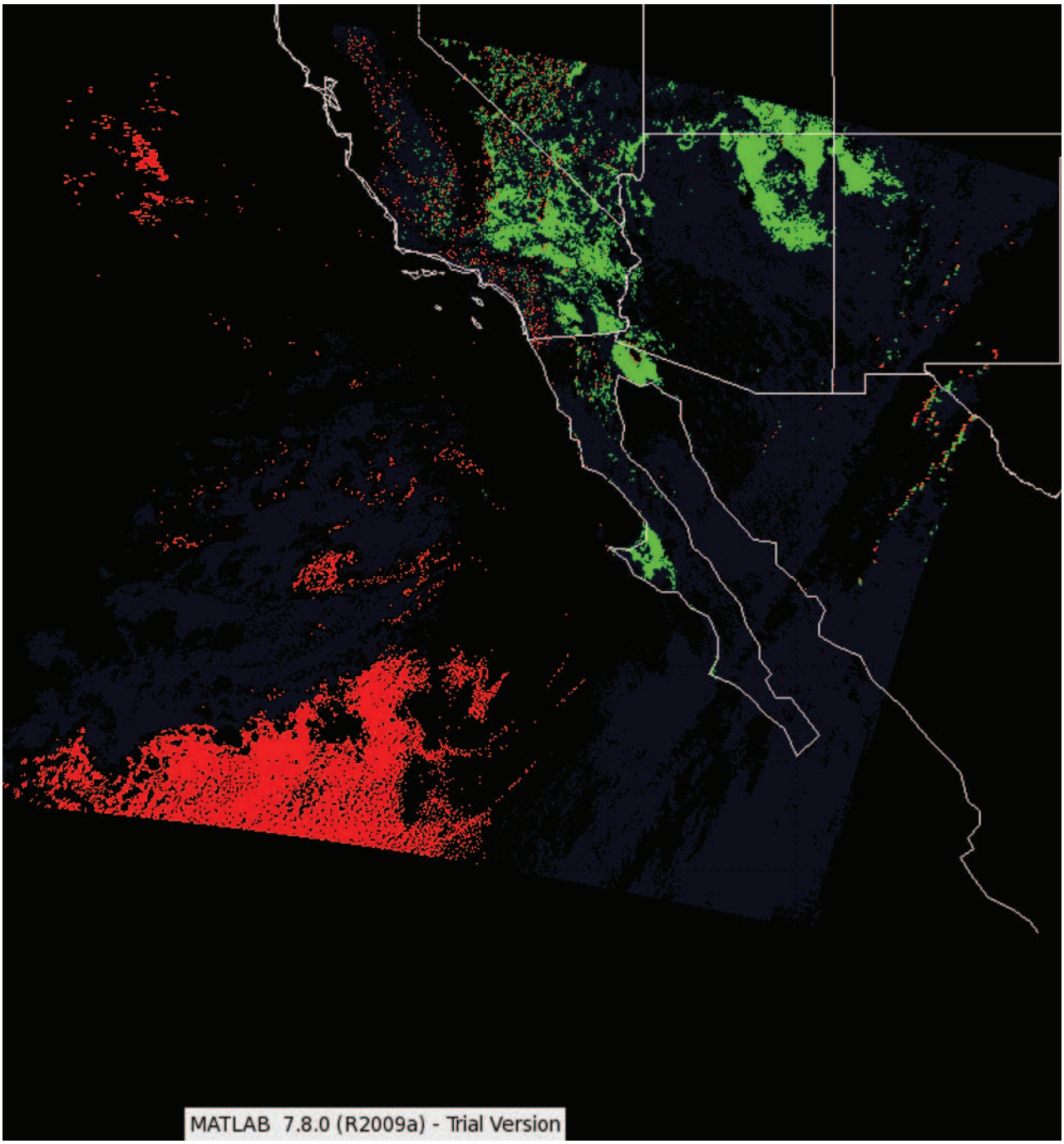

Figure 330. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2002-120 at 18:40 UTC.

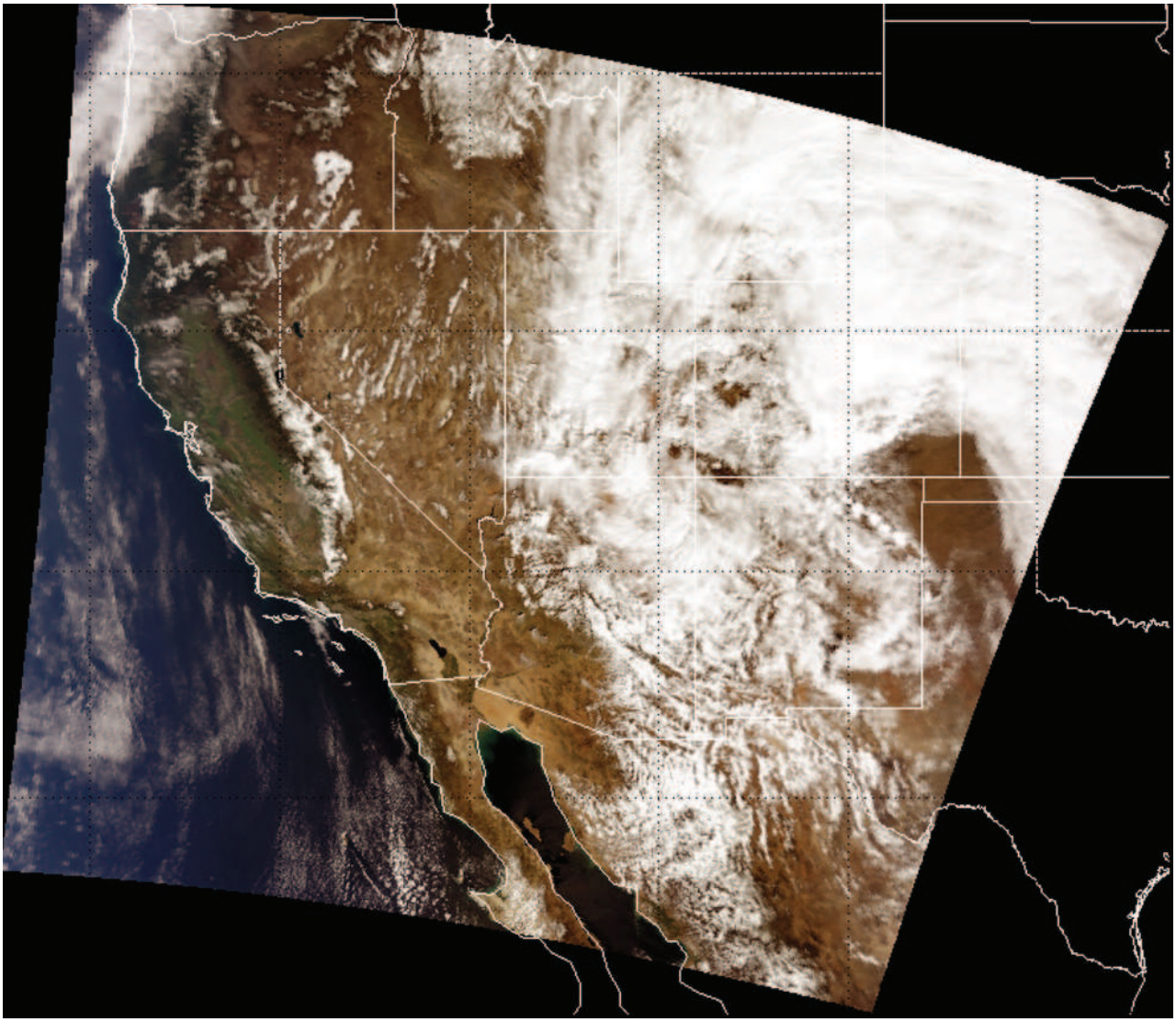

Figure 331. True color image for the event of 2003-077 at 18:25 UTC.

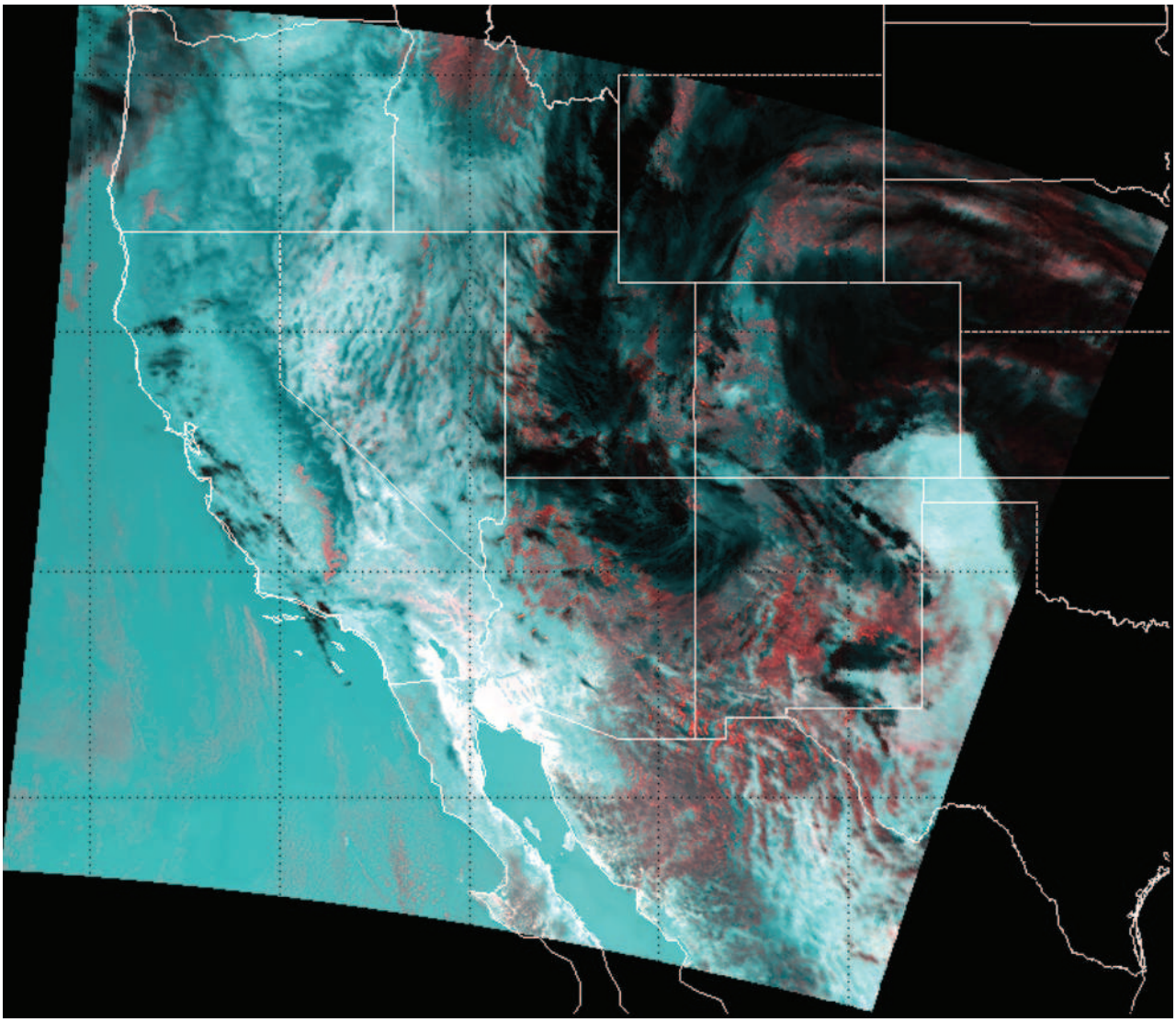

Figure 332. False color image for the event of 2003-077 at 18:25 UTC.

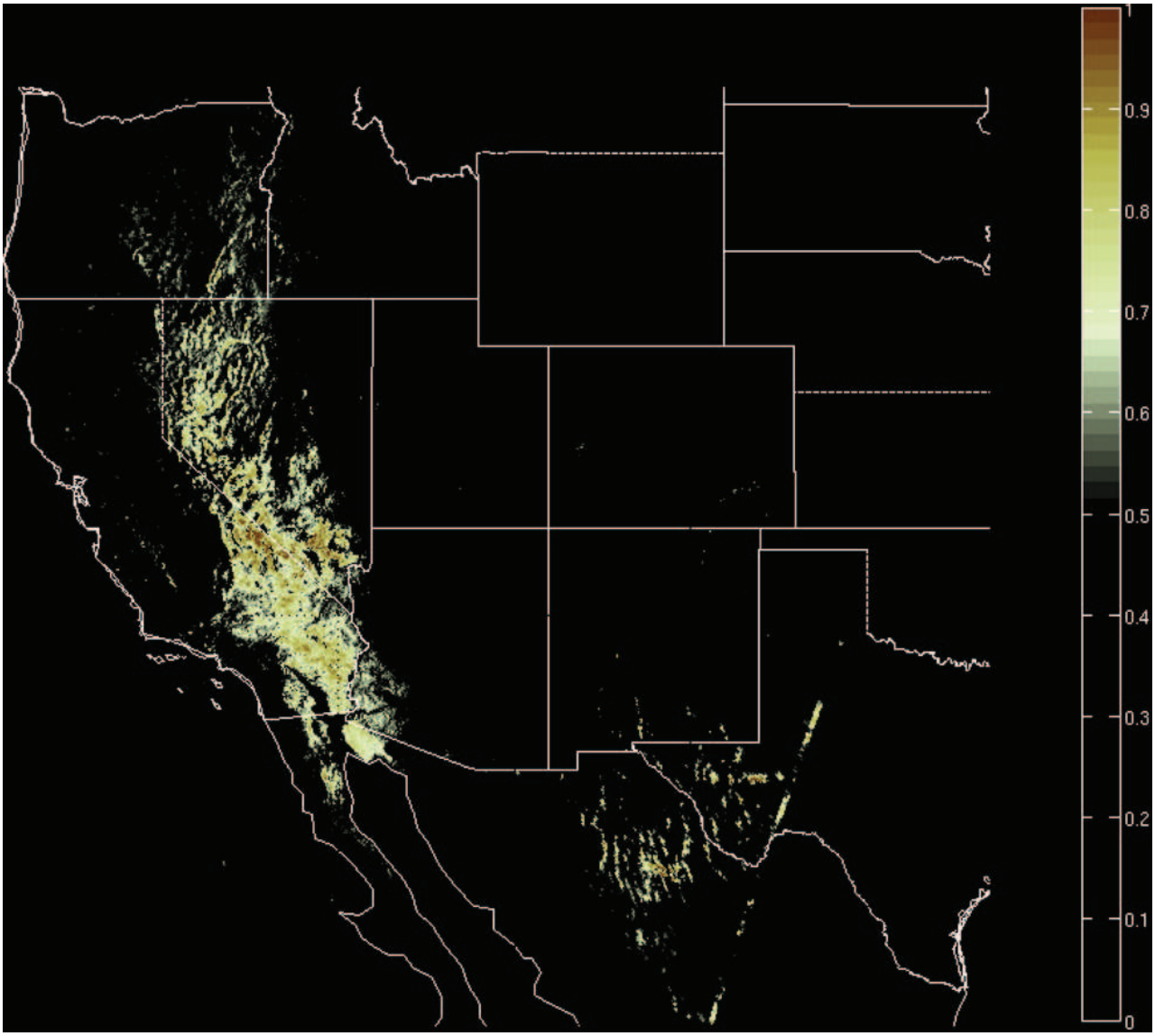

Figure 333. Result of the probabilistic display model (PD) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

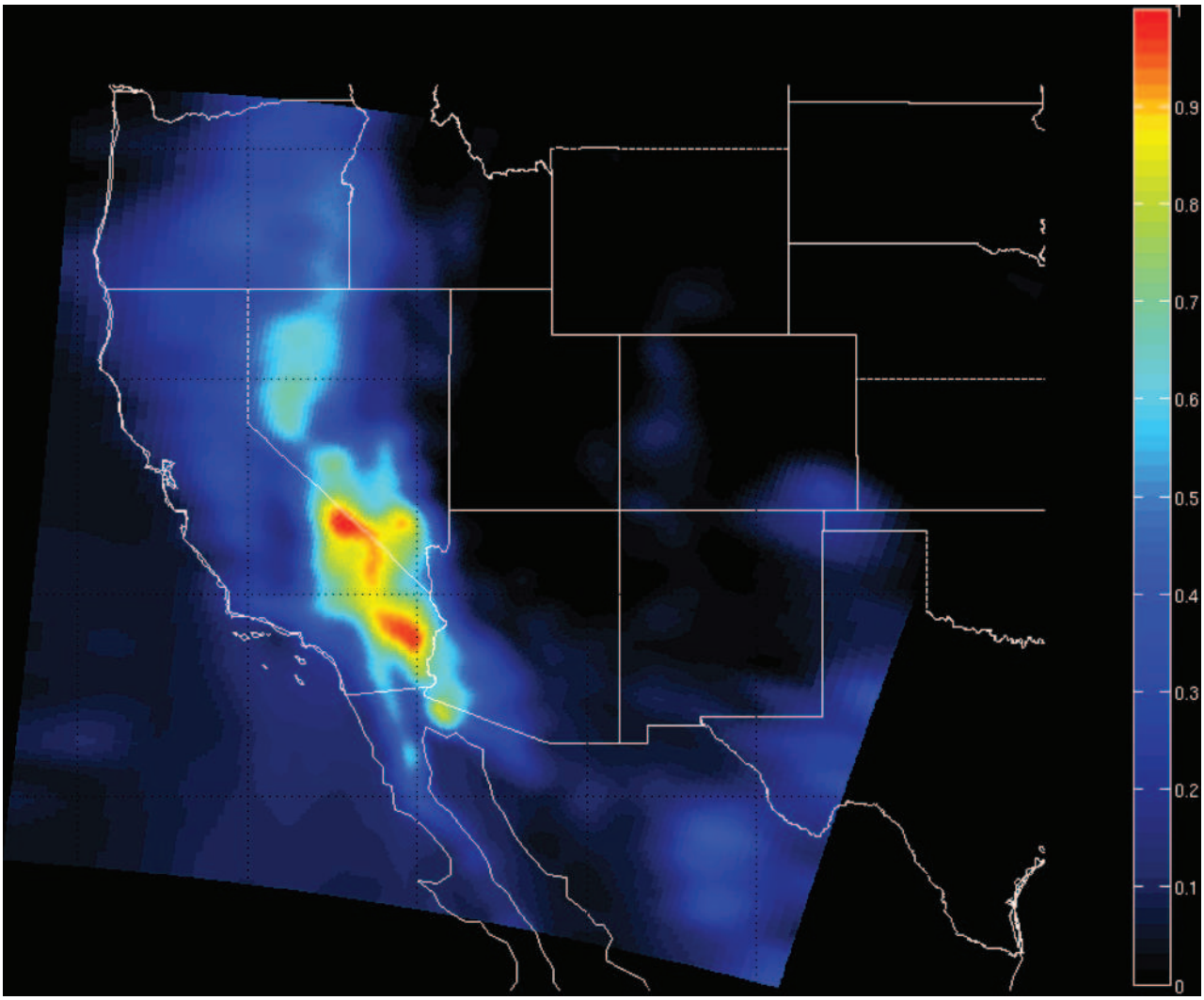

Figure 334. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

Figure 335. Result of the Maximum Likelihood model (MLE) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

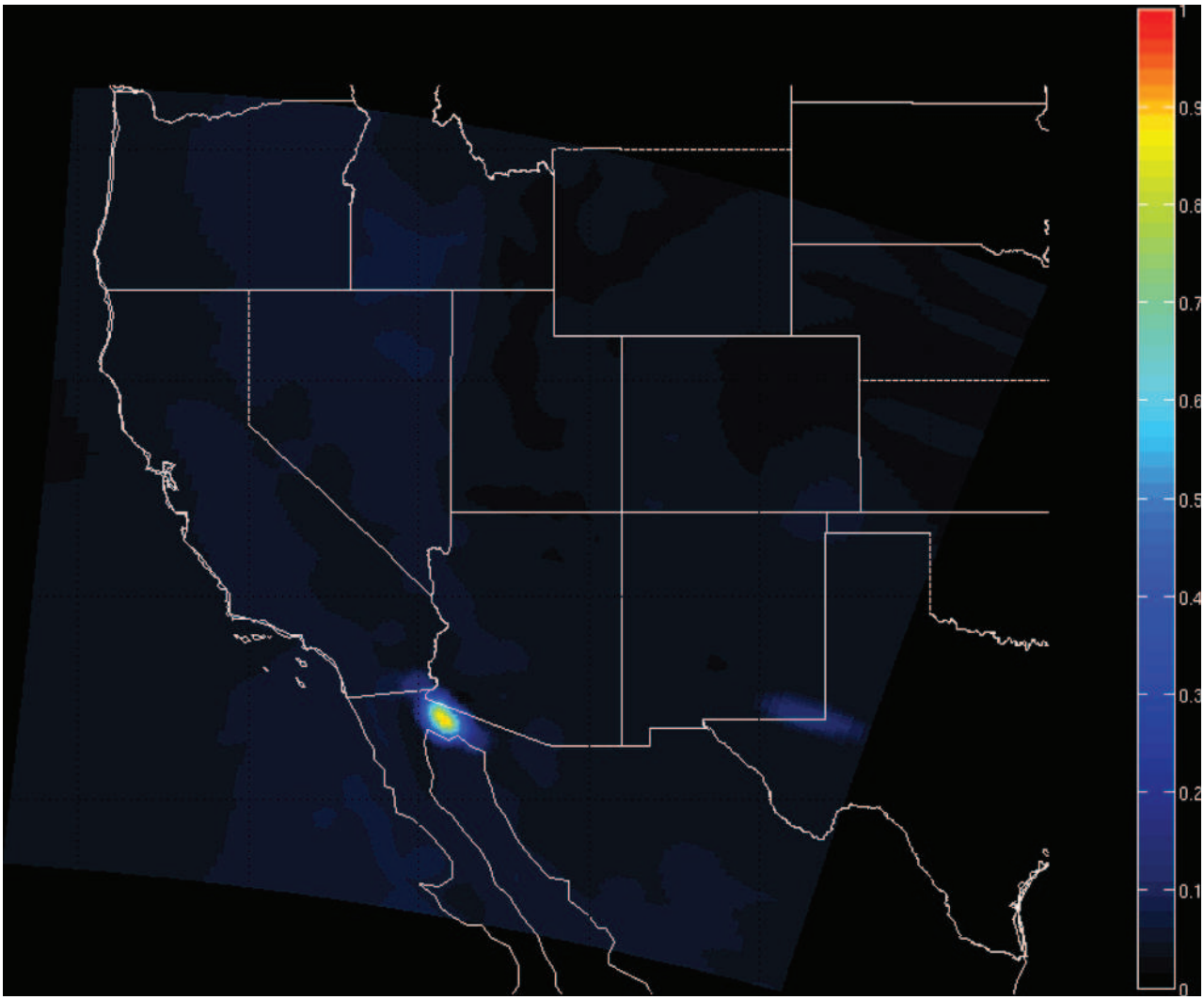

Figure 336. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

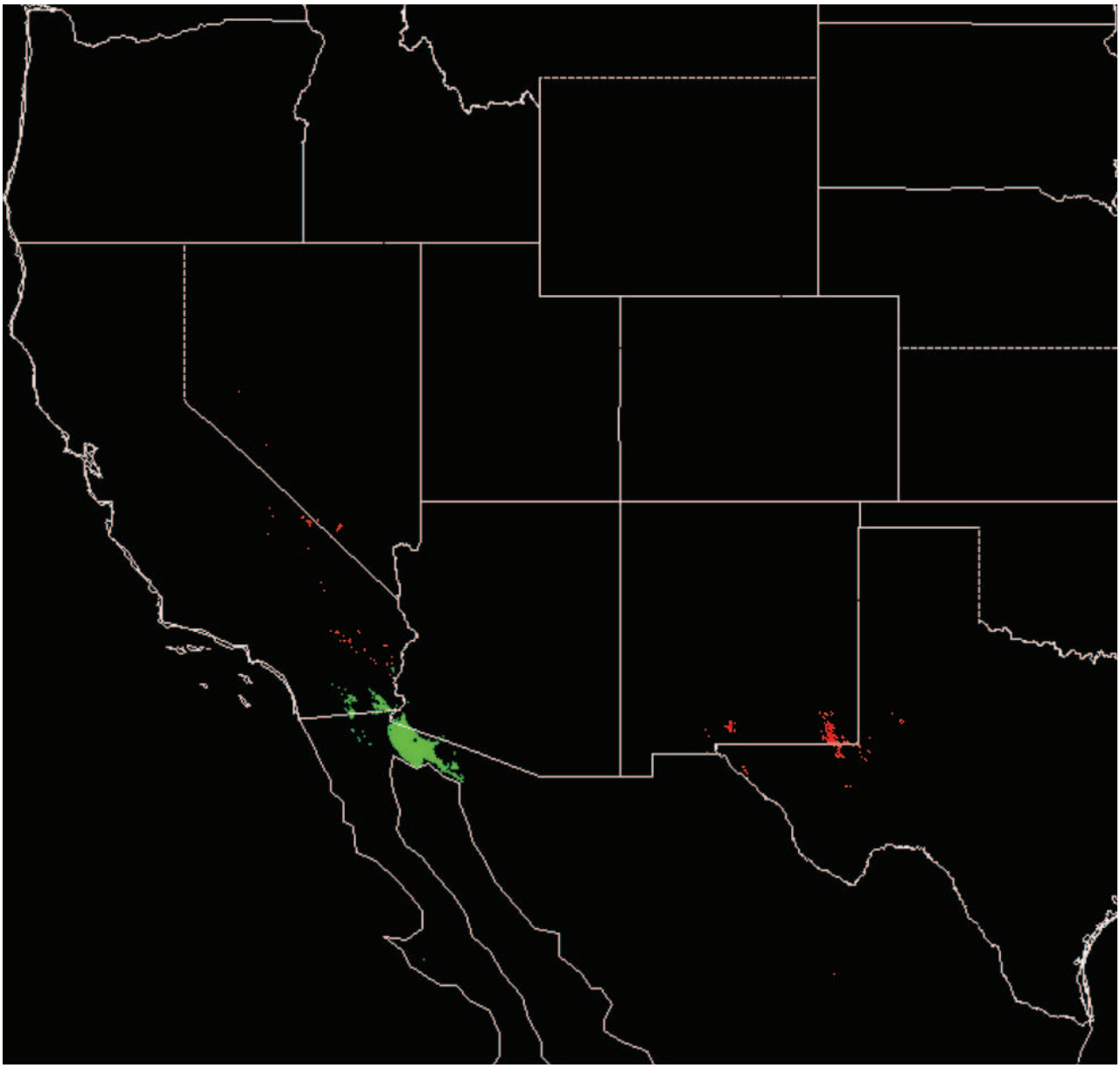

Figure 337. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-077 at 18:25 UTC.

Figure 338. Result of the Maximum a posteriori model (MAP) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

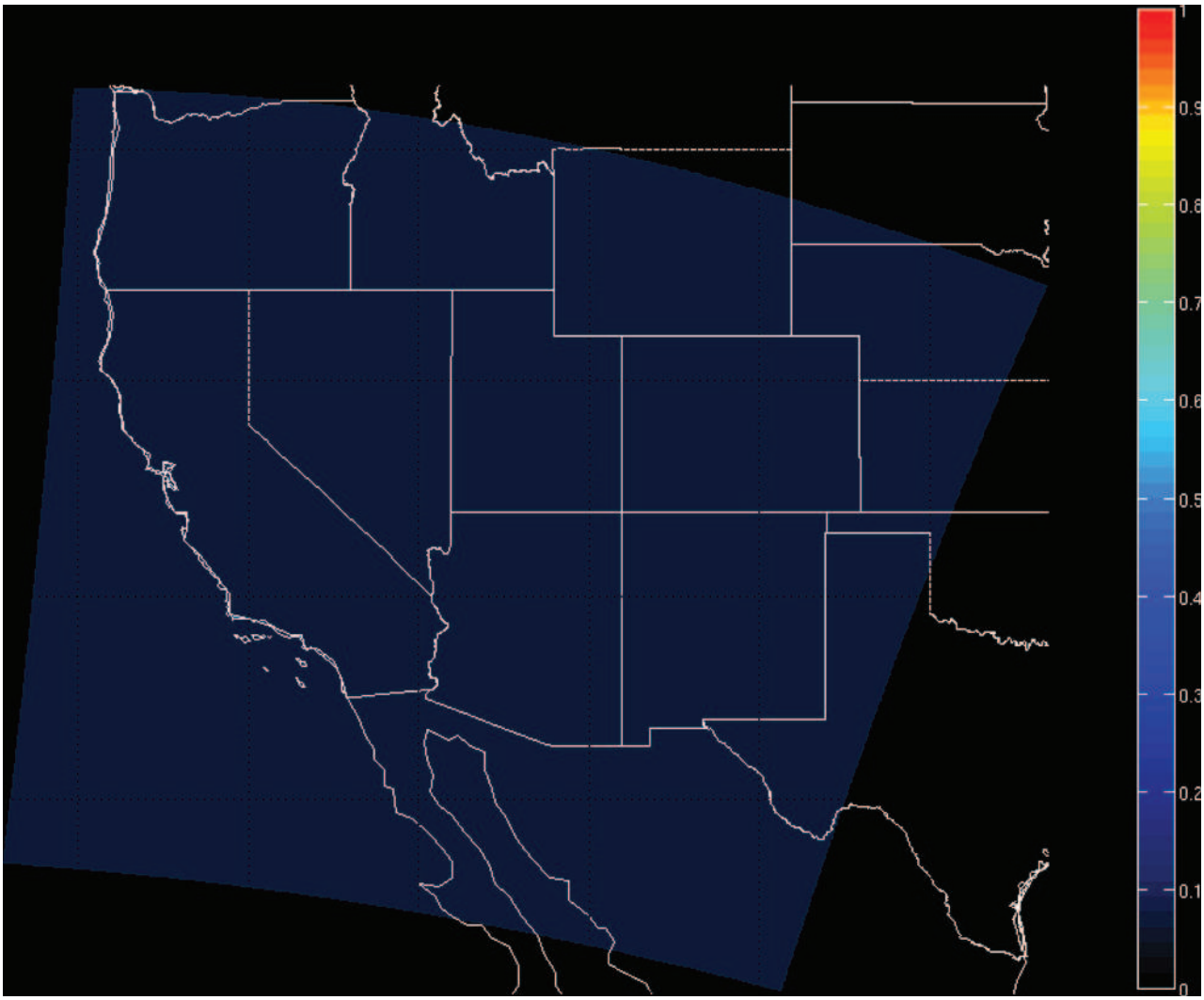

Figure 339. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

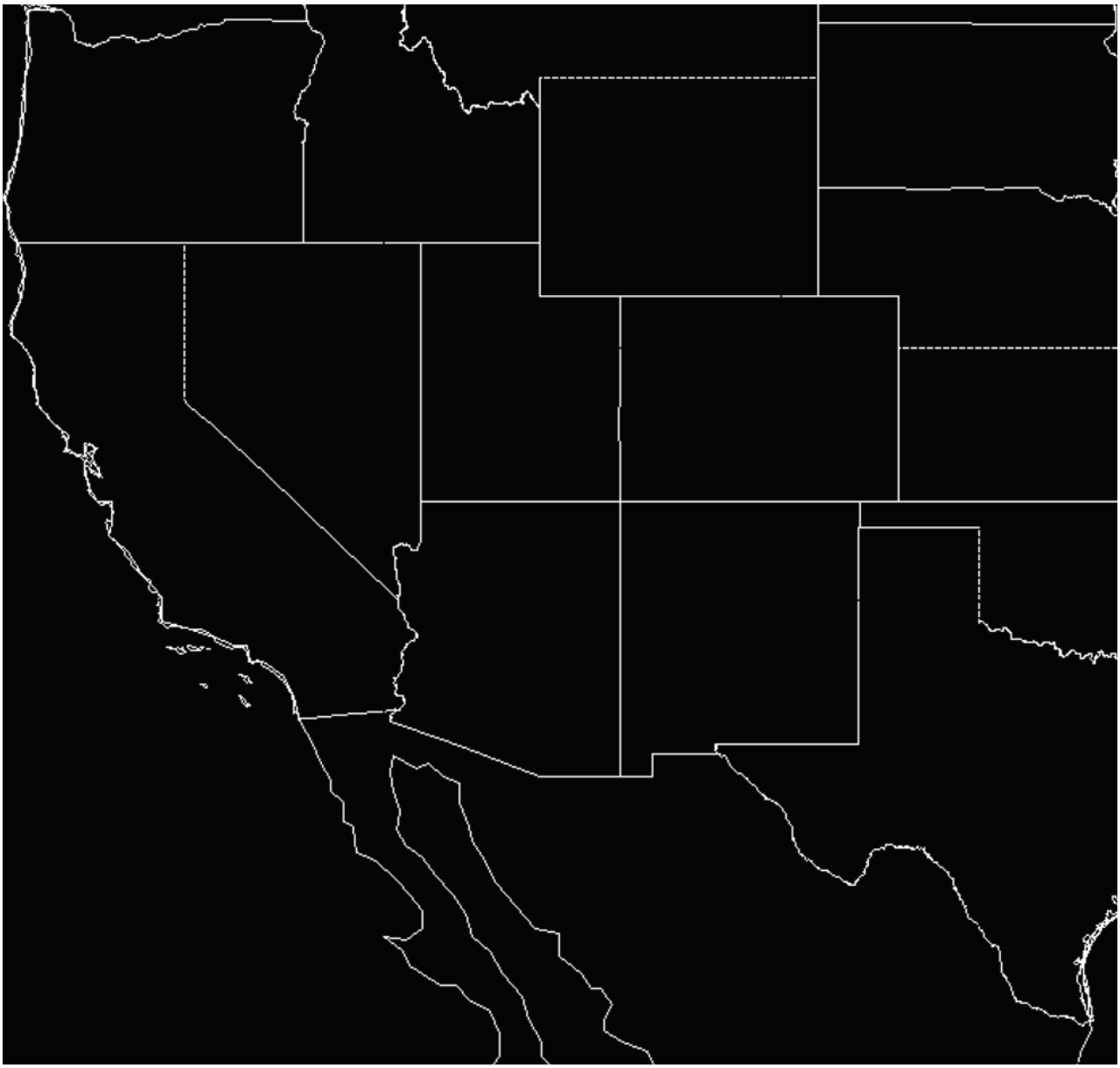

Figure 340. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-077 at 18:25 UTC.

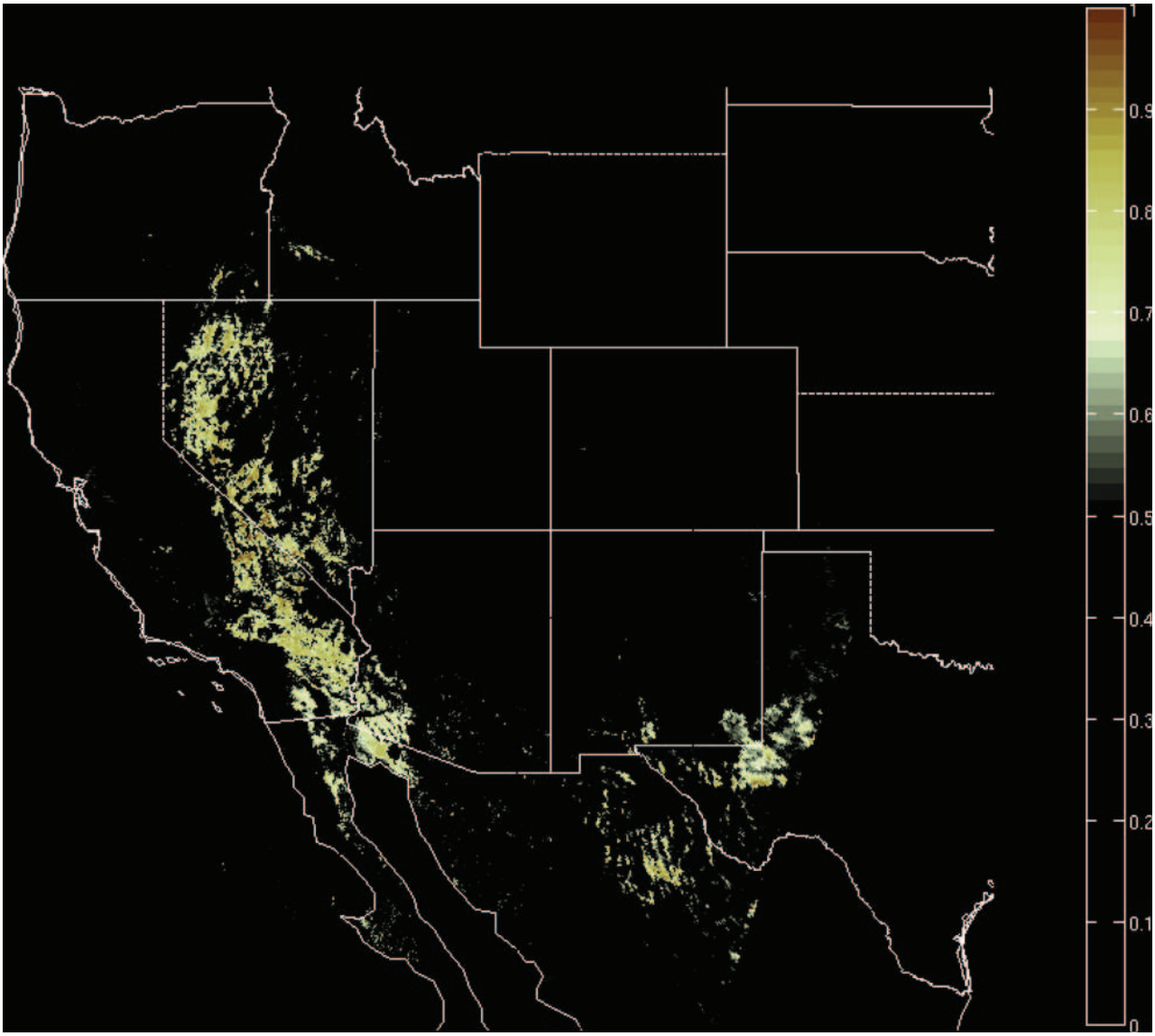

Figure 341. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

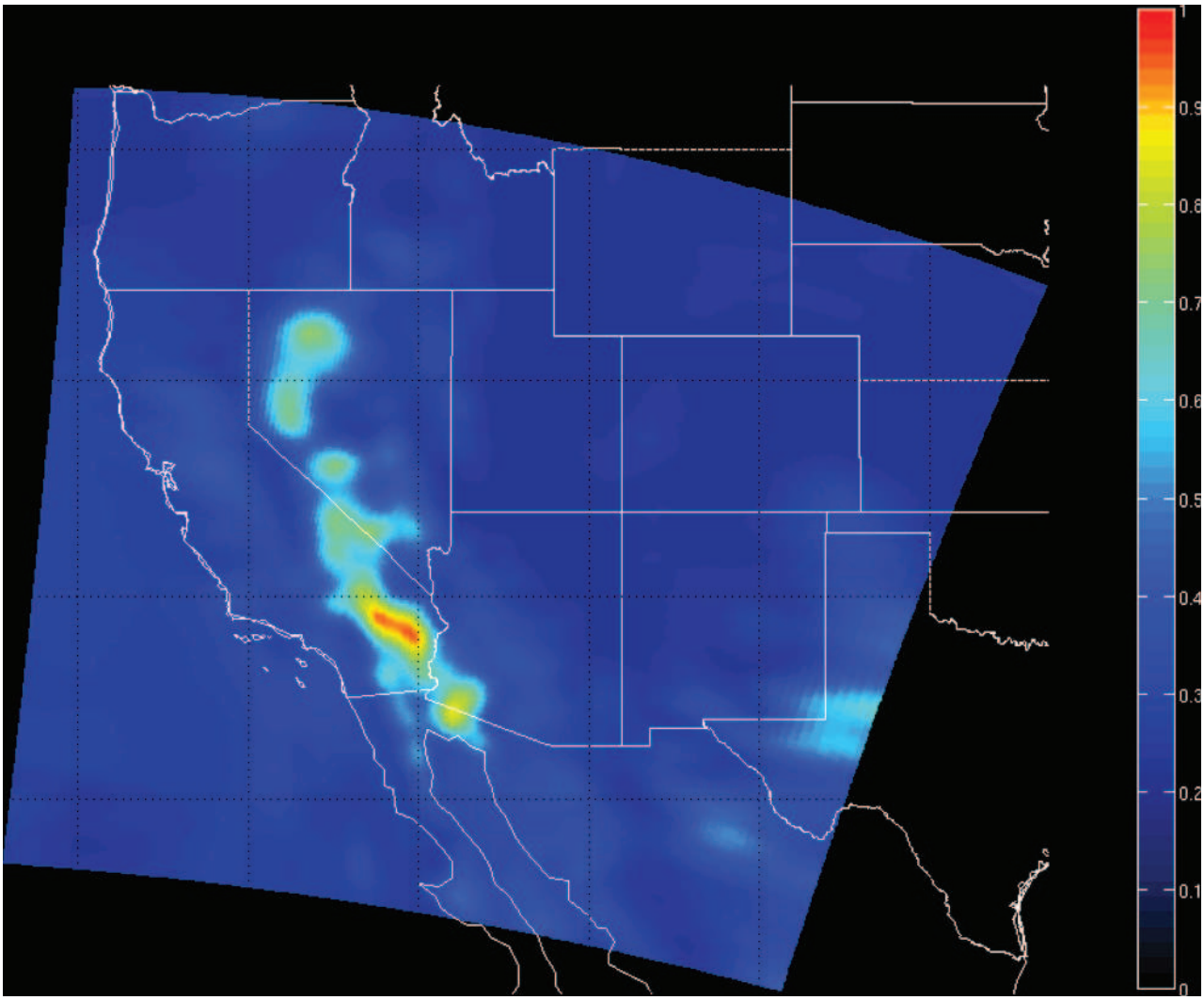

Figure 342. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

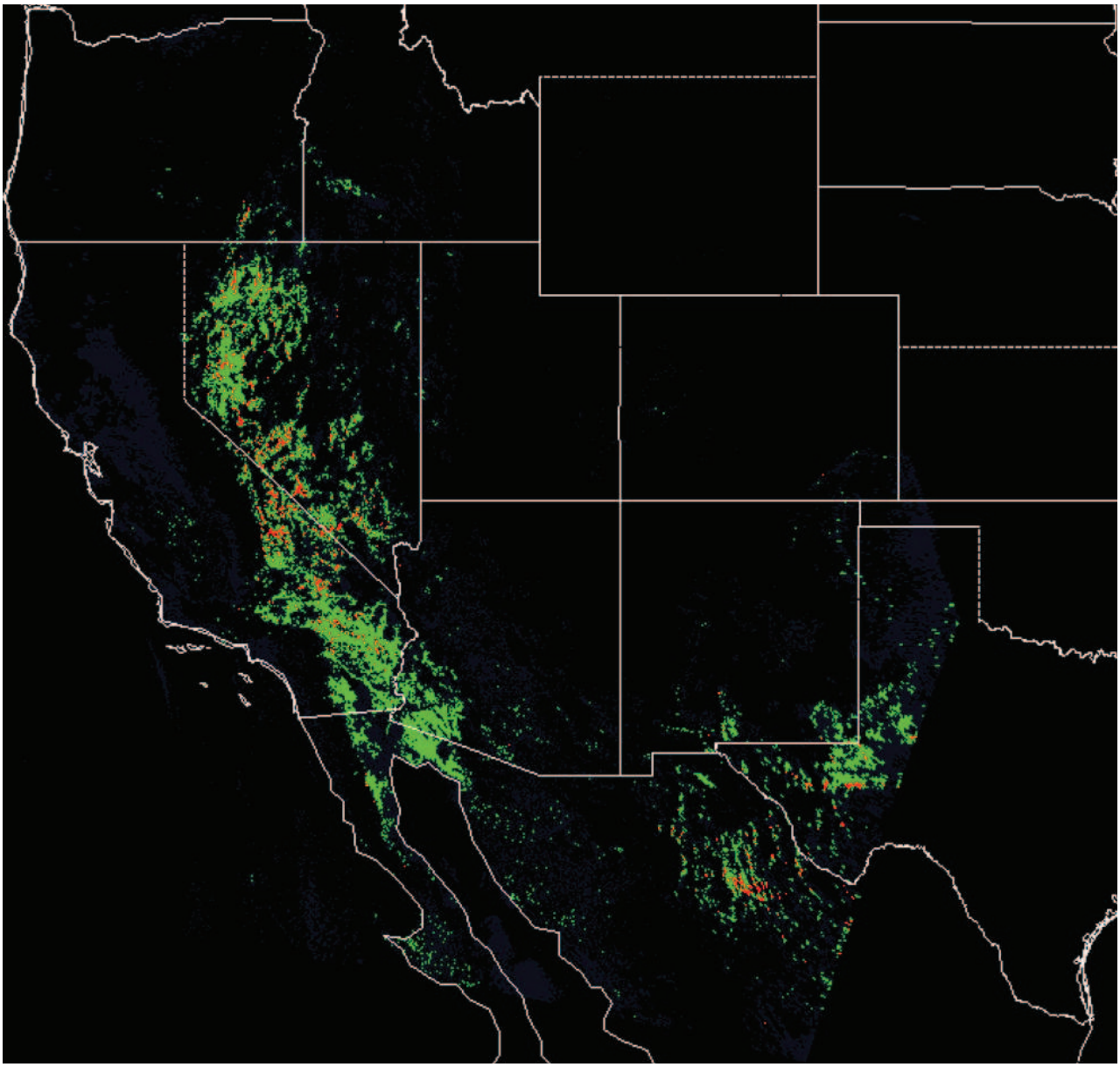

Figure 343. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-077 at 18:25 UTC.

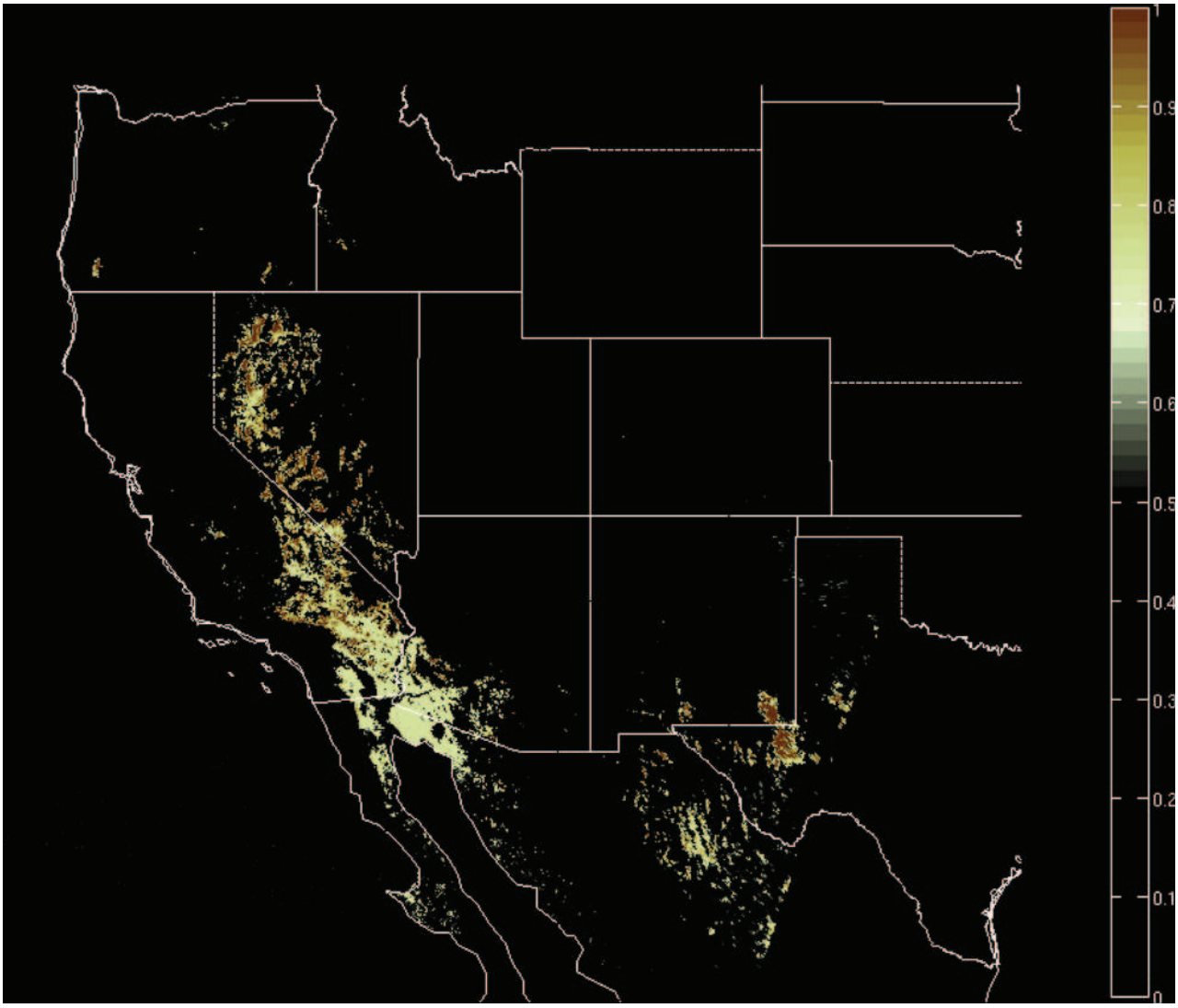

Figure 344. Result of the Probabilistic Neural Network (PNN) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

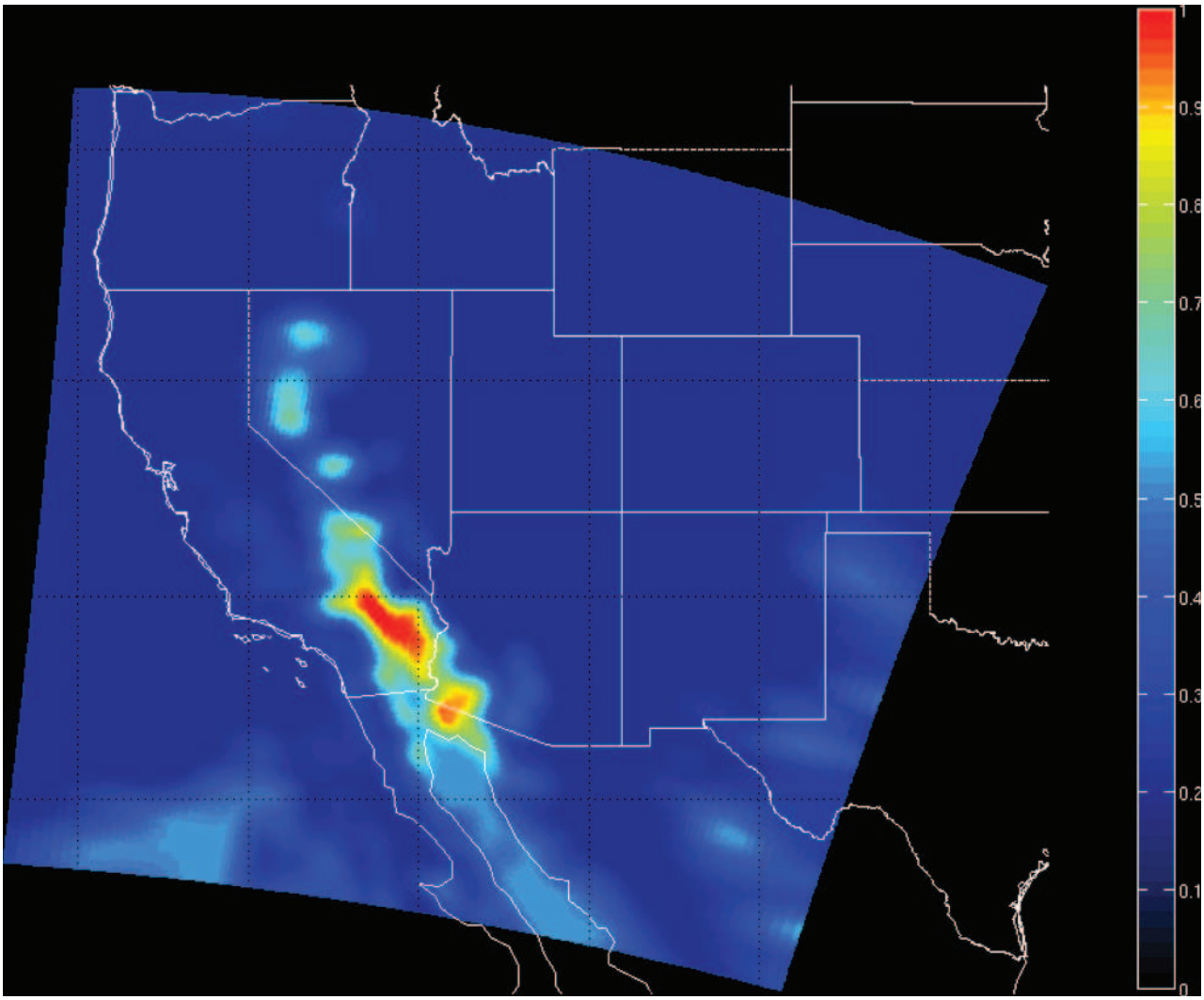

Figure 345. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-077 at 18:25 UTC. The colorbar shows the index of probability.

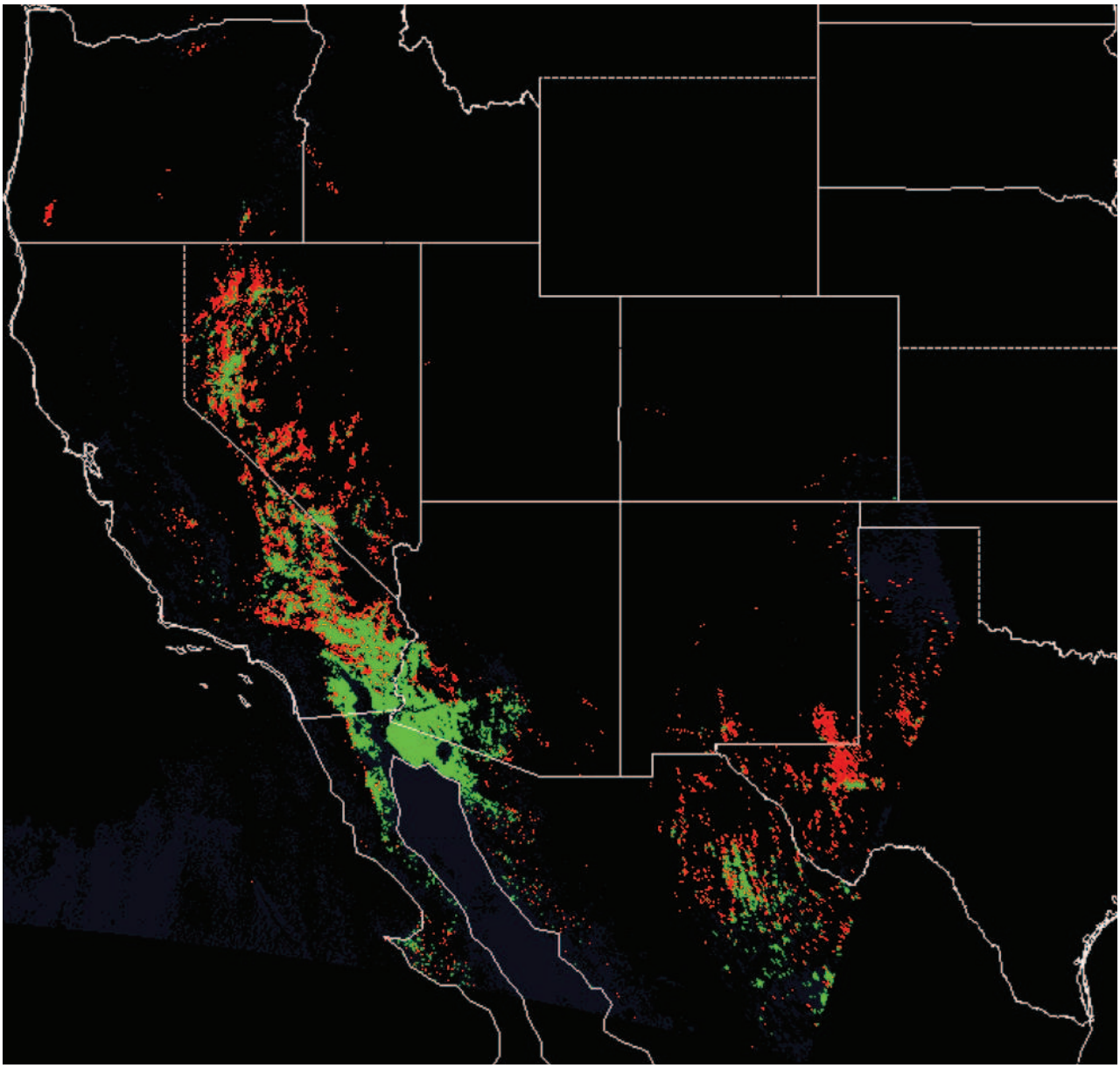

Figure 346. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-077 at 18:25 UTC.

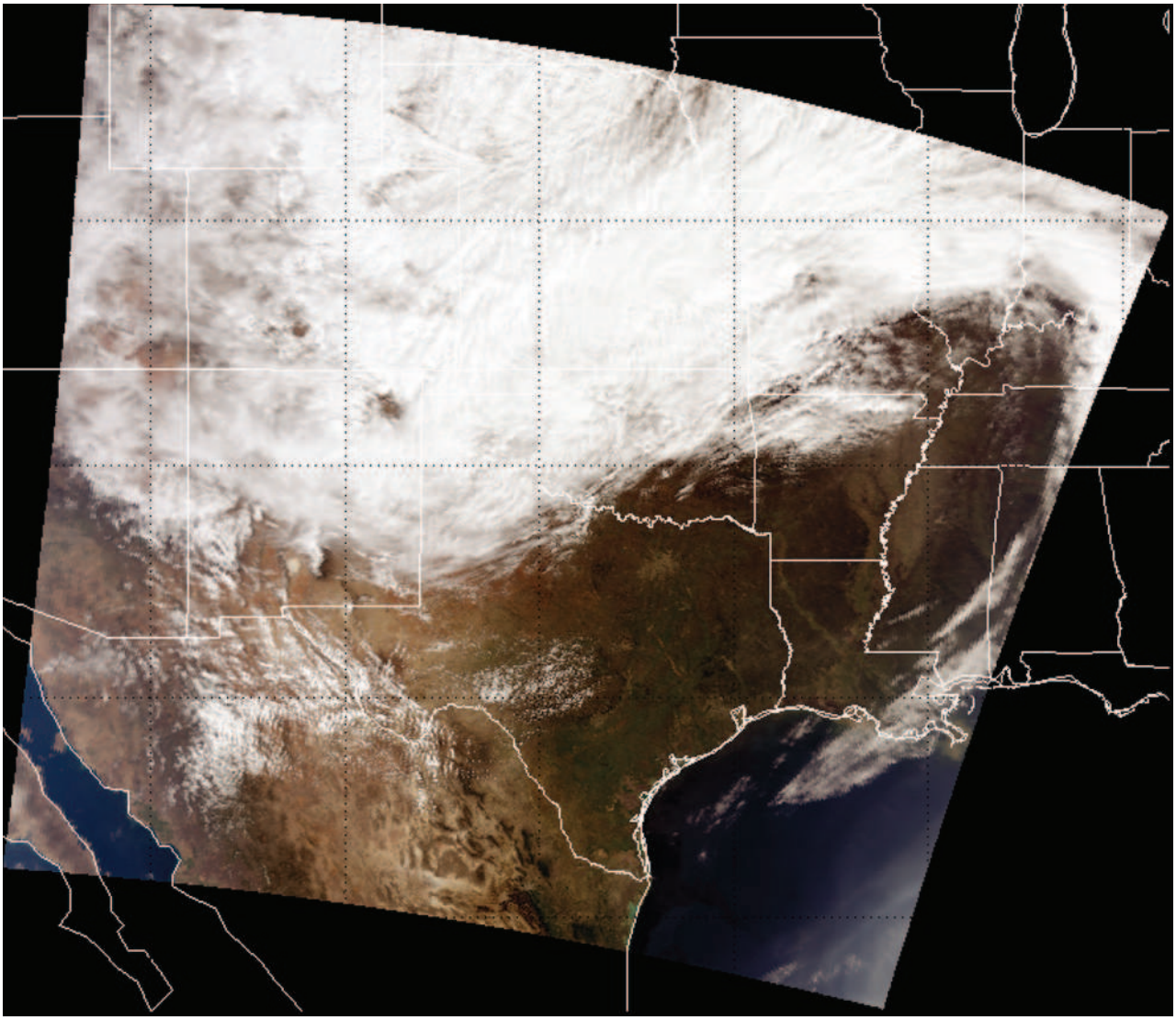

Figure 347. True color image for the event of 2003-078 at 17:30 UTC.

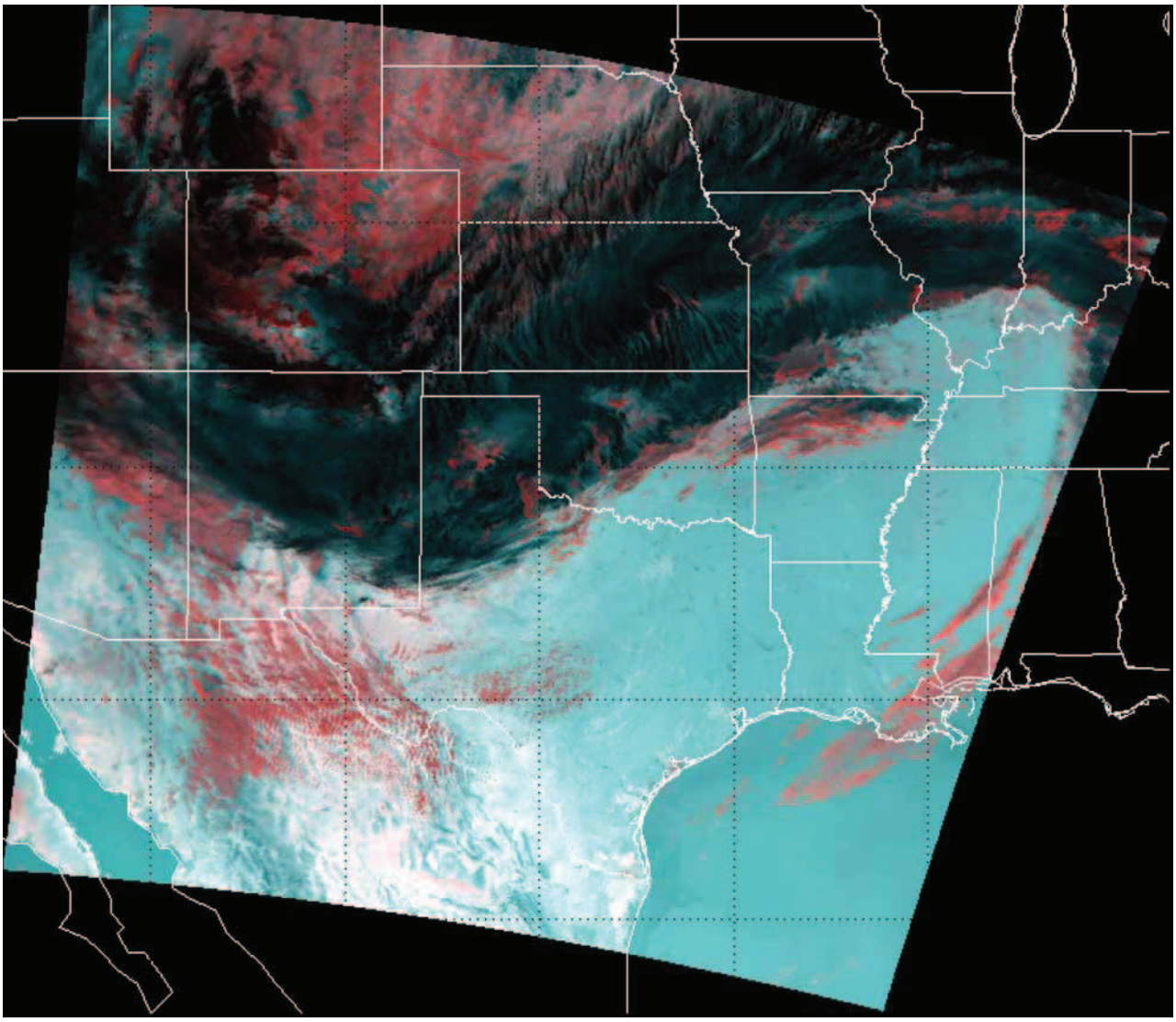

Figure 348. False color image for the event of 2003-078 at 17:30 UTC.

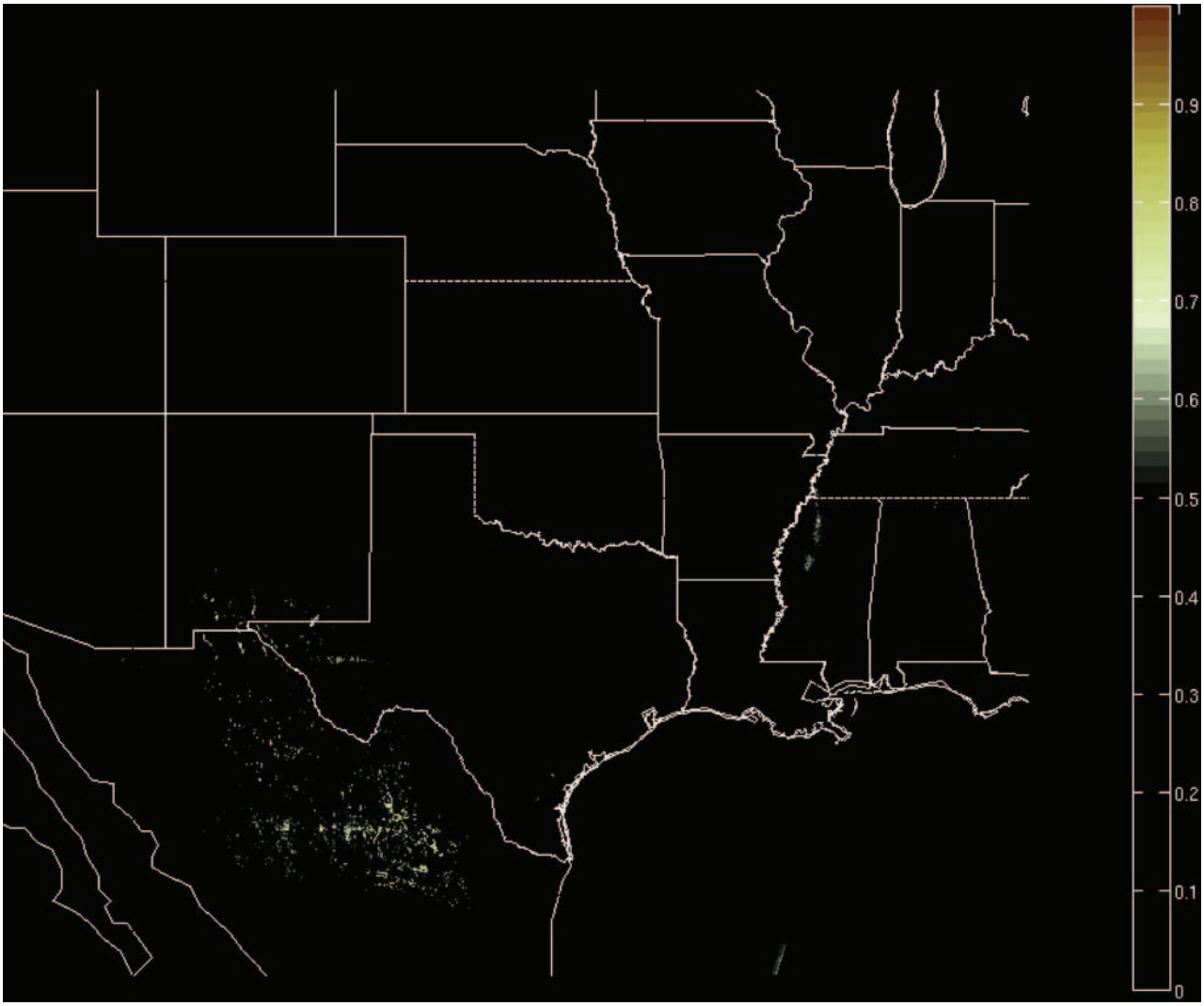

Figure 349. Result of the probabilistic display model (PD) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

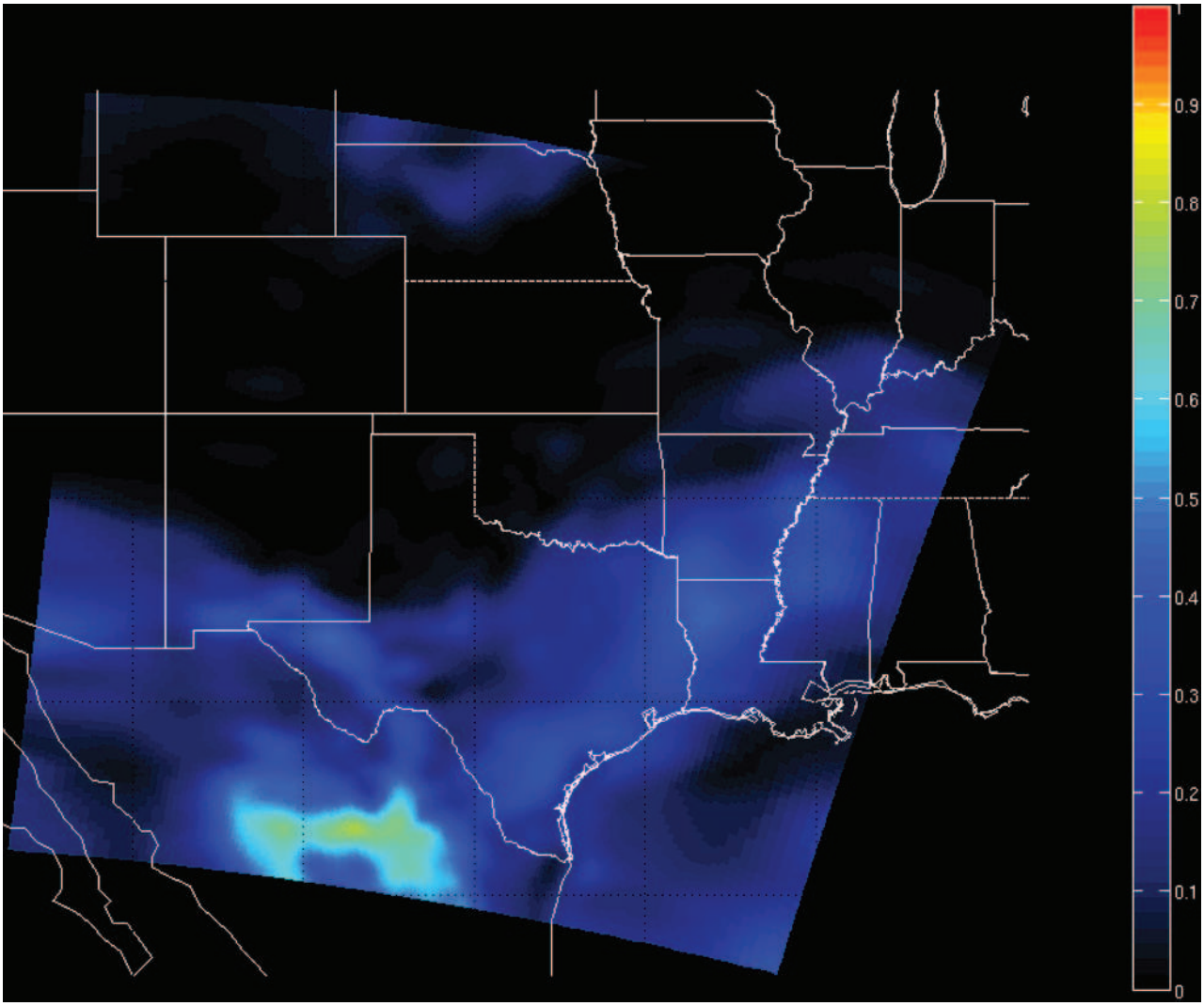

Figure 350. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

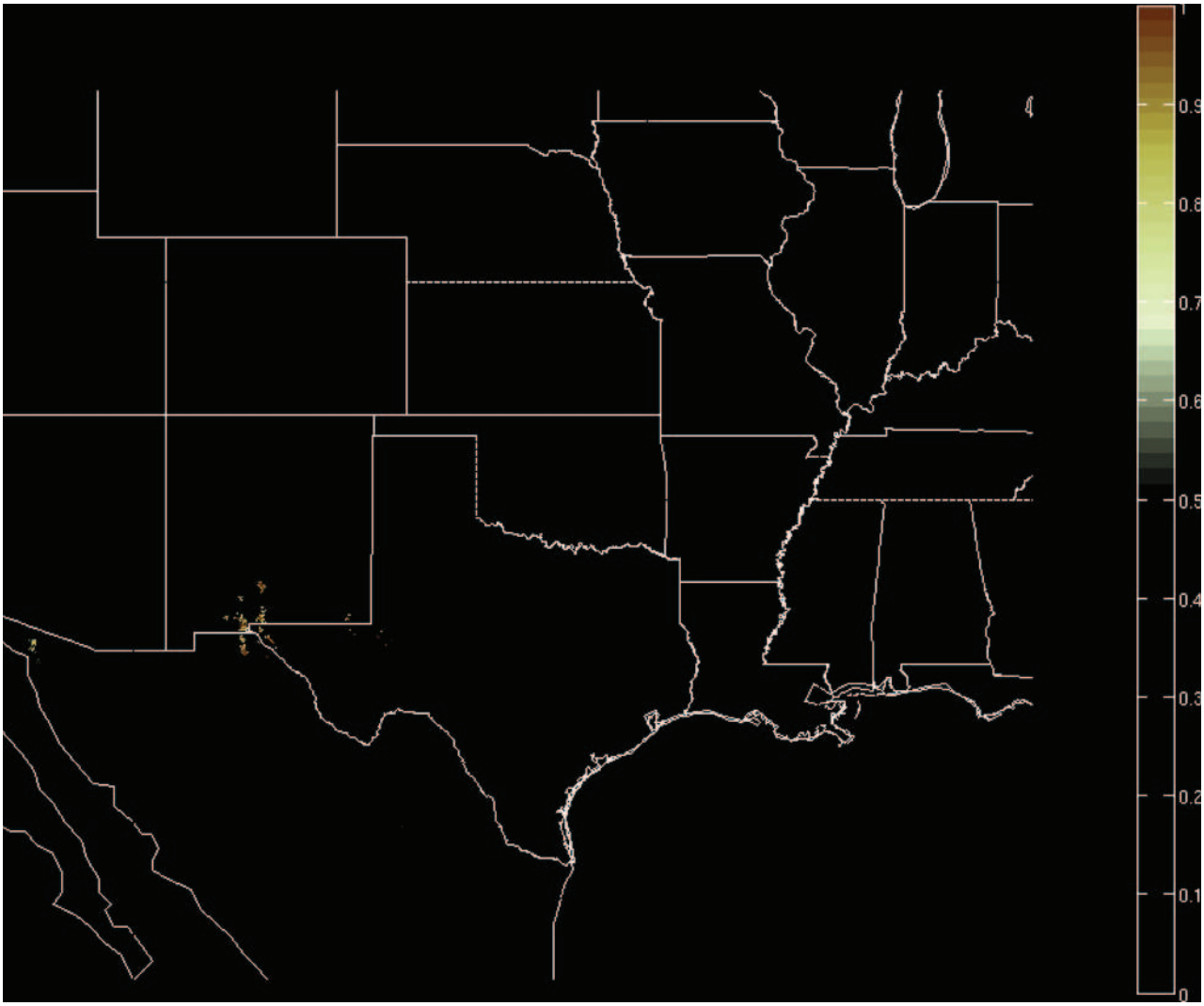

Figure 351. Result of the Maximum Likelihood model (MLE) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

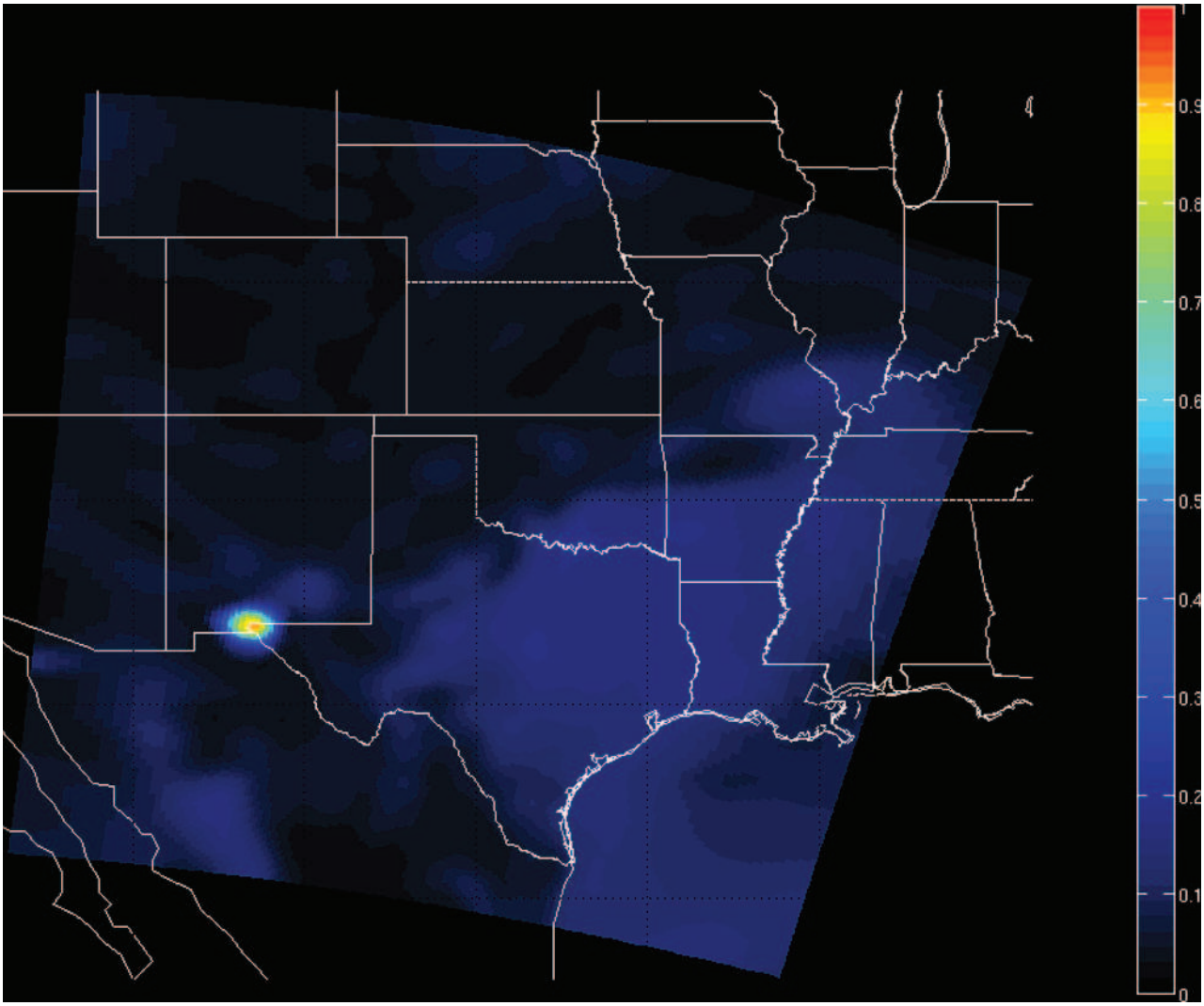

Figure 352. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

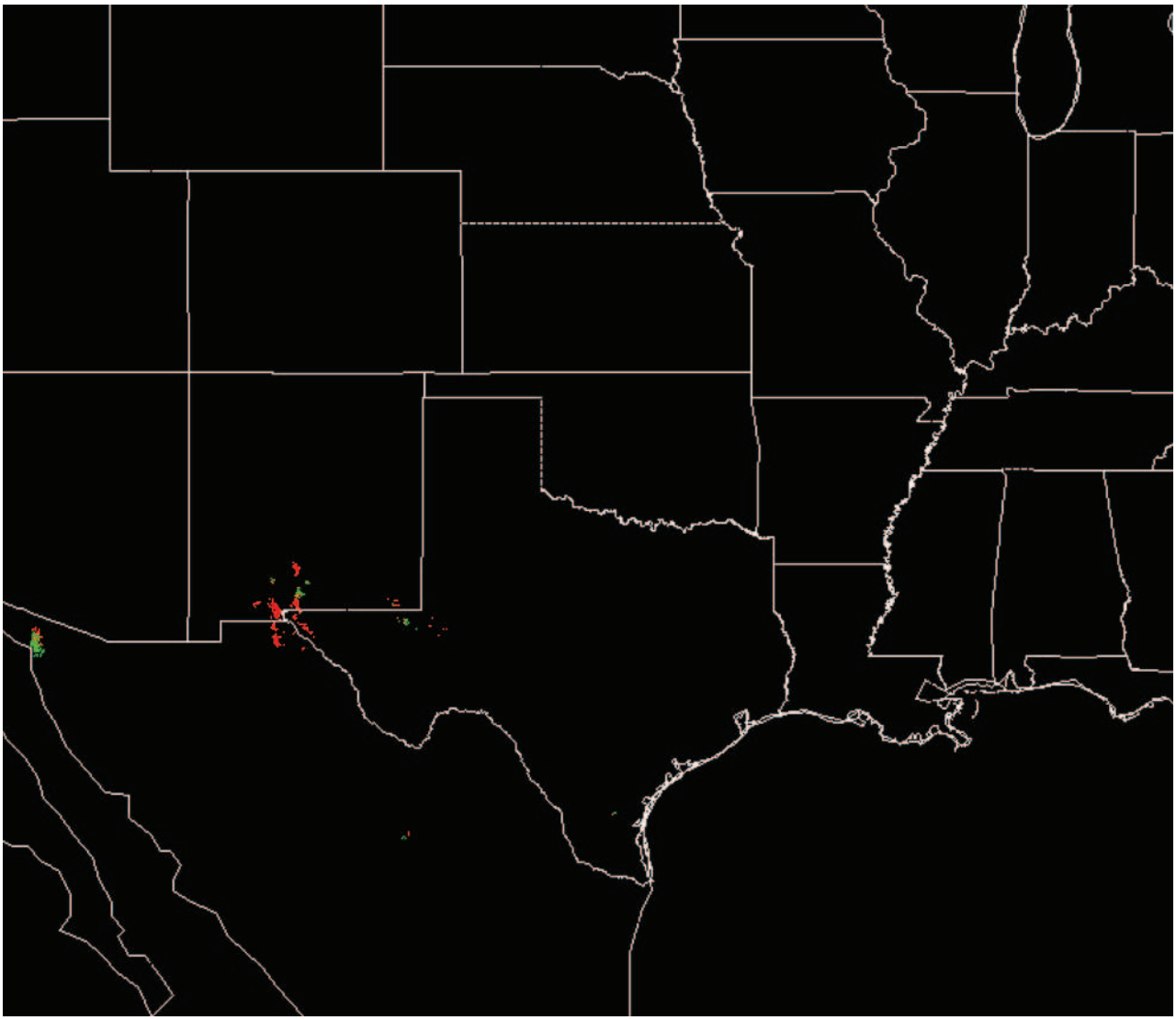

Figure 353. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-078 at 17:30 UTC.

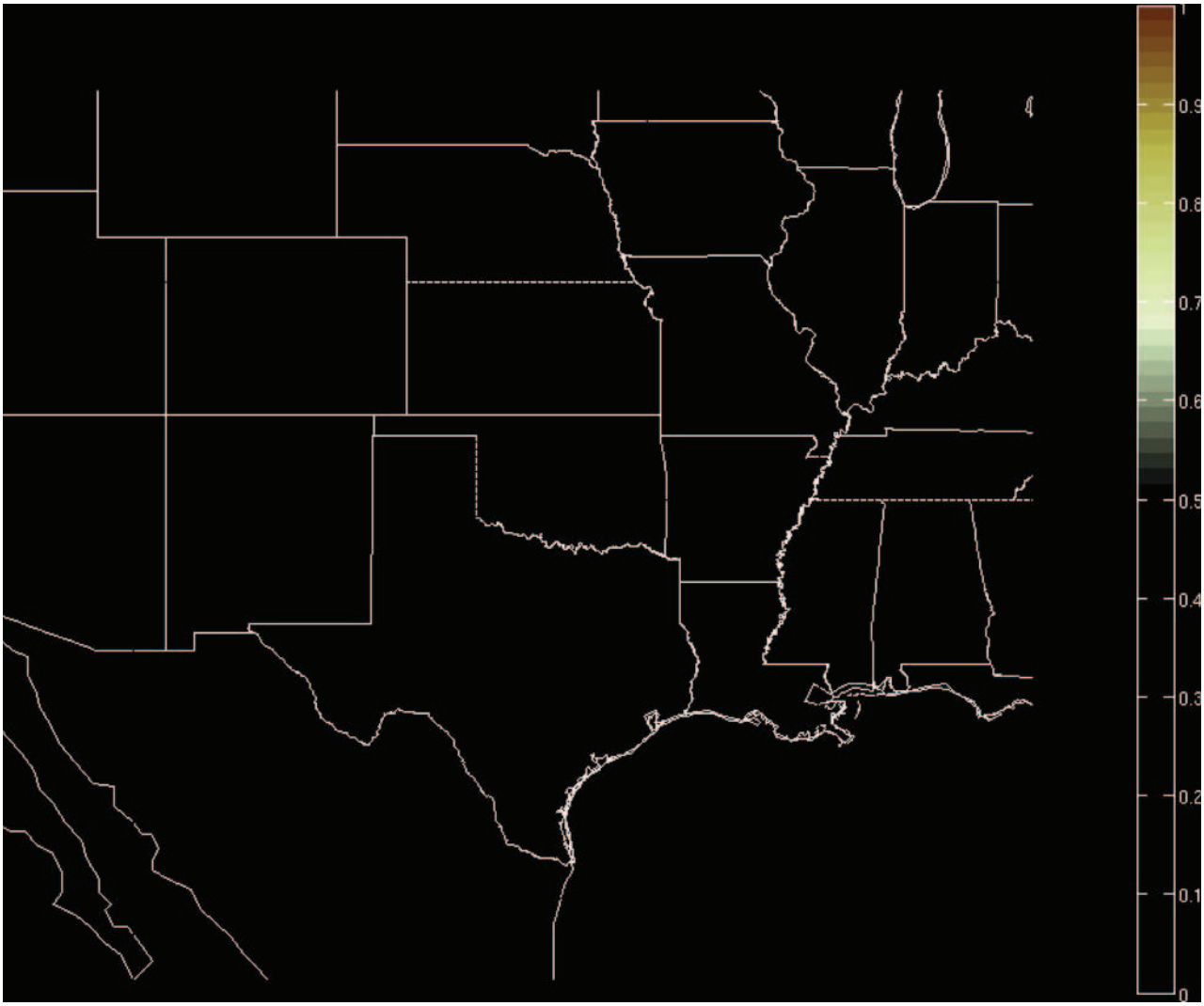

Figure 354. Result of the Maximum a posteriori model (MAP) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

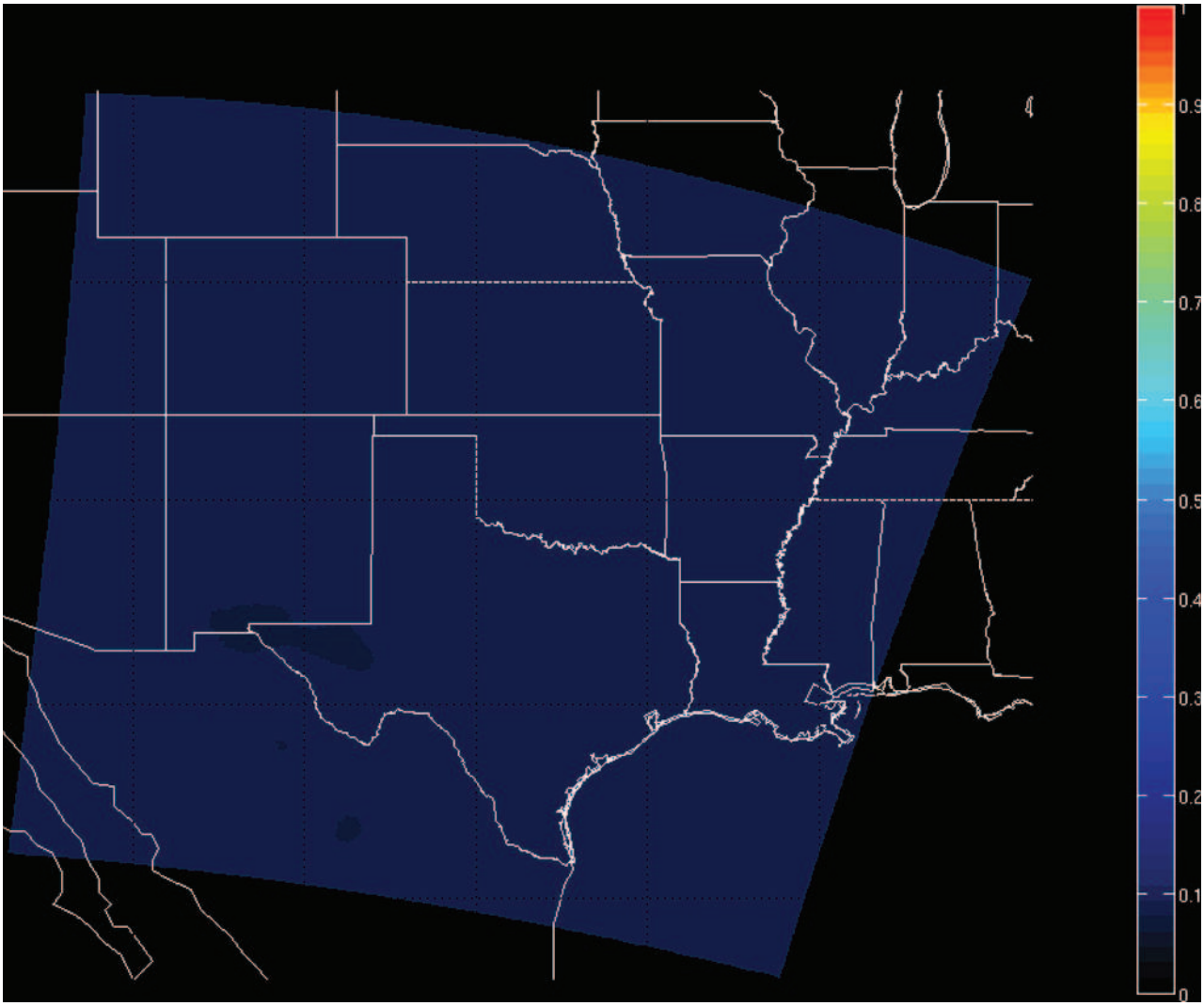

Figure 355. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

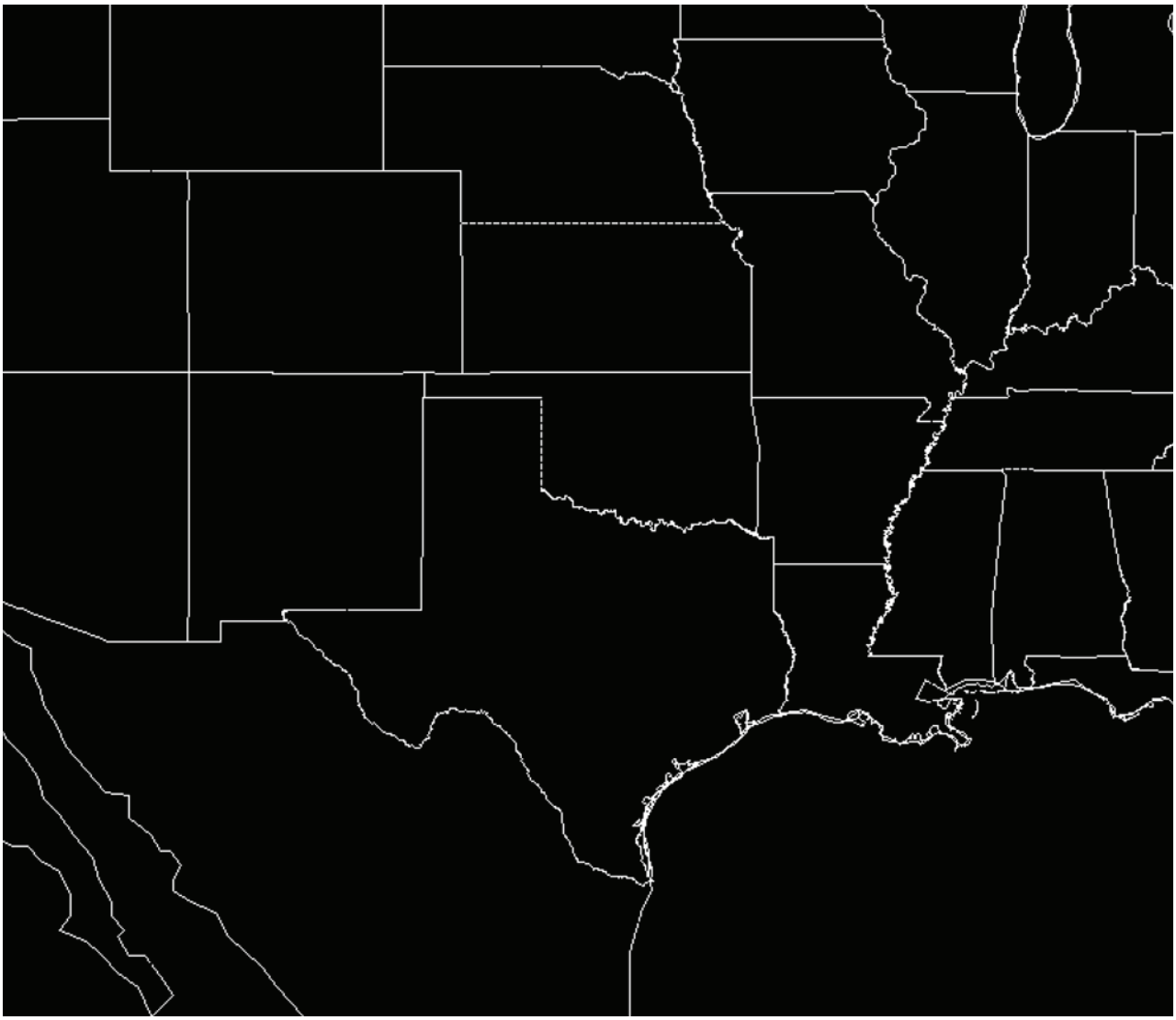

Figure 356. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-078 at 17:30 UTC.

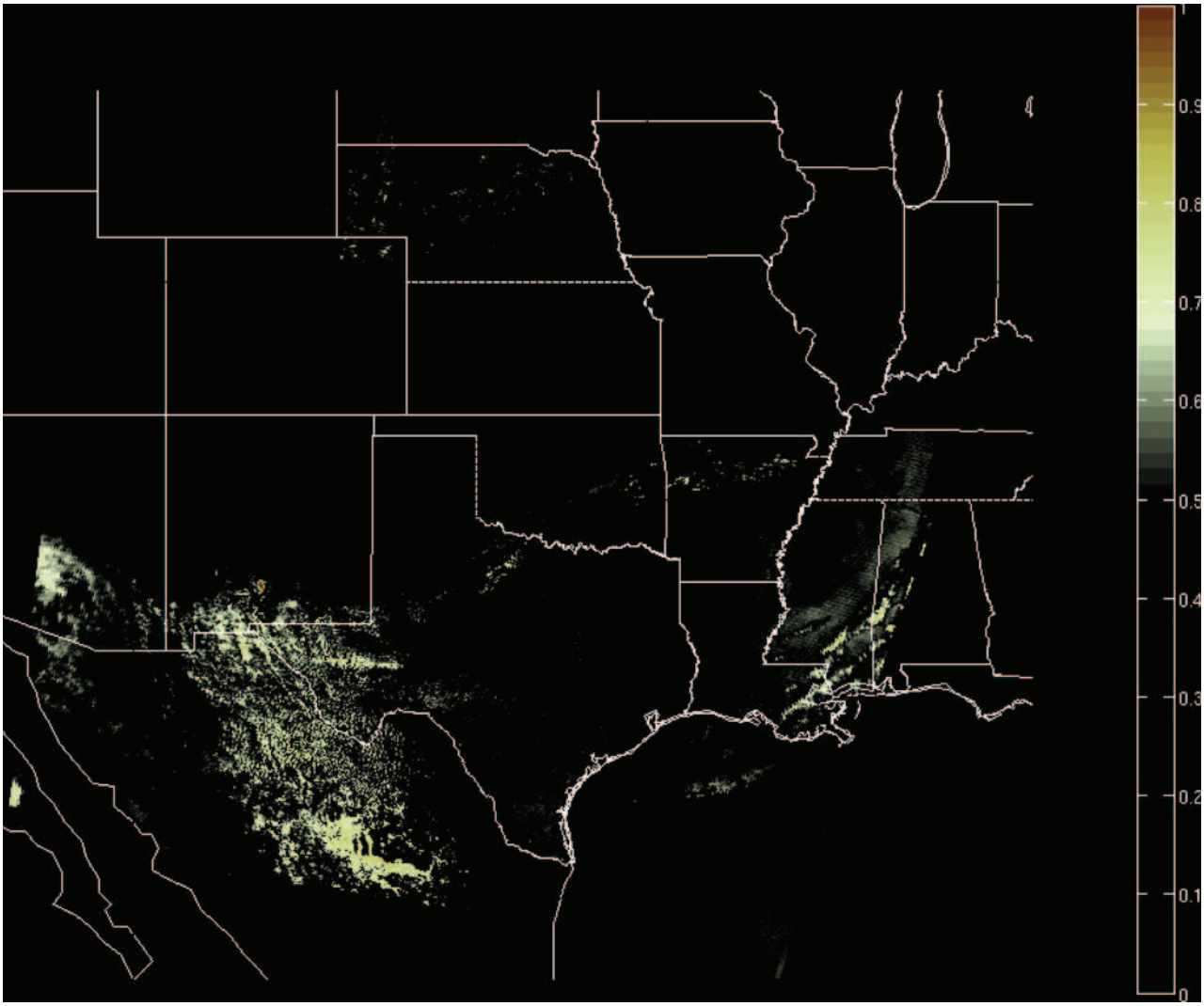

Figure 357. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

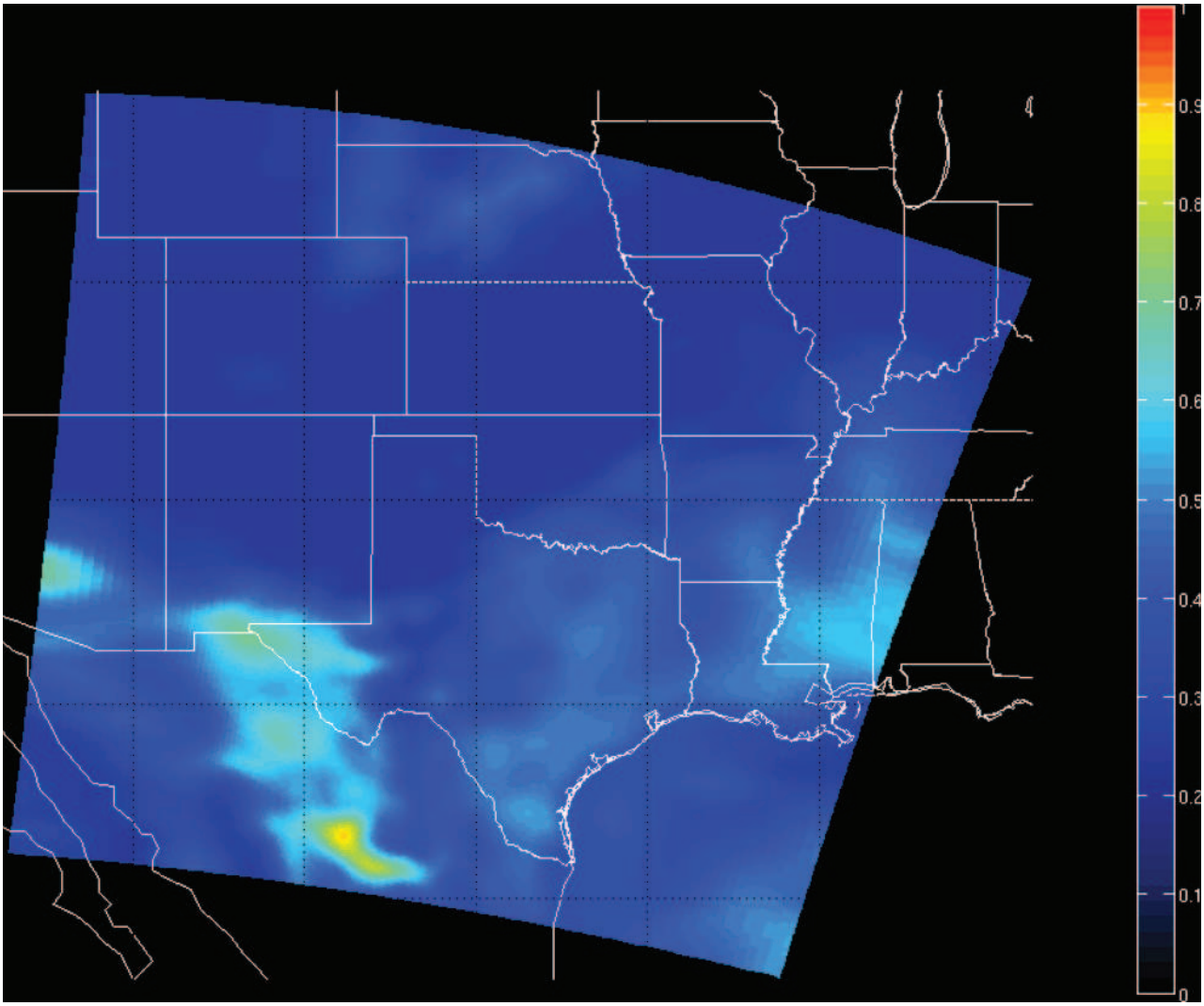

Figure 358. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

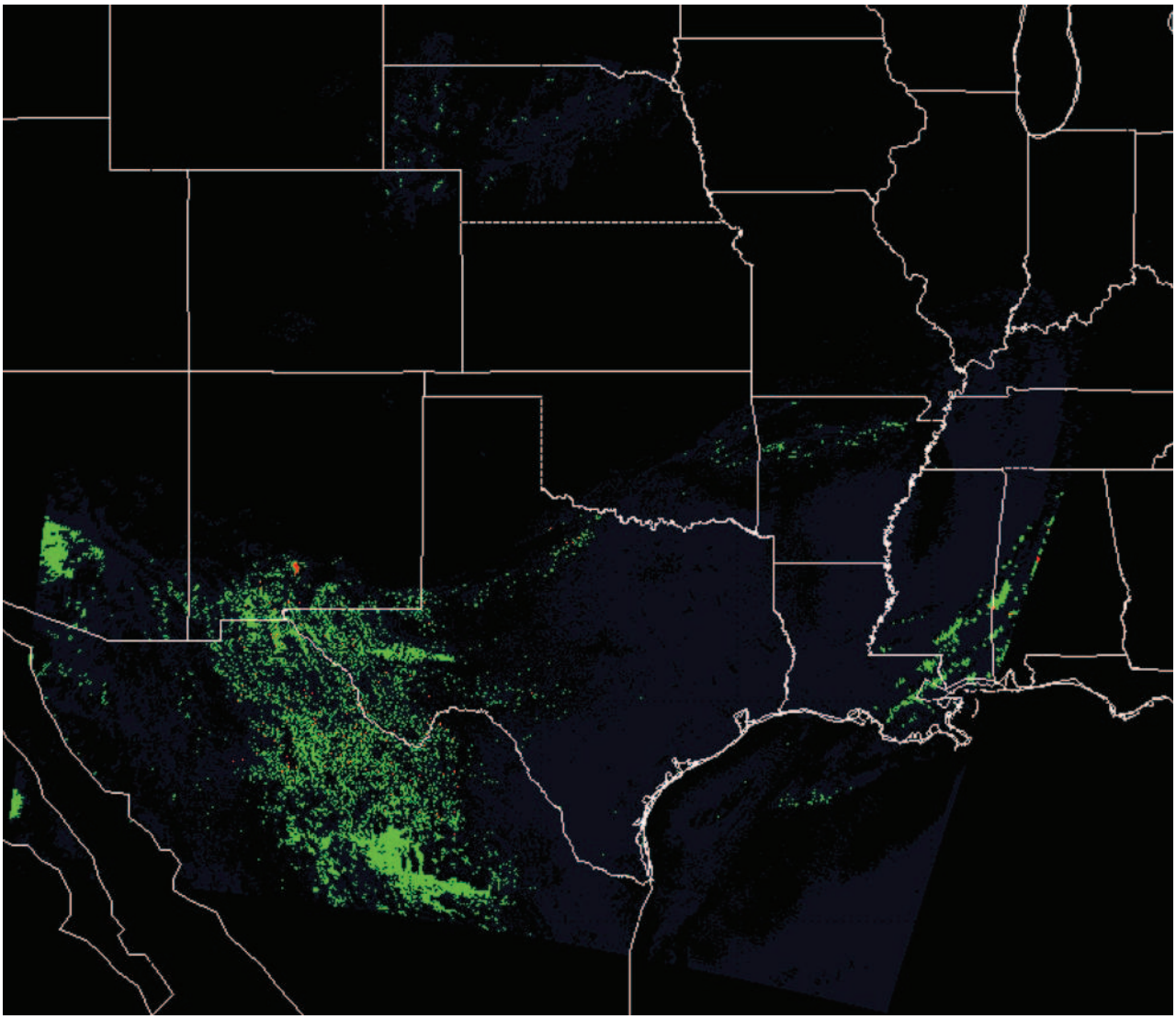

Figure 359. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-078 at 17:30 UTC.

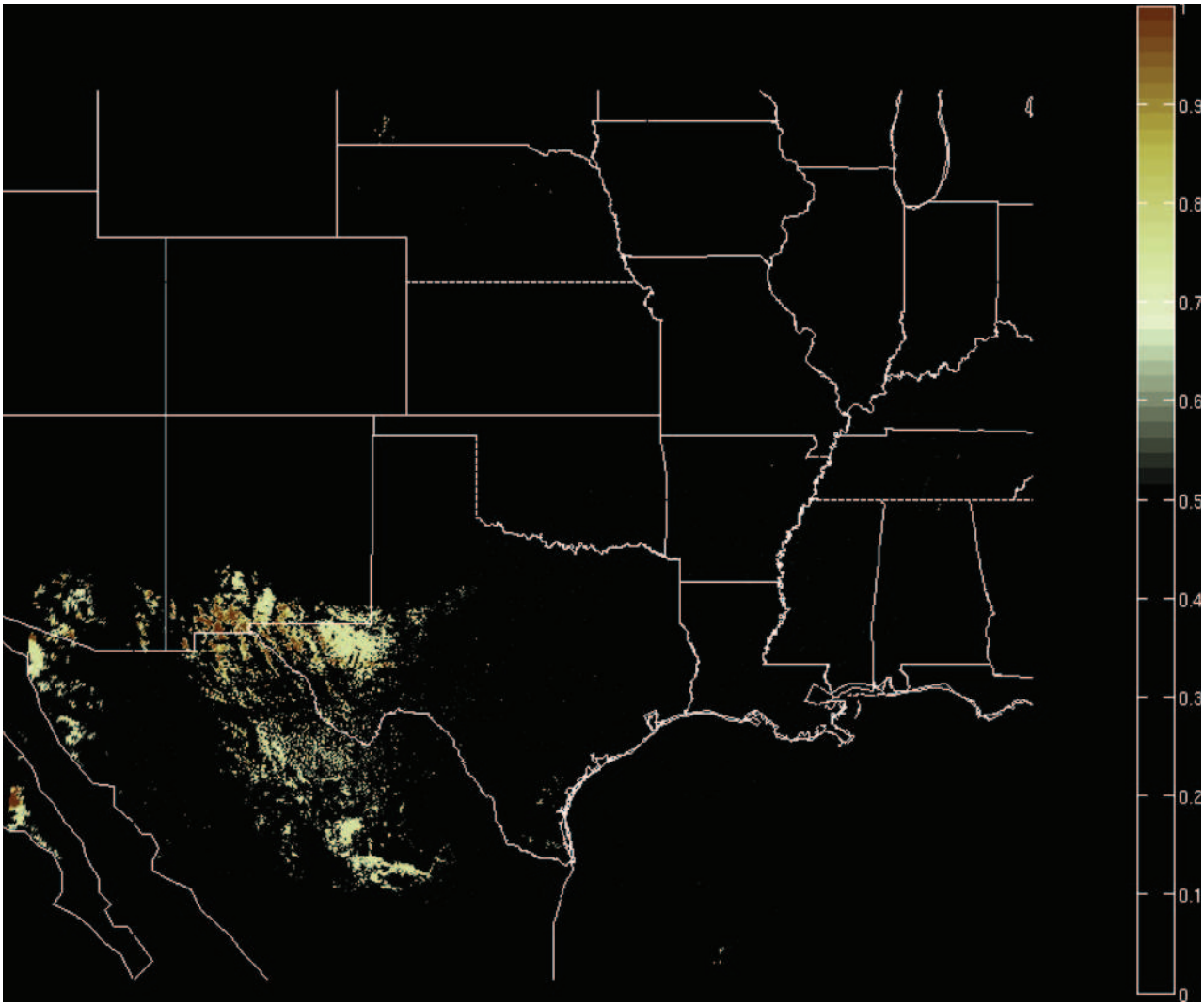

Figure 360. Result of the Probabilistic Neural Network (PNN) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

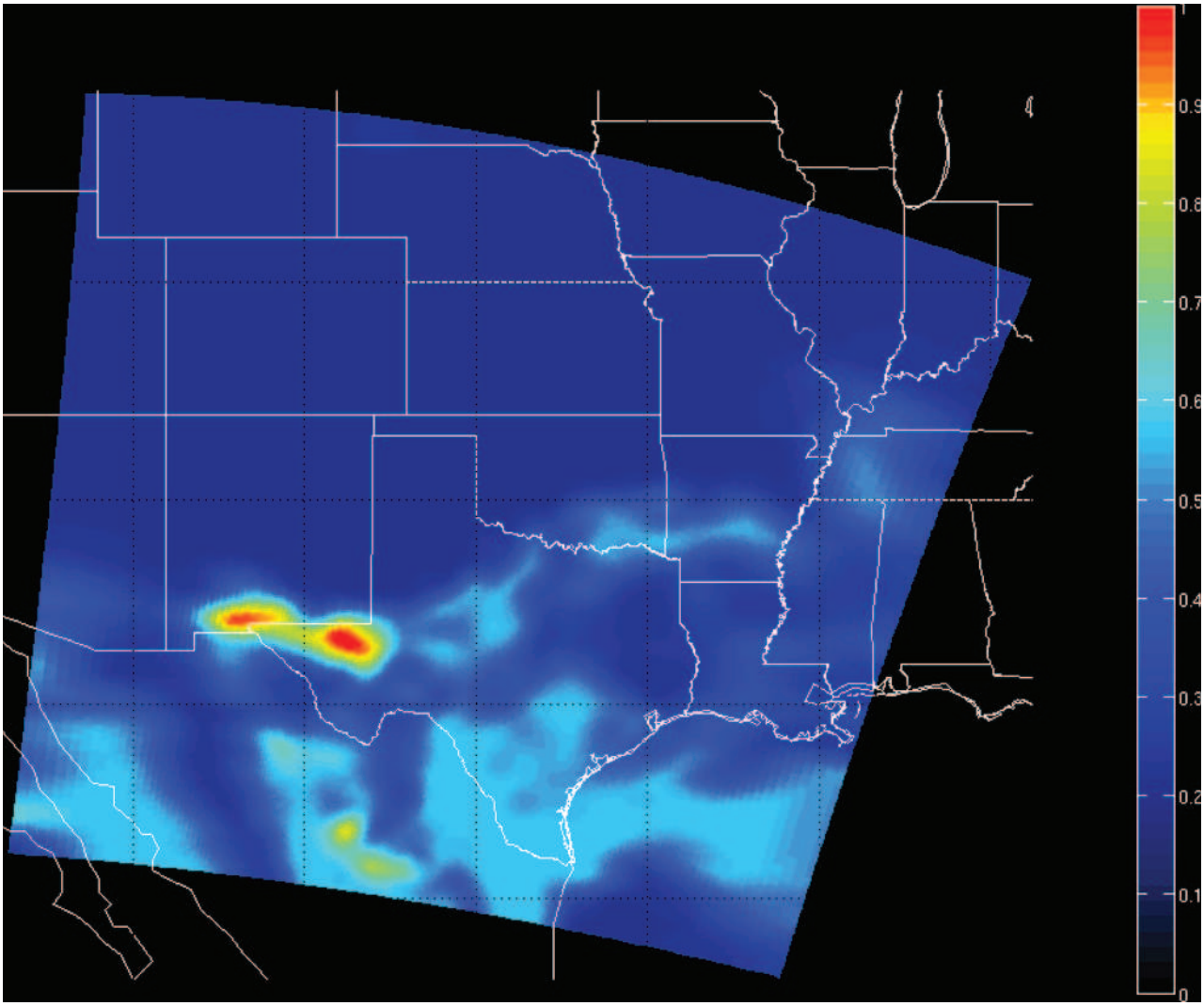

Figure 361. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-078 at 17:30 UTC. The colorbar shows the index of probability.

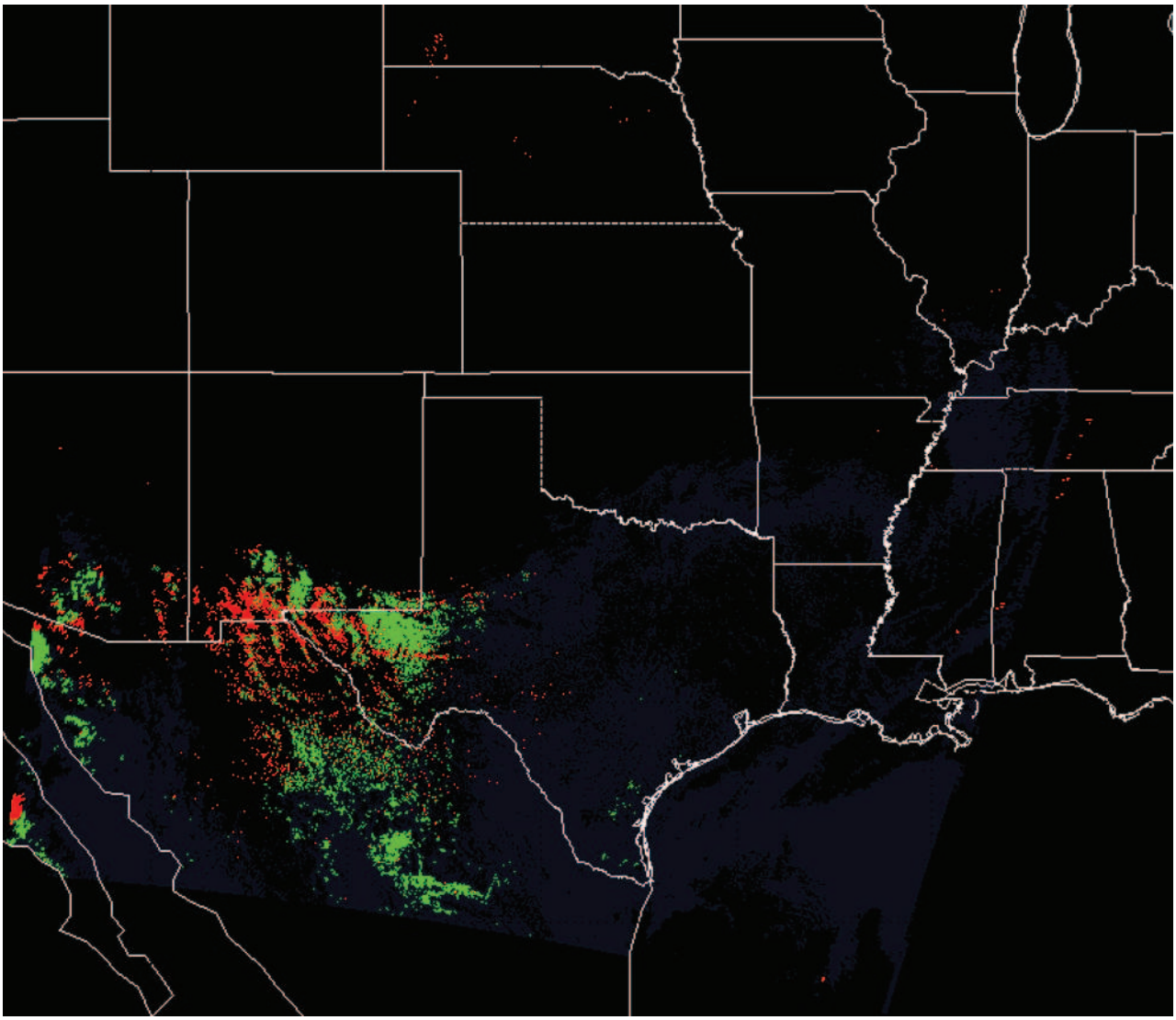

Figure 362. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-078 at 17:30 UTC.

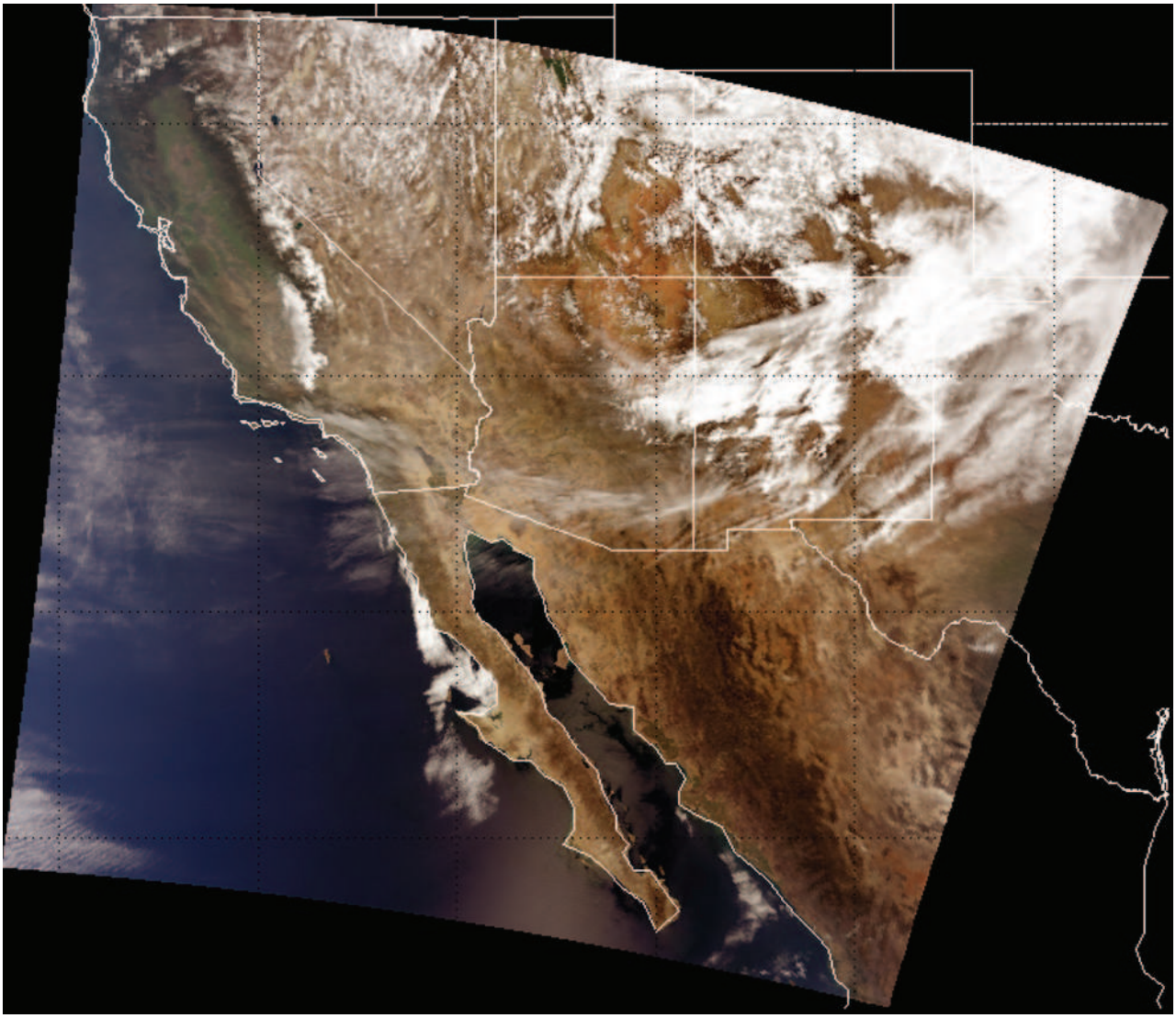

Figure 363. True color image for the event of 2003-086 at 18:20 UTC.

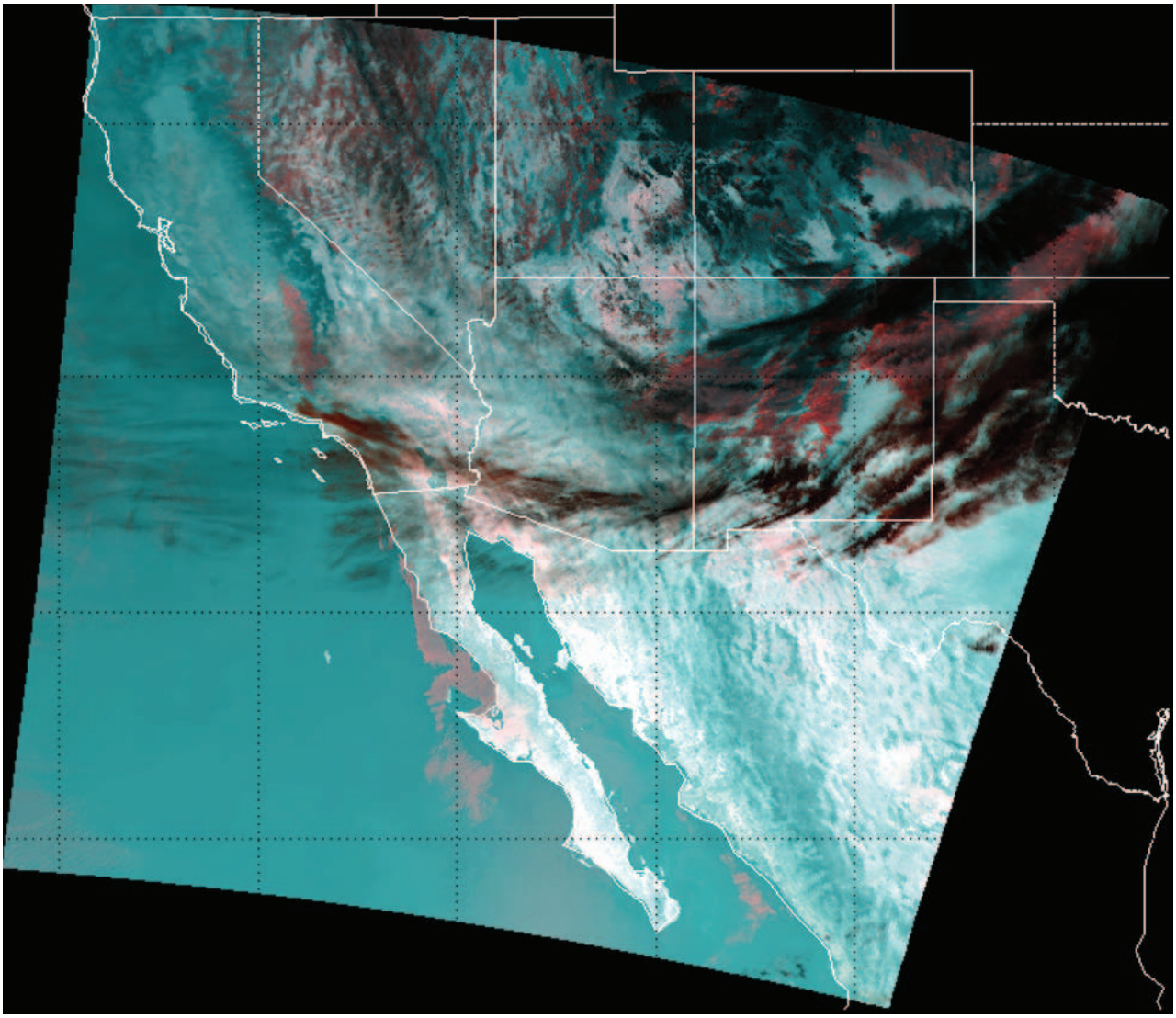

Figure 364. False color image for the event of 2003-086 at 18:20 UTC.

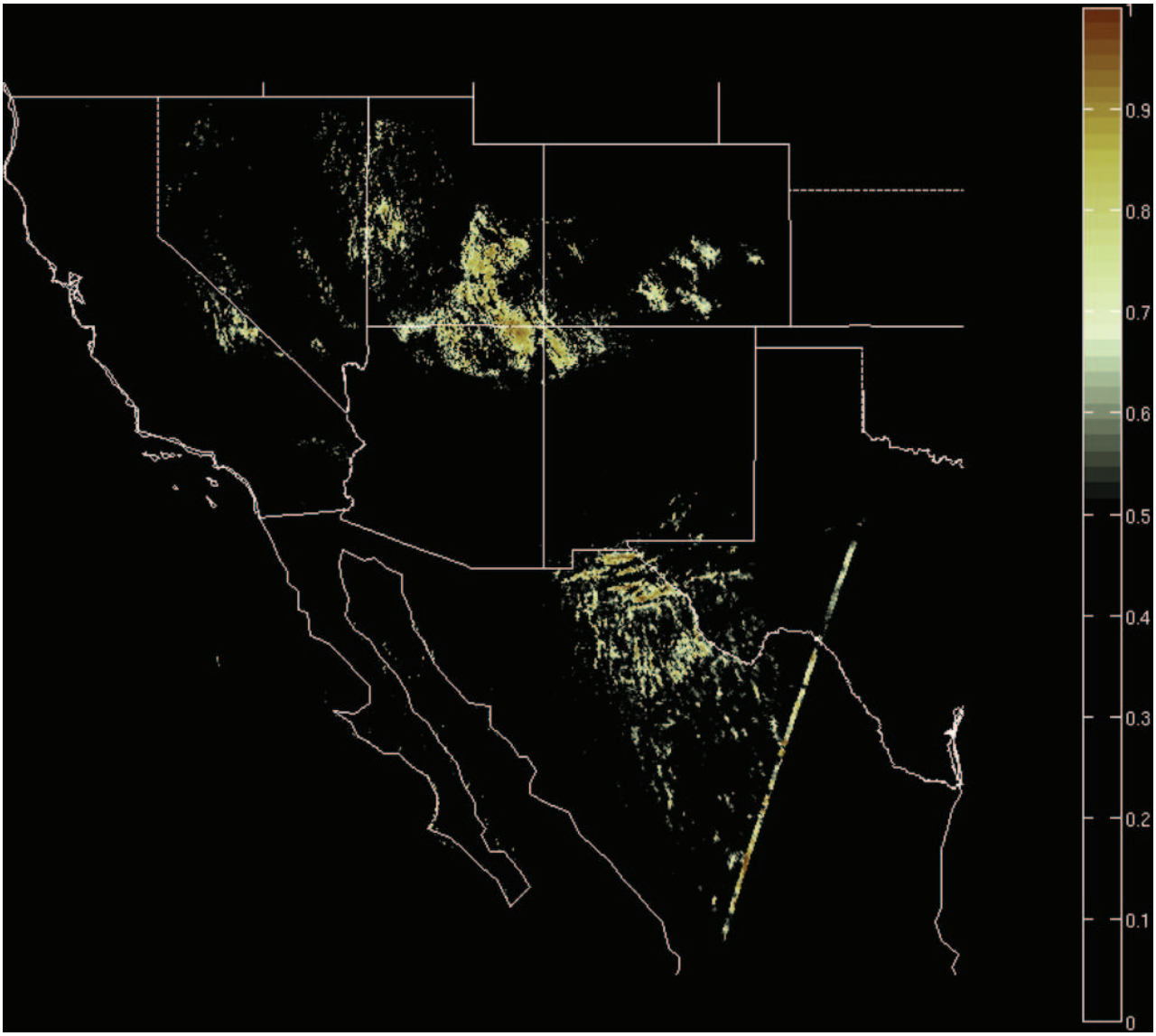

Figure 365. Result of the probabilistic display model (PD) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

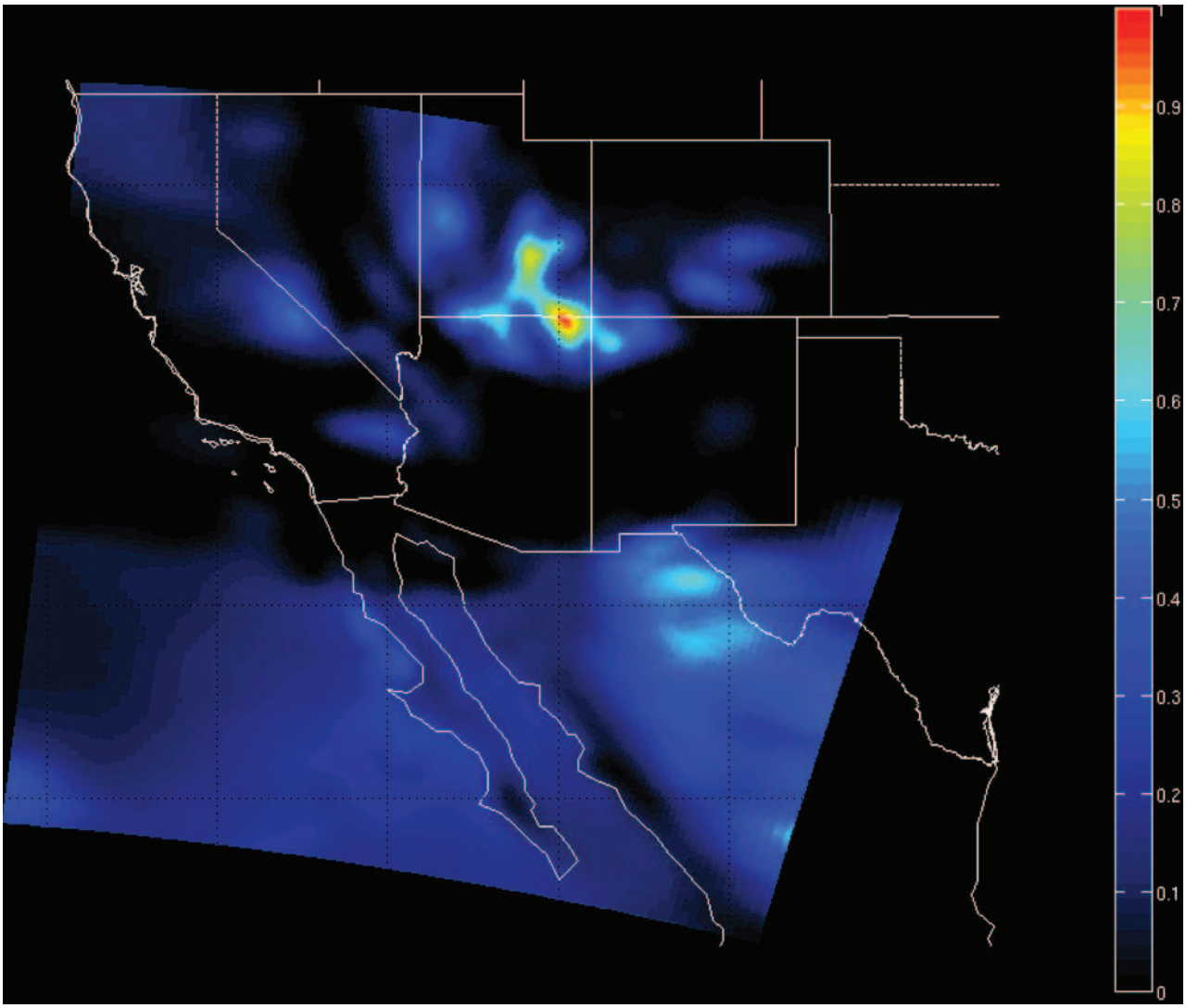

Figure 366. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

Figure 367. Result of the Maximum Likelihood model (MLE) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

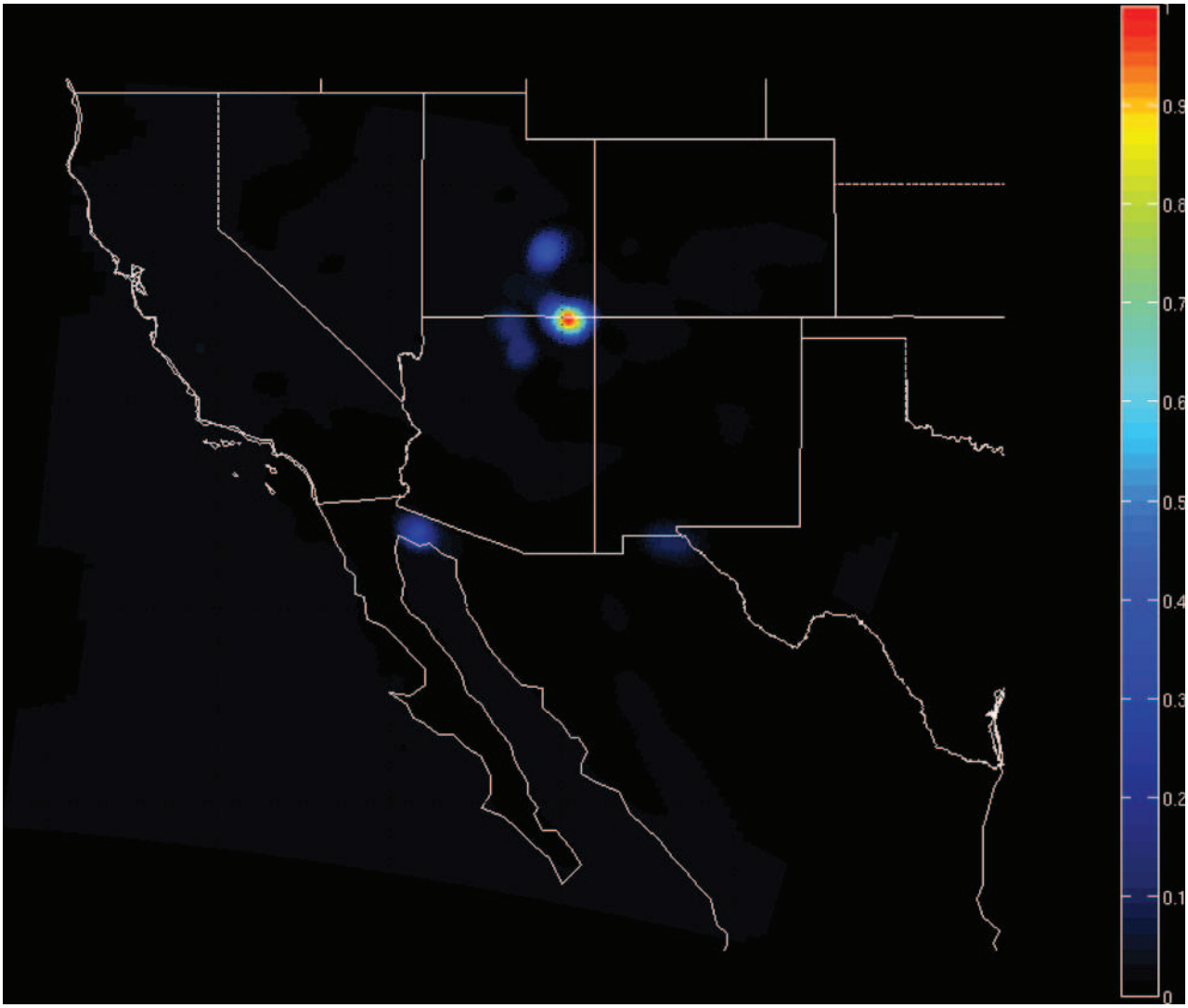

Figure 368. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

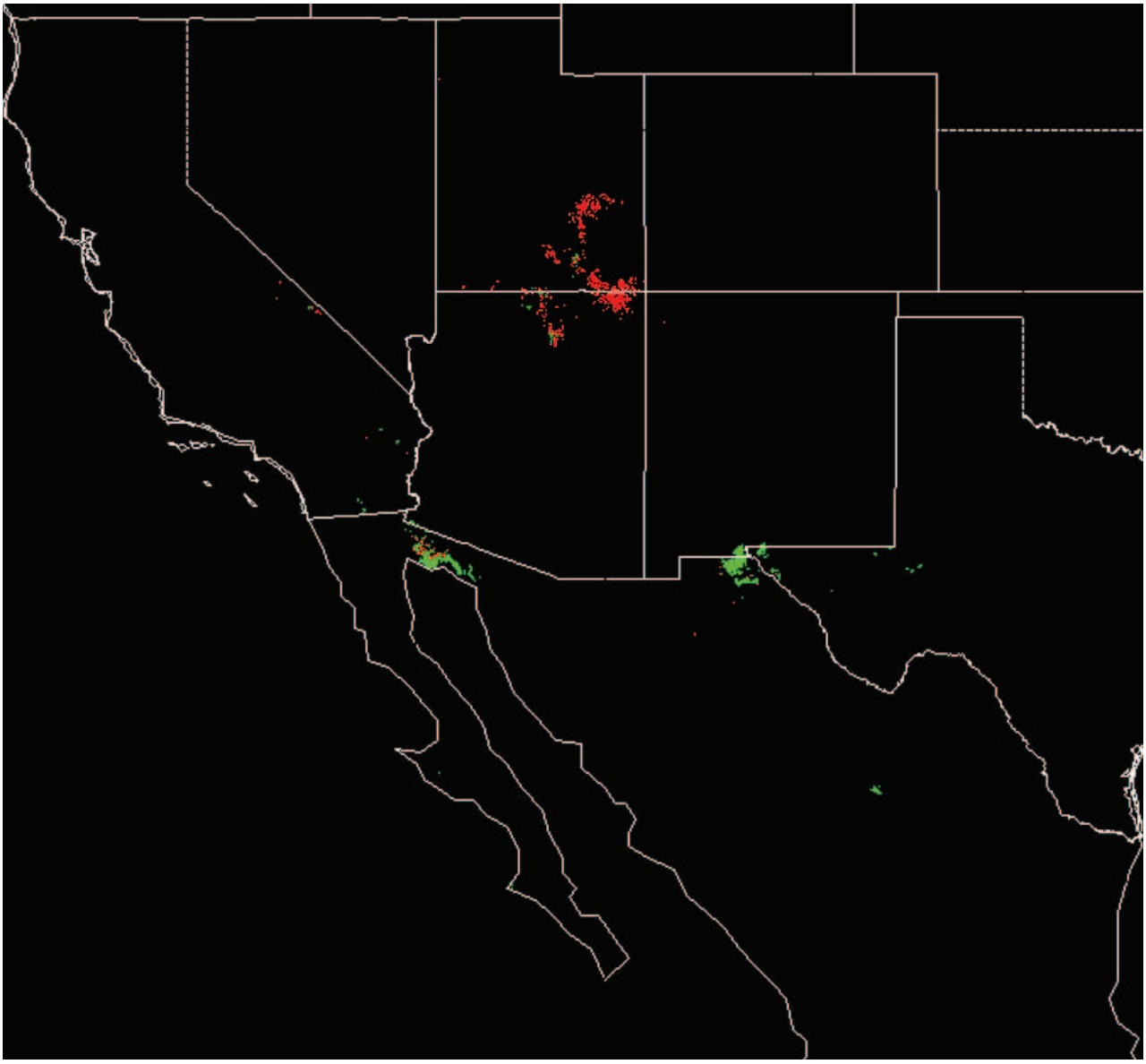

Figure 369. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-086 at 18:20 UTC.

Figure 370. Result of the Maximum a posteriori model (MAP) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

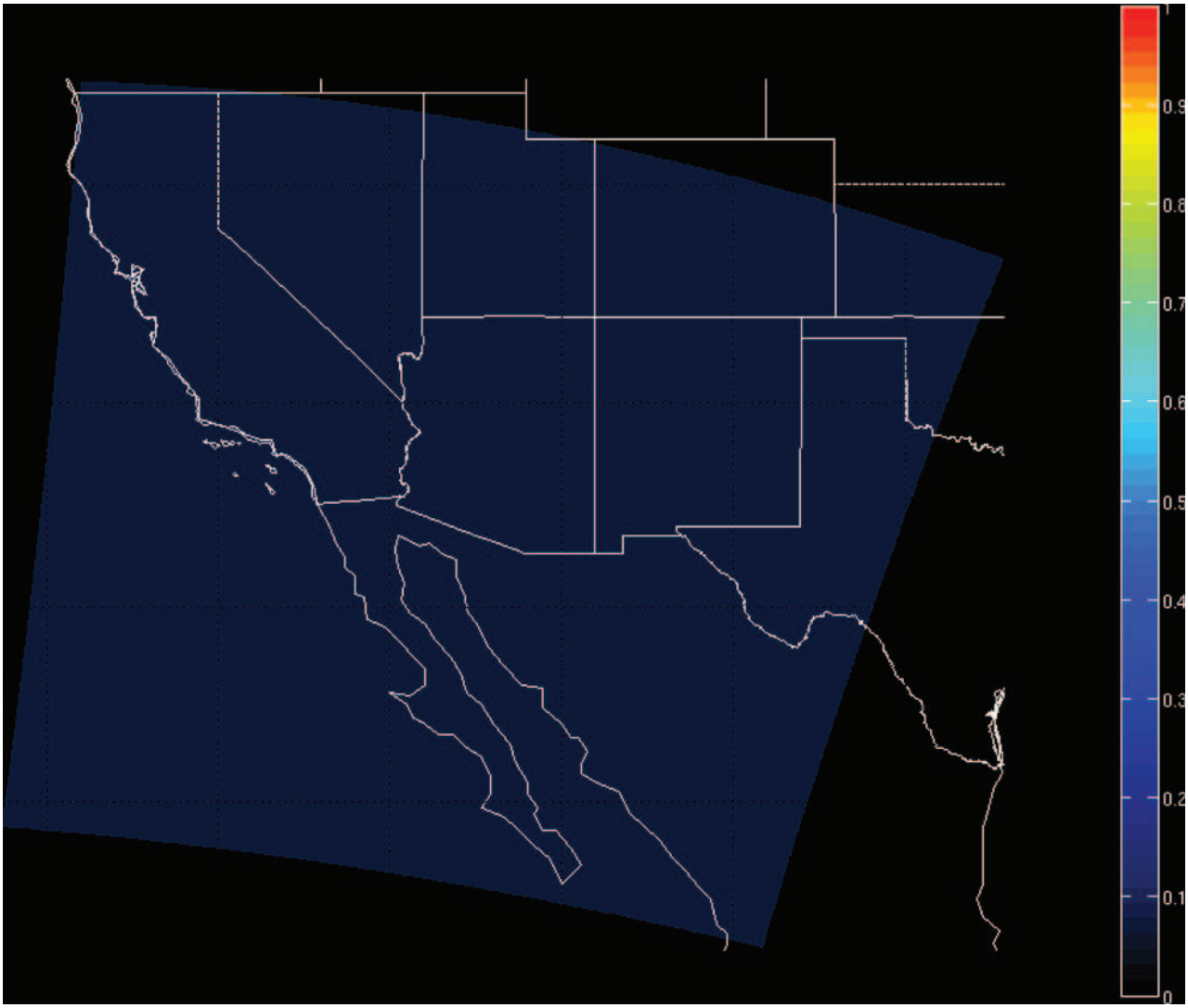

Figure 371. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

Figure 372. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-086 at 18:20 UTC.

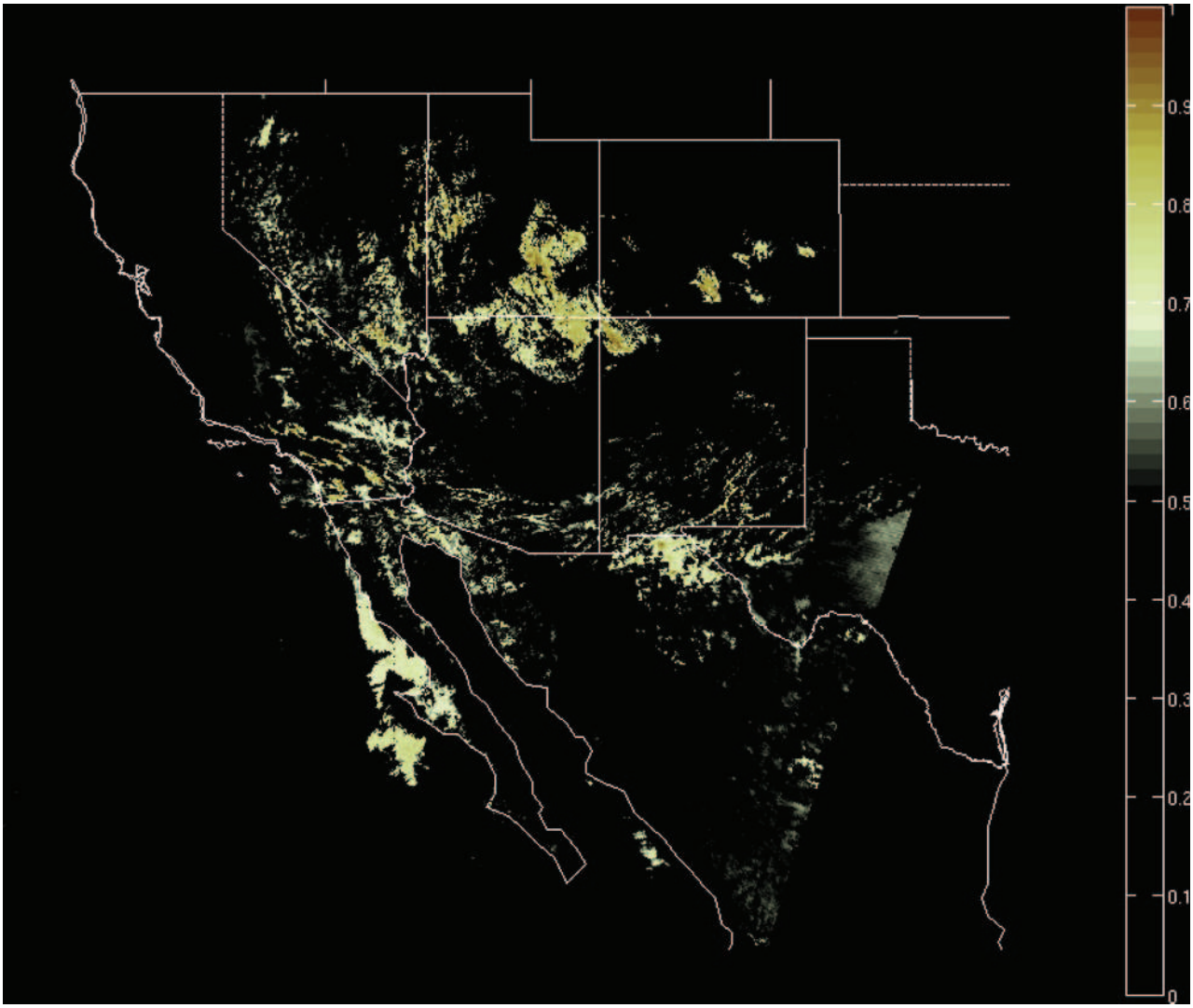

Figure 373. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

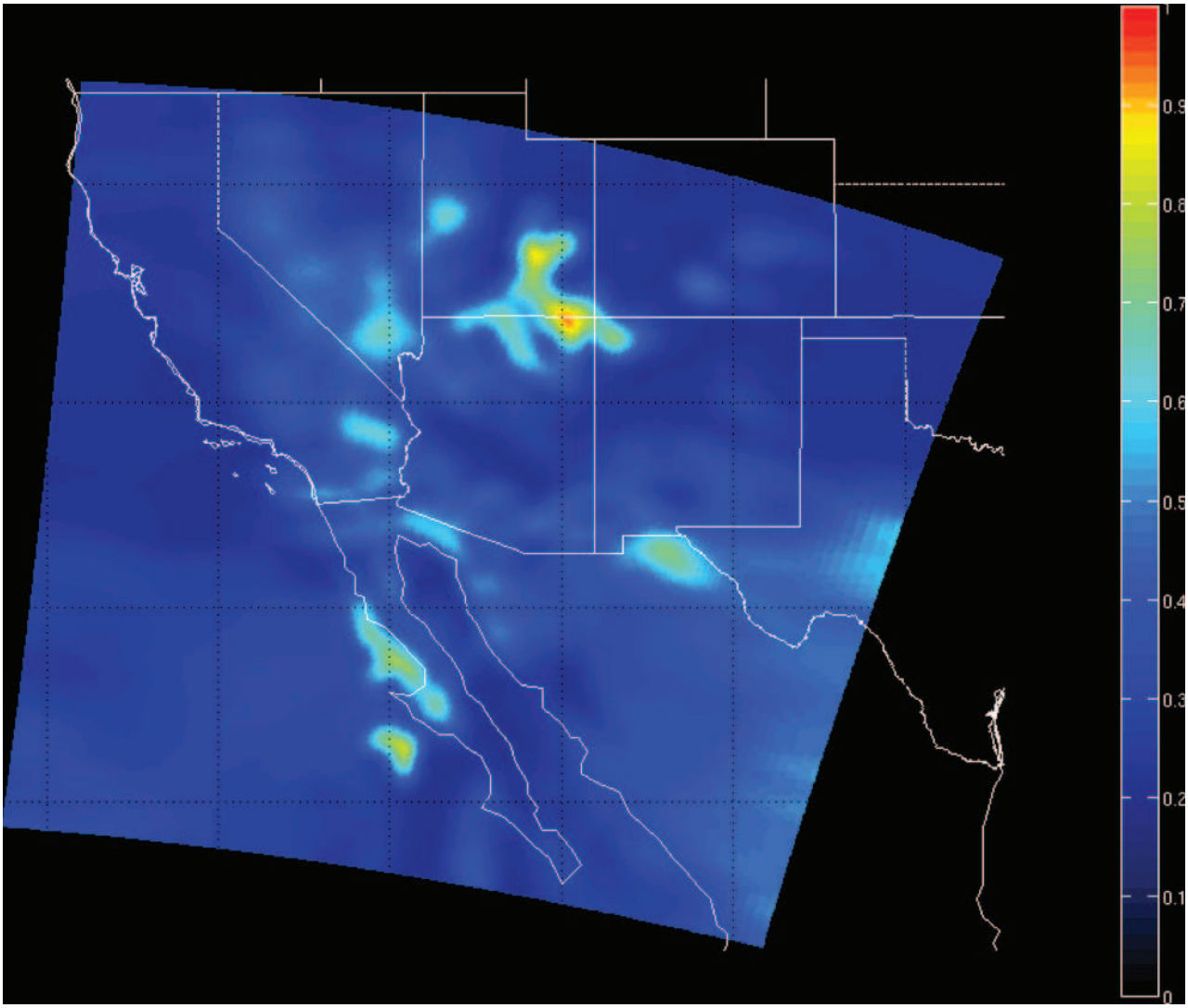

Figure 374. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

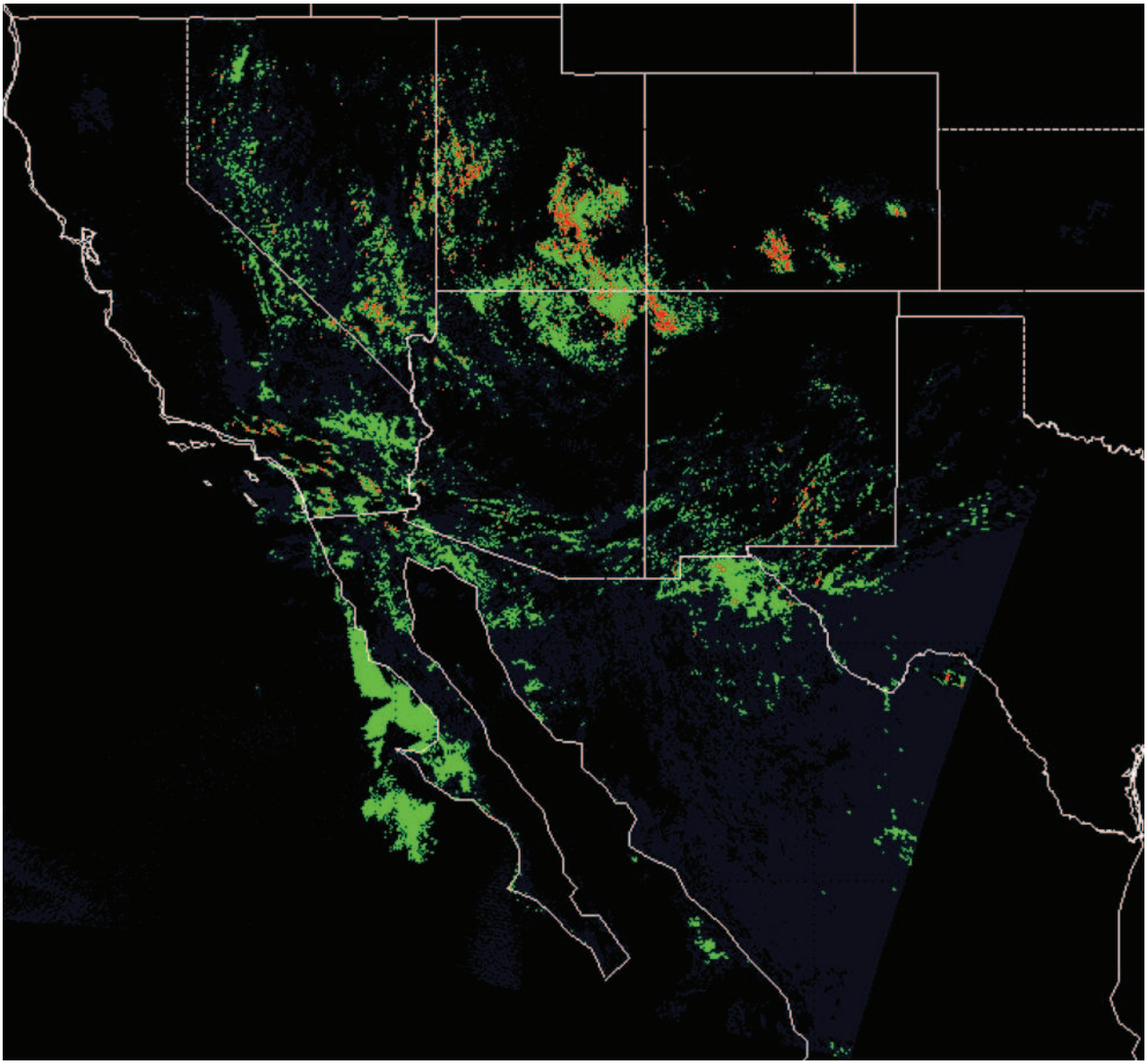

Figure 375. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-086 at 18:20 UTC.

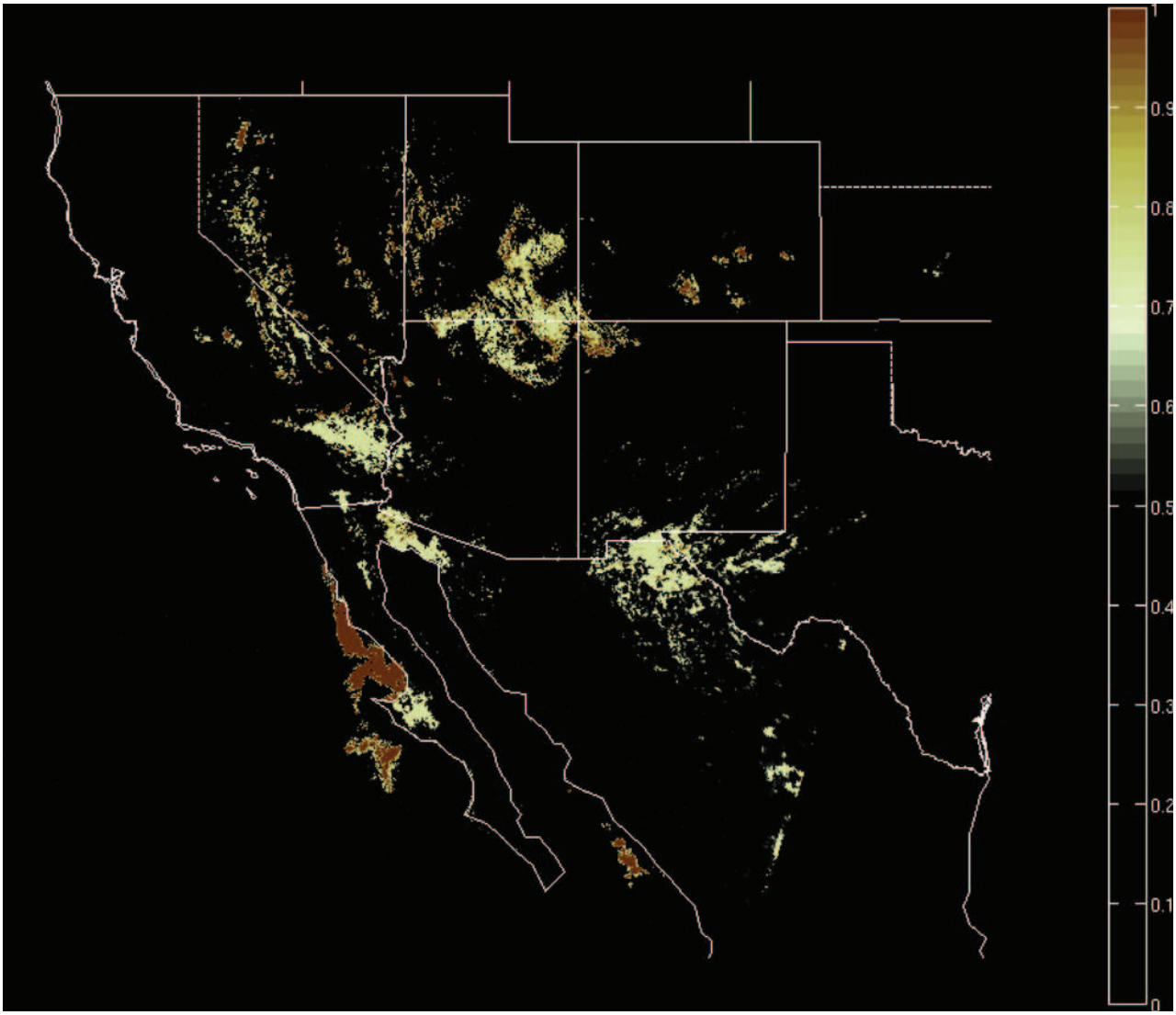

Figure 376. Result of the Probabilistic Neural Network (PNN) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

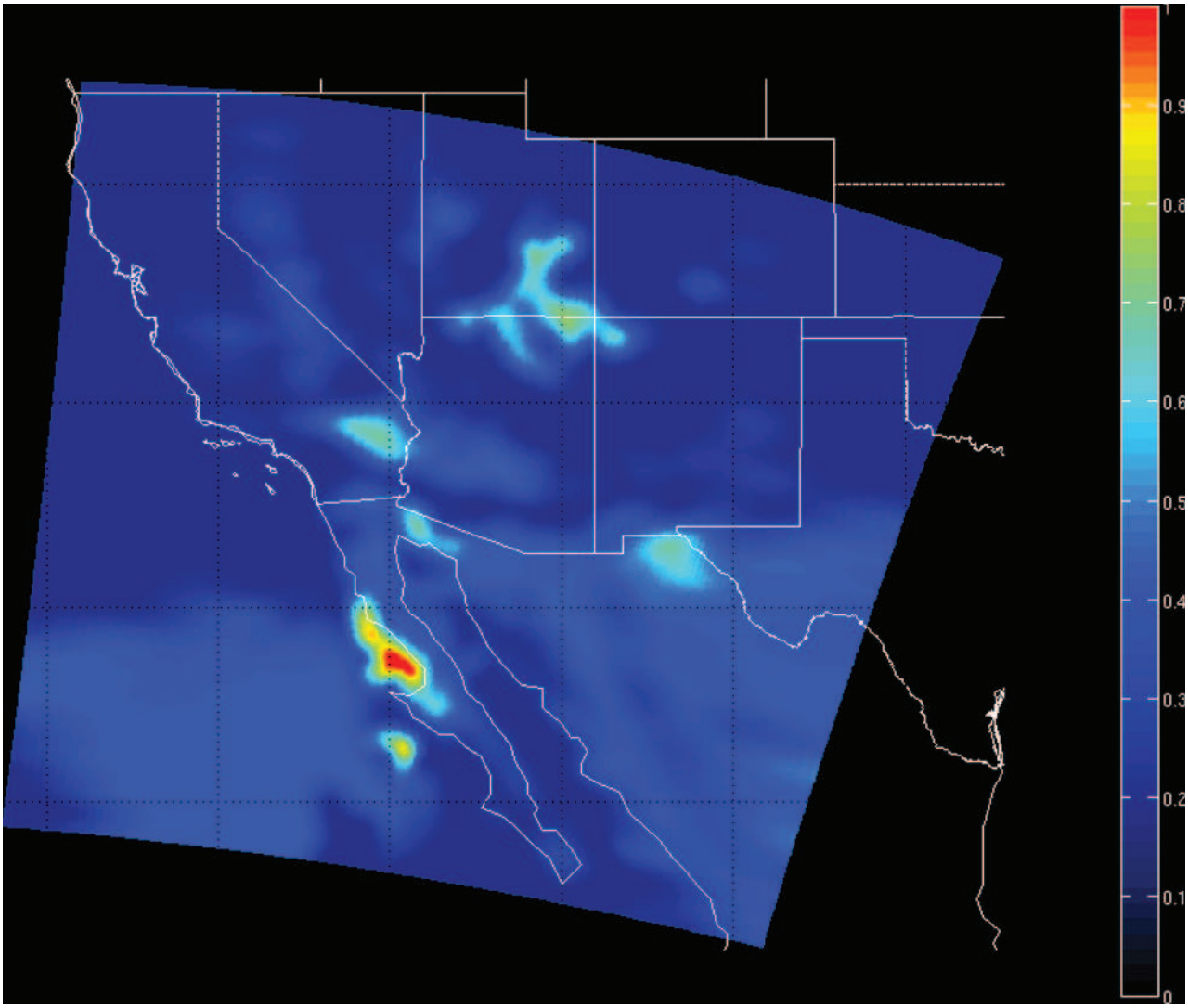

Figure 377. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-086 at 18:20 UTC. The colorbar shows the index of probability.

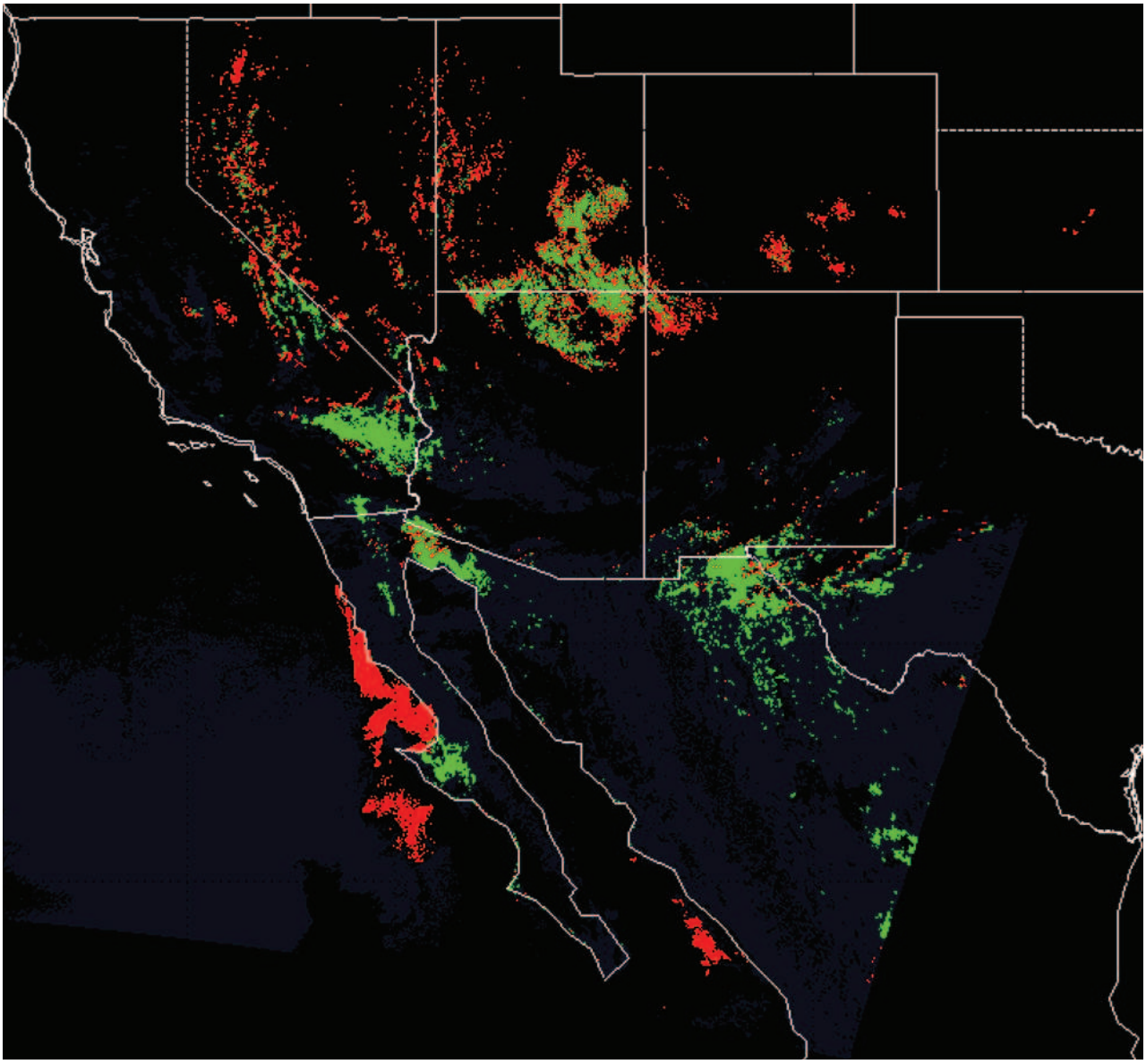

Figure 378. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-086 at 18:20 UTC.

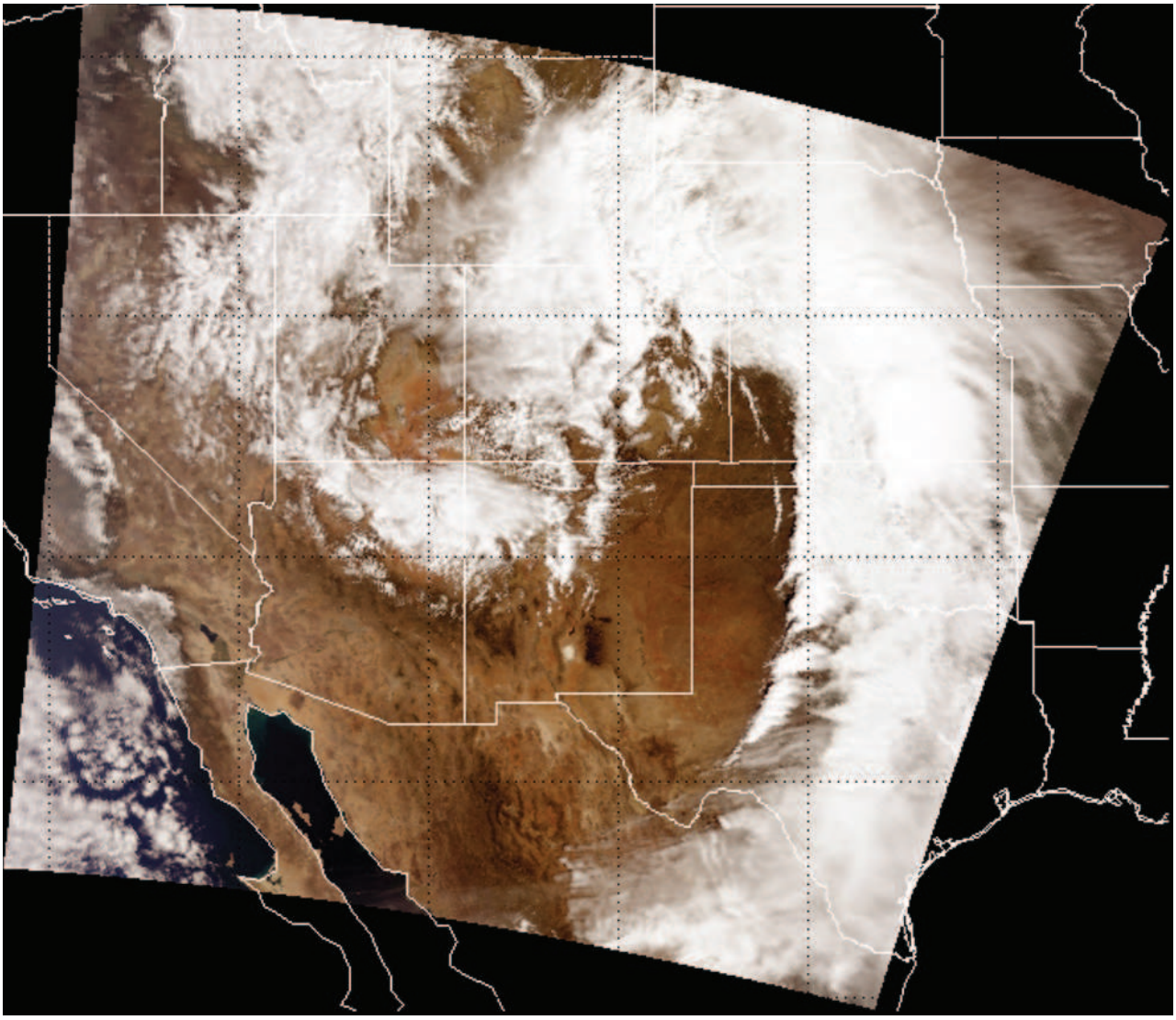

Figure 379. True color image for the event of 2003-113 at 18:00 UTC.

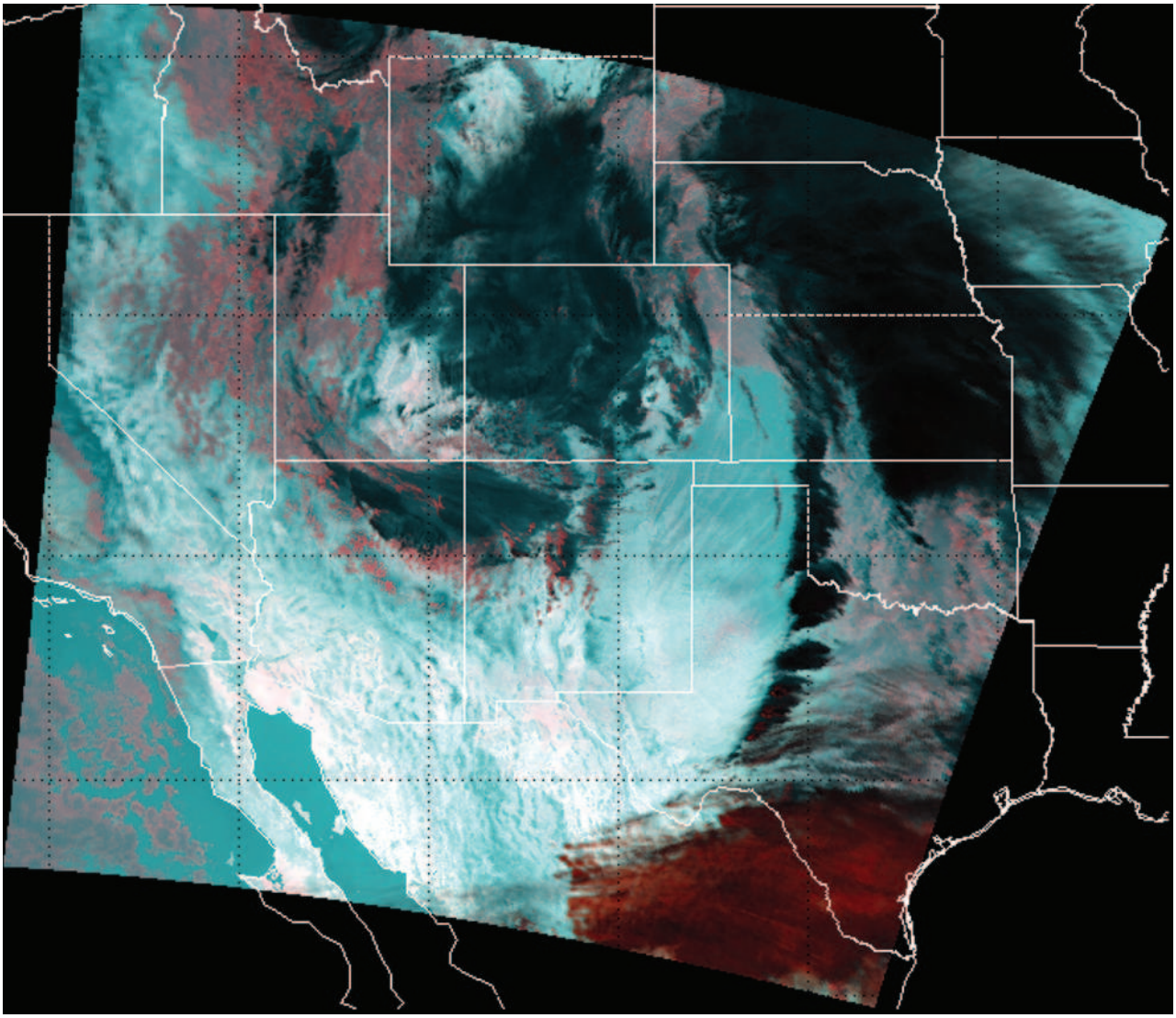

Figure 380. False color image for the event of 2003-113 at 18:00 UTC.

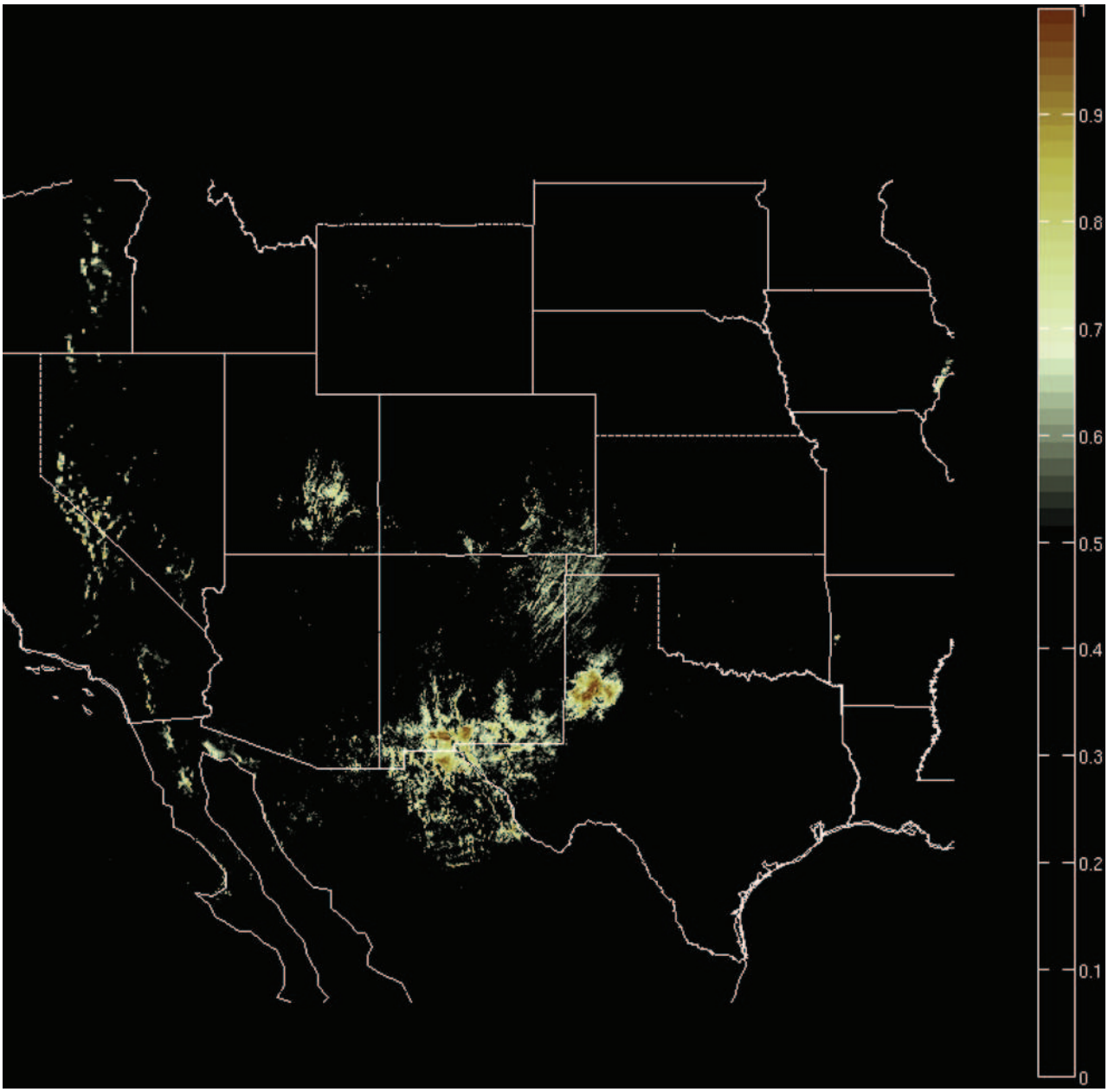

Figure 381. Result of the probabilistic display model (PD) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

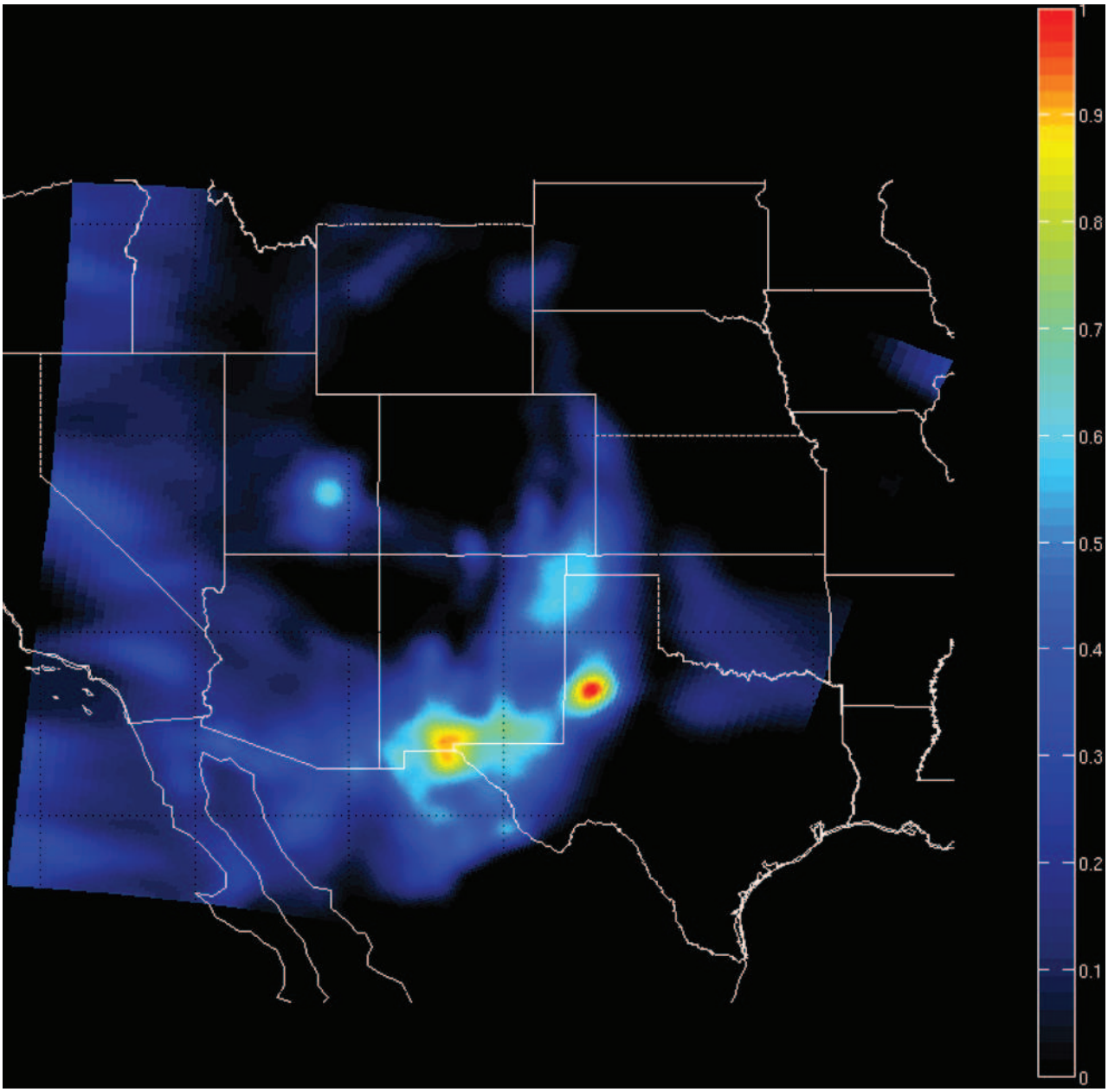

Figure 382. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

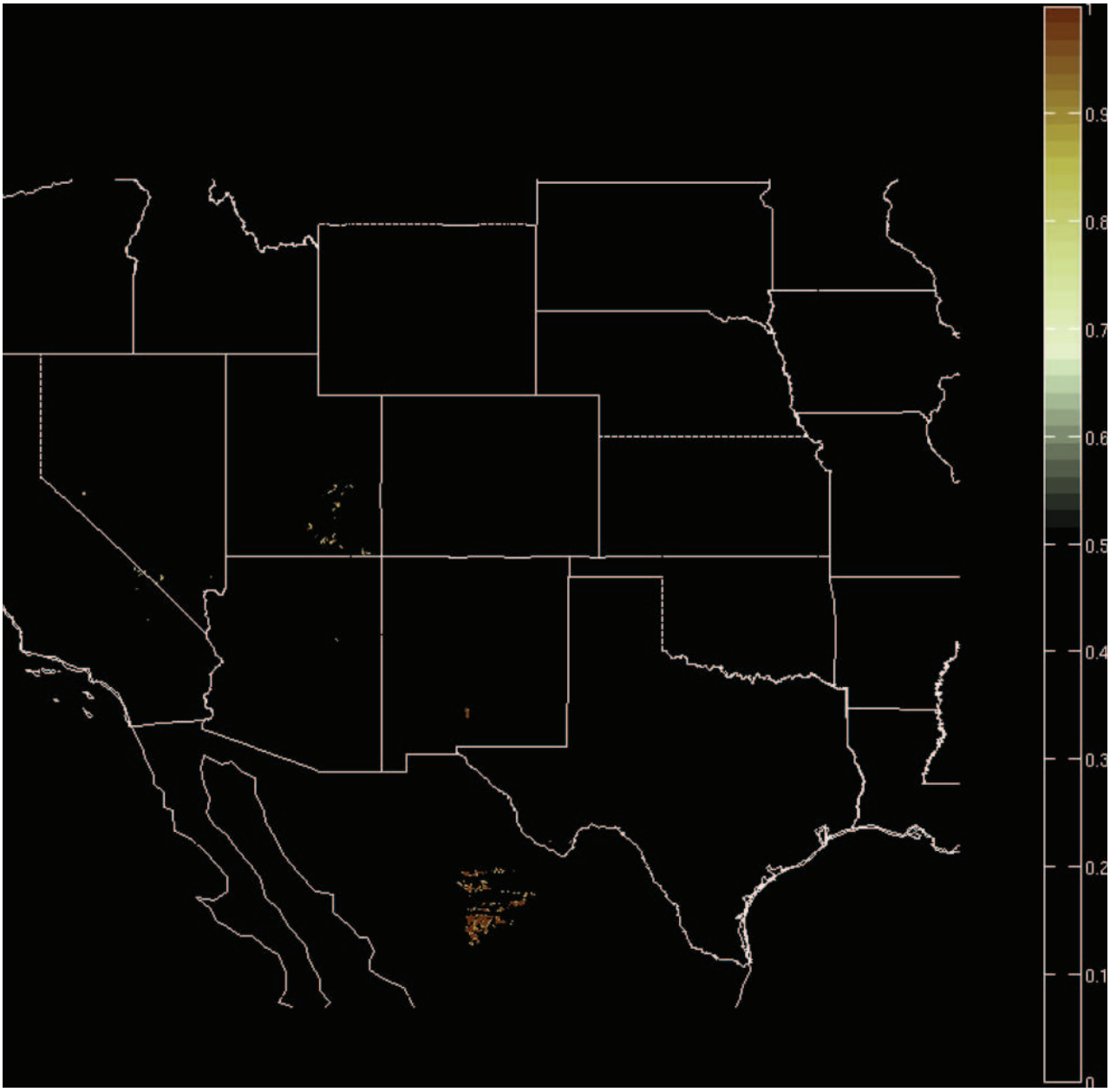

Figure 383. Result of the Maximum Likelihood model (MLE) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

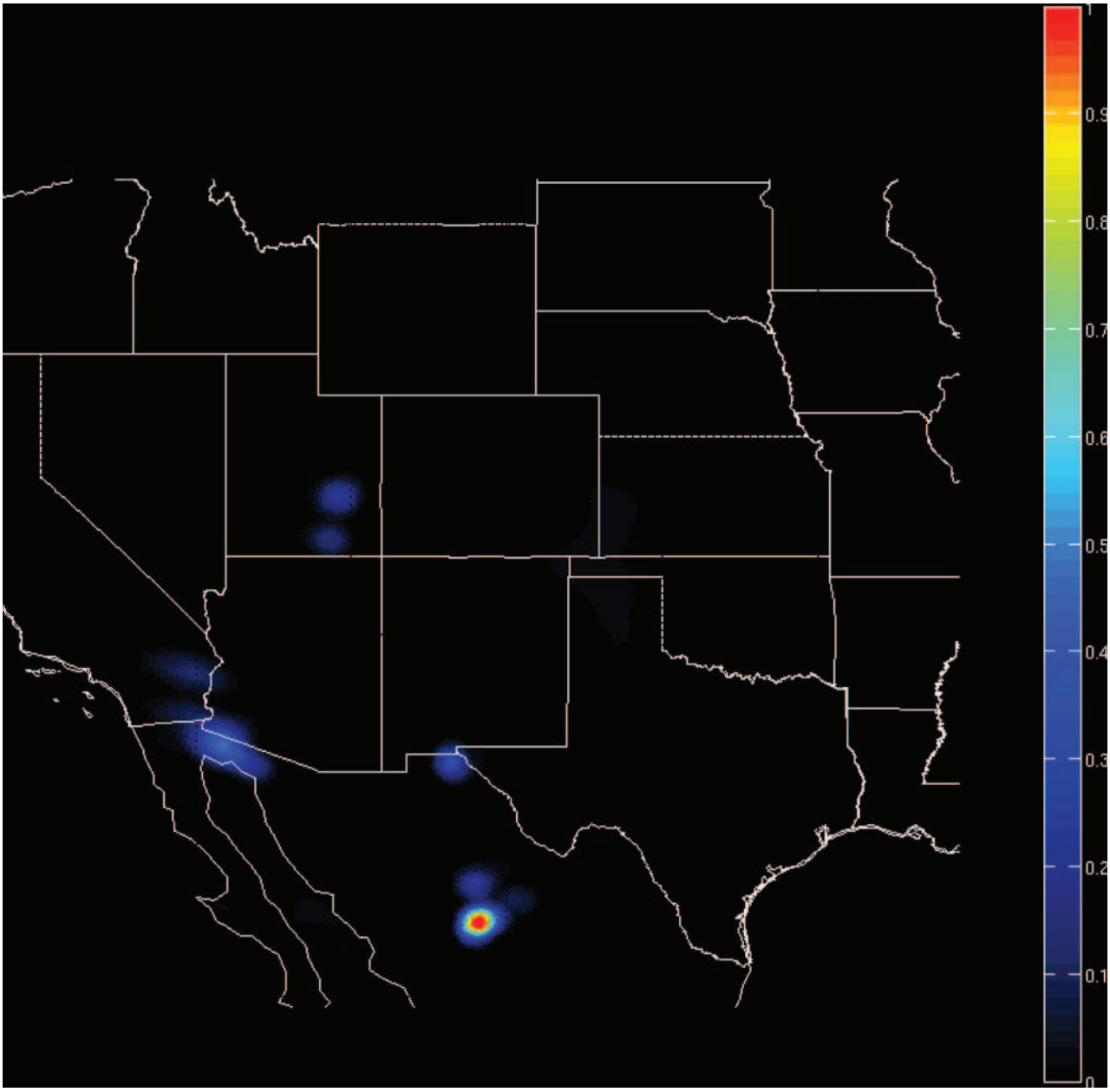

Figure 384. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

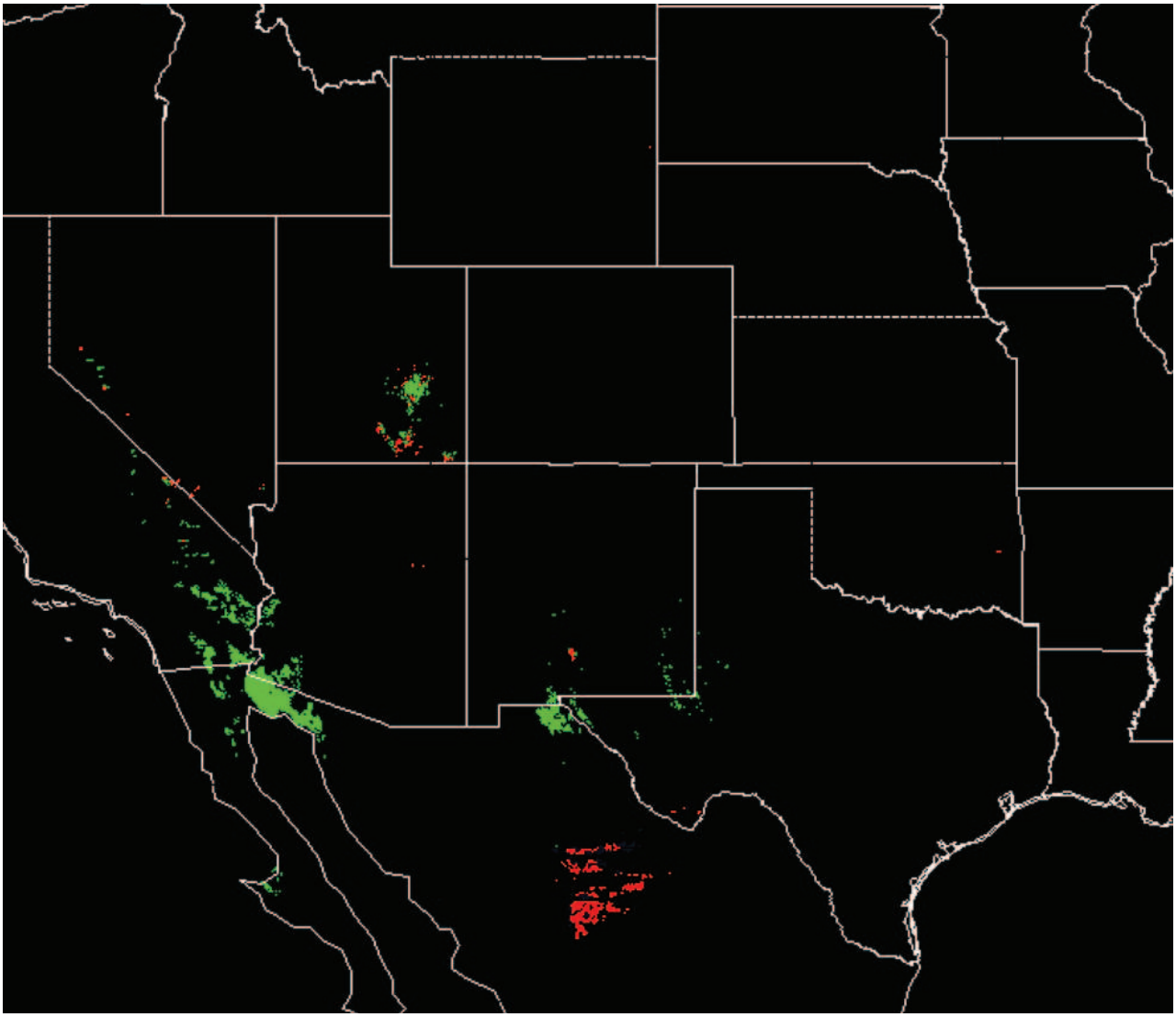

Figure 385. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-113 at 18:00 UTC.

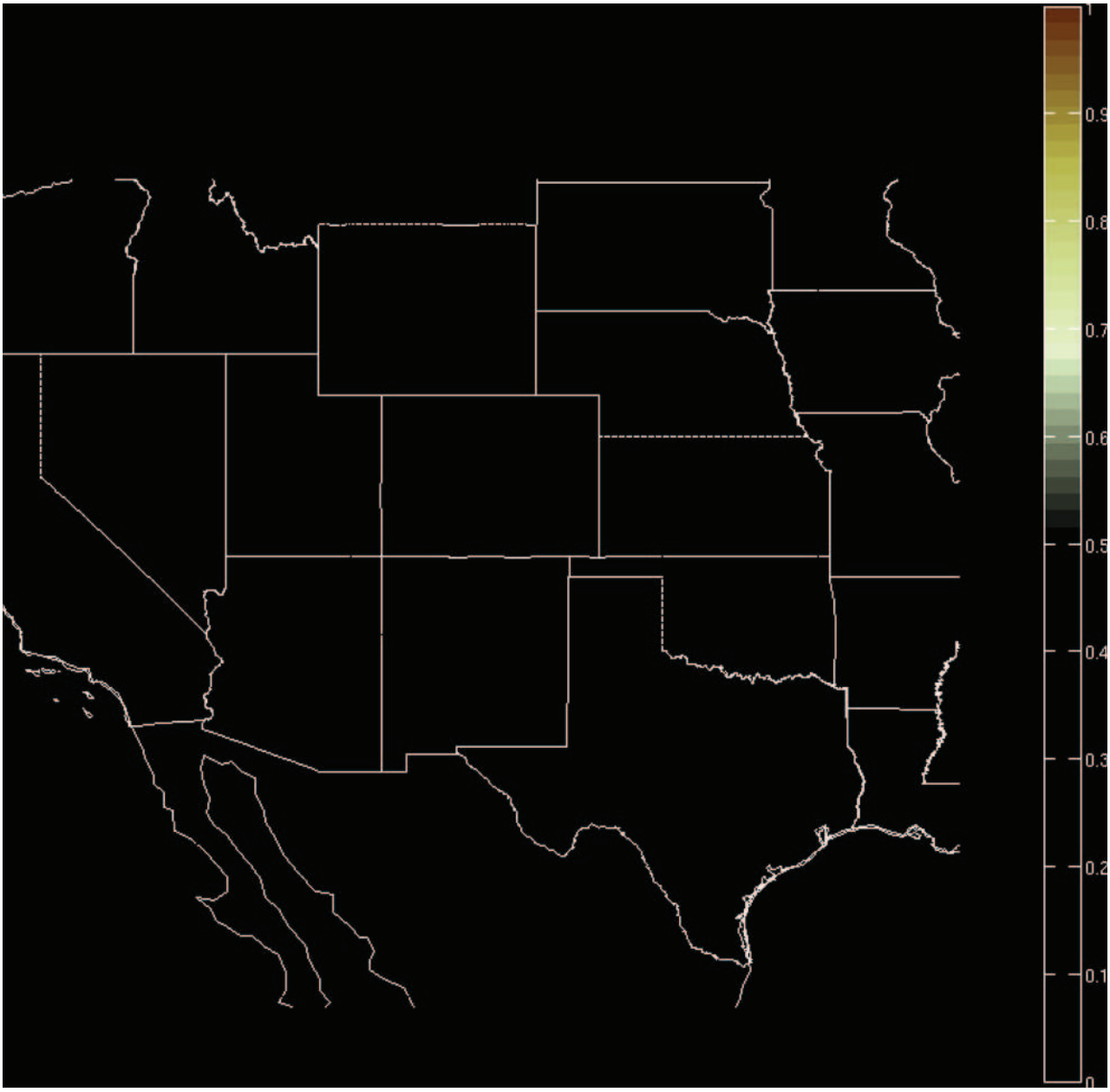

Figure 386. Result of the Maximum a posteriori model (MAP) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

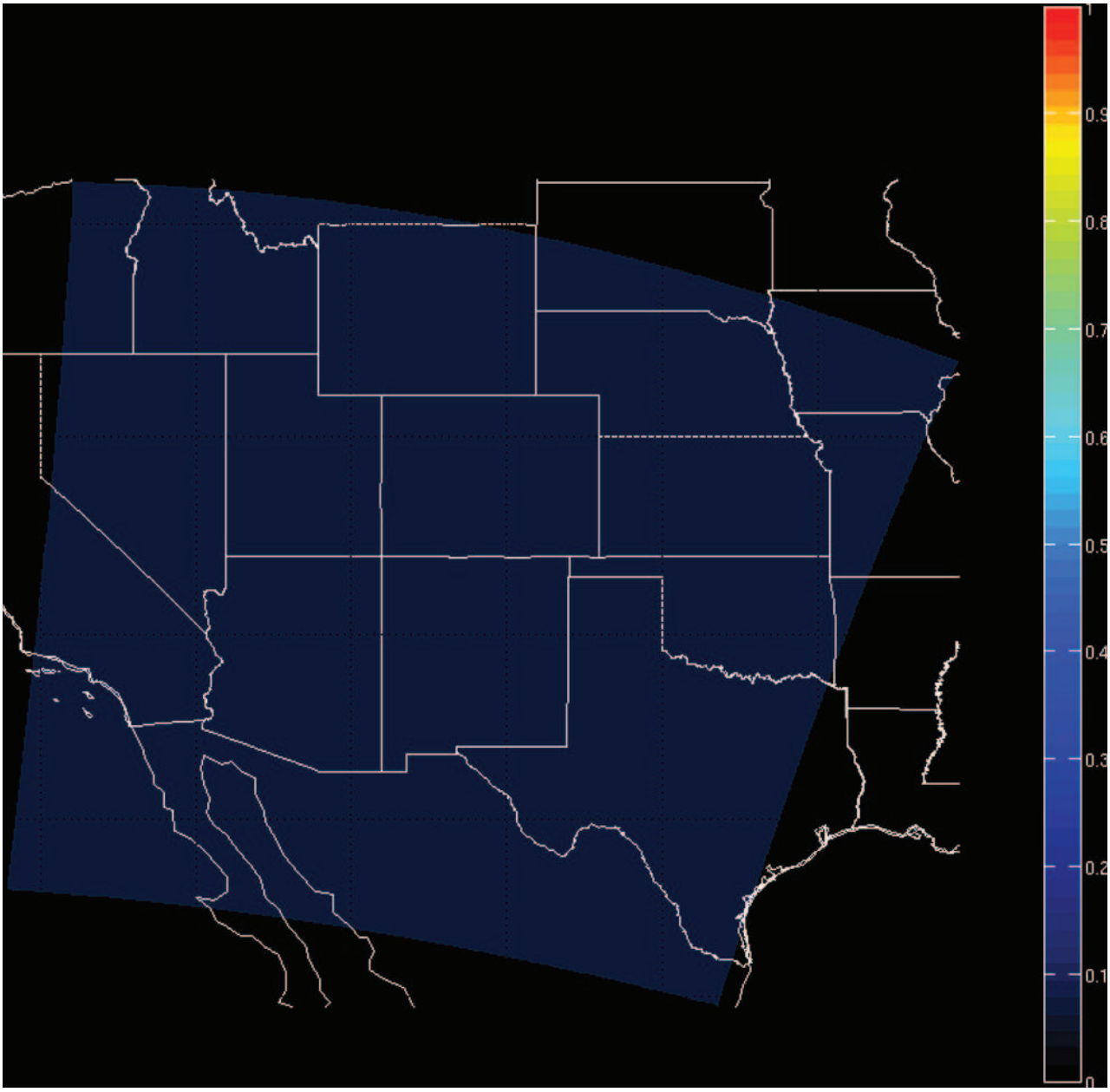

Figure 387. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

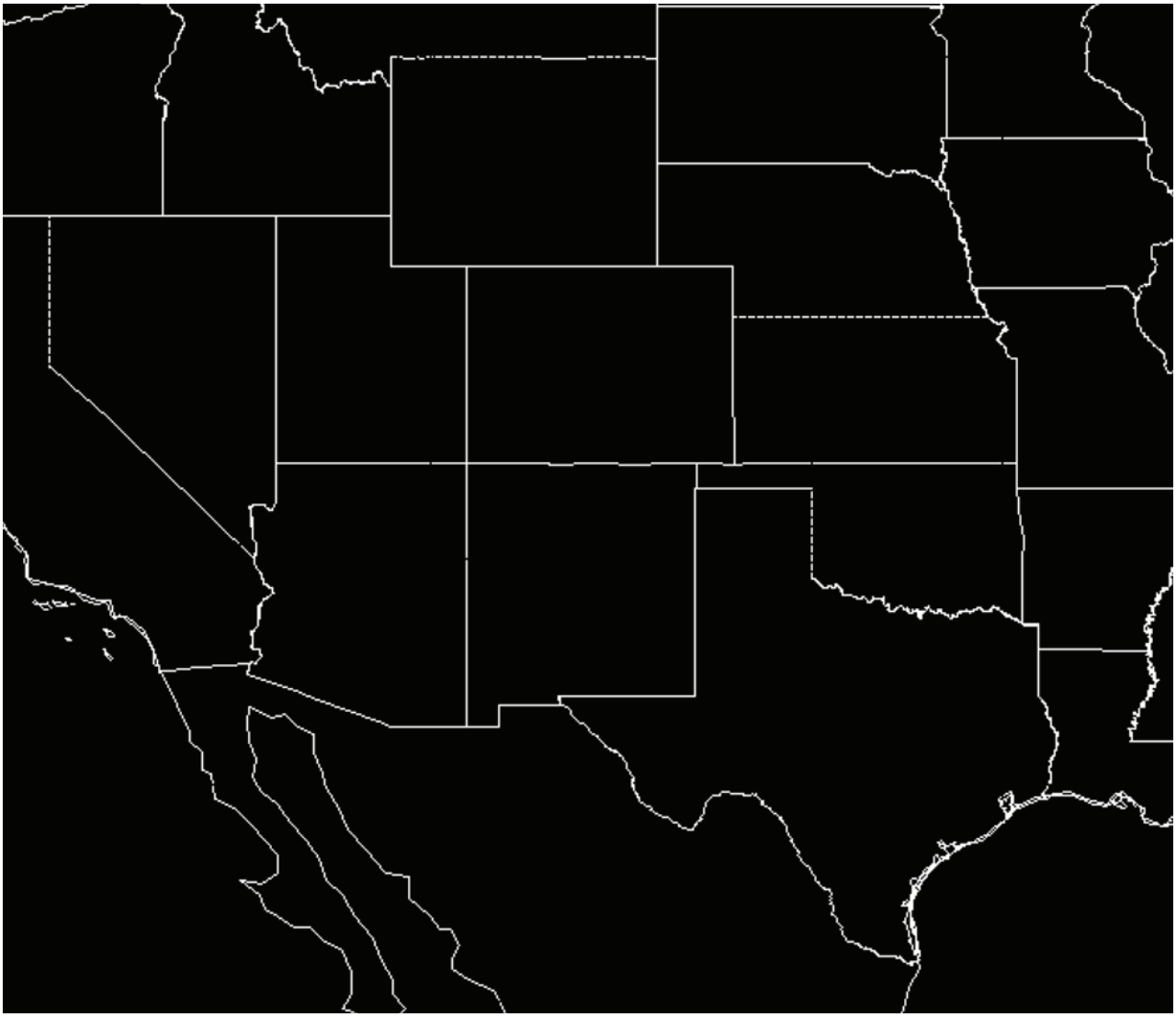

Figure 388. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-113 at 18:00 UTC.

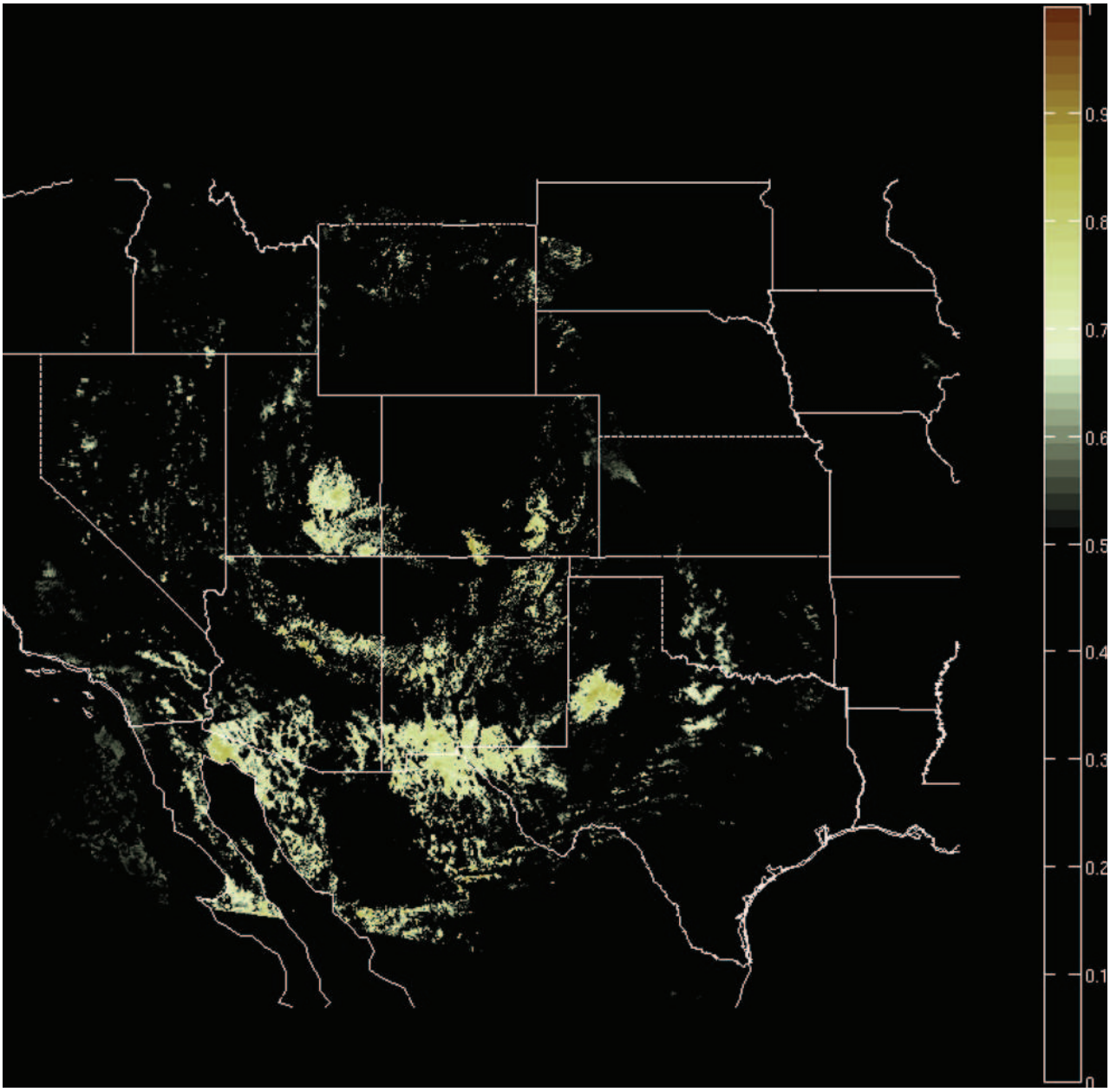

Figure 389. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

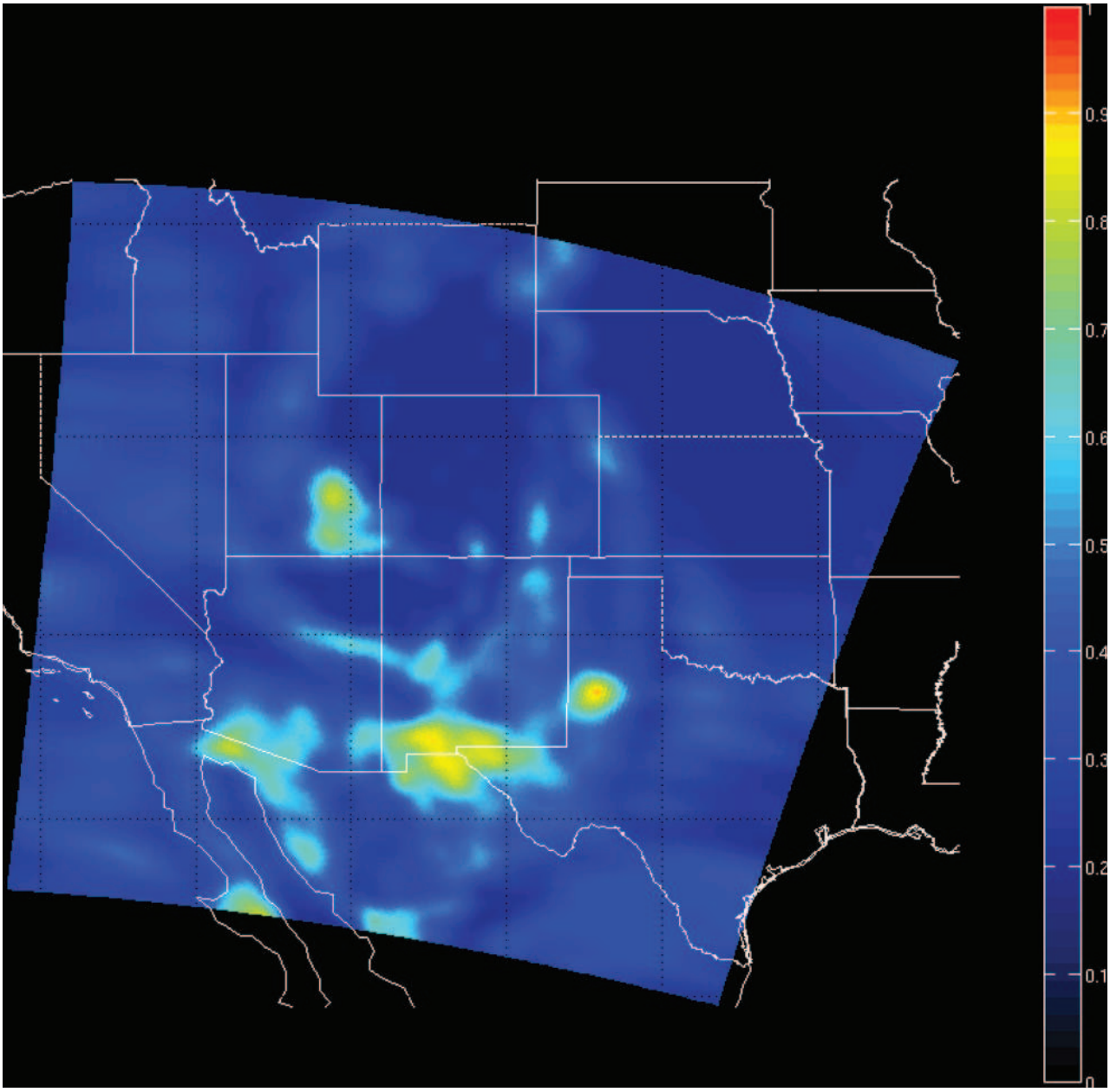

Figure 390. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

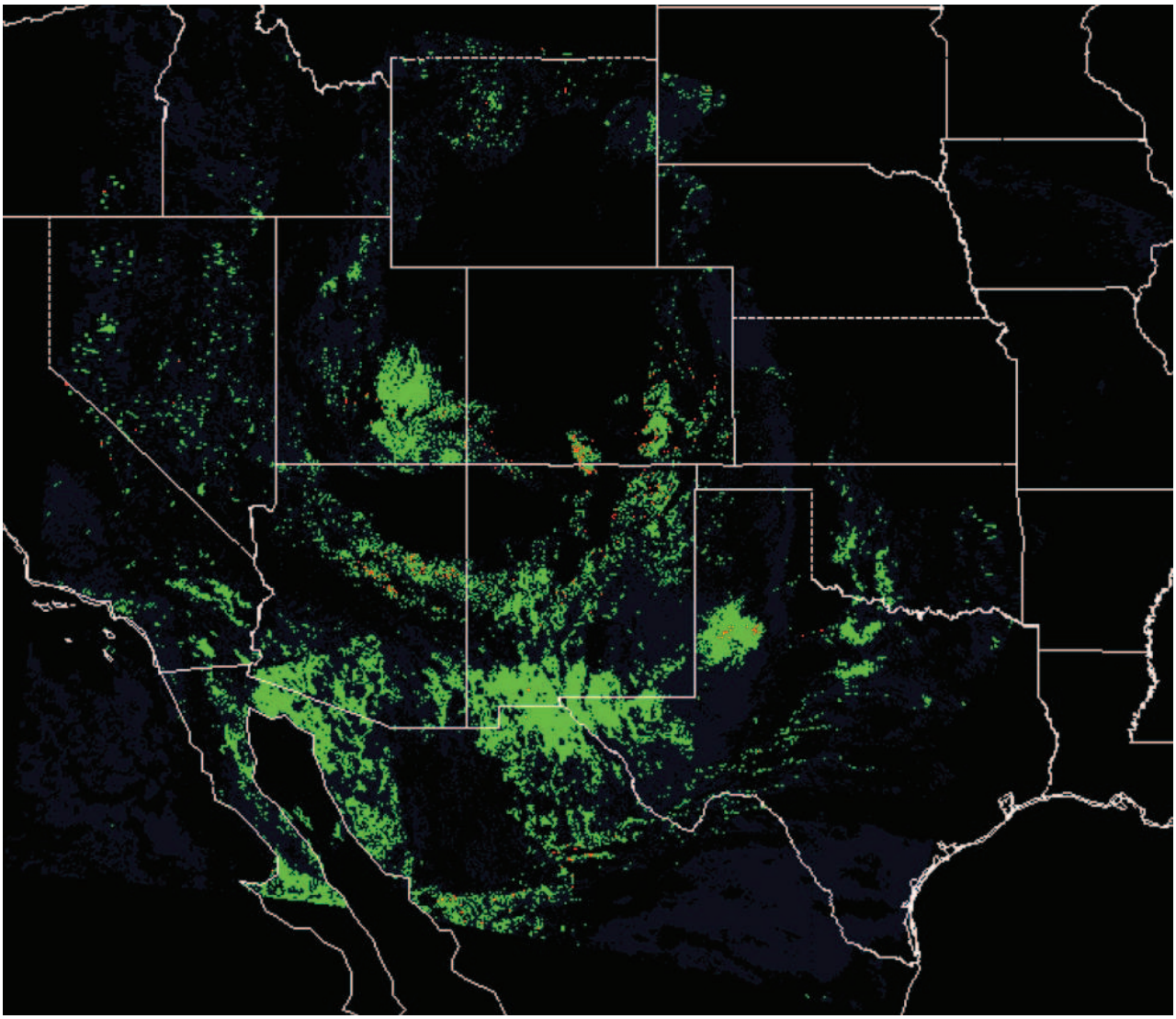

Figure 391. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-113 at 18:00 UTC.

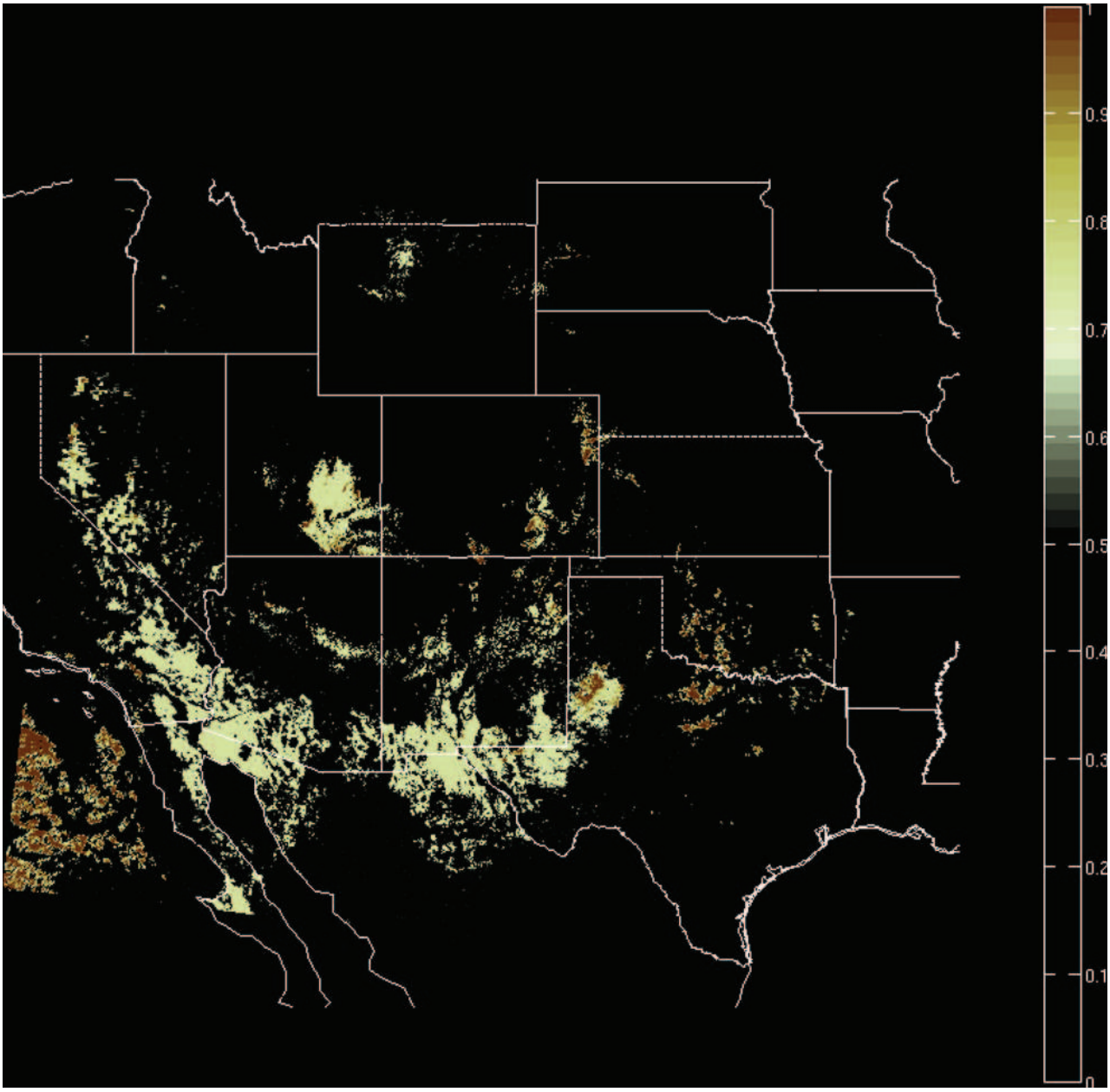

Figure 392. Result of the Probabilistic Neural Network (PNN) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

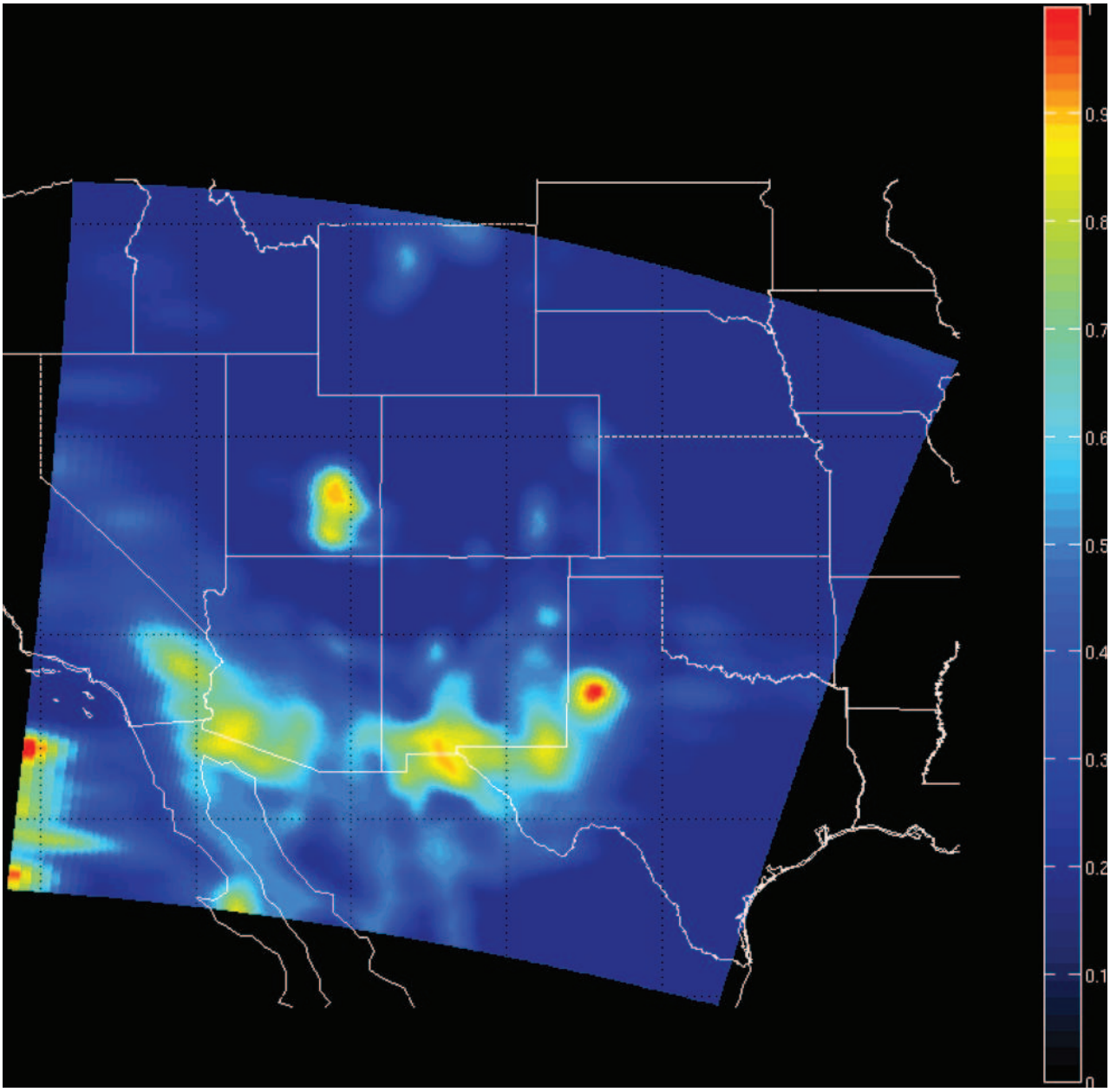

Figure 393. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-113 at 18:00 UTC. The colorbar shows the index of probability.

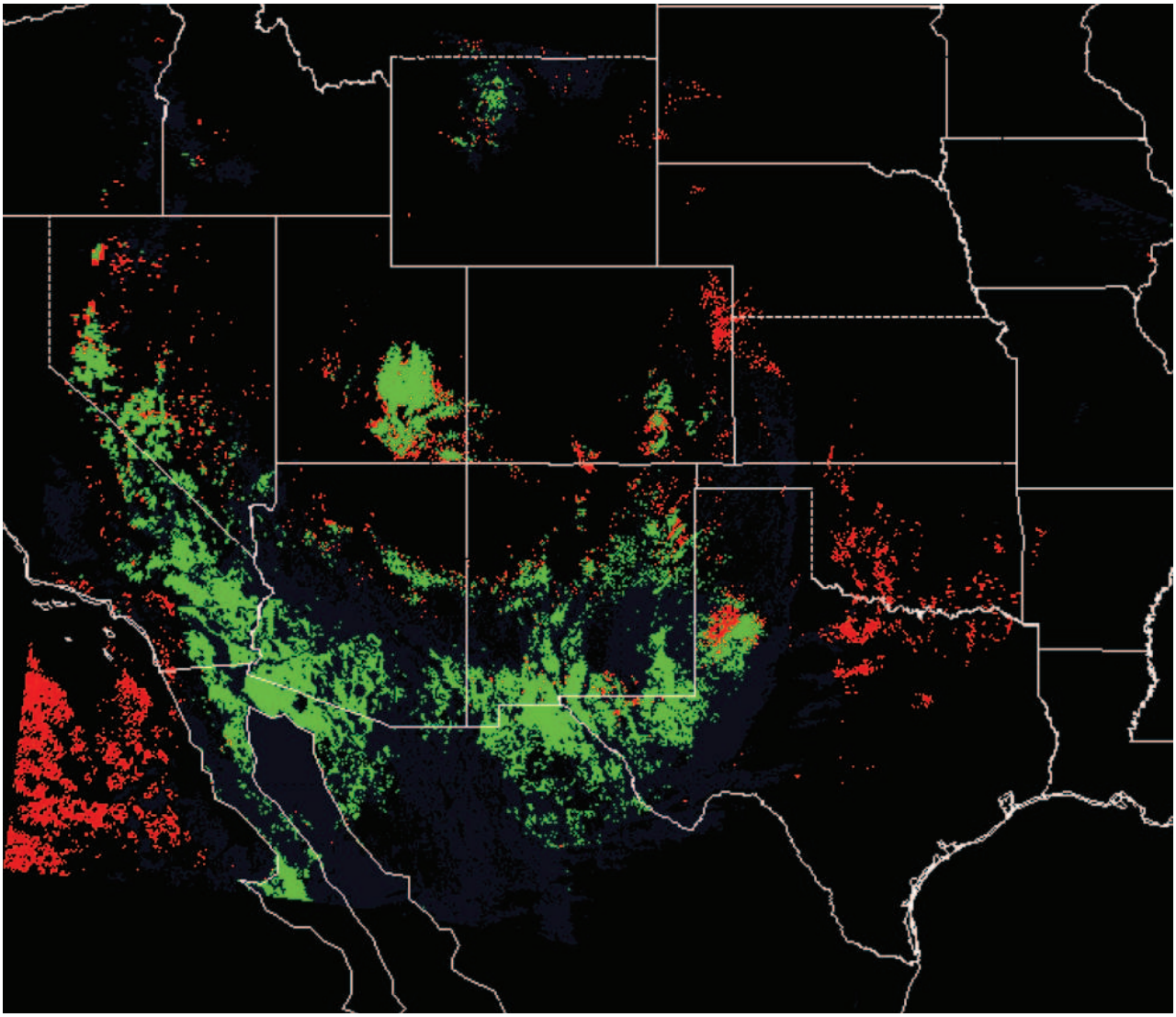

Figure 394. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-113 at 18:00 UTC.

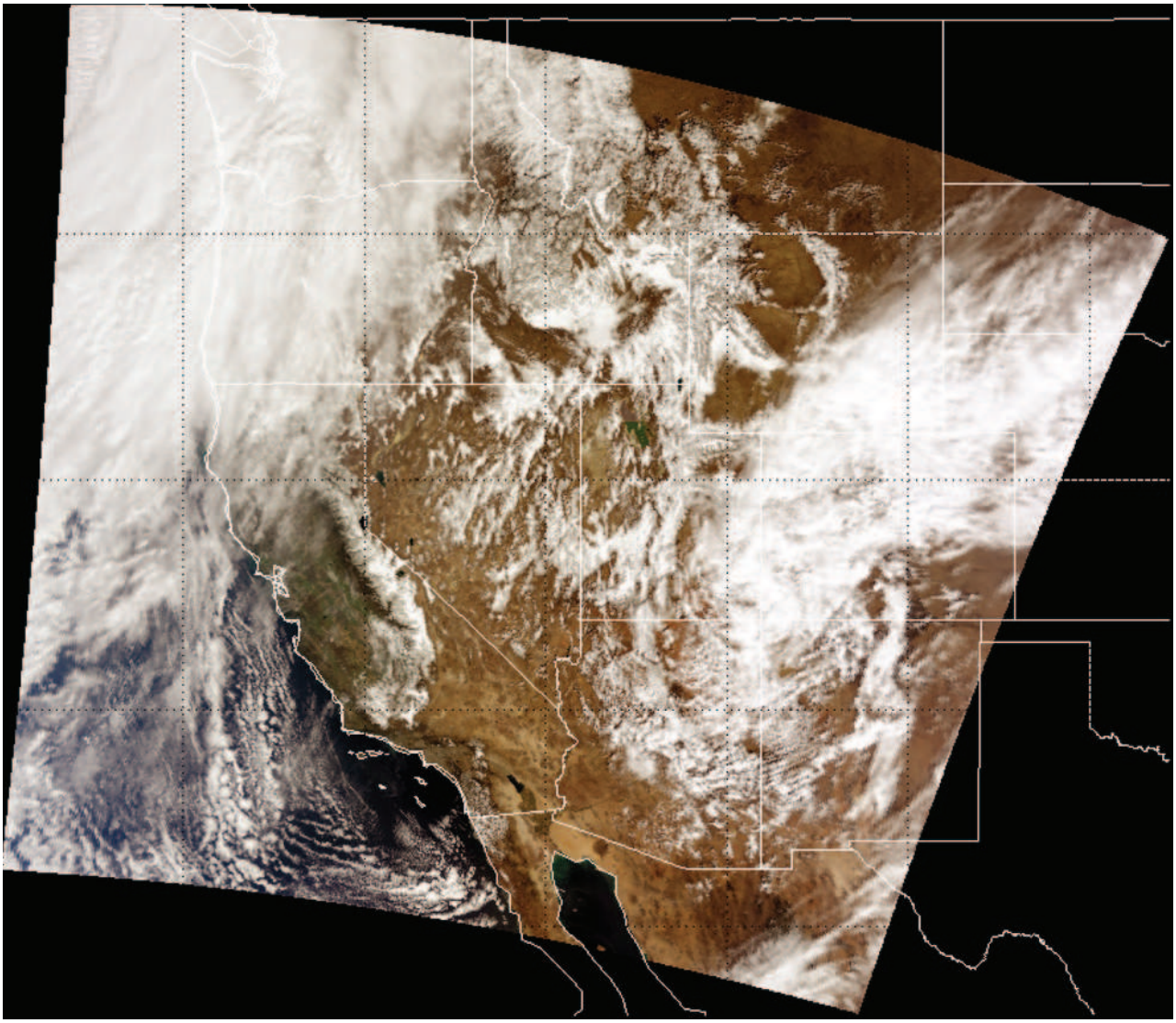

Figure 395. True color image for the event of 2006-067 at 18:35 UTC.

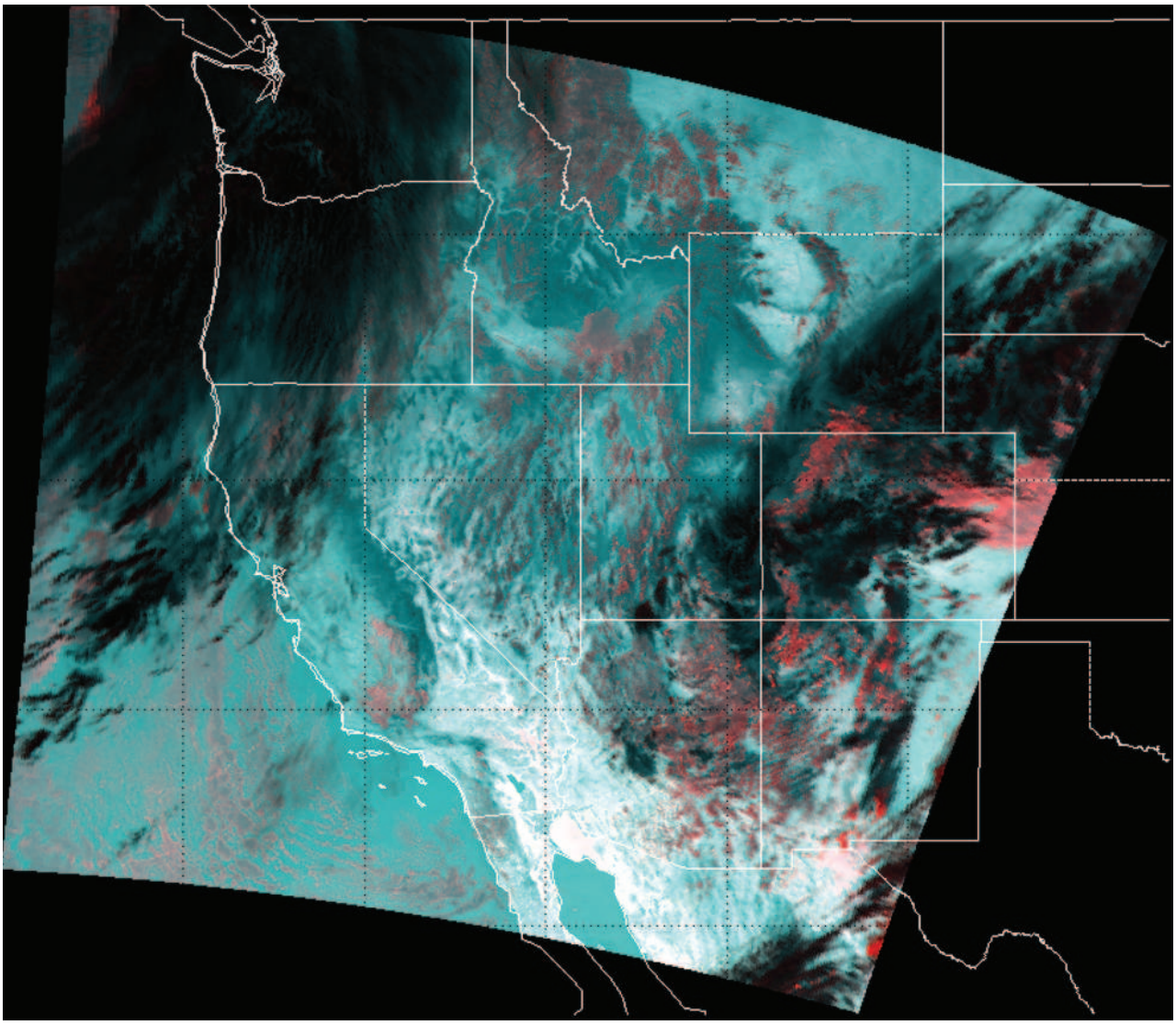

Figure 396. False color image for the event of 2006-067 at 18:35 UTC.

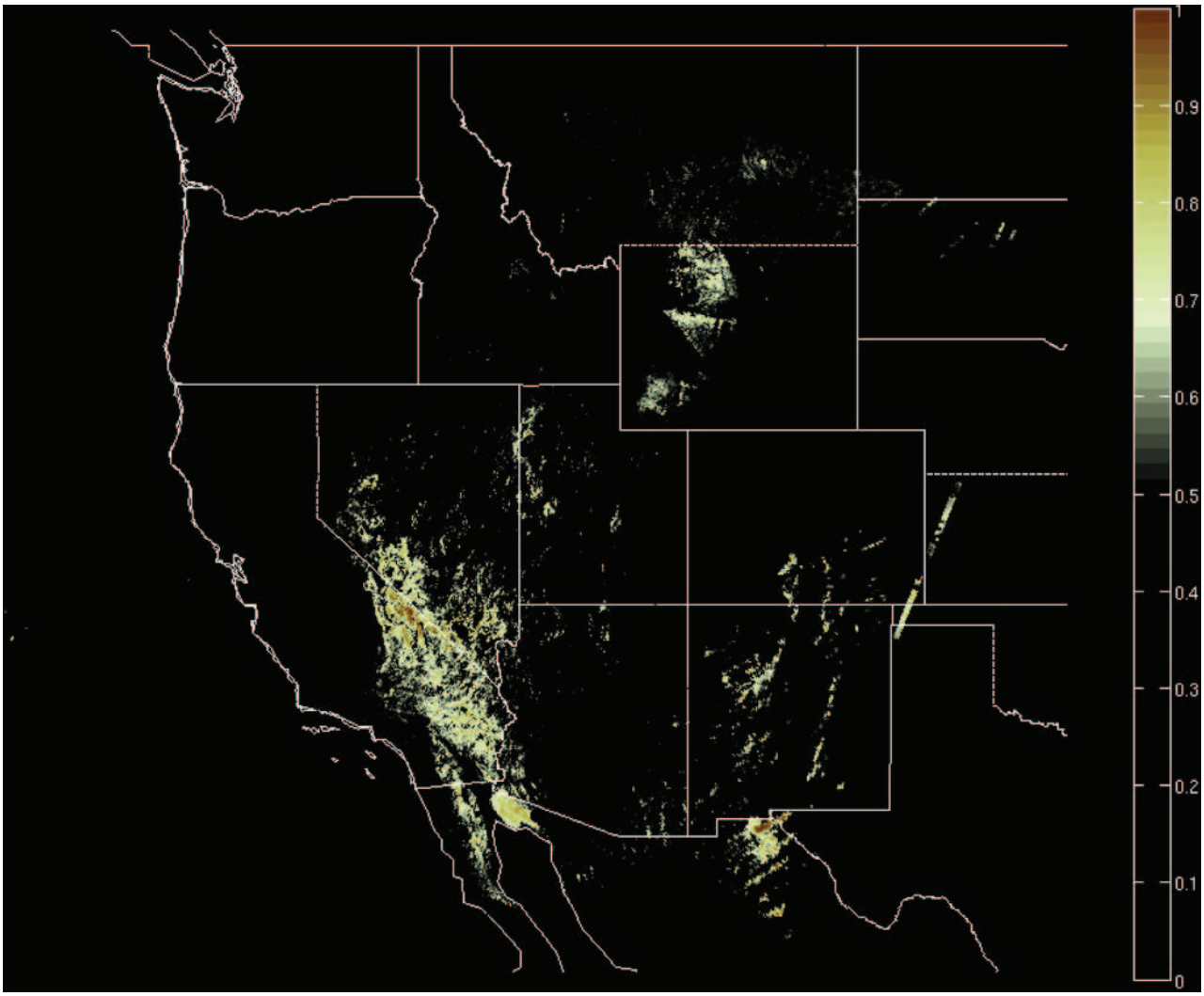

Figure 397. Result of the probabilistic display model (PD) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

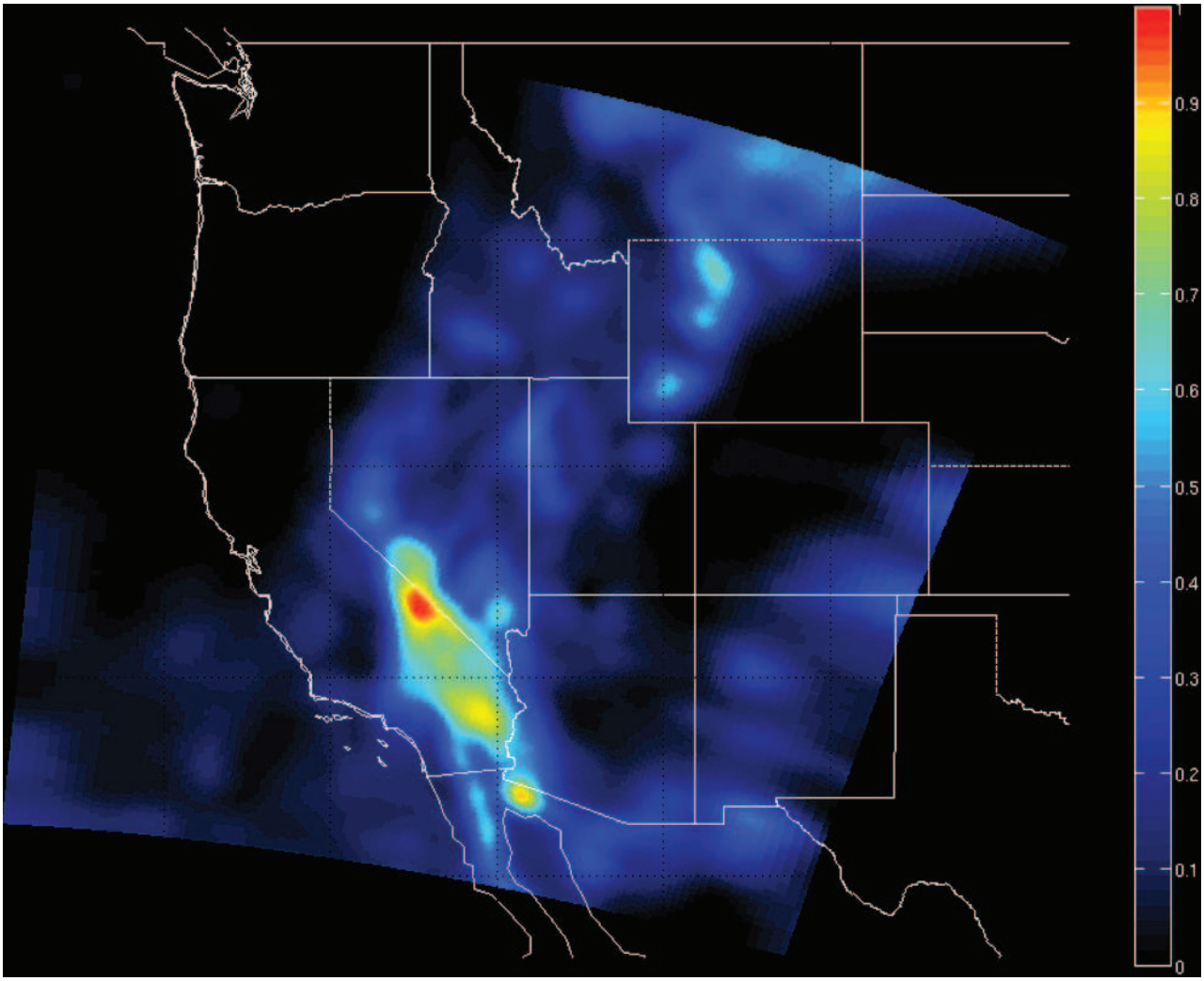

Figure 398. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

Figure 399. Result of the Maximum Likelihood model (MLE) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

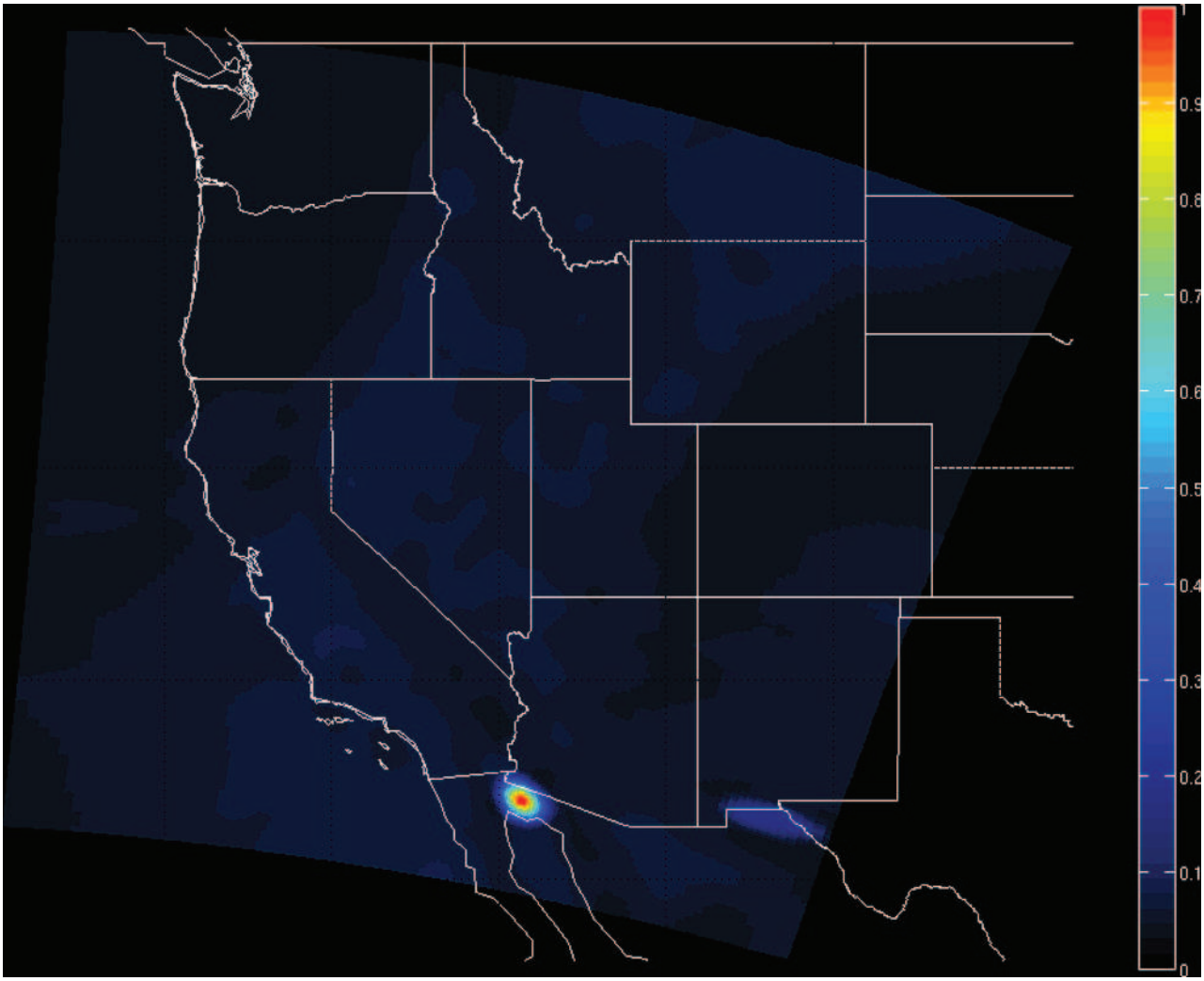

Figure 400. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

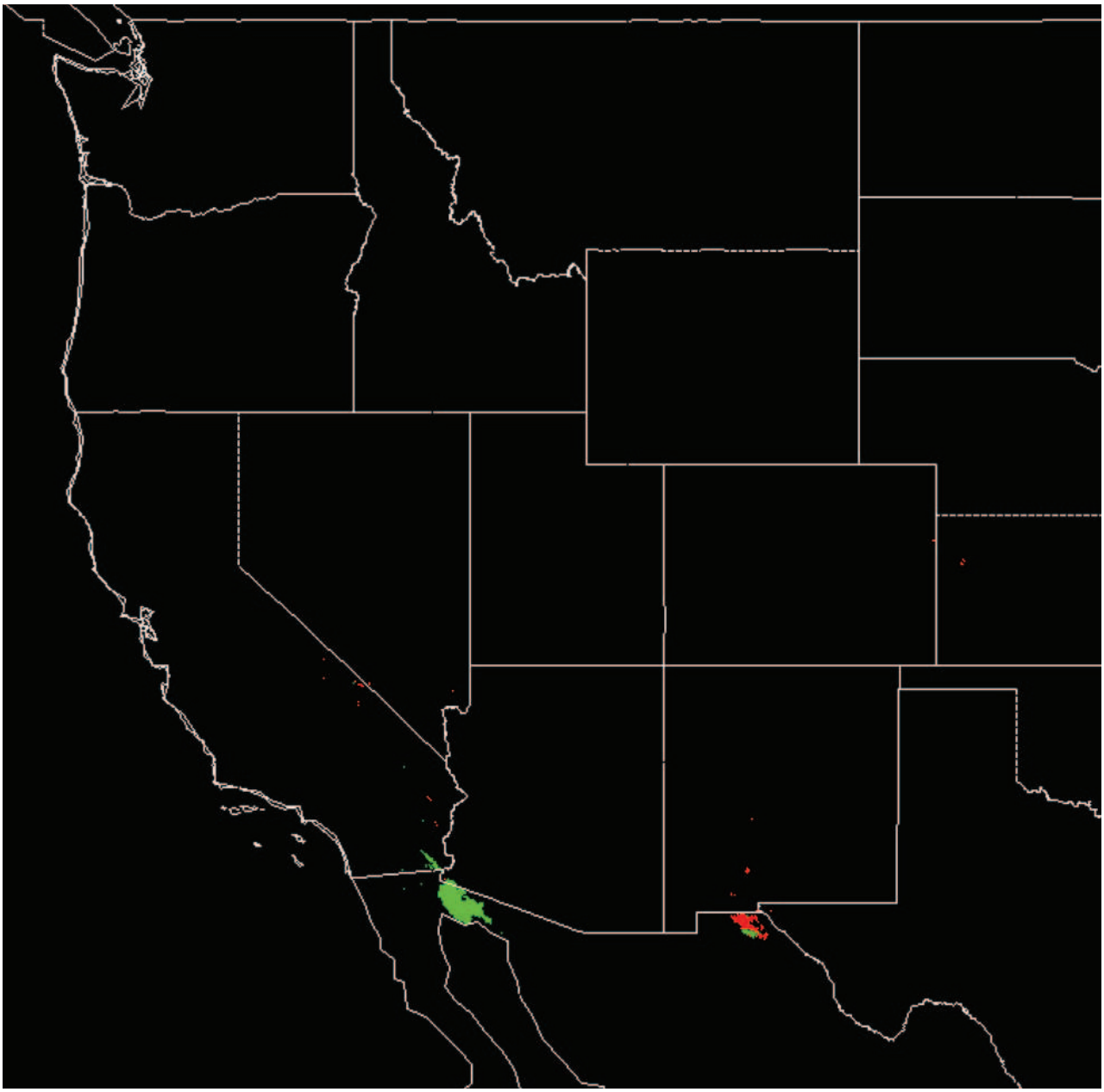

Figure 401. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2006-067 at 18:35 UTC.

Figure 402. Result of the Maximum a posteriori model (MAP) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

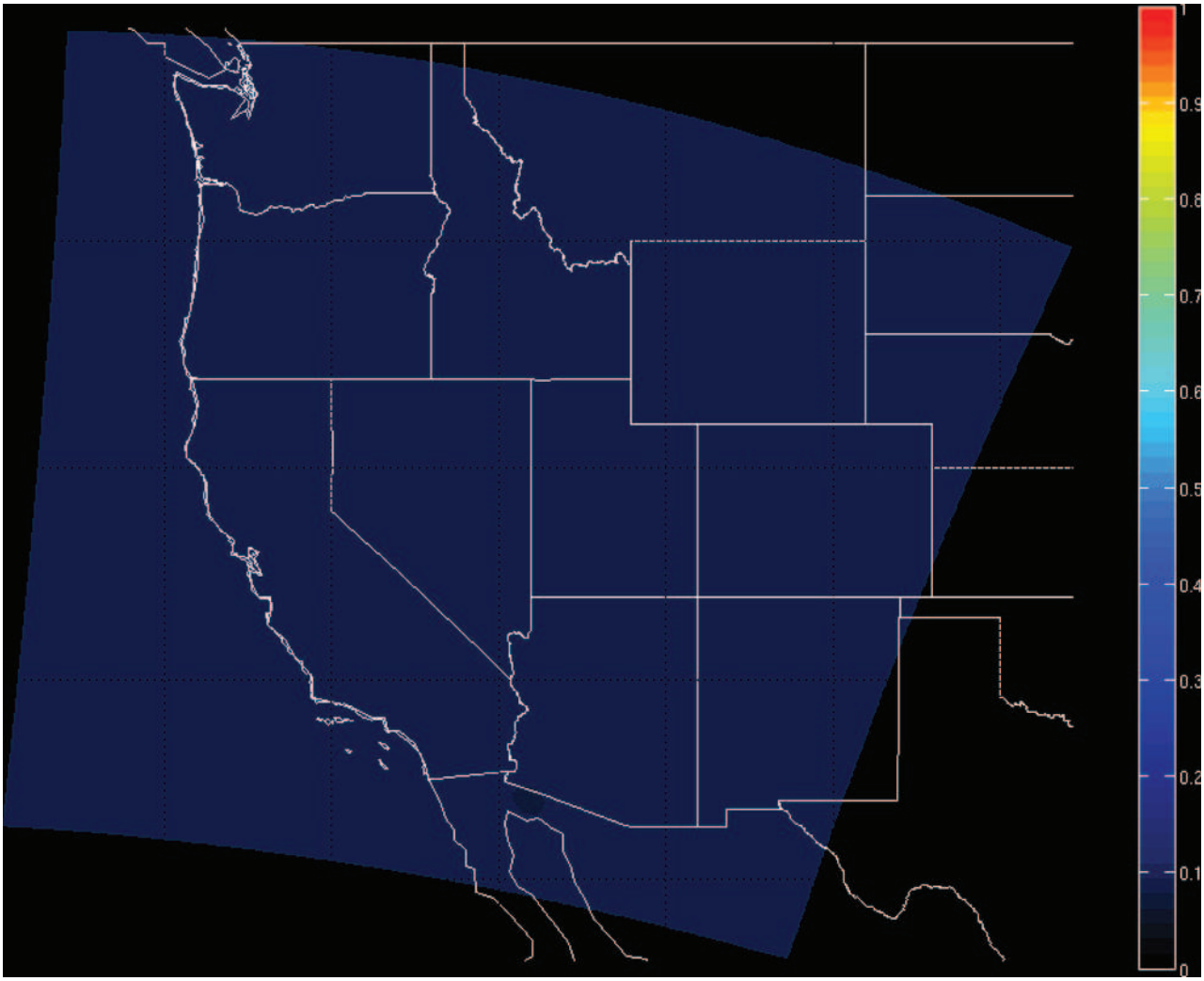

Figure 403. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

Figure 404. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2006-067 at 18:35 UTC.

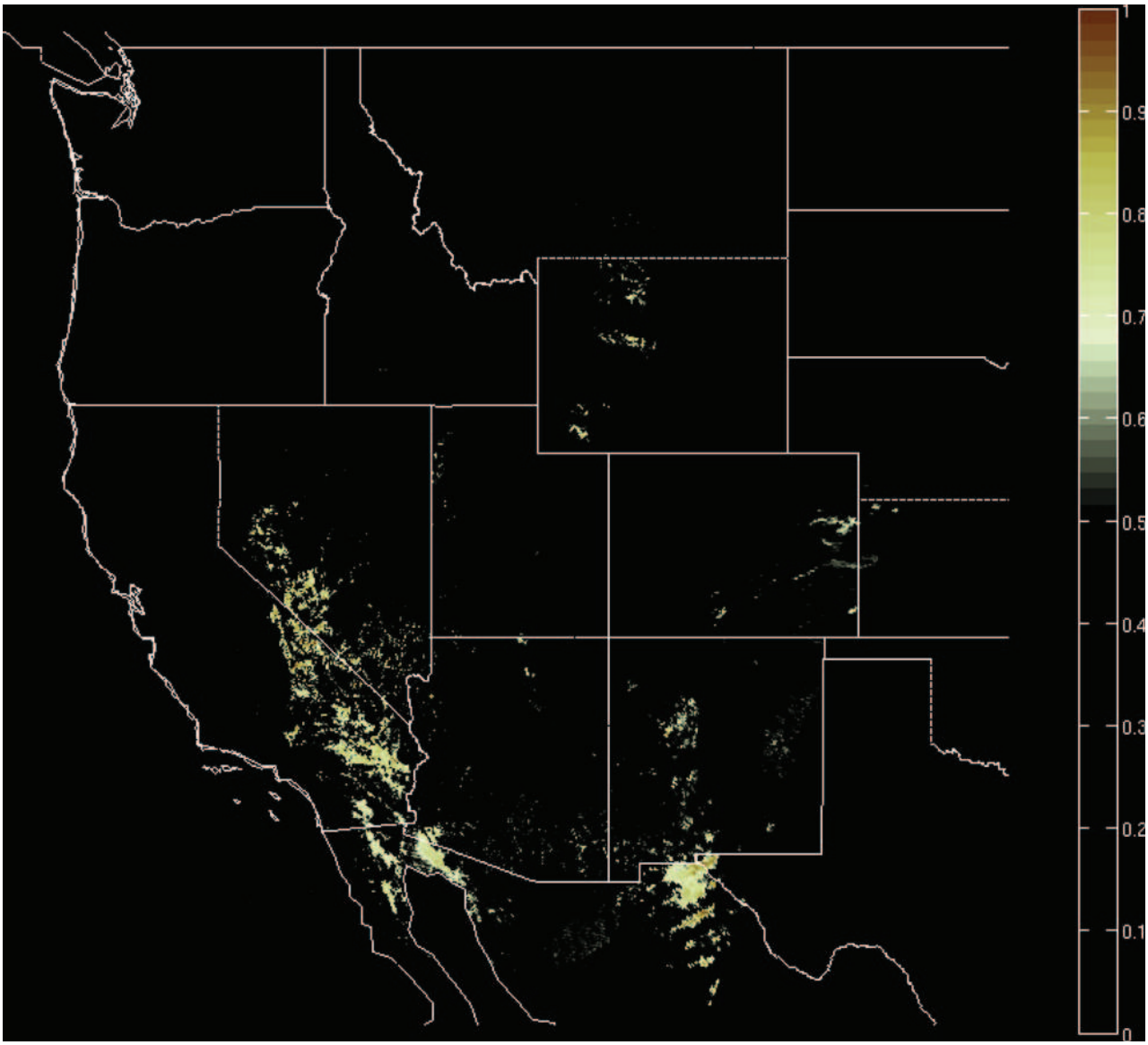

Figure 405. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

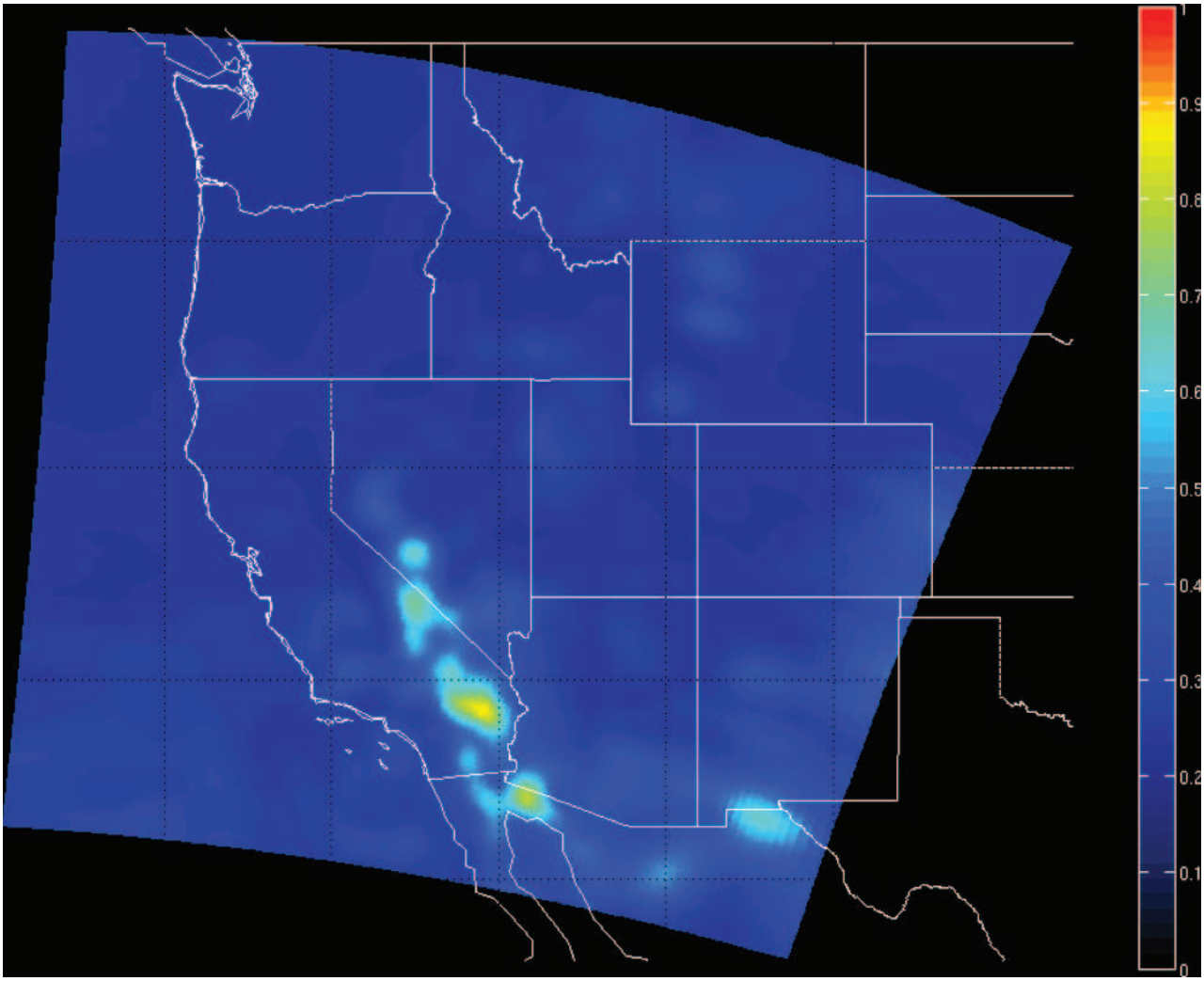

Figure 406. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

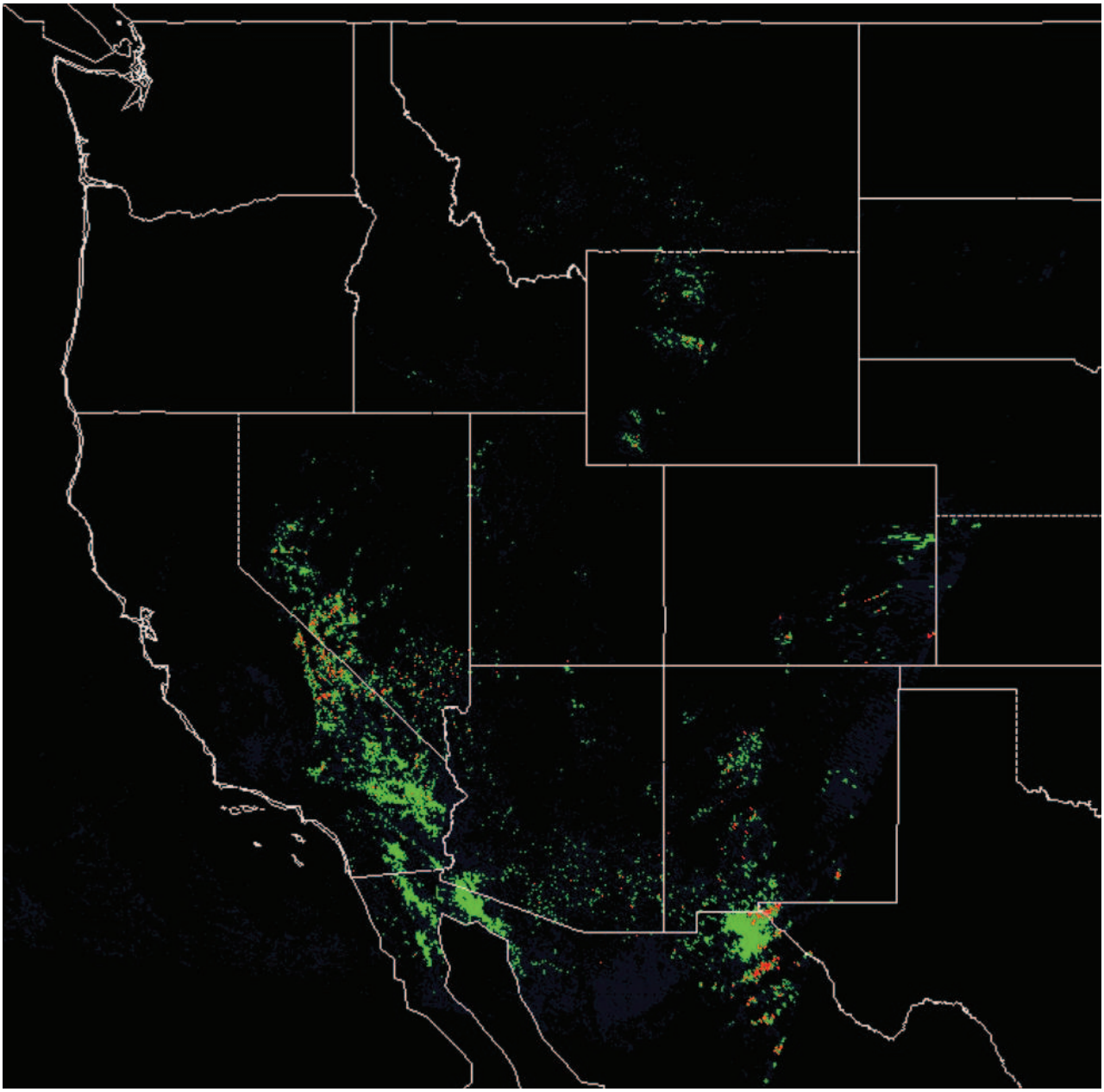

Figure 407. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-067 at 18:35 UTC.

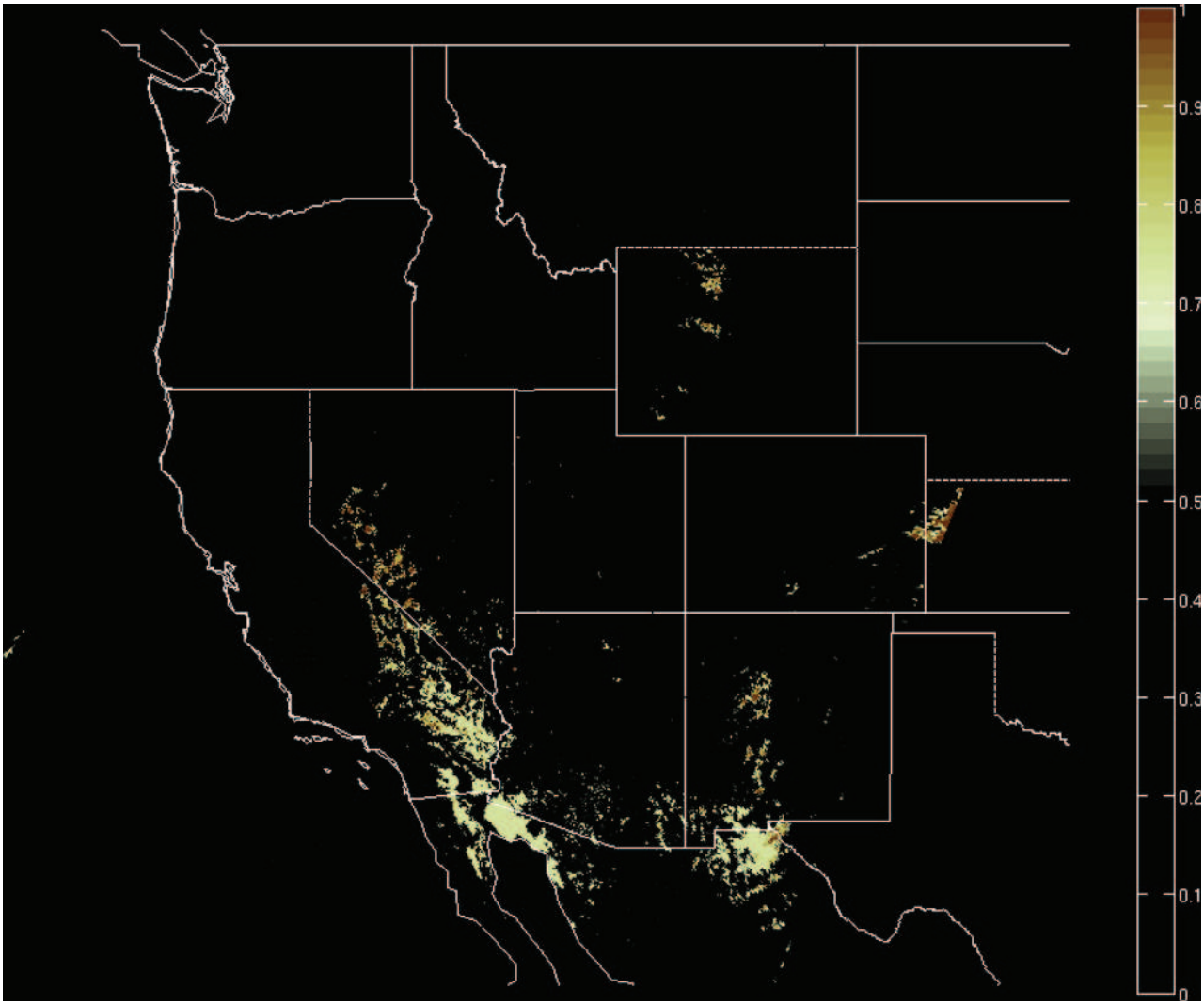

Figure 408. Result of the Probabilistic Neural Network (PNN) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

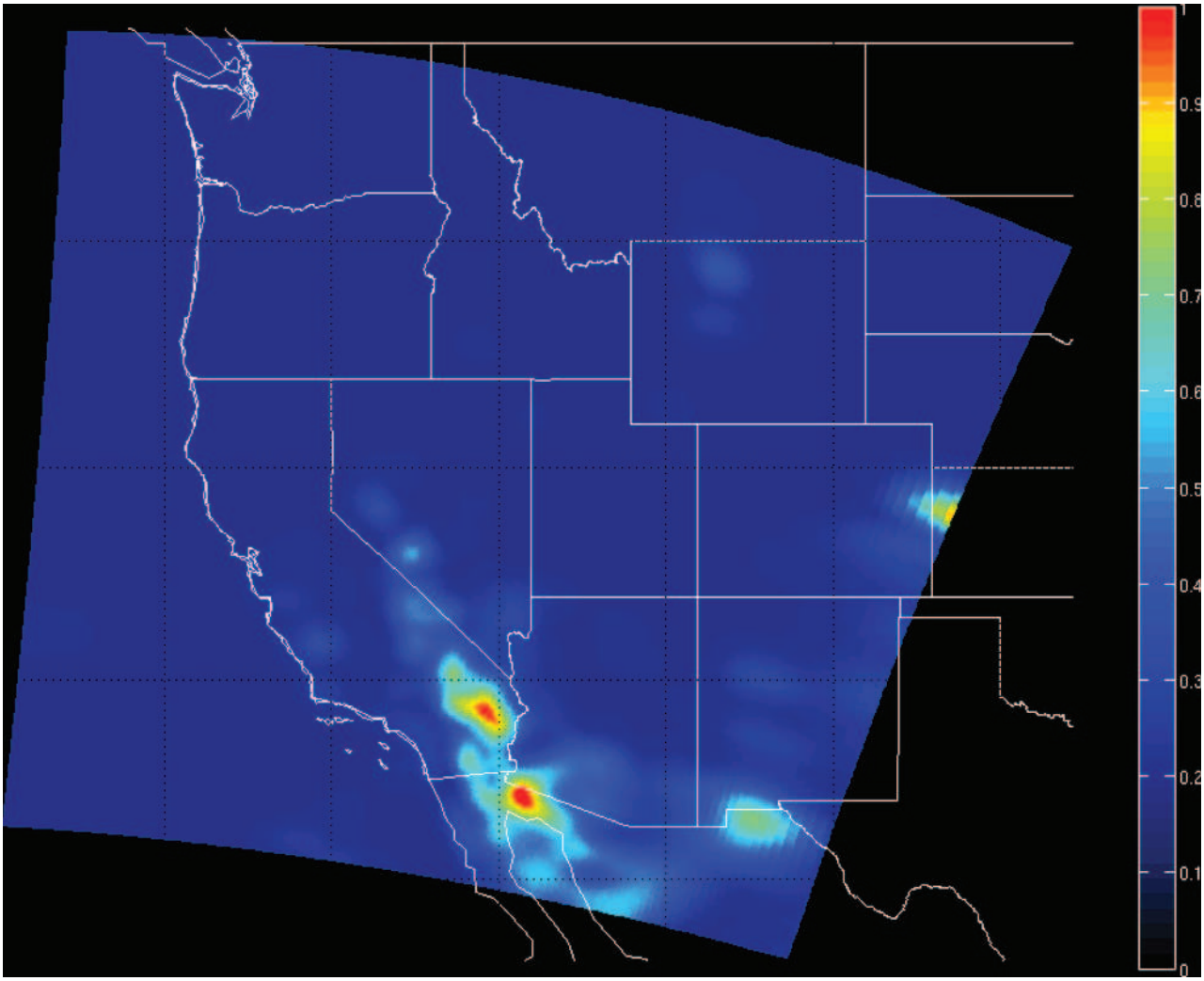

Figure 409. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2006-067 at 18:35 UTC. The colorbar shows the index of probability.

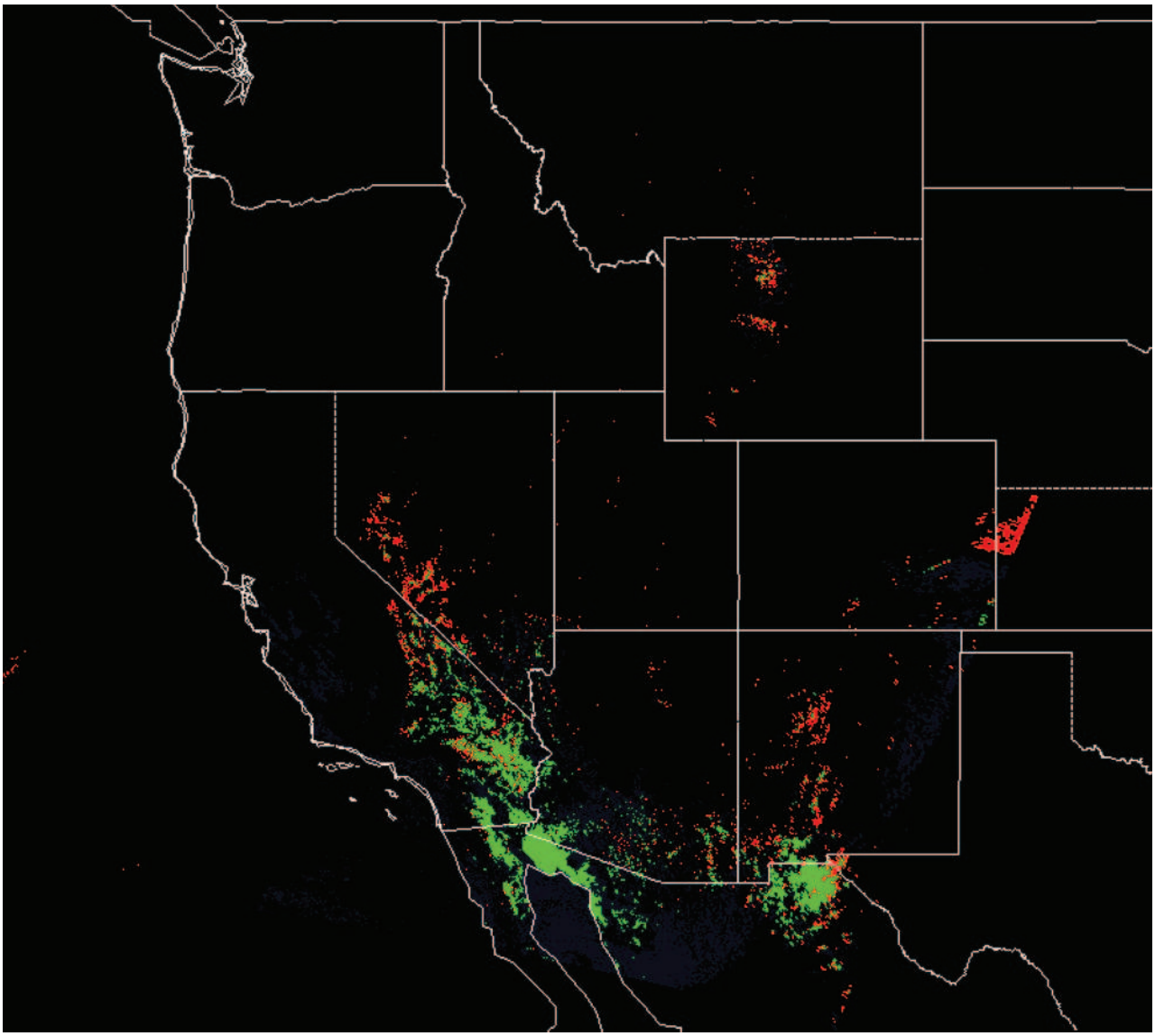

Figure 410. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2006-067 at 18:35 UTC.

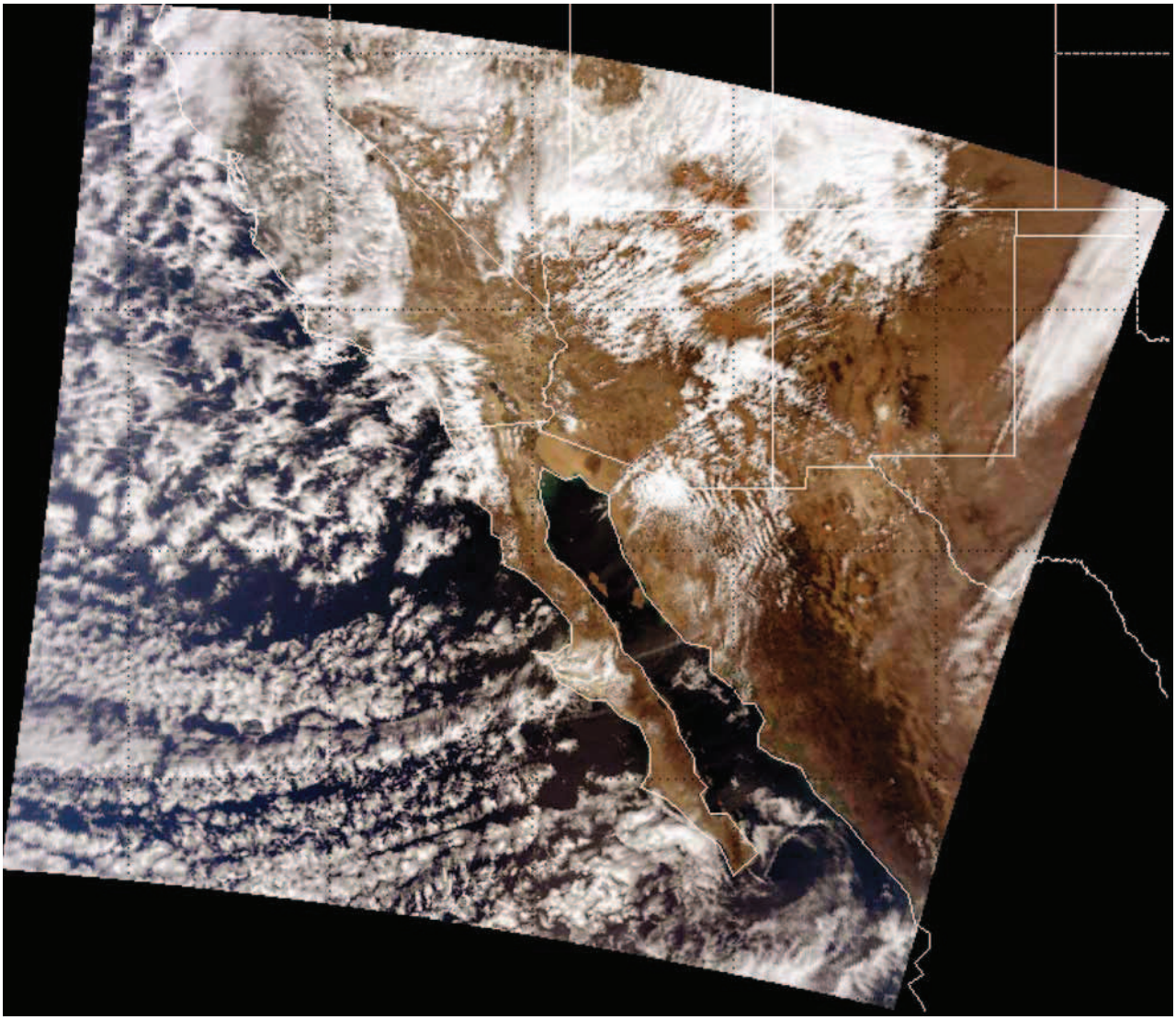

Figure 411. True color image for the event of 2006-069 at 18:25 UTC.

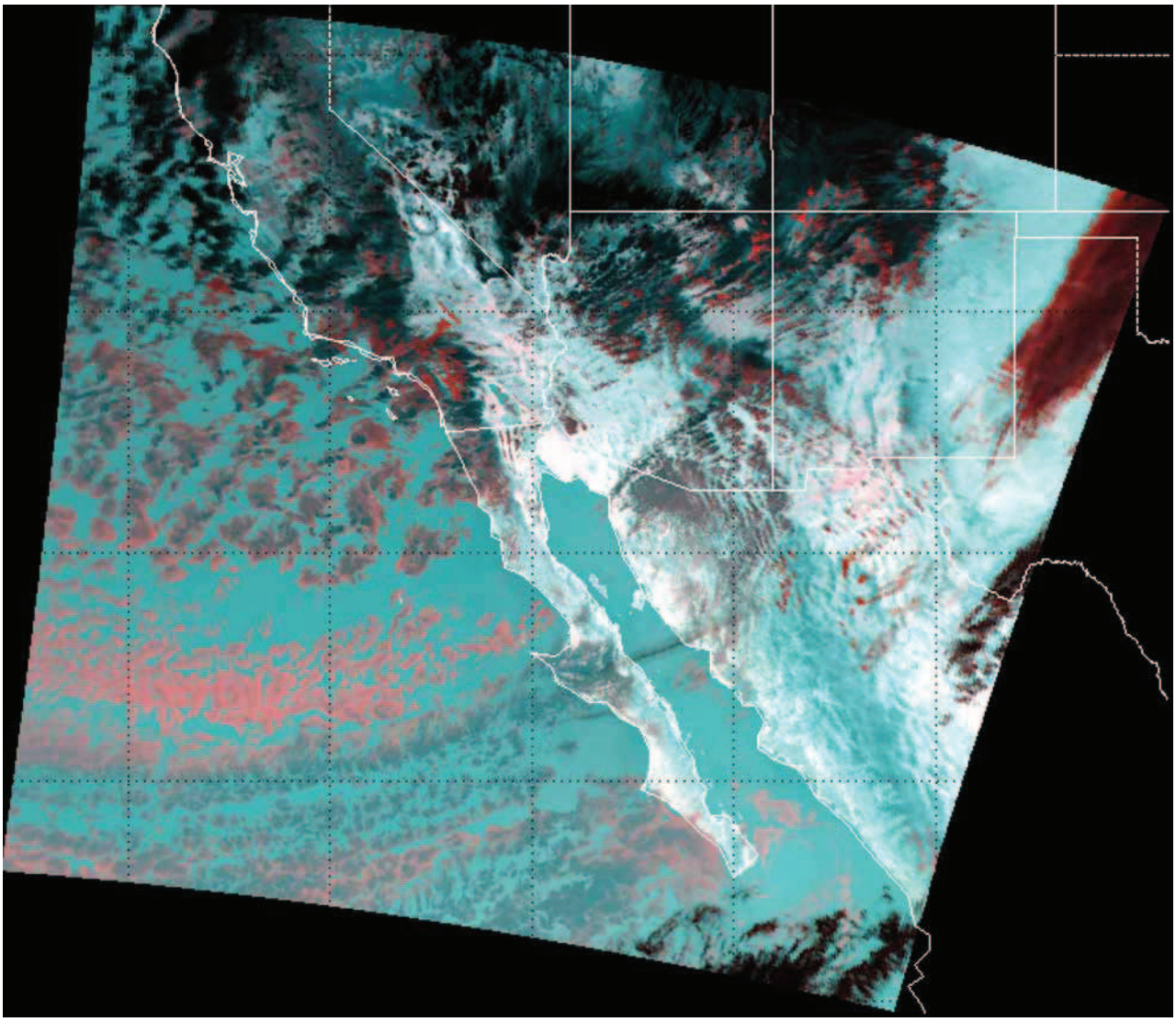

Figure 412. False color image for the event of 2006-069 at 18:25 UTC.

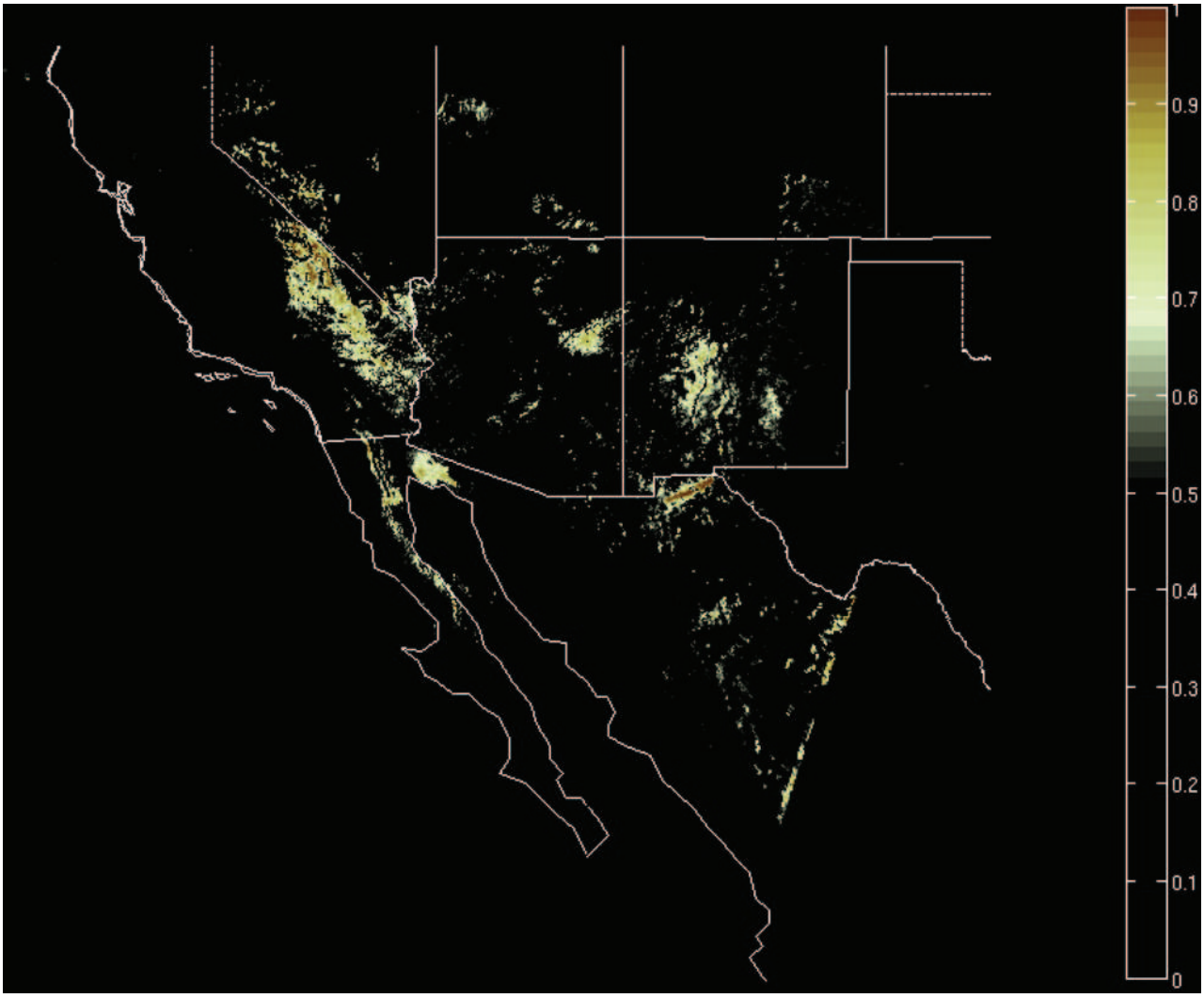

Figure 413. Result of the probabilistic display model (PD) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

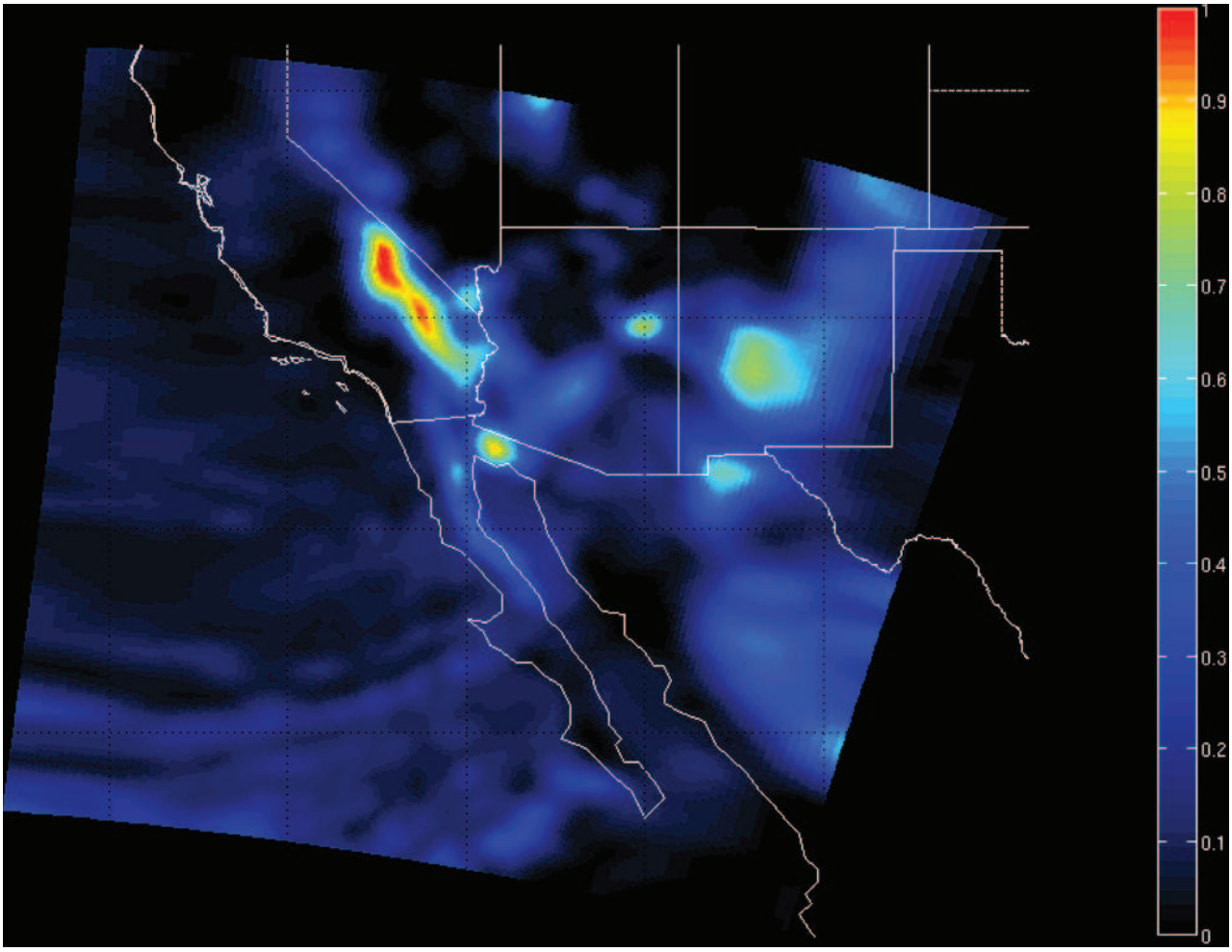

Figure 414. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

Figure 415. Result of the Maximum Likelihood model (MLE) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

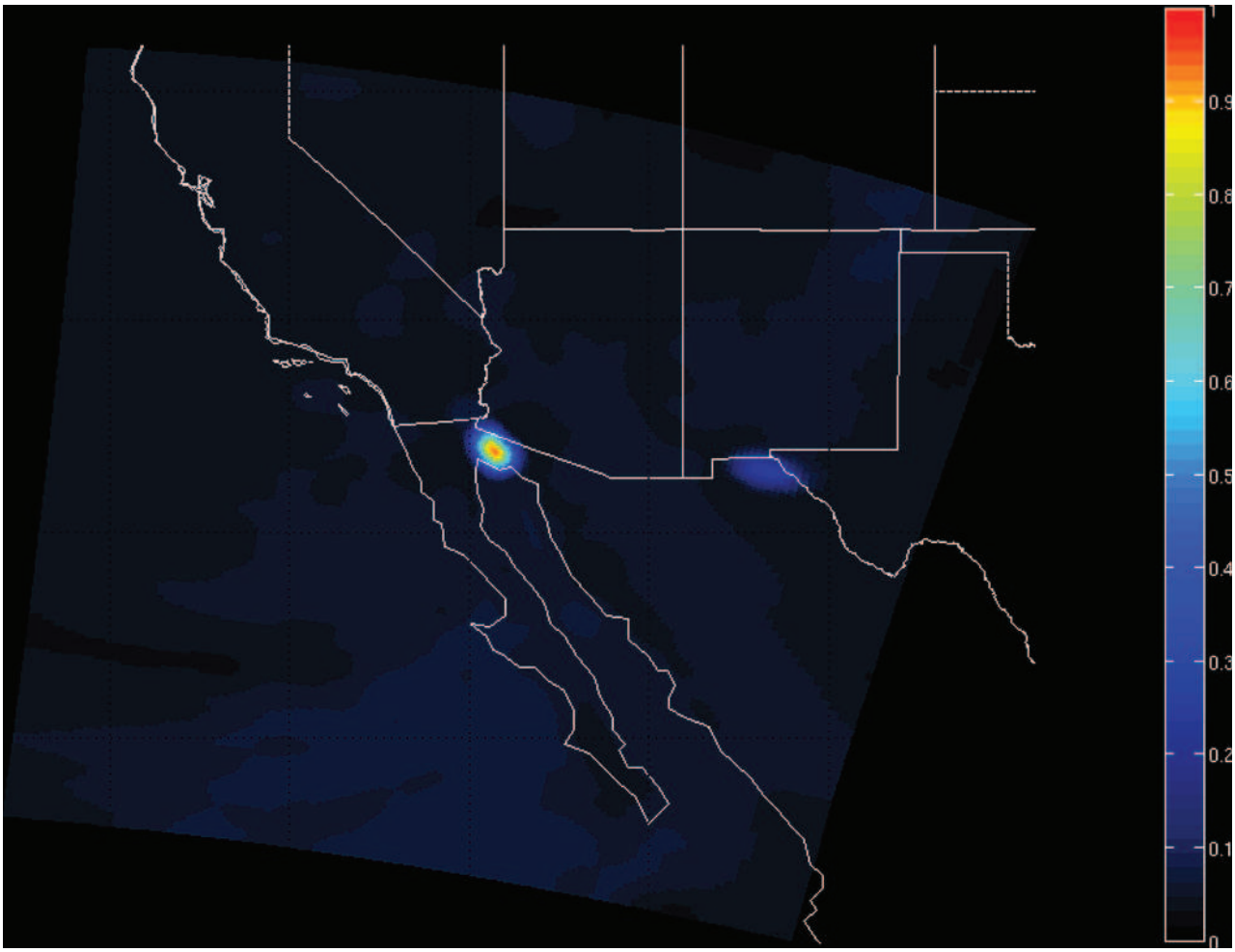

Figure 416. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

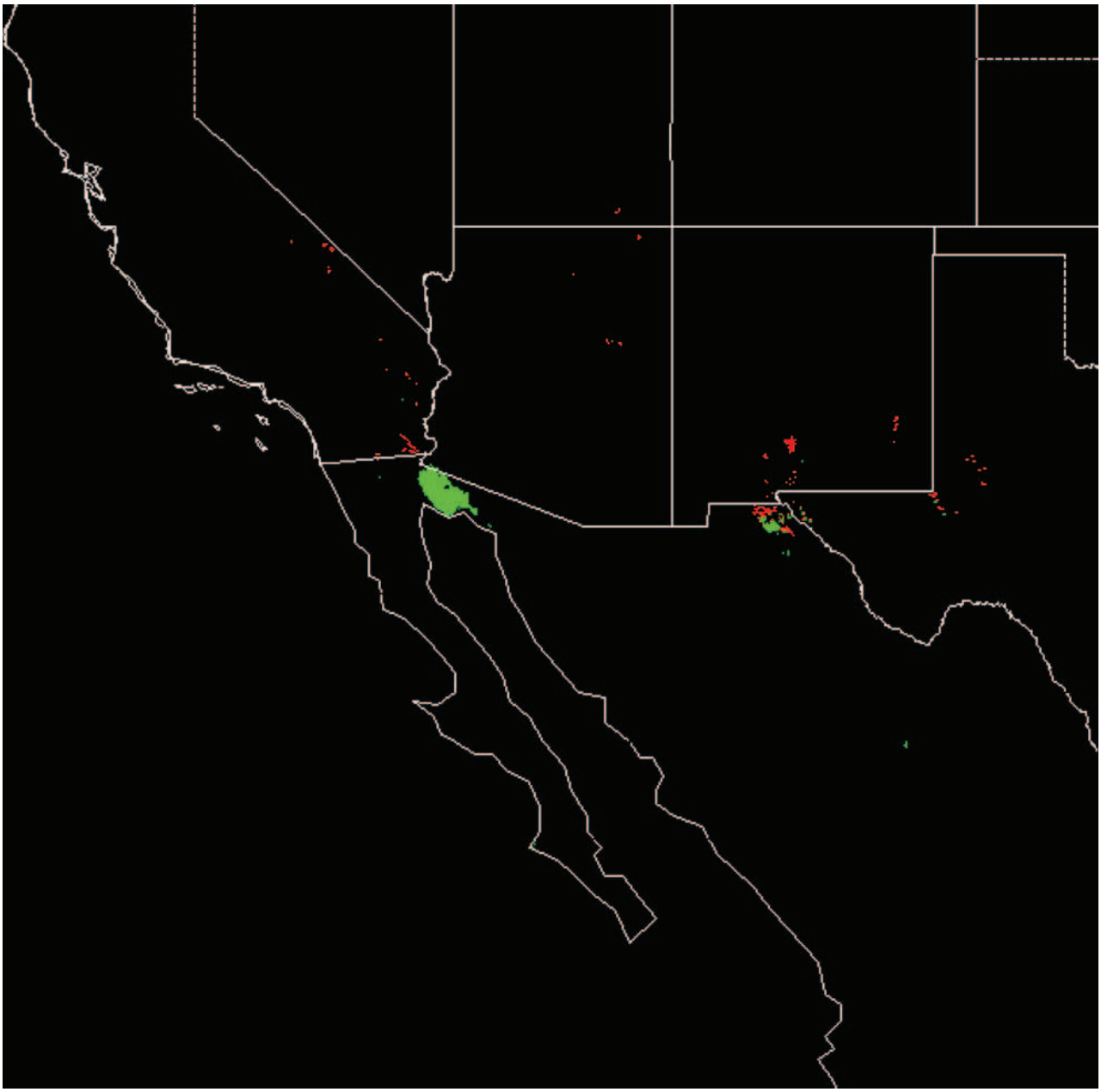

Figure 417. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2006-069 at 18:25 UTC.

Figure 418. Result of the Maximum a posteriori model (MAP) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

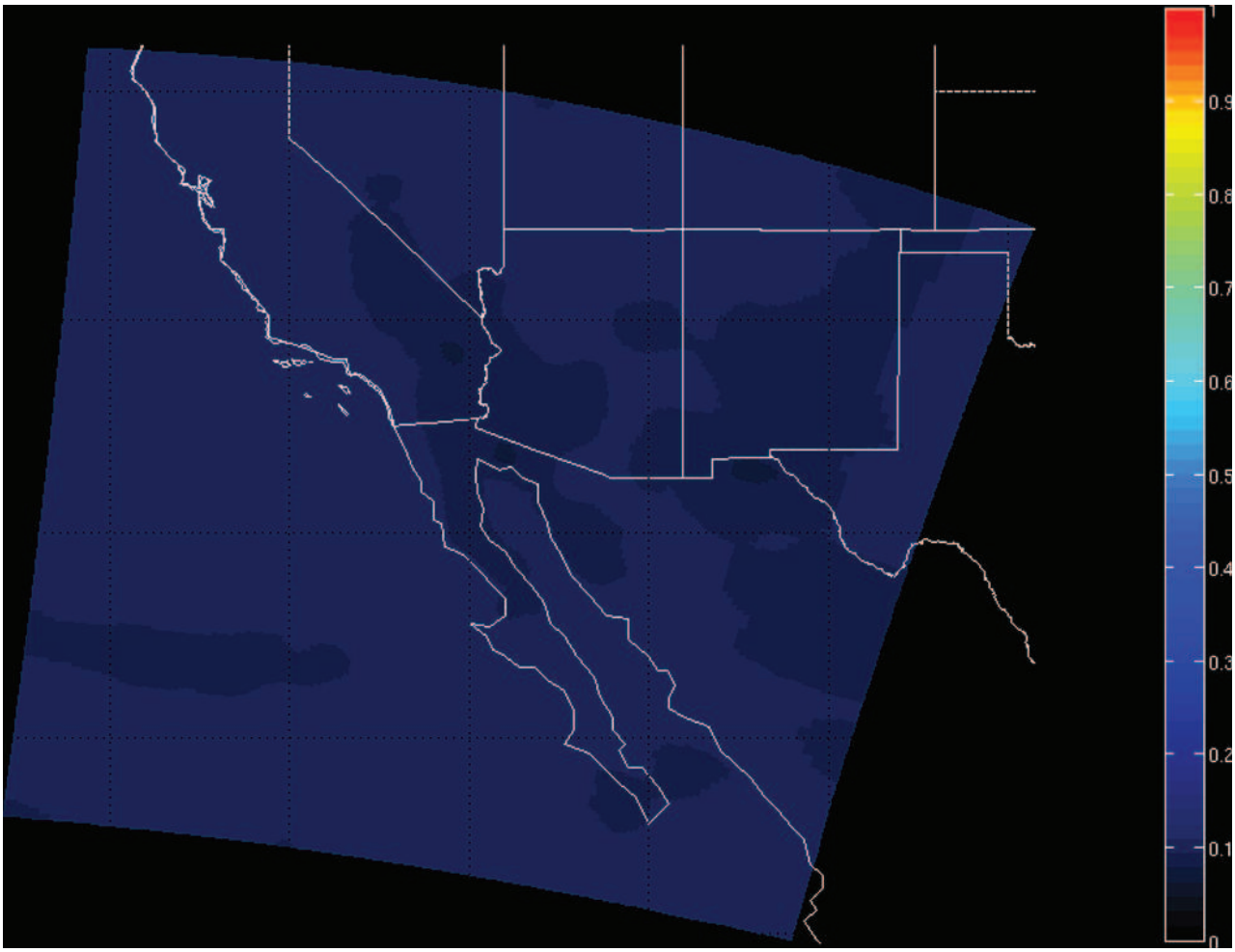

Figure 419. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

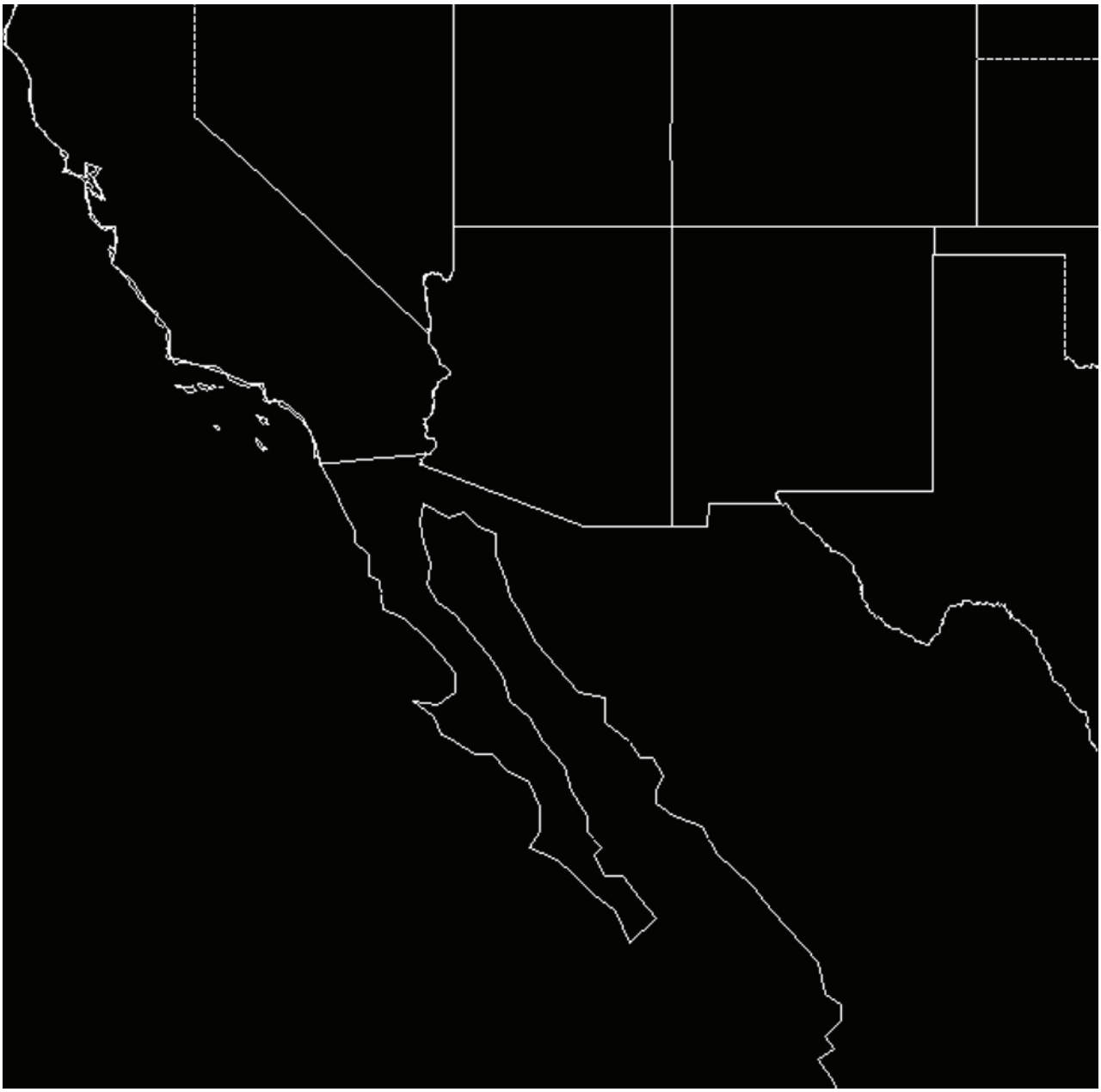

Figure 420. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2006-069 at 18:25 UTC.

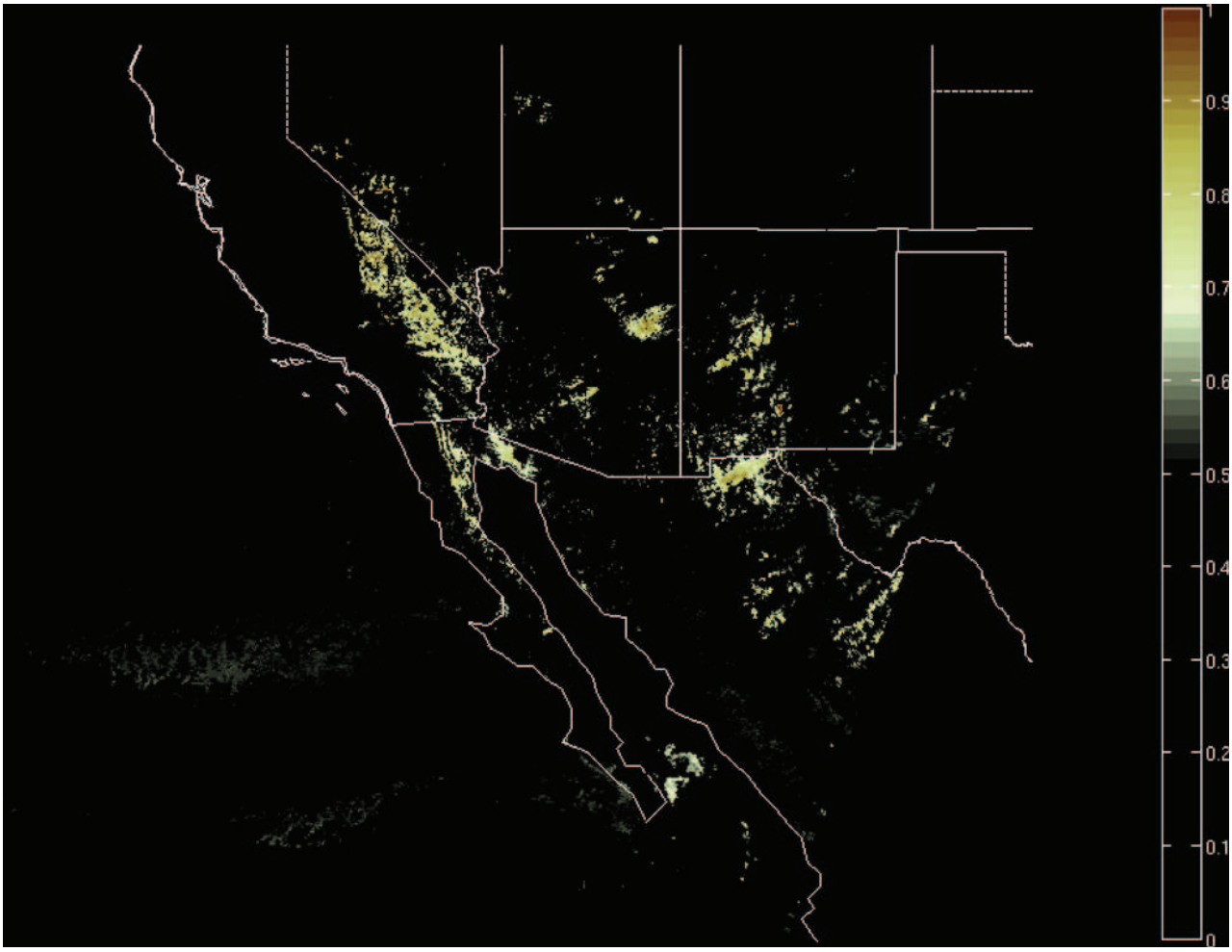

Figure 421. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

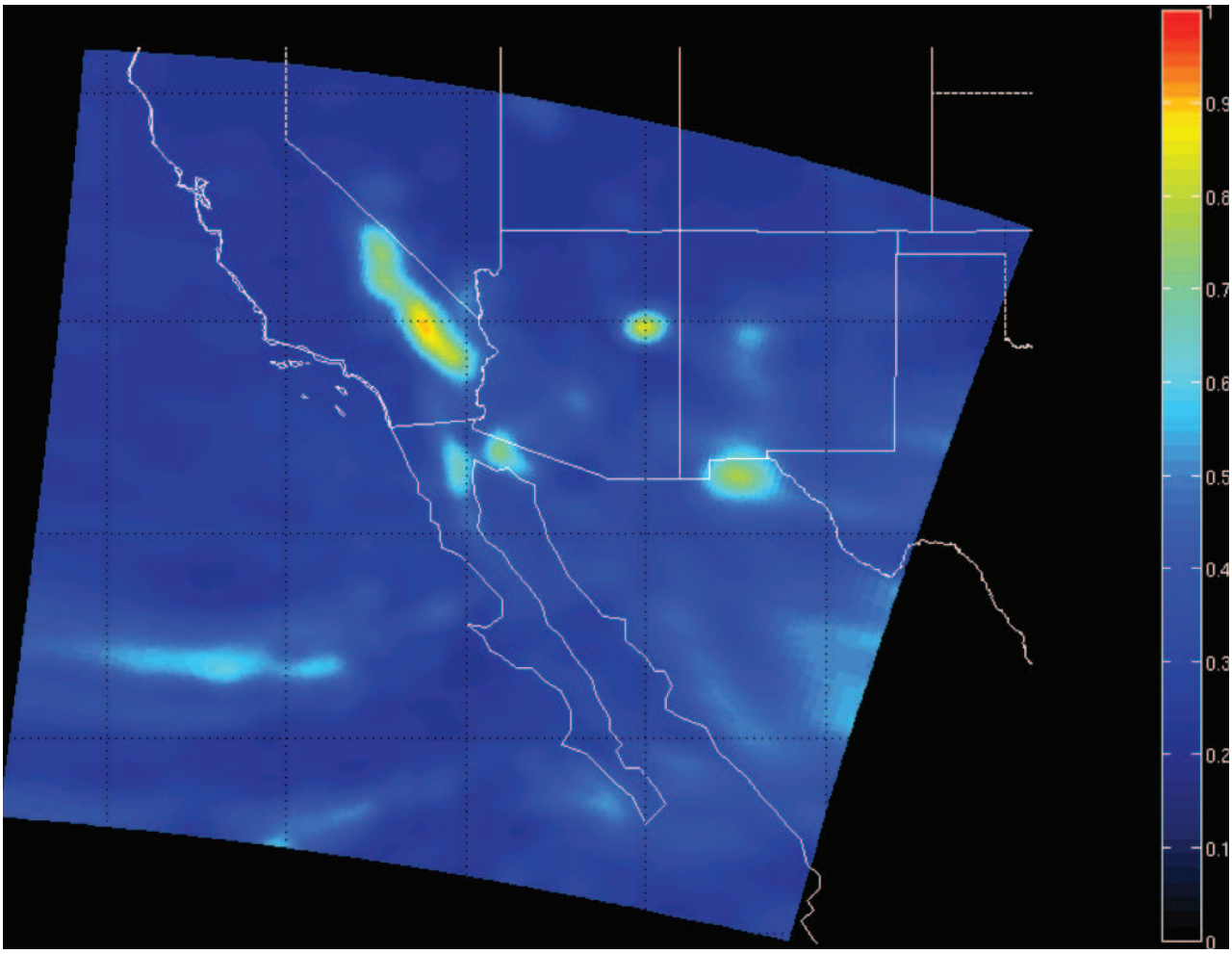

Figure 422. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

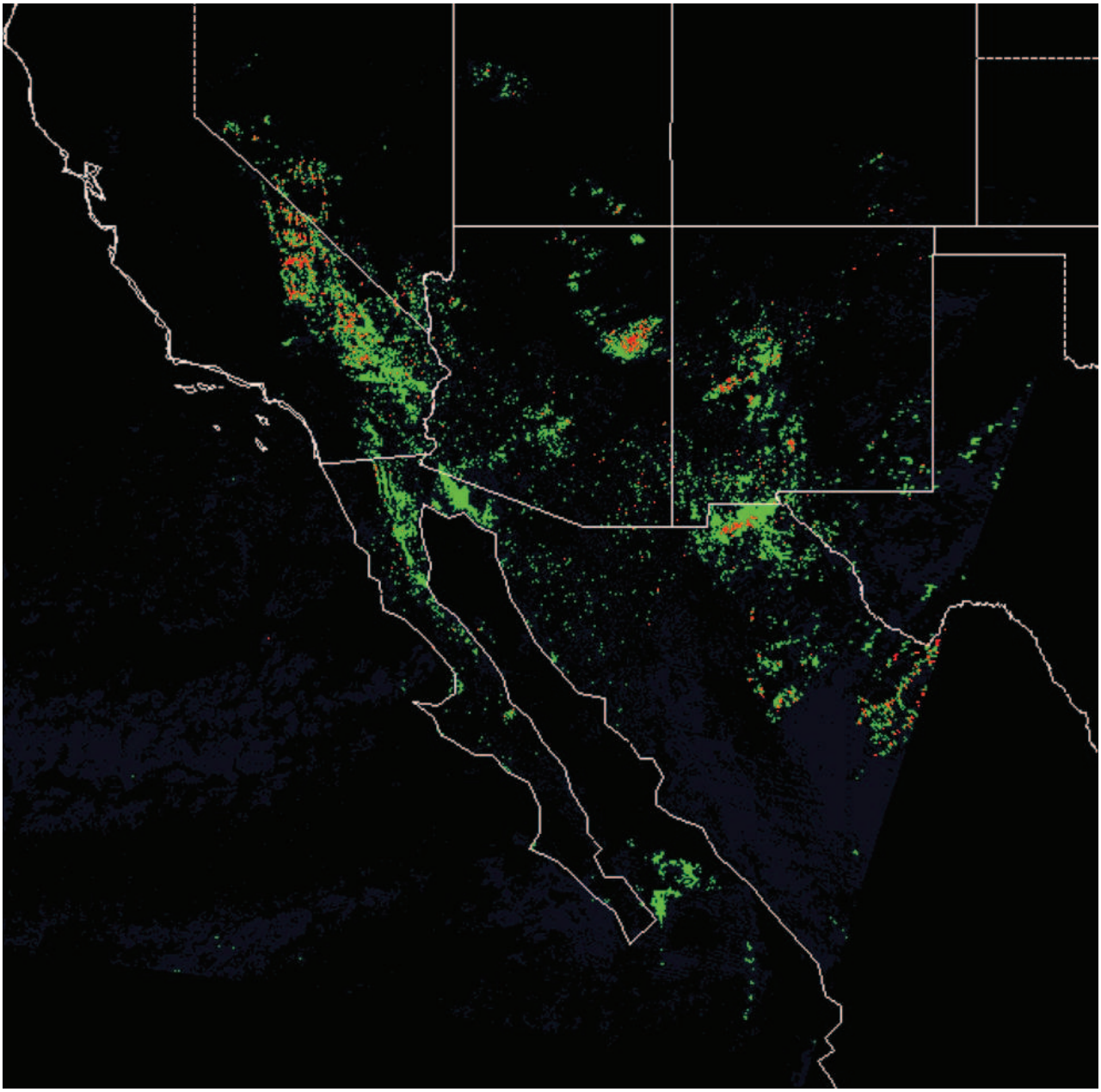

Figure 423. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-069 at 18:25 UTC.

Figure 424. Result of the Probabilistic Neural Network (PNN) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

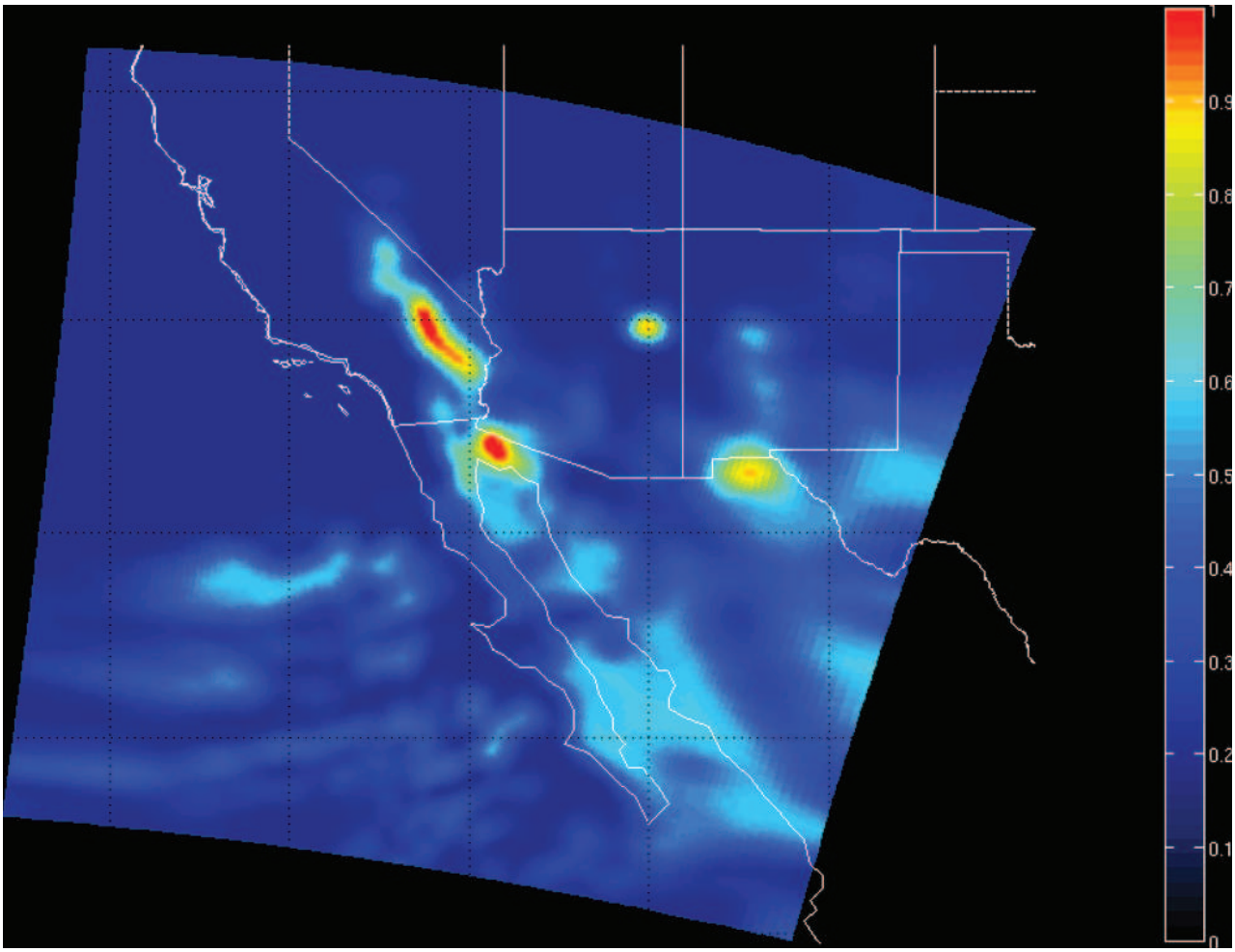

Figure 425. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2006-069 at 18:25 UTC. The colorbar shows the index of probability.

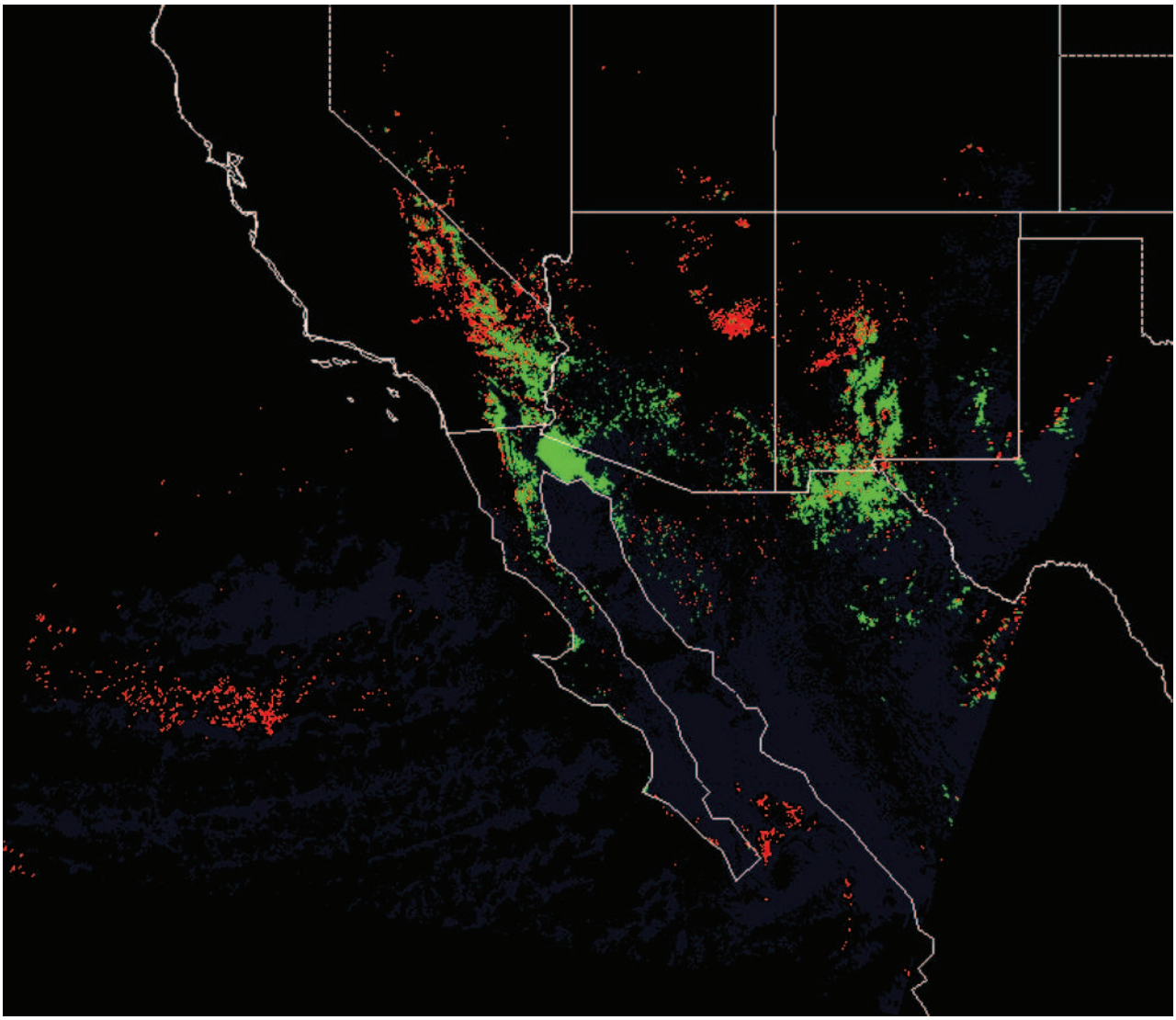

Figure 426. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2006-069 at 18:25 UTC.

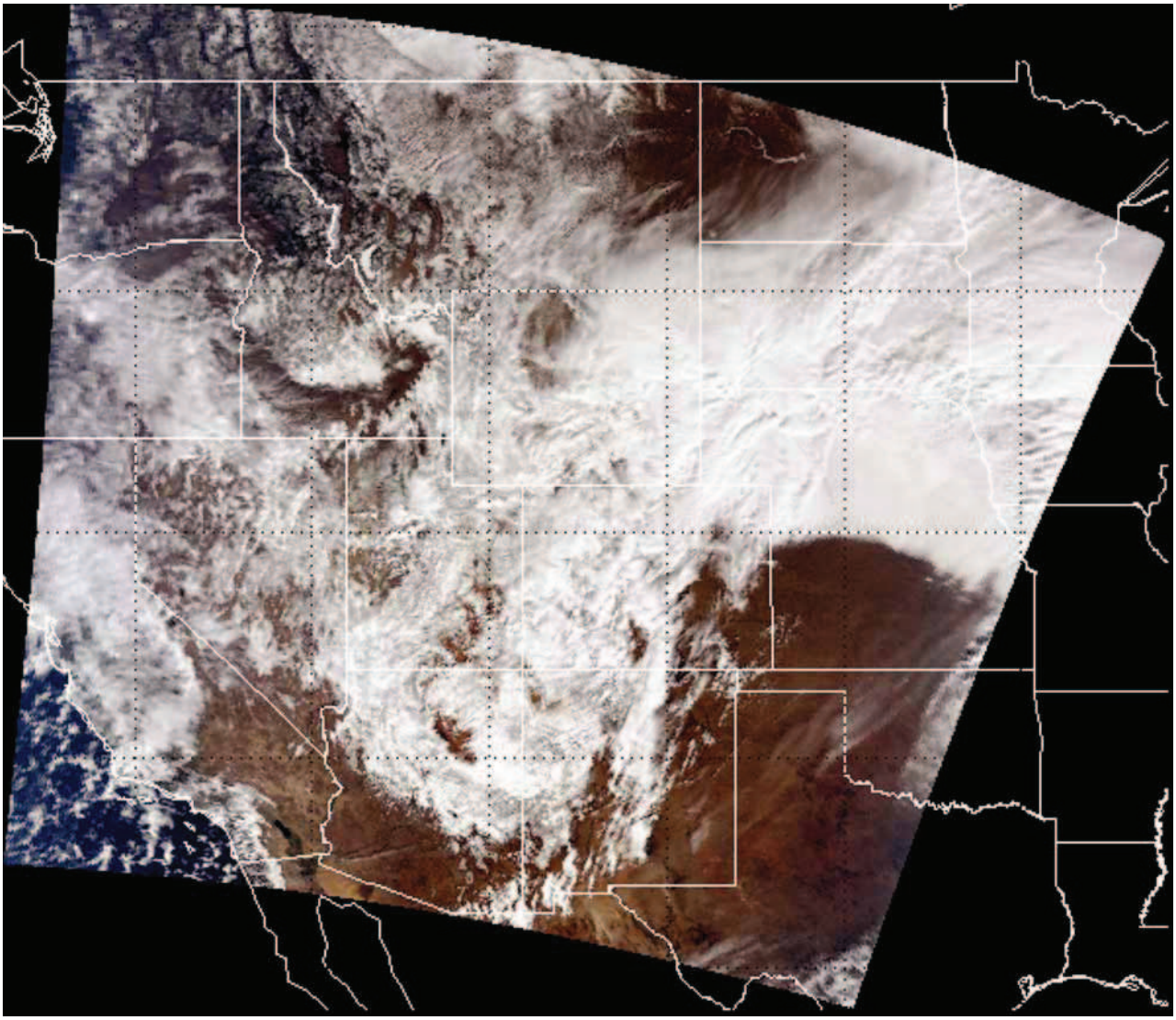

Figure 427. True color image for the event of 2006-071 at 18:10 UTC.

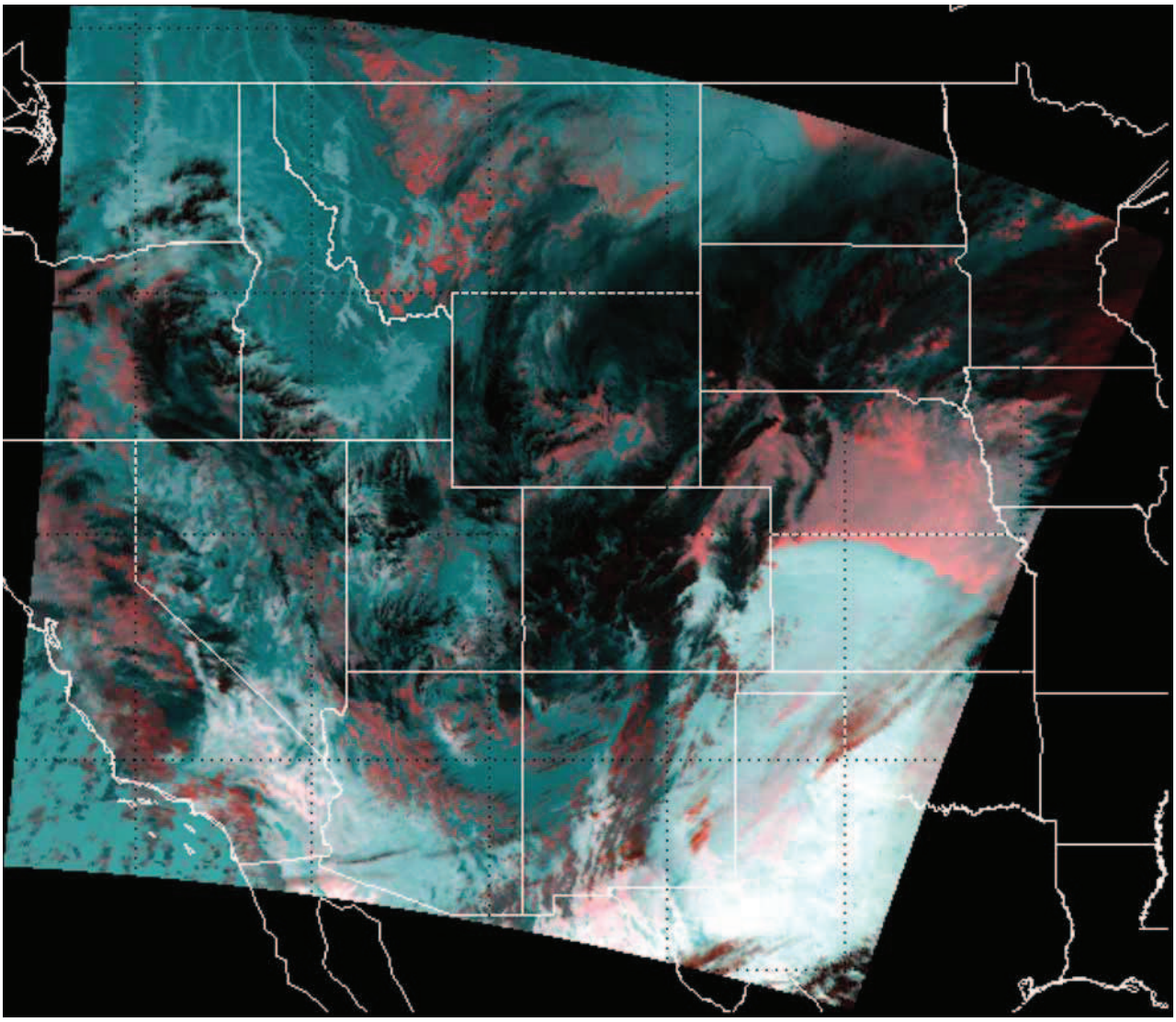

Figure 428. False color image for the event of 2006-071 at 18:10 UTC.

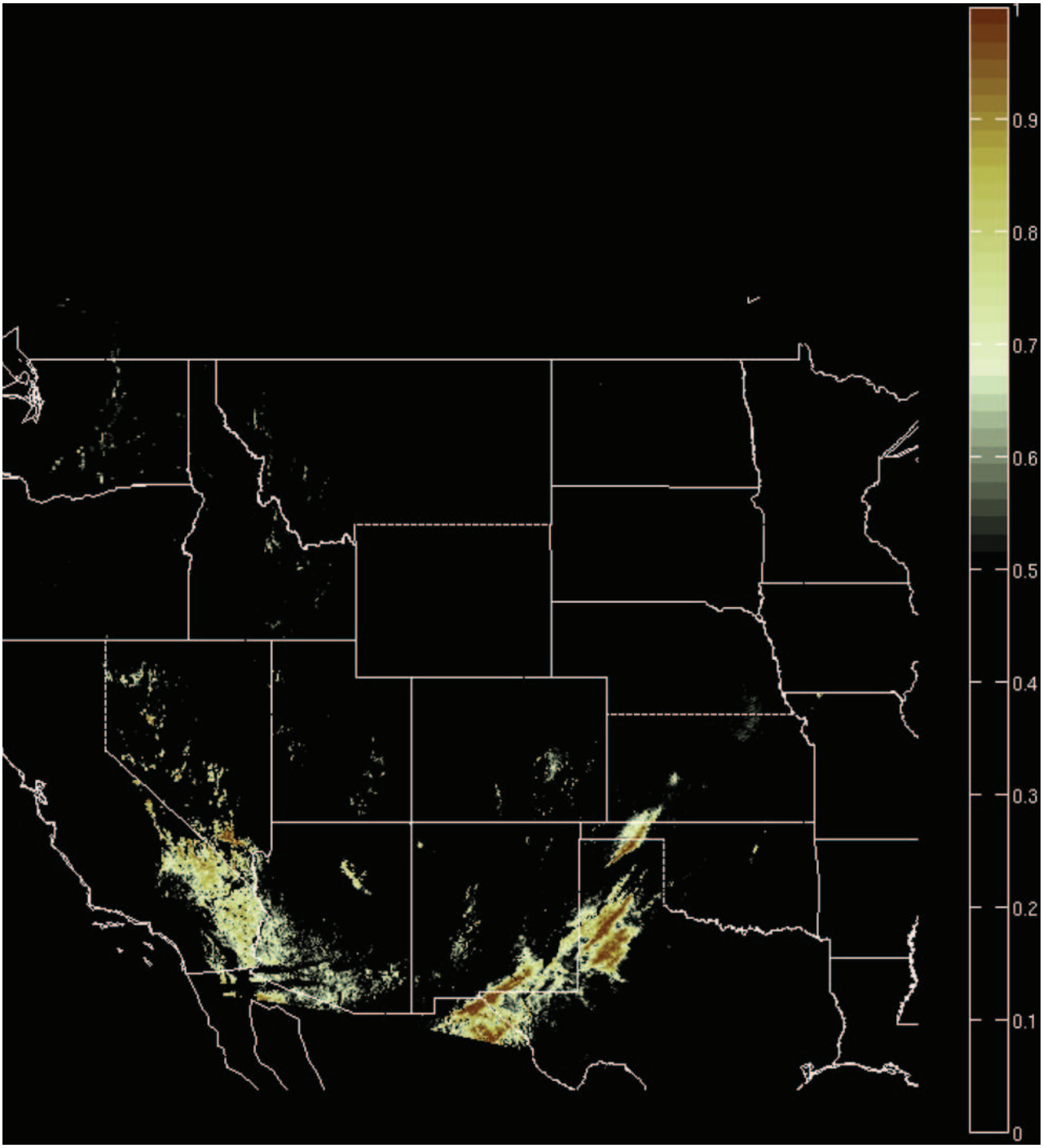

Figure 429. Result of the probabilistic display model (PD) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

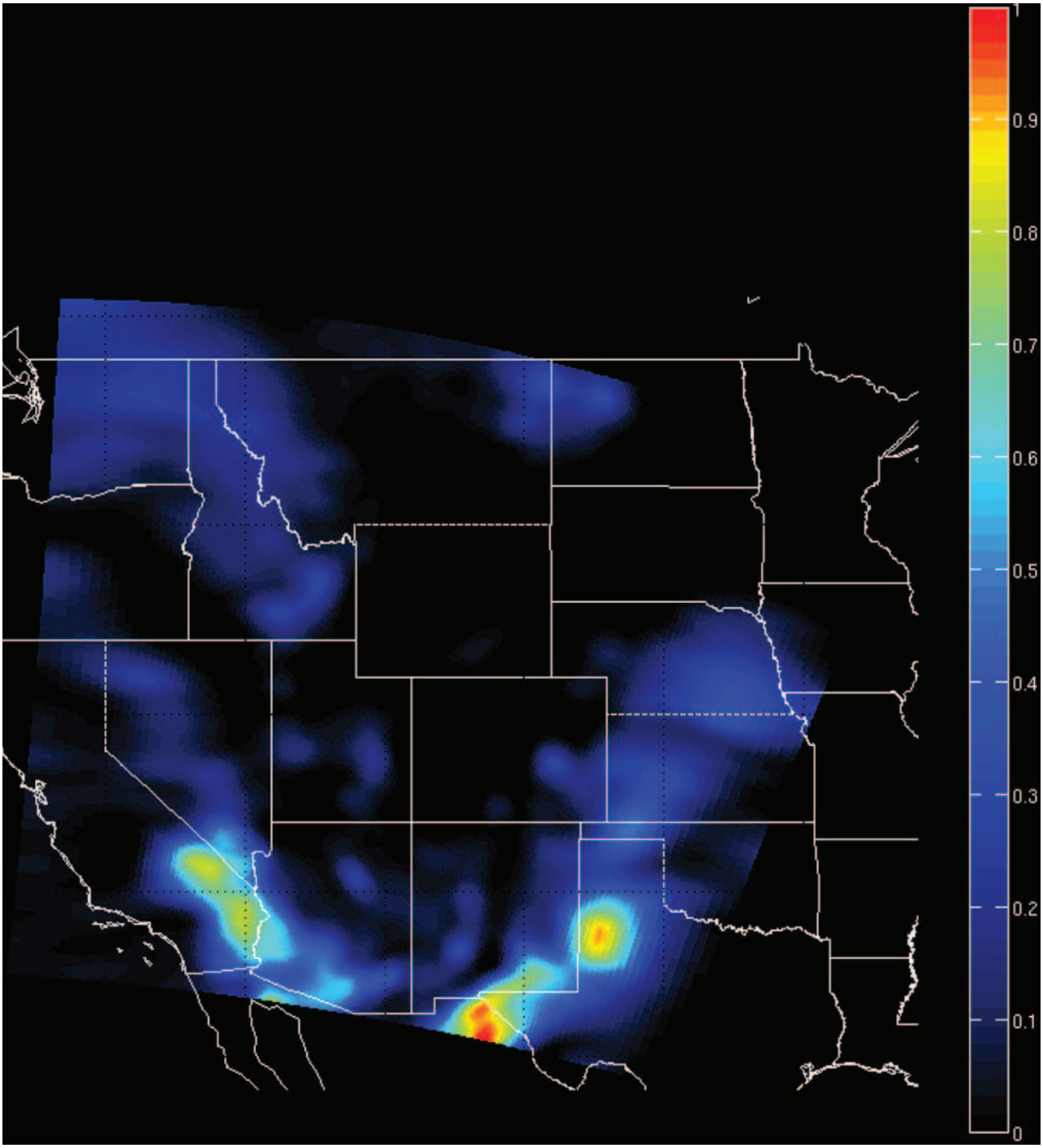

Figure 430. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

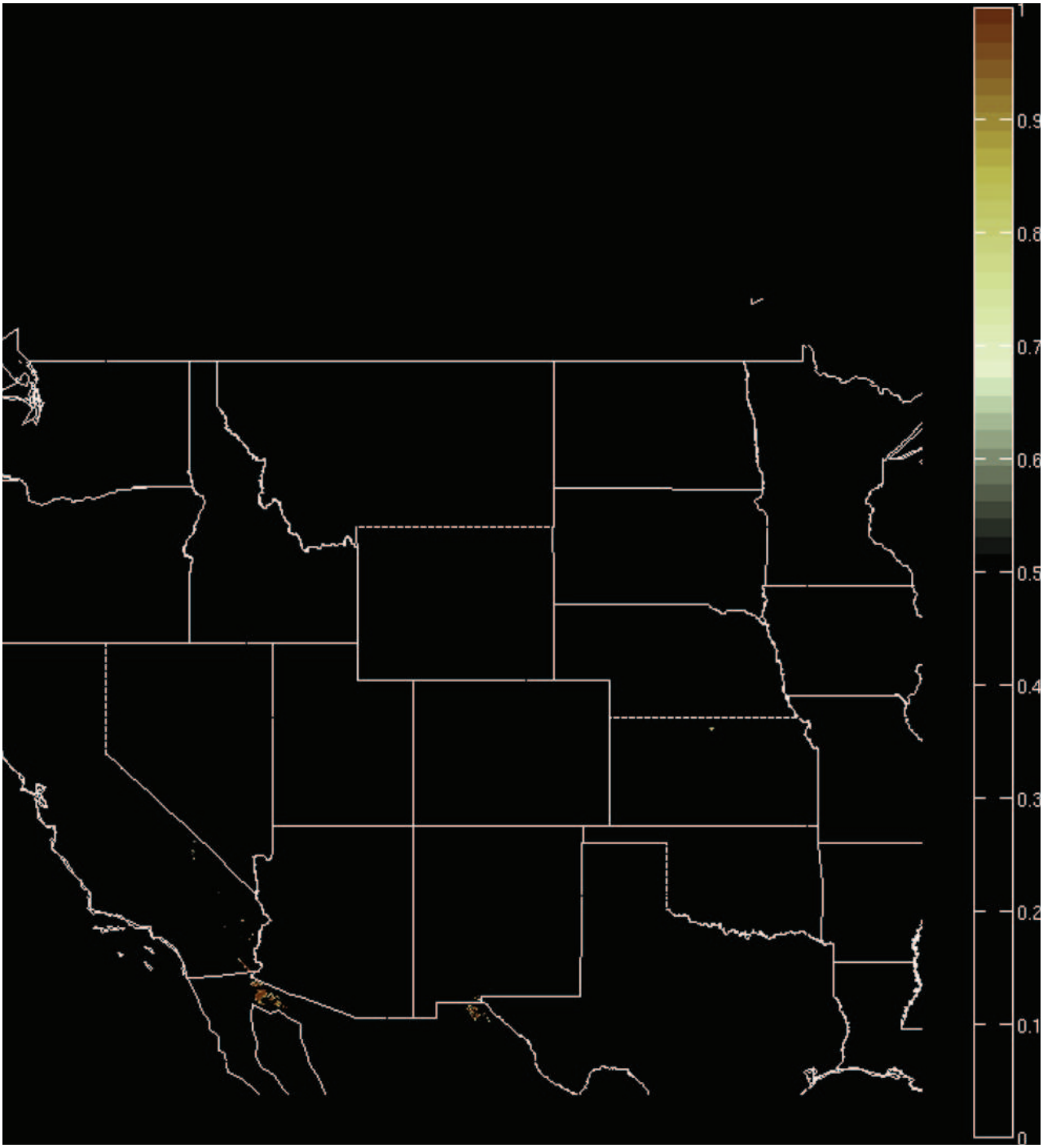

Figure 431. Result of the Maximum Likelihood model (MLE) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

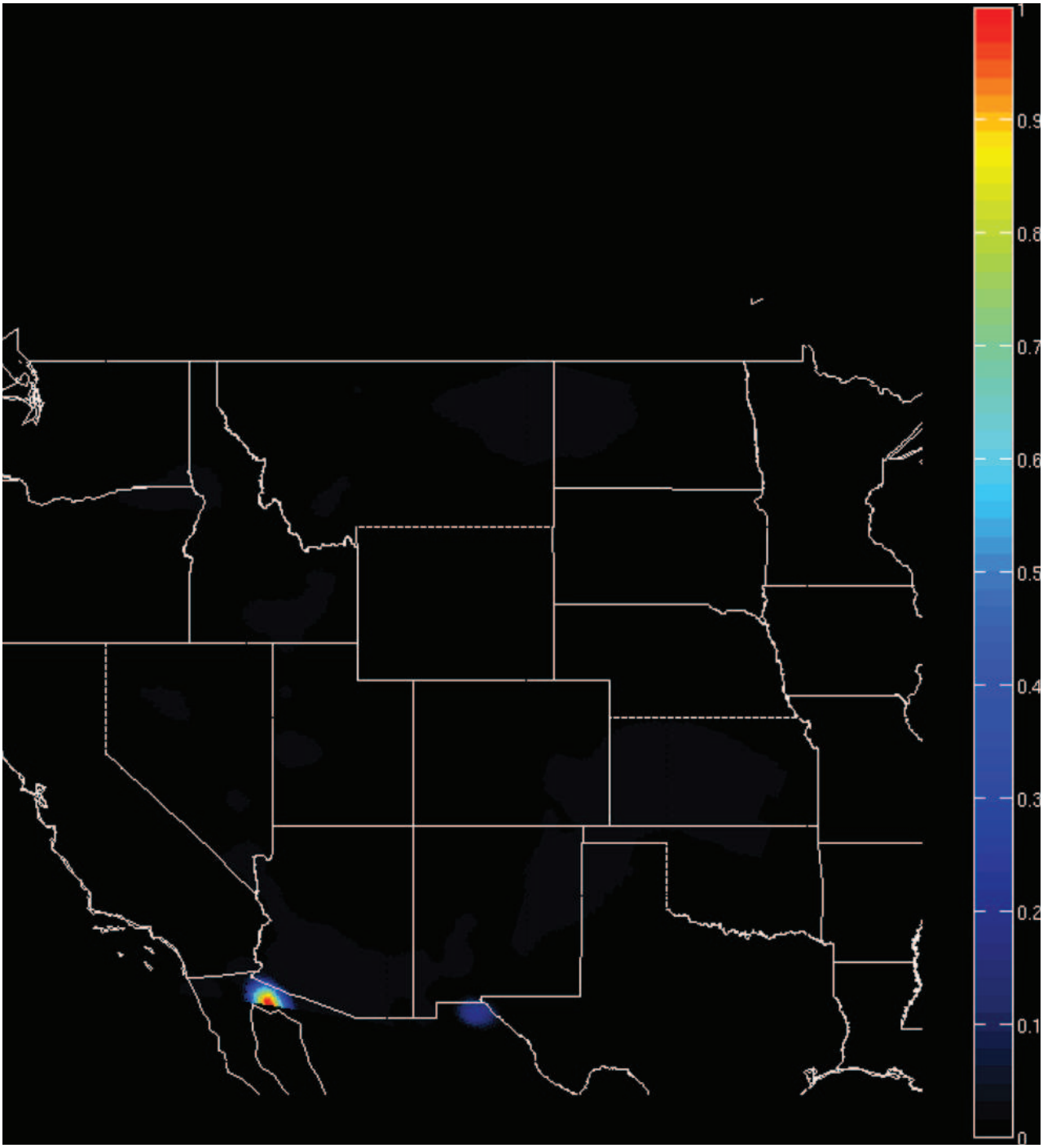

Figure 432. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

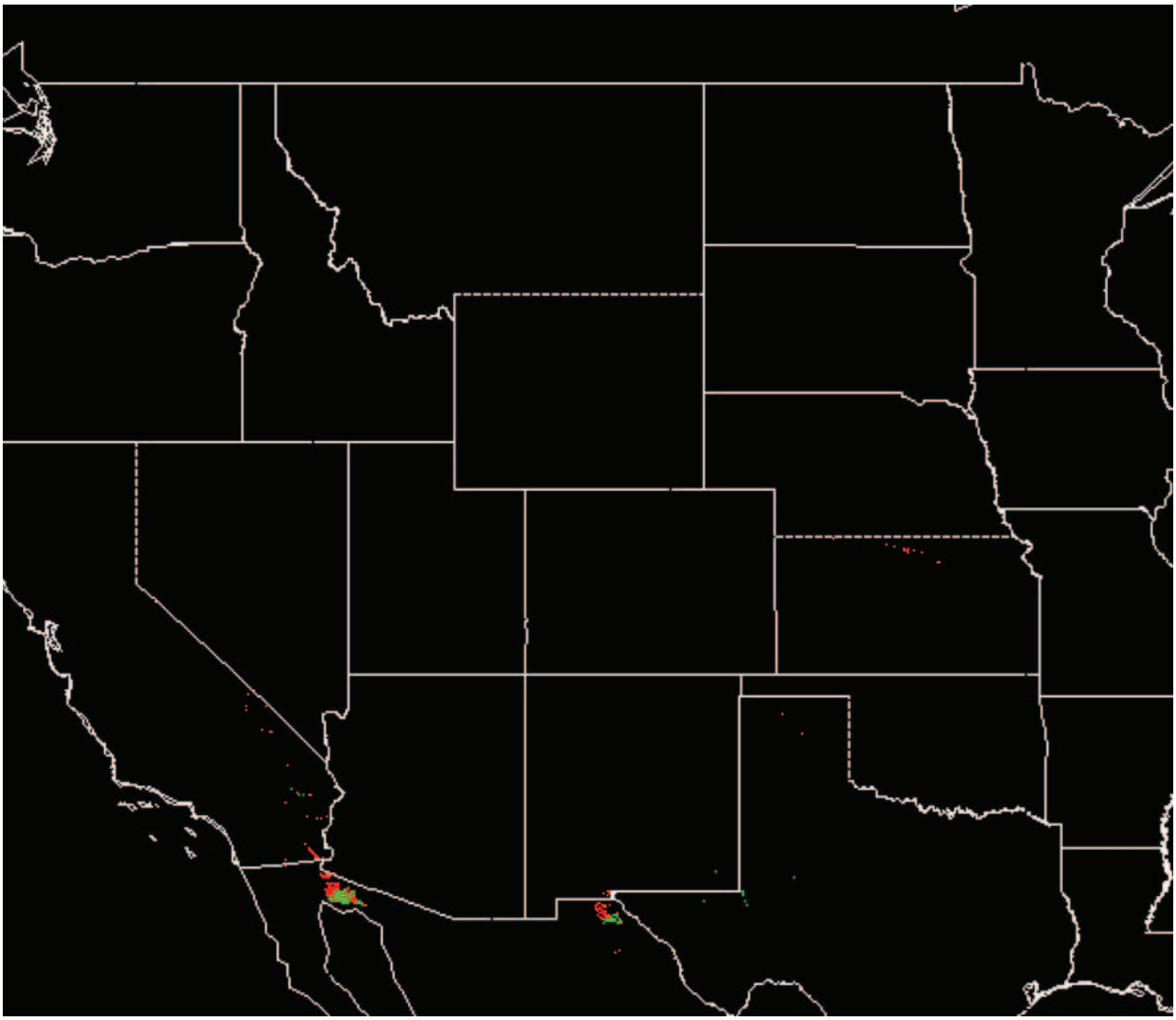

Figure 433. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2006-071 at 18:10 UTC.

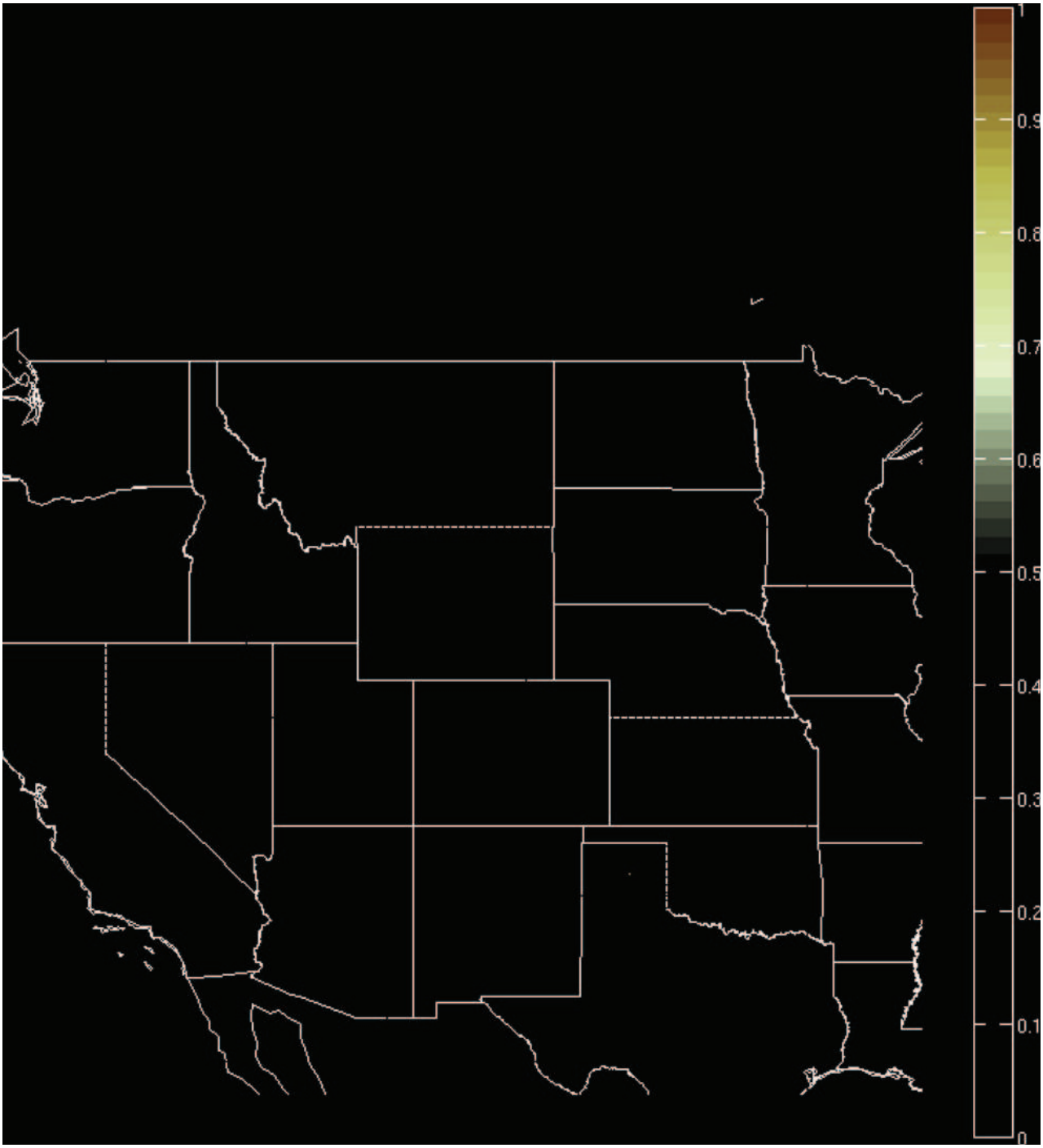

Figure 434. Result of the Maximum a posteriori model (MAP) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

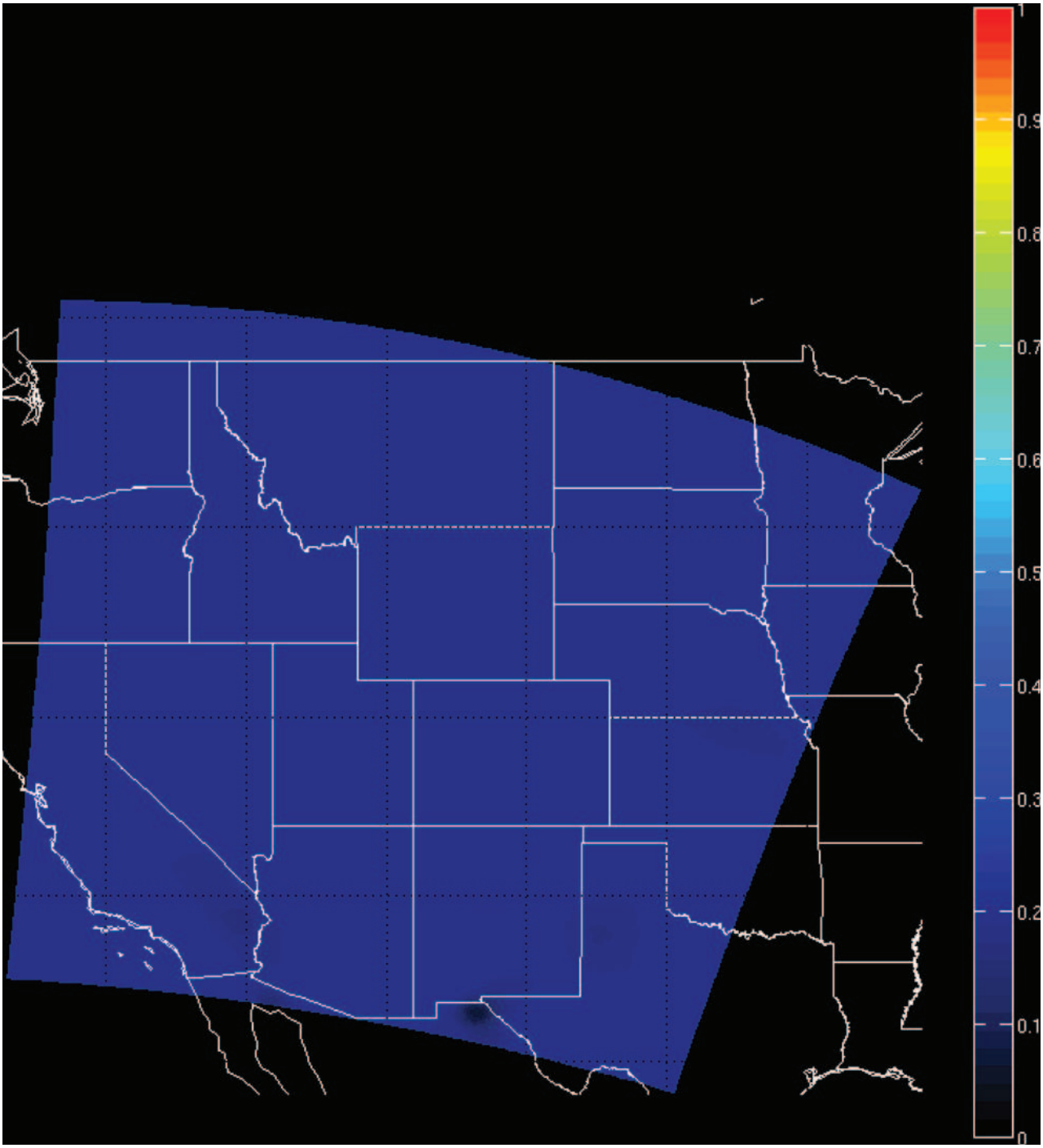

Figure 435. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

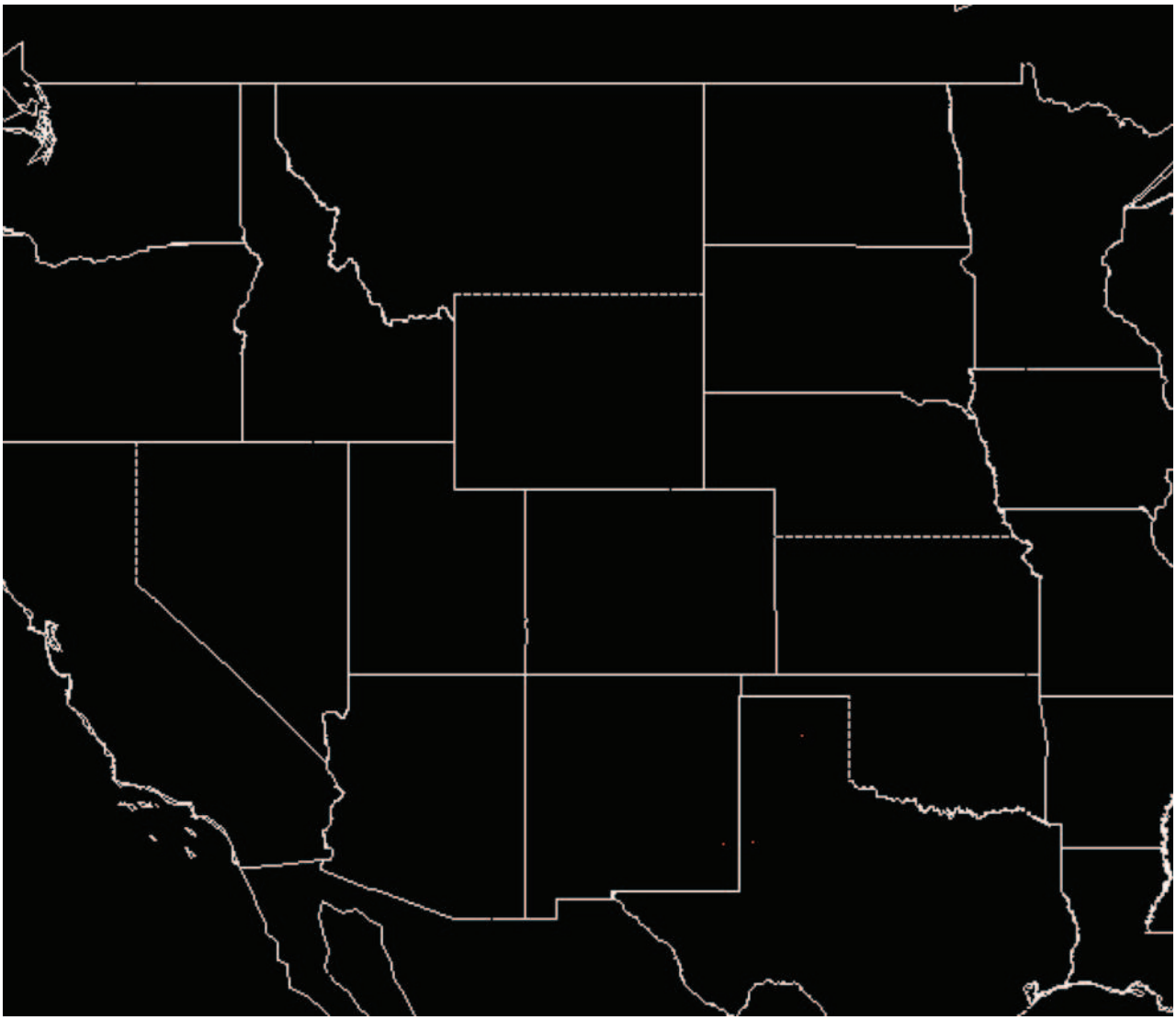

Figure 436. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2006-071 at 18:10 UTC.

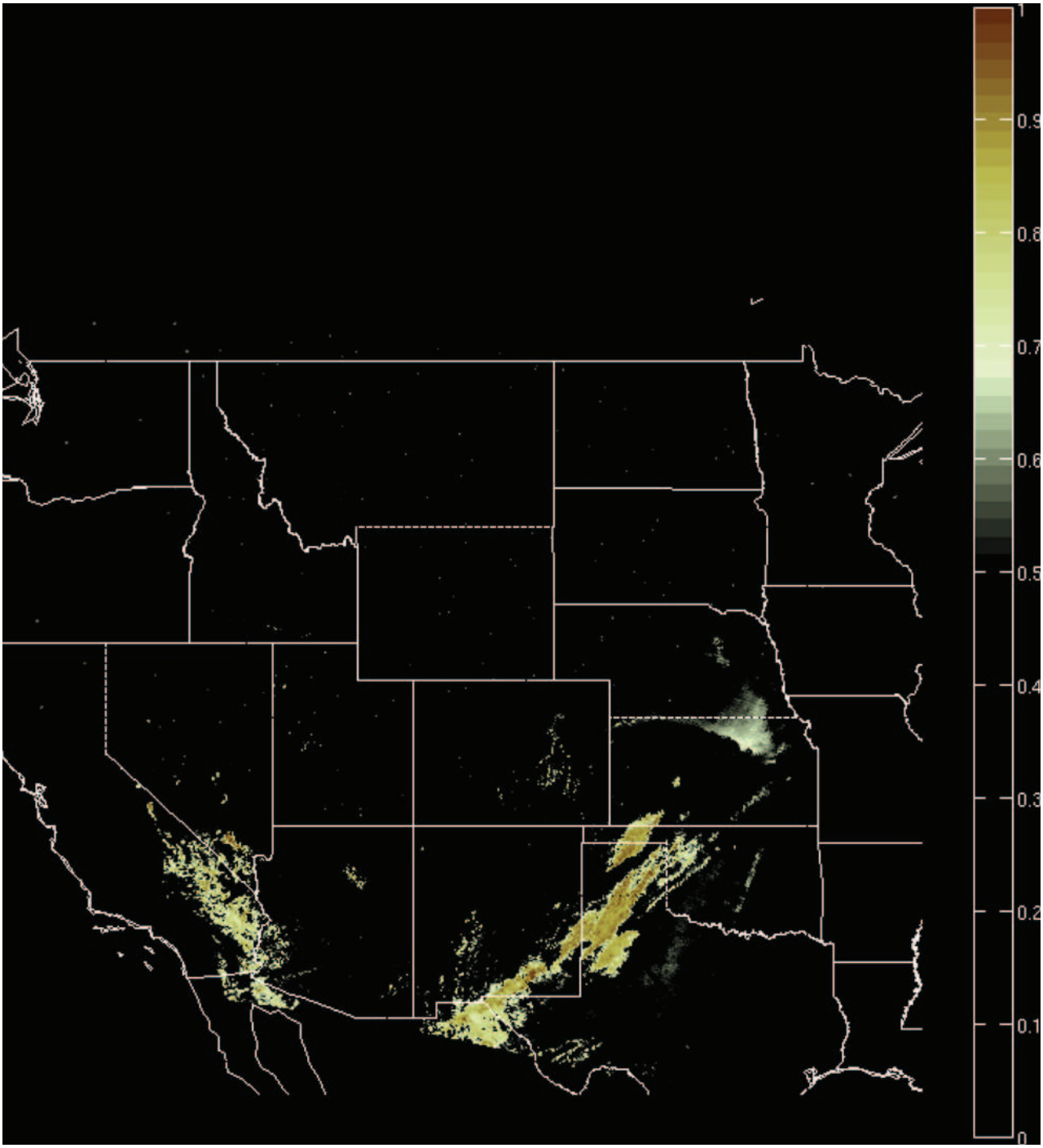

Figure 437. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

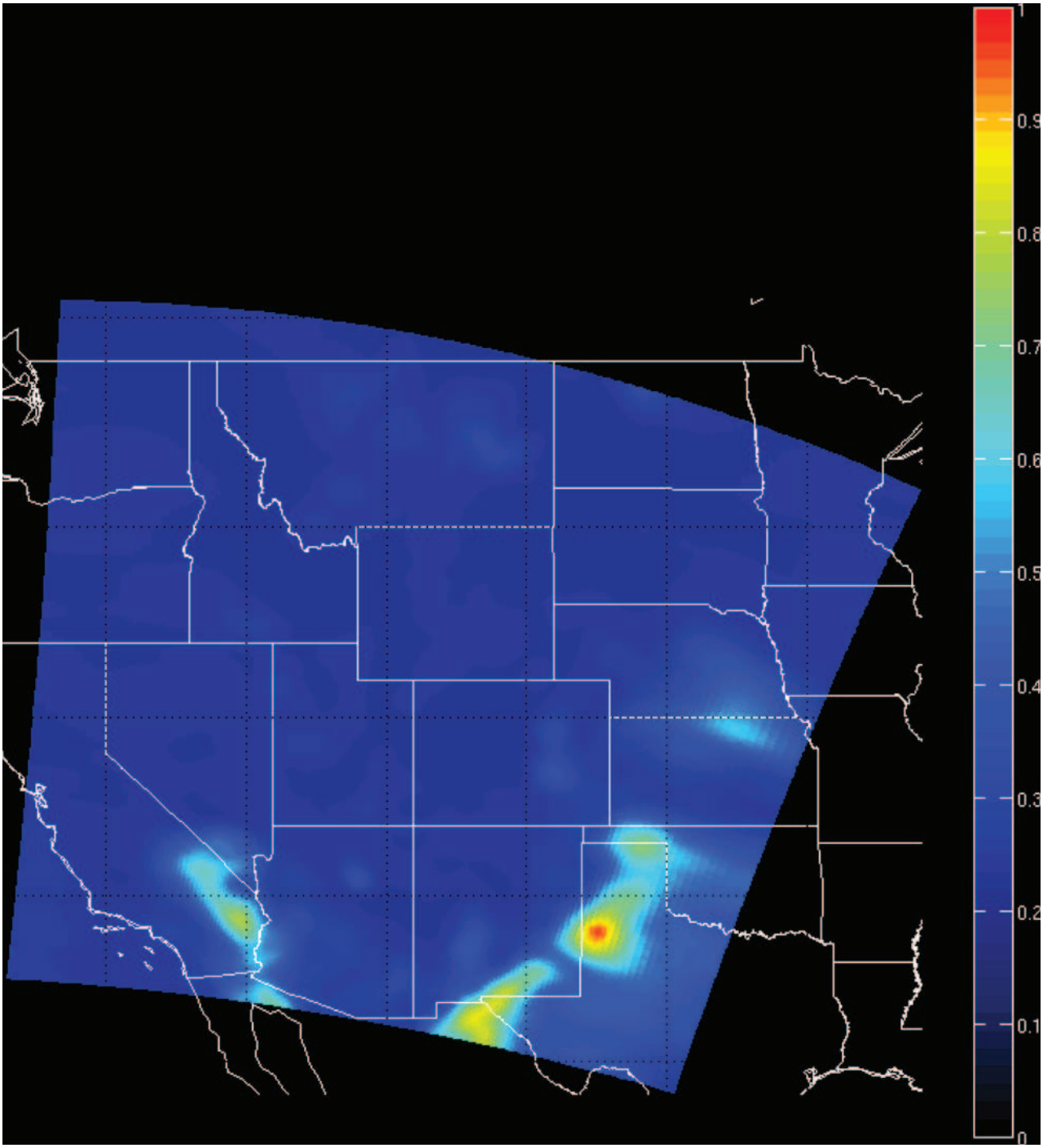

Figure 438. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

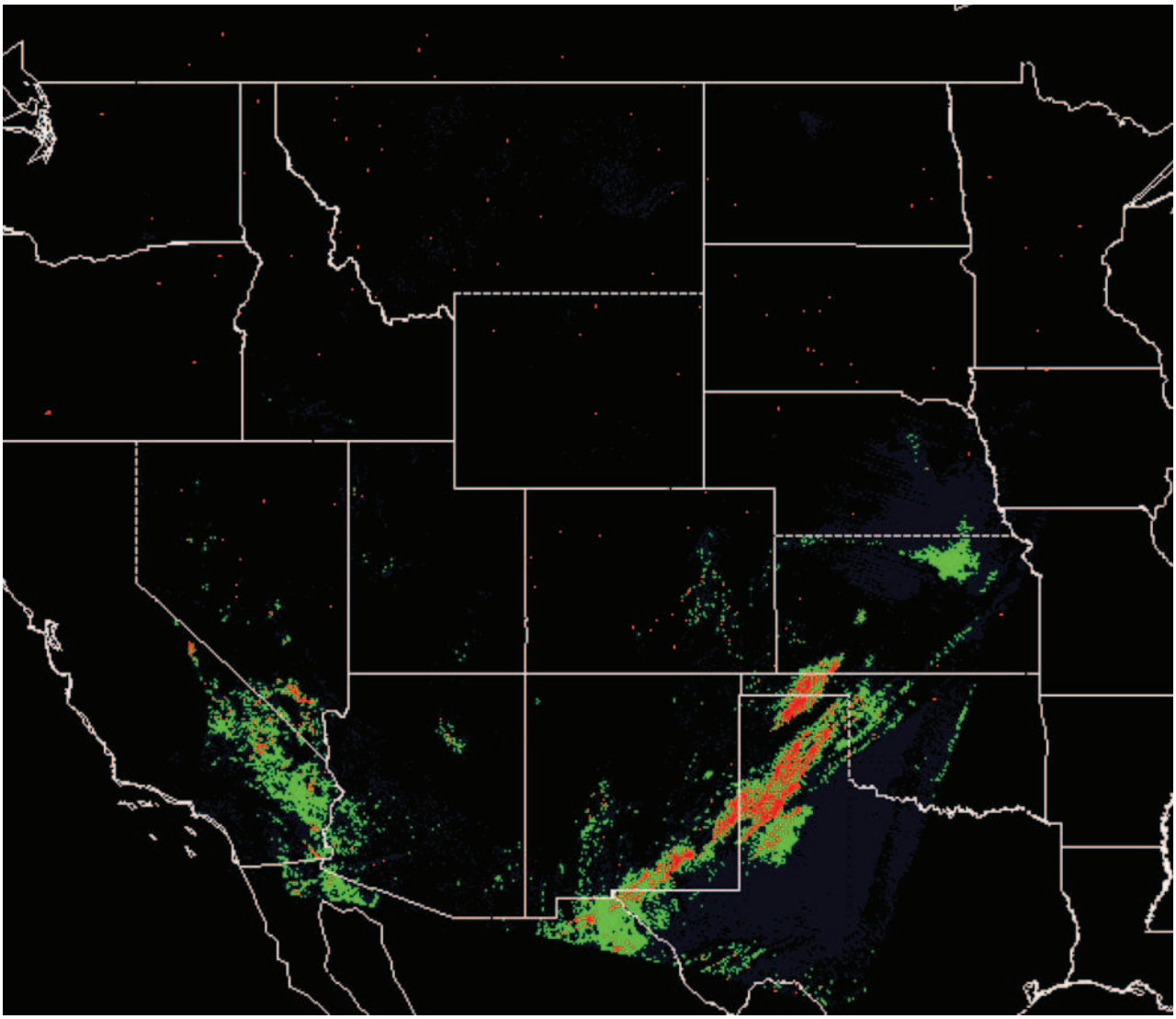

Figure 439. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-071 at 18:10 UTC.

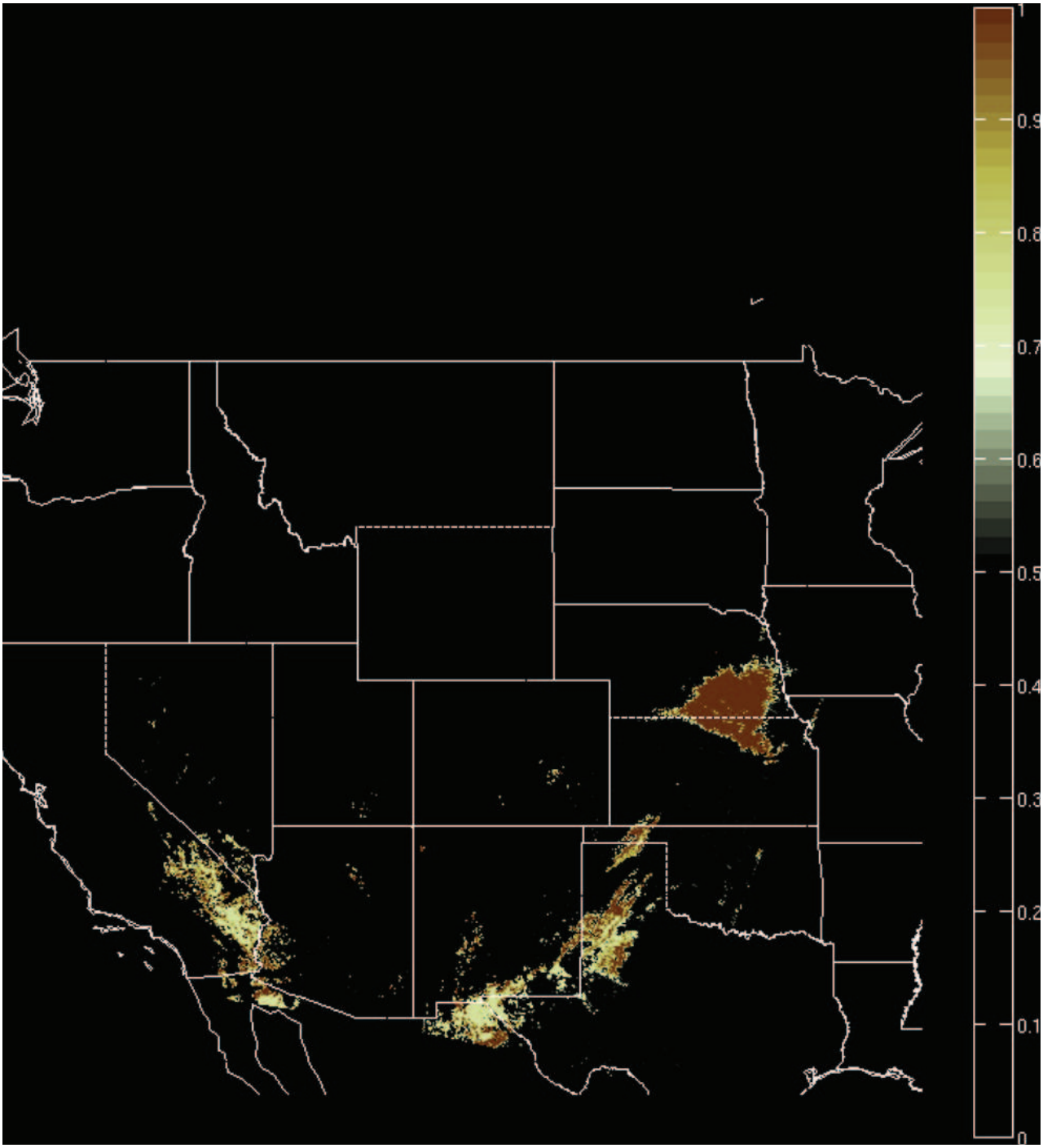

Figure 440. Result of the Probabilistic Neural Network (PNN) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

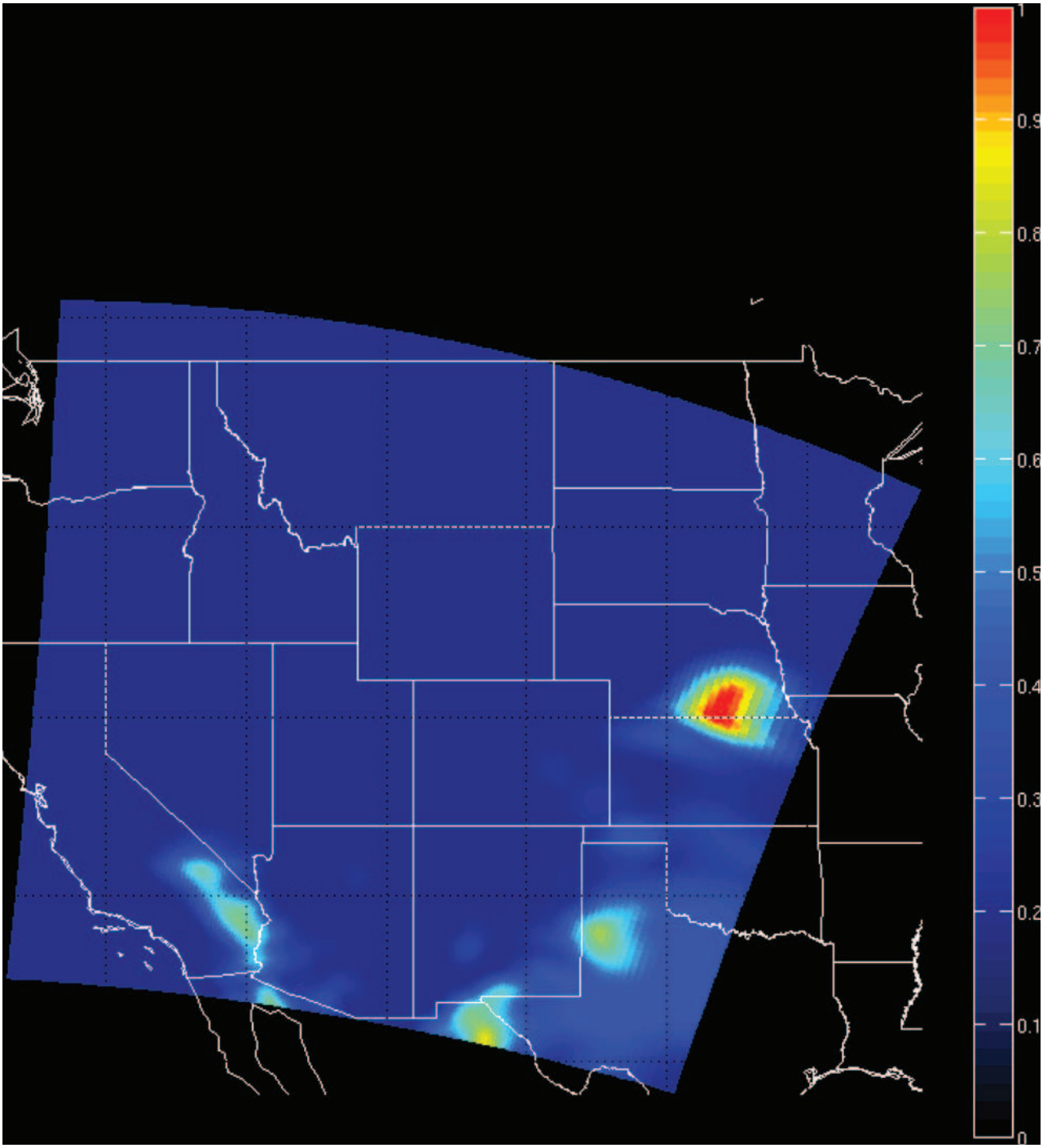

Figure 441. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2006-071 at 18:10 UTC. The colorbar shows the index of probability.

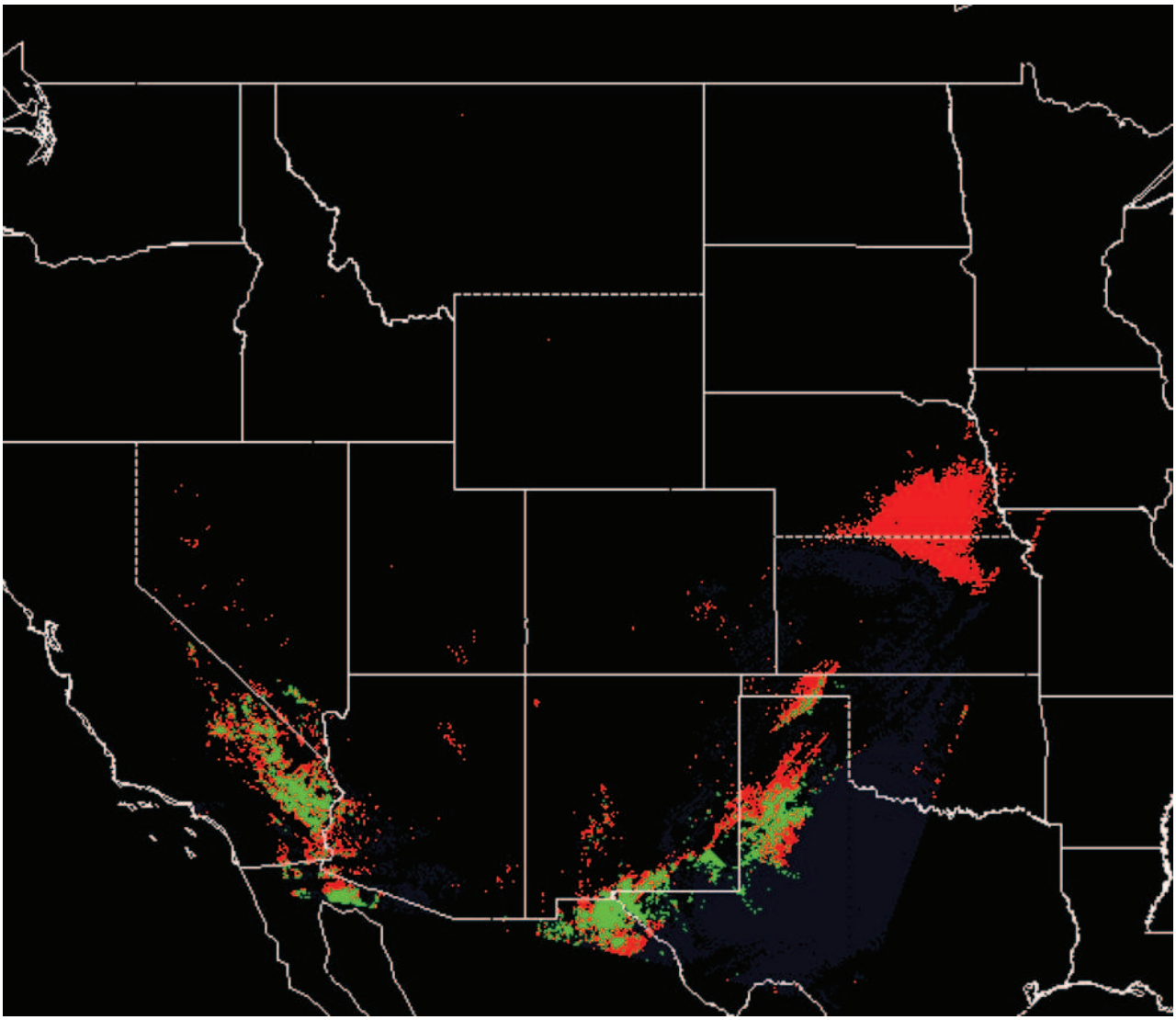

Figure 442. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2006-071 at 18:10 UTC.

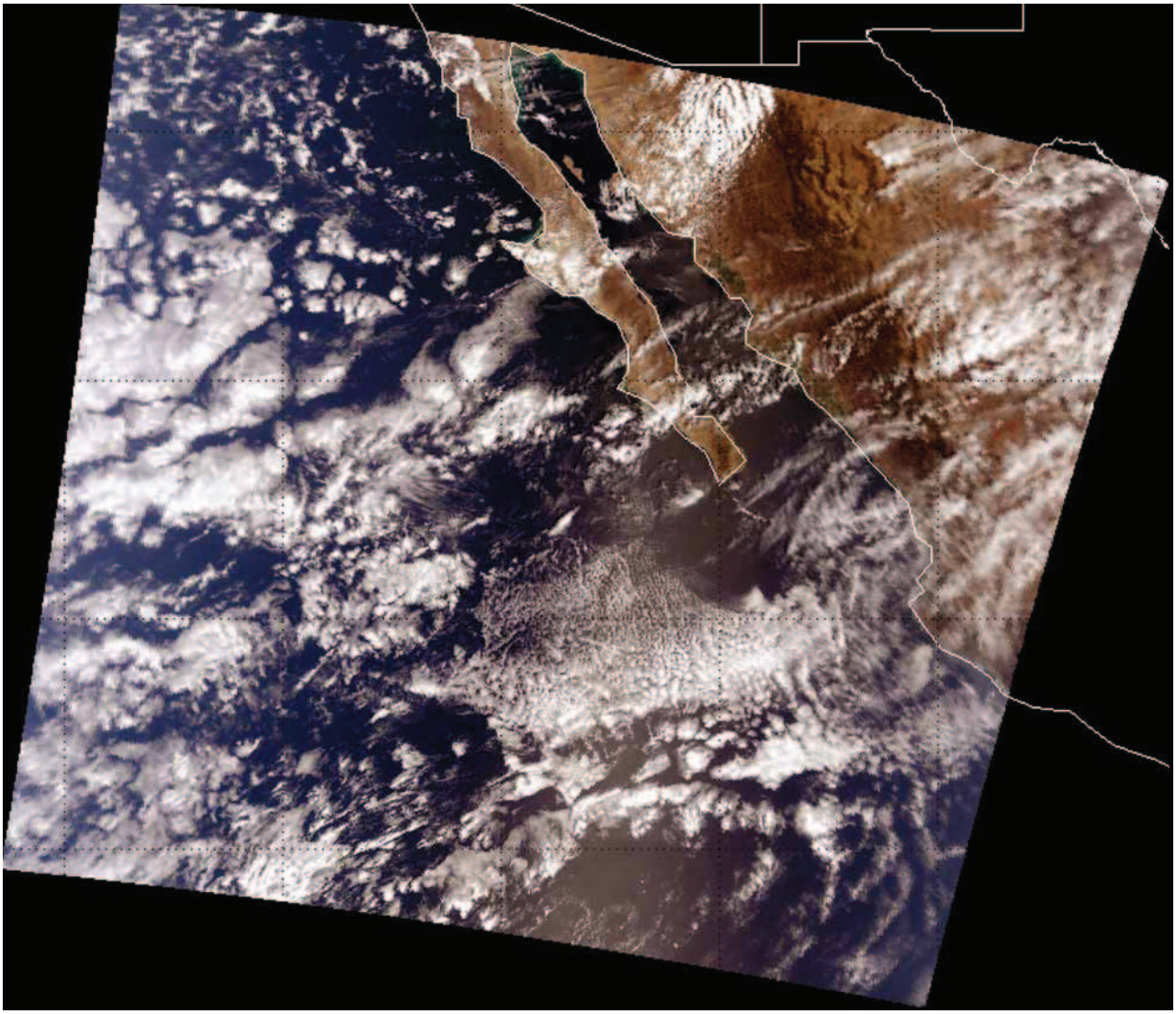

Figure 443. True color image for the event of 2006-071 at 18:15 UTC.

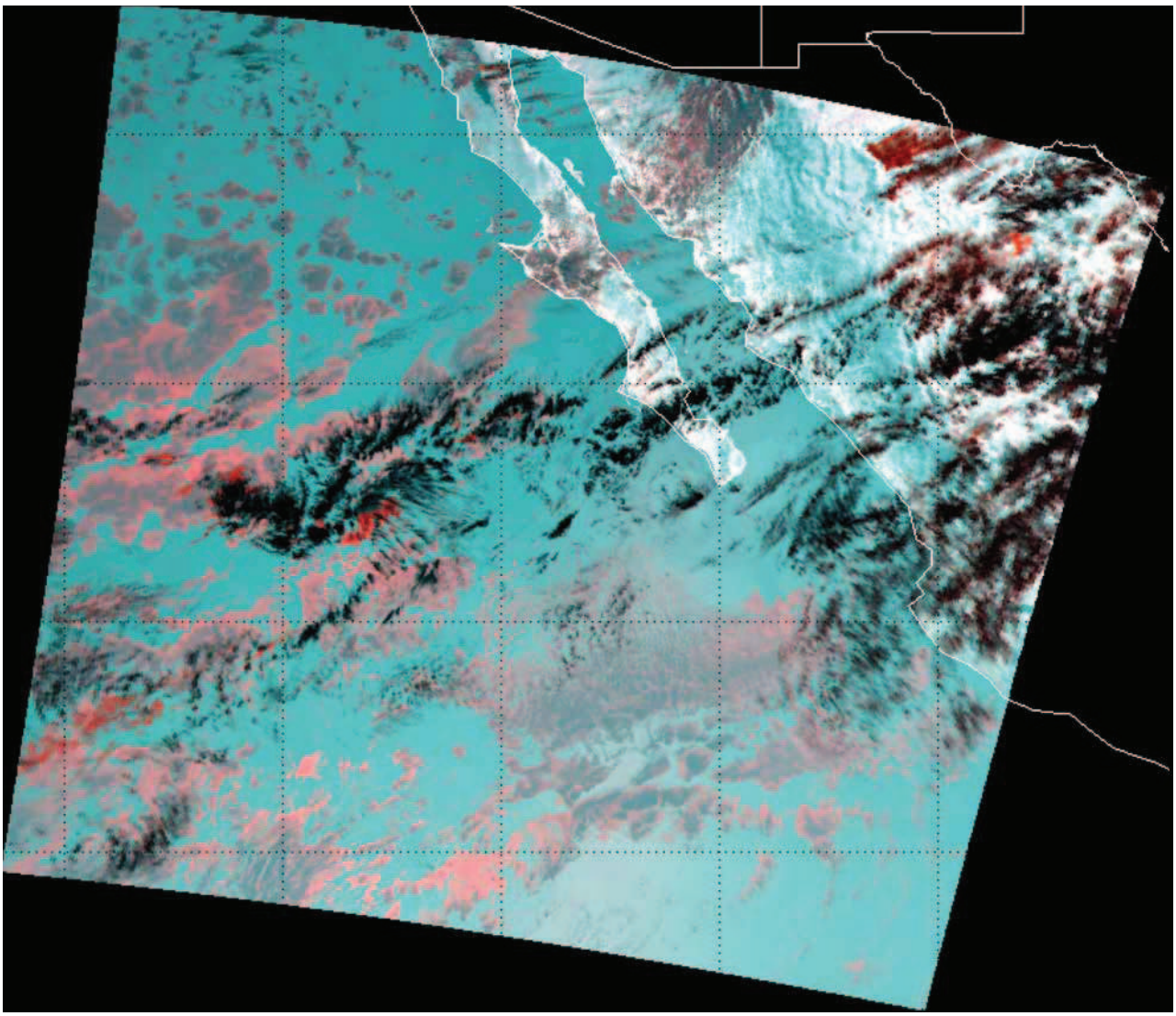

Figure 444. False color image for the event of 2006-071 at 18:15 UTC.

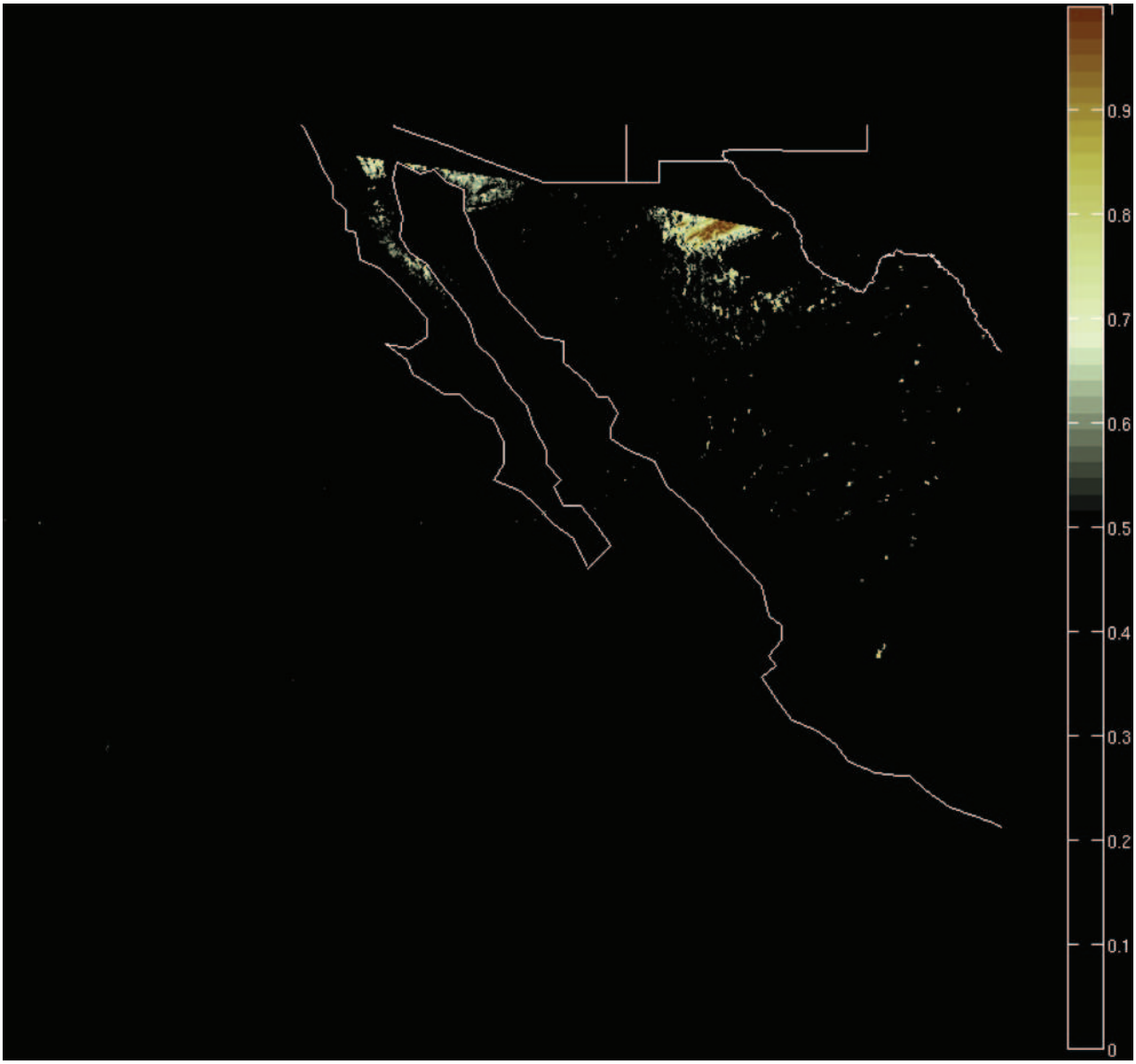

Figure 445. Result of the probabilistic display model (PD) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

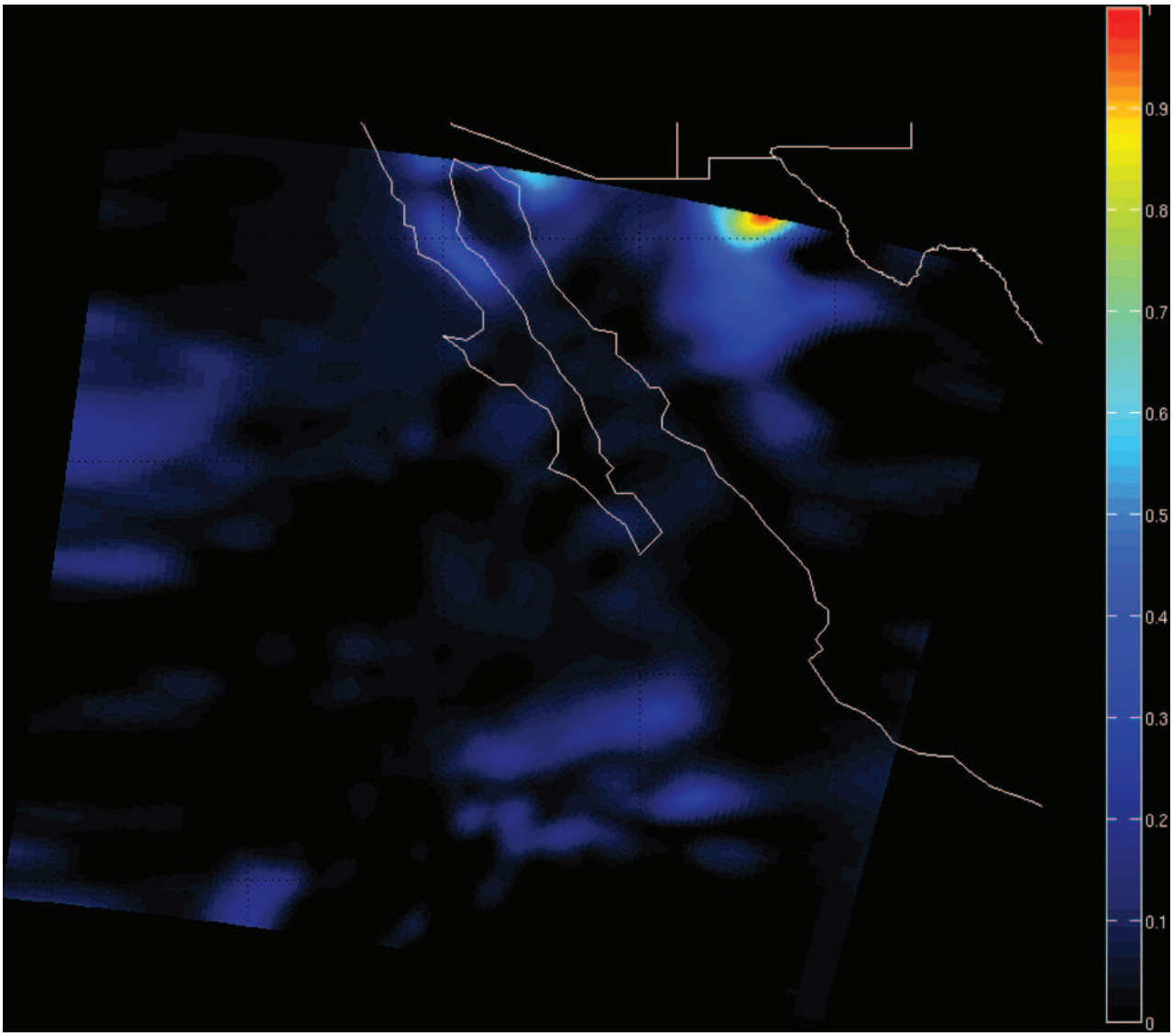

Figure 446. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

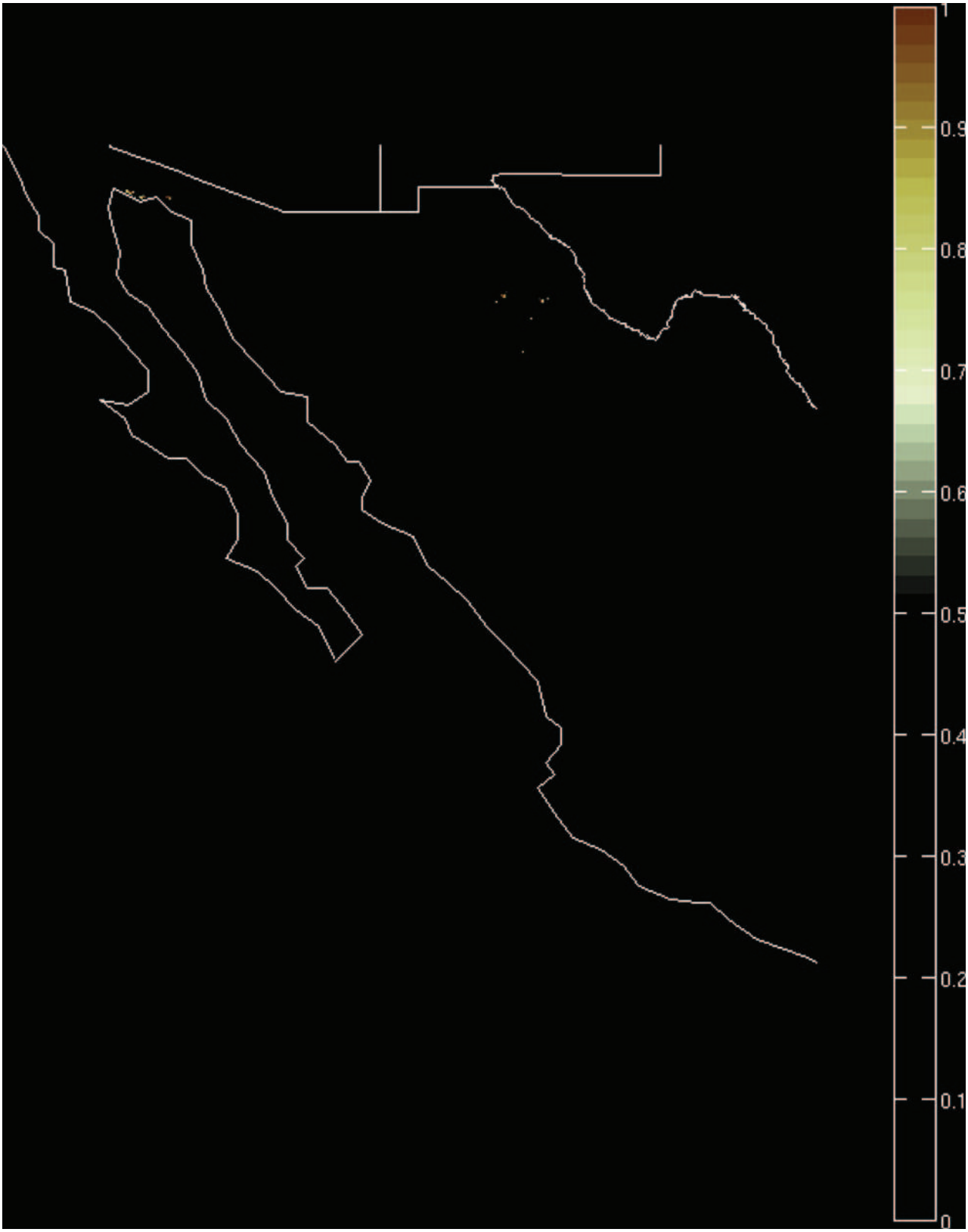

Figure 447. Result of the Maximum Likelihood model (MLE) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

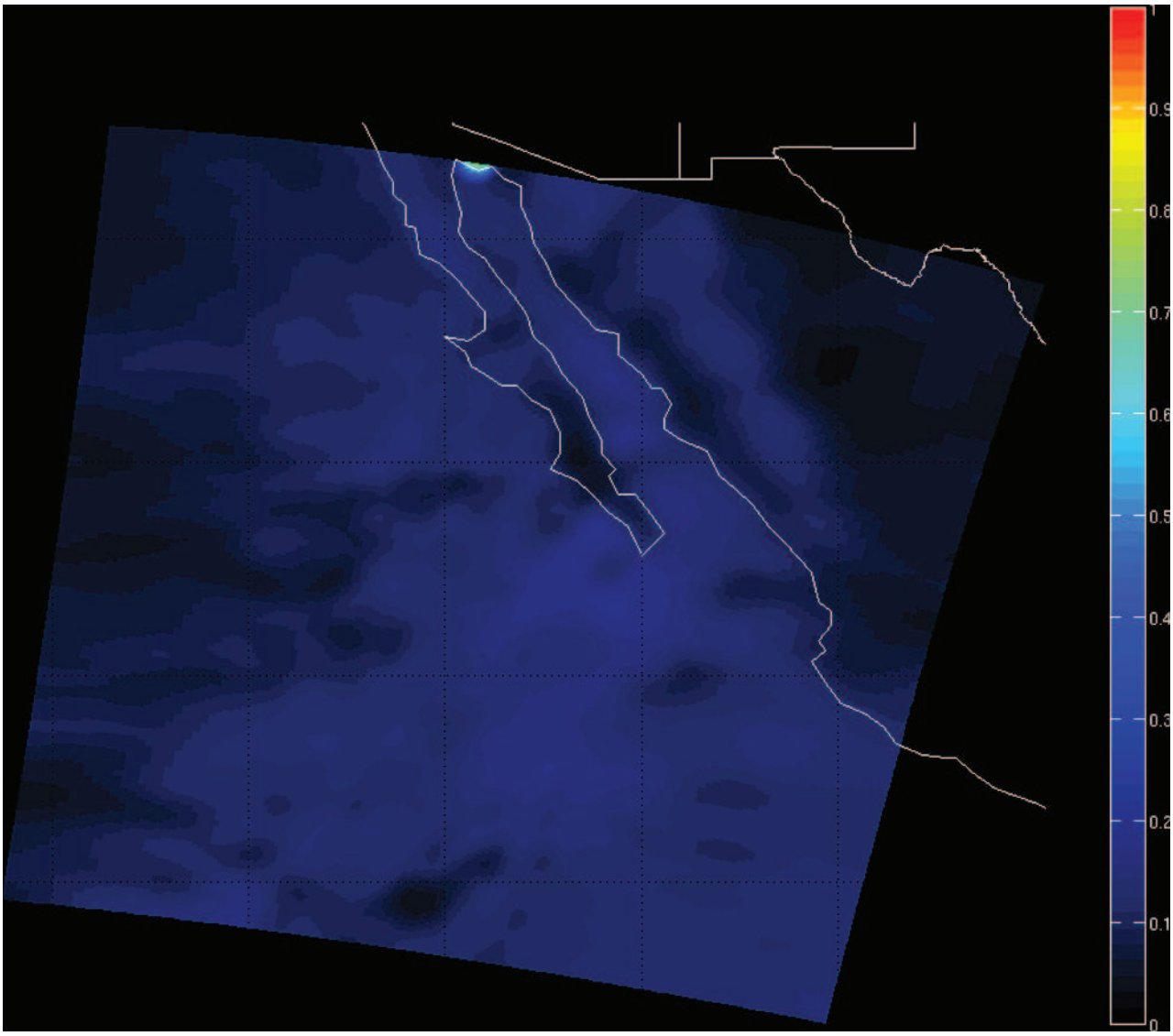

Figure 448. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

Figure 449. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2006-071 at 18:15 UTC.

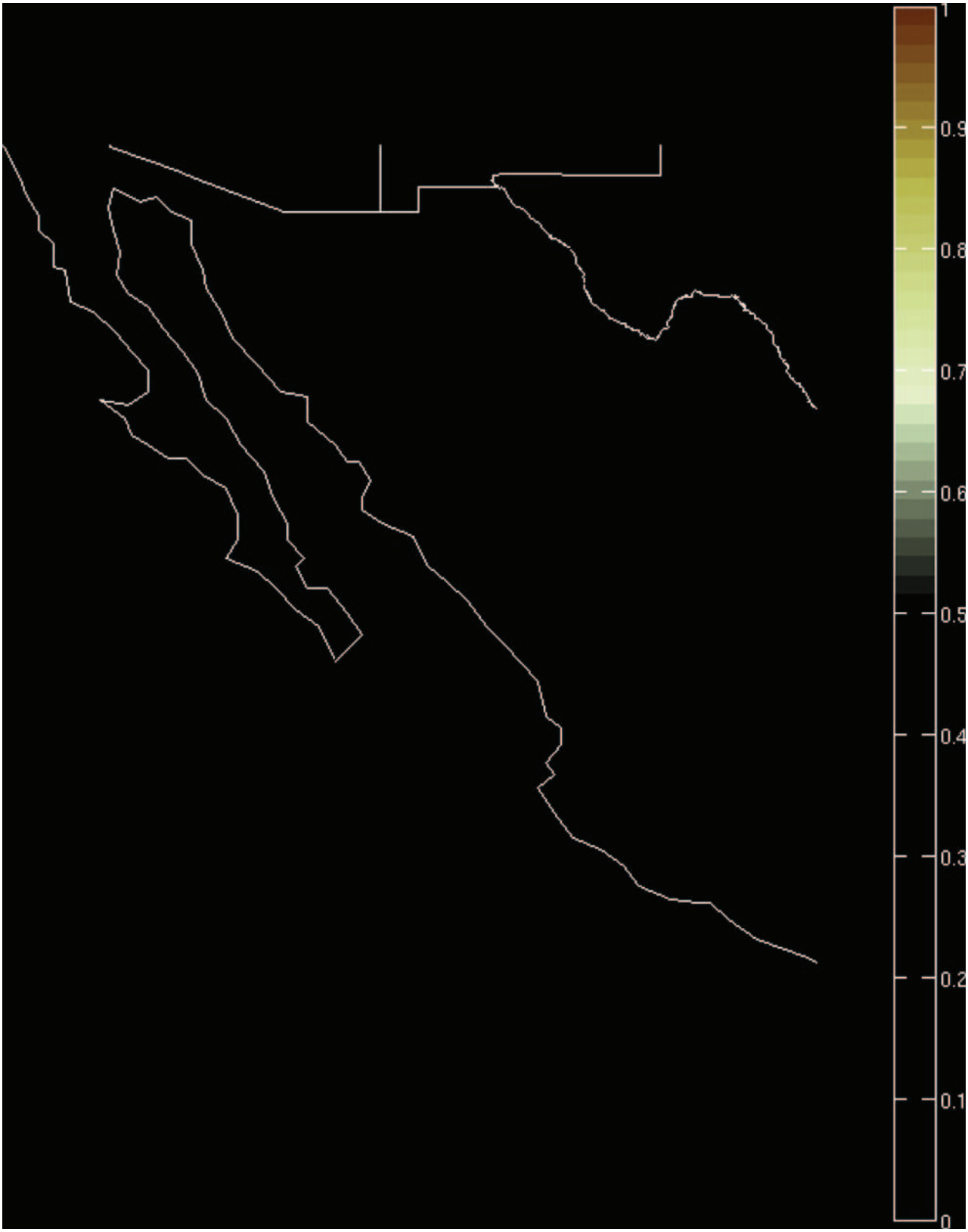

Figure 450. Result of the Maximum a posteriori model (MAP) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

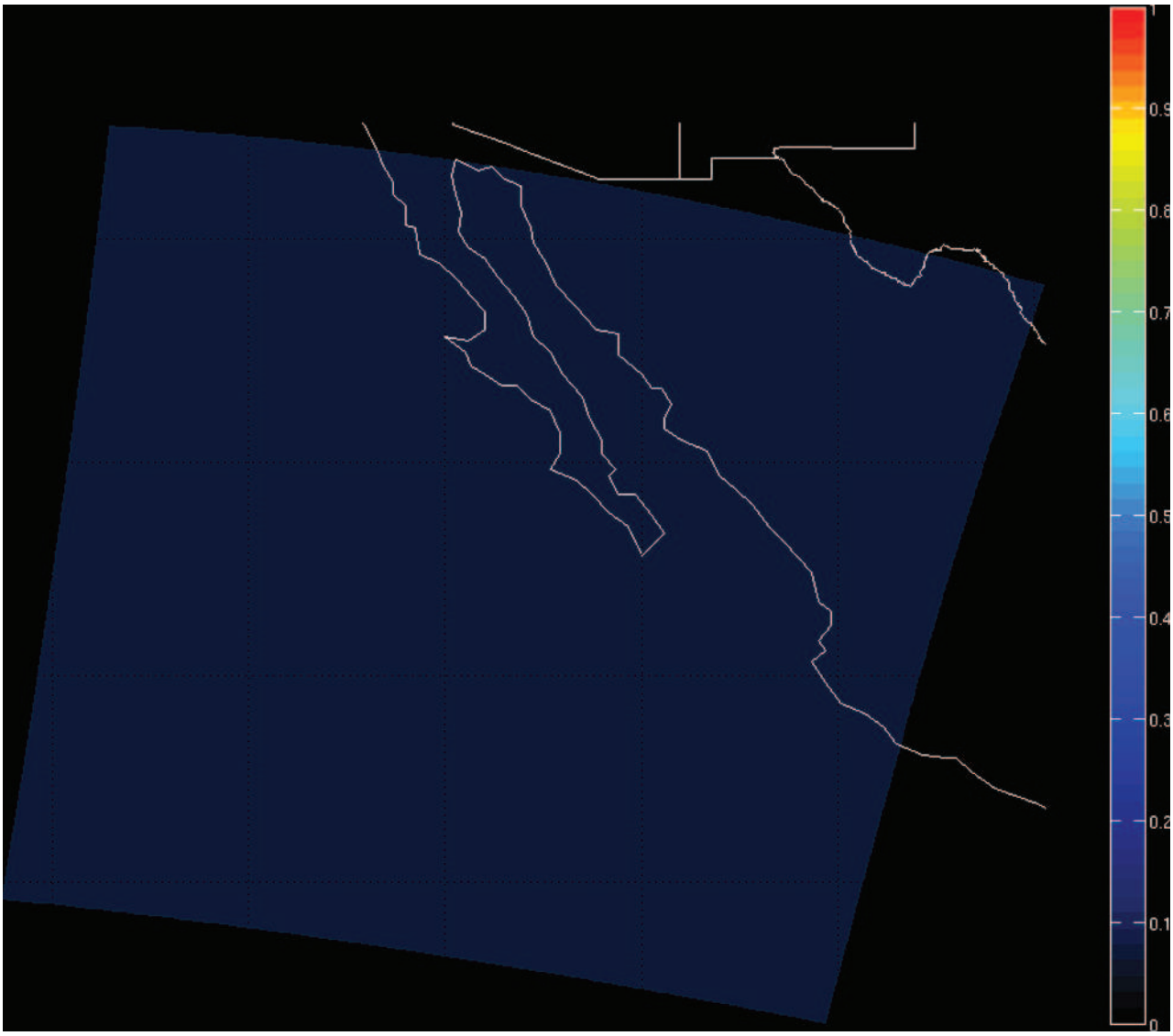

Figure 451. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

Figure 452. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2006-071 at 18:15 UTC.

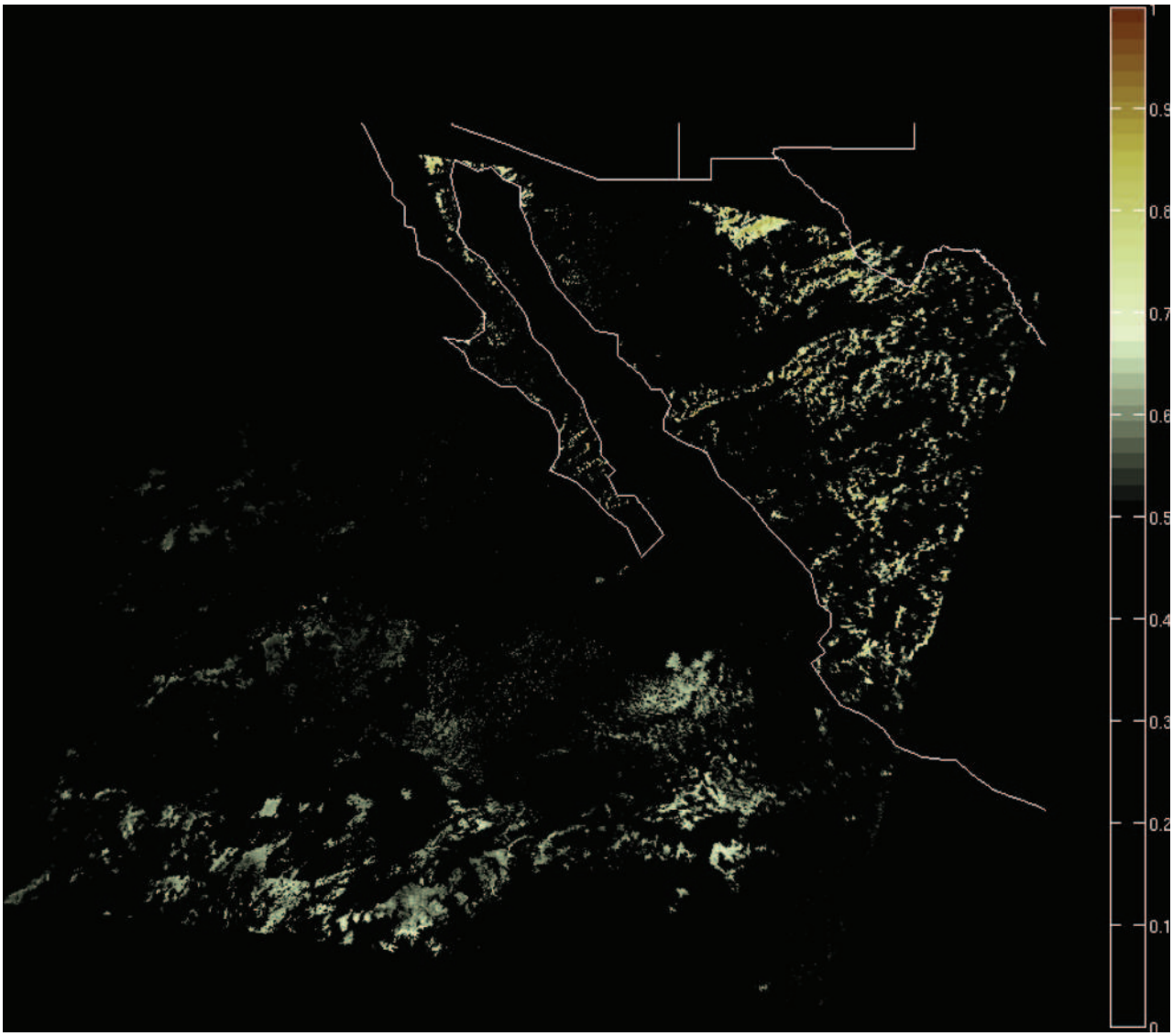

Figure 453. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

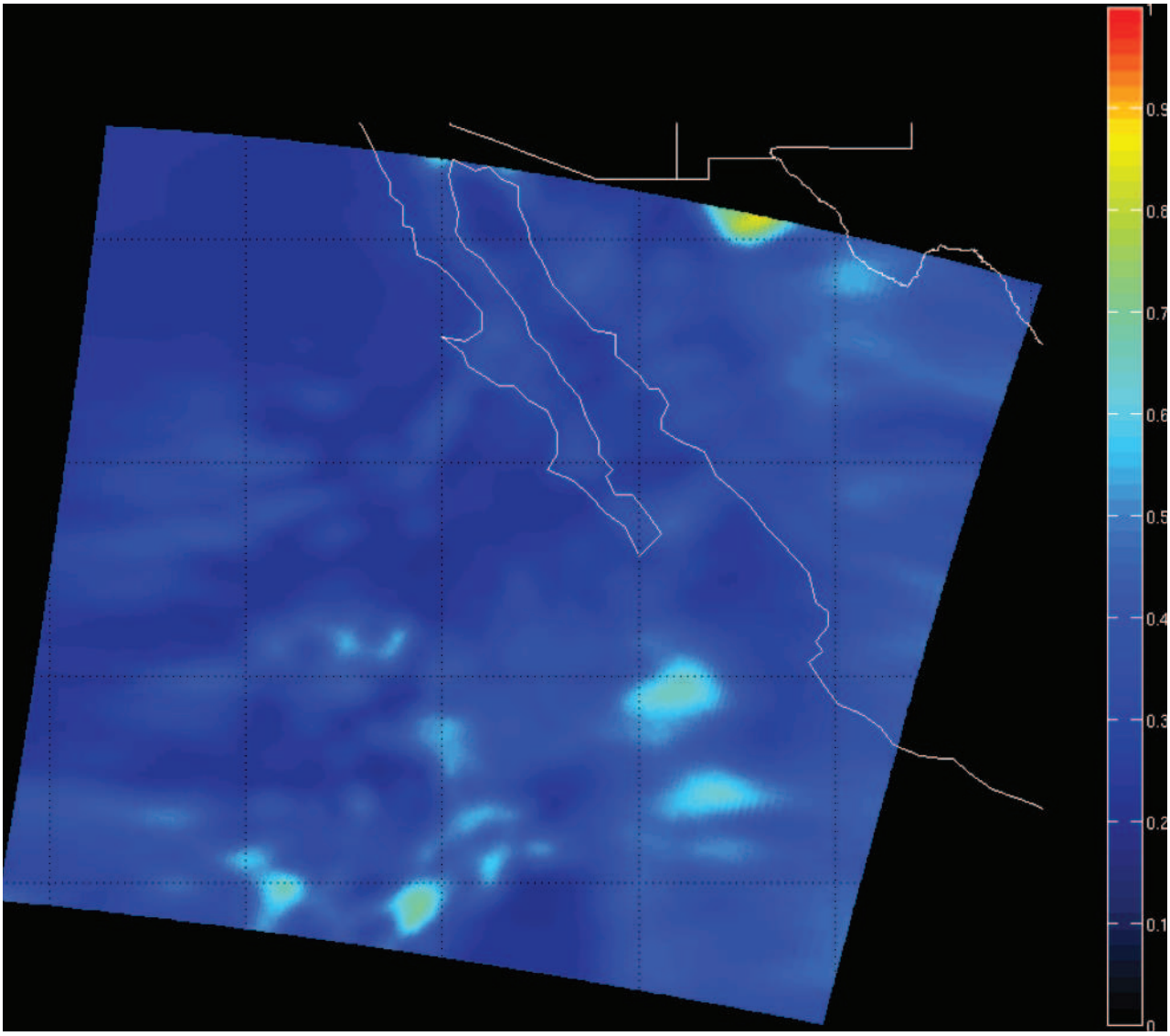

Figure 454. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

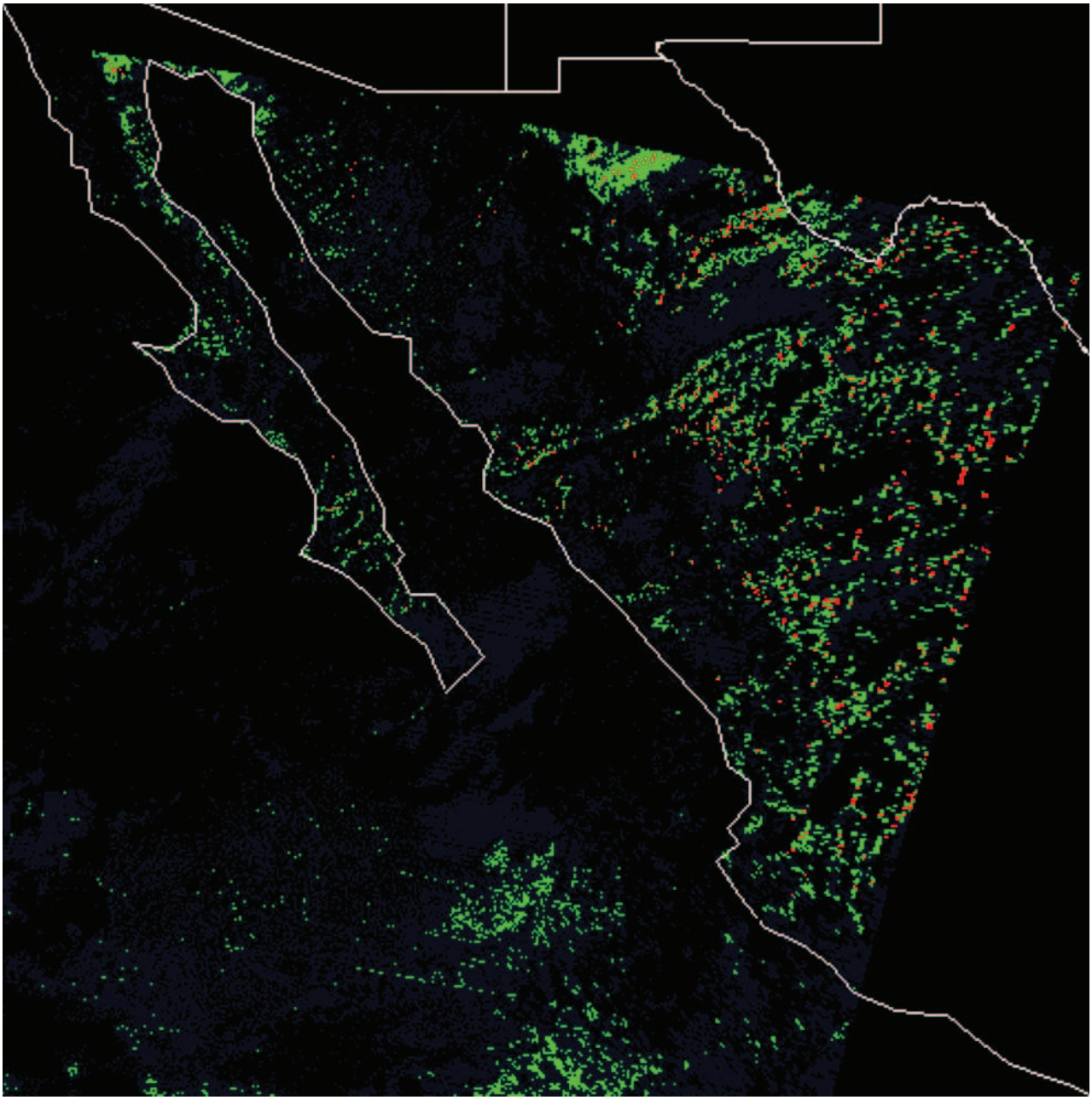

Figure 455. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2006-071 at 18:15 UTC.

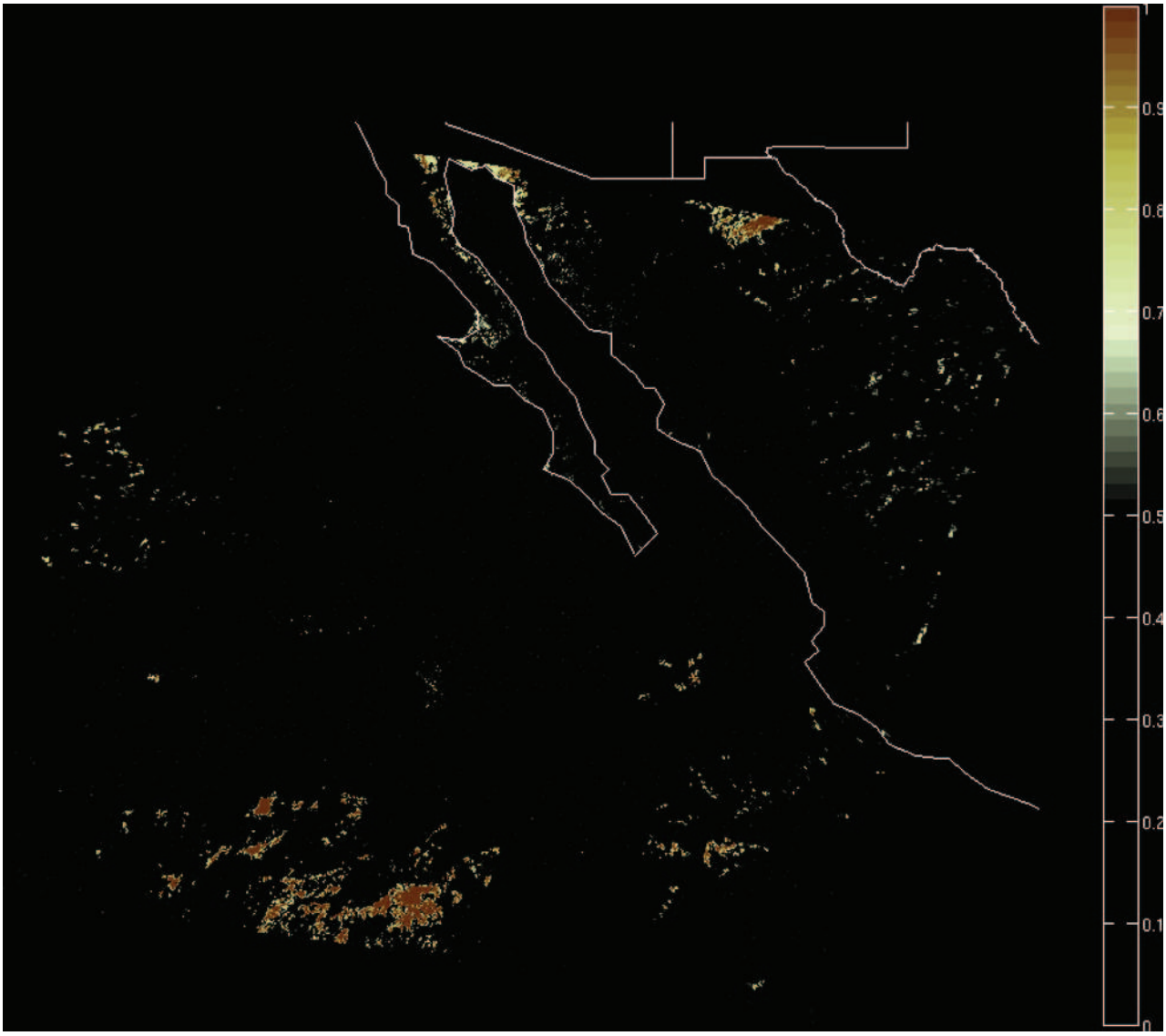

Figure 456. Result of the Probabilistic Neural Network (PNN) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

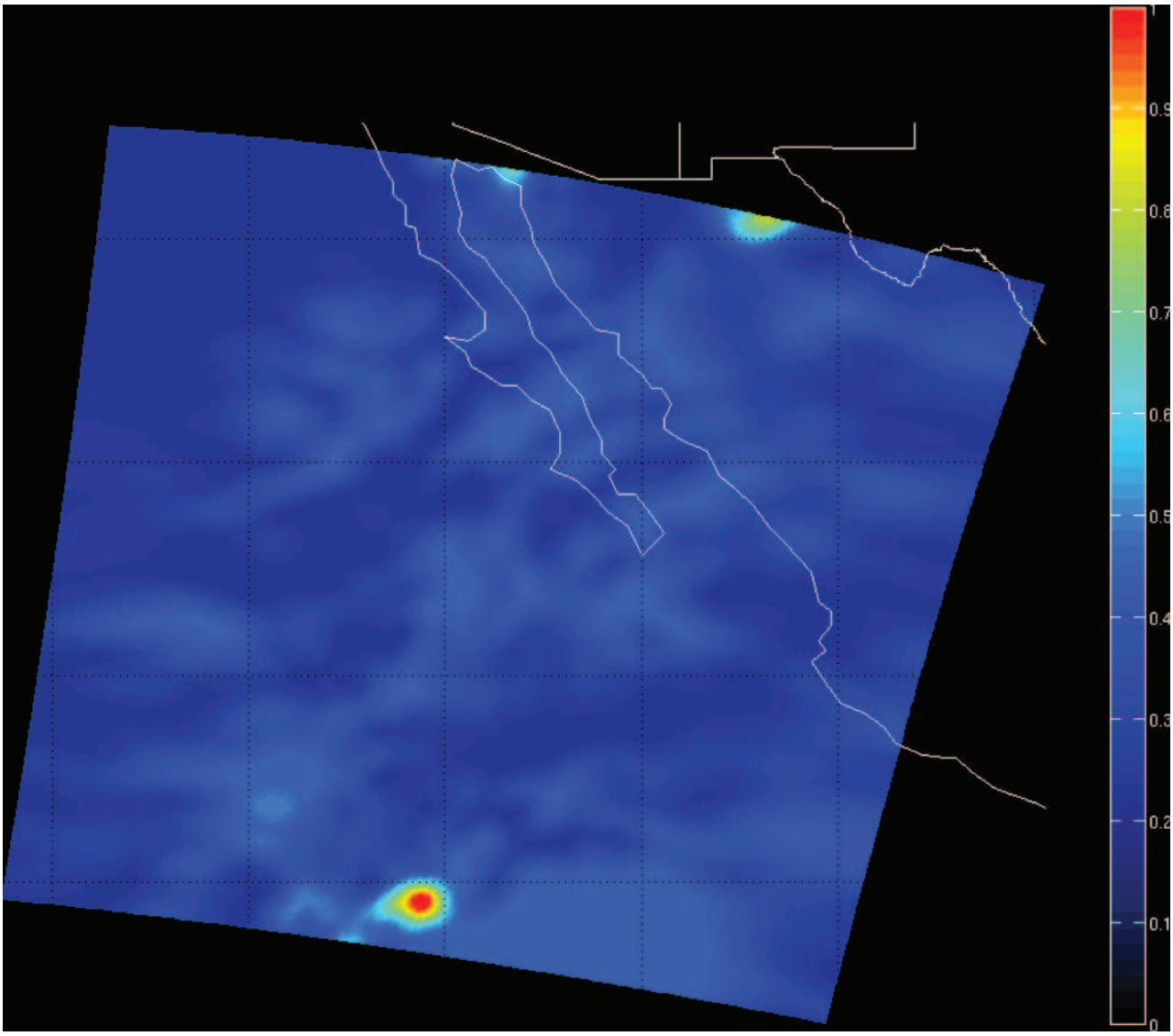

Figure 457. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2006-071 at 18:15 UTC. The colorbar shows the index of probability.

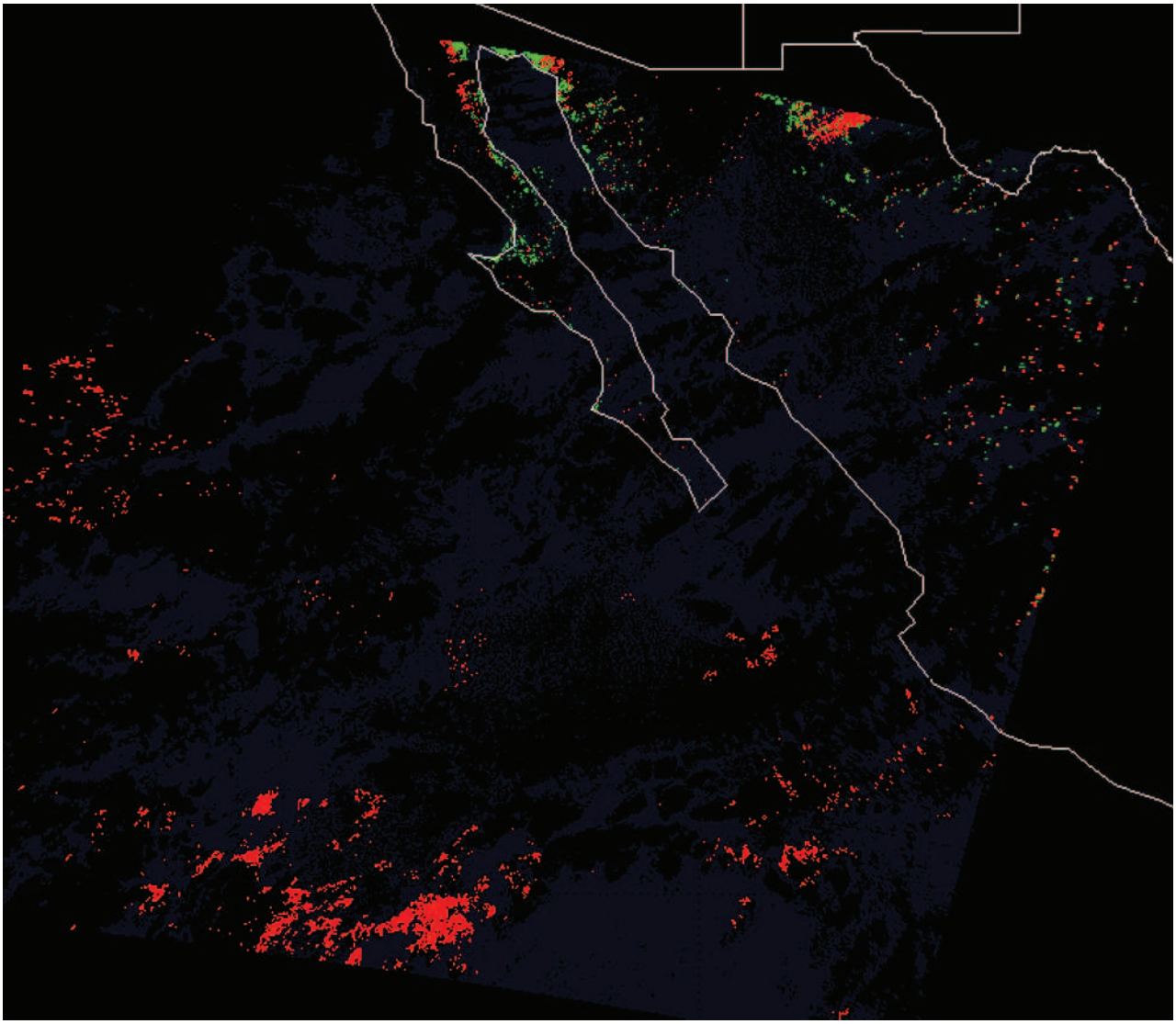

Figure 458. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2006-071 at 18:15 UTC.

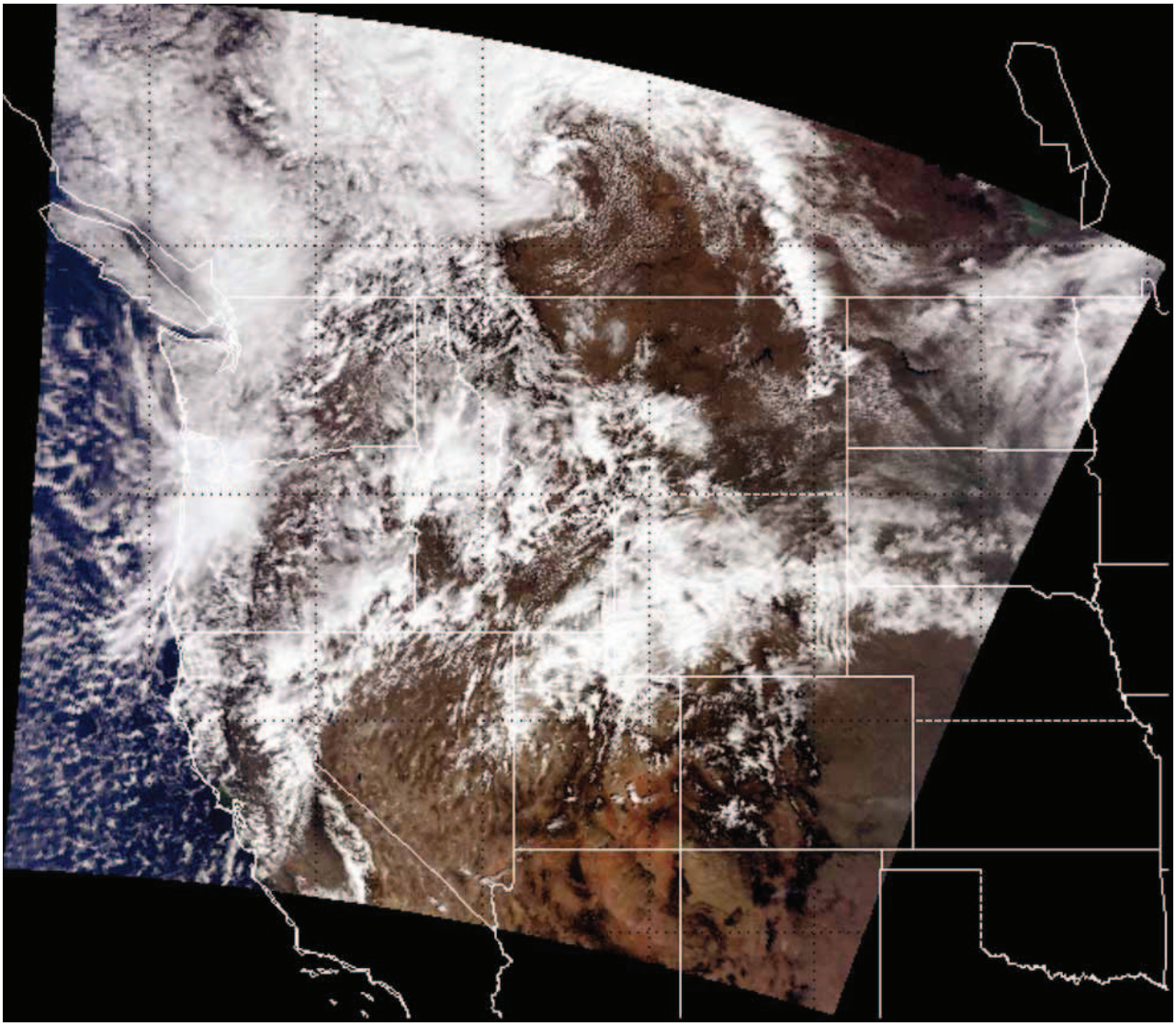

Figure 459. True color image for the event of 2000-131 at 18:50 UTC.

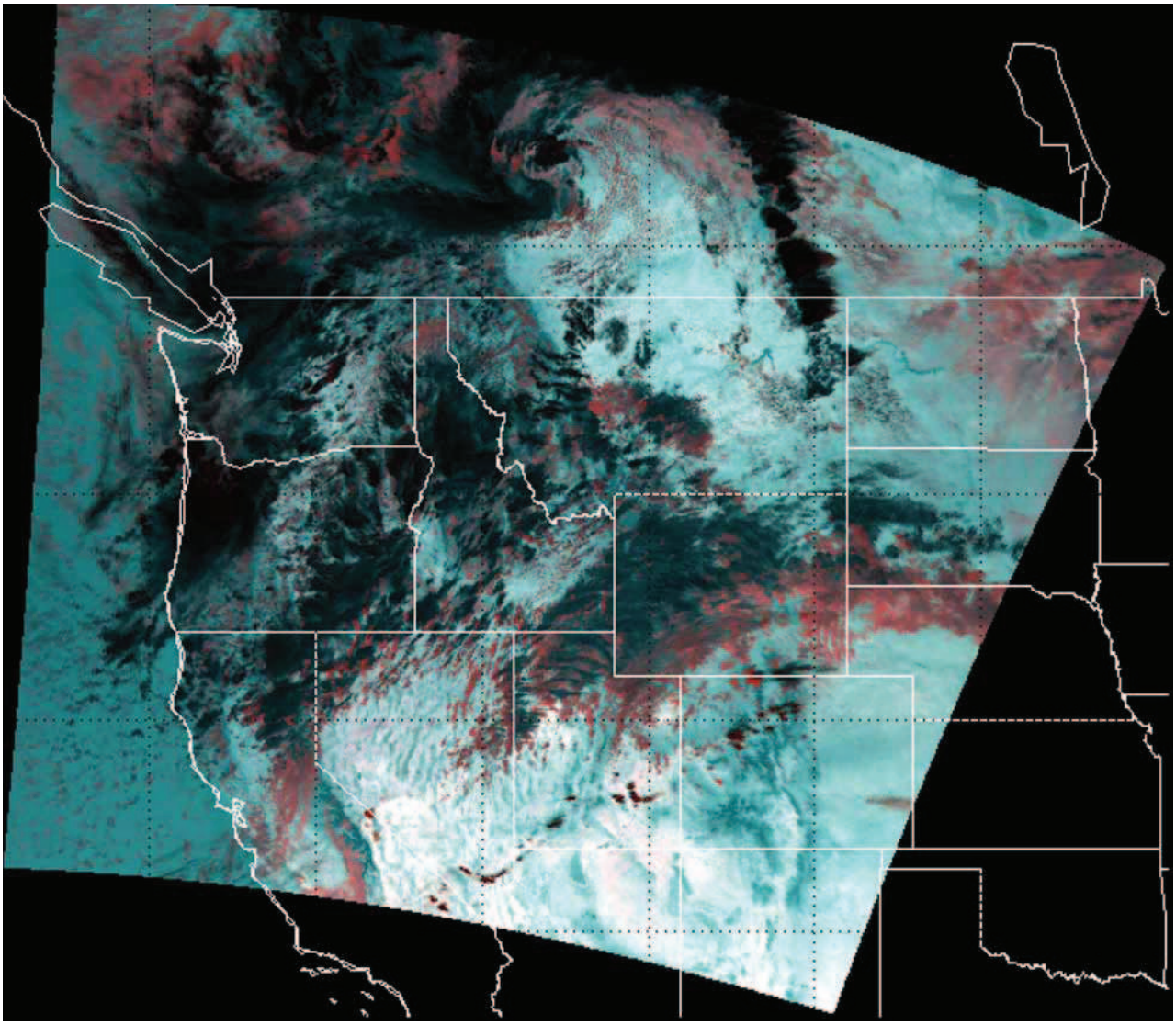

Figure 460. False color image for the event of 2000-131 at 18:50 UTC.

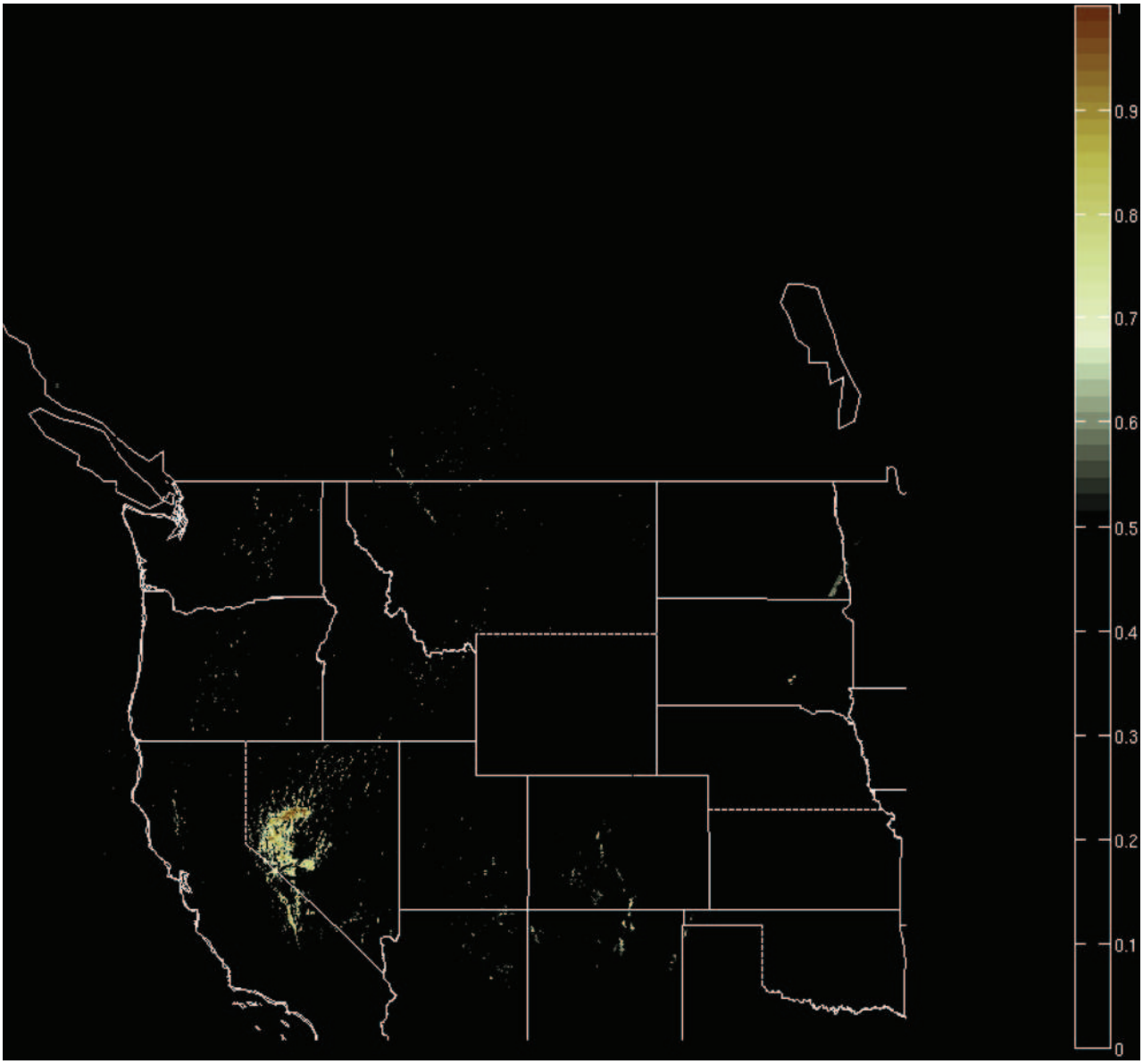

Figure 461. Result of the probabilistic display model (PD) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

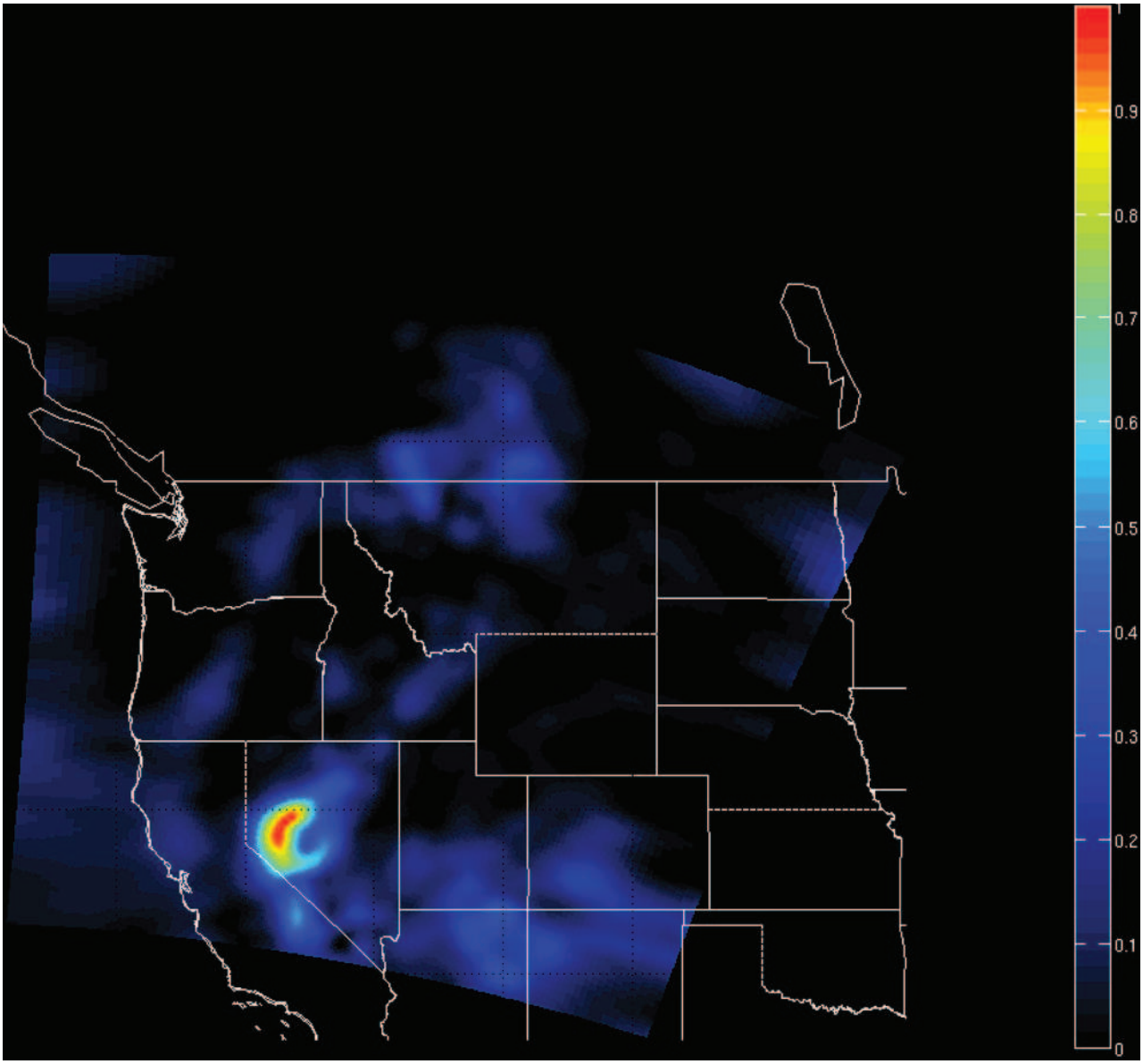

Figure 462. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

Figure 463. Result of the Maximum Likelihood model (MLE) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

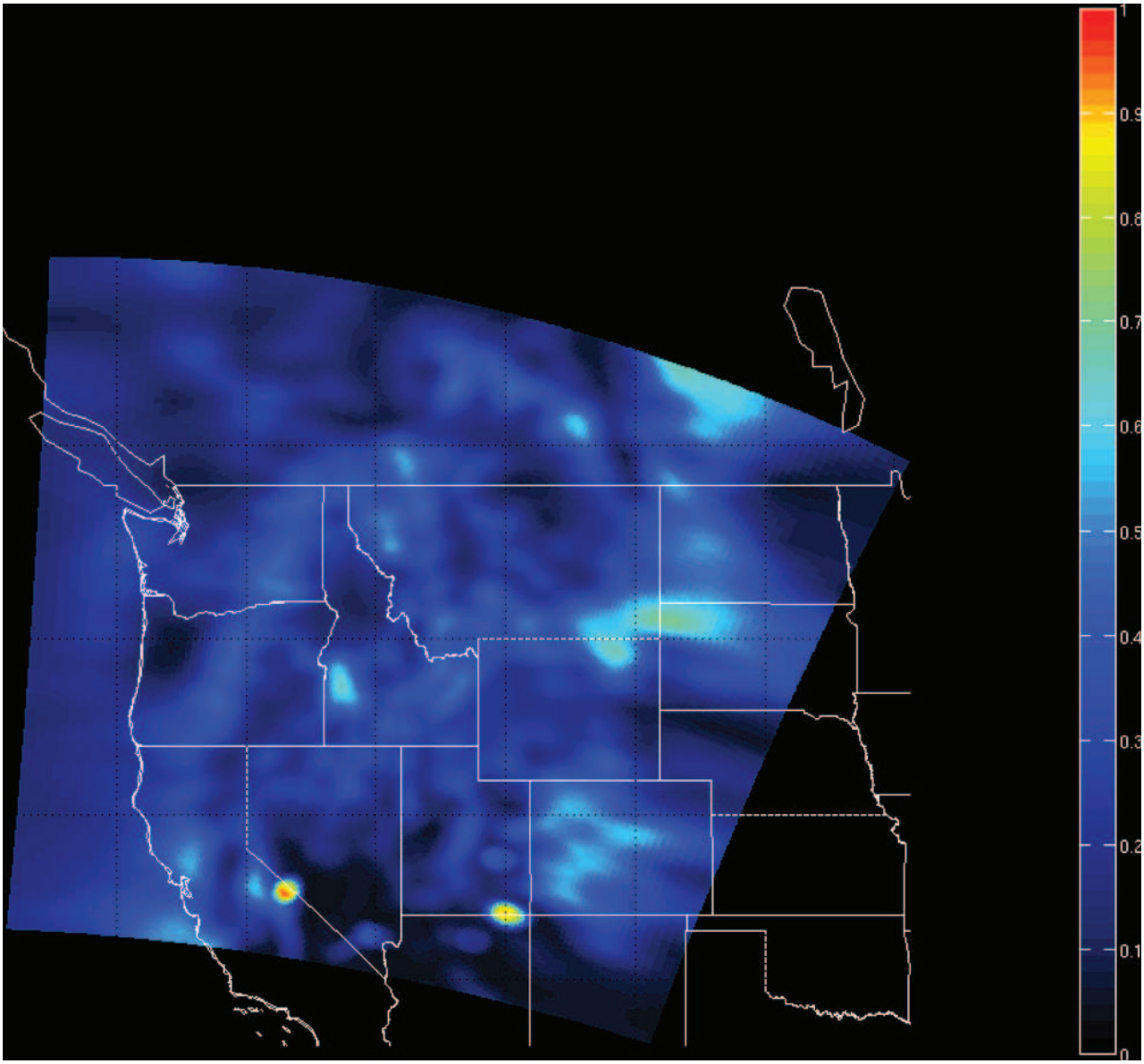

Figure 464. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

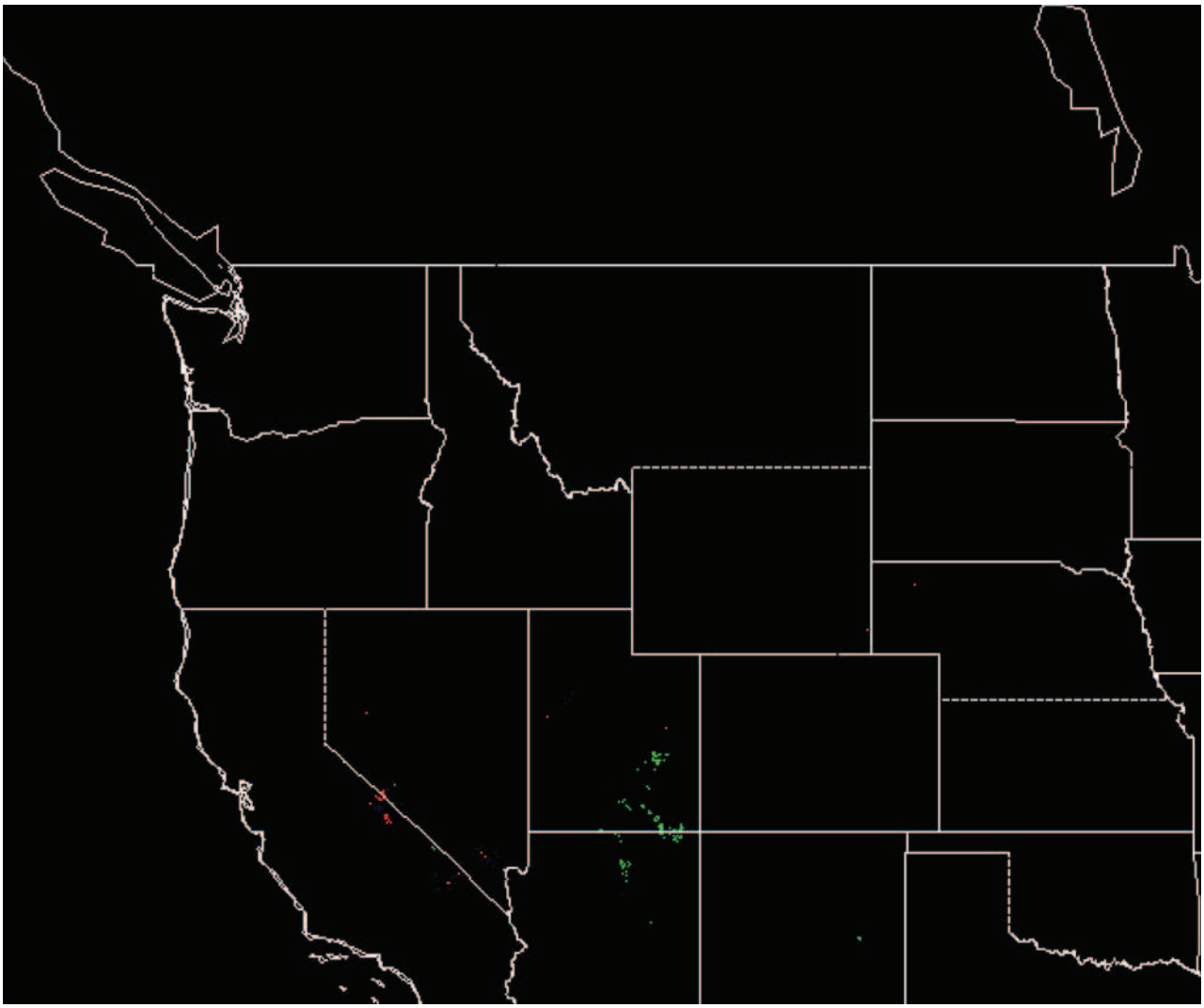

Figure 465. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2000-131 at 18:50 UTC.

Figure 466. Result of the Maximum a posteriori model (MAP) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

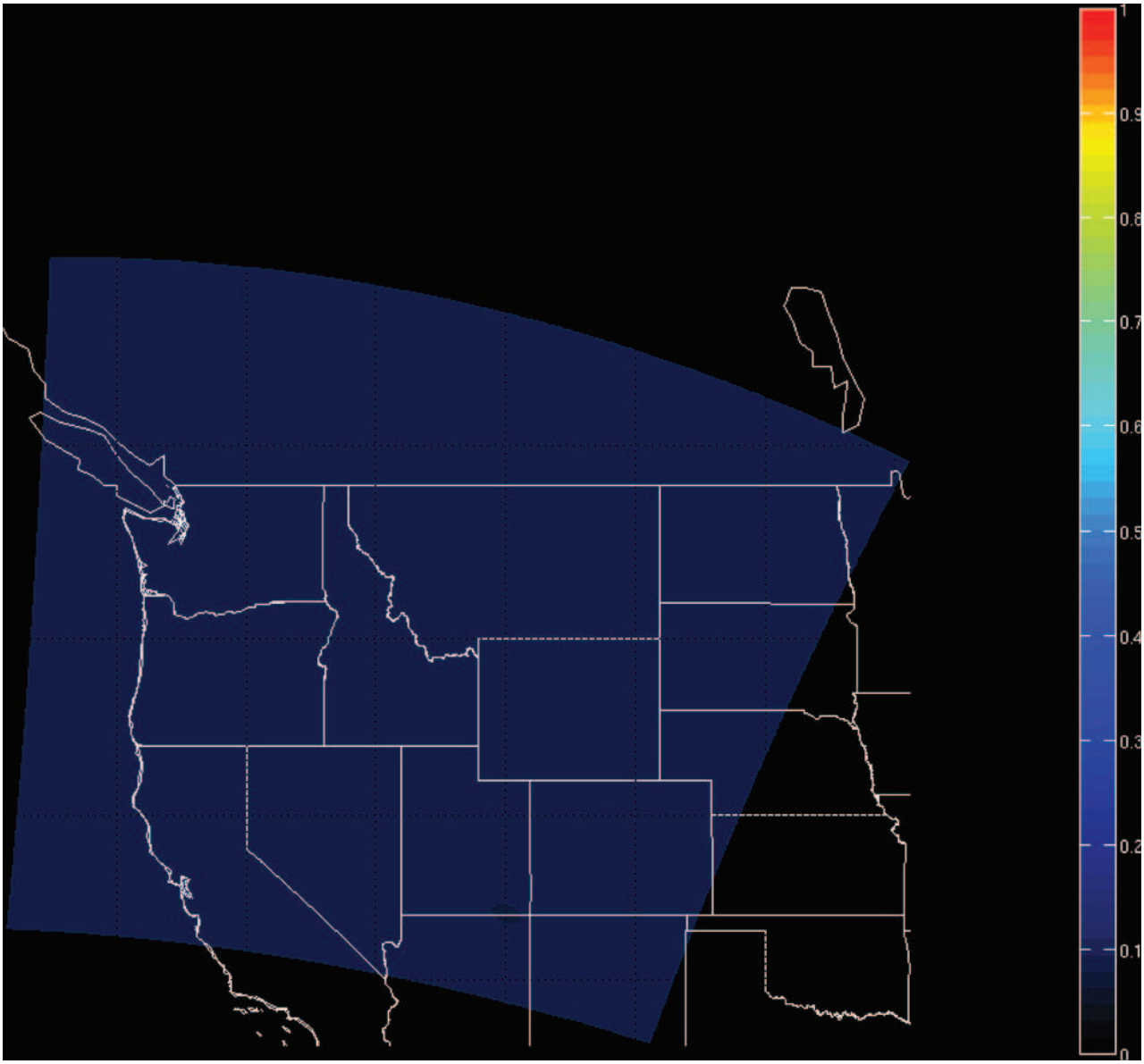

Figure 467. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

Figure 468. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2000-131 at 18:50 UTC.

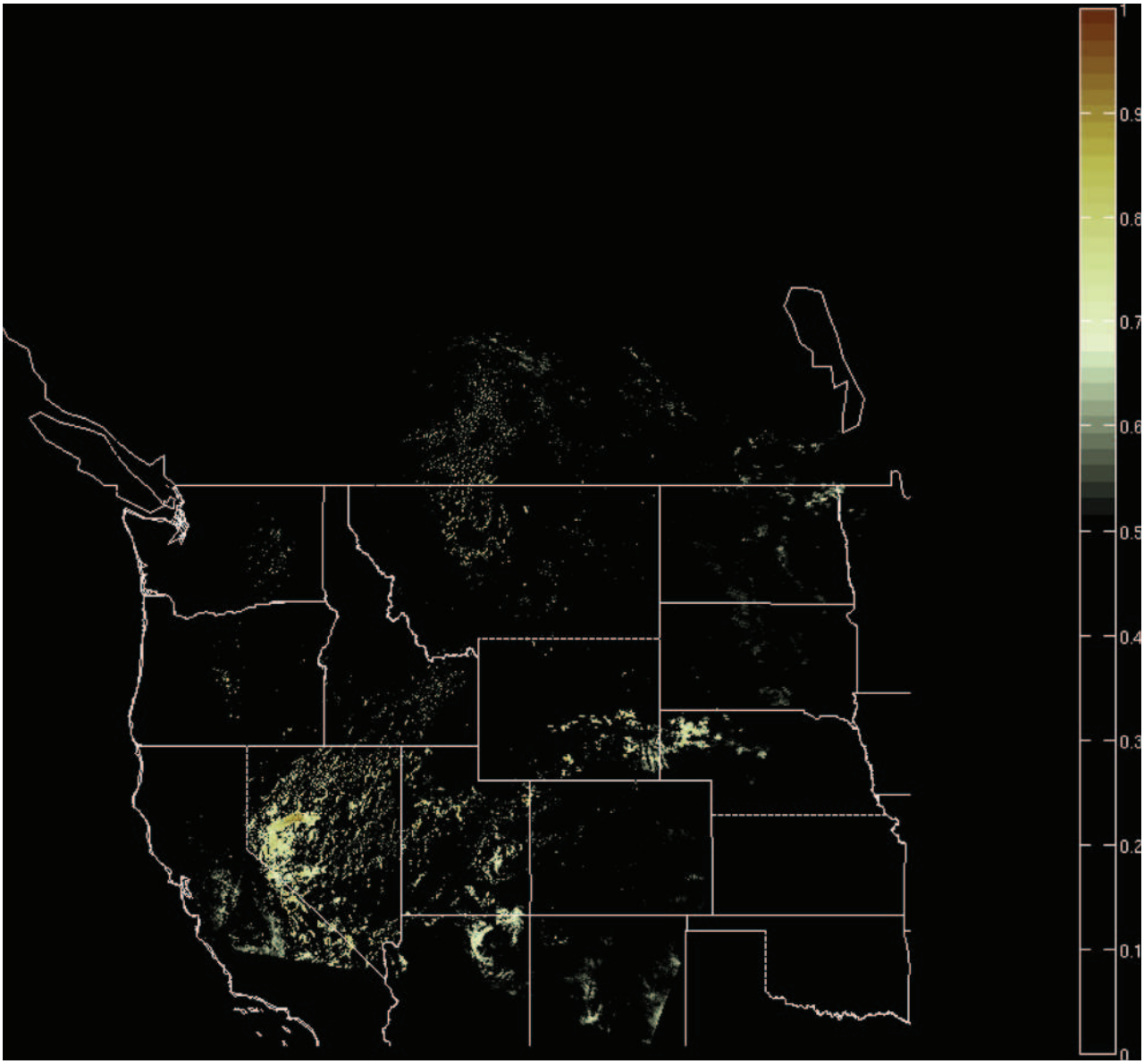

Figure 469. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

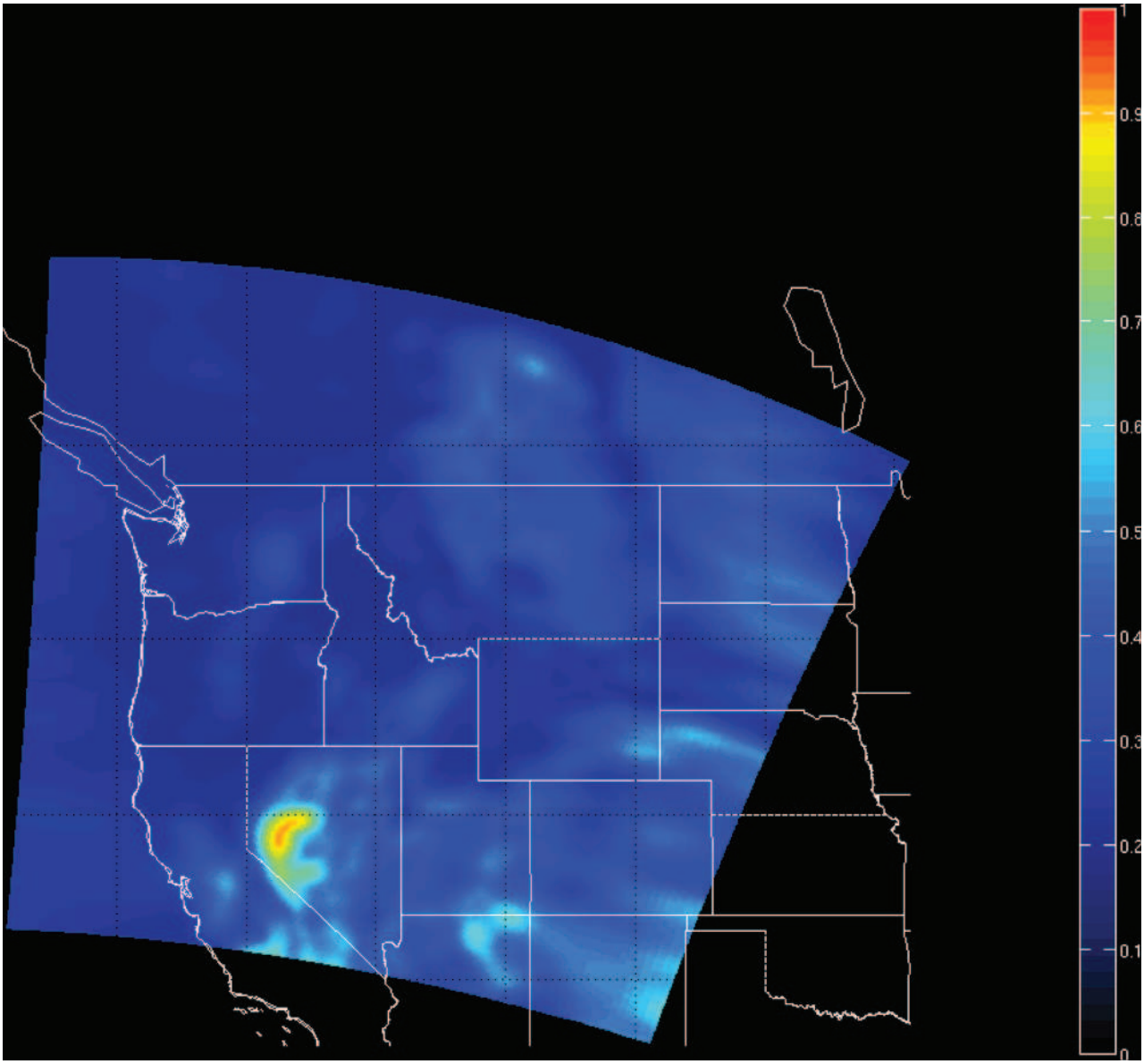

Figure 470. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

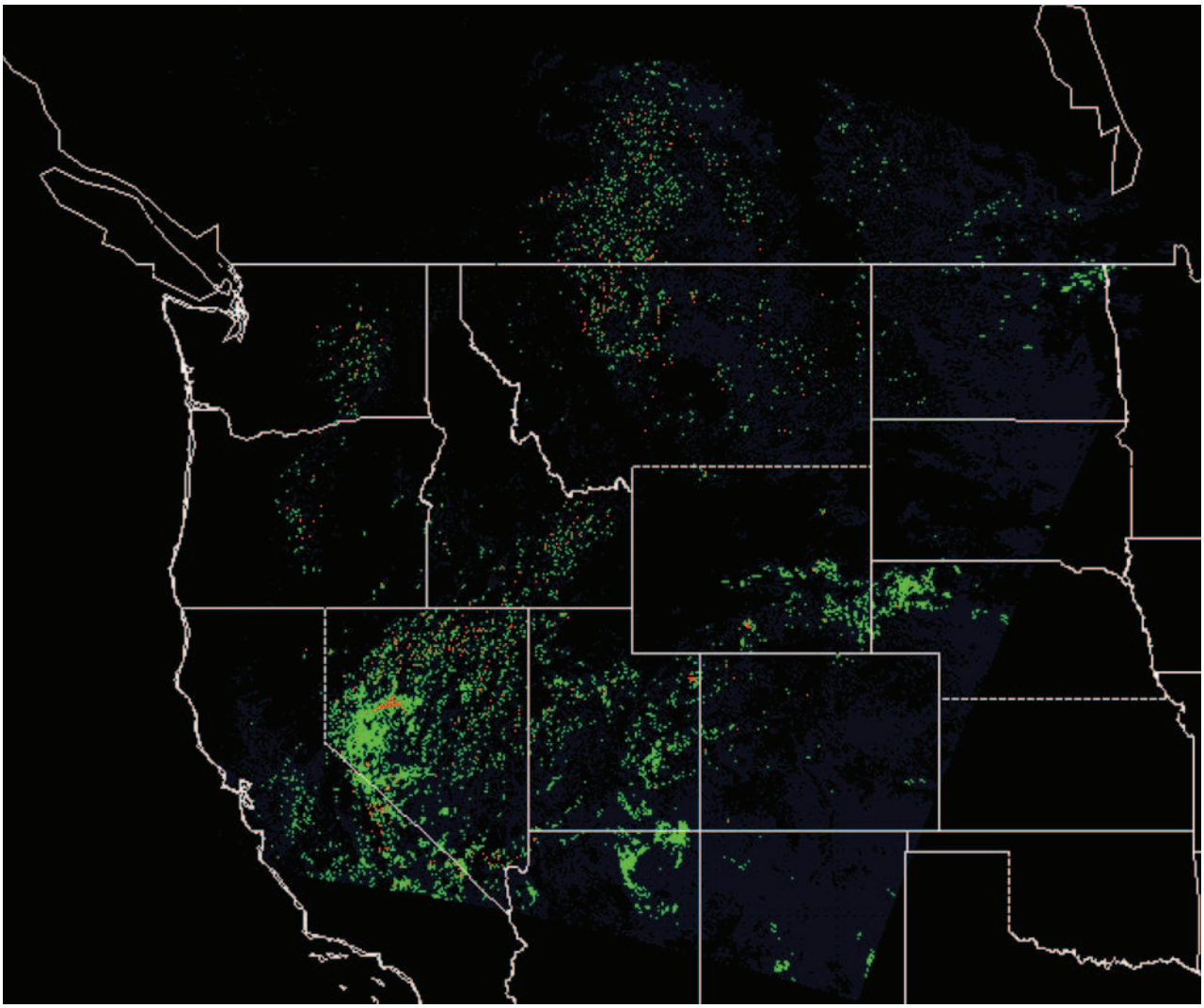

Figure 471. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2000-131 at 18:50 UTC.

Figure 472. Result of the Probabilistic Neural Network (PNN) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

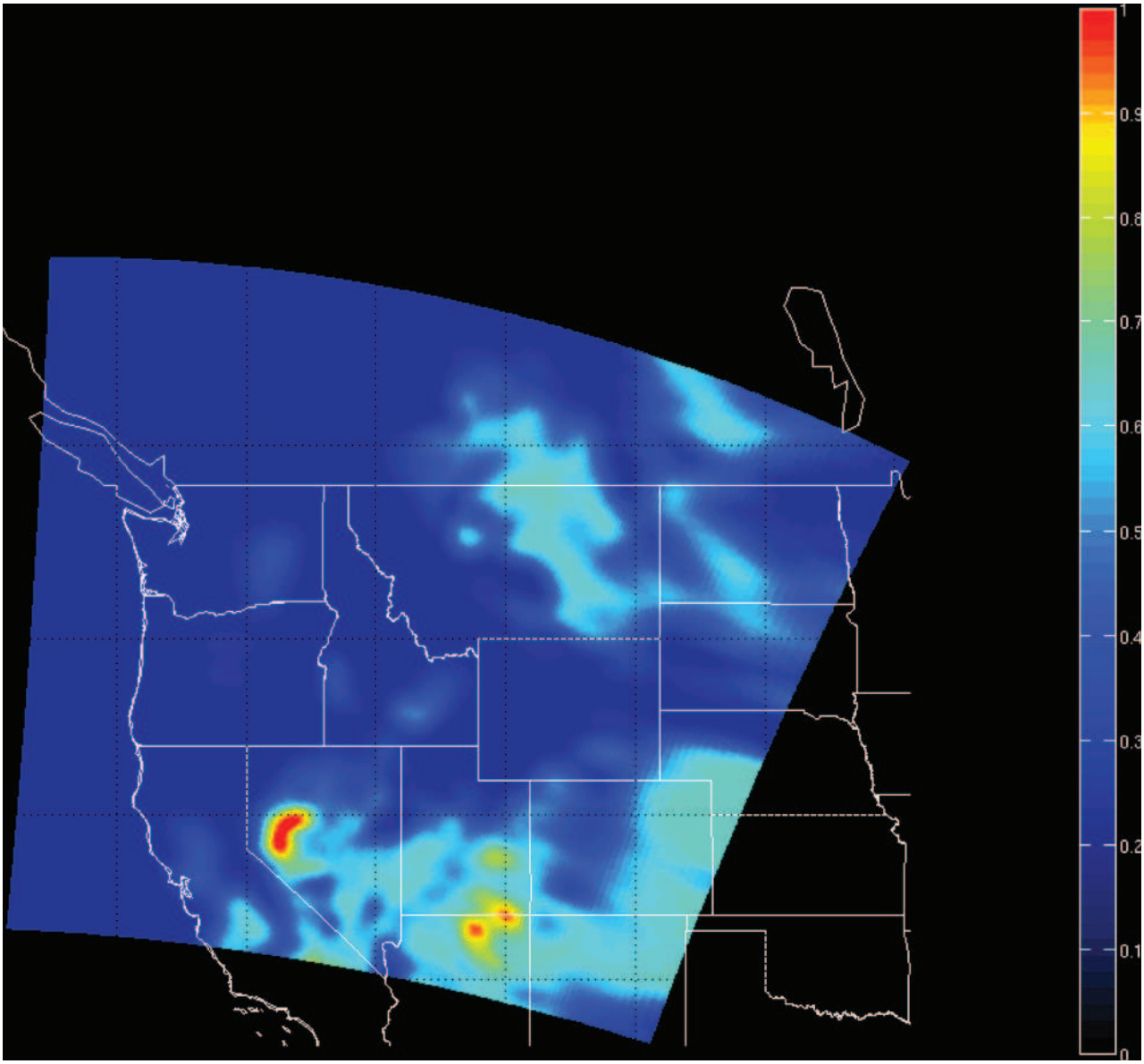

Figure 473. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2000-131 at 18:50 UTC. The colorbar shows the index of probability.

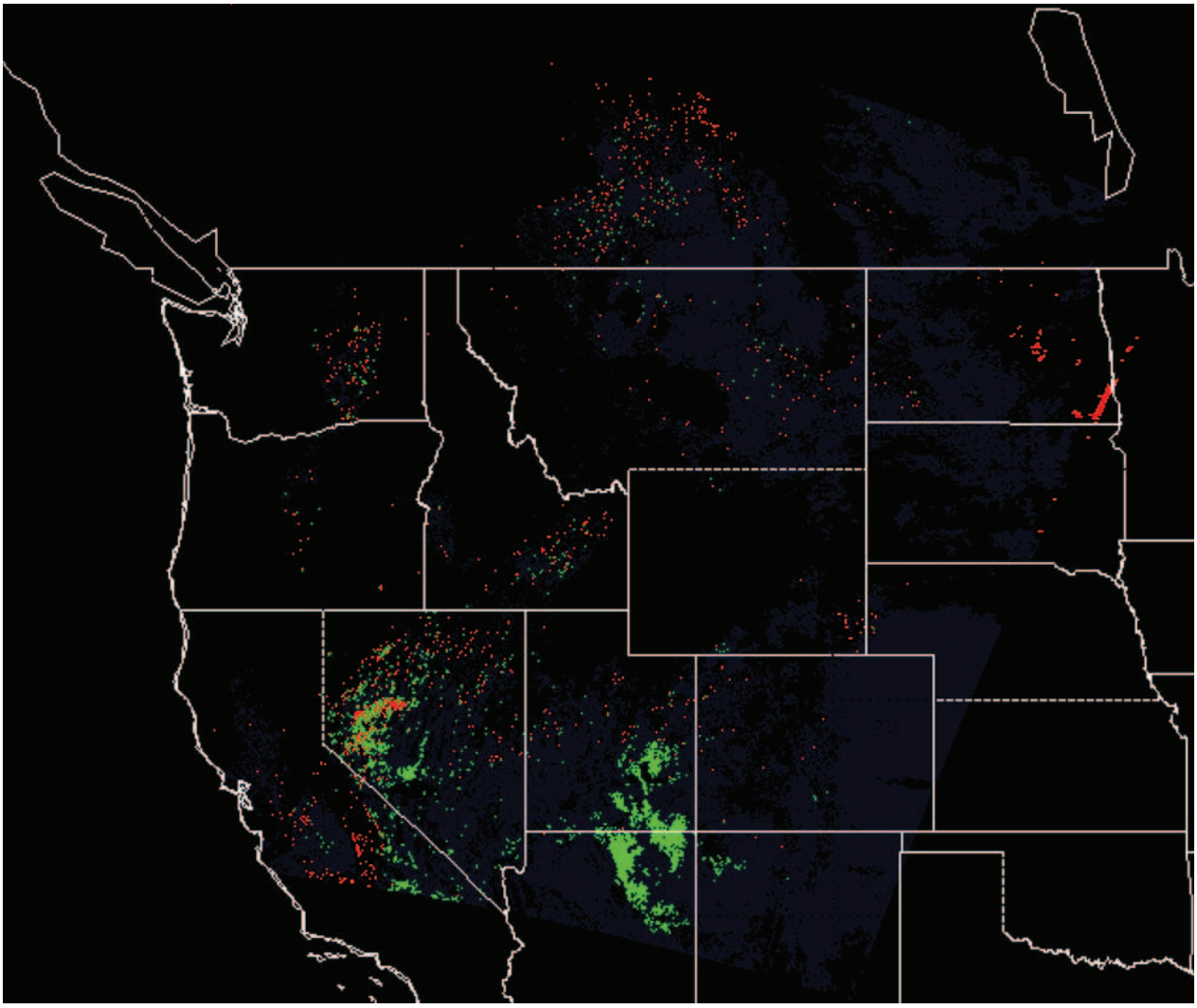

Figure 474. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2000-131 at 18:50 UTC.

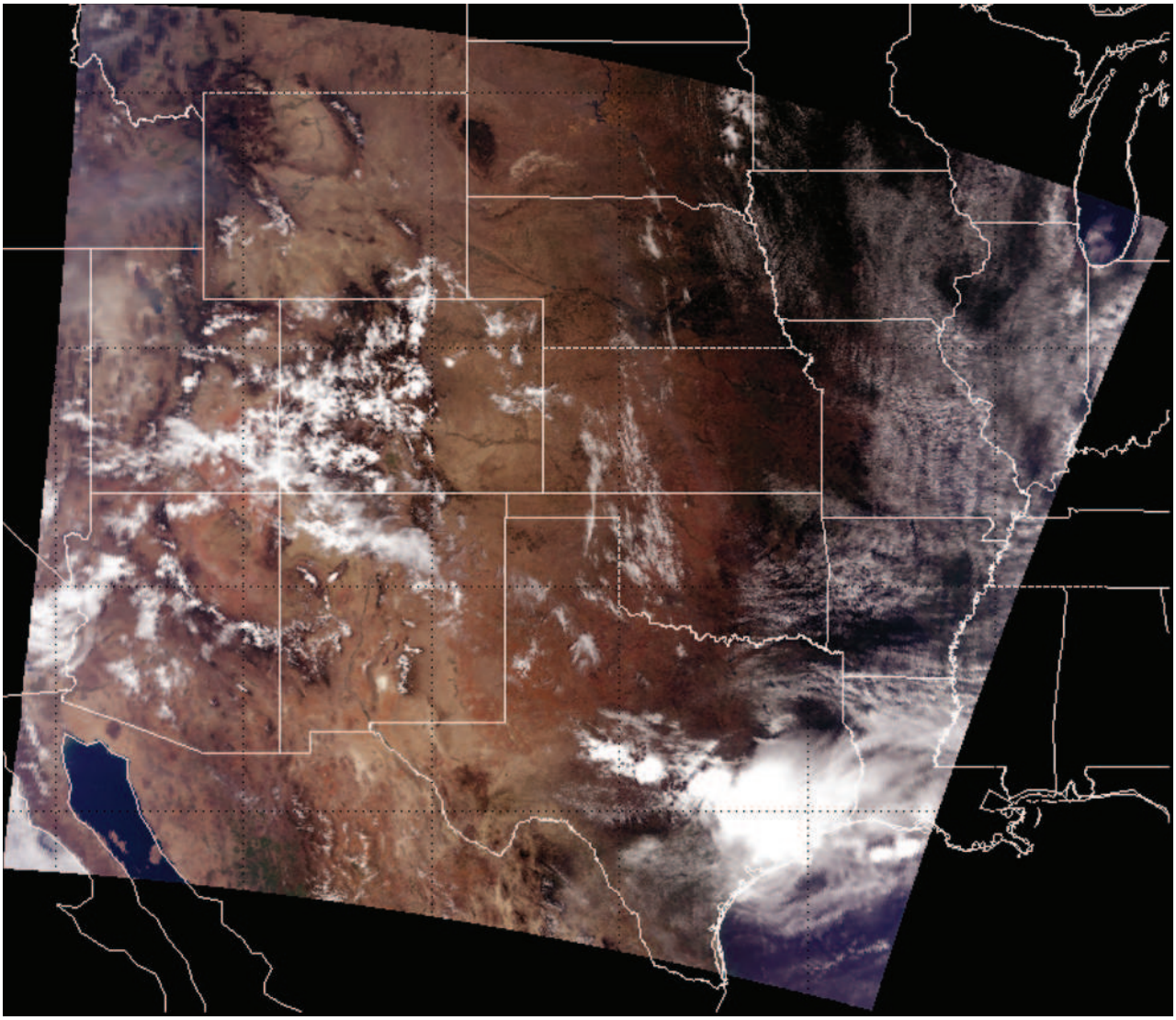

Figure 475. True color image for the event of 2003-204 at 17:40 UTC.

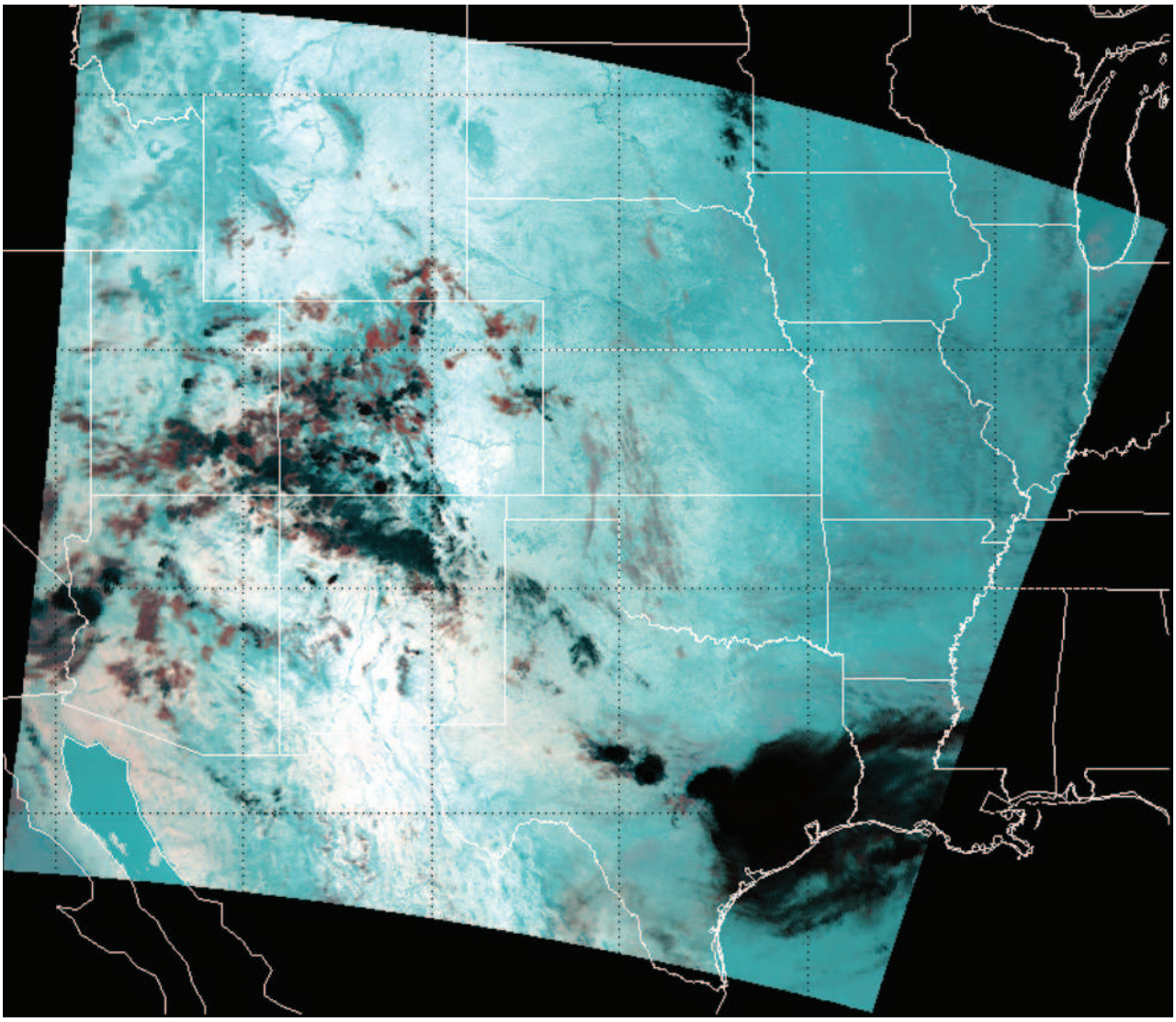

Figure 476. False color image for the event of 2003-204 at 17:40 UTC.

Figure 477. Result of the probabilistic display model (PD) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

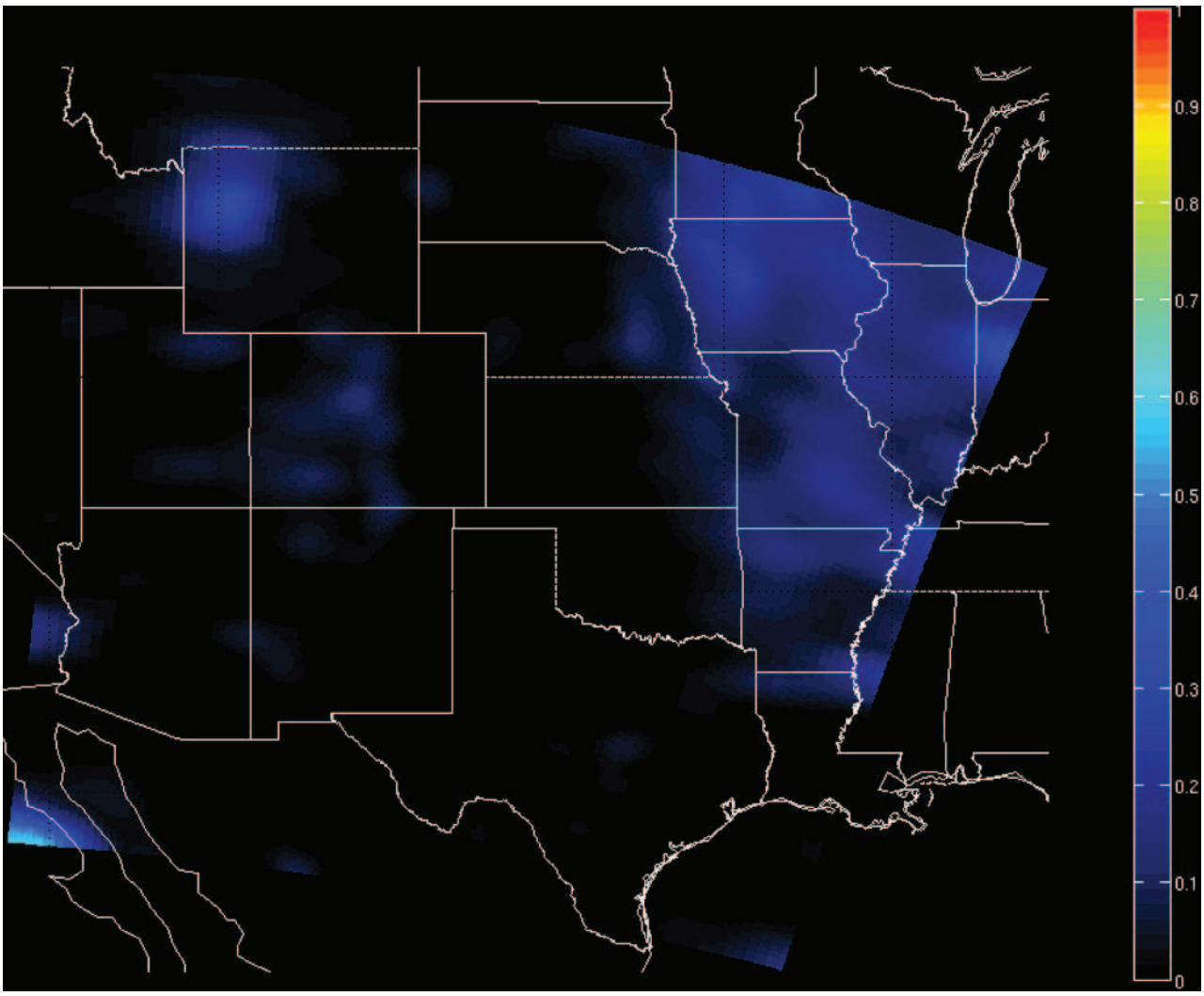

Figure 478. Downsampled and filtered result of the probabilistic display model (PD) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

Figure 479. Result of the Maximum Likelihood model (MLE) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

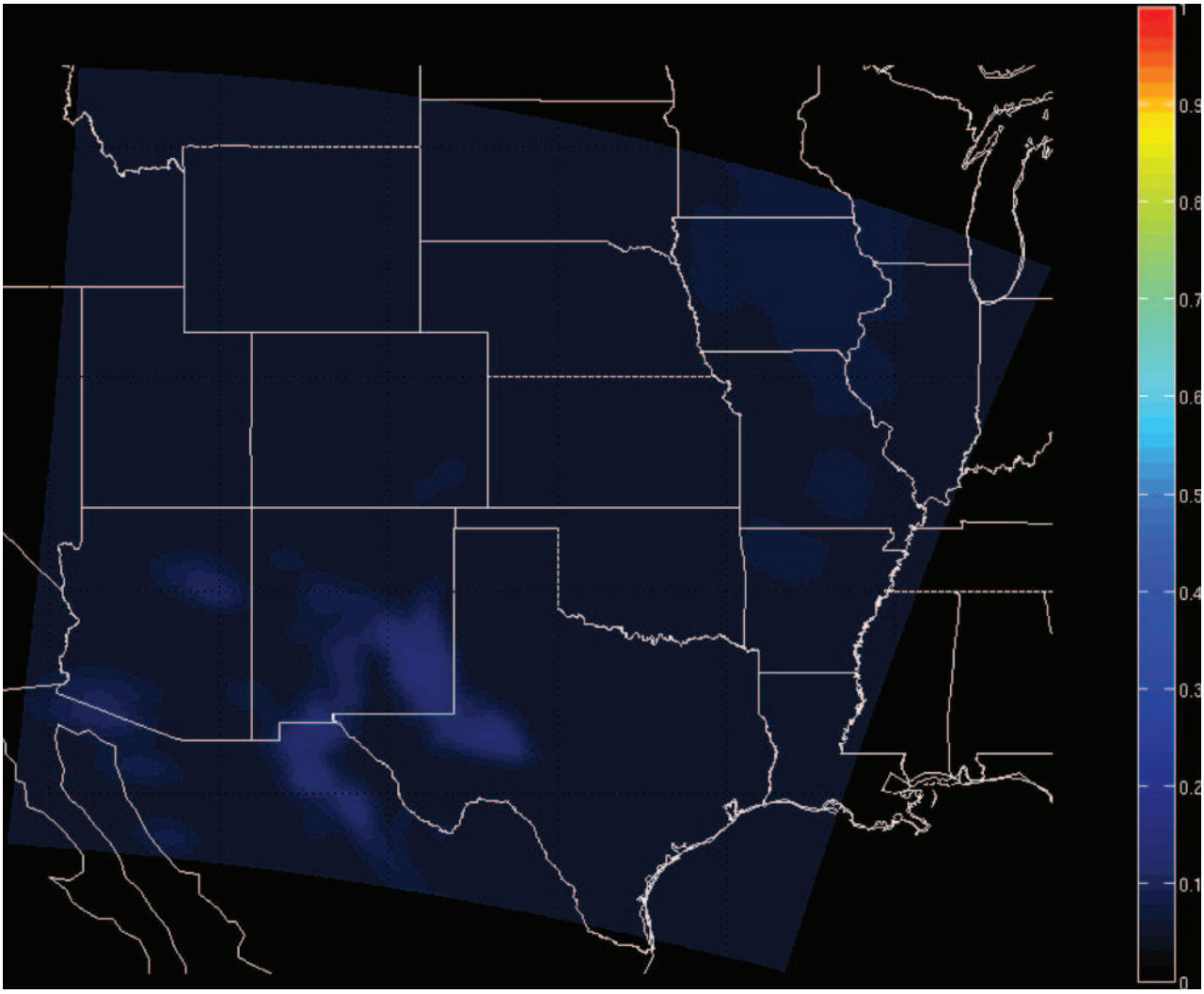

Figure 480. Downsampled and filtered result of the Maximum Likelihood model (MLE) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

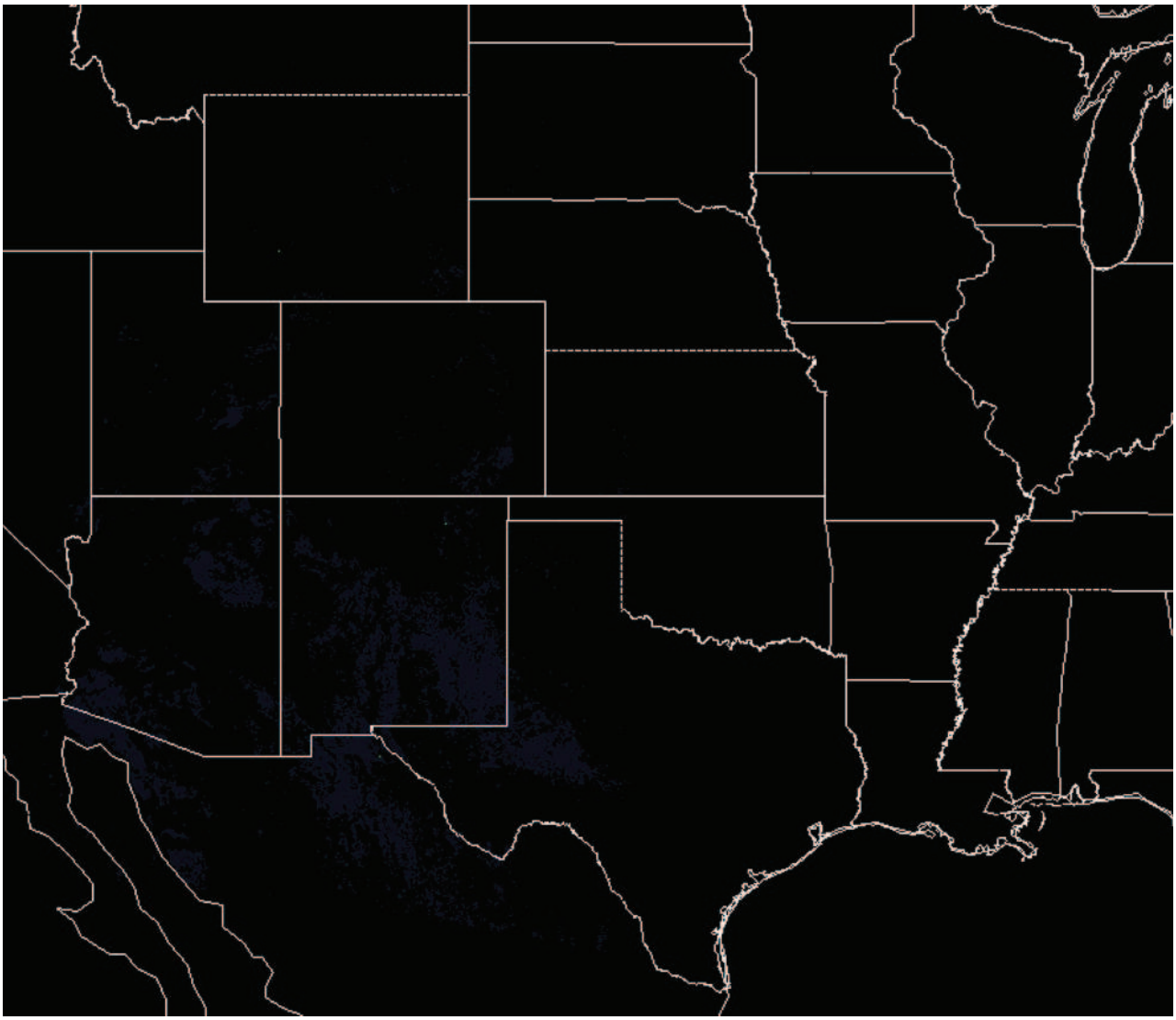

Figure 481. Segmentation result of the Maximum Likelihood model (MLE) for the event of 2003-204 at 17:40 UTC.

Figure 482. Result of the Maximum a posteriori model (MAP) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

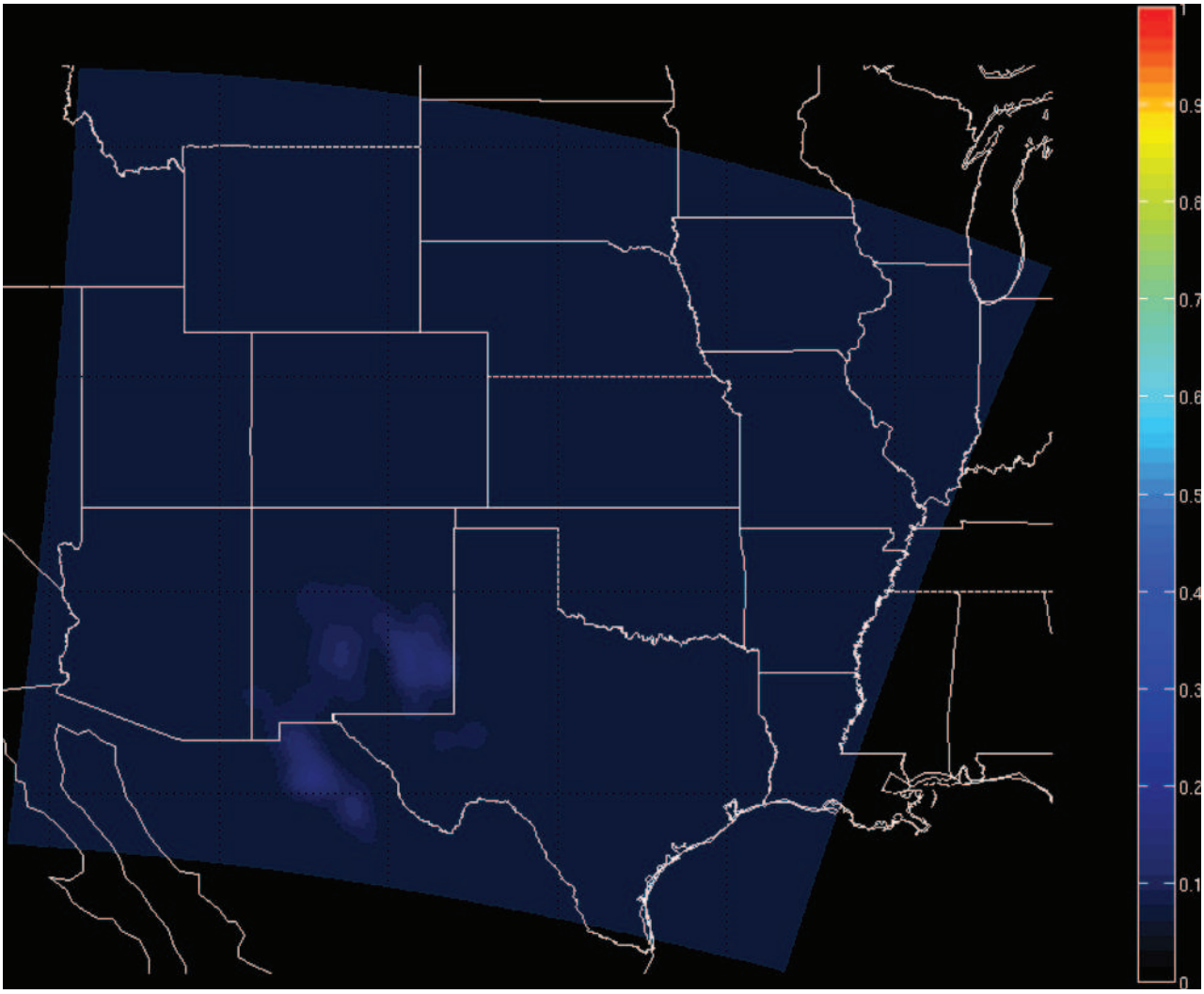

Figure 483. Downsampled and filtered result of the Maximum a posteriori model (MAP) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

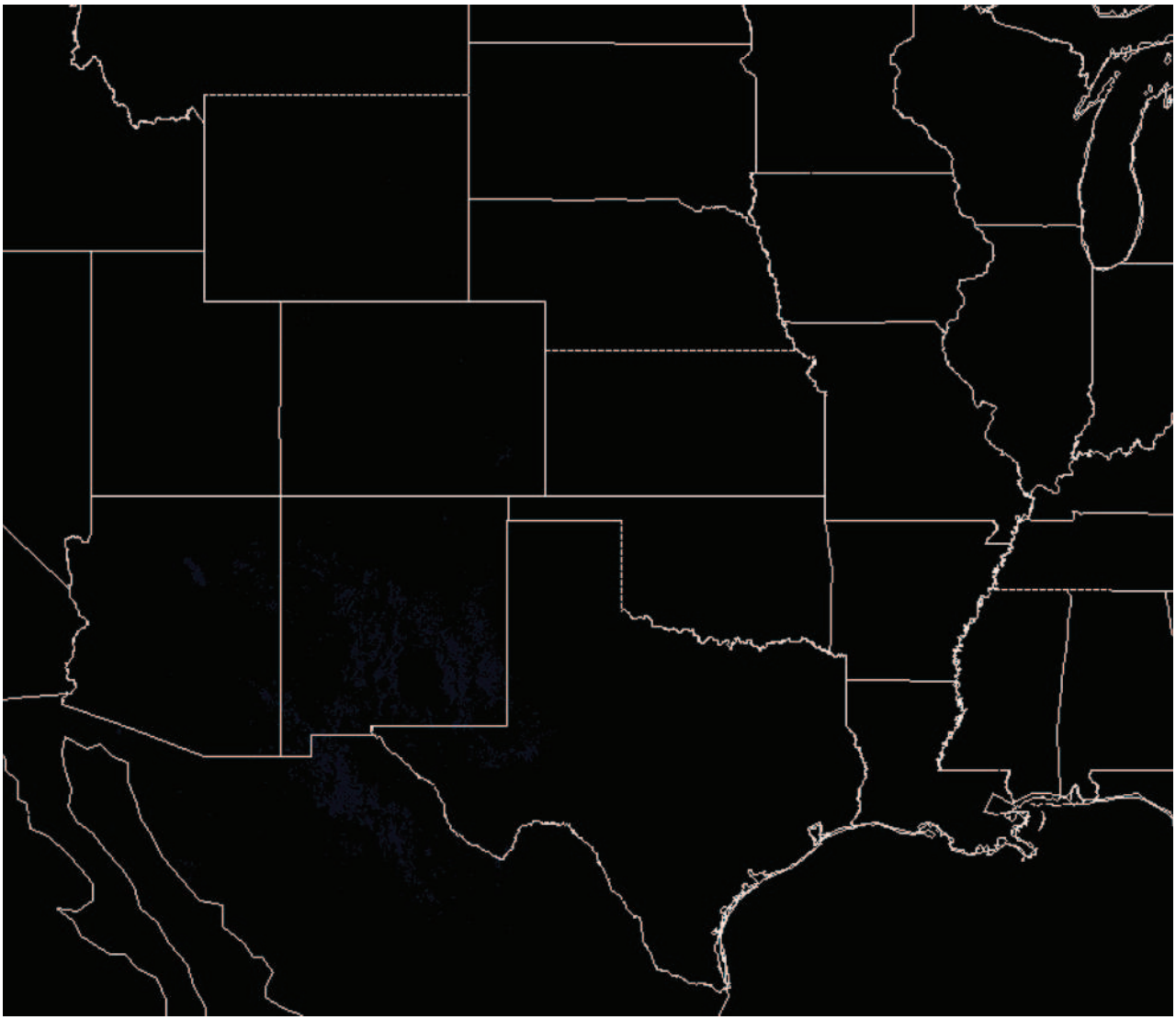

Figure 484. Segmentation result of the Maximum a posteriori model (MAP) for the event of 2003-204 at 17:40 UTC.

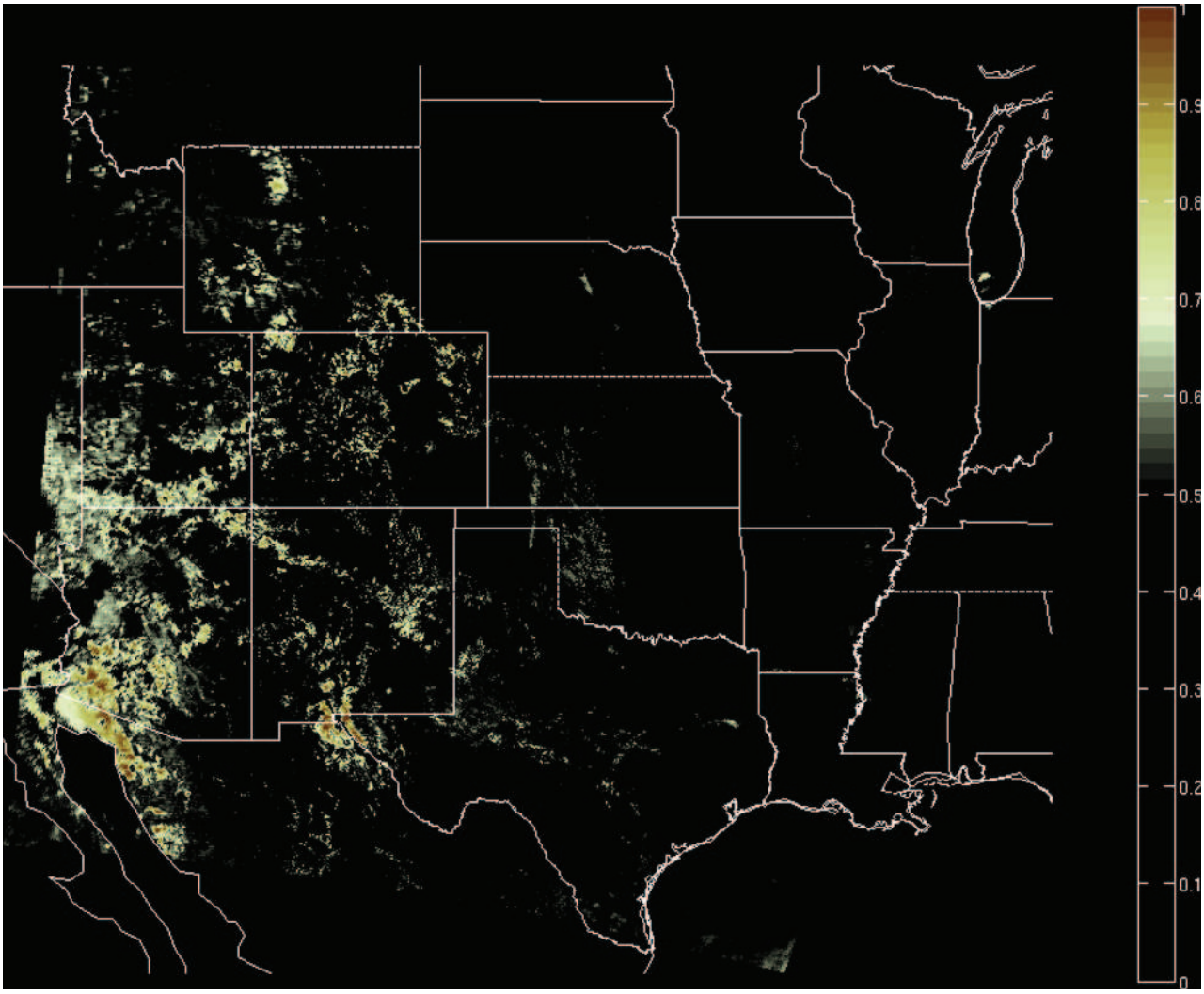

Figure 485. Result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

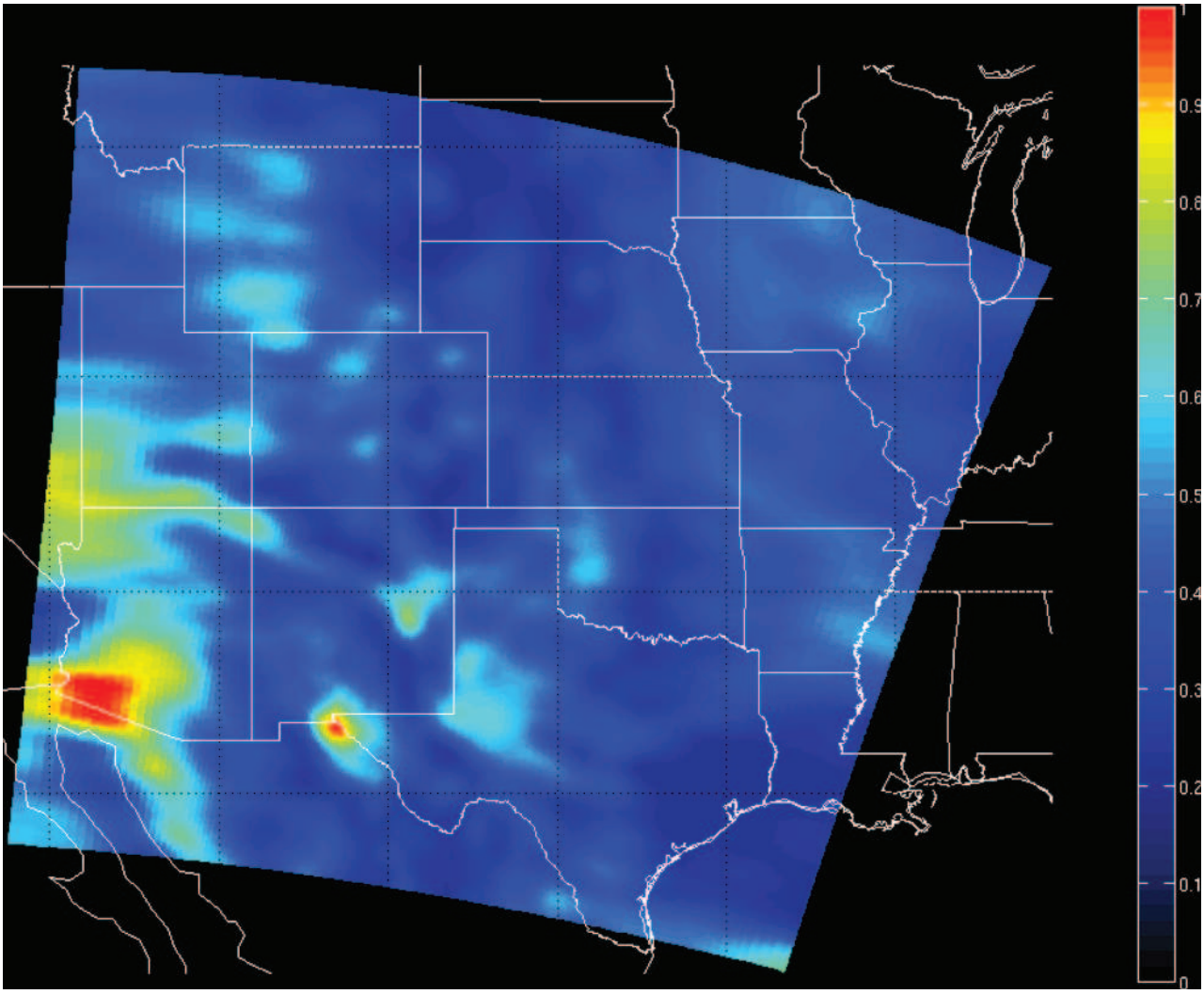

Figure 486. Downsampled and filtered result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

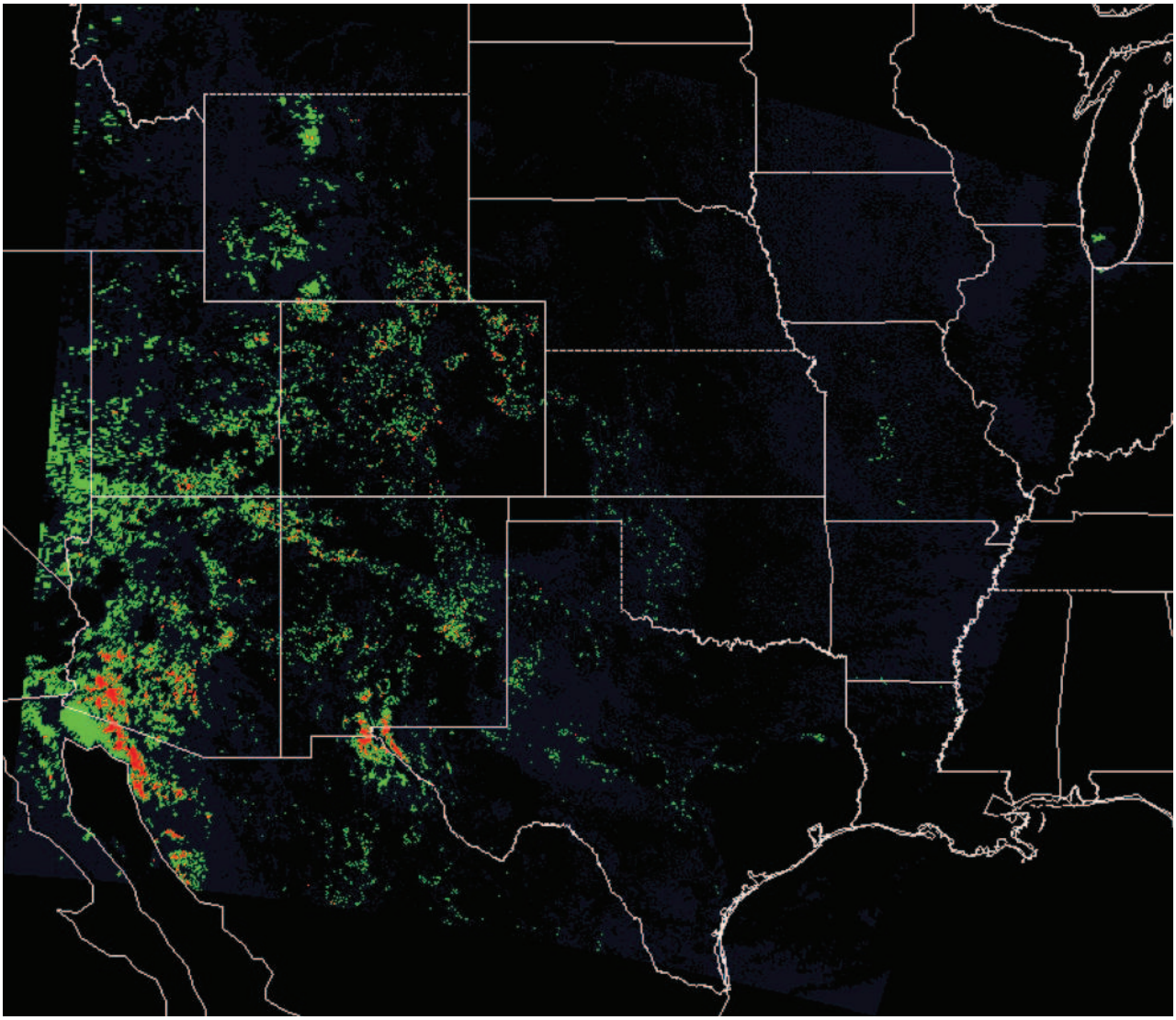

Figure 487. Segmentation result of the Feedforward Backpropagation Neural Network (FFNN) for the event of 2003-204 at 17:40 UTC.

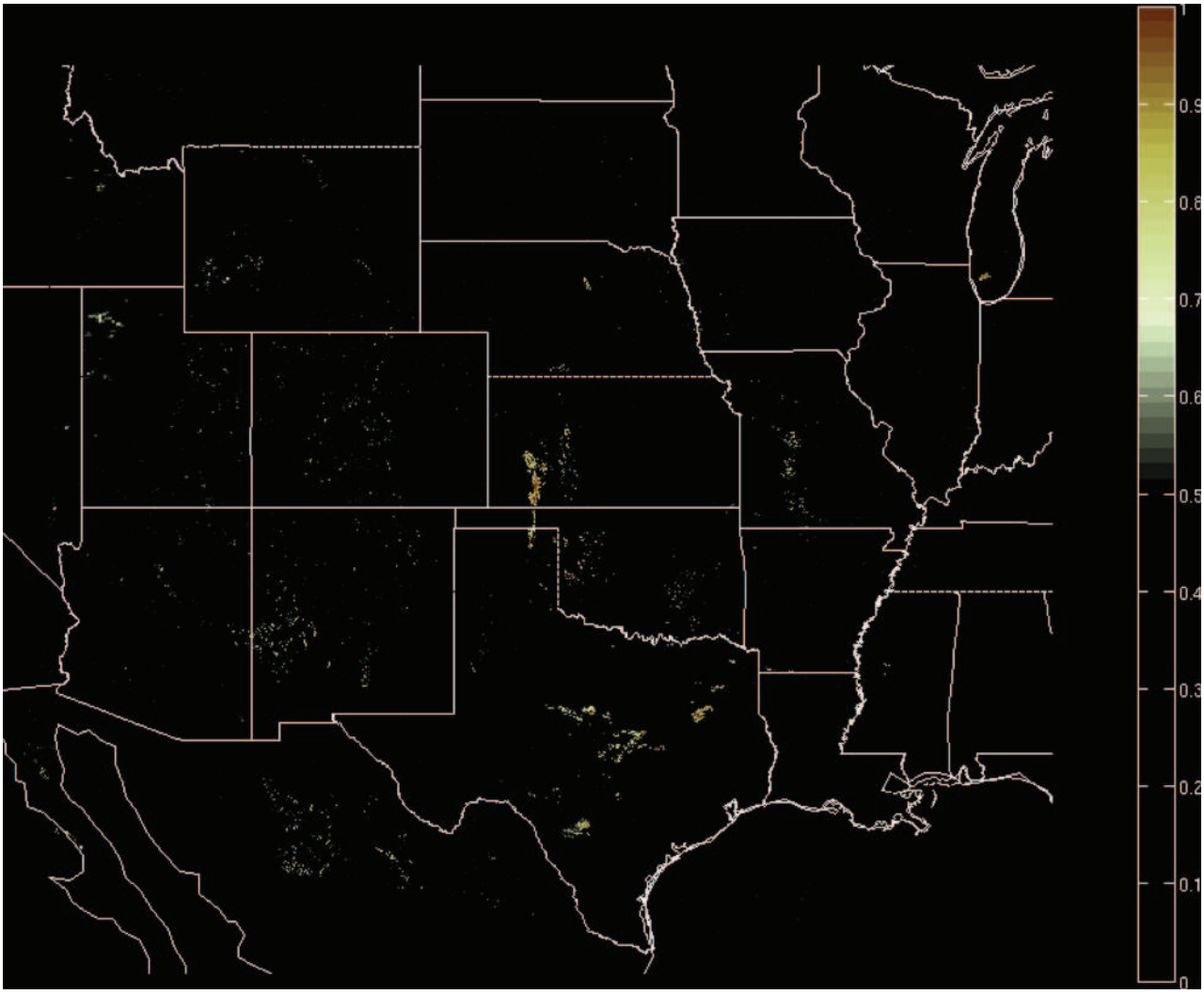

Figure 488. Result of the Probabilistic Neural Network (PNN) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

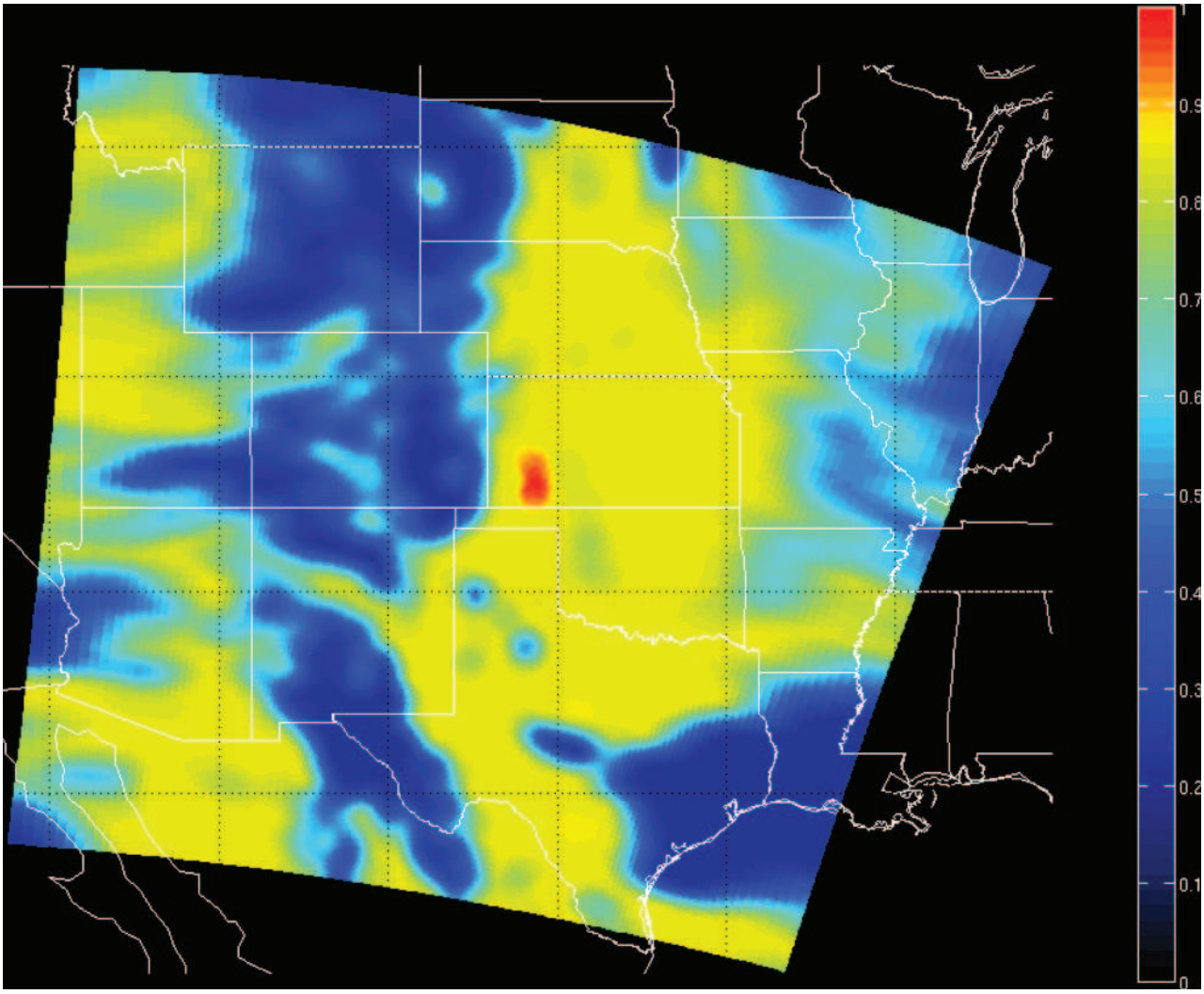

Figure 489. Downsampled and filtered result of the Probabilistic Neural Network (PNN) for the event of 2003-204 at 17:40 UTC. The colorbar shows the index of probability.

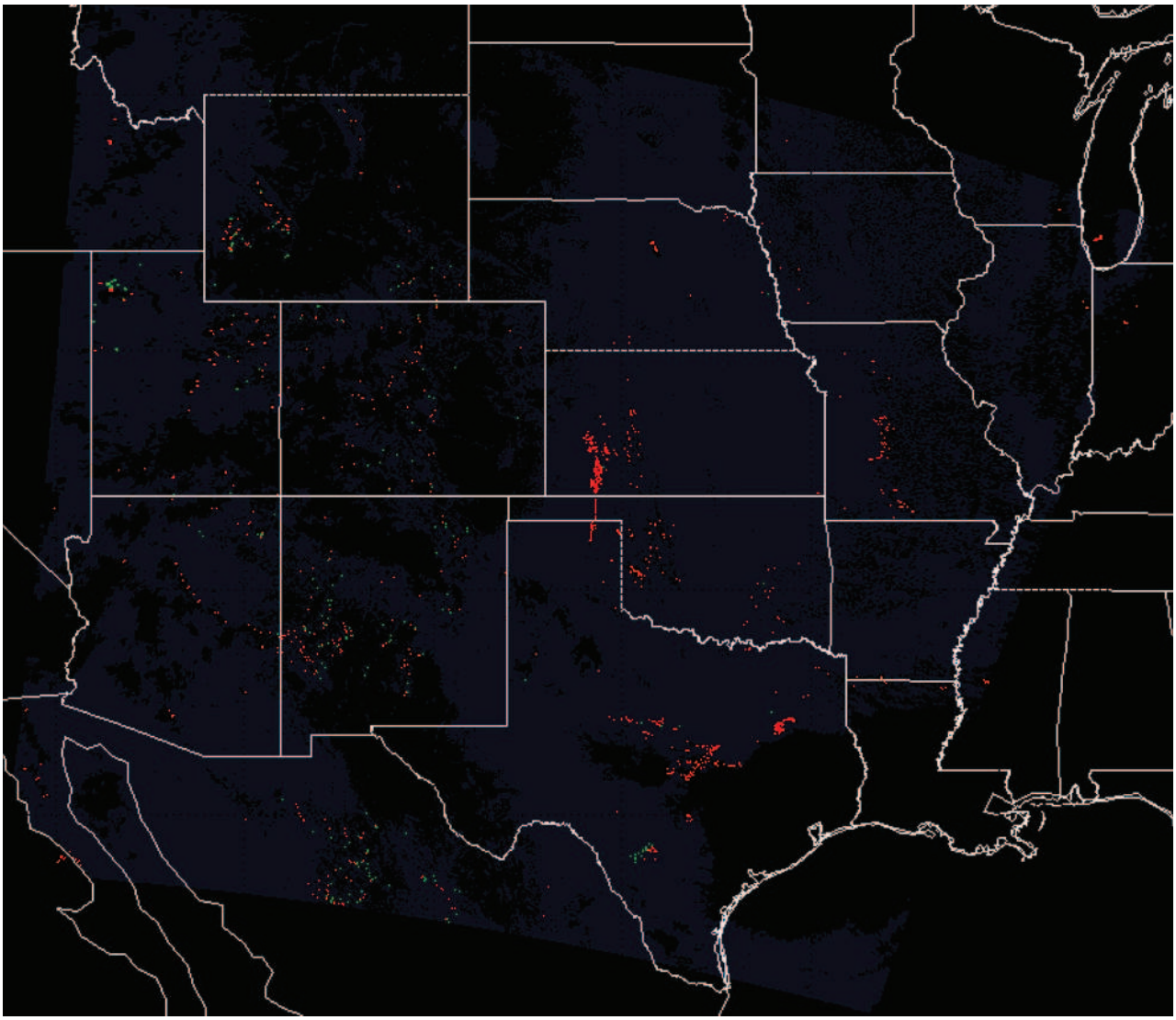

Figure 490. Segmentation result of the Probabilistic Neural Network (PNN) for the event of 2003-204 at 17:40 UTC.

Table 1. List of events collected for Dust Storms (DS), blowing dust ($BLDU$), smoke (SM) and also events with none of this events (background denoted as $C0$). The events listed with (*) are used for testing the algorithms, while the remaining events are used to construct and train the models. The events referenced with [+] are events that (to the best of the knowledge of the author P.R.P.) do not contain dust, therefore, are not reported in previous research work or reports. The latter events were chosen arbitrarily.

Year	Month	Day	Classification	Obs. Time (MST)	Scan Time (UCT)	Reference
2000	05	10	SM	11:50	18:50	[18][16]
2001	04	06	DS	10:56	18:30	[18]
2001	04	10	DS	10:53	18:05	[18]
2001	06	04	$BLDU^*$	0:53	18:10	[18]
2001	06	13	$BLDU$	9:53	18:05	[18]
2001	11	23	$BLDU$	7:53	18:30	[18]
2002	02	09	$BLDU^*$	9:53	18:40	[18]
2002	03	01	$BLDU$	10:53	18:15	[18][10]
2002	03	08	$BLDU$	10:53	18:20	[18][10]
2002	03	14	$BLDU$	2:53	17:45	[18]
2002	04	30	$BLDU$	6:53	18:40	[18]
2002	07	02	DS	17:14	17:55	[18]
2002	12	17	DS	10:52	18:45	[18]
2003	03	04	$BLDU^*$	13:52	18:15	[18][10]
2003	03	18	$BLDU$	11:52	18:25	[18]
2003	03	27	$BLDU$	12:53	18:20	[18][10]
2003	04	15	DS	9:53	17:10	[18][10][6][16]
2003	04	17	$C0^*$	11:40	18:40	[+]
2003	04	23	$BLDU$	1:53	18:00	[18]
2003	07	23	SM	10:40	17:40	[18][6]
2003	11	22	DS	10:53	18:20	[18][10]
2003	12	15	DS	5:17	18:25	[18][10][6]
2004	02	19	DS	7:10	18:15	[18]
2005	04	09	$C0^*$	10:30	17:30	[+]
2005	11	27	DS	0:51	18:20	[18]
2006	03	08	$BLDU$	10:51	18:35	[18]
2006	03	10	$BLDU$	9:51	18:25	[18]
2006	03	12	$BLDU$	7:51	18:10	[18][10]
2006	04	13	$C0^*$	11:10	18:10	[+]
2009	04	04	DS^*	11:15	18:15	[18]
2009	04	07	$C0^*$	11:00	18:00	[+]

Table 2. Collection of granules associated to the list of events described in the previous table to MODIS level 1B files.

Year	Month	Day	Classification	Granule Filename
2000	05	10	SM	MOD021KM.A2000131.1850.005.2007171190825.hdf
2001	04	06	DS	MOD021KM.A2001096.1830.005.2008272181137.hdf
2001	04	10	DS	MOD021KM.A2001100.1805.005.2007006002547.hdf
2001	06	04	$BLDU^*$	MOD021KM.A2001155.1810.005.2007024065302.hdf
2001	06	13	$BLDU$	MOD021KM.A2001164.1805.005.2007027061532.hdf
2001	11	23	$BLDU$	MOD021KM.A2001327.1830.005.2007067003910.hdf
2002	02	09	$BLDU^*$	MOD021KM.A2002040.1840.005.2008290033714.hdf
2002	03	01	$BLDU$	MOD021KM.A2002060.1815.005.2008224212729.hdf
2002	03	08	$BLDU$	MOD021KM.A2002067.1820.005.2008225145618.hdf
2002	03	14	$BLDU$	MOD021KM.A2002073.1745.005.2008226175249.hdf
2002	04	30	$BLDU$	MOD021KM.A2002120.1840.005.2007128042124.hdf
2002	07	02	DS	MOD021KM.A2002183.1755.005.2008238052433.hdf
2002	12	17	DS	MOD021KM.A2002351.1845.005.2007244072834.hdf
2003	03	04	$BLDU^*$	MOD021KM.A2003063.1815.005.2007279011314.hdf
2003	03	18	$BLDU$	MOD021KM.A2003077.1825.005.2008297001154.hdf
				MOD021KM.A2003078.1730.005.2008297223949.hdf
2003	03	27	$BLDU$	MOD021KM.A2003086.1820.005.2008298223751.hdf
2003	04	15	DS	MOD021KM.A2003105.1710.005.2008306182153.hdf
				MOD021KM.A2003105.1715.005.2008306183334.hdf
				MOD021KM.A2003105.1850.005.2008306182854.hdf
2003	04	17	$C0^*$	MOD021KM.A2003107.1840.005.2008306215653.hdf
2003	04	23	$BLDU$	MOD021KM.A2003113.1800.005.2008260222048.hdf
2003	07	23	SM	MOD021KM.A2003204.1740.005.2008311055159.hdf
2003	11	22	DS	MOD021KM.A2003326.1820.005.2008018094422.hdf
2003	12	15	DS	MOD021KM.A2003349.1825.005.2008262193442.hdf
2004	02	19	DS	MOD021KM.A2004050.1815.005.2007241204939.hdf
2005	04	09	$C0^*$	MOD021KM.A2005099.1730.005.2008001182727.hdf
2005	11	27	DS	MOD021KM.A2005331.1820.005.2008260015545.hdf
2006	03	08	$BLDU$	MOD021KM.A2006067.1835.005.2008075230009.hdf
2006	03	10	$BLDU$	MOD021KM.A2006069.1825.005.2008076134942.hdf
2006	03	12	$BLDU$	MOD021KM.A2006071.1810.005.2008077104411.hdf
				MOD021KM.A2006071.1815.005.2008077104359.hdf
2006	04	13	$C0^*$	MOD021KM.A2006103.1810.005.2008088093854.hdf
2009	04	04	$DS*$	MOD021KM.A2009093.1825.005.2009094023122.hdf
2009	04	07	$C0^*$	MOD021KM.A2009097.1800.005.2009098020431.hdf

Table 3. Error measurements for the original cell image vs cell degraded

Performance:	NMSE	SNR	PSNR	WPSNR
Non-restored synthetic	4.37	31.29	24.35	30.50
Restored synthetic	0.91	46.70	35.89	44.64
Non-restored real data	10.96	22.10	23.37	29.86
Restored real data	0.93	46.69	34.59	41.41

Table 4. General details of the training methodology for each classification method.

	PD	MLE	MAP	FFNN	PNN
Features	$B32-31$	$F^{(1)}\dots F^{(4)}$	$F^{(1)}\dots F^{(4)}$	$\widetilde{F}^{(1)}\dots\widetilde{F}^{(8)}$	$\widetilde{\widehat{F}}^{(1)}\quad\dots\widetilde{\widehat{F}}^{(4)}$
Samples	All	All	All	$i=m\times q$	$i=m\times q$
Preprocessing	No	No	No	No	$\widehat{F}^{(l)}=\frac{F^{(l)}-\mu_{F^{(l)}}}{\sigma_{F^{(l)}}}$
Architecture	NA	NA	NA	$\begin{bmatrix} m_{F^{(l)}}, \\ int\left(\frac{\sqrt{m_{F^{(l)}}}}{8}\right), \\ int\left(\frac{\sqrt[3]{m_{F^{(l)}}}}{8}\right), \\ 1 \end{bmatrix}$	$\begin{bmatrix} m_{F^{(l)}}, \\ i\times j, \\ j \\ 1 \end{bmatrix}$
Classify	Visual	$C0, SM,$ $BLDU, DS$	$C0, SM,$ $BLDU, DS$	$C0, SM,$ $BLDU, DS$	$C0, SM,$ $BLDU, DS$

Table 5. General details of the testing methodology for each classification method.

	PD	MLE	MAP	FFNN	PNN
Features	$B32-31$	$F^{(1)}\dots F^{(4)}$	$F^{(1)}\dots F^{(4)}$	$F^{(1)}\dots F^{(8)}$	$\widehat{F}^{(1)}\dots\widehat{F}^{(4)}$
Samples	All	All	All	All	All
Preprocessing	No	No	No	No	$\widehat{F}^{(l)}=\frac{F^{(l)}-\mu_{F^{(l)}}}{\sigma_{F^{(l)}}}$
Classify (1)	Visual	$C0, SM,$ $BLDU, DS$	$C0, SM,$ $BLDU, DS$	$C0, SM,$ $BLDU, DS$	$C0, SM,$ $BLDU, DS$
Classify (2)	Visual	$C0, DS$	$C0, DS$	$C0, DS$	$C0, DS$

Table 6. Results obtained for the different models for dust storms during full testing. Results are totalized by event, then they were averaged and displayed here.

	MLE	MAP	FFNN	PNN
fp rate	0.4688	0.4160	0.4075	0.1816*
tp rate	0.3138	0.3507	0.4554	0.6838*
precision	0.3138	0.3507	0.4554	0.6838*
accuracy	0.4464	0.4951	0.5426	0.7706*
recall	0.3138	0.3507	0.4554	0.6838*
$F-$measure	0.3138	0.3507	0.4554	0.6838*
sensitivity	0.3138	0.3507	0.4554	0.6838*
specificity	0.5312	0.5840	0.5925	0.8184*
AUROC	0.4893	0.4945	0.7402*	0.6477

Table 7. Results obtained for the different models on only two classes: dust and background. Results are totalized by event, then they were averaged and displayed here.

	MLE	MAP	FFNN	PNN
fp rate	0.2446	0.2202	0.0864*	0.1215
tp rate	0.5255	0.5562	0.8080*	0.7664
precision	0.5255	0.5562	0.8080*	0.7664
accuracy	0.6779	0.7061	0.8816*	0.8412
recall	0.5255	0.5562	0.8080*	0.7664
$F-$measure	0.5255	0.5562	0.8080*	0.7664
sensitivity	0.5255	0.5562	0.8080*	0.7664
specificity	0.7554	0.7798	0.9136*	0.8785
AUROC	0.4884	0.4925	0.7117*	0.6293

Table 8. Averaged area under ROC curve (AUC). Feature vectors vs. Classification methods.

	MLE	MAP	FFNN	PNN
$F^{(1)}$	0.4678	0.4926	0.9332*	0.5166
$F^{(2)}$	0.5113*	0.5027	0.4891	0.5000
$F^{(3)}$	0.4934	0.4978	0.8224	0.8253*
$F^{(4)}$	0.4849	0.4849	0.8145*	0.7526
$F^{(5)}$	0.5001	0.5071	0.7041*	0.5047
$F^{(6)}$	0.4921	0.4991	0.5664	0.5978*
$F^{(7)}$	0.4915	0.4913	0.8822*	0.8259
$F^{(8)}$	0.4848	0.4904	0.7096*	0.6583

Table 9. Averaged area under ROC curve (AUC) for two 2 classes. Feature vectors vs. Classification methods.

	MLE	MAP	FFNN	PNN
$F^{(1)}$	0.4683	0.4873	0.8838*	0.5166
$F^{(2)}$	0.5086*	0.5024	0.4953	0.5000
$F^{(3)}$	0.4953	0.4989	0.8400*	0.7907
$F^{(4)}$	0.4816	0.4816	0.7364*	0.6966
$F^{(5)}$	0.4936	0.5000	0.6990*	0.5067
$F^{(6)}$	0.4827	0.4912	0.5060	0.5972*
$F^{(7)}$	0.4936	0.4937	0.8645*	0.7907
$F^{(8)}$	0.4851	0.4880	0.6685*	0.6358

Table 10. Average processing time (in seconds) per each event: Classification method vs. Feature vector.

	MLE	MAP	FFNN	PNN
$F^{(1)}$	53.01	101.10	135.60	1744.49
$F^{(2)}$	49.01	92.37	136.03	2222.61
$F^{(3)}$	52.03	96.97	130.83	1277.77
$F^{(4)}$	38.83*	74.18*	129.80	1272.33
$F^{(5)}$	46.55	89.75	126.35	657.62*
$F^{(6)}$	47.68	91.09	128.76	831.31
$F^{(7)}$	47.61	92.45	129.85	809.18
$F^{(8)}$	46.83	90.54	126.17*	844.48

Table 11. Average processing time per pixel (in mili-seconds): Classification method vs. Feature vector.

	MLE	MAP	FFNN	PNN
$F^{(1)}$	0.0193	0.0368	0.0493	0.6347
$F^{(2)}$	0.0178	0.0336	0.0495	0.8086
$F^{(3)}$	0.0189	0.0353	0.0476	0.4649
$F^{(4)}$	0.0141*	0.0270*	0.0472	0.4629
$F^{(5)}$	0.0169	0.0327	0.0460	0.2393*
$F^{(6)}$	0.0173	0.0331	0.0468	0.3024
$F^{(7)}$	0.0173	0.0336	0.0472	0.2944
$F^{(8)}$	0.0170	0.0329	0.0459*	0.3072

Table 12. Average processing time (in seconds) per multispectral scan: Classification method vs. Feature vector.

	MLE	MAP	FFNN	PNN
$F^{(1)}$	0.0261	0.0498	0.0668	0.8594
$F^{(2)}$	0.0241	0.0455	0.0670	1.0949
$F^{(3)}$	0.0256	0.0478	0.0644	0.6294
$F^{(4)}$	0.0191*	0.0365*	0.0639	0.6268
$F^{(5)}$	0.0229	0.0442	0.0622*	0.3239*
$F^{(6)}$	0.0235	0.0449	0.0634	0.4095
$F^{(7)}$	0.0235	0.0455	0.0640	0.3986
$F^{(8)}$	0.0231	0.0446	0.0622*	0.4160

APPENDIX B. TRUE COLOR EVENT IMAGES WITH ANNOTATIONS

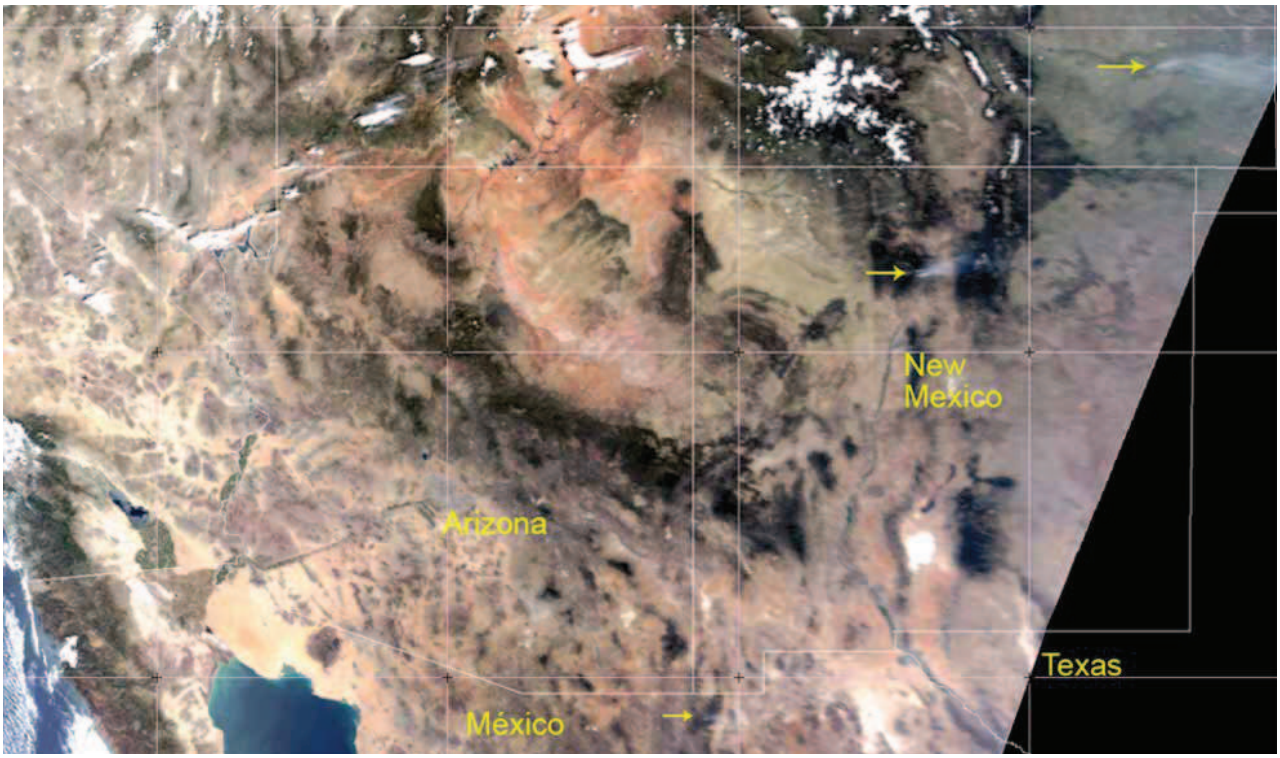

Figure 491. a2000131.1850.png

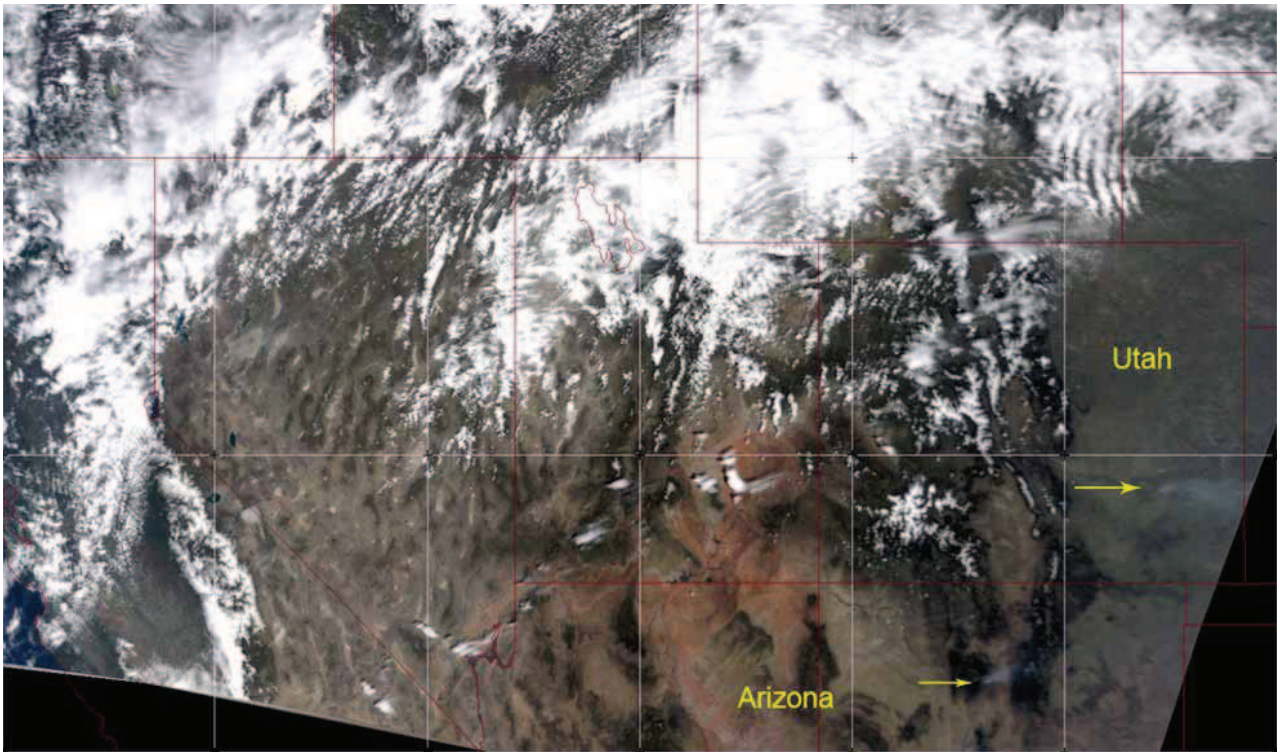

Figure 492. a2003204.1740.png

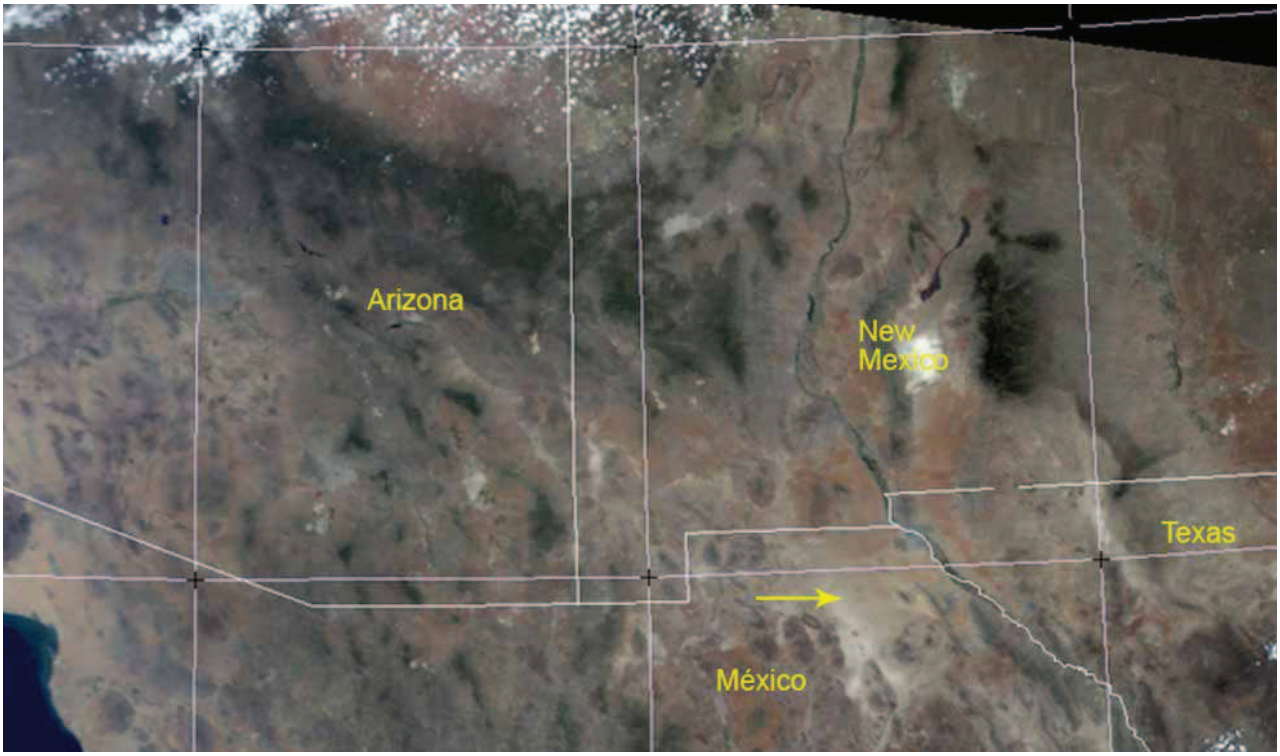

Figure 493. a2001164.1805.png

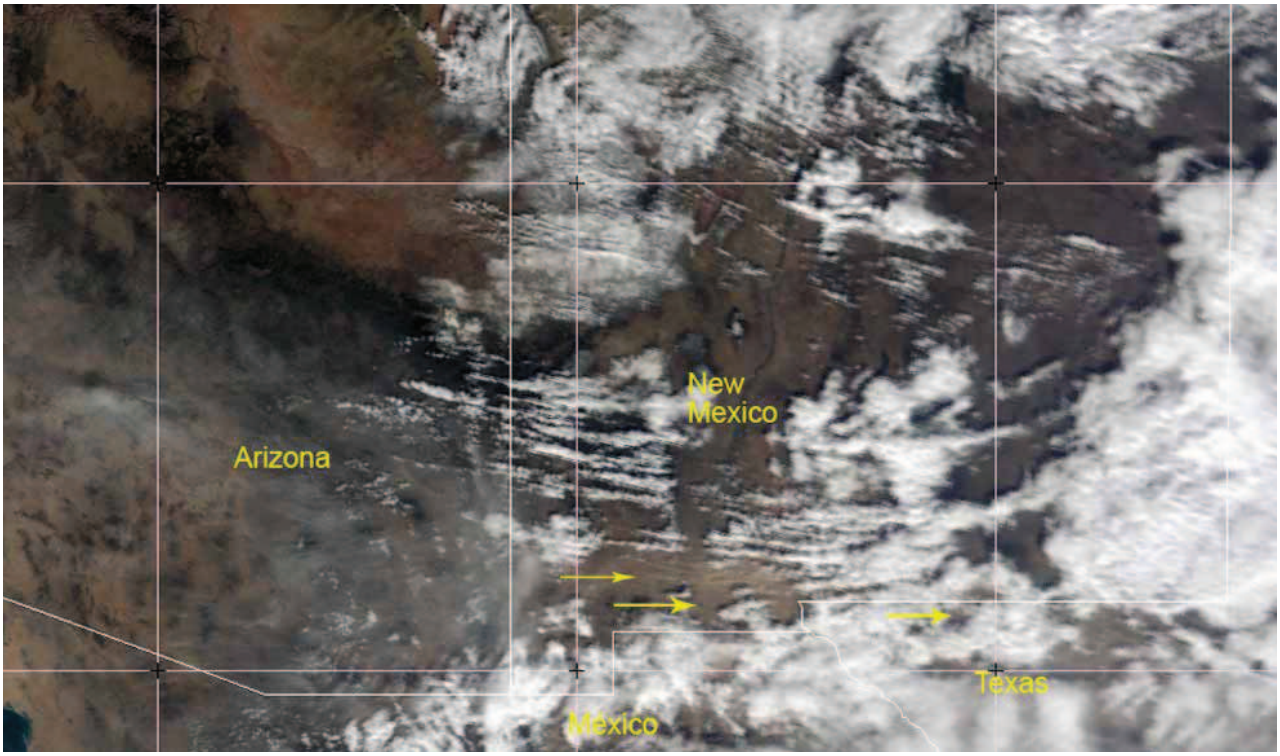

Figure 494. a2001327.1830.png

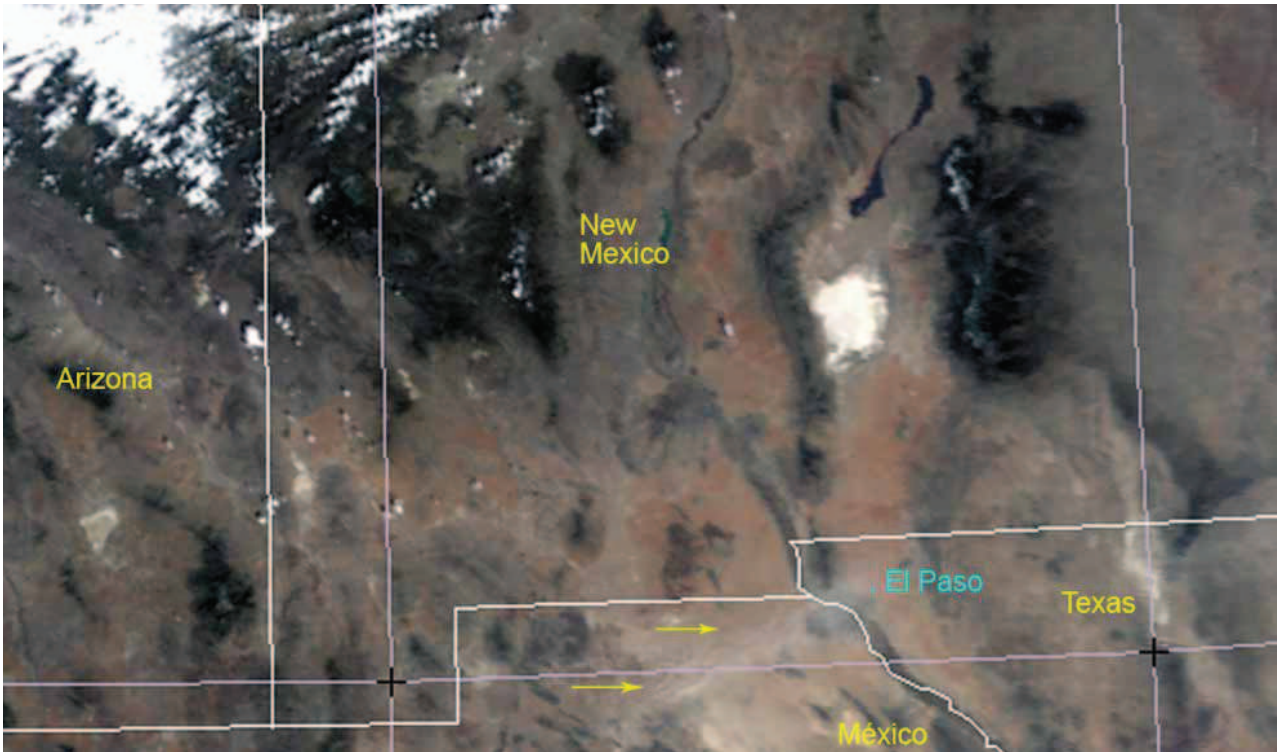

Figure 495. a2002060.1815.png

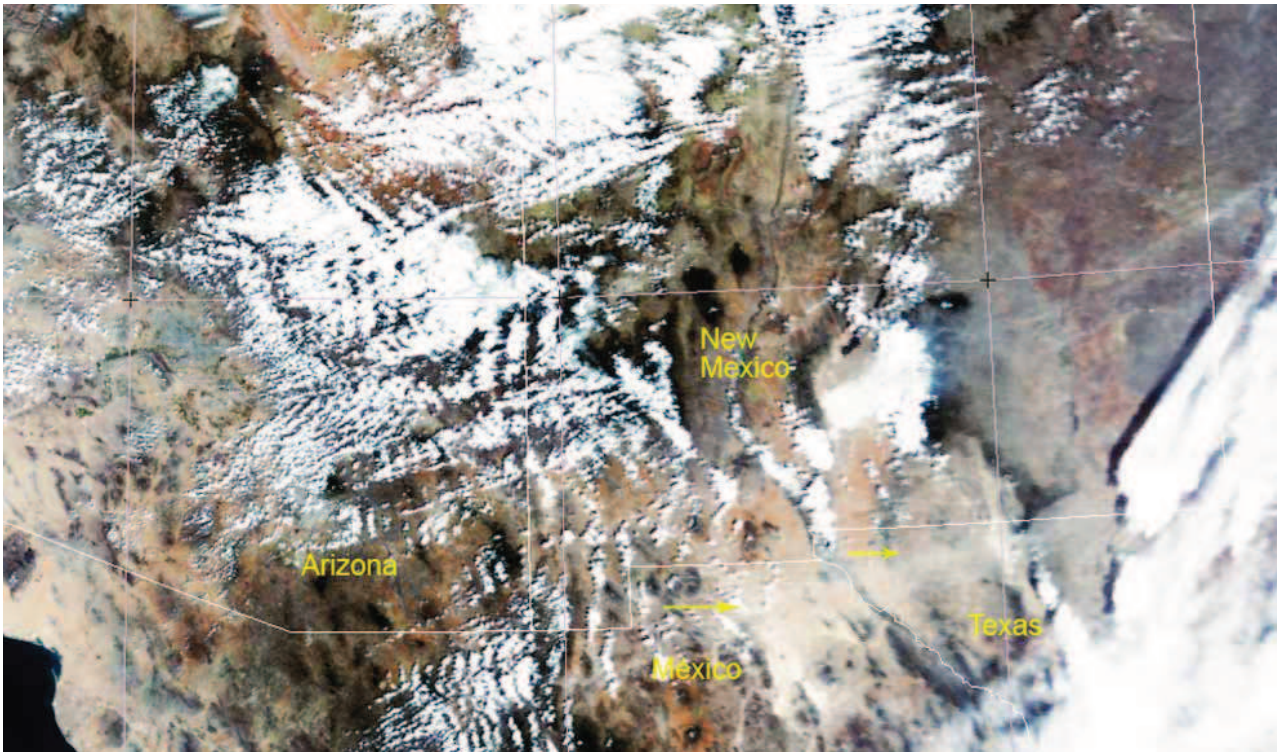

Figure 496. a2002067.1820.png

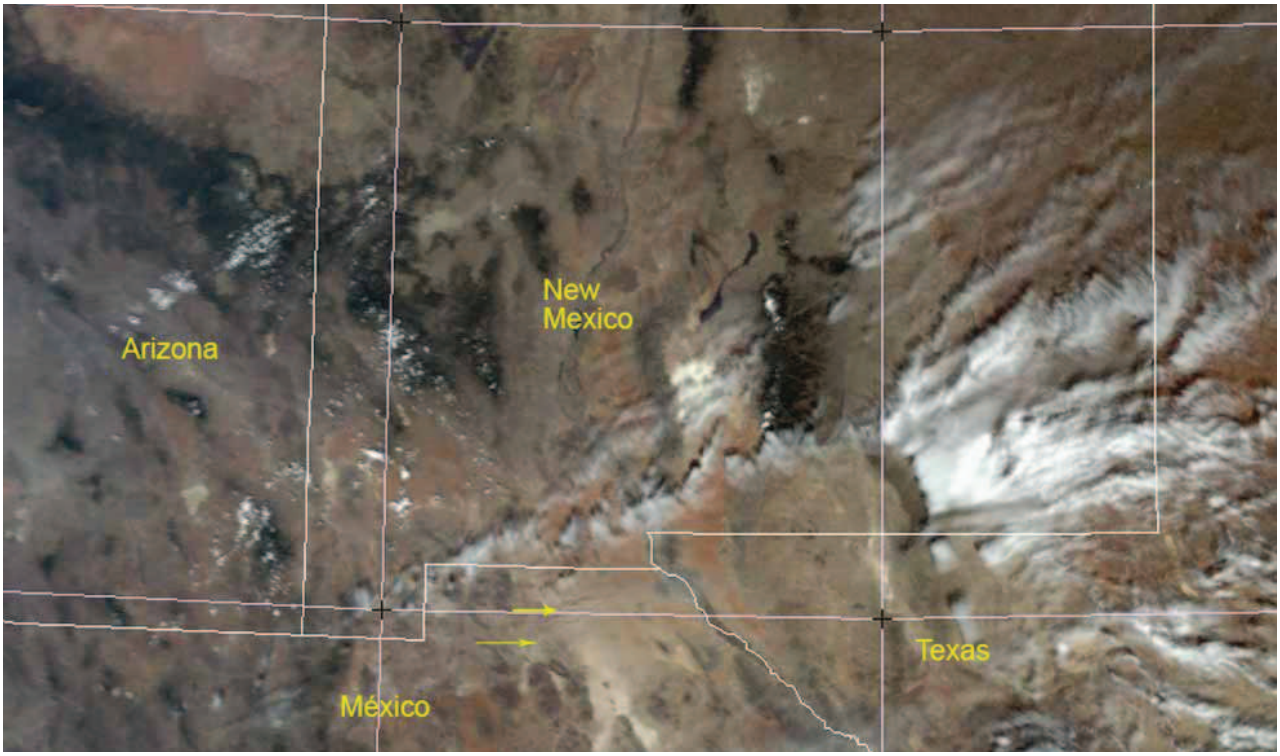

Figure 497. a2002073.1745.png

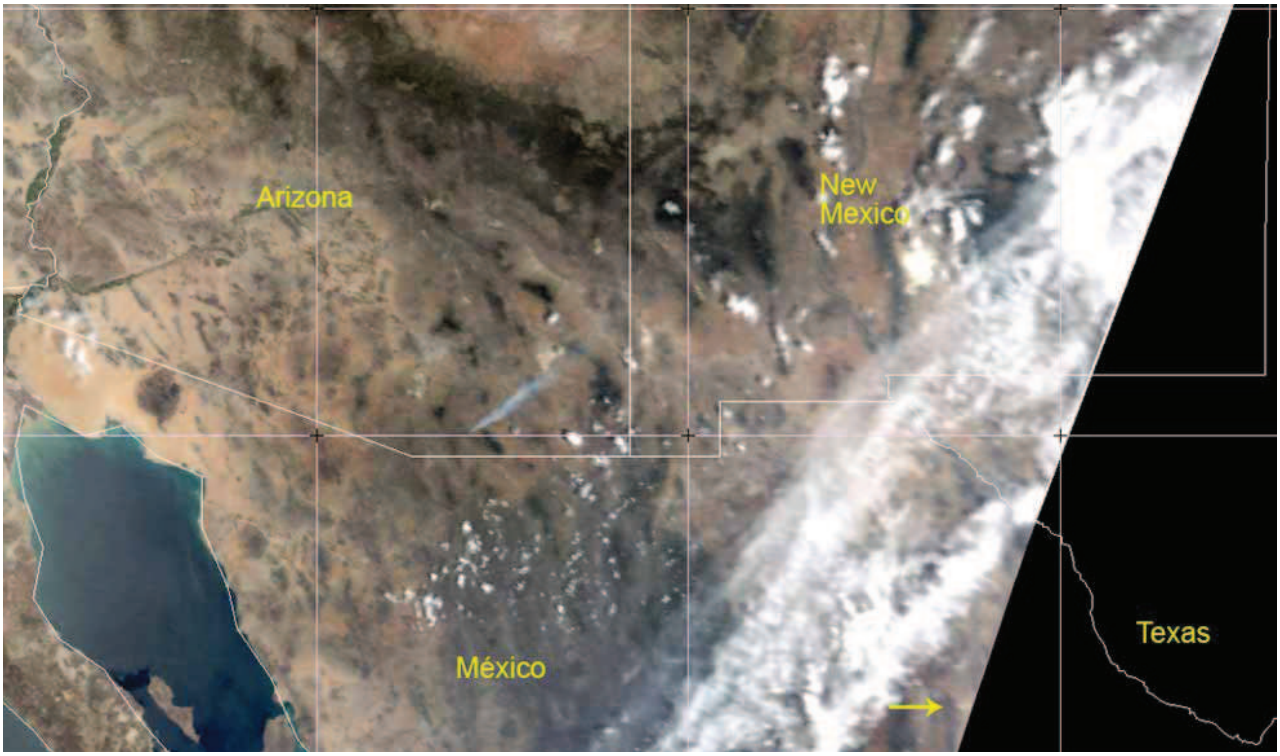

Figure 498. a2002120.1840.png

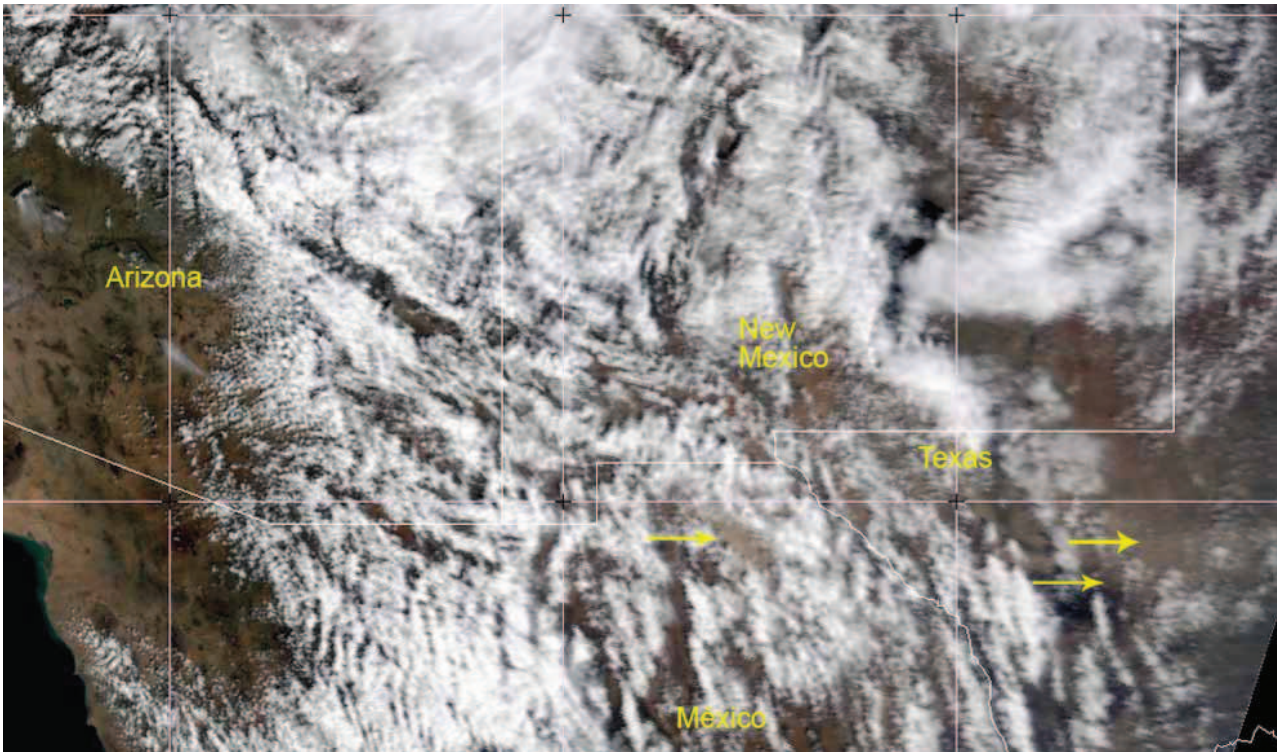

Figure 499. a2003077.1825.png

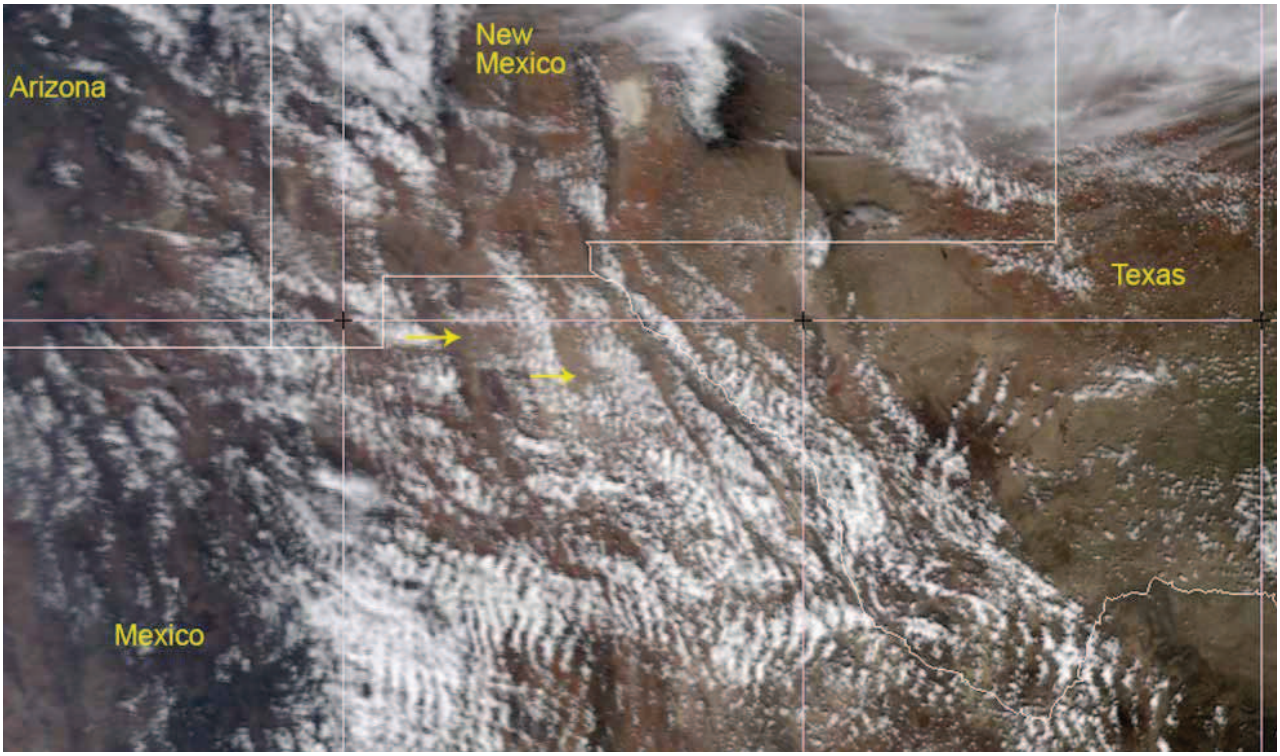

Figure 500. a2003078.1730.png

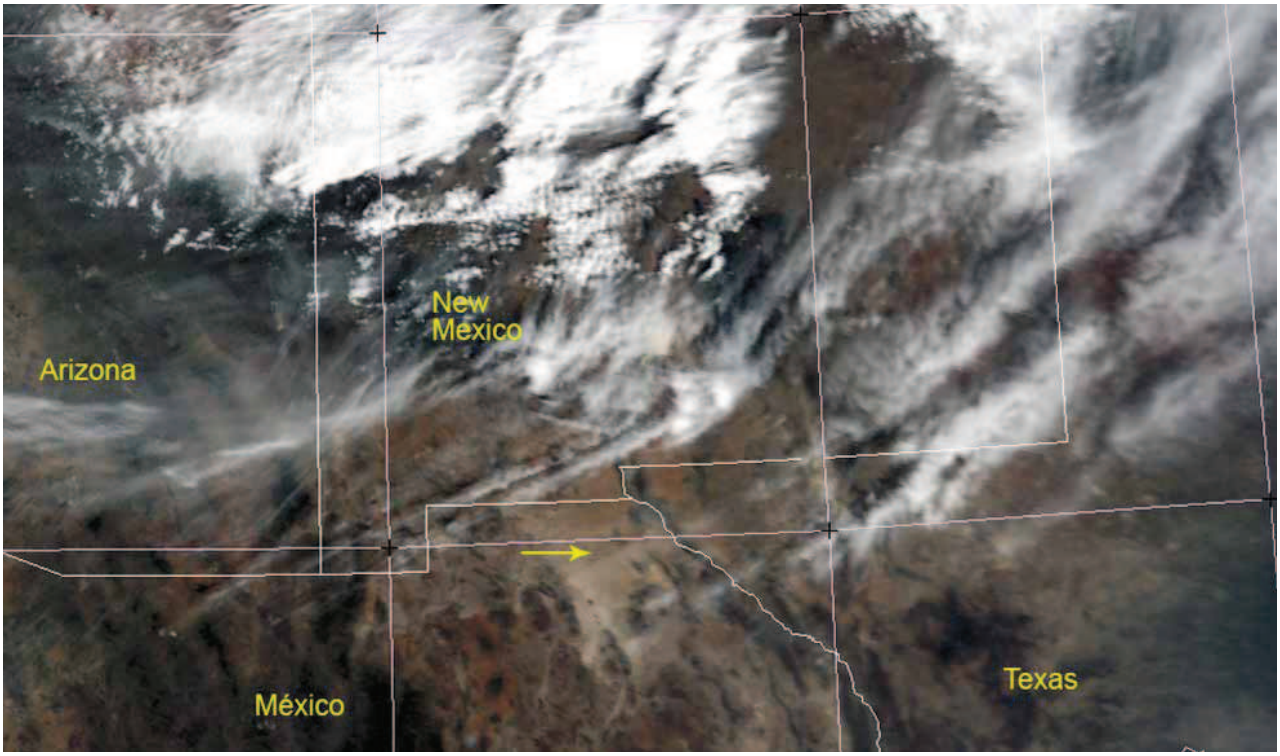

Figure 501. a2003086.1820.png

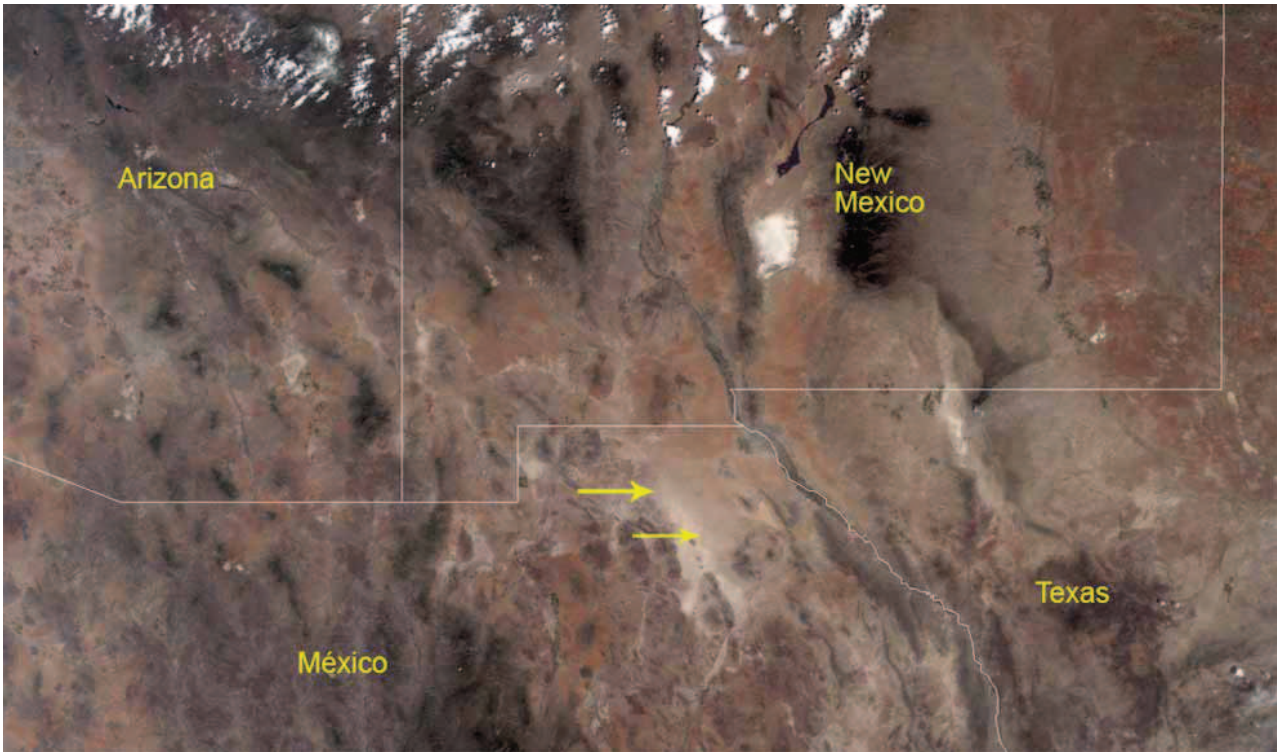

Figure 502. a2003113.1800.png

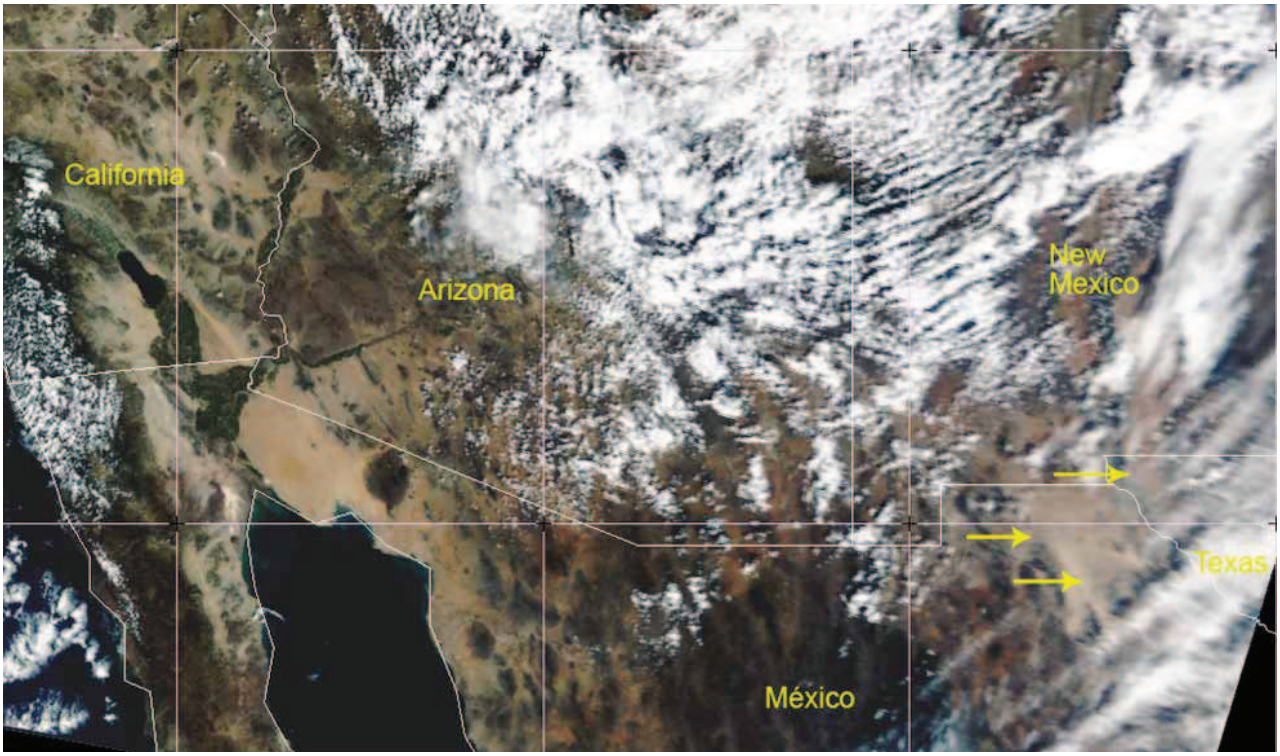

Figure 503. a2006067.1835.png

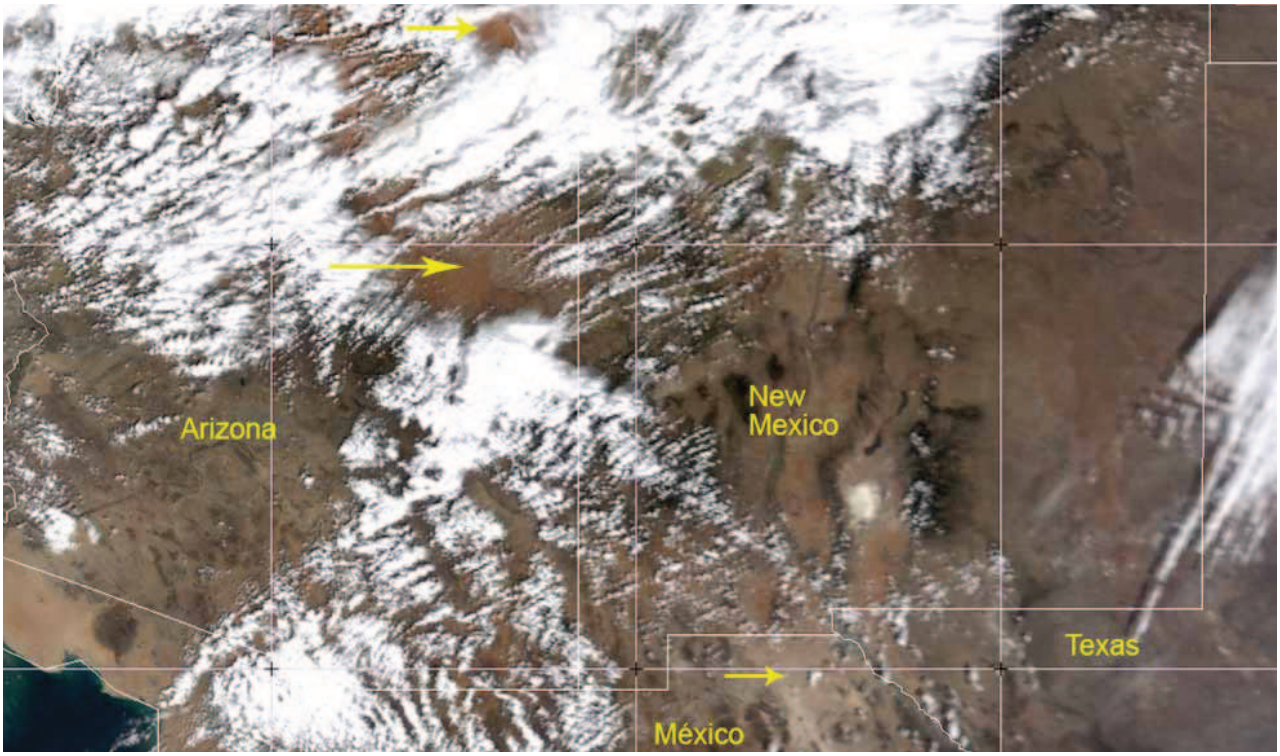

Figure 504. a2006069.1825.png

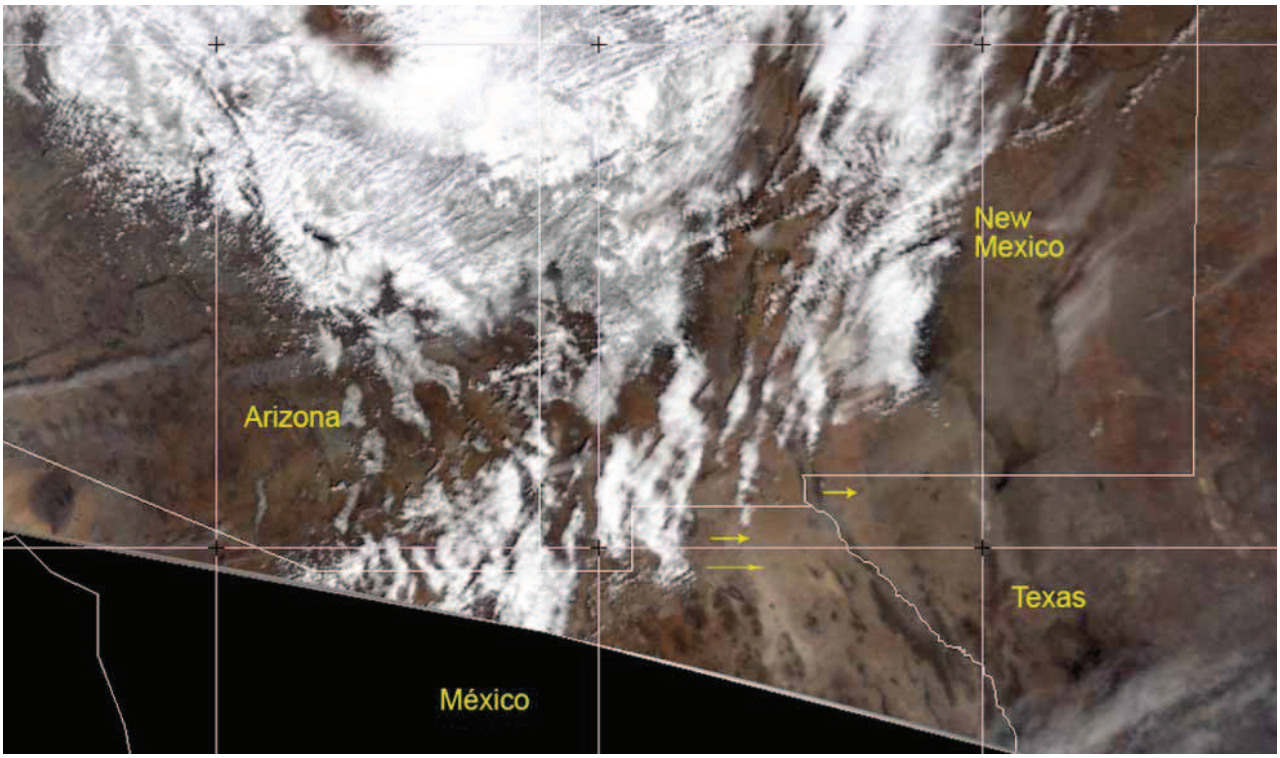

Figure 505. a2006071_1810.png

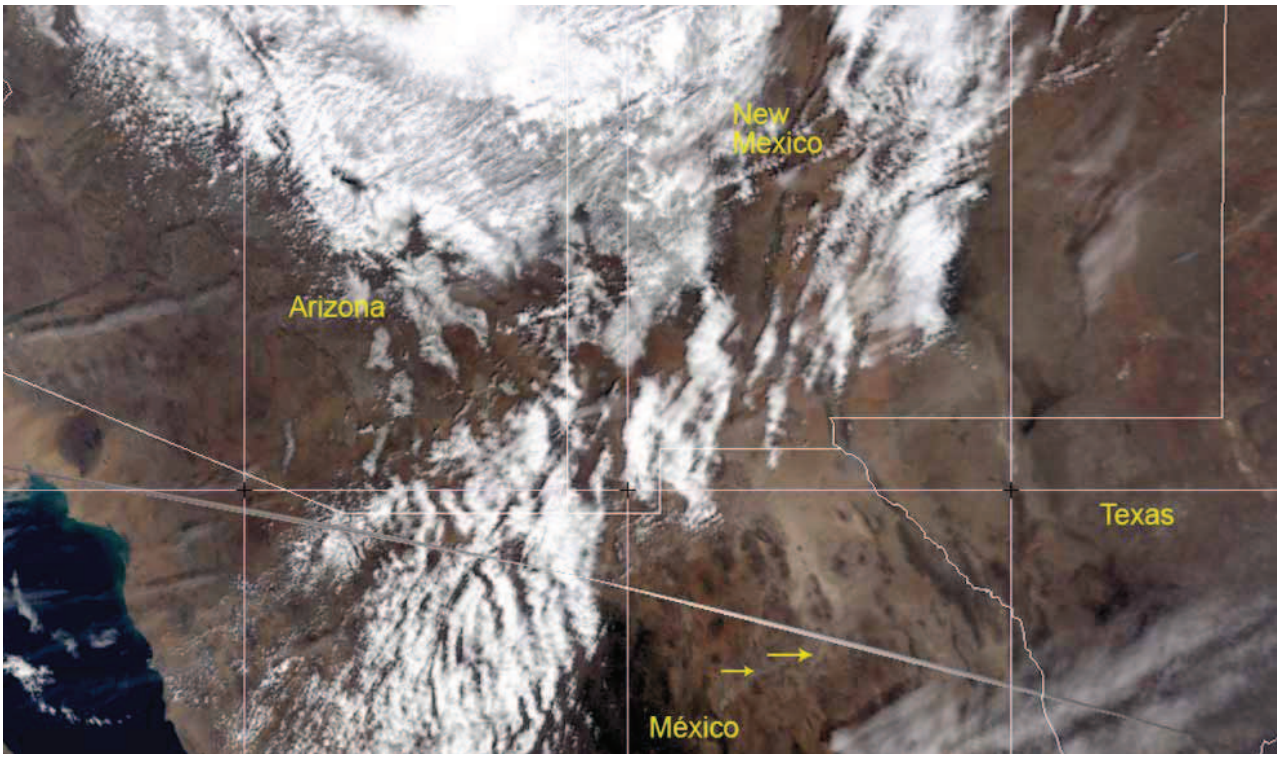

Figure 506. a2006071_1815.png

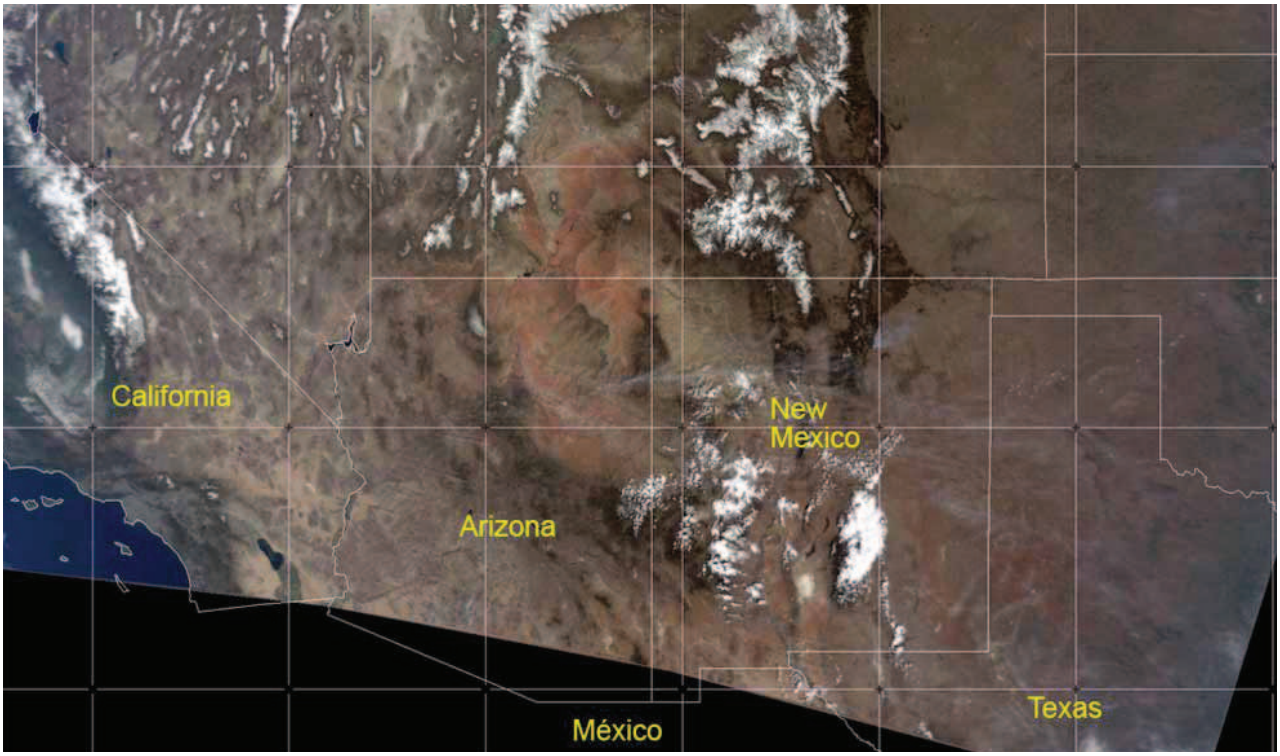

Figure 507. a2006103.1810.png

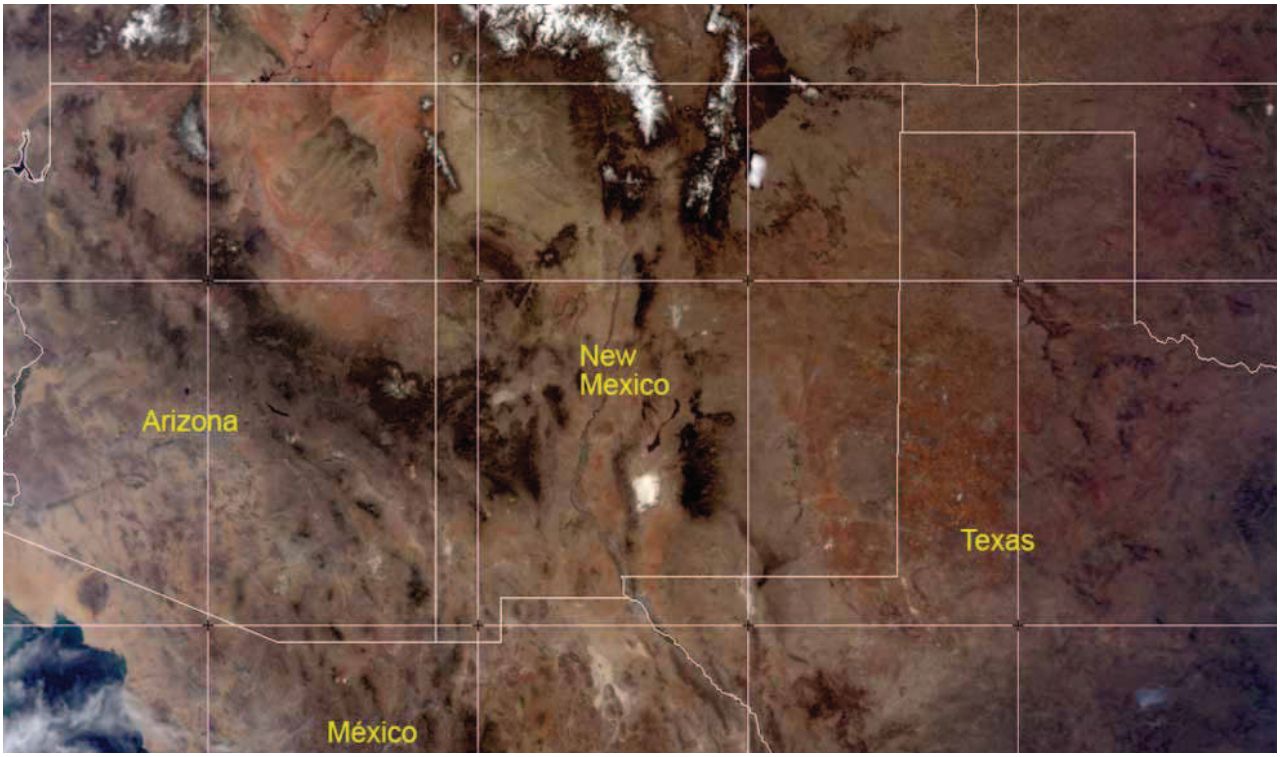

Figure 508. a2009097_1800.png

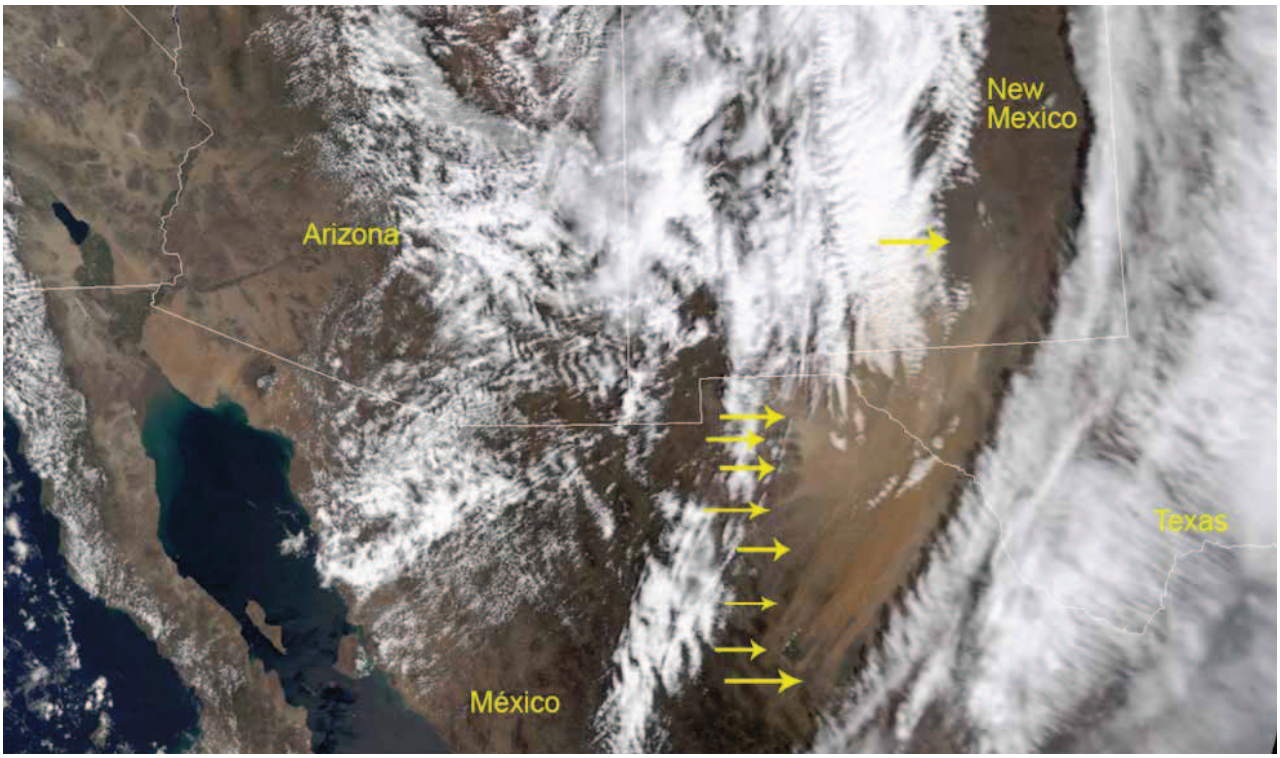

Figure 509. a2001096_1830.png

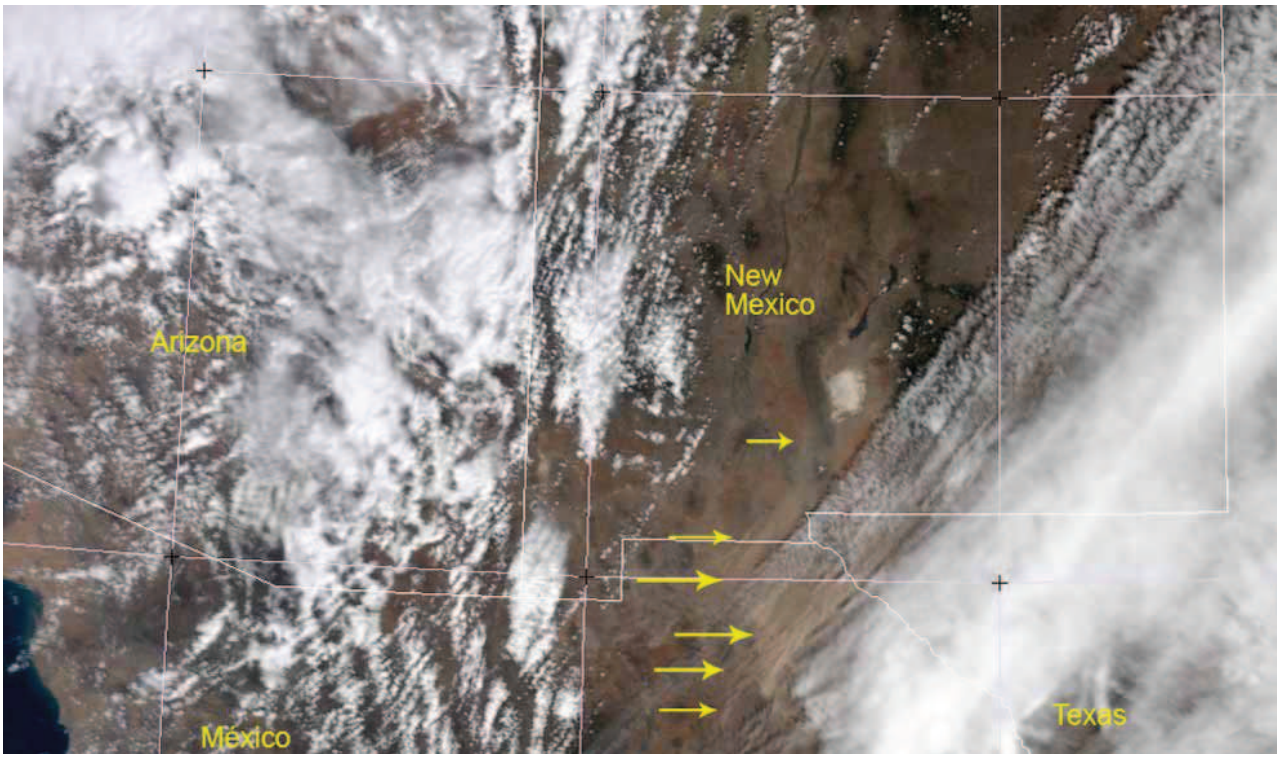

Figure 510. a2001100_1805.png

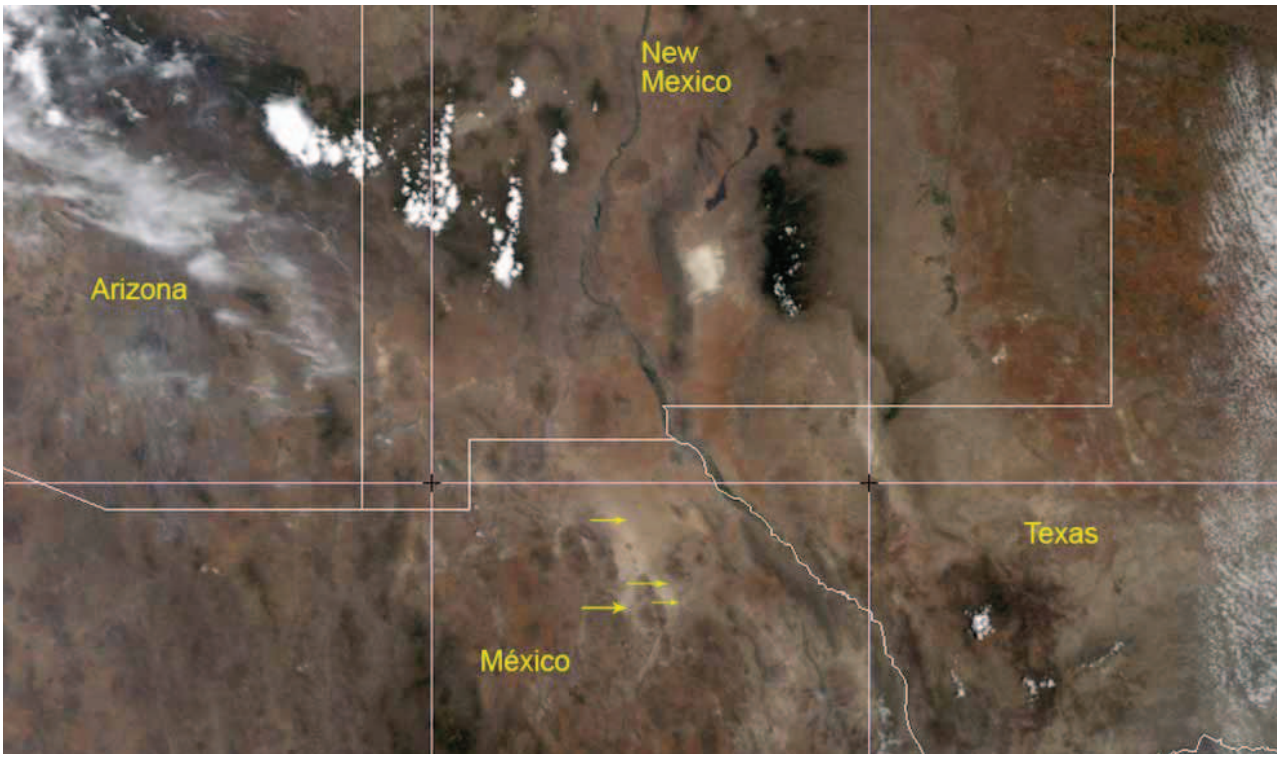

Figure 511. a2002183_1755.png

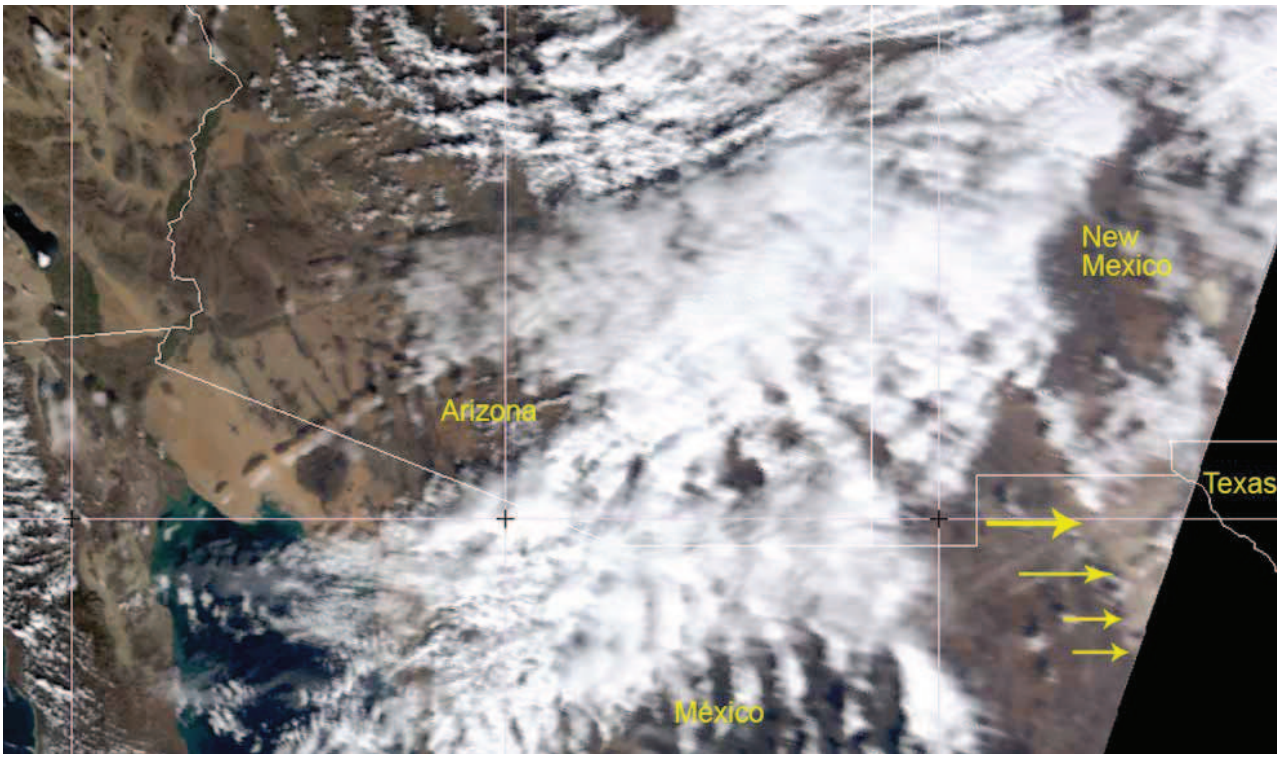

Figure 512. a2002351_1845.png

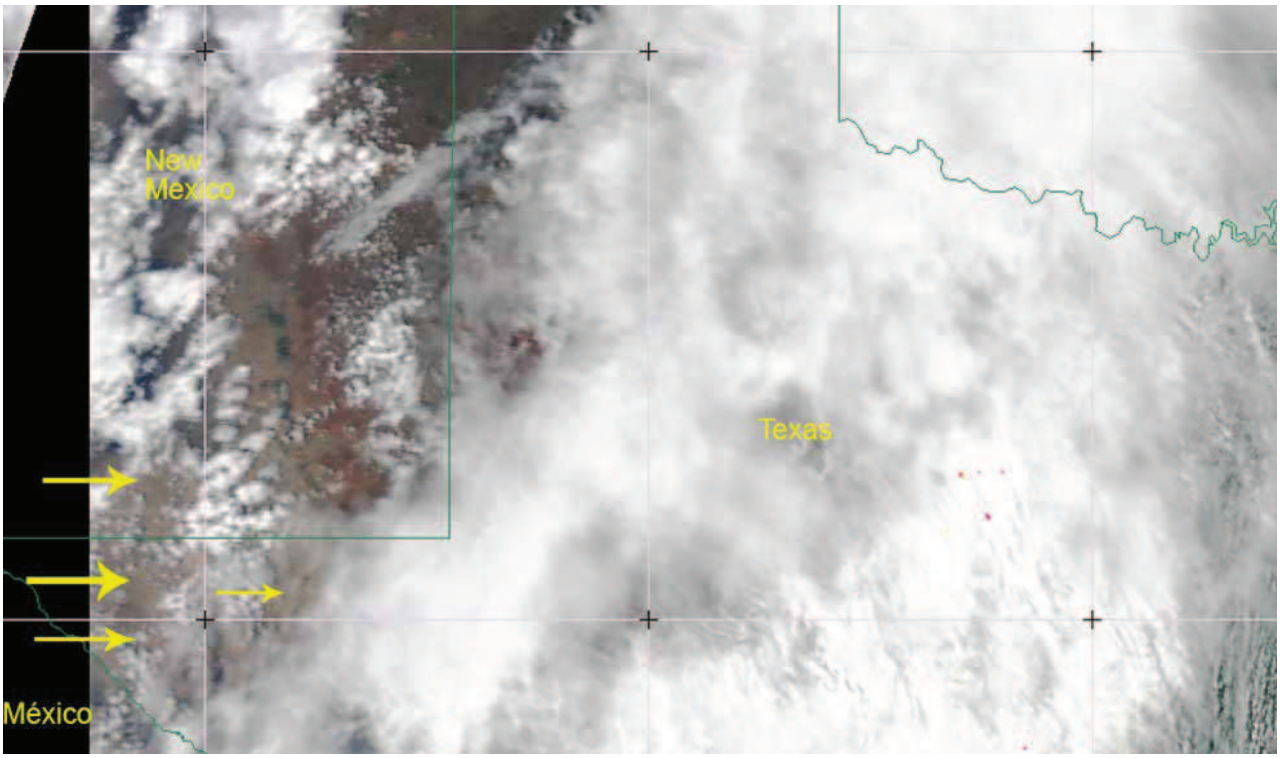

Figure 513. a2003105_1710.png

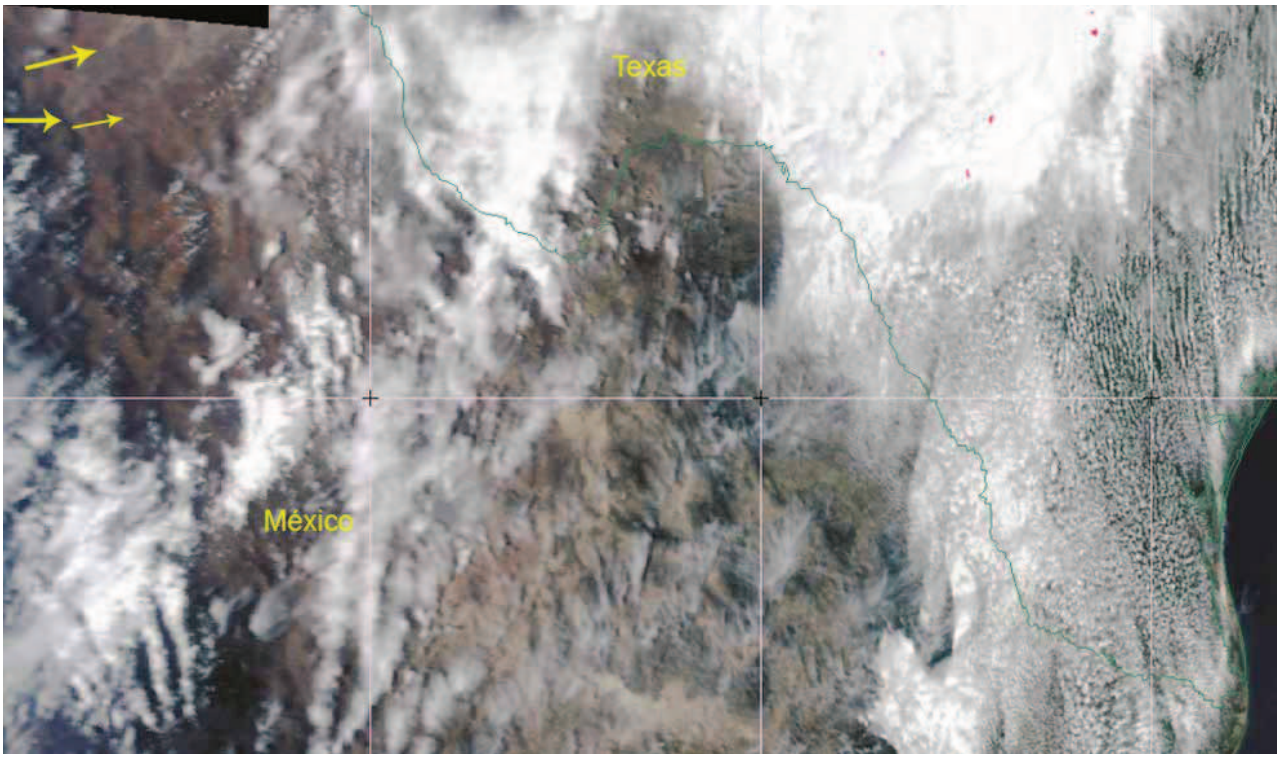

Figure 514. a2003105_1715.png

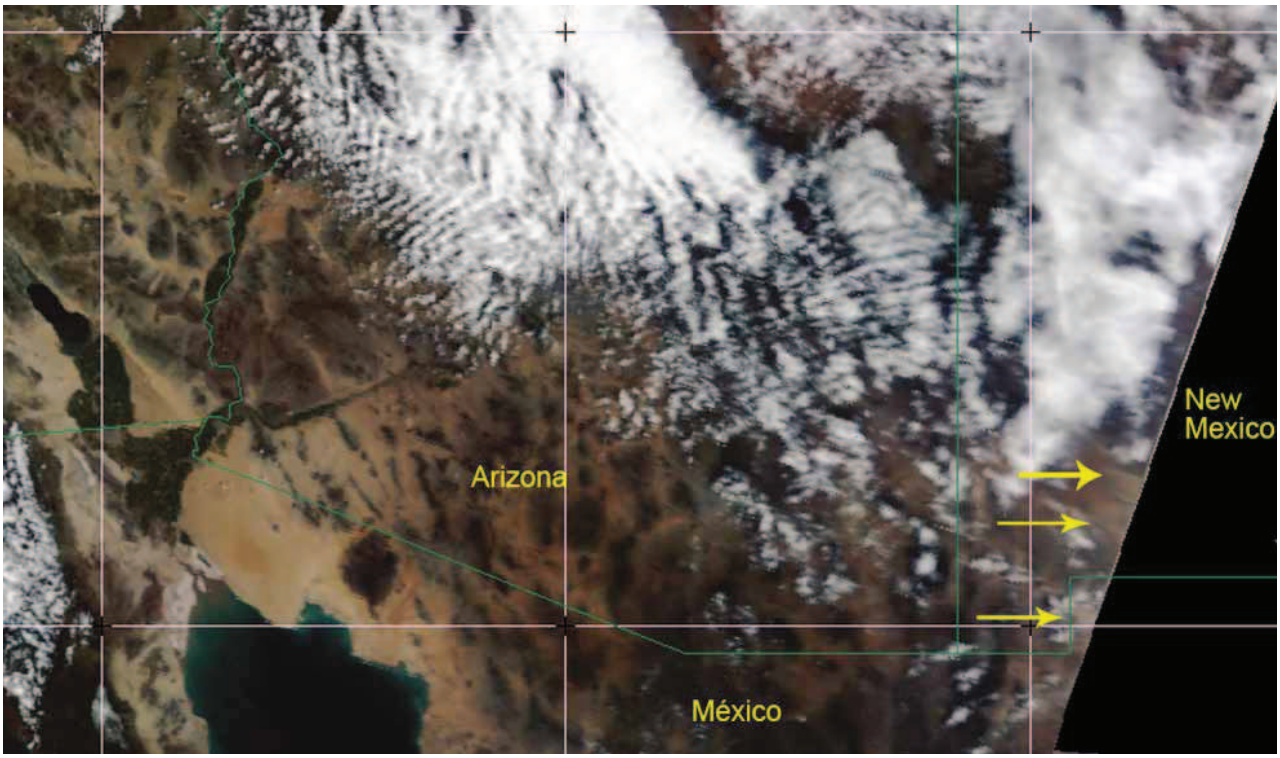

Figure 515. a2003105_1850.png

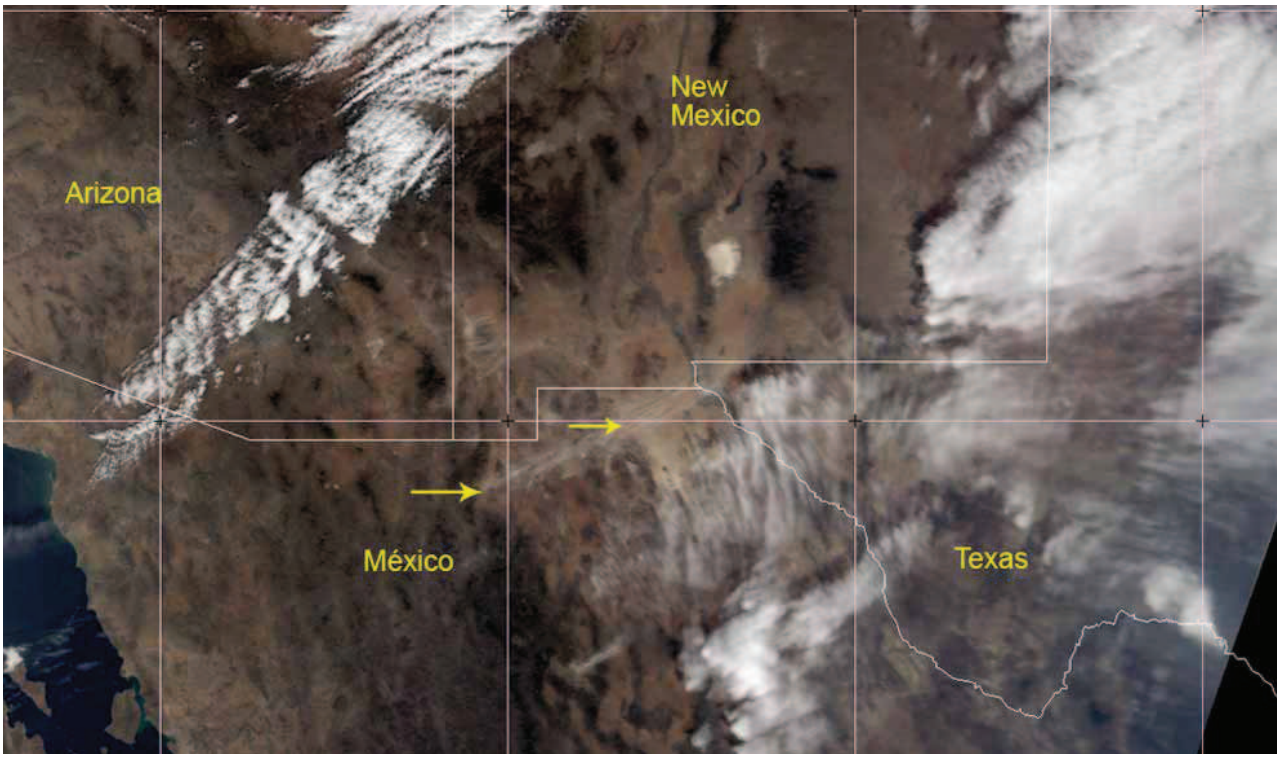

Figure 516. a2003326_1820.png

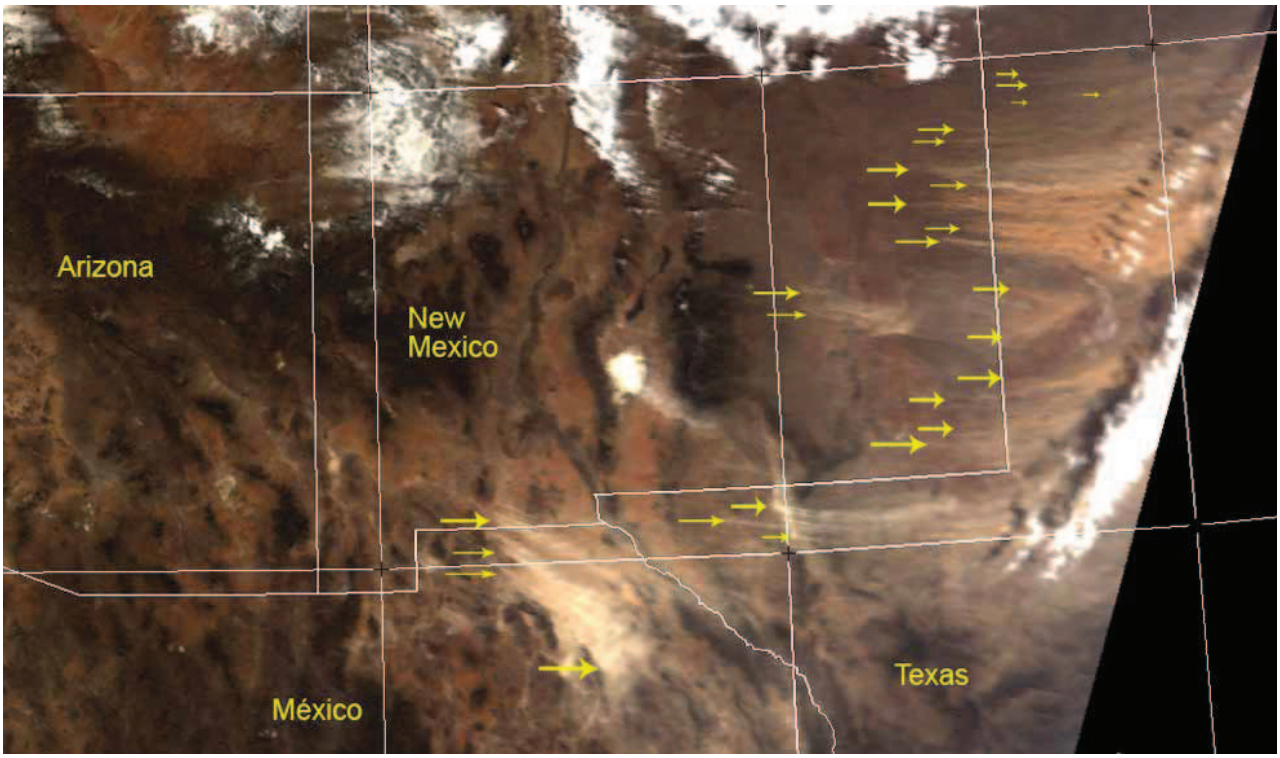

Figure 517. a2003349_1825.png

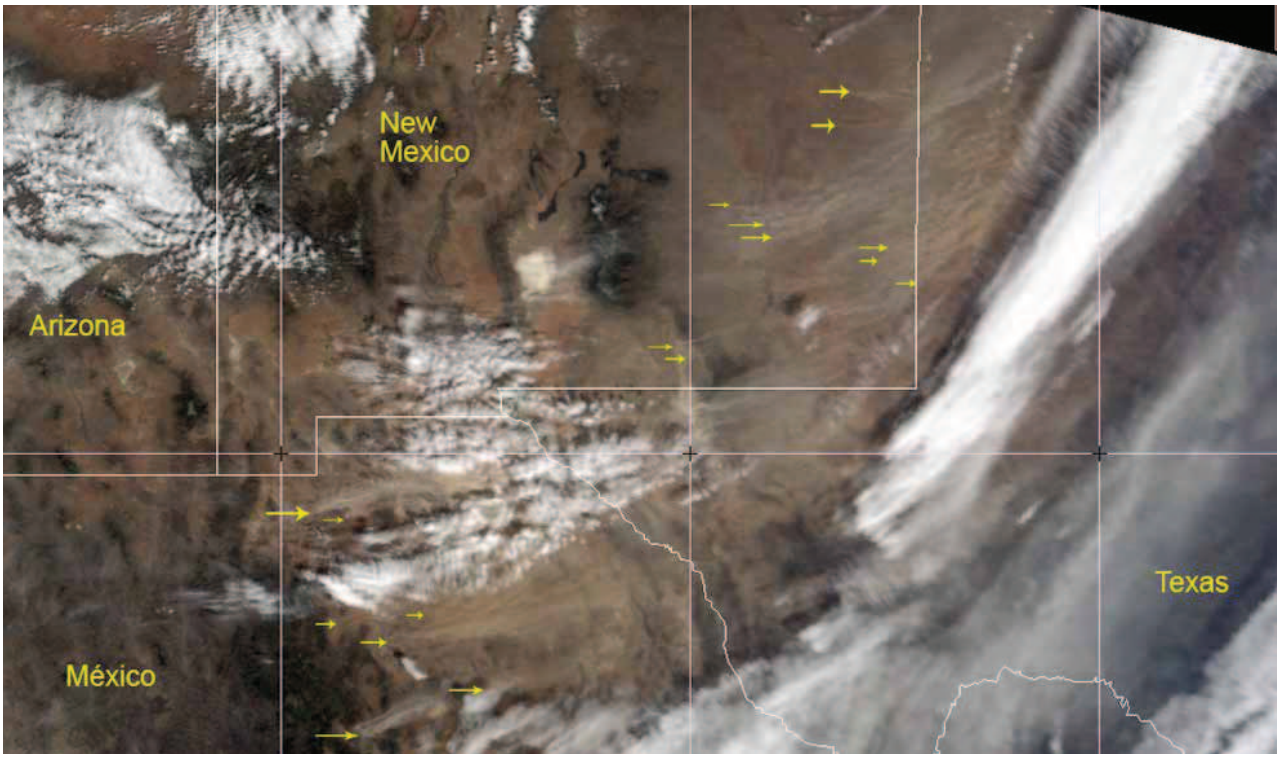

Figure 518. a2004050_1815.png

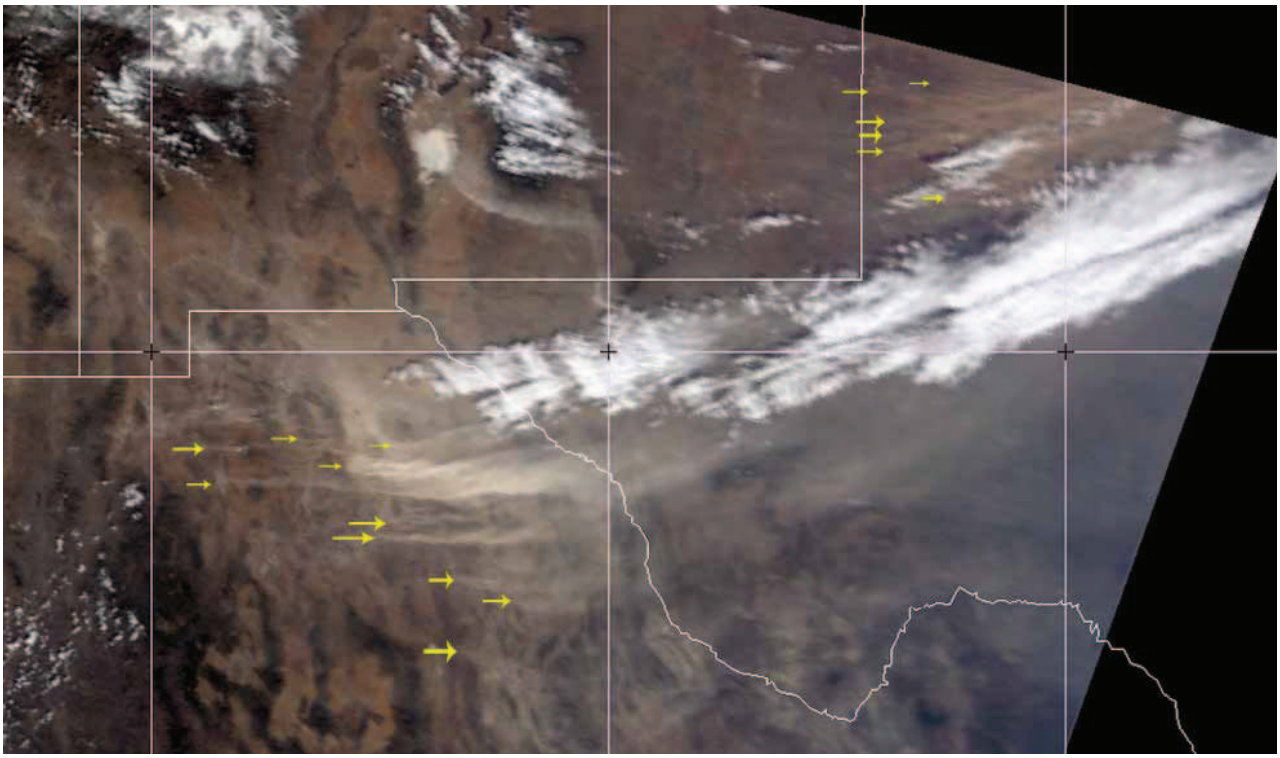

Figure 519. a2005331_1820.png

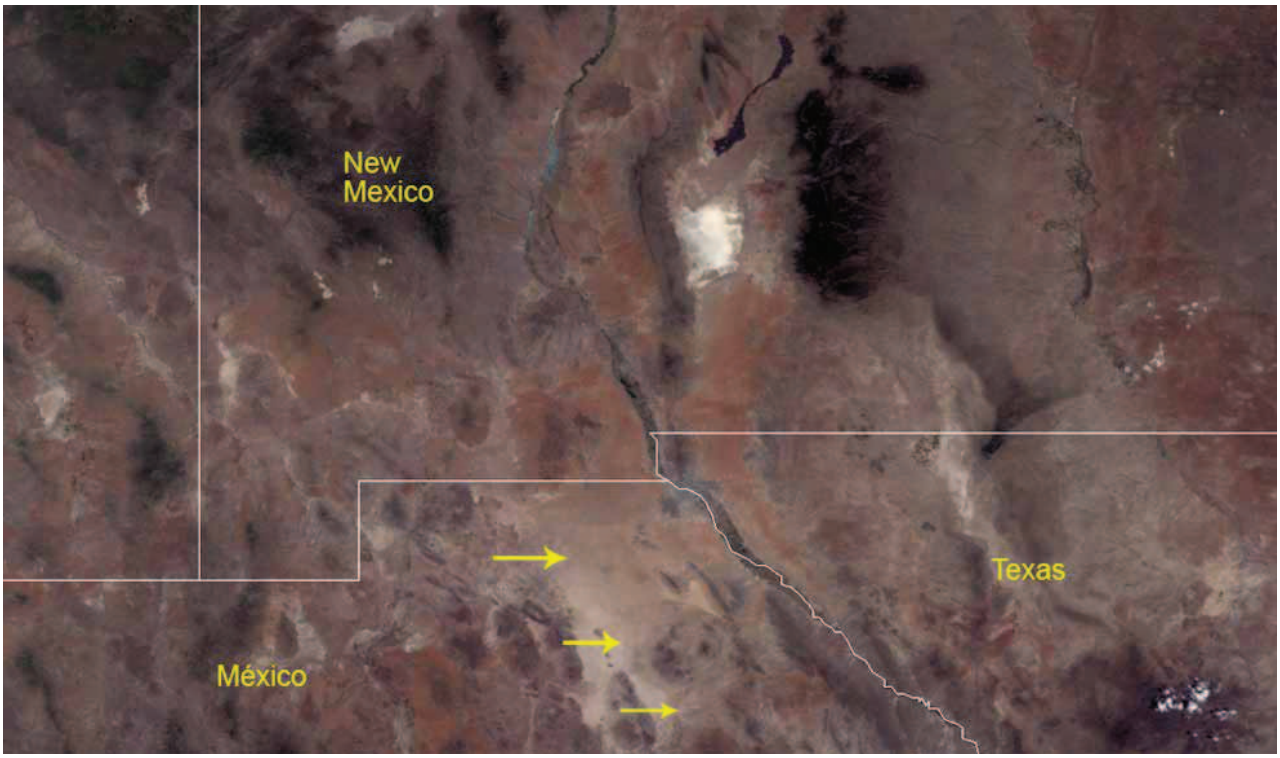

Figure 520. a2001155_1810.png

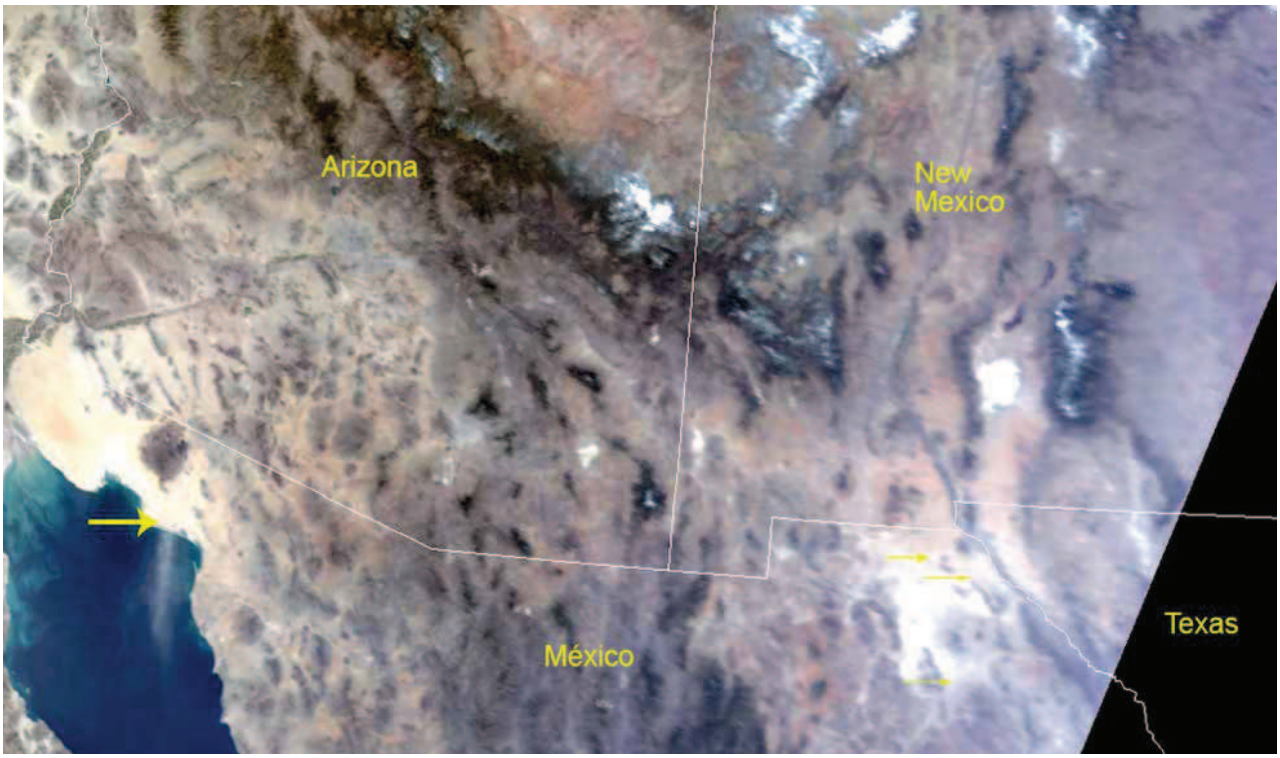

Figure 521. a2002040_1840.png

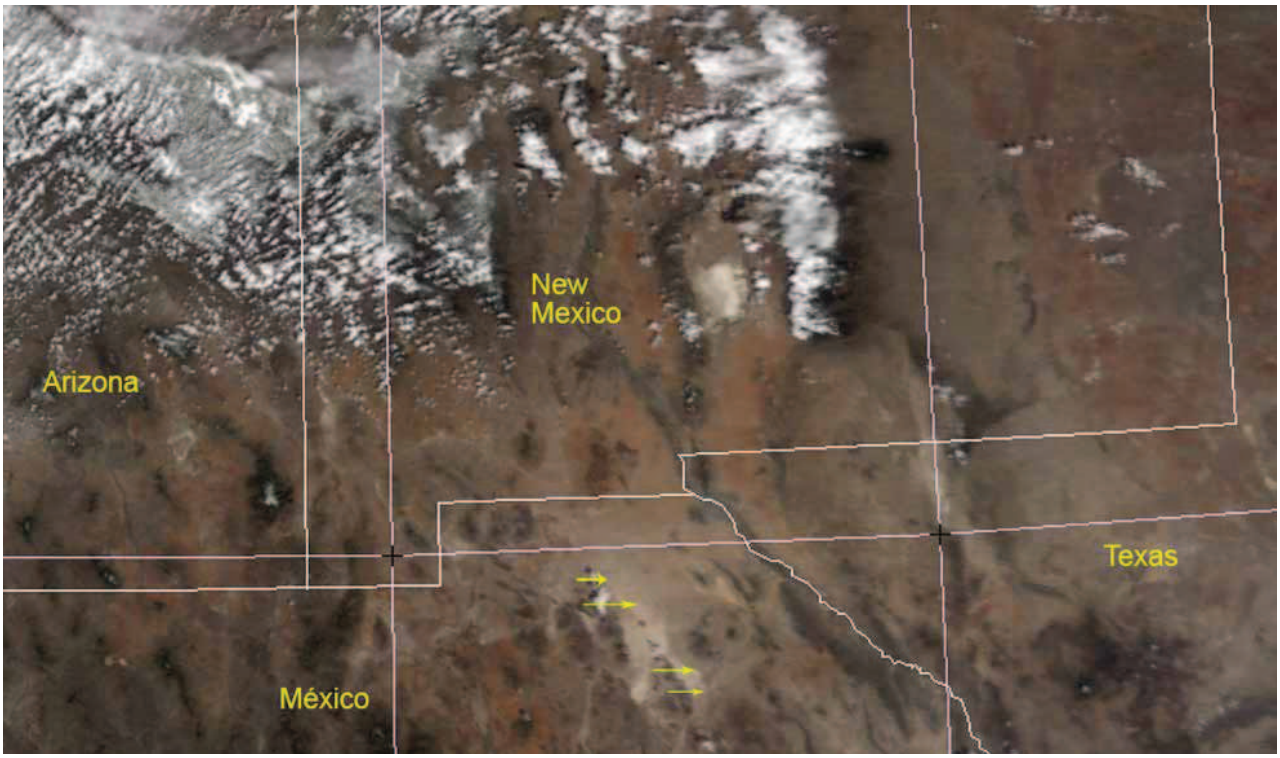

Figure 522. a2003063_1815.png

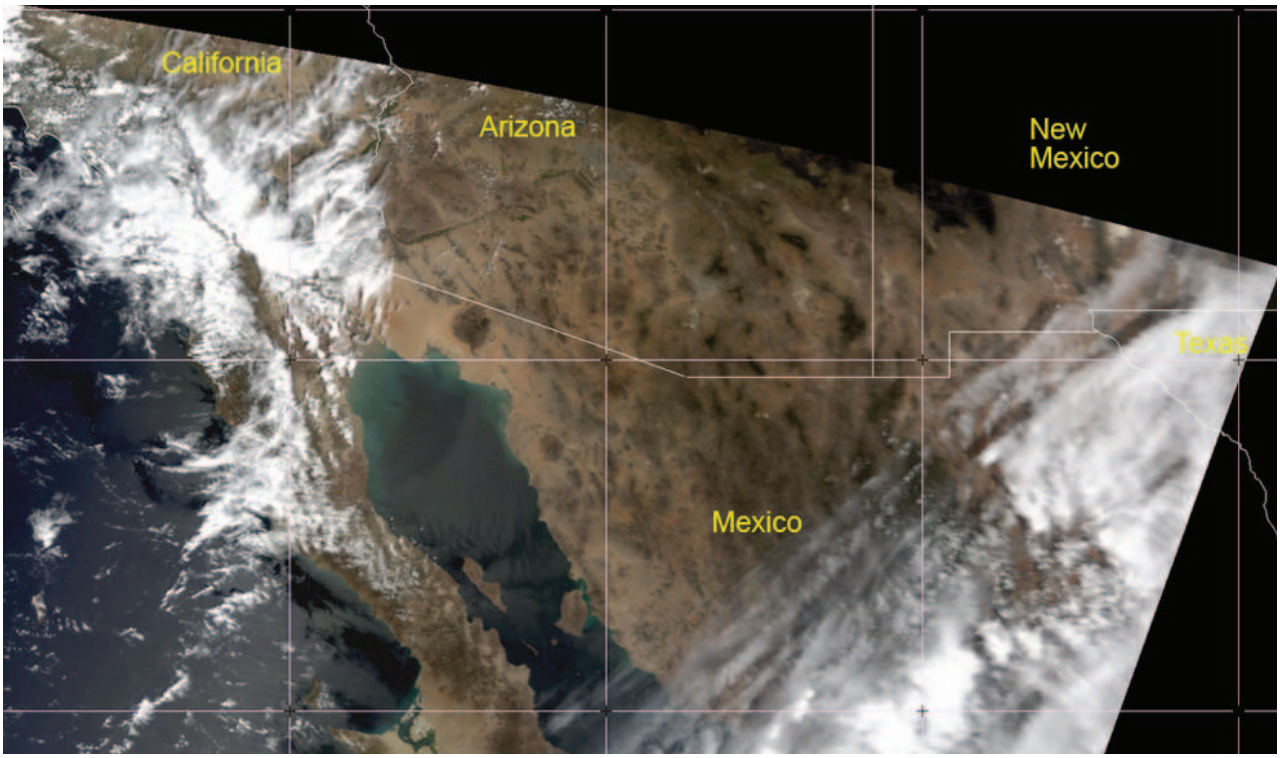

Figure 523. a2003107_1840.png

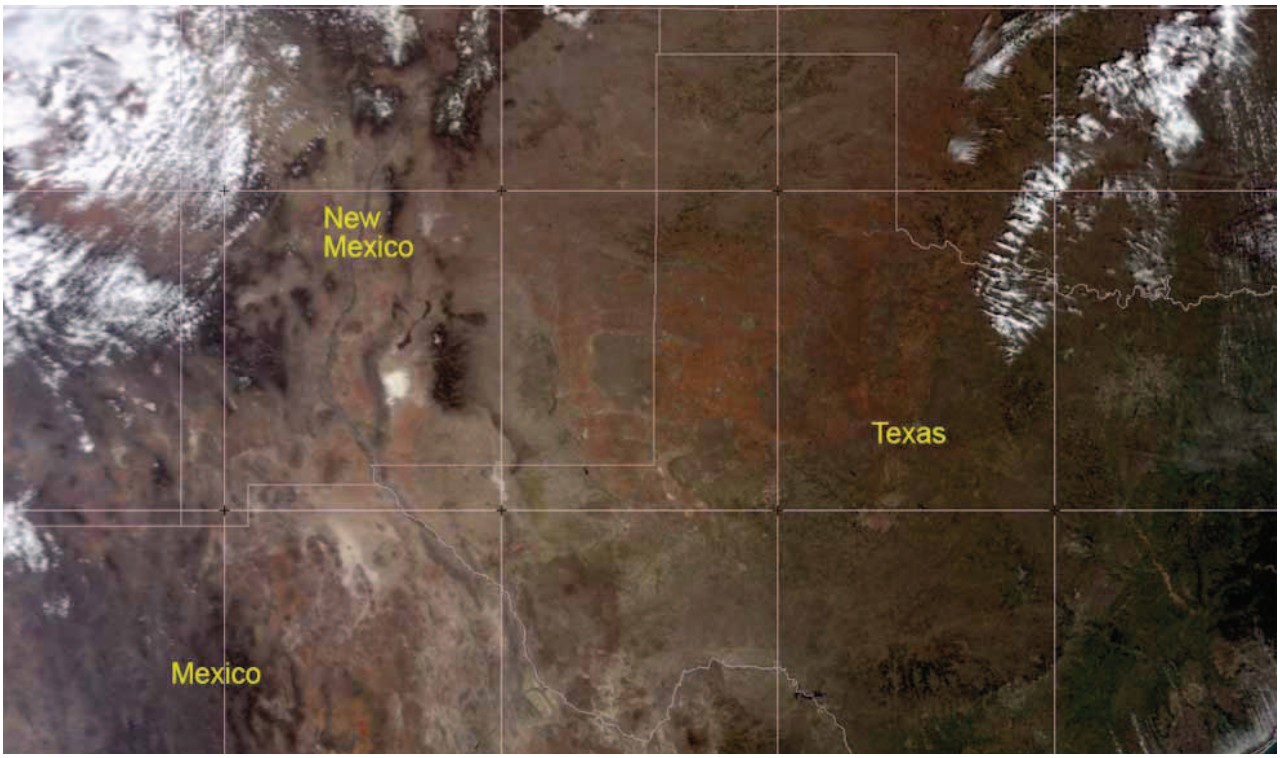

Figure 524. a2005099_1730.png

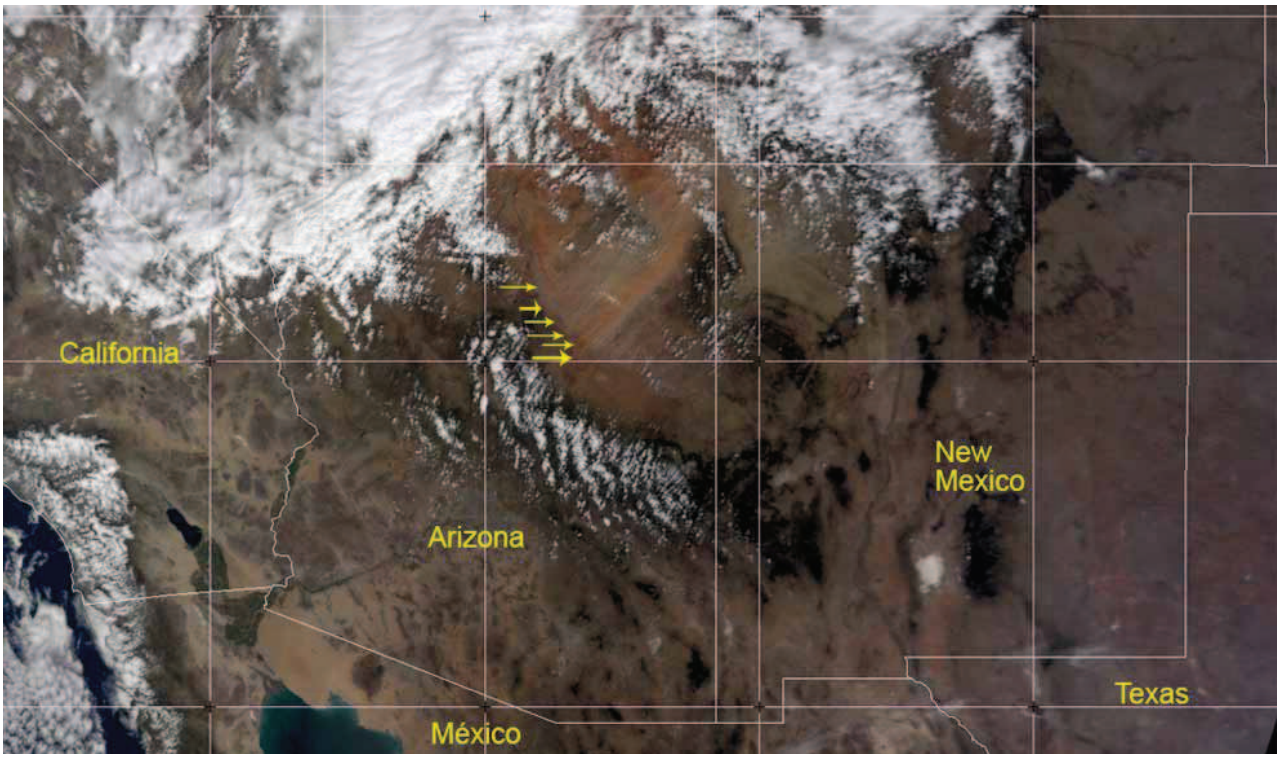

Figure 525. a2009093_1825.png