NEW INDICES FOR THE SPATIAL VALIDATION OF PLUME FORECASTS WITH OBSERVATIONS OF SMOKE PLUMES FROM GRASSFIRES

Joaquin E. Blanco[1] and Guillermo J. Berri [#]

Servicio Meteorologico Nacional, Argentina
[#] Member of Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

ABSTRACT

The purpose of this work is to propose new indices for the spatial validation of hazardous plumes forecast, and apply and test them with data of a case study. One, the Plume-Overlap-Area Hit index, is a modification of a widely used index that considers the overlap area between observed and forecast plumes. The other one, the Plume-Mean-Orientation Hit index, introduces a new concept in plume forecast validation, i.e., the mean direction of plume propagation. These two indices are combined in a new two-dimensional Combined-Direction-Area Hit index. The new indices are applied to the spatial validation of smoke plume forecast for a case study of uncontrolled grassfires that took place during April and May 2008 in the La Plata River region in South America. Operational models at the Argentine National Meteorological Service (SMN) are employed to produce the plume forecast. The HIRHYLTAD dispersion model is used to forecast the smoke plumes, employing the Eta/SMN meteorological forecast model outputs. The forecast plumes are compared to the observed plumes in high resolution MODIS imagery from AQUA and TERRA satellites, from which a total of 59 smoke plumes are identified. The study concludes that the presented methodology that employs operational meteorological models and simplified dispersion models can be used to produce reasonably accurate forecasts of the areas affected by the smoke plumes that originate in forest and grassland fires, particularly in cases when limited information is available about the fires. Although the present study is specifically applied to smoke plumes, the validation technique with the proposed indices can be of utility to study pollutant plumes of diverse nature.

---

1. INTRODUCTION

A wide range of natural and anthropogenic phenomena can derive from hazardous plumes whose impacts affect humans, wild life, vegetation and ecosystems in different time and space scales, and their consequences can become truly severe. Among these we can mention volcanic ash plumes, sand and dust plumes caused by strong winds across deserts and arid regions, and pollution associated with stack emissions from factories or similar point sources. Sometimes they become international in scope for their particular nature, origin or magnitude, like the Kuwaiti oil and gas well fires in 1991 during the Gulf war (Spektor, 1998), and the radionuclide contamination originated by the nuclear disasters in Chernobyl in 1986 (Piedelievre et al., 1990) and Fukushima in 2011 (Lujaniene et al., 2012).

In particular, a common example of hazardous material consists of smoke plumes from biomass burnings, which cause a major impact on population and social activities. When forest and

grassland fires become uncontrolled and extended, the deterioration of ecosystems may require several decades to recover[2]. Every year in Africa, South and North America, the Russian Federation, Southeast Asia and other regions, a variable number of wildfires takes place especially during summer over dry land, and might eventually become uncontrolled and remain so for several days; on occasions they are due to prescribed burnings in rural zones (Schultz, 2002; van der Werf et al., 2010). Emissions associated with these phenomena include different gases and a large spectrum of aerosols and particulate matter of different size and type (Andrae and Merlet, 2001).

Real-time emergency response systems have been developed and directed during the past decades, to prevent, avoid, or mitigate any impact of the hazardous materials, and with growing concern about estimating the long-range dispersion of pollutants. More recently, the need to assess the accuracy of hazardous material transport and dispersion models has arisen as one of the most important issues to be addressed (Petty, 2000; Brost et al., 1988).

Several quantities have been used to perform evaluation analysis. For example, statistical measures of bias, scatter and correlation have been generally applied to point-to-point comparisons (observations and predictions paired in space and time). More recently, the concept of spatial forecast validation was introduced, with the definition of indices that analyze the distribution of both plumes on a single image and, particularly, the areas of overlap. Mosca et al. (1998) and Boybeyi et al. (2001) used different meteorological and dispersion models to perform aerosol and gas hazard prediction, and verified their predictions against the European Tracer Experiment (ETEX) measurements. Warner et al. (2003) used the Project Prairie Grass experiment results for validating his simulations by means of overlap area indices. However, it is the National Oceanic and Atmospheric Administration (NOAA) that routinely produces forecasts and validation of plumes originated from biomass burning. The NOAA's Smoke Forecasting System (SFS) comprises three parts: observation, prediction and validation of smoke plumes. Full details about the descriptions of the components can be found in Rolph et al. (2009) and Ruminski et al. (2006, 2007), while SFS products are available in real time at http://www.arl.noaa.gov/smoke.php. Also, a verification of the 2007 fire season in North America and a sensitivity study of the plume injection height were performed by Rolph et al. (2009), and Stein et al. (2009).

The spatial validation of plume forecasts usually considers observed and predicted plume layout to qualify the forecast. For example, NOAA's SFS uses different indices based on combinations of overlap and non-overlap areas covered by both plumes (Mosca et al., 1998; Boybeyi et al., 2001; Warner et al., 2004), which are discussed in detail in the following section. The objective of this paper is to propose new indices for the spatial validation of plume forecasts, i.e., a modification of a widely used index for the overlap area, and a new plume-mean-orientation index. Furthermore, they are combined in a new two-dimensional index that takes into consideration both the overlap area and the direction of propagation of the observed and forecast plumes. Section 2 discusses different spatial validation techniques and presents the new indices. A case study of smoke plumes originating from grassfires, that took place in the La Plata River region of South America during April and May 2008, is described in Section 3, in which the spatial validation methodology is applied, and the indices are evaluated. Finally, Section 4 presents a summary of results and the conclusions of this work.

---

[2] Fire is an essential ecological process and many ecosystems (particularly prairie, savanna, chaparral and conifer forests) require it as a contributor to habitat vitality and renewal.

2. SPATIAL VALIDATION OF PLUMES

2.1. Overlap area index

Before introducing the overlap area indices, it is convenient to briefly review the different types of areas covered by observed and forecast plumes. The areas covered by the observed (or measured) and forecast plumes (*Aobs* and *Afor,* respectively), determine three other areas, namely: intersection or overlap area (*Aint*), false detection area (*Afal*), and no detection area (*Anod*), which can be seen in Fig. 1 and Table 1. The sum of these three areas corresponds to the union area (*Auni*). In fact, a fourth category could be considered, but of less relative importance since it corresponds to the empty or blank area, that is, all the space in the domain with neither forecast nor observed plumes.
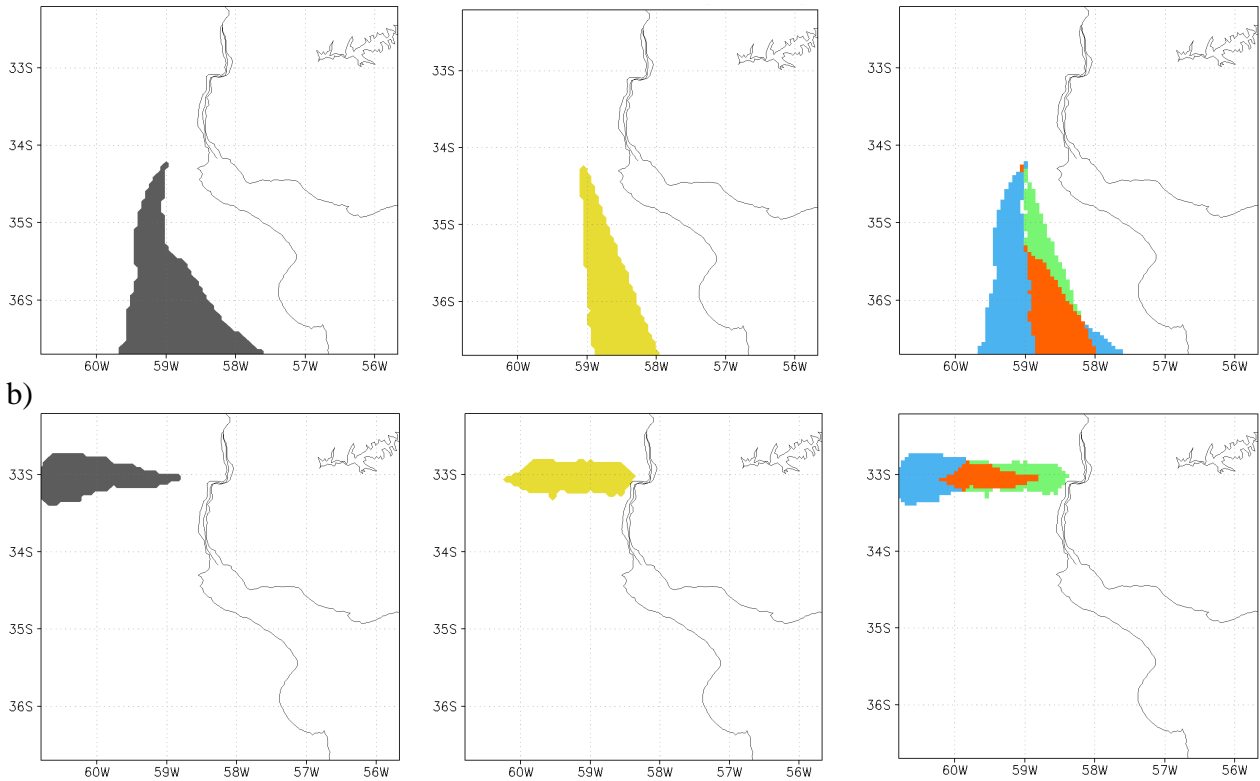

Figure 1: Examples of area types, a) case 1, b) case 2. Left: observed plume. Center: simulated plume. Right: Overlap area in red, no detection area in blue, false detection area in green.

Relationships among these areas can be established in order to define validation indices. For example, the verification analysis at the NOAA's SFS uses two overlapping area indices, namely the Figure of Merit in Space –*FMS*- (Mosca et al., 1998 and Boybeyi et al., 2001), and the two-dimensional Measure of Effectiveness –*MOE*- (Warner et al., 2004).

The *FMS* index is defined as the ratio between the intersection area of observed and forecast plumes and their union area, i.e., *FMS = Aint/Auni*, which can be rewritten as

$$FMS = \frac{Aint}{Aobs + Afor - Aint} = \frac{Aint}{Anod + Afal + Aint} \tag{1}$$

		Observation	
		Yes	No
Forecast	Yes	*Aint*: "hit"	*Afal*: "false alarm"
	No	*Anod*: "miss"	*Ablk*: "correct rejection"

total number of pixels $N = Aint + Afal + Anod + Ablk$

Table 1: Area types in the typical arrangement of a 2x2 contingency table for forecasting a categorical event, in which each box indicates the number of pixels that verify the respective condition. *Aint* is the intersection or overlap area, *Afal* is the false detection area, *Anod* is the no detection area, *Ablk* is the empty or blank area, and *N* is the total number of pixels in the domain.
(Adapted from http://www.swpc.noaa.gov/forecast_verification/Glossary.html)

FMS is also called the threat score or critical success index (*CSI*), for categorical forecast validation. The index is calculated at a fixed time and for a fixed concentration level. Scores range between 0 (no match at all) and 1 (ideal case, all pixels match). However, a limitation of this index is that it cannot distinguish whether the observed plume footprint is smaller or larger than the predicted plume footprint (i.e.: over- or under-prediction of area are "weighted" identically). Therefore, an event with an observed plume fully circumscribed to the simulated plume, can score exactly the same as an event with the latter contained into the former. For some applications, the distinction between these two alternatives can provide insight to a model user for subsequent calibration.

The two-dimensional *MOE* index used by NOAA's SFS to measure the overlap area of the two plumes is given by *MOE(x,y)* = (*Aint/Aobs* , *Aint/Afor*), which can be rewritten as:

$$MOE(x, y) = \left(\frac{Aint}{Aint + Anod}, \frac{Aint}{Aint + Afal} \right) = \left(1 - \frac{Anod}{Aobs}, 1 - \frac{Afal}{Afor} \right) \qquad (2)$$

In terms of categorical forecasts, the first ratio is also called the Probability of Detection (*POD*), and the second one, the Success Ratio (*SR*). Scores for all analyzed events can be plotted in a single *x-y* diagram, and Fig. 2 shows an example with two events. The point (100%:100%) represents the perfect match between the observed and forecast plumes; that is, both shapes have the same size and location. Points along the 1:1 line represent the cases in which the two shapes are identical in area (same number of pixels) but are shifted in space, so that a point at (0:0) means no overlap at all. Points in the upper-left portion of the plot represent cases in which the forecast plume is nearly covered by the observed plume; however, the latter is larger than the former (i.e.: under-prediction). Conversely, the over-prediction cases are located in the lower-right portion of the plot.

A limitation of the aforementioned *FMS* and *MOE* indices that consider the shape matching approach is that their scores are often too low, due to the nature of the *FMS* and *MOE* ratios

themselves. Thus, they could indicate an apparently much poorer performance than what a qualitative examination of the spatial distribution of observed and forecast plumes on a same image would suggest. For this reason, we define a variant of *FMS*, namely the Plume-Overlap-Area Hit (*POAH*) index, which is obtained by adding *Aint* to the numerator and denominator of Eq. (1):

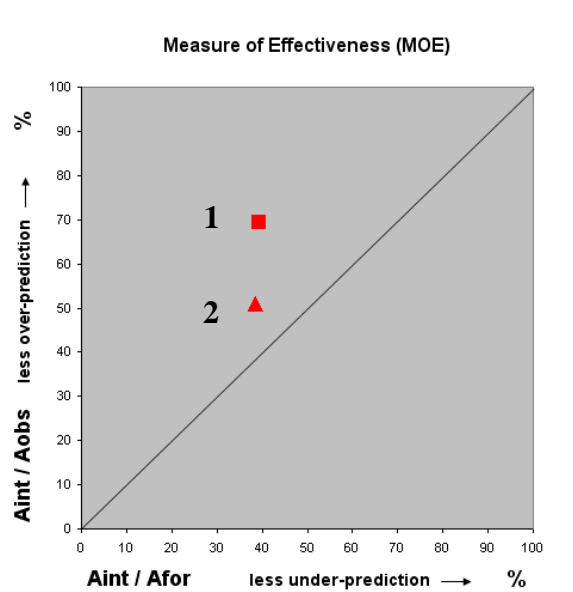$$POAH = \frac{2Aint}{Aobs + Afor} \tag{3}$$

Figure 2: MOE plot associated to the examples of Figure 1. MOE scores are (39.22; 69.41) and (38.56; 51.04), respectively, for cases 1 and 2.
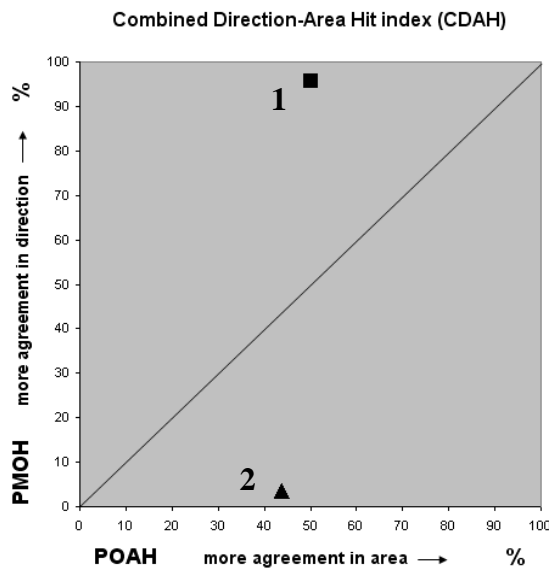

Figure 3: CDAH plot associated to the examples of Figure 1. CDAH scores are (50.12; 95.66) and (43.93; 3.48), respectively, for cases 1 and 2.

Conceptually identical to *FMS*, this index not only ranges between 0 and 1, but also scores better for the intermediate cases since $POAH \geq FMS$, always, as can be easily demonstrated. If errors are to be accounted for, then the Plume-Overlap-Area Error (*POAE*) index can be defined as $POAE = 1 - POAH$.

There is another index used in categorical forecast very similar to *FMS,* with the fourth and last area category included, i.e., "correct rejections" (for our purposes, the blank area), which is summed both at the numerator and denominator. It is called percent correct (*PC*) and is given by: $PC = (Aint + Ablk)/(Aint + Anod + Afal + Ablk)$. This metric, however, is impractical because it is dependent on the size of the domain (the same configuration of the pair observed-predicted plumes in a larger or smaller domain will imply different *Ablk* and hence, different *PC* scores). Moreover, given any domain, *Ablk* will generally be much bigger than the other three area types, and *PC* will tend to 1 for most events. Although *POAH* was defined in order to obtain better numerical results than with *FMS*, the scores behavior with *PC* becomes extreme.

Finally, it is important to mention that if we apply the categorical forecast analogy to the *POAH* index we have $POAH = 2Aint/(2Aint + Anod + Afal)$, that is, a simple function of *FMS*. As a matter of fact, any *k* integer number could replace the "2 factor" in the definition of *POAH*, to establish a generalized ratio $POAH = kAint/(kAint + Anod + Afal)$. As *k* increases, the ratio tends to unity, and gradually the same problem stated for the *PC* metric arises. For this reason, k was set to a value of 2, so that it is large enough to ensure reasonable overlap scores (partially counteracting the *FMS* limitation), and at the same time, small enough to keep the *POAH* expression simple (eq. 3) and avoiding exaggerated and useless scores.

2.2. Direction of propagation index

In recent years, many authors have addressed the topic of inadequacy of the traditional approach for the spatial verification of forecast, based on simple area overlays (i.e.: simple pixel-to-pixel correspondence), generally while evaluating precipitation fields (Zhu et al, 2011). As we explained above for plume analysis, the results of the validation with such technique (metrics such as *CSI, POD, SR, FAR, PC*, etc.), are often not consistent with what a forecaster or analyst might infer by more subjective visual evaluation of a forecast. This tangle has led to the idea that an objective approach that would more closely mimic the subjective approach could provide more useful, diagnostic information about the quality of spatial forecasts (Davis et al, 2006; Wernli et al, 2008). For this reason, object-based verification approach, itself not new, has become widely used, and simultaneously, updated and enhanced (Davis et al, 2009). Object-based assessment identifies "objects" in the forecast and observed fields that are relevant to a human observer. These objects can then be described geometrically, and the attributes of forecast and observed objects can be compared. Because different users may need or have different kinds of information, it is important to allow flexibility in the definition of these attributes. Some object features are listed in Table 2.

There is a distinctive difference between object-based verification for plumes and for other type of meteorological phenomena such as precipitation, icing, turbulence, etc. For the latter, the evaluation can be performed for unmatched objects as well as those that have overlay areas of forecast and observation (see schematic examples on either Fig. 1 of Davis et al, 2006 or Fig 1 of Wernli et al, 2008), while for the former, the two objects are typically matched in the source region. As a matter of fact, the number and position of point sources are inherent object attributes

of smoke plumes, and were not included in Table 2 because forecast will not differ from observation on this issue for the general case.[3]

Object property	Description
Intensity or concentration	Thresholding is useful to define an object's boundary. Different thresholds may lead to changes in the other object's properties. In particular, by narrowing the threshold, the object can be subject to partitioning into two or more pieces or elements, which can eventually be considered as separate objects.
Object boundary and area	Primary properties of an object: the region enclosed by the object's boundary is the area, which is a simple measure of its size.
Centroid or center of mass	Measure of mean location of object, for which the intensity or concentration of the property might be used as a weighting factor.
Overall direction or axis angle	A line drawn through the centroid of an object best characterizes its overall direction or orientation. This line represents the major axis of the object; if the object is bounded by a minimum, enclosing rectangle, the major axis is parallel to the longer sides.
Aspect ratio	Given the major and minor axes of an object, the ratio of the length of the minor axis to the length of the major axis is the aspect ratio.
Curvature	Fitting a circular arc to an object instead of a line gives a measure of the object's overall deviation from straightness.

Table 2. Summary of object attributes relevant to object-based verification (adapted from Davis et al., 2006)

Herein, we apply the object-based approach using some of the attributes listed in Table 2 for the spatial validation of plume forecasts to define the Plume-Mean-Orientation Hit (*PMOH*) index as follows:

$$PMOH = 1 - \frac{\left| \overline{dir}_{obs} - \overline{dir}_{for} \right|}{180°} \tag{4}$$

where $\overline{dir}_{obs}$ and $\overline{dir}_{for}$ are the observed and forecast mean directions of plume propagation, respectively, ranging from 0° to 360°. The numerator is the absolute value of the angle between the two directions, which is always at most 180°. The best score for the index is 1, when both mean directions of propagation are equal, and 0 when they are opposed to each other. The error alternative of this index, namely the Plume-Mean-Orientation Error (*PMOE*), is defined as *PMOE* = 1 – *PMOH*.

The mean direction of propagation $\overline{dir}$ can be estimated subjectively, as was the case in Byrne et al. (2007) for the qualitative verification analysis of volcanic ash plumes (direct comparison of $\overline{dir}_{obs}$ and $\overline{dir}_{for}$). Instead, in the present work we apply the following automated algorithm in

---

[3] We are neglecting here the possibility that a given source position of a plume forecast can result significantly distant from the actual position of the real plume. For real-time operational plume forecasts such as the NOAA SFS, the sources are identified by the time the fire is starting (or afterwards, in the simulations for our case study, i.e.: diagnostic mode). If this is not the case, number and location of sources can be validated separately.

order to calculate $\overline{dir}$. For the case of a single-source, $\overline{dir} = \tan^{-1}\left[(\overline{y} - y_{source})/(\overline{x} - x_{source})\right]$, where (x_{source}, y_{source}) are the coordinates of the fire location, $(\overline{x}, \overline{y})$ is the centroid position of the plume ($\overline{x} = \sum x_i / N$, $\overline{y} = \sum y_i / N$), and N is the total number of pixels comprising the plume ($i = 1, ..., N$). In case of a multiple-source plume, (x_{source}, y_{source}) represents the centroid position of all fire sources, defined in a similar manner as $(\overline{x}, \overline{y})$.

Other criteria for the calculation of $(\overline{x}, \overline{y})$ consider the center of mass, or the maximum concentration pixel. Any of these two would be more convenient than the centroid algorithm in the case of multiple-source plumes, because of the different contributions of various individual plumes onto a single pixel. In such case, the mean direction of propagation will point to the place of highest concentration (where the added effect is greatest), which can be distant from (and non-aligned with) the plume centroid. A clear limitation is that concentrations are required for performing the calculations with these algorithms. In the case of smoke plumes, a quantitative analysis of the observed plume in the satellite image is necessary to account for the concentrations (for instance, automated products derived from the satellites such as Aerosol Optical Thickness -AOT- or the Goes Aerosol/Smoke Product -GASP-). Therefore, when concentrations are unavailable, the centroid approach is the most appropriate method for calculating *PMOH* (or *PMOE*). Since the case study of Section 3 uses visible satellite imagery, the centroid criteria is used for calculating *PMOH* and *PMOE* scores.

The advantage of *PMOH* over *FMS* or *POAH* can be appreciated by comparing the two cases of Fig. 1. The scores for the overlap area and mean direction error indices of Fig 1a (single-source plume) are: 1-*FMS* = 66.56%; *POAE* = 49.88%, and *PMOE* = 4.34%. Therefore, while the overlap area errors are near 50% or worse, the *PMOE* score is less than 5%, more in agreement with a straightforward examination of the plumes layout. The scores for Fig 1b (multiple-source plume) are: 1-*FMS* = 71.85 %; *POAE* = 56.07%, and *PMOE* = 96.52%. Hence, in contrast with the previous case, the almost 50% score of *POAE* index is not really representative of the model performance, since the forecast plume is propagating in the opposite direction with respect to the observed plume, so that *PMOE* is close to 100%.

2.3. A new two-dimensional index

The two spatial indices, *POAH* (Eq. 3) and *PMOH* (Eq. 4), are combined into a single two-dimensional index, namely the Combined-Direction-Area Hit (*CDAH*) index, defined as:

$$CDAH(x,y) = (POAH, PMOH) \tag{5}$$

The *x* component indicates the overlap area between the observed and forecast plumes, while the *y* component accounts for the agreement in direction of propagation. A *CDAH* plot associated to the examples of Fig. 1 is shown in Fig. 3. As in the case of the *MOE* index, the (100%,100%) point corresponds to a perfect forecast, i.e., both plume distributions are equal, while (0,0) is the worst possible case since the forecast plume propagates in opposite direction to the observed one, and there is no area of intersection. Points along the 1:1 line represent the cases in which *POAH* is equal to *PMOH*. By comparing both *MOE* and *CDAH* plots (Fig. 2 and Fig. 3, respectively), it is clear that the scores from the latter represent more adequately and insightfully the plume layout of the cases depicted in Fig. 1, due to the incorporation of the new property, i.e., the mean direction of propagation.

Another interesting feature of the *CDAH* plot is that it allows for discriminating the cases in which the error in overlap area is smaller than the error in direction of propagation (*POAH* > *PMOH*, i.e., points in the lower-right portion of the plot). As can be seen in Fig. 3, case 2 (Fig. 1b) corresponds to an idealized situation in which the error in direction of propagation is greater than the error in overlap area. Moreover, using *POAH* instead of *FMS* as the *x* component of *CDAH*, favors a less-unbalanced difference between the total number of events in the two cases (in other words, the chance of more points under the 1:1 line increases), since *POAH* ≥ *FMS* always. Finally, an alternative Combined-Direction-Area Error (*CDAE*) index can be defined as *CDAE(x,y) = (POAE, PMOE)*.

3. A CASE STUDY

3.1. Description of the event

The case study focuses on the smoke produced by pasture burnings that took place during April and May 2008 in the Paraná River Delta, some 70 km to the northwest of the city of Buenos Aires (the study region is depicted in Fig. 4). The local authorities declared that the fires were intentionally ignited and did not respond to any prescribed plan. The fires soon became uncontrolled and extended, and the smoke propagated over a wide region according to the prevailing atmospheric conditions of each day. Smoke spread northeasterly hundreds of kilometers across the La Plata River into Uruguay and southern Brazil, and southerly as far as the extreme south of Buenos Aires province. The grassfires blackened about 70,000 hectares in the provinces of Buenos Aires and Entre Ríos.

Figure 4: MODIS real-color satellite image (CEILAP-BA subset) in the La Plata River region in South America, with the red box showing the region of the case study. The blue circle indicates the location of the city of Buenos Aires.

The fires that took place between 16 and 20 April 2008 had no historical precedent because of the negative consequences they caused on daily life of 13 million people in the conglomerate of Buenos Aires city and its suburbs. Figs. 5a and 5b illustrate the situation in the afternoon of 16 and 18 April 2008, in which the grassfires in southern Entre Ríos province created a dense smoke

plume that propagated to Buenos Aires city and the La Plata River. Fig. 5c shows the extended smoke plume over the Buenos Aires province and the La Plata River, associated with a wind rotation during the morning of 20 April 2008. The episode resulted in increasing health problems among the population, such as respiratory diseases, skin and eye irritation, etc. CO concentrations in the city of Buenos Aires were 17 ppm on 17 April 2008 and 15,3 ppm on 19 April 2008 (the typical value for a common day is less than 2 ppm); while total suspended carbonate particulate matter concentration on 18 April 2008 was 2,024 mg/m^3 (Clarin newspaper editions, April, 2008). Due to visibility reduction, there were hazardous driving conditions and accidents that forced traffic interruptions of highways, as well as inoperability of airports. The impact of the smoke event was reflected as the main issue addressed by local and regional media, capturing international attention as well (Clarin and La Nacion newspapers editions, April, 2008). The persistence of anomalous northwesterly winds during those days contributed to that situation (Marcuzzi and Hoevel, 2009). Mattio (2009) performed smoke plume dispersion simulations and compared them only qualitatively with visible satellite images. Berbery et al. (2008) used WRF-ARW regional model outputs to study the predictability of the episode, although no smoke plume simulations were made.
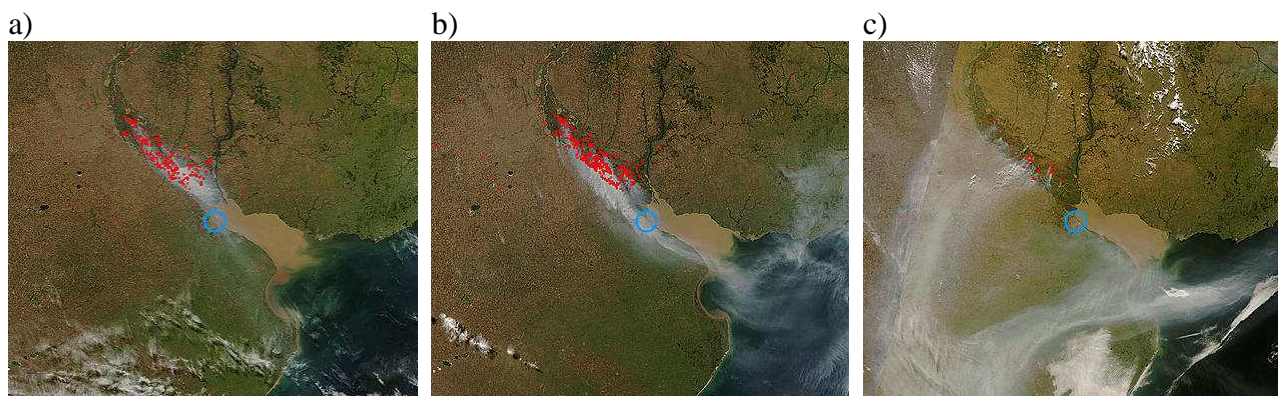
a)   b)   c)

Figure 5: MODIS real-color satellite images (CEILAP-BA subset) from: a) AQUA, 16 April 2008 at 1800 UTC; b) AQUA, 18 April 2008 at 1750 UTC; c) TERRA, 20 April 2008 at 1500 UTC. The red dots correspond to MODIS automated fire detections, and the blue circle indicates the location of the city of Buenos Aires.

3.2. Meteorological and dispersion models

Two operative models at the National Meteorological Service of Argentina are used in this study: the Eta/SMN meteorological forecast model and the HIRHYLTAD dispersion model. The smoke plume forecasts are performed with the HIRHYLTAD puff model (for a detailed description see Blanco and Berri, 2011 and Blanco, 2011), which uses the meteorological forecast of Eta/SMN. HIRHYLTAD is based on the gaussian dispersion model and also on a lagrangian trajectory model, since it first determines smoke lines using lagrangian trajectories, and then simulates dispersion and calculates gaussian concentrations. The model uses the Pasquill stability classes (Pasquill, 1961), the Pasquill-Gifford dispersion coefficients according to the stability class (Gifford, 1976), and the Briggs scheme for buoyant plume rise (Briggs, 1969). HIRHYLTAD is a hybrid model because the input for calculating smoke lines is the forecast wind field at discrete time intervals, while the gaussian model assumes homogeneity and continuity in space, as well as steady-state conditions. Also, it uses a lagrangian (mobile) coordinate system to calculate the trajectories of individual smoke plume elements, and an eulerian (fixed) framework to calculate concentrations in a high resolution grid (see Fig. 6 for a schematic example of the model steps).

The Briggs parameterization scheme for the plume rise Δh requires the following emission parameters: smoke temperature Ts, vertical velocity ws, and 'effective' diameter of the smoke

column *Ds*. Since no such measurements were made during the events, we adopted the following values from the FireFlux experiment (Clements et al., 2007), $Ts = 200°C$ and $ws = 1ms^{-1}$ (both at 10m). To be consistent, the adopted smoke initial emission height is 10m, instead of the surface level, so that the effective emission height is $H = 10m + \Delta h$. A sensitivity test indicated that $Ds = 2.5m$ (slightly larger than a typical stack diameter), was the most appropriate value.
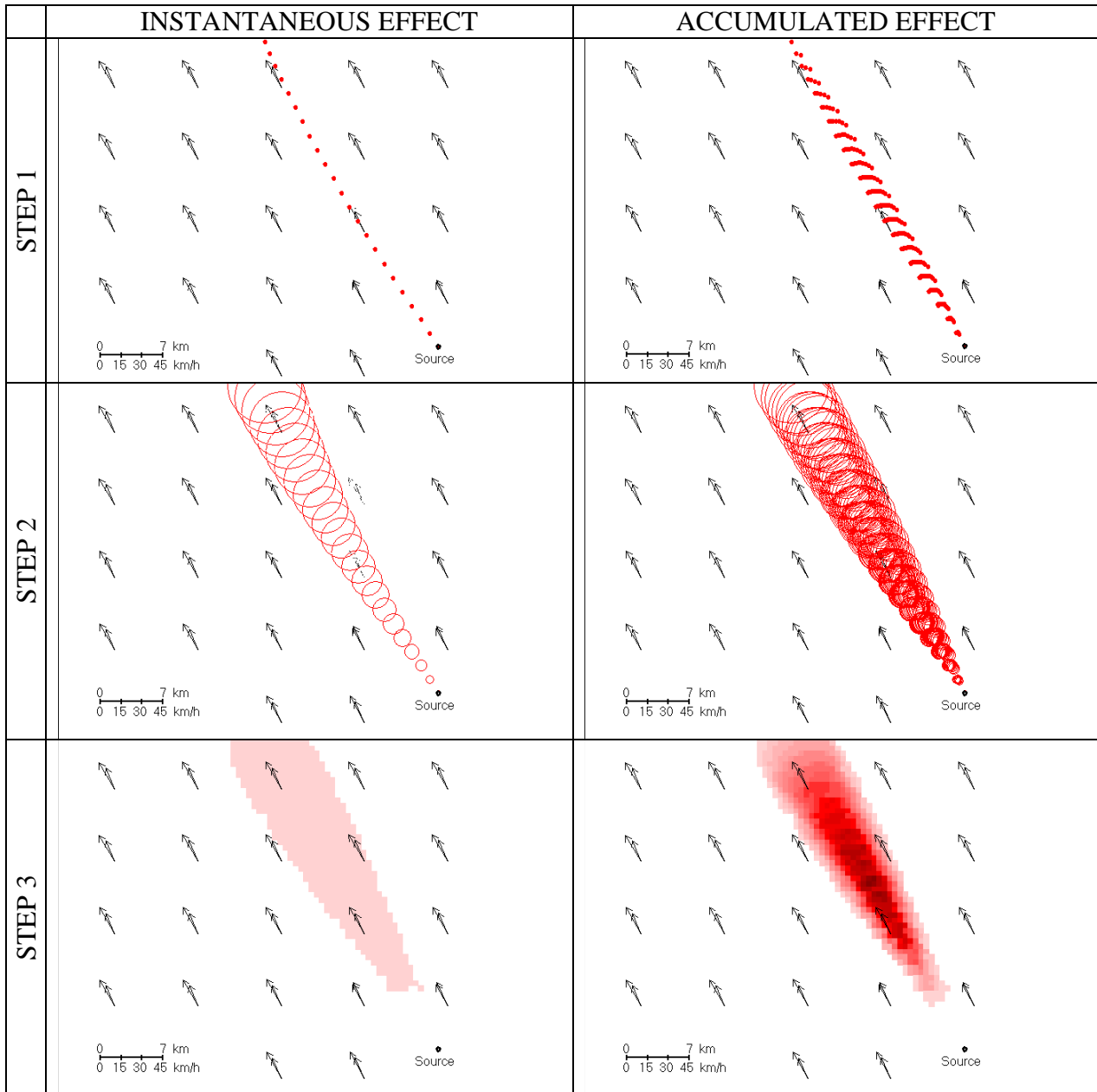

Figure 6: Schematic example of HIRHYLTAD calculations after a 2-hour simulation (5-min time step). Top (step 1): Smoke line set-up at the emission level, using lagrangian trajectories. Center (step 2): Smoke puff dispersion at the emission level, using the Pasquill-Gifford coefficients and Pasquill stability classes. Bottom (step 3): Smoke concentrations at the surface on a high resolution grid. The instantaneous effect is depicted on the left column, and the accumulated effect over the 2-hour period is depicted on the right column. The physical emission height is 120m.

The forecasts of the operational regional Eta/SMN model from the SMN are used as input for HIRHYLTAD. The Eta/SMN model has a horizontal grid spacing of 0.33º (approx. 30km), and

runs twice daily with initial and boundary conditions provided by the GFS model (0000 and 1200 UTC). Time and space interpolation of Eta/SMN outputs are required for the high-resolution HIRHYLTAD simulations. The Cressman scheme (Cressman, 1959), is applied to the original 30-km horizontal resolution grid, in order to obtain a final 5-km grid spacing of HIRHYLTAD. The variables at pressure levels (1000, 975, 950, 925 and 900 hPa) are converted to height levels, and then a linear vertical interpolation is performed to get the following HIRHYLTAD vertical log-linearly spaced levels: 10, 40, 80, 140, 220, 350, 550, 800, 1100 and 1500 m. The Eta/SMN outputs are at 3-hour intervals, so that a linear interpolation in time is applied to match the 5-min time step of HIRHYLTAD.

3.3. Data and methodology

Observed smoke plumes are identified in high resolution imagery from the AQUA and TERRA satellites, which are equipped with the Moderate-Resolution Imaging Spectroradiometer (MODIS) instrument. Visible true-color images (one per day for each satellite) are available from the MODIS Rapid Response System (MRRS) at http://lance.nasa.gov/imagery/rapid-response/.

The AERONET_CEILAP-BA subset is used, as it best covers the study region. The coordinates of the 965km x 720km domain are: 63.7ºW - 53.2ºW, 31.3ºS - 37.8ºS, which is centered over the La Plata River region (see Fig. 4). Because of its high spatial resolution, the MODIS imagery dataset is the most appropriate one for the study of local scale atmospheric phenomena such as smoke dispersion in the atmospheric boundary layer, despite its once daily resolution. The 500-meter resolution MODIS images were systematically cropped to a 467km x 500km domain centered on the La Plata River region (inner rectangle in Fig. 4).

The length of the period of analysis was determined by visual inspection of the smoke distribution in the satellite images, and also considering the fire locations from automated detection algorithms of geostationary and polar-orbiting satellites. For this purpose, two datasets are used, namely, the Global Fire Maps from the MRRS (also available from the aforementioned NASA webpage), with all MODIS detections in 10-day-interval maps, and the 'Banco de Dados Queimadas' from CPTEC-INPE, Brazil (http://www.dpi.inpe.br/proarco/bdqueimadas/; http://sigma.cptec.inpe.br/queimadas/), with more than 20 satellites intervening, including TERRA and AQUA. Based on this information, a 45-day period from 1 April 2008 to 15 May 2008 was finally chosen, totaling 90 MODIS images.

The process followed for selecting the smoke events and digitizing the smoke plumes is completely subjective. During the selection of events the following situations were discarded: no smoke present, excessive cloudiness, sufficiently dispersed smoke with no clear source inside the domain, or well-defined plumes with associated sources outside the domain. A smoke event is defined as one in which a smoke plume is clearly identifiable with at least one associated fire location within the domain, despite the fact that the plume could extend beyond the domain limits. In some cases two smoke events per image were selected, provided that the second plume layout was clearly different from the first one, either in span, length or mean orientation.

A total of 21 events were selected from TERRA images (from 18 days), and 38 events from AQUA images (from 31 days) within the 45-day period of analysis. For each event the plume boundaries were drawn manually, and digitized using a digital-image processing program. No quantitative estimate of the smoke concentration is considered in the process.

The digitization of smoke sources was performed manually, and it revealed that multiple-fire-plumes were common, especially during the uncontrolled fires of 16-20 April 2008. The maximum number of fire locations detected in a single smoke event was 18. The analysis resulted in a total of 165 (103) smoke sources for all the AQUA (TERRA) events. Out of the 59 events, 21 are single-source and 38 are multiple-source smoke plumes. The subjective methodology for determining smoke sources replaced the aforementioned MODIS automated fire detection product, since several cases of false detections and no detections were found.

The digitized fire sources are used as input in the smoke plume simulations with HIRHYLTAD coupled to the Eta/SMN forecasts, for all the events. The HIRHYLTAD simulations assume the dispersion of a passive gaseous substance at a constant emission rate of 1 gr s^{-1}. As mentioned before, the only information available about the fire sources was the satellite imagery with once daily frequency. Since we had no basis for specifying fire duration, each fire source was assumed to remain active during the entire simulation. In the case of multiple-source plumes, simulations are performed individually for each source, and all the contributions of concentrations are added.

The resulting HIRHYLTAD concentrations are vertically-integrated for the subsequent validation with MODIS images. The threshold for establishing the boundaries of the simulated plumes was set equal to zero (i.e., any pixel with a non-zero concentration was considered as part of the smoke plume).
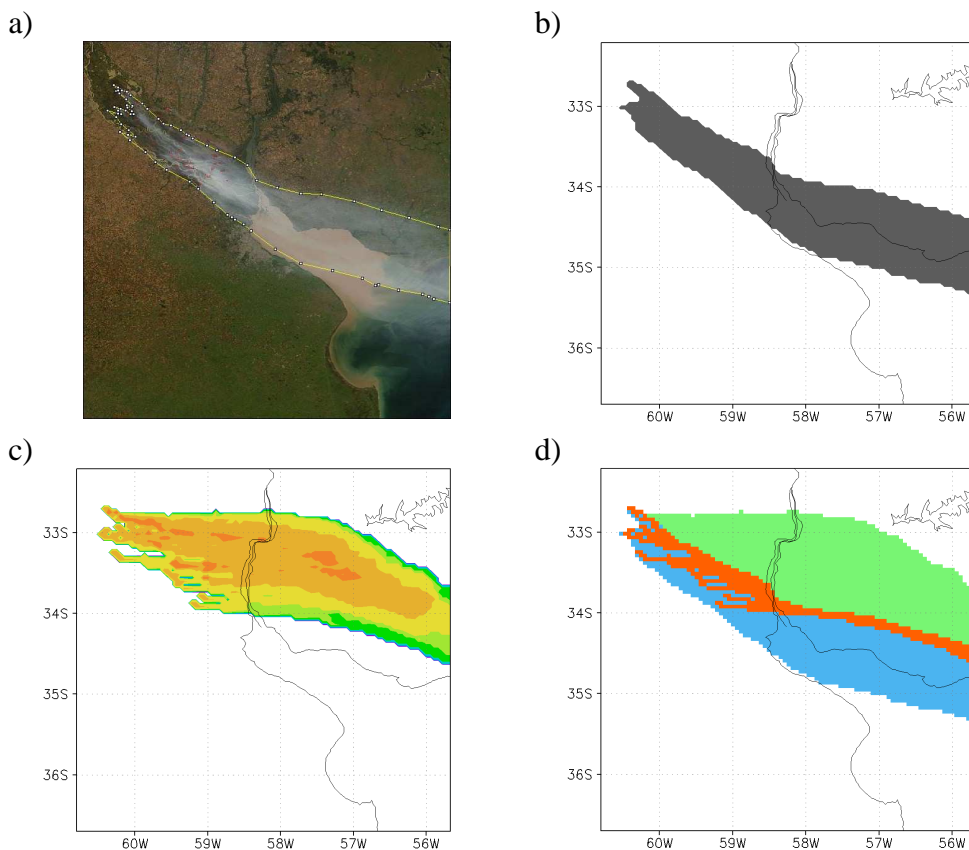
a)

b)

c)

d)

Figure 7: Example of the validation steps for the event of 17 April 2008 at 1430 UTC (satellite: TERRA; number of smoke sources: 18); a) manually-defined smoke plume from the cropped original MODIS image; b) its corresponding digitized image; c) smoke plume simulations with the HIRHYLTAD dispersion model coupled to the Eta/SMN meteorological model (reddish color is indicative of higher concentrations); d) Plume area types as follows: overlap area in red, no detection area in blue, and false detection area in green.

3.4. Results of the case study

A graphical example of the validation steps is shown in Fig. 7, for the 17 April 2008 event, 1430 UTC TERRA image. Panel a) shows the smoke plume boundaries on the MODIS visible image, while panel b) shows its subsequent digitization, for the calculation of the observed area ($Aobs$) and the observed mean direction of propagation ($\overline{dir}_{obs}$) of the plume. The simulation of the plume with HIRHYLTAD is depicted in panel c), with reddish color indicating higher concentrations, from which the forecast area and mean direction of propagation (*Afor* and $\overline{dir}_{for}$, respectively) are calculated; and panel d) shows the area diagram that results from the overlap of the observed and forecast plumes, from which the Plume-Overlap-Area Error is calculated. The associated validation scores for area and direction for this event are: *POAE* = 75.4% and *PMOE* = 12.2%.

The advantage of defining the modified area error *POAE*, and the new direction of propagation error *PMOE*, can be appreciated in Fig. 8. The frequency distribution of *POAE* for the 59 analyzed events is biased towards the lower scores (maximum in the 40-50% category), in contrast with the 1-*FMS* distribution (maximum in the 60-70% category). Also, the low *PMOE* scores (red bars) are highly remarkable: except for three cases, the error in direction of propagation is always smaller than 40% (i.e., an angle less or equal than 72º), while in 63% of the events this error is of less than 10%. Additional statistics of the overall result of all analyzed events are summarized in Table 3, in which we can see that the average overlap error is 51% while in contrast the average error in direction of propagation is only 11.1%. The difference between the two types of error reinforces the importance of considering the direction of propagation for the spatial validation of plume forecasts.
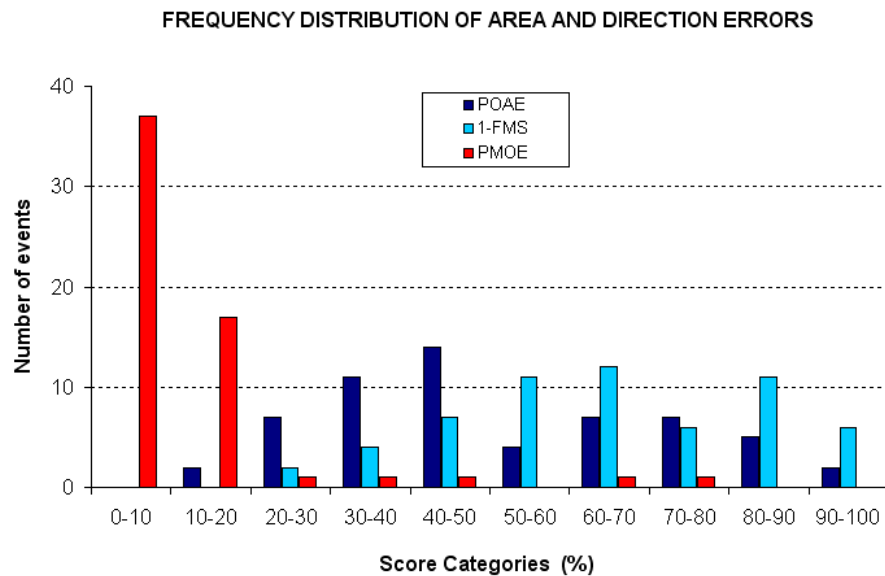

Figure 8: Frequency distribution of overlap area errors *(POAE* and *1-FMS*), and direction of propagation errors (*PMOE*).

Error type	Scores (%)		
	Median	Mean	Standard deviation
Overlap area error *POAE*	46.9	51.0	20.8
Direction of propagation error *PMOE*	7.2	11.1	14.1

Table 3: *POAE* and *PMOE* scores of the 59 analyzed events.

Finally, Fig. 9 shows a comparison between both two-dimensional indices, the *MOE* (Fig. 9a) and our proposed *CDAH* (Fig. 9b). While the *MOE* plot shows more scattered points, indicating variable performance according to the under/over-estimation of overlap area, the *CDAH* plot shows most points confined to the upper portion of the graphic, indicating an overall good agreement between the mean orientation of simulated and observed plumes.

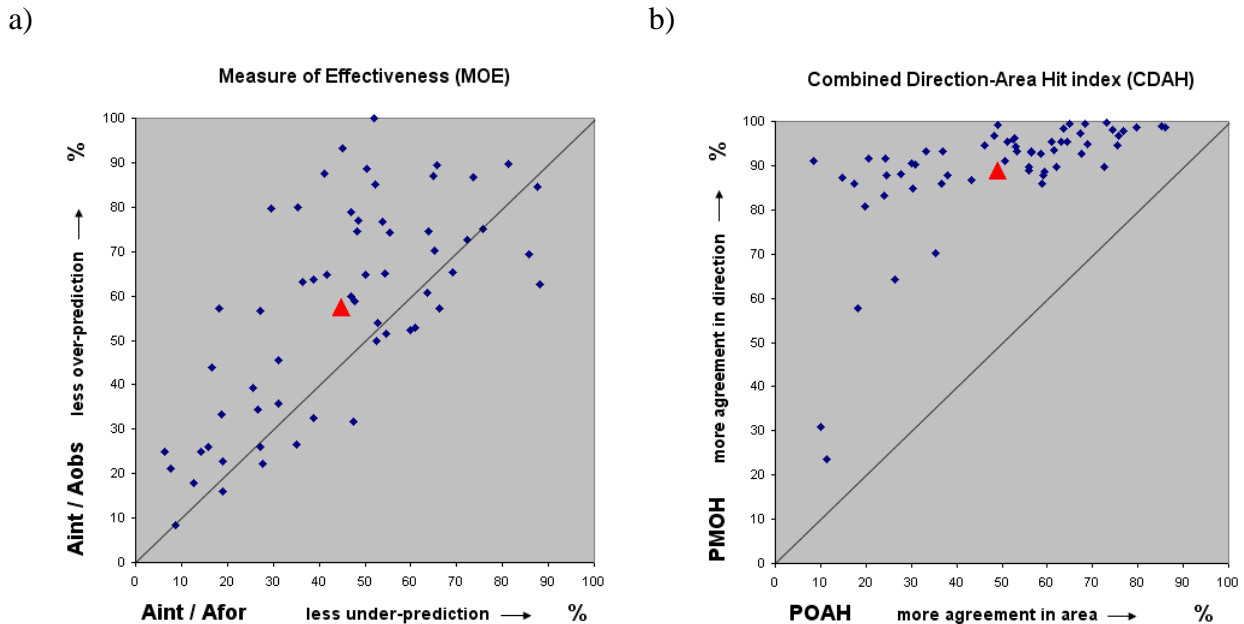a)   b)

Figure 9: Two-dimensional *MOE* and *CDAH* indices for the 59 analyzed events, and mean value of the entire episode (red triangle).

Because of the generally lower errors of *PMOE* in comparison to *POAE*, a natural bias towards the upper-left portion of the plot is expected in the distribution of points within the *CDAH* graphic. Actually, it is remarkable that for our case study, all points are in the upper-left portion of the *CDAH* plot (Fig. 9b), i.e., there was no single event with greater error in direction of propagation than overlap area (*PMOE>POAE*).

An advantage of *MOE* becomes clear when the mean value of the whole distribution is plotted, because it provides the additional information about the possibility of bias towards either under or over-estimation of the observed plumes (distribution of points with respect to the 1:1 line). The results of the case study show a minor bias towards under-estimating the size of the observed plumes, as depicted by the red triangle in Fig. 9a, with a score of (44.96%; 57.43%). This could be due to either to the models employed, the methodology applied or a combination of both.

In contrast, no information about overall under/over-estimation is provided by *CDAH*, and this is because the *x* component of *CDAH* unifies both *x* and *y* components of *MOE*. Nevertheless, the inclusion of the direction of propagation adds a new value to the spatial validation of plume forecasts, which partially compensates for the low scores associated with a qualitatively good performance of the simulation, not clearly evident with *MOE*. Moreover, *CDAH* is more appropriate for situations such as the example of case 2 of Section 2 (please compare Figs. 2 and 3), since the opposite direction of propagation of the forecast plume with respect to the observed plume (i.e., worst possible prediction), would be identified by *CDAH* but not by *MOE*.

4. SUMMARY AND CONCLUSIONS

Two commonly used indices for spatial validation of hazardous plume forecasts are briefly reviewed, namely the Figure of Merit (*FMS)* and the Measure of Effectiveness (*MOE)*; and based on them, new validation indices are proposed and tested with data of a case study.

First, we propose the Plume-Overlap-Area Hit (*POAH*) index, a modification of *FMS*, which we consider appropriate for evaluating plume forecasts. The advantage of *POAH* is that it turns out to be less punitive than *FMS*, which sometimes indicates poor model performance when simple visual inspection indicates an overall good similarity between forecast and observed plumes. We also propose a new index, namely the Plume-Mean-Orientation Hit (*PMOH*) that introduces a new aspect in plume forecast validation by considering the agreement in the mean direction of plume propagation. The plume overlap index *POAH*, and the mean direction of propagation index *PMOH*, are combined in a new two-dimensional Combined-Direction-Area Hit (*CDAH*) index. *CDAH* has the *POAH* index as *x*-component and the *PMOH* index as *y*-component, with individual scores ranging from 0 (worst case) to 100% (best case). The inclusion of the direction of propagation adds a new value to the spatial validation of plume forecasts, sometimes not clearly evident with other commonly used indices.

We apply the new indices to the spatial validation of smoke plume forecasts for a case study of uncontrolled grassfires that took place during April and May 2008 in the La Plata River region in South America. High resolution MODIS imagery from AQUA and TERRA satellites are used for identifying a total of 59 smoke plumes. Two operational models at the Argentine National Meteorological Service, namely the HIRHYLTAD dispersion model and the Eta/SMN meteorological forecast model, are used to simulate the smoke plumes, which are then compared to the observed plumes in the MODIS images.

We conclude that the proposed methodology that employs operational meteorological models and simplified dispersion models can be used to produce reasonably accurate forecasts of the areas affected by the smoke plumes originated in forest and grassland fires, particularly in cases when limited information is available about the fires. The proposed indices are not intended to replace other commonly used indices; instead they add a new value to the spatial validation of smoke plume forecasts by considering the direction of propagation. Finally, we highlight that although the present study is specifically applied to smoke plumes, the validation technique with these indices can be of utility to study pollutant plumes of diverse nature.

16

ACKNOWLEDGEMENTS

REFERENCES

Andreae, M. and Merlet, P., 2001: Emission of trace gases and aerosols from biomass burning, Glob. Biogeochem. Cycle, 15, 955–966.

Berbery, E., Ciappesoni H., Kalnay E., 2008: The smoke episode in Buenos Aires, 15-20 April 2008. Geophys. Res. Let, 35, L21801.

Blanco J., 2011: Desarrollo de un modelo simplificado para el estudio de dispersión de plumas de humo por quema de pastizales, y metodología para la validación con imágenes satelitales. Graduation Thesis at the Department of Atmospheric Sciences, University of Buenos Aires, Sep 2011.

Blanco J. and Berri G., 2011: A simplified model for the study of smoke plume dispersion from grassfires and a methodology for forecasts validation with satellite images. Ninth Symposium on Fire and Forest Meteorology, Palm Springs, CA, Oct 2011.
http://ams.confex.com/ams/9FIRE/webprogram/Manuscript/Paper192583/22_manuscript.pdf

Boybeyi, Z.N.; Ahmad N.; Bacon D.P.; Dunn T.J.; Hall M.S.; Lee P.C.S.; Sarma R.A.; Wait T.R., 2001: Evaluation of the Operational Multi-scale Environment model with Grid Adaptivity against the European Tracer Experiment, J. Applied Meteor, 40, 1541-1558.

Briggs, G.A., 1969: Plume rise. TID-25075. USAEC Critical Review Series, National Technical Information Service, Springfield, VA, 81 pp.

Brost, R. A., P. L. Haagenson, and Y.-H. Kuo, 1988: Eulerian simulation of tracer distribution during CAPTEX. J. Appl. Meteor., 27, 579–593.

Byrne, Marc A., Arlene G. Laing, Charles Connor, 2007: Predicting Tephra Dispersion with a Mesoscale Atmospheric Model and a Particle Fall Model: Application to Cerro Negro Volcano. J. Appl. Meteor. Climatol., 46, 121–135.

Clarin newspaper editions, Argentina, April and May 2008, http://www.clarin.com

Clements C.D and co-authors, 2007: Observing the dynamics of wildland grass fires: FireFlux – a field validation experiment. Bulletin of the American Meteorological Society 88(9), pp.1369-1382.

Cressman G.P., 1959: An operative objective analysis scheme, Mon. Wea. Rev., 87, 367–374 pp.

Davis, C. A., B. G. Brown, and R. G. Bullock, 2006: Object-based verification of precipitation forecasts, Part I: Methodology and application to mesoscale rain areas. Mon. Wea. Rev. 134, 1772 - 1784.

Davis, Christopher A., Barbara G. Brown, Randy Bullock, John Halley-Gotway, 2009: The Method for Object-Based Diagnostic Evaluation (MODE) Applied to Numerical Forecasts from the 2005 NSSL/SPC Spring Program. Wea. Forecasting, 24, 1252–1267.

Gifford F., 1976: Turbulent Diffusion Typing Schemes. A Review, Nucl. Saf., 17: 68-86.

La Nacion newspaper editions, Argentina, April and May 2008, http://www.lanacion.com.ar

Lujanienė G., S. Byčenkienė, P. P. Povinec, M. Gera, 2012: Radionuclides from the Fukushima accident in the air over Lithuania: measurement and modelling approaches. J Environ Radioact. (in press) http://www.happysmile.se/dokument/science_plutonium_litauen.pdf

Marcuzzi E. A. and Hoevel R., 2009: Condiciones Meteorológicas en torno a los incendios ocurridos en el delta del Paraná durante el mes de abril de 2008. Reprints X Congreso Argentino de Meteorología, Buenos Aires, Oct 2009.

Mattio C. A., 2009: Combinación de herramientas para el monitoreo y seguimiento de humo generado por incendios forestales y de pastizales en la República Argentina. Reprints X Congreso Argentino de Meteorología, Buenos Aires, Oct 2009.

Mosca, S.G.; Graziani, W.; Klug, R.; Bellasio; Bianconi R., 1998: A statistical methodology for the evaluation of long-range dispersion models: An application to the ETEX exercise. Atmos. Environ, 32, 4302–4324.

Pasquill, F., 1961: The estimation of the dispersion of windborne material. The Meteorological Magazine, vol 90, No. 1063, pp 33-49.

Petty, R., 2000: User requirements for dispersion modeling. Proc. Workshop on Multiscale Atmospheric Dispersion Within the Federal Community, Silver Spring, MD, Office of the Federal Coordinator for Meteorological Services and Supporting Research, 1-1–1-3.

Piedelievre, J. P., L. Musson-Genon, F. Bompay, 1990: MEDIA—An Eulerian Model of Atmospheric Dispersion: First Validation on the Chernobyl Release. J. Appl. Meteor., 29, 1205–1220.

Rolph, G. D., and Coauthors, 2009: Description and verification of the NOAA Smoke Forecasting System: The 2007 fire season. Wea. Forecasting, 24, 361–378.

Ruminski, M., S. Kondragunta, R. Draxler, and J. Zeng, 2006: Recent changes to the Hazard Mapping System. 15th Internacional Emission Inventory Conf.: Reinventing Inventories—New Ideas in New Orleans, New Orleans, LA, EPA.
http://www.epa.gov/ttn/chief/conference/ei15/session10/ruminiski.pdf

Ruminski, M., S. Kondragunta, R. Draxler and G. Rolph, 2007: Use of environmental satellite imagery for smoke depiction and transport model initialization.16th Annual International Emission Inventory Conf.: Emission Inventories—Integration, Analysis, and Communications, Raleigh, NC, EPA
http://www.epa.gov/ttn/chief/conference/ei16/session10/ruminski.pdf

Schultz, M. G., 2002: On the use of ATSR fire count data to estimate the seasonal and interanual variability of vegetation fire emissions, Atmos. Chem. Phys., 2, 387–395.
http://www.atmos-chem-phys.net/2/387/2002/acp-2-387-2002.pdf

Spektor D. M., 1998: Oil well fires. RAND's National Defense Research Institute, Office of the Secretary of Defense. Volume 6, The Gulf War Illnesses Series.
http://www.rand.org/pubs/monograph_reports/2005/MR1018.6.pdf

Stein, A. F., G. D. Rolph, R. R. Draxler, B. S. Stunder, and M. S. Ruminski, 2009: Verification of the NOAA smoke forecasting system: Model sensitivity to the injection height. Wea. Forecasting, 24, 379–394.

van der Werf, G. R., and Coauthors, 2010: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997-2009). Atmos. Chem. Phys., 10, 11707-11735.
http://www.atmos-chem-phys-discuss.net/10/16153/2010/acpd-10-16153-2010-print.pdf

Warner S., Platt N.; Heagy J.F, 2004: User-oriented two-dimensional measure of effectiveness for the evaluation of transport and dispersion models. J. Appl. Meteor., 43, 58–73.

Wernli, H., M. Paulat, M. Hagen and C. Frei, 2008: SAL--a novel quality measure for the verification of quantitative precipitation forecasts. Mon. Wea. Rev., 136, 4470 – 4487

Zhu, M., V. Lakshmanan, P. Zhang, Y. Hong, K. Cheng, and S. Chen, 2011: Spatial verification using a true metric. Atmos. Res., 102 (4), 408 – 419.