Metrics for vague spatial objects based on the concept of mass

Arta Dilo, Rolf A. de By, Alfred Stein

***Abstract*— Many spatial phenomena exhibit vagueness. Representation of such phenomena requires vague objects. In previous work, we provided definitions for vague objects: vague points, vague lines, and vague regions. Each of these objects is presented as a fuzzy set in $\mathbb{R}^2$ that satisfies well-defined properties. In this paper, we propose a number of geometric measures for vague objects, using the concept of mass distribution. The membership function of a vague object can be seen as a mass distribution. According to this view, a crisp object is a body with constant density, and a vague object is a body of varying density. We provide mathematical definitions for length of a vague line, area of a vague region, centroid of a vague object, as well as a measure for the vagueness of an object. The length of a vague line and the area of a vague region are indeed the mass of the vague line and of the vague region, respectively. Both metrics give an average of the values of the corresponding crisp metric on the α-cuts of the vague object. The centroid of a vague object is its centre of mass associated with a membership degree. The last metric functions as a measure of the degree of vagueness for a vague object.**

I. Introduction

Many natural phenomena exhibit vagueness. This is basically the existence of borderline cases. For example, the boundary between two vegetation classes is often a transition zone instead of a sharp line. That means, there are locations for which we cannot decide with full certainty whether the vegetation belongs to one class or to another. A proper understanding and modelling of these phenomena requires representation of vagueness in spatial information and reasoning under vagueness. There are two principal views on the root cause of vagueness: the first sees vagueness as an inherent property of the phenomena, the second considers it to be linguistic [12]. Many–valued logics, of which fuzzy logic is the most widely used, offer a solution to the first view, considering vagueness as a matter of degree. We subscribe to the first view of vagueness, and use fuzzy theory to handle it.

Current GIS and spatial database systems are constructed under the assumption that spatial objects are crisp. They offer types for representing crisp objects, and operators to perform analysis and reasoning over crisp objects. In previous work [7], [9], we have defined types for vague objects, together with operators like set operators and spatial relations, and have implemented these in an open GIS software package [8]. Object types that we distinguish are vague points, vague lines, and vague regions. Any of these objects is presented as a fuzzy set μ in $\mathbb{R}^2$ that satisfies specific properties. This paper provides metrics for vague objects: average length *AvLength* for a vague line, average area *AvArea* for a vague region, *Centroid* and *Vagueness* for a vague object of any type. We consider the membership function μ of a vague object as a mass distribution over the object extent, i.e. its support set, $supp(\mu)$. Metrics for vague objects are calculated from the concept of body mass. The length *AvLength* of a vague line, and the area *AvArea* of a vague region are calculated as the body mass of the vague object. The *Centroid* of a vague object is its centre of mass, associated with a membership value, calculated as an average of object membership values. The last, *Vagueness*, is a measure for the vagueness of an object, calculated from the mass of the vague object and the mass of its extent.

The paper is structured as follows. Section II summarizes previous work on metrics for fuzzy sets in general, and metrics for fuzzy spatial objects. Section III describes the object types that we work with: vague points, vague lines, and vague regions. Section IV provides metrics definitions: length, area, centroid, and vagueness degree. Section V illustrates these metrics in an example application from air quality in the Netherlands. Section VI closes the paper with discussions and conclusions. At different places we need to use or compare our metrics to geometric measures of crisp objects. We use different fonts to distinguish vague types and metrics from their crisp analogues: *LucidaCasual-Italic* is used for the names of vague types and metrics, LucidaCasual is used for crisp metrics.

II. Previous work

Several metrics have been proposed in fuzzy set theory, and some more specific metrics are proposed in fuzzy image processing research. Distance and fuzziness measures are discussed in fuzzy theory, whereas fuzzy image processing offers geometric measures for vague objects. Different metrics have been proposed for the distance between fuzzy sets [3], [5], [6], [10], [14]. There are also different proposals for geometric measures on vague objects, like area, perimeter, diameter [4], [13], [15], [16]. Here we discuss first the fuzziness measures offered by fuzzy theory, and then discuss the proposals for the geometric measures length and area.

Bandemer and Gottwald [2] presented two fuzziness measures, *entropy* and *energy*, for fuzzy sets in a finite space X, summarizing definitions given by different authors. The *entropy* evaluates the deviation of a fuzzy set from a crisp set. The entropy for a crisp set equals 0, whereas a maximum entropy is reached by a fuzzy set μ if every location x has

Arta Dilo is with the GIS-technology section, OTB, Delft University of Technology, Jaffalaan 9, 2628 BX Delft, The Netherlands (email: a.dilo@tudelft.nl).

Rolf A. de By is with the Department of GeoInformation Processing, International Institute for Geo-Information Science and Earth Observation (ITC), Hengelosestraat 99, 7500 AA Enschede, The Netherlands (email: deby@itc.nl).

Alfred Stein is with the Department of Earth Observation Science, ITC, The Netherlands (email: stein@itc.nl).

1-4244-1210-2/07/$25.00 ©2007 IEEE.

a value $\mu(x) = 0.5$. Three *entropy* measures are presented. Entropy of a fuzzy set $\mu(x)$ is calculated as the maximum value of the intersection of μ with its complement $1-\mu$ over all locations x: $F_1(\mu) = \max\{(\mu \sqcap (1-\mu))(x) \mid x \in X\}$. A second measure for entropy uses the cardinal instead of the maximum value of intersection. For a set X, the cardinal, denoted $\mathsf{card}(X)$, is the number of its elements. The cardinal of a fuzzy set is $\mathit{card}(\mu) = \sum_{x \in X} \mu(x)$. The entropy of a fuzzy set is calculated as: $F_2(\mu) = 2\,\mathit{card}\big(\mu \sqcap (1-\mu)\big) / \mathsf{card}(X)$, where the factor 2 is included to normalize the values of F_2. The third entropy measure is an analogue of the uncertainty measure from probability theory, the Shannon's entropy: $F_3(\mu) = -c \sum_{x \in X} \big(\mu(x) \ln \mu(x) + (1-\mu(x)) \ln (1 - \mu(x))\big)$. The two *energy* measures presented in [2] are also built from the cardinal of a fuzzy set and its maximum value: $E_1(\mu) = \mathit{card}(\mu)$, and $E_2(\mu) = \max\{\mu(x) \mid x \in X\}$.

Proposed geometric measures for vague objects can be divided into two groups: metrics that are functions and metrics that are numbers. Metrics of the first group associate every membership degree $\alpha \in (0, 1]$ with a positive number that is a measure for that degree. For example, Schneider [16] proposes an area measure for a vague region that associates with every α the area of the α-cut of the region. That same idea is applied in [16] for the length of a vague line, as well as for other metrics, like perimeter and diameter.

Metrics of the second group are positive numbers, calculated by integrating over the membership function μ of a vague object. Rosenfeld and Haber [13], [15] calculate the *area* of a vague region μ as the volume between the surface of the function $\mu(x, y)$ and the x, y plane: $\mathit{area}(\mu) = \iint \mu(x, y)\, dx\, dy$. They propose similar formulas for the calculation of other metrics for vague regions: height, width, perimeter, and diameter. Bogomolny [4] modifies their formulas so that they satisfy known interrelations for crisp regions, e.g., the isoperimetric inequality used for characterizing the shape of a region. The area is modified to the integral of the root of μ: $\mathit{area}(\mu) = \iint \mu^{1/2}(x, y)\, dx\, dy$, and other measures are modified similarly. Schneider [16] proposes some additional measures: length and strength for vague lines, and elongatedness and roundedness for vague regions. He calculates the length of a vague line μ from the integral of $\sqrt{\left(\frac{\partial \mu^{1/2}}{\partial x}(x, y)\right)^2 + \left(\frac{\partial \mu^{1/2}}{\partial y}(x, y)\right)^2}$ over the line extent $\mathit{supp}(\mu)$. Measures in common with the previous papers, e.g., area of a vague region, correspond to Bogomolny's measures.

III. Types for vague objects

Vague objects are represented in this study by fuzzy sets in $\mathbb{R}^2$ with specific properties. The collection of fuzzy sets in $\mathbb{R}^2$ is denoted by $\mathcal{F}(\mathbb{R}^2)$. We distinguish between vague points, vague lines and vague regions, each represented by a separate type. These types are *VPoint*, *VLine*, and *VRegion*. They represent objects that are indivisible into components. A vague point is an object of type *VPoint*, a vague line is of type *VLine*, and a vague region is of type *VRegion*. Membership values of a vague object range from 0 to 1 inclusive, often covering the whole range $[0, 1]$. A vague object can also have a finite set of membership values. A crisp object is a special case of a vague object; its membership values are in the set $\{0, 1\}$.

A *vague point* is a site with a known location, but with uncertain membership to a phenomenon of interest. Figure 1(a) illustrates a vague point. Densely populated residential centres are examples of vague points. For each residential centre we know the location precisely, but the density level is a matter of degree. A vague point is defined as a fuzzy set that has positive value at only one location (x, y). This value gives the degree of membership of the site (x, y) to the phenomenon of interest. The set of vague points is

$$\mathit{VPoint} \equiv \Big\{\mu \in \mathcal{F}(\mathbb{R}^2) \mid \exists!(x, y) \in \mathbb{R}^2, \mu(x, y) > 0\Big\}.$$

If the membership value $\mu(x, y)$ is equal to 1, then μ is a crisp point.

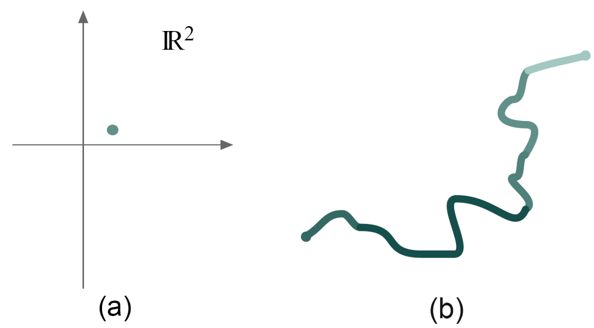

Fig. 1. Vague objects: (a) a vague point, (b) a vague line. Colour saturation is used to show membership values: full saturation correspond to membership value 1, low saturation shows low membership values.

A *vague line* is a linear feature with crisply defined extent, but uncertain membership to some phenomenon of interest for the points on the line. Figure 1(b) illustrates a vague line. It is a simple curve (crisp line) with mostly gradual transitions of membership values between neighbor points on the line. Membership values are positive at every location on the line, except, perhaps, at the end nodes. Stepwise changes of membership values may occur along the line. A traffic congestion on a road network is an example application for vague lines. We know precisely where a road is, but the congestion is a matter of degree. Part of the road is completely blocked, and hence certainly belongs to the traffic congestion, whereas away from the congestion, the car build-up becomes less severe.

A vague line is built from a crisp line by applying a membership function over locations on the crisp line. A crisp line is a continuous, non-self intersecting curve, but possibly closed. The membership function over the crisp line is required to be almost everywhere continuous, allowing stepwise changes in only a finite number of locations. Figure 2(a) is an example of a possible membership function for a vague line. We build a fuzzy set in $[0, 1]$ that satisfies the continuity properties for a membership function, then transfer its membership values to the crisp line via a homeomorphism h. Figure 2(b) illustrates the construction of a vague line from a fuzzy set η in $[0, 1]$, of which the function graph is given in Figure 2(a).

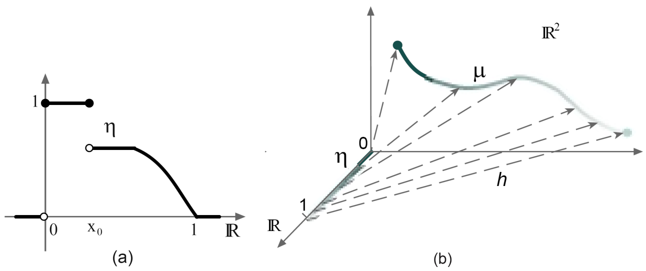

Fig. 2. Vague line construction: (a) membership function η for the vague line, (b) vague line built by transferring the membership values of the fuzzy set η via the homeomorphism h. Observe how membership degrees in η are carried over to μ.

The set of vague lines is defined as

$$\mathit{VLine} \equiv \Big\{\mu \in \mathcal{F}(\mathbb{R}^2) \mid \exists \eta \in \mathcal{F}([0,1]), \eta = \overline{\eta^\circ}, \\ \eta \text{ connected} \\ \exists h : [0,1] \to \mathbb{R}^2 \text{ homeomorphism in } (0,1) \\ \text{and continuous in } \{0,1\}, \\ \mu = \tilde{h}(\eta) \text{ and } \big(h(0) = h(1) \Rightarrow \eta(0) = \eta(1)\big)\Big\}.$$

The first part of the definition imposes conditions for the membership function of the vague line. The regular closure $\eta = \overline{\eta^\circ}$ assures the continuity condition, connectedness assures continuous extent of η over the whole $[0, 1]$. The second part formulates conditions for the crisp line. This is topologically equivalent to the unit interval $[0, 1]$, i.e., there is a homeomorphism h from $[0, 1]$ to the line in $\mathbb{R}^2$, and the line is $l = \{h(t) = \big(x(t), y(t)\big) \mid t \in [0, 1]\}$. To allow closed lines, the homeomorphism is restricted to $(0, 1)$, requiring continuity at the end points 0 and 1. The third part of the formula builds the vague line μ as the image via homeomorphism h of the connected regular fuzzy set η in $[0, 1]$: $\mu = \tilde{h}(\eta)$. When the vague line is closed, the membership values at both end nodes must be equal. If the fuzzy set η has a constant membership value 1 in the interval $[0, 1]$, then the vague line is a crisp line.

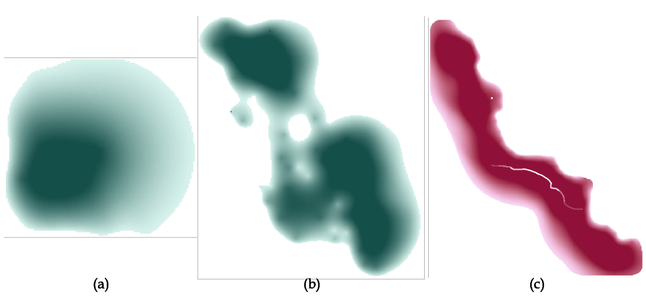

Fig. 3. Fuzzy sets in $\mathbb{R}^2$: (a) and (b) vague regions, (c) a fuzzy set that is not a vague region.

A *vague region* is a single-component fuzzy set that does not have irregularities, like isolated vague points and vague lines, spikes, or punctures and cuts, i.e. removed vague points and vague lines, respectively. The fuzzy set of Figure 3(c) has a puncture and a cut, both being irregularities that are not allowed for a vague region object. The membership values of a vague region change mostly gradually between neighbor points in the region. Stepwise jumps can also happen between sides of a line that is inside the region. Figures 3(a) and 3(b) illustrate vague regions. Air quality is a vague concept. It is related to the concentration of pollutants like ozone, NOx [18], [19]. High polluted areas (therefore of poor air quality) are examples of vague regions. Some locations are certainly polluted, i.e., of poor air quality, whereas others can be considered polluted or not; they are polluted to some degree.

The set of vague regions is then defined as

$$\mathit{VRegion} \equiv \Big\{\mu \in \mathcal{F}(\mathbb{R}^2) \mid \mu \text{ bounded}, \mu = \overline{\mu^\circ}, \mu^\circ \text{ connected}\Big\}.$$

A vague region is bounded, meaning its support set is bounded. The regular closure in $\mathbb{R}^2$ assures continuity properties for the membership function: there can be stepwise jumps along lines, i.e., vertical cliffs on the surface, but no isolated discontinuities are allowed. Connected interior assures the fuzzy set to be single-component. The highest membership value may be less than 1. A crisp region is a specific case of a vague region, when μ has a constant membership equal to 1.

Type *SVSpatial* is the supertype of *VPoint*, *VLine*, and *VRegion*. That is, any simple vague object, a vague point, a vague line, or a vague region, is of the type *SVSpatial*. This last type is useful in the definition of centroid and vagueness degree measures below.

IV. Metrics for Vague Objects

The membership function μ of a vague object can be seen as a mass distribution. A crisp object is a body with constant density, whereas a vague object has a varying density. The total mass of a body consisting of a finite set of locations is calculated from the sum of masses at every location. A total mass distributed over a line is calculated from the integral of the density function over the line extent: $\int_{supp(\mu)} \mu(x, y)\, dl$. The mass distributed over an area is calculated from the integral of the density function over the area: $\int_{supp(\mu)} \mu(x, y)\, dA$ [1]. Metrics presented in this section are based on this concept of mass.

A. A length measure for vague lines

Suppose a vague line is built from a curve l given by the parametric equation $\big(x(t), y(t)\big)$, and a membership function η. The mass of the vague line is

$$\mathit{Mass}(\mu) = \int_0^1 \eta(t) \sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}\, dt. \quad (1)$$

If μ is a crisp line, its membership function is constant, i.e., η has the constant value 1 over the whole interval $[0, 1]$. The mass of the crisp line μ is

$$\mathit{Mass}(\mu) = \int_0^1 \sqrt{\left(x'(t)\right)^2 + \left(y'(t)\right)^2}\, dt.$$

This equals the length of the curve l given by the parametric equations.

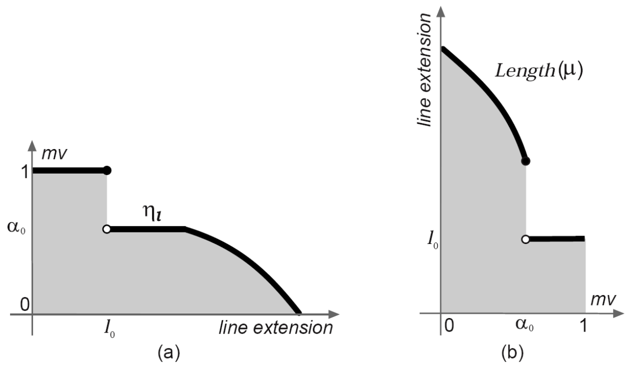

Fig. 4. Two equal areas presenting (a) the mass of a vague line, (b) the average length of a vague line.

The area shown in grey in Figure 4(a) is the integral of $\eta(l)\, dl$ over the line extent, i.e., the mass of the vague line. Figure 4(b) shows the graph of the function $\mathit{Length}(\mu)$, which, for every α in $(0, 1]$, returns the length of the α-cut $\mu_\alpha = \{p \in \mathbb{R}^2 | \mu(p) \geq \alpha\}$. The area in grey in Figure 4(b) is the integral of $\mathit{Length}(\mu)$ over $[0, 1]$. The two integrals present two different ways of calculating the same area. The average value of an integrable function $f(x)$ on an interval $[a, b]$ is given by the formula $\overline{f} = \frac{1}{b-a} \int_a^b f(x)\, dx$ [11]. Based on this property, the integral of $\mathit{Length}(\mu)$ over $[0, 1]$ is the average of all α-cut lengths. From the equality of integrals it follows that the mass of the vague line μ is the average of the lengths of its α-cuts. We define the length measure on vague lines, *AvLength* (for average length), as:

$$\mathit{AvLength} : \mathit{VLine} \rightarrow \mathbb{R}^+$$

$$\forall \mu \in \mathit{VLine},\ \mathit{AvLength}(\mu) = \mathit{Mass}(\mu).$$

When μ is a crisp line, *AvLength* returns the same value as the (crisp) length measure.

B. An area measure for vague regions

The mass of a vague region is

$$\mathit{Mass}(\mu) = \iint \mu(x, y)\, dx\, dy. \tag{2}$$

If μ is a crisp region, the membership function has a constant value 1 over the whole extent, and $\mathit{Mass}(\mu) = \iint_{supp(\mu)} dx\, dy$ equals the area of the extent of μ, $\mathsf{Area}(supp(\mu))$.

Figure 5 applies membership values as the third coordinate to obtain the surface in 3D. Memberships are also used for colouring: low saturation indicates low membership value. The mass of the vague region is the volume below the μ–surface. We define a function $\mathit{Area}(\mu)$ that, for every α in $[0, 1]$, returns the area of the α-cut: $\mathsf{Area}(\mu_\alpha) = \iint_{\mu_\alpha} dx\, dy$.

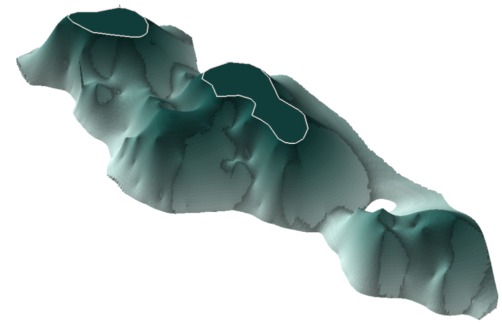

Fig. 5. A vague region shown in 3D with membership values as the third coordinate. Membership values are also used for colour saturation.

The integral of $\mathit{Area}(\mu)$ over $[0, 1]$, $\int_0^1 \mathsf{Area}(\mu_\alpha)\, d\alpha$, produces the volume of Figure 5. This integral is an average of areas of all α-cuts of the vague region μ. Thus, this average is equal to the mass of the vague region. Based on this equality, we name the area measure *AvArea*, for average area. It is defined as

$$\mathit{AvArea} : \mathit{VRegion} \rightarrow \mathbb{R}^+$$

$$\forall \mu \in \mathit{VRegion},\ \mathit{AvArea}(\mu) = \mathit{Mass}(\mu).$$

When μ is a crisp region, *AvArea* returns the same result as the crisp Area measure.

C. The centroid of a vague object

The centroid of a crisp object, i.e., the *centre of mass*, is defined from the mass of the object and its moments [11]. We first show how mass and moments are translated to vague object types, considering their membership function to be the mass density function.

For a body μ consisting of a finite set of locations, the mass and moments, M_x for the x direction, M_y for the y direction, are

$$\mathit{Mass}(\mu) = \sum_{p \in supp(\mu)} \mu(p), \tag{3}$$

$$M_x(\mu) = \textstyle\sum_{(x,y) \in supp(\mu)} y\, \mu(x, y),$$

$$M_y(\mu) = \textstyle\sum_{(x,y) \in supp(\mu)} x\, \mu(x, y).$$

A vague point has only one location, therefore its mass and moments are written more simply as $\mathit{Mass}(\mu) = \mu(x, y)$, $M_x(\mu) = y\, \mu(x, y)$, and $M_y(\mu) = x\, \mu(x, y)$, where (x, y) is the only location with positive membership value.

The mass of a vague line is provided in equation 1. The moments of a vague line are

$$M_x(\mu) = \int_0^1 y(t)\, \eta(t) \sqrt{\left(x'(t)\right)^2 + \left(y'(t)\right)^2}\, dt,$$

$$M_y(\mu) = \int_0^1 x(t)\, \eta(t) \sqrt{\left(x'(t)\right)^2 + \left(y'(t)\right)^2}\, dt.$$

Mass of a vague region has been provided (eq. 2). Moments of a vague region are

$$M_x(\mu) = \iint y\, \mu(x, y)\, dx\, dy, \quad M_y(\mu) = \iint x\, \mu(x, y)\, dx\, dy.$$

The centroid of a vague object is a vague point. Its coordinates are calculated from its mass and moments as the coordinates of the centre of mass. Its membership degree is the ratio between the mass of the object and the mass of its support set. If the object μ is a vague point, the Mass of its support set is 1, as it is the cardinal of a single element set. Mass of the support set of a vague line is the length of its extent, while for a vague region it is the area of the region extent. Using mass functions, and moments we can define *Centroid* as:

$$\mathit{Centroid} : \mathit{SVSpatial} \to \mathit{VPoint}$$

$$\forall \mu \in \mathit{SVSpatial},$$

$$\mathit{Centroid}(\mu) = \left(\frac{M_y(\mu)}{\mathit{Mass}(\mu)}, \frac{M_x(\mu)}{\mathit{Mass}(\mu)}, \frac{\mathit{Mass}(\mu)}{\mathrm{Mass}(supp(\mu))} \right).$$

The assumption is that mass and moment functions will be replaced accordingly to the type of the object they operate with. Centroid returns a crisp point when applied to a crisp object.

D. The vagueness degree of an object

We propose a metric *Vagueness* to measure the degree of vagueness for a vague object. It takes a value in the unit interval $[0, 1]$. A low value indicates low vagueness, and high value signals high vagueness. The metric is determined as the 1-complement of the ratio of the object's mass and the mass of its support set. Vagueness of a vague point is calculated as $\mathit{Vagueness}(\mu) = 1 - \mu(x, y)$, where (x, y) is the location with positive membership. For a vague line, it is $\mathit{Vagueness}(\mu) = 1 - \mathit{AvLength}(\mu)/\mathrm{Length}(supp(\mu))$, and for a vague region it is $\mathit{Vagueness}(\mu) = 1 - \mathit{AvArea}(\mu)/\mathrm{Area}(supp(\mu))$. Using again mass functions, to be replaced according to the object type, we define our *Vagueness* as:

$$\mathit{Vagueness} : \mathit{SVSpatial} \to [0, 1]$$

$$\forall \mu \in \mathit{SVSpatial}, \ \mathit{Vagueness}[\mu] = 1 - \frac{\mathit{Mass}(\mu)}{\mathrm{Mass}(supp(\mu))}.$$

Value 0 is reached when the object is crisp. A maximal value is reached when the 1-cut of an object is empty, or it has a lower dimension than the object.

V. An Application Example

Air quality is an important issue for human health. International and national projects [17] are working on setting up measures to improve air quality. In the Netherlands, it is the task of Milieu en Natuurplanbureau (MNP, the Netherlands Environmental Assessment Agency) to measure and model the air quality. Measurements for different pollutants like NO_2, ozone, etc., are performed at different stations of the air quality monitoring network. The operational priority substances (OPS) dispersion model provides information for NO_2 large scale concentration. Measurements along roads are used to add local concentration caused by traffic to the large scale concentration maps. For this example, we use a constant value $18\mu g/m^3$ over all the highways. Local concentration are calculated by linearly decreasing this value with the distance from a highway, reaching 0 at 3km away. Figure 6 shows large-scale and total NO_2 concentration in 2001.

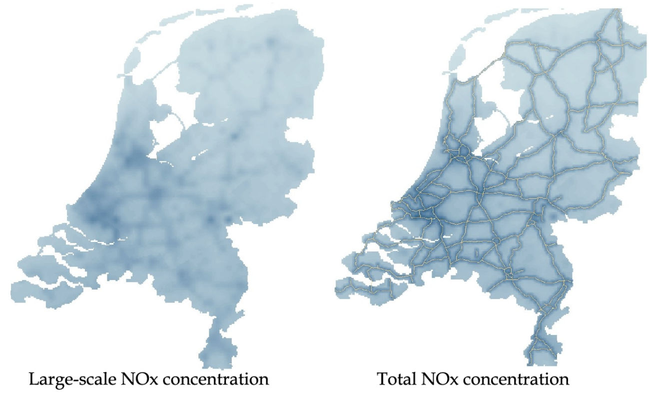

Fig. 6. NO_2 concentration in 2001: large-scale (left), total concentration (right). Darker colour shows higher concentration.

A concentration of $40\mu g/m^3$ is a limit value for air quality [17]. A range about this value is used as transition boundary between high and low concentration regions. To adjust for the uncertainties of the OPS model, we consider the standard deviation of the OPS output to build the transition boundary: $40 \pm 2\sigma$. Figure 7 is the graph of the membership function for high concentration NO_2 in 2001.

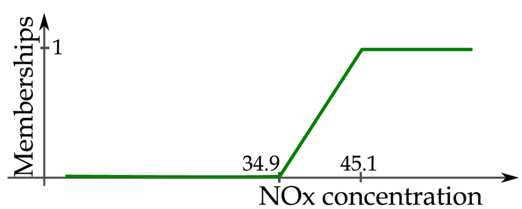

Fig. 7. Membership function for high NO_2 concentration in year 2001.

We have OPS outputs for years 2000–2006, converted to ArcGIS raster data. Total concentration raster maps are created from these, adding local concentration as described above (constant through years). Membership functions are applied to yearly rasters of total concentration to extract high concentration regions for every year. Figure 8 shows high concentration regions for years 2001 and 2006.

It is important to know how air quality changes through years [19]. An indicator for improvement of air quality through years may be the decrease of area size of polluted regions. *AvArea* metric can be used to measure the area of high concentration regions on different years. Vague regions of high NO_2 concentration are created separately from large-scale and total concentration data. The average areas for these regions are calculated with a VisualBasic script in ArcGIS, and results are plotted separately in Figure 9. It may be interesting to see how the centre of pollution is moving. *Centroid* metric on vague regions produces the centre of mass of a vague region. Applying it to regions of high concentration in consecutive years will give an indication on the direction of movement of air pollution.

For this example we considered the NO_2 concentration to be constant along highways. Van Breugel [17] models the

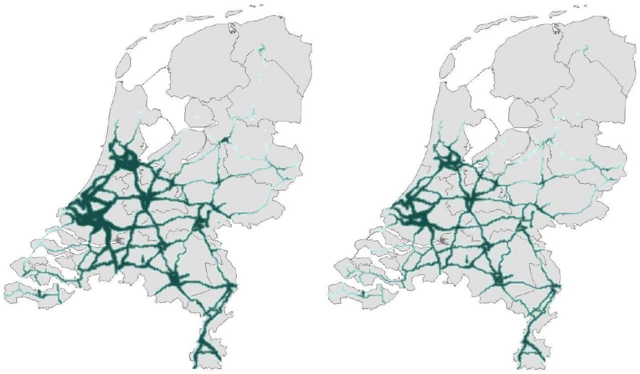

Fig. 8. Objects of high NO_2 concentration in 2001 and 2006. Darker colour shows higher membership value.

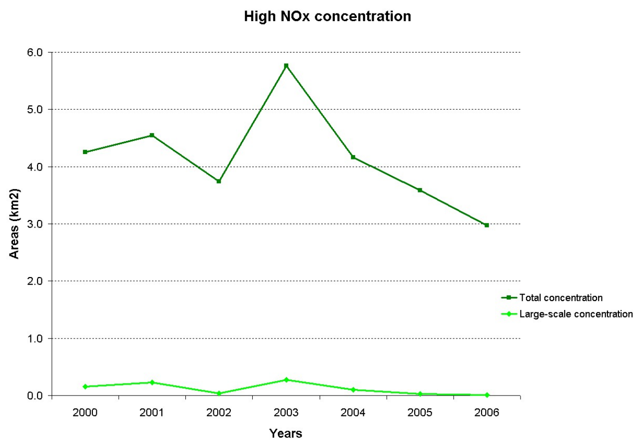

Fig. 9. Graph of areas of high concentration regions from 2000 to 2006.

annual mean NO_2 concentration along highways. This allows to extract polluted highways from the road network as vague lines. Length of polluted highways may be an indicator for traffic contribution to air pollution. *AvLength* metric on vague lines can measure the length of pollution along roads.

VI. Discussion and Conclusions

The metrics proposed in this paper are based on the concept of body mass. We consider the membership function of a vague object to be a density function. We propose a length measure, *AvLength*, for a vague line that is calculated as the mass of the vague line. This measure produces an average of lengths of all α-cuts of the vague line. *AvArea* for a vague region is an area measure, calculated again as the mass of the vague region. It produces an average of areas of all α-cuts of the vague region. *Centroid* is defined for a vague object of any type, and it is always a vague point. The coordinates of the vague point are the centre of mass of the object; its membership value is an average value of all memberships of the object. Area measure *AvArea* that we propose coincides with the area proposed by Rosenfeld [13].

A *Vagueness* measure is defined for a vague object of any type. This metric takes into account membership values and their distribution. It is maximal when the 1-cut of a vague object is of lower dimension that the object itself, or its is empty. *Vagueness* of a crisp object is equal to 0. Rough set theory offers a metric for roughness of a set. Translated to vague objects, the roughness measure of an object μ would be the ratio between the size of the 1-cut of the object and the size of its support set: $\mathsf{Mass}(\mu_1)/\mathsf{Mass}(supp(\mu))$. This metric is coarser than the *Vagueness* measure we propose.

Acknowledgements

The large-scale NO_2 contribution data were made available by MNP. Thanks to Guus Velders and Jan van de Kassteele their help in understanding the data and the application.

References

[1] F. Ayres Jr. *Theory and Problems of Differential and Integral Calculus*. Schaum Publishing Co., New York, 2nd edition, 1964.

[2] H. Bandemer and S. Gottwald. *Fuzzy Sets, Fuzzy Logic, Fuzzy Methods: with Applications*. John Wiley & Sons, 1995.

[3] I. Bloch. On fuzzy distances and their use in image processing under imprecision. *Pattern Recognition*, 32:1873–1895, 1999.

[4] A. Bogomolny. On the perimeter and area of fuzzy sets. *Fuzzy Sets and Systems*, 23:257–269, 1987.

[5] B. B. Chaudhuri and A. Rosenfeld. On a metric distance between fuzzy sets. *Pattern Recognition Letters*, 17:1157–1160, 1996.

[6] B. B. Chaudhuri and A. Rosenfeld. A modified Hausdorff distance between fuzzy sets. *Information Sciences*, 118:159–171, 1999.

[7] A. Dilo. *Representation of and reasoning with vagueness in spatial information: A system for handling vague objects*. PhD thesis, Wageningen University, Enschede, The Netherlands, June 2006. ITC Dissertation No. 135.

[8] A. Dilo, P. Bos, P. Kraipeerapun, and R. A. de By. *Flexible Databases supporting Imprecision and Uncertainty*, volume 203 of *Studies in Fuzziness and Soft Computing*, chapter Storage and manipulation of vague spatial objects using existing GIS functionality. Springer, 2006.

[9] A. Dilo, R. A. de By, and A. Stein. A system of types and operators for handling vague spatial objects. *International Journal of Geographical Information Science*, 21(4):397–426, 2007.

[10] D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York, 1980. pp. 38–40.

[11] C. H. Edwards Jr. and D. E. Penney. *Calculus and Analytic Geometry*. Prentice Hall, Inc., 1982.

[12] P. Fisher. Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, 113(1):7–18, July 2000.

[13] A. Rosenfeld. The diameter of a fuzzy set. *Fuzzy Sets and Systems*, 13:241–246, 1984.

[14] A. Rosenfeld. Distance between fuzzy sets. *Pattern Recognition Letters*, 3:229–233, 1985.

[15] A. Rosenfeld and S. Haber. The perimeter of a fuzzy set. *Pattern Recognition*, 18(2):125–130, 1985.

[16] M. Schneider. Metric Operations on Fuzzy Spatial Objects in Databases. In *Proceedings of 8th ACM Symposium on Geographic Information Systems (ACM GIS)*, pages 21–26, 2000.

[17] P. van Breugel, A. Baum, L. Calovi, M. Hackman, C. de Gier, M. Juneholm, R. Klaeboe, E. Pucher, A. Kampfer, and J. Vinot. Examples of air quality measures within europe, national measures of the international cedr air quality group. Technical Report DWW-2005-043, Rijkswaterstaat/DWW/IM, Delft; The Netherlands, 2005.

[18] J. van de Kassteele. *Statistical air quality mapping*. PhD thesis, Wageningen University, April 2006.

[19] G. Velders, J. Aben, J. Beck, W. Blom, A. Hoen, B. Jimmink, J. Matthijsen, J. de Ruiter, W. Smeets, K. van Velze, H. Visser, W. de Vries, and K. Wieringa. Concentratiekaarten voor grootschalige luchtverontreiniging in nederland. Technical report, Netherlands Environmental Assessment Agency (MNP), 2006.