Geographical Information Management and Applications

Thesis report

The Semantics of Extensive Quantities in Geographical Information

Author:

Eric J. Top, MSc.

E.J.Top@uu.nl

Supervisor:

Dr. Simon Scheider

S.Scheider@uu.nl

Responsible professor:

Prof. dr. Stan Geertman

S.C.M.Geertman@uu.nl

Universiteit Utrecht

TU Delft

University of Twente.

Wageningen UR

The Semantics of Extensive Quantities in Geographical Information

A Master's thesis by:
Eric J. Top, MSc.
https://www.uu.nl/medewerkers/EJTop
e.j.top@uu.nl

Under supervision of:
Dr. Simon Scheider
https://www.uu.nl/medewerkers/SScheider
s.scheider@uu.nl

Abstract

The next generation of Geographical Information Systems (GIS) is anticipated to automate some of the reasoning required for spatial analysis. An important step in the development of such systems is a better understanding and modeling of the decision process about which arithmetic operations can be meaningfully applied to geographic quantities. The concept of extensivity plays an important role in determining precisely what amounts are, when they can be aggregated by summing them, and when this is not possible. However, currently, multiple contrasting definitions of extensivity exist, and none of these suffice for handling the different practical cases occurring in geographical information. As a result, GI-analysts predominantly rely on intuition and ad hoc reasoning to determine whether two quantities are additive. In this paper, we present a novel approach to formalizing the concept of extensivity. In our algebraic definition, we do away with some of the constraints that limit the use of older approaches. By treating extensivity as a relation between quantities, our definition offers the flexibility to relate a quantity to many domains of interest. We show how this new notion of extensivity can be used to classify the kinds of amounts in various examples of geographical information.

Keywords and phrases Extensivity, quantities, Geocomputation, Semantic labeling of geodata

Preface

This work is a thesis submitted in fulfillment of the requirements of the MSc. programme Geographical Information Management & Applications (GIMA). This thesis is an extended and adapted version of a paper that has been submitted to GIScience 2021. Section 4 is new in this work. This work has been primarily produced by E.J. Top, but contains work from others. For the sake of transparency, an author contributions statement has been added after the conclusion.

1 Introduction

An important distinction in geographical analysis is between quantities that can be summed and quantities that should be averaged during aggregation. These are known as respectively extensive and intensive quantities. Human analysts can intuitively tell how a quantity should be processed when two regions are merged. Two temperature values of two spatial regions, for example, should obviously not be summed, but instead averaged when the regions are aggregated. In geographical information systems (GIS), however, the values may be represented by the same concrete data types, and thus cannot be systematically distinguished. Current GIS lack a method for automating aggregations because we lack a theory of extensivity that can tell us under which circumstances we can sum up quantities.

One way to capture extensivity is in terms of a relation between different measurements [29]. This notion of extensivity entails that quantities can be aggregated if they share underlying domains of measurement by which they are controlled and measured. Controlling quantities need to be separated from each other, and both controlling and controlled quantities need to be additive in a certain sense. For example, the population of Europe can be aggregated with the population of Africa, as both populations are part of the world population, but not with the summed GDP of Africa, which does not share the same measurement domain. Here, the spatial administrative units are controlling and population counts are measured. However, Europe's population can also not be aggregated with the population of Utrecht, because even though they share the same measurement domains, Utrecht is already a part of Europe. While such observations may seem intuitive, the sciences still lack a formalization of these kinds of considerations. Most existing definitions that reached prominence are either too restrictive or too vague, leaving room for inadequate interpretation in the context of GIS (See section 2.1).

A concise yet flexible definition of extensivity would enable geodata models to distinguish whether spatial arithmetic is applicable or not. Scheider and Huisjes [29] show the merit of spatial extensivity in this respect and manage to automatically distinguish extensive from intensive quantities with high accuracy. However, their approach is not formalized and does not account for quantities that are additive in domains other than space, such as e.g. time. If we recognize there are multiple dimensions of extensivity, new ways to *categorize* quantities emerge. A water flow accumulation is extensive in space and time, the cost of a stay at a hotel is extensive in time and some monetary currency, and the cost of rental cars is extensive in time (i.e. the duration of renting), space (i.e. the amount of kilometers driven) and the amount of cars (i.e. renting two cars is more expensive than one). Extensivity offers a new semantic dimension by which data can be discovered and processed. A definition of extensivity would also contribute to a data-driven science [15, 12] by allowing to infer which arithmetic operations can be applied to available quantities.

In this paper, we formalize extensivity in a way which is general enough to be applied to the many cases relevant in GIS, and which is distinctive enough to decide on the applicability of aggregation (sum) operations. Based on the formalism, we develop a categorization of quantity measures in geographical information relative to various domains of control. Our contributions are threefold and provide answers to the following questions:

- In what form can extensivity be unambiguously formalized in terms of aggregations of quantities?
- How can extensive quantities be measured across various domains of time, space and content that are relevant for aggregation in GIS?
- How can aggregations in geographical information be systematically categorized according to different kinds of extensive quantities?

The rest of the paper is organized as follows. First, we review what is known about extensivity, quantities and measurement. Second, we reinterpret Sinton's [31] three roles of measurement (i.e. measure, control, constant) and show in an informal manner how they can be used to specify extensivity relations. Third, we present a formalized algebraic theory of extensivity as a relation between a measure and one or more controls. Fourth, we propose twelve categories of measurement of extensive quantities in the context of geographical information. Finally, we shortly discuss the implications of our findings and conclude by answering the posed research questions.

2 Extensivity, quantities, and measurement

We start with reviewing existing literature on extensivity and quantities and scrutinize the underlying approaches for our purpose. Furthermore, we critically examine Sinton's notion of controlled measurement.

2.1 Extensivity and intensivity

The concept of extensivity originates from the fields of Physics and Chemistry where it is used to describe the mathematical nature of properties. Its introduction can be accredited to Tolman [34], who envisions extensivity as a way to describe phenomena whose measures are *naturally additive*. Of all phenomena he identifies only five as extensive in this sense, namely length, time interval, mass, electric charge, and entropy. For a contemporary definition of extensivity, scholars often refer to the green book of the International Union of Pure and Applied Chemistry (IUPAC), which describes an extensive quantity as *"a quantity that is additive for independent, non-interacting subsystems"* [7]. In practice, there seems to be an informal consensus that only properties like volume or mass are considered extensive. Even within this consensus, disagreement exists about what physical properties extensivity depends on. A number of papers from Physics and Chemistry try to address the confusion surrounding the concept [27, 2, 23]. Mannaert [23] finds that the expressions 'extensive quantity' and 'extensive property' are used interchangeably — He favours the use of the term 'extensive quantity' — and that some use additivity to define extensivity (i.e. the sizes of two quantities can be added up during aggregation) while others use proportionality (i.e. a quantity inextricably changes with relation to changes of another quantity). Some scholars limit extensivity to a relation of properties with respect to mass, while others relate them to the amount of substance or volume [23]. This interpretation of extensivity with respect to some specific kind of physical presence deviates considerably from the original theory [34], which holds that properties may be extensive also with respect to time or entropy. Not only do scholars consider different properties as the source of extensivity, they also disagree on the mechanisms of extensivity itself.

The concept of an extensive quantity is opposed by that of an intensive quantity, which has been defined as *"a quantity that is independent of the extent of the system"* [7]. Tolman [34] argues that, except for his five fundamental quantities, all quantities are intensive, because they are in some way derived from the five fundamental quantities. A speed, for example, is found by dividing a length (i.e. the distance) by a time interval. Some scholars hold that not all quantities are either extensive or intensive. They argue that some quantities are expressed as conjugates [1] or composites, which have characteristics of both.

2.2 Quantities

Quantities are described as "*...that by which a thing is said to be large or small, or to have part outside of part, or to be divisible into parts*" [19]. Specifications of quantities are frequently present in spoken language [33]. For example, the sentence 'The flock of birds flew over the wide river' not only specifies two different entities (i.e. 'birds' and 'river'), but also details their quantities (i.e. 'flock' and 'wide') and their interrelation (i.e. 'over').

From a semantic viewpoint, quantities should be distinguished from numbers, which are mathematical objects for counting, and measurement units, which indicate the measurement system a quantity is measured in[1]. In measurement theory, it is common to subdivide measurement systems using measurement levels, which range from nominal through ordinal and interval to ratio [32]. Arguably, these levels encode increasing amounts of information of a quantity, respectively providing information about class membership, order, relative position and absolute effect of a quantity. Chrisman [4] proposes extending the levels with counts, degrees of class membership, cyclical ratio, derived ratio, and immutable absolute measures, like probability.

Quantities can be negative and might be linearly ordered or not. For example, walking backwards for twenty meters can be seen as a negative quantity of forward movement associated with the number -20 and the unit 'meters'. The term magnitude, also called impact or size, is used to measure a quantity on a linear scale. Scholars sometimes distinguish multitudes from magnitudes [21]. Shortly put, multitudes refer to collections of discrete entities (e.g. a collection of cars), while magnitudes capture linearly order-able phenomena (e.g. the length of a road). Plewe [25] vaguely refers to these phenomena as 'geographic masses'. Our approach (see below) can be used to make these notions precise.

Information about the extensivity of a quantity is closely related to its part-whole relations. Such relations are commonly considered *homeomerous* with respect to its parts, meaning that all parts are of the same kind of quantity as the whole [13]. For instance, sectioning a portion of water results in sub-portions of water. According to Guizzardi [14], homeomerous part-whole relations can be modelled as mereological sums (i.e. aggregations of the subquantities) or by means of containment (e.g. a bottle of water). He suggests to conceptually model quantities as maximally self-connected parts. This approach implies that parts of a quantity, also referred to as pieces [22], are only instantiated if there is a need. For example, a body of water may be subdivided into its parts to identify sweet water and salt water if necessary, but this is not required for capturing the water concept. Guizzardi's mereological approach also works for universal properties and classes. For example, a car is a member of the collection of all cars (i.e. the class of 'cars'), and the mass of said car is a part of the set of all mass in the universe (i.e. the 'mass' property).

Mereological relations are essential because they specify whether two quantities are distinct, whether and how much they overlap and whether one quantity value contains another. For example, a university may host multiple lectures at once, meaning they share the same quantity of time. Summing the total time of the lectures may indicate how long it would hypothetically take to attend all lectures (e.g. 400 hours), but this does not correspond to the extent of time that is actually occupied by these lectures (e.g. 3 hours). If two lectures with a duration of 2 hours each overlap for 1 hour, they together occupy 3 hours in time. Claramunt and Jiang [6] show that such relations are not limited to space or time, but also exist between conjunctions of both.

2.3 Measurement of quantities

Sinton [31] is well-known for his idea that the measurement of spatial information requires attribute information about the space, time and theme components of the recording. Sinton argues that during any measurement each one of these three components fills the role of the constant, the control or the measure:

[1] Different measurement systems or reference systems [5] can represent the same kind of quantity. For example, the meter scale and the feet scale both represent the same quantity of lengths.

- The **constant** component, also referred to as the support or the fixed component, does not change at any point in the measurement process.
- The **control** component is allowed to vary over its measurement scale at the observer's discretion.
- The **measure** component is observed and its variation with respect to the control is recorded.

Take the example of a precipitation measurement. Precipitation is commonly measured with a rain gauge. This rain gauge fixes the spatial extent of the precipitation measure, e.g., to 1 m^2. The amount of water falling into the rain gauge is then measured in kg or liters (mm) over a variable amount of time, e.g. an hour or a day [5]. With an established constant (i.e. space) and control (i.e. time), it is possible to measure the theme component, which in this case is rainfall in mm. Sinton's work contains two important messages: 1) a measurement of a phenomenon requires other variables to be controlled, and 2) geographic information contains a combination of spatial information, obtained through the measurement of locations and regions, temporal information, obtained through the measurement of the progress of time, and thematic information, obtained through measuring the properties of a substance or content.

Chrisman [5] argues that apart from space, time and theme, there is another kind of control, namely control by relationship. For example, the measurement of a flow of export products from one country to another first requires establishing a relation between the two countries. Although relevant, we leave the study of network-controlled extensive quantities to future work.

The roles of measure, control and constant are essential for our purpose, because they aptly capture how quantities can play different roles in defining extensivity. Shortly put, it can be said if wo measured quantities are extensive, they share the same kind of control(s). However, Sinton's definition requires some scrutiny before it can be applied to quantities. The roles are by no means fixed to the three components of GIS. Many measurements ignore one or more of the components. For example, when measuring the duration of a given lecture, there is no need to take the size of the lecture room or the didactic ability of the lecturer into account. In fact, only the time at which the lecture happens is required as a control for the duration. Thus, for this example, the time component is both measure and the control at once. Similar examples could be given for the space and theme components. Furthermore, this also shows that the components space, time and theme are too coarse to distinguish different types of quantities within the same component, and thus are too coarse for capturing extensivity. We therefore adopt two alterations of Sinton's idea. Firstly, we interpret Sinton's components as classes of quantities which might play a role or not in a given measurement. We thus allow for arbitrary combinations of quantities filling the roles in a single measurement. Secondly, we assume that quantities exert no influence on measurement (i.e. are kept constant) unless specified otherwise. This prevents the need for explicitly filling the constant role.

3 A formal theory of extensive quantities

In the following we introduce a formal theory of extensivity of quantities in First Order Logic (FOL). FOL is sufficient to reason about a single quantity domain. Strictly speaking, we go beyond FOL when talking about different domains in Def. 4, but this is not part of the current formalism. Free variables in propositions are implicitly all-quantified. Axioms have been checked for consistency and all theorems were automatically proven based on resolution using Prover9[2]. The script is available online[3].

3.1 Quantities, amounts and magnitudes

Intuitively, quantities can be added up to or removed from each other, resulting in a new quantity of which original quantities are parts. For example, a quantity of people can be added up to another quantity of people to form a total sum of people. Furthermore, each of these quantities can be counted and thus compared to each other on a common scale of measurement. In our theory, this means that quantities can be part of each other and can be *summed up* and *subtracted* from each other. Furthermore, quantities can be measured on a *linear scale*. In the following we formalize this intuition based on first distinguishing the notions of quantity, amount and magnitude.

3.1.1 Theory of Quantities

To capture intuition 1, we assume a *quantity domain* to be a domain of values together with algebraic operations that satisfy the following *partially ordered algebra*:

▶ **Axiom 1.** *Partial order*

$$\begin{aligned} & x \leq x && \textit{Reflexivity} \\ & (x \leq y \land x \geq y) \implies x = y && \textit{Antisymmetry} \\ & (x \leq y \land y \leq z) \implies x \leq z && \textit{Transitivity} \end{aligned}$$

▶ **Axiom 2.** *Sums and differences*

$$\begin{aligned} & x + y = y + x && \textit{Commutativity} \\ & (x + y) + z = x + (y + z) && \textit{Associativity} \\ & x + 0 = x && \textit{Identity} \\ & x + -x = 0 && \textit{Inverse} \\ & x \leq y \implies x + z \leq y + z && \textit{Translation invariance} \end{aligned}$$

This introduces operations for adding and subtracting values. Note the identity (empty) element 0 which can be summed up without changing anything and which results from subtractions. Translation invariance logically embeds the partial ordering into the sum operation, by saying that the ordering stays invariant when adding the same quantity on each side[4]. Based on this, we define the following functions and predicates:

▶ **Definition 1.** *strict order and overlap*

$$\begin{aligned} & x < y \iff (x \leq y \land \neg(y \leq x)) && \textit{Strict order} \\ & O(x, y) \iff \exists z (z \leq x \land z \leq y) && \textit{Overlap} \\ & y > x \iff x < y && \textit{Strictly greater than} \\ & y \geq x \iff x \leq y && \textit{Greater than or equal} \end{aligned}$$

---

[2] `https://www.cs.unm.edu/~mccune/prover9/`
[3] `http://geographicknowledge.de/vocab/quantity.txt`
[4] This axiom is also used for a partially ordered group: `https://en.wikipedia.org/wiki/Partially_ordered_group`

Strictly ordered quantities are unequal, and overlapping quantities have a common lower bound in this ordering. However, sometimes, we need to interpret the ordering in terms of parthood. In mereology [3], overlap is interpreted in the way that quantities share a common part (e.g. two regions share a common region), and the neutral (0) element is the empty part (e.g. the empty region). The operators can be interpreted both in an arithmetic as well as a mereological sense. For example, in terms of arithmetic, we can interpret temperature measurements $5°K + 3°K = 8°K$, area measurements $20m^2 - 8m^2 = 12m^2$, and time measurements $2min \leq 5min$. Furthermore, in terms of mereology, two spatial regions can form one bigger region, one interval of time can be deducted from another interval of time, and a collection of two cars may be part of a collection of five cars. Overlap can mean that, for instance, 25 to 29 minutes and 28 to 31 minutes overlap, because they share the range of 28 to 29 minutes.

Many plausible theorems about quantities can be proven in this theory. For example, the fact that when we add a quantity greater than zero to another, the result will always be strictly greater than the original quantity.

▶ **Theorem 2.**

$$x + y = z \wedge y > 0 \implies x < z \quad \textit{Strict order add}$$

3.1.2 Amounts as extensional mereological quantities

The similarity of the above mentioned axioms to mereology[5], e.g. to an algebra of spatial or temporal regions, was already mentioned above. However, this interpretation comes with extra assumptions: in spatial or temporal reference systems, e.g., there always exist regions that are autonomous in the sense that they are not a part of each other. Furthermore, when regions are not part of each other, they give rise to new kinds of (supplemented) region hierarchies that exist in parallel. Thus, these regions are not *linearly ordered* anymore[6]. We make this possible interpretation of a quantity more explicit by adding axioms of a *subtheory* that turn quantities into domains with an *extensional mereology* which are *not totally ordered*:

▶ **Axiom 3.**

$$\neg(y \leq x) \implies (\exists v(v \leq y \wedge \neg O(v, x))) \quad \textit{Strong supplementation}$$

$$\exists x, y(\neg(x \leq y \vee y \leq x)) \quad \textit{Non totality}$$

Strong supplementation turns the theory into an extensional mereology, which is a specialization of a partial order, where $\leq$ can be interpreted as $\subseteq$ and non-overlapping quantities exist if quantities are not part of one another. Non-totality makes sure there are at least two independent quantities that are not part of each other. In this theory, it can e.g. be proven that if you sum up two quantities greater than 0 where one is part of the other, this will always generate the greater one of the two as a result of the operation (reflexivity of sums), which is in apparent contrast to the number line[7]. Furthermore, it can be proven that non-zero quantities with the same proper parts are equal (extensionality):

---

[5] More precisely, the partial order forms a *ground mereology* without supplementation, and the sum axioms introduce a *closure principle* into the mereology, see [3].

[6] They form a *lattice* in the mathematical sense.

[7] This would mean e.g. that 4+4 = 4

▶ **Theorem 3.**

$$(0 \leq x \wedge x \leq y) \implies x + y = y \qquad \textit{Reflexivity of sums}$$
$$(0 < x \vee 0 < y) \implies (\forall z(z < x \iff z < y) \implies x = y) \qquad \textit{Extensionality}$$

The theory contains the most important elements for defining sets in terms of set intersection and union. Note, however, that set theory is, again, only a particular interpretation of this theory[8]. There are also other important interpretations, for example, in terms of parthood of physical things, portions of amounts of matter, or else in terms of intervals in time or space.

We call domains that satisfy the axioms of this subtheory *amounts*. The intuition is that amounts are quantities that can be added to different (independent) piles.

3.1.3 Magnitudes as totally ordered quantities

The other subtheory of quantities that is relevant for us simply assumes a total order. This makes the domain linearly ordered.

▶ **Axiom 4.**

$$x \leq y \vee y \leq x \qquad \textit{Totality}$$

We call such quantities *magnitudes*, and we use this theory to talk about quantities such as size, duration, or the count of a set, which are on a common linear measurement scale. In principle, totality would be compatible with strong supplementation alone, but it can be proven that this would shrink the domain to a single element. To be useful as a subtheory, we therefore also assume non-totality and thus both theories exclude each other.

3.2 Measure and control

Sinton's roles are important for us because they relate different quantities to each other in terms of measurements. In our theory, we assume that *the role of control is always played by amount quantities* whereas the *role of measure can be played any kind of quantity.* For example, a baker may want to measure how much flour he or she uses in a day in kg. Here, we measure a magnitude of flour controlled by an amount of time. Alternatively, a certain volume of soil contains a particular amount of sand. Note that in the latter case, amount means that there might be a different amount of sand with the same weight measured in kilogram.

Using our basic theory of quantities, we introduce a function which maps controlled amount quantities to measured quantities:

▶ **Definition 4.** *Measurement of quantities*
Let X be a domain of amounts, and Y be any domain of quantities. Let m be a function $X \to Y$. Then m is called a measurement function, and $\forall x \in X$ are called controls, and $\forall x$ $m(x) \in Y$ are called measures.

For example, on a particular day of baking we measure 25 kg of flour, or a certain volume of soil has a certain amount of sand.

A note on the formal properties of m: Intuitively, we would always expect that two different controls can exist that have the same measure. This means the measurement

[8] `https://en.wikipedia.org/wiki/Zermelo%E2%80%93Fraenkel_set_theory`

function is expected to be *non-injective* in the case of a magnitude measure. An example would be two different piles of sand with the same weight, or two lectures of the same duration. However, if the measure is also an amount, we would expect, in contrast, that the measurement function is *injective*, at least in many cases. So every quantity in the range of the function has only a single quantity in its domain. For example, two distinct amounts of sand are always contained in different regions of space, even if they have the same weight magnitude. However, we also have counter examples (measurements of amounts that are not injective), like the measurement of a projected area given some region in three dimensions. Therefore we leave injectivity out of the formalism for now.

3.3 Defining extensivity

In this section, we define extensivity as a property of a measurement function between two quantities. In our example of the baker, the daily recordings give the baker the ability to calculate the total use of flour during a week by adding up the daily recordings of used flour. This can only be done because *time intervals as well as amounts of flour both can be added up in a coordinated manner.*

We say that two quantities are additive with respect to a measurement function m, iff the following holds:

▶ **Definition 5.** *Additivity of measurememts*

$$\forall x, x' \in X. (\neg O(x, x') \implies m(x) + m(x') = m(x + x')) \iff \mathit{Additive}(x, x', m)$$

This definition says that the measurement of the mereological sum of two independent quantities (that do not overlap) in the same domain of control should be the same as the sum of their measures. For illustration, consider the weight of the contents of two buckets of ice. Piling up the contents of the two buckets results in a quantity of ice that has the same weight as the sum of the individual weights of each of the buckets of ice (minus the buckets themselves). Note that the two quantities of ice should not mereologically overlap. This avoids errors like the following: One could make two selections of ice from a big pile, each containing two-thirds of the total ice in the pile, and sum them up, which gives 1 2/3 of the total ice pile. However, that would not correspond to the actual amount of ice in the pile.

Whether a quantity (as a domain) is extensive depends on the additivity of its elements in the context of a measurement function:

▶ **Definition 6.** *Extensivity of a quantity domain Y being measured w.r.t. a control domain X*

$$\begin{aligned}&\mathit{ExtensiveWrt}(Y, X, m) \iff \\ &\forall x, x' \in X(\mathit{Additive}(x, x', m) \wedge (m(x) \in Y).\end{aligned}$$

Note that extensive measurements are homeomerous in the sense that every mereological part of a controlling amount can be measured within the same quantity domain Y. If a quantity domain Y is extensive with respect to an amount domain X under measurement m, many additional theorems can be proven. For example, if we remove an amount y from another one x of which it is a part, then the measure of the resulting amount will be the same as when subtracting the measure of y from the measure of x:

▶ **Theorem 7.**

$$y \leq x \implies m(x) + -m(y) = m(x + -y) \quad \mathit{Subtractivity}$$

This notion applies to many examples of quantities. For example, in the specified sense, an amount of sand is extensive with respect to a given volumetric space. In addition, a weight of sand (in kg) is extensive with respect to the corresponding amount of sand. Note that extensivity can also apply in the opposite direction: the volumetric space that sand occupies is extensive with respect to the amount of sand. And a volume of sand (in m^3) is extensive with respect to the corresponding volumetric space it occupies. While it happens to be that volumetric space and mass of sand are both extensive with respect to each other, it should be stressed that extensive relations are not necessarily bi-directional. This depends on whether m is bijective or not (and thus whether there exists an inverse function). In our theory, it can e.g. be proven that m needs to be non-injective in case m maps into a magnitude, under the additional assumption of domain closure (such that there always exist amounts with equal magnitude).

There is also the possibility that a single measure is extensive with respect to multiple controls. For example, a measure of total precipitation is controlled by space (e.g. m^2) and time (e.g. days). At this point only a theory of relations between a measure and a single control has been established. However, the definition can be easily adapted. In the case of multiple controls of a measure, let m be a function $X, A, B, \ldots \to Y$, where $X, A, B, \ldots$ are all domains of controls. We define additivity with respect to one of these controls keeping the others fixed:

▶ **Definition 8.** *Partial additivity*

$$\forall a \in A, \forall b \in B, \ldots, \forall x, x' \in X.\ pAdditive(x, x', m) \iff$$
$$(\neg O(x, x') \implies m(x, a, b, \ldots) + m(x', a, b, \ldots) = m(x + x', a, b, \ldots))$$

For example, precipitation can be considered extensive with respect to its spatial control when its temporal control is fixed. If a measure is partially additive with respect to a single control, we can also say the measure is partially extensive. For example, the measure of total precipitation is partially extensive with respect to its spatial control. If and only if the definition holds for all inputs we can speak of a fully extensive measure. For example, total precipitation is partially extensive with respect to all spatial and temporal controls, thus is fully extensive.

4 Amounts and magnitudes in relation to GIS core concepts

In this section the definition of amounts and magnitudes from the last section are related to prevalent concepts in GI-science. Concretely, we consider the relations between amounts and magnitudes on the one hand and objects, fields, events, and networks on the other. For the latter four concepts, we follow the specifications proposed by Kuhn [20], which are formally defined by Scheider et al. [30]. We argue for a modelling approach where amounts are introduced as an intermediary step for quantifications of properties. Finally, we introduce four types of amounts, which we schematized in context of the core concepts.

4.1 Core concepts

Kuhn [20] proposes ten core concepts of the semantic characteristics of spatial information. Six of these concepts allow spatial reasoning while four allow reasoning with the information represented by the spatial concepts. Unlike the other concepts, four of the spatial concepts identify the existence of some thing or stuff at some location or locations, to which spatial, temporal and thematic information can be attributed. For example, the coordinates 48° 51’

29.1348” N and 2° 17’ 40.8984” E and a weight value of 10,100 tons can be attributed to an object, specifically the Eiffel tower. In general, we refer to the concepts that identify the existence of things or stuff as *entity*[9]. Four of Kuhn’s spatial concepts, namely *object*, *field*, *event* and *network* meet this criterion. We limit our focus to these four concepts.

In spatial information, an **object** is a representation of a spatially discrete entity. Because of their discrete nature, objects can be counted. For example, it is possible to count five cars in a parking lot. Within the context of GIS, a common example of an object is a statistical region, like Europe, the Netherlands, and Utrecht. Objects are characterized by having identity (i.e. each object is unique). To illustrate, two cars of the exact same type and production year may be distinguished by their two different license plates.

An **event** is a representation of a temporally discrete entity. For example, an earthquake does not persist through time; there is an interval of time in which an earthquake took place. Events are similar to objects in the sense that they can be counted. The difference between the concepts is that events are endurant, meaning they have limited existence, while objects are perdurant, meaning their existence is not essentially bounded by time. For example, a car could persist indefinitely if natural decay is disregarded. However, a car crash implies there is a moment in time in which the crash took place. Although events often exist in the spatial domain, this is not necessary. For example, a birthday occupies an interval of time, but not a region of space.

A **network** is a relation between a pair of entities (e.g. two objects) and a qualifier. This could for example be the shortest route (qualifier) between a worker’s home (object 1) and their office (object 2). Kuhn [20] distinguishes *path* networks (e.g. a route from home to work) and *link* networks (e.g. a trade relation). Networks can be counted. It is for example possible to count the number of possible routes between two points or how many segments a route has (The latter could for this purpose be interpreted as a concatenation of networks with the vertices as begin and end points). Although networks derive their meaning from being a relation between entities, their existence is not dependent on it. For example, the original trading posts along the silk routes have all disappeared, but the network of silk routes remains as a historical notion.

A **field** is a representation of a spatially continuous property or entity. In this case, continuous does not mean that the field has a limitless extent per se, but rather that its semantic meaning does not depend on its extent. For example, a temperature field can be split into two. These two new fields are still temperature fields. In contrast, splitting an object in a similar fashion may destroy the integrity of the data representation (e.g. half of a car is not a car). Numeric values of fields are intensive, meaning they should be averaged during aggregation. For example, a distance field contains varying values representing the distance, e.g. to a specified point. Summing these distance values would make little sense for all but very specific purposes and would always result in an abstract sum. There is still a lack of consensus about whether a field just represents continuous properties or also continuous entities [24, 25]. While it seems generally accepted that for example temperatures should be considered as fields, it is not clear whether they should also be used to denote continuous entities like water or sand. A land use dataset may be considered as a field containing different land uses as its values, which shows what kind of land use covers a certain area.

[9] Note that beyond the scope of this paper the term entity is ambiguous (see e.g. [28]).

4.2 Entity, amount and magnitude

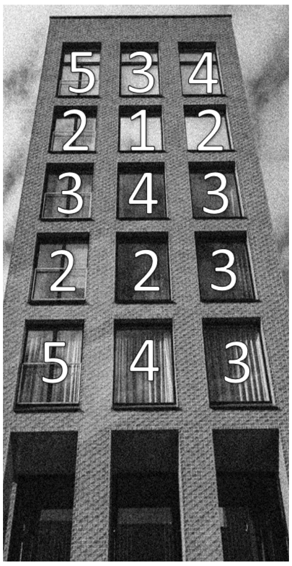
In section **??**, the relations between amounts and magnitudes are defined using a measurement function which assumes a mapping from a control to a measure. A similar mapping function can be defined for relations between entities and amounts or even entities and magnitudes. This function would map from e.g. an object to an amount, which in turn could serve as the control for a magnitude measure in the measurement function. To illustrate, a car may relate to an amount of four wheels, which gives a magnitude of four.

It may be tempting to directly relate an entity like an object or field to a magnitude for the sake of modelling simplicity (e.g. a car is related directly to a magnitude of four wheels without an intermediary amount). This direct mapping is the dominant approach in data modelling, both within GIS and beyond (e.g. [17, 18]). Although not necessarily erroneous, this approach complicates the maintenance of the mereological integrity of entities in data models. Assume for example that there exists an apartment complex with 46 residents and an individual apartment with 5 residents (Figure 1a). A mapping from entity (in this case an object) to magnitude would not capture information about whether the aggregation of the apartment and the complex counts 46 or 51 residents (i.e. are the apartment residents also apartment complex residents?). Adding additional information is thus necessary. A common approach to do so is to add a relation specifying whether the apartment is part of the complex (Figure 1b). However, this would not relate the counts of 46 and 5 residents themselves, but rather the apartment and complex that carry the resident counts as an attribute. Although resident count has the same meaning in relation to each of the two entities, this is not necessarily the case. For example, the number of direct neighbours (residents in rook contiguous apartments of a given apartment) is a count of neighbours for individual apartments, but not for the apartment complex. For the complex, it is a count of the number of neighbour relations between residents of the apartments within the complex. As can be seen in Figure 1b, this subtle information is easily lost during aggregation. Of course, the attribute could be relabeled accordingly, but this would have the adverse effect that the parts do not carry the same label as the whole.

Explicitly identifying the resident counts as amounts circumvents the issues encountered during the aggregation of entities and allows the preservation of homeomerosity (Figure 1c). Using the approach suggested by Guizzardi [14], the parthood of the 5 residents to the 46 residents can be explicated, regardless of which entities relate to these two amounts. Instead of saying that the apartment is part of the complex, it is said that the 5 residents in the apartment are part of the 46 residents in the complex. In addition, the dimensions of parthood become explicit. The amount of residents relates to both the apartment and the apartment complex. However, it is not necessary to use the apartment complex as a carrier for the aggregated amount of neighbours, because this amount can exist independently. In summary, amounts can serve as a focal point for relating quantitative attributes with each other, relations of which the extensivity relation is an example. For instance, a count of residents is mutually extensive with a count of neighbours if the same people are being counted.

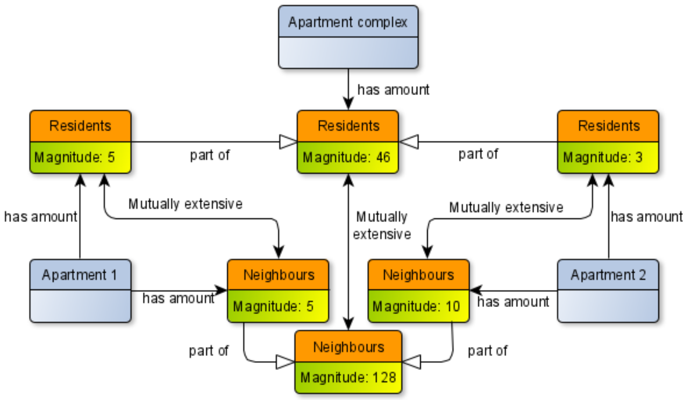
4.3 types of amounts and magnitudes

We find four different types of amounts and magnitudes, based on two distinctions. These distinctions are (1) whether the measured entity is discrete or continuous and (2) whether the value of the measure is nominal or ratio-scaled. The typification bears a resemblance to Talmy's notion of quantity disposition [33], which distinguishes between discrete and



(a) Apartment complex with numbers of residents

(b) Entity-magnitude modelling



(c) Entity-amount-magnitude modelling

Figure 1 Entity-magnitude and entity-amount-magnitude approaches for the apartment example. Apartment 1 is the top-left apartment. Apartment 2 is the top-middle apartment.

continuous and between unbounded and bounded phenomena. An important purpose for these types is to distinguish different methods of aggregation and how these relate to the core concepts. However, we limit our focus to extensive aggregations, which means the following types do not capture statistical aggregates like average, standard deviation, minimum and maximum. The four type pairs are schematized in Figure 2 and can be understood as follows:

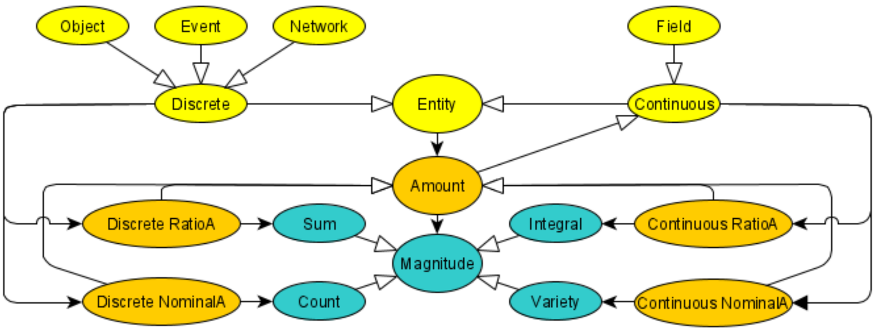


Figure 2 Amounts and magnitudes in relation to the core concepts

Yellow ellipses denote entities; Orange ellipses denote amounts; Blue ellipses denote magnitudes; Black arrows denote measurements; White arrows denote generalization

- A **Discrete RatioA** amount is an amount of a discrete entity (i.e. object, event or network) which represents an attribute with a ratio-scaled measurement scale. For example, the weight of a car is ratio-scaled and attributed to a discrete entity. Although a ratio scale implies some sense of unboundedness of the attribute, the attribute is essentially bounded by the entity by which it is carried. As a result, these amounts can be summed as discrete chunks of continuous stuff. For example, the weights of a 20kg brick and a 50kg metal pipe can be summed, which means we can infer that the combination weighs 70kg. We thus refer to the magnitude of this kind of amount as a **sum**.

- A **Discrete NominalA** amount is a count amount of a collection of homeomerous discrete entities. For example, a collection of five cars can be counted because five discrete entities carry the same nominal value (e.g. 'car') from which they get their homeomerosity. Different selections may lead to different amounts. For example, selecting all entities that are 'car' and 'blue' may give a count of three blue cars. It is also possible to aggregate different counts. For example, a count of five cars can be aggregated with a count of three bicycles, giving a count of 8 vehicles. We refer to the related magnitudes as **counts**.
- A **Continuous RatioA** amount is an amount of a continuous entity representing a ratio-scaled attribute, such as the weight of sand or the size of forested area. These amounts are not an intrinsic part of the entity in the sense that the amounts can increase or decrease in size without changing the essence of the entity. Regardless of whether the amount of water can fill a cup, a swimming pool or an ocean, it is still recognized as water. Because of this characteristic, they cannot be summed like Discrete RatioA amounts. Rather, they need to be measured using an integral function. For example, the distance between a moving point and a static point increases gradually while the first point moves away. Depending on the level of accuracy, the distance measure may yield 1 meter, 1.2 meter, 1.23 meter, 1.232 meter, et cetera. Probst [26] considers the possibility of relating multiple measurements to the same phenomenon. Our entity-amount relation enables such an approach because multiple amounts can represent measurements of the same entity. Simultaneously with the demarcation of the amount, the value of the magnitude is found with the integral function. For this reason, we refer to this magnitude as the **integral**.
- A **Continuous NominalA** amount contains the variety of nominal values within a continuous entity. For example, water can be considered as just water, which gives a variety of one. However, one could distinguish sweet water and salt water, which would give a variety of two. An amount itself can be considered as a continuous entity. In this case, the variety would always take a value of one, because amounts are homeomerous. Continuous NominalA amounts can be considered as the set of nominal values. As a result, there should be no duplication during aggregation. For example, the set of red and blue on the one hand and the set of yellow and blue on the other hand constitute the set of red, blue, and yellow, not the set of red, blue, blue, and yellow. The variety of this kind of amount is thus equal to the cardinality of the measured entity's set of nominal values within the attribute of interest. The magnitude of the kind of amounts discussed here can therefore be typified as **variety**.

Note that in Figure 2 amount is a subclass of entity. This relation entails that amounts both denote continuous entities, which are substances like water and sand, as well as properties of those continuous entities, such as weight and distance. A distinction between amounts as entities and amounts as attributes may be desirable for two reasons. Firstly, an amount 'liter of water' currently has the same class as an amount 'water'. This is despite the former having a unit of measurement and the latter not having one. Secondly, amounts may have nominal values attributed to them. While this could be useful (e.g. it could be desirable to attribute the color 'blue' to water), it does contradict the homeomerosity of the water amount (e.g. 'blue' water is not the same as 'red' water). Note that attributing nominal values to an amount is only sensible if the amount is considered as an entity. As an attribute, the usefulness of an amount is that it quantifies a property of another entity. For example, the question "How much water is in this cup?" inquires about (the magnitude of) the amount of water in e.g. milliliters, not the color of the water in the cup. Entity amounts and attribute amounts could thus be considered separate (sub)concepts, because they may have different identity criteria (See [11]). Entity amounts are identified by means of a nominal value denoting an unbounded substance. This allows for example the distinction between 'sand' and 'water' in common speech. Attribute amounts, on the other hand, are identified by a unit of measurement and some quantifier, like a number. For example, it can be said that a cup contains 20 milliliters of 'stuff'. It is not necessary to know what this stuff is in order to know how much stuff there is. We leave a scrutinous formalization of entity amounts and attribute amounts and their corresponding identity criteria for future work.

5 Classification of geographic amounts

Different classes of extensive geographical measures can be found based on distinguishing *categories of the quantities* being measured. Using the (super-) categories time, space and content, and distinguishing magnitudes and amounts for each, a total of nine classes can be found, where each measurement class is represented as an arrow between domain categories in Fig. 3. Three kinds of measurements map from amounts to magnitudes within the category time, space or content, while six map between amounts of different categories. In the following, we discuss each of the nine relation classes. This is done with the use of nine examples, which are assembled in Figure 4.

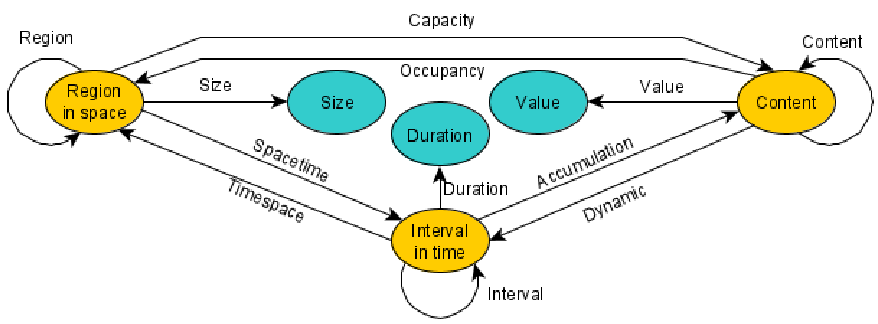
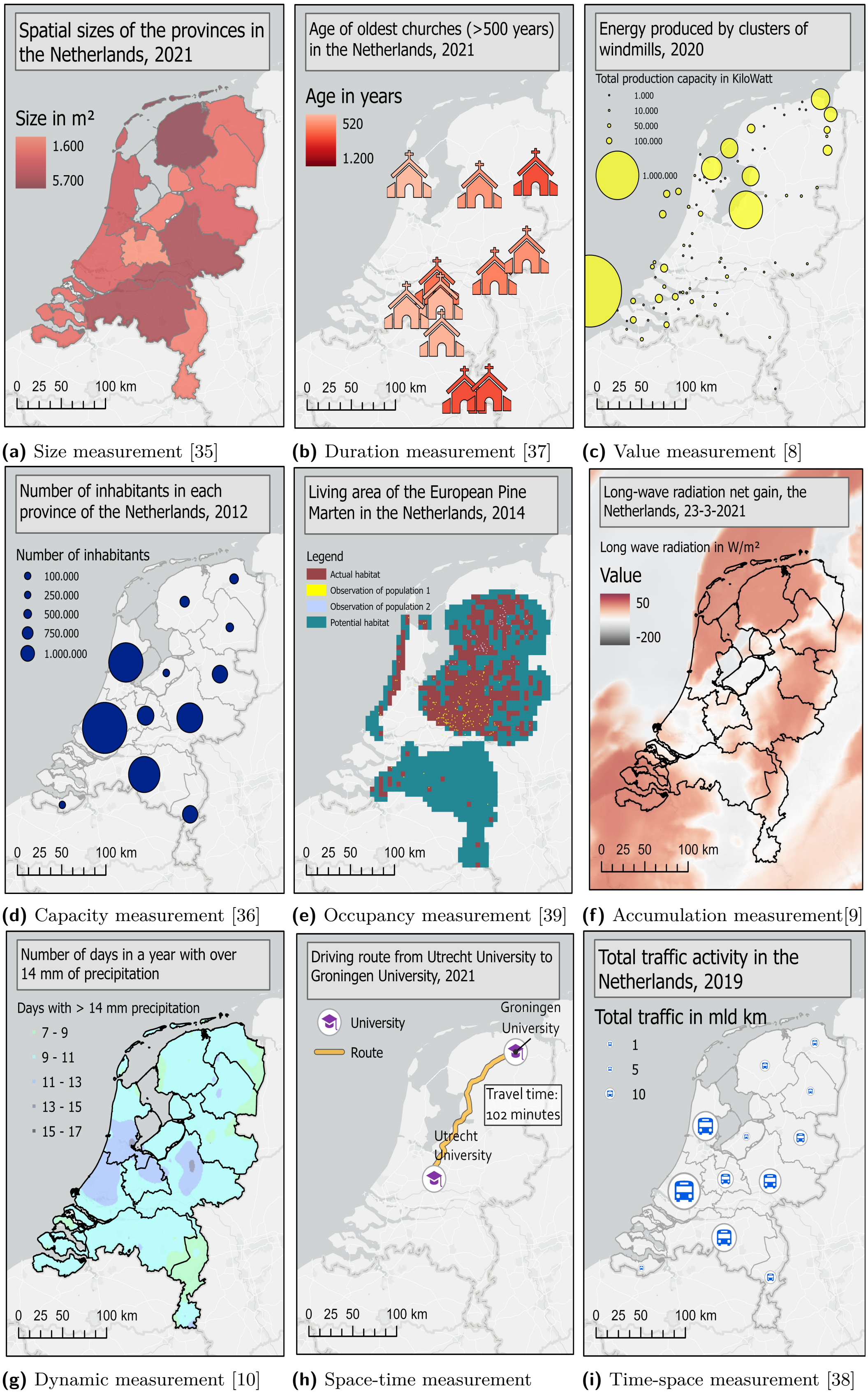


Figure 3 Extensivity triangle, showing possibilities of extensive measurements between three categories of quantities. Blue ellipses denote magnitudes, orange ones denote amounts.

5.1 Magnitude-of-amount measurements

A magnitude-of-amount measurement (measuring a magnitude of an amount) retrieves a magnitude which corresponds to (a selection of) some amount. We distinguish three of these within GIS, namely *size measures* which are spatial magnitudes, *duration measures* which are temporal magnitudes, and *value measures* which are other kinds of magnitudes, such as a count of objects, a monetary value, or a weight measure.

Size measures derive magnitudes from regions of space. Figure 4a provides an example of a size measure. The map depicts the spatial sizes of the provinces of the Netherlands. Clearly the regions of the provinces do not overlap and are partially ordered. They form a spatial lattice with an extensive mereology (regions can be part of one another). The



(a) Size measurement [35] **(b)** Duration measurement [37] **(c)** Value measurement [8]

(d) Capacity measurement [36] **(e)** Occupancy measurement [39] **(f)** Accumulation measurement[9]

(g) Dynamic measurement [10] **(h)** Space-time measurement **(i)** Time-space measurement [38]

Figure 4 Examples of measurement classes amounts are related to their size magnitudes, which in turn are totally ordered. According to our definition of additivity, the sizes of the regions can be directly summed to infer the sizes of mergers, because regions are spatially independent.

Duration measures retrieve magnitudes from intervals in time. Figure 4b shows the age of churches in the Netherlands that exist for at least 500 years. In this example, the intervals of existence of each church overlap for at least the last 500 years, meaning for some time the churches exist at the same time. Just like with sizes, durations of existence can be compared and be added up to derive the duration of existence of all churches. However, when summing up, overlaps need to be taken into account.

Value measures derive magnitudes from amounts of content, which are basically all amounts that are not in space or time. In Figure 4c, each bubble represents a magnitude of energy of an amount of wind turbines. Note that each bubble may contain multiple wind turbines which are implicit here. Another possible value measure would be the number of wind turbines in each cluster.

5.2 Amount-of-amount measurements

An amount-of-amount measurement measures an amount by using another amount as a control. For example, a population can be measured by controlling space and counting the individuals within this space. Also, the space they occupy can be found by measuring the spatial extents of the individuals. Note that the former and latter spatial amounts are opposed to each other[10]. We distinguish nine amount-of-amount measurements. Six of these are mappings between different amount categories, while three of these, namely region, interval, and substance, are mappings between different amounts within the same categories (e.g. from hours to minutes).

We identify six classes of measurements, namely *capacity, occupancy, accumulation, dynamic, space-time, and time-space*, where an amount is extensive with respect to an amount of a different category.

A capacity measurement maps from a spatial amount to a content amount. Figure 4d shows the population amounts of each province (e.g. the ’population of Utrecht’) which has a certain magnitude (e.g. 500,000). The population amounts themselves are measured with the regions as controls. For example, the population of Utrecht is measured with the region of the province of Utrecht as control. An occupancy measurement is the inverse in the sense that it maps from a content amount to a spatial amount. Figure 4e shows the living areas of European Pine Martens in the Netherlands, which is the result of such an occupancy measurement.

An accumulation measurement maps from a temporal amount to a content amount. Such measurements produce accumulations of content within a time interval. Figure 4f shows the net gain of long-wave radiation over one day. For each point in the Netherlands, a magnitude is given of the net radiation gain or loss accumulation over a day. These magnitudes are understood as mappings from radiation content, which is controlled by some time interval. The inverse of the accumulation measurement is the dynamic measurement, which map from content amounts to temporal amounts. The example in Figure 4g shows the amounts of days per region that have exceeded a threshold of >14 mm precipitation in a year.

A space-time measurement maps, as the name suggests, from a spatial amount to some temporal amount. Figure 4h shows the route from Utrecht University to Groningen University,

[10] They correspond to the opposing arrows ”capacity” and ”occupancy” in Fig. 3.

along with an indication of how long traveling this route takes by car. Note that this indication is not just a duration magnitude, but also implies a finite interval in time in which someone actually traveled. A longer path implies a larger time interval. This notion of time is key to Hägerstrand's Time Geography, which tells us that space accessibility is limited by temporal constraints [16]. A time-space measurement maps from a temporal amount to a spatial amount. Figure 4i shows the magnitudes of the amounts of traffic activity in 2019 in traveled kilometers. Note that these amounts are extensive with respect to time intervals, so they can be summed up with the amounts of traffic activity in 2018 to result in the amount for two years.

It is also possible to measure amounts using amounts of the same category as a control. For example, the churches in Figure 4b are selected from a bigger set of churches based on whether they have existed over 500 years. Only the resulting sub-selection is shown in the map with corresponding duration magnitudes. It is also possible to measure with semantically different controls within the same category. For example, an amount of pets can be measured with an amount of households as control. In turn, the amount of pets can be a control for measuring the amount of mice caught by each pet. We refer to all these options as *region measurements*, *interval measurements* and *content measurements* for respectively spatial, temporal and content amounts.

6 Discussion and conclusion

To better understand and to automate decisions on the applicability of arithmetic operations to spatial information, we have suggested that the concept of extensivity should be regarded as a formal property of a measurement function between different kinds of quantities. We have proposed an algebraic formalization of the underlying notions quantity, amount, magnitude, and additivity of a measurement function, and have proven theorems that correspond with our intuition about these concepts. While extensivity is currently primarily used to describe the behavior of physical properties, like mass and volume, our model can be used to generalize the applicability of this concept across various domains of measurement and different cases of information aggregation relevant for GIS. Our definition of extensivity not only lifts the restrictions of a fixed range of properties that can be considered extensive, but also the reliance on system theory. Subsystems are replaced with a simpler notion of quantities with an extensional mereology, similar to [14]. Furthermore, following earlier work [29], we reused Sinton's [31] notion of measure and control to define extensivity with respect to various control domains within the categories time, space and content. This gives rise to an extensivity triangle as a more versatile and succinct model of extensivity (or lack thereof) that is directly applicable to various forms of geographic information. We have tested this assumption by applying the model to a range of map examples, which allowed us to systematically categorize measurements relevant for GIS into twelve classes of extensivity that can be distinguished in principle.

The study is limited in that it ignores intensive, conjugate and composite (derived) quantities. These quantities are often, if not always, the results of multiplicative and divisive operations and they show a distinctly different behavior from extensive quantities, which warrants further research. Regarding our categorization of classes of extensivity, the content class is still a coarse container for many different kinds of geographic amounts. It could thus be further differentiated according to the core concept model [20]. Furthermore, to make the model practically usable, the dependency of various arithmetic operations (like weighted average and sum) on the form of extensive and intensive quantities needs to be investigated [29]. Another area of research is to investigate the role of amounts in natural language processing, such as geo-analytical question answering [40]. Future research should focus on developing conceptual modelling practices involving extensivity relations, on testing our notion of extensivity on empirical data involving analyst behavior, and on intensive and alternative types of quantities.

Author contributions statement

E.J. Top and S. Scheider conceived the original ideas, with the original notion of amounts stemming from S. Scheider and the idea of using extensivity for their definition coming from E.J.Top. E.J. Top wrote the manuscript, at times in collaboration with S. Scheider, and S. Scheider made alterations, especially in sections 2 and 3. The ideas and manuscript of section 4 are exclusively conceived by E.J.Top. The formal logic behind the text of section 3 comes from the hand of S. Scheider with help of E.J.Top. All illustrations come from the hand of E.J.Top. H. Xu, E. Nyamsuren and N. Steenbergen contributed in critical discussions which were essential in the development of the paper.

References

1 Robert A Alberty. Legendre transforms in chemical thermodynamics. *The Journal of Chemical Thermodynamics*, 29(5):501–516, 1997.
2 Sebastian G Canagaratna. Intensive and extensive: underused concepts. *Journal of chemical education*, 69(12):957, 1992.
3 Roberto Casati, Achille C Varzi, et al. *Parts and places: The structures of spatial representation*. Mit Press, 1999.
4 Nicholas R Chrisman. Rethinking levels of measurement for cartography. *Cartography and Geographic Information Systems*, 25(4):231–242, 1998.
5 Nicholas R Chrisman. *Exploring geographic information systems*. Wiley New York, 2002.
6 Christophe Claramunt and Bin Jiang. An integrated representation of spatial and temporal relationships between evolving regions. *Journal of geographical systems*, 3(4):411–428, 2001.
7 E.R. Cohen, T. Cvitas, Jeremy Frey, B. Holmstrom, K. Kuchitsu, R. Marquardt, I. Mills, Franco Pavese, M. Quack, Jürgen Stohner, Herbert Strauss, M. Takami, and A.J. Thor. *Quantities, Units and Symbols in Physical Chemistry, IUPAC Green Book*. 01 2007.
8 Nationale EnergieAtlas. Windturbines. `https://www.nationaleenergieatlas.nl/kaarten`.
9 Rijksinstituut voor Volksgezondheid en Milieu ESRI Nederland. Netto langgolvige straling - knmi (bèta). `https://www.arcgis.com/home/item.html?id=61dc70b9664f4c3c9818427203abdd73`.
10 Rijksinstituut voor Volksgezondheid en Milieu ESRI Nederland, Klimaateffectatlas. Netto langgolvige straling - knmi (bèta). `https://www.arcgis.com/home/item.html?id=2752ef4366d5486388efde9ce3fd75fc`.
11 Kit Fine. Identity criteria and ground. *Philosophical Studies*, 173(1):1–19, 2016.
12 Mark Gahegan. Fourth paradigm giscience? prospects for automated discovery and explanation from data. *International Journal of Geographical Information Science*, 34(1):1–21, 2020.
13 Peter Gerstl and Simone Pribbenow. Midwinters, end games and bodyparts. In *International Workshop on Formal Ontology-93*, pages 251–260, 1993.
14 Giancarlo Guizzardi. On the representation of quantities and their parts in conceptual modeling. In *FOIS*, pages 103–116, 2010.
15 Anthony JG Hey, Stewart Tansley, Kristin Michele Tolle, et al. *The fourth paradigm: data-intensive scientific discovery*, volume 1. Microsoft research Redmond, WA, 2009.
16 Torsten Hägerstrand. What about people in regional science? In *Papers of the Regional Science Association*, volume 24, pages 7–21, 1970.

17 William H Inmon. *Building the data warehouse.* John wiley & sons, 2005.

18 Ralph Kimball and Margy Ross. *The data warehouse toolkit: the complete guide to dimensional modeling.* John Wiley & Sons, 2011.

19 r. a. Kocourek. Quantity, May 2018. URL: https://www.encyclopedia.com/literature-and-arts/language-linguistics-and-literary-terms/language-and-linguistics/quantity.

20 Werner Kuhn. Core concepts of spatial information for transdisciplinary research. *International Journal of Geographical Information Science*, 26(12):2267–2276, 2012.

21 Martin Lachmair, Susana Ruiz Fernández, Korbinian Moeller, Hans-Christoph Nuerk, and Barbara Kaup. Magnitude or multitude–what counts? *Frontiers in psychology*, 9:522, 2018.

22 Edward Jonathan Lowe. Entity, identity and unity. *Erkenntnis*, 48(2):191–208, 1998.

23 Sebastiaan H Mannaerts. Extensive quantities in thermodynamics. *European Journal of Physics*, 35(3):035017, 2014.

24 Donna Peuquet, Barry Smith, and Berit O Brogaard. The ontology of fields. 1998.

25 Brandon S Plewe. A case for geographic masses. In *14th International Conference on Spatial Information Theory (COSIT 2019).* Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

26 Florian Probst. Observations, measurements and semantic reference spaces. *Applied Ontology*, 3(1-2):63–89, 2008.

27 Otto Redlich. Intensive and extensive properties. *Journal of Chemical Education*, 47(2):154, 1970.

28 Bradley Rettler and Andrew M. Bailey. Object. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, Winter 2017 edition, 2017.

29 Simon Scheider and Mark D Huisjes. Distinguishing extensive and intensive properties for meaningful geocomputation and mapping. *International Journal of Geographical Information Science*, 33(1):28–54, 2019.

30 Simon Scheider, Enkhbold Nyamsuren, Han Kruiger, and Haiqi Xu. Geo-analytical question-answering with gis. *International Journal of Digital Earth*, pages 1–14, 2020.

31 David Sinton. The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. *Harvard papers on geographic information systems*, 1978.

32 Stanley Smith Stevens et al. On the theory of scales of measurement. 1946.

33 Leonard Talmy. The relation of grammar to cognition–a synopsis. *American Journal of Computational Linguistics*, pages 16–26, 1978.

34 Richard C Tolman. The measurable quantities of physics. *Physical Review*, 9(3):237–253, 1917.

35 Centraal Bureau voor Statistiek. Cbs gebiedsindelingen. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/cbs-gebiedsindelingen.

36 Centraal Bureau voor Statistiek. Kerncijfers wijken en buurten, 2012. https://www.cbs.nl/nl-nl/maatwerk/2011/48/kerncijfers-wijken-en-buurten-2012.

37 Centraal Bureau voor Statistiek. Rijksmonumenten. https://opendata.cbs.nl/statline/portal.html?_la=nl&_catalog=CBS&tableId=84775NED&_theme=423.

38 Centraal Bureau voor Statistiek. Totale vervoersprestatie in nederland; vervoerwijzen, regio's. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84687NED/table?ts=1617881414400.

39 Rijksinstituut voor Volksgezondheid en Milieu. Genetische structuur van de boommarter populatie (martes martes). https://data.overheid.nl/dataset/abc85851-cede-4ddd-bc4c-cb69ad99b331.

40 Haiqi Xu, Ehsan Hamzei, Enkhbold Nyamsuren, Han Kruiger, Stephan Winter, Martin Tomko, and Simon Scheider. Extracting interrogative intents and concepts from geo-analytic questions. *AGILE: GIScience Series*, 1:1–21, 2020.