Computation and Elicitation of Valleyness

Ralph K. Straumann
Ross S. Purves
Department of Geography, University of Zurich, Switzerland

Existing methods for land surface form characterisation often focus on relatively simple landform element classifications and do not evaluate results with large scale participant experiments. This piece of research takes a semantics-grounded approach to characterising the landform *valley*. Based on definitions three algorithms to characterise valleys in a fuzzy way are introduced. Comparison of the results to assessments regarding the degree of being in a valley gained from over 800 participants in a questionnaire survey yields significant amounts of explained variance ($R^2 = 0.35$–0.37). Furthermore, accounting for very ambiguously perceived stimuli showing vast low places leads to markedly improved regressions ($R^2 = 0.45$–0.49), weighting of the data with a measure of uncertainty in judgment even more so ($R^2 = 0.50$–0.55).

Keywords: Digital terrain modelling, geomorphometry, landform, valley.

1. Introduction

Research into techniques for describing land surface form(s) has its roots primarily in geomorphology (e.g. Maxwell 1870) and in soil science (which developed methods for landform element classification; e.g. Pennock et al. 1987, Irvin et al. 1997). Beside motivations which pertain to the natural sciences, however, an important point can be made that land surface forms are of interest

to humans, for example, in place descriptions and the like. For example, Smith and Mark (2003: 419) posited that "the naïve or folk disciplines appear to work exclusively – or at least overwhelmingly (…) – with object-based representations of reality" and this implicitly suggests that these are favoured over field-based conceptualisations which might equate to computations of continuous terrain parameters like slope gradient or curvatures. This means that the description of land surface form in terms of landforms (i.e. larger regions of similar form and labelled as, for example, *mountain*, *valley* or *plateau*) may be more meaningful to, and easier to understand for, most people. Additionally, we would argue that qualitative measures of such characterisations may be advantageous, for example, when dealing with fuzzy phenomena or when being confronted with uncertainty in digital elevation model (DEM) data which renders computation of accurate terrain parameters at a particular scale difficult or impossible.

With regard to research on landform delineation or characterisation current literature is largely focused on topographic eminences such as mountains and ranges (e.g. Fisher et al. 2004, Chaudhry and Mackaness 2008, Deng and Wilson 2007). After delineating valley floors (Straumann and Purves 2008) a natural extension is to characterise the landform *valley*. We argue that carefully considering the characteristics which a landform term implies is crucial before devising an algorithm (cf. ibid., Deng 2007, Straumann 2009) rather than directly applying a generic algorithm. Thus, our methodological starting point is the term *valley* before we develop and apply techniques for deriving *valleyness* (here related to *being in a valley*; similar to the *peakness* of Fisher et al. 2004). Subsequently, the algorithmic valleyness is evaluated, and since we are interested in deriving a characterisation which fits with 'folk' notions of valleyness, we employ data we gathered in a questionnaire using georeferenced photographs as stimuli. To our best knowledge, this is the first piece of research ever using such an approach to evaluate a computational characterisation of landforms.

Mark and Turk (2003), Mark et al. (2007) and Mark (2009) have argued that landform categories are culture- and language-dependent and coined the term of ethnophysiography. In our methodology we still relied, however, on international geomorphological literature in English for obtaining definitions which were in turn used for developing valleyness algorithms. The questionnaire experiment to derive estimates of being in a valley in order to compare these to algorithmic valleyness measures has received many more responses from German-speaking people than from English-speaking people. However, the decision of participants to fill in the questionnaire in English or in German and the current place of residence (if indicated) is the only information available which might allow inferences on cultural and language backgrounds. Clearly, in the age of globalisation (and considering Switzerland as a country has four official national languages) this information is too weak to allow for any inferences regarding differences related to culture or mother tongue in our study. Thus, in this study it

was assumed that the questionnaire targeted a mostly "western" audience and that the differences in landform perception and appreciation between, for example, people with a Swiss, German, Italian or British cultural background are not as strong as, for example, between people with a "western" American and an aboriginal American background. For simplicity and legibility we use the English terminology throughout the paper, but present the questionnaire phrasing in both the English and the German version.

2. Background

Characterisations of land surface form are used, for example, to analyse and infer large-scale terrain characteristics (Hammond 1954), soil properties, the presence of bundles of geomorphic processes and the movement and distribution of water and soils (Pennock et al. 1987) as well as the description of landscape in human-understandable form for, for instance, place descriptions (cf. Fisher et al. 2004).

Besides the description of land surface form by (a set of) geomorphometric attributes (Wilson and Gallant 2000, Moore et al. 1991) approaches at delineating or characterising landform elements or landforms are numerous. Many approaches cluster or classify DEM cells based on values of various kinds of curvature and slope. A very popular generic classification is that into six morphometric features (peak, ridge, pass, plain, channel, pit) which has been extended by Wood (1996) and Fisher et al. (2004) to allow (fuzzy) multiscale treatment of surface form. Similar curvature- and slope-based schemes were proposed by, for example, Ruhe (1960; cited in Pennock et al. 1987), Troeh (1965), Pennock et al. (1987), Dikau (1989) and have been used to derive both crisp (e.g. Pennock and Corre 2001, Pennock 2003, Bolongaro-Crevenna et al. 2005) and fuzzy (e.g. MacMillan et al. 2000, Schmidt and Hewitt 2004) classifications. Fisher et al. (2004) reasoned about the essence of peaks and their relationship to summits and used fuzzy multi-scale morphometric feature classification resulting in fuzzy areas of peakness associated with summits. However, the classification is very generic and the classes are not *a priori* tied to a specific geomorphologic feature. While in the case of mountain peaks and peakness the analogy may work, we doubt that channelness alone could characterise valley floors (Straumann and Purves 2008) and thus, valleys as a whole. However, methods have been developed to delineate topographic depressions. For example, Tribe automatically delineated valley heads (1991) and valley floors (1992) from DEMs. While the former was done using a region growing algorithm on seed cells near the upper end of drainage branches, the latter included threshold slope gradient to eliminate insignificant depressions and a user-defined analysis neighbourhood to reduce discontinuities in wide valleys. Miliaresis and Argialas (1999) applied a gradient-dependent region growing procedure for the analysis of basins, piedmont slopes and mountains. They operationalised seed cells for the region growing of basins as pixels with higher-

than-mean flow accumulation. Although, through low gradients and undefined flow direction "the seeds for basins did not give the impression of a network" (Miliaresis and Argialas 1999: 720), the final delineation was favourably compared to a physiographic map of the region. Gallant and Dowling (2003) proposed the multi-resolution valley bottom flatness index (MRVBF). This method is based on the application of slope (representing flatness) and elevation percentiles (representing lowness with respect to surroundings). The index is fuzzy and intended for analysing depth of deposit and groundwater constrictions, and delineating hydrologic and geomorphic units. Together with its compliment, the multi-resolution ridge-top flatness index, MRVBF can be used to separate valley floors, hill-slopes and ridge-tops. For validation Gallant and Dowling (2003) compared their results to a range of methods including field-mapping of soil-landscape associations by a surveyor. Wang and Laffan (2009) in turn compared their approach – Multi-Scale Valleyness (MSV; Wang et al. (in press)) – favourably to both flow routing algorithms (D8, D-Inf), which do not characterise valleys as areas, but are stated to perform well regarding valley center lines, and MRVBF. The validity of MSV itself was assessed qualitatively.

Most of the approaches implicitly consider landform semantics in developing algorithms, but the algorithm itself either focus on relatively simple landform elements or a continuous, thresholded value, and none, to our knowledge, evaluates results other than by comparison with other algorithms or assessment by an expert. Hence, in this paper we exemplify a semantics-grounded approach at characterising valleys from DEMs as an example of a description of land surface in terms of landforms. Since the characterisation is explicitly aimed at a human audience the validation of the derived measures compares the results against data gained from a large number of participants in a questionnaire survey.

3. Methods

3.1 Study area and data

The study area comprises a buffered bounding box of Switzerland (Fig. 4) covering a considerable part of the European Alps. The hole-filled Shuttle Radar Topography Mission (SRTM) DEM (Version 3) was used as data. It was obtained at 3 arcseconds resolution from Jarvis et al. (2006), projected into the Swiss national projection system and resampled to 100 metres resolution.

3.2 Conceptualisation of valleyness

The notion of drainage basins was chosen as a starting point for developing a method to extract valley floors in a first step and to eventually compute valleyness or the degree of being in a valley from a DEM. Drainage basins enclose valleys; however, only in small headwater drainage basins is this enclosing relation one-to-one. Drainage basins of higher hydrological order or magnitude usually contain several river reaches and several valleys or valley stretches.

Thus, in order to obtain a tessellation of space related to valleys, we clip drainage basins of a certain Shreve stream magnitude with contributing drainage basins of lower magnitudes (cf. also Straumann and Purves 2008, Demoulin et al. 2007) thus defining a *drainage sub-basin*.

Exploring definitions for and properties of *valley floor* and *valley* in a range of geomorphologic dictionaries and textbooks, and standards (Bates and Jackson 1990, Ahnert 1998, Huggett 2007, Allaby and Allaby 1999, Mayhew 2004, Whittow 2000, Lapidus et al. 2003, ANSI 1998, OED 2009) we find that valleys typically share the following properties:

− Valleys are low areas or depressions relative to their surroundings.
− Valleys are elongated.
− Valleys are (gently) sloping.
− Valleys often contain a stream or a river.

Quite typically, these definitions contain uncertainties and need to be employed cautiously. For example, the criterion of being "(gently) sloping" most probably refers only to the longitudinal axis of valleys and does not necessarily state anything about their side slopes (although the definitions do not mention this explicitly).

Valley floor is defined by Bates and Jackson (1990) as "the comparatively broad, flat bottom of a valley; (...)". In conclusion we can also say that it will inherit the characteristics of valleys listed above. This can be illustrated with the assertion that valleys are low areas relative to their surroundings. Since valley floor is defined to be the lowest part of a valley (or more precisely: the lowest part of the valley cross-sectional profile), it, too, is certainly lower than its surroundings.

Due to these defining characteristics thalwegs (i.e. the lowest lines within valleys) are good candidates for conceptual cores of valleys and their floors (cf. Deng 2007: 412). Consequently, valley floors could be understood as relatively flat areas bordering thalwegs (cf. Straumann and Purves 2008).

We can thus suggest that valleys are elongate, low relative to their surroundings, and – as opposed to topographic eminences – concave areas of the earth's surface at a particular scale. The latter circumstance appears to be attested to in central European languages – consider the following examples of describing one's location with reference to a valley and to a mountain:

English:	"being *in* a valley"	"being *on* a mountain"
German:	"*in* einem Tal sein"	"*auf* einem Berg sein"
French:	"être *dans* une vallée"	"être *sur* une montagne"
Italian:	"essere *in* una valle"	"essere *su* una montagna"

In the above languages, referring to valleys, one is *in*, referring to mountains one is *on*. A valley thus appears to evoke a sense of *containment* which fits well with the notion of concavity, which in turn is a consequence of lowness with respect to the surroundings. Both notions identified regarding valleys (lowness and concavity) can be taken advantage of in operationalising measures of the degree of being in a valley.

Regarding elevation, we hypothesise that relative elevation in each drainage sub-basin above the valley floor is informative regarding that degree.

Regarding concavity, the problem is inverted. Rather than taking concavity as a measure of the degree of being in a valley, convexity was considered as having detrimental effects on the apprehension of a location being in a valley. Specifically, we looked at convexities encountered when going from the valley floor outward towards the drainage divides. Consider the example valley cross-sections in Fig. 1. In Situation A, position 1 is clearly in the valley, even located on the valley floor. Position 2, however, is most certainly not fully in the valley anymore but rather on the adjacent topographic eminence. In Situation B, 3 is certainly in the valley (even on the thalweg). At location 4 (a convex break of slope) something significant happens. Immediately above 4 an observer's view of the valley floor is partly or completely blocked. Above location 4 an observer very probably feels less of the valley's affordance of containment than below 4. At location 5 again, the observer is rather on a topographic eminence than in a topographic depression. Situation C depicting a *U-in-U* example is even more complex. Here, again, location 6 is considered to be in the valley and location 7 is a crucial break of slope. Here, even more clearly than in situation B, above the break of slope an observer has an obstructed view into the valley and very probably feels a bit less in it. At location 8 there is a second profound concavity before the slope rises steeply to the drainage divide, where an observer again would, we suggest, feel more on a topographic eminence, than in a valley.
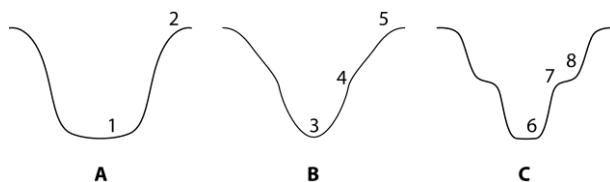

Figure 1. Various valley cross-sections.

It is hoped that the elevation-based approach may be sensible in situations similar to A with a pronounced concave profile almost to the drainage divide. In such a situation the convexity-based approach would assign a high degree of being in a valley even to locations next to the drainage divide. Also, an elevation-based approach could perform better in a valley with nearly straight slopes,

where the convexity-based approach is sensitive even to slight convexities. Additionally to the two approaches, their combination will be explored.

3.3 Operationalisation

The SRTM DEM was filled and D8 flow directions and flow accumulation calculated. In D8, each cell of a raster is defined to drain into the neighbouring cell which lies lower than the cell under consideration and which has the lowest elevation of all neighbouring cells (Jenson and Domingue 1988, ESRI 2010). Using a channel initiation threshold of ≥ 500 cells a network of thalwegs and its Shreve stream magnitudes were generated. The Shreve method assigns all headwater thalwegs a magnitude of 1 and sums magnitudes wherever thalwegs meet (Huggett 2007: 191p, cf. Fig. 2). Thus, the thalweg network was segmented along general flow direction. Subsequently, drainage basins of magnitude x were clipped by all drainage basins of magnitude y < x (Fig. 2).
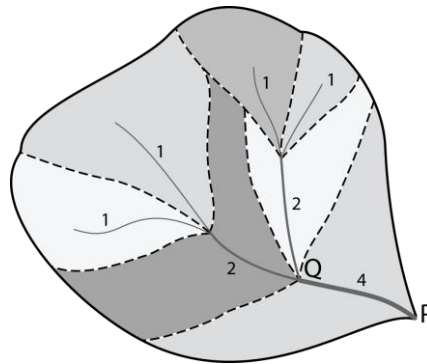

Figure 2. Clipping of drainage basins. Solid outline represents original drainage basin of point P, dashed outlines and different shades of grey represent several drainage sub-basins pertaining to streams/thalwegs (grey lines) of different Shreve magnitude (numbers).

Using the affiliation of pixels to drainage sub-basins and the thalweg raster a region growing procedure was carried out to delineate the flat valley floor bordering the thalwegs. According to the reasoning in the previous section region growing was spatially constrained to drainage sub-basins. A raster cell i was classified as valley floor, when at least one of its neighbours was a seed cell or a grown valley floor cell and at least one of the following conditions was met (Straumann and Purves 2008):

Cardinal neighbours: $\tan(\gamma_{crit}) \cdot \lambda \geq elev_i - elev_{seed} \geq 0$ (1)

Diagonal neighbours: $\tan(\gamma_{crit}) \cdot \sqrt{2}\lambda \geq elev_i - elev_{seed} \geq 0$ (2)

where γ_{crit} is the gradient threshold [°], λ is the cell size [m], $elev_i$ and $elev_{seed}$ are the elevations [m] of cell i and seed cell, respectively.

Region growing was run iteratively until no new valley floor cells were detected. The region growing makes valley floors contiguous and limits them to areas which can be reached from the thalweg with a low slope, thus matching the definitions for valley floors in the previous section. The method does not use a traditional slope gradient computation algorithm but a very simple notion of cell-to-cell gradient thus attaining a smaller footprint (20,000 m^2 versus 90,000 m^2 for a 3 by 3 cells computation). This may be attractive due to the resolution of the SRTM DEM (100 metres) which is quite coarse already anyway.

   A range of gradient thresholds (γ_{crit}) from 0.25° to 3° in steps of 0.25° were tested and the respective results were subjected to qualitative visual examination. Overlaying the delineated valley floor areas onto terrain parameters such as a hillshaded relief and a gradient raster we found the best accordance of the delineation with our subjective judgment for a value of 1.5° and adopted this as the threshold gradient. The chosen method is certainly dependent both on this threshold gradient and on the characteristics of the SRTM dataset, namely its horizontal resolution. Further testing and, potentially, adaptations would be required for changing this method to be used with other elevation datasets.

   The delineated valley floors were assigned a fuzzy membership value of 1 with respect to the class *valley*, reflecting certainty that any point in the valley floor is an essential part of its encompassing valley.

In order to obtain a space-filling measure of the degree of being in a valley, the side slopes of drainage sub-basins must also be characterised. These were allowed to have fuzzy *valley* membership values varying from 1 (definitely part of the valley) to 0 (definitely not part of the valley). In accordance with the conceptualisations in the previous section three algorithms were implemented using the DEM, the valley floor and the drainage sub-basin rasters. The convexity-based approach additionally uses a raster of the Euclidean distances of each non-valley floor pixel to the closest valley floor pixel of the same drainage sub-basin.

   Firstly, the simple elevation-based computation of valleyness in individual drainage sub-basins was devised as:

$$v_{e,j} = 1 - \frac{elev_j - elev_{min}}{elev_{max} - elev_{min}} \quad \in [0, 1]$$
         (3)

where $v_{e,j}$ is the valleyness of pixel j; $elev_j$, $elev_{max}$ and $elev_{min}$ are elevation values [m] of cell j and maximum and minimum elevation within the drainage sub-basin, respectively.

The convexity-based approach requires more computation, trying to mimic the reasoning applied in Fig. 1. Starting from the valley floor, cells in each drainage sub-basin are binned according to their Euclidean distance to the closest valley floor cell in the same drainage sub-basin (Fig. 3, left).
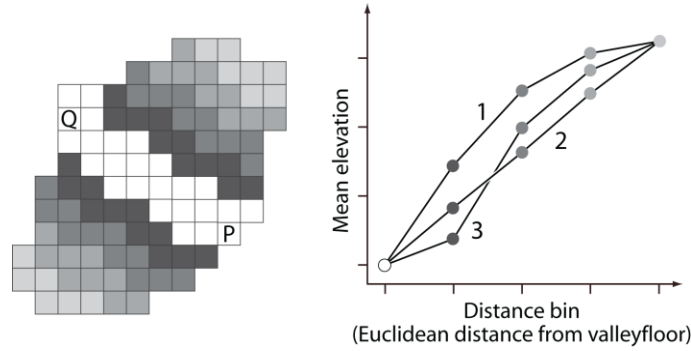

Figure 3. Formation of distance bins and curvature calculation in drainage sub-basin. Left: White cells denote valley floor and different shades of grey represent distance bins in the drainage sub-basin of the thalweg stretch Q–P (cf. Fig. 2). Right: Mean elevations of distance bins are related to each other to compute curvature. Example 1: a convex valley side profile, example 2: a fairly straight profile, example 3: a profile with a marked concavity near the valley floor and convex parts near the drainage divide.

The zoning using drainage sub-basins is a requirement. Firstly, it increases the probability that the algorithm is dealing with very similar forms of the two opposing valley sides. Also, it yields manageable analysis units making the computation of mean elevations (of valley floor and distance bins) sensible. The binning distance was chosen as 150 metres (i.e. 1.5× DEM resolution). This was deemed a sensible distance since it maintains as much information as possible while also avoiding the occurrence of empty distance bins due to known effects of the D8 flow direction algorithm. Again starting from the valley floor, a curvature measure, c_i, is calculated for every distance bin i according to:

$$c_i = (elev_{i+1} - elev_i) - (elev_i - elev_{i-1}) \qquad (4)$$

where $elev_{i-1}$, $elev_i$, $elev_{i+1}$ signify the mean elevation of the respective distance bin (Fig. 3, right). Then each bin is revisited and the curvature value classed as

concave, planar or convex (equation 5). Concave distance bins have their curvature values set to zero. Convex distance bins signifying convex breaks of slope as in Fig. 1 are weighted by the number of pixels, n_i, in the respective distance bin.

$$c_i \begin{cases} > 0 \Leftrightarrow i \text{ is concave} & \Rightarrow c_i = 0 \\ = 0 \Leftrightarrow i \text{ is planar} & \Rightarrow \text{do nothing} \\ < 0 \Leftrightarrow i \text{ is convex} & \Rightarrow c_i = c_i \cdot n_i \end{cases} \tag{5}$$

The weighted (negative) curvature values are then inverted in order to facilitate calculations and are summed up over the whole of the drainage sub-basin to yield total sub-basin convexity, C:

$$C = \sum_{i=0}^{m} - c_i \tag{6}$$

Finally, each distance bin is revisited and assigned a valleyness value which is computed based on the cumulative convexity outward from the valley floor to that distance bin and the total convexity (equation 6) encountered in the respective drainage sub-basin. These two are combined in the convexity-based measure of valleyness for the member cells of distance bin j, $v_{c,j}$:

$$v_{c,j} = 1 - \frac{\sum_{i=0}^{j} - c_i}{C} \quad \in [0, 1] \tag{7}$$

The resulting raster is combined with the binary valley floor raster to yield the convexity-based valleyness measure.

The combination of the two approaches (relative elevation and convexity) is made by simply computing the mean of the two measures (equations 3 and 7) for each raster cell.

3.4 Participant experiment

An experiment in the form of a questionnaire survey was carried out in order to compare the results of the valleyness algorithms described in the previous section to assessments by questionnaire participants.

Participants

Participants were recruited in three different ways. A mass e-mailing was sent to 8,751 members of the University of Zurich who have given their consent to receiving such communication. Similarly, an e-mailing targeted a subsample of the members of the *Geomorphometry* mailing list (Geomorphometry 2009). Lastly, an invitation was sent out to friends, relatives and acquaintances of the

first author. The survey thus purposefully aimed at a diverse audience. Within few weeks over 800 people participated in the survey.
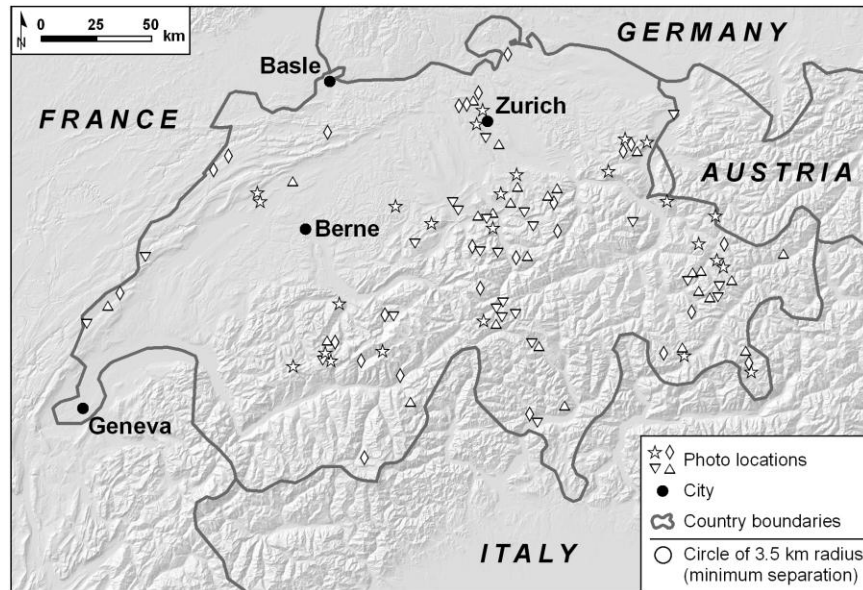
Material

Figure 4. Distribution of 100 remaining photo locations. The four symbols of photo locations represent their affiliation to one of the four image groups.

The initial set of candidate stimuli consisted of 5,503 georeferenced photographs. These were taken by researchers not explicitly for this study but both for a landscape description project (see *Acknowledgements*) and in their spare time on planned trips (which explains some of the spatial autocorrelation). After application of a random sampling scheme picking photographs with a minimum separation distance of 3.5 kilometres (the maximum for retaining a reasonable number of images), 200 photos remained in the dataset. Subsequently, photographs were excluded from the set based on properties which were considered inappropriate for stimulus images and had been defined *a priori*. Reasons for exclusion included close-ups, dense atmosphere hindering sight considerably (e.g. in fog or inside a dense forest) and a predominance of built structures. The manual exclusions yielded a set of 145 photographs which were then randomly reduced to 100 photographs and randomly allocated to four image groups (Fig. 4). Some example photographs can be seen in Fig. 5.

Fig. 5. Example photographs which were used in the questionnaire (images in the questionnaire were in colour).

Design and procedure
The questionnaire was carried out both in English and in German using the dynamic scripting language PHP which produced a single HTML webpage through which participants were guided. Participants were randomly allocated to one of the four image groups. The display order of the stimuli of the respective image group was randomly shuffled for every participant in order to prevent order and context effects (Montello and Sutton 2006: 94).

Fig. 6 shows an example question and stimulus image as displayed in a participant's browser (German version). Beneath the stimulus image there is a Likert scale (Trochim 2006) which the participants used to answer the questions. The question read:

> Bitte bewerten Sie für das obenstehende Bild die Talhaftigkeit des **Standorts des Fotografen**, wobei 5 bedeutet: "ist bestimmt in einem Tal" und 1: "ist bestimmt NICHT in einem Tal".

and

> For the above picture, rate the "valleyness" of the **photographer's location**, where 5 is definitely in a valley and 1 is definitely NOT in a valley.

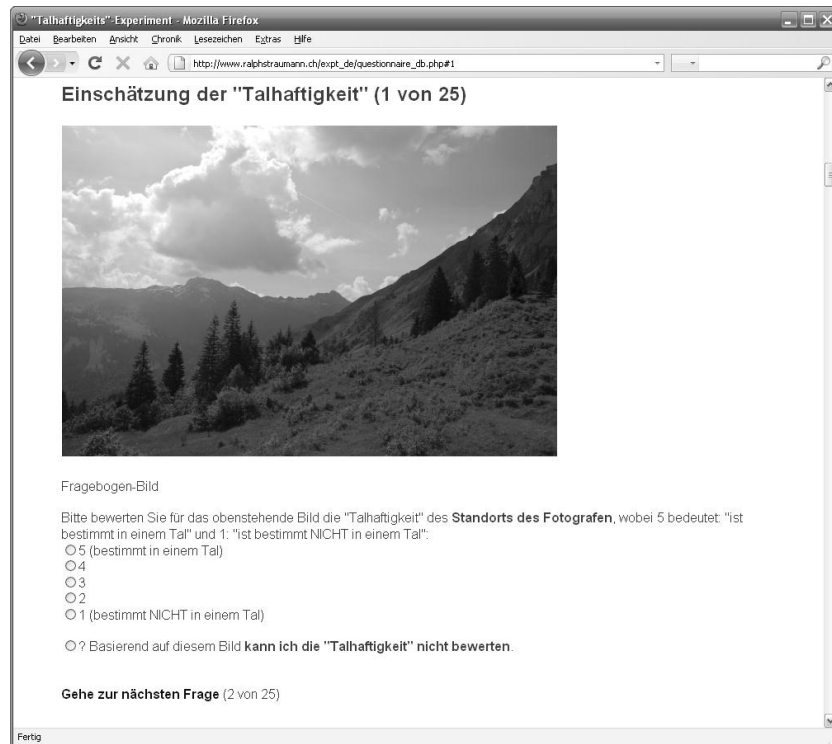in the German and the English version, respectively.

Figure 6. Example of a questionnaire item with stimulus image, question and Likert scale (for the wording please refer to the text).

The Likert scale features 5 levels which were labelled with numbers and the two most extreme levels with the following accompanying texts (respectively):

> (bestimmt in einem Tal)
> (bestimmt NICHT in einem Tal)

and

> (definitely in a valley)
> (definitely NOT in a valley)

The Likert scale was completed by an additional option which read (respectively):

> Basierend auf diesem Bild kann ich die "Talhaftigkeit" nicht bewerten

and

> From this picture I cannot estimate "valleyness"

Since the Likert scale data is discontinuous, the median was used in the subsequent statistical analyses.

4. Results and discussion

4.1 Algorithmic valleyness

Figs. 7, 8 and 9 show the results of the valleyness computations according to equations (3), (7) and the mean of the two. The rasters have been moderately filtered using a low-pass (mean) filter in a circular neighbourhood with a 3 cells radius, since the valleyness computation based on drainage sub-basins naturally exhibits more or less abrupt boundaries at drainage divides and distance bin boundaries. The filtered rasters still have a correlation with the unfiltered ones of > 0.96. In the statistical analyses the raw rasters were employed, however.

Comparing ve (Fig. 7) with vc (Fig. 8) one can see a marked difference between the two mainly in the large valley floors (e.g. locations 1 and 2 in Fig. 7). While in vc valley floors have per definition a valleyness of 1, this is not the case with ve. There, valley floors, too, are subjected to the computation of relative elevation.

In places such as the Rhine basin (location 1 in Fig. 7) where drainage sub-basins are located entirely within relatively flat regions this results in areas with low absolute but high relative elevation having low valleyness values. Of course, through averaging this effect is alleviated somewhat in v. Also, one can see in Figs. 7 and 8 that ve is more continuous than vc, since the latter is computed on distance bins whereas the former is per pixel.
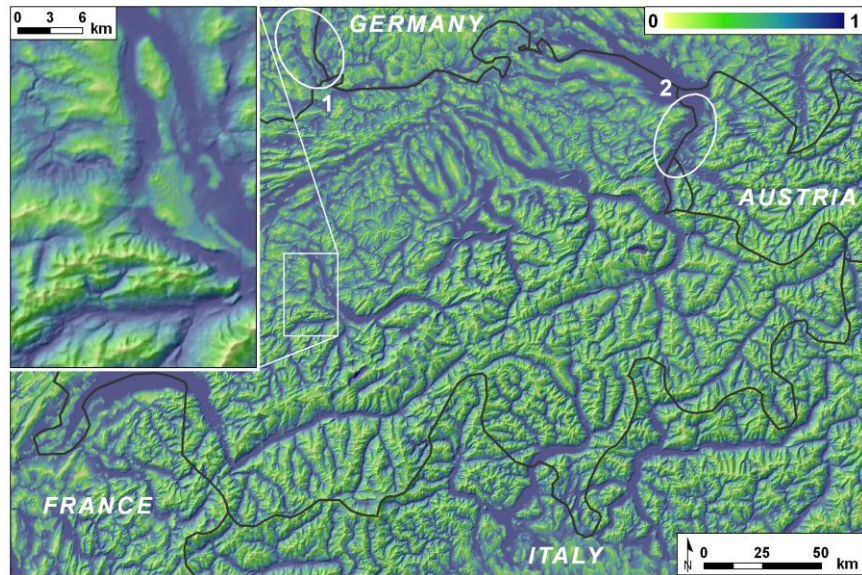
Figure 7. Elevation-based valleyness v_e. The backdrop (also in Figs. 8 and 9) is a hillshaded relief, the black line is the boundary of Switzerland. 1: Rhine graben on French-German border, 2: Rhine valley on Swiss-Austrian border.
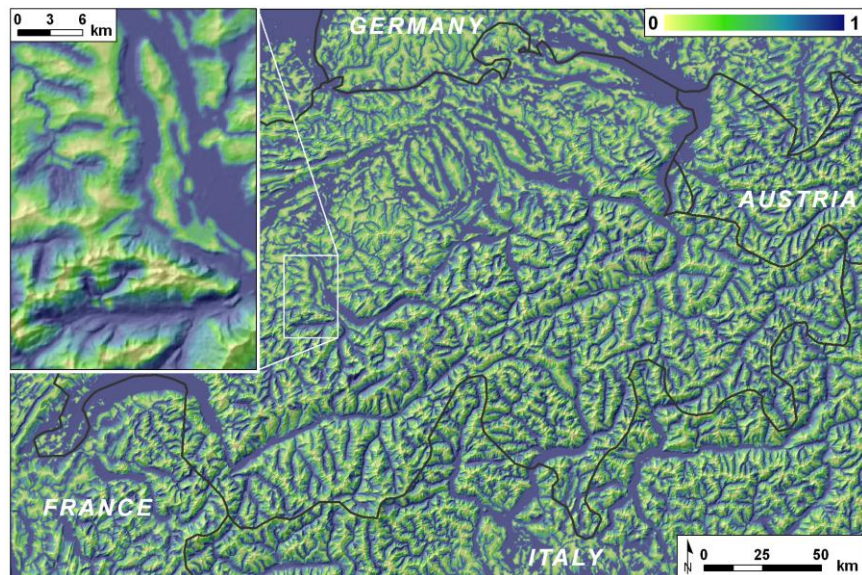

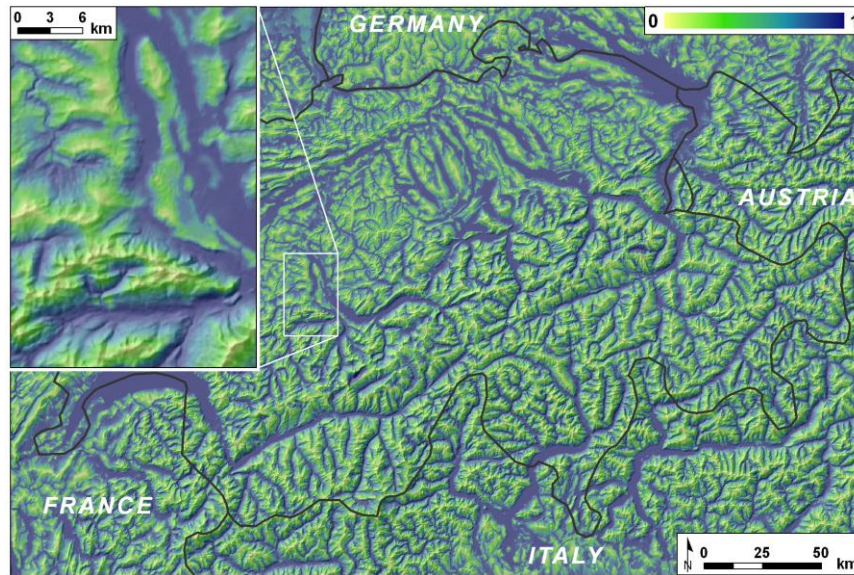Figure 8. Convexity-based valleyness v_c.

Figure 9. Combined valleyness *v*.

4.2 Questionnaire estimates

Questionnaire participants

A total of 810 people answered the questionnaire. Regarding their expertise they chose between the following three options:

− I am a researcher in the field of geosciences (e.g. geography, geomorphology, geomorphometry, ...)
− I am a student in the field of geosciences (e.g. geography, geomorphology, geomorphometry, ...)
− I am neither of the above

These were subsequently considered as *researchers*, *students* and *laypersons*, respectively. The majority of people designated themselves as laypersons ($n = 651$, 80.4%), some as students ($n = 70$, 8.6%) and some as researchers ($n = 47$, 5.8%) (42 persons did not answer the question at all). Unsurprisingly, there was a marked difference regarding languages between these groups. While almost 47% of the researchers answered the English questionnaire, that proportion was drastically lower among laypersons and students (1.4% and 2.7%, respectively). Thus, the proportion of German questionnaires is much higher than that of English ones (95.8% versus 4.2%), since a main source of partici-

pants was students of the University of Zurich (from fields outside the geosciences).

The distribution of participants' age is unequal with a large positive (right) skewness with a peak in the "20–29 years" class which most students belong to.

The online questionnaire had a feature to track and record the time participants spent to answer the 20 questions (more precisely: time span from loading the page initially to pressing the "Submit" button). Correcting for one outlier, the median of this time span was 5.7 minutes, the mean 8 minutes and the standard deviation 11.3 minutes.

Regarding residence there was a strong bias towards Europe and within Europe towards Switzerland; few people participated from outside Switzerland. Within Switzerland the agglomeration and city of Zurich was very heavily represented along with some of the other larger cities of Switzerland. Thus most of the participants could be attributed to what is termed *Mittelland* (the flatter, more populated and crescent-shaped part of Switzerland between the Alps and the Jura mountains).

Overall, the sample is thus biased towards German-speaking, young laypersons – probably mainly towards students from the university mailing list from other departments than Geography and towards other non-academic laypersons.

Subsequent statistical analysis will investigate whether there are differences in the estimates of the degree of being in a valley from different expertise groups, before the estimates are compared to the algorithmic valleyness.

Differences between image and expertise groups
In the analysis of the questionnaire data one cannot assume that, firstly, the image groups are equivalent and, secondly, the three expertise groups answered the questions identically. Thus, the (non)existence of such differences has to be assessed. A scheme is adopted which first tests for differences across the image groups within the expertise groups. If there are no such differences the four image groups can be aggregated within every expertise group.

To analyse the questionnaire data, relative counts which account for the different number of participants allocated to the image groups were computed from the counts associated with Likert scale items of the questionnaire (V_1 (definitely not in a valley), V_2, V_3, V_4, V_5 (definitely in a valley) and V_{99} ("from this picture I cannot estimate valleyness")):

$$rV_i = \frac{V_i}{V_{99} + \sum_{j=1}^{5} V_j} \quad \forall i : i \in \{1, 2, \ldots, 5, 99\} \tag{8}$$

Additionally, the median of the estimates, v_{median}, was computed for every stimulus in every combination of image groups and expertise group. Values of V_{99} were disregarded in the latter computation, since they could not be sensibly

placed on a numerical scale along with the other values. Testing for differences in v_{median} among different image groups within the same expertise group yielded no statistically significant differences (Kruskal-Wallis H tests with $p = 0.221$–0.276). Thus, the individual image groups per expertise group were aggregated for the subsequent investigations of potential differences between expertise groups.
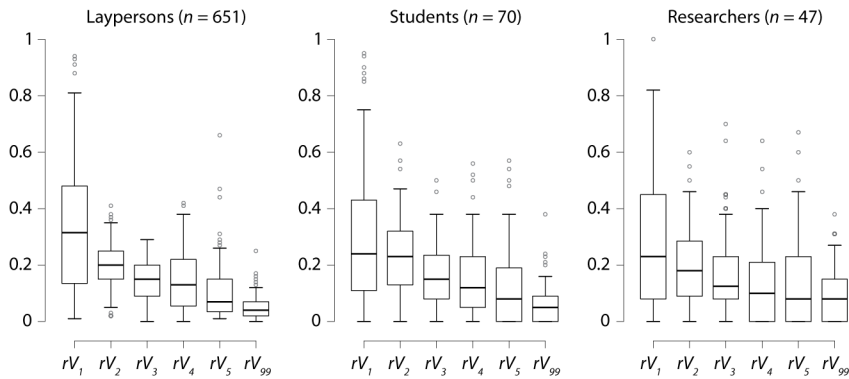

Figure 10.   Boxplots of relative counting variables by expertise group. Outliers (°) lye out > 1 interquartile range.

The boxplots in Fig. 10 show that in all three expertise groups there are lower proportions for higher estimates of being in a valley (i.e. towards V_5). The proportion of V_{99} ("from this picture I cannot estimate valleyness") is lowest in all expertise groups.

Interestingly, researchers more often opted for V_{99} ("from this picture I cannot estimate valleyness") than the other groups, i.e. they were more liable to acknowledge uncertainty in judgment. However, the sample of researchers is with $n = 47$ relatively small and thus a small absolute number of answers V_{99} have a considerable influence on the distribution. Also, the effect was not proven statistically significant (Kruskal-Wallis H test with $p = 0.143$).

The means of all v_{median} values (Table 1) show that researchers tend to have more estimates of high degrees of being in a valley (mean of v_{median} closer to 5) and less estimates of low degrees of being in a valley than the students and especially the laypersons do. That means that overall the researchers estimated the photographer's locations for the stimuli more in a valley than students and these in turn more than laypersons. According to the skewness, researchers and students spread their answers more equally across the spectrum than laypersons did. The broader peak of the distribution is also reflected in bigger negative kurtosis values.

Table 1. Aggregate statistics of v_{median} stratified according to expertise.

Expertise	Range	Median	Mean	Stddev.	Skewness	Kurtosis
Layperson	[1, 5]	2	2.240	.9936	.468	−.524
Students	[1, 5]	2	2.450	1.0384	.419	−.555
Researchers	[1, 5]	2	2.495	1.1180	.328	−.881

However, while the distributions of the relative counting variables are slightly different in different expertise groups, statistical testing shows that differences in terms of v_{median} are not statistically significant (Kruskal-Wallis H test with $p = 0.231$). Thus, while some systematic trends could be pointed out, lack of significance implies that the data pertaining to different expertise groups can also be lumped together for comparison with the algorithmic valleyness measures.

Comparison with algorithmic valleyness measures
As an additional piece of information the uncertainty associated with the judgment of each stimulus is used in our analysis. Preliminary uncertainty $\hat{u}$ of a stimulus (equation 9) is based on rV_{99} (representing inability of judgment) and standard deviation of estimates, v_{std}, (characterising dispersion or ambiguity in judgment). Both can be understood as aspects of uncertainty and are added rather than multiplied in order to prevent either one (being 0) annihilating $\hat{u}$ irrespective of the value of the other. $\hat{u}$ was normalised onto [0, 1] using equation (10) (where $n_{max} = 1$, $n_{min} = 0$) to yield uncertainty u.

$$\hat{u} = \frac{rV_{99}}{\max\limits_{stimuli}(rV_{99})} + \frac{v_{std}}{\max\limits_{stimuli}(v_{std})} \quad \in [0, 2] \tag{9}$$

$$u = (\hat{u} - \min\limits_{stimuli}(\hat{u})) \cdot \frac{n_{max} - n_{min}}{\max\limits_{stimuli}(\hat{u}) - \min\limits_{stimuli}(\hat{u})} + n_{min} \tag{10}$$

For comparing estimates and algorithmic valleyness measures reduced major axis regression (RMA) in the *R* package *lmodel2* (CRAN 2009) was used. RMA (or Model II) regression was used rather than ordinary least squares (Model I), since in the case at hand there is no clear definition of predictand and predictor, the latter of which is assumed to be error free in ordinary regression (Mark and Church 1977).
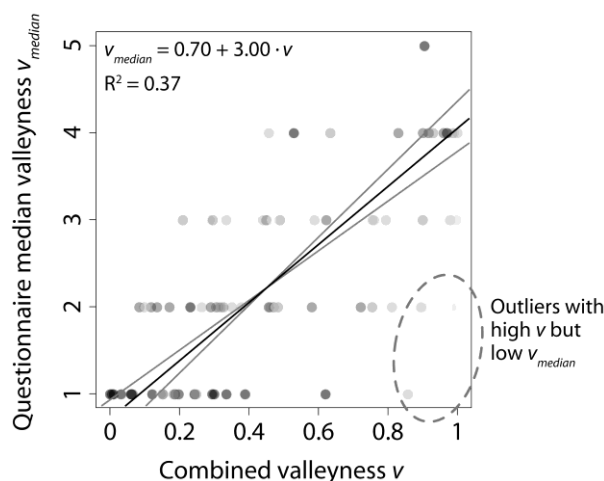
Figure 11.   RMA regression between algorithmically derived valleyness and v_{median} ($n = 100$). Grey lines represent the 95% confidence intervals for the regression's gradient. The dots representing the stimuli are coloured according to the uncertainty of the participants' estimates (five equal intervals); dark: low uncertainty, light: high uncertainty.

Table 2.   RMA regressions involving v_{median}.

Regression against ...	Coefficient	Intercept	R^2
v_e (elevation-based)	3.23	.52	.35
v_c (convexity-based)	2.64	.93	.35
v (combined)	3.00	.70	.37

As can be seen from Fig. 11, there is considerable scatter in the data (also illustrated by the relatively low coefficients of determination, R^2, in Table 2). However, there is a trend pattern in the scatterplots which shows that the regression though far from very clear is certainly substantiated. As a tendency, the less certainly judged stimuli are often found near the fringes of the point cloud. However, there are also some notable exceptions to this.

RMA regression coefficients cannot be tested for statistical significance – but the positions of the 95% confidence intervals for the regression gradient clearly suggest that the gradient is non-zero (cf. Vittinghoff et al. 2005: 42). Legendre and Legendre (1998: 511), however, suggest that in RMA regression the confidence intervals may not be informative either and instead propose to test the correlation coefficient, $R,$ according to McArdle (1988). In fact, all correlations in Table 2 are significant at the 1% confidence level (irrespective of using (para-

metric) Pearson correlation or (non-parametric) Kendall's Tau and Spearman's Rho).

Fig. 12. Photographs which were marked as potential plains (images in the questionnaire were in colour).

In the exploratory analysis of the questionnaire data it was noted that participants' responses showed marked diversity for some images of vast low places for example near large perialpine lakes in Switzerland (Fig. 12). It was hypothesised that while some participants confronted with these stimuli rated the photographer's location as illustrated by the photograph as 'being in a valley' others may have felt that a location was too broad and flat to be a valley. Yet other participants were perhaps unsure because they could not 'see' what lay behind

them; thus, since only one slope was visible, they could not distinguish between an escarpment and a valley. In other words, it was suspected that some stimuli did not confront participants with a dichotomy of *valley – non-valley* (or, for example, *valley – mountain*) – as had been intended in designing the question-naire – but with a trichotomy *"lower and flatter than a valley"* – *valley* – *"higher and less flat than a valley"* (or *plain – valley – mountain*). We also received two feedback e-mails from questionnaire participants hinting at this difficulty.

   Hence, an indicator regarding the presence of a potential plain was introduced into the analysis (Fig. 12). The location associated with most of these stimuli was considered very valley-like by the algorithm, but less so by the participants. 15 out of the 17 affected stimuli were to be found in the lower right half of the graph and below the regression line (Fig. 13). The dashed outline in Fig. 11 marks an area of notable regression outliers and Fig. 13 shows that indeed many of these are due to the above circumstance.
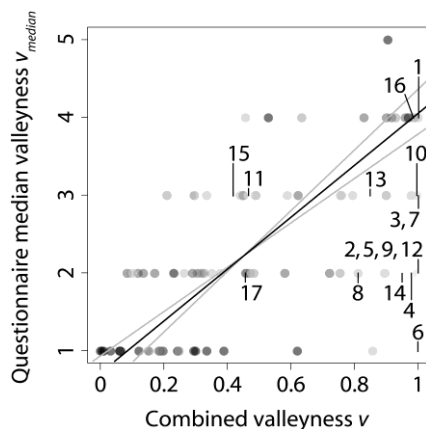

Figure 13.   RMA regression between algorithmically derived valleyness and v_{median} ($n = 100$) with indication of stimuli with *presence of a potential plain*.

Kruskal-Wallis H test found no statistically significant differences in v_{median} when stratified according to the indicator variable *presence of a potential plain* ($p = 0.159$). Nonetheless, the RMA regressions were also carried out using only the stimuli which were regarded as unaffected by the *presence of a potential plain*. Limiting the sample to only such stimuli left a dataset of 83 records. The respective regressions can be seen in Fig. 14. In all depictions the clutter below the regression line at the high end of algorithmically derived valleyness has been considerably lessened in comparison to Fig. 13 or Fig. 11. The removing of potential plains out of the dataset has increased all regressions' gradients and improved all models' fits expressed by R^2 markedly (Table 3).
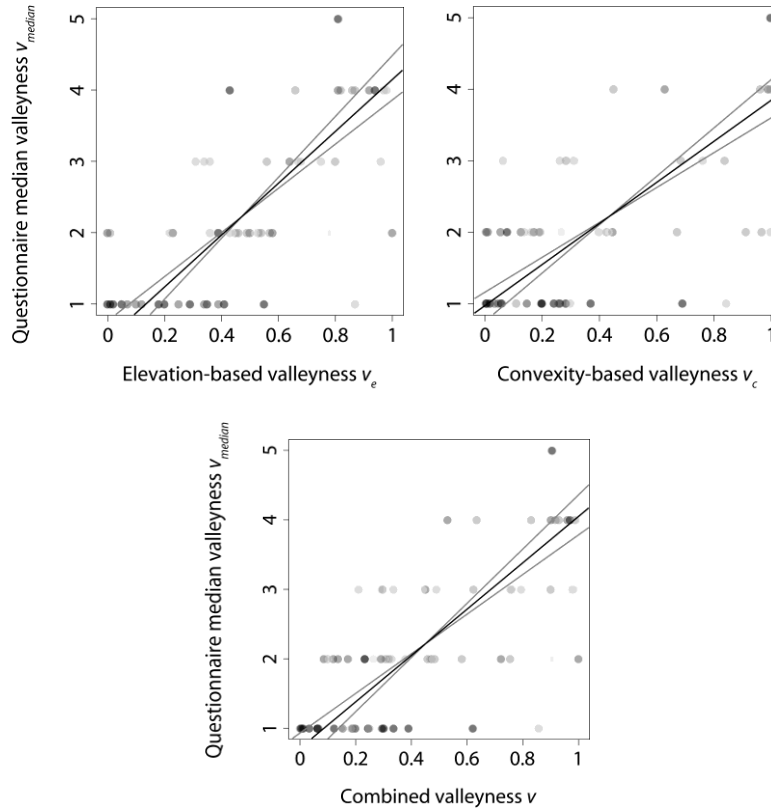
Figure 14.   RMA regressions between algorithmically derived valleyness measures - v_e, v_c and v - and v_{median} ($n = 83$), where stimuli marked as potential plains where excluded from the regression and the display.

The colour-coding of the scatterplots in Fig. 14 suggests that the data points which are less affected by uncertainty tend to be less well represented by the regression line. Thus, the weighted coefficient of determination, R_w^2, has been computed according to Bills and Li (2005: 838) and Greenacre (2007: 229) using inverted uncertainty as weights. Values of R_w^2 vary from 0.50 to 0.55 (Table 3).

Table 3. RMA regressions involving v_{median}.

Regression against ...	Coefficient	Intercept	R^2	R_w^2
v_e (elevation-based)	3.63	.51	.47	.52
v_c (convexity-based)	2.87	.97	.45	.50
v (combined)	3.34	.71	.49	.55

Gravetter and Wallnau in their behavioural sciences statistics textbook (2004: 295) refer to a heuristic by Cohen (1988) which says R^2 values above 0.25 signify large effects. This view is supported by Coolidge (2006: 155). Milton and Arnold (1995: 433) emphasise that the interpretation is subject matter dependent: "An R^2 value of 50% might be considered very large in a social science setting where human subjects are involved; however, the same figure could be considered very small in a designed engineering experiment." Considering the subject and the experimental design involving participants we thus consider the coefficients of determination in Table 3 satisfying.

5. Reflections

Based on definitions and characteristics of *valley floors* and *valleys*, this paper has introduced algorithms to delineate and fuzzily characterise these landform features from a relatively coarse DEM. The evaluation of these algorithms was not performed using subjective judgment or some field or air photo-based expert delineation but employed data gained in a questionnaire experiment from a very broad and large audience. Statistical tests could not find significant differences between different levels of geographic expertise in these people; thus the comparison to the algorithmic valleyness could be done in a lumped manner. While the initial comparison already yielded significant correlations, the amount of explained variance was moderate (R^2 of 0.35 to 0.37). Accounting for very ambiguously perceived stimuli showing vast low places led to markedly improved regressions (R^2 of 0.45 to 0.49), weighting of the data with a measure of uncertainty in judgment even more so (R^2 of 0.50 to 0.55). The evaluation thus showed that the proposed measures of valleyness can explain a substantial amount of variation in the human perception and appreciation of the landform at hand.

We are convinced that, when a landform characterisation is targeted at a human audience (rather than being used in e.g. soil science) this is a valid benchmark, and suggest that, despite the obvious labouriousness of our approach future studies of landform characterisation should consider experiments involving humans as a central part of the research.

As an improvement to the procedure presented here, we suggest that photographs are collected specifically for the purpose. If azimuth, tilt angle and focal length were recorded along with stimuli locations, the image contents rather than the photographer location could be assessed in the questionnaire survey.

Acknowledgements

Mark and Jo Wood for interesting ideas which they contributed to the research described in this paper.

References

Ahnert, F. (1998). *Introduction to Geomorphology*. London: Arnold.

Allaby, A. & Allaby, M. (1999). A Dictionary of Earth Sciences. Oxford Reference Online. Oxford University Press. http://www.oxfordreference.com/views/BOOK_SEARCH.html?book=t13.

ANSI (1998). American National Standards Institute (ANSI) NCITS 320-1998. Draft. Part 2: Spatial Features. http://thor-f5. er.usgs.gov/sdts/standard/latest_draft/pdf/part2.pdf.

Bates, R. L. & Jackson, J. A. (Eds.) (1990). *Glossary of Geology*. Alexandria: American Geological Institute.

Bills, B. C. & Li, G. (2005). Correlating homicide and suicide. *International Journal of Epidemiology, 34*(4), 837–845.

Bolongaro-Crevenna, A., Torres-Rodríguez, V., Sorani, V., Frame, D. & Ortiz, M. A. (2005). Geomorphometric analysis for characterizing landforms in Morelos State, Mexico. *Geomorphology, 67*, 407–422.

Chaudhry, O. Z. & Mackaness, W. A. (2008). Creating mountains out of mole hills: Automatic identification of hills and ranges using morphometric analysis. *Transactions in GIS, 12*(5), 567–589.

Coolidge, F. L. (2006). *Statistics. A Gentle Introduction.* Thousand Oaks: Sage.

CRAN (2009). The Comprehensive R Archive Network, http://cran.r-project.org.

Demoulin, A., Bovy, B., Rixhon, G. & Cornet, Y. (2007). An automated method to extract fluvial terraces from digital elevation models: The Vesdre valley, a case study in eastern Belgium. *Geomorphology, 91*(1–2), 51–64.

Deng, Y. (2007). New trends in digital terrain analysis: Landform definition, representation and classification. *Progress in Physical Geography, 31*(4), 405–419.

Deng, Y. & Wilson, J. P. (2007). Multi-scale and multi-criteria mapping of mountain peaks as fuzzy entities. *International Journal of Geographical Information Science, 22*(2), 205–218.

Dikau, R. (1989). The application of a digital relief model to landform analysis in geomorphology. In J. Raper (Ed.), *Three Dimensional Applications in Geographical Information Systems* (pp. 51–77). London: Taylor & Francis.

ESRI (2010): ArcGIS Desktop Help 9.3: How Flow Direction Works. http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?topicname=how_flow_direction_works.

Fisher, P., Wood, J. & Cheng, T. (2004). Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. *Transactions of the Institute of British Geographers, 29*, 106–128.

Gallant, J. C. & Dowling, T. I. (2003). A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research, 39*(12), 1347.

Geomorphometry (2009). Geomorphometry.org. The science of digital terrain analysis. http://geomorphometry.org.

Gravetter, F. J. & Wallnau, L. B. (2004). *Statistics for the Behavioral Sciences*. Belmont: Wadsworth.

Greenacre, M. J. (2007). *Interdisciplinary Statistics. Correspondence Analysis in Practice*. Boca Raton: Chapman & Hall / CRC.

Hammond, E. H. (1954). Small-scale continental landform maps. *Annals of the Association of American Geographers, 44*(1), 33–42.

Huggett, R. J. (2007). *Fundamentals of Geomorphology*. London: Routledge.

Irvin, B. J., Ventura, S. J. & Slater, B. K. (1997). Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. *Geoderma, 77*(2–4), 137–154.

Jarvis, A., Reuter, H. I., Nelson, A. & Guevara, E. (2006). Hole-filled seamless SRTM data v3. International Centre for Tropical Agriculture (CIAT). http://srtm.csi.cgiar.org.

Jenson, S. K. & Domingue, J. O. (1988): Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing, 54*(11), 1593–1600.

Lapidus, D. F., Coates, D. R., Winstanley, I., MacDonald, J. & Burton, C. (2003). *Collins Dictionary of Geology*. Glasgow: HarperCollins Publishers.

Legendre, P. & Legendre, L. (1998). *Numerical Ecology*. Developments in Environmental Modelling 20, Amsterdam: Elsevier.

MacMillan, R. A., Pettapiece, W. W., Nolan, S. & Goddard, T. W. (2000). A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Sets and Systems, 113*, 81–109.

Mark, D. M. & Church, M. (1977). On the misuse of regression in earth science. *Mathematical Geology, 9*(1), 63–75.

Mark, D. M. & Turk, A. G. (2003): Landscape categories in Yindjibarndi: Ontology, environment, and language. In G. Goos, J. Hartmanis & J. van Leeuwen (Eds.), *Lecture Notes in Computer Science, 2825* (pp. 28–45. Berlin: Springer.

Mark, D. M., Turk A. G. & Stea, D. (2007): Progress on Yindjibarndi ethnophysiography. In S. Winter, M. Duckham, L. Kulik & B. Kuipers (Eds.), *Lecture Notes in Computer Science, 4736* (pp. 1-19). Berlin: Springer.

Mark, D. M. (2009): From land form to landforms: Bridging the quantitative-qualitative gap in a multilingual context. In R. S. Purves, S. Gruber, T. Hengl & R. K. Straumann (Eds.), *Proceedings of Geomorphometry 2009* (pp. 13–16). Zurich: Department of Geography, University of Zurich.

Maxwell, J. C. (1870). On Hills and Dales. *Philosophical Magazine, 2*(30). Reprinted in W. D. Niven (Ed.), *The Scientific Papers of James Clerk Maxwell* (pp. 233–240). New York: Dover Publications.

Mayhew, S. (2004). A Dictionary of Geography. Oxford Reference Online. Oxford University Press. http://www.oxfordreference.com/views/BOOK_SEARCH.html?book=t15.

McArdle, B. H. (1988). The structural relationship: Regression in biology. *Canadian Journal of Zoology, 66*(11), 2329–2339.

Miliaresis, G. C. & Argialas, D. P. (1999). Segmentation of physiographic features from the global digital elevation model/GTOPO30. *Computers & Geosciences, 25*(7), 715–728.

Milton, J. S. & Arnold, J. C (Eds.) (1995). *Introduction to Probability and Statistics. Principles and Applications for Engineering and the Computing Sciences.* New York: McGraw-Hill.

Montello, D. R. & Sutton, P. C. (2006). *An Introduction to Scientific Research Methods in Geography*. Thousand Oaks: Sage.

Moore, I. D., Grayson, R. B. & Ladson, A. R. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes, 5*, 3–30.

OED (2009). Oxford English Dictionary. The definitive record of the English language, http://dictionary.oed.com.

Pennock, D. J. (2003). Terrain attributes, landform segmentation, and soil redistribution. *Soil & Tillage Research, 69*(1–2), 15–26.

Pennock, D. J. & Corre, M. D. (2001). Development and application of landform segmentation procedures. *Soil & Tillage Research, 58*(3–4), 151–162.

Pennock, D. J., Zebarth, B. J. & De Jong, E. (1987). Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. *Geoderma, 40*, 297–315.

Ruhe, R. V. (1960). Elements of the soil landscape. In *Transactions of the 7th International Congress of Soil Science* (pp. 165–170), Madison: International Society of Soil Science.

Schmidt, J. & Hewitt, A. (2004). Fuzzy land element classification from DTMs based on geometry and terrain position. *Geoderma, 121*(3–4), 243–256.

Smith, B. & Mark, D. M. (2003). Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and Design*, *30*(3), 411–427.

Straumann, R. K. (2009). Experiences in developing landform ontologies. In R. S. Purves, S. Gruber, T. Hengl & R. K. Straumann (Eds.), *Proceedings of Geomorphometry 2009* (pp. 17–21). Zurich: Department of Geography, University of Zurich.

Straumann, R. K. & Purves, R. S. (2008). Delineation of valleys and valley floors. In T. J. Cova, H. J. Miller, K. Beard, A. U. Frank & M. F. Goodchild

(Eds.), *Lecture Notes in Computer Science, 5266* (pp. 320–336). Berlin: Springer.

Tribe, A. (1991). Automated recognition of valley heads from digital elevation models. *Earth Surface Processes and Landforms, 16*(1), 33–49.

Tribe, A. (1992). Problems in automated recognition of valley features from digital elevation models and a new method toward their resolution. *Earth Surface Processes and Landforms, 17*(5), 437–454.

Trochim, W. M. (2006). The Research Methods Knowledge Base: Likert Scaling. Web Center of Social Research Methods. http://www.socialresearch methods.net/kb/scallik.php.

Troeh, F. R. (1965). Landform equations fitted to contour maps. *American Journal of Science, 263*, 616–627.

Vittinghoff, E., Glidden, D. V., Shiboski, S. C. & McCulloch, C. E. (2005). *Regression Methods in Biostatistics. Linear, Logistic, Survival, and Repeated Measures Models.* New York: Springer.

Wang, D. & Laffan, S.W. (2009). Characterisation of valleys from DEMs. In R. S. Anderssen, R. D. Braddock & L. T. H. Newham (Eds.), *Proceedings of 18th World IMACS/MODSIM Congress* (pp. 2014–2020). Cairns: IMACS, MSSANZ.

Wang, D., Laffan, S.W., Liu, Y. & Wu, L. (in press). Morphometric characterisation of landform from DEMs. *International Journal of Geographical Information Science*, doi:10.1080/13658810802467969.

Whittow, J. B. (2000). *The Penguin Dictionary of Physical Geography*. London: Penguin Books.

Wilson, J. P. & Gallant, J. C. (2000). Digital Terrain Analysis. In J. P. Wilson & J. C. Gallant (Eds.), *Terrain Analysis* (pp. 1–27). New York: John Wiley & Sons.

Wood, J. D. (1996). The geomorphological characterisation of digital elevation models. PhD thesis, University of Leicester, UK.