DOWN THE RABBIT HOLE: IDENTIFYING PHYSICAL CAUSES OF SINKHOLE FORMATION IN THE UK

Tamsin Brittany Green
School of Geography, University of Leeds, University Road, Leeds, West Yorkshire, LS2 9JZ, gy12tg@leeds.ac.uk

Abstract
Heavy precipitation in the UK in February 2014 induced ground subsidence and consequently a rapid increase in the frequency of sinkhole occurrences. These new sinkhole collapses emphasize the need to further analyze the causes of the increased occurrence by investigating the relative importance of various surficial factors.

Malham and the Mendips are two areas of particular interest, since both are underlain by limestone bedrock and are susceptible to subsidence. This is due to limestone being primarily permeable in joints, and so it dissolves to form an extensive network of karstic caves. It was therefore useful to compare two sites of similar geology, both from the Triassic and Jurassic periods, as this controlled the amount of presently exposed limestone from past glacial retreat, for accurate comparison of susceptibility.

Susceptibility maps of the two areas were created by integrating GIS application and statistical methods to develop algorithms to address the issue of dissolution. The maps aim to identify the physical surficial conditions, in addition to heavy precipitation that exacerbates subsidence development.

Statistical testing of the GIS data indicated that in Malham, slope is the most significant parameter (Kruskal-Wallis, H=29.36, p<0.001; H=14.55, p=0.006, respectively) in sinkhole formation; while in the Mendips altitude is the most significant parameter (Kruskal-Wallis, H= 20.44, p<0.001; H= 86.51, p<0.001, respectively). Curvature appeared less statistically significant with fewer values reported from post-hoc Mann-Whitney U tests. This integrated geological mapping and statistical approach will prove useful in delineating susceptibility zones in areas within the UK.

Introduction
This study investigates the physical and surficial causes of sinkhole formation in both Malham and the Mendips, in the UK. These are two areas underlain predominantly by limestone bedrock, and are highly prone to dissolution, due to the highly soluble nature of the limestone (Waltham et al., 1997). This study explores how surficial features exacerbate dissolution, and aims to demonstrate that bedrock characteristics are not the most important factors. In addition, the aim is to assess the relative importance of the surficial features. Base map data and geological data were obtained from EDINA DigiMap and the British Geological Survey respectively (BGS License number 2014/143 ED British Geological Survey© NERC. All rights reserved. Edina DigiMap© Crown Copyright/database rights 2015. An Ordnance Survey/ (Datacentre) supplied service). This data, together with surface feature data from ArcGIS, enabled the creation of zonation maps with high, considerate, moderate and low susceptibility areas.

Within this report, predictions are estimated on a spatial scale, following the sensitivity associated with temporal prediction. Therefore, chronological data is not accounted for, and the maps only present visual future susceptibility on a two-dimensional level.

This report analyzes the role that surficial factors of slope, curvature, and altitude within the two areas play in exacerbating subsidence, and whether one factor in particular may be more critical. This analysis assesses the relative importance of each variable, to ultimately create a reliable spatial sinkhole susceptibility map.

Due to the topical, public, and media interest in sinkhole collapse, this research is of significant importance in today’s environment, economy, and society. Sinkholes affect 15% of the world’s surface today (Wilson and Beck, 1992). Most relevant papers date back to the early 1900s (Elrod, 1898; Vineyard and Williams, 1967; Purdue, 1907). These predominantly focused on pre-existing cavities, where the limestone cave systems that were once mining sites, initiated subsidence. Recent literature now focuses more upon the range of external factors that exert pressure on these vulnerable locations, due to the increasing availability of modern technological equipment, allowing more in-depth analysis (Sass, 2007; Cooper, 2008; Stecchi et al., 2009).
This paper aims to narrow down the importance of particular surficial factors. Based on a range of literature (Waltham, 2008; Parise et al., 2009; Parise, 2010) it is clear that this hazardous phenomenon has the ability to destroy lives and local communities. The creation of any susceptibility zonation maps based on the current ambiguity of such existing surficial causes will provide insight into avoiding a potentially unsafe environment. Though many variables involving subsidence formation have been previously investigated in research, no definitive answers have been concluded following the arbitrary nature of sinkholes (Upchurch and Littlefield, 1988; Florea et al., 2002) and the fairly novel area of sinkhole research. This paper determines the importance of each factor, rather than concluding the generic causation of multiple factors. Furthermore, many studies focus on evaporite karst areas (Johnson, 1997; Cooper, 2008; Gutiérrez et al., 2008; Galve et al., 2009) due to its higher susceptibility to dissolution, though carbonate karst is more common (Gutiérrez et al., 2008).

This study focuses on limestone carbonate karst areas from the Triassic and Jurassic geological time periods; the two areas have similar geology, with only surficial differences for susceptibility mapping. Though it can be hard to justify the specific causes following such apparent spatial dichotomies even within the UK, it was necessary to focus on a local scale, in order to identify detailed causes, rather than wider, regional causes.

This study identifies gaps in literature by creating a susceptibility map comparing two areas on a local scale, using purely surficial and physical factors in order to obtain as much detail and understanding as possible. Though zonation maps have been previously created, they have been predominantly single-site based and scale-specific (Kaufmann and Quinif, 2002; Stecchi et al., 2009). This paper aims to further this research by creating a comparative map of two areas based upon multiple surficial factors.

This study is meant to be useful in mapping the safety zonation of areas for future building (Gutiérrez et al., 2008), and therefore aims to implement a preventative measure, and create local awareness of specific conditions that may aggravate subsidence (Farrant and Cooper, 2008).

**Study Site**

Two 25km² areas consisting of pre-existing doline points were extracted from the BGS GIS database for each study site; East of Settle around the Craven District in Malham, and North-west Mendip Hills (Figures 5 and 6). Larger sized areas are also used however, to assess future susceptibility based on the slope, aspect and curvature of the surrounding areas, determined by the values as grouped in Table 1. These sites were selected based on the predominant presence of limestone bedrock defining these two areas, in addition to known sinkhole activity.

**Methods of Study**

This research relied purely on secondary data obtained from the British Geological Survey, in order to analyze the physical formation, and spatial distribution of sinkholes. A 1:50,000-resolution, 50m grid cell size digital elevation model and BGS doline data for each area was imported into ArcGIS. The doline points provided extensive information based on count, type, shape and distribution of the points. Distance between the points was calculated using ‘point cluster analysis’, and each point was then corresponded to its bedrock class that it was underlain by. The ‘identity’ tool further enabled the partnering of each point with its related topographical slope, curvature and altitude values, and enabled the integration of statistical testing and GIS. Curvature can be defined as the degree to which a surface is curved, and can be strongly linked with trends of faults and topographical fractures (Stecchi et al., 2007). The curvature is the second derivative of the elevation surface, which was run on a 3x3 cell scale determined by the DEM grid size resolution. The layer was filtered

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Flat slope &lt;3°</td>
</tr>
<tr>
<td></td>
<td>Linear curvature ≈ 0</td>
</tr>
<tr>
<td></td>
<td>On limestone</td>
</tr>
<tr>
<td></td>
<td>High altitude</td>
</tr>
<tr>
<td>Considerate</td>
<td>Gently sloping &lt;6°</td>
</tr>
<tr>
<td></td>
<td>Near linear curvature ≈-1 to 1</td>
</tr>
<tr>
<td></td>
<td>Mudstone</td>
</tr>
<tr>
<td></td>
<td>Intermediate altitude</td>
</tr>
<tr>
<td>Moderate</td>
<td>Slopes &lt;9°</td>
</tr>
<tr>
<td></td>
<td>Linear curvature ≈-2.5 to 2.5</td>
</tr>
<tr>
<td></td>
<td>Siltstone and interbedded rocks</td>
</tr>
<tr>
<td></td>
<td>Intermediate altitude</td>
</tr>
<tr>
<td>Low</td>
<td>Steep slope &gt;9°</td>
</tr>
<tr>
<td></td>
<td>Extremely convex or concave curvature ≈-5 to 5</td>
</tr>
<tr>
<td></td>
<td>Non-porous bedrock e.g. sandstone</td>
</tr>
<tr>
<td></td>
<td>Low altitude</td>
</tr>
</tbody>
</table>
to remove any minor topographical hollows or peaks (Sullivan et al., 2007; Stecchi et al., 2009). Each surficial factor was reclassified into low-high susceptibility groups based on natural breaks, which was used to highlight the values where points naturally clustered. The susceptibility map was consequently created through using the ‘raster calculator’ tool, multiplying each surficial layer together. The output was reclassified into four levels of susceptibility.

Results of Study
Normality tests were executed on each variable; every variable returned as not normal, so non-parametric tests were performed throughout. All Mann-Whitney U tests were carried out with 95% confidence. It is clear from Figures 1 and 2 that there is a higher frequency of sinkholes on flatter slopes than steep slopes, and on linear curvatures than concave or convex. There is less of a correlation present based on altitudinal values. In Malham, the highest frequency of 64 sinkholes occurs at 510m above sea level, though this is not the highest elevation. The second highest peak consists of 52 sinkholes at 380m above sea level. In the Mendips, the frequency also varies, with a peak count of 39 sinkholes at 290m above sea level.

Malham bedrock clustering did however return with a statistical significance (Kruskal-Wallis, $H=23.82$, $p<0.001$) (Figure 3). Bedrock clustering in the Mendips returned with no statistical difference across different bedrock classes (Kruskal-Wallis, $H=8.39$, $p=0.078$) (Figure 4). Slope, altitude and curvature each presented a difference with Kruskal-Wallis testing and a further post-hoc Mann-Whitney U for individual bedrock class pairing. The results were more varied for Malham, with less of a pattern presented; limestone consistently

![Figure 1. Frequency bar charts in Malham (left) presenting the number of sinkholes (total: 400) on each driver of slope, curvature and altitude respectively.](image1)

![Figure 2. Frequency bar charts in the Mendips (right) presenting the number of sinkholes (total: 161) on each driver of slope, curvature and altitude respectively.](image2)
presented a difference. In the Mendips, conglomerate appeared to show the strongest difference between each other class within all tests for the surficial factors. Chert also showed significance when tested for altitudinal difference (Figure 4).

Sandstone presents the widest range of curvature values, ranging from -0.8 to 0.5 1/100 z-units. In contrast, limestone presents the smallest mean range, with 121 sinkholes at 0 1/100 z-units, though with multiple outliers on extreme curvature values, accounting for the large number of sinkholes apparent on limestone bedrock. More outliers are evident in Malham (Figure 3) than the Mendips (Figure 4) due to the wider range of sinkholes that exist outside the predicted common thresholds.

The map above (Figure 5) was created using the raster calculator by combining slope, curvature and altitude. The high-susceptibility values were defined as <5.7° for slope, >413m above sea level for altitude and -0.34 to 0.12 1/100 z-units for curvature; this was based on reclassified natural data breaks with sinkhole frequency. Each reclassified layer was input into the raster calculator where they were combined to provide an output layer. This was reclassified again into the four classes.

**Figure 3.** Malham doline point clustering on each bedrock class presented by boxplots on each surficial factor: slope, curvature and altitude respectively.

**Figure 4.** The Mendips doline point clustering on each bedrock class presented by boxplots on each surficial factor: slope, curvature and altitude respectively.
The high-susceptibility values for slope were defined as <3.7° for slope, -0.37 to 0.16 1/100 z-units for curvature and >176m above sea level for altitude. The same method as stated above was used here also.

Discussion

Susceptibility Maps

The models above (Figures 5 and 6) present high to low susceptibility of sinkhole formation in Malham and the Mendips. It is encouraging that the high frequency of pre-
demonstrating that such local scale features clearly exacerbate sinkhole existence. However it is also clear that some sinkhole points occur on the considerate susceptibility areas. This is because the threshold of the slope, curvature and altitude values do not define the exact points that sinkholes can only occur on, but instead provide an indication of the most vulnerable existing sinkholes is mapped on the highly susceptible red areas, whilst the orange areas have few sinkhole densities, and green and yellow areas have few to none. Based on pre-existing locations of current sinkholes and local topographical features of slope, altitude and curvature, the different areas demonstrate potential wider locations of future sinkhole development. Thus,

**Figure 6.** Spatial susceptibility model for the Mendips based on the values of the surficial factors, and pre-existing doline points.
areas, though considerate to low areas still need to be considered regardless. The threshold values also largely vary on a spatial scale, even within two similar geological areas within the UK, and so these varying boundaries need to be site-specific. Results statistically proved local factors of flat slopes, higher altitudes and linear curvatures to be more susceptible to sinkhole formation, with the exacerbation of precipitation. Though it was expected that concave curvatures would be highly susceptible, my statistical results found linear curvatures to be more so. However, difficulty does arise with curvature analysis through the necessary filtering. These results were also consistent with findings from Farrant and Cooper (2008), Simms and Ruffell, (1989) and Sánchez et al. (2007). The basic theory that these local factors exacerbate subsidence can however provide an approximate and valuable insight into potentially vulnerable locations, with the suitable underground conditions.

Furthermore, in standardising approximations, the high-susceptibility percentage of the total slope and altitudinal values can be quantified. The areas combined present susceptible slopes to occur at 9-11% of the total slope, whilst highly vulnerable elevations occur at 46-64% of the total range.

**Slope**

A difference in sinkhole frequency was clearly evident, with flat slopes containing more sinkholes than steep slopes. This finding was expected. Farrant and Cooper (2008) suggested that dissolution pipes and irregular rockhead form on flat slopes, ideal for karst formation, whereas steep slopes promote erosion. This finding is additionally supported in the context of landslides in Glade (2005) and Cooper (2008), where steep slopes induce ground instability and consequential rockfall, thus also causing erosion indirectly. It is also interesting to note that Farrant and Cooper (2008) claim slope to be an irrelevant factor when considering evaporite subsidence on gypsum and salt bedrock, as the karstic rock is rarely exposed to the surface.

In contrast, Santo et al. (2007) highlights the critical relationship between carbonate karst and local slope stability. This is an issue where widespread presence of carbonate karst and a high availability of dissolution to the slopes through hydrogeology induce subsidence. This is therefore a relevant study, as sinkhole formation can be directly correlated with slope where instability occurs in carbonate karst. Although it is evident gentler slopes are more susceptible than steeper slopes, perhaps the characteristics relating to carbonate karst slopes exacerbate this, in addition to the physical slope angle.

Though my results demonstrated gentle slopes to be more prone to sinkholes, Stecchi et al. (2009) reported building destruction initiated by ground subsidence on these “gentle” slopes. It is interesting to define the thresholds of “gentle” and “steep” slopes; from this study, it is clear that slope boundaries are highly subjective, even on the small-scale, local analysis that this study is built on. For example, sinkhole frequency is high on slopes ranging from 0-6° in Malham, but only 0-2° in the Mendips, thus the “gentle” boundary could not be equally applied. Furthermore, Glade (2005) presents slopes >2° as the most active due to high erosion and weathering processes, though this can only be site-specific to his study.

A critical and potentially useful theory investigated by Sass (2007) looks into the surface depth to bedrock measurements on slopes. This is interesting as although bedrock depth is not considered in this report, it could explain the presence of certain bedrock types in particular locations; if a steep slope is heavily eroded, it would reveal deeper underlying bedrock layers, than a flat slope comprised of the original first layer bedrock, that is a target for further deposited material. This could therefore explain in the Mendips for example, why bedrock such as sandstone, a predominantly low permeability rock (Ward and Morrow, 1987), is more prominent on steep slopes; mudstone and siltstone, a more resistance clayey material, (Franklin and Chandra, 1972) is more prominent on flat slopes, if sandstone was originally at deeper depths than other bedrock.

Sloecechi et al. (2009) also present slope to only be connected with ground movements, and not topography, however multiple papers (Doctor and Young, 2013; Rahimi and Alexander, 2013) note that the visual surface depressions and hollows indicate subsidence and sinkhole development.

**Curvature**

It was anticipated that sinkhole frequency would be higher on concave curvatures than linear or convex curvatures, due to the heavy pooling of precipitation
It is interesting to note the disparities among theories involving the promotion and prohibition of sinkhole development. UWSP (n.d.) explains how higher altitudes promote stronger weathering processes, and therefore expose the karstic bedrock further. In addition to this, it is also noted by Santo et al. (2007) that ice predominantly forms in mountainous areas of low temperature; this can be applicable to the Mendips where average altitude is approximately 700m above sea level, far higher however than Malham. In addition, the Mendips are also comprised of anticlines and periclines, both, which promote accelerated erosion, thus further exposing bedrock below (BGSb, n.d.). The conflicting theory adheres to the concept that exposed karst limits development (Simms and Ruffell, 1989). Research presented a strong variance in the frequency of sinkholes at a range of altitudes, with no distinct pattern, and so both theories are seen to be noteworthy. Whilst it is more strongly believed that erosive processes limit sinkhole development, rather than enrich it, exposed limestone karst is generally rare (Beck, 1986). However, Beck (1986) does note that exposed limestone initiates rapid recharge for sinkholes, so perhaps it is this indirect correlation that can cause a certain type of sinkhole; namely solution sinkholes (UWSP, n.d.). This therefore explains the variety of results returned from ArcMap data extraction. There is a clear local dependence on altitude for sinkhole formation, likely resulting from multiple changes in soluble rock such as land surface cutting and local precipitation. However, the underlying cause for altitude acting as a driver of spatial distribution, is due to sinkholes developing in soluble rock and salt far below sea level, up to high elevations in rock permafrost.

Malham’s range of concave curvature values, in contrast to the Mendips could therefore present a general sinking in the surface; Stecchi et al. (2009) report the negative values to be indicative of sinking.

**Altitude**

Although some high altitudinal values correlated with high sinkhole frequency, there was a wide range of variance in the data. Whilst a pattern is less notable in the Malham (Figure 1), this is most likely due to the wider range of data in the Mendips, and overall higher number of sinkholes present in Malham (Figure 1), so the spread is wider. In the Mendips, the distinct drop in frequency at its highest altitude is questionable; this is still lower than Malham’s lowest altitude, which highlights issues of changeability across spatial scales.

It is further notable link with curvature, are fault lines and fracturing (Murray, 1968; Vendeville, 1991; Tharp, 2002). This is due to the ability of curvature to predict the distribution of deformation (Bergbauer and Pollard, 2003), and its close relationship to geology. Vendeville (1991) points out the criticality of curvature analysis, in how it varies significantly with geology; this reason underlies the study’s choice to compare two sites of similar geology from the Jurassic/Triassic geological time periods.

**Bedrock Clustering**

Whilst some studies have previously identified sinkhole distributions through nearest neighbor analysis (Gao et al. 2005; Gutiérrez et al. 2008), this study analyzes the clustering of points on each bedrock class. It was anticipated that limestone, being the most prevalent bedrock to sinkhole formation, likely resulting from multiple changes in soluble rock such as land surface cutting and local precipitation. However, the underlying cause for altitude acting as a driver of spatial distribution, is due to sinkholes developing in soluble rock and salt far below sea level, up to high elevations in rock permafrost.
to all other bedrocks. However conglomerate consists of limestone fragments (BGSc, n.d.), and so can be said to hold similar characteristics to pure limestone, so the theory is not wholly disproved.

In Malham, slope appears to present a stronger statistical difference across each bedrock class than any other factor; altitude also portrays strong importance. Alternatively, curvature values across both sites appear to have the least importance in determining sinkhole formation, with the least values presenting a statistical significance. This could be explained by the ease of error created, and difficulty in analysing curvature, as noted in Bergbauer and Pollard (2003) and Sullivan et al. (2007).

In the Mendips, the most important driver of subsidence appears to be altitude, presenting the strongest statistical difference across bedrock. Though there is much conflict over whether high altitude influences sinkhole development or not (Simms and Ruffell, 1989; Santo et al., 2007; UWSP, n.d.), perhaps it is more directly the difference in altitude across the different bedrock classes that define development.

Overall, it can be said that slope and altitude have a mutual importance as drivers influencing subsidence, though each portrays different significance within each site. This is due to the change in driver importance across varying spatial gradients, highlighting the need to adapt to the change in site-specific drivers. Though even within the UK Malham and the Mendips present a difference in driver importance, confident conclusions can be drawn, based on the similarities.

**Conclusions**

Three conclusions can be drawn from this investigation:

1. Slope and curvature appear to be the most significant drivers influencing sinkhole formation; given the difficulty in curvature analysis however, slope is the most reliable and critical factor.

2. Spatial zonation maps are still valid even without the detail of temporal data. However, this research has presented the dichotomies present even across a small spatial gradient within the UK, which still creates some ambiguity in the surficial drivers.

3. Susceptibility mapping has proved useful for the future, though studies must be aware of the spatial inconsistencies present. This therefore calls for a wider-scale hazard susceptibility map encompassing larger areas and temporal data.

**References**


