Land subsidence is a result of either gradual or sudden movement of the Earth’s surface due to subsurface movement of the rock or strata. It also refers to the vertical downward movement of a mass (soil or rock) due to loss of underground strength [1]. There are many causes of subsidence, including water ingress; imbalance discharge due to water mining of an aquifer system; drainage of organic soil; underground mining (especially coal mining); consolidation of sedimentary deposits according to geological time scales; subsidence due to tectonic movements; localised collapse of subsurface cavities due to volume change or loss of underground support, and the collapse of the roof of an underground cave system. The collapse of land over natural underground voids can be a frequent phenomenon in underground cave systems. Sinkholes can appear on ground where the roof of an underground cave collapses. The collapse of cave roofs can be due to rock movement, loss of strength of underground soil, and dissolution of minerals with underground water [2]. However, the most common factor for erosion and voids’ formation in caves is limestone erosion. Subsidence can be of three types: funnel-shaped, bowl-shaped, and circular-shaped. In addition, the extent of land subsidence can vary from 1 m to as much as 30 m. Sterkfontein Caves are located in the Cradle of Humankind, which has significant archaeological value. It is undergoing a variety of rheological distresses. There is a need to study the natural occurring surface deformations which can result in the subsidence of land. The caves are frequently visited by the public, making risk assessment and hazard monitoring a requirement to ensure the safety of visitors.

Land-subsidence hazard is common in limestone cave formations such as at Sterkfontein Caves. This type of subsidence is associated with subcircular surface depression or structure collapse when small underground voids collapse [3].

Fig. 1: Extent of shape file for extraction of required data.

Fig. 2: Conceptual model for the geospatial subsidence hazard map of the Sterkfontein Cave area.
Land subsidence is broadly divided into two categories: human activity induced subsidence and geologically induced subsidence. Human-induced subsidence has been extensively investigated. Particularly, it has been found that extensive groundwater withdrawal can result in substantial subsidence and loss of property [4]. Gao and Alexander did subsidence hazard modelling and included rock geology and distance to the nearest sinkhole, but did not include topographical land use factors into their model [5]. Green et al studied Karst mapping using GIS and described hydrology as one of the factors of subsidence and sinkhole formation [6]. They also included bedrock, fractures and faults, and soil depth as major contributing factors. Rawashdeh et al developed a sinkhole probability map based on factors such as sinkhole distribution, bedrock geology and depth to bedrock based on the nearest neighbour analysis in a GIS system [7]. Gao and Alexander suggested that subsidence areas can have increasing clusters because of similar geological and topographic conditions [8].

Gao also designed and developed a GIS-based database for spatial analysis of Karst features and contributing factors [9]. Pirasteh and Rizvi studied the location and mapping of sinkholes on an aerial photograph using remote sensing techniques [10]. Hobbs studied the surface and ground water resources in the Cradle of Humankind [11]. He also discussed the effects of surface and ground water on the Karst system of Sterkfontein Caves. Oosthuizen and Richardson studied the mechanism of sinkhole formation in South Africa in areas underlain by dolomite rock [12]. Durand and Peinke studied the Karst ecology of the Sterkfontein Caves in South Africa [13].

Study area

The Sterkfontein Caves are located about 50 km north-west of Johannesburg within the Cradle of Humankind World Heritage Site. It is situated in the Gauteng province of South Africa and the site straddles the boundary of the Oaktree and Monte Christo formations. The strata dip about 30° to the NW. The passages follow two main directions of sub-vertical joints; the dominant one is mostly W to WNW and the subordinate one varies from N to NNE. The cave system is restricted to a volume of about 200 x 250 x 50 m [14]. Sterkfontein Caves consist of a labyrinth network of passages which are controlled by joints or broad flat chambers - the result of dissolution of chert-free beds sandwiched between cherty dolostone.

Methodology

Data and analysis concept
The data used in this study comprised remotely sensed imagery and related tabulated data. The 10 m spatial resolution SPOT 5 imagery was acquired from South African National Space Agency (SANSA). The imagery was generated on 10 April 2014 with a radiometric resolution of 8 bit. The second imagery is a 2.5 m resolution (re-sampled) aerial photograph produced by GeoTerraImage (GTI) in 2012. The spatial data were...
geo-referenced to conform to WGS 84 projection and then converted to GIS format for multi-criteria analysis.

**Data pre-processing**

Prior to the use of SPOT 5 and aerial photographs, both were enhanced and then geometrically corrected. Some visual screen digitisation was also done for features like roads and small water paths. As the area under investigation is only confined to the boundary of Sterkfontein Caves and its surroundings, it was necessary to extract the area of interest from the complete digital data. This extraction reduced the calculations required to be done by the GIS model and also saved time. Prior to extraction, a new shape file with the extent shown in Fig. 1 was created in ArcGIS environment. The shape file is in the polygon vector format and is in WGS 84 coordinate system.

**Image enhancement**

Image enhancement was done by traversing low pass 3 x 3 cell neighbourhood window filters over the raster image. A low-pass filter smooths the data by reducing local variation and removing noise. It calculates the average value for each 3 x 3 neighbourhood. The effect is that the high and low values within each neighbourhood would average out, reducing the extreme values in the data. The image enhancement technique helps to produce a better-represented scene for clear visual presentation and also helps clear information content for the interpreter.

**Digital elevation model (DEM)**

The DEM used for the spatial analysis was the 30 m global digital elevation model. 20 m spatial resolution DEM produced by GeoTerraImage (GTI) was also used for calculating the slopes.

**Geospatial subsidence modelling design**

The first step in the modelling was to create a conceptual model. The factors responsible for land subsidence were carefully selected so that they corresponded to the area of interest. The conceptual model is given in Fig. 2. The preparation of subsidence maps for the Sterkfontein Caves involved digital data input. Three types of data were used. The land cover map was extracted from the 2.5 m aerial photograph which was further classified into 15 land-use classes. The land-use classifications are given in Table 1.

The existing subsidence areas were obtained from the 10 m spatial resolution SPOT 5 imagery. A shape file was created in WGS 84 coordinate system to mark the present subsidence areas as points. These points were also confirmed through ground observations. Water accumulation areas and slopes were calculated from the digital elevation model (DEM). The spatial analyst slope tool was used to extract the slopes. The water accumulation areas were calculated by the Flow Accumulation hydrological tool in ArcGIS 10.3. The direction and length of flowing body were calculated in ArcGIS and included as an input shape file. Similarly, the contour map was included as a separate shape file in the input digital data.

Spatial Analyst Toolbox and Model Builder (ArcGIS applications) provides powerful tools for comprehensive, raster-based spatial modelling and analysis [15]. In the analysis, Kriging Interpolation technique, reclassification and Map Algebra/Raster calculator (ArcGIS application tools) were used to analyse the data.

Kriging Interpolation is a statistical interpolation technique. The Kriging function assumes that the distance or direction between sample points reflects a spatial correlation that can be used to explain variation in the surface [16]. In this study, Kriging Interpolation

---

Note 1: The Filter tool in ArcGIS was used to eliminate spurious data or enhance features otherwise not visibly apparent in the data. Filters create output values by a moving, overlapping 3 x 3 cell neighbourhood window that scans through the input raster. As the filter passed over each input cell, the value of that cell and its eight immediate neighbours were used to calculate the output value.
SURVEYING

(SURVEYING application tool) was used to create the raster surface from the already present subsidence areas in the Sterkfontein Cave area. However, in the model, the subsidence areas were manually added by extracting the affected cells from SPOT 5 imagery and DEM. A reclassification tool was used to assign weights to different input data in the model.

Weighting of digital data layers

All the factors involved in the model cannot have an equal weighting because each factor has its own importance on the outcome of subsidence hazard map. For weighting a multi-criteria decision analysis approach was adopted, which involved criteria of varying importance. Criterion-weighting expresses the relative importance by assigning a weight to each criterion. There are several approaches to criteria weighting. However, Analytical Hierarchical Process (AHP) is one of the most extensively used in GIS spatial analysis [17]. The AHP is a multi-criteria mathematical evaluation method using a pairwise comparison method. Developed by Saatay in 1980, this method involves pairwise comparisons to create a ratio matrix. It takes pairwise comparisons as input and produces the relative weights as output. A pairwise comparison of factors was carried out to assign weights to these factors in the Map Algebra tool in a Model Builder environment in ArcGIS. The pairwise comparison and the AHP process details are given in Table 2.

Geospatial data and geospatial analysis can be done either in raster or vector format. We selected a raster model because raster models are easy to manipulate and easy to interpret as data is in cellular format. Before we could proceed with the AHP multi-criteria analysis, we required the classification of the individual factor maps. For example, in the slope factor map, all slope angles cannot be equally weighted. Therefore, only slopes between certain angle limits were included. To understand the critical slope for subsidence triggering, we considered a mass of soil lying on a sloping ground. Three forces act on this mass of soil:

- Shear stress (gs) is the force causing movement parallel to a slope which increases with slope angle.
- Shear strength (gp) is the force of internal resistance to movement or force of cohesion and friction.
- Gravitational force (G) is the force with which body of the mass is attracted towards the centre of the Earth.

Slopes between 45° – 60° are considered as critical and were classified in the Reclassification Tool to get a slope factor map of the area. Similarly, the Reclassification Tool in ArcGIS was used to perform the individual classification of each factor maps before assigning the AHP weights.

The calculation unit of the subsidence geohazard model is the cell2 (C), where the cell presents a metric square. Each contributed factor map (Fm) is expressed in a matrix of cells that represent the raster model of that factor. Depending on the number of factor maps (n), the total contributed factors in the subsidence geohazard model are \( \sum F_{mn} \) for each cell of the subsidence geohazard map.

However, the contributed factor maps have different influences. Thus, each factor map \( F_{mn} \) has a coefficient \( Cn \) that determines the influence level on the subsidence geohazard map. Now the total contributed factors with their influence coefficients for each cell are expressed as follows:

\[
C = \sum F_{mn} \times Cn
\]

Where:

\[
C = \text{value of that cell of the hazard factor map } mn,
\]

\[
Cn = \text{influence level coefficient (AHP weight) associated with } F_{mn}.
\]

Each element in Equation 2 was assigned a weight through AHP, and the map algebra tool calculated the final output of the model. Each output cell was the product of all the weighted factors and represented whether \( F_{mn} \) factor was added or not in the final output.

Model Builder in ArcGIS was used to build the model that could house the Map Algebra tool and produce a spatial interpretation of the output map. The formula used in the map algebra tool was as follows:

\[
W_a = 0.489 + SL \times 0.262 + L_a \times 0.155 + SH \times 0.103
\]

Results and discussion

In order to calculate the areas which are hazardous for subsidence, we need to define the working units for calculations. A “cell” was kept as the calculation unit whose dimensions are shown in Fig. 5. Each cell has \( x, y \)

Note 2: Cell is defined as a picture element which has both spatial and spectral properties. The spatial property defines the dimensions of the corresponding ground area and the spectral property defines the intensity of the spectral response for a cell/pixel in a particular band. In the context of the research, one cell on the digital elevation model represents 30 m on the ground and one cell on the Spot 5 imagery represents 10 m on the ground.
coordinates and a height component, to be located in geographic space. The most important factor in the geohazard subsidence model is the spatial distribution of water accumulation areas. Water accumulating in a cell will percolate and reduce the strength of the soil mass, possibly inducing subsidence. Slopes, land-use type and close proximity to the already present sinkholes are factors that directly affect the geo-subsidence model results.

**Water accumulation hazard areas**
Cells with a high-flow accumulation are areas of concentrated flow. Cells with a flow accumulation of 0 are local topographic highs (for example ridges). High-flow accumulating cells are concentrated in the northeast and northwest boundaries of the cave. A total of 111 cells were identified which could accumulate water in case of rainfall. This total was approximately 20% of the total cells in the area of interest (see Table 3). Rain falling anywhere in the DEM had be able to follow a downhill path all the way to the edge of the DEM. In the first step, the flow direction was calculated. There are many different ways to distribute the flow from a cell. Here we opted for a D8 algorithm, used by spatial analyst. In the D8 algorithm, there are eight valid output directions relating to the eight adjacent cells into which flow could travel (Fig. 3). At each cell, a number was assigned based on the direction of the lowest elevation neighbour (steepest descent), as shown in Fig. 3, so that if the elevation was lowest to the left, the output flow direction would be 16. The direction of flow is determined by the direction of steepest descent, or maximum drop, from each cell. This drop was calculated as follows:

\[
\text{minimum drop} = \frac{\text{change in z value}}{\text{distance}} + 10^6
\]

The flow direction and water accumulation for each cell can be calculated by the Flow Accumulation tool in ArcGIS (Fig. 4). The Flow Accumulation tool tells us how many cells are upstream of a given cell. The output corresponds to the upstream drainage area of each cell. This command is recursive and can take a long time to run for large grids. We specified the output type as Integer in order to save space, as the flow accumulation was simply counting cells.

**Slopes hazard potential areas**
The slopes of the study area were calculated in ArcGIS using slope tool in geo-processing (Fig. 4). The analysis of the slope grid revealed that most the cave areas have slopes in the range of 20° to 40°. However, there were substantial slopes in the range of 45° to 60° which were critical for land subsidence. Slopes were then reclassified assigning a value of 1 to slopes between 45° to 60° and 0 values to all other slopes. The reclassified slope map had 140 cells between the ranges of 45° to 60°. This represented approximately 24% of the total cells in the area (Table 3). The area within the red mark in Fig. 6 has a high probability to sink because of their slope angle and high flow accumulation rate.

**Proximity to existing sinkholes**
The angle of repose is used to estimate the maximum radius of a sinkhole of a certain depth. As a sinkhole becomes...
deeper, the subsequent radius will become larger to affect other areas [20]. There are eight sinkholes in the area, with diameters varying from 1.3 to 6.1 m. A 5 x 5 cell neighbourhood window filter was used to calculate the cells affected by the sinkholes. Keeping the sink hole at centre 5 x 5 filter captured cells within 60 m of diameter. There were 136 affected cells in the area due to existing sink holes. This number is approximately 24% of total cells (see Table 3). This 24% is potentially hazardous and can contribute to land subsidence.

**Land-use potential hazard area**
The bare land mass has a low angle of repose as compared to grassy land or forest land which makes it more vulnerable to subsidence. After reclassification of the land-use map, approximately 11.8% of cells were critical for subsidence occurrence. The total number of cells was 67 (see Table 3).

**Land-subsidence risk**
After running the model in the ArcGIS environment in the model builder, land-subsidence risk cells was calculated. The total potential hazard cells were 29 in the case study area. If we only considered the cells in the actual cave area, eleven cells were critical for land subsidence according to the geospatial subsidence hazard model.

**Conclusion**
Land subsidence is a serious geological hazard in terrains hosting caves. GIS and remote sensing technologies were used to build a geospatial subsidence hazard model based on varying. A GIS geospatial reclassification tool was used to reclassify the factor maps for further analysis. GIS table functions were used to analyse the data statistically. The multi-criteria decision analysis technique was applied to the factor maps to weight different factors. A pairwise comparison matrix in AHP was developed and used in the model to find out the subsidence hazardous areas.

The conclusion and recommendations that could be drawn from the study are as follows:
- The potential hazardous area marked on SPOT 5 imagery (see Fig. 8) should be further investigated. The condition of the cave roof must be carefully analysed by ground monitoring for any deformations or rock movement.
- This study included factors whose digital data was easily available. It is recommended that another detailed study incorporating additional geotechnical factors be carried out to increase the accuracy of the results.
- The GIS model built in the Sterkfontein Cave area can be applied to other sensitive areas for ground subsidence.

**Acknowledgement**
This article is based on a paper that was published in the *South African Journal of Geomatics* Vol. 4, No. 3. August 2015 and has been republished here with permission.

**References**
The references for this article are available with the online version of the article: [http://wp.me/p5dDng-Bkm](http://wp.me/p5dDng-Bkm)
Contact Hamid Ashraf, University of the Witwatersrand, 804540@students.wits.ac.za

---

**Note 5:** The Filter tool in ArcGIS is used to eliminate spurious data or enhance features otherwise not visibly apparent in the data. Filters create output values by a moving, overlapping 5 x 5 cell neighbourhood window that scans through the input raster. As the filter passes over each input cell, the value of that cell and its 8 immediate neighbours are used to calculate the output value.

**Note 6:** ArcGIS application is used to create, edit and manage models.