Dear Editor,

Please find attached the revised manuscript and author response to the reviewers' comments of the manuscript "Hydrological characterization of cave drip waters in a porous limestone: Golgotha Cave, Western Australia" by Mahmud et al. for your consideration. We believe we have addressed all questions raised by the second reviewer of the manuscript.

As requested, we have included a detailed response to reviewer's questions (below), indicating page and line numbers where the changes are made in the revised manuscript. A marked-up manuscript version is also provided showing all the modifications.

If you have any additional questions, we would be more than happy to address them.

With kind regards,

For the authors,
Kashif Mahmud

Comments to the Author
Interactive comment on "Hydrological characterization of cave drip waters in a porous limestone: Golgotha Cave, Western Australia" by Mahmud et al.

Reviewer(s) Comments:
Anonymous Referee #1
The authors did a pretty good job responding to my concerns and the concerns of the other referee. Therefore, I feel confident recommending the revised version of the manuscript for publication as is.

We thank the reviewer for the positive feedback.

Reviewer(s) Comments:
Anonymous Referee #2
I must admit, I havn't gone over the revision as thoroughly as in the first round. Generally the authors followed the main theme of changes that were required and elaborated on the new methods and results (to lesser extent) rather than the old work. Nevertheless, I still have some troubling issues that must be corrected before it could published in HESS.

We appreciate reviewer's comment and clarify the issues raised by the reviewer.

I counted 26 references in the text to the single paper Mahmud et al., 2016. So as publishing the same figures was dealt with, the continuous citing of the same work again and again gives the feeling not much is added in this work. This number of citation of the previous study should be reduced significantly (by about 2/3), talk less about previous results.

We have now drastically reduced the references to previous work, which we agree was too heavily referenced. We only kept the references needed to clearly define which parts of the manuscript are novel or can be attributed to our previous publications.

There is a discrepancy in the coefficient of variation values between Figure 2 , Figure 4.
I don't think you have coefficient of variation of 10^4-10^5 like it looks in figure 2. In Figure 4 the highest is 100 (probably in % should be noted on axis title).

These were typos which are now corrected in Figure 2 and % is noted on axis title in Figure 5.

The drip data shown in histogram does not show any trend, and the 4 classes are generally distinct by discharge and the temporal statistics play a secondary role. The auto-correlation figure which was in the original manuscript opens the eye of the reader to some temporal statistics. The authors decided to move it to supplemental material, I think it should be in the main text of the manuscript as it is more informative than the MDS and clustering analysis.

The auto-correlation section is now moved to the main body of the manuscript (line 184-185 and section 4.3).

The authors give a wrong reference to Kurtzman et al., 2009. The following is the correct one:
Daniel Kurtzman, Joseph A. El Azzi, F. Jerry Lucia, Jerome Bellian, Christopher Zahm, Xavier Janson; Improving fractured carbonate-reservoir characterization with remote sensing of beds, fractures, and vugs. Geosphere; 5 (2): 126–139. doi: https://doi.org/10.1130/GES00205.1

This is now corrected (Section Reference).

Hydrological characterization of cave drip waters in a porous limestone: Golgotha Cave, Western Australia

Kashif Mahmud[1], Gregoire Mariethoz[2], Andy Baker[3], Pauline C. Treble[4]

[1]Hawkesbury Institute for the Environment, Western Sydney University, Australia
[2]Institute of Earth Surface Dynamics, University of Lausanne, Switzerland
[3]Connected Waters Initiative Research Centre, UNSW Australia, NSW, Australia
[4]Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW, Australia

Correspondence to: Kashif Mahmud (k.mahmud@westernsydney.edu.au)

Abstract

Cave drip water response to surface meteorological conditions is complex due to the heterogeneity of water movement in the karst unsaturated zone. Previous studies have focused on the monitoring of fractured rock limestones that have little or no primary porosity. In this study, we aim to further understand infiltration water hydrology in the Tamala Limestone of SW Australia, which is Quaternary aeolianite with primary porosity. We build on our previous studies of the Golgotha Cave system and utilize the existing spatial survey of 29 automated cave drip loggers and a LiDAR-based flow classification scheme, conducted in the two main chambers of this cave. We find that a daily sampling frequency at our cave site optimizes the capture of drip variability with least possible sampling artifacts. With the optimum sampling frequency, most of the drip sites show persistent autocorrelation for at least a month, typically much longer, indicating ample storage of water feeding all stalactites investigated. Drip discharge histograms are highly variable, showing sometimes multimodal distributions. Histogram skewness is shown to relate to the wetter than average 2013 hydrological year and modality is affected by seasonality. The hydrological classification scheme with respect to mean discharge and the flow variation, can distinguish between groundwater flow types in limestones with primary porosity, and the technique could be used to characterize different karst flow paths when high-frequency automated drip logger data are available. We observe little difference in the coefficient of variation (COV) between flow classification types, probably reflecting the ample storage due to the dominance of primary porosity at this cave site. Moreover, we do not find any relationship between drip variability and discharge within similar flow type. Finally, a combination of multi-dimensional scaling (MDS) and clustering by k-means is used to classify similar drip types based on time series analysis. This clustering reveals four unique drip regimes which agree with previous flow type classification for this site. It highlights a spatial homogeneity in drip types in one cave chamber, and spatial heterogeneity in the other, which is in concordance with our understanding of cave chamber morphology and lithology.

Keywords: karst aquifers, drip loggers, infiltration, cave drip water

1 Introduction

Karst features in limestone are typically developed from the solutional dissolution of fractures and bedding planes in carbonate rocks (Arbel et al., 2010; Kurtzman et al., 2009). Worldwide, karst regions represent significant geographical areas with potentially high rates of infiltration through fractured and karstified carbonate rocks. The most usual recharge method in karstic aquifers is the faster infiltration through the deep karstic openings (Ford and Williams, 2007). Complex spatial spreading of various karst features such as solutionally widened fractures, caves and conduits, makes the monitoring and precise groundwater recharge modeling very difficult (Lange et al., 2003; Arbel et al., 2010). The upper part of karstified rock (the epikarst zone) has higher permeability than the underlying vadose zone (Klimchouk, 2004). Therefore, infiltration into the epikarst zone is faster compared to the drainage through it, and water is kept stored in this region. This stored water in the vadose zone seeps slowly and finally emerges inside caves as infiltrating drip waters (Williams, 1983).

Karstic features such as speleothems, commonly used to reconstruct paleo-environmental records, are formed due to calcite deposition from cave drip water. Therefore, the knowledge of drip water hydrology is critical to study the paleoclimatic records (Baldini et al., 2006). An early study using tipping bucket loggers formulated a relationship between maximum discharge and coefficient of variation of discharge to categorize cave discharges (Smart and Friederich, 1987), for a fractured-rock limestone system with a vertical range of approximately 140 m (GB Cave, Mendip Hills, UK). They found that the drips close to the surface have extreme coefficient of variations, whereas the drips in depths have fairly constant flow rates over time, with a significant possibility of water storage in vadose zone fractures. Thus the stalagmite record resulting from slower drips may be more closely related to the karst hydrology rather than palaeoclimate (Baldini et al., 2006). This may also be a consequence of the developed connection between the surface and the cave. Quantitative analysis of such stalagmite drip data has, in the past, used manual observations of cave drips (e.g. Baker et al 1997). However, the recent development of automatic cave drip loggers (Collister and Mattey, 2008) has enabled the generation of high temporal resolution and continuous drip discharge time-series (e.g. (Jex et al., 2012; Cuthbert et al., 2014; Markowska et al., 2015; Mariethoz et al., 2012)), providing new opportunities for quantitative hydrological analysis.

Here we present monitoring data from Golgotha Cave located in SW Western Australia that has been extensively monitored since 2005, with the aim of better understanding karst drip water hydrogeology and the relationship between drip hydrology and surface climate. We build on the work of Mahmud et al. (2016), which presented the largest spatial and temporal survey of automated cave drip monitoring with matrix (primary) porosity published to date. This previous study consisted of data from two large chambers within this cave, measured in the period from August 2012 to March 2015, using a highly spatially (29 sites in two separate chambers) and temporally (0.001 Hz, 15 min intervals) resolved dataset. In a separate study, Mahmud et al. (2015) performed morphological analysis of karstic features, based on ground-based LiDAR data, to identify different flow processes in karstified limestone. Based on the findings of these two studies, here we investigate the relationship between drip water hydrology and cave depth, spatial location and stalactite type, and develop a hydrological classification scheme that is appropriate to high-frequency drip logger data and limestones with a

primary porosity. This classification scheme is also compared with previous studies (Smart and Friederich, 1987; Baker et al., 1997) to examine the limitations of these previous schemes. These findings will also help better characterize and understand water movement in highly porous karst formations.

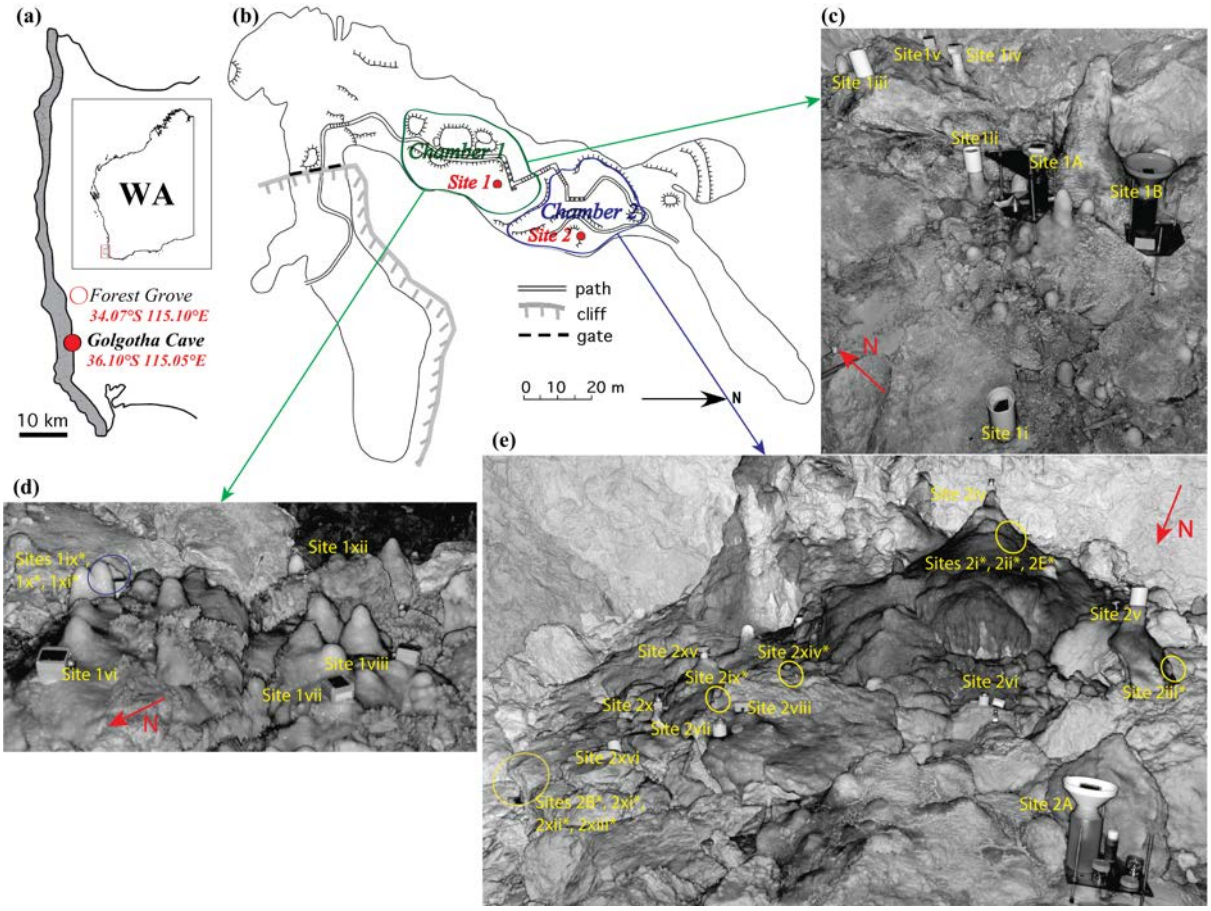
Finally, we use a combination of multi-dimensional scaling (MDS) and the popular K-Means algorithm for clustering similar drip characteristics. Time series clustering has been shown to be effective in providing useful information in various domains (Liao, 2005) and is implemented here to determine the degree of similarity between two drip time series. There seems to be an increased interest in time series clustering as part of the research effort in temporal data mining. The method we use here is suitable for large datasets, has been studied extensively in the past and achieves good results with minimum computational cost (Jex et al., 2012; Scheidt and Caers, 2009; Borg and Groenen, 1997).

2 Site Description

2.1 Studied Cave

The cave site has been explained in detail by Treble et al. (2013). Briefly, the field site, Golgotha Cave is 200 m in length and up to 25 m in width (Figure 1), is developed in Quaternary aeolianite, which consists of wind-blown calcareous sands that were deposited along the southwest coast of Australia (Brooke et al., 2014). Vadose zone water flow, and subsequent widening by ceiling collapse, formed the cave chambers. Treble et al. (2013) described the cave site as being developed in the Spearwood System of the Tamala Limestone and is mantled by a variable thick layer of sand formation having depths of between 0.3 m and 3 m. Diffuse (or matrix) flow is likely to be dominant in the Tamala Limestone formation due to its high matrix porosity as 0.3 – 0.5 (Smith et al., 2012). Karst in this region is also called “syngenetic” (Treble et al., 2013) that implies processes like preferential vertical dissolution and varying morphology of the subsurface caprock. These processes may establish vadose-zone preferential flow extending to the cave ceiling, with occasional rapid delivery of percolating waters deep into the calcarenite which end up seeping through to the cave ceiling. Therefore, this young limestone formation offers various opportunities for preferential flow into the hostrock and storage within it (Brooke et al., 2014). Golgotha Cave was chosen because (a) it is located in an intensively studied karst area (Treble et al., 2015; Treble et al., 2013; Treble et al., 2016), which has over ten years of manual and 3 years of automated drip water monitoring, (b) it contains actively growing speleothems, and (c) it is accessible year-round.

Based on previous studies at this site, we determined previously that Chamber 1 (Figure 1b, c and d) is mostly dominated by matrix flow representing water flowing down and seeping through the rock matrix, characterised by both icicle-shape and soda straw stalactites with slow drip rates of low variability. In contrast, Chamber 2 (Figure 1b and e) is typically controlled by fracture and combined flow, with high drip rates that are shown to vary over time depending upon the mode of water delivery to the preferential flow system. In fracture flow, water moves along the fracture orientation, forming curtain-shape stalactites in the direction of highest fracturing. Finally, combined flow is defined as the combination of conduit, matrix and fracture flow, resulting in a circular pattern of stalactite formation.



110

111 Figure 1: a) Coastal belt of SWWA (South-West Western Australia). (b) Golgotha cave plan view displaying
112 both Chamber 1 (green marked area), which comprises Site 1, and Chamber 2 (blue marked area) containing
113 Site 2. Average limestone thickness from cave ceiling to ground surface over Site 1 and 2 are 32.33 m and 40.24
114 m respectively. LiDAR scans of drip sites on: (c) Chamber 1 north floor, (d) Chamber 1 south floor and (e)
115 Chamber 2 floor. The red arrows show the geographic orientation (c, d and e). * indicates the sites where the
116 stalagmate loggers are not clearly visible in the LiDAR floor images as they are obscured by formations in front
117 of them, however the approximate locations are marked in yellow circles. Additional scans of cave ceiling and
118 photographs of underlying stalagmites are shown in Fig. 3 of Mahmud et al. (2016).

119 **2.2 Climate and Meteorology**

120 A comprehensive description of the climate at our study site has been presented in Treble et al. (2013). To
121 summarize, the site is a Mediterranean climate, associated with wet winters and dry summers. Annual rainfall
122 recorded at Forest Grove weather station (Figure 1a, 5 km away from the study site) is 1136.8 ± 184 mm, among
123 which ~75% occurs between May and September, with an average daily maximum temperature variation from
124 16°C (in July) to 27°C (in February) (BoM, 2017). Typically, the peak rainfall begins in late autumn (May) and
125 the wet season continues until end of September with a median monthly rainfall of ~100 mm. Each hydrological
126 year is defined as April to March, as April has the lowest water budget (precipitation-evapotranspiration).

As reported in previous studies, all hydrological years have water deficit during the dry season (October to April) and significant infiltration during the wet period. Low evaporative conditions during winter should permit increased infiltration to the caves, enhancing the drip discharge response to winter rainfall. The hydrological year 2012 had roughly similar annual rainfall of 1008.6 mm to the long-term annual mean, whereas 2013 was rather wet (total rainfall of 1239.8 mm) and 2014 was a relatively dry year with a total rainfall of 943.8 mm. Recorded rainfall was significantly above average in the 2013 hydrological year for various weather stations in Western Australia (BoM, 2017). Therefore, our site had a wetter winter in 2013 with an estimated annual recharge of 858.67 mm which is very much above average (ten year mean annual recharge is 564 mm).

2.3 Drip data acquisition and characteristics

Data acquisition and pre-processing has been previously described in Mahmud et al. (2016) and is concisely summarized here. Stalagmate drip loggers (www.driptych.com) were set up in approximate transects throughout the two large chambers from higher to lower ceiling elevation in 34 locations and are currently being monitored since August 2012. Each chamber has contrasting discharge, dune facies and karst features of Golgotha Cave (Figure 1). Data loggers were set to record continuously at 15 minute intervals. The notation used for site identification follows the same style as described in previous studies, consisting of a numerical number (representing the chamber) and a letter/roman number (representing a drip site within the given chamber, with a letter indicating the sites having both manual and automatic drip counts and a roman number specifying the sites only having drip logger data). Based on previous studies of the site, 29 sites are considered in the time series analysis although short periods of poor quality data were omitted if they were associated with changes in the mean and variability at the time of fieldwork. This impacted sites 1A, 1B, 2A, 2B, 2E as the logger was temporarily placed aside every 6 weeks in order to sample water from a collection bottle underneath the logger. Time series gaps are filled with synthetic data based on the drip statistics and correlation between drip rates.

As previously reported, drip rates in Chamber 1 are generally very low (the fastest drip rate was 25 drips per 15 mins) consistent with the predominance of matrix flow in this chamber. However, it is obvious that most drip loggers exhibit a clear response to the 2013 wet winter and also indicate the substantial inter-annual variation in discharge between three hydrological years. All Chamber 1 drip sites (except site 1x) show a gradual drip rate decrease during summer 2012 to winter 2013 due to below average rainfall in 2012. Then after displaying the sudden increase in all drip discharges that express the 2013 wet winter, the drip rates further reduce due to the dry 2014 hydrological year. This intra-annual variation is identified much greater than the inter-annual discharge variation of the drip sites, as previously observed in Baker et al. (1997). This suggests that high-resolution intra-annual drip rate data is helpful to obtain a complete picture of changing flow variability with recharge. The high resolution of the data sets includes precise characterization of the temporal behavior of an individual drip, illustrating the differences inherent to the drip sites.

In contrast, Chamber 2 drip rates present more variability between sites both in intra-annual and inter-annual discharges, except few very slow dripping sites. Of the Chamber 2 drips, the slow drip sites have the lowest coefficient of variations (COVs) and lowest discharges, indicative of matrix flow types. The timing of maximum drip rates is generally delayed in Chamber 2 versus Chamber 1: Chamber 1 drip rates typically peak in late

spring/early summer (Oct-Dec) while Chamber 2 drips tend to peak a few months later (Dec-May), reflecting a longer water residence time. This may be a function of the thicker ceiling above Chamber 2 (40.24 versus 32.33 m) but also heterogeneity in flow paths to each chamber. Overall the drip response to the 2013 wet winter is amplified in Chamber 2 versus Chamber 1, consistent with the presence of greater fracture flow in Chamber 2.

By applying morphological analysis of ceiling features acquired by LiDAR data, Mahmud et al. (2015) distinguished three flow patterns (i.e. matrix flow, fracture flow, and a combination of conduit, fracture, and matrix flow) for the observed ceiling morphological features. All the drip sites were then characterized according to this flow classification in Mahmud et al. (2016), which is used here as a reference for clustering similar drip time series.

3 Methods

3.1 Hydrological classification of cave drips

Research involving automated drip monitoring systems is increasing, for example at Cathedral Cave in Wellington (Cuthbert et al., 2014) and Harrie Wood Cave in the Snowy Mountains, Yarrangobilly (Markowska et al., 2015). The variability of the drip discharge might not only be a function of discharge itself, but could also depend on the sampling frequency. We investigate this possibility by plotting the COV versus sampling interval (the original 15 mins and calculated by resampling the data at 1 hour, 1 day, 1 week and 1 month). COV is supposed to be artificially high at the high frequency of 15 mins because of sampling bias that artificially increases the noise. The resampling at low frequencies is only a way of smoothing out this noise. Using the optimum sampling frequency to minimize its effect on drip variability, we plot drip rate histograms to identify the response of drips between the flow classifications and the response to intra and inter-annual variability in infiltration. We also plot the autocorrelation functions (ACFs) to investigate the relationship between the strength of correlation and the LiDAR-based flow type. Finally, we summarize the mean discharge of drip sites in relation to the variability in discharge using the optimum sampling frequency. These are the same drip discharge parameters as used in the classification method proposed by Friederich and Smart (1982), Fairchild et al. (2006) and Baker et al. (1997) that were based on manual drip collection at low frequency.

3.2 Clustering of similar drip time series

We employed multi-dimensional scaling (MDS), which allows data dimensionality reduction i.e., mapping complex multidimensional data on a low-dimensional manifold. MDS is a technique that embeds a set of points in a low-dimensional space, so that the distances between the points resemble as closely as possible a given set of dissimilarities between the objects they represent (Birchfield and Subramanya, 2005). MDS requires a distance matrix to be computed, in which a single scalar number characterizes the similarity between any two time-series. In our case, each drip logger is an object and a specific distance between drip loggers is considered to characterize the similarity between any two loggers. It takes an input matrix giving dissimilarities between pairs of items and outputs a coordinate matrix whose configuration minimizes a loss function. MDS is also known as Principal Coordinates Analysis (PCoA). MDS operates on a distance or dissimilarity matrix (Pisani et al., 2016), which is different than principle-component analysis (PCA) that is based on a covariance matrix.

Even if PCA and MDS methods can return the same results in specific contexts, MDS can be considered more general because it remains validity for non-euclidean distances, such as the distance matrix (*d*) chosen in this study. MDS is used to translate these distances into a configuration of points defined in an *n*-dimensional Euclidean space (Cox and Cox, 1994). A MDS results in a set of points arranged so that their corresponding Euclidean distances indicate the dissimilarities of the time series. According to Birchfield and Subramanya (2005) the basic steps of performing the MDS algorithm are:

i) Construct the distance matrix **D**: One key component in clustering is the function used to measure the temporal similarity (or distance) between any two time-series being compared. To define an appropriate measure of similarity between time series, we determine two factors: firstly, the offset (*O*) to match two time-series based on their maximum correlation, and secondly the complement of the correlation coefficient (1-*R*) between the time series (Jex et al. 2012). Initially, we compute the cross-correlation function, a measure of similarity of two time-series as a function of the displacement of one relative to the other. The cross-correlation function is an estimate of the covariance between two time-series, y_{1t} and y_{2t}, at lags $k = 0$, ±1, ±2, ... The offset (*O*) is defined as the lag time based on the maximum correlation between two time-series. Next, we define R as the correlation coefficient with the time series being moved by the offset amount *O* to have maximum correlation coefficient. Both *O* and *R* are calculated to all *n(n-1)/2* pairs of drip data, where *n* is the number of drip data. Here, we use the original recorded drip counts in 15 mins interval. The sampling bias discussed in section 3.1 is only affecting the drip variability, not the cluster analysis. Moreover, high resolution (15 mins interval) data are more suited for the cluster analysis because it allows better defining the cross-correlation between drips, as sometimes the offset of maximum correlation *O* might be less than a day. Finally, the distance matrix **D** is computed for each pair of loggers using the following equation (Jex et al. 2012):

$$\boldsymbol{D} = O(1 - R)$$

The distance matrix (**D**) is square, symmetric, and has dimension equal to the number of drip loggers.

ii) Compute the inner product matrix $B = -\frac{1}{2}J\boldsymbol{D}J$, where $J = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ is the double-centering matrix and **1** is a vector of ones.

iii) Decompose B as $B = V\Lambda V^T$, where $\Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$, the diagonal matrix of eigenvalues of B, and $V = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$, the matrix of corresponding unit eigenvectors. Sort the eigenvalues in non-increasing order: $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.

iv) Extract the first p eigenvalues $\Lambda_p = \text{diag}\,(\lambda_1, \ldots, \lambda_p)$ and corresponding eigenvectors $V_p = [\mathbf{v}_1, \ldots, \mathbf{v}_p]$.

v) The corresponding Euclidean distances of the set of points, indicating the dissimilarities of the time series, are now located in the $n \times p$ matrix $X = \left[\mathbf{x}_1, \ldots, \mathbf{x}_p\right]^T = V_p {\Lambda_p}^{1/2}$.

The k-Means clustering algorithm is then used to divide these points into k clusters, which corresponds to a categorization of the drip data time series. k-means clustering, or Lloyd's algorithm (Lloyd, 1982), is a method of vector quantization that is popular for cluster analysis in data mining. k-means clustering aims to partition n observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. The algorithm proceeds as follows:

i) Choose *k* initial cluster centers (*centroid*): Here, we use *k*=4 clusters as this was the number of flow categories identified in previous work at this site.

ii) Compute point-to-cluster-centroid distances of all observations to each centroid. There are two steps to follow: first assign each observation to the cluster with the closest centroid. Then individually assign observations to a different centroid if the reassignment decreases the sum of the within-cluster, sum-of-squares point-to-cluster-centroid distances.

iii) Compute the average of the observations in each cluster to obtain *k* new centroid locations.

iv) Repeat steps 2 and 3 until cluster assignments do not change, or the maximum number of iterations is reached.

4 Results and Discussion

4.1 Determining the relationship between sampling frequency and drip discharge COV

We test the variability of drip discharge COV with the sampling frequency in Figure 2, to find the optimum sampling frequency that minimizes sampling artifacts while maximizing the capture of natural variability. For high discharge, COV increases with sampling frequency, which we explain by the smaller sampling interval better capturing the actual drip variability. For low discharges, COV also increases with sampling frequency, which we explain by the variability introduced due to drip rates being less than the sampling frequency. From the data presented in Figure 2, we can conclude that for both chambers and to compare all different types of flow, a sampling frequency of 1 day gives the minimum COV, which does not change significantly with a finer sampling frequency. Therefore, we use a sampling frequency of 1 day that minimizes sampling artifacts while maximizing the capture of natural variability. For Golgotha Cave, this would be to sum the 15 minutes drip rates over a 1-day period. This optimized sampling frequency is used to plot the histograms (section 4.2), ACFs (supplementary section S1) and examine the drip discharge behavior with drip variability for various flow types (section 4.3).

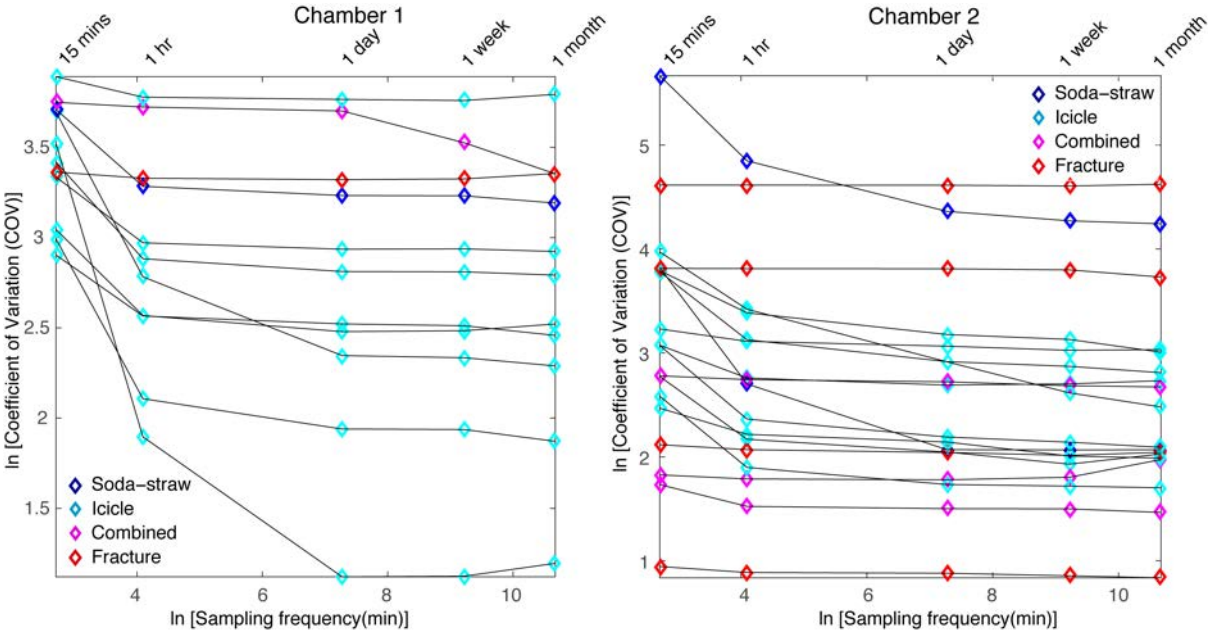


Figure 2: Optimum sampling frequency that minimizes sampling artifacts while maximizing the capture of natural variability.

4.2 Drip rate frequency distributions

Figure 3 shows the drip rate histograms for representative drip sites and different flow categories with optimum sampling frequency of 1-day. Drip sites are organized from lowest to highest discharge in each flow classification. Slow dripping soda-straw flows (e.g. sites 2xi, 2iii and 1v) show variation of drips with seasonality and the response to wetter recharge period with an approximate six-month lag, which suggests the drip water is supplied from storage in the limestone formation. Among these, site 1v displays the response to recharge in much shorter duration, the 6 months following 2013 recharge and then a shift to lower flow rates which may represent flow poaching. The histograms for icicle and combined flow systems represent unimodal skewed to bimodal distributions, indicating the shift to higher drip rates in response to the wetter 2013 hydrological year (except site 2xiii, which shows a shift to lower drip rates). The rest of the fracture sites show bimodal or multimodal distributions. With the limited temporal scale of the analysis, it seems that the histograms with skewed distributions represent the consequences of wetter 2013 hydrological year. These skewed distributions seem to have higher drip rate response to the drier 2014-15 period rather than the earlier normal/wetter years. This clearly denotes potential refilling of storage within the system during the 2013 wet winter, and later supplying drip water in 2014-15 seasons. In contrast, the bimodal distribution of site 2viii indicates the drip response to the annual cycle of wet and dry seasons of each hydrological year with an approximate six-month lag. Several bimodal (e.g. site 1x) and multimodal (e.g. sites 2xvi, 2vi) distributions, characterize as fracture flow, also distinguishes the dry period of 2012 - 2013 (having low drip rates) from the later period of 2013 wet winter (with high drip rates).

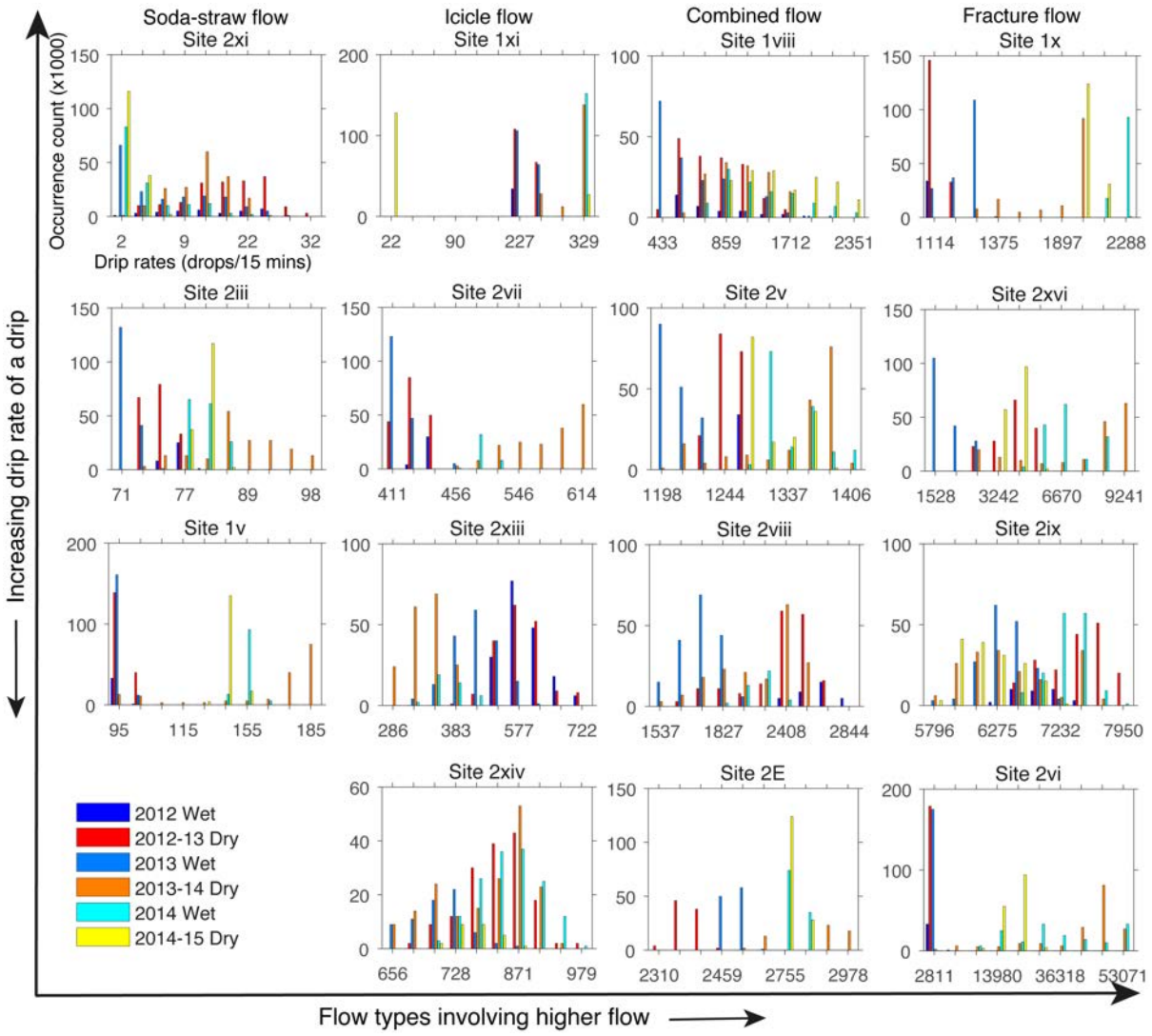


Figure 3: Histogram plots of both chambers drip data according to four flow types identified in Mahmud et al. (2016). Each histogram represents the frequencies of the drip counts per day (The axes labels are shown in the first histogram). Bin size is uniform for all plots and the external tick marks in x-axes delineates the bin intervals. The legend shows all the seasons over the monitoring period (blue to cyan for wet seasons: April to September and red to yellow for dry seasons: October to March, with the color gradually shifting for different years). The 2012 wet season experienced similar rainfall to the long-term annual mean, whereas 2013 was rather wet and 2014 was a relatively dry year. Histogram data for all sites appear in Supplementary Figure S1.

4.3 Autocorrelation functions (ACFs)

We investigate the use of ACFs to analyze drip behavior using the optimum sampling frequency of 1-day and until lags of 365 days. We do not find significant yearly autocorrelation with this limited 3 years of data. In some drips, a negative correlation occurred, but it is very insignificant and no physical process can explain a negative yearly correlation. Therefore, we plot ACFs in Figure 4 for different flow categories with the optimum sampling frequency of 1-day and lag time of 200 days. All sites have an autocorrelation that persists for at least a month, and often much longer. However, there is no relationship between the strength or the temporal decay of

the correlation and the LiDAR-based flow classification. This indicates the presence of ample storage in the system, supplying all stalactite types.

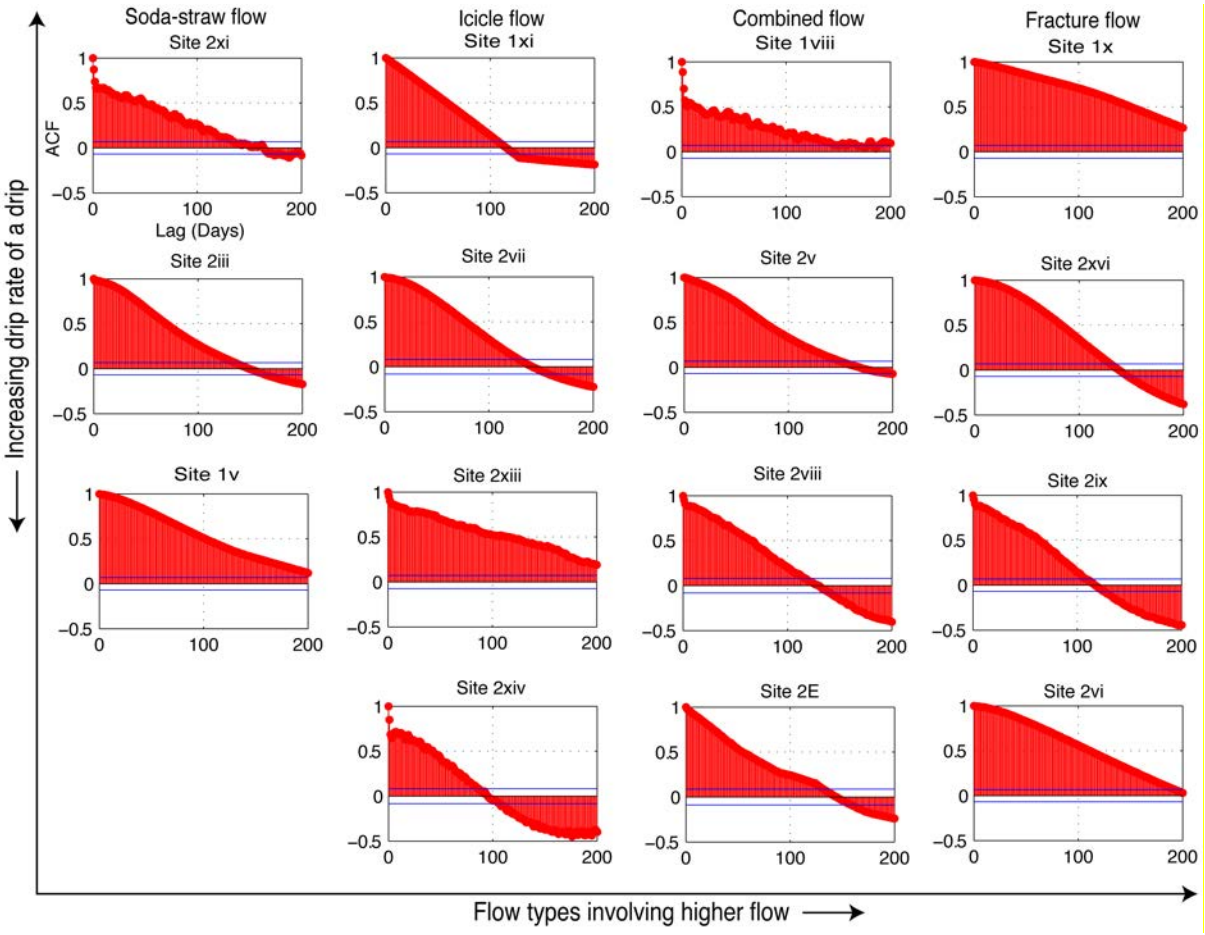


Figure 4: Autocorrelation functions of both chambers drip data according to flow classification of Mahmud et al. (2016). X- and Y-axis of individual plots represents the lag (in days) and ACF respectively (The axes labels are shown in the first ACF plot). ACFs for all sites appear in Supplementary Figure S2.

4.4 Hydrological classification of cave drips

We examine the hydrological behavior of the drips at daily resolution with respect to mean discharge and flow variation in Figure 5. It is clear from Figure 5 that there is no relationship between COV and flow-type. One soda-straw discharge (site 2xi) has a seasonal dryness, a very low discharge, and a very high coefficient of variation due to its irregular dripping. Otherwise, nearly all soda-straw flow, icicle flow, combined flow and fracture flow drips have COV <60%, with the exception of one fracture flow site showing the highest COV (Figure 5). But in general, there is little difference in the COV between classification types, probably reflecting the ample storage (supplementary section S1) due to the dominance of primary porosity at this cave. We do not clearly observe increasing variability with decreasing discharge within similar flow type, in contrast to other studies from older, fractured rock limestones (Smart and Friederich, 1987; Baldini et al., 2006; Baker et al., 1997). This shows that Golgotha Cave drip sites do not fit within the drip classification method proposed by

Smart and Friederich (1987) and Baker et al. (1997), which were based on manual drip counts with limited number of intermittent drip sites. Moreover, we utilize drip data from a cave with primary porosity, capturing the full range of flow types from matrix through to fracture, whereas the previous classifications only captured slow vs fast drips that were likely dominated by fracture flow paths given the host rock setting.

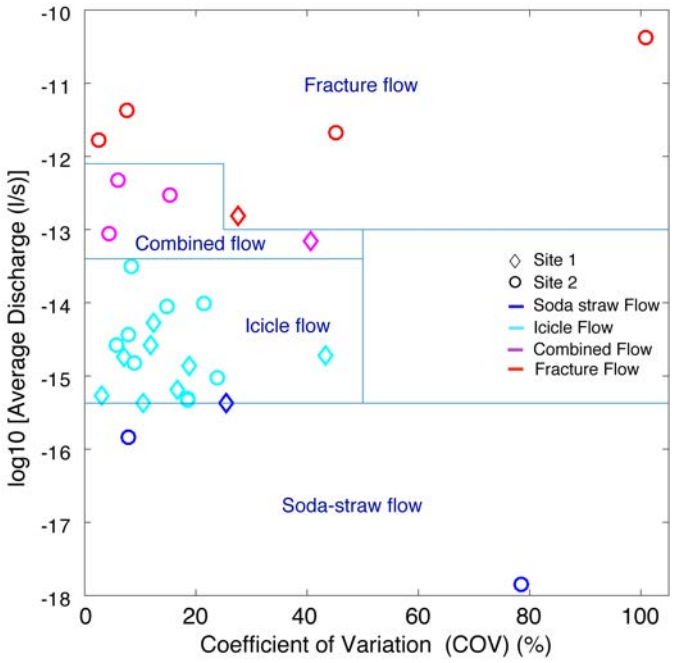


Figure 5: Hydrological behaviour of drip sites expressed in terms of daily mean discharge versus daily discharge variability calculated from the automatic drip rate data for three hydrological years. Measured drip rates are converted to volume units assuming a drip volume of 0.1433 ml (Genty and Deflandre, 1998). Blue lines and symbols reflect flow classification given in Mahmud et al. (2015).

4.5 Clustering of similar drip time series

The clustering results are overlain upon the chamber ceiling images in Figure 6 and also summarized in Tables 1 and 2 with the average drip discharges and flow type classification based on LiDAR. Average drip discharges are calculated from the 15-minute drip rates. As mentioned above, drip logger time series are deemed similar if they are well correlated and only have a small offset with each other, and so these time series should cluster together. Most of the drip sites that are identified as matrix flow (soda-straw and icicle flow) cluster together in C1. However, three of the icicle flow sites with drip rate greater than 4 per 15 minutes fall in C2. The combined flow category and the fracture type usually cluster in C3 and C4 respectively. Therefore we observe that our clustering generally agrees with the morphology-based flow classification of Mahmud et al. (2016). Few of the flow classes show exceptions, for example site 2vi is a fracture type flow and cluster in C1. This site has really high discharge with high variability, showing irregular drip rate.

One consistent feature that appears from the cluster analysis of Figure 6 is the spatial homogeneity of the clusters in Chamber 1, suggesting that they are spatially connected, or that their flow paths are connected to the same hydrological domain (the karst matrix), and supporting the overall dominant matrix flow patterns (both

337 soda-straw and icicle). Chamber 2 presents a completely different situation, where it is obvious that drip sites
338 can have similar behavior (well correlated together with a small lag), and be spatially distinct features, separated
339 by spans of approximately 6 meters (Figure 6). In particular, clusters 3 and 4 are spatially scattered, representing
340 the presence of fractures and combined flow systems throughout the chamber ceiling. This indicates an overall
341 strong heterogeneity of the flow paths between the surface and the cave for Chamber 2. Hence, in Chamber 2 we
342 expect flow paths to be more complex with routing between multiple stores and interconnected fracture
343 networks potentially resulting in non-linear response to infiltration. This is supported by dripwater $\delta^{18}O$ data for
344 this chamber (Treble et al., 2013).

345 Table 1: MDS cluster groups with statistical properties of Chamber 1 drip data.

Site/Stalagmate	MDS Cluster Group	Average drip discharge (l/yr)	Flow type (LiDAR-based)
1A	1	19.8	Icicle
1B	1	12.6	Icicle
1i	1	6.6	Icicle
1ii	1	11.2	Icicle
1iii	1	8.1	Icicle
1v	1	6.7	Soda-straw
1vi	1	7.4	Icicle
1viii	2	60.9	Combined
1ix	1	14.8	Icicle
1x	3	86.2	Fracture
1xi	1	12.7	Icicle

346 Table 2: MDS cluster groups with statistical properties of Chamber 2 drip data

Site/Stalagmate	MDS Cluster Group	Average drip discharge (l/yr)	Flow type (LiDAR-based)
2A	1	9.4	Icicle
2B	1	17.1	Icicle
2E	3	140.3	Combined
2i	4	243.0	Fracture
2iii	1	4.2	Soda-straw
2iv	1	14.6	Icicle
2v	3	67.8	Combined
2vi	1	985.0	Fracture
2vii	2	25.0	Icicle
2viii	3	113.8	Combined
2ix	4	360.2	Fracture
2x	1	7.0	Icicle
2xi	1	0.6	Soda-straw
2xiii	2	26.2	Icicle
2xiv	2	42.8	Icicle
2xv	1	11.6	Icicle
2xvi	3	266.9	Fracture
2xvii	1	7.0	Icicle

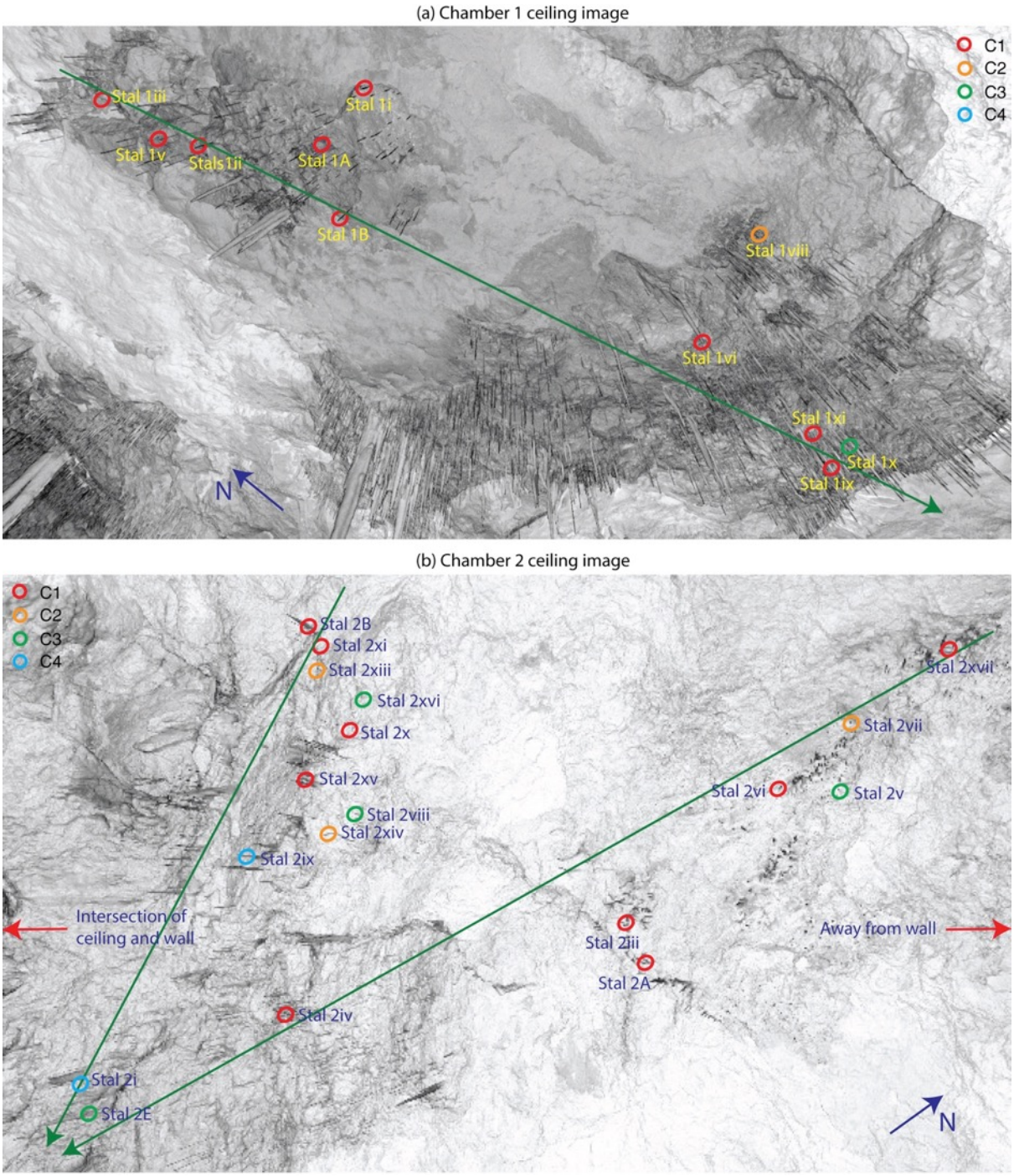


Figure 6: Cluster group plot overlain upon the cave ceiling for both chambers. The ceiling images are captured by LiDAR and the circles represent the ceiling locations of stalactites dripping on various stalagmates in both chambers (shown in Figure 1). The colour of the circles indicates individual MDS cluster group. The blue arrows in both Figures show the geographic orientation and the green arrows represent the approximate transects throughout the chambers from higher to lower ceiling elevation.

5 Implications of the findings and future research

Starting with the time-series analysis, this research presents a methodology that can be applied globally for drip logger data. The results show that some data-integration is necessary to avoid artefacts from slow drip sites. For

356 sites where there is significant matrix flow, our study has demonstrated that the Smart and Friederich
357 classification is not appropriate. Therefore, this study has presented alternative hydrological classification
358 schemes that are suitable for cave sites that include matrix flow. The times series approach adopted in this study
359 also opens the way for improved analysis and classification of hydrology time series in general i.e. tests for
360 histogram, autocorrelation, cluster analysis, and all of these will certainly benefit our understanding of the
361 hydrology of karst systems.

362 In this study, we also extend the analysis of drip time series to multiple sites, whereby we take advantage of the
363 ensemble of loggers to extract common properties by clustering, which would not be possible with single site
364 analysis. The results show that by considering multiple simultaneous time series, one can make better inferences
365 about water flow and unsaturated zone properties. The main impact is to recommend the use of spatial networks
366 of loggers over individual loggers. It should be noted that currently, most researchers deploy only a few loggers
367 to understand the flow to individual sites. This study also proposes a possible methodology for the analysis of
368 such datasets.

369 Regarding application of our findings, we believe that our methodology based on drip logger datasets can
370 provide direct evidence of deep drainage, and therefore the timing of diffuse recharge, which could be used for
371 basic model calibration. Spatial drip data (possibly combined with Lidar) is beneficial to infer flow types (e.g.
372 the proportion of fracture vs matrix, etc.) which could be used for model configuration to produce realistic karst
373 recharge (Hartmann et al., 2012), and hence large-scale groundwater estimation (Hartmann et al., 2015).
374 Another potential application is the integration of flow types in groundwater models through inverse modelling.
375 Such data could also be used to constrain water isotope model configurations used for forward modelling
376 speleothem $\delta^{18}O$ (Bradley et al., 2010; Treble et al., 2013). Overall, the findings of this work will definitely
377 provide a better understanding of processes that control vadose zone flow and transport processes, which would
378 ultimately help develop approaches to incorporate these processes into simulation models (Hartmann and Baker,
379 2017).

380 The analysis, presented here and combined with the findings of previous work at this site, provides valuable
381 information for paleoclimatologists and geochemists wishing to sample stalagmites. While these studies have
382 characterised Golgotha Cave, they could be applied to any other cave system. In our previous work, we have: 1)
383 devised a classification for flow-type based on stalactite morphology; 2) quantified the recharge response of
384 each flow type to infiltration and; 3) combined the findings of points 1-2 to estimate the total volume of cave
385 discharge; 4) compared cave discharge with infiltration to estimate the total recharge volume and identify highly
386 focused areas of recharge. The current study has further developed the spatial and temporal statistical
387 relationships between the flow sites, allowing both quantification and visualisation of the hydrology between the
388 ground surface and the cave ceiling. More generally, these studies illustrate the heterogeneity between flow sites
389 and demonstrate methods that can be applied to any cave system for studying diffuse recharge and paleoclimate
390 records from speleothems.

391 We further propose some ideas for future research that have evolved from this study:

a) Combining a drip logger network with a surface weather station and soil moisture network to constrain the water balance in hydrological models. Additionally, employing sap flow meters could allow constraining tree water use.

b) Combining the logger network, which constrains diffuse recharge, to boreholes measuring groundwater level to understand the relative importance of diffuse and river recharge.

c) Combining cave drip logger data with surface geophysics data to track water movement.

6 Conclusion

Cave drip water response to surface climatic conditions is often complex due to numerous interacting drip routes with varying response times (Baldini et al., 2006). This study explores the relationship between drip water and rainfall in a SW Australian karst, where both intra- and inter-annual hydrological variations are strongly controlled by seasonal variations in recharge. The multi-year drip response data capture the inter-annual drip water variability that are likely to be greater than intra-annual variability as suggested by Baker et al. (1997). Building on previous work, we further analyse a set of statistical properties of three hydrological years of drip data under varying precipitation rates. We test the relationship between drip discharge variability and drip data sampling frequency to determine the optimum sampling frequency that maximizes the capture of natural variability with minimum sampling artifacts. Using the daily optimum sampling frequency, the histogram distributions of various drip data time series illustrate the differences between the flow classifications. Most of the drip sites show persistent autocorrelation for at least a month. The hydrological behavior of the drips is examined with respect to mean discharge and the flow types similar to the classification method proposed by previous researchers (Smart and Friederich, 1987; Baldini et al., 2006; Baker et al., 1997). The drip sites at Golgotha Cave described in this study do not fit within the drip classification method proposed by Smart and Friederich (1987) and Baker et al. (1997). These previous studies were based on manual drip counts with limited number of intermittent drip sites. Here we overcome these limitations with automated drip monitoring system.

Finally, we apply a well-developed clustering method to determine the degree of similarity between drip time series. The clustering indicates one dominating group: C1 (characterized by matrix flow type) with very slow continuous drip discharge indicating matrix porosity in the thick limestone formation. This finding concurs with the observed cave chamber morphology and lithology. Moreover, the cluster analysis agrees with the flow classification of Mahmud et al. (2016) by grouping similar flow type in one single cluster. Overall this study establishes a novel way to characterize cave hydrology, which can be obtained by performing together both methodologies of Mahmud et al. (2015) and Jex et al. (2012). It relies on a metric that defines drip logger time series as similar if they are well correlated and only have a small offset with one another, and therefore these time series should cluster together. The MDS analysis supports this hypothesis and moreover, displays the spatial patterns of the flow paths between the surface and the cave chambers. This technique shows potential to classify, quantify and visualise the observed relationships between infiltration through the fractured limestone rocks and surface climate inputs.

Over the last decade, the automation of cave drip water hydrology measurements has permitted the routine generation of continuous hydrological time series for the first time. This study demonstrates a complete methodology for such datasets, which will help better characterize karst drip water hydrogeology and understand the relationship between drip hydrology and surface climate at any cave site where such measurements are made. We demonstrate that the analysis of the time series produced by cave drip loggers generates useful hydrogeological information that can be applied generally, beyond the example presented here. The time series behaviour integrates a variety of characteristics that combine the properties of the epikarst (storage), fracture configuration, and recharge. The clustering approach can identify which drip behaviour are related to these cave characteristics, and their spatial relationship. Most importantly, information on cave characteristics can now be gathered at a very low cost in terms of measurement and time.

Acknowledgment

This paper is based on work supported by UNSW Australia, UNSW Connected Waters Initiative Research Center and the National Centre for Groundwater Research and Training. The authors wish to thank individuals (Andy Spate, Alan Griffiths, Liz McGuire, Carolina Paice, Anne Wood, Monika Markowska and others) who assisted in data acquisition at Golgotha cave site.

References

Arbel, Y., Greenbaum, N., Lange, J., and Inbar, M.: Infiltration processes and flow rates in developed karst vadose zone using tracers in cave drips, Earth Surface Processes and Landforms, 35, 1682-1693, 10.1002/esp.2010, 2010.

Baker, A., Barnes, W. L., and Smart, P. L.: Variations in the discharge and organic matter content of stalagmite drip waters in Lower Cave, Bristol, Hydrological Processes, 11, 1541-1555, 10.1002/(sici)1099-1085(199709)11:11<1541::aid-hyp484>3.0.co;2-z, 1997.

Baldini, J. U. L., McDermott, F., and Fairchild, I. J.: Spatial variability in cave drip water hydrochemistry: Implications for stalagmite paleoclimate records, Chemical Geology, 235, 390-404, http://dx.doi.org/10.1016/j.chemgeo.2006.08.005, 2006.

Birchfield, S. T., and Subramanya, A.: Microphone Array Position Calibration by Basis-Point Classical Multidimensional Scaling, IEEE Transactions on Speech and Audio Processing, 13, 1025-1034, 10.1109/TSA.2005.851893, 2005.

BoM: Climate Data Online (Station 9547), Bureau of Meteorology Melbourne. http://www.bom.gov.au/climate/data/ (Accessed 26-08-2017), in, 2017.

Borg, I., and Groenen, P.: Modern multidimensional scaling: theory and applications, Springer, New York, 614 pp., 1997.

459 Bradley, C., Baker, A., Jex, C. N., and Leng, M. J.: Hydrological uncertainties in the modelling of cave drip-
460 water δ18O and the implications for stalagmite palaeoclimate reconstructions, Quaternary Science Reviews, 29,
461 2201-2214, 2010.

462 Brooke, B. P., Olley, J. M., Pietsch, T., Playford, P. E., Haines, P. W., Murray-Wallace, C. V., and Woodroffe,
463 C. D.: Chronology of Quaternary coastal aeolianite deposition and the drowned shorelines of southwestern
464 Western Australia – a reappraisal, Quaternary Science Reviews, 93, 106-124,
465 http://dx.doi.org/10.1016/j.quascirev.2014.04.007, 2014.

466 Collister, C., and Mattey, D.: Controls on water drop volume at speleothem drip sites: An experimental study, J.
467 Hydrol., 358, 259-267, http://dx.doi.org/10.1016/j.jhydrol.2008.06.008, 2008.

468 Cox, T., and Cox, M.: Multidimensional scaling, Chapman and Hall, London, 213 pp., 1994.

469 Cuthbert, M. O., Baker, A., Jex, C. N., Graham, P. W., Treble, P. C., Andersen, M. S., and Ian Acworth, R.:
470 Drip water isotopes in semi-arid karst: Implications for speleothem paleoclimatology, Earth Planet. Sci. Lett.,
471 395, 194-204, http://dx.doi.org/10.1016/j.epsl.2014.03.034, 2014.

472 Fairchild, I. J., Tuckwell, G. W., Baker, A., and Tooth, A. F.: Modelling of dripwater hydrology and
473 hydrogeochemistry in a weakly karstified aquifer (Bath, UK): Implications for climate change studies, J.
474 Hydrol., 321, 213-231, http://dx.doi.org/10.1016/j.jhydrol.2005.08.002, 2006.

475 Ford, D., and Williams, P.: Karst Hydrogeology and Geomorphology, Wiley, 576 pp., 2007.

476 Friederich, H., and Smart, P. L.: The classification of autogenic percolation waters in karst aquifers: A study in
477 G.B. cave, Mendip Hills, England, Proceedings of the University of Bristol Speleological Society, 1982, 143–
478 159, 1982.

479 Genty, D., and Deflandre, G.: Drip flow variations under a stalactite of the Pere Noel cave (Belgium). Evidence
480 of seasonal variations and air pressure constraints, J. Hydrol., 211, 208-232, 1998.

481 Hartmann, A., Lange, J., Weiler, M., Arbel, Y., and Greenbaum, N.: A new approach to model the spatial and
482 temporal variability of recharge to karst aquifers, Hydrol. Earth Syst. Sci., 16, 2219-2231, 10.5194/hess-16-
483 2219-2012, 2012.

484 Hartmann, A., Gleeson, T., Rosolem, R., Pianosi, F., Wada, Y., and Wagener, T.: A large-scale simulation
485 model to assess karstic groundwater recharge over Europe and the Mediterranean, Geosci. Model Dev., 8, 1729-
486 1746, 10.5194/gmd-8-1729-2015, 2015.

487 Hartmann, A., and Baker, A.: Modelling karst vadose zone hydrology and its relevance for paleoclimate
488 reconstruction, Earth-Science Reviews, 172, 178-192, http://dx.doi.org/10.1016/j.earscirev.2017.08.001, 2017.

489 Jex, C. N., Mariethoz, G., Baker, A., Graham, P., Andersen, M., Acworth, I., Edwards, N., and Azcurra, C.:
490 Spatially dense drip hydrological monitoring and infiltration behaviour at the Wellington Caves, South East
491 Australia, International Journal of Speleology, 41, 283–296, 2012.

492 Klimchouk, A.: Towards defining, delimiting and classifying epikarst: Its origin, processes and variants of
493 geomorphic evolution, Speleogenesis and Evolution of Karst Aquifers, 2, 1-13, 2004.

494 Kurtzman, D., El Azzi, J. A., Lucia, F. J., Bellian, J., Zahm, C., and Janson, X.: Improving fractured carbonate-
495 reservoir characterization with remote sensing of beds, fractures, and vugs, Geosphere, 5, 126-139,
496 10.1130/ges00205.1, 2009.

497 Lange, J., Greenbaum, N., Husary, S., Ghanem, M., Leibundgut, C., and Schick, A. P.: Runoff generation from
498 successive simulated rainfalls on a rocky, semi-arid, Mediterranean hillslope, Hydrological Processes, 17, 279-
499 296, 10.1002/hyp.1124, 2003.

500 Liao, T. W.: Clustering of time series data-a survey, Pattern Recogn., 38, 1857-1874,
501 10.1016/j.patcog.2005.01.025, 2005.

502 Lloyd, S.: Least squares quantization in PCM, IEEE Transactions on Information Theory, IT-28, 129-137, 1982.

503 Mahmud, K., Mariethoz, G., Pauline, C. T., and Baker, A.: Terrestrial Lidar Survey and Morphological Analysis
504 to Identify Infiltration Properties in the Tamala Limestone, Western Australia, Selected Topics in Applied Earth
505 Observations and Remote Sensing, IEEE Journal of, 8, 4871 - 4881, 10.1109/JSTARS.2015.2451088, 2015.

506 Mahmud, K., Mariethoz, G., Baker, A., Treble, P. C., Markowska, M., and McGuire, L.: Estimation of deep
507 infiltration in unsaturated limestone environments using cave LiDAR and drip count data, Hydrol. Earth Syst.
508 Sci., 20, 359-373, 10.5194/hess-20-359-2016, 2016.

509 Mariethoz, G., Baker, A., Sivakumar, B., Hartland, A., and Graham, P.: Chaos and irregularity in karst
510 percolation, Geophys. Res. Lett., 39, n/a-n/a, 10.1029/2012gl054270, 2012.

511 Markowska, M., Baker, A., Treble, P. C., Andersen, M. S., Hankin, S., Jex, C. N., Tadros, C. V., and Roach, R.:
512 Unsaturated zone hydrology and cave drip discharge water response: Implications for speleothem paleoclimate
513 record variability, J. Hydrol., 529, 662–675, http://dx.doi.org/10.1016/j.jhydrol.2014.12.044, 2015.

514 Pisani, P., Caporuscio, F., Carlino, L., and Rastelli, G.: Molecular Dynamics Simulations and Classical
515 Multidimensional Scaling Unveil New Metastable States in the Conformational Landscape of CDK2, PLoS
516 One, 11, 1-22, 10.1371/journal.pone.0154066, 2016.

517 Scheidt, C., and Caers, J.: Representing spatial uncertainty using distances and kernels, Math. Geosci., 41, 397-
518 419, 2009.

519 Smart, P. L., and Friederich, H.: Water movement and storage in the unsaturated zone of a maturely karstified
520 carbonate aquifer, Proceedings of the conference on Environmenatl Problems in Karst Terranes and their
521 Solutions, Dublin, Ohio, 1987, 59-87,

522 Smith, A. J., Massuel, S., and Pollock, D. W.: Geohydrology of the Tamala Limestone Formation in the Perth
523 Region: Origin and Role of Secondary Porosity, 63, 2012.

524 Treble, P. C., Bradley, C., Wood, A., Baker, A., Jex, C. N., Fairchild, I. J., Gagan, M. K., Cowley, J., and
525 Azcurra, C.: An isotopic and modelling study of flow paths and storage in Quaternary calcarenite, SW Australia:
526 implications for speleothem paleoclimate records, Quaternary Science Reviews, 64, 90-103,
527 http://dx.doi.org/10.1016/j.quascirev.2012.12.015, 2013.

528 Treble, P. C., Fairchild, I. J., Griffiths, A., Baker, A., Meredith, K. T., Wood, A., and McGuire, E.: Impacts of
529 cave air ventilation and in-cave prior calcite precipitation on Golgotha Cave dripwater chemistry, southwest
530 Australia, Quaternary Science Reviews, 127, 61–72, http://dx.doi.org/10.1016/j.quascirev.2015.06.001, 2015.

531 Treble, P. C., Fairchild, I. J., Baker, A., Meredith, K. T., Andersen, M. S., Salmon, S. U., Bradley, C., Wynn, P.
532 M., Hankin, S. I., Wood, A., and McGuire, E.: Roles of forest bioproductivity, transpiration and fire in a nine-
533 year record of cave dripwater chemistry from southwest Australia, Geochim. Cosmochim. Acta, 184, 132-150,
534 http://dx.doi.org/10.1016/j.gca.2016.04.017, 2016.

535 Williams, P. W.: The role of the subcutaneous zone in karst hydrology, J. Hydrol., 61, 45-67,
536 http://dx.doi.org/10.1016/0022-1694(83)90234-2, 1983.

537