1 **Hydrological characterization of cave drip waters in a porous limestone: Golgotha**

2 **Cave, Western Australia**

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

4 [1]Hawkesbury Institute for the Environment, Western Sydney University, Australia
5 [2]Institute of Earth Surface Dynamics, University of Lausanne, Switzerland
6 [3]Connected Waters Initiative Research Centre, UNSW Australia, NSW, Australia
7 [4]Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW, Australia
8
9 *Correspondence to*: Kashif Mahmud (k.mahmud@westernsydney.edu.au)

11 **Abstract**

12 Cave drip water response to surface meteorological conditions is complex due to the heterogeneity of water

13 movement in the karst unsaturated zone. Previous studies have focused on the monitoring of fractured rock

14 limestones that have little or no primary porosity. In this study, we aim to further understand infiltration water

15 hydrology in the Tamala Limestone of SW Australia, which is Quaternary aeolianite with primary porosity. We

16 build on our previous studies of the Golgotha Cave system and utilize the existing spatial survey of 29

17 automated cave drip loggers and a LiDAR-based flow classification scheme, conducted in the two main

18 chambers of this cave. We find that a daily sampling frequency at our cave site optimizes the capture of drip

19 variability with least possible sampling artifacts. With the optimum sampling frequency, most of the drip sites

20 show persistent autocorrelation for at least a month, typically much longer, indicating ample storage of water

21 feeding all stalactites investigated. Drip discharge histograms are highly variable, showing sometimes

22 multimodal distributions. Histogram skewness is shown to relate to the wetter than average 2013 hydrological

23 year and modality is affected by seasonality. The hydrological classification scheme with respect to mean

24 discharge and the flow variation, can distinguish between groundwater flow types in limestones with primary

25 porosity, and the technique could be used to characterize different karst flow paths when high-frequency

26 automated drip logger data are available. We observe little difference in the coefficient of variation (COV)

27 between flow classification types, probably reflecting the ample storage due to the dominance of primary

28 porosity at this cave site. Moreover, we do not find any relationship between drip variability and discharge

29 within similar flow type. Finally, a combination of multi-dimensional scaling (MDS) and clustering by k-means

30 is used to classify similar drip types based on time series analysis. This clustering reveals four unique drip

31 regimes which agree with previous flow type classification for this site. It highlights a spatial homogeneity in

32 drip types in one cave chamber, and spatial heterogeneity in the other, which is in concordance with our

33 understanding of cave chamber morphology and lithology.

34

35 **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

74 primary porosity. This classification scheme is also compared with previous studies (Smart and Friederich,
75 1987; Baker et al., 1997) to examine the limitations of these previous schemes. These findings will also help
76 better characterize and understand water movement in highly porous karst formations.

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

2 Site Description

2.1 Studied Cave

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

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

110

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

2.2 Climate and Meteorology

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

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

135 **2.3   Drip data acquisition and characteristics**

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

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

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

164   spring/early summer (Oct-Dec) while Chamber 2 drips tend to peak a few months later (Dec-May), reflecting a

165   longer water residence time. This may be a function of the thicker ceiling above Chamber 2 (40.24 versus 32.33

166   m) but also heterogeneity in flow paths to each chamber. Overall the drip response to the 2013 wet winter is

167   amplified in Chamber 2 versus Chamber 1, consistent with the presence of greater fracture flow in Chamber 2.

168   By applying morphological analysis of ceiling features acquired by LiDAR data, Mahmud et al. (2015)

169   distinguished three flow patterns (i.e. matrix flow, fracture flow, and a combination of conduit, fracture, and

170   matrix flow) for the observed ceiling morphological features. All the drip sites were then characterized

171   according to this flow classification in Mahmud et al. (2016), which is used here as a reference for clustering

172   similar drip time series.

173   **3   Methods**

174   **3.1   Hydrological classification of cave drips**

175   Research involving automated drip monitoring systems is increasing, for example at Cathedral Cave in

176   Wellington (Cuthbert et al., 2014) and Harrie Wood Cave in the Snowy Mountains, Yarrangobilly (Markowska

177   et al., 2015). The variability of the drip discharge might not only be a function of discharge itself, but could also

178   depend on the sampling frequency. We investigate this possibility by plotting the COV versus sampling interval

179   (the original 15 mins and calculated by resampling the data at 1 hour, 1 day, 1 week and 1 month). COV is

180   supposed to be artificially high at the high frequency of 15 mins because of sampling bias that artificially

181   increases the noise. The resampling at low frequencies is only a way of smoothing out this noise. Using the

182   optimum sampling frequency to minimize its effect on drip variability, we plot drip rate histograms to identify

183   the response of drips between the flow classifications and the response to intra and inter-annual variability in

184   infiltration. We also plot the autocorrelation functions (ACFs) to investigate the relationship between the

185   strength of correlation and the LiDAR-based flow type. Finally, we summarize the mean discharge of drip sites

186   in relation to the variability in discharge using the optimum sampling frequency. These are the same drip

187   discharge parameters as used in the classification method proposed by Friederich and Smart (1982), Fairchild et

188   al. (2006) and Baker et al. (1997) that were based on manual drip collection at low frequency.

189   **3.2   Clustering of similar drip time series**

190   We employed multi-dimensional scaling (MDS), which allows data dimensionality reduction i.e., mapping

191   complex multidimensional data on a low-dimensional manifold. MDS is a technique that embeds a set of points

192   in a low-dimensional space, so that the distances between the points resemble as closely as possible a given set

193   of dissimilarities between the objects they represent (Birchfield and Subramanya, 2005). MDS requires a

194   distance matrix to be computed, in which a single scalar number characterizes the similarity between any two

195   time-series. In our case, each drip logger is an object and a specific distance between drip loggers is considered

196   to characterize the similarity between any two loggers. It takes an input matrix giving dissimilarities between

197   pairs of items and outputs a coordinate matrix whose configuration minimizes a loss function. MDS is also

198   known as Principal Coordinates Analysis (PCoA). MDS operates on a distance or dissimilarity matrix (Pisani et

199   al., 2016), which is different than principle-component analysis (PCA) that is based on a covariance matrix.

200 Even if PCA and MDS methods can return the same results in specific contexts, MDS can be considered more
201 general because it remains validity for non-euclidean distances, such as the distance matrix (*d*) chosen in this
202 study. MDS is used to translate these distances into a configuration of points defined in an *n*-dimensional
203 Euclidean space (Cox and Cox, 1994). A MDS results in a set of points arranged so that their corresponding
204 Euclidean distances indicate the dissimilarities of the time series. According to Birchfield and Subramanya
205 (2005) the basic steps of performing the MDS algorithm are:

206 i) Construct the distance matrix **D**: One key component in clustering is the function used to measure the
207   temporal similarity (or distance) between any two time-series being compared. To define an appropriate
208   measure of similarity between time series, we determine two factors: firstly, the offset (*O*) to match two
209   time-series based on their maximum correlation, and secondly the complement of the correlation coefficient
210   (1-*R*) between the time series (Jex et al. 2012). Initially, we compute the cross-correlation function, a
211   measure of similarity of two time-series as a function of the displacement of one relative to the other. The
212   cross-correlation function is an estimate of the covariance between two time-series, y_{1t} and y_{2t}, at lags $k = 0$,
213   ±1, ±2, ... The offset (*O*) is defined as the lag time based on the maximum correlation between two time-
214   series. Next, we define R as the correlation coefficient with the time series being moved by the offset
215   amount *O* to have maximum correlation coefficient. Both *O* and *R* are calculated to all *n(n-1)/2* pairs of drip
216   data, where *n* is the number of drip data. Here, we use the original recorded drip counts in 15 mins interval.
217   The sampling bias discussed in section 3.1 is only affecting the drip variability, not the cluster analysis.
218   Moreover, high resolution (15 mins interval) data are more suited for the cluster analysis because it allows
219   better defining the cross-correlation between drips, as sometimes the offset of maximum correlation *O*
220   might be less than a day. Finally, the distance matrix **D** is computed for each pair of loggers using the
221   following equation (Jex et al. 2012):

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

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

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

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

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

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

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.

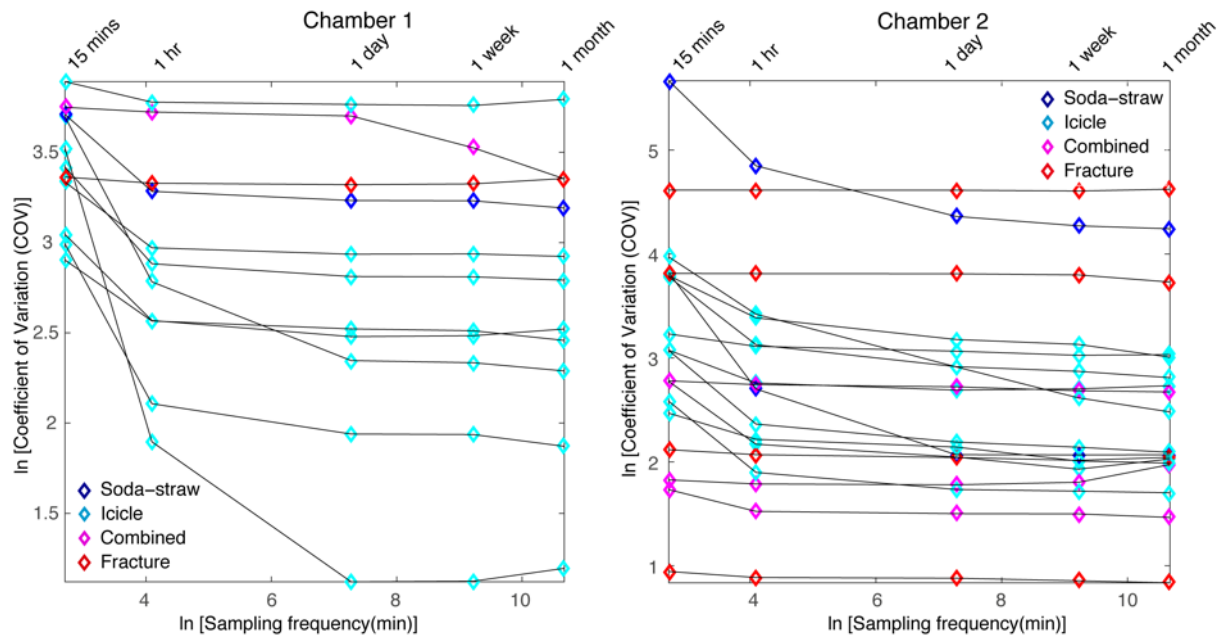
> ii) Compute point-to-cluster-centroid distances of all observations to each centroid. There are two steps to follow: first assign each observation to the cluster with the closest centroid. Then individually assign observations to a different centroid if the reassignment decreases the sum of the within-cluster,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

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

4.2 Drip rate frequency distributions

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

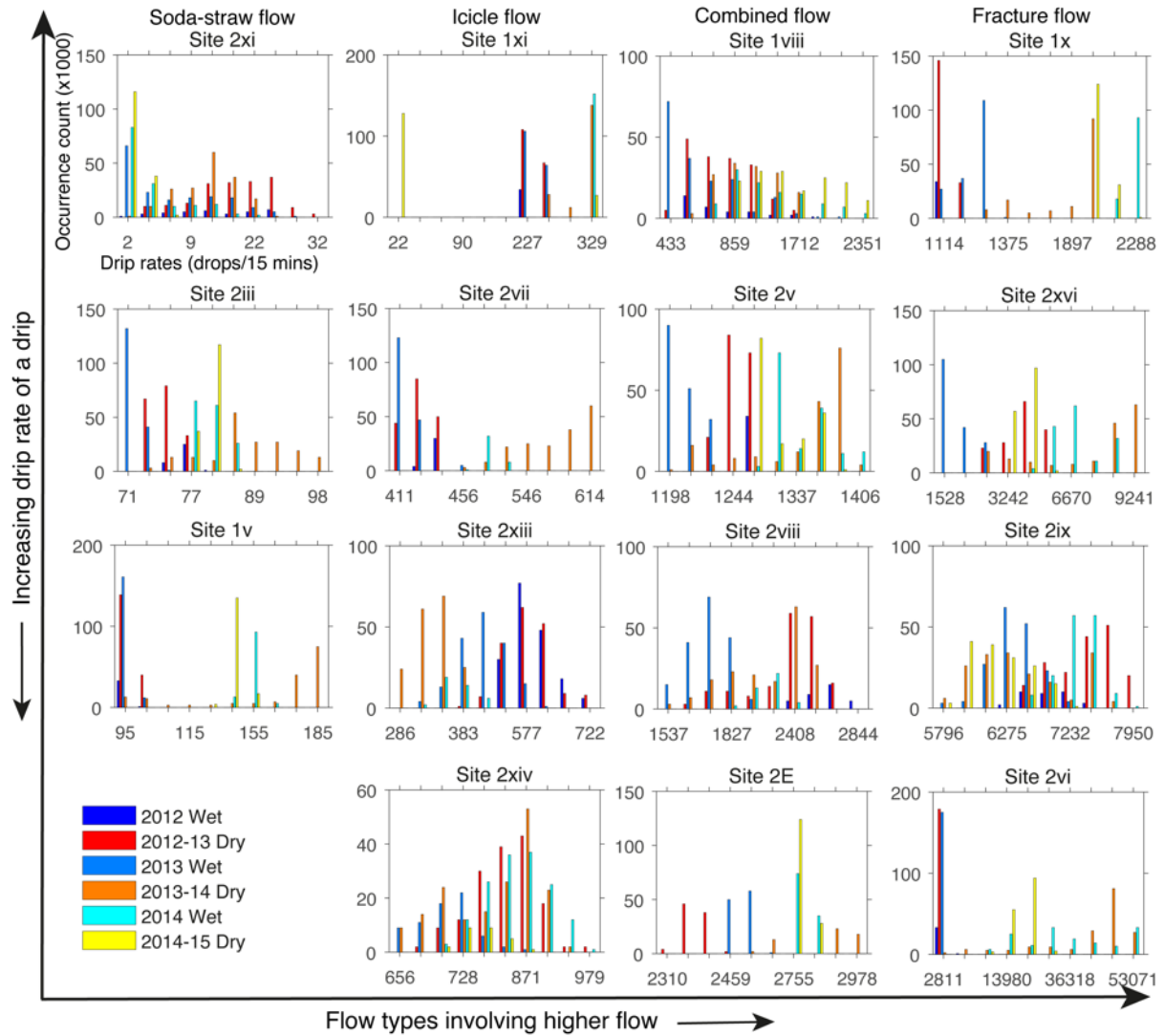



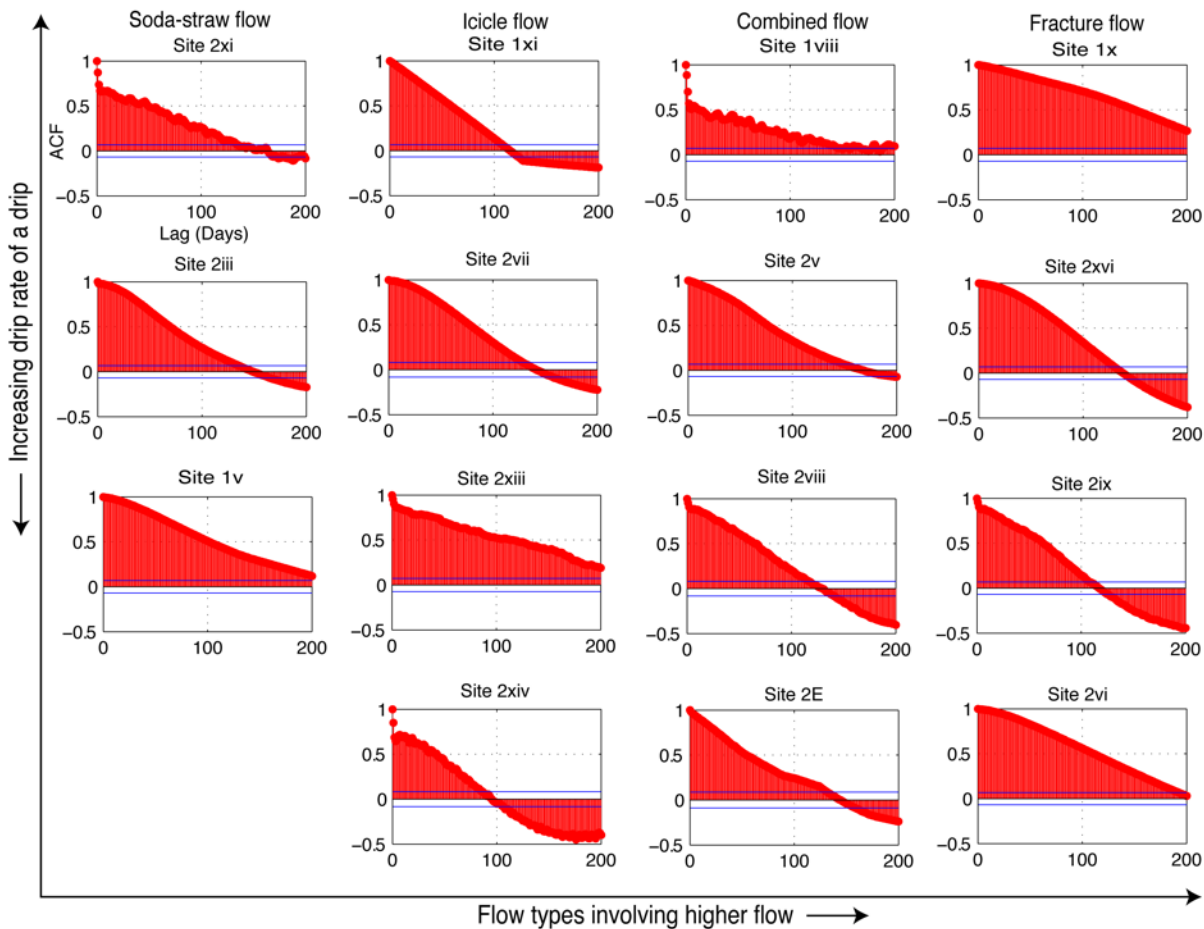
Figure 3: Histogram plots of both chambers drip data according to four flow types identified in Mahmud et al. (2016). Each histogram represents the frequencies of the drip counts per day (The axes labels are shown in the first histogram). Bin size is uniform for all plots and the external tick marks 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

297   the correlation and the LiDAR-based flow classification. This indicates the presence of ample storage in the

298   system, supplying all stalactite types.

299

300   Figure 4: Autocorrelation functions of both chambers drip data according to flow classification of Mahmud et al.

301   (2016). X- and Y-axis of individual plots represents the lag (in days) and ACF respectively (The axes labels are

302   shown in the first ACF plot). ACFs for all sites appear in Supplementary Figure S2.

303   4.4   Hydrological classification of cave drips

304   We examine the hydrological behavior of the drips at daily resolution with respect to mean discharge and flow

305   variation in Figure 5. It is clear from Figure 5 that there is no relationship between COV and flow-type. One

306   soda-straw discharge (site 2xi) has a seasonal dryness, a very low discharge, and a very high coefficient of

307   variation due to its irregular dripping. Otherwise, nearly all soda-straw flow, icicle flow, combined flow and

308   fracture flow drips have COV <60%, with the exception of one fracture flow site showing the highest COV

309   (Figure 5). But in general, there is little difference in the COV between classification types, probably reflecting

310   the ample storage (supplementary section S1) due to the dominance of primary porosity at this cave. We do not

311   clearly observe increasing variability with decreasing discharge within similar flow type, in contrast to other

312   studies from older, fractured rock limestones (Smart and Friederich, 1987; Baldini et al., 2006; Baker et al.,

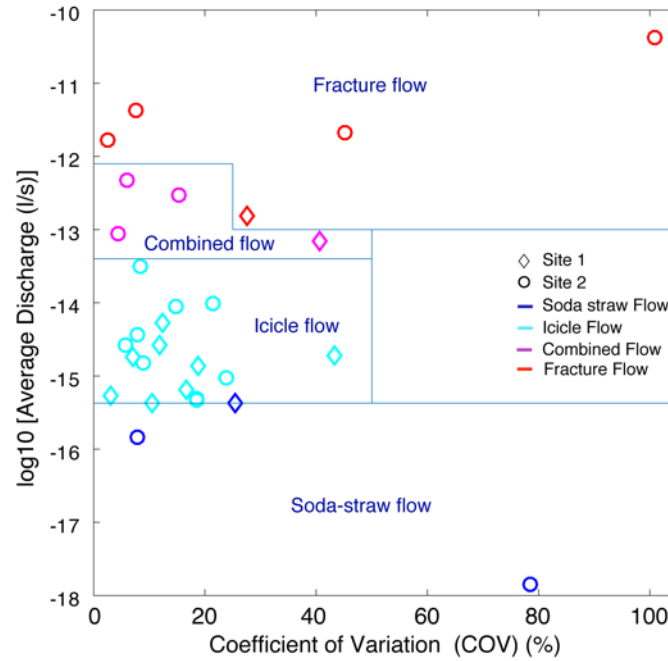
313   1997). This shows that Golgotha Cave drip sites do not fit within the drip classification method proposed by

314   Smart and Friederich (1987) and Baker et al. (1997), which were based on manual drip counts with limited

315   number of intermittent drip sites. Moreover, we utilize drip data from a cave with primary porosity, capturing

316   the full range of flow types from matrix through to fracture, whereas the previous classifications only captured

317   slow vs fast drips that were likely dominated by fracture flow paths given the host rock setting.

318

319   Figure 5: Hydrological behaviour of drip sites expressed in terms of daily mean discharge versus daily discharge

320   variability calculated from the automatic drip rate data for three hydrological years. Measured drip rates are

321   converted to volume units assuming a drip volume of 0.1433 ml (Genty and Deflandre, 1998). Blue lines and

322   symbols reflect flow classification given in Mahmud et al. (2015).

323   **4.5   Clustering of similar drip time series**

324   The clustering results are overlain upon the chamber ceiling images in Figure 6 and also summarized in Tables 1

325   and 2 with the average drip discharges and flow type classification based on LiDAR. Average drip discharges

326   are calculated from the 15-minute drip rates. As mentioned above, drip logger time series are deemed similar if

327   they are well correlated and only have a small offset with each other, and so these time series should cluster

328   together. Most of the drip sites that are identified as matrix flow (soda-straw and icicle flow) cluster together in

329   C1. However, three of the icicle flow sites with drip rate greater than 4 per 15 minutes fall in C2. The combined

330   flow category and the fracture type usually cluster in C3 and C4 respectively. Therefore we observe that our

331   clustering generally agrees with the morphology-based flow classification of Mahmud et al. (2016). Few of the

332   flow classes show exceptions, for example site 2vi is a fracture type flow and cluster in C1. This site has really

333   high discharge with high variability, showing irregular drip rate.

334   One consistent feature that appears from the cluster analysis of Figure 6 is the spatial homogeneity of the

335   clusters in Chamber 1, suggesting that they are spatially connected, or that their flow paths are connected to the

336   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

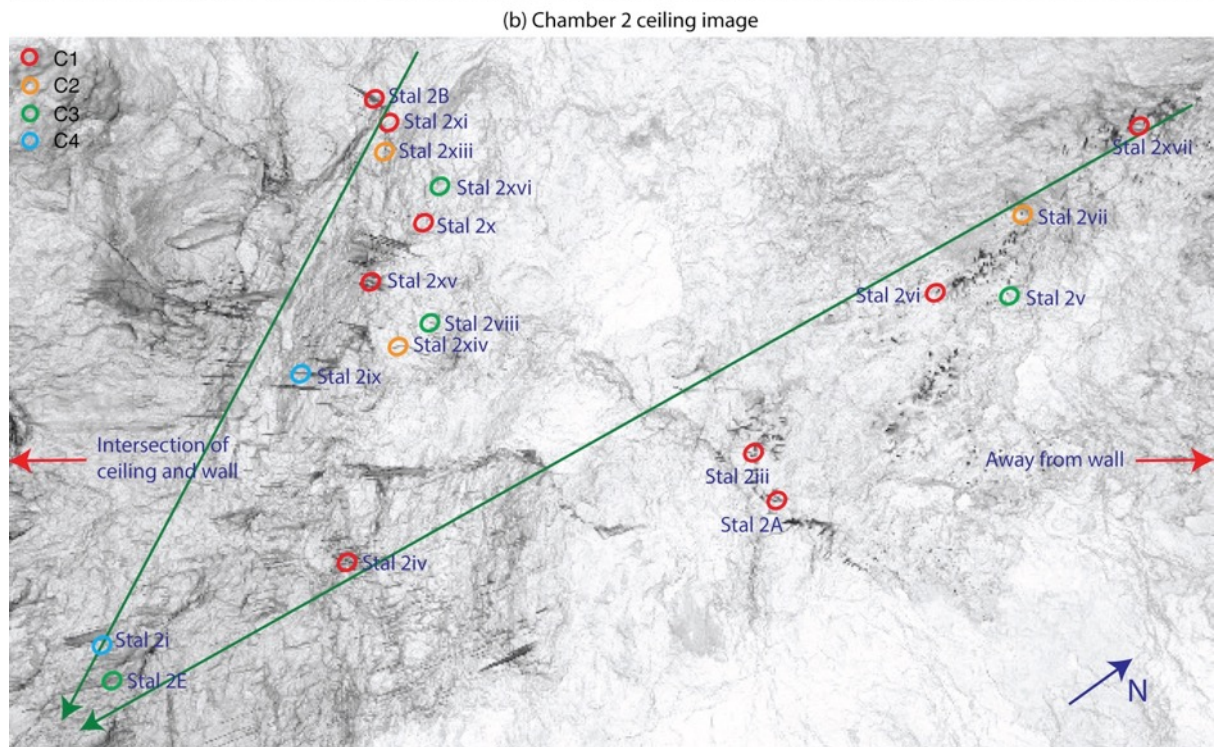
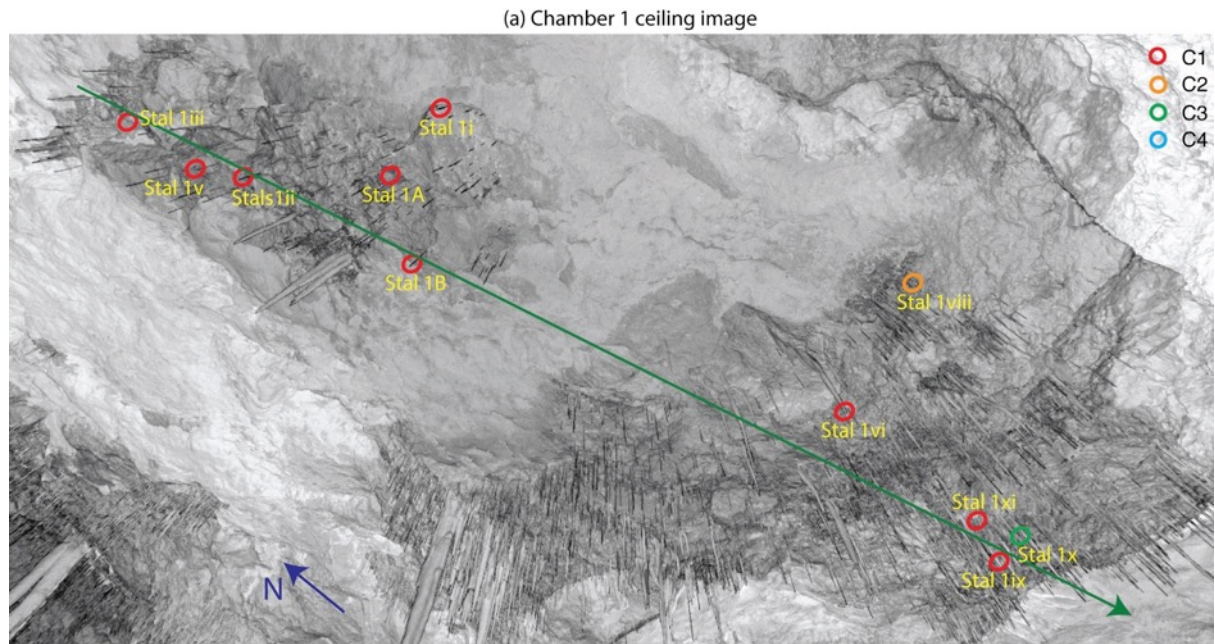



Figure 6: Cluster group plot overlain upon the cave ceiling for both chambers. The ceiling images are captured by LiDAR and the circles represent the ceiling locations of stalactites dripping on various stalagmates in both chambers (shown in Figure 1). The colour of the circles indicates individual MDS 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:

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

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

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

6   Conclusion

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

415 Finally, we apply a well-developed clustering method to determine the degree of similarity between drip time
416 series. The clustering indicates one dominating group: C1 (characterized by matrix flow type) with very slow
417 continuous drip discharge indicating matrix porosity in the thick limestone formation. This finding concurs with
418 the observed cave chamber morphology and lithology. Moreover, the cluster analysis agrees with the flow
419 classification of Mahmud et al. (2016) by grouping similar flow type in one single cluster. Overall this study
420 establishes a novel way to characterize cave hydrology, which can be obtained by performing together both
421 methodologies of Mahmud et al. (2015) and Jex et al. (2012). It relies on a metric that defines drip logger time
422 series as similar if they are well correlated and only have a small offset with one another, and therefore these
423 time series should cluster together. The MDS analysis supports this hypothesis and moreover, displays the
424 spatial patterns of the flow paths between the surface and the cave chambers. This technique shows potential to
425 classify, quantify and visualise the observed relationships between infiltration through the fractured limestone
426 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