Luminescence dating

Luminescence techniques (both thermoluminescence (TL) and optically-stimulated luminescence (OSL)) measure when sediments were either last heated (e.g., burnt or fired, Aitken, 1998) or exposed to sunlight during transport or deposition (e.g., by wind or water; Huntley *et al*., 1998) via the stimulation of minerals using heat (TL) or light (OSL). They are trapped charge techniques whereby a light sensitive signal or energy is “trapped” in defects in the crystal lattice of minerals such as quartz and feldspar. This energy builds up slowly and is stored in the crystal during a period of burial in the sediment column. The amount of trapped energy is proportional to the amount of naturally occurring ionising radiation found in the surrounding sediment and the amount of exposure time during the burial period (Aitken, 1998). As the signal is reset to zero by sunlight – it represents a clock to measure the time since last exposure (Duller, 2004). This makes these techniques unique amongst dating methodologies as they can be used to estimate the time elapsed since sediments were buried (Stokes, 1999), and as most processes or events are based on the deposition of sediment, depositional ages represent an extremely desired commodity in paleontological research.

Sample collection

Two methods for sampling the fossil breccia deposits for luminescence dating were conducted at the cave sites depending on the profusion of clasts and density of the material. Where possible in-situ bulk samples were collected from a cleaned exposure of the excavation using red filtered torch light for illumination (Badak Cave C, Cluster 5, Cistern Cave and Villa Cave) and then stored in a light-safe container. But if the material was not conducive to this technique then a block sample was collected (Badak Cave C, Cluster 4 and Swamp Cave) using a hammer and chisel in similar red light conditions.

Additional samples of sediment were collected from within 30 cm of the sample and subsequently combined for water content and environmental radioactivity determinations. Quartz grains of 90–125 µm in diameter were separated from the matrix under dim red illumination using standard purification procedures (Aitken, 1998), including a final etch in 40% hydrofluoric acid for 45 min to remove the external alpha-dosed rinds. The block samples were rolled in weak (10%) HCl to remove the outer light exposed layers and the inner core was processed according to the methods described above. The acid-etched 90–125 µm quartz grains were mounted on stainless-steel discs using silicone oil spray as adhesive with a total area of either 7 mm (red TL) with each aliquot being composed of ~5000 grains or 0.5 mm (OSL) with each aliquot being composed of ~50 grains

Red TL dating

Very small amounts of quartz were extracted from the breccia samples so the techniques were carefully chosen to maximise the amount of data for minimal quartz requirements. The red TL DAP technique was selected as it was specifically designed to constrain sediments using only 2 aliquots of quartz (Westaway & Roberts, 2006). The techniques were originally applied to volcanic sediments and although these locations are not in a volcanic region per se, they are surrounded by volcanic regions (Sumatra in Indonesia and the Philippines) which may have influenced their sedimentary composition. This combined with the sparse amounts of quartz and the estimated antiquity of the material (the red TL signal saturates at higher doses than the OSL signal) made the red TL signal a suitable choice. Aliquots were analysed using a modified SAR protocol (Murray & Wintle, 2000) to isolate the light-sensitive red TL signal (Westaway & Roberts, 2006).

The dual-aliquot protocol (DAP) described in detail by Westaway & Roberts (2006) was originally designed to exploit the intense red TL emissions typical of quartz from a volcanic province (Hashimoto, 1986; 1987; Fattahi & Stokes, 2000; Scholefield & Prescott, 1999; Morwood *et al*., 2004; 2005). Using this technique, the time elapsed since the quartz grains were last exposed to sunlight was obtained from the light-sensitive (bleachable) component of the red TL signal. The procedure also determines the time elapsed since the unbleachable signal was last reset by a major heating event such as a volcanic eruption (or being located within a close proximity to heated products such as lava).

According to the local geology, it is likely the Malaysian quartz is not volcanic but of granitic origin, however, investigations into the spectral emissions of quartz from different geological settings have revealed that granitic (a form of plutonic) quartz also produces strong and reproducible red emissions (Scholefield *et al*., 1994; Scholefield & Prescott, 1999; Westaway, 2009). When the potential for applying the DAP technique to these non-volcanic (non-heated) sediments was explored using a granitic quartz from Australia (PJIES-75 from Purijarra rockshelter, Northern Territory) it was discovered that the red TL results were comparable with independent age estimates from OSL single-aliquot, OSL single-grain and radiocarbon techniques (Westaway, 2006). Furthermore, recent investigations into quartz emissions has revealed that strong red and blue emissions are a feature of granitic quartz (Westaway and Prescott, under review a) and that both emissions can be used to provide an internal check on the consistency of the D_e (Westaway and Prescott, under review b). Agreement between the OSL and red TL D_e values would signify that the results are valid. Therefore to accommodate the lack 'heating' history – aliquot A is described as the hard-to-bleach TL residual level that is only reset by very long bleaching periods (>1 week, which rarely occur in this cave environment) (Scholefield & Prescott, 1999), while its subtraction from the Aliquot B signal isolates the rapidly easy-to-bleach signal that is used to estimate the timing of sediment deposition.

Aliquots were heated to 260°C at a heating rate of 5 K s^{-1} and then held at 260°C for 1000 s to minimize the unwanted TL from incandescence. The use of these isothermal measurements rather than traditional glow curves measurements (incorporating heating up to 450°C) reduces the potential signal loss from thermal quenching (Spooner & Franklin, 2002) and the potential for temperature induced sensitivity changes (Murray & Wintle, 2000; Huot *et al*., 2006). D_e were estimated from the 20–30 s interval of isothermal decay (which was bleachable by >380 nm illumination), using the final 160 s as background. The isothermal red TL emissions were measured using a red sensitive photomultiplier tube (Electron Tubes Ltd 9658B) and cooling tower (LCT50 liquid-cooled thermoelectric housing) with Kopp 2-63 and BG-39 filter combination (Fattahi & Stokes, 2000), and laboratory irradiations were conducted using a calibrated ^{90}Sr/^{90}Y beta source.

The uncertainty on the D_e includes random errors (include photon-counting errors, errors associated with the variance in the background signal, instrument reproducibility, and errors associated with D_e determination derived by Monte Carlo simulation (Yoshida *et al*., 2003) systematic errors (including a 2% error for beta source calibration) and total errors (the quadratic sum of the random and systematic errors). The precision is much lower for red TL dating in comparison to OSL due to the use of dual aliquots. This technique requires the errors to be propagated through from the first to the second aliquot resulting in precisions of up to ~30 %. However, when compared to independent dating techniques the accuracy of the DAP technique is comparable (Westaway, 2009; Roberts *et al.,* 2009). Multiple dose recovery procedures were performed using the DAP technique and the surrogate dose was recovered to within 10-15% (Westaway, 2006; Westaway and Roberts 2006).

For these samples, the small amounts of quartz present precluded detailed dose recovery procedures, but both Aliquot A and Aliquot B doses could be recovered to within 0.8-12% using at least two heated (used) aliquots for the samples tested (M-VIL, M-SWA).

DAP requires a minimum of two aliquots to determine the bleachable De. To minimise any complications arising from differences in material properties between Aliquots A and B and to improve the signal-to-noise ratio, we used large aliquots composed of many grains. However, not all of these grains are likely to have been fully bleached when last transported, therefore the bleachable TL ages should be viewed as maximum ages for sediment deposition (Westaway & Roberts, 2006). Tests on modern analogue material collected from outside the cave directly above the sinkholes (in Cistern and Swamp Caves in the Batu complex), which is thought to represent the source are for the breccia sediments confirms this assumption. The hard-to-bleach signal retains its D_e value as would be expected, while the easy-to-bleach signal has not been completely reset despite some sunlight exposure, thereby the remaining red TL signal creates a residual value and results in maximum ages. In an effort to improve the confidence in these maximum ages a few multiple estimates of the easy-to-bleach D_e were conducted to test whether the results were reproducible and to assess the distribution of D_e between aliquots. These estimates could only be conducted on samples that contained slightly larger amounts of quartz, e.g., M-SWA and M-VIL, and are presented as separate D_e values, and indicate that despite representing maximum ages there is only a small variation in the range of D_e values between aliquots (Table A3). If the aliquot containing the lowest D_e values is assumed to represent the grains that have experienced the longest bleaching period and is used as a surrogate ‘minimum age’ the D_e values for both samples reduce the maximum age between

7-16 ka, these estimations (symbol R_M) have been included in the final results (Table A2).

If the actual minimum age model (Roberts *et al.*, 1998) is applied to these very small data sets then D_e values of 160 ± 27 and 85 ± 11 are obtained for Villa and Swamp Caves, respectively, which are both very close to the original red TL D_e estimates.

Table A3. Red TL multiple estimate results

Sample	**Original D_e**	**Number of multiple estimations**	**Range of D_e values (Gy)**	**Standard deviation from original**
M-VIL	159 ± 44	6	114-215	35 ± 3
M-SWA	85 ± 17	6	75-98	10 ± 4

OSL dating

For most samples the OSL signal was measured to improve the confidence in the final age estimate, to provide an internal check on D_e consistency and to determine the likely extent of any TL age overestimation arising from incomplete bleaching of quartz grains at the time of sediment deposition. Initial investigations revealed that some of the samples contained saturated OSL signals (M-BADA 4 and 5), while the small amounts of quartz prevented the use of many aliquots for multiple OSL measurements, most of the samples contained very little or none of the larger size fraction (180-212 µm), which prevented single-grain analysis. The remaining samples (M-SWA, M-VIL) were processed using a modified single-aliquot regenerative-dose protocol (Murray & Wintle, 2000; Olley *et al.*, 2004, 2006). Acid-etched 90-125 µm grains were loaded onto stainless steel discs (48) using a 0.5 mm mask and the ultraviolet emissions were detected by an Electron Tubes Ltd 9235QA photomultiplier tube fitted with 7.5 mm of Hoya U-340 filter. A calibrated ^{90}Sr/^{90}Y source was used to perform laboratory irradiations. Quartz grains were preheated to 240°C for 10 s (and a cut heat of 160°C after the test dose) and then heated to 125°C and optically stimulated using blue diodes for 100 s, followed by a hot optical wash at 260°C for 200 s to eliminate the recuperation effect. The most light-sensitive component of the OSL signal was obtained from the initial 3.2 s of stimulation (1-8 channels), and the background count rate was estimated from the last 20 s (200-250 channels). As tests of protocol performance, the regenerated OSL dose-response curve included a zero-dose point to measure the extent of thermal transfer and a duplicate regenerative dose point to examine the adequacy of the test dose sensitivity-correction procedure.

Dose recovery experiments were conducted on fresh aliquots to select appropriate procedures and on used aliquots to test the procedures employed. The natural signal was removed from the fresh aliquots using minimal heat treatments so as not to cause any sensitivity change between the surrogate natural and first regenerative cycles. Fresh aliquots were bleached (using blue diodes at 50% power for 200 s at 50°C) to erase the natural signal and in general, the conditions that recovered a dose closest to the given dose, and that produced an acceptable dose-response curve, were taken to be most appropriate for that sample, and indicated that the luminescence signal could be measured reproducibly under controlled experimental conditions. A beta dose similar to the estimated natural (e.g., 100 Gy) was administered to represent a surrogate natural dose, and an attempt was made to recover this value using a modified SAR protocol. For each aliquot, a different preheat temperature was employed starting at 200°C for 10 s and increasing in 10°C increments up to 270°C all with a 160°C cut heat after each test dose cycle to determine which preheat conditions recovered a dose closest to the given dose. These results formed the basis of a preheat plateau test.

It was determined that the 240°C preheat for 10 s (with a 160°C cut heat for the test dose cycle) recovered a dose closest to the surrogate natural as was located on the 'preheat plateau' so this was chosen as the most suitable preheat conditions. These tests also indicated that a hot optical wash treatment after every regenerative cycle effectively removed a thermally transferred component that potentially causes a recuperated signal >5%. Dose recovery tests were also performed on used discs to test the procedures employed and all recovered D_e values to within 5-6% of the surrogate value. The UV OSL signal was also tested in the modern analogue samples (details above) and it was determined that the degree of resetting is much higher than the red TL signal with most of the aliquots returning a zero age.

Single-aliquots were accepted and rejected according to the criteria of Jacobs *et al*. (2006). Using this criteria aliquots were rejected if 1) the signal was greater than 3σ above the background value, 2) the recycling ratios were greater than 10% either side of unity, 3) there was a greater than 2σ difference between the IR value (run as a final regenerative run) and the signal, 4) aliquots exhibited supersaturing behaviour, 5) aliquots were dominated by an ultrafast component, 6) aliquots displayed recuperation exceeding 5%, 7) excessive increasing or decreasing sensitivity changes in the test dose responses in relation to the natural, 8) the presence of a desensistising slow component which displays a difference between the natural and first regenerative decay curve, and 9) the dominance of a medium component in the shape of the natural shine down (Table A4).

Table A4. OSL Single-aliquot rejections

Sample	**Single-aliquots processed**	**Single-aliquots rejected**	**Single-aliquots accepted**	**Overdispersion (%)**	**Statistical model used**
M-VIL	12	0	12	18	CAM
M-SWA	12	2 (Recycling ratio)	10	23.5	MAM

Statistical models

For each sample, D_e values were estimated for at least ~10 out of the 12 aliquots processed. The statistical overdispersion of the resulting D_e was ~20% for both samples (18, 23%) but as the D_e values were quite large and the overdispersion was very close to the accepted value for the central age model (CAM) (Galbraith *et al*., 1999) it would have been useful to conduct further investigations using single-grains. However, this was prevented by the absence of the larger grain size fraction (180-212 µm) and the small amounts of quartz recovered. Instead, to accommodate the possibility that not all of the grains on the small aliquots were fully bleached in the Swamp Cave environment, a minimum age model (MAM) (Roberts *et al*., 1998) was applied to the single-aliquot data. This interpretation agrees with the structure of the cave (small passages with fluvial activity) and the source area for the sediments (via sinkholes potentially from another connected cave system). In contrast, the source area for sediments for Villa cave is from the wide front entrance, and this source offers a good potential for the bleaching of the sediments. Indeed, a modern sample collected in the entrance returned a zero age (0.6 Gy).

This result combined with the low overdispersion and small range of palaeodoses provides ample justification for the use of the CAM (Galbraith *et al*., 1999). This model is only suitable for samples that are well bleached, with an overdispersion of less than 20%, and is consistent with single-component D_e distributions (Olley *et al*., 2004).

Dose rate

Concentrations of ^{238}U, ^{235}U, ^{232}Th (and their decay products) and ^{40}K were measured by high-resolution gamma spectrometry of dried and powdered sediment samples in the laboratory and a portable gamma spectrometer in the field. Gamma and beta dose rates were both corrected for the measured (field) water contents, allowing for beta-dose attenuation (Mejdahl, 1979) and calculated from radionuclide concentrations using the dose-correction factors of Adamiec and Aitken (1998). Cosmic-ray dose rates were estimated from published relationships (Prescott & Hutton, 1994), making allowance for the density and between 10-20m thickness of limestone roof above the cave deposit, the geometry of the limestone shielding, the sediment overburden at the sample locality (~1-3 m), and the altitude (~20 m above sea level) and geomagnetic latitude and longitude (3° and 101°) of the sampling site. An effective internal alpha dose rate of 0.03 Gy ka^{-1}was assumed for each sample and included in the total dose rate.

Disequilibrium in the cave environment

High resolution gamma spectrometry of the cave sediments is particularly useful as it provides information on the entire ^{238}U and ^{232}Th decay chains, therefore the concentrations of the daughter as well as the parent nuclides can be explored. This information allows as informed assessment of the state of equilibrium or disequilibrium in this sedimentary environment. As these decay chains contribute to 26% and 27% of the total dose rate (Olley *et al*., 1996), disequilibrium in one or both of these chains will have a bearing on the estimated dose rate and age, therefore it is important that any excesses or deficiencies in daughter nuclides relative to the parent nuclides are identified and assessed. In these cave sedimentary environment, three potential disequilibrium states have been identified; 1) in the ^{238}U chain a small deficiency of between 0.6-9.3% and an excess of 5.7% in ^{226}Ra compared to ^{238}U, 2) a ^{210}Pb excess or deficiency of between ~10% compared to ^{226}Ra, and 3) in the ^{232}Th chain a small excesses of between ~7% in ^{228}Th compared to ^{228}Ra. Thorium is thought to be immobile in most sediments (Ivanovitch and Harmon, 1992) and this combined with the short half lives (τ) of the daughter nuclides (^{228}Ra = 5.7 yrs and ^{228}Th = 1.91 yrs) means that this effect is negligible in samples older than 20 years (Olley *et al*., 1996). The excess of ^{210}Pb is probably due to the fallout of lead on surface sediments that are then washed into the cave, but similarly the τ is only 22 years and is unlikely to have any effect on sediments older than 100 years (Olley *et al*., 1996). The deficiency or excess in ^{226}Ra could potentially be the most problematic as ^{238}U is more mobile in the sediment column (Ivanovitch & Harmon, 1992), it contributes 27% of the total dose rate, and ^{226}Ra has a much longer half life (1602 years) so any disequilibrium in this chain will affect the accuracy of the final age. These excesses and deficiencies indicate that radium redistribution is occurring in the sediment column possibly due to the effects of fluctuating ground water levels and leaching in the cave environment, respectively. However, in a worst case scenario with a ^{228}Ra excess or deficiency of ~50% (the sample with the largest deficiency is M-BAD4 with only 9.3%, while the largest excess is M-VIL with only 5.7%) in a closed environment then the deviations from calculated to the true dose rate are only <2.5% and <3% respectively, and as ^{238}U contributes 27% of the total dose rate, <2.5% and <3% of the U chain equates to 0.675 and 0.81% of the total dose rate, respectively.

In these cave locations the excess or deficiency values are at least 5 times smaller than the worst case scenario, therefore the deviation in the true dose rate owing to the excess of ^{226}Ra is negligible (~0.162%) and thus does not need to be taken into account. In addition, if the sediment is buried within closed environment it will return to secular equilibrium within 30-40 ka, which is less than the estimated age of most of the samples. The sample collected from location 4 in Badak Cave C displays sediments close to secular equilibrium demonstrating that the problem of disequilibrium in this cave is not a significant problem.

Estimations of water content

Material collected from around the sampling areas was used for laboratory estimations of the present-day water contents. The material was placed in a (pre-weighed) beaker and weighed, to determine the wet sample weight. The sample was then dried in an oven at 40°C for several days and reweighed to determine the dry sample weight. By deducting the weight of the dry sediment and the empty beaker from the original wet sample weight produced a value for the total weight of water loss during drying. The *in situ* water content was then expressed as the weight of water divided by the weight of the dry sediment. These measurements contributed to the assessment of water content variability during the period of burial. In addition, field observations and assessments of past variability were conducted to assist in the final estimation of the annual dose. The total dose rate was calculated using long-term water contents of between 2 ± 0.2 % and 6 ± 2 %, which is close to the measured (field) water content of between 0.7-5.9 %. The only cave site that contained a higher water content was Villa Cave. The sediments and the cave were very wet during the sampling period providing high field water contents of ~80%, and required further detailed investigations of the site hydrology.

We explored the hydrology of Villa Cave as a means of understanding the recorded high water content values and discovered that although the site itself was not waterlogged the sediments (especially the basal sediments) had high water contents. Villa Cave shows evidence of water activity e.g., active slopewash and saturated clay deposits. But as the samples were collected during the wet season, it seems reasonable to suppose that the measured water contents represent a value above the annual average. Any present-day assessment of water content must, therefore, represent a maximum value of the long-term average. In comparison to the other cave sites, the sampling area in Villa Cave retains most of its sedimentary column, in Swamp, Badak Cave C and Cistern most of the original sediments have been removed allowing the breccia sediments to dry out over the long burial period. Therefore for these latter cave sites the measured field water contents represent minimum values of the long term water contents. For example Badak Cave C site 4 has a field water content of 0.7% but a long term water content of 5 ± 2%.

Water content variability was assessed according to location, sediment depth, and sediment composition. During present-day conditions, the western coast of Peninsular Malaysia usually has 5.5 wet months and 6.5 dry months, is roughly 55% of the year ‘wet’ with an average rainfall of 2500 mm a year. The composition of the sediment in Villa Cave has a bearing on its long-term water content variability; for example, sand tends to drain easily, whereas silt and clay absorb water so that the basal silts are in a state of constant saturation. This variability was considered when estimating the errors associated with the water content for each sample, to ensure that the errors covered the range of likely values. As a starting point surficial cave sediments would be expected to be ~20 ± 5 %; and basal cave sediments 30 ± 6 %. However, the estimated water contents of ~80% for Villa Cave seem a little high. We attribute these high values to the continuous slopewash processes that channel the surface run-off from the limestone into the cave during the rainy season.

This water tends to accumulate in the finer layers as the clays and silts preventing the water from seeping down through the layers. However, we believe that this value is not a true reflection of the long term water content for the cave and has been reduced down to a more conservative 50%.

The TL and OSL ages were calculated by dividing the TL and OSL palaeodoses by the corresponding total dose rate (Table A2). Some representative data are shown in Figure A1.

U-series techniques

Samples were collected from the overlying flowstones using a hammer and chisel and were used to establish the time of initiation of flowstone precipitation on top of the breccia. Thermal ionisation mass spectrometry (TIMS) uranium-series measurements were made using analytical procedures described elsewhere (Zhao *et al*., 2001, 2009b), whereas the multi-collector inductively coupled plasma mass spectrometry (MC-ICP-MS) analyses made using analytical procedures described in Zhou *et al*. (2011). The ^{230}Th/^{234}U ages (Table A1) were calculated using Isoplot EX 3.0 of Ludwig (2003) and decay constants of Cheng *et al*. (2000).

It is worthwhile to note that the two flowstone samples from Badak Cave C in Perak were measured in duplicate by both TIMS and MC-ICP-MS. Although the two samples are pure crystalline from appearance, U/Th isotopic analyses using both methods show they contain unusually high-level Th, with Th/U ratios of 0.75-1.5. We suspect that they are of siderite rather than calcite in composition. Nevertheless, the four duplicate analyses define a good (^{238}U/^{232}Th)-(^{230}Th/^{232}Th) isochron, corresponding to an initial (^{230}Th/^{232}Th) ratio of nearly zero (implying the presence of initial ^{230}Th is negligible) with a slope of 1.02±0.12 (which is the initial ^{230}Th-corrected ^{230}Th/^{238}U activity ratio for the two samples), confirming both samples are analytically indistinguishable from secular equilibrium conditions (i.e. activity ratio = 1). The TIMS measurements for $^{234}U/^{238}U$ appear to be a bit low, which might be due to imperfect peak centring, as U isotopic signals were too low due to poor recovery from column chemistry (large sample size with Fe co-precipitation may not work well for samples of siderite composition). In this regard, we consider that the new MC-ICP-MS measurements are more precise and reliable due to much stronger signals and the ability to perform mass fractionation corrections for U and Th together