OCTOPUS: An open cosmogenic isotope and luminescence database

Alexandru Tiberiu Codilean  
*University of Wollongong, codilean@uow.edu.au*

Henry Munack  
*University of Wollongong, hmunack@uow.edu.au*

Tim J. Cohen  
*University of Wollongong, tcohen@uow.edu.au*

Wanchese Saktura  
*University of Wollongong, wms994@uowmail.edu.au*

Andrew G. Gray  
*University of Wollongong, agray@uow.edu.au*

*See next page for additional authors*

**Publication Details**

OCTOPUS: An open cosmogenic isotope and luminescence database

Abstract
We present a database of cosmogenic radionuclide and luminescence measurements in fluvial sediment. With support from the Australian National Data Service (ANDS) we have built infrastructure for hosting and maintaining the data at the University of Wollongong and making this available to the research community via an Open Geospatial Consortium (OGC)-compliant web service. The cosmogenic radionuclide (CRN) part of the database consists of 10Be and 26Al measurements in modern fluvial sediment samples from across the globe, along with ancillary geospatial vector and raster layers, including sample site, basin outline, digital elevation model, gradient raster, flow-direction and flow-accumulation rasters, atmospheric pressure raster, and CRN production scaling and topographic shielding factor rasters. Sample metadata are comprehensive and include all necessary information for the recalculation of denudation rates using CAIRN, an open-source program for calculating basin-wide denudation rates from 10Be and 26Al data. Further all data have been recalculated and harmonised using the same program. The luminescence part of the database consists of thermoluminescence (TL) and optically stimulated luminescence (OSL) measurements in fluvial sediment samples from stratigraphic sections and sediment cores from across the Australian continent and includes ancillary vector and raster geospatial data. The database can be interrogated and downloaded via a custom-built web map service. More advanced interrogation and exporting to various data formats, including the ESRI Shapefile and Google Earth's KML, is also possible via the Web Feature Service (WFS) capability running on the OCTOPUS server. Use of open standards also ensures that data layers are visible to other OGC-compliant data-sharing services. OCTOPUS and its associated data curation framework provide the opportunity for researchers to reuse previously published but otherwise unusable CRN and luminescence data. This delivers the potential to harness old but valuable data that would otherwise be lost to the research community.

Publication Details

Authors
Alexandru Tiberiu Codilean, Henry Munack, Tim J. Cohen, Wanchese Saktura, Andrew G. Gray, and Simon M. Mudd

This journal article is available at Research Online: https://ro.uow.edu.au/smhpapers1/459
Abstract. We present a database of cosmogenic radionuclide and luminescence measurements in fluvial sediment. With support from the Australian National Data Service (ANDS) we have built infrastructure for hosting and maintaining the data at the University of Wollongong and making this available to the research community via an Open Geospatial Consortium (OGC)-compliant web service. The cosmogenic radionuclide (CRN) part of the database consists of $^{10}$Be and $^{26}$Al measurements in modern fluvial sediment samples from across the globe, along with ancillary geospatial vector and raster layers, including sample site, basin outline, digital elevation model, gradient raster, flow-direction and flow-accumulation rasters, atmospheric pressure raster, and CRN production scaling and topographic shielding factor rasters. Sample metadata are comprehensive and include all necessary information for the recalculation of denudation rates using CAIRN, an open-source program for calculating basin-wide denudation rates from $^{10}$Be and $^{26}$Al data. Further all data have been recalculated and harmonised using the same program. The luminescence part of the database consists of thermoluminescence (TL) and optically stimulated luminescence (OSL) measurements in fluvial sediment samples from stratigraphic sections and sediment cores from across the Australian continent and includes ancillary vector and raster geospatial data. The database can be interrogated and downloaded via a custom-built web map service. More advanced interrogation and exporting to various data formats, including the ESRI Shapefile and Google Earth’s KML, is also possible via the Web Feature Service (WFS) capability running on the OCTOPUS server. Use of open standards also ensures that data layers are visible to other OGC-compliant data-sharing services. OCTOPUS and its associated data curation framework provide the opportunity for researchers to reuse previously published but otherwise unusable CRN and luminescence data. This delivers the potential to harness old but valuable data that would otherwise be lost to the research community. OCTOPUS can be accessed at https://earth.uow.edu.au (last access: 28 November 2018). The individual data collections can also be accessed via the following DOIs: https://doi.org/10.4225/48/5a8367feac9b2 (CRN International), https://doi.org/10.4225/48/5a836cdfac9b5 (CRN Australia), and https://doi.org/10.4225/48/5a836db1ac9b6 (OSL & TL Australia).
1 Introduction

Cosmogenic radionuclide (CRN) exposure dating and luminescence dating are suites of geochronological techniques that have become important for the studying of Earth surface processes (e.g. Rhodes, 2011; Granger et al., 2013). Both permit quantifying the timing of geological events by dating individual landforms. In addition, CRNs can also be used to measure the rate at which landforms or landscapes are being denuded by physical and chemical erosion processes. Thus, the two suites of techniques have been extensively used among others to quantify basin-wide denudation rates (von Blanckenburg, 2005; Granger and Schaller, 2014), to reconstruct the extent of Quaternary glaciations (Spencer and Owen, 2004; Balco, 2011; Ivy-Ochs and Briner, 2014), to study how rivers have adapted to past climate change via incision and aggradation (Schaller et al., 2004; Lewis et al., 2009; Wallinga et al., 2010), and to study the timing of dune construction (Fitzsimmons et al., 2007; Fujioka et al., 2009; Bristow et al., 2010). Both suites of techniques are costly (both in terms of time and money) and require specialised training, laboratories, and equipment. As such, CRN and luminescence studies are often very focused and involve a relatively small number of samples (n < 100). The research questions being addressed by these studies are very specific and study areas are often relatively small. Hence, CRN and luminescence studies will produce small data sets that are unmanaged and that may become forgotten once the study has been completed and results are published. Further, despite there being calls for minimum data reporting standards (e.g. Dunai and Stuart, 2009; Frankel et al., 2010), the published work will often not include appropriate levels of metadata to make the raw data reusable with ease. The latter is especially important in the case of cosmogenic nuclides as procedures used to interpret CRN data are regularly revised and updated, requiring denudation rates and/or exposure ages to be recalculated using updated measurement standards and calculation protocols. Such recalculations are also necessary when comparing results produced by different accelerator mass spectrometry (AMS) facilities that happen to normalise results to different AMS standards. Therefore, without periodic recalulation and maintenance of the data, CRN-based age and rate estimates, for example, can become out of date after a few years. A system and framework for managing CRN and luminescence data and metadata are critical to ensuring the longevity and value of such data collections.

Here we present a database of cosmogenic radionuclide and luminescence measurements in fluvial sediment – OCTOPUS. With support from the Australian National Data Service (ANDS) we have built infrastructure for hosting and maintaining the data at the University of Wollongong and making this available to the research community via an Open Geospatial Consortium (OGC)-compliant web service (http://www.opengeospatial.org, last access: 28 November 2018). The CRN part of the database consists of $^{10}$Be and $^{26}$Al measurements in fluvial sediment samples from across the globe. Sample metadata are comprehensive and include all necessary information for the recalculation of denudation rates using CAIRN, an open-source program for calculating basin-wide denudation rates from $^{10}$Be and $^{26}$Al data (Mudd et al., 2016). To this end, the database also includes a comprehensive suite of geospatial data layers, both vector (e.g. sample site and basin outline) and raster (e.g. elevation, gradient and flow-routing rasters, atmospheric pressure, and CRN production scaling and topographic shielding factors). The luminescence part of the database consists of thermoluminescence (TL) and optically stimulated luminescence (OSL) measurements in fluvial sediment samples from stratigraphic sections and sediment cores from across the Australian continent. Comprehensive metadata and ancillary vector and raster geospatial data are likewise included and available for download. OCTOPUS can be accessed at https://earth.uow.edu.au (last access: 28 November 2018).

2 CRN and luminescence dating in a nutshell

This section briefly describes the two suites of dating techniques and provides information on CAIRN.

2.1 Inferring denudation rates from cosmogenic $^{10}$Be (and $^{26}$Al)

Cosmogenic nuclide exposure dating is based on the study of rare isotopes produced by high-energy cosmic radiation breaking up the atoms that make up the minerals and rocks at the Earth’s surface. The term “in situ” is used to distinguish these isotopes from those that are produced through the same cosmic-ray-induced nuclear reactions in the atmosphere – termed “meteoric” (Dunai, 2010; Granger et al., 2013). Several of the in situ cosmogenic nuclides, including the stable $^3$He and $^{21}$Ne, and the radioactive $^{10}$Be, $^{26}$Al, and $^{36}$Cl, are now routinely measured and have been used in geomorphological studies for the last three decades (Bierman and Nichols, 2004; von Blanckenburg, 2005; Dunai, 2010; Granger and Schaller, 2014). Of these nuclides, however, $^{10}$Be produced in quartz is the workhorse for in situ applications, and most in situ cosmogenic nuclide studies have used $^{10}$Be either alone or in conjunction with other cosmogenic nuclides such as $^{26}$Al and $^{21}$Ne. Given the long half-life of $^{10}$Be ($T_{1/2} = 1.387$ Myr, Chmeleff et al., 2010; Korschinek et al., 2010) and the increasingly low analytical backgrounds that can be realised, it is now possible to analyse samples covering a wide range of temporal settings, including historic times (e.g. Schaefer et al., 2009). The rate at which cosmogenic nuclides are produced is extremely low – a few atoms per gram of rock per year (Borchers et al., 2015) and the rapid attenuation of cosmic radiation with depth confines the production of cosmogenic nuclides to the upper few metres of the crust, production rates decreasing roughly exponentially with depth (Argento et al., 2015a, b). Production rates...
of cosmogenic nuclides are mainly a function of geomagnetic latitude and altitude above sea level (Balco et al., 2008; Lifton et al., 2014). Site-specific cosmogenic nuclide production rates are also subject to several other factors, the most important of these being the geometry of the surrounding topography, which shields part of the incoming cosmic radiation (Dunne et al., 1999; Codilean, 2006; DiBiase, 2018).

The application of $^{10}\text{Be}$ (or any other in situ-produced cosmogenic nuclide) to the study of Earth surface processes is based on the principle that its concentration is directly proportional to the exposure time to cosmic radiation. Cosmogenic nuclides will accumulate in surficial deposits over time such that their concentration will be directly related to not only the exposure age but also the rate at which the surface is eroding (Lal, 1991; Granger et al., 2013; von Blanckenburg and Willenbring, 2014). As a parcel of rock or sediment is brought toward the surface by erosion on a hillslope, its $^{10}\text{Be}$ concentration increases at a rate that depends mainly on the rate of erosion, and the $^{10}\text{Be}$ surface production rate at that locality. When the parcel of rock or sediment reaches the surface, it is transported via hillslope processes to the fluvial system, where it mixes with sediment from other parts of the contributing catchment. Thus, rivers act not only as agents of erosion but also as integrators, collecting sediment from all parts of the catchment in an amount that is proportional to their denudation rate such that, at the outlet of the catchment, the sediment will contain an average concentration of $^{10}\text{Be}$ (and $^{26}\text{Al}$) that is a measure of the catchment’s mean denudation rate (von Blanckenburg, 2005; Granger and Schaller, 2014). The technique of determining basin-wide denudation rates from CRN concentrations in stream sediments was first introduced in the mid-1990s (e.g. Brown et al., 1995; Bierman and Steig, 1996; Granger et al., 1996), and since that time denudation rates have been determined in over 4000 river basins from a wide range of tectonic and climatic settings.

### 2.2 Luminescence dating of sediment

Luminescence dating provides an estimate of the amount of time elapsed since mineral grains (quartz or feldspar) were last exposed to intense heat or sunlight. The suite of techniques includes thermoluminescence dating (TL), in which the luminescence signal is produced by heating mineral grains in the laboratory during measurement (Aitken, 1985; Huntley et al., 1985), and optically stimulated luminescence dating (OSL), in which the luminescence signal is produced by exposing the mineral grains to an intense light source (Aitken, 1998). The suite of techniques can be used to date events as young as a few decades (e.g. Wolfe et al., 1995; Rustomjii and Pietsch, 2007; Pietsch et al., 2015; Croke et al., 2016) to those as old as nearly 1 Ma (Arnold et al., 2015). The basis of both TL and OSL dating resides in measurements of the trapped charge (e.g. electrons) within mineral lattice imperfections which accumulate over time. When electrons are exposed to ionising radiation produced by the decay of radioisotopes contained in the surrounding sediment matrix, and/or via exposure to high-energy cosmic rays, electrons will move from a lower energy level (valence band) to a higher energy level (conduction band). Moving between the two bands, some of the energised electrons will become trapped by defects in the crystal lattice. In a parcel of sediment that is buried and thus shielded from sunlight and/or intense heat, the number of trapped electrons will increase steadily with time in proportion to the intensity of the ionising radiation flux (i.e. dose rate) and water saturation of the sediment. When the irradiated mineral grains are exposed to sunlight (or intense heat) the electrons will escape the traps, and the luminescence clock is zeroed. Thus, TL and OSL provide ages that represent the last time the electron traps were emptied or bleached – either by exposure of the sediment to sunlight (e.g. during sediment transport) or by heating (e.g. during a bush fire or in aboriginal hearths) (Wintle, 2008; Rhodes, 2011).

When luminescence dating techniques are applied to sediments, an often used assumption (when analysing multiple grains) is that the electron traps were completely emptied prior to deposition and so the luminescence clock has been effectively zeroed. In the case of OSL even short exposure to sunlight (< 1 min) is sufficient to bleach the sediment grains and thus zero the luminescence clock; however for TL, a longer exposure to sunlight is required to remove the TL signal. Fine mineral grains (< 63 µm) that are transported by wind or as suspended fluvial sediment should be exposed to sunlight while airborne or in the upper parts of the water column. On the other hand, larger grains that travelled as bedload might only be partially bleached. Independent of grain size, when dealing with fluvial sediment, the bleaching characteristics of the sample need to be assessed in order to determine the use of an applicable age model. Possible strategies for determining bleaching characteristics include using geomorphic models that reconstruct mineral grain pathways and thus predict optimal bleaching regimes (e.g. Fuchs and Owen, 2008), analysing recently deposited sediment to assess their residual OSL–TL signals (e.g. Rhodes and Bailey, 1997; Singarayer et al., 2005), pairing with a second independent dating technique such as radiocarbon dating (e.g. Olley et al., 2004), and analysing different grain size fractions or closely spaced samples with different depositional energies under the assumption that these might have behaved differently during sediment transport (e.g. Richards et al., 2000). Alternatively, various age models can be applied to either single or multi-grain data sets (e.g. minimum age model; Galbraith et al., 1999), which statistically differentiate partially bleached grain populations so as to derive the equivalent dose and subsequent age of the depositional event.
2.3 The CAIRN method for calculating CRN-based basin-wide denudation rates

CAIRN is an automated, open-source method for calculating basin-averaged denudation rates so that inferred denudation rates are reproducible: the method ingests topographic data, cosmogenic $^{10}$Be and $^{26}$Al concentrations and a parameter file and any two users with the same inputs will calculate the same denudation rate (Mudd et al., 2016). CAIRN forward models $^{10}$Be or $^{26}$Al concentrations at every pixel for a given denudation rate, taking into account latitude and altitude scaling of CRN production rates as well as snow, self, and topographic shielding. The obtained concentrations are averaged to predict a basin-averaged $^{10}$Be or $^{26}$Al concentration, and Newton’s method is then used to find the denudation rate for which the predicted concentration matches the measured concentration and to derive associated uncertainties. CAIRN is also capable of ingesting fixed denudation rates in masked portions of the input raster, allowing users to calculate spatially varying denudation rates in nested basins. In addition, CAIRN outputs spatially averaged CRN production scaling and topographic shielding values that can be used with other available CRN calculators that do not provide spatial averaging, including the online calculators formerly known as the CRONUS-Earth online calculators (Balco et al., 2008) and the Microsoft Excel-based COSMOCALC (Vermeesch, 2007). Because there is no graphical interface and because releases of the software are tagged, CAIRN users can simply publish digital elevation model (DEM) metadata, CRN data files, and CAIRN input files, and denudation rates should be reproducible. The open-source framework means that the code can be modified to include updated methods for production rates and scaling factors. Future users can thus recalculate denudation rates using updated versions of the code. CAIRN includes scripts for producing separate basin rasters for each cosmogenic sample from a regional topographic raster so that the denudation rate calculations can be run on multiple processors, meaning that large regional data sets can be processed simultaneously on compute clusters.

3 Accessing data from OCTOPUS

This section provides a description of the software infrastructure behind OCTOPUS. It also describes the ways in which data can be accessed, interrogated, and downloaded. The software infrastructure behind OCTOPUS consists of a combination of off-the-shelf open-source packages, bespoke code for handling the upload and download of data, and a web interface.

3.1 System architecture

The software architecture behind OCTOPUS is illustrated in Fig. 1. The data are stored in two separate locations. First, tabular data and the point and polygon geometries associated with each sample site or study (see Sect. 4) are stored in a PostGIS database. PostGIS (https://postgis.net, last access: 28 November 2018) is a spatial database extender for the PostgreSQL object-relational database management system (https://www.postgresql.org, last access: 28 November 2018), adding support for geographic objects and allowing location-based queries to be run in SQL. Second, all data (tabular, vector, and raster) and auxiliary information (e.g. CAIRN input and output files) (see Sect. 4) are also stored in separate zip archives, with one zip file for each study. This hybrid setup was chosen over having all tabular, vector, and raster data together in the PostGIS database because (i) it offered more flexibility regarding the list of files and file formats that could be included for download, and (ii) it made the coding of data upload and download simpler. The PostGIS database is connected to a GeoServer instance (Fig. 1). GeoServer is an open-source server that allows the sharing, processing, and editing of geospatial data (http://geoserver.org, last access: 28 November 2018), and implements a range of OGC data-sharing standards, including the widely used Web Feature Service (WFS) and the Web Map Service (WMS) standards. GeoServer also produces a variety of commonly used geospatial data formats via WFS, including KML and the ESRI Shapefile, and so can export data for using with popular desktop GIS applications such as Google Earth, ArcGIS, and QGIS (Fig. 1). It is also possible to connect to the GeoServer instance directly in QGIS and

Figure 1. The setup of the OCTOPUS data-storage and -sharing platform. See text for more details.
interrogate the data via the WFS protocol (see below). The OpenLayers (https://openlayers.org, last access: 28 November 2018) JavaScript library is used to display the geospatial data served by the GeoServer instance in a web browser (Fig. 1). OpenLayers also allows for the data to be queried and a selection to be made for download.

3.2 Accessing data using the web interface

The web interface has a simple design and its sole purpose is to enable users to visualise the various data collections and to select data for download. The web interface includes the following elements (Fig. 2): a message box that provides the user with step-by-step help on how to navigate the web page (Fig. 2, #1); a collapsible panel with a list of all available data layers – these are grouped by data collection (see Sect. 4) and can be toggled on or off (Fig. 2, #2); navigation buttons allowing zooming and scrolling (Fig. 2, #3); and the data download button (Fig. 2, #4). The latter opens a dialogue panel and switches the cursor from panning mode to selection mode, allowing for data layers to be selected and added to a download list. The OpenLayers map frame uses Google Terrain as the base layer and the point and polygon data are displayed using different colours for each collection (Fig. 2, #5).

Figure 3 illustrates a typical user interaction with the web interface. First, the user displays the data collection(s) of interest and navigates to the desired geographical area. This can be achieved by using the navigation buttons or simply by clicking (to zoom) and dragging (to pan) on the map area. To query the data, the user clicks on a point or polygon feature. This action displays an information panel that includes a subset of the records available as part of the attribute table for each point or polygon feature. In the case of overlapping features, the information panel displays records for all features (Fig. 3, #1). Displayed information includes sample ID, publication details, and recalculated $^{10}$Be-based denudation rate with uncertainty for CRN data, or published age with uncertainty for OSL–TL data. The dialogue panel closes automatically once the user clicks anywhere outside of the panel in the map display window. The displayed data are only a subset of the available attribute data and are meant to provide the user with basic information about each point or polygon record. To download data, the user clicks on the download button. This action turns the cursor into a selection tool (the user drags a box around desired points and polygons to select) and displays a dialogue panel requesting user information such as name, email address, and the intended use of the data (Fig. 3, #2–3). The user has the option to fine-tune the list of selected studies by toggling on or off each study from the list generated after the selection box is drawn. It is
Figure 3. Screenshots of the OCTOPUS web page illustrating a typical user session. See text for more details.

possible to select multiple studies from multiple collections at the same time. A valid email address is required as links to the data are sent to the user via email immediately after the download button is pressed. There is no verification of who the data requestor is or where that person is from; however, none of the fields can be left empty and all entered information is stored in a log file permanently and is used for reporting purposes. Thus, although not mandatory, providing some meaningful information when downloading the data via the web interface will support future efforts to secure funding for OCTOPUS.
3.3 Accessing data using the Web Feature Service (WFS) capability

The web interface allows users to download all of the data: tabular, vector, and raster. The data are organised in studies – each publication is a “study” – with files belonging to each study stored in separated zip archives (see below). The size of these zip archives ranges from as small as 1 MB to just over 2.5 GB, and so the web interface is meant for downloading a small number of studies per session rather than the entire collection – the size of which at the time of writing was just over 165 GB. Users who want access to a subset or to the entire collection but who do not need access to the raster data, however, can download the tabular and vector data using the WFS capability running on the GeoServer instance, instead. WFS allows geospatial data to be interrogated and requested for download using a URL via a web browser or displayed directly in desktop applications such as QGIS (see below). It is beyond the scope of this paper to provide a manual on WFS. Rather, here we provide a series of examples that users can modify to perform basic queries and download data. For a more comprehensive introduction to WFS and GeoServer, the reader is referred to Iacovella and Youngblood (2013) or to the GeoServer documentation web page at http://docs.geoserver.org (last access: 28 November 2018).

The following example WFS request will download all drainage basins belonging to the CRN International collection in the ESRI Shapefile format:


In the above example, be10-denude:crn_int_basins is the name of the data layer to be downloaded and SHAPE-ZIP refers to the format used to export the data (here, ESRI Shapefile). To obtain a full list of available data layers and export data formats, one should use the following request:


and look under FeatureTypeList and ows:OperationsMetadata in the results displayed for layer names and data formats, respectively. It is possible to request only a subset of the data by using the CQL/ECQL query language. For example, the following WFS request will download all drainage basins (with CRN data in the attribute table) from the CRN Australia collection in the ESRI Shapefile format that were published between 2005 and 2010:


CQL_FILTER=pubyear+BETWEEN+2005+AND+2010

where pubyear is the name of the field containing the publication year (see Table S1, included as part of the Supplement). Similarly, CQL_FILTER=ebe_mmkyr<10 will download all records with a $^{10}$Be denudation rate < 10 mm/kyr$^{-1}$, and CQL_FILTER=studyid='S066' will download all records that belong to study S066. Lastly, it is also possible to subset the data by geographic location:


where 0, 20, 40, 60 are the coordinates of the bounding box used to clip the data (following the format: x1,x2,y1,y2) and EPSG:4326 indicates that the coordinates are WGS84 latitude and longitude in degrees (for a complete list of EPSG references see http://spatialreference.org, last access: 28 November 2018).

For users that wish to display and interrogate the data without actually downloading any files, it is possible to access OCTOPUS from QGIS directly by using the “add WFS layer” function and connecting to http://earth.uow.edu.au:80/geoserver/wfs

QGIS is a free and open-source cross-platform desktop GIS application that supports viewing, editing, and analysis of geospatial data (https://www.qgis.org, last access: 28 November 2018). An example QGIS session with a WFS connection to OCTOPUS is shown in Fig. 4. A WFS connection will provide direct access to the data, and so any data accessed remotely is treated in the same way as data stored locally. Thus, it is possible to modify the symbology of the layers, query the data, and run analysis functions on them. However, all the above come at the cost of much more data being transmitted, and so users wanting to perform analyses on the OCTOPUS collections are recommended to instead download a local copy of the data first using one of the methods described above.

4 The OCTOPUS data collections and data structure

The compiled CRN and OSL–TL data are organised in three collections, namely (i) CRN International, including $^{10}$Be (and $^{26}$Al) measurements in fluvial sediment samples from across the globe but excluding Australia; (ii) CRN Australia, including $^{10}$Be (and $^{26}$Al) measurements in fluvial sediment samples from Australia; and (iii) OSL & TL Australia, including OSL and TL measurements in fluvial sediment samples from stratigraphic sections and sediment cores from across the Australian continent. The aim of OCTOPUS is to
Figure 4. Screenshot of the QGIS application interface displaying data from OCTOPUS accessed through a WFS connection. A WFS connection will provide direct access to the data, meaning that (1) its symbology can be modified, (2) it can be queried, and analysis functions can be run on it.

compile and incorporate all data – both published and unpublished – that is publicly available and we do not think that it is our role to decide on the quality of the data that is already published. However, in some instances, where a publication did not provide sufficient information for the data files to be produced (e.g., insufficient information to be able to confidently locate and delineate drainage basins) and this information could not be obtained from elsewhere, those data were excluded from OCTOPUS. Further, despite our best efforts it is likely that we have missed some studies during our search. Given the above, there are studies that were excluded from the current release of OCTOPUS. This was not an editorial decision except in cases where we had no choice due to a lack of information (see above).

4.1 CRN International and CRN Australia

The CRN International and CRN Australia collections consist of $^{10}\text{Be}$ (and where available, also $^{26}\text{Al}$) basin-wide denudation rates published in the peer-reviewed literature up to 2018. As already mentioned, the data are organised in studies, with files belonging to each study stored in separated zip archives (Fig. 5). The mean sample number per study is $\sim 20$ and the ratio of published $^{26}\text{Al}$ to published $^{10}\text{Be}$ measurements is approximately 1 to 10. For each $^{10}\text{Be}$ data point, there is a point geometry file representing the location of the sample site, and a polygon geometry file representing the outline of the drainage basin from which the sampled material is originating. An attribute table including published and recalculated $^{10}\text{Be}$ (and $^{26}\text{Al}$) data and a comprehensive set of metadata is linked to the polygon geometry file. A complete description of all attribute data entries is provided in Table S1, included as part of the Supplement. For each study, each zip archive also includes seven raster layers: (i) a hydrologically corrected DEM with elevation values in metres (file name suffix: _demhydro), (ii) a flow-direction raster calculated using the D8 flow-routing method (Jenson and Domingue, 1988) (_d8flowdir), (iii) a flow-accumulation raster calculated with the same D8 method (_flowacc), (iv) a slope gradient raster calculated using...
Figure 5. Data organisation diagrams for the CRN and OSL–TL collections. See text for details.
the method described in Horn (1981) with units in m km$^{-1}$\_gradmkm, (v) an atmospheric pressure raster, showing local atmospheric pressure in hPa calculated based on the NCEP2 climate reanalysis data (Compo et al., 2011)\_atmospres, (vi) a cosmogenic nuclide production scaling raster calculated using the method described in Stone (2000)\_prodscale, and (vii) a cosmogenic nuclide production topographic shielding raster calculated using the method described in Codilean (2006)\_toposhield. All raster layers were derived using the Shuttle Radar Topography Mission (SRTM) 90 m Digital Elevation Database (Farr et al., 2007) and extend 20 km beyond the boundaries of the drainage basins in each study. For two studies, namely Henck et al. (2011) and Reber et al. (2017), due to very large basin areas, all raster layers with the exception of slope gradient were calculated from SRTM data resampled to 500 m resolution. Each zip archive also includes a series of text files representing CAIRN configuration and input data files, and CAIRN output files, including files to be used with the online calculators formerly known as the CRONUS-Earth online calculators (Balco et al., 2008).

Published $^{10}$Be concentrations (atoms g$^{-1}$) were renormalised to the Nishizumi 2007 $^{10}$Be AMS standard (Nishizumi et al., 2007), and basin-wide denudation rates were calculated with CAIRN. Basin-averaged nuclide production from neutrons and muons was calculated with the approximation of Braucher et al. (2011) and using a sea-level and high-latitude total $^{10}$Be production rate of 4.3 atoms g$^{-1}$ yr$^{-1}$ (Mudd et al., 2016). Production rates for catchment-wide denudation rates were calculated at every pixel using the SRTM 90 m DEM, with the time-independent Lal/Stone scaling scheme (Stone, 2000). Atmospheric pressure was calculated via interpolation from the NCEP2 reanalysis data (Compo et al., 2011). Topographic shielding was calculated from the same DEM using the method of Codilean (2006) with $\theta = 8^\circ$ and $\phi = 5^\circ$. Following the submission of this paper, a new study by DiBiase (2018) showed that topographic shielding corrections are inappropriate for calculating basin-wide denudation rates, in most settings, and are only required for steep catchments with non-uniform distribution of quartz and/or denudation rates. For this reason, future iterations of the CRN International and CRN Australia collections will also include $^{10}$Be (and where available $^{26}$Al) denudation rates calculated without correcting for topographic shielding. All calculations assumed a $^{10}$Be half-life of 1,387 ± 0.012 Myr (Chmeleff et al., 2010; Korschinek et al., 2010).

For simplicity and consistency across the global compilation, no corrections were made for lithological differences in quartz abundance, glacier cover, and snow shielding. Performing such corrections in a consistent manner on a global scale is impossible. However, all CAIRN input and configuration files are provided and these corrections can be readily applied by end users to individual studies. Further, CRN International and CRN Australia include both the originally published denudation rates and the ones recalculated using CAIRN, and so detailed comparisons can be made by users. Figure 6 shows a first-order comparison between published and recalculated $^{10}$Be denudation rates. With the exception of a small number of data points ($n \sim 10$), there is good agreement between published and recalculated $^{10}$Be denudation rates, with no obvious trends related to elevation or basin size. Where large discrepancies exist, these are due either to differences in drainage basins as published versus drainage basins identified on the SRTM DEM during data recalculation or due to corrections that were applied to the data in the original publication that were not appropriately described in the latter. Discrepancies also exist in the case of studies where substantial portions of a drainage basin consisted on non-quartz-bearing lithologies (e.g. Safran et al., 2005; Croke et al., 2015) and where corrections for quartz abundance were applied to the data in the original publications but were not replicated here. The number of such basins is small, however, and will not impact any regional or larger-scale analyses done with the CRN data. For small-scale studies users should compare published with recalculated denudation rates and determine whether a new recalculation that involves corrections for quartz abundance, glacier cover, and/or snow shielding is warranted.

Approximately 5% of compiled $^{10}$Be measurements – all of which were published in two highly regarded journals – could not be incorporated into OCTOPUS due to information that is insufficient to reproduce drainage basin extents.

In terms of geographical extent, the global CRN compilation exhibits considerable bias (Fig. 7a). The majority of the $^{10}$Be (and $^{26}$Al) measurements are from Northern Hemisphere drainage basins, clustering around distinct, mostly tectonically active, topographic regions, such as the Pacific coast of the United States, the Appalachians, the European

Figure 6. Published versus recalculated $^{10}$Be-based denudation rates. Data points are coloured according to average basin elevation and circle sizes are proportional to basin area. Note the good agreement between the two data sets and the lack of obvious trends related to basin elevation and basin area.
Alps, and the Tibet-Himalaya region. Due to some recent studies, there is also good coverage of the South American Cordillera. However, there is a considerable lack of data from low-gradient and tectonically passive regions, such as large parts of Australia, most of Africa, and most of Asia less the Tibet-Himalaya region. Further, there are no data from latitudes above \( \sim 55^\circ \). The observed geographical bias is a reflection of the intense interest of the geomorphological community in estimating rates of erosion and weathering in tectonically active mountain regions with one of several aims to understand the role of surface processes in the global climate system (e.g. Molnar and England, 1990; Raymo and Ruddiman, 1992; Willenbring and von Blanckenburg, 2010; Herman et al., 2013). Further, the lack of data from high latitudes is partly due to the desire to stay away from formerly glaciated environments. Although the geographical bias does not make the CRN collection less valuable, it may confound studies aiming to infer global-scale trends from these data (cf. Portenga and Bierman, 2011; Willenbring et al., 2013; Harel et al., 2016). Despite the geographical bias, however, the global CRN data sample basins with a wide range of slope gradients, elevations, and basin areas (Fig. 7b–c).

4.2 OSL & TL Australia

The OSL & TL Australia collection consists of thermoluminescence (TL) and optically stimulated luminescence (OSL) measurements in fluvial sediment samples from stratigraphic sections and sediment cores from across the Australian continent. The collection includes data published in the peer-reviewed literature up to 2017 and also previously unpublished data compiled from technical reports and various Honours, MSc, and PhD theses. The majority of the TL data are from sources published from 1986 up to 2005, whereas the majority of the OSL data are from sources less than 10 years old (Fig. 8). In terms of geographical extent both TL and OSL data are concentrated in the south-eastern and eastern parts of the Australian continent, with \( \sim 500 \) measurements from Australia’s largest river basins – Lake Eyre (LEB) and Murray-Darling (MDB) basins – and with an equal amount from rivers draining the eastern seaboard.
The western half of Australia is severely understudied, with one single OSL study for the entire region, namely Veth et al. (2009). Focused interest on river systems is proximal to high-population density areas, where floods are a potential threat (e.g., Brisbane River, after major floods in 2011; Croke et al., 2016), or where rivers are of great agricultural importance, such as the MDB. This well-justified bias, however, leaves a gap in the understanding of regions dominated by rivers that are now dry or ephemeral and yet that could hold information on past climatic regimes now buried under the desert sand. The focus on south-eastern coastal river systems draining the Great Dividing Range could be a source of bias in continent-wide interpretations, where the rivers draining the western intracontinental ranges and plains remain underrepresented.

Similarly to the CRN collections, the data are organised in studies – each publication is a study – with files belonging to each study stored in separated zip archives (Fig. 5). For each OSL or TL data point, there is a point geometry file representing the location of the sample site. An attribute table including published OSL or TL ages and a comprehensive set of metadata is linked to the point geometry file (a complete description of all attribute data entries is provided in Table S2, included as part of the Supplement). The zip archive also includes two separate polygon geometry files: one representing the outline of the drainage basin of the most downstream sample and one representing the area extending 20 km beyond the boundaries of this drainage basin. For studies with basin areas up to $100,000 \text{ km}^2$, each zip archive also includes four raster layers: (i) a hydrologically corrected DEM with elevation values in m (file name suffix: _demhydro), (ii) a flow-direction raster calculated using the D8 flow-routing method (Jenson and Domingue, 1988) (_d8flowdir), (iii) a flow-accumulation raster calculated with the same D8 method (_flowacc), and (iv) a slope gradient raster calculated using the method described in Horn (1981) with units in $\text{m km}^{-1}$ (_gradmkm). All raster layers were derived using the hydrologically enforced SRTM 30 m digital elevation model (DEM-H) obtained from Geoscience Australia (Geoscience Australia, 2011) and were clipped to the extent of the 20 km buffer polygon layer. For studies with basin areas exceeding $100,000 \text{ km}^2$ (e.g., Callen and Nanson,}
1992; Bourman et al., 2010; Jansen et al., 2013) raster layers (i) to (iii) were derived using Geoscience Australia’s GEO-
DATA 250 m digital elevation model and flow-direction grid (Geoscience Australia, 2008), as the SRTM DEM produced files that were too large to transfer online.

4.3 Other collections

In addition to the CRN and OSL–TL collections described above, the current version of OCTOPUS also includes additional CRN data organised under two collections: CRN XXL and CRN In-Prep. These two collections are not officially supported by the OCTOPUS project, and are included here only for completeness. The first collection consists of five studies with samples from the Yangtze (Chappell et al., 2006), Amazon (Wittmann et al., 2009, 2011), Ganga (Lupker et al., 2012), and Brahmaputra basins (Lupker et al., 2017). These studies focused on very large basins that could only be handled by CAIRN when running a 500 m resolution DEM that, however, produced drainage basins that were substantially different to what was published, especially in the case of rivers in the Amazon basin. Further, Chappell et al. (2006) do not report denudation rates – suggesting that calculating these might have little meaning for their samples and both Wittmann et al. (2009, 2011) and Lupker et al. (2012, 2017) perform corrections to the data, some of which (e.g. removing floodplain areas from production rate calculations) we did not wish to replicate. To this end, CRN XXL does not include recalculated values nor does it include any raster layers. CRN In-Prep is an inventory of samples processed at the University of Wollongong where $^{10}$Be and $^{26}$Al have been measured and the data are not yet published. The collection includes sample metadata and point and polygon geometry files.

5 User contributions to OCTOPUS

User contributions to OCTOPUS are welcome. Those wishing to submit data should download a study and use that as the template for data structure, formats, and naming convention (see also Fig. 5). As a minimum, a contribution should include point and polygon geometry files, and an attribute table with all records listed in Tables S1 and S2 with the exception of those records that are output by CAIRN. Data files should be submitted to the contact address listed in the email received from OCTOPUS when downloading data. The data collections making up OCTOPUS have been assigned digital object identifiers (DOIs), and as a consequence, adding new data needs to follow a versioning scheme, with each new version requiring new DOIs. Thus, data contributed by users will be incorporated in the next release of a given collection, rather than being added to the current one.

6 Data availability

OCTOPUS can be accessed at https://earth.uow.edu.au (last access: 28 November 2018). The data collections that make up the 2018 release of OCTOPUS (version 1 – Dooku’s Dilemma) have been assigned the following DOIs: https://doi.org/10.4225/48/5a8367feac9b2 (CRN International; Codilean et al., 2017a), https://doi.org/10.4225/48/5a836cdfac9b5 (CRN XXL; Lapparelin et al., 2017b), and https://doi.org/10.4225/48/5a836db1ac9b6 (OSL & TL; Codilean et al., 2017c). A copy of OCTOPUS has also been deployed to https://earthtest.uow.edu.au (last access: 28 November 2018). This copy is not supported and is used for testing modifications to the website and data collections before deployment to the official site. Users should refer to the DOIs provided to ensure that they are accessing the current and supported version of the data.

7 Conclusions

We have produced a database of cosmogenic radionuclide and luminescence measurements in fluvial sediment and we have built infrastructure for hosting and maintaining the data at the University of Wollongong and making them available to the research community via an Open Geospatial Consortium (OGC)-compliant web service. The database consists of $^{10}$Be, $^{26}$Al, TL, and OSL measurements in fluvial sediment samples along with ancillary geospatial vector and raster layers. Sample metadata are comprehensive and include all necessary information for the recalculation of $^{10}$Be and $^{26}$Al denudation rates using the CAIRN program. The repository and visualisation system enable easy search and discovery of available data. Use of open standards also ensures that data layers are visible to other OGC-compliant data-sharing services. Thus, this project will turn data that were previously invisible to those not within the CRN and luminescence research community into a findable resource. This aspect is of particular importance to industry or local government who are yet to discover the value of geochronological data, for example, in evaluating how human-induced land use practices have accelerated soil erosion and which measures are necessary for restoring these rates to their natural benchmark levels. Our intention is for OCTOPUS to become the default go-to place for CRN and luminescence data. The availability of the repository and its associated data curation framework will provide the opportunity for researchers to store, curate, recalculate and reuse previously published but otherwise unusable CRN and luminescence data. This delivers the potential to harness old but valuable data that would otherwise be lost to the research community. OCTOPUS will enable new research and generate new knowledge by converting a multitude of disconnected data sets into one connected and streamlined database. Current data sets allow local-scale analyses. The streamlined database will allow for regional-scale and
even continental- or global-scale analyses. The transparent data reanalysis framework will also reduce research time and avoid the duplication of effort, which will be highly attractive to other researchers. Ultimately, OCTOPUS will ensure that CRN and luminescence data are reusable beyond the scope of the project for which they were initially collected.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/essd-10-2123-2018-supplement.

Author contributions. ATC, HM, TJC, and WMS compiled the CRN and luminescence data; ATC and HM performed the GIS analyses and the data recalculations using CAIRN with input from SMM; AG designed and built the OCTOPUS platform and web interface with input from ATC. All authors contributed to the writing of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We acknowledge financial support from the Australian National Data Service (ANDS – Project HVC15) and the University of Wollongong (UOW) to Alexandru T. Codilean and Timothy J. Cohen. Henry Munack benefited from an Endeavour Postdoctoral Research Fellowship granted by the Australian Government. A prototype of OCTOPUS was built while Alexandru T. Codilean worked at the Deutsches GeoForschungsZentrum (GFZ). We thank Monika Oakman, Ben Cornwell, Peter Newnam, Clare Job, Gabriel Enge, Heidi Brown, Alan Glixman, and Keith Russell for their support during this project. We also thank the CRN and luminescence communities for generously sharing their data. We thank Greg Balco, Zsófia Ruszkiczay-Rüdiger, Sebastian Kreutzer, and István Gábor Hatvani for careful reviews of the discussion paper. We also thank Ingrid Ward, Todd Ehlers, and Vincent Godard for providing comments on the discussion paper.

Edited by: Attila Demény
Reviewed by: Greg Balco, Sebastian Kreutzer, Zsófia Ruszkiczay-Rüdiger, and István Gábor Hatvani

References

Chmeleff, J., van Blanckenburg, F., Kossert, K., and Jakob, D.: Determination of the 10Be half-life by multicollector ICP-MS and


