

Contributing Data to the IGCP 509 Databases

IGCP 509 project participants will be using data from two databases, both of which are available from the IGCP 509 web site (<http://earth.geology.yale.edu/igcp509/>) and from the host web site at the Saskatchewan Isotope Laboratory (SIL) (<http://sil.usask.ca/databases.htm>). These databases are:

- DateView geochronology database
- IGCP509 lithostratigraphy, deformation and metamorphism database.

The latter is actually an age-restricted view of the StratDB database (also available at the SIL web site).

It is the IGCP509 database that will form the principal mechanism to facilitate the construction of time-space charts for the project. This document provides a brief overview of the database, together with details of how to contribute information. Details of how to construct the time-space charts, utilising information in IGCP509 and DateView, will be provided later.

IGCP509 is a web-enabled extension of an earlier desktop system which captured information for lithostratigraphic units recognised by the South African Committee for Stratigraphy, with several enhancements specifically developed for the IGCP 509 project by Bruce Eglington, Steve Reddy and Dave Evans. The data are stored in a Firebird open-source relational database management system which provides full referential integrity. Although open-source, Firebird provides similar capabilities to major commercial database systems such as Oracle and is very easy to manage. The web interface to the database has been programmed by Bruce Eglington, using the Delphi language and IntraWeb components.

Conceptual Design

The central tables in the IGCP509 database contain essential information for each lithostratigraphic unit (Figure 1) and structural domain (Figure 2). Other tables provide referential, look-up values or linkage fields to draw together information from various of the database tables.

Conceptual Model of IGCP 509 Database (lithostratigraphic components)

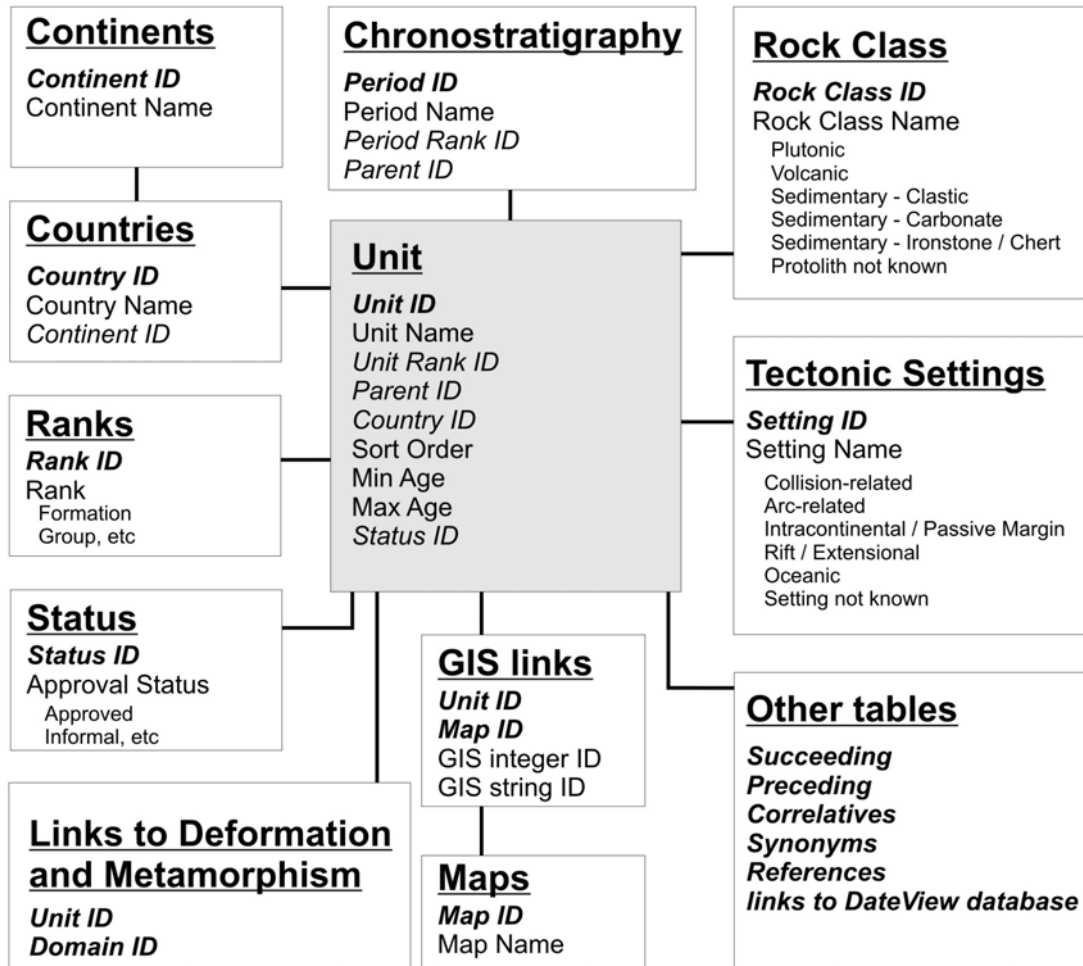


Figure 1

Conceptual Model of IGCP 509 Database (deformation and metamorphism components)

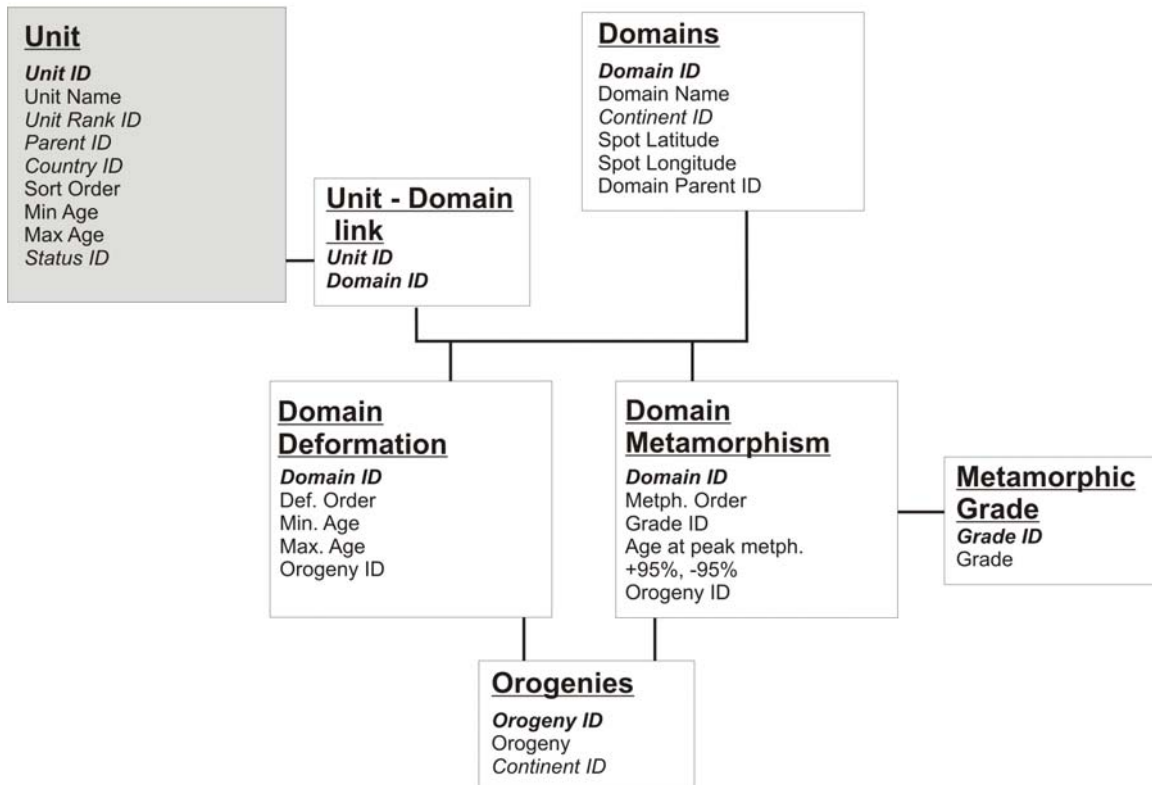


Figure 2

Data Collation and Submission

Contributors will need to populate certain essential information (see later), but are encouraged to also populate information in various of the supplementary data tables if possible. The emphasis of the IGCP 509 project is on the Palaeoproterozoic and immediately time-adjacent time intervals (approximately 1500 Ma to 2600 Ma), but all information contributed will be available via the StratDB database. Obviously, the more information is provided, the greater the benefit will be to future researchers.

Contributors may also choose to add data to the DateView geochronology database, in which case, provided that appropriate records exist in both IGCP509/StratDB and DateView databases for the dated units, users will be able to see geochronological information associated with individual lithostratigraphic units and tectonic/metamorphic domains. Information on how to contribute to the DateView database is provided elsewhere (http://sil.usask.ca/DV_compilations.htm). To ensure that link fields are in place, with a minimum of subsequent editing, contributors should first add lithostratigraphic units to IGCP509/StratDB and then export these to DateView using the

appropriate button on the web form for the IGCP509/StratDB database. Thereafter, all crystallisation, metamorphism and deposition records added to DateView for that lithostratigraphic unit will automatically be visible in IGCP509/StratDB.

Data collation for submission to the IGCP509 database may be performed using either a MicroSoft Access database or MicroSoft Excel spreadsheets. The database option allows contributors to compile all aspects of their contributions offline, whereas the spreadsheet approach will entail a two-stage process. Phase 1 requires the initial submission of certain essential data and the importing of these data to the primary database. Contributors will then need to download this information from the IGCP509 or StratDB web database and populate spreadsheets with data to link units to domains, etc. Other, non-essential data could also be contributed at this stage.

For those contributors that have access to GIS coverages of countries or regions, it is strongly recommended that the information for submission to this IGCP509 database be initially derived from the database associated with the GIS shapefile, etc. In such cases, be sure to extract the unit name and unique ID (not the GIS Object ID, but rather the unique ID associated with each specific unit). If this is done, future users will be able to directly query the IGCP509/StratDB database and link to the GIS maps (if they have the digital GIS files) as well as to time-space charts without any additional data capture. More details of the specifics to be captured are present later in this document.

Even with the various templates in place, it is going to require a substantial amount of manual work to incorporate the various contributions into the primary database. ***If you can, please use the MSAccess method of data collation and submission as this will substantially reduce the work-load for all of us.***

Once data have been imported to the primary Firebird database, the original MS Access database tables and Excel spreadsheets will no longer be valid (some of the key-value information will have changed because of contributions from multiple sources). Contributors should also note that the MS Access database requires that the end user have a legal copy of MS Access and that Apple computers do not have the software to use MS Access (unless you use Windows emulation).

Please **do not** make changes to the values stored in the various look-up tables and don't change the structure or layout of the database tables and Excel spreadsheets. If you do, there is every chance that your data will not be compatible with everybody else's. If changes are required, please first discuss them with the IGCP 509 project leaders (Dave Evans, Steve Reddy and Rajat Mazumder) and with Bruce Eglington, using the IGCP 509 Google Newsgroup.

Essential Information for MS Access Submission

1. First, please rename the empty database file you have downloaded to something like IGCP509_yourname.mdb so that there is no chance of confusion when everybody submits their database compilations.
2. After opening the IGCP509 database, open the *users* table and add a record with your *userid*. This should be the same as the one registered with StratDB and DateView.
3. Open the *unitid* and *domainid* tables and ensure that both contain sufficient records for the number of lithostratigraphic units or domains that you intend to add. You will be limited to the values in these tables when entering data in other tables. You may, however, add more records at any time and may change record values at any stage, in which case the modified values will be propagated throughout the database.
4. Open the *stratunits* table and enter information for each field. If you so desire, you may append new records for each of the lithostratigraphic units you are contributing and accept the default values for all other fields to start with. You will, however, need to fill in appropriate values in all fields before submission. **Be warned**, if you don't, you will be faced with the very tedious process of modifying record values individually through a web browser.
5. Both the lithostratigraphic and domain information is designed to reflect hierarchical relationships. Appropriate fields need to be populated to reflect this and to get units to show up in queries in approximately the correct stratigraphic order. The hierarchical relationship for units is achieved by capturing the integer ID of the parent unit (*Parent ID* in Figure 1) **and** by specifying a *sort order* value for each unit. The convention used for this latter field is to provide the approximate age of the unit in years. This field, along with the minimum age and maximum age (both in Ma) are used to define which records fall within the scope of IGCP 509. Only those records which have Stratigraphic *Sort Order* values between 1,500,000,000 and 2,600,000,000 years and have *Minimum Age* \geq 1500.0 Ma and *Maximum Age* \leq 2600.0 Ma will be visible in the IGCP509 web interface. Users will need to use the StratDB web interface to view records outside these limits. Units which do not have a logical 'parent', should have *Parent ID* set to zero. In the case of tectonic domains, the link is achieved through a *Parent ID* field only, there is no reason to order the sub-domains.
6. Enter appropriate values for the *Rank ID* of each unit. These will be used to define the width of rectangles in the space-time charts created as final products of the IGCP 509 project.

7. Ensure that each unit is associated with the correct *Country ID*. If a unit with the same name occurs in two countries, then each country needs to have a record for the unit with different unique *Unit ID*'s for each.
8. Every unit entered into the *stratunits* table must also be entered into the *unitrockclass* and *unitsetting* tables with appropriate values for *RockClass ID* and *Setting ID*.
9. The *Status ID* field is useful for other users of the database, for instance to distinguish between stratigraphic units which are formally recognised and those that are informal. Contributors may choose to accept the default value provided without impacting on the outputs of the IGCP 509 project.
10. At this point, one needs to open the table *orogenies* and fill in all appropriate values. Make sure that each *orogeny* is associated with the continent to which it applies. One should not create records with the same orogeny name for different continents.
11. Now open table *domains* and enter appropriate values. Make sure that each *domain* is associated with the continent to which it applies. One should not create records with the same domain name for different continents. Please also provide an approximate latitude and longitude value for somewhere appropriate within the geographic extent of each domain. These will be used to facilitate queries looking for information within user-defined geographic boundaries.
12. Open table *domainmetamorphism* and enter values as appropriate for each period of metamorphism you wish to associate with each domain. *Grade ID* should be selected from the drop-down list provided. The field *Metamorphism Order* may be manually incremented to capture multiple episodes of metamorphism for each domain. Please ensure that you provide your best estimate of the age of peak metamorphism and the plus and minus 95% uncertainties for the event. These will be used to construct boxes on the time-space charts later in the project. Please also associate each period of metamorphism with an appropriate *Orogeny*. This field will allow other users, who may not know the local geology well, to search for all episodes of metamorphism associated with a regional orogeny.. You will be able to edit values later, once information has been entered into the primary Firebird database.
13. Open table *domaindeformation* and enter values as appropriate for each period of deformation you wish to associate with each domain. The field *Deformation Order* may be manually incremented to capture multiple episodes of deformation for each domain. Please ensure that you provide your best estimate of the minimum and maximum age of deformation for the event. These will be used to construct boxes on the time-space charts later in the project. Please also associate each period of deformation with an appropriate *Orogeny*. This field will allow other users, who may not know the local geology well, to search for all episodes of

Non-essential Information

If you are importing data from a GIS shapefile or coverage, please first create a record in the table *maps*, to define the map with which the data are associated. Then, open table *stratmaps* and, for each *Unit ID*, enter the *Map ID* and a unique *GISInt* value (an integer) plus, if appropriate, a unique character string for the unit (in field *GISStr*). If your GIS shapefile uses unique integer values for each unit, then all you need to enter is this value in the field *GISInt* and accept the default value of 'nd' for *GISStr*. On the other hand, if your GIS shapefile uses unique character strings for each unit, enter these in the *GISStr* field and accept the default value of 0 (zero) for *GISInt*.

Contributors may also fill in other information, none of which are essential for the IGCP 509 project to succeed. For instance, one can associate various references with individual lithostratigraphic units (tables *sourcelist* and *unitrefs*) or specify the relationships between different units. Table *succeeding* stores the information for units which are younger than the stratigraphic unit of interest whereas table *preceding* stores information for units which predate the unit of interest. Note that one only needs to fill in information for one of these. The database software can derive the reverse relationship.

Table *correlatives* stores information for units which are believed to be equivalent but occur in other countries, whereas table *synonyms* stores information for units which occur in the same country. One of the principal uses of this latter table is to capture information for unit names which were previously used (historical names). In other cases, coeval units in different sub-basins may have different names and it is useful to know that they are considered time-equivalent.

One may also download unique record ID's for information in DateView and associate these with specific lithostratigraphic units, using table *unitrecords*, or with specific domains, using table *domainrecords*. Please note that, if both DateView and IGCP509/StratDB use the same *unitname* or associated unique *unit ID*, then the database software can create these links programmatically. It is thus always best to enter lithostratigraphic unit information into IGCP509 or StratDB first, and then transfer the *unitname* and unique *unit ID* to DateView using the appropriate button on the IGCP509/StratDB web interface.

Essential Information for MS Excel Submission

1. Submission using the Excel spreadsheets will, of necessity, be a multistage process and will thus require more effort from each of us. This is because unique ID's will need to be semi-manually associated with submitted records from different contributors. The contributors will then need to download these updated records and values from the web database before creating linked records to other information. All lookup tables may be downloaded into Excel spreadsheets from the web database at any stage.
2. Download the following information from the IGCP509 or StratDB database web interfaces. You will need them to know what values to enter in some of the fields in the submission spreadsheets. The are:
 - AgeUnits
 - Continents
 - Countries
 - Metamorphic Grade
 - Ranks
 - Rock Classes
 - Settings (tectonic settings)
 - Approval Status

If you are also submitting non-essential information or additional information associated with any of the subsequent files, you will need to download updated versions of one or more of the following:

- Stratigraphic Units
 - Orogenies
 - Domains
 - References
 - Stratigraphic Relationships
 - Relative Positions
3. Please rename the spreadsheets you have downloaded to something like IGCP509_spreadsheetname_yourname.xls so that there is no chance of confusion when everybody submits their spreadsheet compilations. In some cases it might also be appropriate to include a date in the file name.
 4. Open ***IGCP509_submission_stratunits.xls*** and enter information for each stratigraphic unit. You will need to manually ensure that linking values are matched e.g. ParentID with UnitID of another unit and that fields with values derived from other tables in the primary Firebird web database are correct. If values don't match, your information will not be accepted by the primary database, resulting in tedious manual editing to find missing records. The database administrator (Bruce Eglington) will not have time to do this so it will be up to contributors.

5. Both the lithostratigraphic and domain information is designed to reflect hierarchical relationships. Appropriate fields need to be populated to reflect this and to get units to show up in queries in approximately the correct stratigraphic order. The hierarchical relationship is achieved by capturing the integer ID of the parent unit (*Parent ID* in Figure 1) and by specifying a *sort order* value for each unit. The convention used for this latter field is to provide the approximate age of the unit in years. This field, along with the *minimum age* and *maximum age* (both in Ma) are used to define which records fall within the ambit of IGCP 509. Only those records which have Stratigraphic *Sort Order* values between 1,500,000,000 and 2,600,000,000 years and have *Minimum Age* ≥ 1500.0 Ma and *Maximum Age* ≤ 2600.0 Ma will be visible in the IGCP509 web interface. Users will need to use the StratDB web interface to view records outside these limits. Units which do not have a logical 'parent', should have *Parent ID* set to zero.
6. Enter appropriate values for the *Rank ID* of each unit. These will be used to define the width of rectangles in the space-time charts created as final products of the IGCP 509 project.
7. Ensure that each unit is associated with the correct *Country ID*. If a unit with the same name occurs in two countries, then each country needs to have a record for the unit with different unique *Unit ID*'s for each.
8. Be sure to enter values in the *rockclass ID* and *setting ID* columns, using appropriate values for *RockClass ID* and *Setting ID* from the look-up spreadsheets. These values are essential to the functionality of the IGCP509 database and the creation of space-time charts.
9. The *Status ID* field is useful for other users of the database, for instance to distinguish between stratigraphic units which are formally recognised and those that are informal. Contributors may choose to accept a default value of IF without impacting on the outputs of the IGCP 509 project.
10. Open *IGCP509_submission_orogenies.xls* and capture the information for each orogeny you intend to refer to. Make sure that each *orogeny* is associated with the continent to which it applies. One should not create records with the same orogeny name for different continents.
11. Open *IGCP509_submission_domains.xls* and capture the information for each domain you intend to refer to. Make sure that each *domain* is associated with the continent to which it applies. One should not create records with the same domain name for different continents. Please also provide an approximate latitude and longitude value for somewhere appropriate within the geographic extent of each domain. These will be used to facilitate queries looking for information within user-defined geographic boundaries.

12. Open *IGCP509_submission_references.xls* and capture the information for each reference you intend to refer to. Note that references, and subsequent links to these references are not essential to the operation of the IGCP509 database. They are merely a ‘nice-to-have’ feature.
13. At this point, you will need to submit these spreadsheets to Bruce Eglington (bruce.eglington@usask.ca) for uploading into the primary database. You will be notified by email once this is done and can then run a query on the database and download the results as MicroSoft Excel spreadsheets. Copy and paste appropriate unique ID’s (such as *Unit ID*) into other spreadsheets that now need to be completed.
14. Open *IGCP509_submission_domainmetamorphism.xls* and enter values as appropriate for each period of metamorphism you wish to associate with each domain. *Grade ID* should be selected from the appropriate look-up spreadsheet. The field *Metamorphism Order* may be manually incremented so as to capture multiple episodes of metamorphism for each domain. Please ensure that you provide your best estimate of the age of peak metamorphism and the plus and minus 95% uncertainties for the event. These will be used to construct boxes on the time-space charts later in the project. Please also associate each period of metamorphism with an appropriate *Orogeny*. This field will allow other users, who may not know the local geology well, to search for all episodes of metamorphism associated with a regional orogeny. You will be able to edit values later, once information has been entered into the primary Firebird database.
15. Open table *IGCP509_submission_domaindeformation.xls* and enter values as appropriate for each period of deformation you wish to associate with each domain. The field *Deformation Order* may be manually incremented to capture multiple episodes of deformation for each domain. Please ensure that you provide your best estimate of the minimum and maximum age of deformation for the event. These will be used to construct boxes on the time-space charts later in the project. Please also associate each period of deformation with an appropriate *Orogeny*. This field will allow other users, who may not know the local geology well, to search for all episodes of deformation associated with a regional orogeny. You will be able to edit values later, once information has been entered into the primary Firebird database.

Non-essential Information

If you are importing data from a GIS shapefile or coverage, please first create a record in the spreadsheet *IGCP509_submission_maps.xls*, to define the map with which the data are associated. Then, open *IGCP509_submission_stratmaps.xls* and, for each *Unit ID*, enter the *Map ID* and a unique *GISInt* value (an integer) plus, if appropriate, a unique character string for the unit (in field *GISStr*). If your GIS shapefile uses unique integer values for each unit, then all you need to enter is this value in the field *GISInt* and use a

default value of 'nd' for *GISStr*. On the other hand, if your GIS shapefile uses unique character strings for each unit, enter these in the *GISStr* field and use a default value of 0 (zero) for *GISInt*.

Contributors may also fill in other information, none of which are essential for the IGCP 509 project to succeed. For instance, one can associate various references with individual lithostratigraphic units (spreadsheets *IGCP509_submission_references.xls* and *IGCP509_submission_unitrefs.xls*) or specify the relationships between different units. Spreadsheet *IGCP509_submission_succeeding.xls* stores the information for units which are younger than the stratigraphic unit of interest whereas spreadsheet *IGCP509_submission_preceding.xls* stores information for units which predate the unit of interest. Note that one only needs to fill in information for one of these. The database software can derive the reverse relationship.

Spreadsheet *IGCP509_submission_correlatives.xls* stores information for units which are believed to be equivalent but occur in other countries, whereas spreadsheet *IGCP509_submission_synonyms.xls* stores information for units which occur in the same country. This latter spreadsheet also stores information for unit names which were previously used (historical names).

One may also download unique record ID's for information in DateView and associate these with specific lithostratigraphic units, using spreadsheet *IGCP509_submission_unitrecords.xls*; or with specific domains, using table *IGCP509_submission_domainrecords.xls*. Please note that, if both DateView and IGCP509/StratDB use exactly the same *unitname* or associated unique *unit ID*, then the database software can create these links programmatically.

IGCP 509 Outputs from the Database

In addition to providing a long-term resource for researchers, IGCP509/StratDB and DateView will also allow project participants to query the databases and produce outputs for use in the construction of time-space charts. The web interface will provide data output in a format suitable for construction of time-space diagrams using standard GIS software such as ArcGIS (ArcView). These diagrams will be queryable using various attribute information, such as RockClass, Tectonic Setting, etc. A prototype version of the system has already been tested, using ESRI ARCGIS 9.1, and more testing will be done in the next few months. More information on this aspect of the project is provided in separate documentation, available from the SIL database web site.

The general legend for the space-time charts will be something like the image shown in Figure 3. This legend is based on the one used for a compilation of Canadian Palaeoproterozoic terranes (e.g. Ansdell *et al.*, 2006). With the pseudo-GIS approach

proposed, each one of the attributes associated with individual rectangles on the chart is selectable, making for a very flexible resource. Specific colours and symbology for the setting – rock-class matrix are derived from an ESRI ARCGIS style sheet and can thus be easily changed to suite the requirements of IGCP 509 project coordinators without any time-consuming recoding by participants.

Legend

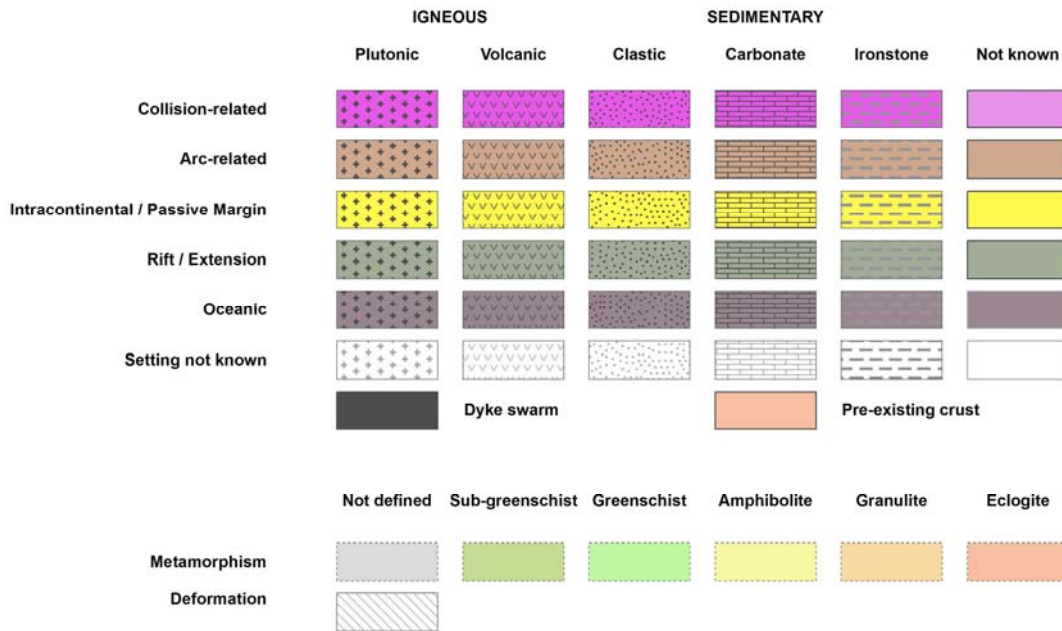


Figure 3