Meeting Notes/Outcomes

Best Practices in Tephra Collection, Analysis, and Reporting: Leading Toward Better Tephra Databases

Saturday August 19, 09:00–18:00

IAVCEI, Portland Oregon

Workshop Presentation Archives: https://vhub.org/groups/tephra2017workshop

Executive Summary

All workshop participants reconfirmed a strong commitment toward standardization of tephra field/core data collection, processing, storage and distribution. The community feels that such an interdisciplinary effort will help to advance and to solve future emerging research problems in tephra studies. While major discussion was focused on geochemical analysis, correlation, and data reporting, there was a mutual understanding that other datasets will benefit from improved interdisciplinary compatibility. Best practice checklists and templates for minimum required data are already being developed. The need for transparent data access across disciplines is a more complicated issue, one that may require a new generation of computer-based research tools. These tools should be integrated into a more complex system that is designed to assist users with solving problems for the particular research area (domain specific), while allowing data streams and tools to be interconnected into a larger framework that is flexible by design to adapt to emerging interdisciplinary problems.

Some specific anticipated products of the workshop include 1) publication of a consensus paper to draw attention to the demand, and to develop a plan for creating comparable datasets across disciplines, 2) continuing to develop multiple open access products, for example, best practice checklists, data collection and processing templates with minimum sets of required data, and 3) collation of already build tools/code/software for data processing.

Purpose

Continue efforts to bring disparate tephra researchers together, highlight commonalities, and discuss ways to work together /share data; list major research goals that a collaborative system may help to address.

What are our motivations?

- Bridging the gap between tephra communities
    - What can the proximal (volcanological) community bring to distal (tephrostratigraphic, CC, etc) communities, and vice versa
- Making tephra data more comparable (facilitating correlations)
- Making tephra data more accessible
- Building tools to make working with tephra data easier

Goals

- Collect and report data that is comparable across disciplines
- Improve or establishing mechanisms for sharing data, data products, analysis tools, code
- Making data more accessible and data workflow more transparent
- Linking tephrostratigraphers with one another (proximal-distal).

- Establish guidelines to make geochemical data potentially usable by all.
- How to allow proximal data and models to inform distal interpretations, and interpretations in other, related domains, e.g., climate change.
- How to allow distal data, e.g., from maps of extent, to inform interpretation of proximal deposits.

Proposed workshop products

- EOS article – short, quick PR piece
- Consensus paper: BV (or other) – with emphasis on scientific rationale!
- Best practices guidelines: examples, templates, code, web links, which can be built on checklists from Tephra 2014 workshop
- Collate and provide a single portal to share already build tools/code/software for data processing

Future directions

- Translate checklists into templates --
    - distribute initially by including as supplements to paper
    - Earthchem and Geochron have templates for different communities that can be adapted-we could work with Earthchem to develop a geochemical template that is tephra specific
- Build on our current collection of known databases;
- Begin to collate links to analysis tools into one place; work on motivations for linking proximal and distal datasets.

BUILDING THE SYSTEM-BRAINSTORMING SESSION

The system itself

- Its core is a relation database management system (RDBMS), or a group of distributed RDBMSs, such as the group that is currently maintained at a number of institutions.
- What do we mean by a RDBMS? Engine, data, front end, interoperability, processing…
- All agree that there is no single database out there that fits all our needs.
- Funding agencies do not often support adding old data into a database…not sexy. We could make the case that this is critical – give case studies, show value, e.g. regional correlations.
- Do we need a system with every shred of data that we collect or a simple core system?
- Would we like it to be question/answer (query) based?
- Must be able to address what's important to YOU.
- Keep it simple – first step, know what is the most important information to collect/report (e.g. template guideline). For publication, Earthchem data entry single step. Host somewhere, point to it.
- Keep it simple – build incrementally, in modules, like AVO's GeoDiva. Have data to populate and test design-more likely to be used
- It is important to focus on the problem and find a solution and not to be concern with software tools.

- Ability to use your own data to compare to a dataset in a database without having to add your data to the database (sandbox) –
- Data permissions- ability to share sample information OR keep it secure until you are ready to share it (publish).
- Catalog of what exists and what is available
    - We have posted a list of tephra databases online
- Proximal sample RDBMS – listing of physical samples that exist for comparison to your sample. Database of occurrences at minimum, so we can request samples (sample ID, quantity of material, etc.). Suggested that all samples are registered in SESAR (system for earth sample registration). SESAR catalogs and preserves sample metadata profiles, and provides access to the sample catalog via the Global Sample Search.
- Global marker horizons would be nice to have in a database
- Is there an existing system that we can use as a framework? Tephrabase, AVO's Tephra Db, RESET?
- Why don't we use Earthchem?
    - We need a simple two-three steps design that immediately help to answer research questions. User should NOT be force into someone else's system vision. Build research tool for the TEPHRA COMMUNITY and arrange how large systems like EarthChem could host and accommodate OUR needs
        - New system needs to work for YOU
    - It is highly possible that systems' like EarthChem could morph to meet our needs. They are great and flexible systems!
- Systems like Earthchem could help to host and design research-specific modules (e.g., geochemistry, regional data systems, grain sizes) provide task specific portals (build user specific) interfaces but under the hood link all data together
- System Name and design Idea: Universal Computer Support System for Tephrochronology UCoSST

Sharing existing data & data manipulation tools

- At this point we could provide a single portal designed to share already build tools/code/software for data processing
    - MS Excel, R, MatLab, Data visualization, stratigraphic sections, geochemistry, geospatial data etc.

Workflow, standardization and process

- Goals and capabilities vs. how do we do it?
- Need a consistent (standardized) work flow - could be facilitated by standard practices/protocols; discipline specific?
- Key terms need to be defined (proximal vs. distal, etc.)
- Data standardization - templates are the way to go (csv or xml) – universal file format
- Workflow diagrams by discipline
    - Physical volcanology has these published (reference needed)
- Deal with other countries/contributions/sharing
- If in an online database – they are available to ALL – no borders
- Best practices in collection, analysis, reporting (publishing) – this could be brief or long and detailed, discipline based – the group wants it to be both

- Research Gate might be a place to advertise checklists
- Chronology
    - More than just the mean values
    - Raw data
    - Report all processing steps (calibration, etc.)
- Constrain physical properties of deposits
- Do you want physical properties in a database – here are some quick field techniques to consider.
    - Use field data to get at ESP (eruption source parameters), and forecast eruptions
    - Systematic approaches for:
        - Plume Height (Carey and Sparks, 1986),
        - Mass Eruption Rate (Wilson & Walker, 1987; Mastin et al., 2009, Degruyter and Bonadonna, 2012)
        - Volume (e.g. Fierstein and Nathenson, 1992; Bonadonna and Houghton, 2005; Bonadonna and Costa 2012)
    - Report on uncertainty and constrain them
    - Seb Biass suggests a workflow (see his workshop presentation for this)
    - TError (Biass et al, 2014)– easy approach to constrain error
- Spatial distribution of tephra deposits
- Include all methodologies for all processing steps
- Need to emphasize protocols for analysis/reporting and make it easy to find, journals and editors, reviewers need to be able to find it and point to it quickly –TEMPLATE. We've been talking about this a long time but people are still not following through – MAKE IT EASIER

Existing datasets & data quality

- How do we deal with regional or goal specific datasets?
- Pre-existing databases that DO NOT talk to each other – point this out-dealing with it! Linkages – how does this work (XML, JASON standards, build libraries and APIs code, etc.)
- Data quality – what do we do? Philosophy
    - Let people know how to identify bad data. Can this be built into a database in a smart way (yes)?
    - Document this issue through well documented case studies (ice vs terrestrial, etc.)
    - Peer review should be a basic cut off for data sharing (must be published first)
    - Include all metadata to make you own decision on data quality