Abstract
Ocean-floor drillings, corings, and dredgings represent millions of years of earth and ocean history distributed along plate tectonic trajectories that may stretch over thousands of kilometers. To correctly locate the data in their context of climate zones and ocean-basin structure, they must be transformed to paleocoordinates along the trajectories, back from the drillsite. This poster describes software to do that transformation on large volumes of sediment, biological and chemical data, suitable for global-scale statistical data analysis.

Introduction
There is a wealth of data on earth and ocean environments in the age-dated materials of the ocean drilling programs, and in corings and dredgings from many other expeditions. They are recorded at their sampling sites of course, but all the deep-time samples have been conveyed by plate tectonic motions to that sampling site. A new combination of software and data allows the such samples to be located to their paleocoordinates for seafloor deep-time data. Data values in the Cretaceous to the Triassic (e.g., von Rad & others 1990). The drillings are somewhat biased to deep-sea scenarios. Fortunately, they can be augmented using dated samples from short cores, dredgings and underwater outcrops (e.g., von Rad & others 1990). The total current in dbSEABED, linked to all sources is 7,196 dated drill sites, all these data was done by institutions such as NGDC, TAMU, and JAMSTEC, requiring very significant time and funding resources.

Application of age-models to the samples was essential for the ocean drilling datasets, but significant time and funding resources. (i) The number of models is very limited. (ii) Considerable work was required to untangle non-standard sample namings (i.e., leg, [] site, hole, mbf) in the various data sources (databases). (iii) It is important to recognize that, in many cases, later phenomena such as consolidation and diagenesis, tectonics and metamorphism will have been imposed at a subsequent time and location. Some parameters are strongly affected, especially physical properties. (iv) The principal source is a 'top-20' set of parameters that address sediment textures and compositions, and also copy data on grain constituents. The total count of dated samples with geologic attributes is currently 1,033,783. The invaluable initial curation of the marine records from cores, drillings, and dredgings are sparse – but constantly being improved. More can be done to add industry data and word-based data. It will also be important to add continental platform data to the set. (Unfortunately deformed continental and oceanic tectonic zones present special problems, yet to be overcome.)

Form of the Input Data
The form of the core sample data is a strong control on the necessary software methods.

• Scientific ocean drilling programs (ODP, IODP) have now drilled over 1300 deep penetration cores into the seabed globally, recovering lithologies of ocean-floor origin dated back to ~220Ma (Trassick, e.g., von Rad & others 1992). The drillings are somewhat biased to deep-sea scenarios. Fortunately, they can be augmented using dated samples from short cores, dredgings and underwater outcrops (e.g., von Rad & others 1990). The total current in dbSEABED, linked to all sources is 7,196 dated drill sites.

• The basic sedimentary-lithologic data accompanying the datings indicate the geologic and environmental conditions at certain times and locations. This type of data is held in databases like dbSEABED and GeoMapApp (Goodwilli & others 2010). In the former the principal source is a 'top-20' set of parameters that address sediment textures and compositions, and also copy data on grain constituents. The total count of dated samples with geologic attributes is currently 1,033,783. The invaluable initial curation of these data was done by institutions such as NGDC, TAMU, and JAMSTEC, requiring very significant time and funding resources.

The Processing Pipeline
• Aggregation of the sedimentary data within a unified and globally extensive database structure. For this project dbSEABED information processing system (Jenkins 1997; Goff & others 2008) performed the necessary harmonization of the formats, units and parameters, and integration of the word-based and numerical data types and fields.

• Identification of the present-day plate based on the sampling location. Two mutually checking methods were employed, based on the plate Static Geometries of Seton & others (2012): (i) spherical geometry point-in-polygons; (ii) raster plate depictions. An unresolved difficulty is the case of sites located in accretionary wedges at consuming plate margins.

• Submitting the collection site information to GPLATES in GML format and computing the paleo-coordinate paths at 1my intervals, back to 100Ma. Site identification could not be carried through the GPLATES process. Positional matching between sampling sites and the reconstruction seed-points was therefore required.

• Assigning age dates to each sample, preserving their serial order in cores, and associating uncertainties. This is far from a trivial operation. Matching published age models (Cervato & others 2009) to data-based site-sample and lab-analysis names is a practical difficulty.

• Align the computer paleocoordinates to the samples. Note that the standard reconstruction path output from GPLATES, in 1Ma steps, is coarse compared to the chronology of the Quaternary but finer than many age determinations older than Miocene.

• To visualize the results, paleogeographic basemaps were obtained from the ‘AgeGrids’ collection of Seton & others (2012). For this exploratory study the sediment attribute data (e.g., carbonate percent) were assigned to gridcells within an X/Y ‘datacube’. This is a good format for data dissemination. In the 2x2dg, 1Ma resolution ‘cubes’ 2,604 cells currently have values.

References