The Grid Rocks

Tectonics, Fossils and users

Mark Mitchell
University of Glasgow

University of Glasgow

GridPP
UK Computing for Particle Physics

Grid at Glasgow

- 2112 Cores
- 1.4 Petabytes of Storages
- 6 GB network Connection
- Primary use ATLAS
- www.scotgrid.ac.uk

Archaeologist on the Grid

- Former Tele-communication Engineer
- Also an Archaeologist
- A GridPP first

University of Glasgow

GridPP
UK Computing for Particle Physics

Non-HEP users and the Grid

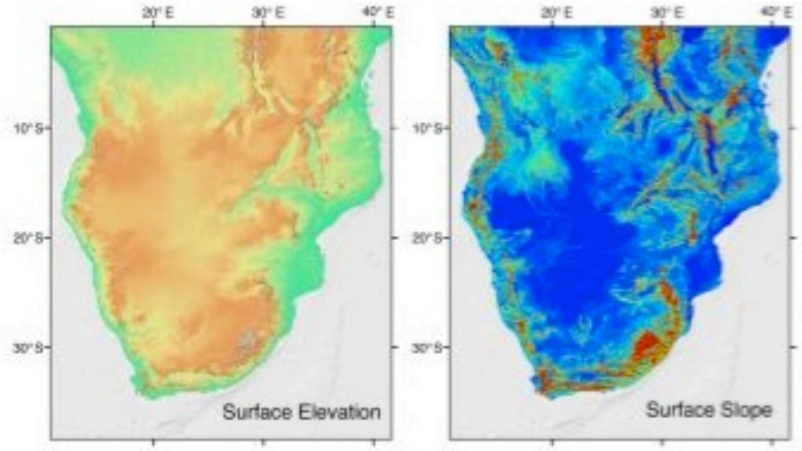
- Utilisation of collaborative environment
- Development of user awareness of the technologies
- Growth of non-physics related user communities with the Grid at Glasgow

Earth Science: Glasgow

- Two Distinct Research Projects on related themes
- Surface processes modelling
  - Dr Romain Buecher
- Paleo Database Analysis
  - Dr Alistair McGowan

Earth Science on the Grid

- Surface processes modelling
  - Utilised Cluster technology elsewhere in EU
  - Aware of the benefits of a collaborative environment
  - Requirement for MPI on the Glasgow Cluster

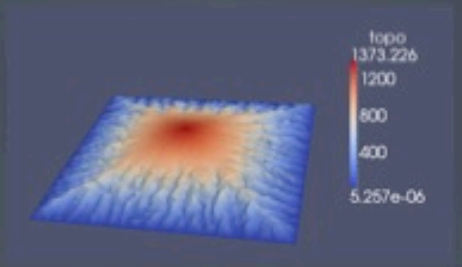


Shaded relief image of southern Africa illustrating the first-order topography characterised by a high elevation interior plateau flanked by steep seaward facing escarpments. b) Image highlighting the steep slopes typical of the major escarpment zone (shown in red) and the low relief of the interior plateau region and coastal plain (blue areas).

Surface processes modelling

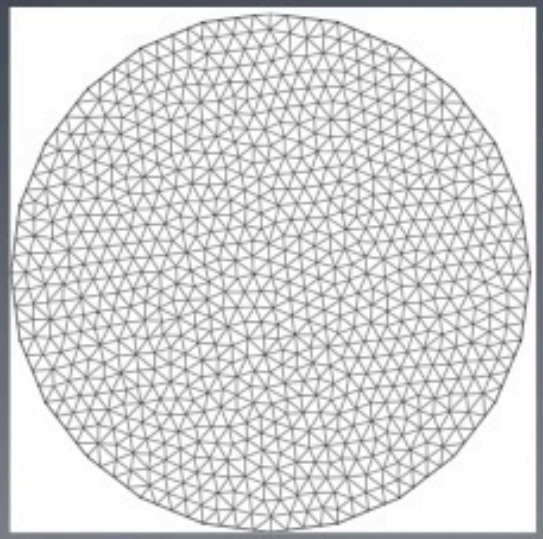
CASCADE (Braun and Sambridge, 1997)

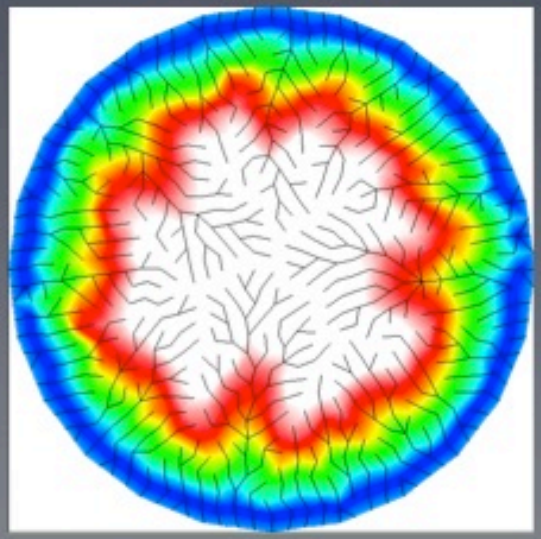
- River erosion.
- Hillslope evolution.
- Sediment transfer.



3000 Km

3000 Km





1 to 10 km resolution
100Myr history

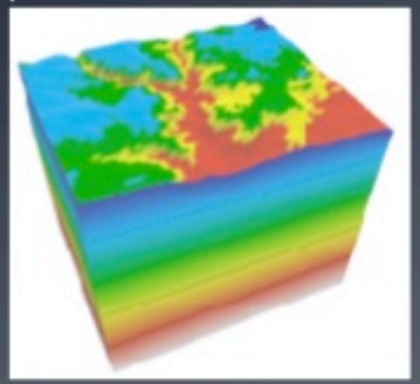
Parallel computation of meshing (for each time step)

Parallel computation of erosion on grid subsets (for each time step)

Typical computing time: Several days to weeks for a million nodes grid depending on resolution)

Surface **Thermal history of rocks: inversion of Fission track and helium ages** modelling

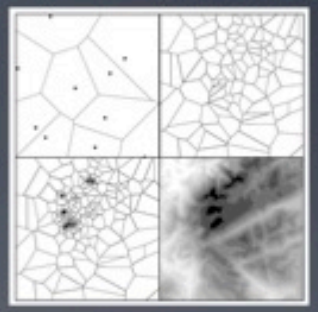
Resolution of heat transport equation in a crustal block.



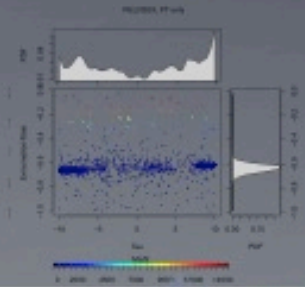
PECUBE (Braun , 2000)

Misfit
(Evaluation function)

NA, (Sambridge et al,1999) Optimisation of parameter space exploration



Information on thermal history, relief evolution etc...

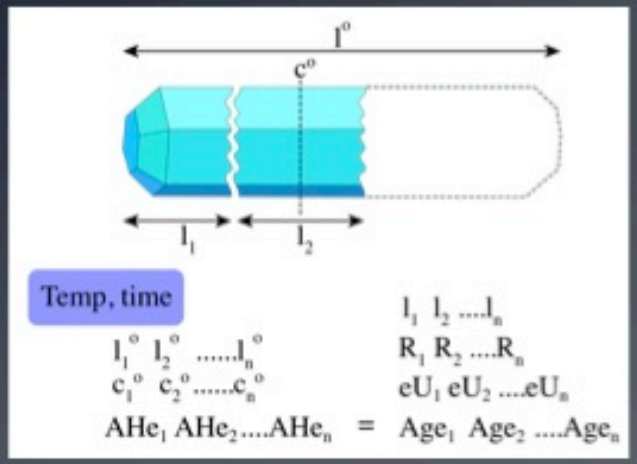


5000 Forward models ran in parrallel through MPI
Typical computing time: hours to days depending on the grid-size... (80 cpus)

University of Glasgow

GridPP
UK Computing for Particle Physics

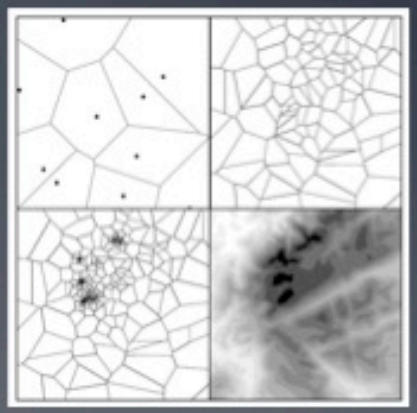
Deriving thermal history information from apatite (U-Th)/He analyses:

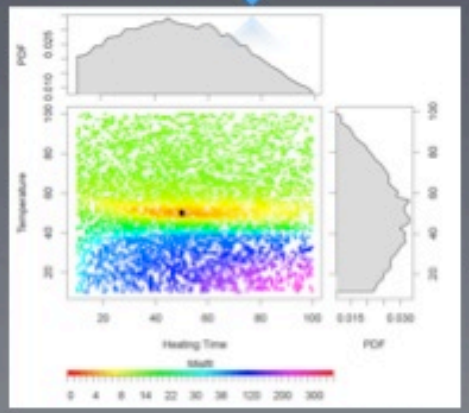


HELFRAG= 3D calculation of helium diffusion in a cylinder.

Misfit
(Evaluation function)

NA: Optimisation of parameter space exploration





5000 Forward models ran in parrallel through MPI

Typical computing time: 2 hours using 20 CPUs

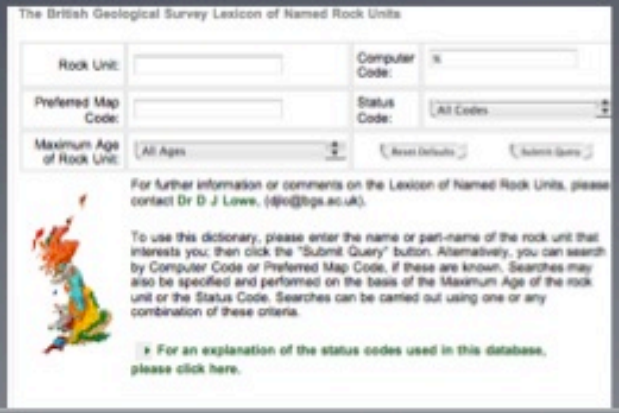
Paleo DataBase

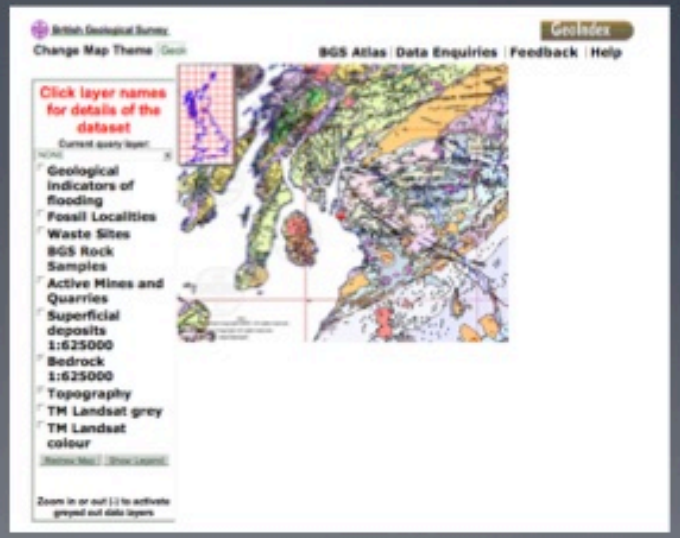
- Web based system: http://www.paleodb.org
- Utilises multi user input through web interface
- Designed to allow analysis of possible prehistoric environmental systems based upon the fossil record

Paleo DataBase at Glasgow

- Read Only Copy of database
- Specifically designed to allow for local Glasgow research
- Investigating the mirroring of the live database with J Alroy et al
- User less familiar with Grid and Cluster technologies

Paleo DataBase





University of Glasgow

GridPP
UK Computing for Particle Physics

Paleo DataBase

Hypothesis: Museum collections are our best sources of data on taxonomic richness, but abundance data need to be recollected systematically in the field.

Tested by:

Comparing collector's curves from museum collections with those from historical and new field collections.

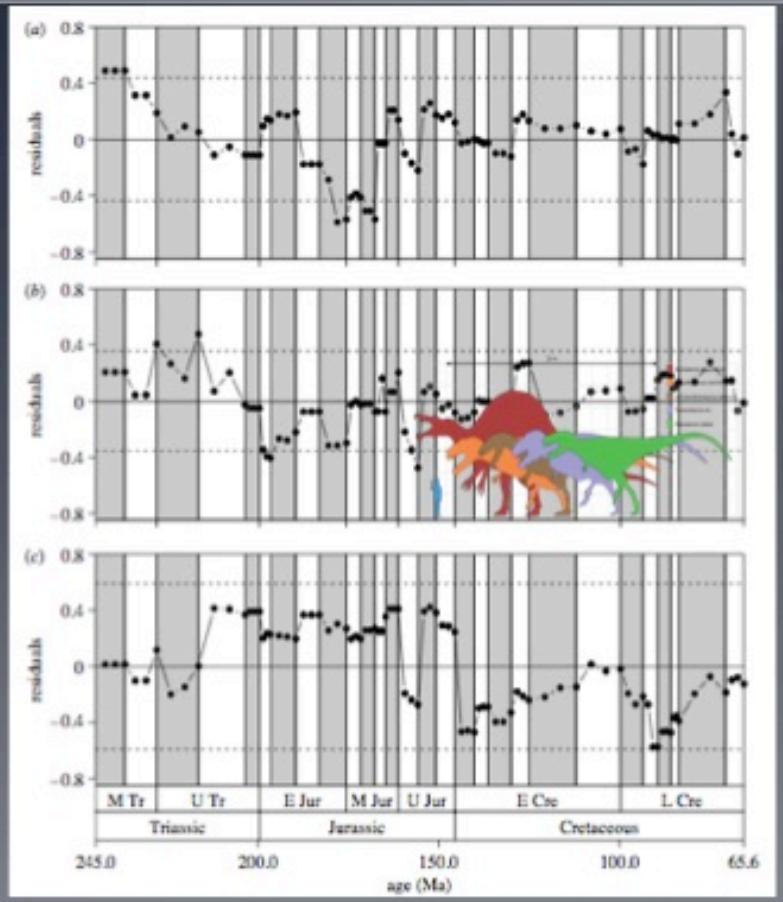
Falsified by:

Field finds of many new taxa after deposits have been known for some time.

University of Glasgow

GridPP
UK Computing for Particle Physics

Residual plots of deviation from a model where ONLY the amount of rock available to sample causes fluctuations in biodiversity (richness)



University of Glasgow

GridPP
UK Computing for Particle Physics

Paleo DB Requirements

- Speed Up Processing of complex data sets
- Investigate 2 year project which was run on single cores and non-grid environment
- Investigate further the potential for analysis on the Grid for this field

Geological Outcomes

- Surface processes modelling
  - Jobs ran successfully and research ongoing
  - Use the cluster at Glasgow to develop a wider geological user community for this field
  - Gain a greater understanding of the technologies and their capabilities and limitations
  - Advance the collaborative environment across the UK and the EU

Geological Outcomes

- Paleo DB
  - Build a wider user community utilising this technology across the UK and EU
  - Investigate the potential of cluster and Grid technologies with a primary dataset which is updated by professionals and the public

User Experience

- Dr Romain Buecheur
  - Used similar systems in the past
  - NA code compiled using Fortran 77
  - Comfortable with the idea of job wrappers and compiling code
  - Comfortable with UNIX and distributed systems

User Experience

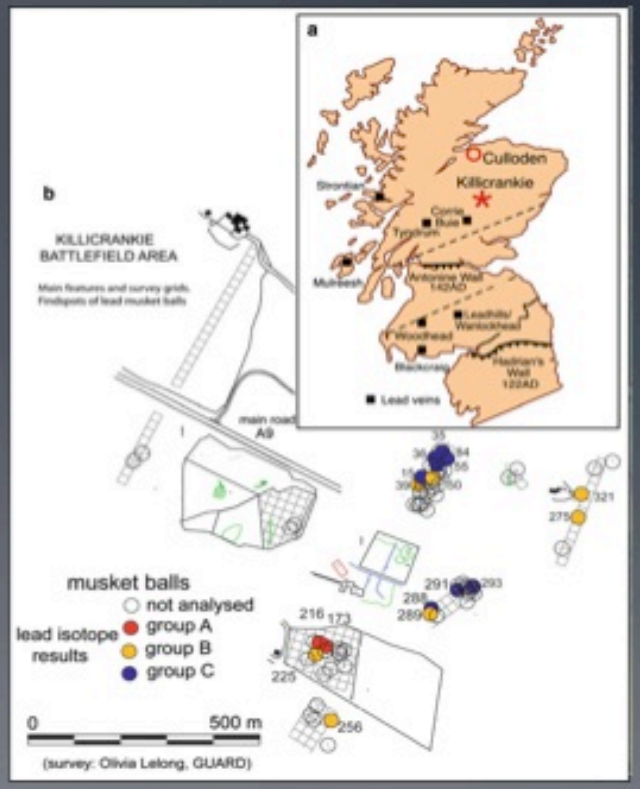
- Dr Alistair McGowan
  - Not familiar with the systems
  - Good R and Python scripts
  - Not familiar with the idea of job wrappers and compiling code
  - Comfortable with UNIX but not distributed systems

Systems Experience

- Complexity around the Paleo DB
  - Database owners at first wary of the cluster install
  - Slow process of building up a relationship with data base owners
  - MySQL install fine
  - User educational experience aided by online help
  - User Accounts and new VO Earthsci.ac.uk

Bigger Challenges

- Archaeological Data
  - Lead Isotope Analysis of musket balls from Scottish Battle fields
  - Building Erosion Modeling combining multiple techniques including LIDAR and material analysis
- Geo-dynamic simulation
  - Highly complex analysis of prehistoric tectonic events



Any Questions ?