JUST-IN-TIME ALTIMETRY:

INTERNATIONAL COLLABORATION IN PROVISION OF ALTIMETRY DATASETS

H.M. Snaith[(1)], R. Scharroo[(2)] and M. Naeije[(3)]

[(1)]Laboratory for Satellite Oceanography, National Oceanography Centre, Southampton, European Way, Southampton. SO14 3ZH, UK;
[(2)]Altimetrics LLC, Cornish, United States;
[(3)]Technical University Delft, Delft, Netherlands

ABSTRACT

Environmental research requires access to quality controlled, calibrated data. Satellite altimeter data are used in a range of environmental research, including oceanography, ice and land surface studies. Users who are not altimeter specialists may not be aware of, or have access to, the latest updates and most appropriate corrections to use for their application.

We propose a GRID based methodology to give all users access to the best possible altimeter data product at the time of the request, tailor made for their specific application. A data portal system would be based on a "Network of Trust" consisting of the data providers and a certificating authority. Data could be served through a fully interactive web 'front-end' or directly from within analysis programmes.

This system would build on the experiences gained in combining two existing Altimeter Data services (GAPS and RADS) to produce a coherent data service with alternative web interfaces and configurable users access.

Keywords: Altimetry, data processing, GRID systems.

1. INTRODUCTION

In many fields of science, there are datasets that undergo regular revision, improvements and enhancement. This is particularly true of satellite altimetry, with its dependence on progress in a wide range of disciplines: geodesy, oceanography, orbitography, atmospheric science, ionospheric modelling and air-sea interactions. The traditional method of controlling data revisions is the release of fixed versions. These versions are then distributed, often on physical media such as CDROM or DVD, or, increasingly, through online systems. There is an inevitable delay between the initial production of a new version and it's general availability, particularly if a large volume dataset needs to be reprocessed. The number of new versions is often limited: to save on reprocessing and production costs; to maintain consistency for a useful period of time; to enable improvements to be grouped to form a single new version of the data set. This leads to further delays in implementing data upgrades. In between versions, new corrections and calibrations may be made available as corrections files or calibration information and algorithms or models through web sites to allow the 'current best' data to be generated. Keeping track of the plethora of such updates for the large number of satellite altimeter products is a non-trivial operation.

Satellite altimeter data are increasingly being used by 'mainstream' researchers for environmental research. In order for data to be used by non-specialists, they are required in easily managed formats, with consistent quality control and appropriate calibration. The 'off the shelf' products, as used by non-specialists, are almost invariably out of date relative to the best products available. In addition, there are several data providers who provide similar products. These different products may be optimised for different applications and there is little ready information on how they compare. In addition, altimeter data may require auxiliary data to get the 'best' product, *e.g.* a regional tide model, and they can be processed by a variety of different optimized algorithms, *e.g.* the interpolation scheme for gridded products.

Altimeter data share a series of processing requirements and complications that are common to many other remote sensing data:

1) Variety of sources for the both raw and processed data
2) Variety of sources for the required auxiliary corrections and choice of correction to apply
3) Variety of processing level, *e.g.* near-real time, intermediate and final high quality, together with improvements in processing giving rise to multiple versions.
4) Variety of sources of data components that can change in time, giving rise to improved versions becoming available. This means that fixed products whether distributed on CD, DVD or available on line through FTP are almost immediately out of date.

Although we here present a possible solution to these issues focused on satellite altimetry, the principles outlined could be adapted to a range of other earth observation data

Proceedings of the Symposium on 15 Years of Progress in Radar Altimetry, 13-18 March 2006, Venice, Italy (ESA SP-614, July 2006)

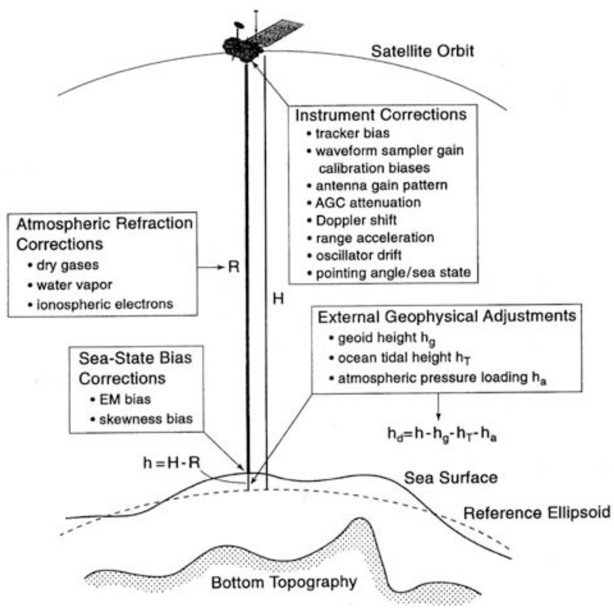

Figure 1 Schematic diagram of a satellite altimeter measurement system (from [1])

Satellite radar altimeters provide three basic measurements (although more can obtained by more detailed analysis): the time taken for a radar pulse to return to the satellite, the returned power of the pulse and the rate of decay of the returned power over time. These measurements are taken along a narrow path below the satellite. Over the ocean, these measurements are used to determine the sea surface height above some reference, the sea surface roughness at microwave length-scales (related to wind speed) and the significant wave height. Over other surfaces, the interpretation of the measurements in terms of geophysical parameters is more complex. The available datasets range from the sensor data records telemetered from the satellite, through the commonly used Geophysical Data Records (GDRs), to the 'improved GDR' products and gridded products produced from these. Differences in products occur not only from the level of processing, but also from the initial re-tracking of the returned radar waveforms, to the selection of an orbit model (itself requiring tracking data from alternative sources), the choice of atmospheric delay corrections, both measured and modelled, tidal models and mean sea surface and/or geoid models. Figure 1 shows the complex system of corrections required to retrieve just one geophysical parameter: sea surface height. The determination of geophysical parameters to sufficient accuracy to allow study of climate variables, such as mean sea surface, and particularly changes in these variables, is at the very limits of the current systems (see *e.g.* [1] and [2]). The expertise to produce the best source models or correction values (*e.g.* tide or orbit model or atmospheric correction) is widely dispersed globally and altimeter processing requires an expert knowledge of what is available, how to apply it and what quality control is necessary at all stages.

In addition, the wide range of research fields where altimetry data are now applied (ocean, land, ice, atmosphere) leads to difficulty in using a single set of predefined parameters. The optimal corrections applied and the quality criteria used, will vary with the application. As alternative altimeter data sources are developed to try and overcome some of these problems, it becomes ever more confusing to compare products and select the most appropriate one.

The proposed data delivery system would allow 'just-in-time' production of tailored altimeter data products, with user selected corrections, using the most appropriate source for the raw altimetry data and each of the data corrections and auxiliary datasets required.

One of the most essential parts of the system is a clearly defined audit trail giving a unique identifier for the final processed data. This identifier would allow the processed data to be reproduced exactly at a future date – either to reproduce an existing dataset or to produce an extension to an existing dataset.

2. A NEW DATA DELIVERY SYSTEM

The Department of Earth Observation and Space Systems of the Technical University of Delft, in collaboration with researchers at NOAA, have developed an 'improved GDR' product called RADS (Radar Altimeter Database System), accessible online at http://www.deos.tudelft.nl/altim/rads/rads.shtml [3]. This system allows for a rolling update of the raw altimeter data – initially from the fast delivery FTP products, later replaced by the high quality offline GDR products. All known corrections to the data are then applied to fields in the database and a consistent set of corrections is added to the database. Some of these corrections are calculated from gridded products, interpolated to the altimeter locations (along the satellite ground track). Other values are calculated using specified algorithms from data included in the altimeter products whilst yet others are calculated using models or algorithms dependent on auxiliary data. The RADS system includes the software to extract data from the database and calculate relevant biases and simple functions from data held in the database. During the data extraction, corrections to the sea surface height are applied according to user-selected criteria, and the appropriate quality control criteria are applied. The system provides default values for all corrections and QC, according to the best knowledge of the experts who set up the system, but any of these can be over-ridden by the user.

At the National Oceanography Centre, Southampton, this output from RADS is already used as the input to the GAPS (Global Altimeter Processing Scheme) data processing system [4]. In order to make the data more useful for oceanographic purposes, the data are collocated to a consistent along-track grid with the calculation of appropriate corrections caused by the cross-track geoid slope.

The primary technological challenge of the proposed system is to combine the functionality of the RADS and GAPS systems, whilst increasing the flexibility to use tailored auxiliary data, models and/or algorithms, through the use of a Grid network of data providers. When data are requested, the system would first provide recommendations for a 'current best' product via a system of 'User lists' supplied by recognised 'Experts'

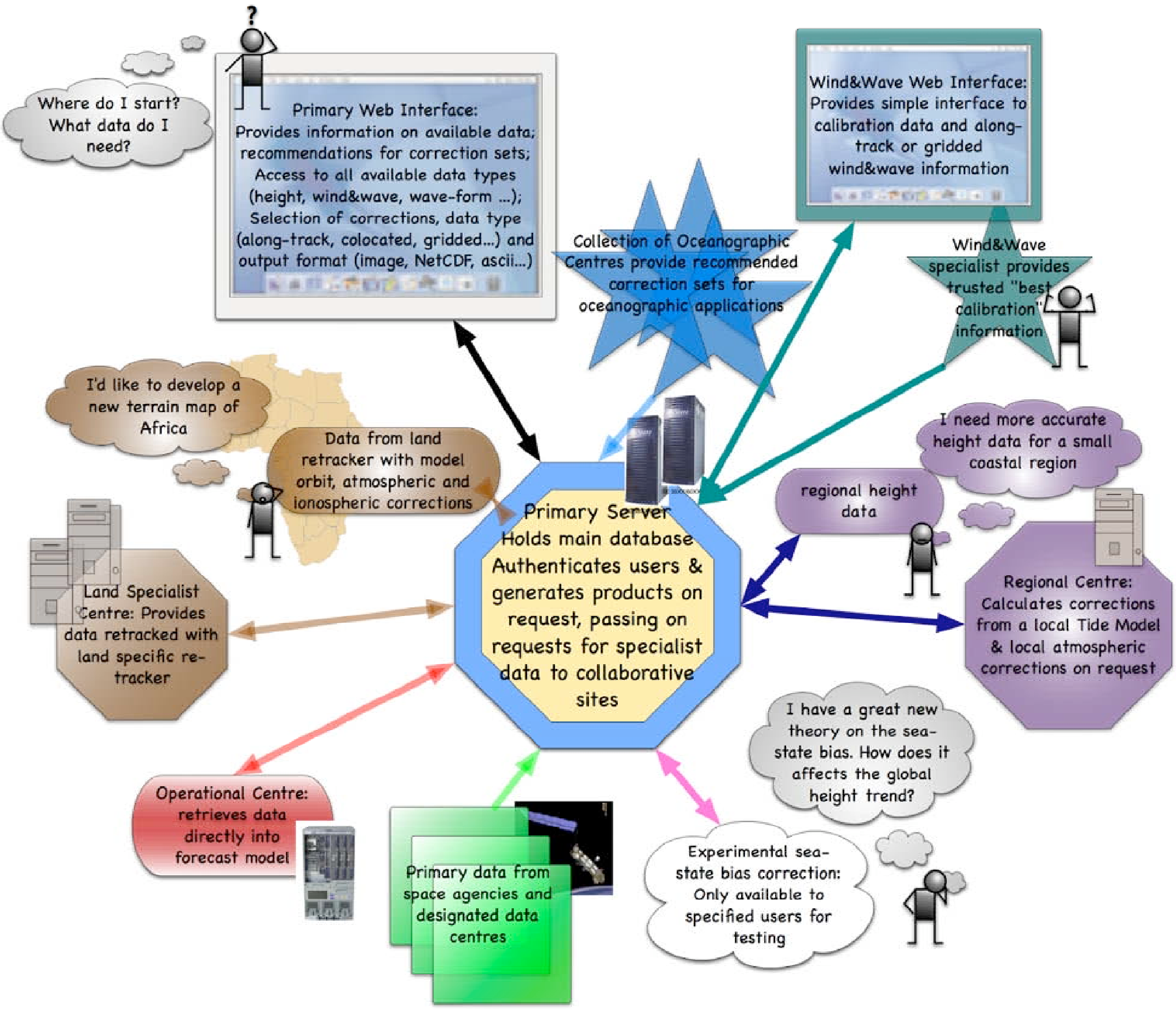

Figure 2 How some aspects of a proposed distributed altimeter system might interact.

that provide guidance on the best correction sets to use. These recommendations may be tailored for a range of applications. The data, both primary and auxiliary, would be extracted and appropriate corrections calculated at the most appropriate site and applied, together with appropriate quality control, before calculation and delivery of the final product to the user.

The system would also incorporate an audit trail, giving a unique product identifier, to allow reproducibility and be able to be overridden by user requirements - *e.g.* to apply a local tide model not currently in the system or to ignore default quality controls to look at extremes.

This system would have clear advantages over the current systems:

- All users have access to the same, high quality product, using the best available data, corrections and models available at the time of data retrieval, without requiring specialist knowledge.
- The user is not limited to the fields that can be feasibly held in a single database (storage limitation).
- Previous datasets can be accurately reproduced, for comparisons with other analysis, by use of the unique product identifier generated by the audit trail.
- There is no delay in implementing improved algorithms and corrections due to compatibility or consistency problems or time taken to re-populate a large database.
- The data storage overhead for the service provider is significantly reduced.
- Only fields actually required for the retrieval need be calculated, *e.g.* in altimetry, many of the corrections are required only for the sea surface height retrieval. If wind/wave products are required, the majority of the processing overhead is not required.

In order to submit data, or 'modules', to be provided through this service, providers would need to be certified. The embryonic International Altimeter Service (IAS) would be the obvious candidate for this role if an operational system were to be developed. The IAS is a proposed service of the International Association of Geodesy, currently in a planning phase, whose terms of reference include "serving the altimeter user community with an utmost long time series of harmonized multi-mission altimeter observations with up-to-date geophysical corrections and consolidated geocentric reference and with related sea level products" (http://iag.dgfi.badw.de/index.php?id=93).

The primary access to the system would be through a web portal that covers the main aspects of the system: selection of correction set and final products type *etc.* However, the interface with the data portal could be tailored to the needs of particular groups of users – *e.g.* wind and wave product users could develop a simplified interface as users would not require access to the full complexity of the system. In addition, it would be possible to access the system directly from within applications – *e.g.* analysis software or a forecast model, using web service interfaces.

The primary server has a key function in maintaining the interactions between the collaborating groups, and providing a central archive, if needed. For example, it could offer a long term archive of corrections no longer used operationally, but potentially useful for testing reliability of analysis systems by reproducing previous results.

3. TECHNOLOGY CHALLENGES

There are clear technological challenges within this proposed system, primarily concerned with designating 'Networks of Trust' for the data providers, allowing the portal system to access data and potentially use remote servers to generate correction sets. A system for certification of 'trusted' suppliers of modules to the system would have to be developed. X.509 digital certificates with GSI (Grid Security Infrastructure) provide authentication, *i.e.* they make a strong statement, to whoever needs to know, that as the holder of a certificate "I am who I say I am". Certificates thus form the basis of trust, in that someone coming in with a certificate should be trusted if you can trust the Certificating Authority that issued the certificate. However, certificates do not do much, of themselves, about authorization, which is handled by the server. In Globus (http://www.globus.org/) there is the gridmap concept which contains information linking authorisation to certificates, but the certificate itself does not carry information about what the user is permitted to access; this comes in at a higher level. Initiating a system in which authorisation may be different for different components of the system is not well researched in environmental *e*-Science. More developments of authorization and its implementation has been carried out within the NERC DataGRID project (http://www.ndg.nerc.ac.uk).

A similar, if potentially simpler, challenge is the development of a users authorisation system for access to each of the multitude of primary and auxiliary data required, *e.g.* ESA altimeter data are only available free of charge to recognised researchers, whilst TOPEX/POSIEDON altimeter data have no restrictions. Alternative data sources and corrections may also have user restrictions and the user's authorisation level will need to be used to restrict the options presented by the interface. The recognition of user limitations can be used to allow the system to be used for distribution of preliminary or novel corrections to a wide audience for validation and testing, whilst withholding them from non-specialist users.

Within this Network of Trust, the service would need to make use of existing toolkits, such as Globus, linked to web and Grid services, to develop a system that will allow calculation of the required correction sets on remote servers, where appropriate. Globus tools would need to be developed to allow access to the required compute resources, *e.g.* to carry out interpolations or calculate a delay correction from atmospheric soundings, but these are not optimised for handling data file transfers. Where data transfer is required, it might be appropriate to operate through existing Grid technologies, such as Storage Resource Broker (SRB), or the suite of 'Live Access Server' type systems.

4. CONCLUSIONS

In this paper, we have outlines a system that allows customised altimeter data products to be produced and delivered to users via a web portal interface. This delivery of highest quality altimeter products, tailored to the needs of individual users would reduce the overhead in processing needed for individuals to gain maximum benefit from altimeter data.

The service would be a data portal that accesses data, corrections, models and a variety of required information from a "Network of Trust". This network would take advantage of the widely distributed areas of expertise required for the successful use of altimeter data across the wide range of applications now common.

Although the proposed system has been described as specific to altimeter data, many of the solutions to technical challenges would applicable to a wide range of science fields, particularly in the field of remote-sensed data provision, but potentially to any field that deals with evolving data sets.

A proposal (AltiCORE - currently under review with INTAS) will be developing a part of this system. Regional specialist centres will be investigation the most appropriate data editing and correction criteria, as well as providing regionally specific corrections (such at atmospheric and tidal) for a series on enclosed and semi-enclosed basins across Europe. These recommendations will be delivered by a distributed, GRID enabled system based on the RADS system.

5. REFERENCES

1. Fu, L.-L. and A. Cazenave, eds. *Satellite Altimetry and Earth Sciences*. 2001, Academic Press. 463.
2. Saxena, N.K., ed. *Marine Geodesy Special Issue on Jason-1 Calibration/Validation Part II*. Marine Geodesy. Vol. 27. 2004. 341.
3. Naeije, M., et al., *Radar Altimeter Database System: exploitation and extension (RADSxx)*. 2002.
4. Snaith, H.M., *Global Altimeter Processing Scheme User Manual: v1*. 2000, Southampton Oceanography Centre: Southampton. p. 44.