MEDSPIRATION : A EUROPEAN CONTRIBUTION TO THE GLOBAL OCEAN DATA ASSIMILATION EXPERIMENT HIGH RESOLUTION SEA SURFACE TEMPERATURE PILOT PROJECT

Ian S. Robinson (1), Jean-Francois Piollé (2), Pierre Leborgne (3), Craig Donlon (4), Olivier Arino (5)

(1) National Oceanography Centre, Southampton, European Way, Southampton, SO14 3ZH, U.K.
(2) CERSAT, IFREMER, BP70, 29280 Plouzane, France
(3) Centre de Météorologie Spatiale, .P. 147, 22302 Lannion Cedex, France
(4)GHRSST-PP International Office, Met Office, Fitzroy Road, Exeter EX1 3PB, U.K.
(5)European Space Agency, ESRIN, Via Galileo Galilei, CP 64 00044 Frascati, Italy

ABSTRACT

The Global Ocean Data Assimilation Experiment (GODAE) has articulated the requirement which operational ocean forecasting models have for improved resolution sea surface temperature (SST) observations using various satellite sensors of different types. The GODAE high resolution SST pilot project (GHRSST-PP) has developed a method and data specification to meet this need. The European Space Agency responded by establishing the Medspiration project to produce such data for the European region. This paper explains and defines the new SST data products. It describes the Medspiration project and system architecture, outlines the successful progress to the present operational phase and reviews the current applications of the new SST data for assimilation into operational forecasting models.

1 INTRODUCTION

1.1 The requirement for high resolution sea surface temperature

Sea surface temperature (SST) is an ocean parameter that is widely used in ocean dynamics, in the study of upper ocean physical and biogeochemical processes, as a boundary condition for meteorological models, as an important factor in studies of air-sea fluxes, and as an indicator for climate and its variation. The records of SST date back 100 years and measurements from satellite infrared sensors have been acquired for nearly 25 years. Several different satellite systems are now in place for monitoring SST continuously, producing updated global maps of SST every few days or monthly averaged composites. Many of these SST products are readily available on the Internet. A casual observer might conclude that the operational monitoring of SST is a mature accomplishment of Earth Observation Science, that sensors in orbit today are more than adequate and the data processing infrastructure is satisfactory; but they would be wrong.

The early success of SST measurement from Space has not kept pace with operational demands based on improved scientific understanding [1]. The required SST resolution (spatial, temporal and radiometric) is steadily becoming finer. The needs of global scale operational oceanography (that is model-based ocean forecasting systems using satellite data for assimilation or validation) were defined by the Global Ocean Data Assimilation Experiment (GODAE). They state that SST observations with a spatial resolution of 10 km and an accuracy of 0.2ºC need to be updated every six hours [2, 3]. In the context of European operational oceanography systems within the framework of Global Monitoring for Environment and Security (GMES) even finer spatial resolution is required, if possible down to 2 km in coastal and regional seas [4]. Climate monitoring demands more accurate measurements of SST, with a target of ±0.1 K absolute.

1.2 The GODAE high resolution SST pilot project

Comparison of the GODAE specification for SST against the capability of individual satellite sensors and *in situ* thermometry shows that no single system by itself can meet the resolution requirements. For example the finest spatial resolution can be achieved by infra-red radiometers but these can never overcome the problem of cloud cover which restricts their temporal sampling capacity. Infrared sensors at present offer the finest radiometric resolution, but the SST calibration approach used over 20 years for NOAA AVHRR data is compromised by dependence on *in situ* buoy measurements which cannot eliminate the uncertainties from unmeasured variable near-surface thermal structure. The more fundamental approach of the AATSR sensor avoids this, its two-view capability providing a robust method for removing atmospheric effects [5], but by itself the AATSR has insufficiently frequent global coverage for some applications. Improvements in SST retrieval accuracy from microwave sensors such as TMI and AMSR-E [6, 7] offer promise of avoiding the limitations of cloud cover, but such sensors are unlikely to provide spatial resolution finer than 25 km for the foreseeable future, remain sensitive to heavy rain and are very noisy within 100-200 km of land. Further improvement in temporal sampling is offered by SEVIRI, the well-calibrated geostationary infrared sensor on Meteosat Second Generation, but this does not give global coverage and is still limited by cloud.

The consensus reached amongst international SST measurement experts is that no single sensor, however well calibrated, can by itself meet the requirements of the diverse and growing SST user community. On the other hand, by combining the data from the different types of sensors and data sources it should be possible to generate SST products which are superior to those derived from a single mission and which approach the specification being sought. This requires international agreement and co-ordination to ensure co-operation between agencies which are, on the face of it, producing competing products. Since 2001 the GODAE High-Resolution SST Pilot Project (GHRSST-PP) has provided a forum in which its international Science Team has worked to define a set of new data products which harmonise the SST datasets from several different providers [8]. In 2003 the GHRSST-PP International Project Office (IPO) was established at Met Office, UK, co-funded by ESA. The GHRSST operational approach is to generate data products in Regional Data Assembly Centres (RDACs), which feed the data into Global Data Assembly Centres (GDAC), the first of which is now established at NASA's Jet Propulsion Laboratory (JPL).

1.3 ESA's Medspiration project

Although from the start of GHRSST-PP there was a strong European input to the Science Team, initially there was no basis for setting up a European RDAC until in 2004 ESA established the Medspiration Project through its Data Utilisation Envelope (DUE) programme. The purpose of the Medspiration SST data service is to fulfil the role of the European RDAC and thereby meet the growing need of operational ocean forecasting models in Europe for a simple, single-point access to all the available SST data being acquired from satellites, in a common format, quality controlled and with error estimates and confidence measures attached. In parallel with RDACs serving other parts of the world, Medspiration produces in near real time a new generation of data products to the GHRSST specification [9], created from the existing high resolution SST data products delivered by other agencies. Formally the function of the Medspiration processing system is threefold:

- To generate a geographically limited set of GHRSST-specified combined SST products in near-real time and to serve them to the GDAC and to European operational ocean models.
- To populate an off-line archive in which to curate them.
- To provide a data product dissemination service for all types of users.

2 THE SST DATA SOURCES USED AS INPUTS

It is not the task of Medspiration to perform any of the original retrievals of SST from raw satellite data. Instead it uses existing SST datasets from a variety of suppliers. Table 1 lists the different SST source products ingested into Medspiration. At present SST data are acquired from four different data providers.

Table 1. SST data products used as inputs to Medspiration

Product	Satellite/Sensor	Input No.	Provider	Resolution	Coverage
ATS_NR_2P	ENVISAT / AATSR	1	European Space Agency	1 km	Atlantic (including Mediterranean)
AVHRR16_G AVHRR17_G	NOAA16 / AVHRR NOAA17 / AVHRR	2	NASA / JPL Physical Oceanography Distributed Active Archive Centre (PO.DAAC)	9 km	Atlantic (including Mediterranean)
AVHRR16_L AVHRR17_L	NOAA16 / AVHRR NOAA17 / AVHRR	3	NASA / JPL PO.DAAC	1-2 km	Eastern Mediterranean
NAR16_G NAR17_G	NOAA16 / AVHRR NOAA17 / AVHRR	4	Eumetsat Oceans & Sea Ice Satellite Applications Facility (O&SI SAF)	1-2 km	European Shelf Seas & Mediterranean
SEVIRI_SST	MSG / SEVIRI	5	Eumetsat O&SI SAF	10 km	Atlantic (including Mediterranean)
AMSRE	AQUA / AMSRE	6	Remote Sensing Systems Inc. (REMSS)	25 km	Atlantic (including Mediterranean)
TMI	TRMM / TMI	7	REMSS	25 km	Atlantic (including Mediterranean

It is important to notice that there are several different types of sensor and satellite orbit represented by the different input datasets. Each of these has different strengths and weaknesses in relation to sampling capability.

First we may consider the sensor type. Inputs 1-5 in Table 1 are derived from infrared radiometers. These operate in wavebands 10 to 12.5 µm and (at night) 3.5 to 3.9 µm. They cannot penetrate cloud, but in clear sky conditions they can resolve in fine spatial detail down to length scales of about 1 km. They respond to the temperature of the ocean skin, within the top tens of microns of the surface. Inputs 6 and 7 are SSTs derived from microwave radiometers operating in the frequency band 6 to 11 GHz. They are less affected by the atmosphere apart from heavy rain. They can therefore penetrate cloud, but their spatial resolution is presently no better than about 50 km, although by over-sampling the SST they are presented on a ¼º lat. × ¼º long. grid. Microwave radiometry detects the temperature in a layer within a few mm of the sea surface, and thus penetrates below the cool skin to record the "subskin" SST.

Input 1 in Table 1 is the ESA Advanced Along-Track Scanning Radiometer (AATSR) which, by virtue of its dual view, onboard black bodies and atmospheric correction strategy, achieves a high absolute accuracy approaching ± 0.1 K which is robust to exceptional atmospheric aerosol loading. Within the GHRSST strategy its ability to provide a measure of SST independent of *in situ* calibration data gives it a unique role, particularly in relation to the production of climate-quality reanalyses, despite its relatively poor coverage which is a consequence of its conical scanning geometry.

Diversity of orbits is also found in Table 1. Inputs 1-4 and 6 are all from polar sun-synchronous orbits, which give global coverage, but sample the same location at best twice per day. Input 5 is from a geostationary orbit, allowing sampling ever hour or more frequently, but limited to a fixed geographical coverage. Input 7 is from a non-sun-synchronous polar orbit which, over time, is able to provide more complete sampling of the diurnal cycle.

By harmonising the data from all sensor types, the weaknesses of each can be mitigated and their benefits combined to provide a more comprehensive SST sampling capability. Moreover the independence of the different SST products leads to improved confidence in the error estimates that are needed when data are assimilated into ocean models.

3 THE GHRSST-SPECIFIED OUTPUT PRODUCTS

Medspiration produces several output products, each adhering to the specification laid down by the GHRSST-PP [9].

3.1 Level-2 Pre-processed data products (L2P)

The primary output from Medspiration is the pre-processed (L2P) product This product is designed to facilitate assimilation of several different level 2 (L2) SST sources into operational models or SST analysis systems. Each L2P output corresponds to an input L2 dataset, containing the same temperature fields on the same sampling grid as the input. The extra processing that makes it an L2P product is to attach to every valid SST pixel an error estimate and confidence flag, along with a number of ancillary fields corresponding to the time of the SST measurement. These include the wind speed, the solar irradiance at the surface, the sea ice cover (if relevant) and deviation from a reference standard or climatology. The L2P product does not blend or merge SST data from different sources. It does however present all the data in a common netCDF format. L2P data are sent in near-real time to the GHRSST GDAC and to selected operational users in Europe for assimilation into ocean forecasting models. They are also archived and served more widely to general European users of SST.

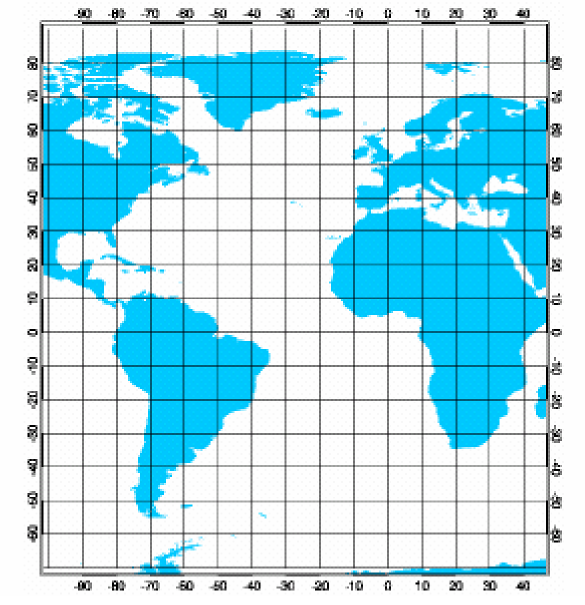

Figure 1. The Medspiration L2P product domain (Note that the Pacific areas are not included).

Figure 1 shows the domain covered by Medspiration products, from 70ºS to 90ºN (to the ice limit) and from 100ºW to 45ºE. It contains the whole Atlantic and adjacent European seas, but excludes the Pacific Ocean. The AVHRR L2 data sources used (see Table 1) have a resolution of 10 km over much of the Atlantic, but finer resolution data are used over European waters and the Mediterranean. The AATSR data are handled at their full resolution of 1 km throughout the Atlantic. In the near future the Medspiration Service will generate L2P products from the global AATSR dataset.

3.2 The Match-up Database (MDB)

Available *in situ* measurements of SST are matched with coincident samples from the L2P SST data, in order to prepare records for delivery to the GHRSST-PP match-up data base (MDB). This is required by GHRSST-PP to determine the sensor-specific error statistics (SSES) needed when assigning accuracy estimates for the L2P data. The geographical scope of the MDB records is the same European area as defined for the L2P products. By intercepting the *in situ* data as soon as they become available in data centres such as CORIOLIS, and then attaching them to cloud-free pixels of satellite data acquired within a given narrow spatial and temporal proximity, a wide set of matched data can be assembled for all sensors. The MDB is an essential preparation for operational assimilation of SST data. Without the MDB it is not easy to assign reliable error statistics to individual pixels, as required by the models.

3.3 Level 4 Analysed SST Product (L4)

The Medspiration L4 data product is an analysed SST field which is derived from all of the available L2P outputs in order to produce a temperature field blended from all sources, weighted according to the characteristics and quality of each input value. GHRSST-PP has defined the "foundation" SST as the temperature at the highest point in the water column each day that is not affected by diurnal warming or the thermal skin effect. It is effectively the same as the subskin temperature at dawn. It corresponds to the underlying surface temperature on which solar heating may build a diurnal thermocline, and is closest to what a typical ocean circulation model would represent as SST. To generate the L4 product requires that each input L2P value should first be converted to the foundation SST by estimating the magnitude of any diurnal warm layer and, in the case of infrared data, the cool skin effect. This is facilitated by the availability of the ancillary data within the L2P product. Then an optimal interpolation scheme is applied to merge data from different sources and to fill gaps caused by cloud. This makes use of the error estimates contained in the L2P products. The resulting L4 product, evaluated daily, corresponds to the foundation SST. Medspiration produces L4 data on a 2 × 2 km grid (see Figure 2) for the Mediterranean Sea only, as a demonstration of the GHRSST principles.

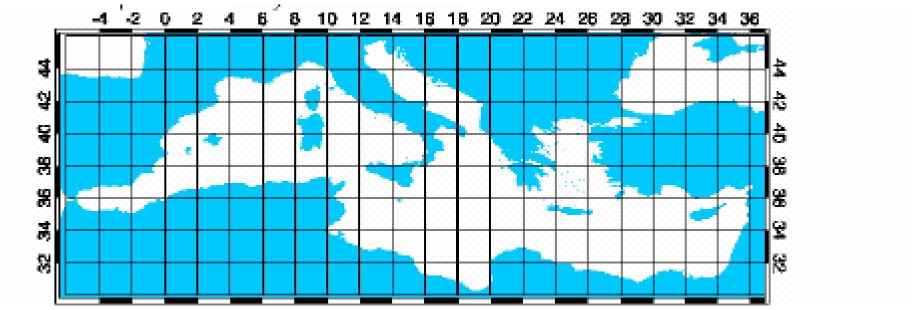

Figure 2. The Medspiration L4 product domain (30N to 46N and 6W to 36.5E at 2 km spatial resolution).

3.4 The High-Resolution Diagnostic Data Set (HR-DDS)

The purpose of the GHRSST-PP high resolution diagnostic data set (HR-DDS) is to facilitate intercomparison of the characteristics of the SST derived from different sensor types, underpinning the research needed to test and improve the data merging methods used in GHRSST-PP. A number of DDS sites are specified, distributed around the world ocean, and typically 2º × 2º in extent. Within Medspiration, for each of these sites data are extracted from every L2P and L4 data product produced by Medspiration, for which valid SST pixels overlap the site. The extracted data are resampled by nearest neighbour substitution onto a 0.01º grid, irrespective of the L2P pixel size, archived and made publicly available. While the DDS contains no information additional to the L2P products, the extracts are more manageable for analysing SST differences between sensors and identifying any dependence on location, season or ancillary data.

4 CREATING THE MEDSPIRATION SYSTEM

4.1 The Medspiration Consortium

The team which was formed to set up the Medspiration system is summarised in Fig. 3. The project was managed by the National Oceanography Centre, Southampton (NOCS, formerly the Southampton Oceanography Centre). The system was designed and implemented by a core group consisting of the Centre Météorologie Spatiale (CMS) of Meteo France in Lannion, the Laboratory of Oceanography from Space at Ifremer in Brest, the Space Oceanography Division of CLS in Toulouse and the Laboratory for Satellite Oceanography at NOCS. Vital contributions were also made during the development phase by the Gruppo di Oceanografia da Satellite at CNR in Rome who provided expertise on applying optimum interpolation schemes to. SST fields in the Mediterranean, the French company Actimar (formerly Avelmor) providing software support and Vega Ltd. who gave broad-based management support and took responsibility for Quality Assurance. CNR and Met.no also served as Test Users during the development phase. During the second phase of Medspiration the operation of the Medspiration Service is the responsibility of NOCS, Ifremer and CMS/Meteo-France, with support from Actimar for some specific software upgrades.

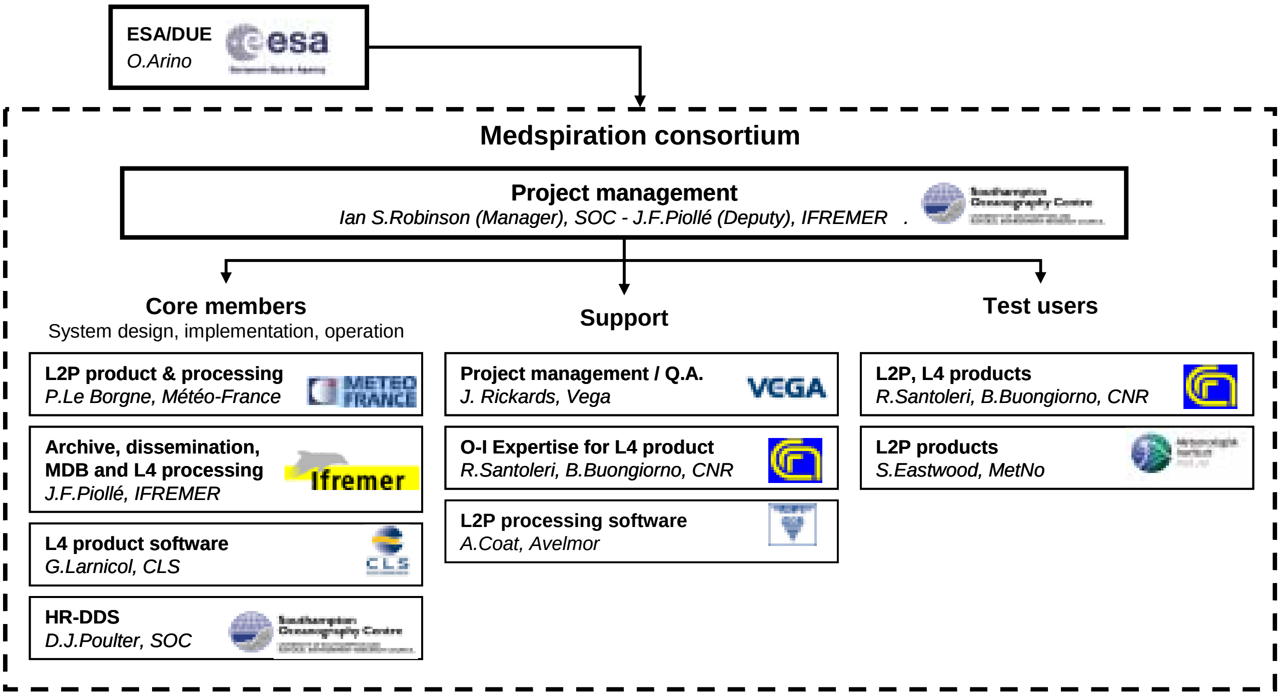

Fig. 3. The organisation established to create the Medspiration System.

The GHRSST IPO developed the User Requirements Document for Medspiration, with input from a wide range of European users of SST data, and worked alongside ESA to oversee the development phase. There is a two-way interaction with the GHRSST-PP. Medspiration products have always been tied to the processing rules laid down by the GHRSST Science Team [9], while the practical experience of developing Medspiration has enabled several of its key players to make valuable contributions to Science Team discussions about those processing rules.

4.2 Medspiration System Architecture

Figure 4 outlines schematically the processing tasks required to produce the four types of SST products described in 3. **Error! Reference source not found.** It shows the main classes of inputs and outputs, and assigns distinct processing stages to different boxes. The system implementation is distributed across the facilities of three contractors, CMS, Ifremer and NOCS (SOC), as indicated in the figure. The underlying data flow through the system is basically simple. Level 2 SST data are obtained from their producers by Internet and ingested into the first main processing centre, hosted in CMS / Meteo-France. This takes advantage of the co-location with the Eumetsat O&SI SAF which produces inputs 4 and 5 in Table 1. The L2P processing facility at CMS also ingests the necessary auxiliary satellite data and evaluates the quality indicators.

As soon as each L2P data set is generated in near real time the products are pushed to the core processing centre at Ifremer, where they are archived, and disseminated to GDAC and to users by ftp push or pull. They are also made available by OpenDAP. Another major process at Ifremer is to use the L2P data in the production of the L4 interpolated product. The OI process is run once per day, during the morning, using all the available L2P data from the previous day's satellite acquisitions. Thus the L4 product for day N is made available to operational users by midday on day N+1. The remaining major task performed at Ifremer is to generate entries for the GHRSST-PP match-up data base (MDB). Because the focal point for the collection, processing, dissemination and archiving of many of Europe's operational *in situ* measurements of SST from buoys and other platforms is already based at Ifremer in the CORIOLIS project, the identification of matches between these data and the L2P datasets is achieved speedily and efficiently.

The final part of the processing chain is located at NOCS, where the DDS granules are extracted from every L2P and L4 dataset, archived and made publicly available. During Phase 1 and part of Phase 2, NOCS has also been responsible for downloading the globally complete near-real time AATSR SST L2 products from the ESA server, expanding the incomplete ancillary fields and reformatting in netCDF to be pulled by CMS for ingestion into the L2P processor (Input 1 in Table 1). However, during Phase 2 the AATSR L2P processing will be expanded to global coverage, and the complete AATSR chain (ingestion from ESA and L2P formation) will be implemented at Ifremer.

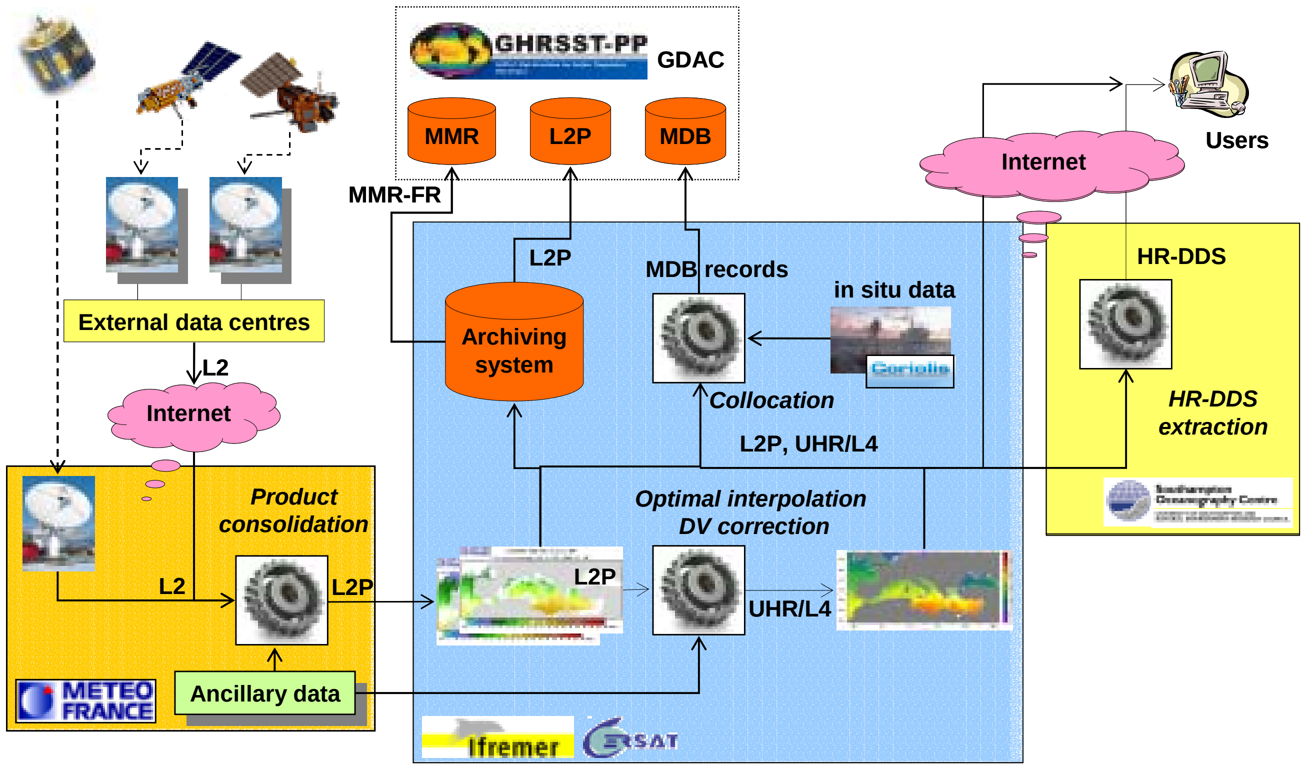

Fig. 4. Schematic of the Medspiration processing chain

5 ACHIEVEMENTS OF MEDSPIRATION

5.1 Project Milestones

The Project commenced in January 2004 and the System Requirements Definition was completed in March. Following software development and implementation the system Qualification Review was passed in November. By this stage it was confirmed that all the data sources were accessible, the data throughput rates were comfortably achievable and the modular system architecture was successful. In fact it was soon apparent that the approach adopted, that of processing L2P data as and when the input L2 datasets became available, is very robust. If one data source is not available, there is no impact that delays the processing of other sources. Compromises between the quality and timeliness of ancillary data had to be found so that this would not unduly delay the L2P process.

In February 2005 a one month long Beta Test was run in order to prove the system's ability to run continuously and to evaluate the data delivery mechanisms. Several organisations using SST data in support of operational ocean forecasting activities were signed up as Medspiration Users for the Beta Test. The input of the users was an essential element of the successful Acceptance Review (AR) of Medspiration in March 2005.

Since the Beta Test the system has continued to operate up to the present. Formally, phase 2 of Medspiration commenced on 1st June, 2005 with the emphasis on operational delivery of data. Minor modifications to the software in response to the AR are logged on the Medspiration website (www.medspiration.org) so that users of the archive can track the history of the products. Medspiration is now scheduled to continue in operation until June 2007 within the original funding envelope, 17 months longer than originally envisaged in the Medspiration Tender Invitation.

Typically the system is now producing about 2600 L2P files per month. The typical time elapsed between ingestion of L2 data at CMS and the availability of L2P products in the Ifremer archive is less than one hour. Differences between the ages of data arriving in the archive are mainly due to the delays of delivery by source providers. Around 70,000 entries to the MDB are made per month from buoy data and another 20,000 entries from other sources of *in situ* data.

L4 products are produced for the Mediterranean once per day. The production run commences at 06:00 and typically takes less than one hour, in which case the L4 analysis is available five hours before the deadline of 12:00 on the day following the measurements to which it refers. Because the emphasis of the project in phase 2 has been refocused on extending the supply of L2P data in order to support operational users (see 5.2) the L4 process has not been as fully optimised as originally intended, but its regular and timely delivery continues to prove the practicality of the GHRSST approach to merging global SST data.

5.2 User Response

The Medspiration project has been driven throughout by the requirements of users. At first this was represented primarily by the GHRSST IPO acting to federate the requirements of a wide range of operational users of SST. GHRSST has continued to provide the formal means by which Medspiration can interact and harmonise with parallel activities taking place in other RDACs around the world, ensuring that Medspiration products are internationally recognised, and allowing the experience of the Medspiration team to be fed back into the international forum. The formal establishment of the GDAC since Medspiration was set up has provided a metadata repository and data archive that is global in scope.

Moreover, since the Beta Test, Medspiration has also acquired a set of registered institutional users of its products, whose experience and opinions are sought regularly. Their feedback is built into the ESA's ongoing supervision of the project. Following the recommendations of the users to the AR, ESA decided to refocus Phase 2 on extending the supply of L2P data for an additional year. They were persuaded by the argument that operational users need a guaranteed stability of supply to make it worth the effort of adapting their operational systems to use Medspiration products. At present, the Met Office in UK is making a parallel run of its forecasting ocean assimilation model (FOAM) which assimilates Medspiration L2P data, and is also developing a new SST analysis which uses L2P as input, in support of numerical weather prediction. The Norwegian and Danish meteorological services are also making trial use of Medspiration products. The French MERCATOR ocean forecasting model is preparing for the assimilation of Medspiration L2P data.

This change in emphasis in response to user demand has meant that the original intention to spend effort on refining the configuration of the optimal interpolation scheme for the L4 processor has not been possible. Neither has it been possible within the funded Medspiration activity to make a fuller analysis of the diurnal variability correction schemes proposed by GHRSST. However, these tasks are now being pursued in parallel to Medspiration by L4 data users such as Ifremer and the EU Framework 6 MERSEA Integrated Programme. The improvement of the analysis schemes is considered to be a more open-ended scientific activity that is better managed through MERSEA, in collaboration with Medspiration. In support of this, Medspiration is making some modifications to the L4 processor to extend the domain beyond that illustrated in Fig. 2 to cover a wider area matching that needed by the Mediterranean Forecasting System.

Finally, in response to the requirements expressed by GHRSST, Medspiration is expanding the L2P coverage of the AATSR product to the global domain, which will make it easier for AATSR data to be used by other operational agencies worldwide.

7 REFERENCES

1. Robinson, I. S. and Donlon, C. J. Global measurement of sea surface temperature from Space: some new perspectives. *Global Atm. Ocean System*, **9** (1-2), 19-37, 2003.
2. Le Traon, P.-Y., Rienecker, M., Smith, N.R., Bahurel, P., Bell, M., Hurlburt, H. and Dandin, P. Operational oceanography and prediction, a GODAE perspective. In: Koblinsky, C. and Smith, N.R. (Eds.), *Observing the Oceans in the 21st Century*, pp 529-545, GODAE Project Office, Bureau of Meteorology, Melbourne, Australia, 2001.
3. Smith, N.R. and Koblinsky, C. The ocean observing system for the 21st Century, a consensus statement. In: Koblinsky, C. and Smith, N.R. (Eds.), *Observing the Oceans in the 21st Century*, pp 1-25, GODAE Project Office, Bureau of Meteorology, Melbourne, Australia, 2001.
4. Robinson, I. S. and Cromwell, D, D. *Observational Requirements for inputs to European Ocean Forecasting System Models*. Report D3.1 of EU FP V project "Marine Environment and Security for the European Area - Strand 1 (MERSEA-S1)", Nansen Environmental and Remote Sensing Center, Bergen, Norway, 57pp, 2003.
5. Merchant, C.J., Harris, A. R., Murray, M. J. and Zavody, A. M. Towards the elimination of bias in satellite retrievals of sea surface temperature 1: Theory, modelling and inter-algorithm comparison. *J. Geophys. Res*., **104** (C10), 23565-23578, 1999.
6. Wentz, F. and Meissner, T. AMSR Ocean Algorithm (Version 2), Algorithm Theoretical Basis Document (ATBD), Remote Sensing Systems, Santa Rosa, CA, Revised edition, November 2, 2000.
7. Donlon, C. J., Gentemann, C. and Wentz, F. Measuring sea surface temperature with microwave sensors, *Backscatter*, **12** (2), 37-39, 2001.
8. Donlon, C. J. *Proceedings of the 6th GHRSST-PP Science Team Meeting*. Document GHRSST/21 (available from GHRSST-PP Project Office, Met Office, Exeter, UK, or on line from www.ghrsst-pp.org), 2005.
9. Donlon, C. J. *GHRSST-PP Data Processing Specification (GDS Version 1, revision 1.5).* Document GHRSST/17 (available from GHRSST-PP Project Office, Met Office, Exeter, UK, or on line: www.ghrsst-pp.org), 2004.