ALBiUS Description of Work

Work package number	WP6	Start date or starting event:				month 1				
Work package title	Advanced Long Baseline interoperable User Software									
Activity Type	RTD									
Participant id	4 JIVE	1 ASTRON	18 UCAM	15 ESO	5 MPG	23 NRAO	19 UOXF	6 UMAN	20 BORD	Total
Person-months per beneficiary:	**26**+18	**13**+13	**12**+10	**18**+12	**6**+0	**12**+9	**12**+9	**15**+12	**12**+9	**126**+92

Objectives

The objectives of ALBiUS (Advanced Long Baseline interoperable User Software) are to develop key algorithms required for the successful exploitation of the upgraded and new generation of RadioNet telescope facilities (e-MERLIN, LOFAR, APERTIF, ALMA etc). These new telescopes will result in an explosion of data rates, and an expansion in the continuum spectral window of one to two orders of magnitude. ALBiUS will produce new software systems and algorithms that are designed to meet these challenges. The focus will lie in the production of new algorithms that address issues of calibration (both in the uv and image plane) and sky modelling. The need for identifying bad data and the issue of data quality control in general, will also be addressed. In addition, ALBiUS aims to make good use of existing software packages - the goal is to make these algorithms available in a modern, distributed computing environment, and to provide transparent interoperability between the different software suites. The latter will encourage a more unified approach to software development in radio astronomy across Europe and beyond.

Description of work

Task 1 Interoperability

Portable Algorithms

A pilot project is carried out to transparently process a dataset using software from multiple packages. There is scope to exploit the commonality in the Python user interface of AIPS, Casa, MIRIAD and GILDAS to implement this. The major issue is the variation in the data formats and calibration models used by the different packages. A simple approach is to convert between the different data formats using an intermediate data format, which is already accessible to all the packages, such as UV FITS. An alternative approach would be to adopt HDF5 as the basis of an interoperable data format. The interoperable packages would then either have to be made HDF5-aware, or else be provided with a conversion routine from HDF5 to their native format. Using HDF5 has the added benefit of giving access to additional HDF5 tools not found in the main interferometry packages, in particular for 3-D data visualisation. This work will be carried out at JIVE and ESO.

Distributed ParselTongue

ParselTongue was developed by the ALBUS project as a Python interface to classic AIPS. This is the main vehicle for making the algorithms developed in ALBUS available to the user community. It has also proved an extremely effective tool for pipeline data processing, and has been incorporated into the production environment of the EVN and MERLIN arrays. In addition, it is used by an increasing number of astronomers for processing datasets which are either large in size or consist of repeated observations, each requiring similar processing. The current ParselTongue functionality will be developed in the context of interoperability with other packages. Given the continuing dependence on AIPS of a large part of the radio astronomy community, even in this era of ever increasing data volumes, there is a necessity for enhancing ParselTongue to allow better exploitation of AIPS on a distributed computing environment. Some initial progress has already been made within ALBUS on creating an infrastructure that allows data distribution for parallel processing, but there are many issues which require more atomic procedures than currently available in the AIPS suite - these problems need to be addressed. This work will be carried out at JIVE.

Task 2 - Calibration algorithms

Global Fringe Fitting

Global fringe fitting is a crucial calibration step for EVN and e-MERLIN, the extended (European) LOFAR baselines and the highest ALMA frequencies, where the residuals to the correlator model of the station-based delays must be determined. The algorithm for this is currently only implemented in AIPS, but will have to be incorporated into the new software packages for these arrays. In addition, the current algorithm produces results which are difficult to interpret in an automated way due to peculiarities of the weighting scheme and resulting anomalies in the reported signal to noise ratio. The procedure is complex and there are many pitfalls for both the user and the potential developer. Alternative methods to the non-linear least squares fit (Levenberg-Marquardt algorithm) used in AIPS are also likely to produce more robust results, less susceptible to converging on local minima. The current fringe-fitting algorithm will be evaluated and based on this evaluation an improved version will be provided. This work will be carried out at NRAO, Manchester and JIVE.

Image Plane Calibration

In ALBUS, methods were developed for calibrating the effects of ionospheric and tropospheric fluctuations on long baseline interferometry data. These methods have been implemented in AIPS with some success. A limitation, however, is that the AIPS calibration model assumes that a single calibration factor is applicable across the instantaneous field of view of the interferometer. This assumption does not hold for the large fields of view and/or large fractions of the primary beam which will be observed by LOFAR, ALMA and APERTIF and which are becoming available for the EVN, e-MERLIN and EVLA as the correlator capabilities of those instruments are enhanced. Polarisation imaging over wide fields of view presents additional challenges for all these instruments. In particular, the polarisation response of the dipole arrays used by LOFAR will have strong direction dependence.

These image plane effects can be corrected using some of the more modern calibration packages, which are currently being developed, but the algorithms will require enhancements to the calibration model to incorporate the required direction dependence. At the same time, some related calibration issues will be addressed, such as correcting for station primary beams, and using mosaicing techniques to produce images covering areas of the sky much larger than is achievable with a single pointing, the latter being particularly important for ALMA with its relatively small instantaneous field of view.

High spectral and temporal resolution is required to achieve large fields of view as well as to allow high quality spectroscopy, and the study of transient phenomena (both astronomical sources and contaminants such as RFI). This implies very large data sets: 100's of GBytes for EVN and several TBytes for LOFAR. Efficient processing of such large data sets requires access to a distributed computing environment and therefore existing algorithms to enable this need to be modified. This will also require investigating the support for distributed computing in the chosen calibration packages. Solving this issue will be important for LOFAR and ALMA, but suitable test datasets from GMRT, Westerbork and the EVN already exist. This work will be carried out at ASTRON and Oxford, Manchester, Cambridge, NRAO and ESO.

Calibration of Astrometric Source Positions

Having highly accurate source positions is essential for almost any astronomy application, and certainly for high-resolution interferometric imaging of weak targets. At present, astrometric measurements are carried out with two different techniques: wide-angle astrometry based on the total interferometric phase response (for major calibrator sources distributed on a 5° grid on the sky) and narrow-angle astrometry based on relative phase measurements (for sources located between the major calibrators). There are ways to develop innovative algorithms and observing techniques that will combine the two approaches to produce an improved, denser, and unified celestial reference frame comprising all sources (whether observed with wide- or narrow-angle astrometry). Ultimately, these new methods should serve as a basis for conducting deep astrometric surveys with instruments such as the EVN, e-MERLIN (or the geodetic IVS2010 network). Algorithms such as those used in traditional geodesy to adjust positions of second-order geodetic markers into a first-order grid of markers may be adapted to realise such a global adjustment. The multi-beam capabilities of future instruments such as the SKA will be taken advantage off. The work will include studies of potential algorithms, simulations to test these algorithms, and analysis of test data acquired with the EVN and e-MERLIN to validate observing strategies. This work will be carried out at Bordeaux in collaboration with JIVE and Manchester.

Task 3 - Tools for Large Datasets

Automated Data Quality Control

The new generation of interferometers currently coming on line (LOFAR, eEVN, e-MERLIN, EVLA, ALMA) will have vastly increased output data rates compared with current instruments. Traditional methods for data inspection will become impractical for these new instruments. It is therefore necessary to develop new techniques, such as subspace decomposition, to allow automated identification (and either correction or excision) of artefacts resulting from poor calibration, poor atmospheric or ionospheric conditions, man-made interference or other (instrumentation) problems. Some methods to mitigate RFI are known to be very effective, but are not yet implemented at many radio observatories. The usefulness of existing algorithms will be investigated and to make them available in interferometry, where appropriate. The most useful algorithms would then be implemented in one (or more) of the mainstream interferometry analysis packages. This work will be carried out at MPG, ASTRON, Cambridge and Oxford.

Source Parameterisation

One of the primary scientific deliverables of LOFAR (and other SKA pathfinders like APERTIF) will be huge catalogues of all the detected sources. This so-called Global Sky Model will also provide the basis for the calibration, as a good sky model will be required to make tractable the complex calibration needs of these instruments. Most sources will be unresolved, so simple point-source models can be used, but a significant fraction will be extended. The vast size of the Global Sky Model means that new and innovative methods must be developed to:

1. automatically extract from the data an accurate description of the sources in a computationally efficient way,
2. store the source descriptions in a form which is compact and yet retains sufficient information to allow an accurate reconstruction of the objects described in it.

Traditional methods for doing this have involved describing sources in terms of collections of point sources (CLEAN components) or as collections of elliptical Gaussians. However, these methods are neither very compact, nor can they be easily extended to incorporate a complex time or frequency dependence. They do not provide sufficient dynamic range for instruments such as LOFAR or APERTIF. A possible solution is to use techniques based on shapelets or pixons. Such techniques can describe extended sources with relatively few parameters to high accuracy and, given their continuous nature, do not suffer from the problems caused by the discreteness of CLEAN-component models and will be able to achieve much higher dynamic range. An optimal means of producing a Global Sky Model based on these techniques will be produced. This work will be carried out at ASTRON.

Deliverables (brief description)

6.1.1 Final report on calibration of pilot experiment using interoperability framework – Month 21

6.1.2 Release of distributed ParselTongue - Month 21

6.2.1 New implementation of Global Fringe Fitting algorithm – Month 36 6.2.2 Direction dependent ionospheric, tropospheric calibration to test data set – Month 21 6.2.3 Software for mosaic imaging including primary beam correction – Month 25 6.2.4 Report on image plane polarization calibration effects – Month 19 6.2.5 Final report on the implementation of algorithms for image plane calibration in a distributed environment – Month 30 6.2.6 Final report on new algorithms and observing strategies for astrometry – Month 28 6.3.1 RFI mitigation software – Month 19 6.3.2 Final report on data quality algorithms and excision methods – Month 36 6.3.3 Final report on models for extended sources – Month 28

Deliverables list

Del. No.	Deliverable name	WP no.	Lead beneficiary	*Estimated indicative person-months*	Nature	Dissemination level	Delivery date (proj.month)
6.1.1	Final report on calibration of pilot experiment using interoperability framework	WP6	4 JIVE	33	R	PU	16
6.1.2	Release of distributed ParselTongue	WP6	4 JIVE	23	D	PU	21
6.2.1	New implementation of Global Fringe Fitting algorithm	WP6	23 NRAO	24	D	PU	36
6.2.2	Direction dependent ionospheric, tropospheric, calibration to test data set	WP6	6 UMAN	12	R	PU	21
6.2.3	Software for mosaic imaging including primary beam correction	WP6	15 ESO	24	D	PU	25
6.2.4	Report on image plane polarization calibration effects	WP6	18 UCAM	12	R	PU	19
6.2.5	Final report on the implementation of algorithms for image plane calibration in a distributed environment	WP6	1 ASTRON	17	R	PU	30

6.2.6	Final report on new algorithms and observing strategies for astrometry	WP6	20 BORD	27	R	PU	28
6.3.1	RFI mitigation software	WP6	5 MPG	6	D	PU	18
6.3.2	Final report on Data Quality algorithms and excision methods	WP6	19 UOXF	31	R	PU	36
6.3.3	Final report on models for extended sources	WP6	1 ASTRON	9	R	PU	28

List and schedule of milestones

Milestone no.	Milestone name	WPs no's.	Lead beneficiary	Delivery month from Annex I	*Comments*
6.1.1	Calibration of pilot experiment using interoperability framework	WP6	4 JIVE	16	
6.1.2	Release of ParselTongue for distributed processing	WP6	4 JIVE	21	
6.2.1	New implementation of Global Fringe Fitting algorithm	WP6	23 NRAO	36	
6.2.2	Direction dependent ionospheric, tropospheric, calibration data set test	WP6	6 UMAN	21	
6.2.3	Primary beam correction and mosaic imaging exercised	WP6	15 ESO	25	
6.2.4	Inclusion of polarization calibration in image plane calibration	WP6	18 UCAM	19	

6.2.5	Evaluate algorithms for image plane calibration in a distributed environment	WP6	1 ASTRON	30	
6.2.6	New algorithms and observing strategies for astrometry	WP6	20 BORD	30	
6.3.1	RFI mitigation software	WP6	5 MPG	18	
6.3.2	Evaluation of data quality characterization algorithms and excision	WP6	19 UOXF	36	
6.3.3	Present models for extended sources	WP6	1 ASTRON	28	