GAIA PHOTOMETRIC DATA ANALYSIS

A.G.A. Brown

Sterrewacht Leiden, P.O. Box 9513, 2300 RA Leiden, The Netherlands

ABSTRACT

The photometric data analysis task encompasses all steps from the raw treatment of the photometric data coming directly from the satellite to the delivery of the final mission-averaged fluxes and magnitudes for each observed object and for each filter. I discuss the role of the photometric data analysis in the context of the overall data analysis for the Gaia mission, the goals and specific challenges of the photometric data analysis and the tasks that can be derived from those, the organisation of the photometric data analysis effort, and finally some results from exploratory studies.

Key words: Gaia; Photometry; Data analysis.

1. INTRODUCTION

The photometric data from the Gaia mission will have a very wide range of scientific applications. The primary information that will be extracted from photometry will be the astrophysical parameters and reddening for individual stars, as well as their variability characteristics in both brightness and colour. Photometry will also be used in discrete object classification in order to separate stars from non-stellar objects, such as galaxies and quasars. Finally, photometry can also provide immediate scientific information during the mission for transient events such as supernovae and micro-lensing. These issues are discussed at length in the contributions to this volume by Jordi (photometric system design and performance), Bailer-Jones (classification and parametrization), Eyer (variability), and Evans (science alerts).

Photometry also plays an essential role in the overall data analysis for the Gaia mission. It is critical for properly treating the astrometric measurements and will provide important information for the processing of the radial velocity instrument (RVS) measurements. In Section 2 I discuss how the photometric data processing fits in with the Gaia data analysis in general and in Section 3 the specific goals of the photometric data analysis are outlined. Section 3.1 describes the specific challenges that make the photometric data analysis for Gaia unique with respect to traditional large scale photometric/imaging surveys. From the photometry goals and the features specific to Gaia's measurement principles a set of tasks can be identified around which one can plan and design the photometric data processing chain that will eventually be used during the mission. This is described in Section 4 followed by a discussion in Section 5 of the organisation of the photometric data analysis effort. In Section 6 I describe some results from exploratory studies of specific problems that have already been carried out in this context.

2. THE ROLE OF PHOTOMETRY IN THE OVERALL GAIA DATA ANALYSIS

To set the stage for the discussion of the photometric data analysis Figure 1 shows a schematic overview of the focal plane arrays for the Astro and Spectro/MBP instruments of Gaia. The Astro instrument will perform the astrometry measurements and also contains the broad band photometer (BBP), which will provide photometry in $\sim$ 5 wide bands ($\sim$ 100–300 nm). The medium band photometer (MBP) is part of the Spectro instrument, which also contains the radial velocity instrument. It will provide photometry in $\sim$ 10 intermediate bands ($\sim$ 20–80 nm). The MBP focal plane array contains red and blue-enhanced CCDs, where the latter use the same optical field as the RVS, the separation being achieved by a dichroic mirror. In addition to the BBP and MBP measurements Gaia will also provide photometry in the 'white-light' bands G and GS of the AF CCDs of the Astro instrument and the sky-mappers of the Spectro instrument, respectively.

Figure 2 schematically illustrates the role of photometry in the Gaia mission by highlighting the three main results that will come out of the photometric data processing and their impact on other parts of the mission. From the overall Gaia data analysis point of view the photometric data processing is closely related to three important aspects.

1. Broad band colours are necessary for making accurate chromaticity corrections for the astrometric measurements. The necessity of these corrections arises because the point spread function for each star crossing the astrometric focal plane will be wavelength dependent. Hence accurate position measurements will require a knowledge

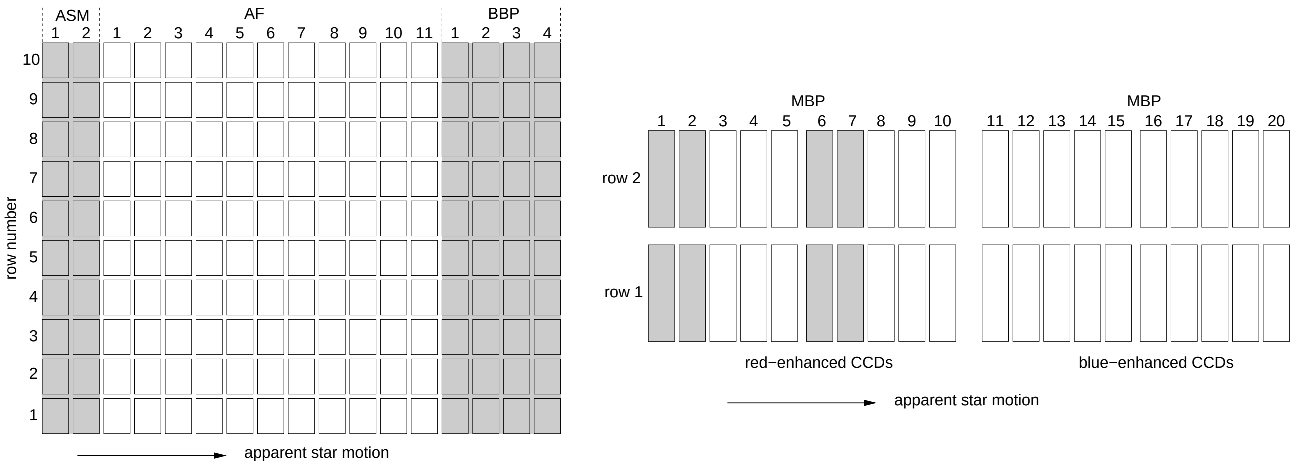

Figure 1. Schematic layout of the focal plane arrays for the Astro (left) and Spectro/MBP (right) instruments. In each focal plane array all CCDs are of the same size and number of pixels. Columns ASM1 and ASM2 of the Astro field form the astrometric sky mapper, i.e., the detection chain. Columns AF1–AF11 form the astrometric field where the astrometric measurements will be done. The broad band photometry is done with CCD columns BBP1–BBP4, which have filters placed in front of them. The filters for the medium band photometry will be placed in front of CCD columns MBP3–MBP5 and MBP8–MBP20. The remaining MBP columns are for the Spectro sky mappers (detection chain operated without filters).

of the colour of each star. In turn the astrometric measurements will feed back into the photometric data processing by providing information on source positions which will be especially important in crowded regions on the sky.

2. The classification and astrophysical characterization of all objects will be based to a large extent on the photometric information and is important for several reasons. (a) The processing of the radial velocity data will require the disentangling of many overlapping spectra in the radial velocity spectrometer focal plane (see Katz et al. this volume). This process will be facilitated very much by having a good indication of the spectral types and astrophysical parameters of the stars involved. (b) The construction of the Gaia reference frame will require the careful selection of a sample of quasars from the Gaia observations. This can be done to a large extent based only on the medium band photometry (see Claeskens et al. this volume). (c) The classification of objects, notably their variability and multiplicity, will also contribute to the important task of maintaining a list of ‘well-behaved’ (single, non-variable) stars that will be used in the processing of the global iterative solution for the astrometry. Finally, the classification of stars and other objects during the mission will also feed back into the photometric data processing itself.

3. As objects are detected during the Gaia mission they will immediately be cross-matched with the existing known objects. This cross-matching combined with a prompt analysis of the photometric data for signs of variability (for which the broad G-band will be especially useful) will lead to so-called science alerts of interesting transients (such as supernovae or gamma-ray burst afterglows), prompting follow-up observations from the ground (see Evans, this volume).

3. PHOTOMETRIC DATA ANALYSIS GOALS

The aims of the photometric data analysis, in terms of ‘data-products’, can be divided into two categories: scientific and calibration products. Both of these will be used during the mission in the manner described in the previous section.

The scientific goals of the photometric data analysis are to provide:

- Mission-averaged calibrated fluxes and magnitudes (and their corresponding measurement errors) for each object in the Gaia data base and each photometric filter. This end-of-mission photometry will be the prime source of object classification and astrophysical parametrization for the scientific exploitation of the Gaia data.
- Calibrated epoch photometry for each object and each filter. This will form the variability data base, which will be used for science applications ranging from studies of specific types of variable stars to the search for planets by using transits. During the mission the variability data base will be used for the science alerts processing chain.

A very important aim of the photometric data analysis is to provide calibration products which include:

- A standard flux scale to which all the photometric measurement will be referred. This includes the calibration of the magnitude scale over the whole range of object magnitudes covered by Gaia.
- Detailed CCD calibrations in terms of their response (differences) on all spatial (from pixel columns to

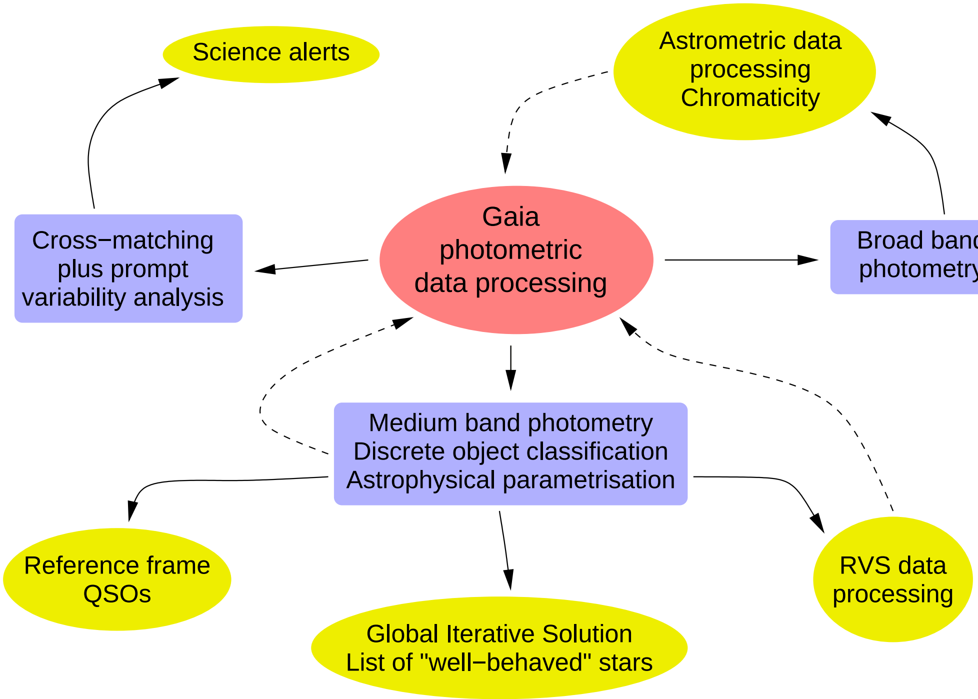

Figure 2. Schematic illustration of the role of photometry in the overall Gaia data analysis. The photometric data analysis will directly result in three main products: broad band photometry, medium band photometry, and variability analysis. Each of these will have an impact on other data analysis processes during the mission, the results of which will in turn affect the photometric data analysis itself (indicated by the dashed lines).

the whole CCD) and temporal scales and as a function of wavelength.

- Calibration of the point spread function (PSF) and, more importantly in the case of Gaia, the line spread function (LSF) as a function of wavelength and position in the focal planes. The calibration of the PSF/LSF involves a detailed understanding of the effects of time-delayed integration (TDI) operation of the focal plane.

- Reconstruction of the photometric pass-bands. This entails the modelling of the evolution of the effective photometric filter bands. The latter are defined by the reflectivities of the mirrors in the Gaia optical train, the CCD quantum efficiency curves and the actual filters themselves. All of these quantities will be subject to change during the mission and the changes have to be monitored and modelled.

3.1. Specific Challenges

The way Gaia will collect the astrometric and photometric measurements is schematically illustrated in Figure 3. The objects that Gaia detects on the sky will travel across the focal plane array where the recorded flux is read out in TDI mode. Due to the rectangular mirrors the PSF for Gaia consists of a core with four prominent spikes, indicated in the figure by the crosses. In order to stay within the telemetry limits only a small region of the sky around each detected object, a window, is read out. The pixels in this window are electronically binned into a set of samples. Before transmission to the ground the samples may be further binned numerically (i.e., summed). Most windows will be transmitted to the ground not as an image but as a one-dimensional set of flux measurements. This way of collecting the data leads to a highly intertwined data set, in both space and time, and makes the Gaia photometric survey very different from the large scale imaging surveys carried out from the ground. As a consequence a number of specific problems arise that need to be taken into account during the photometric data analysis.

The TDI operation across very large focal planes filled with large numbers of CCDs will lead to a complex calibration process. Each transmitted window contains at least a single star but may contain more complicated astronomical objects such as binaries, multiple systems and galaxies. Many objects will be variable in time and the PSF features from bright objects may appear in the windows of neighbouring ones (see Figure 3). In addition the fact that the windows will mostly be transmitted as a one-dimensional set of flux measurements complicates the task of disentangling different sources within one window. The latter is important for removing disturbing sources in order to achieve accurate photometry of the actual target. This requires a mapping of the surroundings of each star which is a task foreseen in the focal plane sampling and windowing strategy.

Further complications include the projection of the two astrometric fields onto a single focal plane and the very

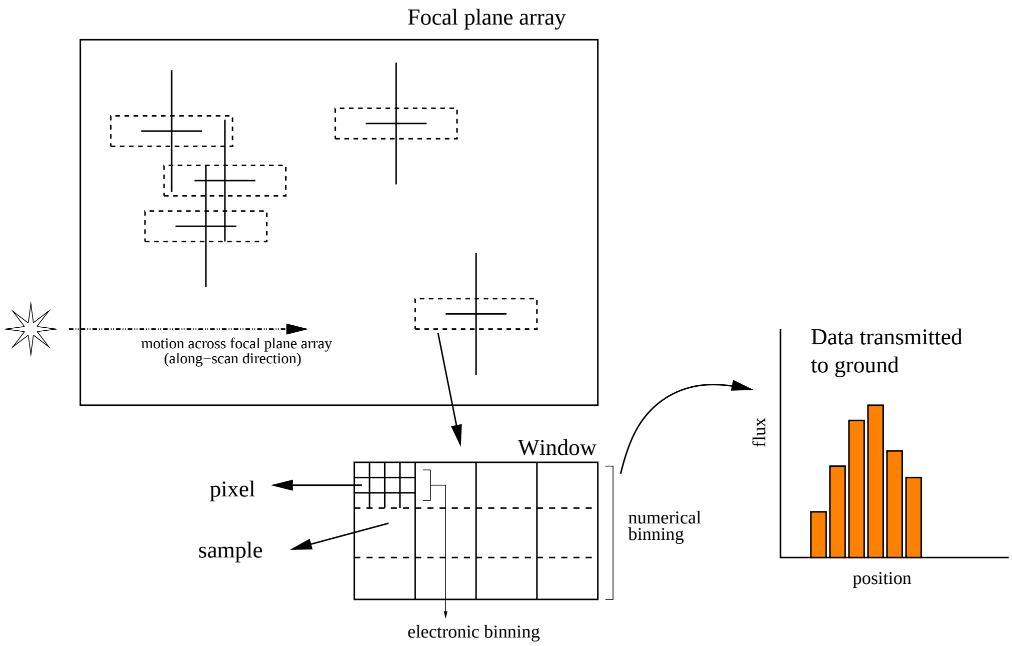

Figure 3. A schematic overview of the astrometric and photometric data collection. The objects that Gaia detects on the sky will travel across the focal plane array where the recorded flux is read out in TDI mode. Due to the rectangular mirrors the PSF for Gaia consists of a core with four prominent spikes, indicated in the figure by the crosses. Only a small region of the sky around each detected object, a window, is read out. The pixels in this window are electronically binned into a set of samples. Before transmission to the ground the samples may be further binned numerically (i.e., summed).

different spatial resolution of the Astro and Spectro fields of view. The former will make the task of background determination for each source much more complicated, while the latter implies a specialised treatment of crowded regions (see Section 6).

4. PHOTOMETRIC DATA ANALYSIS TASKS

The description of the aims and the specific challenges in the previous sections provides the background for the detailed definition and planning of the tasks that have to be carried out as part of the photometric data analysis. A schematic (and non-exhaustive) overview of the tasks that can be identified is given in Figure 4.

In this figure the photometric data analysis is depicted as consisting of a set of core tasks. These are surrounded by a set of other tasks that either provide the interface with the rest of the Gaia data processing system, represent tools that are needed to achieve the goals of the core tasks, or represent data analysis tasks that make use of the results of the photometric processing. An example of an 'interface' is the setting up of a data base of calibration parameters which will be fed to other data analysis tasks that require, e.g., the CCD parameters or standard fluxes. An example of a 'tool' is the analysis of crowded regions which is crucial to the core task of deriving accurate photometry.

4.1. Core Activities

The core task consists of the reduction of the measurements collected in the focal planes to calibrated photometry and can be divided into:

G and GS-band photometry. This concerns the photometry in the 'white-light' bands G of Astro and GS of Spectro which are defined by the mirror reflectivity and CCD quantum efficiency (QE) curves only (no filters). There will be 110 astrometric field CCDs and each star will cross 11 such CCDs as it travels across the focal plane (see Figure 1). The G-band thus provides the most accurately measured photometry and will form the instantaneous 'image' of the sky for the Gaia mission as well as the primary source of the variability analysis for all objects. It will play an essential role in the science alerts processing chain.

Broad band photometry. This task concerns the photometry from the 40 CCDs at the trailing edge (with respect to the scan direction, see Figure 1) of the astrometric focal plane. Filters will be placed in front of these CCDs in order to measure the colours of stars in ~ 5 bands. The main reason for this is to supply colour information to the astrometric data reductions (chromaticity correction) but the pass-bands are chosen such that significant astrophysical information can also be extracted. The latter will be important in those regions of the sky that are too crowded for the medium band photometer. The BBP colours will provided important extra input for variability studies and can be used for issuing science alerts based on spectral changes or peculiarities in objects.

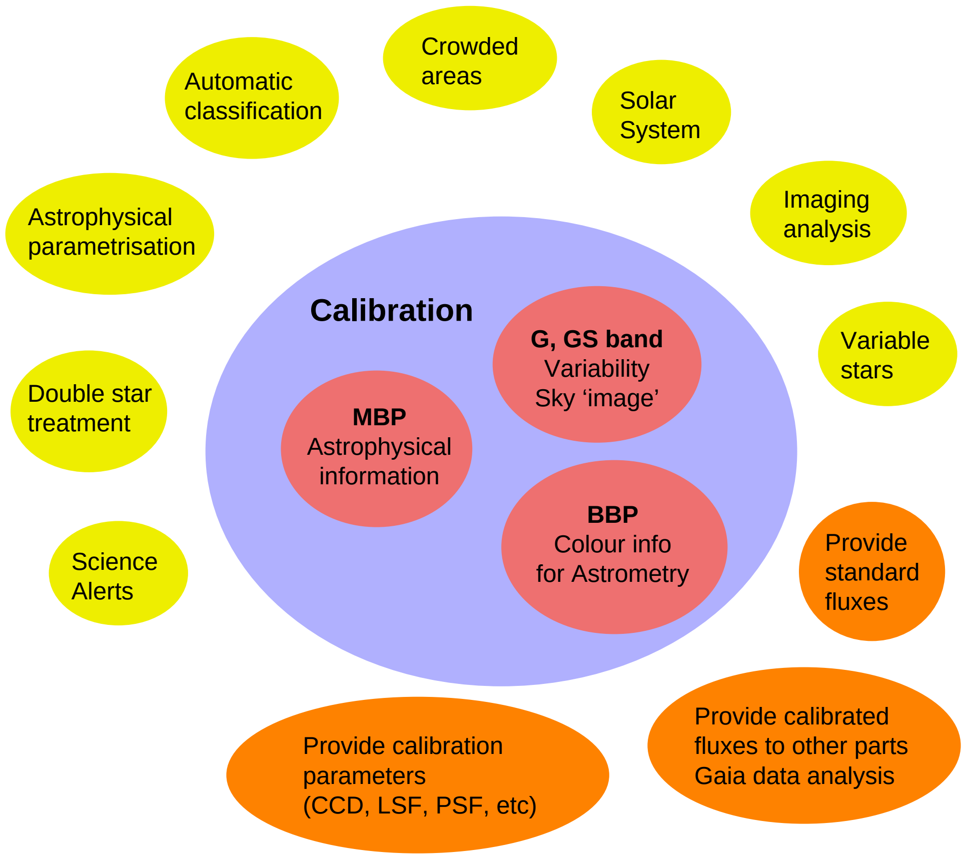

Figure 4. Schematic overview of the tasks that can be identified within the photometric data analysis effort. The core tasks (centre of the diagram) consist of the reduction of the measurements to the calibrated photometry for the medium and broad band photometers as well as in the white light bands G and GS. The other tasks either provide the interface with the rest of the Gaia data processing system, represent tools that are needed to achieve the goals of the core tasks, or represent data analysis tasks that make use of the results of the photometric processing.

Medium band photometry. The medium band photometer is associated with the radial velocity instrument and contains the sky-mappers (detection chain) which serve both MBP and RVS, as well as a passband corresponding to the RVS wavelength range. The information on the brightness of stars in the RVS passband will be used to select the stars brighter than magnitude 17–18, for which spectra will be measured, and to accurately model the contribution to the RVS background by sources fainter than this limit. In addition the MBP contains ~ 10 passbands of intermediate width, which define a photometric system that will provide detailed classification and astrophysical information for each object measured by Gaia.

Calibration. This is in a sense the most important task for both the photometric data analysis and the Gaia mission as a whole since the whole measurement process is intended to be self-calibrating (see Lindegren, this volume). At an elementary level the goal of the photometric data reduction is to accurately extract from each data window (see Figure 3) the object number/photon counts, and to subsequently transform the number counts into calibrated fluxes on a standard scale. The effects that influence the measured flux can be broadly divided into four categories: (1) the mirror and filter reflectivity and transmission profiles; (2) the details of the CCD response; (3) the point spread function; (4) the astronomical sky background. The calibration process concerns the careful control and monitoring of these effects. I will discuss each of the four calibration categories in some more detail here.

The overall spectral response of the Gaia instruments is determined by the wavelength dependence of the reflectivity of the mirrors in the optics, the QE as a function of wavelength of the CCDs, and the transmission profiles of the filters in front of the photometric CCDs. The mirror surfaces/coatings and the filters will undergo changes during the mission as they age due to, e.g., radiation effects. This ageing will change their response as a function of wavelength and this process should be modelled and calibrated as part of the passband reconstruction for each of the photometric bands of Gaia. This cannot be done by measuring spectra of bright stars and thus suitable calibration methods and data will have to be identified or created for Gaia.

The CCD calibrations will form a very large part of the whole calibration process. This involves the detailed monitoring and modelling of the spectral response (QE curve) as well as the overall response of the CCDs, including their noise characteristics. This has to be done for large (CCD to CCD) and small (one pixel column to another) spatial scales and for the corresponding time scales. This task will benefit greatly from the scanning mode in which Gaia will be operated. For average stellar densities on the sky about 60 and 300 stars per CCD per second will cross the focal planes of Astro and Spectro, respectively. This translates to about 660 and 3400 stars per pixel column per 6 hours (which is the spin period of Gaia). The same stars are observed repeatedly on time scales varying from hours to weeks to months, up to the mission life time of 5 years. There will thus be plenty of measurements to perform detailed CCD calibrations on scales from pixel columns to CCDs.

In the case of Gaia for most measurements the two-dimensional point spread function will not be sampled directly. The spatial information for each object is usually only resolved in the direction along scan and therefore what is really needed is the line spread function. The PSF/LSF needs to be measured as a function of position in the focal plane and for each photometric filter. The operation of the CCDs in TDI mode means that the PSF and LSF are convolved with charge transfer and loss mechanisms in the scan direction, variations in the scan rate, the effect of pixel inhomogeneities etc. In addition there are the charge diffusion and modulation transfer function effects. For bright stars there will be a need to deal with CCD saturation problems. This can be done by employing TDI gates or by using the PSF wings. The use of gates will lead to a different integrated response and thus a different LSF, while the use of the PSF wings will require a careful calibration of the full 2D PSF with the possible complication of spectral response differences between the core and wings.

Finally, the astronomical sky background around each source has to be taken into account. Part of the background consists of diffuse emission from the general sky background and zodiacal light. In addition there will be discrete background sources which may be fainter than the Gaia survey limit and PSF features of neighbouring bright stars may also contribute to the background. In Astro there is the added complication that the fields of view from two telescopes overlap on the focal plane, which combined with the across-scan motion may lead to a background that varies as the star travels across the focal plane.

4.2. Further Tasks

Apart from the basic tasks above there are numerous other data processing tasks that either use the result of the photometric analysis as input or provide results that are used in the photometric data analysis. Many examples are given in Figure 4, of which a few are discussed here.

The results of the calibration activities will be stored in a data base that can be accessed by any other part of the Gaia data processing chain that requires, for example, CCD calibration parameters or sets of standard fluxes. The design, maintenance, and data access methods for the calibration data base need to be planned. This includes a definition of the interfaces between the photometric and other data analysis chains, such as astrometric and radial velocity data processing.

Calibrated fluxes will be provided to the variable star processing chain, which will build up the Gaia variability data base during the mission. The calibrated fluxes will also be used for the automatic classification and astrophysical parametrization of objects during the mission. Both these tasks are very large by themselves and dedicated working groups in the Gaia community are planning them (see Bailer-Jones and Eyer, this volume). The results of both the variability analysis and the classification of objects will feed back into the photometric data analysis.

Important data analysis tasks that will go hand-in-hand with the photometric processing are the analysis of crowded regions, the processing of multiple stars, and the imaging of the immediate surroundings of each object measured by Gaia. These tasks will all require their own specialised data processing chains and will provide the photometric data analysis with the means of accurately removing the contributions of background or disturbing sources and accurately assigning fluxes to components of resolved multiple stars.

5. ORGANISATION

The photometric data analysis clearly represents a major effort and thus requires careful planning and organising. The way the organisation is envisioned at the moment is schematically drawn in Figure 5. The planning and design of the photometric data processing chain will involve both the exploration of ideas for suitable algorithms and their actual implementation and integration in an efficient and robust manner. These two sides of the preparatory activities will be the focus of two groups.

The group concentrating on the exploratory studies will be responsible for finding general solutions to specific problems such as how to handle crowded regions or how to produce images of the surroundings of an object from the (mostly) one-dimensional data windows. These studies are more ‘academic’ in nature and are not in the first place aimed at the most efficient or robust implementation of the solutions into algorithms. This group will also deal with the high-level aspects of designing the photometric data processing pipeline and ensuring the integration with the overall Gaia data analysis. Finally this group will also be responsible for identifying the simulated data that will be needed for the photometric data analysis study and making them available (by carrying out specific simulations, for example).

The other group will work on the actual implementation of the results from the exploratory studies into efficient and robust software that will be used during the mission. This group will therefore have more of an information-technology emphasis and concern itself with issues such as: identifying the most suitable data base structure for the calibration and standard flux data bases; deciding on

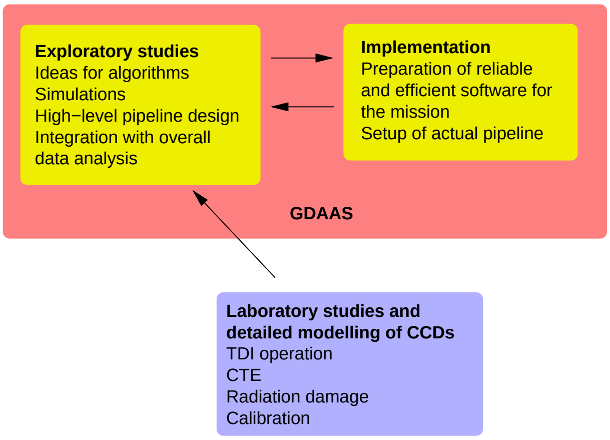

Figure 5. Schematic overview of the organisation of the photometric data analysis effort. The design of the data processing chain has been split up into exploratory studies and the actual implementation of the ideas developed in the latter. Both activities will be embedded in the Gaia Data Access and Analysis System. An important supporting activity will be the laboratory studies and detailed modelling of the CCDs.

the programming language to use; making detailed estimates of computing and storage resources; develop software for monitoring and visualising trends in the calibration parameters etc.

There will obviously be a close collaboration between the two groups. The general algorithmic ideas will come from the 'exploration' group while transforming these into actual software will be done by the 'implementation' group. The results of the latter will in turn have to be verified by the 'exploration' group. The activities of both groups will be embedded in the Gaia Data Access and Analysis System (GDAAS, see Torra et al. and Figueras et al., this volume) as this is the most natural way of ensuring integration with the rest of the Gaia data processing and of having access to simulated data.

An important supporting activity is the detailed study of the characteristics of the CCDs that will be use by Gaia. This includes both laboratory studies of, e.g., radiation damage effects and detailed modelling of the CCDs (see Short, this volume). The results of these studies will feed directly into the CCD and PSF/LSF calibration studies and algorithms.

5.1. Present Commitments

At the moment the exploratory studies are carried out by members of the photometry working group. Examples thereof include: the photometric and PSF/LSF calibration in the context of the simplified GDAAS model; the possibility of calibrating the photometry using the PSF wings; the reconstruction of images of the surroundings of a star by combining the data windows; medium band photometry in crowded regions. These activities now also include a study, funded by the Netherlands Organisation for Scientific Research, of the suitability of the Grid paradigm for the photometric data processing.

In the United Kingdom a group is being set up (with funding from PPARC) which will focus on the implementation of the results of the exploratory studies.

The supporting CCD studies are being undertaken by groups at Leicester and Brunel universities in the UK and these concern laboratory studies and development of CCD calibration algorithms.

6. SOME RESULTS FROM EXPLORATORY STUDIES

A first study of the photometric data analysis problem for Gaia was done by Brown (2001), who wrote down (using simplifying assumptions) the equations that describe the Gaia observations as a sum of overlapping PSFs that are subsequently windowed and binned. A least squares solution for the flux of each source and the local background can then be obtained. This work formed the basis for a study by Evans (2004) of medium band photometry in crowded regions. Crowded in this case means that the stellar densities are so high that overlapping of the stellar images in the focal plane occurs, thus requiring a careful deconvolution in order to derive accurate photometry. This problem is especially relevant for the medium band photometry because of the much poorer spatial resolution of the Spectro instrument (~ 1 arcsec per pixel as opposed to ~ 0.05 arcsec in Astro).

In this study the idea was explored that the accurate positional information from the astrometric processing and a knowledge of the PSF in Spectro will enable a proper deconvolution of the medium band photometer data in crowded regions. The results show that from the photometric data analysis point of view it is possible to process medium band photometry measurements from regions on the sky with stellar densities up to about 400 000 stars per square degree (down to magnitude of 20) without a significant degradation of the photometric precision.

The value of the limiting sky-density is important for assessing how much one can still do in the Galactic plane and bulge regions with the MBP data. This is discussed in the contribution to this volume by Reylé et al. However, keep in mind that the limiting sky density is also determined by how many stars can actually be detected and assigned windows in crowded regions. This limit will be lower than the value quoted above.

The data processing for crowded regions can be improved further if we also have knowledge of disturbing sources around each star to a magnitude limit that goes fainter than the Gaia survey limit. For this purpose the window assigned to object in the last CCD column (AF11) of the astrometric field of the Astro instrument will be longer in the along scan direction. Having these windows for different scan directions will enable a reconstruction of an image of the surroundings (within a ~ 2.5 arcsec diameter) of each star.

This reconstruction process has been studied by Nurmi (2003, 2004) and Dollet et al. (2004) (see also their contributions in this volume). The study by Dollet et al. (2004) shows that one can obtain an all-sky map from the data of the Astro instrument sky-mappers ASM1 and ASM2 at a resolution of ~ 0.1 arcsec, and that this sky map goes 2–3 magnitudes deeper than the Gaia survey limit. The studies by Nurmi show that using the longer AF11 window one can map the disturbing sources in the immediate surroundings of stars to $V \sim 24$ for brightness differences $\Delta V < 8$. The image reconstruction will obviously also be useful in identifying components of multiple stars.

7. CONCLUDING REMARKS

The photometric data analysis study for Gaia has been underway since the middle of 2003 and will now move from the mainly exploratory studies done so far to the actual design and implementation of a photometric data processing system for Gaia. These efforts will lead to the photometric data processing chain that will be used during the Gaia mission.

ACKNOWLEDGEMENTS

I would like to thank the following people for the contributions they have made in one way or another to this work: M. Denby, D. Evans, C. Fabricius, E. Høg, C. Jordi, F. van Leeuwen, P. Nurmi, and members of the Photometry Working Group. I also thank the Netherlands Research School for Astronomy (NOVA) and the Netherlands Organisation for Scientific Research (NWO) for their financial support of this work.

REFERENCES

Brown, A., 2001, Photometric Data Analysis for GAIA: Mathematical Formulation, Gaia technical report GAIA–AB–001

Dollet, C., Bijaoui, A., Mignard, F., 2004, A&A, 426, 729

Evans, D.W., 2004, MBP Profile Fitting in Crowded Regions, Gaia technical report PWG–DWE–001

Nurmi, P., 2003, Combining GAIA windows I: SNR calculations of secondary sources in the ideal case for AF11, Gaia technical report DMS–PN–01

Nurmi, P., 2003, Combining GAIA windows II: SNR calculations of secondary sources using different numerical binnings and samplings for AF11, Gaia technical report DMS–PN–02