Publication Year	2023
Acceptance in OA	2025-02-28T13:25:49Z
Title	Gaia Data Release 3. Processing and validation of BP/RP low-resolution spectral data
Authors	De Angeli, F., Weiler, M., MONTEGRIFFO, Paolo, Evans, D. W., Riello, M., Andrae, R., Carrasco, J. M., Busso, G., Burgess, P. W., Cacciari, C., Davidson, M., Harrison, D. L., Hodgkin, S. T., Jordi, C., Osborne, P. J., PANCINO, Elena, ALTAVILLA, Giuseppe, Barstow, M. A., Bailer-Jones, C. A. L., BELLAZZINI, Michele, Brown, A. G. A., CASTELLANI, Marco, Cowell, S., Delchambre, L., DE LUISE, Fiore, Diener, C., Fabricius, C., Fouesneau, M., Frémat, Y., Gilmore, G., Giuffrida, G., Hambly, N. C., Hidalgo, S., Holland, G., Kostrzewa-Rutkowska, Z., van Leeuwen, F., Lobel, A., MARINONI, Silvia, Miller, N., Pagani, C., Palaversa, L., PIERSIMONI, Anna Marina, Ragaini, S., RAINER, Monica, Richards, P. J., Rixon, G. T., Ruz-Mieres, D., SANNA, Nicoletta, Sarro, L. M., Rowell, N., SORDO, Rosanna, Walton, N. A., Yoldas, A.
Publisher's version (DOI)	10.1051/0004-6361/202243680
Handle	http://hdl.handle.net/20.500.12386/36331
Journal	ASTRONOMY & ASTROPHYSICS
Volume	674

A&A 674, A2 (2023)
https://doi.org/10.1051/0004-6361/202243680

Astronomy & Astrophysics
Special issue

Gaia Data Release 3

Gaia Data Release 3

Processing and validation of BP/RP low-resolution spectral data

F. De Angeli[1,⋆], M. Weiler[2], P. Montegriffo[3], D. W. Evans[1], M. Riello[1], R. Andrae[4], J. M. Carrasco[2], G. Busso[1], P. W. Burgess[1], C. Cacciari[3], M. Davidson[5], D. L. Harrison[1,6], S. T. Hodgkin[1], C. Jordi[2], P. J. Osborne[1], E. Pancino[7,9], G. Altavilla[8,9], M. A. Barstow[10], C. A. L. Bailer-Jones[4], M. Bellazzini[3], A. G. A. Brown[11], M. Castellani[8], S. Cowell[1], L. Delchambre[12], F. De Luise[13], C. Diener[1], C. Fabricius[2], M. Fouesneau[4], Y. Frémat[14], G. Gilmore[1], G. Giuffrida[8], N. C. Hambly[5], S. Hidalgo[15], G. Holland[1], Z. Kostrzewa-Rutkowska[11,16], F. van Leeuwen[1], A. Lobel[14], S. Marinoni[8,9], N. Miller[1], C. Pagani[10], L. Palaversa[17,1], A. M. Piersimoni[13], L. Pulone[8], S. Ragaini[3], M. Rainer[17,18], P. J. Richards[19], G. T. Rixon[1], D. Ruz-Mieres[1], N. Sanna[7], L. M. Sarro[20], N. Rowell[5], R. Sordo[21], N. A. Walton[1], and A. Yoldas[1]

(Affiliations can be found after the references)

Received 30 March 2022 / Accepted 29 June 2022

ABSTRACT

Context. Blue (BP) and Red (RP) Photometer low-resolution spectral data are one of the exciting new products in *Gaia* Data Release 3 (*Gaia* DR3). These data have also been used to derive astrometry and integrated photometry in *Gaia* Early Data Release 3 and astrophysical parameters and Solar System object reflectance spectra in *Gaia* DR3.

Aims. In this paper, we give an overview of the processing techniques that allow raw satellite data of multiple transits per source to be converted into combined spectra calibrated to an internal reference system, resulting in low-resolution BP and RP mean spectra. We describe how we overcome challenges due to the complexity of the on-board instruments and to the various observation strategies. Furthermore, we show highlights from our scientific validation of the results. This work covers the internal calibration of BP/RP spectra to a self-consistent mean instrument, while the calibration of the BP/RP spectra to the absolute reference system of physical flux and wavelength is covered by one of the accompanying *Gaia* DR3 papers.

Methods. We calibrate about 65 billion individual transit spectra onto the same mean BP/RP instrument through a series of calibration steps, including background subtraction, calibration of the CCD geometry, and an iterative procedure for the calibration of CCD efficiency as well as variations of the line-spread function and dispersion across the focal plane and in time. The calibrated transit spectra are then combined for each source in terms of an expansion into continuous basis functions. We discuss the configuration of these basis functions.

Results. Time-averaged mean spectra covering the optical to near-infrared wavelength range [330, 1050] nm are published for approximately 220 million objects. Most of these are brighter than $G = 17.65$ but some BP/RP spectra are published for sources down to $G = 21.43$. Their signal-to-noise ratio (S/N) varies significantly over the wavelength range covered, and with magnitude and colour of the observed objects, with sources around $G = 15$ having a S/N above 100 in some wavelength ranges. The top-quality BP/RP spectra are achieved for sources with magnitudes $9 < G < 12$, with S/N reaching 1000 in the central part of the RP wavelength range. Scientific validation suggests that the internal calibration was generally successful. However, there is some evidence for imperfect calibrations at the bright end $G < 11$, where calibrated BP/RP spectra can exhibit systematic flux variations that exceed their estimated flux uncertainties. We also report that, due to long-range noise correlations, BP/RP spectra can exhibit wiggles when sampled in pseudo-wavelength.

Conclusions. The *Gaia* DR3 data products are the expansion coefficients and corresponding covariance matrices for BP and RP separately. Users are encouraged to work with the data in this format, with full covariance information showing that correlations between coefficients are typically very low. Documentation and instructions on how to access and use BP/RP spectral data from the archive are also provided.

Key words. instrumentation: photometers – instrumentation: spectrographs – catalogs – surveys – techniques: photometric – techniques: spectroscopic

1. Introduction

The European Space Agency (ESA) mission *Gaia* (Gaia Collaboration 2016) has already released three catalogues to the astronomical community; of increasing richness in terms of content, precision, and accuracy. Researchers from many branches of astrophysics have shown great interest in the published data, leading to the publication of more than 6000 refereed papers based on *Gaia* data to date[1].

With respect to the previous *Gaia* Early Data Release 3, *Gaia* Data Release 3 (*Gaia* DR3; Gaia Collaboration 2023b) introduces a number of new data products based on the same source catalogue, including a total of 1.8 billion objects and based on a period of 34 months of satellite operations. A large fraction of the objects in the catalogue has astrophysical parameters determined from the medium (Radial Velocity Spectrometer, RVS) and low-resolution (Blue and Red Photometers, BP and RP) spectral data as well as from the photometric data (Andrae et al. 2023; Creevey et al. 2023). For many of these objects, the actual RVS and/or BP/RP data themselves are part of the release; RVS spectra are released for about 1 million sources, while mean low-resolution BP/RP spectra are available

⋆ Corresponding author: F. De Angeli, e-mail: fda@ast.cam.ac.uk
[1] See the list of refereed papers since launch available at https://ui.adsabs.harvard.edu/public-libraries/fWFE_JYLRZG2jwgwKetH8w

for about 220 million objects, which were selected to have a reasonable number of observations and to be sufficiently bright to ensure good signal-to-noise ratio (S/N) at this stage in the mission. New estimates of mean radial velocities, variable-star classification, and epoch photometry are released for a subset of sources. A large set of Solar System objects, including new discoveries, with preliminary orbital solutions and individual epoch observations are available in the *Gaia* DR3 release. A selection of these also have their reflectance spectra estimated from the epoch BP/RP spectral data (Gaia Collaboration 2023a). The release also includes results for non-single stars, quasars, and extended objects. Finally, an additional data set is also released, called the *Gaia* Andromeda Photometric Survey (GAPS), which consists of the photometric time series for all sources located in a 5.5 degree radius field centred on the Andromeda galaxy (Evans et al. 2023). A number of papers have been prepared by the Data Processing and Analysis Consortium (DPAC) describing all aspects of the data processing and the results of the performance verification activities. In this paragraph, we have only included specific citations to papers that have made use of the BP/RP spectral data. A full list is available at[2].

This paper focuses on the BP/RP low-resolution spectral data and on the processing that led to the generation of the BP/RP spectra included in *Gaia* DR3. Some aspects of the BP/RP processing have already been introduced in recent papers which should be considered essential companions to this one. In particular, calibrations that were also required for the generation of the BP/RP integrated photometry are detailed in Riello et al. (2021) and are described only very briefly in this paper. The algorithm adopted for the internal calibration of the BP/RP spectral data is presented in the dedicated paper Carrasco et al. (2021). We refer to Carrasco et al. (2021) for a detailed justification of the model definition and complement that work by providing information on the actual model configuration adopted to generate the *Gaia* DR3 BP/RP spectra. The focus of this paper is the processing leading to the generation of a homogeneous catalogue of source spectra from the raw *Gaia* BP/RP observations. While *Gaia* DR3 does not provide access to individual observations, knowing the complexities related to the instruments, observing strategies, and processing is important to understand the final product. This paper also contains useful information about the representation of the spectra and the strategies adopted to optimise it and minimise the noise in the final spectra. The validation shown in this paper focuses on these aspects. The calibration of the BP/RP spectral data to the absolute reference system (both in terms of flux and wavelength) is detailed in Montegriffo et al. (2023). This latter should be seen as an essential companion to this paper. Users interested in systematic effects present in the final BP/RP products should refer to that paper, which presents the results of the validation of the externally calibrated data with respect to external absolute spectra. Finally, Babusiaux et al. (2023) present the overall results of the independent DPAC validation process, with useful insights into the limitations and recommendations for BP/RP spectral data.

The paper outline is the following: in Sect. 2 we describe the general concept of low-resolution spectroscopic data and the specific aspects of the *Gaia* BP/RP data that are relevant for this paper; Sect. 3 is dedicated to the data processing, with considerations as to the processing strategies, algorithms, and results; a description of the composition of the BP/RP spectral catalogue in *Gaia* DR3 is provided in Sect. 4; highlights from the internal validation activities are given in Sects. 5 and 6 offer some recommendations for the users.

2. Input data

During its operations, the *Gaia* satellite scans the entire sky every 6 months while spinning around its principal axis and precessing around the Earth-Sun direction. The light from two fields of view (FoVs) is focused on the same focal plane. Images of sources crossing the focal plane move over an array of charge-coupled devices (CCDs) operating in time-delayed integration (TDI) mode, such that the charges generated by a point-like astronomical source are clocked through the CCD at the same speed as the apparent motion of the source caused by the satellite scanning motion. In the following, we use transit to refer to a full focal plane crossing of a source and CCD transit when referring to the crossing of a single CCD, generating one observation.

Throughout this paper, time is expressed in on-board mission time (OBMT) in units of satellite revolutions (1 OBMT-Rev = 21 600 s). A formula to convert OBMT to barycentric coordinate time is provided by Eq. (3) in Gaia Collaboration (2016). In the focal plane array (see Fig. 4 in Gaia Collaboration 2016, or Fig. 2 in Carrasco et al. 2021), the CCDs are arranged in rows (in the along-scan direction, AL) and strips (in the across-scan direction, AC). The largest section of the focal plane array (including 62 astrometric field (AF) CCDs, arranged in seven rows of nine CCDs each, except for one row where there are only eight) is dedicated to the collection of the observations in the broad *G*-band which are used for astrometric measurements and photometry. Following these, two strips of seven CCDs each are dedicated to the BP and RP instruments. Finally, four rows and three strips of CCDs collect the RVS observations. Not all sources crossing the focal plane will also cross the RVS CCDs.

Colour information for all sources is essential to achieving the high accuracy that characterises the *Gaia* astrometry. An initial design – where the flux of sources in a variety of medium bands would be measured on different CCD strips to fulfil this requirement (Jordi et al. 2006) – was abandoned in favour of low-resolution aperture prism spectroscopy. This observational technique is frequently used to obtain a large number of spectra with a single exposure in large-scale astronomic surveys, starting from the Draper catalogue in the early 20th century (Pickering 1890) all the way to future applications such as in Euclid (Costille et al. 2016) and NGRST (formerly known as WFIRST Akeson et al. 2019). The BP/RP instruments were added to the satellite payload to collect this data covering the wavelength ranges [330, 680] nm and [640, 1050] nm, respectively, with varying resolution depending on the position in the spectrum and on the CCD (the resolution covers the range 100 to 30 for BP and 100 to 70 for RP in $\lambda/\Delta\lambda$; see Fig. 3 in Carrasco et al. 2021).

In normal operation mode, observations transmitted to the ground from the satellite are cut-outs of a small area surrounding the position where each source was detected on board. In the case of BP/RP observations, because of the need to cover the full range of the dispersed light, these cut-outs (windows in *Gaia* terminology) need to be much longer in the direction in which the light is dispersed, which is aligned with the AL direction. This is why the size of the BP/RP windows is 60 pixels in AL (as opposed to a maximum of 18 pixels for the AF windows assigned to the brightest objects) by 12 pixels in AC direction, corresponding to an area in the sky of

[2] https://www.cosmos.esa.int/web/gaia/dr3-papers

approximately 3.5 by 2.1 arcsec[3]. This affects the possibility to assign different windows to nearby sources in crowded regions. As a consequence, not all detections result in a BP/RP observation and the average number of BP/RP observations is lower than the average number of transits per source on the focal plane. Partly overlapping windows can in some cases be allocated by the on-board software. When this happens, the window of the brightest sources is transmitted fully to the ground, while only the non-overlapping section of the other window is transmitted to the ground. These truncated windows are not included in the data leading to *Gaia* DR3 as they are normally rather disturbed by the nearby brighter source and require special treatment which will only be implemented for future data releases.

Observations on board can be taken in different configurations depending on the on-board magnitude estimate of the source. The activation of a given configuration can also affect simultaneous observations nearby.

Different configuration aspects include the AC resolution within a window which is only achieved for sources brighter than 11.5 mag in the *G*-band, while windows assigned to fainter sources are binned in the AC direction on board before transmission, resulting in a spectrum with 60 AL samples, where each sample contains the overall flux measurement from 12 pixels. Figure 1 shows the case of a 2D spectrum. The top panel shows the 1D spectrum resulting from the binning in the AC direction. The shape of the 1D spectrum is defined by the combined effect of the response curve, the line spread function (LSF), and the dispersion. The flux from a point-like source is dispersed in the AL direction. The flux at each wavelength is further spread according to the LSF at that wavelength. As a result, each sample in the spectrum, in addition to the local photons, will contain alien photons with different wavelengths. The instrument response, including the filter transmission curve, modulates the flux, only allowing light from a given wavelength range to be detected. A well-centred point-like source should have no flux close to the edge of the observed window. The purpose of the different window strategy for sources fainter than 11.5 mag is to limit the volume of the data that needs downloading from the satellite and to reduce the readout noise. In the following, we refer to these as different window classes (WCs) and in particular to 2D (where the AC resolution is preserved) versus 1D (where binning AC occurs) spectra, respectively. It should be noted that all BP/RP spectra available in *Gaia* DR3 are 1D (i.e. flux values corresponding to positions in the AL coordinate or wavelength when the external calibration is applied). Spectra acquired with a 2D configuration on board are flattened to 1D during the calibration process: a simple sum of the samples in the same AC column is adopted for consistency with the on-board AC binning algorithm.

An ad hoc strategy is also available to prevent saturation when observing bright sources. Different gates can be activated at different locations in the CCD to limit the section of the CCD where the charges are accumulated and therefore effectively reduce the exposure time. The exposure time of an ungated observation is approximately 4.4 s, and the shortest gate active in BP/RP (`Gate05`) reduces this to 0.06 s. Each gate is activated on board as required based on a configured set of magnitude ranges and the on-board magnitude estimated for each transit. The configuration changes for different instruments (BP/RP) and across the focal plane (even within a CCD); see Fig. 2 for the distribution of different gate and WC configuration versus on-board magnitude for BP and RP. As already mentioned, the selection of the appropriate gate configuration is based on the on-board magnitude estimate which can show up to 0.5 mag uncertainties at the bright end. This implies that a given source may be observed in different gate configurations in different transits. Some of these gate configurations will be suboptimal and therefore some saturation cannot be excluded. Moreover the activation of a gate will affect all observations taken at the same time (within 60 pixels or 0.06 s AL) in the same CCD, thus generating gated observations for faint sources that would normally be observed without any gate. This can also cause what are called complex gate cases, where different gates are active in different sections of a window. Complex gate cases are also not included in the processing leading to *Gaia* DR3.

[3] The angular dimensions of each pixel are approximately 58.9 and 176.8 mas in the AL and AC directions, respectively.

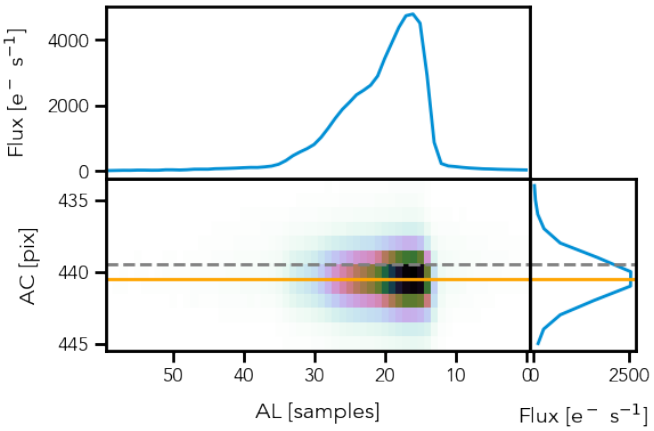

Fig. 1. Example of a 2D BP spectrum. The *central panel* shows the observed spectrum. The dashed and continuous horizontal lines show the AC centre of the window and the AC predicted position based on the source astrometry, the satellite attitude, and the BP CCD geometry. The *top and right panels* show the result of binning in the AC and AL directions, respectively. The AL coordinate is given in units of samples.

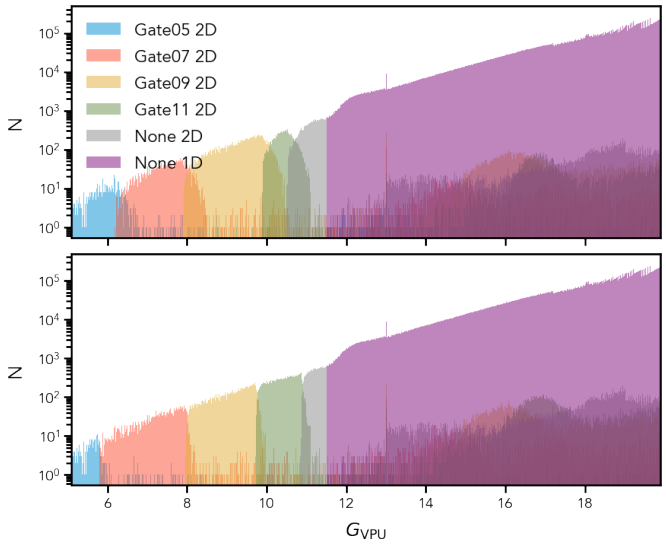

Fig. 2. Distribution of the number of BP/RP observations acquired in BP (*top panel*) and RP (*bottom panel*) with a given gate and WC configuration vs. on-board magnitude labelled as G_{VPU}. The gated observations for sources fainter than ≈11.5 mag in the *G*-band are due to occasional alignment of these sources with brighter objects triggering the activation of a gate.

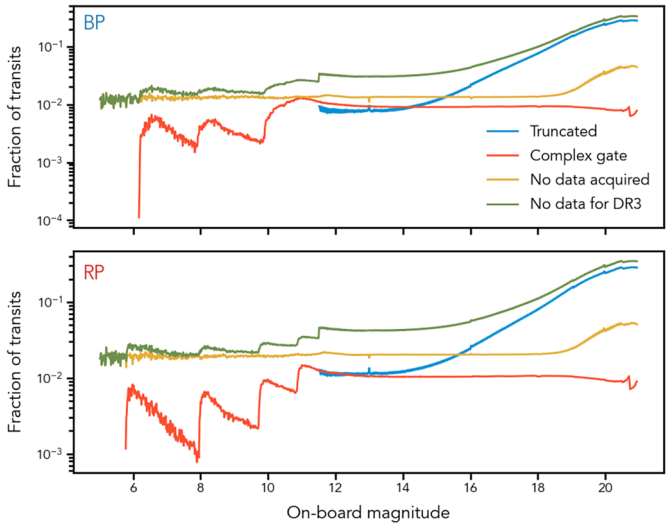

Fig. 3. Fraction of transits that will not contribute a BP/RP observation to the processing leading to *Gaia* DR3 due to either the window not having been acquired (orange line), or to the window being truncated (blue line), or to the window having been observed with multiple gates active within the window (red line). The green line shows the total effect. This is shown as a function of the on-board magnitude estimate as this is the parameter that defines the observation strategy applied to each observation. Truncation for instance is only applied to 1D windows and therefore the corresponding fraction is zero for on-board magnitude brighter than 11.5 mag.

Figure 3 shows the implications for the fraction of BP (top) and RP (bottom) transits available for processing of some of the mission aspects mentioned in this section (size of BP/RP windows, gates, and truncation). The different curves show the fraction of transits that will not contribute a BP/RP observation to the processing leading to the *Gaia* DR3 catalogue for various reasons: the blue curve shows the fraction of BP transits affected by truncation, the red line those acquired with a complex gate, the orange line shows the fraction of transits that do not have a BP or RP window acquired, and the green line simply shows the sum of the three previous quantities and therefore the fraction of transits that will not have an observation that can be processed at this stage. Both fractions of truncated and not-acquired windows increase significantly at the faint end, as expected.

The total number of transits acquired in the period covered by *Gaia* DR3 was almost 78 billion. The processing of the BP/RP spectral data produced calibrated BP/RP epoch spectra (i.e. spectra generated from one single observation) for about 65 billion transits, and mean BP/RP spectra (i.e. spectra averaged over the many observations for a given source) for more than 2 billion sources. Not all transits or sources had a complete set of BP and RP spectra. Section 4 provides more information on the selection criteria that lead to the composition of the *Gaia* DR3 catalogue containing BP/RP spectra for about 220 million sources.

3. Processing

When calibrating the BP/RP data, the characteristics of the various CCDs, the effects introduced by the different optical paths for the two FoVs and by the configuration activated for each observation, and the variation in time of all these elements need to be taken into account. We refer to a set of validity time range (i.e. the interval in time where a given calibration is applicable), CCD, FoV, WC, and gate as a configuration or calibration unit. A set of calibrations per calibration unit (for a total of several tens of thousands of configurations) is produced as part of the instrument calibration process to describe each effect that needs calibrating. Due to the complexity of the system (effectively equivalent to many instruments), the calibration of the data cannot rely on any existing catalogue of standards (all too limited in number and quality), but needs to be solved for internally in the first instance using a large subset of the BP/RP data themselves. This subset is selected to contain data for a sufficiently large catalogue of sources (referred to as calibrators; see Sect. 3.2) covering all calibration units as homogeneously as possible within the limits imposed by nature (e.g., in terms of magnitude and colour distribution). The goal of the internal calibration is to define a reference instrument which is homogeneous across all configurations and time. It is then the responsibility of the external or absolute calibration to define the link between the internal system and the absolute system using a carefully assembled catalogue of spectro-photometric calibrators (Pancino et al. 2021; Marinoni et al. 2016; Altavilla et al. 2015, 2021) and other objects that present features in their spectra that are useful to calibrate specific aspects of the instrument and for which suitable absolute spectra are already available. The internal reference system is defined by the calibrations, that is, the actual calibration coefficients. Once the reference system is established, all the data can be brought to the same system by applying the calibrations. The same approach has been followed for the processing of the *Gaia* photometric data (Carrasco et al. 2016). In this paper, we focus on the internal calibration of the BP/RP spectral data, while the external calibration is the subject of Montegriffo et al. (2023).

The internal calibration includes many different individual calibration steps that are solved for in separate stages of the data processing, often relying on different subsets of calibrators and requiring different strategies for accessing the data in an optimal way. Figure 4 shows a schematic overview of the major steps and dependencies of the process starting from the input raw observed spectra until the output mean spectra.

The two main inputs to the process are the BP/RP observed spectra and the source catalogue containing astrometry and photometry information for all sources observed so far. In the flow diagram, dashed lines are used to represent data flow for calibrators only, while solid lines are used to indicate that the entire set of the observed spectra is used as input into a given stage. The process flows from left to right and top to bottom. The first calibrations are those grouped in the Initial Calibrations block (see Sect. 3.1), which are repeated after the crowding assessment to ensure only the best-suited data are used. The output of these calibrations is part of a database of calibrations that are needed in various stages of the process. The other block of calibrations is the one labelled Flux and LSF Calibration (see Sect. 3.3), which can only start after the initial calibrations are finalised. This is an iterative process that calibrates the effects of differences in response, varying LSF across the focal plane, and small deviations from the nominal differential dispersion functions. When all calibrations are defined, the final steps in the process produce the output catalogues of internally calibrated spectra. In this paper we focus on the mean source spectra (see Sect. 3.4), which are produced using all the observations for a given source, while the process producing the epoch spectra (one calibrated spectrum per observed spectrum) is only briefly described in Sect. 3.3.1. While the epoch spectra are not directly available in *Gaia* DR3, they contributed to the generation of mean reflectances for Solar System objects.

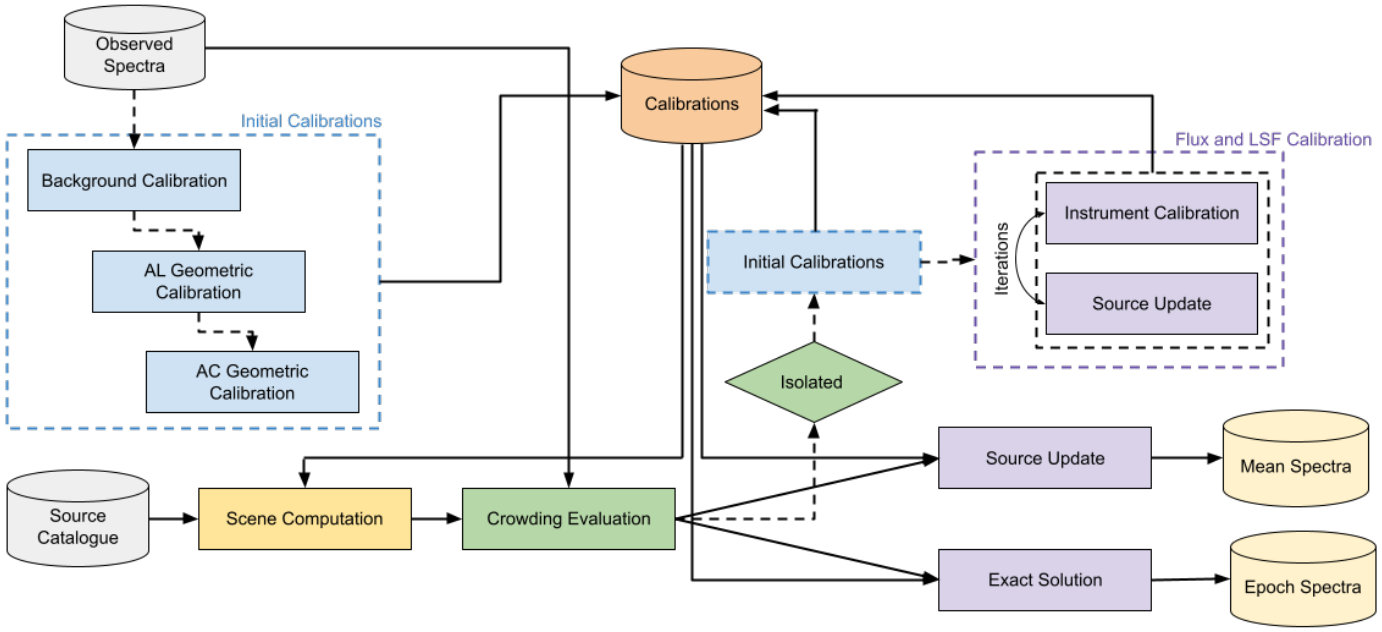

Fig. 4. Schematic view of the processing leading to the generation of the BP/RP mean spectra in *Gaia* DR3.

3.1. Initial calibrations

Starting at the top left corner from the raw BP/RP data, we find a first block of calibrations labelled Initial Calibrations. Some of these have been described in previous papers (Riello et al. 2018, 2021) because they are also required for the photometric processing: the computation of integrated BP/RP fluxes and spectrum shape coefficients (SSCs) – which are the input to the photometric processing together with the corresponding *G*-band fluxes – requires the application of the background and AL geometric calibrations.

The background calibration for *Gaia* DR3 is a two-stage process: high-resolution stray-light maps are first generated to remove effects due to diffraction from loose fibres in the sun shield (Fabricius et al. 2016); a k-nearest neighbour approach is then applied to the map residuals to describe the local astrophysical background (e.g., non-resolved sources, diffuse light from nearby objects, zodiacal light) at a resolution of about 25 arcsec. More details about this calibration and a validation of the results are provided in Riello et al. (2021) (see their Sect. 3.2).

Due to small inaccuracies in on-board detection and window assignment, sources are usually not perfectly centred within the acquired windows. In order to be able to align spectra taken at different times and in different configurations for a given source, we need to rely on a detailed geometric calibration, an accurate attitude reconstruction, and high-accuracy astrometry for all observed sources. Attitude and astrometry are inputs to the BP/RP processing, while the geometric calibration is a product of one of the calibration steps (AL and AC Geometric Calibration in Fig. 4). The AL geometric calibration provides a correction in the AL direction to the location of a reference wavelength within the observed window as computed using our pre-launch knowledge of the CCD geometry. Once the reference wavelength is located within the window, this can also be used as reference position for the application of nominal differential dispersion functions that mitigates the difference in dispersion across the focal plane. More details about the AL geometric calibration can be found in Riello et al. (2018) and Carrasco et al. (2016). The AC geometric calibration is similarly defined as a correction to the predicted location on the source centroid in the AC direction as obtained from pre-launch knowledge of the CCD geometry, the satellite attitude, and the source astrometry.

The two geometric calibrations (AL and AC) are required for the generation of accurate BP/RP transit time and AC coordinate predictions for all sources in the catalogue, that is, the Scene Computation in Fig. 4. An assessment of the crowding status of a given transit (the assessment needs to be done per transit rather than per source because of the overlap of the two FoVs on the focal plane and the varying scan direction) cannot be purely based on the acquired surrounding windows. As we have already mentioned, crowding and priorities imply that a given source may not be assigned a window in the BP/RP CCDs, and therefore such an assessment would be incomplete. This is why the scene is generated starting from the source catalogue containing objects that have been observed at all times during the mission operations so far. The astrometric information from the source catalogue is combined with the satellite attitude and with the geometric calibrations of the CCD of interest to generate the predictions. A detailed description of the scene computation and crowding evaluation has been included in Riello et al. (2021) because of its relevance in the generation of crowding information included in *Gaia* EDR3.

As shown in the schematic view, the Initial Calibrations are repeated after the Crowding Evaluation to include only data that have been assessed as not significantly affected by crowding, thus minimising the disturbing effects of crowding on the calibrations. After this second run of the Initial Calibrations, the spectra are used to generate integrated BP/RP fluxes and Spectrum Shape Coefficients (a set of ad hoc filters designed for the photometric calibration; see Riello et al. 2021). At this point, 2D spectra are marginalised in the AC direction to form 1D spectra and all subsequent processing only deals with 1D spectra.

3.2. Internal calibrators

Each calibration step normally relies on a specifically designed set of calibration data. For the background calibration, for

instance, only Virtual Objects (empty windows acquired on a predefined pattern for calibration purposes) and observations of objects fainter than $G = 18.95$ mag were used to avoid systematic effects due to the target source flux biasing the background measurement obtained from the first and last few samples in the window. For the AL geometric calibration, the need to find the best alignment of the spectra implies a requirement that their shape be similar and therefore that the colour range of calibrators be quite narrow. Finally, for the AC geometric calibration, 2D spectra are essential in order to resolve the location of the peak in the flux distribution in the AC direction.

In the case of the Flux and LSF calibration, the most important requirement is that all configurations are well covered by the set of calibrators. Calibrators covering more than one configuration are particularly valuable. This is naturally the case for time, FoV, and CCD (sources are observed an average of about 40 times in the time range covered by *Gaia* DR3, in different FoVs and CCDs), while in the case of gates and WC, only a limited subset of the calibrators will have observations in two or more observation configurations; these will be sources that have a magnitude close to the boundary of the magnitude range where that strategy is active and that, due to inaccuracies of the onboard magnitude estimate, may therefore be observed in different configurations in subsequent transits. The following criteria were tailored to ensure a clean but well-populated set of calibrators. Only sources in the colour range $-2.0 < (G_{\rm BP} - G_{\rm RP}) < 5.0$ mag and magnitude range $5.0 < G < 17.0$ mag based on the *Gaia* DR2 photometry were considered. Sources with G-band magnitude brighter than 11.5 mag were selected as long as they had more than ten transits in BP/RP, which is to ensure that the magnitude range where gates are activated is well covered. Sources fainter than 11.5 mag with at least one 2D or gated observation were selected as long as their number of usable transits was larger than the median of the distribution of the number of transits in the same HEALPix pixel of level 6 minus the uncertainty estimated as the median absolute deviation of the distribution. This particular criterion was designed to avoid cases of faint sources that happened to be observed in a gated configuration because of their proximity to a bright object: in these cases, a large fraction of the transits of the faint source would be acquired with multiple gates (a case that is not currently processed) and would therefore not be usable. Only the few transits acquired when the two sources were observed at the same time would be usable. These would be likely to be significantly disturbed by the nearby bright source and therefore hardly suitable for calibration purposes. Finally, to enhance the fraction of sources with extreme colours (within the allowed range) with respect to sources of intermediate colours, the distribution of sources fainter than 11.5 mag that are only observed in ungated configuration and in 1D window strategy is flattened in colour as much as possible. Blue sources in particular are essential to constrain the internal calibration at short wavelengths and a poor calibration for blue sources may affect the absolute calibration given that the catalogue of external calibrators contains a large fraction of white dwarfs. The colour flattening is achieved in ranges of magnitude and HEALPix pixels by considering the distribution in colour of the possible calibrators and selecting calibrators from the least populated colour ranges first: each time a number of calibrators are added to the list of selected calibrators from the least populated colour bin, an equal number of calibrators are selected from each of the other colour ranges, giving priority to the sources with the largest number of transits. The process is repeated until the number of selected calibrators has reached the desired number of calibrators per HEALPix. These criteria generated a list of internal calibrators including about 7.6 M objects.

As very blue sources are naturally rare, during the calibration process measurements coming from sources from less populated areas of the colour–magnitude diagram were given larger weight in the least squares solution of the calibration. These additional source-based weights were computed from the density of calibrators in the colour–magnitude diagram and were only applied for the calibration of the BP instrument.

3.3. Flux and LSF calibration

The flux and LSF calibration model is described in detail by Carrasco et al. (2021). This calibration has been defined to take into account sensitivity differences, LSF variations, deviations from the nominal differential dispersion function, and AC flux loss. However, flux-loss terms were not activated for the processing that lead to *Gaia* DR3. The calibration model describes the overall effect of these different aspects on the BP/RP spectra.

Here, it is useful to reiterate Eq. (9) from Carrasco et al. (2021) as the basic formulation of the Flux and LSF calibration:

$$h_{s,k}(u_i) = \sum_{n=0}^{N-1} b_{s,n} \sum_{j=-J}^{J} A_k(u_i, u_{i+j})\, \varphi_n(u_{i+j}), \tag{1}$$

which describes the observed spectrum of source s in calibration unit k, $h_{s,k}$, as a discrete convolution via the instrument model A_k of the mean spectrum. The mean spectrum is in turn defined as a linear combination of some basis functions $\sum_{n=0}^{N-1} b_{s,n}\varphi_n$. In the following, basis functions and bases are used interchangeably. Here, u refers to a pseudo-wavelength system close to the AL coordinate of the samples within a window but adjusted for AL geometry and differential nominal dispersion function. We use u_i to indicate the coordinate of sample i in the pseudo-wavelength system and consequently $h_{s,k}(u_i)$ is the flux measured in the sample i, corrected for effects calibrated in the initial calibration stage (see Sect. 3.1). In this formulation, all the information about the individual source BP/RP spectra is encoded in the $\boldsymbol{b_s}$ coefficients, while the A_k describes the instrument properties. The spectra available in *Gaia* DR3 are in this format (see Sect. 4 for more details on the archive content).

The discrete convolution kernel A_k, the actual calibration, describes the transformation to be applied to the mean spectrum to predict an observation in calibration unit k. Only differential effects between the reference system and the calibration unit it refers to are calibrated in this process. These include contributions from LSF, response, and dispersion. The calibration A_k depends on both the pseudo-wavelength of the sample i that the model is trying to predict and the pseudo-wavelength of the sample $i + j$ that is contributing to the discrete convolution. As explained in Carrasco et al. (2021), given the expected smooth behaviour of A_k across the pseudo-wavelength range, the discrete kernel is replaced by a linear combination of polynomial bases. A smooth variation of the calibration with AC coordinate (within a CCD) is ensured by defining the coefficients of the polynomial in pseudo-wavelength as a polynomial in AC coordinate (see Eq. (13) in Carrasco et al. 2021). A quadratic dependency with the pseudo-wavelength and a cubic dependency in AC coordinate were used for *Gaia* DR3, where the AC coordinate refers to the centre of the window for both 1D and 2D spectra. Given the size of the LSF (see Fig. 5 in Carrasco et al. 2021) and of the expected deviations from the nominal dispersion function, only contributions from neighbouring samples are expected to be significant. Two adjacent neighbours on each side (i.e. $J = 2$ in

Eq. (1)) were considered in the processing leading to *Gaia* DR3. The number of neighbours and the possibility of introducing a step between neighbours were adjusted during trial runs to offer the best balance between residuals and the number of calibration parameters.

At the start of the calibration process, both the mean spectra for the internal calibrators (the $\boldsymbol{b}_s$ coefficients) and the instrument calibrations (A_k) are unknown. An identity calibration is therefore assumed to compute a first set of reference mean spectra for the internal calibrators, effectively solving the following simplified equation for the $b_{s,n}$ parameters:

$$h_{s,k}(u_i) = \sum_{n=0}^{N-1} b_{s,n}\, \varphi_n(u_i). \tag{2}$$

The resulting mean spectra are then used to solve for a first set of calibrations, A_k, using Eq. (1). With these in hand, we can update the reference mean spectra by solving the same Eq. (1) again for the $\boldsymbol{b}_s$ coefficients. The process then proceeds via iterations. The step where the mean spectra are solved for is called Source Update, while the one where the calibrations are computed is the Instrument Calibration. When solving for the BP or RP mean spectrum for a given calibrator, all its observed spectra in that calibration unit need to be collected and used to set up the least squares problem. When solving for the instrument calibration of a specific calibration unit instead, all observed spectra for the calibrators that happened to be observed in that calibration unit and their corresponding mean spectra need to be combined to form the least squares problem. This iterative algorithm was developed using the Map/Reduce paradigm (Dean & Ghemawat 2008) which provides a simple parallelisation model; the Hadoop implementation provided a very efficient horizontally scalable I/O and processing capacity (see e.g., Riello et al. 2018). As the algorithm described above requires grouping the data in two different ways (by source when producing the mean spectrum, and by calibration unit when solving the instrument model), the implementation required two Map/Reduce jobs to perform a single iteration. Although the execution time of individual iterations was quite reasonable, the cost of running a large number of iterations and testing different configuration parameters for the instrument model proved to be the main limitation of this approach. For iterative algorithms, such as the one required for the instrument model computation, a better alternative to Map/Reduce has proven to be Apache Spark[4] which was used for the *Gaia* EDR3photometric processing. For *Gaia* DR4, the iterative instrument model solution will be ported to Spark, allowing for in-memory iterations between source update and instrument model which will dramatically reduce the cost of running a large number of iterations.

Given the large systematic effects present in the data due to water-based contamination in the payload (Gaia Collaboration 2016), particularly at the start of the mission, and the discontinuities caused by the various decontamination campaigns designed to reduce those effects, the iterations designed to initialise the BP/RP reference system were restricted to use only data collected during a specific time period, which was chosen to have the lowest and most stable contamination level. The same strategy was followed for the *Gaia* EDR3photometry (Riello et al. 2021). We refer to this time period as INIT. The periods adopted are approximately [2574.7, 2811.7] and [4121.4, 5230.1] in OBMT-Rev (these are the same used for the photometric processing; see Riello et al. 2021). This effectively implies that the set of calibrators is defined not as a list of sources but as a list of observations, restricted to a specific time period and to a specific subset of sources. A consequence of this is that at the end of the iterative process described above, only instrument calibrations covering the INIT period will be available. Calibrations for all the other periods (collectively called CALONLY) can be computed with a final Instrument Calibration step using all the observed spectra from the CALONLY time ranges for the sources used as calibrators combined with their reference mean spectra. This is shown in the flow diagram in Fig. 5 where dashed lines are used for calibrators' data and the labels INIT and CALONLY indicate the time periods covered by each calibration step.

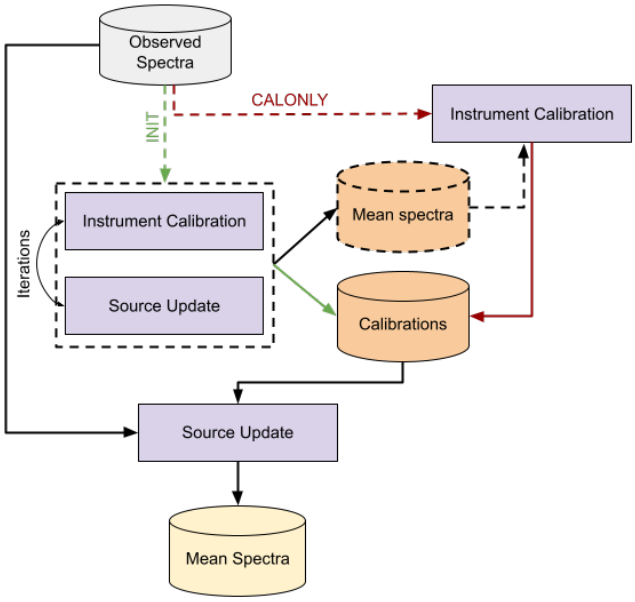

Fig. 5. Flow diagram of the flux and LSF calibration process. Dashed arrows show the flow of calibrator data (also the corresponding mean spectra dataset is shown with dashed borders). When applicable the labels INIT or CALONLY have been added to indicate that only data from the corresponding time periods are being used by a given process.

When calibrations are available to cover the entire time period, a final Source Update using all observed spectra for all sources – not only calibrators – produces the catalogue of mean spectra.

It should be mentioned that in all steps of this process, weighted least squares solutions are obtained via QR-decomposition using Householder reflection to ensure numerical robustness (van Leeuwen 2007). Each solution is computed iteratively: at a given iteration, we use the solution computed at the previous iteration to reject observations that have residuals larger than 5σ. Sample flux measurements are weighted by the inverse variance computed from the flux error for each sample. In the last run of the source update – the one that applies the instrument calibration to all observations to generate the catalogue of mean spectra –, sample flux errors are re-scaled taking into account the scatter in the normalised residuals to mitigate the effects of error underestimation in the wings of the spectra.

3.3.1. Exact solution

Calibrations can also be applied to a single observed spectrum to obtain an internally calibrated epoch spectrum. This process

[4] https://spark.apache.org

appears as Exact Solution in the schematic overview in Fig. 4. In this case the system of equations to be solved is

$$h_{s,k}(u_i) = \sum_{j=-J}^{J} A_k(u_i, u_{i+j})\, g_s(u_{i+j}) \tag{3}$$

where $\boldsymbol{g_s}$ is the output internally calibrated epoch spectrum and A_k is the instrument calibration for the calibration unit k of the observed spectrum $\boldsymbol{h_{s,k}}$ being calibrated. In this case, the solution is simply obtained by inverting the matrix representing the instrument calibration and the resulting spectrum has the same sampling (in terms of number of samples and their location in pseudo-wavelength space) as the observed spectrum, as opposed to the mean spectrum that, being defined as a linear combination of some analytic bases, is effectively a continuous function in pseudo-wavelength. The instrument calibration matrix A_k was generally non-singular and the inversion could be done successfully. Only very few epoch spectra could not be calibrated using this procedure.

Epoch spectra are particularly valuable for objects that vary in time (either due to intrinsic variability or due to different distance or orientation such as is the case for Solar System objects). For these types of objects, the mean spectrum will be ill-defined. Although epoch spectra are not included in *Gaia* DR3, they are relevant here because they have been the input to the generation of the reflectances for Solar System objects.

3.3.2. Calibrations

Calibrations are obtained in time intervals or scopes of about 20 OBMT-Rev (corresponding to about 5 days) for most calibration units. Only for the shortest-exposure configurations, with Gate 05 or Gate 07 active, was it necessary to extend the length of the time intervals to about 100 OBMT-Rev (~25 days) because of the much smaller number of calibrators in these magnitude ranges. The length of the time intervals will vary slightly between calibrations due to the few events that cause discontinuities in the calibrations (such as decontamination campaigns and refocus events; see also Riello et al. 2021). As within a time scope the calibration is assumed to be constant in time, time scopes need to be defined so that such events happen at the boundary between two subsequent intervals.

A set of calibration parameters was solved for each of the 31 860 calibration units. For Gate 05 and Gate 07, the number of nominal calibration units was 1064 per gate configuration, while for other gate configurations or in the ungated case the number of nominal calibration units was 5708 (the ungated case having twice as many as the others because of the two possible window strategies active for objects with magnitude fainter than 11.5 mag). This implies a total of 24 960 nominal calibration units, but there are often cases of non-nominal configurations that get a sufficient number of observations to allow a robust calibration. These are cases of faint sources being observed with a gate triggered by a nearby bright source being observed at the same time (see also Fig. 2).

Displaying detailed information for such a large number of calibrations is challenging. To facilitate this we define two parameters describing each calibration. One is defined as the sum over j of the $A_k(u_i, u_{i+j})$ values weighted by the distance between u_i and u_{i+j}. In the case of a perfectly symmetric calibration (seen here as a convolution kernel) this sum would be equal to zero. In general, it indicates the location of the *peak* of the kernel. A skewed kernel might be caused by small deviations from the nominal dispersion. The second parameter is given by the sum over j of all $A_k(u_i, u_{i+j})$ values, that is, the *integral* of the kernel. Variations in this parameter show differences in the response across the focal plane and between different calibration units.

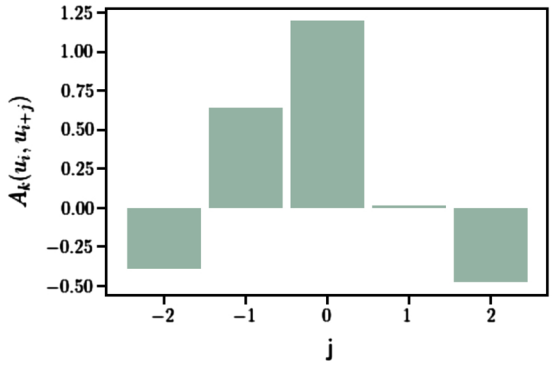

Fig. 6. $A_k(u_i, u_{i+j})$ values defining the instrument calibration for one specific configuration (RP, CCD row 1, preceding FoV, ungated, 1D) in the time range including OBMT-Rev 5000 evaluated at $u_i = 30.0$ and AC coordinate 1000.

Figure 6 shows an example of the calibration for a given calibration unit, evaluated in the central part of the spectrum and of the CCD. This particular case has the peak parameter equal to −0.80 and the integral parameter equal to 0.98.

The plots shown in Fig. 7 offer a quick view of the calibrations computed for all ungated and 1D configurations for the preceding and following FoVs, in terms of the two parameters defined above. The first row of plots refers to the preceding FoV while the second shows the following FoV calibrations. Starting from left, the first two sets of 14 panels show the variation of the peak parameter with the AL position u_i (and therefore wavelength) and time in OBMT-Rev or AC coordinate in all the BP (first seven panels, one panel per CCD) and RP (second column of seven panels) CCDs; the following two sets of 14 panels show the variation of the integral parameter with respect to the same dependencies. Several discontinuities can be observed in the time variation of these parameters. Most of these can be traced back to particular events during the mission, such as decontamination campaigns and refocus activities. The strong variations in the BP calibrations and in particular in the integral parameter versus AL position and time are mostly linked to the varying level of contamination from water-based contaminants present in the payload (Gaia Collaboration 2016), which affects BP more strongly than other instruments (see Fig. 8 in Riello et al. 2018, where the effect of contamination on the photometry in G-band, $G_{\rm BP}$, and $G_{\rm RP}$ is compared).

Relative residuals computed for a random subset of the calibrators (about 50 000 sources) are shown in Fig. 8 for BP and RP. For each observed spectrum, relative residuals are computed as the difference between the observed flux value and the predicted value (computed applying the calibration to the source mean spectrum) divided by the observation flux error. Residuals from all observations and all sources in this dataset are accumulated in a grid in u_i, magnitude, and colour to analyse residual dependencies. From these plots, it is evident that the performance of the internal calibration for the BP data varies significantly over the wavelength range covered and with magnitude and colour. Sources brighter than $G = 12.5 - 13$ and in particular red bright sources show a much larger spread in relative residuals. Performances in RP show a much smoother behaviour across all parameters. The additional weights based