SF2A 2012
S. Boissier, P. de Laverny, N. Nardetto, R. Samadi, D. Valls-Gabaud and H. Wozniak (eds)

GAIA INITIAL QSO CATALOGUE: THE VARIABILITY AND COMPACTNESS INDEXES

A.H. Andrei[1], S. Anton[2], C. Barache[3], S. Bouquillon[3], G. Bourda[4], J.-F. Le Campion[4], P. Charlot[4], S. Lambert[3], J.J. Pereira Osorio[2], J. Souchay[3], F. Taris[3], M. Assafin[5], J.I.B. Camargo[6], D.N. da Silva Neto[7] and R. Vieira Martins[6]

Abstract. The manifold Gaia scientific output relies on precise astrometry accurate to submas standards. This depends on building a fundamental reference frame formed by pointlike, position stable, and allsky homogeneous grid points. In one word, quasars. The Gaia CU3 Initial Quasar Catalogue Working Package was established to beforehand produce one such list, although ultimately the satellite multiband photometry aided by astrometric monitoring has the potential to pick up a clean sample of quasars.

Keywords: Gaia, QSO, catalogues, variability, galaxy morphology

1 Introduction

The Gaia mission (Mignard 2010) ranks among the most ambitious undertakings of modern astronomy. Operating from space, free from atmospheric, thermal, and gravity constraints and with full sky visibility, it will provide completeness to more than one billion objects in survey mode, to a nearly even eighty transits of each one. The launching is planned to 2013, for a five years mission. The measurements include astrometry, multi broad band photometry, and low dispersion spectroscopy for the 20% brighter objects. The astrometric precision will attain 7μas to magnitudes up to G =12, 25μas to magnitudes from G =12 to 15, and 250μas to magnitudes beyond G =15. The astrometry primary objective is the definition and materialization of the celestial reference frame, including the stellar distance scale and standard radial velocities.

The survey mode means that will be observed solar system, galactic, and extragalactic objects. Among the latter, an expected number of about 500,000 QSOs. By QSO we express active galactic nuclei objects (AGN) at large, that is radio loud quasars, Blazars, radio quiet qsos, BL LACs, Seyfert galaxies, LINERS, or as expressed by Souchay et al. (2012), QSO in the sense of an object which can be seen as an extragalactic quasi stellar source from a certain point of view and a specific set of parameters. The satellite observations imply in proper, in the relativistic sense, reference systems to which the measurements are initially referred. These are the described in Bastian (2007) as the Center-of-Mass Reference System (CoMRS), the Scanning Reference System (SRS), the Field-of-View Reference Systems (FoVRS), and the Reference Great-Circle Systems (RGCS), but the final catalogue will comply to the IAU's sanctioned Barycentric Celestial Reference System (BCRS), resulting in the Gaia Celestial Reference Frame (GCRF) materialized by a dense mesh of fiducial QSOs. Notwithstanding, it is also worth to mention that two other quite robust extensions of the GCRF will be produced, to brighter regimes. The one formed by the unresolved galaxies (some 10 million of objects) and the QSOs that didn't make it to be in the GCRF (which would presumably contain several radio-loud quasars). And the one formed the approximately half a billion of stars with highly accurate position and proper motions.

[1] ON/MCTI, Brasil; OATo/INAF, Italy; SYRTE/OP, France; OV/UFRJ, Brasil; present address: SHAO/CAS, China
[2] CICGE/FCUP, Portugal
[3] SYRTE/Observatoire de Paris, France
[4] LAB2, France
[5] OV/UFRJ, Brasil
[6] ON/MCTI, Brasil
[7] UEZO, Brasil

2 Gaia Initial QSO Catalog

QSOs are thus crucial targets to define the GCRF, and accordingly on board means are capable of classifying them. The QSO classification contains three major orientations: getting the cleanest QSO sample to determine the GCRF; deriving the most complete QSO sample based on the full Gaia data; and determining the astrophysical parameters for each QSO. The determination itself of a Gaia source as a QSO is planned to rely primarily on comparison of the photometric output against a template of spectral energy distributions (SED), ands secondarily on astrometric observables, variability analysis and a reliable initial list of known QSOs.

The study by Claeskens et al. (2006) shows that, based on the end-of-mission colour information, supervised Artificial Neural Networks (ANNs) can virtually reject all contaminating stars (including white dwarfs), although the completeness drops to about 20% at G =20. Variability on the color and space domains is efficient to remove stellar contaminants, whereas zero proper motion and parallax constraining can single out the late type red stars and white dwarfs contaminants. An alternative method relying on the QSOs and stars different predominant spaces in the color-color spaces can also be employed providing maybe 95% efficiency to a larger sample, although the crucial u-band will not be available in Gaia, as it was for the SDSS candidates scheme (Richards et al. 2002). The template adherence method can deliver a tighter but smaller sample. Notice, however, that a sample as small as 10,0000 quasars can stabilize the GCRF to a residual rotation of less than 0.5 μas per year, provided they are well distributed over the sky.

The relatively small number of points actually required to constitute a robust GCRF brings particular relevance for an initial list of known QSOs. This is exactly the purpose of the Gaia work package Initial QSO Catalogue for Gaia (GIQC), under the CU3, Core Processing Coordination. The aim of which is to obtain a clean sample of at least 10,000 quasars, distributed all sky above $b = \|20°\|$ of galactic latitude, with magnitude brighter than V=20 and point-like PSF. This bona fide initial clean sample is useful both for the actual orientation of the GCRF and to enlarge the templates of the recognition scheme. For that, catalogue and published QSO determinations are inspected. The starting number of publications handled amounted to a few thousands, from which the main contributors were the Véron-Cetty and Véron catalogue (12th edition), the 2dF catalogue, and the SDSS (DR5) catalogue. The collection kept expanding, and a recent status of the compilation can be found in Souchay et al. (2009); Andrei et al. (2009a,b). Several criteria were enforced to acknowledge a given source, the most importants of which were at least two independent determinations and the existence of spectroscopic redshift. In the most recent versions also the aspects of morphology and variability are addressed, which is reviewed below.

Presently the GIQC contains 187,505 objects. 187, divided in thre categories - defining, candidates, and other. The defining objects are 136,643 well documented QSOs, being 103,422 from the SDSS/DR8 (Schneider et al. 2010). The candidate sources are 24,227 objects, 1.075 from the ICRF2 for which there is no reliable optical counterpart 15,373 optically point like AGNs, and 7.779 QSOs of poor record. Finally, as other there 26.635 objects, being 385 radio-quasars, 23.178 objects fainter than the 20th magnitude, 2.985 unreliable detections, and 87 empty fields. Figure 1 brings an excerpt of the catalogue (without including the still provisional morphology and variability indexes).

RA (deg)	DEC (deg)	MAG	z	Rshr	Rsrn	Rgrn	Bshr	Bsrn	Bgrn	Ishr	Isrn	Igrn	Class
0.000000	-0.032778	19.40	1.560										C P
0.002083	-0.450833	20.09	0.250										O F
0.005291	-2.033269	19.29	1.356	0.75	0.10	0.13							D
0.005735	-30.607458	19.18	1.143	0.20	0.01	0.91							D
0.007326	-31.373790	19.74	1.331	0.73	0.44	0.00				1.82	1.14	1.37	D
0.011279	-25.193609	21.56	1.314										O F
0.012178	-35.059062	17.09	0.508	0.59	0.20	0.27				0.39	0.80	0.07	D
0.022792	-27.419533	19.11	1.930	0.12	1.01	0.41							D
0.027500	0.515278	20.37	1.823										D S
0.033333	-63.593333	17.00	0.136										C A
0.034167	0.276389	20.03	1.837										D S
0.038604	15.298477	19.40	1.199	0.92	0.02	0.30	0.36	0.92	0.08	1.11	1.51	1.46	D S
0.039089	13.938450	18.29	2.240	0.59	0.23	0.14	0.63	0.91	0.09	2.07	0.16	1.43	D S

Fig. 1. Excerpt from the GIQC (version 3, 2010). The morphological indexes, columns `Rshr` to `Igrn` (from the DSS plates B, R, and I) are presented in section 2. The first Class column signals the Defining, Candidate or Other classification, while the second one details the classification.

The catalog compiles reliable redshift for 183,543 objects (97.87%), and there are reliable optical images for 159,701 objects (85.17%). The space density distribution is shown in Figure 2a.

The astrometric coordinates compiled for GIQC are presented in the LQAC2 (Souchay et al. 2012). They are the most precise ones available for each source, but their accuracy is not consistent as they originate from widely different catalogs. Their sky density (4.5 per sq.deg) and sky distribution (about 75% of the minimum distances from adjacent sources are between 1.5°and 5°) prompted to combine the set into an optical materialization of the ICRS. For this we followed the precepts established for building the LQRF (Andrei et al. 2009a,b), but now nearly doubling the number of grid points.

Obviously, only those QSOs for which there are available optical images were used. The distribution is homogeneous, with no empty zones, though there is a quantity decreasing onwards the inner galactic disk, whereas there is an enhancement on the SDSS zones. By comparing with the common ICRF-2 quasars, we find the systematic offsets $\Delta\alpha\cos\delta$=+1.6 mas (σ=154.9 mas) and $\Delta\delta$=+3.6 mas (σ=158.8 mas). Excluding the optical minus radio residuals deviating by more than 2σ, the adherence to the ICRF2 is obtained as 55.5mas on right ascension and 59.7 mas on declination. Figure 2b shows the zonal warps. They are mostly (67%) smaller than 50 mas, with just 6 zones showing systematics larger than 100 mas.

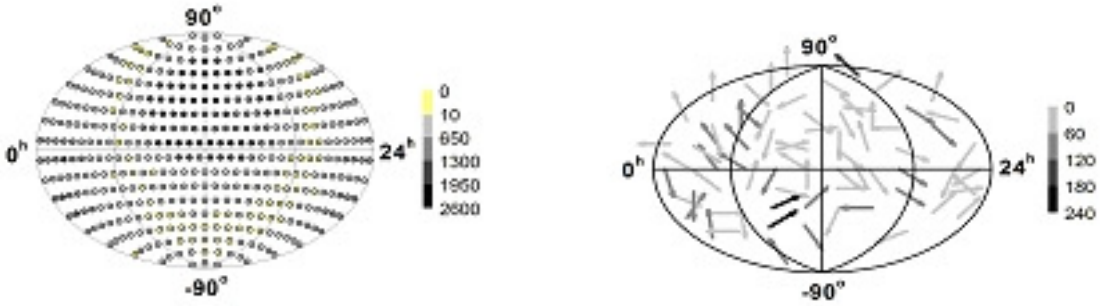

Fig. 2. Left: (a) The sky density of the GIQC based extragalactic reference frame. The counts are in bins of 10°. The zones in the inner galactic plane are less populated, while there is a density enhancement on the SDSS region. **Right:** the vectorial distribution of the systematic deviations (North up, East right) of the ICRF2 to the GIQC based extragalactic reference frame within bins of 30°. 67% of the deviations are smaller than 50 mas, while only 6 of them are larger than 100 mas.

3 Morphology

In the framework of the unified theory the classification of AGNs is largely a matter of line-of-sight perspective. Naturally a less schematic classification must consider the object interaction and merger history, dust contents, star formation waves, the rate and gyro direction of the central massive black hole spin, and the enrichment of the off accretion disk shells or regions. Nonetheless the entrusted relationship between the mass of the central black hole and the energy output from the inner regions versus the mass and luminosity of the host galaxy must generally hold. Therefore, since by optical ground observations quasars are essentially quasi stellar objects, an apparent paradox arises by which the more massive and luminous a host galaxy is, the more luminous the quasar tends to be, thus making the more invisible by contrast the host galaxy. Hubble images and active optics are now enabling the study of the host galaxy, what increases the knowledge about the quasar itself too. The presence of the host galaxy can be inferred also from color studies (Sakata et al. 2010) and from departures of the compound point spread function (PSF) of quasar and host galaxy to the purely pointlike stellar PSFs (Falomo et al. 2001). This methodology can be used to classify morphologically the quasars observed out from the atmosphere, as will be the case of the Gaia mission, hence deriving a centroid astrometrically more precise than if a stellar PSF would have been applied for the centroid determination.

We derived a PSF-based morphological classification of the GIQC objects using the available the B, R, and I DSS images. The morphological classification derives from comparing the target PSF against the local PSF. To that, neighborhoods of 5 × 5 arcmin around the QSO are obtained from the DSS plates. From the total LQAC-2 sample were obtained images of 114,606 fields from the B plates, 191,030 fields from the R plates, and 183,421 fields from the I plates. The incompleteness in most cases is due to the sky coverage, and lack

of available uncompressed digitalization for the DSS2 blue plates, but also to a residual number of faulty file transfers. In the retrieved plates there were cases where the quasar was not present or was too faint to provide a meaningful PSF (29,497 objects in the B plates, 37,790 objects in the R plates, and 29,799 objects in the I plates). Finally, there was not enough adequate comparison stars in 6 B plates, in 134 R plates, and in 86 I plates. As expected, due to the more complete sky coverage and brighter magnitude, the largest quantity of quasars corresponds to the R plates.

The IRAF task `DAOFIND` is used to detect both stars and target, as well as to derive the PSF parameters. Additionally the tasks from the `IMMATCH` IRAF's package are used to match stars and target to their catalog positions and magnitudes. Stars are collected within one magnitude from the QSO's magnitude, but in case less than five stars are picked up the magnitude limits are progressively enlarged at one magnitude steps, except to stars brighter than the tenth magnitude. Stars must be isolated from each other by an inner radius of ten pixels, and within the frame by the same threshold. If less than five comparison stars are found no morphological index is derived for that QSO on that plate. Three estimators of the PSF are used: SHARP (probing skewness), SROUND (probing roundness), and GROUND (probing normalness).

To test the power and efficiency of the above procedure applied to the DSS Schmidt plates, a comparative test was made using 1,343 objects present both on the R DSS images and on the r' SDSS images (Andrei et al. 2011). The large number of comparison objects enabled to sample regularly the SDSS quarter of sky space, as well as to collect the extreme examples on low and high redshift, bright and faint magnitudes, and the tails of the color distribution. The large number also allowed retaining only those QSOs for which at least 20 comparison stars were found. The analysis showed that there is no degeneracy of the indexes with magnitude. All nine morphological indexes, namely three parameters in three colors, behaved alike on the DSS and SDSS. Only 1% of the stars were misidentified by the morphological indexes, while for the quasars the correlation between the morphological classification and the SDSS catalogue classification was of 0.86 for the SDSS images and of 0.72 for the DSS images. Due to the better quality of the SDSS images and pixelization, the number of QSOs reckoned as extremely non-pointlike is larger from the SDSS fields (144 objects) than from the DSS fields (86 objects), but 50% of those are common. This trial thus supports using the DSS images to derive a morphological index, that describes the degree of agreement or disagreement of the quasar PSF to the local mean stellar PSF. Owing to the limited resolution, the morphological indexes are interpreted as presenting the signature of the host galaxy.

The relative distributions of the morphological indexes are shown in Figure 3. It is evident that the number of non-pointilike QSOs is small but by no means negligible. We found the least of non-pointlike quasars on the B plates and progressively more into the R and I plates, which is expected from the redder emission from the host galaxy than from the inner QSO sources of optical emission.

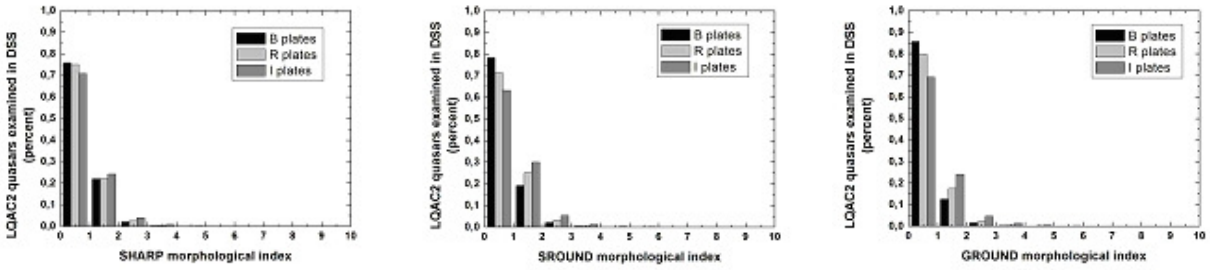

Fig. 3. From left to right the Sharp, Sround, and Ground morphological indexes, for the B, R, I plates from the DSS. Although in all cases most of the QSOs appear as pointlike sources, notice that the fraction of extended sources is non neglible, and the proportion increases from B to I, i.e., from the outshining bluer central central source to the redder host galaxy.

4 Variability

The intrinsic QSO position stability at the sub-mas level will be over important for the establishment of the GCRF. At this point, one should take into account that at the same time that QSOs are the ideal (or rather the only) choice for fiducial grid points in the establishment of a quasi-inertial celestial reference frame, as

prescribed by the ICRS/ICRF paradigm, they are the most energetic and violent, large scale structures in the Universe, active galactic nuclei powered by a super-massive black hole (SMBH). Therefore, there can be several effects that can affect precise photometric center determination of the quasars (see, e.g. Porcas 2009) for the case of chromatic AGN core positions). Additionally, the Gaia mission measurements are planned with to an unprecedent precision in photo-center position, which will allows us to investigate the astrometric stability of QSOs and the possible physical consequences. Always according to the standard model of AGNs, a QSO consists of a SMBH (10^7 to 10^{10} solar masses) surrounded by a X-ray and optical continuum emitting region, probably with an accretion disk geometry (see Sulentic et al. 2000), a broad line region (BLR) and a larger region that can be resolved in several nearby AGN, which usually is referred to as the narrow line region (NLR). These central regions are surrounded by a torodial structure of dust. The regions emit in different wavelength bands and are supposed to have different dimensions. Flux ratios observed in the different part of the QSO continuum are sensitive to substructure and geometries of the emitting regions. Due to accretion that occurs close to the SBMH, one can expect instabilities and consequently variations in brightness and spectral distribution of these objects (see, e.g., Shields et al. 2003; Popović et al. 2012a,b). There are several mechanisms which can cause variation (see Andrei et al. 2009a,b): instabilities in the accretion disc around the central black hole: supernova bursts; jet instabilities, and gravitational microlensing. On the other hand, the dusty torus is illuminated from the accretion disk and also re-emits and absorbs, and some variability can be expected in photo-center position due to different illumination of the torus. Recently, Taris et al. (2011) reported about the magnitude variations of quasars and the potentially correlated motions of their centroids, finding that in one QSO there is a correlation between the centroid motion and magnitude variation.

As discussed, one thus can expect an accretion disk emission. Its variation can be caused by an outburst from the central (compact) continuum source, but also the variation can be connected with some kind of perturbations in the accretion disk (see, e.g. Jovanović et al 2010; Popović et al. 2012a,b). In the GIQC the accretion disk and the torus dimensions are being used as variability indexes, in the sense of indicating which objects are apt to suffer a variation on their photocenter along the 5 years of Gaia mission, due to existence of an angularly extended geometry.

Popović et al. (2012a,b) elaborate this investigation of spectro-photocentric variability of quasars caused by changes in their inner structure. They consider perturbation in a relativistic accretion disk assumed to be around a SMBH in quasars; and changes in the pattern of radiation scattered by the dust particles in the surrounding torus, due to the variations in the accretion disk luminosity and dust sublimation radius. As a result it is derived how much these effects may contribute variability of photo-center, i.e. to quantify "noise" with goal to better characterize any resulting error on the position determination; as well as the estimation of the possibility to observe this effect during Gaia mission, and the group of quasars for which these effects may be dominant.

One can conclude that perturbations (or bright spots) in an accretion disk may bring offsets of the photo-center, and this effect will be resolved by Gaia observations. The best candidates are low-redshifted AGNs with massive black hole, which are in principle very bright objects. One can expect a maximal offset of the center (in the case bright spot located at disk-edge) order of a few mas. The photocenter offset can also be caused by changes in the torus structure due to different illumination of the torus, in the case when inclination of the torus is larger than 30deg. In this case, the maximal offset can be of several mas, therefore also detected by Gaia. Both types of photocenter offset can also appear connected. Causing smaller photometric variation, can also be listed variations in the BLR, in the NLR, and in the optical jet.

In principle, to avoid the possibility of the photocenter variation due to perturbations in accretion disk, or in the BLR, one may estimate the dimensions of the BLR and choose for the GCRF objects with small dimensions of the BLR. On the other hand, to avoid variations of the photocenter due to filaments in the torus, it is preferable to choose quasars where the torus is face on orientated. In short, QSOs with high variability are not good objects for construction of the GCRF. Nonetheless, if the optical variability, that is likely to be sensed by Gaia's typical sampling (about 1 month on average), can be linked to the size, if not the preferred direction, of the astrometric jitter, this can be modeled and accounted for. If this is so, the astrometric error budget is alleviated and some variable quasars can be brought back to the GCRF. Conversely, it is important to remark that Gaia astrometric measurements will be very useful for the investigation of the inner quasar structure and physical processes, especially in low redshif variable sources.

5 Conclusions

There are very few aspects of the Gaia program that are tagged as "mission critical", but for satellite, instrument, and data handling issues. This testifies of the length and breadth, and depth, of the observations along the five years of the mission. By the same token, there is no exaggeration to state that although the aspects of morphology and variability of quasars might impact on the realization of the astrometric reference frame, which is a central to the Gaia concept, the information that will be gathered on those aspects will have just as important an impact on the astrophysics of quasars.

In the case of variability is evident that a data base must be formed to keep record and to make sense of the departure from the norm, either regular or as fluctuations. But also in the case of morphology it is crucial to form a database, due to the line spread function assessment of the centroids which will be differently affected from pose to pose taken at different directions. Such data bases ought to be organized in a way to ease the evaluation object by object along the mission to account for and compensate the effects regarding a degradation of the precision of the quasar's astrometry.

Such databases might also play a role on bringing in or dispose of objects misclassified by the color loci fundamental recognition scheme. Although this will probably be an adequation percentually minor, it is a major facility for the astrophysical studies pos-mission. A natural, aggregative increment of the data base, observation by observation, with all the observational circumstances in hand, is clearly preferable than the massive task of building it from scratch afterwards the mission by mining the archives.

AHA thanks the PARSEC International Incoming Fellowship within the Marie Curie 7th European Community Framework Programme, the Visiting Professorship for Senior International Scientists of the Chinese Academy of Sciences, and the CNPq grant PQ-307126/2006-0. J.I.B.C. acknowledges CNPq financial support 477943/2007-1. D.N.S.N. thanks FAPERJ grant E-26/110.177/2009.

References

Andrei, A.H., Souchay, J., Zacharias, N., Smart, R.L., Vieira Martins, R., da Silva Neto, D.N., Camargo, J.I.B., Assafin, M., Barache, C. 2009a, A&A, 505, 385

Andrei, A.H., Bouquillon, S., Camargo, J.I.B., Penna, J.L., Taris, F., Souchay, J., Silva Neto, da D.N., Vieira Martins, R., Assafin, M. 2009b, *Proc. of the Journées 2008 Systemes de refrence spatio-temporels*, Observatoire de Paris, ed. M. Soffel and N. Capitaine, p. 199

Andrei, A.H., Gontier, A-.M., Barache, C., da Silva Neto, D.N., Taris, F., Bourda, G., LeCampion, J-.F., Souchay, J., Camargo, J.I.B., Pereira Osório, J.J., Assafin, M., Vieira Martins, R., Bouquillon, S., Anton, S. 2011, *Journées 2010 Systemes de reference spatio-temporels*, Observatoire de Paris, ed. N. Capitaine p. 125

Bastian, U. 2007, *GAIA-CA-SP-ARI-BAS-003-06*, Version 6.0.

Claeskens, J.-F., Smette, A., Vandenbulcke, L., Surdej, J. 2006, MNRAS, 367, 879

Falomo, R., Kotilainen, J., Treves, A. 2001, ApJ, 547, 124

Jovanović, P., Popović, L.Č ., Stalevski, M., Shapovalova, A.I. 2010, ApJ, 718, 168

Mignard, F. 2010, *The Gaia mission objectives, description, data processing*, ADA 6 - Sixth Conference on Astronomical Data Analysis, p. 10

Popović, L.Č., Shapovalova, A.I., Ilic, D., et al. 2012, A&A, 528, 130

Popović, L.Č., Jovanović, P., Stalevski, M., Anton, S., Andrei, A.H., Kovacevic, J., Baes, M. 2012, A&A, 538, 107

Porcas, R.W. 2009, A&A, 505, L1

Richards, G.T., Fan, X. Newberg, H.J., et al. 2002, ApJ, 123, 2945

Sakata, Y., Minezaki, T., Yoshii, Y., et al. 2010, ApJ, 711, 461

Schneider, D.P. et al. 2010, ApJ, 139, 2360

Shields, G. A., Gebhardt, K., Salviander, S., et al. 2003, ApJ, 583, 124

Souchay, J., Andrei, A.H., Barache, C., Bouquillon, S., Gontier, A.-M., Lambert, S.B., Le Poncin-Lafitte, C., Taris, F., Arias, E.F., Suchet, D., Baudin, M. 2012, A&A, 537, 995

Souchay, J., Andrei, A.H., Barache, C., Bouquillon, S., Gontier, A.-M., Lambert, S.B., Le Poncin-Lafitte, C., Taris, F., Arias, E.F., Suchet, D., Baudin, M. 2009, A&A, 494, 799

Sulentic, J. W., Marziani, P., Dultzin-Hacyan, D. 2000, ARA&A, 38, 521

Taris, F., Souchay, J., Andrei, A. H., Bernard, M., Salabert, M., Bouquillon, S., Anton, S., Lambert, S. B., Gontier, A.-M., Barache, C. 2011, A&A, 526, 25