Telemetric flows for the RVS from GSC-2.2 star counts taking into account overlap of spectra in crowded FoVs
RVS-YV-002

Y.P. Viala, F. Arenou, D. Katz

July 19, 2004

Abstract

Combining the GAIA Nominal Scanning Law and star counts from the GSC-II (version 2.2) catalogue, we have refined our previous estimations (RVS-YV-001) of the telemetric flows expected from the Radial Velocity Spectrometer (RVS). We now take into account the overlap of spectra between neighbouring stars in crowded fields of view. This reduces the number of pixels effectively occupied by spectra and that must be sent to the ground. Mean telemetric flows integrated over one day are determined all along an assumed 5 years (1800 days) GAIA mission. The telemetry budget is computed for various magnitude limits, two possible values for the area of single spectrum. The amount of data that can be saved and downloaded to the ground is estimated for different data compression factors and as a function of the storage capacity of data in the solid state mass memory implemented on board. Further improvements of the RVS telemetry budget determination are discussed at the end of this note.

1 Introduction

In previous notes (RVS-YV-001) [1] and [2] we presented a preliminary investigation on the telemetry budget that can be expected from the Radial Velocity Spectrometer (RVS). This first estimation of the telemetric flow was obtained by combining 1) the calculation of the attitude of GAIA (derived from the Nominal Scanning Law (NSL) and its parametrization using quaternions as described in SAG-LL-30 [3] and SAG-LL-35 [4]), which provides the RVS-FoV position on the sky as a function of time, and 2) the determination of the star density versus magnitude at that position, which is obtained from star counts from the GSC-II catalogue (version 2.2).

Daily telemetric flows were computed all along a 1800 days mission starting from J2000.0 for two magnitude limits (in the GSC2.2 F-band at $\lambda \sim 0.71\mu$) $F_{lim} = 17$ and $F_{lim} = 18$ and for various areas of a single spectrum.

In these preliminary calculations, we implicitely assumed that the individual spectra of stars in the FoV do not overlap on the CCD. This is of course not true, especially in dense regions of the sky where spectra of neighbouring stars can easily overlap ; in very dense areas, due to this overlapping, a CCD pixel can be occupied by several spectra belonging to different stars.

The main aim of this note is to take into account this crowding effect in the determination of the telemetric flow. We consider, for the moment, that a single star spectrum has a unique possible area on the CCD.

We conclude this note by summarizing the work what can be done to improve further the estimations of the telemetric flows from RVS. All suggestions from inside and outside the RVS-WG are of course most welcome.

2 Spectra overlap : Fraction of pixels occupied in a 1x1 square degree of the RVS-FoV versus stellar density

In the RVS-FoV, assuming a focal length of 2.1 m for the mirror, the angular size of one pixel is 0.982" along scan and 1.473" across scan. A field of view of 1x1 deg hence contains 3666 pixels along scan and 2444 pixels across scan, i.e. a total of 8 959 704 pixels. The RVS adopted resolution (Paris 5th RVS-WG meeting in november 2002) is 11500 (corresponding to a sampling of 0.375 Å/pixel) so that a single spectrum spreads over $\sim$ 694 pixels along scan (along dispersion). According to the position of a star entering the FoV with respect to the CCD rows, we consider the possibility that a spectrum can extend over one or, at most, two CCD rows perpendicular to dispersion. The area S of a single spectrum is hence 694 or 1388 (694 x 2) pixels, corresponding to $S = 7.75 \times 10^{-5}$ or 1.55×10^{-4} square degree, respectively.

Adopting one of these two possible single spectrum area, we have performed Monte Carlo simulations of the occupation of the CCD pixels of a FoV of 1x1 deg for stellar densities N ranging from 0 to 10^6 stars per square degree. These simulations confirmed that the probability that a pixel is occupied by n distinct star spectra obeys a Poisson law :

$$P(n) = e^{-\lambda} \frac{\lambda^n}{n!} \tag{1}$$

where

$$\lambda = N \times S \tag{2}$$

the density N is the number of stars per square degree and the size S of a single spectrum is in square degree.

As far as telemetry is concerned, we want to send to the ground all pixels that are covered by the spectrum of one star or more. The probability that a pixel is occupied by one star ore more is :

$$P(n \geq 1) = 1 - P(0) = 1 - e^{-\lambda} = 1 - e^{-N \times S} \tag{3}$$

Figure 1 displays, as a function of the star density, the number of pixels occupied by at least one star spectrum in a FoV of 1x1 deg. For the two adopted single spectrum area, plots are given for two cases : 1) individual spectra do not overlap, in which case the total number of occupied pixels (each of them being occupied by only one star) is merely the product of the single size spectrum by the stellar density and 2) we take into account spectrum overlap, in which case the fraction of pixels occupied by at least one star spectrum is given by relation 3 above and the number of occupied pixels that must be sent to the ground is obtained by multiplying this fraction by the total number of pixels in a FoV of 1x1 deg.

As can be seen on the figure, reduction of the telemetric flow, which occurs when overlap of stellar spectra is taken into account, is important only at intermediate star densities. Indeed, at low densities, spectra overlap is not important and the number of occupied pixels in a FoV of 1x1 deg is nearly equal to the product of the star density by the size of a single spectrum. As the density increases, overlap of spectra becomes more and more important, reducing the mean number of occupied pixels per object : as a consequence the total number of occupied pixels increases with density less rapidly than in the case of no overlap. At even larger densities, all pixels of the FoV are occupied, the CCD is full and the telemetry budget is the same as the one computed in the case of no spectra overlap. The critical star density that fills the CCD is much larger if spectra overlap is taken into account : $N_{crit} \sim 32000$ and $N_{crit} \sim 65000$ stars per square degree for single spectrum area of 1388 and 694 pixels, respectively ; this must be compared to the values of 6455 and 12910 stars per square degree derived when spectra are assumed not to overlap.

To summarize, it can be seen from figure 1 that, when overlap of spectra is taken into account, the telemetry budget will be noticeably reduced in all FoVs in which the star density lies in the range 4000 to 20000 stars per square degree, if a single spectrum has an area of 1388 pixels, and in the range 8000 to 40000 stars per square degree, if a single spectrum extend over 694 pixels only.

3 Telemetry budget from the RVS

The procedure adopted to compute telemetric flows from RVS is described in [1] (section 2). The attitude of GAIA is computed, all along a 1800 days mission arbitrarily starting from J2000.0, every 120 s which corresponds to one transit time of the RVS-FoV along scan. The matrix attitude allows to derive the celestial coordinates (centre and extremities) of the RVS-FoV ; with the adopted timestep, we build a series of successive juxtaposed RVS-FoV (1 296 000 for a 1800 days mission). Each FoV is divided in 4 smaller cells of size 1 deg along scan and 0.8 deg across scan. At each time step, the matrix attitude is used to compute the galactic coordinates of each cell. The number of stars within each cell is the product of the star density (per square degree) at those coordinates, derived from star counts in the

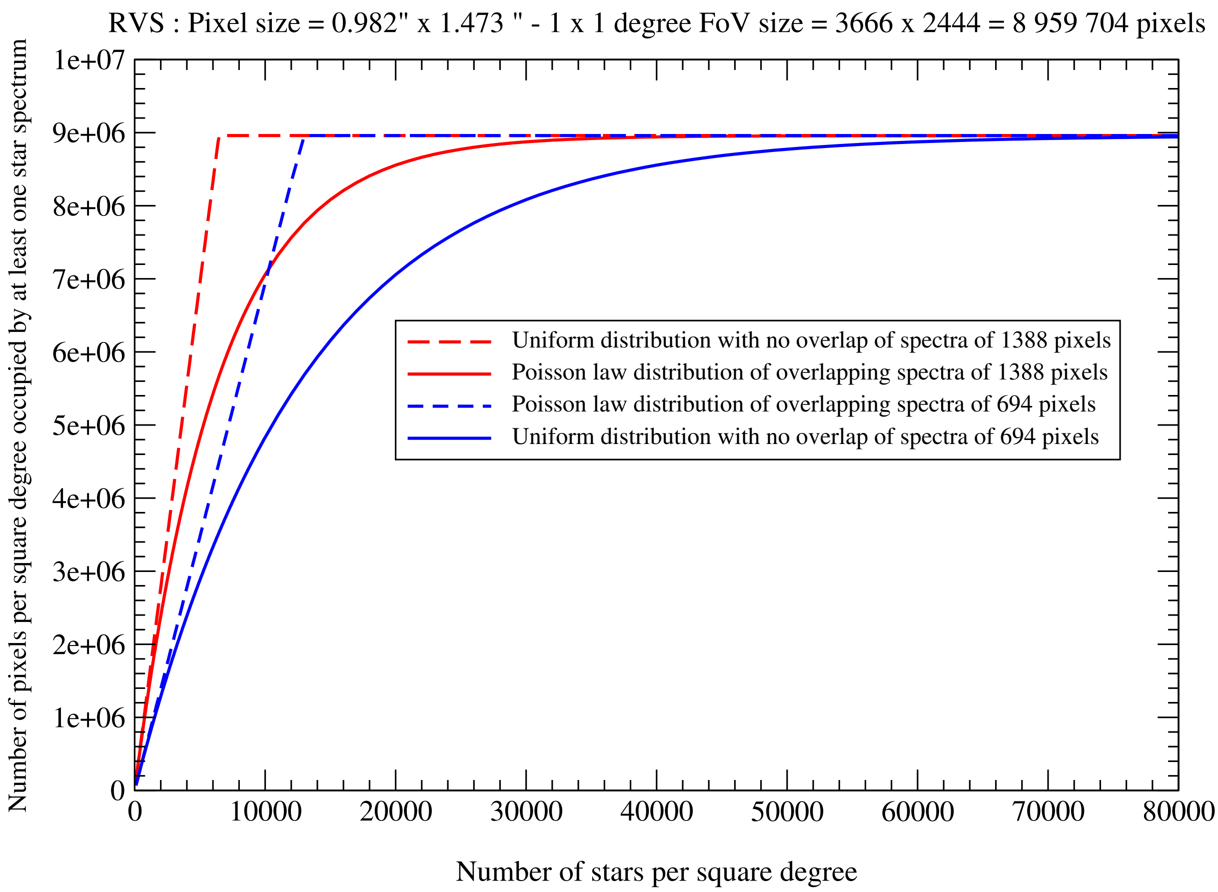

Figure 1: Number of pixels per square degree of the RVS-FoV occupied by at least one star spectrum versus the star density.

GSCII (version2.2) star catalogue, by the surface of the FoV cell (normalization to the area of the cell).

The number of pixels within each cell that must be sent to the ground has been computed in both cases : without and with spectra overlap. In the first case, the number of occupied pixels in a cell is the product of the number of stars within the cell by the size of a single spectrum until the cell is full ; beyond, it is the number of pixels within the cell. In the second case, the number of occupied pixels in a cell is the product of the total number of pixels within the cell by the fraction of occupied ones given by relation 3 above. The number of bits per FoV to be downloaded is obtained by summing over the 4 FoV cells, each pixel corresponds to 16 bits of data.

3.1 Comparison of the telemetry budget computed with and without spectra overlap

The telemetric flows integrated over one day are plotted on figure 2 versus time for a magnitude limit $F_{lim} = 17$ and a single spectrum area of 1388 pixels (2 CCD rows). The blue line corresponds to flows computed with no overlap of spectra while the red line is obtained when spectra overlap is taken into account. In all our calculations, it is assumed that the CCD (6 in the present design) are summed on-board before sending data to the ground.

The reduction of the telemetric flow when overlap of spectra is taken into account ranges from 10 to 30 % all along the mission ; integrated over the whole mission, the telemetry budget is reduced by 23 %. This value depends slighly on the magnitude limit of observation : the reduction factor is 24 % for $F_{lim} = 16$ and 20 % for $F_{lim} = 18$.

3.2 Telemetric flows versus magnitude

Variations of the telemetry budget versus the magnitude limit of observation is shown on figure 3. Daily telemetric flows, computed by taking into account spectra overlap, are plotted versus time for four magnitude limits of observation $F_{lim} = 14, 16, 17, 18$. Telemetric flows were calculated with a single spectrum area of 1388 pixels. We also show on the figure the telemetric flow per day that would be allocated to RVS, assuming an average instantaneous rate of 0.25 Mbits/s (derived by assuming a visibility of 8 hours per day of the ground station, a telemetric power of 3 Mbits/s during the visibility window, and a quarter of this power allocated to RVS) and 4 values for the data compression factor : 1, 3, 6 and 10, which correspond to an allocated telemetry budget per day of $2.2 \times 10^{10}, 6.5 \times 10^{10}, 1.3 \times 10^{11}, 2.2 \times 10^{11}$ bits, respectively.

If no compression factor is available, it is possible to download nearly all the data, exept during the telemetry peaks, for all stars with $F \leq 14$. A compression factor of 3 allows to download all data outside the telemetric peaks up to $F = 16$. Saving nearly all data up to $F = 17 - 18$ requires a data compression factor of 10, beyond the present state of the art ; furthermore, even in that case, a substantial fraction of data would be lost during all (23) peaks of telemetry.

3.3 Mass memory on board : Data transmitted and lost

The telemetric flows plotted on on figure 3 are flows integrated over one day of scanning. The instantaneous telemetric flow integrated over much shorter periods of times presents fairly large variations with time. This is clealy illustrated on figure 4 where is plotted, for a magnitude limit of observation $F = 17$ and a single spectrum area of 1388 pixels, the variation of the telemetric flow integrated over a period of 2 minutes, one FoV transit, during the day 34 (from J2000.0) which corresponds to the minimum value of the daily telemetric flow for the whole mission. As in figure 3, the telemetric flow allocated to RVS is also plotted, still assuming an instantaneous rate of 0.25 Mbits/s and 4 values for the data compression factor : 1, 3, 6 and 10, which correspond to an allocated telemetry budget per FoV transit of $3 \times 10^7, 6 \times 10^7, 1.2 \times 10^8, 3 \times 10^8$ bits, respectively.

The figure clearly shows that, if a single spectrum extends over two CCD rows (1388 pixels), the input telemetric flow from RVS is always above the limit (0.25 Mbits/s) allocated to the instrument. A rather efficient data compression process is required to increase the amount of data that can be saved and sent to the ground. If the input telemetric flow is reduced, due to data compression, there are several periods during the day where this input flow is below the allocated flow. Provided that data can be stored on board, on a mass memory for instance, it is hence possible to take advantage of these periods to

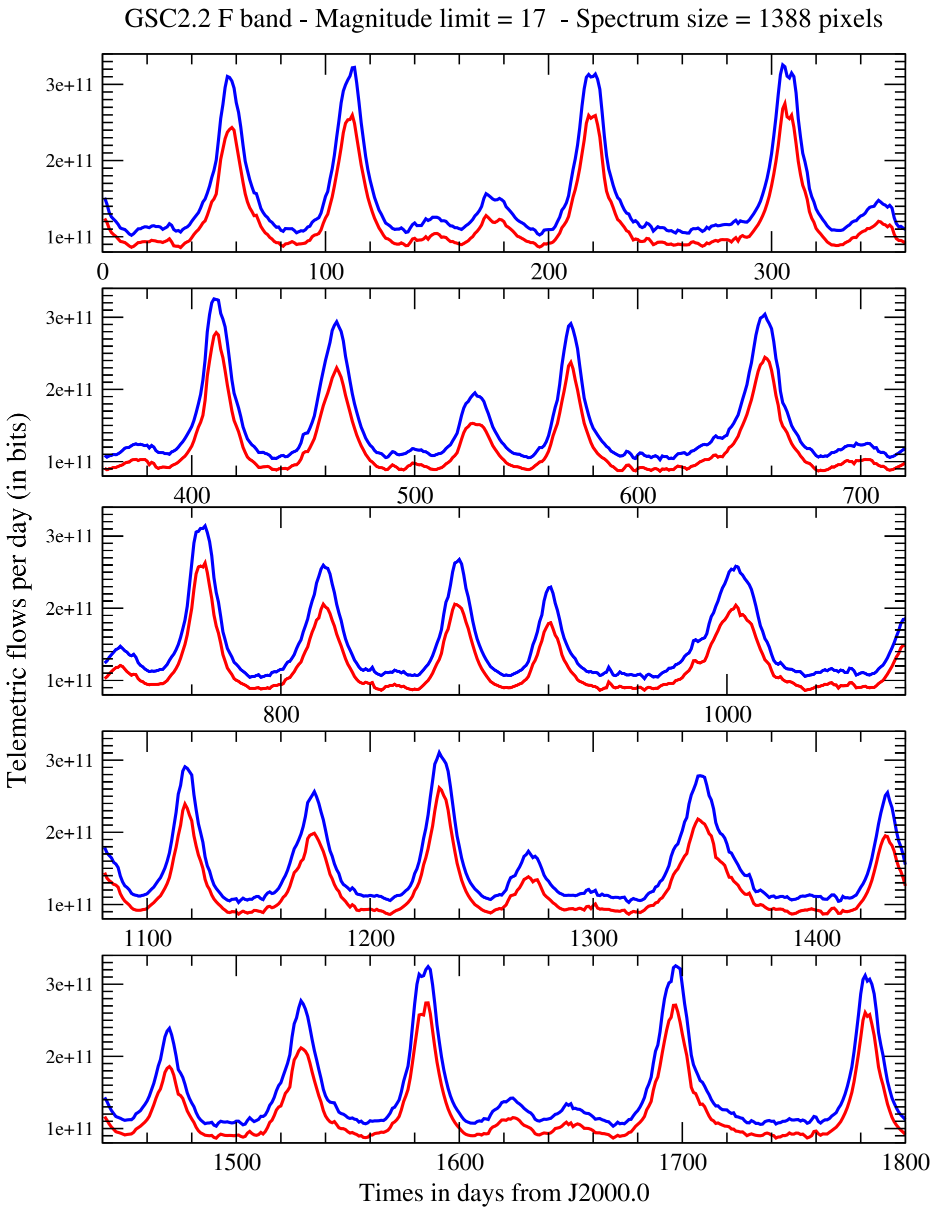

Figure 2: RVS telemetric flows integrated over one day from counts in the GSC-2.2 F-band up to a magnitude limit of 17 and for a spectrum size of 1388 pixels. Blue curve : no overlap of spectra. Red curve : overlap of spectra taken into account.

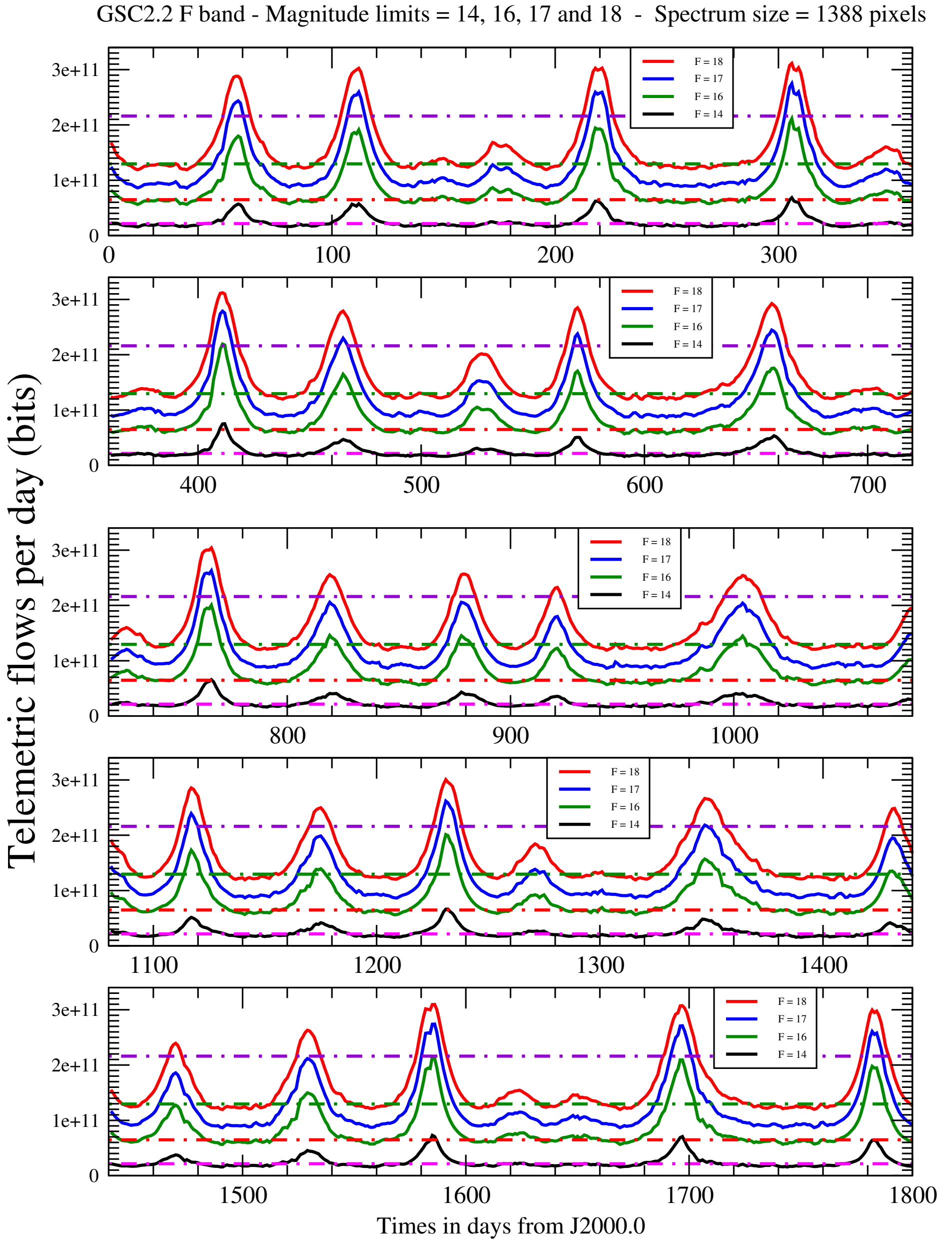

Figure 3: RVS telemetric flows integrated over one day from counts in the GSC-2.2 F-band, taking into account overlap of spectra a and assuming a single spectrum size of 1388 pixels. Plots are given for 4 magnitude limits $F = 14, 16, 17, 18$. The 4 dash-dotted lines show daily telemetric flows of $2.2\times10^{10}, 6.5\times 10^{10}, 1.3\times10^{11}, 2.2\times10^{11}$ bits, corresponding to a telemetry rate of 0.25 Mbits/s allocated to RVS and available data compression factors of $1, 3, 6$ and 10.

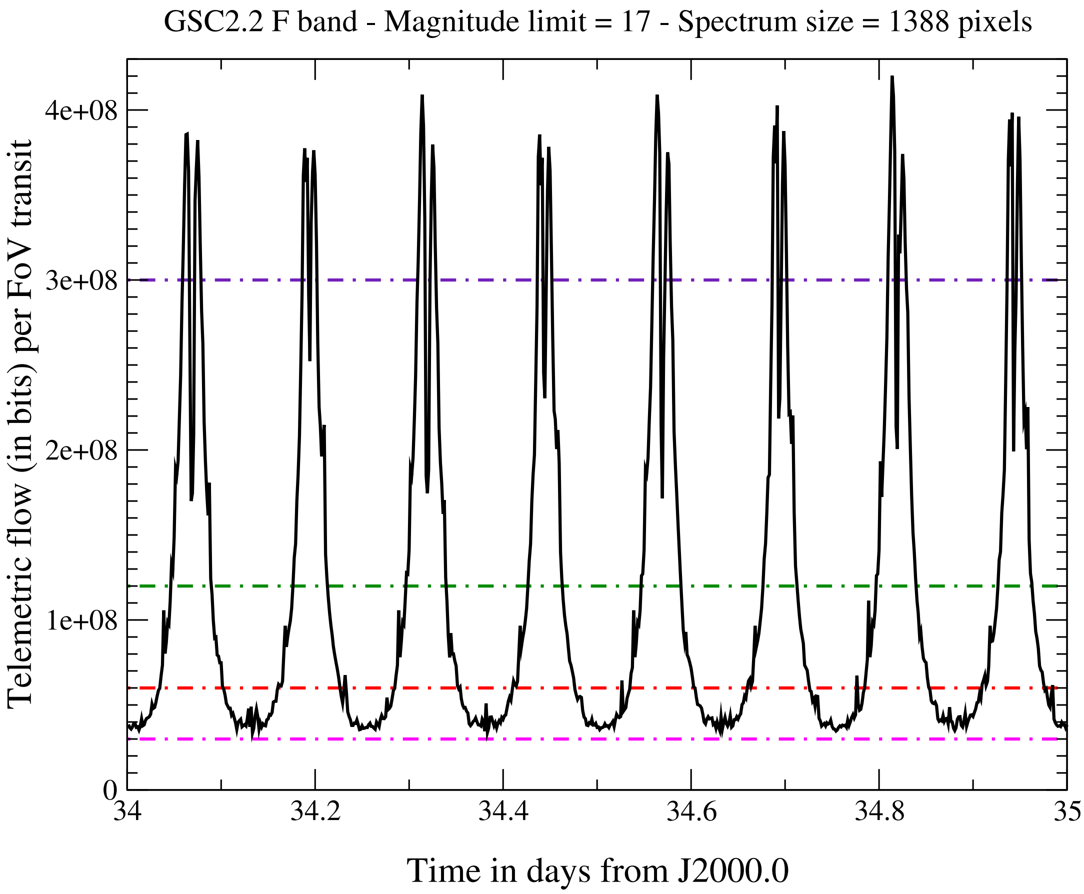

Figure 4: RVS telemetric flows integrated over one FoV transit (2 minutes) during day 34 (from J2000.0). Star counts are performed up to magnitude limit $F = 17$ from the GSC-2.2 F-band and the telemetric flow is computed by taking into account overlap of spectra a and assuming a single spectrum size of 1388 pixels. The 4 dash-dotted lines show daily telemetric flows of $3 \times 10^7, 6 \times 10^7, 1.2 \times 10^8, 3 \times 10^8$ bits, corresponding to a telemetry rate of 0.25 Mbits/s allocated to RVS and available data compression factors of $1, 3, 6$ and 10.

download data stored as the telemetric flow was above the limit. At any time during these "low telemetric flow" periods, the surplus of data that can be saved is the difference between the allocated flow and the true input flow. Of course, the larger the compression factor, the larger the amount of data saved.

We have determined the fraction of data produced by RVS that can be saved (or lost) all along a 1800 days mission with the following assumptions :

1) A storage capacity on board is allocated to RVS data : 4.32×10^{10} bits (two days).

2) Data are tranferred from Gaia to the Earth every day, but only during one third of the day (8 hours), when the satellite is visible from the antenna which receives the data.

3) Data transmission, with a rate of 3 Mbits/s for scientific data, is possible only during the last third of the day (in reality, the period of the day favourable to transmission changes with time).

4) During the time interval t and t + dt, the flow of input data either feeds the mass memory, if it is not full, or is entirely lost (it can of course replace older data, but this equivalent in terms of data losses) ; furthermore, if this time interval belongs to the period where downloading is allowed, the data from RVS are saved with the rate of 0.75 Mbits/s assumed to be allocated to the instrument.

The amount of data produced by the RVS and the fraction of data saved and lost are obtained through summation over the whole mission ; this has been done by using a timestep of integration equal to a FoV transit time, i.e. 2 minutes.

We present in table 1 the fraction of data saved throughout the whole mission for magnitude limits $F = 14, 16, 17, 18$ and 4 values of the data compression factors : 1 (no compression), 3, 6 and 10. We assume that a mass memory of 4.32×10^{10} bits is allocated to RVS, which corresponds to 2 days of storage at a mean rate of 0.25 Mbits/s.

Table 1: Fraction of data saved (in %) for various compression factors and magnitude limits. Single spectrum area : 1388 pixels. Allocation to RVS : a telemetry rate of 0.25 Mbits/s and a storage capacity on board of 4.32×10^{10} bits (two days).

Compression factor	$F = 14$	$F = 15$	$F = 16$	$F = 17$	$F = 18$
1	89.3	44.2	26.1	17.8	13.5
3	100	100	78.3	53.5	40.6
6	100	100	100	100	81.2
10	100	100	100	100	100

A compression factor of 3 allows to save all data up to magnitude $F = 15$, but less than half of them if one wants to observe stars as faint as F = 17, as required by some of the scientific issues addressed by the RVS. Saving all data up to these faintest objects would require either a compression factor of 6, not likely in the present state of the art, or a more poweful antenna and/or a larger visibility window of the ground station.

Another way to reduce the telemetric flow consists to limit, within each FoV cell, the number of stars whose spectra are sent to the ground. A glance at figure 1 shows that, with a single spectrum area of 1388 pixels, the CCD is full as soon as the star density reaches 30 000 stars per square degree. Table 2 displays the fraction of data that can be saved if, within crowded FoV cells, the number of stars whose spectra are sent to the ground are limited to 20 000 and 10 000 stars per square degree. The fraction of data saved in the case where the star density is not limited (see Table 1) is also given for purpose of comparison. Limiting the density in crowded FoVs to 20 000 has practically no influence on the fraction of data saved ; a minor increase of this fraction by only a few percent (depending on the magnitude and the compression factor) is obtained if the density limit is reduced to 10 000 stars per square degree.

Figures 5 and 6 illustrate well why reducing to 10 000 the density of stars to be downloaded has only a small impact on the amount of data that can be saved throughout the whole mission. Telemetric flows per FoV transits are plotted for days 34 (top) and 411 (bottom) which, in our model, correspond to the minimum and the maximum telemetric flow per day, respectively. Telemetric flows are plotted both cases : the star density within the FoV is not limited (black curve) and is limited to 10 000 (red dotted curve). Figure 6 shows a zoom in abcissa (i.e. time) of figure 5. As can be seen on both figures, the two flows differ only in the FoVs with density larger than 10 000 stars per square degree, i.e. along or

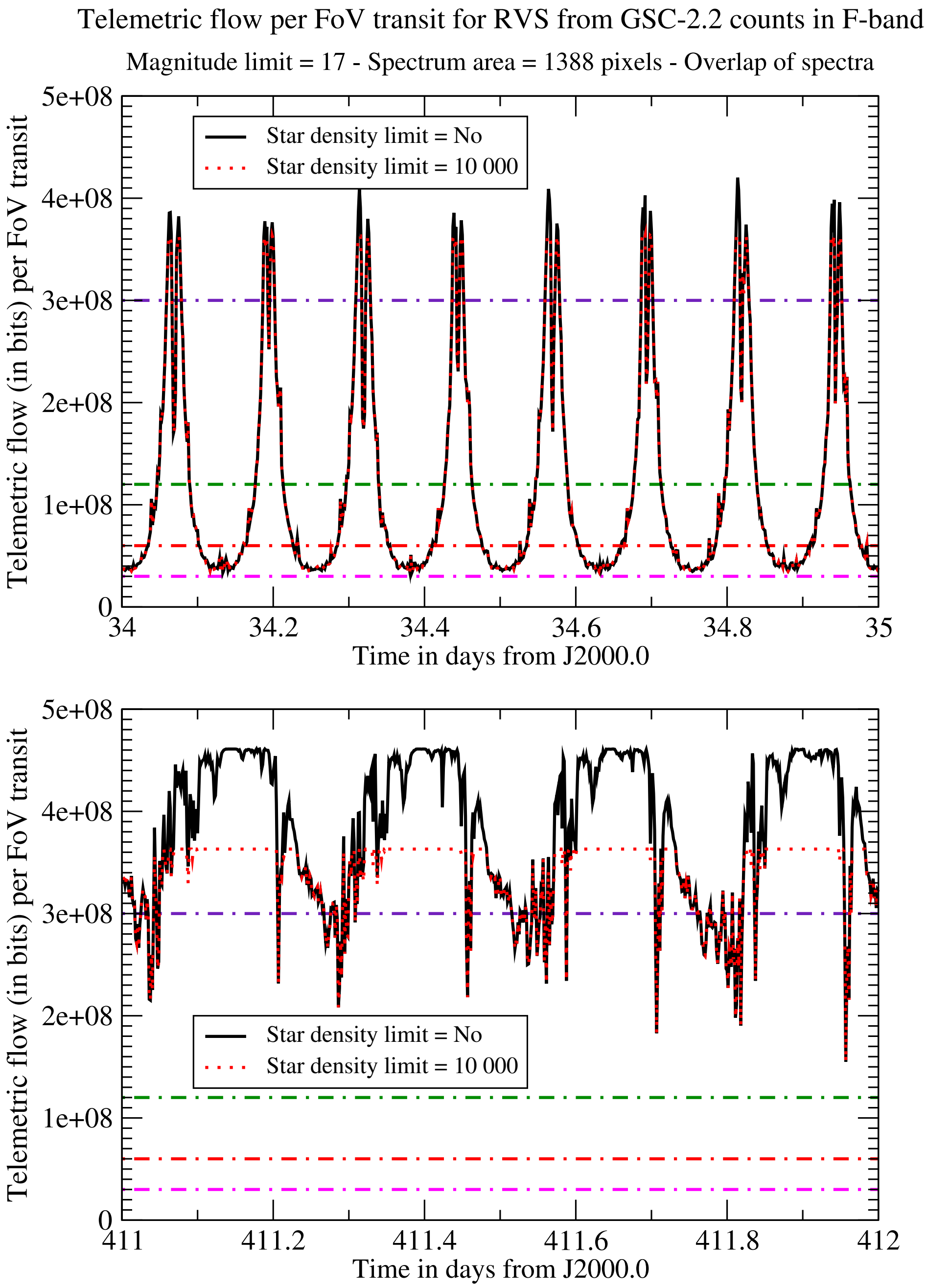

Figure 5: RVS telemetric flows integrated over one FoV transit (2 minutes) during days 34 and 411 (from J2000.0). Star counts are performed up to magnitude limit $F = 17$ from the GSC-2.2 F-band and the telemetric flow is computed by taking into account overlap of spectra a and assuming a single spectrum size of 1388 pixels. As in previous figures, the 4 dash-dotted lines show daily telemetric flows of $3 \times 10^7, 6 \times 10^7, 1.2 \times 10^8, 3 \times 10^8$ bits, corresponding to a telemetry rate of 0.25 Mbits/s allocated to RVS and available data compression factors of $1, 3, 6$ and 10.

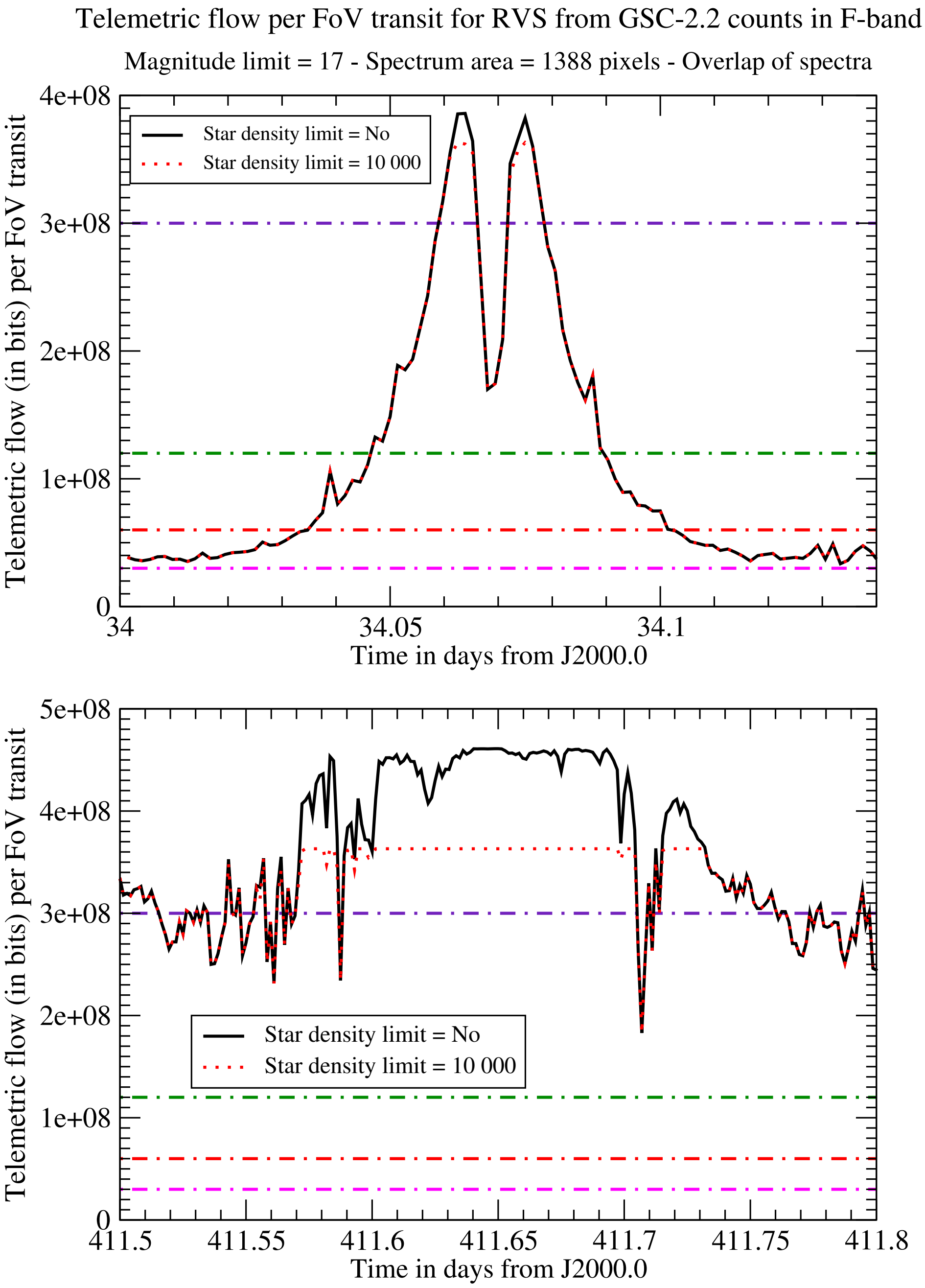

Figure 6: Zoom of previous figure : RVS telemetric flows integrated over one FoV transit (2 minutes) during parts of days 34 and 411 (from J2000.0). Star counts are performed up to magnitude limit $F = 17$ from the GSC-2.2 F-band and the telemetric flow is computed by taking into account overlap of spectra a and assuming a single spectrum size of 1388 pixels. As in previous figures, the 4 dash-dotted lines show daily telemetric flows of $3 \times 10^7, 6 \times 10^7, 1.2 \times 10^8, 3 \times 10^8$ bits, corresponding to a telemetry rate of 0.25 Mbits/s allocated to RVS and available data compression factors of $1, 3, 6$ and 10.

Table 2: Fraction of data saved (in %) for various compression factors and maximum values of the number of star per square degree for which spectra are send to the ground. Data are given for three magnitude limits. Other parameters are the same as used in table 1

Compression factor	Density limit	$F = 16$	$F = 17$	$F = 18$
1	10 000	26.6	18.6	14.6
1	20 000	26.1	17.9	13.6
1	No limit	26.1	17.8	13.5
3	10 000	79.7	55.9	43.7
3	20 000	78.4	53.7	40.9
3	No limit	78.3	53.5	40.6
6	10 000	100	100	87.3
6	20 000	100	100	81.8
6	No limit	100	100	81.2
10	No limit	100	100	100

close to the galactic plane. This occurs only during part of the observational time ; the fraction of the time during which the telemetry rate is reduced being larger when scanning proceeds along or near the galactic plane (day 411) than when it proceeds nearly perpendicular to it (day 34). Averaged over a 1800 days mission, the gain in telemetry is only a few percent.

4 Conclusion and future work

This note mainly shows that taking into account overlap of spectra between neighbouring stars in the FoV reduces the number of pixels that must be sent to the ground and hence, the instantaneous telemetric flow from the RVS, with respect to the situation where spectra are assumed not to overlap.

However, the reduction of the telemetric flow is important only at intermediate star densities in the FoV. The range of densities for which this effect is significant depends on the area of a single spectrum on the CCD ; typically it extends from 4000 to 20000 stars per square degree for a single spectrum extending over two CCD rows (1388 pixels), or in the range 8000 - 40000 stars per square degree if a single spectrum extends to only one row.

Integrated over a 1800 days GAIA mission duration, the gain in telemetry lies in the range 20-25 % for magnitude limits of observations between 16 and 18. If a mass memory corresponding to 2 days of storage is implemented on board and if data can be compressed by a factor 3, it is possible to save, throughout the whole mission, 78%, 53% and 40% of data for magnitude limits of observations F_{lim} = 16, 17 and 18, respectively. A further small increase between 2 and 3% of the amount of data saved can be obtained if one limits the density of stars whose spectra must be sent to the ground to 10 000.

If no solution exists to either increase the compression factor and/or the ground visibility window and/or the antenna power, loosing part of the data would be unavoidable and a source/pixel selection strategy would have to be implemented to download the most interesting data on scientific grounds. Among possible strategies, we could envisage to download a single CCD row per spectrum (when the stars are well centered on a pixel) or send only part of the spectra for certain magnitude range.

These assumptions lead of course to a decrease of the telemetric flow. In this case however, the CCDs in the RVS-FoV are covered by spectra having different sizes ; the densities of objects will also be different for each class of spectrum. Treatment of spectra overlap is more complicated than presented in the above section and has to be refined to allow for different size spectra and object densities. This will be done in a further note.

Several other improvements of our telemetry model, mainly concerning star counts, are planned in the near future :

- Performing star counts versus magnitude in the RVS wavelength band so as to reach the real observing

conditions. As a first step, because the RVS band is close to the photometric Cousin's I-band, we will use the point source DENIS catalogue when it will be implemented in the CDS in Strasbourg. The main drawback is that the DENIS survey does not covers the entire sky.
- A way to overcome this gap can be achieved, as a second step, through conversion of F-magnitudes into I magnitudes for the whole GSC-2.2 catalog before performing star counts. This can be done, and work is in progress, by establishing correlations between I and F magnitudes for selected sets of well identified photometric standards. This will however introduce some uncertainties in magnitude determinations because color-color relations sometimes depend on spectral types of stars, an information not available from star catalogues.
- It would be also very useful, for purpose of comparison, to determine RVS telemetric flows by using star counts from other catalogues. We are intending to do it in the future by using the the USNO-B1 catalogue version implemented in the CDS.
- At last, whatever the star catalog used to perform star counts on the sky, it would be worth to do it with a better spatial resolution to take into account the fact that the star distribution is highly non uniform at a scale lower than a square degree. Counts within boxes of 6x6' are beeing undertaken from the GSC-2.2 catalog : they are limited to low galactic latitudes, in the range $-10\,\text{deg} \leq b \leq +10\,\text{deg}$, to save computing time.

Whatever the improvements that can be brought to the telemetry model for RVS (crowding within the FoV, star counts...), it is clear that, under the assumptions adopted in the present note, the flow of data produced by the spectrometer will strongly exceed the capacity of of saving and sending data to the ground, even if a mass memory is implemented on board. This is particularly true during the peaks of telemetry which occur when the instruments scan the sky along great circle close to the galactic plane. Because a limited fraction of data can be saved and downloaded, according to the available compression factors, it will be necessary to define one (or several) observational strategy (selection of objects to be observed). This addresses the question of the scientific objectives of RVS and the priorities among them and must be one of the main task of the RVS-WG for the future. The telemetry model presented here can help to check several possible selection algorithms and their impact on the RVS telemetry budget.

Acknowledgements Star counts have been performed by using the version of the GSC-II catalogue settled at the Centre de Données Stellaires de Strasbourg and the software of the Vizier database environment.

References

[1] Viala Y.P., Katz D., Morin D., Ochsenbein F. 2003, Telemetric flows for the RVS computed by using star counts from the GSC-2.2 catalogue, Technical Report RVS-YV-001

[2] Viala Y.P., Katz D., Morin D., Ochsenbein F. 2004, Telemetric flows on GAIA-RVS estimated from star counts from the GSC-2.2 catalogue, ASP Conference Series Vol. 298, "GAIA Spectroscopy, Science and Technology", ed. U. Munari, p. 105

[3] Lindegren, L. 2000, Attitude parametrization for GAIA, Technical Report SAG-LL-030, Lund Observatory

[4] Lindegren, L. 2001, Calculating the GAIA scanning law, Technical Report SAG-LL-035, Lund Observatory