GAIA TELEMETRY RATE SIMULATIONS: A FIRST LOOK AT THE COMPLETE PICTURE

U. Lammers

Research and Scientific Support Department, ESTEC/ESA, Keplerlaan 1, 2200AG Noordwijk, The Netherlands

ABSTRACT

The first study on science telemetry data rates and volumes considering all instruments (Astro, RVS, MBP) has been carried out. We present results from simulations over a nominal 5 year mission duration which give for the first time an overall picture and assessment of the telemetry rate situation in view of expected massive payload-generated data volumes on the one hand, and limited downlink capabilities on the other hand. Two central quantities of fundamental importance for the total amount of downlinkable data are the daily satellite-ground station contact times and the size of the onboard Solid-State-Mass-Memory. A first attempt has been made to optimise both quantities in the context of this work and those results are likewise presented. The calculations depend on a large number of parameters and assumptions many of which are not frozen but will evolve with time as the satellite's design and mission profile consolidate in subsequent project phases. Thus, the presented results and conclusions can not be regarded as static or final but repeated studies are warranted in due time.

Key words: Gaia; Telemetry; Data Rates; Data Volumes; Solid State Mass Memory; Data Loss; Compression.

1. INTRODUCTION

Relatively little effort has gone into detailed telemetry (TM) rate studies so far. Past works have considered either Astro or RVS in isolation and exhibited a number of other shortcomings. For MBP merely very crude estimates existed up until now. This contribution describes the main results of a recently completed study (Lammers 2004) that considers for the first time all instruments and uses the latest mission and design parameters to investigate the TM rate and volume situation from a global angle. In addition, the simulations take into account the presence of a Solid State Mass Memory (SSMM) as an important component in the onboard data acquisition and downlink chain. In a first group of calculations the SSMM size is fixed at 396 Gbit which is the mid-2004 Gaia Parameter Database (de Bruijne et al. 2005) value for this quantity. This and other existing size estimates exhibit a clear and unsatisfactory *ad hoc* character. Here, an attempt is made to arrive at a concrete size recommendation based on minimising the total end-of-mission data loss under given boundary conditions.

A factor with predominant influence on the amount of downlinkable data is the visibility of the receiving ground station from the satellite. This aspect has not been investigated in detail yet but various sources just quote "a minimum duration of 8 hours". In the context of this work positions and lengths of the visibility windows over the 5 year nominal mission duration have been optimised and used as input to the TM rate simulations.

2. SIMULATION FRAMEWORK

A new and dedicated framework to study TM rates and volumes has been designed and implemented. We present here only a very brief outline of the scheme and refer to Lammers (2004) for more detailed information.

A central element is a galaxy model that allows the prediction of stellar densities in the three telescopes' FOVs at any time t. We employ a simple, numerically efficient model (Lindegren 1998) that provides cumulative star densities for any given limiting magnitude and celestial direction in the two standards bands V (Johnson) and I (Cousin). Using this and known colour-colour transformations, star counts in the three relevant Gaia photometric bands G, GS, and $GM861$ can be derived. The model has been normalised such that it yields an all-sky integrated number of 1 billion objects with the Gaia-1 design (ESA 2000) which then results in a total number of objects to be observed with the more sensitive Gaia-2 design as given in Table 1. These numbers have been selected as baseline for the simulation runs.

The variations of the telescopes' viewing directions over time is dictated by the Nominal Scanning Law (NSL) which forms the second main element of the simulation framework. We use here the formulation of Lindegren (2001) and solve the differential equations by standard numerical methods.

Each detected and confirmed/cross-matched object will generate a distinct amount of data onboard during its transit over the focal planes. Sizes and structures of this

Table 1. Total number of objects in the three bands as predicted by the galaxy model using the Gaia-2 design and the indicated limiting magnitudes.

Instrument	band	limit	number of objects
Astro	G	20.0	1 458 153 266
Spectro-MBP	GS	20.15	1 454 670 069
Spectro-RVS	$GM861$	17.5	542 864 601

science data stream are determined by a number of factors and summarised in telemetry data models for Astro, MBP, and RVS respectively. A fixed number of object attributes constitutes a first data block for each object. A second block is then made up from CCD readout sample data the size of which varies with the object's apparent magnitude. As an example, a G =–20 star will generate about 2700 bits of data in each Astro field (ASM+AF+BBP+attributes) - see Lammers (2004) for further details.

Finite detection and cross-matching probabilities in all instruments as well as the effect of crowding in RVS and MBP is taken into account by suitable means. This is an essential pre-requisite for calculating realistic TM rates and data volumes especially in regions of high and extreme stellar densities.

Given a chosen time granularity, the NSL is integrated accordingly and the telescope viewing directions calculated. The number of objects in each FOV is then derived with the galaxy model and a total per-object data size derived for each focal plane transit. The simulation proceeds then by compressing each acquired data packet by a given fixed factor and storing it in the SSMM. This unit is characterised by its size and an assumed strict FIFO-driven operation flow: During periods of non-visibility of the ground station data packets are stored sequentially in the buffer up to its maximum capacity. Beyond that point any further incoming data will be considered as lost. When the ground station becomes visible the buffer is gradually emptied by downlinking data packets in the order of their arrival with a nominal downlink speed. This continues until the next non-visibility period starts.

The simulations are dependent on a large number of scalar parameter values all of which are obtained from the Gaia Parameter Database. A few key value are: Downlink rate: 5 Mbit/s, Common compression factor: 3, SSMM size: 396 Gbit

3. OPTIMISATION OF GROUND STATION COVERAGE

In addition to the above listed key parameters, the amount of data that can be downlinked is largely determined by the visibility of the receiving ground station from the location of the satellite over time. Cebreros, Spain (−4.248° E longitude, +40.430° N latitude, ≈ 1000 m altitude) is baselined to act as Gaia's primary ground station with an often quoted minimum visibility of 8h per day.

Given Gaia's orbit around L2, we optimise the durations and positions of the visibility windows and obtain the following results: Cebreros possesses a mean visibility of 11.03 ± 1.84 h over the entire mission duration almost independent of initial orbit phase parameters. The usage of a station at a geographical location other than Cebreros but closer the the equator (New Norcia) improves the situation but only marginally. A clear advantage would be the use of two (or more) stations simultaneously with a flexible switch-over policy. For the simulations we have chosen to use the optimised Cebreros visibility windows.

4. RESULTS

We describe here the main simulation results that have been obtained using the framework, parameters, and assumptions described in the preceding sections.

The instantaneous total data rate (not shown here but see Figure 10 in Lammers 2004) over the mission duration varies with time in a complex manner. A number of distinct, relatively wide peak structures can be identified which are associated with scans along the galactic plane. In these areas of extreme stellar density the instantaneous total rate reaches maximal values in excess of 25 Mbit/s which can last for several days. Assuming a compression factor of 3 and a sustained mean downlink rate of 5 Mbit/s × 11 h/24 h = 2.29 Mbit/s it is clear that the data rate exceeds the available downlink capabilities by more than a factor 20 Mbit/s/3/2.29 Mbit/s ≈ 3 in those periods. For moderate sizes of the SSMM it is thus anticipated that data loss will be inevitable. On the other hand, a mean total rate of around 5 Mbit/s can be identified which corresponds to roughly 3/4 of the available downlink budget (5 Mbit/s/3/2.29 Mbit/s ≈ 0.75). Consequently, the total available downlink budget is not exhausted and, hence, with a SSMM of sufficient size it shall be possible to avoid data loss completely: During galactic plane scans the buffer's fill grade shall monotonically grow and in the periods following a galactic plane scan the buffer is successively emptied again by making use of excess downlink capabilities.

Figure 1 shows the individual data rates of one Astro unit, RVS, and MBP over the 5 year nominal mission duration using an integration time of $T_{\text{int}} = 12$ h. The single curve below each of the four rate graphs depicts on the same time axis the galactic latitude of the spin axis in units of degrees. The two dotted horizontal red lines mark the (arbitrary) ±75° limits. If one defines a galactic plane scan as the period of time in which the spin axis is located in a cone centred around either of the galactic pole directions ($b = \pm 90°$) with a full opening angle of 15° then galactic plane scans occur 14 times, viz. around the days 55, 110, 219, 307, 409, 464, 570, 657, 764, 1118, 1230, 1583, 1696, and 1783. There is a clear correlation between galactic plane scans and maxima in the total as well as individual data rates.

During some galactic plane scan periods the data rate

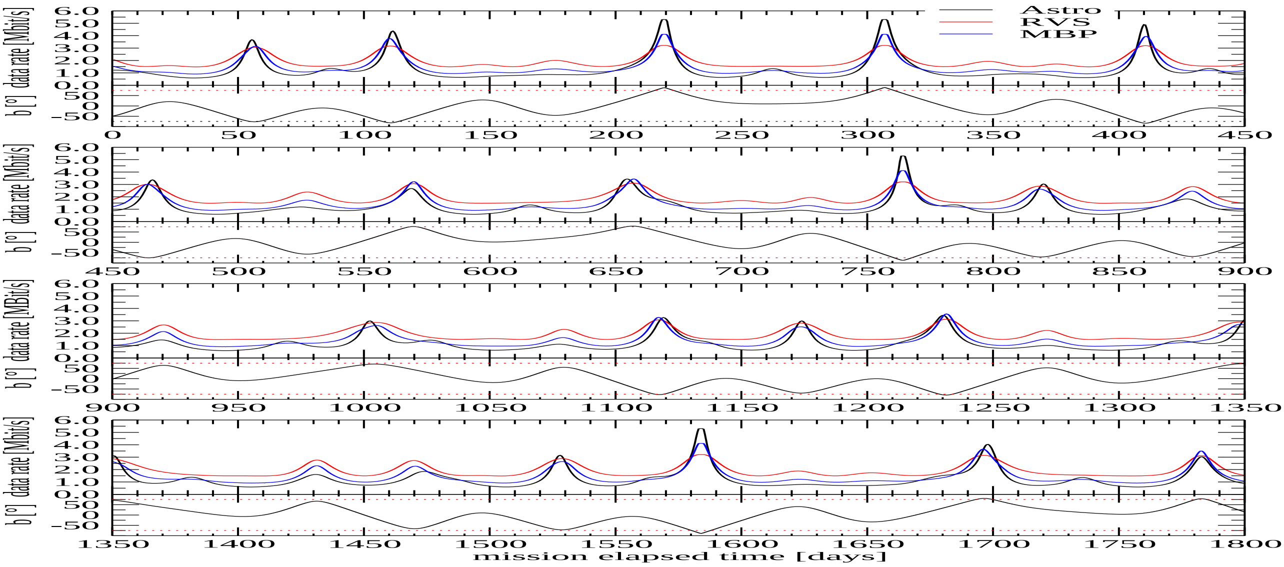

Figure 1. Mean data rates of one Astro unit, RVS, and MBP over the entire 5 year nominal mission duration. The plots beneath each of the four partial rate curve graphs shows the galactic latitude of the spin axis on the same time axis. The dotted horizontal red lines mark the ±75° limits.

from each Astro units reaches almost 6 Mbit/s and peak values for RVS and MBP are around 3–4 Mbit/s. Outside galactic plane scan periods mean values for Astro/MBP are well below 1 Mbit/s (e.g., $\approx$ 0.9 Mbit/s around day 490) while RVS is at a constant high average of $\approx$ 1.5 Mbit/s. This reflects the fact that each star spectrum occupies a constant large, and magnitude independent amount of data in the RVS telemetry. In contrast, faint stars in Astro, and MBP generate less data because of much smaller readout windows.

Table 2 summarises the mean uncompressed data rates and associated standard deviations together with the total, accumulated uncompressed data volumes at the end of the mission and these numbers referred to the Astro-1 unit. A total size of >110 TB demonstrates vividly that Gaia will generate an unprecedented amount of data compared to any other space science mission. RVS's data rate lies about 60% above Astro's (per unit) but the rate is extremely variable and in low-density regions the RVS rate is typically 2–2.5 times higher than Astro/MBP (see Figure1). Despite of the TM-limiting effect of crowding MBP generates about 20% more telemetry than each Astro unit, a clear consequence of the larger faint star windows of MBP compared to Astro.

Module	mean rate [kbit/s]	sigma [kbit/s]	total volume [TB]	rel. to Astro-1
Astro-1	1216.8	2057.6	23.98	1.00
Astro-2	1216.9	2060.9	23.98	1.00
RVS	1906.2	946.6	37.57	1.57
MBP	1434.4	1427.0	28.27	1.18
Total	5774.3	4155.3	113.81	4.75

Table 2. Mean uncompressed data rates and total data volumes for a 5 year nominal mission duration.

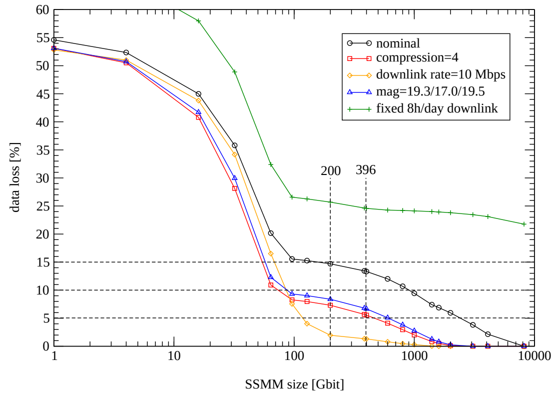

Figure 2. Total data loss as a function of the SSMM size for a 5 year mission duration depicted for a number of different scenarios. The horizontal lines mark the 5%, 10%, and 15% limits. The vertical lines show the current industrial baseline and Parameter Database values.

Figure 2 shows the accumulated, end-of-mission total data loss L as a function of the SSMM size (abscissa is logarithmic) for a number different scenarios. The black line denoted 'nominal' represents L calculated under the assumption that all relevant parameters have values currently considered as baseline in the current mission phase. The other curves show the evolution of L when one (and only this one) of the key parameter values is changed from its baseline value to the value indicated in the figure legend. The blue curve refers to reducing the limiting G and GS magnitudes to 19.3 and 19.5 required to arrive at a Gaia-2 total number of observed objects of 1 billion in Astro/MBP with the galaxy model of Lindegren (1998). Also, the RVS limiting magnitude has been relaxed from 17.5 to 17.0 which roughly corresponds to reducing the number of observed objects by 1/3. Using the assumptions that

1. a total data loss of less than 5% would be acceptable from a scientific point of view
2. SSMM sizes larger than 1000 Gbit appear unfeasible to realise from a technological, cost, and programmatic point of view

the following conclusions can be drawn:

- Any SSMM size of less than 100 Gbit results in an unacceptable data loss of more than 10% in all cases considered. It may be possible to bring L down to below 10% for sizes not significantly smaller than 100 Gbit through a *combination* of parameter changes (e.g., a higher compression factor and a lowering of the limiting magnitudes), however, realising a 100 Gbit SSMM appears to pose no technological problem and, hence, this should be regarded as the absolute minimal size.
- It is theoretically possible to reach $L = 0$, i.e., avoid data loss completely, but this would require prohibitively large sizes in all considered cases. For the nominal case, for instance, a size of around 8000 Gbit would be needed.
- In the nominal case the data loss with currently envisaged SSMM sizes of 200 Gbit (industrial baseline) and 396 Gbit (Parameter Database) is larger than 10% (14.7% and 13.3% respectively) and, thus, unacceptable. Increasing the compression factor to 4 or lowering the limiting magnitudes reduces the data loss considerably but only a combination of the two would bring it below the acceptance level of 5%.
- The red and blue curve are almost identical, hence, for all SSMM sizes increasing the compression factor from 3 to 4 has the same effect as reducing the limiting magnitudes down to a level that yields 1 billion objects. This coincides with expectations as both measures would result in roughly the same reduction of data that have to be downlinked, i.e., 1.5/4=0.375 $\approx$ 1/3.
- Clearly, the parameter with the most beneficial influence on the data loss is the downlink data rate. Doubling the nominal value of 5 Mbit/s to 10 Mbit/s would result in reaching the data loss acceptance level already with SSMM sizes only slightly larger than the minimum of 100 Gbit. At 200 Gbit L is even as low as 2% and drops further to 1.3% at 396 Gbit. The dominant influence of the downlink rate on L can be understood by realising that once the SSMM is full, increasing the downlink speed is the only measure than can minimise further data loss. Increasing the compression factor or doubling the SSMM's size in this situation will only postpone the point when the memory becomes full and data loss starts to occur.
- The green curve depicts the total data loss as a function of SSMM size for an assumed fixed, non-optimised ground station visibility of 8 h/day. It can be seen that for all considered SSMM sizes the data loss remains at an unacceptably high level of $> 20\%$. This, at first seems surprising. However, it can be understood by realising that, assuming a compression factor 3, a mean total data rate of 5.7743 Mbit/s/3 $\approx$ 1.92 Mbit/s (Table 2) *exceeds* a sustained downlink rate of 8/24 $\times$ 5 Mbit/s $\approx$ 1.67 Mbit/s by about 15%.

5. CONCLUSIONS

Over the nominal 5 year mission duration the Gaia payload will generate a total uncompressed data volume of roughly 113.8 TB composed of $\approx$ 24 TB from each Astro unit, $\approx$ 37.6 TB from RVS, and 28.3 TB stemming from MBP. The average data rates exhibit the same ratio of 1:1:1.6:1.2. Absolute values are in the range of 1.2 Mbit/s (each Astro), 1.9 Mbit/s (RVS), and 1.4 Mbit/s (MBP) with very large standard deviations. Those mean values are significantly exceeded when regions of extreme stellar density in the galactic plane are scanned. Scans through the galactic plane at steep angles will nominally take place every 3h for every instrument and thus result in frequent short-term peaks of the corresponding data rates. At irregular intervals throughout the mission, in-plane scans will occur with spin axis pointings close the galactic poles. The total data rate in these periods will reach several 10 Mbit/s and data loss will be inevitable unless a SSMM with a size of several 1000 Gbit is present. This appears unrealistic assuming a unit of this dimension would pose unsurmountable problems in terms of cost, reliability, and risk. Assuming that a total end-of-mission data loss of the order 5% is acceptable from a scientific viewpoint an SSMM of size 200–400 Gbit will suffice. This, however, necessitates that a lossless compression factor of 4 is attained and that limiting G/GS magnitudes are lowered from their current values 20/20.15 to yield the original 1 billion Gaia-1 total number of observed objects with Gaia-2. The telemetry downlink rate through the MGA is of predominant importance for the acquired total end-of-mission data loss. Having any value larger than the current baseline of 5 Mbit/s will be highly advantageous. Assuming a compression factor of 3, a mere fixed 8 h/day ground station visibility is clearly unacceptable since the mean rate would exceed the sustained rate and lead to an unacceptably large end-of-mission data loss.

REFERENCES

Lammers, U., 2004, Gaia technical report GAIA-UL-008

de Bruijne, J., et al., 2005, ESA SP-576, this volume and references therein.

Lindegren, L., 1998, Gaia technical report SAG_LL_012

ESA, 2000, Concept and Technology Study Report, ESA-SCI(2000)4.

Lindegren, L., 2001, Gaia technical report SAG-LL-35