Methods for constellation design in projects in early phases, based on studies carried out at CNES in recent years

Romain Pinède [(1)], Alain Lamy [(1)], Francesco Margiotta [(2)], Thierry Martin [(1)]

[(1)] Centre National d'Études Spatiales (CNES), 31400 Toulouse, France
Email: romain.pinede@cnes.fr

[(2)] Thales Services Numériques, 31670 Labège, France

Abstract

Some key elements to consider when designing a constellation will be presented, based on the studies carried out at CNES for Low Earth Orbit projects in early phases over the recent years. Designs can become more robust by paying close attention to the mission criteria and requirements, the simulations modelling, the different design methods, the performance trade-offs, and the links with other domains than spaceflight dynamics.

I. INTRODUCTION

A. Motivations

Various kinds of flight dynamics studies are performed in the department in charge of mission analyses for projects in early phases at CNES, the French space agency: orbit and manoeuvre design, geometrical computations, etc. Some of the projects involve the design of constellations of satellites, which expand the range of available performance and possibilities: with a single satellite design, ambitious mission objectives that are unattainable with a single spacecraft can be achieved. One limitation to the use of constellations is the overall cost of the mission. As a result, designing a constellation with just the required number of satellites is critical to the feasibility of such projects.

This paper will not cover the whole process of a constellation design. Its objective is to explain some key elements to consider when designing a Low Earth Orbit (LEO) constellation, illustrating each step with examples based on several studies carried out at CNES over the recent years, which cover diverse fields of interest: meteorology, ground observation, augmentation of navigation services, electromagnetic monitoring, internet of things, etc.

The paper starts by exposing in section II the criteria and requirements that the mission must satisfy, continue with some considerations on the modelling in section III, present different design methods in section IV. It will then show some performance trade-offs in section V, and finally in section VI some results linked to other domains than spaceflight dynamics.

B. Tools

CNES studies rely on adaptable and extensively validated software developed by the mission analysis team. The constellation simulations are performed by the VISIR tool, developed in Scilab and Java. It notably relies on the Scilab toolbox Celestlab, freely available at [1] or [2].

C. Conventions

We will use the following abbreviations or symbols in the paper:

a	Semi-major axis
e	Eccentricity
i	Inclination
Ω or $raan$	Right ascension of the ascending node
α	Argument of latitude

II. ESTABLISHING REQUIREMENTS

A. Criteria Types

The mission requirements need to be translated into a combination of performance criteria for the mission analysis. These criteria can be of very different types. Some of these types are addressed in the paper:

- Global or partial coverage of the Earth's surface.
- Revisit time, defined from any location on the Earth's surface, as the duration between the end of an Earth/satellite visibility and the next one.
- Number of satellites in simultaneous visibility of any location on the Earth's surface.
- Data latency, defined from any location on the Earth's surface, as the duration between an Earth/satellite visibility and the next station network/satellite visibility. This last visibility can either be with the same satellite or with another one linked through Inter-Satellite Links (ISL).
- Access time, defined from any location on the Earth's surface and at any time, as the duration between the current time and the next Earth/satellite visibility.

Other types can be used:

- Access latency, defined from any location on the Earth's surface, as the duration between the previous station network/satellite visibility and an Earth/satellite visibility.

- Mission latency, defined from any location on the Earth's surface, as the duration between two successive station network/satellite visibilities that surround an Earth/satellite visibility.
- Global response time, defined from any location on the Earth's surface and at any time, as the delay between the current time and the next station network/satellite visibility that follows an Earth/satellite visibility.

Note that to compute the Earth/satellite visibility passes, we partition the Earth ground into a mesh grid which contains several mesh cells. More details are given in section III.

B. *Statistical quantities to consider*

The criteria can be specified in terms of worst-case values, mean values, median values, or other percentiles such as "90% of the cases", as illustrated on Fig.1.

The mean values are useful to compare constellations between each other, because the values are generally smooth. However, they rarely are adequate to satisfy the mission constraints.

The worst-case values (maximum or minimum, depending on the criterion) are generally the first studied statistics. Unfortunately, they often lead to highly populated constellations.

Compromises are then needed between the constellation performance and its cost, and are mathematically perform through the use of intermediate percentiles (with values at 90% or 95% for example). However, there are several subtleties with this approach. In particular, it is important to fully understand what the percentile refers to. Let us illustrate by a case where we are interested in the revisit time, and consider the non-visibility durations for a specific mesh cell on the ground. The 90%-percentile value (written L from now on) means that 90% of the non-visibility durations are below L. It does not mean that L is worth 90% of the maximum non-visibility duration, nor that the non-visibility durations are below L during 90% of the simulation time. The statistics only applies to the sample of the non-visibility quantities, which means that it is possible to have a rare non-visibility hole but particularly long. By the way, such large durations are observed at the junction of the ascending and descending orbit planes in Walker Star constellations on high latitudes, as represented of the top right of Fig.1. In this example case, should we prefer a compromise on the maximum non-visibility durations instead of the number of non-visibility holes, we could use the access time as criterion instead of the revisit time; however, it brings other subtleties as the visibility passes are then counted in the statistics.

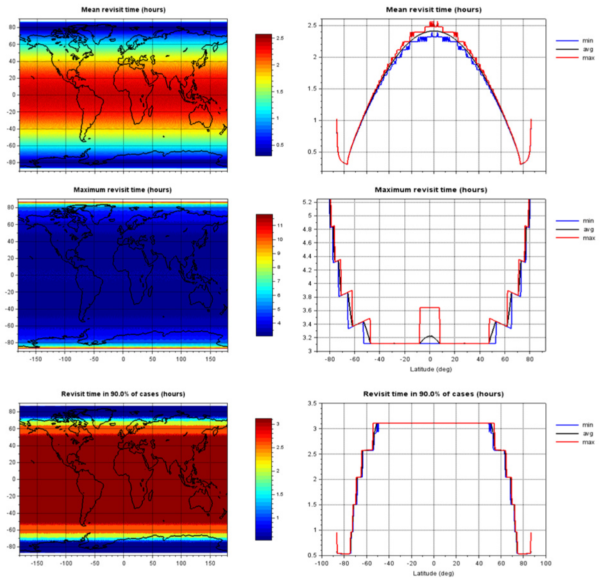

Figure 1 - Criteria representation example: revisit time. Walker Star constellation with 14 satellites, 4 planes, phasing factor 1, altitude 630 km, lateral field of view of 40°.

C. *Criterion representation*

For some criteria (such as revisit time), the main variations are generally a function of latitude and results are hardly dependent on longitude. Two-dimensional plots are redundant and graphs which are a function of latitude (see the right column of Fig.1) are easier to interpret. For other criteria (such as data latency), results depend on the station network and two-dimensional plots are unavoidable.

Surprises occasionally happen. Fig.2 illustrates some unexpected variations that may be observed when studying the mean revisit time.

- Inconsistencies in the modelling (mesh grid sizes, simulation step size, and so on) may produce unwanted signals. Small non-visibility holes appear between the satellites locations at each simulation step and disrupt the final statistical distribution.
- The studied simulation time is not long enough to produce a smooth output. The ground tracks of the orbits are still visible on the final graph, and introduce longitude-dependent results.
- The geometry coincidentally makes the satellites from different planes get very close to each other over the latitude 60°. When it happens, the non-visibility duration is twice the usual duration at these locations. Consequently, we obtain a very noticeable augmentation on the mean revisit time around this latitude.

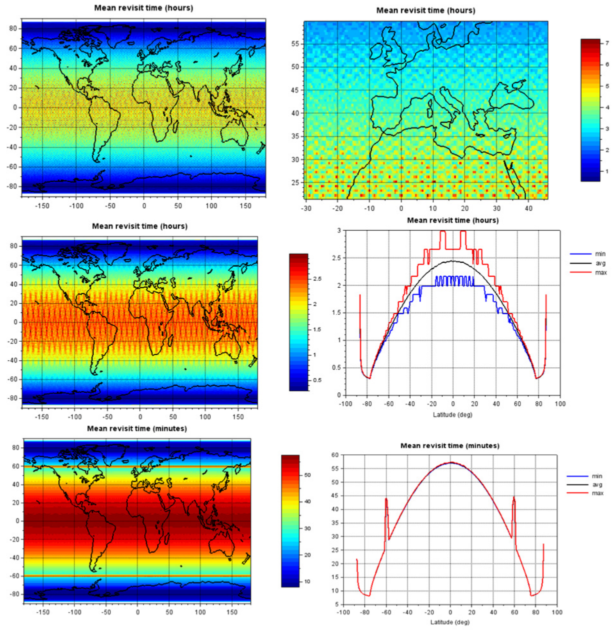

Figure 2 - Sensitivity on mean revisit time.
Top: same case as Fig.1, with a poor choice of settings (longitudinal field of view is too small compared to the simulation step).
Middle: same case as Fig.1 with a new altitude at 665 km on a three-day-long cycle duration repeat orbit.
Bottom: Walker Delta constellation with 25 satellites, 5 planes, phasing factor 2, altitude 767 km, lateral field of view of 40°.

III. MODELLING

For projects in early phases, precise numerical propagation of the satellite orbits is generally not needed. Simpler analytical models offer a sufficient precision and a much faster computation time. However, numerical discretization is by all means still needed. The Earth is partitioned into a mesh grid which contains several mesh cells; the simulation time period is discretized in small time steps. Combining these steps, the Earth/satellite and station network/satellite visibilities can be known from any location at any time.

A. Iso-surface mesh grid

Instead of an intuitive iso-longitude mesh definition, where mesh cell sizes are of a fixed longitude and latitude range, one can save a subsequent number of meshes by using an iso-surface mesh grid as seen on the right of Fig.3. The definition is a bit more complex mathematically, but for the same mesh cell sizes at the equator, it allows to drastically reduce the number of mesh cells at high latitudes for an overall gain of 36% (see Tab.1).

Table 1 - mesh cells amount

Number of cells in latitude	25 (as in Fig.3)	60	90	180
Iso-longitude	**1250**	**7200**	**16200**	**64800**
Iso-surface	**792**	**4584**	**10312**	**41252**

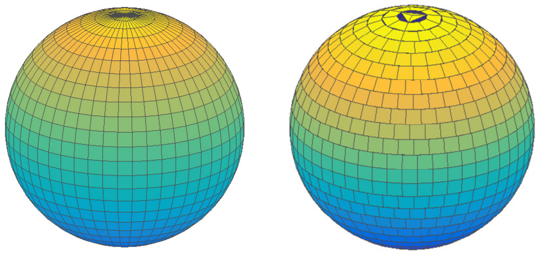

Figure 3 - Mesh grid possible definitions.
Iso-longitude (left) or iso-surface (right).

B. Mesh cell visibility

A mesh cell/satellite visibility at a specific time can be defined with three possible rules, given here from the most restrictive one to the least restrictive one:

- The satellite sees every corner of the cell.
- The satellite sees the center of the cell.
- The satellite sees at least one corner of the cell.

Depending on the size of the cells, this setting may have a strong impact. We generally choose the second rule (center of the cell), because of its simpler approach and its sensitivity to small variations of the visibilities even with large mesh cell sizes.

IV. CONSTELLATION DESIGN METHODS

The mission analysis outputs for a constellation design include the orbits to consider and the constellation characteristics: number of satellites, number of orbital planes, position of the satellites in each plane, position of each plane. There are numerous methods to obtain those.

A. Orbits

Repeat orbits are orbits for which ground tracks replicate after the so-called repeat cycle duration. Their existence and design can be computed by linking the orbital period to the Earth rotation period, and can even encompass the drift due to J2 (see [3]). Some of them are shown on Fig.4. They are almost systematically used by our mission analysis team. It means that when a project provides a specific altitude to work on, a close repeat orbit is often found to replace it. The reason is that the use of a repeat orbit guarantees the stability and representability of the results: when simulating a constellation during the repeat cycle duration, all possible ground tracks combinations are encountered. In the general case longer simulation times would be necessary to reach the results' stability, which extends the computation time and adds the needs of a preliminary analysis to find the adequate simulation time.

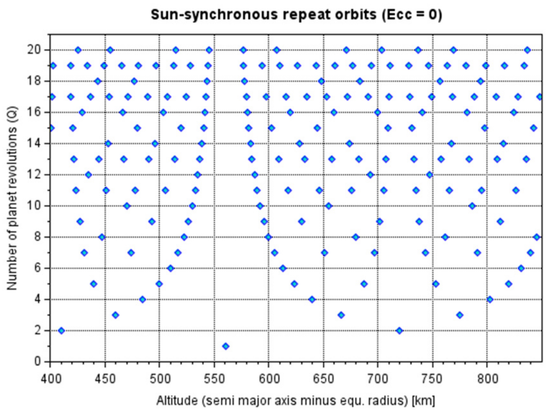

Figure 4 - Repeat cycles at low altitudes (Demo from Celestlab [2])

B. Constellation design types

To limit the number of parameters, simple designs are used.

Homogeneous constellations (circular orbits with the same semi-major axes and same inclinations) are considered most of the time. They provide stable results at an orbital period scale, as different semi-major axes would lead to different orbital periods (through Kepler's third law). They enable to lower the station keeping budget by reducing differential perturbations effects between satellites [4], as the long-term effects of J2 on the orbit are then the same for all satellites (see [3]).

Among homogeneous constellations, Walker constellations [5] are preferred. They enable to limit the number of free parameters in the design. The only degrees of freedom left with the classical Walker Delta constellations are: the number of satellites, the number of orbital planes, the phasing factor, respectively written T, P, and F. These constellations are homogeneously distributed in space. Walker Star constellations extend the definition of Walker Delta. The orbital planes are not defined over 360° like before, but on a hemisphere, close to 180°.

The distribution angle of 180° is not always the adequate choice when using orbits with a non-polar inclination, such as Sun-Synchronous Orbits for which the inclination is close to 100° ($\pm 4°$). The junction between ascending and descending planes is imperfect, visibility holes appear in the coverage at high latitudes, and a higher distribution angle value such as 200° or 220° is sometimes preferable. The distribution angle considered in our simulations is generally found by trial and error (see the iterative approach in next section). However, some systematic studies of the optimal distribution angles also indicate that values lower than 180° can occasionally be suitable (for very small constellations) as seen in Fig.5.

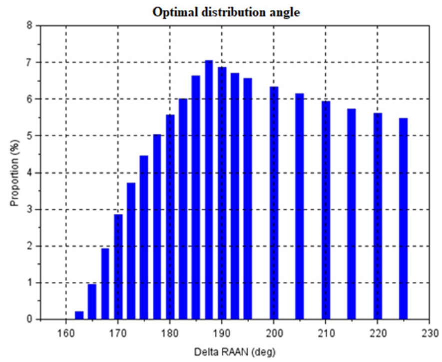

Figure 5 - Optimal distribution angle found when studying all possible small Walker Star constellations with 9 or less orbital planes, 9 or less satellites per plane, near altitude 623 km and inclination 80°.

As far as the phasing factor F is concerned, systematic studies for small constellations have not demonstrated a strict rule in favour of specific F values (see Fig.6). However, we determined that starting with a $F = 0$ or $F = 1$ value is a decent choice when starting to check a constellation performance, because it often happens to be an optimal value.

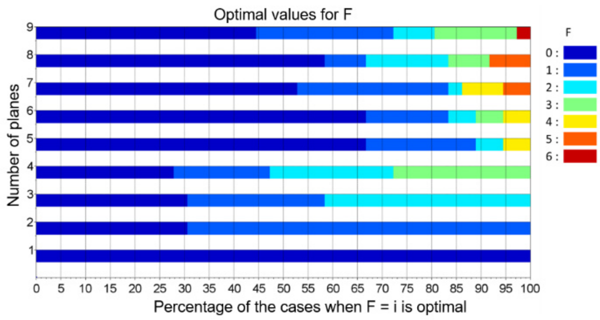

Figure 6 - Optimal phasing factor distribution when studying all possible small Walker Star constellations with 9 or less orbital planes, 9 or less satellites per plane, near altitude 623 km and inclination 80°.

Walker Star constellations are not spatially homogeneous. We have discovered that it is possible to outperform the revisit time performance of a Walker Star constellation, by leaving the Walker formalism and slightly modifying the $raan$ and α of one satellite in the constellation. Such a study has been realized for example for a constellation with 6 satellites and 3 planes, near altitude 687 km and with Sun-Synchronous Orbits. However, the benefits are very limited (a few minutes of maximum revisit time, to be compared to an approximate performance of 2.3 hours in the example) and do not counterbalance the advantages of the design simplicity.

C. Constellation design approaches

To find out the optimal number of satellite and/or orbital planes, various approaches can be taken:

The geometric approach relies on the orbit altitude and the field of view, and estimates how many orbital plans and satellites per plane are needed to comply with the mission objectives. It generally works with criteria such as global and permanent coverage. However, the results with simple hypotheses are quite approximate; this approach is mainly used to initialize the iterative approach.

Reference results stored in tables or graphs (an example is presented of Fig.7) can be methodically produced beforehand for given problems. When projects arise with constraints close to the ones considered for the reference results, the mission analysis team can then:

- Instantaneously provide some approximate numbers (number of satellites or orbital planes).
- Appreciate the sensitivity of the results, to know how much performance is changed by adding or removing a few satellites or one orbital plane.
- Use these results as a basis to initiate the iterative approach.

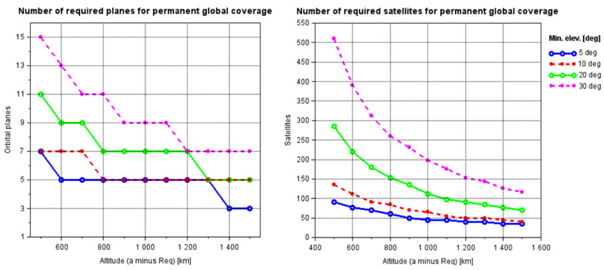

Figure 7 - Reference results to reach a permanent global coverage, with Walker Delta constellations (limited to odd numbers of plans), at different minimum elevations.
Left / Right (respectively): optimized number of planes / satellites. The two graphs do not always refer to the same constellations (see section V.A).

A more general way is the iterative approach: a preliminary constellation is simulated and its parameters are adjusted based on the resulting performance. Tools need to be as fast as possible (a lot of constellations may be tested) and adaptable to enable both automatic scanning and manual tuning, depending on the mission criteria.

V. PERFORMANCE ANALYSIS

A constellation design output is hardly a single constellation with characteristics perfectly matching the mission constraints. For projects in early phases, mission constraints are rather imprecise. The constellations may be drastically reduced with a small constraint relief. In contrast, some constraints will have no impact on the constellation performance whatsoever. Results should include several concepts to help us understand this sensitivity. The objective is to guarantee the best performance for the minimum completion cost.

A. *Number of satellites or number of orbital planes*

A choice needs to be made between the optimization of the number of satellite or the optimization of the number of orbital planes. In fact, these objectives may be contradictory as illustrated on Tab.2.

Table 2 – Trade-off between two Walker Star constellations at altitude 630 km. The second constellation has 6 fewer satellites, but 2 more planes. Its maximum revisit time is slightly better, but its mean revisit time is a bit worse.

Number of satellites	**30**	**24**
Number of planes	**6**	**8**
Phasing factor	4	3
Distribution angle	185°	185°
Maximum revisit time at equator	1.75 h	1.5 h
Maximum revisit time at latitude 60°	2 h	1.75 h
Mean revisit time at equator	0.8 h	1 h
Mean revisit time at latitude 60°	0.4 h	0.5 h

What mainly drives the completion cost is typically the number of orbital planes, directly linked to the number of required orbital launches. In many cases, one launch will only populate one orbital plane as transfers between planes would take a heavy toll either on the satellites fuel or on the deployment schedule time. However, it is difficult to assert a general rule. Sometimes it can be preferred to add more satellites on every plane rather than adding a new orbital plane. In other cases, a new plane is mandatory to improve the performance, especially for small constellations.

B. *Criteria to dismiss*

Sometimes, the sensitivity of the constellation size to the mission constraints and criteria is very low. In such situations, one needs to separate the constellation design problem from these constraints. This kind of event happens for example with a constraint on the worst data latency value as shown in Fig.8. At some point, no matter how many satellites or orbital planes are added to the constellation, the performance will be capped and can only be improved by a change in the distribution of the station network.

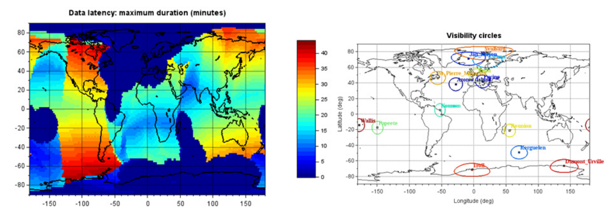

Figure 8 - Maximum data latency (left), for a Walker Star constellation with 1400 satellites and 25 planes (the requirement was to see every point of the Earth with a sufficient elevation with 4 different satellites at altitude 650 km), with a European station network (right).

C. *Heterogeneous constellations*

An idea frequently arises in discussions about reducing the size of a constellation: to try to use heterogeneous constellations. Two main reasons explain such attempts.

- When deciding between two contradictory tendencies that both present advantages. For example, on one hand Sun-Synchronous Orbits (SSO) lead to easier satellite design because of the stability of the lighting characteristics on the orbit over the year, but on the

other hand missions interested in populated area and subsequent limited latitude range (below 60°) would be interested in orbits with inclinations limited to 60°, as the covered latitudes are linked to the inclination value. A heterogeneous constellation concept can be evoked, with both SSO and 60°-inclination orbits.
- To take advantage of pre-existing satellites or constellations with different orbital characteristics. For example, a constellation design is needed to fulfil a revisit time objective, and some pre-existing satellites could be associated to the mission, even if they are at an altitude which is unattainable for the next satellites. We could hope that their use would be advantageous, even if different altitudes mean different orbital periods.

In spite of these hopes, we never found a situation where such an idea would turn to be profitable. What happens is that the heterogeneous constellations indeed improve the best situations, but also preserve the worst-cases. As we usually define worst-case criteria or close-to-worst-case criteria for the performance, the benefits of this idea are usually insufficient.

VI. ITERATIONS WITH OTHER DOMAINS

Before validating the constellation design, some iterations always need to be performed with other domains than spaceflight dynamics to challenge the design.

A. Launcher capacity

This is a very simple but very important example of such a useful iteration. Assuming that one launch will only populate one orbital plane, if the launcher has a capacity of C satellites per launch, the number of satellites to launch per orbital plane (including the possible redundancies) should be as close as possible to an integer number multiple of C. As an example it would be regretful to have a constellation design with 24 satellites per orbital plane, for a launcher with a 22-satellite capacity. Two launches per orbital plane would be necessary, the latter with a very small occupancy.

B. Reliability and redundancies

When the constellation design is optimal, a satellite withdrawal will lower the performance and the mission requirements will not be fulfilled anymore. Unfortunately, many malfunctions can lead to such withdrawal, so this eventuality needs to be taken into account. Redundancies are necessary, both inside the platforms (which is not the concern of the spaceflight dynamics teams) and by adding more satellites to the initial constellation design. Two approaches are possible:

First we may consider a cold redundancy, where redundant satellites are launched into orbit but do not contribute to the constellation performance until they need to replace another satellite. When a satellite malfunction occurs, the mission performance is not ensured until a redundant satellite has taken place of the defective one. The main drawbacks of this approach are that it can be seen as a waste of resources (the redundant satellites are operational and age like the others but do not really fully contribute to the constellation performance), and the replacement can take some time which will be deduced from the mission uptime. These redundant satellites should be injected to the same orbital plane. Otherwise, the replacement would be too expensive (in time, or in fuel). If the redundant satellites are placed on a different altitude than the mission altitude, the orbital periods will differ. The J2 perturbation will be endured differently so the orbital plane will drift and will have to be maintained by manoeuvres. If all satellites are placed on the same plane and altitude, an adequate geometry needs to be considered. If all satellites are evenly distributed, the operational satellites geometry will not match the initial constellation design and performance needs to be assessed again. Another solution is to keep the initial operational distribution and to add redundant satellites where it is possible, raising the need to station keeping studies.

Secondly, we may consider a warm redundancy, where more working satellites are launched than deemed necessary by the initial constellation design. Some additional performance analyses are then necessary to confirm the suitability of the degraded heterogeneous constellations that appear (see Fig.9). Three choices are possible when a satellite malfunction occurs:
- Adjust the whole orbital plane geometry to obtain a homogeneous angle of separation between working satellites, which requires manoeuvres for all satellites in the plane. It is as more expensive as there is a lot of satellites per plane.
- Keep the orbital plane geometry as it is. Note that even if the plane satellite population has been augmented with the warm redundancy, the worst separation angle on the plane after a withdrawal will be superior to the initial design separation angle. The resulting constellation is heterogeneous.
- A compromise could be reached between both previous options: to only adjust the position of a few satellites around the malfunction position. It would reduce the worst-case separation angle, but also slightly increase a few nominal separation angles around the missing spacecraft.

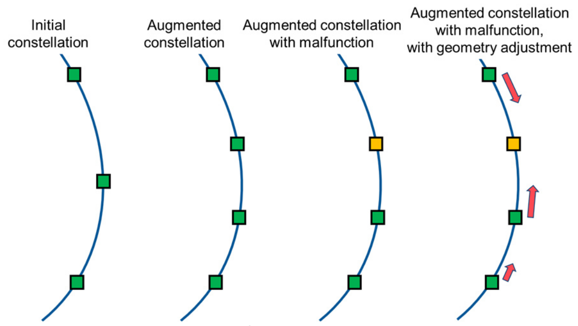

Figure 9 - Illustration for the warm redundancy strategy.
SP is the nominal number of satellites per plane.
R is the number of redundant satellites.
First: initial constellation with a $360°/SP$ separation.
Second: Augmented constellation with a $360°/(SP+R)$ separation.
Third: Malfunction with a "steady geometry", $2 \times 360°/(SP+R)$.
Four: Malfunction with compromise, the highest separation angle is lower than in the third case, but the other one is a bit higher.

C. *Local times and satellite design*

For the CMIM study [6], Walker Star constellations (4 planes, 2 satellites per orbit) with Sun-Synchronous Orbits have been considered. Once the design has been performed, the Walker parameters (T, P, F, and the distribution angle) are to remain unchanged so that the constellation fulfils the required mission performance. With Sun-Synchronous Orbits, the mean local times at the ascending node (MLTAN) will stay rather constant for each orbital plane. Still, a degree of freedom remains. All local times can be shifted by the same value without changing the performance. In our case it has been possible to tune this shift in order to improve the power efficiency and then facilitate the satellite design, as presented below.

Solar panels can be fixed or rotating, the latter enabling to achieve a better illumination over an orbit. A simple satellite design requires that solar panels need to be installed on the platform in the same way for every orbital plane of the constellation. However, it is then difficult to reach good lighting conditions for every plane. Even if it is possible to "tilt" a little the way the solar panels are attached to the platform structure, the accessible variation range is only of ~20° because of the mechanical constraints involved.

The lighting conditions per orbit are derived from the angle between the normal to the solar panels and the Sun direction. The suitable quantity to study this is the angle β between the orbital plane and the Sun direction. Over the year, for a specific local time at the ascending node, β fluctuates because of the Sun apparent motion as seen on Fig.10.

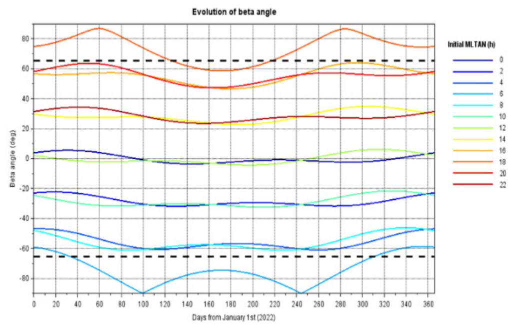

Figure 10 – Evolution of β angle depending on the period of year and the mean local time at the ascending node (Demo from Celestlab [2])

The constellation has then been shifted in $raan$ (equivalently, in local times) to minimize the overall β variations reached during a year for the constellation (see Fig.11), so that they become compatible of the accessible "tilt" range. In the end it meant putting the local times of every plane as far away as possible as the 6h or 18h values, because they are the local times with the most β variations as seen in Fig.10. Afterwards, each plane "tilt" has been adjusted to maximize the lighting ratio. The lighting ratio with rotating solar panels finally reached at least 62% for each plane (taking the eclipses into account) as understandable on Fig.12.

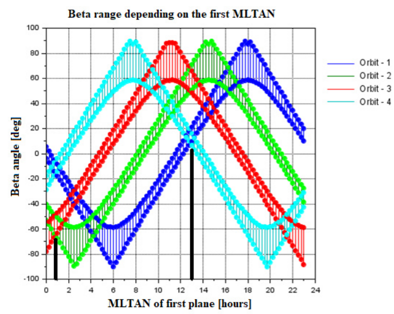

Figure 11 - Beta range of each plane over a year. To minimize the overall β variation (maximum value minus minimum values) leads to choose the first MLTAN at 1h or 13h.

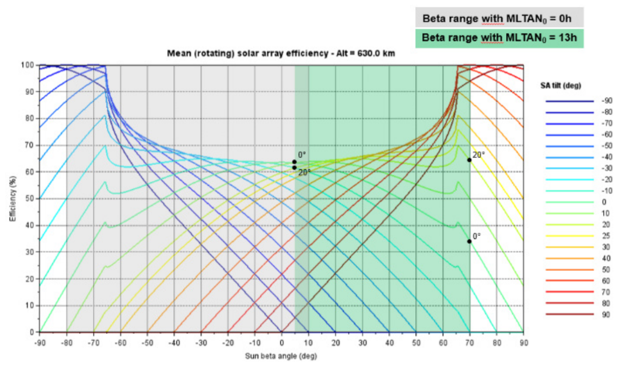

Figure 12 - Solar array efficiency with rotating panels, depending on β and the solar array tilt (Demo from Celestlab [2]). In the original beta range (grey area), no tilt in [-20;20]° would allow an efficiency higher than 50%. With the new beta range (green area), it is possible to stay over 62% efficiency when navigating in tilt between 0° and 20°.

VII. CONCLUSION

Some key elements to consider when designing a constellation have been shown and illustrated in this paper. Still, the mission analysis of a constellation-related project in early phase implies many more elements that have not been evoked here. As more and more constellations appear in today's space program schedules, each one with different mission constraints or technical advances, it is not possible to be exhaustive; new possibilities and new design methods will continue to emerge.

VIII. REFERENCES

[1] https://connectbycnes.fr/celestlab

[2] https://atoms.scilab.org/toolboxes/celestlab/

[3] CNES, J.-P. Carrou, "Spaceflight Dynamics", Editions Cépaduès, ISBN 978-2854283785, 1995.

[4] A. Lamy, S. Pascal, "Station keeping strategies for constellations of satellites", *Advances in the Astronautical Sciences, Vol. 84, American Astronautical Society*, ISSN 0065-3438, p. 819-833, 1993.

[5] J. G. Walker, "Satellite constellations", *Journal of the British Interplanetary Society*, vol. 37, pp. 559-571, 1984.

[6] V. Pascal, J. Michaud, A. Deschamps, L. Le Barbier, L. Hermozo, F. Bernard, T. Martin, R. Pinède, N. Fourrié, P. Chambon, J. Vidot, T. Carrel-Billard, O. Audouin, N. C. Cabrera Gutierrez, "CMIM, a constellation of mini sounders for meteorology", *Eumetsat Meteorological Satellite Conference*, n°155, 2023 September 11th.