Exploiting Standardized Metadata for GNSS SDR Remote Processing: a Case Study

A. Favenza, Istituto Superiore Mario Boella
N. Linty, F. Dovis, Politecnico di Torino

BIOGRAPHIES

Alfredo Favenza is a senior researcher and project manager at Istituto Superiore Mario Boella in Turin. In 2007, he received his Master of Science in Computer Science and started his research activity in ISMB in the area of Navigation technologies developing strong expertise on GNSS Software Defined Radio, and Augmentation Systems (EGNOS/EDAS). Nowadays, Alfredo is a researcher of Mobile Solutions Research Area of ISMB interested in cross-technologies domains where Location-Based services, Big Data and Navigation technologies can be combined to generate added-value innovation.

Nicola Linty is a post-doc researcher at the Department of Electronics and Telecommunications of Politecnico di Torino (Italy). In 2010 he received his M.S. in communication engineering and in 2015 his PhD in electronics and communication engineering, both at Politecnico di Torino. His research interests cover the field of signal processing and simulation, applied to telecommunication and satellite positioning. His work is mainly focused on the design and development of innovative architectures, techniques and algorithms for GPS and Galileo professional and mass-market receivers and on ionospheric scintillation monitoring.

Fabio Dovis is an associate professor at the Department of Electronics and Telecommunications of Politecnico di Torino as a member of the Navigation Signal Analysis and Simulation (NavSAS) group. His research interests cover the design of GPS and Galileo receivers and advanced signal processing for interference and multipath detection and mitigation. He has a relevant experience in European projects in satellite navigation as well as cooperation with industries and research centers.

ABSTRACT

The proliferation of ad-hoc, highly customized, GNSS Software-defined radio (SDR) receivers and raw data collection systems is opening new opportunities for research focused on the analysis of new approaches in GNSS data processing. Especially when looking at the post-processing of GNSS signals for monitoring purposes, the availability of big-sized raw digital samples allows for the application of different algorithms and processing strategies on the same data-set, thus enabling valuable comparisons and proper tuning of the algorithms themselves.

A side effect is that the huge quantity of GNSS data collections, grabbed in disparate scenarios and environmental conditions and distributed all around the world, is making hard to share this data between different research labs, thus limiting the possibility of conducting extensive analysis of relevant ionospheric phenomena affecting the quality to the GNSS signal.

In this paper, the authors propose a novel methodology to deal with the processing and the management of this huge amount of data, and guarantee interoperability of different data collection and data processing systems leveraging on the adoption of a new emerging and open standard for GNSS metadata which is now under standardization by the ION GNSS SDR Metadata Standardization Working Group.

INTRODUCTION

These days, the growing number of Global Navigation Satellite System (GNSS) Software-defined radio (SDR) receivers and customized raw data collection systems is producing a huge, unprecedented mass of GNSS data collections, grabbed in disparate scenarios and environmental conditions and distributed all around the world [1]. These data collections are often characterized by what in the Big Data domain is called the “3 Vs Model” [2, 3]: Volume (Gigabyte per data collection), Velocity (number of data collection systems and products around the world) and Variety (different formats).

The creation of a framework to guarantee the access to this huge amount of experimental data is thus becoming of crucial relevance for the scientific community, in equal measure to the processing of the data itself. Furthermore, GNSS-based monitoring is often implemented in remote areas, where resources are limited and installation or maintenance and replacement of GNSS receivers may be critical. For instance, providing researchers with the opportunity to use their own SDR, right in their laboratory, to process and analyse GNSS data grabbed in a far country, with not replicable environmental conditions, may offer valuable opportunities for the study of environmental phenomena affecting the GNSS signal. At the same time, the laboratories would be able to release and make accessible their data to the world helping other researchers in GNSS receiver testing activities.

Furthermore, this multitude of different SDR receivers and data collection systems is pushing the demand for a standard framework to standardize the access to these data. In order to fulfil this need, in 2014 [4, 5] a working group was established by the Council of the Institute of Navigation in order to define a free and open standard for the exchange of GNSS software radio data in order to promote and guarantee the interoperability between different GNSS SDR receivers, data collection systems and SDR processors.

In addition, this approach is paving the way to find a solution to the relevant issue of the complexity of the data transfer, which may be impossible due to the limited bandwidth available in remote sites. As a direct consequence, this approach also enables saving in time and cost operation of moving Gigabytes of GNSS data collections from a remote *Data Collection Site* to the *Research Lab*.

This main objectives of this paper are:

1. to describe the process to prepare and asses a GNSS SDR receiver to be compliant with the ION GNSS SDR Metadata Standard under definition [5];
2. to provide an example of real GNSS data ingestion and processing based on the use of the metadata file and related Application Programming Interfaces (APIs) [6, 7] to configure the SDR;
3. to pave the way to face the big data issues, proposing an approach based on paradigm shift from "moving data" to "moving software", in the particular case of isolated and hardly accessible installations.

In this paper, the authors provide a detailed description of a case-study as early adopters of this common standard, for the remote processing of SDR data collected in a remote site in the Antarctic region, where the connection bandwidth and the logistic issues do not allow for an easy transfer of the SDR data collected. The paper assesses the advantages in terms of bandwidth saving and implementation, thus proving the effectiveness of having a common baseline for the data cataloguing.

I – THE DATA PROCESSING CHAIN

The GNSS receiver architecture commonly considered when working with SDR GNSS receivers consists of a GNSS antenna, a Radio Frequency (RF) front-end, a data grabber, a communication link and a data processing system. The block scheme of this overall GNSS data processing framework is reported in Figure 1. Each blocks is described in more details in the following sections.

A – The front-end

The RF front-end performs signal conditioning. It is normally connected in input to an antenna and it converts the RF analog signal in a sequence of digital samples, either at Intermediate Frequency (IF) or in Base Band (BB).

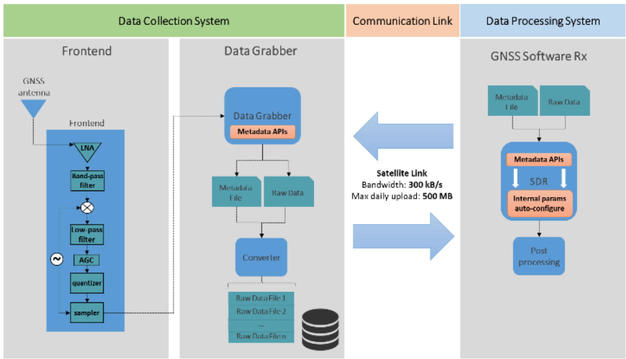


Figure 1. The overall data processing framework block scheme.

A detailed block scheme of a typical frond-end for GNSS data grabbing is reported in the left part of Figure 1. A Low Noise Amplifier (LNA) can be optionally used as a first step to amplify the input Signal In Space (SIS), in some cases too weak to be processed directly. The signal is then band-pass filtered, to reduce the out-of-band components and down-converted to IF or to BB, to allow SDR processing. A further filtered stage is needed to remove the high frequency term. In the case in which multi/bit resolution is employed, an Automatic Gain Control (AGC) should be used in order to adjust the signal dynamic. The final step is the digitization of the analog signal: an Analog-to-Digital Converter (ADC) performs sampling and quantization. The signal can then be stored in a memory for a successive post-processing or directly elaborated by the SDR receiver.

B – Raw GNSS data

The concept of GNSS raw data is of paramount importance when dealing with SDR receivers. GNSS raw data, or IF data, are the output of the RF front-end. They consist of a sequence of digital samples, at a certain sampling frequency f_s and represented on a certain number of bits (from 1 bit up to 2 bytes). Raw data can be stored as binary files in memories, transferred, shared, and then used for post processing.

It is important not to confuse raw data with GNSS observables, such as the IQ correlation outputs, the Carrier-to-Noise power density ratio (C/N_0), the pseudo-ranges and the Position, Velocity and Time (PVT) solution, which correspond to the output of most GNSS commercial receivers.

C – The GNSS software receiver

SDR refers to an ensemble of hardware and software technologies enabling reconfigurable radio communication architectures [1]. All the components traditionally developed in dedicated hardware, such as correlator, are realized in software, on programmable platforms, such as high-performance General Purpose Processors (GPPs). At this point, SDR receivers are very common tools in satellite navigation research and development areas, not only because of their lower cost and faster implementation, but mainly because of the new and still un-explored opportunities offered by the availability of digital signals. Among the wide variety of SDR solutions, employing different combination of processing platforms, in this paper we refer to the fully software architecture, in which all the GNSS receiver, except the RF front-end, is completely implemented in software, on a Personal Computer (PC) [1, 8].

D – The SDR Metadata Standard

This research leverages on the work conducted by the ION GNSS SDR Metadata Group, which is proposing a standard for an Xml-based GNSS SDR metadata file format to describe the content of the data collections to help the automated interpretation process by a processing software. Based on this standard, a SDR receiver can automatically configure internal parameters (i.e., RF and IF centre frequencies, sample rate, quantization bits, etc.) enabling the playback and post-processing of the collected data. The standard is designed to promote the interoperability of GNSS SDR data collection systems and processors and includes a formal XML Schema Definition (XSD). A fully compliant open source C++ API is also officially supported to promote ease of integration into existing SDR systems. [6]

These APIs provide open source libraries and code which can be easily used to integrate these functionalities at different levels of data processing:

- in any data collection system (e.g. GNSS data grabber) that, embedding the metadata creation libraries into grabber code, can generate, during the data capturing process, also the needed metadata file, accordingly to a predefined and fine customization of the APIs (accordingly to the features of the data collection itself);
- in any GNSS software receiver that, integrating the metadata libraries at the first step of the data processing chain, can automatically configure its internal parameters (i.e., RF and IF centre frequencies, sample rate, quantization bits, etc.) enabling the playback and post-processing of the collected data

All the detail of the integration of the APIs in the data grabber and in the SDR code will be provided in Section III.D, System Setup.

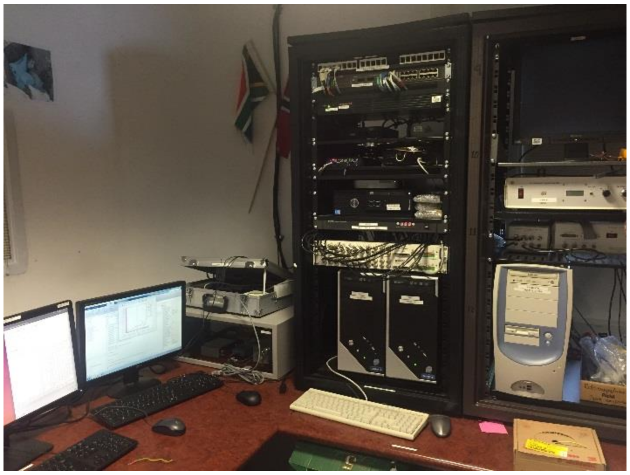
II – CASE STUDY

In this section, a real case study is introduced. The theoretical concepts outlined in Section I. are contextualized in the frame of a real implementation of a data processing framework exploiting metadata.

A – Scenario overview

The Demogrape project [9], funded by the Italian National Antarctic Research Program (Programma Nazionale di Ricerche in Antartide, PNRA) under the coordination of the Italian national institute of geophysics and volcanology (Istituto Nazionale di Geofisica e Vulcanologia, INGV), has the objective to improve the quality of the GNSS position solution in Polar Regions. Ionosphere is indeed the major source of errors in GNSS processing, especially in Equatorial and in Polar regions. In the frame of the project, a demonstrator will assess the corruption induced by ionospheric propagation on GNSS signals and share data and results, by means of a federated cloud infrastructure, thus enabling the development of new scientific and technological applications in Antarctica. By monitoring the amplitude and phase variations of the received GNSS signals in such locations, it is possible to better understand the behavior of ionosphere. Furthermore, a better knowledge of the ionospheric behavior in these singular regions could improve the fidelity of global models, thus extending the advantages of this research to GNSS users worldwide.

a) Location of the GNSS antenna in the EACF station.



b) Installation of the GNSS-based ionosphere monitoring station in the SANAE IV station.

Figure 2. Installation of the DemoGRAPE system in two Antarctica research stations.

While ionosphere monitoring was traditionally a prerogative of commercial hardware Professional Scintillation Monitoring Receivers (ISMR), recent trends consider also SDR technology as a valuable, low cost and flexible alternative [8]. GNSS raw data acquisition systems and software receivers have indeed been successfully used for monitoring ionosphere in Equatorial regions for few years [10, 11]. On the other hand, the installation and employment of a SDR-based receiver in Antarctica represents a novelty; it enables innovative and interesting solutions in ionosphere monitoring, at the expenses of more severe problems in data handling and processing.

Between October 2015 and January 2016, in the frame of the DemoGRAPE project, two GNSS-based ionosphere monitoring stations were installed in Antarctica, respectively in the Brazilian research station EACF (Figure 2a) and in the South African research station SANAE IV (Figure 2b). In parallel to a Septentrio PolaRxS receiver, an acquisition system, able to grab raw GNSS data at high sampling rate and high resolution and to post-process them exploiting a SDR receiver, was installed in both stations [12, 13].

B – The bottleneck of data storage and transfer

Accessibility, space, energy, network and bandwidth resources are indeed extremely limited in environments such as a research station in Antarctica. This is also the case of some installations in Equatorial regions where resources are very limited, and monitoring stations are implemented in remote areas not easily accessible. Reaching Antarctica to install and maintain the station requires months of preparation and logistic issues. In most of the cases, once installed, the set-up is left unattended for several months, during which nobody is present at the station and can operate on the devices. Moreover, particular privileges are required to transfer data through the slow and expensive satellite links. Therefore, a well-planned strategy for data management and sharing is mandatory.

While the typical output data of a Septentrio receiver, such as Rinex files, amounts to about 8 MB per day, GNSS raw data can grow up to about 1 GB per minute [8]. For example, the South African station is equipped with a satellite link, offering a bandwidth of 300 kB/s and a maximum daily upload capacity of 500 MB, shared by all the scientific projects and logistic need of the station. It is straightforward that in the second case no data transmission is possible.

C - Novel metadata-based approaches to raw data post-processing

The exploitation of the metadata file and the flexibility of SDR implementation of the GNSS receivers provide a solution to the issue of the transfer of large amounts of collected data. The fully software implementation offers, in fact, the possibility to move the processing engine, enabling the new paradigm of moving the computer to the data rather then moving the data to the computers. This approach consists in transferring pieces of code on the machine hosting the raw data in the *Data Collection Site*, rather than moving the data itself to the *Research Lab*.

This paradigm has been applied to the Antarctic installation at EACF and SANAE, where the data collection systems have been specifically implemented for scintillation monitoring purposes [14]. In this case, the SDR receiver is configured to have a proper robust architecture for the tracking loops, and optimized to provide as outputs the typical indexes used for scintillation monitoring, such as the amplitude scintillation index (S4) and the phase scintillation index (σ_{Φ}). The presence of additional nuisances such as multipath, in some data collections, requires the tuning of the software receiver, for example changing the spacing of the early-late correlators, or the activation of a signal processing unit to mitigate the multipath effect. In this way more reliable results for the assessment of the intensity of the scintillation event can be obtained.

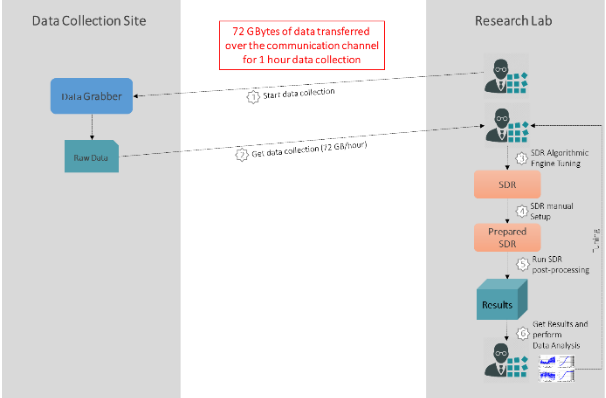


Figure 3. Scenario 1: traditional approach for the post-processing of remote data collections

In order to better clarify the proposed methodology, Figure 3 reports a scheme of the traditional approach: the *Data Collection Site* hosts the data grabber, while the *Research Lab* hosts the SDR receiver. The user runs a command (1) to start the data collection. The acquisition system grabs the raw data, which are then transferred from the *Data Collection Site* to the lab through the network (2). The user can then tune the SDR receiver (3), manually configure it according to the settings of the grabber (4), and finally post-process the data (5) to obtain the results (6). In this particular case, with the parameters summarized in Table 1 about 72 GB of data are transferred for about 1 hour data collection:

$$5 \text{ Msamples/s} \cdot 2 \text{ B/sample} \cdot 2 \cdot 3600 \text{ s/hour} = 72 \text{ GB/hour}$$

Similarly, Figure 4 reports the flowchart of the innovative approach, in which both the grabber and the receiver are hosted by the *Data Collection Site*. In this case, the user in the *Research Lab* runs a command to start the data collection (1); the acquisition system grabs the GNSS data, generating the raw data file and the related metadata file (2). Meanwhile, the user tunes the software receiver (3) and uploads it to the *Data Collection Site*, exploiting the network link (4).

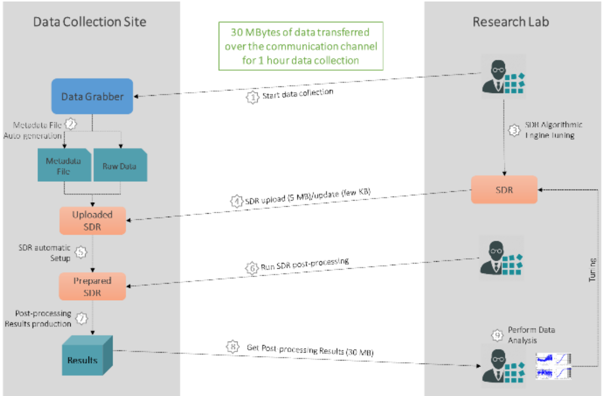


Figure 4. Scenario 2: innovative metadata-based approach for the post-processing of remote data collections – upgrade and reconfiguration of the SDR receiver

Once the SDR has automatically configured itself using the metadata (5), the user can run the post-processing (6) and the SDR receiver starts processing the raw data remotely (7). Finally, the SDR results are transferred to the *Research Lab* via the communication link (8) and the user can analyse them (9). In this case, about 30 MB of data are transferred, up to 2000 times less than in the traditional approach for the processing of 60 minutes of data, as reported in Table 1.

Furthermore, thanks to the modularity feature of an SDR receiver, the complete transfer of the SDR is needed just once. In the case of a second data collection, no SDR upload (4) is required, while in the case of a second post-processing with a different tuning, only the modified software blocks must be transferred.

An example of results achieved exploiting the full metadata-based chain is reported in Figure 5. The figure reports the scintillation statistics (C/N_0, S4 and σ_Φ) related to a one-hour data collection. Raw GPS L1 data have been remotely grabbed on March 7, 2016, in SANAE Antarctica station. The signals have been post-processed with the SDR receiver, exploiting the metadata file. Finally, the results have been transferred through the network and plotted. It has to be noted that moderate phase scintillation activity is present ($\sigma_\Phi > 0.4$ rad). An example of software block upgrade is reported in Section III.B.

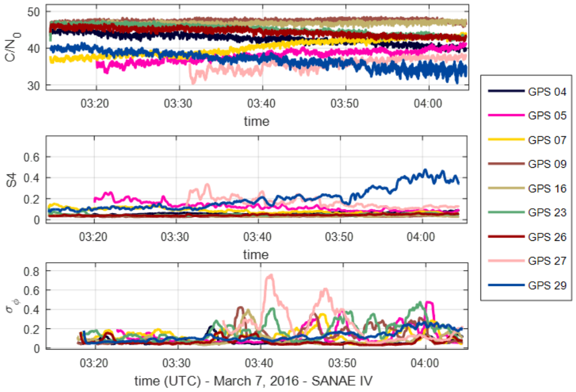


Figure 5. Example of results of remote post-processing exploiting the innovative metadata-based approach

A third approach, derived from the previous one, is depicted in Figure 6. In this case, the user in the *Research Lab*, after the analysis of the data received through the network (9), can decide to switch to another software receiver, denoted SDR-2 (10). Similarly, the same data can be accessed by a second user, in a different *Research Lab*, and with its personal SDR receiver. Also in this case, the overhead of the transfer of raw data is reduced to the upload of the SDR-2, limiting the bandwidth requirements to a few Megabytes. An example of this approach is provided in Section III.A.

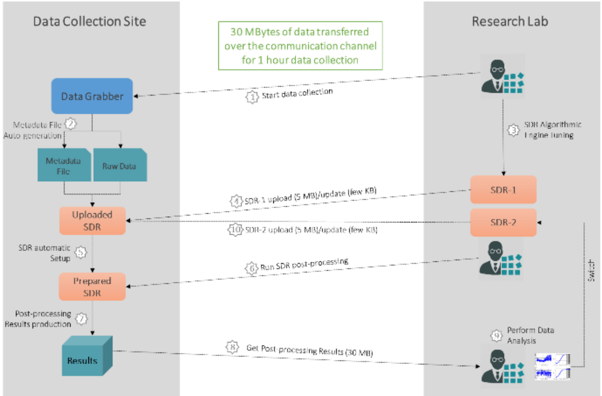


Figure 6. Scenario 3: innovative metadata-based approach for the post-processing of remote data collections – upload of different SDR receivers.

The availability of a meta-data file besides the GNSS raw data represents an added value in both scenarios 2 and 3, helping in automatically re-configuring the SDR without need of human intervention. It is sufficient to design an interface between metadata and software receiver, able to interpret the content of the XML file and to properly initialize the SDR.

Table 1: Size and description of the items transferred through the network in the traditional and in the metadata-based approaches.

	SIZE	DESCRIPTION
Raw data	72 GB/hour	GPS L1 signals, IQ sampling (2 Bytes) at 5 Msamples/s
Post-processing results	30 MB/hour	50 Hz data (I and Q correlators, C/N_0, carrier phase) 1 Hz data (S4, σ_Φ, C/N_0) Rinex file, Scintillation summary
SDR receiver	5 MB	L1 C/A and L2C GPS SDR fully software receiver and ionosphere monitor

An example of metadata file, generated by the GNSS acquisition systems in the remote Antarctica station, used to configure the SDR receiver and compliant with the standard is reported in *Figure 7*. It includes only a stream of data, on E1-L1 frequency. The description of the different fields is out of the scope of this paper; more information can be found in the standard official documentation [4, 5, 6].

```
<?xml version="1.0" encoding="UTF-8"?>
<metadata xmlns="http://www.ion.org/standards/sdrwg/schema/metadata.xsd">
    <lane id="Galileo E1 and GPS L1 Data">
        <bandsrc idband="L1-E1" idsrc="Multiband"/>
        <session id="0">
            <comment format="text">DemoGRAPE ionosphere monitoring station.</comment>
            <toa>1970-01-01T00:00:0.000000000Z</toa>
            <contact>N. Linty</contact>
            <campaign>DemoGRAPE project</campaign>
            <scenario>EACF Antarctica station</scenario>
        </session>
        <system id="NavSAS USRP SDR”/>
        <block>
            <cycles>1</cycles>
            <chunk>
                <sizeword>1</sizeword>
                <countwords>1</countwords>
                <endian>Undefined</endian>
                <padding>None</padding>
                <wordshift>Left</wordshift>
                <lump>
                    <stream id="L1-E1">
                        <ratefactor>1</ratefactor>
                        <quantization>14</quantization>
                        <packedbits>16</packedbits>
                        <alignment>Undefined</alignment>
                        <shift>Undefined</shift>
                        <format>IQ</format>
                        <encoding>INT16</encoding>
                        <band id="L1-E1">
                            <centerfreq format="MHz">1.5754200000000001e+03</centerfreq>
                            <translatedfreq format="MHz">0.0000000000000000e+00</translatedfreq>
                            <delaybias format="sec">0.0000000000000000e+00</delaybias>
                            <bandwidth format="MHz">5.0000000000000000e+00</bandwidth>
                        </band>
                    </stream>
                </lump>
            </chunk>
        </block>
    </lane>
    <system id="NavSAS USRP">
        <comment format="text">USRP with clock from the 4tuNe.</comment>
        <freqbase format="MHz">5.0000000000000000e+00</freqbase>
        <equipment>URSP N210</equipment>
        <types>Processor</types>
        <source id="Multiband">
            <type>UndefinedType</type>
            <polarization>UndefinedPolarization</polarization>
            <idcluster>Antenna</idcluster>
        </source>
        <cluster id="Antenna">
            <vendor>Septentrio</vendor>
            <model>700031</model>
        </cluster>
    </system>
    <file>
        <url>EACON_160614_135052_USRP_L1.bin</url>
        <timestamp>2016-06-14T13:50:52.000000000Z</timestamp>
        <owner>NavSAS/PNRA</owner>
        <lane id="Galileo E1 and GPS L1 Data"/>
    </file>
</metadata>
```

Figure 7. Example of metadata compliant with the ION GNSS SDR Metadata standard

D – System Setup

As previously mentioned, a first version of an open source C++ API which is fully compliant with the standard under definition is officially supported and available to promote ease of integration into existing SDR systems [6].

In this research, we have embedded the provided APIs within the data grabber the software receiver operating at different steps of the processing chain.

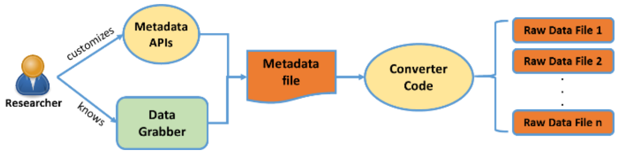


Figure 8. Metadata and data creation scheme.

As shown in Figure 8, the metadata creation and data collection creation code is integrated just at the end of the data collection phase, where the configuration provided by the user is used to automatically generate:

- a **metadata xml-based file** describing all signals/sources present in the data collection;
- a number of **raw data files**, as described in the metadata file on the basis of the different of data sources (e.g. different GNSS signals) in the original data collection. Datasets containing GNSS samples may in fact also contain other information such as data from other sensors and data from RF bands other than GNSS bands.

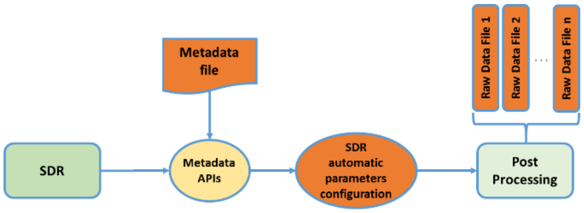


Figure 9. Metadata and data utilization scheme.

On the data post-processing side (Figure 9), the GNSS SDR receiver has been prepared an assessed to be compliant with the ION GNSS SDR Metadata standard. In this particular case, the fully software ionosphere monitoring receiver developed by Politecnico di Torino and installed in Antarctica is considered. The task was relatively easy, thanks to the modularity structure proper of SDR implementations. It was enough to adapt the original initialization function of the receiver, requiring manual configuration of several parameters at each run, with the same values read from the metadata file. The APIs have been integrated just at the beginning of the processing chain performing the following operations:

- reading of the metadata xml-based file and automatically configuration of the receiver parameters;
- post-processing of the data collection accordingly to the previous configuration.

This structure guarantees that the SDR receiver can ingest any type of raw data, no matter which their features are, and in a completely transparent way, without the need to manually specify any parameter, as long as the core processing functions support that signal and constellation grabbed. What is more, this makes the front-end and the receiver parts completely independent, thus easing re-configurations of the data collection system. Often non-specialist technicians are required to operate in remote data collection sites; the availability of metadata-compliant front-ends and receivers facilitates their tasks and reduces the risk of errors of failures.

III – EXPERIMENT RESULTS

This section describes two experiments, related to the DemoGRAPE [9] project and underlines some of the advantages of using SDR metadata, in particular for remote location data processing. The experiments and the solutions illustrated trace the points outlined in the Introduction. In addition, they refer to particular case studies, which conveniently suit the features and opportunities offered by the ION GNSS Metadata Standard, confirming the validity of the tool on field.

A – GNSS data ingestion with two different SDR receivers

This section provides an example of real GNSS data ingestion and processing, in a remote and non-accessible scenario, using two different SDR receivers properly adapted to use the metadata files generated by the GNSS front-end.

The example comes from a real issue experienced at EACF station. Each block of raw IF GNSS data grabbed in both Antarctica stations by the DemoGRAPE acquisition system is immediately post-processed by a basic SDR receiver, to quickly compute the scintillation indices. This is a first rough elaboration, useful to assess the presence of interesting ionospheric events and thus to trigger the permanent storage of the raw data. Such results, including the C/N_0, the amplitude scintillation index (S4) and the phase scintillation index (σ_ϕ), are transferred through the network and plotted in the *Research Lab*, as described in Section II.C and depicted in Figure 6. The plots related to the particular dataset collected on June 14, 2016, around 14:50 UTC are reported in Figure 10. No relevant amplitude scintillation activity is detected: there are moderate values of the S4 (up to 0.4) for satellites with PRN 03 and 31, but this is likely due to multipath reflection, as confirmed by the low values of the C/N_0 for the same satellites. On the contrary, a spike in the σ_ϕ is present for satellite with PRN 17 around 16:19 UTC. From a first analysis, this could be declared as phase scintillation. However, the short duration of the anomaly, and the very steep leading and trailing edges can raise suspicions.

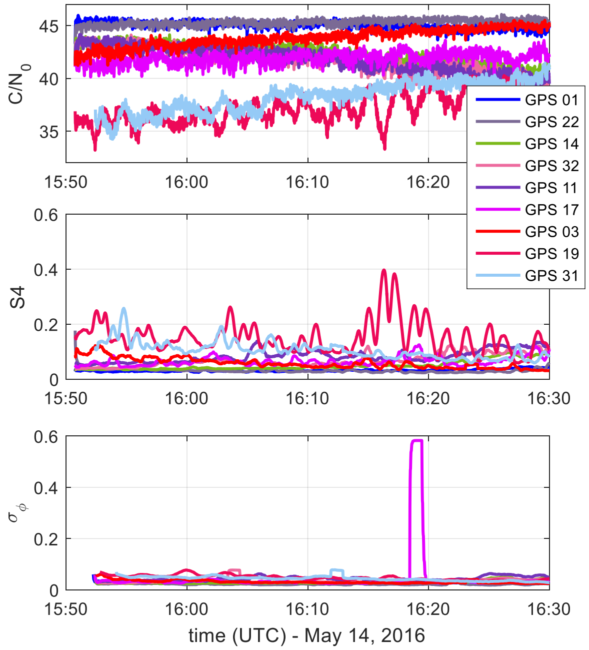


Figure 10. C/N_0, S4 and σ_ϕ as computed by the basic SDR receiver after grabbing at EACF Antarctica station.

The presence of raw IF data and of its respective metadata at the *Data Collection Site* allows to perform a deeper analysis of the signal. Following the approach described in Section II.C and exploiting the presence of the metadata co-located with the GNSS raw data, a custom version of a different and more advanced SDR receiver, enabling a deeper analysis phase scintillation, has been transferred to the *Data Collection Site*. The block of raw data has then been post-processed a second time, with the second receiver, and the results transferred through the network towards the *Research Lab*. The plot of the scintillation indices of the second SDR receiver confirm the doubts: the σ_ϕ spike of the first run is not actually a phase scintillation, but an artefact due to the lower processing capabilities of the first software receiver, motivated by the need to perform a quick and rough data analysis. Scintillation indices related to the second run with the advanced software receiver are reported in Figure 11.

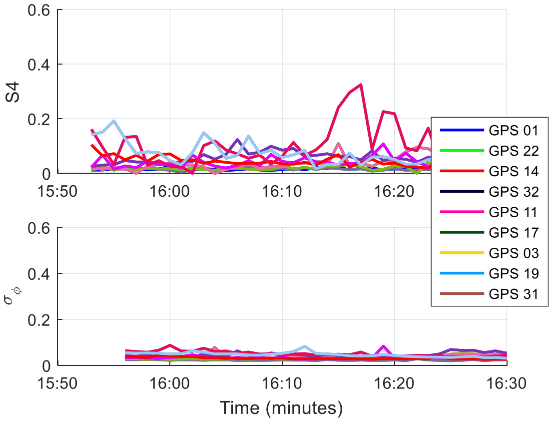


Figure 11. C/N_0, S4 and σ_ϕ as computed by the advanced SDR receiver after grabbing at EACF Antarctica station.

B – From “moving data” to “moving software”

This section deals with a case study of the “moving software approach”, described in Section II and depicted in Figure 4.

This example is the description of a real operation performed on the DemoGRAPE data collection installation at EACF. Before the installation of the monitoring station, we had a limited knowledge of the ionospheric scintillation at the two Antarctic sites where EACF and SANAE IV are located, especially at the signal level, as well as on the installation environment and characteristics. The installation itself, in fact, may introduce artifacts in the collected signal, such as multipath or radio frequency interference, which in some cases might be negligible for the estimation of the usual parameters (C/N_0, S4 and σ_Φ) but cannot be overlooked when collecting data that either might be used to re-play the scenario, or used to extract different pre-correlation pieces of information. For this reason, it was not possible to optimize the SDR receiver and to adapt the processing and the scintillation analysis to the particular scenario. Since one of the purposes of the installation was to grab GNSS raw data in correspondence of interesting events for a further analysis in the lab, a proper detection and data storage strategy had to be set-up, avoiding false alarms while minimizing missed detections.

A continuous monitoring over time was set-up at EACF, along with a simple scintillation detection rule. The front-end was configured to automatically grab a block of raw data, quickly process them with a SDR receiver, computing scintillation indices for the GPS L1 signals, comparing them to predefined scintillation thresholds, and according to the results, either discard or store the raw data on an external hard drive.

After a few weeks of monitoring, it was observed that a large amount of data was frequently saved and that the detection strategy and storage rules had to be modified, to reduce the number of false alarms, which has a dramatic impact on the use of the storage resources. In particular, multipath effects were identified as the main reason of amplitude scintillation false alarms. To solve the problem, an improved version of the SDR receiver and of the detection routine was prepared off line in the *Research Lab*. Afterwards, thanks to the modularity of the SDR implementation, only the new and updated modules were successfully moved to the PC located in Antarctica, exploiting the data link, even if slow and unreliable.

Similarly, after a few months of operations, it was noticed that the detection rule could be improved, to reduce the missed detections of phase scintillation events, especially at the SANAE station, which is located at higher latitudes. In order to give more weight to phase scintillations, the detection rule could was then modified. However, as discussed in Section II.B data transfer of polar data was not possible and the new detection routines were directly tested at the station site, by relying again on the transfer of the software through the low rate data link. Finally, a third upgraded version of the SDR receiver, tailored to Polar scintillation monitoring, was moved to the Antarctic stations.

The total size of the data amount to 359 KB in the first case, and to 947 KB in the second case; as a matter of comparison, the full SDR receiver upgrade amounts to about 12 MB.

This innovative aspect of this approach resides in the new paradigm of “moving software” rather than “moving data”. This example proved how this approach is particularly convenient for isolated and hardly accessible installations, and to fully exploit the advantages of the availability of GNSS metadata. In fact, it can be replicated in case of availability of raw data samples at different sites, provided that the corresponding metadata file is available as well. The use of a standardized metadata file allows to cope with the potential variety of data formats produced at different sites.

IV – A PATH TOWARDS A SMARTER GNSS DATA PROCESSING

In this paper, the issues related to the management of GNSS data were emphasized, with a specific focus on the bottleneck of the communication channel, proposing a novel approach based on transferring the processing code of the receiver instead of the data themselves. This approach has demonstrated an enormous reduction in the use of bandwidth and consequently costs reduction.

However, the volume problem is not completely solved. In fact, at collection sites, data grabbing systems are already collecting everyday tons of GNSS data for scientific study purposes. The collections of raw data samples is taking this issue even to a further level of complexity. As far as the scintillation analysis is concerned, a relevant portion of dataset and collected and stored in each site is completely irrelevant at informative level, since the events they highlight are often not correlated with real scintillations but with other types of interference (e.g. multipath). At this moment, this fact becomes evident only after the data processing, thus when the data are stored, dramatically increasing the data management and data storage costs. The update of the detection rule, as previously discussed, might not be enough, and more complex detection rules might introduce delays in the process that make it either unreliable or useless.

These remarks highlight the need of a smart data processing approach in the data domain (pre and post correlation) which could open new research paths at the crossing of data science and GNSS domains. In particular, the adoption of machine learning and deep learning approaches for the automatic classification of data collection able to trigger the storage phase showed some very interesting preliminary results and the authors aim at exploiting them as further development of the research activity presented in this paper.

V – CONCLUSIONS

Several experimental tests, based on the approach described in Figure 4 and in Figure 5, being performed on the Antarctica installations, have been described in the paper. They have been critically analysed, to assess the performances of the new post-processing paradigm and to evaluate the pros and the cons of adopting the ION GNSS Metadata Standard currently under definition. Metadata and on-site post processing can also foster the collaboration among research teams around the World, especially in Antarctica, where scientific cooperation is a key aspect. Metadata can facilitate the sharing of data and tools allowing different institutions to access the data and run their own software receiver, which automatically configures itself using the metadata. Different front-ends and acquisition systems configurations could be tested, as long as different metadata are provided. Coupling the use of metadata with a smart data processing approach, leveraging the use of advanced data analysis techniques, may represent a novel way to approach the GNSS data collection management as well as for the data processing.

ACKNOLEDGEMENTS

This study was carried out with financial support from the Italian Ministry of Education, Universities and Research (Ministero dell'Istruzione, dell'Università e della Ricerca, MIUR) and the Italian National Research Programme for Antarctica (Programma Nazionale di Ricerche in Antartide, PNRA), in the framework of the DemoGRAPE project, under contract 2013/C3.01.

The authors would also thank the Joint Research Center of the European Commission at ISPRA for their support to the DemoGRAPE project, providing assistance and hardware support. The present study was made possible also by the cooperation with the Brazilian Institute of Space Research (INPE), the Centro de Rádio Astronomia e Astrofísica Mackenzie (CRAAM) and the South African National Space Agency (SANSA).

REFERENCES

[1] L. Lo Presti, E. Falletti, E., M. Nicola, and M. Troglia Gamba, “Software Defined Radio technology for GNSS receivers,” in Metrology for Aerospace (MetroAeroSpace), 2014 IEEE. p. 314-319

[2] “Big Data A New World of Oppurtunities,” NESSI White Paper, December 2012.

[3] R. Magoulas, and B. Lorica, "Introduction to Big Data". Release 2.0 (Sebastopol CA: O'Reilly Media)

[4] S. Gunawardena, and T. Pany, "Initial Report of Activities of the GNSS SDR Metadata Standard Working Group", in Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014), Tampa, Florida, September 2014, pp. 1426-1432.

[5] S. Gunawardena, and T. Pany, "GNSS SDR Metadata Standard Working Group Report," in Proceedings of the 28th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2015), Tampa, Florida, September 2015, pp. 3218-3221.

[6] ION GNSS SDR Metadata Working Group Open Source code repository for GNSS SDR Metadata XML interchange standard: https://github.com/IonMetadataWorkingGroup

[7] S. Gunawardena, “A Universal GNSS Software Receiver Toolbox for Education and Research”, *Inside GNSS*, July/August 2014, pp. 58-67.

[8] N. Linty, R. Romero, F. Dovis and L. Alfonsi, "Benefits of GNSS software receivers for ionospheric monitoring at high latitudes," in Proceedings of Radio Science Conference (URSI AT-RASC), 2015 1st URSI Atlantic, Las Palmas, 2015, pp. 1-6

[9] Istituto Nazionale di Geofisica e Vulcanologia (INGV), “Demogrape project website.”. Available: http://www.demogrape.net

[10] C. Cesaroni, L. Alfonsi, S.V. Veettil, P. Jihye, R. Romero, N. Linty, F. Dovis, D. Barroca, M. Cobacho Ortega, and R. Orus Perez, "Monitoring Ionosphere Over South America: The MImOSA and MImOSA2 projects," in Proceedings of Navigation World Congress (IAIN), 2015 International Association of Institutes of , pp.1-7, 20-23 Oct. 2015.

[11] G. Povero, M. Pini, L. Alfonsi, L. Spogli, D. Di Mauro, F. Dovis, R. Romero, M. Le Huy, P. Abadi, V. La The, and N. Floury, “Ionosphere Monitoring in South East Asia in the ERICA Study,” in Proceedings of the International Technical Meeting ION ITM. Monterey, California, January 25-28 2016.

[12] N. Linty, and I. Hunstad, “Installation and configuration of an ionospheric scintillation monitoring station, based on GNSS receivers in Antarctica,” under publication in Rapporti Tecnici INGV, ISSN 2039-7941, pp. 1-20, 2016.

[13] N. Linty, R. Romero, C. Cristodaro, F. Dovis, M. Bavaro, J. T. Curran, J. Fortuny-Guasch, J. Ward, G. Lamprecht, P. Riley, P. Cillers, E. Correia, and L. Alfonsi, “Ionospheric scintillation threats to GNSS in polar regions: the DemoGRAPE case study in Antarctica,” in Proceedings of European Navigation Conference ENC-2016. Jun 2016, pp. 1-7

[14] N. Linty, F. Dovis, R. Romero, C. Cristodaro, L. Alfonsi, and E. Correia, “Monitoring ionosphere over Antarctica by means of a GNSS signal acquisition system and a software radio receiver,” in Proceedings of 2016 International Technical Meeting of The Institute of Navigation. Manassas, VA (USA): Institute of Navigation, Jan 2016, pp. 549 - 555.