Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 58 (2016) 1177–1200

ADVANCES IN SPACE RESEARCH
(a COSPAR publication)

www.elsevier.com/locate/asr

The ESPAS e-infrastructure: Access to data from near-Earth space

Anna Belehaki [a,b,*], Sarah James [c], Mike Hapgood [c], Spiros Ventouras [c], Ivan Galkin [d], Antonis Lembesis [e], Ioanna Tsagouri [a,b], Anna Charisi [a,b], Luca Spogli [f], Jens Berdermann [g], Ingemar Häggström [h], the ESPAS Consortium [1]

[a] *National Observatory of Athens, Greece*
[b] *National and Kapodistrian University of Athens, Greece*
[c] *RAL Space, STFC Rutherford Appleton Laboratory, UK*
[d] *University of Massachusetts Lowell, USA*
[e] *ATHENA Research Center, Greece*
[f] *Istituto Nazionale di Geofisica e Vulcanologia, Italy*
[g] *Deutsches Zentrum für Luft- und Raumfahrt e.V., Germany*
[h] *EISCAT Scientific Association, Sweden*

Received 10 March 2016; received in revised form 7 June 2016; accepted 8 June 2016
Available online 17 June 2016

Abstract

ESPAS, the "near-Earth space data infrastructure for e-science" is a data e-infrastructure facilitating discovery and access to observations, ground-based and space borne, and to model predictions of the near-Earth space environment, a region extending from the Earth's atmosphere up to the outer radiation belts. ESPAS provides access to metadata and/or data from an extended network of data providers distributed globally. The interoperability of the heterogeneous data collections is achieved with the adoption and adaption of the ESPAS data model which is built entirely on ISO 19100 series geographic information standards. The ESPAS data portal manages a vocabulary of space physics keywords that can be used to narrow down data searches to observations of specific physical content. Such content-targeted search is an ESPAS innovation provided in addition to the commonly practiced data selection by time, location, and instrument. The article presents an overview of the architectural design of the ESPAS system, of its data model and ontology, and of interoperable services that allow the discovery, access and download of registered data. Emphasis is given to the standardization, and expandability concepts which represent also the main elements that support the building of long-term sustainability activities of the ESPAS e-infrastructure.

Keywords: Near-Earth space data; Data infrastructure; Data model; Ontology

1. Introduction

Near-Earth space is the region that extends from the middle atmosphere up to the outer radiation belts. This region is of significant interest because of its potentially undesired effects on human life and on technological systems, whose understanding, modeling and prediction requires continuous scientific exploration and advances. Consequently a number of observing systems have been set up to acquire observations from the near-Earth space, producing a wealth of diverse types of data which still need to be homogenized and organized in order to become widely accessible.

The exploitation of multi-instrument data from a large number of distributed observing sites is the requirement

* Corresponding author at: National Observatory of Athens, Metaxa and Vas. Pavlou, Palaia, Penteli 15236, Greece.
E-mail address: belehaki@noa.gr (A. Belehaki).

[1] http://www.espas-fp7.eu/trac/wiki/PublicPages/ESPASConsortium.

http://dx.doi.org/10.1016/j.asr.2016.06.014

for accurate predictions of the near-Earth space environment. As the near-Earth space is part of the complex Sun-Earth system, data from the Sun, the interplanetary medium but also from the upper and lower layers of the atmosphere, are needed to drive near-Earth prediction models. In space physics, predictions are made via physical, semi-empirical or empirical models. The models are fed with observed values (e.g., measured solar wind speed and density) or with typical values for specific environmental parameters (e.g., average speed and density of the slow solar wind during solar minimum), and the model output provides values which can be compared to other observations or parameters derived from observations (e.g., local or global geomagnetic activity index). A comprehensive comparison between model results and observed data enables the community to distinguish between models with good and with poor performance under certain geophysical conditions. Space physics models with good predictive capabilities may be used to forecast accurately the state of the space environment and to enable the end user communities to mitigate the effects of major disturbances on humans and technological systems. Results obtained from model runs depend to a large extent on the boundary conditions. Sometimes the problem can be solved by specifying boundary conditions over the entire globe and running the model on a global scale. However, specification of global boundary conditions requires data from many observational sites. Ionospheric total electron content (TEC) maps are a typical example for the dependency of maps on global data coverage in order to be realistic.

This specific need has led the space science community to work intensively for the development of systems that can facilitate data discovery and processing.

The Inter-university Upper atmosphere Global Observation NETwork (IUGONET) has been implemented by Japanese universities and institutes and aims at providing new research platforms, metadata database and analysis software tools, to facilitate the use and distribution of the long-term observation data for upper atmospheric physics (Hayashi et al., 2013). In addition to the open search service based on the metadata database, IUGONET provides: automatic data download; data analysis without regard to the file format of the data; parallel display of different types of data; utilization of various analysis functions (e.g., frequency analysis, filtering); output into an ASCII file or image files. The generated metadata are archived as XML files for interoperability with other metadata databases and future expandability. As the base of the metadata format, IUGONET selected the Space Physics Archive Search and Extract (SPASE) data model/metadata format (Merka et al., 2008), that has been modified to best match the upper atmosphere data, to create the IUGONET common metadata format.

The Automated Multi Dataset Analysis (AMDA) is provided by the Centre de Données de la Physique des Plasmas (CDPP) supported by CNRS, CNES, Observatoire de Paris and Université Paul Sabatier, Toulouse (Jacquey et al., 2010). AMDA is a web-based facility for on line analysis of space physics time series data coming from either its local database or distant ones. AMDA offers functionalities to access and analyze multi-point and multi-instrument data in a transparent way by the user. More precisely, AMDA provides functionalities for: performing search of events; performing automated and semi-automated characterization of events; extracting sub-database from an input time table; performing basic data treatment in order to provide to the user data ready to use with her/his favorite software. AMDA provides direct access to data from distant databases in a transparent way and includes a connection layer compliant with the SPASE standards.

The HELiophysics Integrated Observatory (HELIO) has been developed in the framework of an EU-FP7 research infrastructure project (Bentley et al., 2010). HELIO adopts the concept of distributed network of services that addresses the needs of a broad community of researchers in heliophysics. It coordinates access to the resources needed by the community, and provides services to mine and analyse the data. HELIO has been developed as a set of independent services. Several ways are provided to access them. The services can be used individually, within a workflow or scripting language, or through the HELIO Front-End web user interface. HELIO provides the scientist with an operational scenario for heliophysical data handling (Pérez-Suárez et al., 2012). It relieves the user from the burden of data source identification and data integration, as its web interface makes it possible to place complex searches on multiple data repositories relevant to heliospheric data in a unified, user-transparent way. This facilitates research, and creates a favorable operational environment for knowledge discovery.

The Integrated Space Weather Analysis System (iSWA) is a U.S. Government Computer Server that provides access to space weather data products and tools for both real-time as well as historical analysis (http://iswa.gsfc.nasa.gov/iswa/iSWA.html). Users are provided with the capability to specify and/or forecast the large scale and local space environment. Certain data products may be in experimental or evaluation phases of development. iSWA is customer configurable and adaptable for use as a powerful decision making tool, providing mission managers and decision-makers with personalized "quick look" space weather information, detailed insight into space weather forecasts, and tools for historical impact analysis. iSWA data management is based on a comprehensive data model that drives the system and it is supported by the Cygnet software.

Driven from these developments and in order to reach further advances in the discovery and uniform access to data, the European Commission has funded the near-Earth space data infrastructure for e-science (ESPAS) project from Framework Programme 7. ESPAS provides a one-stop shop for researchers and users of research results who wish to exploit multi-instrument multi-point science data for analysis.

Data handled by ESPAS originate from a variety of different sensors including space-borne and ground-based observations. Currently data accessed through ESPAS are provided by several spacecraft instruments including CLUSTER, DEMETER, NOAA-POESS, ACE, MAGION3, IMAGE, CHAMP, GRACE, and by ground-based instruments such as ionosondes, magnetometers, Fabry–Pérot *interferometers*, Incoherent Scatter Radars, SUPERDARN and GNSS receivers. ESPAS also provides access to model output data such as EDAM[2], CMAT[3], IRI[4], TaD[5] and SIRM[6]. A detailed list of the observation collections that are registered in ESPAS, or are under registration, is given in Appendix A. These are very heterogeneous data regarding their format and resolution, supplied by a number of widely distributed institutions with different data acquisition and processing philosophies and different data distribution policies.

ESPAS offers an homogenised mean of searching for data and provides common access to all data identified from a query. In addition the system can extract parameters from data files and make them available to the user in a common format, independently of the format of the source file. This facilitates greatly the operations of key Research Infrastructures like EISCAT which is a representative example. EISCAT is a major European research infrastructure with radar facilities in Northern Scandinavia and at Svalbard. For funding reasons, the current radars operate in campaign mode, rather than continuously, with data-taking periods of some hours to several days. The incoherent scatter radar technique is the most powerful ground-based tool to conduct research on the whole atmosphere and ionosphere. The basic parameters are the derived narrow field profiles of plasma parameters. To get a better understanding of the physics the data represent it is often needed to gather data from other instruments to complete the picture in several dimensions like time, space and depth of physical parameters. ESPAS with its functionalities, is able to investigate what data is available and suitable for this, by giving lists of available complementing observations covering the same time periods, nearby or overlapping locations, with direct links to the data. This allows also a wealth of further parameters, describing the ionosphere and neutral atmosphere, to be derived. On the other hand, ESPAS will also significantly increase the user base of EISCAT. The EISCAT data is normally available in the Madrigal database, but the user needs to be familiar with the EISCAT data acquisition techniques, in order to use efficiently this facility. ESPAS enables the data to be visible, searchable and easily accessible by other communities. In the long run, this can open up new funding possibilities for the radars enabling extended operations over longer times bringing the science further with increased speed.

In this article, we present an overview of the architectural design of the ESPAS system, we analyse its data model and ontology, we describe the set up of interoperable services that allow the discovery, access and download of registered data, we summarize the main functionalities of the ESPAS platform towards the end-user and finally we provide an outlook of our plans for the sustainability and integration of the system to the global e-infrastructures network.

2. ESPAS overview

ESPAS is a data e-infrastructure facilitating discovery and access to observations and model predictions of the near-Earth space environment (Belehaki et al., 2014). Through the ESPAS portal, the user can have access to a large number of repositories with heterogeneous data from ground and space, in situ and remote sensors. The user can perform searches for observations using specific criteria (e.g. time, instrument) and then download data files or data values.

From its inception, ESPAS was designed to meet the following requirements:

- Integration of heterogeneous data from multiple providers, ranging from ground-based observations acquired with multiple instruments to data from space borne sensors. To do so, specific policies were established on identification, access, availability, quality, sharing and re-use of the data (or metadata) of the participating content providers.
- Enablement the metadata/data search and access across multiple data sources through a central platform. This was made possible through the establishment of a variety of workflows (data flows) initiated on the ESPAS platform.
- Provision of value-added services to explore metadata, visualize or manipulate the integrated data, and eventually mine metadata.
- Acting to serve as a test-bed for proposed methodologies and standards for validation and optimization of metadata of a specific data collection.
- Provision of the possibility to allow for extensive testing through several test and use cases, designed to serve the needs of the wide and interdisciplinary user and provider communities for computationally intense science carried out in a highly distributed data environment.
- Integration into the wider European scientific infrastructures. To do so, it carries on and extends the policies and procedures issued at its participating data providers on data openness, quality, identification, etc.

ESPAS uses and adapts well-established techniques from related data e-Infrastructures disciplines and from

[2] Electron Density Assimilative Model.
[3] Coupled Middle Atmosphere and Thermosphere general circulation model.
[4] International Reference Ionosphere.
[5] Topside Sounders Model assisted by Digisonde.
[6] Simplified Ionosphere Reference Model.

the Digital Libraries domain to provide the required semantic integration of the participating data sources. Main elements towards building interoperability are: the data model, which is compatible with and encapsulates all underlying data formats; the definition and employment of XML schemas for metadata exchange format; the domain-specific vocabularies used to describe the near-earth science data and phenomena; the services (*wrappers)* installed at the data nodes to support metadata publishing and data exposition; the central coordinating platform that showcases the metadata aggregation and data access through a portal that provides tools for data registration and validation and customized search workflows for more complex queries. An illustration of ESPAS architecture is given in Fig. 1. The ESPAS platform adopts the principles of the Service Oriented Architecture (SOA), based on the D-NET system (Manghi et al., 2014), and it is structured in the following layers:

– *Enabling Layer:* provides the necessary components that glue, manage and organize all the services in an SOA. It implements service registration and a notification based communication system, making the overall system extensible and extendible.
– *Data & Semantic Integration Layer:* focuses on the metadata harvesting and the basic data management services of the ESPAS system: storage, database, index.
– *Value-added Services Layers:* include advanced search, statistics capabilities, and enable the development of models and visualizations tools.
– *Web Layer:* provides the User Interface service, i.e., the ESPAS portal.

3. ESPAS data model

The ESPAS data model aims at enabling interoperability among heterogeneous data registered in ESPAS. The data model is built entirely on ISO 19100 series geographic information standards, particularly the ISO 19156 Observations and Measurements (O&M) standard. This standardisation facilitates interoperability with other information systems and provides freedom to mix and match information system components without compromising overall success [*ISO 19101:2002*].

The general structure of the ESPAS data model gives a central place to the concept of "observation". According to (Fowler, 1998) an "observation is an act that results in the estimation of the value of a feature property using a designated procedure, such as a sensor, instrument, algorithm or process chain". An observation is associated with a discrete time instant or period through which a number, term or other symbol is assigned to a phenomenon. The result of an observation is an estimate of the value of a property of some feature, so the details of the observation are metadata concerning the value of the feature property.

Example. Measuring (*the act of the observation*) that the F2-layer Critical Frequency (foF2) of the Ionosphere above Athens is 5.2 MHz on 1/1/2015 at 16:00 GMT by

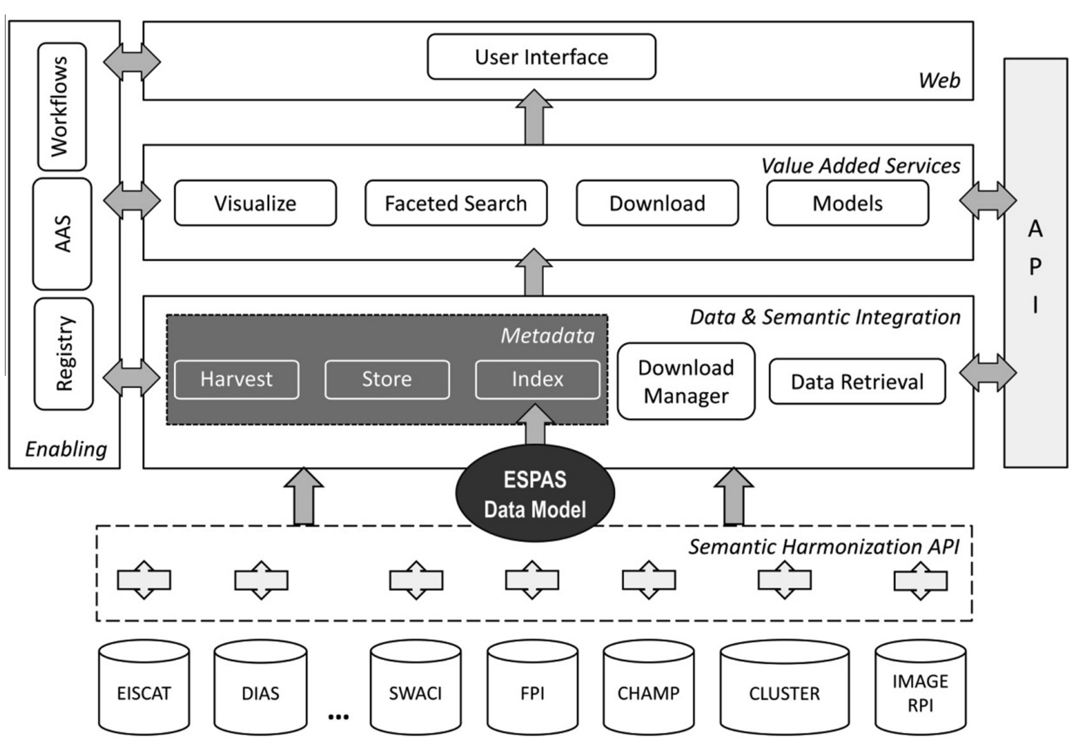

Fig. 1. A high level representation of the ESPAS architecture.

Athens Digisonde. The *featureOfInterest* is the Ionosphere, the *observedProperty* is the foF2, the *procedure/process* is the acquisition made by the *instrument* (Athens Digisonde) operated by the National Observatory of Athens (NOA) which is the *platform* and the *result* of the observation is 5.2 MHz.

Following this approach, the data which ESPAS data model is aimed at describing is always considered as observation results and the observation together with its properties provide relevant metadata.

Besides the main concept of "observation", the other related concepts that are used in ESPAS data model are listed below, while a high-level overview of the relationships among them is presented in Fig. 2.

- *Feature of Interest:* a real-world object, carrying the properties which are under observation. It is the subject of the observation. In ESPAS, this corresponds to the region of space (e.g. Ionosphere, Magnetosphere, etc.) of the observed property of the observation. A controlled vocabulary (ESPAS Space Physics Ontology) is used for its values.
- *Observed Property:* a phenomenon associated with the feature of interest for which the observation result provides an estimate of its value. It is the object of the observation (e.g. Temperature, Electron Density) and a controlled vocabulary (ESPAS Space Physics Ontology) is used for its values.
- *Observation Result:* the act of the observation produces a numerical artefact, i.e. the observation result. ESPAS is not particularly concerned with details of observation result structures, but is aimed at providing the required metadata in order to make this result fully understandable and exploitable. Therefore, in ESPAS, observation result is regarded the set of metadata for accessing and obtaining the actual values of the observed property obtained by the action of observation.
- *Process:* a designated procedure used by the action of observation in order to assign a number, term or other symbol to a phenomenon generating the observation result. A procedure is often an instrument or sensor but may be a process chain, human observer, an algorithm, a computation or simulator [*ISO 19156*]. Therefore, a procedure may consist of more than one component. A component shall be either an *acquisition* or a numerical *computation*.
  - An Acquisition process interacts with the feature of interest/sampling feature to provide a result and involves the use of an Instrument which is mounted

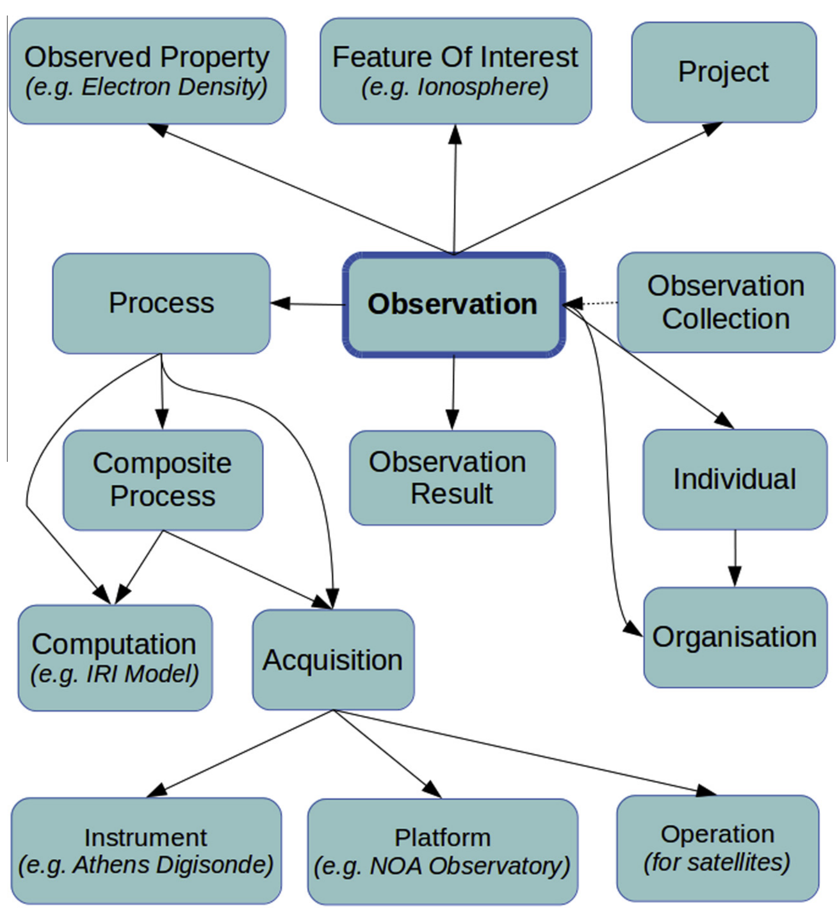

Fig. 2. High-level overview of ESPAS data model concepts.

on a Platform that may have an Operation (for satellites, aircrafts).
  - A Computation process involves only pure computation (no Instrument is involved), as in the case of Models (e.g. EDAM, SIRMUP) or specific software (e.g. ARTIST for the autoscaling of the ionograms).
  - A Composite process consists of more than one components of type Acquisition or Computation.

- *Instrument:* designations for the measuring instruments/sensors which interact with the feature of interest in order to obtain an estimate of the observed property. In ESPAS an instrument is assigned to a specific type of Instrument obtained from the relative controlled vocabulary (ESPAS Space Physics Ontology).
- *Platform:* an identifiable object which brings the acquisition instrument(s) to the appropriate environment (e.g. satellite, ground-based station) in order data to be acquired according to the observation objectives. In ESPAS the concept of platform is used to distinguish ground-based from space-based platforms (ESPAS Space Physics Ontology).
- *Operation:* Information about a platform operation – e.g. orbit of a satellite – needed for the data acquisition. The concept of platform operation applies only to platforms that are in motion during the acquisition e.g. a satellite – geostationary or orbital, an aircraft, a ship and not to the operation of a static platform such as a ground station.
- *Project:* an identifiable activity/project designed to accomplish a set of objectives in order to produce datasets.
- *Collection:* any set of existing observations. The organisation of observations into collections is based on specific criteria, e.g. common observed property, common instrument, common process. An observation may be aggregated in more than one observation collections.
- *Organisation:* a body/organization having a particular role associated with a real world object.
- *Individual:* an individual having a responsibility regarding a real world object.

One of the main ESPAS extension to the ISO 19156 Observations and Measurements (O&M) model is the definition of the Process Capability as a property of a Process. This property has been added in order to describe specific and usually limited capability of the sensing instrumentation or models to produce information about the Observed Property. Therefore, the "capability" in this context is the description of an Observed Property that this Process is capable of measuring and its related information: units, dimensionality (instance and timeline), valid minimum, valid maximum, fill value, vector representation (crs, component, projection), qualifier, compressed representation, extracted Parameter.

A detailed description of ESPAS Data Model is provided by Ventouras (2013). The definition of the ESPAS controlled vocabularies is given in Appendix B.

4. Space physics ontology for ESPAS

The ESPAS data portal manages a vocabulary of space physics keywords that can be used to narrow down data searches to observations of specific physical content. Such *content-targeted* search is an ESPAS innovation provided in addition to the commonly practiced data selection by time, location, and instrument. The ESPAS Space Physics Ontology is the cornerstone of the domain-specific data search functions. The ontology is organized in several hierarchies of *keywords* connected to each other via a "broader-narrower" relationship. Understanding the ontology hierarchies is critical for efficient data search and discovery in ESPAS.

4.1. Observation, observed property, and process capability

Space physics keyword vocabularies in ESPAS ontology are tailored to the OGC standard for Observations and Measurements (O&M), in which it is the *Observed Property* that describes the underlying physical phenomena, while the *Observation* is an act that results in evaluation of the feature property (ISO, 2011). Most of the design efforts to build the ESPAS ontology have focused on the Observed Property vocabulary that is closely related to physical concepts and relationships specific to the near-Earth space.

O&M standard makes a clear, important distinction to separate the Observed Property (as occurring naturally) from the Process Capability (that describes specific and usually limited capability of the sensors or models to evaluate the property). For example, one of the common Observed Properties of the ionospheric plasma is its bulk velocity that naturally exists as a vector quantity everywhere in the 3D volume of the Earth's ionosphere. However, the bulk velocity can be observed with various spatial coverages, completeness of the vector representation, choices of coordinate system and units, background assumptions, approximations, etc. For example, certain high-frequency ionosondes can detect velocity of the bottomside ionospheric plasma under assumption that the plasma is drifting across the sky over the observatory location as a single entity. In this case, the Process Capability points to the Observed Property "drift velocity of electrons", but defines that its dimensionality as 0D.Point (rather than the natural 3D.Volume), along with descriptions of simplifications and other assumptions used in the process of computing the velocities.

The stated relationship between Observed Property, Process Capability and Observations is illustrated in Fig. 3 showing key constituents of the O&M data model. In this figure, with "1" we annotate the sequence between two model constituents when the target has unique capability (for example "*Result*"), while the "1+" annotation is used as a symbol when the target constituent has multiple capabilities (for example "*Process Capability*").

The ESPAS search engine uses the vocabulary of Observed Properties rather than the Process Capabilities.

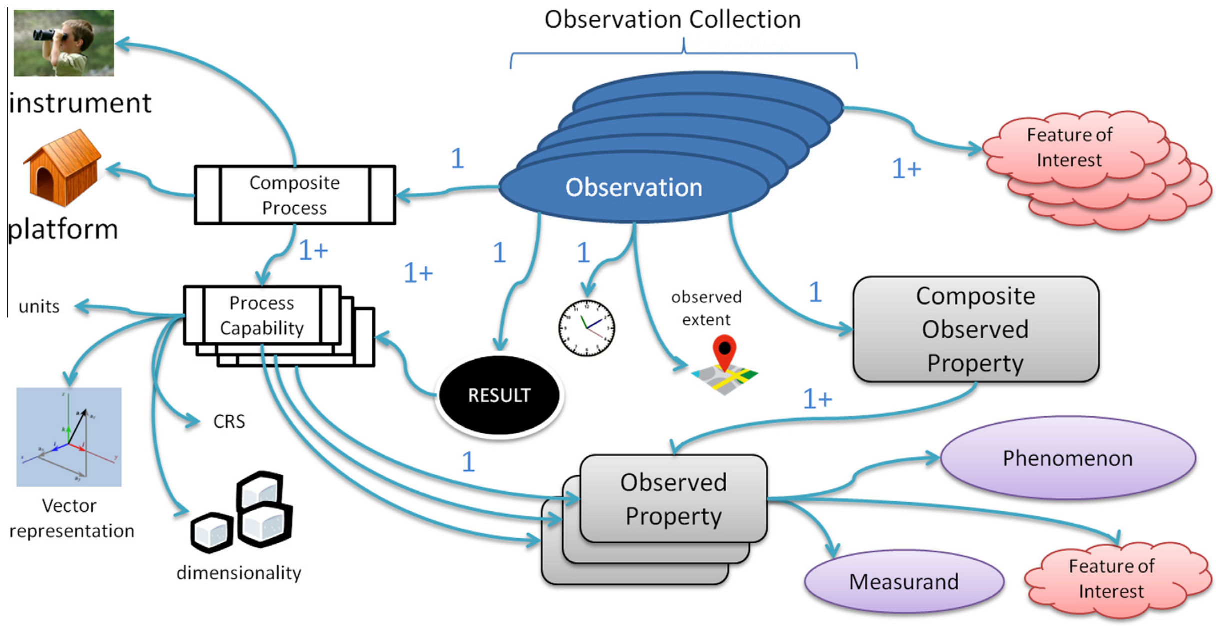

Fig. 3. Relationship between observation, observed property, and process capability.

Search for the "drift velocity of electrons" will hit several Observation Collections produced by incoherent scatter radars, coherent radars, ionosondes, with a variety of Process Capabilities associated with different sensors in their instrument categories.

4.2. Observed property vocabulary

A large number of Observed Properties are defined in the space physics vocabularies, ranging from particle fluxes to critical frequencies of plasma layers. Given the overwhelming variety of observations managed by ESPAS, it is inefficient to present scientist with an alphabetically sorted list of all registered Observed Properties.

For a more structured approach to select the Observed Property criteria for the content-targeted data searches in ESPAS, the Observed Property vocabulary is built as a set of hierarchies describing various aspects of the properties. While some of the resulting definitions are quite elaborate, especially in the category of wave phenomena, all Observed Properties have two defined components: Phenomenon and Measurand.

def. Phenomenon: (not to be confused with Event): underlying physical phenomenon for which the Observation provides an estimate of its value.
def. Measurand: measurable quantity of the Observed Property, whose value is estimated in Observation.

ESPAS provides hierarchical lists of Observed Properties sorted by their Phenomenon and Measurand aspects so as to allow rapid access to the search criteria. The Phenomenon vocabulary provides a greater structural organization of the concepts by using one master criterion of classification:

- Waves are sorted by oscillating agent.
- Particles are sorted by particle kind.
- Fields are sorted by field type.
- Photons are sorted by frequency band.

Correspondingly, the organization of the Phenomenon vocabulary is given in Fig. 4.

The Measurand terms are arranged in a simple list with only a minor structurization (Fig. 5).

5. ESPAS interoperability services

To facilitate access to heterogeneous data offered by data providers a set of basic supporting services has been defined and implemented by ESPAS. Each of these services essentially implements a client–server protocol, with the server part residing at the data provider side (in the Semantic Harmonization Layer, bundled within the ESPAS wrapper) and the client side at the ESPAS platform Data Management Layer:

- An OGC compliant Catalogue Service of the Web (CSW), which supports the identification of ESPAS resources offered by each data provider.
- A Download Service, that facilitates the download of data bundles in terms of data collections offered by each provider.
- An OGC Compliant Sensor Observation Service (SOS) which facilitates the collection of selected data parameters or values from the observations of each data provider.

The visualization of the internal workflows for different user queries is given in Fig. 6.

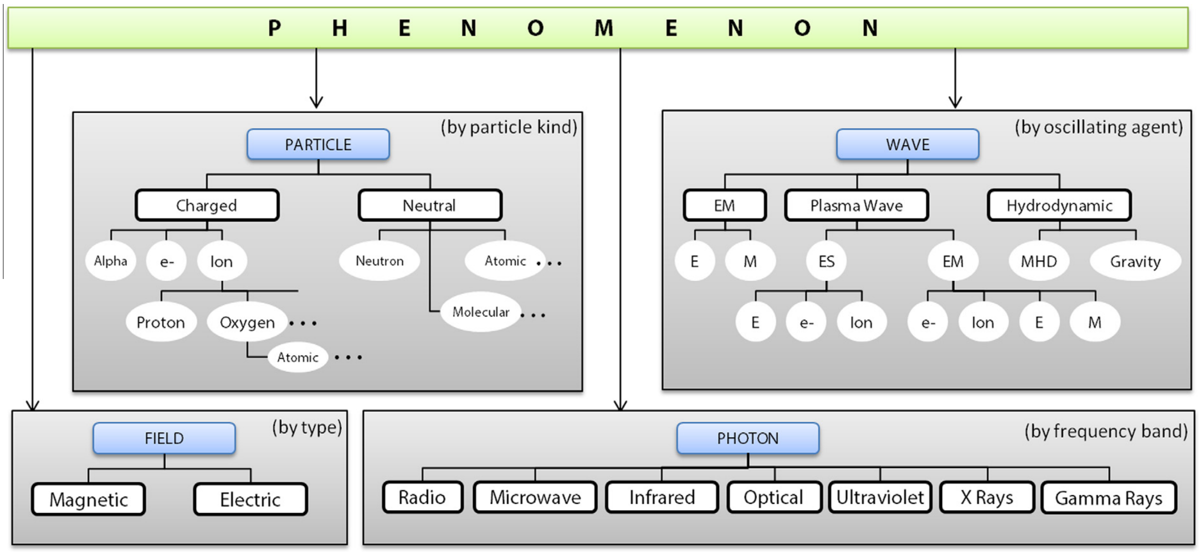

Fig. 4. Phenomenon vocabulary.

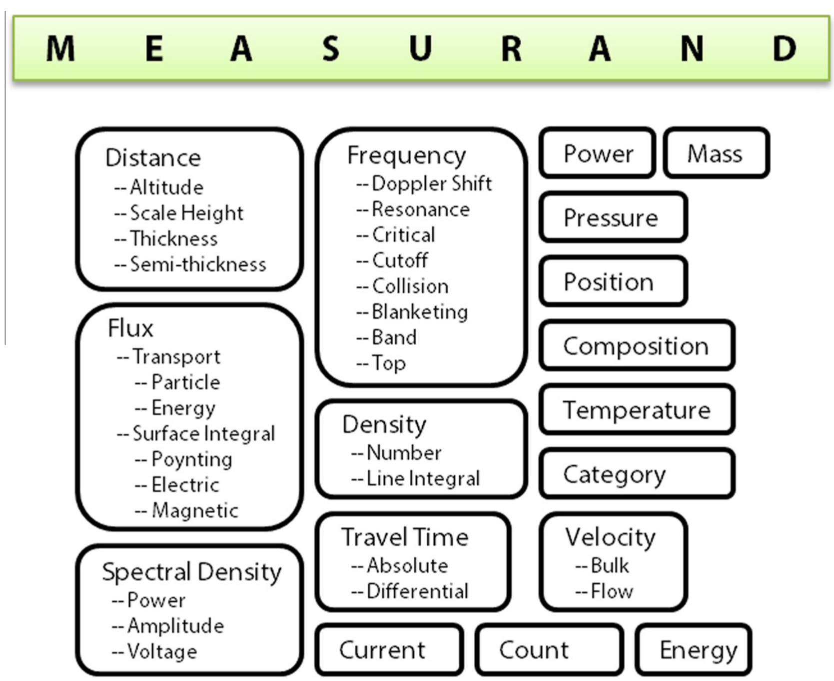

Fig. 5. Measurand vocabulary.

5.1. Harvesting metadata: OGC catalogue service

OGC Web services (OWS) are the prevailing types of services in the geospatial application domain (Doyle et al., 2001). Based on the OGC Catalogue Service for the Web (CSW), i.e., a profile of the catalogue service with the goal to expose the catalogue functionality over the web (Nebert et al., 2007), ESPAS has partially implemented the specified CSW interface specification to accommodate the getRecords operation. Through the getRecords operation, clients are able to submit queries for the discovery of metadata records for each of the major metadata entities specified by the ESPAS model. Constraints on the expected results are expressed as Contextual Query Language – CQL queries over a list of specific properties, which include, the type and the modification date of the expected metadata entity. The supported binding includes HTTP Key Value Pairs (KVPs).

5.2. Retrieving datasets: download service

The download service is an ESPAS proprietary service that enables ESPAS clients to retrieve data collections in formats defined by the data providers and specified in the metadata descriptions. It is implemented on top of basic D-NET services (Manghi et al., 2014), which provide

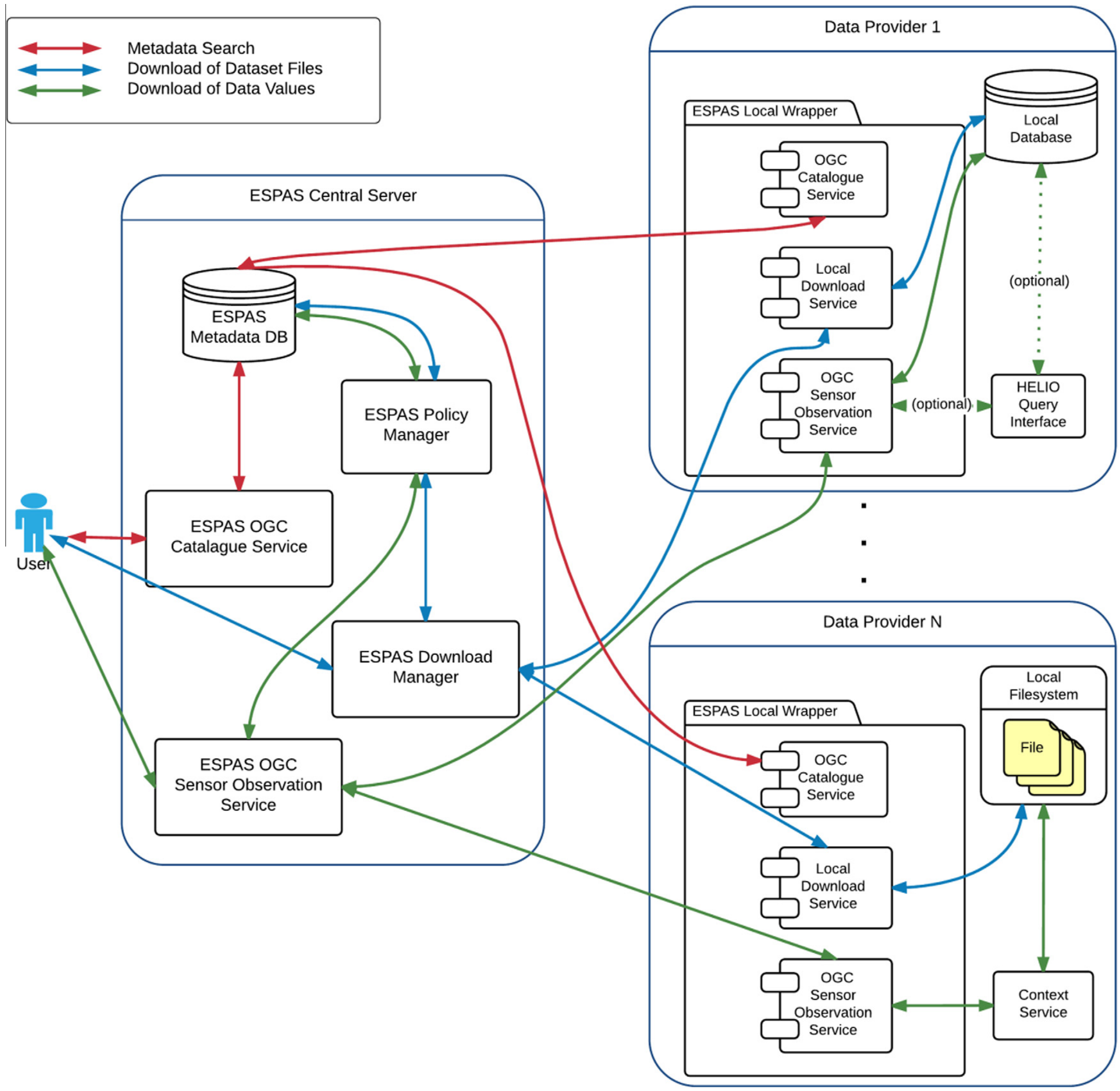

Fig. 6. A visualization of the internal workflows for different user queries.

subscribe-notify and database interaction operations, as well as on other services that facilitate policy enforcement functionality. The download service operates in an asynchronous manner so as to avoid blocking clients from performing other activities while the download request is submitted and processed by the associated data providers. In terms of architecture, the download service may be logically split into three distinct components, namely the Central Download Manager, local Download Service and Policy Manager. The Central Download Manager component is responsible for checking conformance of each request against the policies defined in the Policy Manager and coordinating the interaction with the local Download Services. The Policy Manager is an XACML v3.0 (Rissanen, 2013) based policy server, responsible for maintaining the policies defined by the data providers and validating each request against related policies. Local Download Services are deployed on each provider and are responsible for processing the download requests related to each distinct provider.

5.3. Retrieving data values: OGC sensor observation service

One of the well-known types of OGC Web services is the Sensor Observation Service (SOS). SOS is a web service that is used to query real-time sensor data and sensor data time series and is part of the Sensor Web vision. The offered sensor data comprises descriptions of sensors themselves, which are encoded in the Sensor Model Language (SensorML) and the measured values in the Observations and Measurements (O&M) encoding format (Cox, 2006).

The preliminary implementation offered by ESPAS is supporting partially the specified SOS service interface (Bröring et al., 2012). It includes a customized implementation of the getResult and getResultTemplate operations, which facilitate the retrieval of specific observed property values and value templates respectively, for observations performed by data providers. Among the set of optional query criteria defined for the getResult operation the provided implementation supports only time related ones, whereas in terms of bindings only HTTP Key Value Pair (KVPs) based requests.

Moreover, the actual implementation requires a layer that retrieves data from the data providers, whether this might reside in databases or files. Using the HELIO approach, and bundled in the ESPAS wrapper, the ESPAS implementation of the SOS service is able to do one of the following through the use of generic services that rely on configuration: convert standard SQL queries, HELIO-compliant databases, or retrieve from files through the CSX generic services.

6. ESPAS functionalities for the end-user

6.1. Basic functionalities

The basic functionalities provided by ESPAS can be grouped into five main categories:

1. *Metadata search* for observations that satisfy the following specific criteria:
   - Time period: the time period when the observations were acquired
   - Assets: the Instruments and Models that were used for the generation of the observations
   - Observed properties: the observed properties that were measured in the observations
   - Observation collections: the collections that the observations belong to
   - Location: the location of the platforms (ground-based observatories or satellites) on which the instruments are mounted and used for the generation of the observations
2. *Download of data files*: The result of a metadata search is a list of the observation collections (that contain the observations) that satisfy the query criteria. Then, the user can proceed to download data files residing in the ESPAS data providers nodes. The data files are provided in the original format as defined by each ESPAS data provider.
3. *Download of data values* (extracted parameters): Following the metadata search the user can request to download data values of specific observed properties, as extracted from the data files. A subset of the observed properties is available for download as extracted data values. After a data value download request, the user gets as a result a text file (csv or XML format) that contains the values of the selected observed properties.
4. *Plotting tools*: Downloaded data values can be plotted using either the quick plot of the ESPAS platform or the most advanced IDL-based plotting tool for OGC data files.
5. *Registration and validation of data*: Data collections from space missions and ground-based instruments can be registered in ESPAS following the standards of the ESPAS data model and domain ontology. The service is available to the scientific community upon request.

The search for metadata is the fundamental service based on which the user can proceed to the data queries. This service is available to all and no registration is required. However, in order the user to proceed with the use of more advanced functionalities provided by the system, registration is required. The search for metadata is based on the execution of workflows whose design is dynamic and depends on the user needs. The user can start its query with any of the criteria: time, asset, observed properties or observed collections, and finish its workflow design at any criterion. If for example it is supposed that the chosen criterion is "observed properties", a list of all the observed properties (in alphabetical order) that are associated with observations appears on the right side of the screen (see Fig. 7a). The filters (Phenomenon, Measurand, Qualifier) which appears on the left side serve to narrow down or facilitate the selection of observed properties presented in the right part. A hierarchical view of each filter is provided. The filters provide an alternative search option based on the ESPAS space physics ontology definition. It should be noted that the "`OR`" relationship is implied among the options of the same filter, and an "`AND`" relationship is implied between the filters. If one selects the "`ion`" and "`electron`" as phenomenon (see Fig. 7b), then the observed properties with phenomenon "`ion`" OR "`electron`" are presented on the right. A high level query could be represented as: phenomenon = "`ion`" OR phenomenon = "`electron`". However, if also the "`density`" is selected at the "`measurand`" filter, then the observed properties with phenomenon "`ion`" OR "`electron`" AND measurand equal to "`density`" are presented. A high level query could be represented as: measurand = "`density`" AND (phenomenon = "`ion`" OR phenomenon = "`electron`").

Contrary to the search described above which is progressive and non-ordered, the Metadata Search by Location is ordered. In this case, the search is using as criteria the time period (up to a maximum of 30 days) and the location of the instrument (ground-based observatories and/or satellites).

6.2. Value added services

In the framework of the ESPAS project various kind of possible Value Added Services (VAS) have been discussed in order to support the user in identifying interesting time periods or via additional high level operations on the data. In general, one can identify two fundamentally different types of value added services depending on the availability of real data or only metadata within a data infrastructure system.

The two VAS types are based on entirely different system requirements. VAS using metadata stored within the ESPAS system can be accessed very fast and are stable since no further dependencies on data provider platforms with special access requirements and/or safety policies have

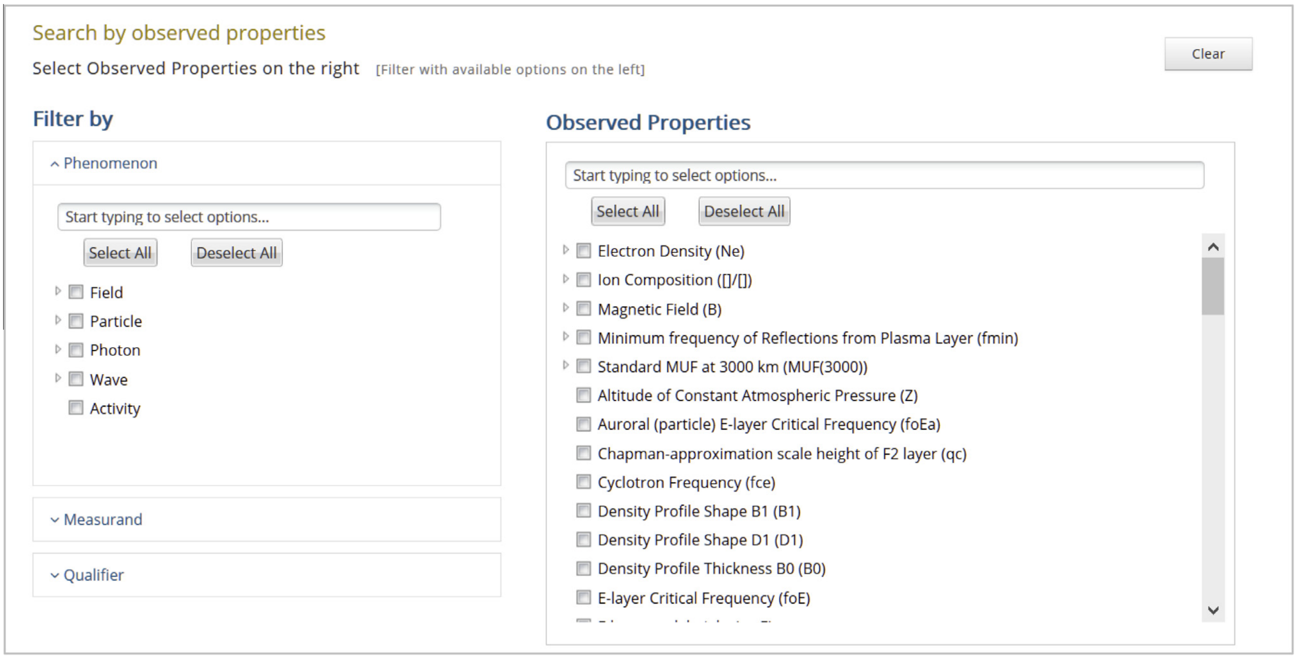

Fig. 7a. A screenshot from the ESPAS platform user interface showing the response of the "Metadata search query" when the user is searching by "Observed Properties". Without any further filtering.

to be taken into account. The complexity of VAS operations using metadata is low and once established, easy to maintain, but these services are limited in their capabilities.

Value Added Services based on real data are flexible and allow a broad range of possible supporting operations, but are difficult to maintain if the data operations are executed remotely at the different external data provider sites via the ESPAS infrastructure system as interface. To realize real time VAS within ESPAS with minimal maintenance and service efforts, the following two preferred concepts have been identified and demonstrated within the project:

(a) To retain the measuring values at least partially within a meta-data format using the sensor observation service (SOS), which allows to query real-time sensor data.
(b) To develop a JavaScript web-application tool, which allows to run the service client based and is therefore independent from system requirements.

In the ESPAS VAS demonstrator "TEC Time Series Plotter (TTSP)" a combination of extraction, scaling and comparison service is given using real data provided via the ESPAS platform.

The TTSP is a demonstrator for an ESPAS value added service generated by the German Aerospace Center (DLR) one of the ESPAS data providers. It allows to plot, analyse and download the time series of Total Electron Content (TEC) values and the corresponding range errors for different radio frequencies at selected locations worldwide for a given time period from 2D TEC maps provided by the Space Weather Application Center Ionosphere (SWACI). The temporal resolution of the plotted time series can be selected too, but depends on the maximal temporal resolution of the data downloaded via ESPAS. The path to the downloaded data has to be given in the data location field. One click at the progress bar starts the processor. The progress of the VAS data processing is visible via the progress bar. The generated time series plots can be downloaded as PNG and the corresponding data in JSON data file format.

Fig. 8 shows the graphical interface of the demonstrator, where the GPS L2 frequency range error at different location during the time period from 23rd February till 1st March 2014 is plotted in 15 min resolution. The time series shows the benefit of the VAS for data analysis. It supports the detection of space weather effects since the latitude dependent influence of a moderate geomagnetic storm on the ionosphere during 27–28th February is clearly visible.

The ESPAS user can download TEC map data and the latest version of the service tool from the ESPAS portal. Since the service uses only HTML and JavaScript technology, it is straightforward to integrate it into existing web services. The caveat is the implementation of data provision for this service. The easiest solution is a static data hosting next to the application. A standalone version of this service was built by using node-webkit, an app runtime based on Chromium browser and node.js, which enables writing and distributing native applications with the common web technologies supporting Windows and Linux. Finally the ESPAS user will have the opportunity to download an executable of this service next to the SWACI TEC maps which can be acquired from the ESPAS system. Since the app contains the web browser and all required HTML and JavaScript files, no installation is needed. The maintenance of such services in case of a data format change can be done directly at the data provider site, where the VAS

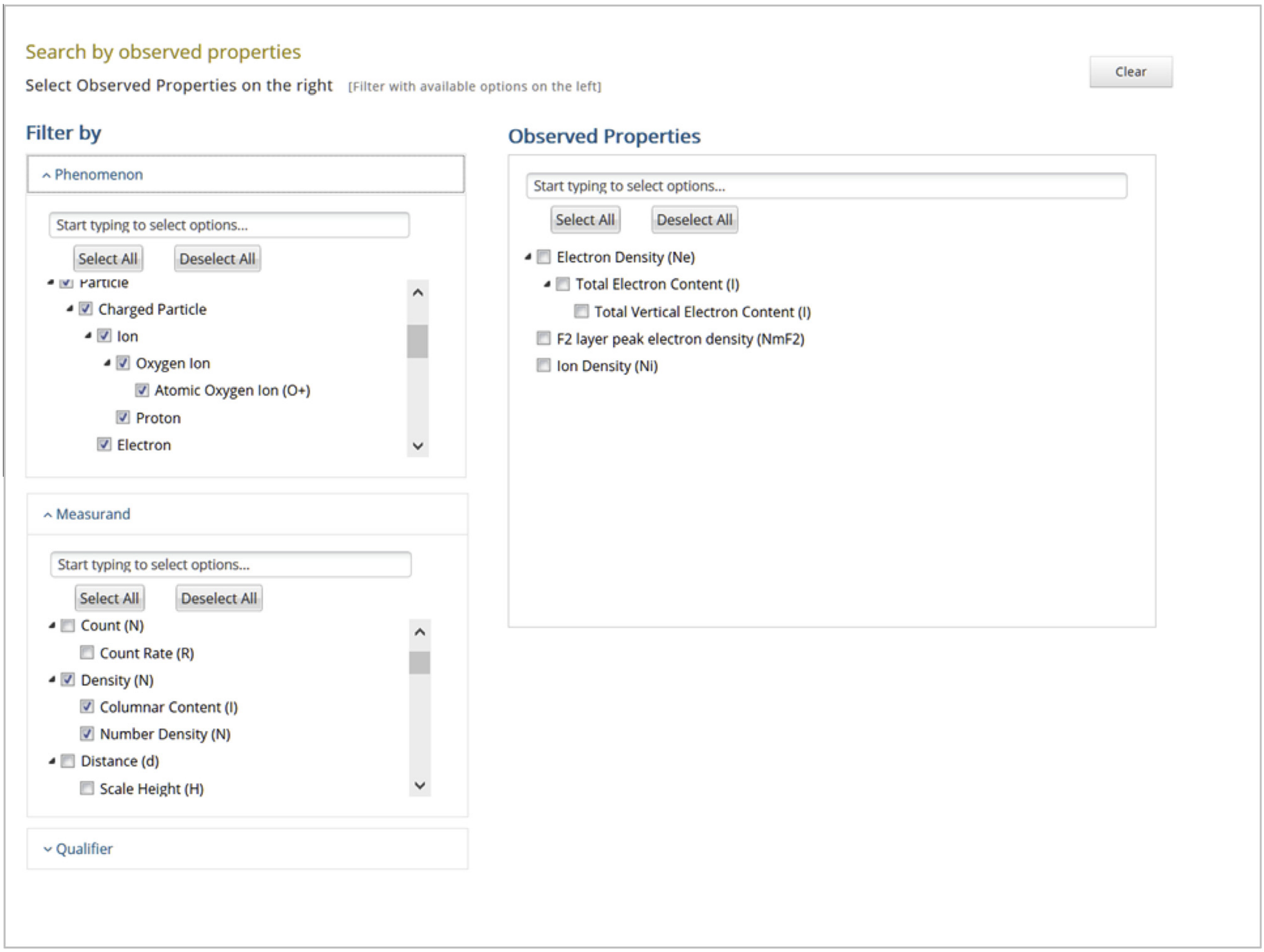

Fig. 7b. A screenshot from the ESPAS Platform User Interface showing the response of the "Metadata search query" when the user is searching by "Observed Properties", but further filtered by "Phenomenon" and "Measurand" components.

was generated. Therefore, no cost intensive and time consuming changes at the ESPAS system or communications with the administrators are needed. New real data service tools can be easily made available by data providers at the ESPAS portal.

7. Discussion

ESPAS aims at facilitating discovery and access to data from near-Earth space. Its basic principles are interoperability of the nodes holding heterogeneous data and expandability, to serve the evolving needs of the users.

To achieve interoperability, the basic and innovative achievement was the adoption and adaption of the ISO 19156 Observations and Measurements (O&M) standard data model to the requirement of near-Earth space data and the development of the ontology for near-Earth space data. As described in the introduction, the choice of the data model is not unique for systems working with space data. It depends very much on the specific region of space that the data are describing. SPASE (Space Physics Archive Search and Extract) model for example is widely used by the heliospheric community since it specifies terms and fields with which heliophysics data and related resources are to be described and provides for each of the terms, a list of valid values. However, for the full description of near-Earth space data, the solution of the ISO 19156 O&M standard data model has been considered much more flexible and inclusive since it provides the possibility for a full description of all heterogeneous data from the Sun down to the Earth's upper atmosphere. Furthermore, the ISO 19156 O&M standard data model provides an excellent basis for interoperability of ESPAS with other systems providing access to data from the Earth's environment such as the Infrastructure for Spatial Information in Europe (INSPIRE) and the Ocean Data Interoperability Platform (ODIP). It is worth noticed that a number of communities relevant to the ESPAS disciplines use ISO 19100 series geographic information standards. We can refer to the Australian Ocean Data Network Portal, the British Antarctic Survey Polar Data Centre, the British Geological Survey National Geoscience Data Centre, the Common Data Index Discovery Metadata Service, the GoGeo Portal, the Gulf of Mexico Research Initiative Information and Data Cooperative, and the World Meteorological Office Core Metadata Profile. This extensive use of the ISO 19100 series

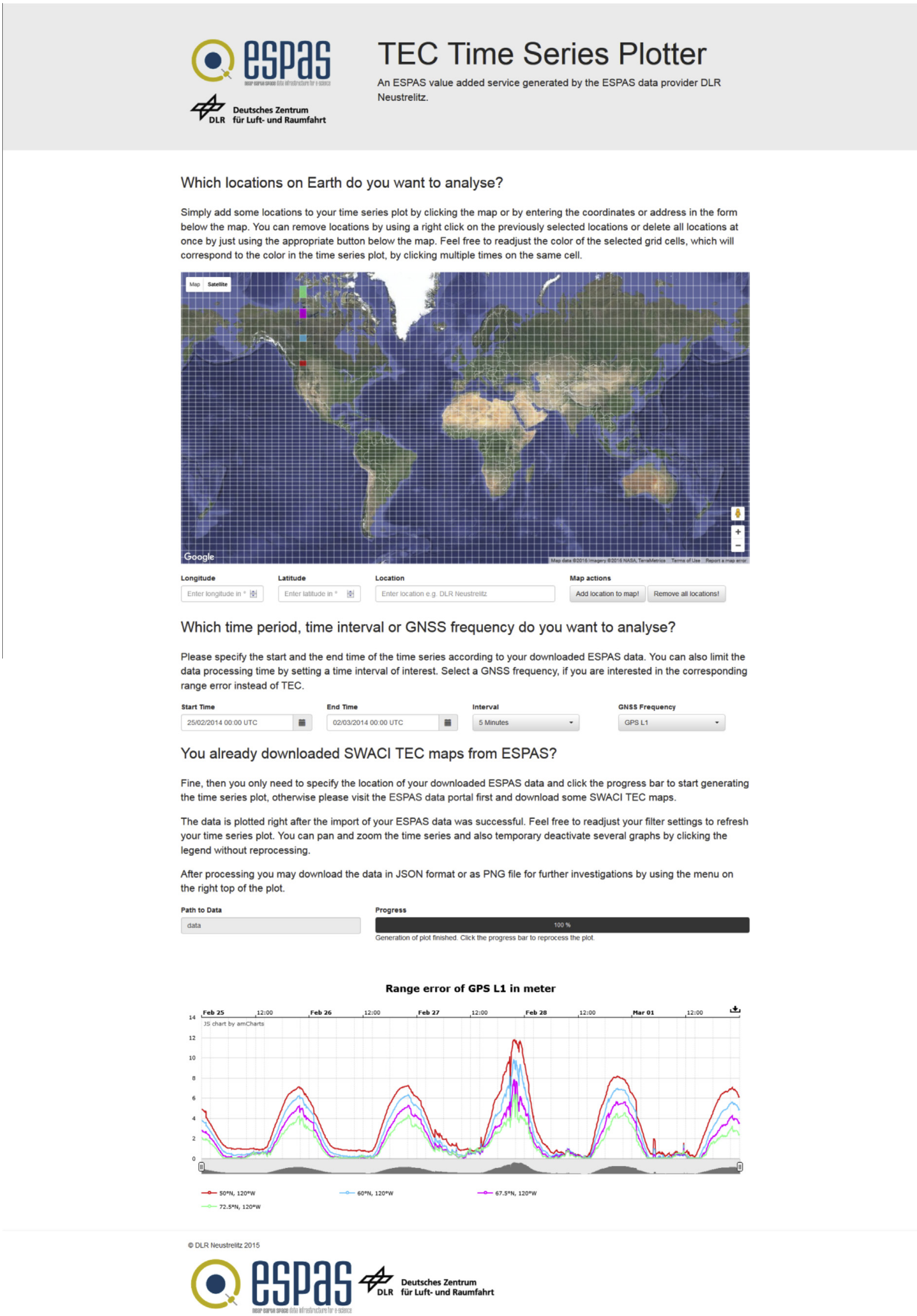

Fig. 8. Graphical interface of the TEC Time Series Plotter VAS showing the range error for the L2 GPS frequency during the time period from 23.02–01.03.2014 in 5 min resolution.

geographic information standards from the Earth's environmental and space communities gives excellent prospects for the expandability of the platform and its eventual integration to a global e-infrastructures ecosystem.

Furthermore, ESPAS expandability principles are driven by the users requests. Building the ESPAS system has meant also building up a community around ESPAS, fostering different scientific and cultural backgrounds and taking strength from them to convey the project aims. Moreover, a state of the art cutting-edge, system like ESPAS, aiming at being a point of reference for an entire scientific and technological community, must attract and include potential new users and stakeholder.

To fulfil the aforementioned scopes, proper community interaction activities were carried out along the project lifetime, relying on three mutually interacting pillars (Fig. 9):

(1) the establishment and consolidation of the ESPAS community;
(2) the dissemination and exploitation of the project results;
(3) the training of the ESPAS users.

The establishment and consolidation through time of the ESPAS community were achieved through different tools. One of these tools was the constant release of information through the ESPAS web portal (www.espas-fp7.eu), in its both parts: the public and restricted-access (wiki) one.

In particular, the setting and use of a dedicated wiki section, with access restricted to the project participants, was a powerful and fast tool to support both the administration activities of the project, and the spreading of information and participation to the ESPAS progress within the whole ESPAS community.

The public part of the portal (http://www.espas-fp7.eu), regularly updated by all ESPAS partners with all the latest news, events and development of the ESPAS related issues and the production of dissemination material (leaflet, brochures, papers, etc.) and of guidelines for the ESPAS portal (data model and ontology description, system manual, public reports etc), were of support both for the establishment of the ESPAS community, and for the dissemination of the project scopes and results to the stakeholders.

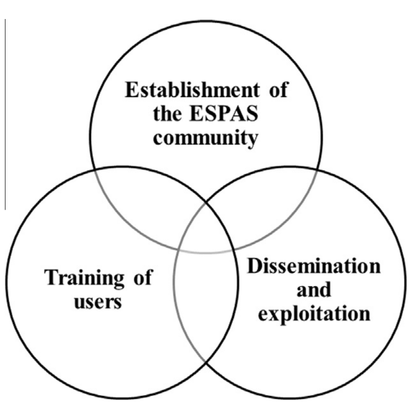

Fig. 9. The three pillars of the ESPAS community building.

The dissemination activities have been carried by organizing of ESPAS special sessions and fairs in the European Space Weather Weeks (ESWW), by delivering presentations in all the major scientific conferences and workshops to make aware the related communities about the possibilities offered by the use of ESPAS platform.

The involvement of the stakeholders in the assessment of ESPAS has been organized in the format of the Coordinated Data Analysis Workshops (CDAWs), in which selected internal and external experts have been invited to participate. The aim of the CDAW's was to ensure the easiness of use of ESPAS, to meet the real needs of scientists, engineers and operators and to give feedbacks to the ESPAS system developers.

The training aspect of the community interaction strategy has converged into the organization of a 5-days training school, held in November 2015. Such school allowed to present the ESPAS system to PhD students and early career scientists, to let them develop the competence to exploit the ESPAS e-infrastructure and to interface it into their own operating systems. The school was tuned on the use of ESPAS based on specific use cases, but it also included lectures and tutorials concerning the physical problems that can be further explored with the ESPAS system, together with insights into the system technical features, architecture, data model and ontology. Big emphasis has been given to the practical sessions, in which students were asked to experiment with the system functionalities, and to prepare case studies using ESPAS resources in complementarity with other systems such as HELIO and AMDA, in order to address real science problems.

8. Outlook

ESPAS started in 2011 as an idea and it has been evolved to a mature system with high technological readiness level, demonstrating a stable and robust system operation in the academic environment. For the years to come, the ESPAS consortium has committed to continue its activities with emphasis on the following priorities:

- Maintenance of the access to ESPAS platform and to data nodes of the e-infrastructure;
- Maintenance, upgrade and evolution of the ESPAS platform and the data nodes;
- Quality control and validation of ESPAS services;
- Promotion of scientific activities linked to bring in more data; Networking activities to develop links with European and international projects and organizations;
- Dissemination activities to raise awareness of different users about ESPAS functionalities and specification of new services tailored to users' needs;
- Performance of RTD activities for the release of new services;

- Support and training activities especially towards students and young researchers.

The legacy of ESPAS is the standardization of metadata and data, the organized access to data, the protection of Intellectual Property Rights of the data providers and the operation of a mature and stable e-infrastructure able to serve a number of multidisciplinary communities. The major challenge for the years to come is to work more closely with the industrial and operational sector to develop new services able to meet their specific needs. It is important to promote multi-domain community-driven approaches to fully exploit the ESPAS core e-infrastructure services with high economic innovation potential, but also aware the commercial sector of the potential of services provided by e-infrastructures, providing the support to become active innovators as users and or suppliers of e-infrastructures. These stronger links between e-infrastructure operators and other actors in the innovation chain will pave the way to the long-term sustainability of ESPAS and of relevant e-infrastructure systems.

Acknowledgements

The ESPAS project is funded through EU-FP7 under grant agreement 283676.

Appendix A. Appendix 1

A.1. ESPAS observation collections

A detailed list of the observation collections that are registered in ESPAS, or are under registration, is given in Tables A1–A3.

Table A1
Near-Earth space datasets from ground-based experiments and related applications.

1. Radars	
1. 1 Coherent radar	
SuperDARN	Observed properties • Electric potential [resolution: 40 × 40 grid, 2 min cadence] ○ MLT grid – *SuperDARN (ULEIC)* ○ 49 Spherical harmonic coefficients – *SuperDARN (ULEIC)* ○ Associated information – *SuperDARN (ULEIC)*
1.2 Incoherent scatter radar	
EISCAT	Observed properties [coverage: sporadic, 1000 km range, resolution: few Km^3, 1 min] • Line-of-sight profiles (Standard ISR and derived parameters) – *EISCAT ISR (EISCAT)*
2. Receivers	
2.1 GNSS	
NOA GNSS	Observed properties [receiver location: Athens, Greece] • Raw data (RINEX files) – *Athens GNSS (NOA)* • TEC ○ Local estimates – *Athens GNSS (NOA)*
GFZ GNSS	Observed properties [sample rate 1 Hz (1 h-files per station)] • Raw data (RINEX files) – *CHAMP: High rate GPS ground tracking data (GFZ – ISDC)*
INGV GNSS	Observed properties [receiver location: Chania (Greece – 35.51°N, 24.02°E)] • Raw data (RINEX files) [time resolution: 60 min (since June 2007)] – *CHA1 GPS RINEX data (INGV)* • Scintillation indices [time resolution: 15 min (since June 2007)] – *CHA1 GPS Scintillation and TEC Parameters (INGV)* ○ TEC – Local estimates [time resolution: 60 min (since 2007)] – *CHA1 GPS Scintillation and TEC raw data (INGV)* – *CHA1 GPS Scintillation and TEC Parameters (INGV)*

(*continued on next page*)

Table A1 (*continued*)

DLR GNSS	Observed properties • TEC ○ Maps/Grids ■ European [spatial/time resolution: 2,5 × 2,5°, <1 h] – *SWACI (DLR)* ■ Global [spatial/time resolution: 5 × 5°, <1 h] – *SWACI (DLR)* • TEC derivatives ○ Latitudinal gradients ■ European maps [spatial/time resolution: 2,5 × 2,5°, <1 h] – *SWACI (DLR)* ■ Global maps [spatial/time resolution: 5 × 5°, < 1 h] – *SWACI (DLR)* ○ Longitudinal gradients ■ European maps [spatial/time resolution: 2,5 × 2,5°, <1 h] – *SWACI (DLR)* ■ Global maps [spatial/time resolution: 5 × 5°, <1 h] – *SWACI (DLR)* ○ Rate of change ■ European maps [spatial/time resolution: 2,5 × 2,5°, <1 h] – *SWACI (DLR)* ■ Global maps [spatial/time resolution: 5 × 5°, <1 h] – *SWACI (DLR)* • TEC Median ○ European maps [spatial/time resolution: 2,5 × 2,5°, <1 h] – *SWACI (DLR)* ○ Global maps [spatial/time resolution: 5 × 5°, <1 h] – *SWACI (DLR)* • Scintillation indices [time resolution: <1 h] – *SWACI (DLR)* • Slab Thickness [location/time resolution: Over Juliusruh and Tromso, <1 h] – *SWACI (DLR)*
2.2 Beacon	
SGO beacon	Observed properties [height range 0–700 km, latitudes: 60–70°N] • Electron density ○ Tomography reconstructed matrix ■ Chapman regularization – *Finnish Ionospheric tomography (SGO – Uoulu)* ■ IRI regularization – *Finnish Ionospheric tomography (SGO – Uoulu)*
3. Sounders	
3.1 Ionosonde	
Athens digisonde	Observed properties [range: 500 km around Athens, time resolution: 15 min (or higher under special campaigns)] • Ionograms ○ Raw data – *Athens Digisonde Data (NOA)* ○ Images – *Athens Digisonde Data (NOA)* • Ionospheric characteristics (autoscaled) – *Athens Digisonde Data (NOA)* • Electron density (profiles) – *Athens Digisonde Data (NOA)* • Ionospheric plasma drifts ○ Skymaps – *Athens Digisonde Data (NOA)* ○ Drift velocities (data files and daily plots) – *Athens Digisonde Data (NOA)*

Table A1 (*continued*)

DIAS network	Observed properties • Ionograms (Images) [DIAS locations, time resolution: 15 min] – DIAS (NOA) • Ionospheric characteristics ○ Autoscaled [DIAS locations, time resolution: 15 min] – *DIAS (NOA)* ○ Predictions [DIAS locations, time resolution: 15 min] – *DIAS (NOA)* ○ Maps [the European Region from −5°W to 40°E (longitude) and from 32°N to 60°N (latitude)] ▪ Nowcasting [time resolution: 15 min] – *DIAS (NOA)* ▪ Long-term predictions [from 1 to 3 months ahead] – *DIAS (NOA)* ▪ Short-term predictions [from 1 to 24 h ahead] – *DIAS (NOA)* • Electron density ○ Profiles [DIAS locations, time resolution: 15 min] – *DIAS (NOA)* ○ Maps (nowcasting) [the European Region from -5°W to 40°E and from 32°N to 60°N, time resolution: 15 min] – *DIAS (NOA)* • Ionospheric alerts – DIAS (NOA)
Rome digisonde	Observed properties [digisonde location: Rome (41.8°N, 12.5°E)] • Ionograms [time resolution: 15 min (from 1998)] ○ Raw data – *Rome raw Ionograms – DPS4 (INGV)* ○ Images – *Rome Ionogram images – DPS4 (INGV)* • Ionospheric characteristics (Autoscaled) [time resolution: 15 min (from 1998)] – *Rome ionogram autoscaled data – DPS4 (INGV)* – *Rome Ionogram images – DPS4 (INGV)* • Electron density (Profiles) [time resolution: 1 hour (from 1976)] – *Rome ionogram autoscaled data – DPS4 (INGV)*
Rome AIS.INGV	Observed properties [ionosonde location: Rome (41.8°N, 12.5°E), time resolution: 15 min (from 2005)] • Ionograms ○ Raw data – *Rome raw Ionograms – AIS.INGV (INGV)* ○ Images – Rome Ionogram images – AIS.INGV (INGV) • Ionospheric characteristics (Autoscaled) – *Rome ionogram autoscaled data – AIS.INGV (INGV)* – *Rome Ionogram images – AIS.INGV (INGV)* • Electron density (Profiles) – *Rome ionogram autoscaled data – AIS.INGV (INGV)*
Gibilmanna AIS. INGV	Observed properties [ionosonde location: Gibilmana (37.9°N, 14.0°E), time resolution: 15 min (from 2002)] • Ionograms ○ Raw data – *Gibilmanna raw Ionograms – AIS.INGV (INGV)* ○ Images – *Gibilmanna Ionogram images – AIS.INGV (INGV)* • Ionospheric characteristics ○ Autoscaled – *Gibilmanna ionogram autoscaled data – AIS.INGV (INGV)* – *Gibilmanna Ionogram images – AIS.INGV (INGV)* ○ Manually scaled – *Gibilmanna ionogram validated data – AIS.INGV (INGV)* • Electron density (Profiles) – *Gibilmanna ionogram autoscaled data – AIS.INGV (INGV)*

(*continued on next page*)

Table A1 (*continued*)

GIRO	Observed properties [location: at 60 + ground GIRO locations, typical cadence of 5 min, some older locations still using 15 min cadence] • Ionograms (Images) – *GIRO Ionogram display products (GIRO)* • Ionospheric characteristics (Autoscaled) – *GIRO Ionogram display products (GIRO)* – *GIRO Ionogram-derived numerical data (GIRO)* • Ionogram echo traces – *GIRO Ionogram display products (GIRO)* – *GIRO Ionogram-derived numerical data (GIRO)* • Electron density (Profiles) – *GIRO Ionogram display products (GIRO)* – *GIRO Ionogram-derived numerical data (GIRO)* • Ionospheric Plasma Drifts ○ Skymaps – *GIRO Doppler skymap display products (GIRO)* ○ Drift velocities (data files and daily plots) – *GIRO Plasma Drift display products (GIRO)* – *GIRO Plasma Drift numerical data (GIRO)*
SGO ionosonde	Observed properties [ionosonde location: at 67°22’N, 26°38’E] • Ionograms ○ Raw data – *Ionosonde of SGO (SGO – Uoulu)* ○ Matrices [Time resolution: 1 hour (since 1957)] – *Ionosonde of SGO (SGO – Uoulu)* • Ionospheric characteristics [Time resolution: 10 min (since 2005) and 1 min (since 2007)] – *Ionosonde of SGO (SGO – Uoulu)*
Warsaw ionosonde	Observed properties [Typical cadence: 15 min] • Ionograms (images) – *Warsaw automatically scaled ionograms (SRC – PAS)* – *Warsaw manually scaled ionograms (SRC – PAS)* • Ionospheric characteristics ○ Manually scaled – *Warsaw manually scaled ionograms (SRC – PAS)* ○ Autoscaled – *Warsaw automatically scaled ionograms (SRC – PAS)* • Electron density (profiles) – *Warsaw automatically scaled ionograms (SRC – PAS)* – *Warsaw manually scaled ionograms (SRC – PAS)*
Hornsund ionosonde	Observed properties [Typical cadence: 15 min] • Ionograms (images) – *Hornsund automatically scaled ionograms (SRC – PAS)* • Ionospheric characteristics (autoscaled) – *Hornsund automatically scaled ionograms (SRC – PAS)* • Electron density (profiles) – *Hornsund automatically scaled ionograms (SRC – PAS)*
3.2 Dynasonde	
3.2.1 EISCAT dynasonde	Observed properties [Time resolution: 6 min] • Standard vertical sounding and derived parameters – *Dynasonde (EISCAT)*
3.3 Oblique Sounding	
Inskip – Rome	Observed properties [Time resolution: 15 min (Available dates under special campaigns)] • Ionograms (Raw files) – *Oblique ionograms Inskip-Rome (INGV)*

Table A1 (*continued*)

Inskip – Chania	Observed properties [Time resolution: 15 min (Available dates under special campaigns)] • Ionospheric characteristics (MUF, fmin) (manually scaled) – *Observed Maximum Usable Frequencies – of the radio link Inskip-Chania (INGV)*
4. Ground-based magnetometers	
DTU array	Observed properties [Spatial resolution: locally fixed (200–1000 km distance in Greenland), time resolution: 1–20 s] • Geomagnetic field (vector – variations) – *DTU ground magnetometer array (DTU)*
NORMAG	Observed properties [magnetometer locations: 14 sites in Norway, time resolution: 10 s averages (since 1986)] • Geomagnetic field (vector – variations) – *NORMAG – TGO magnetometer array (TGO)*
INGV	Observed properties [Observatories: Castello Tesino, L' Aquila and Lampedusa, time resolution: 1 s, digital data available for: 1-min, 1-h, 1-month, 1-year] • Geomagnetic field ○ Vector – *Definitive values of the geomagnetic field elements (INGV)* ○ Vector (variations) – *Variations of the geomagnetic field vector along its elements (INGV)* ○ Total intensity – *Geomagnetic field total intensity (INGV)*
IMAGE	Observed properties [magnetometer locations: 30 stations, time resolution: 10 s] • Geomagnetic field (vector – variations) – *IMAGE magnetometer network (FMI)*
Finish magnetometer chain	Observed properties • Geomagnetic field (pulsation) – *Finnish pulsation magnetometer chain (SGO-Uoulu)*
5. Riometers	
Finnish network	Observed properties [European high latitudes, time resolution: 1 min absorption values from widebeam (60°) riometers at fixed frequencies (30.0, 32.4, 51.4 MHz) calculated from 10 s raw data] • Ionospheric absorption – *Finnish riometer network (SGO-Uoulu)*
6. Neutron monitors	
Oulu NM	Observed properties [time resolution: 1 min] • Cosmic ray ○ Cosmic ray (Count rate) – *Oulu NM cosmic ray (SGO – Uoulu)* ○ Barometric Pressure – *Oulu NM cosmic ray (SGO – Uoulu)*
7. Indices	
Observed properties • Solar indices (International Sunspot number) – *International Sunspot number (ROB)*	
8. FPIs	
UCL	Observed properties [FPIs: time series with some spatial coverage (8 points around a circle centred on the FPI, SCANDI: time series and spatial coverage (variable nos of points in a 500 km radius circle)] • Thermospheric neutral wind (Vector) – *FPIs (UCL)* • Thermospheric temperature – *FPIs (UCL)*

Table A2
Near-Earth space datasets from space based experiments and related applications.

1. Radio occultation data	
CHAMP	Observed properties • Raw data (RINEX files) – *High rate CHAMP GPS-SST (CH-AI-1-HR) (GFZ – ISDC)* [sample rate 50 Hz] – *Medium rate CHAMP GPS-SST data (CH-AI-1-MR) (GFZ – ISDC)* [sample rate 1 Hz] • Atmospheric excess path delay – *Atmospheric excess path delay (CH-AI-2-PD) (GFZ – ISDC)* • Occultation tables – *Occultation tables (CH-AI-2-TAB) (GFZ – ISDC)* [Tables of the occultation events, per day] • Atmospheric parameters (Vertical profile) – *Atmospheric parameters (CH-AI-3-ATM) (GFZ – ISDC)* • Electron density ○ Vertical profile – *Vertical electron density profiles (CH-AI-3-IVP) (GFZ – ISDC)* – *SWACI (DLR)* [Daily update] • TEC (Occultation Link related) – *Occultation link related TEC data (CH-AI-3-TCR) (GFZ – ISDC)*
TerraSAR-X	Observed properties • Raw data – *TerraSAR-X GPS-SST occultation data (TSX-OCC-1-HR) (GFZ – ISDC)* [sample rate of 0.02 sec (=50 Hz)] – *TerraSAR-X GPS-SST occultation data (TSX-OCC-1-MR) (GFZ – ISDC)* [sample rate of 1.0 sec]
2. Solar data	
SOHO/LASCO	Observed properties • Coronographic white-light images (Outer solar corona images) – *SOHO/LASCO (ROB)*
SOHO/EIT	Observed properties • Coronal EUV images (Full disk images) – *SOHO/EIT (ROB)*
PROBA2/SWAP	Observed properties [FOV = 54′, spatial resolution 3.16″, Wavelength = 17 nm, temporal cadence ~ 2 min] • Coronal EUV images (Full disk images) – *PROBA2/SWAP (ROB)*
PROBA2/LYRA	Observed properties • UV and EUV solar irradiance – *PROBA2/LYRA (ROB)*
3. Plasma sounding/monitoring	
IMAGE/RPI	Observed properties [Spatial information: elliptical polar orbit (apogee 45,922 km, perigee 1000 km, 13.5 h period). Data coverage: selected times during mission period from May 2000 to Dec 2005] • Plasmagrams (Images) [Temporal information: typical plasmagram cadence of 5 minutes, typical measurement duration of 1 minute] – *Plasmagram display products (GIRO)* • Plasma frequencies [Temporal information: typical plasmagram cadence of 5 minutes, typical measurement duration of 1 minute] – *Plasmagram display products (GIRO)* – *Plasmagram-derived data products (GIRO)* • Echo traces [Temporal information: typical plasmagram cadence of 5 minutes, typical measurement duration of 1 minute] – *Plasmagram display products (GIRO)* – *Plasmagram-derived data products (GIRO)* • Electron density (Profiles along the signal propagation path) [Temporal information: typical plasmagram cadence of 5 minutes, typical measurement duration of 1 minute] – *Plasmagram display products (GIRO)* – *Plasmagram-derived data products (GIRO)* • Spectograms (images) [Temporal information: typical spectrogram measurement cadence of 3 minutes, typical measurement duration of 1 minute] – *Spectrogram display products (GIRO)*

Table A2 (*continued*)

CLUSTER/WHISPER	Observed properties [Space locations: magnetosphere and solar wind, Frequency range: 2–80 kHz, Frequency resolution: 163 Hz, Density range: 0.05–80 cm^{-3}, Time resolution: 2 s] • Electron density – *CLUSTER/WHISPER density (BISA)*

CLUSTER/Cluster Active Archive	Observed properties
	[Coverage: 2001/02 ~ 50% from 2003% ~ 100%. Cadence is typically at spin resolution (~4 s) with some instruments (particularly fields) providing much higher sampling rates. Duration of the mission ~ 10 years]
	• Magnetic field
	  – *Cluster Active Archive (ESA – CAA)*
	• Electric field
	  – *Cluster Active Archive (ESA – CAA)*
	• Magnetic waves
	  – *Cluster Active Archive (ESA – CAA)*
	• Electric waves
	  – *Cluster Active Archive (ESA – CAA)*
	• Plasma electrons and ion composition
	  ○ few eV to ~10 keV
	    – *Cluster Active Archive (ESA – CAA)*
	  ○ 40 keV to few hundred keV
	    – *Cluster Active Archive (ESA – CAA)*

ACE/SWEPAM	Observed properties
	[Interplanetary medium: L1 point, Daily files, Time resolution: 1 min, Data latency: 5 min]
	• Solar wind proton density
	  – *ace_swepam_1m (DH consultancy)*
	• Bulk speed
	  – *ace_swepam_1m (DH consultancy)*
	• Ion temperature
	  – *ace_swepam_1m (DH consultancy)*

ACE/MAG	Observed properties
	[Interplanetary medium: L1 point, Daily files, Time resolution: 1 min, Data latency: 5 min]
	• IMF components
	  – *ace_mag_1m (DH consultancy)*
	• IMF total field strength
	  – *ace_mag_1m (DH consultancy)*

DEMETER	Observed properties
	[Spatial information: Geographical longitude range: -180:180°, Geographical latitude range: −60:60°, Temporal information: July 2004–December 2010 (each map is one month average)]
	• Particle density
	  ○ Electron (maps)
	    – *Demeter maps (SRC – PAS)*
	  ○ Ion (maps)
	    – *Demeter maps (SRC – PAS)*
	• Particle temperature
	  ○ Electron (maps)
	    – *Demeter maps (SRC – PAS)*
	  ○ Ion (maps)
	    – *Demeter maps (SRC – PAS)*
	• Drift velocity (maps)
	  – *Demeter maps (SRC – PAS)*

Magion-3	Observed properties
	[Spatial information: Geographical longitude range: −180:180°, Geographical latitude range: −83:83°, Spatial resolution of sigle measurements m-km, Altitude 440–3070 km, Temporal information: December 1991–August 1992, time resolution of singe measurements s-ms]
	• Electron density (maps)
	  – *Magion-3 (SRC – PAS)*
	• Electron temperature (maps)
	  – *Magion-3 (SRC – PAS)*
	• HF emission (maps)
	  – *Magion-3 (SRC – PAS)*

(*continued on next page*)

Table A2 (*continued*)

• HF plasma instabilities and electromagnetic emissions – *Magion-3 (SRC – PAS)*	
FORMOSAT-3/COSMIC	Observed properties [Spatial information: Geographical longitude range: -180:180°, Geographical latitude range: −85:85°, 2000–2500 osculation points per day which correspond to 200 km spatial resolution, Temporal information: January 2007 till now, 2 h – 1 day time resolution] • Electron density (profiles) – *COSMIC (SRC – PAS)*
4. Energetic particle data	
NOAA/POES:MEPED	Observed properties [Averaged fluxes at 102 min orbit period resolution, including time, and orbit plane MLT orientation, from 11 NOAA/POES satellites since 1979] • Energetic particle fluxes ◦ Electrons (corrected and calibrated) (From 30 keV to 2.5 MeV) – *NOAA/POES MEPED (Uoulu)* ◦ Protons (corrected and calibrated) (30 keV upwards) – *NOAA/POES MEPED (Uoulu)*

Table A3
Models.

Models	
International reference ionosphere (IRI)	Modeled properties • foF2 (global grids) – IRI foF2 grids – CCIR F peak model This collection contains the global grids (0 to 360° in longitude and −90 to 90° in latitude) of the foF2 critical frequency predicted by IRI (International Reference Ionosphere) model using the CCIR (Comite Consultatif International des Radiocommunications) F peak model. The foF2 predictions were obtained through the IRI-2012 version of the model available for online computations at http://irimodel.org. The grids are provided with hourly time resolution in ASCII format.
International reference ionosphere (IRI)	Modeled properties foF2 (global grids) – IRI foF2 grids – URSI F peak modelThis collection contains the global grids (0 to 360° in longitude and −90 to 90° in latitude) of the foF2 critical frequency predicted by IRI (International Reference Ionosphere) model using the URSI (International Union of Radio Science) F peak model. The foF2 predictions were obtained through the IRI-2012 version of the model available for online computations at http://irimodel.org. The grids are provided with hourly time resolution in ASCII format.
Simplified ionospheric regional model updated in real-time (SIRMUP)	Modeled properties • foF2 (grids for the European region) – DIAS SIRMUP nowcasting maps of foF2 This collection contains the ionospheric maps of the foF2 critical frequency provided by DIAS for the European region (from −5 to 40° E in longitude and from 34 to 60° N in latitude). The maps are based on the Simplified Ionospheric Regional Model Updated in real-time (SIRMUP) with autoscaled ionospheric parameters obtained by DIAS Digisondes. The maps are provided as images (PNG format) as well as numerical grids (ASCII format).
Simplified ionospheric regional model updated in real-time (SIRMUP)	Modeled properties • M(3000)F2 (grids for the European region) – DIAS SIRMUP nowcasting maps of M(3000)F2 This collection contains the ionospheric maps of the M(3000)F2 propagation factor provided by DIAS for the European region (from −5 to 40° E in longitude and from 34 to 60°N in latitude). The maps are based on the Simplified Ionospheric Regional Model Updated in real-time (SIRMUP) with autoscaled ionospheric parameters obtained by DIAS Digisondes. The maps are provided as images (PNG format) as well as numerical grids (ASCII format).
Electron density assimilative model (EDAM)	Modeled properties [Derived electric potential on a 40x40 grid evenly spaced on a pole-centred MLT-latitude grid such that at midnight and noon, dawn and dusk the grid extends to 5° latitude, Cadence: 2 min] • Electron density (3D grids) – *EDAM (UBIRM)*

Table A3 (*continued*)

CMAT2	Modeled properties • Neutral parameters (3D grids) – *CMAT2 (UCL)* • Ionospheric parameters – *CMAT2 (UCL)*

Appendix B. Appendix 2

B.1. Controlled vocabularies used in ESPAS data model

Many of the properties of the concepts of ESPAS Data Model use values from controlled vocabularies. As described in Section 4, the ESPAS data portal manages a vocabulary of space physics keywords that can be used to narrow down data searches to observations of specific physical content. A list of these controlled vocabularies used in ESPAS Data Model is given, along with a short definition:

- *Component:* For vector properties, it describes which of the three components is provided in the data only in those cases when observation does not specify the vector property in full. Typical components are X, Y, Z. The Component has to be accompanied by a suitable description of the Coordinate Reference System (Crs).
- *Compressed Representation:* Describes the formalism of compressed representation of voluminous or complex 3D, 2D, and 1D data. Typical examples are spherical harmonics for 2D maps on the sphere, truncated Fourier transforms (harmonics) for diurnal time series, Empirical Orthogonal Functions (EOF).
- *Crs:* Corresponds to the Coordinate Reference Systems (e.g. GSE, GSM, ..) used to describe a vector Component (observed property definition), the location of a Platform and the geographic extent of an Observation.
- *Dimensionality Instance:* Dimensionality is a compact description of the domain X spanned by the independent (input) variables x1, x2, x3... of the Observation result (output dependent variable Y): $Y = f(x1, x2, x3...)$. The independent variables x1, x2, x3... are tested in the course of the Observation to acquire values of the dependent variable Y. For example, an Observed Property "NeutralWindVelocity" is a vector field variable with a natural presentation as a Vector (magnitude and direction) defined in 3D space (latitude, longitude, altitude). The Dimensionality Instance describes the Single instance of the acquired Y values of the observed property in time (time is not included in the list of independent variables) (e.g. 1D.point, 1D.Profile, 2D.Map, 2D.image).
- *Dimensionality Timeline:* Dimensionality is a compact description of the domain X spanned by the independent (input) variables x1, x2, x3... of the Observation result (output dependent variable Y): $Y = f(x1, x2, x3...)$. The independent variables x1, x2, x3... are tested in the course of the Observation to acquire values of the dependent variable Y. For example, an Observed Property "NeutralWindVelocity" is a vector field variable with a natural presentation as a Vector (magnitude and direction) defined in 3D space (latitude, longitude, altitude). The Dimensionality Timeline describes the timeline of the acquired Y values of the observed property (time is one of the independent variables) (e.g. Timeseries, Animation).
- *Licence:* It is the element of an agreement describing the terms under which data registered in ESPAS can be used.
- *Platform Type:* Describes the type of a Platform (e.g. ground-based station, satellite).
- *Projection:* For vector properties, it describes a plane or a line on which the vector projection is observed by the Instrument. The Projection is provided in the data only in those cases when Observation does not specify the vector property in full. Typical projections are horizontal, line of sight, orbital, perpendicular. The Projection has to be accompanied by a suitable description of the Coordinate Reference System (crs).
- *Related Observation Role:* Describes the role of the related Observation (e.g. location information...).
- *Related Party Role:* Describes the role (owner, principal investigator, researcher, etc) of a related party for an object (Project, Observation Collection, Instrument etc).
- *Result Accumulation:* Describes the frequency with which additions are/were made to the Observation's result (e.g. daily, monthly, hourly...).
- *Result Data Format:* Describes the data format of a resulting file of an Observation.
- *Service Function:* Describes the function of a service offered by a Data Provider at the Observation Result level. So, it specifies whether the Observation Result files (data files) are available for download or for view only from the end user.
- *Status:* Describes the status of a Project, an Observation, an Operation of a Platform (e.g. historical, ongoing,..).
- *Unit:* Describes the unit of the Observed Property as measured in a specific process (e.g. km, MHz).

For a detailed view of the vocabularies, one should visit the Browse → ESPAS Supplementary Vocabularies of the ESPAS Portal (https://www.espas-fp7.eu/portal/).

References

Belehaki, A., Hapgood, M., Manola, N., Ventouras, S., James, S., Athanasopoulos, G., Lempesis, A., Marziou, S., Watermann, J., and the ESPAS team, 2014. THE ESPAS E-INFRASTRUCTURE: ACCESS TO DATA FROM NEAR-EARTH SPACE, 350, Proc. of the 2014 conference on Big Data from Space (BiDS'14) 2014, European Space Agency-ESRIN, Frascati, Italy, 12–14 November, pp. 350–353. http://dx.doi.org/10.2788/1823.

Bentley, R.D., Csillaghy, A., Aboudarham, J., Jacquey, C., Hapgood, M. A., Bocchialini, K., Messerotti, M., Brooke, J., Gallagher, P., Fox, P., Hurlburt, N., Roberts, D.A., Duarte, L. Sanchez, 2010. HELIO: The Heliophysics Integrated Observatory. Adv. Space Res. 47 (12), 2235–2239. http://dx.doi.org/10.1016/j.asr.2010.02.006.

Bröring, A. et al., 2012. OGC Sensor Observation Service Interface Standard v2.0.0, Open Geospatial Consortium Standard, OGC 12-006.

Cox, 2006. Simon, "Observations and measurements." Open Geospatial Consortium Best Practices Document. OGC.

Doyle, A. et al., 2001. Introduction to OGC Web Services. OGC-Open Geospatial Consortium, OGC White Paper.

Fowler, M., 1998. Analysis Patterns: reusable object models. Addison Wesley Longman, Menlo Park, CA.

Hayashi, H., Koyama, Y., Hori, T., Tanaka, Y., Abe, S., Shinbori, A., Kagitani, M., Kouno, T., Yoshida, D., UeNo, S., Kaneda, N., Yoneda, M., Umemura, N., Tadokoro, H., Motoba, T.IUGONET project team, 2013. Inter-university Upper Atmosphere Global Observation NETwork (IUGONET). Data Sci. J. 12, WDS179-WDS184, link.

ISO 19156:2011 Observations and measurements, http://www.iso.org/iso/catalogue_detail.htm?csnumber=32574, OGC Observations and measurements Schema, http://www.opengeospatial.org/standards/om.

Jacquey, C., Génot, V., Budnik, E., Hitier, R., Bouchemit, M., Gangloff, M., Fedorov, A., Cecconi, B., André, N., Lavraud, B., Harvey, C., Dériot, F., Heulet, D., Pallier, E., Penou, E., Pinçon, J.L., 2010. AMDA, Automated Multi-Dataset Analysis: A Web-Based Service Provided by the CDPP. In: Laakso, H., Taylor, E., Escoubet, M.E., Philippe, C. (Eds.), The Cluster Active Archive, Astrophysics and Space Science Proceedings. Springer, Netherlands, pp. 239–247. http://dx.doi.org/10.1007/978-90-481-3499-1_16.

Manghi, P. et al., 2014. The D-NET software toolkit: a framework for the realization, maintenance, and operation of aggregative infrastructures. Program: Electronic Library and Information Systems 48 (4), 322–354.

Merka, J., Narock, T.W., Szabo, A., 2008. Navigating through SPASE to heliospheric and magnetospheric data. Earth Sci. Inform. 1, 35–42. http://dx.doi.org/10.1007/s12145-008-0004-5.

Nebert, D., et al., 2007. OpenGIS® Catalogue Services Specification v2.0.2, Open Geospatial Consortium Standard, OGC 07–006r1.

Pérez-Suárez, D., Maloney, S.A., Higgins, P.A., Bloomfield, D.S., Gallagher, P.T., Pierantoni, G., Bonnin, X., Cecconi, B., Alberti, V., Bocchialini, K., Dierckxsens, M., Opitz, A., Le Blanc, A., Aboudarham, J., Bentley, R.B., Brooke, J., Coghlan, B., Csillaghy, A., Jacquey, C., Lavraud, B., Messerotti, M., 2012. Studying sun-planet connections using the heliophysics integrated observatory (HELIO). Sol. Phys. 280 (2), 603–621.

Rissanen, E., 2013. eXtensible Access Control Markup Language (XACML) v3.0. OASIS standard, Available at: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-en.html.

Ventouras, S., 2013. Common Data Model for ESPAS (version 2). Available at: http://www.espas-fp7.eu/index.php/overview/deliverables.