ESTEC/Contract No. 14069/99/NL/SB
ESA Space Weather Study (ESWS)
WP432. Space Weather Service

Space Weather Service

ESWS-RAL-TN-0003
Issue 1.1, November 6, 2001

Author: **Richard Stamper, CLRC Rutherford Appleton Laboratory**

ESA technical Officer: **A Hilgers D/TOS, TOS-EMA**

Contents

List of Tables

List of Figures

List of Abbreviations

AE	Auroral Electrojet
AMPTE	Active Magnetospheric Particle Tracer Explorer
AO	Auroral Oval
Ap	Planetary A index (geomagnetic activity index)
ASCII	American Standard Code for Information Interchange
CARI	Civil Aeromedical Research Institute
CDF	Common Data Format
CDPP	Le Centre de Données de la Physique des Plasmas
CERN	Conseil Européen pour la Recherche Nucléaire
CME	Coronal Mass Ejection
CORBA	Common Object Request Broker Architecture
CREME	Cosmic Ray Effects on Micro-Electronics
CSMR	Consolidated System Measurement Requirement
CTIP	Coupled Thermosphere Ionosphere Plasmasphere model
DoD	Department of Defense
Dst	Disturbance storm time (geomagnetic activity index)
DTD	Document Type Definition
ECMWF	European Centre for Medium-range Weather Forecasting
EPO	Education and Public Outreach
ESA	European Space Agency
ESTEC	European Space Research and TEchnology Centre
ESWP	European Space Weather Programme
EU	European Union
EUV	Extreme UltraViolet
FTP	File Transfer Protocol
GCR	Galactic Cosmic Rays
GIC	Ground Induced Current
GIM	Global Ionospheric Maps
GNSS	Global Navigation Satellite System
GNU	Gnu's Not Unix
GUMICS	Grand Unified Ionosphere-Magnetosphere Coupling Simulation
HF	High Frequency
HLA	High Level Architecture
HTML	HyperText Markup Language
HTTP	Hypertext Transfer Protocol
IACG	Inter-Agency Consultative Group
IDL	Interactive Data Language
	Interface Description Language
IEEE	Institute of Electrical and Electronics Engineers
IMF	Interplanetary Magnetic Field
IPS	InterPlanetary Scintillation
ISO	International Standards Organisation
ISTP	International Solar-Terrestrial Physics
Kp	Planetary K index (geomagnetic activity index)
LEO	Low Earth Orbit
LET	Linear Energy Transfer
LHC	Large Hadron Collider
MSFM	Magnetospheric Specification and Forecast Model
NASA	National Aeronautics and Space Administration

NGDC	National Geophysical Data Center
NOAA	National Oceanic and Atmospheric Administration
PHP	PHP: Hypertext Processor
PIM	Parameterized Ionospheric Model
POES	Polar Operational Environmental Satellite
PR	Public Relations
PRISM	Parameterized Real-time Ionospheric Specification Model
RAID	Redundany Array of Identical Disks
s/c	Spacecraft
SCR	Solar Cosmic Rays
SEC	Space Environment Center
SFE	Service Functional Element
SFR	Service Functional Requirement
SIP	Service Implementation Proposal
SLAC	Stanford Linear Accelerator Center
SPENVIS	SPace ENVironmenet Information System
SUR	Service User Requirement
SIDC	Solar Influences Data analysis Center
SMS	Short Message Service
SMTP	Simple Message Transfer Protocol
SQL	Structured Query Language
SRD	System Requirements Definition
SSN	SunSpot Number
STP	Solar-Terrestrial Physics
SWS	Space Weather Service
TEC	Total Electron Content
UCL	University College London
UK	United Kingdom
US	United States (of America)
UV	UltraViolet
WBMOD	WideBand MODel
WDC	World Data Centre
XDF	eXtensible Data Format
XML	eXtensible Markup Language
XSIL	eXtensible Scientific Interchange Language

1 Introduction

1.1 Purpose

This document presents a proposal for implementation of the service element of an operational European Space Weather Service (SWS). Specifically, it describes a generic architecture for providing such a service, considers how the service should acquire, process and disseminate data, and outlines the operating procedures required to manage it.

Use is made of findings presented in other reports that have been written as part of a wide-ranging study [1] of the prospects for a future European Space Weather Programme (ESWP). The principal components of the study that have provided input to this document are the Market Analysis report [2], the System Requirements Definition Document [3] and the prototyping activity carried out in Work Package WP434 [4]. The Market Analysis presents a broad assessment of the interest in and potential market for space weather services and is used to provide general input to the desired features of the SWS. The System Requirements Definition document provides a wealth of specific details of the data and modelling needs that the SWS should meet, and the prototyping work gives some insights into the practical obstacles that the SWS would need to overcome during development.

1.2 Scope

1.2.1 Service vs Data provision

A key distinction is made in the Market Analysis report [2] between Space Weather *service* providers and *data* providers. The main function of a data provider is to distribute raw, or nearly raw, data to end-users. Service providers, on the other hand, process data with a targeted customer base in mind in order to provide a tailored and value-added facility focused on the needs of each customer or class of customers.

The Market Analysis showed that the data provider market is dominated by, and has largely been defined by, NOAA/SEC. Any future European Space Weather Programme would certainly provide data streams that both complemented and duplicated those provided by existing data providers such as NOAA. A notable feature of the data provider market is that there is a history of free data provision, so that the expectation of the user market is that raw or nearly raw data should be free. A significant finding of the market analysis was that 'it is doubtful that another broad-scale data provider is needed or wanted in the marketplace' [2, section 8.1]. Therefore, although an ESA Space Weather programme would undoubtedly contain an element of raw data provision, the SWS should offer more.

Service providers are more numerous than data providers and have hitherto concentrated on specific classes of end-users. Because of the extra value they bring to users in sifting, organising and presenting space weather data, customers are happy to pay for the services offered even when they can and do retrieve raw data free of charge.

1.2.2 Target audience

In the light of these observations and others, the assumption in this report is that the Space Weather Service should aim to be a coordinator, supporter and advocate of service providers, and an educator and source of information for a wider public audience. It should *assist* service providers in providing services for end-users, rather than providing them directly because of the amount of sector-specific knowledge required to do so effectively. This is not to say that end-users would not be able to use the SWS, but that to make use of the detailed data or modelling services a certain level of background knowledge would be required, so that if an end-user organisation had this expertise in-house it could be regarded as acting as its own service provider.

This approach is similar to that adopted by national and international meteorological offices in relation to terrestrial weather. Meteorological offices issue forecasts of various parameters of the tropospheric environment; the selection of parameters to forecast is made with a whole range of potential end-users in mind, but the focus of a meteorological office is the environment not its specific effects on users. For example, the likelihood of future warm dry weather is important to a supermarket company in deciding the volume of salad vegetables it needs to purchase, but a meteorological office would not presume to instruct the company on its purchasing policy.

The same focus on describing and forecasting the environment should apply to the SWS, only more so because of the fact that the impact of space weather is mediated by relatively complex technological systems. Without detailed knowledge of the operating characteristics of the systems involved, a Space Weather Service could only give the broadest type of 'at risk' warning which is unlikely to provide a sufficient basis for taking potentially costly operational decisions. The vulnerability of a power distribution network to problems arising from Geomagnetically Induced Currents, for example, will depend on factors such as the orientation of the network, the characteristics of the transformers and the position and orientation of the electrojets. The business of the SWS should be to provide reliable and accurate descriptions and forecasts of the environment, in this case the electrojets, rather than attempt to provide a specific decision-making system for power companies.

Note that the Space Weather Service is not targeted at 'scientists' in general – its focus is space *weather* not space *physics.* It is assumed that data relevant to scientific work would be better placed in long-term archives of scientific data, for example the World Data Centre system or the CDPP at Toulouse. Where individual scientists or scientific teams are concerned with space weather *per se* then their requirements of the SWS are likely to match those of companies or institutions providing services to end-users.

The two sets of target users – space weather service providers and the broader interested community – are very different. There is a great distance between them, both in terms of the sophistication of their understanding of the physics of space weather and of their data needs. The SWS must therefore be designed with careful attention paid to the different needs of these two target user groups. It is on the basis of this analysis that structure of the proposed Space Weather Service presented in this document is developed.

1.3 Structure of the document

Section 2 describes the generic system objectives in more detail and outlines the overall system architecture needed to meet these objectives. The Service User Requirements (SURs) are identified and enumerated as are the Service Functional Elements (SFEs) needed to meet them.

The issues raised during this overview are then explored in more detail in subsequent sections: Section 3 deals with all data-handling issues – the data formats and the data flows into, through and out of the system; Section 4 considers how simple data products should be enhanced to provide a more valuable service to users; Section 5 deals with the issues of the user interface and user-support; Section 6 considers the outreach role of the SWS in promoting a wider understanding of space weather; Section 7 covers the management and operation of the service and identifies the resources required to implement the system, in terms of equipment and personnel. Throughout these sections functional properties that are required of the SFEs are identified and tabulated at the end of each section as Service Functional Requirements (SFRs) numbered within the relevant SFE; where specific proposals are made as to how these functions should be provided, these Service

Implementation Proposals (SIPs) are also listed.

Section 8 considers possible future developments of the service, whether scientific, technical or organisational level. Finally there are two appendices – Appendix A presents the details of an estimate of the daily data flows into the SWS, and Appendix B collates the SIPs from the earlier sections of the document into a single table with cross-references back to the earlier discussion.

2 Architecture overview

2.1 Philosophy

On the basis of the analysis summarised in the previous section, it is clear that the Space Weather Service should aim to satisfy two, largely distinct, sets of target users: space weather service providers and a broader interested public. The most productive use of resources will be to develop a system that is well suited to meet the specific needs of these two very different groups, without deliberately trying to develop a universal system that seamlessly delivers what both groups require.

Section 2.2 considers the needs of the two sets of target users. These are used to identify a collection of service functional elements that the SWS should provide. These service elements are then drawn together in Section 2.3 into a single schematic illustrating the proposed system architecture.

2.2 User Requirements

We briefly enumerate here the principal requirements of the target users, together with functional elements of the Space Weather Service that meet these needs; these are summarised in Table 2, which assigns numbers to the Service User Requirements (SURs) and their associated Service Functional Elements (SFEs).

2.2.1 General

Although the intention is to cater specifically for two very different target groups of users, some needs are common across all classes of user. These include:

- Data from multiple sources worldwide available in a reliable and timely manner.
- Sufficient documentation to support understanding of how to use the service, and how to interpret the data products.
- A service that continues to develop.

Service elements to meet these needs include:

- Data supply must be via the network from high-availability computing systems.
- A flexible scheduler to fetch data when appropriate.
- A comprehensive set of documentation, including dataset metadata.
- Access to some level of human support when pre-prepared documentation is not enough.
- Monitoring of prediction performance.
- Procedures for user feedback.

2.2.2 Service providers

Service providers are assumed to be knowledgeable about the physics of space weather and can be expected to be running operational models of specific phenomena of their own, driven by space weather data. This user group therefore has the more demanding set of needs, since they are likely to need a wide range of data regularly and promptly, and will have well-defined interests in models and predictions of specific parts of the space weather system. Their needs include:

- A straightforward and uniform interface to multiple data sets.
- Ease of identifying relevant data sets by class of application, type of measurement, temporal resolution of observation, promptness of data acquisition.
- Access to sophisticated enhanced products e.g. altitude-varying models of energetic particle fluxes, spatially resolved forecasts of GICs.
- Historical information needs to be accessible for post-incident analysis and for model development.
- Ability to tailor a personalised list of regular data retrievals
- Access to informed advice and scientific technical support.

Further service elements are required to meet these needs:

- A generic but comprehensive and accessible data format.
- A Data Dictionary to enable data searches to be made by class of application, region of interest, feature of interest, promptness of data acquisition, prediction period etc.
- The service must provide a core of enhanced products e.g. data aggregation, model outputs, forecasts, appropriate visual representation.
- A data archive covering a number of months or years.
- Personal profiles for FTP or e-mailing of data.
- Human support that commands the respect of users by being demonstrably competent to handle detailed queries about data products at both the scientific and technical levels.

2.2.3 Wider public audience

The other class of user that the SWS should target is the broader interested community. This class covers many groups: informed commentators such as technical journalists, decision-makers in politics and industry, potential industrial users, the education sector ranging from schools through to universities and technical colleges, and the general public. The needs of this range of groups are rather different from those of service providers, since they are likely to need considerable background information on the science of space weather and are unlikely to have an ongoing interest in detailed modelling work. Their principal needs are:

- Extensive and accessible background information on the science and applications of space weather.
- Attractive presentation of selected data products.

The service elements required for these needs are:

- On-line material introducing the science of space weather and the impacts it has on various technological sectors: airlines, electric power, geological/offshore, HF radio, insurance and financial services, military, satellites, tourism.
- Outreach materials e.g. posters, stickers, CD-ROMs, curriculum packs for schools, museum exhibits, newsletters.
- A graphics engine to generate attractive plots from data.

The last of these elements would have general value for all users, but the first two are very different in nature from the others covered earlier. Moreover, much of the material will already exist so the development of these system elements could be largely independent from the rest of the SWS and developed separately or in parallel. The challenge will be to bring the science and space environment data closer to the public in a readily appreciated away (e.g. the AuroraWatch UK web-site [5]), without slipping into the role of being just ‘another broad-scale data provider’.

2.3 Generic Architecture

A high-level design of the computer and database systems needed to meet the user needs outlined above is shown in Figure 1. There are five large components:

Retriever Those elements that fetch relevant data from data providers.

Database Storage for parameters, enhanced products and associated metadata.

Modeller The collection of modelling and forecasting processes that work on data to generate more sophisticated data products.

Interface The modules concerned with handling interactions with users.

Scheduler A generic module that schedules the many system processes that need to happen at specific times.

The detailed functionality of these components is dealt with in more detail in Sections 3, 4 and 5. Throughout these detailed considerations three underlying assumptions about the appropriate computing technologies apply:

- For ease of access and ease of use, by data providers, service providers and the broader interested community, all systems should be based on the public internet using open standard protocols and programming languages wherever possible. This facilitates interopability and maintenance of systems and software.
- Any web sites should be database driven, using technologies such as PHP, Perl or Java Servlets; all three have public standards and interfaces with all major relational databases. Static web pages are harder to maintain.
- To save on development time and cost the use of off-the-shelf software should be encouraged. Much of this may be open source software, which should not be neglected provided it is well-supported and has a robust pedigree. Prominent examples are the Apache web server, Perl for scripting, the GNU *wget* tool for retrieval or mirroring of data from data suppliers.

3 Data Handling

The core of the SWS is the data describing the space weather environment, plus ancillary data describing the impact of space weather phenomena on the immediate environment of ground-based systems. There are three roles that data plays in the SWS: as an input to the system, as it is used and stored within the system, and as an output from the system; each role is considered in turn.

SUR	User Requirement	SFE	Service Functional Element	Comments
1	Timely and reliable data from multiple sources	1 2	Networked and reliable data access Retrieval scheduler	Delivery by magnetic media is too slow. Must be able to carry out retrievals automatically.
2	Good documentation of system and data	3 4 5	On-line help Comprehensive metadata Human support	 A prerequisite for sophisticated data provision to users. For when all else fails.
3	Consistent interface to multiple datasets	6	Generic, comprehensive and accessible data output format	
4	Easy to identify relevant datasets	7 8	Data dictionary Yellow pages system	Having metadata is not enough – it must be query-able. For locating datasets at remote data providers.
5	Access to enhanced products	9 10	Data aggregation. Models and forecasts.	
6	Access to past data	11	A local archive of relevant data	Important for post-incident analysis and monitoring of quality of warnings.
7	Personalised regular data retrieval	12	User accounts with personal profiles	
8	Access to informed advice and scientific technical support	13	Technically and scientifically competent personnel	A consultancy role
9	Background information on science and impact of space weather	14 15	On-line introduction to space weather. Outreach materials	Much pre-existing material exists. Posters, stickers, curriculum materials, CD-ROMs, displays, museum exhibits; see Section 6.
10	Graphical presentation of selected data products	16	Graphics engine	
11	Continuous service development.	17 18 19	Regular service monitoring User feedback facilities Medium and long-term strategic planning	 On-line and face-to-face.

Table 2: Service User Requirements and Service Functional Elements

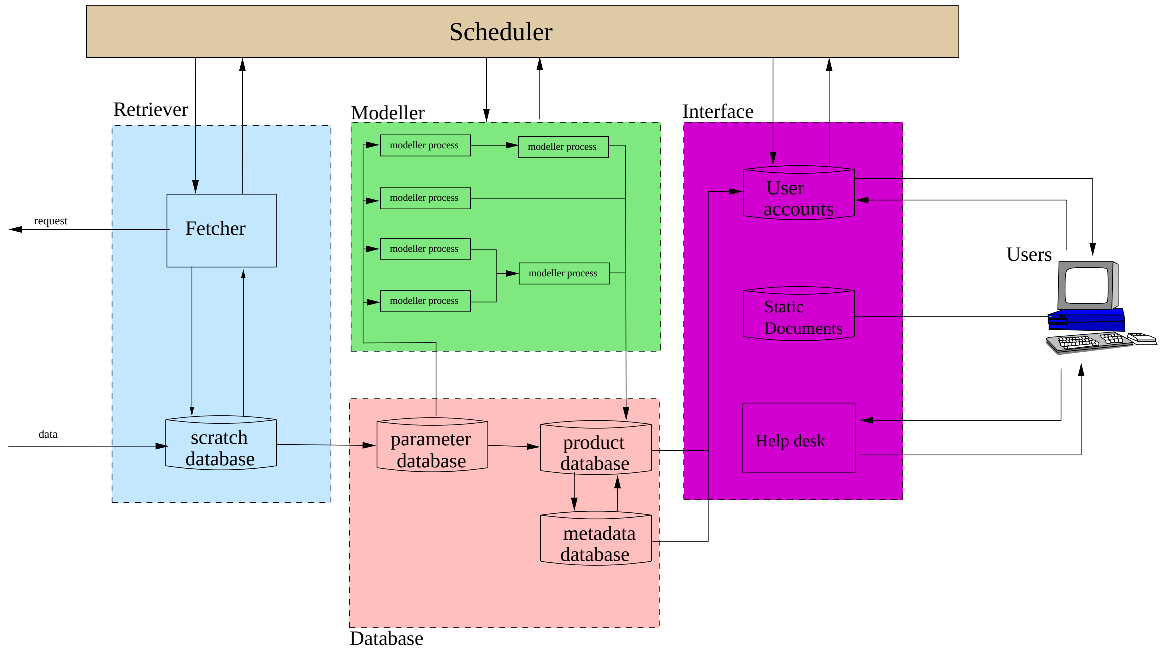


Figure 1: High-level design for the Space Weather Service

3.1 Data input

The essential input to the SWS is the flow of raw data from the space environment. For the SWS to serve its purpose it needs to acquire the data necessary for modelling and prediction, and it needs to do so in a timely manner to meet the operational needs of its users. This section is concerned with the data to acquire; the matter of its timeliness is dealt with in section 3.2.3. The data feed to the SWS would need to satisfy the Consolidated System Measurement Requirements (CSMRs) identified in the System Requirements Definition (SRD) [3] and summarised in Table 10, Appendix A.

3.1.1 Data sources

One important group of data inputs to the SWS are those from ESWP space weather instruments. As part of a full programme it is expected that there will be a number of instruments flown on ESWP missions, whose purpose is primarily to provide data for the Space Weather programme. The interface between these instruments and the ground segment is covered by another document [6], which describes a conventional front-end service that generates calibrated physical parameters from the raw instrument data. It is the calibrated physical parameters that are of direct interest to the users of the SWS, and with which we are concerned here.

It is assumed that each instrument front-end service will provide its own archive of the raw instrument data and a short-term collection of the derived calibrated physical parameters that result. A generic interface has been specified [7] for the delivery of these parameters to the SWS. The responsibility for long-term archives of the calibrated physical parameters from dedicated ESWP space weather instruments could lie either with the instrument front-end service or with the SWS. Which option is chosen may depend on the nature of the missions on which the instruments are flown; a dedicated mission is more likely to have an autonomous archive than a 'piggyback' instrument on another mission.

The other class of data inputs to the SWS are those from other data providers, principally by the delivery of data over the internet. Because of the variety of ways in which other remote sites provide data there are a variety of data retrieval mechanisms possible, several of which were implemented in the prototype service [4] developed as part of this study. The data retrieval capabilities that the SWS should offer are considered in the following section.

3.1.2 Data Retrieval

Data delivery (and request) mechanisms can be classified usefully according to two different schemes. Both schemes regard the delivery of data as a dialogue between a local and a remote site, and both analyse this dialogue using a pair of classifiers. The first scheme classifies methods by who starts the dialogue and what determines the data content; the second scheme classifies the methods according to how the two parties communicate the request and the data.

The first classification is by Starter/Determiner. In this we consider

- Who starts the dialogue – either the Local site (needing data) or the Remote site (providing data).
- What determines the data that is sent – this can be one of three possibilities: the Time at which the dialogue was begun, some instruction from the Local site, or at the discretion of the Remote site

The six possible combinations of these two classifiers are shown in Table 3, where each is paraphrased as a question and an example is given of how that mode has been used in practice.

		Determiner		
		Time	Local	Remote
Starter	Local	**"Tell me about now"** The method used when retrieving a file that is updated regularly – the time you access the file determines the data you get. (ACE data from SEC)	**"Tell me about this"** This is the usual approach where the site needing data specifies, using an HTML form or by constructing a filename, what data to fetch. (Geophysical parameters from WDC for STP, Chilton)	**"Tell me what you know"** A retrieval of opportunity where, for example, a file is updated intermittently. A retrieval will get the latest data, but 'latest' need bear no fixed relation to the time of retrieval. (Cluster Science Data System)
	Remote	**"This is what I know about now"** This is the mode of regular e-mail distributions. (SIDC Monthly mailing of sunspot numbers)	**"What do you want to know?"** The requester could put a request on their local FTP site, which would be retrieved by the data provider at their discretion to determine what data to transfer by FTP PUT, for example. (No known examples)	**"This is what I know"** This is the mode of Alerts and Warnings distributed by e-mail. (SIDC PRESTO alerts)

Table 3: Starter/Determiner classification of data retrieval

The second classification is by Request/Delivery. This divides data retrieval methods according to the protocols used to make the initial contact (data request) and to deliver the data. This classification is open-ended in the sense that new messaging and data transfer protocols are developed over time and existing protocols change in popularity. Table 4 shows the combinations of request and delivery for the three most widely-used protocols at the present – FTP, HTTP (web) and SMTP (e-mail). Note that the two dimensions of the table

		Delivery		
		FTP	HTTP	SMTP
Request	FTP	Standard FTP retrieval	Not used	Not Used
	HTTP	Typically a large data request is made by HTTP then retrieved from an FTP site	Standard web retrieval	Data e-mailed following a web request
	SMTP	E-mail request then FTP retrieval e.g. NDADS-ARMS at NSSDC	E-mail triggers HTTP retrieval	Mail responder e.g. Geomagnetic Information Nodes

Table 4: Request/Delivery classification of data retrieval

are only the protocols used for the start and end of the dialogue between local and remote sites. For example, the HTTP-FTP combination would typically be a two-stage process in which the requester completes a web form to request data and is returned a web page giving the location of a newly created file on an FTP site. Parsing this web page to extract the FTP location would enable the requester to start an FTP session to retrieve the desired data.

Two of the combinations are shown as ‘Not used’, not because they are impossible but because there are better ways to achieve the same ends. For example, one could imagine an FTP-SMTP scheme in which a file indicating the data required was deposited on a remote FTP site, and subsequently the remote site would e-mail the data to the requester. This is entirely feasible but the same effect is more likely to be implemented in practice using the conventional SMTP-SMTP combination.

The SWS will need to be able to retrieve data using many of the methods described in Tables 3 and 4, and consequently the SWS must include several capabilities:

- Use of an FTP client to retrieve data.
- Use of an HTTP client to retrieve data.
- An e-mail system for sending data requests and parsing received data e-mails,

All of these must be able to function without operator intervention and be capable of being invoked automatically on a predefined or adaptive schedule. The e-mail requirements can be met by commonly used Unix tools such as *sendmail* or *procmail* in conjunction with the standard mail client *mail*. It is worth emphasising that the widespread use of e-mail as a data exchange mechanism is not always appreciated by corporate managements. This brings the danger of corporate e-mail policies or systems that assume that e-mail is used only for person-to-person communication and prevent the use of e-mail for automatic data exchange.

The FTP and HTTP retrieval requirements also include automatic invocation and processing of retrieved results. These can be met by various tools in the Unix environment; a prominent example is the open source software tool *wget* that provides recursive FTP and web site download and mirroring facilities. The scheduling demands can be met by standard operating system scheduling tools such as Unix *cron* or the Windows NT *schedule* service. For the highest time resolution retrievals operating at minute or sub-minute frequencies, a continuous monitor or server process would be more suitable.

SFR	SFR Description	SIP	SIP Description
1.1	Allow data retrieval using the FTP, HTTP and SMTP protocols	1.1.1	Use the FTP/HTTP agent GNU *wget* for FTP and HTTP retrieval and mirroring
		1.1.2	On Unix, the requirements are met by standard Mail Delivery Agents such as *sendmail* in conjunction with *procmail* and the mail client *mail*.
1.2	Data retrieval mechanisms should be sufficiently flexible to cope with the systems of many different data suppliers.	1.2.1	Provide for data acquisition using all methods from Table 3 except the Remote/Local combination.
2.1	Data retrieval must be able to proceed on a pre-determined schedule without human intervention.	2.1.1	The Unix *cron* or Windows *schedule* facilities should be used to start retrieval jobs as required.

3.2 Data Storage and Processing

There are a number of key data storage and processing issues that the SWS must address:

- Local versus Remote data
- Data history
- Scheduling
- Database systems
- Data formats
- Metadata

3.2.1 Local versus Remote data

The SWS is not primarily intended to be an archive of space weather data. There are several good reasons, however, for keeping copies of retrieved data locally for an extended period:

- Necessity – For any dedicated ESWP space weather instruments it is assumed that the SWS will act as the definitive archive of the calibrated physical parameters that they generate. Note that the presence of these local archives is not material to the modelling and forecasting functions of the SWS, since the data retrieval interfaces should not discriminate between local and remote data.

- Buffering – To guard against network delays impacting the prediction functions of the SWS, it would be useful to fetch data as soon as it is available rather than on a just-in-time basis. This could be regarded as a caching rather than archiving function, but given the range of models and predictors that the service will run on different timescales, the cached data might need to be kept for weeks or months.

- Review – Some user requirements are for post-event analysis up to several months after a space weather event. Examples include the evaluation of insurance claims on satellite failures and airline monitoring of radiation doses to crew members following extreme events.

- Development – Models and prediction tools need data for their development, and for them to be effective under operational conditions they need to be developed and tested on data that resembles the environment in which they will be run in earnest. Having access to realistic snapshots of the data environment over a range of times is therefore of great value important when developing predictors.

For all these reasons it would be prudent for the SWS to maintain substantial local databases of physical parameters retrieved from remote sources. Appendix A presents an estimate of the daily volume of data required to satisfy the system measurement requirements identified in the System Requirements Document [3]. The conclusion is that the daily data rate is of the order 200-1000Mb, equivalent to roughly 70-400Gb per year of archived data. These data rates and volumes are manageable with relatively modest modern hardware; for reliable and quick on-line access a RAID (Redundant Array of Identical Disks) system can be employed, with a larger volume of near-on-line data stored on a modular and expandable tape storage system.

The SWS would provide access to routine runs of various models and predictors, executed either locally or at co-operating institutions. The outputs of these model runs might also be worth preserving in many cases, particularly if their production was computationally expensive. The issues regarding models are considered in depth in Section 4.2.

SFR	SFR Description	SIP	SIP Description
11.1	Maintain local databases of retrieved data, estimated to accumulate at 200-1000Mb/day.	11.1.1	Use a RAID system with capacity of at least several months of data for immediate on-line access, with near-on-line tape storage holding the remaining data.

3.2.2 Data History

Related to the extent of local archiving of parameters and products is the ability to provide a data history. In order to assess the effectiveness of prediction algorithms and operational decision tools, service providers and their users will need to be able to recreate the data environment at the time that predictions and operational decisions were made.

For example, neural networks frequently require complete sets of desired inputs on which to operate, and also that the inputs be of good quality. Clean and comprehensive historical data is often available, but if the operational data environment in which a neural network predictor has to execute has many data dropouts or poor quality values then the predictor may perform much less well. Indeed, another predictor developed with explicit allowance made for the likely quality of operational input data may prove superior in practice, although inferior under ideal conditions.

Another example would be a post-mortem on an operational decision made by an end-user in the light of a space weather prediction. If that prediction, or other related predictions, were to be re-run after the event, then if the relevant input data had become more complete or had been corrected then the prediction results would not match those available when the original operational decisions were made. This would clearly invalidate any post-mortem investigation.

Being able to recreate the data environment is a non-trivial requirement since the availability and quality of data are likely to change as time passes. There is a tendency for data providers in the space data field to divide into two camps – those which concentrate on recent data (e.g. SEC) and those which provide long-term archives (e.g. NGDC). When a long-term archive is constructed it is usually the case that only the best quality and most complete dataset is stored. This will not be the case for a Space Weather archive.

To be able to recreate the operational data environment imposes two significant conditions on the capabilities of the SWS data storage systems:

- Data updates must be non-destructive. In particular, if a parameter has its value corrected at a later date, or a new version of a dataset is issued, the uncorrected or older versions must be preserved.
- All data items retrieved and archived by the SWS must have a timestamp of retrieval associated with them, in addition to the timestamp of observation that they customarily have.

Provided these conditions are met, it should be possible to recreate the data environment at any given time. Note that this provides another reason for maintaining a local archive, since it is only with local data that these conditions can be guaranteed.

SFR	SFR Description	SIP	SIP Description
11.2	Store data so that the operational data environment at any point in time is recoverable at any later time.	11.2.1	Update data non-destructively.
		11.2.2	Timestamp all data files or values with their time of retrieval.

3.2.3 Scheduling

The timeliness of data retrieval, processing and distribution is vital for a space weather service. Data must be retrieved on a variety of timescales, ranging ideally from every few seconds for magnetometer data up to only a couple of times a year for space debris distributions. The quantity and diversity of the data required (see Appendix A) necessitates automatic retrieval for the majority of datasets. Many users will want to receive data from the SWS on a regular basis as well, so automatic distribution must also be possible.

Besides retrieving data for onward distribution, space weather parameters will be used to drive prediction tools and other models. One can envisage these being invoked both on demand and routinely. In the latter case, the model or prediction processes will also need to be started and managed automatically.

These scheduling demands can mostly be met by standard operating system scheduling tools such as Unix *cron* or the Windows NT *schedule* service.

SFR	SFR Description	SIP	SIP Description
2.2	Data processing and distribution must be able to proceed on a pre-determined schedule without human intervention.	2.2.1	The Unix *cron* or Windows *schedule* facilities should be used to start data processing or distribution jobs as required.

3.2.4 Database systems

The internal organisation of the data stored by the SWS should not have any functional consequences for users of the system. The choice of database system will have significant impact, however, on the ease with which data are ingested, processed and distributed. Three broad approaches are possible: the use of a relational database system, an object-oriented database system, or storing a collection of structured data files.

Relational databases Relational databases have the advantage of being a mature technology with a well-defined and widely used query language, SQL [8, 9]. By placing all data in a set of tables they render straightforward the process of selecting data from multiple sources according to complex criteria based on the values of data items. This capability has some potential benefit for the SWS given its role in managing data acquired from numerous sources, although such content-lead queries are more likely to be a feature of a research activity than an operational service. A drawback to the relational approach is that there is little direct provision for metadata in relational systems. This is not necessarily a problem for simple data or data that is well understood by all users, but this is not the case for space weather. Space weather data is relatively complex and has extensive metadata necessary for its proper interpretation, and for this reason relational databases are not the best data storage method for the SWS.

Object-oriented databases One method for binding metadata and data more closely is to use an object-oriented database, which represents data not as tuples in tables but as objects with associated behaviours and properties. This is a much newer technology but one with several advantages. The data model corresponds better to the world than does the relational data model and because information about an object is maintained within a single data entity, data access can be much quicker for standard queries. Moreover, the object-oriented approach to data storage fits well with modern object-oriented programming languages, thus simplifying data access code. This would be an advantage to the SWS in the development of code for modelling and prediction. There are some general drawbacks, prominent being the absence of a standard query language along the lines of SQL. In the case of the SWS, however, there are other reasons why object-oriented databases might not be appropriate. The SWS will mainly be retrieving data solely on the basis of its time of observation, and feeding it either to model and predictor processes or direct to users. The mode of data query is thus one driven by the *value* of an attribute of an observation, rather than the *relationship* between objects. This is a form of query that object-oriented databases do not handle well, since they can involve a trawl through many objects. Moreover, the matter of presenting data to users in a consistent and easily comprehended fashion is not addressed, and this is a critical consideration for the usability of the service. Although object-oriented databases present a better option than the relational model, they are still not ideal.

Structured datafiles The final option is the use of structured data files. Although this might be regarded as an old-fashioned approach, for the purposes of the SWS it is the most satisfactory. The data retrievals necessary for the operation of the SWS would not usually be data-driven; almost all data would be requested by time of observation, so data files that can be identified as holding observations from an instrument for a specific time range are sufficient. Another advantage of holding data internally in files is that data output from the SWS will need to use a structured file format. With either a relational or object-oriented system extracted data would need to be appropriately packaged before distribution, but if structured files are used internally by the SWS then preparing data for distribution is greatly simplified. One key requirement, as discussed previously, is that the metadata be easily retrieved with the data. There are several candidate file formats that enable one to keep metadata and data together, and their relative merits are discussed in the following section. On balance, the use of structured files is the most suitable data storage method, with an object-oriented database system as the second choice.

SFR	SFR Description	SIP	SIP Description
11.3	Store data in a manner that simplifies ingestion, processing and distribution.	11.3.1	Store data in structured files rather than relational or object-oriented database systems
4.1	Store data in a manner that facilitates the association of data with the necessary metadata for its proper use.	4.1.1	Store data in structured files with integrated provision for metadata

3.2.5 Data formats

If a structured file format is to be used for internal data storage in the SWS, then a choice of format must be made. The oldest and currently most widely used structured file format in the space weather field is the binary Common Data Format, CDF [10]. Recently other formats have been developed based on XML, the Extensible Markup Language which is becoming widely used for data exchange in many disciplines. The formats proposed that are relevant to science data in general and space physics in particular are the Extensible Data Format, XDF [11] and the Extensible Scientific Interchange Language, XSIL [12]. Some of the advantages and disadvantages of the CDF, XDF and XSIL formats are given in Table 5.

The various advantages and disadvantages listed in Table 5 vary in importance. The fact that XSIL is a lightweight format consequently making much greater demands on data access programs is crucial, particularly if XSIL were to be used as an output format to users. On that consideration alone, XSIL is probably not suitable. The decision between CDF and XDF is much less clear cut. CDF has the advantage of being well-established; XDF, being based on XML, can be seen as 'coming technology' and is likely to be widely adopted given its origin (like CDF) with NASA. The XML basis of XDF also makes data more easily accessible by non-proprietary software, which might be a consideration for the service providers that the SWS is intended to serve. For CDF, and potentially for XDF, the NASA pedigree is significant since it follows that associated software is free and is likely to be supported over many years once in widespread use for NASA datasets.

SFR	SFR Description	SIP	SIP Description
11.3	Store data in a manner that simplifies ingestion, processing and distribution.	11.3.2	Store data in CDF or XDF format files. At the time of writing CDF is the better option, but this may change rapidly over time.

3.2.6 Metadata

The need for good metadata has been noted previously. Metadata is necessary for the proper understanding of data so should accompany data when stored in or exported from the SWS. Sufficiently comprehensive metadata is also needed when automating data processing or display procedures. For example, the CDAWlib display and retrieval tools that work with the CDF data format rely on the metadata for a datasets complying with the ISTP/IACG guidelines [13]. Without the degree of metadata completeness enforced by these guidelines the CDAWlib software will not be able to handle the data properly. This aspect of metadata functionality has been covered by SIP 4.1.1.

Metadata is also a valuable resource to enable users to locate the data they need by specifying properties of the data rather than where it can be found, for example selecting by a region of interest, type of particle, energy range, or by name of parameter. To be able to use metadata in this fashion it is not sufficient for it to be stored with the data in CDF or XDF files, it must also be databased so that it is searchable.

The Space Weather Service prototype [4] implements a database of metadata called the Space Environment Yellow Pages. This demonstrates two forms of metadata that are important to store in a readily retrievable format: metadata describing the form and content of the data, and metadata describing how to retrieve the data – its location. In the full SWS being outlined in this document, much of the location metadata would point to a local source within the SWS itself, but this would not eliminate the need for it. It would be valuable both as an internal resource to the SWS when retrieving data from remote data providers, and as a resource for users to enable them to identify these remote data sources.

Format	Advantages	Disadvantages
CDF	– Widely used. – Many tools and libraries for manipulating files on many architectures and operating systems. – Support for Java and IDL access to data. – Standards for space science metadata defined (ISTP/IACG guidelines [13]. – Developed and used widely by NASA. – Widely user in Europe e.g. Cluster, Equator-S, AMPTE, SPENVIS.	– The file structure is relatively complicated. – A private binary format, so tied to proprietary data access libraries. – Format has changed between versions, forcing changes in installed software.
XDF	– Based on the open standard of XML and the metadata (at least) is ASCII, so easily inspected. – Defined by an XML DTD (Document Type Definition) so files are parse-able by general XML software. – Software libraries for Perl and Java to access data. – XML feature of reference to or inclusion of the DTD allows the format to evolve in a manner visible within individual data files, hence visible to software. – Developed and supported by NASA.	– Not widely used yet. – No standards defined for required metadata. – The file structure is relatively complicated.
XSIL	– Based on the open standard of XML and the metadata (at least) is ASCII, so easily inspected. – Defined by an XML DTD so files parse-able by general XML software. – Well coupled with a visualisation tool. – Software library for Java to access data. – XML feature of reference to or inclusion of the DTD allows the format to evolve in a manner visible within individual data files so visible to software. – Lightweight format to be extended on a per discipline basis.	– Not widely used yet. – Lightweight format imposes greater demands on processing code.

Table 5: Structured File Formats: Advantages and disadvantages

Metadata queries are typically value driven, unlike those to be expected of the space environment data itself. It is therefore appropriate to use relational database technology for storing metadata, following the same arguments rehearsed earlier in Section 3.2.4 when discussing the data itself. Using a centralised database for the location metadata is also advisable, since it is the form of metadata that is most volatile so needs to be easy to amend as data sources are moved or change their format.

SFR	SFR Description	SIP	SIP Description
7.1	Allow the data holdings to be searched by metadata attributes	7.1.1	Extract datafile metadata and store in a relational database – a 'Data Dictionary'
8.1	Allow easy configuration and access to location metadata	8.1.1	Store location metadata in a relational database – a 'Yellow Pages Directory'.

3.3 Data Output

3.3.1 Formats

The SWS will not have control over the format of data from other data providers but can determine the format of internal data and data it distributes. One of the key user needs given in Table 2 is for a consistent and accessible data format in which to distribute the diverse data holdings of the SWS. The leading candidate formats are again CDF and XDF since they integrate data and content metadata and are well supported by libraries and tools for reading and manipulating the data. These formats are well suited to professional users of space weather data, but there may be demand from more casual users for a more 'approachable' ASCII format. Provision of simplified access might be worthwhile for a selection of data products, but should not be the first priority. The use of the XML-based, and hence ASCII, XDF format might be sufficient to meet the needs of both professional and casual users, particularly since the XDF format does provide the facility for referencing a subsidiary data file from the formal XDF document. This subsidiary file could then be a plain ASCII data file suitable for casual retrieval or inspection. More sophisticated users who require metadata would fetch the more complete XDF file instead.

3.3.2 Methods

As with data input to the SWS, it is necessary to consider the methods by which users would need to retrieve data. The analysis of data retrieval methods presented in Section 3.1.2 applies equally when viewed from the other side, with the SWS as the 'remote' data-providing site rather than the 'local' data retriever. The same reasoning applies – the SWS should make it possible for users to retrieve data by a variety of methods, to accommodate their various needs and operational constraints.

Examining first the Starter/Determiner classification of data retrieval (Table 3), the SWS should offer all of the combinations except Remote/Local. The Local started combinations are all simple to provide; the distinction between Time or Remote (i.e. SWS) determined retrieval will depend on the type of data being retrieved, and the Local determined option is that commonly provided by data access driven by web forms. The Remote started combinations of Time and Remote determined data, corresponding to regular mailings and irregular alerts respectively, should be offered for appropriate data products such as daily Kp forecasts and CME warnings.

A number of the data delivery mechanisms shown in Table 4 should also be offered. Given the dominance of the web for data delivery the principal method should be to use HTTP for request and delivery. Thus the core of the SWS will be a web site of dynamic pages, offering a a range of forms-driven access to the data holdings and other services. The market-leader for web servers is the Apache server [14], which has well-integrated support for a range of scripting languages such as Perl CGI and PHP.

Providing an e-mail interface to the same data request software would be straightforward and should also be done. The delivery of data by e-mail can be accomplished using standard Unix tools such as *sendmail* and *mail*, and use of these or similar tools can be abstracted from by using, for example, the Perl Mail::Mailer module.

Using an FTP client to deliver data to a user's FTP site is conceptually no different from delivery by e-mail, and can be achieved with various degrees of sophistication: simple shell invocation of a command-line FTP client, managed use of a command-line FTP client with the *expect* tool, or programmatic use of the FTP protocol with the Perl LWP or Net::FTP modules (or similar for other languages).

The provision of passive FTP access via an FTP site should be a lesser priority. The nature of the FTP protocol is that it provides access to data via a hierarchy. Given that it is proposed to store the SWS data in structured files such access would be possible, but the data storage hierarchy does not necessarily correspond with the conceptual route by which users might want to reach the data. Providing more than one route to the same data is also made difficult.

There are certain classes of data product that might require more specialised data retrieval options. One example from the Market Analysis report is the application of auroral tourism; users, whether individuals or tour companies, might welcome a facility by which an SMS message could be sent to their mobile phones giving short-term alerts that an auroral display was likely.

3.3.3 Bandwidth

Estimating the volume of data that will flow out of the SWS is more difficult than estimating the inflows, as carried out in Appendix A, because it is entirely dependent on the number and type of users of the service. It is possible to make conservative estimates however, to arrive at an estimate of the network bandwidth required to support likely usage.

Assuming all access to data is via a nominal 'hit' on a web server, a generous estimate of the number of hits/day would be 10^5 and a similarly generous estimate of the transfer per hit would be 50 KBytes. This implies $5.0e6$ KBytes per day which is roughly 60 KBytes/s or 500 Kbits/s. Allowing a factor of 10 variation multiplier between average and peak requirements would imply a bandwidth of 5Mbits/s. This level of network capacity can be met through commercial leased lines, which are widely available at capacities up to 155 MBits/s. The cost of leased lines varies widely in the EU (see [15]) but illustrative costs are included in Section 7.3.

SFR	SFR Description	SIP	SIP Description
1.3	Data distribution mechanisms should be sufficiently flexible to cope with the needs of many different users.	1.3.1	Provide for data distribution using all methods from Table 3 except the Remote/Local combination.
1.4	Allow data distribution using the FTP, HTTP and SMTP protocols	1.4.1	Offer active FTP delivery of data (PUT operation) by using appropriate software, such as *expect* or the Perl LWP and Net::FTP modules.
		1.4.2	Offer dynamic HTTP access by running a forms-driven web service on the *apache* webserver [14].
		1.4.3	Offer e-mail data distribution using a standard Mail Transport Agent such as *sendmail*, Mail User Agent such as *mail*, or wrapped by higher-level software such as the Perl Mail::Mailer module.
		1.4.4	Consider novel data distribution channels (e.g. SMS) for selected products (e.g. alerts)
1.5	Provide sufficient network capacity to allow prompt access to SWS data products by users	1.5.1	Procure a leased line connection with bandwidth of at least 5Mbits/s.
6.1	Output all data using a generic and accessible data format, for consistency of access.	6.1.1	Output data in CDF or XDF formats.
		6.1.2	Facilitate simple ASCII output by structuring XDF files so that the data element has sufficient internal structure for it to be intelligible in the absence of the metadata elements.

4 Enhanced Data Products

The SWS as described so far would provide a consistent interface to multiple datasets providing observations taken between the surfaces of the Sun and the Earth. To offer a quality service to space weather service providers and thence to commercial users more is required; it must be possible for simple observational and index data to be combined to produce more sophisticated data products. Three generic type of enhancement are discussed in this section: data pipelining, modelling, and graphical display.

4.1 Data pipelining

There are a number of simple operations that can be applied usefully to many numerical data streams. While it is not the purpose of the SWS to provide generic data processing facilities for its data outputs, some operations are likely to be wanted sufficiently often as to be desirable to offer. The two classes of operations that should be provided are filtering and aggregation.

Filtering by time will be implemented as a matter of course; most data requests will be for a specific data product describing some feature of the space weather environment at a specific time or during a specific time interval, whether past, present or future. Adding the facility to filter by data values would allow users to fashion personalised 'Alerts'; for example, to e-mail the Ap values to the user when they exceed a given threshold. More complicated scenarios can be imagined, where users could configure the system so that one data stream would be retrieved contingent on some threshold being reached in another data stream.

Various data aggregation functions should also be offered. These include taking the mean, median, minimum, maximum or sum over some regular repeating interval of the domain of interest. This aggregation could be over time for a simple scalar time series, or on a spatial grid for observations distributed over an area or volume.

SFR	SFR Description	SIP	SIP Description
9.1	Enable users to filter data by time	9.1.1	All data requests should specify a time range.
9.2	Enable users to apply simple aggregating operations to data streams.	9.2.1	Implement a suite of aggregation functions to compute mean, median, minimum, maximum and sum on arbitrary data fields over regular repeating intervals of time or space.

4.2 Models

The provision of models, of varying complexity, is the most significant way by which the SWS can enhance the simple data streams that are at its disposal. In the course of determining the measurement requirements for the Space Weather Programme, the System Requirements Definition (SRD) document [3, Appendix B] considers the model processes needed to generate the desired information. These models or processes are classified as Mature, Immature or Speculative; a model described as 'Mature' can be regarded as being already in a position to give information of value. It is therefore incumbent on the SWS to provide these mature models as a minimum initial offering. The other models mentioned by the SRD should be considered for addition to the system when they are sufficiently mature, and there is some discussion of this in Section 8.

4.2.1 Required models

Table 6 presents details of all the mature model or processes identified in the SRD document. The table breaks these down in to four groupings: plasma and radiation environment, solid body environment, magnetic environment, atmospheric environment. For each model the table shows:

- Time of operation : F(orecast), N(owcast), H(indcast)
- Mode of operation : S(tandard), C(ustom). Standard outputs would be generated routinely and stored; custom outputs would require user input to generate the desired outputs on demand.
- Update period
- The numbers of the SRD User Requirements that the model is designed to help meet.

SFR	SFR Description	SIP	SIP Description
10.1	Offer a wide range of models and forecasting tools.	10.1.1	The models in Table 6 should all be provided.

Table 6: Initial set of models for the SWS to provide. The columns headed 'T', 'M' and 'UR' show the operation Time, operation Mode and the SRD User Requirement numbers respectively.

Model	**T**	**M**	**Update period**	**UR**
Plasma and Radiation Environment				
Climatic models of solar protons from Solar Proton Events e.g. JPL-91	F	S	6 monthly	1
Climatic models of GCR e.g. CREME [16] or CARI [17]	H	S	6 monthly	3
Magnetic field rigidity models to map observed fluxes of GCR, SCR, solar protons, energetic ions and neutrons to arbitrary orbits or spacecraft positions.	NH	C	30 minutes	13 15 19 20
Rigidity cut-off models (e.g. Störmer[18] or Shea and Smart [19]) to map observed fluxes of $>$10MeV ions to altitudes, routes and locations of aircrew, avionics equipment and spacecraft launches.	NH	C	hourly	2 3 22
Radiation transport model (e.g. GEANT[20]) to map $>$10MeV ions to aircraft altitudes and routes.	NH	C	hourly	13 15
Magnetic field model(s) (e.g. those of Tsyganenko [21]) to map observed fluxes of electrons, ions and protons of all energies along field lines to arbitrary orbits and spacecraft locations.	NH	C	1 minute to hourly	13 15
Prediction of Solar Proton Events from CME and flare detection.	F	S	1 minute	1 18
Plasma environment modelling (e.g. Salammbo [22] or MSFM) to fill gaps where good plasma measurements are not available.	NH	C	3 hourly	13 15
Solid Body Environment				
Climatological models of debris (e.g. MASTER-97 [23] or IDES) and random meteoroids.	FNH	S	6 months	13 15 18 19
Gravitational modelling of debris and random meteoroids to Low Earth Orbit.	FNH	S	6 months	13 15 18 19
Orbit propagation modelling of large debris.	F	S	6 months	14
Models of meteoroid streams [24].	FNH	S	daily	13 14 15 18 19
Magnetic Environment				
Predictions of Kp from solar rotation out to $\sim$27 days.	F	S	daily	7
Predictions of Kp from CMEs or flares out to 2 or 3 days.	F	S	hourly	4 7
Predictions of Kp from solar wind 1 or 2 hours (Lund [25, 26] or Costello [27] models).	F	S	15 minutes	7
Predictions of Dst from solar rotation out to $\sim$27 days.	F	S	daily	7
Predictions of Dst from CMEs or flares out to 2 or 3 days.	F	S	hourly	7
Predictions of local dB/dT from CME detection	F	S	15 minutes	4
Estimates of dB/dt by interpolation from local field measurements. This is probably better done directly by end users with their service providers because of the high temporal resolution required.	NH	C	10 seconds	5

Table 6: (continued)

Model	**T**	**M**	**Update period**	**UR**
Atmospheric Environment				
Predicted N_e profiles from climatic models driven by SSN, $F_{10.7}$, EUV flux and Kp (e.g.IRI [28]).	F	C	daily	10
Derived N_e profiles from ionosonde observations and interpolations of global or regional models e.g. ITU-R [29].	N	C	15 minutes	11
Derived N_e profiles by scaling static models with dynamic TEC data. e.g. PRISM [30], PIM [31] updated by JPL GIM [32].	N	C	5 minutes	11
Ionospheric scintillation from the NwRA model WBMOD [33, 34] driven by Kp and SSN forecasts.	F	S	15 minutes	10
Polar Cap absorption driven by CME/flare detection.	F	S	5 minutes	10
SW fadeouts from flare detection.	F	S	5 minutes	10
Predictions of decreased S/N ratio from solar radio emissions	F	S	5 minutes	10
Ionospheric storm forecasts from Kp predictions based on solar rotation.	F	S	3-hourly	10
Ionospheric storm forecasts from Kp predictions based on CME and solar flare detection.	F	S	hourly	10
Ionospheric storm forecasts from Kp predictions based on solar wind (Lund [25, 26] or Costello [27] models).	F	S	15 minutes	10
TEC from GNSS.	N	C	5 minutes	12
Atmospheric model (e.g. UCL CTIP) modulated by solar activity to get atmospheric drag in LEO.	N	S	daily	16
Statistical model of auroral location (e.g. NOAA POES [35]) from geomagnetic activity using Kp forecasts.	F	S	3 hourly	17

4.2.2 Model interconnection

Given that the SWS will need to use a wide range of models there arises the questions of model location and model interaction; where should a model reside and be executed, and how should models communicate if they need to?

The modelling expertise that has resulted in the list of models in Table 6 is distributed across a number of scientific institutions in Europe and elsewhere. The choice facing the SWS, for any given model, is whether to implement it locally or to access it remotely from another site. With sufficiently complex models it is generally advantageous to site an operational model at an institution where the model is well understood, typically where it was developed. This arrangement encourages the model to be well-maintained, and trouble-shooting can be done by staff experienced in the model's development. The advantages of local implementation of a model at the SWS are that there is full control over it and the overhead of setting up communication protocols between two sites is eliminated. Many of the models listed in Table 6 are sufficiently mature and well documented for local implementation to be possible, but distributed model implementation should still be considered.

Figure 2 illustrates three ways in which two distinct models could be run by the SWS. By combining these three forms all possible interactions of multiple models can be built up. The first and second forms are simple serial and parallel execution; the third form is more complicated since the two models interact pass information between each other during execution. This might occur when the models call each other iteratively to converge on a mutually satisfactory outcome. Another example is where a model can be run by making assumptions about the operating conditions, but these assumptions can be removed by using outputs from another model.

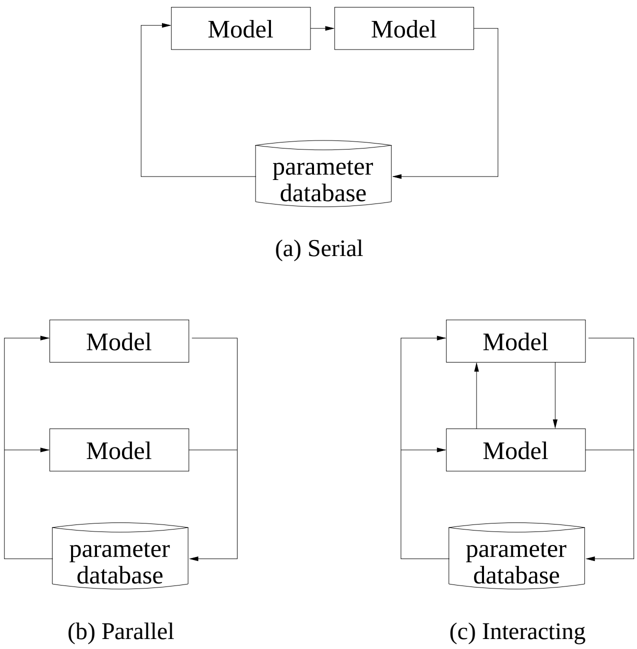


Figure 2: Modes of execution of two models

For the models listed in Table 6 only the serial form of interaction is needed; for example, the forecasts of ionospheric storms rely on forecasts of Kp. This case of simple chaining poses relatively few problems of communication, since a full model run is the natural level of granularity. The output of the first model can be passed in its entirety to the second and this provides the correct synchronisation between them. The case of independent but parallel execution offers no difficulty, since the two models can both run to completion. In these cases, whether the models are local or remote to the SWS, communication between models is naturally achieved by passing the structured data files (CDF or XDF) that contain the model outputs.

The case of interacting models is more complicated and raises all the classical problems of synchronisation and coordination between parallel processes. Some of these problems, such as deadlock avoidance, are specific to the models that are communicating so must be dealt with on a case by case basis. The question of how models should communicate is generic and the SWS should adopt a standard approach. This is considered in detail in Section 8.2.

SFR	SFR Description	SIP	SIP Description
10.2	Allow models to be interconnected	10.2.1	For the models in Table 6, passing standard output data files in CDF or XDF format is sufficient.
		10.2.2	Run models at the most appropriate location, which may be at a remote institution.
		10.2.3	Maintain awareness of developments in the Grid field, and on possible applications within space weather modelling.

4.2.3 Model execution

The data inputs to the models in Table 6 are updated at a wide range of time intervals from every ten seconds to only twice a year. The models also vary widely in the computer processing and data storage resources required for a full model run, and the frequency with which they are likely to be consulted. These two factors will determine whether it is most appropriate to run a model routinely or on demand.

An example of routine execution would be forecasts of Kp for hours ahead to 27 days ahead on timescales from 15 minutes to daily, which would be of general interest and relatively light in computing demands. Other model outputs are of interest only to specific parties; an example would be mapping high energy ion fluxes to specific aeroplane flight paths using a radiation transport model. These are specific to particular operators and the possible flight paths are too numerous to generate routinely. Instead, a user would invoke the model on the most current set of data on demand, although there might be pre-processing of input data carried out routinely to prime the model for user invocation.

If the frequency of update becomes too great or the model results too user-specific, then it may be better for the SWS not to offer a model directly. The clearest example from Table 6 is the estimation of dB/dT by interpolation from local field measurements with a ten second update period. In such cases the SWS should act as a source of advice or consultancy with the model run locally by the user.

SFR	SFR Description	SIP	SIP Description
10.3	Offer models to be run routinely or on demand as appropriate	10.3.1	The models with S(tandard) mode in Table 6 should generate standard outputs that are archived routinely.

4.3 Graphics

With the exception of some indices of generic interest, such as Kp or indicators of the likelihood of auroral activity for auroral tourism, the SWS should not make the graphical display of data products its priority. Since it is anticipated that the SWS would be acting principally in support of service providers, the greater priority is providing data in a readily accessible form rather than images of that data.

The option of providing graphical data presentation should be included, however, for which the SWS should make use of a generic display package such as IDL from Research Systems Inc. (where IDL here stands for Interactive Data Language, as opposed to the CORBA Interface Description Language). For the CDF data format there are additional data display libraries available, such as the CDAWlib library of IDL routines. Provided the data has appropriate metadata this can plot arbitrary time series data using appropriate line or spectrogram plots and spatial data on an appropriate grid. A capability like this is a necessary feature of the SWS. A couple of examples of the types of plots required are presented in Figure 3.

SFR	SFR Description	SIP	SIP Description
16.1	Provide the capability of generating plots from data outputs	16.1.1	Use the generic data visualisation package IDL.
		16.1.2	If using CDF, use the CDAWlib plotting routines on top of IDL.

5 User Interface

The interface between SWS users, whether service providers or the general public, consists of three parts:

- On-line data services on the internet.
- Off-line materials for distribution.
- Human contact for support or consultancy.

5.1 On-line data services

The majority of interaction with users for the SWS would be using on-line data services provided over the internet. The main access would be through a web service enabling users to set up data product retrievals. Data distribution would then be carried out using the appropriate methods, for example HTTP, FTP, e-mail or SMS.

A key feature of this service is that users should be able to configure a personal set of data retrievals. The components they would need to specify are:

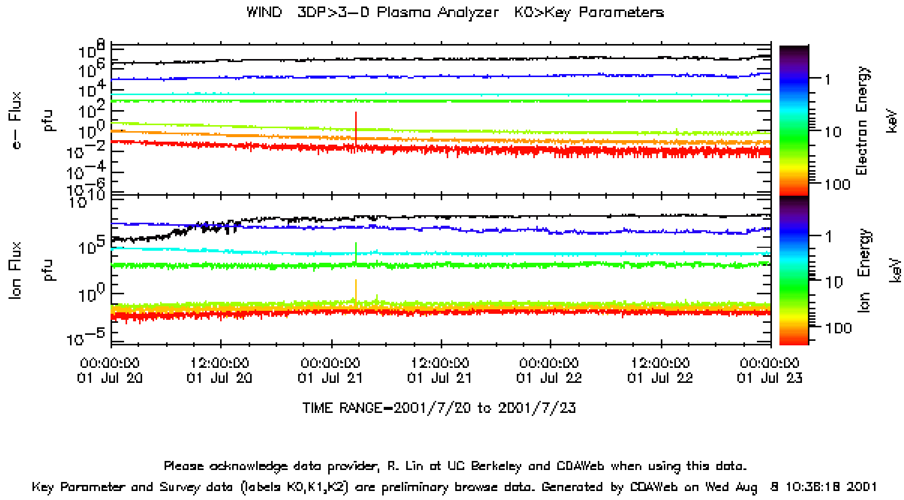
- Timing of data retrievals.
- Which data products to retrieve.
- Parameters for model evaluation or data selection e.g. time intervals, aircraft routes, orbits, geographical areas.
- Format of data products retrieved e.g. CDF/XDF data file, plots.
- Mode of data retrieval e.g. HTTP, FTP, e-mail.

'Data retrieval' is here assumed to encompass actions ranging from fetch a simple index such as Kp, through generating a plot of some parameter and viewing it on a web browser, to initiating a model run and returning a set of model results. Each user would need a personal password-protected 'account' that would store the details of their data retrieval profile and any other persistent information. It is possible that these user accounts would be most easily implemented as real user accounts on the SWS computer system, but in any event they should be administered with a similar degree of care because of the resource implications of each additional user.

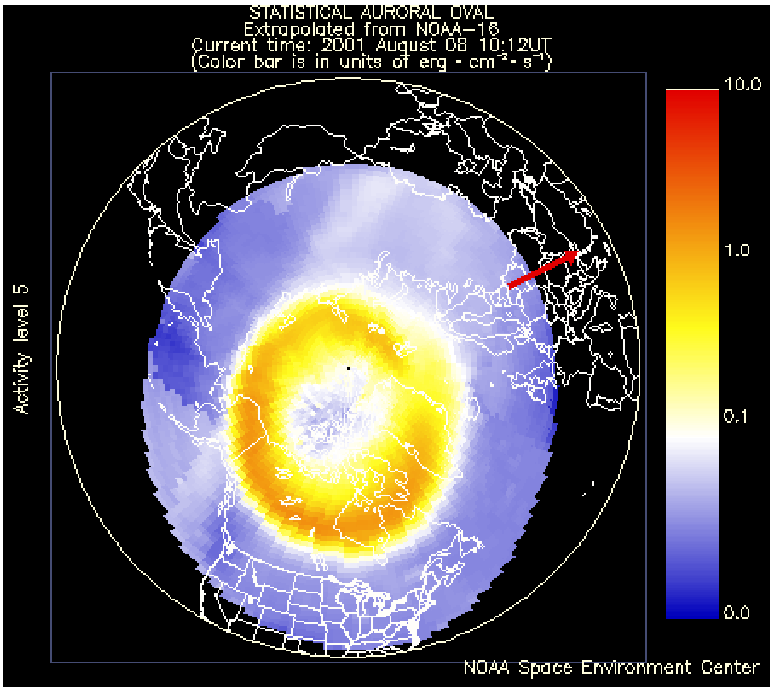
All the facilities listed here are targeted at the relatively sophisticated community of service providers. For general public access there could be a 'dummy' account that carried out a selection of actions of general interest and posted resulting plots or data listings to part of the SWS web site.

Other on-line facilities should include

- On-line help – this includes general information on how to use the system plus structured access to the metadata database.



(a) Plasma data from the WIND spacecraft, plotted using CDAWlib



(b) Statistical auroral oval, derived from NOAA-16 data by SEC

Figure 3: Example data plots

- A bug reporting and suggestions system.
- A frequently updated source of relevant Space Weather news, along the lines of ESA's existing SWEN newsletter.
- A system for requesting further help from SWS personnel.

SFR	SFR Description	SIP	SIP Description
12.1	Enable users to build personal sets of data retrieval actions	12.1.1	For all actions, users must be able to specify and store when the action should occur and its repeat interval.
		12.1.2	Allow the format of retrieved data products to be user-specifiable (CDF/XDF/plot)
		12.1.3	Allow the data transfer protocol to be user-specifiable (HTTP/FTP/SMTP).
		12.1.4	For custom model invocation, allow users to supply a model-specific set of parameter values.
		12.1.5	User profiles must be password protected so that they are accessible only to the controlling user.
3.1	Provide on-line help for users	3.1.1	Provide context-specific help on how to set up data retrievals
		3.1.2	Provide an on-line route to SWS support staff to request further help with operating the system.
18.1	On-line mechanisms for user feedback	18.1.1	Provide a bug-reporting on the on-line system
		18.1.2	Provide a user suggestions or discussion forum

5.2 Static materials for distribution

As discussed in Section 2.2, besides supporting service providers the SWS should have a role in informing and educating a broader community of interested parties. The consideration of on-line data services above suggested generating standard Space Weather outputs of general interest and posting them to a portion of the SWS web service. There would also need to be some static web pages introducing the science of space weather and the impacts it has on various technological sectors: airlines, electric power, geological/offshore, HF radio, insurance and financial services, military, satellites, tourism.

This ground is covered by several other web sites such as the SEC [36], the Space Weather Center [37] and the Lund Space Weather Science Center [38], but if the SWS is to be a serious player on the global Space Weather scene then this sort of general material must be presented. Since much relevant material has already been assembled on-line, it would make sense to build on that where possible, with the Lund site being an obvious starting point. As a European initiative the SWS should aim to emphasise European instances of technological impacts and to highlight ESWP spacecraft that provide relevant data – the existing general sites tend to be oriented towards a US readership.

There is also a need for static material for distribution such as paper newsletters and flyers, educational packs for schools, exhibits and displays. These materials and accompanying activities are covered in detail in Section 6.

SFR	SFR Description	SIP	SIP Description
14.1	Easy on-line access to introductory material on the science and impact of space weather.	14.1.1	Provide a set of tutorial web pages; the Lund site [38] is a good starting point.
		14.1.2	Construct a general user account that generates a selection of mainly graphical data products to accompany a static tutorial.

5.3 Human contact

Although the primary emphasis of this document has been on describing how to provide automated access to data products, there will be occasions when users will need direct contact with SWS personnel. The most frequent cases are likely to be when some assistance is needed in interpreting a data product or instruction required on how to use a model. These examples are relatively undemanding and could probably be dealt with by support staff whose primary role was in managing the computing systems.

More significant will be the less frequent requests from service providers needing assistance in developing specific tailored products for users or in understanding some aspect of the physics of the space weather environment. Meeting such requests needs a greater depth of understanding of the science than computing specific support personnel are likely to possess, so this calls for staff with scientific support as their main focus.

Another role for scientific support staff would be in providing a personal interface for the user feedback essential to proper service development. It would be valuable for SWS personnel to make regular personal contact with both the scientific institutions in Europe developing (and possibly operating) the models needed by the SWS, and representatives of the service providers and end-users. In addition to making regular contact visits to individual institutions or companies, an effective way of doing this would be to have a number of scientific and applications working groups, each focusing on a part of the science or applications of space weather. Their role would be to provide a forum for discussing the existing SWS services and for feeding in the views of the user community as to how the services should be developed in the future. Should such a working group already exist for a particular sector, relevant SWS personnel should liaise with it rather than create a duplicate body. The forward-looking role of the scientific support staff would also involve monitoring the relevant scientific literature and attending pertinent events, such as the SEC ‘Space Weather Week’.

SFR	SFR Description	SIP	SIP Description
5.1	Human backup to the on-line help systems.	5.1.1	Provide a telephone and e-mail help-desk with defined response time characteristics
18.2	SWS development to reflect the interests and concerns of the user community and the evolving state of STP.	18.2.1	SWS staff to make regular visits to scientific institutions and service providers
		18.2.2	Run working groups focused on specific parts of the science and applications of space weather.

6 Outreach

Education and Public Outreach (EPO) has become a growing concern within ESA, partly as a result of a Saatchi & Saatchi report that revealed that the general public interest in space was not matched by an awareness of the activities of ESA. Even in Europe awareness of ESA fell far short of that achieved by NASA.

Space weather is a natural topic in which to interest the public, since it is concerned precisely with how the space environment has direct physiological effects on people and how it impacts the technological systems that people rely on every day. The SWS would need to take a key role in promoting awareness and understanding of Space Weather in Europe. Since ESA already undertakes significant Public Relations (PR) and Outreach activities much of the outreach work of the SWS would be delivered through these channels. The SWS role would be in developing appropriate materials and in co-ordinating ESA activities with those of the scientific institutions in member countries that are engaged in space science work. There are four broad sectors that would need to be targeted: the general public, schools, universities and industry.

6.1 General public

General public awareness of space weather would be best addressed through central ESA channels, in the form of advertising, PR campaigns tied to mission launches or significant space weather events. The SWS PR staff would need to feed relevant facts and graphics to the central ESA PR function. In the case of significant space weather events, there would need to be material pre-prepared so that it was available at short notice.

There are also opportunities to present space weather in more depth. These include

- Temporary exhibits for use in museums or other venues. NASA's Space Weather Center has developed such an exhibit [39] that can be hired out to any interested parties at the rate of $9,500 for 12 weeks.
- Making presentations at suitable events. For example, the UK has a National Science week organised by the research councils with an extensive program of talks and events for the public at many venues nationwide. The SWS could provide speakers or provide materials and encouragement for local scientific institutions to put forward speakers for such events.

SFR	SFR Description	SIP	SIP Description
15.1	To raise general public awareness of space weather and its impacts.	15.1.1	Prepare template information for press releases to be used when a space weather events is newsworthy.
		15.1.2	Commission/design portable exhibits for public display.
		15.1.3	Prepare material and personnel for giving public presentations.

6.2 Schools

There are obvious opportunities to develop short study units based on space weather topics. There is scope for both stand-alone activities (such as the PopMagNet project [40] run by the University of York in the UK) and for activities tied to the SWS on-line services. In both cases they should be supported by classroom materials for teachers and students.

SFR	SFR Description	SIP	SIP Description
15.2	To promote awareness of space weather and space science in schools.	15.2.1	Develop short study units based on space weather topics.

6.3 Universities

ESA already has an outreach effort based at ESTEC termed the 'Education Office' [41]. This has a mission 'to promote and reinforce European educational excellence in the space field'. It does this by promoting space-related topics in curricula, encouraging awareness of space technologies amongst students, organising work experience opportunities at ESA, national space agencies and industry, and seeking to coordinate general information technology to facilitate educational outreach. The SWS should work to ensure that space weather features in the work of this office.

SFR	SFR Description	SIP	SIP Description
15.3	To promote awareness of space weather in universities.	15.3.1	Collaborate with the ESA Education Office.

6.4 Industry

The most effective path to develop industry interest would be to work with the space weather service providers that target various technological sectors. These organisations already have the contacts and the experience to do this, but for them to be able to state that they were working in partnership with an supra-national SWS would probably be advantageous. The sector-specific working groups established by the SWS (Section 5.3) would provide one route for making connections with relevant people and organisations; more general opportunities would arise at industry conventions, for which display materials could be made available.

7 Operations and Management

This section considers the management and operations of the SWS and includes some estimates of the cost of setting up and operating the various elements of the system discussed in previous sections of this document.

7.1 Management

The management of the SWS should be based on the well-understood principles of quality system management as outlined in the ISO 9000:2000 standard [42]. This standard describes various elements that the management of a system is well-advised to contain if the system is to deliver a quality product, which includes the delivery of data products and services such as those offered by the SWS. The requirements of the standard fall into five broad categories: Systemic, Management, Resource, Realization and Remedial. The Systemic and Management requirements relate to the establishment and operation of a Quality Management System; the general requirement is that there be a documented quality system requiring operations to be carried out as described in documented procedures.

The other categories relate more directly to product delivery, and some of their implications for the SWS are drawn out below. Section numbers given in parentheses indicate the sections of the ISO 9001 standard that apply.

7.1.1 Resource requirements

The requirements on resources can be summarised as the providing adequate resourcing for the SWS to be effective. In particular:

- Competent personnel are vital (6.2.1). There are three broad flavours of competency required by SWS staff – technical, scientific and promotional. The computing support staff must be technically proficient in the relevant areas, there must scientific support personnel who are able to handle scientific queries from service providers, and the promotion of public understanding of space weather by the SWS should be carried out by staff selected with that role in mind.
- Competence must be maintained (6.2.2), so a programme of staff development and training is necessary.
- A quality infrastructure must be provided (6.3). One implication for the SWS is that there must be a commitment to resource the on-going operation and development of the service over an extended period. The service will not gain the confidence and respect of users if its ability to develop and respond to their concerns does not match its initial scope.

7.1.2 Realization requirements

These requirements deal with the substance of service and product delivery and consequently generate a number of specific requirements on the SWS:

- Interactions with users must be controlled (7.2). In particular data product requests must be logged and a record kept of the system activities that occurred to meet them.
- Data products generated should be validated, either individually or on a sample basis (7.5.2).
- Data inputs from other data providers should be validated (7.4.3).
- The operations of models run remotely for the benefit of the SWS need to be controlled (7.4.1).
- User-supplied information, such as contact details and their profiles of data retrievals must be kept private from other users (7.5.4). The account details that are protected by password should not be divulged to other users.

7.1.3 Remedial requirements

- User satisfaction needs to be monitored and measured (8.2.1).
- Problem reporting and correction (8.3). The principal reporting mechanism is the on-line bug reporting procedure, and there must be a system in place to track the progress of work on reported bugs up and report back bug fixes

7.2 Operations

The consequences of the management requirements of the previous section are drawn together here with additional elements from elsewhere in the document, to present an overview of the required operational elements and procedures of the SWS.

7.2.1 Computing Infrastructure

The SWS should be a high-availability system operating 24 hours a day. This requires a clustered computing system with built-in fail-over capabilities, and the hardware should be under a same-day attendance maintenance contract with the manufacturer. Reliable and high-capacity network links are also essential given the constant flow of data to and from the system. Where a particular data feed is important it is not usually the case that a dedicated link is the most effective or cost-effective means of guaranteeing a reliable connection. It is usually better either to determine fixed routes for data traffic that can be relied on, or to seek to invest in those parts of the general network where lack of capacity has the greatest impact on the end-to-end capacity.

For security of local data, data will be stored on disk and tape, with regular daily backup of volatile and new data to tape from disk.

The computer systems will need to provide the standard internet data exchange protocols of HTTP, FTP and SMTP, with unnecessary communication services disabled for security reasons. The enabled servers need to be configured with due regard for security, with firewalling to prevent unauthorised access and sufficient logging to enable detection of intrusion attempts.

SFR	SFR Description	SIP	SIP Description
1.6	The on-line data must be available 24 hours.	1.6.1	Use clustered computing systems with intrinsic fail-over capabilities.
		1.6.2	Seek to eliminate network bottlenecks by supporting capacity upgrades.
		1.6.3	Purchase a same-day call-out maintenance contract for computing hardware.
1.7	The data archive must be secure.	1.7.1	Carry out daily back-up of volatile data.
		1.7.2	Apply the usual security measures applicable to publicly accessible networked computers.

7.2.2 Personnel

The computing systems should be supported by personnel competent to deal with software problems and identify when external hardware support is required. The local support staff should be contactable 24 hours; on the assumption that the computing systems would be generally reliable, out-of-hours cover would best be provided by an on-call system. The call-out system can be automated by using local software procedures to monitor system functions, and by using commercial remote services to monitor machine state. A minimum of three full-time staff members would be required to support the computing operations, allowing for planned or unplanned absence. Additional staff would be required while the service was being set up because of the greater demands for software development during that phase.

To support the sophisticated users such as the existing space weather service providers in Europe will require more than just a fully computerised system. There is also a need for scientifically expert support staff who can liaise with external organisations to develop the suite of models and forecast services, as described in Section 5.3. Given the range of models listed in Table 6 as necessary for a minimum initial service multiple personnel would be required:

- 3 for the plasma and radiation environment
- 1 for the solid body environment
- 2 for the magnetic environment

- 3 for the atmospheric environment

Section 6 presented a wide range of possible education and outreach activities and materials that the SWS could usefully generate. The nature of this outreach work is such that the greater part of the production of materials would be carried out either by other sections of ESA or external sub-contractors. The SWS must be able to act as a source of ideas and material and to provide co-ordination for the outreach activities and this would require dedicated staff with an interest and flair for PR activities, accompanied by a thorough understanding of the science.

All staff must have timely access to relevant training opportunities determined by developments in computing technologies, the state of solar-terrestrial physics and the information needs of the wider community.

SFR	SFR Description	SIP	SIP Description
13.1	Be able to support the operation of the computing systems	13.1.1	Employ at least three staff members dedicated to computer support
		13.1.2	Computing support staff should be contactable 24 hours.
		13.1.3	Use local and remote automatic systems to monitor system availability
13.2	Be able to handle scientifically sophisticated user queries and develop the service appropriately.	13.2.1	Have a suitable number of staff dedicated to scientific support and consultancy: a minimum of 9 is suggested (Plasma and radiation environment – 3, Solid body environment – 1, Magnetic environment – 2, Atmospheric environment – 3)
13.3	Be able to promote public understanding of space weather	13.3.1	Employ at least one staff member whose primary focus is public outreach, in tandem with some scientific knowledge.
13.4	Staff competencies must develop in line with changes in science, technology and the public environment.	13.4.1	Provide on-going training for all staff relevant to their roles.

7.2.3 Data inputs

The SWS will retrieve data from numerous on-line sources world-wide, with an initial estimate of data types and volumes made in Appendix A. While not being in a position to determine the data formats from these sources, the SWS should make data providers aware that their data is being used, and ask to be given prior notice of changes in data format or content.

Data retrievals should be logged, the retrieved data should be checked for the correct format and the data values screened for plausibility. Any check failures should be recorded and the system operator notified at an appropriate time.

For all retrieved data values it must be possible to determine the time at which the values were retrieved, so that the data environment at any given time can be recovered. On the structured files model of data storage proposed in Section 3.2, this can be achieved by adding retrieval time to the metadata accompanying the datafile.

SFR	SFR Description	SIP	SIP Description
1.8	Aim to acquire all relevant space weather data for input to the SWS.	1.8.1	Maintain good relationships with data suppliers by keeping them informed of data usage.
11.4	Ensure that retrieved data is of appropriate quality	11.4.1	Log all data retrievals
		11.4.2	Only ingest incoming data that has been checked as free of format errors
		11.4.3	Endeavour to check incoming data for plausible values

7.2.4 Data outputs

The SWS should seek to output data in a single standard format, although a plain ASCII output might need to be accommodated in addition to a richer structured format. Data outputs should be logged to maintain a record of the use of the SWS. Since some data outputs can be generated on demand it may not be practicable to check all data products before dispatch, but there should be a program of validating all model outputs. For large data products or for those that are only sampled by user requests, only representative samples of model or predictor outputs need be checked. The checking should ensure that the data values are reasonable given the data inputs and, in the case of predictors, regular assessments of accuracy should be made after the event to provide feedback for model enhancement.

SFR	SFR Description	SIP	SIP Description
12.2	Users should be able to determine what data products they have retrieved.	12.2.1	Log all data outputs from the SWS, centrally and by user.
		12.2.2	Allow users to inspect their personal retrieval logs.
17.1	Ensure that generated data products from models and forecasts are of appropriate quality	17.1.1	Validate at least some of all model outputs against data
		17.1.2	Periodically assess the accuracy of predictions

7.2.5 Data processing

The principal data processing activity of the SWS is the operation of models or predictors, either locally or remotely at cooperating institutions. The configuration of each model should be documented, with both this document and the model's source code under revision control.

The models should be set up to ingest and deliver data from and to the SWS in the standard SWS structured data format. In the case of remotely run models this will require a formal interface specification to be agreed between the SWS and the institution providing the remote model; this should be done under the auspices of the SWS quality policy, with formal change control.

The other data processing activity is handling user accounts and data requests. As indicated previously, the user accounts should be accessible over the web under password protection. Because the SWS would allow users to set up a personal profile of data retrieval activities, thus imposing additional load on the SWS computer systems, it would be necessary to have user registration with these approved by SWS staff rather than automatically. User data requests and changes to their personal profiles should be logged. The only exception would be for a generic public account that extracted and displayed selected model outputs on a regular basis for unrestricted access.

SFR	SFR Description	SIP	SIP Description
10.4	Ensure that generated data products from models and forecasts are of appropriate quality	10.4.1	Document the configuration of all models, whether run locally or remotely.
		10.4.2	Maintain model documentation and source code under formal revision control

7.3 Costing

This section present illustrative costings for establishing and running the SWS in the form presented in this document. The figures should be considered only as a rough guide to the costs involved – some part of the real costs are dependent on the level of usage of the service and on the costs of procuring computer and networking facilities, which historically have decreased substantially in real terms over time. All costs are shown in KEuro.

7.3.1 Start-up costs

The set-up costs are composed of the purchase costs of the computing equipment and the initial costs of software development. The example computing system is a cluster of four Compaq DS20 AlphaServers running Tru64 Unix, with disk storage provided by a 1400Gb RAID system and tape storage by a modular AIT-2 tape library. The initial software development costs would depend on the degree to which the necessary models were out-sourced to co-operating institutions. The majority of the initial software development work would be carried out by the permanent staff but the initial development would require additional manpower, estimated as 5 man-years of effort. For this estimate, total staff costs are given as 130 KEuro/year which is the same as the development rate for spacecraft ground segment [6]. It is assumed that for the SWS the development and operational staff would be the same personnel, so there is no difference in the rates for the two phases. All costs are shown in KEuro.

The RAID system has the capacity to hold several years' supply of input data, as estimated in Appendix A. No estimate has been made of the volume of generated data products that would be stored, but the disk storage option has been selected with plenty of spare capacity. The AIT tape system is modular, allowing up to 5 modules; each tape stores 50Gb, so a 19 slot module represents roughly 1Tb of storage. The cost of IDL licenses is the nominal list price – discounts are usually negotiable for government or academic users.

7.3.2 Operating costs

The principal running costs are staff salaries, the lease of an internet connection,, the maintenance contract on the computer systems, and the depreciation on the computing hardware. Prompt call-out maintenance contracts typically cost 10% of the hardware costs so the maintenance contract is costed appropriately, based on the initial computing equipment spend of 250 KEuro. Computer equipment is customarily written off over no more than 5 years, so 20% of the initial spend is designated for capital depreciation.

The cost of a leased line of a given capacity and length varies widely across the EU and has typically been decreasing in cost by 10% a year for the last three years [15]. The figure quoted in the table below should

Item	Cost/Item	Number	Total
Hardware			
Server – Compaq DS20E (2 CPUs)	35.0	4	140.0
Cluster kit – Base unit	13.0	1	13.5
Cluster kit – server cards	2.0	4	10.0
Disks – 1400Gb Compaq FibreChannel RAID 5400	65.0	1	65.0
Tape library – AIT-2, 2 drive, 19 slots, base unit	20.0	1	20.0
Tape library – additional module	6.5	1	6.5
Tapes – 50Gb capacity	0.1	40	4.0
Software			
IDL licenses – unlimited	37.5	1	37.5
Staff			
Software development	130.0	5	650.0
Miscellaneous			
Portable public exhibits	20.0	2	40.0
Total			986.5

Table 7: Set-up costs

therefore be regarded as illustrative rather than definitive. The final line, 'Miscellaneous consumables', allows 3 KEuro per person per year to cover any miscellaneous expenses for travel, photocopying etc.

Item	Cost/Item	Number	Total
Staff – Scientific	130.0	9	1170.0
Staff – Computing	130.0	3	390.0
Staff – PR	130.0	1	130.0
Leased line rental	25.0	1	25.0
Computer maintenance	25.0	1	25.0
Depreciation	50.0	1	50.0
Miscellaneous consumables	3.0	13	39.0
Total			1829.0

Table 8: Running costs

8 Future Development

The form of Space Weather Service outlined in this document has been one which could be established in the near-future to serve a present need. Over time the needs of the user community will develop as will the capacity of the science of space weather to meet them. This section sketches out some of the ways in which the SWS should evolve, scientifically, technically and organisationally.

8.1 Scientific development

Section 4.2.1 presented the initial set of models that the SWS should offer, derived from the set of models identified in the SRD. The models listed in Table 6 are those classified as 'immature' in the SRD that would have wide application – models to predict Kp and other indices using other models, a numerical model of solar proton events based on time series analysis, and a probabilistic model of geo-effective solar wind structures.

Model	**T**	**M**	**Update period**	**UR**
Numerical models of solar protons based on time series, e.g. work by Southampton University [43]	F	S	hourly	1 9 13 17 20
Probabilistic models of geo-effective solar wind structures e.g. the Cargill method [44]	F	S	1 minute	4 7 9 13
Numerical models for Kp forecasting	F	S	3 hourly	4 7 9

Table 9: Future models for the SWS to provide

The relatively small number of such models is testament to the obstacles facing better space weather prediction at present. These include poor understanding of the detailed dynamics of the processes in the sun that drive the space weather chain and the difficulty of adequately instrumenting the region of space between the sun and the earth. Both of these areas of difficulty will only be addressed satisfactorily on long time-scales. This poses problems for some prediction techniques, in particular the artificial intelligence approach using neural networks.

Present neural network prediction tools have proved successful at characterising the behaviour of magnetic indices on timescales of a few hours. It would appear that the magnetosphere exhibits characteristic modes of behaviour on these timescales which the networks are able to identify from the time series of the indices; beyond a few hours, the significance of the external drivers of the system becomes ever greater. There is a tendency for such models, therefore, to be good at prediction under quiet conditions or at tracking the results of a disturbance, but poor at predicting the onset of disturbed conditions. The presence of solar wind monitors at L1 has alleviated this to some degree, since there is now up to a couple of hours warning of changes in the solar wind conditions. Further improvement of the same models will require instrumentation closer to the sun, but this approach still limits prediction to at most 2-3 days ahead. To go beyond that will require a better understanding of the solar processes leading to CMEs, flares and other disturbances.

What the SWS *can* do is to encourage the evolution of models currently used for scientific purposes towards a state where they can be operationally useful. Two leading candidates for this are global models of

- the neutral and ionised atmosphere, such as CTIP [45]
- the magnetosphere, such as GUMICS [46].

Both of these cited examples are already leading the way in their respective fields in Europe, or even the world, and the concerted support of the SWS would maintain or enhance their eminence. The CTIP model has been included in an earlier section to provide gross atmospheric characteristics such as a daily estimate of atmospheric drag on LEO spacecraft (Table 6); the role envisaged here is more demanding, requiring prompt modelling of more dynamic behaviour driven by fluctuating magnetospheric conditions.

With several good models of different regions of the space weather environment readily available, a further natural development would be to encourage the interconnection of these various models as the next step in developing a coordinated model of the full space weather environment, from the solar surface (or below) through the solar corona, solar wind, Earth's magnetosphere and ionosphere down to the Earth's surface or crustal current systems. For a task of this complexity, involving feedback between models of different regimes operating under a variety of modelling assumptions, sophisticated technical computing facilities would be required, as discussed in the following section. A valuable first stage would be the connection of the CTIP and GUMICS models.

8.2 Technical development

The volumes of data required and generated by the SWS are small by comparison with those from the international scientific collaborations that are driving the current 'Grid' initiatives in Europe and the US [47], such as the LHC at CERN and SLAC. The greatest technical challenges faced and posed by the SWS are in the area of model interconnection.

There are numerous protocols and software libraries available for facilitating inter-process communication and synchronisation, whether local or distributed. Some of the more prominent of these are the High Level Architecture (HLA) [48], Globus [49], CORBA [50] and Jini [51] :

HLA Developed by the US DoD originally to allow individual simulations to cooperate to simulate a larger system. It works by building a 'federation', where each federate is a distinct simulation or model. The HLA stipulates a set of rules that a federate must obey, provides a set of run-time services and interfaces to these services in many languages (C++, Java, Ada, CORBA IDL). There is comprehensive provision for exchanging data between federates and for specifying complex rules for inter-federate communication and synchronisation. The HLA has been adopted by the IEEE as an open standard, and as the Facility for Distributed Systems by the Object Management Group (OMG) which has developed CORBA.

Globus The Globus project provides software tools to facilitate resource sharing among distributed computers. They provide a number of tools and libraries that address issues of security, resource and data management, communication and fault detection. The toolkit APIs are all in C but there are Java class wrappers for the most important components.

CORBA An infrastructure for distributed systems. It is a lower-level architecture than the HLA, and is also largely computer architecture and software independent. Its Interface Definition Language (IDL) has been accepted as an ISO standard and has mappings to C, C++, Java, COBOL, Smalltalk, Ada, Lisp, Python and IDLscript.

Jini A Java-based distributed systems architecture. It is specific to the Java language and covers much of the same ground as CORBA. The principal conceptual difference is that CORBA offers access to distributed objects resident on CORBA servers, whereas Jini allows one to offer distributed services, with proxy objects running on the clients.

All of these architectures for inter-process communication impose obligations and constraints on systems that use them. Their use is not widespread in STP modelling work at present, so their adoption would require a large measure of coordination from and encouragement by the SWS. The HLA is the most sophisticated technology described above but it is also the only one aimed specifically at simulation and modelling. Thus, although implementing it makes the greatest demands on model developers it offers the best prospect for connecting large-scale models because it has the rigour and sophistication to handle the complex rules on feedback between such models that would be necessary. As a first step, the SWS should support the provision of the relevant HLA interface routines in the CTIP and GUMICS models at their home institutions.

For relatively simple models the overhead of a sophisticated architecture like the HLA would be excessive. It would be better for such models to be provided with appropriate interfaces using Jini or CORBA to allow remote interrogation and invocation. The SWS should again act as an advocate of these relatively simple technologies.

Many of the issues raised by operating distributed models are also being tackled by the so-called 'Grid' initiatives in Europe and the US [47]. The lower-level technologies listed above (Globus, CORBA, Jini) are all being used actively in Grid work, and any developments from this field should be considered when developing the modelling environment of the SWS.

SFR	SFR Description	SIP	SIP Description
19.1	Foster collaboration between relevant groups in Europe.	19.1.1	Support the use of the HLA [48] in the CTIP [45] and GUMICS[46] models.
		19.1.2	Develop simple protocols for model invocation and interrogation and support their implementation using appropriate technologies (e.g. CORBA [50], Jini [51]).

8.3 Organisational development

The proposed SWS as presented here is a relatively small-scale operation, relying on the input and involvement of data suppliers and model developers to provide the services it offers. As the understanding of the physics of space weather improves, and in particular if it becomes possible to model the full interaction chain from the solar surface or below to the Earth, the complexity of useful models will increase. It will therefore become increasingly important to have a single European centre where definitive versions of such models reside, supported by the efforts of a multinational team of scientists and computing specialists. In this case, the size of the SWS operation would need to be increased significantly, with enhanced computing infrastructure and more personnel, some on secondment from national institutions. This approach is similar to that of the European Centre for Medium-range Weather Forecasting (ECMWF), which has close links with the national weather forecasting services of many European nations.

Given the European situation of geographically dispersed and complementary expertise in the various subdisciplines of space physics contributing to space weather, it is important for the SWS to provide a focus for all parties interested in space weather phenomena and effects. The details of the organisational structure which houses the service are of much less importance than the role it plays. Two possible scenarios are:

- An organisation along the lines of the ECMWF, whose role is entirely devoted to forecasting and modelling the space weather environment, or
- An organisation analogous to EUMETSAT, which runs the entire European space weather programme, from satellites through ground-support for these satellites through to dissemination of data and forecast products the the SWS.

Either approach would be satisfactory, although in practice the looser collaborative approach of the ECMWF would be simpler to establish in the short term.

SFR	SFR Description	SIP	SIP Description
19.1	Foster collaboration between relevant groups in Europe.	19.1.3	Aim to establish a programme of secondment of personnel from national institutes to the SWS, along the lines of the ECMWF.

9 Summary

This document has presented a detailed proposal for a European Space Weather Service. The key parts of the analysis are:

- A set of Service User Requirements that any Space Weather Service should meet (Table 2, page 13).
- A high-level design schematic of the components necessary to meet these requirements, and how they are interrelated (Figure 1, page 14).
- The Consolidated System Measurement Requirements from the SRD that constitute the 'raw' data feed to the SWS, with estimates of the data volumes involved (Table 10 in Appendix A).
- The initial set of models that the SWS should provide (Table 6, page 27).
- A set of detailed Service Functional Requirements and associated Service Implementation Proposals describing the features that the service must provide and suggested means of implementation (Table 11 in Appendix B).
- Estimates of the set-up and running costs of the service including suggested hardware purchases and staffing levels (Tables 7 and 8 on page 42).

References

[1] Study for an ESA Space Weather Programme, Proposal in response to ESA ITT AO/1-3533/99/NL/SB, Issue 1.1, 28th October 1999.

[2] Space Weather Market Analysis, Astrium Ltd. Report produced for ESTEC under contract no. 14069/99/NL/SB, Draft Issue 5.0, 31st January 2001.

[3] Lesley M Murphy and David J Rodgers. European Space Weather Programme System Requirements Definition. Technical Report DERA/KIS/SPACE/CR010157, DERA, February 2001. ESTEC/Contract No. 14069/99/NL/SB.

[4] Michel Kruglanksi and Daniel Heynderickx. Space weather service prototype. Technical Report WP 434 – BIRA-IASB, BIRA, July 2001. ESTEC/Contract No. 14069/99/NL/SB.

[5] http://www.aurorawatch.york.ac.uk/.

[6] Spacecraft Interface. Report produced for ESTEC under contract No. 14069/99/NL/SB, July 2001.

[7] Mike Hapgood. The interface between instrument ground-segments and the Space Weather Service. Technical Report ESWS-RAL-TN-0002, RAL, June 2001. ESTEC/Contract No. 14069/99/NL/SB.

[8] ISO. *ISO/IEC 9075-1:1999 Information technology – Database languages – SQL – Part 1: Framework (SQL/Framework)*, 2000.

[9] ISO. *ISO/IEC 9075-2:1999 Information technology – Database languages – SQL – Part 2: Foundation (SQL/Foundation)*, 2000.

[10] http://nssdc.gsfc.nasa.gov/cdf/cdf_home.html.

[11] http://sml.gsfc.nasa.gov/XDF/XDF_home.html.

[12] http://www.cacr.caltech.edu/SDA/xsil/.

[13] http://spdf.gsfc.nasa.gov/istp_guide/istp_guide.html.

[14] http://httpd.apache.org/.

[15] Commission of the European Communities, Brussels. *Sixth Report on the Implementation of the Telecommunications Regulatory Package, COM(2000) 814*, December 2000. See http://europa.eu.int/comm/information_society/policy/telecom/6threport/index_en.htm.

[16] http://crsp3.nrl.navy.mil/creme96/cm/refs.htm.

[17] http://www.cami.jccbi.gov/aam-600/610/600Radio.html.

[18] C Stormer. . *Z. Astrophys.*, 1:237, 1930.

[19] M A Shea and D F Smart. . In *18th International Comsic Ray Conference, Conference Papers*, volume 3, page 415, 1983.

[20] http://wwwinfo.cern.ch.asc/genat4/geant4.html.

[21] http://nssdc.gsfc.nasa.gov/space/model/magnetos/tsygan.html.

[22] S Bourdarie, D Boscher, M Blank, and J A Sauvaud. A physical 4D radiation belt model including a time dependent magnetic field. *Advances in Space Research*, 25(12):2303–2306, 2000.

[23] H Klinkrad, J Bendisch, H Sdunnus, P Wegener, and R Westerkamp. An introduction to the 1997 ESA MASTER model. In *Proceedings of the Second European Conference on Space Debris*, volume ESA SP-393, pages 217–224, May 1997.

[24] P Jenniskens. Meteor stream activity. *Astronomy and Astrophysics*, 287:990–1013, 1994.

[25] F Boberg, P Wintoft, and H Lundstedt. Real time Kp predictions from solar wind data using neural networks. *Physics and Chemistry of Earth, Part C*, 25(4):275–280, 2000.

[26] http://sol.irfl.lu.se/~fredrikb/Kp/.

[27] http://sec.noaa.gov/rpc/costello/index.html.

[28] http://nssdc.gsfc.nasa.gov/space/model/ionos/iri.html.

[29] International Telecommunications Union. *ITU-R Methods of basic MUF, operational MUF and ray-path prediction*, Recommendation P.1240 (05/97) edition, May 1997.

[30] http://www1.primushost.com/~cpibos/. Computational Physics, Inc., vendor of PRISM.

[31] R E Daniell, L D Brown, D M Anderson, and M W Fox. Parametrized Ionospheric Model: A global ionospheric parametrization based on first principle models. *Radio Science*, 30(5):1499–1510, Sep-Oct 1995.

[32] http://iono.jpl.nasa.gov/gim.html.

[33] E J Fremouw and J A Secan. Modeling and scientific application of scintillation results. *Radio Science*, 19:687–694, 1984.

[34] http://www.nwra-az.com/ionoscint/wbmod.html.

[35] http://www.sel.noaa.gov/pmap/index.html.

[36] http://www.sec.noaa.gov/Education/.

[37] http://www.spacescience.org/SWOP/1.html.

[38] http://www.irfl.lu.se/HeliosHome/spacew2.html.

[39] http://www.spacescience.org/SWOP/Exhibits/Mini_Exhibit/overview/.

[40] http://www.aurorawatch.york.ac.uk/popmagnet/.

[41] http://www.estec.esa.nl/outreach/office.html.

[42] ISO. *ISO 9000:2000, Quality management systems – Requirements*, 2000.

[43] http://www.ses.soton.ac.uk/projects/Astronautics/Gpatrick/gpatrick.html.

[44] J Chen, P J Cargill, and P J Palmadesso. Real-time identification and prediction of geoeffective solar wind structures. *Geophysics Research Letters*, 23:625, 1996.

[45] http://alpha.apg.ph.ucl.ac.uk/Docs/ctip_index.htm.

[46] http://www.geo.fmi.fi/~pjanhune/gumics3/.

[47] R Buyya and M Baker, editors. *Grid computing – GRID 2000 : 1st IEEE/ACM International Workshop. Proceedings, Bangalore, India, 17 December 2000*, volume 1971 of *Lecture Notes in Computer Science*. Springer, Berlin, 2000.

[48] http://www.dmso.mil/index.php?page=64/.

[49] http://www.globus.org/.

[50] http://www.omg.com/gettingstarted/corbafaq.htm.

[51] http://www.javacommerce.com/tutorial/jini.

[52] Mike Hapgood and Mike Oliver. A definition of instruments needed for space weather measurements. Technical Report ESWS-RAL-TN-0001, RAL, May 2001. ESTEC/Contract No. 14069/99/NL/SB.

A Data Storage Requirements

Table 10 lists the Consolidated System Measurement Requirements (CSMR) from the SRD and provides estimates of the daily data rate that each measurement entails. Where a CSMR duplicates another except for its temporal sampling requirement, only the highest sampling rate is listed; all lower rates can be derived from the highest rate sequence by temporal averaging or accumulation, so no further bulk storage is required. CSMR 18 is omitted because it was decided, after completion of the SRD, that these observations were no longer required (see section 2.5 of the instrument definition report [52]).

Two assumptions have been made about the size of data elements. Images are presumed to use 1Mb, this being derived as a 1024 $\times$ 1024 pixel array with one byte per pixel. All other numeric fields are assumed to be long real values needing 8 bytes of storage. This may be an overestimate for simple integer parameters such as $F_{10.7}$ or Sunspot number, but on average it is unlikely to be in error by a factor of more than two, whether numbers are stored in binary or ASCII.

The data storage regime employed would also affect the daily requirement for extra storage. For example, if a relational database were used then there would be some overhead per table row associated with the base data and any related indexes. If instead data were stored in a file tree in CDF or XDF format, then no allowance has been made for the metadata associated with the data items; it is assumed that data will be databased in sufficiently large units that any metadata overhead imposed by the storage format is small as a proportion of the total.

The final column of Table 10 shows the percentage of the total data flow that each set of measurements accounts for. It is readily apparent that the data rates are dominated by only a few of the measurements, namely those that are images, to the extent that many of the other percentages have rounded to zero. Over 70% of the data is attributed to the ground-based auroral imaging for which there is an assumed need for 30 auroral cameras producing hourly images of 1Mb in size to meet CSMR 5 (SMR 16.5). If the measurement requirement that these auroral images are intended to meet could be satisfied by a sparser network of imagers, or by using a meridian scanning photometer rather than an imager (as suggested in [52]), then this figure could be reduced by a factor of up to 1000 if MSPs were substituted for imagers. This is the only one of the imaging requirements where such economies may be possible.

The conclusion to be drawn from the figures presented in the table is that the daily data rate implied by the need to retrieve the space environment data is between 200Mb and 1000Mb per day.

Further data storage requirements are imposed if the SWS also has to serve as the archive for calibrated physical parameters produced by dedicated ESWP Space Weather missions. Many of the space weather requirements on physical parameters are for relatively low rate temporal sampling, for example, the various solar images are required only hourly and the $>$10MeV ion flux is required only half-hourly. It is likely, however, that any instrument flown to produce these parameters will sample at a much higher rate. The figures in the table can be used to estimate the daily data rates implied by greater temporal resolution and the implications for data storage calculated.

CSMR	Data item	Channels	Sites	Bytes	#/day	Total Bytes	% of data
1	Solar EUV images	2	1	1.05e+6	24	5.03e+7	4.88
1	Solar X-Ray images	2	1	1.05e+6	24	5.03e+7	4.88
2	Coronagraph image	1	1	1.05e+6	24	2.52e+7	2.44
3	Stereo UV/visible images	2	2	1.05e+6	24	1.01e+8	9.77
4	Auroral imaging (s/c)	1	1	1.05e+6	24	2.52e+7	2.44
5	Auroral imaging (ground)	1	30	1.05e+6	24	7.55e+8	73.27
6	Auroral oval description	1	1	4.80e+1	24	1.15e+3	0.00
7	AO equatorward bdy	1	24	8.00e+0	24	4.61e+3	0.00
8-11	X-ray flux	2	1	8.00e+0	1440	2.30e+4	0.00
12	UV flux	1	1	8.00e+0	288	2.30e+3	0.00
13	EUV flux	1	1	8.00e+0	288	2.30e+3	0.00
14-17	F10.7	1	1	8.00e+0	288	2.30e+3	0.00
19-21	2ndary neutrons (ground)	1	40	8.00e+0	24	7.68e+3	0.00
22	2ndary neutrons (air)	1	30	8.00e+0	288	6.91e+4	0.01
23-25	V_{sw}	1	2	8.00e+0	1440	2.30e+4	0.00
26-27	N_{sw}	1	1	8.00e+0	96	7.68e+2	0.00
28	Kp	1	1	8.00e+0	8	6.40e+1	0.00
29	Kp*	1	1	8.00e+0	288	2.30e+3	0.00
30	Ap	1	1	8.00e+0	1	8.00e+0	0.00
31	Dst	1	1	8.00e+0	24	1.92e+2	0.00
32	Dst*	1	1	8.00e+0	288	2.30e+3	0.00
33	AE index	1	1	8.00e+0	1440	1.15e+4	0.00
34-35	SSN	1	1	8.00e+0	1	8.00e+0	0.00
36-38	IMF	1	2	8.00e+0	1440	2.30e+4	0.00
39-43	B (magnetosphere)	1	20	2.40e+1	1440	6.91e+5	0.07
44	B (ground)	1	100	2.40e+1	8640	2.07e+7	2.01
45	IPS	1	3	8.00e+0	24	5.76e+2	0.00
46-47	foF2, foE, foF1	3	40	8.00e+0	288	2.76e+5	0.03
48-49	TEC	1	40	8.00e+0	288	9.22e+4	0.01
50	Cross-tail E-field	1	1	8.00e+0	8	6.40e+1	0.00
51	Ion drift velocity	1	10	8.00e+0	8640	6.91e+5	0.07
52	Cold ions	1	4	8.00e+0	1440	4.61e+4	0.00
53	1-10keV electrons	10	4	8.00e+0	1440	4.61e+5	0.04
54-55	10-100keV electrons	10	4	8.00e+0	1440	4.61e+5	0.04
56-58	$>$10MeV ions	25	1	8.00e+0	48	9.60e+3	0.00
59-61	$>$10MeV protons	5	3	8.00e+0	96	1.15e+4	0.00
62-65	$>$100MeV ions (spectra)	4	1	8.00e+0	24	7.68e+2	0.00
66-67	Relativistic electrons	10	3	8.00e+0	24	5.76e+3	0.00
68	Atmospheric scale height	1	20	8.00e+0	24	3.84e+3	0.00
69	Debris size + v-dist	1	1	8.00e+2	1	8.00e+2	0.00
70-71	Meteoroid size + v-dist	1	1	8.00e+2	1	8.00e+2	0.00
72	Dose rate + LET spectrum	5	10	8.00e+0	288	1.15e+5	0.01
73	Total dose	1	1	8.00e+0	1	8.00e+0	0.00
74	Satellite position	1	20	2.40e+1	48	2.30e+4	0.00
	GRAND TOTAL (Mb)					1.03e+9	100.00

Table 10: Measurement data rate

B Service Implementation Proposals

Table 11 brings together all the Service Functional Requirements (SFRs) and Service Implementation Proposals (SIPs) from throughout the document. The last column is a cross-reference to the document section where each SFR and SIP was identified and discussed. The major number in each SFR is the Service User Requirement (SUR) of which the SFR forms a part.

Table 11: Consolidated Service Implementation Proposals

SFR	SFR Description	SIP	SIP Description	Sec
1.1	Allow data retrieval using the FTP, HTTP and SMTP protocols	1.1.1	Use the FTP/HTTP agent GNU *wget* for FTP and HTTP retrieval and mirroring	3.1.2
		1.1.2	On Unix, the requirements are met by standard Mail Delivery Agents such as *sendmail* in conjunction with *procmail* and the mail client *mail*.	
1.2	Data retrieval mechanisms should be sufficiently flexible to cope with the systems of many different data suppliers.	1.2.1	Provide for data acquisition using all methods from Table 3 except the Remote/Local combination.	3.1.2
1.3	Data distribution mechanisms should be sufficiently flexible to cope with the needs of many different users.	1.3.1	Provide for data distribution using all methods from Table 3 except the Remote/Local combination.	3.3
1.4	Allow data distribution using the FTP, HTTP and SMTP protocols	1.4.1	Offer active FTP delivery of data (PUT operation) by using appropriate software, such as *expect* or the Perl LWP and Net::FTP modules.	3.3
		1.4.2	Offer dynamic HTTP access by running a forms-driven web service on the *apache* webserver [14].	
		1.4.3	Offer e-mail data distribution using a standard Mail Transport Agent such as *sendmail*, Mail User Agent such as *mail*, or wrapped by higher-level software such as the Perl Mail::Mailer module.	
		1.4.4	Consider novel data distribution channels (e.g. SMS) for selected products (e.g. alerts)	
1.5	Provide sufficient network capacity to allow prompt access to SWS data products by users	1.5.1	Procure a leased line connection with bandwidth of at least 5Mbits/s.	3.3

continued . . .

Table 11: (continued)

SFR	SFR Description	SIP	SIP Description	Sec
1.6	The on-line data must be available 24 hours.	1.6.1 1.6.2 1.6.3	Use clustered computing systems with intrinsic fail-over capabilities. Seek to eliminate network bottlenecks by supporting capacity upgrades. Purchase a same-day call-out maintenance contract for computing hardware.	7.2.1
1.7	The data archive must be secure.	1.7.1 1.7.2	Carry out daily back-up of volatile data. Apply the usual security measures applicable to publicly accessible networked computers.	7.2.1
1.8	Aim to acquire all relevant space weather data for input to the SWS.	1.8.1	Maintain good relationships with data suppliers by keeping them informed of data usage.	7.2.3
2.1	Data retrieval must be able to proceed on a pre-determined schedule without human intervention.	2.1.1	The Unix *cron* or Windows *schedule* facilities should be used to start retrieval jobs as required.	3.1.2
2.2	Data processing and distribution must be able to proceed on a pre-determined schedule without human intervention.	2.2.1	The Unix *cron* or Windows *schedule* facilities should be used to start data processing or distribution jobs as required.	3.2.3
3.1	Provide on-line help for users	3.1.1 3.1.2	Provide context-specific help on how to set up data retrievals Provide an on-line route to SWS support staff to request further help with operating the system.	5.1
4.1	Store data in a manner that facilitates the association of data with the necessary metadata for its proper use.	4.1.1	Store data in structured files with integrated provision for metadata	3.2.4
5.1	Human backup to the on-line help systems.	5.1.1	Provide a telephone and e-mail help-desk with defined response time characteristics	5.3
6.1	Output all data using a generic and accessible data format, for consistency of access.	6.1.1 6.1.2	Output data in CDF or XDF formats. Facilitate simple ASCII output by structuring XDF files so that the data element has sufficient internal structure for it to be intelligible in the absence of the metadata elements.	3.3
7.1	Allow the data holdings to be searched by metadata attributes	7.1.1	Extract datafile metadata and store in a relational database – a ‘Data Dictionary’	3.2.6
8.1	Allow easy configuration and access to location metadata	8.1.1	Store location metadata in a relational database – a ‘Yellow Pages Directory’.	3.2.6

continued . . .

Table 11: (continued)

SFR	SFR Description	SIP	SIP Description	Sec
9.1	Enable users to filter data by time	9.1.1	All data requests should specify a time range.	4.1
9.2	Enable users to apply simple aggregating operations to data streams.	9.2.1	Implement a suite of aggregation functions to compute mean, median, minimum, maximum and sum on arbitrary data fields over regular repeating intervals of time or space.	4.1
10.1	Offer a wide range of models and forecasting tools.	10.1.1	The models in Table 6 should all be provided.	3.3
10.2	Allow models to be interconnected	10.2.1	For the models in Table 6, passing standard output data files in CDF or XDF format is sufficient.	4.2.2
		10.2.2	Run models at the most appropriate location, which may be at a remote institution.	
		10.2.3	Maintain awareness of developments in the Grid field, and on possible applications within space weather modelling.	
10.3	Offer models to be run routinely or on demand as appropriate	10.3.1	The models with S(tandard) mode in Table 6 should generate standard outputs that are archived routinely.	4.2.3
10.4	Ensure that generated data products from models and forecasts are of appropriate quality	10.4.1	Document the configuration of all models, whether run locally or remotely.	7.2.5
		10.4.2	Maintain model documentation and source code under formal revision control	
11.1	Maintain local databases of retrieved data, estimated to accumulate at 200-1000Mb/day.	11.1.1	Use a RAID system with capacity of at least several months of data for immediate on-line access, with near-on-line tape storage holding the remaining data.	3.2.1
11.2	Store data so that the operational data environment at any point in time is recoverable at any later time.	11.2.1	Update data non-destructively.	3.2.2
		11.2.2	Timestamp all data files or values with their time of retrieval.	
11.3	Store data in a manner that simplifies ingestion, processing and distribution.	11.3.1	Store data in structured files rather than relational or object-oriented database systems	3.2.4

continued . . .

Table 11: (continued)

SFR	SFR Description	SIP	SIP Description	Sec
11.4	Ensure that retrieved data is of appropriate quality	11.4.1	Log all data retrievals	7.2.3
		11.4.2	Only ingest incoming data that has been checked as free of format errors	
		11.4.3	Endeavour to check incoming data for plausible values	
12.1	Enable users to build personal sets of data retrieval actions	12.1.1	For all actions, users must be able to specify and store when the action should occur and its repeat interval.	5.1
		12.1.2	Allow the format of retrieved data products to be user-specifiable (CDF/XDF/plot)	
		12.1.3	Allow the data transfer protocol to be user-specifiable (HTTP/FTP/SMTP).	
		12.1.4	For custom model invocation, allow users to supply a model-specific set of parameter values.	
		12.1.5	User profiles must be password protected so that they are accessible only to the controlling user.	
12.2	Users should be able to determine what data products they have retrieved.	12.2.1	Log all data outputs from the SWS, centrally and by user.	7.2.4
		12.2.2	Allow users to inspect their personal retrieval logs.	
13.1	Be able to support the operation of the computing systems	13.1.1	Employ at least three staff members dedicated to computer support	7.2.2
		13.1.2	Computing support staff should be contactable 24 hours.	
		13.1.3	Use local and remote automatic systems to monitor system availability	
13.2	Be able to handle scientifically sophisticated user queries and develop the service appropriately.	13.2.1	Have a suitable number of staff dedicated to scientific support and consultancy: a minimum of 9 is suggested (Plasma and radiation environment – 3, Solid body environment – 1, Magnetic environment – 2, Atmospheric environment – 3)	7.2.2
13.3	Be able to promote public understanding of space weather	13.3.1	Employ at least one staff member whose primary focus is public outreach, in tandem with some scientific knowledge.	7.2.2

continued . . .

Table 11: (continued)

SFR	SFR Description	SIP	SIP Description	Sec
13.4	Staff competencies must develop in line with changes in science, technology and the public environment.	13.4.1	Provide on-going training for all staff relevant to their roles.	7.2.2
14.1	Easy on-line access to introductory material on the science and impact of space weather.	14.1.1	Provide a set of tutorial web pages; the Lund site [38] is a good starting point.	5.2
		14.1.2	Construct a general user account that generates a selection of mainly graphical data products to accompany a static tutorial.	
15.1	To raise general public awareness of space weather and its impacts.	15.1.1	Prepare template information for press releases to be used when a space weather events is newsworthy.	6.1
		15.1.2	Commission/design portable exhibits for public display.	
		15.1.3	Prepare material and personnel for giving public presentations.	
15.2	To promote awareness of space weather and space science in schools.	15.2.1	Develop short study units based on space weather topics.	6.2
15.3	To promote awareness of space weather in universities.	15.3.1	Collaborate with the ESA Education Office.	6.3
16.1	Provide the capability of generating plots from data outputs	16.1.1	Use the generic data visualisation package IDL.	4.3
		16.1.2	If using CDF, use the CDAWlib plotting routines on top of IDL.	
17.1	Ensure that generated data products from models and forecasts are of appropriate quality	17.1.1	Validate at least some of all model outputs against data	7.2.4
		17.1.2	Periodically assess the accuracy of predictions	
18.1	On-line mechanisms for user feedback	18.1.1	Provide a bug-reporting on the on-line system	5.1
		18.1.2	Provide a user suggestions or discussion forum	
18.2	SWS development to reflect the interests and concerns of the user community and the evolving state of STP.	18.2.1	SWS staff to make regular visits to scientific institutions and service providers	5.3
		18.2.2	Run working groups focused on specific parts of the science and applications of space weather.	

continued . . .

Table 11: (continued)

SFR	SFR Description	SIP	SIP Description	Sec
19.1	Foster collaboration between relevant groups in Europe.	19.1.1	Support the use of the HLA [48] in the CTIP [45] and GUMICS[46] models.	8.2
		19.1.2	Develop simple protocols for model invocation and interrogation and support their implementation using appropriate technologies (e.g. CORBA [50], Jini [51]).	
		19.1.3	Aim to establish a programme of secondment of personnel from national institutes to the SWS, along the lines of the ECMWF.	8.3