Environment Canada / Environnement Canada
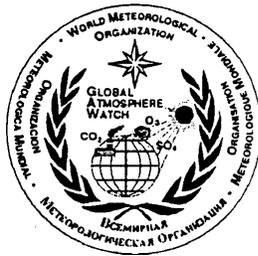
Guide to the WMO/GAW
World Ultraviolet Radiation Data Centre

Version 6.0

Environment Canada

September 2006

Canada

Guide to the WMO/GAW
World Ultraviolet Radiation Data Centre

Environment Canada*

Version 1 - December 10, 1996
Version 2 - January 2, 1997
Version 3 - May 10, 1997 (Version 3.2 - June 9, 1997)
Version 4 - November, 1997 (Version 4.2 - February 1998)
Version 5 - April-August 1998 (Version 5.1 - January 1999
Version 6 – September 2006)

FOREWORD

This document describes the procedures for data submission and information extraction at the World Ultraviolet Radiation Data Centre (WUDC), which is a subset of the World Ozone and Ultraviolet Radiation Data Centre (WOUDC). This document is intended as a guide for data originators and clients. The WOUDC web site is the primary medium for access to data, documentation and reports related to ozone and UV studies. As the WUDC is further developed, this guide will be updated accordingly and will continue to posted electronically on the WOUDC web site.

The current version, Version 6.0, represents a considerable re-writing of the previous document. Version 6.0 has been condensed and makes use of web links to include information that is independently updated, thus keeping the guide as dynamic a document as possible. Valuable input from the WMO Science Advisory Group for UV (SAG-UV) along with the originators of data and the data client communities have been incorporated into this document.

The use of terminology throughout this text will interchange between the acronym WOUDC and the acronym WUDC. Issues that refer to the general aspects of the Data Centre such as data reception or access, will use the acronym WOUDC. Issues specific to ultraviolet radiation such as form and content, will use the WUDC acronym.

The WOUDC web site is used to update information and serve as a virtual "bulletin board". As this guide is further developed, use the web site to retrieve the latest version.

E.W. Hare*, V.E. Fioletov*, D.I. Wardle* and E.J. Carty[++]

[++] Private Consultant, Guelph, Canada

TABLE OF CONTENTS

Guide to the WMO/GAW World Ultraviolet radiation Data Centre

1. INTRODUCTION

The World Ozone and Ultraviolet Data Centre (WOUDC) began as the World Ozone Data Centre (WODC), starting in 1961. In 1992, Environment Canada agreed to accept ultraviolet radiation data as part of the data centre and the WODC changed its name to the WOUDC. The WOUDC continues to be funded and operated by the Canadian Government (Environment Canada) and has been since its inception. Although the WOUDC is operated by Environment Canada, it works under the auspices of the Global Atmosphere Watch (GAW) programme of the World Meteorological Organization (WMO). The WOUDC gets guidance from the WMO Science Advisory Groups (SAGs) for both UV (WUDC) and Ozone (WODC).

For the purposes of this guide book, the acronym, WUDC will be used with the understanding that it represents the UV portion of the WOUDC. When referencing issues pertinent to both ozone and UV, such as data file formats, then the acronym WOUDC will be used.

1.1 HOW TO USE THIS GUIDE

This guidebook is intended for data originators who wish to submit data sets to the WOUDC and for users of the data who want to understand more about the data formats and content. The WOUDC uses the Internet as its main tool for data dissemination and so any updates to data products or changes occurring to the data centre as a whole will be posted on the WOUDC web site. Although this guide is intended to be used as a stand-alone document, changes are inevitable and the web site should be consulted for the most current version.

This Guide addresses, in general, questions about content, submission and conditions of access. Data content and particularly submission procedure will change and the Guide will be revised and updated, when necessary.   Data types are defined in Section 2 and Data levels in Section 3. The rules and rationale for data submission are discussed in Section 4 which will be of interest to data originators. The data quality control procedures are then outlined in Section 5, which describes what procedures the WUDC uses to accept and flag data. Finally in Section 6, data access is discussed with all the relevant Internet addresses for easy access to the data sets. This section will be of interest to data clients. Appendix A is brief overview of the development of the WUDC and some general background information.

1.2 ACCESS TO THE WOUDC WEB AND FTP SITES

The WOUDC web site is at: http://www.woudc.org and the WOUDC ftp site can be accessed either from the WOUDC web site or directly.

Address: ftp.tor.ec.gc.ca
Username: woudc
Password: woudc*

2 DATA TYPES

The WUDC terminology for data types is as follows:

Basic data: Spectrally resolved or weighted spectral integrals of UV radiation such as radiance, irradiance, irradiation as functions of time and wavelength or of time alone and possibly of direction (these data may be level 0, 1 or 2, refer to section 4).

Ancillary data: Data describing atmospheric variables and variable site characteristics such as temperature, ozone, pressure, cloud conditions for example by a cloud index or by a sky image, albedo, pyranometer data. Information on permanent characteristics of the measuring site.

Auxiliary data: Characteristics of the instrument that made the measurements, the way in which the instrument was operated, the algorithms that were used to compute the measured values from the instrument signals, etc. Auxiliary data may be potentially of value to the user or absolutely necessary as when the basic data is raw (Level 0).

Meta data: Information that uniquely describes sections (files) of basic data ('data about the data'), such as instrument and agency identification, location and category of the measurements, etc.

2.1 BASIC DATA

The following types of ultraviolet radiation measurements are considered acceptable for submission to the WUDC archive.

- Spectral global UV irradiance at the Earth surface expressed in $Wm^{-2}nm^{-1}$ by spectroradiometer-type instruments. The recommended spectral interval is 290-400 nanometres (nm), with a minimum requirement 300-315 nm. Possible subsets are 290-360 nm and 290-325 nm The spectral resolution should be 1.0 nm full width at half maximum (FWHM) or better.

- Spectral global UV irradiance at the Earth surface measured in $Wm^{-2}nm^{-1}$ at selected wavelengths in the UV part of the spectrum measured by Multi-band filter instruments. The requirement is for measurements at two or more wavelengths.

- Spectrally integrated global UV irradiance expressed in Wm^{-2} measured by a broadband filter instrument. The originator should provide information on the weighting function which should approximate one of the standard weighting functions (e.g. CIE).

- If available, independent measurements of the direct and diffuse components of the global irradiance should also be submitted.

- Actinometric (actinic) irradiance(flux), spectrally resolved or integrated, may also be submitted.

2.2 ANCILLARY DATA

- Pyranometer global solar radiation data in the broadband spectral range from 280 to 3000 nm in Wm^{-2}. The data file can include direct solar, downwelling and upwelling global solar irradiance for the same broadband spectral range from 280 to 3000 nm, albedo, and aerosol optical depth. Recommended temporal resolution is 1-10 minutes. Pyranometer data is defined as a WOUDC UV Category. Refer to Section 5 for details PAR (Photosynthetically Active Radiation) data are acceptable.

- Meteorological data that includes: surface temperature (in °C), surface pressure (hPa), cloud amount (in tenths), snow on the ground (in cm), stratospheric temperatures at 100, 50, and 30 hPa (in °C). Recommended temporal resolution is 12-24 hours.

- Atmospheric composition data such as total ozone, NO_2, SO_2. Recommended temporal resolution 12-24 hours.

- 2D (e.g., sky images in UV part of the spectrum) and 3D (e.g., spectrally resolved sky images) data files also can be submitted to the WUDC and can be distributed by the WOUDC, although algorithms for their control and processing by the WUDC are not established yet.

- Other types of UV-related data can be included in the WOUDC database. This may be negotiated between the data originator and the WOUDC.

2.3   AUXILIARY DATA

- The originator of the data is encouraged to submit any auxiliary data that he or she may consider relevant. The WUDC has defined required auxiliary data that will be discussed later in Section 4, Data Submission.

- Measurements of instrument characteristics are valuable regardless of whether the characteristics are described in text elsewhere. Such measurements include scans of radiation from lasers or spectral lamps that characterise the spectral sampling function of the measuring instrument and measurements of the directional sensitivity of any irradiance measuring instrument.

- The Agency Profile (which includes Station Information) and the Scientific Sponsorship Statement (see Section 4) are Auxiliary data.

- For the special case when "raw" Brewer UV files are submitted, the instrument responsivity files are required.

2.4   META DATA (FILE IDENTIFICATION)

- Meta data ('data about the data') are those data used by the WOUDC for file identification and tracking. Typically, the Agency Profile and the Station or Site information data are considered examples of 'metadata'. The requirements for meta data are given later in Section 4, Data Submission.

3   DATA LEVELS AND VERSIONS

The WOUDC has attempted to maintain as much consistency as possible with usage in the satellite data community and with other networks such as the Network for the Detection of Atmospheric Composition Change (NDACC), formerly known as the Network for the Detection of Stratospheric Change (NDSC), in choosing the following definitions of Data Levels for UV radiation data [2].

Level 0   Raw data. E.g. signals from radiation detectors in counts or converted voltage, signals from sensors or digital data that describe instrument configuration, as functions of time.

Level 1   Radiometric data, i.e. irradiance vs. wavelength (and time). The spatial, temporal and spectral characteristics of each datum are those of the original measurement.

Level 2   Spectrally gridded.
2a On a standard wavelength sampling interval such as (integer nm)/2 or (integer nm)/5.
2b With a standard passband e.g. trapeziod 1.0 nm full width at half intensity with 0.5 nm slope. (Note: this implies a base width of 1.5 nm and a flat peak of width 0.5 nm.)

Level 3   Spatially and/or temporally gridded and other high level products. These are not yet being addressed by WUDC.

In principle, Level 0 data are unchangeable. In practice they might be changed to correct gross mistakes like an assignment of the wrong time or location to a file. There are no versions of Level 0 data.

Versions and levels are entirely distinct. The level refers to the product while the version specifies the procedure used to generate the product from the Level 0 data. In the normal course of progress, research in measurement and modeling generates better algorithms and better instrument characterization. The non-zero level data can then be upgraded and then given incremented Version numbers (c.f. TOMS Version 7) and/or version dates.

Preliminary data are the result of processing that may be over-simple or lacking key information due to lack of time. The WOUDC, in common with other data agencies, asks that Preliminary data be designated as Version 0.

3.1   LEVEL 0 RAW DATA

- An instrument's output in the numerical form which cannot in principle be altered or corrected, is known as "raw data". Examples include photon counts, step counts, voltages and temperatures. An example of a raw data file is the UV-file generated by the Brewer instrument.
- In some instruments the signals are internally processed to provide output voltages or digital signals that directly represent the measurement. Also the internal processing may involve discarding some of the data. In either event whether such data should be classed as Level 0 or Level 1 is debatable.
- WOUDC encourages the submission of Level 0 data as an alternative or in addition to Level 1 data. For directions on submission of raw data to the WOUDC refer to Section 4, Data Submission.

3.2   LEVEL 1 PROCESSED DATA

- Level 1 data are the result of processing instrument signals to give numbers describing physical quantities in standard units such as $Wm^{-2}nm^{-1}$. Examples include irradiance, total column ozone and airmass. The processing may be simple or it may involve several stages including corrections for instrument stability and response, conversion of raw parameters to standard physical units or any other adjustment required to assure the data is of high quality.

- The spatial, temporal and spectral characteristics of Level 1 data are the same as those of the Level zero data from which they are generated. Thus gridded spectra and temporal or wavelength integrations of measured spectra are on higher levels.
- The data originator is responsible for applying appropriate, documented corrections and quality control coding to raw data in order to obtain the required processed parameter such as irradiance. If the output of an instrument is already the required processed parameter (i.e. irradiance calculated internally by the instrument), the quality control still remains the responsibility of the originator.

3.3   LEVEL 2 SPECTRALLY GRIDDED DATA

- It is expected that data will not usually be submitted in a "gridded' form. Gridding data may be done mostly by the WUDC, subject to approval by the originator, to produce UV data in a uniform format with the standard spectral range 290-400 nm and with a sampling interval of 0.5 nm. It is planned that data measured in restricted ranges (such as 290-360 nm or 290-325 nm) will be supplemented with simulated (and flagged) data in order to make a similar product for all spectra. With the data represented in a uniform format, there is an additional QA/QC step which provides value-added products such as quality flags and statistical summaries (e.g. daily means).
- Processing algorithms will be described in the WUDC documentation. If certain data transformations are required (e.g. re-calculation to a different wavelength grid), the WUDC will compute and distribute these data.
- File formats for the output data products are easy-to-use and in-line with other WOUDC data formats. Although some slight changes to the formats are possible, the WUDC expects that major changes will be infrequent.
- The WOUDC would be glad to accept Level 2 or higher data provided that the data set is accompanied by the appropriate Level 1 version of the data. In all cases format and quality control checks will be done by WUDC.

4   DATA SUBMISSION

This section will present a step-by-step data submission procedure which is essential to the data originator, but may also be of use to data clients who are interested in understanding how ultraviolet data are forwarded to the WOUDC.

The logical development of data submission is as follows:

1. Write a Scientific Sponsorship Statement (SSS) or "data passport" using the models given here (if appropriate). Contact the WOUDC for assistance or clarification if needed.
2. Submit to the WUDC a completed SSS along with an Agency (data originator) Profile. A template for the Agency Profile is available on the Web. If submission by means of the Web is not possible, data originators are requested to contact the WOUDC to arrange an alternative method of submission. Refer to section 4.2.1.
3. Upon approval of the SSS and Agency Profile, a data submission agreement will be established whereby the details of the file format for submission may be negotiated between the data originator and the WUDC. If special agreements, provisos, or security restrictions are required, they will be established at this point.
4. The WUDC will assign an ftp account (with a personal username/password) to be used exclusively for data submission. Data may also be submitted on computer media such as CD-ROM or diskettes.
5. It is recommended that data submissions be made monthly.

4.1   STEP 1: THE SCIENTIFIC SPONSORSHIP STATEMENT

The Scientific Sponsorship Statement (SSS) is a condensed scientific description of the origin, processing and quality of the data, including estimates of measurement uncertainty. The writer would normally be the scientist responsible for making the measurements but might be some other scientist who could fully describe the data. It is an ASCII file that is intended to be read by WUDC staff and data users, not by computers. Thus, there are no format rules.

The SSS should include the following:

- Instrument calibration information such as sources, frequency, stability, etc.
- Detailed data processing information about algorithms, data corrections, interpolation, and smoothing etc.
- Estimation of measurement uncertainties.
- References to publications that describe instruments, observation program, error estimations, results. (The WUDC will be grateful for receiving reprints or copies of these publications.)
- Conditions under which the originator permits use of the data.

A comprehensive list of sources of uncertainties and techniques for their estimation is provided in the WMO Guidelines for Site Quality Control of UV Monitoring [3]. The writer of the SSS is strongly encouraged to consult this QC document before starting his or her task. Another document that may be worth consulting is the WMO Guidelines on Instruments to Measure Solar Ultraviolet Radiation [4], which describes characteristics and categories of UV instruments. The SSS may include references to other ASCII files with data that are relevant to the SSS, e.g., instrument laser scans, instrument's filter characteristics, directional sensitivity scans, etc. These files will be stored in the same directory as the SSS file.

The QC document includes forms for the description of the site and the estimation of measurement uncertainty. Submitting these forms is not a requirement for data to be accepted by the WUDC, but the forms can be very useful adjuncts to the SSS, even if only partially completed. As stated in the preface of

the QC document, *"In all cases it is better to have a little information than none at all. The WUDC will post all relevant QC information with the corresponding data so it is easily available to users."*

Typically, the SSS file would be named Agency.SSS. For example, the SSS for the agency MSC is named MSC.SSS. The agency name will typically be an acronym such as MSC, JMA or U_Manchester. An example of an SSS for MSC is given in Appendix B. Further examples of the SSS may be found on the WOUDC web site at the following link: *http://www.woudc.org/data/UsingArchive/statements_e.html.* One can also view an individual agency folder, such as the MSC, at the following link: *ftp://ftp.tor.ec.gc.ca/Metadata/agencies/MSC/*.

4.2 DATA IDENTIFICATION

5.2.1 STEP 2: THE AGENCY PROFILE

The Agency Profile provides the necessary auxiliary and meta data required by the WOUDC to establish details about the measurement program, instruments, contact person(s) and address(es). This information establishes the necessary link between the data originator and the WOUDC. The profile will provide the following information:

- Agency name, mailing and electronic address, phone/fax numbers, contact person(s).
- A list of all stations from which the Agency plans to submit data.
- The instruments in use by each Agency. Each instrument must be specified by type, model and serial numbers and the main characteristics such as pass-band and spectral range must be provided.
- Details about the observation program and schedule such as type and frequency of observations, along with sources of ancillary information.

On-line forms for the Agency profile and registration of instruments are available on the WOUDC web site at: *http://www.woudc.org/data/UsingArchive/agency_form_e.html*.

Each agency will be represented by it's known acronym or be assigned an acronym by the WUDC. The country of origin must be expressed as the 3-character ISO-3166 standard [5]. This information must be included with each data file submission.

For examples of agency acronyms and ISO country codes, refer to the WOUDC web site links:

http://www.woudc.org/data/Metadata/agency_e.html
http://www.woudc.org/data/iso3166_e.html

4.2.2 STEP 3A: SPECIAL AGREEMENTS

The WUDC expects to receive quality controlled, Level 1 processed data. However, data originators are encouraged to submit Level 0, raw data. Currently, the data submission options are as follows:

Level 0 (Raw) Data - Submission

1. Level 0 data submitted in it's "native" format.
2. File format requirement - none.

Distribution options for the release of Level 0 data are presented in Table 4.2.2.

Table 4.2.2 Distribution Options for release of Level 0 data by the WOUDC.

Option	Description
1	Archive only. No distribution.
2	Distribute the data as submitted with data originator's permission.
	a. Restricted distribution for up to 24 months
	b. Unrestricted, public distribution

These options are discussed in more detail in the security section. The WUDC will archive and document the existence of these data.

Level 1 (Processed) Data - Submission

     A basic requirement of the WUDC is to be a "processed data" archive and distribution centre. Therefore, the usual method for sending data to the WUDC will involve the submission of Level 1 (processed) data. The remaining sections of this chapter will examine the required format(s) and the minimum expected content for Level 1 data.

Level 2 (Spectrally Gridded) Data - Submission

     Should a data originator decide to submit Level 2 (spectrally gridded) data then the data originator is encouraged to contact the WUDC to negotiate the terms of a submission and distribution agreement.

4.3   STEP 3B: SUBMISSION FILE FORMAT

This section will discuss in detail the required data submission file format by first giving the rationale for the simple file format directives, followed by a description of the general format rules and finally the required data content.

> *The WUDC is aware that large volumes of UV "legacy" data exist in formats which are unique to an instrument or perhaps to an Agency. **The file format used for the data submission to the WUDC may be negotiated between the originator and the WUDC. It should be clearly noted, that a certain minimum data requirement must be met for data to be accepted.***

4.3.1   RATIONALE OF THE WOUDC FILE FORMAT

The WOUDC Data Inventory, relational database uses the "metadata" (or data about the data) at the beginning of each file as an index for locating files based on search criteria such as location, time, data type etc. A direct benefit of this approach is that the same metadata rules apply for complimentary data sets from the WODC such as total ozone and vertical profile data which in turn makes data extraction a relatively simple procedure.

The format for data submission to the WOUDC addresses the following design requirements:

1. Ease of use by the submitting agency.
2. Efficient processing, archiving and quality control of data by the WOUDC
3. Providing manageable and consistent output products in a simple, readable text (ASCII) format.

4.3.2 THE EXTENDED COMMA SEPARATED FILE FORMAT FOR DATA SUBMISSION

Comma Separate Value (CSV) files can be directly imported into database, spreadsheet and analysis applications. As they are written in ASCII, they are portable to differing computer platforms. The WOUDC has extended these standard CSV syntax rules to support comments and multiple data content (tables) within individual files and this format has been given the name extended CSV or extCSV for short. An extCSV file consists of two sections: a metadata header and data tables. The metadata header has rigid format and content rules, but the data tables are designed to be more flexible. Level 1 data submissions are expected in the extCSV format, however, Level 2 and 3 data may also be submitted using this format.

EXTENDED CSV (extCSV) SYNTAX RULES

1. Fields are separated by COMMA characters, and the field width is variable.
2. Records end with some combination of CARRIAGE-RETURN and/or LINE-FEED.
3. If a data field contains either COMMA or DOUBLE-QUOTE characters, the entire field must be enclosed within DOUBLE-QUOTE characters when written. *Note:* These added characters are removed when read.
4. DOUBLE-QUOTE characters within a field are doubled (two DOUBLE-QUOTE characters) when written. *Note:* These added characters are removed when read.
5. NULL values are represented as *empty* fields.
6. Blank lines are ignored.
7. Lines beginning with an ASTERISK are comments.
8. Lines beginning with a POUND (#) symbol are table names.
9. The first record following the #TABLENAME record specifies column (field) names. Each data column (field) requires a name separated by a comma.

Example
```
* The following data are simulated...
#TABLENAME
X,Y,Comment
12,35.6,Clear sky.
12.5,,Thunderstorm (can't measure Y).
13,55.5,"It's raining, it's pouring!"
13.5,70,"Better start ""The Ark""."
```

WOUDC CONTENT RULES
1. Table names are written in UPPERCASE. Field names are case-sensitive.
2. Trailing NULLS are assumed for unspecified data attributes.
3. Each file must contain only one instance of the following static metadata tables;
    ```
    #CONTENT
    #DATA_GENERATION
    #INSTRUMENT
    #PLATFORM
    ```
5. Each file must contain at least one instance of the following dynamic metadata tables;
    ```
    #LOCATION
    #TIMESTAMP
    ```
6. *Static* metadata (#CONTENT, #DATA_GENERATION, #INSTRUMENT and #PLATORM) apply to all data within the file.
7. *Dynamic* metadata (#LOCATION and #TIMESTAMP) apply until they are re-stated. Class content that use the fields of these tables can modify the attributes (**i.e., time can be updated by individual samples**).

REQUIRED METADATA IN *extCSV*.

Styles:	**Bold**:	Required attributes.
	Italic:	Optional attributes.
	`Courier`:	Sample data.

Example	Description
#CONTENT **Class**,**Category**,**Level**,**Form** `WOUDC,Spectral,1.0,1`	*Static metadata - one per file* • WOUDC definitions required • Form refers to the current version of the data form and content of each Category
#DATA_GENERATION **Date**,**Agency**,**Version**,*ScientificAuthority* `1997-05-21,AES,1.5,Jim Kerr`	*Static metadata - one per file* • Version is specific to an Agency, referencing algorithm and instrument revisions. Refer to the *Special Note* below for details • Agency is coded by the Agency acronym. • The ScientificAuthority is the person(s) who either authored the SSS or is qualified to author the SSS and who has undertaken to adhere to the standards described in the SSS and thus be responsible for the data quality.
#INSTRUMENT **Name**,**Model**,**Number** `Brewer,MKII,15`	*Static metadata - one per file* • WOUDC definitions required (Refer to Appendix G)
#PLATFORM **Type**,**ID**,**Name**,**Country**,*GAW_ID* `STN,065,Toronto,CAN,71638`	*Static metadata - one per file* • Three letter Country code per ISO-3166 [5] • Station number (ID) assigned by WOUDC • STN-stationary, FLI-flight, SAT-satellite
#LOCATION **Latitude**,**Longitude**,**Height** `43.78,-79.47,198`	*Dynamic metadata - at least one per file* • Decimal-Degrees, (North/East are positive) • Height is in meters above sea-level

#TIMESTAMP
UTCOffset,**Date**, *Time*
`-05:00:00,1996-17-29,12:50:00`
	Dynamic metadata - at least one per file
	• Date and Time per ISO-8601 [6]
	• UTCOffset is the amount of time SUBTRACTED from the local time in order to obtain UTC time (where east is positive). Local time is defined by the user and is equivalent to *UTC =Time-UTCOffset*. From the sample data given,
	`UTC=12:50:00-(-05:00:00)=17:50:00`
	`Time given in UTC has an offset of +00:00:00.`

Special Note for the table" #DATA_GENERATION" and Category, "Version", the number conventions are as follows.
These version numbers have the form ***major.minor (i.e. 3.2)*** where;

- Major values are incremented with changes to the processing algorithm.
- Minor values are incremented when the characterisation or calibration values are changed
- Minor values are reset to zero with changes in the processing algorithm.

4.3.3 THE WUDC METADATA CONTENT: CLASS AND CATEGORY: FORM 1

WUDC data is defined as the WOUDC Data CLASS - "WOUDC". Within the CLASS WOUDC, there are several categories of UV data: Spectral, Multi-band, Broad-band and Pyranometer.

Tables 4.3.1 through 4.3.4 contain definitions for table and field names, for WUDC data content. Additional table and field types can be defined as required. These table definitions constitute the data file format - Form 1. Changes to the data file submission format are viewed as infrequent. However, should a change be necessary to either file format (structure) or data content, then a Form 1+*i* would be defined.

The use of the following table names is required for submission to the WUDC database. The tables provide the logical structure for separating (and organising) the different physical data types.

Required Table and Field Names for the WOUDC CLASS & Categories

Table 4.3.1 Required Table names based upon data content (i.e. Category) definitions.

CLASS	Category	Table Names
WOUDC	Spectral	GLOBAL, DIRECT, DIFFUSE, ACTINOMETRIC
WOUDC	Multi-band	GLOBAL, DIRECT, DIFFUSE, ACTINOMETRIC, SIMULTANEOUS
WOUDC	Broad-band	GLOBAL, DIRECT, DIFFUSE, ACTINOMETRIC, SIMULTANEOUS
WOUDC	Pyranometer	GLOBAL, DIRECT, DIFFUSE, SIMULTANEOUS

Content Rules for WOUDC Data Tables

A minimum set of physical parameters are defined for each table. The number of defined parameters for each table can vary, that is to say, many more fields can be added, but a minimum number is required. Table 4.3.2 gives the minimum set of required field names for the listed tables within the CLASS WOUDC. Table 4.3.3 gives a required set of table and field names for ancillary and auxiliary data. Table 4.3.4 lists a set of defined field names representing physical parameters relevant to the WUDC. **The spelling of table and field names must be exactly as shown in Tables 4.3.2 through 4.3.4**.

Ancillary and Auxiliary data are best submitted as separate files. For example, pyranometer data have been defined as a separate category (refer to Table 4.3.1) and therefore, require storage in separate files.

TABLE 4.3.2 Required Category, Table and Field names for the data CLASS -*WOUDC*. **The following category, table and field names must be specified as shown. (Form 1).** *Note:* Column 3 (or field name) *Time* for the Spectral and Multi-band tables is denoted in italics to indicate that field position is **reserved,** but **not required**.

Category	Table Name(s)	Field names (in order)
Spectral	GLOBAL	Wavelength,S-Irradiance,*Time*
Spectral	DIRECT	Wavelength,S-Irradiance,*Time*
Spectral	DIFFUSE	Wavelength,S-Irradiance,*Time*
Spectral	ACTINOMETRIC	Wavelength,S-Irradiance,*Time*
Multi-band	GLOBAL	Wavelength,S-Irradiance,*Time*
Multi-band	DIRECT	Wavelength,S-Irradiance,*Time*
Multi-band	DIFFUSE	Wavelength,S-Irradiance,*Time*
Multi-band	ACTINOMETRIC	Wavelength,S-Irradiance,*Time*
Multi-band	SIMULTANEOUS	Wavelength,GLS-Irradiance,*DFS-Irradiance,DRS-Irradiance,Time*
Broad-band	GLOBAL	Time,Irradiance
Broad-band	DIRECT	Time,Irradiance
Broad-band	DIFFUSE	Time,Irradiance
Broad-band	ACTINOMETRIC	Time,Irradiance
Broad-band	SIMULTANEOUS	Time,GL-Irradiance,*DF-Irradiance,DR-Irradiance*
Pyranometer	GLOBAL	Time,Irradiance
Pyranometer	DIRECT	Time,Irradiance
Pyranometer	DIFFUSE	Time,Irradiance
Pyranometer	SIMULTANEOUS	Time,GL-Irradiance,*DF-Irradiance,DR-Irradiance*

Note: The table SIMULTANEOUS refers to a single detector whereby the measurements are made in a near-simultaneous mode, but reported on the same time scale. For example, measurements are recorded in 15 second intervals, first for global then for diffuse and an integrated value is reported every 15 minutes. This integrated value is considered "simultaneous" for inclusion in this table. Simultaneous measurements made using two or more independent detectors, requires a data file for each detector. GL/GLS refers to Global/Global spectral, DF/DFS refers to Diffuse/Diffuse spectral and DR refers to Direct/Direct spectral.

General Table Names for Ancillary and Auxiliary Data

Table 4.3.3 General Table Names for Ancillary and Auxiliary Data. **Note**: *the tables are shown with sample fields which may be included in any of the above Categories.*

Table name	Suggested field names
CALIBRATION	Specified by the data originator (Scientific Authority)
METEOROLOGY	Temperature,Pressure,Relative Humidity
METEOROLOGY_SUMMARY	Time,Temperature,Pressure,Relative Humidity
SURFACE_CONDITIONS	Albedo,
IMAGE	Filename,Comments (Example: sky97001.gif)

Table 4.3.4 lists all the required and optional field names. Those field names specified in **bold** are required, while those in *italics* are optional. (Abbreviations of the units are given in brackets.)

Table 4.3.4. Accepted Field Names for WUDC data submissions [7], including units.

Field Name	Physical Units	Comments
Albedo	Dimensionless	
Cloud Amount		special codes
Date	yyyy-mm-dd	ISO-8601 standard[6]
DF-Irradiance	Wm^{-2}	Diffuse Irradiance
DFS-Irradiance	$Wm^{-2}nm^{-1}$	Diffuse Spectral Irradiance
DR-Irradiance	Wm^{-2}	Direct Irradiance
DRS-Irradiance	$Wm^{-2}nm^{-1}$	Direct Spectral Irradiance
GL-Irradiance	Wm^{-2}	Global Irradiance
GLS-Irradiance	$Wm^{-2}nm^{-1}$	Global Spectral Irradiance
Height	Metres (m)	Above sea level
Image filename and -path	No units	*.bmp, *.gif, *.jpeg etc.
Irradiance	Wm^{-2}	
Irradiation	$KJ\ m^{-2}\ nm^{-1}$	Daily integrated irradiance
Latitude	Decimal degrees	North is positive
Longitude	Decimal degrees	East is positive
Pressure	Hecto-Pascal (hPa)	
RelativeHumidity	Dimensionless (%)	
S-Irradiance	$Wm^{-2}nm^{-1}$	Spectral Irradiance
S-Irradiation	KJm^{-2}	Spectral Irradiation
Time	hh:mm:ss	ISO-8601 standard [6]
Temperature	Degrees Celsius (°C)	
Wavelength	Nanometres (nm)	
WindSpeed	ms^{-1}	
WindDirection	Decimal degrees	

4.3.4 EXAMPLES

Examples are given below of the use of the WUDC metadata to describe UV data in the extCSV file format.

The examples of spectral UV data are provided by the NSF (USA) the University of Manchester (GBR), and the MSC (CAN) as shown in Figures 4.3.4a-c.

Figure 4.3.3a NSF Data (USA) in extCSV

```
#CONTENT
Class,Category,Level,Form
WOUDC,Spectral,1.0,1

#DATA_GENERATION
Date,Agency,Version,ScientificAuthority
1995-08-12,NSF,2.3,R. Booth

#INSTRUMENT
Name,Model,Number
Biospherical,SUV-100,5

#PLATFORM
Type,ID,Name,Country,GAW_ID
STN,239,San Diego,USA

#LOCATION
Latitude,Longitude,Height
32.82,-117.13,124

#TIMESTAMP
UTCoffset Date,Time
+00:00:00,1995-08-12

*High resolution spectral scan between 280-315 nm with a total of 176 points
*Increment is 0.2 nm
*Zenith angle is 58.38 and the azimuth angle is 268.25

#GLOBAL
Wavelength,S-Irradiance,Time,TSI sensor
280.27,2.6E-6,0:01:24,1.9049
280.48,0.4E-5
280.68,-0.2E-6
280.88,1.2E-6
281.08,-1.5E-6
281.28,0.4E-5,,1.9049
…
314.93,7.58828E-2
315.13,7.62414E-2
315.33,7.38935E-2, 0:03:49,1.87805
```

Figure 4.3.3b University of Manchester Data (GBR) in extCSV. *Note in this example, that this file represents a single a day and individual spectral steps are time stamped and concatenated into one large day file of many spectra..*

U_Manchester sample data in extCSV

```
#CONTENT
Category,Category,Level,Form
WOUDC,Spectral,1.0,1

#DATA_GENERATION
Date,Agency,Version,ScientificAuthority
1994-01-28,UMIST,1.0,Ann Webb

#INSTRUMENT
Name,Model,Number
Optronic,742,89304061

*Although the serial number for this instrument is unavailable for this example,
*it is required within the data file.  In the #PLATFORM table, the station
*number for Reading is assigned by the WOUDC.  TBA means To Be Assigned.

#PLATFORM
Type,ID,Name,Country,GAW_ID
STN,353,Reading,GBR

#LOCATION
Latitude,Longitude,Height
51.45,-0.93,66

#TIMESTAMP
UTCoffset,Date,Time
+00:00:00,1994-01-27

#GLOBAL
Wavelength,S-Irradiance,Time,Undefined
280.00,2.06699E-07,10:00:00,0.29
281.00,1.02418E-08,10:00:02,0.31
282.00,0.00000E+00,10:00:04,0.32
…
399.00,,10:01:48,,
400.00,,10:01:50,,
280.00,2.06699E-07,11:00:00,0.29
281.00,1.02418E-08,11:00:02,0.31
282.00,0.00000E+00,11:00:04,0.32
…

#METEOROLGY_SUMMARY
Time,Temperature,Pressure,RelativeHumidity
10:00:00,19.2
11:00:00,21.0,
12:00:00,23.0,
…
```

Figure 4.3.3c. Meteorological Service of Canada (CAN) data in extCSV. In this example, Spectral data are given with supporting Ancillary data (pyranometer) for both Global and Diffuse radiation. Note that these data are contained within three individual files since each data sets is supplied by a separate instrument.

```
*Data file 1: spectral

#CONTENT
Name,Category,Level,Form
WOUDC,Spectral,1.0,1

#DATA_GENERATION
Date,Agency,Version,ScientificAuthority
1997-07-02,AES,1.0,Bruce McArthur

#INSTRUMENT
Name,Model,Number
Brewer,MKII,71

#PLATFORM
Type,ID,Name,Country,GAW_ID
STN,338,Regina,CAN,72863

#LOCATION
Latitude,Longitude,Height
50.21,-104.71,592

#TIMESTAMP
UTCOffset,Date,Time
-06:59:08,1997-06-08,06:45:00
*The data are reported in solar (apparent) time.
#GLOBAL
Wavelength,Irradiance,Time
290.0,0.00,06:45:00
290.5,0.00,06:45:03
291.0,0.00,06:45:06
…
325.0,6.395E-02,06:48;00

#METEOROLOGY
Temperature,Pressure,RelativeHumidity
15,972

#INSTRUMENT_CONDITIONS
Temperature
16.2

* The sequence of a #TIMESTAMP, #GLOBAL spectral scan, site and instrument
*conditions (#METEOROLOGY, INSTRUMENT_CONDITIONS) continues throughout the file

#TIMESTAMP
UTCOffset,Date,Time
-06:59:08,1997-06-08,21:45:00

#GLOBAL
Wavelength,Irradiance,Time
290.0,0.00,20:05:00
290.5,0.00,20:05:03
291.0,0.00,20:05:06
…
325.0,4.669E-01,20:08;00
```

```
#METEOROLOGY
Temperature,Pressure,RelativeHumidity
18,976

#INSTRUMENT_CONDITIONS
Temperature
19.4
```

----- Start of New data file -----

```
* Data file 2 Pyranometer Data (Global)

#CONTENT
Name,Category,Level,Form
WOUDC,Pyranometer,1.0,1

#DATA_GENERATION
Date,Agency,Version,ScientificAuthority
1997-07-02,AES,1.0,Bruce McArthur

#INSTRUMENT
Name,Model,Number
Kipp and Zonen,CM21,920065

#PLATFORM
Type,ID,Name,Country,GAW_ID
STN,338,Regina,CAN,72863

#LOCATION
Latitude,Longitude,Height
50.21,-104.71,592

#TIMESTAMP
UTCOffset,Date,Time
-06:59:08,1997-06-08

* In this example, the pyranometer data are concatenated into one table
*(#GLOBAL), with a single #TIMESTAMP table. The #METEOROLOGY and
*#INSTRUMENT_CONDITIONS_SUMMARY tables are a summary.

#GLOBAL
Time,Irradiance
06:45:00,201
07:00:00,212
07:15:00,243
…
22:00:00,244

#METEOROLOGY_SUMMARY
Time,Temperature,Pressure,RelativeHumidity
07:00:00,15,972
08:00:00,16,973
…

#INSTRUMENT_CONDITIONS_SUMMARY
Time,Temperature
06:45:00,15.8
07:00:00,16.0
…
```

----- Start of New data file -----


```
* Data file 3 Pyranometer Data (Diffuse)

#CONTENT
Name,Category,Level,Form
WOUDC,Pyranometer,1.0,1

#DATA_GENERATION
Date,Agency,Version,ScientificAuthority
1996-07-02,AES,1.0,Bruce McArthur

#INSTRUMENT
Name,Model,Number
Kipp and Zonen,CM21,920066

#PLATFORM
Type,ID,Name,Country,GAW_ID
STN,338,Regina,CAN,72863

#LOCATION
Latitude,Longitude,Height
50.21,-104.71,592

#TIMESTAMP
UTCOffset,Date,Time
-06:59:08,1997-06-08

* In this example, the pyranometer data are concatenated into one table
*(#DIFFUSE), with a single #TIMESTAMP table. The #METEOROLOGY and
*#INSTRUMENT_CONDITIONS_SUMMARY tables are a summary.

#DIFFUSE
Time,Irradiance
06:45:00,99
07:00:00,101
07:15:00,111
…
22:00:00,112

#METEOROLOGY_SUMMARY
Time,Temperature,Pressure,RelativeHumidity
07:00:00,15,972
08:00:00,16,973
…

#INSTRUMENT_CONDITIONS_SUMMARY
Time,Temperature
06:45:00,15.8
07:00:00,16.0
…
```

4.3.5 REPRESENTATION OF TIME

Typically, spectral data (where each record is Wavelength,Irradiance,Time) are represented in a single file for each day, but how time is represented within the file may be decided by the data originator. A single day may have one #TIMESTAMP table for every scan, especially if each spectral step does not have an associated time. Or the file may have one #TIMESTAMP table for the entire file, in which case some form of a time stamp (i.e. a start time, midpoint time stamp, end time or a time stamp for each spectral step) is to be indicated in the data table (#GLOBAL, #DIFFUSE, #DIRECT etc.). If a file represents a month of data, then typically, there would be a #TIMESTAMP table for each day that data are reported as well as the time stamp indicated within the data table itself as previously mentioned. If the amount of monthly data is such that these data can be accommodated in a single file, this is recommended. Otherwise, daily data are best submitted in a single file.

4.4 STEP 4: ASSIGNMENT OF PERSONAL DATA SUBMISSION ACCOUNT

This is the final step before data submission can proceed. Data originators are assigned a unique account to access the WOUDC ftp site through either conventional ftp or the Web. It is recommended that the time between data submissions should be between one week and several months. The optimum frequency is either bi-weekly or monthly. This allows adequate time to process the data.

File naming is at the discretion of the data originator. The personal ftp accounts allow write permissions. Data originators have the flexibility to re-submit and overwrite existing data for a period of one week. *The onus is on the data originator to maintain a record of data submissions for use in the WOUDC Data Inventory database*. Every Sunday at 00:00 UTC, all contents of the account are extracted, archived and then removed.

5   IDENTIFICATION OF DATA QUALITY

Data quality control and assurance is primarily the responsibility of the data originator. The purpose of the SSS is to ensure and illustrate that the appropriate measures have been undertaken to provide high quality data for the inclusion into the WOUDC database. Once a file has been submitted to the WOUDC, data quality is monitored at various stages during data processing. Database quality control consists of checking for file format and metadata content, inspection of data value ranges and data redundancy checks.

Initially, a file will be tested for adherence to the file format specifications, but later, higher level statistical summaries may be generated, which identify anomalous values, calibration problems etc. Although the later examples involve more detailed knowledge of the data, the archive already includes enough data to make some critical judgments and general assessments about the data. As the volume and statistical knowledge increase, the scope of these assessments will be broadened.

This section will focus on the initial stages of the quality control path, how are the files checked, what happens to the original data and what sort of data flagging occurs.

5.1   GENERAL PRINCIPLES OF THE WUDC IDENTIFICATION OF DATA QUALITY

There are four basic elements in the quality control of WUDC data which are:

1. Acceptance of the Data Passport or SSS.
2. Examination by the WUDC. Data are checked for file format, presence of required meta data and valid value ranges.
3. Examination by External groups
4. Additional QA/QC by the WMO and AES Advisory Committees.

The first is the joint responsibility of the originator and the WUDC while the second is an examination of the data solely by the WUDC. The third and fourth elements are not strictly defined and are expected to evolve. Data that have passed through the first two elements are published on the WOUDC ftp server.

5.2   THE DATA PASSPORT OR SSS

The single most important step in the quality control process is the acceptance of the Data Passport or SSS. Guidelines for quality control with regards to the SSS have been established and published by the WMO/GAW Scientific Advisory Group on Ultraviolet radiation (SAG_UV) [3]. In addition, examples of SSS documents can be examined on the WOUDC ftp server and Web site. The subjects to be covered are, therefore, well established. However, the writer of the SSS need not follow a rigid prescription. The contents of the SSS are discussed in Section 4.1.

5.3   EXAMINATION BY THE WOUDC

Once a file arrives at the WOUDC, the data are first checked for adherence to the correct format and content rules established in Section 5.3. Only those files that meet the metadata requirements are to be indexed in the WUDC database. It is not necessary for all files to conform to the extCSV file format, provided that the required metadata can be extracted from the file.

Examination of data values includes the identification of obvious errors such as misprints or data transmission errors. Data values are range checked and any anomalous values are further examined. When the data have been processed through the initial stages, output files ready for posting on the ftp site may include an additional data index (flag) which is aimed at providing the data client with information about any unusual or anomalous values of the data itself or derived properties of the spectrum.

5.3.1   ADDITIONAL DATA QUALITY IDENTIFICATION PROCEDURES

An additional data index (flag) has been established for WUDC output data products (Appendix C). The aim is to provide information about unusual or anomalous values or a more detailed description of other useful characteristics. Each flag is a separate digit and can have values in the range 0-3 inclusive. Code 0 indicates that the value of the relevant property is in the usual range of values for that property while the codes 1, 2 and 3 indicate increasingly unusual values. The interpretation is as follows: Code 0 means no problems or potential problems were found, Code 1 means that there may be a minor problem with the data. Normally, 3-5% of all data have Code 1. Code 2 shows either unusual observing conditions or some problem with the instrument. The percentage depends on the criteria and is different for the various flags. Code 3 shows a major disturbance in the data (property) that the "user" should definitely know about. While Codes 1, 2 and 3 raise the possibility of instrument malfunctions, it is important to note that spectra with known significant instrument problems have already been rejected. However, instrument problems cannot be considered to be eliminated from these data.

5.4   EXAMINATION BY OTHERS

The WOUDC has benefited considerably from research done by users of the data and expects to continue to do so. The user may find aspects of the data and deficiencies that are entirely unexpected by the originator. Notification of the WOUDC can produce better data processing and comments; references and/or papers by users can be included in the WOUDC.

5.5   ROLE OF WMO AND MSC (FORMERLY AES) ADVISORY COMMITTEES

Advisory Committees can help the WUDC in a number of ways. The MSC (formerly AES) committee has special expertise in the management of related data centres. The WMO/GAW SAG_UV has a wide base of experience in many aspects of UV radiation and has the task of coordinating UV measurement activities for the WMO. The SAG_UV has a subcommittee on UV data archiving which is intended to interface with the WUDC. As well as providing advice on techniques and requirements, the people on these committees can assist WUDC by being model or test clients, that is, originators and users of data, and by promoting the use of the centre. Finally, the committees can ratify or not ratify critical scientific decisions made by WUDC staff or more generally to help WUDC make such decisions. This might be particularly important with regard to decisions to accept or reject submitted data.

6 ACCESS AND DATA RETRIEVAL

Data access and retrieval are core functions of the WUDC. The primary search elements will be those fields defined in the metadata tables and content rules such as: the data CONTENT and Category, the instrument in use and it's location, the Source Agency, Location and Time stamp. The format for data outputs will generally appear the same. The basic output format will be the extCSV which will provide data clients an easy means to search a broader set of data by using the CLASS and Category fields. It is likely that enhancements to existing data sets will occur as well as the inclusion of value added data products such as statistical summaries. Therefore, some variations in the representation of data output should be expected.

The WOUDC, on a point of principle , will provide access to all data in their native, as submitted, format, subject only to possible restriction required by the data originator. The data will normally be on-line but, in the case of large volumes, special arrangement may be needed.

Regular updates and enhancements to data outputs will be posted at the WOUDC Web site.

The WOUDC data archive is available at an Internet ftp site and is accessible by means of conventional ftp or through the Web. The Internet site provides two levels of access to the posted data.

1. A generic username/password access that presently exists, and will be maintained. This level permits access to all data output products posted by the WOUDC on the Internet.
2. The second type of access is defined as "Restricted" and will be available to those data originators and clients who wish to negotiate special access agreements.

Both levels of data access are described in more detail later in this chapter.

6.1 ADDRESSES

Mail:	World Ozone and Ultraviolet Radiation Data Centre (WOUDC) c/o Environment Canada 4905 Toronto Street North York, Ontario, CANADA M3H 5T4
E-Mail	woudc@ec.gc.ca
FTP Access	ftp.tor.ec.gc.ca
WWW Access	www.woudc.org

6.2 GENERAL ACCESS

General access is a quasi-anonymous access to the entire suite of WOUDC data output products posted on the Internet. The generic username: **woudc** and password: **woudc*** will continue to be used. There are no restrictions imposed on the reading or viewing of data other than the use of the generic username and password. This access is essentially universal and conforms to what WOUDC staff believe to be the intent of WMO data agreements and to certain provisions in the Vienna Convention of 1985 [8], [9]. Nevertheless, users are asked to read relevant SSS files and to abide by any specified restrictions.

At present, the WOUDC Data Inventory will not allow searches to "drill down" to individual file attributes that are related to data content such as Temperature, Irradiance or Total ozone value. Instead, a

selection of files will be returned which meet the metadata query and it is the task of the data client to develop higher order relational database searches based upon individual requirements. ***These files allow relatively easy parsing out of elements to load conventional relational databases because the form is the same for all output files, only the data content varies***.

KEY ELEMENTS OF THE WOUDC INVENTORY DATABASE

- The archive is organised as an ASCII flat-file system using a hierarchical directory structure of raw, processed or gridded data files based on the metadata.
- Instrument and site information are stored at the beginning of each file as part of the metadata.
- No formal requirements for file formats, with respect to data submission, aside from the required metadata. **ExtCSV is suggested since data output files will be represented using this format.**
- Database queries are based on metadata (what files are available, from when and where, etc.).
- Files from the same CLASS, in this case WOUDC, "belong together", but have different "contents"; spectral, Multi-band, broad-band or pyranometer.

Use the following form to search the archive using these met data parameters:
http://www.woudc.org/data/MetaQuery/metaquery_e.cfm

6.3   DATA SECURITY: RESTRICTED ACCESS AND PROTOCOL

The second level of data access is defined as "restricted" and will be available to those data originators and clients who wish to negotiate special access agreements. Data originators and clients must apply to the WOUDC to view the restricted data archive.

Although believing strongly that the merits of fast universal distribution of data far outweigh any advantage in not distributing data, the WUDC respects the positions of individual data originators and agencies and has attempted to accommodate them as recommended by the WMO/GAW Scientific Advisory Group on UV Radiation in May, 1995.

WOUDC continues to offer data originators the flexibility to generate a second volume of data that can only be accessed by special users who have signed a protocol agreement, which is posted on the WOUDC ftp server, and who would be issued with personal passwords of limited duration.

The Scientific Sponsorship Statement can be used to prescribe certain uses of the data. Someone who violated that prescription could be censured by the WUDC as well as by the originator. In such an instance his or her password could be cancelled in addition to the censure.

The staff at the WOUDC have created special "secure" directories and user access through a specified username/password will be made available to interested individuals or groups. The general procedure to access these new directories by means of the Internet file transfer protocol (ftp) will remain the same, just the username and password will be different. Files and documentation may then be freely exchanged within this second level and be protected from "public" viewing.

The submission of data either to the "secure" or public directories is acceptable, the latter being preferred. Originators should note that users of public data are advised by the WOUDC to abide at all times by whatever stipulations the originator has written in the sponsorship file. Refer to the readme.wudc file for

further details of the sponsorship file format. Data originators may choose the level of security and access by indicating the option and signing the OPTIONS.TXT form.

To get special directory access, the data client will have to complete the information and sign that the security and protocol undertaking in PROTOCOL.TXT will be accepted. After mailing the signed PROTOCOL.TXT form to the WOUDC the data client (user) can contact ***woudc@ec.gc.ca*** to obtain a password. The form is available on the WOUDC web site at:
http://www.woudc.org/data/UsingArchive/restrict_data_form_e.html

> *Although believing strongly that the merits of fast universal distribution of data far outweigh any advantage in not distributing data, the WODC respects the positions of individual data originators and agencies and has attempted to accommodate them as recommended by the WMO Scientific Advisory Group on UV Radiation in May, 1995.*

6.4   DATA ARCHIVE STRUCTURE

The present form of the WOUDC ftp site, directory structure is outlined on the WOUDC web site at the following link: ***http://www.woudc.org/data/UsingArchive_e.html#retrieve.***   Updates or information regarding any changes to WOUDC procedures will be posted on the WOUDC Web site.

In some instances (example, large volume of data) these data files may not be kept on line, however, these data will be made available if requested. Also, if the WOUDC or the data originator(s) reformats these data files into the standard, extCSV format, then the original data files (in non-standard format) may be removed from the on-line archive, but would always be made available on request.

7   ACKNOWLEDGMENTS

The authors would like to acknowledge the contributions of the following individuals: C.T. McElroy, J.B. Kerr, L.J.B. McArthur and D.V. Barton of Environment Canada, G. Blair from Vision Works, Canada, Ann Webb of the University of Manchester Institute of Science and Technology, England and Arve Kylling (formerly of Norut It) NILU, Norway.

8   REFERENCES

[1]   NSF Polar Programs UV Spectroradiometer Network 1994-1995 Operations Report. C.R. Booth et al., Biospherical Instruments, Inc., 5340 Riley Street., San Diego, CA.., USA.

[2]   Network for the Detection of Stratospheric Change, A Status and Implementation Report, Prepared by M.J. Kurylo and S. Solomon, NASA Upper Atmosphere Research Program and NOAA Climate and Global Change Program, January, 1990.

[3]   WMO, Ann Webb ed.: "Guidelines for Site Quality Control of UV Radiation Measurements", *Global Atmospheric Watch*, WMO/GAW Report 126, 1998.

[4]   WMO, Gunther Seckmeyer ed.:, "Guidelines for Instruments to Measure Solar Ultraviolet Radiation" *Global Atmospheric Watch*, to be published, WMO/GAW Report 1998.

[5]   ISO 3166 Country and Currency Codes, International Standards Organization (ISO), Switzerland, (1995). *http://www.iso.ch/infoe/otherpub.html and ftp://ftp.ripe.net/iso3166-countrycodes.*

[6]   ISO 8601 Numeric presentations of Date and Time, International Standards Organization (ISO), Switzerland, (1997). *http://www.iso.ch/infoe/otherpub.html and http://www.iso.ch/cate/d15903.html for information on how to obtain an official version of this publication. http://www.mcs.vuw.ac.nz/technical/SGML/doc/iso8601/ISO8601.html or http://www.ft.uni-erlanger.de/~mskuhn/iso-time.html for abbreviated versions of this document.*

[7]   R. Dogniaux, et al., "Solar Meteorology: Units and Symbols": Int. J. Solar Energy, 1984, 2, p249-255.

[8]   United Nations Environment Programme, Vienna Convention of the Protection of the Ozone Layer, Final Act., United Nations Environment Programme, (1985).

[9]   World Meteorological Organization, Twelfth World Meteorological Congress, Abridged Final report with Resolutions, WMO-No.827, Resolution 40, Secretariat of the WMO, Geneva, Switzerland, 30-May-21 June, 1995.

APPENDIX A   BACKGROUND INFORMATION

A.1   HISTORY

A.1.1   ORIGIN

The World Ozone Data Centre (WODC) in Toronto was established as a WMO data centre in 1961. It is operated by Environment Canada as part of the WMO Global Atmospheric Watch (GAW) program. In 1992, a WMO sponsored Experts meeting on UV data took place in Toronto and in early 1993 the World Ultraviolet Data Centre (WUDC) was formed and began collecting spectral ultraviolet data by the end of the year. The WODC and WUDC components now constitute the one data centre which is called the World Ozone and Ultraviolet Radiation Data Centre (WOUDC). Although matters relevant to the WOUDC will be discussed, the primary focus of this document is WUDC related topics.

Although the WOUDC is operated by Environment Canada, it works under the auspices of the Global Atmosphere Watch (GAW) programme of the World Meteorological Organization (WMO). The WOUDC gets guidance from the WMO Science Advisory Groups (SAGs) for both UV (WUDC) and Ozone (WODC).

A.1.2   PURPOSE

The purpose of the WOUDC was identified in the 1992 agreement between Environment Canada and the WMO as to facilitate research on ozone and ultraviolet radiation and the effects of the latter. As with other WMO Data Centres the basic activities are the collection, archiving and distribution of geophysical data and information. The early examination of measurements by researchers in addition to those responsible for the measurements is a valuable form of quality control that can lead to improvements in measurement and analysis techniques. This is a particularly valid process for atmospheric UV radiation because the measurement technique and theory are not yet fully developed in many aspects. The WUDC, therefore, encourages originators to make their data so available. Research is also facilitated by providing uniform sets of ultraviolet radiation data which can be easily used by the scientific community. This is done by maintaining an up-to-date, well-qualified archive which involves applying simple quality control procedures, organising the data logically and providing value added output products such as data gridded on wavelengths and bandwidths, or time series analysis. The WUDC also offers centralized secure archiving of raw data that may be of value in the long-term.

A.1.3   DEVELOPMENT

At present, there remains little uniformity in ultraviolet radiation data. UV instruments not only have different characteristics, but often make measurements which differ in both sampling and resolution, and in both time and wavelength. Individual instrument types and data originators represent data in unique output formats. These formats, in turn, become the easiest form in which to submit data. Although desirable, it is not easy nor perhaps practical to establish a fixed spectral range and resolution for data submission. Conversely, users like uniformity in data, and it is reasonable that a data centre provide clients with output products which are made more uniform, for example by being gridded to a standard set of wavelengths, and presented in the same format. It may be stressed that the WUDC will not in any way process submitted data without the specific consent of the originators.

A.2 CURRENT STATUS

The guide is currently on its sixth version issued in March 2006.

A.2.1 SUBMISSION OF DATA (see section 5 for current procedure)

- Data have been accepted in the originator's format if a detailed description of the format is provided.
- The submitted data are used by the WUDC to create uniform output products.
- The data can be submitted electronically (by e-mail, FTP) or on computer media (diskettes, CD-ROMs).
- The data have been processed by the submitting agencies, i.e. the data are not Level 0 (Refer to Section 4.0 Data Levels and Versions).

A.2.2 CURRENT DATA OUTPUT PRODUCTS

- The WUDC has converted the submitted processed spectral data to a uniform, gridded wavelengths set (from 290 to 325 nm with 0.5 nm increments).
- The output files also contain calculated integrated (spectral and daily) UV characteristics and data quality flags (see section 6).
- Separate summary files list daily statistics and monthly statistics of the spectral irradiance at a few wavelengths and of two weighted integrals of the UV irradiance (CIE and ACGIH-NIOSH.)
- Data are available through the Internet at the WUDC FTP site and on the first edition, WOUDC CD-ROM (Volume 1).

Data are now available in the prototype extended comma separated values format (refer to Section 5.0 Data Submission).

APPENDIX B AN EXAMPLE OF A SCIENTIFIC SPONSORSHIP STATEMENT (SSS)

The following example of a Scientific Sponsorship Statement (SS) is for the Meteorological Service of Canada (MSC.SSS).

```
           Scientific Sponsorship Statement - MSC-SSS_v2.TXT
       for Brewer spectral data from Canadian Stations (Updated 2006-01-02)
                          Version 2.0

Toronto  1989-         Resolute  1991-        Montreal 1993-2005
Edmonton 1992-         Churchill 1992-        Regina       1994-
Winnipeg 1992-2004     Saturna   1991-        Alert        1995-
Halifax  1992-2005     Saskatoon 1990-2002    Eureka       2001-


The Brewer angular responce error correction is implemented in Version 2
of the spectral UV irradiance data processing algorithm. The correction
largely removes the bias in Brewer UV measurements mentioned below for
Version 1 data. The correction procedure is described in ref. 5.



           Scientific Sponsorship Statement - SSSAES.TXT
       for Brewer spectral data from Canadian Stations (Issued 1999-04-15)
                          Version 1.0


General Methodology.
The data are from UV-B scans of the single monochromator Brewer instrument
(ref 1).

The calibration source comprises calibrated DHX 1000 watt lamps that are
traceable to the US National Institute of Standards and Technology (NIST).
These lamps were either calibrated by Optronics Inc., based on NIST standards,
or directly by NIST.

Daily stability checks are done on the Brewer by reading the signals from an
internal 20 watt quartz halogen lamp.  Prior to 1990, a 200 watt external lamp
was used roughly every month. These stability checks are done primarily to
detect malfunctions and for ozone measurements
rather than to establish UV responsivities

The Toronto Brewer #014, and two traveling standard Brewers #017 & #039 have
been calibrated roughly six times per year with 1000 watt lamps at Toronto.

The field stations (Halifax, Edmonton and Winnipeg etc.) have been calibrated
approximately each year by bringing two 1000 watt lamps and one traveling
Brewer instrument to the site.

Laboratory tests have shown that there is a wavelength-dependent temperature
effect on responsivity of up to +/- 2% over the operating range of temperature.
The measurements (at this time)  have not been corrected for any temperature
effect.

The wavelength on the Brewer is checked several times per day by reference to
a mercury discharge lamp.  Although the control of the wavelength setting is
in steps of about 0.007nm and the finding can be reproduced to 0.002nm,
changes in the lamp and the accuracy of the dispersion equation limit the
wavelength accuracy.

The pass-band is approximately triangular with a 0.55 nm full width at half
intensity.( See note 1).
```

The spectra are corrected for instrumental stray light which affects the measurements at wavelengths less than 305 nm where the light intensity is very small. The method of stray light correction was revised in August, 1996 from the "5 wavelength" correction which subtracts the signal measured between 290 and 292 nm from the rest of the spectra (ref. 3) to the "laser scan" correction which uses a scan of the HeCd laser (325.03 nm) to define the stray light characteristics of the single monochromator Brewer instrument. The Toronto data are corrected using a laser scan measured on the Toronto instrument (#14). As of August 1996, because laser scans from the other field instruments are not yet available, the data are derived using estimated laser scans with parameterized characteristics specific to each instrument.

The spectra are not corrected for the dependence of responsivity on incidence direction. The responsivity of Brewer instruments is greatest at normal incidence and decreases with increasing incidence angle. Calibration is done at normal incidence. Charaterizations of the Toronto instrument and a new MKIII instrument (ref. 4) indicate that measurements usually under estimate the horizontal global irradiance by 4-8% and 2-7% respectively depending on the solar zenith angle and sky conditions. The revised uncertainty limits given below include bias due to this effect.

Results.
The Toronto instrument and the traveling Brewers show a month-to-month reproducibility of
about 2% over the five years period.

The Toronto Brewer shows long-term spectral responsivity drifts ranging from -0.3% per year to +1.0% per year based on 1000 watt lamp readings.

On field visits results from the lamps are usually within 2% of the result derived from side-by-side
measurement with the traveling standard Brewer.

Uncertainty estimates (bias, +/- 2 sigma):
Wavelength   0 +/- 0.05nm
Intensity   Field stations   +6% +/- 6%
               Central station +6% +/- 4%.

Notes:
1   A scan of the Helium-Cadmium laser emission at 325 nm has been submitted to WOUDC as L3250141.DAT.

2   New 1000 watt lamps and lamp housings are now being used. This will enable more frequent calibrations at all sites.

3   Various new methods of quality control and checking calibrations are being developed.
     These include correlation with satellite measurements and with local measurements of spectrally integrated downward irradiance.

4   The variability in the Brewer readings on the 1000 watt lamps is currently treated as entirely due to the instrument. However, there is clear evidence of changes in the lamps and of change in results due to small differences in procedures. It is hoped that a more sophisticated treatment, based on a larger number of lamp calibrations, can be developed.

5   It is expected that the uncertainty of new and existing data can be reduced in future.

D.I. Wardle and J.B. Kerr 95 04 28, revised 96 04 02, 96 08 30, 99 04 15

Ref1.   J.B. Kerr, C.T. McElroy, D.I. Wardle, R.A. Olafson and W.F.J. Evans:1984 The automated Brewer spectrophotometer. Atmospheric Ozone: Proceedings of the Quadrennial Ozone Symposium held in Halkidiki, Greece, 3-7 September 1984. D. Reidel Publishing Company

Ref2.   D.I. Wardle, J.B. Kerr and C.T. McElroy: Operational ozone and spectral UV-B monitoring in Canada: 13th UOEH International Symposium on the Impact of Increased UV-B Exposure on Human Health and Ecosystems. 13-15 October 1993, Kitakyushu, Japan.

Ref3.   J.B. Kerr and C.T. McElroy: Evidence for Large Upward Trends of Ultraviolet-B Radiation linked to Ozone Depletion. Science 262, 1032-1034, 12 November 1993.

Ref4.   A.F. Bais, S. Kazadzis, D. Balis, C. Zerefos and M. Blumthaler: Correcting global solar ultraviolet spectra recorded by a Brewer spectroradiometer
for its angular response error, Applied Optics, 37, 6339-6344, 1998.

Ref5.   V.E. Fioletov, J. B. Kerr, D. I. Wardle, N. A. Krotkov, and J. R. Herman Comparison of Brewer UV irradiance measurements with TOMS satellite retrievals, Optical Engineering, 41 (12), 3051-3061, 2002.

APPENDIX C   WUDC DATA FLAGS

For the spectral UV irradiance data properties associated with the six flags are as follows:

1. The irradiance at 324nm which is not significantly affected by ozone and which is compared with the "clear-sky" value for the same solar elevation. In this context, the "clear-sky" value is defined as the 95th percentile of all measurements except those made when there was snow on the ground. (The actual value is not very sensitive to the percentile level, the 99th percentile being ~1.08 of the adopted value.).
2. An estimate of total ozone derived from the spectrum and solar elevation (Ref.).
3. The standard error in the total ozone estimate (Ref.).
4. An estimate of sulfur dioxide from the spectrum (Ref.).
5. The ratio of irradiances at individual wavelengths to those in an average UV spectrum for the given solar zenith angle and derived ozone amount.
6. The difference between the raw spectra recorded with increasing and decreasing wavelength. (Each spectrum in the processed Version 1 and 2 data is the average of an upward and a downward scan.)

Reference: Fioletov, V.E., J.B. Kerr and D.I. Wardle, The relationship between total ozone and spectral UV irradiance from Brewer spectrophotometer observations and its use for derivation of total ozone from UV measurements, *Geophys. Res. Lett.*, **24**, 2997-3000, 1997.

The formal criteria are shown in Table AD.1.

Table AD.1. The formal criteria used for error flag determination.

Digit	Code	Criterion	Comment
1	1	UV324 is either less than the 5th percentile (i.e. <13% of "clear sky" conditions) or 1.5 times higher than "clear sky" conditions	Rare but possible cases: low UV324 indicates heavy clouds. High UV324 can be related to clear sky and fresh snow.
	2	UV324 is either less than the 1st percentile (i.e. <6% of clear sky conditions) or 2 times higher than "clear sky" conditions	Low UV324 indicate very heavy clouds, rain, or snow. Very high UV324 can be observed in high latitudes at low sun conditions.
	3	UV324 is either at least 10 times less than the 1st percentile (<0.6% of clear sky conditions) or 2.5 times higher than "clear sky" conditions	It is still possible to observe these low UV324 in the case of a very heavy rain. Otherwise it may be caused by snow covering the instrument.

2	1	Total ozone derived from UV is less than 210 DU or more than 550 DU with small errors	Rare but possible total ozone amounts.
	2	Total ozone derived from UV is less than 150 DU or more than 650 DU with high errors	Very rare total ozone amounts. High errors may indicate variable cloud conditions during the measurements
	3	Total ozone derived from UV is less than 100 DU or more than 750 DU	Unusual total ozone may be an indicator of variable cloud conditions. For example, O_3 will be low if clouds covered the Sun only during the measurements of UV at 320-325 nm, i.e. in the middle of the measurement cycle. Very high ozone can also be observed with heavy clouds
3	1	The error in total ozone retrieved from the spectrum is above the 99th percentile of all such errors at the given solar elevation	Indicates variable conditions during the measurement
	2	The error in total ozone retrieval is twice that specified for Code 1	Indicates highly variable conditions during the measurement
	3	The error in total ozone retrieval is six times that specified for Code 1	Very high variations of the conditions during the measurement, the data probably cannot be used for O_3 or SO_2 estimations
4	1	SO_2 derived from UV is 5-10 DU	Traces of SO_2.
	2	SO_2 derived from UV is 10-20 DU	SO_2 component is seen in UV spectrum
	3	SO_2 derived from UV is more than 20 DU	High total amount of SO2. (It has to be accounted for when ozone - UV relationship is studied).

5	1	At 1 or 2 wavelengths the irradiance is outside the +/-6 sigma envelope	Can be a random error. The appearance of that code in a large number of spectra may indicate, a wavelength shift.
	2	As for Code 1 but at 3 or 4 wavelengths	Indicates variable conditions during the measurement or an wavelength shift
	3	As for Code 1 but at 5 or more wavelengths	Indicates highly variable conditions during the measurement. Total ozone and SO_2 probably cannot be estimated correctly

6	1	At 1 or 2 wavelengths the ratio between the irradiance measurements on the upward and downward scans exceeds the factor 2.	Another criterion that indicates variable conditions during the measurement. Applies to Brewer type double scans.
	2	As for Code 1 but at 3 or 4 wavelengths	
	3	As for Code 1 but at 5 or more wavelengths	Indicates highly variable conditions during the measurement. Total ozone and SO_2 probably cannot be estimated correctly. The data may be useful only for calculation of integrated UV characteristics.

APPENDIX D LIST OF ACRONYMS

Acronym	Name
ACGIH	American Conference of Governmental Industrial Hygienists
ASCII	American Standard Code for Information Interchange
BDMS	Brewer Data Management System
CIE	Commission Internationale de L'Eclairage (Int'l Com. of Illumination)
CSV	Comma Separated Value
extCSV	extended Comma Separated Value
FTP	File Transfer Protocol
FWHM	Full Width at Half Maximum
ISO	International Organization for Standardization
JMA	Japanese Meteorological Agency
MSC	Meteorological Service of Canada (Environment Canada)
NASA	National Aeronautical and Space Administration (USA)
NDSC	Network for Detection of Stratospheric Change
NIOSH	National Institute of Occupational Safety and Health
NIST	National Institute of Standards and Technology (USA)
NSF	National Science Foundation (USA)
PAR	Photosynthetically Active Radiation
SAG_UV	WMO/GAW Scientific Advisory Group on Ultraviolet radiation (formerly the Scientific Steering Committee on Ultraviolet radiation or SSC_UV)
SSS	Scientific Sponsorship Statement
TOMS	Total Ozone Mapping Spectrometer
UTC	Coordinated Universal Time
WMO/GAW	World Meteorological Organization - Global Atmosphere Watch
WOUDC	World Ozone and Ultraviolet Radiation Data Centre
WODC	World Ozone Data Centre
WUDC	World Ultraviolet Radiation Data Centre