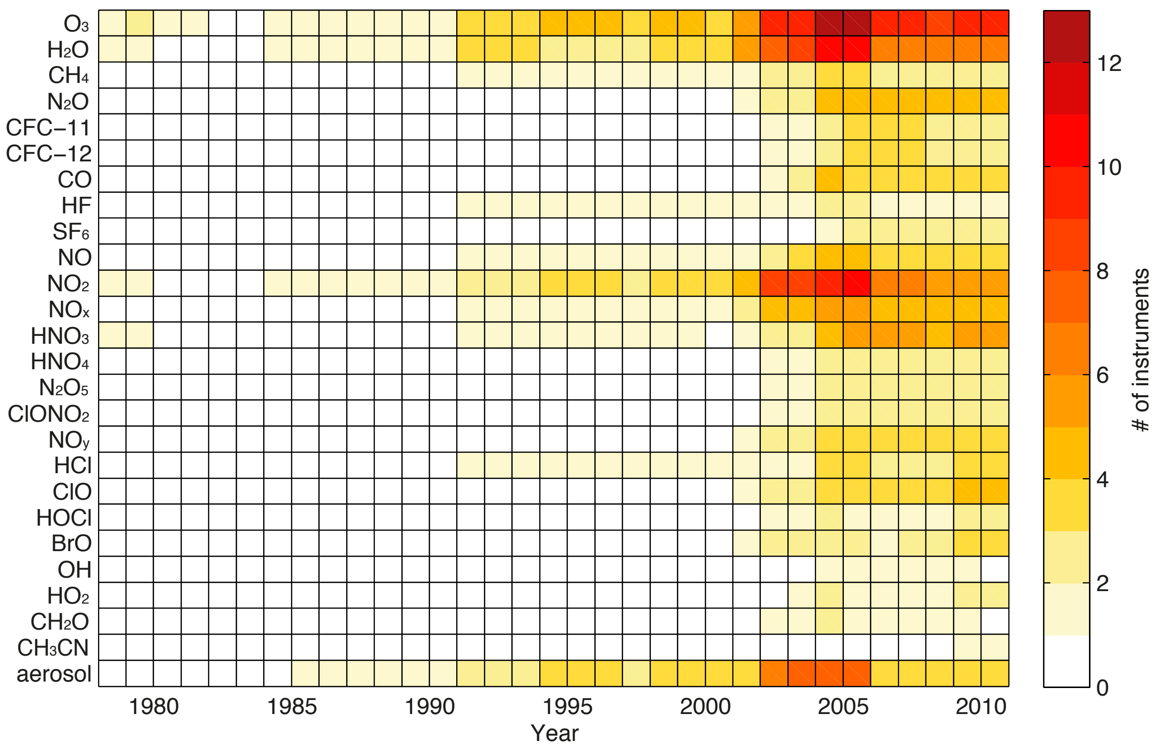
Chapter 1: Introduction

The past 30 years have been a 'golden age' for satellite measurements and have provided a wealth of knowledge regarding chemical trace gas abundances in the stratosphere. There is a danger that in the future the stratosphere will not be as well measured and it is therefore important to capture existing knowledge of current and recent instruments, retrievals and datasets before this knowledge is lost.

Satellite instruments from CSA, ESA, JAXA, NASA, SNSB, and other national space agencies provide a large number of trace gas datasets, which differ in terms of measurement method, geographical and seasonal coverage, spatial and temporal sampling and resolution, time period, and retrieval technique. These datasets of chemical trace gases are widely used for empirical studies of stratospheric climate, trends, and variability, and for the evaluation of the representation of transport and chemistry in numerical models. However, the validity of such studies strongly depends on the quality and representativeness of the datasets used, and it is often difficult for a user to determine which is the most reliable or useful dataset for a particular application. Hence, it is essential that the characteristics of the datasets be known prior to their use and prior to the interpretation of results. For example, comparing numerical model output to different chemical datasets can lead to conflicting results, which limits the value of model-measurement intercomparison studies.

Issues arising when using observational datasets for model evaluations have been identified in the SPARC CCMVal report [*SPARC*, 2010], which undertook a comprehensive assessment of model performance in the stratosphere. The report's recommendations directly motivated the work for the SPARC Data Initiative. The recommendations included: (1) '*Long-term vertically resolved datasets of constituent observations in the stratosphere are required to assess model behaviour and test model predictions. This includes ozone, but also other species that can be used to diagnose transport and chemistry. The current set of GCOS* [note at the time of writing] *Essential Climate Variables is not sufficient for process-oriented evaluation of CCMs.*' (2) '*More global vertically resolved observations are required, particularly in the UTLS. As CCMs evolve towards including tropospheric chemistry, lack of observations in this region will become a major limitation on model evaluation.*' (3) '*A systematic comparison of existing observations is required in order to underpin future model evaluation efforts, by providing more accurate assessments of measurement uncertainty.*'

***Table 1.1:** Number of instruments within the SPARC Data Initiative measuring a particular chemical trace gas species or aerosol in a given year.*



There is also a strong need to characterise instrument differences as a prelude to data merging activities. These activities aim to merge various data sources into homogeneous climate data records suitable for trend studies, evaluation diagnostics, or climate forcings in global climate models. Merging of data for such purposes is only meaningful if differences between datasets are systematic and consistent.

Finally, the atmospheric trace gas datasets are not always available in a standard form, or with appropriate documentation. To enable the best possible use of the satellite datasets it is important to provide easy access to the datasets in a common format as well as to the information on the different instrument techniques and retrieval procedures.

The SPARC Data Initiative helps to address these issues by having performed the first comprehensive multi-instrument comparison of stratospheric chemical trace gas climatologies. It thereby provides a user guide to the different datasets, along with easy access to the data in a common format, and recommends future studies that would enhance the quality and usefulness of the existing data. In order to attain these goals, the SPARC Data Initiative assessed, *in a first step*, the current availability of vertically resolved, chemical trace gas and aerosol datasets from a suite of multi-national space-based instruments. *In a second step*, chemical trace gas and aerosol monthly zonal mean time series were compiled in a common and simple-to-use NetCDF data format. *In a third step*, these trace gas time series underwent detailed comparisons, which identified strengths and shortcomings of all datasets and differences between them. Where possible, an expert judgment on the source of those differences is provided.

Assessment of trace gas availability: Middle atmospheric trace gas observations are available from an international suite of satellite limb sounders, with the first measurements starting in 1979. Some of the instruments launched after 2000 are presently still taking regular measurements, despite being already past their expected lifetimes. All instruments have been measuring different sets of chemical species depending on the measurement technique applied. Earlier instruments were mostly based on the solar occultation technique, measured in the UV/VIS range and focused on ozone, water vapour and some nitrogen species. Instruments launched after 2000 were more often scattering and emission sounders, the latter extending measurements into microwave and sub-mm wavelengths, and covered a wider range of measured species. For each trace gas the number of satellite datasets within the SPARC Data Initiative is given as a function of time in **Table 1.1**.

Compilation of zonal monthly mean time series: The observational datasets have been compiled into a common data format, which is easy to handle by data users. To this end, zonal monthly mean time series of each trace gas species (in volume mixing ratio, VMR) and aerosol (as extinction ratio) have been calculated for each instrument on the SPARC Data Initiative climatology grid, using 5 degree latitude bins (with mid-points at 87.5°S, 82.5°S, 77.5°S, ..., 87.5°N) and 28 pressure levels (300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1, 0.7, 0.5, 0.3, 0.2, 0.15, and 0.1 hPa) corresponding to the CCMVal pressure levels. The data therefore encompass the atmospheric region from the upper troposphere up to the lower mesosphere. Along with the monthly zonal mean value, the standard deviation and the number of averaged data values are given for each month, latitude bin and pressure level. Furthermore, the mean, minimum, and maximum local solar time (LST), the average day of the month, and the average latitude of the data within each bin for one selected pressure level are provided.

Evaluation diagnostics: In contrast to traditional data evaluation techniques based on coincident profiles, the SPARC Data Initiative compares climatologies in order to reduce geophysical variability and to obtain an assessment of our knowledge of the mean atmospheric state. Different standard evaluation diagnostics are used, such as single- or multi-year annual or monthly mean climatologies, vertical and meridional profiles, and seasonal cycles. In addition, time-latitude or time-altitude evolutions are assessed in order to test the physical consistency of the datasets. These include the tropical tape recorder in water vapour, polar dehydration, polar ozone loss, or the Quasi-Biennial Oscillation (QBO). The general approach taken is to compare the instruments to the multi-instrument mean, as explained in **Box 1**.

The notations for different atmospheric and geographical regions that are being used throughout this report are listed in **Tables 1.2** and **1.3**, respectively. **Table 1.4** defines the naming convention for the level of agreement between the instruments used in this report.

Table 1.2: Definitions and abbreviations of different atmospheric regions referred to in the report*. Note that the notations UTLS and USLM refer accordingly to the total extent of the sub-regions (i.e., 300-30 hPa and 5-0.1 hPa).*

Region	Abbreviation	Lower boundary	Upper boundary
Upper Troposphere	UT	300 hPa	Tropopause
Lower Stratosphere	LS	Tropopause	30 hPa
Middle Stratosphere	MS	30 hPa	5 hPa
Upper Stratosphere	US	5 hPa	1 hPa
Lower Mesosphere	LM	1	0.1 hPa

Table 1.3: ***Definitions of different geographical regions referred to in the report.***

Region	Latitude range
Tropics	30°S-30°N
Subtropics	20°S-40°S and 20°N-40°N
Mid-latitudes	30°S-60°S and 30°N-60°N
High/polar latitudes	60°S-90°S and 60°N-90°N

Table 1.4: ***Definition of levels of agreement between a given climatology and the multi-instrument mean.***

%-differences	Level of agreement
Up to ± 2.5%	Excellent agreement
Up to ± 5%	Very good agreement
Up to ± 10%	Good agreement
Up to ± 20%	Reasonable agreement
Up to ± 50%	Considerable disagreement
Up to ± 100%	Large disagreement

An approximate measure of random uncertainty in each climatological mean is the standard error of the mean (SEM); calculated from n measurements and a standard deviation, SD, as $SEM=SD/\sqrt{n}$. Due to its ease of computation and frequent use in past studies and despite its shortcomings (see *Chapter 3* for details), the SEM will be used as an approximate measure of uncertainty in each climatological mean, graphically illustrated by error bars of ±SEM, which can be loosely interpreted as a 68% confidence interval of the mean.

The analysis of O_3, aerosol and H_2O climatologies in the report is intended to support other ongoing SPARC activities focused on characterising long-term changes such as WAVAS II (for H_2O), SI^2N (for O_3), and SSiRC (for aerosol), and also to provide valuable information on data quality to "data merging" activities currently being carried out by NASA and ESA.

The zonal mean climatologies of the different chemical trace gas and aerosol products that were compiled during the SPARC Data Initiative can be downloaded from the SPARC Data Centre website (**http://www.sparc-climate.org/data-centre/**). In general, the results of this report depend on the specific level-2 data versions on which the climatologies are based, and future data versions might give different results. The goal is to provide updated climatologies whenever new data versions become available. The improvements achieved in moving to the next data version will be explained in meta-data or references provided. Interested users of the SPARC Data Initiative climatologies are asked to follow the data policy instructions posted in the same directory.

The report is structured as follows. *Chapter 2* comprises detailed information on the instruments participating in the SPARC Data Initiative, including measurement techniques and retrieval descriptions. *Chapter 3* gives an overview of the methodology used by the SPARC Data Initiative to create the climatologies and the approach used to evaluate them. *Chapter 4* features all comparisons of the chemical trace gases and aerosol, while *Chapter 5* summarises some general interpretation and higher-level conclusions of the results.

Box 1: *Multi-Instrument Mean Reference*

The approach of the SPARC Data Initiative is to use the multi-instrument mean (**MIM**) as a common point of reference. The choice of the MIM is by no means based on the assumption that it is the best estimate of the atmospheric trace gas field, but is motivated by the need for a reference that does not favor a certain instrument. It should be stated that the MIM *is not* a data product and *is not* provided as part of the SPARC Data Initiative datasets.

The MIM is calculated by taking the mean of all available instrument climatologies within a given time period of interest. The time periods can vary for the different trace gases and are chosen to ensure maximum spatial and temporal data coverage for each instrument and to limit the impact of sampling bias. In general, all available instrument datasets are included in the MIM regardless of their quality and without any weighting applied to them. Only if measurements from a particular instrument are deemed completely unrealistic, or if the same instrument is providing two versions of a specific trace gas data product, are they not included in the MIM.

The SPARC Data Initiative evaluations are based on relative differences between the trace gas mixing ratios of an instrument (X_i) and the MIM (X_{MIM}) given by:

$$diff[\%] = 100 * (X_i - X_{MIM}) / X_{MIM}$$

One has to keep in mind when interpreting relative differences with respect to the MIM that the composition of instruments from which the MIM was calculated may have changed between time periods. Hence, changes in derived differences are not to be interpreted as changes in the performance (or drifts) of an individual instrument. Also, if there is unphysical behaviour in one instrument, the MIM and thus the differences of the other instruments with respect to the MIM will most certainly reflect this unphysical behaviour as well. Finally, if one instrument does not have global coverage for every month some sampling biases may be introduced into the MIM. A detailed assessment of the uncertainty introduced due to inhomogeneous temporal or spatial sampling in the SPARC Data Initiative climatologies is provided in *Chapter 3*.