TITLE:

ESA-SPARC Initiative (SPIN)

User Requirements Document (URD)

for the Essential Climate Variables (ECVs)

Temperature and Water Vapour

Reference: SPIN_URD_v1.2

Date of issue: 22/10/2012

Distributed to:

ESA
SPIN consortium

This work is supported by the European Space Agency

DOCUMENT PROPERTIES

Title	User Requirements Document (URD) for the Essential Climate Variables (ECVs) Temperature and Water Vapour
Reference	SPIN_URD_v1.2
Issue	01
Revision	02
Status	Final
Date of issue	22/10/2012
Document type	Deliverable

	FUNCTION	NAME	DATE
LEAD AUTHOR	WP Manager	Greg Bodeker	22/10/2012
CONTRIBUTING AUTHORS	Deputy Science Leader	Ted Shepherd	22/10/2012
REVIEWED BY	WP Manager	Michaela Hegglin	
APPROVED BY	Science Leader	Michel van Roozendael	22/10/2012
ISSUED BY	Project Manager	Nathalie Kalb	22/10/2012

DOCUMENT CHANGE RECORD

Issue	Revision	Date	Modified items	Observations
01	01	31/05/2012	New document – initial draft for submission	Main purpose is to solicit input for discussion
01	02	22/10/2012	Significant revisions and additions	Final version for submission to ESA

Table of Contents

List of tables

List of figures

Executive summary

- Detecting changes in temperature and water vapour in the atmosphere is a cornerstone of climate change research. Long-term trends in these essential climate variables, and in particular the vertical structure of the trends, provide a fingerprint of anthropogenic origin for those changes. However, no single measurement system, either for temperature of for water vapour, provides the long-term homogeneous climate data records required for robustly detecting such trends. Combining measurements from several sources is therefore an imperative. It is also essential that the additional uncertainty introduced when splicing together measurement series from different instruments is faithfully captured in the homogenized record.
- Different uses of temperature and water vapour data products have different requirements. Providing a single set of measurement requirements to cover all intended uses is likely to result in requirements that are too stringent for some uses.
- Measurement requirements change with location and season. Ideally measurement requirements should be tailored to each location to determine e.g. the measurement random error required to ensure that a trend of some expected magnitude can be detected over some sampling period. Alternatively, for a measurement system with fixed characteristics, the vertical, seasonal and geographic coverage over which those measurements are useable for some intended purpose must be determined.
- Changes in observation schedule also affect trend estimates. Reducing the number of observations in a month, or changing the timing of a single observation each day, has a greater potential to produce errors in trends than reducing the number of days per month on which observations are made.

Applicable and reference documents

Austin, J.; Wilson, J.; Li, F. and Vömel, H., Evolution of Water Vapor Concentrations and Stratospheric Age of Air in Coupled Chemistry-Climate Model Simulations, *J. Atmos. Sci.*, 64, 905-921, 2007.

Bodeker, G.E.; Boyd, I.S. and Matthews, W.A., Trends and variability in vertical ozone and temperature profiles measured by ozonesondes at Lauder, New Zealand: 1986-1996, *J. Geophys. Res.*, 103, 28661-28681, 1998.

Dee, D. P.; et al., The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.

Foelsche, U.; Scherllin-Pirscher, B.; Ladstädter, F.; Steiner, A.K. and Kirchengast, G., Refractivity and temperature climate records from multiple radio occultation satellites consistent within 0.05%, *Atmos. Meas. Tech.*, 4, 2007-2018, doi:10.5194/amt-4-2007-2011, 2011.

Forster, P.M.; Bodeker, G.E.; Schofield, R.; Solomon, S. and Thompson, D.W.J., Effects of ozone cooling in the tropical lower stratosphere and upper troposphere, *Geophys. Res. Lett.*, 34, L23813, doi:23810.21029/22007GL031994, 2007.

Fueglistaler, S. and Haynes, P.H., Control of interannual and longer-term variability of stratospheric water vapor, *J. Geophys. Res.*, 110, D24108, doi:24110.21029/22005JD006019, 2005.

Gettelman, A.; Randel, W.J.; Wu, F. and Massie, S.T., Transport of water vapor in the tropical tropopause layer, *Geophys. Res. Lett.*, 29, 10.1029/2001GL013818, 2002.

Hegglin, M.I.; Boone, C.D.; Manney, G.L.; Shepherd, T.G.; Walker, K.A.; Bernath, P.F.; Daffer, W.A.; Hoor, P. and Schiller, C., Validation of ACE-FTS satellite data in the upper troposphere/lower stratosphere (UTLS) using non-coincident measurements. *Atmos. Chem. Phys*., 8, 1483-1499, 2008.

Immler, F.J.; Dykema, J.; Gardiner, T.; Whiteman, D.N.; Thorne, P.W. and Vömel, H., Reference Quality Upper-Air Measurements: guidance for developing GRUAN data products, *Atmos. Meas. Tech.*, 3, 1217–1231, 2010.

Knudsen, B.M., Accuracy of arctic stratospheric temperature analyses and the implications for the prediction of polar stratospheric clouds, *Geophys. Res. Lett.*, 23, 3747-3750, 1996.

Po-Chedley, S.; Fu, Q., A bias in the mid-tropospheric channel warm target factor on the NOAA-9 microwave sounding unit, *J. Atmos. Oceanic Tech.*, doi:10.1175/JTECH-D-11-00147.1, 2012.

Randel, W.J.; Shine, K.P.; Austin, J.; Barnett, J.; Claud, C.; Gillett, N.P.; Keckhut, P.; Langematz, U.; Lin, R.; Long, C.; Mears, C.; Miller, A.; Nash, J.; Seidel, D.J.; Thompson, D.W.J.; Wu, F. and Yoden, S., An update of observed stratospheric temperature trends, *J. Geophys. Res.*, 114, D02107, doi:02110.01029/02008JD010421, 2009.

Rodgers, C.D. and Connor, B.J., Intercomparison of remote sounding instruments, *J. Geophys. Res.*, 108, 4116, doi:4110.1029/2002JD002299, 2003.

Rosenlof, K.H.; Oltmans, S.J.; Kley, D.; Russell III, J.M.; Chiou, E.-W.; Chu, W.P.; Johnson, D.G.; Kelly, K.K.; Michelsen, H.A.; Nedoluha, G.E.; Remsberg, E.E.; Toon, G.C. and McCormick, M.P., Stratospheric water vapour increases over the past half-century, *Geophys. Res. Lett.*, 28, 1195-1198, 2001.

Scherllin-Pirscher, B.; Steiner, A.K.; Kirchengast, G.; Kuo, Y.-H. and Foelsche, U., Empirical analysis and modeling of errors of atmospheric profiles from GPS radio occultation, *Atmos. Meas. Tech.*, 4, 1875-2011, doi:10.5194/amt-4-1875-2011, 2011.

Seidel, D.J. and Free, M., Measurement Requirements for Climate Monitoring of Upper-Air Temperature Derived from Reanalysis Data, *J. Climate*, 19, 854-871, 2006.

Soden, B.J.; Turner, D.D.; Lesht, B.M. and Miloshevich, L.M., An analysis of satellite, radiosonde, and lidar observations of upper tropospheric water vapor from the Atmospheric Radiation Measurement Program, *J. Geophys. Res.*, 109, D04105, doi:10.1029/2003JD003828, 2004.

Solomon, S.; Rosenlof, K.H.; Portmann, R.W.; Daniel, J.S.; Davis, S.M.; Sanford, T.J. and Plattner, G.-K., Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, *Science*, 327, 1219-1223, 2010.

SPARC, Assessment of Upper Tropospheric and Stratospheric Water Vapour, SPARC Report No. 2, WCRP No. 113, WMO/TD No. 1043, 2000.

Thomason, L.W.; Moore, J.R.; Pitts, M.C.; Zawodny, J.M. and Chiou, E.W., An evaluation of the SAGE III version 4 aerosol extinction coefficient and water vapor data products. *Atmos. Chem. Phys.*, 10, 2159-2173, 2010.

Vaisala: Vaisala Radiosonde RS92-SGP, available at: www.vaisala.com, last access: 13 April 2010, 2006.

Wang, L.; Zou, C.-Z. and Qian, H., Construction of stratospheric temperature data records from Stratospheric Sounding Units. *J. Clim.*, 25, 2931-2946, 2012.

Zhou, S.; Geller, M.A. and Zhang, M., Cooling trend of the tropical cold point tropopause temperatures and its implications, *J. Geophys. Res.*, 106, 1511-1522, 2001.

Acronyms and abbreviations

AMSU - Advanced Microwave Sounding Unit
BIPM - International Bureau of Weights and Measures
CDR - Climate Data Record
CEOS - Committee on Earth Observation Satellites
ECMWF - European Centre for Medium-Range Weather Forecasts
ECV - Essential Climate Variable
ESA - European Space Agency
GCOS - Global Climate Observing System
GHG - Greenhouse Gas
GOMOS - Global Ozone Monitoring by Occultation of Stars
GPS - Global Positioning System
GRUAN - GCOS Reference Upper Air Network
JMA - Japan Meteorological Agency
MSU - Microwave Sounding Unit
NASA - National Aeronautics and Space Administration
NCEP - National Centers for Environmental Prediction
NCEPCFSR - NCEP Climate Forecast System Reanalysis
NWP - Numerical Weather Prediction
OLR - Outgoing Longwave Radiation
RO - Radio Occultation
SAGE - Stratospheric Aerosol and Gas Experiment
SPARC - Stratospheric Processes And their Role in Climate
SPIN - ESA-SPARC Initiative
SSU - Stratospheric Sounding Unit
URD - User Requirements Document
UT/LS - Upper troposphere/lower stratosphere
WAVAS II - Second SPARC Water Vapour Assessment
WMO - World Meteorological Organisation

1. Introduction

1.1. Purpose

The purpose of this document is to review the requirements for the measurement of vertical profiles of temperature and water vapour as detailed in GCOS-107 and its more recent update (GCOS-154). This user requirements document is a deliverable within the ESA-SPARC Initiative (SPIN).

The document begins by summarizing the envisaged uses of stratospheric temperature and water vapour profile measurements since these uses will guide the requirements of the data. The document also recognizes that different uses of the data will have different requirements. In this way the document ensures that the needs of the climate community are being adequately addressed.

With these requirements in mind the document goes on to present current measurement requirements as documented by various ECV data user groups. The document closes with an expert review of the measurement requirements, with a particular focus on whether these requirements can be realistically achieved with current satellite-based instrumentation.

The user requirements presented in this document are based on peer-reviewed publications, other documents where user requirements have been formulated, and user consultation focussing on a set of key ECV data users. A close cooperation between SPIN and the SPARC temperature trends activity and water vapour activity (WAVAS II) has been established for this purpose.

1.2. Scope

This user requirements document focusses exclusively on the extent to which satellite-based measurements of stratospheric temperature and water vapour can meet the needs of the climate community.

2. Background

2.1. Stratospheric temperature

Stratospheric temperatures represent the first order connection between natural and anthropogenically driven changes in radiative forcing and changes in other climate variables at the surface. Furthermore, the vertical structure of temperature trends is important information for climate change attribution since increases in atmospheric long-lived greenhouse gas (GHG) concentrations warm the troposphere but cool the stratosphere, steepening vertical temperature gradients in extra-tropical regions. Other drivers of stratospheric temperature changes, e.g. changes in solar output, would not have the same vertical profile fingerprint. Resolving discrepancies between temperature trends derived from satellite-based measurements and from radiosondes strengthens the attribution of changes in temperatures to changes in climate forcing agents (Po-Chedley and Fu, 2012).

Temperatures measured by satellite-based instruments are needed to:

- Monitor the vertical structure of temperature trends.
- Correlate changes in other parameters, especially water vapour (see below), with changes in temperature.
- Validate temperature trends simulated by climate models.
- Provide input to global meteorological reanalyses such as NCEP, ECMWF, NASA, JMA.
- Provide input to numerical weather prediction (NWP) models. Stratospheric measurements of temperature and water vapour are two of the basic measurements used in the initialization of NWP models for operational weather forecasting.

The requirements for random error, bias and long-term stability are detailed below and are guided, in part, by the needs of the four end-user communities described in Section 3. A particular focus is the use of the measurements, which include the effects of natural, unforced climate variability, in detecting stratospheric temperature trends. This becomes a signal-to-noise ratio problem and climate models can be used to guide the measurement requirements given expectations of future trends in temperature and natural variability.

It is particularly important that trends in tropical cold point tropopause temperatures are accurately detected since these temperatures in large part control the flux of water vapour into the stratosphere (Gettelman et al., 2002; Fueglistaler and Haynes, 2005) and changes in stratospheric water vapour influence radiative forcing and temperatures both in the lower stratosphere but also in the upper troposphere (Forster et al., 2007; Solomon et al., 2010). At present, temperature trend uncertainties in the lower stratosphere remain large, particularly in the tropics. For this ECV specifying measurement requirements sufficient to robustly detect tropical cold point temperatures is essential.

2.2. Stratospheric water vapour

Water vapour is the primary natural GHG and is central to global water and energy cycles. It acts primarily as a feedback, amplifying the effects of increases in other GHGs. Water vapour is also a source of OH in the upper troposphere and stratosphere, influencing methane, ozone and halogenated GHGs. High clouds due to water vapour in the upper troposphere/lower stratosphere (UT/LS) affect both the planet's shortwave albedo and its longwave greenhouse effect, and both cloud particles and water molecules are involved in chemical reactions that govern stratospheric ozone concentrations. Fully quantifying the Earth's radiation budget depends on an accurate assessment of the radiative properties of the water vapour continuum.

Changes in water vapour in the UT/LS exert a greater radiative forcing than changes elsewhere (Solomon et al., 2010). A number of factors, many linked to changes in climate, are likely to affect the flux of water vapour into this climatically important region of the atmosphere, viz.:

- Changes in the cold-point tropopause temperature (Zhou et al., 2001).

- Changes in convection. Convective transport of ice particles into the UT/LS can provide a path which bypasses the limitation imposed by the cold-point tropopause temperature.
- Changes in the Brewer-Dobson circulation (Austin et al., 2006).

Monitoring atmospheric water vapour presents many difficulties. Because the residence time of water vapour in most of the atmosphere is short, about 10 days, its distribution is horizontally, as well as vertically, heterogeneous. Stratospheric water vapour is more spatially homogeneous but because its concentrations there are measured in parts per million it is difficult to measure accurately. The satellite data record of upper troposphere and lower stratosphere water vapour measurements to date is not sufficiently accurate to be useful for climate applications (Soden et al., 2004). However, accurate water vapour measurements in the stratosphere are critical, especially for radiative transfer modelling. Understanding the stratospheric water vapour budget is also necessary for interpreting measurements of outgoing longwave radiation (OLR).

Satellite-based solar occultation and limb-sounding instruments can measure water vapour in the upper troposphere and stratosphere but inter-satellite differences preclude the use of earlier data in long-term trend analyses (Rosenlof et al., 2001). High precision measurements of water vapour profiles will provide valuable input data to global meteorological reanalyses and data for validating global climate models.

2.3. Terminology

The following terminology, as specified by the International Bureau of Weights and Measures (BIPM) in the *Guide to the Expression of Uncertainty in Measurement*[1], is used throughout this user requirements document to describe the uncertainty components of a measurement:

True value: This is a value consistent with the definition of a given particular quantity that would be obtained by a perfect measurement. True values, by nature, cannot be determined.

Measurement accuracy: Every measurement has imperfections that cause it to differ from the true value. The measurement accuracy describes the closeness of the agreement between the result of a measurement and a true value of the measurand.

Measurement uncertainty: A parameter, associated with the result of a measurement, which characterizes the dispersion of the values that could reasonably be attributed to the measurand. Measurement uncertainties may be time dependent.

Measurement error: The result of a measurement minus a true value of the measurand.

Random error: The result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions. The random error component of any measurement is the re-

---

[1] http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

sult of stochastic variation in quantities that influence that measurement. While random errors cannot be designed out of a system, the random error on the mean of multiple measurements is reduced since, by definition, the expected value for the random error is zero. The term 'random error' is preferred over the term 'precision' since precision is often used to designate the number of bits or significant digits to which a value is specified.

Systematic error: The mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand. It results from systematic biases that do not average to zero as the number of measurements increases. However, if these systematic biases can be identified and quantified, they can be corrected for. The term 'systematic error' is preferred over the term 'accuracy' since it denotes more clearly that the deviation is systematically in one direction.

Stability: Stability refers to the consistency of random errors and systematic errors with time. Undetected changes in systematic errors induce artificial trends in measurement time series.

3. Envisaged uses of ECV data

3.1. Climate change detection and attribution

Long-term climate data records (CDRs) of stratospheric temperature and water vapour, created by combining measurements from, preferably, overlapping satellite-based instruments, are essential for detecting and attributing changes in the climate of the stratosphere, and for validating model simulations of long-term changes. For this envisaged use of satellite-based measurements of temperature and water vapour, the following measurement requirements are key:

Long-term stability: For detecting long-term changes (sometimes called trends) in stratospheric temperature and water vapour, reducing the systematic error on the measurement is less important than ensuring that the systematic errors and random errors remain consistent in time. Undetected changes in systematic errors induce artificial trends in measurement time series.

Accuracy of altitude registration: In the presence of steep vertical gradients in temperature or water vapour, long-term drifts in the altitude registration of the measurement can alias into artificial trends in the measurement series. Splicing measurements with different systematic errors to create a single CDR relies on the stability of each measurement over the overlap period.

Minimized orbital drift: In the presence of a strong diurnal cycle in the ECVs, a drift in local solar time of the measurement can alias into artificial trends in the measurement series.

3.2. Satellite calibration/validation

Measurements from one satellite-based instrument may be used to validate measurements from another satellite-based instrument, or to provide the input needed for radiative transfer calculations required to calibrate another satellite-based instrument. For this envisaged use of satellite-based measurements of temperature and water vapour, the following measurement requirements are key:

Small systematic and random errors: The measurement needs to be as close to the true value as possible and must have as small a measurement uncertainty as possible. Measurement random and systematic errors must be well characterized to allow a valid comparison between measurements from two different systems (Immler et al., 2010).

High vertical and horizontal resolution: This allows for the measurements to be smoothed to match the resolution of the measurements being validated (Rodgers and Connor, 2003).

High spatial coverage: This ensures that the likelihood of finding coincident measurements in space and time between the two systems is high.

3.3. Reanalysis and numerical weather prediction (NWP)

Satellite-based measurements of stratospheric temperatures and water vapour constitute an important data source for reanalyses and for assimilation into NWP models. The direct assimilation of radiances (now standard practice in data assimilation) has reduced many of the systematic errors associated with retrieved temperatures. For this envisaged use of satellite-based measurements of temperature and water vapour, the following measurement requirements are key:

Long-term stability: For reanalyses, long-term stability is key, although variational bias correction (Dee et al. 2011) can overcome some issues (especially rapid changes in systematic errors) provided there are data from another, stable data record (so-called 'anchoring' data) which can be used to detect changes in systematic errors. This is the case for temperature in the lower stratosphere (radiosondes, and increasingly GPS) but not for upper-stratospheric temperature or for water vapour. The assimilation of measurements at key locations that are stable over the multiple decades of the assimilation ensures that the assimilation products as a whole exhibit the same level of stability.

Small systematic and random errors: Long-term homogeneity of measurement series being assimilated into reanalyses ensures that the reanalysis products do not exhibit any discontinuities. For NWP, while minimizing both systematic and random errors is important, this is less critical than having a large number of measurements (so long as the accuracy is accurately known), which are better able to constrain the analysis. Current NWP data assimilation systems correct for systematic errors by comparing the satellite measurements with the same variables simulated by the model. Mean differences are assumed to result from the systematic error in the satellite observations (the model and its radiative transfer calculation are assumed to have no bias).

Measurements in key locations: For numerical weather prediction, measurements in key locations can significantly reduce the uncertainty in the forecast. However, the location of such key sites varies with synoptic condition and cannot always be determined *a priori*.

3.4. Atmospheric process studies

A wide range of studies of atmospheric processes require accurate and high vertical resolution measurements of stratospheric temperature and water vapour. Such data are also needed for process-oriented evaluation of climate models. For such applications the covariation of temperature and water vapour is important (e.g. to test hypotheses regarding dehydration) and so it is important to characterize any correlations of errors in the two fields. Generally speaking the dynamic range of process variability helps increase the signal compared to that typical of long-term changes, making accuracy less critical than for other applications, but high vertical resolution is required to capture the processes in detail. High temporal resolution ensures that measurements are available at the time that the process to be studied occurs.

4. Documented ECV requirements from data user groups

4.1. Global Climate Observing System (GCOS)

Measurement requirements for stratospheric temperature, as detailed in GCOS-107 and later updated in GCOS-154 are listed in Table 1.

Variable/parameter	Horizontal resolution	Vertical resolution	Temporal resolution	Systematic error[2]	Stability
Stratospheric temperature profile	100km along-track	2km	4 hours	0.5K	0.05K/decade
Temperature of deep atmospheric layers	100km	5km	Monthly averages	0.2K	0.02K/decade

Table 1: Measurement requirements for stratospheric temperature as detailed in GCOS-107 and GCOS-154.

Measurement requirements for stratospheric water, as detailed in GCOS-107 and later updated in GCOS-154 are listed in Table 2.

Variable/parameter	Horizontal resolution	Vertical resolution	Temporal resolution	Systematic error[2]	Stability
Tropospheric and lower-stratospheric profiles of water vapour.	25km in the troposphere & 100-200km in the stratosphere	2km	4 hours in the troposphere and daily in the stratosphere	5%	0.3%/decade
Upper-tropospheric humidity	25km	N/A	1 hour	5%	0.3%/decade

Table 2: Measurement requirements for stratospheric water vapour as detailed in GCOS-107 and GCOS-154.

---

[2] GCOS-154 uses the term ‘accuracy’ which they define as the requirement for closeness of agreement between product values and true values. This is equivalent to our term ‘systematic error’.

4.2. GCOS Reference Upper Air Network (GRUAN)

GRUAN differentiates between requirements consistent with state-of-the-art capability and GRUAN measurement targets. Measurement requirements for stratospheric temperature as detailed in the *GRUAN Guide to Operations* are listed in Table 3.

Target	**Vertical resolution**	**Random error**	**Systematic error**	**Stability**
State-of-the-art capability	100m or better below 30km altitude, 500m above 30km altitude.	≤0.2K	1K	0.05K/decade
GRUAN goal	100m or better below 30km altitude, 500m above 30km altitude	≤0.2K	≤0.2K	Better than 0.05K/decade

Table 3: Measurement requirements for stratospheric temperature as detailed in the *GRUAN Guide to Operations*.

GRUAN distinguishes between different potential uses of the water vapour measurements in defining the water vapour measurement requirements:

Attribute	**Trend detection**		**Satellite validation and radiation studies**		**Process studies**
	Upper troposphere	**Lower stratosphere**	**Radiance comparisons**	**Comparisons in retrieval space**	
Vertical resolution	<1 km	<1 km	N/A	< 2km	10-100 m
Systematic error	profile: 5-10%	profile: 5-10% or better	column: 3% profile: 5% in lower and mid-troposphere, 10% in upper troposphere	column: 3% profile: 10% in 2 km thick layers	profile: 10%
Random error	up to 50%[3]	<10%	many comparisons: 10-20% individual comparison: ≤5%		<10-25%
Stability	0.05/decade	0.1K/decade	N/A	N/A	N/A
Temporal resolution	<1 hour	no data	high as possible		1 minute

Table 4: Measurement requirements for stratospheric water vapour as detailed in the *GRUAN Guide to Operations*.

4.3. WMO/CEOS Rolling Review of Requirements (WMO)

An online version of the WMO Observing Requirements Database is available at http://www.wmo-sat.info/db/. Measurement requirements for stratospheric temperature extracted from this database are listed in Table 5[4].

---

[3] For measurements made 2-3 times per week and assuming that systematic errors have been randomized using appropriate procedures.

[4] The WMO/CEOS database of rolling review of requirements does not list values for stability.

Layer	Application Area	Horizontal resolution	Vertical resolution	Temporal resolution	Random error[5]
Lower stratosphere	Global modelling	Goal[6]: 50km Min: 500km	Not listed	Goal: 3h Min: 12h	Goal: 0.5K Min: 3K
	Global numerical weather prediction	Goal: 15km Min: 500km	Goal: 300m Min: 3km	Goal: 1h Min: 12h	Goal: 0.5K Min: 3K
	High resolution numerical weather prediction	Goal: 10km Min: 100km	Goal: 1km Min: 3km	Goal: 15min Min: 6h	Goal: 0.5K Min: 3K
	SPARC	Goal: 50km Min: 500km	Goal: 500m Min: 2km	Goal: 6h Min: 3d	Goal: 0.5K Min: 1K
	Synoptic meteorology	Goal: 20km Min: 200km	Goal: 100m Min: 2km	Goal: 3h Min: 12h	Goal: 0.5K Min: 3K
	Climate-AOPC	Goal: 100km Min: 500km	Goal: 2km Min: 3km	Goal: 3h Min: 6h	Goal: 0.5K Min: 2K
Upper stratosphere	Global modelling	Goal: 50km Min: 500km	Not listed	Goal: 3h Min: 12h	Goal: 1K Min: 3K
	Global numerical weather prediction	Goal: 50km Min: 500km	Goal: 300m Min: 3km	Goal: 1h Min: 24h	Goal: 0.5K Min: 5K
	SPARC	Goal: 50km Min: 500km	Goal: 500m Min: 2km	Goal: 6h Min: 3d	Goal: 0.5K Min: 1K
	Climate-AOPC	Goal: 100km Min: 500km	Goal: 2km Min: 3km	Goal: 3h Min: 6h	Goal: 0.5K Min: 3K

Table 5: Measurement requirements for stratospheric temperature extracted from the WMO Observing Requirements Database.

Measurement requirements for stratospheric water vapour have been extracted from the specific humidity tables from the WMO/CEOS Rolling Review of Requirements database and are listed in Table 6.

---

[5] The WMO/CEOS database of rolling review of requirements uses the term ‘uncertainty’ but do not define anywhere specifically what the uncertainty includes. It has been assumed that this is the random error.

[6] The WMO/CEOS database of rolling review of requirements lists a target goal as well as a minimum standard.

Layer	Application Area	Horizontal resolution	Vertical resolution	Temporal resolution	Random error
Lower stratosphere	Climate	Goal: 50km Min: 200km	Goal: 2km Min: 3km	Goal: 3h Min: 6h	Goal: 2% Min: 20%
	Atmospheric chemistry	Goal: 50km Min: 500km	Goal: 1km Min: 5km	Goal: 12h Min: 3d	Goal: 5% Min: 20%
	Global climate modelling	Goal: 50km Min: 250km	Not listed	Goal: 3h Min: 12h	Goal: 5% Min: 20%
	SPARC	Goal: 50km Min: 500km	Goal: 500m Min: 2km	Goal: 6h Min: 3d	Goal: 2% Min: 5%
Upper stratosphere	Climate	Goal: 50km Min: 200km	Goal: 2km Min: 5km	Goal: 3h Min: 6h	Goal: 2% Min: 20%
	Atmospheric chemistry	Goal: 50km Min: 500km	Goal: 1km Min: 5km	Goal: 12h Min: 3d	Goal: 5% Min: 20%
	Global modelling	Goal: 50km Min: 250km	Not listed	Goal: 3h Min: 12h	Goal: 5% Min: 20%
	SPARC	Goal: 50km Min: 500km	Goal: 500m Min: 2km	Goal: 6h Min: 3d	Goal: 2% Min: 5%

Table 6: Measurement requirements for stratospheric temperature extracted from the WMO Observing Requirements Database.

5. Currently achievable measurement attributes

The discussion in this section is restricted to global data sets, hence to operational radiosondes and satellite data sets. Although the radiosonde network is not truly global, with large data gaps especially over the oceans, it can provide large-scale averages which are representative of atmospheric domains such as the tropics.

5.1. Temperature

Operational radiosondes have long provided the backbone of the upper-air temperature network, and are critical for anchoring reanalyses (Dee et al., 2011). With a vertical resolution of tens of metres, a random error of ≤0.2 K (Vaisala, 2006; Knudsen, 1996) and systematic error ≤0.5 K below ~30 km (Immler et al., 2010), radiosonde measurements generally meet the vertical resolution and measurement error requirements for temperature profiles listed in Section 4. However, operational radiosonde measurements are typically only archived at WMO standard levels, not at the intrinsic resolution of the profile, and so the vertical resolution of the profile is generally no better than 3 km in the stratosphere, which is not adequate. The temporal sampling is generally every 12 hours, which is not adequate. The horizontal sampling is extremely inhomogeneous and definitely not compliant with the requirements. Finally, radiosondes are limited to about 30 km altitude, and generally only provide reasonable coverage up to 25 km (Randel et al., 2009). With regard to long-term stability, this can be compromised by changes in instrument or operating practice, and as a result there are various quality-controlled radiosonde data sets which can be compared to assess their consistency (Randel et al., 2009). The resulting uncertainties in potential drifts certainly exceed the GCOS requirement of 0.05 K/decade.

The space-based GPS RO network is providing a critical supplement to the radiosonde network for lower stratospheric temperature. With a measurement uncertainty of ≤0.1 K for altitudes below 20 km, and 0.2 K in dry temperature between 4 and 35 km (Scherllin-Pirscher et al., 2011), GPS RO measurements meet the error requirements for temperature profiles listed in Section 4 for the upper troposphere and lower stratosphere. However the errors increase with altitude, especially above 35 km, and so the useful data are probably limited to the lower stratosphere. With a vertical resolution of better than 1.5 km, GPS RO measurements meet the climate-related vertical resolution requirements of GCOS, but not the process or NWP-related requirements of GRUAN or WMO. The horizontal and temporal resolution of the global network of GPS RO measurements depends on the number of receivers; although it is not currently adequate, it is improving. The GPS RO method has exceptional long-term stability (Foelsche et al., 2011).

Limb viewing research satellites provide an inhomogeneous record of stratospheric temperature profiles. The solar occultation technique cannot provide global coverage and is especially restrictive for temperature, which is highly variable in time. Thus useful records are limited to those derived from thermal emission or stellar occultation techniques. For thermal emission sounders, the vertical resolution is generally no better than 3-4 km so is not adequate to meet the requirements described in Section 4, although such measurements could be used to anchor the operational nadir-sounding temperatures of deep atmospheric layers. Much higher vertical resolution is in principle possible with stellar occultation (GOMOS), although this has yet to be demonstrated. Horizontal resolution is generally around 300 km due to horizontal smearing in the limb. Horizontal and temporal sampling is generally inadequate since these measurements are usually obtained from single satellites and are thus constrained by orbital considerations (e.g. not even 1000 km longitudinal sampling once per day). Measurement uncertainty is generally around 0.5 K so, in principle, is adequate, although long-term stability is not assured since it is generally not a requirement of research satellites, and would need to be confirmed in each case.

Temperatures of deep atmospheric layers are provided by nadir-viewing operational satellites in the MSU/SSU/AMSU series. These measurements inherently have relatively coarse vertical resolution (about 5 km, after over-sampling), but reasonably high horizontal and temporal resolution (roughly 500 km global sampling every 6 hours). Apart from their vertical resolution, the main issue with the operational nadir measurements is their long-term stability, given the rapid drift in the orbits and the short (several years) lifetime of individual instruments (Wang et al., 2012). In the lower stratosphere, there are various retrievals of MSU Channel 4 (representative of the 15-20 km layer) which can be compared with homogenised radiosonde or GPS RO measurements, so there is some understanding of the long-term stability of the derived climate data records (Randel et al., 2009). In the middle and upper stratosphere, however, there are only two retrievals of SSU Channels 1 (25-35 km), 2 (35-45 km), and 3 (40-50 km), only one of which is documented, which disagree significantly with each other (Wang et al., 2012). Unfortunately, neither the radiosonde nor the GPS RO measurements can be used to calibrate the temperature measurements in this altitude range. Thus, the long-term stability of nadir-based middle and upper stratospheric temperature measurements is highly questionable.

5.2. Water vapour

Radiosonde measurements of water vapour in the upper troposphere and stratosphere are not considered reliable (SPARC, 2000) and are not assimilated in NWP or reanalysis systems (Dee et al., 2011). The only useful global observations of water vapour in this region therefore come from limb viewing research satellites. For thermal emission sounders, the vertical resolution is generally no better than 3-4 km so is not adequate to meet the requirements described in Section 4. Higher vertical resolution (1 km) is possible from solar occultation because of the higher signal-to-noise ratio which allows oversampling (Hegglin et al., 2008); this meets the climate-related vertical resolution requirements of GCOS, but not the process or NWP-related requirements of GRUAN or WMO. However, in contrast to thermal emission, solar occultation does not provide global coverage on a daily (or even monthly) basis. Horizontal resolution is generally around 300 km due to horizontal smearing in the limb. Even for thermal emission sounders, the horizontal and temporal sampling is generally inadequate since these measurements are usually obtained from single satellites and are thus constrained by orbital considerations (e.g. not even 1000 km longitudinal sampling once per day). Measurement uncertainty is generally around 10-15%, of which perhaps 5% is systematic error (SPARC 2000; see Thomason et al. 2010 for

SAGE III), so is borderline in terms of the requirements discussed in Section 4. However, long-term stability is not assured since it is generally not a requirement of research satellites, and would need to be confirmed in each case. The difficulty with respect to the latter is that there is no accepted ‘gold standard’ for water vapour measurements in the upper troposphere and stratosphere.

6. Requirements, rationale and traceability

6.1. Introduction

The temperature and water vapour profile data products developed in SPIN are monthly mean products. Few, if any, of the measurement requirements specified by existing ECV data user groups (see Section 4) list target measurement requirements for monthly means; these are primarily for instantaneous measurements. It is therefore necessary to link these instantaneous measurement requirements to requirements on monthly means. The requirements on the monthly means will, in turn, depend on the intended uses of the data.

6.2. Temperature profile data product

Seidel and Free (2006), using a reanalysis of the climate of the past half century as a model of temperature variations over the next half century, tested various data collection protocols to develop recommendations for observing system requirements for monitoring upper-air temperature. The analysis focussed on accurately estimating monthly average temperature and its standard deviation, and multi-decadal trends in monthly temperatures at specific locations, from the surface to 30 hPa. Because the analysis of Seidel and Free did not extend above 30 hPa, the analysis has been repeated here, but now based on NCEPCFSR reanalyses to extend the results to 1 hPa.

The effects of increasing the random error of temperature measurements, incomplete sampling of the diurnal cycle, incomplete sampling of the days of the month, imperfect long-term stability of the observations, and changes in observation schedule were assessed by Seidel and Free. It was found that to ensure accurate monthly climate statistics, observations with random error ≤0.5 K, made at least twice daily, at least once every two or three days is sufficient. Using these same criteria, and maintaining long-term measurement stability to within 0.25 K for periods of 20 years, errors in trend estimates can be avoided in at least 90% of cases. Maintaining stability to within 0.1 K for 50 years ensures that errors in trend estimates can be avoided in 95% of cases. This requires no more than one intervention (e.g., instrument change) over measurement period, and its effect must be to change the measurement systematic error by no more than 0.25 K. The effect of the first intervention dominates the effects of subsequent, uncorrelated interventions.

To corroborate the findings of Seidel and Free, and to extend the analysis into the upper stratosphere, a similar approach has been followed here where, for a number of selected locations around the globe, the uncertainty on monthly mean temperatures is determined as a function of sampling frequency, random error on instantaneous measurement (which can then be related to the measurement requirements outlined in Section 4), season, and pressure. This analysis is based on sampling of NCEPCFSR reanalyses, assuming that sampling at the highest possible frequency (6 hourly) produces the ‘true’ monthly mean, and then by simulating different sampling strategies, with different simulated random errors on each measurement, and doing this in a Monte Carlo framework, the standard deviation of the differences between the calculated monthly means and the true monthly means can be determined.

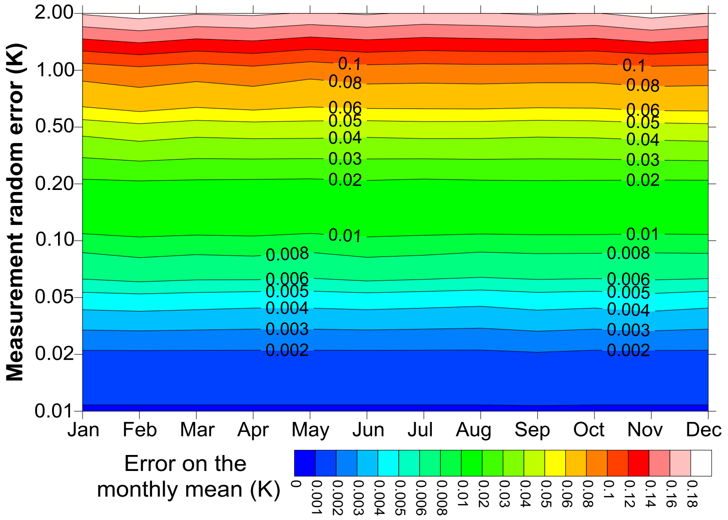

Figure 1: The resultant error on the monthly means at 50 hPa, 85°N, 135°W, for 6 hourly sampling, as a function of the random error on each instantaneous measurement.

Figure 1 shows the random error on the monthly mean as a function of season and random error on each instantaneous temperature measurement for 6 hourly sampling of the temperature through the month. Since there is no contribution to the uncertainty on the monthly mean from sampling (the 6 hourly sampling is the same as that used to derive the ‘true’ monthly mean), the error on the monthly mean is about an order of magnitude smaller than the error on each instantaneous measurement which is what is expected when averaging $\sim$120 measurements through the month i.e. $1/\sqrt{120} \approx 0.1$.

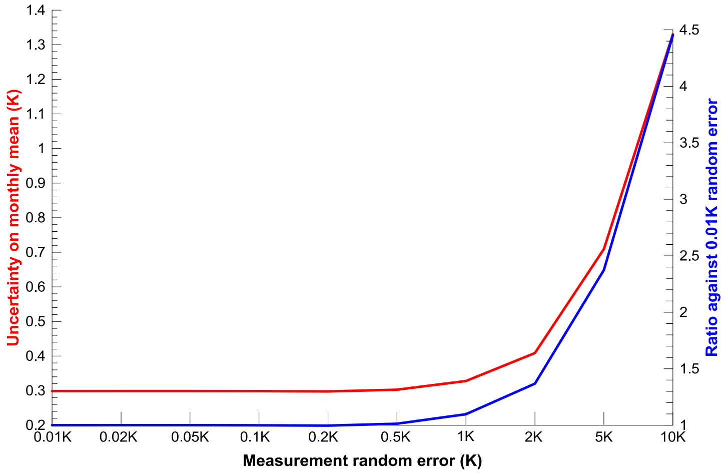

Figure 2: An example of how the uncertainty on the monthly mean varies with measurement random error (red trace). The blue trace shows the ratio of the monthly mean random errors with respect to the monthly mean random errors at 0.01K sampling.

Figure 2 shows an example of how the random error on the monthly mean typically varies with measurement random error. In this particular example it is clear that reducing the measurement random error below 0.5 K has little effect on the random error on the monthly means, corroborating the results of Seidel and Free. However, this behaviour is likely to be altitude, season, and location dependent.

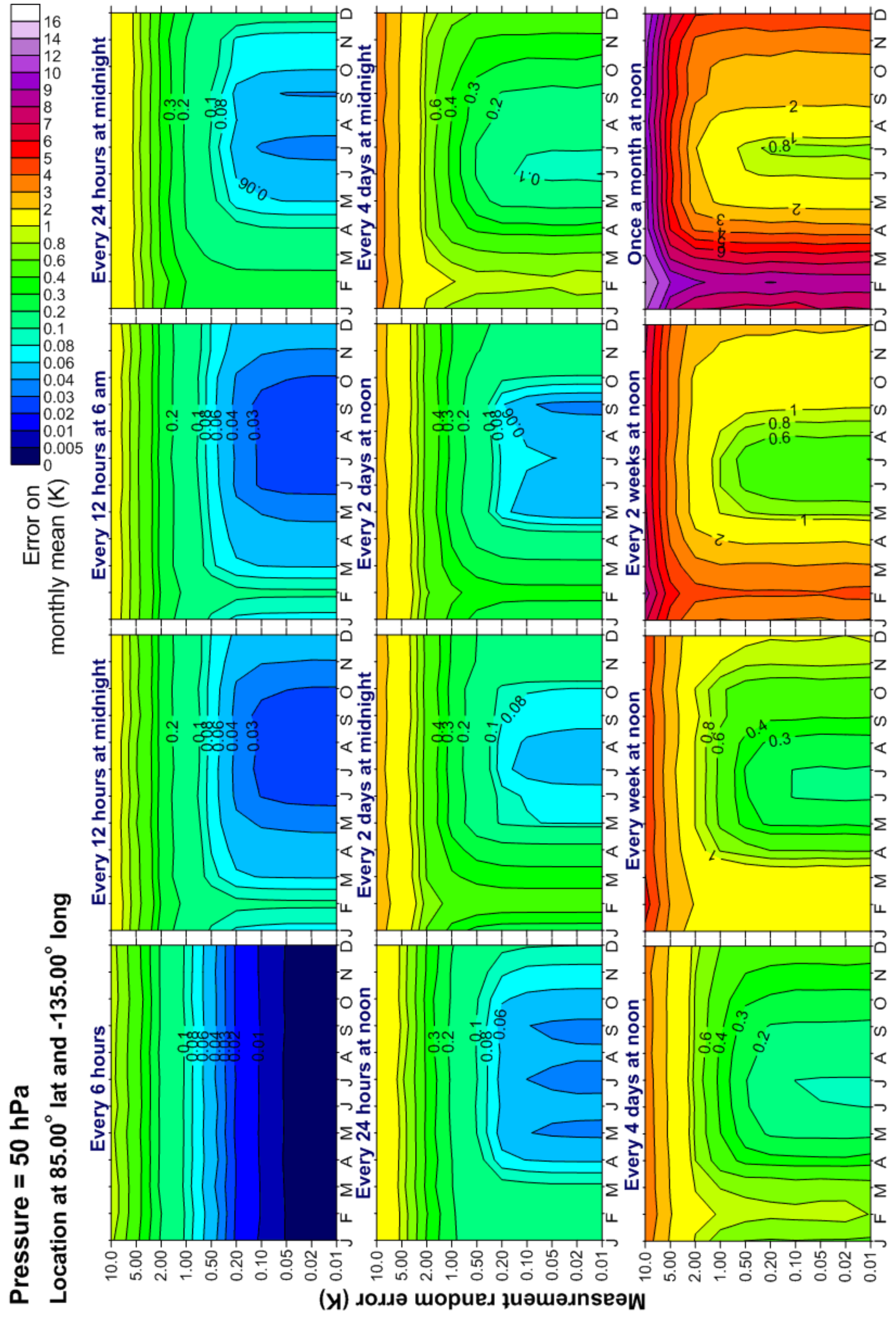

Figure 3: The uncertainty on the monthly means at 50 hPa, 85°N, 135°W, for a range of sampling strategies, as a function of the random error on each instantaneous measurement.

Figure 3 shows the random errors on the monthly means for the same location as shown in Figure 1, but now for a range of different sampling strategies. We reiterate that this is the *random error* on the monthly means, ignoring any systematic errors (offsets) since these are less important for trend analysis. The random error on the monthly mean now shows a clear seasonal cycle for 12 hourly sampling or courser since the temperatures show a higher degree of variability in the winter months. At this pressure level (50 hPa), reductions in measurement random errors below 0.2 K have little effect on the uncertainty on the monthly means since it is the uncertainty resulting from incomplete sampling that dominates. It is only for measurement random errors of greater than 0.2 K where the measurement random error begins to make an appreciable contribution to the random error on the monthly mean. This 0.2 K threshold lends support to the GRUAN target of ≤0.2 K random error on instantaneous stratospheric temperature measurements. Clearly the permissible measurement error is likely to vary with pressure and season.

When sampling every 12 hours at midnight, Figure 4 shows the permissible random errors on individual measurements required to avoid increasing the uncertainty on the monthly means by more than 10% above what would be achieved when sampling at 0.01K random error. While it is clear that in the upper stratosphere sampling at 0.5 K random error is sufficient to avoid affecting the uncertainty on the monthly means, this reduces to 0.25 K at ~20 hPa and to 0.15 K in the free troposphere.

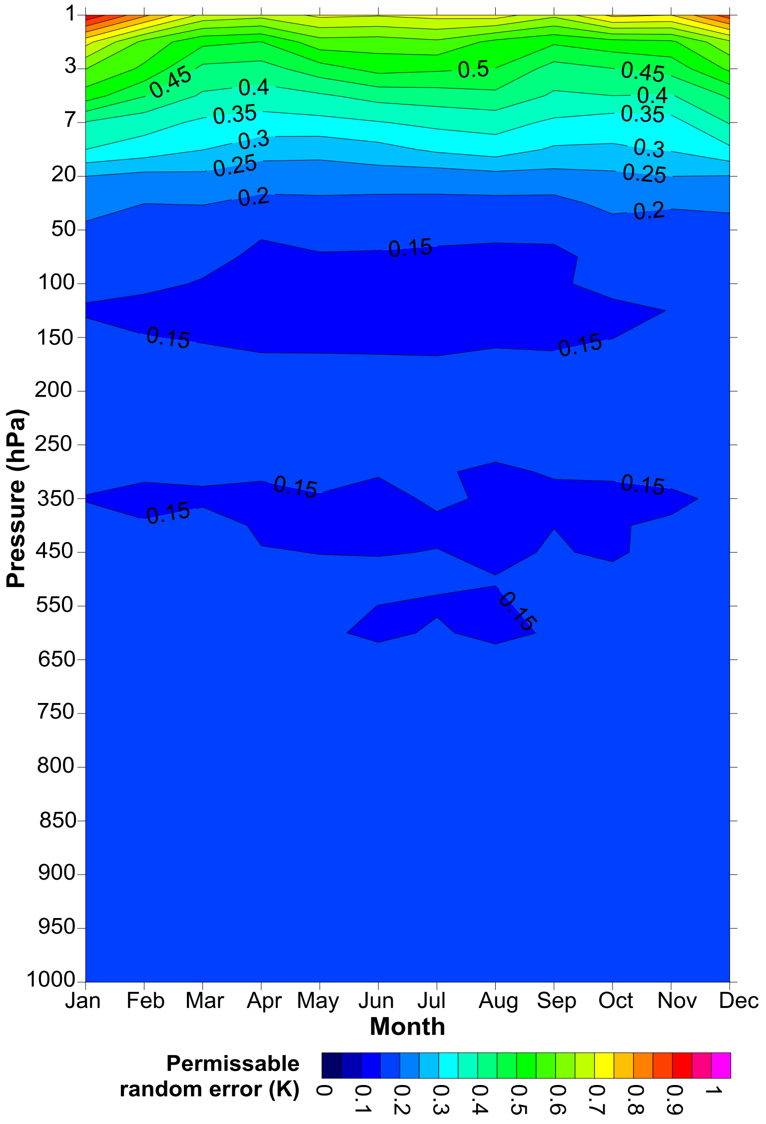

Figure 4: The permissible random error on temperature measurements, when measured every 12 hours at midnight, as a function of pressure and season, required to avoid a more than 10% increase in the uncertainty on the monthly means compared to the uncertainty that would result from sampling with 0.01 K random error. Results from all 87 sites selected for this analysis were averaged to produce this figure.

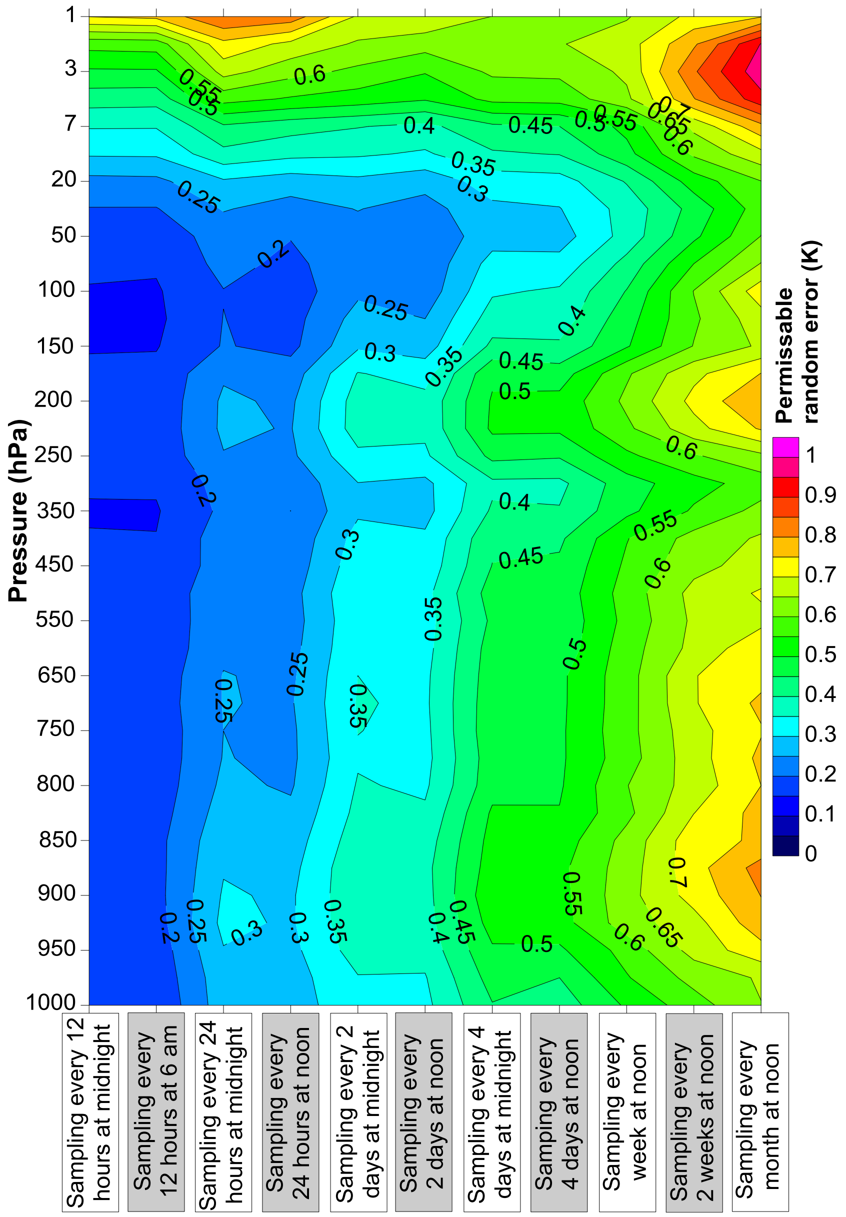

Figure 5: The permissible random error on temperature measurements across a range of sampling strategies required to avoid a more than 10% increase in the random error on the monthly means compared to the monthly mean random error that would result from sampling with 0.01 K random error. Results from all 87 sites selected for this analysis and for all months were averaged to produce this figure.

Results such as those shown in Figure 4 are summarized in Figure 5. As the frequency of sampling decreases, so the sampling random error comes to dominate and less stringent random error requirements on each measurement result. For operational radiosonde sounding stations making temperature profile measurements twice daily, there is something to be gained by reducing the random error on each measurement to 0.2 K or better since this minimizes the random error on the resultant monthly means, thereby allowing for more robust estimates of upper air temperature trends.

For sites where sampling is every 3 days or so, as discussed in Seidel and Free, measurement random errors of 0.5 K are sufficient to ensure that there is no additional increase in the random error on the resultant monthly means.

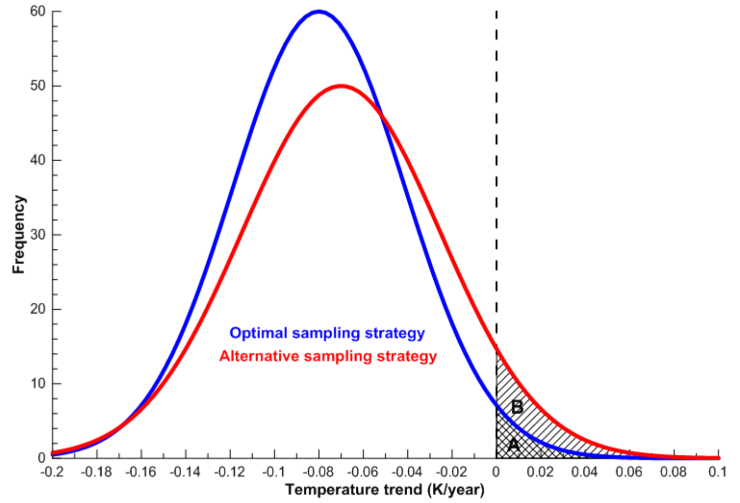

Figure 6: A pedagogical example of the metric used to quantify the degradation in trend resolution as a result of different sampling strategies (measurement frequency and assumed random error on each measurement). The ratio of area A to area B defines the extent to which an alternative sampling strategy compromises the extent to which a trend can be statistically significantly differentiated from zero.

The effects of individual measurement random error and sampling strategy on the ability to detect upper air temperature trends has also been investigated using the NCEPCFSR reanalyses. Temperature trends were calculated at each of the 37 pressure levels, for each of the 87 locations used in the analyses presented above, using a state-of-the-art regression model (Bodeker et al., 1998). Residuals from the regression model fit were then used in a Monte Carlo bootstrap resampling to create 1000 statistically identical time series, each of which was then passed through the regression model to create a histogram of trends. Blocks of residuals are selected so as to preserve the autocorrelation structure in the original time series. This method was applied to each of the monthly mean time series, as generated above, based on different assumptions about the random error on each of the individual temperature measurements, and the 12 different sampling strategies.

Figure 6 shows an example of two hypothetical histograms of temperature trends. The optimal sampling strategy is taken to be sampling every 6 hours with 0.01 K random error on each measurement and is then used to calculate the monthly means which form the input to the trend analysis. The extent to which any alternative measurement strategy might compromise the ability to statistically differentiate the trend from zero can be calculated from the ratio of area A to area B.

Figure 7 shows an example of the effects of random error on individual measurements and sampling strategy on the resolution of temperature trends at 85°N, 135°W and 1 hPa. In some cases, by chance, a ‘weaker’ measurement strategy may lead to better (in terms of being statistically significantly different from zero) resolution of a trend. At this location and pressure, sampling less frequent than once weekly, and with measurement random error ≥2 K significantly degrades the quality of trend detection.

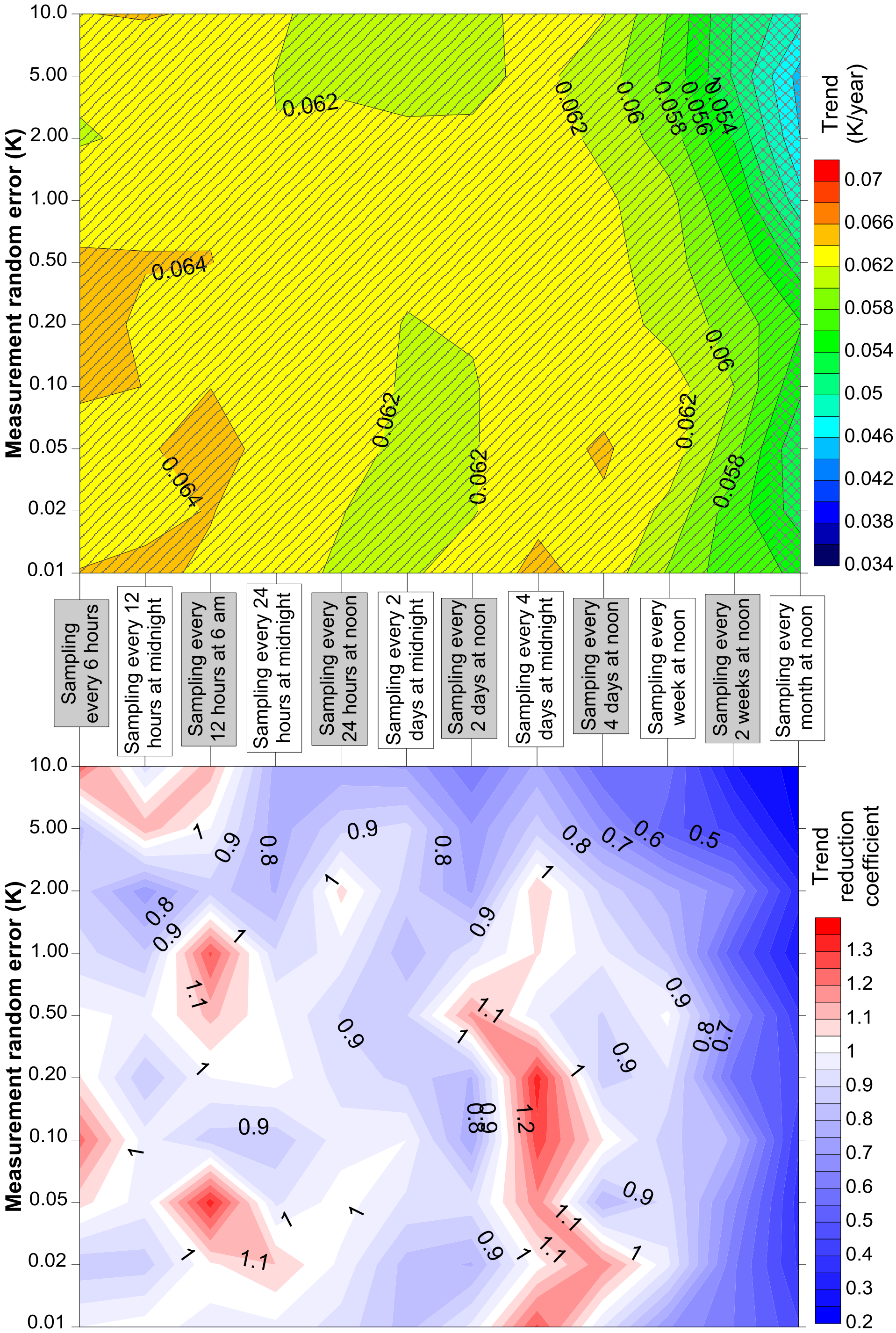

Figure 7: Upper panel: Annual mean trends at 1 hPa, 85°N, 135°W, as a function of random error on individual measurements used to calculate the monthly means used as input to the regression analysis, and sampling strategy. Regions with single hatching show where trend are statistically significantly different from zero at between 1σ and 2σ. Regions with double hatching show where the trend is not statistically significantly different from zero at the 1σ level. Lower panel: the trend reduction coefficient. Red regions show where a sampling strategy leads to improved resolution of the trend (compared to 6 hours sampling at 0.01 K random error) and blue regions show where a sampling strategy leads to a degradation of the trend resolution.

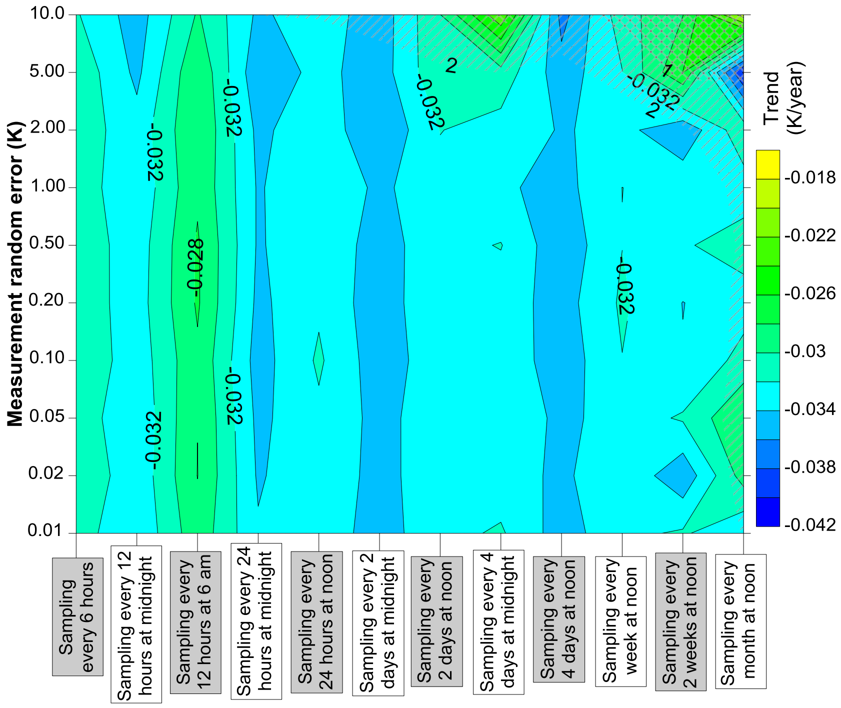

Figure 8: Upper panel: Annual mean trends at 50 hPa, 39.95°N, 105.2°W, as a function of random error on individual measurements used to calculate the monthly means used as input to the regression analysis, and sampling strategy. Regions with single hatching show where trend are statistically significantly different from zero at between 1σ and 2σ. Regions with double hatching show where the trend is not statistically significantly different from zero at the 1σ level.

A second example is given for 39.95°N, 105.2°W at 50 hPa in Figure 8. At this location the temperature trends of ~-0.32 K/decade are highly statistically significant (such that none of the 1000 Monte Carlo simulations produced positive trends) and are robust against almost all of the combination of measurement random error and sampling strategy. It is only at when measurement random errors exceed 2 K and measurements are made only once or twice a month that the quality of the trend determination is compromised.

6.3. Water vapour profile data product

An analysis similar to that done for temperature has not been done for water vapour since it is not clear whether the quality of the NCEPCFSR water vapour product is sufficiently good to capture water vapour trends in the stratosphere. A thorough assessment of the reanalyses for water vapour is required before an analysis similar to that done for temperature can be undertaken.