Executive summary

- The assessment of long-term observations by LOTUS confirms the significant decline of ozone concentrations in the upper stratosphere (at altitudes above the 10–5 hPa level) between January 1985 and December 1996. The strongest trends are observed near 2 hPa (~42 km) with values of 5.9–6.2 % per decade at mid-latitudes and 4.8 % per decade in the tropics. Trends are significant at more than 5 standard deviations in this altitude range.

- Trends derived from satellite and ground-based records in the pre-1997 time period agree with climate model simulations within respective uncertainties thus confirming our understanding of ozone loss processes in the upper stratosphere during that period.

- Between January 2000 and December 2016, positive trends are obtained throughout the upper stratosphere for satellite and ground-based records. The combined trends from six merged satellite records are larger in the Northern Hemisphere mid-latitudes (2–3 % per decade between ~5–1 hPa) than in the tropics (1–1.5 % per decade between ~3–1 hPa) and Southern Hemisphere mid-latitudes (~2 % per decade near 2 hPa). Statistical confidence is largest for trends in the Northern Hemisphere mid-latitudes.

- For altitudes below the 4 hPa level, ozone trends in the post-2000 time period are not significant. Though not significant, negative ozone trends of 0.5–1.5 % per decade are consistently detected by multiple satellite combined records in the 50–15 hPa altitude range over the tropics. Trends derived from ground-based data and Chemistry-Climate Model Initiative (CCMI) model simulations are generally consistent but more variable in this region. The mean CCMI model trend is negative at altitudes below 30 hPa, but the range of individual model trends is large; trends in ground-based records tend to be negative at 20 hPa but increase at lower altitudes (except trends from the microwave records). At mid-latitudes, the trends are close to zero down to 50 hPa.

- Larger differences in post-2000 trends from the various records are observed in the lowermost stratosphere (100–50 hPa) in all latitude bands. Non-significant negative trends are derived from merged satellite records over the tropics and the Northern Hemisphere mid-latitudes. Model simulations show positive trends in the mid-latitudes in both hemispheres in this altitude range, although the trends are not statistically significant.

- LOTUS estimates of past and recent ozone trends are in fairly good agreement with results from previous studies. For the post-2000 period, the largest differences are found throughout the middle stratosphere. These differences stem primarily from extensions of and revisions to existing data records, the addition of new data records, and in some cases the use of a different trend model.

- While trend values in recent studies are fairly similar, the uncertainties and hence significances of the combined trends in broad latitude bands differ substantially. The LOTUS approach, based on both error propagation and standard error of the mean, also explicitly accounts for correlation between the data sets, which results in more conservative uncertainties and thus lower, but more realistic, confidence in positive upper stratospheric trend values compared to the most recently published assessment of merged satellite data set trends.

Have ozone concentrations in the stratosphere significantly increased since the end of the 1990s when levels of ozone depleting substances (ODSs) started to decline? Finding an answer to this question is of great societal importance to ensure that the measures taken by the Montreal Protocol and subsequent amendments to reduce ODSs continue to adequately protect the ozone layer. However, the confidence with which we can assess changes in stratospheric ozone since the mid-1990s has been the subject of considerable scientific debate in recent years, as it depends on the data sets and the analysis methods used. Settling this scientific debate is one of the main objectives of the LOTUS activity, short for Long-term Ozone Trends and Uncertainties in the Stratosphere.

Below, we summarise the main results obtained during the first phase of LOTUS, which was primarily targeted at providing timely input to the 2018 World Meteorological Organization (WMO) Ozone Assessment (*WMO*, 2018). During this phase we reevaluated the satellite and ground-based data records as well as the time series analysis methods commonly used to derive long-term trends. Using a single "LOTUS regression" model, we reassessed past and recent trends in the vertical distribution of stratospheric ozone from the updated individual data records. We then developed a new approach for combining the individual trend estimates from satellite-based records into a single best estimate of ozone profile trends with associated uncertainty estimates. Finally, we compared the satellite-based profile trends in broad latitude bands to trends from ground-based data, from the collection of CCMI-1 model simulations, and from past evaluations of satellite-based trends in peer-reviewed literature.

Some regions in the stratosphere have not been considered (*e.g.*, polar) or have not been analysed in full detail (*e.g.*, lower stratosphere) because of the timeline for the 2018 WMO Ozone Assessment (*WMO*, 2018).

ES.1 New and improved data sets for trend analyses

The work performed in LOTUS and the resulting trends are based on the latest observations from single and merged satellite records as well as from ground-based instruments. In addition to including four additional years of data compared to the results published in the 2014 WMO Ozone Assessment and in the framework of the SI2N activity (*i.e.*, *Harris et al.*, 2015; and references therein), many of the records utilised in LOTUS have been improved for trend analyses (*i.e.*, new methods to combine/homogenise data sets, sampling corrections, and revised calibration and pointing stability). With nearly global coverage and durations spanning at least 30 years (1985–2016), the combined satellite records constitute the backbone of LOTUS trend analyses. With respect to ground-based measurements, LOTUS used individual records from passive and active remote sensing techniques as well as from ozonesondes, including the few homogenised sonde records that exist. In addition to observations, CCMI model simulations were used to test our understanding of ozone profile trends and results shown here represent the first analysis of trends in the vertical distribution of ozone from the CCMI-1 REF-C2 simulations. The details of all observational and model data used in LOTUS, as well as the methods for averaging and merging data records, are discussed in *Chapter 2* of this Report.

ES.2 Addressing the challenges with data for trend estimation

Any assessment of trends and uncertainties necessitates the investigation of the data sets themselves as well as the nuances of their creation, in the case of merged data, and utilisation. The intercomparisons of the satellite and ground-based ozone time series reported in *Chapter 3* reveal a number of measurement artifacts (*e.g.*, drifts, discontinuities, and spikes) but generally show good agreement. In fact, the agreement between observational records is better than for earlier versions of merged ozone records used for previous assessments (*i.e.*, *WMO*, 2014; *Harris et al.*, 2015; and references therein), which lends increased confidence in derived trends. Acknowledging and understanding potential anomalies is important for explaining differences in the trends and trend uncertainties and provides guidance on how to improve the data sets. Availability of data records by a number of complementary instruments is key in singling out these issues and attributing them to one of the data records. Recently proposed Bayesian analyses (*Section 3.1.5*) may further help to systematically identify artifacts in particular data sets, while Monte Carlo (MC) methods (*Section 3.1.4*) can help assess the impact of remaining uncertainties in the records on the final trend estimates. Sampling biases also come into play, since the ozone time series are regressed at an aggregate level and lead to systematic changes in derived trends by up to 1–2 % per decade in parts of the stratosphere, which constitute a considerable fraction of the estimates of post-2000 trends.

ES.3 Sensitivity testing for a consensus regression model

One of the primary goals of LOTUS is to assess the impact of analysis methods on derived ozone trends and their uncertainties. In that regard, a test of 15 previously published multiple linear regression (MLR) models applied to a common data set was performed to evaluate the sensitivity of derived trends to different methodologies. The results showed good agreement in the shape of retrieved trends but a general spread in derived trend values after 2000 of 1–2 % per decade, with overall differences as high as 3 % per decade, which revealed the need for additional sensitivity tests in order to create a consensus analysis method.

Chapter 4 of this Report details a series of sensitivity tests pertaining to the impact of different geophysical and empirical proxies used in MLR analyses on the derived trends and their uncertainties. Since work during the first phase of LOTUS was optimised towards estimating middle and upper stratospheric trends in satellite data sets, we focused on the sensitivity of proxies to ozone variability at these altitudes. Accordingly, results showed that short period proxies (*e.g.*, AO[1] , AAO[1], NAO[1], and EHF[1]) had negligible effects on trends and variably small impacts on uncertainties while excluding the solar cycle, QBO[1], or ENSO[1] proxies from the regression model had significant effects on the trend (1–2 % per decade difference) and uncertainty (around 1 % per decade) estimates. Different long-term trend proxies were also investigated, revealing the complexity of attempting to capture changes in ozone stemming from the influences of both ODSs and greenhouse gases. Ultimately a single consensus "LOTUS regression" model, based on a simple yet appropriate set of geophysical proxies and a trend proxy designed to capture mean trends in satellite data sets, was chosen for our analyses and was also packaged for public use (**https://arg.usask.ca/docs/LOTUS_regression**).

Lastly, while a more traditional MLR-based approach was chosen for the majority of work in LOTUS, a newer statistical approach, namely a dynamic linear model (DLM),

[1] Abbreviations for proxies: AO = Arctic Oscillation; AAO = Antarctic Oscillation; NAO = North Atlantic Oscillation; EHF = Eddy Heat Flux; QBO = Quasi-Biennial Oscillation; ENSO = El Niño Southern Oscillation.

was also tested. The DLM technique, as its name implies, does not constrain the shape of the likely nonlinear long-term trend and thus may ultimately be better suited for this type of work. However, this method was not fully evaluated during the first phase of LOTUS and a more comprehensive comparison of MLR and DLM for ozone trend estimates is still needed.

ES.4 Examination of trend results from individual data sets

The "LOTUS regression" model was first applied to the individual data sets at their native resolution. Results display expected patterns of ozone decline in the upper stratosphere prior to the late 1990s and a subsequent smaller increase since 2000 in generally good agreement with models, though the magnitude and statistical significance of these results vary between the different data sets. Most data records and model simulations point to continuously declining ozone levels in the tropical lower stratosphere in a fairly coherent yet generally not statistically significant way, but results at mid-latitudes in the lower stratosphere are variable and inconclusive. Additionally, the differences between trend results are discussed as they pertain to the differences in the various data sets and merging techniques (see *Chapter 3*). In order to place these results in context with previous comprehensive studies, the regressions are repeated with the data sets averaged over broad latitude bands (*i.e.*, 60°S–35°S and 35°N–60°N representing southern and northern mid-latitudes respectively and 20°S–20°N representing the tropics) prior to continuing with the overall analysis.

ES.5 Revised approach to combine trends

The typical desire for a single set of spatially resolved trend results (*e.g.*, as in previous Ozone Assessments) when faced with an ensemble derived from the analyses of multiple data sets creates the unique challenge of merging not only the ensemble of trend results but also their uncertainties. With an aim towards determining if these results are statistically significant, it is the latter component that is more important and often more complicated. Historically, a variety of techniques have been used to merge the uncertainties and the LOTUS Report introduces a new, statistically more robust method. This includes not only components involving simple error propagation, which captures uncertainties introduced from the data and the analyses, and the standard error of the mean, which captures systematic uncertainties such as those introduced by drifts between data sets, but also the correlations between the data sets themselves, estimated from the correlation of the fit residuals. Consequently, it is the nature of the independence of the data sets and their resulting trends that is the most important aspect of merging the trend uncertainties. *Section 5.3* in this Report details the complexities of this merging and the necessary assumptions chosen for the LOTUS work, which are balanced between not wanting to overestimate or underestimate the combined uncertainties. The results presented here have uncertainties that fall between previous comprehensive works but err towards a more conservative estimate. Ultimately, this work concludes that the most meaningful way to improve the uncertainties in future analyses would be to reconcile the discrepancies between the data sets themselves prior to the merging process.

ES.6 Assessment of combined ozone profile trends

Estimates of combined satellite trends are summarised in **Figure ES.1** and in **Table ES.1** (see *Section 5.6*). For this work, results are separated into two distinct time periods with "pre-1997" being defined as the period from January 1985 to December 1996, while "post-2000" refers to the period from January 2000 to December 2016. Comparisons of LOTUS trends (hereafter L19) with previously published trends (WMO, 2014, hereafter W14; Harris et al., 2015, hereafter H15; Steinbrecht et al., 2017, hereafter S17) are shown in **Figure ES.1** as well.

ES.6.1 1985–1996 trends

Negative trends are found across nearly the entire stratosphere in the pre-1997 period for almost all satellite and ground-based data records. Individual and combined satellite data show highly statistically significant evidence of declining ozone concentrations in the upper stratosphere (at altitudes above the 10–5 hPa level) since the mid-1980s and well into the 1990s. The depletion reaches a maximum rate near 2 hPa (~42 km) of 5.9–6.2 % per decade at mid-latitudes and 4.8 % per decade in the tropics (see **Table S5.1** in the *Supplement*). Ozone decline rates in the middle stratosphere (30–15 hPa) are considerably smaller, with statistically insignificant values of at most 1–2 % per decade. Negative trends are found across the lower stratosphere (down to 50 hPa), while in the lowermost stratosphere (down to 100 hPa) trends differ according to latitude, with large significant negative trends of about 5 % per decade in the Northern Hemisphere. However, confidence in trend results is reduced in the lower stratosphere due to large natural variability, low ozone values, and decreased sensitivity of satellite observations. Trends derived from ground-based measurements generally corroborate satellite trend results. However, due to their larger sparseness in space and time, especially during this early period, the significance of the trends is not as high and trend values differ. Results agree well with those of model simulations (within 1 % per decade) throughout the middle and upper stratosphere at all latitudes, lending confidence that these losses in ozone were the result of chemical forcing from ODSs according to model predictions. However, larger differences exist between satellite and model results in the lowermost stratosphere, with disagreements outside the large uncertainties only in the Southern Hemisphere.

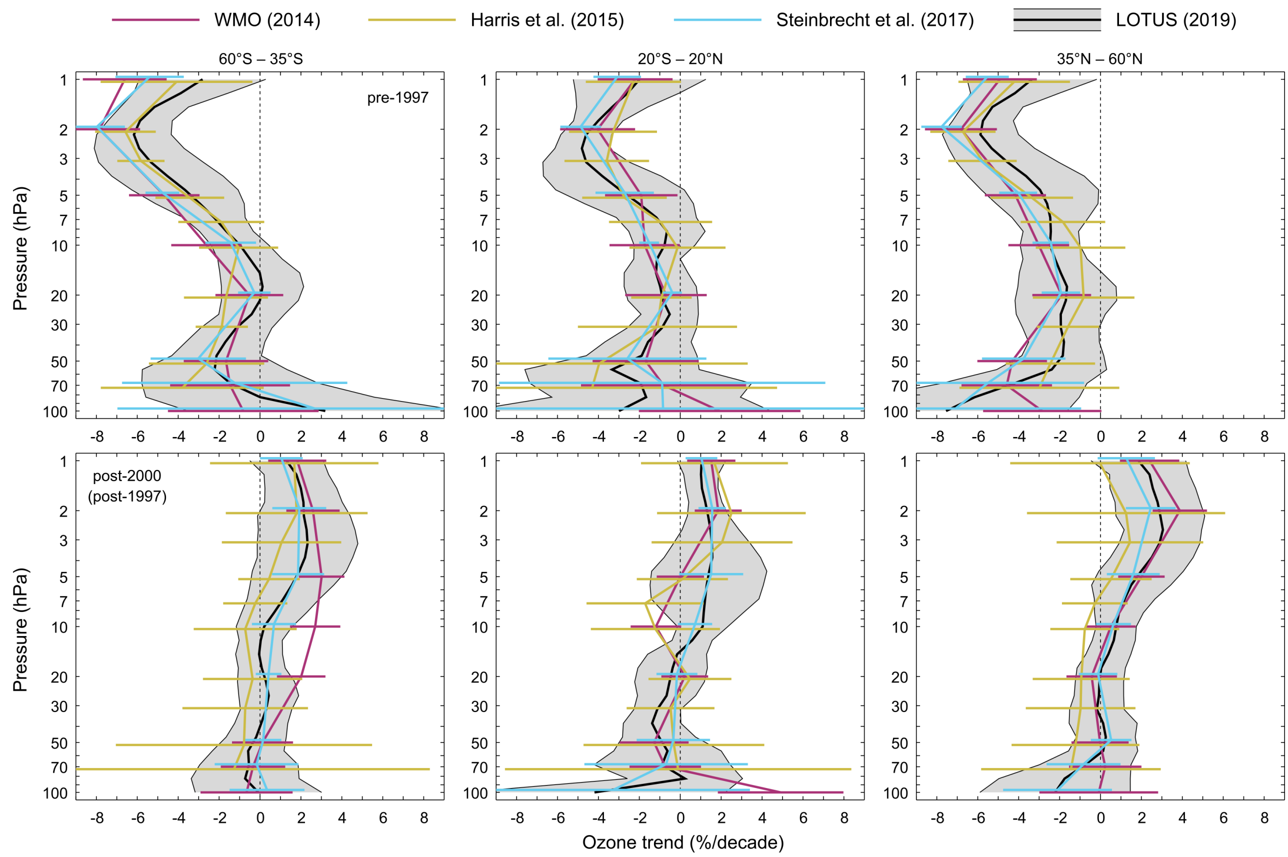

Figure ES.1: *Overview of ozone profile trends from past and recent assessments: WMO (2014), Harris et al. (2015), Steinbrecht et al. (2017), and LOTUS (this work) are shown in red, orange, blue, and black respectively. Top row shows trends before the turnaround of ODSs and bottom row since the turnaround (analysis time period differs by assessment). Shaded area and error bars represent the 95 % confidence interval for the combined trend. Coloured profiles are slightly offset on the vertical axis for display purposes. This figure is also shown in Chapter 5 as **Figure 5.12**. LOTUS results are tabulated for each pressure level in **Table S5.1** in Supplement. Steinbrecht et al. (2017) did not report or discuss pre-1997 trends, but results shown here were obtained from that work (private communication).*

	Ozone trend Jan 1985 - Dec 1996 (% per decade, ± 2 σ)			**Ozone trend Jan 2000 - Dec 2016** (% per decade, ± 2 σ)		
Pressure (hPa)	**60-35°S**	**20°S-35°N**	**35-60°N**	**60-35°S**	**20°S-35°N**	35-60°N
1	-2.8 ± 3.1	-2.0 ± 3.2	-3.3 ± 3.2	1.3 ± 1.8	1.0 ± 1.2	1.9 ± 2.3
2	-6.0 ± 1.7	-4.4 ± 1.2	-5.8 ± 1.7	2.1 ± 2.1	1.3 ± 0.9	2.9 ± 2.1
5	-3.4 ± 2.4	-2.6 ± 3.0	-2.8 ± 2.7	1.8 ± 2.2	1.4 ± 2.8	1.8 ± 2.0
7	-2.2 ± 1.5	-1.1 ± 1.9	-2.5 ± 1.8	1.1 ± 1.7	1.2 ± 2.5	1.1 ± 1.4
10	1.0 ± 1.4	-0.8 ± 1.5	-2.5 ± 1.5	0.2 ± 1.4	1.1 ± 1.2	0.8 ± 1.0
20	0.0 ± 1.9	-0.9 ± 1.7	-1.7 ± 2.5	0.2 ± 1.3	-0.5 ± 1.7	0.0 ± 1.2
50	-2.2 ± 2.7	-2.4 ±3.3	-2.1 ± 2.3	-0.3 ± 1.6	-0.9 ± 2.1	0.2 ± 1.5
70	-1.2 ± 4.5	-1.9 ± 5.3	-4.5 ± 3.8	-0.6 ± 2.4	-0.7 ± 3.3	-0.9 ± 2.4

Table ES.1: *Overview of LOTUS combined satellite trends in three latitude bands and two time periods. Central values and uncertainties representing the 95 % confidence interval are listed in the table. Trend results that are statistically significant at the 2-sigma level are highlighted in grey cells. See also **Figure ES.1**. Please note that trends and uncertainties are interpolated onto pressure levels that are common to other studies (e.g., WMO, 2014; Steinbrecht et al., 2017) to facilitate comparisons between these studies and LOTUS. Trends discussed in Chapter 5 are presented on the LOTUS pressure levels, which have a higher vertical resolution, and these are tabulated in **Table S5.1** in Supplement.*

ES.6.2 2000–2016 trends

Positive trends are found throughout the upper stratosphere and part of the middle stratosphere in the post-2000 period for both satellite and ground-based trends, though results vary for ground-based data depending upon the observation technique. Results from satellites show statistically significant positive trends in the Northern Hemisphere at mid-latitudes of 2–3 % per decade in the upper stratosphere (between ~5–1 hPa) and 1–1.5 % per decade in the tropics (between ~3–1 hPa).

Positive trends of ~2 % per decade are also found in the Southern Hemisphere near 2 hPa at mid-latitudes though the statistical confidence is smaller. In the Northern Hemisphere mid-latitudes, trends in the upper stratosphere are significant down to 4 hPa. At altitudes below 4 hPa, mid-latitude trends are no longer statistically significant, dropping from positive 1.8 % per decade at 5 hPa to near zero between 50–20 hPa. In the tropics, trends become negative below 15 hPa though the estimates of 0.5–1.5 % per decade are statistically insignificant as well. Generally, these satellite-based results are in agreement with ground-based observations and model simulations. The persistent negative trends in the middle and lower stratosphere over the tropics are likely the consequence of radiative and dynamical forcing from greenhouse gases according to model predictions (*WMO*, 2014; and references therein). Derived trends differ considerably in the lowermost stratosphere, below 50 hPa, depending on the data set and latitude. For example, satellite-based results show statistically insignificant negative trends (or near zero in the Southern Hemisphere) and ground-based trends agree in sign except in the tropics where there are significant positive trends. Model simulations, however, predict positive trends in mid-latitudes in both hemispheres in this altitude range.

ES.6.3 Comparison with previous assessments

LOTUS estimates of past and recent ozone trends are in fairly good agreement with results from previous assessments (*e.g.*, W14, H15, S17, and references therein). L19 and S17 trends differ by less than 0.5 % per decade in the post-2000 period, which is expected since similar data sets and a similar regression model were used for both studies. Trends by W14 and by H15 are in reasonable agreement with L19 as well, though larger differences are noted for the post-2000 period at Southern mid-latitudes, in the tropical middle stratosphere, and, for H15, in the Northern mid-latitudes as well. These differences stem primarily from extensions of and revisions to existing data records, the addition of new data records, and the use of a different trend proxy (*e.g.*, H15 assumed an inflection point at 1997).

While trend values in W14, H15, S17, and L19 are fairly similar, the uncertainties and hence significances of the combined trends differ substantially. This is the most critical component for the detection of the relatively small post-2000 trends. Even though both S17 and L19 use similar data records, the L19 approach, based on both error propagation and standard error of the mean, yields different uncertainties as compared to S17. In the upper and middle stratosphere, uncertainties from the standard error dominate and thus the estimate of the independence of the data records is critical in testing the trend null hypothesis. S17 trends are statistically significant across the entire upper stratosphere while in the L19 trend analysis, which derives a smaller degree of data independence from the correlation of fit residuals, high significance is found only at Northern mid-latitudes, and less significant trends are found in the tropics and at Southern mid-latitudes. There is not sufficient information in the trend analyses that can help determine exactly how independent the different data sets are. However, it is concluded that the real trend uncertainty lies in between S17 and L19 uncertainty estimates in the upper stratosphere. On the other hand, analysis of results of the two other assessments considered in the study suggested that they used either a too conservative (H15) or too optimistic (W14) approach to estimate combined uncertainties.

In the lower stratosphere, ozone trends are affected by large atmospheric variability and decreased sensitivity of satellite measurements. The L19 approach, which includes a term for error propagation from the regression coefficients, is not capable of capturing all sources of uncertainty and most importantly measurement drift, which leads to the conclusion that uncertainties derived from the analysis may be underestimated.

ES.6.4 Open issues and future work

The LOTUS Report assessment of satellite and ground-based ozone data sets (*Chapter 2*) builds the foundation for reconciling the discrepancies in ozone trends estimated from the individual climate data records. Understanding the causes of these differences would create improvements not only in the internal consistency of data sets, but also in the uncertainties of overall ozone trends. Further, development of techniques to directly assess uncertainties in the merged records resulting from discrepancies that cannot be completely reconciled, such as small relative drifts and differences resulting from coordinate transformations and sampling differences, would allow for a more precise estimate of significance of the mean trend.

For the satellite and ground-based data used in the LOTUS Report, information on stability and drifts of the measurement is still incomplete (*Chapter 3*).

The homogenisation of ozonesonde records was not finished prior to their use in the LOTUS assessment, and thus the ozonesonde trends and their uncertainties (especially in the lower stratosphere) may change in the future. In addition, in order to properly combine instrument-specific trends, a common matrix for providing error budget information for each ozone record is needed. Work developing a common approach to assessing errors in Level 2 satellite data is ongoing under the SPARC "Towards Unified Error Reporting (TUNER)" activity and ozone record uncertainties are addressed in other SPARC (Stratosphere-troposphere Processes And their Role in Climate) activities. Standardised error budgets have also been defined within the Network for the Detection of Atmospheric Composition Change (NDACC) and are in the process of being included in the data records.

The common statistical linear regression trend model (*Chapter 4*) used in the LOTUS Report was optimised for analyses of the zonally averaged satellite data sets. However, analyses of the ground-based data require reconsideration of additional proxies (*i.e.*, lag for ENSO, AO, AAO, NAO, EHF, *etc.*) and optimisation methods that can improve interpretation of the processes that impact ozone changes over the limited geophysical region and reduce trend uncertainties (*Chapter 5*).

The first attempt to evaluate representativeness of the ground-based station records for the middle and upper stratosphere using Solar Backscatter Ultraviolet Radiometer (SBUV) data was done under the LOTUS Report activity and discussed in *Chapter 4*. Comparisons of trends derived from satellite data selected under overpass criteria against zonally averaged trends will help with interpretation of stability in all observing systems and determine ozone recovery with high confidence.

There is a clear need for future activities of the CCMI modeling community, with experiments designed with the view on the verification of simulated trends. A large number of models is absolutely necessary in order to be able to assess the ozone variability associated with chemistry and dynamical transport mechanisms. Moreover, an assessment of model sensitivity to uncertainties in the volcanic aerosols, solar cycle, QBO, ENSO and other mechanisms is considered of great importance in order to advance our understanding of the ozone layer variability and associated response to natural variability.

In this Report, the ozone trends are analysed at low and middle latitudes, with a focus on the upper and middle stratosphere. Future works would explore trends in polar regions and in the lower stratosphere, which can be done in conjunction with the SPARC activity Observed Composition Trends And Variability in the Upper Troposphere and Lower Stratosphere (OCTAV-UTLS), dedicated to the assessment of the composition of the upper troposphere and lower stratosphere (UTLS) and identification of atmospheric processes that impact UTLS changes on the decadal scales. Similarly, the trends derived from total column data are also left for future work.

Assessments similar to the LOTUS activity need to be regularly repeated, preferably in collaboration with other SPARC and WMO/GAW (Global Atmospheric Watch) activities.