Workshop Report

I Introduction

IHOP scientists gathered for its 1st Intercomparison Workshop at the National Center for Atmospheric Research on October 2/3, 2003. This workshop was focusing on the specification of water vapor measurements performed during IHOP. 27 participants attended the workshop coming from the instrumentation but also from the modeling community.

The workshop was run by Volker Wulfmeyer of the Institute of Physics and Meteorology (IPM) at the University of Hohenheim (UHOH). Wulfmeyer pointed out that the goals of the IHOP Instrumentation Component (IC) are described in detail in chapter 7 of the IHOP Operations Plan. In short, the goals are to

- demonstrate and characterize new moisture measuring instruments and techniques,
- perform new types of intercomparisons between previously unavailable combinations of sensors, improve the knowledge about performance of water vapor instrumentation under different atmospheric conditions.

It is obvious, that the Intercomparison Component plays a fundamental role in connection with IHOP projects, particularly to perform studies on

- model validation,
- atmospheric processes,
- and data assimilation.

Therefore, it was an important step to hold this workshop in order to permit these kind of studies as soon as possible in the future in the framework of IHOP.

The workshop was organized as follows: The first day of the workshop was divided into three sessions of talks. The first session was the introduction and the next session was on “Instrument Specification and Performance”. This session was supposed to give the workshop participants an introduction in the measurement methodologies of the various sensors before going into the details and presenting the intercomparisons. The latter was performed in the final session “Comparisons” . This session contained an overview of the present state of the intercomparisons and the characterization of different moisture measuring instrumentation. On the next day, an in-depth discussion concerning the results and future activities took place including the definition of several action items.

II The talks in a nutshell

1. First session: Introduction

a) The workshop was opened by Dave Parsons from NCAR/ATD/MMM who welcomed the participants and noted the importance of this workshop.

b) The opening remarks were followed by an overview presentation of Volker Wulfmeyer from IPM/UHOH who welcomed the participants and summarized the IHOP Intercomparison Component (IC) goals. To reach these goals, Wulfmeyer suggested an extension of the description of the archived IHOP data. The data should be specified with respect to their resolution and a detailed error analysis should be added.
This list of specifications should include (if applicable)

- horizontal resolution including the definition of the weighting function,
- vertical resolution including the definition of the weighting function,
- systematic errors or the bias in the data (the remaining overall error which cannot be removed by averaging),
- noise errors in dependence of range and time resolutions,
- non-representativeness of the data.

The participants acknowledged this list and discussed mainly the latter specification. It was clarified that this error must be considered with respect to the area of the interest (e.g. model resolution) where atmospheric studies are performed. On the other hand, it was agreed that the analysis of the non-representativeness cannot be performed by the user of the data alone. An interface between data provider and user is essential to assess this kind of error as accurate as possible. Finally, this specification list should lead to the development of error covariance matrixes of the instruments under all atmospheric conditions.
Wulfmeyer also reviewed the intercomparisons which are currently available. These included extended comparisons of lidar systems with each other and with other instruments. Also a large data set of comparisons between different types of radiosondes exists. GPS data have been compared with microwave radiometers. Also the radar refractivity technique has been compared with several types of sensors. However, no representatives of the NAST or satellite instruments participated in the workshop so the status of these instruments should be requested.

Wulfmeyer finished this presentation with a summary of expectations from this workshop. He suggested further efforts for coordinating data intercomparisons. This could be ensured by summarizing the performed data intercomparisons on the IHOP web page. This would also help to identify missing intercomparisons which should be performed to improve the specification of the IHOP water vapor instrumentation. The IHOP web page should also be expanded by sections describing water vapor instrumentation and their specifications. This should include flow charts of the data analysis and a clear definition and characterization of all errors (see above). Finally, he suggested also a “blind test” where simulated water vapor DIAL data are sent to all DIAL groups in order to compare their data processing algorithms.

c) In the next presentation “The UHOH ESA WALES/IHOP Project”, Volker Wulfmeyer introduced a basic intercomparison activity which has been funded by the European Space Agency (ESA). The motivation of this project, which is called “WALES Verification”, is the proposed Earth Explorer Mission Water Vapour Lidar Experiment in Space (WALES). The goal of this mission is the measurement of water vapor profiles with global coverage using a space borne differential absorption lidar (DIAL) system. After giving an overview of this mission, Wulfmeyer presented the mission concept and summary. He demonstrated that analytical error analyses of this mission have been performed which indicate excellent performance and resolution. For instance, it can be expected that measurements can be performed with a horizontal resolution of 100 km, a vertical resolution of 1-2 km resulting in an noise error of less than 20% from ground to the upper troposphere/lowermost stratosphere. However, a refinement of the performance analysis of WALES is considered essential. This includes a thorough characterization of state-of-the-art DIAL systems. This is the goal of the WALES Verification project. An ideal data set are the IHOP measurements. Intercomparisons between different DIAL systems and other sensors were performed for verifying the expected performance of WALES and potentially for refining the WALES Mission Requirements. Nearly all IHOP lidar intercomparisons activities have been supported by this ESA project.

d) Andreas Behrendt from IPM/UHOH introduced the comparison efforts performed so far within the ESA project "WALES verification". He gave an overview of all possibilities for intercomparisons of the lidar data obtained during IHOP. In connection with the WALES Verification project, detailed tables were derived taking all data into account where airmasses were sampled at the same time which have a (minimum) distance less than 20 km. As a result, for the airborne water vapor DIAL systems the following possibilities for intercomparisons were determined:

- for DLR-DIAL and LEANDREII there are two formation flights and 3 additional profile-to-profile intercomparisons
- for DLR-DIAL and LASE there are 7 profile-to-profile intercomparisons
- for LASE and LEANDREII there are no direct intercomparisons.

In the same way, the ground-based Raman lidar data were analyzed showing that the number of possible intercomparisons between DLR-DIAL and SRL is 36, between LEANDREII and SRL 10 intercomparisons are possible, whereas 25 intercomparisons can be made between LASE and SRL. For CARL, 2, 4, and 7 intercomparisons are possible with DLR-DIAL, LEANDREII, and LASE, respectively.

At present, not all lidar data have been released. With the presently available data sets, the numbers of possible intercomparisons between the lidar systems are the following:

- DLR-DIAL and LEANDREII: 1 formation flight
- DLR-DIAL and LASE: 3 profile to profile intercomparisons
- DLR-DIAL and SRL: 9 profile to profile intercomparisons
- LASE and SRL: 19 profile to profile intercomparisons (some of which are not independent).

It should be noted, that in a second step it remains to be investigated for each of these intercomparison cases whether the boundary set with the 20 km minimum distance is stringent enough, i.e., whether the natural heterogeneity of the atmosphere is obviously too large to allow the data to be compared.

Other than lidar-to-lidar intercomparisons, intercomparisons with the Snow White reference radiosondes have been made. Several reference sondes were launched in Homestead near the SRL. One intercomparison is possible with one airborne lidar (LEANDREII, the data of which, however, were not yet available). Furthermore, several comparisons of the DLR-DIAL data with Falcon dropsondes are possible which have not yet been included in the study.

Finally, Behrendt discussed the formula he used to specify differences between the measured data (root-mean-squared difference and mean difference) and illustrated the methods with a set of artificial data.

e) Steve Williams of JOSS presented the current status of the IHOP data management. He gave examples how to access specific data using the IHOP field catalog. The main activities are focused on the production of composite data sets. This is an effort to provide consistent and quality controlled data sets from different sources for the user community. He demonstrated how these data sets are constructed. For instance, these efforts will be applied to provide maps of meteorological data such as precipitation and radiosondes. A difficulty is the access and the combination of data sets provided from different sources.

Presently, 215 of 254 (85%) data sets have been submitted or linked to the IHOP web page. 204 of these are available on line. Data submission guidelines and instructions are also available via the IHOP web page. Whereas the upper air composites have been completed and are available, the surface composites are planned to be released in the Fall 2003. Further details are found on http://www.joss.ucar.edu/ihop/dm/.

2. Second session: Instrument Specifications and Performance

a) Junhong Wang of NCAR/ATD started the session presenting the talk entitled "IHOP radiosonde and dropsonde data: highlights and problems". She discussed the performance of radiosondes, dropsondes and reference radiosondes as well as first science results. She demonstrated that an amount of more than 2600 different radiosonde measurements are available for comparisons. Particularly interesting are measurements with the RS90 as well as with the Snow White (SW) sensors. In addition, 420 sondes were dropped from aircrafts. Comparisons between dropsondes and DLR DIAL showed good agreement.

Mainly during LLJ and CI missions, frequent drops of sondes have been used to determine 2-d cross sections demonstrating the high spatial variability of the water vapor field.

She showed procedures for correction of radiosonde data. In the reference radiosonde, pressure and wind data have to be carefully interpreted. In many cases, the hypsometer used on this sensor produced erroneous pressure measurements. Wind measurements may have errors due to large balloon swings. Dropsonde data may use incorrect geopotential height. Corrections have been developed. Implementation of these corrections are on the way.

b) Subsequently, John Braun of COSMIC/UCAR presented the "GPS instrumentation specification and performance". Data are available from NOAA/FSL and OU/ARM networks. Additionally, during IHOP, 8 additional ground-stations have been set up in the northern part of OK. Operation times are available for each station. Data gaps occurred mainly at the beginning of the campaign but for nearly the full duration of IHOP, all stations performed very well.

There are three kinds of data which can be provided by GPS. PW data are available every 30 min. Here, an accuracy of 1.5 mm can be expected. Slant path measurements (SW) have a resolution of 2 min and a similar accuracy. Tomographic GPS is under development and the expected time resolution is 15 min. The errors are unknown yet and should be determined by comparisons with DIAL data. The PW data are already available on the JOSS web site. For accessing SW and/or tomographic data, John Braun can be contacted directly.

c) In her presentation "Refractivity during IHOP_2002", Tammy Weckwerth from NCAR/ATD introduced S-Pol refractivity measurements. After explaining the measurement techniques, she described the procedure to map refractivity fields. In the warm season, the refractivity changes are primarily sensitive to moisture changes.

The refractivity data are available with a time resolution of 5 min, a horizontal resolution of about 4 km and a range of ~50 km. Preliminary data are available on
http://www.atd.ucar.edu/rtf/projects/ihop_2002/spol/summary/

d) Dave Whiteman of NASA GSFC SRL discussed the "Scanning Raman lidar (SRL) error characteristics and calibration for IHOP". After introducing the Raman lidar system parameters, he described the determination of noise errors using error propagation of Poisson photon statistics. He showed major improvements of Raman lidar measurements during daytime. During daytime, using resolutions of 2 min and 60-210 m, the noise error is typically less than 10 % in the boundary layer.

Several case studies such as a bore event on June 19/20 confirm the excellent resolution of SRL measurements. It followed a detailed analysis of the procedure for assessing the calibration constant of SRL. Whiteman showed that to date it is not feasible to calibrate water vapor Raman lidar without using comparisons with atmospheric water vapor measurements performed with other instruments. This is mainly due to the poor knowledge of the water vapor Raman scattering cross section (10 % uncertainty of the ration of water vapor to nitrogen Raman scattering cross sections). Following this procedure, the potential exists to calibrate SRL with an accuracy of a few percent. For the SRL data analysis, a calibration constant derived using GPS PW achieved during the Aqua validation campaign was used. Comparisons with GPS PW during IHOP revealed that the SRL data are wet by 4 %. Though this deviation is low, reasons for this deviation were discussed during the workshop but no definite conclusion was obtained. It is presently under investigation whether the SRL water vapor data have to be readjusted for the final release.

e) Christoph Kiemle of DLR Oberpfaffenhofen presented "High Resolution Airborne DIAL Measurements of Water Vapor and Vertical Humidity Fluxes during IHOP". He showed that the DLR Falcon collected 51 h of lidar data during 21 IHOP missions. 40 % of these data could be used to retrieve water vapor profiles. After presenting the set up of the DIAL system and the payload of the Falcon aircraft, several results of case studies demonstrated the outstanding resolution of boundary layer measurements. Typically, using a horizontal resolution of 73 m and a vertical resolution of 100-200 m, a noise error of <7 % is maintained. This performance is sufficient to retrieve water variance and flux profiles in the convective boundary layer.

The overall accuracy of the results was mainly influenced by limited spectral purity and frequency stability of the laser transmitter. Attempts have been performed to determine spectral purity for applying correction to the water vapor profiles. For instance, drop sonde profiles were taken as ground truth and the value of spectral purity was determined by fitting the DIAL water vapor profile to the dropsonde data. , when necessary, by adapting water vapor profiles, which have been corrected for spectral impurity, to dropsonde data. Currently, the overall accuracy is estimated to be 4 % in the near-range increasing to 9 % in the far-range. Diagnostics are available to monitor the performance of the laser transmitter. In cases where the transmitter fulfilled the expected performance, excellent agreement with dropsonde data was achieved. He also presented the first airborne flux measurements using collocated water vapor DIAL and NOAA HRDL data. Applying the eddy correlation technique, and averaging the absolute humidity and vertical wind covariance a resulting flux profile showed a value of about 500 W/m^2 and a low flux divergence.

f) As representatives of the LEANDREII DIAL system were not able to make it to the meeting, Volker Wulfmeyer of IPM/UHOH gave a short presentation concerning the "LEANDREII Design, Performance and Operation". He summarized the system parameters of the system and discussed the set up of the laser transmitter. In contrast to other existing DIAL systems, LEANDREII achieves the challenging specifications with respect to the spectral properties of the laser transmitter using spectral narrowing optical filters in a linear alexandrite laser resonator. If the transmitter specifications are maintained, up to the middle troposphere water vapor measurements should be possible with a systematic error of less than 5 % from low-level flights looking up.

LEANDREII collected 142 h of data during 24 IHOP missions. For a major part of the missions, the data quality and the scientific interest in the results is excellent. Efforts are ongoing to perform error analysis of the water vapor data.

3. Third session: Comparisons

a) The session was started by Ed Browell, NASA LaRC, who gave a presentation about "LASE measurements of water vapor during IHOP". He demonstrated the measurement capabilities of the NASA LASE system which includes profiling of the particle scattering ratio, water vapor and – by combination with scanning HIS or radiosondes temperature profiles- relative humidity profiles.

The accuracy of LASE water vapor measurements is typically better than 6 % or 0.01 g/kg, whatever is lower, using resolutions of 330 m in the vertical and 14 km in the horizontal, respectively. Water vapor measurements can be extended to the ground using the ground return. For these data, the accuracy is of the order of 10 %. LASE participated in 8 IHOP missions collecting about 43 h of data.

A large amount of intercomparisons have been analyzed. Mainly data have been considered with a delay of less than 20 min and a distance of less than 20 km. Comparisons between LASE and CARL show very good agreement with a bias of less than 5 %. However, it was pointed out that the CARL calibration constant may be subject of change in the future so that these comparisons may have to be reconsidered. Comparisons of LASE with SRL show similar results even excluding data of flight 11 on June 14, as in the second data release LASE was calibrated using SRL on this day. Also the calibration constant of SRL may be subject to change on a 4 % level in the future (see II 2. c) . But even taking these modifications into account, LASE and Raman lidar data showed a very good agreement.

Comparisons of LASE with drop sondes show larger deviations. For instance, comparisons with DLR Falcon dropsondes indicate that the dropsonde data are drier by about 10 % above the boundary layer. Comparisons with radiosonde data are also available. Deviations between AERI and LASE are of the order of 5 %. **Difference between previous and comparison.

Generally, the comparisons indicate that an overall accuracy of 6 % can be specified for LASE water vapor measurements.

A lively discussion arose concerning the procedure of how to compare the performance of two different instruments. First of all, it is clear that only independent data sets should be compared. Otherwise, the results will be meaningless. Then two different statistical variables can be determined. The mean bias between two instruments can be calculated which was the approach used in these LASE comparisons. However, certain atmospheric conditions may still exist which could cause a bias in a certain direction. The same holds for specific locations where errors can occur such as in regions with large aerosol particle backscatter gradients. If the mean bias is averaged under different atmospheric conditions, this specific bias will be averaged and basically eliminated. Therefore, it was recommended to calculate simultaneously rms deviations which maintain information about specific deviations. It is a topic of the intercomparisons working group to come up with a specific path how the intercomparisons should be compared and interpreted (see below).

b) Andreas Behrendt from IPM/UHOH showed in three presentations the results of the comparisons which were so far possible between DLR-DIAL and the other lidar systems:

I) LEANDREII/DLR DIAL formation flight:

On 7 June 2002, the Falcon and P3 aircrafts with DLR-DIAL and LEANDREII on-board, respectively, performed a formation flight with parallel tracks with minimum distances between 0.3 and 3.0 km. The 48 profile-to-profile intercomparisons, each profile of 10-s integration time, show rms differences of 15 to 40 % for the height intervall of 1.3 to 3.5 km. The mean of all rms differences is (24 ± 7) % or (1.0 ± 0.2) g/kg. The largest deviations are found near the boundary layer top and within the boundary layer where also the natural heterogeneity of the atmosphere is largest. Behrendt analyzed also the variability within the LEANDREII data set and found that here the rms difference between the profiles with ~3 km distance is – depending on height – between 5 and 25 % (0.2 to 1.0 g/kg). For the height intervall of 1.3 to 3.5 km the mean rms difference is ~15 % (~0.6 g/kg), i.e., the variability within the LEANDREII data (caused by both natural heterogeneity and intrumental noise) is here smaller than the differences between the two instruments. In conclusion, up to 10 % larger than "natural" differences are found here between the DLR-DIAL and LEANDREII water vapor data. It remains to be investigated which fraction of the differences can be attributed to any uncorrected spectral impurities or systematic differences due to data processing algorithms.

II) DLR DIAL/LASE

With the data available so far, three profile-to-profile intercomparisons could be performed with DLR-DIAL and LASE data (May 24, June 9, June 14). The released LASE data have a resolution of 330 m vertical and 14 km horizontal. The DLR data have an effective vertical resolution of 195 m and were averaged horizontally between 4 and 60 s. For the first case (May 24, distances 6 – 13 km), the two profiles show clearly different structures due to the presence of a dryline and cannot be used to assess the performances of the instruments. In the second case (June 9), the two airplanes were on parallel tracks with opposite directions (~ 1.5 km distance of the parallel tracks). Here the DLR-DIAL data are nearly constantly 20 % drier between 1.5 and 3.5 km a.s.l.., larger deviations are found close to the ground. The Falcon in-situ measurement agrees very well with the LASE data in 4.85 km height a.s.l. (relative difference 2.3 %). Possible reasons for the differences between the lidar data were discussed at the meeting but no explanation could be found agreed upon. In the third case (June 14, ~ 1.5 – 10 km distance) the DLR-DIAL data and the readjusted LASE data (see below) agree well in the free troposphere (< 5 % rms for 2.0 to 3.0 km). The larger differences below (up to 30 % {2.0 g/kg} at the boundary layer top) can be explained by the natural variability of the atmosphere.

III) DLR DIAL/SRL

At present, 9 comparisons between DLR-DIAL and SRL can be made (minimum distances 0.3 – 18.5 km, SRL: effective temporal resolution of ~ 2 min, DLR-DIAL: 5-s average which is equivalent to a horizontal resolution of ~ 710 m). The mean rms difference between 1.3 and 3.8 km a.s.l. is (0.74 ± 0.33) g/kg or (11.1 ± 3.4) %. This agreement is good taking the different measurement locations and periods into account and that neither the cross-section employed by the DLR-DIAL nor an exact calibration of the SRL during IHOP2002 have finally been determined. There is no clear dependency of the RMS deviation on spatial distance of the measurement locations: Small and large deviations are found for both large and small distances. The largest differences occur within the boundary layer and at the strong water vapor gradient at the boundary layer top which displays the temporal and spatial differences of the acquired data. Looking at the relative differences in the upper part of the intercomparable height region there is an indication (at least for some intercomparison cases) for a systematic ~10% bias between both instruments (DLR-DIAL data drier than SRL data).

c) Paolo Di Girolamo from UNIBAS gave two presentations concerning comparisons. The first one entitle "Airborne/ground-based sensor intercomparison: SRL/LASE" and "SRL/Reference Sonde".

I) In his first presentation, comparisons of SRL with LASE were presented and discussed. SRL operated for approximately 35 days during IHOP. Most of the measurements were carried out in vertically pointing mode. Present availability of analyzed data covers 11 days (29 May 2002, 30 May 2002, 31 May 2002, 3 June 2002, 4 June 2002, 9 June 2002, 10 June 2002, 12 June 2002, 14June 2002, 19June 2002, 20 June 2002), with both daytime and night-time cases. Six dedicated DC-8 flights were made in the IHOP 2002 region and LASE collected profile data over more than 40 flight hours. Comparison between SRL and LASE data are possible on 30 May, 3 June, 9 June and 14 June, with 24 comparisons available. Comparisons are based on 10 minute data averaging for SRL and 1 minute data averaging for LASE. Based on 10 minute data averaging, 10 % random uncertainty for SRL is reached between 2.8-4.3 km, while LASE measurement accuracy for 1 minute integration time is better than 6 % across the troposphere. In the comparison discussed at the workshop, data of flight 11 of LASE after the second data release were included. However, later on it was realized that in this data release, flight 11 data of LASE were calibrated using SRL. Also the data of flight 10 were calibrated using dropsondes. Therefore, if the second data release of LASE is used, flight 10 and 11 should not be included. In the meantime, updated comparisons are available which are discussed below, too.

Agreement between SRL and LASE was very good, with larger deviations between the two sensors revealed at the top of the boundary layer, where effects associated with inhomogeneities in the humidity field are larger. In order to come to a quantitative estimate of the agreement between the two sensors, total root mean square (RMS) deviations and BIAS in the altitude region 1.3-2.8 km and 1.3-3.8 km, together with contributions from different height intervals (1.3-1.8 km, 1.8-2.3 km, 2.3-2.8 km, 2.8-3.3 km and 3.3-3.8 km) were estimated. The average RMS deviation between SRL and LASE results did not exceed 8.4 % in the altitude region 1.3-2.8 km, while the average RMS deviation between each of the two sensors and the mean of the two does not exceed 4.2 % in the same altitude region. The average BIAS between SRL and LASE results do not exceed 1.2 % between 1.3 and 3.3 km, while the average BIAS between each of the two sensors and the mean of the two does not exceed 0.6 % in the same altitude region.

Comparisons have been updated ignoring 14 June 2002. In this case, the average RMS deviation between SRL and LASE results do not exceed 8.1 % in the altitude region 1.3-2.8 km, while the average RMS deviation between each of the two sensors and the mean of the two does not exceed 4.1 % in the same altitude region. The average BIAS between SRL and LASE results do not exceed 0.6 % between 1.3 and 3.3 km, while the average BIAS between each of the two sensors and the mean of the two does not exceed 0.3 % in the same altitude region.

It should be pointed out that when distances between LASE footprint and SRL are larger than 2.5 km, differences in air mass properties are appreciable and deviations between LASE and SRL are larger. Considering only those comparisons between SRL and LASE when distance between LASE footprint and SRL does not exceed 2.5 km (10 cases), even better results in terms of RMS deviation between the two sensors are reached (5.8 % in the altitude region 1.3-2.8 km, average RMS deviation between each of the two sensors and the mean of the two not exceeding 2.9 % in the same altitude region).

II) In the second presentation, comparisons of SRL with the reference radiosonde were presented. Nine reference radiosondes were launched from the Homestead Profiling Site during IHOP. Eight of these were launched at times when SRL was operational. Only three reference radiosondes were launched during the 11 days for which SRL data have been released (28 May 2002, Z20020528_162645; 9 June 2002, Z20020609_104618; and 20 June 2002, Z20020620_025610). Some data points are missing in the reference sonde profiles: specifically, sonde Z20020528_162645 has no humidity information between 2200 and 2700 m, while sonde Z20020609_104618 has large gaps of humidity information (2.9-3.6 km; 3.8-5.5 km a.s.l.) and pressure information is missing. However, pressure information from the reference sonde was not used for all three cases since it was considered to be not reliable. Pressure data from co-located or simultaneous Vaisala RS80 was used instead.

SRL provided direct measurements of mixing ratio, while reference sonde and Vaisala RS80 provided measurements of relative humidity. Hence, sonde data were converted into mixing ratios based on WMO formulation for saturation vapor pressure.

Agreement between SRL and reference sonde was good for 28 May 2002, and worse for 9 June and 20 June 2002. On 28 May 2002 differences between the two sensors did not exceed

10 % up to 3.5 km a.s.l., while on 9 June 2002 SRL is moister than the reference sonde up to 4.8 km (differences ~ -20 %) and on 20 June 2002 SRL is drier than the reference sonde below 2.3 km, moister between 2.3 and 3 km, and drier above (differences ~ -30 to +20 %).
However, it is to be mentioned that reference sonde data on IHOP server are indicated with the time of start of preparation of sonde instead of the time of sonde release. A misinterpretation of this issue during the data analysis led on 9 and 20 June 2002 to an effective separation of 30-60 minutes between sonde launch time and lidar start time, so that for these two cases non simultaneous measurements were compared. This time lag was very likely responsible for the large deviation between SRL and reference sonde on 9 and 20 June 2002. This was confirmed to be the case in a later analysis. It was suggested at the workshop that the launch times of the reference sonde stored on the IHOP server should be corrected.
In fact, comparisons between SRL and reference sonde were updated after the workshop in order to properly match the integration time of the two sensors. An average BIAS between SRL and reference sonde of 3.5 %, 9.5 % and -1.6 % was then found in the altitude region 1.3-3.8 km, respectively, on 28 May, 9 and 20 June 2002, with SRL resulting on average to be 4 % moister than the reference sonde. This 4 % wet BIAS is in agreement with the 4 % deviation mentioned above (Dave Whiteman) between SRL and simultaneous GPS PW, and may suggest the use of a wrong calibration constant for SRL during IHOP.

d) In his talk "GPS comparisons", John Braun from COSMIC/UCAR presented comparisons between GPS PW and MWR PW. He selected GPS stations which were set up for IHOP. In all cases, deviations are less than 1.5 mm confirming the expected accuracy. The GPS measurements appear to be consistently drier with respect to MWR PW measurements.
The audience suggested further comparisons between GPS and radiosonde as well as lidar data. These comparison will be particularly useful to specify also the accuracy of GPS SW and tomographic measurements.

e) Junhong Wang from NCAR/ATD presented detailed "Intercomparisons of Water Vapor Measurements during IHOP_2002 - Radiosonde and Dropsonde". After a discussion of error sources and types, she presented several comparisons of SW, RS80 and hygristor sensors. Obviously, both the RS80 and the hygristor show a severe dry bias in the upper troposphere.
Further data have been compared using radiosondes which have been launched from 1996-2002 from the Norman, OK, and the ARM B6 site. The VIZ sonde seems to be drier by about

15 % in RH in comparison with the RS80 sonde. The RS80 seems to be slightly drier (4%) than the RS90 in the upper troposphere. Also temperature measurement performance between different radiosondes was discussed.

f) In her presentation “Refractivity during IHOP_2002”, Tammy Weckwerth from NCAR/ATD discussed several comparisons of Radar refractivity measurements. Many comparisons with surface and airborne in-situ sensors, AERIBAGO as well as SRL have been performed. All results demonstrated excellent correlation between different measurement techniques. The comparisons achieved with in-situ sensors indicate a dry bias of refractivity. However, it was suggested that this may also be due to a different vertical ranges covered by these measurement techniques. The refractivity seems to be representative for a range from ground up to 300 m. The vertical weighting function which will depend on range and azimuth angle is under investigation. The range resolution is of the order of 4 km.

g) Tammy Weckwerth, on behalf of Cyrille Flamant (CNRS), also gave a presentation on “Comparison of sideways LEANDREII measurements with sounding, airborne in situ and DLR DIAL measurements”. Two cases from June 14 and June 19, 2003, were used where LEANDREII was operated in a horizontally pointing mode. During these days, comparisons with the UKWA and the P-3 in-situ sensors as well as the DLR DIAL could be performed. In general, the results revealed similar structures and trends in the water vapor field.

In summary, a large set of intercomparisons has been collected after one year of the field campaign. Though the data base was limited, a sufficient number of comparisons between active remote sensing systems could be performed to investigate their performance. Data gaps exist in LEANDREII and DLR DIAL data which was mainly due to problems with the temperature control of their laser cavity of DIAL system. Two flights of LASE had to be recalibrated using SRL and dropsondes. Regardless of these difficulties, the DIAL systems confirmed their high accuracy and resolution, if the DIAL systems were running maintaining all required system specifications. The SRL performed very-well and showed excellent agreement with LASE.

Now, it seems to feasible to develop detailed lists of specifications of different water vapor measurement techniques. The large interest in these specifications was demonstrated by the lively discussions how to understand and to improve the performance analyses of the IHOP instrumentation.

III Discussion and results

Fourth session: Discussion

To prepare the discussion, Volker Wulfmeyer summarized the results of the previous day. He showed that several activities reaching the goals of the IHOP IC are ongoing. Mainly, extended analyses of active remote sensing systems (lidar, GPS, S-Pol refractivity) as well as in-situ sensors have been performed. Further comparisons of passive systems (AERI, NAST) and satellite data should be performed. He suggested the set up of an instrument performance analysis web page which should contain the following information:

1) Status of intercomparisons: Steve Williams of JOSS and other participants proposed to put intercomparisons tables on this web page in order to give an overview on intercomparisons activities. Also tables showing possible intercomparisons times should be made available. The latter can be based on the WALES Verification intercomparisons tables. This would avoid double work and allowed the suggestion of further activities.

2) Then specific web pages for each instrument could be developed. This would be very helpful for data users. These pages could contain the following information

a) Measurement methodology or corresponding references
b) Description of retrieval technique (e.g. flow diagram and/or references)
c) Expected errors, critical components, references to error analyses
d) Units of the data (e.g. height asl including height errors)
e) Approach used for error analyses, definition of errors, lessons learned during error analyses such as during intercomparisons
f) Most important error sources, critical components, suggestions for future operation of the instrument
g) All these activities are resulting in a instrument specification table
h) Suggestions for future efforts to close remaining gaps in the data analysis

To demonstrate the utility and feasibility of this activity, Wulfmeyer demonstrated how a corresponding DIAL web page could look showing the methodology, the retrieval technique, data processing flow diagrams and techniques for error analyses. For instance, systematic errors can be assessed by intercomparisons or by error propagation analyses. Noise errors can often be derived using the instrument itself but also here intercomparisons of noise errors were suggested for verifying the performance of the instruments. Finally, a water vapor lidar web page should contain the following specifications (if applicable):

a) Horizontal and time resolution including weighting function
b) Vertical resolution including weighting function
c) Systematic error profile including description of error analysis procedure
d) Noise error profile in dependence of vertical and time resolutions
e) Specific error analysis at boundaries
f) Overall accuracy covariance matrix, probably broken down with respect to specific missions
g) Ideas to assess non-representativeness of the data. Here it must be distinguished between representativeness along the flight track due to missing data (clouds, detector saturation, system problems) and representativeness with respect to different grid boxes of atmospheric models.
h) Suggestions for improved error analyses: Improve measurements of water vapor Raman cross section at 355 nm and water vapor absorption cross section at 930 nm.

Similar lists were set up for S-Pol refractivity, GPS, microwave radiometer and in-situ measurements. For refractivity, the main difference is that 2-d error maps must be derived. Error may be related to topography. Steve Koch of NOAA/FSL stated that noise error may show a threshold effect which should also be investigated in detail. A main remaining issue seems to be the definition of the vertical weighting function determining for what height above ground the data are representative. Further comparisons with radiosondes and lidar systems were suggested.

For GPS measurements, more comparisons with lidar data were suggested, particularly for investigation of the representativeness of ZPD (zenith path delay) measurements. In-situ sensors, AERIs and microwave profilers need the same specification list as vertically pointing lidar systems. Also here, errors at boundaries such as clouds and temperature inversions need special attention. If intercomparisons are performed, the atmospheric conditions affecting radiosondes measurements such as horizontal drift should be taken into account.

The discussion of these specifications highlighted the role of intercomparisons. These are mainly a tool to compare systematic errors. However, to specify a systematic error of a single instrument, several intercomparisons are needed including an assessment of single-instrument systematic errors. The latter can only be derived by applying the methodology of the measurement process including a detailed analytical error propagation.

Another important topic was the definition and comparison of rms error or mean bias. Mean bias was understood as a means to determine systematic errors which are independent of atmospheric conditions and can be averaged over different comparisons. However, bias occurring under special atmospheric conditions are averaged, can be eliminated and may not show up anymore in the average mean bias profile. Therefore these errors (e.g. larger errors in DIAL systems due to gradients in the particle backscatter coefficient) need special attention. Some indication of these deviation can be maintained by calculation of rms errors. They extend this error analysis and seem to provide a more detailed description of the different performance of two instruments. Therefore, it was suggested to calculate both mean bias and rms deviation, if intercomparisons are presented.

III Results and action items

It was agreed on the following actions items including a distribution of the responsibilities (in brackets) :

- A IHOP IC web page shall be set up (Williams, Wulfmeyer). Both will interact to design this page and corresponding specific pages for single instruments. This activity shall be completed in March 2004. This activity shall be started immediately.
- An IC web page with tables showing possible intercomparisons and intercomparisons performed shall be set up (Williams in collaboration with workshop participants).
- Remaining data processing activities shall be completed and the data shall be put in its final version on the JOSS server (all instrument PIs).
- Instrument representatives shall put flow charts of the data processing procedures on the web page (each instrument PI).
- Error shall be defined clearly and the procedures for the determination of errors shall be described (after interaction of instrument PIs).
- In this connection, agreement of the use of mean bias or rms errors for the characterization of systematic errors shall be achieved (Behrendt, Browell, Di Girolamo, Whiteman).
- Instrument specification tables shall be filled out (all PIs).

- Suggest further intercomparisons activities (refine lidar comparisons, GPS, refractivity, others).
- Everybody is urged to develop strategies for assessing the non-representativeness of the data.
- A blind test shall be performed providing simulated DIAL signals to DIAL groups including all required information for data processing.
- The invitation list shall be extended to passive and satellite remote sensing representatives (NAST, AERI, Satellite, microwave profiler)
- The next meeting shall take place at NCAR/ATD in the week of April 5-9, 2004.

Acknowledgement:

The support of the intercomparisons activities by ESA grants 16993/03/NL/FF and 16669/02/NL/FF is highly appreciated.