science softCon

Data Validation with Examples

K. E. Wolff (1), M. L. Richards (2), G. K. Hartmann (3), G. Kirchengast (4),
A. v. Engeln, J. Ramsauer, Ch. Jarchow, H. Heinrichs, W. Boogaerts

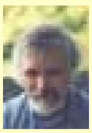
Formal Concept Analysis

Since Adam accepted an apple from Eve ...

...the conception of concept is in the mind of the humans:

Prof. Dr. Karl Erich Wolff
FH Darmstadt

Aristoteles introduced the hierarchy of concepts...



... and in 1982 Prof. Dr. Rudolf Wille (Technical University Darmstadt) introduced a mathematical theory for conceptual knowledge processing, called Formal Concept Analysis, based on formal concepts and their hierarchies. The following line diagrams of concept lattices are drawn with the program TOSCANA-CERNATO developed by NaviCon GmbH Frankfurt.

Introduction

Formal Concept Analysis has been applied to visualize ozone and water vapor data from the Earth's atmosphere as part of a research project at MPAE. Specific results for the ozone data obtained during the ATLAS 3 mission in 1994 are presented here.

Conceptual Knowledge Representation for Ozone Data

A data table with 4752 measurements of the ozone concentration at 216 locations around the Earth is evaluated with the purpose of understanding multidimensional dependencies between all measured variables. For a meaningful conceptual interpretation a suitable granularity has to be introduced. The chosen granularity for the values of the ozone concentration is represented by certain attributes, for example O3: >10, applied to all ozone measurements. That yields the concept lattice in Figure 1 (showing the cardinality of the extent of each concept):

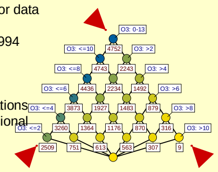
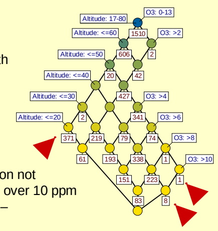


Fig.1: From low to high ozone concentration

Reading examples ( ) in the line diagram in Figure 1:
Among the 4752 ozone measurements there are 2509 measurements with an ozone concentration under 2 ppm and 9 measurements over 10 ppm.

Are you interested to know more about the dependencies of the ozone concentration with altitude, longitude, latitude ... ?
Here is the dependency with altitude (Figure 2):

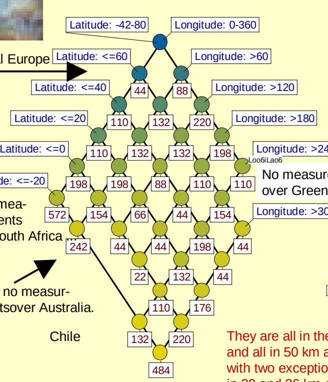


Up to 20 km altitude there are 371 measurements with ozone concentration not greater than 2 ppm. The 9 measurements having an ozone concentration over 10 ppm (see Figure 1) can be seen in Figure 2 again under the concept O3: > 10 – now splitted into two parts: 8 of them between 30 and 40 km altitude, one between 40 and 50 km.

Fig.2: Ozone concentration depending on the altitude

A conceptual map of the world (of these ozone measurements)

This conceptual map shows the number of measurements taken around the Earth (the latitude ranging from –42° to 80°).



Central Europe

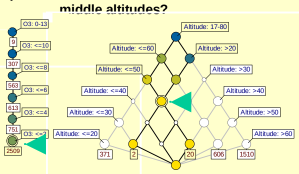
No measurements over Greenland.

Many measurements over South Africa

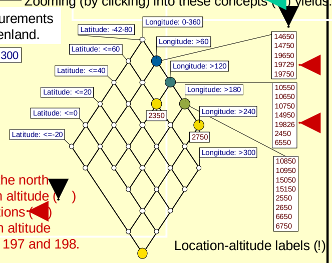
...and no measurements over Australia.

Chile

Low ozone concentration in middle altitudes?



Low ozone concentration (O3:<=2) occurs very often, but only 22 times between 20 and 50 km altitude. Where are their locations?
Zooming (by clicking) into these concepts ( ) yields:



They are all in the north and all in 50 km altitude ( ) with two exceptions ( ) in 29 and 26 km altitude at the locations 197 and 198.

Location-altitude labels (!)

Use of Neural network analysis in geophysical data validation,

(Richards and Heinrich)

This is a pilot study of the application of neural network methods to comparison of different remote sensing data sets. The data sets chosen were DMSP-SSM/T1 (temperature) radiances and radiosonde temperature profiles for the period March and April, 1993. The region was restricted to the 11 German stations of the European Upper Air Network. Radiosonde data were obtained from the Deutsche Wetter Dienst. SSM/T1 data tapes for satellites F10 and F11 came from the U.S. National Geophysical Data Center in Boulder CO.

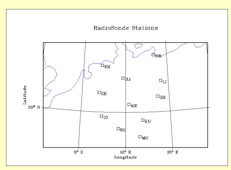
Data selection: Spatial colocation was obtained when a ground station lay within the field of view (14 degrees) of the down looking satellite sensor. However, the orientations of the circular, solar synchronous satellite orbits made difficult the temporal colocation for central Europe ground stations, with the radiosonde launch times of 0 and 12 h GMT. A time difference up to 3 h was allowed.

NN Method: Multi-layer perceptrons with one hidden layer were trained. The quickprop algorithm (eps=1e-8, and mu=2) was used for optimization since it leads to faster convergence [Fahlman 1988] than the standard back-propagation algorithm [Werbos 1974, Rumelhart et al. 1986]. The number of iterations was 300.

Network topology: There were 9 input nodes (radiances at the 7 frequencies: 51.5, 53.2, 54.9, 58.4, 58.8 and 59.4 GHz, and field of view latitude and longitude), about 50 nodes in the hidden layer, and 9 output nodes (radiosonde temperatures at the 9 pressure heights: surface, 600, 300, 200, 100, 60, 30, 20 and 10 hPa).

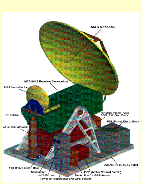
Training and testing discrepancies: In the learning phase an rms discrepancy of temperature profile of 1.2K was attained. In the verification phase, the discrepancy was 2.1K. The pressure height dependence is given in the table.

References:
Fahlman SE (1988), An empirical study of learning speed in back-propagation networks. Tech Report CMU-CS-88-162. School of Computer Science, Carnegie Mellon University, Pittsburgh
Rumelhart D, Hinton GE, Williams RJ (1986), Learning representations by back-propagating errors. Nature 323:533-536
Werbos P (1974), Beyond regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. thesis. Harvard Univ., Cambridge, MA



Height Variation of Discrepancy

	Learning Phase	Testing Phase
Pressure [hPa]	rmd [K]	rmd[K]
Surface	1.47	3.03
600	1.42	2.23
300	1.09	1.82
200	1.44	2.56
100	0.83	1.27
60	0.79	1.19
30	1.03	1.63
20	1.46	1.62
10	1.14	2.32



MAS Follow-On on Express Pallet on International Space Station (ISS)

MAS/GRAS Sensor Combination and Data Analysis

3.1 The MAS Follow-on concept and the addition of GRAS

*G.K. Hartmann (3), M.L. Richards (2), Ch. Jarchow *, W. Boogaerts**
** Max-Planck-Institut für Aeronomie, Katlenburg-Lindau, Germany*

The Millimeter-wave Atmospheric Sounder (MAS) has been successfully flown as a core payload of the three NASA ATLAS Space shuttle missions. It has recorded microwave emission spectra of the Earth's atmosphere allowing to retrieve temperature, pressure, ozone, water vapor, and the anthropogenically produced chlorine monoxide which mainly causes the depletion of the stratospheric ozone layer. A realistic simulation study at the University of Graz investigating joint retrieval based on both MAS data and GRAS (GPS/GLONASS Receiver for Atmospheric Sounding) radio occultation data showed that this allows to achieve very favorable accuracy of temperature profiles of the Earth's atmosphere. See Part 3.2 for results.

Thus a combination of a MAS Follow-on experiment – i.e., a modified MAS with second generation radiometers and electronics – together with a GRAS receiver on the EXPRESS Pallet of the International Space Station (ISS) is highly recommended. This would allow to obtain not only more accurate temperature and water vapor profiles but also provide simultaneously liquid water data below 17 km. All these quantities are very important for climatological research.

Measured quantities by combining MAS Follow-on (M) with GRAS (G):

Quantity	MAS frequency	Height Range [km]
Temperature	61 GHz	0–50 (G) + 15–90 (M) = 0 – 90
Pressure	62 GHz	0–50 (G) + 15–90 (M) = 0 – 90
H_2O	180 GHz	< 7 km (G) + > 7 km (M)
H_2O	183 GHz	17 – 95 (M)
O_3	184 GHz	17 – 95 (M)
ClO	204 GHz	17 – 45 (M)

Combined Optimal Estimation Retrieval for the MAS/GRAS Sensor

3.2 MAS/GRAS combined retrieval results

G. Kirchengast (4), A. v. Engeln †, J. Ramsauer ‡, G.K. Hartmann (3)
† *Institute for Environmental Physics, University of Bremen, Germany*
‡ *Inst. for Geophysics, Astrophysics, and Meteorology, University of Graz, Austria*

We investigated, with the aid of the optimal estimation method (Rodgers, 2000), the retrieval accuracy of temperature profiles from MAS-only and GRAS-only data analysis and, of particular interest, from combined MAS/GRAS data analysis. The retrieval accuracy of water vapor profiles from GRAS data analysis was studied as well.

MAS and GRAS sensor data were simulated based on a realistic sounding geometry assuming the sensors mounted on an ISS EXPRESS Pallet and probing the same air volumes toward the backward limb. A representative sample of 30 globally distributed GRAS occultation events and co-located MAS limb sounding scans were "forward-modeled" using enhanced versions of the MSIS-90/CIRA-86 models (Fig. 1). Retrievals were then performed on a case-by-case basis and retrieval error statistics were produced.

For temperature we found a modern MAS/GRAS sensor to achieve < 1 K accuracy below 35 km at ~ 1 km resolution and < 4 K accuracy throughout the mesosphere at ~ 5 km resolution (Fig. 2). An individual case is illustrated in Figure 3, where the right panel shows the MAS/GRAS "averaging kernels" indicating the high synergy in information contribution by the two sensors. Water vapor can be retrieved from GRAS data to < 10–20% accuracy below 5 – 8 km (Fig. 4; assumed a priori water vapor uncertainty: 25%).

It seems highly worthwhile to place a MAS/GRAS sensor on the ISS. In fact it were a role model of synergistic use of sensors in the best sense.

Detailed Information:
v.Engeln, A., J. Langen, T. Wehr, S. Bühler, and K. Künzi, Retrieval of upper stratospheric and mesospheric temperature profiles from MAS data, *J. Geophys. Res.* **103,** 31,735–31,748, 1998.
Kirchengast, G., and J. Ramsauer, Accuracy of temperature and water vapor profiles from GRAS and from combined MAS and GRAS profiles, *Report MPAE-MASGRAS No. 1/1998*, 20pp., Inst. for Meteorol. and Geophys., Univ. of Graz, Austria, 1998.
v.Engeln, A., G. Kirchengast, and J. Ramsauer, Accuracy of temperature and water vapor profiles derived from MAS and GRAS data by optimal estimation, *Report MPAE-MASGRAS No. 1/1999*, 37pp., Inst. for Meteorol. and Geophys., Univ. of Graz, Austria, 1999.
v.Engeln, A., S. Bühler, G. Kirchengast, and K. Künzi, Temperature profile retrieval from surface to mesopause by combining GNSS radio occultation and passive MW limb sounder data, *Proc. 10th Conf. on Sat. Met. & Ocean.*, 240-243, AMS'2000, Long Beach, USA, 2000.
Rodgers, C.D., *Inverse Methods for Atmospheres: Theory and Practice,* World Scientific Publ., Singapore, 2000.

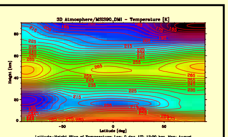


Fig. 1: Example cross section through MSIS-90 model

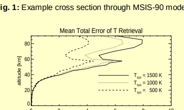


Fig. 2: Statistical T error for MAS/GRAS sensor

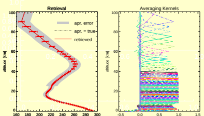


Fig. 3: T retrieval example – Profiles and Av. Kernels

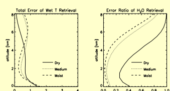


Fig. 4: Moist air retrieval accuracy – 3 scenarios

Contact:

Prof. Dr. Karl Erich Wolff, (Part 1) (1)
FH Darmstadt, FB Mathematik und Naturwissenschaften
Schöfferstr. 3
D-64295 Darmstadt
Web: www.mathematik.tu-darmstadt.de

Dr. Michael L. Richards, (Part 2) (2)
Copernicus Gesellschaft e.V.
Max-Planck-Straße 13
D-37191 Katlenburg-Lindau
www.copernicus.org

Prof. Dr. Gerd K. Hartmann, (Part 3, PI) (3)
Max-Planck-Institut für Aeronomie
Max-Planck-Straße 2
D-37191 Katlenburg-Lindau
www.linmpi.mpg.de

Prof. Dr. Gottfried Kirchengast, (Part 3) (4)
IGAM – University of Graz
Universitätsplatz 5
A-8010 Graz
www.kfunigraz.ac.at/igamwww