Publishing Data

Featuring
“Earth System Science Data”
– A Data Publishing Journal

Hans Pfeiffenberger, David Carlson, Sünje Dallmeier-Tiessen,

Alfred-Wegener-Institute for Polar and Marine Research, Helmholtz Association - Germany
British Antarctic Survey, Great Britain

Contact: suenje.dallmeier-tiessen@awi.de

Academic Publishing in Europe 2009, Berlin, Germany

Agenda

- Why publish data ... and: What is the problem?
  - Developments in the arena of science politics/policy
  - State of the art and missing elements

- ESSD - “Earth System Science Data”, a journal
  - A practical contribution to an emerging genre of scholarly communication
  - Aims and scope; structure of articles, review criteria

- Conclusion and Outlook:
  - Specific: On ESSD
  - General: Contribution of “classical” academic publishing to data publishing
Interestingly, there is no mention of a world class publishing industry ....

Or is this industry a research infrastructure ? !!

We will show how publishing can help comply with the requirement for quality assured research data.

The venerable Royal Society

... organized a workshop in April 2008 bearing the title:

**INTRODUCTION**

The environmental eScience revolution

... recognizing the changes brought about through computing, modelling and analysis of massive amounts of observed data.
What is the problem? Consider Ozone data sources:


 ESA / other gov. agencies as stewards
 Elaborate technical infrastructure (Grid)

- ESA(IT) - KNMI(NL)
  Processing of raw GOME data to ozone profiles
  2 alternative algorithms
  ~28000 profiles/day

- IPSL(FR)
  Validate some of the GOME ozone profiles (~10⁴/y)
  Coincident in space and time with Ground-Based measurements

- Visualization & Analyze

- LIDAR data (7 stations, 2.5 MB per month)


**Well known instruments**

**Well documented procedures**

**Well defined products**

**Historical ozone profiles from Antarctica**


- Ozone soundings (balloon-carried sonde profiles) in the years when the “ozone hole” first developed

- Satellite data provided “total column” values only

- => balloon data needed for **calibration** of satellite data and **verification** of models
Handling of Ozone data as State of the Art

- These two “datasets” exemplify the two prevailing modes of handling data at present:

- Either at the “Petascale”, where largely homogeneous mounds of data are handled in an industrial fashion, and collated into one super-dataset, comparable to a book holding the work of a lifetime.

- Or at the “Megascale”, where large numbers of heterogeneous datasets are handled as in a factory (manufaktur), by a craftsperson or an artisan. They are communicated on demand through mail or via obscure ftp-server, comparable to the letter from scholar to scholar.

- There is almost no in-between, yet, to handle the bulk of information at the Giga- to Terascale, which needed to be comparable to the system of academic journals for textual information.

Who is who…

Advisory Board:
Paul J. Crutzen
Sydney Levitus
Alexander Petrovich Lisitzin

Editors in Chief:
David Carlson
Hans Pfeifferberger

Publishing House
Copernicus Publishers – OA Publisher, EGU

Managing Editor
Suenje Dallmeier-Tiessen
Compilation of ozone profiles from the Antarctic Georg-Forster-Station from 1985 to 1992

G. König-Langlo and H. Gernandt

Alfred Wegener Institute for Polar and Marine Research, Bussassestraße 24, 27570 Bremerhaven, Germany

Received: 29 July 2008 – Accepted: 5 September 2008 – Published: 22 September 2008

Correspondence to: G. König-Langlo (gerit.koenig-langlo@awi.de)

Published by Copernicus Publications.

Abstract

On 22 May 1985 the first balloon-borne ozone profile was successfully launched by the staff of Georg-Forster-Station (70°46' S, 11°41' E). The following weekly ozone soundings mark the beginning of the continuous investigation of Germany to study the vertical ozone distribution in the southern hemisphere. In 1985 these ozone soundings have been the only record showing the change of vertical ozone distribution in the southern polar stratosphere in September and October. The regular ozone soundings from 1985 until 1992 are a valuable reference data set since the chemical ozone loss became a significant feature in the southern polar stratosphere.

The balloon-borne soundings were performed at the upper air sounding facility of the neighbouring station Novolazarevskaya, just 2 km apart from Georg-Forster-Station. Till 1992, ozone soundings were taken without interruption. Afterwards, the ozone sounding program was moved to Neumayer-Station (78°39' S, 8°15' W) 750 km further west.

Data coverage and parameter measured

Repository Reference: doi:10.1594/PANGAEA.547983
Available at: http://dx.doi.org/10.1594/PANGAEA.547983

Coverage: East: 11.8300; South: -70.7700
Location Name: Georg-Forster-Station, Antarctica
Date/Time Start: 1985-05-22T06:19:00
Date/Time End: 1992-01-29T01:19:00
Estimate of Error and Data Provenance

For balloon-borne ozone profile measurements a pump correction has to be applied in order to compensate the decreasing pump efficiency with increasing height and changing air temperature. Both, an inadequate pump correction and an erroneous estimate of residual ozone above the height of balloon burst may contribute to the overall measurement error of the ozone profile. Usually an independent column ozone observation $X_D$ by spectrometer measurement is compared with the integrated column ozone $X_S$ between the ground level and the height of balloon burst plus estimated residual ozone above that level to adjust the recorded profile values. The correction factor is

$$C = \frac{X_D}{X_S}.$$ 

Systematic differences and random errors of the electrochemical ozone sonde, type OSR, has been estimated by analysing 20 tandem ozone soundings at the Aerological Observatory Lindenberg in 1982 (Feister et al., 1985). Random errors are at their maximum of about 10 to 13% in the troposphere and above 32 km, and reach a minimum of 2 to 5% between 20 and 28 km. The mean random error is 11.5% in the troposphere, 7% in the stratosphere beneath the ozone maximum height (ca. 22 km), and 5.6% above that height.

Data Provenance and Structure

The first permanently operated German research base – later named Georg-Forster-Station – was established in 1976 in the Schirmacher Oasis at 70° 46' S, 11° 41' E. Since then the station was permanently used and operated as an annex to the Russian station Novolazarevskaya until 1987, and then as a German Antarctic station named after

Review Guidelines

- **Originality:**
  Are the data or methods new - i.e., never measured or employed before

- **Significance:**
  Is there any potential of the data being useful?

  **Uniqueness**
  **Usefulness**
  **Completeness**

- **Data Quality**
  The data must be presented readily available in a usable format.
  Accuracy, methods, instrumentation and processing as state of the art
Today’s Data Reuse, Citation and Quality Control

Reuse, Citation and Quality Assessment with ESSD

Submission:
1. Data to Data Repository
2. Data-Paper to ESSD
3. Traditional Paper
Summary - Outlook (Specific: ESSD)

- **Reward** for data publication, citable (impact factor)
- **Quality assured** data and data documentation to facilitate future reuse
- First article online – first experiences

**Outlook**

- Special Issue with 18 papers to be published soon
- Development of more specialized manuscript templates and review guidelines for other types of research data

---

Summary - Outlook (General)

**Preservation and (Open) Access to Data**

**Aim: Reuse & Reproduce**

- **Quality Assessment**
- **Persistent (and Open) Access**
- **Digital Longterm Preservation**

Data publishing provides **Data provided and described by researchers**

**Basic and advanced data infrastructure, provided by ???**
Summary - Outlook (General)

- Text has been with us for 5,000 years
- The printing press, 500 years
- Digital data, as preserved items, 50 years (World Data Centres)
- Online access to massive amounts of data, 5 years

=>

- Do not expect perfect, final “layout” publishing of data anytime soon
- However, let us identify and take steps which bring us forward

Thank you!