Reprocessing of 14 years of GTS data at CPTEC/INPE for Distribution

Contact:
waldenio.almeida@cptec.inpe.br
AMS 2010

Presentation Roadmap:

1) Some Informations about the CPTEC/INPE
2) The GTS data-processing systems
3) Improvements on the systems
4) Reprocessing tape-archived data
5) Comparisons between the systems
6) Preparing products for distribution
7) Conclusions and remarks

A bit about CPTEC/INPE...

- Center of Weather Forecast and Climate Analysis, under the National Institute of Space Research

- Supercomputing began in 1994 with a NEC SX-3. (after the signing of the non-proliferation of nuclear weapons treaty).

- An operational and research center for Numerical Weather Forecast.

- Distributes freely data, imagery and outputs from several models

Some marks of CPTEC history:

- Following a path of acquiring capability and know-how, models and systems became operational over the time:
  - ☐ 1994: NWP Global model (from COLA)
  - ☐ 1995: Global Data processing system (from ECMWF)
  - ☐ 1996: NWP ETA model for all South America
  - ☐ 1998: Oracle Database for GTS meteorologic data (from Meteofrance)
  - ☐ 2004: CATT-BRAMS Environmental model
  - ☐ 2006: Data assimilation schemes (PSAS from NASA)
  - ☐ 2008: One of 10 institutions to provide the WMO TIGGE Project with Global Ensemble NWP data (CPTEC EPS)
- => This presentation is focused on the data processing...

GTS Data Processing system...

- From the starting, GTS data were received from Brasilia RTH by FTP

- The Pre-processing system from ECMWF was adapted to decode the basic types of GTS data in CPTEC
  - Data: Synop,Ship,buoy,temp,pilot,airep,metar,satob,satem
  - Note: Because of modifications, it became a bit incompatible with ECMWF's original system.

- This system became operational mid-1995 to feed all CPTEC necessities:
  - Weather Forecasters and Climate analysis
  - Researchers and Numerical modeling
  - Webpages and products

Following a "mobile target"...

- Meteorological data are always changing:
  - Bigger data volumes and quantity
  - Changes on coding standards
  - New techniques, computers and requirements
  - New Data types, and so on...
- Data-processing systems needs constant improvements and maintenance... But this could not be easy:
  - Requires dedicated permanent people
  - Eventually, key people moves out.
  - New tasks and projects drain resources and attention...
    - ===> Very important factor in a new developing institution !!

In need of an improved system

- In 2003, institutional GTS data requirements were not being met by the processing system.
    - Need of more observations
    - New data types
    - Need of faster results and more operational stability
    - Lower internal losses of data

- Identified reasons:
    - Some needed data were not available at RTH Brasilia
    - Need of more decoders for new data types..
    - Old machines and techniques were slow and unreliable
    - Not enough human resources to maintain legacy codes.

Action Strategy

- In 2003 the responsibility of the data-processing system moved to Operational Division. A new team was formed.

- Other tools and solutions were considered and evaluated

- Starting of more international collaborations: (UNIDATA/UCAR was an important one)

- Revitalization of existent collaborations (Brazilian Met. Service, Navy, universities, ECMWF,...)

- Starting of an active participation in WMO commissions and expert teams.

- New ideas and practices:
    - Preference for open-source software and LINUX.
    - Investment in a collaborating informal community.
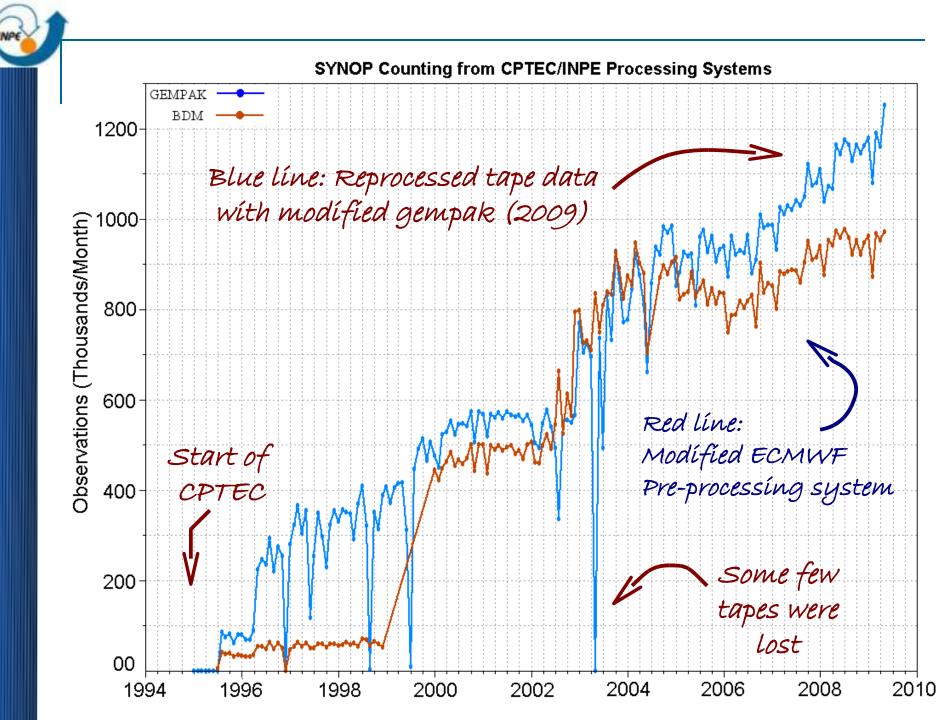   - Distribution of tools and support to universities

An Improved Data-processing

- 2004 to 2009: New servers, new tools, new concepts
- An improved data processing system was set-up:
  - All script-based data-transfer schemes were replaced by unidata's LDM and DWD's AFD.
  - The original CPTEC/ECMWF pre-processing system was migrated to new machines, feeding the oracle database.
  - A second data-processing line based on GEMPAK was implemented, creating a database in gempak-format .
  - Other tools and systems were locally developed for operational verifications and raw data organization.
- Some of these solutions were also implemented in other Brazilian, Portuguese and African organizations - a collaborative work with CPTEC.

Reprocessing stored GTS data

- Some modifications were made to produce more data:
    - Small modifications on some gempak decoders and graphical interfaces
    - Correction and expansion of station tables
- Finally a new system is on place with superior performance !
- *Now it is time to reprocess all the GTS data stored on the tapes !*
    - A database for raw GTS data was developed to help with the tape data organization
    - Comparisons were made with the data from the previous system, archived in the oracle database.
- *As a example, let's see the comparisons for SYNOP ..*

SYNOP Counting from CPTEC/INPE Processing Systems

Blue line: Reprocessed tape data with modified gempak (2009)

Start of CPTEC

Red line: Modified ECMWF Pre-processing system

Some few tapes were lost

Preparing a product for distribution

- The comparisons for the others data types shows a pattern similar.

- Comparisons with ECMWF's global data availability are underway, with good preliminar results, in particular for the South-America region.

- The 14 years of reprocessed data (1995-2009) in gempak format are being written in DVDs. Shall be ready for distribution in 2010.

- Initial release: SYNOP, METAR and TEMP/PILOT data

- Following: BUOY, SHIP and AIREP data.

- A LINUX self-installation distribution bundled with a gempak version ready to read the data DVDs also will be available.

Next steps under planning ...

- Old archived data could be acquired from international centers to fill and complete the raw database

- New reprocessing tasks shall be done on a regular basis to generate improved and more complete datasets

- Better quality-control procedures can be implemented..

- Similar products could be generated with datatypes stored in archive tapes, but not included in this work.

- The improvement work on the data-processing systems need to continue in order to meet new requirements (like the WMO's BUFR migration project and new datatypes).

Conclusions and Remarks

- The CPTEC is being build from scratch since 1994 to be a center of numerical weather modeling.

- Since its beginning competence and capability is being developed and accumulated.. in systems and peoples !

- These achievements came through a strong policy of international cooperation and human resource training.

- To allow the improvement of data assimilation systems for NWP, good data-processing systems are needed.

- A new level of competence at this field was reached with the improvements from the last 5 years.

- International cooperation and free data policy is a prized local value !

Thank You !

....
Questions ?

Contact:

waldenio.almeida@cptec.inpe.br

CPTEC/INPE - Brazil