SCANTEC: A Community System for Evaluation of Numerical Weather and Climate Prediction Models

João Gerd Zell de Mattos*, Luiz Fernando Sapucci, Carlos Frederico Bastarz, Ariane Frassoni, Wanderson Santos, Arletis Carrasco

*Center for Weather Forecasting and Climate Studies, Brazilian National Institute for Space Research, Cachoeira Paulista, SP, Brazil

1. Introduction

The Brazilian meteorological community has few model verification tools following the World Meteorological Organization (WMO) recommendations. The National Institute for Space Research (INPE), Center for Weather Forecasting and Climate Studies (CPTEC) operational Numerical Weather Prediction (NWP) center is providing operationally the Community System for Evaluation of Numerical Weather and Climate Prediction Models – SCANTEC (from the acronym in Portuguese to ***S**istema **C**omunitário de **A**valiação de modelos **N**uméricos de **T**empo **E** **C**lima* – de Mattos and Sapucci 2017) in order to contribute with the improvement of model quality assessment. SCANTEC project aims to offer for the community and operations a unified, standardized, and flexible tool for forecast verification.

The SCANTEC project is under management in a flexible institutional project management web application at CPTEC. SCANTEC Version 1.0 includes traditional measures for categorical and continuous variables, like the Root Mean Square Error (RMSE), bias, and Anomaly Correlation (AC). Besides, the package provides advanced spatial forecast evaluation techniques in research mode, like the Method for Object-based Diagnostic Evaluation (MODE). SCANTEC advantages span in the flexible integration of modeling systems employed by different institutions in Brazil, or different versions of the same modeling system. SCANTEC is also flexible to receive new statistical metrics (de Mattos and Sapucci 2017).

This paper aims to describe the main components of SCANTEC and its potential as an open-source, community-based development software for model verification. Section 2 presents the basic structure of SCANTEC Version 1.0, in which the main features are described. Section 3 presents the statistical metrics available, and finally, Section 5 describes the planning of future developments, including the capability for broad use in supercomputing environments and future applications.

2. Basic structure of SCANTEC

SCANTEC is a system based on open-source tools and can be run on different operating systems, such as UNIX, Linux, Windows, and macOS. The system structure includes a kernel fully modular to facilitate the implementation of new features. The kernel is developed in Fortran 2003 programming language, following the ANSI standard. Besides, SCANTEC is userfriendly and configurable through American Standard Code for Information Interchange (ASCII) files. The main components of SCANTEC are illustrated in Figure 1.

SCANTEC was designed by software development practices that encourage the reuse and community sharing of algorithms among the scientific community. The components were designed as functional abstractions using flexible object-oriented programming paradigms to facilitate reuse and the development of future implementations. Interoperable features in SCANTEC also include reuse and joint development with other numerical modeling groups. Similar to the nature of object-based in structured programming, SCANTEC provides standard functionalities for model evaluation and allowing the user to fill variable functionalities according to their needs. The number of variable functionalities in SCANTEC includes interfaces to facilitate the incorporation of (1) domains, (2) numerical models interfaces, (3) types of observations, and (4) statistical methods for evaluation. A set of abstract functions are incorporated to represent the variable functionalities. These interfaces called "plugins" (model-plugins, obs-plugins, stat-plugins), contain access points or extensible interfaces to incorporate routines to read new models and statistical metrics not included in the stat-plugins component. The model-plugin component is responsible for model data reading; obs-plugin is responsible for the reference dataset reformatting and access a dataset server, and stat-plugin performs the computation of statistical metrics; visualization and post-processing tools provide iterative access to SCANTEC products.

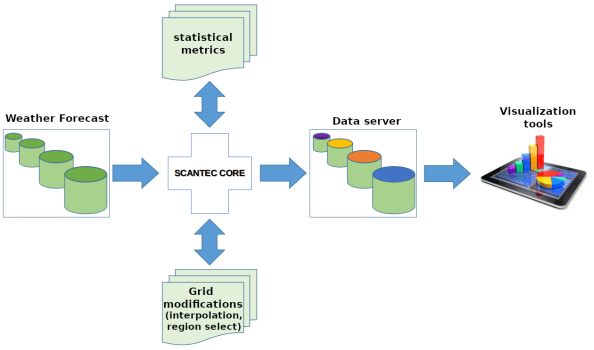

Fig. 1. The main components of SCANTEC.

3. Statistics and visualization

SCANTEC includes two modules that perform statistical computations: Method For Object-Based Diagnostic Evaluation (MODE) and basic-statistic. Both are components of stat-plugin. The basic-statistic module includes standard statistical metrics for comparing forecasts and grid point references. This functionality is suitable for comparing model outputs with its analysis, and also perform intercomparison between different models or experiments performed with the same model. In addition to MODE, SCANTEC also offers methods for dichotomous forecast evaluation, extracting information from the contingency table after considering, for the precipitation field, cer-

* *Corresponding author address:* João Gerd Zell de Mattos, Center for Weather Forecasting and Climate Studies, Rodovia Presidente Dutra Km 40, Cachoeira Paulista, SP, Brazil

E-mail: joao.gerd@inpe.br

tain thresholds commonly used by the community. The metrics available in Version 1.0 are: Anomaly Correlation (CC), Root Mean Square Error (RMSE) and Mean Error (ME), metrics for specific precipitation assessment such as frequency histogram and contingency table.

SCANTEC produces an output file in ASCII, which contains the average statistical results on the selected domain as a function of lead time or time of day. A file in a sequential binary format is produced, containing the statistical results for each grid point of the domain and period. SCANPLOT tool performs the visualization of statistics tables provided by SCANTEC. This tool consists of a set of scripts written in Python, where graphical outputs include visualization of statistics in the scorecard and Taylor Diagram format. Figure 2 shows an example of a scorecard provided by SCANTEC, highlighting the RMSE improvement in different variables and levels evaluated (Sapucci et al. 2016).

4. Spatial verification method

Traditional metrics are not sufficiently informative to evaluate numerical models, especially those with high horizontal resolution. MODE is an object-based method to verify properties of spatial forecasts of entities, where an entity is anything that can be defined by a closed contour (Ebert and McBride 2000). This technique emulates the visual identification of a forecaster analyzing the meteorological field, identifying matched objects, and then comparing each other (Davis et al. 2006, 2009). MODE has been implemented by (Carrasco 2017) and was applied to evaluate precipitation forecasts and intercompare the Brazilian developments on the Regional Atmospheric Modeling System (BRAMS) and the Weather Research and Forecasting (WRF) Model (Carrasco et al. 2020). Recent work at CPTEC has used MODE to identify and evaluate forecasts of heatwaves predicted by two versions of BRAMS model using ECMWF ERA5 reanalysis and GFS analysis as the reference database (Garcia 2020).

5. Future developments

SCANTEC is a project in progress, and future developments consider community needs and contributions. A graphical web interface based on the Python Jupiter notebook tool is under development, which should be migrated to a stand-alone graphical interface. NetCDF (network Common Data Form) format will be included as one of the data formats supported in SCANTEC. Observational datasets provided in PrepBuffer and ASCII, which include conventional data such as SYNOP, SHIP, METAR, among others, will also be supported and included in the list of references database available in obs-plugin. As many centers work in a high-performance computing environment, parallel processing is desired for SCANTEC to provide faster and more efficient statistical computation. To meet this aim, the parallelization of SCANTEC is required and will be available in future versions. Scientific applications restricted to meteorological variables are considered to be extended to air quality variables, as CPTEC is a producer of operational air quality forecasting for South America, and recognizes the need for an operational procedure for air quality forecasting verification.

6. Summary and Conclusions

SCANTEC has been developed at INPE/CPTEC for use by the internal community in NWP assessment. SCANTEC Version 2.0 is under release in the operational CPTEC NWP environment. The tool is applied to evaluate meteorological variables and offers a flexible environment for user needs under a userfriendly configuration to another modeling system other than those currently available. SCANTEC has been applied over the last year in scientific studies and is under the GNU General Public License. As a community tool, SCANTEC is open-source software that will be available to community contributions to enhance the tool and keep it relevant for scientific applications.

Acknowledgments. SCANTEC has been sponsored by The Brazilian National Council for Scientific and Technological Development (CNPq) and Foundation for Research Support of the State of São Paulo (FAPESP). We acknowledge Dr. Chou Sin Chan, head of Modeling and Development Branch at CPTEC, for her support and effort to release the operational version of SCANTEC for the INPE/CPTEC NWP community.

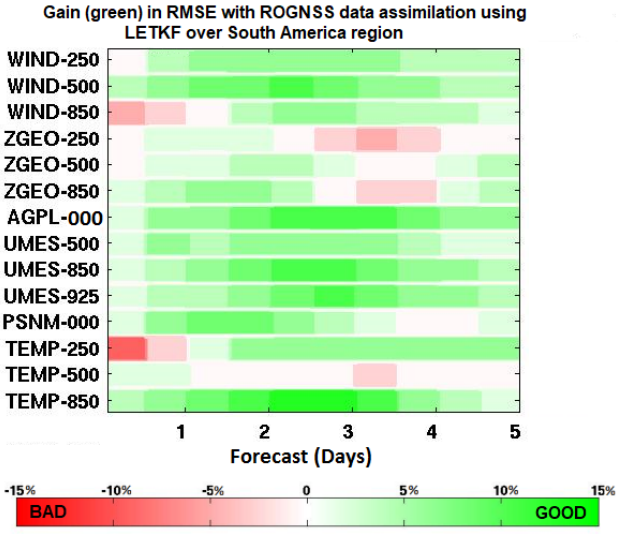

FIG. 2. Type of analysis provided by SCANTEC exploring scorecard shows the gain in RMSE over South America after the assimilation of radio occultation data. More detail about this study is available in Sapucci et al. (2016).

References

Carrasco, A. R., 2017: Método de avaliação orientada a objeto aplicado às previsões de precipitação sobre a américa do sul. M.S. thesis, Instituto Nacional de Pesquisas Espaciais (INPE), 116 pp., São José dos Campos, URL http://urlib.net/rep/8JMKD3MGP3W34P/3NH9KMB.

Carrasco, A. R. C., L. F. Sapucci, J. G. Z. de Mattos, M. S. Lorenzo, and I. B. Monteiro, 2020: Exploring the particularities of the method object-based in the precipitation forecast evaluation, In publication process. *Revista Brasileira de Meteorologia*.

Davis, C., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. part i: Methodology and application to mesoscale rain areas. *Monthly Weather Review*, **134 (7)**, 1772–1784.

Davis, C. A., B. G. Brown, R. Bullock, and J. Halley-Gotway, 2009: The method for object-based diagnostic evaluation (mode) applied to numerical forecasts from the 2005 nssl/spc spring program. *Weather and Forecasting*, **24 (5)**, 1252–1267.

de Mattos, J. G. Z., and L. F. Sapucci, 2017: BR 51 2017 000576-1. Scantec - sistema comunitÁrio de avaliaÇÃo de modelos numÉricos de tempo e clima. 13-junho-2017.

Ebert, E., and J. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *Journal of Hydrology*, **239 (1–4)**, 179–202.

Garcia, G. R., 2020: Object-based evaluation of the impact of burning aerosols on heat waves forecast in south america. Master dissertation: Meteorology, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, document in Portuguese.

Sapucci, L. F., F. L. R. Diniz, C. F. Bastarz, and L. A. Avanço, 2016: Inclusion of GNSS radio occultation data into CPTEC Local Ensemble Transform Kalman Filter (LETKF) using the ROPP as an observation operator. *Meteorological Applications*, **23 (2)**, 328–338, doi:10.1002/met.1559.