3-15-2016

SeFo: A Package for Generating Probabilistic Forecasts from NMME Predictive Ensembles

Nir Krakauer
CUNY City College

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/cc_pubs

Part of the Civil and Environmental Engineering Commons, and the Software Engineering Commons

Recommended Citation
https://academicworks.cuny.edu/cc_pubs/758

This Article is brought to you for free and open access by the City College of New York at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
Overview

Introduction

Long-range weather forecasts based on output from ensembles of computer simulations are attracting increasing interest. A variety of methods have been proposed to convert the ensemble outputs to calibrated probabilistic forecasts. The package presented here (SeFo, for Seasonal Forecasting) implements a number of methods for producing forecasts of monthly surface air temperature anomalies up to 9 months in advance using output from the North American Multi-Model Ensemble (NMME). The package contains modules for downloading and reading past observations and ensemble output; producing forecast probability distributions; and verifying and calibrating a user-determined subset of methods using arbitrary past periods. By changing individual modules, the package could be extended to use other model ensembles, forecast other weather variables, or apply other forecast methods. SeFo is written in the numerical computing language Octave and is available on Bitbucket under the GNU General Public License (Version 3 or later).

Implementation and architecture

SeFo is written as a package for the free numerical computing language and environment GNU Octave [21–22], exploiting built-in Octave functions and capabilities wherever possible. (Although Octave is deliberately designed to be able to run programs written for the proprietary environment MATLAB, no effort has been made to make SeFo usable in MATLAB.) The package has a modular architecture that isolates components of the workflow for probabilistic seasonal forecasts using model ensembles. Each module is an Octave function that writes its output to a file, from which it can be accessed by other modules. An input structure is used to hold options that need to be passed on to different modules, such as which data sets, dates, and forecast methods to use. (The list of possible options is given in a README text file.)

The core SeFo functions (modules) all have names starting with "sefo_", and their interrelationships are diagrammed in Figure 1. There are also a number of ancillary functions included in the package, for example for obtaining regression models used for particular forecast methods, regridding to the 1 degree grid used for the forecast ensemble output, and making maps. The capabilities of the core functions are as follows:

- sefo_obs_read: Download and regrids observational data (currently surface air temperature, either from Berkeley Earth...
Surface Temperature [23]; NCEP/NCAR reanalysis [24]; or Climate Prediction Center [25]) for a specified month.

deso_obs_assemble: Collect the observations for a sequence of months.

deso_fcst_read: Download and store ensemble forecasts from a given climate model and month (currently the data source is the North American Multi-Model Ensemble (NMME) Phase I [26] accessed via the IRI Data Library [27]).

deso_fcst_assemble: Collect ensemble forecasts for a sequence of months.

deso_predict: Apply one of several (currently 23) available prediction methods to estimate a probability distribution values for a given month from current ensemble predictions plus a set of past prediction-observation pairs. Currently the implemented forecast methods all return t distributions as the forecast probability distributions.

deso_adj: Apply an optional calibrating adjustment to the forecast t distribution to better match the distribution of verifying observations over some specified past period.

deso_cdf: Calculate, and optionally map, requested quantiles of a forecast probability distribution.

deso_verify: Compare probabilistic forecasts for a past period against observations using several metrics, including forecast root mean square error, bias, mean negative log likelihood, and Kolmogorov-Smirnov statistic.

deso_time_methods: Compare the computation times for selected forecast methods.

deso_example: Exercise the key components of the package by generating a sample forecast for next month (Figure 2).

Basic installation instructions are provided in the README file.

Quality control
Each of the core functions (all the functions with names beginning in deso, except deso_example) has a demonstration script that tests and illustrates its basic capabilities.

"demo function_name" will run this script in Octave. Some of the ancillary functions have their own unit tests defined ("test function_name"). Development and testing was carried out in a Linux environment, specifically the Debian distribution (versions Jessie (Stable) and Unstable), and in Mac OS X with a Macports Octave installation.

Limitations and potential improvements
Currently, the package routines are not fully generalized. For example, NMME is currently the only supported source of ensemble predictions.

Documentation for the package and unit tests and demos for non-core functions are also not complete.

More information could be provided while the software is running, such as percentage progress of the downloads and data analysis.

Better input checking for the options structure could be provided with analogues of the odeset and odeget functions used for supplying parameters to Octave’s differential equation solvers.

While the current data sources for SeFo, referenced above, are, to the author’s knowledge, available without restrictions on use, abilities to handle and display different data licenses could potentially be added.

Once the functionality has been extended to more use cases and the documentation is more complete, it is envisioned that SeFo might be added to the Octave Forge repository, from which it might be accessed by a wider user base.

Users are encouraged to submit bugs and patches to the repository issue tracker on Bitbucket.

(2) Availability

Operating system
While in theory the package should run in any operating system for which Octave is available, including Windows, it has only been tested in Unix-like environments (Linux and Mac OS X).

Programming language
The package requires GNU Octave (Version 3.8 or newer) with the linear-algebra [28], nan [29], netcdf [30], and splines [31] packages installed.

Additional system requirements
An Internet connection is required to download observational data and numerical weather prediction model ensemble output. Data and intermediate files are stored locally, which will typically require one to several gigabytes of space, depending on the use case.

Dependencies
There are no dependencies beyond those for Octave with the indicated packages.

Software location

Archive (e.g. institutional repository, general repository)

Name: figshare

Persistent identifier: https://dx.doi.org/10.6084/m9.figshare.3114844.v1 (tarball of version 0.0.2)
Given the modular structure of SeFo, it could be extended with comparatively little additional work within the seasonal forecast context to accommodate alternative weather variables (such as precipitation or sunniness, although because these are farther than temperature from a normal distribution, some modification in the forecast methods would be advisable [32-33]), observation data sources, sources of ensemble outputs (besides NMME), and methods for generating forecasts from ensemble outputs and past observations. Many of the components of SeFo, including the specific forecast and verification methods implemented in the functions called by sefo_predict and sefo_verify, could also be reused for forecasting applications in fields outside of weather prediction.

Competing Interests
The author declares that they have no competing interests.

References
1. National Research Council Assessment of Intrasen-

to to Interannual Climate Prediction and Predict-


org/10.17226/12878

2. Troccoli, A 2010 Seasonal climate forecasting. Me-

teorological Applications, 17: 251–268. DOI: http://

dx.doi.org/10.1002/met.184

3. Alemu, E T, Palmer, R N, Polebitski, A and Meaker, B

2011 Decision support system for optimizing res-

ervoir operations using ensemble streamflow pre-

dictions. J. Water Resour. Plng. and Mgmt., 137 (1):

72–82. DOI: http://dx.doi.org/10.1061/(ASCE)

WR.1943-5452.0000088


is the current state of scientific knowledge with regard

to seasonal and decadal forecasting? Environmental

Research Letters, 7 (1): 015602. DOI: http://dx.doi.

org/10.1088/1748-9326/7/1/015602

5. Doblas-Reyes, F J, García-Serrano, J, Lienert, F, Bie-

escas, A P and Rodrigues, L R L 2013 Seasonal cli-

eate predictability and forecasting: status and pros-

pects. Wiley Interdisciplinary Reviews: Climate Change,

4 (4). DOI: http://dx.doi.org/10.1002/wcc.217

6. Arribas, A, Glover, M, Maidens, A, Peterson, K,

Gordon, M, MacLachlan, C, Graham, R, Fereday, D,

Camp, J, Scaife, A A, Xavier, P, McLean, P, Colman,

A and Cusack, S 2011 The GloSea4 ensemble pre-

diction system for seasonal forecasting. Monthly


dx.doi.org/10.1175/2010MWR3615.1


