

Standards for EUROSIP forecast and hindcast production

Initial Proposal: 21st July 2008

Multi-model products can be constructed in various ways from diverse sources of input data. In particular, if only rough estimates of ensemble mean anomalies or tercile probabilities of seasonal means are needed, the requirements on the input data are not very strong. We expect that international exchange of basic forecast data will enable us to construct a very basic set of "global multi-model" forecast products in the future, using robust software to handle the variety of data inputs.

The EUROSIP multi-model system, by contrast, is designed to deliver a comprehensive and coordinated set of multi-model forecast data to the user. This supports the creation of a very wide range of products, based on the use of both monthly and daily data, and using the full hindcast information for proper calibration of the forecast products, including corrections based on past forecast skill. As in the very successful DEMETER and ENSEMBLES projects, ease of use of the data is considered a key factor - the EUROSIP multi-model dataset is both comprehensive and easy to use. We do produce a basic set of graphical products, and we are about to disseminate a basic set of multi-model "EUROSIP products" in grib form, but the basic multi-model data archive is the key input that we provide to the ECMWF member states.

The multi-model EUROSIP products are produced for those who are unable or unwilling to access the raw data of the individual models, in particular commercial customers. To allow sensible use of the forecast products, we provide the equivalent "hindcast" multi-model products. The length of this hindcast product set is determined by the **shortest** record length of the contributing models. Thus to preserve the overall usefulness and value of the multi-model EUROSIP products, we need to set standards for the minimum hindcast set for a system used to produce the multi-model products.

In principle, bigger ensemble sizes are always better. However, at least with our present software, there are limits as to how much data can be processed at once in a multi-model combination. The standards below are the limits of what will actually be processed at the present time. If larger ensemble sizes are available, we will be happy to archive them, assuming the total data volumes are not excessive. Larger ensemble sizes might be useful in the future.

Proposed standards

1. All forecast data must be supplied so as to fit the common data structure. For the time being, this means that the forecast and hindcast data must be **in the form of** an ensemble starting at 0Z on the 1st of every month, and extending to the required length. Note that the ensemble might still be generated by a lagged average technique (eg collecting a set of runs made over the preceding 4 weeks). The forecasts and hindcasts must be produced in a compatible way. As long as this is done, and for the purposes of seasonal prediction, the

statistical discrepancies between lagged average and burst mode ensemble are expected to be second order, compared to issues such as model error.

2. Real-time forecast ensemble size should be equal to 41 if possible, and must be at least 30.
3. Hindcasts should cover the period 1981-present if possible. If this is not achievable, a minimum period of 20 years (eg 1989-2008) is requested.
4. Hindcast ensemble size should be 11 if possible, and must be at least 5.
5. The model used for operational seasonal forecasts should not normally be changed without a reasonable level of testing to ensure that the forecast skill is at least maintained. Three months' notice of any proposed changes should be given, together with a brief summary of the changes and the expected impacts.
6. Ideally, most or all of the hindcast data should be provided before a new model version becomes operational. If this is not possible, hindcast data must be provided at least 3 weeks prior to the corresponding real-time forecast.

It is expected that these standards will be met before a new model becomes part of the operational EUROSIP system, and that contributors will strive to maintain them. Nonetheless, if there are specific reasons as to why a model might need to miss these standards for a limited period of time, the EUROSIP system will attempt to cope with the situation as best it can. Advance notice of any such issues will be much appreciated, to allow work-arounds to be explored.