

On the ensemble mode in HirLam 7.1

Kai Sattler, DMI

1 Introduction

This note describes some aspects of the Hirlam 7.1 script system development for the Hirlam ensemble prediction system. In the framework of the GLAMEPS initiative of the HIRLAM A program, work has been started to implement into the HIRLAM reference system the capability to run ensemble forecasts. The HIRLAM A scientific plan for 2006 states among others:

- As a first step, existing LAM EPS systems in member countries should be the starting point for a careful design of a computational framework for short-range HIRLAM EPS.
- The HIRLAM EPS framework should include probabilistic validation and verification tools that should be applied for evaluation of initial, as well as new when the become available, components of the system
- A possible starting point for HIRLAM EPS could be the met .no downscaling approach combined with INM multi-boundary/multi-model approach and possibly also lagged averaging techniques

The scientific plan includes the goal that the Hirlam ensemble prediction system is to be capable of serving as a laboratory for investigation and development of methods for numerical short range ensemble prediction. The work on the HIRLAM ensemble prediction system (HIREPS) started with a working week in Norrköping in June 2006. Beyond other work, the following tasks concerning the Hirlam EPS script system were achieved:

- LAM EPS scripts from met .no as starting point
- script adaptation to run in downscaling mode from ECMWF-EPS data
- MARS retrieval of online EPS boundaries
- 3-member test case
- work on SLAF (Scaled Lagged Average Forecasts) initial perturbation method started

Another working week followed in January 2007 (Reading) with the following content:

- update to trunk version
- 3DVAR option for control run
- random choice of convection scheme for perturbed members
- GLAMEPS version 0 domain defined
- work on TEPS
- archiving strategy

In addition, work on post-processing and SLAF was performed some weeks later in March 2007 by staff from INM.

The following sections describe some ideas and details on the current state of the Hirlam EPS script system, as of the time of the ASM 2007. They should be seen as an add-on to the ASM 2007 presentation with the same title as this note, explaining especially the concept of random lists in more detail. The scripts system is still under development, and the newest description and documentation can be found under the *wiki* pages on <https://www.hirlam.org/trac/wiki>.

2 HIRLAM EPS system

2.1 Some basic ideas

During the planing phase of Hirlam EPS, some ideas were formulated with respect to the design of the system. One of these ideas is the possibility to use different methodologies for ensemble prediction. This way, investigation and development of short range ensemble methods should be facilitated.

Another idea is the capability of the system to run on different platforms. Even though the ECMWF computer platform is regarded as the major platform for Hirlam EPS, installation on local computers must be possible. This is a basic necessity in order to perform distributed production within GLAMEPS.

The current design of the script system attempts to account for these demands from the start.

2.2 The current HIRLAM EPS system

The Hirlam EPS system is still under development. The main characteristics of the system as of the time of the Hirlam-A All Staff meeting 2007 comprise the features:

- allow for deterministic as well as ensemble runs
- use of online ECMWF-EPS boundaries
- analysis in control run
- ensemble member specific environment settings (currently implemented: choice of convection/condensation scheme). Settings may be fixed or random.
- processing single member or group of members
- administration of up to 999 members

The only available ensemble methods are currently ensemble downscaling of ECMWF-EPS data (boundary perturbation), simple perturbed analysis (initial perturbation) and a random choice between KFRK and STRACO scheme. Use of targeted ensembles (TEPS) at the boundaries is to be implemented as a next task, and the SLAF initial perturbation method is to be implemented soon, too.

The Hirlam EPS system allows for the following run modes:

1. deterministic, using `hireps.tdf` (mSMS for LAM EPS) or `hirlam.tdf` (standard mSMS)
2. control member only
3. one single perturbed member

4. multiple perturbed members
5. full ensemble mode (control plus perturbed members)

2.3 Random lists

A specific feature of the Hirlam EPS system tries to manage random events that are implemented in the system, by establishing random event lists specifically for a certain type. The creation of the random lists happens during initialization of the ensemble, and it is used when running the ensemble members. This way the random event is *determined* beforehand for later use. Even though it may sound like a paradox, this predetermined random event still retains all stochastic properties. It is similar to pseudo random numbers, but with the difference that the numbers are listed beforehand, which should retain reproducibility even across different platforms.

An example of where the concept of random lists is applied in the current Hirlam EPS is the random choice of the convection/condensation parameterization scheme. HIRLAM includes two of these schemes: KFRK and STRACO. A random choice between these two is a random event of size 2. Instantly one would determine what scheme to use during configuration of each ensemble member, by calling a random number generator, and then mapping the result to the event type of size 2. However, this call can be done already at an earlier stage, and the information can be *stored* until needed. The storage makes it possible to reproduce the run in order to investigate runtime problems or bugs etc.

The basic idea behind the use of random lists is the possibility to fully reproduce Hirlam EPS simulations, even though random events are included in the runs. By use of random lists the random choice can even be retained over different computer platforms, a feature, which would not necessarily be possible when making use of the same start values in random number generators.

The concept of random lists is implemented in a first version in the Hirlam script system within the two shell-scripts `EnsInit` and `Env_ensmbr`.

`EnsInit` is responsible for preparing ensemble related actions. Currently this script just includes the creation of random lists. The only random list currently implemented is the one for the random choice of the convection/condensation parameterization scheme.

`Env_ensmbr` is an environment definition script (as indicated by the name), which is called each time an ensemble member is processed. It is responsible for establishment of the ensemble member specific environment, like e.g. setting environment variables. Choice of the convection/condensation scheme can be specified here, for example. This can be done either explicitly (Fig. 1), or using random choice by exploitation of the appropriate random list (Fig. 2).

```

#===== explicit settings

case $ENSMBR in
  001) COND_SCHEME=KF
      ;;
  002) COND_SCHEME=STRACO
      ;;
  *)  COND_SCHEME=STRACO
      ;;
esac

```

Figure 1: Example for explicit ensemble member specific environment settings in the script `Env_ensmbr`.

```

#----- random settings

for eventtype in `echo $RND_EVENTTYPES | tr : ' '; do
  random_list="${HL_DATA}/${CYCLEDIR}/../ensemble/random_${eventtype}_${ENSSIZE}.list"
  if [ -f $random_list ]; then
    case $eventtype in
      con) i=`expr $ENSMBR + 0` # choice of convection/condensation scheme
          random_number=`cat $random_list | cut -d ' ' -f $i`
          case $random_number in
            0) COND_SCHEME=STRACO
              ;;
            1) COND_SCHEME=KF
              ;;
            ...
          esac
          ...
        ...
      esac
    fi
  done

```

Figure 2: Example for randomly chosen ensemble member specific environment settings in the script `Env_ensmbr`.

2.4 Computer resources

Performing Hirlam ensemble experiments can be very expensive with respect to computing resources. Some first measurements have been undertaken using the GLAMEPS 0.2 degree domain (EPS71_20: 306x250x40 grid, 20+1 members over 3 days with 12h cycling), as well as using a GLAMEPS test domain (EPS71T_15: 60x52x40 grid, 15+1 members over 7 days with 12h cycling).

The results are shown in Table 1. They show that the use of the GLAMEPS 0.2 degree domain takes significant computer resources. For technical testing and during the setup phase of an experiment, it is therefore recommended to make use of the GLAMEPS test domain **EPS71T** in order to get reasonable responses and to save computing resources.

experiment	elapsed time				SBU's			
	forecast	cycle	3 d	7 d	forecast	cycle	3 d	7 d
EPS_71T_15	35 – 47s	≈ 45M		10h	5	80		1200
EPS71_20	550 – 650s	≈ 160M	≈ 15h	1.5d*	60 – 70	1400	8500	20000*

*estimated

Table 1: Consumption of computing resources on the ECMWF computer platform for two Hirlam EPS experiments.

2.5 Outlook

The next steps in the development of the Hirlam EPS script system contain inclusion of processing TEPS boundaries, enhanced cycling of data assimilation and of course inclusion of more perturbation methodologies.

Whether the idea of full reproducibility of the ensemble simulations can be applied also for model internal random events like stochastic physics has to be seen. A full reproducibility even of perturbed Hirlam ensemble runs with stochastic properties would be a desirable feature in the EPS system.