# Introductory Information for ASP Summer Colloquium 2003 Exercises

## 1. Overview

The colloquium exercises are designed to help students gain a better understanding of data assimilation issues that are being discussed in formal presentations. The exercises cover a variety of assimilation techniques, models, and observation types. While most of the exercises are interactive, allowing students to create and analyze their own assimilation experiments, there is also an attempt to provide access to information about pre-existing assimilation systems and data sets. Because of resource limitations, many of the exercises will be completed by small groups of two or three individuals. In the spirit of the colloquium, we hope that this will lead to useful interactions among members of the group. Computer exercises will be undertaken in a common room, and a set period is available on most days of the colloquium during which exercises will be the primary focus of all participants. At these times, experts on both the exercise sets and data assimilation methods will be freely available to consult with student teams about both implementation and theory of the exercises.

It will quickly become obvious that the exercises depart from many of the lectures in two ways. First, almost all the exercises make use of models and observation sets that are orders of magnitude smaller than the operational prediction systems that will be the focus of most lecturers. Almost by definition, operational prediction centers make use of the largest and most complicated models that can be supported on state-of-the-art high end computing systems. Since data assimilation can be the most computationally taxing part of an operational prediction suite, it follows that allowing a group of students to perform real-time data assimilation experimentation in models even remotely approaching the size of operational models is impractical. In addition, the complexity and size of the models can often obscure important aspects of the assimilation methodology. We will address this problem by using a hierarchy of models in our exercises, ranging from a model with only 3 state variables through a low-order atmospheric general circulation model and a state-of-the-art regional atmospheric model. With the smaller models, students will be able to generate large sets of experiments to explore the sensitivity of assimilation algorithms to a number of parameters and to characteristics of observational sets. In the larger models, only a limited set of experiments can be performed and, in some cases, results may not be available until the next day's exercise period. Students will have to use care in generalizing their results from low-order models before performing experiments in the larger models. It is also important for students to keep in mind that not all aspects of low-order exercises can be cleanly extended to the operational prediction systems that are the primary focus of the lectures.

A second way in which the exercises differ from the lectures is in the variety of assimilation algorithms that will be examined. Although the exercises try to touch on a variety of the data assimilation algorithms that are discussed in presentations, they are nevertheless predominantly related to ensemble filtering assimilation algorithms. The first reason for this is that ensemble filters are remarkably easy to implement and modify for low-order model application. This should give students a chance to explore many issues related to data quality, density and distribution that might be more difficult to address in a flexible way with variational algorithms. In addition, two of the three organizers share expertise in ensemble filtering algorithms and their natural inclination has

been a bias toward these methods in the exercises. Again, students should keep this in mind and be careful not to let the inevitable bias towards filter methods in the exercises mislead them about the capabilities of other more established methods that will dominate the lectures.

The basic exercise sets are designed to fill most of the seven 2-hour periods set aside for exercises at the colloquium. However, students desiring to pursue more aggressive projects (some of which are suggested in the exercise descriptions) will be free to spend additional time in the evenings or on the weekend working with the exercise assimilation system. It is also possible to obtain copies of the exercise system which could be installed on computer systems at the students' home institutions for continuing use after the colloquium.

## 2. Computer accounts

ASP summer colloquium students have guest accounts on two of NCAR's MMM division computing platforms. The account names are mmm01, mmm02, ..., mmm28 and the password for all accounts is asp.symp. Students will be assigned a particular account number during the first day of exercises.

The guest accounts are valid on the DEC workstations in the MMM Layton computer laboratory where exercise sessions will be held. There are 14 of these workstations named service01, service02, ..., service14. They are also valid on the Aspen Linux cluster 'ocotillo' which can be accessed by telnet, ssh, or ftp from the DEC workstation. There is no domain-wide capability to change passwords on the DEC machines or 'ocotillo', so please be careful to avoid accidentally logging into another users account.

Each mmm?? account has a quota of 200 MB on the home directory /users/mmm??. There is also about 1.4 GB available in the directory /usr/tmp on the DECs. As some of the later exercises are extremely data intensive, students will have to be very careful to manage this limited disk space.

On 'ocotillo', students have home directories in /ocotillo/users/mmm??. There is very limited space in the /ocotillo disk and it should not be used for storing large datasets or executables. Students also have space in /ocotillo2/mmm??; this disk has a total capacity of about 70 GB and should be the primary storage area used by students. Even here, students must be extremely conservative to avoid overflowing the disk during the second week of the exercise set.

Students do have access to mail accounts on the DEC machines and are advised to use the mail utility 'elm' to access their mail. It is possible that the colloquium organizers will send mail with updated information about the exercise sets to the student accounts during the process of the colloquium. If so, this will also be mentioned during the lectures.

Students will also require web access for some of the exercise sets and to access on-line journals and other materials. The supported web browser is Netscape which should be started by issuing the command

```
netscape &
```

from a terminal window on the DECs. Do not use any of the menu-driven browsers from the DEC as they access seriously out-dated browser versions.

If students require printed output, the default printer is ‘mango’ which is located on the south side of the third floor in the same building as the Layton computing laboratory. Ask one of the organizers to direct you to the printer.

## 3. Summary of exercise sets

The colloquium will include a series of independently developed exercise sets, each with a principal developer.

A. Weather Research and Forecast (WRF) 3D-Variational overview (Dale Barker)

This exercise set is a tutorial for the Weather Research and Forecasting mesoscale model’s 3D-variational assimilation facility. It was created by the WRF data assimilation team and we are just borrowing it for the colloquium. It is web-based and can be found at as a link off the MMM assimilation home page:

http://www.mmm.ucar.edu/mm53dvar/

on the link on-line 3dVar tutorial.

B. Fundamentals of data assimilation (Chris Snyder)

C. 4D-Variational Assimilation in a low order model (Jenny Sun, Dale Barker)

D. Survey of Ensemble Filtering in the Data Assimilation Research Testbed (Jeff Anderson)

E. Students may also be interested in surveying a variety of web-sites that are related to operational data assimilation. A list of interesting sites will be provided.

F. Additional reading: Students may also wish to do directed reading on a specific data assimilation topic and are encouraged to consult with any of the organizers or lecturers to formulate a list of relevant papers which can be obtained from NCAR’s library (either on-line or paper).