Tutorial Evaluation Data Report

Migration from the Desktop: HPC application of R and other codes for biological research
May 9-11, 2011

Pamela Bishop
Program Evaluation Coordinator
National Institute for Mathematical and Biological Synthesis
June, 2011

Table of Contents

Table of figures

Migration from the Desktop: HPC application of R and other codes for biological research

Evaluation Design

Evaluation Questions

The evaluation of the tutorial was both formative and summative in nature, in that the data collected from respondents was intended to both gain feedback from respondents about the quality of the current tutorial and also to inform future similar events. The evaluation framework was guided by Kirkpatrick's Four Levels of Evaluation model for training and learning programs (Kirkpatrick, 1994[1]). Several questions constituted the foundation for the evaluation:

1. Were participants satisfied with the tutorial overall?
2. Did the tutorial meet participant expectations?
3. Was the tutorial appropriate to the participants' levels of expertise?
4. Did participants feel they learned an appropriate amount of information?
5. Were participants satisfied with the amount of content and format of the tutorial?
6. Were participants satisfied with the accommodations offered by NIMBioS?
7. What changes in accommodations, group format, and/or content would participants like to see at future similar meetings?

Evaluation Procedures

An electronic survey aligned to the evaluation questions was designed by the NIMBioS Evaluation Coordinator with input from the NIMBioS Director and Deputy Director. The final instrument was hosted online via the University of Tennessee's online survey host mrInterview. Links to the survey were sent to 26 tutorial participants on May 12, 2011 (tutorial organizers and participants associated with NIMBioS were excluded from the evaluation). Reminder emails were sent to non-responding participants on May 19 and 24, 2011. By May 31, 2011, 24 participants had given their feedback, for a response rate of 92%.

---

[1] From Kirkpatrick, D.L. (1994). *Evaluating Training Programs: The Four Levels.* San Francisco, CA: Berrett-Koehler.

Evaluation Findings

Respondent Satisfaction

Figure 1. Satisfaction with various aspects of tutorial

Scale: -2 = Strongly disagree to 2 = Strongly agree
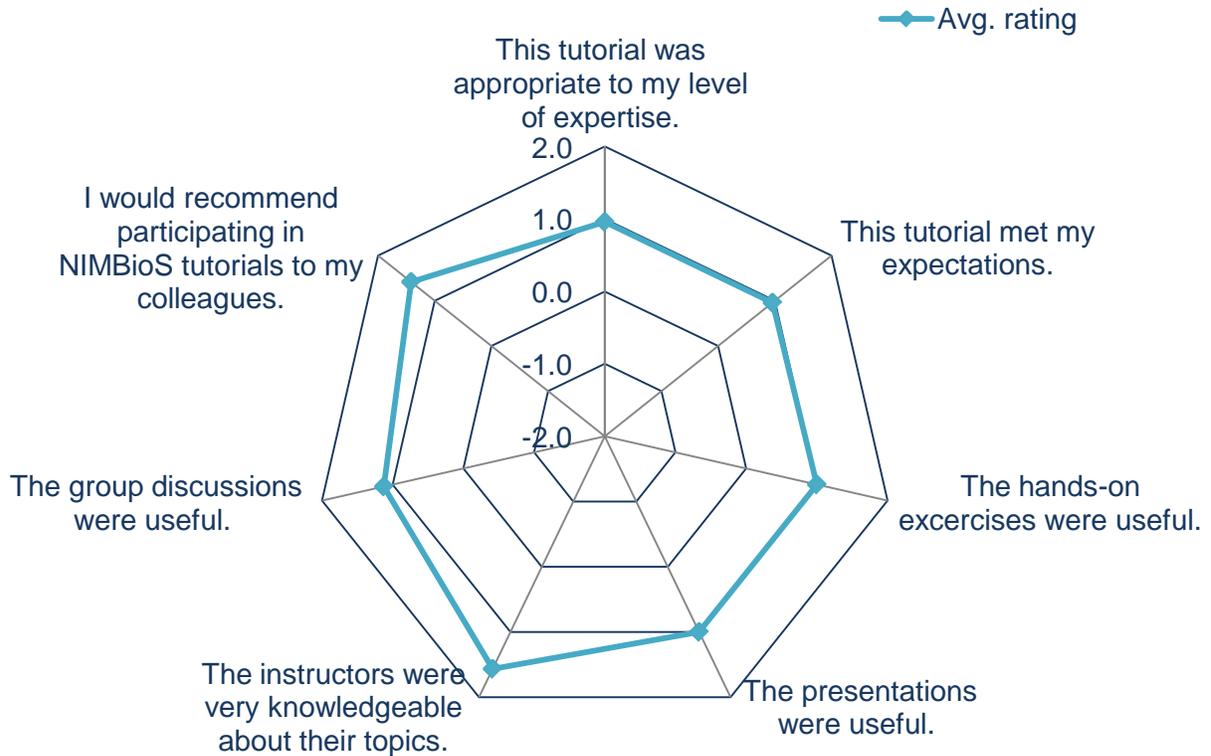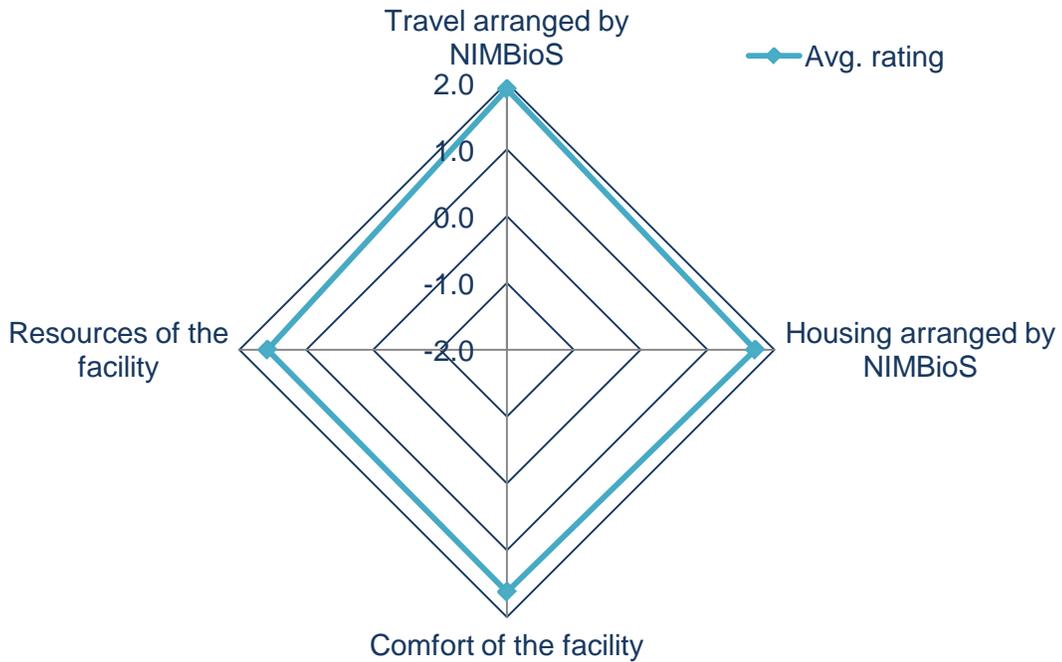
Figure 2. Satisfaction with accommodations
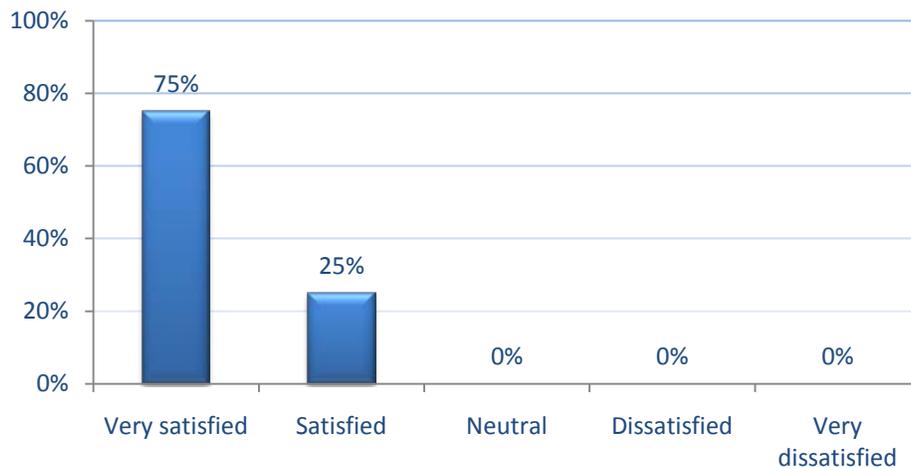
Scale: -2 = Very dissatisfied to 2 = Very satisfied

Tutorial Format and Content

Format

Figure 3. Satisfaction with opportunities provided during tutorial presentations and discussions to ask questions and/or make comments

Please indicate any suggestions you have for facilitating communication among participants during the tutorial:

Allocating time for a brief discussion/presentation about individual research work and how HPC and application of R would help achieve our goals, with some time for discussion on each of these aspects by other participants/organizers.

Dividing the group in sub groups that would work on common interests.

I think the tutorial was very well organized and provided a good venue for communication between participants.

It might be nice to, before the seminar, ask if any participants would like the opportunity to give short (<5 slides, <5 minutes) presentations about their relevant research and then have the instructors pick a few that seem useful for the class. A half hour for 5 or so such mini-presentations might help students see some of the potential uses and expand discussion possibilities.

Some sort of email list put together post meeting would be great. Finally, I'd really like to thank Amy, Eric and others for all the effort that went into this meeting. Thanks!

The members of NIMBioS were very friendly, helpful and I learned greatly. Thank you.

Figure 4. Effectiveness of tutorial format
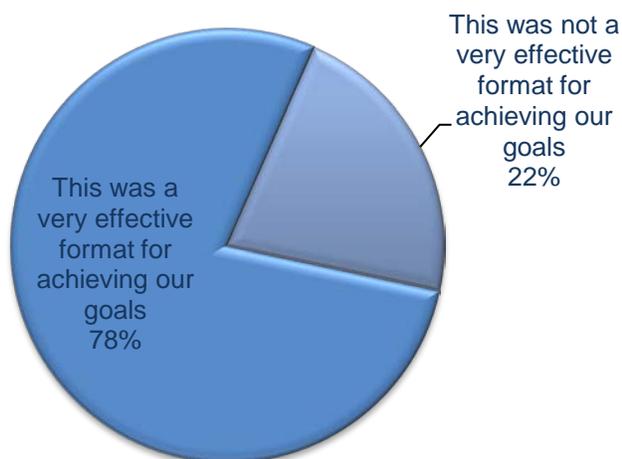
**The tutorial format would have been more effective if:**

A couple of days more could have been useful for more details in using the tools presented.

A mini project with handling HPC was offered.

Two broad things: a) One thing to realize is that we are mostly biologists who really only need to know as much about computer chips as it is to get the science done. So some of the talks were a bit too jargony and would have been helpful if it was geared to the right audience (so examples are Pragnesh's talk and the first talk after Lou). b) If we had a smaller class size, and participants were selected from a range of backgrounds, we could easily split into groups and run (albeit small job) something to completion on EDEN; compare how some trivial code we wrote runs on a cluster to what it would do on a smaller desktop etc. That really is the only way to learn. My last night with a few participants, people mostly said, I still don't know what is going on but at least I heard a few more terms.

Content
Figure 5. Amount of content provided
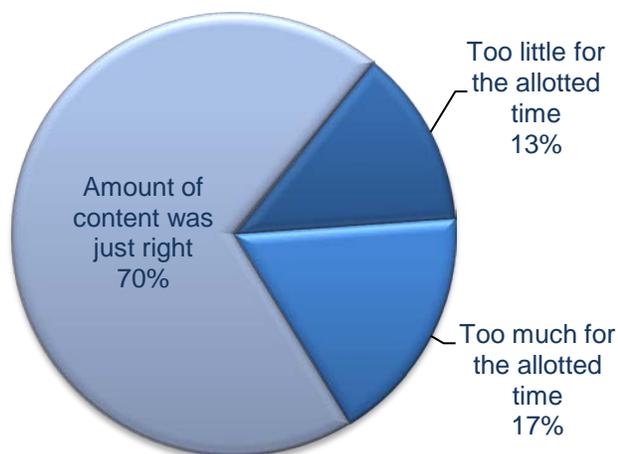
What do you feel was the most useful aspect of the tutorial?

(1) Hands on training sessions on 'R' and VISIT. (2) Meeting like minded group of researchers.

The tutorial provided an excellent combination of application and methodology. Overall, excellent work.

Cluster access and Unix. Visualization tools.

Focused introduction with knowledgeable instructors, access to code and computers, and the time to ask questions and have them answered to the depth needed - all of that together is quite rare, and this class was therefore quite useful for a quickly getting up to speed in this area.

For me, the most useful aspects were hearing about Teragrid, and how to push R to do things on the grid. In that respect, Pragnesh's talk was the most useful one. However, he breezed through a ton of material in no time at all and he wasn't all that easy to follow. So perhaps allot more time to him (say an entire morning session) so we can absorb the material, have plenty of Q&A and maybe do a demo.

General information on resources and knowledge of programming language for HPC.

Group discussions.

Hands-on experience with HPC tools, from Vista to simple details about how to login and access the queue.

Having access to the expertise of staff at NIMBioS.

How to access to the HPC resources. Introduction to R.

I got an incredibly clear sense of what is involved with running an HPC system, and what is necessary to support scientific users (all the different levels of support).

Learning how to log onto nautilus, and learning how to get funding to use it in the future.

R graphics

The interaction with instructors and NIMBioS postdocs that are indeed very knowledgeable in the use of HPC resources. The instruction about the resources that are available presently to perform HPC.

*The packages and tools we learned about were useful, and how to use a cluster and write Unix code.*

The R-specific tips and tricks! Also, being able to discuss computing issues with people from a lot of different backgrounds

Participant Learning

Figure 6. Participant knowledge gains

Scale: -2 = Strongly disagree to 2 = Strongly agree

As a result of attending the tutorial, I have a better understanding of:
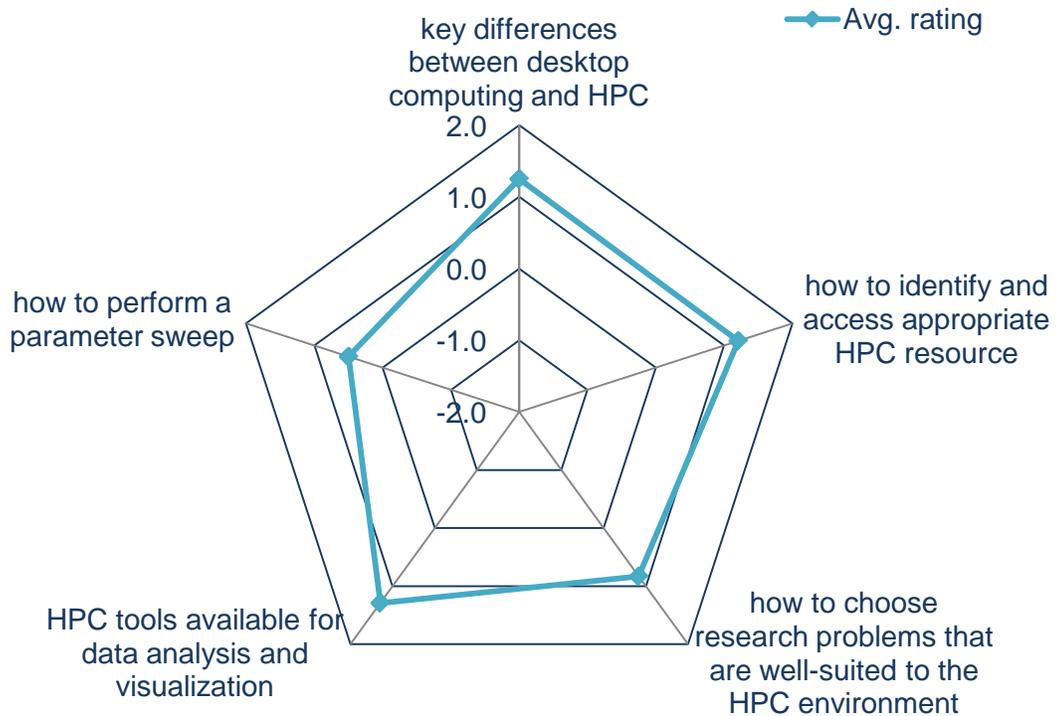
Suggestions for Future Tutorials

What would you change about the tutorial?

A bit more time in order to give more details about the topics discussed.

Actually give us things to do after you show us something. I fully appreciate the difficult in organizing such an event and the first time is always a learning experience. Scott's presentations were poorly organized and most of the time he was just thinking out loud. Perhaps next time it could be more focused with a hands on section that followed. I found it difficult to get the attention of organizers when I needed help. This may be, in part, due to the fact that we had so many participants.

Add less info to the Visit tutorial, and have it be more step-by-step for people

who aren't familiar with it.

I expected to be coding and running some R code using HPC. There was too much time allocated to presentations of applications and too little to actual HPC. Though the science was interesting, the presentations were tangential to learning HPC. The group is already motivated to learn HPC; there was no need for in-depth exploration of population genetics, for example.

I thought the tutorial was really useful but at times (not being a computer scientist) I felt that the pace was a little too fast. It might be better to extend the workshop by a day or two and give a slower introduction to topics that for a biologist might be new. Also, I found the use of acronyms also to be a little hard to follow so explicitly spelling those out (at least for the first few slides) would be helpful. It would be useful to run an example in R too, to test out the packages we learned about.

I would add a mini project.

I would like more technical details of the material presented.

I would like to see more hands on that would actually use the cluster. Using the visualization softwares (visit and paraview) in the desktop was just like using any other visualization software developed for desktop use (there are so many). The instructors should have used a huge large data set to exemplify an interactive use of visualization. This would have clearly shown what is particularly different about this software. The same is valid for other sections, like using R code with big problems or big files directly in nautilus.

It would have been helpful to schedule a problem session for those with specific questions and seeking guidance.

Less lecture, especially those that did not directly deal with the HPC environment. There was hardly any R coding within the HPC environment which was a disappointment and would like to see a lot more of that.

More hands on for R. How to install a program. Some more examples for a batch job.

Personally, I would have liked less time on data visualization and more time on HPC technical issues, coding, etc.

Possibly more handouts before we started the tutorial

R tutorial started abruptly and looked more like for someone who already knew how to use it. I had never used it and was clueless about what was happening.

Scott Simmerman had much useful information, but he did not seem that

concerned about connecting clearly with the audience.

There didn't seem to be enough vegetarian-friendly food- it ran out long before the other things did.

To have more time on each sessions (particularly the hands on training sessions).

What topics would you have liked to have covered in this tutorial if given more time?

Analysis of sequencing data using HPC

Further hands on training on use of 'R' in an HPC environment.

Hardware choices and set up of small clusters - 30-60 cores

If enough time might have permitted, I would like to have small project so that we might learned better

Instead of just presenting functions and commands in some language for HPC, I would like to see a full code and how to execute from the beginning to the end on HPC architecture with at least on real problem in bioinformatics.

More about parallel programming

More details and exercises on PBS scripts and R parallel programming.

More hands-on coverage and more exercises for attendees to work on their own.

More hands-on HPC with R.

More on using R for hp computing, more background on the topic

More specific examples of R code using some of the packages, specifically developed for HPC, shown in presentations.

The hands on tutorials were poorly organized. We spent far too much time just getting a majority of people set up with software and access to the system. This should have been done very early on. As a result, the actual demo was rushed and not very useful. If you do this next time, please hand out a list of software and a questionnaire to gauge participant expertise so you can plan accordingly. I am not a Unix expert, but I use HPC resources frequently enough that using vi, ssh'ing into a system, install software on a Unix/osx environment are very easy. However, I do understand that these things can be rather intimidating for someone who hasn't used such resources. So there should be some balance such that more experienced users aren't just sitting

around doing nothing.

The material related to the technical analysis could be expanded.

Worked examples of HPC data analysis. Most of the hands-on exercises involved relatively simple code, rather than demonstrating how to move on to statistical procedures.

Additional Comments

Please use this space for any additional comments:

Attending the tutorial was a very positive experience. I really appreciate how friendly everyone at NIMBioS was and the extra effort taken by Eric, Amy, Jennifer and Chandra. I was impressed by how well the visualization tools and cluster access were taught.

I really liked the workshop; I would just suggest that it is either extended in the future and the pace a little slower.

It is good if we have a communication platform in the future

Thank you - I learned greatly and would like to thank the organizers. This tutorial was extremely helpful.

Thank you for your wonderful work.

Thank you so much for this great opportunity to meet an excellent group of people that are advancing with the great needed development of quantitative Biology.

Thanks for giving me the opportunity to learn with you!

Thanks!

The travel agency was very helpful in dealing with flight cancellations and rebooking.

This tutorial was timely for me and my research, thank you!

This was a supremely well-organised meeting and I commend the administrative team for ensuring that everything ran so smoothly.

Appendix

HPC Application of R Tutorial Evaluation Survey

HPC Application of R Tutorial Evaluation Survey

Thank you for taking a moment to complete this survey. Your responses will be used to improve the tutorials hosted by the National Institute for Mathematical and Biological Synthesis. Information supplied on the survey will be confidential, and results will be reported only in the aggregate.

Please check the appropriate box to indicate your level of agreement with the following statements about this tutorial: (Very satisfied, Satisfied, Neutral, Dissatisfied, Very dissatisfied)

       The tutorial was appropriate to my level of expertise.
       The tutorial met my expectations.
       The hands-on exercises were useful.
       The presentations were useful.
       The instructors were very knowledgeable about their topics.
       I would recommend participating in NIMBioS tutorials to my colleagues.

How do you feel about the amount of content offered during the tutorial?

       Too little for the allotted time
       Too much for the allotted time
       Amount of content was just right

Please check the appropriate box to indicate your level of agreement with the following statements. As a result of participating in this tutorial, I have a better understanding of: (Strongly agree, Agree, Neutral, Disagree, Strongly disagree)

       key differences between desktop computing and HPC
       how to identify and access appropriate HPC resource
       how to choose research problems that are well-suited to the HPC environment
       HPC tools available for data analysis and visualization
       how to perform a parameter sweep

What topics would you have liked to have covered in this tutorial if given more time?

What do you feel was the most useful aspect of the tutorial?

What would you change about the tutorial?

How do you feel about the format of the tutorial?
       This was a very effective format
       This was not a very effective format
          The tutorial format would have been more effective if:

Please indicate your level of satisfaction with the tutorial accommodations:
(Very satisfied, Satisfied, Neutral, Dissatisfied, Very dissatisfied, Not applicable)

       Travel arranged by NIMBioS
       Housing arranged by NIMBioS
       Comfort of the facility in which the tutorial took place
       Resources of the facility in which the tutorial took place

How satisfied were you with the opportunities provided during tutorial presentations and discussions to ask questions and/or make comments?

    Very satisfied
    Satisfied
    Neutral
    Dissatisfied
    Very Dissatisfied

Please indicate any suggestions you have for facilitating communication among participants during the tutorial:

Please provide any additional comments about your overall experience with the tutorial: