

USE OF HTML AND WEB TOOLS IN ATMOSPHERIC SCIENCES RESEARCH

Brian F. Jewett and Robert B. Wilhelmson¹
Department of Atmospheric Sciences and
National Center for Supercomputing Applications (NCSA)
University of Illinois at Urbana-Champaign

1. MOTIVATION

It has long been true that carrying out research requires a great deal of organization. This includes keeping track of relevant published literature, archiving raw and processed data for later perusal, and preparing conclusions and publications based on that data. What is different today is the vast amount of information created in one's own research as well as in joint research with others. Finding an efficient means to access this data is essential to make the most of your work.

This data comes in a variety of forms: notes, graphics, images, animations, model code, data files, and publications. The usual means of managing this data are varied: desk file folders, paper stacks on desks and floors, computer files, video, 8 mm data tapes, etc.

Our research involves numerical simulations of cold fronts and squall lines. As we reviewed our own management system for studying model results, writing research notes, viewing images, and keeping a (sometimes written, sometimes online) log of what distinguishes one simulation from another, it became apparent that an improvement was possible through the use of the relatively simple to learn

HTML (HyperText Markup Language) and network information browsers (see Schatz and Hardin² for an overview). Using this strategy we can access our information using HTML provided that this information is stored electronically using hyperlinks.

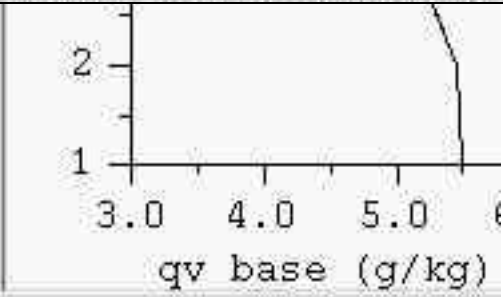
These clickable hyperlinks can be organized in one or more HTML documents in ways meaningful to our research and provide easy access to other HTML documents, images, animations, and voice recording stored on various networked compute servers including our own desktop workstations. This "electronic notebook" has been a substantial help in maintaining our research data and in browsing and carrying out further analysis of our model results.

2. TYPES OF INFORMATION

Table 1 lists the type of data we keep online in our electronic research notebook along with examples of the displays we find useful.

¹Corresponding author addresses:
Brian F. Jewett and Robert B. Wilhelmson
NCSA, 5249 Beckman Institute, 405 N. Mathews,
Urbana, IL 61801
E-mail: bjewett@ncsa.uiuc.edu,
bw@ncsa.uiuc.edu
Phone: (217) 244-1998 (Jewett),
(217) 244-6833 (Wilhelmson)
FAX: (217) 244-2909

² Schatz, B. R., and J. B. Hardin, 1994:
NCSA Mosaic and the world wide web:
Global hypermedia protocols for the Internet.
Science, **265**, 895-901.

Table 1. Information kept online																
Type of information	Example (taken from online Mosaic display)															
Model textual output (HTML preformatted text) is taken directly from the model and placed online. This is ideal for data added often and needing little formatting.	>>Frontal collapse by time enter time (DAYS) for qv ev qv evolution will begin at Collapsed: time= 939200.0.															
Simulation logs are lists of simulations, including parameters that differentiate the runs from one another as well as links to images from each experiment. This is now done using HTML tables which can also contain hyperlinks.	Simulation Data <table border="1"> <thead> <tr> <th>Run</th> <th>dx</th> <th>dR</th> <th>QT</th> <th>Min/Max</th> </tr> </thead> <tbody> <tr> <td>30W</td> <td>32</td> <td>25</td> <td>12</td> <td>summary</td> </tr> <tr> <td>31c</td> <td>32</td> <td>15</td> <td>12</td> <td>summary</td> </tr> </tbody> </table>	Run	dx	dR	QT	Min/Max	30W	32	25	12	summary	31c	32	15	12	summary
Run	dx	dR	QT	Min/Max												
30W	32	25	12	summary												
31c	32	15	12	summary												
Reference lists are kept of journal articles of interest. References are listed alphabetically and by topic. Articles which have been read may have short summaries typed in and linked to the reference lists.	Emanuel, K. A., M. Fantini, instability in an enviro moist convection. Part <i>Sci.</i> , 44 , 1559-1573. Farrell, B., 1985: Transie															
Publications ³ such as abstracts and preprints are stored online, along with other material destined for later publication. Formatting includes use of unnumbered lists, headings and rules along with equations, charts and other figures (converted to GIF and stored as in-line images). We as well as others ⁴ have found it useful to make publications easily available to the wider web community in this way.	Solving for the grid 2 damping coefficie $K_2 = \frac{1 \pm \sqrt{1 - \alpha K_1}}{\alpha} \quad (B.7)$ <p>I choose the negative root above, as the stability range in B.5). This expression Nonetheless, K2 is usually one-half K1 K1. Using the series expansion for the s</p>															
Simulation summaries discuss conclusions reached from examining model runs. We use hyperlinks to other documents and unnumbered lists. Small images are often placed in-line within an HTML document near related text whereas larger images are referenced instead through hyperlinks. HTML tables prove useful here in organizing many related hyperlinked images.	 <p>Discussion: The qv perturbation is mc advection (behind the front) there. The q</p>															

³Ours may be found on our group home page:
<http://redrock.ncsa.uiuc.edu/AOS/home.html>

⁴This is done extensively at the National Severe
Storms Laboratory: <http://www.nssl.uoknor.edu/>

3. METHODS

We want to emphasize that putting our results online is quite simple. Once the necessary tools are acquired, the process from data-in-hand to formatted-documents-online is fairly rapid. How we carry this out is described below. Depending on the computer you are using, different editors and conversion programs can be used.

Familiarity with HTML⁵ is required to put documents on the web. Introductory documents already exist online⁶, and learning basic HTML takes only an afternoon. You must have a client⁷ (web browser) and access to a Web server⁸. Note that basic HTML can be created with a simple text editor.

Plain text can be put on the web immediately. To place model text output online, we make a copy of the text output file, edit it with a Unix text editor to save only the desired lines, and name the resulting file "anyfilename.txt". The .txt suffix identifies the file as plain text without formatting - and any web server will be able to make this text available without any further effort.

HTML-formatted text needs slightly more work. A plain-text file can be made into a formatted HTML file by, at a minimum, entering a title and a heading and renaming the file to have a suffix of .html rather than .txt. Titles appear as names of the document window. Headers help organize the text. Unnumbered lists () are an easy way to list items. Paragraphs are separated with a <p>. An example follows, where <> denotes a HTML "tag", and </> the ending tag:

```
<TITLE>Model simulation X</TITLE>
<H3>Heading; H2=larger, H4 smaller</H3>
Text text text ... <p>
Parameters for this simulation:
<UL>
<LI>Unnumbered list - first item
<LI>Unnumbered list - second item
</UL>
```

⁵ Web terms: <http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Glossary/index.html>

⁶ See, e.g., <http://www.ncsa.uiuc.edu/General/Internet/WWW/HTMLPrimer.html>

⁷ See <http://www.w3.org/hypertext/WWW/Clients.html>

⁸ See <http://www.w3.org/hypertext/WWW/Servers.html>

This translates into the following:

```
Heading; H2=larger, H4 smaller
Text text text ...
Parameters for this simulation:
• Unnumbered list - first item
• Unnumbered list - second item
```

Links (or "anchors", <A>) refer to other files (text, HTML, images, etc.) from the current HTML file. Anchors may appear within a paragraph, a list, etc. HREF refers to the actual file location. This example defines 3 links to the image file "file5.gif", text file "info.txt" and HTML file "simB.html" with link names of "image", "summary" and "case B", shown underlined by the web browser. In the usual browser display, underlined text is clickable, allowing the user to "follow" the link.

```
<UL>
<LI>day5: <A HREF="file5.gif">image</A>
<LI>info: <A HREF="info.txt">summary</A>
<LI>also: <A HREF="simB.html">case B</A>
</UL>
```

Tables⁹ are a powerful feature of HTML. We use them extensively to organize links and data. The basic elements include: a caption, table headers (<th>), and table data. Table data (<td>) and headers are all parts of table records (<tr>). For example:

```
<table border>
<caption>Simulation results</caption>
<tr> <th>Data </th> <th>T=2days</th> </tr>
<tr> <td>Max W </td> <td>2 cm/s </td> </tr>
</table>
```

results in the following:

Simulation results	
Data	T=2days
Max W	2 cm/s

We use images widely, both in-line and referenced through links. Images placed online are first converted to GIF since it is widely used on the web. We convert images

⁹ See <http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Tables/tutorial.html>

to GIF using the Unix imtools¹⁰ package or the Macintosh program GraphicConverter¹¹. In the following example the image in "plot1.gif" will appear within the HTML document, while "plot2.gif" will appear as a link, displayed only when the users clicks on the text "click here":

In-line image:
External: click here

Finally, there are programs that will convert from word processor documents to HTML, such as the Macintosh program rftohtml¹². We use rftohtml as a first step, followed by further manual editing.

4. CONCLUSION

We find our desire to keep and review results of past and current numerical simulations to be well matched with capabilities of HTML to format and web browsers to display this information. Model data received frequently is often stored first online as text only, while other information is formatted and combined with image data and links as the need arises. The amount of time spent preparing data for availability over the web is proportional to the importance we attach to that data set. For simple text, this takes only a few minutes; to convert many images and format the data may take an hour or two. Simulations we find important and worthy of later detailed examination are cataloged extensively (Fig. 1).

This method of storage and retrieval has proven to be invaluable for comprehensive, easy point-and-click access to our data. While our motivation for this approach was entirely for our own archival and retrieval of our data, the same process makes the information available to others with no additional effort.

We have begun storing 2D animations online, and look forward to use of 3D data tools. In the meantime, HTML and the web are allowing us to make the most of our simulations by keeping the most salient information readily available.

5. ACKNOWLEDGMENTS

Our modeling research is being supported by NSF through ATM 92-14098 and with computing and other support from NCSA.

¹⁰See <ftp://ftp.sdsc.edu/pub/sdsc/graphics/imtools>

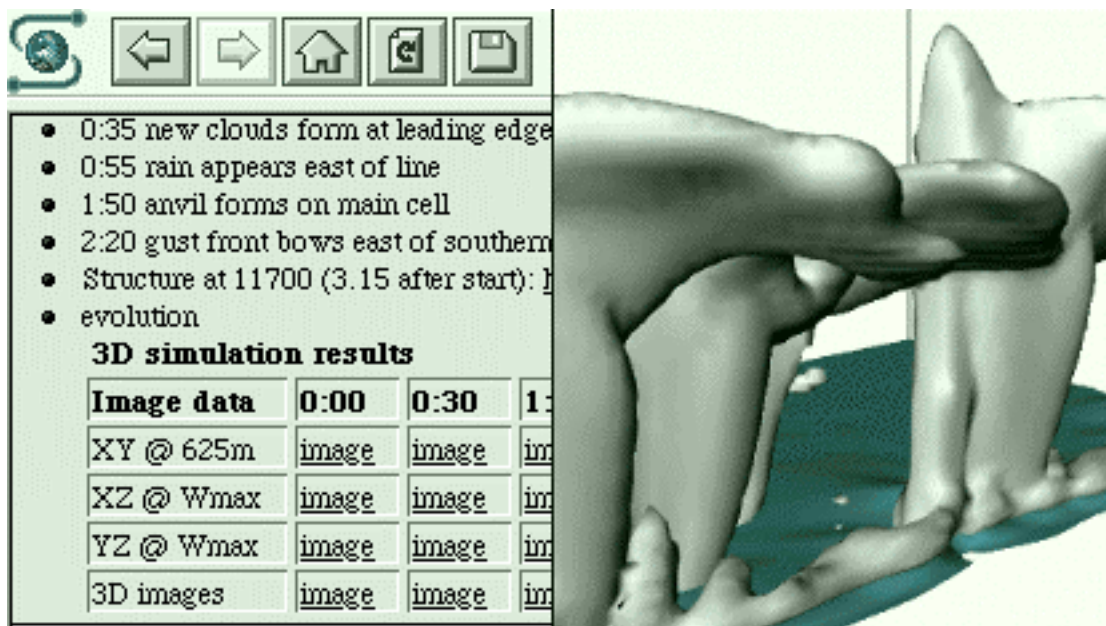


Figure 1: A recent simulation viewed via the web. Here three dimensional rendered surfaces of potential temperature and (cloud+rain)water are combined with text information in forming part of our permanent record of a model experiment.

¹¹Available through <http://wwwhost.ots.utexas.edu/mac/pub-mac-graphics.html>

¹²Available through <http://wwwhost.ots.utexas.edu/mac/internet-www.html>