ATM 693 Analysis Methods in Meteorology and Climate (3 Credits), Spring 2011

- 10-11 MWF, with some make up classes due to travel
- Is this an OK time?
- Go through Syllabus, other topics of interest?
- Helpful to know UNIX, NCL, GrADS, & FORTRAN.
- Also helpful to know C++ & Python

Today we begin with

Chapter 1 Data Storage Methods

1.1 Common Formats

1.2 Less Common Formats

1.3 I/O Tools

1.1 Common Formats

- Ascii, here
- Binary, here
- GRIB, here
- NetCDF, here
- HDF, self-describing format

Number Representation on Computers

- BIT - Binary digIT, ones and zeros (basic unit of information on a computer
- 1011 Binary number
  (1*2^3)+ (0*2^2) + (1*2^1) + (1*2^0) = 8+0+2+1 = 11 (base 10 numbers)
- 8 bits = 1 byte
- 8 bits represent 256 values, 0 to 255
  0=00000000, 1=00000001, 2=00000010, ...
  254=11111110, 255=11111111
- ASCII 0==> 127, each character is 7 bits

ASCII

- American Standard Code for Information Interchange is a character-encoding scheme based on the ordering of the English alphabet
- Read with a text editor, easily moved around (BEWARE a few things... Mac returns). Do Example
- Not very efficient way to store large data sets.

GrADS does not work easily with ASCII while NCL does.

http://en.wikipedia.org/wiki/ASCII

Binary

- Data encoded to store on a computer. Many different kinds of binary, thus complication.
- **Ask for code to read binary data.**
- Direct access, sequential access FORTRAN open statement will specify.
- Big endian and little endian, refers to byte order.

FORTRAN, GrADS, & NCL work.

Little and Big Endian Issues

Little-endian

[edi

Atomic element size 8-bit, address increment 1-byte (octet)

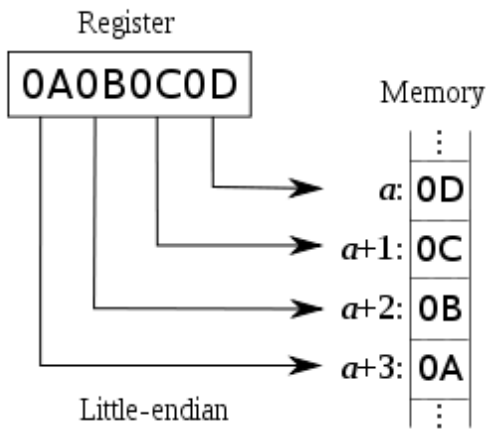
[edit]

increasing addresses →

...	$0D_h$	$0C_h$	$0B_h$	$0A_h$	...

The least significant byte (*LSB*) value, $0D_h$, is at the lowest address. The other bytes follow in increasing order of significance.

[edit]



Big-endian

[e

Atomic element size 8-bit, address increment 1-byte (octet)

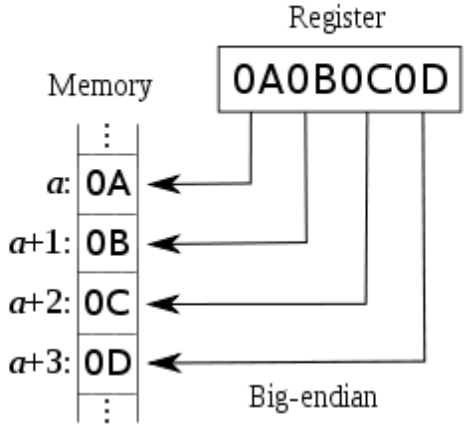
[edit]

increasing addresses →

...	$0A_h$	$0B_h$	$0C_h$	$0D_h$	...

The most significant byte (*MSB*) value, which is $0A_h$ in our example, is stored at the memory location with the lowest address, the next byte value in significance, $0B_h$, is stored at the following memory location and so on. This is akin to Left-to-Right reading in hexadecimal order.

[edit]



Fortran solution, Grads and NCL solutions exist

```
OPEN(unit,CONVERT='BIG_ENDIAN',...)
```

or

```
OPEN(unit,CONVERT='LITTLE_ENDIAN',...)
```

Intel CPUs are little-endian, while Motorola 680x0 CPUs are big-endian.

GRIB

- GRIB (GRIdded Binary) is a mathematically concise data format commonly used in meteorology to store historical and forecast weather data.
- Historical comments
- Each individual GRIB record has two components - the part that describes the record (the header), and the actual binary data itself.

GrADS & NCL work.

NetCDF

- (Network Common Data Form) is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data.
- Hosted by UNIDATA program at UCAR (originated in the NASA CDF format)

FORTRAN, GrADS & NCL work plus R (ncdf [2] and ncvar packages), Perl, Python, Ruby, Matlab, IDL, and Octave.

http://en.wikipedia.org/wiki/NetCDF

1.2 Less Common Formats

- HRIT/LRIT, HRPT/LRPT (satellite data formats, will not talk about here)
- BUFR here
- METAR - format of encoded hourly reports from Airports & permanent weather stations. here

METAR

Python, PERL for parsing data

North American METAR codes [edit]

North American METARs deviate from the WMO (who write the code on behalf of ICAO) FM 15-XII code. Details are listed in the FAA's Aeronautical Information Manual (AIM), but the non-compliant elements are mostly based on the use of non-standard units of measurement. This METAR example is from Trenton-Mercer Airport near Trenton, New Jersey, and was taken on 5 December 2003 at 18:53 UTC.

METAR KTTN 051853Z 04011KT 1/2SM VCTS SN FZFG BKN003 OVC010 M02/M02 A3006 RMK AO2 TSB40 SLP176 P0002 T10171017=[9]

- **METAR** indicates that the following is a standard hourly observation.
- **KTTN** is the ICAO identifier for the Trenton-Mercer Airport.
- **051853Z** indicates the day of the month is the fifth and the time of day is 1853 Zulu/UTC, 6:53PM GMT, or 1:53PM Eastern Standard Time.
- **04011KT** indicates the wind is from 040° true (north east) at 11 KT (20 km/h; 13 mph). In the United States, the wind direction must have a 60° or greater variance for variable wind direction to be reported and the wind speed must be 7 KT (13 km/h; 8 mph) or higher.
- **1/2SM** indicates the prevailing visibility is 0.5 SM (800 m).
- **VCTS** indicates a thunderstorm in the vicinity (within 10SM (16 km), but beyond 5SM (8 km)).
- **SN** indicates snow is falling at a moderate intensity. If any precipitation begins with a minus or plus (-/+), it's either light or heavy.
- **FZFG** indicates the presence of freezing fog.
- **BKN003 OVC010** indicates a broken cloud layer at 300 ft (91 m) above ground level and a solid or overcast layer at 1,000 ft (300 m). The last two zeros are omitted.
- **M02/M02** indicates the temperature is –2 °C (28 °F) and the dewpoint is –2 °C (28 °F). An M in front of the number indicates that the temperature/dew point is below zero (0) Celsius.
- **A3006** indicates the altimeter setting is 30.06 inHg (1,018 hPa).

Note that what follows are not part of standard observations outside of the United States and can vary significantly.

- **RMK** indicates the remarks section follows.
- **AO2** indicates that the station is automated with a rain/snow precipitation sensor. Stations that aren't equipped with a rain/snow sensor are designated **AO1**.[10]
- **TSB40** indicates the thunderstorm began at 40 minutes past the hour at 1840 Zulu/UTC, 6:40PM GMT, or 1:40PM Eastern Standard Time.
- **SLP176** indicates the current barometric pressure extrapolated to sea level is 1017.6 hPa.
- **P0002** indicates that 0.02 inches (0.5 mm) of liquid-equivalent precipitation accumulated during the last hour.
- **T10171017** is a breakdown of the temperature and dew point in 8 digits separated into two groups of four. the first four digits (1017) is the temperature. The first digit (1) designates above or below zero Celsius (0=above zero 1=below zero). The next three digits in the group '017' is the temperature in degrees and tenths of a degree, –1.7 °C (29 °F). The last four digits '1017' is the same as the first group but for dew point, –1.7 °C (29 °F).
- **=** indicates the end of the METAR report.

http://en.wikipedia.org/wiki/METAR

BUFR

- BUFR (Binary Universal Form for the Representation of meteorological data) is a World Meteorological Organization (WMO) standard for storing observational data (aka sequence or in-situ data).
- BUFR is self-describing data format and can store a large amount of data and metadata in a small amount of disk space by using look-up tables and bit-by-bit packing.

GrADS reads in this type of data. NCL does not at present.

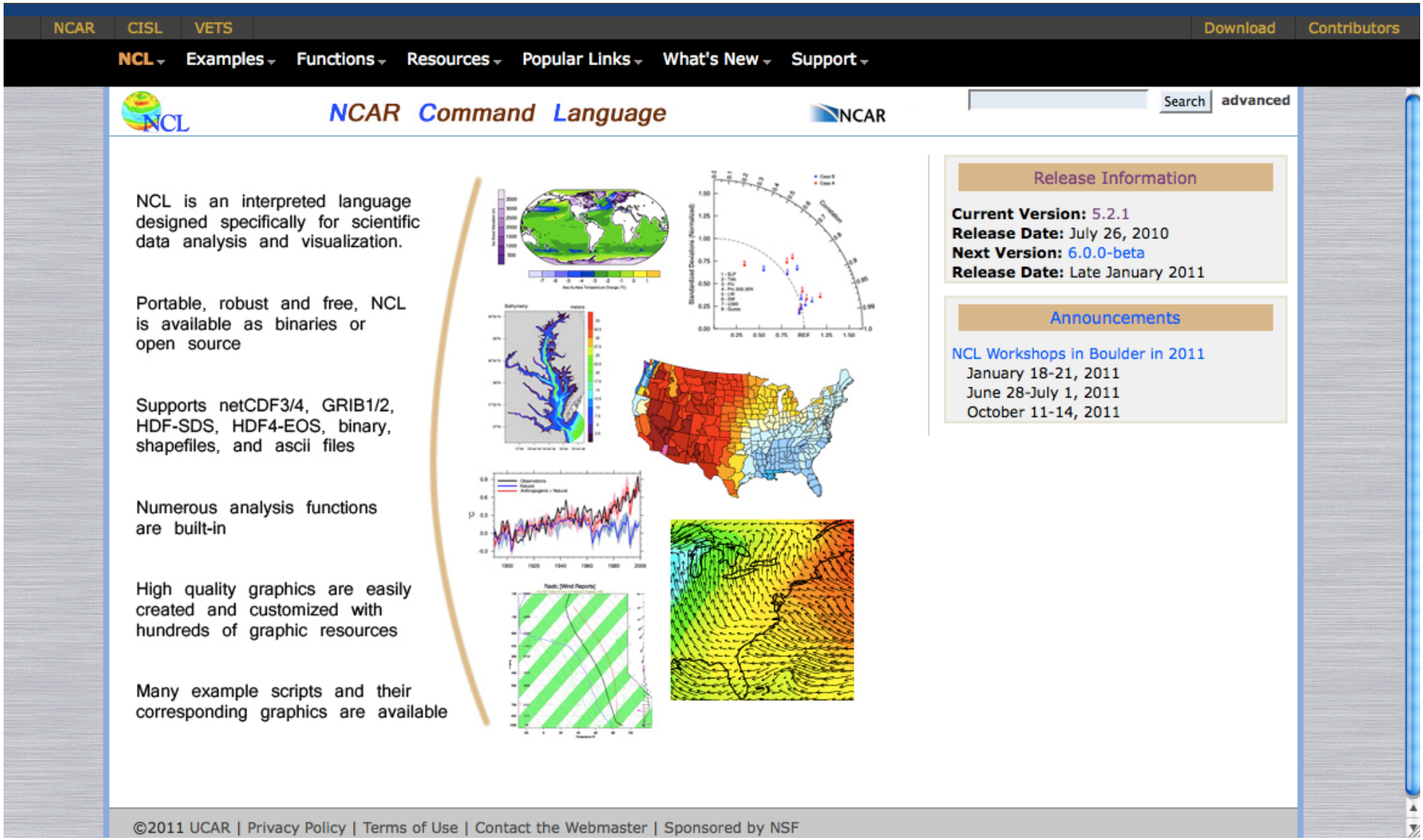
http://grads.iges.org/grads/gadoc/bufrformat.html

1.3 I/O Tools

- Common
  - NCL, here
  - Grads, here (Compare and contrast)
  - NCO, here
  - CDO, Climate Data operators -command line, https://code.zmaw.de/projects/cdo/
- Less Common
  - McIdas (satellite data viewer and analyzer, U. Wisconsin SSEC)
  - OpenDAPS (framework for scientific data networking)
  - XML (is a set of rules for encoding documents in machine-readable form, ex: RSS, XHTML, facilitates sharing of data)
- GIS - real strength is in spatial analysis! Example

NCL

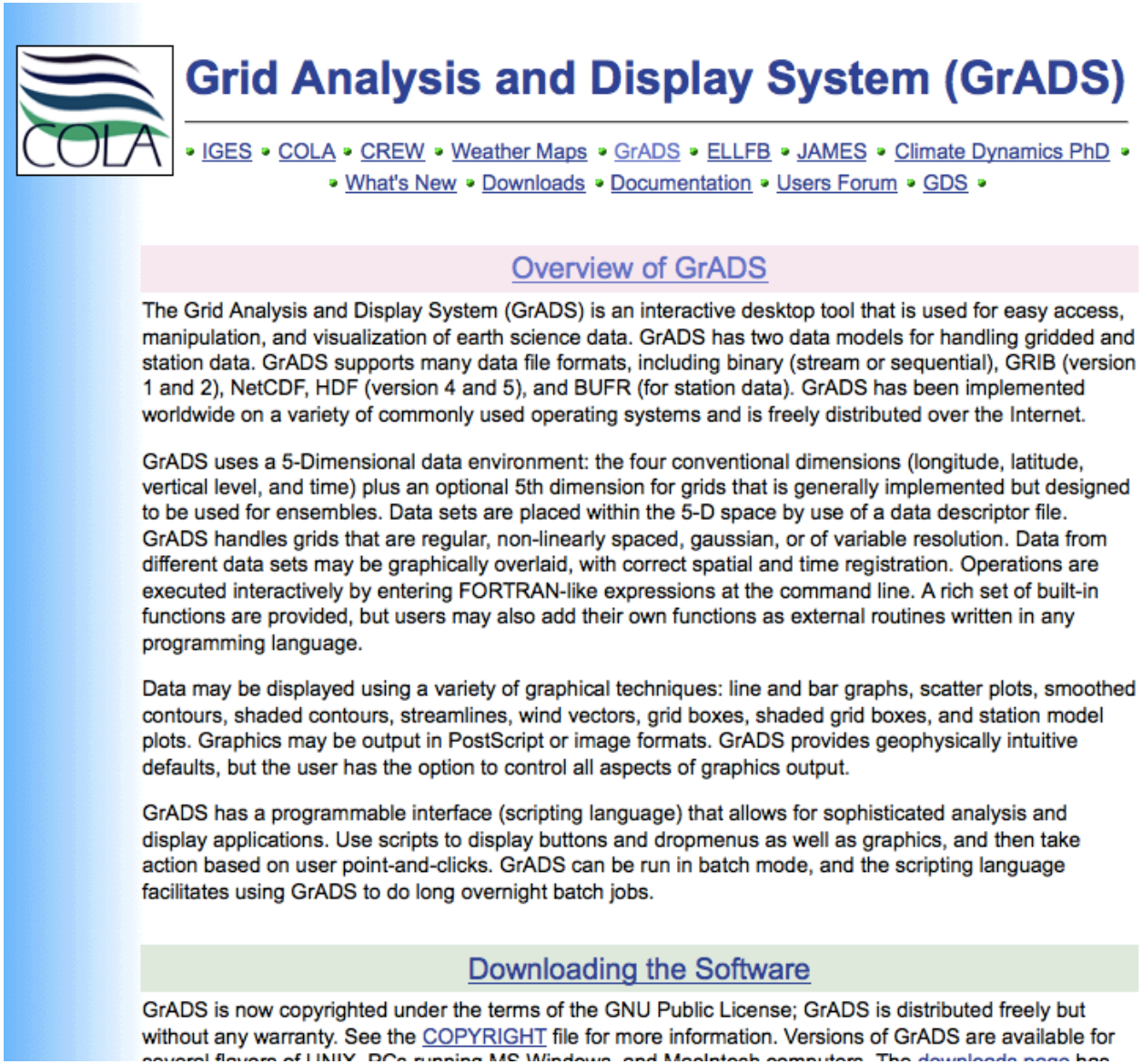
http://www.ncl.ucar.edu/



Data analysis and visualization tool, Example

GrADS

http://grads.iges.org/grads/



Grid Analysis and Display System (GrADS)

• IGES • COLA • CREW • Weather Maps • GrADS • ELLFB • JAMES • Climate Dynamics PhD •
• What's New • Downloads • Documentation • Users Forum • GDS •

Overview of GrADS

The Grid Analysis and Display System (GrADS) is an interactive desktop tool that is used for easy access, manipulation, and visualization of earth science data. GrADS has two data models for handling gridded and station data. GrADS supports many data file formats, including binary (stream or sequential), GRIB (version 1 and 2), NetCDF, HDF (version 4 and 5), and BUFR (for station data). GrADS has been implemented worldwide on a variety of commonly used operating systems and is freely distributed over the Internet.

GrADS uses a 5-Dimensional data environment: the four conventional dimensions (longitude, latitude, vertical level, and time) plus an optional 5th dimension for grids that is generally implemented but designed to be used for ensembles. Data sets are placed within the 5-D space by use of a data descriptor file. GrADS handles grids that are regular, non-linearly spaced, gaussian, or of variable resolution. Data from different data sets may be graphically overlaid, with correct spatial and time registration. Operations are executed interactively by entering FORTRAN-like expressions at the command line. A rich set of built-in functions are provided, but users may also add their own functions as external routines written in any programming language.

Data may be displayed using a variety of graphical techniques: line and bar graphs, scatter plots, smoothed contours, shaded contours, streamlines, wind vectors, grid boxes, shaded grid boxes, and station model plots. Graphics may be output in PostScript or image formats. GrADS provides geophysically intuitive defaults, but the user has the option to control all aspects of graphics output.

GrADS has a programmable interface (scripting language) that allows for sophisticated analysis and display applications. Use scripts to display buttons and dropmenus as well as graphics, and then take action based on user point-and-clicks. GrADS can be run in batch mode, and the scripting language facilitates using GrADS to do long overnight batch jobs.

Downloading the Software

GrADS is now copyrighted under the terms of the GNU Public License; GrADS is distributed freely but without any warranty. See the COPYRIGHT file for more information. Versions of GrADS are available for

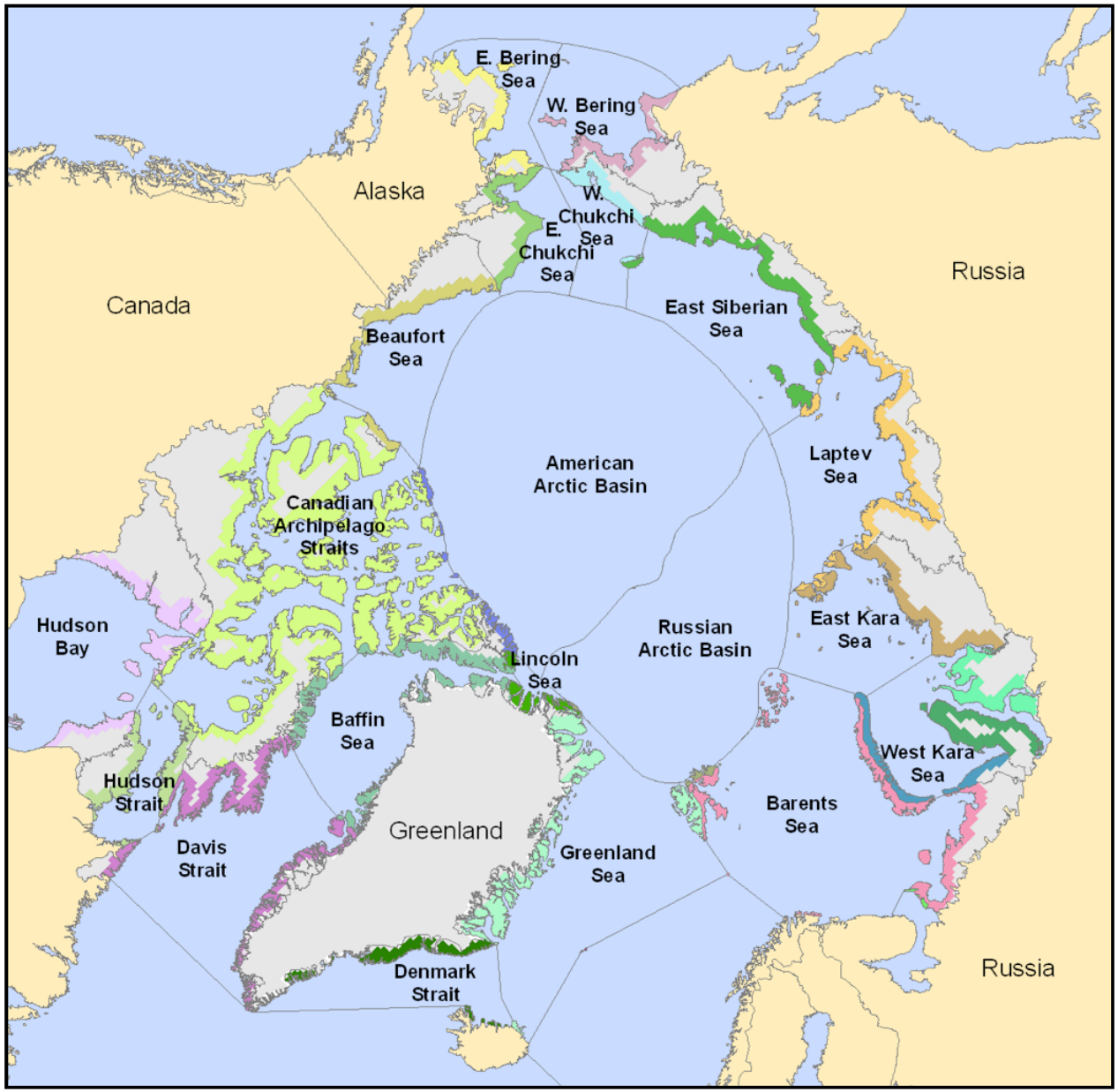
History, costs
Example

NCO

- NetCdf Operators - is of UNIX commands (programs) designed to simplify manipulation and analysis of self-describing data stored in the netCDF format.
- averaging data, math/statistics, extract variables from files, concatenate data, ensemble averaging
- VERY FAST!

http://nco.sourceforge.net/

Usefulness of GIS



Bhatt et al. 2010