Working with Forecast Models in PyFerret and Ferret

June 2015
A plot to illustrate the data output by a Forecast Model Run Collection (FMRC). Every colored cell is a model snapshot.

Every 2 months, another run

Each run, 12 monthly snapshots
Each model output (starting at one initialization time) is a “dataset”. How to aggregate these many individual datasets into a single dataset of higher dimensionality?

• use the Ferret command
  \texttt{DEFINE AGGREGATION/F}
  or
• serve the collection via OPeNDAP using Unidata’s THREDDS Data Server (TDS)

many 4D files $\rightarrow$ become a single 5D dataset
Native (“compact”) form of FMRC as run ...
In the native, compact form the time step values are a 2D variable. Note that the same forecasted date/time occurs at lags in multiple forecast runs.

<table>
<thead>
<tr>
<th>Model Run Time Axis</th>
<th>Forecast Series Axis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RUN 1</strong></td>
<td><strong>RUN 2</strong></td>
</tr>
<tr>
<td>1</td>
<td>744.</td>
</tr>
<tr>
<td>2</td>
<td>1440.</td>
</tr>
<tr>
<td>3</td>
<td>2184.</td>
</tr>
<tr>
<td>4</td>
<td>2904.</td>
</tr>
<tr>
<td>5</td>
<td>3648.</td>
</tr>
<tr>
<td>6</td>
<td>4368.</td>
</tr>
<tr>
<td>7</td>
<td>5112.</td>
</tr>
<tr>
<td>8</td>
<td>5856.</td>
</tr>
<tr>
<td>9</td>
<td>6576.</td>
</tr>
<tr>
<td>10</td>
<td>7320.</td>
</tr>
<tr>
<td>11</td>
<td>8040.</td>
</tr>
<tr>
<td>12</td>
<td>8784.</td>
</tr>
</tbody>
</table>
netCDF CDL markup language showing how ‘auxiliary coordinates’ are defined in a CF file

```c
dimensions:
    ntime = 10
    lev = 15; // number of levels
float temp(ntime, lev);
    temp: long_name = "temperature";
    temp: coordinates = "depth";
float depth(lev);
    depth: units = "meters";
    depth: axis = "Z";
```

Depth in a layered model

```
2D times in an FMRC
```

```
dimensions:
    ntime = 10
    nrun = 15; // number of runs
float temp(nrun, ntime);
    temp: long_name = "temperature";
    temp: coordinates = "times";
float times(nrun, ntime);
    times: units = "days since 01-01-1996";
```
Ferret syntax: how to regrid using an auxiliary coordinate

- **Depth:** \texttt{temp[ GZ(density) = My\_density\_axis]}
- **Time2d:** \texttt{temp[ GT(tf\_times)= TF\_CAL\_T @FMRC]}

Regridding in a manner analogous to a depth-to-density transformation, we convert the FMRC compact form into diagonal form.
Similarly also replace the Forecast time axis with a lead (lag) time axis

temp[GT(tf_times)=TF_CAL_T, GF(tf_times)=TF_LAG_F]
That’s the background.

Now, what does it feel like to analyze forecast model outputs using Ferret?
Start up the dataset as you would any other. “N” is the forecast index
> pyferret
yes? USE “http://server/path/myfiles/TAUX_fmrc.ncd” ! aggregation via TDS
yes? FILL tau_x[L=1,N=1]

How is the variability distributed?

I/X = longitude  
J/Y = latitude  
K/Z = depth  
L/T = time  
M/E = ensemble member  
N/F = forecast run
Another standard Ferret analysis -- take the time-standard deviation to visualize how variability is distributed.
SHADE/X=180/Y=0  \texttt{tau}_x^{\text{GT(times)}=\text{TF\_CAL\_T}}

View the forecast series at 180W, 0N in ‘diagonal’ form

Each run, 12 months

Every month, another run
forecast skill visualization: as we look towards the right, we see how long in advance we were able to forecast the variable.

Can we quantify how good our forecasts were?
LET tau_tf = tau_x[gt(times)=TF_CAL_T, gf(times)=TF_LAG_F]
LET tau_fe = tau_tf - tau_tf[N=1]
FILL/X=180/Y=0  tau_fe

Is there a season cycle to the errors?

We define and view the forecast error, tau_fe, relative to the N=1 initialization state.

Lets normalize this by the variability at each XY point.
LET tau_stddev = tau_tf[N=1,L=@std]
LET/TITLE=...  tau_nfe = tau_fe/tau_stddev
FILL/Y=180/Y=0/T=...  tau_nfe

We normalize the forecast error, dividing by its time-std dev at every location.

How long in advance were our forecasts “good”?
Let's say “good” == abs. val. of error within 0.5 std dev of final value

LET \( \tau_{\text{abs}} = \text{ABS}(\tau_{\text{nfe}}) \)

LET \( \tau_{\text{fcst lead}} = \tau_{\text{abs}}[F=@loc:.5] / 730 \)

LET/TITLE="…” \( \tau_{\text{skill}} = \text{MISSING}(\tau_{\text{fcst lead}},8) \) ! cap at 8 mo

PLOT/X=180/Y=0 \( \tau_{\text{skill}} \)

Lead time (months) for achieving “good” forecast error

What is the average lead time to get a “good” forecast?
The average lead time at which the forecast becomes accurate to 0.5 std dev at X=180, Y=0

LIST /X=180/Y=0 tau_skill[L=@ave]

VARIABLE : # of months lead achieving 0.5 std dev
FILENAME : TAUX_fmrc.ncd
LONGITUDE: 178.8E
LATITUDE : 1S
TIME : 17-JAN-1996 12:00 to 16-DEC-1999 00:00 (averaged)

4.165

Now lets see how our skill is distributed globally
Global distribution of “good” forecast -- lead time in months

# of months lead achieving 0.5 std dev