What to do about salinity when a Assimilating XBT data?

Results for estimating salinity in the Gulf of Mexico and in the Northwestern Atlantic Ocean

I’m Carlisle Thacker

and I approve of this message.

What to do when temperature is observed, but density is important?

- When XBT data are assimilated,
  salinity must be corrected along with temperature.
- Incorrect salinity causes incorrect density and currents.
- How to correct salinity without observing salinity?

Two regions as examples.

- Gulf of Mexico
  - Loop current, eddies.
  - Broad shelves with deep central basin.
  - River inflow.
  - Many bad data.
- Large North Atlantic region containing the Gulf Stream
  - Very large T and S variability.
  - Shelf in north but mostly deep ocean.
  - Gulf Stream inflow.
  - Few bad data.

Gulf of Mexico

3485 CTD stations – many redundant.

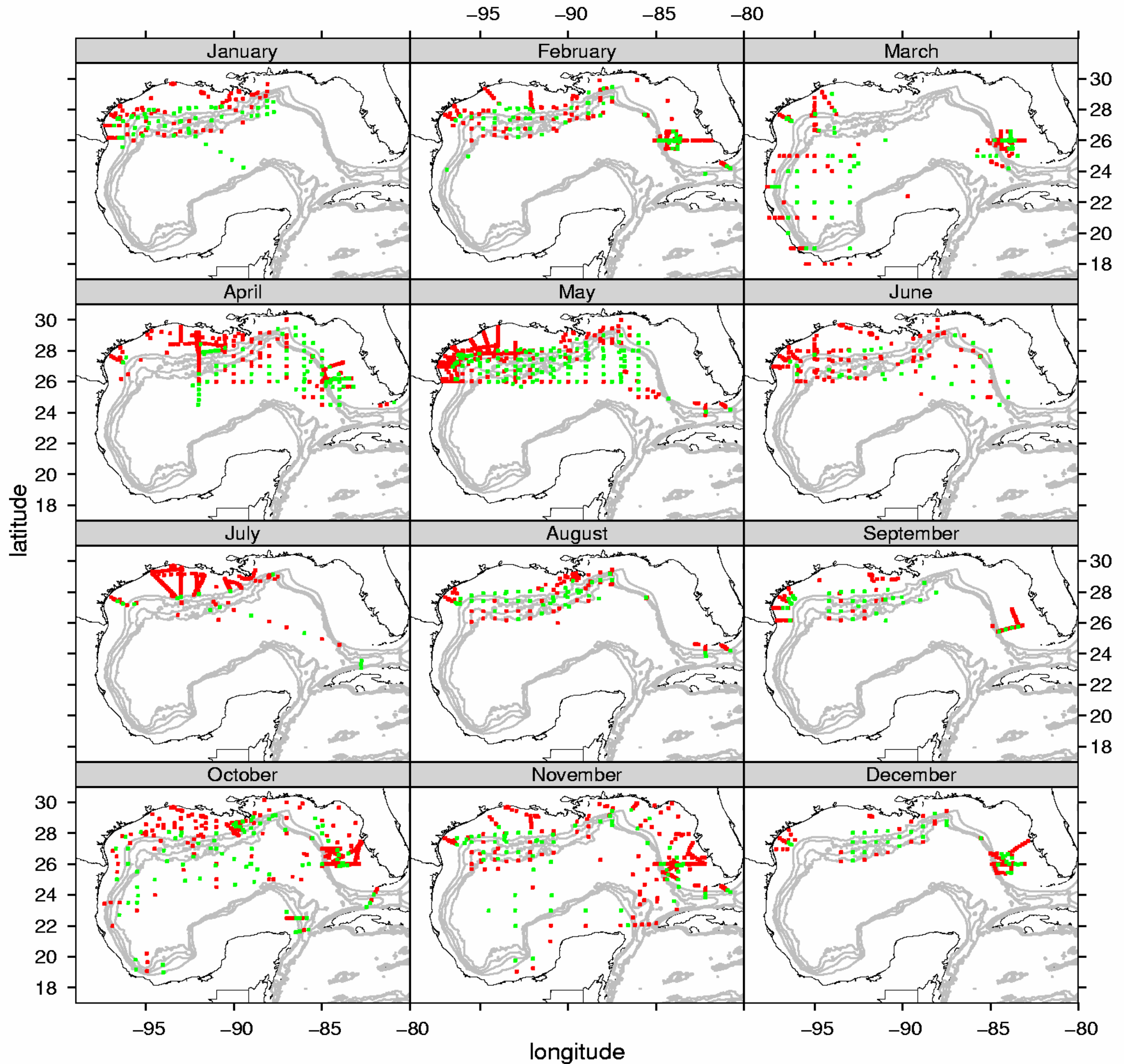
Most stations in shallow water.

Few in south.

No problem.

Sub-sampled to avoid near duplicates.

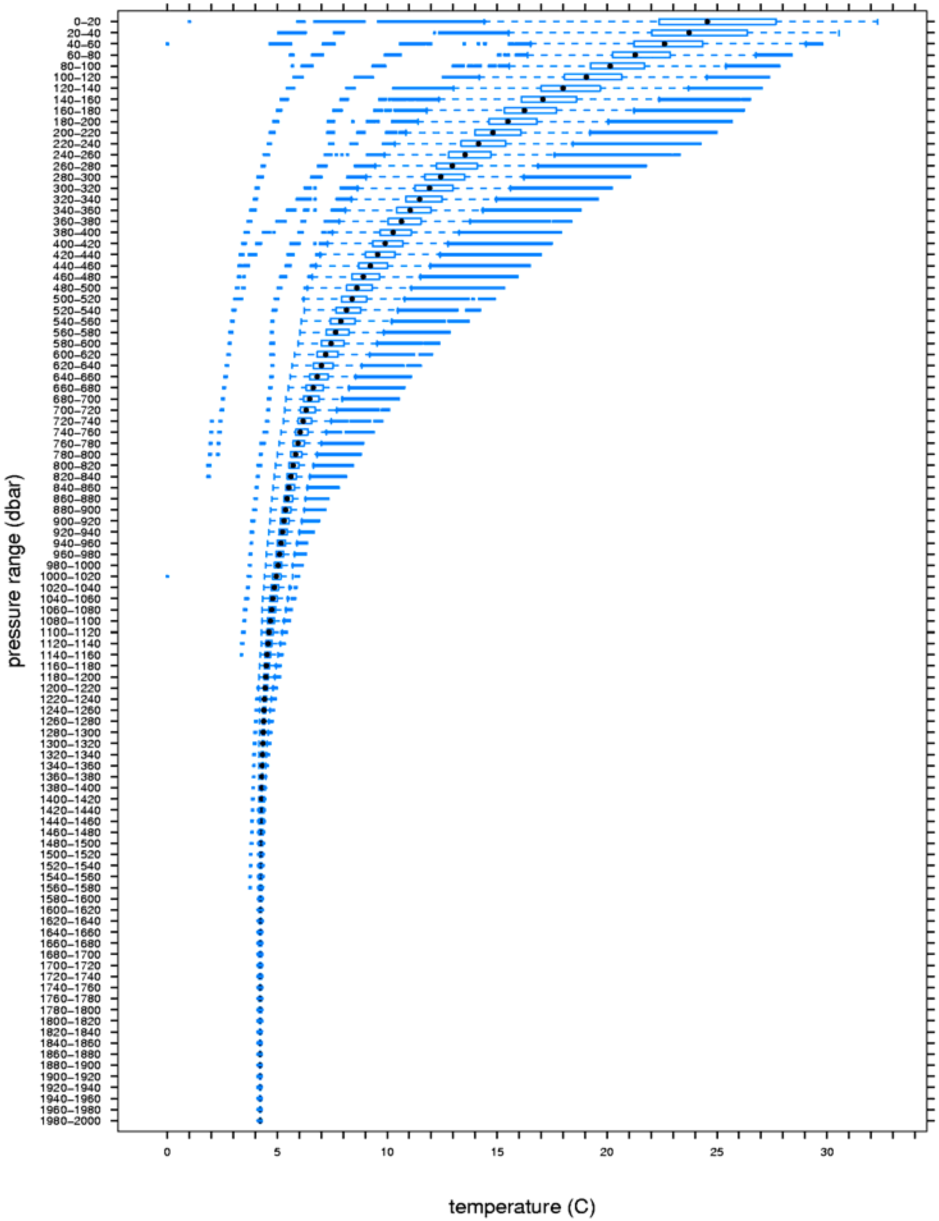
- 739 stations used.

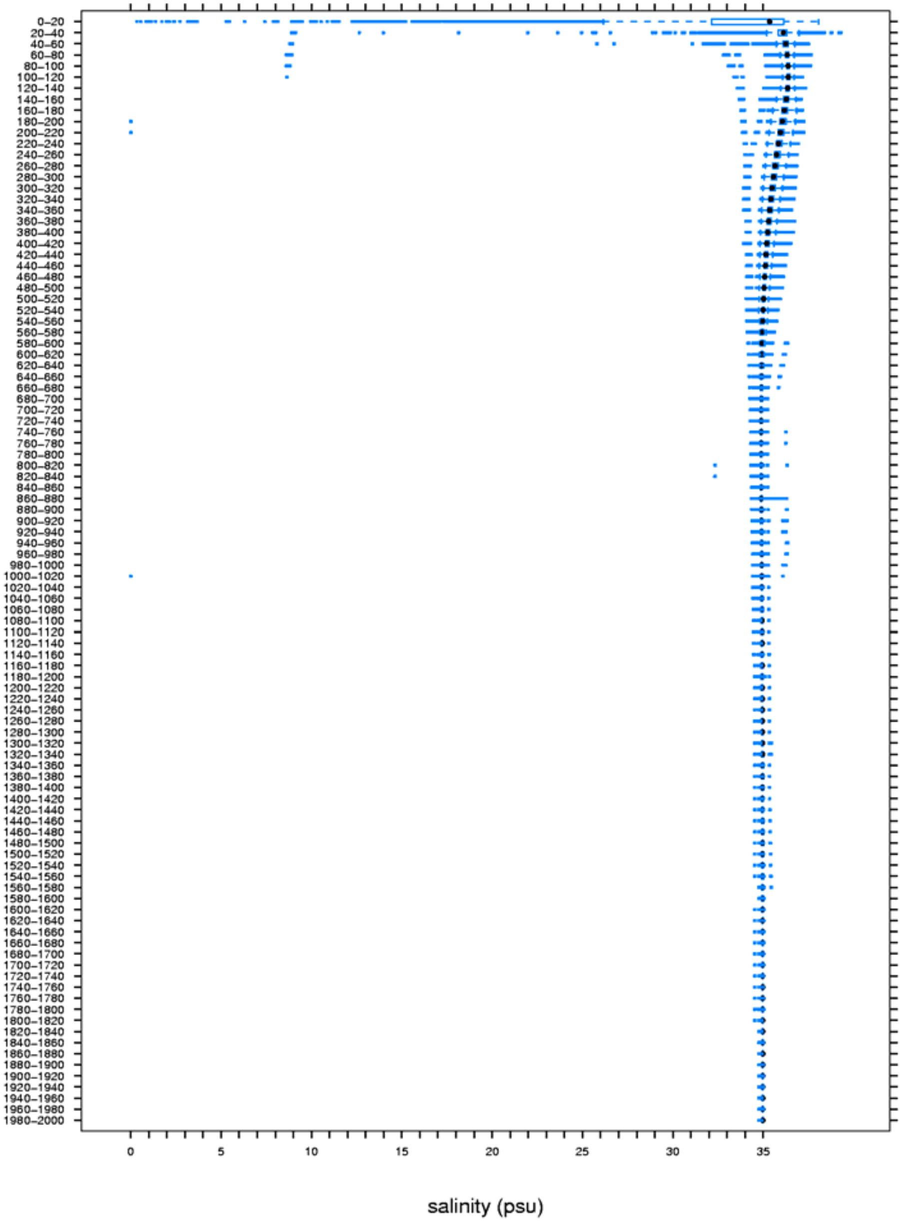


Problems with archived data:

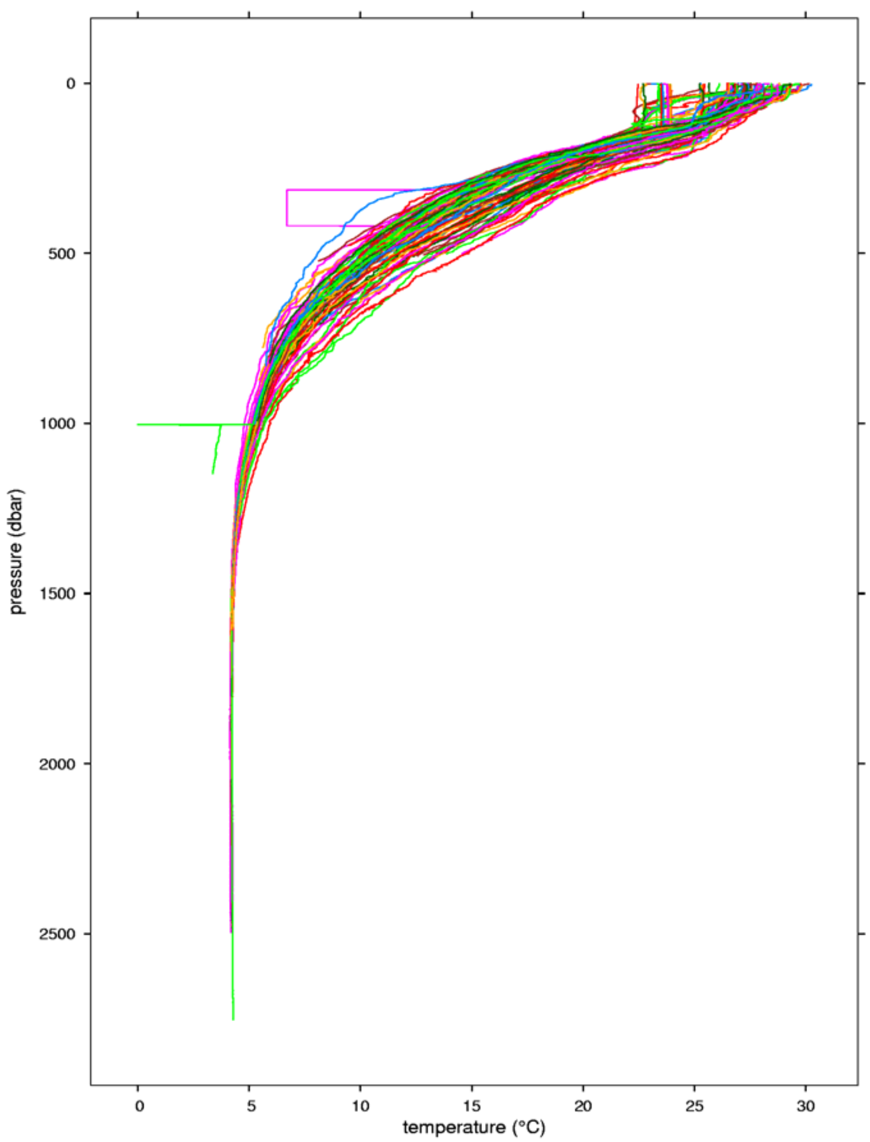
- Sampling is not uniform.
  - Local high-density sampling.
  - Few samples in south.
- Some data are bad.
  - Flags are not very helpful.
- Distributions are not Gaussian.
  - How to distinguish bad data from heavy tails?
  - Box and whisker plots are helpful.
  - TS plots also show outliers.

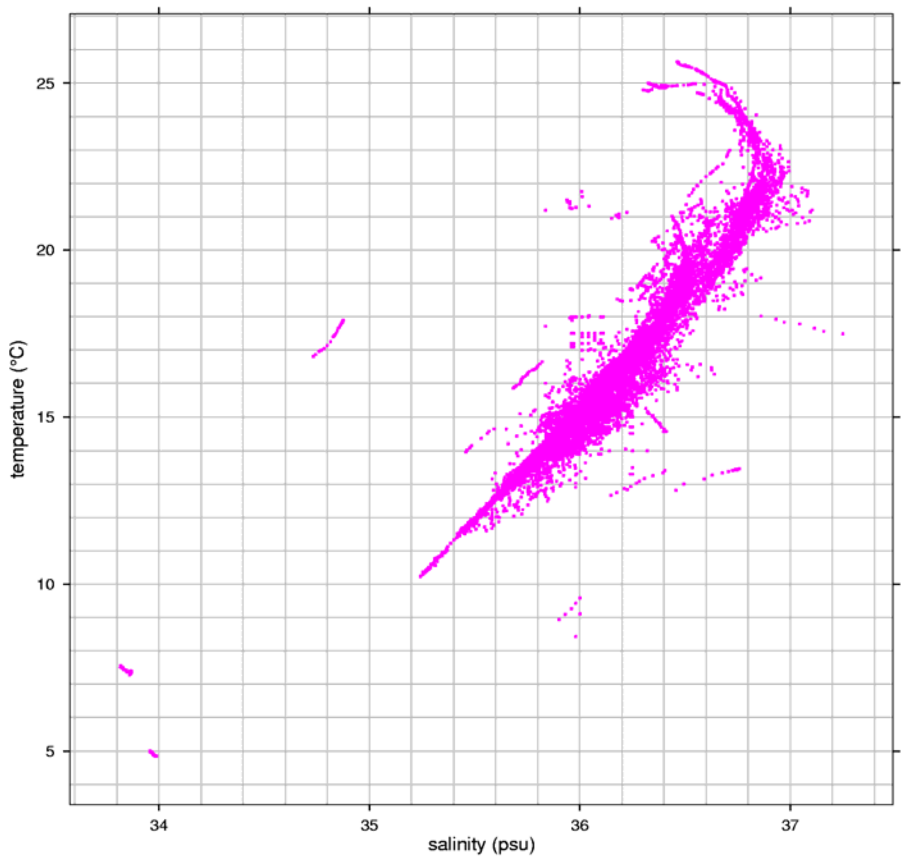
A first look at the 3489 CTD profiles for Gulf of Mexico.





Distributions of T and S data in 20 dbar intervals.

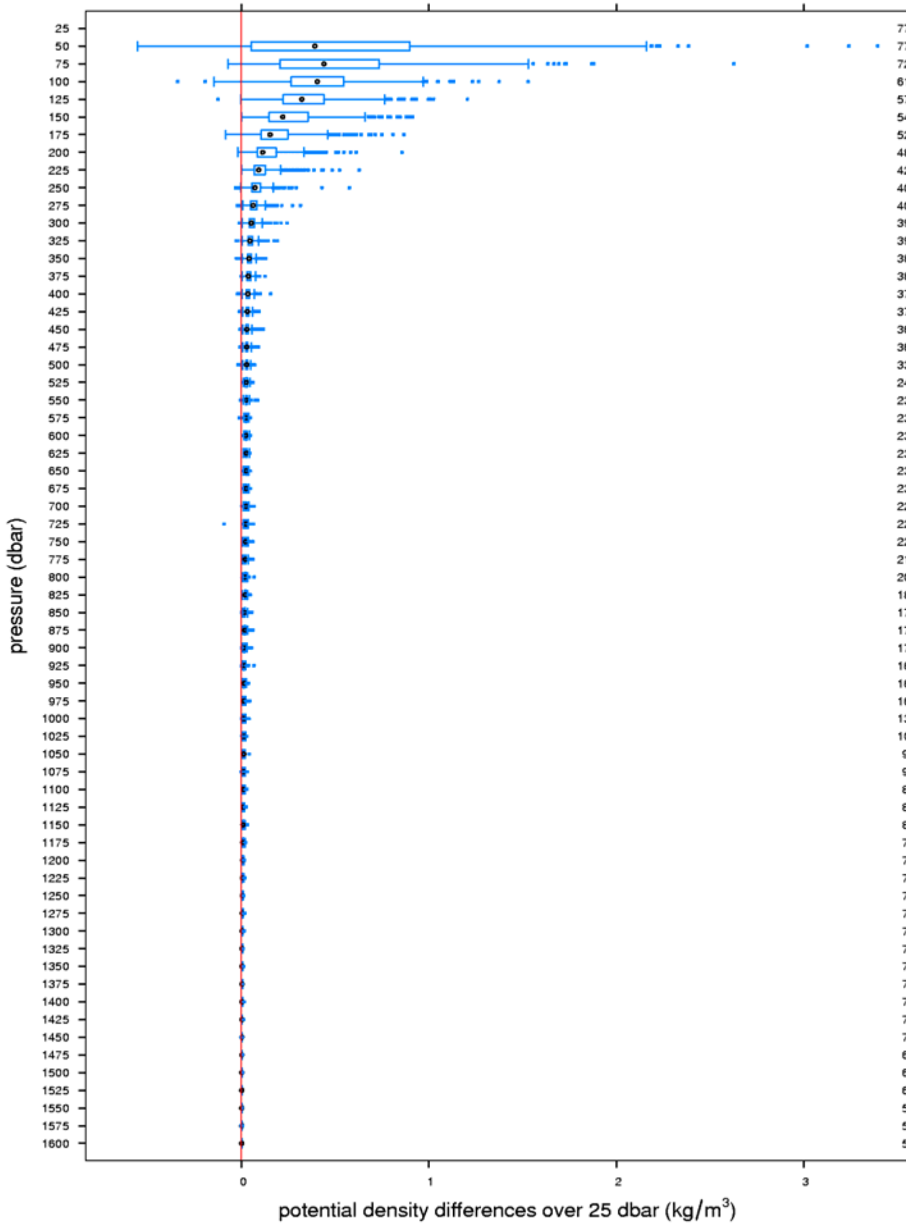


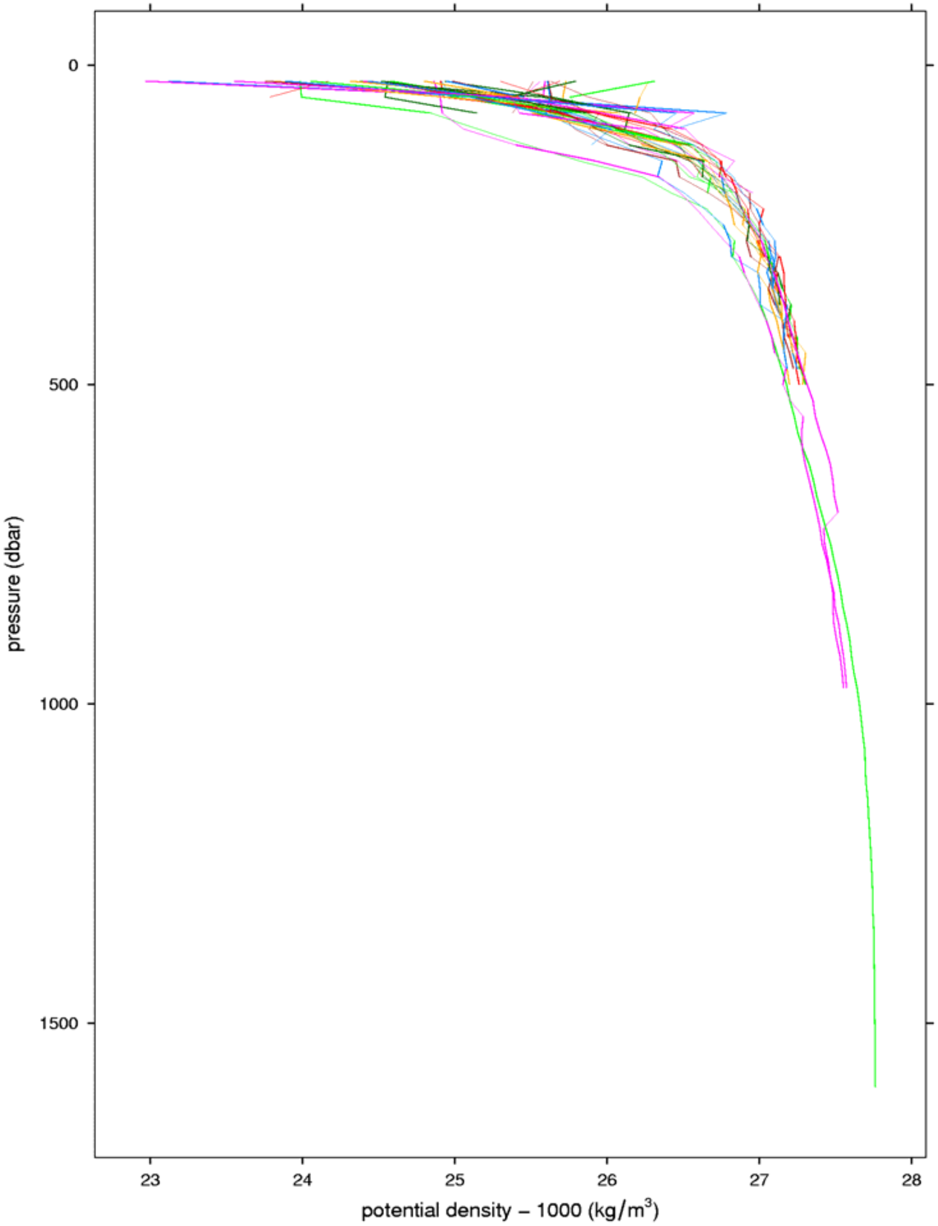


All data between 180 dbar and 200 dbar

Notice bad data!

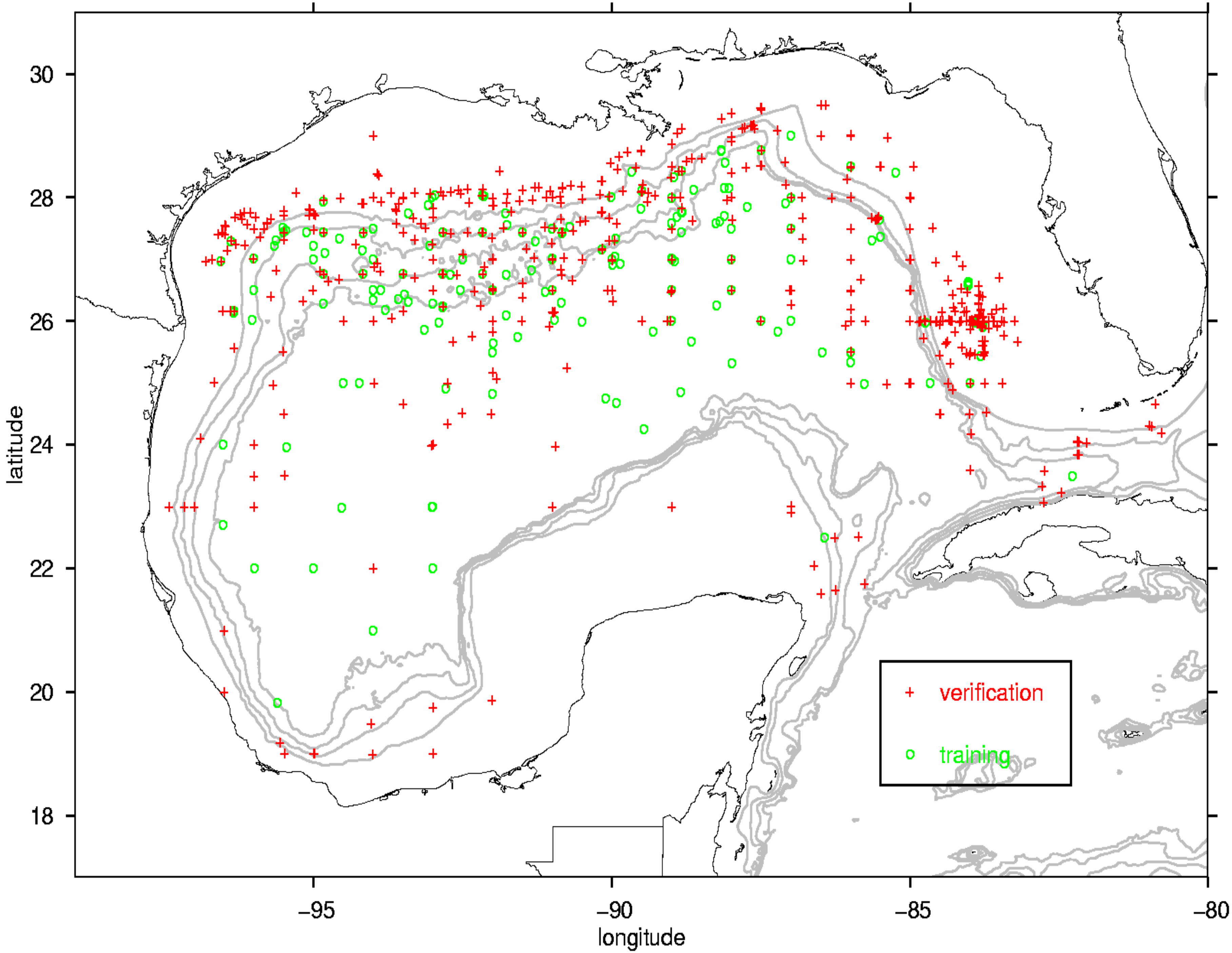
Warm outliers between 180 dbar and 200 dbar
Mostly good loop-current data.



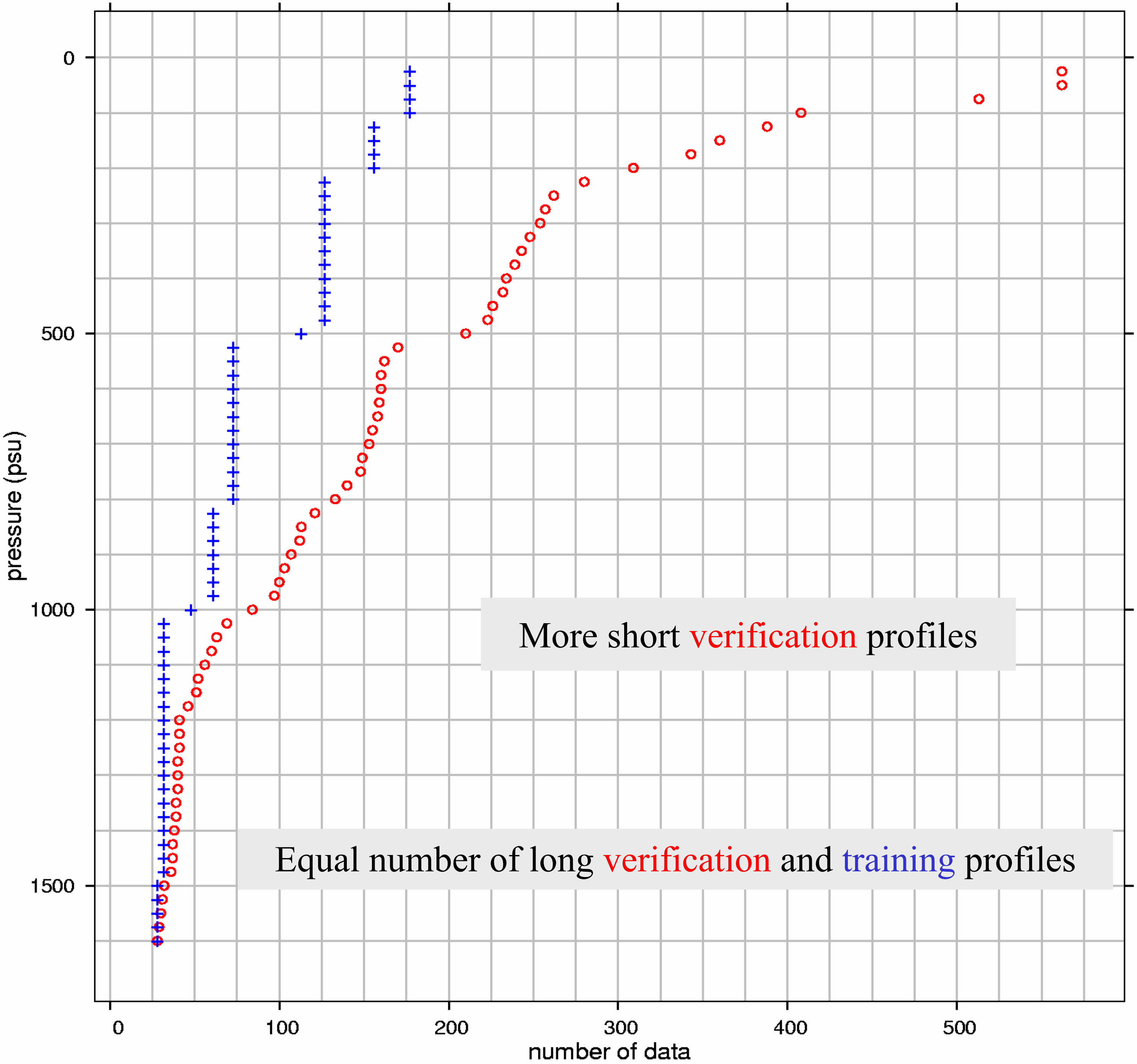


Some profiles have density inversions

37 with inversion greater than 0.01 kg/m^3



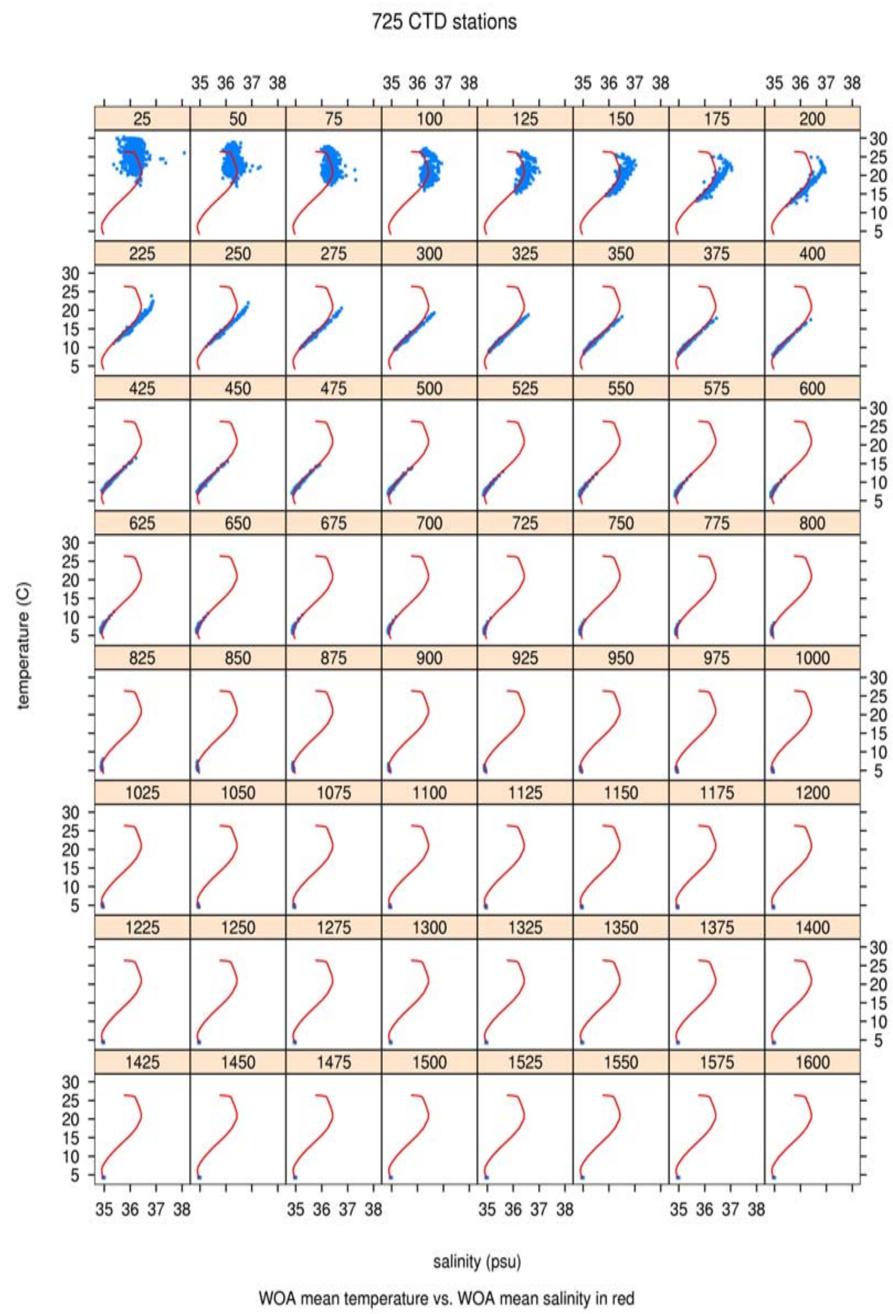
30
28
26
24
22
20
18
latitude
−95
−90
−85
−80
longitude
verification
training

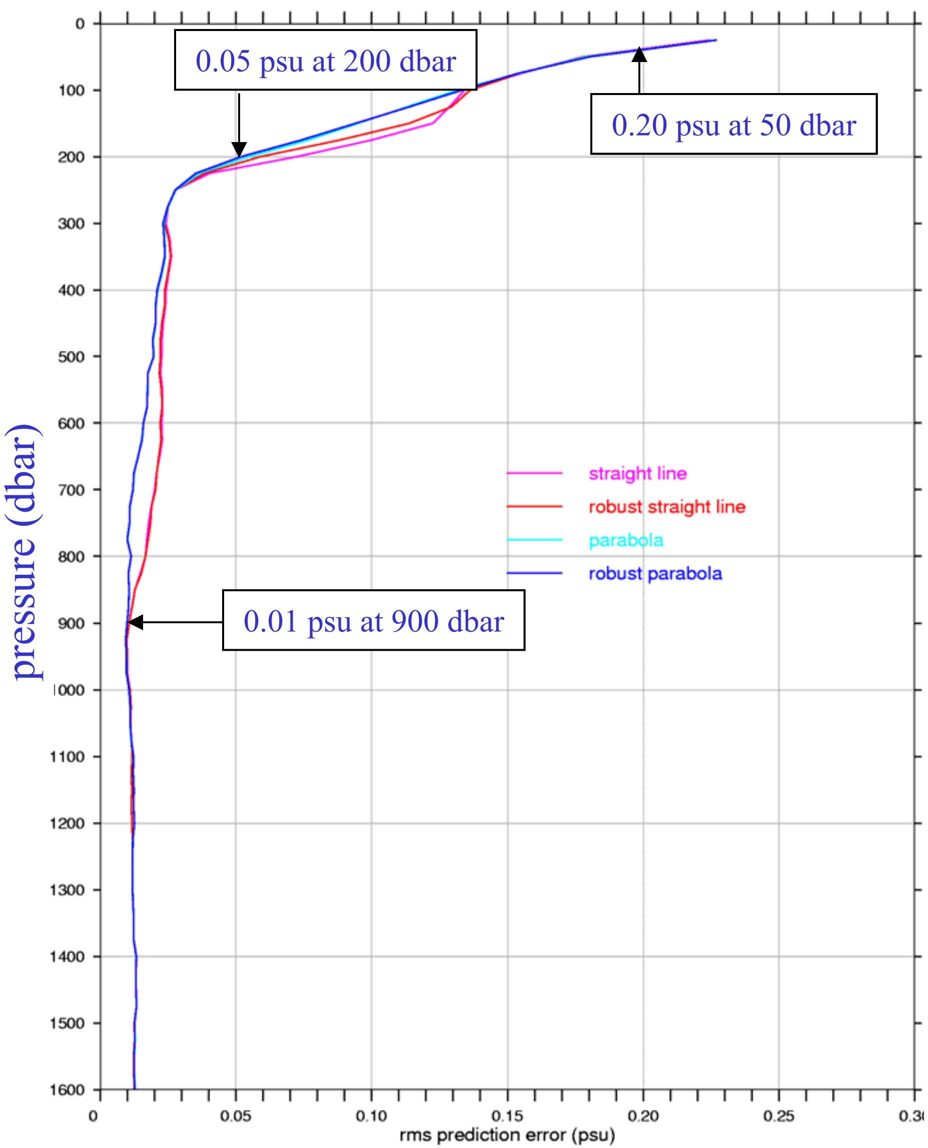


0
500
1000
1500
pressure (psu)
0
100
200
300
400
500
number of data
More short verification profiles
Equal number of long verification and training profiles

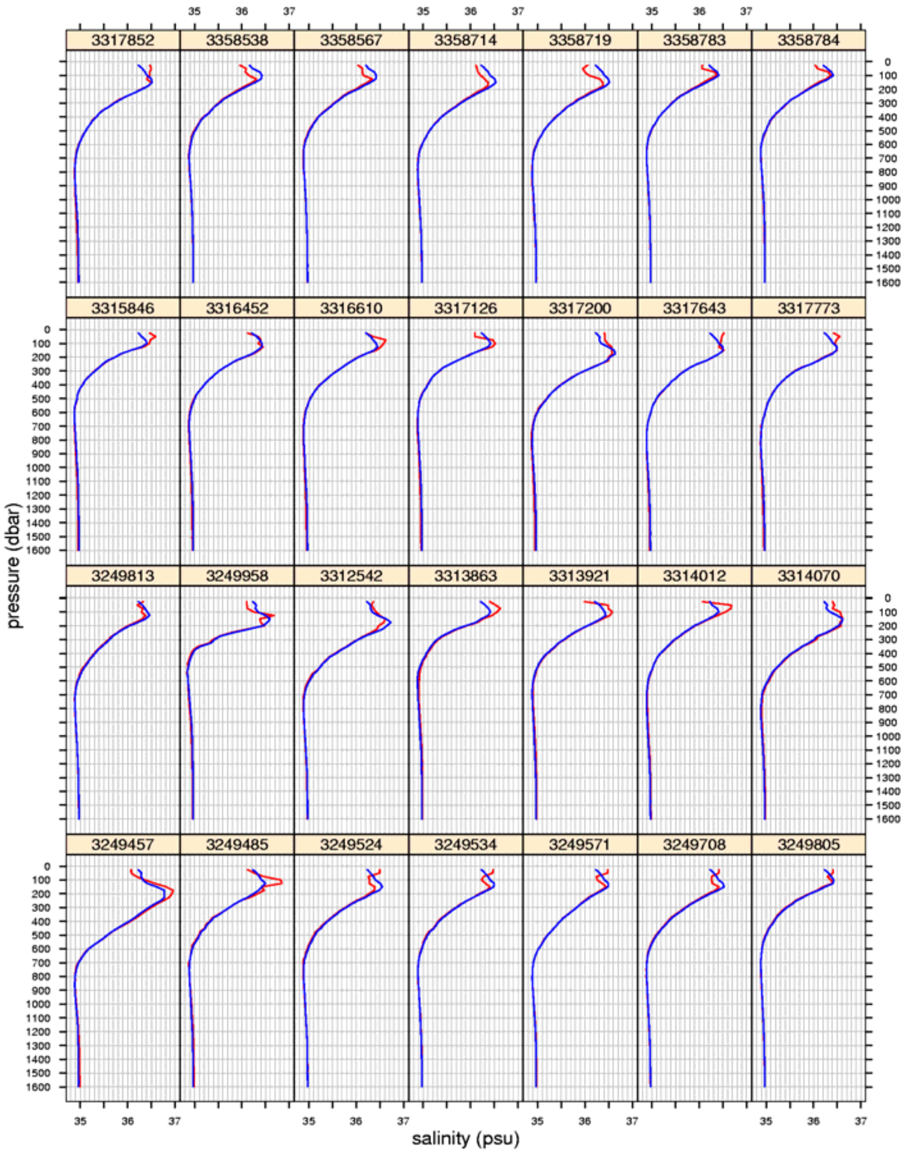
TS plots
training + verification data

- Data interpolated at 25 dbar pressure intervals.
- Mean T vs. mean S at all levels indicated in red on each panel.
- Warm-salty Loop Current values are not on mean TS curve.





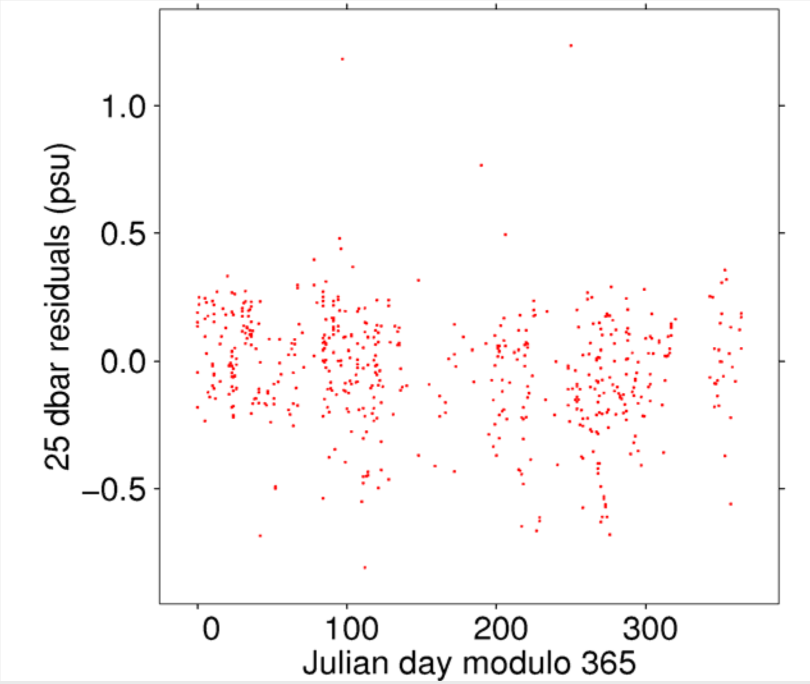
rms prediction error (psu)

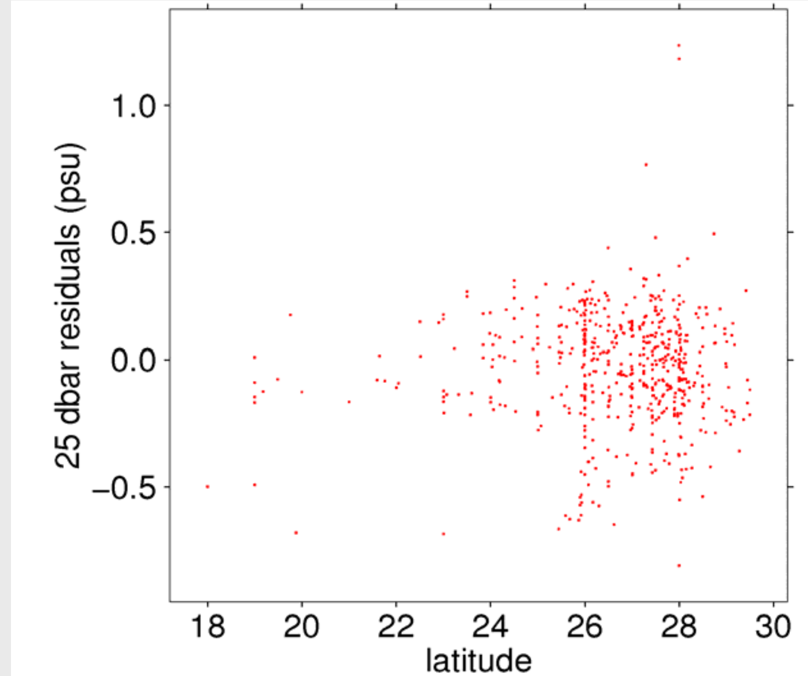


Skill explaining independent data

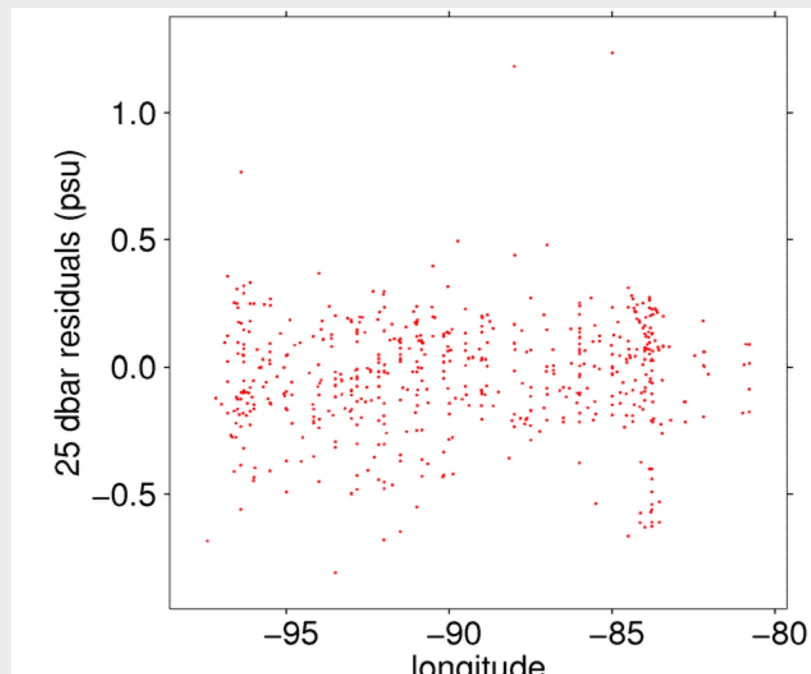
Estimated and observed salinity

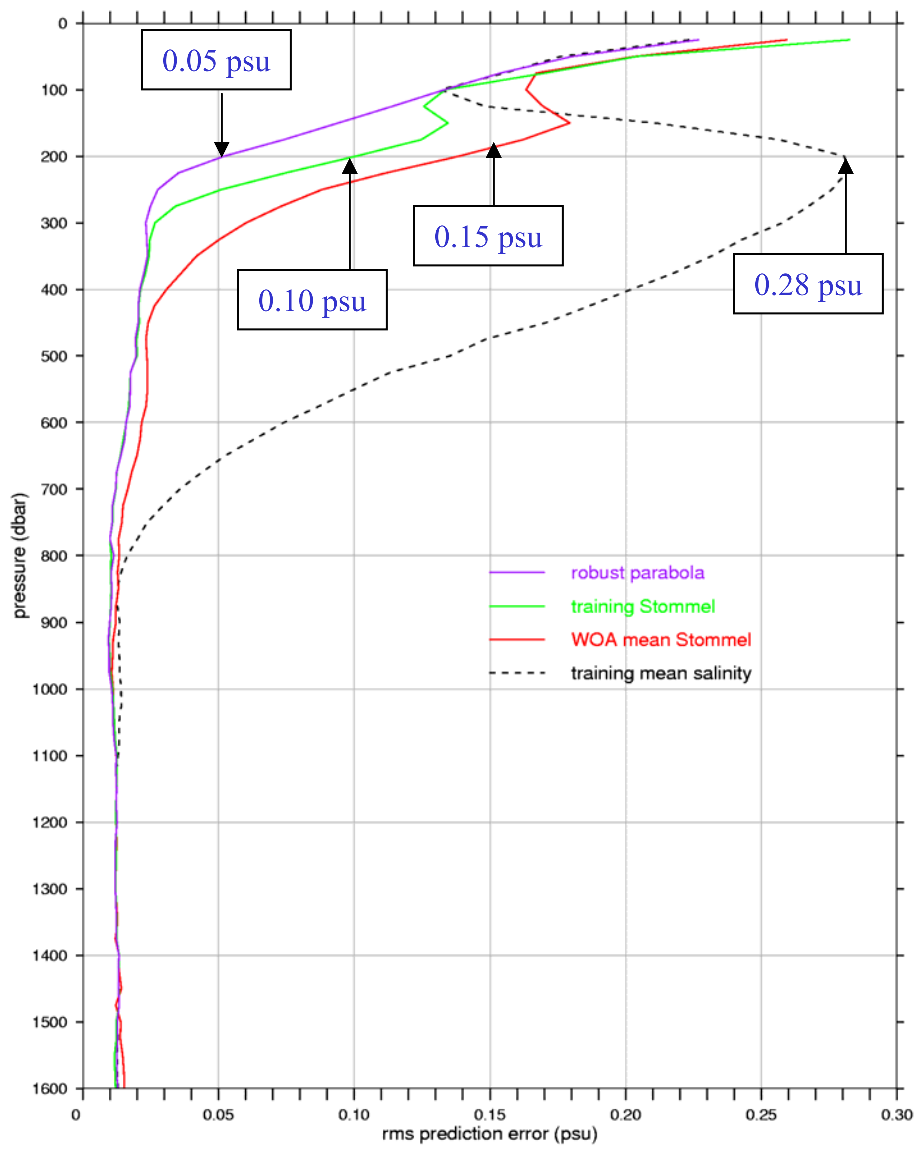
robust parabola

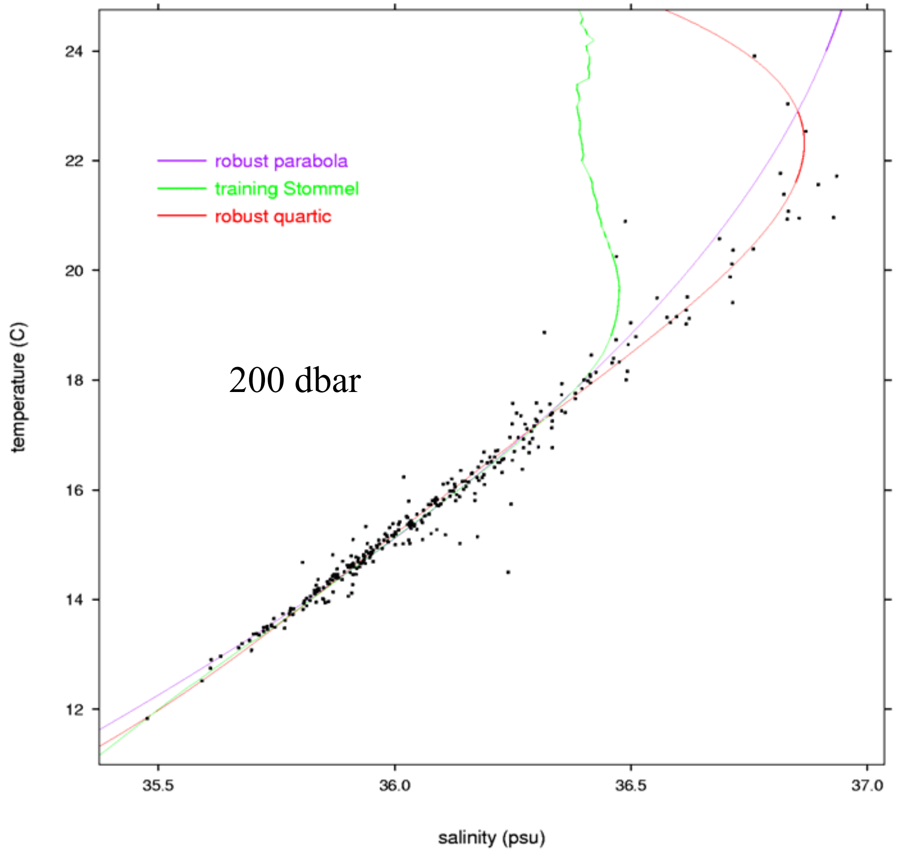




Residuals of robust parabola show no systematic seasonal or spatial behavior at 25 dbar.

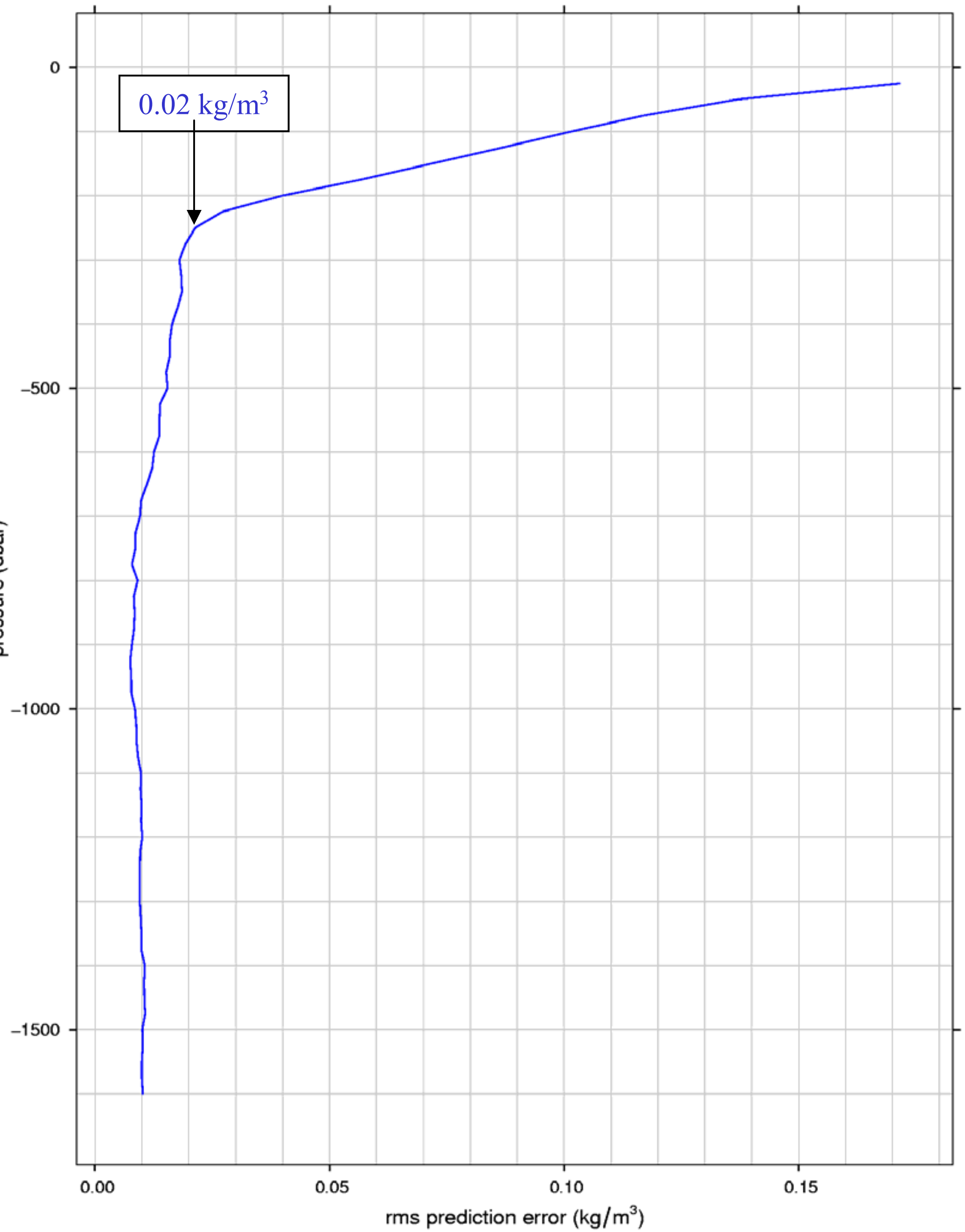






Stommel's method misses Loop Current

Models derived from CTD data are considerably more accurate than those inferred from climatology.

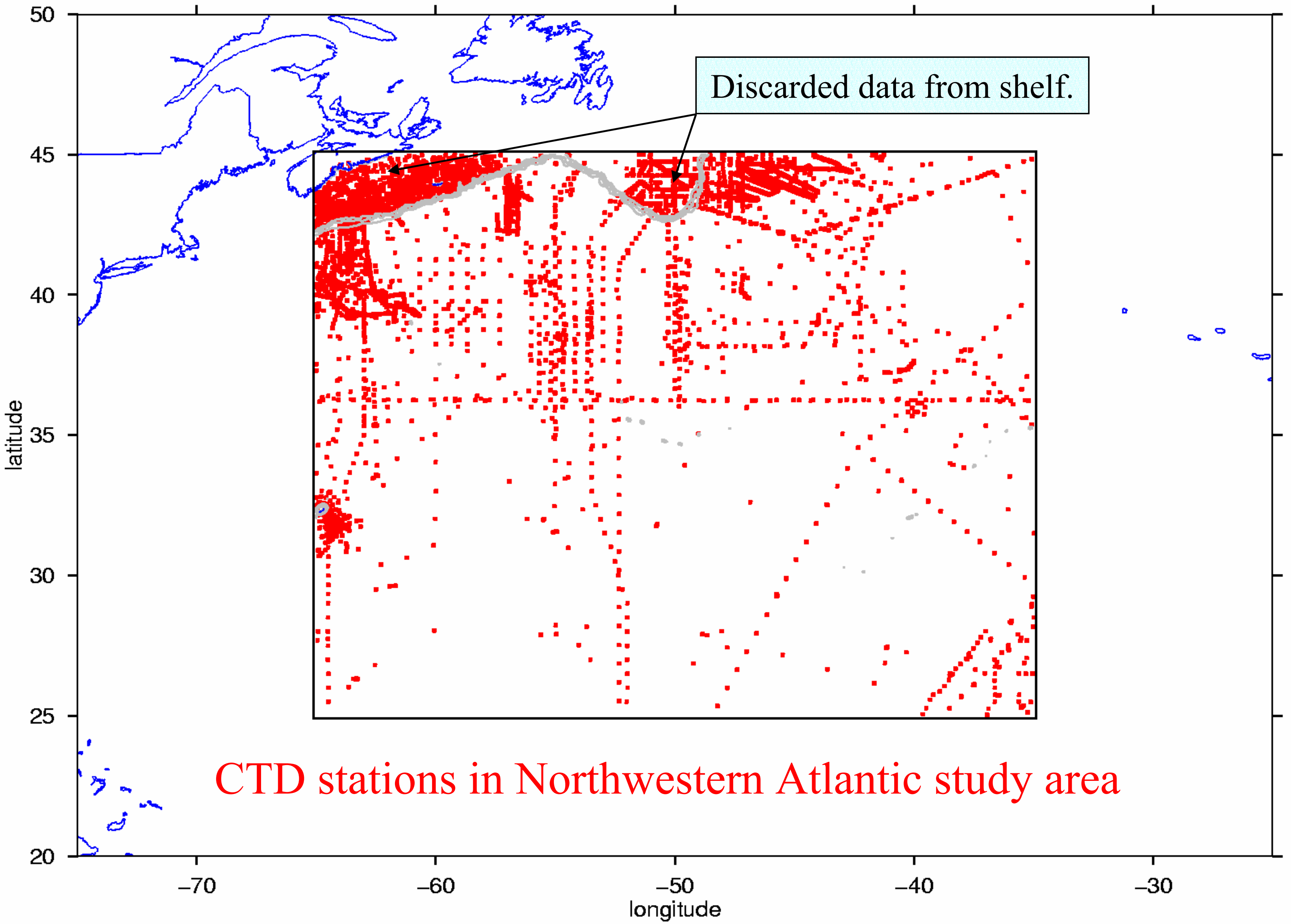


How good is density ?

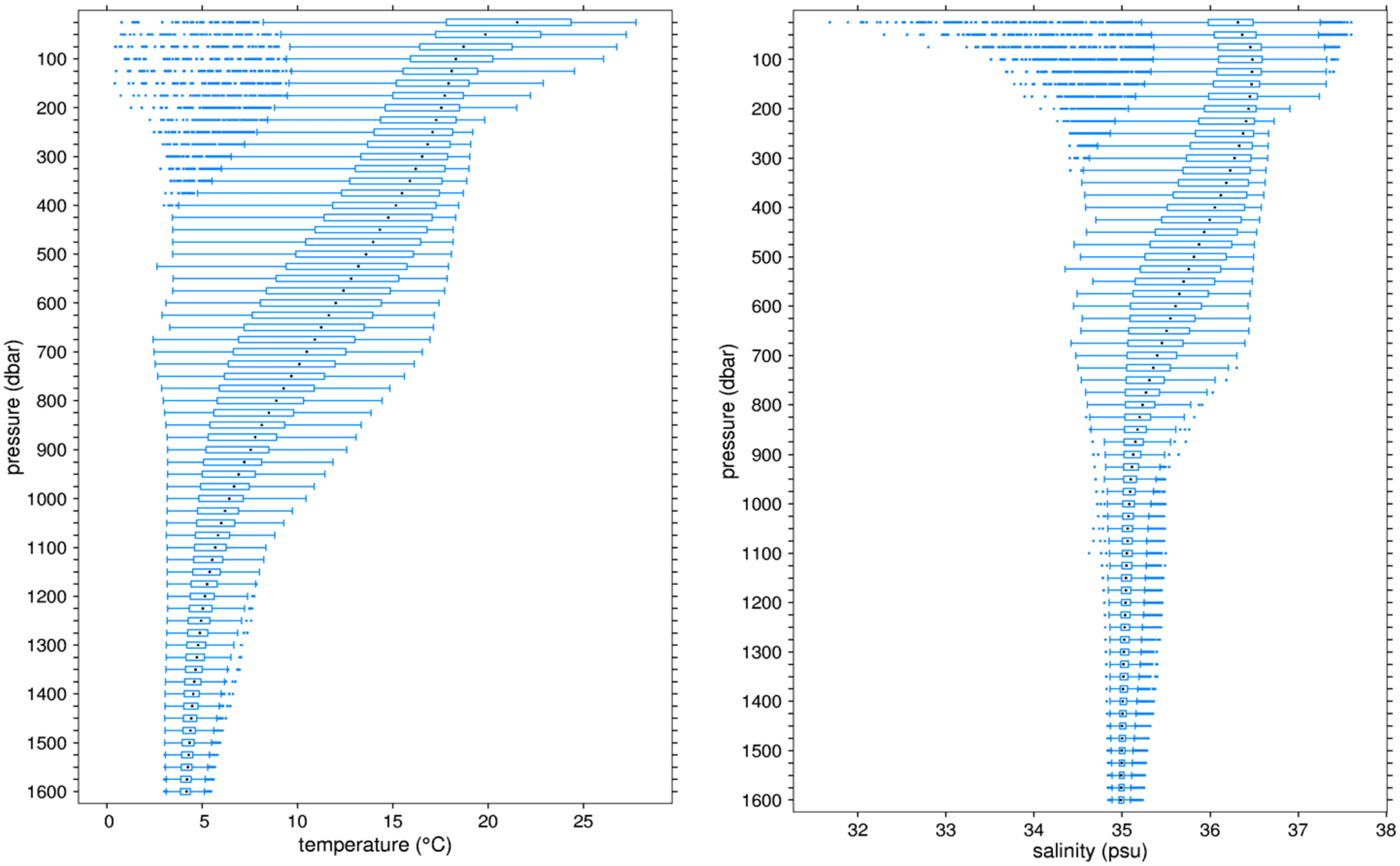
Might want better near the surface.

Get same accuracy with regression models for density.

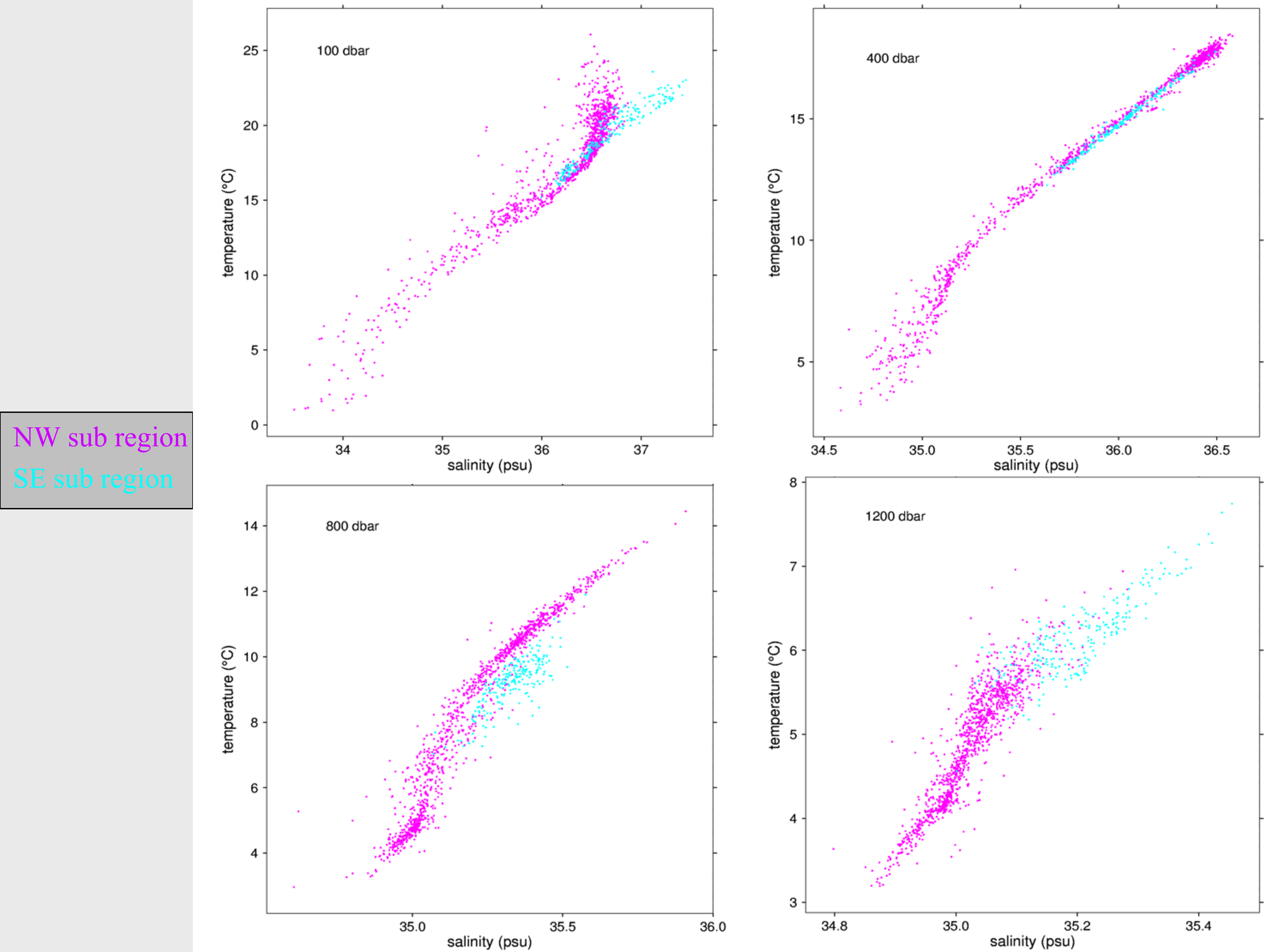
Northwest Atlantic



Discarded data from shelf.
CTD stations in Northwestern Atlantic study area
latitude
longitude
50
45
40
35
30
25
20
−70
−60
−50
−40
−30

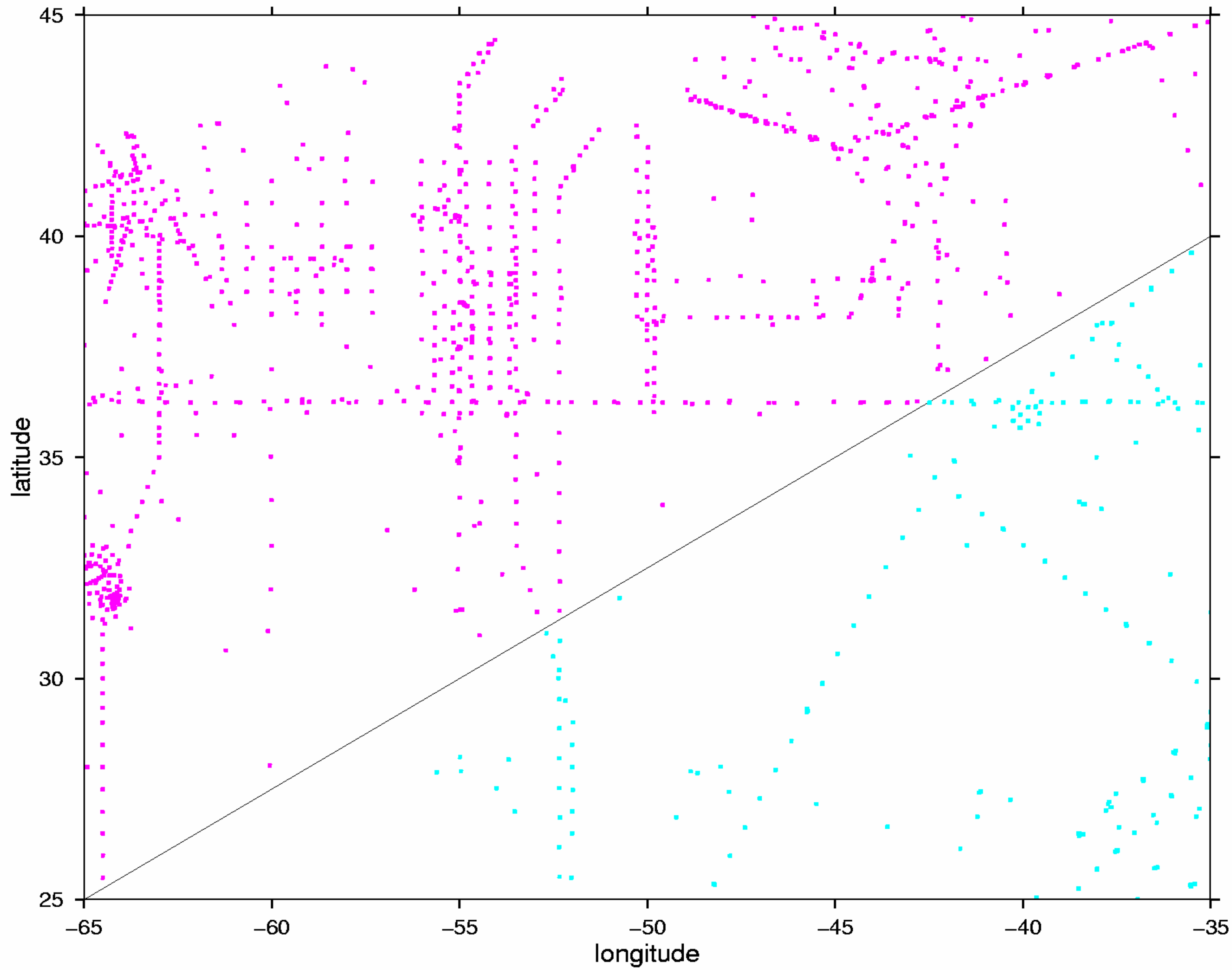


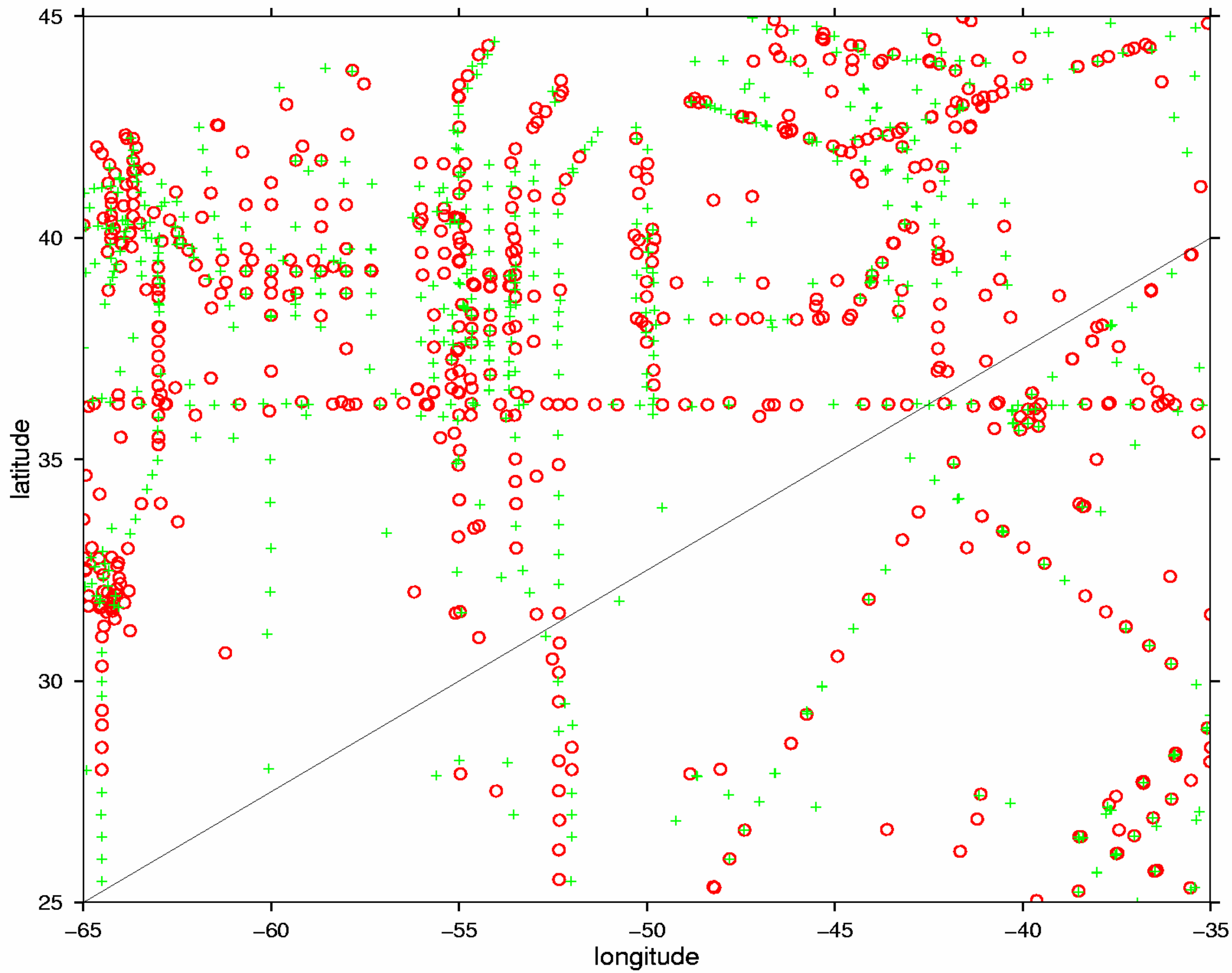
Distributions of CTD observations in NW Atlantic



NW sub region
SE sub region
100 dbar
400 dbar
800 dbar
1200 dbar
temperature (°C)
salinity (psu)

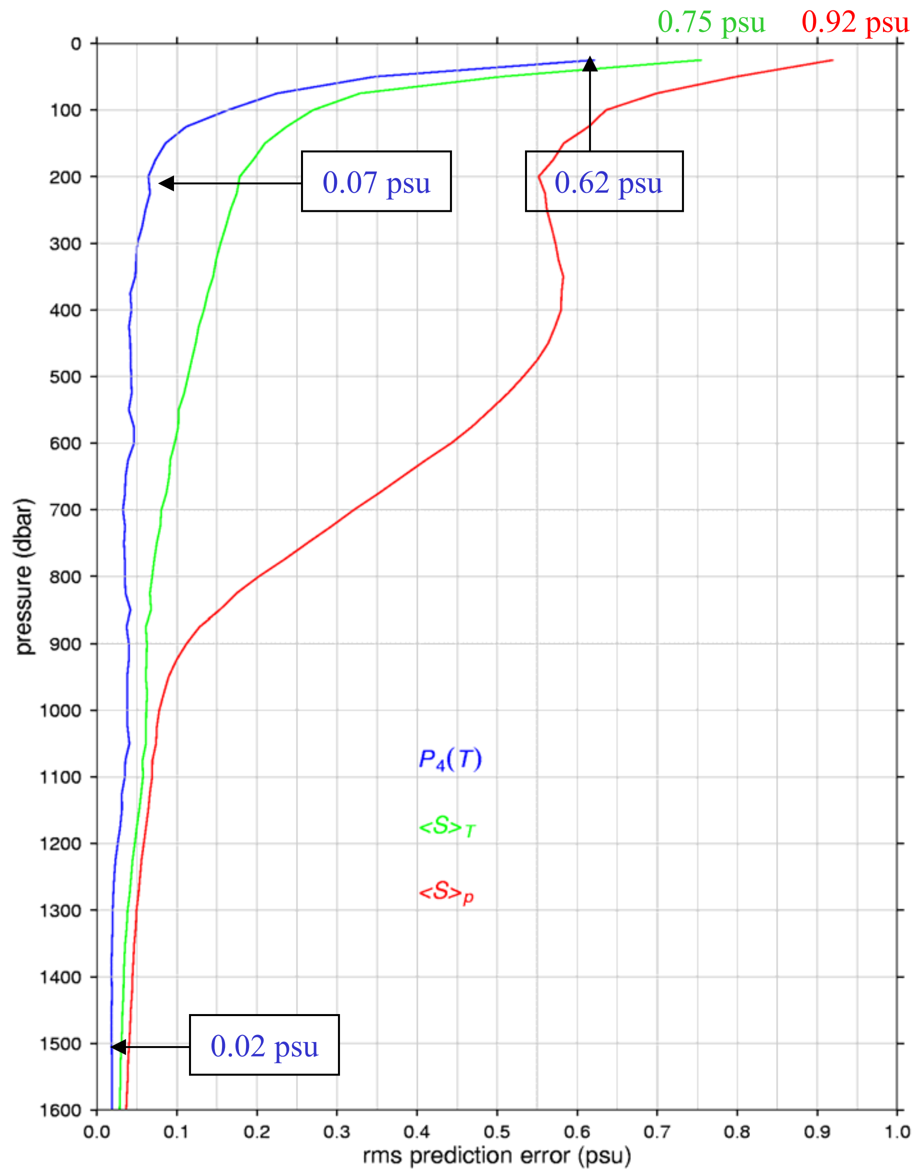
1390 stations with long profiles in northwestern Atlantic

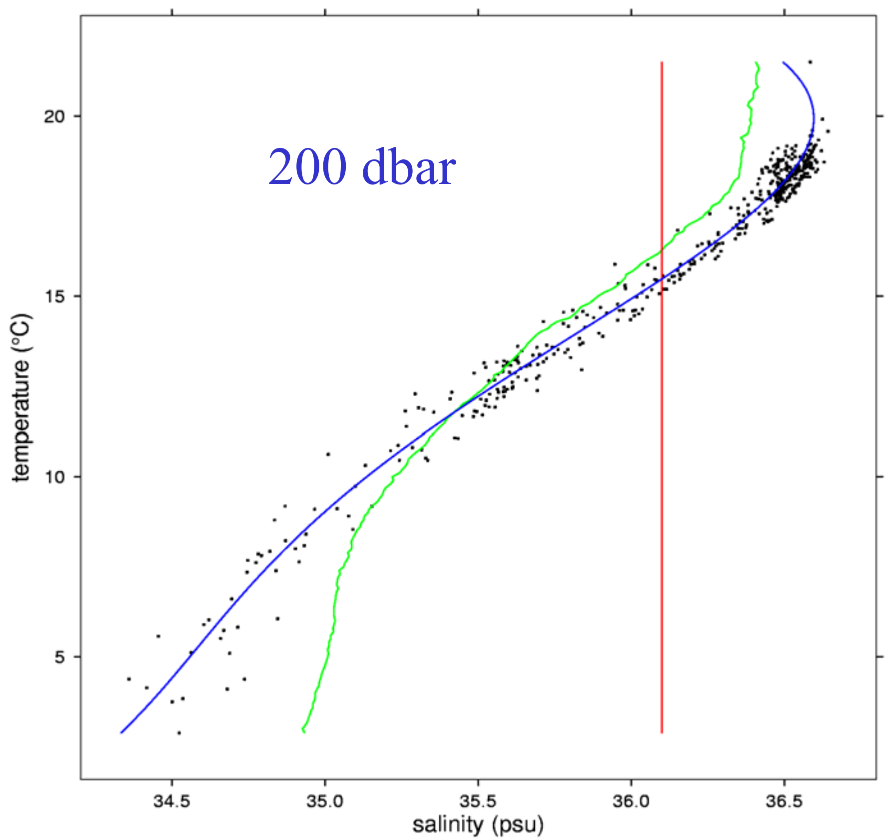




Training and verification stations

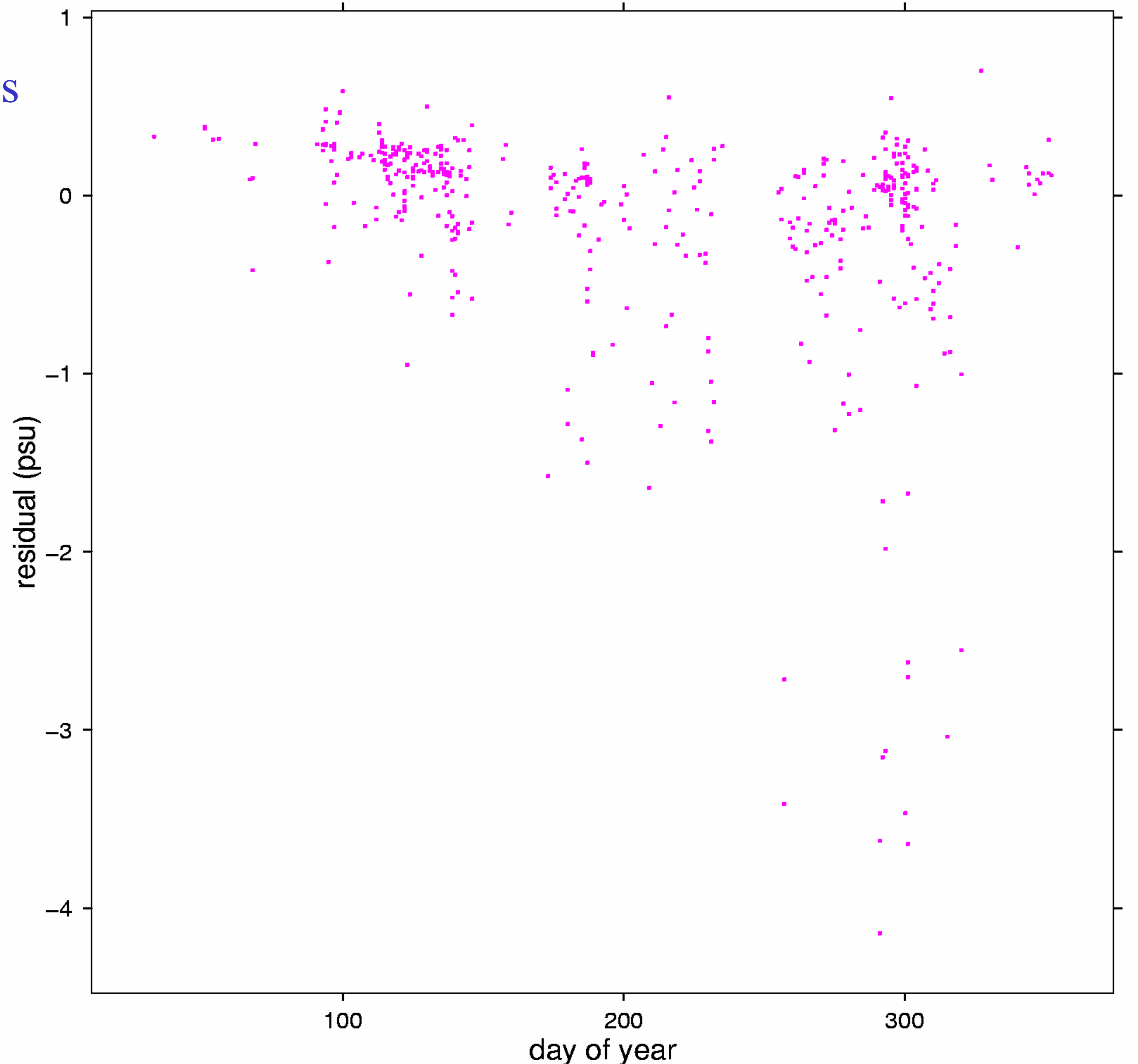
NW sub-region





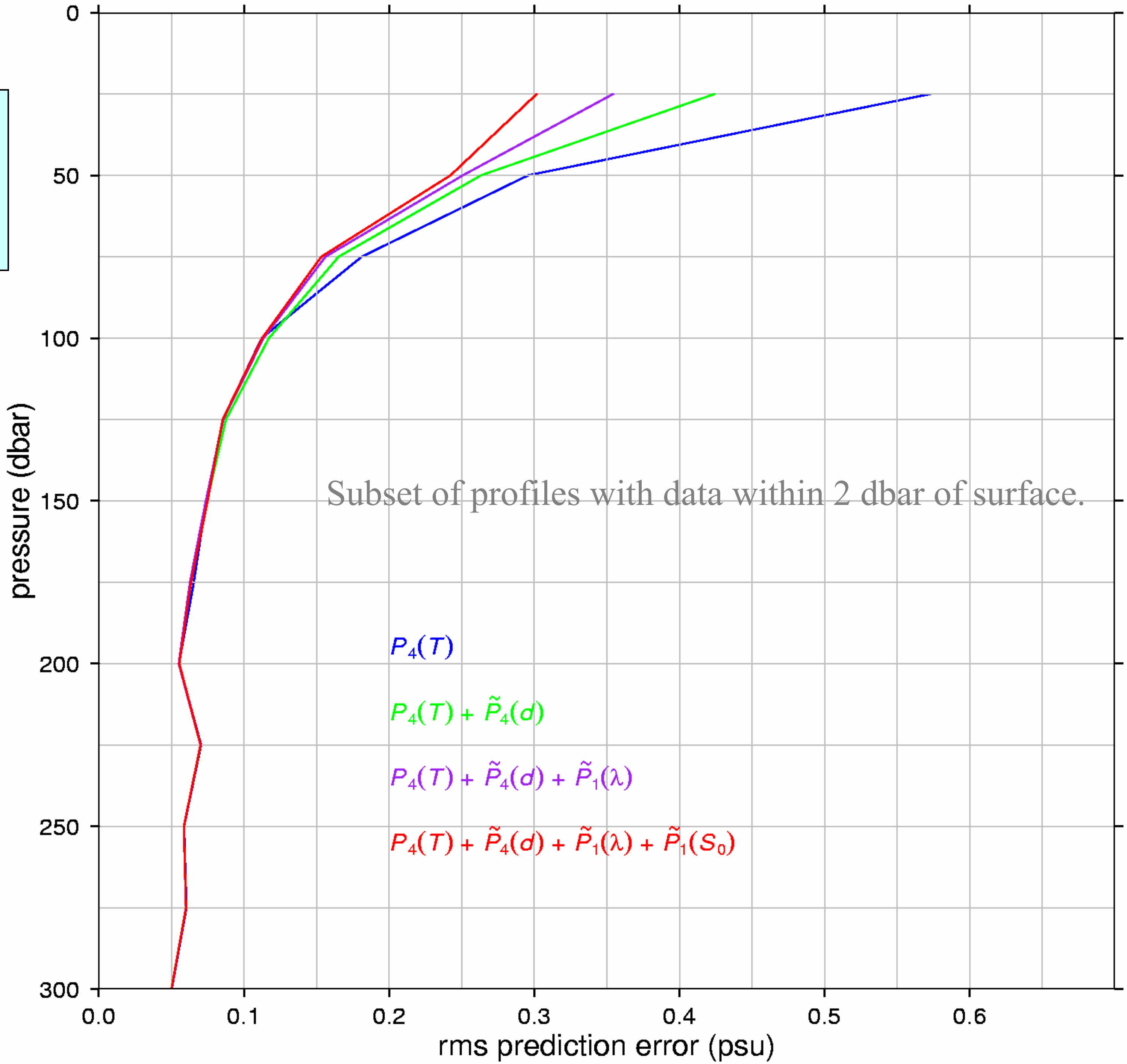
Skill explaining independent data

$P_4(T)$ residuals at 25 dbar



NW sub-region

Other variables help near the surface



T temperature

d day-of-year

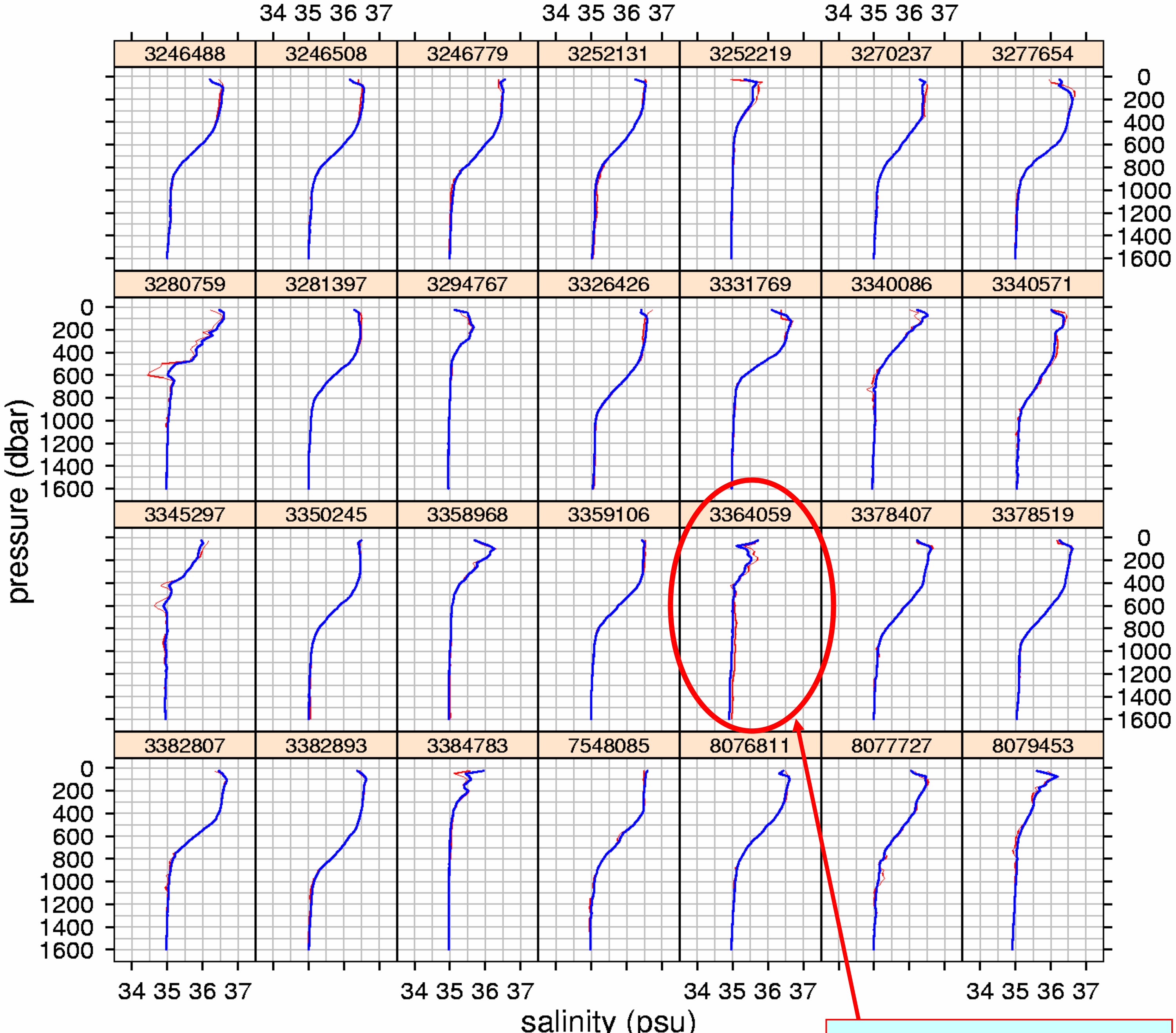
λ Longitude

S_0 surface salinity

Estimated

vs.

observed



28 randomly selected NW profiles

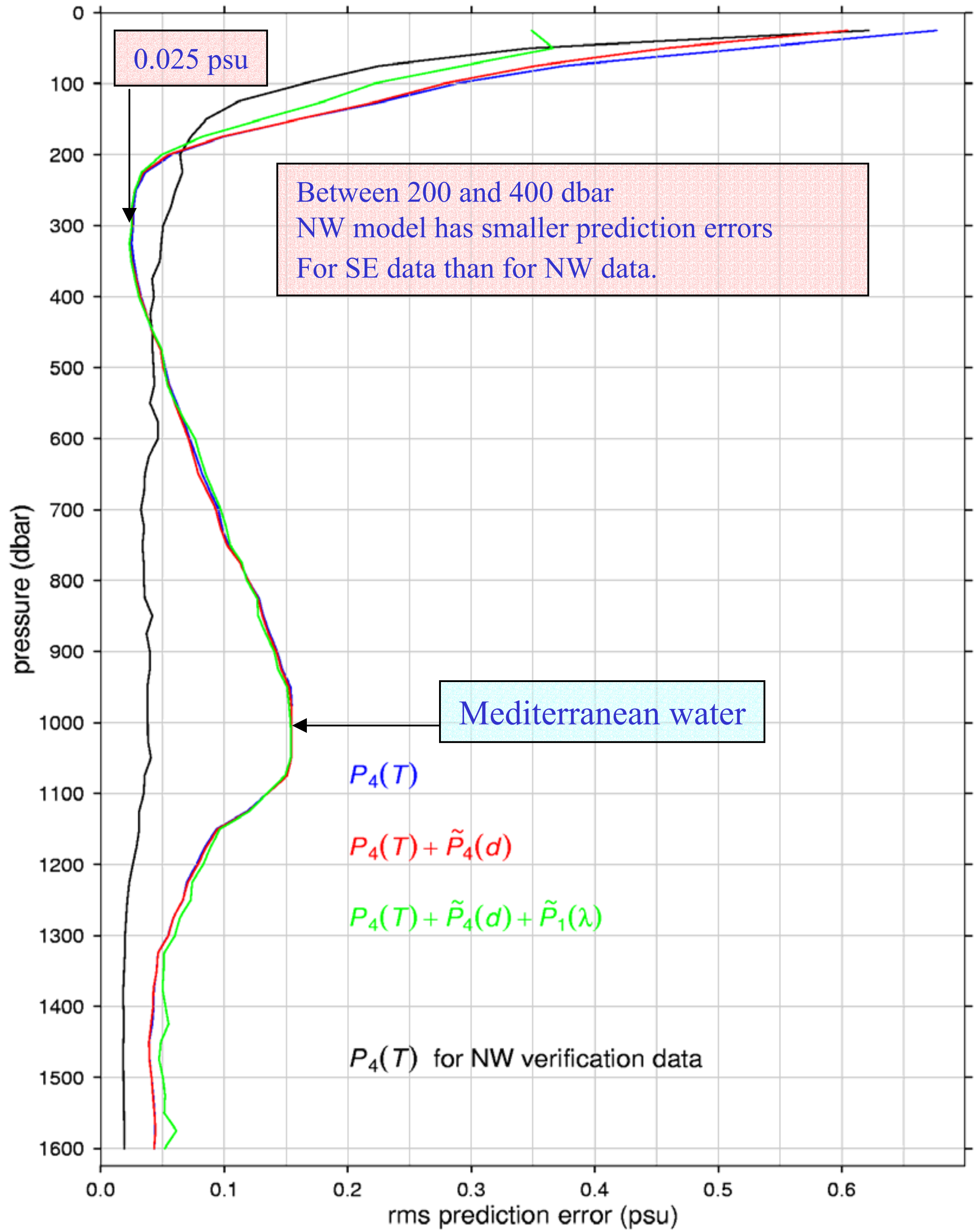
Calibration error?

NW models applied to SE data

T temperature

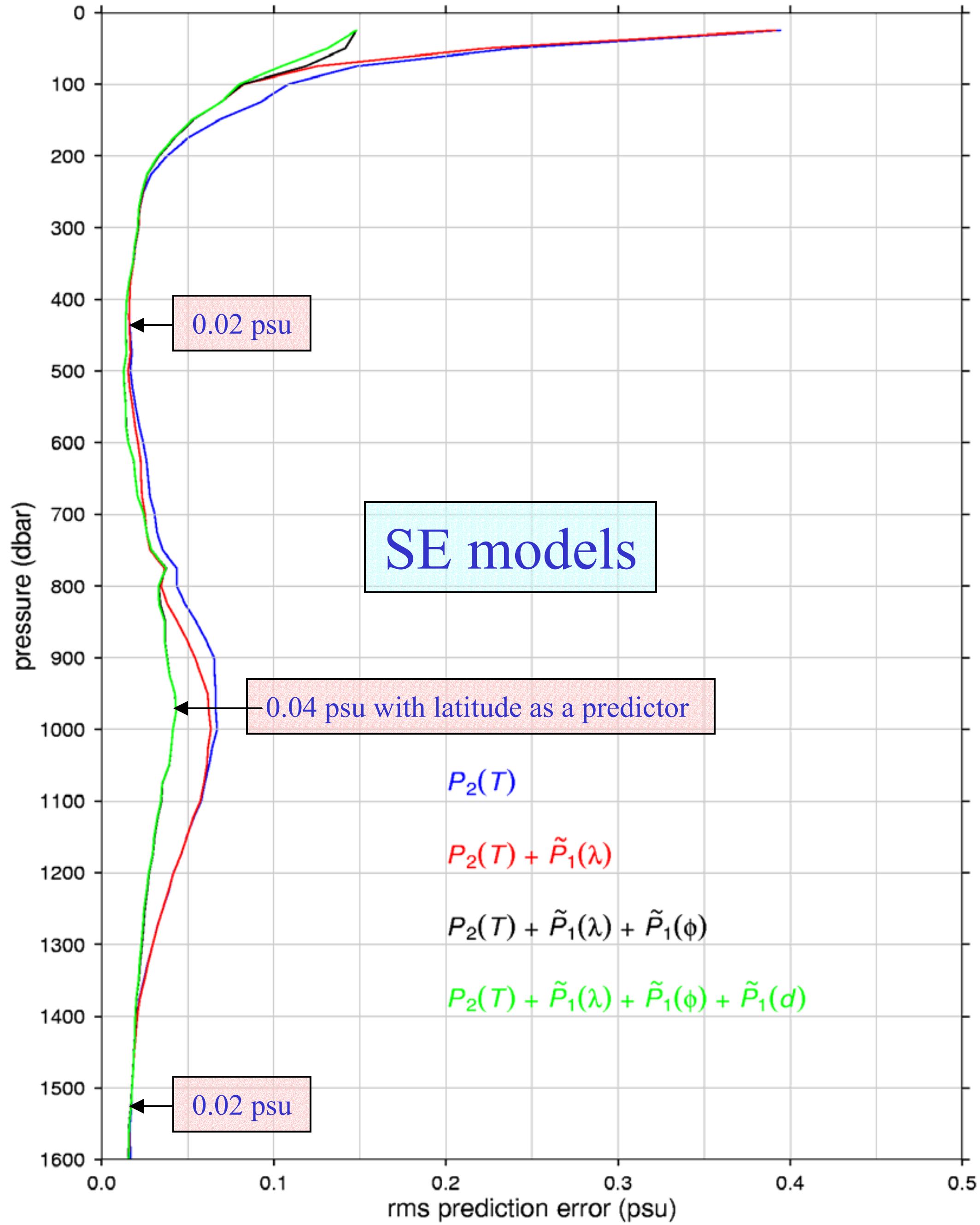
d day-of-year

λ longitude



SE sub-region

Skill explaining independent SE data



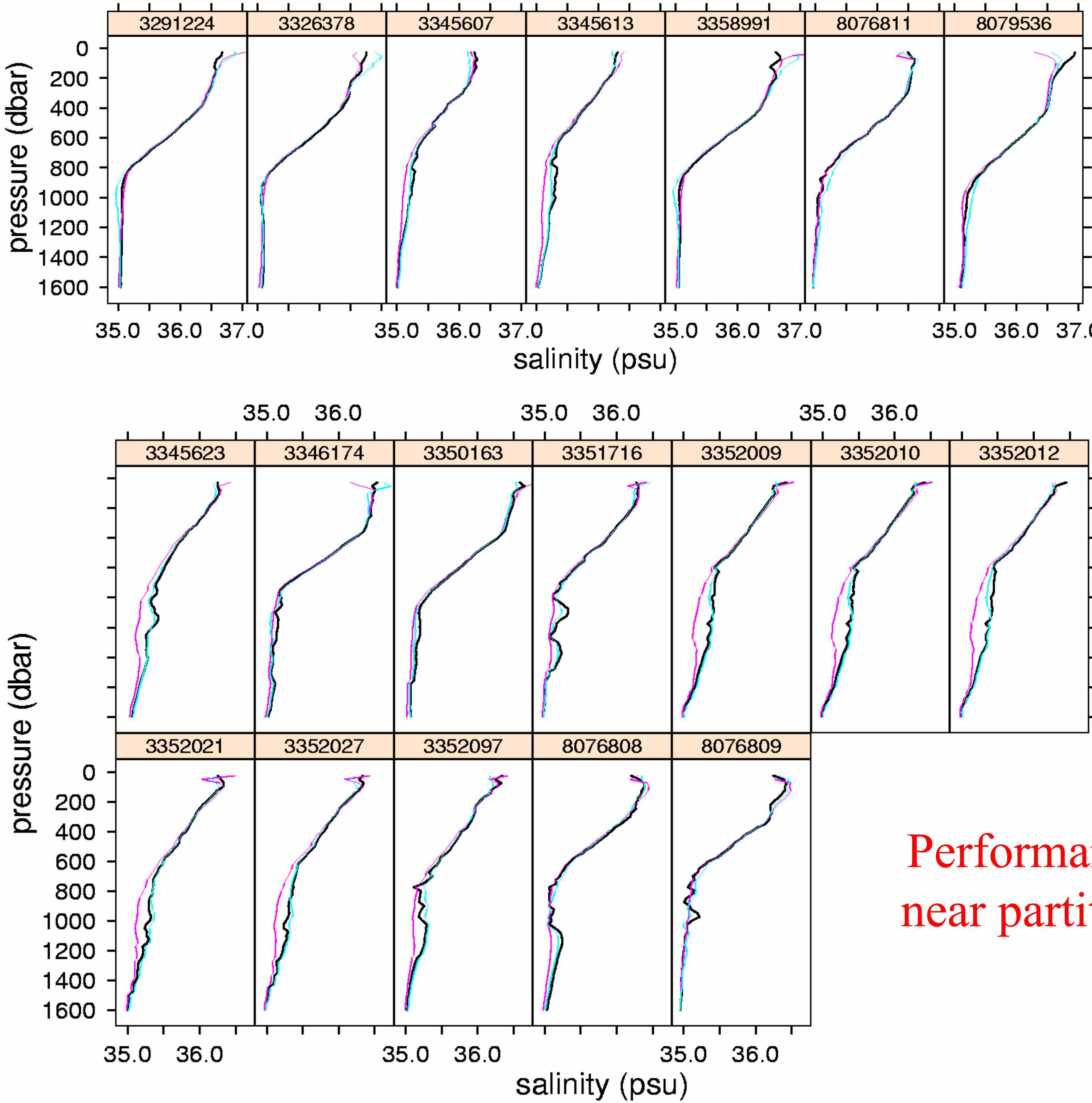
T temperature
λ longitude
ϕ latitude
d day-of-year

NW profiles

NW model

SE model

SE profiles

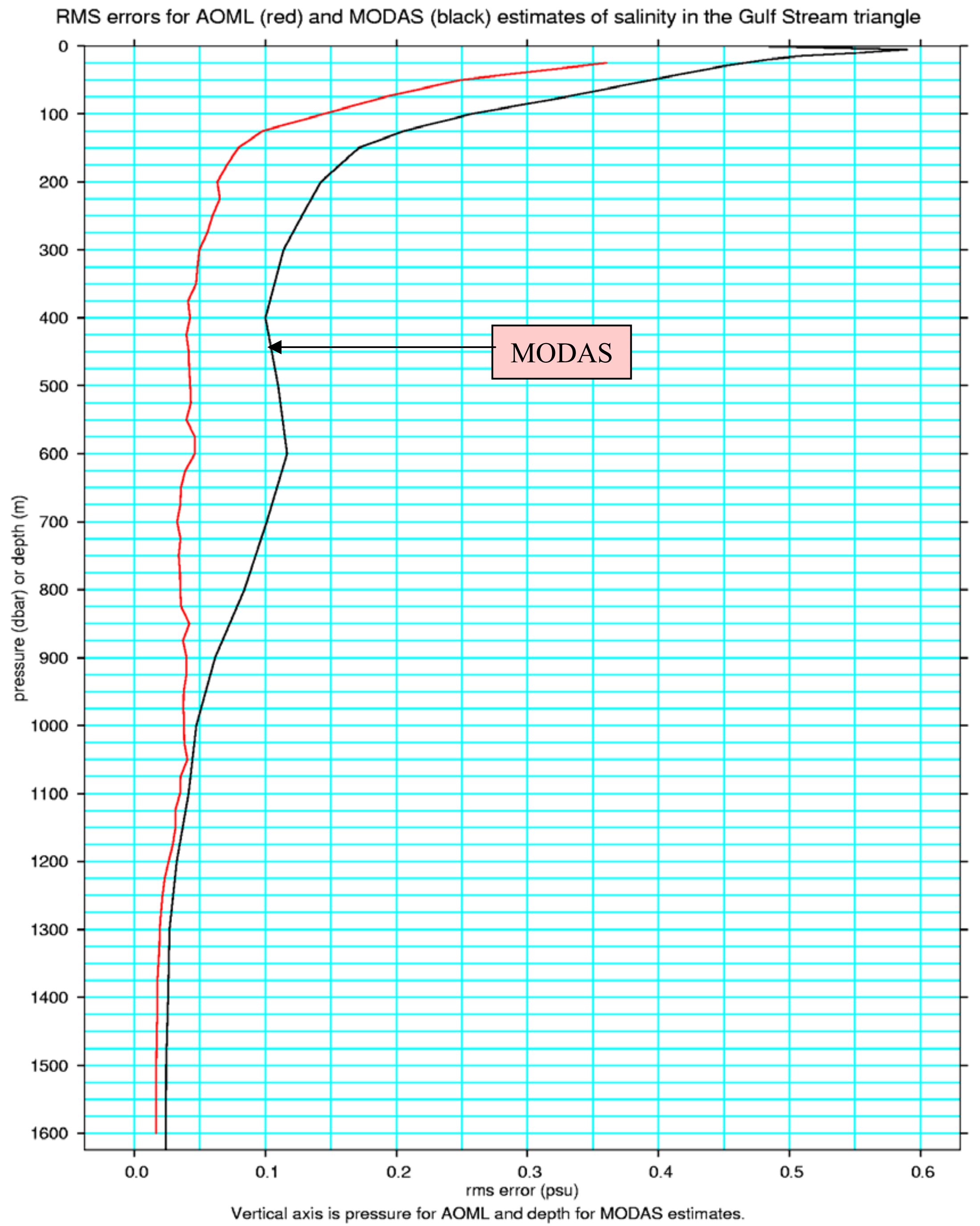


Performance near partition

Regression beats Navy's MODAS system in Gulf Stream triangle.

Best regression model
for NW sub-region
(Gulf Stream and its eddies)
$P_4(T)+P_4(d)+P_1($🖰$)$

4^{th} degree in temperature
4^{th}in day of year
1^{st} in longitude



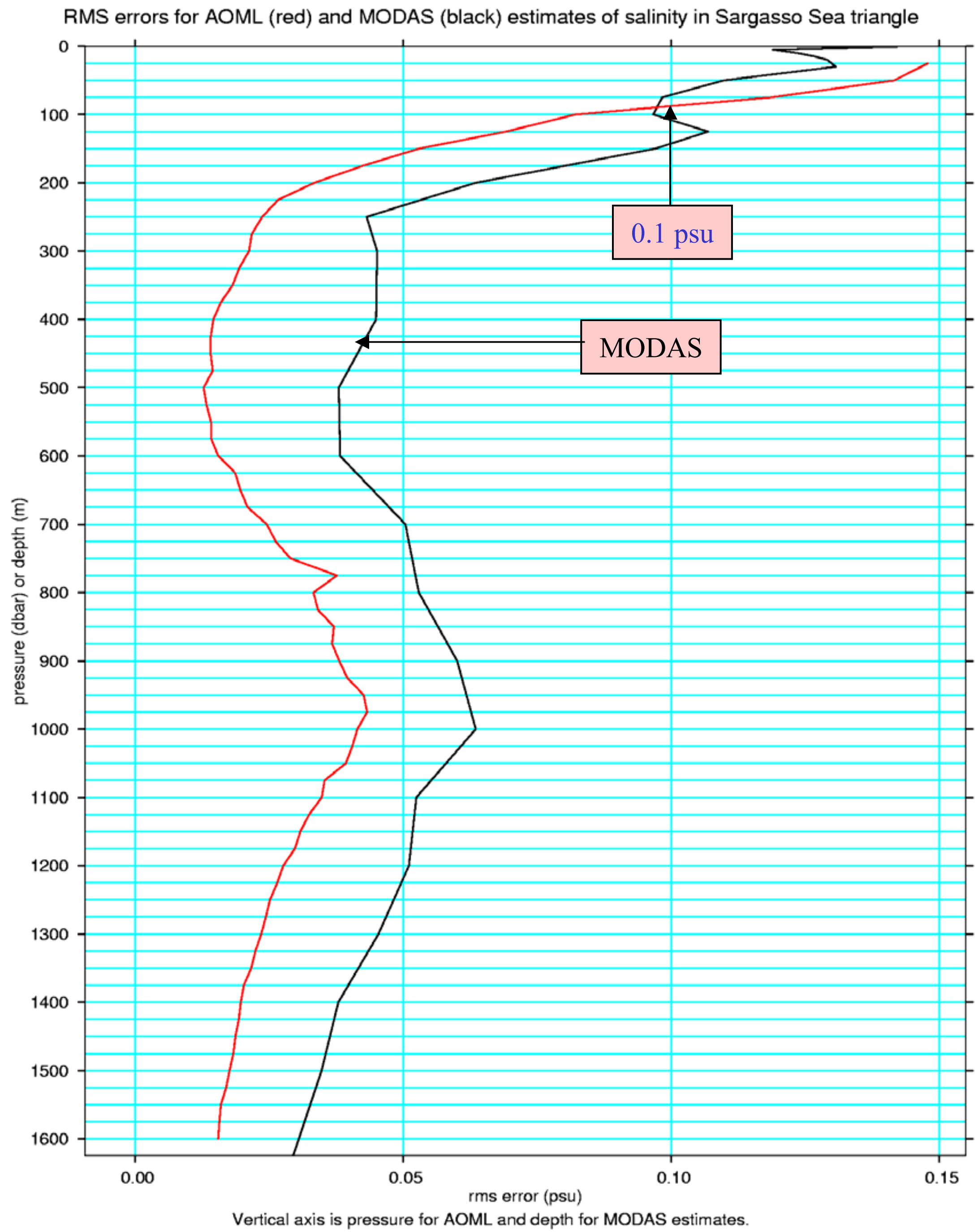
Except near the surface regression beats Navy's MODAS system in Sargasso Sea triangle.

Best regression model for SE sub-region (Sargasso Sea)

$P_2(T)+P_1(\varphi)+P_1(\lambda)$

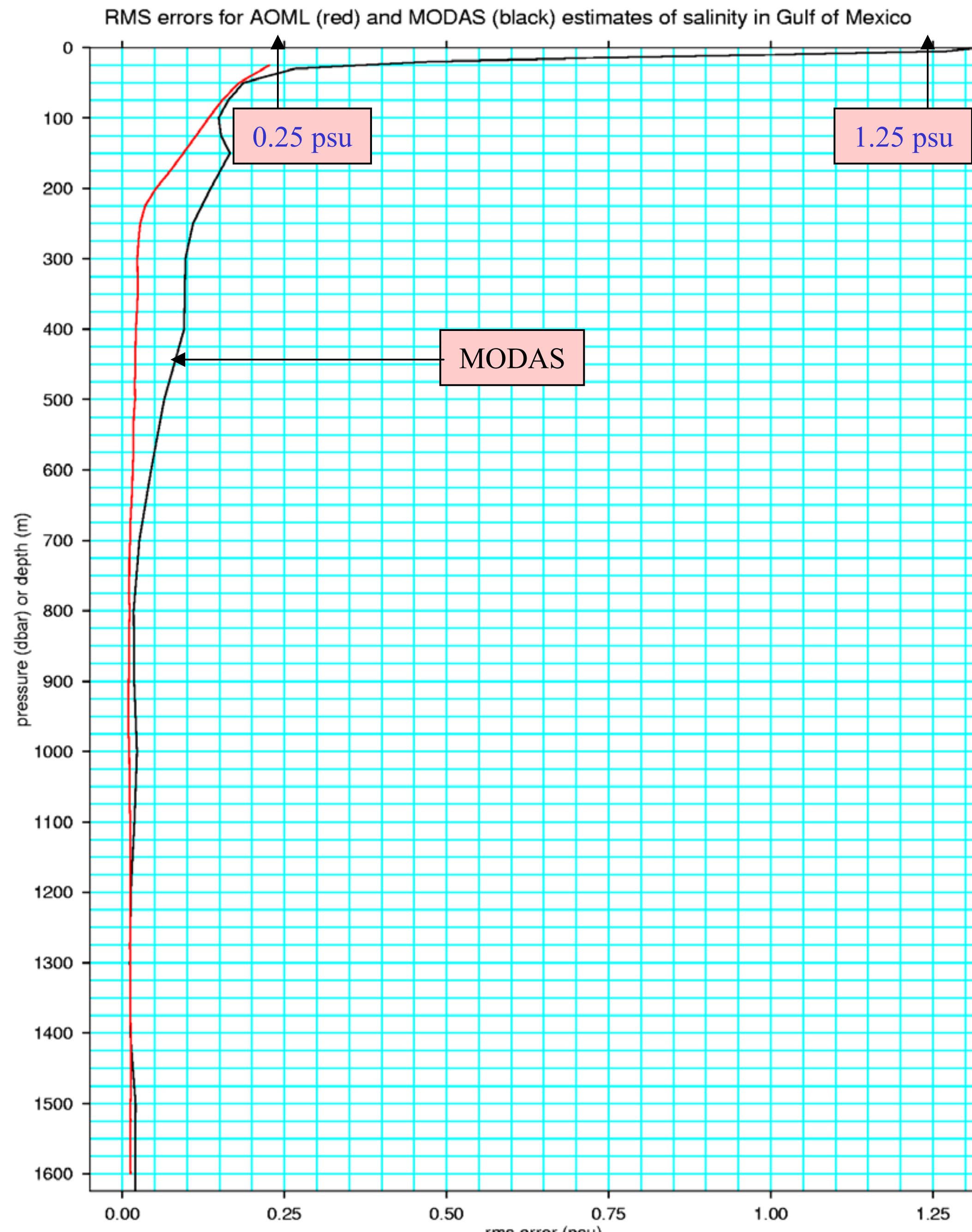
2nd degree in temperature

1st in latitude and longitude



Regression beats Navy's MODAS system in Gulf of Mexico.

Best regression model
for Gulf of Mexico
$P_4(T)$
4^{th} degree in temperature



Vertical axis is pressure for AOML and depth for MODAS estimates.

Conclusions:

- Regression beats using climatological T and S.
- Can handle fronts.
- Where to draw regional boundaries?
- Accurate near-surface estimates are difficult.
- Can use to check salinity calibration in CTD archives.
- Can also check ARGO float calibration.
- Big ocean – still lots of work to do.
- Can use MODAS until more regions are modeled.
- MODAS is being reworked.