#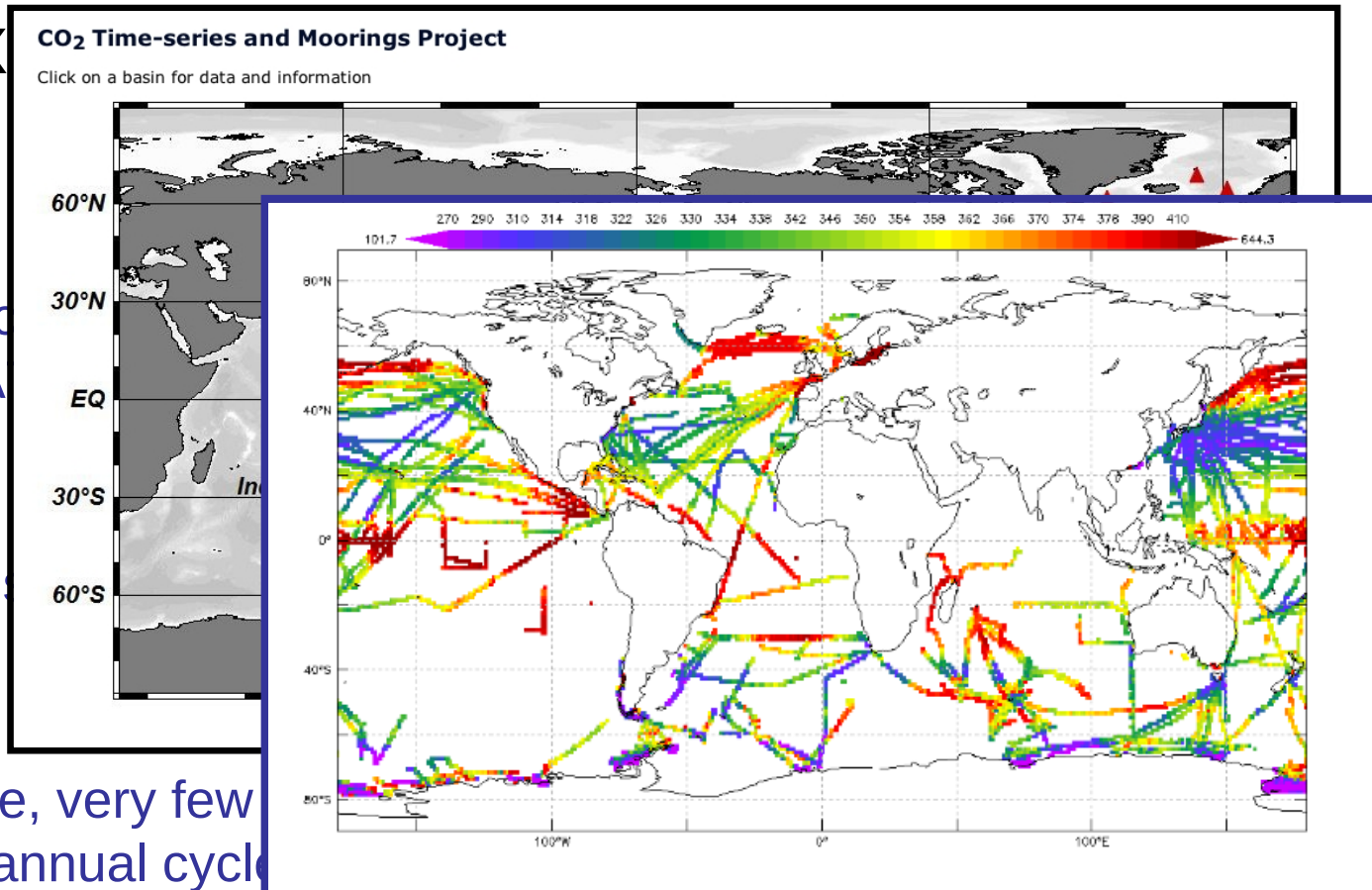 Prediction of ocean carbon dioxide levels at observational time series nodes using satellite ocean remote sensing products

D. Vandemark[1], T. Moore[1] , J. Salisbury[1] , B. Chapron[2]

[1]*University of New Hampshire, EOS/OPAL, Durham, NH*
[2]*IFREMER /Centre de Brest, LOS, Plouzane, France*

UNIVERSITY *of* NEW HAMPSHIRE

Ifremer

Mass flux

CO₂ Time-series and Moorings Project

Click on a basin for data and information

- Standard approach
  Takahashi/SOCA
  CTracker)
- **Pros:** increasing
  mean global and
  points)
- **Cons:**
  - time coverage, very few
    for even one annual cycle
  - spatial coverage, many basins (e.g. SO, Indian
    Ocean, and S. Pacific & Atlantic are very sparsely
    sampled)

UNIVERSITY of NEW HAMPSHIRE

Ifremer
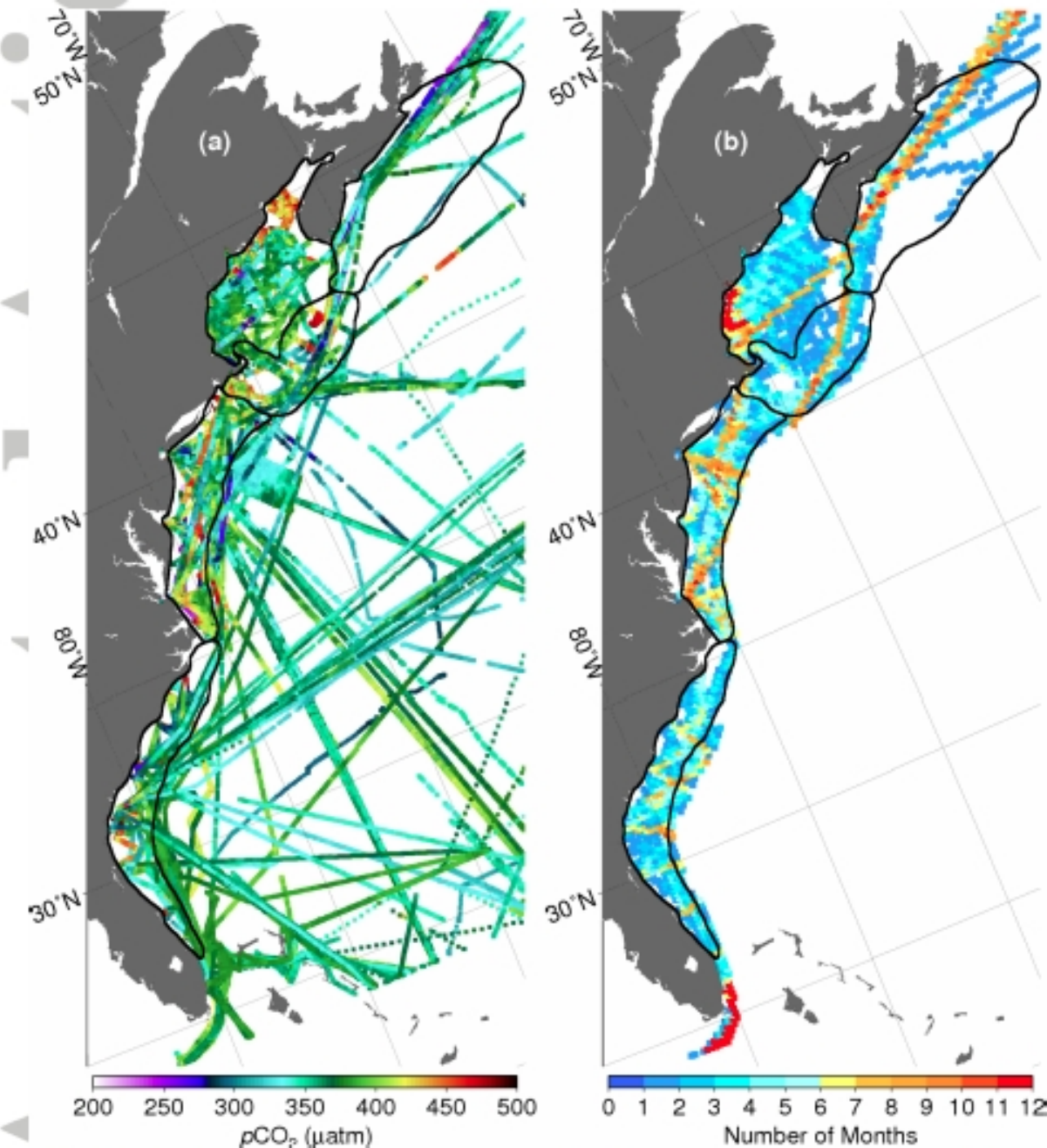
US East Coast example from Signorini et al 2013

More than 800,000 indiv. samples, 1987-2011

Few pixels with monthly coverage

200 250 300 350 400 450 500
pCO_2 (μatm)

0 1 2 3 4 5 6 7 8 9 10 11 12
Number of Months

Fig. 3. Color-coded SOCAT surface ocean pCO_2 cruise tracks (a) and corresponding coastal binned data (b) with associated color-coded temporal coverage in months. The highest temporal coverage corresponds to the most travelled routes (in orange to red), i.e., most frequent destination ports (Boston, New York, Norfolk, Miami) used by

Ifremer

Objective – global space/time ΔpCO2 estimator at daily-to-weekly time step

- Current state of the art:
  - Using satellite or model estimates of key controlling factors (SST (1st order), Chl/SSS/ML) + ship-based CO2 measurements to develop empirical models
  - Recent examples: Signorini et al. 2013; Liu-JPL with SVR; Chierici et al. 2012; Hales et al., 2012; Lohrenz & Cai, 2006; Zhu et al., 2009; Gledhill et al., 2009
  - Regionally < 5-25 uAtm rms, globally higher

UNIVERSITY of NEW HAMPSHIRE

Ifremer

Objective – global space/time ΔpCO2 estimator

- Issues:
  - Is the ocean pCO2 data sampling coverage (time & space) sufficient to train a model
  - Validation can be tenuous (how to determine sample independence?)
  - Choice of inputs: often not mechanistic (e.g. lat, long, time_day, time_year) nor sufficient
  - Is upper ocean pCO2 too variable for such global or even regional models?
  - What level of accuracy needed for gas transfer rate investigations and flux products?

UNIVERSITY of NEW HAMPSHIRE

Ifremer

Our approach – focus on time series

- Assumption: without time resolved algorithm training datasets, pCO2 prediction is likely under-constrained

- Desired training sets - continuous in situ pCO2 data with sub seasonal sampling and limited regional span

Ocean pCO2 time series data for algo input

- Goal: Obtain virtual and actual time series data
- Requirements
  - sub seasonal resolution
  - at least one complete annual cycle
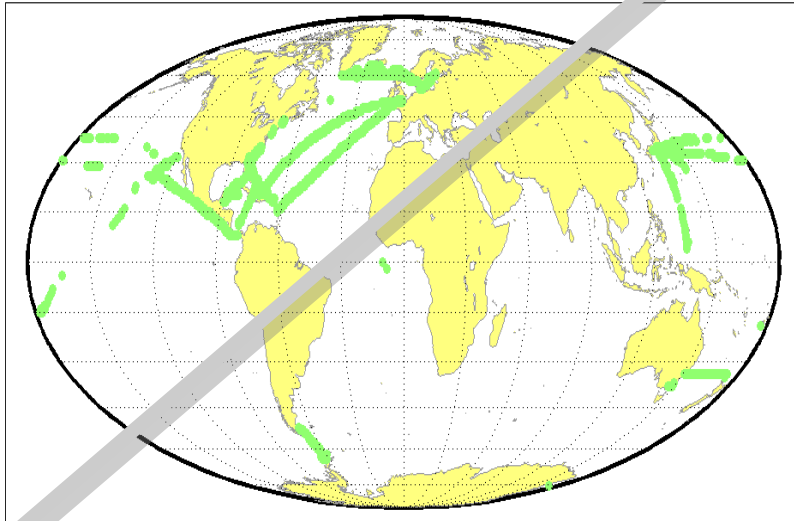 - select several diverse ocean regions and apply consistent algorithm development approach

VIRTUAL (SOCAT)

MOORED

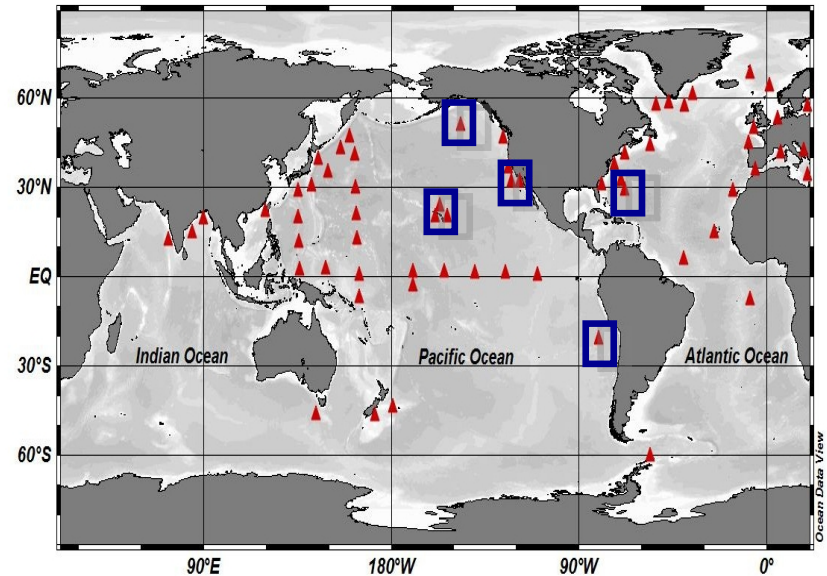Very few regions yielding a time series, even using recent SOCAT

Several (few) of the global network have multi-month time sets online

Pixels with at least 8 months data for 2 years: 2005 and 2006

CO₂ Time-series and Moorings Project

Click on a basin for data and information

Indian Ocean   Pacific Ocean   Atlantic Ocean

UNIVERSITY
of NEW HAMPSHIRE

Ifremer

Methods overview

5 site training sets: Satellite matchups + buoy pCO2

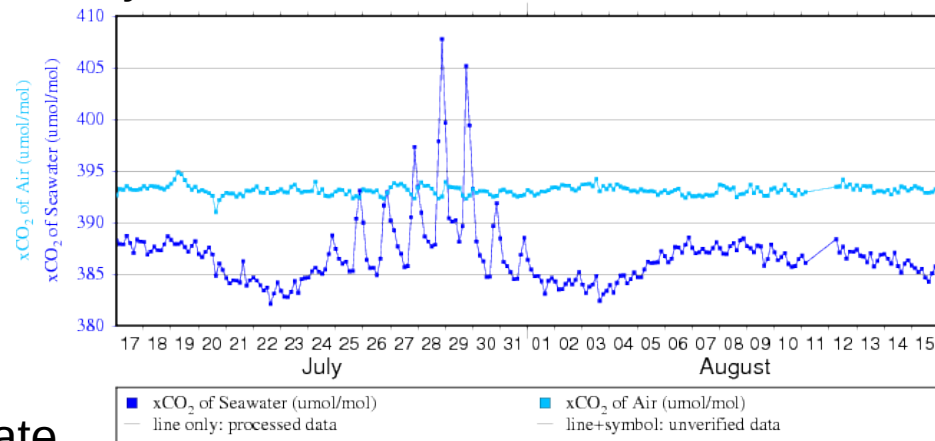LS regression modeling
       Multiple Linear Regression
       Neural Network

Mechanistic modeling (e.g. 1-D temp-normalized pCO2 vs. DIC_bio, carbonate closure with TA(SSS)) NOT DISCUSSED HERE

Similar methodology at each sites, 3 N Pac, 1 S Pac, 1 N. Atlantic + an overall model for experimentation

Validation

xCO_2 of Seawater & xCO_2 of Air @ Stratus (85W,20S)
[Date: 2013–07–17 to 2013–08–16]

- ■ xCO_2 of Seawater (umol/mol)   ■ xCO_2 of Air (umol/mol)
- line only: processed data   line+symbol: unverified data

Overall matchup dataset creation, site-by-site

Type	Source	Time	Spatial Res.
In situ SST	Mooring	sub-daily	m
In situ SSS	Mooring	sub-daily	m
In situ pCO2	Mooring	sub-daily	m
Sat. Chl_a	GlobColour	daily	9 km
Sat. SST & PAR	Aqua-MODIS	daily	9 km
Sat. NPP	OSU/MODIS	8-day	9 km
Model MLD	OSU/hycom	8-day	9 km

pCO2, in situ and satellite data used for algorithm training

Total buoy measurements and coincident cloud-free satellite data

Note: from 7 M pCO2 samples down to few hundred samples per site

Source	Type	Region	Time period	# total points (daily/native)	# satellite extractions (all reqd)
Stratus	Buoy	S. Pacific	2006-2008	1185/9543	339
Papa	Buoy	N. Pacific	2007-2010	809/6343	68
CCE1	Buoy	West coast	2008-2010	559/4468	162
WHOTS	Buoy	HOTS	2007-2009	598/4803	287
BATS	Buoy	N. Atl.	2005-2007	709/5345	384
SHIPBOARD					
Clivar	Ship	Atlantic	1997,2007	4556	701
Carina	Ship	N. Atl.	...	...	...
SOCAT	Ship	N. Atl.	1998-2001		30040
VOCALS	Ship	S. Pac.	2008-2009		6094

UNIVERSITY of NEW HAMPSHIRE

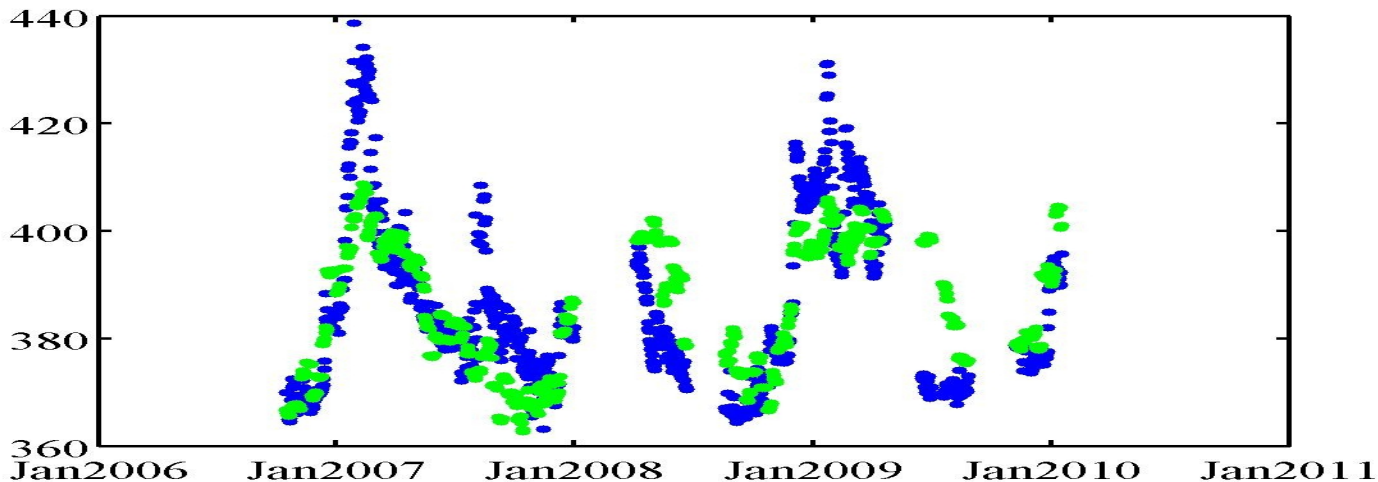*OceanFlux GHG Science Workshop, Plouzané, Sept. 2013*

Ifremer

Site-specific multiple linear regression – what inputs help?
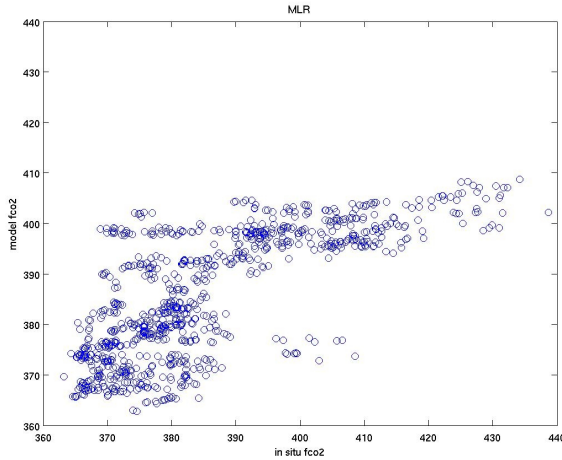
Stratus Time series pCO2 modeling

Inputs	RMS (uatm)	# pts *
SST, MLD	17.8	339
SST, MLD, SSS	12.3	339
SST, MLD, SSS, Chl	14.8	339
SST, MLD, SSS, NPP	11.3	339
SST, MLD, SSS, NPP, bbp/Chl	10.0	339
SST, MLD, SSS, NPP, bbp/Chl, CDM	8.7	62

Salinity drives largest variance reduction

Ocean color also contributes

Similar at each site PAPA, BATS, HOTS

Neural Network – factor 2 improvement

Stratus Time series pCO2 modeling - NN

| Inputs	RMS (uatm)	# pts
SST, MLD	7.4	339
SST, MLD, SSS	4.9	339
SST, MLD, SSS, Chl	4.4	339
SST, MLD, SSS, NPP	3.8	339
SST, MLD, SSS, NPP, bbp/Chl	4.2	339
SST, MLD, SSS, NPP, bbp/Chl, CDM	4.8	62

• Neural network solutions superior at each site

• Similar rms reduction as variables added up to 4 inputs (chosen best case)
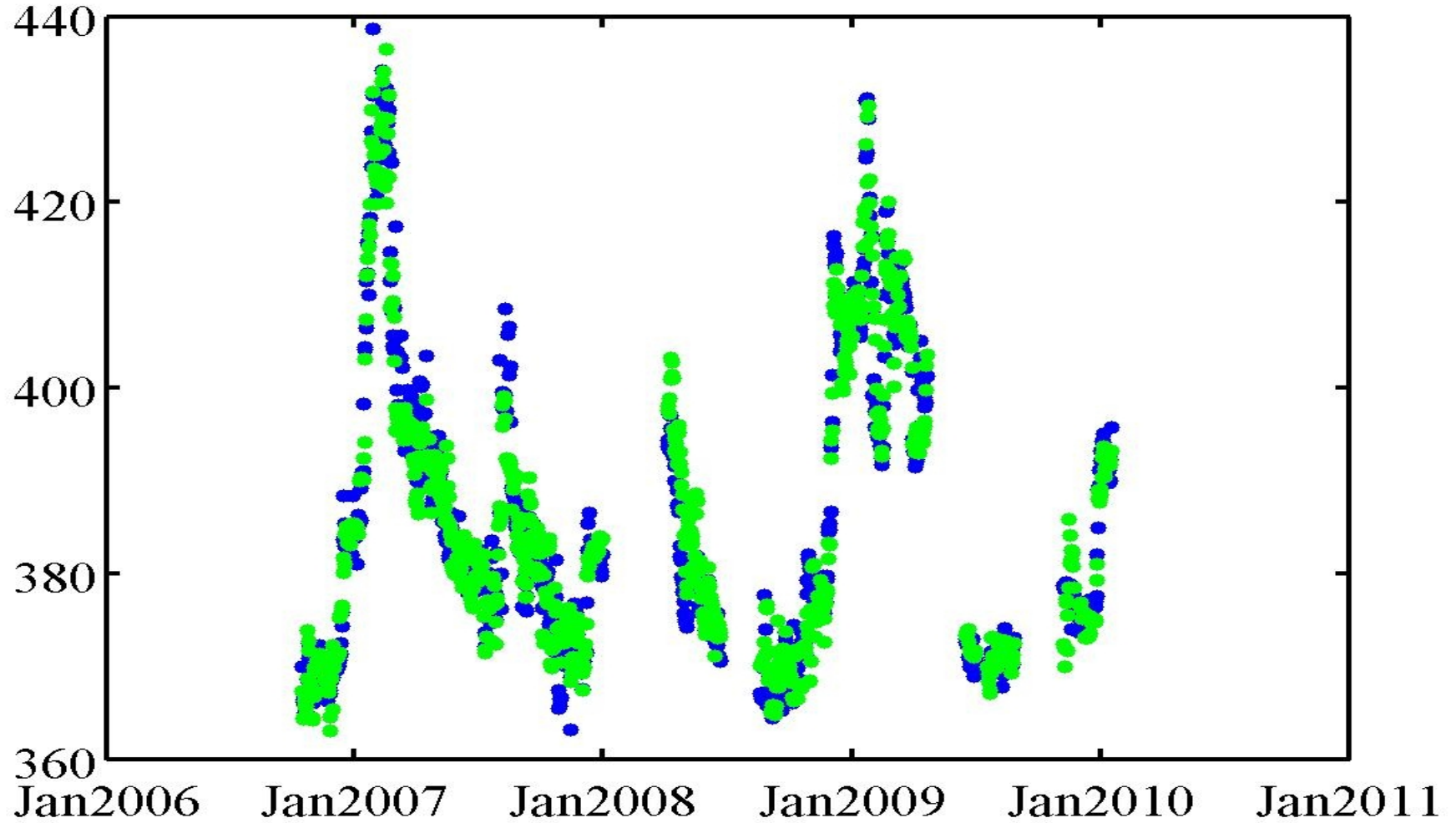
• Similar reduction and results site-by-site

Ifremer

Stratus Time series pCO2 model   NN (SST, MLD)

Ifremer
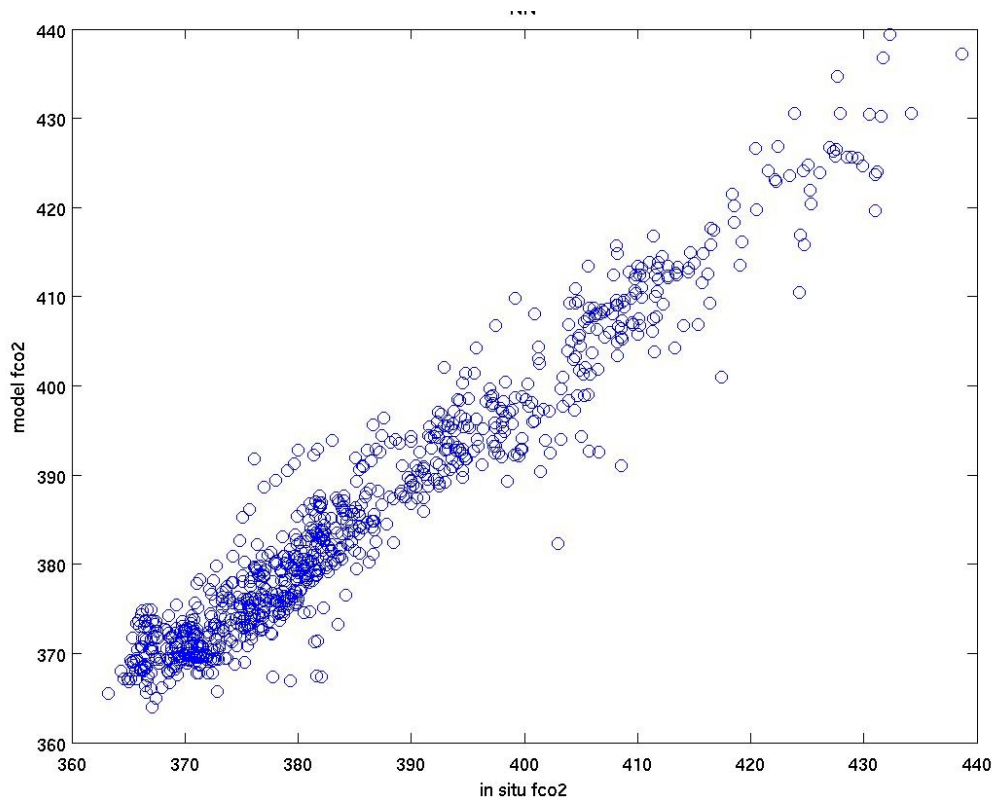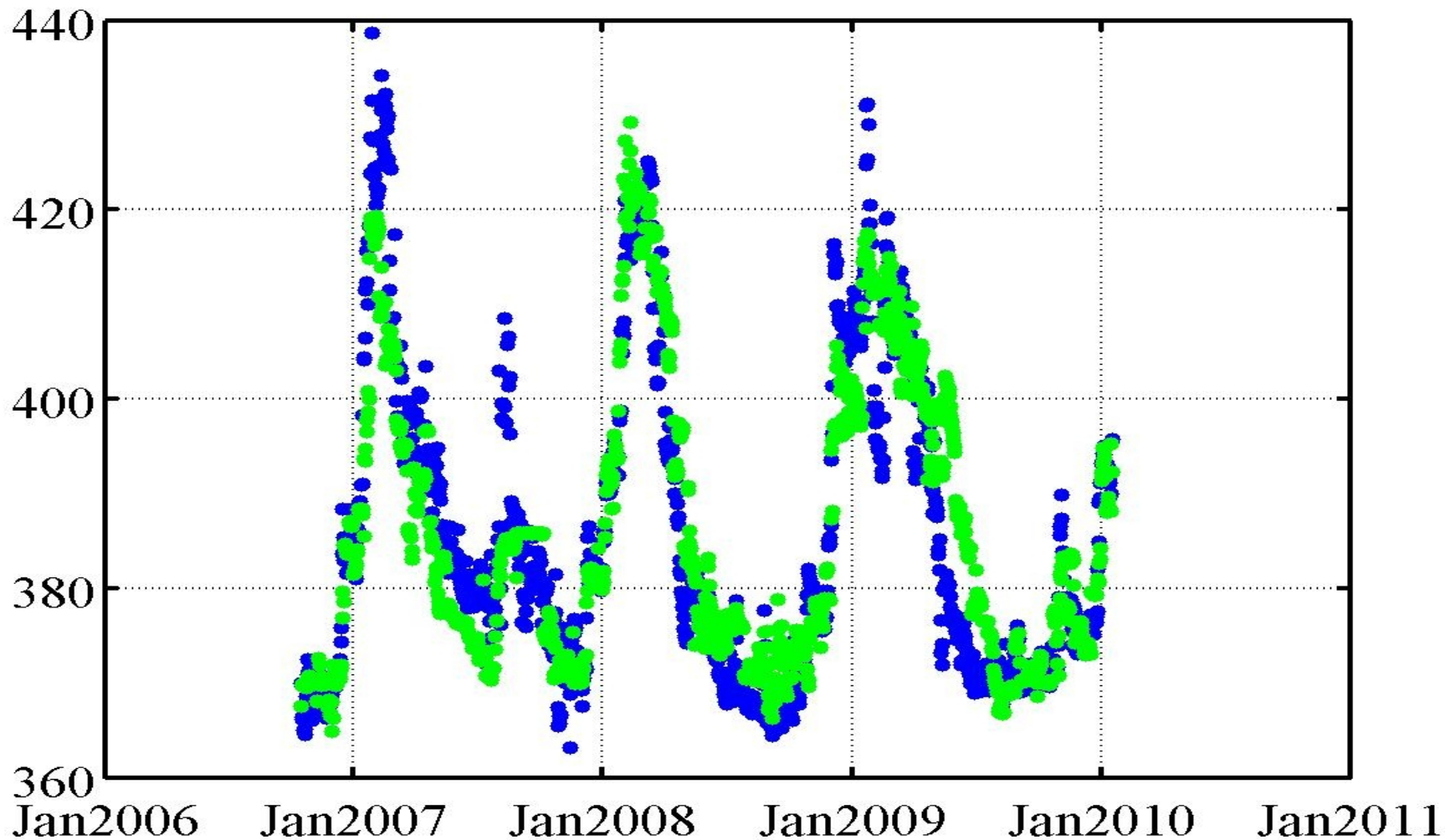
Stratus Time series pCO2 model   NN (SST, MLD, SSS, NPP)
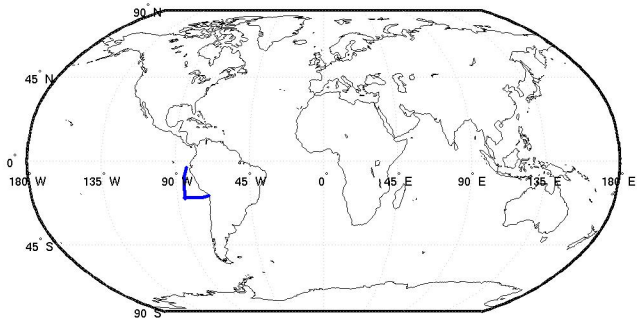
UNIVERSITY
of NEW HAMPSHIRE

Ifremer

assessment with ship data – near Stratus
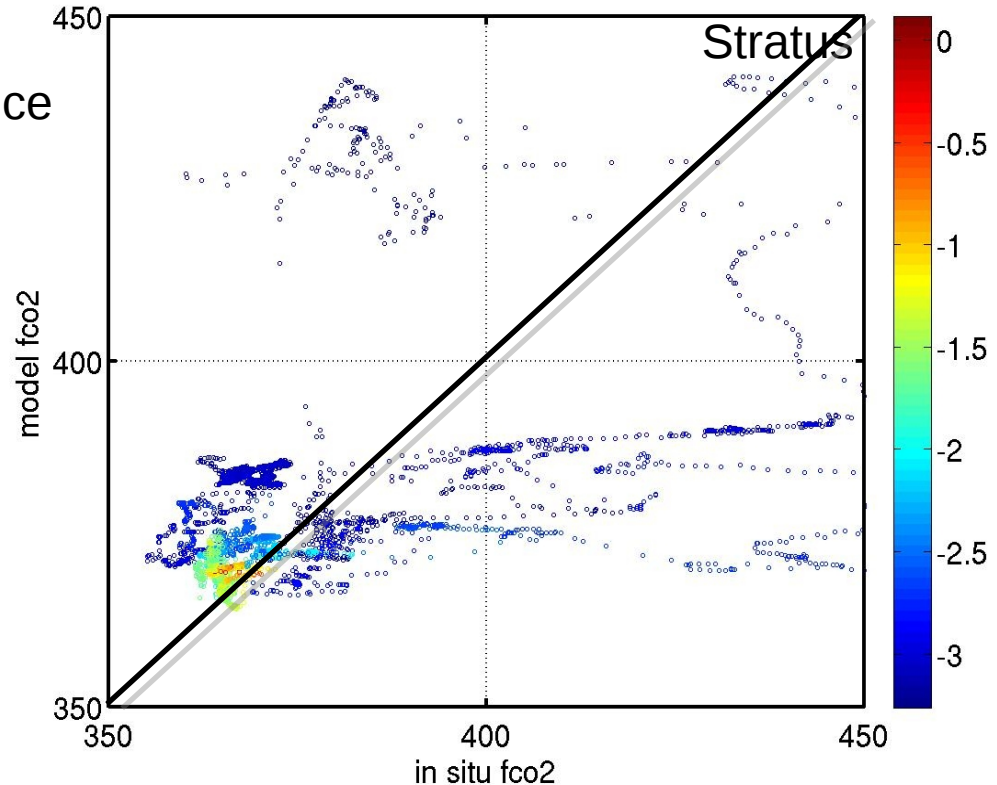
Use independent ship data surrounding the site with same satellite matchups

In situ vs. predicted colored by distance (log10(distance^-1)

agreement not impressive

VOCALS cruise data

University of New Hampshire

Ifremer

assessment with ship data – near BATS

Use independent ship data surrounding the site with same satellite matchups

In situ vs. predicted colored by distance (log10(distance^-1)

SOCAT + Clivar cruise data

BTM
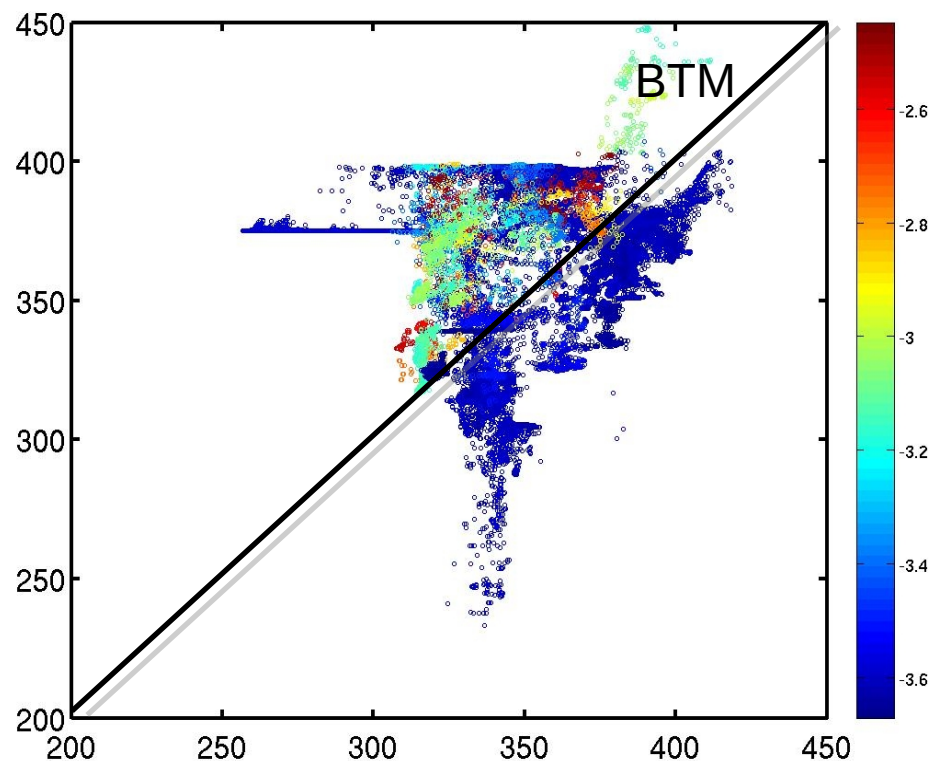
further assessment required

UNIVERSITY of NEW HAMPSHIRE

Ifremer

Future? Multi-site algorithm
why, why not, how successful at each site and elsewhere?

Pooling time series data and model training to create a single multi-site algorithm

RMS goes up factor of 1.5-2 at each site (Still all less than 8 uAtm)

NN - all data: N=2707
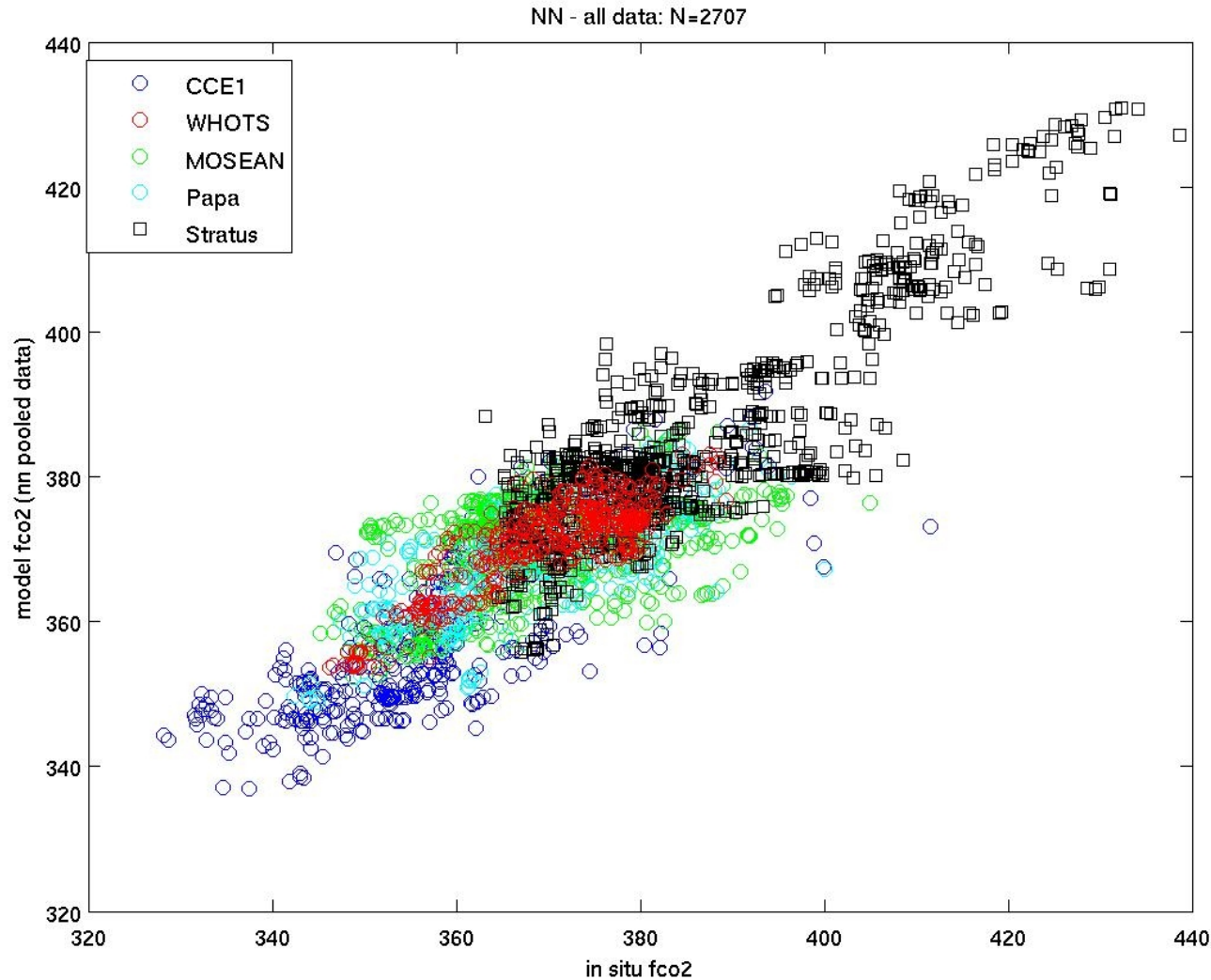
Future? - Liu et al. support vector regression algorithm**

- Adding geography and time in the input suite

- 16 uAtm rms against withheld validation data

- Very dependent on good pCO2 data coverage in space/time (200 k + SOCAT samples used)

❑Input (3-day): sin(day), cos(day), lat, sin(lon), cos(lon), SST (AMSR-E), Chl-a (SeaWiFS+MODIS TERRA+MODIS Aqua), SSS (Levitus climatology), Mixed layer depth (GODAS).

** http://aquarius.umaine.edu/docs/aqsci2012_WGC-02-Timothy.pdf

Future? - Liu et al. support vector regression algorithm**
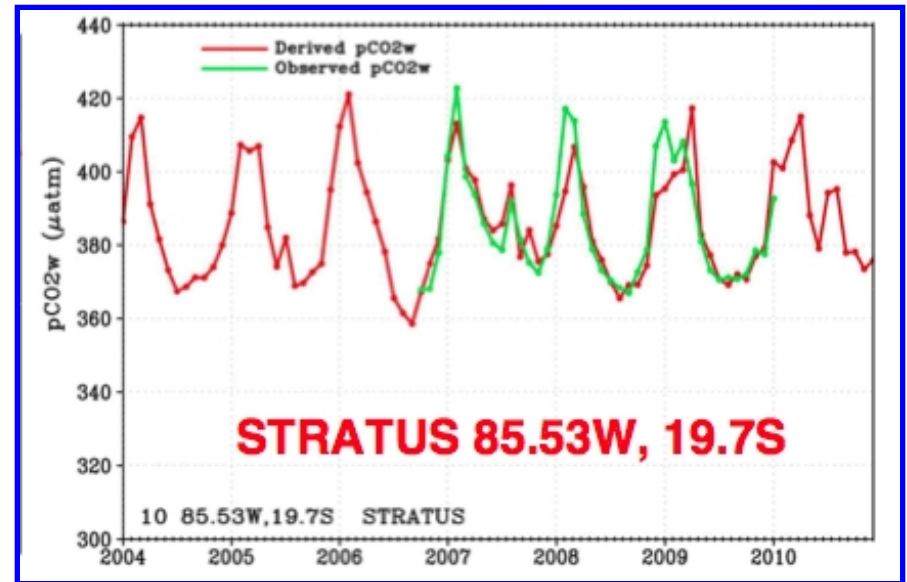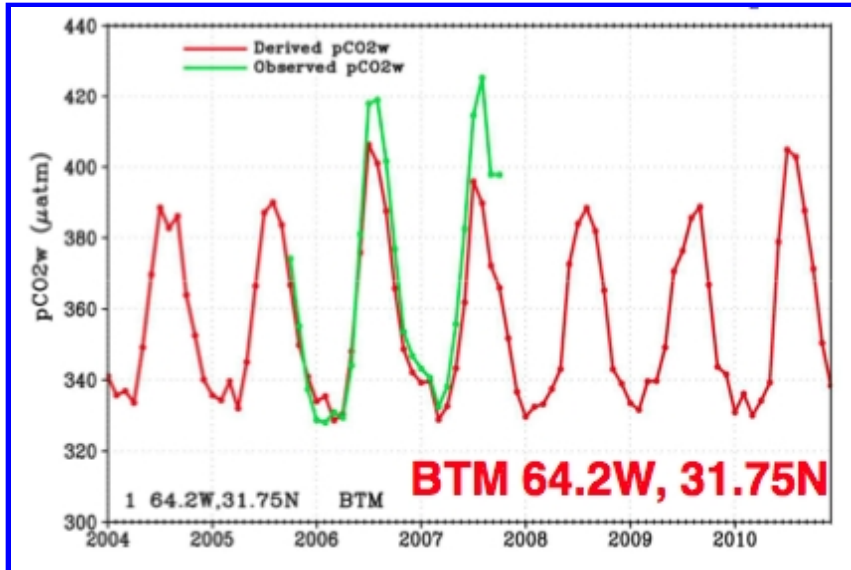
- Doing quite well even on time series nodes that are not in the data set
- Some 10-15 uAtm overshoot in summer

** http://aquarius.umaine.edu/docs/aqsci2012_WGC-02-Timothy.pdf

Summary

- First steps at time series pCO2 sites and their utility for improving ocean pCO2 prediction models
- Contributions - a) methodical means for evaluating satellite inputs and their value in the inversion, b) revisiting pCO2 training and validation datasets
- Neural networks outperforming Multiple Linear regression
- After SST and MLD, improvements with SSS, satellite-derived NPP (PAR)
- Our site-based algorithms yield < 5-6 uAtm rms, all-site NN < 8 uAtm
- BUT do they not appear to travel well, even within the region **???**

FUTURE

- Can apply/expand any aspect within the OAFlux Cloud matrix to test further (e.g. SMOS SSS, GlobCoulour)
- Bringing in geography & time inputs + more data (cf. Liu approach) may indeed be the path forward at global scale (but is it better than climatology for SOLAS?)
- More gas flux evaluation at time series nodes?

*** We wish to acknowledge the use of field and satellite data from NOAA, WHOI, SOCAT, and the space agencies. ****

UNIVERSITY of NEW HAMPSHIRE

Ifremer