CALIBRATION OF MODIS SATELLITE IMAGES FOR CHLOROPHYLL-*a* ESTIMATES AGAINST AUTOMATED FLUOROMETER RECORDS WITH PLS ANALYSIS

Seppo Kaitala, Pasi Ylöstalo, Jukka Seppälä

Finnish Institute of Marine Research
Alg@line/Ferrybox

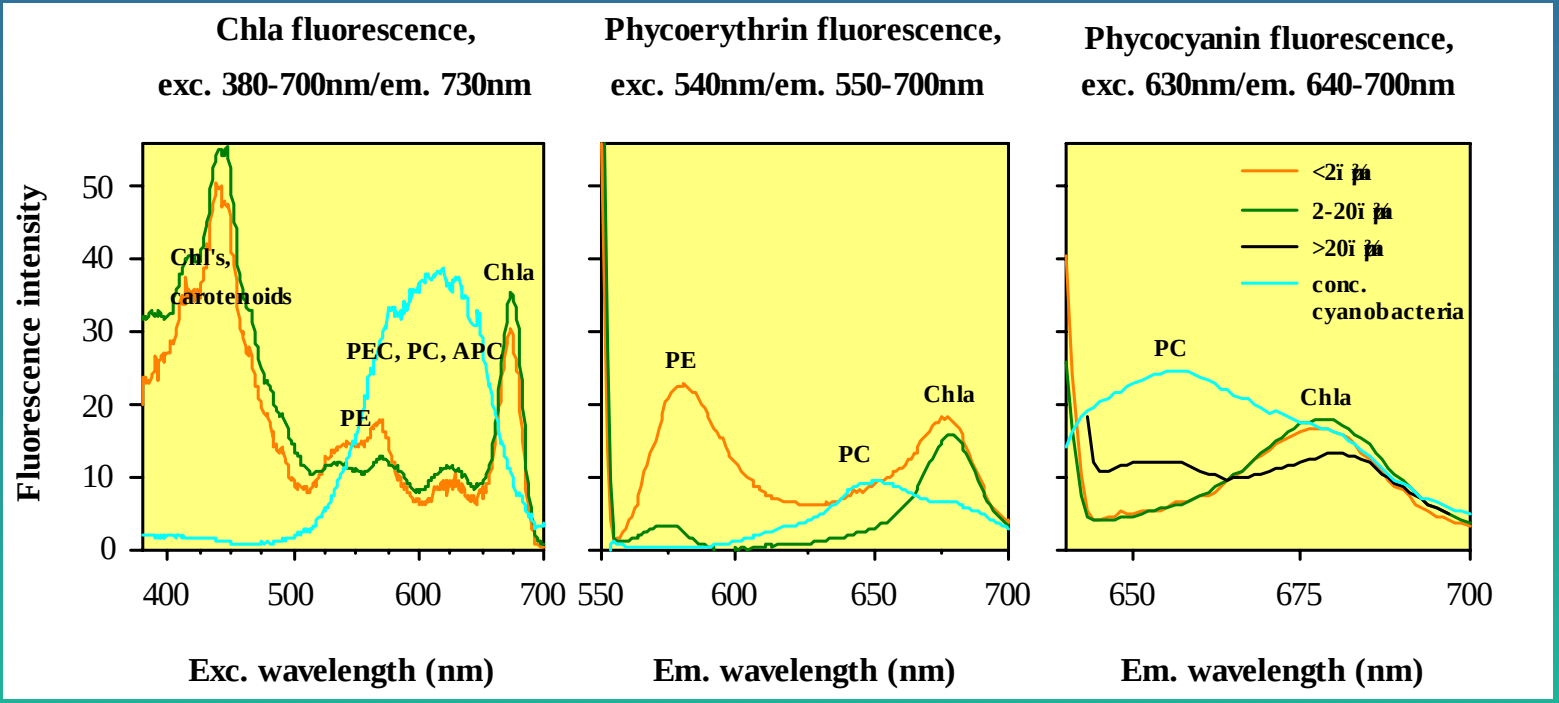
Conventionally chlorophyll-a estimates are obtained using empirical reflectance ratios

High concentrations of coloured dissolved organic matter (CDOM) and high turbidity due to high contents of suspended matter make the predicti more difficult

Specific structure of phytoplankton communities as blue-green algae blooms creates extra challenge for predictions

In these multicomponent cases more complex hyperspectral models are needed

Aranda cruise 2000, origins of phycobilin signals in the Baltic Sea



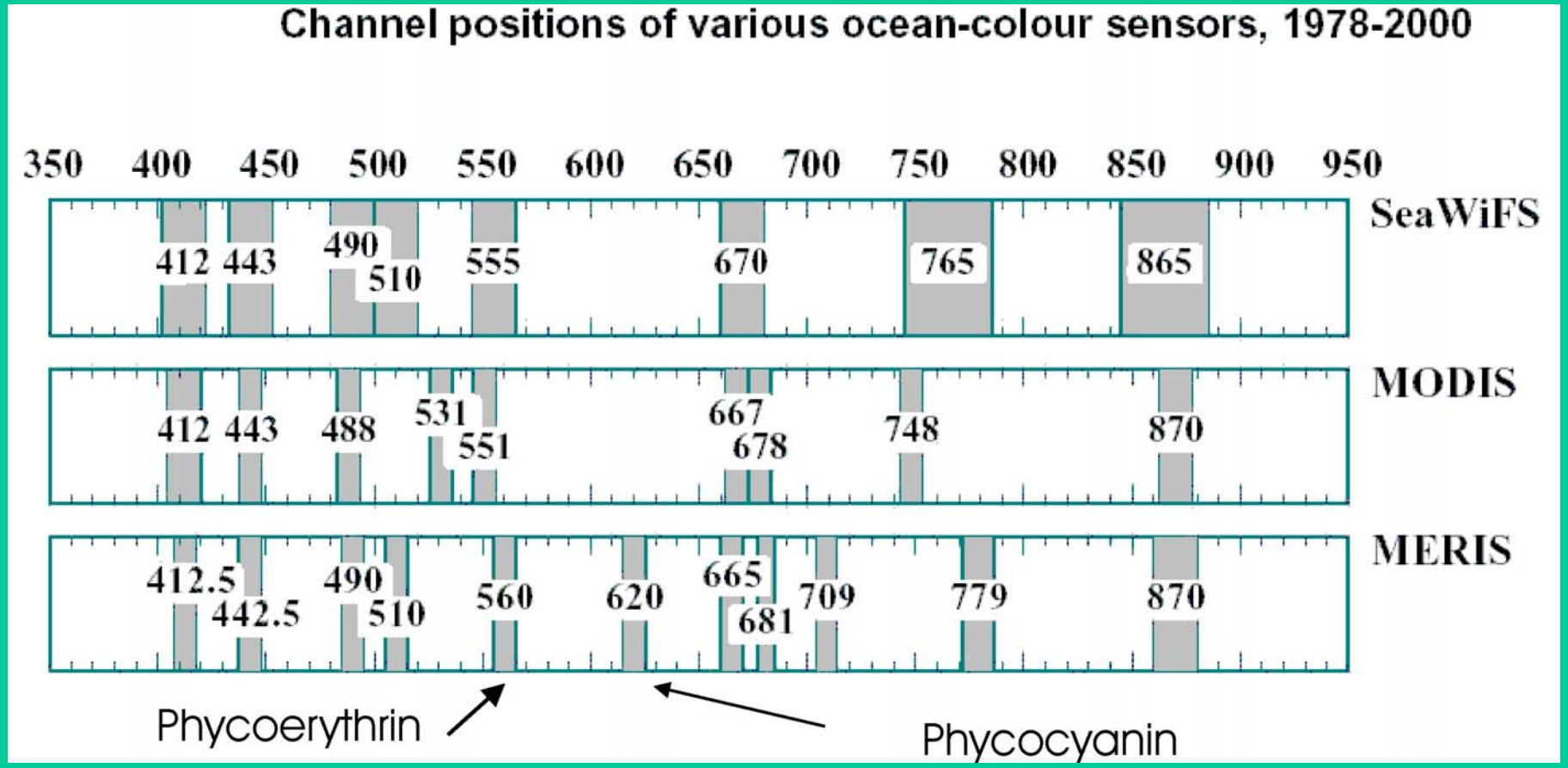
-Pico-sized cyanophytes, predominated by phycoerythrin rich cells, were main source of the phycoerythrin signal (average 74%, range 35-100%).
-Large filamentous cyanobacteria, in turn, were responsible for the phycoerythrocyanin and phycocyanin signals.

Multivariate calibration was applied to validate Modis satellite data against automated fluorescence records of chlorophyll-a on board the ferry Finnpartner with regular route from Travemünde to Helsinki (Alg@line data).

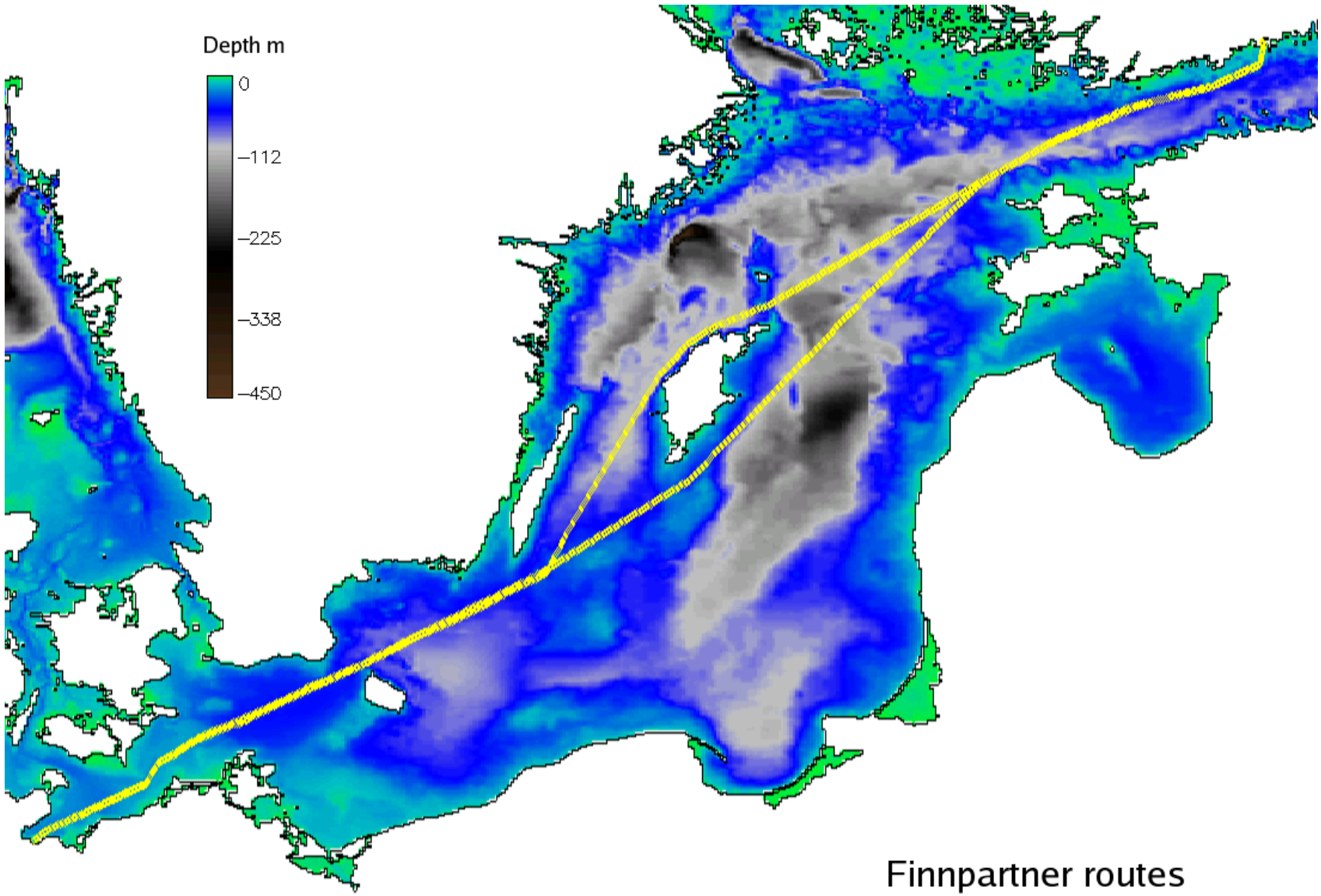
Partial least square (PLS) regression analysis was used to validate chlorophyll-a records against 1 km resolution bands. Satellite data was received from NASA GES Distributed Active Archive Center (DAAC) Data Pool through Internet.

http://daac.gsfc.nasa.gov/data/datapool/

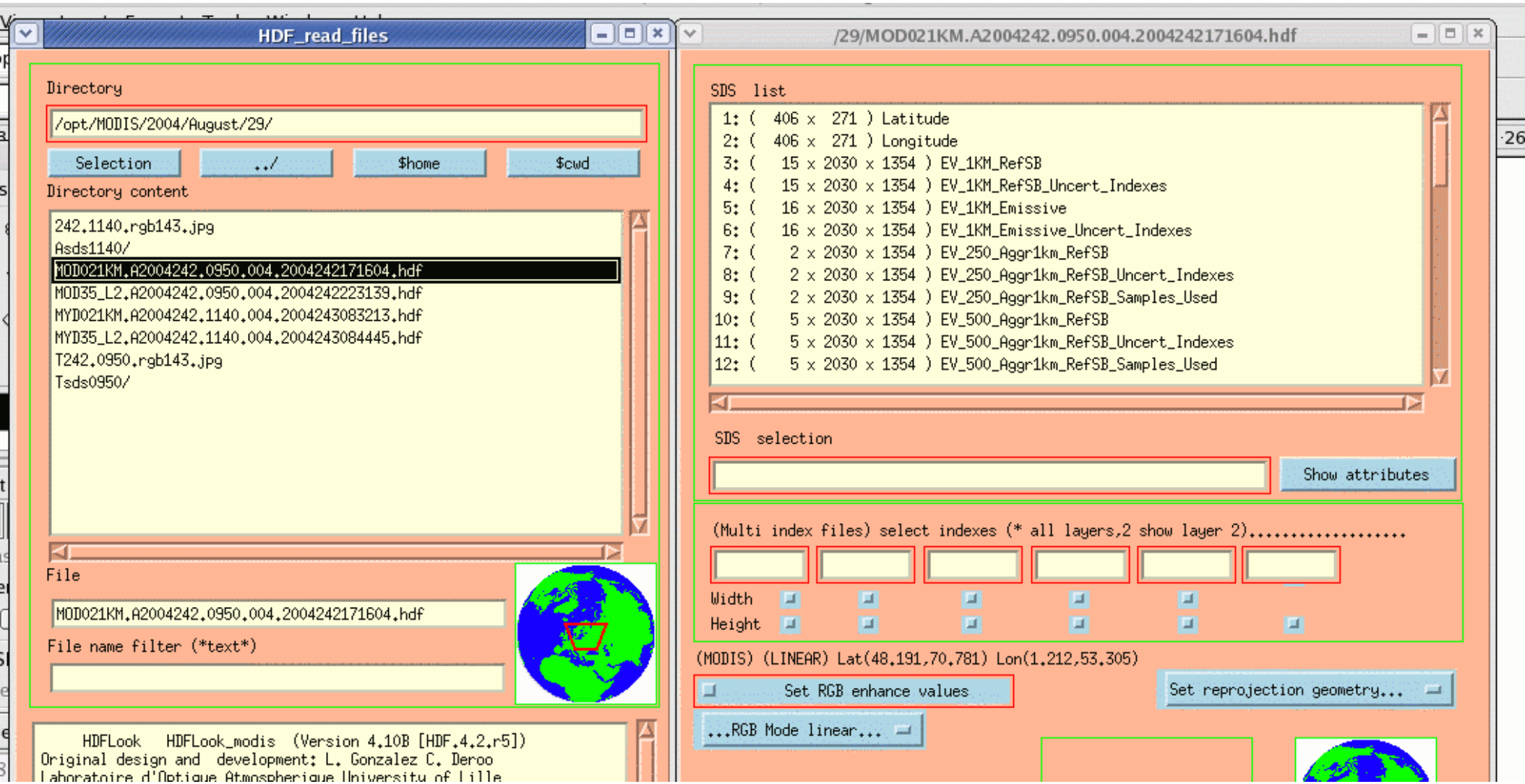
Seppo.Kaitala@FIMR.FI



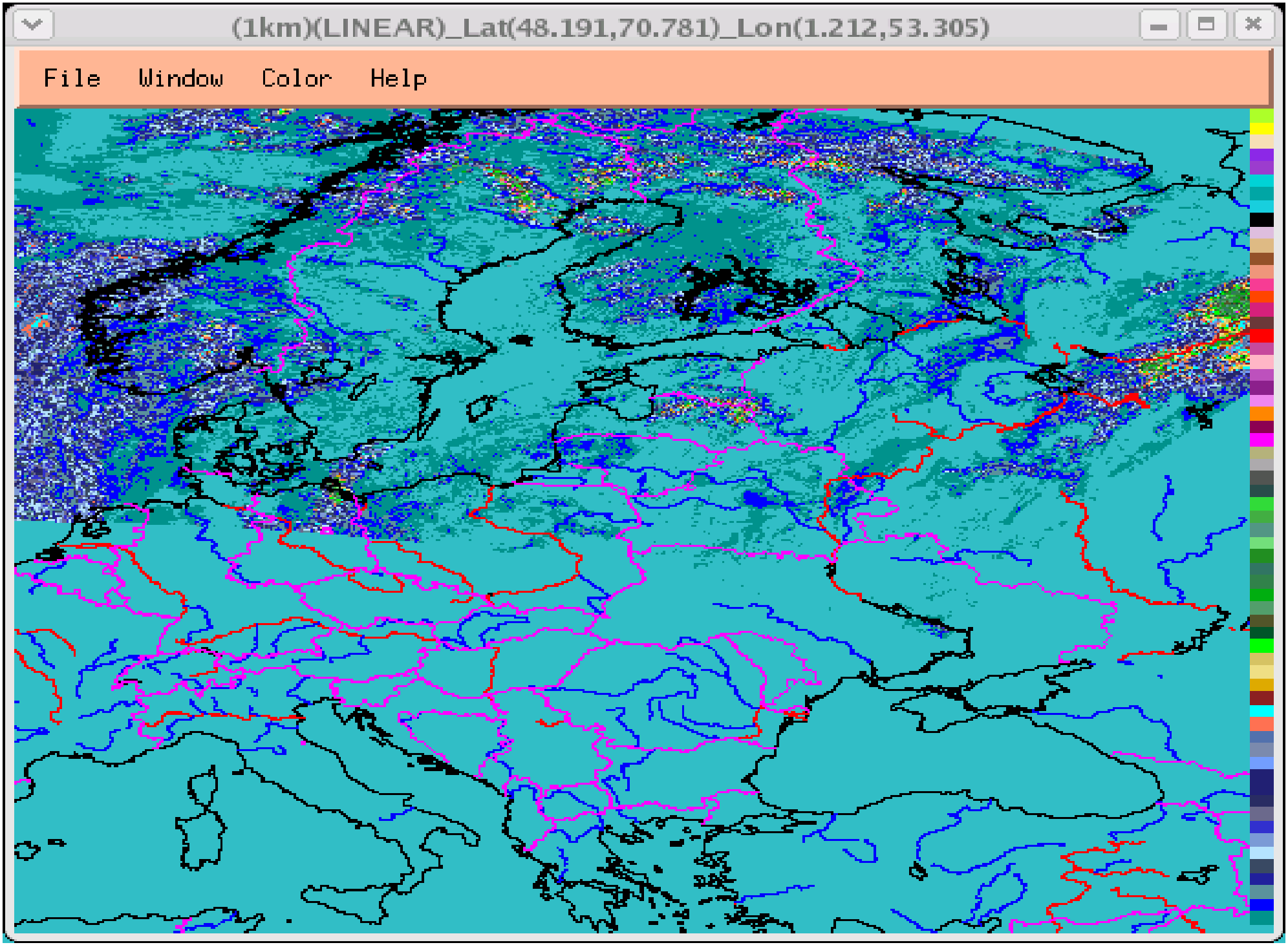
For validation the Modis/Terra data on the 28 of July, 2004 at 9:50 UTC was used and the chlorophyll-a records along the route ± 3 hours the satellite recording.



HDFLook, Version 4.1B, February 2004



http://daac.gsfc.nasa.gov/www/tools_services/HDFLook/

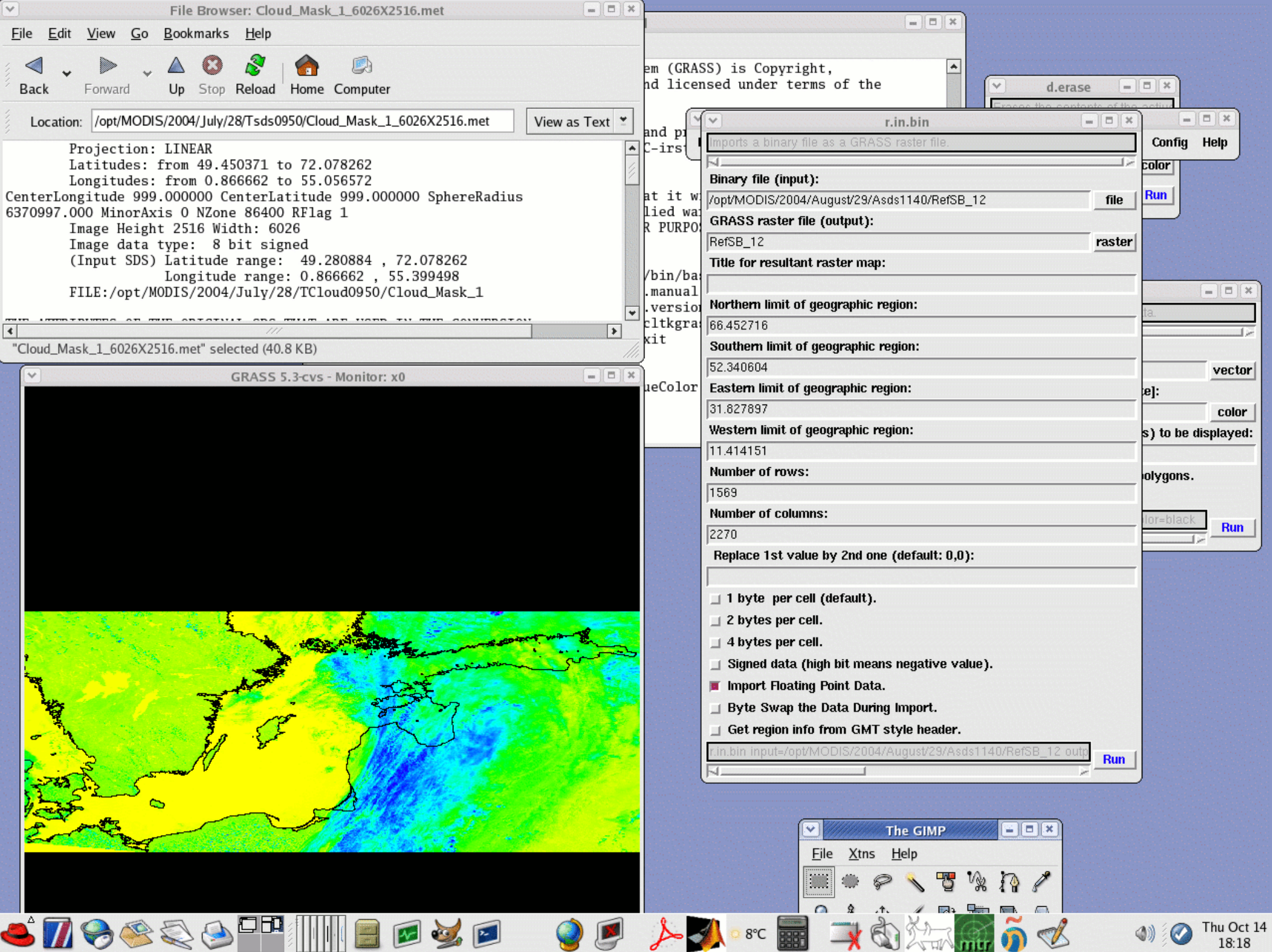

(1km)(LINEAR)_Lat(48.191,70.781)_Lon(1.212,53.305)
File
Window
Color
Help

Data for each band was extracted with HDFLook-Modis software and further analyzed together with chlorophyll-*a* data with GRASS GIS (Geographic Resources Analysis Support System) software.

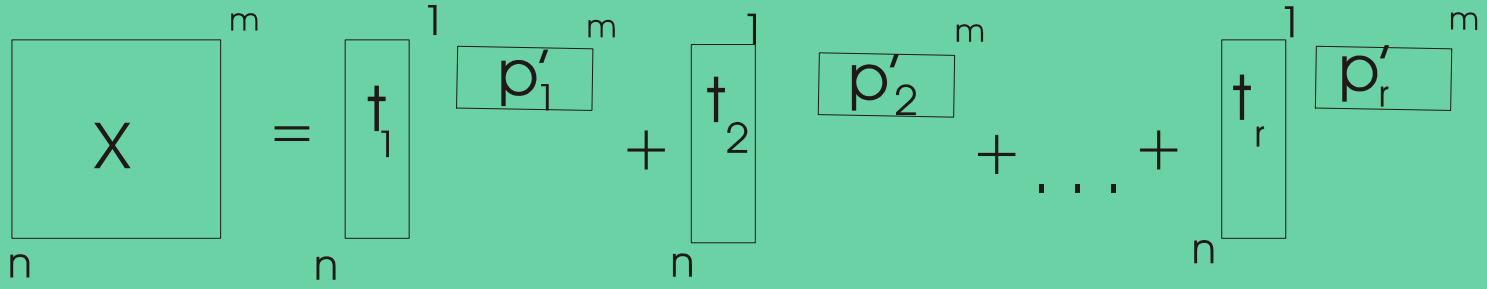
http://grass.navicon.dk/index.html

Statistical analysis was done with PLS and PCR package in the R statistical software.

http://www.r-project.org/


File Browser: Cloud_Mask_1_6026X2516.met
File Edit View Go Bookmarks Help
Back Forward Up Stop Reload Home Computer
Location: /opt/MODIS/2004/July/28/Tsds0950/Cloud_Mask_1_6026X2516.met
View as Text
Projection: LINEAR
Latitudes: from 49.450371 to 72.078262
Longitudes: from 0.866662 to 55.056572
CenterLongitude 999.000000 CenterLatitude 999.000000 SphereRadius
6370997.000 MinorAxis 0 NZone 86400 RFlag 1
Image Height 2516 Width: 6026
Image data type: 8 bit signed
(Input SDS) Latitude range: 49.280884 , 72.078262
Longitude range: 0.866662 , 55.399498
FILE:/opt/MODIS/2004/July/28/TCloud0950/Cloud_Mask_1
"Cloud_Mask_1_6026X2516.met" selected (40.8 KB)
GRASS 5.3-cvs - Monitor: x0
em (GRASS) is Copyright,
nd licensed under terms of the
d.erase
Config Help
color
Run
r.in.bin
Imports a binary file as a GRASS raster file.
Binary file (input):
/opt/MODIS/2004/August/29/Asds1140/RefSB_12
file
GRASS raster file (output):
RefSB_12
raster
Title for resultant raster map:
Northern limit of geographic region:
66.452716
Southern limit of geographic region:
52.340604
Eastern limit of geographic region:
31.827897
Western limit of geographic region:
11.414151
Number of rows:
1569
Number of columns:
2270
Replace 1st value by 2nd one (default: 0,0):
1 byte per cell (default).
2 bytes per cell.
4 bytes per cell.
Signed data (high bit means negative value).
Import Floating Point Data.
Byte Swap the Data During Import.
Get region info from GMT style header.
r.in.bin input=/opt/MODIS/2004/August/29/Asds1140/RefSB_12 outp
Run
vector
color
Run
The GIMP
File Xtns Help
8°C
Thu Oct 14
18:18

Partial least squares regression



t_1 = score

p_1 = loading (as cos *B*)

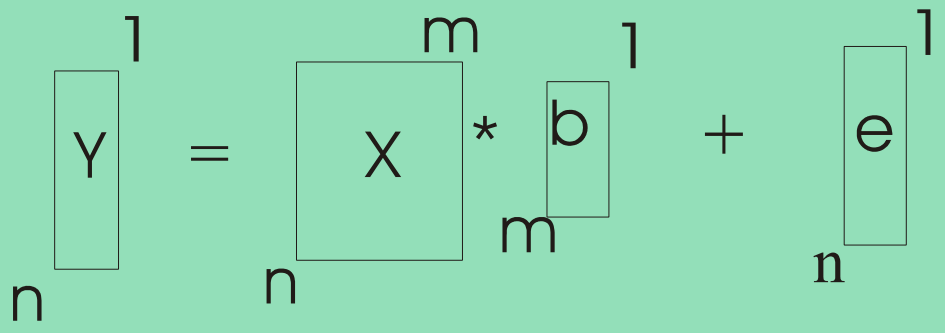
$$X = t_1 p'_1 + t_2 p'_2 + \ldots + t_r p_r$$

$$X = T P' + E$$

$$Y = U Q' + F$$

Least squares regression

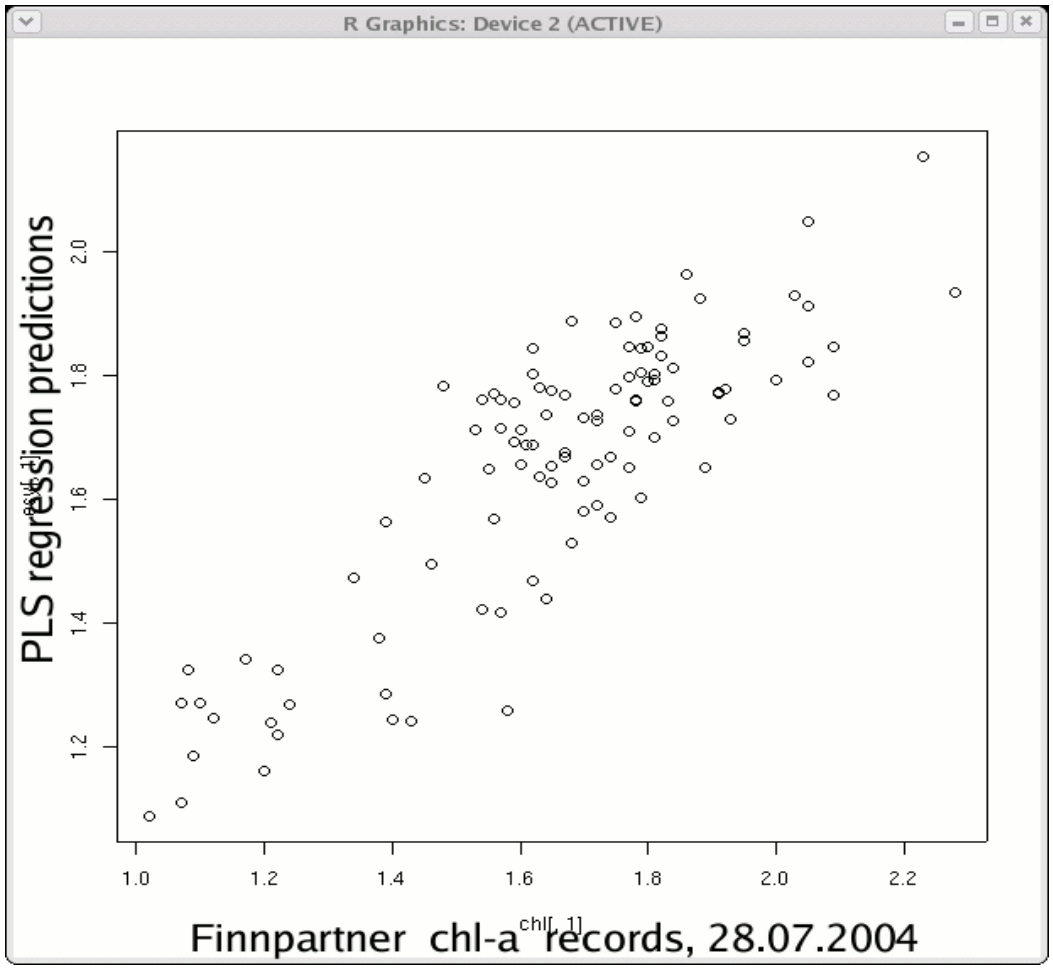
Y= observations (emission)
X=factors (exitation)
b=coefficients
n=no of observations
m= no of factors



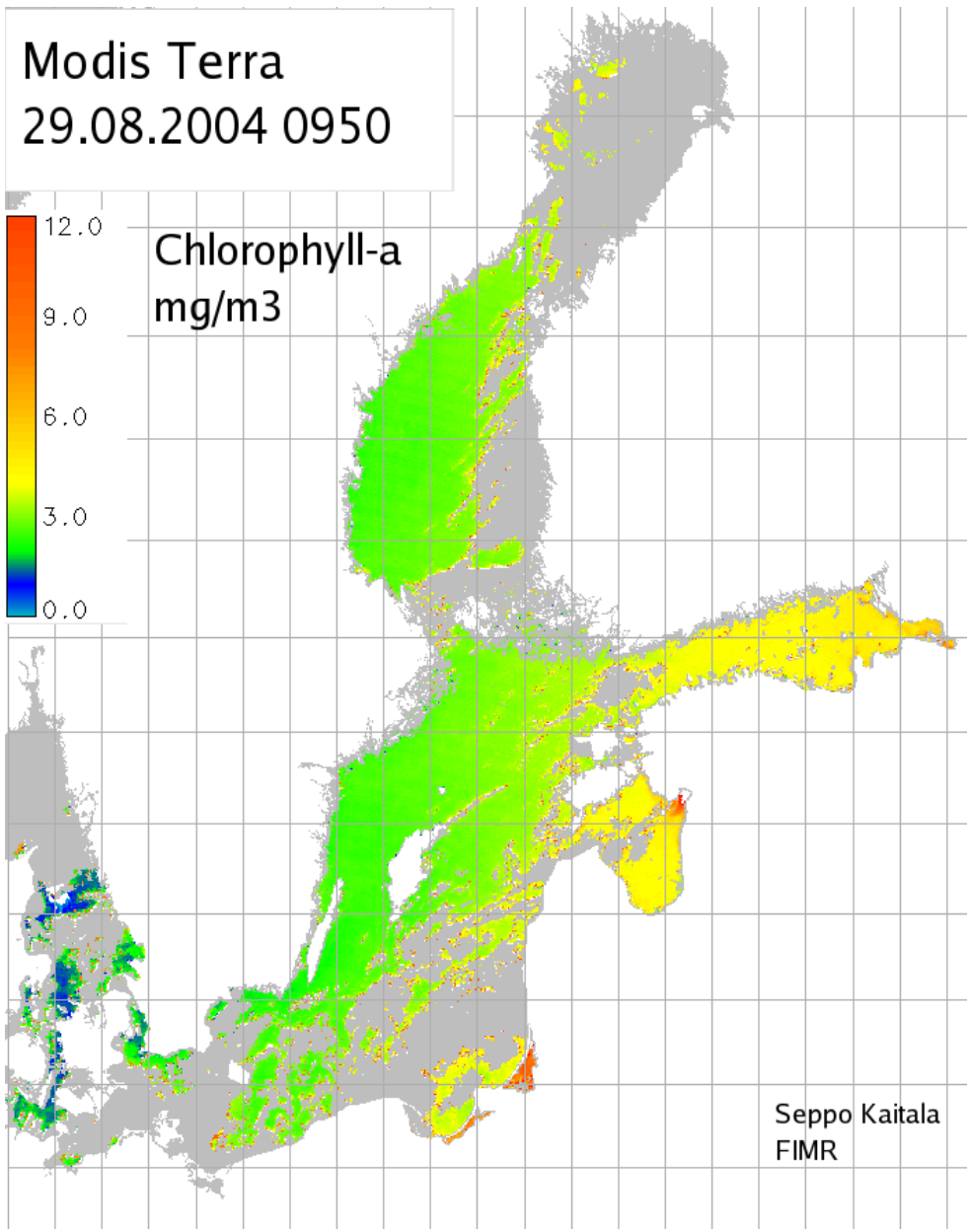
$$Y = XB + e$$

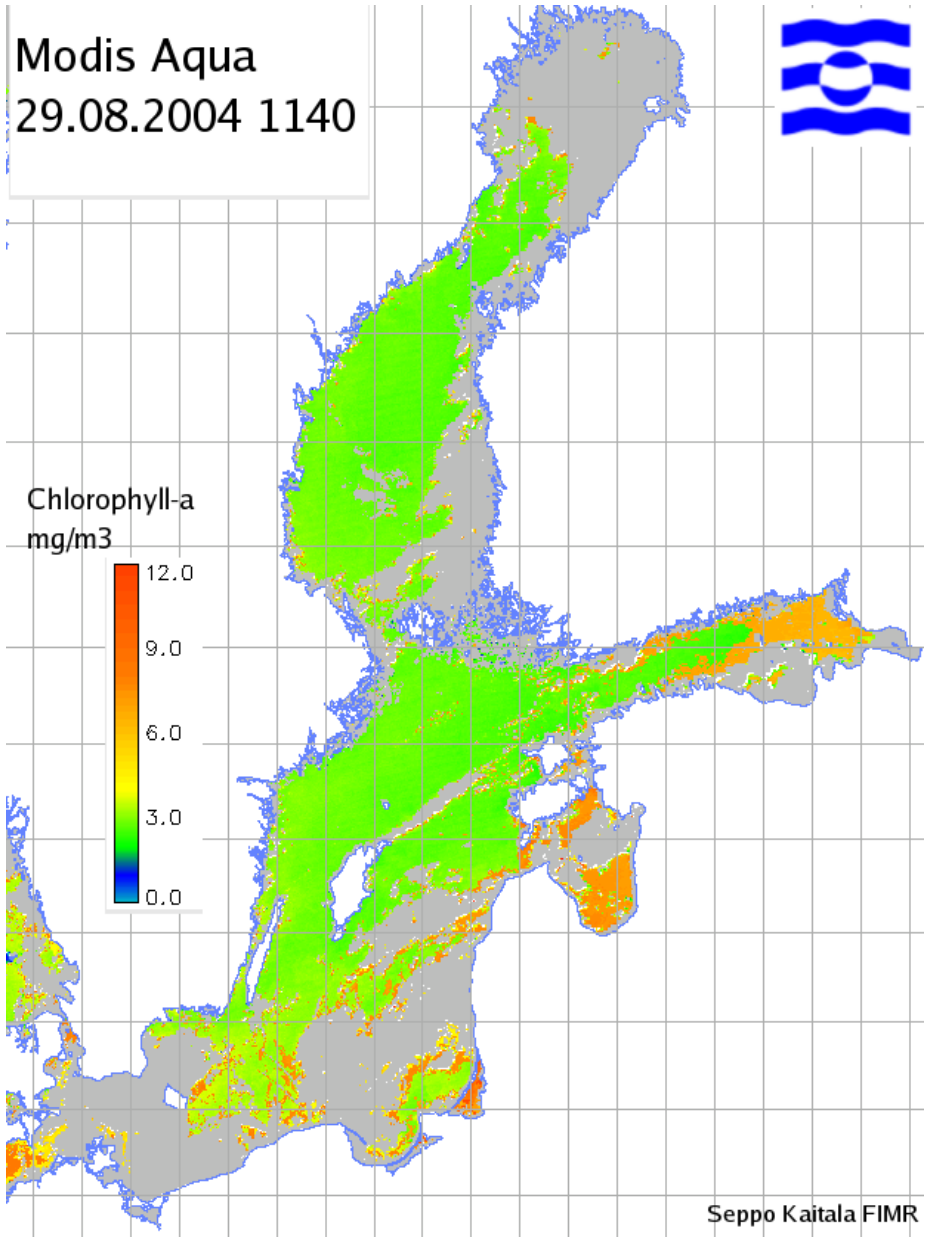
PLS analysis showed that only the bands with the wavelengths from 562 to 920 nm (ie b11, b12, b13L, b13H, b14L, b14H, b15, b16, b17) had contribution to chlorophyll-avariance.

The R^2 reached 72 % with 6 latent variables recommended for modelling.



The chlorophyll-a distribution maps evaluated according to the model are are shown for 29 August 2004.





Lesson learned

* multivariate validation of satellite data for chlorophyll a is needed

* Combination of GPL software is

- sufficient
- flexibe
- cost effective