EARSeL Lissig
7-9 Feb 2017
Bern

EVALUATION OF NORTHERN HEMISPHERE SNOW EXTENT PRODUCTS USING IN-SITU DATA WITHIN ESA SNOWPEX PROJECT

Sari Metsämäki[1)], *Elisabeth Ripper*[2)], *Olli-Pekka Mattila*[1)] , *Gabriele Schwaizer (nee. Bippus)*[2)] **,** ***Kari Luojus***[4)], *Thomas Nagler*[2)], *Richard Fernandes*[3)], *Bojan Bojkov*[5)] , *Michael Kern*[6)]

1) Finnish Environment Institute
2) ENVEO IT GmbH
3) Canada Center of Remote Sensing
4) Finnish Meteorological Institute
5) Eumetsat (earlier at ESA ESRIN)
6) ESA/ESTEC

SnowPEx general mission

- **Is to evaluate the performance of Northern Hemisphere and regional snow products by different providers**
  - by comparison against in-situ observation networks
  - by intercomparing the products
  - by compiling the analyses for different snow zones and separately for forests and non-forests
  - by investigation the trends in snow evolution when applying different products in the analyses
- **Several data providers from U.S., Canada, Asia and Europe have contributed by providing their snow products for the evaluations**
- **In-situ observations network comprises the extensive dataset from ECMWF, RIHMI (Russian Research Institute for Hydrometeorological Information), from U.S. (different sources collected by Rutgers University) and from Europe ECA&D and SMHI.**
  - In addition, data from smaller field campaigns were available for SnowPEx (Norway, Canada)

Snow Extent (SE) products evaluated in this study

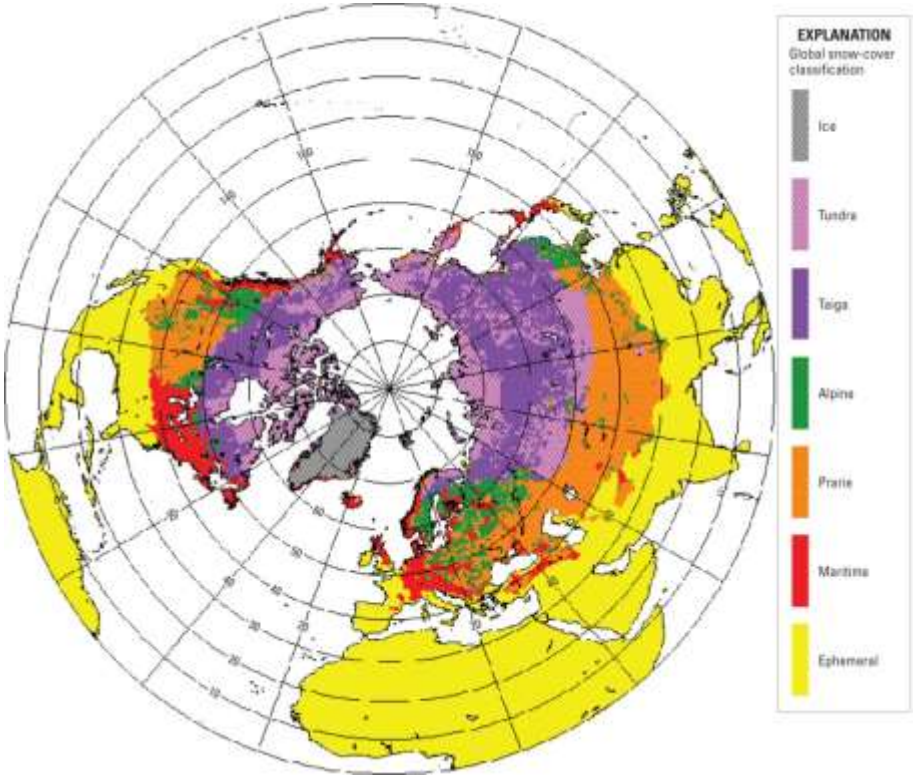
SE may be binary (snow/no-snow) or snow coverage (SCF = Snow Cover Fraction)

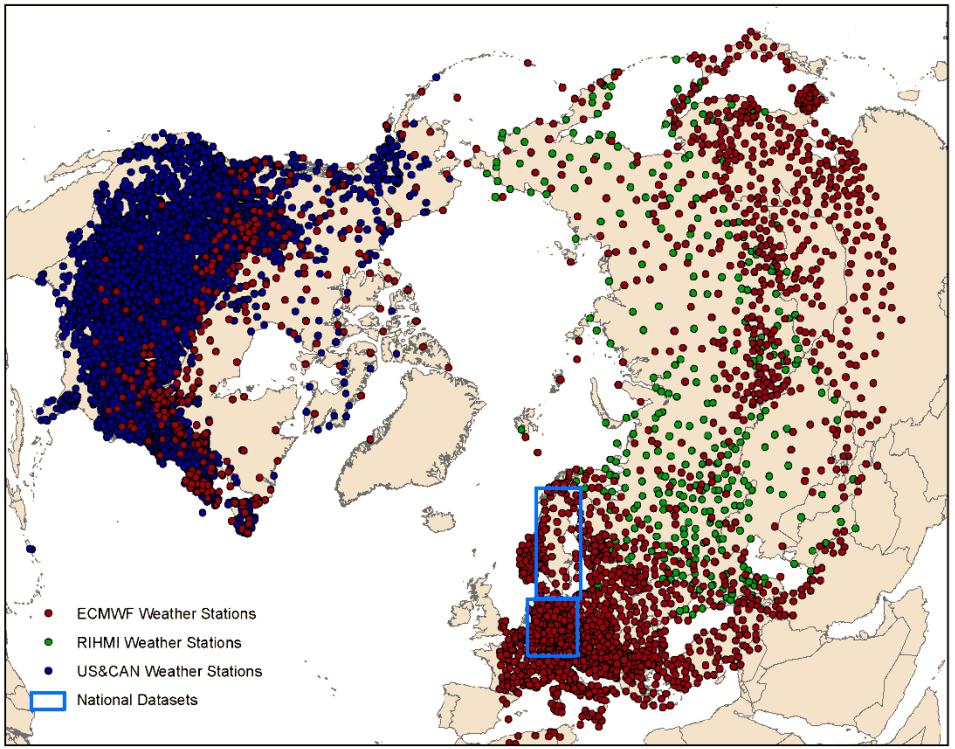
Product acronym	Product	Resolution	Thematic inform.
ASNOW	Autosnow	4km	Binary
CRCLIM	CryoClim	5km	Binary
GLSSE	GlobSnow v2.1	1km	SCF
IMS04	NOAA IMS	4km	Binary
JXAM5	JASMES GHRM5C	5km	Binary
JXM10	JASMES MDS10C	5km	Binary
M10C05	MOD10_C5	0.5 km	SCF
SCAG, SCAGCA	ModSCAG, ModSCAG canopy adjusted	0.5 km	SCF

in-situ observations (SD = Snow Depth) used in the study

Dataset	SnowPEx-ID	Region	Snow Class	Method	Available Time Period	Temporal Resolution	Parameter
Pointwise data							
ECMWF Weather stations	snpx.ws.ecmwf	Europe	All	Sonic snow dept, Manual surveys	1978-2014	Daily	SD
RIHMI Weather stations	snpx.ws.rihmi	Russia and former USSR	All	Manual surveys	1966-2011	Daily	SD, FSC
North American Weather stations (Unite States & Canada)	snpx.ws.uscan	United States and Canada	All	Sonic snow depth, Manual surveys	1978-2014	Daily	SD
ECA&D Weather stations (Germany)	snpx.ws.ecad	Germany (+ Europe)	All	Sonic snow depth, Manual surveys	2000-2012	Daily	SD
SMHI Weather station data (Sweden)	snpx.ws.smhi	Sweden	Mountains, taiga	Sonic snow depth, Manual surveys	1980-2015	Daily	SD
NVE snow stations (Norway)	snpx.snst.nve	Norway	All	Automated stations	1967-2015	Hourly/ Daily	SD, SWE

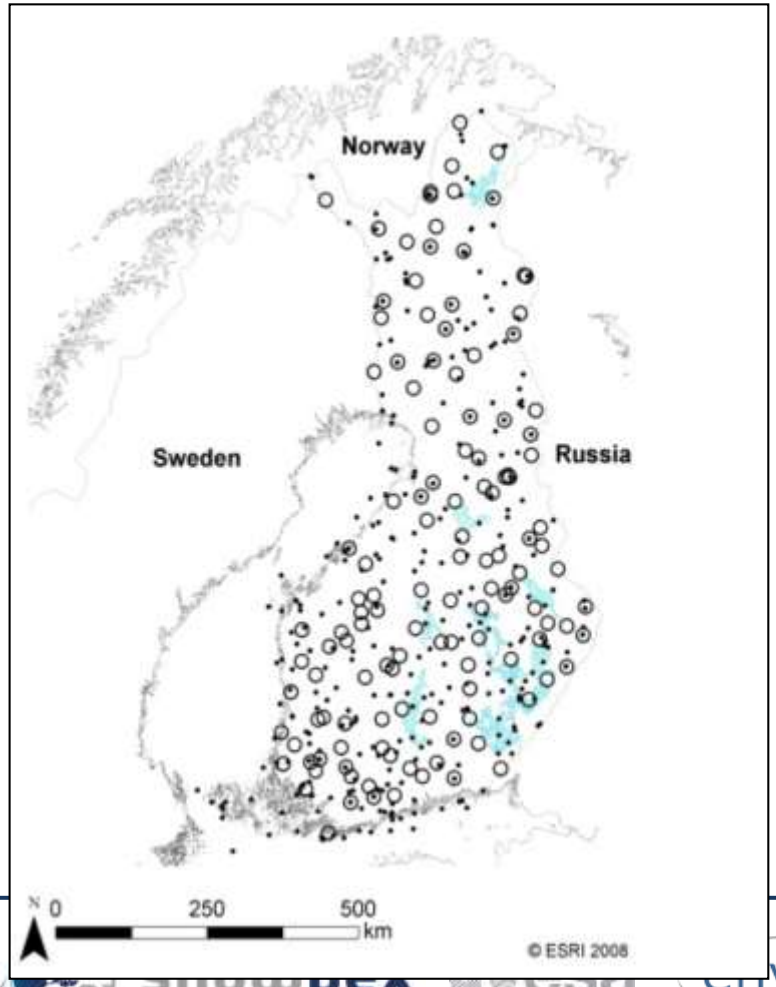
In-situ data distribution and Snow Zones (According to Sturm snow classes)



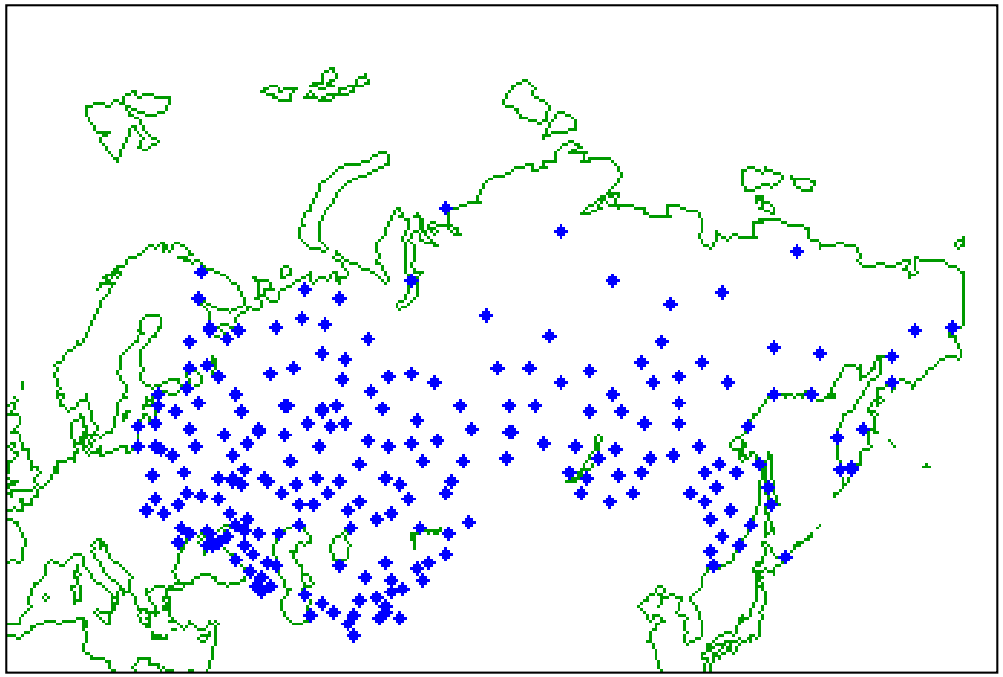


SCF Datasets
(SYKE Snow course and RIHMI snow coverage)

SYKE snow courses



RIHMI weather stations



1st trial on evaluation method: a statistical approach (by CCRS) based on sampling the Cumulative distribution functions of 1) in-situ observation and 2) snow product

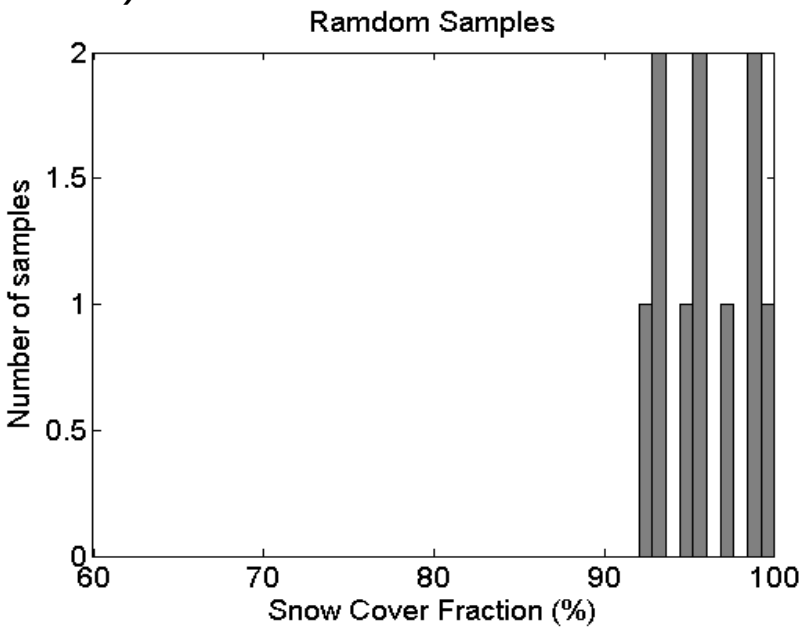
from the snow course SCF (ID 1510, April 15, 2004):

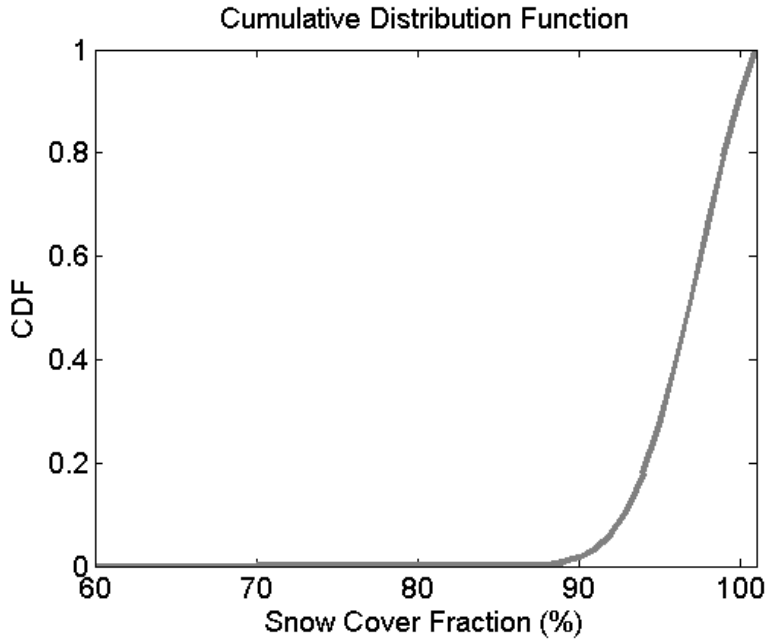
The provided average SCF and standard deviation are used to generate

1) probability distribution function (pdf)

2) Cumulative distribution function (cdf)

from cdf N samples are randomly generated (10 samples according to the protocol)

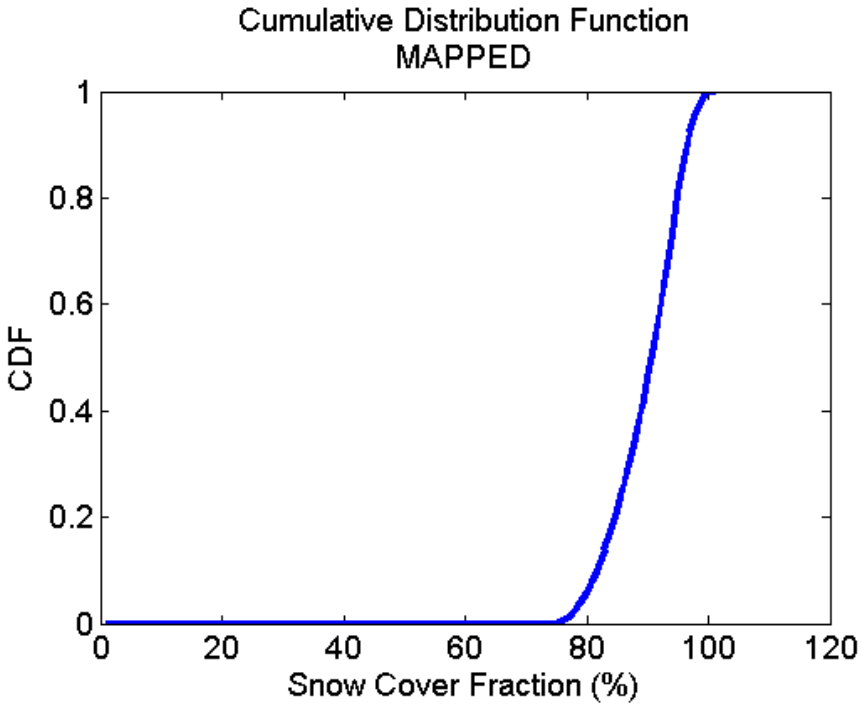


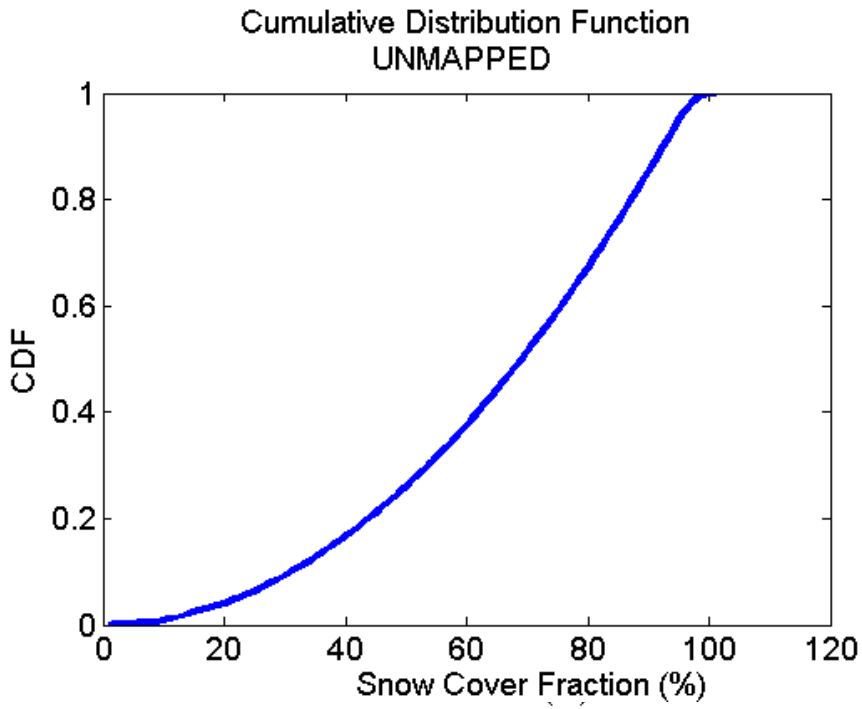


Generation of samples from direct SCF snow product for a validation cell

- **PDF for MAPPED** portion is generated using triangular function with mean at product SCF, and vertices according provider-gven precision
- **PDF for UNMAPPED** portion is generated using triangular function with mean at 50%, and vertices at 0 and 100%
- In the example case, mapped portion = 0.8889 → unmapped = 0.1111

Cumulative Distribution Functions are generated from PDF's, then 10 random samples are generated from these, weighting proportionally to %mapped and %unmapped



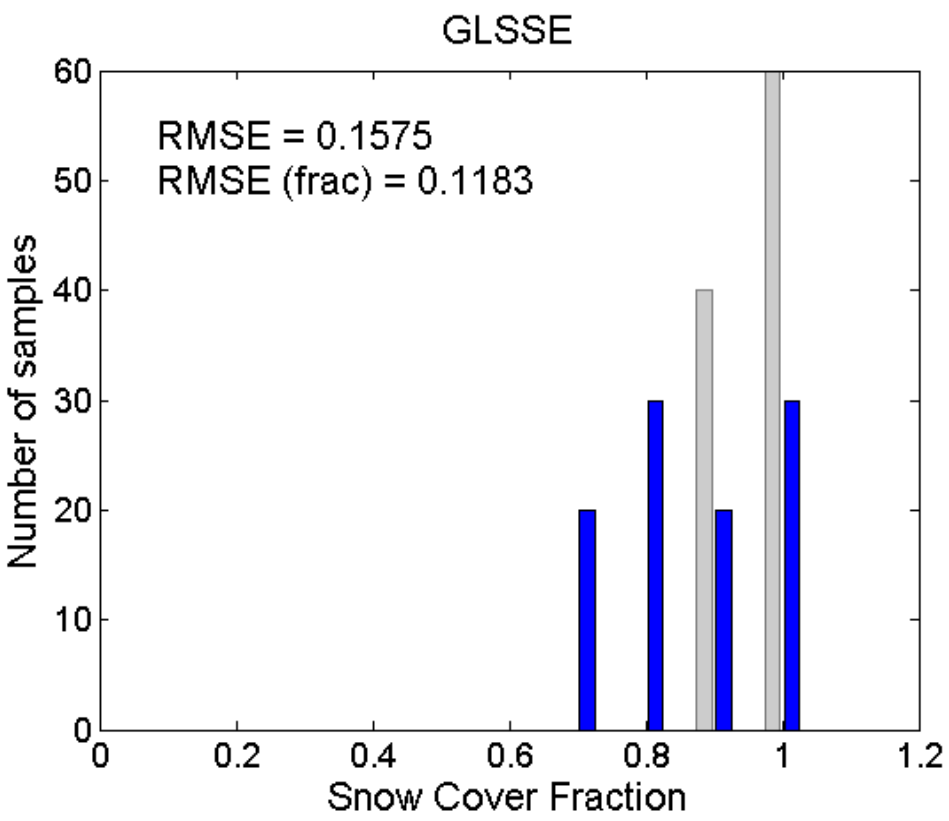


Generation of N*N cases from N in-situ realizations and N product realizations

10*10 combinations are taken from in-situ samples and product samples → 100 datapairs are generated.

Product samples	In-situ samples
0.8900	0.9200
0.7522	0.9700
0.6978	0.9500
0.7778	0.9900
0.8333	0.9300
0.9311	0.9900
0.8478	0.9600
0.8489	1.0000
0.9044	0.9300
0.7944	0.9600

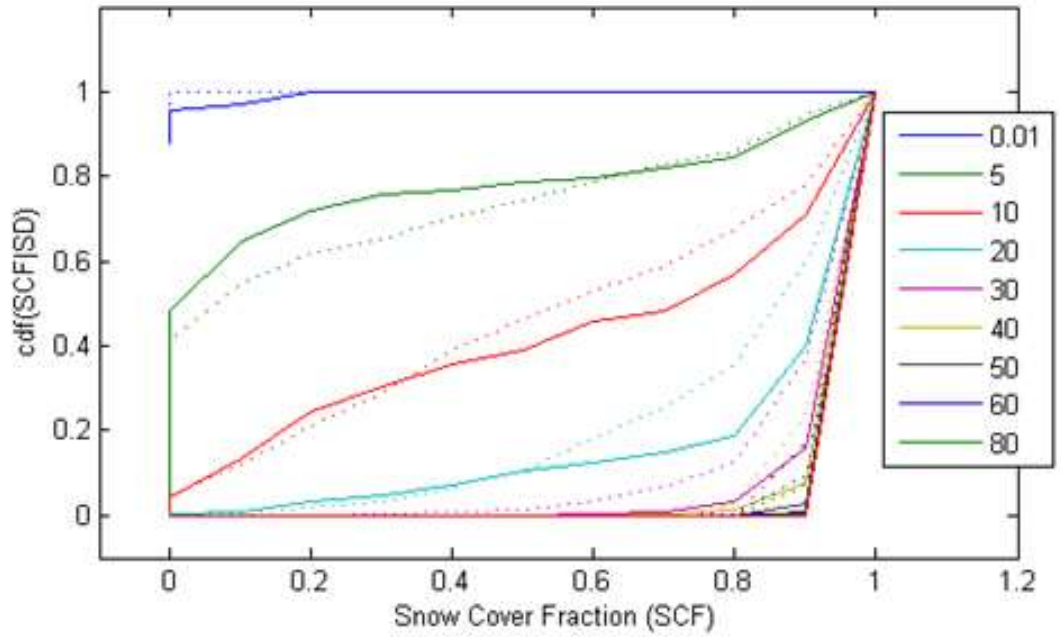
100 Samples
(binned in the figure)



Generation of in-situ samples from Snow Depth Observations

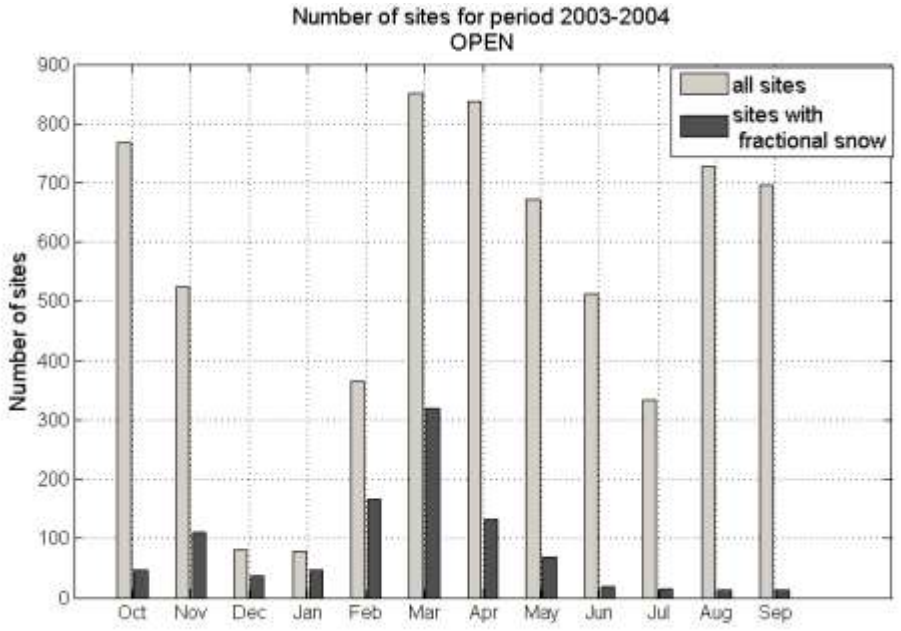
The ultimate majority of in-situ snow observations are on Snow Depth (SD) not on Snow Cover Fraction.

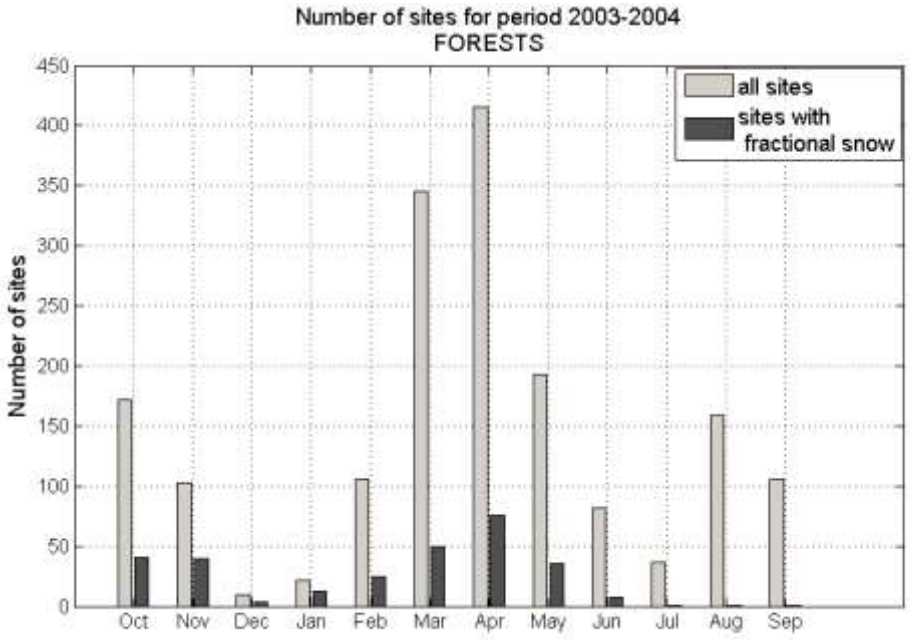
- SD is converted to FSC using the statistical relationship of these two, based on Finnish Snow course observation covering both these variables



Relevance to COST snow; reasons to select a more conventional validation

Example: Number of intersecting sites for the analysis with ECMWF in-situ data (period 2003-2004)





Most NH in-situ sites record Snow Depth: is the SD → SCF conversion valid outside Finland?? Obviously not!

We would need more information on the relationship between SCF and SD, for different snow types, different climate zones etc.

The finally chosen validation technique: a more conventional approach based on pixel-pixel comparison

- All snow products and in-situ data are treated as binary (snow/non-snow) information.
- For originally binary products no conversion is needed, but for products featuring Snow Cover Fraction, SCF is converted to binary by thresholding.
- In-situ snow depth is converted to binary information also by thresholding.
- Finally, a contingency table for binary data are formed and several binary metrics are calculated.
- To avoid the effect of the varying geolocation accuracy and of the location of in-situ site within a product pixel, only those cases where the 4-neighbourhood of in-situ site is homogenous (represents the same binary value), are accepted for the analyses

Validation techniques

Criteria for products
if SCF≥25% then "snow" else "no snow"
if SCF≥15% then "snow" else "no snow"

Criteria for in-situ
if SD>0cm then ”snow” else ”no snow"
if SD≥2cm then "snow" else "no snow"
if SD≥15cm then "snow" else "no snow"

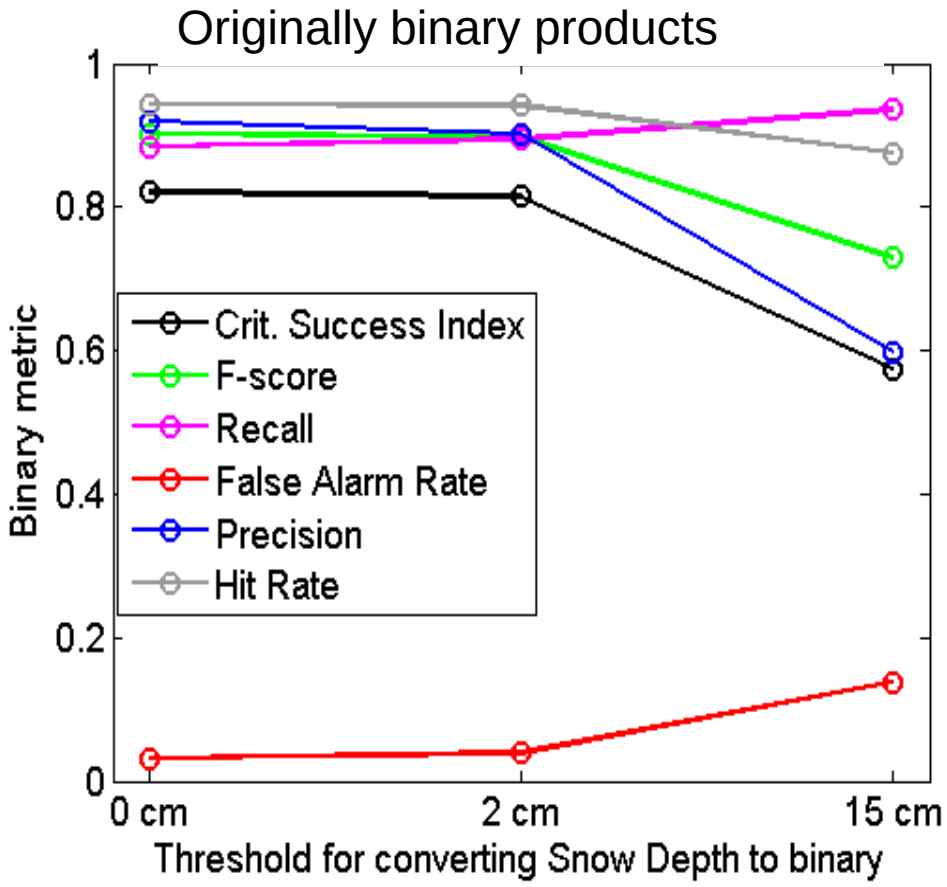
in-situ	Product	
	snow	No-snow
snow	True Positives (TP)	False Negatives (FN)
no-snow	False Positives (FP)	True Negatives (TN)

To binary metrics…

Binary metrics presented here

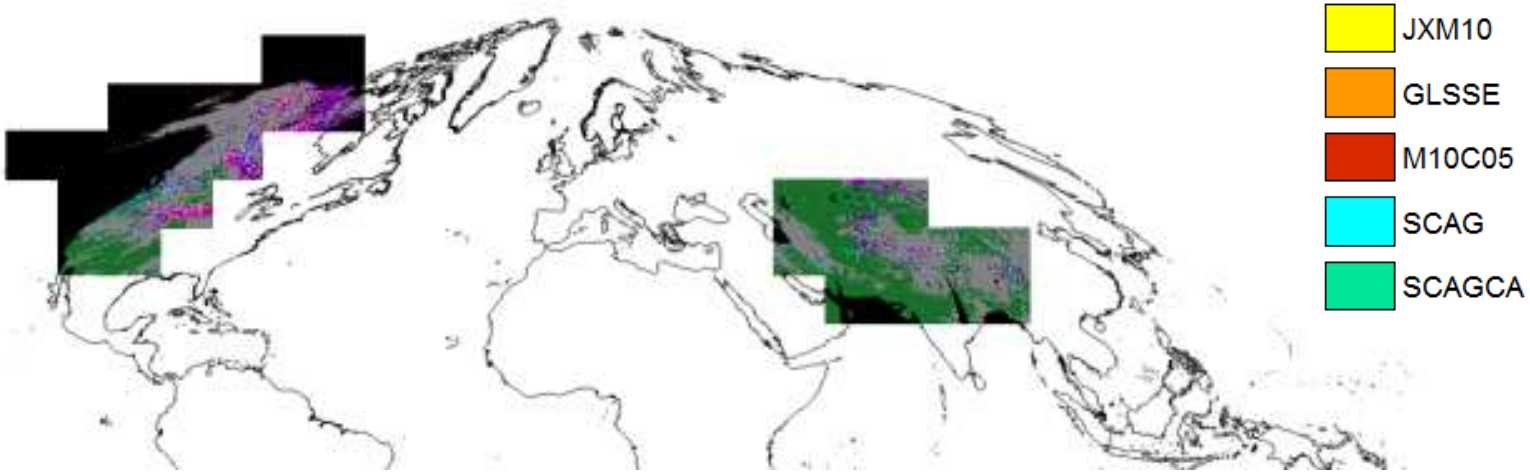
BINARY METRICS		
False Alarm Rate	**FP / (FP+TN)**	the tendency of SE-product to interpret snow-free areas as snow
F-score	**2*TP / (2*TP + FP + FN)**	the ability of SE-product to recognize truly existing snow, and a the same time, the ability to avoid interpreting snow-free areas as snow
Recall	**TP / (TP+FN)**	The ability of SE-product to identify all real snow cases

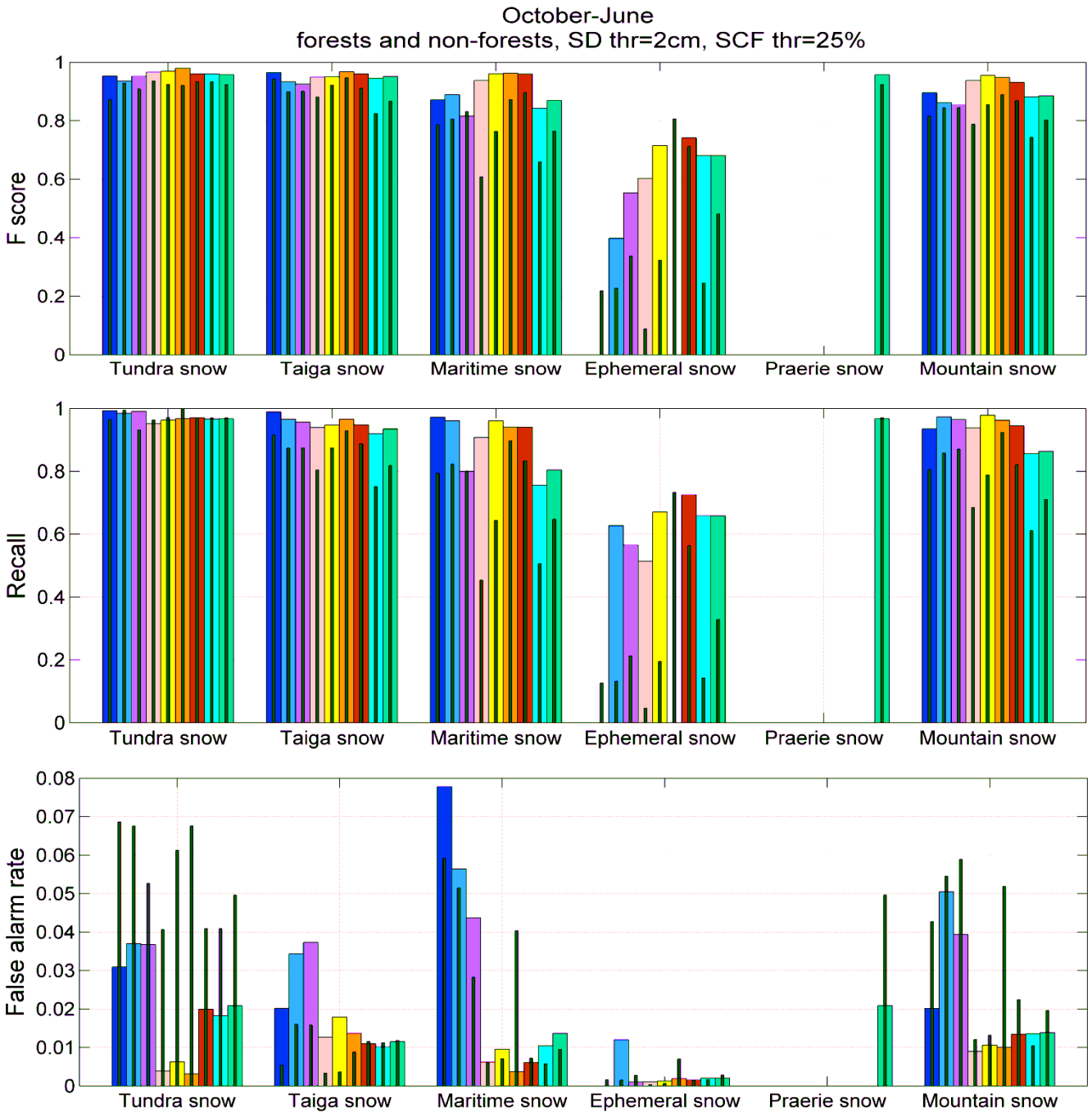
Effect on Snow Depth-threshold in SCF-> binary conversion for the in-situ data



Results for NH areas covered by MODSCAG products (Painter et al.)

- SCAG and SCAGCA are relatively new products and are provided only part of the NH.
- Included in the SnowPEx due to the uniuque new method *and* because SCF (not binary) is provided
- Only spatially and temporaly coinciding data are analyzed in the following





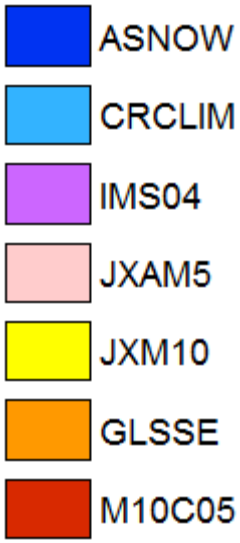
- Coloured bar: non-forest
- Black bar: forest

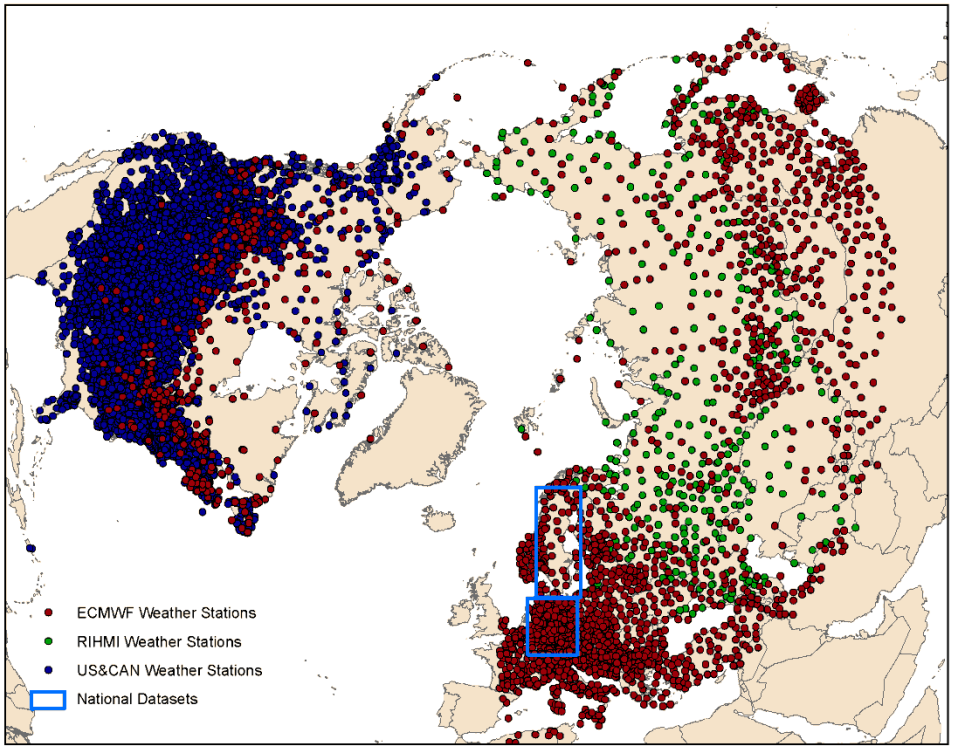
Summary and Conclusions for NH areas covered by SCAG and SCAGCA

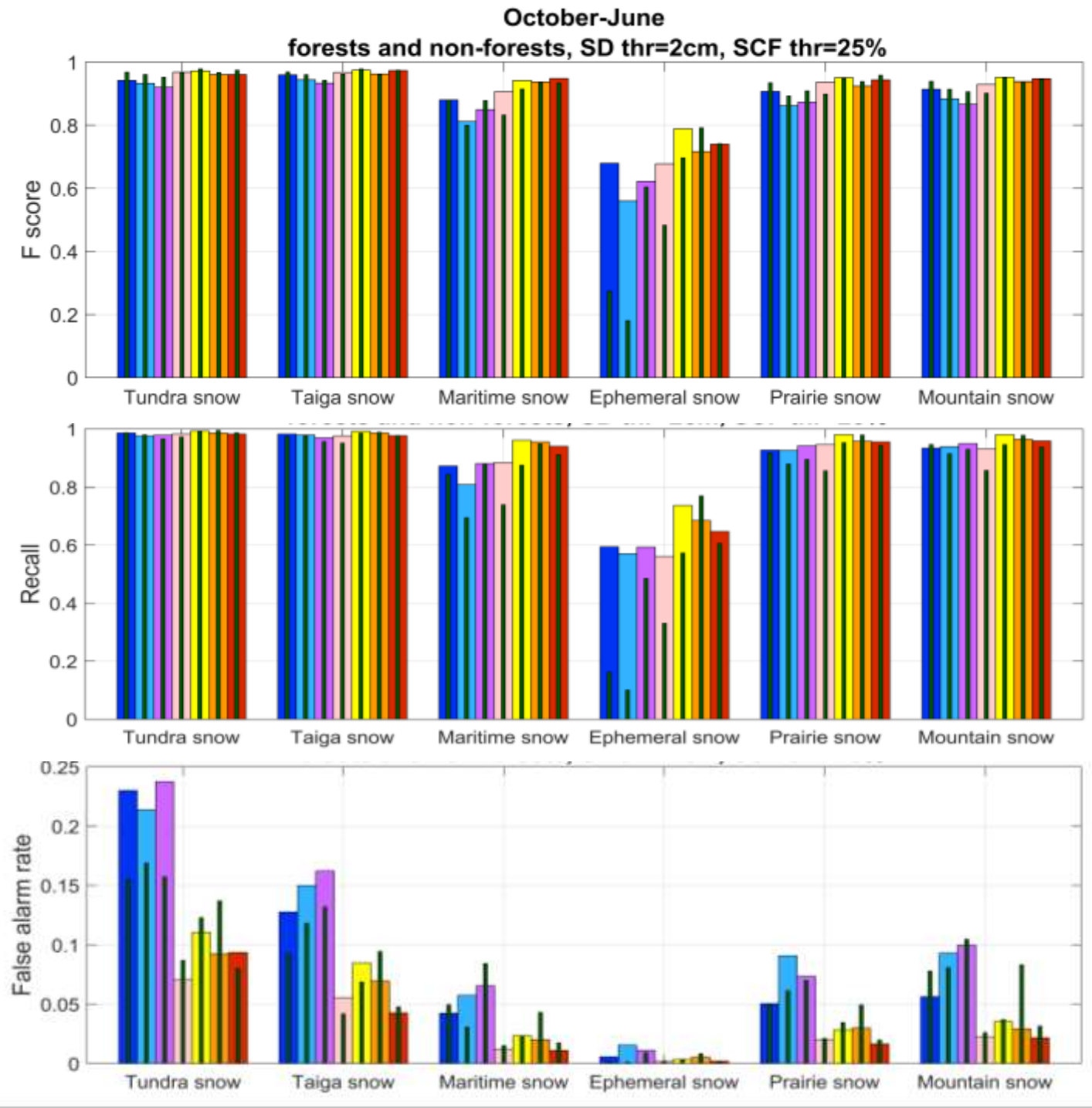
- Several Snow extent products were compared against in-situ data in different snow zones
- **Tundra**, **Taiga**, **Mountain snow** are handled very well by all the products (results are missing for Prairie snow due to too low number of cases). SCAG and SCAGCA have the largest difficulties in Taiga forests.
- **Maritime snow and Ephemeral snow** are most challenging (this applies in all the analyses also outside this presentation) .
  - For Maritime snow, JXM10, M10C05 and GLSSE do the best job.
  - For Ephemeral snow, GLSSE provides the best performance in forests (not enough data for non-forests). Considering both forets and non-forests, M10C05 provided the best performance, accompanied by JXM10, SCAG and SCAGCA, the latter three showing clear difficulties in forests though.

Validation oh NH products, now excluding products based on MODSCAG, as well as European-scale product.

- Period 2000-2001, 2003-2004, 2005-2006 and 2007-2008 analyzed. 2011-2011 excluded as Nothern America dataset is limited in that period.







- Coloured bar: non-forest
- Black bar: forest

ASNOW
CRCLIM
IMS04
JXAM5
JXM10
GLSSE
M10C05

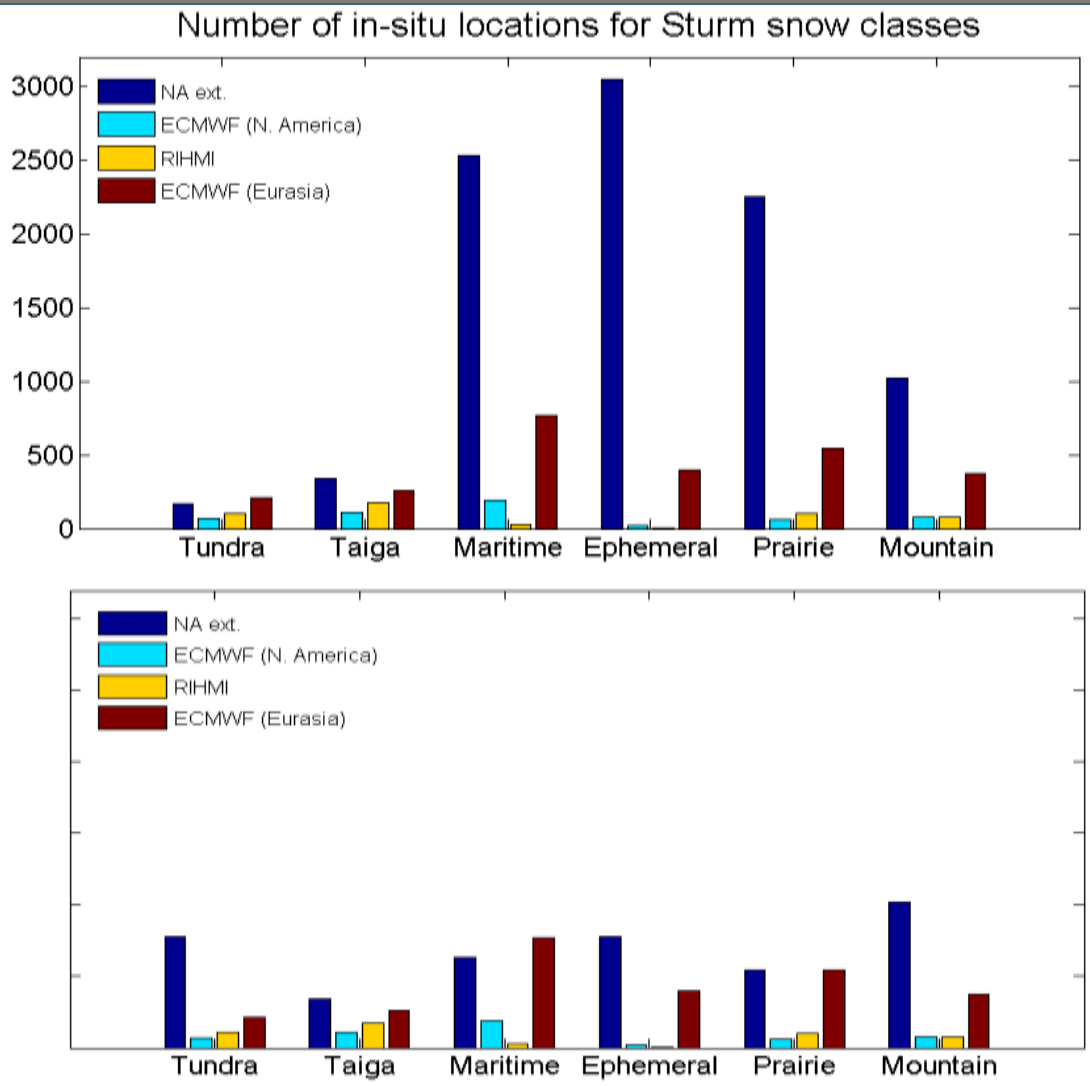
Summary and Conclusions for NH

- Several Snow extent products were compared against in-situ data in different snow zones
- Tundra, Taiga, Prairie and Mountain snow are handled very well by all the products. Distinctively high False snow alarms are associeted t to ASNOW, CRCLIM and IMS04. These high FAR's are gained in accumulation and melting periods; mid-winter FAR's are as low as with the other products.
- Maritime snow and Ephemeral snow are most challenging areas: these are handled clearly best by JXM10, M10C05 and GLSSE. In forest, GLSSE provides the best performance.
- The results presented here are in-line with the comparisons vs. Landsat-based high-resolution snow maps also carried out within SnowPEx

Thank you for your attention!



Analyses on in-situ data; balancing of the dataset: Northern America (NA) data dominates the dataset



Original number of sites

Number of sites after reduction of Northern America dataset

Effect on SCF-threshold in SCF-> binary conversion for snow products

- Threshold of SCF=25% seems to be most feasible

