ALPHA v2.0 Status: Application of Machine Learning Technology

Allyson Rugg and Julie Haggerty

ALPHA 3-Input v1.0: Deployed in Field Campaigns

Satellite	Model	3D Radar Mosaic
Find highest, coldest, thickest clouds from Total Water Path, Cloud Top Height and Cloud Top Temperature – *2D field*	Find deep cloud layer, heavy precipitation, high condensate, updrafts, temperature below -15°C – *3D field*	Find active updrafts, high reflectivity in column along with heights of 10 and 30 dBz echo tops – *2D field*
Total Satellite Interest	**Total Model Interest**	**Total Radar Interest**

Calculate Total HIWC Interest

If Total Satellite Interest is > 0

Model 3D Temperature Interest * [45% Total Satellite Interest + 10% Total Model Interest + 45% Total Radar Interest]

= **Total HIWC Interest**

Objective Re-Design of the ALPHA Fuzzy Logic Algorithms Using Field Campaign Data

- Fuzzy logic methodology allows for adjustment of multiple parameters in the algorithms including:
  - Input variables used
  - Shape of membership function for each variable
  - Weight given to each variable in the blending process

- Optimization of parameter set
  - Need a performance metric that defines “optimal”
  - Apply machine learning tool to our data set
  - Many iterations later, we have a new algorithm

Input Variables Considered for Use in ALPHA

Satellite	NWP Model	Groundbased Radar
Effective Cloud Top Temperature	Temperature	Maximum Reflectivity in Column
Effective Cloud Top Height	Surface Precipitation	Maximum Height of 30 dBz Reflectivity
Total Water Path	Total Condensate	Maximum Height of 10 dBz Reflectivity
Optical Depth	Total Water Path	Vertically Integrated Liquid
Brightness Temperature Difference (6.7 – 10.8 um)	Vertical Velocity	Volume Averaged Height Integrated Reflectivity
Brightness Temperature Difference (10.8 -12 um)	Tropopause Height	Precipitation Ice Mass
	Convective Available Potential Energy, Convective Inhibition	
	Divergence/Convergence	
	Vorticity	

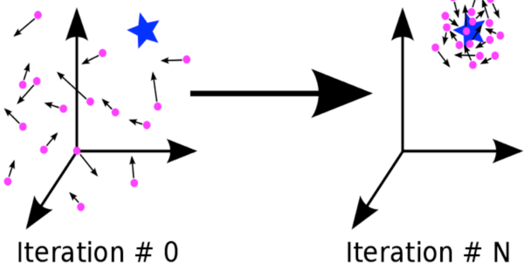
Methods/PSO Summary

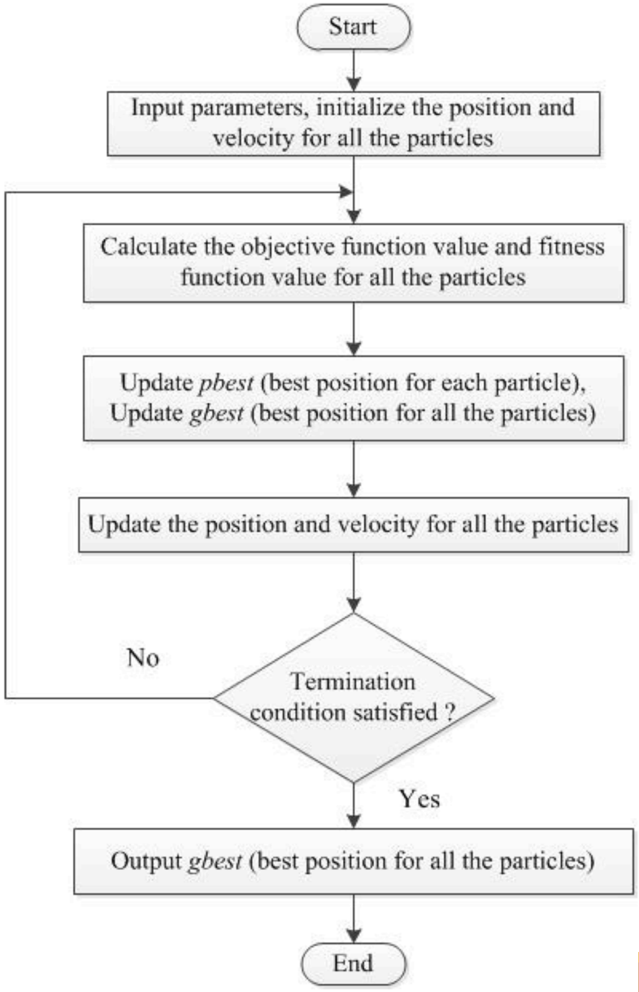
Used a “Particle Swarm Optimization” to optimize membership functions and weights

- Stochastic, supervised machine learning algorithm
- Tests sets of parameters for performance and automatically adjusts them to get progressively better performance
- Objective way to choose ALPHA membership functions and weights
- Works in any number of dimensions

Particle Swarm Optimization (PSO)

- Supervised machine learning algorithm to find optima
- Easily to implement in any number of dimensions
- Each “particle” in the swarm “remembers” where it and its neighbors have been
  - The location of a particle represents a set of parameters by which we could compute HIWC interest
  - Particles move stochastically towards better solutions while searching the space around them
  - Eventually the particles converge on a local optimum
  - Running several times increases likelihood of finding the global optimum





PSO Implementation

3 degrees of freedom per input field

- Location of first inflection point
- Distance between inflection points (enforced to be positive)
- Weight (enforced to be positive and sum to 1)

Membership functions are piecewise linear with exactly 2 inflection points

- Sign of each membership function also determined by user

PSO can remove inputs by assigning zero weight, but cannot add new inputs

Radar, satellite, model, and temperature interests optimized separately

- Blended using weights from another PSO

Optimization Metrics Tested

The PSO algorithm needs to know how to score each location so that it can determine how the particles should move in each iteration. We tried several metrics:

1) **Correlation** between IWC and interest
2) **Histogram Error** where we bin interest values, look at sum of squared difference from an "ideal" histogram where interest = fraction of MOG IWC measurements
3) **Fraction of Correct** assignments using a 0.5 IWC and 0.5 interest thresholds
4) **Sum of PoDno and PoDyes** using 0.5 IWC and 0.5 interest thresholds
5) **Products** developed to try to balance two or more of the other metrics at once
    - (1 – correlation)*(histogram error)*(fraction of wrong assignments)
    - (1 – correlation)*(histogram error)*(2 – PoDyes - PoDno)

Choice of Metric

The **correlation** between IWC and interest was the preferred metric

- Utilized the most information
  - No need to bin or categorize the IWC or the interest
  - All of the other metrics relied on defining bins or thresholds for IWC and/or interest, so some information is lost in the process of doing this
  - Many other metrics put a lot of weight on just a handful of inputs
- Can still look at histogram shape and PoD statistics after running PSO for further verification

ALPHA 2.0 Summary

All membership functions and weights adjusted based on “training” with IWC data from IKP2

Satellite Interest

- Incorporate two new brightness temperature difference fields
- Replace total water path with optical depth

Radar Interest

- Remove 30 dbz height
- Add new VAHIRR field (volume averaged height integrated radar reflectivity)

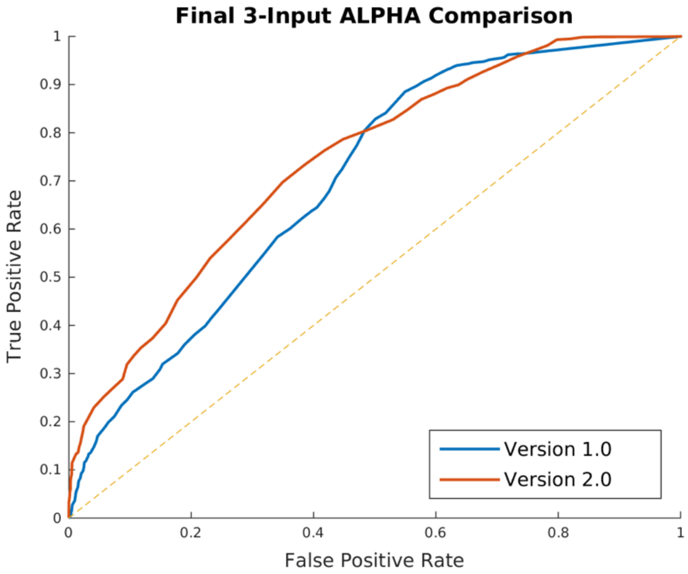
Model Interest

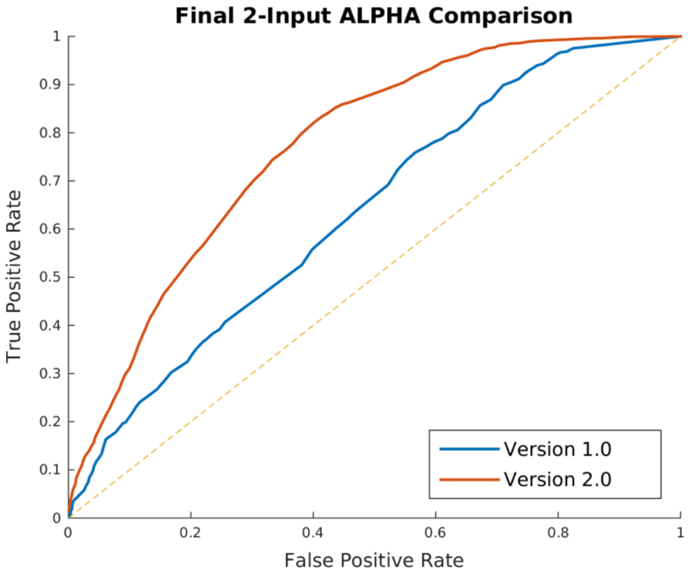
- Remove total water path and precipitation
- Add surface wind curl and divergence
- Only permitted to increase final interest

Temperature Interest

- Include warmer temperatures

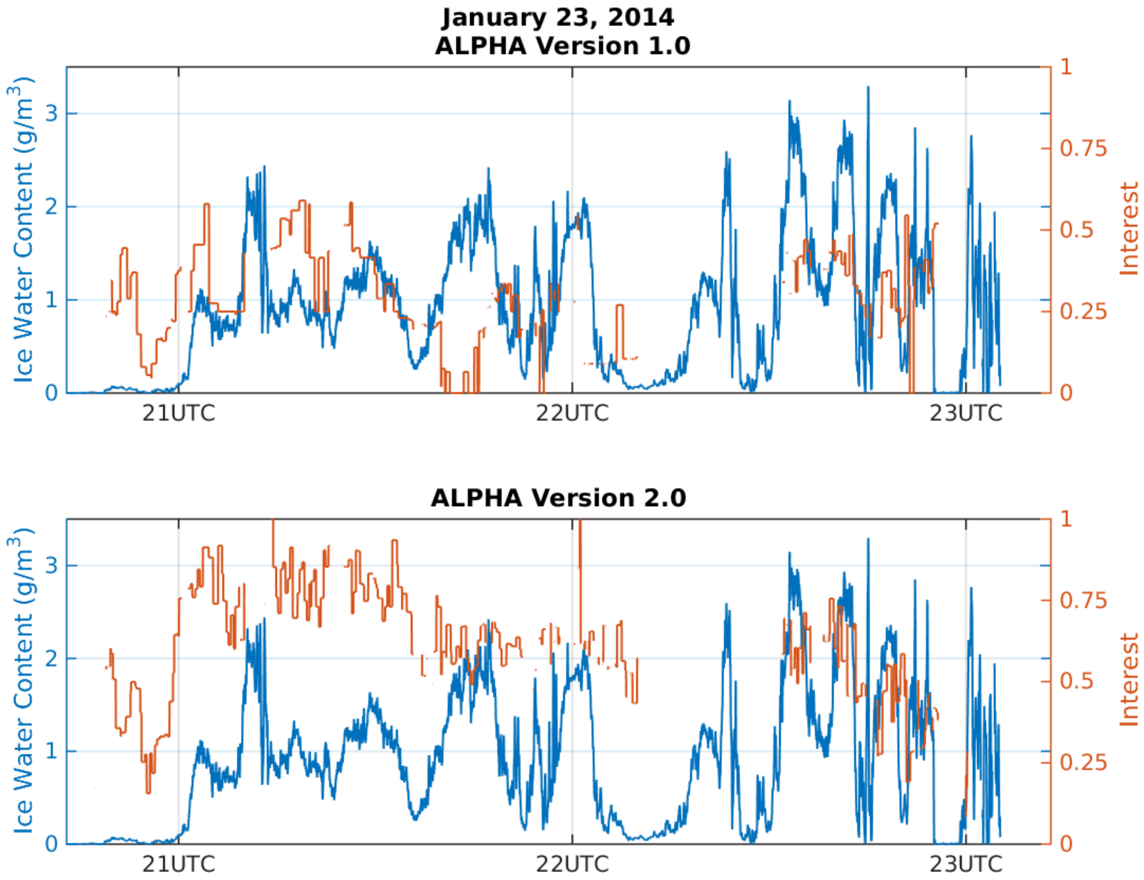
Performance Comparison





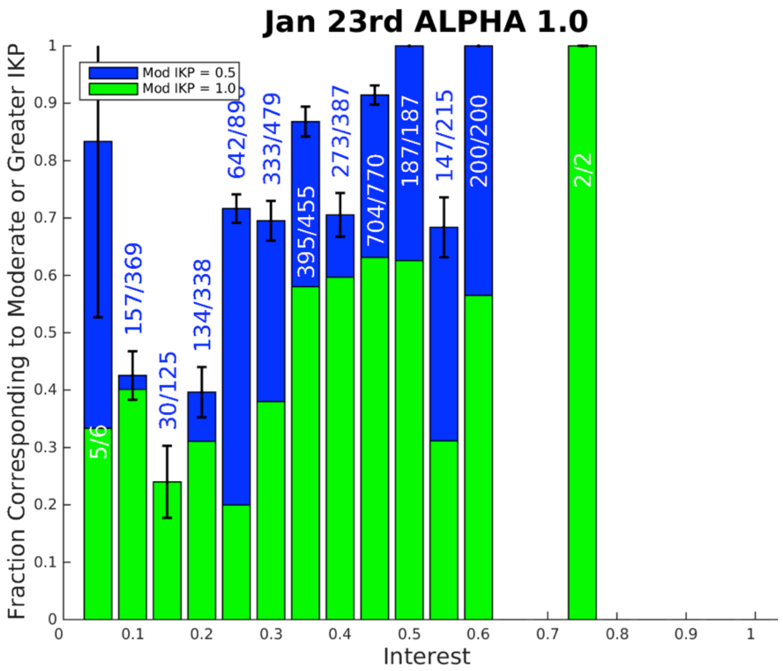
These ROC curves are created by setting a constant HIWC threshold of 0.5 g/m^3 and letting the HIWC interest threshold vary between 0 and 1

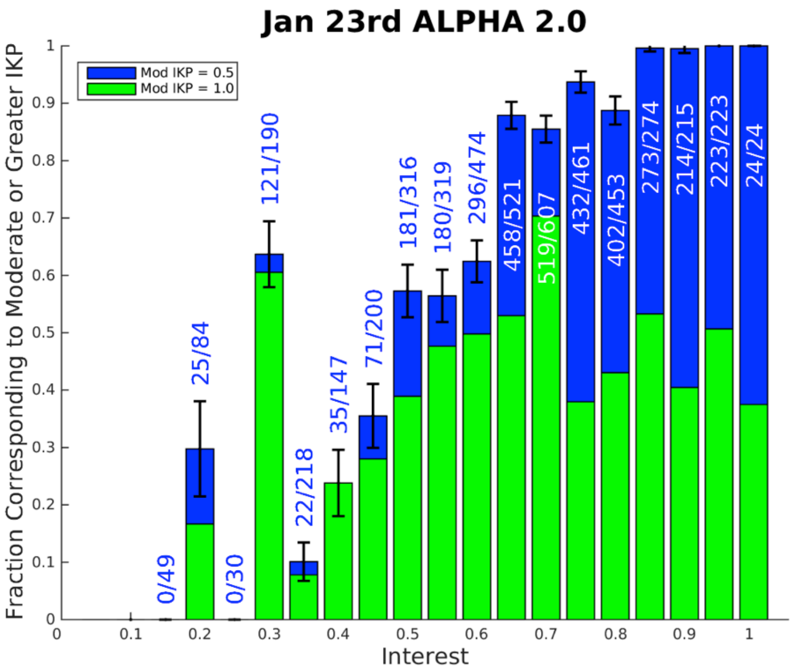
Note: The 3-Input interest has a much smaller sample size than the 2-Input. If we only consider point where both interests are available, the 3-Input performs better than the 2-Input interest.



January 23, 2014
ALPHA Version 1.0
Ice Water Content (g/m3)
Interest
0
1
2
3
0
0.25
0.5
0.75
1
21UTC
22UTC
23UTC
ALPHA Version 2.0
Ice Water Content (g/m3)
Interest
0
1
2
3
0
0.25
0.5
0.75
1
21UTC
22UTC
23UTC

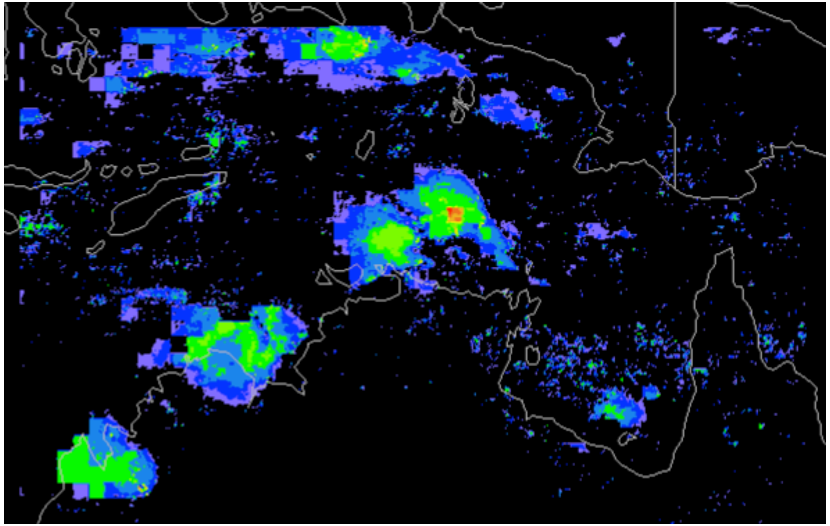
Case Study: Jan 23rd Darwin Flight



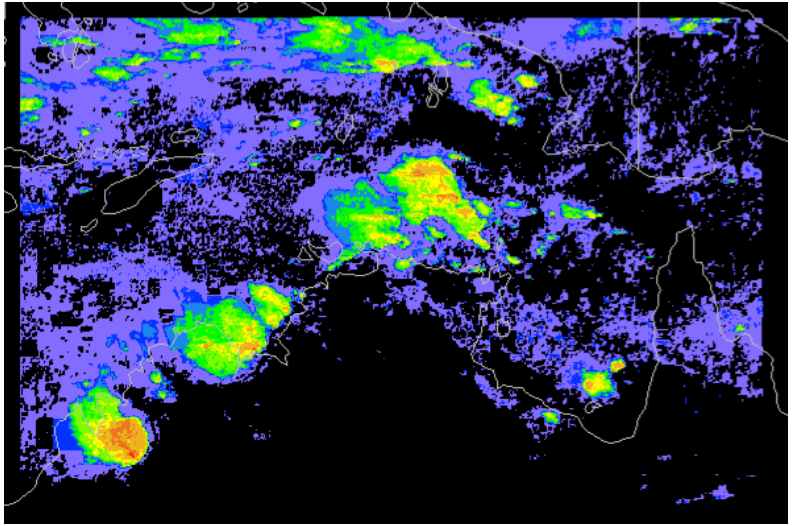


Case Study: Jan 23rd 2015, 22:45 UTC

ALPHA 1.0



ALPHA 2.0



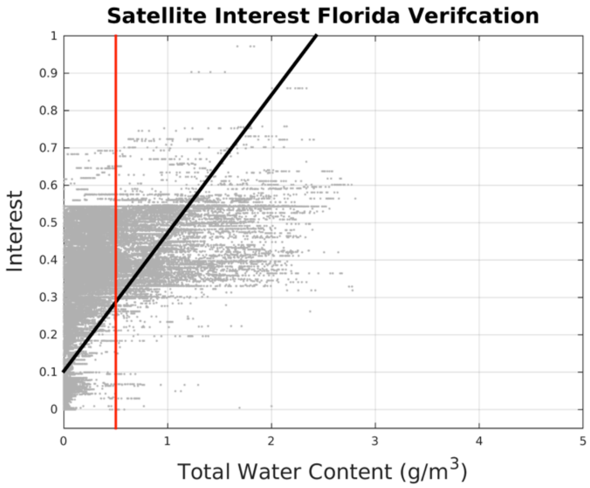
Florida Verification: Satellite

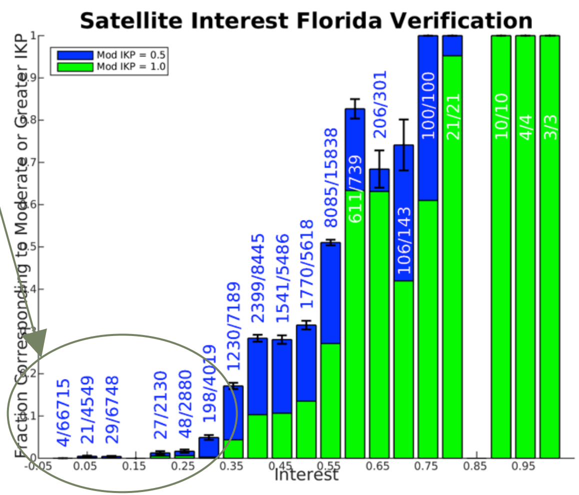
Stronger correlation between interest and IWC than training set (Darwin and Cayenne)

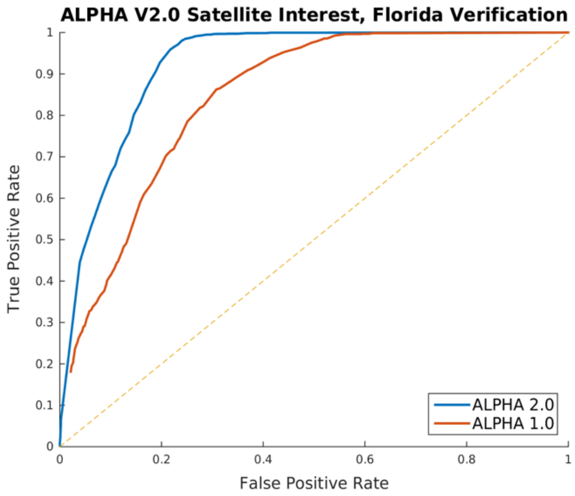
- Florida satellite correlation: 0.6671
- Training set satellite correlation: 0.4394

Few interest values above ~0.55

Very few false negatives







Next Steps with ALPHA v2.0

- Continue with comparison of ALPHA product with IKP2 IWC measurements from HIWC-Florida experiment
  - Independent assessment
- Update ALPHA-CONUS real-time product with ALPHA v2.0; implement a version in Australia
- Use ALPHA v2.0 to characterize horizontal variation and time duration of HIWC features in ALPHA products
- Airborne cloud radar (RASTA) IWC retrievals for comparison with ALPHA vertical variation
- Advection of HIWC features using TITAN (Thunderstorm Identification Tracking and Nowcasting)

Planned presentations and publications

AMS ARAM Conference - Jan 2017

1. Haggerty, Rugg, McCabe, Kessinger, Strapp, Potts, Palikonda: Detection of High Ice Water Content (HIWC) conditions: Status of nowcasting tool development for avoidance of ice crystal icing events, submitted.
2. Rugg, Haggerty, McCabe, Kessinger, Strapp, Delanoe: Evaluation of the Algorithm for Prediction of High Ice Water Content Areas (ALPHA): Methods and Results, submitted

AIAA Atmosphere and Space Environment Conference - June 2017

1. Rugg, Haggerty, Palikonda, Potts: High Ice Water Content Conditions around Darwin: Frequency of Occurrence and Duration as Estimated by a Nowcasting Model, submitted.

Journal Articles in Preparation

1. Haggerty and HIWC co-authors: Development and Verification of a Detection Method for High Ice Water Content Regions, planned submission to an AMS journal, early 2017
2. Haggerty, Jensen, and Yost: High Ice Water Content and Airborne Temperature Measurement Anomalies near Tropical Convection, planned submission to an AMS journal, early 2017

Additional Slides

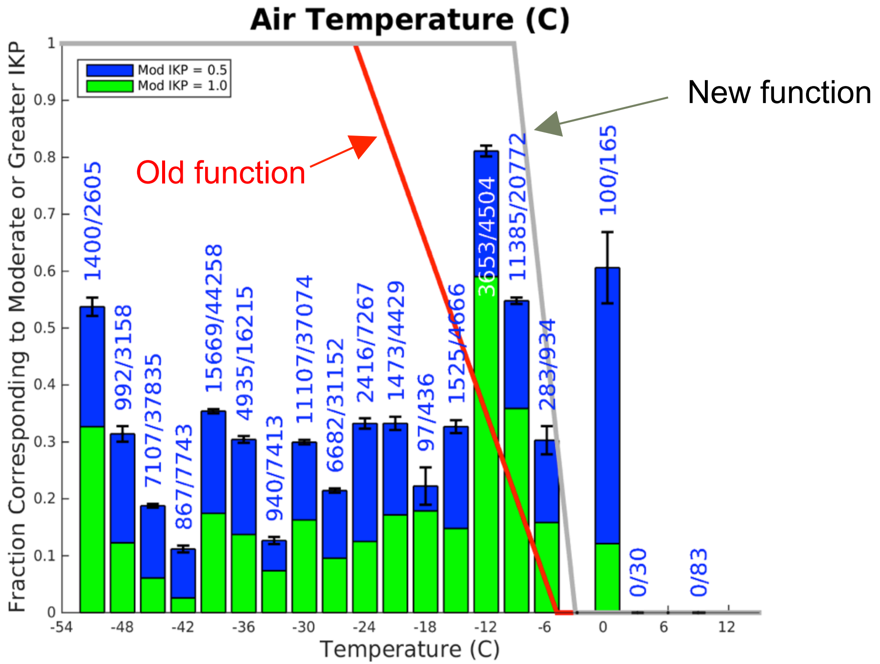
ALPHA 2.0: Temperature Interest

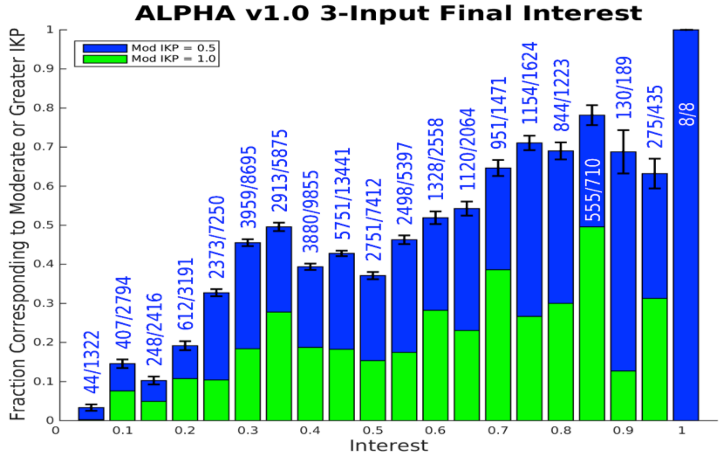
The blended interest is multiplied by the temperature interest to mask out areas too warm for HIWC
PSO runs resulted in a membership function that includes warmer temperatures

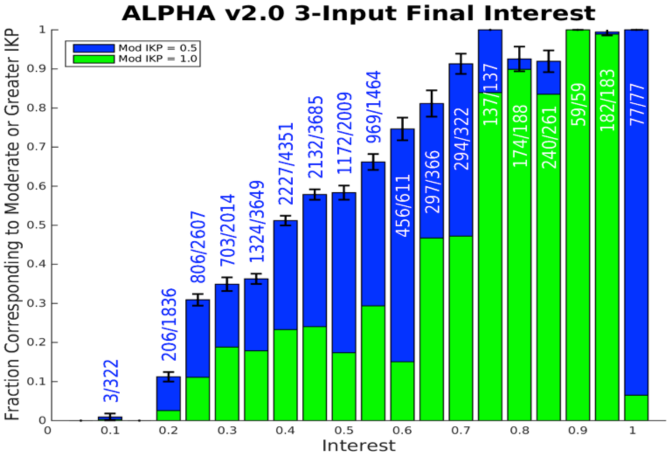
Big improvement in final ALPHA performance between old and new membership functions

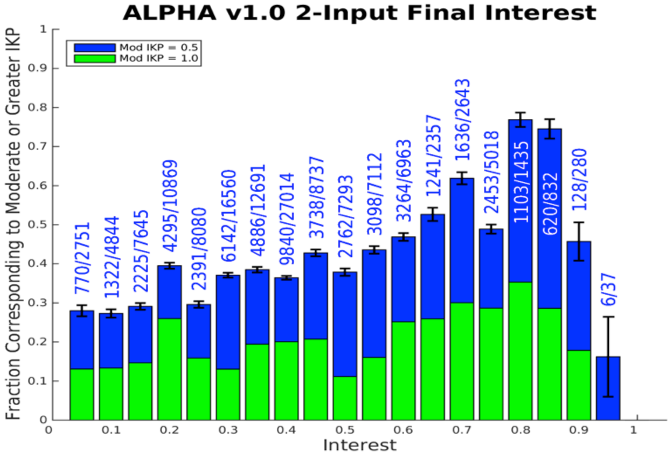
Little difference in performance between new algorithm and using no temperature masking

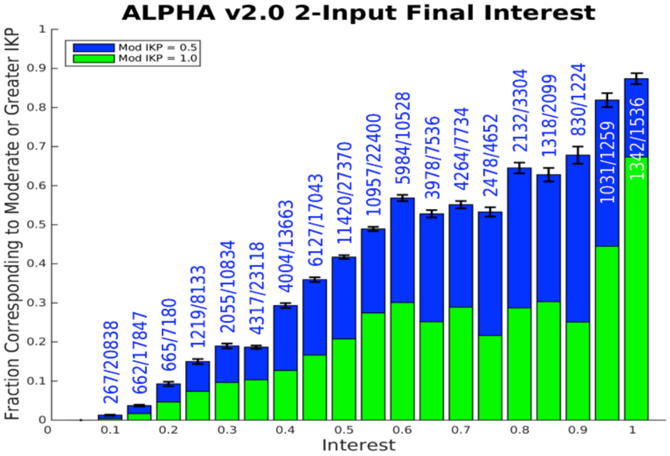
- Sampling bias
- Likely prevents over-forecasting in above freezing temperatures

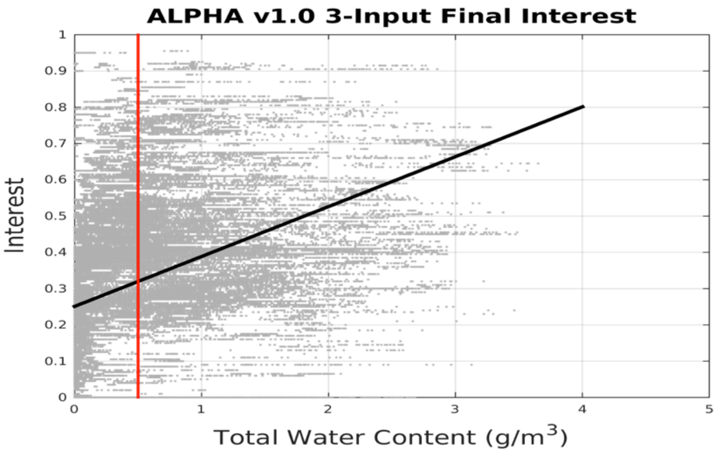


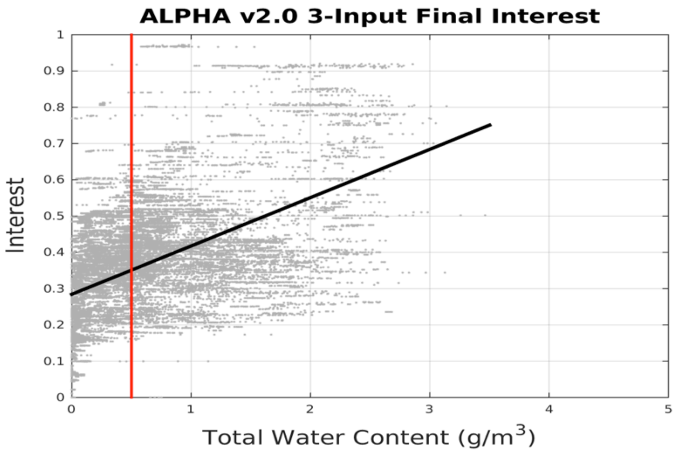

ALPHA v1.0 3-Input Final Interest
Mod IKP = 0.5
Mod IKP = 1.0
Fraction Corresponding to Moderate or Greater IKP
Interest
44/1322
407/2794
248/2416
612/3191
2373/7250
3959/8695
2913/5875
3880/9855
5751/13441
2751/7412
2498/5397
1328/2558
1120/2064
951/1471
1154/1624
844/1223
555/710
130/189
275/435
8/8

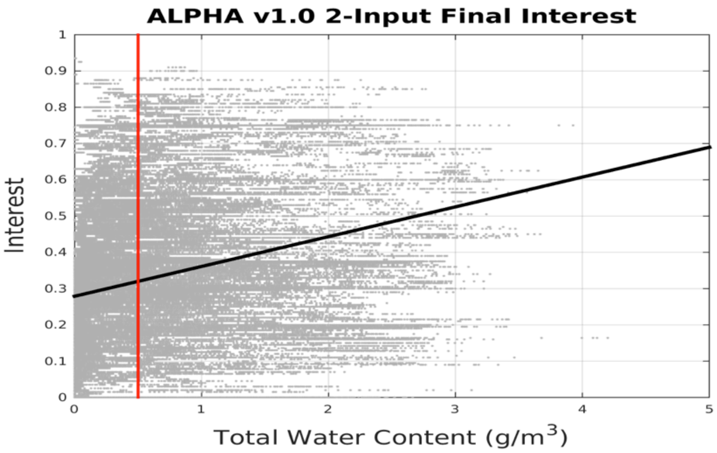

ALPHA v2.0 3-Input Final Interest
Mod IKP = 0.5
Mod IKP = 1.0
Fraction Corresponding to Moderate or Greater IKP
Interest
3/322
206/1836
806/2607
703/2014
1324/3649
2227/4351
2132/3685
1172/2009
969/1464
456/611
297/366
294/322
137/137
174/188
240/261
59/59
182/183
77/77

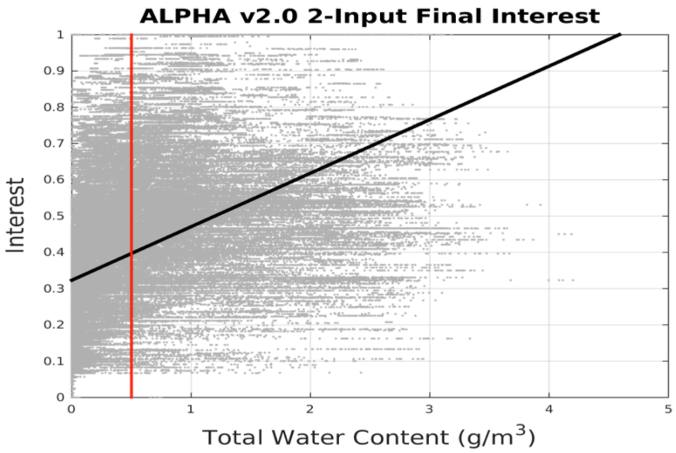

ALPHA v1.0 2-Input Final Interest
Mod IKP = 0.5
Mod IKP = 1.0
Fraction Corresponding to Moderate or Greater IKP
Interest
770/2751
1322/4844
2225/7645
4295/10869
2391/8080
6142/16560
4886/12691
9840/27014
3738/8737
2762/7293
3098/7112
3264/6963
1241/2357
1636/2643
2453/5018
1103/1435
620/832
128/280
6/37


ALPHA v2.0 2-Input Final Interest
Mod IKP = 0.5
Mod IKP = 1.0
Fraction Corresponding to Moderate or Greater IKP
Interest
267/20838
662/17847
665/7180
1219/8133
2055/10834
4317/23118
4004/13663
6127/17043
11420/27370
10957/22400
5984/10528
3978/7536
4264/7734
2478/4652
2132/3304
1318/2099
830/1224
1031/1259
1342/1536

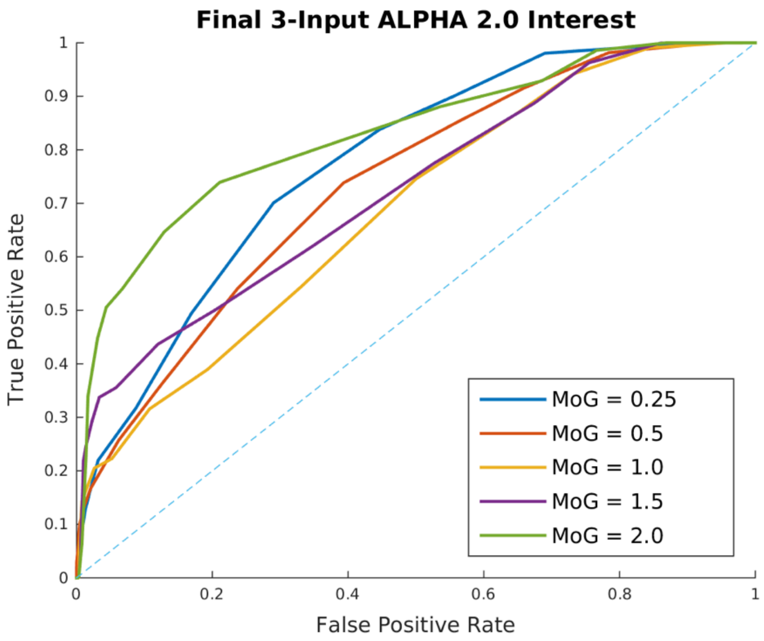

ALPHA v1.0 3-Input Final Interest
Interest
Total Water Content (g/m³)

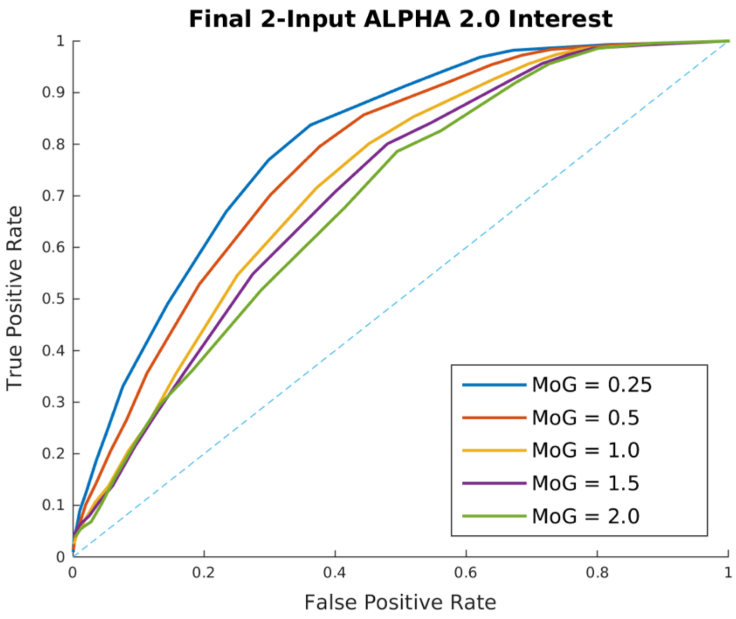

ALPHA v2.0 3-Input Final Interest
Interest
Total Water Content (g/m³)


ALPHA v1.0 2-Input Final Interest
Interest
Total Water Content (g/m³)


ALPHA v2.0 2-Input Final Interest
Interest
Total Water Content (g/m³)

V2.0 Performance by IWC Threshold





Remember sample size for 3-Input Interest is much smaller than for the 2-Input

Brightness Temperature Differences

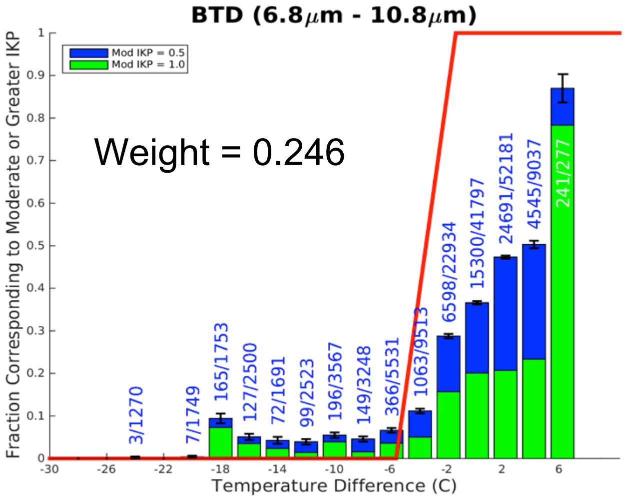
Assigned high weight in ALPHA 2.0 satellite interests

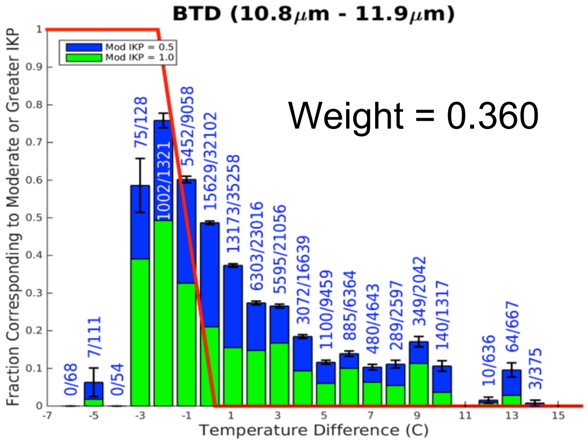
- Over 60% combined

Water vapor minus infrared (right)

- Indicates moist stratosphere
- Associated with overshooting tops

Two different infrared channels (below)





Schmetz, J., S. A. Tjemkes, M. Gube, and L. Van De Berg. "Monitoring Deep Convection and Convective Overshooting with Meteosat." *Adv. Space Res.* 19.3 (1997): 433-441.