

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

Evan Manning and George Aumann

18 October 2008



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

The AIRS level 2 retrieval accuracy is specified as
1 K in 1 km thick layers for global statistics.

This may have been a good requirement 10 years ago, but in
the current environment we should rethink this specification.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

The AIRS level 2 retrieval accuracy is specified as
1 K in 1 km thick layers for global statistics.

This may have been a good requirement 10 years ago, but in
the current environment we should rethink this specification.

A product which is more accurate than the best
competing product is more important



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

The AIRS level 2 retrieval accuracy is specified as 1 K in 1 km thick layers for global statistics.

This may have been a good requirement 10 years ago, but in the current environment we should rethink this specification.

A product which is more accurate than the best competing product is more important

Retrieval Skill quantifies the ability of one retrieval to be more accurate than the best forecast relative to another retrieval with the same or another sounder.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Forecasting

The skill for weather forecasting compares the accuracy of the forecast for day 1, 2, 3, ... to the difference between the actual conditions and the conditions expected from climatology.

The skill is measured using the anomaly correlation

$$AC(t) = \text{Cor}(\text{forecast}(t)\text{-climatology, analysis-climatology})$$

Where $t = 1, 2, 3, \dots$ are days for the forecast.

For day=0 $AC(0)=1$.

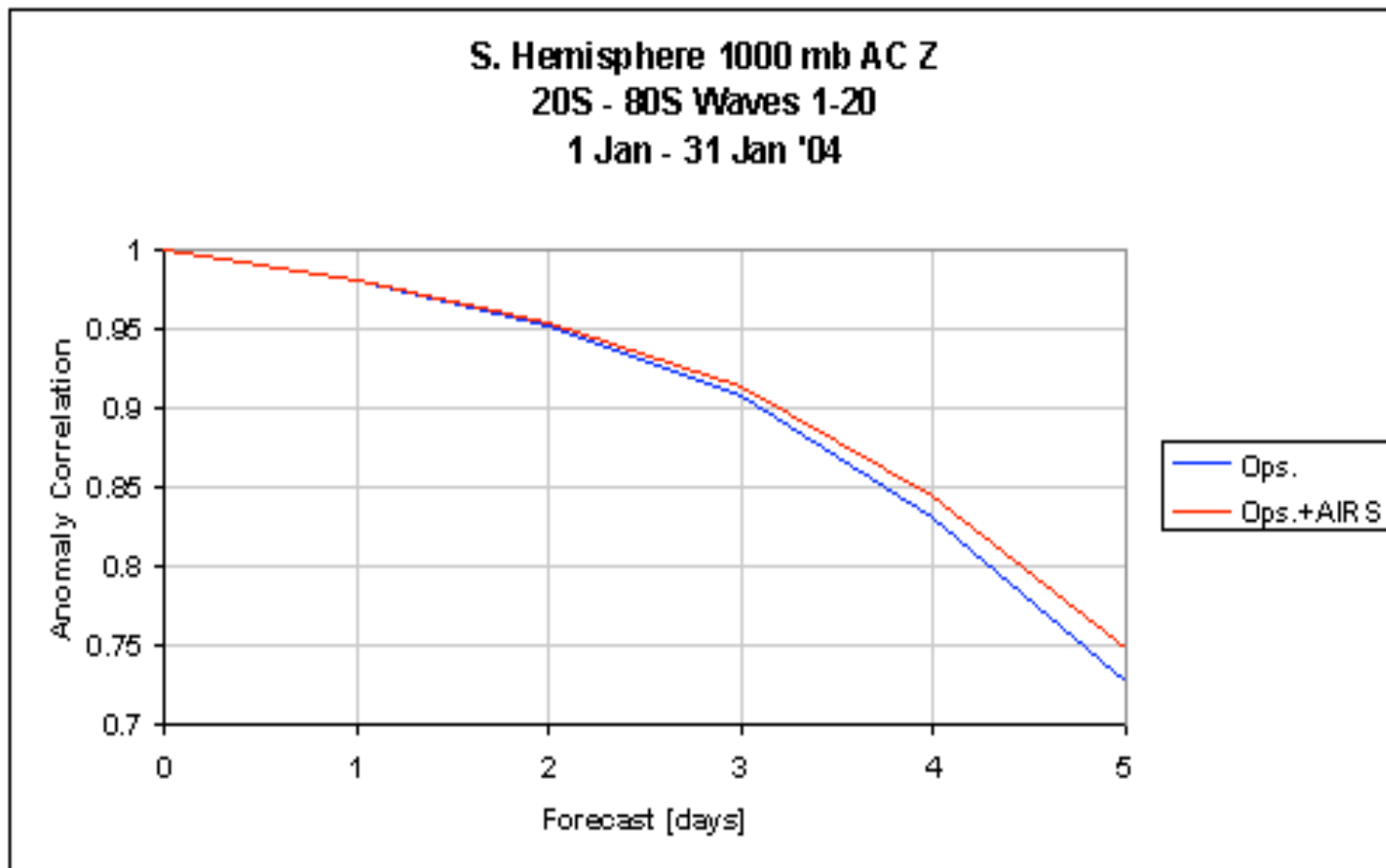
The length of the useful forecast is the number of days with $AC > 0.6$



National Aeronautics and
Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Forecasting





National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

We want to use a similar approach for the evaluation of the skill of retrievals of $T(p)$, $q(p)$, ozone, etc.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

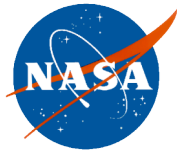
The skill score has a range from zero to 1

Skill is zero when the product

- matches the background
- gives no answer
- has zero correlation with the truth

Background = the best solution obtainable in real time

For $T(p)$ and $q(p)$ this is the NCEP or ECMWF forecast



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

Retrieval Anomaly Skill Score

$$\text{RASS} = \text{cor}(\text{retrieved-background}, \text{truth-background}) * \text{sqrt}(f)$$

Where f = the ratio of accepted to the possible retrievals
i.e. the fractional yield.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

Retrieval Anomaly Skill Score

$$\text{RASS} = \text{cor}(\text{retrieved-background}, \text{truth-background}) * \text{sqrt}(f)$$

Where f = the ratio of accepted to the possible retrievals
i.e. the fractional yield.

An optimum likelihood retrieval returns a solution at all times.
This is automatically taken into account in RASS.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

Retrieval Anomaly Skill Score

$$\text{RASS} = \text{cor}(\text{retrieved-background}, \text{truth-background}) * \text{sqrt}(f)$$

Where f = the ratio of accepted to the possible retrievals
i.e. the fractional yield.

An optimum likelihood retrieval returns a solution at all times.
This is automatically taken into account in RASS.

For retrieval evaluation truth=RAOB,
ground truth or NCEP analysis
depending on the quantity



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

We have tested this scheme using the V4, V5, and SCNN for
2002-09-06

$RASS = \text{cor}(\text{retrieved-background}, \text{truth-background}) * \text{sqrt}(f)$



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

We have tested this scheme using the V4, V5, and SCNN for
2002-09-06

$RASS = \text{cor}(\text{retrieved-background}, \text{truth-background}) * \text{sqrt}(f)$

For the truth we use the ECMWF T(p)

The background the NCEP 15 year reanalysis (1988-2002)



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Skill in Retrievals

We have tested this scheme using the V4, V5, and SCNN for
2002-09-06

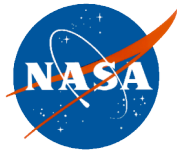
$RASS = \text{cor}(\text{retrieved-background}, \text{truth-background}) * \text{sqrt}(f)$

For the truth we use the ECMWF T(p)

The background the NCEP 15 year reanalysis (1988-2002)

The NCEP 15 year reanalysis was selected because it is readily available for quick results and should be reasonable good for the tropical ocean.

It is adequate for relative skill comparisons, but not for absolute skill.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Features of RASS

- We've been struggling to balance yield vs. error levels in comparing versions. RASS automatically includes yield.
- But RASS gives zero weight to bias and scaling. It does not replace scoring retrievals for accuracy and bias.



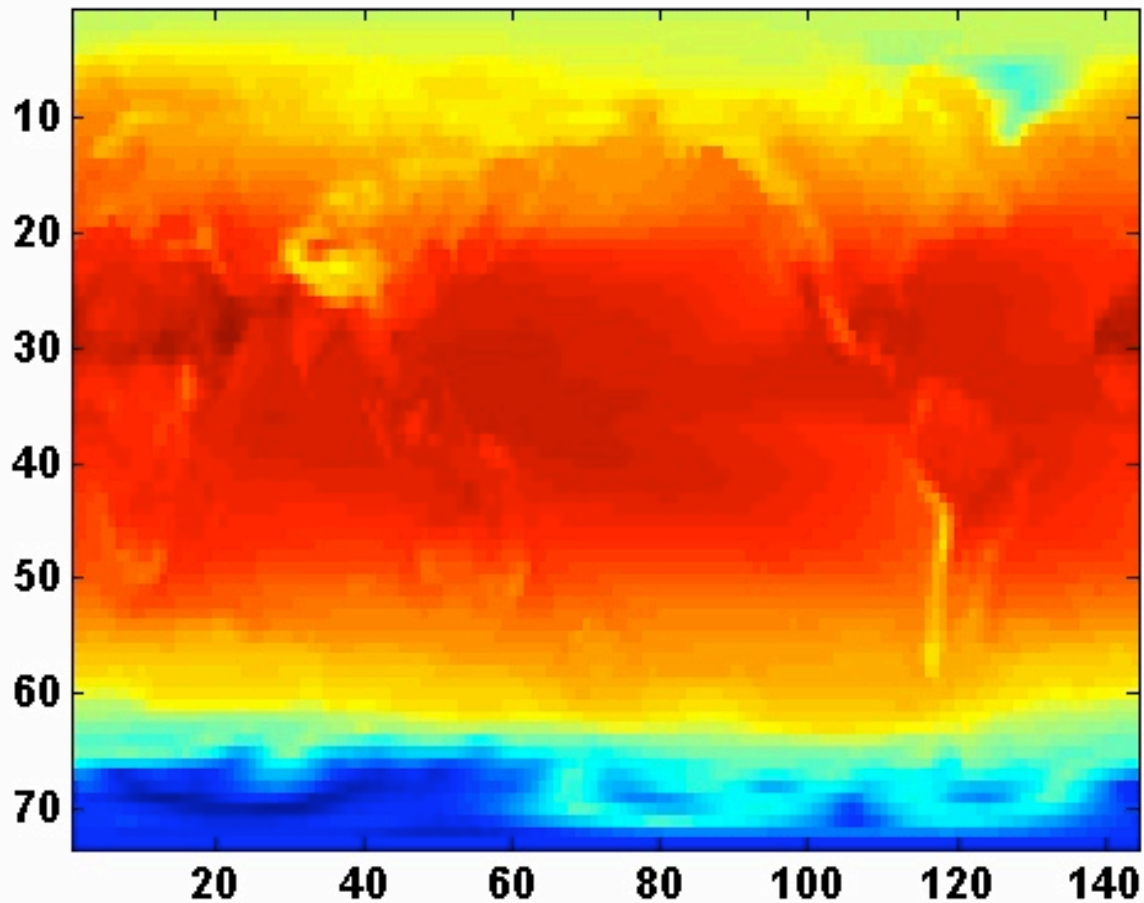
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The September mean T_{surf} from the 1978-1998 NCEP Reanalysis

Atmospheric Infrared Sounder

NCEP 20 year reanalysis september mean T_{surf}



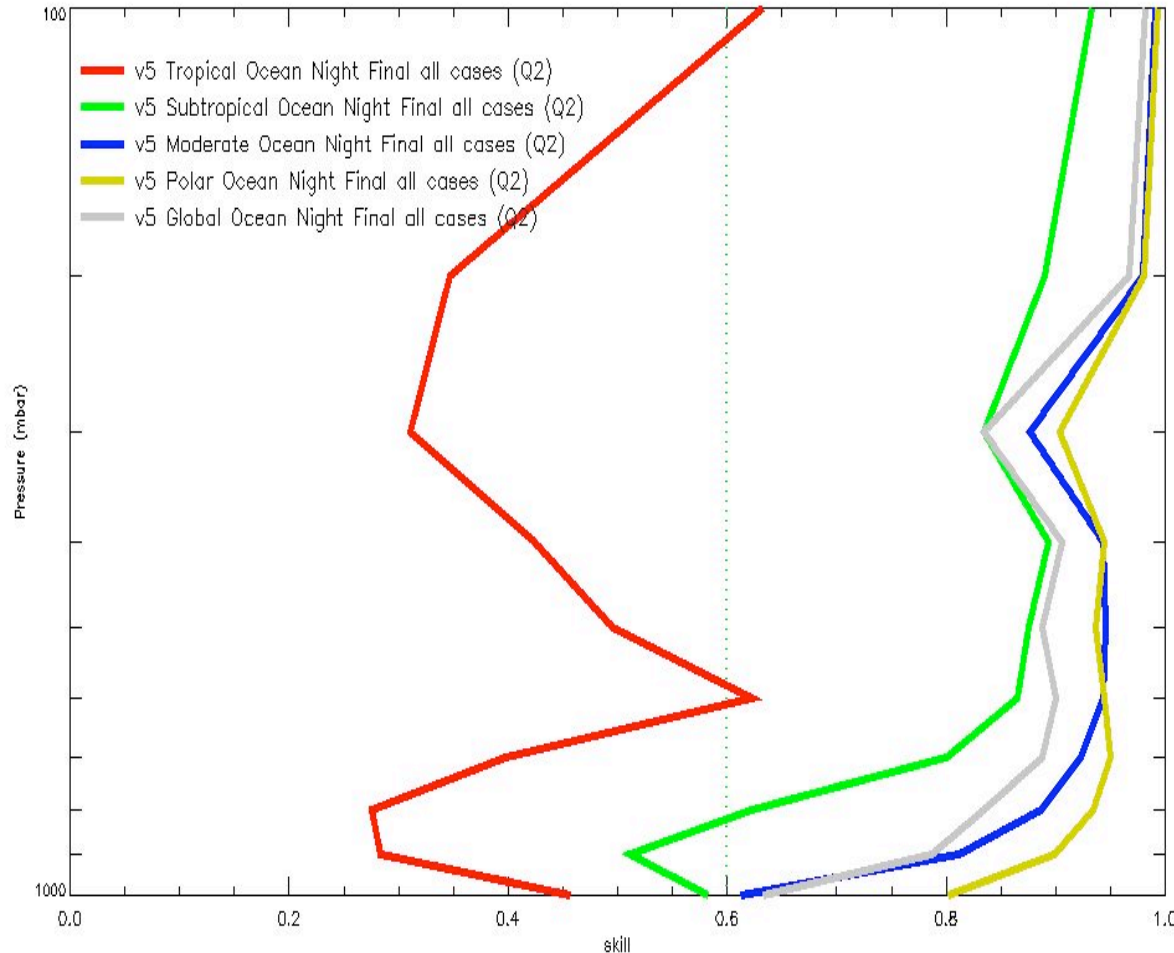
RTGSST.sept2002 -Tsurf_1995 bias = 0.00 stdev=1.1 K for all non-frozen oceans.



National Aeronautics and
Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

RASS -- Regional Variation



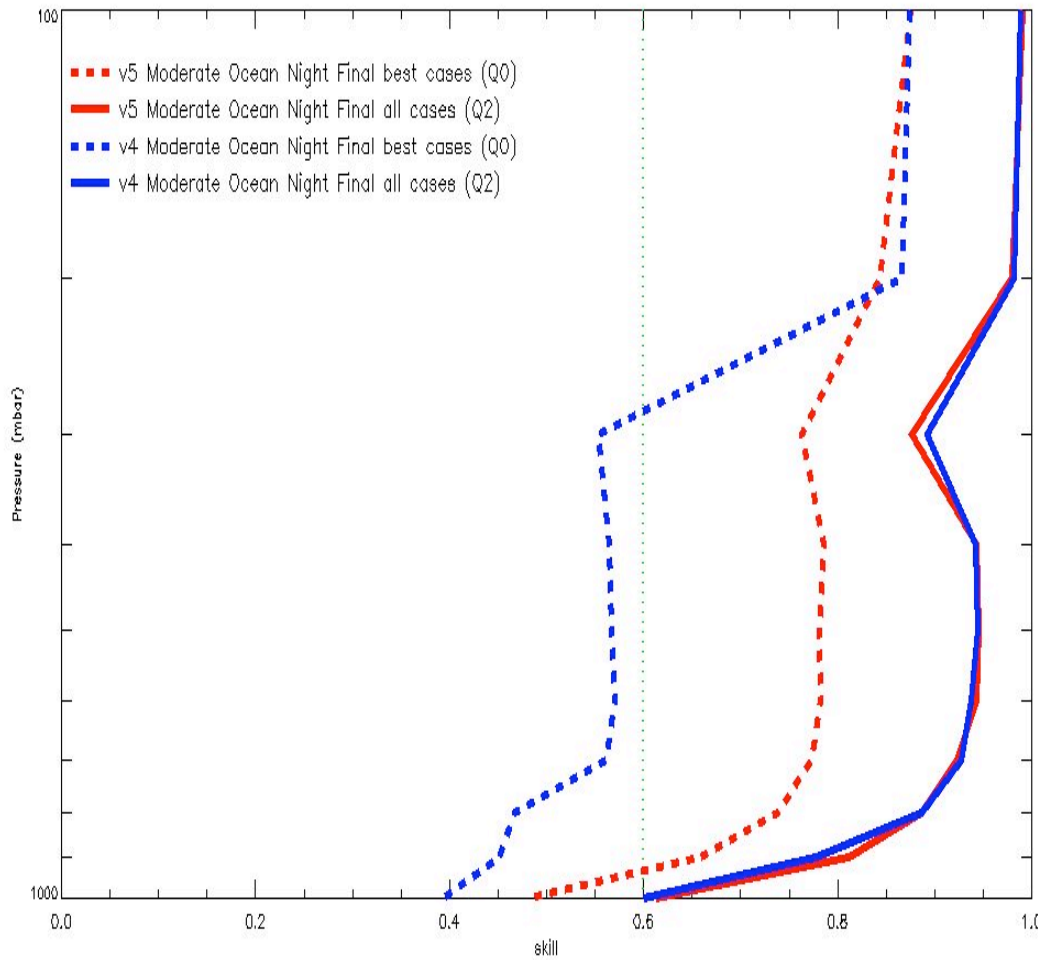
Retrieval
Anomaly Skill
Score is
generally
lowest in the
tropics
because
climatology is
a good first
guess there.



National Aeronautics and
Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

RASS -- Skill vs. Yield



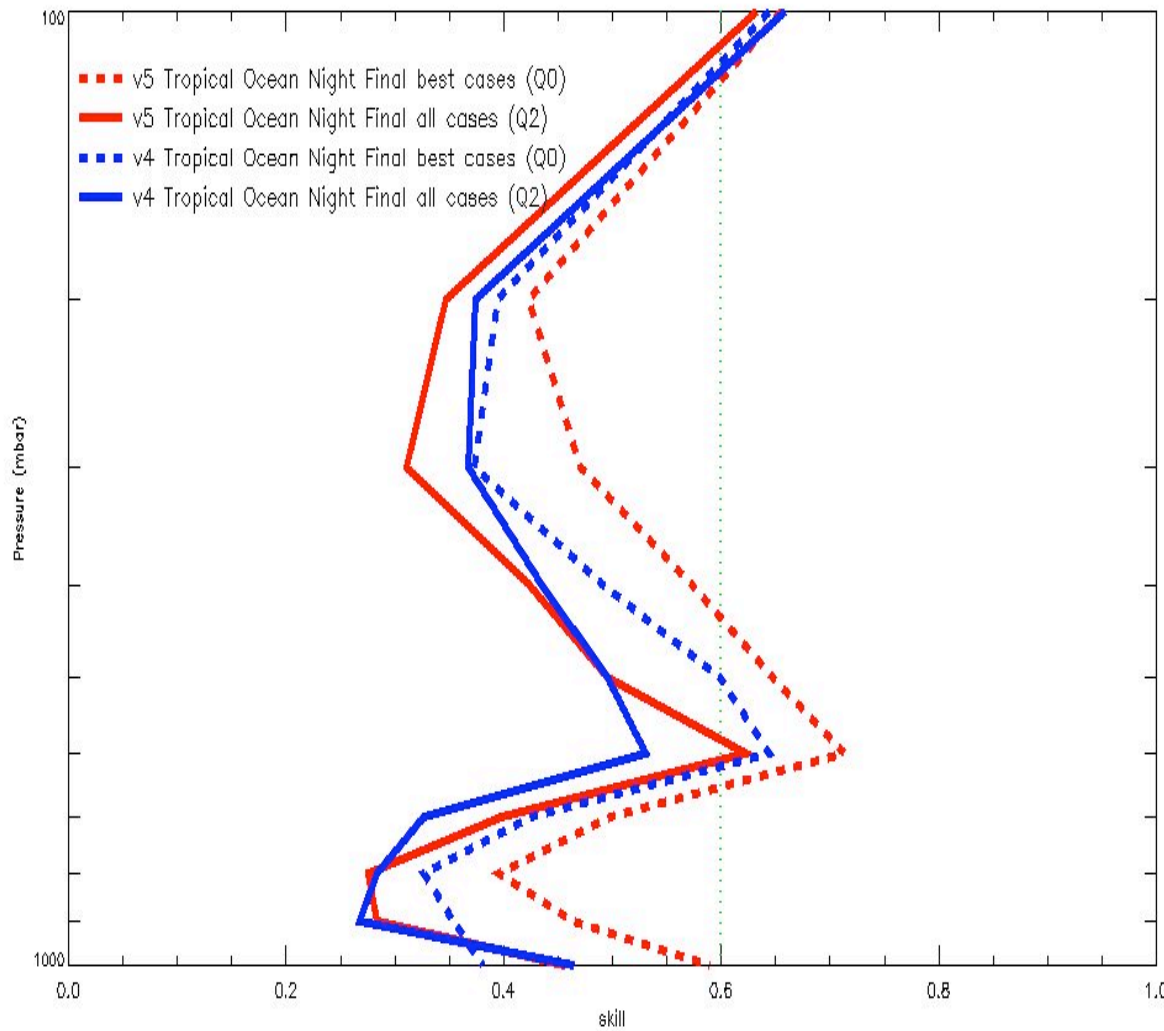
The lower yield of Q0 (best data only) gives Q2 (all data) generally a higher Retrieval Anomaly Skill Score



National Aeronautics and
Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

RASS -- Skill vs. Yield



But in the
Tropics Q0
(Best Data
Only) Gets
the Best
Retrieval
Anomaly Skill
Score

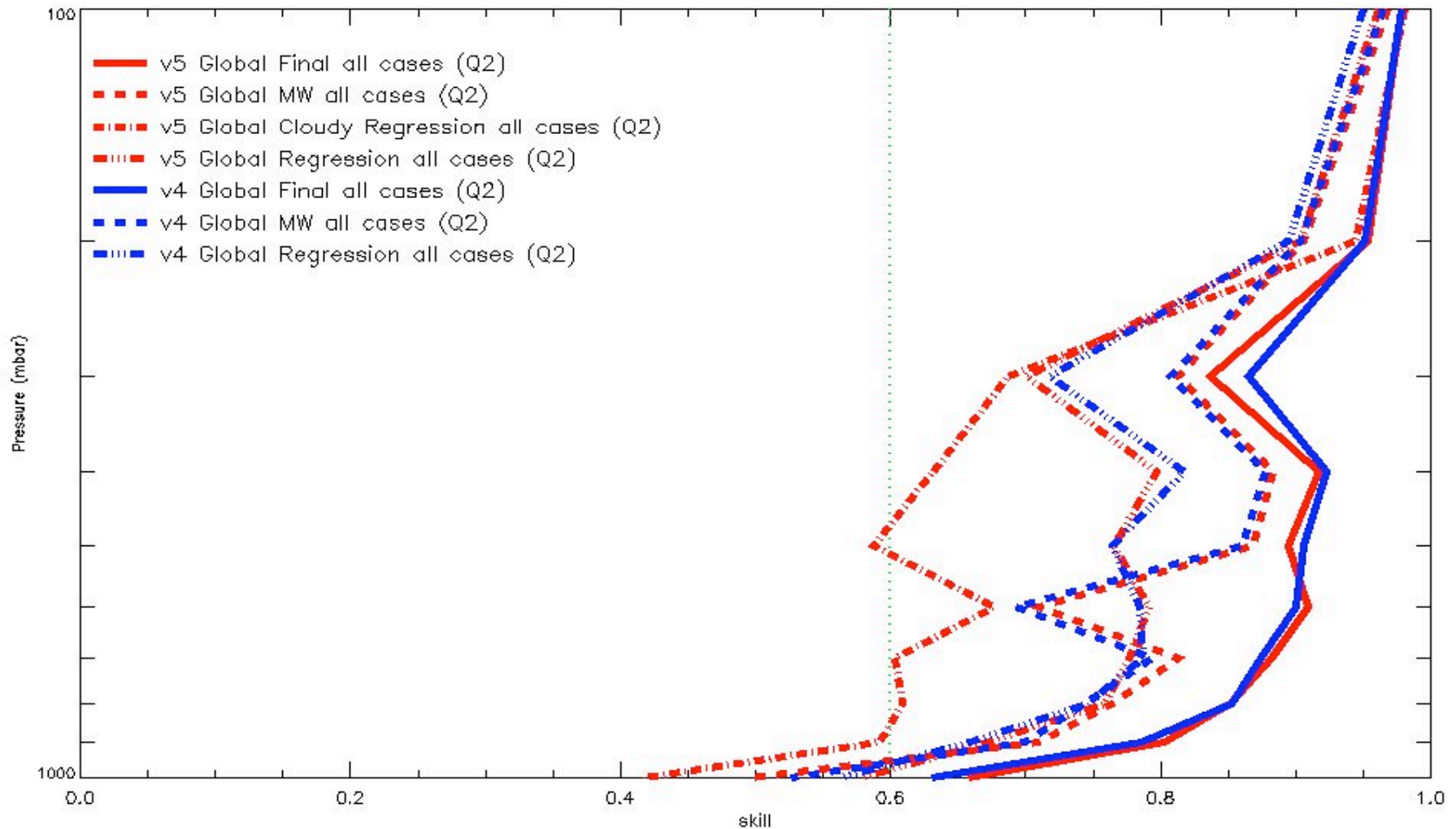


National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

The Physical Retrieval gets Consistently Higher RASS than v4 & v5 Regressions





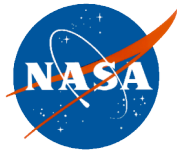
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

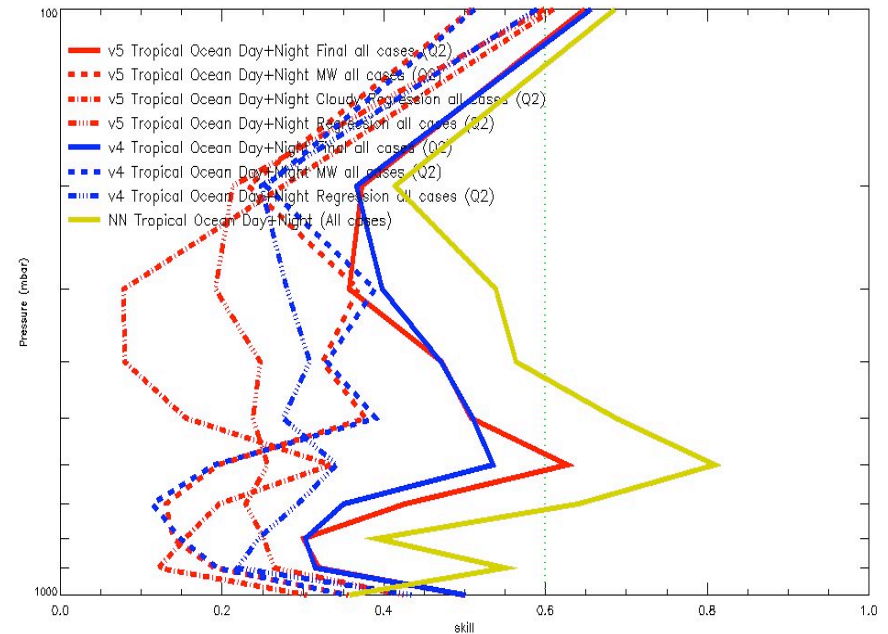
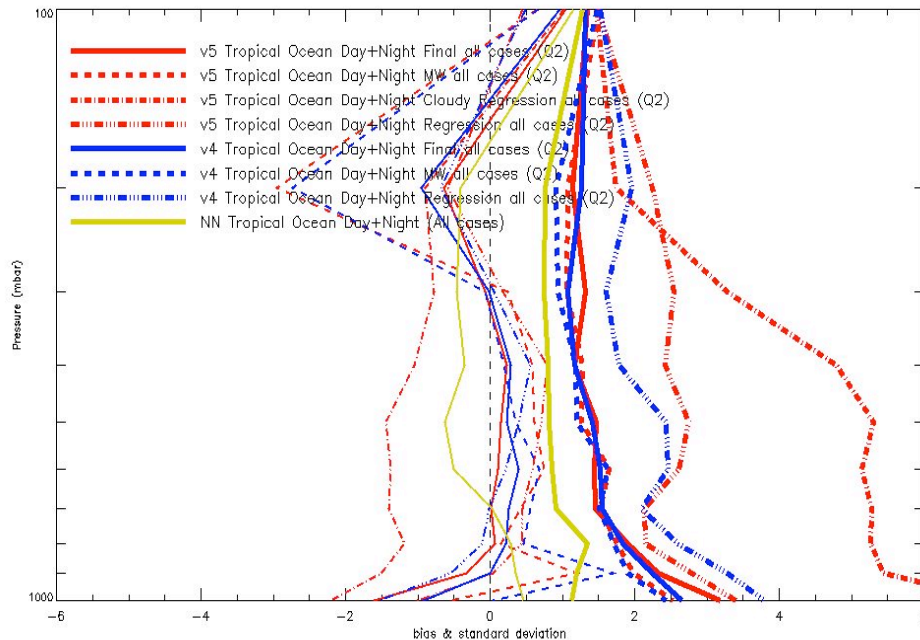
Using RASS to Investigate SCNN as a Possible Regression Replacement

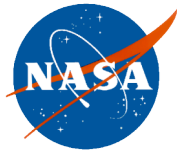
- Bill Blackwell of MIT has produced a stochastic cloud clearing plus neural network retrieval (SCNN).
- SCNN is a candidate to replace the current regression as a first guess into the physical retrieval for v6.
- A bug in the output translator for SCNN cuts out the lowest levels of the atmosphere.
- We compare temperature profile performance of SCNN with the current first guess algorithm using traditional bias & standard deviation and with skill score.
- All cases are included (Q2) because no quality control is available for SCNN.



Comparing SCNN with Current Algorithms in the Tropics

- SCNN looks best by either methodology
- RASS punishes SCNN near the surface for its low yield
- MW-Only is almost as good as final physical retrieval according to deviation but not by skill
 - Optimal estimation gives good statistics but AMSU-A provides relatively little information

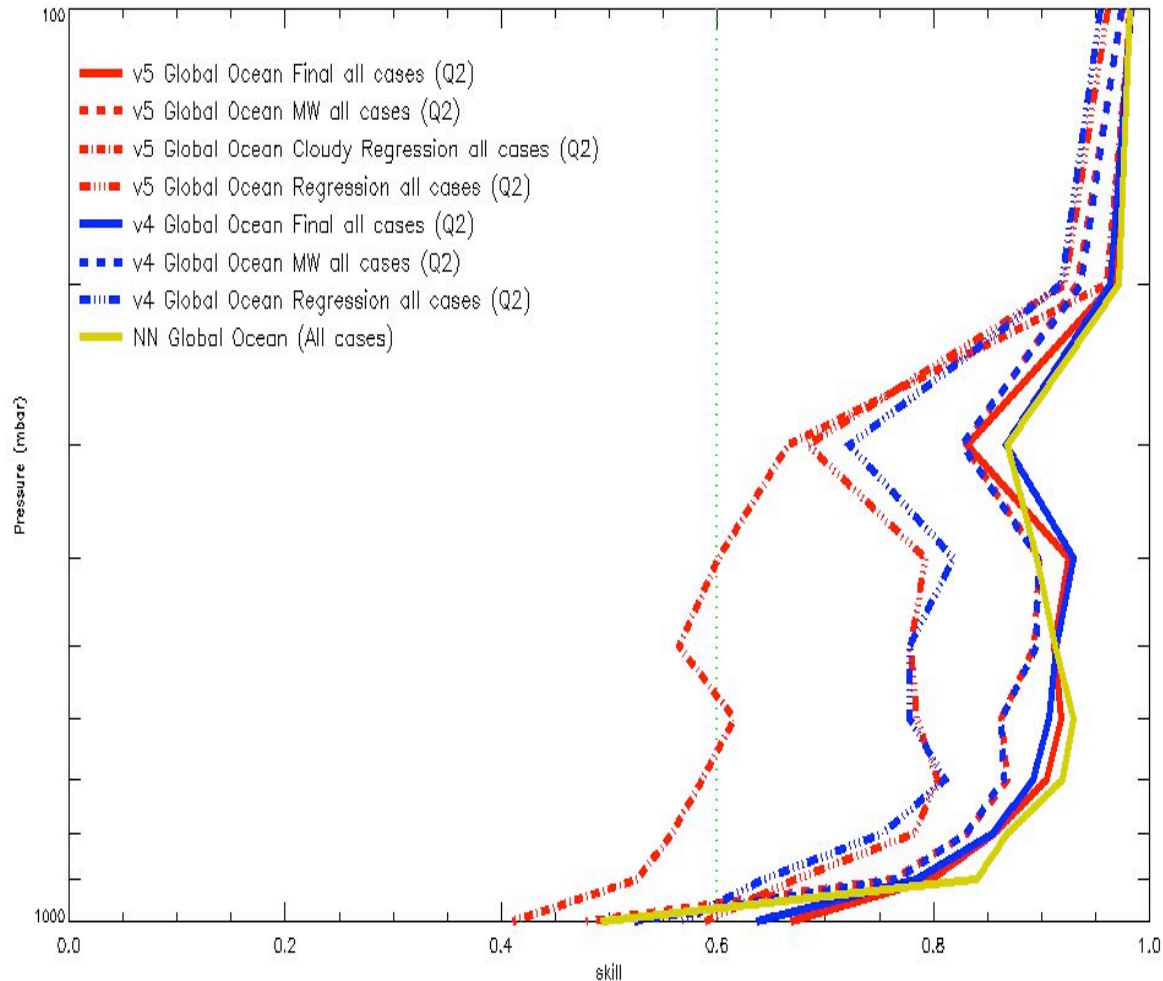




National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

RASS for SCNN Globally



Globally
SCNN is better
than all other
first guesses
and
competitive
with physical
retrieval



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

SCNN Next Steps

- From this analysis SCNN seems like a promising candidate to replace regression as a first guess.
- But this is a case study of SCNN, not a full evaluation.
- Bill Irion is leading an effort to do a robust evaluation including algorithmic review and implications for trends.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

RASS Next Steps

- The Retrieval Anomaly Skill Score methodology should be extended to other AIRS products:
 - Surface temperature
 - Lapse Rate
 - Water vapor
- RASS cannot be used globally for these because truth is rare:
 - Trace gasses
 - Clouds
- Test RASS on other AIRS retrievals:
 - Allen Huang of UW
 - Xu Liu and Daniel Zhou of LARC
- For a realistic assessment of absolute skill replace the NCEP 15 year reanalysis with the NCEP forecast as background and replace ECMWF analysis with Radiosondes as truth.
- Test the skill of retrievals with IASI



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Atmospheric Infrared Sounder

Conclusions

- With the operational availability of very accurate forecasts improvement in skill is more important than improvements in accuracy.
- RASS is an important tool for comparing retrieval versions.
- RASS with climatology background is straightforward to implement.
- Bias and accuracy remain important metrics