LIGHTNING CONTRIBUTION TO IMPROVEMENT OF PASSIVE MICROWAVE VERTICAL STRUCTURE AND RAINFALL ESTIMATION

6.4

Dennis J. Boccippio*
NASA / Marshall Space Flight Center XD-11, Huntsville, Alabama

1. Introduction

Most of the attention given to retrieval of geophysical parameters from passive microwave (PM) observations has focused on surface rainfall estimation (e.g., Kummerow et al, 2001). Diagnosis of convective / stratiform (C/S) state and vertically integrated hydrometeor content (VIHC) have also been performed (Hong et al, 1999; Grecu and Anagnostou, 2001). The techniques employed have either been blended physical/statistical (essentially generating a brightness temperature Bayesian "lookup table" from ensembles of Cloud Resolving Model [CRM] runs) or empirical/parametric, fitting linear or mildly nonlinear parameterizations using a subset of the available brightness temperature frequency / polarization observation ensemble (typically at or near 10H/V GHz, 19H/V, 21V, 37H/V and 85 H/V on prior, current and planned suborbital and orbital passive microwave sensors (DMSP/SSM-I, TRMM/TMI, Aqua/AMSR-E, NPOESS/CRiM, GPM/TMI). In the CRM/Bayesian approach, diagnosed profiles cloud water, cloud ice, precipitation water and precipitation ice content are also retrieved (Kummerow et al, 2001). These retrievals are only "as good" as the CRM physics, representativeness of the modeled CRM ensemble, and knowledge of confounding variates such as surface emissivity. We hypothesize that significantly more comprehensive and accurate estimation of geophysical parameters related to vertical structure is possible using a purely empirical retrieval with highly multivariate inputs (all available frequency / polarization pairs, texture variates, and nonlinear transformations of these such as VIHC estimates) and given ample degrees of freedom and nonlinear "capability".

Classification and regression neural networks (NNs) (essentially, multivariate, nonlinear categorical and continuous regressions with highly "flexible" basis functions) are empirical models that may provide the versatility needed to empirically extract subtle vertical structure information from PM observations. As (potentially) highly nonlinear models, they are susceptible to overfitting, unless a sufficiently large training / validation (T/V) dataset is available, and unless proper steps are taken to ensure regularization (excessive growth of nonlinearity in the models, and hence overfitting). The Tropical Rainfall Measuring Mission (TRMM) provides precisely the sort of T/V dataset needed for this purpose, if we choose co-located TRMM Microwave Imager (TMI) PM observations (and derived parameters) as model inputs and Precipitation Radar (PR)-derived quantities as outputs (targets). The available T/V dataset from the first 3 mission years alone is several times larger than current computational capacity (e.g., from a high-end desktop workstation) can utilize in NN fitting. Using the first 8 months of data from 1998 (as we do here), we are able to fit models with O(10^3) free parameters with several hundred training observations per free parameter, and still have ample unused co-located data to reserve for validation. As a bonus, co-located lightning observations from the Lightning Imaging Sensor (LIS) can be used as additional inputs to the neural networks, and the incremental benefit in retrieval from their inclusion can be assessed.

Fitting PM (and optionally lightning) observations to radar-derived parameters differs from the conventional technique of retrieval of "pure" geophysical state variables (such as rainfall or latent heating). However, there are compelling reasons to estimate radar-like parameters. These are already the "lingua franca" for many existing applications, such as forecast Data Assimilation (DA) (where assimilation modules already exist to ingest radar data) and end-user Decision Support Systems (DSS), e.g., in aviation. For both DA and DSS applications, absolute geophysical calibration is not necessarily critical (e.g., in some DSS' employing fuzzy logic / expert systems, the key step is to create an input which captures the salient geophysical *variability*, which is later transformed by an ad-hoc membership function anyway).

In this proof-of-concept study, we retrieve the following radar parameters. Wherever possible, we attempt to adhere to either "standard" NEXRAD parameter definitions, or "standard" TRMM PR 2A23 / 2A25 definitions: (1) Convective / stratiform state (2A23), (2) Bright band detection (2A23), (3) Overhanging anvil state (2A23 "other" category), (4) Surface rainfall (2A25), (5) Vertically Integrated Liquid (NEXRAD), (6) Ice Water Content (colder than 12C; Petersen and Rutledge, 2001), (7) 20 dBZ Echo Tops, (8) Probability of Hail / 45 dBZ altitude above 0C (NEXRAD), (9) Severe Hail Index (Witt, 1998), (10) Vertical Profile Type (Boccippio et al, 2005), (11) Full vertical reflectivity profile. At the time of this preprint's authorship, models retrieving (1)-(10) had been completed for both land and ocean, and models retrieving (11) were in the process of being trained.

Corresponding author address: Dennis J. Boccippio, National Space Science and Technology Center, 320 Sparkman Drive, Huntsville AL 35805;
email: Dennis.Boccippio@nasa.gov

2. Methodology

The retrieval methodology is motivated by the following hypotheses: (1) Significant “untapped” capacity to explain variability exists in the “full” (highly multivariate) PM observation ensemble and derived parameter dataset, (2) since PM observations are path-integrated measurements weighted by hydrometeor type and altitude (temperature), optimal extraction of variability may depend on isolation of “residual” signals between “primary” observations which often covary strongly, (3) the highly multivariate dataspace contains important subregions which must be identified (such as bright band effects, high frequency emission effects of supercooled liquid water, unique signatures from overhanging anvils, etc).

The retrieval process is thus stepwise. As a first step, a Principal Components Analysis (PCA) is performed on 33 available inputs (9 PM brightness temperature frequency / polarization pairs, 15 texture variables, 8 nonlinear transformations such as VIHC, and a composite lightning metric from LIS). This helps filter highly covarying inputs and reduce the dimensionality of the problem. 25 principal components (PCs) are used as the primary NN inputs. This is significantly more than are conventionally retained after PCA; however, pairwise analysis of radar profile vertical structure (using the profile typing scheme of Boccippio et al, 2005) reveals residual covariance of vertical structure with PCs as high as 20-25. Training data are “oversampled”” (multiple PR pixels per larger TMI 85 GHz pixel are included in the training sample), allowing the regressions to perform their own resolution blurring.

The next step is to identify important subregions of the highly multidimensional input parameter space. This is achieve by fitting classification neural networks to parameters (1)-(3) above (C/S, BB and Anvil state) and to binary classifiers for Hail Probability > (0, 50, 100%) and for Severe Hail Index > (1, 10). In principle, these classifiers help describe regions of important variations in physics within the dataspace. Once fitted, these 8 parameters are included with the the original 25 PCs as new input parameters (33 total) to the third step.

The third step is retrieval (via regression neural networks) of integrated and boundary parameters (4)-(9) (R, VIL, IWC, Tops, P(H), SHI), as the “next most complicated”. Once retrieved, these 6 parameters are included with the original PCs and 3 state classifiers (P(H) and SHI replacing the binary predictors), yielding a new set of 34 inputs to the 4^{th} step. This successive refinement of input parameters is conceptually no different from applying a physically-based nonlinear transformation to input variates in “conventional” (non-NN) regression. To indulge in an anthropomorphization, the physically-based transformations make it “easier” for NN’s to correctly navigate the multidimensional input data space and find physically-based local minima closer to the “correct” global minimum (a challenge in any nonlinear optimization).

The fourth step is prediction of radar vertical profile type, using the scheme of Boccippio et al (2005) (a cluster analysis of 3 years of TRMM PR column data, which identified 25 total [9 convective, 7 stratiform, 2 mixed and 6 anvil] “archetypal” vertical profile types). The output of this retrieval is probability of membership in each of the 25 type categories. Since each type corresponds to a Probability Distribution Function (PDF) of radar vertical reflectivity profiles in the full PR dataset, the 25 predictions can be combined to provide a “first guess” at the full vertical reflectivity profile (extracted at 40 temperature levels).

The 40-level first guess, 3 classifiers and 6 geophysical parameters provide 49 inputs to the final retrieval, a regression neural network for the actual vertical reflectivity profile (essentially, a “correction” of the type-based estimate, which tends to yield median values and miss extrema). *[Note – at time of writing; this network was still being trained].*

During each major step, a *different* training data subset was extracted from an 8-month (Mar-Nov 1998) full data sample. This helps prevent cumulative overfitting in the stepwise retrieval process. Furthermore, during the training of each individual network, an iterative process was used in which the training sample was split 50/50 8 times, with the first half used to fit model weights, and the second half used to diagnose overfitting. After 8 fits, an appropriate “stopping point” was then estimated, and the full 100% subset used to estimate the final model weights. Finally, a small weight-decay factor was included as a standard regularization technique to further prevent overfitting (excessive nonlinearity). Between the three techniques, we are confident that the resultant models are not overfitted (as confirmed by error analysis of predictions using the unused, or validation, dataset).

3. Results

Figure 1 shows the skill scores for three of the “first-step” classifiers (C/S, BB and Anvil state). Each NN outputs a 0-1 probability of membership in the given category (for C/S, 1.0=convective). The left panels are standard Receiver Operating Characteristic (ROC) plots, which show the POD and FAR as a call/no-call decision is applied to each threshold probability from 0-1. Dark solid, solid and dashed curves correspond to Convective, Bright Band and Anvil models, respectively. Curves closer to the upper left corner represent “better” (more robust across their entire range of output probability) models. The right panels show, for the same range of threshold probabilities, the model Heidke Skill Score (HSS), with the peak score marked with a diamond (and with that threshold overlaid on the corresponding ROC curve). Overall, C/S separation is easier over land than ocean. (For comparison, but not shown, the ROC curves of conventional texture or polarization rule-based PM C/S classifiers (Hong et al, 1999) lie very close to the diagonal, or no-skill line.)

Conversely, bright band detection is easier over ocean. Anvil status can be diagnosed with fairly high POD / low FAR; however, the HSS curve shows this is largely due to the rare incidence of anvils with > 17 dBZ echo in the PR dataset; the actual model skill is fairly low.

Note that Fig. 1 (and subsequent plots) do not formally show retrieval "errors", since the models are overtrained / oversampled (multiple PR pixels used for each lower resolution TMI pixel). The plots show the *combined* effects of intrinsic PM resolution blurring *and* the inherent ability (or inability) of PM observations to infer the corresponding radar parameter.

Figures 2-3 show the "errors" in retrieval of 4 of the integrated / boundary parameters (VIL, IWC, Echo Tops, P(Hail)). The plots are standard box-and-whisker plots of the distribution of predicted (TMI+LIS) vs true (PR) values. In general, the models tend to slightly overestimate parameters at the low end and underestimate at the high end, consistent with the expected effects of PM resolution degradation. Crudely, model errors (based on their RMSE and / or quintile widths in the boxplots) are approximately 10% over the parameters' dynamic range (note: VIL and IWC are in log [dB] units).

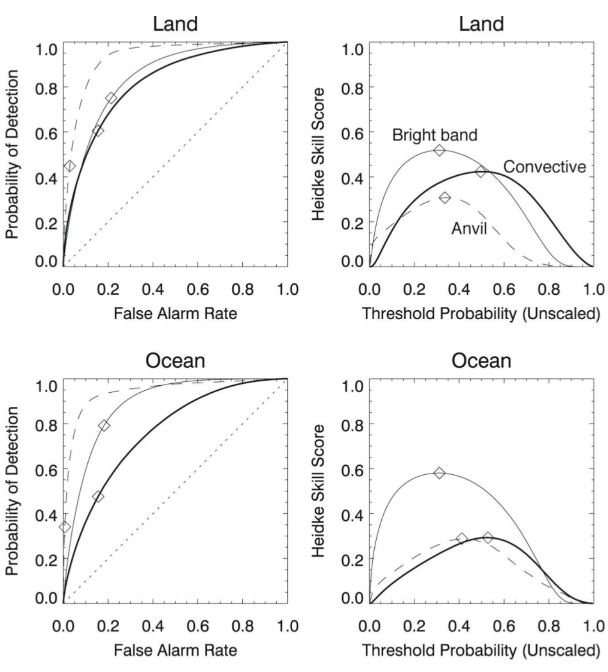

Figure 1 – Stage 1 (C/S, BB, Anvil state classifier) retrieval performance.

Prediction of the radar vertical profile type (using the 25-category classification scheme of Boccippio et al, 2005) is illustrated with several sample. Figure 4 shows two TRMM orbit PR swaths, with radar truth shown in the left columns and PM/lightning retrievals shown in the right columns. Recall that vertical profile types are predicted probabilistically; i.e., the probability of membership in each of the 25 type categories is predicted. As such, the entries in the profile type color table are weighted by these probabilities and combined for the TMI/LIS retrieval type subplot panel. For reference, the C convective categories correspond to warm (C1), mixed-phase tops (C2), deep (C3), and very deep / wet growth (C4) profiles, and the S stratiform categories correspond to warm (S1), cold with bright band (S2) and "MCS-type" deep with bright band (S3). Clearly the overall agreement of the results in these cases demonstrates that the retrievals are suitable, e.g., for inclusion as interest fields in an expert-system based DSS.

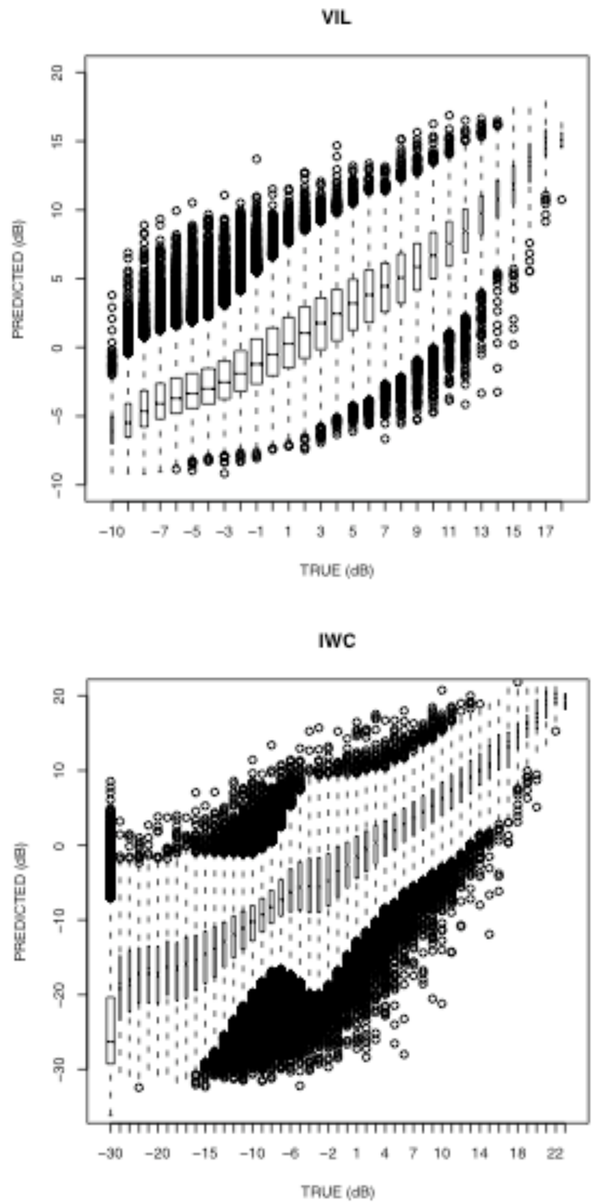

Figure 2 – Stage 2 retrievals of Vertically Integrated Liquid (NEXRAD defintion) and Ice Water Content above -12C (Petersen et al, 2003).

Figure 5 shows an example of the retrieval applied across an entire TMI swath (759 km, compared with the 215 km PR swath from Figure 4) and sampled at the native TMI 85 GHz pixel resolution (5x7 km) rather than oversampled at the PR pixel resolution (4x4 km).

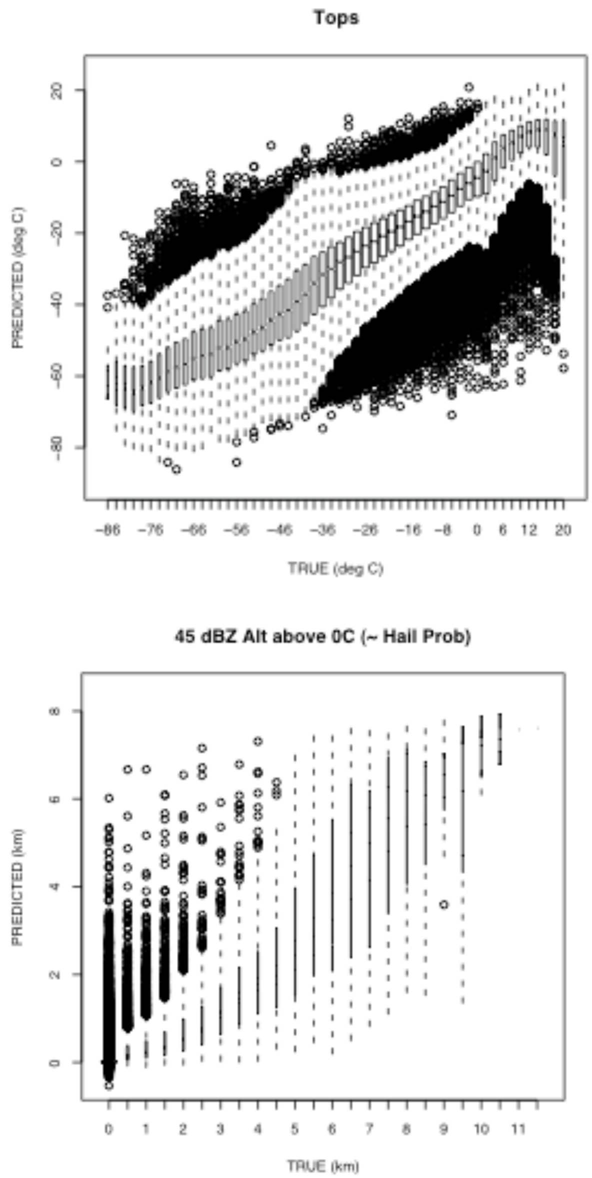

Figure 3 – Stage 2 retrievals of 20 dBZ Echo Tops and 45 dBZ echo top altitude above 0C (varies as Hail Probability).

4. Lightning Benefit

Results shown thus far are for models using both observed PM *and* lightning as inputs during their training. Fortuitously, the input lightning metric is isolated in either one (over land) or two (over ocean) PC's of the original input data (its loading on other PC's is negligible). One further regression NN is thus trained, a *prediction* of the lightning PC(s) from the non-lightning PC's. (The fact that these are all principal components does not preclude inference of one from other, it simply means that a *linear* prediction model does not exist [by definition of PCA/EOF analysis]. NN's, on the other hand, are intrinsically *nonlinear* models). The same series of trained NN's can thus be used either with observed lightning as inputs, *or* predicted lightning. By comparing these, the incremental benefit of including actual lightning observations in the retrieval can be assessed. (Note that this does not fully document the lightning "benefit", as part of this benefit is "hard-wired" into the models' weights during the original training process.)

At the time of writing, this comparison has only been performed for the ocean retrievals, with results:

P(Convective)	+6%	+10%
P(Bright Band)	-8%	-33%
P(Anvil)	-13%	-14%
Rain	-2%	-1%
VIL	-1%	+5%
IWC	+10%	+19%
Echo Tops	+13%	+13%
P(Hail)	+18%	+19%
SHI	+13%	+14%

The value range considers either all, or only cold-topped, profiles; % gain is computed from the models' RMSE (continuous parameters) or cross-entropy (categorical parameters) on the validation dataset.

The greatest benefit is found, expectedly, for retrieval of Ice Water Content, Hail Probability (45 dBZ altitude above 0C) and Severe Hail Index, all of which strongly covary, and Echo Tops. Observed lightning also contributes nontrivially in convective / stratiform discrimination. Interestingly, no gains are found in rainfall retrieval; either observed lightning does not provide a strong enough signal, or its benefits have already been incorporated during the training phase and the predicted lightning is "sufficient" to incorporate them. For two parameters (bright band and anvil probability), observed lightning worsens the retrievals. To some extent, this is understandable; when lightning occurs in or extends to anvils or trailing stratiform regions, the NN's attempt to "over-convectify" their predictions. This effect is muted when predicted, rather than observed, lightning is used (the prediction "smooths out" the discrete, binary occurrence/non-occurrence of a single flash during the 83 second LIS observation window).

It is useful to note that the lightning "benefit" (up to 20%) is comparable to the cost differential of simple optical orbital lightning detectors vs. passive microwave radiometers. This would argue that lightning sensors supplementing, e.g., NPOESS constellation radiometers would provide a reasonable return on investment.

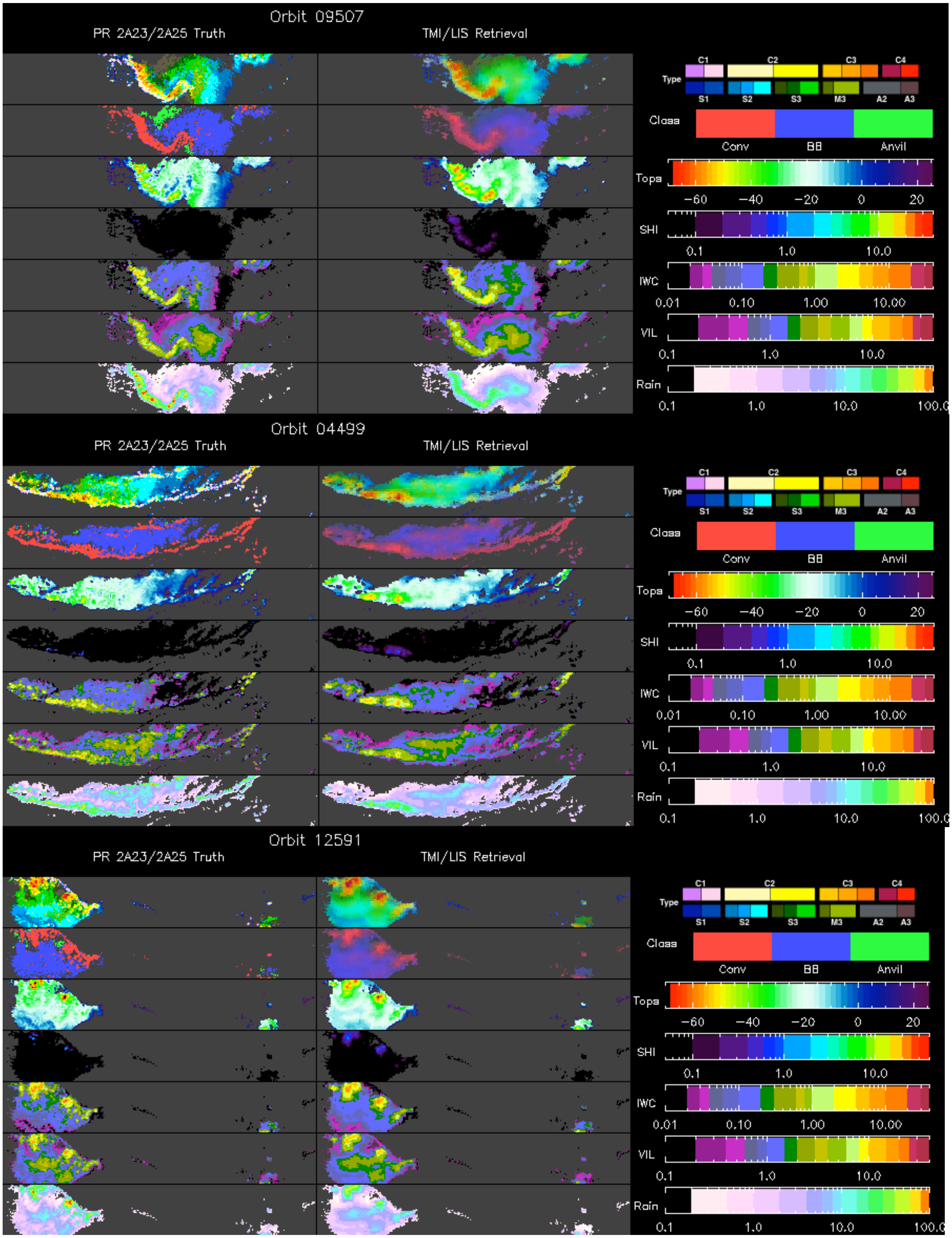

Figure 4 – Sample retrieved orbit swaths (radar truth, left; passive microwave + lightning retrieval, right).

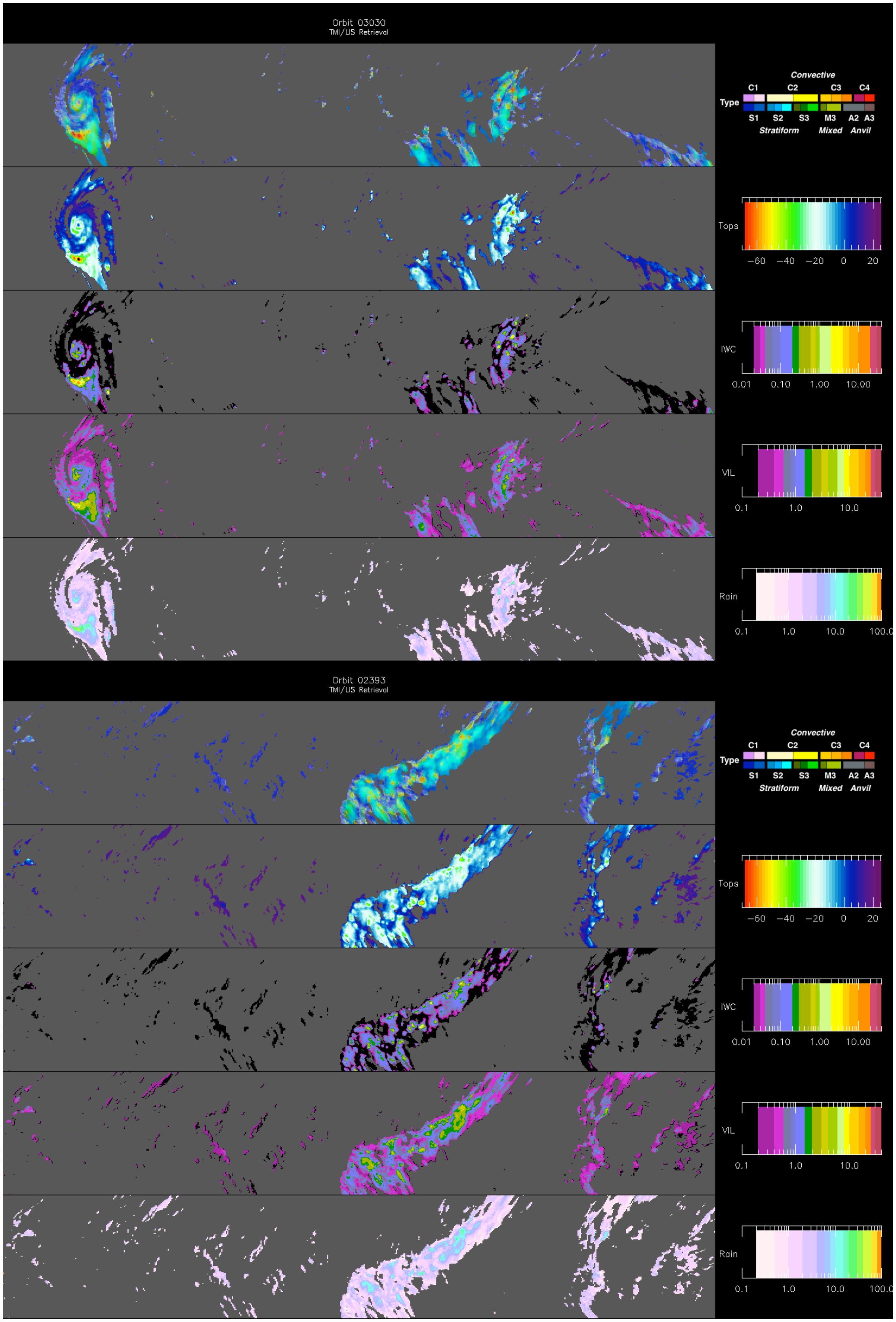

Figure 5 – Full TMI-swath retrievals. Insets show profile type, echo tops, IWC, VIL and rain, respectively.

5. Discussion and Conclusions

This proof-of-concept study demonstrates that direct, purely empirically-based retrieval of radar parameters (including state classifiers, geophysical integrated and boundary parameters, and vertical profile structure) is possible from passive microwave observations. Concurrent lightning observations can optionally be used to improve ice water content, hail threat and echo top estimation as well as convective / stratiform discrimination, with error reductions of up to 20%.

While not shown exhaustively, retrievals are possible over both ocean and land. Interestingly, the land retrievals make significant use of low frequency PM inputs, despite the fact that these are often presumed to too contaminated by unresolved surface characteristics to be useful for rainfall retrieval. In general, land retrievals perform <5% worse than ocean retrievals.

Retrieval quality appears to be adequate for use in Data Assimilation applications (where ad-hoc nudging schemes must be applied to the data anyway) or Decision Support Systems (where, e.g., in fuzzy logic-based systems, ad-hoc membership functions must be applied to the data anyway). Predictors for important vertical structure characteristics such as bright band existence or high supercooled liquid water contents (via the hail predictors) may also be useful inputs to blended physical/statistical rainfall retrievals.

The capability of generating a moderate-quality “virtual radar” retrieval from Low Earth Orbit passive microwave observations may be useful over the next 20 years. While orbital radars such as the TRMM PR are rare (the PR itself, and a possible Global Precipitation Mission follow-on), passive microwave radiometers have been deployed operationally for over 20 years (DMSP/SSM-I), are experimentally deployed now (Aqua/AMSR-E) and in the future (GPM), and will continue to be operationally deployed and enhanced (NPOESS/CRiM). Current NPOESS constellation plans call for significant improvements in passive microwave swath revisit time at individual ground locations. Particularly in regions (coastal, offshore, deep tropics) where volumetric radar coverage is absent, this can significantly improve the usage of passive microwave observations in DA and DSS tools. Retrieval of quantities as “standard” radar products further improves the likelihood of research-to-operations transition, as existing assimilation modules or membership functions can be “re-used” to accommodate the virtual radar retrievals. The incremental retrieval benefit of co-located lightning observations suggests that low-cost optical lightning detectors would make suitable candidates for supplementary sensors on operational (e.g., NPOESS) platforms.

6. References

Boccippio, D.J., W.A. Petersen and D.J. Cecil, 2005: The tropical convective spectrum. I: Archetypal vertical structures. *J. Climate*, in press.

Grecu, M. and E.N. Anagnostou, 2001: Overland precipitation estimation from TRMM passive microwave observations. *J. Appl. Met.*, **40**, 1367-1380.

Hong, Y., C.D. Kummerow and W.S. Olson, 1999: Separation of convective and stratiform precipitation using microwave brightness temperature. *J. Appl. Met.*, **38**, 1195-1213.

Kummerow, C.D., Y. Hong, W.S. Olson, S. Yang, R.F. Adler, J. McCollum, R. Ferraro, G. Petty, D-B Shin and T.T. Wilheit, 2001: The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Met.*, **40**, 1801-1820.

Olson, W.S., Y. Hong, C.D. Kummerow and J. Turk, 2001: A texture-polarization method for estimating convective-stratiform precipitation area coverage from passive microwve radiometer data. *J. Appl. Met.,* **40**, 1577-1591.

Petersen, W.A. and S.A. Rutledge, 2001: Regional variability in tropical convection: Observations from TRMM. *J. Climate*, **14**, 3566-3586.

Witt, A., M.D. Eilts, G.J. Stumpf, J.T. Johnson, E.D. Mitchell and K.W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286-303.