Status of the Community Long-term Infrared Microwave Coupled Atmospheric Product System (CLIMCAPS), *formerly known as CHIMERA*

Presented by: Chris Barnet

With contributions from: Antonia Gambacorta and Mitch Goldberg

AIRS Science Team Meeting

Thursday, Mar. 24, 2016 (Session 05, 9:10 am)

Topics

- Why we changed the name
- Discussion of the philosophy of our approach
  - Essentially this is the requirements for CLIMCAPS
- Discuss how this led to certain design features of the AIRS v6, NUCAPS
  - Essentially this is the justification for selecting the AIRS science team approach at NOAA
- Key differences between NUCAPS and CLIMCAPS
- Schedule

Why the name change

- Michael King asked that we cease using the acronym CHIMERA since it conflicts with the Cross-platform High resolution Multi-instrument AtmosphEric Retreival Algorithms CHIMAERA) system in use since 2010
- As a result we have renamed our algorithm CLIMCAPS (Community Long-term Infrared Microwave Coupled Atmospheric Processing System) and will use that name from this point forward.
  - NUCAPS is now the "NOAA Unique Combined Atmospheric Processing System"
    - To indicate that now both IASI/AMSU/MHS and CrIS/ATMS are supported by NOAA into AWIPS-II
  - CLIMCAPS is the "climate" version of NUCAPS

Common Philosophy behind Joel's v6, NUCAPS, CLIMCAPS

- Algorithm and code should be open source
- Algorithm should function on all operational modern hyperspectral and advance microwave sounder space-borne instruments
  - Minimize instrument dependent features
  - Exploit the full information content of the measurements
    - Ability to discriminate between physical correlations (e.g., climate sensitivity of dq/dT) and spectral correlations induced by measurements (e.g., dq/dT induced by spectroscopy)

Philosophy 2/3

- Minimize dependence of things we don't know
  - Minimize sensitivity to clouds
    - Exploit microwave information
    - IR cloud forward models are not robust (in my opinion, but some day soon this will be false)
    - Sensitivity of radiances to cloud parameters (particle sizes and shapes, vertical density) are highly non-linear
    - cloud parameters are not well constrained by infrared or microwave sounder measurements alone
  - Minimize Sensitivity of products to interfering signals
    - dT, dq, dTskin co-varies with cloud signals
    - dT co-varies w/ dq, dO3, ... for IR, dT co-varies with O2 for MW
    - dq co-varies w/ dT, dCH4, dSO2, ... for IR
      - dq is significantly more linear for MW
    - If ignored, this “spectral” covariance can confound measurement of natural correlations (*i.e.*, Earth physical correlations)

Philosophy 3/3

- Desire a global, all season, all sky, all regions, retrieval
  - avoid regional or highly tailored a-priori terms
  - avoid datasets that are not available or not skillful in remote regions
- Derive formal and traceable error estimates
  - have either averaging kernels or error covariance output for each product
  - algorithm should fully characterize inter-correlation of products induced by retrieval
- Desire a real time (weather) and re-processing (climate) capability
  - avoid algorithms that are computationally intensive (either in CPU or memory requirements) if they do not add sufficient skill
  - avoid datasets with high latency (weather)

Algorithm approach, Slide 1/4

- Cloud clearing radiance (CCR) approach
  - high global yield
  - can parameterize clouds with a small number (~4) highly linear parameters
  - CCR sacrifices spatial resolution or vertical information content
  - Full spectrum is derived from a small subset of infrared channels
- Sequential algorithm
  - very fast
    - matrix inversions done on small # of channels
  - more robust globally
    - solves for the most linear components 1st
  - minimizes cross-terms in prior covariance by removing channels that have strong interference
    - very strong signal-to-noise for terms being solved
    - minimal sensitivity on a-priori
  - use off-diagonal elements in the measurement (*i.e.*, obs-calc) error covariance
    - bring in *a-priori* information inherent in the laboratory spectroscopy (*i.e.*, the spectral “fingerprint” of the gas being measured)
    - mitigate inter-parameter correlations induced by spectroscopy

Simultaneous vs sequential O-E approach

Simultaneous OE	Sequential OE
Solve all parameters simultaneously	Solve each state variable (*e.g.*, T(p)), separately.
Error covariance includes only instrument model.	Error covariance is computed for all *relevant* state variables that are held fixed in a given step. Retrieval error covariance *must be* propagated between steps.
Each parameter is derived from all channels used (*e.g.*, can derive T(p) from CO2, H2O, O3, CO, … lines).	Each parameter is derived from the best channels for that parameter (*e.g.,* derive T(p) from CO2 lines, q(p) from H2O lines, etc.)
A-priori must be rather close to solution, since state variable interactions can de-stabilize the solution.	*A-priori* can be simple for hyperspectral.
This method has large state matrices (all parameters) and covariance matrices (all channels used). Inversion of these large matrices is computationally expensive.	State matrices are small (largest is 25 T(p) parameters) and covariance matrices of the channels subsets are quite small. Very fast algorithm. Encourages using more channels.

Algorithm approach, 2/4

- employ vertical functions
  - allows use of highly accurate vertical model (i.e., 100 level radiative transfer needed to prevent systematic forward model errors) without a speed penalty
  - reduces dimensionality of Jacobians
    - faster computations
    - further reduces size of matrices, faster inversions
    - Perturbations approximate the vertical resolution of the instrument
      - Derivatives of forward model are more physical (more accurate?)
  - Act as a smoothing constraint
    - reduces the need for other *a-priori* terms to stabilize the solution

Algorithm approach 3/4

- Employ an SVD information content analysis
  - minimal sensitivity to a-priori
  - "do no harm", do not make up data we don't know
  - retrieval only moves where we have signal
    - can produce un-physical vertical structures
    - but can capture vertical structures that are a-typical
    - this may be undesirable and we are exploring this aspect with the NASA grant
  - alternative is a shape preserving retrieval by using *a-priori* constraints (i.e., O-E)
    - so ... this also “makes up” data in regions of low information content domains
  - both SVD and O-E provide answers that reside within the retrieval null space (*i.e.*, both minimize the radiances are are part of the family of solutions of this ill-determined inversion)
    - but models tend to properly digest O-E data
    - Theoretically, the O-E a-priori can be replaced with model background

CLIMCAPS: Will use a formal a-priori (sequential O-E)

- NUCAPS uses a regression operator for an ”a-priori” and does not use covariance as a constraint
  - Have an *ad hoc* “background” term for stabilization
- For CLIMCAPS we are going to explicitly include a formal *a-priori* for all state variables
  - Initially run with both a simple climatology and MERRA-2 for T(p) and q(p)
  - GSFC climatology for O3(lat,season)
  - Simple climatologies for CO2(p,t), CH4(lat,p,season,t), NHO3(p,lat), N2O(p,t), and SO2(p)
- We are going to have targeted discussions with selected members of the community to decide what a-priori is best for the majority of applications
  - Originally I wanted a “panel of expert users,” but there has been a resounding resistance to that idea.
  - Antonia is going to fulfill that role and prepare a one-on-one discussion with colleagues that have used AIRS products for climate process studies.
  - There is a trade-off between including sub-resolved structures versus being independent of any model
  - Could run ensemble of *a-priori* , if users wanted (it is just CPUs)

CLIMCAPS: Explore other options for the a-priori

- NUCAPS uses a regression *a-priori*
  - I consider this is a flaw, in that (a) measurements are used twice and (b) statistical operators are not <u>*constrained*</u> to the observations
    - As a result, we cannot derive a meaningful error estimate
    - Requires *ad-hoc* QC to remove errant cases
  - But it does provide high vertical resolution structure that will require more channels to be used in the physical approach (hence slower)
    - This can be a good thing (when correct) and a bad thing (when incorrect)
  - NUCAPS uses a single, globally derived, eigenvector solution (i.e., is regularized) for 4 scan angle regions
    - I believe this is preferable to a regionally trained regression or neural network because it has a significantly smaller number of degrees of freedom (and reduces spurious signals into the a priori that the physical algorithm can't remove)

CLIMCAPS: Improve the propagation of retrieval errors

- NUCAPS only maps the diagonal component of the retrieval error covariance into down-stream steps.
  - We have an ad-hoc oscillatory term
- It has been shown that there is a robust way to pass the full 2D retrieval error covariance from one step to the next (Eric Maddy Mar. 23 2007 AIRS meeting, and Eric Maddy, AIRS meeting, Apr. 27, 2011)
  - compressing the retrieval error estimate covariance matrix and only propagating the significant eigenvectors to the next step.
  - reconstruct the retrieval error covariance on new set of channels and use it to compute the measurement error covariance.
  - We believe 3-6 eigenvectors are sufficient (effectively current algorithm employs 1 sub-optimal eigenvector)

Open Source for NUCAPS and CLIMCAPS

- Antonia Gambacorta is leading the CLIMCAPS effort and will be providing documentation through ATBD, a set of peer-reviewed papers, and informal pathways.
- Nadia Smith and Jonathan Smith will be leading NUCAPS maintenance and will also be providing theoretical documentation and upgrades, if NOAA accepts them, from CLIMCAPS development.
- My goals is that CLIMCAPS and NUCAPS will be the same code, simply run with different namelists
  - Unless this makes the code too complicated.
- Today, NUCAPS can be easily run on any Unix machine and any compiler that I have encountered (e.g., it runs on Apple laptops using gfortran, LINUX boxes, CSPP at UW, etc.)
- CLIMCAPS will follow and enhance that philosophy.

Schedule

- NUCAPS code currently operates on Aqua/AIRS, S-NPP CrIS/ATMS and Metop-A, -B IASI/AMSU/MHS
  - Can be delivered to SIPS at any time for CrIS/ATMS
  - Other systems require separate preprocessors (all in IDL)
- CLIMCAPS is only funded for CrIS/ATMS
  - We would like to test it for AIRS if SIPS framework is robust enough
- Summer 2016: install a-priori and error propagation in T(p) and q(p) steps
- Late summer 2016: Propagate error through the cloud clearing step
  - Reduce or eliminate iteration of cloud clearing
- September 2016: install new emissivity a-priori and error propagation into surface retrieval
  - Right now leaning towards MEASURES product, to be delivered in Aug. 2016
- September 2016: Deliver of algorithm to SIPS for evaluation of T(p), q(p), surface products
- Oct-Dec 2016: convert downstream trace-gas steps from SVD to O-E

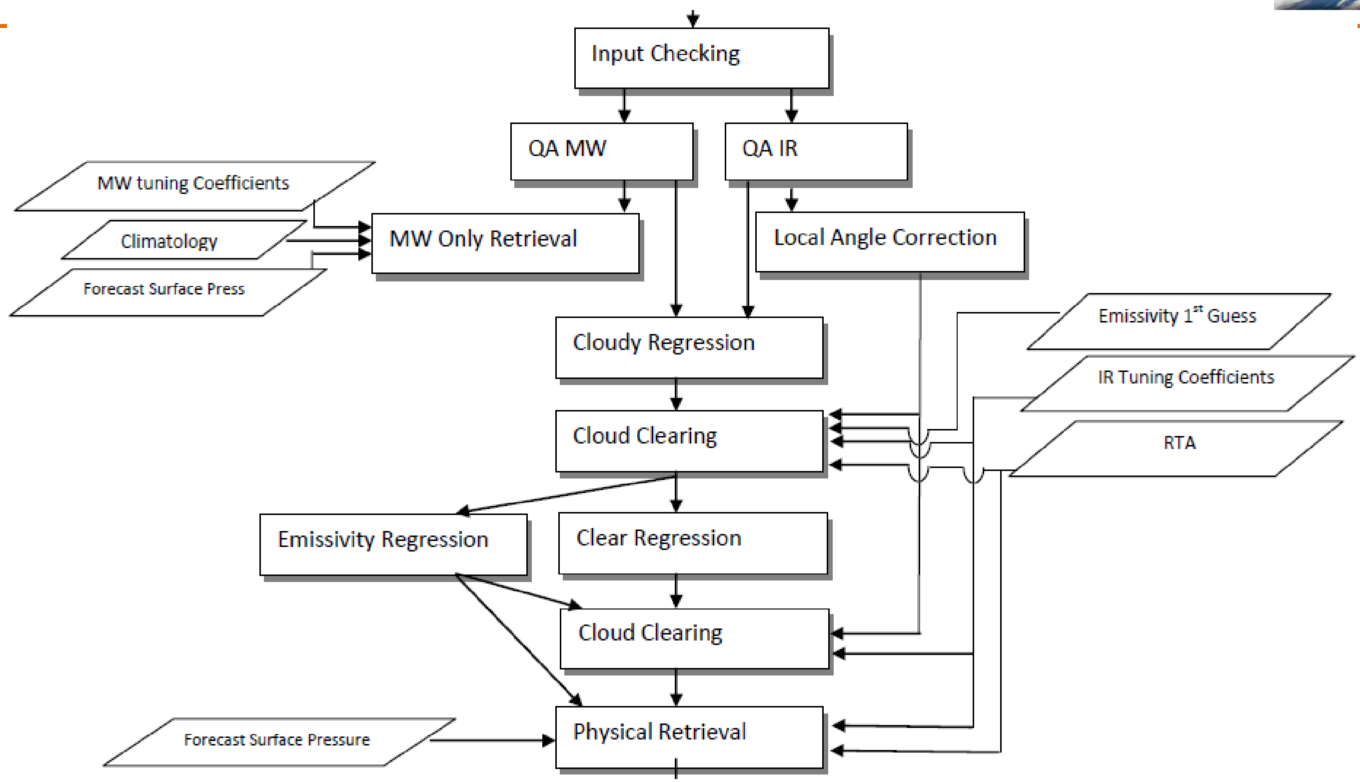
Summary

	AIRS v6	NUCAPS	CLIMCAPS	Irion
Regularization	SVD	SVD	O-E	O-E
Alg. type	sequential	sequential	sequential	simultaneous
Clouds	Cloud clearing	Cloud clearing	Cloud clearing	Solve f/cloud
A-priori for T(p), q(p)	Neural Net	Linear Regression	Climatology or MERRA-2	?
A-priori for emissivity	Borbas	Constant (v5-like)	MEASURES	?
A-priori for trace gases	Climatology, No covariance	Climatology, No covariance	Climatology w/ covariance	?
Error Propagation	Single term	Single term	3-6 eigenvector	n/a
Error estimates	Parametric fit to ECMWF	Propagated	Propagated O-E	O-E

NOTE: question marks reflect my ignorance and does not reflect uncertainty by the developer

QUESTIONS?

NUCAPS Flow Chart



- **I. A microwave retrieval module** which computes Temperature, water vapor and cloud liquid water (Rosenkranz, 2000)
- **II. A fast eigenvector regression** retrieval that is trained against ECMWF and CrIS all sky radiances which computes temperature and water vapor (Goldberg et al., 2003)
- **III. A cloud clearing module** (Chahine, 1974)
- **IV. A second fast eigenvector regression** retrieval that is trained against ECMWF analysis and CrIS cloud cleared radiances
- **V. The final infrared physical retrieval** based on a regularized iterated least square minimization: temperature, water vapor, trace gases (O3, CO, CH4, CO2, SO2, HNO3, N2O) (Susskind, Barnet, Blaisdell, 2003)