

2021-04-29

CRTM Monthly Meeting Protocol

Core Topic of the Meeting: CRTM v2.4.1 release planning

Date: 2021-04-29 **Time:** 15:02h

Location: Virtual (Google Hangouts)

Invited Speakers: -

Meeting Chair: Benjamin Johnson (JCSDA)

Keeper of the Minutes: Patrick Stegmann (JCSDA)

Attendees: Benjamin Johnson, Patrick Stegmann, Cheng Dang, Shih-Wei, Scott Sieron, Jim Jung, Andrew Tangborn, Hui-Ya Chuang, Haixia Liu, Yingtao Ma, Kevin Garrett, Daniel Abdi, Sarah Lu, Yiangqu Zhu, Nick Nalli, Quanhua Liu, Bob Potash

A g e n d a l t e r n a t i v e	CRTM v2.4.1 release planning
D i s c u s s i o n	<p>Introduction by Ben:</p> <p>Ben: Couple of items on the list today. I have some slides on Release 2.4.1 and then go through AOP2021 and then we can have an open discussion about AI and where people see potential for improvement. Hui-Ya has to leave early, so she can give an update early.</p> <p>My question for her is whether we should update UPP?</p> <p>Hui-Ya: We only want to keep two versions of GSI on WCAS, so we have to update. Is this the one with the OMP over channels?</p>

Ben: Yes, we want to have a consistent, solid basis for the 2.x version for the next couple of years. Ok, let me share my screen real-quick. 2.4.1 is a bugfix release, so there's not much code change. Not a lot of these are implemented yet, but work is ongoing. OMP over channels was implemented by Daniel Abdi and it's pretty intelligent on how to do that. Ming is working on the snow emissivity, hopefully he can give some updates. The biggest challenge for 2.4 is the netCDF implementation. That's because it's using a module file that needs to be consistent with your compiler. Cheng has been working on updating the coefficient files, specifically NAPS. Still ongoing effort to modernize the binary system. We have moved all the binary files to git LFS and gzip'ed them in order to minimize the download burden. I'm working on a method to do this automatically. We're continuing the work on test codes and also to create standardized test output. There's sort of a release standard output. In terms of sensors, there's GOES ABI delivered by STAR. Patrick is working on IASI-NG and also released the new updated TROPICS and is working on GEMS-1 and 2. We also provided the v2 ACCoeff for Metop-C AMSUA to EMC. The other part is version 3. As we finish up v2.4.1, this will be the basis for the v3.0-beta. That will require a lot of documentation. This will be a big work for us in the upcoming months. The testing framework is pretty comprehensive already. We want to add a few new aerosol tests for the new coefficients. Haixia has been working on the evaluation of the ABI_G17_81K coefficients and she found some differences. The differences are large enough to be of interest. Maybe STAR wants to comment. Ming and Haixia, and Patrick have been working on a comparison using the GSI. Haixia, did you plan to show any slides today?

Haixia: Yes, sure. This is just simply comparing the new GOES-17 coefficients compared to the old 60K coefficients. For just one cycle all the 10 IR channels are shown with BT differences on the y-axis. For channel 9 my differences are larger than from Yingtao and the university of Wisconsin. For Ch. 12 and 14 Yingtao doesn't show much difference, but for my work, it shows 0.2K differences. And I am wondering where these differences are coming from. For Ch. 9 the biases are actually smaller.

Yingtao: When you make these comparisons, are you using the old coefficients?

Ben: Yes.

Yingtao: Part of the difference between 81K is attributed to the SRF. The new 81K is using a different SRF. The other part is because the original 60K coefficient was trained using the old version of LBLRTM. That may also contribute to the difference. When I check the difference, I use the UMBC48 training profiles and I noticed nearly the same difference. Haixia, I remember you did another case.

Haixia: Yes, I did some other cycles, but the results are similar. My understanding is that 81K and 60K are using different SRFs, but the other difference is LBLRTM training?

Yingtao: Yes.

Ben: Did you touch base with the Wisconsin people at all?

Yingtao: We actually did a comparison early and found close differences. They are using a standard profile.

Ben: When I look at the mean values in the plot, the means line up pretty well with what they predicted. I guess what I would like to find out about is additional information they came up with since it's been about a year since they created that. I can reach out to Tim.

Yingtao: What did you expect in terms of differences?

Kevin: Can I share my screen for a second?

Ben: Sure.

Kevin: This is the comparison we did when we trained the 81K with the new version of LBLRTM. This is consistent with what Yingtao said. These are the differences between 81K and 60K. The middle one is PFAAST, that's what Wisconsin used. So, I haven't gone back to reconcile with what Haixia is showing. But these are at least consistent in terms of SRF response.

Haixia: How many datasets did you use?

Kevin: For PFAAST it's only one, and for the ECMWF it's about 30000.

Haixia: For the third channel all these results show 0.1 Kelvin, but my differences are much larger, about 0.2 Kelvin. We're not actively assimilating those two channels right now, but if we can find out the reason that would be great.

Yingtao: Although your number is much larger than what Kevin just showed, your version of coefficients gives similar mean and bias when you are using a similar atmospheric configuration.

Haixia: You see channel 12, the ozone channel...

Yingtao: So, it's above 0.9?

Haixia: No, Kevin's figure shows the difference between red and blue bar, so it's about 0.2 Kelvin.

Kevin: I guess you can attribute it to the SRF and the LBLRTM version.

Ben: That's probably the primary reason here.

Kevin: Are the changes with the new LBLRTM version meaningful and helpful. I would be interested in seeing the difference in the standard deviation.

Quanhua: The only difference here is SRF data?

Yingtao: No, the training is also difference.

Quanhua: And the ensemble is more closely to 81K?

Haixia: Yes.

Ben: And that has been confirmed by a couple of groups.

Quanhua: Looks like for most channel not so big, only channel 16.

Haixia: Yeah.

Quanhua: That looks like strange with temperature difference. Other channel very similar.

Haixia: Is channel 14 affected most? That's why the 81K coefficients perform most differently.

Ben: Will these channels be used?

Haixia: No, that's an ozone channel.

Hui-Ya: Quick question for Haixia, did you try to play around with threading to help with speed?

Haixia: Not threading, just OpenMP. I had email communication with Ben and I didn't see a lot of speedup.

Ben: Yes, in 2.4 we're only parallelizing over profiles and GSI is only calling one profile at a time.

Hui-Ya: Sounds like UPP should wait then.

Haixia: Ok, this slide is just a IASI Metop-A near-surface channel. Left is old CRTM and right is new CRTM. The histograms look similar, but the number of profiles that pass quality control actually increase. This is due to the surface emissivity bug fix in v2.4.

Quanhua: But the other thing looks like slightly worse. Very small change.

Haixia: Minimum and maximum are similar and the mean is very small.

Ben: We'll keep monitoring it.

Haixia: Not just IASI. All the clearsky instruments have more profiles passing QC. But I see that the O-F stdev is getting smaller.

Yingtao: What did you fix for this?

Haixia: I looked at the code. It's in the Common_RTSolution module to compute the emissivity.

Ming: That's general, not only in the IR. That's what I did, it's very important.

Ben: Ming, can you just review the fix.

Ming: Sorry, I forgot.

Yingtao: That's what Ming modified if certain channels are invalid.

Ming: We have a lengthy documentation on this.

Haixia: Let me share my screen. Is this too small? Let me zoom in. On the left is v2.3 and on the right is 2.4. There are a lot of changes. The surface emissivity goes to this logic block. Currently we do not have this IF, so v2.4 is actually using this emissivity calculation.

Ming: I think it's better, we had a long discussion before. Ben should have that documentation.

Yingtao: So, this bug has existed for a long time.

Ming: Yes, since two years.

Ben: Yes, the fix was supposed to be in v2.3.1.

Ming: Mark I think you also have that documentation.

Quanhua: Yes, I remember there was a discussion but after two years I forgot the details.

Ben: Ming, If you still have it please send it to me.

Ming: Yes, in the sentiment of close cooperation I think we should have <unintelligible>.

Ben: I put Haixia on the spot, so I'm not gonna do that anymore. Cheng, do you have any additional slides?

Cheng: No, I'm just gonna give a summary. So basically, we are providing more aerosol models now and the interface provides support for both binary and netCDF. For 2.4 did a lot of effort to include CMAQ and I added the new GOCART. Basically, for the v2.4.1 release we provide more flexibility for the user.

Ben: Where are we with the CMAQ tables? I remember there were large differences.

Cheng: Yes, Hongli showed about two orders of magnitude lower AOD. I'm not sure how he is going to proceed. The slides just show a comparison of AODs. You can see with the default LUT and with different LUTs you can have slightly or very different AODs. This is something we can tune in the future.

Yingtao: Cheng, I remember the CMAQ coefficients, the conclusion is mostly because CMAQ is mostly for dry environment. Is that the final conclusion?

Cheng: I have concerns, because I see differences in dust, which is hydrophobic. There are more things we need to consider. But you are right, GOCART GEOS-5 has humidity.

Quanhua: Cheng, there is code change from 2.3 to 2.4 on netCDF. After 2.4 do we need code change?

Cheng: Yes, we need to do some code changes, because each table has their own specifications.

Quanhua: Because right now we have two tables. For CMAQ, nitrate shouldn't be different.

Cheng: I didn't generate tables for nitrate. If you want to assimilate nitrate you can select GOCART GEOS-5 as a LUT.

Quanhua: Ok, that's because of how the table generated by Boulder people might be different. That's what I suggest, that you and Yingtao think whether in the future we have more generic code, rather change the code by different table. In the future when the Navy has table and other people have table it's not easy to change code.

Cheng: We don't need to change the code at this point if we are ok with having very larger aerosol tables. The reason we didn't do this is because we would end up having extremely large tables.

Quanhua: Maybe in the future.

Ben: That is true for all CRTM tables. It puts a burden on the people generating new tables. It's a really nice way of doing it like in JEDI. Okay, Yingtao, any updates on NLTE?

Yingtao: I uploaded the coefficients. Do I have to compress the coefficients?

Ben: No.

Yingtao: Ok, for documentation I will put the AMS slides.

Ben: We also need the little-endian files.

Yingtao: Ok.

Ben: Ok, work with Patrick. I think he can do that pretty quickly.

Quanhua: Usually we generate coefficients from netCDF to binary.

Yingtao: I only work with binary coefficients directly.

Ben: Just be sure to use CRTM 2.4 to use netCDF4.

Yingtao: I will.

Ben: Just on that point because of the AOP meeting that the NESDIS contribution is lower than it should be...

Yingtao: That is what I am concerned about. When Mark is doing development, he is doing that on his own computer with thousands of commits, but on Github he is only uploading the end result, so it counts as one commit.

Ben: Yes, I know it's not easy and I have made that comment with management. It is required for in-kind. The alternative is if you are using your own repository is to merge your commits into Github.

Yingtao: Last time I looked at the slides, it looks like NESDIS is at the bottom of the contribution.

Quanhua: I agree.

Yingtao: But Mark, for that you need to learn how to use github.

Quanhua: I think it's not good that it only depends on how many commits you make.

Ben: I agree, but it's kind of a new way. When you look at the UK Metoffice they put everybody to shame with less people. But I am happy to provide more training. Thanks. Alright, I already talked about the spectral coefficients. The 431 stands for what?

Yingtao: It's only a subset, not a full set. Andrew selected that from EMC.

Quanhua: Maybe Haixia can confirm are you still reading BUFR data with 431?

Haixia: I think it's full spectrum BUFR.

Jim: Yes, the 431 is a subset of the full resolution. Paul van Delst wrote some code so you can subset on the fly. That's the way the code is set up right now. If you want to change the channel things get messy. Both CrIS, IASI, and AIRS use subsets, and they are subsetted within the GSI.

Ben: That's the channel subset code within the CRTM.

Quanhua: Looks like from the CRTM you provide all the channels and the user is choosing.

Kevin: But it's the 431 set that is actually assimilated?

Jim: Yes, we originally tried the full set and it worked for one instrument, but not for two. This really helps with the research side.

Quanhua: So right now, NCEP is using the full set for CrIS?

Jim: In operations they're using 431. But they're using the CrIS FSR full resolution coefficient files.

Ben: But the tanks are 431?

Jim: Yes.

Ben: Jim, any news on IASI-NG subsetting?

Jim: I can't talk to IASI-NG subsetting. For subsetting you don't save any memory. For AIRS there are several subsets with multiple coefficient files, but here we have only the one.

Ben: Any more comments?

R
e
s
u
l
t:

N/A

T
a
s
k
s:

N/A

R e s p o n s i b l e P e o p l e:	N/A
D e a d l i n e:	N/A

A g e n d a l t e r n a t i v e:	AOP2021 and CRTM v3.0 Introduction
D i s c u s s i o n:	<p>Ben:</p> <p>Just briefly about AOP2021. There are three basic core items here. AOP2021 starts on April 1st, 2021. We have a big focus on applications in the JCSDA to highlight the massive amount of work that has been ongoing behind the scenes. I don't have time go through all these things here. This list is what we have to do to stay current with our task lists. What you can see here are tasks that are kind of overhead. The Simobs is basically an hofx Jacobian simulator, similar to pycrtm, but inside the UFO. The other part is of course continuing the transmittance work by Patrick in coordination with GMAO and STAR. CRTM-AI is sort of a high-flying task to see what is possible. In 2021 this is not a big part of our effort, which is on request of the MOB. We have generic interfaces for aerosols and clouds and transmittance and prototype development. Just to look at Simobs, it's an ambitious plan. The idea is to have a python interface and create and hofx plot depending on whatever input data are available, quickly followed by the ability to compute Jacobians. The goal is to create computed radiances in conjunction with observations, which is very powerful for doing science, especially when looking at statistical observations. As we ramp up this Python side, that helps us understand instrument health.</p> <p>Quanhua: What is ICBS?</p> <p>Ben: It's not really ICBS, it's not really about health monitoring. My ultimate goal is to have external model validation. This could be a really useful tool for validation.</p> <p>Quanhua: Will this be part of the current ICBS?</p> <p>Ben: That's up to the contributors. I don't want to step on any toes.</p> <p>Yingtao: Can this tool be used for outside users, like a student?</p> <p>Ben: This will not be a replacement for the UPP, but I would hope that the UPP would take some of the ideas. It would replicate some things that the UPP does.</p>

Quanhua: Just my thought, maybe at this time we should not emphasize this for 2.4.1. There are lots of places that are already doing this, like NESDIS and RTTOV. That could be confusing. If e.g. ATMS has a bias of 0.1K, but other centers are saying 0.2K...

Ben: Yes, it's not meant as an operational tool. But if this is something EMC wants they can have it. It's not just us who are developing this. There is sort of a broader application called Skylab that is sort of showing of everything that we have created in the Joint Center, starting with the observer – which is us – and going down to quality control and so on. The technology for all of the yellow and green part exists already, so we are heading towards a cycled forecast. You can see the similarities from the Simobs application and we are developing generic elements, so that they can be moved from application to application. There is a lot of work done in the background.

CRTM3 is just continuing the science work we have already been doing. The work with Greg Thompson is the microphysics work. A piece we are missing is the UV impact in the CRTM. The other thing is making sure that we are computing accurate coefficients, Zeeman code updates, NLTE updates. CSEM, we are working closely with Ming.

The final and critical part is the validation part. Often, we just say O-B looks good, or things like that. I think we can come up with a very strong framework to quantify the uncertainty.

The application outcome is just stat analysis and polarization impact. That's where we are getting some feedback from operations. As we are developing the AI methods, we want to see what's the cost-benefit analysis.

Just to wrap it up, things that have been integrated, transmittance work is ongoing, Tom Greenwald did some solver development even though he is not funded. We want to continue with additional optimizations like OMP from Daniel. Questions?

Quanhua: For Tom Greenwald, we should talk to management to support him. For other people, we have lots of projects, but for him that should be considered.

Ben: Right, but this is above my pay grade. But absolutely, we need that capability to support Tom and external contributors. I have advocated for this from day one.

Kevin: Yes, these are discussions to be had within the executive team. Right now, pretty much all the funding is going to the core team. It could be a more complicated process. But just FYI, Ping is funded.

Ben: Yes, I was looking for the Joint Centers capability to fund external partners consistently.

Kevin: It's maybe a discussion for the executive team to have some money set aside.

Ben: Anything else anybody wants to talk about?

R
e
s
ul
t:

N/A

T
a
s
k
s:

N/A

<i>R e s p o n s i b l e P e o p l e:</i>	N/A
<i>D e a d l i n e:</i>	N/A

Agenda Item 3:	Open discussion on ML/AI approaches toward accelerating CRTM computations
Discussion:	<p>Kevin: I think you wanted to talk about the AI. We have someone focused on passive MW to generalize Eric Maddy's work. Jin Ming is working on this. Probably at some point we can show those results. And he made a lot of progress to generalize the results and we are hoping to see big improvements to all-sky DA.</p> <p>Quanhua: For AI we still need two models in the future. We need to have people working on both. Otherwise we might find RTTOV making big progress in one direction.</p> <p>Ben: The key part I want to see if we are speeding up the CRTM, we can make the CRTM a little bit slower and add more science.</p> <p>Quanhua: Yes, but before we thought AI is just fast, but AI might also provide a way to handle complicated problem like emissivity.</p> <p>Ben: The possibilities with AI are improving every month. Ok, any final comments? Thanks a lot.</p>
Result:	N/A
Tasks:	N/A
<i>Responsible People:</i>	N/A
<i>Deadline:</i>	N/A

16:30h Final end of meeting.