The Cryosphere Discussions

The Cryosphere Discuss., referee comment RC2
https://doi.org/10.5194/tc-2020-363-RC2, 2021

Comment on tc-2020-363

Anonymous Referee #2

Referee comment on "Estimating subpixel turbulent heat flux over leads from MODIS thermal infrared imagery with deep learning" by Zhixiang Yin et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2020-363-RC2, 2021

This study evaluates the use of "convolutional neural networks" (CNN) to improve calculations of turbulent heat flux (THF) in Arctic leads from 1km resolution MODIS Terra infrared satellite imagery. It represents a novel application of machine learning to both MODIS imagery, which is widely used and invaluable in studies of polar regions, and THF in leads, which is an important parameter in our understanding of climate change in Arctic regions. Though the discussion section is lacking and editing is needed with regards to the English language, this is a exciting paper that should be published following revision.

Overall comments

- I agree with the first reviewer in that "Ice Surface Temperature (IST)" would be a more appropriate term than SST. "IST" is used in studies of polynyas using MODIS data, even when describing open water pixels. One of many literature examples is: https://www.mdpi.com/2072-4292/10/3/366. It is also the term used by NASA for their level-2 product.
- The discussion section would be suitable if this were submitted to a journal purely focused on machine learning techniques, but because this is The Cryosphere and readers will be interested in the DeepSTHF methodology in the context of polar science, this section should be expanded to include discussion of:
  - The significance of the improvement in calculated THF using the DeepSTHF vs other methods, eg. is the magnitude of improvement in W/m^2 significant relative to the overall heat budget in the Beaufort Sea or similar study area? Is it worth the additional computing resources to utilize DeepSTHF over CubicSTHF?
  - How applicable do the authors think DeepSTHF would be to other areas outside the Beaufort Sea in the Arctic, or even the Antarctic?Does the fact that DeepSTHF was unable to capture very narrow leads mean it may be less applicable in certain areas or times of year, when narrow leads are more common? A discussion of the nature of leads in the study area/broader Arctic would add some helpful context
  - DeepSTHF is promising, but exactly how far is it, or neural networks in general, from being widely implemented in studies of THF in leads? What are the next steps?

- Also agreed with the first reveiwer in that the paper would benefit from professional English editing

Line/Section/Figure Notes

Line 50: Would be helpful to list out specific (satellite-related?) examples instead of just citing the papers

Line 53: Is there a way to explain how CNN is used to produce super-resolution imagery to those of us without a extensive technical understanding of its mechanics or machine learning in general?

Line 60: It would be nice to explain these names so they are easier to remember- eg. DeepSTHF is based on deep learning and THF, but I am unsure what the S is for. Likewise, does the "Ori" in OriSTHF stand for "original" image?

Line 93: Why is SST calculated manually instead of using the Level 2 MODIS IST product with the same 1km resolution?

Line 93: Should also cite original work of Key et al. (1994, 1997), which Hall and Riggs (2001) followed for IST calculations

Line 95: Because clear sky is so important to the split-window algorithm, can you explain how you chose the 10% cloud cutoff?

Line 120: You introduce these formulae here but don't show them until lines 221-222. To avoid confusion, perhaps just describe the meterological variables collected and don't describe the formulae yet.

Line 167-168: Move to beginning of section

Line 169: Move "The following subsection explains the two CNNs fmore fully) to Line 162, after the sentence "Therefore, we used two CNNs... to achieve generation of a fine-resolution SST image and lead map."

Line 225: Though the Goosse paper does describe everything in detail, because the latent

and sensible flux calculations are so essential to this study, it would be good to write out more detail in this section, especially regarding the calculation of transfer coefficients, which can be parameterized many ways.

Line 342: DeepSTHF method achieved the most accurate THF, but there still appears to be a considerable amount of scatter in Figure 10c. Related to my overall comment regarding the discussion section, further discussion of the magnitude of the descrepency between reference and estimated THF and what the nature of this descrepency means for a potential broader application of DeepSTHF is important

Line 400: Why were these specific dates chosen?

Line 408: Please describe this additional layer of error caused by temporal correction

Section 3.3: I don't think the second paragraph (Lines 237-243) is necessary, as you describe these statistics in the results section. Instead, it would be more helpful for the reader if, after describing the three methods DeepSTHF, CubicSTHF, and OriSTHF in Lines 230-236, you explain that you will run two experiments: One with simulated MODIS imagery and one without, and why you do this. I was unaware at start that there are two separate experiments, which led to confusion while reading the long results section.

Section 4.2.2: Why are the THF algorithms compared to Landsat for the lead maps but not the overall THF calculation as they are in Section 4.1.2 with the simulated MODIS imagery? Quantifying how well DeepSTHF performs relative to the reference is important, especially when applied to real MODIS imagery.

Figure 6: The caption suggests it is a simulated super resolution image when it is instead a coarse resolution MODIS image simulated from Landsat imagery (I think). Please clarify this

Figure 14 - Why is THF calculated from the reference Landsat imagery not shown as well?