**Responses to Comments on the Manuscript:**

**"Estimating subpixel turbulent heat flux over leads from MODIS thermal infrared imagery with deep learning"**

**(ID: tc-2020-363)**

We sincerely thank the anonymous referee #1 for his/her detailed and useful comments and suggestions during the whole review process. We have carefully studied these comments, and made corrections or changes according to the comments and suggestions to improve our paper, and we are now resubmitting a revised manuscript which we hope will meet with your approval. The major revised portions are marked in green in the revised manuscript. The item-by-item responses to the reviewers' comments are listed as follows:

**The authors present a study to solve the mixed pixel problem in the remote sensing of ice surface temperature and ice leads by using convolutional neural network. Then the finer resolution data facilitate the further lead heat flux estimation at a more detailed level. The proposed deep learning-based method outperforms other methods mainly due to its capability of capturing complex nonlinear spatial pattern/relationship between images on different scales. Overall, the study provides a new prospects of lead mapping, but the manuscript in its current state does not meet the standard of the TC. I suggest major revision and the language needs further improvements.**

**General Comments:**

**Comment 1: Most of the study area cover the ice zones, having temperature lower than 2â„ƒ(Fig.2 and Fig.12). This might not be appropriate to use the term "sea surface temperature". I suggest to use "Ice surface temperature".**

**Response:** Thanks very much for your suggestion. Indeed, "ice surface temperature" is more appropriate for this paper, we have changed all "sea surface temperature" to "ice surface temperature" in the revised manuscript.

**Comment 2: This experiment was conducted on the Beaufort Sea. Would the model be suitable for other Arctic sea ice regions such as the central Arctic Ocean where the Landsat imagery is lacking? Although the reconstructed SR IST is hard to validated there, it is possible to assess the accuracy of leads map through other source of high resolution dataset such as SAR image.**

**Response:** Thanks very much for your valuable advice. According to your suggestion, we have tested the trained leads mapping network in the Barents Sea of the Arctic and used Sentinel-2 imagery to assess the accuracy. The visual performance and corresponding quantitative evaluation were shown in Fig.1 and Table 1, from which we can see that the model has a good generalization ability and performed well in the other Arctic sea ice region besides the Beaufort Sea. The detailed experiment result has been demonstrated in the discussion part (Lines 548-571).

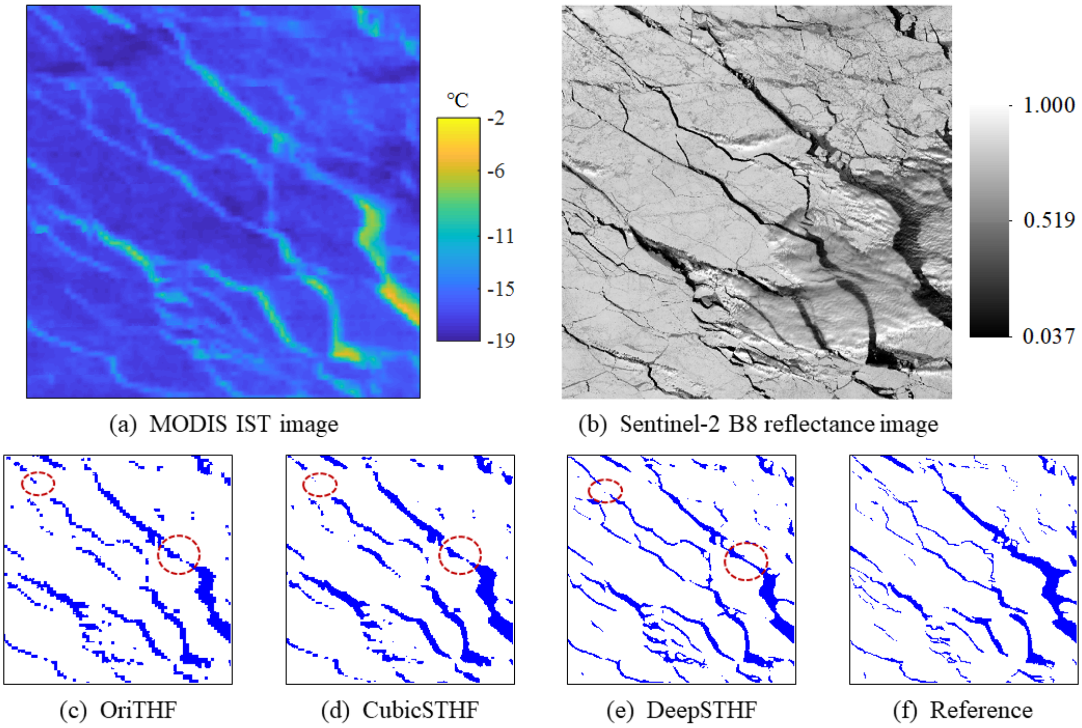


**Figure 1. (a) The MODIS IST image of a subarea in the Barents Sea; (a) the Sentinel-2 B8 reflectance image; (c) the lead map obtained from the MODIS IST image by the OriTHF method; (c) the lead map obtained from the MODIS IST image by the CubicSTHF method; (c) the lead map obtained from the MODIS IST image by the DeepSTHF method; (f) the reference lead map extracted from the Sentinel-2 image. The red ellipse in (e) represents the area impacted by the drifting snow.**

**Table 1. The lead mapping results of the OriTHF, CubicSTHF, and DeepSTHF methods.**

| Method | Overall accuracy | Commission error | Omission error | MIOU |
|---|---|---|---|---|
| OriTHF | 0.918 | 0.049 | 0.350 | 0.686 |
| CubicSTHF | 0.916 | 0.054 | 0.333 | 0.684 |
| DeepSTHF | **0.941** | **0.035** | **0.265** | **0.753** |

**Note: MIOU stands for the mean intersection over union. The most accurate results are highlighted in bold text.**

**Comment 3: In the introduction, has CNN-based SR method ever been used in downscaling thermal infrared images in other regions, for example, in middle latitude areas? I suggest adding some background about it.**

**Response:** Thanks for your question and suggestion. CNN model has been applied for MODIS and

AMSR2 sea surface temperature super resolution in the middle latitude sea areas (Ping et al., 2021). We are sorry for not mentioning this background in the original manuscript, and it has been added in the revised manuscript (Lines 58-59).

Reference:

Ping, B., Su, F., Han, X., and Meng, Y.: Applications of Deep Learning-Based Super-Resolution for Sea Surface Temperature Reconstruction, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 887-896, 10.1109/JSTARS.2020.3042242, 2021.

**Comment 4: The wind and air temperature are referred at different altitude, any measure on solving the inconsistence? on which height is the turbulent heat flux calculated? Also, the hourly air temperature from ERA5 reanalysis is provided on 0.25°grid (which is not mention in the manuscript). The scale of air temperature data doesn't match with those of MODIS or Landsat images, therefore potential influence of warm lead surface on the bottom air might be neglected. Uncertainty in this case should be noted.**

**Response:** Thanks for your questions and advices. In the experiment, we focused on the turbulent heat flux on 2 m height. Unfortunately, 2 m wind speed data is not available the European Center for Medium-Range Weather Forecasts ERA5 reanalysis dataset. Therefore, like previous study (Qu et al., 219), the logarithmic wind profile equation (Tennekes, 1973) was used to calculate wind speed at 2 m height based on 10 m wind speed . We are sorry for not providing related description about this in the original manuscript, and we have added detailed turbulent heat flux calculating process including this conversion in the revised manuscript.

Indeed, the air temperature from ER5 reanalysis is provided on 0.25°grid, which does not match those of MODIS or Landsat images, we apologize for not mentioning in the manuscript. Theoretically, the scale of the air temperature should be consistent with those of Landsat imagery in the experiment, however we cannot find such data at present. Alternatively, we downscaled the air temperature imagery using cubic convolution interpolation method. The air temperature above the warm lead surface (especially for those small leads) might be influenced, and it would bring about uncertainty in the experiment. However, constrained by the lack of fine spatial resolution air temperature data, existing approaches (Qu et al., 2019) can only use the air temperature from ER5 reanalysis datasets to calculated turbulent heat flux over leads. Additionally, our study mainly aims to improve the accuracy of THF estimation using CNN-based method (DeepSTHF) under present conditions, and the results showed the potential of DeepSTHF. Therefore, we also used the hourly air temperature from ERA5 reanalysis. To make readers have a comprehensive understanding of our work. The uncertainty of this has been mentioned in the revised manuscript (Lines 631-635).

Reference:

Qu, M., Pang, X., Zhao, X., Zhang, J., Ji, Q., and Fan, P.: Estimation of turbulent heat flux over leads using satellite thermal images, The Cryosphere, 13, 1565-1582, 10.5194/tc-13-1565-2019, 2019.

Tennekes, H.: The Logarithmic Wind Profile, Journal of Atmospheric Sciences, 30, 234-238, 10.1175/1520-0469(1973)030, 1973.

**Comment 5: Please rewrite the conclusion section, it looks to me that it is more like a discussion.**

**Response:** Thanks very much for your suggestion. Indeed, the conclusion part is not appropriate in the original manuscript. In the revised manuscript, it has been rewritten to "This paper proposes the DeepSTHF method for MODIS thermal infrared imagery. Specifically, the proposed DeepSTHF method includes two CNN models that are used to generate a finer spatial resolution IST image and the corresponding finer resolution lead map from the MODIS IST image. The finer spatial resolution data are used for THF estimation. The proposed DeepSTHF method is compared with a pixel-based method, the OriTHF, and a cubic interpolation-based method, the CubicSTHF, in two experiments using real and simulated data. The results showed that the proposed DeepSTHF acquired more accurate and reliable THF results than the other two methods, which was because it could detect more narrow leads and generate more accurate temperature in the leads area than the OriTHF and CubicSTHF methods. This study demonstrates the potential of deep learning in the field of THF estimation over leads, where the deep learning-based methods can represent a favorable tool for analyzing fine variations in leads and the corresponding impact on the climate in the Arctic region".

**Comment 6: I found some long sentences such as Line 60-62, Line 519-520, hard to understand. Suggest authors do professional English editing.**

**Response:** Thanks for kindly suggestion. In the revised manuscript, these long sentences have been modified. Furthermore, according to your suggestion, the revised manuscript has been polished by a professional English editing service. The certificate of English language editing is shown in Fig.2.

# Certificate of English Language Editing

## Manuscript Title:

Estimating subpixel turbulent heat flux over leads from MODIS thermal infrared imagery with deep learning

## Date of Revision

April 16, 2021

Abstract:

The turbulent heat flux (THF) over leads is an important parameter for climate change monitoring in the Arctic region. Currently, the THF over leads is often calculated from satellite images, but the accuracy of the estimated THF is low for images consisting of mixed pixels that include both ice and leads because the existence of mixed pixels along lead boundaries decreases the measuring accuracy of the surface temperature over leads and the corresponding lead map. To address this problem, this paper proposes a deep residual convolutional neural network (CNN)-based framework to estimate THF over leads at the subpixel scale (DeepSTHF) based on remotely sensed images. The proposed DeepSTHF provides an ice surface temperature (IST) image and the corresponding lead map with a finer spatial resolution than the two-CNN model so that the subpixel scale THF can be estimated from them. The proposed approach is verified using...

This document certifies that the manuscript listed above was copy edited for proper English language at LetPub. All of our language editors are native English speakers with long-term experience in editing scientific and technical manuscripts. We are committed to leveling the playing field for researchers whose native language is not English.

- Neither the research content nor the authors' intended meaning were altered in any way during the editing process.
- Documents receiving this certification should be considered ready for publication where language issues are concerned. *However, the authors may accept or reject LetPub's suggestions and changes at their own discretion.*
- If you have any questions or concerns about this edited document, please contact us at support@letpub.com

LetPub is an author service brand owned and operated by Accdon LLC. Headquartered in the Boston area, we are a full-spectrum author services company with a large team of US-based certified language and scientific editors, ISO 17001 accredited translators, and professional scientific illustrators and animators. We advocate ethical publication practices and are an official member of the Committee on Publication Ethics (COPE).

For more information about our company, services, and partnership programs, please visit www.letpub.com.



**Figure 2. The certificate of English language editing**

- - - - - - - - - - - - - - - - - -

**Specific Comments:**

**Comment 1: Figure 1: The light color rather than black area represents leads.**

**Response:** Thanks very much for your kindly reminding. It has been modified in the revised manuscript.

**Comment 2: Line 92~94: could the authors elaborate on their decision to not use the NSIDC MOD29 sea-ice surface temperature product directly but instead calculate it themselves?**

**Response:** Thanks very much for the question. In the NSIDC MOD29 sea-ice surface temperature product, pixels labeled as cloud according to a cloud mask from MOD35 are removed. However, from visual inspection (especially by comparing with the corresponding Landsat imagery), some lead areas with ocean fog or plume (Qu et al., 2019; Fett et al., 1997) are mistakenly marked as cloud in MOD35 product, which would influence the experiment. Therefore, to preserve potential leads, we calculated it from MOD021KM product. Additionally, the cloud was determined by using MOD35 and visual inspection. We have provided the above-mentioned reason for not using the NSIDC MOD29 sea-ice surface temperature product directly in the revised manuscript (Lines 86-90).

Reference:

Fett, R. W., Englebretson, R. E., and Burk, S. D.: Techniques for analyzing lead condition in visible, infrared and microwave satellite imagery, Journal of Geophysical Research: Atmospheres, 102, 13657-13671, 10.1029/97JD00340, 1997.

Qu, M., Pang, X., Zhao, X., Zhang, J., Ji, Q., and Fan, P.: Estimation of turbulent heat flux over leads using satellite thermal images, The Cryosphere, 13, 1565-1582, 10.5194/tc-13-1565-2019, 2019.

**Comment 3: The document "Hall and Riggs, 2001" is not cited properly. "Algorithm Theoretical Basis Document (ATBD) for the MODIS Snow and Sea Ice-Mapping Algorithms" has three main contributors and another seven co-authors, thus you should cite this paper as following: Hall, D.K.; Riggs, G.A.; Salomonson, V.V.; Barton, J.; Casey, K.; Chien, J.; DiGirolamo, N.; Klein, A.; Powell, H.; Tait, A. Algorithm Theoretical Basis Document (ATBD) for the MODIS Snow and Sea Ice-Mapping Algorithms; NASA GSFC: Greenbelt, MD, USA, 2001.**

**Or Hall D.K., Riggs G.A. and Salomonson V.V., 2001. Algorithm Theoretical Basis Document (ATBD) for the MODIS Snow and Sea Ice-Mapping Algorithms. NASA's Goddard Space Flight Center, Greenbelt, MD.,1-45.**

**Response:** We are sorry for not properly citing this document and thanks very much for your kindly reminding. It has been modified in the revised manuscript. Additionally, other reference documents have been carefully checked as well.

**Comment 4: Line 105: As for choosing the retrieval algorithm for sea ice, I recommend to cite a related publication:**

**Fan, P., Pang, X., Zhao, X., Shokr, M., Lei, R., Qu, M., Ji Q, Ding, M. (2020). Sea ice surface temperature retrieval from Landsat 8/TIRS: Evaluation of five methods against in situ temperature records and MODIS IST in Arctic region. Remote Sensing of Environment, 248(January), 111975. https://doi.org/10.1016/j.rse.2020.111975.**

**Response:** Thanks very much for your suggestion. We have added this related publication in the revised manuscript.

**Comment 5: Line112: "manually drawn", what is the criteria in producing the reference lead maps? Is there any physical threshold used here, like that in Lindsay et al. (1995)?**

**Response:** Thanks very much for raising the question. In the experiment, the recommended iterative threshold method used in Qu et al. (2019) was applied to produce lead maps from IST imagery. Additionally, some outliers had been eliminated through visual inspection with visible spectral bands spectral bands. We are sorry for not introducing the detailed process in the original manuscript and it has been added in the revised manuscript (Lines 119-121).

Reference:

Qu, M., Pang, X., Zhao, X., Zhang, J., Ji, Q., and Fan, P.: Estimation of turbulent heat flux over leads using satellite thermal images, The Cryosphere, 13, 1565-1582, 10.5194/tc-13-1565-2019, 2019.

**Comment 6: Line115: why don't you use the 100 m raw data instead of relying on the up-sampled 30 m product?**

**Response:** Thanks for the question. Though the raw data acquired from Landsat-8 thermal infrared sensor (TIRS) is at a spatial resolution of 100 m, it was officially resampled to 30 m to match the data from the OLI spectral bands. Unfortunately, the 100 m raw data is not provided. Therefore, we have to use the up-sampled 30 m product instead of the 100 m raw data in the experiment.

**Comment 7: Line 174: Is the layer number of very deep residual CNN model a prescriptive constant, or we can adjust them?**

**Response:** Thanks for raising this question. The layer number of very deep residual CNN model can be adjusted according to different tasks. Generally, CNN models with more layers would have a better performance. In real applications, however, the layer number of CNN models is determined by many factors, such as the complexity of the problem, the number of training samples. At present, the feasible way is to set the number of layers via a series of experiments. In our study, we set the layer number through lots of experiments as well, and the final results demonstrated that the used CNNs with the set layer numbers were able to achieve accurate estimation of THF over leads.

**Comment 8: Line 211~213: Does the lead map show consistency with the assumption that the surface temperature all above the freezing point?**

**Response:** Thanks for the question. For the super resolution lead mapping CNN, each pixel value of the output is the probability that it is lead, thereby the threshold 0.5 in Lines 211-213 of the original manuscript represents this probability. In practice, the threshold can be set in 0~1, and 0.5 (which was usually used in the mapping task with CNN models) is empirically set in this study. In the experiment, we found that the temperature of lead area (including the reference lead map produce by the iterative threshold method (Qu et al., 2019) and segmented lead maps by the three methods) was not always above the freezing point. The major reason for this may be that the open water of leads usually comprises several pieces of ice (especially along the boundaries of leads), which would lower the surface temperature of leads to some extent. Additionally, seawater is a complex mixture of water, salts, and smaller amounts of other substances, salts and other substances would influence the surface temperature of leads as well.

Reference:

Qu, M., Pang, X., Zhao, X., Zhang, J., Ji, Q., and Fan, P.: Estimation of turbulent heat flux over leads using satellite thermal images, The Cryosphere, 13, 1565-1582, 10.5194/tc-13-1565-2019, 2019.

**Comment 9: Line 239: The Pearson coefficient is generally represented as "r" instead of "R".**

**Response:** Thanks very much for kindly reminding. It has been modified in the revised manuscript. Note that, according to the suggestion of referee #2, description on quantitative evaluation indices (Lines 237-243 in the original manuscript) has been deleted in the revised manuscript, the changed representation of Pearson coefficient was in the results section.

**Comment 10: 6: Suggest to mark the name of corresponding methods in the sub-images or subtitles.**

**Response:** Thanks very much for kindly suggestion. The name of corresponding methods has been marked in corresponding subtitles.

**Comment 11: Line 284~286: the subsentence after "because" is not the cause, please rephrase this sentence.**

**Response:** Thanks very much for the kindly suggestion. We have modified the sentence as "Additionally, the results in Fig. 7b show that the CubicSTHF method underestimated most pixels with a reference temperature higher than −6 °C, which is indicated by the substantial number of data points below the diagonal line" in the revised manuscript.

**Comment 12: 7: In this figure the labels of both X and Y-axis are not appropriate. As the scatter plots represents the IST between Landsat and SR images, the X and Y-label can be "Reference IST" and "IST from xxx".**

**Response:** Thanks very much for your suggestion. They have been modified in the revised manuscript.

**Comment 13: Line309: what about the surface temperature distribution of the lead in the map? Are they all above the freezing point? Same question for the Fig 13.**

**Response:** Thanks for raising the question. No, they are not all above the freezing point. The major reason for this is that the open water of leads usually comprises several pieces of ice (especially along the boundaries of leads), which would lower the surface temperature of leads to some extent. Additionally, seawater is a complex mixture of water, salts, and smaller amounts of other substances, salts and other substances would influence the surface temperature of leads as well.

**Comment 14: Line 397: What is "at a step size of 40"? please clarify.**

**Response:** We are sorry for this confusion. Generally, in producing of training data, each two neighboring image subsets are partially overlapped to increase the number of training samples. In the experiment, the overlap size was empirically set to 40 pixels. A step size means the sliding size in the clipping procedure. To make it more clearly, we have modified the sentence in the revised manuscript (Line 279 and Line 415).

**Comment 15: 12: Note that Landsat images acquired on 25 April 2018 is partly contaminated by cloud. You also mentioned the red dashed ellipse in Fig. 13c.Could you discuss the impact of the cloud on your method?**

**Response:** Thanks very much for the suggestion. Yes, the Landsat imagery acquired on 25 April 2018 is partly contaminated by cloud. If the satellite imagery was covered by cloud, the surface temperature of the contaminated region does not represent the real case and will therefore influence the result. Note that, the impact of cloud on DeepSTHF may be different in the training and testing stages. The specific impacts have been discussed in the revised manuscript (Lines 626-631).

**Comment 16: 13: It seems the sub-plots m to x do not maintain the same sizes with the black rectangle r1 to r3. Please make sure they have same size and do not stretch them.**

**Response:** Thanks for the advice. They have been modified to maintain the same sizes without stretching in the revised manuscript.