Answer to Referee 1 (J. Inoue)

We would like to thank the referee for his constructive criticism which helped to improve the manuscript.

1. My concern is that the analysis period is different between Alert/Barrow stations (4 or 6 years) and Tara station (1 year). Although I understand that Tara provided invaluable data, the statistics (variances, biases, RMSE, and frequencies) shown in Figs. 5-10 might be difficult to compare with those from Alert/Barrow stations. Please discuss and mention the limitation in Tara data set.

In the Data section we now introduced Barrow and Alert as the main considered stations supplemented by one month of Tara data which is valuable because in situ measurements from the Central Arctic are sparse. We added the following sentence (Sect. 2, par. 1):

"As the thermal differences between sea ice and open water surfaces are small in summer, only one month (April 2007) of Tara data was used in the analysis. Despite the short timeseries, Tara provides invaluable data since measurements from the Central Arctic are sparse."

We also pointed out in Sect. 4.5 (Radius of impact) and in the Conclusions section that results using Tara data are less significant than those from the other two stations. We write now:

Sect. 4.5:
"Results for Tara with respect to R values are ambiguous. The curves for bias and RMSE using the IST method indicate a radius of impact of about 5 hr, while the results using the AT method hardly change with decreasing trajectory length. However, only one month of data is used for Tara which might not be long enough to draw reliable conclusions concerning the comparison of results for Tara and both other stations."

Conclusions:
"In general, all results for Tara are less relevant compared with those for Alert and Barrow since only one month of data is used for the calculations."

2. P3013 Line 11: It should be noted that ice concentration during summer is worse than winter due to melt ponds (Inoue et al. 2008).

The citation has been added.

3. P3015 line 10: Please specify the height of surface wind in JRA-25. It should be the same height as ERA-Interim 10-m wind; otherwise it is difficult to compare the results. This links to the conclusion part as 'problems of ERA with the wind field in the Central Arctic.' (P3031 line15)

Both wind fields are from 10 m height. The sentence has been rewritten for clarification and reads now:
"Backward-trajectories arriving at the stations are calculated from the 10 m-wind fields of the Japanese 25-year reanalysis (JRA) and of the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA-Interim)."

4. P3022 line 22: Ice concentration in JRA-25 is only used for judgment to allocate an ice-covered or ice-free grid (i.e., 100 % or 0%) by using cut-off threshold of 55% ice concentration. Therefore, leads and polynyas do not exist in ice-covered grid in JRA-25. The authors cited Inoue et al. (2011) and should have already known these problems.

In this context we refer to the four ice concentration data sets obtained from remote sensing, not to the ice concentration of the reanalyses. A rephrasing hopefully cleared this misunderstanding.

"The frequency distributions of ice concentrations obtained from remote sensing data for the trajectories resulting from ERA or JRA wind fields are very similar."

5. P3023 line 17: Please mention the value of correlation coefficient quantitatively.

Done. The sentence reads now:

"The correlation between the observed 2-m air temperatures and the mean ice surface temperatures along the trajectories (IST) and modeled temperatures (AT) is positive, exceeding values of 0.6, and significant at the 95 % level for all combinations of reanalyses, ice concentration data sets and BL 350 depths."

6. P3026 line 19: is largest -> is the largest

Done

7. P3028 line 2: I can not understand why JRA-25 reproduced the surface wind direction relatively well.

Thanks for this comment. We checked this once more and found indeed a programming error in ERA trajectories which influenced especially Tara results. As a consequence, the mean separations between ERA and JRA trajectories decreased (Sect. 4.1) and correlations for Tara are now even slightly higher using ERA trajectories than using those calculated from JRA wind fields. Therefore, we deleted the sentence above. The changed results for Tara are described in Sect. 4.4:

"Tara shows the largest sensitivity to different reanalyses for the AT method. Explained variances are about 70 % using JRA trajectories with RMSE of about 3.6 C (Fig. 8). Using ERA trajectories gives larger explained variances in the order of 80 % with RMSE of 3.3 C (100 m) to 3.9 C (350 m)."

Abstract:
"Trajectories based on 10-m wind fields from both reanalyses show large spatial differences in the Central Arctic, which leads to differences in the correlations between modeled and observed 2-m air temperatures. They are most pronounced at Tara where explained variances amount to 70 % using JRA and 80 % using ERA."

Conclusions:
"All methods give slightly better results using ERA trajectories than using those derived from the JRA reanalysis."