Replies to the comments of Referee 1

We are grateful to the referee for the constructive criticism, which helped to improve the clarity of the manuscript. Please find below the replies to the comments and an account of the modifications implemented.

General comments

1. *The objectives as formulated in the introduction are not followed by a corresponding structure and sequence in the methods and results section. This makes the overall manuscript very hard to follow as a reader, as one needs to search for the corresponding information.*

We apologize for this inconvenience. We have now organized the structure of the manuscript (in Result and Discussion sections) following the objectives in the introduction (Lines 59-64, p. 3). First, we calculate climatic indices to show the possibility of vegetation shift; second, we study dynamics of fires; third, we study dynamics of vegetation and its links to fires and topography.
"The main objectives of the study are:
1) to study dynamics of regional climatic factors in order to assess the possibility of vegetation shifts due to climate change, in particular tundra-forest transition;
2) to quantify burned surface areas and calculate frequency of wildfires;
3) to study the link between wildfires and dry tundra transition to woodlands and forests.
Finally, we take into account physiographic characteristics of the landscape and study the effect of the topographic slope on the transition."

The structure of the manuscript is now organized as follows:
1 Introduction
2 Materials and methods
 2.1 Field sites
  2.1.1 General description
  2.1.2 In-situ observations of vegetation and permafrost state
 2.2 Calculation of climatic indices
 2.3 Wildfires
 2.4 Vegetation dynamics
 2.5 Topographic slopes
3 Results
 3.1 Temperature, precipitation and climatic indices
 3.2 Dynamics of fires
 3.3 Vegetation dynamics and its links to fires and topography
  3.3.1 Estimates of recovery time after fire using NDVI
  3.3.2 Vegetation shift using visual method, its connection to fires and topography
4 Discussion
5 Conclusion

2. *For example, there is not a dedicated methods section that explains how the first objective (to quantify burned surface areas and assess frequency and causes of wildfires) was addressed and the sequence is changing between methods and results. Apart from structural problems, some of the objectives are not directly followed at all or in a qualitative way only*

We have described how we quantified burned areas (please find in the answer to the next question) and frequency of wildfires in Methods, section 2.3 Wildfires (Lines 137-156).

'We studied the percentage and distribution of the burned sites and calculated the frequency of fire return. Corona and Landsat images showed that some years were characterized by particularly large-scale fires in the study areas (see an example for 1990 in SM, Fig. SB1). These years are referred to as the years of major fires. The burned areas can be detected in the satellite images during a few years after the fire. Landsat mosaics from 1988, 2001, 2016 and 2018 largely reflect the state of the study areas after the major fire years 1976, 1990, 2012 and 2016 (Table 5). Burned areas in Corona mosaic from 1968 were partially dated back to the fires in the period 1953-1964 based on geological surveys and early Corona images (the sources are listed below Table 5). The period between fires was calculated as the difference in years between the major fires.'

In order to avoid qualitative results, we decided not to address causes of wildfires in the current manuscript and leave this topic for a more careful quantitative study in future. We have removed all the information pertaining to possible causes of fires from the manuscript. We have also removed section 'Qualitative observations of the vegetation dynamics'.

3. *Methods: The manuscript contains tables with data sets, but it remains unclear which data sets were used for which objectives/results specifically, how the imagery was preprocessed given so many different data sets of highly varying spectral, spatial resolution and quality were used. Also, details on the processing of data (esp. remote sensing data, e.g. atmospheric correction) are largely missing (indicating a software without even the version or parameters used for the algorithm is not sufficient for reproducible methods). Further, there is little to no information about validation of the classification results or reference to uncertainties of results.*

We have added the information about data sets and preprocessing of the data in Methods (sections 2.3 Wildfires and 2.4 Vegetation dynamics).

Wildfires

Lines 126-129: 'The initial state of the study areas was obtained from 'Corona' images. We identified 21 frames under clear-sky conditions from 21 August 1968. Each frame consisted of 4 scanned fragments. The cropped fragments without color correction were georeferenced to the chosen orthomosaic (SPOT layer, see section 2.4) using polynomial method (3rd order polynomial) in software ArcGIS (v. 10.4.1).'

Lines 132-136: 'Further, we used Landsat data to study dynamics of burned areas. The data providing the best coverage of the study areas were available from the following years: 1988, 2001, 2016 and 2018 (Table 4). The images were synthesized using near- and mid-infrared channels (Landsat 5 and 7: 1.55-1.75μm, 0.76-0.90μm and 0.63-0.69μm; Landsat 8: 1.57-1.65μm, 0.85-0.88μm, 0.64-0.67μm) as burned areas are visible in the infrared range of wavelengths. Landsat mosaics for all years were formed after application of color correction using mosaic operator Blend in ArcGIS.'

Lines 137-149: 'Mapping and quantification of the burned areas were performed by means of an object-based image analysis, successfully used for studies of landscape dynamics (Blaschke, 2010). On the first stage, we performed segmentation of mosaics using algorithm 'Multiresolution segmentation' in software eCognition (v. 9.0). The segmentation was done using parameters 40 for Scale and 0.5 for Color. The second stage, classification, was different for Corona and Landsat mosaics. For Landsat mosaics, we used unsupervised classification ISODATA (15 classes, a change threshold 5%). Further, we identified visually one or two classes corresponding to burned areas.

The segments containing more than 90%of pixels within these classes were identified as burned areas. In addition, we visually checked the segments with lower percentage of pixels (down to 40-50%) belonging to these classes and they were manually added to burned areas when necessary. In Corona mosaics, the spectral information was absent and we had to rely on the contrast of colors between background tundra and burned areas. In this visual check, we used two criteria. First, background tundra is lighter due to the presence of lichen in vegetation community, whereas recently burned areas are dark. Second, burned areas are characterized by well-defined boundaries often coinciding with river coastlines. An example illustrating segmentation and the visual choice of burned areas is shown in Fig 3. Calculation of the areas of segments classified as burned areas were performed using standard instrument Calculate geometry in ArcGIS.'

Vegetation dynamics

Lines 162-164: 'The initial state of vegetation was obtained from Corona imagery and topographic maps. The compilation of mosaic using Corona images is described in Section 2.3. The topographic maps were used mainly to develop the forest mask using automatic tracing in software EasyTrace, v. 8.7. The resulting vector layer was checked and corrected using Corona mosaic.'

Lines 165-169: Assessment of the current state of vegetation was based on Resurs-P and SPOT data (Table 6). Majority of the territory was covered by the mosaic of SPOT-6,7 imagery synthesized in the visible range without color correction. Ca 10% of the territory were covered by three paths of Resurs-P (the product level 2A, including four channels B, G, R and NIR). The SPOT mosaic was used as a pluggable webmap layer without additional processing. We co-registered Resurs-P data to the SPOT mosaic in ArcGIS.

Classification was performed using the visual method described in Section 2.4 'Vegetation dynamics'. Overall, visual methods are not rare in scientific studies. For example, classification of clouds performed by observer is typically taken as an etalon when automatic methods are developed. It is also not the first time when the visual methods are used for quantification of the vegetation shift. In the pioneering study, Frost and Epstein (Global Change Biology (2014) 20, 1264–1277) used a similar visual method for the same purpose. They state that 'Gambit and Corona are well suited for land-cover change studies in tundra ecotones because tall shrubs and trees form abrupt transitions in vegetation structure that create unambiguous, readily interpreted photo-signatures. These photo-signatures result from the shadowing projected by the canopies of tall shrubs and trees, which greatly overtop tundra vegetation and create areas of high contrast in panchromatic imagery.'

We have added a figure illustrating different decisions about the vegetation shift (Fig. 4) and added reference to Frost and Epstein (2014) study in the methods (Lines 160-161).

4. *Overall the presentation of the manuscript is really not sufficient – as indicated in more detail below, graphs are poor (missing legends (esp. Fig 1 & 2), scale, missing reference of Figure in main text). Consider a more rigorous selection of graphs and information displayed in tables. Also, thorough revision of language (esp. articles) and checking of consistency are needed to make this manuscript more accessible.*

We would like to thank reviewer for valuable comments regarding the figures and tables of the manuscript. Here is the list of modifications/adding done in the figures (numbers of figures are from the previous version of the manuscript):

Fig. 1: we added the legend. We added boundaries 'southern tundra - forest-tundra –northern taiga' in the figure.

Fig. 2: we removed the figure.

Fig. 4a: we added linear trends of temperature.

Fig. 5b: we marked years of major fires by dashed lines to emphasize connection between fires and evapotranspiration.
Fig. 6 was erroneously referenced as Fig. 5 in the previous version of the manuscript. In the present version, we have removed the section where this figure should have been referenced, as well as the figure.
Fig. 7a: increased fonts, added 'background' in the legend.
Fig. 8a: added boundaries 'southern tundra - forest-tundra –northern taiga' in the figure.
Fig. 8b is replaced by the figure with NDVI distributions.
Fig. 11: only the figures with the mean slopes are retained.

From the tables, we have removed the information related to causes of fires and classification of vegetation zones based on recent 'Atlas…', 2004 (Tables 1 and 4 in the current version, Tables 2 and 3 in the previous version).

We have made revision of language throughout the text and reorganized paragraphs in several sections to make them more consistent (e.g., In-situ observations of vegetation and permafrost state, Discussion). The situation with articles will be further improved if the manuscript is accepted, because all EGU journals including BG support English correction before the manuscript is published.

Specific comments

1. *Temp., precip, climatic indices - Methods for GDD5 (line 144) – provide reference for this formula.*

We added the reference: 'We calculated growing degree-days following Tchebakova et al. (1994)' (Line 115).

Line 158 – what is the 3◦ increase based on – a trend fitted to the climate data in Fig. 4? If yes, show the trend and related statistical information.

Yes, and we added the trends and the corresponding statistical information in Fig. 4a (Fig. 5a in the current version).

- Data from 3 meteorological stations are presented. It remains unclear how these are linked to the 3 selected study sites as the stations are located outside of the study sites and not in obvious pairing to the 3 sites.

For climate, latitudes play an important role. These three stations cover the whole range of latitudes of our study areas. The stations are located reasonably close to the study areas. One of them (Nyda) is located within study area 1, but its latitude is also close to that of the northern part of study area 2. Two stations (Novy Port and Nadym) are indeed located outside the study areas. Station Nadym is located near the southern borders of two study areas, between areas 1 and 3 (50 km from the study area 3, 150 km from the study area 1). It should be representative of the climatic conditions in the southern parts of study areas. Station Novy Port is located 70 km to the north of study area 1 and this is the closest station to the northern boundaries of the study areas. The northern part of study area 1 is located between stations Nyda and Novy Port.

Growing degree-days, an index based on air temperature (2 m height), should be largely the same along the latitude. Precipitation can significantly vary from station to station, which is also seen in

our Fig. 6, specifically on the example of Novy Port. However, the dryness index is not a limiting parameter for the vegetation shift at any of the stations and we do not expect this parameter to prevent vegetation change anywhere within our study areas.

We added the information about location of stations in the manuscript (Lines 105-108): 'Station Nadym **(**65°32'N, 72°32'E, 7 m a.s.l.**)** is located near the southern borders of the study areas (50 km from the study area 3). Station Nyda (66°37'N, 72°57'E, 10 m a.s.l.) is located within the study area 1. Station Novy Port (67°41'N, 72°52'E, 12 m a.s.l.) is located to the north of the study areas and this is the closest station to the northern boundaries of the study areas.'

> *For example on line 180 it is stated that based on the climate data analysis, the vegetation class in Novy Port changed from forest-tundra to dark needled northern taiga - how are the climate data linked to vegetation classes, and the station data to the study sites?*

The vegetation classes obtained from the topographic maps do not follow the classification based on the climatic indices. For example, in Nyda latitudes the maps indicate southern tundra, while the climate-based classification suggests dark needed northern taiga. This is likely the consequence of the insufficient precision of vegetation classification based on climatic indices, which can be too rough in the transitional zones. However, climatic indices illustrate that from the point of view of climate, conditions over all the study areas are suitable for forests.

2. *Qualitative assessment of vegetation dynamics. Overall this section is not convincing as it is largely missing a corresponding reproducible methods section. Quantitative results are hard to reach based on Corona as reference data set. But even if only qualitatively assessed, methods need to be clearly outlined.*

We have removed this section.

> *- How were these transitions qualitatively assessed? Some information can be found in section 2.1, some in the field sites general description, but nowhere is clearly formulated how the transitions were visually/qualitatively assessed, what classes were followed.*

We added Fig. 4 in the current manuscript to illustrate how vegetation change was assessed.

> *Also, it remains unclear how the topographic map was used for this (does it contain forested area? Burned area?).*

It contains forested areas, and we used it to create our forest mask. We added the following information in the Section 2.4 Vegetation dynamics (Lines 163-164):
'The topographic maps were used mainly to develop the forest mask using automatic tracing in software EasyTrace, v. 8.7. The resulting vector layer was checked and corrected using Corona mosaic.'

> *The graphs that are mentioned to highlight how this was done are not conclusive (e.g. Fig2 misses a color legend, also it is not clear from this graph which of the layers show the most reliable forest cover.*

> *Fig SB3 – without clear indication in the imagery it is hard to understand where the active afforestation mentioned in the figure title is located – this is certainly due to the very different quality of the Corona versus Yandex map layers, but as presented does not*

convince the reader that this active afforestation has happened). Also, how many sites (burned and background sites) were assessed in total? Are the different conditions statistically balanced (for several of the assessed transitions only a single reference site is mentioned, does it mean that this condition was only observed once)? What was the exact sampling design? - What is active afforestation? Define in the related methods section - L. 191- fig 5 is wrongly referenced (Fig 5 displays potential evaporation, not tundra after wildfire) - L 204 – removal of vegetation cover – define in related methods section.

Certainly, the sites were not statistically balanced as some conditions are quite specific (e.g., river valleys) but the conclusion was never made based on a single reference site. We have removed this section and this part of Supplementary material.

3. *Dynamics of fires - Methods: classification to identify burned areas: how was the initial state determined in the Corona images?*

Added in section 2.3 Wildfires (Lines 144-149), see also answer to Q3 of General comments.

In table 2 you also list Sentinel, Modis and VIIRS data – how were these data used (you only mention Corona and Landsat in the fire methods section)

We used these data to study qualitatively hot spots of the fires, but we removed this information from the current version of the manuscript.

4. *Dynamics of vegetation and fires - NDVI is NOT the normalized digital vegetation index*

We apologize for this error, which was unfortunately copied several times. We changed 'digital' to 'difference' throughout the text.

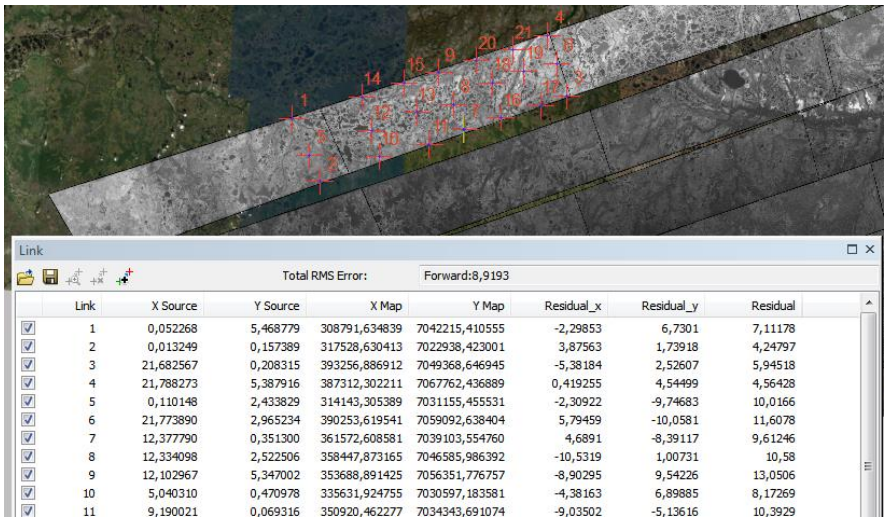
- Which imagery was used to calculate NDVI? Any preprocessing performed? Georegistration issues discovered? Explanations on remote sensing data in methods are insufficient. –

We added this information in Section 2.4, Vegetation dynamics (Lines 188-190):
'NDVI was calculated based on two scenes from one path of Landsat-8 from 30 June 2018 (path/row 160/013 and 160/014).We used Level 2 data (CEOS) after atmospheric correction by the standard Landsat 8 OLI Atmospheric correction algorithm (Vermote et al., 2016). NDVI was calculated in ArcGIS using standard tools'.

For ArcticDEM and satellite data sets except Corona, georegistration was performed by the data owners during orthotransformation of the satellite images. For topographic maps, georegistration was performed using coordinates of the check points in the field. The rms errors of georegistration for different types of data are reported below:
Resurs-P 1,2 – 5-10 m,
SPOT-6,7 – 5-8 m,
ArcticDEM – 5-7 m,
topographic maps – 20-25 m,
Landsat – below 15 m,
Corona/KH-4B – 10-15 m.

Corona images were georegistered to the mosaic of SPOT images in ArcGIS using polynomial method (3rd order). According to the data provider (CNES, National Centre for Space Studies,

France), for SPOT images, the rms of georegistration is 5-8 m and this was confirmed in different studies (e.g. Parage et al. New sensors benchmark report on SPOT 7, 2014).
For each of 88 frames of 'Corona', there is a number of check points (see figure below) for which we estimated rms errors along the latitude (NS), along the longitude (EW) and 2D rms error. For all the frames, the rms error was below 10-15 m.



Link	X Source	Y Source	X Map	Y Map	Residual_x	Residual_y	Residual
1	0,052268	5,468779	308791,634839	7042215,410555	-2,29853	6,7301	7,11178
2	0,013249	0,157389	317528,630413	7022938,423001	3,87563	1,73918	4,24797
3	21,682567	0,208315	393256,886912	7049368,646945	-5,38184	2,52607	5,94518
4	21,788273	5,387916	387312,302211	7067762,436889	0,419255	4,54499	4,56428
5	0,110148	2,433829	314143,305389	7031155,455531	-2,30922	-9,74683	10,0166
6	21,773890	2,965234	390253,619541	7059092,638404	5,79459	-10,0581	11,6078
7	12,377790	0,351300	361572,608581	7039103,554760	4,6891	-8,39117	9,61246
8	12,334098	2,522506	358447,873165	7046585,986392	-10,5319	1,00731	10,58
9	12,102967	5,347002	353688,891425	7056351,776757	-8,90295	9,54226	13,0506
10	5,040310	0,470978	335631,924755	7030597,183581	-4,38163	6,89885	8,17269
11	9,190021	0,069316	350920,462277	7034343,691074	-9,03502	-5,13616	10,3929

Fig. 1. Georegistration of Corona frames to the SPOT mosaic

Fig 8 – what is displayed here exactly? This remains unclear based on the corresponding methods section and figure title. Is the standard deviation based on spatial variation for the background sites? How is temporal variation in NDVI of the background sites accounted for? Are the different years and the background areas statistically balanced for their size?

In the previous version of the manuscript, Fig. 8a showed the distribution of NDVI over study area 1 based on the Landsat data from 3 July 2019. The curves of different colors indicated boundaries of the background tundra sites and burned tundra sites detected in the Landsat mosaics from different years (legend). Fig. 8b showed mean NDVI indices and standard deviations calculated for the background areas and the areas burned in different years as indicated in panel (a). The standard deviation was based on the spatial variation of NDVI (determined by the number of pixels with different NDVI within each area) and temporal dynamics was not accounted for. Instead of mean values and standard deviations, in the current version of the manuscript we use distributions of NDVI. We also show NDVI based on the Landsat data from 30 June 2018 instead of data from 3 July 2019 (the reason is described below). Therefore, Fig. 8 has changed (see Fig. 2 in this document).

In order to make the size of areas more balanced, we merged the data sets for 1968 and 1988, for which we might expect that vegetation recovered after fires, based on the mean NDVI values. Currently, all study areas are larger than 1000 km^2 (1968+1988 – 1265 (807+458) km^2, 2001 – 2320 km^2, 2018 – 1331 km^2, background – 1109 km^2).
Fig. 2b (Fig. 8b in the new version of the manuscript) shows the distributions of NDVI for the burned areas detected in 1968+1988, 2001, 2018, and for background areas. Using the dates of major fires, we can assume that 1968+1988 data show the state of vegetation in the site burned more than 42 years ago, 2001 data – 28 years ago, 2018 data – 2 years ago and background data refer to tundra not affected by fires during the whole study period.

The distributions are close to Gaussian ones for the background tundra and recently burned sites. Interestingly, the distributions are bimodal in the sites burned 28 and >42 years ago and they moved to higher NDVI values as compared to the background site. We fitted the distributions by the sums of two Gaussian functions and determined mean values and standard deviations for all the peaks (Table 1 in this document, Table 7 in the new version of manuscript). We found that bimodal distributions had almost the same two peaks. However, in the distribution from the sites burned 28 years ago, the peak with lower NDVI was more pronounced as compared to the peak with higher NDVI. Oppositely, in the sites burned >42 years ago, the peak with higher NDVI became more pronounced.

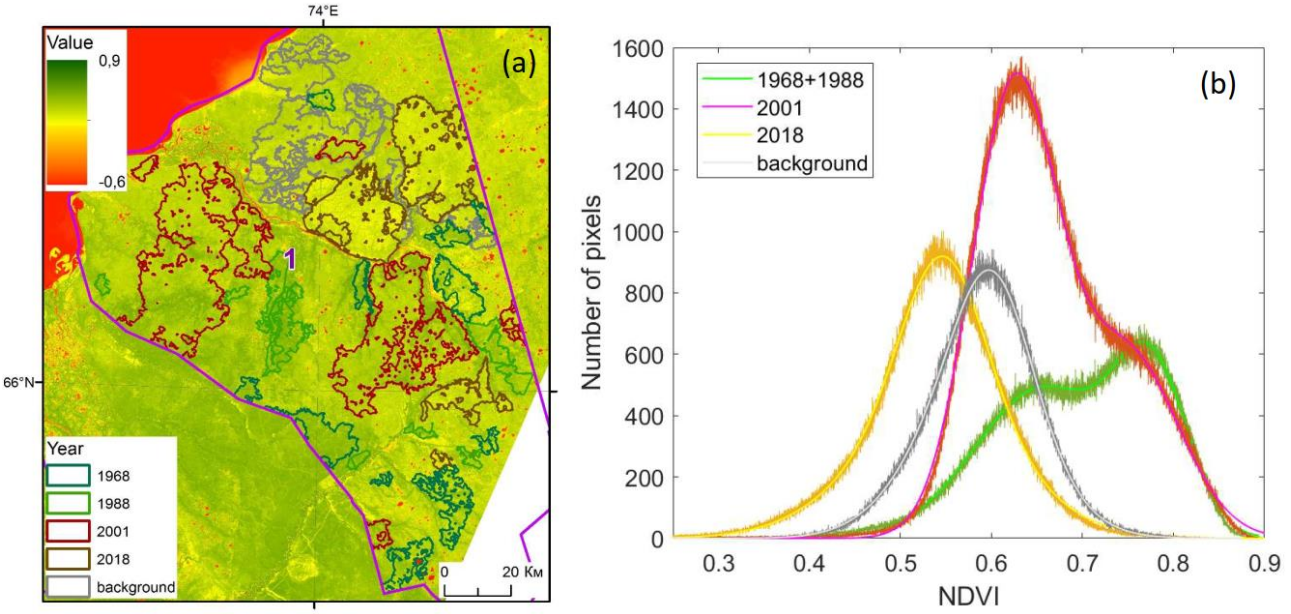


Fig. 2. (a) The distribution of NDVI over study area 1 in 2018. The curves of different colors indicate boundaries of the background tundra sites and burned tundra sites detected in the Corona and Landsat mosaics from different years (see legend). Burned areas in the mosaics from 1968 and 1988 are mainly due to fires from >42 years ago, in 2001 – due to fires from 28 years ago, in 2018 – due to fires from 2 years ago. (b) Distributions of NDVI based on the data from background sites and the sites burned in different years.

Table 1. Parameters of fits of the NDVI distributions in Fig. 2b by Gaussian functions

$$N_{pix} = A_1 \exp\left(-\frac{(NDVI-NDVI_{max,1})^2}{2\sigma_1^2}\right) + A_2 \exp\left(-\frac{(NDVI-NDVI_{max,2})^2}{2\sigma_2^2}\right).$$

Year	A_1	$NDVI_{max1}$	σ_1	A_2	σ_{max2}	$stdv_2$
1968+1988	489	0.6557	0.0741	496	0.7762	0.0393
2001	1459	0.6251	0.0472	598	0.7493	0.0566
2018	366	0.5385	0.0890	554	0.5471	0.0463
Background	853	0.5934	0.056	-	-	-

Further, we used the mean values and standard deviations to identify vegetation associated with the peaks of the distributions in the satellite images. For illustration, we chose an image containing all representative examples of vegetation (Fig. 3 in this document, Fig. 9 in the new version of manuscript).

Green color in Fig 3, right panel, indicates the pixels which have NDVI in the interval ($NDVI_{max,2}$-σ_2; $NDVI_{max,2}$+σ_2) corresponding to the upper peak of the distribution based on the data from

1968+1988. This peak is mainly associated with forests. The lower peak (pixels in blue color, Fig. 3b), as can be seen from Fig. 3, corresponds to woodlands and tundra. This lower peak has a large intersection with the peaks in the distributions based on the data after recent fire and background data. However, interestingly, there is a significant decrease in the pixels with NDVI below ca 0.52 in the bimodal distributions. These sites are marked in pink in Fig. 3, right panel. They correspond to the tundra sites lightest in color due to the presence of lichen in the vegetation community. The fraction of such pixels decreases in bimodal distributions, meaning that lichen does not recover to previous state. Instead, bimodal distributions gain a large fraction of pixels with high NDVI corresponding to forests, absent in background tundra.

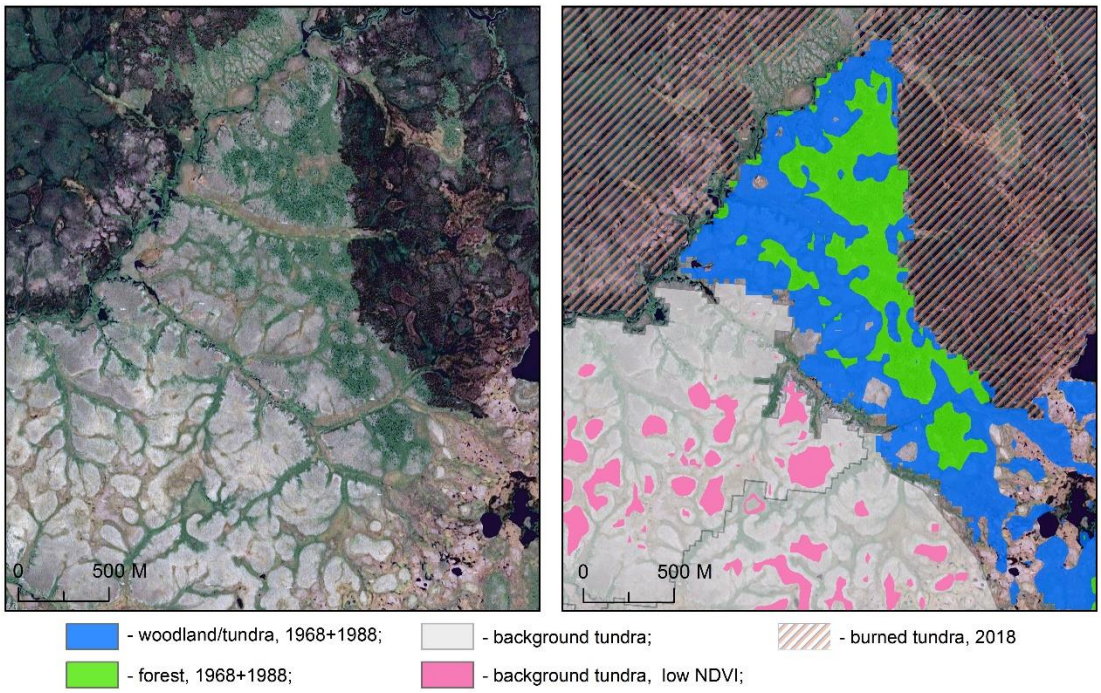


Fig. 3. Representative types of vegetation associated with different state of the sites and NDVI. Left panel: an image without mask, right panel: the same image colored according to the state of the site (burned in mosaics from 1968+1988 or 2018, background) and NDVI. In the right panel, green color corresponds to the upper peak and blue color corresponds to the lower peak in the bimodal distribution from the sites burned in 1968/1988 (Fig. 2b). Pink color marks pixels with NDVI lower than 0.52 in the background site.

We compared NDVI distributions based on Landsat 8 data from 30 June 2018 and 03 July 2019 that we used in the previous version of the manuscript (Fig. 4 in the current document). The distributions of NDVI are similar for both years, but in 2019, the upper peak in the bimodal distributions is less pronounced. This could be due to different phenological states of vegetation, dependent on temperature and precipitation from year to year. The peaks corresponding to tundra vegetation and woodlands almost did not change their position in both figures, but the peak of the distribution after the recent fire (2018 in the legends) and the forest peaks in bimodal distributions have larger mean NDVI in 2018.

Finally, we estimated the fraction of forest in bimodal distributions. We used NDVI data from 2018, as the separate peaks were better pronounced in bimodal distributions. Using standard deviations of the two peaks, the boundary approximately separating forest peak from tundra peak corresponds to the threshold value of NDVI=0.73. We assume that pixels with NDVI > 0.73 represent mainly forest, and pixels with NDVI<0.73 – tundra and woodlands. We integrated the distribution to find the fraction of pixels with NDVI>0.73 in the total number of pixels. In the areas burned 28 years

ago, forests occupied 24% of the total area. In the areas burned more than 42 years ago, the forest fraction increased to 55% of the total area. This number is comparable to our estimates of the vegetation shift within forest-tundra zone (56%) and exceeds the estimates for the northern taiga zone (14%).

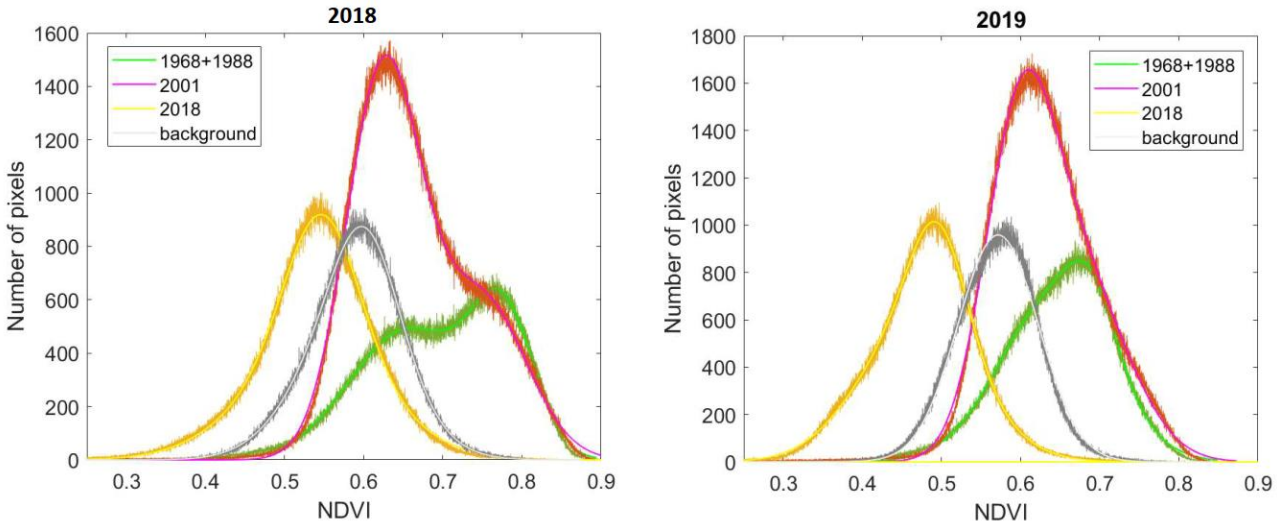


Fig. 4. Distributions of NDVI in the burned and background areas based on the images from 30 June 2018 (left panel) and 3 July 2019 (right panel).

While precise estimates of forest fraction based on NDVI are challenging, the main results following from the distributions in Fig. 2b can be summarized as follows:

1. The NDVI distributions based on the data from background tundra and areas burned 2 years ago are predominantly unimodal, whereas the distributions based on the data from areas burned 28 years ago and earlier are bimodal.
2. The low NDVI pixels corresponding to vegetation communities in tundra characterized by relatively high amounts of lichen and thus having lightest colors in the images largely disappear from the distributions corresponding to vegetation communities recovered after fires.
3. Instead, the new state of vegetation recovered after fires is characterized by a higher mean NDVI due to a new peak associated with forest. The fraction of pixels representing forest increases with time after the fire.

We thank again the referee for the useful suggestions. We hope that the present manuscript addresses all the comments raised.