**Review of the manuscript**

**The use of remote sensing for studying environmental changes in a palsa mire zone of Northern Finland in the period 1985–2024**

Małgorzata Szczepańska, Sebastian Czapiewski, Magdalena Suchora, Danuta Szumińska

This manuscript addresses the pressing issue of assessing the informative value of remote sensing (RS) data for studying landscape changes in subarctic tundra with sporadic permafrost. The importance of research in this area is primarily due to the need to identify areas (regions) of the most likely and intense landscape transformation and permafrost degradation due to climate change. Clearly, large-scale, detailed, and systematic observations can only be carried out using RS, while ground-based observation data from localized areas is necessary for training and validating satellite imagery results.

Using a local area in northern Finland as a case study, the authors attempted to link long-term Landsat satellite observations with ground-based observations characterizing landscape heterogeneity, with a focus on identifying the characteristics of the palsa mire zone, beneath which permafrost is observed. The stated objective is too general and does not fully align with the content and conclusions of the study, as, in particular, there is no information indicating the representativeness of the small study area (402 $km^2$) in relation to the whole of northern Finland. The authors should clarify the knowledge gap this study aims to address.

The study area's descriptions include excessively detailed descriptions of lithological features and soil properties, which are not clearly related to the analyzed parameters. At the same time, no information is provided on the characteristics, presence, absence, or transformation of permafrost in areas grouped into three soil types (palsa, organic soils, and mineral soils), making it impossible to relate the obtained results to the heterogeneity of geocryological conditions.

Satellite imagery data was analyzed using the NDVI, NDWI, MNDWI, NDMI, and LST indices for areas without water bodies. Unfortunately, the authors poorly understand the meaning and information content of these indices. The NDWI and MNDWI indices are not intended to characterize the surface moisture of land areas and vegetation. Interpreting them as indicators of vegetation or soil moisture can lead to erroneous conclusions. These indices (NDWI and MNDWI) were developed to identify water bodies and are calculated based on the differences in the reflectance spectrum of water compared to the spectra of other surface types. Higher NDWI and MNDWI values are attributed to water bodies, while lower values are attributed to all other surface types.

The NDWI index (Green-NIR/Green+NIR) is essentially an inverted (reversed) version of the NDVI index (NIR-Red/NIR+Red), since the Green and Red channels are located close to each other in the overall spectrum and differ only slightly in reflectance. This is confirmed by the authors' results, according to which all average values and long-term distributions of the NDWI index with a high negative correlation invert the values and distributions of the NDVI index. The NDWI index will always be negatively correlated with the

NDVI, **including regardless of vegetation moisture**. The authors' analysis of NDWI as an indicator of vegetation moisture led to a contradiction.

The authors write: (341) *«… the median NDVI in each period is notably higher for palsa mires (0.46–0.52) than for other organic soils (0.40–0.48) or mineral soils (0.37–0.43) (Figures 7 and S18). This indicates that biomass productivity is highest within palsa mires. There is a similar increase through P1, P2 and P3 for the soil moisture index (NDMI)…».* Hence, the conclusion is that palsa mires are characterized by high moisture (as measured by NDMI). According to the authors' results, the NDVI index correlates well and positively with the surface moisture index NDMI, which is confirmed by the results in earlier publications by other authors.

At the same time, the authors write: (570) *«It is also interesting that NDWI values were lowest on palsa mires as compared to the other substrate types, which may be due to their unique relief. Their ice core elevates these formations above the mire level, which forces a natural slope of the terrain. As a result, **water does not accumulate on the tops of the palsas but flows down to the lower depressions**.»* From these two fragments of text, it is unclear whether the humidity in palsa bogs is high or low.

In fact, the NDWI values in the palsa swamps are the lowest for the same reason as the highest NDVI values. The vegetation here has higher green phytomass and, consequently, higher surface moisture (including NDMI). Low NDWI values simply indicate **the absence of a water surface**.

It should be noted that the NDMI index clearly characterizes moisture content only at the surface of the vegetation cover. The NDMI index clearly cannot characterize the moisture content of the soil beneath the vegetation, since the penetration depth of solar radiation at these wavelengths does not exceed 0.6 cm. The relationship between moisture content at the surface of the vegetation and in the soil is not always clear, so the NDMI index cannot be considered a soil moisture indicator, but only an indicator of surface moisture.

The authors write that at the stage of processing remote sensing data, spectral differences between generations of Landsat satellites were corrected, and for this purpose, the intersensor calibration equations proposed by (Roy et al. 2016) were used. However, the work (Roy et al. 2016) presents formulas for correcting data obtained from the OLI (Landsat 8) and ETM+ (Landsat 7) radiometers and there is no information on intersensor correction between the data of these radiometers and the data of the TM radiometer (Landsat 5). At the same time, there are publications indicating a difference in the spectral response of the ETM+ and TM radiometers, which is manifested in particular by higher values of the NDVI and NDMI indices according to the ETM+ radiometer data compared to the TM radiometer. It is also known that these differences in the spectral response depend on the types of surface and vegetation of the analyzed areas. The authors do not provide information on the correction of the TM radiometer data. In most cases, neglecting this type of correction **leads to overestimated values of the slopes of long-term trends** in the NDVI and NDMI indices when using data from all three types of radiometers - TM, ETM+ and OLI. Thus, the average values obtained by the authors NDVI and NDMI indices and the slopes of the trends of these indices may be erroneous due to the lack of a stage of intersensor correction of TM radiometer data.

Unfortunately, the authors do not provide specific formulas from (Jimenez and Sobrino, 2003) by which they calculated the LST temperature and the LSE emissivity of the surface.

To statistically analyze and characterize long-term trends in index changes, the authors propose linear mixed-effects models (LMM), justifying this approach by the hierarchical nature of the data (the nature of the hierarchy is not explained) and the absence of data for some years. The authors do not provide specific formulas for converting the original data into data corresponding to LMM, which precludes the possibility of verifying the results of the statistical analysis and long-term trends using LMM. The authors' LMM results do not allow for direct comparison with those of other authors using more common statistical and regression analysis methods (t-statistics, Mann-Kendall tests, ordinary least squares (OLS), and the slope of the trend line using Sen's method), which are used, among other things, to analyze long-term trends with missing data for some years.

According to the authors' results, all index trends (Fig. 5, Table 2) have very high significance levels ($p < 0.001$), including for LST, which raises doubts given the observed data scatter. In this regard, the reviewer conducted a statistical analysis, in particular, of the LST trend (Fig. 5) based on the attached remote sensing dataset (https://doi.org/10.5281/zenodo.20817134) using OLS and t-statistics to assess the significance level. The result for LST: $\beta = -0.016$; $0.1 < p < 1$. The LST trend is insignificant and cannot be analyzed as a pattern. Elimination of one LST value (1985) leads to a change in the sign of $\beta$. Such a significant difference in the significance level of the LST trend based on the t-statistic from that shown in Table 2 ($p < 0.001$) cannot be attributed to the specifics of the LMM analysis. In this regard, the results of the statistical analysis for other indices also raise doubts.

According to the reviewer, the manuscript is overloaded with statistical analysis methods, graphics, and tables that duplicate fairly simple and obvious conclusions about the relationships between the indices themselves, as well as between the indices and predictors. The text is excessively replete with repetitions and unexplained terms (*environmental gradient, environmental control, fixed effect, sampling point, random intercept, continuous predictor, predictor redundancy, heat load*). The text also contains many complex phrases, which significantly complicate understanding and obscure the essence of the conclusions (e.g. (409),… *the joint variation in standardised point-level mean values of RS indices was strongly structured by environmental conditions;* (414) *…This indicates that the environmentally structured component of RS variability was organised mainly along a single dominant gradient.)*

The "Discussion" section provides overview information that is clearly unrelated to the results obtained. The authors do not provide specific information on long-term permafrost transformation trends in the study area, so repeating well-known conclusions about these processes is unnecessary. The "Results" section provides an interpretation of the data, which should be included in the discussion.

General information about the potential and limitations of RS data application in the Arctic is well known. Authors should consider the potential and limitations of the methods used in their work if they are original or are being applied for the first time in similar conditions.

**Comments on the authors' conclusions:**

*(637) 1. Temporal trends in the vegetation index (NDVI) and moisture indices (NDMI, MNDWI) increased throughout the entire study period, in contrast to the water index (NDWI) and surface temperature (LST), which decreased towards the end of the study period.*

Comm: The NDWI index mirrors the NDVI index; LST trends are insignificant and cannot characterize a decrease in surface temperature.

*(640) 2. A strong positive correlation between the vegetation index (NDVI) and the moisture index (NDMI) was found and a negative correlation between the vegetation index (NDVI) and the water index (NDWI).*

Comm: The positive correlation between NDVI and NDMI is well known for both tundra vegetation and other landscape types. The explanation for the negative correlation between NDVI and NDWI is given above.

*(642) 3. Spatial differentiations in individual indices and in trends over the period 1985–2024 were found.*

Comm: Spatial differentiation of indicators is obvious, as it is related to landscape heterogeneity. Differentiation in trends (direction and rate of change) by soil category is not obvious (Fig. 6). Numerical values are not provided.

*(643) 4. Discrepancies were found between the air temperature trend (continuous increase) and the surface temperature index trend (decrease at the end of the period).*

Comm: The LST trend is insignificant. Overall, the conclusion is trivial, as it is known that LST depends not only on air temperature but also on evapotranspiration, humidity, and surface thermophysical properties.

*(645) Distinct environmental factors were identified that determined the multivariate structure of remote sensing indices. Of these, soil categories, DTW2, and DEM were the strongest and most consistent predictors across all three subperiods from 1985 to 2024.*

Comm: The conclusion is unclear. Do these predictors influence the average index values or their long-term trends?

*(647) The importance of soil categories and the DTW2 parameter increased significantly in the last period (P3: 2013–24), indicating the progressive intensification of environmental control over the spectral signal.*

Comm: The conclusion is unclear. The spectral signal always depends on surface conditions. If there is no information on changes in soil conditions and DTW2, then how can we talk about an increase in their importance?

*(648). These results also confirmed the clear thermal and hydrological distinctiveness of palsa mires, which throughout the study period retained an environmental character that was entirely distinct from the other organic soils (i.e., dystric histosols, sapric histosols) and the mineral soils (dystric leptosols, endogleyic podzols, hypersceletic leptosols).*

Comm: According to the authors, palsa swamps differ from other organic and mineral soils only by their higher average NDVI and NDMI values, which is known and typical for any peatlands with a high content of ice inclusions.

*(652) The use of Landsat images, being characterised by high spatial resolution (30 metres), enabled the detection of even small spatial changes, which is important due to the heterogeneous and mosaic landscape of the young glacial landscape of northern Finland.*

Comm: It is well known that remote sensing data characterizes the heterogeneity of any landscape.

(654) *At the same time, the Landsat mission provides a long time series of images, which allowed for the analysis of changes over as much as 40 years. The approach used can be easily applied to other study sites, which can help understand and quantify changes in dynamically changing permafrost areas, including those at the margins of its range.*

Coom: The authors' approach is well-known, but what's new is unclear?

**Summary**

Overall, of all the results obtained by the authors, the differences in the average NDVI and NDMI indices for the three soil categories, with the highest index values for palsa, are considered the most reliable and informative. Palsa bog zones represent typical peatlands with a higher content of ice inclusions. This result is not original, as similar results by other authors have been reported, confirming anomalously high NDVI and NDMI indices in peatlands with higher moisture/ice content in the thawed soil layer and higher ice content below the seasonal thaw layer. No obvious differences in the trends in index change for the three soil types were identified. In the reviewer's opinion, this work does not contain fundamentally new results and is not recommended for publication, including due to errors and inconsistencies in the remote sensing data analysis.