

## ABSTRACT

The Taimyr wild reindeer herd is the largest and the longest monitored reindeer herd in Eurasia. The Extreme North Agricultural Research Institute in Norilsk, Russia carried out monitoring since 1969. However, monitoring data accumulated in the form of paper hard copies and the documentation was kept in Russian. The reports were not digitized and processed using modern geospatial technologies. Also, the only limited amount of the data was published, so these works remained inaccessible to most scientists. The goal of TAMARA the project is to digitize the data contained in the reports, produce digital maps and derivative products, identify the herd migration patterns, disseminate the derived datasets and publish the analysis results. Now we nearly completed the digitization of the data contained in the reports and created digital derived products showing the spatial distribution of reindeer herds and their migration process. This poster presentation will present the initial results of the digitization and analysis of the distribution of herds based on aerial census reports.

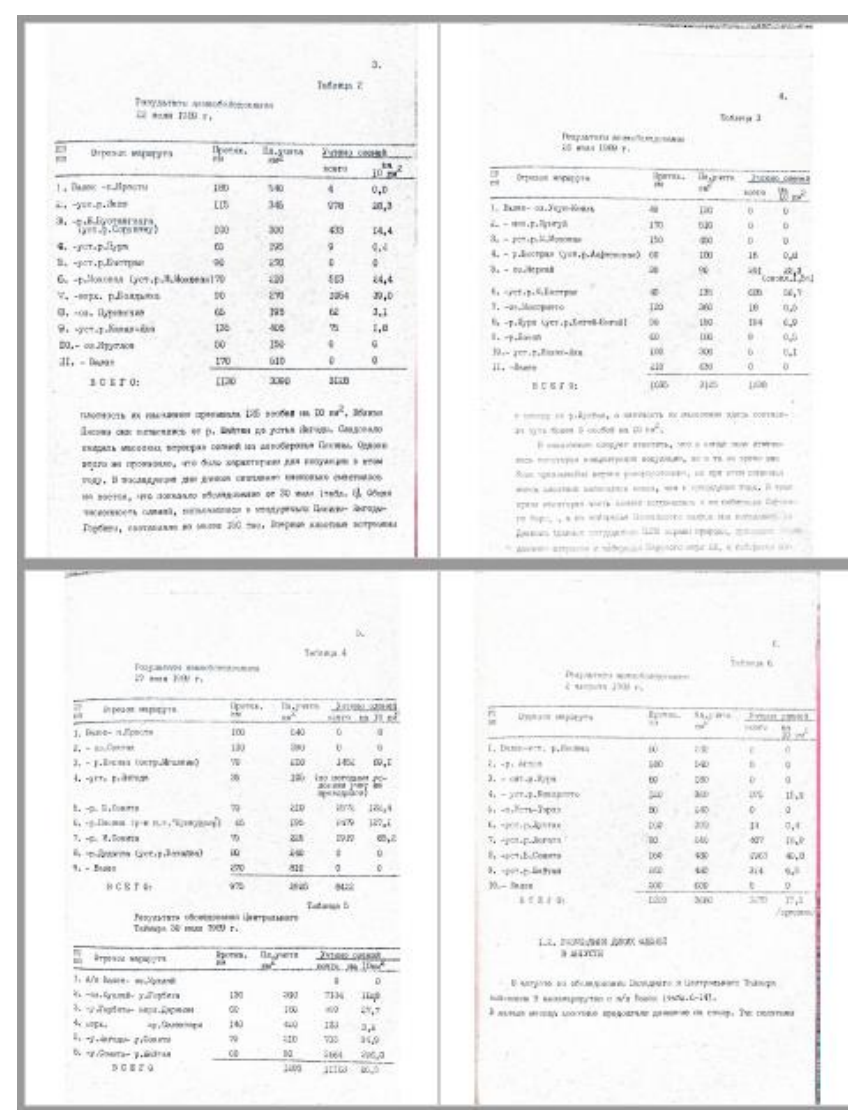
## RESEARCH OBJECTIVES

- Retrieve, digitize, archive, transfer, process and publish historical observational data collected between 1969 and 2009, document relevant methodologies, and develop appropriate metadata;
- Develop value-added data products resulting from the (re)analysis of historical spatial migration patterns of the TRH;
- Contribute to developing polar spatial cyberinfrastructure by creating a comprehensive and open data visualization and dissemination system.

## STUDY AREA & SOURCE DATA



Study area location, small and large scale



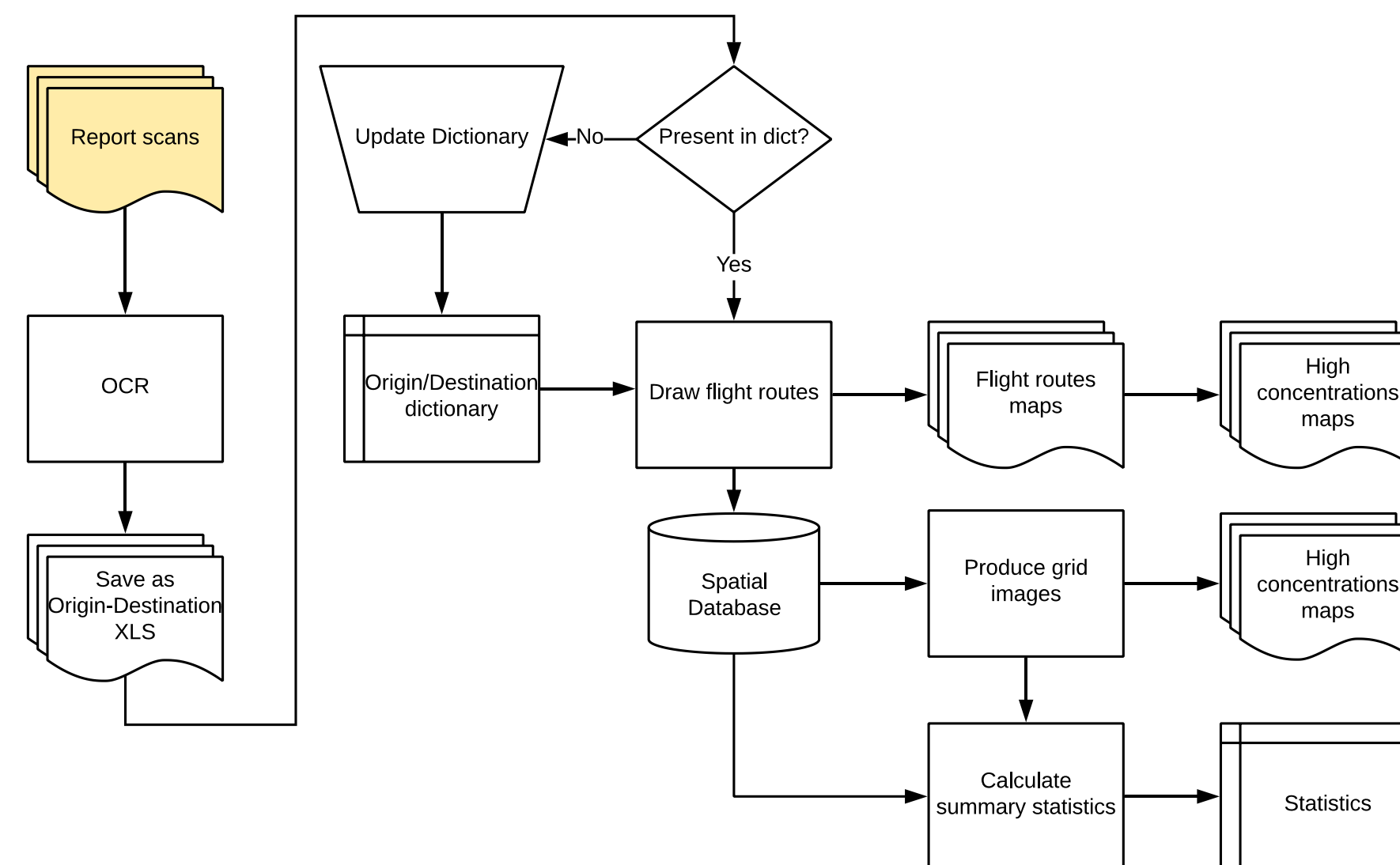
Scanned pages of original report(1989)

The study area covered Taimyr Dolgan-Nenets district and Evenki district, Russia. The source data for our study were aerial census reports on the reindeer distribution. Deer counts were produced during flights on small aircrafts with trained observers. Data was collected from 1969 to 1991. A typical report contained air census tables with data on the point of departure, destination, length of the flight segment, number of dispersed deer and the deer within concentrations.

## LIMITATIONS

- Data on flight segments is not reported for every year, for some years only a summarized text description of the distribution was given or year skipped entirely.
- Often no accurate location of origin-destination is reported, for instance, just name of the river is given.
- Hard copies were produced using typewriter and carbon paper and likely stored in inappropriate environmental conditions. It greatly obstructed OCR text extraction.

## METHODOLOGY



Text from scanned reports was recognized using Finereader OCR software, and after manual proofreading and correction, we saved data as Excel files. Excel file fields are the same as in the final feature classes in the spatial database.

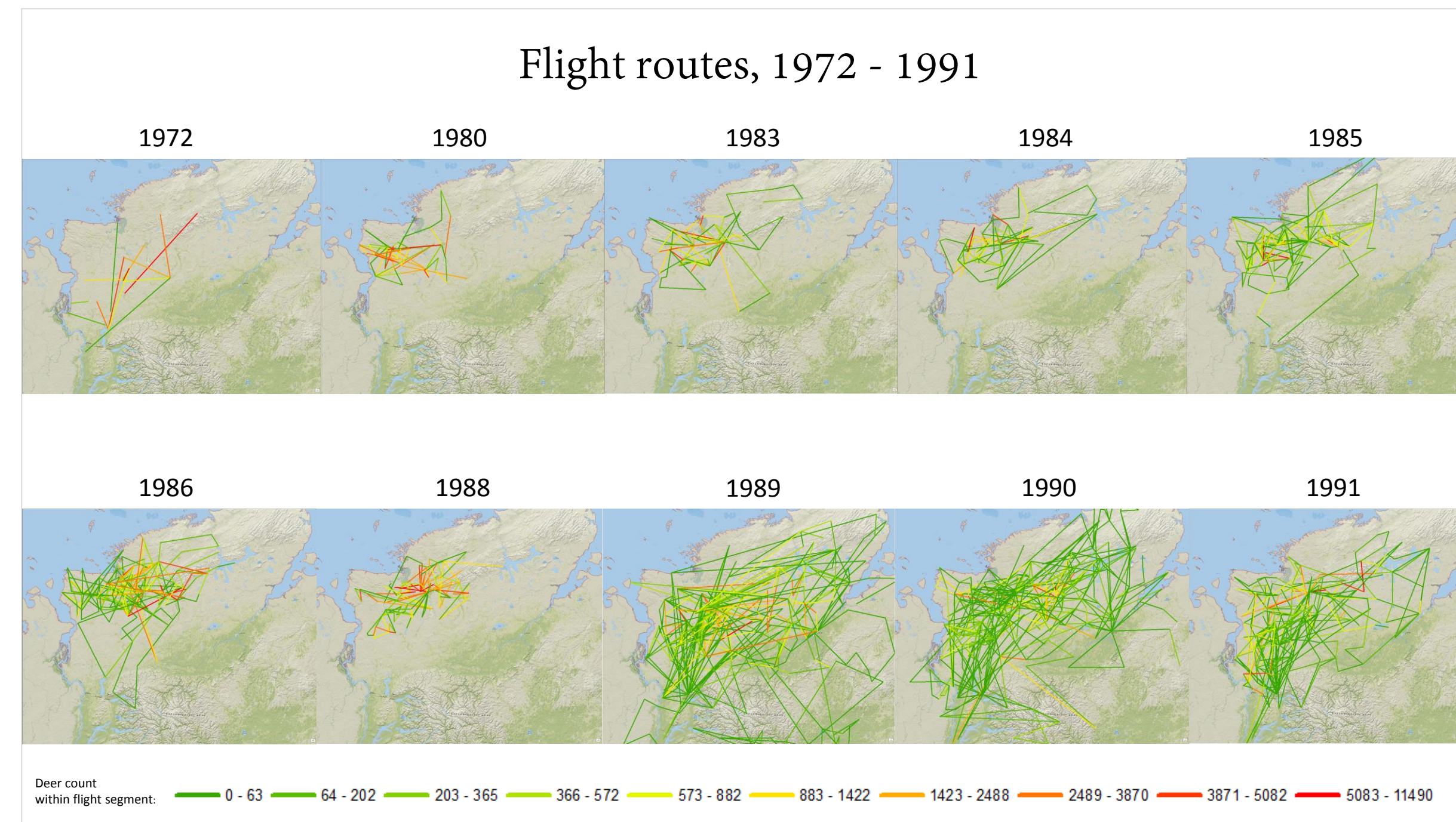
To automatically create the geometry of flight segments, we created the location dictionary class which contains all origins and destinations to be linked with geographic names, mentioned in the source Excel file

Most geographic names in the dictionary (70%) were geocoded with Yandex geocoder (<http://maps.yandex.ru>) which is the most accurate geocoding solution for post-soviet territory. The remaining names were found manually on a 1: 200000 scale topographic map. High concentrations maps were produced by selecting flight segments from flight routes class, where the count of reindeer in herds was provided. In most cases, point feature created in the middle of the flight segment. If detailed information to pinpoint herds was provided in the report, point feature created accordingly.

## RESULTS

After we made sure that the dictionary contains all the origin and destination geographical names, we matched the names from the dictionary with data in the Excel file. For reports with a small number of flight segments (up to 150), matching was done manually. Then we developed a mapping in Altova Mapforce ETL software, to automatically match dictionary entries with Russian geographic names in the mentioned file.

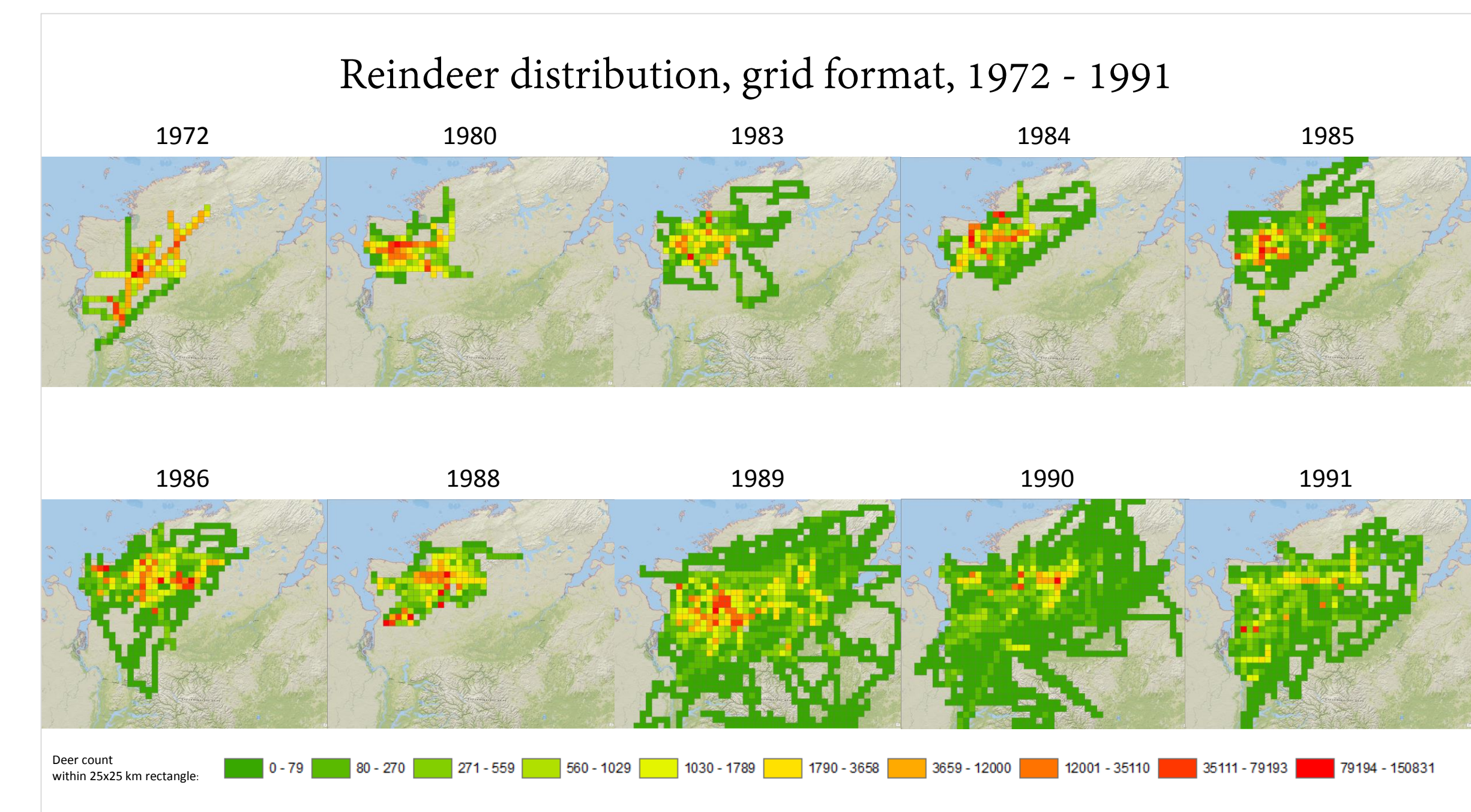
Having all the origin and destination entries matched with the dictionary, the flight routes were automatically created using FME Desktop. Data was saved as an ArcGIS Feature Class in a file GDB using UTM84-47N projection. The resulted flight routes are depicted below:



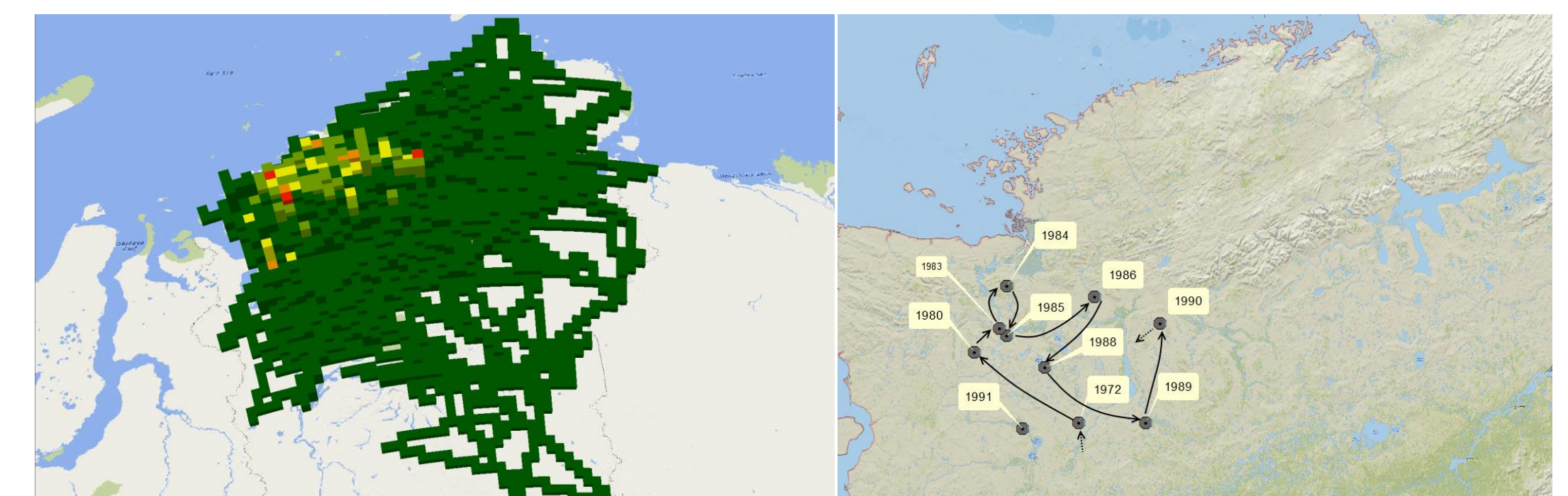
## RESULTS

Derived products were created based on flight segment data: a set of high concentration maps and series of deer distribution map in a grid format. We used the following steps:

- Create polygons of a 25x25km size covering each flight route.
- Cut flight segment by polygon borders.
- Calculate reindeer number within every piece of cut flight segments as  $number = (piece\_length / total\_length) * observed\_number$
- For every polygon - Sum up reindeer number from every piece falling into polygon and assign this number to polygon attribute.



To get a generalized picture of the long-term distribution of herds, we superimposed all the layers in the grid format and calculated how often polygons overlap and summed up the number of deer in overlapping polygons of the grid. The extrusion height represents how often the particular location was surveyed, color - the accumulated number of reindeer inside overlapping polygons, sharing the same location. Additionally, mean center map reflecting the displacement of the population core during the years was created.



Superimposed grid datasets

1972-1991 mean centers map

## CONCLUSION

- We are completing the digitization of reindeer aerial census data. We created a spatial database of aerial census data for 1969 to 1991 with flight segments, high concentrations, and grid representation per each available report.
- The created data provide more accurate localization of the spatial distribution of wild reindeer than previously published data without a detailed geographic reference.
- We are planning to perform a more in-depth analysis of the movement patterns of reindeer taking into account the influence of climatic, natural and anthropogenic factors.
- All created data will be uploaded to the ARCTIC Datacenter website for free access.

## ACKNOWLEDGEMENTS

We are grateful to National Scientific Foundations for the award #1504934