**Supplementary information**

Declining human pressure and opportunities for rewilding in the steppes of Eurasia

Appendix S1: Mapping cropland abandonment

We mapped cropland abandonment across our study region by classifying Landsat image composites for the years circa 1990 (representing the end of the Soviet era), 2000 (one decade after the beginning of the transition period, representing the time when land-use intensity was lowest), and 2015 (current situation, following a partial revival of the agricultural sector). We produced Landsat image composites, which summarize all the spectral information for the time period of interest and have several advantages over more traditional classifying approaches that rely on single Landsat images. Composites provide typically gap-free, wall-to-wall coverage across large study region, composites can be derived specifically for one or several phenologically distinct time points of interest, and provide detailed information via spectro-temporal metrics calculated for all available images (e.g., mean or standard deviation of the reflectance of all cloud-free observations) as well as the meta-information for these images (e.g., the number of cloud-free images used to calculate the metrics).

For this study we calculated three components for each of our three time points (i.e. 1990, 2000, 2015). First, we calculated three so-called best-pixel composites centered around the Julian days 121, 180, and 260, representing the best individual observations from the spring, summer and fall of each time point, which is important for mapping cropland-grassland dynamics and farmland abandonment using Landsat imagery (Kuemmerle et al. 2008; Prishchepov et al. 2012). Second, we derived a set of ten spectro-temporal metrics per band (for more details on how these metrics were calculated, see Griffiths et al. (2013)). Third, we derived information on the number of clear-sky observations as well as the zenith and azimuth for each of the best-pixel composites. This resulted in a layer set of 85 bands overall for each of the three time points. All components of all three time points were merged into one multi-temporal stack.

We gathered training data for our classification via on-screen digitization of areas representative for our land-cover classes based on high-resolution imagery from Google Earth, extensive expert-based knowledge from field visits and local collaborators, and the Landsat image composites themselves. The most important classes in this study were “croplands” and “grasslands” and changes between them. An area was defined as cropland in one year, when clear signs of open soil were visible during spring, and a clear vegetation signal during summer. Grasslands were characterized by a clear vegetation signal in spring, summer and fall. Training data were gathered in form of larger polygons to cover the spectral variability within classes.

Once we had collected a sufficient number of polygons for each class, we randomly sampled 5,000 points per class and used this as input to a random forest classification algorithm. We classified our entire Landsat image composite stack and gathered more training polygons iteratively in areas that were misclassified. Finally, we applied a minimum mapping unit of 10 Landsat pixels (equaling 0.9 ha).

We validated our change map using independent data not used in the classification process. This dataset consisted of 100 randomly sampled points per class (Stehman 2009), which we labeled according to our class definition based on visual inspection of the Google Earth imagery and the Landsat composites. We generated an error matrix and calculated overall classification accuracy as well as class-wise user’s and producer’s accuracies (Foody 2002). We corrected for potential sampling bias and calculated confidence intervals around our error and area estimates (Olofsson et al. 2014). Our maps had an overall accuracy of 86.3%. User’s and producer’s accuracies were generally high, and highest for the stable classes (e.g. stable cropland with 94.6% and 81.9%, Tab SI1.1).

Tab SI1.1: per class user’s and producer’s accuracies

|  |  |  |
| --- | --- | --- |
|  | User’s accuracy | Producer’s accuracy |
| Stable steppe | 94.00% | 93.58% |
| Stable cropland | 81.91% | 94.63% |
| Cropland abandoned 1990-2000 | 82.00% | 69.10% |
| Cropland abandoned 2000-2015 | 81.00% | 89.71% |
| Cropland abandoned 1990-2000  and recultivated 2000-2015 | 70.00% | 75.30% |
| Other | 85.11% | 90.91% |
| **Overall accuracy** | **86.34%** |  |

Appendix S2: Mapping changes in livestock distribution and human settlements

We anticipated that human population density and the associated density of herded (free-ranging) livestock were directly related to disturbances for wildlife. We therefore quantified patterns of change in the density of settlements and livestock breeding and herding infrastructure between Soviet times and now.

Livestock stations, across the former Soviet Union, are outposts on summer (‘Letovki’) and winter pastures (‘Zimovki’), where livestock are concentrated and where they do not need to be kept in sheds overnight, or taken back to the stables for milking or other uses. Such stations usually consist of 1-3 houses or felt tents (‘yurts’) for shepherd accommodation and a corral to confine the livestock during the night. To estimate rates and patterns of change in livestock station and settlements, we created a vector layer of both for ca. 1984 (representing the still intact Soviet infrastructure) and ca. 2012 (representing the current situation). For the Soviet period, we manually digitized settlementsand all livestock infrastructure (stations but also single building housings, places were single families lived in yurts, or processing milk and meat factories, which are all contained on Soviet topographic maps) across the study region (n = 6,482)from official, georeferenced Soviet topographic paper maps scaled 1:200,000 (VTU 1989). For the period around 2012, we estimated from publicly available very high-resolution satellite images, whether the settlements and livestock stations identified on the maps were still used. We assessed Google and Bing maps using the open layer plugin in QGIS 2.18. We first drew polygons around each settlement and livestock station on the Soviet maps. These were then superimposed on the contemporary satellite images. We visually inspected the settlements and stations and estimated which proportion (in %) of the buildings were derelict ca. 2012 compared to the end 1980s (0%: no intact buildings left, settlement/station fully abandoned, values >100% growth in the area of the settlement/station). Abandoned buildings were easily identified as they collapse quickly due to the predominance of clay bricks used for construction. Further evidence of complete abandonment of the livestock stations were the absence of recently used dirt tracks and the absence of larger livestock and fresh corrals. Active dirt tracks appear as pale lines due to the light surface soil color and the sandy or silty soil texture. Active corrals are easily detected on the satellite images due to their quadratic shape and the dark color resulting from overnight animal dung deposition and soil tramping.

TheSoviet topographic maps were issued between 1970 and 1988 (with one exception, a map that was produced in 1962 and never updated). Issue dates were given on the map margins. Ca. 90% were updated and published in the period 1981–1985. Satellite images used to evaluate the contemporary situation were mostly from 2010 to 2014, with a peak availability in 2012. For the Bing images, information on the image date was extracted from <http://mvexel.dev.openstreetmap.org/bing/>. For the Google images, the date was read directly from Google Earth. Figs. S2.1 to S2.4 depict examples of changes in settlement and livestock station infrastructure.



Figure S2.1: Almost completely abandoned settlement (>90% abandoned); only two intact, roofed houses remain, roads only partly used (view altitude: 1 km). Summer 2014.

****

Figure S2.2: Intact settlement (100% intact), no derelict houses visible, large cattle stables at northeastern end of village (view altitude: 1 km). Summer 2016.



Figure S2.3: Former “Zimovka” (livestock station), summer 2013 (view altitude: 1 km). No signs of current use, no intact building, road overgrown and not used, no active corrals visible (0% intact).



Figure S2.: Used “Zimovka” (livestock station), summer 2015 (100% intact). Buildings maintained, road network in use, corrals (with dark surface due to dung deposition) visible (view altitude: 1 km). Different corrals are visible here: the small quadratic ones close to the stable are for keeping sheep overnight, while the large, circular one is used to concentrate horses for milking and veterinary treatment.

Appendix 3: Assessing landscape connectivity

We modelled landscape connectivity based on circuit flow theory, using Cirtcuitscape 4.0. One common problem of this approach is that current flow between close nodes, protected areas in our case, are disproportionally high (Koen et al. 2014). To circumvent this problem, we assessed landscape connectivity independently from node placement, following the approach outlined by Koen et al. (2014). In this approach nodes are placed randomly in a buffer around the region of interest. Resistance values within the buffer are randomly assigned, but drawn from the distribution of resistance values within the study region. This procedure requires testing two parameters: buffer width and the number of nodes and we followed the general guidance by Koen et al. (2014) and Leonard et al. (2017) to find these parameters. Specifically, we determined the optimal number of nodes by first sampling an initial number of nodes (i.e., 20, (Koen et al. 2014)), then increasing the number of nodes by 10 and comparing the two current density maps using Pearson’s correlation coefficient. If r was below 0.98, we again increased the number of nodes by 10, until the correlation coefficient between the two maps exceeded 0.98 (Leonard et al. 2017). For this study here, the number of nodes at which r saturated was 60. Once we determined the optimal number of nodes, we identified the optimal buffer width around the study area. Following (Koen et al. 2014), we increased the buffer width in 5% increments of the study area extent, while keeping the number of nodes constant (at 60). Again, we compared current density maps until r>0.98 saturated, which was at a buffer width of 50% of the study area extent.

We then used this parameter combination and calculated landscape connectivity of our study area for both years of interest (i.e., 1990 and 2015). Last, we calculated the differences in landscape connectivity between 2015 and 1990. As we calculated combined human influence across our study region in three ways (i.e. by calculating the product, the maximum, and the mean of our three input layers—percent cropland, distance to livestock stations, distance to settlements), we calculated current density maps for each of these scenarios.



Figure S3.1: Human influence index and steppe rewilding index in comparison between the two ways of how the three input layers were combined: (A) the product of the three layers; (B) the mean of the three layers.

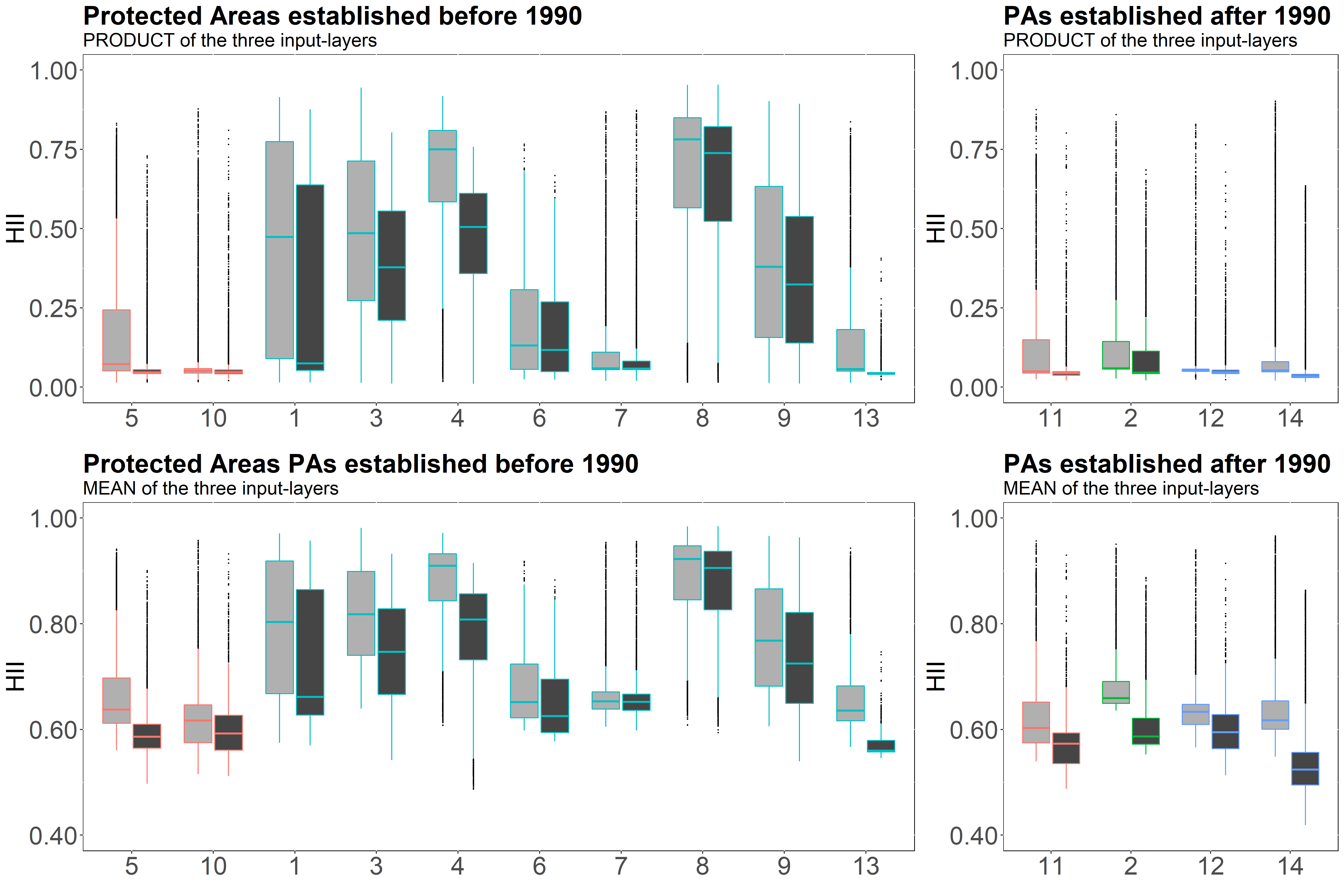


Figure S3.2: Comparison of human influence values within the protected areas when combining the three input layers through multiplying them (top row) or calculating the average (bottom row).

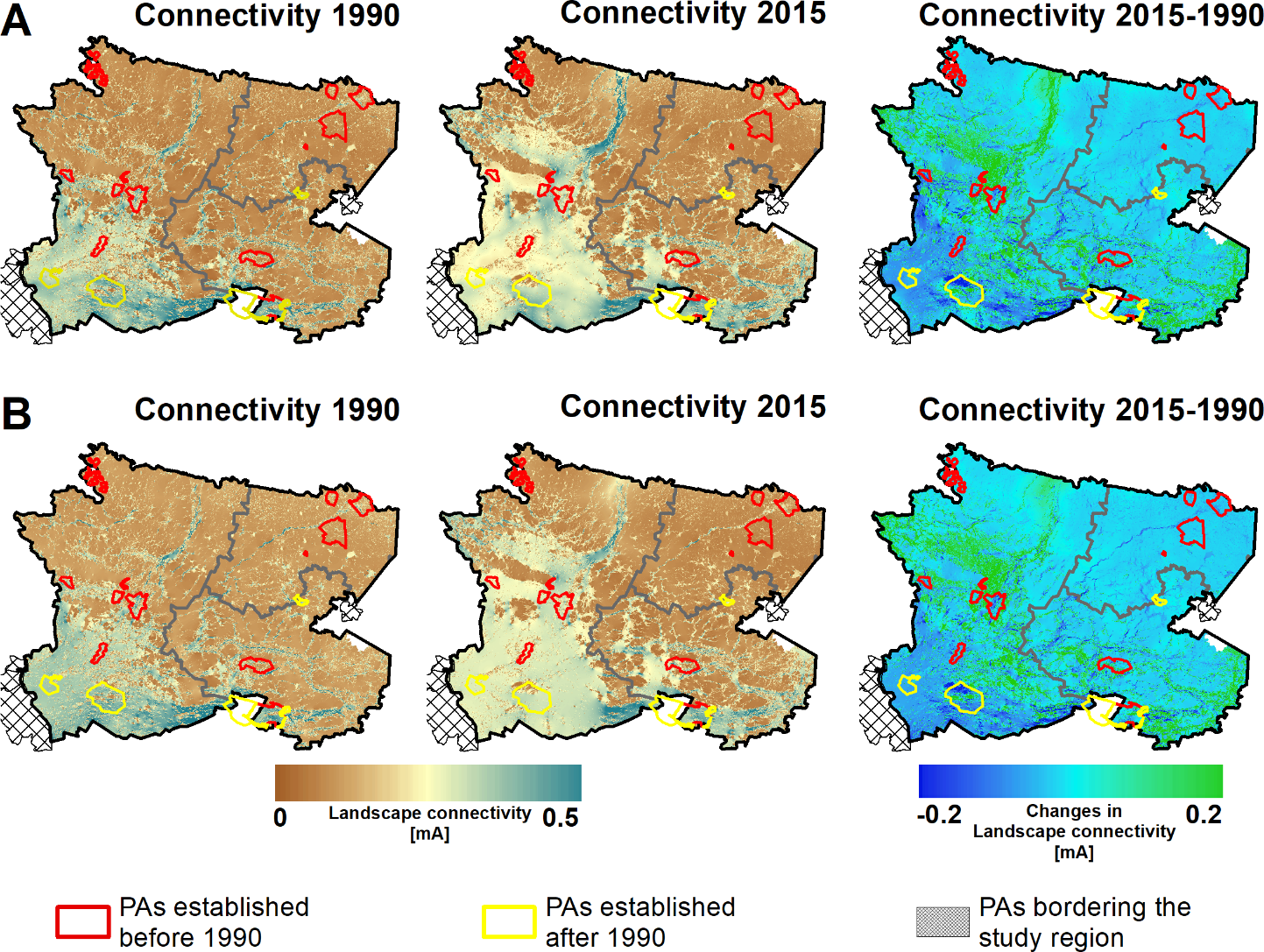


Figure S3.3: Landscape connectivity compared for the two ways of how the input layers were combined: (A) the PRODUCT of the three layers; (B) the MEAN of the three layers.

Table S3.1: Protected areas in our study region, and their protection status. The numbering refers to the numbering in figure 3 of the main manuscript.

|  |  |  |
| --- | --- | --- |
| # | Protected area name | IUCN-Category |
| 1 | Atbasarskii Zakaznik | IV |
| 2 | Kokshetau National Park | II |
| 3 | Mamlyutskii Zakaznik | IV |
| 4 | Mikhailovskii Zakaznik | IV |
| 5 | Naurzumskii Gosudarstvennyi Zapovednik | I |
| 6 | NN (unknown name) | IV |
| 7 | Sarykopinskii Zakaznik (now part of Altyn Dala Gosudarstvennyi prirodnyi reservat) | IV |
| 8 | Smirnovskii Zakaznik | IV |
| 9 | Sogrovskii Zakaznik | IV |
| 10 | Tengiz-Korgalzhynskii Gosudarstvennyi Zapovednik | I |
| 11 | Tengiz-Korgalzhynskii Gosudarstvennyi Zapovednik (extension) | I |
| 12 | Tosynkumskii (part of Altyn Dala Gosudarstvennyi prirodnyi reservat) | VI |
| 13 | Tounsorskii Zakaznik | IV |
| 14 | Uly-Zhylanshikskiii (part of Altyn Dala Gosudarstvennyi prirodnyi reservat) | VI |

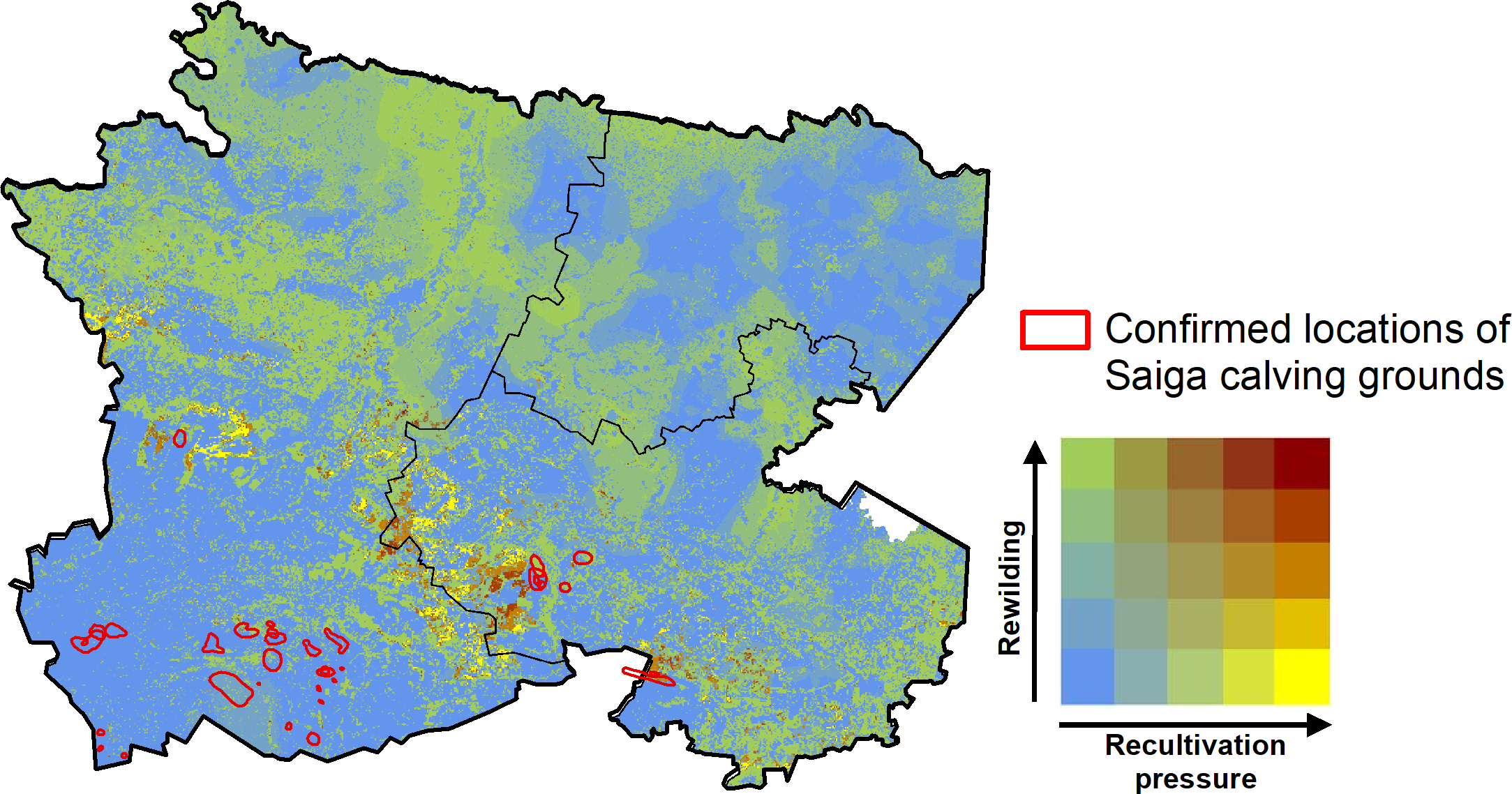


Figure S3.4: Confirmed locations of Saiga calving grounds, overlaid over our bivariate representation of SRI and recultivation trends. For representation purposes only one version is provided (i.e., the product of the three input layers, see Figure 4 for reference). The calving grounds represent locations that were either directly located in the field, come from aerial surveys or telemetry data. Data source: (Robinson et al. 2017).

References

Foody GM. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment **80**:185-201.

Griffiths P, van der Linden S, Kuemmerle T, Hostert P. 2013. A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **PP**:1-14.

Koen EL, Bowman J, Sadowski C, Walpole AA. 2014. Landscape connectivity for wildlife: development and validation of multispecies linkage maps. Methods in Ecology and Evolution **5**:626-633.

Kuemmerle T, Hostert P, Radeloff VC, van der Linden S, Perzanowski K, Kruhlov I. 2008. Cross-border Comparison of Post-socialist Farmland Abandonment in the Carpathians. Ecosystems **11**:614-628.

Leonard PB, Duffy EB, Baldwin RF, McRae BH, Shah VB, Mohapatra TK. 2017. gflow: software for modelling circuit theory‐based connectivity at any scale. Methods in Ecology and Evolution **8**:519-526.

Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA. 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment **148**:42-57.

Prishchepov AV, Radeloff VC, Dubinin M, Alcantara C. 2012. The effect of Landsat TM/ETM+ image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. Remote Sensing of Environment **126**:195-209.

Robinson SJ, Singh N, Zuther S. 2017. Saiga antelope die-off and calving sites in Kazakhstan. NERC Environmental Information Data Centre.

Stehman SV. 2009. Sampling designs for accuracy assessment of land cover. International Journal of Remote Sensing **30**:5243 - 5272.