Rufin, Philippe; Peña-Guerrero, Mayra Daniela; Umirbekov, Atabek; Wei, Yanbing; Müller, Daniel

Article — Published Version

Post-Soviet changes in cropping practices in the irrigated drylands of the Aral Sea basin

Environmental Research Letters

Provided in Cooperation with:
Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Halle (Saale)

Suggested Citation: Rufin, Philippe; Peña-Guerrero, Mayra Daniela; Umirbekov, Atabek; Wei, Yanbing; Müller, Daniel (2022) : Post-Soviet changes in cropping practices in the irrigated drylands of the Aral Sea basin, Environmental Research Letters, ISSN 1748-9326, IOP Publishing, Bristol, Vol. 17, Iss. 9, https://doi.org/10.1088/1748-9326/ac8daa , https://iopscience.iop.org/article/10.1088/1748-9326/ac8daa

This Version is available at:
https://hdl.handle.net/10419/264382

Post-Soviet Changes in Cropping Practices in the Irrigated Drylands of the Aral Sea Basin

Philippe Rufin[1,2*], Mayra Daniela Peña-Guerrero[2,3,6], Atabek Umirbekov[2,3,6], Yanbing Wei[3,4], and Daniel Müller[2,3,5]

[1] Earth and Life Institute, Université Catholique de Louvain, Place Pasteur 3, 1348 Louvain-la-Neuve, Belgium

[2] Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

[3] Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany

[4] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences (CAAS), 12 Zhongguancun South St., Beijing 100081, China

[5] Integrative Research Institute on Transformations of Human-Environment Systems (IRI THESys), Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

[6] National Research University TIIAME, 39 Kari Niyazov St., Tashkent, 100000, Uzbekistan

*corresponding author: philippe.rufin@geo.hu-berlin.de

Supplementary Material

S1: Landsat data and pre-processing

We used all available Landsat TM, ETM+, and OLI Collection 1 L1TP surface reflectance images from the period 1987 to 2019, except images that failed to meet the highest quality standards (as indicated by image quality flags) and those that had more than 70% cloud cover. All images were cloud masked using pixel-wise quality flags to indicate the presence of clouds, cloud shadows, snow and ice, and cirrus clouds in the case of Landsat OLI (Zhu *et al* 2015). We further removed clouds detected with medium confidence to maintain an aggressive cloud removal. For all remaining images, we calculated the enhanced vegetation index (EVI). All image classification and pre-processing steps were performed in Google Earth Engine (Gorelick *et al* 2017).

For each year, we generated spectral temporal metrics (STM) that represent descriptive statistics of the reflectance signal in predefined seasonal time windows and have been widely used in the context of land-cover and agriculture mapping (Waldner *et al* 2015, Phalke and Özdoğan 2018, Rufin *et al* 2021a). We defined three seasonal windows: April through July ("wet season"), August through October ("dry season"), and April through October ("growing season"), following the literature on satellite-based crop phenology in the study region (Conrad *et al* 2011, Conrad *et al* 2014). For each season, we generated median, 25th, and 75th percentiles, interquartile ranges, and standard deviations for all six Landsat spectral bands plus EVI. Interquartile range and standard deviation in the wet season were sensitive to several years' data scarcity and thus removed from further analysis.

We derived additional features that represent the texture and local contrast within selected STM. Concerning texture, we selected 25th and 75th percentile EVI bands for the dry season and the growing season window and calculated median values in a circular radius of 150, 300, and 900 meters for each pixel. Local contrasts were included by creating a normalized difference between the STM and the texture metrics. These features normalize each pixel value with their surrounding areas and emphasize contrasting spectral-temporal behavior on the land surface, e.g., field boundaries, roads, or patches of woody vegetation.

We included additional layers representing elevation, slope, latitude, and longitude, which are useful in large area mapping studies (Pflugmacher *et al* 2019, Rufin *et al* 2019). Combining the STM, texture, and contrast metrics with these additional layers yielded 119 input features for each year. We further calculated the annual availability of clear-sky observation for the entire season, April through October, which we used in the post-processing of the image time series (Figure S 1).

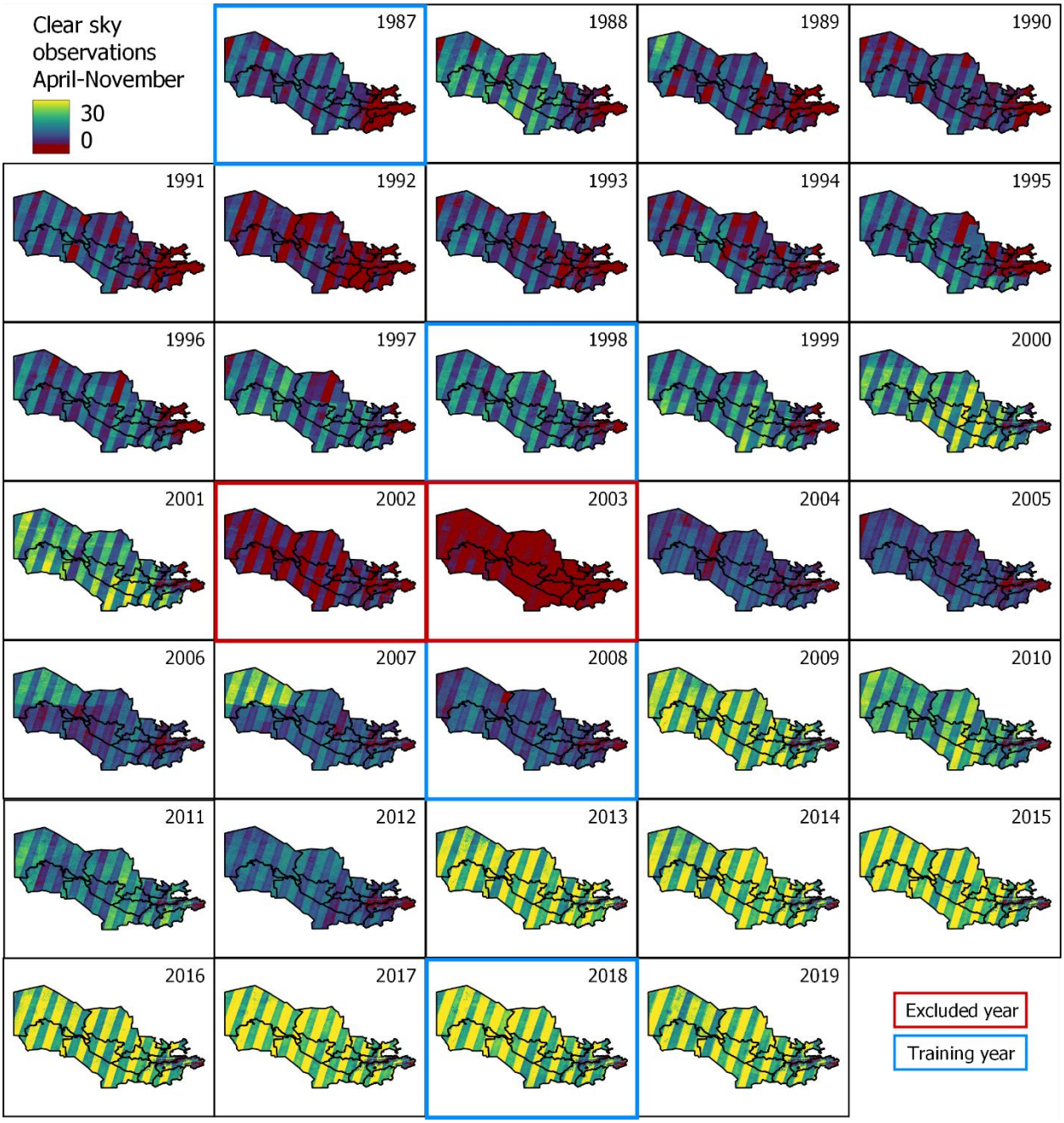

Figure S 1: Pixel-wise sum of clear sky observations for April through October in each year in the study period. Individual years used for training the classification model highlighted in blue, the years excluded from the analyses due to insufficient data availability highlighted in red. Color gradient shows pixels with less than 6 observations in red.

S2: Training data

We defined the target class catalog to capture the most prominent cropping practices (Table 1 in the main text). We collected training data for 1987, 1996, 2008, and 2018. These years covered

different Landsat sensor constellations and acquisition plans. A team of eight trained and experienced interpreters collected training polygons based on intra-annual time series graphs of various spectral indices (soil-adjusted vegetation index [SAVI], EVI, tasseled cap components [TC], and normalized difference vegetation index [NDVI]), custom visualizations of multiple seasonal STMs, and Google Earth VHR images, where available. As such, the data used to label the training dataset were different from the features used for classification (described above).

In addition, we included field data for the years 2008 and 2018 on crop types from previous field campaigns (Remelgado *et al* 2020). Training areas were digitized as polygons, from which we sampled proportionally to the polygon area, with a maximum of five points per polygon for the final training dataset. The non-cropland class was trained and predicted in the subclasses of bare land, urban environment, vegetated land (including woody vegetation and unmanaged herbaceous vegetation), and water, which were aggregated after prediction. Similarly, the model was trained to separate fodder crops and double cropping, which were merged into a single class for further processing. Combined, all training data yielded a training database with 31,718 points.

S3: Accuracy assessment and area estimation

We assessed map accuracy based on current state-of-the-art principles. We created a stratified random sample based on the classification maps and ancillary datasets. In our analyses, we considered only areas below an elevation of 2,000 meters. We created province-level substrata for all 14 provinces within the study area and calculated class occurrence frequency across time to 1) mask areas classified as bare or urban in all years and 2) produce substrata of areas dominantly classified as wet season crop, dry season crop, other vegetation, unvegetated land, or water in most years. Due to the rarity of double cropping and fodder crops, we created sampling strata

for these rare classes with an occurrence above three years. The procedure resulted in 98 strata (14 provinces and 7 classes), two of which could not be identified because the combination of province and dominant class did not occur (wet season crops in Tadzhikistan Territories, and dry season crops in Khorezm, Uzbekistan). Consequently, we created 96 strata and sampled up to 40 points (some strata were small and the number of sample points thus lower) per stratum. Note that despite the lack of the respective stratum, we identified 35 samples corresponding to wet season cropping in Tadzhikistan Territories, and 153 samples corresponding to dry season cropping in Khorezm.

Overall, our sample contained 2,784 points, for which the interpreter team assessed class labels across all 33 years using the GEE Timeseries Explorer plugin in QGIS (Rufin *et al* 2021b). Overall, 11.9% (n = 10,968) of all labels (n = 91,872) could not be determined, mostly due to data scarcity in parts of the study region in the years 1987 (31.3%) and 2003 (27.2%). We used the remaining reference samples (n = 80,904) to calculate area-adjusted overall and class-wise user´s and producer´s accuracies and derived class-wise and province-level area estimates with confidence intervals for every year.

S4: Classification and post-processing

We trained a global random forest (RF) classification model (Breiman 2001) with the training data and the respective input features (seasonal STM, texture, contrast, latitude, longitude, elevation, and slope, as described in section S1). We used 250 trees and considered the square root of the number of features (n = 11) at each split. The RF model was used to predict land cover and cropping practices across the study region and for each year, yielding a set of 33 annual maps.

We masked all pixels with fewer than six clear-sky observations between April and October in each year to exclude misclassifications due to limited data availability. To maintain a focus on the croplands in irrigated lowlands, we restricted our analyses to regions that were classified, for at least 5% of the years with sufficient clear-sky observations, as dry season cropping, double cropping, or fodder crops. We thereby excluded highly variable rain-fed agriculture, mostly present in the hillslopes of the southeastern parts of the Amu Darya basin. For the remaining pixels, we generated layers representing the percentage of a pixel being labeled as 1) cropland, 2) dry season cropping, and 3) double cropping, by normalizing the occurrence of each class at the pixel level in relation to the number of years with a valid classification i.e., excluding years with low data availability. The resulting data revealed spatial patterns of land use history at the 30-meter level provided by Landsat. For comparing differences over time, we created these layers for the periods 1987 to 2000 and 2001 to 2019, respectively.

S5: Land use change trajectories and processes

We derived trends in three land use indicators for grid cells of 3 km by 3 km. For each grid cell and year, we calculated 1) cropland extent as the share of cropland per grid cell, 2) dry season cropping fraction as the share of dry season crops, double cropping, and fodder crops within all cropland, and 3) cropping frequency as the average number of harvests per year, where double cropping and fodder crops were counted as two harvests per year. We manually removed misclassifications of wetlands, lakes, and reservoir shores. We also removed grid cells with less than 5% cropland from the analysis.

Due to the aggressive masking of regions with low data availability, the time series of our annual land use indicators at the 3x3km level had gaps. We filled the gaps by linear interpolation with a

maximum gap size of five years, using the imputeTS (Moritz and Bartz-Beielstein 2017) package in R. Based on the aggregated metrics, we inspected pixel-level time series and assessed the distribution of the metrics across the study region and per country for every year to observe changes over time. We then used autoregressive (AR) models to estimate linear time trends in the gap-filled time series using the AR function implemented in the remotePARTS package for R (Ives *et al* 2021). While the resulting trend estimates indicate the spatial patterns of overall change directions, we want to highlight that the trend estimates should be interpreted with caution, because of the nonlinearities and multiple change processes inherent to the data at hand.

We combined expert knowledge with the time series and trend estimates to investigate characteristic trajectories of change in the study region. We did this by visually inspecting the time series of land use metrics and the trend estimates. We then examined the observed changes in detail and provided a qualitative explanation of the underlying causes from the available literature and expert knowledge.

S6: Validation, accuracy assessment, and robustness checks

The maps including the classes of wet season crop, dry season crop, double and fodder crop, and non-cropland had acceptable error rates, with a median overall accuracy of 91.4% across all years (Figure S 2). In our case, the area-adjusted overall accuracy is strongly influenced by the non-cropland class which has high accuracies and covers a large extent of the study region. Class-specific accuracies should be considered instead. The separation between cropland and non-cropland classes was accurate with a median F1 score of 96.3% for the non-cropland class. In line with previous studies, the detailed discrimination of cropping practices suffered higher error rates, with median F1 scores of 70.2%, 73.8%, and 32.3% for wet season cropping, dry season

cropping, and double cropping, respectively (Figure S 3). Classification results for 2003 and 2013 were less reliable, due to the highly constrained availability of data, and were subsequently removed from the analyses.

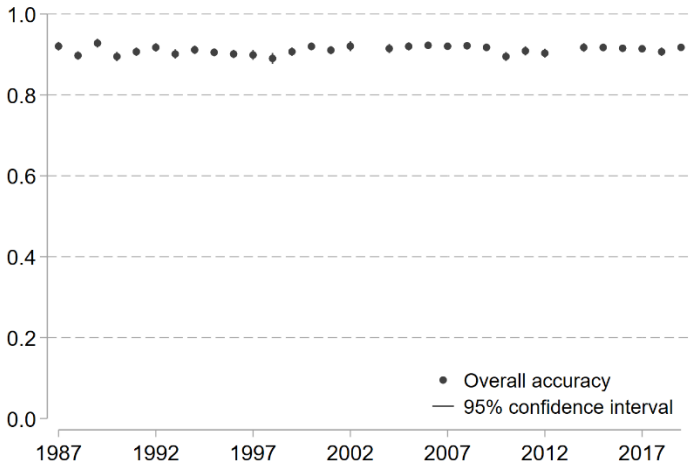

Figure S 2: Area-adjusted overall accuracy with 95% confidence interval. The years 2003 and 2013 were removed from the analyses due to data gaps.

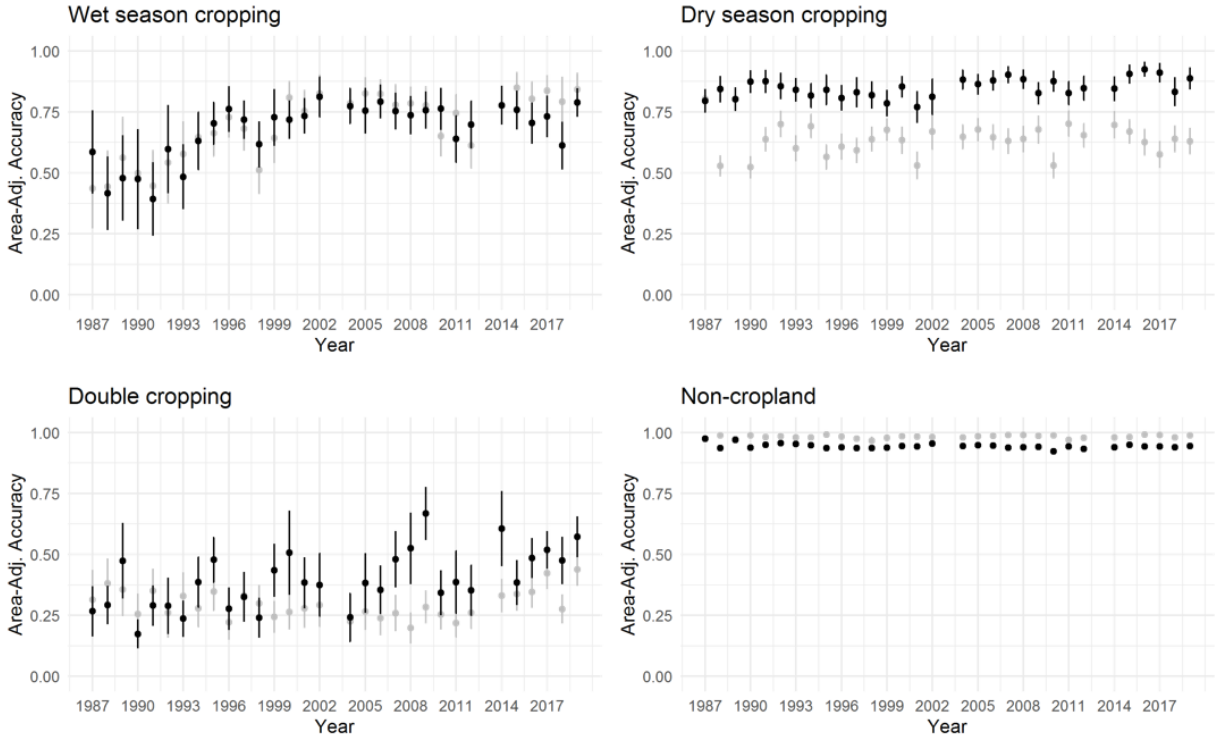

Figure S 3: Area-adjusted user (black) and producer (gray) accuracy with 95% confidence interval. The years 2003 and 2013 were removed from the analyses due to data gaps.

We created a temporally averaged confusion matrix populated with class probabilities (Table S 1), with columns representing the reference labels and rows representing the map labels. The dominant error types in our maps were the omission of dry season cropping (mapped as non-cropland or double cropping), and the omission of double cropping (mapped as non-cropland).

Table S 1: Confusion matrix populated with probabilities averaged across the study period. Values in brackets represent standard deviations.

		Reference			
		Wet Season Cr.	Dry Season Cr.	Double Cr.	Non-Cr.
Map	Wet Season Cr.	0.024 (0.011)	0.003 (0.001)	0.003 (0.001)	0.004 (0.002)
	Dry Season Cr.	0.002 (0.001)	0.070 (0.015)	0.004 (0.002)	0.006 (0.004)
	Double Cr.	0.001 (0.001)	0.011 (0.006)	0.010 (0.004)	0.004 (0.003)
	Non-Cr.	0.006 (0.003)	0.025 (0.006)	0.016 (0.005)	0.807 (0.016)

Using the validation sample, we compared the spectral-temporal signatures of the mapped classes (Figure S 4), to assess the consistency of the class behaviour between training years and the remaining years, and in relation to the seasonal windows chosen for the calculation of STMs.

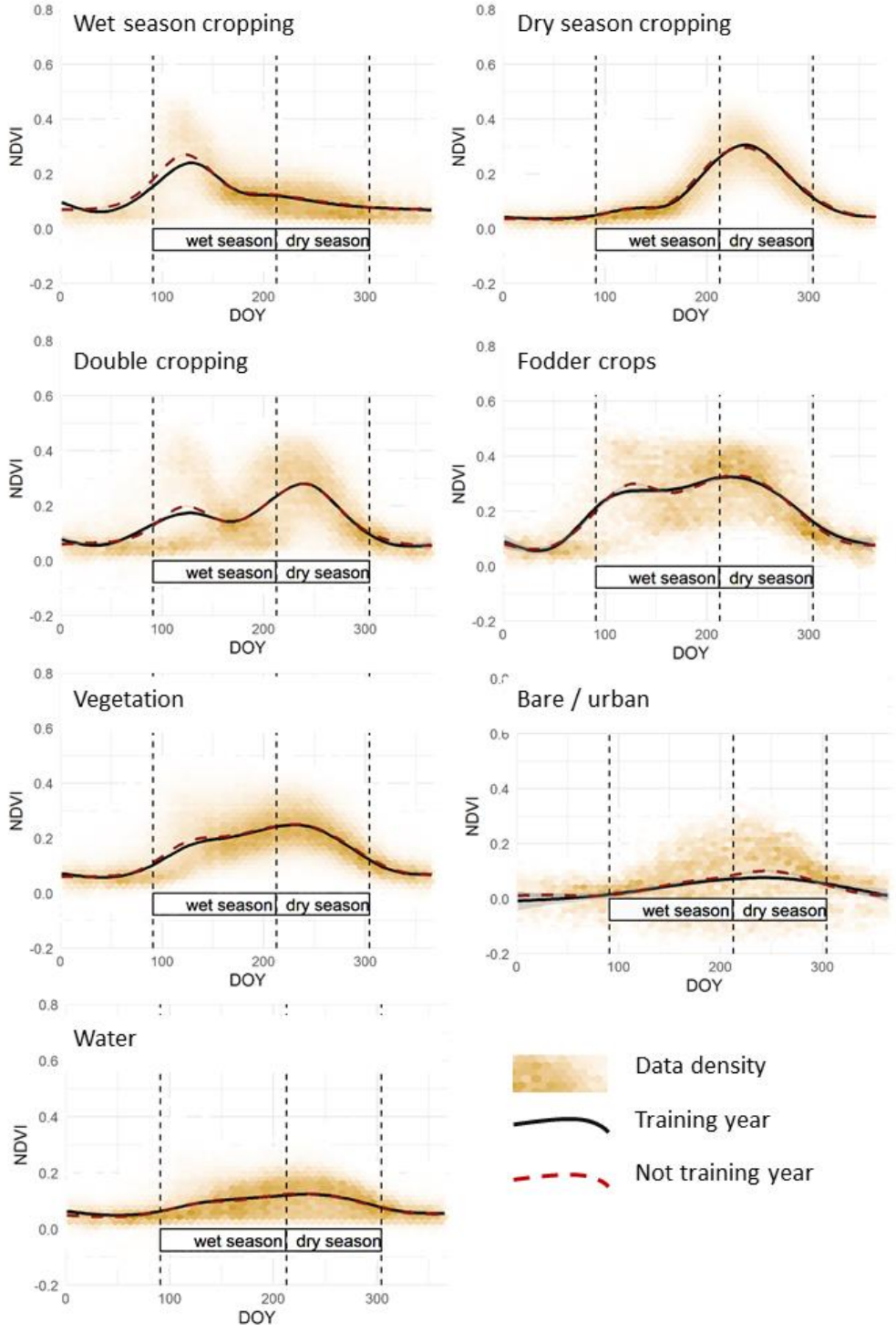

Figure S 4: Spectral-temporal signatures of the seven classes considered in the initial classification. Smoothed curves represent data obtained for years included in training (black) and those outside of the training years (red dashed). Data density represents the number of clear sky observations per hexagon.

S7: Cropland and harvested area estimates

We used the class-wise area estimates to calculate the overall cropland extent and harvested area (Figure S 5, Figure S 6). Cropland extent was calculated as the sum of the area under wet season cropping, dry season cropping, and double cropping. The harvested area included wet season cropping, dry season cropping, and two times the area of double cropping, assuming two harvests per year.

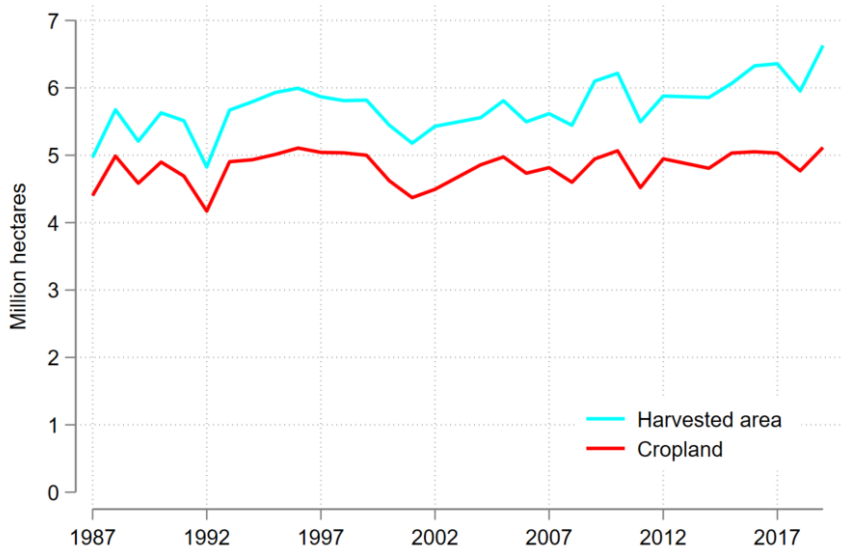

Figure S 5: Area estimates of cropland extent and harvested area per year.

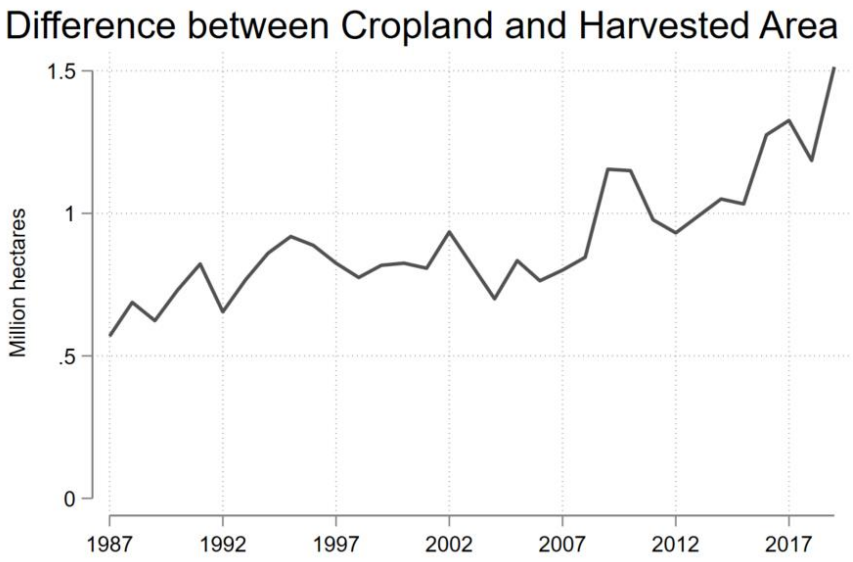

Figure S 6: Difference between cropland and harvested area, derived from area estimates.

Literature

Abdullaev I, Giordano M and Rasulov A 2007 *Cotton in Uzbekistan: Water and Welfare: The cotton sector in Central Asia economic policy and development challenges.* (London)

AghaKouchak A, Norouzi H, Madani K, Mirchi A, Azarderakhsh M, Nazemi A, Nasrollahi N, Farahmand A, Mehran A and Hasanzadeh E 2015 Aral Sea syndrome desiccates Lake Urmia: Call for action *Journal of Great Lakes Research* **41** 307–11

Aldaya M M, Muñoz G and Hoekstra A Y 2010 *Water footprint of cotton, wheat and rice production in Central Asia (Research Report Series No. 41)* (UNESCO-IHE)

Bann C, Rakhmon S, Boziev L and Rakhmanova D 2012 The Economics of Land Degradation for the Agriculture Sector in Tajikistan - A Scoping Study *UNDP-UNEP, Dushanbe*

Bégué A, Arvor D, Bellon B, Betbeder J, Abelleyra D de, P. D. Ferraz R, Lebourgeois V, Lelong C, Simões M and R. Verón S 2018 Remote Sensing and Cropping Practices: A Review *Remote Sensing* **10** 99

Beurs K M de, Henebry G M, Owsley B C and Sokolik I 2015 Using multiple remote sensing perspectives to identify and attribute land surface dynamics in Central Asia 2001–2013 *Remote Sensing of Environment* **170** 48–61

Blickensdörfer L, Schwieder M, Pflugmacher D, Nendel C, Erasmi S and Hostert P 2022 Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany *Remote Sensing of Environment* **269** 112831

Breiman L 2001 Random Forests *Machine Learning* **45** 5–32

Conrad C, Colditz R R, Dech S, Klein D and Vlek P L G 2011 Temporal segmentation of MODIS time series for improving crop classification in Central Asian irrigation systems *International Journal of Remote Sensing* **32** 8763–78

Conrad C, Dech S, Dubovyk O, Fritsch S, Klein D, Löw F, Schorcht G and Zeidler J 2014 Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of Uzbekistan using multitemporal RapidEye images *Computers and Electronics in Agriculture* **103** 63–74

Conrad C, Schönbrodt-Stitt S, Löw F, Sorokin D and Paeth H 2016 Cropping Intensity in the Aral Sea Basin and Its Dependency from the Runoff Formation 2000–2012 *Remote Sensing* **8** 630

Conrad C, Usman M, Morper-Busch L and Schönbrodt-Stitt S 2020 Remote sensing-based assessments of land use, soil and vegetation status, crop production and water use in irrigation systems of the Aral Sea Basin. A review *Water Security* **11** 100078

d'Andrimont R, Verhegghen A, Lemoine G, Kempeneers P, Meroni M and van der Velde M 2021 From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations *Remote Sensing of Environment* **266** 112708

Dara A, Baumann M, Kuemmerle T, Pflugmacher D, Rabe A, Griffiths P, Hölzel N, Kamp J, Freitag M and Hostert P 2018 Mapping the timing of cropland abandonment and recultivation in northern Kazakhstan using annual Landsat time series *Remote Sensing of Environment* **213** 49–60

Deng H and Chen Y 2017 Influences of recent climate change and human activities on water storage variations in Central Asia *Journal of Hydrology* **544** 46–57

Edlinger J, Conrad C, Lamers P A J, Khasankhanova G and Koellner T 2012 Reconstructing the Spatio-Temporal Development of Irrigation Systems in Uzbekistan Using Landsat Time Series *Remote Sensing* **4**

FAO 2012 *AQUASTAT Irrigation in Central Asia Survey* (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS)

FAO 2020 FAOSTAT data *http://faostat.fao.org*

Fick S E and Hijmans R J 2017 WorldClim 2: new 1 - km spatial resolution climate surfaces for global land areas *Int. J. Climatol* **37** 4302–15

Jiang L, Jiapaer G, Bao A, Li Y, Guo H, Zheng G, Chen T and Maeyer P de 2019 Assessing land degradation and quantifying its drivers in the Amudarya River delta *Ecological Indicators* **107** 105595

Kariyeva J and van Leeuwen W JD 2012 Phenological dynamics of irrigated and natural drylands in Central Asia before and after the USSR collapse *Agriculture, Ecosystems & Environment* **162** 77–89

Khasanov S *et al* 2022 Evaluation of the perennial spatio-temporal changes in the groundwater level and mineralization, and soil salinity in irrigated lands of arid zone: as an example of Syrdarya Province, Uzbekistan *Agricultural Water Management* **263** 107444

Khaydar D, Chen X, Huang Y, Ilkhom M, Liu T, Friday O, Farkhod A, Khusen G and Gulkaiyr O 2021 Investigation of crop evapotranspiration and irrigation water requirement in the lower Amu Darya River Basin, Central Asia *J. Arid Land* **13** 23–39

Kulmatov R, Khasanov S, Odilov S and Li F 2021 Assessment of the Space-Time Dynamics of Soil Salinity in Irrigated Areas Under Climate Change: a Case Study in Sirdarya Province, Uzbekistan *Water Air Soil Pollut* **232** 8

Lerman Z, Sedik D, Yusupov Y, Stanchin I and Kazakevich I 2016 *Wheat production and regional food security in CIS: The case of Belarus, Turkmenistan, and Uzbekistan (FAO Policy Studies on Rural Transition* No. 2016-1*)* (FAO)

Lombardozzi L and Djanibekov N 2021 Can self-sufficiency policy improve food security? An inter-temporal assessment of the wheat value-chain in Uzbekistan *Eurasian Geography and Economics* **62** 1–20

Löw F, Biradar C, Dubovyk O, Fliemann E, Akramkhanov A, Narvaez Vallejo A and Waldner F 2018 Regional-scale monitoring of cropland intensity and productivity with multi-source satellite image time series *GIScience & Remote Sensing* **55** 539–67

Micklin P 2010 The past, present, and future Aral Sea *Lakes & Reservoirs: Research & Management* **15** 193–213

Micklin P 2014 Irrigation in the Aral Sea Basin *The Aral Sea* ed P Micklin *et al* (Berlin, Heidelberg: Springer Berlin Heidelberg) pp 207–32

Olofsson P, Foody G M, Herold M, Stehman S V, Woodcock C E and Wulder M A 2014 Good practices for estimating area and assessing accuracy of land change *Remote Sensing of Environment* **148** 42–57

Pokhrel Y *et al* 2021 Global terrestrial water storage and drought severity under climate change *Nat. Clim. Chang.* **11** 226–33

Qadir M, Noble A D, Qureshi A S, Gupta R K, Yuldashev T and Karimov A 2009 Salt-induced land and water degradation in the Aral Sea basin: A challenge to sustainable agriculture in Central Asia *Natural Resources Forum* **33** 134–49

Remelgado R, Zaitov S, Kenjabaev S, Stulina G, Sultanov M, Ibrakhimov M, Akhmedov M, Dukhovny V and Conrad C 2020 A crop type dataset for consistent land cover classification in Central Asia *Scientific data* **7** 250

Reyer C PO *et al* 2017 Climate change impacts in Central Asia and their implications for development *Reg Environ Change* **17** 1639–50

Rufin P, Frantz D, Ernst S, Rabe A, Griffiths P, Özdoğan M and Hostert P 2019 Mapping Cropping Practices on a National Scale Using Intra-Annual Landsat Time Series Binning *Remote Sensing* **11** 232

Rufin P, Müller D, Schwieder M, Pflugmacher D and Hostert P 2021a Landsat time series reveal simultaneous expansion and intensification of irrigated dry season cropping in Southeastern Turkey *Journal of Land Use Science* **16** 94–110

Rufin P, Rabe A, Nill L and Hostert P 2021b GEE TIMESERIES EXPLORER FOR QGIS – INSTANT ACCESS TO PETABYTES OF EARTH OBSERVATION DATA *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **XLVI-4/W2-2021** 155–8

Siebert S, Kummu M, Porkka M, Döll P, Ramankutty N and Scanlon B R 2014 A global dataset of the extent of irrigated land from 1900 to 2005 *Hydrology and Earth System Sciences Discussions* **11** 13207–58

Singh A 2015 Land and water management planning for increasing farm income in irrigated dry areas *Land Use Policy* **42** 244–50

Stehman S V 2014 Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes *International Journal of Remote Sensing* **35** 4923–39

Stone R 2008 Engineering. A new great lake--or dead sea? *Science* **320** 1002–5

UN 2013 *Environmental Performance Reviews: Turkmenistan (ECE Environmental Performance Reviews Series* no.35*)* (s.l.: United Nations)

Uzcomstat 2020 *Data from the State Committee of the Republic of Uzbekistan on Statistics: District level crop production and harvest area dataset* (unpublished: Uzcomstat)

Vörösmarty C J *et al* 2010 Global threats to human water security and river biodiversity *Nature* **467** 555–61

Wang J *et al* 2018 Recent global decline in endorheic basin water storages *Nature geoscience* **11** 926–32

White C J, Tanton T W and Rycroft D W 2014 The Impact of Climate Change on the Water Resources of the Amu Darya Basin in Central Asia *Water Resources Management* **28** 5267–81

Wine M L and Laronne J B 2020 In Water‐Limited Landscapes, an Anthropocene Exchange: Trading Lakes for Irrigated Agriculture *Earth's Future* **8** 1

World Bank 2020 *Costs of Environmental Degradation in the Mountains of Tajikistan* (World Bank)

World Bank 2021 *World Development Indicators* (Washington, D.C.: The World Bank)