Global mapping of soil salinity change

Ivushkin, K., Bartholomeus, H., Bregt, A. K., Pulatov, A., Kempen, B., & De Sousa, L.

Global mapping of soil salinity change

Konstantin Ivushkin[1], Harm Bartholomeus[1], Arnold K. Bregt[1], Alim Pulatov[2], Bas Kempen[3], Luis de Sousa[3]

[1]Laboratory of Geo-Information Science and Remote Sensing, Wageningen University & Research, Droevendaalsesteeg 3, 6708 PB Wageningen, the Netherlands

konstantin.ivushkin@wur.nl, harm.bartholomeus@wur.nl, arnold.bregt@wur.nl

[2]EcoGIS Center, Tashkent Institute of Irrigation and Agricultural Mechanization Engineers, Qari Niyoziy 39, 100000 Tashkent, Uzbekistan

alimpulatov@mail.ru

[3] ISRIC — World Soil Information, Droevendaalsesteeg 3, 6708 PB Wageningen, the Netherlands

bas.kempen@wur.nl, luis.desousa@wur.nl

Abstract

Soil salinity increase is a serious and global threat to agricultural production. The only database that currently provides soil salinity data with global coverage is the Harmonized World Soil Database, but it has several limitations when it comes to soil salinity assessment. Therefore, a new assessment is required. We hypothesized that combining soil properties maps with thermal infrared imagery and a large set of field observations within a machine learning framework will yield a global soil salinity map. The thermal infrared imagery acts as a dynamic variable and allows us to characterize the soil salinity change. For this purpose we used Google Earth Engine computational environment. The random forest classifier was trained using 7 soil properties maps, thermal infrared imagery and the ECe point data from the WoSIS database. In total, six maps were produced for 1986, 2000, 2002, 2005, 2009, 2016. The validation accuracy of the resulting maps was in the range of 67-70%. The total area of salt affected lands by our assessment is around 1 billion hectares, with a clear increasing trend. Comparison with 3 studies investigating local trends of soil salinity change showed that our assessment was in correspondence with 2 of these studies. The global map of soil salinity change between 1986 and 2016 was produced to characterize the spatial distribution of the change. We conclude that combining soil properties maps and thermal infrared imagery allows mapping of soil salinity development in space and time on a global scale.

Keywords: soil salinization, Google Earth Engine, Landsat, SoilGrids

1. Introduction

Soil salinity increase is a serious and global threat to agricultural production. It affects an area of more than 1 billion hectares in more than 100 countries all over the world and these numbers are constantly growing (Abbas et al., 2013; FAO and ITPS, 2015; Squires and Glenn, 2004; Szabolcs, 1989). Besides this estimate of the affected area globally, several others exist, which sometimes quite dramatically differ in the extent of the affected area (IAEA, 1995; Oldeman et al., 1991). Therefore, only a rough approximation of salt affected area globally can be given. FAO (2018) recognises this issue and stresses that the divergence of current estimations of the extent of salt affected areas are quite often the result of differences in methods for collecting and aggregating statistics. They specifically state that there is a need for data on the rate of change in areas affected by salinization at regional and global level (FAO, 2018). Status of the World’s Soil Resources report by FAO and ITPS (2015) also mentions that information on the extent and characteristics of salt-affected soils is very scattered.

The only database that currently provides soil salinity data with global coverage is the Harmonized World Soil Database. This database is an important source of soil data for global studies, but it has several limitations when it comes to soil salinity assessment. First, the database consists of soil mapping units, rather than a continuous grid with soil properties’ values unique for each pixel. It has over 15,000 mapping units and have only a single soil salinity value per unit, while some of these units are stretching for hundreds of kilometres. Although the spatial resolution of the maps produced from this database is around 1 km, the actual spatial resolution is much coarser. Second, though the database was updated several times in the past (last time in 2012; version 1.2), most of it relies on the FAO/UNESCO Soil Map of the World created in 1970-1981, which can be considered outdated given the highly dynamic nature of soil salinity. Lack of spatial detail and outdated data illustrate the need for an updated global soil salinity map.

Having up to date information on spatial distribution and severity of soil salinity is crucial for agricultural management of affected areas. It allows to take necessary measures to reduce, or even avoid, economical losses and restore the productivity of the soil. Mapping dynamic soil properties like salinity has challenges compared with other, less dynamic properties. Soil salinity can rapidly change after irrigation or a rainfall event. Drought, on the other hand, might increase salinity in the course of several weeks. Therefore, monitoring by traditional methods will require sampling frequently in time, which can be cost-prohibitive. That is one of the reasons why remote sensing methods are now used more and more often for soil salinity monitoring and mapping (Allbed and Kumar, 2013; Hasanlou and Eftekhari, 2019).

Remote sensing is used for soil salinity mapping already for years (Metternicht and Zinck, 2009). Nevertheless, there are still no universally acceptable methods to derive soil salinity parameters from remote sensing data that can be used for different environments. On field and local scales many studies proposed conversion models from remote sensing variables to soil salinity levels on the ground. Nevertheless, these models usually do not demonstrate the same high accuracy in different parts of the world (Allbed et al., 2014a; Allbed et al., 2014b; Douaoui et al., 2006), which means that scaling up to a global scale is problematic.

Recently, thermal infrared imagery were used to distinguish between different levels of soil salinity on agricultural lands (Ivushkin et al., 2017; Ivushkin et al., 2018). The principle behind this approach is that canopy temperature of the plants grown in affected areas will be higher than of plants growing in non-affected areas (Gómez-Bellot et al., 2015; Urrestarazu, 2013). The approach was tested on regional and local scales and showed its robustness in different climatic conditions and on areas covered with different crops. Therefore, it seems promising for use on a global scale.

We foresee, however, that scaling up to a global scale will bring additional challenges like the issue of different climatic zones. The thermal approach was previously applied on areas small enough to presume constant air temperature per single image acquisition scene, therefore there was no need to normalise the values. On a global scale this will be impossible because of the different climatic zones and extreme temperature differences between regions, and use without normalisation will just lead to characterisation of climate, rather than soil salinity. But even with some kind of normalisation, using only thermal data on a global scale will be insufficient because of other factors that will influence the temperature.

Here we propose to tackle this challenge by using auxiliary data. It is known that other soil properties are correlated with soil salinity. For example, Al-Busaidi and Cookson (2003) described the interrelations of pH and soil salinity, Setia et al. (2013) studied the influence of soil salinity on the soil organic carbon content. A connection between cation exchange capacity and soil salinity has also been reported (Saidi, 2012). Moreover, bulk density and soil texture can have some auxiliary information for soil salinity monitoring. Often saline and alkaline soils are affected by compaction, and lower water retention in sandy soils will make them less prone to salinity problems. Global maps of properties relevant for soil salinity mapping are available from the SoilGrids portal[1] (Hengl et al., 2017).

[1] https://soilgrids.org

We hypothesize that combining these maps together with thermal infrared imagery and a large set of field observations on soil salinity indicators, such as electrical conductivity, within a machine learning framework we can produce a global soil salinity map. Moreover, since the SoilGrids data is static, using thermal data from different time periods will enable us to assess soil salinity change in an area of interest over time. Therefore, the overall aim of this study is to investigate if combination of soil properties maps and thermal imagery will allow us to map the development of soil salinity in space and time on a global scale and measure how accurate these estimates will be.

2. Methods and materials

Because our study was implemented on a global scale, we decided to use Google Earth Engine (GEE) as freely available platform specially tailored for analysis and processing of geodata on a global scale. Among GEE advantages are the extensive library of geospatial datasets, including widely used satellite imagery, and computational power enough to process these data on a global scale. GEE has already been used for soil properties mapping. For example for soil moisture mapping (Sazib et al., 2018) or soil type and soil organic carbon mapping (Padarian et al., 2015). Therefore, it became our platform of choice for further analysis.

2.1. Ground truth data

As ground truth we used the WoSIS Soil Profile Database (Ribeiro et al., 2015), which is maintained by ISRIC – World Soil Information and includes over 100,000 georeferenced soil profiles. For our study we selected the upper layer of soil profiles for which electrical conductivity (ECe) values are available. The thickness of this layer varied from 0-5 cm to 0-60 cm. In total, 15,188 data points were selected and used in further analysis. Figure 1 shows the spatial distribution of the data points. It shows that the distribution is unequal and depends first of all on the geographical location. For example, only few points are available for higher latitudes, which can be explained by the fact that agriculture is limited at these latitudes and therefore the demand for soil analyses. The second reason for the unequal distribution is the willingness of local data holders to make the data publically available. For example, Figure 1 shows that while there are thousands of samples available in the USA and Mexico, there are hardly any in Russia or Central Asia. Nevertheless, with all its limitations, the WoSIS database is the richest available on a global scale, and therefore used in this research.

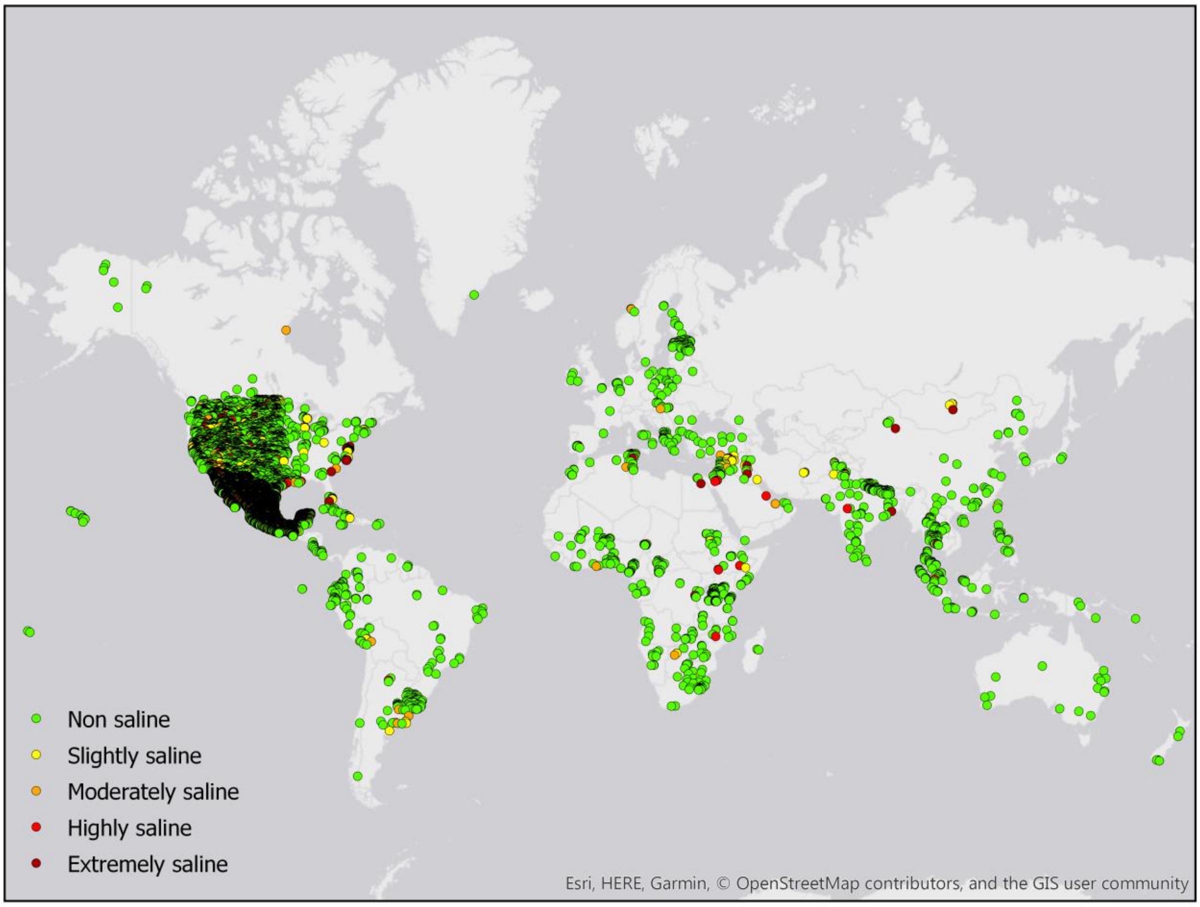


Figure 1. Distribution of ground truths sampling data

The ECe values were classified into Non saline (12,160 points), Slightly saline (2,106 points), Moderately saline (440 points) , Highly saline (232 points) and Extremely saline (250 points) classes according to widely used classification of Abrol et al. (1988) (Table 1).

Table 1. Soil salinity classification used in this paper.

Salinity class	Non-saline	Slightly	Moderately	Highly	Extremely
ECe, ds/m	<2	2-4	4-8	8-16	>16

2.2. Data processing and analysis

2.2.1. Thermal remote sensing data pre-processing

Two thermal datasets were used. The first one is the USGS Landsat 5 Surface Reflectance Tier 1 collection and second is the USGS Landsat 8 Surface Reflectance Tier 1 collection, both of which are available from the GEE data catalogue. Both collections provide orthorectified brightness temperature acquired in wavelength range from 10.4 to 12.5 micrometres. Landsat 8 data were used in this study for the year 2016, Landsat 5 data were used for 1986, 2000, 2002, 2005, 2009. For these years, mosaics from available cloud-free images in the period from March to September were averaged on per-pixel basis and used in further analysis. The mosaicking was done using capabilities of Google Earth Engine, where in the first step the whole image collection of interest was filtered based on the date of interest, using function .filterDate(), and on the second step the average of all images fulfilling the previous requirement was calculated by applying .reduce(ee.Reducer.mean()), which produced final mosaics used in the analysis.

As an input variable for our modelling we chose to work with the temperature anomaly instead of the absolute temperature to harmonise the data for the global analysis. This means that for each pixel the recorded temperature value was subtracted from the long-term temperature average for this pixel. This was done for each global layer in our thermal time series. The long term average grid was constructed from the Landsat 5 GEE dataset mentioned before, by calculating the average in the period from 1999 to 2012 from all available cloud-free images on per-pixel basis.

2.2.2. Data modelling

We used the temperature difference layers together with several SoilGrids layers. SoilGrids is a collection of global soil class and soil properties maps (Hengl et al., 2017). In our analysis we used seven grids that contain information indirectly connected with soil salinity: sand content, silt content, clay content, pH in H_2O, cation exchange capacity, bulk density, organic carbon content. These grids are available for seven depths up to two metres. Here we used the top layer (0 cm). SoilGrids were produced using a large set of covariates, including relief characteristics derived from a digital elevation model like slope, profile curvature and others that can affect development of soil salinity. That is why we consider adding these variables into our analysis as redundant and choose not to do so. The Landsat thermal images have been resampled during the processing to 250 m by a built-in functionality of GEE to correspond in resolution with the SoilGrids layers and have a common basis for modelling and prediction.

The Google Earth Engine contains several machine learning classifiers. The three most often used are Support Vector Machines, Classification and Regression Trees (CART) and Random Forest. For our study we chose random forest because trial runs of other two showed that these were not suitable for our purposes well enough. This is described in the 'Results' section in more detail. The random forest classifier was trained using the eight variables mentioned and the ECe data from the WoSIS database. The random forest algorithm constructs an ensemble of decision trees and lets them "vote" for the most probable class (Breiman, 2001; Strobl et al., 2009). We set the number of trees parameter to 50 and mtry parameter (variablesPerSplit argument in GEE) to the square root of the number of variables. This number of trees was chosen after a set of trial runs during which we established that further increase in the amount of trees does not bring any significant increase in the map validation accuracy. There are no widely accepted guidelines in the literature on the selection of the mtry parameter. Conflicting opinions led to a practice where importance and sensitivity of the mtry parameter in case of each model should be investigated by modellers. Therefore, we tried different values of this parameter and observed no significant difference in the results, except for the increase of computation time when a mtry value close to the maximum number of variables was used. Therefore, we chose the default setting which is the square root of variables used.

In total six models were trained and six maps at 250 m resolution were produced. These models differ in thermal image uses: for each model we used thermal imagery from a different year. The maps were produced for six time steps: 1986, 2000, 2002, 2005, 2009, 2016. These years were selected to correspond with other studies describing temporal changes in soil salinity with which further comparisons are made (Fan et al., 2012; Taghadosi and Hasanlou, 2017; Wang et al., 2008).

In the learning stage we used around 3500 points from the WoSIS database. They were selected by random stratified sampling, preserving the relative salinity class distribution in the ground-truth dataset. Meaning that the non-saline class will be the most abundant and the highly and extremely saline class will be less abundant in the training dataset. The final learning dataset consisted of 2,000 points of Non saline class, 1,000 of Slightly saline, 210 of Moderately saline, 105 of Highly saline and 110 of Extremely saline classes. The trained classifier was applied to the eight layers mentioned earlier to produce the final global map of soil salinity.

The map was validated by selecting randomly 100 points of each class from the WoSIS database. The 100 was selected as a maximum because of the limited amount of points in Highly and Extremely saline classes. A higher number would lead to significant overlap between learning and validation points in these classes. For the selection of validation points a different randomisation seed was used than for the learning stage. The equal amount of points for each class ensures that the final validation accuracy represents the accuracy throughout the entire range of salt affected areas. We expected that the Non-saline class will have the highest classification accuracy and using non-stratified selection of validation points will unjustly overestimate the accuracy. During the validation we compared the salinity class at the validation site with the modelled value. The same validation set was used for maps of all years by using the same seed in the random stratified sampling function. The main accuracy metrics calculated are confusion matrix, overall accuracy, user's accuracy and producer's accuracy which are provided further in the results section.

We did not select more than 3500 points because attempts to increase the number of training points lead to critical errors in most runs, and for runs where the computation did finish the increase in validation accuracy was not significant. Therefore 3500 has been selected as a number of data points for all further runs.

The computation times for most of the runs were below ten minutes, depending on the load of the servers. However, preparation of thermal mosaics took up to five hours for each time step, mainly because it required export of the mosaics into the Google Earth Engine asset, rather than proceeding with the analysis directly after computation of a mosaic.

3. Results and discussion

3.1. Global distribution of soil salinity

Figure 2 shows a global map of soil salinity classes using the thermal image of 2016. It highlights main salt affected areas in North America, Central Asia, Middle East.

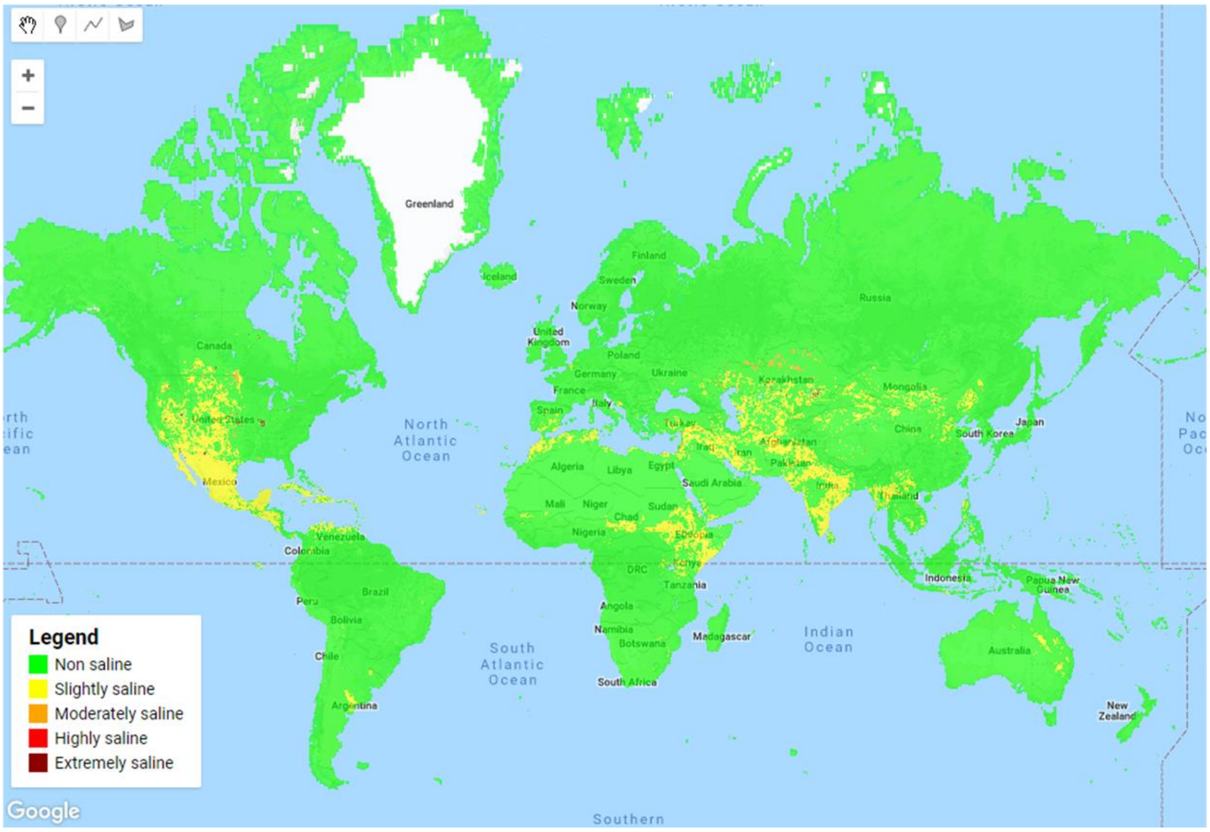


Figure 2. Resulting global soil salinity map for 2016

Global statistics of affected area for all six time steps are presented in Figure 3 and Table 2. Our analysis shows that the total area of salt affected lands increased with more than 100 Mha between 1986 and 2016, though some natural variation is present. The majority of the increase is the increase in slightly saline area. This suggests that more and more previously unaffected areas are starting to suffer from soil salinisation. This is supported by the fact that the total area of affected lands is continuing to increase. The actual area of Moderately saline lands has decreased, while Highly and Extremely saline are more volatile in time.

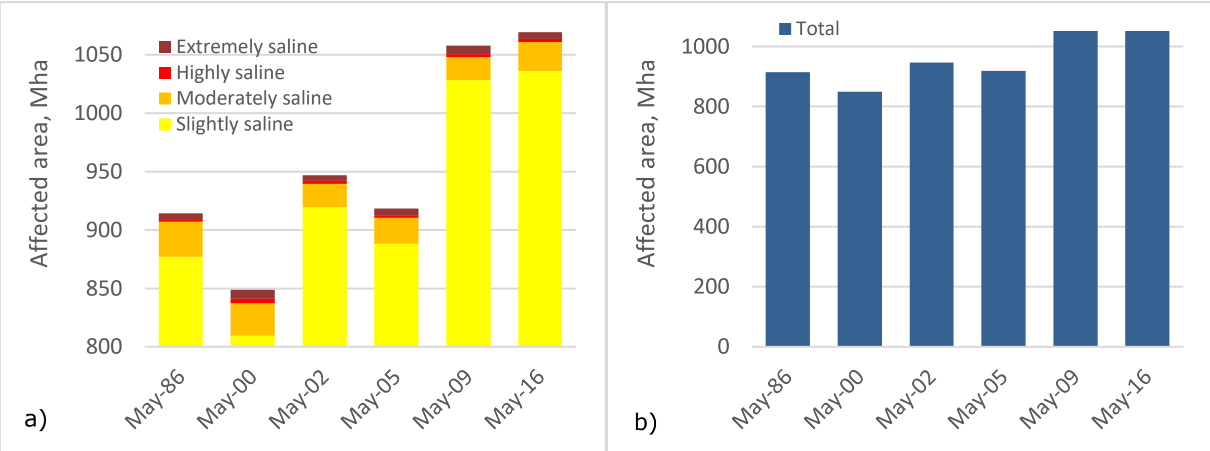


Figure 3. Salt affected land area for different years; a) shows division per salinity class (note the y-axis uses an offset), b) shows the affected area

Table 2. The world salt affected area as predicted from ground truth data, thermal satellite imagery and soil property maps for different years, Mha

	Slightly saline	Moderately saline	Highly saline	Extremely saline	Total
May/86	877.9	30.3	2.1	5.2	915.5
May/00	809.6	27.7	3.8	7.7	848.8
May/02	919.5	20.0	2.7	4.5	946.7
May/05	888.3	22.2	2.0	5.8	918.3
May/09	1,028.2	19.8	2.3	7.4	1,057.7
May/16	1,036.2	24.8	2.5	5.8	1,069.3

We found two sources referring to a global distribution of salt affected lands. Szabolcs (1989) assessed the total area of salt affected lands globally to be around 955 Mha, which is not far from our assessment of 914 Mha in 1986. The second source is the review by Squires and Glenn (2004) where the salt affected area approximately covers 1 billion hectares. We consider correspondence of other studies with our assessment quite encouraging, since 68% validation accuracy (Table 3) and unequal distribution of training and validation data might suggest bigger discrepancy with other assessments that were based more on field studies.

Another interesting observation from the global map in Figure 2 is that affected areas in Central Asia have been captured. We had almost no training points in that area (Figure 1), but the region is known to be one of the most severely affected by soil salinisation. In our opinion, this finding is supporting the principal validity of the method. However, we acknowledge that comparison with ground truth data from this area is required to further assess how well the maps produced here represent the spatial soil salinity patterns in Central Asia.

The map generally captures known hotspots in salinity-affected regions, which we further discuss towards the end of this section, but also shows overestimation of salt-affected areas. For example, the map shows that Mexico is almost completely salt affected, which is an overestimation. Szabolcs (1989) states that 1.65 Mha is the area of salt affected lands in Mexico. This number would increase to this time, but still would be far from the total area of the country. We supposed that one of the reasons for this overestimation is an underrepresentation of Non saline class data points in the samples collected in Mexico. But after scrutinizing the sample dataset this appeared not to be the case. A vast majority of these samples (77%) belong to the Non saline class, which is comparable with the distribution in the dataset for the whole world. Moreover, trial maps produced with different seed numbers still had the same overestimation for Mexico. Therefore, it is not the result of a sampling bias, but probably the specific combination of values in soil properties maps we used for prediction that lead to this overestimation. In global affected area assessments this overestimation was probably negated by some cases of underestimation, like in Australia, where only few patches of salt affected lands are shown.

The validation accuracy of this map is 68%. For different time steps classification is in a range of 67-70%, depending on the date of thermal images used. In general, highest accuracy of 70% has been achieved when thermal images of 2000-2002 were used.

Most of the classification errors appear in highly and extremely saline classes (Table 3). Those are the less common classes globally and they represented only a small fraction of WoSIS database, therefore we presume that using a larger number of highly and extremely saline training points might increase the accuracy. The influence of the amount of training points is especially visible if you compare the accuracy of two classes. The highly saline class is even less abundant in WoSIS database than Extremely saline. Therefore, less training points for the Highly saline class have been used and accuracy for it is less than for Extremely saline, though in reality Highly saline areas are more widespread than Extremely saline ones.

The important result is that when a point is misclassified, in most cases this point is still in a saline class, though maybe of a different degree, and only rarely it is assigned to the Non saline class. When only two classes are considered (saline and non-saline) producer's accuracy raises up to 89%. That suggests that the approach is quite useful in distinguishing between salt affected and non-affected lands, and only the definition of the degree of salinity remains challenging.

Table 3. Confusion matrix and accuracy statistics of 2016 map

		Predicted						
	Salinity class	Non saline	Slightly saline	Moderately saline	Highly saline	Extremely saline	Total	Producer's accuracy, %
Observed	Non saline	**90**	10	0	0	0	100	90
	Slightly saline	10	**88**	1	0	1	100	88
	Moderately saline	11	28	**61**	0	0	100	61
	Highly saline	15	34	4	**47**	0	100	47
	Extremely saline	18	29	1	0	**52**	100	52
	Total	144	189	67	47	53	500	
	User's accuracy, %	62.5	46.6	91	100	98.1		

Together with random forest algorithm we checked two other classifiers available in GEE that are based on machine learning principles. The Support Vector Machine did not produce any meaningful results in our case. Almost the whole map has been classified as non-saline area. Classification and Regression Trees (CART) algorithm showed somewhat better results, but still worse than the Random Forest algorithm. The accuracy was around 50% and the map unrealistically overestimated highly and extremely saline areas.

As we mentioned previously in this section, the global maps captured known soil salinisation hotspots. One of them is Grand Forks county on the border of North Dakota and Minnesota in the United States. It is a known salt affected area (Seelig, 2000) and it has been depicted on the map we produced (Figure 4). In Seelig (2000) this area is marked as an area of frequent inclusion in productive land, which correspond to areas marked as Moderately saline in Figure 4.

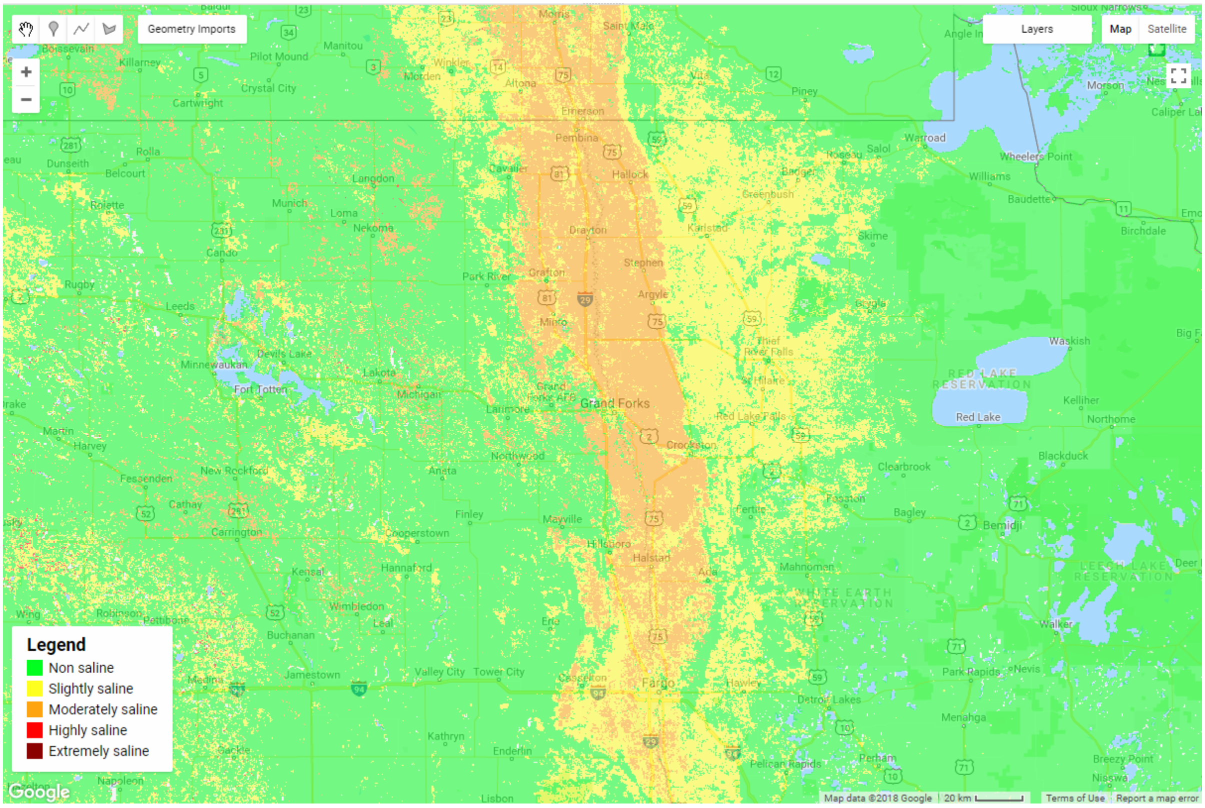


Figure 4. Soil salinity map of Grand Forks county and surroundings (2016)

One of a few countries in Europe affected by inland soil salinity problem is Hungary. Our map in Figure 5 shows some slightly affected lands, which is correct for the area where ECe values are just slightly higher than 2 ds/m (Kovács et al., 2006). Though some areas were correctly identified, the big areas in the east of Hungary have been missed. The probable cause is that most of the areas with higher salinity are grasslands and croplands with more tolerant species, therefore our method, which includes crop canopy temperature metric, did not capture those areas.

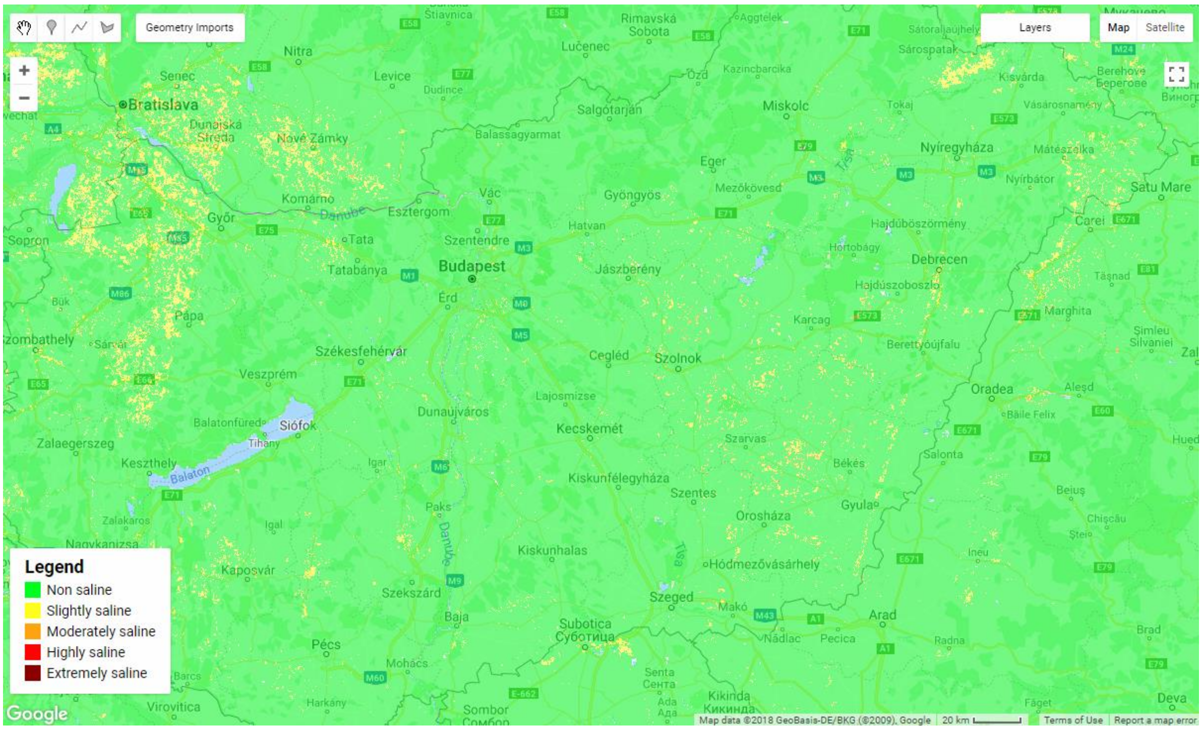


Figure 5. Hungary map of soil salinity (2016)

3.2. Local soil salinity change

To verify our hypothesis that integration of thermal infrared imagery from different periods of time will allow us to asses temporal change in soil salinity, we compared our maps with outcomes of several studies where such change is assessed.

Figure 6 shows the soil salinity map for study area in Xinjiang Province, China. According to Wang et al. (2008) this area in a period from 1983 to 2005 suffered an increase in soil salinity due to irrigation and a rise in the shallow water table. A similar increase is shown by the maps.

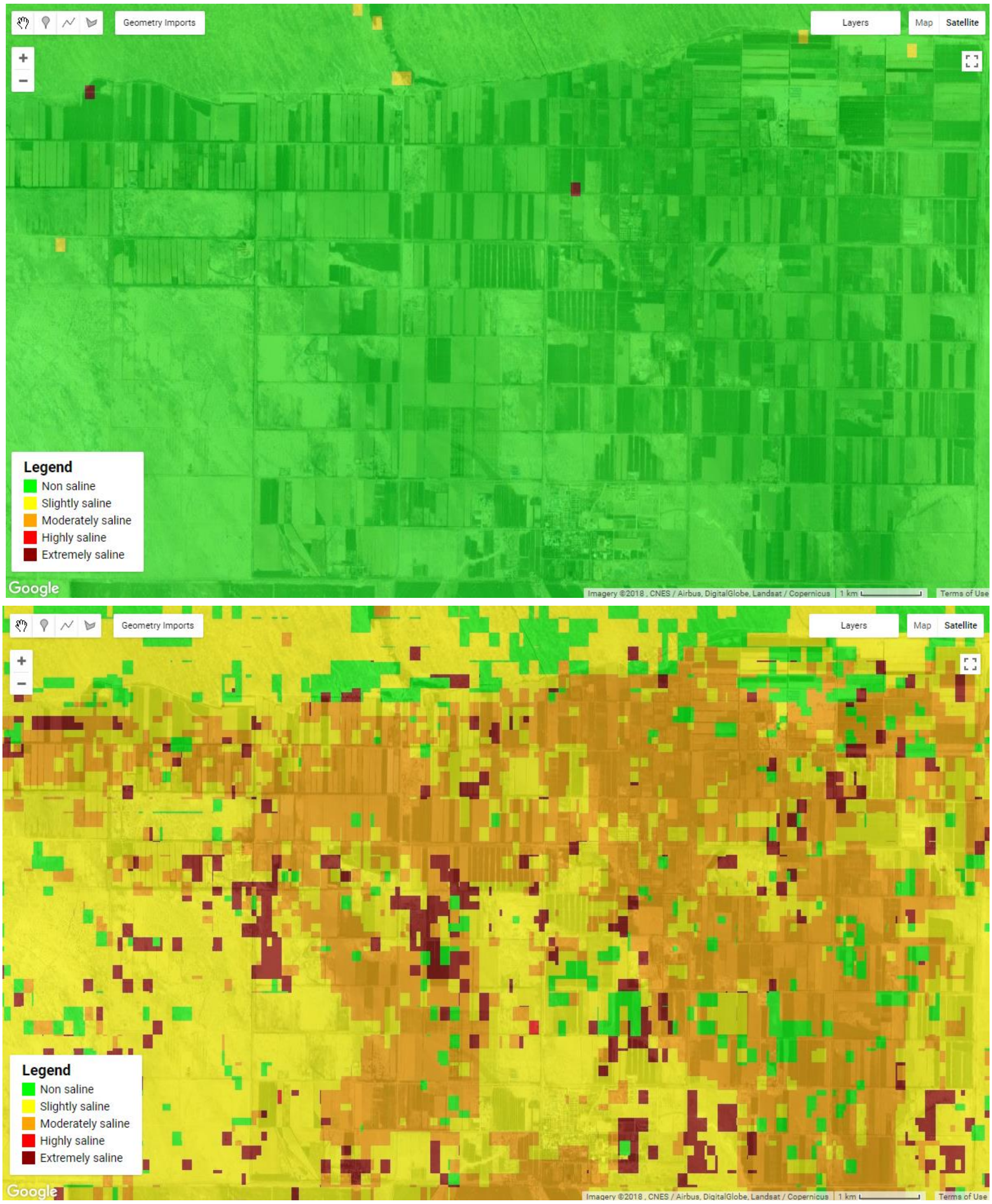


Figure 6. Soil salinity maps (upper from 1986 and lower from 2005) of the Fubei region of Xinjiang Province, China. According to Wang et al. (2008) soil salinity increased in this area.

Another area of interest we found data on is the Bakhtegan Salt Lake region in Iran. According to Taghadosi and Hasanlou (2017) more that 76% of vegetated areas of this region experienced increase in soil salinity from 2000 to 2016. This is in accordance with the maps shown in Figure 7, where the map from 2016 shows significantly more salt affected areas compared with the map from 2000.

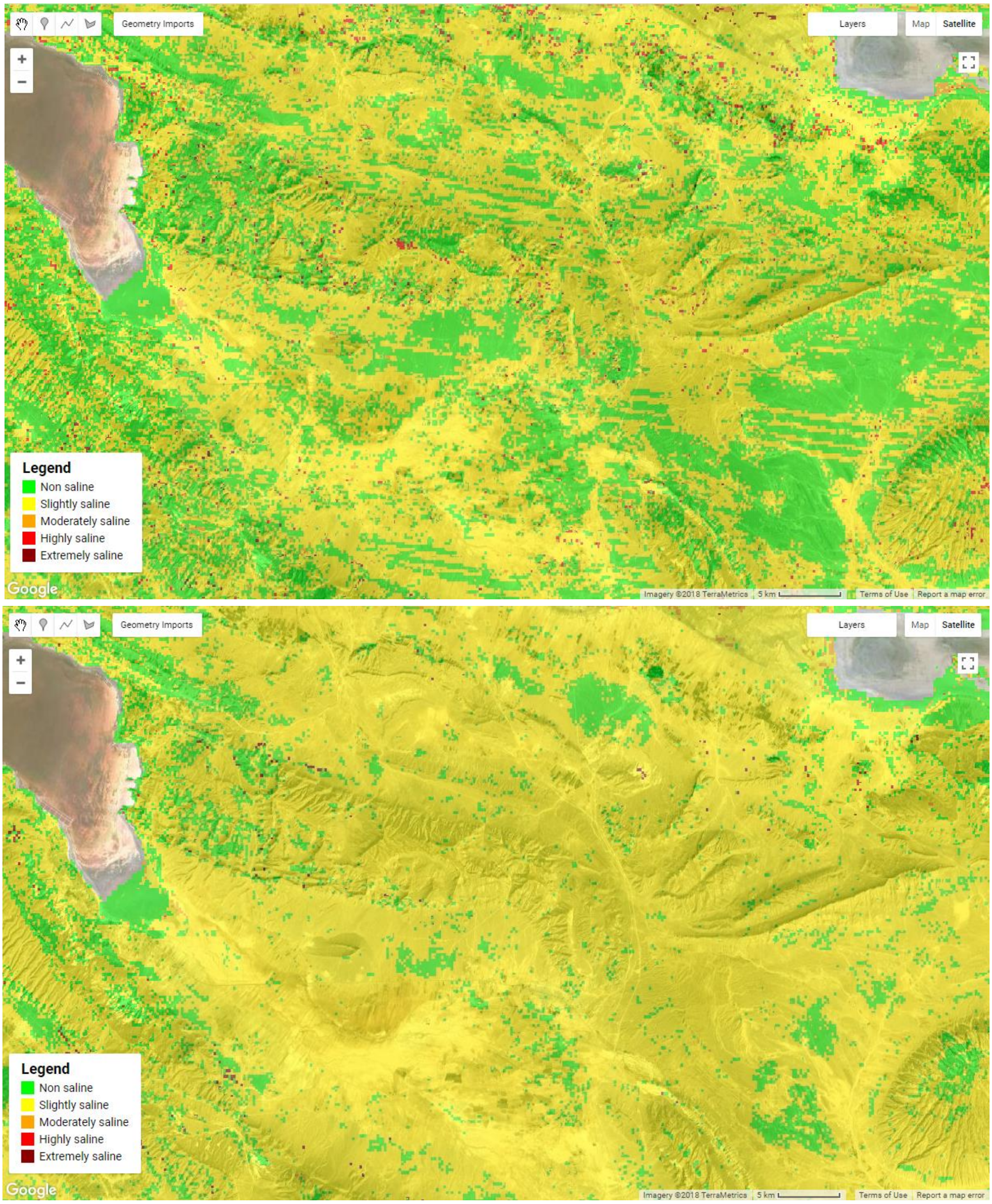
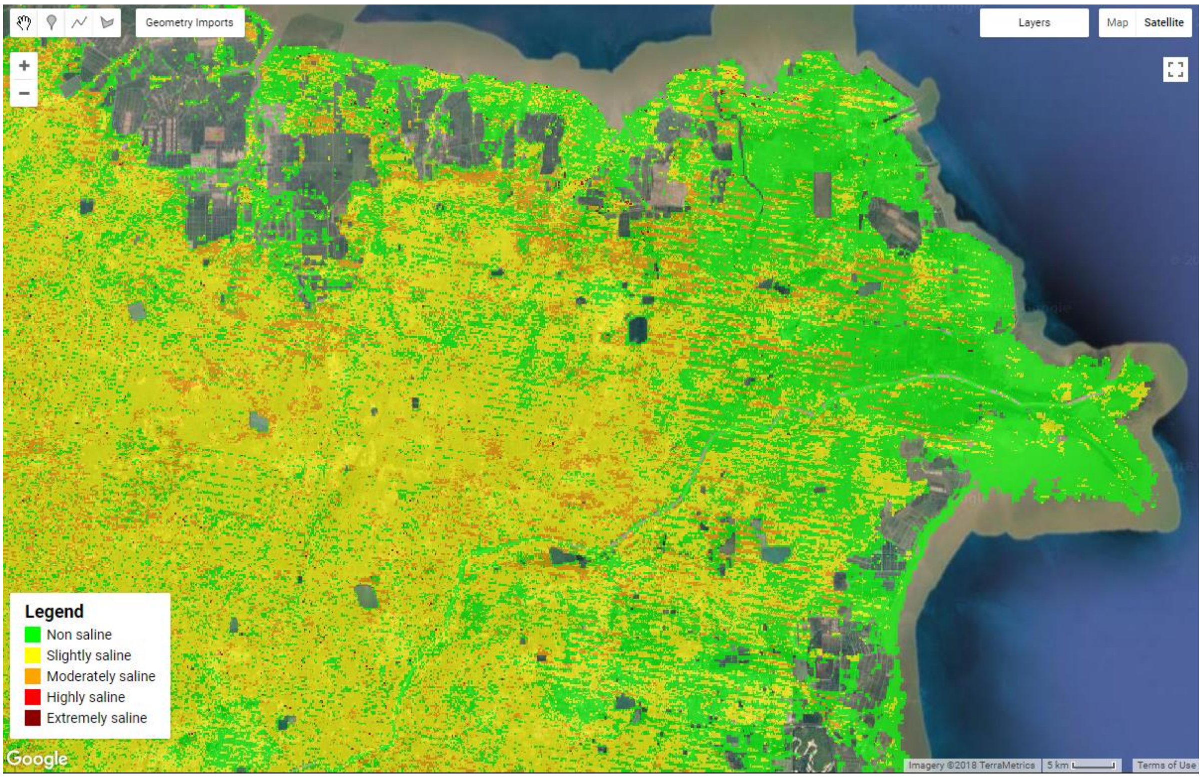


Figure 7. Soil salinity maps (upper from 2000 and lower from 2016) of the Bakhtegan Salt Lake region in Iran. According to Taghadosi and Hasanlou (2017) soil salinity increased in this area.

The next area we investigated is the Yellow river delta in China. Fan et al. (2012) have researched the dynamics of soil salinisation in the region for the period from 1985 to 2006. Their results show that while in 1985 salt affected areas were mostly located in the immediate vicinity of the river, in 2006 the majority of salt affected areas were mapped around the coast. In general, during those two decades the area suffered rapid increase in soil salinity.



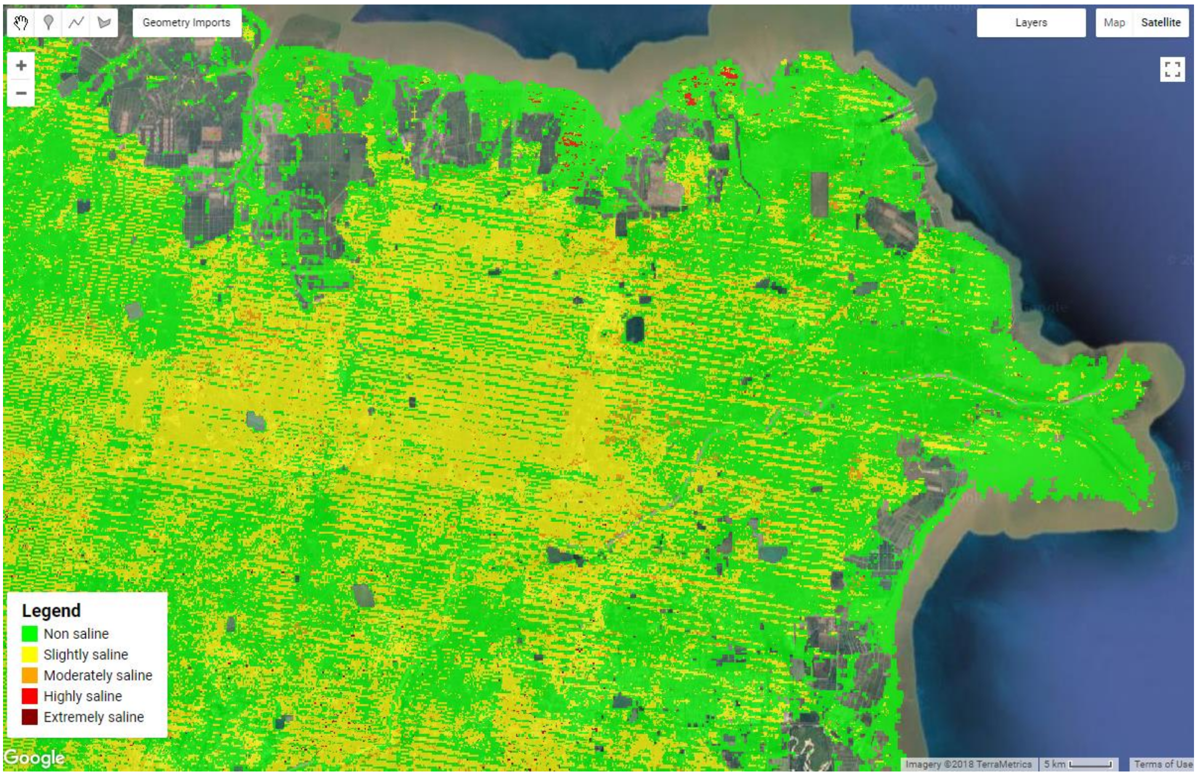


Figure 8. Soil salinity maps (upper from 1986 and lower from 2005) of the Yellow River Delta, China. According to Fan et al. (2012) the soil salinity increased in this area.

On contrary to the cases described previously, our map of the area is not in complete accordance with the reference. In Figure 8 the map from 1986 shows visibly more salt affected areas compared with the map from 2005. Nevertheless, some of the changes seem to be captured. For example, from 1986 to

2005 salt affected areas in the immediate vicinity of the Yellow river have decreased. Moreover, some highly affected spots appeared on the coastal area in the north of the 2005 map. Both of which is in accordance with the results of Fan et al. (2012).

The probable reason of the discrepancy in this result is the specifics of soil salinity development in this area. Here the main reason is the seawater intrusion, while in previous cases we looked into the problem of inland, dryland salinity. Moreover, close proximity of the sea could also influence the thermal data results.

3.3. Global changes map

To understand the spatial distribution of soil salinity change we produced a soil salinity change map (Figure 9). That is a difference map between 1986 and 2016 maps. In accordance with the statistics presented earlier (Table 2, Figure 3) the map shows mainly an increase in soil salinity. Yellow colours, representing the increase, are prevalent, while only few areas of the decrease can be seen. The majority of the salt affected areas experienced a change to a neighbouring class (i.e. from Non saline to Slightly saline or from Extremely saline to Highly saline) that is why only two colours are shown in the map. However, there are some areas of interest where more dramatic changes appeared. Those are marked by circles on the map. The area in Kazakhstan experienced an increase in soil salinity of up to 3 classes and areas in the North of the US have experienced a decrease of up to two classes.

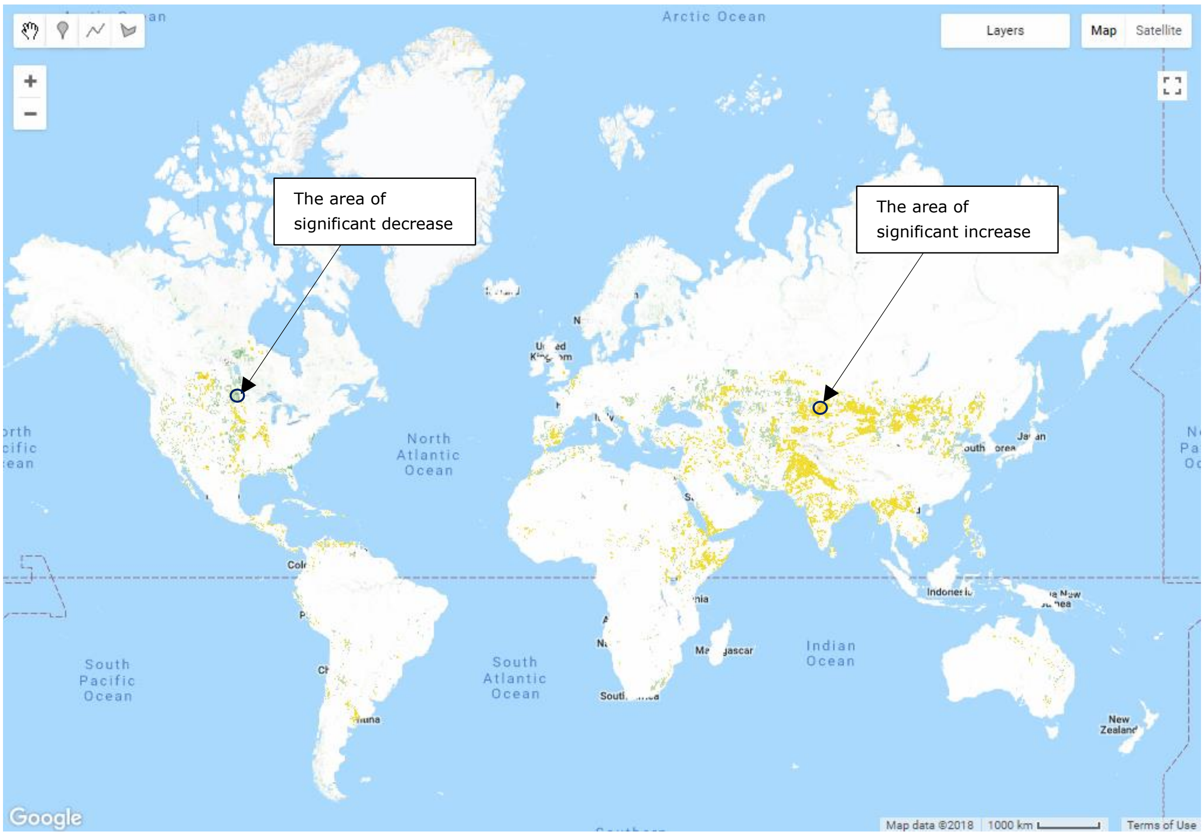


Figure 9. Global soil salinity changes map from 1986 to 2016 (yellow shades show soil salinity increase and green shades show soil salinity decrease).

3.4. Discussion of the method and implications

The trend in soil salinity increase over time we showed was in most of the cases in accordance with other studies (Taghadosi and Hasanlou, 2017; Wang et al., 2008). However, a comparison with studies that describe soil salinity decrease over time would provide better validity of the method. Though areas where soil salinity decreased in time might exist, they definitely would not be widespread. Overall consensus among experts is that soil salinisation is expanding on a global scale, probably at a rate of 2 Mha per year (Abbas et al., 2013). As a result, we could not find a study describing soil salinity decrease through time. Without it, the trend might also represent general trend of global warming. Interestingly enough, even if so, our change maps still might be valid. Climate change is promoting soil salinisation through more frequent drought events, seawater intrusion in coastal areas and general increase in temperature (Dasgupta et al., 2015; Várallyay, 1994). Therefore, we can assume that many areas suffering from climate change would be prone to soil salinisation.

The basis of thermal infrared imagery approach we used is described in Ivushkin et al. (2017); Ivushkin et al. (2018); Ivushkin et al. (2019). In those studies certain pre-processing was done to ensure that the thermal infrared data used in the analysis were coming from cropped areas vegetated above a specified threshold. Therefore, thermal infrared data used in these studies could be related to canopy temperature. In our case we did not do such a pre-processing because of issues that are the consequence of a global scale study, like vegetation season spanning all year round. Nevertheless, since increase in canopy temperature is a universal response to salt stress for a vast majority of plants (Munns, 1993, 2002), we assume that it will hold for other vegetation covers. In case of extremely saline areas where no vegetation is present, the surface temperature will be affected anyway, because open soil at a day time will have higher temperature than vegetated areas. Though we may presume that applying some comprehensive algorithm to select the thermal signal of vegetated areas only would increase validation accuracy, this will complicate the thermal analysis. As we described earlier, absolute temperature will not be a suitable indicator by itself, therefore some normalisation is required. That would be hardly possible if different parts of the single scene would have thermal information captured in different time of the year (because of several vegetation seasons, crop rotations, etc.). One solution might be using the deviation from air temperature, instead of the deviation from the long term mean surface temperature, which we used. But this will require global air temperature dataset of unprecedented detail, both temporal (30-60 minutes) and spatial, and synchronisation of this dataset with remote sensing observations.

The thermal infrared data used was obtained by Landsat 5 and 8 satellites. While there are other sensors exist, our choice was deliberate. The main requirement that ruled out almost all other sensors is the time span of the research. There were only few openly available satellite sensors in 80s, even few capable of capturing images in thermal infrared band. Therefore, taking into account requirement for global coverage, Landsat program satellites are an obvious choice.

Soil salinity is a quite dynamic soil property, both spatially and temporally, and it can vary on the scale of few meters. Therefore, high spatial resolution of remote sensing data is desirable for soil salinity assessment. While for visible bands such resolution is achievable (though mainly in commercial sensors), acquiring thermal images in high resolution is hindered by much lower signal intensity in the thermal infrared band. Moreover, other covariates in our study have spatial resolution of 250m. Therefore, having thermal images with resolution higher than 250m will bring only limited improvement for the final map. However, using higher resolution images in further studies would be undoubtedly beneficial due to the dynamic nature of soil salinity.

One of the limitations of the WoSIS dataset we used is the spatially unequal sample distribution. That might be one of the reasons why the amount of salt affected lands in Mexico is overestimated. On the other hand, our approach was able to map salt affected areas in regions where training data were absent, like Central Asia. Therefore, we conclude that the unequal spatial distribution of training points did influence the results, but did not influence them significantly.

Though the thermography approach is more universal compared with other remote sensing techniques for soil salinity assessment, it still has some drawbacks. One of them is the different degree of the thermal response among plants. More salt tolerant plants would exhibit less increase in canopy temperature compared with not tolerant plants. Therefore, in the areas where more and less salt tolerant plants are growing in vicinity of each other, an assessment error is possible. Though in reality such a situation can rarely be encountered.

Though Google Earth Engine is a powerful tool that provides access to the biggest library of open earth observation data and computational power to process it, the scientific community is quite cautious in adopting it. The main reason is that exact implementation of different functions, including random forest we used, is not always known. Moreover, these implementations can be changed at any moment, leading to different results even if you use the same functions to compute these results later. We recognise this issue. Nevertheless, its free of charge access and rich earth observation data archive makes GEE a useful tool for global assessments of different kinds.

One of the directions for a further research might be the application of different machine learning algorithms, numerous selection of which exists. We intentionally limited ourselves to three main algorithms present in Google Earth Engine and, after discovering that two of them are not useful for our dataset/model combination, we continued the analysis using random forest.

As we mentioned before, existing assessments of salt affected soils on a global scale are quite limited and approximate. Though the knowledge about the total affected area and its change would be an important information to improve global food security. The economic costs of soil salinity are also impressive. For example, just 2 million hectares of salt affected lands are costing Uzbekistan about US $1 billion annually (UNDP, 2009; World Bank, 2007). On a global scale the economic losses are just tremendous. A proper inventory of the affected lands would allow proper mitigation measures to be applied and cut the losses to the minimum. We hope that our study will contribute to this cause.

4. Conclusions

The results show that GEE random forest classifier is a useful tool for the global assessment of soils salinity. The resulting global soil salinity maps have a validation accuracy of up to 70% with several known hotspots captured. The assessment of global area affected is comparable with the assessments of other authors. The addition of thermal infrared imagery into the analysis can act as a dynamic variable that allows to capture the trend of soil salinity change. That was confirmed in 2 out of 3 investigated cases. The one case where our results were different from the referred study had soil salinity of a different origin and we suspect that this might be the reason why the method did not perform well in this case. The method we applied allowed to predict affected areas even in the regions where training data were unavailable. Moreover, even in cases of misclassification in Highly and Extremely saline classes, misclassified points were still attributed to a saline class and only rarely to Non-saline, which means that areal extent of salt affected lands can be successfully mapped, and only definition of degree of salinity still represents a challenge. Therefore, we conclude that a combination of soil properties maps and thermal infrared imagery can allow mapping of soil salinity development in space and time on a global scale.

The code and data used to produce the global soil salinity maps can be accessed by registered Google Earth Engine users at https://code.earthengine.google.com/d43e5a92ae1deed32a0929f57b572756.

References

Abbas, A., Khan, S., Hussain, N., Hanjra, M.A., Akbar, S., 2013. Characterizing soil salinity in irrigated agriculture using a remote sensing approach. Physics and Chemistry of the Earth, Parts A/B/C 55-57, 43-52. 10.1016/j.pce.2010.12.004.

Abrol, I.P., Yadav, J.S.P., Massoud, F.I., 1988. Salt-Affected Soils and their Management. FAO.

Al-Busaidi, A.S., Cookson, P., 2003. Salinity–pH Relationships in Calcareous Soils. Journal of Agricultural and Marine Sciences 8, 6. 10.24200/jams.vol8iss1pp41-46.

Allbed, A., Kumar, L., 2013. Soil Salinity Mapping and Monitoring in Arid and Semi-Arid Regions Using Remote Sensing Technology: A Review. Advances in Remote Sensing 02, 373-385. 10.4236/ars.2013.24040.

Allbed, A., Kumar, L., Aldakheel, Y.Y., 2014a. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region. Geoderma 230-231, 1-8. 10.1016/j.geoderma.2014.03.025.

Allbed, A., Kumar, L., Sinha, P., 2014b. Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques. Remote Sensing 6, 1137-1157. 10.3390/rs6021137.

Breiman, L., 2001. Random Forests. Machine Learning 45, 5-32. 10.1023/a:1010933404324.

Dasgupta, S., Hossain, M.M., Huq, M., Wheeler, D., 2015. Climate change and soil salinity: The case of coastal Bangladesh. Ambio 44, 815-826. 10.1007/s13280-015-0681-5.

Douaoui, A.E.K., Nicolas, H., Walter, C., 2006. Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. Geoderma 134, 217-230. 10.1016/j.geoderma.2005.10.009.

Fan, X., Pedroli, B., Liu, G., Liu, Q., Liu, H., Shu, L., 2012. Soil salinity development in the yellow river delta in relation to groundwater dynamics. Land Degradation and Development 23, 175-189. 10.1002/ldr.1071.

FAO 2018. Salt-affected soils. Accessed on 04-10-2018. http://www.fao.org/soils-portal/soil-management/management-of-some-problem-soils/salt-affected-soils/more-information-on-salt-affected-soils/en/.

FAO and ITPS, 2015. Status of the World's Soil Resources (SWSR) – Main Report. Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, Rome, Italy.

Gómez-Bellot, M.J., Nortes, P.A., Sánchez-Blanco, M.J., Ortuño, M.F., 2015. Sensitivity of thermal imaging and infrared thermometry to detect water status changes in Euonymus japonica plants irrigated with saline reclaimed water. Biosystems Engineering 133, 21-32. 10.1016/j.biosystemseng.2015.02.014.

Hasanlou, M., Eftekhari, K., 2019. Soil salinity mapping using dual-polarized SAR Sentinel-1 imagery AU - Taghadosi, Mohammad Mahdi. International Journal of Remote Sensing 40, 237-252. 10.1080/01431161.2018.1512767.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLOS ONE 12, e0169748. 10.1371/journal.pone.0169748.

IAEA, 1995. Management strategies to utilize salt affected soils : isotopic and conventional research methods. IAEA, Vienna.

Ivushkin, K., Bartholomeus, H., Bregt, A.K., Pulatov, A., 2017. Satellite Thermography for Soil Salinity Assessment of Cropped Areas in Uzbekistan. Land Degradation & Development 28, 870-877. 10.1002/ldr.2670.

Ivushkin, K., Bartholomeus, H., Bregt, A.K., Pulatov, A., Bui, E.N., Wilford, J., 2018. Soil salinity assessment through satellite thermography for different irrigated and rainfed crops. International Journal of Applied Earth Observation and Geoinformation 68, 230-237. 10.1016/j.jag.2018.02.004.

Ivushkin, K., Bartholomeus, H., Bregt, A.K., Pulatov, A., Franceschini, M.H.D., Kramer, H., van Loo, E.N., Jaramillo Roman, V., Finkers, R., 2019. UAV based soil salinity assessment of cropland. Geoderma 338, 502-512. https://doi.org/10.1016/j.geoderma.2018.09.046.

Kovács, D., Tóth, T., Marth, P., 2006. Soil Salinity between 1992 and 2000 in Hungary. AGROKÉMIA ÉS TALAJTAN 55, 89-98.

Metternicht, G., Zinck, J.A., 2009. Remote sensing of soil salinization impact on land management. CRC Press, Boca Raton, FL.

Munns, R., 1993. Physiological processes limiting plant growth in saline soils: some dogmas and hypotheses. Plant, Cell & Environment 16, 15-24. 10.1111/j.1365-3040.1993.tb00840.x.

Munns, R., 2002. Comparative physiology of salt and water stress. Plant, Cell & Environment 25, 239-250. 10.1046/j.0016-8025.2001.00808.x.

Oldeman, L.R., Hakkeling, R.T.A., Sombroek, W.G., 1991. World map of the status of human-induced soil degradation: an explanatory note, 2nd. rev. ed. ISRIC, UNEP, Wageningen [etc.].

Padarian, J., Minasny, B., McBratney, A.B., 2015. Using Google's cloud-based platform for digital soil mapping. Computers & Geosciences 83, 80-88. 10.1016/j.cageo.2015.06.023.

Ribeiro, E., Batjes, N.H., Leenaars, J.G.B., van Oostrum, A.J.M., Mendes de Jesus, J., 2015. Towards the standardization and harmonization of world soil data: Procedures manual ISRIC World Soil Information Service (WoSIS version 2.0). Report 2015/03. ISRIC - World Soil Information, Wageningen, p. 100.

Saidi, D., 2012. Importance and Role of Cation Exchange Capacity on the Physicals Properties of the Cheliff Saline Soils (Algeria). Procedia Engineering 33, 435-449. https://doi.org/10.1016/j.proeng.2012.01.1223.

Sazib, N., Mladenova, I., Bolten, J., 2018. Leveraging the Google Earth Engine for Drought Assessment Using Global Soil Moisture Data. Remote Sensing 10, 1265.

Seelig, B.D., 2000. Salinity and Sodicity in North Dakota Soils. NDSU Extension Service, North Dakota State University of Agriculture and Applied Science, Fargo, ND, p. 16.

Setia, R., Gottschalk, P., Smith, P., Marschner, P., Baldock, J., Setia, D., Smith, J., 2013. Soil salinity decreases global soil organic carbon stocks. Science of The Total Environment 465, 267-272. https://doi.org/10.1016/j.scitotenv.2012.08.028.

Squires, V.R., Glenn, E.P., 2004. Salination, desertification, and soil erosion, in: Squires, V.R. (Ed.), The role of food, agriculture, forestry and fisheries in human nutrition. UNESCO, EOLSS Publishers, Oxford ,UK.

Strobl, C., Malley, J., Tutz, G., 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychological methods 14, 323-348. 10.1037/a0016973.

Szabolcs, I., 1989. Salt-affected soils. CRC Press, Boca Raton, Fla.

Taghadosi, M.M., Hasanlou, M., 2017. Trend analysis of soil salinity in different land cover types using Landsat Time Series Data (Case Study Bakhtegan salt lake). Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-4/W4, 251-257. 10.5194/isprs-archives-XLII-4-W4-251-2017.

UNDP, 2009. National irrigated land reclamation fund capacity development, project document, p. 54.

Urrestarazu, M., 2013. Infrared thermography used to diagnose the effects of salinity in a soilless culture. Quantitative InfraRed Thermography Journal 10, 1-8. 10.1080/17686733.2013.763471.

Várallyay, G., 1994. Climate Change, Soil Salinity and Alkalinity. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 39-54.

Wang, Y., Xiao, D., Li, Y., Li, X., 2008. Soil salinity evolution and its relationship with dynamics of groundwater in the oasis of inland river basins: case study from the Fubei region of Xinjiang Province, China. Environmental monitoring and assessment 140, 291-302. 10.1007/s10661-007-9867-z.

World Bank, 2007. Integrating environment into agriculture and forestry. Progress and prospects in Eastern Europe and Central Asia. Uzbekistan, p. 12.

List of Figure Captions

Figure 1. Distribution of ground truths sampling data

Figure 2. Resulting global soil salinity map for 2016

Figure 3. Salt affected land area for different years; a) shows division per salinity class (note the y-axis uses an offset), b) shows the affected area

Figure 4. Soil salinity map of Grand Forks county and surroundings (2016)

Figure 5. Hungary map of soil salinity (2016)

Figure 6. Soil salinity maps (upper from 1986 and lower from 2005) of the Fubei region of Xinjiang Province, China. According to Wang et al. (2008) soil salinity increased in this area

Figure 7. Soil salinity maps (upper from 2000 and lower from 2016) of the Bakhtegan Salt Lake region in Iran. According to Taghadosi and Hasanlou (2017) soil salinity increased in this area

Figure 8. Soil salinity maps (upper from 1986 and lower from 2005) of the Yellow River Delta, China. According to Fan et al. (2012) the soil salinity increased in this area

Figure 9. Global soil salinity changes map from 1986 to 2016 (yellow shades show soil salinity increase and green shades show soil salinity decrease)