SCIENTIFIC DATA

OPEN

Data Descriptor: Top soil physical and chemical properties in Kazakhstan across a north-south gradient

Received: 17 May 2018
Accepted: 10 September 2018
Published: 13 November 2018

Vadim Yapiyev[1,2,*], Charles P. Gilman[3,6], Tolganay Kabdullayeva[3], Akmaral Suleimenova[3], Aizhan Shagadatova[3], Azat Duisembay[3], Sanzhar Naizabekov[3], Saule Mussurova[2,3], Kamilya Sydykova[3], Ilyas Raimkulov[3], Ilyas Kabimoldayev[3], Ainagul Abdrakhmanova[2,3], Symbat Omarkulova[4], Dastan Nurmukhambetov[4], Aliya Kudarova[4], Daniyar Malgazhdar[2], Christian Schönbach[3,5,*] & Vassilis Inglezakis[4,*]

Kazakhstan's soil properties have yet to be comprehensively characterized. We sampled 40 sites consisting of ten major soil types at spring (wet) and late-summer (dry) seasons. The sample locations range from semi-arid to arid with an annual mean air temperature from 1.2 to 10.7 °C and annual precipitation from less than 200 to around 400 mm. Overall topsoil total (STC), organic (SOC), and inorganic (SIC) carbon did not change significantly between spring and late summer. STC and SOC show a wave like pattern from north to south with two maxima in northern and southern Kazakhstan and one minimum in central Kazakhstan. With a few exceptions SIC content at northern sites is generally low, whereas at Lake Balkhash SIC can exceed 75% of STC. Independent of the seasons, SOC significantly differed among soil types. Total nitrogen content distribution among our sampling sites followed a similar pattern as SOC with significant differences between soil types occurring in northern, central and southern Kazakhstan.

Design Type(s)	data collection and processing objective
Measurement Type(s)	recorded image • soil
Technology Type(s)	photography • data collection method
Factor Type(s)	sampling time • geographic location
Sample Characteristic(s)	North Kazakhstan Province • Aqmola Province • Qaraghandy Province • Zhambyl Province • area of scrub • area of cropland • area of deciduous forest • steppe • area of developed open space • area of dwarf scrub • cold desert • desert scrubland • area of emergent herbaceous wetland

[1]Department of Civil and Environmental Engineering, School of Engineering, Nazarbayev University, Astana, Kazakhstan. [2]National Laboratory Astana, Nazarbayev University, Astana, Kazakhstan. [3]Department of Biology, School of Science and Technology, Nazarbayev University, Astana, Kazakhstan. [4]Environmental Science & Technology Group (ESTg), Chemical Engineering Department, School of Engineering, Nazarbayev University, Astana, Kazakhstan. [5]International Research Center for Medical Sciences, Kumamoto University, Kumamoto, Japan. [6]Present address: United States Agency for International Development, Washington, D.C., USA. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to C.S. (email: christian.schoenbach@nu.edu.kz) or V.I. (email: vasileios.inglezakis@nu.edu.kz)

Background & Summary

Kazakhstan is the largest land-locked country in the world. Its massive land area of 2.725×10^6 km^2 represents a key reservoir for soil organic carbon that is thought to play an important role in global climate-carbon modelling[1,2]. Yet SOC data pertaining to Central Asia[3] and particularly Kazakhstan[4] are scarce, resulting in considerable ignorance of the extent to which forests and steppe soils in the north and east, and arid areas in the south and west, contribute to sequestering atmospheric carbon or yielding carbon into the atmosphere. Modelling approaches, such as the arid ecosystem model (AEM), yielded high soil organic carbon densities of up to 34 Pg for top soils of 1 m depth and 12–14 Pg for 30 cm deep top soils in temperate deserts of Central Asia[5]. Sub-regional SOC studies utilizing satellite imagery of farm land in northern Kazakhstan and temporal carbon variations lacked predictive power[6]. Normalized difference vegetation index (NDVI) values that reflect vegetative sensitivity to climate change were found to be high in northern Kazakhstan, and in central Kazakhstan, and NDVI changes were positively correlated with the annual temperature[7]. Net ecosystem CO_2 exchange (NEE) studies of two sites representing alkaline desert soils near Lake Balkhash and the Aral Sea revealed CO_2 flux dependence on moisture, pH and light[8]. The authors reported net CO_2 release nocturnally and on cloudy days with precipitation, whereas on sunny and dry days, CO_2 was taken up. Overall, the current picture of carbon cycling in Kazakhstan is supported by only a few sites and little data. In addition numerous SOC centric studies[9–11] make extrapolations on regional and global carbon cycles without taking into account soil inorganic carbon (SIC), a likely sink of secondary carbonates, particularly in arid areas. According to Lal *et al.*[12], SIC accumulation is high in arid and semiarid regions, for example, grass lands which are thought to harbour one fifth of global soil carbon stocks[13]. SIC accumulation in top soils (15 cm) was shown to be largely dependent on soil pH[14].

More data and improved coverage of fundamental soil properties including pH, moisture content, particle size, and cation and anion composition that influence microbial activity, and thus the rate of SOC decomposition, are needed to synthesize better agro-economic and ecological strategies, especially regarding climate change predictions. Towards this end we assessed the current soil properties of Kazakhstan by sampling top soil (15 cm) at 40 sites between Petropavlovsk (north) and Taraz (south) during wet and dry seasons and determined the physical and chemical compositions, related vegetation, land cover and climate properties.

Methods

Study sites

The study has 40 sampling sites that were located a minimum of 50 m from the nearest road. Sample sites were approximately 50 km apart with flat topographic conditions (toeslope) (Fig. 1). No permits were required for the sampling site locations and the sites did not harbour endangered or protected species. Geographical coordinates (WGS-84) were recorded using a Garmin T650 hand-held global positioning system. The coordinates and geographical annotations are shown in Table 1 (available online only). Sampling was conducted in 2015: one in the "wet" season after the snowmelt (May), and the other in the "dry" season at the end of the growing season (September).

Field Methods

Sampling site documentation. Pictures of the landscape and vegetation were taken at each location (see Supplement 1). The soil sampling procedure was divided into physical properties sampling, chemical sampling, and biome sampling. Soil sample preparation conditions for chemical and physical analyses (including depth of cores, particle size and milling) were chosen to be harmonized with future analyses of soil microbiomes[15,16]. Samples for biome (not subject of this paper), physical and chemical analysis parts (see Laboratory Methods) were transported on ice and transferred to 4 °C refrigerators for extended storage.

Samples for physical analyses. Samples for gravimetric moisture determination were obtained by digging a 10 cm diameter hole with a spade. The spade was rinsed well with distilled water between samples. Approximately 1 kg of soil was excavated, mixed and sealed in a 1 L plastic bag. For bulk density determination, a 50 ml conical tube (Corning Inc.) was filled with undisturbed soil derived from 15 cm deep cores and weighed on a portable balance (Maxx-412, Denver Instrument). At five sample sites additional samples for bulk density measurements were obtained to verify the cone method against the traditional soil ring method.

Samples for chemical analyses. After removing the litter layer if present, (approximately top 2 cm) soil cores of 1.5 cm in diameter were taken to depth of approximately 15 cm. The cores were transferred into 50 ml falcon tubes that were sealed with Parafilm M (Bemis Company, Inc.).

Laboratory Methods

Soil moisture and bulk density. Soil moisture was calculated gravimetrically as wet and dry soil weight ratios. Briefly, triplicates of 10 g of soil were placed on aluminium paper, weighed and transferred to a 105 °C ventilated oven (Heraeus LUT 6050, Therma Fisher Scientific Inc.). After 24 h, samples were

Figure 1. **Locations of soil sampling sites.** Made with Natural Earth (http://www.naturalearthdata.com).

removed, immediately weighed and the dry mass recorded. Volumetric soil moisture content was calculated based on oven-dry bulk densities assuming 15 cm soil depth and by subtracting the mass of evaporated water from the wet bulk density. The volumetric soil moisture and calculated carbon stock values represent averages from two oven-dried bulk densities values obtained from two independent samples.

Soil milling for pH, conductivity and elemental analysis. Ten grams of dried and sieved soil of smaller than 150 µm particle size[15] was transferred into the jar of a vibrating ball mill (MM 400, Retsch GMBH) for 2 min milling at 30 Hz. The milled sample was collected in a 15 ml Falcon tube. The vibrating ball mill processing reduces the sample particle size to approximately 1 µm.

Soil pH and conductivity. Soil pH and electrical conductivity were measured in triplicate in aqueous solution suspension (supernatant) with a 1: 2.5 (soil : water) ratio according the protocol of Pansu and Gautheyrou (2006) using a 8107UWMD Ross Ultra pH/ATC triode (Thermo Fisher Scientific, USA) and Orion 013010MD conductivity cell (Thermo Fisher Scientific, USA) electrodes.

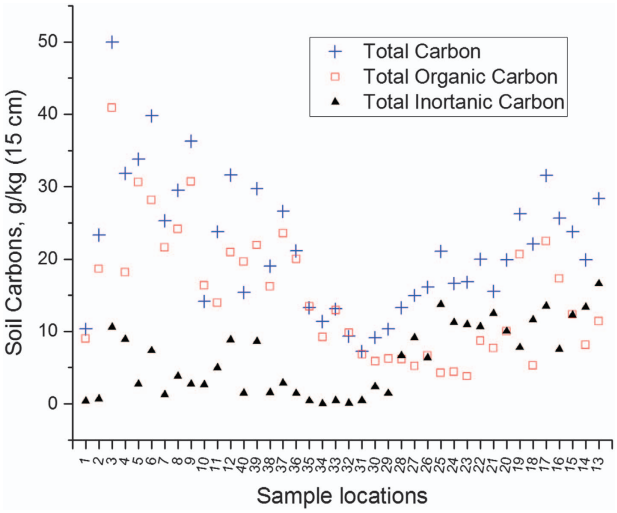

Figure 2. Soil carbon as total soil cabon (blue crosses), soil organic carbon (red squares), soil inorganic carbon (black triangles). See also Table 2.

Total soil carbon (SC), organic (SOC) and inorganic carbon (SIC) based on elemental analysis. Total SC and SOC were measured using a CNHS-O dry combustion elemental analyzer Multi N/Cb 3100, (Analytik Jena, Germany). For total SC 100 mg milled soil was placed in a ceramic combustion boat and combusted under pure oxygen at a flow rate of 2.8 L/min at 950 °C. The CO_2 emitted during combustion is detected by a thermal conductivity detector. Milled soil samples for SOC estimation were pre-treated in combustion boats with 100 µl of H_3PO_4 (30–40%) to dissolve carbonates. Samples were dried overnight at 70 °C and subjected to combustion at 950 °C under 14 L/min oxygen (Multi N/Cb 3100, (Analytik Jena, Germany)). SIC was calculated as the difference between total SC and SOC (Fig. 2 and Table 2).

Loss-on-ignition (LOI) procedure for soil organic matter (SOM). The soil samples were dried and sieved through a 2 mm sieve. We adapted the LOI method described by Emmett *et al.* (2007; Countryside Survey: Soils Report from 2007). In brief, crucibles were washed and then rinsed three times with distilled water, dried for 40 min at 105 °C in an oven (Carbolite PN 60) and then cooled to room temperature (RT) in a desiccator for 30 min. Each crucible weight was recorded (Wc) and 10.00 g of crushed soil sample (< 2 mm) was weighed in the crucible, dried for four hours at 105 °C and cooled to RT in a desiccator for 30 min and the dry sample weight (Ws) recorded. Dried samples were loaded into a muffle furnace (Carbolite ELF 11/6B), heated to 375 °C for 16 h and allowed to cool down to 150 °C before being transferred to a desiccator for 30 min to cool to RT. The weight of the samples was recorded as Wa and LOI was calculated as $(Ws - Wa)/(Ws - Wc) \times 100$.

Total nitrogen (TN) measurement by elemental analysis. The total nitrogen content of each sample was analyzed by quantitative combustion in excess oxygen using DuMaster D-480 analyzer (Büchi Labortechnik AG). L-glutamic acid was used for calibration (N-factor) of the sample measurement series. Dried and milled samples were weighed in portions of 700 mg, packed in tin foil and loaded on the sample carousel for total nitrogen measurement according to the manufacturer's instructions (Fig. 3).

Climate data

The Climate Research Unit high-resolution dataset (CRU TS v. 3.24.01 https://crudata.uea.ac.uk; January 26, 2017 release) contains air temperature and precipitation data ranging from 1901 until 2015 at 0.5° resolution of grid-boxes[17] Google Earth Interface Pro was used to download the raw monthly temperature and precipitation data for the sample locations. If multiple sample locations (1–3, 4–5, 32–33, 28–29, and 19–20) mapped to one grid-box, the same temperature and precipitation data were used (e.g., sample sites 1–3 were assigned the same climate data). The climate data are presented in Table 1 (available online only).

Carbon and nitrogen stocks calculation

To infer soil total organic carbon (TOC) and nitrogen (TN) stocks in tons of TOC and TN per hectare (tC/ha, tN/ha) we followed the procedure outlined by Rowell (1994)[18], (section 3.7, p. 55). The soil layer

Sample #	TC	TOC	TIC
	g/kg	g/kg	g/kg
1	10.3	9.0	0.4
2	23.3	18.7	0.7
3	50.0	40.9	10.6
4	31.9	18.2	8.9
5	33.8	30.6	2.7
6	39.8	28.1	7.4
7	25.3	21.6	1.3
8	29.5	24.1	3.8
9	36.3	30.7	2.7
10	14.2	16.3	2.7
11	23.8	14.0	5.0
12	31.6	21.0	8.8
40	15.4	19.6	1.5
39	29.7	21.9	8.6
38	19.0	16.2	1.6
37	26.6	23.6	2.9
36	21.1	20.0	1.5
35	13.3	13.4	0.4
34	11.4	9.2	0.1
33	13.1	13.0	0.5
32	9.3	9.8	0.1
31	7.2	6.9	0.5
30	9.2	5.9	2.4
29	10.3	6.2	1.5
28	13.2	6.1	6.7
27	15.0	5.2	9.1
26	16.1	6.7	6.4
25	21.1	4.3	13.7
24	16.6	4.5	11.3
23	16.8	3.8	10.9
22	19.9	8.7	10.7
21	15.5	7.7	12.5
20	19.9	10.0	10.1
19	26.2	20.7	7.8
18	22.1	5.3	11.6
17	31.5	22.5	13.5
16	25.6	17.3	7.5
15	23.8	12.3	12.3
14	19.9	8.1	13.4
13	28.3	11.4	16.6

Table 2. Soil carbon across the sites as total cabon (TC), total organic carbon (TOC), total inorganic carbon (TIC) in g/kg. See also Fig. 2.

was assumed to represent the 0–15 cm depth. We used oven dry bulk densities averages from the two samplings (May and September). For TOC and TN absolute content (g/kg or mg/g) we used mean values of averaged duplicate measurements (elemental analysis) obtained from two samplings (May and September) see Fig. 4.

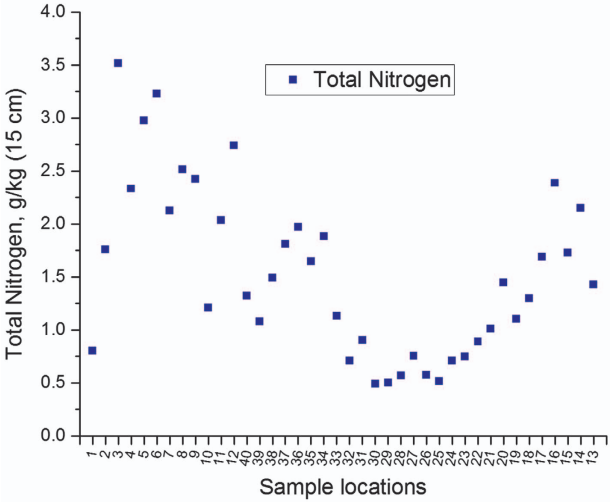

Figure 3. **Total soil nitrogen.**

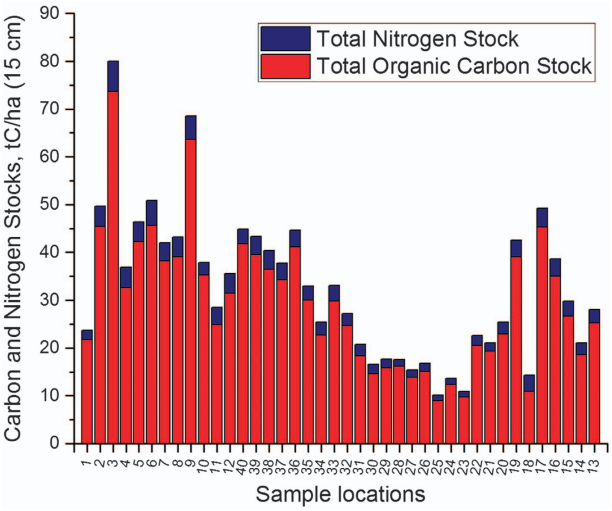

Figure 4. **TOCand TN stocks in 15 cm of top soil.**

Data Records

The data were deposited at the Mendeley data repository (Data Citation 1) as Supplement 1 and 2. Supplement 1 contains a folder with photos of landscape and vegetation for each sample location. The naming convention is location number. photo number (date the sample was taken). For example, 1.1 (25.05.15) corresponds to location 1 and photo 1 on May 25, 2015. For the images presented in our data records which feature identifiable human participant(s) the informed consent was obtained from the participants prior to publication of the images. Supplement 2 is a Microsoft Excel file that contains spreadsheets with data on 1) soil TC, soil total organic carbon (TOC), soil total inorganic carbon (TIC), 2) Loss on Ignition (LOI), 3) soil TN, 4) soil dry bulk density, 5) gravimetric and volumetric soil moisture 6) soil suspension and supernatant pH, 7) soil electrical conductivity, and 8) soil TOC and TN 9) annotation table (Table 1 (available online only)). Values of sample measurements from each location were reported with standard deviation (STDEV) and standard error (SE) when applicable.

Technical Validation

Soil carbon and nitrogen concentrations were averaged from duplicate sample preparations and validated using reference standards. If the concentrations deviated significantly the measurements were repeated.

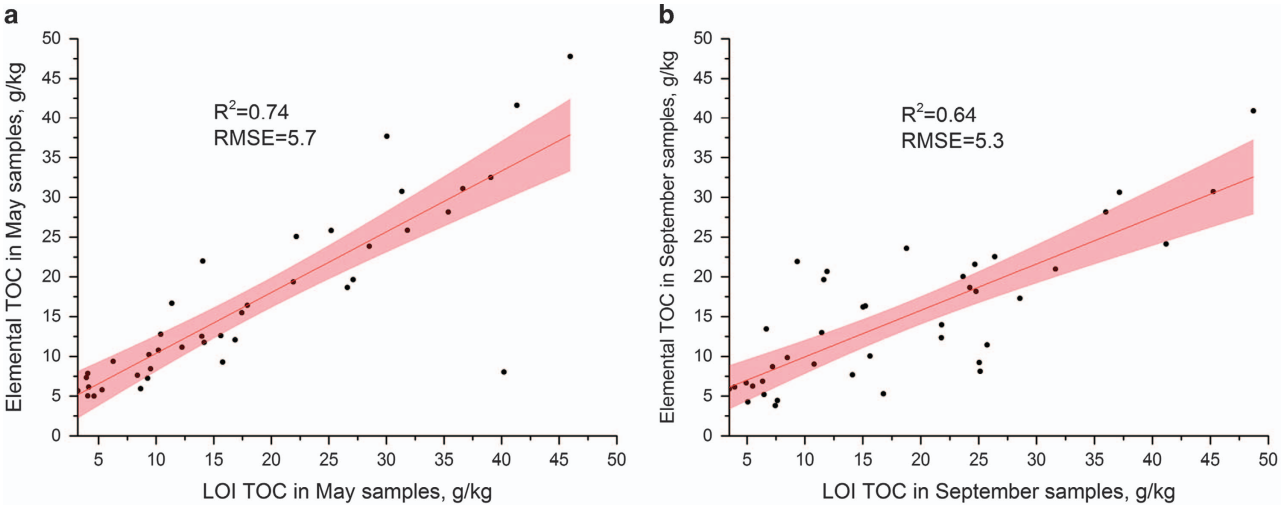

Figure 5. Comparison of TOC values derived from elemental analysis and LOI. (**a**) May samples. (**b**) September samples. R^2: coefficient of determination, RMSE: root mean squared error, and red-shaded area: 95% confidence interval.

SOC results were validated independently by conducting LOI measurements (see Fig. 5). Soil organic matter (SOM) (g/kg) data obtained from LOI analyses were converted to TOC (g/kg) by multiplying the values with the coefficient 0.58[19].

References

1. Propastin, P. A. & Kappas, M. In *Land Degradation and Desertification: Assessment, Mitigation and Remediation* 297–306 Springer (2010).
2. Yapiyev, V., Sagintayev, Z., Inglezakis, V. J., Samarkhanov, K. & Verhoef, A. Essentials of endorheic basins and lakes: A review in the context of current and futurewater resource management and mitigation activities in Central Asia. *Water (Switzerland)* **9** 798 (2017).
3. Groisman, P. *et al.* Northern Eurasia Future Initiative (NEFI): facing the challenges and pathways of global change in the twenty-first century. *Prog. Earth Planet. Sci* **4,** 41 (2017).
4. Takata, Y., Funakawa, S., Akshalov, K., Ishida, N. & Kosaki, T. Influence of land use on the dynamics of soil organic carbon in northern Kazakhstan. *Soil Sci. Plant Nutr.* **53,** 162–172 (2007).
5. Li, C. *et al.* Carbon stock and its responses to climate change in Central Asia. *Glob. Chang. Biol* **21,** 1951–1967 (2015).
6. Takata, Y. *et al.* Influence of crop rotation system on the spatial and temporal variation of the soil organic carbon budget in northern Kazakhstan. *Soil Sci. Plant Nutr.* **54,** 159–171 (2008).
7. Yin, G., Hu, Z., Chen, X. & Tiyip, T. Vegetation dynamics and its response to climate change in Central Asia. *J. Arid Land* **8,** 375–388 (2016).
8. Li, L., Chen, X., van der Tol, C., Luo, G. & Su, Z. Growing season net ecosystem CO2 exchange of two desert ecosystems with alkaline soils in Kazakhstan. *Ecol. Evol* **4,** 14–26 (2014).
9. Song, B. *et al.* Light and heavy fractions of soil organic matter in response to climate warming and increased precipitation in a temperate steppe. *PLoS One* **7,** e33217 (2012).
10. Lal, R., Negassa, W. & Lorenz, K. Carbon sequestration in soil. *Current Opinion in Environmental Sustainability* **15,** 79–86 (2015).
11. Yigini, Y. & Panagos, P. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. *Sci. Total Environ.* **557–558,** 838–850 (2016).
12. Lal, R. Carbon sequestration. *Philos. Trans. R. Soc. B Biol. Sci* **363,** 815–830 (2008).
13. Wang, W. & Fang, J. Soil respiration and human effects on global grasslands. *Glob. Planet. Change* **67,** 20–28 (2009).
14. Shi, Y. *et al.* Organic and inorganic carbon in the topsoil of the Mongolian and Tibetan grasslands: Pattern, control and implications. *Biogeosciences* **9,** 2287–2299 (2012).
15. Burt, R. *Kellogg soil survey laboratory methods manual.* (United States Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center, Kellogg Soil Survey. *Laboratory* (2014).
16. Zhou, J., Bruns, M. A. & Tiedje, J. M. DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* **62,** 316–322 (1996).
17. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int. J. Climatol.* **34,** 623–642 (2014).
18. Rowell, D. L.. *Soil Science: Methods and Applications*. (Routledge, 1994).
19. Pribyl, D. W. A critical review of the conventional SOC to SOM conversion factor. *Geoderma* **156,** 75–83 (2010).

Data Citations

1. Yapiyev, V. *et al. Mendeley Data* https://doi.org/10.17632/5ZMVX4VH28.1 (2018).

Acknowledgements

This research was supported in part by the grant "Kazakhstan soil microbiome: agricultural characteristics and perspectives" in the framework of the agreement #104 dated 12.02.2015 of the program #055 "Scientific and/or Scientific-Technical Activities" within subprogram #101 "Research Grant Funding" from the Ministry of Education and Science of the Republic of Kazakhstan and by Nazarbayev University. We thank Jong Kim for letting us use the equipment of Department of Civil and Environmental Engineering, Almagul Kushugulova and Samat Kozhakhmetov for administrative and managerial support and Nurlan Ongdas for helping to collect field samples. We are grateful for the valuable input and assistance provided by Drs. Joanna Clark, Tom Sizmur, Maria Shahgedanova, and Anne Verhoef at University of Reading.

Author Contributions

C.S., C.G. and V.Y. conceived the project and led the work in collaboration with V.I. C.G., V.Y. and V.I. established the sampling and analysis methodologies. C.G., V.Y., D.M. and T.K. collected the soil samples and performed the field measurements. V.Y. prepared the map, climate data and calculated carbon stocks. A. Sh. conducted pH and EC measurements. A.D. and D.M. performed the physical analyses of samples. S.N., S.M. and K.S. conducted the L.O.I. analyses. I.R., I.K., C.G., D.M., A. Su., A.A., and T.K. processed the field samples. S.O., D.N., A.K. and C.S. performed the chemical analyses of samples. C.S. and V.I. analysed the chemical data. V.Y. and C.S. took the lead in writing the manuscript. All authors provided critical feedback and contributed to the final version of the manuscript.

Additional Information

Table 1 is only available in the online version of this paper.

Competing interests: The authors declare no competing interests.

How to cite this article: Yapiyev, V. *et al*. Top soil physical and chemical properties in Kazakhstan across a north-south gradient. *Sci. Data*. 5:180242 doi: 10.1038/sdata.2018.242 (2018).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.