Copper Content and Export in European Vineyards Soils Influenced by Climate and Soil Properties

Boris Droz,[†] Sylvain Payraudeau,[†] José Antonio Rodríguez Martín,[‡] Gergely Tóth,[§] Panos Panagos,[‖] Luca Montanarella,[‖] Pasquale Borrelli [○,¥] & Gwenaël Imfeld [†*]*

[†] Institut Terre et Environnement de Strasbourg, Université de Strasbourg, ENGEES, CNRS, UMR 7063, 5 rue Descartes, Strasbourg F-67084, France,

[‡] Department of the Environment, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaría (INIA), Carretera de la Corunã 7.5, 28040 Madrid, Spain

[§] Institute of Advanced Studies, Chernek str. 14., 9730 Kőszeg, Hungary

[‖] European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability, Via E. Fermi 2749, 21027 Ispra, Italy

[○] Department of Earth and Environmental Sciences, University of Pavia, Via Ferrata, 1, 27100 Pavia, Italy

[¥] Department of Biological Environment, Kangwon National University, Chuncheon 24341, Republic of Korea

* Corresponding author.

e-mail: bodroz@bluewin.ch; imfeld@unistra.fr

phone: + 33 3 6885 0474

Supporting Information

24 Pages

5 Figures

5 Tables

Table of contents

Supporting Materials and Methods

Total Copper in the Topsoil Datasets

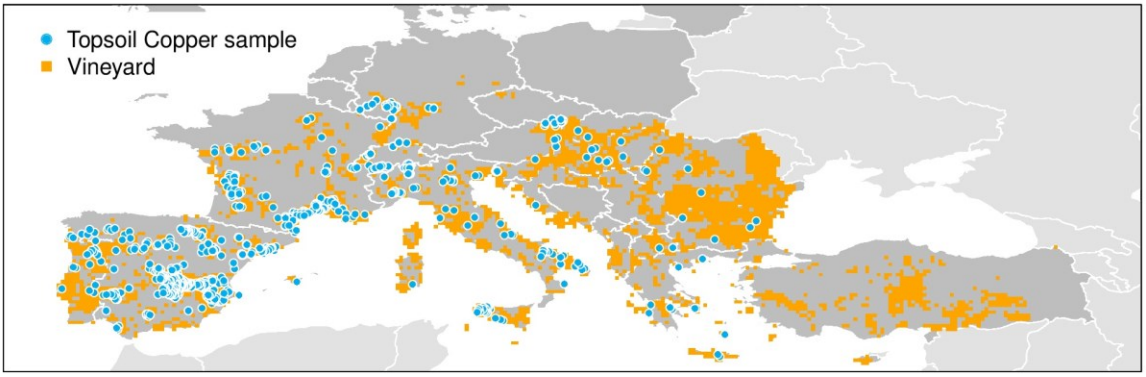

Figure S1. Distribution of the total copper content in the topsoil dataset (1,202 $Cu_{t,\,dat,}$ represented by blue dots) across European vineyards (Orange area; CLC land cover of vineyard was aggregated to a pixel of 20 km^2 to improve readability); Dark grey: Extent of land use Corine Land Cover (CLC 2012; v18.5.1; http://land.copernicus.eu).

Table S1. Total copper content in the topsoil dataset ($Cu_{t,\ dat}$; i.e., soils within 0 – 30 cm depth). The region, references, the sampling year(s), the analytical method used to quantify Cu, the total number of data (Nb data within parentheses the data located outside vineyards), and the minimal (min) and maximal (max) Cu_t, are reported for each dataset.

Region	Dataset name / Owner Ref.	Sampling Year(s)	Analytical Method [a]	Nb data [b]	Cu_t (min-max) (mg Cu kg^{-1})
Europe	Europe Land Use and Cover Area frame Statistical survey (LUCAS) / European Environment Agency (EEA) [1]	2009	AR / ICP-OES	380 (20,692)	0.54 - 496
Europe	Geochemical Mapping of Agricultural and Grazing land Soil (GEMAS) [2]	2008	AR / ICP-MS	43 (4,131)	2.6 - 395
Germany	Heavy metal and trace element background values in German Soils / Bundesanstalt für Geowissenschaften und Rohstoffe (BGR; https://produktcenter.bgr.de)	2003	HF-HCl-$HClO_4$ / XRF	10 (818)	15 - 239
France	Réseau de Mesure de la Qualité des Sols (RMQS) / Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE) [c]	2001-08	HF-$HClO_4$ / ICP-OES	38	11 - 491
Spain	Heavy metal in Spanish agricultural soil;data from Ministries of Agriculture and Environment[3]	2003-07	AR/ICP-OES	280 (4,013)	1.0 - 772
Bulgaria	Monitoring of degraded soil / Institute of Soil Science "Nikola Poushkarov"	2011	AR / ICP-OES	1	28
Czech Republic	Register of Contaminated Plots and Basal Soil Monitoring, Central Institute for Supervising and Testing in Agriculture (ÚKZÚZ) [4]	1992-1997	AR / ICP-MS	46	4.1 - 35
Catalogna (Spain)	[5, 6]	2001-13	AR / AAS	22	27 - 87
Mosel (Germany)	[7]	2003		60	19 - 656
Slovenia	[8, 9]	2005	AR / ICP-MS	72	2.4 - 166
Neuchâtel (Switzerland)	Service de l'énergie et de l'environnement, département du développement territorial et de l'environnement, canton de Neuchâtel (SENE)	2004-15	MicrO-HNO_3 (2M) / ICP-OES	7	157 - 497
Aargau (Switzerland)	Kantonales Bodenbeobachtungsnetz (KABO) vom Departement Bau, Verkehr und Umwelt, Abteilung für Umwelt Kanton Aargau	2006		4	269 - 278
Schaffhausen (Switzerland)	Umweltschutz Schaffhausen (SCHA)	1999-2011		10	71 - 436
Genève (Switzerland)	Département de l'environnement, des transports et de l'agriculture de Genève	2010-13		9	31 - 486
Tecino (Switzerland)	Sezione della protezione dell'aria, dell'acqua e del suolo	2016		200	11 - 774
Valais (Switzerland)	Département des transports, de l'équipement et de l'environnement (DTEE) du Valais	1992-2013		20	47 - 350
Total				**1,202 (29,654)**	**0.54 - 774**

[a] Acronyms of extraction method stand for aqua regia (AR), microwave (MicrO), Fluoridric acid (HF), Perchloric acid ($HClO_4$), nitric acid (HNO_3) and/or chloridric acid (HCl). Acronyms of methods stand for Inductively coupled plasma optical emission spectrometry (ICP-OES), ICP – mass spectrometry (ICP-MS), X-ray fluorescence spectrometry (XRF) and Atomic Absorption Spectroscopy (AAS). [b] Number of data used in the model; the total number of data in the survey are also provided in parenthesis if different. [c] Due to a confidential policy, the RMQS data were provide by the INRAE with a standard deviation of the geolocation of 206 m compare to an overall accuracy between 1 to 10 m for the other dataset. However, the deviation is systematically lower than the grid cell used for further prediction (250 m), therefore the deviation of the geolocation is considered as acceptable for predictions. Depth of the sampling is up to 20 cm for all data exception made for the RMQS, BGR and ÚKZÚZ dataset who contained data up to 30 cm depth. All Cu content were above the quantification limit (LQ)(highest LQ for the different dataset: 0.5 mg Cu kg^{-1}).

Building Spatial Variables

Table S2. Hypothetic predictive variables used as proxies of processes governing the total copper content in vineyard topsoil (Cu_t). All variables are averaged yearly.

Predictive variables	Original resolution (m)	Year of coverage*	Reference
Soil properties			
Clay, silt, sand, BD, CEC, C_{org}, soil pH, available soil water capacity, soil moisture	250	2000-2010	http://soilgrids.org [10]
Topography/Climate			
DEM	90	2000	http://srtm.csi.cgiar.org
LS, $S_{rad,}$ Slope	250	2000	DEM
P_{sum}, T_{av}	1000	1960-1990	www.worldclim.org, v1.4 [11]
PET, ET, AI	1000	1960-1990	www.cgiar-csi.org [12]
R-factor	500	2000-2010	[13, 14]
Land cover			
NDVI, EVI	250	2014-2016	MODIS [15]
LAI, FAPAR, FCover	300	2014-2016	http://land.copernicus.eu

***** Long-term average values within the period of collecting the soil data (Table S1) integrate the seasonal and inter-annual variation of the surface reflectance made by satellite recording. Average values can be considered as a better soil proxy than a single year of surface reflectance.[10]

Soil Physicochemical Properties

Soil texture (i.e. fraction of clay, silt and sand in %), bulk density (BD; kg m^{-3}), cation exchange capacity (CEC; cmol kg^{-1}), soil organic carbon content (C_{org}; %), soil pH in H_2O (unitless), available soil water capacity (%) and soil moisture content (moisture; %) of topsoil (0-50 cm depth). These properties were extracted from the SoilGrids project (http://soilgrids.org, series M_sl1, v18.4.2017)[10] at 250 m resolution.

Topography/Climate

The Digital Elevation Model (DEM) was generated from the NASA's Shuttle Radar Topography Mission (SRTM, v4.1, http://srtm.csi.cgiar.org/) at 90 m resolution for the year 2000. The DEM was previously aggregated by averaging the values to 250 m resolution. The aggregate DEM was used to compute a slope length (LS) factor using the Universal Soil Loss Equation (USLE) and the solar radiation input in kWh m^{-2} using SAGA-GIS 2.2.2+ through RSAGA

package v0.94-5. Prior to calculating LS, slopes in gradient were computed using the Zevenbergen and Thorne [16] algorithm and the catchment area (CA) using the O'Callaghan and Mark [17] algorithm. Based on the slope and CA, LS was computed using the Desmet and Govers [18] algorithm. The solar radiation input for each month for one year was computed using the Kumar, et al. [19] algorithm with basic parameters (solar constant = 1,367 W m^{-2}, height of the atmosphere = 12,000 m, atmospheric pressure = 1.013 bar, the vertical slide of the atmosphere containing water = 1.68 cm, transmittance of the atmosphere for Europe = 70%). Finally, the annual average potential solar radiation input (S_{rad};;W m^{-2}) was calculated as the day-weighted mean of monthly data and was used in further calculations.

The annual rainfall (rainfall; mm) and the annual average temperature (T_{av}; °C) were obtained from worldclim (www.worldclim.org, v1.4) at 1 km resolution for the period 1960-1990.[11] The annual average of potential evapo-transpiration (PET; mm), evapo-transpiration (ET; mm) and aridity index (AI; unitless) were calculated based on the wordlclim data at 1 km resolution.[12] PET is essentially a measure of the ability of the atmosphere to remove water through ET processes. PET was calculated by Hargreaves, et al. [20] following the eq. S1 where, S_{rad} is the annual average potential incoming solar radiation previously computed, T_{av} and T_{range} (°C) are the average and the range of annual temperatures, respectively, also computed based on worldclim data.

$$PET = 0.0023 \cdot S_{rad} \cdot (T_{av} + 17.8) \cdot 0.5 \cdot T_{range} \quad (S1)$$

The AI is used to quantify rainfall availability for atmospheric water demand and to characterize the locate climate (i.e. <0.03 hyper arid, 0.03 – 0.2 arid, 0.2 – 0.5 semi-arid, 0.5 – 0.65 dry sub-humid, >0.65 humid).[21] The AI was calculated by dividing the annual sum of rainfall by the average PET.

Rainfall erosivity (R-factor) as used by the Universal Soil Loss Equation (USLE) was computed by Panagos et al.[13] using datasets from 2000 to 2010 at 500 m resolution for Europe excluding Balkan region and Turkey, for which a 1 km resolution [14] was adopted. We merged both grid cells using the highest resolution available for each pixel as the R-factor value for further analysis.

Land Cover

Vegetation indexes for the period 2014-2016 were obtained from two specific sources. First, normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) were

obtained from the moderate resolution imaging spectro-radiometer (MODIS)[15] from the NASA Land Processes Distributed Active Archive Center (LP DAAC; https://lpdaac.usgs.gov) at 250 m resolution and a temporal coverage of 16 days. Second, leaf area index (LAI), fraction of absorbed photosynthetically active radiation (FAPAR) and fraction of vegetation cover (FCover) v1 for the same period at 300 m resolution and a temporal coverage of 10 days were obtained through Copernicus Global Land Service (CGLS; http://land.copernicus.eu). MODIS and Copernic data were averaged over a year, serving as a basis for further modeling.

Total Copper Background in Topsoil

Total copper background content in topsoil (Cu_{bgd}) was defined as the geogenic and atmospheric non-agricultural anthropogenic Cu_t. The Cu_{bgd} across Europe was estimated using 25,122 field samples (LUCAS, GEMAS, BGR and data from Rodriguez Martin, Arias and Grau [3]) collected from non-vineyard areas across Europe ($Cu_{bgd,obs}$; Figure S2.A) delimited from the Corine Land Cover (CLC 2012; v18.5.1; http://land.copernicus.eu). 0.4% of these field samples showed more than 200 mg Cu kg^{-1}, and were consequently considered as suspiciously contaminated and subsequently removed from the Cu_{bgd} calculation. The 25,022 remaining field samples were used to calculate the Cu_{bgd}. For each vineyard pixel, a stepwise buffer increase of 5 km radius was computed until the buffer contained a minimal number of three $Cu_{bgd,obs}$. Then $Cu_{bgd,obs}$ within the buffer were averaged and used as the Cu_{bgd} for the corresponding pixel. The mean Cu_{bgd} is 26.7 mg Cu kg^{-1} with a minimum and maximum of 1.12 and 196 mg Cu kg^{-1} for Europe, respectively. Because Cu_{bgd} was estimated using variable buffer sizes, comparison of the Cu_{bgd} might be subsequently affected by this spatial discrepancy leading to an underestimation or overestimation in Cu_{bgd}. To evaluate if Cu_{bgd} are stable for the vineyard pixels, we calculated a modified $Cu_{bgd,i}$ for every stepwise buffer increase I (i.e. from 5 km to 100 km; Figure S2.B). Then, the modified $Cu_{bgd,i}$ was compared to the original one (ΔCu background = Cu_{bgd} -$Cu_{bgd,i}$; Figure S2.B). The ΔCu_{bgd} did not change much over space (ΔCu_{bgd} = 5.3 ± 13.8 mg Cu kg^{-1}). As a result, Cu_{bgd} was assumed to not be significantly affected by the buffer size and could be compared across European vineyards.

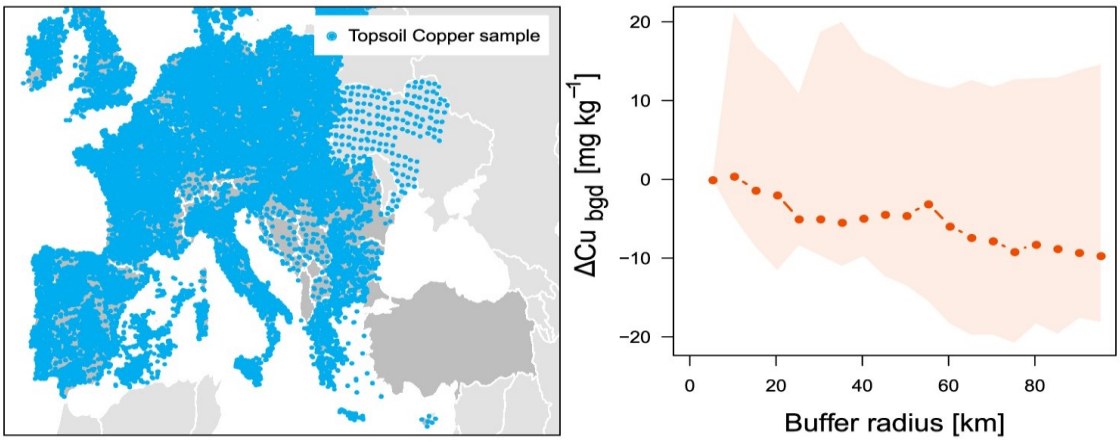

Figure S2: A) Total copper datasets in topsoil from non-vineyard area across Europe (n = 25,122; LUCAS, GEMAS, BGR and data from Rodriguez Martin, Arias and Grau [3]) used to compute the Cu_{bgd} estimation. B) ΔCu_{bgd} average as a function of the stepwise buffer radius. $\Delta Cu_{bgd} = Cu_{bgd} - Cu_{bgd,i}$, where Cu_{bgd} is the original estimation computed at one pixel and $Cu_{bbgd,i}$ the stepwise estimation for the modified buffer radius.

Soil Loss Rate

The soil loss rates in European vineyards were estimated using the RUSLE-based[22] modeling platform Global Soil Erosion Modeling (GloSEM),[23] a semi-empirical model which is an example of a detachment limited model. This means that although the overland flow may theoretically transport an infinite amount of sediment, the quantity of sediment available to be moved is actually limited by the soil detachment capacity defined by the erosivity of the rainfall. The soil erosion (Mg ha^{-1} yr^{-1}) resulting from sheet and rill erosion processes is given by the following multiplicative eq. S2,

$$A = R \times L \times S \times K \times C \times P \quad (S2)$$

where: A (Mg ha^{-1} yr^{-1}) is the annual average soil erosion. R (MJ mm h^{-1} ha^{-1} yr^{-1}) is the rainfall-runoff erosivity factor, a numerical descriptor of the rainfall's ability to erode soil. It expresses the kinetic energy of the raindrop's impact and the rate of associated runoff and was estimated using hourly and sub hourly rainfall data acquired from 1,728 samples across the study area. The soil erodibility factor K (Mg h MJ^{-1} mm^{-1}) is an empirical parameter that is measured based on intrinsic soil properties such as texture, organic matter, structure and permeability of the topsoil profile.

The LUCAS top soil data were used to compute the K factor.[24] The slope length factor L (unitless) and the slope steepness factor S (unitless) were derived from the 90m SRTM DEM. The land cover and management factor C (unitless) for European vineyards was computed by taking into consideration a potential range from 0.15 to 0.45 based on the annual ground vegetation density. The status of the vegetation was quantified according to biophysical parameters derived from MERIS satellite images, i.e., FCover. The soil conservation or prevention practices factor P (unitless) was spatially described using the data provided by Panagos, et al. [25]

Predicted No-Effect Noncentration (PNEC) on Soil

Predicted no-effect concentration (PNEC) of Cu_t in soil was previously measured by Smolder et al.[26, 27] using toxicity tests under laboratory conditions on 75 species across three terrestrial trophic levels (plants, invertebrates and microorganisms) and for six different soils. The six standard soils had specific soil properties (i.e., pH, % C_{org} and % Clay) covering 96% of the variation observed in European vineyard soils for pH, C_{org} and Clay respectively (Figure S3). For each of the 75 species and soils, the toxicity test measured concentrations yielding 50% of the inhibition (EC_{50}). PNEC was corrected to account for i) the difference between laboratory and field conditions (further referred to as LF with a value of two for Cu_t)[28, 29] and ii) the toxicity threshold as a function of soil properties, made according to Smolder et al.[26, 27] For each pixel of the European vineyards, the PNEC was estimated using soil standards with the minimum Euclidian distance between soil characteristics of the pixel and one of the six standard soils used by Smolder et al.[26, 27]. The estimated PNEC ranged from 30 to 290 (average = 127 mg Cu kg^{-1}) for the European vineyard soils.

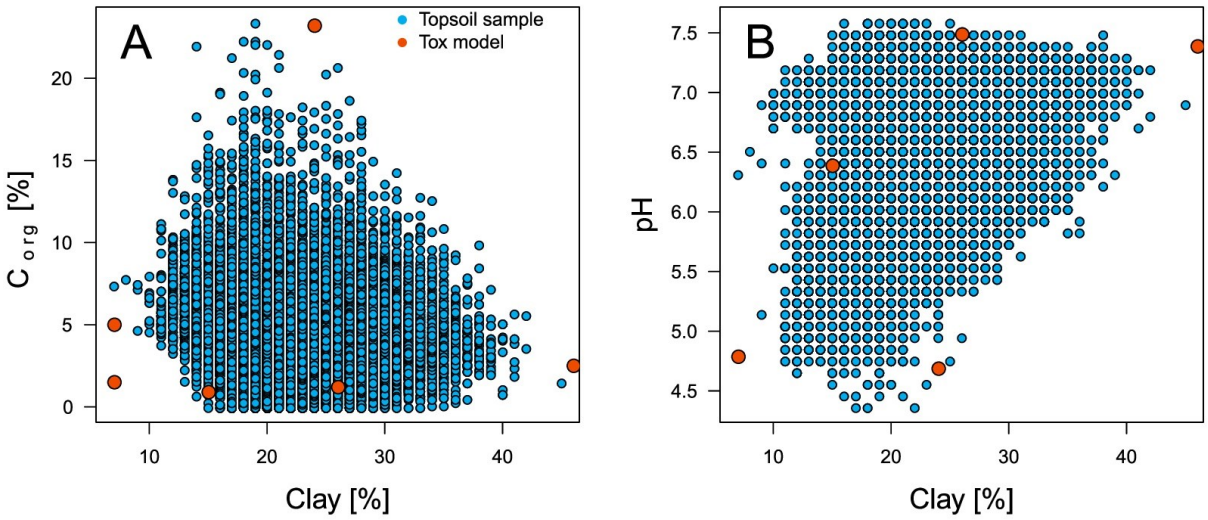

Figure S3: Topsoil properties used to evaluate the predicted no-effect concentration (PNEC) of total copper content in vineyard topsoil (Cu_t). A) Soil organic carbon content (C_{org}) as a function of the clay content. B) H_2O pH as a function of the clay content. In blue, topsoil sample properties associated with the 1,202 data of total copper content in topsoil ($Cu_{t,\,dat}$; Table S1). In orange, properties of the six reference soils used to build the PNEC model on Cu_t by Smolder et al.[27] Note in B, two soils have similar pH and clay content, by consequence only five dots appear on the graph.

The Soluble Fraction (f_{sol}) of the Cu

The relationship between the soluble fraction (f_{sol}) of Cu_t, C_{org} and pH was established by McBride et al. [30] The "semi-empirical" model was developed under environmental condition using soil with the following properties: (i) pH between 5 to 8 and (ii) a range of Cu_t less than the loading limit established in the USEPA 503 (<1,500 kg Cu ha^{-1}), (iii) long-term aged equilibrate soil , more than 10 years, and (iv) dissolved copper was measured in water or diluted salt ($CaCl_2$ 0.01M) to avoid unrealistic extraction procedure. Briefly, the model was built as follow: In our study, the soluble fraction (f_{sol}) of the Cu_f application was defined following eq. S3,

$$\begin{aligned} f_{sol} &= Cu_{sol}/Cu_{Pred} \mid t = 0 \\ f_{sol} &= Cu_{sol}/Cu_{t,in} \mid t > 0 \end{aligned} \qquad \text{(S3)}$$

where Cu_{sol} and $Cu_{t,in}$ represent the concentration of the dissolved and the total copper in topsoil (mg Cu kg^{-1}) per year of Cu_f application t, respectively. Cu_{Pred} represent the average predictions of Cu_t, previously calculate. Cu_{sol} was estimated using the semi-empirical equation based on the metal complexation theory [30], eq. S4:

$$\log(Cu_{sol}) = 1.42 - 0.10 \times \mathrm{pH} + 094 \times \log(Cu_{t,in}) - 0.68 \times \log\left|C_{org}\right| \quad \text{(S4)}$$

where $\left|C_{org}\right|$ represents the organic carbon in g kg^{-1}. The goodness of the S4 fit was $R^2 = 0.851$ and the interval of confidence (2σ) estimated on the McBride data was 0.617. However, in soils where $\left|C_{org}\right|$ <9 g C kg^{-1} the f_{sol} were overestimated. Therefore, the eq S5 was used:

$$\log(Cu_{sol}) = 0.699 - 0.11 \times \mathrm{pH} + 0.86 \times \log(Cu_{t,in}) \quad \text{(S5)}$$

The goodness of the S5 fit was $R^2 = 0.877$ and the interval of confidence (2σ) estimated on the McBride data was 0.507.

We used eq. S4 & S5 to compute a soluble fraction for all grid cell using the grid soil pH and C_{org} (https://soilgrids.org/, series M_sl1, v18.4.2017).[10] We hypothesize that Cu_t is homogeneously distributed within the 0 – 30 cm soil depth. As a result, the relationship between Cu application dose (Cu_{input}; kg Cu ha^{-1} $year^{-1}$) and $Cu_{t,in}$ is expressed in eq. S6,

$$Cu_{t,in} = Cu_{input}/(BD \times h_{topsoil}) \quad \text{(S6)}$$

where BD is the soil bulk density (https://soilgrids.org/, series M_sl1, v18.4.2017)[10] and $h_{topsoil}$ the average topsoil thickness (30 cm) where Cu_t mainly accumulates. The estimated soluble fraction (f_{sol}) ranges between 29 and 95% (95% of the confidence interval), depending on soil properties, according to the semi-empirical model.

Data Pre-Processing

P_{sum}, T_{av}, PET, ET, AI and R-factor and LAI, FAPAR, F_{Cover} were only available on a larger resolution (between 0.3-1 km), which was lower than the 250 m resolution of most other predictive variables. This resolution discrepancy is acceptable when considering the climatic variability in space. P_{sum}, T_{av}, PET, ET, AI and R-factor and LAI, FAPAR, F_{Cover} were thus resampled at a resolution of 250 m. $Cu_{t,\,dat}$ geolocation (i.e. x and y coordinates) were used to extract values of

each predictive variable on the corresponding grid cell pixel. Then extracted values were used to select variables and calibrate the model. Because machine learning algorithm and principal components analysis (PCA) assume normality, normality was tested for all predictive variables (examination of the skewness of the extracted values). Variables were deemed normally distributed if the |skew| <1.[31] Variables that were not normally distributed were either inverse, log10 or square root transformed, which also reduced the influence of outliers. Finally, all predictive variables were centered using z-scores (eq. S7),

$$z_i = \frac{x_i - \bar{x}}{SD} \quad \text{(S7)}$$

where *SD* is the standard deviation, $\bar{x}$ the average and x_i each data point of the predictive variable.

Table S3. Values of predicted and predictive variables for the model. a) Predicted total copper content in the topsoil (Cu_{pred}) for the model calibration extracted from the 1,202 field samples. b) List of the 10 predictive variables used in the models.

Variable	Unit	Resolution (m)	Years of coverage	transf. [a]	$\bar{x}$ [b]	SD [b]	Ref.
a) Predicted variable							
Total copper content in topsoil	mg kg^{-1}	1-100	1989-2016	log	1.46	0.52	Table S1
b) Predictive variables							
Annual rainfall (rainfall)	mm	1000	1960-1990	log	2.73	0.25	[11]
Aridity Index (AI)	unitless	1000	1950-2000	10000×log	−2.24	0.34	[12]
Soil organic carbon content (C_{org})	%	250	1990-2016	no	75.86	60.70	[10]
Silt fraction	%	250	1990-2016	no	37.87	9.42	[10]
Soil moisture content (moisture)	%	250	1990-2016	no	47.19	6.10	[10]
Annual average temperature (T_{av})	°C	1000	1960-1990	no	11.16	4.42	[11]
Enhanced vegetation index (EVI)	unitless	250	2013-2016	1/x	4.59	3.72×10^{3}	[15]
Soil pH	unitless	250	1990-2016	no	67.46	8.76	[10]
Slope	degree	250	2000	log	0.21	1.18	RSAGA v0.94-5[16]
Clay fraction	%	250	1990-2016	no	21.40	7.41	[10]

[a] transformation used to satisfy a normal distribution of the variable, [b] $\bar{x}$ average and SD standard deviation of the transformed variable used to center the normal (Z-score; eq. S7).

Variable Selection

We statistically evaluate whether each predictive variable should be retained to compose the most appropriate set of predictive variables for the final model following the criteria detailed below. Four independent analyses were run in parallel to evaluate the multi-collinearity and spatial variability of the topsoil dataset. (i) Pearson correlation was used to compare topsoil properties obtained from Soilgrid with the field data collected from 26,324 field measurements across Europe assembled from different datasets (LUCAS, GEMAS, BGR and data from Rodriguez Martin, Arias and Grau [3]; Table S1). This comparison concerned H_2O pH values, soil texture content (% clay, silt and sand), CEC and C_{org}. Soilgrid variables that were weakly correlated (i.e., $|R^2| < 0.6$) with field Cu measurements were not considered for further analysis. (ii) Pearson correlation was used to test the collinearity between all possible pairs of predictive variables within the topsoil extracted dataset. One predictive variable for each highly correlated pair (i.e., $|R^2| > 0.7$) was taken for further analysis. (iii) Paired t-test was used to compare the representativeness of predictive variables between extracted values from the 1,202 $Cu_{t,\,dat}$ geolocation and 50,000 randomly extracted data from each grid cell within the European vineyard domain. Predictive variables that significantly differed among the two datasets (t-test; $p > 0.05$) were not considered as spatially representative of the European vineyard domain and were consequently not used in further analyses. (iv) PCA is a commonly used variable selection technique to reduce the number of variables.[32] Contribution and relationship analyses of each predictive variable on multiple scales were made under a distance biplot using the two first principal's axes. Comparison of the Euclidean distance (D) allowed to remove variables that contain little information ($|D| < 0.3$).[33]. PCA correlation matrix (CO) was also used to evaluate the contribution of predictive variables. First, the number of relevant principal axes (n_{PA}) were calculated based on the Kaiser method.[34] Then the values of importance ($V.imp_j$) for each predictive variable j were calculated following eq. S8,

$$V.imp_j = \sqrt{\sum_{i=1}^{n_{PA}} CO_{j,i}} \quad \text{(S8)}$$

where $CO_{j,i}$ are the values on the CO matrix for the j predictive variable on the principal axes i.

Machine Learning Building, Performance and Evaluation

Three independent machine learning techniques i.e., neural networks (R package nnet v 7.3-12), random forest (R package randomForest v.4.6-12) and bagging tree (R package ipred v.0.9-5) were averaged into an ensemble (ENS) to provide the final prediction. First, parameters extracted from random forest and bagging tree were tuned with the extracted dataset to optimize parameter values with respect to the out-of-bag (OOB) error estimate.[35] Otherwise recommended parameter values provided by the machine learning package were used. For each technique, 1000 independent models were generated with as many as 10,000 iterations for fitting the Neural Networks and 1,000 trees for the "tree" techniques (i.e., random forest and bagging tree).

Our final calibrated ensemble model (ENS) is constituted by 1,000 independent models for each of the three machine learning techniques (i.e., 3,000 in total). For all grid cells of the European vineyard (Figure S1), Cu_{pred} was calculated over the 3,000 models and averaged to provide the final prediction. Model performance was accomplished using a 10-fold cross-validation (CV) and the result evaluated on the regression of $Cu_{t,\,dat}$ in vineyard soils in function of the Cu_{pred} to compare slope and intercept parameters against the 1:1 line as recommended by Piñeiro, et al.[36] to avoid potential misinterpretations and more accurately apply an evaluation of uncertainty.

Model Transferability in Space

Transferability in space evaluation of the model was computed to determine if the model was consistent in space or changed across different domains represented by various environmental conditions.[37] Good transferability in space is defined when a model built in one domain is able to accurately predict the field data values outside of the domains.[38] First, topsoil samples (–9.2° to 27° E; 35° to 50° N) were divided into four equivalent geographical domains defined by a cut off at the mean x and y coordinate of the extent. For each domain, a sub-ensemble model using the three techniques was calibrated and averaged similarly to the predictive model. Then, the sub-ensemble model of one domain was used to predict the Cu_t inside the three other domains. Transferability was then evaluated four times based on the R^2 of the regression of $Cu_{t,\,dat}$ in vineyard soil as a function of the Cu_{pred}.

Influence of heterogeneous sampling depths and time on the Cu prediction

Influence of the various sampling depths (0-20 vs 0-30 cm) and different times of sampling (1989-1999, 2000-2009 vs 2010-2016) was evaluated to determine if these heterogeneous datasets might induce over- or underestimation of the final prediction. Topsoil data were classified independently by depth or by time. For each class, a sub-ensemble model using the three techniques was calibrated and averaged similarly to the predictive model. However, the number of samples for the classes 0-30 cm depth (n=87), 1989-99 (n=76) and 2010-2016 (n=222) were not sufficient to build a robust model. Then, two sub-ensemble model, 0-20 cm and 2000-2009 were used to predict the Cu_t inside the other class, for example model 0-20 cm was used to predict data of 0-30 cm and in the entire dataset. Over- or underestimation of the final prediction were then evaluated based on the R^2 of the regression of $Cu_{t,\,dat}$ in vineyard soil as a function of the Cu_{pred}.

Averaged Relative Importance and Sensitivity Analysis

The averaged relative importance (ARI) of each variable in the model was calculated with an input permutation technique.[39] Sensitivity analysis was used to investigate the impact of each predictive variable on Cu_{pred}. We investigated the response of 1,000 independent models for the three machine learning techniques (i.e., 3,000 models in total). The models were previously built by computing a one-factor-at-a-time (OFAT) sensitivity analysis. In OFAT, only one input predictive variable is changed at one time, holding the remaining variables at their median. 10,000 regular steps were allowed for each input variable between the minimum and maximum observed value across the European vineyard range.

Scaling Analysis.

Sensitivity of model predictions to grid resolution was evaluated using systematic resampling. All predictive variable and topsoil data were successively resampled and averaged on a coarser grid cell using a stepwise resolution increase from 400 to 1000 m. At each resolution, we built an ensemble of 1,000 independent models for the three machine learning techniques (i.e., 3,000 models in total). The ensemble was calibrated and averaged similar to the predictive model. Calibrated models were evaluated using R^2, as a metric of precision, and normalized root-mean-square error (RMSE) as a metric of bias from the regression of $Cu_{t,\,dat}$ in vineyard soil as a function of the Cu_{pred}.

Scenarios of Copper-Based Fungicide Application

Four scenarios of Cu_f application dose were compared using a simple mass balance approach defined as,

$$m_{Cu,t+1} = m_{Cu,t} + m_{Cu,f} - m_{Cu,leach} - m_{Cu,runoff} \quad \text{(S9)}$$

where the end of the year mass of copper ($m_{Cu,t+1}$) is egal to the current mass ($m_{Cu,t}$) within the topsoil plus the masse of Cu_f applied over the year ($m_{Cu,f}$) minus the mass export by leaching ($m_{Cu,leach}$) and runoff ($m_{Cu,runoff}$). Assuming no change over time of the considered topsoil volume and a homogenous topsoil layer, the change of mass could be considered as a change of Cu_t and rewrite based on eq 1-2 and S3-6 as,

$$Cu_{pred,t+1} = Cu_{pred,t} + (APP_{Cu,f} - Cu_{pred,t} \times net\ export - Cu_{input} \times f_{sol})/(BD \times h_{topsoil}) \quad \text{(S10)}$$

where Cu_{Pred} is the average predictions of Cu_t (mg Cu kg^{-1}), $APP_{Cu,f}$ the yearly application dose of Cu_f (kg Cu ha^{-1} $year^{-1}$), f_{sol} the soluble fraction previous calculated in eq S3-6, *net export* of Cu (kg of Cu ha^{-1} yr^{-1}) assumed to occur during runoff events in the form of soil particle-bound, *BD* the soil bulk density (kg m^{-3})[10] and $h_{topsoil}$ the average topsoil thickness (30 cm). Note that f_{sol} from the Cu_{Pred}t is considered once at t=0 (eq. S3). The eq S10 was used to compute scenario with $APP_{Cu,f}$ equal to 2, 4 and 8 kg Cu ha^{-1} $year^{-1}$ over a large time scale. For each, pixel, the scenario was applied until the Cu_{pred} at year t+1 reach the predicted no-effect concentration (see building spatial variables - PNEC).

Uncertainty analysis of the scenarios

Uncertainties associated to the four scenarios were estimated by computing a numerical propagation of individual uncertainty associated to the 10 variables governing Cu_t in vineyard. The table S4 summarizes the uncertainty calculation.

Table S4. Uncertainties calculation associated the variables used in our modelling framework and selected to numerical calculation of uncertainty analysis on Cu_{pred}, Cu_{bgd}, net acc, net exp and PNEC.

Variables group or calculation step	Uncertainty calculation
Soil grid (pH, C_{org}, clay, silt)	Lower and upper limits of a 90% prediction interval.[10]
Aridity, moisture, EVI, slope	Not available
Rainfall, T_{av}	SD of the year between 1960-1990
Runoff	Estimated using a Markov Chain Monte Carlo (MCMC) approach.[23]
Cu_{pred}	Lower and upper limits of all combinations of lower and upper values for the selected variables, i.e. $2^6 = 64$, a 90% prediction interval was considered.
PNEC	Uncertainty of pH, C_{org}, clay is propagated trough the model of Smolder et al.[26, 27]
Cu_{bgd}	SD as calculation described in the section above.
net acc., net exp.	Uncertainty of corresponding variables is propagated trough eq. 1 and 2

Then, the final uncertainty is propagated trough the scenario using Eq. 3 & S9-10 with all combination of lower and upper values for the selected variables, by producing a lower and upper estimations of the scenario's uncertainty.

Supporting Results and Discussion

Model Performance, Evaluation and Comparison to Previous Copper Prediction

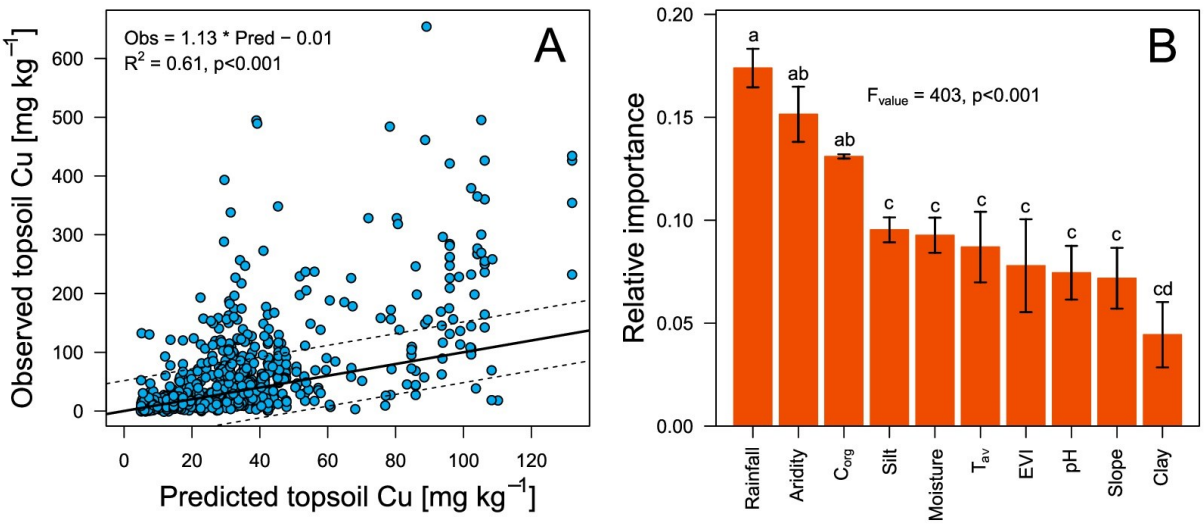

Figure S4. Performance and variable contribution to the predictive models. (A) Scatter plot of the observed (i.e, measured data) and predicted total copper content in the topsoil (n = 1,202) based on the ensemble model including model performance. The black solid line represents the slope at a ratio of 1:1 with 95% (2σ) interval confidence in dash lines. (B) Average variable contributions for the ensemble of model (i.e., 3,000 independent models) derived from three machine learning techniques. The error bar represents two times the standard deviation across all model iterations. A one-way ANOVA was used to determine statistical differences of average contribution of variables (statistical values are given in the figure). A Tukey test was used to evaluate statistical differences among variables (α = 0.05) (the significative differences between groups are indicated by different letters; $p \leq 0.01$ for all variables).

On a reduced data set within the range 0 to 130 mg Cu kg^{-1}, the linear relationship of $Cu_{t,dat}$ as a function of Cu_{pred} was high ($R^2 = 0.83$). However, Machine learning algorithms are known to be more robust but with lower reproducibility of observed outliers than classic regression algorithms.[40] The algorithms of machine learning might consider outliers as false or noise data

during the training stage. One possible improvement strategy consists in removing outliers from the dataset to enhance the prediction [41] however more advanced algorithms might have a better capacity of handling this problem especially with the next generation of machine learning,[42] e.g., deep machine learning. Despite an underestimation, our predictions were significantly higher than the two previously reported Cu_{pred} (average of 15.3 [43] and 19.5 [44] mg Cu kg^{-1}) when comparing common pixels locations (t-test; p-value <0.01). These differences can be explained by previous predictions targeting the overall Cu_t for different land uses. The small proportion of vineyard area (0.9% of the European land use) likely reduces the weight of vineyard land use in the model leading to an underestimation of Cu_{pred} in vineyards. In addition to this, the two previous predictions were computed with 1 km and 500 m grid cell resolution. Coarse spatial resolution has been demonstrated to limit accuracy of the prediction in previous works.[10, 45, 46] The grid resolution impact was clearly shown by the scaling analysis (see above method) with a significant decrease in both accuracy and precision of the model when the data were degraded to the lowest resolution (Figure S5).

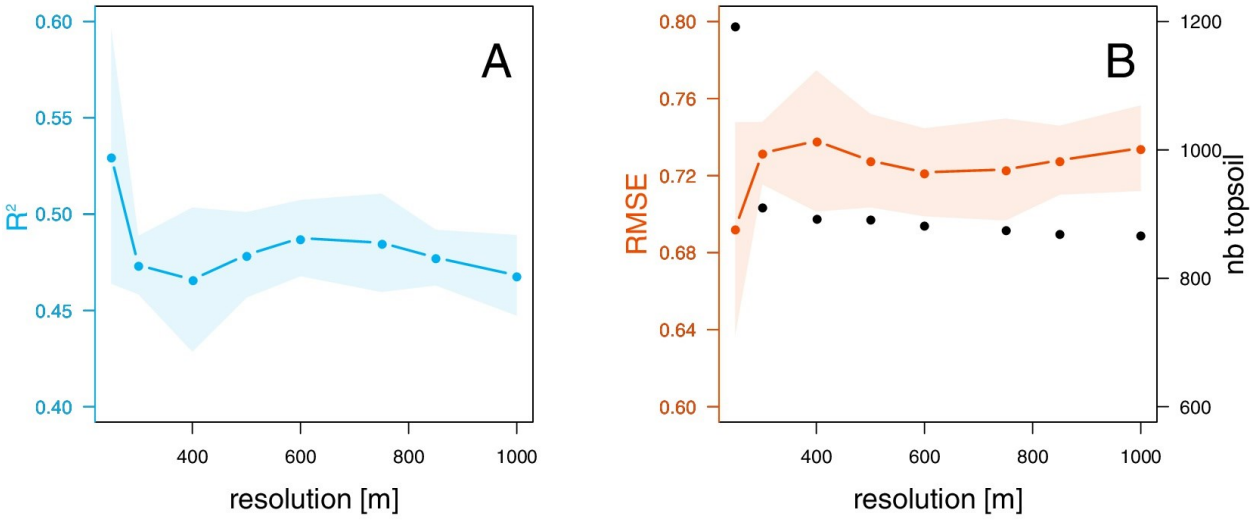

Figure S5. Grid resampling impact on the model results. Resampling was performed by averaging data under the initial resolution to a coarser resolution. Average performances of the 3,000 models for the A) Accuracy with R^2 and B) Precision with root mean square error (RMSE) were reported as a function of the resolution. The standard deviations for these two metrics under the 3,000

models are represented by the shaded polygon. Decrease of the number of topsoil data due to averaging is also reported by the black points in graph B.

Regional Predictions

Vineyard proportion is highest in the Mediterranean region (Table S5) where the climate is the most suitable for viticulture (warm and dry summers; cool and wet winters).[47] Compared to the Alpine area, the Mediterranean is not affected by high annual rainfall, consequently low to moderate doses and application frequencies of Cu_f are required to prevent fungal disease.

Table S5. Predicted Cu_t (Cu_{pred}) averaged per European region. Regions follow the European biogeographic regions designated in 2016 (https://www.eea.europa.eu). Corresponding proportions of each region to the total European vineyard area are reported.

Region	Cu_{pred} average (min-max) (mg Cu kg^{-1})	Vineyard Proportion (%)
Alpine	57 (11 - 130)	1
Anatolian	16 (6 - 37)	2
Atlantic	34 (15 - 86)	10
Black Sea	19 (9 - 49)	1
Continental	32 (8 - 130)	23
Mediterranean	22 (4 - 127)	56
Pannonian	20 (8 - 80)	5
Steppic	14 (9 - 35)	2

Limits of the Approach

In the last decade, significant improvement was seen in the resolution, accuracy and coverages of spatial data compared to previous work on the field, with even more advancements in the near future.[48] However, current Corine Land Cover data are the only large scale land cover containing vineyard land use information. We estimated on a regional land use cover (Switzerland) that the Corine Land Cover, originally presented at a 100 m grid cell resolution, presents a ± 150 m inaccuracy in border delineation between land use types in some regions. To resolve this issue, regional land use cover, sometimes more precise, could be merged together to produce a global

land cover. However, only few countries (e.g., France, Germany, Switzerland, Austria) have accurate regional land use covers including vineyard as a land use type. Moreover, some climate variables are only available at 500 or 1000 m grid cell resolution which might negatively impact prediction reliability in areas where local climate variations are important, i.e. mountain areas.

As previously demonstrated,[10, 37] predictions were also consolidated by using a compilation of different datasets significantly increasing the amount of training data (i.e., Cu_t data), improving the learning experience to build the algorithm and thus the accuracy of predictions. For instance, the global land use stratified scheme corrected per area of country, such as the LUCAS topsoil database,[1] gives a comprehensive representation of the situation and likely captures the extremes of Cu contamination per area for the entirety of European vineyards. However, from a modeling point of view the number of training data could be not enough 'experiences' derived from the main vineyard regions, i.e. the Mediterranean region, to accommodate predictions of the overall Cu_t.

Supporting References

1. Tóth, G.; Jones, A.; Montanarella, L. *LUCAS topsoil survey methodology, data and results*; Joint Research Centre: Ispra, Italy, 2013; p 154.
2. Reimann, C.; Birke, M.; Demetriades, A.; Filzmoser, P.; O'Connor, P., *Distribution of elements/parameters in agricultural and grazing land soil in Europe. Chemistry of Europe's agricultural soils. Part A: methodology and interpretation of the GEMAS data set*. Bundesanstalt für Geowissenschaften und Rohstoffe: Hannover, Germany, 2014; p 523.
3. Rodriguez Martin, J. A.; Arias, L.; Grau, J., *Metales pesados, materia orgánica y otros parámetros de los suelos agrícolas y pastos de España*. 2009.
4. Bednářová, Z.; Kalina, J.; Hájek, O.; Sáňka, M.; Komprdová, K., Spatial distribution and risk assessment of metals in agricultural soils. *Geoderma* **2016,** *284*, 113–121.
5. Ramos, M. C., Metals in vineyard soils of the Penedès area (NE Spain) after compost application. *J. Environ. Manage.* **2006,** *78*, (3), 209–215.
6. Ramos, M. C.; Romero, M. P., Effects of soil characteristics and leaf thinning on micronutrient uptake and redistribution in Cabernet Sauvignon. *Vitis* **2016,** *55*, (3), 8.
7. Tülp, H. C. *Die Kupfergehalte in Böden und Vegetation Bewirtschafteter und Aufgelassener Weinberge der Mosel*; University of Trier: Trier, 2004; p 66.
8. Rusjan, D.; Strlič, M.; Pucko, D.; Korošec-Koruza, Z., Copper accumulation regarding the soil characteristics in Sub-Mediterranean vineyards of Slovenia. *Geoderma* **2007,** *141*, (1–2), 111–118.

9. Rusjan, D.; Strlič, M.; Pucko, D.; Šelih, V. S.; Korošec-Koruza, Z., Vineyard soil characteristics related to content of transition metals in a sub-Mediterranean winegrowing region of Slovenia. *Geoderma* **2006,** *136*, (3–4), 930–936.

10. 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., SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One* **2017,** *12*, (2), e0169748.

11. Hijmans, R. J.; Cameron, S. E.; Parra, J. L.; Jones, P. G.; Jarvis, A., Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **2005,** *25*, (15), 1965–1978.

12. Antonio, T.; Robert, Z. Global aridity index and potential evapotranspiration (ET0) climate database v2. https://doi.org/10.6084/m9.figshare.7504448.v3

13. Panagos, P.; Ballabio, C.; Borrelli, P.; Meusburger, K.; Klik, A.; Rousseva, S.; Tadić, M. P.; Michaelides, S.; Hrabalíková, M.; Olsen, P.; Aalto, J.; Lakatos, M.; Rymszewicz, A.; Dumitrescu, A.; Beguería, S.; Alewell, C., Rainfall erosivity in Europe. *Sci. Total Environ.* **2015,** *511*, 801–814.

14. Panagos, P.; Borrelli, P.; Meusburger, K.; Yu, B.; Klik, A.; Jae Lim, K.; Yang, J. E.; Ni, J.; Miao, C.; Chattopadhyay, N.; Sadeghi, S. H.; Hazbavi, Z.; Zabihi, M.; Larionov, G. A.; Krasnov, S. F.; Gorobets, A. V.; Levi, Y.; Erpul, G.; Birkel, C.; Hoyos, N.; Naipal, V.; Oliveira, P. T. S.; Bonilla, C. A.; Meddi, M.; Nel, W.; Al Dashti, H.; Boni, M.; Diodato, N.; Van Oost, K.; Nearing, M.; Ballabio, C., Global rainfall erosivity assessment based on high-temporal resolution rainfall records. *Sci. Rep.* **2017,** *7*, (1), 4175.

15. Didan., K. MOD13Q1 MODIS/Terra vegetation indices 16-day L3 global 250m SIN grid V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD13Q1.006

16. Zevenbergen, L. W.; Thorne, C. R., Quantitative analysis of land surface topography. *Earth Surf. Proc. Land.* **1987,** *12*, (1), 47-56.

17. O'Callaghan, J. F.; Mark, D. M., The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process.* **1984,** *28*, (3), 323–344.

18. Desmet, P. J. J.; Govers, G., A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. *J. Soil Weter Conserv.* **1996,** *51*, (5), 427–433.

19. Kumar, L.; Skidmore, A. K.; Knowles, E., Modelling topographic variation in solar radiation in a GIS environment. *Int. J. Geogr. Inf. Sci.* **1997,** *11*, (5), 475–497.

20. Hargreaves, G. L.; Hargreaves, G. H.; Riley, J. P., Irrigation water requirements for Senegal River basin. *J. Irrig. Drain. Eng.* **1985,** *111*, (3), 265–275.

21. Zomer, R. J.; Trabucco, A.; Bossio, D. A.; Verchot, L. V., Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agric., Ecosyst. Environ.* **2008,** *126*, (1–2), 67–80.

22. Renard, K. G.; Foster, G. R.; Weesies, G. A.; McCool, D. K.; Yoder, D. C. *Predicting soil erosion by water: A guide to conservation planning with the revised universal soil loss equation (RUSLE)*; USDA: 1997; p 49.

23. Borrelli, P.; Robinson, D. A.; Fleischer, L. R.; Lugato, E.; Ballabio, C.; Alewell, C.; Meusburger, K.; Modugno, S.; Schütt, B.; Ferro, V.; Bagarello, V.; Oost, K. V.; Montanarella, L.; Panagos, P., An assessment of the global impact of 21st century land use change on soil erosion. *Nat. Commun.* **2017,** *8*, (1), 2013.

24. Panagos, P.; Meusburger, K.; Ballabio, C.; Borrelli, P.; Alewell, C., Soil erodibility in Europe: A high-resolution dataset based on LUCAS. *Sci. Total Environ.* **2014,** *479–480*, 189–200.

25. Panagos, P.; Borrelli, P.; Meusburger, K.; van der Zanden, E. H.; Poesen, J.; Alewell, C., Modelling the effect of support practices (P-factor) on the reduction of soil erosion by water at European scale. *Environ. Sci. Policy* **2015,** *51*, 23–34.

26. ECHA, Characterisation of dose [concentration]-response for environment. In *Guidance on Information Requirements and Chemical Safety Assessment*, European Chemicals Agency: Helsinki, Finland, 2008.

27. Smolders, E.; Oorts, K.; Van Sprang, P.; Schoeters, I.; Janssen, C. R.; McGrath, S. P.; McLaughlin, M. J., Toxicity of trace metals in soil as affected by soil type and aging after contamination: Using calibrated bioavailability models to set ecological soil standards. *Environ. Toxicol. Chem.* **2009,** *28*, (8), 1633–1642.

28. Oorts, K.; Schoreters, I., Use of monitoring data for risk assesment of metals in soil under the european REACH regulation. In *Distribution of elements/parameters in agricultural and grazing land soil in Europe. Chemistry of Europe's agricultural soils. Part B: General Background Information and Further Analysis of the GEMAS Data Set*, Bundesanstalt für Geowissenschaften und Rohstoffe: Hannover, Germany, 2014.

29. De Schamphelaere, K. A. C.; Janssen, C. R., Bioavailability models for predicting copper toxicity to freshwater green microalgae as a function of water chemistry. *Environ. Sci. Technol.* **2006,** *40*, (14), 4514–4522.

30. McBride, M.; Sauve, S.; Hendershot, W., Solubility control of Cu, Zn, Cd and Pb in contaminated soils. *Eur. J. Soil Sci.* **1997,** *48*, (2), 337–346.

31. McCune, B. P.; Grace, J., *Analysis of ecological communities*. MJM Software Design: Gleneden Beach, OR, 2002; Vol. 289.

32. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P., *Numerical recipes in C: The art of scientific computing*. Cambridge University Press: 2002.

33. Legendre, P.; Legendre, L., *Numerical ecology*. Elsevier: 1998.

34. Kaiser, H. F., The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* **1960,** *20*, (1), 141–151.

35. James, G.; Witten, D.; Hastie, T.; Tibshirani, R., *An introduction to statistical Learning with applications in R*. Springer: 2015.

36. Piñeiro, G.; Perelman, S.; Guerschman, J. P.; Paruelo, J. M., How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecol. Model.* **2008,** *216*, (3–4), 316–322.

37. Jones, G. D.; Droz, B.; Greve, P.; Gottschalk, P.; Poffet, D.; McGrath, S. P.; Seneviratne, S. I.; Smith, P.; Winkel, L. H. E., Selenium deficiency risk predicted to increase under future climate change. *Proceedings of the National Academy of Sciences* **2017,** *114*, (11).

38. Randin, C. F.; Dirnbock, T.; Dullinger, S.; Zimmermann, N. E.; Zappa, M.; Guisan, A., Are niche-based species distribution models transferable in space? *J. Biogeogr.* **2006,** *33*, (10), 1689–1703.

39. Satizábal M, H. F.; Pérez-Uribe, A., Relevance metrics to reduce input dimensions in artificial neural networks. In *Artificial Neural Networks – ICANN 2007*, de Sá, J.; Alexandre, L.; Duch, W.; Mandic, D., Eds. Springer Berlin Heidelberg: 2007; Vol. 4668, pp 39–48.

40. Zhu, J.; Ge, Z.; Song, Z.; Gao, F., Review and big data perspectives on robust data mining approaches for industrial process modeling with outliers and missing data. *Annu. Rev. Control* **2018,** *46*, 107–133.

41. Elith, J.; Leathwick, J. R.; Hastie, T., A working guide to boosted regression trees. *J. Anim. Ecol.* **2008,** *77*, (4), 802–813.

42. Fadlullah, Z. M.; Tang, F.; Mao, B.; Kato, N.; Akashi, O.; Inoue, T.; Mizutani, K., State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems. *IEEE Commun. Surv.* **2017,** *19*, (4), 2432–2455.

43. Tóth, G.; Hermann, T.; Szatmári, G.; Pásztor, L., Maps of heavy metals in the soils of the European Union and proposed priority areas for detailed assessment. *Sci. Total Environ.* **2016,** *565*, 1054–1062.

44. Ballabio, C.; Panagos, P.; Lugato, E.; Huang, J.-H.; Orgiazzi, A.; Jones, A.; Fernández-Ugalde, O.; Borrelli, P.; Montanarella, L., Copper distribution in European topsoils: An assessment based on LUCAS soil survey. *Sci. Total Environ.* **2018,** *636*, 282–298.

45. Lado, L. R.; Hengl, T.; Reuter, H. I., Heavy metals in European soils: A geostatistical analysis of the FOREGS Geochemical database. *Geoderma* **2008,** *148*, (2), 189–199.

46. Winkel, L.; Berg, M.; Amini, M.; Hug, S. J.; Johnson, C. A., Predicting groundwater arsenic contamination in Southeast Asia from surface parameters. *Nat. Geosci.* **2008,** *1*, (8), 536–542.

47. Hannah, L.; Roehrdanz, P. R.; Ikegami, M.; Shepard, A. V.; Shaw, M. R.; Tabor, G.; Zhi, L.; Marquet, P. A.; Hijmans, R. J., Climate change, wine, and conservation. *Proceedings of the National Academy of Sciences* **2013,** *110*, (17), 6907–6912.

48. Hengl, T.; de Jesus, J. M.; MacMillan, R. A.; Batjes, N. H.; Heuvelink, G. B. M.; Ribeiro, E.; Samuel-Rosa, A.; Kempen, B.; Leenaars, J. G. B.; Walsh, M. G.; Gonzalez, M. R., SoilGrids 1km? Global soil information based on automated mapping. *PLoS One* **2014,** *9*, (8), e105992.