Supplement to Modeling dust mineralogical composition: sensitivity to soil mineralogy atlases and their expected climate impacts

María Gonçalves Ageitos[1,2], Vincenzo Obiso[3,2], Ron L. Miller[3,4], Oriol Jorba[2], Martina Klose[2,5], Matt Dawson[2,6], Yves Balkanski[7], Jan Perlwitz[3,8], Sara Basart[2], Enza Di Tomaso[2], Jerónimo Escribano[2], Francesca Macchia[2], Gilbert Montané[2], Natalie Mahowald[9], Robert O. Green[10], David R. Thompson[10], and Carlos Pérez García-Pando[2,11]

[1]Projects and Construction Engineering Department. Universitat Politècnica de Catalunya - BarcelonaTECH, Terrassa, Spain
[2]Barcelona Supercomputing Center, Barcelona, Spain
[3]NASA Goddard Institute for Space Studies, New York, NY, USA
[4]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA
[5]Department Troposphere Research, Institute of Meteorology and Climate Research (IMK-TRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[6]Currently at Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA
[7]Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France
[8]Climate, Aerosol,and Pollution Research, LLC, Bronx, NY, USA
[9]Cornell University, Ithaca, NY, USA
[10]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[11]ICREA, Catalan Institution for Research and Advanced Studies, Barcelona, Spain

Correspondence: María Gonçalves Ageitos (maria.goncalves@upc.edu,maria.goncalves@bsc.es)

1 Evaluation metrics and uncertainty estimates

The evaluation of the modelled dust surface concentration, dust deposition and mineral fractions against observations uses three different evaluation metrics: the normalized mean bias (nMB), eq. 1, the normalized root mean square error (nRMSE), eq. 2, and the Pearson's correlation (r), eq. 3, where M_i represents the modelled value at the observation location and time, O_i the corresponding observed value, N the number of observations considered, and $\bar{M}$ and $\bar{O}$ the mean values for the model and observations, respectively.

$$nMB = \frac{\sum_i M_i - O_i}{\sum_i O_i} \cdot 100 \tag{1}$$

$$nRMSE = \frac{\sqrt{\frac{1}{N}\sum_i (M_i - O_i)^2}}{\bar{O}} \cdot 100 \tag{2}$$

$$r = \frac{\sum_i (M_i - \bar{M}) \cdot (O_i - \bar{O})}{\sqrt{(\sum_i (M_i - \bar{M})^2 \cdot \sum_i (O_i - \bar{O})^2)}} \tag{3}$$

We use the three evaluation metrics described above with the aim of providing a quantitative assessment of our model skills when compared to observations. However, these statistical parameters are often derived from a limited set of observations, particularly when we address the evaluation of the mineral fractions. A key element of our work is the intercomparison of two model experiments that rely on different soil mineralogy maps, and the evaluation against mineralogy observations provides relevant information to assess the strengths and weaknesses of these datasets. Relying exclusively on the mean values of the nMB, nRMSE and r to compare the two model experiments could be misleading, specially for those cases with a low number of observations. Therefore, we incorporate to our assessment the information of the number of data points that are available to produce these metrics and their uncertainty ranges at a 95% confidence level, as described below. We assume that the performance of our two experiments is statistically distinct for an evaluation metric (nMB, nRMSE, or r) whenever its mean value for one of the experiments falls outside the uncertainty range estimated for the other.

For the nMB we estimate the uncertainty range (nMB_{int}) using a two tailed Student t-test, as in eq. 4, with sd_{nMB} the standard deviation of the nMB and N the number of points considered. The degrees of freedom (df) to define the critical value (qt) are N-1, and the probability levels (p), 0.025 and 0.975.

$$nMB_{int} = qt(p, df) \cdot \frac{sd_{nMB}}{\sqrt{N}} \tag{4}$$

For the nMRSE, a chi-squared statistic is used, thus the uncertainty levels ($nRMSE_{int}$) can be calculated as in eq. 5, where qx is the critical value for N degrees of freedom (df) and the probability levels (p) of 0.025 and 0.975.

$$nRMSE_{int} = \sqrt{\frac{N}{qx(p, df)}} \tag{5}$$

Finally, for the correlation, a Fisher transformation is applied to create a normal variable. The uncertainty range is assessed for the normally distributed values assuming a two-sided confidence of 95% (critical value of 1.96). Then the transformation is reversed for the calculated bounds to obtain the uncertainty range of r.

2 Processing the observational database of mineral fractions in dust samples

The observations of mass mineral fractions compiled by Perlwitz et al. (2015) are used in this study for model evaluation purposes. To avoid using observations that might be contaminated with minerals of anthropogenic origin (i.e., from combustion or industrial activities) we apply a spatial filter to the dataset. The mask to remove the anthropogenic influence is derived from a multi-annual experiment conducted with the Community Atmosphere Model version 5 (Hamilton et al., 2019) at the global scale, that represents the atmospheric iron cycle and considers dust, anthropogenic combustion and biomass burning iron sources. We define as prominently dusty those areas where the fractional contribution of dust to the total iron is of 75% or larger (see Figure S1). Observations that fall in regions with contributions of combustion sources larger than 25% to the total iron are removed from our comparison.

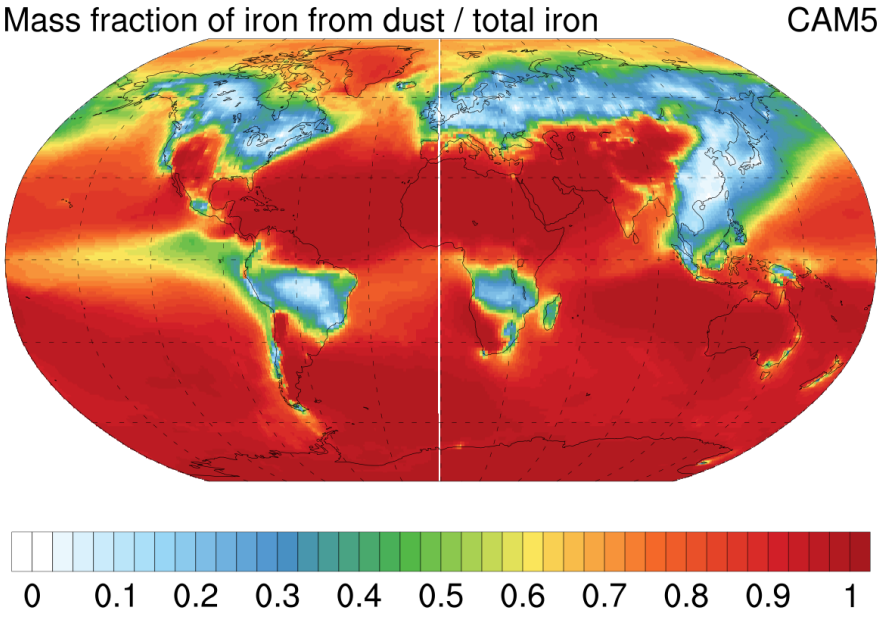

Figure S1. Mass fraction of iron from dust relative to total iron as simulated by the CAM5 model. The 0.75 threshold is used to filter the mineral fraction observations, values below 0.75 represent regions with a relevant contribution of anthropogenic sources to mineralogy, values above are assumed to happen in regions primarily influenced by dust

.

In addition, the observational database includes for each sample a number of minerals that are not explicitly represented in MONARCH, but that are in some cases similar in structure and/or physico-chemical properties to those modelled, i.e. they are considered as equivalent and belonging to the same mineralogical family.

Therefore, a first step in the evaluation process is to create a correspondence between the observed and the modelled minerals and re-calculate the mean mass fraction and standard deviation for each of the samples reported. We select as modelled minerals those in C1999-SMA and J2014-SMA, and add an additional category of *Other* to include the remaining observed minerals that are not traced by the model. In the case of some clays (e.g. kaolinite and chlorite) that are reported together, the values are kept whenever those are modelled. For the iron oxides, we decide to keep the values of hematite when reported, independently.

We further assume that iron oxides may include hematite, goethite, magnetite or others, and we will compare them with the iron oxides traced in our model. The minerals' correspondence is shown in table S1.

To estimate the mean mass fraction of the lumped minerals we sum up the mass fraction of the different individually reported minerals. To estimate the standard deviation of the re-calculated mean, we assume the mineral measurements to be non correlated, and therefore the new standard deviation for a mineral group (*i*) is calculated as the square root of the sum of variance of the *n* independently reported minerals (*j*), as in equation 6.

$$\sigma_i = \sqrt{\sum_{j=1}^{n} \sigma_j^2} \tag{6}$$

3 Results of the bulk dust evaluation

Table S2 presents the evaluation statistics for the comparison of the modelled dust surface concentration against RSMAS and AMMA climatologies, and dust deposition fluxes against observations for present climate compiled in Albani et al. (2014).

4 Annual mean budgets for minerals at emission, atmospheric burden and deposition

Table S3 presents the annual average of the global emission, burden, dry, wet and total deposition per mineral as estimated in the C1999 and J2014 experiments presented in the main paper.

5 Relative differences in surface concentration of phyllosilicates between C1999 and J2014

Figure S2 shows the comparison of phyllosilicate mass fractions between the experiments of C1999 and J2014, including mica as part of the illite-like minerals.

6 C1999-SMA and J2014-SMA Soil Mineralogy Atlas

Figures S3 to S8 represent the global distribution of the mineral mass fractions in the clay and silt sizes of the soil as derived from C1999-SMA and J2014-SMA at 0.5x0.5º resolution. The description of these datasets can be found in the Appendix A of the main article.

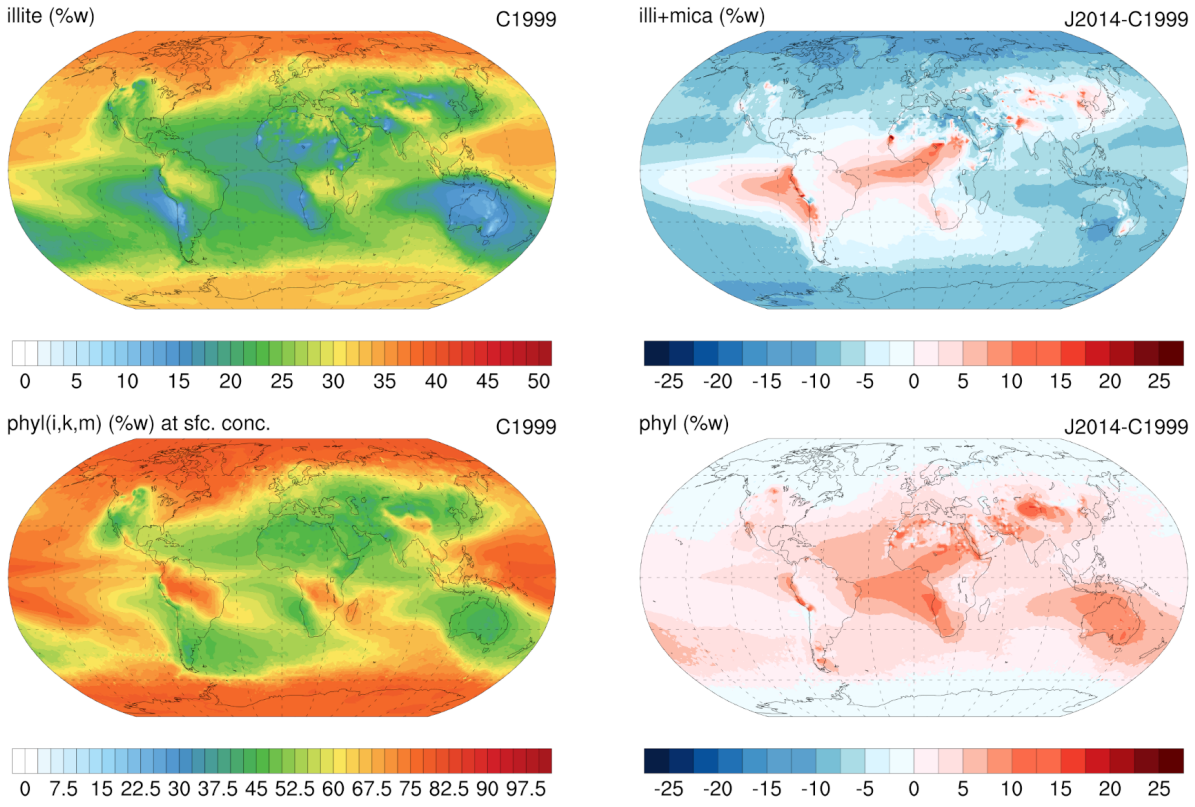

Figure S2. Mass fraction of illite and phyllosilicates:illite, kaolinite and smectite (%w) at the surface in C1999 (left) and differences (%w) bewteen J2014 illite and mica (up) and phyllosilicates: illite, kaolinite, chlorite, vermiculite and smectite (right).

References

Albani, S., Mahowald, N. M., Perry, A., Scanza, R. A., Zender, C. S., Heavens, N., Maggi, V., Kok, J. F., and Otto-Bliesner, B.: Improved dust representation in the Community Atmosphere Model, Journal of Advances in Modeling Earth Systems, 6, 541–570, https://doi.org/10.1002/2013MS000279.Received, 2014.

Hamilton, D. S., Scanza, R. A., Feng, Y., Guinness, J., Kok, J. F., Li, L., Liu, X., Rathod, S. D., Wan, J. S., Wu, M., and Mahowald, N. M.: Improved methodologies for Earth system modelling of atmospheric soluble iron and observation comparisons using the Mechanism of Intermediate complexity for Modelling Iron (MIMI v1.0), Geosci. Model Dev, 12, 3835–3862, https://doi.org/10.5194/gmd-12-3835-2019, 2019.

Perlwitz, J. P., Pérez García-Pando, C., and Miller, R. L.: Predicting the mineral composition of dust aerosols - Part 2: Model evaluation and identification of key processes with observations, Atmospheric Chemistry and Physics, 15, 11 629–11 652, https://doi.org/10.5194/acp-15-11629-2015, 2015.

Table S1. Association between reported minerals in observations and modelled MONARCH minerals for evaluation purposes.

Modelled mineral	C1999	J2014	Observed minerals
Illite	x	x	Illite, Fe-illite, Iron rich illite, Mica-illite
Smectite	x	x	Smectite, Montmorillonite
Kaolinite	x	x	Kaolinite
Chlorite		x	Chlorite, Clinochlore, Al-Chlorite
Vermiculite		x	Vermiculite
Illite-montmorillonite	x	x	Illite Smectite Mixed layers, Interstratified illite-smectite, Illite and mixed layer illite-smectite
Kaolinite-Chlorite		x	Kaolinite-Chlorite
Feldspars	x	x	Feldspars, Albite, Ca-Feldspar, Ca-Na-Feldspar, K-Feldspar, Microcline, Na-Feldspar, Orthoclase, Plagioclase, Other feldspar
Quartz	x	x	Quartz
Calcite	x	x	Calcite, Dolomite, Carbonates, Calcite and dolomite, Mg calcite
Hematite	x	x	Hematite
Goethite	x	x	Goethite
Iron oxides	x	x	Iron oxides, Hematite, Goethite, Fe-Ti oxides, Magnetite
Mica		x	Mica, Muscovite
Gypsum	x	x	Gypsum

Table S2. Number of stations (n), normalized mean bias (nMB,%), normalized root mean square error (nRMSE, %), and correlation (r) from the comparison of modelled annual mean dust surface concentration against RSMAS and AMMA observations, and modelled annual mean dust deposition against observations in Albani et al. (2014). To match the observed size range estimates of modelled values below 10 mum in diameter are used. Numbers in parenthesis represent the ranges of the statistical scores with a 95% confidence level (see section 1 in this document for details). Regions are identified in Figure 2 of the main article.

Region	**Dust surface concentration**				**Dust deposition flux**			
	n	nMB (%)	nRMSE (%)	r	n	nMB (%)	nRMSE (%)	r
NAmer	1	-32.7	32.7		7	-85.9 (-98.4,-73.3)	118.9 (78.6,241.9)	0.60 (-0.28,0.93)
CAmer	2	28.5	43.5		3	-22.5	50.4	0.97
SAmer					1	-93.8	93.8	
Europe	2	-95.2	119.0		14	-51.3 (-55.5,-47.1)	105.5 (77.2,166.3)	0.46(-0.09,0.80)
NAfri	4	17.3	54.7	0.92	23	-50.1 (-52.7,-47.5)	127.6 (99.1,178.9)	0.66(0.34,0.84)
SAfri		4	-60.8	86.1	-0.66			
WAsMe		5	-77.8 (-114.6,-40.9)	159.6 (99.6,391.4)	-0.40 (-0.95,0.74)			
EAsia	2	-43.2	43.3		14	-48.3 (-59.6,-37.0)	171.9 (125.8,271.0)	1.00 (0.99,1.00)
AusOc	3	-77.8	115.8	0.96	9	-97.5 (-113.6,-81.5)	199.4(137.1,364.0)	0.88 (0.53,0.98)
SPac	3	-70.4	73.0	0.95	2	-90.9	92.5	
NPac	4	-3.4	18.8	0.95	13	-84.2 (-90.1,-78.3)	141.5 (102.6,227.9)	0.60 (0.08,0.87)
SOce	2	-98.4	113.0		15	-30.8 (-37.2,-24.3)	108.4 (80.1,167.7)	0.96 (0.88,0.99)
Global	**23**	**11.7 (8.0, 15.4)**	**116.1 (90.2,162.8)**	**0.95 (0.88, 0.98)**	**110**	**-57.3 (-57.7,-56.9)**	**210.7 (186.2,242.8)**	**0.80 (0.72, 0.86)**

For regions with 4 or less points the scores are to be interpreted as qualitative and not statistically significant according to our analyses.

Table S3. Emission, atmospheric burden and deposition global budgets (Tg). Lifetime (days). Annual mean over 2006-2010.quar: quartz, calc: calcite, feld: feldspar, gyps: gypsum, illi: illite, kaol: kaolinite, smec: smectite, irox: iron oxides, chlo: chlorite, verm: vermiculite, mica: mica, hema: hematite, goet: goethite, and othr: other.

C1999	**quar**	**calc**	**feld**	**gyps**	**illi**	**kaol**	**smec**	**chlo**	**verm**	**mica**	**hema**	**goet**	**phyl**	**irox**	**othr**	**dust**
Emission	1359.6	181.8	451.9	47.2	617.6	428.9	339.5				56.0		1386.0	56.0		3482.6
Burden	9.8	1.3	3.3	0.4	5.8	4.3	3.1				0.5		13.2	0.5		28.5
Dry deposition	820.1	109.1	273.3	27.1	342.4	229.2	188.7				31.4		760.4	31.4		2021.3
Wet deposition	511.2	68.5	169.5	18.9	255.7	186.4	140.1				23.2		582.2	23.2		1373.6
Total deposition	1331.3	177.6	442.9	46.0	598.1	415.6	328.8				54.6		1342.5	54.6		3394.8
Lifetime	2.7	2.7	2.8	3.0	3.5	3.7	3.5				3.2		3.6	3.2		3.1
J2014	**quar**	**calc**	**feld**	**gyps**	**illi**	**kaol**	**smec**	**chlo**	**verm**	**mica**	**hema**	**goet**	**phyl**	**irox**	**othr**	**dust**
Emission	1097.0	289.8	373.0	3.9	463.1	430.5	338.0	179.9	43.2	173.3	19.1	47.9	1628.1	67.0		3458.8
Burden	7.8	2.3	2.7	0.0	4.4	4.1	3.2	1.5	0.4	1.2	0.2	0.4	14.8	0.6		28.3
Dry deposition	659.6	170.1	228.1	2.1	251.2	237.4	185.8	105.0	23.9	107.2	10.8	27.1	910.5	38.0		2008.4
Wet deposition	414.8	112.7	137.6	1.7	196.8	179.5	141.8	70.8	17.9	62.9	7.7	19.4	669.7	27.1		1363.7
Total deposition	1074.4	282.8	365.7	3.7	448.0	416.9	327.6	175.8	41.9	170.1	18.5	46.6	1580.2	65.1		3372.1
Lifetime	2.7	3.0	2.7	2.6	3.6	3.6	3.6	3.0	3.7	2.6	3.6	3.2	3.4	3.4		3.1
J2014NN	**quar**	**calc**	**feld**	**gyps**	**illi**	**kaol**	**smec**	**chlo**	**verm**	**mica**	**hema**	**goet**	**phyl**	**irox**	**othr**	**dust**
Emission	971.5	256.2	334.5	3.0	424.0	405.2	304.3	162.1	41.5	154.4	18.4	43.8	1491.6	62.2	339.7	3458.8
Burden	7.0	2.1	2.5	0.0	4.0	3.9	2.9	1.3	0.4	1.1	0.2	0.4	13.7	0.6	2.6	28.3
Dry deposition	584.5	151.0	204.5	1.6	230.6	224.3	168.1	94.6	23.1	95.8	10.5	24.8	836.4	35.3	195.1	2008.5
Wet deposition	366.9	99.0	123.4	1.3	179.6	168.2	126.9	63.7	17.2	55.8	7.4	17.8	611.4	25.1	136.4	1363.5
Total deposition	951.4	250.0	328.0	2.9	410.2	392.5	295.0	158.3	40.2	151.5	17.8	42.6	1447.8	60.4	331.5	3372.0
Lifetime	2.7	3.0	2.7	2.6	3.6	3.6	3.6	3.1	3.7	2.6	3.6	3.3	3.4	3.4	2.8	3.1

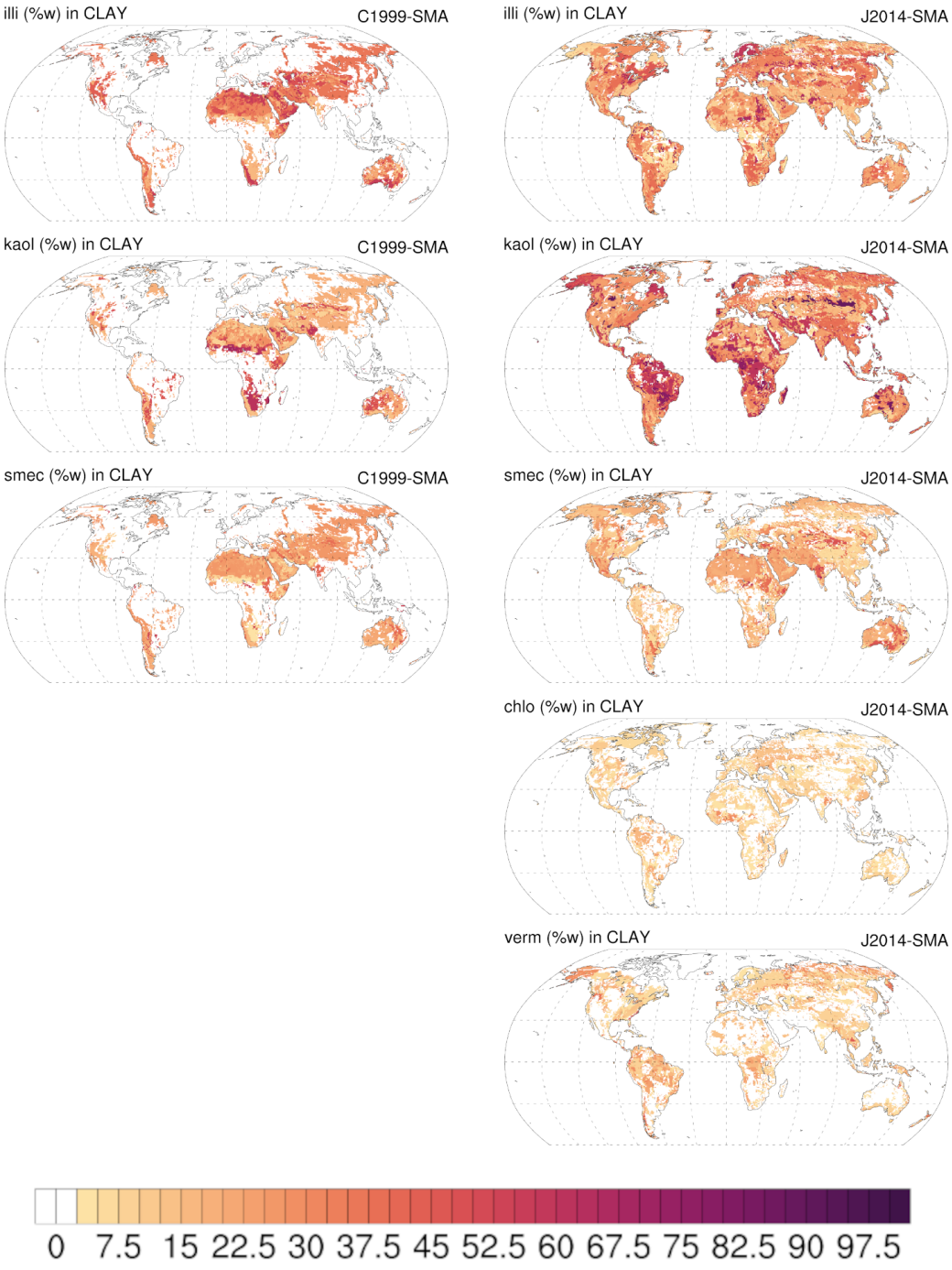

Figure S3. Mass fraction of phyllosilicates (%w) in the clay sizes of the soil according to C1999-SMA -left- and J2014-SMA -right-.

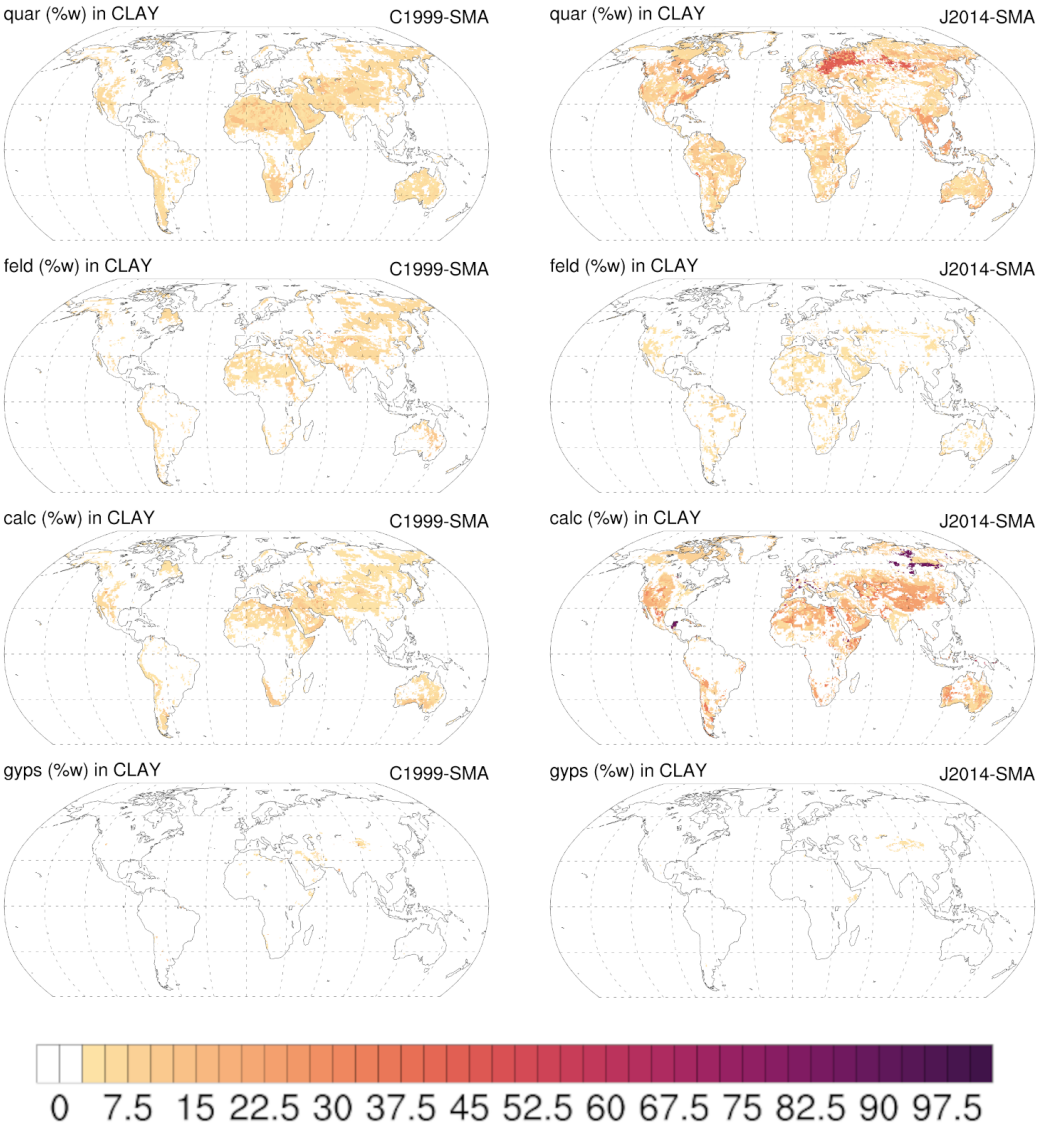

Figure S4. Mass fraction of quartz, feldspars, calcite and gypsum (%w) in the clay sizes of the soil according to C1999-SMA -left- and J2014-SMA -right-.

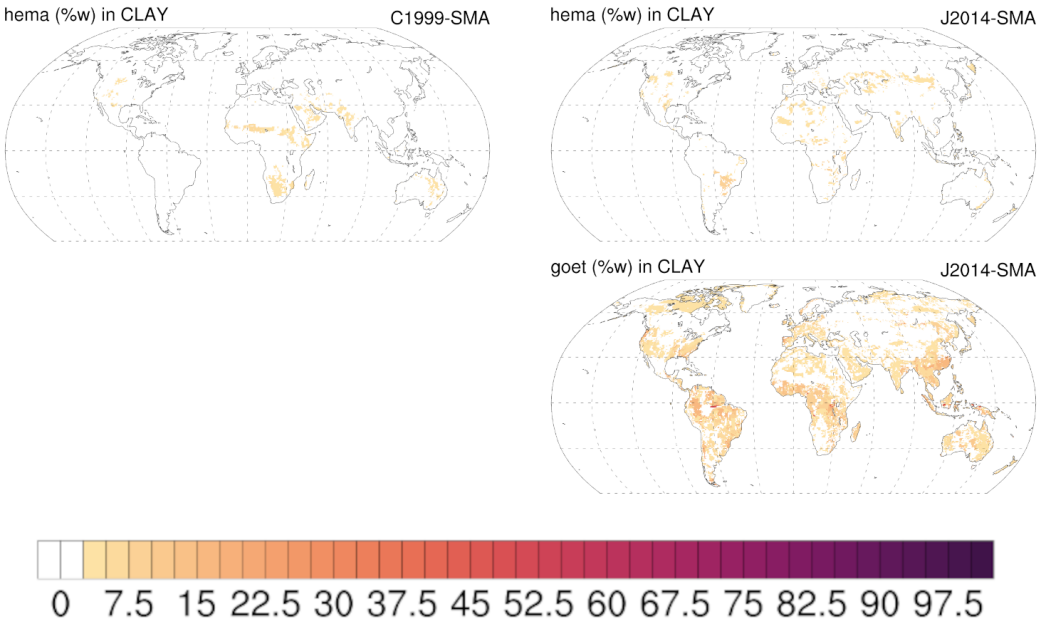

Figure S5. Mass fraction of hematite (iron oxides) and goethite (%w) in the clay sizes of the soil according to C1999-SMA -left- and J2014-SMA -right-.

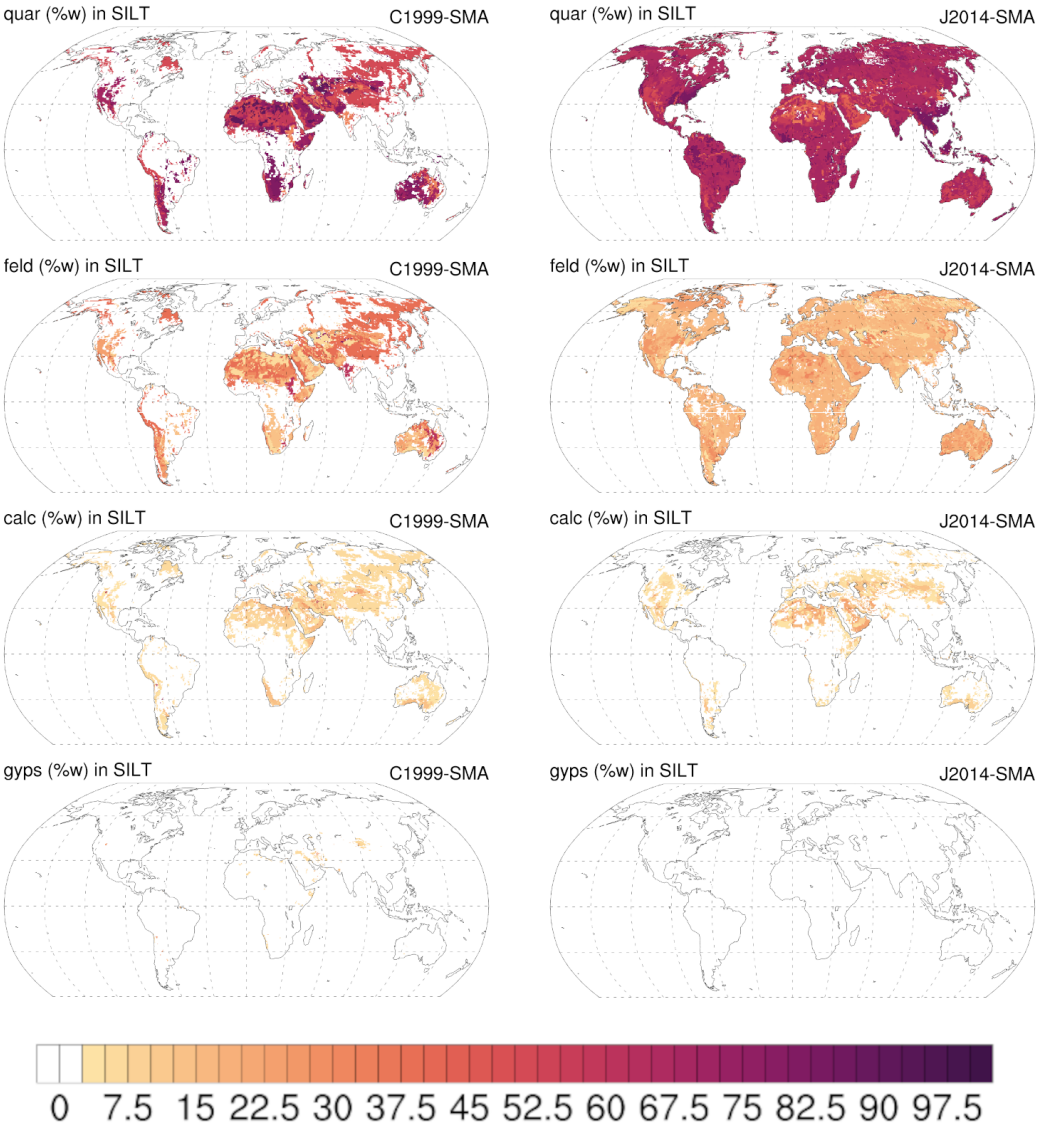

Figure S6. Mass fraction of quartz, feldspars, calcite and gypsum (%w) in the silt sizes of the soil according to C1999-SMA -left- and J2014-SMA -right-.

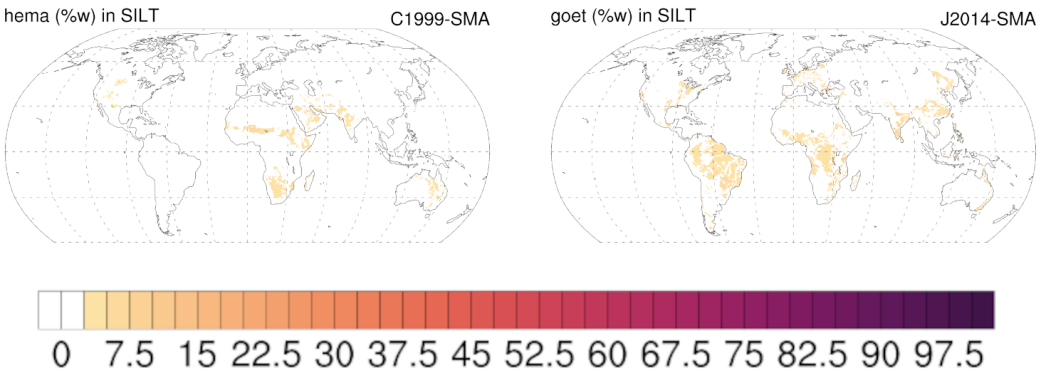

Figure S7. Mass fraction of hematite and goethite (%w) in the silt sizes of the soil according to C1999-SMA -left- and J2014-SMA -right-.

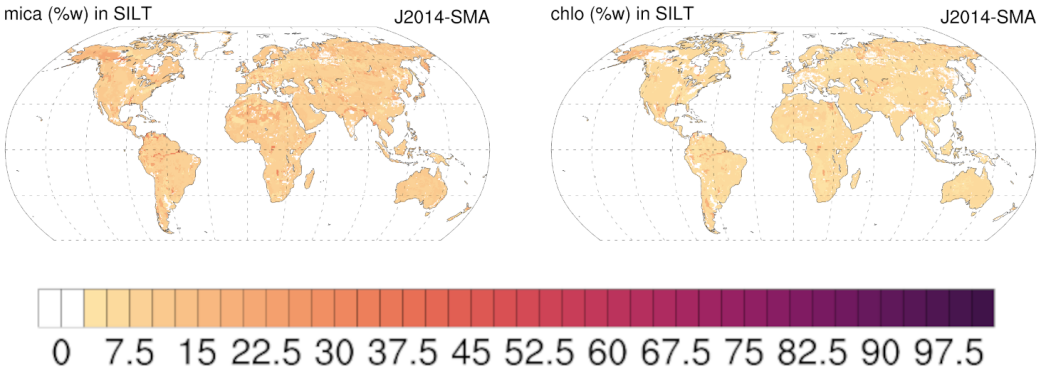

Figure S8. Mass fraction of mica and chlorite (%w) in the silt sizes of the soil according to J2014-SMA.