SUPPLEMENTARY DATA

Levoglucosan Analyses: Materials, reagents and instrumentation

All pre-analytical steps including decontamination of LDPE bottles, vials and sample bags were performed under a Class-100 clean bench located in a Class-10,000 clean room at the University of Venice. Sampling procedures were optimized to analyze trace elements and levoglucosan on the same samples. In order to minimize interference for trace element analysis the storage bottles were subject to strict decontamination procedures. Suprapur grade HNO_3 (65%, Merck) was used for all cleaning. We used a Purelab Ultra system (Elga, High Wycombe, U.K.) to produce the ultrapure water (18.2 MΩ cm, 0.01 TOC) utilized in all the analytical and pre-analytical procedures (i.e. cleaning and decontamination procedures, standard solution preparation) (Barbante et al., 1999; Gambaro et al., 2008)

All bottles were left for one week in each of the 3 subsequent solutions (5%, 2% and 1%) of HNO_3 and ultrapure water. The bottles were rinsed 3 times with ultrapure water between each bath. Bottles were stored filled with water in three layers of pre-cleaned polyethylene plastic bags. These cleaned LDPE 15 mL-bottles (Nalgene Corporation, Rochester, NY) were sent to the NEEM camp to store the collected melted ice samples. LDPE bottles used to contain the levoglucosan standard solutions and the polyethylene vials (Agilent Technologies, Wilmington, DE) for the chromatographic analysis were washed in ultrapure water, sonicated in an ultrasonic bath with ultrapure water (3 times for 14 minutes each) and rinsed with water. The cleaned bottles and vials were stored filled with water in pre-cleaned LDPE bags and were rinsed again with ultrapure water before using.

Samples and standards were transferred using Eppendorf pipettes and polyethylene tips (Eppendorf AG, Hamburg, Germany). The levoglucosan standard (purity of 99.7%) used for response factors was obtained from Sigma-Aldrich (Steinheim, Germany). Labeled levoglucosan ($^{13}C_6$ enriched to 98%, purity of 98%) was purchased from Cambridge Isotope Laboratories Inc. (Andover, MA). Standard solutions were prepared through successive dilutions with ultrapure water. Standard solutions were stored at +4 °C in pre-cleaned PE bags until the sample preparation. HPLC/MS - grade methanol was purchased from Romil Ltd. (Cambridge, U.K.). Ammonium hydroxide (≥ 25%) analytical grade was purchased from Sigma-Aldrich (Steinheim, Germany). The 13 mM ammonium hydroxide solutions were prepared by adding ultrapure water.

We used a standard solution concentration of 1.4 ng mL^{-1} for the labeled levoglucosan compound ($^{13}C_6$ enriched to 98%, purity of 98%) and a concentration of 0.4 ng mL^{-1} for the native levoglucosan (purity of 99.7%). Samples were prepared under a Class 100 clean bench in a Class 10,000 clean room by transferring 675 µL of the melted ice from the storage bottle and adding 25 µL (35 ng) of the labeled levoglucosan internal standard into the 700 µL pre-cleaned LDPE vials. Levoglucosan quantification was performed by Isotope Dilution Mass Spectrometry (IDMS) using labeled levoglucosan, and comparing the native compound peak area with that of $^{13}C_6$ isotopomer. Instrumental response factors were analysed before, during and at the end of each sample analysis set in order to evaluate instrumental response deviations. Response factors contained combined levoglucosan and $^{13}C_6$-labeled levoglucosan at a concentration of 50 pg mL^{-1} in ultrapure water. Chromatographic separations were conducted on an Agilent 1100 series liquid chromatography system (Agilent, Waldbronn, Germany). The HPLC system consists of a vacuum degasser unit, a binary pump, autosampler, and thermostatted column unit. Separation was performed injecting 300 µL (LOOP Multidraw Upgrade Kit G1313 - 68711 for Agilent 1100 series autosampler) in a C18 Synergy Hydro column (4.6 mm i.d. × 50 mm length, 4 µm particle size, Phenomenex, Torrance, CA). For the off-line post column addition of the ammonium hydroxide solution we used a Waters 515 HPLC pump (Waters Corporation, Milford, MA). The mass analyser detector used to determine and quantify levoglucosan in Arctic ice was an API 4000 triple quadrupole mass spectrometer (Applied Biosystems/MDS SCIEX, Toronto, Ontario, Canada) equipped with Turbo V ion spray source (ESI). The ion source was operated in the negative mode and three characteristic transitions for levoglucosan and isotopic enriched internal standard were monitored by multiple reaction monitoring with a 200 ms dwell time/transition. The transitions 161/113 *m/z* for levoglucosan and 167/118 *m/z* for labeled levoglucosan were used for the sample quantification.

DATA ANALYSIS AND VALIDATION

NEEM levoglucosan concentrations varied from 9 pg mL^{-1} to 1767 pg mL^{-1}. The data exhibit high variance with abrupt changes between points, resulting in a high percentage variation coefficient (or relative standard deviation), defined as $(\sigma/\bar{x})\times 100$, equal to 167.4%. Including all data results in mean of 92 pg mL^{-1} where most of samples (205/273, 75.1%) then become negative anomalies.

Statistical analysis

We calculated the correlation between the levoglucosan data and major ions measured by continuous flow analysis (CFA) at the NEEM camp. This correlation includes all available data (from 98.45 m to 450.45 m, from AD 1657 to BCE 144) from the deep ice core. CFA major ion data are available for each 55-cm bag, but the levoglucosan data are for two consecutive bags, or 110-cm samples. We therefore calculated average values of the major ion data to correspond with the levoglucosan depths. If one variable (a major ion or levoglucosan) was not recorded over a 110-cm interval, we did not include any of the other data over this interval. As BC was measured on a parallel core, and as BC data have much higher resolution, we do not include these data in our analysis in order to avoid error resulting in the attempt to calculate BC averages from NEEM-2011-S1 core over the same temporal interval covered by the deep NEEM core.

We examined the normality of our dataset as well as the normality of each variable in order to determine if we should use the raw levoglucosan data or transformed (i.e. logarithmic) values. We applied a Shapiro-Wilk normality test and we also calculated skewness and kurtosis. Levoglucosan data distribution is asymmetric rather than normal (Shapiro-Wilk normality test, p-value < 2.2 10^{-16}), with a long upper tail resulting from few strong levoglucosan spikes, yet it is not log-normally distributed. We can determine with an $\alpha = 0.05$ that all the variables are not normally distributed with exception of H_2O_2. This consideration is confirmed looking at skewness and kurtosis values that are, respectively, -0.37 and kurtosis = 0.055 (it is known that a normal distribution has kurtosis = 0). We then tested if our data (not including H_2O_2) are log-normally distributed. After applying a log transformation to our original dataset, we reapplied the Shapiro-Wilk normality test. The variables Ca^{2+}, NH_4^+ and HCHO are log-normally distributed with $\alpha = 0.05$. Therefore we cannot assume a normal distribution for our entire set of variables, even if they are log-transformed.

We then applied Pearson and Spearman correlations. The Pearson correlation is computed on *true values* and benchmark linear relationships between variables. The Spearman correlation is a non-parametric analog and is calculated on ranked data (Table S1). In a second step, in order to avoid a misleading interpretation of correlation values, we decided not to include sulphate measurements as sulphate has a large number of missing values (Table S2) that limit the amount of available data for calculating correlation.
In both cases (presence/absence of sulphate) levoglucosan does not correlate with the crustal markers Ca_{2-} and dust, but does slightly correlate with ammonium (Table 1 and 2).

Pearson correlation

	Na^+	Ca^{++}	dust	NH_4^+	NO_3^-	$SO_4^=$	H_2O_2	HCOH	levo
Na^+	1.00	0.11	0.06	0.01	0.10	-0.19	-0.12	0.03	-0.01
Ca^{++}	0.11	1.00	0.04	0.06	0.02	-0.51	-0.13	0.17	-0.06
dust	0.06	0.04	1.00	0.00	0.16	-0.09	-0.03	0.04	0.02
NH_4^+	0.01	0.06	0.00	1.00	0.23	0.09	0.22	-0.01	0.42
NO_3^-	0.10	0.02	0.16	0.23	1.00	0.06	-0.09	0.06	0.02
$SO_4^=$	-0.19	-0.51	-0.09	0.09	0.06	1.00	0.24	-0.11	0.07
H_2O_2	-0.12	-0.13	-0.03	0.22	-0.09	0.24	1.00	0.01	0.06
HCOH	0.03	0.17	0.04	-0.01	0.06	-0.11	0.01	1.00	-0.02
levo	-0.01	**-0.06**	**0.02**	**0.42**	0.02	0.07	0.06	-0.02	1.00

Spearman correlation

	Na^+	Ca^{++}	dust	NH_4^+	NO_3^-	$SO_4^=$	H_2O_2	HCOH	levo
Na^+	1.00	0.21	0.23	0.02	0.05	-0.21	-0.17	-0.01	-0.17
Ca^{++}	0.21	1.00	0.39	0.06	0.03	-0.47	-0.22	0.22	-0.02
dust	0.23	0.39	1.00	0.06	0.16	-0.07	-0.11	0.06	-0.09
NH_4^+	0.02	0.06	0.06	1.00	0.21	0.13	0.27	0.14	0.45
NO_3^-	0.05	0.03	0.16	0.21	1.00	0.06	-0.07	0.09	0.03
$SO_4^=$	-0.21	-0.47	-0.07	0.13	0.06	1.00	0.17	-0.21	0.17
H_2O_2	-0.17	-0.22	-0.11	0.27	-0.07	0.17	1.00	0.03	0.20
HCOH	-0.01	0.22	0.06	0.14	0.09	-0.21	0.03	1.00	0.04
levo	-0.17	**-0.02**	**-0.09**	**0.45**	0.03	0.17	0.20	0.04	1.00

Table S1 Pearson (above) and Spearman (bottom) correlation matrix of all data.

Pearson's correlation

	Na^+	Ca^{++}	dust	NH_4^+	NO_3^-	H_2O_2	HCOH	levo
Na^+	1.00	0.08	0.06	0.00	0.09	-0.13	0.01	-0.02
Ca^{++}	0.08	1.00	0.02	0.05	0.00	-0.08	0.18	-0.05
dust	0.06	0.02	1.00	0.00	0.17	-0.03	0.05	0.02
NH_4^+	0.00	0.05	0.00	1.00	0.21	0.21	0.02	0.41
NO_3^-	0.09	0.00	0.17	0.21	1.00	-0.04	0.04	0.02
H_2O_2	-0.13	-0.08	-0.03	0.21	-0.04	1.00	-0.01	0.04
HCOH	0.01	0.18	0.05	0.02	0.04	-0.01	1.00	-0.03
levo	-0.02	**-0.05**	**0.02**	**0.41**	0.02	0.04	-0.03	1.00

Spearman's correlation

	Na^+	Ca^{++}	dust	NH_4^+	NO_3^-	H_2O_2	HCOH	levo
Na^+	1.00	0.19	0.24	0.02	0.03	-0.19	0.00	-0.16
Ca^{++}	0.19	1.00	0.34	0.07	0.02	-0.16	0.24	0.02
dust	0.24	0.34	1.00	0.10	0.18	-0.11	0.10	-0.06
NH_4^+	0.02	0.07	0.10	1.00	0.19	0.24	0.18	0.45
NO_3^-	0.03	0.02	0.18	0.19	1.00	-0.02	0.05	0.00
H_2O_2	-0.19	-0.16	-0.11	0.24	-0.02	1.00	0.02	0.18
HCOH	0.00	0.24	0.10	0.18	0.05	0.02	1.00	0.05
levo	-0.16	**0.02**	**-0.06**	**0.45**	0.00	0.18	0.05	1.00

Table S2 Pearson (above) and Spearman (bottom) correlation matrix of all data without sulphate.

Smoothing analysis and the Global Charcoal Database (GCD) record

We tested different approaches in order to determine multi-decadal fire activity from levoglucosan concentrations. In order to compare levoglucosan data with decadal to centennial trends in other paleoclimate records, we first applied standardized statistical procedures based on those used to analyze the Global Charcoal Database (GCD) (Marlon et al., 2008; Power et al., 2008), which is a robust method for summarizing different datasets from various environmental archives. These techniques are described in detail elsewhere (Marlon et al., 2008; Power et al., 2008), where the procedure is summarized with the following steps:

a) Box Cox transformation to homogenize variance of the record.
b) Mini-max transformation to rescale data to a range between 0 and 1
c) Z-score with standard deviation and mean calculated over the period AD 1000 - 1800
d) APPROX package for R software (linear interpolation that creates output data equally spaced over time)
e) LOWESS (Locally Weighted Scatterplot Smoothing) model (Cleveland and Grosse, 1991).

This approach minimizes the influence of outliers, which helps filter noise from the data. It uses every data point, including anomalous values.

We differ from the GCD procedure in our treatment of individual spikes, as these strongly affect multi-decadal trends, even when using a LOWESS regression model. As shown in Fig. S1 C, century-long peaks were generated by single levoglucosan spikes, i.e. around AD 340. The centennial peak was an artifact since it is produced by only one sample with a high levoglucosan concentration in a period of unexceptional fire activity. We examined other solutions to solve the "smoothing problem" (i.e. use of pre-smoothing, median-based approach) but we preferred avoiding further approximation of the real behavior of levoglucosan concentrations. In order to minimize the influence of high levoglucosan spikes on the general trend, we applied LOWESS to our data after omitting peaks above a fixed threshold (Fig. S1). Excluded peaks were studied separately. Using suggestions in the literature (Tukey, 1977), we selected the following threshold: 3*rd*_Q + 1.5 x IR, which corresponds to a concentration of 168 pg mL^{-1} in our NEEM levoglucosan record. 3*rd*_Q is the third quartile and IR is the interquartile range calculated as the difference between the third quartile and the first quartile (Fig. S1 A).

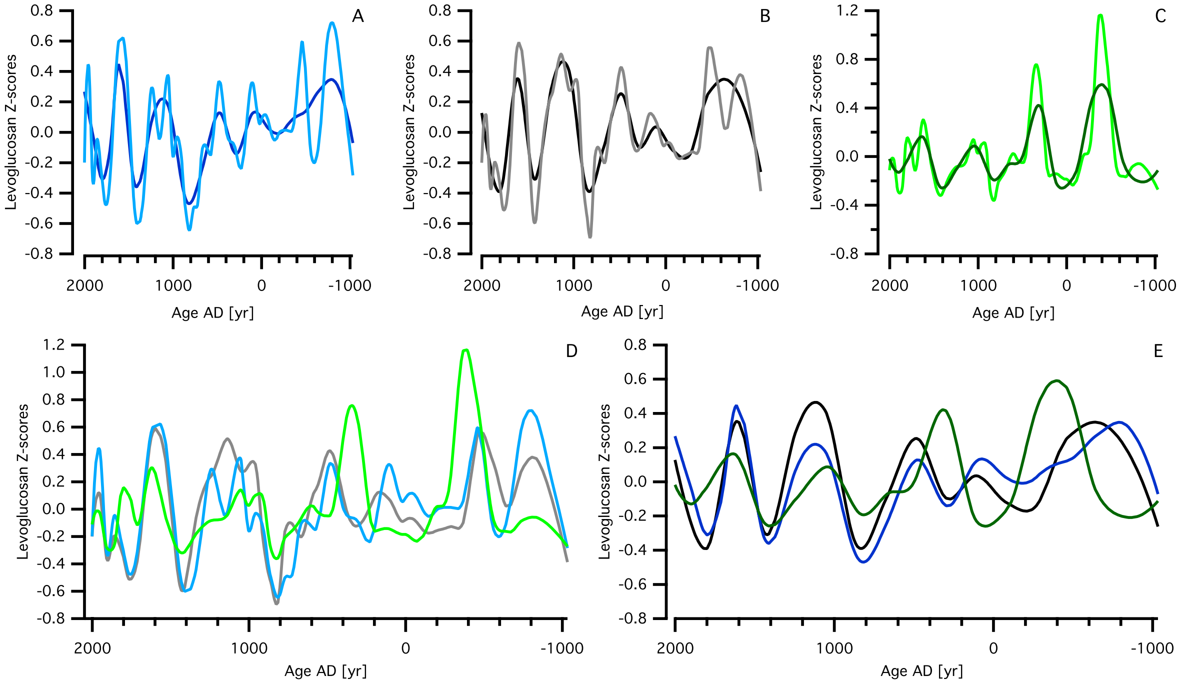

Fig. S1 Effect of levoglucosan spikes. LOWESS smoothing with SPAN parameter (f) 0.1 (light blue) and 0.2 (blue) of levoglucosan Z-scores without peaks above the threshold 3*rd*_Q + 1.5 x IR (where 3*rd*_Q is the third quartile and IR is the interquartile range calculated as the difference between the third quartile and the first quartile) (A); same transformation of A but using the threshold $\bar{x} + \sigma$, with f = 0.1 (gray) and f = 0.2 (black) (B); LOWESS smoothing including spikes with SPAN parameter (f) 0.1 (light green) and 0.2 (green) (C); comparison between LOWESS with f = 0.1 presented in A (light blue), B (gray) and C (light green) (D); comparison between LOWESS with f = 0.2 presented in A (blue), B (black) and C (green) (E).

One of the intrinsic problems of ice core analyses is that the samples are often equidistant in depth, but not equidistant in time. We tried using statistics to create data output that is equally spaced over time. However statistical interpolations (linear or other) generate “unreal” data during periods not covered by the analyses. The quality of the final smoothed function becomes less reliable, in the sense that the final smoothed data are less similar to the measured data. We hesitate to apply this technique as this approach results in data that are only interpolated rather than actually measured, and the resulting smoothed data set is yet another step farther away from the measured values.

We tested using a moving window for the definition of thresholds. We divided the whole dataset in ten subsets, where each subset had 10% of the data. We decided to fix the amount of data in each subset to guarantee the possibility to calculate significant statistical indicators (i.e. mean, deviation standard, etc.). We calculated the threshold (the 3rdQ +1.5IR) and we individuated the outliers for each subset. Using the fixed threshold results in 24 outliers, while using a moving window results in 25. Of the outliers identified by the moving window, 22 of these are the same as the 24 outliers using the fixed threshold.

We compare the smoothed data after the outliers were removed using the two methods in Figures S2 and S3. No major differences occur from using the different threshold calculation forms.

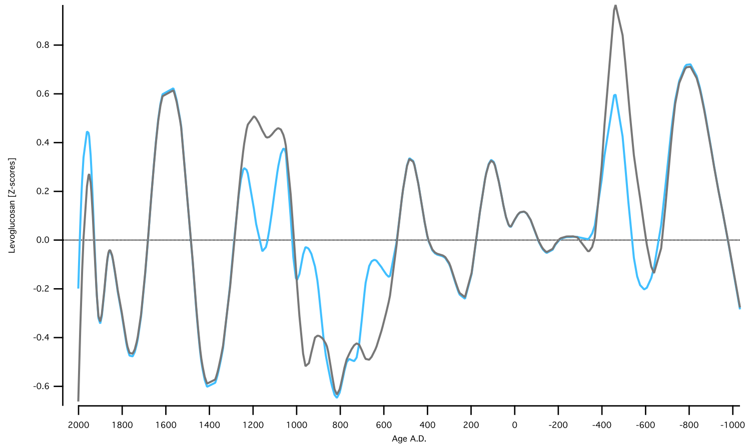

Fig. S2 Effect of different threshold calculations. LOWESS smoothing with SPAN parameter (f) 0.1 of levoglucosan Z-scores without peaks above the fixed threshold 3*rd*_Q + 1.5 x IR (where 3*rd*_Q is the third quartile and IR is the interquartile range calculates as the difference between the third quartile and the first quartile) (light blue); LOWESS smoothing with SPAN parameter (f) 0.1 of levoglucosan Z-scores without peaks above the thresholds 3*rd*_Q + 1.5 x IR calculated in ten subsets, each one containing 10% of data (gray).

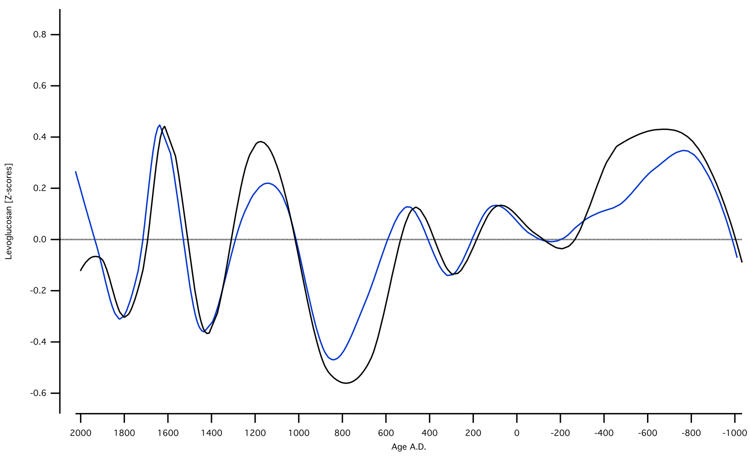

Fig. S3 LOWESS smoothing with SPAN parameter (f) 0.2 of levoglucosan Z-scores without peaks above the fixed threshold 3*rd*_Q + 1.5 x IR (where 3*rd*_Q is the third quartile and IR is the interquartile range calculates as the difference between the third quartile and the first quartile) (blue); LOWESS smoothing with SPAN parameter (f) 0.2 of levoglucosan Z-scores without peaks above the thresholds 3*rd*_Q + 1.5 x IR calculated in ten subsets, each one containing 10% of data (black).

We tested the effect of the Box Cox transformation by observing the LOWESS results after applying a Box Cox transformation on the levoglucosan dataset. The results from this comparison of statistical techniques demonstrate that the Box Cox transformation does not appear to change the data distribution. Since the Box-Cox transformation is not necessary (Fig. S4), as we have a single dataset, we prefer to avoid this additional transformation on our data.

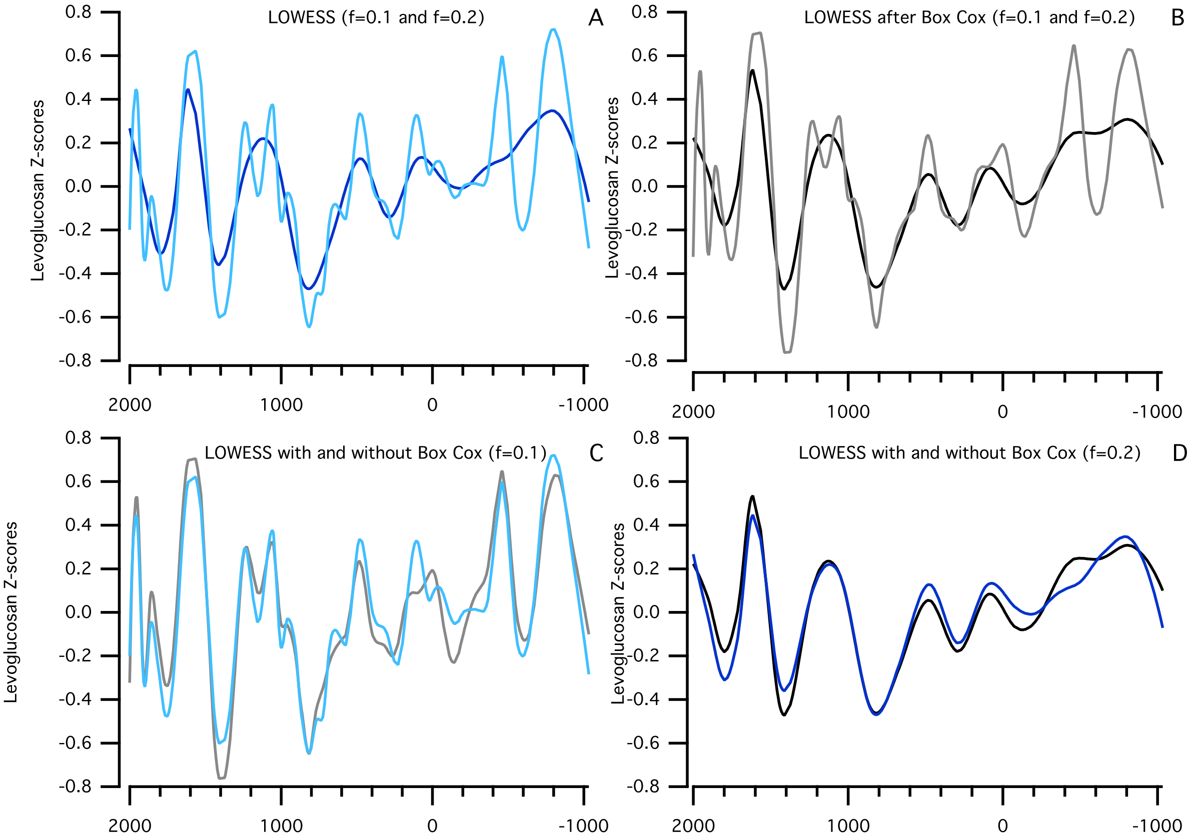

Figure S4 Effect of Box Cox Transformation. LOWESS smoothing with SPAN parameter (f) 0.1 (light blue) and 0.2 (blue) of levoglucosan Z-scores without peaks above the threshold 3*rd*_Q + 1.5 x IR (where 3*rd*_Q is the third quartile and IR is the interquartile range calculates as the difference between the third quartile and the first quartile) (A); same transformation of A but with Box Cox transformation with f = 0.1 (gray) and f = 0.2 (black) (B); comparison between LOWESS with f = 0.1 with Box Cox transformation (gray) and without Box Cox transformation (light blue) (C); comparison between LOWESS with f = 0.2 with Box Cox transformation (black) and without Box Cox transformation (blue) (D).

Using the linear interpolation APPROX to obtain equally spaced data strongly influences the multi-decadal trends and we prefer to use the original data rather than interpolated points.

In this work we used the following steps/approach:

a) Isolation of "outliers"
b) Z-score with standard deviation and mean calculated over the entire period covered by the dataset
c) Linear locally weighted polynomial regression model with tricube weight function commonly called LOESS, the later generalization of LOWESS.

The smoothing parameter (α or SPAN) is set to 0.1 and 0.2. These values give a nearest-neighbor based bandwidth covering 10% and 20% of the data.

In order to compare the high-resolution BC records with the long-trend levoglucosan profile, we applied the same statistical treatments as we used for the levoglucosan record. Ammonium has multiple anthropogenic and natural sources, and background values are linked to temperature changes (Fuhrer et al., 1993; Fuhrer et al., 1996; Legrand et al., 1992). Individual ammonium peaks correspond with levoglucosan peaks (Fig. 2, Table 1), but due to the incorporation of multiple sources in the ammonium record, we do not compare multi-decadal ammonium variability to the smoothed levoglucosan record.

MEGAFIRES

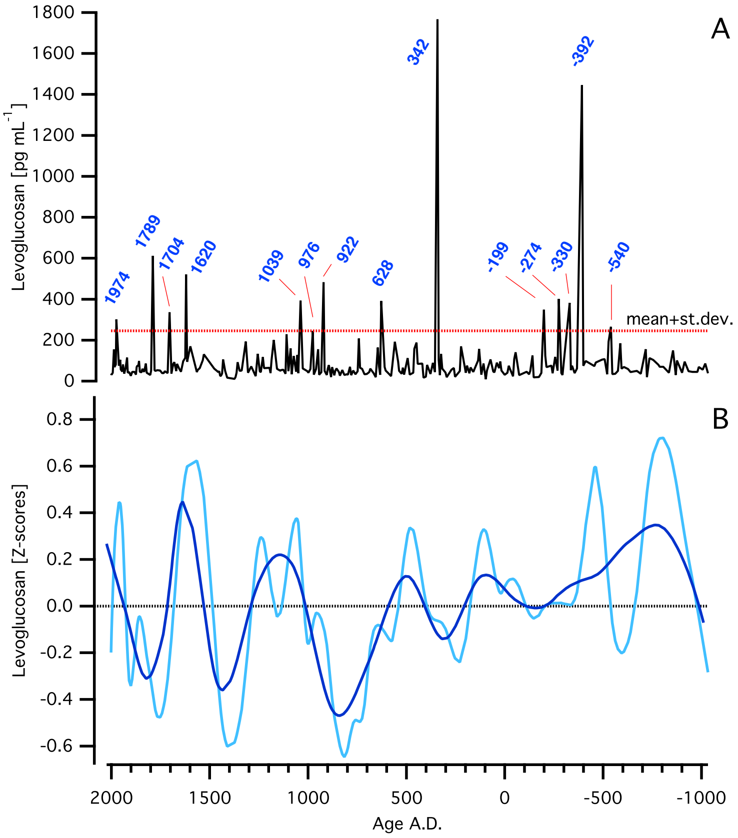

Figure S5 Levoglucosan concentration profile and megafires (peaks with concentration above the average plus one standard deviation) (A); LOWESS smoothing with SPAN parameter (f) 0.1 (light blue) and 0.2 (blue) of levoglucosan Z-scores without peaks above the threshold 3*rd*_Q + 1.5 x IR (where 3*rd*_Q is the third quartile and IR is the interquartile range calculated as the difference between the third quartile and the first quartile) (B).

THE GCD SAMPLING SITES

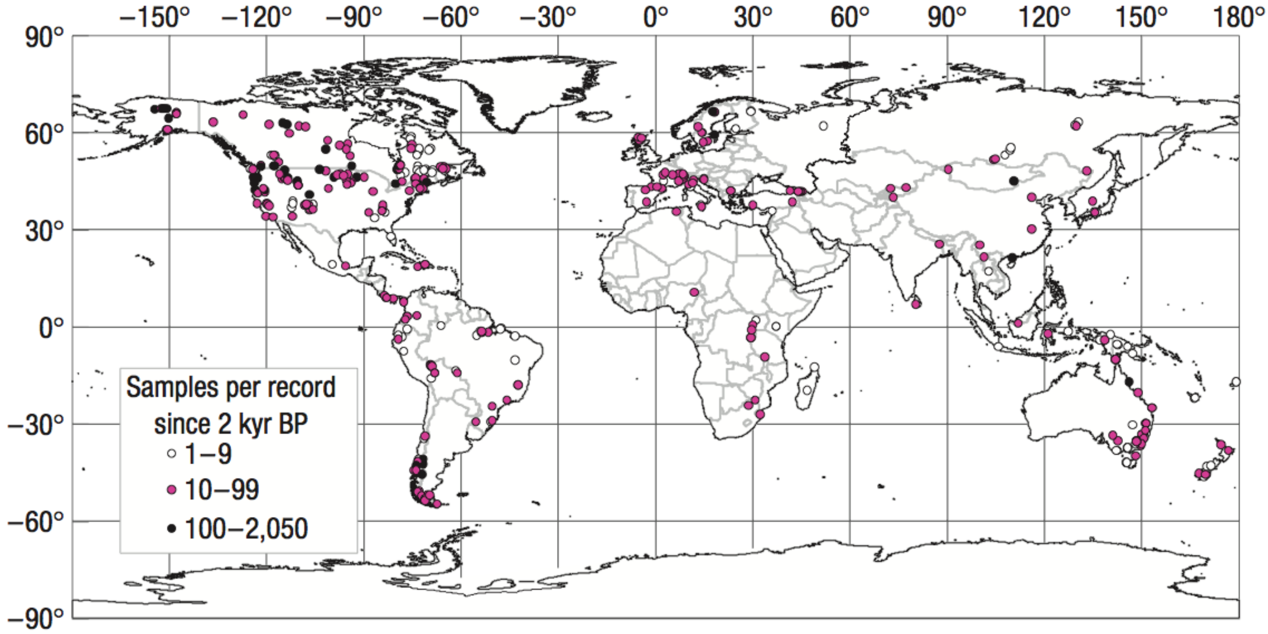

Figure S6 Locations of charcoal records and samples number. Extracted from Marlon (2008).

References

Barbante, C., Cozzi, G., Capodaglio, G., Van de Velde, K., Ferrari, C., Veysseyre, A., Boutron, C. F., Scarponi, G., and Cescon, P.: Determination of Rh, Pd, and Pt in polar and alpine snow and ice by double focusing ICPMS with microconcentric nebulization, Anal. Chem., 71, 4125-4133, 10.1021/ac981437g, 1999.

Cleveland, W., and Grosse, E.: Computational methods for local regression, Stat Comput, 1, 47-62, 10.1007/bf01890836, 1991.

Fuhrer, K., Neftel, A., Anklin, M., and Maggi, V.: continuous measurements of hydrogen-peroxide, formaldehyde, calcium and ammonium concentrations along the new grip ice core from summit, central greenland, Atmospheric Environment Part a-General Topics, 27, 1873-1880, 10.1016/0960-1686(93)90292-7, 1993.

Fuhrer, K., Neftel, A., Anklin, M., Staffelbach, T., and Legrand, M.: High-resolution ammonium ice core record covering a complete glacial-interglacial cycle, J. Geophys. Res.-Atmos., 101, 4147-4164, 1996.

Gambaro, A., Zangrando, R., Gabrielli, P., Barbante, C., and Cescon, P.: Direct determination of levoglucosan at the picogram per milliliter level in Antarctic ice by high-performance liquid chromatography/electrospray ionization triple quadrupole mass spectrometry, Anal. Chem., 80, 1649-1655, 10.1021/ac701655x, 2008.

Legrand, M., Deangelis, M., Staffelbach, T., Neftel, A., and Stauffer, B.: Large perturbation of ammonium and organic-acids content in the Summit-Greenland ice core - Fingerprint from forest-fires, Geophys. Res. Lett., 19, 473-475, 1992.

Marlon, J. R., Bartlein, P. J., Carcaillet, C., Gavin, D. G., Harrison, S. P., Higuera, P. E., Joos, F., Power, M. J., and Prentice, I. C.: Climate and human influences on global biomass burning over the past two millennia, Nat. Geosci., 1, 697-702, 10.1038/ngeo313, 2008.

Power, M. J., Marlon, J., Ortiz, N., Bartlein, P. J., Harrison, S. P., Mayle, F. E., Ballouche, A., Bradshaw, R. H. W., Carcaillet, C., Cordova, C., Mooney, S., Moreno, P. I., Prentice, I. C., Thonicke, K., Tinner, W., Whitlock, C., Zhang, Y., Zhao, Y., Ali, A. A., Anderson, R. S., Beer, R., Behling, H., Briles, C., Brown, K. J., Brunelle, A., Bush, M., Camill, P., Chu, G. Q., Clark, J., Colombaroli, D., Connor, S., Daniau, A. L., Daniels, M., Dodson, J., Doughty, E., Edwards, M. E., Finsinger, W., Foster, D., Frechette, J., Gaillard, M. J., Gavin, D. G., Gobet, E., Haberle, S., Hallett, D. J., Higuera, P., Hope, G., Horn, S., Inoue, J., Kaltenrieder, P., Kennedy, L., Kong, Z. C., Larsen, C., Long, C. J., Lynch, J., Lynch, E. A., McGlone, M., Meeks, S., Mensing, S., Meyer, G., Minckley, T., Mohr, J., Nelson, D. M., New, J., Newnham, R., Noti, R., Oswald, W., Pierce, J., Richard, P. J. H., Rowe, C., Goni, M. F. S., Shuman, B. N., Takahara, H., Toney, J., Turney, C., Urrego-Sanchez, D. H., Umbanhowar, C., Vandergoes, M., Vanniere, B., Vescovi, E., Walsh, M., Wang, X., Williams, N., Wilmshurst, J., and Zhang, J. H.: Changes in fire regimes since the Last Glacial Maximum: an assessment based on a global synthesis and analysis of charcoal data, Clim. Dyn., 30, 887-907, 10.1007/s00382-007-0334-x, 2008.

Tukey, J. W.: Exploratory Data Analysis, Addison-Wesley, 1977.