Biogeosciences Discuss., 8, 10119–10154, 2011
www.biogeosciences-discuss.net/8/10119/2011/
doi:10.5194/bgd-8-10119-2011

Biogeosciences Discussions

Intra-versus inter-site macroscale variation in biogeochemical properties along a paddy soil chronosequence

C. Mueller-Niggemann[1], A. Bannert[2], M. Schloter[3], E. Lehndorff[4], and L. Schwark[1]

[1]Institute of Geosciences, Christian-Albrechts-University of Kiel, Kiel, Germany
[2]Chair of Soil Ecology, Technische Universität München, Neuherberg, Germany
[3]Research Unit for Environmental Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany
[4]Institute of Crop Science and Resource Conservation, Bonn University, Bonn, Germany

Received: 18 August 2011 – Accepted: 9 September 2011 – Published: 14 October 2011

Correspondence to: L. Schwark (ls@gpi.uni-kiel.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

In order to assess the intrinsic heterogeneity of paddy soils, a set of biogeochemical soil parameters was investigated in five field replicates of seven paddy fields (50, 100, 300, 500, 700, 1000, and 2000 yr of wetland rice cultivation), one flooded paddy nursery, one tidal wetland (TW), and one freshwater site (FW) from a coastal area at Hangzhou Bay, Zhejiang Province, China. All soils evolved from a marine tidal flat substrate due to land reclamation. The biogeochemical parameters based on their properties were differentiated into (i) a group behaving conservatively (TC, TOC, TN, TS, magnetic susceptibility, soil lightness and colour parameters, $\delta^{13}C$, $\delta^{15}N$, lipids and *n*-alkanes) and (ii) one encompassing more labile properties or fast cycling components (N_{mic}, C_{mic}, nitrate, ammonium, DON and DOC). The macroscale heterogeneity in paddy soils was assessed by evaluating intra- versus inter-site spatial variability of biogeochemical properties using statistical data analysis (descriptive, explorative and non-parametric). Results show that the intrinsic heterogeneity of paddy soil organic and minerogenic components per field is smaller than between study sites. The coefficient of variation (CV) values of conservative parameters varied in a low range (10 % to 20 %), decreasing from younger towards older paddy soils. This indicates a declining variability of soil biogeochemical properties in longer used cropping sites according to progress in soil evolution. A generally higher variation of CV values (>20–40 %) observed for labile parameters implies a need for substantially higher sampling frequency when investigating these as compared to more conservative parameters. Since the representativeness of the sampling strategy could be sufficiently demonstrated, an investigation of long-term carbon accumulation/sequestration trends in topsoils of the 2000 year paddy chronosequence under wetland rice cultivation was conducted. The evolutionary trend showed that the biogeochemical signatures characteristic for paddy soils were fully developed in less than 300 yr since onset of wetland rice cultivation. A six-fold increase of topsoil TOC suggests a substantial gain in CO_2 sequestration potential when marine tidal wetland substrate developed to 2000 year old paddy soil.

BGD

8, 10119–10154, 2011

1 Introduction

On global scale rice (*Oryza sativa*) is the most important staple crop feeding more than 50 % of the World's population. Cultivation of rice thus affords large proportion of arable land, amounting to app. 156×10^6 hectare, of which >90 % is used for wetland or paddy rice cultivation in 2008 (IRRI, 2010; Jahn et al., 2011). A critical factor in paddy rice cropping is the periodic flooding of soils and the associated fluctuations in soil redox conditions, biogeochemical cycling of essential and trace elements, and microbial community structure. Rice paddy fields are assumed to contribute significantly to the emission of potent greenhouse gases CH_4 and N_2O (e.g. IPCC, 2007; Conrad, 2009) and to the loss of nitrate into the environment (Koegel-Knabner et al., 2010). Consequently, the investigation of biogeochemical processes in paddy soils is of critical importance in order to assess environmental impact and initiate reduction strategies.

A major problem in the design of biogeochemical studies of paddy fields is the intrinsic heterogeneity of paddy soils in the spatial as well as the temporal realm. Spatial variability may occur on the micro (nm–mm), meso (cm) and macro (m–tens of m) scale level. In paddy soils microscale variability has been described for soil aggregates and within the rhizosphere, whereas mesoscale variations occur within paddy soil profiles on cm or decimetre scale and can be related preferentially to changes in redox conditions (Koegel-Knabner et al., 2010). Macroscale heterogeneity in paddy soils occurs over distances of meters or tens of meters and is less well studied than micro or mesoscale variability. The focus on such heterogeneity investigations has been placed on soil fertility, crop yields and nutrient levels in paddy fields (Tatsuya et al., 2004; Wang et al., 2009; Wei et al., 2009; Yanai et al., 2001; Zhao et al., 2009). However the heterogeneity of bulk organic, molecular and isotopic biogeochemical parameters used to interpret paddy soil processes has not yet been investigated on the macroscale. The objectives of this study thus were, first to evaluate intra- and inter-site spatial variability of geochemical properties indicative for soil organic matter (SOM),

BGD
8, 10119–10154, 2011

minerogenic substrate and nutrients in paddy fields. Hereby, a differentiation of parameters assumed to behave conservatively by reflecting time-integrated properties (averaging over years or decades) versus fast reacting or labile parameters (reflecting daily, weekly or seasonal changes) was performed. Biogeochemical properties assumed to behave conservatively comprised soil TOC, TN, δ^{13}C, δ^{15}N and extractable lipid content and composition (reflecting time-averaged influx and composition of crop biomass and microbial re-mineralization), soil magnetic susceptibility and spectral soil colour (reflecting time-averaged soil mineralogy and redox conditions). As labile parameters microbial biomass carbon and nitrogen, nitrate, ammonium, dissolved organic N, and dissolved organic C were considered.

Factors influencing macroscale paddy soil heterogeneity can be either linked to natural variability of the substrate on which paddy soils evolved and/or management practices that locally affect influx and efflux of various components into soils, which in turn regulates the soil microbial community. Management practices can cause very localized and arbitrary enrichment (spots of 1–3 m diameter) of fertilizers, pestizide application, vegetation waste, or biomass combustion residues (heaps of burning rice straw) on paddy fields. Additionally, more systematic in-field variations in soil properties may result from flow pathways of irrigation water and its suspended materials load. Puddling of rice fields (ploughing under flooded conditions) is considered a key factor in the homogenization of locally constrained inputs and when repeated oftentimes may finally lead to the establishment of homogeneous distribution of conservative soil parameters, whereas the labile components may still exhibit severe spatial variability on the field scale.

Depending on the methodological approach applied, challenges to obtain representative paddy soil samples may vary considerably. This may lead to incompatible results, if e.g. microbial ecology conducted by genomics or proteomics targeting labile components is compared to lipidomics (analysis of phospholipid fatty acids or other microbial membrane lipids) employing conservative components.

BGD
8, 10119–10154, 2011

Secondly, if it could be proven that inter-site variations exceed intra-site variability for specific parameters, the biogeochemical trends over up to 2000 yr of rice cultivation could be evaluated for a chronosequence from the Zhejiang Province, China. Here rice cultivation started in coastal regions following land reclamation after dyke building at well dated times (Cheng et al., 2009; Jahn et al., 2011), which allows for investigation of long-term evolutionary trends in rice paddy biogeochemistry. It is postulated that ongoing paddy soil evolution will continuously diminish the intrinsic heterogeneity of young paddy soils and ultimately establish homogeneous soil biogeochemical conditions. Verification of paddy soil homogeneity in this investigation will contribute to validating pedogenic and biogeochemical studies of the same chronosequence conducted previously (Cheng et al., 2009; Bannert et al., 2011a, b; Jahn et al. 2011; Roth et al., 2011; Wissing et al., 2011). All biogeochemical investigations were carried out using 5 field replicates that were treated statistically and allow assessing whether a composition in one field or a trend over several fields is robust and representative.

2 Material and methods

2.1 Study sites

The sites are located in the coastal Cixi area (Hangzhou Bay) in the north-east of the Zhejiang province, China, as shown in Fig. 1. The Bay is affected by river runoff and tide from the East China Sea. The Yangtze (Changjiang) River with an average water runoff of 925×10^9 m^3 yr^{-1} and sediment load of 480×10^9 kg yr^{-1} supplied the dominant amount of sediment to the Hangzhou Bay (Li et al., 2009; Wang et al., 2008), where it is re-deposited by southward coastal currents and tides (Jilan and Kangshan, 1989; Xie et al., 2009). The climate is subtropical monsoonal with annual average temperature and rainfall of 16.3 °C and 1418 mm, respectively. The coastal plain of Cixi is densely covered by rivers, lakes, as well as urban and agriculture areas with main crops being wetland rice, rape, barley, and cotton (Hua and Zhu, 2000).

BGD
8, 10119–10154, 2011

Step by step land reclamation on marine tidal mudflat sediments (continuous alluvial plain of Andong Beach) through the building of protective dykes over the past 2000 yr, allows the investigation of a soil chronosequence with different stages of development and well known starting dates of cultivation (Edit Committee of Chorography of Cixi County, 1992; Yu et al., 2003; Cheng et al., 2009; Jahn et al., 2011). Wetland rice cultivation generally started when salt concentration decreases to tolerance levels, commonly after <5 yr. Based on the time of dyke construction and information's of the Edit Committee of Chorography of Cixi County (1992) sites with ongoing wetland rice cultivation for 50, 100, 300, 500, 700, 1000 and 2000 yr were identified.

In this region the generally cropping system constitutes one wetland rice season and one dry inter-crop (vegetables, cotton or cereals) season per year, called paddy-upland rotation. Soils with wetland rice cultivation represent anthraquic anthrosols, or briefly paddy soils. These are exposed to longer phases of irrigation influenced by oxygen deprivation and establishment of reducing conditions. The sampled paddy soils can be differentiated into a stagnic gleyic cambisol (50 year old paddy), a gleyic cambisol (100 to 500 yr old paddy soils) and an endogleyic stagnosol (700 to 2000 yr old paddy soils) as outlined in Jahn et al. (2011).

2.2 Sampling

Sampling was conducted in June 2008 after the harvest of the upland crop from seven paddy sites (P50, P100, P300, P500, P700, P1000, P2000) before flooding. In addition, sediment from a flooded paddy nursery site (P50N), a marine site (TW for tidal wetland), and a lacustrine site (FW for freshwater sediment) were analyzed. From each site the top soil/sediment (roughly 0–20 cm) was sampled. The sample representativeness was ensured by collecting five field replicates at each site. The sampling strategy in Fig. 2 shows that each field replicate constitutes a composite sample of seven subsamples. All soil and sediment samples were freeze dried and homogenized by grinding to fine powder. Samples were stored in glass vials in a freezer prior to further analysis.

BGD
8, 10119–10154, 2011

2.3 Laboratory analysis

Total organic carbon (TOC) concentrations of the soils and sediments were determined with a LECO CS-225 analyzer after decarbonatisation with HCl (10 % v/v) and neutralization with distilled water. The total carbon (TC), total nitrogen (TN) and total sulfur (TS) were measured directly with a CNS analyzer (Elementar Vario EL-III). Bulk magnetic susceptibility was analyzed at room temperature with a Kappabridge (KLY-2, noise level 4×10^{-8} SI) to characterize the magnetizability of ferromagnetic particles in the sample. Soil colour was quantified using a Minolta (CM-700d/600d) spectrophotometer by measuring the colour parameters on the surface of air-dried samples as described in Wiesenberg et al. (2006). Determinations of water extractable organic carbon (DOC) and nitrogen (DON) were conducted after extraction of the samples with 0.01 M $CaCl_2$ (solid to liquid ratio of 1:3) with a total organic carbon analyzer DIMA-TOC 100 (Jörgensen and Brookes, 1991). For the detection of microbial biomass carbon (C_{mic}) and nitrogen (N_{mic}) content, aliquots of soils/sediments were fumigated with chloroform (24 h) prior to $CaCl_2$ extraction. The nitrate and ammonium concentrations were measured in $CaCl_2$ extracts by a photometric autoanalyzer (CFA-SAN Plus/Skalar Analytic) using the commercial kits NANOCOLOR Nitrat-50 and Ammonium-3.

Bulk elemental analysis-isotope ratio mass spectrometry (EA-IRMS) was conducted with an elementary analyzer (FlashEA1110, ThermoFisherScientific) coupled to a mass spectrometer (DeltaV Advantage, ThermoFisher Scientific). The isotopic compositions were expressed as $\delta^{13}C$ and $\delta^{15}N$ permil units in relation to internal standards V-PDB and air nitrogen.

The hydrocarbon extractable lipids of soils and sediments were obtained by pressurized solvent extraction (Dionex ASE 200). Briefly, lipids from ca. 8 g dry soil were extracted with a dichloromethane/methanol (3/1; v/v) solvent mixture at 100 °C and 7×10^6 kg m^{-1} s^{-2}. Elemental sulphur was removed from the total lipid extracts by addition of activated copper. For quantification known amounts of perdeuterated *n*−tetracosane were added as internal standard. Total extracts dissolved in *n*−hexane

BGD
8, 10119–10154, 2011

were separated into apolar and polar compounds using small scale chromatography (Bastow et al., 2007). The aliphatic hydrocarbons were eluted with *n*−hexane after passing the glass column (4 mm × 8 cm) filled with activated silica gel (2 h at 130 °C). The sample volume was reduced via evaporation prior to transfer to GC-MS vials. Aliphatic hydrocarbon fractions were analyzed on a 30 m, ZB-1ms fused silica capillary column (0.25 mm internal diameter, film thickness 0.25 µm; Phenomenex) in a HP 5890 Series II gas chromatograph equipped with a split/splitless injector coupled to a HP 5971A mass spectrometer operated in EI-mode at 70 EV.

2.4 Statistical analysis

All individual data sets were subjected to a statistical evaluation including calculation of various descriptive statistics such as the average (AV), the standard deviation (SD) and the coefficient of variation (CV), whereby the latter describe the spread and relative proportion of variation in the data set. Non-parametric statistical analyses were applied to compare all measured soil parameters among the different sampling sites. For identification of significant variations between the sites a Kruskal-Wallis Test, suitable for non-Gaussian distributed populations, was operated (null hypothesis was all medians are equal), where the asymptotic significances (p-values) <0.05 indicate one or more medians are different. Additionally, multivariate techniques such as a principal component analysis and a cluster analysis were conducted to identify inter-site variability. The analyses were performed using MS Office Excel and PASW Statistics 18.

3 Results and discussion

3.1 Soil parameters

For the control of sample representativeness five field replicates (each a composite sample of seven individual subsamples covering an area of 2 square meters) were taken from every site and investigated for 23 different biogeochemical parameters. The

BGD
8, 10119–10154, 2011

summary of the descriptive statistics with average, standard deviation and the coefficients of variations of all parameters determined for the individual sites are listed in Table 1.

3.2 Macroscale intra-site variability

According to their properties the parameters chosen could be pre-differentiated in two groups. The first group termed as "conservative" parameters contained the soil properties that were assumed to represent time-averaged and well homogenized (puddled) soil properties. The second group termed as "labile" parameters encompassed all fast reacting properties (reflecting daily, weekly or seasonal changes) as shown in Table 1.

3.2.1 Bulk organic and minerogenic parameters

The results show much lower variation of individual parameters in the conservative group with non-uniform distribution pattern over different paddy soil cultivation times (Table 1, Fig. 3a). A coefficient of variation (CV) for soil parameters lower than 20 % generally indicates insignificant variability of these soil properties. The lowest spreading of CV values (< 1.4 %) within all soils and sediments was observed for the bulk soil organic matter (SOM) δ^{13}C isotope values and the lightness parameter L^* (CIE axis ranging from black to white). A slightly higher but generally low variation of CV values (between 1 to 13 %) was detected for the δ^{15}N values of bulk SOM, magnetic susceptibility (χ), and soil colour parameters a^* and b^* (Fig. 3a). These parameters and their marginal variation within a sampling site, comparable in soils and sediments provide information about the homogeneity of the parent material on which the respective paddy soils developed. Only the 1000 yr old paddy soil presents an exception (Fig. 3a) with a higher variation in magnetic susceptibility (18.4 %), which can be explained by a mixture of the paddy soil with adjacent upland soils of different mineralogy caused by topsoil removal and mixing in the course of dyke maintenance work (Jahn et al., 2011; Roth et al., 2011).

BGD
8, 10119–10154, 2011

Other conservative parameters (TOC, TN, C/N) influenced predominantly by organic matter input at a given site also show minor dispersion with coefficients of variation generally <10 % in soils with more than 100 yr of paddy cultivation. Only in younger paddy soils and in the reference substrates the CV values of these parameters varied in a slightly broader range up to approximately 20 % (Table 1). A range in the CV close to or less than 10 % at older paddy sites is better than expected and indicates a decreasing variability in soil organic matter parameters in longer used cropping sites. In reference substrates the CV values of conservative parameters are generally higher than in paddy soils, exemplified by variation of TN and C/N in marine sediments and of TN and TS in lacustrine sediments (Table 1).

3.2.2 Lipid and alkane concentration and composition

The highest variations of conservative parameters within a sampling site were noted for the concentration of total extractable lipids and the lipid class of *n*–alkanes. The latter is derived from land plant wax coatings, limnic macrophytes, marine/limnic algae and cyanobacteria. The *n*–alkane distributions reveal a maximum CV of 22 % in paddy soil sites and of 10 % or 30 % in the TW and FW reference substrates, respectively (Table 1). The reason for the higher variation in the limnic environment could be attributed to sampling in a shallow water environment. This favoured mixed organic matter input from submerged aquatic macrophyte biomass and terrestrial plant matter supplied by the catchment to the near-shore limnic setting.

The range of alkane concentrations in paddy soils is caused by diverse organic matter input from actual crop or weed vegetation, products from incomplete biomass combustion, or fossil fuel contaminants at different “hot-spots” on a site. In general, total extractable lipids in paddy soils represent 5.6 % of the total soil organic carbon and are mainly composed of *n*–fatty acids, *n*–alcohols, sterols, long-chain wax esters, sugars and other functionalized lipid classes. On average 2.0 mg kg^{-1} of these soil lipids are composed of source-diagnostic *n*–alkanes (Table 1) but in the 700 yr old paddy soil substantially higher proportions of *n*–alkanes (6.5 mg kg^{-1}) were observed, which

BGD
8, 10119–10154, 2011

could be attributed to fossil fuel contamination.

A partial origin of *n*−alkanes from fossil contamination is evident from the presence of a pronounced "unresolved complex mixture" (UCM) and a high abundance of thermally mature tricyclic and pentacyclic triterpenoids (hopanoids) dominating over the recent microbial triterpene diploptene (Fig. 4). Recent bacteria biosynthesize the unsaturated 17β(H)-, 21β(H)-hope(22,29)ene also termed diploptene, which is only stable under near-surface conditions (Ourisson et al., 1987). This compound is diagenetically transformed into saturated analogues upon sediment burial when reaching thermal maturity (Peters et al., 2005). In petroleum thermally stable hopanoids with 17α(H)-, 21β(H)-isomer configuration and a predominance of the 22S over 22R isomers are found. Such a petroleum derived hopanoid distribution has been encountered in the P700 topsoils (Fig. 4). Fossil fuel contamination in a paddy field could originate from a point source in the field, e.g. caused by breakdown of motorized farming machinery associated with spillage of lubricants or fuels. In such a case, only a small area of a few square meters would be contaminated, due to hydrocarbon hydrophobicity preventing further dispersal. The spatially continuous presence of fossil fuel derived hydrocarbons in the P700 field argues against such a localized point spill, but points towards a diffuse contamination, e.g. by inflow of contaminated irrigation canal waters.

The compositional variation of *n*−alkanes in the paddy soils can be evaluated using standardized parameters describing the preferential enrichment of individual alkanes. The carbon preference index (CPI) established by Bray and Evans (1961) is used to highlight the predominance of odd-over even numbered *n*−alkane homologues. High CPI_{long} values for long-chain components (> nC_{23}) derive from fresh plant waxes and tend to decline with increasing biodegradation and thermal maturity. The same accounts for short chain *n*−alkanes (< nC_{22}) derived from algae or cyanobacteria. Fossil fuels exhibit CPI values close to unity. Variation in CPI values thus reflects recent diagenetic progress or fossil fuel origin. The short chain alkanes for paddies and TW reveal CPI values <1.7 indicative of minor algal and/or cyanobacterial input with only the FW site giving a higher CPI_{short} of 2.1 pointing to more enhanced algal/cyanobacterial

BGD
8, 10119–10154, 2011

contributions. In conjunction with a small average CV of <10 % the overall proportion of aquatic microbial biomass has been low. More substantial variation was observed for CPI_{long} values around 6.0 for the young paddy sites (P50–P300) and around 4.0 for the older paddies and reference sites, indicative of progress in diagenetic overprinting. Exceptionally high CPI_{long} values were observed for P500 (Table 1) and indicate an origin of plant waxes from crops other than rice. Based on comparison with recent crop plants, the *n*–alkane distribution at this site is governed by input of wax lipids from the upland crop rape (*Brassica napus*). Very low CPI_{long} values of 2.5 for P700 are due to admixture of fossil fuel with a CPI near unity and thus support the UCM and hopanoid data. For all CPI_{long} values CV values are between 5 to 20 % showing no preference for source or degree of diagenesis.

The P_{aq}-ratio established by Ficken et al. (2000) has been used to determine the relative contribution of submerged aquatic macrophytes to the sedimentary *n*–alkane load. Values for all paddies except for P500 are close to that of the TW substrate and only the FW reference sites reveal enhanced values of 0.34 (Table 1). The exceptionally low P_{aq} of P500 is due to an origin of alkanes from non-aquatic plants and emphasizes an origin from upland crops growing under dry conditions. During time of sample acquisition the P500 site was used as a paddy soil and had been under this utilization for more than 3 yr. The molecular composition of lipids, in particular *n*–alkanes, from this site however clearly reflects its previous long-term use as a non-paddy upland cropping site. The time-integrative manner of conservative molecular biogeochemical indicators unravelling the temporally dominating land-use of soils is well illustrated in the P500 case. Despite the coexistence of older (<3 yr) and fresh lipids, the CV for various molecular parameters is only 12 to 17 %.

For all sites the CV values are below 20 % and confirm that application of molecular proxies for source identification, degree of diagenetic overprinting and fossil fuel contamination are very robust and reliable.

BGD
8, 10119–10154, 2011

3.2.3 Integrating conservative and labile parameters over cultivation times

Calculation of averages for the most important conservative parameters (TC, TOC, TN, TS, lipid yield, alkane yield, $\delta^{13}C$, $\delta^{15}N$, χ, L^*, a^*, and b^*) resulted in CV values of <20 % for all cultivation times as shown in Fig. 3c. A prominent outlier (exceeding the 95 percentile) was the *n*−alkane yield at P700 that is controlled by addition of fossil fuel contaminants to this site.

A comparison of the P50 and P50N site reveals differential behaviour that can be explained by management practices. Soil at the P50 site has experienced a long series of redox cycles like all other paddy soils, whereas the P50N site as a rice seedling nursery is kept under flooded conditions for longer times longer and thus has been going through less frequent and less dramatic redox cycles. As a result paddy soil evolution at the P50N site proceeded further compared to other sites of comparable overall cultivation time. A notable exception is the P_{aq} ratio that shows a CV comparable to the P50 or P100 sites (Fig. 3c), indicating that water table fluctuations at the P50N site affect wax lipid composition of rice seedlings, even if the soils do not pass through completely dry cycles.

Except for the P1000 site where interferences due to dyke maintenance have been reported (Jahn et al., 2011; Roth et al., 2011), a decline in the CV values over cultivation time is noted (Fig. 3c). This can be explained by an increasing degree of paddy soil evolution and homogenization of SOM and minerogenic composition, accompanied by stabilization in soil microbial community structure. This establishment of quasi-continuous composition in conservative paddy soil biogeochemical parameters was established after only 300 yr of cultivation time.

The averages for labile parameters according to cultivation age depicted in Fig. 3d show a much higher degree of variability, with the lower and upper boundaries of CV values for the 75 percentile ranging between 20 and 40 % (Fig. 3d). Outliers exceeding the 95 percentile are nitrate for the P50, DON for the P700, and ammonium for the P2000 site. All of these spatially highly variable parameters are associated to reactive

BGD
8, 10119–10154, 2011

compounds of the nitrogen cycle and are highly influenced by spatially non-systematic human manipulation such as fertilization. Additionally, the presence of biopores and cracks in the plough pan could contribute to irregular leaching processes coupled with a high variability of these water soluble parameters within a field (Sander and Gerke, 2007). The approximately 20 times higher ammonium content in P2000 could have been induced by uneven manual application of nitrogen fertilizer and an inefficient field management practices just prior to sampling (see also Roth et al., 2011). Other indicators of nitrogen cycling in paddy soil including microbial N (CV_{avg} = 33 %), C/N (CV_{avg} = 5 %), or $\delta^{15}N$ (CV_{avg} = 3 %) behave stable and demonstrate the establishment of a well controlled nitrogen cycle in paddy soils.

3.3 Inter-site variability

Reliable identification of differences in biogeochemistry between individual paddy fields and interpretation of evolutionary trends according to cultivation time, physiogeographical properties, management practises, etc. can only be achieved, if the intra-site heterogeneity is smaller than inter-site differences in biogeochemistry. We thus tested individual parameter relationships and applied statistical approaches to the entire data set employing PCA and non-parametric tests for variance analysis as well as the Kruskal-Wallis test to verify that inter-site exceeds intra-site variation. Finally, a cluster analysis was performed to elucidate, if duration of paddy soil management and associated soil evolution leads to establishment of robust clusters of comparable soil properties for the different paddies.

Examination of pairs of individual parameters revealed that in general values for each site clustered closely and only moderate overlap between site clusters occurred. This is exemplarily shown in Fig. 5a, b, c for $\delta^{13}C$ vs. $\delta^{15}N$, χ vs. a^* and $\delta^{15}N$ vs. P_{aq}, respectively. The binary plots demonstrate that individual single parameters often do not show a clear separation between sites, whereas addition of a second dimension allows full discrimination of the site clusters. The bulk isotope parameters show no

BGD
8, 10119–10154, 2011

overlap between site clusters and clearly separate the soils from the marine and limnic substrates (Fig. 5a). Soil magnetic susceptibility and colour parameters depending on minerogenic composition and expression of redoximorph features also show clear separation of site clusters with little to no overlap (Fig. 5b), except for one outlier in soil colour at the FW site. Combinations of molecular compositional and isotope parameters are suitable for site differentiation also exhibiting less intra-site versus inter-site variability (Fig. 5c). Hereby the variance in P_{aq} ratios is substantially higher than for $\delta^{15}N$ signatures. In general, overlap due to spread in one parameter is more frequent in young, less well developed or in disturbed soils (P1000).

The Kruskal-Wallis test is employed in ecological, biogeochemical, and environmental quality studies to evaluate, whether variance between sites exceeds variance within sites (Gratton et al., 2000; Katsaunos et al., 2007; Lehndorff and Schwark, 2008). Including all sites and parameters the asymptotic significance gave $p < 0.001$, except for the C/N ratio where a value of 0.004 was reached (Table 2), indicating that sites are less well distinguishable. Nevertheless the critical H values for all sites did not exceed the H-values proposed for the null-hypothesis (Table 2), implying that a full discrimination of all sites using median values of any of the selected parameters was possible. If the data set was reduced to contain only the paddy sites, i.e. P500, TW, and FW excluded (Table 2), asymptotic significance values for the C/N, CPI_{short} and P_{aq}-ratios for $p > 0.01$ could not be met. Furthermore, the median-referred critical H-values exceeded the H-values for the following parameters: TS, C/N, TOC, extract yield and CPI_{short} (Table 2), indicating that the intra-site variance was comparable to or exceeded inter-site variance. As most of these parameters represent concentrations that are primarily related to the absolute amount of soil organic matter rather than its compositional differences, a discrimination of sites using these such indicators is not feasible.

Application of principal component analysis allows evaluating the entire data set and was carried out on all paddy soils using all parameters determined (Fig. 5d), and for all paddy soils employing conservative parameters only (Fig. 5e). A full discrimination of all sites was achieved (Fig. 5d), when using the 1st and 2nd regression factors

BGD
8, 10119–10154, 2011

of all parameters, explaining 39.8 and 17.3 % of variance in the data set (Table 3). Factor 1 exhibits highest loading by parameters associated with soil organic matter concentration, whereas the 2nd factor is primarily controlled by minerogenic composition parameters (Table 3). No overlap of parametric values between the individual sites was observed, which was taken as direct evidence that the intrinsic heterogeneity of paddy soil does not exclude detailed interpretation of biogeochemical differences between sites.

If the data set was reduced to the conservatively behaving parameters, the discriminative power was reduced (Fig. 5e), whereby preferentially overlap for younger sites was observed. The 1st and 2nd regression factors for the data restricted to conservative parameters explain 50.6 and 14.2 % of the variability and are controlled by organic matter concentration and *n*−alkane compositions, respectively (Table 4). The similarity in biogeochemical properties concerning the conservative parameters, in particular for the P50 and P100 sites, can be attributed to the low evolutionary stage of the paddy soils. All sites under paddy cultivation for 300 yr and more have developed individual soil characteristics as mentioned above when discussing CV for individual age classes. Not only the duration of paddy soil utilization is of critical importance but also the individual management practice. Two sites used for 50 yr of rice cultivation were investigated, whereby one of these sites was used as nursery (P50N) for growing rice seedling prior to transplantation. As the P50N site is consistently kept under flooded conditions, soil evolution proceeds differently from the P50 site. This allows full discrimination of the P50N from the P50 and P100 sites, whereas the latter two do show considerable overlap, when PCA is conducted (Fig. 5e).

Including the reference substrates in PCA for the conservative parameters reveals a more pronounced differentiation of the substrates from the soils that develop on them when 2nd factor and 3rd factor extracted by PCA, explaining 18.3 % and 17.0 percent of the variance (Table 5) are used for discrimination (Fig. 5f). The abandonment of factor 1 explaining 37.93 percent of variance in the regression analysis leads to incomplete separation of the individual paddy sites emphasizing the importance of this factor in

BGD
8, 10119–10154, 2011

discriminant analysis. The properties exhibiting the highest loading scores on the 1st factor are organic matter concentration-related, those for the 2nd factor are governed by alkane composition, isotope signature and soil brightness, those for the 3rd factor are dominated by properties linked to redox-conditions (Table 5).

Cluster analysis performed to evaluate whether the statistical approach would group individual sites in clusters based on the entirety of all biogeochemical parameters determined is presented in Fig. 6. All older paddy soils not affected by secondary alteration (P500: extended non-paddy use, P1000: removal of surface soil for dyke maintenance) are clustered appropriately, whereas the younger soils exhibit insufficient development of individual biogeochemical paddy soil characteristics. Contamination of the P700 site did not lead to a significant change in time-integrated basic soil biogeochemical parameters but preferentially affected the aliphatic hydrocarbon composition. This indicates that the addition of the petroleum contaminants did not detrimentally affect the soil microbial community or inhibited plant growth by adding toxic substances or providing alternative substrates for microbial utilization. Seen from a temporal perspective, the P700 site acquired its biogeochemical profile over a time span of about 700 yr, whereas the minor petroleum contamination is assumed to have occurred only a few years ago and thus has negligible influence on the overall biogeochemical status. In a similar but opposite fashion, the P500 site reveals the cumulative biogeochemical characteristics of almost 500 yr of use as upland cropping site and only recently (app. 3 yr before sampling) was converted to a paddy field. Consequently, this site still exhibits the time-integrated features of a non-paddy land management and clusters with the very young paddies (P50, P50N, P100) developed on a marine tidal substrate (Fig. 6). Similarly, the P1000 site, though continuously utilized as paddy field, groups with the young soils due to repetitive removal of surface soils and dilution with soil material of non-paddy origin. Both of these sites, the P500 as well as P1000 exhibit four subsamples of close similarity and one subsample of largely deviating character, indicating the large intra-site variation caused by human interference.

BGD
8, 10119–10154, 2011

Although several sites of the chronosequence studied have been affected by anthropogenic perturbations, the cluster analysis indicates that paddy soil evolution over time led to the establishment of stable biogeochemical properties and conditions, even if permanent human intervention via soil management and utilization prevailed.

3.4 Organic matter accumulation and sequestration trends

The environmental budget of paddy soils is under debate and considered to show a negative balance due to the emission of greenhouse gases and intensive nitrate loss from paddies (IPPC, 2007; Conrad, 2009; Koegel-Knabner et al., 2010). On the other hand a positive balance could be attributed to paddy soils based on intensive atmospheric CO_2 sequestration via surface soil accumulation and preservation of fresh photosynthate. The chronosequence studied here offers the opportunity to evaluate CO_2 sequestration in paddies, comparison with non-paddy sites (P500) and interferences via intentional management (P1000) or unintentional contamination (P700). The accumulation trends for TOC, lipids and *n*−alkanes over cultivation time are shown in Fig. 7a–c, complemented by the accumulation of lipids and alkanes normalized to TOC (Fig. 7d, e). The TOC concentrations of paddy soil reach app. 1 % after 50 yr of cultivation, i.e. more than double the concentration of the substrate (Table 1). Increase in TOC continues to be rapid until about 300 yr and levels off to reach maximum concentrations after 2000 yr of cultivation. Severely lower TOC concentrations are noted for the P500 site, which might indicate a use as non-paddy for a longer period and thus has accumulated much less TOC compared to a paddy soil. The P1000 site is assumed to have accumulated TOC continuously but has lost about 50 to 70 % of this TOC due to human interference.

The lipid concentration of the paddy soils reveals a similar accumulation pattern though lipids are less stable than recalcitrant TOC (Wiesenberg et al., 2004; Marschner et al., 2006) that includes non-extractable fractions, e.g. black carbon from incomplete combustion of rice straw. Normalization of lipid yield to soil TOC indicates that these constitute app. 5 to 6 % of the organic carbon (Fig. 7d) and that the relative

BGD
8, 10119–10154, 2011

concentration increases with cultivation age. This indicates that the labile proportion of organic input into paddy soils and/or the standing microbial biomass is better preserved in long cultivated paddy soils.

The *n*−alkane accumulation trends exhibit not only negative perturbations as did TOC and lipid yield but also a pronounced increase at the P700 site due to fossil fuel contamination (Fig. 7c). This externally added *n*−alkane fraction even increases in proportion, if normalized to TOC concentrations (Fig. 7e). Exclusion of the P700 site still results in an increase of *n*−alkanes over time, which is explained by the lower mineralization of *n*−alkanes compared to functionalized lipids, e.g. fatty acids or alcohols. These components undergo oxidation and decarboxylation reactions upon diagenesis, finally leading to generation and accumulation of *n*−alkanes.

Aberrations in the accumulation of organic matter and organic matter fractions can be sufficiently explained by applications of molecular proxies, unravelling deviating sources of organic input or fossil fuel contamination. A different input of plant material, preferentially rape based on comparison of *n*−alkane distributions with recent reference crops, at the P500 site revealed this site to have been used as upland field for prolonged periods. Exceptional concentrations and compositional differences in aliphatic hydrocarbons, in particular *n*−alkanes, hopanes and UCM identify human perturbations of the soil ecosystem by petroleum contamination.

4 Conclusions

Biogeochemical proxies determined on five field replicates of paddy soils differing in cultivation age and two substrates on which these paddy soils evolved showed that the intrinsic heterogeneity of paddy soil organic and minerogenic components is smaller than differences in biogeochemical properties between study sites. This conclusion was drawn based on interpretation of individual parameters, descriptive and non-parametric statistical analysis, PCA and cluster analysis. The coefficient of variation for conservative parameters determined in pentuplicate and reflecting time-integrated

BGD
8, 10119–10154, 2011

evolution of soil properties in general was 10 % or lower. Based on this study, collection of one composite field sample is considered sufficient for generation of representative biogeochemical data in paddy soils. In field heterogeneity of fast cycling and anthropogenically amended nutrients was found very high with coefficients of variation usually between 20 and 40 % and frequent outliers. Sampling strategies covering the heterogeneity of such parameters will require much higher sampling frequency and spatial resolution.

Biogeochemical properties acquired by paddy soils over centennial periods of time behave conservative and do adapt not rapidly, if management conditions or practices are altered. Hence, previous historic land use or management practices can be reconstructed, even after new utilization has been established.

Duration of cultivation as paddy soil leads to establishment of specific soil characteristics that become increasingly stable with cultivation time. For paddies evolving on marine tidal substrates as in this study, the full development of paddy biogeochemical signatures was completed in less than 300 yr.

The environmental/ecological budget of paddy soils in this study revealed a positive balance, when sequestration of atmospheric CO_2 was considered. Perturbation of paddy soils leading to severely reduced sequestration potential can be identified by application of molecular source proxies. Thus the integrity of the carbon accumulation history of paddy soils in unknown areas can be critically evaluated.

Acknowledgements. We thank the German Research Foundation (DFG) for financial support (Schw554/20). Chinese and German partners of Research Initiative FOR 995 are thanked for field work collaboration. We appreciate analytical assistance by laboratory staff at Cologne University.

BGD
8, 10119–10154, 2011

References

Bannert, A., Mueller-Niggemann, C., Kleineidam, K., Wissing, L., Cao, Z-H., Schwark, L., and Schloter, M.: Comparison of lipid biomarker and gene abundance characterizing the archaeal ammonia-oxidizing community in flooded soils, Biol Fertil Soils, 47, 834–843, doi:10.1007/s00374-011-0552-6, 2011a.

Bannert, A., Kleineidam, K., Wissing, L., Mueller-Niggemann, C., Vogelsang, V., Welzl, G., Cao, Z.-H., and Schloter, M.: Changes in diversity and functional gene abundances of microbial communities involved in nitrogen fixation, nitrification and denitrification comparing a tidal wetland to paddy soils cultivated for different time periods, Appl. Environ. Microbiol., 77, 6109–6116, doi:10.1128/AEM.01751-10, 2011b.

Bastow, T., van Aarssen, B. G. K., and Lang, D.: Rapid small-scale separation of saturate, aromatic and polar components in petroleum, Org. Geochem., 38, 1235–1250, 2007.

Bray, E. E. and Evans, E. D.: Distribution of *n*−paraffins as a clue to recognition of source beds, Geochim. Cosmochim. Ac., 22, 2–15, 1961.

Cheng, Y.-Q., Yang, L-Z., Cao, Z-H., Ci, E., and Yin, S.: Chronosequential changes of selected pedogenic properties in paddy soils as compared with non-paddy soils, Geoderma, 151, 31–41, 2009.

Conrad, R.: The global methane cycle: recent advances in understanding the microbial processes involved, Environ. Microbiol. Reports, 1, 285–292, 2009.

Edit Committee of Chorography of Cixi County: Chorography of Cixi County, Zhejiang. Peoples Express House, Hangzhou, 1992.

Ficken, K. J., Li, B., Swain, D. L., and Eglinton, G.: An *n*−alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes, Org. Geochem., 31, 745–749, 2000.

Gratton, W. S., Nkongolo, K. K., and Spiers, G. A.: Heavy metal accumulation in soil and Jack Pine (Pinus banksiana) needles in Sudbury, Ontario, Canada, B. Environ. Contam. Tox., 64, 550–557, 2000.

Hua, X. L. and Zhu, H.: Ningbo Yearbook, Publishing House of China (Zhong Hua Shu Ju), Beijing , 2000 (in Chinese).

IPPC: Fourth Assessment Report: Climate Change 2007, Website http://www.ipcc.ch/publications_and_data/publications_and_data_reports.shtml, 2007.

IRRI: Website http://www.knowledgebank.irri.org, 2010.

BGD
8, 10119–10154, 2011

Jahn, R., Schad, P., Amelung, W., Cao, Z-H., Fiedler, S., Kalbitz, K., Koelbl, A., Schwark, L., Vogelsang, V., Wissing, L., and Koegel-Knabner, I.: Genesis of paddy soils from marshlands (Zhejiang Province, China), Geoderma, 2011.
Jilan, D. and Kangshan, W.: Changjiang river plume and suspended sediment transport in Hangzhou Bay, Cont. Shelf Res., 9, 93–111, 1989.
Jörgensen, R. and Brookes, P.: Die Bestimmung der mikrobiellen Biomasse des Bodens mit der Fumigations-Extraktions-Methode, VDLUFA-Schriftenreihe, 666–671, 1991.
Katsaounos, C. Z., Giokas, D. L., Vlessidis, A. G., and Karayannis, M. I.: Identification of longitudinal and temporal patterns of phosphorus fractionation in river sediments by non-parametric statistics and pattern recognition techniques, Desalination, 213, 311–333, 2007.
Koegel-Knabner, I., Amelung, W., Cao, Z.-H., Fiedler, S., Frenzel, P., Jahn, R., Kalbitz, K., Koelbl, A., and Schloter, M.: Biogeochemistry of paddy soils, Geoderma, 157, 1–14, 2010.
Lehndorff, E. and Schwark, L.: Accumulation histories of major and trace elements on pine needles as function of air quality, Atmos. Environ., 42, 833–845, 2008.
Li, X., Mander, Ü., Ma, Z., and Jia, Y.: Water quality problems and potential for wetlands as treatment systems in the Yangtze River Delta, China, The Society of Wetland Scientists, 29, 1125–1132, 2009.
Marschner, B., Brodowski, S., Dreves, A., Gleixner, G., Grootes, P.M., Hamer, U., Heim, A., Jandl, G., Ji, R., Kaiser, K., Kalbitz, K., Kramer, C., Rethemeyer, J., Schmidt, M. W. I., Schwark, L., and Wiesenberg, G. L. B.: Is recalcitrance a relevant property for the stabilization of organic matter in soils?, J. Plant Nutr. Soil Sc., 178, 191–210, 2008.
Ourisson, G., Rohmer, M., and Poralla, K.: Prokaryotic hopanoids and other polyterpenoid sterol surrogates, Ann. Rev. Microbiol., 41, 301–333, 1987.
Peters, K. E., Walters, C. C., and Moldowan, J. M.: The biomarker guide, volume 1 and 2: biomarkers and isotopes in the petroleum exploration and earth history, Cambridge University Press, Cambridge, 2005.
Roth, P. J., Lehndorff, E., Cao, Z.-H., Zhuang, S., Bannert, A., Wissing, L., Schloter, M., Koegel-Knabner, I., and Amelung, W.: Accumulation of nitrogen and microbial residues during 2000 years of rice paddy and non-paddy soil development in the Yangtze River Delta, China, Global Biogeochem. Cy., doi:10.1111/j.1365-2486.2011.02500.x, 2011.
Sander, T. and Gerke, H. H.: Preferential Flow Patterns in Paddy Fields Using a Dye Tracer, Vadose Zone J., 6, 105–115, 2007.
Tatsuya, I., Goto, K., Iida, M., Nonami, K., Inoue, H., and Umeda, M.: Geostatistical analysis of

BGD
8, 10119–10154, 2011

yield, soil properties and crop management practices in paddy rice fields, Plant Prod. Sci., 7 , 230–239, 2004.
Wang H., Zuosheng Y., Saito Y., and Liu J. P.: Reconstruction of sediment flux from the Changjiang (Yangtze River) to the sea since the 1860s, J. Hydrol., 349, 318–332, 2008.
Wang, L., Wu, J. P., Liu, Y. X., Huang, H. G., and Fang, Q. F.: Spatial Variability of Micronutrients in Rice Grain and Paddy Soil, Pedosphere, 19, 748–755, 2009.
Wei, Y.-C., Bai, Y.-L., Jin, J.-Y., Zhang, F., Zhang, L.-P., and Liu, X.-Q.,: Spatial variability of soil chemical properties in the reclaiming marine foreland to Yellow Sea of China, Agricultural Sciences in China, 8, 1103–1111, 2009.
Wiesenberg, G. L. B., Schwarzbauer, J., Schmidt, M. W. I., and Schwark, L.: Sources and turnover rates of soil organic matter derived from *n*-alkane/*n*-carboxylic acid compositions and C-isotope signatures, Org. Geochem., 35, 1371–1393, 2004.
Wiesenberg, G. L. B., Schwark, L., and Schmidt, M. W. I.: Extractable lipids and colour in particle-size separates and bulk arable soils, Eur. J. Soil Sci., 57, 634–643, 2006.
Wissing, L., Koelbl, A., Vogelsang, V., Fu, J.-R., Cao, Z.-H., and Koegel-Knabner, I.: Organic carbon accumulation in a 2000-year chronosequence of paddy soil evolution, Catena, doi:10.1016/j.catena.2011.07.007, 2011.
Xi, D., Wang, Z., Gao, S., and De Vriend, H. J.: Modeling the tidal channel morphodynamics in a macro-tidal embayment, Hangzhou Bay, China, Cont. Shelf Res., 29, 1757–1767, 2009.
Yanai, J., Lee, C. K., Kaho, T., Iida, M., Matsui, T., Umeda, M., and Kosaki, T.: Geostatistical analysis of soil chemical properties and rice yield in a paddy field and application to the analysis of yield-determining factors, Soil Sci. Plant Nutr., 47, 291–301, 2001.
Yu, Q., Yu, Y., You, A., Song, L., and Cao, Y.: Contemporary evolution process of the Hangzhou Bay, International Conference on Estuaries and Coasts, 9–11 November, Hangzhou, China, 2003.
Zhao, Y., Xu, X., Darilek, J. L., Huang, B., Sun, W., and Shi, X.: Spatial variability assessment of soil nutrients in an intense agricultural area, a case study of Rugao County in Yangtze River Delta Region, China, Environ. Geol., 57, 1089–1102, 2009.

BGD
8, 10119–10154, 2011

Table 1. Descriptive statistics of all biogeochemical parameters determined for the 10 study sites. P50N designates the seedling nursery paddy, TW designates the marine tidal flat substrate and FW designates the freshwater limnic substrate. AV = average value, SD = standard deviation, CV = coefficient of variation. Conservative parameters were grouped TC to b^*(D65), labile parameters were grouped N_{mic} to DOC.

	P50			P50N			P100			P300			P500		
	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV
TC [%]	1.49	0.17	11.3	1.33	0.11	8.1	1.40	0.27	19.2	2.41	0.10	4.1	1.39	0.07	4.8
TOC [%]	0.99	0.14	14.2	1.04	0.07	6.3	1.39	0.28	20.0	2.25	0.12	5.3	1.33	0.11	8.4
TIC [%]	0.50	0.06	12.8	0.29	0.08	26.6	0.02	0.02	98.2	0.17	0.10	63.0	0.06	0.08	140.1
TN [%]	0.118	0.017	14.5	0.122	0.014	11.1	0.157	0.026	16.7	0.261	0.010	3.9	0.158	0.012	7.5
TS [%]	0.032	0.006	19.6	0.031	0.001	4.3	0.030	0.005	16.9	0.056	0.010	17.7	0.030	0.004	14.1
C/N	8.37	0.19	2.3	8.54	0.78	9.1	8.85	0.59	6.6	8.63	0.40	4.6	8.44	0.37	4.4
Lipids [mg kg^{-1} dw]	553	115	20.8	544	42	7.7	735	154	21.0	1210	51	4.3	651	50	7.6
Lipids [g kg^{-1} TOC]	56	5	9.0	52	3	6.1	53	4	8.4	54	4	6.7	49	3	5.5
n-Alkanes [µg kg^{-1} dw]	1531	310	20.3	1271	96	7.6	2141	446	20.8	3440	221	6.4	2346	186	7.9
n-Alkanes [mg kg^{-1} TOC]	155	30	19.0	123	15	12.0	154	16	10.5	153	13	8.6	176	10	5.7
CPI^{b}_{short}	1.8	0.2	11.9	1.6	0.2	10.1	1.4	0.1	3.6	1.3	0.1	9.2	1.4	0.2	12.9
CPI^{c}_{long}	5.7	0.8	14.6	4.2	0.9	21.4	5.8	1.0	18.1	6.1	0.5	8.4	10.4	1.8	17.0
P^{d}_{aq}	0.24	0.05	19.2	0.22	0.05	22.2	0.19	0.03	16.1	0.19	0.01	3.4	0.07	0.01	17.6
$\delta^{13}C$ [‰]	−28.0	0.4	−1.4	−27.1	0.3	−1.2	−28.2	0.2	−0.6	−28.5	0.1	−0.4	−28.0	0.1	−0.4
$\delta^{15}N$ [‰]	3.1	0.1	2.9	5.7	0.3	4.4	4.6	0.1	2.9	2.7	0.3	11.4	5.6	0.1	2.6
χ [10^{-8} m^3 kg^{-1} dw]	26.1	1.1	4.1	24.7	0.3	1.4	14.8	1.1	7.2	11.8	0.9	7.7	21.3	0.8	3.9
L^* (D65)	55.6	0.4	0.7	56.3	0.6	1.0	55.3	0.4	0.7	54.1	0.6	1.1	55.0	0.6	1.0
a^* (D65)	3.7	0.1	2.8	3.6	0.1	4.1	3.7	0.1	3.5	3.3	0.1	2.0	3.2	0.0	1.3
b^* (D65)	15.3	0.2	1.3	15.1	0.5	3.0	14.8	0.2	1.1	14.0	0.3	2.1	13.6	0.4	2.6
N_{mic} [µg g^{-1} dw]	33.4	8.5	25.5	45.1	5.8	12.9	37.6	3.3	8.8	4.1	1.1	27.9	15.6	5.3	33.7
C_{mic} [µg g^{-1} dw]	293	103	35.0	558	61	11.0	522	108	20.6	167	33	19.5	490	93	18.9
Nitrate [µg N g^{-1} dw]	2.7	1.8	67.1	3.6	1.1	30.8	7.5	3.9	52.1	27.5	12.0	43.5	12.0	3.2	27.0
Ammonium [µg N g^{-1} dw]	0.4	0.1	33.3	0.5	0.4	83.8	0.2	0.1	30.8	0.6	0.3	59.0	0.1	0.0	13.1
DON [µg g^{-1} dw]	2.2	0.9	42.1	3.0	0.6	19.9	5.2	1.3	24.3	5.8	1.4	23.3	12.0	2.6	21.8
DOC [µg g^{-1} dw]	21.5	7.7	36.0	9.3	2.8	29.9	14.7	4.6	31.2	16.4	5.2	31.5	23.6	4.2	17.9

[a] $\sum$ *n*-alkanes = $n\text{-}C_{13-40}$

[b] CPI_{short}: 0.5* ([($n\text{-}C_{15} + n\text{-}C_{17} + n\text{-}C_{19} + n\text{-}C_{21}$)/($n\text{-}C_{14} + n\text{-}C_{16} + n\text{-}C_{18} + n\text{-}C_{20}$)]+ [($n\text{-}C_{15} + n\text{-}C_{17} + n\text{-}C_{19} + n\text{-}C_{21}$)/($n\text{-}C_{16} + n\text{-}C_{18} + n\text{-}C_{20} + n\text{-}C_{22}$)])

[c] CPI_{long}: 0.5* ([($n\text{-}C_{25} + n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31}$)/($n\text{-}C_{24} + n\text{-}C_{26} + n\text{-}C_{28} + n\text{-}C_{30}$)]+ [($n\text{-}C_{25} + n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31}$)/($n\text{-}C_{26} + n\text{-}C_{28} + n\text{-}C_{30} + n\text{-}C_{32}$)])

[d] P_{aq}: ($n\text{-}C_{23} + n\text{-}C_{25}$)/($n\text{-}C_{23} + n\text{-}C_{25} + n\text{-}C_{29} + n\text{-}C_{31}$)

BGD
8, 10119–10154, 2011

Table 1. Continued.

	P700			P1000			P2000			TW			FW			all soils		
	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV
TC [%]	2.10	0.19	8.9	1.16	0.11	9.4	3.05	0.13	4.3	1.28	0.02	1.6	1.62	0.17	10.3	1.80	0.65	36.2
TOC [%]	2.00	0.09	4.7	1.10	0.10	8.9	2.85	0.10	3.5	0.49	0.06	12.4	1.51	0.21	14.0	1.71	0.70	41.0
TIC [%]	0.12	0.12	98.8	0.06	0.03	60.7	0.20	0.13	67.7	0.79	0.05	6.3	0.10	0.11	101.2	0.09	0.12	126.2
TN [%]	0.208	0.022	10.5	0.128	0.017	13.2	0.361	0.018	4.9	0.052	0.011	21.4	0.148	0.028	19.1	0.191	0.082	42.9
TS [%]	0.050	0.006	12.2	0.026	0.004	17.0	0.057	0.005	9.6	0.047	0.004	8.0	0.035	0.007	20.1	0.039	0.013	34.1
C/N	9.68	0.77	8.0	8.67	0.57	6.5	7.93	0.61	7.7	9.63	1.99	20.6	10.38	1.34	12.9	8.99	0.74	8.2
Lipids [mg kg^{-1} dw]	1358	143	10.5	639	54	8.4	1764	73	4.2	862	248	28.7	1438	177	12.3	941	440	46.8
Lipids [g kg^{-1} TOC]	68	8	11.7	59	10	16.7	62	2	3.4	176	49	28.1	95	6	5.9	54	8	13.9
n-Alkanes [µg kg^{-1} dw]	6557	1456	22.2	1690	146	8.7	5806	961	16.6	975	102	10.5	4551	1341	29.5	3.138	2.017	64.3
n-Alkanes [mg kg^{-1} TOC]	330	80	24.4	153	8	5.1	204	34	16.9	200	11	5.6	299	74	24.8	175	65	37.1
CPI^{b}_{short}	1.6	0.1	5.7	1.5	0.2	16.3	1.7	0.2	10.4	1.3	0.1	6.6	2.1	0.2	10.5	1.5	0.2	13.5
CPI^{c}_{long}	2.5	0.1	5.0	4.2	0.9	20.3	4.1	0.2	4.1	3.4	0.3	9.5	4.0	0.4	9.6	5.4	2.4	44.9
P^{d}_{aq}	0.24	0.01	5.8	0.24	0.04	17.5	0.25	0.01	5.2	0.26	0.02	6.0	0.34	0.03	7.6	0.21	0.06	30.2
δ^{13}C [‰]	–28.0	0.0	–0.2	–25.9	0.2	–0.6	–29.4	0.1	–0.3	-24.4	0.1	-0.2	–22.6	0.4	–2.0	–27.9	1.0	–3.6
δ^{15}N [‰]	4.9	0.2	4.0	5.3	0.2	3.7	2.1	0.3	13.4	4.6	0.2	3.7	1.6	0.2	11.6	4.3	1.3	31.3
χ [10^{-8} m^3 kg^{-1} dw]	15.5	0.8	4.9	15.5	2.9	18.4	12.0	1.0	8.7	62.8	0.7	1.1	16.0	2.1	13.1	17.5	5.3	30.3
L^{*} (D65)	55.1	0.7	1.2	57.1	0.8	1.4	57.2	0.2	0.4	56.4	0.4	0.7	58.7	0.2	0.3	55.7	1.1	2.1
a^{*} (D65)	2.8	0.1	2.9	2.9	0.1	4.8	1.8	0.2	8.5	3.7	0.2	4.1	2.7	0.3	12.4	3.1	0.6	19.9
b^{*} (D65)	14.5	0.2	1.2	15.2	0.5	3.4	12.4	0.7	5.7	14.5	0.3	2.2	14.3	0.7	5.2	14.3	1.0	6.9
N_{mic} [µg g^{-1} dw]	30.0	6.2	20.8	37.7	3.8	10.1	32.6	12.4	38.0	27.4	5.6	20.3	57.2	26.6	46.4	29.4	14.1	47.8
C_{mic} [µg g^{-1} dw]	535	43	8.1	331	71	21.4	4043	1591	39.4	255	41	16.2	1006	616	61.3	882	1340	151.9
Nitrate [µg N g^{-1} dw]	7.0	2.8	40.0	6.5	0.8	12.3	15.1	5.0	33.5	3.9	2.3	60.9	21.9	20.9	95.5	10.4	9.0	85.9
Ammonium [µg N g^{-1} dw]	0.2	0.0	19.3				6.6	5.8	87.9	0.6	0.2	32.6	12.4	8.4	68.2	1.1	2.9	262.7
DON [µg g^{-1} dw]	9.7	6.4	65.4	4.4	0.9	20.6	10.7	3.0	27.7	0.7	0.5	72.8	21.0	9.5	45.5	6.8	4.3	63.5
DOC [µg g^{-1} dw]	21.1	4.0	18.9	8.1	2.1	26.0	34.7	7.7	22.1	13.6	2.8	20.3	30.7	20.4	66.3	18.6	9.4	50.5

BGD
8, 10119–10154, 2011

Table 2. Non-parametric variance analysis by Kruskal-Wallis test, suitable for non-normal distributed data sets, performed for all sites ($n = 49$) and for paddy sites only ($n = 34$). Significant variation between sites is indicated, H values from Chi-squared test exceed the critical H-values of the null-hypothesis. Parameters indistinguishable between sites because intra-site variance exceeds inter-site variance are plotted in italic.

	all sites				all paddy sites			
	H	df	p	Hcrit	H	df	p	Hcrit
TC [%]	40.824	9	0.000	28.992	29.524	6	0.000	27.800
TOC [%]	43.880	9	0.000	39.396	27.995	6	0.000	29.200
TIC [%]	33.971	9	0.000	28.191	22.712	6	0.001	14.800
TN [%]	43.005	9	0.000	36.195	29.282	6	0.000	29.200
TS [%]	38.561	9	0.000	37.995	26.502	6	0.000	27.800
C/N	24.233	9	0.004	17.187	16.390	6	0.012	16.400
TOC/S	30.601	9	0.000	23.389	19.565	6	0.003	16.400
Lipids [mg kg^{-1} dw]	41.197	9	0.000	40.997	28.920	6	0.000	29.200
Lipids [g kg^{-1} TOC]	37.719	9	0.000	29.992	17.955	6	0.006	18.200
n-Alkanes [µg kg^{-1} dw]	44.868	9	0.000	39.396	30.076	6	0.000	29.200
n-Alkanes [mg kg^{-1} TOC]	41.218	9	0.000	36.195	26.051	6	0.000	18.200
CPI_{short}	30.716	9	0.000	28.191	16.890	6	0.010	18.000
CPI_{long}	39.956	9	0.000	36.195	25.216	6	0.000	22.800
P_{aq}	35.325	9	0.000	25.190	14.853	6	0.021	14.000
δ^{13}C [‰]	45.219	9	0.000	37.995	30.028	6	0.000	23.600
δ^{15}N [‰]	46.165	9	0.000	40.997	31.663	6	0.000	29.200
χ [10^{-8} m^3 kg^{-1} dw]	44.184	9	0.000	36.195	29.069	6	0.000	21.200
L^* (D65)	42.415	9	0.000	40.997	27.689	6	0.000	23.600
a^* (D65)	44.108	9	0.000	39.396	30.311	6	0.000	29.200
b^* (D65)	37.053	9	0.000	28.191	26.286	6	0.000	21.200
N_{mic} [µg g^{-1} dw]	33.908	9	0.000	24.190	21.355	6	0.002	16.600
C_{mic} [µg g^{-1} dw]	41.157	9	0.000	36.195	28.730	6	0.000	26.000
Nitrate [µg N g^{-1} dw]	35.111	9	0.000	28.191	25.720	6	0.000	19.600
Ammonium [µg N g^{-1} dw]	41.891	9	0.000	37.995	26.021	6	0.000	15.000
DON [µg g^{-1} w]	43.210	9	0.000	39.396	27.202	6	0.000	21.200
DOC [µg g^{-1} dw]	31.265	9	0.000	28.992	25.586	6	0.000	21.400

BGD
8, 10119–10154, 2011

Table 3. Factor loadings table obtained from PCA performed with all paddy soils and all parameters illustrated in Fig. 5d.

analysis 1	factor loadings			
	1	2	3	4
TC	0.932	0.306	−0.054	−0.061
TN	0.882	0.427	−0.062	0.127
TS	0.836	0.267	−0.102	−0.336
C/N	−0.028	−0.122	0.108	−0.790
TOC/S	0.373	0.611	−0.041	0.396
TOC	0.885	0.437	−0.063	−0.056
CPI_{short}	0.274	−0.488	0.509	0.005
CPI_{long}	−0.028	−0.264	−0.842	0.265
P_{aq}	0.100	−0.155	0.687	0.017
$\delta^{15}N$	−0.860	0.085	0.317	−0.127
$\delta^{13}C$	−0.874	0.021	0.222	−0.035
L^*(D65)	−0.075	0.044	0.651	0.556
a^*(D65)	−0.646	−0.522	−0.439	−0.157
b^*(D65)	−0.817	−0.369	−0.019	−0.149
χ	−0.370	−0.839	0.263	−0.096
Lipids	0.835	0.502	0.122	−0.117
n-Alkanes	0.677	0.460	0.320	−0.389
TIC	−0.170	−0.926	0.078	0.002
N_{mic}	−0.410	−0.183	0.571	0.415
C_{mic}	0.711	0.143	0.318	0.541
Nitrate	0.464	0.370	−0.634	−0.033
Ammonium	0.648	0.065	0.186	0.533
DON	0.529	0.507	0.205	0.038
DOC	0.845	−0.026	0.218	0.155

BGD
8, 10119–10154, 2011

Table 4. Factor loadings table obtained from PCA performed with all paddy soils and all conservative parameters illustrated in Fig. 5e.

analysis 2	factor loadings			
	1	2	3	4
TC	0.987	0.009	0.006	−0.018
TN	0.964	0.028	−0.191	−0.111
TS	0.897	0.001	0.304	−0.051
C/N	−0.019	0.137	0.815	0.035
TOC/S	0.533	0.093	−0.512	−0.383
TOC	0.983	0.052	−0.018	−0.133
CPI_{short}	0.094	0.195	−0.025	0.812
CPI_{long}	−0.064	−0.909	−0.173	−0.101
P_{aq}	−0.008	0.608	−0.133	0.532
$\delta^{15}N$	−0.799	0.383	0.170	−0.249
$\delta^{13}C$	−0.840	0.322	0.006	−0.158
L^*(D65)	−0.106	0.579	−0.617	0.233
a^*(D65)	−0.748	−0.543	0.283	0.047
b^*(D65)	−0.899	−0.081	0.233	0.036
χ	−0.614	−0.054	0.230	0.654
Lipids	0.950	0.247	0.031	−0.141
n-Alkanes	0.796	0.445	0.306	−0.106

BGD
8, 10119–10154, 2011

Table 5. Factor loadings table obtained from PCA performed with all sites and all conservative parameters illustrated in Fig. 5f.

analysis 3	factor loadings			
	1	2	3	4
TC	0.942	−0.161	0.207	0.061
TN	0.829	−0.245	0.477	−0.022
TS	0.871	−0.211	−0.345	0.174
C/N	0.055	0.629	−0.391	0.075
TOC/S	0.271	0.050	0.894	−0.200
TOC	0.871	−0.139	0.451	0.001
CPI_{short}	0.131	0.758	0.302	0.061
CPI_{long}	−0.154	−0.207	0.152	−0.910
P_{aq}	0.093	0.568	−0.085	0.773
$\delta^{15}N$	−0.625	−0.465	−0.114	−0.131
$\delta^{13}C$	−0.307	0.771	−0.365	0.162
L^*(D65)	0.003	0.798	0.089	0.279
a^*(D65)	−0.750	−0.293	−0.406	−0.047
b^*(D65)	−0.813	−0.038	−0.039	0.416
χ	−0.261	0.067	−0.904	0.033
Lipids	0.905	0.260	0.100	0.187
n-Alkanes	0.797	0.124	0.286	0.170

BGD
8, 10119–10154, 2011

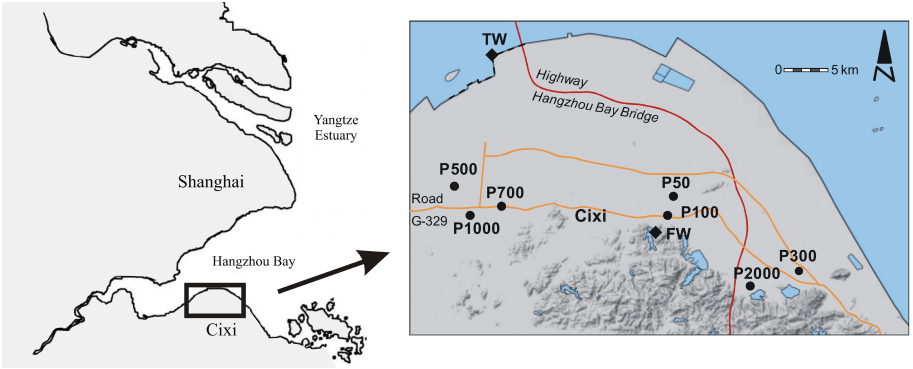

Fig. 1. Map of study area, depicting generations of dykes constructed for land reclamation purposes (adapted from Jahn et al., 2011) and sampling locations.

BGD
8, 10119–10154, 2011

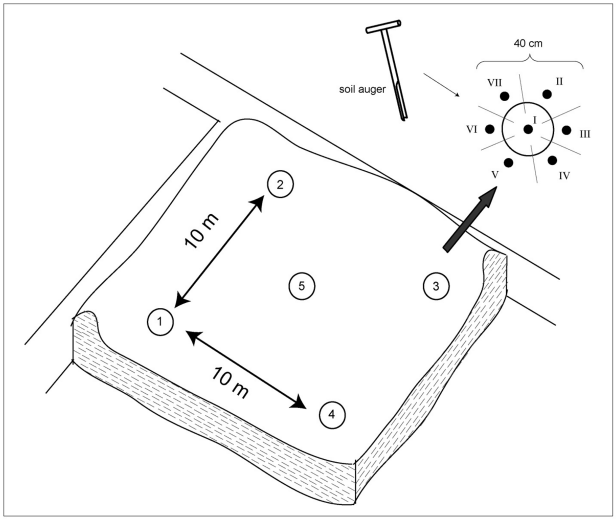

Fig. 2. Design for recovery of field replicates at 10 m regular spacing, each of which is a composite of 7 subsamples taken at 40 cm regular spacing.

BGD
8, 10119–10154, 2011

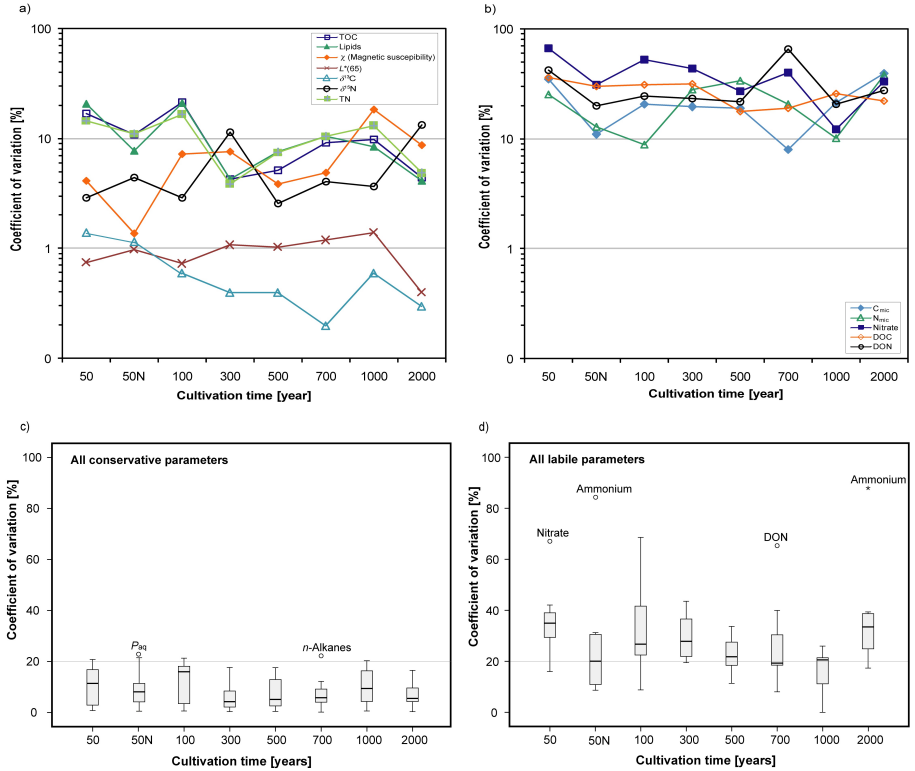

Fig. 3. Coefficients of variation for paddy soil sites sampled in pentuplicate with **(a)** conservative parameters, **(b)** labile parameters, **(c)** box and whisker-plots showing median value, 75 percentile, 90 percentile and outliers for conservative parameters, **(d)** same as **(c)** but for labile parameters.

BGD
8, 10119–10154, 2011

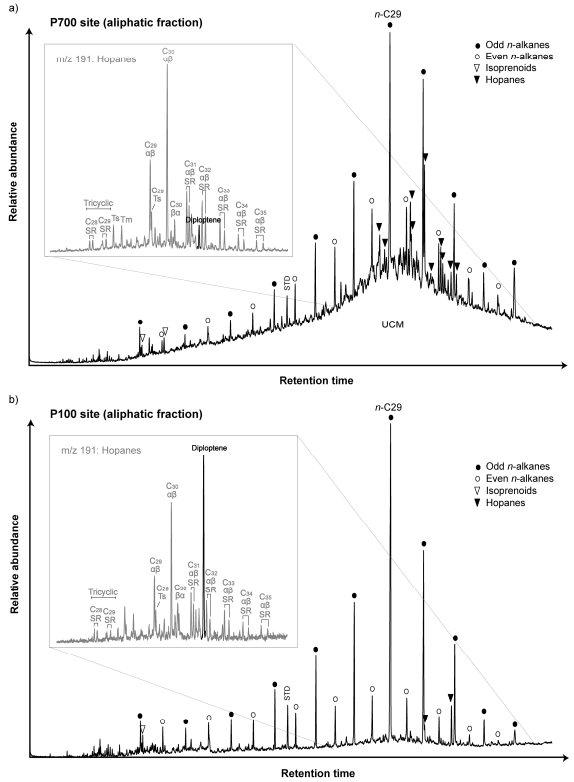

Fig. 4. GC/MS total ion traces obtained from aliphatic hydrocarbon fractions, with major peaks labelled for identification. The inset shows the extracted mass fragmentogram of *m/z* = 191, indicative for tri- and pentacyclic triterpenoids. Diploptene marked black is indicative of recent bacteria, hopanes and tricyclics in grey derive from fossil fuel contamination. Peaks are labelled according to number of carbon atoms per molecule and isomerisation at position C17, C21 and C22. Ts = trisnorhopane, Tm = Trisnorneohopane. Note high abundance of fossil fuel hopanes vs. recent diploptene in P700.

BGD
8, 10119–10154, 2011

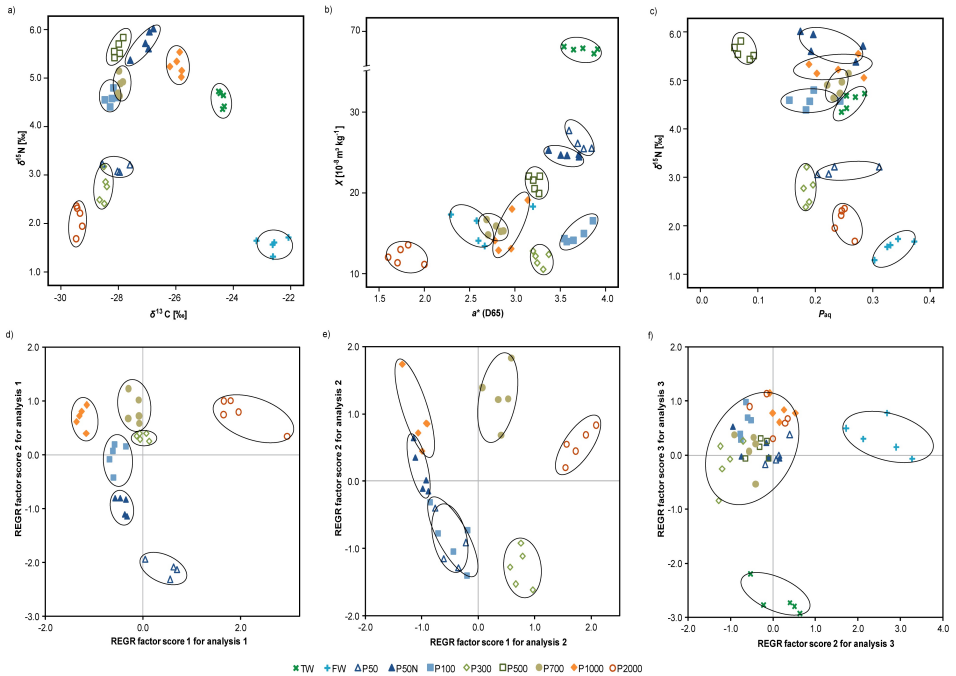

Fig. 5. Discrimination of variance between sites versus in-site using descriptive statistics **(a)** to **(c)** and non-parametric and multivariate methods **(d)** to **(f)**. The factor plots obtained from PCA are shown for **(d)** all paddy soils using all parameters, **(e)** all paddy soils using exclusively conservative parameters, and **(f)** using all paddy soils, non-paddy soil P500 and substrates FW and TW. Note that discrimination of substrates was achieved best, when using the 2nd and 3rd factor rather than 1st and 2nd factor as in **(d)** and **(e)**.

BGD
8, 10119–10154, 2011

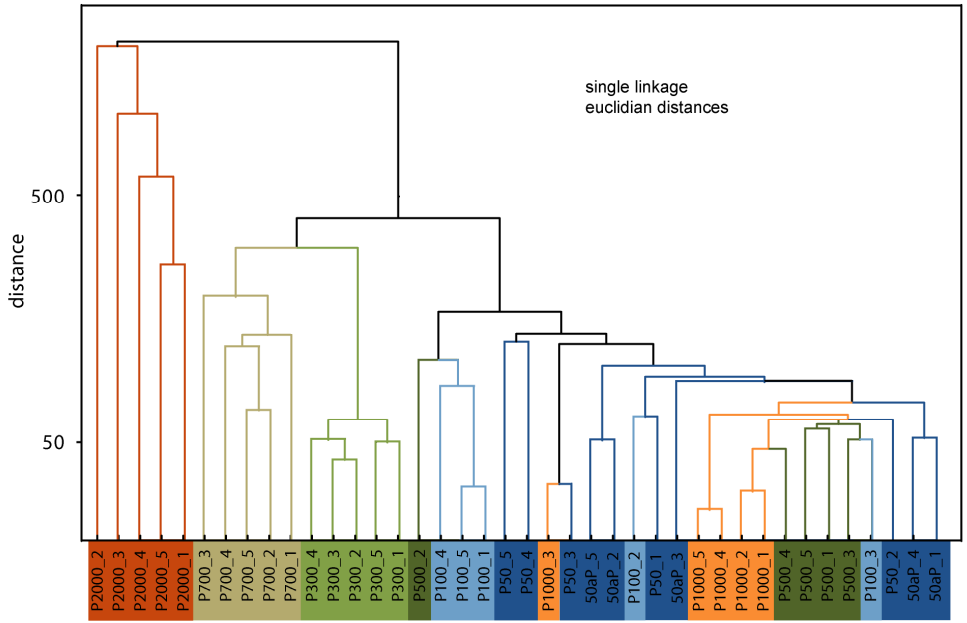

Fig. 6. Cluster diagram for paddy soils and non-paddy soil P500, constructed using all parameters.

BGD
8, 10119–10154, 2011

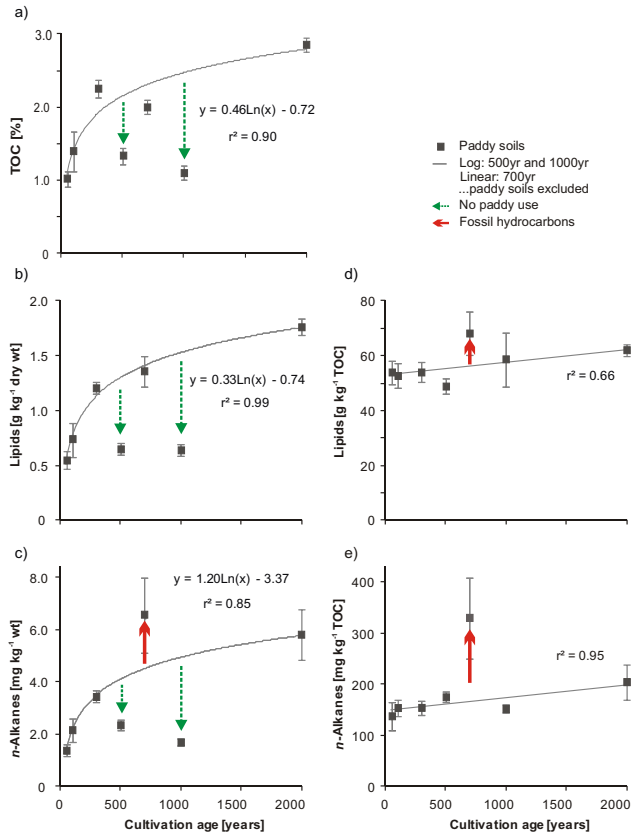

Fig. 7. Accumulation trends of **(a)** TOC, **(b)** lipid yield and **(c)** *n*-alkane yield normalized to dry sample weight and **(d)** lipid yield and **(e)** *n*-alkane yield normalized to TOC over cultivation time. Arrows denote deviations from natural accumulation trends due to human disturbance of the paddy soil system. The P500 was used as upland field and only recently converted to paddy soil use, the P1000 site experienced topsoil removal and admixture of other soil material in the course of dyke maintenance work, the P700 site suffers from petroleum contamination.

BGD
8, 10119–10154, 2011