Report 2004/06

Soil data derived from SOTER for studies of carbon stocks and change in the Indo-Gangetic Plains (India)

(version 1.0)

(GEF-SOC Project: GFL-2740-02-4381)
(VROM Project: DGM-200207282)

Niels H Batjes
Tapas Bhattacharyya
C. Mandal
Koos Dijkshoorn
Dilip K. Pal
E. Milne
K.S. Gajbhiye

(December 2004)

Global Environment Facility

UNEP

United Nations Environment Programme

VROM

Netherlands Ministry of Housing, Spatial Planning and the Environment

I C A R

National Bureau of Soil Survey and Land Use Planning (ICAR)

ISRIC World Soil Information

Disclaimer:

Correct citation:

Batjes NH, Bhattacharyya T, Mandal C, Dijkshoorn K, Pal DK, Milne E and Gajbhiye, KS 2004. Soil data derived from SOTER for studies of carbon stocks and change in the Indo-Gangetic Plains (India) (ver. 1.0; GEFSOC Project). Report 2004/06. ISRIC – World Soil Information, Wageningen and NBSS&LUP (ICAR), India.

Inquiries:

c/o Director, ISRIC – World Soil Information
PO Box 353
6700 AJ Wageningen
The Netherlands
Telefax: +31-(0)317-471700
E-mail: soil.isric@wur.nl

Contents

List of Tables

List of Figures

SUMMARY

This report presents a harmonized set of soil parameter estimates for the Indo-Gangetic Plains (IGP) of India at scale 1:1 000 000. It has been derived from soil and terrain data collated in SOTER format by staff of the National Bureau of Soil Survey and Land Use Planning (NBSS&LUP), India. The data set has been prepared for use in the project on "Assessment of soil organic carbon stocks and change at national scale" (GEFSOC), which has IGP-India as one of its four case study areas.

The land surface of IGP-India has been characterized using 36 unique SOTER units, corresponding with 497 polygons. The major soils of these units have been described using 36 profiles, selected by national soil experts as being representative for these units. The associated soil analytical data have been derived from soil survey reports.

Gaps in the measured soil profile data have been filled using a scheme of taxotransfer rules. Parameter estimates are presented by soil unit for fixed depth intervals of 0.2 m to 1 m depth for: organic carbon, total nitrogen, pH(H_2O), CEC_{soil}, CEC_{clay}, base saturation, effective CEC, aluminum saturation, $CaCO_3$ content, gypsum content, exchangeable sodium percentage (ESP), electrical conductivity of saturated paste (ECe), bulk density, content of sand, silt and clay, content of coarse fragments, and available water capacity. These attributes have been identified as being useful for agro-ecological zoning, land evaluation, crop growth simulation, modelling of soil carbon stocks and change, and analyses of global environmental change.

The current parameter estimates should be seen as best estimates based on the current selection of soil profiles and data clustering procedure. Taxotransfer rules have been flagged to provide an indication of the possible confidence in the derived data.

Results are presented as summary files and can be linked to the 1:1M scale SOTER map in a GIS, through the unique SOTER-unit code.

The secondary SOTER data set for IGP-India is considered appropriate for exploratory studies at regional scale. Correlation of soil analytical data, however, should be done more rigorously when more detailed scientific work is considered.

Keywords: soil parameter estimates, IGP-India, environmental modelling, WISE database, SOTER database, secondary data set

1. INTRODUCTION

This study[1] has been carried out in the framework of the GEF co-funded project, *Assessment of soil organic carbon stocks and change at national scale* (GFL-2740-02-4381). The project will develop and demonstrate generic tools to quantify the potential impact of land management and climate scenarios on change in soil carbon stocks at national and sub-national level. It involves participation from national scientists in Brazil, India, Jordan and Kenya working closely with data management and modeller groups in Austria, France, the Netherlands, the United Kingdom and the USA.

The main research objectives, summarized on the project website[2], are:

1. To identify and use long-term, plot-scale, experimental datasets to systematically evaluate and refine modelling techniques to quantify carbon sequestration potential in tropical soils;
2. To define, collate and format national-scale soils, climate and land-use datasets and to use them in the development of coupled modelling-GIS tools to estimate soil carbon stocks;
3. To demonstrate these tools by estimating current soil organic carbon stocks at country-scale – using the Indo-Gangetic Plains (India), Jordan, Kenya and Amazon-Brazil as case studies – and to compare these estimates with the existing techniques of combining soil mapping units and interpolating point data;
4. To quantify the impact of defined changes in land use and climate on carbon sequestration in soils with a view to assisting in the formulation of improved policies to optimise resource use in the four case-study countries.

This report presents parameter estimates for the major soils of IGP-India, at scale 1:1M, in the common format required for the modelling phase. The materials and methods are described in Chapter 2, with special focus on the procedure for preparing the

---

[1] Having the same scope, the structure and body of the report for IGP-India are similar to those prepared for the other case study areas: Brazil, Jordan and Kenya.

[2] http://www.reading.ac.uk/GEFSOC

secondary SOTER sets. Results are discussed in Chapter 3, while concluding remarks are drawn in Chapter 4. The structure of the various output tables is documented in the Appendices, which also include a brief description of the contents of the secondary SOTER file for IGP-India (Appendix 5).

2. MATERIALS AND METHODS

2.1 Biophysical setting

A detailed account of regional geography in India may be found in Singh (1971). The country is bound by Pakistan to the northwest, China, Bhutan and Nepal to the north, Myanmar and Bangladesh to the east, the Bay of Bengal to the southeast and the Arabian Sea to the southwest. In view of India's large area (~3.2 million km^2), the GEFSOC case study area has been limited to the Indo-Gangetic Plains (IGP-India). This region is particulary important to food production, particularly cereals such as rice and wheat, and is home to about 40% of India's population. Over the years, IGP has been considered as the most productive agricultural region of the country. IGP-India played an important role during the Green Revolution, in the mid 1960's, during which crop yields increased significantly (Aggarwal *et al.* 2004; Ladha *et al.* 2003). However, in view of the ever-increasing human and animal population the demand on soils for food production continues to increase.

The case study area covers about 480,000 km^2 (15%) of the country. It includes parts of the states of Bihar, Haryana, Punjab, Rajasthan, Uttar Praddesh, West Bengal, Tripura and Delhi.

Elevation in IGP-India ranges from about 0 m to 330 m above mean sea level. Annual rainfall increases from 300 to 1600 mm, from west to east. IGP-India thus encompasses both hot arid and humid regions (Velayutham *et al.* 1999). The soils of IGP-India are of great importance for the agricultural production of the country

(Velayutham *et al.* 2002) — a wide range of information is available about their properties and recommended management (e.g., Bhattacharyya and Pal 2003; Bhattacharyya *et al.* 2004b; Bhattacharyya *et al.* 2000; Lal *et al.* 1994; Mandal *et al.* 1999; Murthy *et al.* 1982; Nambiar 1994; Pal *et al.* 2000; Pal *et al.* 2003; Velayutham *et al.* 1999).

Ladha *et al.* (2003) investigated the extent of yield stagnation or decline in the Indo-Gangetic Plains of South Asia to identify possible causes of this decline. Important challenges associated with maintaining and increasing food production in IGP-India include the increasing competition for land resources by non-agricultural sectors and by the deterioration of agri-environments and water resources (Aggarwal *et al.* 2004). The indiscriminate use of inputs, such as chemical fertilizers and irrigation, and the change of climate over the years in this part of India, are causing degradation of formerly fertile lands, reducing crop production (Aggarwal *et al.* 2004). Other potential causes of yield decline include depletion of soil carbon, nitrogen and zinc, and a reduced availability of phosphorus (Ladha *et al.* 2003). In the semi-arid part of the IGP natural chemical degradation in terms of pedogenic calcium carbonate and concomittant development of sodicity has been identified as a principal impediment of sustained agricultural production (Bhattacharyya *et al.* 2004b; Bhattacharyya *et al.* 2000; Pal *et al.* 2000; Pal *et al.* 2003; Velayutham *et al.* 2000).

The IGP, India with about 13% geographical coverage produces nearly 50% of the food grains for 40% of the total population of India. Soils under arid and semi-arid climates in the IGP cover 16.4 million ha and lack in organic carbon due to high rate of decomposition. The total carbon stock in the IGP, India has been estimated as 0.76 Pg (Pg= 10^{15} g) in the first 30 cm depth of soils of which 0.63 Pg is contributed by soil organic carbon (SOC) and 0.13 Pg by soil inorganic carbon (SIC) (Bhattacharyya *et al.* 2004b). The SOC, SIC and total C stock in the IGP, India account for 6.4%, 2.3% and 5.5% of the total SOC, SIC and total carbon stock of India (Bhattacharyya *et al.* 2004b; Bhattacharyya *et al.* 2000). It is in this respect soil database documentation through SOTER appears important for modeling future changes in carbon level in soils.

According to Velayutham *et al.* (2000), about 70% of India falls in the "SOC deficient zone" based on the low content of organic carbon

(SOC) in the topsoil (<1%; 0-30 cm); this zone includes a significant part of IGP-India. Decreasing SOC levels were often due to intensive agricultural practices, with limited returns of crop residues and nutrients to the soil. Increases in SOC, upon adoption of "best management" practices on SOC-depleted (degraded) soils have been reported for parts of IGP-India (Bhattacharyya *et al.* 2004b; Dwivedi *et al.* 2003; Goswami *et al.* 2000).

Land use change, the main driving factors of which have been described by Lambin *et al.* (2003), in combination with climate change will have a varying impact on net primary production and soil carbon turnover, and thus soil carbon stocks, in the various natural regions of IGB-India. A selection of these aspects will be studied by the consortium during the modelling phase of the GEFSOC project.

2.2 Soil data compilation

ISRIC staff organized a SOTER-training (19-30 November 2003) at the National Bureau of Soil Survey and Land Use Planning (NBSS&LUP), Nagpur, as part of the capacity building component of the GEFSOC project (Dijkshoorn 2003). About ten scientists, technical officers and research fellows attended the training programme. The International Project Coordinator greatly contributed to the logisitics (Figure 1).

The subsequent compilation of the actual soil and terrain data for IGP-India, at scale 1:1M, was the responsibility of NBSS&LUP (Bhattacharyya *et al.* 2004a). Three regional teams undertook the work, which was carried out between December 2003 and August 2004.


Figure 1. SOTER training at NBSS&LUP, Nagpur

2.3 SOTER methodology

The SOTER methodology allows mapping and characterization of areas of land with a distinctive, often repetitive, pattern of landform, lithology, surface form, slope, parent material, and soils (van Engelen and Wen 1995). The approach resembles physiographic or land systems mapping. The collated materials are stored in a SOTER database linked to GIS, permitting a wide range of environmental applications (e.g., Batjes and Dijkshoorn 1999; Falloon *et al.* 1998; Mantel *et al.* 2000; Nachtergaele *et al.* 2002). The SOTER methodology is mainly applied at scales ranging from 1:250 000 to 1:5M.

Each SOTER database is comprised of two main elements, a geographical component and an attribute data component (Figure 2). The *geographical database* holds information on the location, extent and topology of each SOTER unit. The *attribute database* describes the characteristics of the spatial unit and includes both area data and point data. A geographical information system (GIS) is used to manage the geographic data, while the attribute data are handled in a relational database management system (RDBMS).

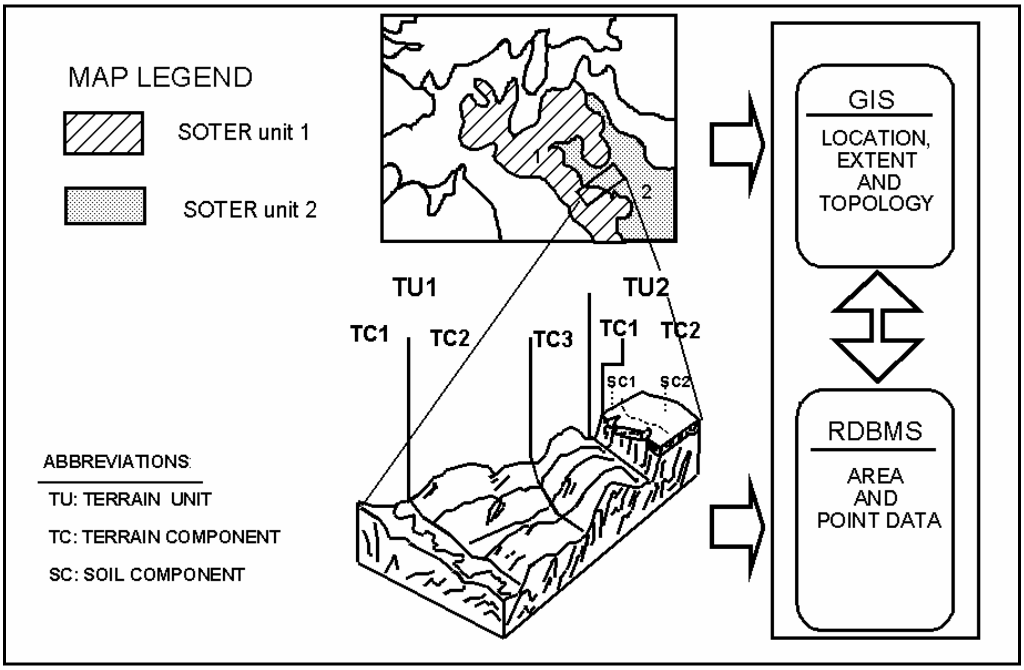

Figure 2. Schematic representation of two SOTER units and their terrain and soil components

Each SOTER unit in the geographic database has a unique identifier, called SOTER unit-ID (SUID). This primary key provides a link to the attribute data for its constituent terrain, terrain component(s) (TCID) and soil component(s) (SCID) (see Appendix 4).

Each soil component within a SOTER unit is described by a profile (PRID), identified by the national soil experts as being regionally representative. This selection is largely based on purposive sampling (Webster and Oliver 1990). Profiles in SOTER are characterised according to the Revised Legend of FAO (1988) or World Reference Base for Soil Resources (WRB 1998). Representative profiles are mainly selected from available soil survey reports as the SOTER program does not involve new ground surveys. Batjes (1999) reviewed issues of data acquisition, quality control and sharing in the context of SOTER projects, some of which proved of particular importance for the Indian data (Shrikantia 1999).

A comprehensive description of the methodology and coding conventions is given by Van Engelen and Wen (1995). The SOTER attribute data can be managed with an automated data entry facility (Tempel 2002). SOTER uses commercially available MS Access® and ArcView® software.

2.4 Preparation of secondary SOTER data sets

2.4.1 List of soil parameters

Special attention has been paid to the key attributes required for the spatial runs of the two organic carbon models considered in the GEF-SOC project: RothC and Century. These are: the extent and type of soil, soil drainage status, content of clay, content of organic carbon, and bulk density per depth layer (Falloon *et al.* 1998; Paustian *et al.* 1997). This limited set has been expanded to include 18 soil parameters (Table 1) commonly required in studies of agro-ecological zoning, food productivity, soil gaseous emissions/sinks and environmental change (see Batjes *et al.* 1997; Bouwman *et al.* 2002; Cramer and Fischer 1997; Fischer *et al.* 2002; Scholes *et al.* 1995).

Table 1 does not consider soil hydraulic properties. Although essential for many simulation studies, these properties are seldom measured during soil surveys. As a result, the corresponding records are lacking in databases such as SOTER and WISE. Information on soil hydraulic properties and pedotransfer functions for Western Europe and the USA may be found in auxiliary databases (see Nemes *et al.* 2003; Wösten *et al.* 1998) but similar work for tropical soils has just begun (Tomasella and Hodnett 1998; van den Berg *et al.* 1997).

Table 1. List of soil parameters

Organic carbon
Total nitrogen
Soil reaction (pH_{H2O})
Cation exchange capacity (CEC_{soil})
Cation exchange capacity of clay size fraction (CEC_{clay}) • ‡
Base saturation (as % of CEC_{soil}) ‡
Effective cation exchange capacity (ECEC) † ‡
Aluminum saturation (as % of ECEC) ‡
$CaCO_3$ content
Gypsum content
Exchangeable sodium percentage (ESP) ‡
Electrical conductivity of saturated paste (ECe)
Bulk density
Coarse fragments (volume %)
Sand (mass %)
Silt (mass %)
Clay (mass %)
Available water capacity (AWC; from -33 to -1500 kPa; % w/v) ‡ □

‡ Calculated from other measured soil properties.
† ECEC is defined as exchangeable (Ca^{++}+Mg^{++}+K^{+}+Na^{+}) + exchangeable (H^{+}+Al^{+++}) (van Reeuwijk 2002).
• CEC_{clay} was calculated from CEC_{soil} by assuming a mean contribution of 350 $cmol_c$ kg^{-1} OC, the common range being from 150 to over 750 $cmol_c$ kg^{-1} (Klamt and Sombroek 1988).
□ The soil water potential limits for AWC conform to USDA standards (Soil Survey Staff 1983). Values shown have not been corrected for the presence of coarse fragments.

2.4.2 Consistency and integrity checks of the primary data

A preliminary version of the soil attribute and GIS shape-files for IGP-India was made available to ISRIC, through GEFSOC headquarters in Reading, for comments in September 2004 (Bhattacharyya *et al.* 2004a). Subsequently, the primary data were re-processed at ISRIC to generate a secondary SOTER data set for the region.

In India, soil series are classified according to USDA Soil Taxonomy (Lal *et al.* 1994; Murthy *et al.* 1982; Soil Survey Staff 1975, 1998).

NBSS&LUP therefore re-classified the selected soil series according to WRB (1998) to conform with SOTER standards.

Complete sets of analytical data, as required for SOTER (van Engelen and Wen 1995), were not always available for the representative profiles, as is the case in many other countries. Hence it was necessary to apply a taxotransfer rule-based approach to fill gaps in the measured data (Batjes 2003). To this avail, the representative profiles were re-classified according to the Revised Legend (FAO 1988). The guiding principle in this conversion were the USDA and WRB classifications, as provided by NBSS&LUP expert staff, complemented with the available soil morphological, physical and analytical data.

For some of the WRB-units encountered, such as sodic Cambisols, there was no suitable equivalent in the Revised Legend so that sodic phases were used (e.g. Eutric Cambisol, sodic phase). Information on soil phases, however, is not considered explicitly in the taxotransfer scheme.

2.4.3 Procedure for filling gaps in the measured data

In several instances there were gaps in the measured data, soil physical attributes in particular as these were seldom measured during the ground surveys (Lal *et al.* 1994; Murthy *et al.* 1982). The occurrence of such gaps limits the direct use of primary SOTER data in models. Therefore, the GEFSOC project developed a standardized procedure to fill gaps in key measured data in three main stages (Batjes 2003):

a) Collating additional measured soil data where these exist, in the uniform SOTER format;
b) Using national expert estimates and common sense to fill selected gaps in a secondary data set;
c) Using taxotransfer rule (TTR) derived soil parameter estimates for similar FAO soil units, as derived from the global WISE profile database (Batjes 2002; Batjes and Bridges 1994).

The desirability of the above stages decreases from highest (a) to lowest (c).

a) Collating additional measured data

The main soils of IGP-India were characterized using 36 representative profiles for well documented Benchmark series. This corresponds with a density of about 0.08 profiles per 1000 km^2. All profiles were sampled and analyzed between 1972 and 1983.

b) Using expert-based estimates

There was no need to define "synthetic" profiles for IGP-India, unlike for Jordan and Kenya where it was necessary to estimate all soil parameters for an entire (synthetic) profile in some instances (Batjes and Gicheru 2004; Batjes *et al.* 2003). Thus possible gaps in the primary data were only filled using taxotransfer rules.

c) Application of taxotransfer rules

The taxotransfer (TTR) approach was developed initially for application with the Soil Map of the World, in collaborative studies with FAO and IIASA, using soil analytical data held in ISRIC's WISE database (Batjes 2002; Batjes *et al.* 1997). The methodology has been modified in the framework of the GEFSOC project for use with national scale SOTER databases, as detailed by Batjes (2003).

3. RESULTS AND DISCUSSION

3.1 General

IGP-India has been described using 36 unique SOTER units, each characterised by a regionally representative soil profile, corresponding with 497 mapped polygons.

Although having the same database structure, the Indian data set differed from the other SOTER databases prepared for the GEFSOC project (Batjes *et al.* 2004; Batjes and Gicheru 2004; Batjes *et al.* 2003) in that all SOTER units were characterised using one soil series. In the other case study areas, however, most map units were defined as soil associations or complexes.

3.2 SOTER unit composition

The main soil unit of each SOTER unit has been characterised in an MS Access® table called *SOTERunitComposition* (see Appendix 1). For example, the SOTER or map unit with NEWSUID number IN19 is coded as SNh1 (Figure 3). The 19th[th] map unit for IGP-India is comprised of 100 per cent (coded as 1: 80-100%) of haplic Solonetz (SNh).

SOTERunitComposition : Table

NEWSU	SoilMapUnit	SOIL1	PROP1	Profile-ID1	SOIL2	PROP2	Profile-ID2
IN10	SCn1 [10]	SCn	100	IN-HRPR			
IN11	SNh1 [11]	SNh	100	IN-SKIT			
IN12	CMe1 [12]	CMe	100	IN-DHD			
IN13	CMe1 [13]	CMe	100	IN-JGJT			
IN14	ARh1 [14]	ARh	100	IN-BHNR			
IN15	CMe1 [15]	CMe	100	IN-BRPR			
IN16	CMe1 [16]	CMe	100	IN-BSRM			
IN17	LVg1 [17]	LVg	100	IN-ITW			
IN18	CMe1 [18]	CMe	100	IN-SMR1			
IN19	SNh1 [19]	SNh	100	IN-AKBP			

Record: 1 of 36

Figure 3. Characterization of SOTER units in terms of their main component soils – with their representative profile – and their relative extent

3.3 Soil parameter estimates

The depth-weighted primary and TTR-derived data, by layer, for the 18 soil properties under consideration (Table 1) have been stored in a secondary SOTER data set (Figure 4); the cut-off point for applying any TTR is $n_{WISE} < 5$. Appendix 2 shows the structure of the corresponding file.

SOTERparameterEstimates : Table

CLAF	PRID	Drain	Laye	TopDep	BotDep	CFRAC	SDTO	STPC	CLPC	PSCL	BULI	TAWC	CECS	BSAT	CECc	PHAQ	TCEQ	GYPS	ELCO	TOTC
LVv	IN-EKCH	W	D1	0	20	0	12	48	40	F	1.30	14	21	76	49	6.5	0	0	0	3.7
LVv	IN-EKCH	W	D2	20	40	0	11	43	46	F	1.38	10	24	77	50	6.7	0	1	0	2.5
LVv	IN-EKCH	W	D3	40	60	0	8	42	50	F	1.37	14	27	81	53	6.9	0	1	0	1.5
LVv	IN-EKCH	W	D4	60	80	0	4	47	49	F	1.45	14	27	81	54	7.2	0	1	0	1.0
LVv	IN-EKCH	W	D5	80	100	0	4	46	50	F	1.50	12	27	81	54	7.2	0	1	0	1.0
CMe	IN-FTHPR	W	D1	0	20	0	82	12	6	C	1.50	3	4	90	52	7.8	0	0	1	1.2
CMe	IN-FTHPR	W	D2	20	40	0	79	13	8	C	1.50	5	4	82	48	7.8	0	0	1	0.6
CMe	IN-FTHPR	W	D3	40	60	0	79	12	9	C	1.40	4	5	82	53	7.7	0	0	1	0.7
CMe	IN-FTHPR	W	D4	60	80	1	79	13	8	C	1.44	5	5	88	61	7.8	0	0	1	0.4
CMe	IN-FTHPR	W	D5	80	100	4	77	16	7	C	1.51	5	4	85	54	7.7	0	0	1	0.3
SNg	IN-GBDN	P	D1	0	20	1	36	42	22	M	1.48	9	11	99	45	9.9	9	0	5	1.9
SNg	IN-GBDN	P	D2	20	40	0	34	37	29	M	1.39	14	13	99	43	9.6	7	0	3	1.5
SNg	IN-GBDN	P	D3	40	60	1	35	36	29	M	1.39	10	14	100	46	9.6	5	0	2	1.4
SNg	IN-GBDN	P	D4	60	80	1	36	35	29	M	1.40	10	14	100	47	9.5	4	0	1	1.3
SNg	IN-GBDN	P	D5	80	100	1	37	33	30	M	1.40	10	13	100	43	9.4	2	0	1	1.0

Record: 50 of 176

Figure 4. Example of ultimate result of the application of the TTR-scheme and depth weighing for three profiles

The type of TTR used, if any, has been flagged by profile and depth layer in a separate table (Figure 5, Appendix 3). The field TTRsub indicates that the data substitution for a given attribute, in the secondary SOTER set, is based on WISE-derived parameter estimates for similar soil units. Otherwise, should the corresponding population in WISE be too small ($n_{WISE} < 5$) for a meaningful substitution, the rules are flagged under TTRmain (see Batjes 2003).

Each flag consists of a sequence of letters followed by a numeral (see under TTRsub and TTRmain in Figure 5). The letters indicate soil attributes for which a TTR has been applied (Figure 6). The number code reflects the size of the sample population in WISE, after outlier rejection, on which the statistical analyses were based (Table 2).

SOTERflagTTRrules : Table

CLAF	PRID	Layer	Newtopdep	Newbotdep	TTRsub	TTRmain
SNg	IN-GBDN	D1	0	6	c3h3r2	o3
SNg	IN-GBDN	D1	6	20	c3h3r2	o3
SNg	IN-GBDN	D2	20	24	C2h3r3	o3
SNg	IN-GBDN	D2	24	40	C2h3r3	o3
SNg	IN-GBDN	D3	40	48	C2H2r3	A4o3
SNg	IN-GBDN	D3	48	60	C2H2r3	A4o3
SNg	IN-GBDN	D4	60	73	C2H2R3	A4o3
SNg	IN-GBDN	D4	73	80	C2d3H2R3	A4o3
SNg	IN-GBDN	D5	80	98	C3d3h3R3	A4o3
SNg	IN-GBDN	D5	98	100	C3d3h3R3	A4o3
GLe	IN-HNGR	D1	0	11	h2o3r1	A3
GLe	IN-HNGR	D1	11	20	c1h2o3r1	A3
GLe	IN-HNGR	D2	20	25	c1h1o3r1	A4
GLe	IN-HNGR	D2	25	40	h1o3r1	A4
GLe	IN-HNGR	D3	40	56	h1o3r1	A4
GLe	IN-HNGR	D3	56	60	c1h1o3r1	A4
GLe	IN-HNGR	D4	60	80	c1h1o3r1	A4
GLe	IN-HNGR	D5	80	84	c1h1o3r2	A3
GLe	IN-HNGR	D5	84	100	h1o3r2	A3

Record: 102 of 303

Figure 5. Flagging of taxotransfer rules by profile, depth zone and attribute

Table 2. Criteria for defining confidence in the derived data

Code	Confidence level	n_{WISE}
1	Very high	> 30
2	High	15-29
3	Moderate[†]	5-14
4	Low	1-4
-	No data	0

* n_{WISE} is the sample size after the screening procedure (see Figure 6)
[†] The cut-off point in the TTR-approach is $n_{WISE} < 5$

When a small letter is used (figure 5), the substitution considered median data for the corresponding textural class (for example for fine textured and $n_{WISE} > 5$). Otherwise, when a capital is used, this indicates that the substitution is based on the whole set for the corresponding soil unit and depth layer, irrespective of soil texture (i.e. undifferentiated or #). The same coding conventions apply for TTRmain. This is depicted schematically for the upper 0 to 20 cm of a hypothetical profile (IN<u>hyp</u>04):

CLAF	PRID	LAYER	Newtopdep	Newbotdep	TTRsub	TTRmain
Fle	INhyp04	D1	0	18	b3c2j3o3r2	a2h1
Fle	INhyp04	D1	18	20	C3j1	A3h2

Soil parameter estimates based on WISE-derived data, using data for the corresponding major grouping and either the same textural class (small letter) or undifferentiated textural class (capital).

Soil parameter estimates based on WISE-derived data, using data for the corresponding soil unit and same textural class:
- b: Base saturation, 3 (n_{WISE} = 5 –14)
- c: Bulk density, 2 (n_{WISE} = 15 – 29)
- j: Exchangeable sodium percentage, 3 (n_{WISE} = 5 –14)
- o: Volumetric water content, 3 (n_{WISE} = 5 –14)
- r: Total Nitrogen, 2 (n_{WISE} = 15 – 29)

The overall assumption is that the confidence in a TTR-based parameter estimate should increase with the size of the sample populations present in WISE, after outlier rejection. In addition, the confidence in soil parameter estimates listed under TTRsub, will be higher than for those listed under TTRmain.

SWcodes : Table

SOTnam	WISnam	SoilVariable	TTRflag	Comments
---	---	---	y	PSCL estimated from PTR-derived sand, silt and clay
ALSA	ALSA	ALSAT	a	exchangeable Aluminum percentage (% of ECEC)
BSAT	BSAT	BSAT	b	base saturation (% of CECs)
BULK	BULK	BULKDENS	c	bulk density
CECC	CECC	CECCLAY	d	cation exchange capacity of clay fraction (corr. for org. C)
CECS	CECS	CECSOIL	e	cation exchange capacity
CFRAG	GRAV	GRAVEL	f	coarse fragments
CLPC	CLAY	CLAY	g	clay % (see also Y for texture (g, m & n))
ECEC	ECEC	ECEC	h	effective CEC
ELCO	ECE	ECE	i	electrical conductivity
ESP	ESP	ESP	j	exchangeable Na percentage (% of CECs)
GYPS	GYPS	GYPSUM	k	gypsum content
PHAQ	PHH2	PHH2O	l	pH in water
SDTO	SAND	SAND	m	sand %
STPC	SILT	SILT	n	silt %
TAWC	TAWC	TAWC	o	volumetric water content (-33 to - 1500 kPa)
TCEQ	CACO	CACO3	p	carbonate content
TOTC	ORGC	ORGC	q	organic carbon content
TOTN	TOTN	TOTN	r	total nitogen content

Record: 9 of 20

Figure 6. Conventions for coding the various attributes used in the taxotransfer scheme.

A high confidence rating does not necessarily imply that the soil parameter estimates shown will be representative for the soil unit under consideration. Profile selection for SOTER, as for any other database that depends on historic data, is not probabilistic but based on available data and expert knowledge. Several of the soil attributes under consideration in Table 1 are not diagnostic in the Revised Legend (FAO 1988). In addition, some properties are readily modified by changes in land use or management, for example soil salinity upon irrigation and soil organic matter content upon changes in tillage practices.

Table 3. Type and frequency of taxotransfer rules applied

Parameter	Code	Frequency of occurrence (%)		
		TTRsub	TTRmain	Total
ALSA	A	3	80	83
BSAT	B	41	0	41
BULK	C	85	0	85
CECC	D	28	0	28
CECS	E	12	0	12
CFRAG	F	3	0	3
CLPC	G	0	0	0
ECEC	H	98	0	98
ELCO	I	15	0	15
ESP	J	30	0	30
GYPS	K	57	7	64
PHAQ	L	0	0	0
SDTO	M	0	0	0
STPC	N	0	0	0
TAWC	O	74	6	80
TCEQ	P	29	0	29
TOTC	Q	0	0	0
TOTN	R	99	1	100

Note: For definitions of abbreviations see text and Figure 6; also see the footnote in Appendix 3.

Irrespective of the region or country under consideration, a standard list of 18 soil parameters (Table 1) was generated in view of the global scope of the SOTER project. Thus, effective CEC and Al-saturation, for example, while not considered pedo-chemically relevant parameters for the soils of IGP have nonetheless been included in this assessment to have comparable data sets for the four case study areas.

Table 3 summarizes how often a given TTR has been applied as a percentage of the total number of horizons (up to a depth of 100 cm) in the secondary SOTER database; details may be found in table *SOTERflagTTRrules* (see Appendix 3). For example, measured data were always available for the content of sand, silt and clay, soil pH, and organic carbon. For bulk density (BULK), however, measured data were available for 15% of the horizons under consideration — so taxotransfer rules were applied in 85% of the cases. Table 3 also shows that there were no measured data for total nitrogen so that taxotransfer rules were used in 100% of the cases.

3.4 Linkage to GIS

Aggregated information about the SOTER unit composition and results of the TTR-work can now be linked to the SOTER map using GIS. At the national scale, this can be done via the unique SOTER unit identifier (SUID, see Appendix 4). In transnational databases, however, linkage will be through the NEWSUID, which is a combination of the country's ISO code plus the SUID code.

In the primary database, the soil geographic and attribute data for each soil component is stored in a range of relational databases, mainly to enhance data storage and management efficiency. To assist end-users, a new table has been created that incorporates data held in the primary SOTER database and the present information on soil parameter estimates (Figure 7 Appendix 4). Clearly, this wealth of information, although needed for the modelling work, complicates linkage to GIS.

SOTERsummaryFile : Table

NEWSUI	TCID	SCID	Propoı	PRID	CLAF	Drain	Layer	TopDe	BotDep	CFRAG	SDTO	STPC	CLPC	PSCL	BULK	TAWC	CECS	BSAT	PHAQ	TCEC
IN1	1	1	100	IN-MSTWLI	FLc	W	D1	0	20	0	70	16	14	C	1.57	5	4	94	8.0	1
IN1	1	1	100	IN-MSTWLI	FLc	W	D2	20	40	0	64	19	17	M	1.56	6	4	96	8.0	2
IN1	1	1	100	IN-MSTWLI	FLc	W	D3	40	60	0	61	21	18	M	1.57	7	4	95	8.1	2
IN1	1	1	100	IN-MSTWLI	FLc	W	D4	60	80	0	60	21	19	M	1.57	7	4	93	8.2	2
IN1	1	1	100	IN-MSTWLI	FLc	W	D5	80	100	0	69	17	14	M	1.54	6	4	92	8.2	6
IN2	1	1	100	IN-NHLKRA	CLh	I	D1	0	20	1	63	21	16	M	1.30	9	17	100	8.1	5
IN2	1	1	100	IN-NHLKRA	CLh	I	D2	20	40	1	61	21	18	M	1.35	9	15	100	8.2	6
IN2	1	1	100	IN-NHLKRA	CLh	I	D3	40	60	8	62	22	16	M	1.35	9	15	100	8.3	9
IN2	1	1	100	IN-NHLKRA	CLh	I	D4	60	80	14	61	24	15	M	1.34	9	15	100	8.3	10
IN2	1	1	100	IN-NHLKRA	CLh	I	D5	80	100	28	58	28	14	M	1.42	10	15	100	8.3	14
IN3	1	1	100	IN-JSPW	ARc	E	D1	0	20	0	90	5	5	C	1.50	3	6	100	8.2	2
IN3	1	1	100	IN-JSPW	ARc	E	D2	20	40	1	90	6	4	C	1.50	3	7	100	8.1	4
IN3	1	1	100	IN-JSPW	ARc	E	D3	40	60	1	90	6	4	C	1.49	4	6	100	8.2	5
IN3	1	1	100	IN-JSPW	ARc	E	D4	60	80	1	90	5	5	C	1.49	4	4	100	8.3	5
IN3	1	1	100	IN-JSPW	ARc	E	D5	80	100	1	90	5	5	C	1.50	4	4	100	8.3	5

Record: 1 of 176

Figure 7. Excerpt of a SOTER summary file for units IN1, IN2 and IN3.

For visualization and analysis in GIS, it will often be necessary to make an extra selection. For example, in the case of the RothC and Century models, information may be required about the properties of the topsoil – that is layer D1: 0-20 cm – for the dominant soil. In this example, the necessary selection would be for the first Terrain Component (TCID=1), first Soil Component (SCID= 1) and the upper most layer (D1). The corresponding selection is included as a separate table in the secondary database for IGP-India. The database structure is detailed in Appendix 4.

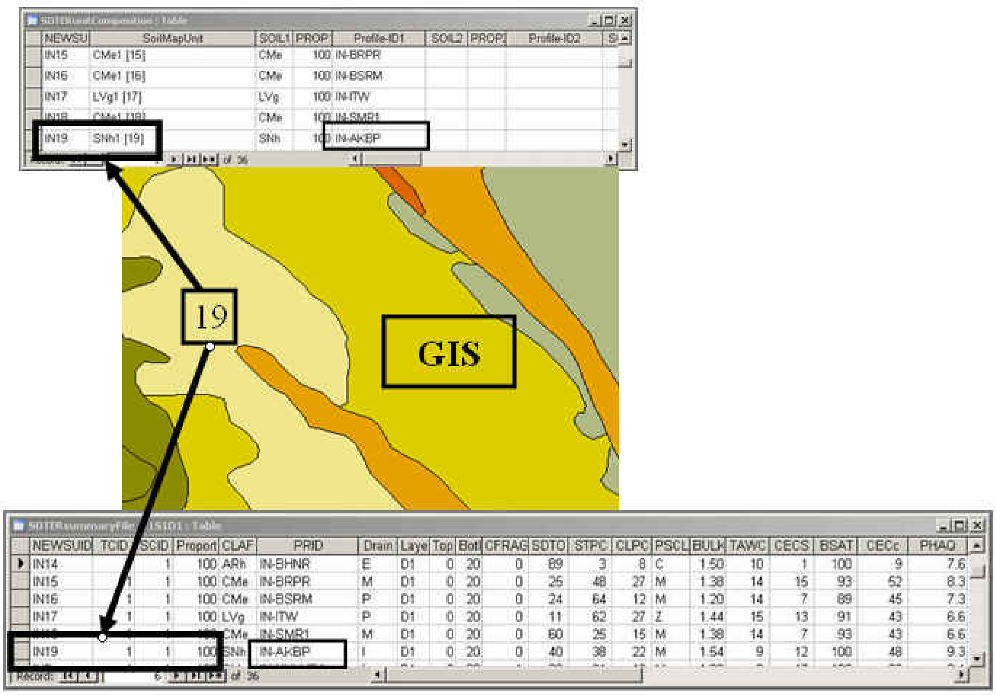

Figure 8. Linking soil parameter estimates for the top 20 cm of the dominant soil of SOTER unit IN19 with the geographical component of SOTER

Figure 8 schematically shows the procedure for linking the various secondary attribute data to the geographical SOTER data held in the GIS. For ease of visualization, it considers only the upper layer (D1) of the spatially dominant (first) soil component of SOTER unit IN19.

All geographic data in SOTER are presented in vector format. However, should grid-based soil layers be required, these can be generated using the convert-to-grid module of the spatial analyst extension to ArcView (ESRI 1996) — the minimum legible delineation implied by the current scale of 1:1M is about 25 km^2. Gridding should be based on the NEWSUID field to permit subsequent linkage with the various attribute tables discussed in this report. The procedure will be the same as depicted in Figure 8.

4. CONCLUSIONS

- Linkage between soil profile data and the spatial component of a SOTER map, for environmental applications, requires generalisation of measured soil (profile) data by soil unit and depth zone. This involves the transformation of variables that show a marked spatial and temporal variation and that may have been determined in a range of laboratories according to various analytical methods.
- A pragmatic approach to the comparability of soil analytical data has been adopted. This is considered appropriate at the present scale of 1:1M but must be done more rigorously when more detailed scientific work is considered.
- The present set of soil parameter estimates for IGP-India should be seen as best estimates, based on the currently available selection of profile data held in IGP-SOTER and WISE.
- Modellers should familiarize themselves with the assumptions and taxotransfer rules used to develop the set of soil parameter estimates, before using these in their models.
- The detail and quality of primary information available within the study area may result in a variable resolution of the products presented.

- The primary and secondary data sets for IGP-India, will be useful for agro-ecological zoning, land evaluation, and modeling of carbon stocks and changes at a scale of 1:1M.

ACKNOWLEDGEMENTS

This work contributes to the GEF co-financed project *Assessment of Soil Organic Carbon Stocks and Change at National Scale* (GFL-2740-02-4381). ISRIC – World Soil Information receives funding from the Netherlands Ministry of Spatial Planning, Housing and the Environment (VROM) Project DGM-200207282.

The contributions of the scientific and technical staff from NBSS&LUP, ICAR, in compiling the primary SOTER data for IGP-India are gratefully acknowledged. Jan Huting (ISRIC) provided GIS support.

REFERENCES

Aggarwal PK, Joshi PK, Ingram JSI and Gupta RK 2004. Adapting food systems of the Indo-Gangetic plains to global environmental change: key information needs to improve policy formulation. *Environmental Science & Policy* 7, 487-498

Batjes NH 1999. Soil vulnerability mapping in Central and Eastern Europe: Issues of data acquisition, quality control and sharing. In: Naff T (editor), *Data Sharing for International Water Resource Management: Eastern Europe, Russia and the CIS*. NATO Science Series 2: Environmental Security (Vol. 61). Kluwer Academic Publishers, Dordrecht, pp 187-206

Batjes NH 2002. Revised soil parameter estimates for the soil types of the world. *Soil Use and Management* 18, 232-235

Batjes NH 2003. *A taxotransfer rule-based approach for filling gaps in measured soil data in primary SOTER databases (GEFSOC Project)*. Report 2003/03, ISRIC - World Soil Information, Wageningen

Batjes NH and Bridges EM 1994. Potential emissions of radiatively active gases from soil to atmosphere with special reference to methane: development of a global database (WISE). *Journal of Geophysical Research* 99(D8), 16479-16489

Batjes NH and Dijkshoorn JA 1999. Carbon and nitrogen stocks in the soils of the Amazon Region. *Geoderma* 89, 273-286

Batjes NH and Gicheru P 2004. *Soil data derived from SOTER for studies of carbon stocks and change in Kenya (ver. 1.0; GEFSOC Project)*. Report 2004/01, ISRIC - World Soil Information, Wageningen

Batjes NH, Rawajfih Z and Al-Adamat R 2003. *Soil data derived from SOTER for studies of carbon stocks and change in Jordan (ver. 1.0; GEFSOC Project)*. Report 2003/04, ISRIC - World Soil Information, Wageningen

Batjes NH, Bernoux M and Cerri CEP 2004. *Soil data derived from SOTER for studies of carbon stocks and change in Brazil (ver. 1.0; GEFSOC Project)*. Report 2004/03, ISRIC - World Soil Information, Wageningen

Batjes NH, Fischer G, Nachtergaele FO, Stolbovoy VS and van Velthuizen HT 1997. *Soil data derived from WISE for use in global and regional AEZ studies (ver. 1.0)*. Interim Report IR-97-025, FAO/ IIASA/ ISRIC, Laxenburg

Bhattacharyya T and Pal DK 2003. Carbon sequestration in soils of the Indo-Gangetic Plain, *Addressing resource conservation issues in rice-wheat system of south Asia, A Resource Book*. RWC-CIMMYT, New Delhi, India, pp 68-71

Bhattacharyya T, Pal DK, Mandal C and Velayutham M 2000. Organic carbon stocks in Indian soils and their geographical distribution. *Current Science* 79, 655-660

Bhattacharyya T, Mandal C, Pal DK, Gajbhiye KS and co-workers 2004a. *Compilation of available soil and other data sets for the Indo-Gangetic Plains, India (Digital data set prepared for GEFSOC Project; Nov. 2004 version)*, National Bureau for Soil Survey and Land Use Planning (NBSS&LUP), Nagpur

Bhattacharyya T, Pal DK, Chandran P, Mandal C, Ray SK, Gupta RK and Gajbhiye KS 2004b. *Managing soil carbon stocks in the Indo-gangetic Plains, India*, Rice-Wheat Consortium for the Indo-Gangetic Plains (http://www.rwc.cgiar.org/pubs/104/SoilCarbon.pdf), New Delhi, India

Bouwman AF, Boumans LJM and Batjes NH 2002. Modeling global annual N_2O and NO emissions from fertilized fields. *Global Biogeochemical Cycles* 16, 1080, doi:10.1029/2001GB001812

CEC 1985. *Soil Map of the European Communities (1:1,000,000)*. Report EUR 8982, Office for Official Publications of the European Communities, Luxembourg

Cramer W and Fischer A 1997. Data requirements for global terrestrial ecosystem modelling. In: Walker B and W Steffen (editors), *Global Change and Terrestrial Ecosystems*. Cambridge University Press, Cambridge, pp 529-565

Dijkshoorn JA 2003. *Soil and terrain database training - National Bureau of Soil Survey and Land Use Planning, Indian Council of Agricultural Research (19-30 November 2003)*. Consultancy Report (unpubl.), ISRIC - World Soil Information, Wageningen

Dwivedi BS, Shukla AK, Singh VK and Yadav RL 2003. Improving nitrogen and phosphorus use efficiencies through inclusion of forage cowpea in the rice-wheat systems in the Indo-Gangetic Plains of India. *Field Crops Research* 84, 399-418

ESRI 1996. *ArcView GIS*. Environmental Systems Research Institute, Redlands CA, 350 p

Falloon PD, Smith P, Smith JU, Szabó J, Coleman K and Marshall S 1998. Regional estimates of carbon sequestration potential: linking the Rothamsted carbon model to GIS databases. *Biology and Fertility of Soils* 27, 236-241

FAO 1988. *FAO/Unesco Soil Map of the World, Revised Legend (with corrections and updates)*. FAO World Soil Resources Report 60 (reprinted with updates as ISRIC Technical Paper 20 in 1997), ISRIC, Wageningen

Fischer G, van Velthuizen HT, Shah M and Nachtergaele FO 2002. *Global Agro-ecological Assessment for Agriculture in the 21st Century: Methodology and Results*. RR-02-02, International Institute for Applied Systems Analysis (IIASA) and Food and Agriculture Organization of the United Nations (FAO), Laxenburg

Goswami NN, Pal KK, Narayanasamy G and Bhattacharya T 2000. Soil organic matter - Management issues, *Proceedings International Conference on managing natural resources for sustainable agricultural production in the 21st century*, New Delhi, pp 87-96

Klamt E and Sombroek WG 1988. Contribution of organic matter to exchange properties of Oxisols. In: Beinroth FH, MN Camargo and H Eswaran (editors), *Classification, characterization and utilization* *of Oxisols. Proc. of the 8th International Soil Classification Workshop (Brazil, 12 to 23 May 1986)*. Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA), Soil Management Support Services (SMSS) and University of Puerto Rico (UPR), Rio de Janeiro, pp 64-70

Ladha JK, Dawe D, Pathak H, Padre AT, Yadav RL, Singh B, Singh Y, Singh Y, Singh P and Kundu AL 2003. How extensive are yield declines in long-term rice-wheat experiments in Asia? *Field Crops Research* 81, 159-180

Lal S, Deshpands SB and Sehgal J 1994. *Soil series of India*, NBSS Publication 40. National Bureau of Soil Survey and Land Use Planning, Nagpur

Lambin EF, Geist HJ and Lepers E 2003. Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources* 28, 205-2041

Mandal C, Mandal DK, Srinias CV, Seghal J and Velayutham M 1999. *Soil-climatic database for crop planning in India*, NBSS Publ. No. 53. National Bureau of Soil Survey and Land Use Planning, Nagpur, 994 p

Mantel S, van Engelen VWP, Molfino JH and Resink JW 2000. Exploring biophysical potential and suitability of wheat cultivation in Uruguay at the national level. *Soil Use and Management* 16, 270-278

Murthy RS, Hirekepur LR, Deshpande SB and Venkata Rao BV 1982. *Benchmark soils of India - Morphology, characteristics and classification for resource management*. National Bureau of Soil Survey and Land Use Planning (ICAR), Nagpur, 374 p

Nachtergaele FO, van Lynden GWJ and Batjes NH 2002. Soil and terrain databases and their applications with special reference to physical soil degradation and soil vulnerability to pollution in central and eastern Europe. In: Pagliaia M and R Jones (editors), *Advances in GeoEcology 35*. CATENA Verlag GMBH, Reiskrirchen

Nambiar KKM 1994. *Soil fertility and crop productivity under long-term fertilizer use in India*. Indian Council of Agricultural Research, New Delhi, 144 p

Nemes A, Schaap MG and Wosten JHM 2003. Functional Evaluation of Pedotransfer Functions Derived from Different Scales of Data Collection. *Soil Science Society of America Journal* 67, 1093-1102

Pal DK, Srivastava P, Durge SL and Bhattacharyya T 2003. Role of microtopography in the formation of sodic soils in the semi-arid part of the Indo-Gangetic Plains, India. *CATENA* 51, 3-31

Pal DK, Dasog GS, Vadivelu S, Ahuja RL and Bhattacharyya T 2000. Secondary calcium carbonate in soils of arid and semi-arid regions of India. In: Lal R, JM Kimble, H Eswaran and BA Stewart (editors), *Global climate change and tropical ecosystems*. Lewis Publishers, Boca Raton (FL), pp 149-185

Paustian K, Levine E, Post WM and Ryzhova IM 1997. The use of models to integrate information and understanding of soil C at the regional scale. *Geoderma* 79, 227-260

Scholes RJ, Skole D and Ingram JS 1995. *A global database of soil properties: proposal for implementation*. IGBP-DIS Working Paper 10, International Geosphere Biosphere Program, Data & Information System, Paris

Shrikantia S 1999. Map restrictions conflict with today's needs in India: modern maps but outrageous regulation. *Geomatics Info Magazine* 13, 68-71

Singh RL 1971. *India - A regional geography*. National Geographic Society of India, Bhargava Bhushan Press, Varnassi (India), 992 p

Soil Survey Staff 1975. *Soil Taxonomy: A basic system of soil classification for making and interpreting soil surveys*. Soil Conservation Service, USDA Agriculture Handbook 436, USDA, Washington

Soil Survey Staff 1983. *Soil Survey Manual (rev. ed.)*. United States Agriculture Handbook 18, USDA, Washington

Soil Survey Staff 1998. *Keys to Soil Taxonomy (8th. ed.)*. United States Department of Agriculture, Natural Resources Conservation Service, Blacksburg (Virginia), 326 p

Tempel P 2002. *SOTER - Global and National Soil and Terrain Digital Databases: Database structure (Ver 3.1)*, ISRIC - World Soil Information, Wageningen

Tomasella J and Hodnett MG 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Science* 163, 190-202

van den Berg M, Klamt E, van Reeuwijk LP and Sombroek WG 1997. Pedotransfer functions for the estimation of moisture retention of Ferralsols and related soils. *Geoderma* 78, 161-180

van Engelen VWP and Wen TT 1995. *Global and National Soils and Terrain Digital Databases (SOTER): Procedures Manual (rev. ed.)*. (Published also as FAO World Soil Resources Report No. 74), UNEP, IUSS, ISRIC and FAO, Wageningen

van Reeuwijk LP 2002. *Procedures for soil analysis (6th ed.)*. Technical Paper 9, ISRIC, Wageningen

Velayutham M, Pal DK and Bhattacharya T 2000. Organic carbon stocks in soils of India. In: Lal R, JM Kimble and BA Stewart (editors), *Global climate change and tropical ecosystems*. CRC Press, Boca Raton, pp 71-95

Velayutham M, Mandal DK, Mandal C and Seghal J 1999. *Agro-ecological subregions of India for planning and development*, NBBS Publ. No. 53. National Bureau of Soil Survey and Land Use Planning, Nagpur, 372 p

Velayutham M, Pal DK, Bhattacharya T and Srivastava P 2002. Soils of the Indo-Gangetic Plains India - The historical perspective. In: Abrol Y, S Sangwan and MK Tiwari (editors), *Land use - Historical perspective: Focus on Indo-Gangetic Plains*. Allied Publisher PVT. Ltd., New Delhi, pp 61-70

Webster R and Oliver MA 1990. *Statistical methods in soil and land resource survey*. Spatial Information Systems. Oxford University press, Oxford, 316 p

Wösten JHM, Lilly A, Nemes A and Le Bas C 1998. *Using existing soil data to derive hydraulic parameters for simulation models in environmental studies and in land use planning*. Report 156, DLO-Staring Centre, Wageningen

WRB 1998. *World Reference Base for Soil Resources*. World Soil Resources Report 84, ISSS, ISRIC, and FAO, Rome

APPENDICES

Appendix 1: SOTER unit composition file

This summary table gives the full composition of each SOTER unit in terms of its main soil units (FAO, 1988), their relative extent, and the identifier for the corresponding representative profile. It contains information aggregated from a number of primary SOTER tables, *viz.* SoilComponent and Profile. It can be easily linked to the SOTER geographical data in a GIS through the unique SOTER unit code – NEWSUID, a combination of the fields for ISO and SUID – and linked to the table holding the soil parameter estimates through the unique profile identifier (PRID, see Appendix 2 and Figure 8).

Structure of table SOTERunitComposition

Name	Type	Size	Description
ISOC	Text	2	ISO-3166 country code (1994)
SUID	Integer	2	The identification code of a SOTER unit on the map and in the database
NEWSUID	Text	10	Globally unique SOTER code, comprising fields ISOC plus SUID
SOIL1	Text	3	Characterization of the first (main) according to the Revised Legend (FAO, 1988)
PROP1	Integer	2	Proportion, as a percentage, that the main soil occupies Within the SOTER unit
PRID1	Text	15	Unique code for the corresponding representative soil profile (as selected by the national soil experts)
SOIL2	Text	3	As above but for the next soil component
PROP2	Integer	2	As above
PRID2	Text	15	As above
SOIL3	Text	3	As above but for the next soil component
PROP3	Integer	2	As above
PRID3	Text	15	As above
SOIL4	Text	3	As above but for the next soil component
PROP4	Integer	2	As above
PRID4	Text	15	As above
SOIL5	Text	3	As above but for the next soil component
PROP5	Integer	2	As above
PRID5	Text	15	As above
SOIL6	Text	3	As above but for the next soil component
PROP6	Integer	2	As above
PRID6	Text	15	As above
SOIL7	Text	3	As above but for the next soil component

(*cont.*)

PROP7	Integer	2	As above
PRID7	Text	15	As above
SOIL8	Text	3	As above but for the next soil component
PROP8	Integer	2	As above
PRID8	Text	15	As above
SOIL9	Text	3	As above but for the next soil component
PROP9	Integer	2	As above
PRID9	Text	15	As above
SOIL10	Text	3	As above but for the next soil component
PROP10	Integer	2	As above
PRID10	Text	15	As above

Note: Generally, not all 10 available fields for $SOIL_i$ will be filled in SOTER.

Appendix 2: Taxotransfer rule-based soil parameter estimates

This table lists soil parameters estimates for all representative profiles considered in a given SOTER database. This information can be linked to the geographical component of the SOTER database – in a GIS – through the unique profile code (PRID, see Appendix 1).

Structure of table SOTERparameterEstimates

Name	Type	Size	Description
CLAF	Text	3	Revised Legend FAO (1988) code
PRID	Text	15	profile ID (as documented in table SOTERunitComposition)
Drain	Text	2	FAO soil drainage class
Layer	Text	8	code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm)
TopDep	Integer	4	depth of top of layer (cm)
BotDep	Integer	4	depth of bottom of (cm)
CFRAG	Integer	2	coarse fragments (> 2mm)
SDTO	Integer	2	sand (mass %)
STPC	Integer	2	silt (mass %)
CLPC	Integer	2	clay (mass %)
PSCL	Text	1	FAO texture class (see note at end of this report for codes)
BULK	Single	4	bulk density (kg dm^{-3})
TAWC	Integer	2	available water capacity (cm m^{-1}, -33 to -1500 kPa conform to USDA standards)
CECS	Single	4	cation exchange capacity ($cmol_c$ kg^{-1}) for fine earth fraction
BSAT	Integer	2	base saturation as percentage of CECsoil

(cont.)

CECc	Single	4	CECclay, corrected for contribution of organic matter ($cmol_c\ kg^{-1}$)
PHAQ	Single	4	pH measured in water
TCEQ	Single	4	total carbonate equivalent ($g\ kg^{-1}$)
GYPS	Single	4	gypsum content ($g\ kg^{-1}$)
ELCO	Single	4	electrical conductivity ($dS\ m^{-1}$)
TOTC	Single	4	organic carbon content ($g\ kg^{-1}$)
TOTN	Single	4	total nitrogen ($g\ kg^{-1}$)
ECEC	Single	4	effective CEC ($cmol_c\ kg^{-1}$)

Note: These are depth-weighted values. In view of the TTR-rules applied and depth weighting, the parameters listed for TOTC and TOTN should not be used to compute C/N ratios!

The above table should be consulted in conjunction with table SOTERflagTTRrules which documents the taxotransfer rules that have been applied (see Appendix 3).

Appendix 3: Flagging taxotransfer rules

The type of taxotransfer that has been used when creating the table SOTERparameterEstimates (Appendix 2) is documented in table SOTERflagTTRrules. Further details on coding conventions may be found in the text (Section 3.3).

Structure of table SOTERflagTTRrules

Name	Type	Size	Description
CLAF	Text	3	Revised Legend (FAO, 1988) code
PRID	Text	15	Unique identifier for representative profile
Newtopdep	Integer	2	Depth of top of layer (cm)
Newbotdep	Integer	2	Depth of bottom of layer (cm)
TTRsub	Text	50	Codes showing the type of taxotransfer rule used (based on data for soil *units*; see text)
TTRmain	Text	50	Codes showing the type of taxotransfer rule used (based on data for *major units*; see text)
TTRfinal	Text	25	Additional flags (based on expert knowledge)

Note: The exchangeable aluminum percentage (ALSA) has been set at zero when pH_{water} is higher than 5.5. Similarly, the electrical conductivity (ELCO), content of gypsum (GYPS) and content of carbonates (TCEQ) have been set at zero when pH_{water} is less than 6.5. Finally, the CEC of the clay fraction (CEC_{clay}) has always been re-calculated from the depth-weighted measured and TTR-derived data for

CEC_{soil} and content of organic carbon, assuming a mean contribution of 350 $cmol_c$ kg^{-1} OC (Klamt and Sombroek 1988). When applicable, this has been flagged in the field TTRfinal; the coding conventions are given in Figure 6.

Appendix 4: SOTER summary file

Interpretations of a SOTER database, in combination with the current set of soil parameter estimates requires a good knowledge of relational database handling systems and a sound understanding of the SOTER database structure. This may be an obstacle to end-users with limited programming expertise. Therefore, to facilitate access to the data and its ultimate linkage to GIS, a SOTER summary file has been created. The structure of the corresponding table is shown below.

Information on landform, lithology and slope has been derived from Bhattacharrya *et al*. (2004a).

Structure of table SOTERsummaryFile

Name	Type	Size	Description
ISOC	Text	2	ISO-3166 country code (1994)
SUID	Integer	2	The identification code of a SOTER unit on the map and in the database
NEWSUID	Text	10	Globally unique SOTER code, comprising fields ISOC Plus SUID
TCID	Integer	1	Number of terrain component in given SOTER unit
SCID	Integer	1	Number of soil component within given terrain component and SOTER unit
PROP	Integer	3	Relative proportion of above in given SOTER unit
CLAF	Text	3	Revised Legend FAO (1988) code
PRID	Text	15	Profile ID (as documented in table SOTER-unitComposition)
Drain	Text	2	FAO soil drainage class
Layer	Text	8	Code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm)
TopDep	Integer	4	Upper depth of layer (cm)
BotDep	Integer	4	Lower dept of layer (cm)
CFRAG	Integer	2	Coarse fragments (> 2mm)
SDTO	Integer	2	Sand (mass %)
STPC	Integer	2	Silt (mass %)
CLPC	Integer	2	Clay (mass %)

(cont.)

PSCL	Text	1	FAO texture class (see Figure 9)
BULK	Single	4	Bulk density (kg dm^{-3})
TAWC	Integer	2	Available water capacity (cm m^{-1}, -33 to -1500 kPa, USDA standards)
CECS	Single	4	Cation exchange capacity ($cmol_c$ kg^{-1}) of fine earth fraction
BSAT	Integer	2	Base saturation as percentage of CECsoil
CECc	Single	4	CECclay, corrected for contribution of organic matter ($cmol_c$ kg^{-1})
PHAQ	Single	4	pH measured in water
TCEQ	Single	4	Total carbonate equivalent (g kg^{-1})
GYPS	Single	4	Gypsum content (g kg^{-1})
ELCO	Single	4	Electrical conductivity (dS m^{-1})
TOTC	Single	4	Organic carbon content (g kg^{-1})
TOTN	Single	4	Total nitrogen (g kg^{-1})
ECEC	Single	4	Effective CEC ($cmol_c$ kg^{-1})

Notes:
1) These are depth-weighted values, per 20 cm layer.
2) Terrain Components, and their constituent Soil Components, within a given SOTER unit are numbered starting with the spatially dominant one (see Figure 7). The sum of the relative proportions of all Soil Components within a SOTER unit is always 100 per cent.
3) A condensed file showing only soil parameter estimates for the main Terrain Component (TCID= 1) and Soil Component (SCID =1) for the upper layer (D1) is attached as table SoterSummaryFile_T1S1D1 (see Figure 8). This type of table can be created directly in the GIS, in the table mode, using the SQL-connect option.

Appendix 5: Contents of GIS-folder

The SOTER-GIS shapefiles for IGP-India and soil parameter estimates are provided in one single zip file called: SOTWIS_IGP-India_ver1.zip (about 7.5 Mb zipped and 11 Mb de-compressed).

By default, the compressed file should be unzipped to folder *X:\SOTWIS_IGP-India_ver1.0*. This folder contains:

1) The project's apr-file, called *sotwis_IGP-India_01.apr*. This file can best be accessed from within ArcView.
2) The shape, legend and documentation files for IGP-India, in four subfolders.
3) The access database containing the soil parameter estimates (SOTWIS_IGP-India_1.mdb; see Appendices 1 to 4).

The first time the project is opened on a new system, the new folder information will be automatically updated in the project-file.

The current project file only shows a limited number of selections for the upper soil layer (D1= 0 to 20 cm or less for shallow soils) of the dominant soil of a SOTER unit. These include: content of organic carbon; content of inorganic carbon; bulk density; content of clay; content of coarse fragments, and soil drainage class.

Should any other selections be needed, the underlying MS Access® database can be easily queried via the SQL-connect option of ArcView®.

If grid-based soil layers are required, these can be generated using the convert-to-grid module of the spatial analyst extension to ArcView® (ESRI 1996). Gridding should be based on the NEWSUID field to permit subsequent linkage with the various attribute tables discussed in this report.

The project file was developed for a 17 inch screen.

Appendix 6: Limits for soil textural classes

The textural classes (PSCL, see Appendix 2 and 4) used in this study follow the criteria of FAO (1988) and CEC (1985). The following abbreviations are used: C–coarse, M–medium, Z–medium fine, F–fine and V–very fine. The symbol # is used for undifferentiated (i.e. C + M + F + Z + V). The class limits are shown in Appendix 6.

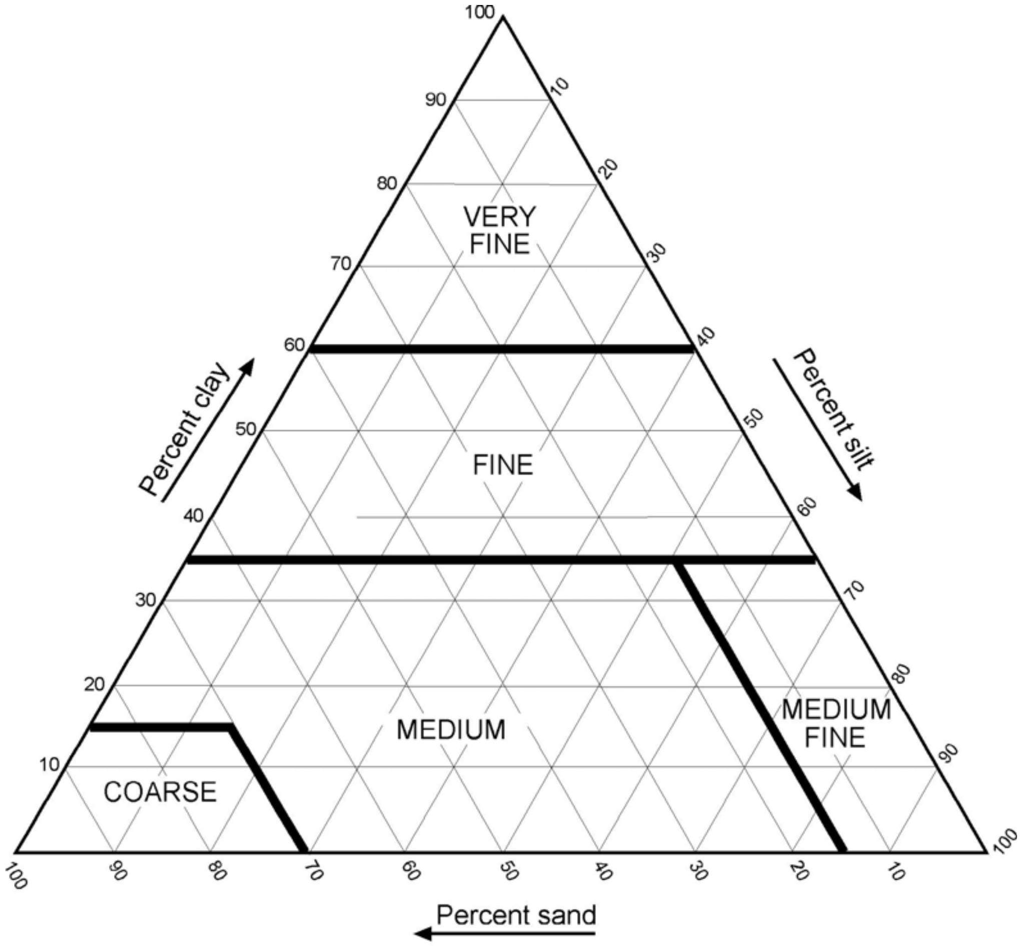

Figure 9. Soil texture classes

————————————