

# Modelling the impact of increasing soil sealing on runoff coefficients at regional scale: a hydrogeological approach

Fabrizio Ungaro<sup>1,2\*</sup>, Costanza Calzolari<sup>1</sup>, Alberto Pistocchi<sup>3</sup>, Francesco Malucelli<sup>4</sup>

<sup>1</sup> CNR IBIMET - Via Madonna del Piano 10. I-50019 Sesto F.no (FI), Italy.

<sup>2</sup> Leibniz Centre for Agricultural Landscape Research, Institute of Socio-Economics, Müncheberg, Germany.

<sup>3</sup> GECOSistema srl - Botengasse, 27 I-39050 Jenesien (BZ), Regione Emilia Romagna - Autorità dei bacini regionali romagnoli, via Biondini, 1 I-47100 Forlì (FC), Italy.

<sup>4</sup> Regione Emilia Romagna - Servizio Geologico, Sismico e dei Suoli, Viale A. Moro, Bologna (BO), Italy.

\* Corresponding author. Tel.: +39 055 5226557. Fax: +39 055 5226550. E-mail: fabrizio.ungaro@cnr.it

**Abstract:** Soil sealing is the permanent covering of the land surface by buildings, infrastructures or any impermeable artificial material. Beside the loss of fertile soils with a direct impact on food security, soil sealing modifies the hydrological cycle. This can cause an increased flooding risk, due to urban development in potential risk areas and to the increased volumes of runoff. This work estimates the increase of runoff due to sealing following urbanization and land take in the plain of Emilia Romagna (Italy), using the Green and Ampt infiltration model for two rainfall return periods (20 and 200 years) in two different years, 1976 and 2008. To this goal a hydrogeological approach was adopted in order to characterize soil hydraulic properties via locally calibrated pedotransfer functions (PTF). PTF inputs were estimated via sequential Gaussian simulations coupled with a simple kriging with varying local means, taking into account soil type and dominant land use. Results show that in the study area an average increment of 8.4% in sealed areas due to urbanization and sprawl induces an average increment in surface runoff equal to 3.5 and 2.7% respectively for 20 and 200-years return periods, with a maximum > 20% for highly sealed coast areas.

**Keywords:** Land use change; Soil functional groups; Urbanization; Runoff; Emilia-Romagna.

## INTRODUCTION

Soil sealing is here defined as the permanent covering of the land surface by buildings, infrastructures or any impermeable artificial material. It has been identified as one of the major threats to soil in the Soil Thematic Strategy (CEC, 2006) of the European Commission, both in terms of the permanent loss of soil as a resource and for its important impacts on soil functionality. Beside the loss of fertile soils (Tóth, 2012), soil sealing strongly modifies the hydrological cycle, reducing infiltration, evapotranspiration and groundwater recharge, and increasing runoff (Poelmans et al., 2010). The reduction of infiltration capacity, which can be partly mitigated by a correct urban planning and by the adoption of semi permeable materials (Scholz and Grabowiecki, 2006), is the main cause of the increased runoff, which in turn increases flooding risk. Moreover, the flooding risk may be enhanced by the increased vulnerability of the urbanized and settled land. Different modelling approaches and scenario settings can be adopted to quantify the effect of urbanization on surface runoff generation. Using a physically based groundwater model coupled with a land-use change allocation model to assess the effects of different land use scenarios in a Belgian catchment with 11% of built-up areas, Dams et al. (2008) estimated a runoff increase of 2.8% following a projected increase of 14% in sealed surfaces in 20 years, while the projections to 2031 of the actual rate of sprawling of the Canadian city of Calgary (representing 5.9% of the case study area) resulted in an increase in runoff equal to 7.3% due to a 65% increase of the built-up areas in the watershed (Wijesekara et al., 2012). Depending on rainfall duration and return period, a 50% increase of built-up areas (from 7.4 to 11.1% of the case study area) in a 115 km<sup>2</sup> catchment in southwestern Germany resulted in a runoff increase between 0.8 and

3.7% on a volume basis as predicted by coupling a land-use change model with a physically based distributed hydrological model (Niehoff et al., 2002). Using a physically based model under different land-use change scenarios projected to 2040, Franczyk and Chang (2009) projected a 5 to 9% expansion in urban land use throughout a basin in the Portland metropolitan area (nearly 60% of the case study area) which resulted in an increase in runoff between 2.3 and 2.5%, with the highest change observed for low-density, unchecked urban growth.

The plain areas of Emilia Romagna (northern Italy) have been subject to significant expansion of urban and settled areas over the last decades. About 14.7% of the regional plain area territory was estimated to be covered by urban and industrial areas (Regione Emilia Romagna, 2011), as compared to the Italian average of 7.3% in 2009 (Istat, 2012, quoting LUCAS, Eurostat data). Apart from the lack of a common procedure for estimating the sealed areas, the percentage of soils lost at EU level every year, estimated at ~1,000 km<sup>2</sup> (Prokop et al., 2011), causes a serious concern.

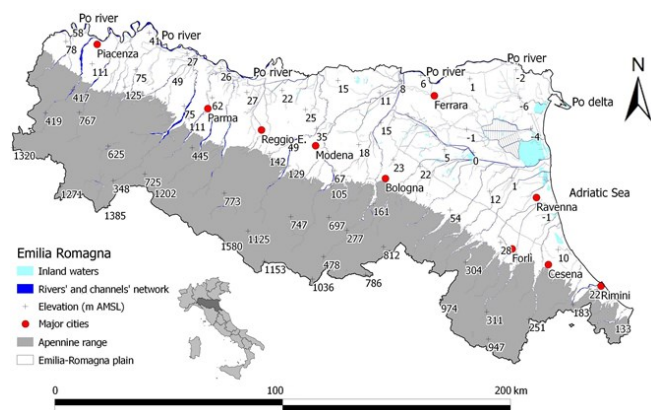
The goal of this work is to present a hydrogeological approach to predict spatially exhaustive runoff, in order to assess quantitatively the impact of soil sealing on runoff in the plain of Emilia Romagna, Italy. A novel methodology is here presented which couples an event based infiltration model under different land use scenarios with a set of locally calibrated pedotransfer functions. PTF's inputs are derived via a Scorpan kriging which combines the trend component of soil properties, as derived from the 1:50,000 soil map, with the geostatistical simulations of the stochastic, locally varying but spatially correlated component. The methodological steps were as follows: i) continuous mapping of basic soil properties; ii) estimation of continuous parameter fields (retention properties, saturated conductivity) using a set of locally calibrated PTFs; iii) calculation of soil

infiltration and runoff using the Green and Ampt infiltration model for two rainfall return periods (RP; 20 and 200 years) and in two different years, 1976 and 2008, in order to assess the effect of land use changes. The following conditions were assumed: i) an antecedent soil moisture content corresponding to the average between field capacity and wilting point; ii) a constant soil depth of 1 m; iii) the matric potential in soils equal to the air entry potential; iv) a runoff coefficient equal to 1 in sealed areas.

**MATERIALS AND METHODS**

**Study area: description**

Emilia Romagna (lat 43°50'N – 45°00'N; long 9°20'E – 12°40'E Greenwich, approx.) is situated in Northern Italy and has a total area of 22.124 km<sup>2</sup>. The main agricultural area, covering slightly more than half of the region (~12.002 km<sup>2</sup>), is the continuous plain stretching south of the Po river and delimited by the Apennines range in the south and by the Adriatic sea in the east (Fig. 1).



**Fig. 1.** Study area: simplified sketch of Emilia Romagna region, showing the Apennine range, the Emilia-Romagna Plain, the major cities, and the rivers' and channels' network. Elevations (+) are in m AMSL.

The climate is temperate-sub oceanic, with a maximum and minimum average annual temperature of 19.3 and 8.2°C, respectively, and a mean annual precipitation ranging from 520 to 820 mm. Elevation ranges between -3 and 150 m a.m.s.l. The soils of the area, mainly on quaternary alluvial deposits, are characterized by a high degree of heterogeneity, ranging from coarse textured Arenosols (Calcaric Arenosols (Gleyic), FAO-ISRIC-ISSS, 1998) to fine textured Vertisols (Calci Hyposalic Vertisols, FAO-ISRIC-ISSS, 1998), leached paleosols along the Apennine border (Chromic Luvisols, FAO-ISRIC-ISSS, 1998), and soils characterized by the presence of organic horizons in the reclaimed areas of the Po delta (Humi Thionic Fluvisols Thapthohistic, FAO-ISRIC-ISSS, 1998). The soils of the Emilia Romagna plain sustain intensive agricultural productions that range from orchards, vineyards and industrial crops, to cereals, grassland and fodder crops. In the period 1954–2003 the agricultural land uses lost 1,181 km<sup>2</sup> mainly due to urbanisation (Di Gennaro et al., 2010)

**Data sets**

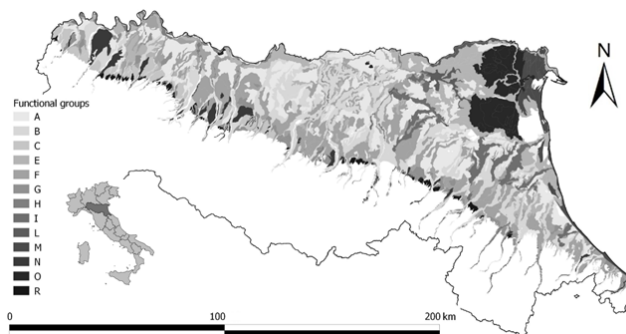
**Soil and land use data**

Soil data are routinely collected and analysed by the Regional Soil Survey and by Agricultural Extension Services. At pre-

sent about 3,302 soil profiles (17,652 soil horizons) over 10,734 km<sup>2</sup> of cultivated land are identified by a complete set of physical and chemical parameters. For each site the textural fractions (%), USDA limits) and soil organic matter content (%), modified Walkley-Black method; Nelson and Sommers, 1982) are available for a reference depth of 100 cm. These sites are linked to a regional catalogue of 237 soil typological units (STUs) mapped in the available 1:50,000 soil map of the region (Regione Emilia Romagna, 2006). For technical applications, the regional Soil Survey has aggregated the 237 STUs in a limited number of “functional groups” defined according to the top-soil textural family, drainage class, slope, presence of horizons with organic carbon >2.5%, and flooding occurrence (Table 1; Guermandi and Tarocco, 2007). The distribution of the functional groups in the study area is shown in Fig. 2.

**Table 1.** Soil Functional Groups and their characteristics. C, clay; CL, clay loam; SC, sandy clay; SiL, silt loam; Si, silt; SiCL, siltyclay loam; SCL, sandy clay loam; SL, sandy loam; L, loam; S, sand; LS, loamy-sand (USDA Textural classes, USDA, 1993).

Functional Groups	Soil characteristics for Functional Groups
A	Ap texture fine (C, CL, SC)
B	Ap texture medium (SiL, Si), well drained
C	Ap texture medium (SiL, Si), moderately well drained
E	Ap texture medium-fine (SiCL, CL, SCL with clay <35%), moderately well drained to poorly drained
F	Ap texture medium-fine (SiCL, CL, SCL with clay <35%), well drained
G	Ap texture moderately fine (SiCL, CL, with clay > 35%) to fine (C, CL, SC), moderately well drained
H	Frequent flooding
I	Ap texture moderately coarse (SL, L) to coarse (S, LS), well drained
L	Ap texture moderately coarse (SL, L), moderately well drained
M	Ap texture coarse (S, LS)
N	Soils with coarse fragments (>2 mm) ≥ 35%
O	High SOC (>2.5% in Ap/Op or at any depth within 100 cm)
R	Slope > 6%



**Fig. 2.** Study area: simplified soil map (scale 1:50,000) of the Emilia Romagna plain; the characteristics of the soil functional groups are summarised in Table 1.

In Table 2 the land use trends are summarised for the period 1861–2008 (Regione Emilia Romagna, 2011). From the data it is clear that the expansion of artificially sealed areas represents the major land use change occurred in the study area (13.5%), with the greatest increase between 1976 and 2003 (ca. 7%). In the two reference periods considered in this study, i.e. 1976 and 2008, the increase in artificially sealed areas results greater than 8%.

**Extreme rainfall events**

Runoff coefficients represent an indicator of soil or catchment response to rainfall; they are well known to vary following soil moisture conditions antecedent to rainfall, and need to

be defined with reference to a specific rainfall event and initial soil state.

**Table 2.** Land use shares in the Emilia-Romagna plain between 1861 and 2008.

Land use	1861	1954	1976	2003	2008
Urban and Industrial	1.5%	2.4%	6.3%	13.1%	14.65%
Cropland	73.1%	91.3%	89.3%	81.2%	77.36%
Woodland	1.5%	0.8%	0.7%	0.7%	1.06%
Grassland, bush and natural areas	12.6%	0.9%	1.5%	0.6%	0.78%
Wetlands and river	11.3%	4.6%	2.2%	4.4%	6.15%

In assessing the effects of soil sealing on runoff generation, we refer to extreme events of precipitation defined as design values by the river basin authorities in the region. These design values are given in the form of precipitation depth-duration-frequency curves (DDFCs), i.e. relationships between rainfall amount and rainfall duration for a given return period in the following form:

$$h_{t,T} = a(T) t^{n(T)}, \quad (1)$$

where  $h$  is the cumulative rainfall (mm) and  $t$  is its duration (hr),  $T$  is the return period and parameters  $a(T)$  and  $n(T)$  of the rainfall probability curves are given in the form of maps as a function of the return period, that was selected here as either 20 or 200 years. Parameters  $a(T)$  and  $n(T)$  were obtained from previous hydrological studies at regional scale (Autorità di Bacino del Fiume Po, 2001; Autorità dei Bacini Regionali Romagnoli - Regione Emilia Romagna, 2003; Pistocchi, 2001;) and are shown in Fig. 3.

A rainfall event is defined with reference to its duration. In this study, we are interested in extreme events that may be critical at the local scale, such as short and intense rainfall events over small catchments. For this type of events, a typical duration is one hour or less. Therefore, we refer to a 1 hour rainfall duration throughout of the paper.

## INFILTRATION MODEL

In order to calculate the runoff coefficient,  $\phi$ , for a given rainfall event, we used the popular Green-Ampt infiltration model (Green and Ampt, 1911) as presented in Mein and Larson (1973):

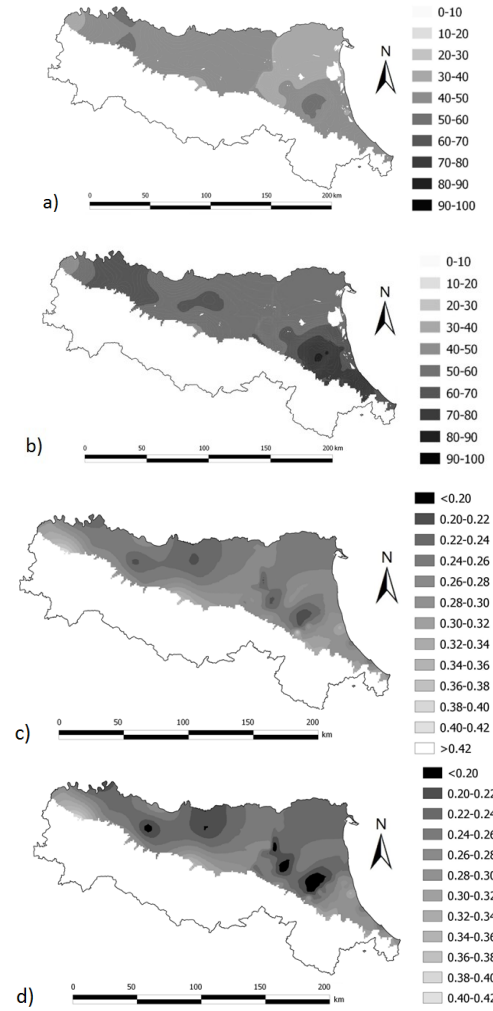
$$I(t) = K_{sat} \left[ 1 + \frac{\theta_s - \theta_i}{f(t)} \psi \right], \quad (2)$$

where  $I(t)$ , is the infiltration rate (mm h<sup>-1</sup>) at time  $t$  from the beginning of the rainfall event,  $K_{sat}$  is the saturated hydraulic conductivity (mm h<sup>-1</sup>),  $\theta_s$  is the water content at saturation (cm<sup>3</sup> cm<sup>-3</sup>),  $\theta_i$  is the initial water content (cm<sup>3</sup> cm<sup>-3</sup>),  $\psi$  is the matric suction potential and  $f(t)$  is the cumulated infiltration at time  $t$  from the beginning of the rainfall event (mm). The runoff coefficient  $\phi$  for a rainfall duration  $t$ ,  $\phi(t)$ , is computed as:

$$\phi(t) = 1 - I(t)/P t, \quad (3)$$

where  $P$  is the rainfall amount (mm), and  $t = 1$  hr. In this study, the antecedent soil moisture content was set to the average

between field capacity and wilting point, and the matric potential at the wetting front was assumed to be equal to the air entry potential. A constant soil depth of 1 m was also assumed. A sensitivity analysis of the model to soil depth and initial soil moisture is presented in Appendix A, together with a discussion of the assumed values of depth and soil moisture.



**Fig. 3.** Rainfall parameters  $a$  (a, b) and  $n$  (c, d); return periods 20 years (a, c) and 200 years (b, d); rainfall duration 1 hr.

## PEDOTRANSFER FUNCTIONS AND SPATIALISATION OF THE HYDROLOGICAL PROPERTIES

The hydraulic parameters required by the infiltration model, total porosity, volumetric soil water contents at field capacity and at wilting point, and air entry potential (cm) were calculated with a set of locally calibrated pedotransfer functions (PTFs, Ungaro et al., 2005). As concerns the hydraulic conductivity at saturation ( $K_{sat}$ ), a new PTF was calibrated on a new data set ( $N = 268$ ). This was then validated on an independent data set (UNSODA, Leij et al., 1996;  $N = 212$ ). The  $K_{sat}$  was determined in laboratory on undisturbed soil cores with the constant head permeameter method (Klute and Dirksen, 1986). The descriptive statistics of the calibration and validation data sets are shown in Table 3, along with the correlations among variables. The signs of the correlation coefficients are nearly the same in both data sets, with Log  $K_{sat}$  positively and significantly (at  $p < 0.05$ ) correlated with sand content ( $r = 0.40$  and  $0.31$

respectively for the calibration and the validation data set). Of opposite sign and significant for both data sets, are the correlations with clay (−0.39 and −0.20 for the calibration and the validation data set respectively) and silt (−0.26 and −0.28 for the calibration and the validation data set respectively) contents and with bulk density (−0.31 and −0.17 for the calibration and the validation data set respectively). Although not significant, for the calibration data set the correlation between Log  $K_{sat}$  and organic carbon (OC) is negative ( $r = -0.04$ ), while it is positive and significant, in the validation data set ( $r = 0.21$ ). This difference is likely due to the fact that in the calibration data set OC is not significantly correlated with any textural fractions while in the validation data set it is significantly correlated with sand ( $r = -0.29$ ) and with clay ( $r = 0.45$ ) contents. Furthermore, the case study area is characterised by the widespread occurrence of swelling soils (Vertisols and vertic intergrades), with higher OC contents but low saturated conductivity, and by a local occurrence of recently reclaimed organic soils with appreciable levels of soil salinity. These findings are in agreement with those observed by other authors, i.e. that organic matter content and  $K_{sat}$  are not positively correlated in all soil types (Nemes et al. 2005).

The PTFs were developed using a group method of data handling (GMDH) technique (Ivakhnenko, 1968), also known as polynomial neural networks, using the commercial software Model quest (AbTech, 1996). A complete overview of the methods is given by Müller and Lemke (2000) and a description of the software can be found in Pachepsky and Rawls (1999), Pachepsky et al. (1998) and Rawls et al. (2003). Along with artificial neural networks (ANNs) and genetic programming (GP), GMDH is a data driven technique suitable for automatic generation of models linking input and output variables. ANNs have been extensively applied in hydrology and soil science (Minasny et al., 1999; Motaghian and Mohammadi, 2011; Schaap et al. 2001;) and a comprehensive review is given by Rogiers et al. (2012). Differently from the other machine learning approach, GMDH is a self-organising inductive modelling technique based on explicit polynomial regression algorithms. The method selects automatically the essential input variables and builds a hierarchical, polynomial regression given a degree of complexity specified by the user. Differently from NNs techniques, which are essentially deductive in their nature, the GMDH technique, as GP (Parasuraman and Elshorbagy, 2007), does not require an arbitrary, a priori structure of the network connecting inputs and outputs: the structure of the model and the dependence of outputs on the most significant inputs are found generating a network structure based on the characteristics of the data set during the estimation process itself; the GMDH networks can then be considered as universal structure identifiers. An example of their application to the development of PTF for estimating  $K_{sat}$  is given by Nemes et al. (2005), and a comparison of their performance as compared to ANNs in predicting soil water properties is given by Ungaro et al. (2005). Although GMDH had been proved as an alternative technique to the most widely used ANNs, it is not the focus of this study to compare their performance, rather to further explore the efficacy of GMDH in developing PTFs to estimate saturated hydraulic conductivity. The GMDH algorithm developed in this work is reported in the Appendix B. To evaluate the performance of the new PTF, the observed and the estimated hydraulic conductivity were compared calculating the following error indexes (Tietje and Hennings, 1996): the geometric mean error ratio (GMER) and the geometric standard deviation of the error ratio (GSDER), defined as follows:

$$GMER = \exp\left(\frac{1}{n} \sum_{i=1}^n \ln(\varepsilon_i)\right), \quad (4)$$

$$GSDER = \exp\left[\left(\frac{1}{n-1} \sum_{i=1}^n [\ln(\varepsilon_i) - \ln(GMER)]^2\right)^{1/2}\right], \quad (5)$$

where  $\varepsilon$  is the error ratio of measured ( $K_{sat m}$ ), vs. predicted ( $K_{sat p}$ ) values, calculated as follows:

$$\varepsilon = \frac{K_{sat p}}{K_{sat m}}. \quad (6)$$

When values are generally overestimated, GMER is  $>1$ , while the GMER  $<1$  indicates a general underestimation, with GMER equal to 1 corresponding to an exact matching between measures and predictions. A value of GSDER equal to 1 indicates a perfect matching and it increases with the distance of the estimates from the measurements. An optimal model would then have a GMER close to 1 and a small GSDER. These indexes were computed also for a number of widely used PTFs from literature.

Uncertainty associated to PTF input data (McBratney et al., 2002) was calculated using Monte Carlo simulation, i.e. by sampling repeatedly the observed distribution of input data followed by the evaluation of PTF results. The approach followed here is the same followed by Ungaro et al. (2005), and Monte Carlo simulations were implemented following the multivariate normal Johnson transformation approach (Johnson, 1987) as described in Carsel and Parrish (1988) and using the MS Excel add-in developed by Barreto and Howland (2005a). The simulations return 10,000 PTF input variable vectors; those within calibration range were then inserted into the PTF providing the mean and the standard deviation of the saturated hydraulic conductivity for the different textural classes.

The uncertainty due to the PTF modelling (McBratney et al., 2002) was assessed with the non-parametric bootstrap approach (Efron and Tibshirani, 1993) using the MS Excel add-in developed by Barreto and Howland (2005b). Bootstrap sampling was performed to create a set of 10,000 PTFs outputs and model residuals; 1,000 bootstrap replications were generated for each subset of the original data set defined in terms of USDA textural class (USDA, 1993). New  $K_{sat}$  values were then calculated adding the model residuals to the PTF estimates in order to assess the confidence intervals for each textural class. Unlike standard asymptotic confidence intervals, confidence intervals computed from bootstrap estimates are generally asymmetric around the estimated mean. For this reason, 95% confidence intervals are assessed calculating the 2.5 and the 97.5 percentiles of the bootstrap estimates distribution (Li et al., 2010).

The basic soil attributes used as inputs of the PTFs, (i.e. sand, silt and clay %, organic carbon content and bulk density) were estimated via sequential Gaussian simulations (median values,  $N = 1,000$ ) at the nodes of a  $1 \times 1$  km regular grid over the entire area, using a Scorpan kriging approach (McBratney et al., 2003). We used the approach, applied by Ungaro et al. (2010) to topsoil, considering 100 cm depth which combines the trend component of soil properties as derived from the 1:50,000 soil map with geostatistical modelling of the stochastic, locally varying but spatially correlated component. The trend component is described in terms of varying local means, calculated from the available soil profile ( $N = 3,302$ ) taking into account soil functional groups and dominant land use.

**Table 3.** Descriptive statistics for the PTF calibration (N = 268) and validation data set (N = 212).  $K_{sat}$ , hydraulic saturated conductivity; OC, organic carbon. Correlations in bold are significant at  $p < 0.05$ .

	Sand %	Silt %	Clay %	C org. %	Bulk density $Mg\ m^{-3}$	$K_{sat}$ $mm\ h^{-1}$	Log $K_{sat}$
<b>Calibration data set (N = 268)</b>							
Mean	36.59	40.51	22.9	0.76	1.35	42.33	0.95
Standard deviation	23.4	14.57	13.81	1.84	0.21	82.05	1.1
Minimum	1	4.38	1.66	0.04	0.24	0.0001	-4
Median	35.47	41.28	20.48	0.39	1.37	13.46	1.13
Maximum	91.77	79.5	70.1	20.52	1.77	838.53	2.92
<b>Correlations</b>							
Sand %	1						
Silt %	-0.82	1					
Clay %	-0.80	0.30	1				
OC %	-0.05	-0.01	0.08	1			
Bulk density $Mg\ m^{-3}$	-0.01	0.05	-0.03	-0.49	1		
$K_{sat}$ $mm\ h^{-1}$	0.18	-0.18	-0.11	-0.01	-0.24	1	0.51
Log $K_{sat}$	0.40	-0.26	-0.39	-0.04	-0.31	0.51	1.00
<b>Validation data set (N = 212)</b>							
Mean	45.80	35.89	18.31	1.8	1.42	173.75	1.46
Standard deviation	30.25	22.57	14.89	1.45	0.22	549.20	0.89
Minimum	1.81	1.00	0.00	0.01	0.59	0.03	-1.57
Median	43.45	34.40	14.35	0.56	1.47	35.54	1.55
Maximum	98.00	90.00	63.30	12.44	1.80	5400.00	3.73
<b>Correlations</b>							
Sand %	1.00						
Silt %	-0.88	1.00					
Clay %	-0.70	0.27	1.00				
OC %	-0.29	0.10	0.45	1.00			
Bulk density $Mg\ m^{-3}$	0.33	-0.17	-0.40	-0.52	1.00		
$K_{sat}$ $mm\ h^{-1}$	-0.04	-0.08	0.20	0.50	0.03	1.00	
Log $K_{sat}$	0.31	-0.28	-0.20	0.21	-0.17	0.52	1.00

## RESULTS AND DISCUSSION

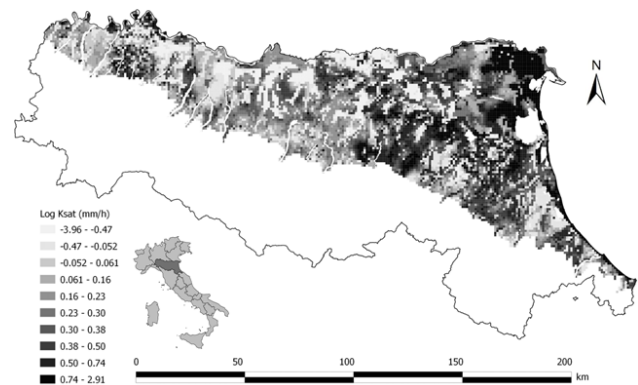
### PTF validation and uncertainty assessment

Results of applying the GMDH in order to estimate the logarithm of saturated hydraulic conductivity are shown in Table 4 along with the same error statistics for a number of PTFs from literature widely used to estimate saturated hydraulic conductivity. The overall poor performance of the tested PTFs, with the only possible exception of that of Rawls and Brakensiek (1989), justifies the development of a new and locally calibrated PTF.

**Table 4.** GMDH results and error statistics for the predicted hydraulic conductivity at saturation for a number of selected PTFs from literature. GMER: geometric mean error ratio; GSDER: geometric standard deviation of the error ratio; C: clay content (%); S: sand content (%); Si: silt content (%); BD: bulk density ( $Mg\ m^{-3}$ ); OC: organic carbon content (%). The slash separates the variable included in the second layer of the GMDH network (see Appendix B1).

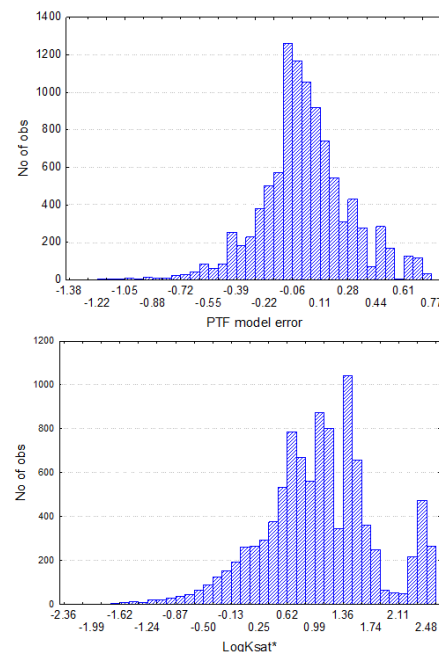
	New PTF GMDH	Rawls & Brakensiek 1989	Cosby et al. 1984	Dane and Puckett 1994	Saxton et al. 1986	Vereecken et al. 1990
<b>Calibration data set (N = 268)</b>						
Average absolute error	0.54	0.75	0.68	0.74	0.77	1.11
Absolute error std. dev.	0.53	0.68	0.76	0.82	0.81	0.98
Root mean square residuals	0.80	1.1	1.2	1.10	1.12	1.48
$R^2$	0.53	0.35	0.18	0.16	0.13	0.14
GMER	1.00	1.65	0.68	0.80	0.37	10.4
GSDER	5.73	9.77	10.12	12.55	10.75	12.19
<b>Validation data set (N = 212)</b>						
Average absolute error	0.85	0.89	0.70	0.85	0.75	0.85
Absolute error std. dev.	0.84	0.83	0.55	0.79	0.58	0.64
Root mean square residuals	1.19	1.22	0.89	1.16	0.94	1.6
$R^2$	0.08	0.15	0.09	0.04	0.05	0.03
GMER	0.40	0.33	1.56	1.32	0.70	1.77
GSDER	13.37	13.15	7.40	14.15	8.50	10.79
<b>Input variables</b>	S, OC, BD/C	S, C, BD	S, C	C	S, C	S, OC, BD/C

The values of GMER and GSDER for the calibration of the new PTF are 1.00 and 5.73 respectively. The established GMDH statistical network was eventually validated on the UNSODA data set (N = 212): in this case GMER and GSDER are equal to 0.40 and 13.4 respectively, indicating an underestimation of  $K_{sat}$ . These figures are almost coincident with those obtained through the validation of a widely used PTF (Rawls and Brakensiek, 1989) on the same data set, which resulted in a GMER equal to 0.33 and in a GSDER of 13.2. The same PTF of Rawls and Brakensiek (1989) applied to the calibration data set returned a GMER value equal to 0.6 and in a GSDER of 9.8. The map of the estimated  $K_{sat}$  ( $mm\ h^{-1}$ ) is shown in Fig. 4; these values were eventually used as input for the Green and Ampt infiltration model.



**Fig. 4.** Saturated hydraulic conductivity of bare soil (Log,  $mm\ h^{-1}$ ).

The uncertainty of the new PTF was quantified in terms of both model and input uncertainty. Fig. 5 shows the marginal frequency distribution of model residuals and of hydraulic conductivity obtained via bootstrapping; the results for the ten textural classes are reported in Table 5.



**Fig. 5.** Marginal frequency distribution of model error (left,  $mm\ h^{-1}$ , log scale) and log  $K_{sat}$  (right,  $mm\ h^{-1}$ ).

Model errors are on average positive for sandy, silty clay loam, silt loam and silty clay, negative for the remaining classes. In absolute terms, the error is greater for the sandy and

clayey classes. In terms of bootstrap variance, larger uncertainties are observed for clay, clay loam and silty clay, with the 95% confidence interval of the estimates varying over one order of magnitude or nearly two in the case of clays. Narrower 95% confidence intervals are observed for loam, sandy loam and sandy textural classes, with estimates varying over less than 0.5 order of magnitude (Fig. 6).

In the case of the Monte Carlo output only the observations within the limits of observed distribution were retained, which led to a total of 7291 sets of input variables to be inserted as inputs in the GMDH PTF; the distributions of resulting conductivities were estimated to calculate the uncertainties in predicted  $K_{sat}$  values. For all textural classes, model uncertainty resulted smaller than input uncertainty, suggesting that the PTF model was adequately calibrated (McBratney et al., 2002).

**Soil sealing effect on soil infiltration capacity**

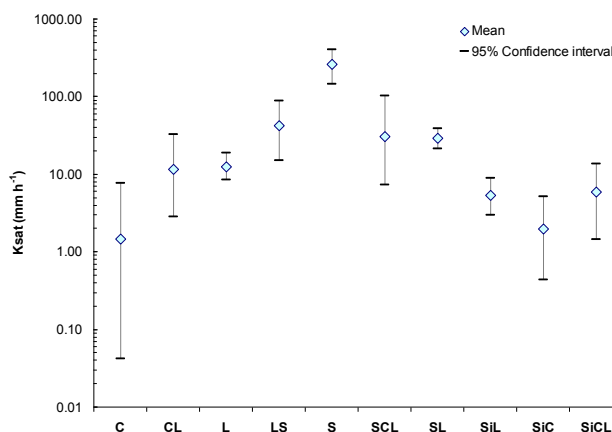
Assuming initial intermediate soil moisture conditions, i.e. volumetric water contents equal to the average of those at field capacity and wilting point, between 1976 and 2008, an increment of 8.4% in sealed areas due to urbanization and sprawl results in an average increment in surface runoff equal to 3.5 and 2.7% respectively for 20 and 200 years return periods (Fig. 7).

**Table 5.** Bootstrap and Monte Carlo statistics for the ten textural classes (values in  $mm\ h^{-1}$ , log scale). StdDev: standard deviation; CI: confidence interval; C: clay; CL: clay loam; L: loam; LS: loamy sand; S: sand; SCL: sandy clay loam; SL: sandy loam; SiL: silt loam; SiC: silty clay; SiCL: silty clay loam.

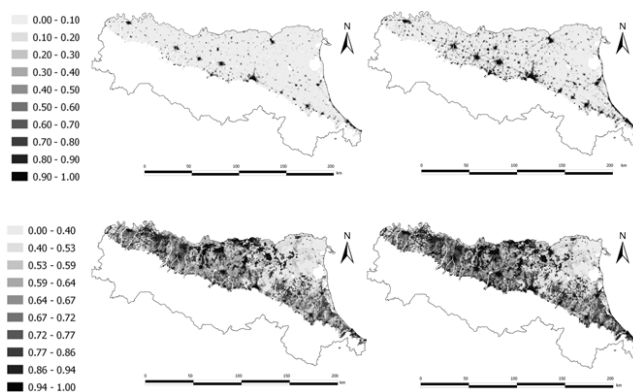
Textural classes	Bootstrap statistics				
	Num	Mean	Std.Dev.	95% CI lower	95% CI upper
C	1000	-0.138	0.554	-1.387	0.875
CL	1000	0.978	0.278	0.448	1.509
L	1000	1.090	0.089	0.920	1.263
LS	1000	1.576	0.202	1.167	1.941
S	1000	2.406	0.111	2.154	2.604
SCL	1000	1.393	0.274	0.853	2.004
SL	1000	1.454	0.066	1.325	1.579
SiL	1000	0.710	0.119	0.468	0.947
SiC	1000	0.214	0.281	-0.366	0.710
SiCL	1000	0.697	0.253	0.153	1.127
All Grps	10000	1.038	0.737	-0.490	2.514

Textural classes	Monte Carlo statistics				
	Num	Mean	Std.Dev.	95% CI lower	95% CI upper
C	293	0.178	1.320	0.027	0.330
CL	690	0.852	0.816	0.791	0.913
L	1319	1.071	0.838	1.026	1.116
LS	180	1.489	0.919	1.354	1.625
S	18	1.652	0.601	0.906	2.399
SCL	289	1.128	0.823	1.033	1.223
SL	1565	1.280	0.947	1.233	1.327
SiL	1028	0.689	0.860	0.636	0.741
SiC	783	-0.370	1.367	-0.466	-0.274
SiCL	1126	0.433	0.808	0.386	0.480
All Grps	7291	0.765	1.082	0.740	0.790



**Fig. 6.** Bootstrap 95% confidence intervals and mean values of  $K_{sat}$  estimates for USDA textural classes. C: clay; CL: clay loam; L: loam; LS: loamy sand; S: sand; SCL: sandy clay loam; SL: sandy loam; SiL: silt loam; SiC: silty clay; SiCL: silty clay loam.

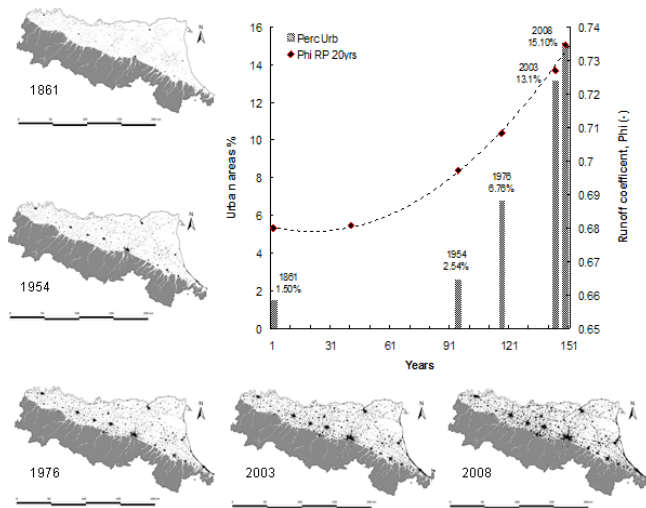


**Fig. 7.** Share of urban areas (top) and runoff coefficient Phi (RP 20 years, bottom) in 1976 (left) and 2008 (right).

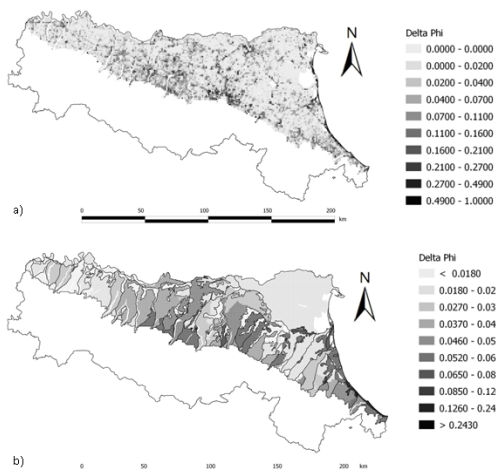
Under both rainfall scenarios, break-down ANOVA results show that the average values for the two reference years are significantly different at a probability level  $p < 0.05$ . It is interesting to point out that the increase in surface runoff between 1976 and 2008 is greater than that observed between an hypothetical time 0 with “bare soil condition” (i.e. no sealing) and 1976 (Fig. 8) corresponding to +3.05% and +2.34% respectively for 20 and 200 years return periods.

As an example, in Fig. 7 the changes in runoff between 1976 and 2008 for 20 years return period are shown. The differences are more evident in the urbanised surroundings of the main towns and along the coast. Aggregating the data at watershed level, an increment of runoff above 5% is recorded in 15 basins, for more than 2,000  $km^2$ , with a peak of 25% in the area surrounding the coast town of Rimini.

In order to assess the relevance of model parameters and soil sealing in affecting the value of the final runoff coefficient at the 1 x 1 km grid scale, we calibrated multiple linear regressions under the different simulation scenarios and analysed the corresponding beta coefficients, i.e. the regression coefficients of the standardized variables, which represent the change in terms of standard deviations in the dependent variable following a change of one standard deviation in an independent variable (Table 6).



**Fig. 8.** Phi values (RP 20 years) and urban areas % between 1861 and 2008.



**Fig. 9.** Difference between the Phi values at 1976 and 2008 (20-years RP): a) 1\*1 km grid; b) mean values for drainage basins.

**Table 6.** Multiple linear correlation results: Beta coefficient values under the different scenarios.  $Psi_e$ : air entry potential (cm).

Land use scenario	1976	1976	2008	2008
Rainfall scenario	RP 20	RP 200	RP 20	RP 200
Sealed area	0.326	0.269	0.471	0.398
Log $K_{sat}$	-0.649	-0.611	-0.624	-0.597
$Psi_e$	0.020	-0.125	0.014	-0.120
Total porosity	-0.397	-0.545	-0.359	-0.506
Initial moisture content	0.259	0.430	0.201	0.365

Under all scenarios the hydraulic conductivity explains the larger portion of the observed variability in runoff coefficient. In the case of the 20-years RP the figures are equal to 39.3% and 37.4% under the 1976 and the 2008 land use scenarios respectively, while for the 200-years RP the percentages drop, as expected for rainfall events generally exceeding soil hydraulic conductivity, to about 30% under both land use scenarios (30.8% in 1976 and 30.1% in 2008). In most cases, total soil porosity ranks second in affecting runoff generation, explaining part of observed variability ranging between 24.1% (land use 1976) and 21.5% (land use 2008) for the 20-years RP scenarios

and between 27.5% (land use 1976) and 25.5% (land use 2008) for the 200-years RP scenarios. The incidence of sealed area ranks third in most scenarios, and is greater under the 20-years RP rainfall scenarios and for the 2008 land use scenarios. It ranges from 19.8 to 28.2% for the 20-years RP scenarios respectively for 1976 and 2008, while under the 200-years RP rainfall scenarios the values drop to 13.6% and 20.1% for 1976 and 2008 land use scenarios respectively. The incidence of initial moisture content ranges between 15.7% (land use 1976) and 12.0% (land use 2008) for the 20-years RP scenarios and rises up between 21.7% (land use 1976) and 18.4% (land use 2008) for the 200-years RP scenarios. Although statistically significant at  $p < 0.05$ , the pressure head at the wetting front explains less the variation in runoff coefficient, with values ranging between 1.8% (land use 1976) and 0.9% (land use 2008) for the 20-years RP scenarios and between 6.3% (land use 1976) and 6.0% (land use 2008) for the 200-years RP scenarios.

At regional scale, between 1976 and 2008, the variation of the spatial pattern of urban sprawl mainly affected the southern (+15.2%) and the northern (+8.3%) coast, the Apennine margin (+12.4%) and the relieves of the alluvial plain (+9.4%). Along the coast, where the coarse textured soils of the functional group M are diffused, the increase of runoff is of about 13% with an increase of 9.5% in the sealed area, for the 20-years RP scenario. Over the fine textured soils of group G, along the Apennine margin, an increase of 14.7% in the sealed area results in an increase of 5% in runoff, for the same scenario; while the increased sealed areas of 11.7% result in a runoff increase of 5% in well drained, medium textured soils of the functional group B, which dominates the relieves of the alluvial plain. The share of sealed areas for those three groups in 2008 is greater than 23%, well above the regional mean of 14.7%, for the same reference year. On the other side, fine textured soils of functional groups A and the organic soils of group O are the least affected by sealing, with increments in runoff of 1.10 and 0.54% respectively for the 20 years RP scenario, following an increase of sealing equal to 5.3 and 0.8% respectively.

## CONCLUSIONS

This work presents a methodology to assess the effect of soil sealing on runoff generation and a case study in a densely populated agricultural area of Northern Italy characterised by a high urban sprawl. The modelling results indicate that assuming rainfall duration of 1 hour, an increment of 8.4% in sealed areas due to urbanization and sprawl between 1976 and 2008 results in an average increment in surface runoff equal to 3.5 and 2.7% respectively for 20- and 200-years return periods. These figures, strongly based on a hydrogeological approach and on locally calibrated PTFs, provide the basic information for the assessment of the hydraulic risk associated with increasing soil sealing and urbanisation. Due to differences in soil infiltration capacity, the local impact of sealing on soil ecosystem services depends upon the type of soil being sealed and the proposed methodology is then tailored to take explicitly into account such differences. The outcomes of this study provide a significant step towards the quantification of increasing of flood risks in the drainage networks of areas with significant soil sealing.

*Acknowledgement.* The authors gratefully acknowledge Regione Emilia-Romagna, Servizio Geologico, Sismico e dei Suoli for kindly providing basic soil data. The authors wish to thanks Dr. B. Rogiers, and an anonymous reviewer for their precious suggestions. Neither the Regione Emilia-Romagna,

nor any person acting on behalf of the Regione is responsible for the use which might be made of this publication.

## REFERENCES

- AbTech, 1996. Modelquest User's Manual. Version 4.0. AB-TECH, Charlottesville, VA.
- Autorità di Bacino del Fiume Po, 2001. Sub-plan for the geo-hydrological asset (PAI) – Interventions on the hydrographic network and on slopes. 7. Implementation rules - Directive on the project flood to be considered for designing and testing the hydraulic compatibility. (In Italian.)
- Autorità dei Bacini Regionali Romagnoli - Regione Emilia Romagna, 2003. Sub-plan for the geo-hydrological risk - Directive concerning hydraulic tests and technical measures to be implemented to achieve the safety objectives defined by the Master Plan for the Geo-hydrological Risk. Autorità dei bacini Regionali Romagnoli, Forlì, 83 pp. (In Italian.)
- Barreto, H., Howland, F.M., 2005a. Installing and Using the Monte Carlo Simulation Excel Add-in Software for Introductory Econometrics. Available at the following URL: <http://www3.wabash.edu/econometrics/EconometricsBook/Basic%20Tools/ExcelAddIns/MCSim.htm>. Last accessed: 07.05.2013.
- Barreto, H., Howland, F.M., 2005b. Installing and Using the Bootstrap Excel Add-in Software for Introductory Econometrics. Available at the following URL: <http://www3.wabash.edu/econometrics/EconometricsBook/Basic%20Tools/ExcelAddIns/bootstrap.htm>. Last accessed: 10.10.2012.
- Calzolari, C., Ungaro, F., 2012. Predicting shallow water table depth at regional scale from rainfall and soil data. *J. Hydrol.*, 414–415, 374–387.
- Carsel, R.F., Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. *Water Resour. Res.*, 24, 755–769.
- CEC, 2006. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions. Commission of the European Communities. Thematic strategy for Soil Protection. COM (2006) 231 final, Brussels.
- Cosby, B.J., Hornberger, G.M., Clapp, R.B., Ginn, T.R., 1984. A statistical exploration of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.*, 20, 682–690.
- Dams, J., Woldeamlak, S.T., Batelaan, O., 2008. Predicting land-use change and its impact on the groundwater system of the Kleine Nete catchment, Belgium. *Hydrology and Earth System Sciences*, 12, 1369–1385.
- Dane, J.H., Puckett, W., 1994. Field soil hydraulic properties based on physical and mineralogical information. In: van Genuchten, M.Th. et al. (Eds.), *Proceedings of the International Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*. University of California, Riverside, 389–403.
- Di Gennaro, A., Malucelli, F., Filippi, N., Guandalini, B., 2010. Dossier “Land use dynamics: analysis of Emilia Romagna between 1850 and 2000”. *Rivista Territori*, 1/2010, 60–71 (In Italian.)
- Efron, B., Tibshirani, R.J., 1993. An introduction to the bootstrap. *Monographs on Statistics and Applied Probability*, 57. Chapman & Hall, London, UK.
- FAO-ISRIC-ISSS, 1998. World reference base for soil resources.
- Franczyk, J., Chang, H., 2009. The effect of climate change and urbanization on the runoff of the Rock Creek basin in the Portland metropolitan area, Oregon, USA. *Hydrol. Process.*, 23, 805–815.
- Green, W.H., Ampt, G.A., 1911. Studies of soil physics, part I. The flow of air and water through soils. *J. Agr. Sci.*, 4, 1–24.
- Guermendi, M., Tarocco, P., 2007. Metadata on soil organic carbon contents in the soils of the Emilia-Romagna plain. Regione Emilia-Romagna, Geological, seismic and soil survey, 14 pp. (In Italian.)
- Istat, 2012. The problems associated to soil consumption. Hearing of Enrico Giovannini, President of the 13th Commission of the Senate of the Republic "Territory, Environment and Environmental Heritage", Rome January 18th, 2012 (<http://www.istat.it/it/archivio/51331>) (In Italian.)
- Ivakhnenko, A.G., 1968. The Group Method of Data Handling - a rival of the method of stochastic approximation. *Soviet Automatic Control*, vol. 13 c/c of *automatika*, 1, 3, 43–55.
- Johnson, M.E., 1987. *Multivariate Statistical Simulation*. John Wiley, New York
- Klute, A., Dirksen, C., 1986. Hydraulic Conductivity and Diffusivity: Laboratory Methods. In: Klute, A. (Ed.), *Methods of Soil Analysis: Part 1. Physical and Mineralogical Methods*, 2nd ed. Soil Science Society of America Book Series, vol. 9. Soil Science Society of America, Madison, WI., 444–461.
- Leij, F.J., Alves, W.J., Van Genuchten, M.Th., Williams, J.R., 1996. *Unsaturated Soil Database, UNSODA 1.0 User's manual*. Report EPA7600/R96/095. US Environmental Protection Agency, Ada, OK, 103 pp.
- Li, Z., Shao, Q.M., Xu, Z., Cai, X., 2010. Analysis of parameter uncertainty in semi-distributed hydrological models using bootstrap method: A case study of SWAT model applied to Yingluoxia watershed in northwest China. *J. Hydrol.*, 385, 76–83.
- McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. *Geoderma*, 117, 3–52.
- McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W., 2002. From pedotransfer functions to soil inference systems. *Geoderma*, 109, 41–73.
- Mein, R.G., Larson, C.L., 1973. Modelling infiltration during a steady rain. *Water Resour. Res.*, 9, 2, 384–394.
- Minasny, B., McBratney, A.B., Bristow, K.L., 1999. Comparison of different approaches to the development of pedotransfer functions for water retention curves. *Geoderma*, 93, 225–253.
- Motaghian, H.R., Mohammadi, J., 2011. Spatial estimation of saturated hydraulic conductivity from terrain attributes using regression, kriging, and artificial neural networks. *Pedosphere*, 21, 2, 170–177.
- Müller, J.A., Lemke, F., 2000. *Self-Organising Data Mining. Extracting Knowledge from Data*. Herstellung, Berlin.
- Nelson, D.W., Sommers, L.E., 1982. Total Carbon, Organic Carbon and Organic Matter. In: *Methods of Soil Analysis. Part 2: Chemical and Microbiological Properties*, Wiscosin, A.L. (Ed.). 2nd Edn., ASA and SSSA, Madison, WI., 539–579.
- Nemes, A., Rawls, W.J., Pachepsky, Ya.A., 2005. Influence of Organic Matter on the Estimation of Saturated Hydraulic Conductivity. *Soil Sci. Soc. Am. J.*, 69, 1330–1337
- Niehoff, D., Fritsch, U., Bronstert, A., 2002. Land-use impacts on storm-runoff generation: scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany. *J. Hydrol.*, 267, 80–93



- Pachepsky, Ya.A., Rawls, W., Giménez, D., Watt, J.P.C., 1998. Use of soil penetration resistance and group method of data handling to improve soil water retention estimates. *Soil and Tillage Research*, 49, 117–126.
- Pachepsky, Ya.A., Rawls, W.J., 1999. Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Sci. Soc. Am. J.*, 63, 1748–1757.
- Parasuraman, K., Elshorbagy, A., 2007. Estimating saturated hydraulic conductivity using genetic programming. *Soil Sci. Soc. Am. J.*, 71, 1676–1684.
- Pistocchi, A., 2001. The hydrological evaluation of urban plans: a simplified method for the hydraulic invariance in hydraulic zoning. *Ingegneria Ambientale*, XXX, 7/8, 407–413. (In Italian.)
- Poelmans, L., Van Rompaey, A., Batelaan, O., 2010. Coupling urban expansion models and hydrological models: How important are spatial patterns? *Land Use Policy*, 27, 965–975.
- Prokop, G., Jobstmann, H., Schönbauer, A., 2011. Overview on best practices for limiting soil sealing and mitigating its effects in EU-27. Technical Report - 2011-50, ISBN 978-92-79-20669-6.
- Rawls, W.J., Brakensiek, D.L., 1989. Estimation of soil water retention and hydraulic properties. In Morel-Seytoux, H.J., (Ed.), *Unsaturated Flow in Hydrologic Modelling - Theory and Practice*, Kluwer Academic Publishing, Dordrecht, 275–300.
- Rawls, W.J., Pachepsky, Ya.A., Ritchie, J.C., Sobek, T.M., Bloodworth, H., 2003. Effect of soil organic carbon on soil water retention. *Geoderma*, 116, 61–76.
- Regione Emilia Romagna, 2011. Land use 2008, edition May 2011. Available on line at the following URL: <http://geoportale.regione.emilia-romagna.it/it/catalogo/dati-cartografici/pianificazione-e-catasto/uso-del-suolo/2008-coperture-vettoriali-delluso-del-suolo-edizione-2011>, Last accessed: 09.05.2013.
- Regione Emilia Romagna, 2006. Map of the soils of Emilia-Romagna plain, scale 1:50,000. Regione Emilia Romagna, Geologic, seismic and soil survey; Available on line at the following URL: [http://www.regione.emilia-romagna.it/cartpedo/Delineazioni/Scelta\\_provincia\\_D.htm](http://www.regione.emilia-romagna.it/cartpedo/Delineazioni/Scelta_provincia_D.htm). Last accessed: 09.05.2013.
- Rogiers, B., Mallants, D., Batelaan, O., Gedeon, M., Huysmans, M., Dassargues, A., 2012. Estimation of hydraulic conductivity and its uncertainty from grain size data using GLUE and artificial neural networks. *Math. Geosci.*, 44, 739–763.
- Saxton, K.E., Rawls, W.L., Rosenberger, J.S., Papendick, R.I., 1986. Estimating generalized soil – water characteristics from texture. *Soil Sci. Soc. Am. J.*, 50, 1031–1036.
- Schaap, M.G., Leij, F.J., van Genuchten, M.Th., 2001. ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.*, 251, 765–779.
- Scholz, M., Grabowiecki, P., 2006. Review of permeable pavement systems. *Build. Environ.*, 42, 3830–3836.
- Tietje, O., Hennings, V., 1966. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes, *Geoderma*, 69, 71–84.
- Tóth, G., 2012. Impact of land-take on the land resource base for crop production in the European Union. *Sci. Total Environ.*, 435–436, 202–214.
- Ungaro, F., Calzolari, C., Busoni, E., 2005. Development of pedotransfer functions using a group method of data handling for the soil of the Pianura Padano-Veneta region of North Italy. Water retention properties. *Geoderma*, 124, 293–317.
- Ungaro, F., Staffilani, F., Tarocco, P., 2010. Assessing and mapping topsoil organic carbon stock at regional scale: a scorpan kriging approach conditional on soil map delineation and land use. *Land Degradation and Development*, 21, 565–581.
- USDA, 1993. Soil Survey Manual. USDA Handbook vol. 18. Government Printer, Washington, DC, 136–40.
- Vereecken, H., Maes, J., Feyen, J., 1990. Estimating unsaturated hydraulic conductivity from easily measured soil properties. *Soil Sci.*, 149, 1–12.
- Wijesekara, G.N., Gupta, A., Valeo, C., Hasbani, J.G., Qiao, Y., Delaney, P., Marceau, D.J., 2012. Assessing the impact of future land-use changes on hydrological processes in the Elbow River watershed in southern Alberta, Canada. *J. Hydrol.*, 412–413, 220–232.

## Appendix A

### A.1. Sensitivity analysis

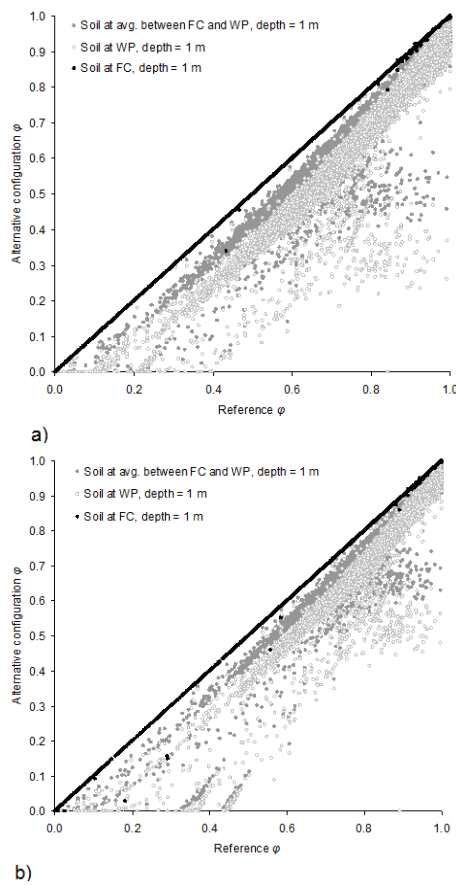
The runoff coefficient was computed as described in the paper using the Green-Ampt model where the initial soil moisture was set to the average between field capacity and wilting point and soil depth equal to 1 m. A sensitivity analysis of the model was conducted starting with a reference configuration where the initial soil moisture was set to field capacity and soil depth equal to the 95<sup>th</sup> percentile of the average depth of the water table as estimated by Calzolari and Ungaro (2012).

The following configurations were simulated: i) Soil at field capacity (FC), soil depth of 1 m; ii) Soil at the average of FC and wilting point (WP), soil depth of 1 m; and iii) soil at WP, soil depth of 1 m. These configurations were used to simulate runoff coefficients ( $\phi$ ) corresponding to rainfall events of 20 and 200 years return period (RP). The graphs in

Fig. A1 show the scatter diagram of  $\phi$  in the reference configuration (x-axis) and in each of the above alternative configurations (y-axis) for the reference year 2008.

As it can be seen, for both 20- and 200-years RP, introducing the depth to the water table does not yield appreciable variations compared to a constant 1 m deep soil (points aligned on the 1:1 line). Assuming initial soil moisture at the average of WP and FC yields a somehow intermediate result between soil at FC and soil at WP, as logically expected. Based on experience from field surveys, it may be argued that when extreme rainfall events occur (typically in autumn) soils are seldom fully replenished after a usually dry summer season, so assuming them to be at FC might yield overestimation of  $\phi$ , while the average of WP and FC seems more realistic, although there is no conclusive evidence at present to support this argument. It should be considered that assuming soil at FC would generally increase  $\phi$  in unsealed soil

based on this calculation, thus yielding estimates of lower impacts of sealing on soils than obtained under the assumption of soil at average between WP and FC, as presented in the paper.



**Fig. A1.** Scatter diagrams of phi in the reference configuration (x-axis, soil at field capacity, FC, and water table, WT, depth equal to the 95<sup>th</sup> percentile of the average depth of the water table) and with a constant soil depth of 1 m, and soil at FC (black dots); soil at wilting point (WP, white dots), and; soil at the average between FC and WP (light grey dots). a) 20 years RP; b) 200 years RP.

## Appendix B

### B.1 Pedotransfer function for the calculation of Log K<sub>sat</sub>

Figure B1 shows the structure of the GMDH network for the estimation of the saturated hydraulic conductivity (Log K<sub>sat</sub>, mm h<sup>-1</sup>); the inputs to the network are: sand content (S%), organic carbon content (OC%), clay content (C%) and bulk density (BD, g/m<sup>3</sup>). In case of K<sub>sat</sub> estimates out of the calibration ranges (4.5% of total, 598 out of 13215 grid cells), the mean measured values of the K<sub>sat</sub> for each textural class are used (Table B1).

### Input data normalisation 1

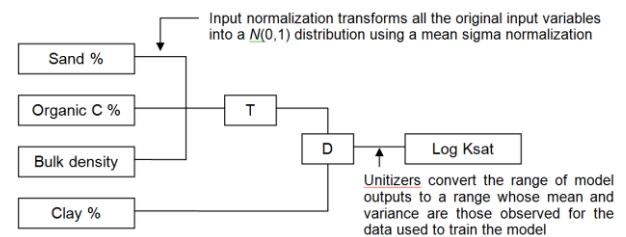
$$\begin{aligned}
 \mathbf{A} &= -1.58807939026797 + 0.043399261792585 * \mathbf{S}\% \\
 \mathbf{B} &= 0.895704956226224 + 1.90436429623194 * \text{Log}(\text{OC}\%) \\
 \mathbf{C} &= -6.56853764508901 + 4.86964821073382 * \text{BD} \\
 \text{First net node output calculation (T)} \\
 \mathbf{D} &= 0.0836015967122676 + 0.416153421876942 * \mathbf{A} + 0 * \mathbf{B} \\
 &+ -0.445772230372808 * \mathbf{C} + -0.134334562427775 * \mathbf{A}^2 + - \\
 &0.0544032825090131 * \mathbf{B}^2 + 0 * \mathbf{C}^2 + -0.198584356240143 \\
 &* \mathbf{A} * \mathbf{B} + 0.096017083549925 * \mathbf{A} * \mathbf{C} + -0.144971973570487 \\
 &* \mathbf{B} * \mathbf{C} + -0.125314845712128 * \mathbf{A} * \mathbf{B} * \mathbf{C} + \\
 &0.0716205065345566 * \mathbf{A}^3 + -0.123673920290893 * \mathbf{B}^3 + \\
 &0.0587033757513834 * \mathbf{C}^3 + 0.152071065904968 \\
 &* \mathbf{A}^2 * \mathbf{B} + -0.102414907323155 * \mathbf{A} * \mathbf{B}^2 + 0 * \mathbf{A} * \mathbf{C}^2 + 0 * \\
 &\mathbf{A}^2 * \mathbf{C} + -0.156148957009503 * \mathbf{B}^2 * \mathbf{C} + - \\
 &0.0305947339564729 * \mathbf{B} * \mathbf{C}^2
 \end{aligned}$$

### Input data normalisation 2

$$\begin{aligned}
 \mathbf{E} &= -1.65784483580252 + 0.0723996897397788 * \mathbf{C}\% \\
 \text{Second net node output calculation (D)} \\
 \mathbf{F} &= 0.174732633027132 + 1.03829358346795 * \mathbf{D} + \\
 &0.151865200913679 * \mathbf{E} + -0.341867233805586 * \mathbf{D}^2 + - \\
 &0.11074767165566 * \mathbf{E}^2 + -0.13351716106058 * \mathbf{D} * \mathbf{E} + - \\
 &0.461153694809156 * \mathbf{D}^3 + 0 * \mathbf{E}^3 + -0.367403924274197 \\
 &* \mathbf{D}^2 * \mathbf{E} + 0.218709934220076 * \mathbf{D} * \mathbf{E}^2
 \end{aligned}$$

### Final output calculation

$$\text{Log } K_{\text{sat}}^* = 0.951757462686567 + 1.10472133391165 * \mathbf{F}$$



**Fig. B1.** Structure of the GMDH network for the estimation of the saturated hydraulic conductivity.

**Table B1.** Tabular values for estimates out of the PTF calibration range (Log K<sub>sat</sub> > 2.924, Log K<sub>sat</sub> < -4.00): measured mean values for (USDA) textural classes.

USDA textural class	Log K <sub>sat</sub> (mm hr <sup>-1</sup> )
Clay	-0.16
Clay loam	0.97
Loam	1.9
Loamy sand	1.57
Sand	2.41
Sandy clay loam	1.4
Sandy loam	1.46
Silt loam	0.72
Silty clay	0.2
Silty clay loam	0.68

Received 30 November 2012  
Accepted 15 July 2013