



# Estimation of infiltration rate from readily available soil properties (RASPs) in fallow cultivated land

P. K. Pandey<sup>1</sup> · Vanita Pandey<sup>1</sup>

Received: 30 September 2017 / Accepted: 9 July 2018 / Published online: 13 July 2018  
© Springer Nature Switzerland AG 2018

## Abstract

The information about water transmission rate into the ground is vital concerning the productivity of water system and seepage, advancing the accessibility of water for the plants, enhancing the yield of harvests, limiting degradation of soil and wastage of the water. Infiltration rate can successfully be measured using double ring infiltrometer. However, measurement of infiltration in the field is labour and time consuming and difficult especially in mountainous sites. As an alternative, RASP-based infiltration models can be developed. The study was carried out in the field near the NERIST campus (Nirjuli Complex), Arunachal Pradesh, India. Twenty sites were identified at the grid of 10 m interval, and the field measurement of infiltration was performed. The soil was analysed for properties, namely, soil texture, bulk density (BD), particle density (PD), moisture content (MC), and organic carbon content (OC) for each site. The basic infiltration varied from 1 to 4.84 cm/h. The Scatter plot between RASPs and infiltration rate revealed that there is a positive correlation with OC, PD, and sand, and a negative correlation with BD, MC, silt, and clay. The partial least square regression (PLSR) analysis was carried out to develop predictive models for five different groups of inputs of soil properties. The influential variable projection (VIPs) analysis revealed sand as a highly influential factor, while silt as a reluctant predictor of infiltration characteristics of the study site. It was found that to predict the soil infiltration rate based on RASPs with seven independent variables (Eq. 13) with coefficient of determination ( $R^2$ ) 0.92, root mean square error (RMSE) 0.378 cm/h, mean absolute error (MAE) 0.143 cm/h, and standard error (SD) 0.398 cm/h is strongly recommended for the prediction of infiltration characteristics.

**Keywords** Infiltration rate · Soil physical properties (RASPs) · Organic carbon · PLSR · VIP · Prediction

## Introduction

In the context of changing climatic conditions, soil and water are considered as essential natural resources for crop production are degrading due to lots of pressure on this two natural reservoir. Effective managing of this water will require prominent control of infiltration processes of soil, which in turn help to minimise numerous issues such as upland submerging, contamination of water resources (surface and groundwater), declining water tables, and wasteful water system of rural terrains.

The infiltration rate of soil is a necessary hydrological phenomenon that influences the surface water distribution

consistency due to its component of exchange and disseminates water from the landscape to the subsurface soil profile. In addition, an adequate understanding of site-specific infiltration characteristics of soil is vital to understand and regulate of soil and water-related natural perils.

Information about the soil infiltration process is essential for management and monitoring of hydrological processes and water-induced hazards in the watershed. The design and evaluation of farm irrigation delivery systems likewise depend on the intake rate of soil, since intake rate influences different design parameters of the irrigation system.

Water transmission inside the soil is guided by Darcy's law, which expresses that the flux is a function of hydraulic conductivity and hydraulic gradient. The hydraulic gradient is consisting of gravity potential, osmotic potential, and matric potential (transmission of water from wet to dry soil) (Singer and Munns 1999). The infiltration rate is higher at the initial and decreases in elapsed time as it approaches the saturated hydraulic conductivity. The

✉ Vanita Pandey  
vanita2@gmail.com

<sup>1</sup> Department of Agricultural Engineering, North Eastern Regional Institute of Science and Technology, Nirjuli, Itanagar, Arunachal Pradesh 791109, India

constant rate at which water enters into the soil after attaining saturation is termed basic infiltration rate. The intake rate relies upon the physical, chemical, and organic properties of the surface soil, the underlying dispersion of water in soil preceding irrigation, the transmission and redistribution of water over the surface.

Different physical properties of the soil are influencing infiltration processes. Out of these, soil texture has a significant impact on infiltration (Huang et al. 2011; Mousavi 2015). Soil texture is one of the primary factors affecting infiltration characteristics. Water rushes through substantial pores of sandy soil than through the little pores of clay soil, particularly if clay is compressed and has a practically no structure upon the sum and kind of clay minerals, some clay soil create shrinkage breaks as they dry. Osuji et al. (2010) reported that soil organic matters, particle density, bulk density, porosity, moisture content, texture influence the infiltration. The depthwise classification of soil organic matter content has been recommended as a pointer of soil quality, because the organic matter is necessary to control soil loss, water intake to the soil, and protection of nutrients (Franzluebbers 2002). The bulk unit weight of soil has an inverse relation with infiltration rate. An increase in soil bulk density decreases soil porosity resulting in a decrease in soil infiltration characteristics. Initial moisture content is one of the critical parameters, which influences water retention and transmission characteristics of the soil. Pandey et al. (2009) reported that the hydraulic conductivity of sandy soil decreases with depth. Therefore, initial moisture content is one of the potential input variables to model infiltration characteristics of the soil.

The physical properties of soil and condition change over and altogether make infiltration characteristics itself show a large variety at the scale. The soil physical properties and condition have sizeable spatial variability driving considerable variability in infiltration characteristics of the soil. Field-scale characterisation of infiltration is tedious due to the necessity of various measurements (Khatr and Smith 2005). Estimation of the features of infiltration is hard due to significant spatiotemporal variability (Pandey and Pandey 2010).

Several studies around the globe (Mirzaee et al. 2014; Parhi 2014; Tuffour and Bonsu 2015) recommended the use of infiltration models as an alternative to field-based infiltration measurement. The application of a specific model may be contingent on several factors, such as the type of application, the desired level of accuracy, and user expertise (Clausnitzer et al. 1998). Application of particular model requires detail understanding of assumptions and boundary conditions of the specific model. Most of this model based on the fundamental assumption of uniform absorption of soil moisture, rapid surface ponding, and saturation of soil at final infiltration. These assumptions rarely observed under

actual field conditions, which may lead to the inaccurate prediction of infiltration characteristics.

An alternative approach to measurement and infiltration models is to develop site-specific pedotransfer functions (PTFs) of infiltration rate as a function of soil properties. Many successful applications of PTFs in infiltration estimation reported in the literature such as (Arshad et al. 2010; Ghorbani-Dashtaki et al. 2010; Ekhmaj 2010; Kashi et al. 2014; Mahdian et al. 2009; Parchami-Araghi et al. 2013). Ghorbani-Dashtaki et al. (2016) concluded that soil physical properties derived pedotransfer functions could be successfully used to predict infiltration rate with reasonable accuracy. Mousavi (2015) successfully used PTFs to predict infiltration characteristics in Iran. Hence, it is reasonable to correlate infiltration characteristics with less spatially easily measurable soil properties to model infiltration characteristics at field scale (Rawls et al. 1983; Saxton et al. 1986; Wosten and Genuchten 1988; Mbagwu 1994; Van de Genachte et al. 1996; Saxton and Rawls 2006).

As literature suggests, several techniques could be applied to develop infiltration PTFs. Linear and non-linear regression are widely used approaches (Rahmati 2017). To develop reliable PTFs also necessitate correct identification of the most influential variable.

Partial least square regression is a robust multivariate regression technique that enables the operator to play out an extensive variety of investigations (Martens and Martens 2000). Developing accurate and reliable models to predict soil infiltration rate is one of the most apprehensive topics in hydrology and selecting predictors that are more relevant is a critical judgement in model development. PLSR, which finds an expedient connection between a set of predictors and output factors, not just give a particular methodology to choose essential input variables yet, also, brings about more exact and reliable estimates than other generally used approaches.

PLSR is reasonable for selectivity improvements of scientific instruments (Martens and Martens 2000). PLSR is a technique for relating two information frameworks, A and B, by a straight multivariate model, yet goes past the customary regression in that it demonstrates likewise the structure of A and B. PLSR gets its convenience from its capacity to investigate variables with numerous, noisy, collinear, and even deficient factors in both A and B (Wold et al. 2001). Details of the hypothesis, standards, and utilisation of PLSR can be found in the literature (Abdi 2010). The nonlinear models created utilising PLSR has a conditional implementation of multiple modern data-driven methods like ANN (Abudu et al. 2011). Moreover, PLSR has an advantage over the other machine learning system as a modelling procedure was direct, and the conditions were more straightforward and more expressive than the other data-driven methods (Abudu et al. 2011).

Despite the considerable amount of research works have been done to understand the relationship between readily available soil properties (RASPs) and soil infiltration characteristics (Rashidi et al. 2013; Lake et al. 2009), insufficient work has been led to model soil infiltration using different physical properties of soil.

As reviewed above, an attempt was made to develop predictive models as an alternative to field measured infiltration rate of fallow agricultural land located near the NER-IST campus (Nirjuli Complex) using the PLSR technique based on RASPs, and to verify the models predictively with measured infiltration rate. However, no application of PLSR to model soil infiltration characteristics based on RASPs is reported in the study area (best of our knowledge). Therefore, applying PLSR to develop reliable and accurate models to predict soil infiltration seems to be interesting.

## Materials and methods

### About study area

The study was carried out on the sandy field near the NER-IST campus (Nirjuli Complex) campus. Locations of each measurement station were recorded with the help of global positioning system (GPS) instrument. Geographical details (longitude, latitude, and altitude) were recorded at each site (Table 1). The sampling distance was maintained at 10 m interval, as depicted in Fig. 1.

### Measurement of infiltration rate

In this study, the infiltration rate was measured using double ring infiltrometer (Perroux and White 1988) which consists of two concentric metal rings and gauge. The distance across of the inward and outward ring was 25 and 35 cm individually, and both had an equivalent height of 25 cm. Rings are put concentrically and hammered with the help of rammer to ensure 12 cm insert of rings into the soil consistently. Measuring gauge has been put on rings. A thin sheet of plastic was used inside the ring to avoid the surface soil disturbance due to splashing action of poured water. The water poured over the plastic sheet up to a depth of 15–20 cm. The plastic sheet was removed slowly and the outer ring filled with water up to the same level as inside or more but not less than inside water level. Observations were recorded using point or hook gauge fixed over the gauging stand. The rate of fall was measured at a constant time interval of 5, 10, and 15 min. The change of time was done when we got three or four consecutive constant reading for same time interval and time interval was increased gradually to get more accurate basic infiltration rate.

**Table 1** Location of sampling sites

Site	Latitude (°N)	Longitude (°E)	Elevation (m)
1	27°08'5.1"	93°44'35.2"	114
2	27°08'4.8"	93°44'35.1"	113
3	27°08'4.4	93°44'35.2"	115
4	27°08'4.1"	93°44'35.4"	116
5	27°08'5"	93°44'34.8"	114
6	27°08'4.7"	93°44'34.8"	114
7	27°8'4.4"	93°44'34.8"	114
8	27°8'4.1"	93°44'34.8"	114
9	27°8'5"	93°44'34.5"	115
10	27°8'4.8"	93°44'34.5"	116
11	27°8'4.4"	93°44'34.5"	115
12	27°8'4.1"	93°44'34.6"	113
13	27°8'5.6"	93°44'34.1"	112
14	27°8'4.7"	93°44'34.1"	111
15	27°8'4.4"	93°44'34.1"	112
16	27°8'4.1"	93°44'34.2"	115
17	2°8'5.7"	93°44'34.2"	116
18	27°8'4.7"	93°44'34.2"	117
19	27°8'4.4"	93°44'34.2"	116
20	27°8'4.1"	93°44'34.2"	116

### Soil physical properties

The soil sample was collected and analysed in the laboratory for the determination of soil properties viz. soil texture (sand %, silt %, and clay %), soil textural class, bulk density, particle density, moisture content, and organic carbon content, respectively.

### Measurement of moisture content

A soil sample collected at each site at 30 cm depth before conducting the infiltration experiment were analysed for percentage moisture content using hot air oven-drying method. In this method, soil samples are kept for 24 h at 150 °C temperature. Each sample was kept in three dishes for determining the moisture content. After getting all the three dishes moisture content, the average was taken as the moisture content of the sample. The moisture content (MC) in percentage calculated using the formula suggested by Punamia et al. (2005):

$$MC = \frac{(M_2 - M_1) - (M_3 - M_1)}{(M_3 - M_1)}, \quad (1)$$

where  $M_1$  = mass of empty container with a lid (g),  $M_2$  = mass of container with wet soil and lid (g), and  $M_3$  = mass of container with dry soil and lid (g).





Fig. 1 Map of selected study site

### Measurement of BD (bulk density)

The bulk density of soil samples determined using Grossman and Reinsch (2002) and Punamia et al. (2005) methodology. The cylindrical core cutter of 130 mm long and 100 mm internal diameter was used to calculate soil bulk density. The bulk density (BD) in  $\text{g/cm}^3$  was calculated as suggested by Punamia et al. (2005):

$$\text{BD} = \frac{(M_2 - M_1)}{V}, \quad (2)$$

where  $M_1$  = mass of cylinder (g),  $M_2$  = mass cylinder + mass of wet soil (g), and  $V$  = volume of cylinder ( $\text{cm}^3$ ).

### Measurement of PD (particle density)

Soil particle density was measured with the help of pycnometer (Punamia et al. 2005; Flint and Flint 2002). A 10-g screened soil sample with 200- $\mu\text{m}$  sieve was used for analysis. The following formula used to determine the value of particle density (PD) in  $\text{g/cm}^3$ :

$$\text{PD} = \frac{(M_2 - M_1)}{(M_2 - M_1) - (M_3 - M_4)} \times \rho, \quad (3)$$

where  $M_1$  = mass of empty pycnometer (g),  $M_2$  = mass of dry soil, and pycnometer (g),  $M_3$  = mass of dry soil, water, and pycnometer (g),  $M_4$  = mass of water and pycnometer (g), and  $\rho$  = water density ( $\text{g/cm}^3$ ).

### Soil texture analysis

The textural class of soil samples was analysed by the hydrometer method (Gee and Or 2002). A dispersing reagent was prepared by adding 50 g of sodium hexametaphosphate into 1 l of distilled water and stirred until it completely dissolved. The oven-dried soil samples of 5 g were put into 500 ml beaker in which 100 ml of dispersing solution and 500 ml of distilled water was added. The mixture was stirred for 15 min and was kept for 24 h. After keeping 24 h, the mixture was carefully transferred into 500-ml measuring cylinder without any slippage and distilled water added to make it 500 ml. Afterwards, the hydrometer was dipped into the

mix, and the reading was taken at 2 min, 12 min, and 24 h in the meantime the temperature of the mixture was taken. The process based on the sand particles settles entirely within 12 min, silt settles within 24 h, and remaining clay as suspended. Afterwards, percentage of sand, silt, and clay contents was determined. The soil triangle diagram (USDA 1951) demonstrating the percentage of sand, silt, and clay in principal textural class was used to determine soil type based on primary particle percentage.

### Determination of soil organic carbon (OC) content

The wet oxidation technique (Nelson and Sommers 1982) was used to determine the percentage of organic carbon (OC) in the site-specific soil samples. The 0.48-g dry soil sample screened with the help of 200- $\mu\text{m}$  sieve was taken for analysis. The soil sample was put into a 50-ml beaker in which 11 drops of distilled water was added. After that, 1 vial of  $\text{OC}_1$  and  $\text{OC}_2$  was added. It was kept for 30 min, then after 47 ml of distilled water was added again. The mixture was stirred carefully and was put in a test tube. At the same time, new sample, i.e., without soil was prepared in the same process. The test tube was placed in the device, and OC values were recorded.

### Partial least square regression (PLSR) analysis

Regression analysis is a numerical method for the examination of connections between related factors or variables. Partial least square regression (PLSR), which is based on linear regression algorithm commonly used in soil analysis. PLSR especially useful when the number of variables is more when compared to the number of samples. This advantage of PLSR leads to its application in numerous fields including soil science. In the design of PLSR, the measured basic infiltration rate was taken as a dependent variable. The different RASPs (OC, PD, BD, sand, silt, and clay) were used as independent variables. Using the parameters PLSR, based basic infiltration rate models, was formed for different combinations of inputs. The developed models were verified using an independent data set based on various statistical indices.

### Description of PLSR

Multivariate regression method like partial least squares regression (PLSR) enjoys immense popularity in a wide range of fields, including the soil sciences. The main reason is that they have been designed to confront the situation that there are many, possibly correlated, predictor variables, and relatively few samples situation that is common, especially in soil physics. The equation of the PLS regression model writes

$$Y = T_h C_h + E_h = X W_h C_h + E_h = X W_h (P_h W_h)^{-1} C_h + E_h, \quad (4)$$

where  $Y$  is the matrix of the dependent variables,  $X$  is the matrix of the input factors in this case RASPs [sand (%), silt (%), clay (%), BD ( $\text{g}/\text{cm}^3$ ), OC (%), and moisture content of the soil (%)].  $T_h$ ,  $C_h$ ,  $W_h^*$ ,  $W_h$ , and  $P_h$ , are the matrices generated by the PLSR algorithm, and  $E_h$  is the matrix of the residuals.

### Statistical evaluation of developed models

The different statistical indices were used to evaluate the predictive accuracy of developed models are followings:

#### Root mean square error (RMSE)

RMSE considered better for showing bigger deviations. By squaring the errors, we can get results that are more accurate as the negative and positive errors do not cancel out each other and stay in existence until the end of the computation, thus adding more accuracy to the result. RMSE was calculated as explained by Pandey et al. (2014):

$$\text{RMSE} = \left( \frac{1}{n} \sum_{i=1}^n (M_i - E_i)^2 \right)^{0.5}, \quad (5)$$

where  $M_i$ =measured value  $E_i$ =estimated value,  $n$ =number of samples.

#### Mean square error (MSE)

MSE has the easiest interpretation and useful to compare the precision between different volumes under study. The mean square error was calculated as explained by Pandey et al. (2014)

$$\text{MAE} = \frac{\sum_{i=1}^n |M_i - E_i|}{n}. \quad (6)$$

#### Standard deviation (SD) and coefficient of variation (CV)

Both SD and CV indices are used to measure variability in observations from the mean value. The values of SD CV was calculated as follows:

$$\text{SD} = \left( \frac{1}{n} \right) \sqrt{\sum_{i=1}^n (m_i - m')^2}, \quad (7)$$

$$\text{CV} = \frac{\text{SD}}{m'} \times 100, \quad (8)$$

where  $n$  = total number of observations,  $mi$  = sample observation,  $m'$  = mean of  $n$  observations.

## Results

### Infiltration characteristics of sites

The summary of the average infiltration rate measured (cm/h) and the accumulated infiltration depth (cm) against the elapsed time for each location are presented (Fig. 2a, b). We notice that initially, the infiltration rate was high for all the selected sites ranged from (2.4–14.4 cm/h), and it reduces as soil wet, and it becomes almost constant after the soil is saturated. The corresponding accumulated infiltration rate ranged (1.8–17.8 cm). The basic infiltration rate at selected sites was ranged between 35 and 90 min from the start of infiltration. The site-specific details of infiltration characteristics, as depicted in Fig. 2a, b. Based on the analysis of results it may infer that there is high spatial variability in the infiltration rate among the sites.

### Soil physical properties

The site-specific different soil properties were estimated along with their descriptive statistics and depicted in Tables 2 and 3, respectively. From the analysis of results, it was observed that the sand percentage is 92.73%, silt percentage is 2.25%, and the clay percentage is 5.018%. The high proportion of sand was observed, because the field was sandy soil. The BD and PD range from 1.162 to 1.789 and 1.996–2.99 g/cm<sup>3</sup>, respectively, with the average bulk density 1.5531 g/cm<sup>3</sup> and particle density 2.447 g/cm<sup>3</sup>. The moisture content varies from 19.96–26.91% with a mean value of 23.22%. The OC ranges from 0.264 to 0.364% with the average value 0.322%. The textural class of the study area was sandy soil.

### Scatterplot of the measured infiltration rate (IR) vs physical properties of soil

Figure 3 depicts the association between per cent sand vs IR, per cent silt vs IR, per cent clay vs IR, bulk density against IR, per cent moisture content (MC) against IR, and per cent, organic carbon content vs IR. From the analysis in Fig. 3, it can be depicted that the soil property is either positively or inversely proportional to the infiltration rate. Sand content of the soil is the crucial factor, which affects the soil infiltration rate. Figure 3 shows that sand content of the soil is directly proportional to the soil infiltration rate. Silt content is inversely related (Fig. 3). The clay content of

the soil is the most crucial factor, which affects the soil IR. According to Fig. 3, the clay content of the soil is indirectly proportional to the IR. The BD content of the soil is another crucial factor, which affects the soil IR. It was observed that BD the content of the soil is inversely proportional to the soil IR (Fig. 3). The PD content of the soil has a similar relation to sand (Fig. 3). The MC and infiltration rate are negatively related (Fig. 3). The OC content of the soil has the significant effects on the soil IR. The OC content and infiltration rate of soil has a positive relationship (Fig. 3).

### Correlation between the infiltration rate (IR) and independent variables

The relationship between the measured IR and RASPs is presented in Table 4. The analysis of results presented in Table 4 demonstrates that infiltration rate has a positive relationship with sand, PD, and OC by 0.883, 0.691, and 0.335 separately, which shows cumulative the sand, PD, and OC, will improve the IR of the soil. The clay, silt, BD, and MC have an adverse (negative) association with the measured IR of  $-0.775$ ,  $-0.563$ ,  $-0.743$ , and  $-0.673$  separately. It demonstrates that expanding the clay, silt, BD, and MC will have diminished the IR of the soil. Among all, the sand has a maximum positive correlation of 0.883 took after by clay as a negative relationship ( $-0.775$ ) which cause a substantial effect on infiltration characteristics of the soil.

### Prediction model using partial least square regression (PLSR) method

For the development of prediction models of IR, the analysis was categorising into five different input groups. The soil texture was considered the necessary input of all the group. Group A has only soil texture as input. The group B has one additional input as bulk density along with texture. The group C has two additional input as bulk density and particle density along with texture. The group D has three additional inputs as bulk density, particle density, and moisture content along with texture. The group E has four additional inputs bulk density, particle density, moisture content, and organic carbon content along with texture. In all the developed models of different groups, independent variables have a measurement unit [% (sand, silt, clay, MC, OC), g/cm<sup>3</sup> (BD, PD)].

### Development of predictive model for group A

The infiltration rate model for group A was developed using sand, silt, and clay as the independent variable. It was observed that the average percentage of sand, silt, and clay across the selected sites was 93.07, 2.34, and 5.187, respectively. The developed equation for the estimated infiltration rate (EIR) is given below:

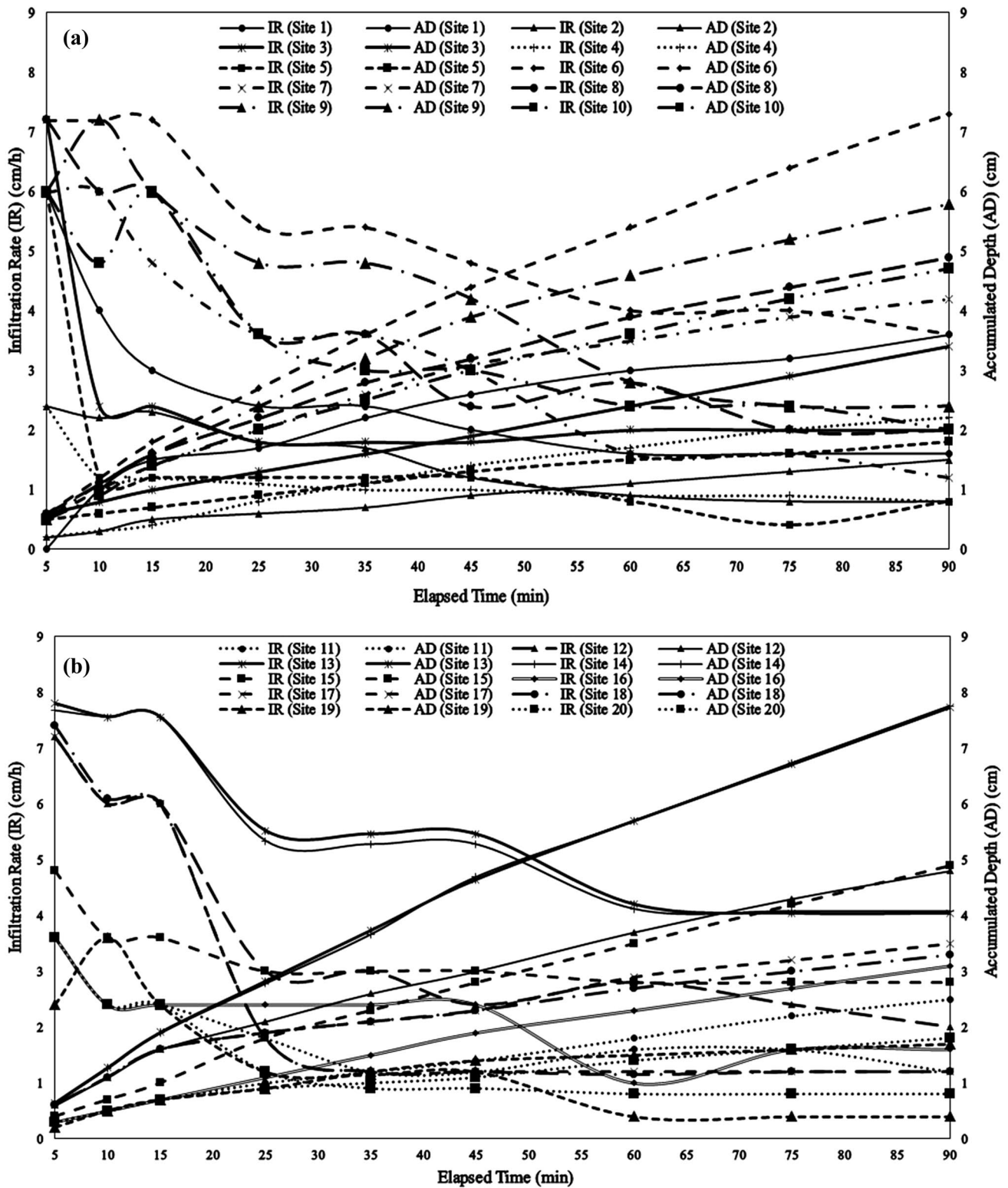


Fig. 2 a Plot of measured infiltration rate and accumulated depth against time for site 1 to site 11. b Plot of measured infiltration rate and accumulated depth against time for site 11 to site 20



**Table 2** Measure RASPs and infiltration rate at study area

Site	Sand (%)	Silt (%)	Clay (%)	Textural class	BD (g/cm <sup>3</sup> )	PD (g/cm <sup>3</sup> )	MC (%)	OC (%)	IR (cm/h)
1	93.13	2.21	5.56	Sandy soil	1.671	2.51	25.9	0.3	2.73
2	92.98	2.5	5.15	Sandy soil	1.76	2.07	21.8	0.31	1.57
3	93.029	2.38	5.6	Sandy soil	1.379	2.18	25.2	0.34	2.60
4	92.84	2.4	6.05	Sandy soil	1.789	2.32	26.91	0.29	1.16
5	92.78	2.51	6.46	Sandy soil	1.553	2.45	23.3	0.29	1.55
6	93.326	2.11	4.12	Sandy soil	1.306	2.76	22	0.33	5.42
7	93.04	2.37	4.8	Sandy soil	1.416	2.71	20	0.32	3.49
8	93.093	2.18	4.4	Sandy soil	1.406	2.73	23.7	0.31	3.95
9	93.29	2.4	5	Sandy soil	1.441	2.702	21.17	0.34	4.51
10	93.15	2.41	5.47	Sandy soil	1.501	2.32	25.2	0.31	3.68
11	92.76	2.32	5.95	Sandy soil	1.44	2.48	24.72	0.32	1.88
12	93.22	2.38	4.87	Sandy soil	1.626	2.43	21.96	0.33	3.86
13	93.353	2.1	4.4	Sandy soil	1.4	2.63	22.2	0.34	5.73
14	93.466	2.26	3.94	Sandy soil	1.251	2.74	18.1	0.34	5.66
15	93.23	2.52	5.75	Sandy soil	1.653	2.65	24.4	0.34	3.27
16	93.04	2.29	6.2	Sandy soil	1.649	2.18	24.67	0.35	2.20
17	92.965	2.16	4.47	Sandy soil	1.434	2.456	22.7	0.35	3.00
18	93.23	2.47	4.7	Sandy soil	1.662	2.15	22.6	0.33	2.99
19	92.7	2.59	5.7	Sandy soil	1.68	2.41	26.01	0.36	1.46
20	92.974	2.3	5.15	Sandy soil	1.781	2.24	27.7	0.33	1.53

**Table 3** Descriptive statistics of measured RASPs

Statistic	Sand (%)	Silt (%)	Clay (%)	BD (g/cm <sup>3</sup> )	PD (g/cm <sup>3</sup> )	MC (%)	OC (%)
Nos. of observations	20	20	20	20	20	20	20
Minimum	92.700	2.100	3.940	1.251	2.070	18.100	0.292
Maximum	93.466	2.590	6.460	1.789	2.760	27.700	0.361
1st quartile	92.972	2.248	4.643	1.414	2.300	21.990	0.318
Median	93.067	2.375	5.150	1.527	2.453	23.500	0.335
3rd quartile	93.230	2.425	5.713	1.664	2.663	25.200	0.343
Mean	93.080	2.343	5.187	1.540	2.456	23.512	0.330
Variance ( $n - 1$ )	0.044	0.020	0.522	0.026	0.049	5.702	0.000
Standard deviation ( $n - 1$ )	0.210	0.141	0.722	0.163	0.221	2.388	0.020

$$\text{EIR} = 2.87 (\text{SAND}) - 2.743 (\text{SILT}) - 0.732 (\text{CLAY}) - 254.43. \quad (9)$$

The details of predictive accuracy of the developed model are shown in Fig. 4. Analysis of results revealed that the EIR ranged from 0.779 to 5.89 cm/h with the average rate of 3.112 cm/h and the average measures IR was 3.11 cm/h. The  $R^2$  was 0.81. The RMSE, MSE, and SD were 0.623, 0.389, and 0.659 cm/h, respectively. The sand has the highest variable impact projection (VIPs) value = 1.173 followed by clay with 1.028. However, silt with VIP = 0.753 which is lower than (Wold (1995) and Wold et al. 2001 recommended threshold value of 0.8). They also categorised VIP in two groups [moderate influence if ( $0.8 < \text{VIP} < 1$ ), and highly influential ( $\text{VIP} > 1$ )]. Therefore, silt is not an important predictor.

### Development of predictive model for group B

The infiltration rate model was developed using the inputs (group A, and BD). The observed values of that the bulk density vary from the 1.37 to 1.789 gm/cm<sup>3</sup> with the average of 1.53 g/cm<sup>3</sup>. The established equation of estimated infiltration rate (EIR):

$$\text{EIR} = 2.349 (\text{SAND}) - 2.241 (\text{SILT}) - 0.598 (\text{CLAY}) - 2.546 (\text{BD}) - 203.289. \quad (10)$$

Analysis of results revealed that the EIR ranged from 1.001 to 4.3 cm/h with the mean value of 3.112 cm/h, against the average measures IR 3.112 cm/h.  $R^2$  was 0.876. The RMSE, MSE, and SD, were 0.505, 0.254, and 0.533 cm/h, respectively. The sand has the highest variable



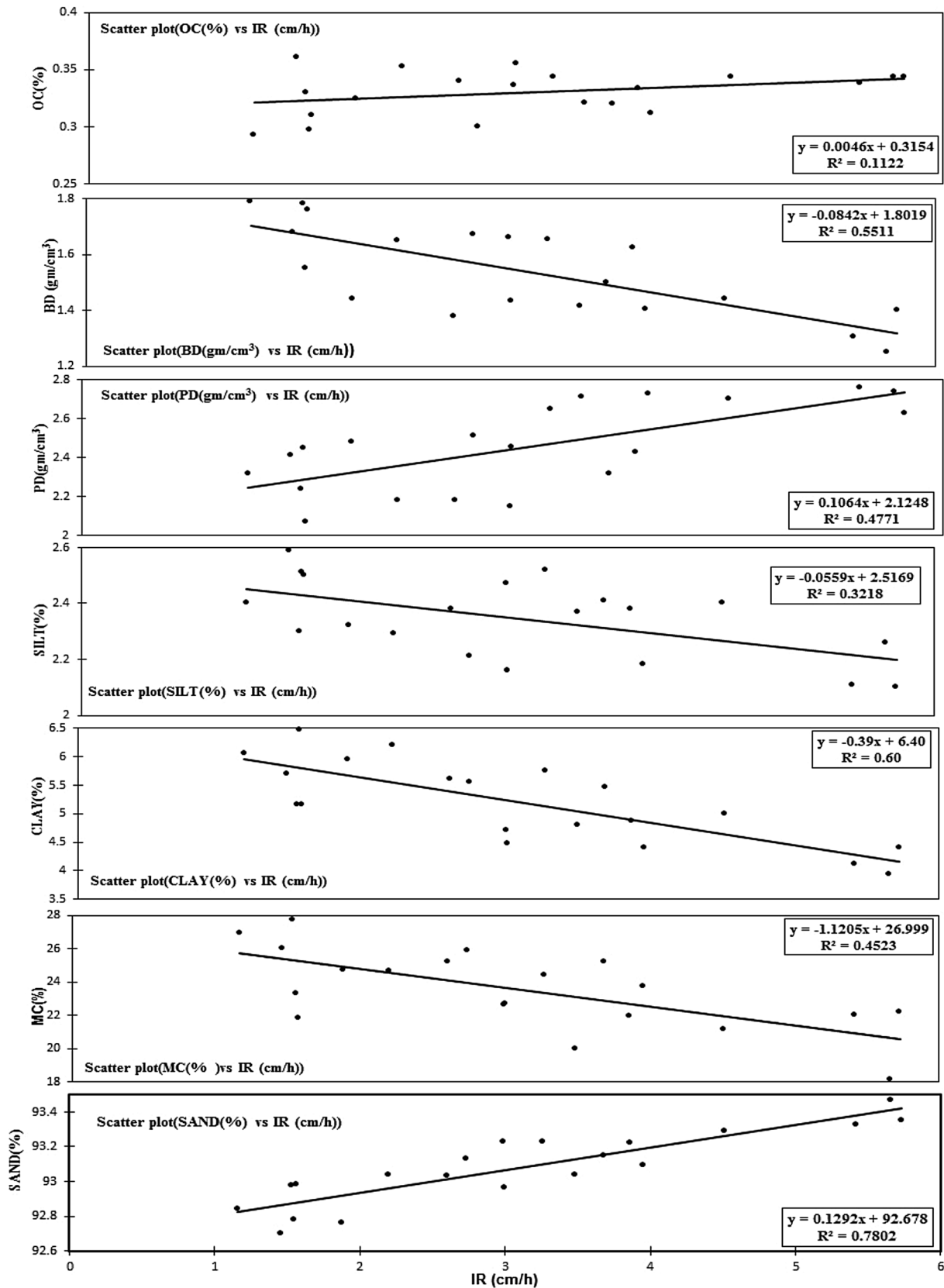
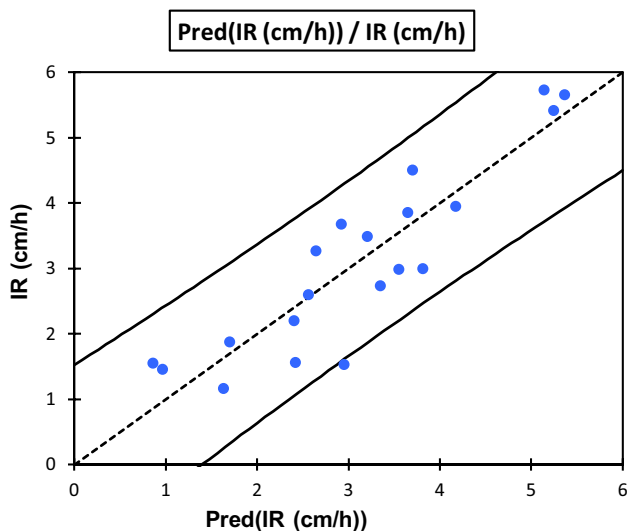


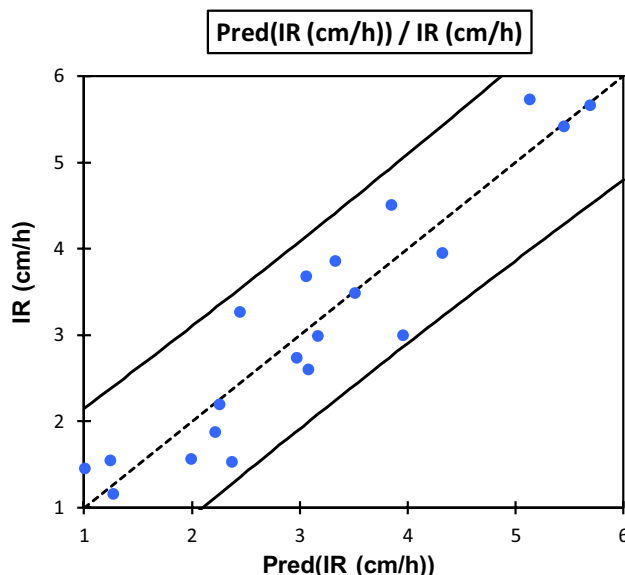
Fig. 3 Relation between studied soil properties and measured infiltration rate

**Table 4** Correlation matrix among RASPs

Variables	Sand (%)	Silt (%)	Clay (%)	BD (g/cm <sup>3</sup> )	PD (g/cm <sup>3</sup> )	MC (%)	OC (%)	IR (cm/h)
SAND (%)	1.000	-0.424	-0.691	-0.448	0.433	-0.584	0.275	0.883
SILT (%)	-0.424	1.000	0.571	0.525	-0.422	0.170	-0.088	-0.567
CLAY (%)	-0.691	0.571	1.000	0.554	-0.472	0.643	-0.312	-0.775
BD (g/cm <sup>3</sup> )	-0.448	0.525	0.554	1.000	-0.647	0.609	-0.287	-0.742
PD (g/cm <sup>3</sup> )	0.433	-0.422	-0.472	-0.647	1.000	-0.482	0.090	0.691
MC (%)	-0.584	0.170	0.643	0.609	-0.482	1.000	-0.208	-0.673
OC (%)	0.275	-0.088	-0.312	-0.287	0.090	-0.208	1.000	0.335
IR (cm/h)	0.883	-0.567	-0.775	-0.742	0.691	-0.673	0.335	1.000



**Fig. 4** Measured vs predicted IR using sand, silt, and clay



**Fig. 5** Measured vs predicted IR using sand, silt, clay, and BD

impact projection (VIPs) value = 1.173 followed by clay (VIP = 1.032), and BD (VIP = 0.989). The BD influences infiltration process moderately. However, the importance of silt was almost similar as in Eq. 9 with VIP = 0.756. Therefore, silt again not an important predictor.

With the consideration of BD, as additional input, improvement in model efficiency was observed as compared group A model which can be seen through the increased of *R*<sup>2</sup> and reduction in error statistics. Furthermore, Fig. 5 shows measured and predicted IR, which also support Eq. 10 has better productivity compared to Eq. 9.

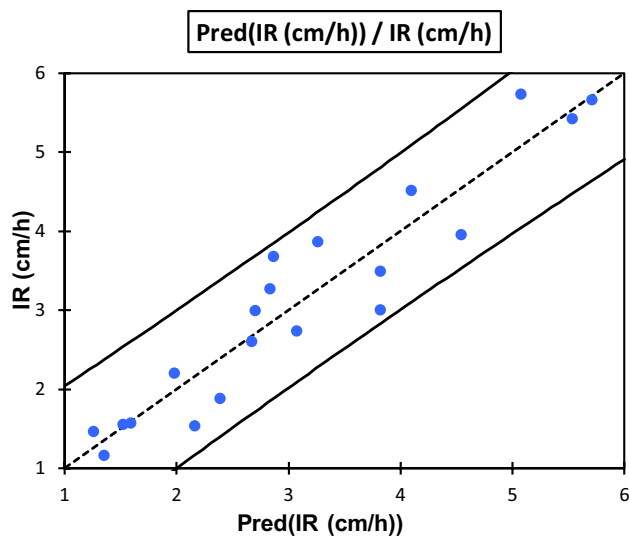
**Development of predictive model for group C**

A predictive model was developed using inputs (group B and PD). It was observed that the particle density varies from 2.07 to 2.74 g/cm<sup>3</sup> with the average value of 2.455 g/cm<sup>3</sup>. The established equation for the estimated infiltration rate (EIR):

$$EIR = 1.988 (SAND) - 1.896 (SILT) - 0.507 (CLAY) - 2.154 (BD) + 1.476 (PD) - 175.144. \tag{11}$$

Findings revealed that the EIR ranged from 1.25 to 5.53 cm/h with the mean value of 3.112 cm/h. The *R*<sup>2</sup> was 0.899. The RMSE, MSE, and SD were 0.456, 0.208, and 0.483 cm/h, respectively. The sand has the highest variable impact projection (VIPs) value = 1.19 followed by clay (VIP = 1.048), BD (VIP = 1.04), PD (VIP = 0.935). The PD influences infiltration process moderately. Results revealed that the importance of BD improved from moderate influence to high influence by the addition of PD as an input variable. However, the importance of silt was almost similar as in Eqs. 9, and 10 with VIP = 0.768. Therefore, silt again not an important predictor.

With the consideration of two additional inputs, namely, BD and PD, the modelling efficiency further enhanced compared to the previously developed models for the group A and B, which can be seen through the increased



**Fig. 6** Measured vs predicted IR using sand, silt, clay, BD, and PD

$R^2$  and reduction in error statistics. Figure 6 depicts the better correlation between measured vs predicted IR for group C inputs.

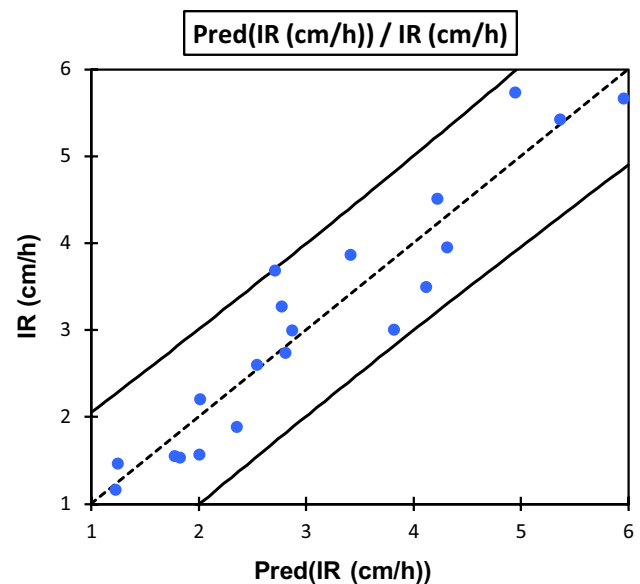
#### Development of predictive model for group D

A predictive model was developed using inputs (group C, and MC). Findings showed that the moisture content ranged (18.1–27.7%) with the average value of 23.512%. The developed equation for the EIR infiltration rate is given below:

$$\text{EIR} = 0.252 (\text{SAND}) - 0.541 (\text{SILT}) - 0.33 (\text{CLAY}) - 1.724 (\text{BD}) + 0.984 (\text{PD}) - 0.132 (\text{MC}) - 14.052 \quad (12)$$

The findings revealed that the EIR ranged from 1.22 to 5.95 cm/h with the mean value of 3.112 cm/h, against the mean measured IR as 3.116 cm/h.  $R^2$  was 0.897. The RMSE, MSE, and SD were 0.461, 0.212, and 0.487 cm/h, respectively. The sand has the highest variable influential projection (VIPs) value = 1.213 followed by clay (VIP = 1.064), BD (VIP = 1.01), PD (VIP = 0.948), and MC (VIP = 0.810). Results revealed that the importance of MC close to the threshold. Therefore, model predictively was almost similar to the previous case. However, the importance of silt was almost similar as in previous developed equations with VIP = 0.779. Therefore, silt again not an important predictor.

The consideration of moisture content (MC) does not improve predictability, resulted in almost similar performance compared to the previously developed models for the group C which can be seen through the similar statistical performance indices. According to Fig. 7 measured vs predicted IR using all selected RASPs except organic carbon



**Fig. 7** Measured vs predicted IR using sand, silt, clay, BD, PD, and MC

(OC) input showed almost similar correlation compared to developed equation 11 of (group A, group B, group C).

#### Development of predictive model for group E

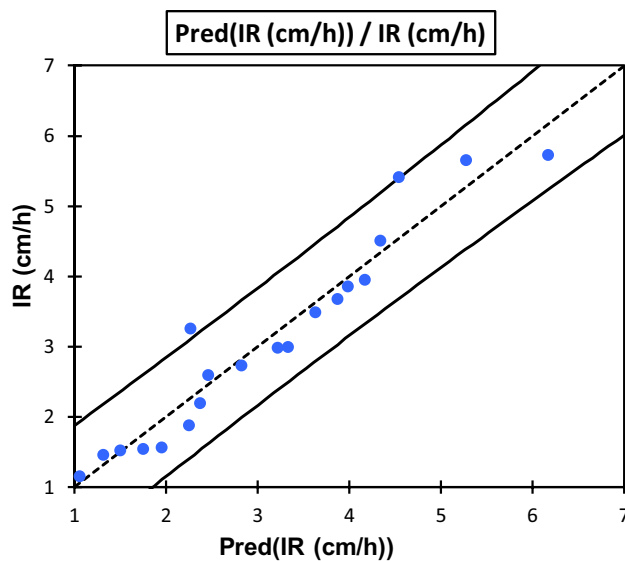
A predictive model was developed using the inputs of (group D, and OC). Findings resulted that the OC varies from the 0.29–0.36% with the average value of 0.329%. The developed equation for the EIR infiltration rate is given below:

$$\text{EIR} = 1.652 (\text{SAND}) - 1.576 (\text{SILT}) - 0.42 (\text{CLAY}) - 1.79 (\text{BD}) + 1.22 (\text{PD}) - 0.11 (\text{MC}) + 0.66 (\text{OC}) - 144.64 \quad (13)$$

From the analysis of results revealed that the estimated EIR ranged from 1.02 to 5.34 cm/h with the average value of 3.112 cm/h, and the mean observed IR was 3.112 cm/h.

The  $R^2$  was 0.927. The RMSE, MSE, SD, were 0.378, 0.212, 0.143 and 0.398 cm/h, respectively. The sand has highest variable impact projection (VIPs) value = 1.288 followed by clay (VIP = 1.129), BD (VIP = 1.082), PD (VIP = 1.008), MC (VIP = 0.827), OC (VIP = 0.980). Findings revealed that the OC moderately influence the model. Furthermore, the addition of OC improved VIP values of all the inputs except MC that remains close to the threshold. Therefore, model predictively was almost similar to the previous case. However, the importance of silt was reduces compared to previous cases with VIP = 0.658. Therefore, silt again not an important predictor.

We found that modelling efficiency enhanced due to the addition of organic carbon content (OC) as an input



**Fig. 8** Measured vs predicted IR using sand, silt, clay, BD, PD, MC, and OC

variable in comparison with recent analysis for the other groups, which can be seen through the increased  $R^2$ . Figure 8 measured vs predicted IR using all the RASPs as independent variables shows a better correlation compared to other developed Equations.

## Discussion

For the development of predictive models, it is essential to understand dependency among dependent and independent variables. In the present study, it was observed most of the selected soil properties have either highly positive or profoundly negative correlation with observed infiltration rate. The sand and clay content is the most crucial factor, which affects the soil infiltration rate. The bulk density (BD) and moisture content showed another possible input as these properties inversely related to infiltration characteristics (Pandey et al. 2009). The organic carbon (OC) content of the soil showed a positive correlation with infiltration characteristics of the study area. The findings of present investigation correlate with studies (Van de Genachte et al. 1996; Pandey et al. 2009; Dabral and Pandey 2016; Sarmadian and Taaghizadeh-Mehrjardi 2014; Rahmati 2017; Saleh et al. 2017) reported that easily measurable soil characteristics could successfully predict field measured infiltration characteristics.

From the analysis of all the developed PLSR equations, it was observed that the predictability of models improved (with the increased  $R^2$ , and decreased error indices) by increasing numbers of inputs. Rahmati (2017) reported similar conclusion in development of PTFs of cumulative

infiltration depth based on radially observed soil characteristics in Iran. The developed equation 9 showed the worst statically fitted value of the  $R^2$  and RMSE and SD. The developed equation 13 depicted best fitting statistic maximum value of the  $R^2$  and minimum value of error indices. Thus, Eq. 13 is the best amongst all the equations. It may be concluded that an increasing number of independent variables better productivity in the prediction of infiltration characteristics. However, silt and MC were not a powerful predictor for the study site. Infiltration characteristics of the soil are one of the most complex hydrological process dependent on various parameters; it may be possible by identifying any more suitable parameter its productivity further enhanced.

## Conclusions

The present study aims to develop predictive models of IR based on radially available soil properties (RASPs) as an alternative to measured IR, for the study area. Twenty sites soil samples were used for the establishing model (partial least square regression based). Analysis of results depicted that infiltration rate has a positive relationship with sand, particle density (PD) and organic carbon (OC). However, silt, clay, bulk density (BD), and moisture content (MC) have a negative correlation with the infiltration.

The variable influential projection (VIP) analysis revealed sand as a highly influential factor, while silt as a reluctant predictor of infiltration characteristics of the study site. In addition, the organic carbon, bulk density, and particle density were moderately influencing variable of infiltration characteristics of the soil. The highest improvement in models performance was observed by adding organic carbon as compared to other moderately influencing inputs. The moisture content has limited influence on infiltration characteristics of the study site.

The results revealed that PLSR is satisfactory to develop suitable models and linking soil properties to infiltration rate. It also identifies suitable predictor based on influential variable projection. The model developed with all the identified soil properties was a best-fitted with the maximum value of  $R^2$  and the least value of mean error statistics. In addition, the value  $R^2$  enhancing some inputs increases in the model, which confirm that infiltration characteristics of study can be predicted successfully with readily observed soil properties. The infiltration process modelled with readily observed soil properties, save the resources in the measurement of time-consuming infiltration characteristics and could be applied under limited data conditions.



## References

- Abdi H (2010) Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Comput Stat* 2:97–106
- Abudu S, Cui CL, King JP, Moreno J, Bawazir AS (2011) Modelling of daily pan evaporation using partial least squares regression. *Sci China Technol Sci* 54(1):163–174. <https://doi.org/10.1007/s1143.1-010-420.5-z>
- Arshad R, Sayad G, Mazlum M, Jafarnejadi A, Mohammadi Safarzadeh V (2010) Pedo transfer functions application to estimate the infiltration rate of the soil using neural network and linear regression methods. *J Crop Improv* 2(5):55–62
- Clausnitzer V, Hopmans JW, Starr JL (1998) Parameter uncertainty analysis of common infiltration models. *Soil Sci Soc Am J* 62:1477–1487
- Dabral PP, Pandey PK (2016) Models to estimate soil moisture retention limits and saturated hydraulic conductivity. *J Indian Water Resour Soc* 36(1):50–55
- Ekhmaj AI (2010) Predicting soil infiltration rate using artificial neural network. In: 2010 International Conference on Environmental engineering and applications (ICEEA). IEEE, pp 117–121
- Flint AL, Flint LE (2002) Particle Density. In: Dane JH, Topp CG (eds) *Methods of soil analysis: part 4 physical methods*, SSSA Book Ser. 5.4. SSSA, Madison, pp 229–240. <https://doi.org/10.2136/sssabookser5.4.c10>
- Franzluebbers AJ (2002) Water infiltration and soil structure related to organic matter and its stratification with depth. *Soil Tillage Res* 66:197–205. [https://doi.org/10.1016/S0167-1987\(02\)00027-2](https://doi.org/10.1016/S0167-1987(02)00027-2)
- Gee GW, Or D (2002) Particle-size analysis. *Methods of soil analysis. Part 4*:255–293
- Ghorbani-Dashtaki S, Araghi F, Mirlatifi M (2010) Soil cumulative infiltration prediction using PTFs in calcareous soil. *J Water Soil Conserve* 17(3):25–44
- Ghorbani-Dashtaki S, Homae M, Loiskandl W (2016) Towards using pedotransfer functions for estimating infiltration parameters. *Hydrol Sci J* 61(8):1477–1488. <https://doi.org/10.1080/0262667.2015.1031763>
- Grossman RB, Reinsch TG (2002) Bulk density and linear extensibility. In: Dane JH, Topp CG (eds) *Methods of soil analysis: part 4, physical methods* SSSA Book Ser. 5.4. SSSA, Madison, pp 201–228
- Huang M, Barbour SL, Elshorbagy A, Zettl JD, Si BC (2011) Infiltration and drainage processes in multi-layered coarse soils. *Can J Soil Sci* 91(2):169–183
- Kashi H, Emamgholizadeh S, Ghorbani H (2014) Estimation of soil infiltration and cation exchange capacity based on multiple regression, ANN (RBF, MLP), and ANFIS models. *Commun Soil Sci Plant Anal* 45(9):1195–1213
- Khatr KL, Smith RJ (2005) Evaluation of methods for determining infiltration parameters from irrigation advance data. *Irrig Drain* 54(4):467–482
- Lake HR, Akbarzadeh A, Mehrjardi RT (2009) Development of pedo transfer functions (PTFs) to predict soil physicochemical and hydrological characteristics in southern coastal zones of the Caspian Sea. *J Ecol Nat Environ* 1(7):160–172
- Mahdian M, Oskoee R, Kamali K, Angoshtari H, Kadkhodapoor M (2009) Developing Pedo Transfer Functions to predict infiltration rate in flood spreading stations of Iran. *Res J Environ Sci* 3(6):697–704
- Martens H, Martens M (2000) Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Qual Prefer* 11:5–16
- Mbagwu J (1994) Soil physical properties influencing the fitting parameters in Philip and Kostiakov infiltration models. International Centre for Theoretical Physics, Trieste
- Mirzaee S, Zolfaghari AA, Gorji M, Dyck M, Ghorbani-Dashtaki S (2014) Evaluation of infiltration models with different numbers of fitting parameters in different soil texture classes. *Arch Agron Soil Sci* 60(5):681–693. <https://doi.org/10.1080/03650340.2013.823477>
- Mousavi SB (2015) Employment of Pedo-transfer functions for predicting initial and final infiltration rates using horton model and soil readily available characteristics. *Indian J Agr Sci* 49(4):333–337. <https://doi.org/10.5958/0976-058X.2015.00060.8>
- Nelson D, Sommers LE (1982) Total carbon, organic carbon, and organic matter. In: Page AL (ed) *Methods of soil analysis. Part 2. Chemical and microbiological properties*, Agron Monogr 9.2. ASA, SSSA, Madison, pp 539–579. <https://doi.org/10.2134/agronmonogr9.2.2ed.c29>
- Osuji GE, Okon MA, Chukwuma MC, Nwarie II (2010) Infiltration characteristics of soils under selected land use practises in Owerri, Southeastern Nigeria. *World J Agric Sci* 6(3):322–326
- Pandey V, Pandey PK (2010) Spatial and temporal variability of soil moisture. *Int J Geosci* 1(2):87–98
- Pandey PK, Dabral PP, Kumar A, Kumar N (2009) Determination of physical, water retention and transmission properties of soil of paddy fields in the vicinity of Nirjuli (Arunachal Pradesh). *J Soil Water Conserv (India)* 8(4):3–10
- Pandey V, Pandey PK, Mahanta AP (2014) Calibration and performance verification of Hargreaves–Samani equation in a humid region. *Irrig Drain* 63:659–667
- Parchami-Araghi F, Mirlatifi SM, Dashtaki GS, Mahdian MH (2013) Point estimation of soil water infiltration process using artificial neural networks for some calcareous soils. *J Hydrol* 481:35–47
- Parhi PK (2014) Another look at Kostiakov, modified Kostiakov and revised modified Kostiakov infiltration models in water resources applications. *Int J Agric Sci* 4(3):138–142
- Perroux K, White I (1988) Designs for disc permeameters. *Soil Sci Soc Am J* 52(5):1205–1215
- Punamia BC, Jain AK, Jain AK (2005) *Soil mechanics and foundations* (16th thoroughly revised and enlarged edition). Laxmi Pub. House, New Delhi, p 8170087910. ISBN: 9788170087915
- Rahmati M (2017) Reliable and accurate point-based prediction of cumulative infiltration using soil readily available characteristics: a comparison between GMDH, ANN, and MLR. *J Hydrol* 551:81–91
- Rashidi M, Hajiaghahi A, Amel-Sadeghi M (2013) Prediction of soil infiltration rate based on sand content of soil. *Agric Eng Res J* 3(2):20–24
- Rawls WJ, Brakensiek DL, Miller N (1983) Green–Ampt infiltration parameters from soils data. *J Hydraul Eng* 109(1):62–70
- Saleh I, Kaviani A, Jafarian Z, Ahmadi R (2017) Estimation of final soil infiltration rate using fuzzy clustering algorithm, neuro fuzzy (ANFIS) and fuzzy inference system (FIS) (a case study: Behshahr Plain, Galougah, Mazandaran, Iran). *Appl Soil Res* 4(2):47–59
- Sarmadian F, Taaghizadeh-Mehrjardi R (2014) Estimation of the infiltration rate and deep percolation water using feed-forward neural networks in the Gorgan province. *Eur J Soil Sci* 3(1):1–6
- Saxton KE, Rawls WJ (2006) Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci Soc Am J* 70:1569–1578. <https://doi.org/10.2136/sssaj2005.0117>
- Saxton KE, Rawls WJ, Romberger JS, Papendick RI (1986) Estimating generalized soil water characteristics from texture. *Trans ASAE* 50:1031–1035
- Singer MJ, Munns DN (1999) *Soils: an introduction*, 4th edn. Prentice Hall, Upper Saddle River

- Tuffour HO, Bonsu M (2015) Application of Green–Ampt equation to infiltration with soil particle phase. *Int J Sci Res Agric Sci* 2(4):076–088
- U.S. Department of Agriculture (USDA) (1951) Soil survey manual, USDA agricultural handbook. No. 18. U.S. Government Printing Office, Washington, DC
- Van de Genachte G, Mallants D, Ramos J, Deckers JA, Feyen J (1996) Estimating infiltration parameters from basic soil properties. *Hydrol Process* 10:687–701
- Wold S (1995) PLS for multivariate linear modeling. In: van der Waterbeemd H (ed) *Chemometric methods in molecular design: methods and principles in medicinal chemistry*. Verlag-Chemie, Weinheim, pp 195–218
- Wold S, Sjostrom M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. *Chemometrics Intell Lab Syst* 58:109–130
- Wosten JHM, Van Genuchten MTH (1988) Using texture and other soil properties to predict the unsaturated soil hydraulic functions. *Soil Sci Soc Am J* 52:1762–1770. <https://doi.org/10.2136/sssaj1988.03615995005200060045x>