Surface runoff from arable land - a homogenized data base of 726 rainfall simulation experiments

Compiled by

Seibert, S.P.[1, 2], Auerswald, K.[2*], Fiener, P.[3,4], Disse, M.[5], Martin, W.[6], Haider, J.[7], Michael, A.[8] and Gerlinger, K.[9]

*Corresponding author: Tel.: +49 (0)8161-713965, E-mail address: auerswald@wzw.tum.de

[1] Present address: Lehrstuhl für Hydrologie, Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany (s.seibert@bv.tum.de)
[2] Lehrstuhl für Grünlandlehre, Technische Universität München, Alte Akademie 12, 85350 Freising-Weihenstephan, Germany (auerswald@wzw.tum.de)
[3] Geographisches Institut, Universität zu Köln, Albertus Magnus Platz, 50923 Köln, Germany (peter.fiener@uni-koeln.de)
[4] Present address: Indo-German Centre of Sustainability, Indian Institute of Technology Madras, Chennai 600036, India
[5] Lehrstuhl für Wasserwirtschaft und Ressourcenschutz, Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany (markus.disse@unibw.de)
[6] Present address: Bayerisches Landesamt für Umwelt, Bürgermeister-Ulrich-Straße 160, 86179 Augsburg, Germany (Walter.Martin@lfu.bayern.de)
[7] Present address: KfW Entwicklungsbank, Palmengartenstraße 5-9, 60325 Frankfurt, Germany (Josef.Haider@kfw.de)
[8] Technische Universität Bergakademie Freiberg, Fachgebiet Boden- und Gewässerschutz, Agricolastr. 22, 09599 Freiberg, Germany
[9] Hydron Ingenieurgesellschaft für Umwelt und Wasserwirtschaft mbH, Haid-und-Neu-Str. 7, 76131 Karlsruhe, Germany (kai.gerlinger@hydron-gmbh.de)

Abstract

A data base of rainfall simulations was compiled comprising the experimental work of Martin (1988), Haider (1994), Gerlinger (1997), Schröder and Auerswald (2000) and Michael (2001). The data base covers 726 rainfall simulations (24384 runoff measurements) which were carried out on 209 arable plots in Germany under a wide range of different conditions. The data base is freely available and may be used as an extensive source to examine specific questions. Fiener *et al.* (2011) provide detailed information on the origin of the individual data sets, their compilation, the range of the data base and on the quality of the included data. This supplement provides meta data on the structure of the relational data base, on its identifiers SEQ, RUN and SOIL, on the coding of missing and flagged data (see below) as well as a brief introduction of the available files and variables.

Introduction

We compiled and harmonized experimental conditions and hydrographs from five different research groups, which used artificial rainfall from different rainfall simulators based on Veejet nozzles (Figure 1) to study infiltration and erosion processes on arable land in a relational data base. The individual data sets (FB, FS, SY, WB and WS) originate from different regions in Germany (Figure 2).

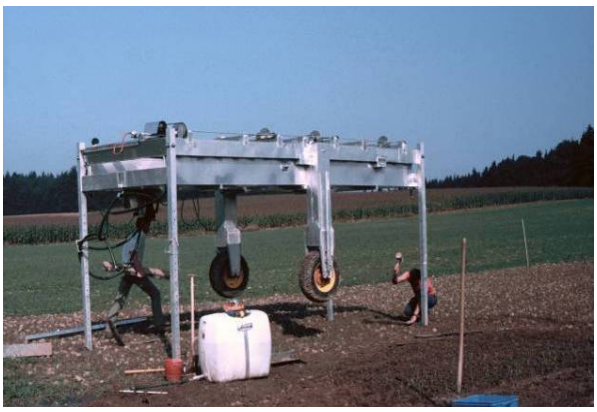

Figure 1: Top left: The 'Swanson' rainfall simulator was used in the WS data set. **Top right:** The 'Kainz Regner' was used in the FS and SY data sets. **Bottom right**: Gerlinger's (1997) rainfall simulator from the WB data set which in a slightly different set-up was also used for the FB data set (Michael 2001).

Surface runoff was measured as accumulated runoff during the rainfall simulation experiments. These measurements, which included also the measurement of time, will be called *runoff measurements*. They resulted from 726 simulated rain events, which we call *runs*. One to three runs were applied to the same plot within less than 48 h. Such a combination of runs on the same plot will be called *sequence* (n = 370). The runs within a sequence are named according to the time span to the preceding simulated rain. Sequences start with a *dry run* (n = 336), which is the first simulation on a plot or a simulation following a preceding one after a time span much larger than 48 hours. Although this denotation is commonly used, it is misleading, because a dry run may be on a soil, which is wet due to natural precipitation. A run following more than 12 hours and less than 48 hours after the

preceding run is a *wet run* (n = 126). Again this term is conventionally used but it has to be noted that it ignores that the most consistent difference compared to a dry run is not the difference in soil moisture but the existence of erosion features like surface sealing or rills resulting from a preceding extraordinary large erosive rain. A run following after less than 12 hours was called *very wet run* (n = 264) and can be immediately preceded by either a dry run or a wet run. In general, a sequence can consist of a dry run only or any combination starting with a dry run, except for some rare cases (n = 34) where the dry run was discarded due to equipment failure. In several cases of bare-fallow plots rains followed on the same plot with a rain spell of several months and soil tillage in between. Hence the number of rained *plots* (n = 209) is smaller than the number of sequences.

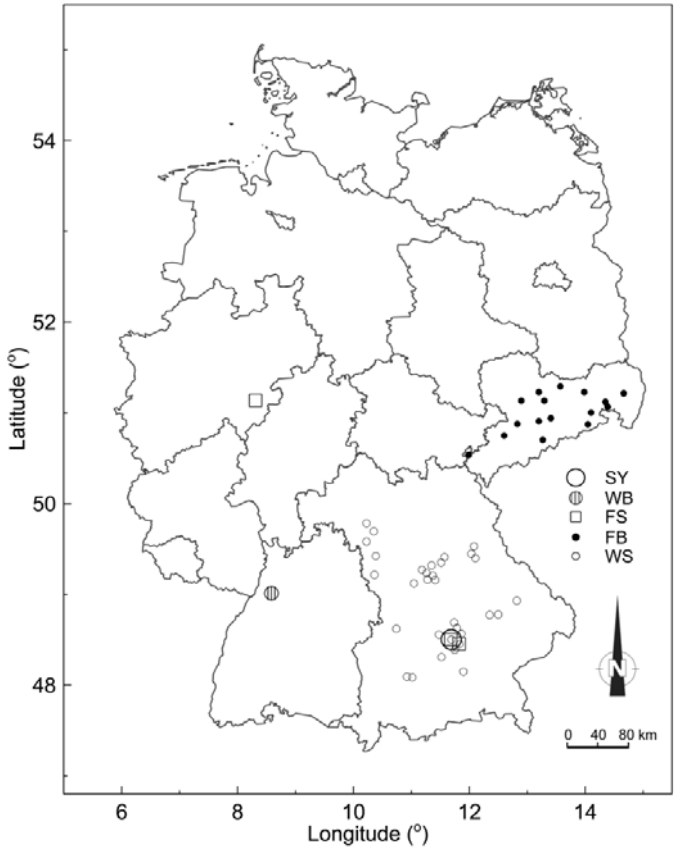

Figure 2: Map of all experimental sites for the data sets FS, FB, SY, WB and WS. Note that the locations of the data set WS denote the origin of the top soils, which where all rained at one location (near SY) after long-term bare fallow. Lines show boundaries of the German federal states.

Coding and structure of the data base

The data base comprehends runoff measurements and several data on plot, land use and experimental conditions. In total it consists of five relational tables that are linked by three different identifiers (Figure 3). The five tables are:

(i) ***REL-TAB***, which is the central relational linking table. It contains the identifiers of all sequences (SEQ), dry, wet and very wet runs (RUN) and plots (SOIL) (details below). *REL-TAB* provides the relational data model. It links all runs to sequences and provides the corresponding identifiers RUN and SOIL that link to four other tables outside of *REL-TAB*. These are:

(ii) ***RUN-TEMP***, which contains temporally resolved runoff measurements (linked via RUN)

(iii) ***RUN-PROP***, which contains run-specific data on rain and runoff characteristics like rain intensity or duration (linked via RUN).

(iv) ***LAND-PROP***, which contains land management information like surface cover, crop or tillage (linked via RUN) and

(v) ***SOIL-PROP***, which stores soil and plot properties (linked via SOIL).

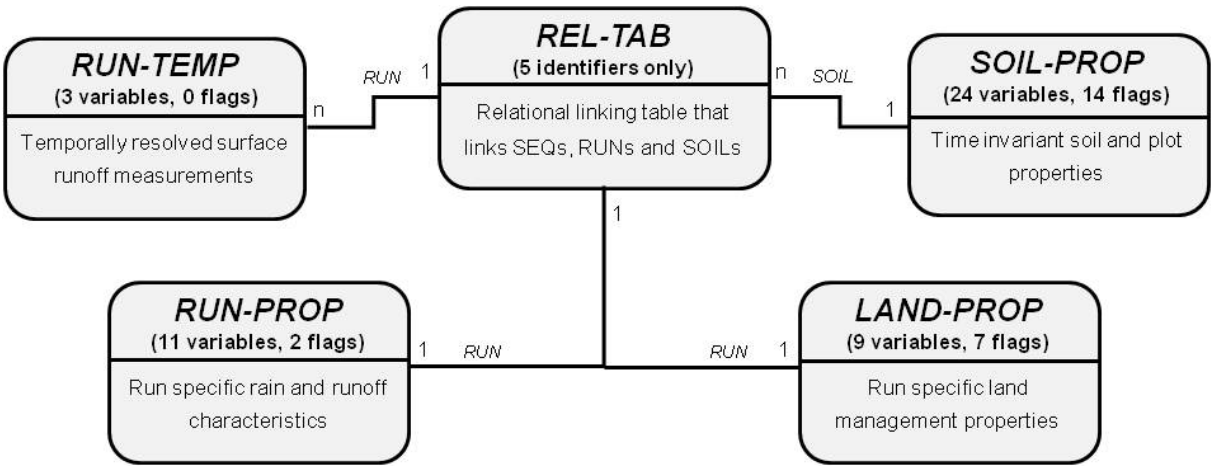

Figure 3: Relational data structure: *REL-TAB* is the central linkage table. It contains the unique identifiers of all sequences (SEQ) and links them to the identifiers of the dry, wet and very wet rainfall experiments (RUN) and the corresponding plot identifiers (SOIL). The respective identifiers used to link between tables are given on the connecting lines. Additionally, the notations 1 and n on the lines represent the relationship between the tables. E.g. the n:1 (many-to-one) relationship between *REL-TAB* and *SOIL-PROP* indicates that *REL-TAB* may contain several identical SOILs while SOIL is unique in *SOIL-PROP*. The number of variables and flags within each table are provided in parentheses underneath the table name. Flags indicate the data quality.

Rainfall und runoff properties (in *RUN-PROP*), plot and soil properties (in *SOIL-PROP*) and land management properties (in *LAND-PROP*), were stored separately to avoid redundancies (details on the tables are provided below). The identifiers RUN, SOIL and SEQ are unique numeric codes. They support a flexible handling and querying of the data base.

RUN codes the individual rainfall experiments. It is a 5-digit number. The first digit denotes the data set (1=FB, 2=FS, 3=SY, 4=WB and values ≥ 5 correspond to runs from WS), the 5th digit in RUN codes the run type (1=dry run, 2=wet or 3=very wet run). Digit two, three and four of RUN code the individual rainfall experiments within the different data sets. They lack superior importance.

SOIL is a 4-digit number. The first digit again codes the dataset (1=FB, 2=FS, 3=SY, 4=WB and 5=WS). The second digit was intentionally left blank, the third and fourth digit represent the experimental plot within the dataset.

SEQ, the sequence identifier, links runs that were conducted on the same plot within less than 48 h. Thus, one SEQ refers to one to three RUNs. SEQ is a 7-digit number with a more complex structure. SEQ contains information about the dataset, the year of the experiment, the experimental period within the year, SOIL and the RUNs within the sequence (Table 1). For example SEQ=1930014 codes a sequence of rainfall experiments from the FB data set carried out in year 1993 on plot 1001. The last digit indicates that the sequence starts with a dry run and that there must be at least two runs in this sequence. Note that a sequence can consist of a either a dry, wet or very wet run only, or any combination of these.

Table 1: Coding of the seven digits comprising the sequence identifier (SEQ), which links runs and sequences.

Digit in SEQ							Meaning
1	2	3	4	5	6	7	complete identifier SEQ, which is a unique 7-digit number
1							codes the data set: 1=FB, 2=FS, 3=SY, 4=WB, 5=WS
	2	3					codes year of the experiment: 19XX
			4				codes the number of the experimental period within a year. It is 0 in all datasets except WS.
1				5	6		code the SOIL: digit 1 times 1000 plus digit 5 multiplied by 10 plus digit 6 equals SOIL.
						7	codes the run-type(s) that form the sequence. A SEQ can consist of a: 1 = dry run only, 2 = wet run only*, 3 = very wet run only*, 4 = sequence of a dry run + wet and/or very wet run, 5 = wet and very wet run only*

* the dry run was discarded due to any reason

Tables, variables and flags

The tables of the data base are available as tab separated ascii files. They contain identifiers, variables and flags. The identifiers link the tables *REL-TAB, RUN-PROP, RUN-TEMP, LAND-PROP* and *SOIL-PROP* (Figure 3). The variables in *REL-TAB, RUN-PROP, RUN-TEMP, LAND-PROP* and *SOIL-PROP* are introduced in tables 2, 3, 4, 5 and 6. More information on theses variables regarding their distribution within the range, which is covered by each variable, regarding their quality and regarding their collinearity can be found in Fiener *et al.* (submitted). Flags indicate whether a variable was measured or estimated during the compilation of the data base. They were added only in those cases, where both possibilities exist. They are named "flag_" followed by the name of the variable, to which they apply. Flag "0" indicates a measured value while "1" indicates an estimate. Missing values, which could not be harmonized or gap-filled, are given as "NA".

Table 2: Variables in the relational table ***REL-TAB*** that links the runs within each sequence. Note, SEQ and RUN are unique in *REL-TAB*, while the same SOIL may appear in several SEQ.

Variable *(Abbreviation)*	**Description**
SEQ	It provides the identifier SEQ for a sequence of rainfall simulations that were conducted on the same plot within less than 48 h.
RUN_D	It provides the identifier RUN of the dry run within a sequence. NA codes a missing dry run.
RUN_W	It provides the identifier RUN of the wet run within a sequence. NA codes a missing wet run.
RUN_vW	It provides the identifier RUN of the very wet run within a sequence. NA codes a missing very wet run.
SOIL	It provides the identifier SOIL that links to soil and plot properties.

Table 3: Variables in the relational table ***RUN-TEMP*** that provides the temporally resolved runoff measurements of all runs.

Variable *(Abbreviation)*	**Description**	**Unit**	**Range**
RUN	Identifier RUN that links to experimental conditions and land management properties	-	-
SEQ	Identifier SEQ that links to RUN and SOIL	-	-
time	Time after start of the rainfall simulation	s	0-6180
q	discharge	$l\ s^{-1}$	0-0.5
seq_time	Consecutive time of rain within a sequence starting from the beginning of the dry run. The break between dry/ wet/ very wet run remained unconsidered, also the duration of the afterflow. Thus, the times of the individual runs of the sequence were seamlessly joined. Note seq_time was set to NA if a preceding run of a sequence had been deleted.	s	0-9660

Table 4: Variables in the relational table ***RUN-PROP*** that provides the experimental conditions of the individual runs.

Variable (Abbreviation)	**Description**	**Unit**	**Range**
RUN	Identifier RUN, which links the tables *REL-TAB*, *RUN-TEMP*, *LAND-PROP* and *RUN-PROP*	-	-
SOIL	Identifier SOIL which links *REL-TAB* and *SOIL-PROP*	-	-
Date	Date of the rainfall simulation	DD.MM. YYYY	1985-1995
DWvW	Indicates the type of the run (1=Dry, 2=Wet, 3=very Wet)	--	1-3
eP	Specific kinetic energy	J m^{-2} mm^{-1}	12-20
duration	Duration of a rainfall simulation	s	590-6180
p	Rain intensity	mm h^{-1}	29-99
PTot	Total rainfall	mm	11-99
QTot	Total runoff without afterflow	mm	0-59
tP	Time to ponding	s	4-1074
tR	Time to runoff	s	6-3588
Theta_plough	Volumetric antecedent soil moisture in the plough layer (0-30 cm depth)	%	8-44
Theta_surf	Volumetric antecedent surface soil moisture (0-3 cm depth)	%	2-26
flag_tP	Flag of *tP* (0 = measured, 1 = estimated)	--	0 or 1
flag_tR	Flag of *tR* (0 = measured, 1 = estimated)	--	0 or 1
run_remarks	Additional remarks on the individual rainfall simulation. If available these were taken from the experimental protocols.	--	--

Table 5: Variables in the relational table ***LAND-PROP*** that provides the land management conditions of the individual runs.

Variable (Abbreviation)	**Description**	**Unit**	**Range**
RUN	Identifier RUN, which links the tables *REL-TAB*, *RUN-TEMP*, *LAND-PROP* and *RUN-PROP*	--	--
TsT	Time since tillage	d	0.04-227
CovTot	Total surface cover (sum of *CovVeg*, *CovRes* and *CovStone*)	%	0-93
CovVeg	Cover by vegetation	%	0-90
CovRes	Cover by residues not covered by vegetation	%	0-18
CovStone	Cover by stones not covered by vegetation or residues	%	0-35
CovResTot	Total residue cover including also those residues which are covered by vegetation	%	0-20
CovStoneTot	Total stone cover including also those stones which are covered by vegetation or residues	%	0-38
Crop	Crop type	--	bare fallow (=seed bed), winter wheat, winter barely, rape, summer rye, winter rye, sugar beet, maize
flag_TsT	Flag of *TsT* (0 = measured, 1 = estimated)	--	0 or 1
flag_CovTot	Flag of *CovTot* (0 = measured, 1 = estimated)	--	0 or 1
flag_CovVeg	Flag of *CovVeg* (0 = measured, 1 = estimated)	--	0 or 1
flag_CovRes	Flag of *CovRes* (0 = measured, 1 = estimated)	--	0 or 1
flag_CovStone	Flag of *CovStone* (0 = measured, 1 = estimated)	--	0 or 1
flag_CovResTot	Flag of *CovResTot* (0 = measured, 1 = estimated)	--	0 or 1
flag_CovStoneTot	Flag of *CovStoneTot* (0 = measured, 1 = estimated)	--	0 or 1
crop_remarks	Additional remarks on the land management	--	--

Table 6: Variables in the relational table ***SOIL-PROP:*** Properties of the individual soils and plots

Variable (Abbreviation)	**Description**	**Unit**	**Range**
SOIL	Identifier SOIL, which links *REL-TAB* and *SOIL-PROP*	--	--
Length	Length of the simulation plot	m	4-22
width	Width of the simulation plot	m	1-2
slope	Slope of the simulation plot	%	1.6-23.6
ClTot	Total clay content (0.004-2 µm) in bulk soil (w/w)	%	2-61
SiTot	Total silt content (2-63 µm) in bulk soil (w/w)	%	6-86
SaTot	Total sand content (63-2000 µm) in bulk soil (w/w)	%	2-87
vfSi	Very fine silt (2-6.3 µm) in bulk soil (w/w)	%	0-27
fSi	Fine silt (6.3-20 µm) in bulk soil (w/w)	%	1-33
mSi	Medium silt (20-36 µm) in bulk soil (w/w)	%	1-31
cSi	Coarse silt (36-63 µm) in bulk soil (w/w)	%	1-35
vfSa	Very fine sand (63-100 µm) in bulk soil (w/w)	%	0-19
fSa	Fine sand (100-200 µm) in bulk soil (w/w)	%	0-49
mSa	Medium sand (200-630 µm) in bulk soil (w/w)	%	0.4-61
cSa	Coarse sand (630-2000 µm) in bulk soil (w/w)	%	0-35
vfSt	Very fine stones (2-6.3 mm) in bulk soil (w/w)	%	0-31
fSt	Fine stones (6.3-20 mm) in bulk soil (w/w)	%	0-16
mSt	Medium stones (20-63 mm) in bulk soil (w/w)	%	0-16
cSt	Coarse stones (63-200 mm) in bulk soil (w/w)	%	--
Skeleton	Total sum of stones (2-200 mm) in bulk soil (w/w)	%	0-63
dg	Geometric mean particle diameter of bulk soil	µm	1-737
Sg	Standard deviation of the geometric mean particle diameter of bulk soil	µm	5-44
BD	Air-dry bulk density	kg m^{-3}	1070-1750
Corg	Soil organic carbon content in fine earth fraction	%	0.5-3.5
pH	pH	--	4.5-7.5
flag_vfSi	Flag of *vfSi* (0 = measured, 1 = estimated)	--	0 or 1
flag_fSi	Flag of *fSi* (0 = measured, 1 = estimated)	--	0 or 1
flag_mSi	Flag of *mSi* (0 = measured, 1 = estimated)	--	0 or 1
flag_cSi	Flag of *cSi* (0 = measured, 1 = estimated)	--	0 or 1
flag_vfSa	Flag of *vfSa* (0 = measured, 1 = estimated)	--	0 or 1
flag_fSa	Flag of *fSa* (0 = measured, 1 = estimated)	--	0 or 1
flag_mSa	Flag of *mSa* (0 = measured, 1 = estimated)	--	0 or 1
flag_cSa	Flag of *cSa* (0 = measured, 1 = estimated)	--	0 or 1
flag_vfSt	Flag of *vfSt* (0 = measured, 1 = estimated)	--	0 or 1
flag_fSt	Flag of *fSt* (0 = measured, 1 = estimated)	--	0 or 1
flag_mSt	Flag of *mSt* (0 = measured, 1 = estimated)	--	0 or 1
flag_cSt	Flag of *cSt* (0 = measured, 1 = estimated)	--	0 or 1
flag_Skel	Flag of *Skeleton* (0 = measured, 1 = estimated)	--	0 or 1
flag_BD	Flag of *BD* (0 = measured, 1 = estimated)	--	0 or 1

References

Fiener, P. Seibert, S., Auerswald, K. (2011). A compilation and meta-analysis of rainfall simulation data on arable soils. Journal of Hydrology. doi:10.1016/j.jhydrol.2011.08.034.

Gerlinger, K., 1997. Erosionsprozesse auf Lössböden: Experimente und Modellierung. PhD thesis, Universität Karlsruhe (TH), Karlsruhe, Germany.

Haider, J., 1994. Herbizide in Oberflächenabfluß und Bodenabtrag - Feldversuche mit simulierten Regen. PhD thesis, TU München, Freising-Weihenstephan, Germany.

Martin, W., 1988. Die Erodierbarkeit von Böden unter simulierten und natürlichen Regen und ihre Abhängigkeit von Bodeneigenschaften. PhD thesis, TU München, Freising-Weihenstephan, Germany.

Michael, A., 2001. Anwendung des physikalisch begründeten Erosionsprognosemodells EROSION 2D/3D. Empirische Ansätze zur Ableitung der Modellparameter. PhD thesis, Technische Universität Bergakademie, Freiberg, Germany.

Schröder, R., Auerswald, K., 2000. Modellierung des Jahresgangs der verschlämmungsinduzierten Abflussbildung in kleinen landwirtschaftlich genutzten Einzugsgebieten. Z. Kulturtechnik Landentw. 41**,** 167-172.