IMPLEMENTATION OF LEARN-ALGORITHMS TO IMPROVE AN EXPERT-SYTEM FOR THE DETERMINATION OF CONTROL STRATEGIES IN UDS

Khelil A.*, Heinemann A.**, Müller D.**

* Institut of Water Resources, University Hannover, Appelstr. 9A, 3000 Hannover 1, Germany

** Institut of Computer Sciences, University Hannover, Welfengarten 1, 3000 Hannover 1, Germany

KEYWORDS

RTC of UDS, Learning Algorithms, Classification Algorithms, Expert System (resp. Knowledge Based System), Evaluation of Control Strategies in UDS

ABSTRACT

The requirements which UDS are submitted to in respect of pollutants emission are ever increasing. Alone or in combination with sanitation/restructuration measures, RTC becomes ever more widespread in Europe. The automatic determination of an reliable online control strategy constitutes however a difficult problem, whose methodological approach is still object of research.

In our configuration, a strategy is determined by a knowledge based system (so called expert system) and is proposed for validation to the operator. The propagation/acceptation of such a system is hindered due to the lack of straightforward methods to develop a knowledge base. In Urban Hydrology, the situation is made even worse, because experience/knowledge about the true behaviour of UDS under loading is generally not available or if existing, difficult to put into the required form. This situation may however completly change, if efficient learn-algorithms exist, by means of which a gradual and <u>automatic refinement of a first implemented knowledge base</u> is made possible through simulation.

In our study, two learn-modules (one based on the use of so called meta-rules, the second one based on classification algorithms) have been implemented into a knowledge based system, which itself is connected to a hydraulic simulation program. For an artificial UDS, long term simulations have examplarily been performed. The modifications in decision making are traced back and discussed for both learn-processes.

REFERENCES

BOCK H., "Automatische Klassifikation", Studia Mathematica Band XXIV, ISBN 3-525-40130-2, edited by Vandenhoek & Ruprecht in Göttingen, 1974

FUCHS L., MÜLLER D., NEUMANN A., "Learning Production Systems for the Control of UDS", 1st IAWPRC International Symposium on Systems Analysis in Water Quality Management, London, 1987

KHELIL, "Steuerung eines Mischsystems zur Verbesserung der Gewässergüte und zur Verminderung der Betriebskosten -Wissenschaftliche Erarbeitung von Steuerungskonzepten-", final Report of Research Project 02-WA86470, BMFT-Germany, 1990

KHELIL, "Development of a Control Strategy to Reduce Combined Sewer Discharges of a UDS, Case of Bremen-Germany", WATERMATEX '91, 2nd International Conference on Systems Analysis in Water Quality Management, Durham, USA, 1991, Proceedings edited in Pergamon Press

SIXTH INTERNATIONAL CONFERENCE ON URBAN STORM DRAINAGE

National Water Research Institute
P.O. Box 5050
Burlington, Ontario, Canada
L7R 4A6

July 7, 1992

Dear Dr. Khelil:

We are pleased to inform you that your paper (Control Number **115**), entitled:

"Implementation of Learn-Algorithms to Improve an Expert-System for the Determination of Control Strategies in UDS"

---

has been accepted for presentation in a Regular Session.

Please fill out and return to me the attached **AUTHOR'S COMMITMENT FORM** at your earliest convenience. Your paper should be completed following the enclosed TYPING FORMAT and sent so that it will arrive in Canada by January 1, 1993. Late papers may not be included in the Proceedings (no supplements will be printed).

You will receive a conference brochure with some information about a preliminary program, registration, hotels, and travel in October, 1992.

If you have submitted paper abstracts on behalf of your colleagues, please make sure that all co-authors receive copies of this letter with enclosures. Please refer to notes below and the attachments for further information and requirements.

Sincerely,
Jiri Marsalek

Jiri Marsalek
Chairman, ICUSD '93

NOTES:

1. All participants are expected to pay their own travel expenses and registration fee.

2. If you need a formal letter (other than this letter) to obtain travel approval from your superiors, please write us with the necessary details and we will furnish such a letter.

3. As noted in the original announcements, all authors are required to present their papers at the conference. If for some reason you are unable to do this, please notify the Conference Chairman as soon as possible.

4. Reviewers' comments received and attached: (No) Yes

Telephone: (416) 336-4899 FAX: (416) 336-4989

LEARNING ALGORITHMS IN A RULE BASED SYSTEM FOR CONTROL OF UDS

Khelil A.*, Heinemann A.**, Müller D.**

* Ingenieurgesellschaft für Stadthydrologie mbH,
Vahrenwalder Str. 7, 3000 Hannover 1, FRG
** Institut of Computer Sciences, University Hannover,
Welfengarten 1, 3000 Hannover 1, FRG

KEYWORDS
RTC of **UDS**, Learn Algorithms, automatic Classification, Expert System (resp. Knowledge Based System), Evaluation of Control Strategies in **UDS**

I. Introduction

To improve the performances of **UDS**, **RTC** in combination with other measures is becoming widespread. The determination of set points in the control elements is difficult during a rainfall event, even if experienced people are on command. A rule based system (**RBS**) could facilitate the decision making at least by delivering its recommendations, but the acceptation of such a control configuration is hampered by uncertainties. For a given **UDS**, reliable knowledge about its behaviour is often not directly available; and if existing, it must be put into the required form of a consistent and complete set of rules. If efficient learning algorithms could be implemented on line, they would insure a gradual and automatic refinement of the initial control rules. Several learning modules were developed in Hannover and tested in off line simulations. This paper briefly present the basic principles of learning and describe a study case.

II. Short description of a learning cycle

Fig. 1 shows the sequence of treatments in a learning **RBS**. It is differentiated between knowledge and meta-knowledge. Knowledge data is stored in files and programmes and copes with the determination of a control strategy at a given time step. Meta-knowledge is stored in files and programmes and describes the "learning path".
While learning, the **RBS** constantly evaluates the **UDS** performances and modify the control rules if, according to the meta-knowledge, improvement is expected. Evaluation criteria are:

1. occurrence, amount and frequency of flooding
2. occurrence, amount and frequency of **CSO´s**
3. occurrence, duration and frequency of pressure flow in the pipes
4. energy consumption of the control gages or other security requirements.

A learning cycle consists of three tasks:

1. detection and collection of situations, where a questionable control strategy is determined
2. Modification of control set points at the relevant control elements
3. construction of new control rules

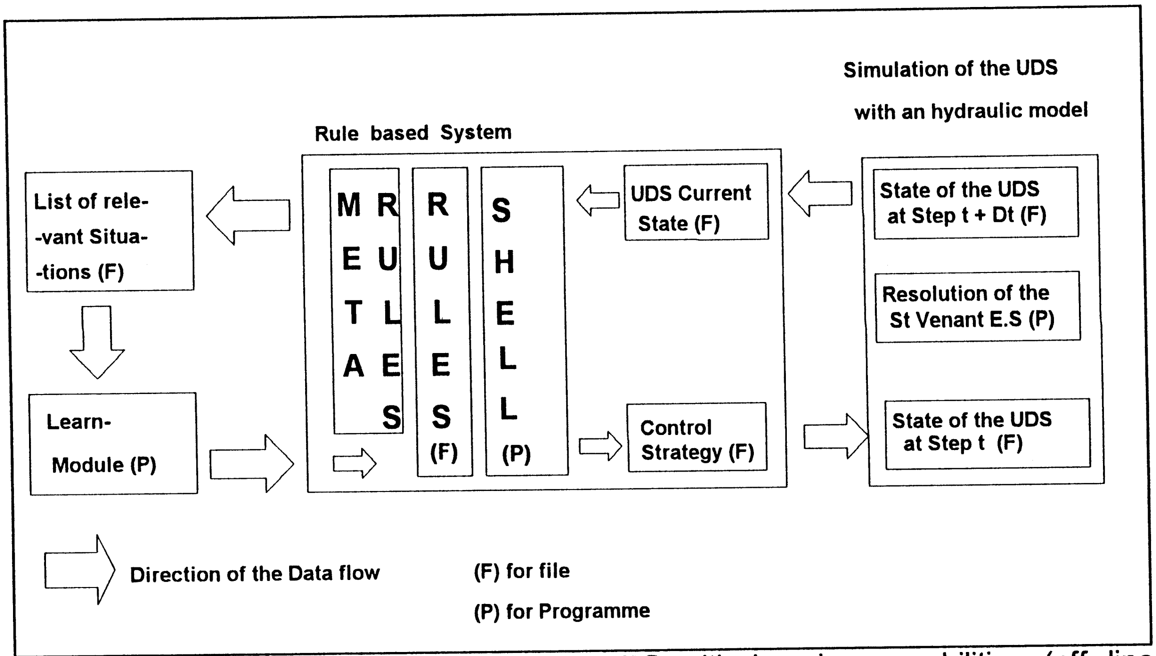

Fig. 1: Data Treatment Sequences in a **RBS** with learning capabilities (off line simulation)

At every decision step, the current situation in the **UDS** is evaluated by meta-rules. They check up, whether disturbances take place, which are possibly caused by inappropriate control decisions. If need be, the values of relevant state variables are stored in a so-called "situation-file". Parallely, the control set points are analysed, modified and stored in the same file. The construction of new control rules is only achieved, after many situations have been collected. This prevents from "learning in the false direction".
First modules were developed by Neumann, Fuchs and Müller (1987). The present one has following characteristics, which ensure a greater independence between the treatment steps.

- Meta-rules only intervene in the detection of questionable situations
- Modifications of the set points rely on information separately stored in a file "logic of control".
- For generation of new rules, the canal situations are classified by means of an algorithm first described by Wishart (see Bock, 1974), which is controlled by two parameters: the minimal class density and the minimal distance between two separate classes. If these parameters are correctly evaluated, one obtains a good recognition of the data clouds. A very accurate pattern recognition may however produce the generation of many (very specific) rules, whose validity domains are so small that they become almost insignificant.

III. Study case

III.1 Description of the UDS

The selected **UDS** is inspired from real combined **UDS** in Northern Germany. Its central part (see **Fig. 2**) is composed of two pipe storage capacities **PSC1** and **PSC2**. Two retention basins **RB1** and **RB2** are connected to **PSC1**. Every pump has at most 5 flow rates values (= discontinuous flow control). The pipe transport is controlled by 3

pump stations (**P1**, **P2**; **P3**). **P3** controls the inflow rates to the Treatment Plant (**TP**). If it pumps more than the **TP** maximal capacity **CSO** occurs. Pump station **P4/P5** controls the access to retention basin **RB2**.
The runoff-inflows are generated by 5 identical subcatchments (**C1-C5**). Each one has 10 ha impervious area and 10 ha pervious area.

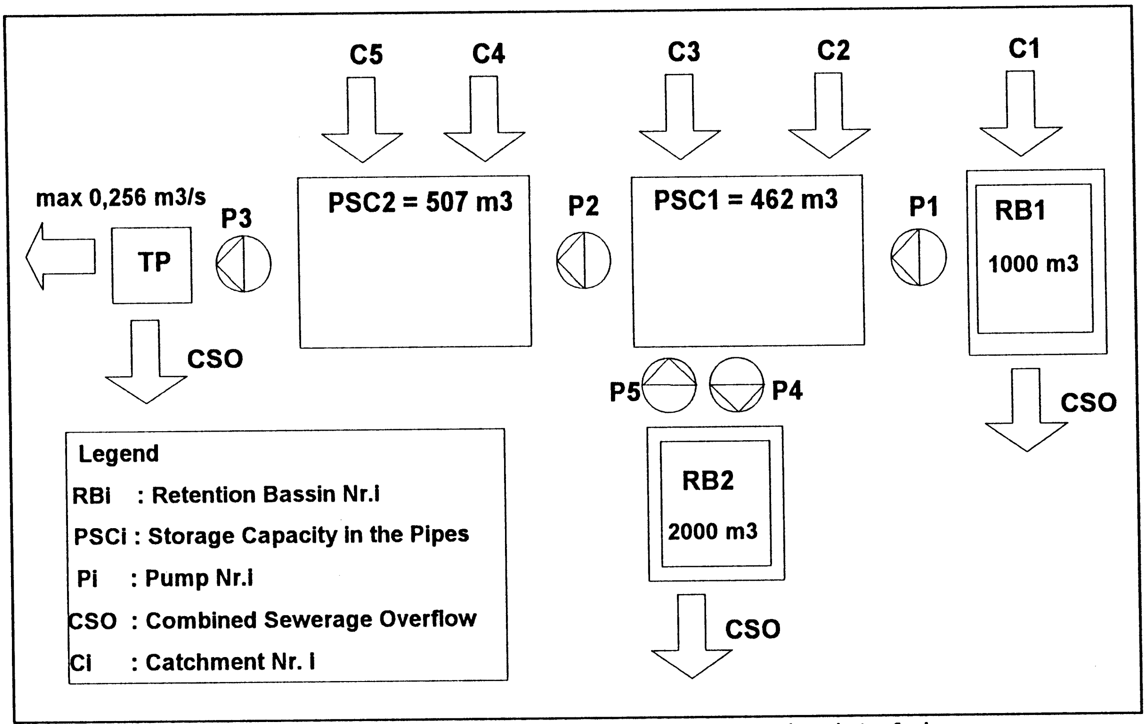

Fig. 2: simplified representation of the **UDS** from the control point of view

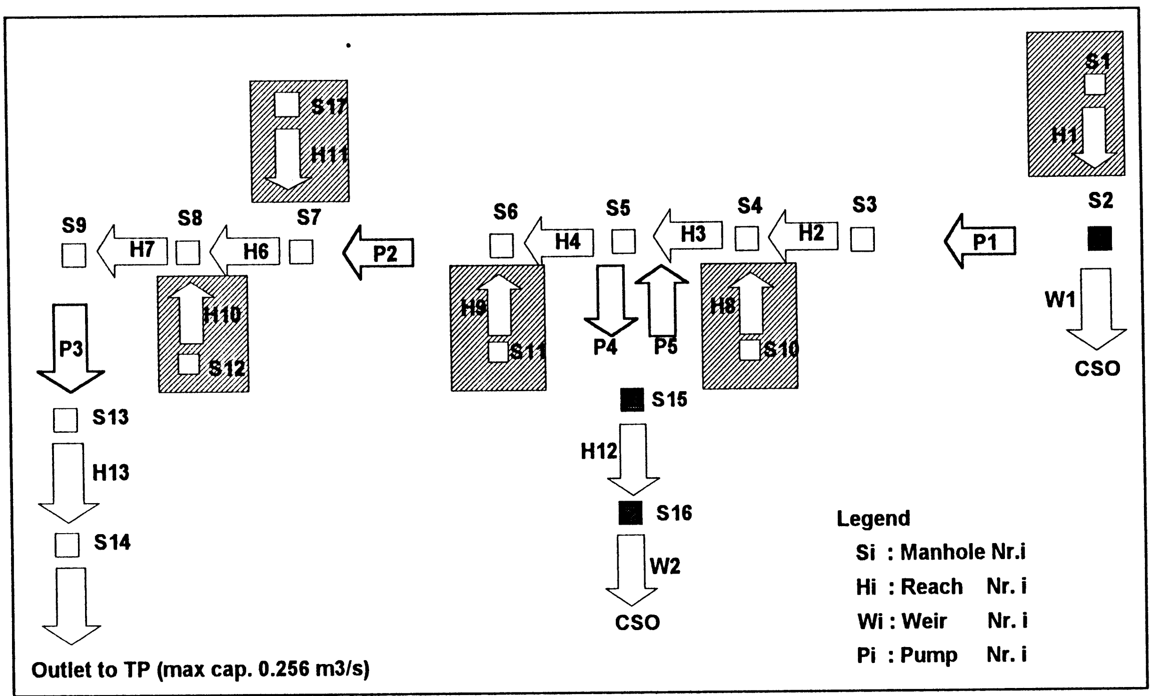

Fig. 3: Description of the **UDS** for the connected hydraulic transport model

III.2 Modelling of the flow processes

The rainfall and surface runoff processes are simulated with an hydrologic model according to the "method of the standard unit hydrograph" (**HYSTEM**). The generated inflow waves (not influenced by the control strategy) are stored in one file, which is read by the hydraulic model (**EXTRAN**) for a detailed simulation of the transport processes.

In order to reduce the amount of computation (without much distorting the results), the pipe system was divided into a central part and a peripheral part. In the central part the transport processes (see **Fig. 2** and **Fig. 3**) are simulated with the detailed hydraulic model. The "peripheral" transport, which takes place in each of the 5 subcatchments is globally accounted for in the hydrologic simulation of the inflow waves. This simplification required an analysis of the transport rainfall-runoff-transport characteristics in the subcatchments and a calibration of the runoff simulation.

III.3 Description of the local strategy (= initial strategy)

First simulations were done with **EXTRAN** not connected to the **RBS**. In this case, only one state variable can be considered for the control of each pump **Pi**, i=1,5: the water level in its pump pit (see **Tab. 1**). A local strategy was then elaborated according to following principles:

- the more water in the pipes, the higher the flow rates in the transport pumps
- if there is any danger of flooding in **PSC1**, the retention basin **RB2** should be filled up.

This local strategy was translated into a set of control rules, which describes the initial strategy in simulation with the **RBS.**

Pump name	Function	Flow rate (m3/s)	switch[1] point 1 (m)	switch point 2 (m)	pressure flow level (m)	flooding level (m)
P1	transport in the pipes	0.032 0.200 0.400 0.700 1.000	1 2 3 4	 0.8 1.8 2.8 3.8	1.0	6.0
P2	transport in the pipes	0.096 0.192 0.500 0.800 1.000	0.3 0.5 0.7 0.9	 0.15 0.35 0.55 0.75	0.90	3.5
P3	transport to the **TP**	0.128 0.256 0.500 0.700 1.000	2.0 3.0 3.5 3.9	 1.8 2.8 3.3 3.7	1.0	4.0
P4	Fill **RB2**	0.000 0.500 1.000 1.500 2.000	0.5 1.0 1.2 1.5	 0.3 0.8 1.0 1.3	0.90	3.0
P5[2]	no emptying of **RB2**	0.000	-	-	0.90	3.0

Tab. 1: Description of the local strategy

[1]Every water level (m) in **Tab.1** is related to the bottom level of the connected pipes. Switch point 1 is valid, if the water level in the pump pit increases and switch point 2, when it decreases.

[2]The hydraulic programme can not itself simulate the emptying of a retention bassin, because other state variables as the water level in the "pump pit" (=bassin) are required to make a decision (i.e rainfall intensities, water levels in the transport pipes). In the initial strategy, however, such conditions have been included so that the emptying phase may also be simulated.

IV. Analysis of the initial local strategy with design rainfalls

The selected rainfall data were recorded in Bremen (North Germany) during 1955-1985. After statistical analysis, four design rainfalls **DRi** i=1,4 were constructed and used to evaluate the local strategy (see **Tab. 2** and **Tab. 3**).

Design Rainfall	return period (a)	duration (min)	height (mm)	runoff vol. [3](m3) imp+perv=tot	max. runoff flow rate (m^3/s)
DR1	1	15	9.3	608+ 41= 649	0.422
DR2	5	15	14.5	1035+561=1596	0,762
DR3	1	30	11.7	805+101= 906	0.423
DR4	5	30	17.9	1320+721=2041	0,765

Tab. 2: Rainfall runoff characteristics of selected design rainfall

Design Rainfall	total vol. to the TP (m^3)	max. flooding vol. in **S8** (m^3/s)	max stored vol. in RB1 (m^3)	max. stored vol. in RB2 (m^3)
DR1	4944	818	316	0
DR2	5881	2801	1088	884
DR3	6204	1399	400	52
DR4	6181	3788	600	1505

Tab. 3: Evaluation of the **UDS** performances with selected design rainfall

The results show that according to German standards the pipes are not under-dimensioned . For events **DR1** and **DR3** (return period of 1 year), flooding at the weakest point of the system (**S8** in **PSC2**) could have been avoided, if the available storage capacity had been fully used. For events **DR2** and **DR4** (return period of 5 years) the available storage capacity (3969 m^3) is anyway insufficient. But, even in this case, the retention basins are not fully filled. In term of control decision, it can be concluded that:

- **P1** pumps too much in **PSC1** (**RB1** is not fully used)
- **P2** pumps too much in **PSC2** (flooding occurs)
- **P4** pumps too little in **RB2** (unused storage capacity)

V. Learning process

In accordance with the previous analysis, learning concentrates on **P1** (5 meta-rules) , **P2** (6 meta-rules), **P4** (7 meta-rules). The simulated real events must be so selected, that they generate a significant number of situations, in which local control is not appropriate. On the whole period 1955-1985, every rainfall was listed with following separation criterion: rainfall intensity < 0.1 mm/5min during more than 2 hours. The 261 detected events were ordered according to the mean rainfall intensity and 16 were finally retained: 12 among the 20 strongest events, the remaining 4 ones are medium (places 76-82). The shortest event lasts 45min, the longest 2h40min. In **Tab. 4** and **Tab. 5,** the rule basis before and after learning (70 cycles) is evaluated.

In all 16 events, flooding volumes and durations of pressure flow decrease[4]. Except event No2, the number of flooding spots is also reduced. However, improvement is not be achieved for every evaluation criteria:

- In all events the necessary delay to empty the **UDS** after rainfall is longer (**VOL2** increases)
- In the strongest events, less flooding implies more **CSO** volumes.

---

[3]Runoff volumes and maximal flow rates in **Tab. 1** are related to a single subcatchment **Ci**, i=1,5

[4]A learning step takes place each time all 16 events have been simulated.

It must also be noticed, that the inflow volumes to the **TP** have been hardly affected. No "learning" has taken place for **P3**.

No.	**QINF** (m^3)	**VOL1** (m^3)	**VOL2** (m^3)	Nb of flooding pts (-)	corresp. pres. flow duration (min)	max. **FLVOL** (m^3)	**CSO** vol (m^3)	**VTP** vol. (m^3)
1	28 191	614	1 286	6	265	12 759	9 194	18 301
2	25 952	614	1 295	5	272	10 803	7 937	17 321
3	24 287	614	1 075	5	295	6 858	7 026	16 789
4	21 202	614	808	5	200	4 767	5 519	15 490
5	**20 703**	**613**	**1 012**	**5**	**259**	**4 727**	**5 186**	**15 124**
6	6 134	612	315	1	0	0	0	6 521
7	14 608	614	601	5	120	2 485	2 078	12 565
8	12 836	614	604	5	96	1 696	1 178	11 710
9	11 919	614	591	5	67	1 064	728	11 236
10	10 895	615	545	4	33	375	0	11 007
11	9 738	614	595	2	4	70	0	9 823
12	11 143	613	575	5	52	779	385	10 827
13	9 482	612	580	1	0	1	0	9 565
14	9 433	613	588	1	9	116	0	9 505
15	8 217	614	587	1	3	3	0	8 338
16	9 487	613	590	2	14	241	0	9 557

Tab. 4: Evaluation[5] of the control strategy <u>before</u> learning (simulation interval: 10h)

QINF : total volume of runoff
VOL1 : sewage volume stored in the pipes, when the event begins
VOL2 : sewage volume in the system after simulation
pres. flow duration : addition of the durations of pressure flow at every flooding spots
max. **FLVOL** : addition of the max. flooding volumes at every flooding spots
CSO : **CSO** volume without consideration of the possible **CSO´s** in the **TP**
VTP : total sewage volume pumped into the TP during the simulation (10h)

No.	**QINF** (m^3)	**VOL1** (m^3)	**VOL2** (m^3)	Nb of flooding pts (-)	corresp. pres. flow duration (min)	max. **FLVOL** (m^3)	**CSO** vol (m^3)	**VTP** vol. (m^3)
1	28 191	615	1 746	6	179	8 561	9 412	17 688
2	25 952	615	2 066	6	180	6 590	8 046	16 503
3	24 287	613	1 679	5	81	2 171	7 946	15 347
4	21 202	613	1 427	4	69	1 638	6 130	14 355
5	**20 703**	**614**	**1 535**	**3**	**66**	**1 179**	**6 028**	**13 844**
6	6 134	613	509	1	0	0	0	6 406
7	14 608	612	1 233	5	46	1 083	2 206	11 891
8	12 836	612	1 216	2	24	561	1 281	11 102
9	11 919	612	799	2	17	287	1 207	10 766
10	10 895	612	883	1	10	146	554	10 324
11	9 738	613	959	1	4	29	0	9 629
12	11 143	613	950	2	14	363	391	10 685
13	9 482	615	960	0	0	0	0	9 336
14	9 433	613	707	2	4	95	0	9 567
15	8 217	612	865	1	0	0	0	8 193
16	9 487	614	828	1	6	114	0	9 495

Tab. 5: Evaluation of strategy <u>after</u> 70 learning steps (simulation interval: 10h)

A detailed analysis of the learning process shows that most improvements are achieved in the first 20 learning steps. As an example, **Fig. 4** shows the evolution of

---

[5]For all 16 events, a period of 10 hours after beginning of the rainfall has been simulation.

flooding volumes and pressure flow durations at **S8** during rainfall event No 5. After the 20th learning step, the **UDS** behaviour at the critical spots (**S3**, **S4**, **S5**, **S6**, **S8**) "oscillates" without significant trends. The limited size of the knowledge basis is mainly responsible for this. In its constant effort to improve the performance, more and more new rules are generated, until the storage capacity is exhausted. At this time, a certain amount of stored rules is deleted. **Fig. 5** traces the number of stored rules in the **RBS** during learning. From the 20th learning step onward, the maximal computer storage capacity is reached. Now and then, existing rules must be erased to make some place for the newcomers. But even if the computer storage were unlimited, oscillations would probably still happen because of the necessary restrictions in the storage-transport capacity of the **UDS**.

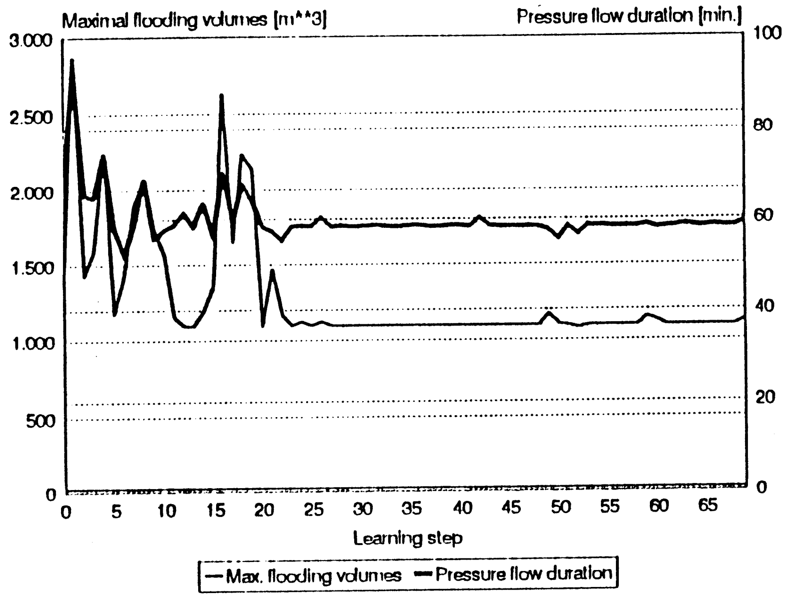

Fig. 4: Evolution of flooding vol. and pressure flow dur. at **S8** in event No 5

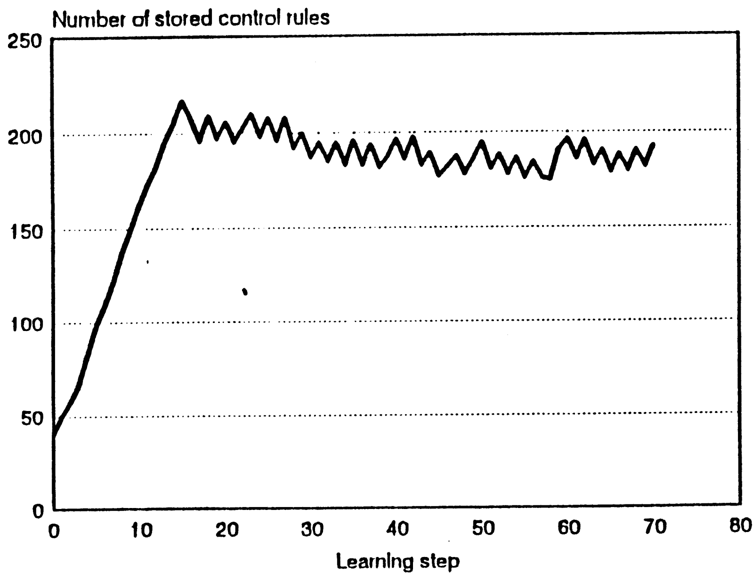

Fig. 5 : Evolution of the number of stored rules during learning

VI. Conclusions

Learning algorithms were developed to automatically improve control rules. The necessary meta-knowledge does not require much computer storage capacity and proves efficient. However, its formulation necessitates a careful analysis of the **UDS** behaviour at the initial state. The paper focuses on a specific learning algorithm, which can very accurately classify the state variables. The advantage is that no undue generalisation happens. The shortcoming is that many rules are generated, so that the computer storage capacity may be exhausted and learning oscillate.

REFERENCES

BOCK H., "Automatische Klassifikation", Studia Mathematica Band XXIV, ISBN 3-525-40130-2, edited by Vandenhoek & Ruprecht in Göttingen, 1974

FUCHS L., MÜLLER D., NEUMANN A., "Learning Production Systems for the Control of UDS", 1st IAWPRC International Symposium on Systems Analysis in Water Quality Management, London, 1987

KHELIL, "Development of a Control Strategy to Reduce Combined Sewer Discharges of a UDS, Case of Bremen-Germany", WATERMATEX '91, 2nd International Conference on Systems Analysis in Water Quality Management, Durham, USA, 1991, Proceedings edited in Pergamon Press

- A.J. Cole "Numerical taxonomy" ...